Universität Passau
Lehrstuhl für Mathematik mit Schwerpunkt
Digitale Bildverarbeitung
Prof. Dr. Tomas Sauer

Bachelorarbeit

Nash Equilibrium and Economic Applications

Betreuer: Prof. Dr. Tomas Sauer

Name: Lea Rohleder
Matrikelnummer: 102803

rohled02@ads.uni-passau.de
Bahnhofstraße 16B
94032 Passau

Abgabe am 26. Januar 2024

Abstract

What John Forbes Nash Jr. found out in his landmark work – "Equilibrium Points in N-Person Games" – is that even in complex non-cooperative games, there exists an equilibrium where no side would benefit by changing its course. By proving that there always exists at least one such equilibrium point in any such game, Nash contributed very much to the work of economists, politicians, business strategists and many more, who understand the world around us in a better way now. The primary objective of this Bachelor's Thesis is to provide two proofs to the Nash equilibrium using the fixed point theorems by Brouwer and Kakutani. We are going to subsequently apply this theoretical framework to analyse two classical economic scenarios – the Cournot Competition and the Bertrand Competition. The first part of the thesis focuses on establishing a comprehensive understanding of the Nash equilibrium. The proof of the Nash equilibrium serves as the cornerstone, explaining the conditions under which rational decision-makers, each pursuing their self-interest, converge to a stable and optimal outcome. The second part of the thesis delves into the application of the Nash equilibrium in the context of two prominent economic models: the Cournot model, featuring firms that compete in quantities, and the Bertrand model, where firms compete in prices.

Acknowledgements

I would like to express my gratitude to all those who have contributed to the completion of this thesis. First and foremost, I extend my appreciation to my thesis advisor, Prof. Dr. Tomas Sauer, for his guidance and expertise throughout the research process. I would also like to thank Thomas Kirz, Elina Heinloth, Claudia Kleiner, Johann Sebastian Schicho, Kassian Köck and Lea Heußer for their proofreading and editing assistance. Their attention to detail has significantly improved the clarity and coherence of this thesis.

Contents

1 Introduction **1**

2 Games and Game Theory **2**

2.1 Non-Cooperative and Cooperative Games 2

2.2 Pure Strategies 3

2.3 Games and Pay-Off 5

2.4 Mixed Strategies 6

3 A Special Case **12**

3.1 The Two-Person Zero-Sum Game 12

3.2 Minimax-Theorem 14

4 Nash Equilibrium in Non-Cooperative Games **18**

4.1 Equilibrium Points 18

4.2 Best Response 21

5 Digression into the World of the Fixed Points **25**

5.1 Fixed Points 25

5.2 Brouwer Fixed-Point Theorem 26

5.3 Kakutani's Gereralised Fixed Point Theorem 37

6 Nash's "One-Page-Proof(s)" **44**

7 Economic Applications **50**

7.1 Cournot Competition 50

7.1.1 The Easier Duopoly Case 51

7.1.2 The General Cournot Model 54

7.2 Bertrand Competition 56

8 Conclusion **58**

References **59**

List of Tables **61**

Eigenständigkeitserklärung **62**

1 Introduction

Let us imagine ourselves on a sunny afternoon in a vibrant metropolis. The hustle and bustle of the city is reflected in the determined steps of passers-by.

The city streets are lined with a colourful mosaic of shops, each with its unique character and offerings. There are boutiques with quaint storefronts showcasing the latest in fashion, from high-end designer labels to locally crafted, one-of-a-kind pieces. Trendy streetwear shops draw in the fashion-forward crowd with their innovative displays and hip atmospheres.

In this city, you'll find bustling markets that fill the air with the aroma of freshly cooked international cuisine. Stalls piled high with exotic fruits, spices, and street food tempt passers-by to indulge their taste buds. Food trucks and pop-up eateries offer a wide variety of culinary delights, from gourmet burgers to authentic sushi, creating a foodie's paradise.

The shops themselves are not just places to buy goods but destinations in their own right. Bookstores invite you to get lost among the shelves of novels, while record shops let you explore the rich history of music. Quirky speciality stores cater to niche interests, whether it's vintage vinyl records, rare collectables, or handmade crafts.

Unknowingly or maybe knowingly, we have become participants in a fascinating game of strategic interaction. A game based on the principles of game theory, in which every decision influences the actions of others. In this highly complex game, every shop wants a perfect price for their goods. Hence they seek an equilibrium state in which neither party has an incentive to deviate from their chosen course - a state known as the Nash equilibrium [1].

In what follows, we examine the early evolution of non-cooperative game theory and equilibrium points. Though both are associated with John Nash, the idea originated a lot earlier. In the following we will look at some creative mathematical work and how it was interpreted by mathematical economics.

2 Games and Game Theory

We all know games. They play a huge part in our culture – humans have always played games throughout their evolutionary history [2]. Awari, Chess, Uno, Dots and Boxes, Backgammon, Breakthrough, Schafkopf, different sports games and so many more are certainly no strangers to us.

So if games play such a huge role in the history of humankind, the fact that we analyse these games does not seem very surprising. Maybe we can go so far and describe the analysis of games as a natural instinct of a mathematician like John D. Beasley did in [3, p. 1]? What is a game? Are there any characteristics which are shared by all games? What is the best move in a certain situation? Who wins? These are the questions we want to explore – as simple as they may seem.

When we compare games, there are some evident differences. In Chess, Dots and Boxes or Backgammon, for example, we know everything about the position of our opponent – so nothing is hidden, whereas in card games like Uno or Schafkopf (when played fairly and according to the rules) we don't know the cards of our fellow players with certainty. Awari is a thousand year old game, while Breakthrough was invented very recently in 2001. Backgammon, Awari, Chess, Dots and Boxes and Breakthrough are games that are played by only two players. Uno for example can be played by multiple players. In Schafkopf, there have to be exactly four competitors. In sports, the outcome of a match is not completely determined by the player's decisions, but also by their physical skills. On the other hand, board games are only influenced by the decisions of the player. So it is pretty obvious that sports is a very difficult game to analyse.

For our interest – the analysis of the competition between companies – board-game-like-situations with a finite amount of players seem very fitting. Multiple players interact and their decisions determine some kind of outcome [4].

2.1 Non-Cooperative and Cooperative Games

Game theory was defined by Holler and Illing in [5, p. 1] as the analysis of strategic decision situations. These are situations where the following properties hold:

1. The result is highly dependent on the decisions of many decision makers[1].
2. Every decision maker knows about this dependence.
3. Every decision maker assumes that every other decision maker knows about this dependence.

[1] this means a result can not be independent of the chosen move of the fellow players

4. Every decision maker considers the three properties above.

Hence we can state that game theory deals with the relations between a decision-maker and the other players. When we fokus on these relations, we can distinguish between non-cooperative and cooperative game theory [4].

Von Neumann and Morgenstern shaped the subject of cooperative game theory in their book *Theory of Games and Economic Behaviour* like no other and laid the foundation for game theory in general. Their book is based on the analysis of the relations with coalitions. Those can be formed by the players. Nash mentions von Neumann and Morgenstern explicitly in [1].

In this thesis however, we will focus on non-cooperative games. These are defined by the absence of coalitions between players. So we assume that everybody acts purely independently and in their own interest [1].

We could now debate, if the assumptions of non-cooperative or cooperative game theory are more accurate. However, for our interest, non-cooperative games are more relevant. Firms are not allowed to make binding agreements over the price or the supplied quantities. In Germany price fixing and other agreements between companies that restrict competition are generally prohibited by the *Gesetz gegen Wettbewerbsbeschränkungen (§ 1 GWB)* [6].

2.2 Pure Strategies

Before we can dive into analysing non-cooperative games, we first need to establish a basis and define some concepts, which will be useful to relay on later on.

Von Neumann and Morgenstern define some fundamental components of games in [7]. A game Ω needs n players, which can be humans or organisations or in fact any decision maker – for example firms[2]. These will be denoted by $P_1, P_2, \ldots, P_n$.

Karlin defines these in [8, p. 15] as in the following definition:

Definition 2.2.1 (pure strategies)**.** A strategy for a player P_k is a sequence of all the actions P_k will take for all eventualities that could arise, whether by chance or by creation by one of the opponents. If P_k decides to play exactly one strategy, we call this a pure strategy $s_{k,i}$. Here k indicates the player and i which strategy the player chooses.

Remark 2.2.2. We note that the pure strategies result from looking at a single game run. In reality a player normally only plans a few steps ahead. In game theory on the

[2]this will play a significant role later on

other hand we assume that the player has already decided on a strategy that includes the entire game before they started to play [9, p. 4].

We can illustrate this seemingly abstract concept with an example, which we find in [9, p. 2].

Example 2.2.3 (Matching Pennies)**.** In this game there are two players. Player 1 – hence P_1 – chooses "Heads" (H) or "Tails" (T). Player 2 (P_2) does not know the choice of Player 1 and chooses Heads or Tails, too. This game only consists of a single move. So these two players have two strategies each. Each one can choose between Heads or Tails. Hence $s_{1,1} = H$ is the first pure strategy of P_1. This assignment is arbitrary.

We see that there could be more than just one pure strategy for each player[3]. So the pure strategies of a certain player can be accumulated to a set:

Definition 2.2.4 (strategy sets)**.** We call S_k the set of all pure strategies of P_k. So it holds that $S_k = \{s_{k,1}, \ldots, s_{k,e_k}\}$ with e_k being the number of strategies P_k has available.

To demonstrate this definition, strategy sets can be applied to our Example 2.2.3.

Example 2.2.5. The set of all pure strategies of player P_1 in this game is $S_1 = \{s_{1,1}, s_{1,2}\} = \{H, T\}$ and accordingly the set of all pure strategies of player P_2 in this game is $S_2 = \{s_{2,1}, s_{2,2}\} = \{H, T\}$. So in this case it applies that $S_1 = S_2$.

In general, strategy sets are not necessarily equal for all players. Consider any game with cards. Usually each player has a very different set of cards[4] and hence very different strategies. Another way more dangerous example is russian roulette, which is described in [10, p. 21].

Example 2.2.6 (Russian roulette)**.** Russian roulette is a lethal[5] game which is played by two players P_1 and P_2. Additionally, the players need a revolver loaded with only a single round[6]. The game is played as follows:

1. Each player places a winning unit (this could be anything).
2. P_1 has the choice: Either they could place two more winning units and pass the revolver to P_2 ("Pass") or they could place only one winning unit and place the muzzle against their own head and pull the trigger ("Play").
3. If P_1 survives their move, P_2 has the same choice (but they are not allowed to change the revolver in any way).

[3] otherwise our game would be very boring
[4] in some games the quantity of the cards might even differ
[5] or potentially lethal
[6] the other five places are not filled with cartridges

4. Now the game is over and the surviving players split the stake.

Hence P_1 has two strategies: Pass (PA) and Play (PL). P_2 on the other hand has four strategies: Pass, Play, play the same as P_1 (S) and do exactly the opposite of P_1 (O). Accordingly, our strategy sets are as follows:

$$S_1 = \{PA, PL\} \text{ and } S_2 = \{PA, PL, S, O\}. \tag{2.1}$$

2.3 Games and Pay-Off

But where is all the fun, when the players have nothing to win? This means we need a final concept: a pay-off [8, p. 16]:

Definition 2.3.1 (pay-off)**.** The pay-off is a rule that tells how much a player P_k is expected to win from choosing a certain strategy while the other players each choose a certain strategy, too.

This can be applied to our Pennies Example 2.2.3.

Example 2.3.2. The pay-off of our game is defined in the following way: If both players choose the same side of the penny, P_2 wins one monetary unit[7] from P_1. In the other case, where the players choose different sides, P_1 wins one monetary unit from P_2.

Now we are finally ready for a formal definition of a game. Nash identifies in [1, p. 286] the most important components and puts them together in one big definition.

Definition 2.3.3 (n-person game)**.** An n-person game Ω consists of a set of n players $P_1, P_2, \ldots, P_n$. Each P_k has a set of pure strategies S_k with $k \in \{1, \ldots, n\}$, which they choose from, and a payoff function p_k. This function maps the set of all n-tuples of pure strategies into real numbers, as we see in Equation 2.2:

$$p_k : S_1 \times \cdots \times S_n \to \mathbb{R},\ k = 1, \ldots n \tag{2.2}$$

We call the game finite if $\#S_k < \infty$ for $k = 1, \ldots, n$.

We can organise the pay-off function into schemes and get the so called normal-form game as defined in [9, p. 5]:

Remark 2.3.4. The pay-off function $p_k(s_1, s_2, \ldots, s_n)$ can now be tabulated for all possible tuples $(s_1, s_2, \ldots, s_n)$. In this way, we can see very easily which pay-off occurs given a specific tuple. This can happen with a relation or with an n-dimensional scheme[8].

[7] the monitery unit may be anything

[8] in the $n = 2$ case this is a matrix – in higher dimensions this is called a tensor [11, p. 455]

2.4 Mixed Strategies

Let us look again at our Example 2.3.2. If Player 1 considers which pure strategy would be best, it could look somewhat like this: "Player 2 surely expects me to play Heads, so I'm going to play Tails. But maybe they see that coming. So I should play Heads." This could go on and on, but Player 1 would not get to a satisfying result – much like Player 2.

With these considerations in mind, we see that pure strategies will not get us very far. In reality, a player often does not decide on exactly one strategy, but will decide to choose a certain order of executions. For example, this can look like this: Player 1 chooses two times Heads, then two times Tails and then again two times Heads. If this pattern is repeated several times, the opponent will figure out the strategy. So we need a strategy, where the actions are randomised. However, we need a concept where we can decide, how often actions are chosen in average. Thus, we assign probabilities to each action and choose a strategy by chance every time. We call this phenomenon mixed strategies, which are defined by Owen in [9, p. 15]:

Definition 2.4.1 (mixed strategies)**.** A mixed strategy of a player is a probability distribution on the set of their pure strategies.

Nash defines mixed strategies in the following way: A mixed strategy of player k is a collection of non-negative numbers which sum up to one[9] [1, p. 287]. Hence we can write $m_k = \sum_i^{e_k} a_{k,i} s_{k,i}$ with $a_{k,i} \in \mathbb{R}_{\leq 0}$ and $\sum_{i=1}^{e_k} a_{k,i} = 1$ and $S_k = \{s_{k,1}, \ldots, s_{k,e_k}\}$[10], where the $a_{k,i}$'s are the probability that the strategy $s_{k,i}$ is chosen.

The m_k's can be interpreted as points in an e_k-dimensional simplex $\mathbb{S}_{e_k}$, which is a convex subset of a real vector space.

Thus $\mathbb{S}_{e_k}$ is defined as

$$\mathbb{S}_{e_k} = \{\sum_{i=1}^{e_k} a_{k,i} s_{k,i} \in \mathbb{R}^{e_k} | \sum_i a_{k,i} = 1 \text{ and } a_{k,i} \geq 0 \text{ for all } i \in \{1, \ldots, e_k\}\} \subseteq \mathbb{R}^{e_k}. \quad (2.3)$$

But what about our pure strategies? Consider now a player k with e_k pure strategies. The probability distribution vector for the i^{th} pure strategy, denoted as $s_{k,i}$, is defined as:

$$s_{k,i} = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{S}_{e_k} \quad (2.4)$$

[9] this should make sense since we are looking at a probability distribution

[10] so P_k has e_k strategies to choose from

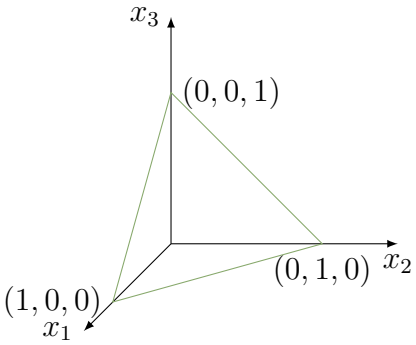
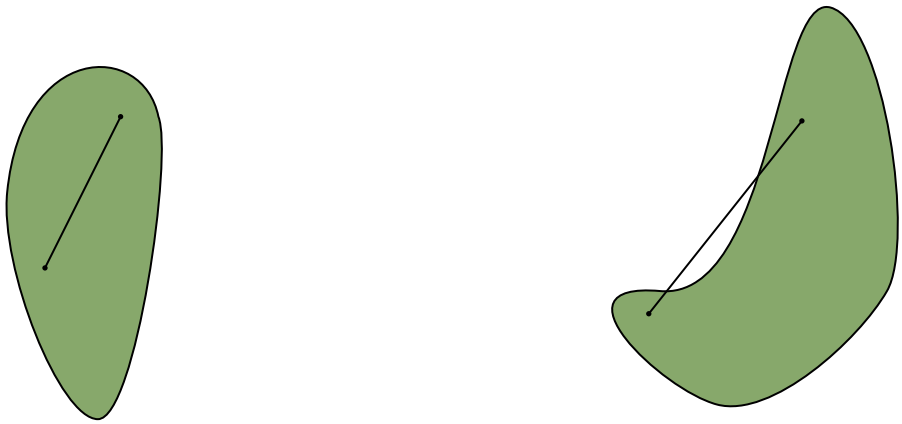


Figure 2.1: 3-dimensional simplex $\mathbb{S}_3$



(a) convex (b) non-convex

Figure 2.2: Two different sets in $\mathbb{R}^2$

with $a_{k,i} = 1$ and $a_{k,j} = 0$ for all $j \neq i$. Why is that? We see that the i^{th} entry is set to 1, reflecting the certainty with which the i^{th} pure strategy is chosen. All other components are set to 0, indicating that the player does not select the corresponding strategies.

It is easy to see that in the 3-dimensional case our simplex $\mathbb{S}_3$ looks like Figure 2.1.

Next we introduce the concept of a convex set and show that the simplex from Equation 2.3 is in fact convex. Deitmar defines this concept in [12, p. 234]:

Definition 2.4.2 (Convex Set)**.** A set $B \subseteq \mathbb{R}^m$ is called *convex* if and only if the connecting line of two points lying in B is also in B. Meaning the following:

$$b_1, b_2 \in B \;\Leftrightarrow\; \alpha b_1 + (1-\alpha) b_2 \in B \text{ for all } \alpha \in [0,1] \tag{2.5}$$

In Figure 2.2 we can see two graphical examples of the concept of convexity. In Figure 2.2a all two points can be connected by a line that also lies in the subset. Thus this subset is convex. Figure 2.2b is not convex, because we can find two points in the subset whose connecting line does not lie in the subset, as illustrated.

Additionally, we can show that our simplex is compact, where compactness is a property that combines closed and bounded subsets of $\mathbb{R}^{e_k}$ [13, p. 95].

Lemma 2.4.3. *The simplex $\mathbb{S}_{e_k}$ is*

1. *convex and*
2. *compact.*

Proof. Let us look at the convexity of $\mathbb{S}_{e_k}$:

Let $x, y \in \mathbb{S}_{e_k}$ with $x = \sum_{i=1}^{e_k} b_{k,i} s_{k,i}$ and $y = \sum_{i=1}^{e_k} c_{k,i} s_{k,i}$. Therefore we have: $\sum_{i=1}^{e_k} b_{k,i} = 1$ and $\sum_{i=1}^{e_k} c_{k,i} = 1$. Our aim is to show that

$$\begin{aligned} \alpha x + (1-\alpha)y &= \alpha \sum_{i=1}^{e_k} b_{k,i} s_{k,i} + (1-\alpha) \sum_{i=1}^{e_k} c_{k,i} s_{k,i} \\ &= \sum_{i=1}^{e_k} (\alpha b_{k,i} + (1-\alpha) c_{k,i}) s_{k,i} \in \mathbb{S}_{e_k} \text{ for any } \alpha \in [0,1]. \end{aligned} \tag{2.6}$$

This means we have to show that $\alpha b_{k,i} + (1-\alpha) c_{k,i} \geq 0$ for all $i \in \{1, \ldots, e_k\}$ and $\sum_{i=1}^{e_k} \alpha b_{k,i} + (1-\alpha) c_{k,i} = 1$.

Let us start by looking at $\alpha b_{k,i} + (1-\alpha) c_{k,i}$ with $i \in \{1, \ldots, e_k\}$ and $\alpha \in [0,1]$. We know that $b_{k,i} \geq 0$ since $x \in \mathbb{S}_{e_k}$. Because the same holds for $c_{k,i}$, we can conclude the following:

$$\underbrace{\alpha}_{\geq 0} b_{k,i} + \underbrace{(1-\alpha)}_{\geq 0} c_{k,i} \geq 0 \tag{2.7}$$

Since Equation 2.7 holds for any $i \in \{1, \ldots, e_k\}$, the first property of the convexity of $\mathbb{S}_{e_k}$ is proven.

The second property can be shown in the following way:

$$\begin{aligned} \sum_{i=1}^{e_k} (\alpha b_{k,i} + (1-\alpha) c_{k,i}) &= \alpha \underbrace{\sum_{i=1}^{e_k} b_{k,i}}_{=1 \text{ because } x \in \mathbb{S}_{e_k}} + (1-\alpha) \underbrace{\sum_{i=1}^{e_k} c_{k,i}}_{=1 \text{ because } y \in \mathbb{S}_{e_k}} \\ &= \alpha \cdot 1 + (1-\alpha) \cdot 1 = 1 \end{aligned} \tag{2.8}$$

Thus, the second property is proven and $\mathbb{S}_{e_k}$ is convex.

Now we can look at the compactness of $\mathbb{S}_{e_k}$:

1. Indeed it is closed. First, we show $\mathbb{S}_{e_k} = \mathbb{R}^{e_k}_{\leq 0} \cap \{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$.
   "$\subseteq$": If $x \in \mathbb{S}_{e_k}$ than $x \in \mathbb{R}^{e_k}_{\leq 0}$ and $x \in \{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$. Therefore $x \in \mathbb{R}^{e_k}_{\leq 0} \cap \{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$.

"$\supseteq$": On the other hand if $x \in \mathbb{R}^{e_k}_{\leq 0} \cap \{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$ we know that $x \in \mathbb{R}^{e_k}_{\leq 0}$ and $x \in \{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$. So $x_i \leq 0$ for all $i \in \{1, \ldots, e_k\}$ and $\sum_{i=1}^{e_k} x_i = 1$. This is exactly the definition of our simplex. It follows that $x \in \mathbb{S}_{e_k}$.
We know that the intersection of closed sets is again closed [13, p. 93]. With $\mathbb{S}_{e_k}$ being the intersection between $\mathbb{R}^{e_k}_{\leq 0}$ and $\{l \in \mathbb{R}^{e_k} | \sum_{i=1}^{e_k} l_i = 1\}$ (which are closed sets), we can conclude that $\mathbb{S}_{e_k}$ is closed.

2. To prove that $\mathbb{S}_{e_k}$ is bounded we start by stating that $\mathbb{S}_{e_k}$ is a subset of $[0,1]^{e_k}$. This is because we defined $\mathbb{S}_{e_k}$ as all mixed strategies of player k, which are a collection of non-negative numbers which have unit sum. $[0,1]^{e_k}$ is obviously bounded. Hence $\mathbb{S}_{e_k}$ must be bounded, too.

Thus we get the compactness of $\mathbb{S}_{e_k}$. □

A simplex is often defined as the convex hull as in [14, p. 931]. For the definition of a convex hull we are using [15, p. 37].

Definition 2.4.4. The convex hull of $y^1, \ldots, y^d \in \mathbb{R}^m$ is described as the set of all finite convex combinations:

$$\begin{aligned} \operatorname{conv}(y^1, \ldots, y^d) &= [\![y^1, \ldots, y^d]\!] \\ &= \left\{ \sum_{i=1}^{d} a_i y^i | \sum_{i=1}^{d} a_i = 1 \text{ and } a_i \geq 0 \text{ for all } i \in \{1, \ldots, d\} \right\}. \end{aligned} \tag{2.9}$$

When considering $y^i = s_{k,i}$, it is easy to see that our simplex is a convex hull.

But where do our pure strategies fit into these simplices or convex hulls? To answer this question, we first have to consider a special points in our simplex.

The *vertices* y of $\mathbb{S}_{e_k}$ or of $[\![s_{k,1}, \ldots, s_{k,e_k}]\!]$ are the points that can not be calculated with a strict convex combination of two points of $\mathbb{S}_{e_k}$ [10, p. 39]. Hence y is a vertex, if and only if the following holds true:

$$y = \alpha x_1 + (1-\alpha) x_2,\ \alpha \in (0,1) \Rightarrow x_1 = x_2 \tag{2.10}$$

This formulation helps to state the next lemma:

Lemma 2.4.5. *The pure strategies $s_{k,i}$ are the vertices of $\mathbb{S}_{e_k}$.*

Proof. Remind that $s_{k,i} = (0, \ldots, 0, 1, 0, \ldots, 0)$ with $i = 1, \ldots, e_k$, where the 1 occurs in the i^{th} entry.

Let us choose $x_1, x_2 \in \mathbb{S}_{e_k}$ with $s_{k,i} = \alpha x_1 + (1-\alpha) x_2$ for a fitting $\alpha \in (0,1)$ and with $x_1 = \sum_{i=1}^{e_k} b_{k,i} s_{k,i}$ and $x_2 = \sum_{i=1}^{e_k} c_{k,i} s_{k,i}$.

Now let us look at a $j \in \{1, \ldots, e_k\}$ with $j \neq i$. Hence it must hold that

$$0 = \underbrace{\alpha}_{>0} \underbrace{b_{k,j}}_{\geq 0} + \underbrace{(1-\alpha)}_{>0} \underbrace{c_{k,j}}_{\geq 0}. \tag{2.11}$$

The only possible solution for this equation is $b_{k,j} = c_{k,j} = 0$ for all $j \in \{1, \ldots, e_k\}$ with $j \neq i$.

Now there is only one entry left. This is the i^{th} entry, which is 1. It must hold that $b_{k,i} = c_{k,i} = 1$. This is because of $x_1, x_2 \in \mathbb{S}_{e_k}$ and thus $\sum_{l=1}^{e_k} b_{k,l} = \sum_{l=1}^{e_k} c_{k,l} = 1$.

Hence $b_{k,l} = c_{k,l}$ for all $l \in \{1, \ldots, e_k\}$. And for that reason it must hold that x_1 is equal to x_2. □

With u_i being the i^{th} unit vector we know the following:

$$\mathbb{S}_{e_k} = [\![s_{k,1}, \ldots, s_{k,e_k}]\!] = [\![(1,0,\ldots,0), \ldots, (0,\ldots,0,1)]\!] = [\![u_1, \ldots, u_{e_k}]\!] \tag{2.12}$$

So, our simplex $\mathbb{S}_{e_k}$ is built on the unit vectors u_i for all $i \in \{1, \ldots, e_k\}$ where i points to the entry which contains the 1. Because of this, $\mathbb{S}_{e_k}$ is often called *standard simplex* [10, p. 10].

In this context, we move beyond the traditional notion of a pay-off, as we now deal with probabilities associated with mixed strategies. In this case the pay-off is not guaranteed.

For that reason we need a function that tells us which pay-off is to be expected given the mixed strategy chosen by the player. So an extension to the tuples of mixed strategies is much needed, which we can find in [1]. Although this extension is something different, we will denote this by p_k. This function is linear in the mixed strategy of each player. This is called n-linear (since it is linear in all n components) or multi-linear. By

knowing that we can write the *expected pay-off* with $S_k = \{s_{k,1}, \ldots, s_{k,e_k}\}$ and $m_k = \sum_{i_k=1}^{e_k} a_{k,i} s_{k,i}$ in the following way:

$$\begin{aligned} p_k(m_1, \ldots, m_n) &= p_k(\sum_{i_1=1}^{e_1} a_{1,i} s_{1,i}, m_2, \ldots, m_n) \\ &= a_{1,1} p_k(s_{1,1}, m_2, \ldots, m_n) + \cdots + a_{1,e_1} p_k(s_{1,e_1}, m_2, \ldots, m_n) \\ &= \sum_{i_1=1}^{e_1} a_{1,i_1} p_k(s_{1,i_1}, m_2, \ldots, m_n) \\ &= a_{2,1} \sum_{i_1=1}^{e_1} a_{1,i_1} p_k(s_{1,i_1}, s_{2,1}, m_3 \ldots, m_n) + \ldots \\ &\quad + a_{2,e_2} \sum_{i_1=1}^{e_1} a_{1,i_1} p_k(s_{1,i_1}, s_{2,e_2}, m_3 \ldots, m_n) \\ &= \cdots = \sum_{i_1=1}^{e_1} \cdots \sum_{i_n=1}^{e_n} (\prod_{j=1}^{n} a_{j,j_k}) p_k(s_{1,i_1}, \ldots, s_{n,i_n}). \end{aligned} \tag{2.13}$$

We got over the first hurdle and defined what a game means to us!

Thus in a finite n-game each player has a finite set of pure strategies. When each of the n players choose their strategies, there is a certain pay-off for each player. So every n-tuple of pure strategies corresponds to an output to the n players [16, p. 48].

3 A Special Case

3.1 The Two-Person Zero-Sum Game

In Section 2 we examined the dynamics of n-person games defining their fundamental elements, including pure strategies, mixed strategies and (expected) pay-offs. Now, our focus narrows to a specific subset — the two-person zero-sum game, a concept we will navigate through with insights drawn from [9] and [8].

So, what does this term mean? The answer is elegantly intuitive. A two-person zero-sum game involves two participants, where the victory of one mirrors the loss of the other [8, p. 15]. This dynamic is captured in Equation 3.1, where all pay-offs sum up to zero.

$$p_1 + p_2 = 0 \tag{3.1}$$

Because in a zero-sum two-person game a win of one player implies the loss of the other player, it is enough to state the pay-off for one player. The pay-off of the second player is simply the negative of the pay-off of the first player.

Now we can take our findings from Section 2 and apply them to our specific case. There is a player P_1 with a set of strategies S_1 and a player P_2 with a set of strategies S_2. P_1 chooses a strategy $s_{1,i} \in S_1$, P_2 a strategy $s_{2,j} \in S_2$. The pay-off of the first participant is $p_1(s_{1,i}, s_{2,j})$, the pay-off for the second $p_2(s_{1,i}, s_{2,j})$.

Like stated in the previous section, the normal-form game is a matrix in the two-person case. Let us call this matrix P. The number of rows matches the number of strategies of player 1 – whereas the number of columns corresponds to the number of strategies of player 2. Player 1 and 2 choose their j^{th} and i^{th} strategy, respectively. The pay-off in this situation is equal the element p_{ji} of the matrix P.

We can explore this with our Matching Pennies Example 2.3.2. First we organise the pay-offs in a table.

		Player 2	
		Heads	Tails
Player 1	Heads	-1	1
	Tails	1	-1

Table 3.1: Pay-off in the game Matching Pennies for Player 1 from the Example 2.3.2

This table can now easily be transformed into a matrix P_1, which states the pay-offs that Player 1 gets given a certain strategy chosen by the two players.

$$P_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3.2}$$

And now we have the normal-form game of our Penny Example 2.3.2 in which p_{ji} is equal to the pay-off of Player 1, when Player 1 chooses their strategy $s_{1,j}$ and Player 2 chooses $s_{2,i}$. In this case it holds that $i, j \in \{1, 2\}$, because there are only two strategies for each player.

Because we can easily infer from the pay-off of one player to the pay-off of the other player, from now on we will only write P for the pay-off matrix of one of the two players.

We will see that some games have a strategy combination that is almost evident that it will be chosen. This is a so called saddle point[11]. Not every matrix has such a saddle point, as we see in the next examples which are acquired from [9, p. 13].

Example 3.1.1. The following normal-form game has a saddle point in the second row and the second column.

$$P = \begin{bmatrix} 5 & 1 & 3 \\ 3 & \underline{2} & 4 \\ -3 & 0 & 1 \end{bmatrix} \tag{3.3}$$

Player 1 will decide, to play their first or second strategy. This is because these two strategies are in all cases better than the third one. Hence we can assume that Player 2 will take this circumstance into account and will prefer their second strategy, because they will lose the least big amount with this strategy. Thus Player 1 would be well advised, to play their second strategy. That is because they will get a higher pay-off (2) than by playing their first strategy (1).

Example 3.1.2. The normal-form game in this example is a representation of the Example 2.3.2.

$$P = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3.4}$$

This matrix has no saddlepoint.
If Player 1 considers which strategy would be best, it could look like this: “Player 2 surely expects me to play Heads, so I’m going to play Tails. But maybe they see that

[11] We will refrain from a mathematically correct definition here, as this is not the subject of this thesis.

coming. I should play Heads." This could go on and on, but Player 1 would not get to a satisfying result – much like Player 2.

As was shown in the last section, the example presented justifies the use of mixed strategies. The strategic mixing of pure strategies is not just a theoretical consideration, but a practical necessity that forms the basis for the analysis in the next subsection.

3.2 Minimax-Theorem

Now that we have explored the field of games with and without saddle points a little, a question arises: How do we deal with games in which these clear advantages do not exist?

Unlike their straightforward counterparts, games without saddle points demand a deeper analysis. Here, a crucial property comes into play: complete information. In these games, each player possesses knowledge of all the decisions made by their opponent before making their own move. To illustrate this, let us take a look at an example from [9, p. 14].

Example 3.2.1. Assume there is a game with the following pay-off-matrix

$$P = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \tag{3.5}$$

Suppose we are Player 1. Of course we want to maximise our profit, while keeping our opponent in mind. Hence we will definitely choose a strategy where we at least get 2. However with the second strategy we will win only 1 because our opponent will know which strategy we are going to choose and wants want to minimise our pay-off. So we will minimise in the rows and then maximise over these entries. We will call this minimum profit v_1^{*}[12]. This means[13]:

$$v_1^* = \max_k(\min_i \ p_{ki}) \tag{3.6}$$

Suppose now we are Player 2. Here we want to minimise the profit of the opponent. By choosing the first strategy, our opponent would also choose their first strategy and we would loose 4. We will choose the second strategy because in this case we would

[12] which has the value 2 in our example

[13] here we need to keep in mind that p_{ki} represents an element of our matrix P

loose at most 3. Now we have the following for v_2^*[14], which can be regarded as the maximum loss:

$$v_2^* = \min_i(\max_k p_{ki}) \tag{3.7}$$

Now we can show the following result, as it is stated in [9, p. 15]:

Lemma 3.2.2. *Let P be the pay-off matrix of a zero-sum two-person game. Assume player 1 has the strategies $s_{1,1}, \ldots, s_{1,e_1}$ and player 2 has the strategies $s_{2,1}, \ldots, s_{2,e_2}$. Then the following holds:*

$$\max_{k\in\{1,\ldots,e_1\}}(\min_{i\in\{1,\ldots,e_2\}} p_{ki}) \leq \min_{i\in\{1,\ldots,e_2\}}(\max_{k\in\{1,\ldots,e_1\}} p_{ki}). \tag{3.8}$$

Proof. It holds that for every $i \in \{1, \ldots, e_2\}$ and $k \in \{1, \ldots, e_1\}$:

$$\max_{j\in\{1,\ldots,e_1\}} p_{ji} \geq p_{ki}. \tag{3.9}$$

Adding a minimum to each side does not change truth of the Equation 3.9 for any of the $k \in \{1, \ldots, e_1\}$. Hence we get the following:

$$\min_{i\in\{1,\ldots,e_2\}}(\max_{k\in\{1,\ldots,e_1\}} p_{ki}) \geq \min_{i\in\{1,\ldots,e_2\}} p_{ki}. \tag{3.10}$$

Equation 3.10 is also true for the largest element of the right side. Why is that? As mentioned above our Equation 3.10 holds true for every $k \in \{1, \ldots, e_1\}$. Thus we have the following:

$$\min_{i\in\{1,\ldots,e_2\}}(\max_{k\in\{1,\ldots,e_1\}} p_{ki}) \geq (\max_{k\in\{1,\ldots e_1\}}\min_{i\in\{1,\ldots,e_2\}} p_{ki}). \tag{3.11}$$

And we are done. □

As Owen notes in [9, p. 15], sticking to rational thinking alone can make our game somewhat predictable and a little mundane. To add a touch of unpredictability to the game, we turn to a more irrational approach to strategy selection: mixed strategies. In this context, let $\mathbb{S}_{e_1}$ represent the set of all mixed strategies for player 1, where e_1 denotes the quantity of pure strategies at player 1's disposal. Analogously, this should hold true for $\mathbb{S}_{e_2}$ and player 1.

Now, let us navigate through a transformation, relying on [9, p. 16] unless specified otherwise. Imagine player 1 and player 1 are engaging in a game with the pay-off matrix

[14] which has the value 3 in our example

P, each one choosing strategies m_1 and m_2 respectively. According to Equation 2.13, the expected pay-off is given by:

$$p(m_1, m_2) = \sum_{k=1}^{e_1} \sum_{i=1}^{e_2} (m_{1_k} m_{2_i}) p_{ki} = \sum_{k=1}^{e_1} \sum_{i=1}^{e_2} m_{1_k} p_{ki} m_{2_i}. \tag{3.12}$$

Or we can simply transfer this into a notation with our pay-off matrix P:

$$p(m_1, m_2) = m_1 P m_2^T. \tag{3.13}$$

As discussed above player 1 always has to consider that player 2 knows what they are about to do. Therefore player 1 is going to minimise $p(m_1, m_2)$. This means that when player 1 plays m_1, their expected pay-off is:

$$v(m_1) = \min_{m_2 \in \mathbb{S}_{e_2}} m_1 P m_2^T \tag{3.14}$$

To determine this, could be quite a challenge. Hence we are going to simplify $v(m_1)$ by the following Lemma, which can be found in [17, p. 308]. Note here that $P^{(j)}$ points to the j^{th} column of the matrix P.

Lemma 3.2.3. *For a fixed strategy $m_1 \in \mathbb{S}_{e_1}$ for Player 1, the following holds true:*

$$v(m_1) = \min_{m_2 \in \mathbb{S}_{e_2}} m_1 P m_2^T = \min_{j \in \{1,\dots,e_2\}} m_1 P^{(j)}. \tag{3.15}$$

Proof. Choose a fixed strategy $m_1 \in \mathbb{S}_{e_1}$ for Player 1. Then let

$$x = \min_{j \in \{1,\dots,e_2\}} m_1 P^{(j)}. \tag{3.16}$$

With Equation 3.16 we can conclude for any $m_2 \in \mathbb{S}_{e_2}$[15]:

$$\begin{aligned} m_1 P m_2^T &= \sum_{j=1}^{e_2} (m_1 P^{(j)}) m_{2_j} \\ &\geq \sum_{j=1}^{e_2} x m_{2_j} \\ &= x \sum_{j=1}^{e_2} m_{2_j} \\ &= x \cdot 1 = x \end{aligned} \tag{3.17}$$

The outcomes $m_1 P^{(j)}$ are contained in $\{m_1 P m_2^T \mid m_2 \in \mathbb{S}_{e_2}\}$. This is, because $m_1 P^{(j)}$ is the outcome which will be achieved by Player 2 using their j^{th} pure strategy. And

[15] here it is very necessary to keep in mind that $\sum_{j=1}^{e_2} m_{2_j} = 1$

the j^{th} pure strategy of course lies in $\mathbb{S}_{e_2}$. With $m_1 P m_2^T \geq \min_{j\in\{1,\dots,e_2\}} m_1 P^{(j)}$ we have:

$$\min_{m_2\in\mathbb{S}_{e_2}} m_1 P m_2^T = \min_{j\in\{1,\dots,e_2\}} m_1 P^{(j)} \tag{3.18}$$

□

This leads us to the expression:

$$v(m_1) = \min_{j\in\{1,\dots,e_2\}} m_1 P^{(j)}. \tag{3.19}$$

Naturally, player 1 seeks to maximise their pay-off, and thus selects m_1 in such a way to achieve the highest possible pay-off. Consequently, we express player 1's optimal value as:

$$v_1 = \max_{m_1\in\mathbb{S}_{e_1}} \min_{j\in\{1,\dots,e_2\}} m_1 P^{(j)}. \tag{3.20}$$

The m_1 is called *maximin-strategy of player 1*, referring to the order of the maximum and minimum.

Analogously, we can state that player 2 gets the following expected loss by choosing $m_2 \in \mathbb{S}_{e_1}$ where $P_{(i)}$ represents the i^{th} row of P:

$$v(m_2) = \max_{i\in\{1,\dots,e_1\}} P_{(i)} m_2^T. \tag{3.21}$$

Now player 2 wants to minimise their loss, such that they get:

$$v_2 = \min_{m_2\in\mathbb{S}_{e_2}} \max_{i\in\{1,\dots,e_1\}} P_{(i)} m_2^T \tag{3.22}$$

Here m_2 is called the *minimax-strategy for player 2*, again referring to the order of the minimum and maximum.

Now we are ready for one of the most important theorems in game theory - also known as the Minimax Theorem, which was proven by von Neumann in [7]. We will not prove this theorem since we are about to prove a much more general version.

Theorem 3.2.4 (Minimax Theorem)**.** *In every two-person zero-sum game with finite (pure) strategy spaces S_1 and S_2, there exists a constant v and for each player at least one (mixed) equilibrium strategy with which they can guarantee an expected pay-off of at least v. The v satisfies the following condition:*

$$v_1 = v = v_2 \tag{3.23}$$

4 Nash Equilibrium in Non-Cooperative Games

The concept of a *Nash equilibrium*, a cornerstone in non-cooperative game theory as defined in section 2, finds its roots in the work of John Nash, formalised in [16] and [1]. However, its origin can be traced back even further to Cournot's contributions in [18].

To illustrate the importance of the Nash equilibrium, let us start with an example - an example that encapsulates the essence of what this chapter is about. This example, taken from [19], serves as a showcase for the topics we are about to explore.

Example 4.0.1 (Mathematicians Go to Dinner)**.** Twenty mathematicians decide to go out to dinner. Because they are mathematicians[16], they agree to split the bill completely equally. The restaurant only offers a $ 10 and a $ 20 dinner. When every person acts on the basis of profit maximising, what will be ordered?

At first, we would probably guess that any mathematician would order the $ 10 meal to maximise their utility. But if we think about it a little longer, we actually come to a very different conclusion. In fact everybody will order the $ 20 dinner. Every player thinks that the others will definitely choose the $ 10 dinner. So if the player chooses the $ 20 one, they will get it for just $ 10.50. However, everyone else has exactly the same thought. Now every mathematician orders the $ 20 dinner - even though nobody actually wants to pay the $ 20.

The example shown is a so called Nash equilibrium. It seems perfectly fitting for our attempt to explain the competition between firms and find an equilibrium.

Within the context of companies competing for market dominance, the Nash equilibrium reflects a state of stability where no single firm can unilaterally improve its position by changing its strategy. It mirrors the delicate balance of competition, where each firm, acting in its own self-interest, collectively contributes to a state of equilibrium.

But how can we define this formally? Under which circumstances does a Nash equilibrium exist? What does this imply in the world outside of mathematics? These are the questions we will discuss in the next sections.

4.1 Equilibrium Points

But first we need to ask, what makes an equilibrium point an equilibrium point. The Oxford English Dictionary states in [20] that an equilibrium is "the state of equal

[16] probably not very fair ones

balance between powers of any kind". Hence it can be understood as the exact state, where all players do not see any inducement to change their strategy. This has to be formalised in a mathematical way, which again leads us to Nash. He uses the concepts from the previous section to define an equilibrium point in [1, p. 287][17].

Definition 4.1.1 (Equilibrium Point)**.** An n-tuple $\wp \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$ is an equilibrium point if and only if for every player k the following holds[18]:

$$p_k(\wp) = \max_{m_k \in \mathbb{S}_{e_k}} p_k(m_k, -m_k) \text{ with } k \in \{1, \ldots, n\}. \tag{4.2}$$

Nash also expressed this definition in [1] in other words: An equilibrium point is an n-tuple $\wp$, so that the mixed strategy of each player maximise their own pay-off, while the strategies of the other players are fixed.
Or we say: Each player's strategy is optimal against those of the others.

Remark 4.1.2. We note that this definition on equilibrium points is very pessimistic. If all of the other players' strategies are fixed and only player k changes their strategy, they will worsen their own position. We can actually see that in Example 4.0.1. If any person would deviate from the Nash equilibrium - so choose the 10$ dinner - they would have to pay more than they get.

Let us also use Example 4.0.1 to explore why the non-cooperation is so important in this context. If the mathematicians would have cooperated[19], they would have seen that everybody is going for the more expensive dinner. Hence they could have agreed to each of them buying the cheaper one[20].

Another very famous example for Remark 4.1.2 is the so called Prisoner's Dilemma, which is taken but slightly changed from [19, p. 53].

Example 4.1.3 (Prisoner's Dilemma)**.** Let us imagine being in custody[21] together with a partner in crime. Together we have committed a serious crime which, if proven, will land us each in prison for ten years. For now there is only enough evidence to convict us for less serious crimes to only three years. The police offers the following deal: If one talks, there only will be a sentence of one year - of course only for the

[17] an *n-tuple* $\wp$ shall be regarded as a point in a vector space or rather the product space of the vector spaces containing the mixed strategies (the simplices)

[18] The notation in Equation 4.2 stands for

$$p_k(m_k, -m_k) = p_1(m_1, \ldots, m_{k-1}, m_k, m_{k+1}, \ldots m_n). \tag{4.1}$$

In this notation $-m_k$ represents the mixed strategies of the opponents of P_k. Hence it holds: $-m_k \in \mathbb{S}_{e_1} \times \ldots \mathbb{S}_{e_{k-1}} \times \mathbb{S}_{e_{k+1}} \times \cdots \times \mathbb{S}_{e_n}$.

[19] hence in this case talked

[20] or they might find another agreement, which is fair to all of them

[21] admittedly that is very difficult, but let us try it anyway

talking one. The other one will get a higher sentence for not talking. Hence we will have the following pay-off matrix

$$\begin{bmatrix} (-3,-3) & (-11,-1) \\ (-1,-11) & (-10,-10) \end{bmatrix} \tag{4.3}$$

At the first glance it appears that no prisoner should talk. In this case nobody gets free, but each of them gets a relatively light sentence. However, if one prisoner is absolutely convinced that their partner in crime is truly loyal, they will be better of by talking to the police. Exactly the same holds true for their partner. Hence both will talk to the police and serve the longer sentence. This is a Nash equilibrium, because nobody has any inducement to change their strategy. If they did so, they would only worsen their position.

These two previous Examples 4.0.1 and 4.1.3 both have a pure strategy equilibrium. Again this kind of equilibrium will not get us very far as we have seen in section 3. Let us go back to our Pennies Example 2.3.2 again, which has the following pay-off matrix for player 1:

$$P_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{4.4}$$

We have already established in section 3 that this game has no equilibrium point in pure strategies. But we already know that there is an equilibrium point in mixed strategies. Von Neumann and Morgenstern proved this for zero-sum two-person games in [7]. We stated this in Theorem 3.2.4.

Nash generalised this statement to all finite games – no matter if zero-sum or not and with any number of players. Hence we have the following statement as in [1]:

Theorem 4.1.4 (Nash)**.** *Every finite game has an equilibrium point in mixed strategies.*

Remark 4.1.5. The Minimax Theorem is obviously a special case for two-person zero-sum games of Theorem 4.1.4.

This little sentence in Theorem 4.1.4 – as simple as it may seen – had a huge impact on the field of economics and the social sciences. Some – for example economist Roger Myerson in [21, p. 1067] – go so far and compare the proof of the Nash equilibrium to the discovery of the DNA double helix.

4.2 Best Response

In his paper from 1950, Nash characterised his equilibrium in a slightly different way, too [16]. In this section, we will briefly explore this alternative before we get to the proof of Theorem 4.1.4.

Definition 4.2.1. An n-tuple $\wp = (\tilde{m}_1, \ldots, \tilde{m}_n) \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$ *counters* another n-tuple $m = (m_1, \ldots, m_n) \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$ if the strategy of each player in the countering n-tuple $\wp$ yields the highest obtainable pay-off expectation for its player against the $n-1$ strategies of the other players in the countered n-tuple m. This means that if $\wp$ counters m, for every player k it holds that

$$p_k(\tilde{m}_k, -m_k) \geq p_k(l_k, -m_k) \text{ for all } l_k \in \mathbb{S}_{e_k}. \tag{4.5}$$

Now it is easy to see how our equilibrium point fits into this new concept:

Definition 4.2.2 (Equilibrium point with countering strategies)**.** A self-countering n-tuple $\wp$ is an equilibrium point.

In other publications, for example in [4, p. 24], this phenomenon applied to one player is often regarded as best response or best reply[22].

Definition 4.2.3 (best response)**.** For any player $k \in \{1, \ldots, n\}$ in an n-person game and any strategy combination by k's opponents, $-m_k \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_{k-1}} \times \mathbb{S}_{e_{k+1}} \times \cdots \times \mathbb{S}_{e_n}$, a *best response* of P_k to $-m_k$ is a strategy $\tilde{m}_k \in \mathbb{S}_{e_k}$ such that

$$p_k(\tilde{m}_k, -m_k) \geq p_k(l_k, -m_k) \text{ for all } l_k \in \mathbb{S}_{e_k}. \tag{4.6}$$

It is easy to see, that there can be multiple best responses for P_k. Thus, there is a set of best responses to a fixed $-\tilde{m}_k$. We will define it the following way:

$$\begin{aligned} C_k(-m_k) &= \{\tilde{m}_k \in \mathbb{S}_{e_k} | p_k(\tilde{m}_k, -m_k) \geq p_k(l_k, -m_k), \forall l_k \in \mathbb{S}_{e_k}\} \\ &= \arg \max_{m_k \in \mathbb{S}_{e_k}} p_k(m_k, -\tilde{m}_k) \end{aligned} \tag{4.7}$$

Hence we can once again define an equilibrium.

Definition 4.2.4 (Equilibrium point with best response)**.** A strategy combination $\wp = (m_1, \ldots, m_n) \in \mathbb{S}_{e_1} \times \ldots \mathbb{S}_{e_n}$ is an equilibrium point if, for every player P_k,

$$m_k \in C_k(-m_k). \tag{4.8}$$

[22] we changed the definition from [4, p. 24] here a little bit, because we have to consider mixed strategies

Theorem 4.1.4 tells us that there is at least one equilibrium point if the game is finite.

Let us explore this property with our Matching Pennies example from Example 3.1.2 with the following pay-off matrix:

$$P = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{4.9}$$

We already considered why this matrix has no pure-strategy Nash equilibrium. We now ask: does it have a mixed-strategy Nash equilibrium[23]?

To answer this question, we have to find mixed strategies for both players that are their mutual best responses. For the purpose of simplification, we define mixed strategies for the players 1 and 2 as follows: Let p be the probability that player 1 plays H and $1-p$ the probability for playing T. Therefore, a mixed strategy for player 1 looks like this: $m_1 = (p, 1-p)$. Similarly, let q be the probability that player 2 plays H and $1-q$ the probability for playing T. So the mixed strategies for player 2 are defined by $m_2 = (q, 1-q)$. Using the formulas for expected pay-offs in this game, we can write player 1's expected pay-off from each of their two pure actions as follows:

$$\begin{aligned} p_1((1,0),(q,1-q)) &= q \cdot (-1) + (1-q) \cdot 1 = -2q + 1 \\ p_1((0,1),(q,1-q)) &= q \cdot 1 + (1-q) \cdot (-1) = 2q - 1. \end{aligned} \tag{4.10}$$

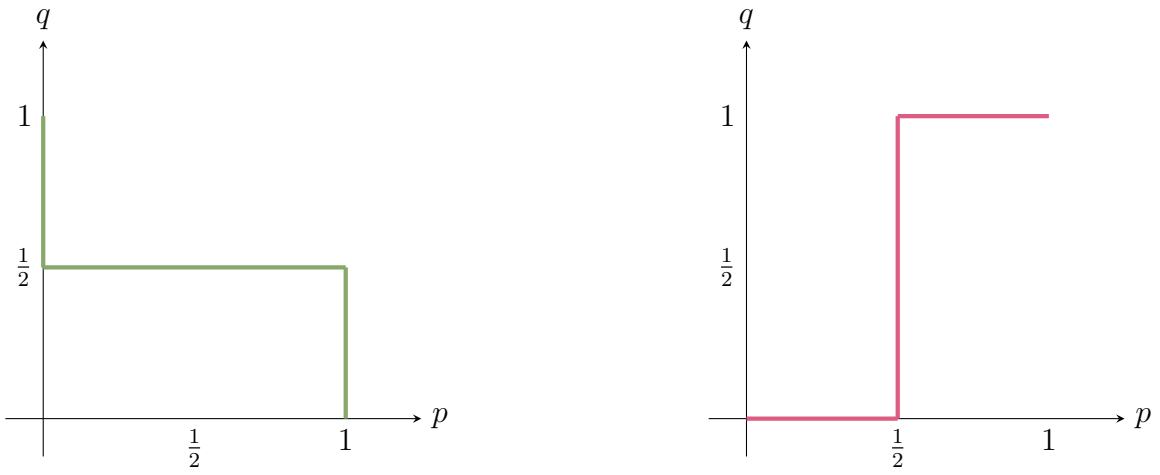
With these equations, we can calculate the best response of player 1 for any choice q of player 2.

Of course, player 2 will prefer to play H over playing T if and only if $p_1((1,0),(q,1-p)) > v_1((0,1),(q,1-q))$. When we use Equation 4.10, this is true if and only if

$$\begin{aligned} -2q + 1 &> 2q - 1 \\ \iff q &< \frac{1}{2}. \end{aligned} \tag{4.11}$$

Analogously we can proceed with T. Playing T will be strictly better than playing H for player 1 if and only if $q > \frac{1}{2}$. When $q = \frac{1}{2}$, player 1 will be indifferent between

[23] Well, we already know that is has at least one (but because we did not prove Nash's Theorem, it is reasonable to ask this question).



(a) C_1 (b) C_2

Figure 4.1: Best-response of player 1 and player 2

playing H or T. This analysis leads to the following best-response correspondence of player 1:

$$C_1((q, 1-q)) = \begin{cases} p = 1 & \text{if } q < \frac{1}{2} \\ p \in [0,1] & \text{if } q = \frac{1}{2} \\ p = 0 & \text{if } q > \frac{1}{2} \end{cases} \tag{4.12}$$

Now we can do exactly the same for player 2. Again, we can calculate the pay-offs of player 2 given a mixed-strategy p of player 1 and we get:

$$\begin{aligned} v_2((p, 1-p), (1, 0)) &= p \cdot 1 + (1-p) \cdot (-1) = 2p - 1 \\ v_2((p, 1-p), (0, 1)) &= p \cdot (-1) + (1-p) \cdot 1 = -2p + 1. \end{aligned} \tag{4.13}$$

As we have seen above, this implies that player 2's best response is given with:

$$C_2((p, 1-p)) = \begin{cases} q = 0 & \text{if } p < \frac{1}{2} \\ q \in [0,1] & \text{if } p = \frac{1}{2} \\ q = 1 & \text{if } p > \frac{1}{2} \end{cases} \tag{4.14}$$

To find a Nash equilibrium, we are looking for a pair of choices (p, q) for which the two best-response correspondences cross. This is why we are now going to map these two best responses in Figure 4.1.

If we superimpose the best response of player 2 onto Figure 4.1a, we will see that the two best-response correspondences cross at $p = q = \frac{1}{2}$. So there is indeed a mixed-strategy Nash-equilibrium.

Our primary objective in this thesis is to prove the very theorem that says there is always at least one equilibrium point. Nash proved this theorem by using both the Brouwer Fixed-Point Theorem in [1] and Kakutani's Generalized Fixed Point Theorem in [16]. Hence we need to gather some background information on Fixed Point Theory.

5 Digression into the World of the Fixed Points

So let's dive into the fixed point theory to prove the Nash equilibrium in Theorem 4.1.4. But when we dip into the theory of fixed points, a fascinating world unleashes itself. In these sections we will mostly use [19] and [22].

5.1 Fixed Points

First, we need to clarify what a fixed point is. The answer to this question is pretty intuitive as we can see in [19, p. 3].

Definition 5.1.1 (Fixed Point)**.** Given a function $f : M \to M$, a fixed point of f is a point $\tilde{x} \in M$ with $f(\tilde{x}) = \tilde{x}$.

Remark 5.1.2. We note that the identity function is a function with infinite fixed points, whereas $f : (0, 1) \to (0, 1)$ with $f(x) = x^2$ has no fixed points.

In this section we will learn which conditions must be met for a function to have a fixed point.

At first sight having a fixed point is a property of a mapping, but as seen in [19, p. 41] we can actually generalise it to spaces.

Definition 5.1.3 (Fixed Point Property)**.** A topological space X has a fixed point property if every continuous mapping $f : X \to X$ has at least one fixed point.

This may seem abstract. But what does Definition 5.1.3 exactly mean? Let us make this property clear with an example.

Example 5.1.4. The interval $[0, 1] \subset \mathbb{R}$ has the fixed point property.
Consider a continuous function $g : [0, 1] \to [0, 1]$ and another mapping $h(x) = g(x) - x$. It follows: $h(0) = g(0) \geq 0$ and $h(1) = g(1) - 1 \leq 0$. With the intermediate value theorem – which we can take from [23] – there is a $\varepsilon \in [0, 1]$ with $0 = h(\varepsilon) = g(\varepsilon) - \varepsilon$, so $g(\varepsilon) = \varepsilon$.

Do we always have to find the fixed points in every function to prove that a space has the fixed point property? This seems very time-consuming. In fact we do not. We will see why in the next Corollary 5.1.5.

Corollary 5.1.5. *Let X and Y be homeomorphic topological spaces, i.e. there is a mapping from Y to X which is continuous and bijective and whose inverse function is also continuous. If Y has the fixed point property so does X.*

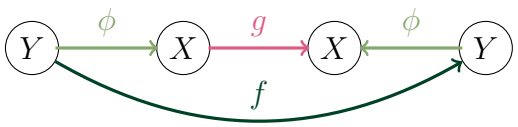


Figure 5.1: Construction for the proof of Corollary 5.1.5

Proof. Assume that Y has the fixed point property. Let X be a topological space and $g : X \to X$ a continuous function and let Y be homeomorphic to X. Therefore there is a continuous and bijective function $\phi : Y \to X$, whose inverse function is also continuous.

Now we define another function $f = \phi^{-1} \circ g \circ \phi : Y \to Y$. The construction can be seen in Figure 5.1.

f in fact is continuous as a composition of continuous functions and has according to our assumption a fixed point x_0. Because ϕ is bijective, it holds that $\phi(x_0)$ is a fixed point for the function g. So X has a fixed point property. □

Thus, if we find one space with the fixed point property, all the homeomorphic spaces have the fixed point property, too. Let us look at an example for this.

Example 5.1.6. Let $a, b\mathbb{R}$ with $a < b$. The interval $[a, b] \subset \mathbb{R}$ has the fixed point property.
Consider the function $f : [0, 1] \to [a, b]$ with $x \mapsto (b - a)x + a$. Since this function is of course bijective and continuous, this is a homeomorphism. As we can see in Example 5.1.4, $[0, 1]$ has the fixed point property. Hence according to Corollary 5.1.5, so does $[a, b]$.

Corollary 5.1.5 is already a pretty powerful statement. Nevertheless, we cannot prove the existence of the Nash equilibrium with this yet. For the proof of Theorem 4.1.4, we need a procedure to find fixed points or rather conditions, in which fixed points occur. This is provided by the fixed point theorems of Brouwer and Kakutani.

5.2 Brouwer Fixed-Point Theorem

Now let us imagine we are enjoying our morning coffee or tea[24] before we devote ourselves to the joy of mathematics. We are looking at the liquid in our mug while wondering what happens to the coffee or tea while stirring. The Brouwer Fixed Point Theorem is a mathematical concept that surprisingly connects our coffee (or tea) mug to the world of topology and fixed points.

[24] or in fact any other beverage

The Brouwer Fixed Point Theorem was proven by the Dutch mathematician Luitzen Egbertus Jan Brouwer in 1911 in [24]. In its initial form, this theorem was defined by Brouwer in [24] within the closed unit ball $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$[25] of the Euclidian space $\mathbb{R}^n$.

Theorem 5.2.1 (Brouwer fixed-point theorem in its initial form)**.** *Every continuous function $f : D^n \to D^n$ has a fixed point.*

We will not prove the theorem in its original form at, but only deal with the more general form.

Nevertheless, we can briefly discuss why a generalisation from the closed unit ball to all compact convex subspaces of $\mathbb{R}^n$ is possible. That is because of Corollary 5.1.5. If D^n has the fixed point property, so do all subsets of $\mathbb{R}^n$ which are homeomorph to D^n. We will see later on that in fact all convex compact subsets of $\mathbb{R}^n$ are homeomorph to D^n.

Thus, our first statement can be generalised. We will actually be able to prove the next generalised statement, which we get from [19, p. 41].

Theorem 5.2.2 (Brouwer fixed-point theorem)**.** *Every continuous function from a standard simplex to itself has a fixed point.*

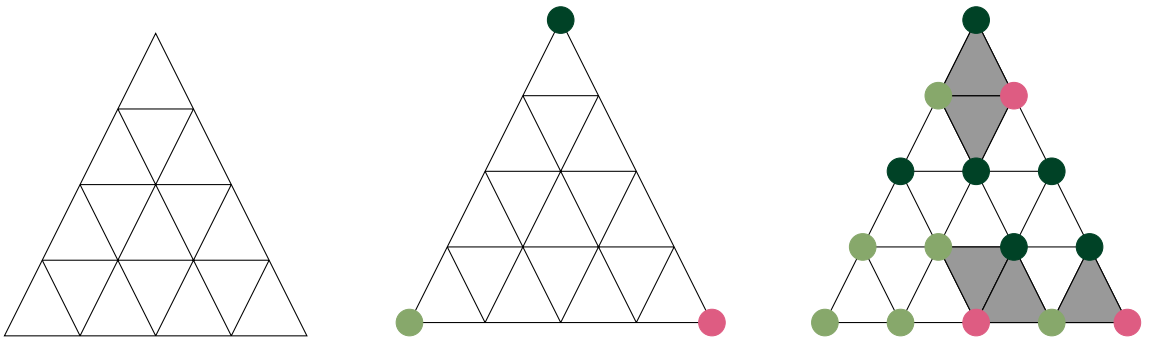
To prove Theorem 5.2.2, we again need some more information. This time it is based on a simple combinatorial lemma. This lemma was established by Emanuel Sperner in [25]. However, this little lemma is as powerful as it is simple. Here we will stick to [22, pp. 223–225] and [14].

Let us explore a particular case of Sperner's lemma before extending it to a more general context. Consider a triangle denoted as W, which is subdivided into smaller triangles, as illustrated in Figure 5.2a. These may be regarded as *elementary triangles* [14, p. 930].

Now, we assign colours to the vertices of the larger triangle according to the following rules: Each corner vertex of the triangle must be coloured with a unique colour, ensuring that the colours at these corners are different from one another. This is illustrated in Figure 5.2b [14, p. 930].

Along the edges of triangle W, the colours of the vertices must match the colours of the adjoining corner vertices. In other words, the colours along any edge must be drawn from the set of colours assigned to the corner vertices spanning that edge. As for the interior of triangle W, the colours assigned to its vertices are arbitrary [14, p. 930].

[25] which norm we consider here is arbitrary, as all norms in $\mathbb{R}^n$ are equivalent



(a) triangles (b) colouring of the triangle (c) labelled triangles (shaded grey)

Figure 5.2: Sperner's Lemma for dimension 2

These two rules are to be called *Sperner's labelling.*

Next, we identify and mark the smaller triangles within W where all three colours are present. Figure 5.2c provides an example of these marked triangles.

In summary, this construction represents a specific instance of Sperner's lemma. It asserts that in any triangulation of a triangle where vertices are coloured with distinct colours, there must be at least one smaller simplex whose vertices are coloured with all three of these colours.

Considering this construction, Sperner came up with the following statement [14, p. 930].

Lemma 5.2.3 (Sperner's Lemma for Triangles)**.** *Any Sperner-labelled triangulation of a triangle W must contain an odd number of elementary triangles which contain all three colours. In particular, there is at least one.*

In fact – as we mentioned – Lemma 5.2.3 is only a specified case. We have to look at the portrayed circumstances with a much broader view. Why stick to triangles? What about points, lines or tetrahedrons? To expand our theory, some more terminology is needed. Firstly, we have to define the environment in which we are going to operate: the d-simplices [14, p. 931].

Definition 5.2.4 (d-simplex)**.**

1. A *d-simplex* or a *d-dimensional simplex* $\Lambda \in \mathbb{R}^k$ may be regarded as the convex hull $[\![y^0, y^1, \dots y^d]\!]$ of $d+1$ affinely independent points in $\mathbb{R}^k$, for $k \geq d$.

2. Every y_j is called a *vertex* of Λ.

3. For a $L \subset \{0, \dots, d\}$ $[\![y^l : l \in L\,]\!]$ with $r = \#L - 1$ is called a *r-face.*

4. Any face that spans exactly d of the $d+1$ vertices of Λ is called a *facet.*

Thus a 0-simplex is equal to a point, a 1-simplex to a line, a 2-simplex to a triangle, a 3-simplex to a tetrahedron and so on. These simplices now have to be triangulated. Again we use [14, p. 931] to define this concept.

Definition 5.2.5 (Triangulation of a d-simplex Λ)**.** A *triangulation* of a d-simplex Λ is considered to be a set of smaller d-simplices $\mathcal{L}$. These simplices will be called *elementary simplices*. Their vertices will be called *vertices of the triangulation*. The union of all those simplices is Λ as seen in Equation 5.1.

$$\Lambda = \bigcup_{\lambda \in \mathcal{L}} \lambda \tag{5.1}$$

Plus, we need the property that the intersection of two of the smaller simplices is empty or the complete $(d-1)$-face common to both.

Now we need the rules for the labelling to be generalised for any d-simplex like in [14, p. 931].

Definition 5.2.6 (Sperner labelling)**.** A *Sperner labelling* or rather colouring of a triangulation of a d-simplex Λ is an assignment of $d+1$ colours to the vertices of the triangulation. This must meet the following conditions: the vertices of Λ must all have different colours and the vertices of the facet l, which lies opposite the vertex with the colour l, of Λ do not use the colour l. The interior vertices can be coloured in any way.

Now we are ready to state the d-dimensional Sperner's lemma and prove it as described in [14, 932f] and [22, p. 225].

Lemma 5.2.7 (d-dimensional Sperner's Lemma)**.** *Any Sperner-labelled triangulation of a d-simplex must contain an odd number of elementary simplices whose vertices all have different colours. We call this a rainbow simplex. In particular, there is at least one.*

Proof. We prove this theorem by induction over the dimension d.

For $d = 1$ the following holds: A triangulated 1-simplex is a segmented line (a, b). The two endpoints a and b have different colours. Therefore the colour must change at least one time along the line. If we add another colour change, the colour will have to change back. Hence, we have an odd number of colour changes.

Assume now that the theorem is fulfilled for all dimensions up to $d-1$.

In the induction step we show that the theorem holds for a triangulated, Sperner-labelled d-simplex using $d+1$ colours. Suppose this triangulated, Sperner-labelled d-simplex is called Λ and we are using $d+1$ colours.

Let R define the number of our so called rainbow simplices which use all of the $d+1$ colours and K the number of elementary simplices which are only labelled with the colours $\{1,\ldots,d\}$. This means there is no colour $d+1$ in K. So because the simplices have dimension d, exactly one of these colours is used twice and the other colours is used once.

Furthermore, we look at the $(d-1)$-dimensional faces or facets of the elementary simplices which use exactly the colours $\{1,\ldots,d\}$. Let B denote the number of these faces on the boundary of Λ, I the number of such faces inside of Λ.

Now we claim the following: Each simplex of type R contributes exactly one face coloured $\{1,\ldots,d\}$ plus each simplex of type K contributes exactly two faces coloured $\{1,\ldots,d\}$. Why is that? Every elementary simplex with at least one facet with the colours $1,\ldots,d$ is either a rainbow simplex (this means there is only one such facet because no colour repeats itself) or a simplex with exactly $1,\ldots,d$ colours. In the second case, there is one colour which repeats itself. Thus, there are exactly two facets with the colours $\{1,\ldots,d\}$ of these elementary simplices.

Plus we can observe that the inside faces of the type I appear in two simplices while the boundary faces of the type B appear in just one simplex. Hence, we get the Equation 5.2.

$$2K + R = B + 2I \tag{5.2}$$

Here we make a second claim: B is odd. Why is that? The only $(d-1)$-dimensional facets of the elementary simplices on the boundary of Λ coloured $\{1,\ldots,d\}$ can be on the facet of Λ whose vertices are coloured $\{1,\ldots,d\}$. This is because of the Sperner labelling. Here we use the inductive hypothesis for this facet of Λ, which forms a Sperner-labelled $(d-1)$-dimensional simplex. According to the induction hypothesis, the facet contains an odd number of rainbow $(d-1)$-dimensional simplices. Hence, B is odd.
Because of Equation 5.2, we conclude that R is odd as well. □

We proved Sperner's lemma in an combinatorical way. We did not actually use any high-level mathematics here. Instead, we just drew some pictures and thought about the implications. And now we got this really nice claim.

Let us try to return to Brouwer's Fixed-Point Theorem and use our knowledge of Sperner's Lemma to prove it. Again, we will stick to [22], although we will present the proof in a bit more detail which we gained from [14, 932f].

Proof. (of Theorem 5.2.2)
Let Λ be a d-dimensional standard simplex and suppose $g : \Lambda \to \Lambda$ is a continuous function. Because Λ is a d-dimensional simplex, for every $x = (x_0, x_1, \dots, x_d) = \sum_{j=0}^{d} x_j v_j \in \Lambda$ it holds that $\sum_{j=0}^{d} x_j = 1$ and $x_j \geq 0$ for all $j \in \{0, \dots, d\}$ with $v_1, \dots, v_d$[26] being the vertices of our simplex Λ.

Consider a sequence of triangulations such that each triangulation triangulates every simplex of the previous triangulation. Then the distances between the vertices of each elementary simplex converge to zero.

Suppose there are no fixed points. So for all $x \in \Lambda$ it holds that $g(x) \neq x$.

Then there is at least one $j* \in \{1, \dots, d\}$ with $g(x)_{j*} \neq x_{j^*}$ and therefore $g(x)_{j^*} < x_{j^*}$ for at least one j^*. To show that the last statement is true, we have to remember that g maps Λ into itself and that Λ is a d-simplex. So the following holds: $\sum_{j=0}^{d} x_j = \sum_{j=0}^{d} g(x)_j = 1$. Therefore if $g(v) \neq v$, there must at least be $k_1, k_2 \in \{0, \dots, d\}$ with $g(x)_{k_1} < x_{k_1}$ and $g(x)_{k_2} > x_{k_2}$.

Now that we know, there is always a $j* \in \{0, \dots, n\}$ with $g(x)_{j^*} < x_{j^*}$ for any point of any triangulation of Λ, we colour it by the colour j^*.

Is this colouring a Sperner labelling? The answer to this question is yes! Every vertex of Λ v_i will be coloured by the colour i. Why is that? The only j^* for which $g(v)_{j^*} < v_{j^*}$ is – due to the definition of v_i as the i^{th} unit vector – $j^* = i$. So every vertex of Λ is coloured in another colour. This fact satisfies the first criterion of a Sperner-labelling.

Now let us look at a vertex of the triangulation that lies on a r-face of Λ. These are formed by all but $d + 1 - r$ of the vertices of Λ. We call these unused vertices $v_{l_1}, \dots, v_{l_{d+1-r}}$. So the l_ith is 0 for all $i \in \{1, \dots, d+1-r\}$. So it is in fact possible that $g(x)_{j^*} < x_{j^*}$, as long as $j^* \neq l_1, \dots, l_{d+1-r}$. It follows that any point of a triangulation on an r-face of Λ is coloured by one of the $k + 1$ colours of the vertices of that face. This satisfies the second criterion of a Sperner-labelling.

Hence we know that our constructed labelling is in fact a Sperner-labelling and so Sperner's lemma is satisfied.

This means that there must exist at least one rainbow simplex R_m in every triangulation – where m points to the m^{th} triangulation. Let us call the i^{th} vertex of R_m $v^{m,i}$. Because Λ is compact, the sequence $(v^{m,1})_{m \in \mathbb{N}}$ has a convergent subsequence. So we can assume that the sequence converges to a $v^* \in \Lambda$. Note here that because of the infinite triangulation – where the distances between the vertices are going smaller and

[26] with $v_i = (0, \dots, 0, 1, 0, \dots)$ where the 1 occurs in the i^{th} component

smaller – the sequence $(v^{m,i})_{m\in\mathbb{N}}$ converges to the same point v^* for all i. Therefore for all i that the following holds true:

$$\lim_{m\to\infty} v^{(m,i)} = v^*. \tag{5.3}$$

As we have seen before, for any i, there is a colour j_i^* such that $g(v^{(m,i)})_{j_i^*} < v_{j_i^*}^{(m,i)}$. Let us remember that $\{v^{(m,i)} \mid i \in \{1,\dots,d+1\}\}$ is the set of all vertices of a rainbow cell for any m. Therefore j_i^* takes on all the colours $0, 1, 2, \dots, d$. Since g is continuous, the following holds:

$$\begin{aligned} g(v^{(m,i)})_{j_i^*} &< v_{j_i^*}^{m,i} \\ \lim_{m\to\infty} g(v^{(m,i)})_{j_i^*} &\leq \lim_{m\to\infty} v_{j_i^*}^{(m,i)} \\ g(\lim_{m\to\infty} v^{(m,i)})_{j_i^*} &\leq \lim_{m\to\infty} v_{j_i^*}^{m,i} \\ g(v^*)_{j_i^*} &< v_{j_i^*}^* \end{aligned} \tag{5.4}$$

So we have $g(v^*)_{j_i^*} < v_{j_i^*}^*$ for all $j_i^* \in \{0,1,\dots,d\}$. But let us once again remember that by the definition of the simplex we have: $\sum_{j=0}^{d} v_j = \sum_{j=0}^{d} g(v)_j = 1$. This leads to a contradiction to $g(v^*)_{j_i^*} < v_{j_i^*}^*$ and therefore g must have at least one fixed point. □

Actually it can be shown that every compact convex subset of $\mathbb{R}^n$ has the fixed-point property. This is, what we want to do next. For this we use [26, pp. 49–53].

To prove this, we need to remember what an interior point is. u is an interior point if there is a radius $r > 0$ such that $\|u\| \leq r$ implies $u \in M$ [13, p. 93]. We denote this by $u \in int(M)$.

Definition 5.2.8. We define a so called *Minkowski functional*. For each $u \in \mathbb{R}^n$, the Minkowski functional of the compact convex set $M \subset \mathbb{R}^n$ with 0 as an interior point is:

$$p : \mathbb{R}^n \to \mathbb{R}_{\geq 0};\ u \mapsto \inf\{\lambda \mid \lambda^{-1}u \in M, \lambda > 0\}. \tag{5.5}$$

Intuitively, for any $u \in \mathbb{R}^n$, the ray through the point u and the origin intersects the boundary of M at the point $p(u)^{-1}u$. This intuition is depicted in Figure 5.3.

Now we are ready to show some important properties of the Minkowski functional.

Lemma 5.2.9. *Let p be a Minkowski functional and M a nonempty, convex and compact subset of $\mathbb{R}^n$, where M has 0 as an interior point. We want to show the following:*

1. $a\|u\| \leq p(u) \leq b\|u\|$ for all $u \in \mathbb{R}^n$ and fixed $a, b > 0$,

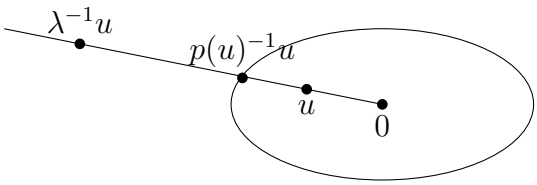


Figure 5.3: Illustration of the Minkowski functional

2. $p(\alpha u) = \alpha p(u)$ for all $a \geq 0$,

3. $p(u+v) \leq p(u) + p(v)$ for all $u, v \in \mathbb{R}^n$,

4. $p : \mathbb{R}^n \to \mathbb{R}^n$ *is continuous,*

5. *For* $C = \{u \in \mathbb{R}^n \mid p(u) \leq 1\}$ *it holds that* $C = M$.

Proof. 1. We know that $0 \in int(M)$. Therefore there is a radius $r > 0$ such that $\|u\| \leq r$ implies $u \in M$. Plus, we can state $p(0) = 0$. Now let us choose a $x \in \mathbb{R}^n$ with $x \neq 0$. Then choose a λ with $\|\lambda^{-1}x\| = r$. Hence $\lambda = r^{-1}\|x\|$. With $\|\lambda^{-1}x\| = r$ we know that $\lambda^{-1}x \in M$. Reconsidering our definition of the Minkowski functional we get:

$$p(x) \leq r^{-1}\|x\| \text{ for all } x \in \mathbb{R}^n. \tag{5.6}$$

With M being compact, we know that M is bounded and therefore:

$$\|u\| \leq R \text{ for all } u \in M \text{ and a } R > 0. \tag{5.7}$$

Thereby we know:

$$\begin{aligned} \|p(u)^{-1}u\| = p(u)^{-1}\|u\| &\leq R \\ \Leftrightarrow p(u)^{-1} &\leq R\|u\|^{-1} \\ \Leftrightarrow p(u) &\geq R^{-1}\|u\|. \end{aligned} \tag{5.8}$$

Hence it holds $R^{-1}\|u\| \leq p(u) \leq r^{-1}\|u\|$ for all $u \in \mathbb{R}^n$.

2. For $\alpha = 0$ the equality is clear. So let us choose a $\alpha > 0$. It holds that:

$$\begin{aligned} p(\alpha u) &= \inf\{\lambda \mid \lambda^{-1}\alpha u \in M, \lambda > 0\} \\ &= \inf\{\alpha\alpha^{-1}\lambda \mid \lambda^{-1}\alpha u \in M, \lambda > 0\} \\ &= \alpha \inf\{\alpha^{-1}\lambda \mid \frac{u}{\frac{\lambda}{\alpha}} \in M, \lambda > 0\} \\ &= \alpha \inf\{\lambda \mid \lambda^{-1}u \in M, \lambda > 0\} = \alpha p(u). \end{aligned} \tag{5.9}$$

3. Let $u, v \in \mathbb{R}^n$ and consider an $\epsilon > 0$ and choose λ_1 and λ_2 with

$$p(u) < \lambda_1 < p(u) + \frac{\epsilon}{2} \text{ and } p(v) < \lambda_2 < p(v) + \frac{\epsilon}{2}. \tag{5.10}$$

Therefore it must hold that $\lambda_1^{-1}u, \lambda_2^{-1}v \in M$. And it follows that

$$\lambda_1 + \lambda_2 < p(u) + p(v) + \epsilon. \tag{5.11}$$

With the convexity of M we know:

$$\begin{aligned}(\lambda_1 + \lambda_2)^{-1}(u + v) &= (\lambda_1 + \lambda_2)^{-1}u + (\lambda_1 + \lambda_2)^{-1}v \\ &= (\lambda_1 + \lambda_2)^{-1}\lambda_1 \underbrace{(\lambda_1^{-1}u)}_{\in M} + (\lambda_1 + \lambda_2)^{-1}\lambda_2 \underbrace{(\lambda_2^{-1}v)}_{\in M} \in M.\end{aligned} \tag{5.12}$$

Therefore we know by the definition of p:

$$p(u + v) \leq \lambda_1 + \lambda_2 < p(u) + p(v) + \epsilon. \tag{5.13}$$

Since our $\epsilon > 0$ was arbitrary:

$$p(u + v) \leq p(u) + p(v). \tag{5.14}$$

4. This follows from the third property. Suppose again that $u, v \in \mathbb{R}^n$. The following holds:

$$\begin{aligned}p(u) &= p(v + (u - v)) \leq p(v) + p(u - v) \\ p(v) &= p(u + (v - u)) \leq p(u) + p(v - u).\end{aligned} \tag{5.15}$$

Now we have to differentiate two cases:
First case ($p(u) \geq p(v)$): In this case it holds that

$$|p(u) - p(v)| \leq |p(v) + p(u - v) - p(u)| = |p(u - v)| = p(u - v). \tag{5.16}$$

Second case ($p(u) < p(v)$): In this case it holds that

$$|p(u) - p(v)| \leq |p(u) - p(u) - p(v - u)| = |p(v - u)| = p(v - u). \tag{5.17}$$

If we combine these two cases we can conclude the following by using the first property from Lemma 5.2.9

$$|p(u) - p(v)| \leq \max\left(p(u - v), p(v - u)\right) \leq b\|u - v\| \text{ for all } u, v \in \mathbb{R}^n. \tag{5.18}$$

Therefore p is Lipschitz continuous [13, p. 151]. Thus p is continuous [13, p. 151].

5. Choose a $u \in C$. This means that $p(u) \leq 1$. If $u = 0$, then obviously $u \in M$. Now assume $u \neq 0$. Because of the first property it holds that $p(u) > 0$. Hence $u_n = \frac{u}{1+\frac{1}{n}} \in M$ for all $n \in \mathbb{N}$. So we can state that $u_n \to u$ and since M is closed it holds that $u \in M$. Therefore $C \subset M$. Now choose a $u \in M$. M is convex so $\alpha u \in M$ for any $\alpha \in [0, 1]$. Therefore $\lambda^{-1}u \in M$ for all $\lambda \geq 0$. And thus $p(u) \leq 1$. Through this we know that $M \subset C$.

□

Now we are ready to show the following.

Proposition 5.2.10. *Every nonempty, compact and convex subset $C \subset \mathbb{R}^n$ with an interior point u_0 is homeomorphic to the closed unit ball $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$.*

Proof. We assume here that $0 \in int(M)$. Otherwise we simply replace x with $x - x_0$. And now we may assume $u_0 = 0$.

Let us look at two functions:

$$\begin{aligned} f : \mathbb{R}^n \longrightarrow \mathbb{R}^n;\ x \mapsto \begin{cases} \frac{p(u)}{\|u\|}u & u \in \mathbb{R}^n,\ u \neq 0 \\ 0 & u = 0 \end{cases} \\ g : \mathbb{R}^n \longrightarrow \mathbb{R}^n;\ x \mapsto \begin{cases} \frac{\|u\|}{p(u)}u & u \in \mathbb{R}^n,\ u \neq 0 \\ 0 & u = 0 \end{cases} \end{aligned} \tag{5.19}$$

We want to show that f is a homeomorphism. After that we will restrict that function to $f : C \to D^n$. By this we have a homeomorphism from a nonempty, compact and convex subset to the closed unit ball.

1. First, we need to show that f is continuous. If $u \neq 0$ we know by the fourth property in Lemma 5.2.9 that $p(u)$ is continuous. And since the identity and the norm functions are continuous on $\mathbb{R}^n\{0\}$, f is continuous for $u \neq 0$. Therefore for every sequence with $u_n \in M$ and $u_n \to u$ it holds that $f(u_n) \to f(u)$ [13, p. 145].
   Now we show that f is continuous for $u = 0$. For this let $u_n \to u$ as $n \to \infty$. It follows $f(u_n) \to 0$. This is because of property 1 of Lemma 5.2.9. Therefore we have:

$$\|f(u)\| = \|\frac{p(u)}{\|u\|}u\| = \frac{p(u)}{\|u\|}\|u\| = p(u) \leq b\|u\| \text{ for all } u \in \mathbb{R}^n. \tag{5.20}$$

   Hence f is continuous.

2. Now we have to show that f is a bijective function. We will show this by proving that f and g are inverse. If $u = 0$ we have: $f(g(u)) = f(0) = 0$ and $g(f(u)) = g(0) = 0$. Therefore f and g are inverse if $u = 0$. When $u \neq 0$, we can see that

$$f(g(u)) = f(\frac{\|u\|}{p(u)}u) = \frac{p(\frac{\|u\|}{p(u)}u)}{\|\frac{\|u\|}{p(u)}u\|}\frac{\|u\|}{p(u)}u = \frac{\frac{\|u\|}{p(u)}p(u)}{\frac{\|u\|}{p(u)}\|u\|}\frac{\|u\|}{p(u)}u = u \tag{5.21}$$

and

$$g(f(u)) = g(\frac{p(u)}{\|u\|}u) = \frac{\|\frac{p(u)}{\|u\|}u\|}{p(\frac{p(u)}{\|u\|}u)}\frac{p(u)}{\|u\|}u = \frac{\frac{p(u)}{\|u\|}\|u\|}{\frac{p(u)}{\|u\|}p(u)}\frac{p(u)}{\|u\|}u = u. \tag{5.22}$$

Thus f and g are inverse and therefore must be bijective.

3. The last property we have to show is that g is continuous. If $u \neq 0$ we know again by the fourth property in Lemma 5.2.9 that $p(u)$ is continuous. And since the identity and the norm functions are continuous on $\mathbb{R}^n\{0\}$, g is continuous a $u \neq 0$. Therefore for every sequence with $u_n \in M$ and $u_n \to u$ it holds that $g(u_n) \to g(u)$ [13, p. 145].
Now we show that g is continuous an $u = 0$. For this let $u_n \to u$ as $n \to \infty$. It follows $g(u_n) \to 0$. This is again because of property 1 of Lemma 5.2.9. Therefore we have:

$$\|g(u)\| = \|\frac{\|u\|}{p(u)}u\| = \frac{\|u\|}{p(u)}\|u\| \leq b\|u\| \text{ for all } u \in \mathbb{R}^n. \tag{5.23}$$

And hence g is continuous.

4. With the fifth property in Lemma 5.2.9 and the definition of f we have that $f(C) \subseteq D^n$. Plus we know from the second property in Lemma 5.2.9 that $g(D^n) = f^{-1}(D^n) \subseteq C$. Therefore we have $f(M) = D^n$. Hence the restriction $f : C \to D^n$ is also homeomorph.

Now we have constructed a homeomorphism between M and D^n and we are done. □

We can generalise this statement.

Corollary 5.2.11. *Let C_1 and C_2 be any nonempty, compact and convex subsets of $\mathbb{R}^n$. Then C_1 is homeomorphic to C_2.*

Proof. Proposition 5.2.10 has shown that there exist homeomorphisms $f_1 : C_1 \to D^n$ and $f_2 : C_1 \to D^n$. The composition $f_2 \circ f_1^{-1} : C_1 \to C_2$ is again a homeomorphism and so the result follows immediately. □

And now we are ready to generalise Theorem 5.2.2 to compact and convex sets.

Theorem 5.2.12 (Brouwer fixed point theorem for compact and convex sets)**.** *Let $C \subset \mathbb{R}^n$ be a compact and convex set. Then every continuous function $f : C \longrightarrow C$ has a fixed point.*

Proof. This result follows directly through Corollary 5.2.11 and Corollary 5.1.5. We know with Theorem 5.2.2 that the standard simplex $\mathbb{S}_k \subset \mathbb{R}^n$ has the fixed point property. By Lemma 2.4.3 we know that these simplices are compact, convex and non-empty. With Corollary 5.2.11 there is a homeomorphism $f : C \to \mathbb{S}_k$. Therefore by Corollary 5.1.5, C has the fixed point property, too. □

Now we know that after stirring our morning coffee or tea – which can be considered a convex compact subset of $\mathbb{R}^3$, there is always one particle of the liquid that stays in exactly the same place. So this is a fascinating example of how abstract mathematics can have unexpected real-world implications, even in something as simple and everyday as enjoying our morning coffee!

5.3 Kakutani's Gereralised Fixed Point Theorem

In 1941 Shizuo Kakutani generalised Brouwer's fixed point theorem in [27]. David Gale actually suggested the proof with Kakutani's fixed point theorem to Nash [16, p. 49]. In this section we mostly will follow [27].

Since Kakutani's Theorem is an extension of the Brouwer Fixed Point Theorem we have to look at the generalisation part. This expresses itself in multifunctions, firstly. The definition we take from [19, p. 65]

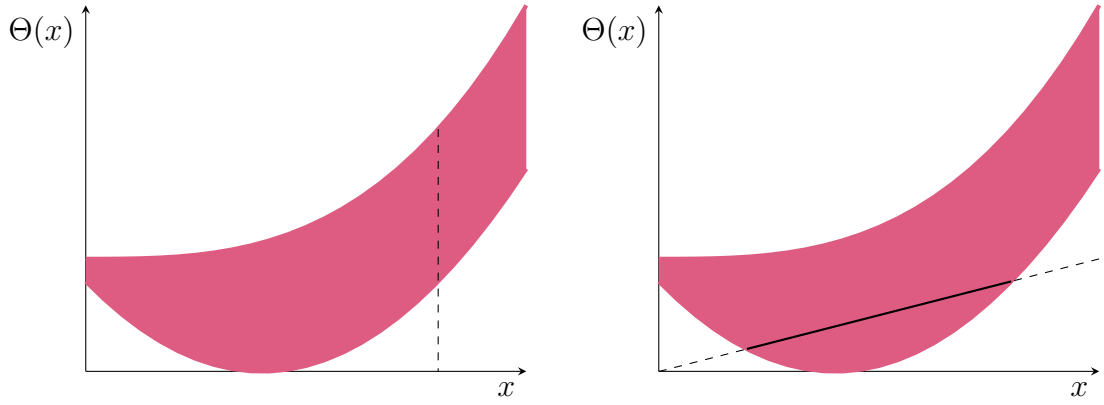
Definition 5.3.1 (Multifunction)**.** A *multifunction*[27] Θ from one set Y to a set Z is a function Θ defined on Y whose values are nonempty subsets of Z.

Remark 5.3.2. We shall denote such a multifunction with $\Theta : Y \rightrightarrows Z$.

Remark 5.3.3. Note that now instead of saying $y \in Y$ is a fixed point of $\Theta : Y \rightrightarrows Y$, we can say $y \in \Theta(y)$.

For example a multifunction $\Theta : \mathbb{R} \to \mathbb{R}$ can look like Figure 5.4a. We see here that $\Theta(\tilde{x})$ has not only one, but many values. Plus, we can explore the fixed points of our multifunction in Figure 5.4a. For that purpose we add the identity function to Figure 5.4a. And hence we get Figure 5.4b. In this example we have infinitely many fixed points, which are marked with the thick black line.

[27] In more recent literature multifunctions are often regarded as correspondences or set-valued functions, but we will stick to the this term.



(a) multiple values (b) fixed points

Figure 5.4: Example of a mulitfunction

Example 5.3.4. Now we are ready to look at more specified examples.

1. Let us consider the function $f : \mathbb{R}_+ \rightrightarrows \mathbb{R},\ x \mapsto \sqrt{x}$. $f(4)$ does in fact have two possible solutions: $\{-2, 2\}$, which is a subset of $\mathbb{R}$. Hence f is a multifunction.

2. Now we come back to our definition of best response from Equation 4.7. We can actually interpret this concept as a multifunction.

$$C_k(-\tilde{m}_k) = \{m \in \mathbb{S}_{e_k} \mid p_j(m_k, -\tilde{m}_k) \geq p_j(l_k, -\tilde{m}_k), \forall l_k \in \mathbb{S}_{e_k}\} \tag{5.24}$$

   Like in [4, p. 42] we can define another function. Suppose the following: $\mathbb{S} = \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$. Now let s be in $\mathbb{S}$ and $C(s) = C_1(-m_1) \times \cdots \times C_n(-m_n)$. Hence C is a function from $\mathbb{S}$ to $\mathbb{S}$ and the values are nonempty subsets of $\mathbb{S}$. Thus C is a multifunction. This will be important later on.

Remark 5.3.5. In this section we will now concentrate on functions where $\Theta(Y) \subseteq \mathcal{R}(Y)$. $\mathcal{R}(Y)$ is the family of all closed convex subsets of Y. Then our multifunction Θ can be written without the double arrow: $\Theta : Y \to \mathcal{R}(Y)$.

Before we can state the Fixed Point Theorem of Kakutani, we need one more definition. The fixed point theorem of Brouwer needed a continuous function as an requirement. So Kakutani needed something similar for multifunctions. We can take the definition from [27, p. 457].

Definition 5.3.6. Let $\Theta : Y \to \mathcal{R}(Y)$ be a multifunction. The correspondence is *upper semi-continuous*[28] if, for every sequence $(y_n)_{n\in\mathbb{N}}$ in Y which converges against a y_0 and every sequence $(z_n)_{n\in\mathbb{N}}$ in $\mathcal{R}(Y)$ which converges against a z_0 with $z_n \in \Theta(y_n)$, it holds that $z_0 \in \Theta(y_0)$.

[28] In some other literature this phenomenon is often regarded upper hemi-continuous. Since Kakutani uses the term upper semi-continuous, we will stick with it, too. But note that there is difference between upper semi-continuity for correspondences and for point-to-point-functions!

To make this property clear, we explore another example, which we find in [28, p. 559].

Example 5.3.7. Let us look at the following multifunction $\Phi : [0,1] \rightrightarrows [0,1]$.

$$\Phi(x) = \begin{cases} [0,1] & \text{for } x < 1 \\ \{0\} & \text{for } x = 1 \end{cases} \tag{5.25}$$

Φ is not upper semi-continuous.
For this we consider the sequence $(y_n)_{n\in\mathbb{N}}$ with $y_n = -\frac{1}{n} + 1$ for all $n \in \mathbb{N}$ and $(z_n)_{n\in\mathbb{N}}$ with $z_n = 1$ for all $n \in \mathbb{N}$. It holds that $\lim_{n\to\infty} y_n = 1$ and $\lim_{n\to\infty} z_n = 1$. Plus, we know that $z_n \in \Phi(y_n)$, because of $1 \in [0,1]$. But $z_0 = 1 \notin \Phi(y_0) = \Phi(1) = \{0\}$ shows us that Φ is not upper semi-continuous.

Kakutani writes in [27, p. 457]: "It is easy to see [upper semi-continuity] is equivalent to saying that the graph of Θ [...] is a closed set". But let us dig a little bit deeper into why this holds true. First we have to ask, what a closed graph is. We are able to take this from [27, p. 457].

Definition 5.3.8. The correspondence $\Theta : Y \rightrightarrows Z$ has a closed graph if and only if $graph(\Theta) = \{(y,z) : y \in Y, z \in \Theta(y)\}$ is closed.

This can also be expressed with sequences, to match our definition of upper semi-continuity. Here we get the inspiration from [29, p. 4] and apply this to our special case.

Lemma 5.3.9. *The graph of a correspondence $\Theta : Y \rightrightarrows Z$ $graph(\Theta) = \{(y,z) : y \in Y, z \in \Theta(y)\} \subset Y \times Z$ is closed if and only if for any sequences $((y_t),(z_t))_{t\in\mathbb{N}}$ in $graph(\Theta)$ with $y_t \to y^*$ and $z_t \to z^*$ it holds that $(y^*, z^*) \in graph(\Theta)$.*

Proof. We prove both directions.
"$\Rightarrow$": Suppose $(y^*, z^*) \in (Y \times Z) \backslash graph(\Theta)$. $(Y \times Z) \backslash graph(\Theta)$ is open. Therefore there exists a $\epsilon > 0$ with $B((y^*, z^*), \epsilon) \subseteq (Y \times Z) \backslash graph(\Theta)$. Since the sequence converges, all but a finite number of sequence elements are in $B((y^*, z^*), \epsilon) \subseteq (Y \times Z) \backslash graph(\Theta)$. This leads to a contradiction.

"$\Leftarrow$": Now suppose that for any sequences $((y_t),(z_t))_{t\in\mathbb{N}}$ in $graph(\Theta)$ with $y_t \to y^*$ and $z_t \to z^*$ it holds that $(y^*, z^*) \in graph(\Theta)$. And assume that $graph(\Theta)$ is not closed. Therefore $(Y \times Z) \backslash graph(\Theta)$ is not open. Hence there is a $(u,v) \in (Y \times Z) \backslash graph(\Theta)$ with the following property: There is no ϵ with $B((y^*, z^*), \epsilon) \subseteq (Y \times Z) \backslash graph(\Theta)$. Now let us consider $\epsilon_t = \frac{1}{t}$. Since $B((y^*, z^*), \epsilon) \not\subseteq (Y \times Z) \backslash graph(\Theta)$, there is a $(y_t, z_t) \in graph(\Theta)$ with $(y_t, z_t) \in B((y^*, z^*), \epsilon_t)$. Then it holds that $y_t \to y^*$ and $z_t \to z^*$ and $(y_t, z_t) \in graph(\Theta)$, but $(y^*, z^*) \notin graph(\Theta)$. This again leads to a contradiction. □

Now we have exactly the same definition for closed graph and upper semi-continuity. Thus it is enough to only show the closed graph property when wanting to prove upper semi-continuity.

Remark 5.3.10. We note that the Definition 5.3.6 is of course only valid for $\Theta(y) \in \mathcal{R}(Y)$ for all $y \in Y$, where $\mathcal{R}(Y)$ is referred to all closed convex subsets of Y.

Now we are finally ready to define the Fixed Point Theorem of Kakutani as in [27, p. 457].

Theorem 5.3.11 (Kakutani's set-valued fixed-point theorem)**.** *Let Λ be a d-dimensional closed simplex in $\mathbb{R}^n$. Suppose $\Theta : \Lambda \longrightarrow \mathcal{R}(\Lambda)$ is a multifunction with a closed graph and $\Theta(x)$ is non-empty and convex for all $x \in \Lambda$. Then Θ has a fixed point.*

Remark 5.3.12. We note that the Brouwer fixed-point theorem is a special case of the Kakutani fixed-point theorem, when each $\Theta(x)$ consists only of one point $g(x)$. In this case the compactness and the convexity is nothing but the continuity of $g(x)$.

For the proof of Theorem 5.3.11 we also need some triangulation of simplices as we have seen in the last section. Our idea to prove this Theorem is to break the multifunction down to a „normal“ function and use the Brouwer fixed point theorem of the previous section.

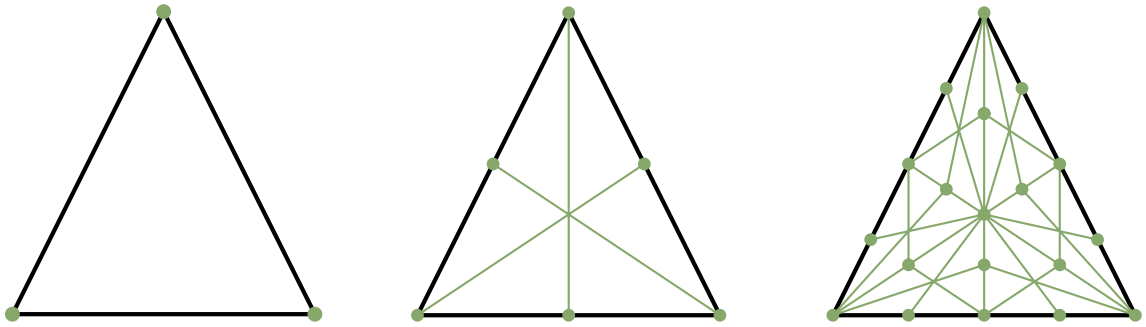
Proof. (to Theorem 5.3.11)
We are again looking at a triangulation of the d-simplex Λ. Kakutani here uses the barycentic simplical triangulation. This is an algorithm to divide a simplex into simplices with the same dimension. This is achieved by connecting the barycenters of their faces. The barycenter of an r-dimensional face is basically the middle, as defined in the following way[29]:

$$\frac{1}{(r+1)} \sum_{j=0}^{r} y_j \quad \text{with } y_j \text{ being the vertices of the face.} \tag{5.26}$$

Let Λ_k be the k^{th} barycentic simplical triangulation. The 2-dimensional case is illustrated in Figure 5.5.

Since Λ is a d-dimensional simplex, there are many smaller d-dimensional simplices after the triangulation. We are now choosing one of these smaller simplices and denote it as Ω_j and each vertex of Ω_j as $j_0, \ldots, j_d$. With the definition of the simplex we know that very point in Ω_j can be reached by $\omega_0 j_0 + \omega_1 j_1 + \cdots + \omega_d j_d$ with $\omega_k \geq 0$ for every $k \in \{0, \ldots, d\}$ and $\sum_{k=1}^{d} \omega_k = 1$.

[29] this is taken from [30, 179f]



(a) Λ_0 (b) Λ_1 (c) Λ_2

Figure 5.5: Barycentric Simplical Triangulation

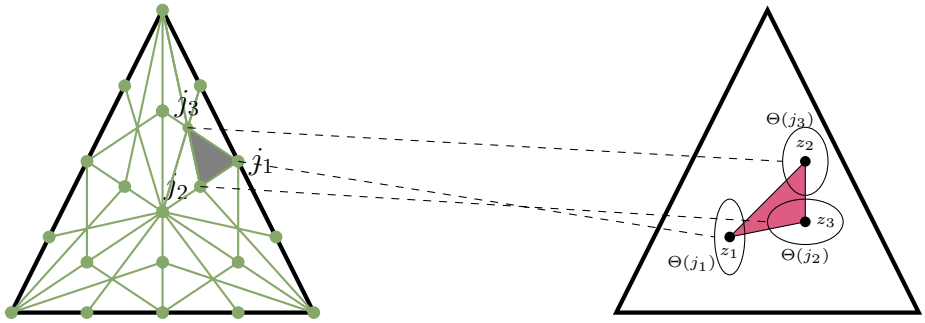


Figure 5.6: Point-to-point mapping

When we now look at any of these j_k, then $\Theta(j_k)$ is a set (since Θ is a multifunction). Now choose an arbitrary point z_j of $\Theta(j_k)$. This point also lies in a d-dimensional simplex. We call this simplex Ω_z and let it be built on the vertices $z_0, \ldots, z_r$. Again we can reach every point in Ω_z through $\omega_0 z_0 + \omega_1 z_1 + \cdots + \omega_d z_d$ with $\omega_k \geq 0$ for every $k \in \{0, \ldots, d\}$ and $\sum_{k=1}^{d} \omega_k = 1$.

Note here that we deliberately use ω_k for both Ω_j and Ω_z, because now we can construct a function[30] between Ω_j and Ω_z. Again we see the 2-dimensional case in Figure 5.6.

Now we need to do the presented construction for all other d-dimensional small simplices of the triangulation. But since it is exactly the same as above, we will do without that, assume we did that and get on with the proof. Now if we repeat our method, we have a point-to-point mapping. We will denote this as $f_k : \Lambda \to \Lambda$, where k stnads for the k^{tn} triangulation. We know that f_k is continuous. With Brouwer and Theorem 5.2.2 we can conclude that there is a fixed point $\tilde{y}_n \in \Lambda$ such that $\tilde{y}_k = f_k(\tilde{y}_k)$.

If we do this process over and over again for $\Lambda_0, \Lambda_1, \Lambda_2, \ldots$[31], there will always be fixed-points $\tilde{y}_0, \tilde{y}_1, \tilde{y}_2, \ldots$, which is an infinite sequence. Because Λ is a closed and bounded set in $\mathbb{R}^n$, we know due to the Bolzano–Weierstrass theorem[32] that the

[30] which is point-to-point
[31] barycentric triangulation can be done infinitely times
[32] the proof of this theorem can be read in [13, p. 58]

sequence $\tilde{y}_1, \tilde{y}_2, \ldots$ has a subsequence $\tilde{y}_{k_0}, \tilde{y}_{k_1}, \tilde{y}_{k_2}, \ldots$ which converges against some $\tilde{y} \in \Lambda$. Our goal now is to prove that $\tilde{y} \in \Theta(\tilde{y})$[33].

Now take a d-dimensional simplex Ω_k in the barycentric triangulation Λ_k which contains the point $\tilde{y}_k$. In fact $\tilde{y}_k$ could be part of more than just one simplex, if it lies on the vertices or the faces. In that case, we simply choose just one of the simplices in question. Let the $d+1$ vertices of Ω_k be $v_0^k, v_1^k, \ldots, v_d^k$. By the definition of a simplex we know that we can reach every point in Ω_k through $\omega_0^k v_0^k + \omega_1^k v_1^k + \cdots + \omega_d^k v_d^k$ with $\omega_j^k \geq 0$ for every $j \in \{0, \ldots, d\}$ and $\sum_{j=1}^d \omega_j^k = 1$. Hence we have $\tilde{y}_k = \sum_{j=0}^d \omega_j^k v_j^k$ for suitable ω_j^k's.

Then define a h_j with $h_j^k = f_k(v_j^k)$ for all $j \in \{0, 1, \ldots, d\}$. Hence we know that the following holds true: $h_j^k \in \Theta(v_j^k)$. This is, because $f_k(v_j^k)$ was defined to choose one element of $\Theta(v_j^k)$. Since $\tilde{y}_k$ is a fixed point of f_k, this means: $\tilde{y}_k = f_k(\tilde{y}_k) = \sum_{j=0}^d \omega_j^k h_j^k$.

We take a further subsequence of $k_0, k_1, k_2, \ldots$ and name it $k_0', k_1', k_2', \ldots$ such that $h_j^{k_0'}, h_j^{k_1'}, \ldots$ and $\omega_j^{k_0'}, \omega_j^{k_1'}, \ldots$ converge for $j = 0, 1, \ldots, d$. Here let us consider $\lim_{v\to\infty} h_j^{k_v'} = h_j^*$ and $\lim_{v\to\infty} \omega_j^{k_v'} = \omega_j^*$. It is obvious that we will have $\omega_j^* \geq 0$, $\sum_{j=0}^k \omega_j^* = 1$ and $\tilde{y} = \sum_{j=0}^d \omega_j^* h_j^*$.

Since the subsequence $\tilde{y}_{k_0}, \tilde{y}_{k_1}, \tilde{y}_{k_2}, \ldots$ converges against $\tilde{y}$ and $h_j^{k_v'} \in \Theta(v_j^{k_v'})$ and $\lim_{v\to\infty} h_j^{k_v'} = h_j^*$ for $j = 0, \ldots, d$ we must have with the upper semi-continuity of Θ: $h_j^* \in \Theta(\tilde{y})$. This implies with the convexity of $\Theta(\tilde{y})$ that $\tilde{y} = \sum_{j=0}^d \omega_j^* h_j^* \in \Theta(\tilde{y})$. And now we have our fixed point. □

Actually we can very easily generalise this statement to all bounded closed convex sets in a Euclidean space. Again we can take this from [27].

Corollary 5.3.13. *Theorem 5.3.11 is also valid even if Λ is an arbitrary bounded closed*[34] *convex set in a Euclidean space.*

Proof. Let us look at a closed simplex Δ which contains Λ as a subset. Now take a continuous retracting mapping[35] $x \mapsto \omega(x)$ of Δ into Λ. This is a point-to-point mapping. Therefore it holds that $\omega(x) = x$ for all $x \in \Lambda$ and $\omega(x) \in \Lambda$.

Plus, let $x \mapsto \Theta(x)$ be a upper semi-continuous multifunction from Λ into $\mathcal{R}(\Lambda)$.

Thus $x \mapsto \Theta(\omega(x))$ is an upper semi-continuous point-to-set mapping of Δ into $\mathcal{R}(\Lambda) \subseteq R(\Delta)$. $\mathcal{R}(\Lambda) \subset \mathcal{R}(\Delta)$ is fulfilled since $\Lambda \subset \Delta$. Therefore $x \mapsto \Theta(\omega(x))$ is also a mapping from Δ into $\mathcal{R}(\Delta)$. So we fulfil Theorem 5.3.11.

[33] of course there could be multiple $\tilde{y}$

[34] hence compact in $\mathbb{R}^m$

[35] A retracting function is a continuous mapping from a topological space into a subspace that preserves the position of all points in that subspace. This can be read in [31, p. 50].

By Theorem 5.3.11, there exists an $x_0 \in \Delta$ such that $x_0 \in \Theta(\omega(x_0))$. Plus we know $\Theta(\omega(x_0)) \subseteq \Lambda$, because we defined the mapping $x \mapsto \omega(x)$ as of Δ into Λ and $x \mapsto \Theta(x)$ as Λ into Λ. Therefore if $\Theta(\omega(x^*)) \subset \Lambda$ so $x^* \in \Theta(\omega(x^*)) \subset \Lambda$. Since $x^* \in \Lambda$ and each point of the subspace is fixed after retraction, we have $x^* = \omega(x^*)$. So we have: $x^* \in \Theta(\omega(x^*)) = \Theta(\omega(x^*)) \subset \Lambda$. Therefore there exists x^* such that $x^* \in \Lambda$ and $x^* \in \Theta(x^*)$. □

Now we showed the Kakutani fixed point theorem, which is a generalisation of Brouwer's fixed point theorem. Finally we are ready to prove the Nash equilibrium in Theorem 4.1.4. This will be our task in the next section.

6 Nash's "One-Page-Proof(s)"

Nash initially proved his theorem in his dissertation [16] with Kakutani's gereralized fixed point theorem which is quoted in Theorem 5.3.11. One year later he published in [1] another proof with the Brouwer fixed point theorem in Theorem 5.2.2. The second proof he himself considered as an improvement over the earlier version. In this section we will explore both options to prove the Nash equilibrium in Theorem 4.1.4.

Let us start with the proof of Theorem 4.1.4 with the help of Katutani's fixed point theorem. We are following the proof given in [19, 70f]. We will supplement this proof, since it is only given for two players in [19, 70f]. Plus, we will use [32, p. 29] for some properties.

Proof. (of Theorem 4.1.4 with Kakutani)
Suppose there is a point $\wp = (\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_n) \in \mathbb{S}_{e_1} \times \mathbb{S}_{e_2} \times \cdots \times \mathbb{S}_{e_n}$. Let us recall that $\wp$ is a Nash equilibrium if the following holds for each P_j:

$$\tilde{m}_j \in C_j(-\tilde{m}_j). \tag{6.1}$$

Next we define the following:

$$\begin{aligned} C : \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n} &\rightrightarrows \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n} \\ C(m_1, \ldots, m_n) &= C_1(-m_1) \times \cdots \times C_n(m_n). \end{aligned} \tag{6.2}$$

Now $\wp$ is a Nash equilibrium if and only if $\wp$ is a fixed point of the multifunction C. Let us briefly consider, why this holds. C having a fixed point means $C(m_1, \ldots, m_n) = (m_1, \ldots, m_n)$ and therefore $m_1 \in C_j(-m_1)$. This is exactly the property from Equation 6.1.

Thus if we can show the properties of Kakutani's fixed point theorem, we are finished. These are the conditions that must be met:

1. $C(\sigma)$ is convex with $\sigma \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$
2. $C(\sigma)$ is nonempty
3. $C(\sigma)$ is upper semi-continuous
4. $\mathbb{S} = \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$ is compact, convex and nonempty

Let us start with Property 4: We know already that the simplices $\mathbb{S}_{e_j}$ are compact and convex, which we proved in Lemma 2.4.3. The nonempty part means that all players

have at least one strategy, which we assume here[36]. Under the Cartesian product compactness[37] and convexity[38] are preserved, which is why property 4 is fulfilled.

Property 1: Of course $C(\sigma)$ is convex if $C_j(-\sigma_j)$ is convex – as we have seen in the property above. Let $\tilde{\sigma}_j, \overline{\sigma}_j \in C_j(-\sigma_j)$. Then we have:

$$\begin{aligned} p_j(\tilde{\sigma}_j, -\sigma_j) \geq p_j(m_j, -\sigma_j) \text{ for all } m_j \in \mathbb{S}_{e_j} \\ p_j(\overline{\sigma}_j, -\sigma_j) \geq p_j(m_j, -\sigma_j) \text{ for all } m_j \in \mathbb{S}_{e_j}. \end{aligned} \tag{6.4}$$

These equations imply that for all $\alpha \in [0,1]$ the following holds:

$$\alpha p_j(\tilde{\sigma}_j, -\sigma_j) + (1-\alpha) p_j(\overline{\sigma}_j, -\sigma_j) \leq p_j(m_j, -\sigma_j) \text{ for all } m_j \in \mathbb{S}_{e_j}. \tag{6.5}$$

With the multi-linearity of p_j we can conclude:

$$\begin{aligned} \alpha p_j(\tilde{\sigma}_j, -\sigma_j) + (1-\alpha) p_j(\overline{\sigma}_j, -\sigma_j) &= p_j(\alpha\tilde{\sigma}_j + (1-\alpha)\overline{\sigma}_j, -\sigma_j) \\ &\leq p_j(m_j, -\sigma_j) \text{ for all } m_j \in \mathbb{S}_{e_j}. \end{aligned} \tag{6.6}$$

Hence it holds that $\alpha\tilde{\sigma}_j + (1-\alpha)\overline{\sigma}_j \in C_j(-\sigma_j)$ for all $j \in \{1, \ldots, n\}$. And therefore $C(\sigma)$ is convex.

Property 2: $C_j(-m_j) = \arg\max_{x \in \mathbb{S}_j} p_j(x, -m_j)$ where $\mathbb{S}_j$ is non-empty and compact. We know that p_j is multi-linear. It follows that p_j is continuous. By Weierstrass's theorem[39] $C(\sigma)$ is non-empty.

Property 3: Since $\mathbb{S}$ is compact as seen in property 2, it is enough to show that C has a closed graph. Assume here that this is not fulfilled. So there is a sequence $(\sigma^n, \tilde{\sigma}^n)$ which converges to $(\sigma, \tilde{\sigma})$ with $\tilde{\sigma}^n \in C(\sigma^n)$, but $\tilde{\sigma} \notin C(\sigma)$. Then we know that $\tilde{\sigma}_j \notin C_j(\sigma)$ for at least one player j. Thus there is a $\alpha > 0$ and a $\overline{\sigma}_j$ such that $p_j(\overline{\sigma}_j, -\sigma_j) > p_j(\tilde{\sigma}_j, -\sigma_j) + 3\alpha$. While keeping in min, that p_j is continuous, we have the following for sufficiently large n:

$$p_j(\overline{\sigma}_j, -\sigma_j^n) \geq p_j(\overline{\sigma}_j, -\sigma_j) - \alpha > p_j(\tilde{\sigma}_j, -\sigma_j) + 2\alpha \geq p_j(\tilde{\sigma}_j^n, -\sigma_j^n) + \alpha \tag{6.7}$$

Hence $\overline{\sigma}_j$ is strictly better against σ_{-j_n} than $\tilde{\sigma}_{j_n}$. This does contradict $\tilde{\sigma}_j^n \in C_j(\sigma^n)$.

[36] otherwise the player without any strategies could simply be excluded

[37] The mathematician Tychonoff showed this, which can be read in [33, p. 109].

[38] Here we can briefly consider, why this holds. Let us look at any $(a_1, a_2), (b_1, b_2) \in \Omega_1 \times \Omega_2$ and $\lambda \in (0,1)$, where Ω_1, Ω_2 are convex. Then it holds that $a_1, b_1 \in \Omega_1$ and $a_2, b_2 \in \Omega_2$. By the convexity of Ω_1 and Ω_2, we have that

$$\lambda(a_1, a_2) + (1-\lambda)(b_1, b_2) = (\lambda a_1 + (1-\lambda)b_1, \lambda a_2 + (1-\lambda)b_2) \in \Omega_1 \times \Omega_2. \tag{6.3}$$

Therefore, $\Omega_1 \times \Omega_2$ is convex

[39] The statement of this theorem is the following: A continuous real function on a compact set takes its maximum (or minimum) in at least one point [13, p. 155].

Since all properties are fulfilled, we can reach for the fixed point theorem of Kakutani, which we have seen in Theorem 5.3.11. Thus the mapping has a fixed point.

So there is at least one equilibrium point. □

Next up is the proof of Theorem 4.1.4 with Brouwer's fixed point theorem. This is acquired from [1] with some changes for a better understanding. Additionally, we use [19, 61f].

Proof. (of Theorem 4.1.4 with Brouwer)
Let $\wp = (m_1, \ldots, m_n) \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$ be a tuple of mixed strategies. Plus, let $p_k(\wp)$ be the pay-off for player k and $p_{k,i}(\wp) = p_k(s_{k,i}, -m_k)$ be the pay-off when P_k changes their strategy from m_k to their i^{th} pure strategy $s_{k,i}$ while the others continue with their strategies. Now we define:

$$\rho_{k,i}(\wp) = max(0, p_{k,i}(\wp) - p_k(\wp)). \tag{6.8}$$

For each mixed strategy m_k of $\wp$ we define:

$$\phi_{k,i}(\wp) = \frac{m_{k_i} + \rho_{k,i}(\wp)}{1 + \sum_{i=1}^{e_k} \rho_{k,i}(\wp)} \geq 0. \tag{6.9}$$

$\phi_{k,i}$ is continuous since the denominator is always larger than 0. Additionally, we know the following:

$$\begin{aligned} \sum_{i=1}^{e_k} \phi_{k,i}(\wp) &= \frac{\sum_{i=1}^{e_k}(m_{k_i} + \rho_{k,i}(\wp))}{1 + \sum_{i=1}^{e_k} \rho_{k,i}(\wp)} \\ &= \frac{1 + \sum_{i=1}^{e_k} \rho_{k,i}(\wp))}{1 + \sum_{i=1}^{e_k} \rho_{k,i}(\wp)} = 1. \end{aligned} \tag{6.10}$$

This means we have shown that $\phi_k(\wp) = \sum_{i=1}^{e_k} \phi_{k,i}(\wp) \in \mathbb{S}_{e_k}$. Therefore $\phi(\wp) = (\phi_1(\wp), \ldots, \phi_n(\wp))$ is a continuous function as a composition of continuous functions with $\phi : \mathbb{S} = \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n} \to \mathbb{S}$.

With the proof of Theorem 5.3.11, we know that $\mathbb{S}$ is convex and compact. Through Brouwer's Fixed Point Theorem in Theorem 5.2.12 we know that ϕ has at least one fixed point. Let us call this fixed point m with $m = \phi(m)$ with $m = (v_1, \ldots, v_n) \in \mathbb{S}_{e_1} \times \cdots \times \mathbb{S}_{e_n}$.

Now we have to show that the fixed point m of the function ϕ is in fact an equilibrium point.

With the definition of a fixed point we get:

$$\begin{aligned}
& \phi(m) = m \\
\Leftrightarrow \quad & \phi_{k,i}(m) = v_{k_i} \text{ for } k \in \{1,\dots,n\} \text{ and } i \in \{1,\dots,e_k\} \\
\Leftrightarrow \quad & \frac{v_{k_i} + \rho_{k,i}(m)}{1 + \sum_{i=1}^{e_k} \rho_{k,i}(m)} = v_{k_i} \text{ for } k \in \{1,\dots,n\} \text{ and } i \in \{1,\dots,e_k\} \\
\Leftrightarrow \quad & v_{k_i} + \rho_{k,i}(m) = v_{k_i}(1 + \sum_{i=1}^{e_k} \rho_{k,i}(m)) \text{ for } k \in \{1,\dots,n\} \text{ and } i \in \{1,\dots,e_k\} \\
\Leftrightarrow \quad & \rho_{k,i}(m) = v_{k_i}(\sum_{i=1}^{e_k} \rho_{k,i}(m)) \text{ for } k \in \{1,\dots,n\} \text{ and } i \in \{1,\dots,e_k\}.
\end{aligned} \tag{6.11}$$

Now we need to distinguish two cases: $\rho_{k,i}(m)$ is equal to 0 for all $i \in \{1,\dots,e_k\}$ and $\rho_{k,i}(m)$ is not equal to 0 for at least one $i \in \{1,\dots,e_k\}$.

Let us look at the second case first:
If $\rho_{k,i}(m)$ is not equal to 0 for at least one $i \in \{1,\dots,e_k\}$, we can divide our last equation in Equation 6.11 by $\sum_{i=1}^{e_k} \rho_{k,i}(m)$ and get:

$$v_{k_i} = \frac{\rho_{k,i}(m)}{\sum_{i=1}^{e_k} \rho_{k,i}(m)}. \tag{6.12}$$

Let us divide this case in two more cases: $\rho_{k,i}(m)$ is not equal to 0 for exactly one $i \in \{1,\dots,e_k\}$ and $\rho_{k,i}(m)$ is not equal to 0 for more than one $i \in \{1,\dots,e_k\}$.

If $\rho_{k,i}(m) \neq 0$ for exactly one $i \in \{1,\dots,e_k\}$ it must hold that $\rho_{k,i} > 0$. This is because of the definition of $\rho_{k,i}$. Plus, we know $v_{k_i} = \frac{\rho_{k,i}(m)}{\sum_{i=1}^{e_k} \rho_{k,i}(m)} = 1$ for this one i. For all the other $j \in \{1,\dots,e_k\}$ with $j \neq i$ it holds that $\rho_{k,j}(m) = 0$ and therefore $v_{k_j} = \frac{\rho_{k,j}(m)}{\sum_{i=1}^{e_k} \rho_{k,i}(m)} = 0$. Hence we have $v_k = s_{k,i}$. This leads to the following contradiction with the assumption:

$$0 < \rho_{k,i}(m) = max(0, p_{k,i}(m) - p_k(m)) = max(0, p_{k,i}(m) - p_{k,i}(m)) = 0. \tag{6.13}$$

If $\rho_{k,i}(m) \neq 0$ for more than one $i \in \{1,\dots,e_k\}$, we pick one i in the following way:

$$p_{k,i}(m) = \min_{l \text{ with } \rho_{k,l}(m)>0} p_{k,l}(m). \tag{6.14}$$

Let us remember that $\rho_{k,i}(m) > 0$ holds and therefore we can use Equation 6.12 and the multi-linearity of p and conclude:

$$\begin{aligned}
0 < \rho_{k,i}(m) &= p_{k,i}(m) - p_k(m) \\
&= p_{k,i}(m) - \sum_{l=1}^{e_k} v_{k_l} p_{k,l}(m) \\
&= p_{k,i}(m) - \sum_{l=1}^{e_k} \frac{\rho_{k,l}(m)}{\sum_{i=1}^{e_k} \rho_{k,i}(m)} p_{k,l}(m) \\
&= \frac{1}{\sum_{i=1}^{e_k} \rho_{k,i}(m)} \big[p_{k,i}(m) \sum_{i=1}^{e_k} \rho_{k,i}(m) - \sum_{i=1}^{e_k} \rho_{k,i}(m) p_{k,l}(m) \big] \\
&= \frac{1}{\sum_{i=1}^{e_k} \rho_{k,i}(m)} \big[p_{k,i}(m) \sum_{l \neq i} \rho_{k,l}(m) - \sum_{l \neq i} \rho_{k,l}(m) p_{k,l}(m) \big].
\end{aligned} \tag{6.15}$$

And therefore we have:

$$p_{k,i}(m) > \frac{1}{\sum_{l \neq i} \rho_{k,l}(m)} \left(\sum_{l \neq i} \rho_{k,l}(m) \right) p_{k,l}(m) \geq p_{k,i}(m). \tag{6.16}$$

This again leads to a contradiction.

Thus we have only one more case to look at: $\rho_{k,i}(m)$ is equal to 0 for all $i \in \{1, \ldots, e_k\}$. This means we have the following equation:

$$\begin{aligned}
& p_{k,i}(m) \leq p_k(m) \\
\Leftrightarrow \quad & p_k(s_{k,i}, -v_k) \leq p_k(v_k, -v_k) \text{ for all } i \in \{1, \ldots, e_k\} \\
\Leftrightarrow \quad & \max_{i \in \{1, \ldots, e_k\}} p_k(s_{k,i}, -v_k) \leq p_k(v_k, -v_k).
\end{aligned} \tag{6.17}$$

Now we only need to show the following:

$$\max_{i \in \{1, \ldots, e_k\}} p_k(s_{k,i}, -v_k) = \max_{q_k \in \mathbb{S}_{e_k}} p_k(q_k, -v_k). \tag{6.18}$$

After that we are finished. This is because all the requirements of Definition 4.1.1 are met, which means m is an equilibrium point.

Let us prove our claim from Equation 6.18. With the n-linearity of p_k, we can conclude the following:

$$\begin{aligned}\max_{q_k \in \mathbb{S}_{e_k}} p_k(q_k, -v_k) &= \sum_{j=1}^{e_k} q_{k_j} p_k(s_{k,j}, -v_k) \\ &\leq \max_{j \in \{1,\dots,e_k\}} p_k(s_{k,j}, -v_k) \underbrace{\sum_{j=1}^{e_k} q_{k_j}}_{=1} \\ &= \max_{j \in \{1,\dots,e_k\}} p_k(s_{k,j}, -v_k).\end{aligned} \tag{6.19}$$

Because $s_{k,j}$ is a mixed strategy and therefore lies in $\mathbb{S}_{e_k}$, we have proven our claim from Equation 6.18.

This means we have proven the following for every $v_k \in \{1, \dots, n\}$:

$$p_k(v_k, -v_k) \geq \max_{i \in \{1,\dots,e_k\}} p_k(s_{k,i}, -v_k) = \max_{q_k \in \mathbb{S}_{e_k}} p_k(q_k, -v_k). \tag{6.20}$$

That means all requirements of Definition 4.1.1 are met. Therefore m is in fact an equilibrium point. □

In contrast to Myerson's interpretation of the Nash equilibrium as a DNA double helix of economics [21, p. 1067], Von Neumann's alleged response was much more critical. He dismissed the proofs of the equilibrium as trivial fixed point theorems, as documented by Shapiro in [19, p. 64].

Despite von Neumann's initial characterization, it is safe to say that Nash's equilibrium, and the proofs associated with it, represent a monumental achievement. Nash's contributions to equilibria in game theory were duly recognized with the Nobel Memorial Prize in Economic Sciences in 1994, shared with John Harsanyi and Reinhard Selten [21, p. 1067].

Having looked at the mathematical intricacies of Nash equilibrium from various angles, we are now ready to examine a concrete application of non-cooperative game theory. In particular, we are concerned with the role of game theory in shaping decisions about prices and quantities made by firms in the market for a given good. This application involves a detailed analysis of strategic interactions between firms, where each firm tries to optimise its decisions by taking into account the actions of the others. The implications of non-cooperative game theory go beyond theoretical constructs and find practical relevance in the dynamic landscape of economic decision making.

7 Economic Applications

Now we can apply our mathematical insights to the competition among firms. In doing so, we look at markets in which more than one, but not infinitely many firms operate. This situation is called oligopoly [34]. The competition in an oligopolistic market can be accurately attributed to a strategic interaction. So we can apply the game theory we learned previously.

7.1 Cournot Competition

Well before Nash, Antoine-Augustin Cournot determined a Nash equilibrium using a concrete example [35, p. 135]. From today's perspective, Cournot could in fact be considered as one of the founders of game theory[40]. But here we have to add that the Nash equilibrium that Cournot found is "only" a pure one.

The example that Cournot considered in [18] is an oligopoly, but not only that. We need more properties. There is only one product to consider. This product is homogeneous. So all companies on the market for this product produce exactly the same thing. The companies do not work together. Another rule is that the companies must have market power. This means that the production of each company influences the price of the good. Last but not least, the companies behave rationally. In other words, they all strive to maximise their own profits.

In his model Cournot describes a market behaviour of two or more competitors – so let us simply say n firms – in a market in which the supply quantity is the variable which the companies change. So the Cournot-Model describes how to find a Nash equilibrium with infinitely many strategies [4, p. 27]. This is, because every firm can choose some quantity $q \in \mathbb{N}$[41].

At first the concept of only choosing quantities may seem odd, because we are used to talking more about prices than quantities. Imagine a scenario where a farmer must determine the daily quantity of crops to harvest and transport to a local market. In this market, the final selling price is determined by the total quantity of crops supplied by all farmers, and this price applies to all the crops sold on that day [34, p. 389]. Another example is the production and transportation of crude oil. The transport from Venezuela to Hamburg is time consuming. So the companies can only decide on the quantity to be delivered. The price then depends on the quantity loaded and the current demand [35, p. 136].

[40] some economists are in fact suggesting that, rather than *Nash equilibrium*, the equilibrium should be called *Cournot-Nash equilibrium* or maybe even *Cournot equilibrium* [21, p. 1070]

[41] for us: $0 \in \mathbb{N}$

Based on the examples, we can see that it makes no sense to assign probabilities for selecting a quantity. After all, the companies have to decide directly what quantity of products to send and how to send them. This means that the equilibrium we are considering is a pure one.

Now we can proceed in defining the Cournot Game. First Cournot looked at the markets with only two firms in [18]. Hence this is called a Duopoly which we will look at in the next subsection. Furthermore we will explore the generalisation to n firms.

7.1.1 The Easier Duopoly Case

Let us now assume there are only two companies: firm A and firm B. Both firms will strive to maximise their own profit, while knowing that the other firm is doing the same. As mentioned above we base the theory on very rationally thinking managers [36].

Firm A and B play a simultaneous-move game. In this game each firm chooses a quantity of the good to supply to the market [4, p. 27]. The game of course has infinitely many strategies since the sets of strategies for both firms are all natural numbers. Hence the following holds true: $S_A = S_B = \mathbb{N}$ [4, p. 27].

So we have two companies who supply the same market with exactly the same product. In this case the price is the same for firm A and B. This holds true because of the following situation: Let us imagine firm A is offering the same good as firm B to a higher price. Now everybody would buy the product form firm B and firm A would have zero sales [18]. Hence the two firms can only decide on the quantity which they are offering, q_A and q_B. Given these, the price adjusts to the level that clears the market - meaning that all units of the product are sold. Hence the two firms in total have the following price: $p(q_A + q_B)$ [34, p. 389]. In this case $p(\cdot) = d^{-1}(\cdot)$ is the inverse demand function [34, p. 389], which is defined in the following way:

$$d : \mathbb{R}_{>0} \to \mathbb{N}, p \mapsto d(p) \tag{7.1}$$

The demand function takes a price and maps it to the corresponding quantity, which is demanded by the costumers.

In this discussion we need the condition that $p(\cdot)$ is differentiable with $p'(\cdot) < 0$ [34, p. 390]. This should not be a problem anyway. Common sense predicts that when we have a higher price the demand should drop.

Of course we have to consider that the firms have some kind of costs. We assume here that both companies produce their output at a cost of $c > 0$ per unit[42]. Plus, we assume that $p(0) > c$ and that there is a $\tilde{q} \in \mathbb{N}_{>0}$ with $p(\tilde{q}) = c$.

Thus the profit of firm A and B is the price multiplied with the quantities minus the costs:

$$q_A \cdot p(q_A + q_B) - c \cdot q_A \text{ and } q_B \cdot p(q_A + q_B) - c \cdot q_B \tag{7.2}$$

Firm A does not have a direct influence on q_B. All it can do is to choose the best value for q_A. Hence we need a kind of "reaction function", which gives us each firm's optimal output for each fixed output level of their opponent [32, p. 14]. The reaction function of firm A against all possible outputs of firm B is defined by:

$$r_A : \mathbb{N} \to \mathbb{N} \tag{7.3}$$

Analogously the reaction function of firm B against all possible outputs of firm A is defined by:

$$r_B : \mathbb{N} \to \mathbb{N}. \tag{7.4}$$

In fact these are the best response functions of firm A and firm B, which we already defined in subsection 4.2.

Given the differentiability of $p(\cdot)$, we can now find an optimal q_A using the first-order condition. Our reaction function $r_1(\cdot)$ must satisfy the following:

$$\begin{aligned} &\frac{\partial p(r_1(q_B) + q_B) r_1(q_B) - c r_1(q_B)}{\partial r_1(q_B)} \\ &= p(r_1(q_B) + q_B) + r_1(q_B) \cdot p'(r_1(q_B) + q_B)) - c = 0 \\ &\iff p(r_1(q_B) + q_B) + r_1(q_B) \cdot p'(r_1(q_B) + q_B) = c. \end{aligned} \tag{7.5}$$

Analogously we can now find an optimal q_B using the first-order condition and our reaction function $r_2(\cdot)$ must satisfy the following:

$$\begin{aligned} &\frac{\partial p(q_A + r_2(q_A)) r_2(q_A) - c r_2(q_A)}{\partial r_2(q_A)} \\ &= p(q_A + r_2(q_A)) + r_2(q_A) \cdot p'(q_A + r_2(q_A)) - c = 0 \\ &\iff p(q_A + r_2(q_A)) + r_2(q_A) \cdot p'(q_A + r_2(q_A)) = c. \end{aligned} \tag{7.6}$$

[42] note that this is in the general case a simplification since normally there would be some kind of fixed costs

Now let us get on to our (pure strategy) Nash equilibrium. The tuple (q_A^*, q_B^*) is a Nash equilibrium if and if only $q_A^* \in r_1(q_B^*)$ and $q_B^* \in r_2(q_A^*)$. Thus our tuple (q_A^*, q_B^*) must satisfy the following condition to be a Nash equilibrium:

$$p(q_A^* + q_B^*) + q_A^* \cdot p'(q_A^* + q_B^*) = c \tag{7.7}$$

and

$$p(q_A^* + q_B^*) + q_B^* \cdot p'(q_A^* + q_B^*) = c \tag{7.8}$$

The property $q_A^* \in r_1(q_B^*)$ and $q_B^* \in r_2(q_A^*)$ can also be interpreted as the intersection between the two reaction functions. So there are intersection points of r_1 and r_2 if and if only there are equilibrium points. In this point none of the two firms has any incentive to change their supplied quantity, since this would worsen their position.

This may seem very abstract. Let us make it clear with a more specific example which we take from [32, p. 14]:

Example 7.1.1. Let us imagine that we have a linear demand. Hence we have the following for $p(\cdot)$:

$$p : \mathbb{N} \to \mathbb{N},\ q \mapsto \max(0, 1 - q). \tag{7.9}$$

Suppose we have costs per unit c for both firms with $0 \leq c \leq 1$. In the relevant range $0 \leq q \leq 1$ we get the reaction function for firm A by the following computations in which we use Equation 7.5:

$$\begin{aligned} p(r_1(q_B) + q_B) + r_1(q_B) \cdot p'(r_1(q_B) + q_B) &= c \\ \iff 1 - (r_1(q_B) + q_B) + r_1(q_B) \cdot (-1) &= c \\ \iff 1 - r_1(q_B) - q_B + r_1(q_B) \cdot (-1) &= c \\ \iff 1 - 2 \cdot r_1(q_B) - q_B &= c \\ \iff r_1(q_B) &= \frac{1 - q_B - c}{2} \end{aligned} \tag{7.10}$$

For firm B we compute the reaction function for the relevant range $0 \leq q \leq 1$ in the exact same way and we use Equation 7.6:

$$\begin{aligned} p(q_A + r_2(q_A)) + r_2(q_A) \cdot p'(q_A + r_2(q_A)) &= c \\ \iff 1 - (q_A + r_2(q_A)) + r_2(q_A) \cdot (-1) &= c \\ \iff 1 - q_A - r_2(q_A) + r_2(q_A) \cdot (-1) &= c \\ \iff 1 - q_A - 2 \cdot r_2(q_A) &= c \\ \iff r_2(q_A) &= \frac{1 - q_A - c}{2} \end{aligned} \tag{7.11}$$

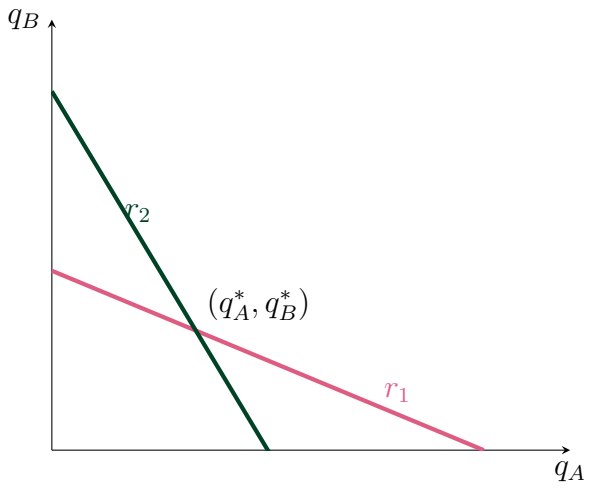


Figure 7.1: Nash equilibrium in the Cournot Competition for two firms

We already know that the tuple (q_A^*, q_B^*) is a Nash equilibrium if and if only $q_A^* = r_1(q_B^*)$ and $q_B^* = r_2(q_A^*)$. Thus we can calculate the following:

$$\begin{aligned} q_A^* = r_1(q_B^*) &= \frac{1 - q_B^* - c}{2} \Leftrightarrow q_B^* = -2q_A^* + 1 - c \\ q_B^* = r_2(q_A^*) &= \frac{1 - q_A^* - c}{2} \end{aligned} \tag{7.12}$$

Now we can equate these two equations. Hence we get:

$$\begin{aligned} \frac{1 - q_A^* - c}{2} &= -2q_A^* + 1 - c \\ \Leftrightarrow 1 - q_A^* - c &= -4q_A^* + 2 - 2c \\ \Leftrightarrow q_A^* &= \frac{1 - c}{3} \end{aligned} \tag{7.13}$$

Thus with our transformation of $q_A^* = r_1(q_B^*)$ in Equation 7.12 and our result in Equation 7.13 we calculate:

$$q_B^* = -2q_A^* + 1 = -2\frac{1 - c}{3} + 1 - c = \frac{1 - c}{3} \tag{7.14}$$

So our Nash equilibrium is $q_A^* = q_B^* = \frac{1-c}{3}$.

We visualise these two reaction functions and the Nash equilibrium in the relevant range in Figure 7.1.

7.1.2 The General Cournot Model

Now let us generalise our exploration to $n > 2$ identical firms, which are facing the same costs and demand functions as we have seen before. We are again using [32, p. 14], but are going to generalise the two firm case to n firms.

The profit of a firm $j \in \{1, \ldots, n\}$ can be described by the following:

$$q_j \cdot p(\sum_{i=1}^{n} q_i) - c \cdot q_j. \tag{7.15}$$

Again we need our reaction function $r_j : \mathbb{N} \to \mathbb{N}$, which gives us the optimal output of firm j given aggregated output of the others. With Equation 7.15 and the first-order condition we get the following[43]:

$$\begin{aligned} &\frac{\partial p\left(r_j(q_{-j}) + q_{-j}\right) r_j(q_{-j}) - c r_j(q_{-j})}{\partial r_j(q_{-j})} \\ &= p\left(r_j(q_{-j}) + q_{-j}\right) + r_j(q_{-j}) \cdot p'\left(r_j(q_{-j}) + q_{-j}\right) - c = 0 \\ &\iff p\left(r_j(q_{-j}) + q_{-j}\right) + r_j(q_{-j}) \cdot p'\left(r_j(q_{-j}) + q_{-j}\right) = c. \end{aligned} \tag{7.16}$$

Our Nash equilibrium again is the tuple $(q_1^*, \ldots, q_n^*)$ if and only if $q_j^* \in r_j(q_{-j}^*)$ for all $j \in \{1, \ldots, n\}$. Thus the tuple $(q_1^*, \ldots, q_n^*)$ must satisfy the following condition for all $j \in \{1, \ldots, n\}$:

$$p(q_j^* + q_{-j}^*) + q_j^* \cdot p'(q_j^* + q_{-j}^*) = c. \tag{7.17}$$

Like before we can visualise this condition with an example:

Example 7.1.2. Let us imagine that we have a linear demand. Hence we have the following for $p(\cdot)$:

$$p : \mathbb{N} \to \mathbb{N},\ q \mapsto \max(0, 1 - q). \tag{7.18}$$

For the response functions for all firms $1, \ldots, n$ we again use our Equation 7.16:

$$\begin{aligned} p(r_j(q_{-j}) + q_{-j}) + r_j(q_{-j}) \cdot p'(r_j(q_{-j}) + q_{-j}) = c \\ \iff r_j(q_{-j}) = \frac{1 - p_{-j} - c}{2} \end{aligned} \tag{7.19}$$

Now our Nash equilibrium is the tuple $(q_1^*, \ldots, q_n^*)$ if and only if $q_j^* = r_j(q_{-j}^*)$ for all $j \in \{1, \ldots, n\}$. Hence we have:

$$q_j^* = r_j(q_{-j}^*) = \frac{1 - p_{-j}^* - c}{2} \tag{7.20}$$

This can be transformed into:

$$0 = 1 - 2q_j^* - q_{-j}^* \tag{7.21}$$

[43] q_{-j} here stands for the aggregated output of all firms but firm j

If we now subtract $0 = 1 - 2q^*_{(j+1)} - q^*_{-(j+1)}$ from $0 = 1 - 2q^*_j - q^*_{-j}$, we get $q^*_{j+1} = q^*_j$ for all $j \in \{1, \ldots, n-1\}$. So it holds that $q^*_1 = q^*_2 = \cdots = q^*_n = q^*$. And if we insert q^* in Equation 7.21 we have:

$$q^* = \frac{1-c}{n+1}. \tag{7.22}$$

We finally have found our Nash-equilibrium!

7.2 Bertrand Competition

In this subsection we will explore the Bertrand model. This model might seem more intuitive, because it's about prices rather than quantities. Joseph Bertrand criticised Cournot for his focus on the quantities [35, p. 136].

Bertrand looked at markets with only two profit-maximising companies. The two companies are firm 1 and 2. Again these companies sell exactly the same product and do not cooperate.

Let the demand function, which is of course dependent on the price p, be $d(p)$. We can assume that $d(p)$ is strictly monotonic decreasing and continuous for all prices p with $d(p) > 0$. Furthermore we suppose that there is a $\hat{p} < \infty$ with $d(p) = 0$ for all $p \geq \hat{p}$. This holds, because when the price of a certain product is increased not everybody will be willing to buy this product again. So the demand will abate. Nevertheless the demand cannot be negative [34, 387f].

In this model we neglect the fixed costs, too. Instead we propose that the two firms have the same constant costs per unit produced. We call this cost $c > 0$ [34, p. 388].

The competition will pass as it follows: the two regarded firms simultaneously name their prices p_1 and p_2. Sales by firm i, where $i, j \in \{1, 2\}$, are given by the following Equation 7.23 [34, p. 388]:

$$G_i(p_i, p_j) = \begin{cases} d(p_i) & if\ p_i < p_j \\ \frac{1}{2}d(p_i) & if\ p_i = p_j \\ 0 & if\ p_i > p_j \end{cases} \tag{7.23}$$

Presume the prices are given with p_i and p_j. Now the profit of firm i can be calculated by $(p_i - c)G_i(p_i, p_j)$. [34, p. 388]

Now we have everything for the so called Bertrand equilibrium in economics. This is actually a Nash equilibrium, which we are going to show in the next proposition. Here we follow mainly the remarks in [34, 388f].

Proposition 7.2.1. *There is a unique Nash equilibrium* $(\tilde{p_1}, \tilde{p_2})$*. In this equilibrium, both firms set their prices in the following way:*

$$\tilde{p_1} = \tilde{p_2} = c. \tag{7.24}$$

Proof. First we have to show that $\tilde{p_1} = \tilde{p_2} = c$ is indeed a Nash equilibrium. In this point no company has an inducement to change their price stetting. Assuming one firm raises its price, they will make no sales and thus zero profit. If the firm in contrast lowers its prices below c it will of course increase its sales but bring in a negative profit.

Secondly, we have to think about why this is the only equilibrium. Assume there is an other equilibrium with the prices (p_1, p_2). We now have distinguish some cases.

1. $p_1 < q_2$ and $p_1 < c$: In this case the first firm will get losses. This can be prevented by raising the price above c. The worst it can get is earning zero, which is better than making negative profit. Hence this can't be a Nash equilibrium.

2. $p_1 = c$ and $p_2 > c$: Now firm 1 will supply the whole market but has no profit at all. By raising the price a little bit - for example to $\tilde{p_1} = \frac{p_2 - c}{2}$ - firm 1 will still supply the entire market, but will in fact make a positive profit. Thus this case is not a Nash equilibrium, too.

3. $p_1 > c$ and $p_2 > c$: Without loss of generality we suppose $p_1 \leq p_2$. Now firm 2 at most can have the profit $\frac{1}{2}(p_2 - c)d(p_2)$. Suppose the company lowers p_2 to $\tilde{p_2}$, which is lower than p_1. In this case firm 2 can earn $(\tilde{p_2} - c)d(p_2)$. Given that $(\tilde{p_2} - c)d(p_2) > \frac{1}{2}(p_2 - c)d(p_2)$ for small $\tilde{p_2}$, firm 2 has now increased its profit. From this follows that this situation is not a Nash equilibrium.

These are without loss of generality all possible cases. We have now seen that each of these cases is not a Nash equilibrium. The only remaining one is $\tilde{p_1} = \tilde{p_2} = c$, which is a Nash equilibrium as seen above. Hence it is the only one and we are finished. □

So we found out that in the Bertrand equilibrium, companies set their prices as low as possible in order to attract demand and increase their market share. Since the two companies offer the same product, they are forced to lower their prices to the level of their competitors in order to maintain their market share. The result of the Bertrand equilibrium is that prices in the market fall to the minimum that companies need to cover their production costs.

8 Conclusion

In conclusion, this thesis has delved into the intricate world of the Nash equilibrium, a pivotal concept in game theory with profound implications across diverse disciplines.

Our analysis has demonstrated that the Nash equilibrium serves as a powerful tool for modelling and understanding interactions among rational agents. It provides a framework to predict outcomes in various real-world scenarios, ranging from economics and politics to biology. In this thesis we concentrated on some or rather a few of the economic applications.

Despite its powerful statement, the Nash equilibrium has some boundaries, too. It is only a framework and cannot model any irrational thinking. In our thesis, we continuously assumed that all participants act rationally and are striving for the best possible outcome for themselves. But what about altruism? What about cheating or ignoring the rules? The Nash equilibrium gives no answers to those questions.

Let us look at this from an economical point of view. Although companies in most countries are not allowed to form a cartel, there are some legal ways to cooperate. Imagine a car company gives a huge press conference on why they are about to raise the prices of their cars. Of course this influences the other car companies. They might raise their prices, too. With this decision all companies are better off than they would be in the Nash equilibrium (assuming of course they started in an equilibrium point). This state is a result of cooperation, which is prohibited in the theory of non-cooperative games.

The portrayed example underscores that not all the nuances of human interaction stand up to mathematical scrutiny. The elegance of the Nash equilibrium finds its limits in the subtleties of real-world dynamics, where cooperation, even within legal limits, can change outcomes contrary to theoretical predictions.

References

[1] J. Nash, "Non-cooperative games", *Annals of mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[2] G. Koenig. "Was macht Spielen zum Kulturgut?" (Nov. 15, 2018), [Online]. Available: `https://museenblog-nuernberg.de/2018/11/15/kulturgut-spiel/` (visited on 01/16/2024).

[3] J. D. Beasley, *The mathematics of games*. Courier Corporation, 2013.

[4] T. Fujiwara-Greve, *Non-cooperative game theory*. Springer, 2015.

[5] M. J. Holler and G. Illing, *Einführung in die Spieltheorie*. Springer Berlin, Heidelberg, 1996.

[6] Bundeszentrale für politische Bildung, "Duden Wirtschaft von A bis Z: Grundlagenwissen für Schule und Studium, Beruf und Alltag", *Aufl. Hg. v. Bibliographisches Institut*, 2016.

[7] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. Princeton university press, 2007.

[8] S. Karlin, *Mathematical methods and theory in games, programming, and economics: Matrix games, programming, and mathematical economics*. Elsevier, 2014.

[9] G. Owen, *Spieltheorie*. Springer-Verlag, 2013.

[10] T. Sauer, *Spieltheorie*. Logos Verlag Berlin GmbH, 2017.

[11] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications", *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[12] A. Deitmar, *Analysis*. Springer-Verlag, 2016.

[13] S. Hildebrandt, *Analysis 1*. Springer-Verlag, 2006.

[14] F. Edward Su, "Rental harmony: Sperner's lemma in fair division", *The American mathematical monthly*, vol. 106, no. 10, pp. 930–942, 1999.

[15] R. Fenn, *Geometry*. Springer Science & Business Media, 2007.

[16] J. F. Nash Jr, "Equilibrium points in n-person games", *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[17] P. R. Thie and G. E. Keough, *An introduction to linear programming and game theory*. John Wiley & Sons, 2011.

[18] A. A. Cournot, *Researches into the Mathematical Principles of the Theory of Wealth*. New York: Macmillan Company [c1897], 1927.

[19] J. H. Shapiro, *A fixed-point farrago*. Springer, 2016.

[20] O. E. Dictionary. "Equilibrium, n." (Sep. 1, 2023), [Online]. Available: `https://doi.org/10.1093/OED/1111299704` (visited on 11/17/2023).

[21] R. B. Myerson, "Nash equilibrium and the history of economic theory", *Journal of Economic Literature*, vol. 37, no. 3, pp. 1067–1082, 1999.

[22] P. Subrahmanyam, *Elementary fixed point theorems*. Springer, 2018.

[23] K. Meyberg and P. Vachenauer, *Höhere Mathematik 1: Differential-und Integralrechnung Vektor-und Matrizenrechnung*. Springer-Verlag, 2013.

[24] L. E. J. Brouwer, "Über Abbildung von Mannigfaltigkeiten", *Mathematische annalen*, vol. 71, no. 1, pp. 97–115, 1911.

[25] E. Sperner, "Neuer Beweis für die Invarianz der Dimensionszahl und des Gebietes", in *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Springer, vol. 6, 1928, pp. 265–272.

[26] E. Zeidler, *Applied functional analysis: applications to mathematical physics*. Springer Science & Business Media, 2012, vol. 108.

[27] S. Kakutani, "A generalization of brouwer's fixed point theorem", *Duke Mathematical Journal*, vol. 8, no. 3, pp. 457–459, 1941.

[28] D. Charalambos and B. Aliprantis, *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin and Heidelberg GmbH & Company KG, 2013.

[29] K. Königsberger, *Analysis 2*. Springer-Verlag, 2013.

[30] M. De Longueville, *A course in topological combinatorics*. Springer Science & Business Media, 2013.

[31] W. Fulton, *Algebraic topology: a first course*. Springer Science & Business Media, 2013, vol. 153.

[32] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.

[33] B. Von Querenburg and G. Bengel, *Mengentheoretische Topologie*. Springer, 1973, vol. 3.

[34] A. Mas-Colell, M. D. Whinston, J. R. Green, *et al.*, *Microeconomic theory*. Oxford university press New York, 1995, vol. 1.

[35] W. Pfähler and H. Wiese, *Unternehmensstrategien im Wettbewerb: eine spieltheoretische Analyse*. Springer-Verlag, 2008.

[36] T. Riechmann, *Spieltheorie*. Vahlen, 2014.

List of Figures

2.1 3-dimensional simplex $\mathbb{S}_3$ 7

2.2 Two different sets in $\mathbb{R}^2$ 7

4.1 Best-response of player 1 and player 2 23

5.1 Construction for the proof of Corollary 5.1.5 26

5.2 Sperner's Lemma for dimension 2 28

5.3 Illustration of the Minkowski functional 33

5.4 Example of a mulitfunction 38

5.5 Barycentric Simplical Triangulation 41

5.6 Point-to-point mapping 41

7.1 Nash equilibrium in the Cournot Competition for two firms 54

Eigenständigkeitserklärung

Hiermit versichere ich, Lea Rohleder, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen wörtlich oder sinngemäß übernommenen Gedanken sind als solche gekennzeichnet. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Passau, den 26. Januar 2024

_____

Lea Rohleder