

## Sind 7 Mischungen genug?

### *Teilnehmer:*

Leo Batzke	Herder-Oberschule
Dietmar Funck	Herder-Oberschule
Philipp Kähler	Georg-Forster-Oberschule
Verena Klös	Herder-Oberschule
Christian Renau	Heinrich-Hertz-Oberschule
Jana Schulz	Andreas-Oberschule

### *Gruppenleiter:*

Elke Warmuth	Humboldt-Universität zu Berlin, Mitglied im DFG-Forschungszentrum MATHEON „Mathematik für Schlüsseltechnologien“
--------------	--

Nach wie vielen Mischungen kann ein Kartenspiel mit 52 Karten als „gut gemischt“ gelten? Dieser Frage sind wir auf der Grundlage eines mathematischen Modells für eine verbreitete Form des Mischens nachgegangen. Zufällige Permutationen, kombinatorisches Zählen, Wahrscheinlichkeitsverteilungen und deren Abstände haben dabei eine Rolle gespielt. Als Ausblick haben wir untersucht, was dieses Problem mit Markov-Ketten und deren stationären Zuständen zu tun hat.

# 1 Der Kartentrick

## 1.1 Beschreibung des Kartentricks

Jeder hat schon einmal einen Kartentrick vorgeführt bekommen. Wenn man sich mit dem Thema „Mischen von Spielkarten“ beschäftigt, kommt man etwas sehr Interessantem auf die Schliche. Dabei wird ein geordnetes Kartenspiel zunächst zweimal per Riffle Shuffle, wie im weiteren Text beschrieben, gemischt. Anschließend hebt einer der Zuschauer die oberste Karte des Stapels ab, merkt sie sich, ohne sie dem „Magier“ zu zeigen, und schiebt sie in den Stapel. Das geübte Auge erkennt anschließend an der Reihenfolge der Karten, trotz zweimaliger Mischung, die eingeschobene.

## 1.2 Erklärung des Kartentricks

Um den Kartentrick erklären zu können, muss zuvor der Begriff „aufsteigende Sequenz“ eingeführt werden. Dafür wird ein Beispiel mit acht nummerierten Karten, die mit Hilfe des Riffle Shuffle gemischt werden, verwendet. Eine mögliche Reihenfolge, die durch diesen Mischvorgang entstehen kann, ist: 51263784. In dieser Aufeinanderfolge der Karten gibt es zwei aufsteigende Sequenzen: 1234 und 5678. Man kann hierbei erkennen, dass der Kartenstapel zwischen der vierten und fünften Karten geteilt wurde. Da die Karten in den einzelnen Stapeln noch in ihrer ursprünglichen Reihenfolge vorhanden sind und die Reihenfolge des einen Stapels nur durch die Karten des anderen unterbrochen ist, sind die aufsteigenden Sequenzen noch zu erkennen.

Bei dem Kartentrick benutzt man die Existenz der aufsteigenden Sequenzen nach dem Mischen durch den Riffle Shuffle. Wird die obere Karte vom Stapel genommen und dann in eine beliebige Stelle des Kartenstapels hineingesteckt, passt diese nicht mehr zu den aufsteigenden Sequenzen. Denn nach einem Mischvorgang sind nur maximal zwei aufsteigende Sequenzen im Stapel vorhanden und die hineingesteckte Karte würde dann eine dritte aufsteigende Sequenz, die aus einer einzigen Zahl besteht, bilden. Wenn also bei dem Beispiel die Karte vier von oben in den Stapel gesteckt wird, dann könnte folgende Reihenfolge entstehen: 51264378. Da nun die vier in keine der beiden aufsteigenden Sequenzen hineingeht, kann man sofort erkennen, dass diese Karte gezogen wurde.

Den Kartentrick kann man natürlich noch optimieren, indem man ein vollständiges Skatblatt benutzt und den Stapel dreimal mit Hilfe des Riffle Shuffle mischt. Nach diesem Vorgang können maximal  $2^3 = 8$  aufsteigende Sequenzen vorhanden sein.

Die Wahrscheinlichkeit, dass der Kartentrick funktioniert, ist sehr hoch. Es kann

jedoch auch vorkommen, dass man die gesuchte Karte nicht herausfindet.

## 2 Der Riffle Shuffle

Der Riffle Shuffle ist die gebräuchlichste Art ein Kartendeck zu mischen. Man teilt den vorhandenen Kartenstapel in zwei Teile und verteilt diese auf die linke und die rechte Hand. Nun greift man mit den jeweiligen Daumen an eine Ecke, führt diese Ecken zueinander und zieht die Daumen so nach oben, dass sich die Karten von der linken und der rechten Hand möglichst einzeln ineinander ordnen. Diesen Vorgang kann man beliebig oft wiederholen, bis man glaubt, dass der Kartenstapel ausreichend gemischt ist.

## 3 Grundlagen

### 3.1 Permutationen

Beim Mischen verändert man die Reihenfolge der Karte. Im mathematischen Modell betrachtet man dafür Permutationen.

Was ist das?

Als Permutation bezeichnet man die Abbildung einer  $n$ -elementigen Menge auf sich selbst. Eine Permutation wird durch explizite Angabe von Urbild und Bild beschrieben. Eine häufig verwendete Darstellungsform ist eine zweizeilige und  $n$ -spaltige Matrix. Dabei wird das Urbild in die obere Zeile und das Bild in die untere Zeile geschrieben.

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 4 & 5 & 1 & 7 & 6 \end{pmatrix}$$

Die Menge aller Permutationen von  $n$  Elementen nennt man  $S_n$ . Beim Verknüpfen zweier Permutationen  $\pi \circ \tau$  wendet man zuerst die rechte Permutation  $\tau$  und anschließend  $\pi$  auf das Ergebnis an.

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$S_n$  ist eine Gruppe, das heißt, dass sie folgende Eigenschaften erfüllt:

1) Die Verknüpfung ist eine innere Komposition. Das Ergebnis der Verknüpfung

ist also ebenfalls ein Element von  $S_n$ .

2) Es gibt ein universell neutrales Element, so dass bei der Verknüpfung mit diesem die Permutation nicht verändert wird.

3) Es gibt für jedes Element ein individuell inverses Element, so dass die Verknüpfung eines Elementes mit seinem inversen das neutrale Element ergibt.

4) Das Assoziativgesetz muss gelten.

Am Beispiel von  $S_3$  kann man die Erfüllung dieser Eigenschaften überprüfen:

1) s. obiges Beispiel

2)

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

3)

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

4)

$$\begin{aligned} \left( \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right) \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \left( \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right) &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \end{aligned}$$

### Weitere Eigenschaften

- $\text{ord } S_n = n!$
- $S_n$  ist nicht kommutativ für  $n \geq 3$ .

## 3.2 Wahrscheinlichkeitsverteilungen und Zählungen

Das Mischen der Spielkarten stellt einen Prozess mit (mehr oder weniger) zufälligem Ergebnis dar, es ist daher notwendig, sich mit Wahrscheinlichkeiten zu befassen.

Als Beispiel sei hier ein perfekter Würfel angenommen, der einmal geworfen wird. Seine Ergebnismenge umfasst die möglichen Augenzahlen.

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Jede Teilmenge  $A$  von  $\Omega$  wird als Ereignis bezeichnet.

Beispiele:

- 1)  $A$ : Würfeln einer 3 ( $A = \{3\}$ )
- 2)  $B$ : Würfeln einer geraden Zahl ( $B = \{2, 4, 6\}$ )
- 3) Zweistufiges Zufallsexperiment.  $C$ : nacheinander werden 2 und 5 gewürfelt. ( $C = \{(2, 5)\}$ )

Für das letzte Zufallsexperiment wäre die Ergebnismenge die Menge aller geordneten Paare von (1, 1) bis (6, 6).

Als Wahrscheinlichkeitsverteilung bezeichnet man die Zuordnung der entsprechenden Wahrscheinlichkeiten zu den Elementarereignissen, den Elementen der Ergebnismenge. Für einen perfekten Würfel wäre das:

Augenzahl	1	2	3	4	5	6
Wahrscheinlichkeit	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Für allgemeine und unbekannte Verteilungen nimmt man üblicherweise  $p_1, p_2, \dots, p_n$  für die Wahrscheinlichkeiten der Elementarereignisse  $\omega_1, \omega_2, \dots, \omega_n$ , die auch nicht unbedingt gleichgroß sein müssen, womit die Verteilung nicht zwangsläufig eine Gleichverteilung ist.

Aus obenstehender Tabelle ist sofort die Wahrscheinlichkeit für  $A$  ersichtlich, nämlich  $P(A) = \frac{1}{6}$ . Die Wahrscheinlichkeit für  $B$  ist, da die einzelnen Elementarereignisse in  $B$  einander ausschließen, die Summe der Einzelwahrscheinlichkeiten, mithin  $P(B) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$ . Allgemein gilt bei Gleichverteilung, d.h. wenn alle Elementarereignisse aus  $\Omega$  die gleiche Wahrscheinlichkeit haben, und wenn alle Ergebnisse in  $A$  paarweise disjunkt sind, also einander ausschließen:

$$P(A) = \frac{\text{ord } A}{\text{ord } \Omega}$$

Wie man mit Hilfe eines Baumdiagramms leicht verdeutlichen kann, ist die Wahrscheinlichkeit des Ereignis  $C$  gerade das Produkt der einzelnen Wahrscheinlichkeiten jedes Wurfs, da beide Würfe voneinander unabhängig sind, also  $P(C) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$ .

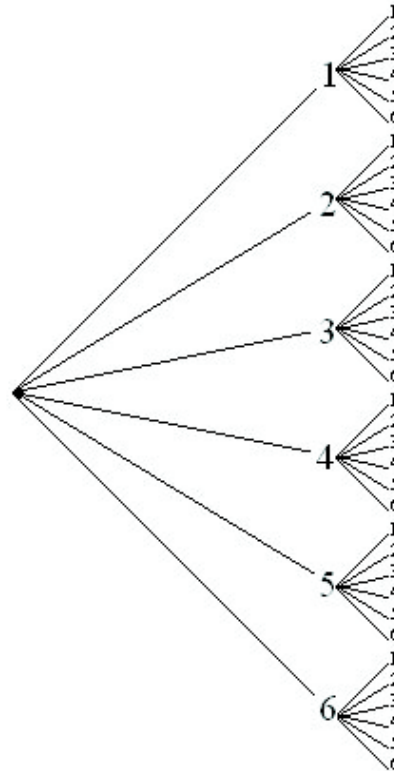


Abbildung 1: Baumdiagramm für zweimaliges Würfeln

Zwei Ereignisse sind, anschaulich gesprochen, dann voneinander unabhängig, wenn das Eintreten des ersten keinen Einfluß auf das Eintreten des zweiten hat. Wie aus Beispiel  $B$  ersichtlich ist es für die Berechnung von Wahrscheinlichkeiten unerlässlich ein Zählsystem zu entwickeln. Ein Ansatz ist bereits bei den Permutationen erkennbar. Wir benötigen hier vor allem den sogenannten Binomialkoeffizienten  $\binom{n}{k}$ . Dieser gibt die Anzahl aller  $k$ -elementigen Teilmengen einer  $n$ -elementigen Menge an. Dies ergibt sich wie folgt:

Es gibt  $n!$  Möglichkeiten  $n$  Elemente einer Menge anzuordnen. Es interessieren uns jedoch nur  $k$  Elemente, deren Ordnung untereinander, gegeben durch  $k!$  egal ist. Genausowenig interessiert uns jedoch die Reihenfolge der verbleibenden  $n - k$  Elemente, die sich mit  $(n - k)!$  bemisst. Es verbleiben also noch  $\frac{n!}{k!(n-k)!} = \binom{n}{k}$  mögliche Mengen.

## 4 Mathematische Modellierung des Mischvorgangs

Man kann einen Mischvorgang an einem Kartenspiel nun als eine Wahrscheinlichkeitsverteilung über alle möglichen Permutationen des Decks auffassen. Ausgehend vom jungfräulichen, geordneten Deck erhält man je nach Art der Mischung mit einer bestimmten Wahrscheinlichkeit eine bestimmte Kartenfolge. Als Beispiel soll hier der denkbar einfachste Mischvorgang dienen, das Abheben der obersten Karte und das Hineinschieben derselben irgendwo in den Stapel.

Permutation	123	132	213	231	312	321
Kartenfolge	123	132	213	312	231	321
Wahrscheinlichkeit	$\frac{1}{3}$	0	$\frac{1}{3}$	0	$\frac{1}{3}$	0

Die drei überhaupt möglichen Anordnungen haben natürlich nur unter der Annahme diese Wahrscheinlichkeit, dass alle drei Möglichkeiten die Karte 1 in den Stapel zurückzuschieben gleichermaßen wahrscheinlich sind. Es ist jedoch eher ungewöhnlich die Karte an ihren ursprünglichen Platz zu legen. Auch unser Modell des Riffle Shuffle weist solche kleinen Mankos auf, doch dazu später mehr. Die verbleibenden beiden können nicht auftreten, da es auch noch notwendig wäre die Karten 2 und 3 zu vertauschen.

Wenden wir uns nun dem Riffle Shuffle zu.

Im ersten Schritt wird der gesamte Kartenstapel mit  $n$  Karten in zwei Stapel zu  $k$  und  $n-k$  Karten geteilt. Auf Basis der Annahme, dass der Mischende sich bemüht mittig zu teilen, erscheint eine Binomialverteilung sinnvoll. Wie sicher bekannt ist, steigen die Werte des Binomialkoeffizienten  $\binom{n}{k}$  bis  $k = \frac{n}{2}$  bzw.  $k = \frac{n-1}{2}$  an und sinken dann wieder ab. (Dies sieht man sehr gut am Pascalschen Dreieck.)

Da die Summe  $\sum_{i=0}^n \binom{n}{i}$  gleich  $2^n$  ist ergibt sich eine Wahrscheinlichkeit von  $\frac{\binom{n}{k}}{2^n}$  für die Teilung des Stapels an der  $k$ -ten Stelle. Wie man leicht sieht, ist es in diesem Modell möglich, dass einer der Stapel keine Karten enthält. Realistisch betrachtet ist es eigentlich mehr als ungebräuchlich einen Kartenstapel auf diese Weise zu teilen.

Für das Zusammenschieben der beiden Stapel ergeben sich  $\binom{n}{k}$  Möglichkeiten. Dies kann man sich plausibel machen, indem man sich einen „leeren“ Stapel mit  $n$  freien Plätzen vorstellt. Von diesen sucht man sich  $k$  aus an die, unter Beibehaltung der internen Reihenfolge, die  $k$  Karten des einen Stapels einsortiert werden. Damit sind auch die Plätze der übrigen  $n-k$  Karten festgelegt, da auch ihre Reihenfolge nicht geändert wird. Wir nehmen in unserem Modell an, dass alle diese Möglichkeiten dieselbe Wahrscheinlichkeit besitzen.

Da die Teilung und das Zusammenschieben voneinander unabhängig sind kann man die Einzelwahrscheinlichkeiten multiplizieren.

$$\frac{\binom{n}{k}}{2^n} \cdot \frac{1}{\binom{n}{k}} = \frac{1}{2^n}$$

Eine Folge aus einer bestimmten Teilung und einer bestimmten Zusammenschiebung hat also diese Wahrscheinlichkeit. Allerdings ist es möglich, die identische Permutation auf mehreren Wegen zu erreichen, so dass wir uns dies genauer ansehen sollten.

Als anschauliches Beispiel soll hier wieder der Stapel aus drei Karten dienen.

Teilungsposition k	Teilung	Wahrscheinlichkeit	Zusammenschiebung
0	123	$\frac{1}{6}$	123
1	1 23	$\frac{1}{6}$	123, 213, 231
2	12 3	$\frac{1}{6}$	123, 132, 312
3	123	$\frac{1}{6}$	123

Die Kartenreihenfolge umzukehren ist so nicht möglich, da die Reihenfolge der Karten im Stapel nicht geändert wird. Wie leicht erkennbar ist, kann man nun folgende Wahrscheinlichkeitsverteilung aufstellen.

Permutation	123	132	213	231	312	321
Kartenfolge	123	132	213	312	231	321
Wahrscheinlichkeit	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0

## 4.1 $a$ -Shuffle

Um zu sehen, was sich bei einer Hintereinanderausführung von Riffle Shuffles für Wahrscheinlichkeiten ergeben, nimmt man die  $a$ -Shuffles zu Hilfe.

Man teile das Kartendeck in  $a$  Kartenstapel, wobei sie die Größen ( $\geq 0$ )  $s_1, \dots, s_a$  haben mit  $s_1 + s_2 + \dots + s_a = n$ . Anschließend fügt man die Kartenstapel wieder zu einem Stapel zusammen. Dabei ist lediglich darauf zu achten, dass die relative Ordnung zwischen den Karten eines jeden Kartenstapels erhalten bleibt. Diese Form des Mischen bezeichnen wir als  $a$ -Shuffle.

Nun betrachten wir die Wahrscheinlichkeiten. Jede Aufteilung in  $a$  Stapel mit den Größen  $s_1, \dots, s_a$  hat die Wahrscheinlichkeit:

$$\frac{\binom{n}{s_1, \dots, s_a}}{a^n}$$

Bei dem Term im Zähler handelt es sich um dem sogenannten Multinomialkoeffizienten. Dieser gibt die Zahl der Möglichkeiten an, eine  $n$ -elementige Menge in  $a$  Teilmengen mit den jeweiligen Größen  $s_1, \dots, s_a$  zu teilen.

Von einer beliebigen Aufteilung ausgehend, haben alle Arten des Zusammenfügens der  $a$  Kartenstapel die gleiche Wahrscheinlichkeit. Dazu berechnen wir die Anzahl der Möglichkeiten die  $n$ -Positionen des Decks auszuwählen. Die Anzahl ist wie folgt gegeben:

$$\binom{n}{s_1, \dots, s_a}$$



Wieder nehmen wir an, dass jede dieser Möglichkeiten dieselbe Wahrscheinlichkeit besitzt, also:

$$\frac{1}{\binom{n}{s_1, \dots, s_a}}$$

Nun ergibt sich die Wahrscheinlichkeit für die Kombination einer Aufteilung und einer Zusammenfügung aus dem Produkt der beiden Einzelwahrscheinlichkeiten:

$$\frac{\binom{n}{s_1, \dots, s_a}}{a^n} \cdot \frac{1}{\binom{n}{s_1, \dots, s_a}} = \frac{1}{a^n}$$

Daher ergibt es sich, dass alle Kombinationen die gleiche Wahrscheinlichkeit haben. Dies ist damit genauso wie beim Riffle Shuffle. Betrachtet man das  $a$ -Shuffle, stellt man fest, dass der Riffle Shuffle ein Spezialfall des  $a$ -Shuffles mit  $a = 2$  ist.

Es gibt eine Codierungsmethode, um zu erkennen, wie der  $a$ -Shuffle funktioniert. Dazu nimmt man eine Zahl mit  $n$  Ziffern zur Basis  $a$ . Wir nennen sie  $A$ . Als erstes zählt man die Anzahl der Nullen in der Zahl  $A$ . Diese Anzahl ist die Größe des ersten Stapels. Danach zählt man die Anzahl der Einsen in der Zahl  $A$  und diese Anzahl ist die Größe des zweiten Stapels. So verfährt man mit allen Ziffern der Basis- $a$ -Zahl. Dabei entstehen  $a$  Stapel. Nun schreibt man sich die Zahlen von 1 bis  $n$  auf und unterteilt sie entsprechend der Stapelgrößen. Danach kommen an die Stellen wo bei  $A$  die Nullen sind die Zahlen des ersten Stapels hin. Die Zahlen des zweiten Stapels kommen da hin, wo die Einsen in  $A$  sind. So verfährt man mit allen Stapeln. Bei dem Ersetzen ist darauf zu achten, dass die relative Ordnung zwischen den Zahlen eines jeden Stapels erhalten bleibt. Als Ergebnis erhält man das gemischte Kartendeck bezüglich des speziellen Aufteilens und Zusammenfügens bezüglich  $A$ . Durch die Codierungsmethode wird allerdings  $\pi^{-1}$  und nicht  $\pi$  beschrieben.

## 4.2 Das Multiplikationstheorem

Das Multiplikationstheorem besagt, dass ein  $a$ -Shuffle gefolgt von einem  $b$ -Shuffle das gleiche ist wie ein  $a \cdot b$ -Shuffle. Zum Beweis lässt sich die eben verwendete Codierungsmethode gebrauchen.

**Satz:**

- Die durch  $A^B \& B$  erzeugten Codes sind  $n$ -stellige Basis- $ab$ -Zahlen.
- Die Zuordnung  $(A, B) \leftrightarrow A^B \& B$  ist bijektiv.

Beweis:

$$1) \quad \begin{array}{l} A^B x_1 \dots x_n \mid \cdot b \\ B y_1 \dots y_n \end{array}$$

An der  $i$ -ten Stelle gilt:

$$0 \leq x_i \cdot b + y_i \leq (a-1) \cdot b + b - 1 = a \cdot b - 1$$

2)

$$A^B \& B = \sum_{i=0}^{n-1} (A_i^B \times b + B_i) \times (a \times b)^i$$

Der Beweis ergibt sich durch den Satz über die Division mit Rest.

Beispiel:  $a = 3, b = 4, n = 6$

$A = 012210, B = 310100$

$$\begin{array}{r} A \quad 0 \ 1 \ 2 \ 2 \ 1 \ 0 \\ A^B \quad 0 \ 2 \ 0 \ 1 \ 1 \ 2 \mid \cdot b \\ B \quad 3 \ 1 \ 0 \ 1 \ 0 \ 0 \\ \hline A^B \& B \quad 3 \ 9 \ 0 \ 5 \ 4 \ 8 \\ \hline \pi^{-1} : \quad 2 \ 6 \ 1 \ 4 \ 3 \ 5 \end{array}$$

$A^B$  bedeutet die Ziffern von  $A$  permutiert gemäß der durch  $B$  beschriebenen Permutation  $\pi$ .

## 5 Euler-Zahlen

Definition:  $\langle n \rangle_k$  ist definiert als die Anzahl der Permutationen von  $\{1, 2, \dots, n\}$ , die  $k$  Aufstiege haben. Ein Aufstieg ist eine Stelle in einer Permutation mit den Elementen  $\pi_i$  und  $\pi_{i+1}$ , an der  $\pi_i < \pi_{i+1}$  gilt. Entsprechend gilt bei einem Abstieg  $\pi_i > \pi_{i+1}$ .

Beispiel: Alle Permutationen von  $\{1, 2, 3, 4\}$ , die 2 Aufstiege haben, wir wählen also  $n = 4, k = 2$ :

1324, 1423, 2314, 2413, 3412, 1243, 1342, 2341, 2134, 3124 und 4123.

$\langle 4 \rangle_2$  ist also 11.

Es können bei  $n$  Elementen höchstens  $n - 1$  Aufstiege vorkommen, deshalb ist  $\langle n \rangle_n = 0$  und  $\langle n \rangle_0 = 1$ , weil  $\{1, 2, \dots, n\}$  genau eine Permutation ohne Aufstieg hat, nämlich die umgekehrte (nicht die inverse!) Permutation:  $n \dots 1$

Da die Elemente von  $\{1, 2, \dots, n\}$  nicht gleich sein können, hat jede Permutation an  $n-1$  Stellen entweder einen Aufstieg oder einen Abstieg. Hat eine Permutation  $k$  Aufstiege, dann hat sie  $n-1-k$  Abstiege.

Bei Umkehrung einer Permutation werden aus Aufstiegen Abstiege und andersrum. Da es genau  $\binom{n}{k}$  Permutationen von  $\{1, 2, \dots, n\}$  mit  $k$  Aufstiegen gibt, gibt es genauso viele mit  $n-1-k$  Abstiegen, nämlich genau die Umkehrungen, es gilt also:

$$\binom{n}{k} = \binom{n}{n-1-k}$$

Die Euler-Zahlen lassen sich rekursiv berechnen, man definiert daher als Anfang:

$$\binom{0}{k} = 0 \text{ und } \binom{n}{k} = 0 \text{ für } k < 0$$

Man nimmt sich eine Permutation  $\rho = \rho_1 \dots \rho_{n-1}$  von  $\{1, 2, \dots, n-1\}$ . Dann stellt man das Element  $n$  in allen möglichen Positionen dazu. Stellt man es in Position  $j$ , erhält man Permutation  $\pi = \rho_1 \dots \rho_{j-1} n \rho_j \dots \rho_{n-1}$ . Die Anzahl der Aufstiege bleibt für  $\rho_{j-1} < \rho_j$  und  $j = 1$  gleich, bei  $\rho_{j-1} > \rho_j$  und  $j = n$  erhöht sie sich um 1.

Es gibt  $k$  Stellen, an denen  $\rho_{j-1} < \rho_j$  gilt und eine für  $j = 1$ , und das bei allen  $\binom{n-1}{k}$  Permutationen, analog gibt es in allen  $\binom{n-1}{k-1}$  Permutationen  $n-1-k$  Abstiege und  $j = n$  als Möglichkeit, Permutationen mit  $k$  Aufstiegen zu erzeugen. Also erhält man für  $\binom{n}{k}$ :

$$\binom{n}{k} = (k+1) \binom{n-1}{k} + (n-k) \binom{n-1}{k-1}$$

## 6 Variationsabstand

### 6.1 Abstände/Metriken

Beispiele für Abstände:

1.  $\mathbb{R}^1$ ,  $x, y \in \mathbb{R}$ ,  $d(x, y) = |x - y|$  beschreibt einen Abstand zwischen zwei reellen Zahlen.

#### 2.a Euklidische Metrik:

$\mathbb{R}^2$ ,  $P(x_1, y_1)$   $Q(x_2, y_2)$   $d(P, Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$  beschreibt einen Abstand zwischen zwei Punkten P und Q.

b Taximetrik:

$\mathbb{R}^2$ ,  $d_1(P, Q) = |x_1 - x_2| + |y_1 - y_2|$  beschreibt ebenfalls einen Abstand zwischen zwei Punkten P und Q.

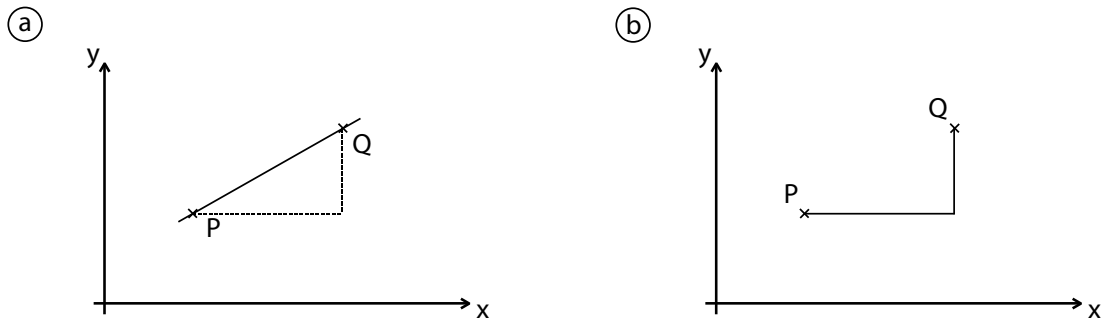


Abbildung 2: a) Euklidische Metrik b) Taximetrik

c diskrete Metrik:

$\mathbb{R}^2$   $d_2(P, Q) = \begin{cases} 1 & P \neq Q \\ 0 & P = Q \end{cases}$  beschreibt ebenfalls einen Abstand zwischen zwei Punkten P und Q, sagt jedoch nichts über dessen Größe aus.

3. Sei  $X$  die Menge aller stetigen Funktionen auf  $[a, b]$   $f, g \in X$

$$d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$$

beschreibt den größten Abstand zweier Funktionen im Intervall  $[a, b]$ .

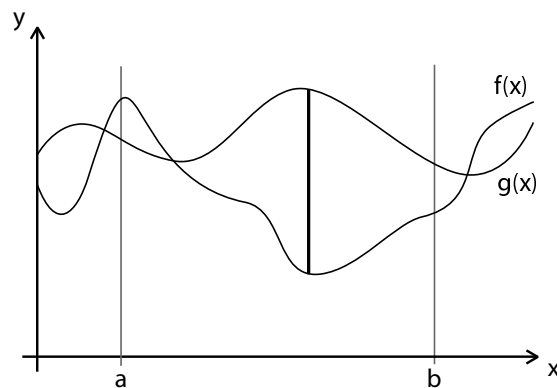


Abbildung 3: Metrik im Funktionenraum

Definition: Metrik

Sei  $X$  eine Menge und  $x, y \in X$ . Dann heißt  $d : X \times X \rightarrow \mathbb{R}$  Metrik, wenn gilt:

- (a)  $d(x, y) > 0$ , falls  $x \neq y$ , und  $d(x, y) = 0$ , falls  $x = y$
- (b)  $d(x, y) = d(y, x)$
- (c)  $d(x, z) \leq d(x, y) + d(y, z) \quad \forall \quad x, y, z \in X$  (Dreiecksungleichung)

Definition: Metrischer Raum

Die Menge, zwischen deren Punkten man einen Abstand definieren kann, heißt Metrischer Raum.

## 6.2 Variationsabstand

Es seien:

- $W_n$  Menge aller Wahrscheinlichkeitsverteilungen auf  $\{1, \dots, n\}$
- $P = (p_1, p_2, \dots, p_n) \in W_n$
- $Q = (q_1, q_2, \dots, q_n) \in W_n$

dann gilt für den Variationsabstand zwischen  $P$  und  $Q$ :

$$2 * \|P - Q\| \rightarrow 2 * d(P, Q) = |p_1 - q_1| + |p_2 - q_2| + \dots + |p_n - q_n|$$

$$\|P - Q\| = \frac{1}{2} \sum_{i=1}^n |p_i - q_i| \leq \frac{1}{2} \sum_{i=1}^n (|p_i| + |q_i|) = \frac{1}{2} \sum_{i=1}^n p_i + \frac{1}{2} \sum_{i=1}^n q_i = \frac{1+1}{2} = 1$$

$$\boxed{\|P - Q\| \leq 1}$$

Mit:

- $S_n$  ist die Gruppe aller Permutationen mit  $n$  Elementen
- $\pi \in S_n$
- $R^{(n)}$  ist die Verteilung nach  $n$  mal Riffle Shuffle
- $R$  ist die Gleichverteilung auf  $S_n$ , d.h.: jede mögliche Permutation hat die Wahrscheinlichkeit  $\frac{1}{n!}$ .

lässt sich der Variationsabstand nach  $n$ -maligem Riffle Shuffle durch

$$\|R^{(n)} - R\| = \frac{1}{2} \sum_{\pi \in S_n} \left| R^{(n)}(\pi) - \frac{1}{n!} \right|$$

beschreiben.

Was bedeutet das für unsere Karten?

Um ein Kartendeck als gut gemischt anzusehen, müssen alle Permutationen ungefähr gleich wahrscheinlich sein. Unser Ziel ist es herauszufinden, nach wie vielen Mischungen der Variationsabstand hinreichend klein ist. Man betrachte den Variationsabstand als Funktion von  $r$  (Anzahl der aufsteigenden Sequenzen):

$$\|R^{(k)} - R\| = \frac{1}{2} \sum_{r=1}^n \left\langle \begin{matrix} n \\ r-1 \end{matrix} \right\rangle \left| \frac{\binom{2^k+n-r}{n}}{2^{nk}} - \frac{1}{n!} \right|$$

Jetzt stellt sich nur die Frage, wie man auf diese Funktion kommt:

Bei einem  $a$ -Shuffle teilt man den Kartenstapel in  $a$  verschiedene Stapel. Man hat den Stapel also  $(a-1)$ -mal geteilt. Von diesen  $a-1$  Teilungen kann man nach dem Mischvorgang  $r-1$  Teilungen, aufgrund der aufsteigenden Sequenzen des Kartenstapels erkennen, wobei  $r$  die Anzahl der aufsteigenden Sequenzen ist. Die restlichen  $a-1-(r-1) = a-r$  Schnitte kann man nach dem Mischen der  $a$  Stapel nicht mehr erkennen. Man hat bei dem Zusammenschieben der Karten  $\binom{a-r+n}{n}$  Möglichkeiten,  $n$  Karten auf  $a-r+n$  Plätzen zu verteilen. Die  $a-r+n$  Plätze ergeben sich dadurch, dass man jeden dieser Plätze entweder mit einem Schnitt oder einer Karte besetzen kann. Die Wahrscheinlichkeit einer Permutation mit  $r$  aufsteigenden Sequenzen ergibt sich aus der Formel  $\frac{\binom{a-r+n}{n}}{a^n}$ , wobei es  $a^n$  Möglichkeiten gibt, die Karten zu mischen.

Für einen Riffle Shuffle mit  $a = 2$  Stapeln, der  $k$ -mal durchgeführt wird ( $2^k$ -Shuffle) ergibt sich folgender Variationsabstand:

$$\|R^{(k)} - R\| = \frac{1}{2} \sum_{i=1}^{n!} \left| R^{(k)}(\pi_i) - \frac{1}{n!} \right| = \frac{1}{2} \sum_{r=1}^n A_{nr} \left| \frac{\binom{2^k+n-r}{n}}{2^{nk}} - \frac{1}{n!} \right|$$

$A_{nr}$  ist die Anzahl der Permutationen von  $n$  Elementen mit  $r$  aufsteigenden Sequenzen und ergibt sich aus der Formel:  $A_{nr} = \left\langle \begin{matrix} n \\ r-1 \end{matrix} \right\rangle$

Das Einsetzen in die Gleichung für den Variationsabstand ergibt:

$$\|R^{(k)} - R\| = \frac{1}{2} \sum_{r=0}^n \left\langle \begin{matrix} n \\ r-1 \end{matrix} \right\rangle \left| \frac{\binom{2^k+n-r}{n}}{2^{nk}} - \frac{1}{n!} \right|$$

Wie man sieht, beträgt der Variationsabstand nach 7 mal Mischen nur noch  $\frac{1}{2}$ . Um eine noch bessere Mischung zu erhalten, sollte man noch 2 bis 3 mal mehr mischen.

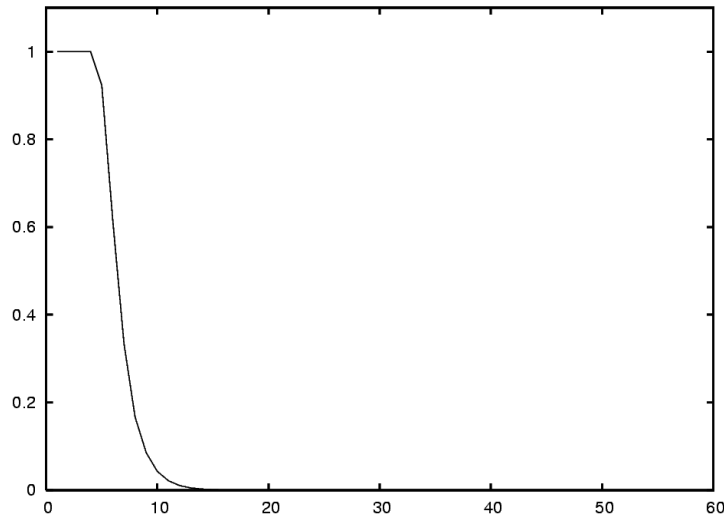


Abbildung 4: Variationsabstand nach  $k$  mal Mischen von 52 Karten

## 7 Modellierung mit Markov-Ketten

$T = \{0, 1, 2, 3, \dots\}$  sei eine Menge von diskreten Zeitpunkten und  $I = \{1, 2, 3, \dots, n\}$  eine endliche Menge diskreter Zustände. Dann heie der Zustand des Systems zum Zeitpunkt  $k$   $X_k$ . Auerdem sei  $P(X_0 = j)$  mit  $j \in I$  die Wahrscheinlichkeit, dass das System im Zustand  $j$  startet. (Man kann eine Markov-Kette auch fur eine stetige Zeit bilden, aber fur diskrete ist sie einfacher zu berechnen und zu erklaren.)

Der zufallige Prozess  $(X_t)$ , also die Folge von Einzelzustanden im Verlaufe der Zeit, heit markovscher Prozess, falls fur alle  $t_i$  und  $j_0, j_1, \dots, j_i$  gilt:

$$P(X_{t_i} = j_i | X_0 = j_0, X_1 = j_1, \dots, X_{t_{i-1}} = j_{i-1}) = P(X_{t_i} = j_i | X_{t_{i-1}} = j_{i-1})$$

Anschaulich bedeutet das, dass es fur den folgenden Schritt nicht wichtig ist, was in samtlichen vorhergehenden Schritten geschehen ist, sondern lediglich, was im letzten Schritt passiert ist. Man bentigt also lediglich die Anfangsverteilung  $P(X_0 = j)_{j=1,2,3,\dots,n}$  und die bergangsverteilungen  $P(X_{t+1} = j | X_t = i)_{i,j=1,2,3,\dots,n}$  fur alle Zeitpunkte  $t$ . Man kann weiter vereinfachen, indem man, was beim Mischen ja auch der Fall ist, annimmt, dass diese bergangswahrscheinlichkeiten zeitunabhngig sind, also es lediglich vom Ausgangspunkt und nicht von der aktuellen Zeit abhngt, wohin man als nchstes geht:  $P(X_{s+1} = j | X_s = i) = p_{i,j}$  mit  $i, j = 1, 2, 3, \dots, n$  und beliebigen  $s \in T$ . Dies liest man:  $p_{i,j}$  ist die Wahrscheinlichkeit vom Zustand  $i$  in den Zustand  $j$  zu wechseln.

Nun kann man eine Übergangsmatrix  $\mathcal{P}$  definieren:

$$\mathcal{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

Die Verteilung im  $m$ -ten Schritt  $p_m$  errechnet sich nach der folgenden Formel:  $p_m = p_0 \cdot \mathcal{P}$ , wobei  $p_0$  die Anfangsverteilung darstellt.  $p_m$  wird hierbei als  $n$ -elementiger Vektor aufgefasst. Wie Folgendes, welches selbiges man sich wiederum anhand eines Baumdiagramms plausibel machen kann, zeigt, entspricht dies auch der erwarteten Formel.

$$\begin{aligned} (P(X_1 = j)) &= \sum_{i=0}^n P(X_0 = i) \cdot P(X_1 = j | X_0 = i)_{j=1,2,3,\dots,n} \\ (P(X_2 = j)) &= \sum_{i=0}^n P(X_1 = i) \cdot P(X_2 = j | X_1 = i)_{j=1,2,3,\dots,n} \\ (P(X_3 = j)) &= \sum_{i=0}^n P(X_2 = i) \cdot P(X_3 = j | X_2 = i)_{j=1,2,3,\dots,n} \\ &\dots \\ (P(X_{t+1} = j)) &= \sum_{i=0}^n P(X_t = i) \cdot P(X_{t+1} = j | X_t = i)_{j=1,2,3,\dots,n} \end{aligned}$$

Obenstehendes sind Wahrscheinlichkeitsverteilungen, wobei die jeweilige Formel bedeutet, dass die Wahrscheinlichkeit zum Zeitpunkt  $t + 1$  im Ort  $j$  zu sein, die Summe der Wahrscheinlichkeiten von jedem der möglichen vorangegangenen Orte  $i$  nach  $j$  zu kommen, die jeweils mit der Wahrscheinlichkeit, zum vorangegangenen Zeitpunkt am Ort  $i$  gewesen zu sein, gewichtet wird.

### Beispiel

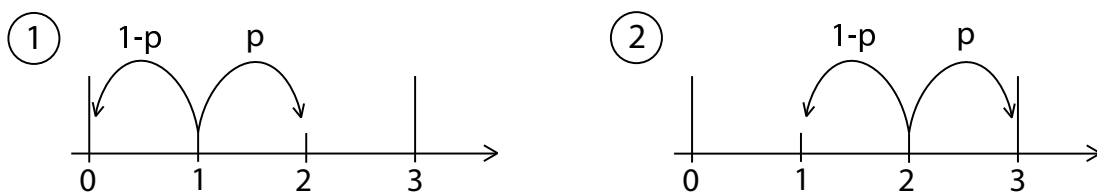


Abbildung 5: Irrfahrt mit reflektierenden Wänden

Anhand des obenstehenden Beispiels lässt sich eine einfache Markov-Kette darstellen. Als Anfangsverteilung sei  $p_a = \{0, 1, 0, 0\}$ . Wir starten also im Punkt



(1). Die Übergangswahrscheinlichkeiten hängen nur vom Ort ab, an dem wir uns befinden.

$$p_0 = \{0, 1, 0, 0\}, p_1 = \{1 - p, 0, p, 0\}, p_2 = \{0, 1 - p, 0, p\}, p_3 = \{0, 0, 1, 0\}$$

Es ist nicht möglich, innerhalb eines Schrittes am selben Ort zu verharren. Da bei (0) und (3) reflektierende Wände stehen, ist man gezwungen, von dort aus zum Herkunftsort zurückzukehren.

Die Formel  $p_m = p_{m-1} \cdot \mathcal{P}$  mit  $p_0$  lässt sich als Fixpunktproblem betrachten und mit den entsprechenden Verfahren lösen.

Allgemein lässt sich beweisen, dass für eine sogenannte stochastische Matrix (einer Matrix wie obenstehendes  $\mathcal{P}$  mit den Eigenschaften  $0 \leq p_{i,j} \leq 1$  und  $\sum_{j=0}^n p_{ij} = 1$  für alle  $i, j$  bzw.  $i$ ), dass wenn sie nach der Erhebung in eine ausreichend hohe Potenz nur noch positive Elemente aufweist, das zugehörige System einen sogenannten stationären Zustand besitzt, was einem Fixpunkt der linearen Abbildung entspricht.

Für das obenstehende Beispiel existiert ein solcher stationärer Zustand mit  $p^* = \{\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\}$

Zur Anwendung auf den Riffle Shuffle mit drei Karten gehen wir von einer Anfangsverteilung aus, die der identischen Permutation eine Wahrscheinlichkeit von 1 zuweist (was aber im Prinzip irrelevant ist), und benötigen nun noch die Übergangsmatrix  $\mathcal{P}$ .

Permutation	$j$	123	132	321	312	213	231
$i$	Kartenfolge	123	132	321	231	213	312
123	123	$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
132	132	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$
321	321	0	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
312	231	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$
213	213	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	0
231	312	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$

Wie man leicht überprüfen kann, hat bereits  $\mathcal{P}^2$  nur noch positive Elemente, was bedeutet, dass es einen stationären Zustand gibt. Durch Lösung des Gleichungssystems  $p^* = p^* \cdot \mathcal{P}$  lässt sich dieser ermitteln und stellt sich auch wirklich als die Gleichverteilung heraus.

Mit Hilfe des Variationsabstands ergibt sich für sieben Mischungen mit dem Riffle Shuffle bei einem Deck von 52 Karten ein Abstand von etwa 0,5 von der Gleichverteilung, für zehn Mischungen rund 0,04, wie es sich auch schon durch Betrachtung der Funktion ergab.

## Fixpunkte und ihre Anwendungen

### *Teilnehmer:*

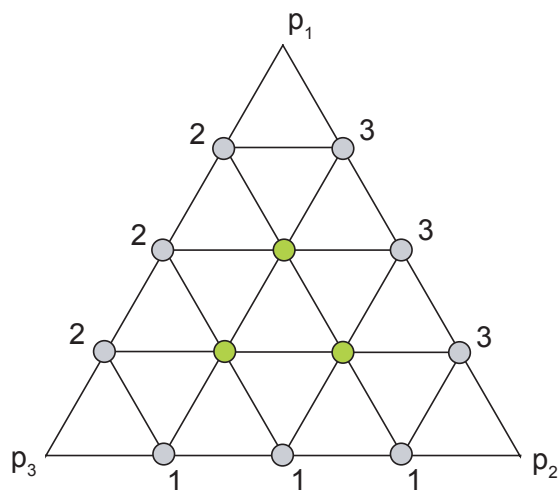
Thomas Bünger	Heinrich-Hertz-Oberschule
Erik Esche	Herder-Oberschule
Felix Günther	Herder-Oberschule
Arne Müller	Herder-Oberschule
Andrej Stepanchuk	Heinrich-Hertz-Oberschule

### *Gruppenleiter:*

Bernd Kummer	Humboldt-Universität zu Berlin
--------------	--------------------------------

Die Existenz von Lösungen in Gleichungen und anderen Problemen wird oft über Fixpunktsätze bewiesen (wann ist  $f(x) = x$  lösbar?). Wir haben zwei Beweisprinzipien (von Banach und Brouwer) für solche Sätze studiert und auf konkrete Aufgaben angewandt.

Stichworte: konvexe Mengen, metrischer Raum (Eigenschaften, Beispiele), iterative Lösungsverfahren, Gleichungen für gesuchte Funktionen, Lösungen einfacher Spiele.



# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>22</b>
<b>2</b>	<b>Vorbereitende Definitionen und Aussagen</b>	<b>22</b>
2.1	Metrik . . . . .	22
2.1.1	Beispiele . . . . .	22
2.2	Stetigkeit . . . . .	23
2.3	Konvergenz . . . . .	23
2.3.1	Beispiele . . . . .	24
2.4	Vollständige Räume . . . . .	24
<b>3</b>	<b>Banach's Fixpunktsatz</b>	<b>25</b>
<b>4</b>	<b>Beweis zur Vollständigkeit von <math>C[a, b]</math></b>	<b>26</b>
<b>5</b>	<b>Methode der sukzessiven Approximation</b>	<b>26</b>
5.1	Beispiel . . . . .	28
<b>6</b>	<b>Brouwer's Fixpunktsatz</b>	<b>28</b>
6.1	Beweis nach Scarf und Hoang Tuy . . . . .	28
6.1.1	Fixpunktbedingung . . . . .	29
6.1.2	Unterteilung des Simplex . . . . .	30
6.2	Erweiterung auf beliebige konvexe Mengen . . . . .	33
<b>7</b>	<b>Sperners Lemma</b>	<b>35</b>
7.1	Konvexe Hülle . . . . .	35
7.2	Lemma . . . . .	35
7.2.1	Sehen . . . . .	36
7.2.2	Gesehen werden . . . . .	37
<b>8</b>	<b>Minimax-Satz</b>	<b>37</b>
8.1	Beispiel . . . . .	37
8.2	Konvexität . . . . .	38
8.3	Beweis des Minimax-Satzes . . . . .	38
8.4	Beweis der Stetigkeit von $X^*$ und $Y^*$ . . . . .	39
8.5	Allgemeiner Fall ohne Striktheit . . . . .	39
<b>9</b>	<b>Retrakte</b>	<b>39</b>
9.1	Satz . . . . .	39
9.2	Äquivalenz . . . . .	39
9.3	Beweis über den Brouwer'schen Fixpunktsatz . . . . .	40
9.4	Beweis von Brouwers Fixpunktsatz über den Retraktsatz . . . . .	40

# 1 Einführung

Eine *stetige* reelle Funktion  $f = f(x)$  möge das Intervall  $X = [0, 1]$  in sich selbst abbilden, d.h.  $f(x) \in X$  für alle  $x \in X$  und kurz  $f : X \rightarrow X$ .

Dann erkennt man leicht, dass es stets ein solches  $x^* \in X$  gibt, welches

$$f(x^*) = x^*$$

erfüllt. Es heißt *Fixpunkt* von  $f$ .

Wenn  $f$  hinreichend "flach" wächst oder fällt, ist der Fixpunkt sogar eindeutig bestimmt.

Offenbar kann es im Allgemeinen viele Fixpunkte geben (z.B. wenn  $f(x) = x$ ), und für Funktionen, die unsere Voraussetzungen der Stetigkeit bzw.  $f(x) \in X$  nicht erfüllen, auch gar keine.

Weiter kann man auch dann Fixpunkte haben, wenn  $X$  irgendeine andere Menge ist. Allerdings - da man Stetigkeit zu brauchen scheint - sollte man diesen wichtigen Begriff schon auf Funktionen  $f : X \rightarrow X$  übertragen können.

## 2 Vorbereitende Definitionen und Aussagen

### 2.1 Metrik

Eine Menge  $X$  mit der Funktion  $d$  heißt *metrischer Raum* (anschaulich Menge mit Abstand), wenn  $d$  jedem Paar von Elementen  $x, y \in X$  einen "Abstand"  $d(x, y) \geq 0$  zuordnet, sodass gilt:

- (i)  $d(x, y) = 0 \Leftrightarrow x = y$ ,
- (ii)  $d(x, y) = d(y, x) \quad \forall x, y \in X$ ,
- (iii)  $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in X$  (Dreiecksungleichung).

Die Funktion  $d$  heißt auch *Metrik*. Den Trivialfall  $X = \emptyset$  schließen wir aus. Der metrische Raum selbst wird oft als Paar  $(X, d)$  bezeichnet. Er ändert sich, wenn man  $d$  ändert.

#### 2.1.1 Beispiele

1.  $X = \mathbb{R}$ ,  $d(x, y) = |x - y|$ .
2.  $X = \mathbb{R}$ ,  $d(x, y) = 0$  wenn  $x = y$ ,  $d(x, y) = 4711$  sonst.
3.  $X = \text{Ebene}$ ,  $d(x, y) = \text{euklidischer Abstand von } x \text{ und } y$ .

## 2.2 Stetigkeit

Sind  $(X_1, d_1)$  und  $(X_2, d_2)$  zwei metrische Räume, so heißt  $f : X_1 \rightarrow X_2$  *stetig in*  $x^* \in X_1$ , wenn es zu jedem (noch so kleinen)  $\varepsilon > 0$  ein  $\delta > 0$  gibt, sodass

$$d_2(f(x), f(x^*)) < \varepsilon \quad \forall x \in X_1 \text{ mit } d_1(x, x^*) < \delta.$$

Ist  $f$  in jedem  $x^* \in X_1$  stetig, so heißt  $f$  *stetig*.

Gibt es eine Konstante  $L$ , sodass

$$d_2(f(x), f(y)) \leq L \cdot d_1(x, y) \quad \forall x, y \in X_1, \quad (1)$$

so heißt  $f$  *Lipschitz-stetig*. Bekanntlich ist jede Lipschitz-stetige Funktion auch stetig. Dann reicht es, oben einfach  $\delta < \varepsilon \cdot L^{-1}$  zu wählen.

Wir haben es zumeist mit dem Fall zu tun, dass beide metrischen Räume gleich sind;  $(X_1, d_1) = (X_2, d_2) = (X, d)$ . In diesem Fall definiert man:

Eine Lipschitz-stetige Funktion  $f : X \rightarrow X$  heißt *kontraktiv*, wenn (1) mit einer Konstanten  $L < 1$  gilt. Dann liegen die Bilder von  $f$ , nämlich  $f(x)$  und  $f(y)$ , stets näher zusammen als die Urbilder  $x$  und  $y$  (falls  $x \neq y$ ).

Für reelle Funktionen  $f$  mit dem üblichen Betragsabstand bedeutet dies, daß  $f$  hinreichend "flach" wächst oder fällt.

## 2.3 Konvergenz

Analog zum Reellen definiert man die *Konvergenz einer (unendlichen) Folge*  $\{x_n\}_{n=1,2,\dots}$  von Elementen  $x_n \in X$  gegen ein  $x \in X$  im metrischen Raum  $(X, d)$ .

Die Folge  $\{x_n\}$  konvergiert gegen  $x$  (kurz  $x_n \rightarrow x$  für  $n \rightarrow \infty$  oder  $x = \lim_{n \rightarrow \infty} x_n$ ), wenn zu jedem  $\varepsilon > 0$  ein  $n(\varepsilon)$  existiert, sodass

$$\forall n > n(\varepsilon) : d(x_n, x) < \varepsilon.$$

Mit anderen Worten:  $x = \lim_{n \rightarrow \infty} x_n$  bedeutet, dass die reelle Folge  $\{a_n\}_{n=1,2,\dots}$  der Abstände  $a_n = d(x_n, x)$  gegen Null konvergiert.

Dann folgt stets aus der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < 2\varepsilon \quad \forall n, m > n(\varepsilon). \quad (2)$$

Nimmt man in (2) das spezielle  $\varepsilon' = \frac{1}{2}\varepsilon$  und setzt  $n(\varepsilon) = n(\varepsilon')$ , so folgt:

Falls

$$x = \lim_{n \rightarrow \infty} x_n,$$

so gibt es zu jedem  $\varepsilon > 0$  ein  $n(\varepsilon)$  mit

$$\forall n, m > n(\varepsilon) : d(x_n, x_m) < \varepsilon. \quad (3)$$

Eine Folge  $(x_n)$  mit dieser (schwächeren) Eigenschaft heißt *Cauchy-Folge*, oft auch *Fundamentalfolge* genannt.

### 2.3.1 Beispiele

1.  $X$  bestehe aus allen reellen Zahlen ohne die Null,  $d(x, y) = |x - y|$  sei der normale Abstand. Dann ist die Folge  $\{x_n\}_{n=1,2,\dots}$  mit  $x_n = \frac{1}{n}$  eine Cauchy-Folge, sie hat aber keinen Limes  $x$ , weil es den im Reellen existierenden Limes  $x = 0$  in  $X$  nicht gibt.

2. Auch die rationalen Zahlen bilden mit ihrem normalen Abstand einen metrischen Raum, der nicht-konvergente Cauchy-Folgen besitzt. Die Elemente der Folge  $(x_n) = (1 + \frac{1}{n})^n$  sind alle rationale Zahlen, allerdings liegt der Grenzwert in den irrationalen Zahlen:

$$\lim_{n \rightarrow \infty} x_n = e.$$

## 2.4 Vollständige Räume

Das motiviert die Definition:

Ein metrischer Raum  $(X, d)$  heißt *vollständig*, wenn jede seiner Cauchy-Folgen auch konvergiert.

**Lemma 1** *Die reellen Zahlen bilden mit  $d(x, y) = |x - y|$  einen vollständigen metrischen Raum.*

**Beweis:** Der Beweis ist durchaus nicht trivial, weil man zunächst auf einer geeigneten Definition der reellen Zahlen (auf Grundlage der rationalen Zahlen) aufbauen muss. Wir helfen uns aus Zeitgründen mittels der Dezimaldarstellung (ohne uns um deren genauen Bezug zur reellen Zahl zu kümmern): Hat man eine Cauchy Folge reeller Zahlen  $x_n$ , so kann man  $\varepsilon = 10^{-2k}$  setzen, um zu sehen, dass für  $n, m > n(\varepsilon)$  die ersten  $k$  Kommastellen von  $x_n$  und  $x_m$  gleich sein müssen. Macht man  $k$  beliebig groß ( $k \rightarrow \infty$ ) und bildet  $x$  aus den dann festen Kommastellen, folgt  $x = \lim_{n \rightarrow \infty} x_n$ .

! Aber: Dieser Schluss ist nur dann richtig, wenn man vorher die Dezimaldarstellung eindeutig macht, indem man (z.B.) vereinbart, dass Ziffernfolgen (mit oder ohne Komma dazwischen) wie etwa 2000... durch 1999... ersetzt werden!

Ansonsten bekämen wir schon mit der Folge  $x_n = 1,000\dots(n \text{ gerade}); x_n = 0,999\dots(n \text{ ungerade})$  Probleme.  $\square$

Auch die metrischen Räume zu Beispiel 2 und 3 sind vollständig.

**Lemma 2** *(geometrische Reihe)*

Sei  $0 < q < 1$  und  $s_n = 1 + q + q^2 + \dots + q^n$ . Dann gilt  $s_n = \frac{1-q^{n+1}}{1-q} < \frac{1}{1-q}$ .

**Beweis:** Es ist nach Ausmultiplizieren  $s_n(1 - q) = (1 + q + q^2 + \dots + q^n)(1 - q) = 1 - q^{n+1}$ . Division durch  $1 - q$  ergibt die Behauptung.  $\square$

### 3 Banach's Fixpunktsatz

**Theorem 1** (*Banach's Fixpunktsatz*)

Es sei  $(X, d)$  ein vollständiger metrischer Raum und  $f : X \rightarrow X$  kontraktiv. Dann besitzt  $f$  genau einen Fixpunkt  $x^*$ .

Zusatz: Mit jedem  $x_0 \in X$  gilt  $x^* = \lim_{n \rightarrow \infty} x_n$ , wenn  $x_{n+1} = f(x_n) \forall n = 0, 1, 2, \dots$ .

**Beweis:** Wir wissen, dass mit einem  $L = q \in ]0, 1[$  wegen der Kontraktivität gilt

$$d(x_{n+2}, x_{n+1}) = d(f(x_{n+1}), f(x_n)) \leq q \cdot d(x_{n+1}, x_n) \quad , \text{ falls } n \geq 0.$$

Also erhält man  $d(x_2, x_1) \leq q \cdot d(x_1, x_0)$ ,  $d(x_3, x_2) \leq q \cdot d(x_2, x_1) \leq q^2 \cdot d(x_1, x_0)$  und per vollständiger Induktion  $d(x_{n+1}, x_n) \leq q^n \cdot d(x_1, x_0)$ . Das bringt nach mehrmaliger Anwendung der Dreiecksungleichung

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \dots + d(x_{n+1}, x_n) \\ &\leq q^{n+k-1} \cdot d(x_1, x_0) + q^{n+k-2} \cdot d(x_1, x_0) + \dots + q^n \cdot d(x_1, x_0) \\ &= (q^{k-1} + q^{k-2} + \dots + 1) \cdot q^n \cdot d(x_1, x_0) \\ &\leq \frac{1}{1-q} \cdot q^n \cdot d(x_1, x_0). \end{aligned}$$

Ist nun  $\varepsilon > 0$  beliebig gegeben, findet sich ein natürliches  $N$ , sodass

$$\frac{1}{1-q} \cdot q^N \cdot d(x_1, x_0) < \varepsilon.$$

Dies zeigt uns mit  $n(\varepsilon) = N$  und  $m = n + k$  in (3), dass die Folge  $\{x_n\}$  eine Cauchy-Folge ist. Wegen der Vollständigkeit von  $(X, d)$  existiert also der Grenzwert  $x^* = \lim_{n \rightarrow \infty} x_n$ .

**Fixpunkteigenschaft von  $x^*$ :** Es gilt:

$$\begin{aligned} d(f(x^*), x^*) &\leq d(f(x^*), x_{n+1}) + d(x_{n+1}, x^*) \\ &= d(f(x^*), f(x_n)) + d(x_{n+1}, x^*) \\ &\leq q \cdot d(x^*, x_n) + d(x_{n+1}, x^*). \end{aligned}$$

Die rechte Seite wird beliebig klein (für  $n \rightarrow \infty$ ) wegen  $x^* = \lim_{n \rightarrow \infty} x_n$ . Also muss  $d(f(x^*), x^*) = 0$  und damit  $f(x^*) = x^*$  sein.

**Eindeutigkeit des Fixpunktes  $x^*$ :**

Gibt es einen zweiten Fixpunkt  $x^{**} = f(x^{**})$ , so folgt

$$d(x^*, x^{**}) = d(f(x^*), f(x^{**})) \leq q \cdot d(x^*, x^{**}).$$

Das kann mit  $0 < q < 1$  aber nur gelten, wenn  $d(x^*, x^{**}) = 0$ , also  $x^* = x^{**}$  ist. Damit ist alles gezeigt.  $\square$

Es ist von Nutzen, dass der Beweis gleich die im Zusatz angegebene Konstruktionsmethode für den Fixpunkt mitliefert. Die Methode heißt auch *Methode der sukzessiven Approximation*.

## 4 Beweis zur Vollständigkeit von $C[a, b]$

**Definition**  $C[a, b]$  besteht aus allen auf  $[a, b]$  definierten, stetigen, reellen Funktionen  $x = x(t)$ . Es wird folgende Abstandsfunktion  $d$  vereinbart:

$$d(x, y) = \max_{t \in [a, b]} |x(t) - y(t)|.$$

**Behauptung:**  $C[a, b]$  ist vollständig.

**Beweis:** Offensichtlich genügt es zu zeigen: Wenn  $(x_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge ist, so existiert  $x = \lim_{n \rightarrow \infty} x_n$  als Element von  $C[a, b]$ .

Es ist Folgendes bekannt:

$$\forall \varepsilon > 0 \exists n(\varepsilon) : d(x_n, x_m) < \varepsilon \quad \forall n, m > n(\varepsilon).$$

Wir betrachten bei festem  $t \in [a, b]$  die reellen  $x_m(t)$ . Dann ist  $|x_n(t) - x_m(t)| < \varepsilon$  mit  $n > n(\varepsilon), m \geq n(\varepsilon)$ . Wenn  $n$  fest ist und wir  $m$  gegen Unendlich gehen lassen ( $m \rightarrow \infty$ ), dann ist die linke Seite eine Cauchy-Folge in  $\mathbb{R}$  und damit existiert:  $x(t) = \lim_{m \rightarrow \infty} x_m(t)$ . Wegen  $|x_n(t) - x_m(t)| \leq \varepsilon$  und  $|x_m(t) - x(t)| \rightarrow 0$  für  $n > n(\varepsilon)$  und  $m \rightarrow \infty$ :

$$|x_n(t) - x(t)| \leq |x_n(t) - x_m(t)| + |x_m(t) - x(t)| \leq \varepsilon + |x_m(t) - x(t)| \leq 2 \cdot \varepsilon. \quad (4)$$

Hieraus folgt:

$$|x_n(t) - x(t)| \leq 2 \cdot \varepsilon, \text{ falls } n > n(\varepsilon).$$

Mit Hilfe dieser Gleichungen beweisen wir nun die Stetigkeit von  $x = x(t)$ :

Sei  $t, s \in [a, b]$ . Da jede auf einem abgeschlossenen Intervall stetige Funktion gleichmäßig stetig ist, gilt:  $|x_n(t) - x_n(s)| < \varepsilon$ , falls  $|t - s| < \delta = \delta(\varepsilon)$ . Mit (4) erhalten wir dann:

$$|x(t) - x(s)| \leq |x(t) - x_n(t)| + |x_n(t) - x_n(s)| + |x_n(s) - x(s)| \leq 2 \cdot \varepsilon + \varepsilon + 2 \cdot \varepsilon = 5 \cdot \varepsilon.$$

Für  $s \rightarrow t$  geht  $\varepsilon \rightarrow 0$ , somit haben wir die Stetigkeit von  $x$  bewiesen und damit auch die Vollständigkeit von  $C[a, b]$ .  $\square$

## 5 Methode der sukzessiven Approximation

Gegeben sei die Differentialgleichung  $\dot{x}(t) = f(x(t), t) \quad \forall t$  mit dem gegebenen reellen Anfangswert  $x(0) = x_0$ . Desweiteren sei bekannt, dass  $f$  ganz  $\mathbb{R}^2$  nach  $\mathbb{R}$  abbildet. Außerdem soll  $f$  bezüglich  $x$  eine einheitliche Lipschitz-Bedingung erfüllen. Dies bedeutet für  $f$ , dass ein  $L > 0$  für alle  $\xi$  mit  $|f(\xi_1, s) - f(\xi_2, s)| \leq L \cdot$



$|\xi_1 - \xi_2|$  existiert. Die Lösung dieser Differentialgleichung führt zu einem Fixpunkt der folgenden Funktion:

$$y = y(t) := x_0 + \int_0^t f(x(s), s) ds,$$

dabei sollte  $x \in C[a, b]$  und  $0 \in ]a, b[$  sein. Es gilt also:  $\dot{y}(t) = f(x(t), t)$ . Sei  $F : C[a, b] \rightarrow C[a, b]$  so gewählt, dass für jeden Index  $k \geq 0$  gilt:

$$F(x_k) = x_{k+1} = x_{k+1}(t) = x_0 + \int_0^t f(x_k(s), s) ds. \quad (5)$$

Im Folgenden zeigen wir, dass  $F$  *kontraktiv* ist, hierfür betrachten wir  $y_1 = F(x_1)$  und  $y_2 = F(x_2)$ :

$$\begin{aligned} |y_2(t) - y_1(t)| &\leq \left| \int_0^t f(x_2(s), s) ds - \int_0^t f(x_1(s), s) ds \right| \\ &\leq \left| \int_0^t |f(x_2(s), s) - f(x_1(s), s)| ds \right| \\ &\leq \left| \int_0^t L \cdot |x_2(s) - x_1(s)| ds \right| \quad (\text{Lipschitz - Bedingung}) \\ &\leq \left| \int_0^t L \cdot d(x_2, x_1) ds \right| \\ &\leq L \cdot |d(x_2, x_1) \cdot (b - a)|. \end{aligned}$$

Hieraus folgt:

$$d(y_2, y_1) \leq L \cdot (b - a) \cdot d(x_2, x_1) =: q \cdot d(x_2, x_1).$$

Damit ist  $F$  eine kontraktive Abbildung, falls  $b - a$  hinreichend klein ist und damit  $q \in ]0, 1[$ . Analog erhalten wir mit dem Banach'schen Fixpunktsatz, dass genau ein Fixpunkt von  $F$  existiert. Diesen erhalten wir mit der Iterationsvorschrift (5). Falls zwei Funktionen  $x = x(t)$  und  $y = y(t)$  anhand der Differentialgleichungen:

$$\dot{x}(t) = f(x(t), y(t), t) \text{ und } \dot{y}(t) = g(x(t), y(t), t)$$

gesucht werden, dann funktioniert die Methode der sukzessiven Approximation analog. Es wird ein  $\vec{x}(t) = (x(t), y(t))$  gesucht. Desweiteren gilt:  $x_{k+1}(t) = x_0 + \int_0^t f(x_k(s), y_k(s), s) ds$  und  $y_{k+1}(t) = y_0 + \int_0^t g(x_k(s), y_k(s), s) ds$ . Für  $n$  gesuchte Funktionen funktioniert die Methode ebenfalls analog.

## 5.1 Beispiel

Gegeben ist die Differentialgleichung  $\dot{x}(t) = x(t)$  mit dem Anfangswert  $x(0) = 1$ . Sukzessives Verwenden der Iterationsvorschrift (5) ergibt:

$$\begin{aligned}x_1(t) &= 1 \\x_2(t) &= 1 + \int_0^t 1 \, ds = 1 + t \\x_3(t) &= 1 + \int_0^t (1 + s) \, ds = 1 + t + \frac{t^2}{2} \\&\vdots \\x_k(t) &= 1 + \cdots + \frac{t^{k-1}}{(k-1)!} \\ \\ \lim_{k \rightarrow \infty} x_k(t) &= e^t\end{aligned}$$

Die  $e$ -Funktion erfüllt also die gegebene Differentialgleichung.

## 6 Brouwer's Fixpunktsatz

Dieser Satz stellt unter anderen Voraussetzungen die Existenz mindestens eines Fixpunktes fest. Das Beweisprinzip ist völlig anders.

Man denke sich in der Ebene  $X$ , deren Punkte  $x$  wir mit kartesischen Koordinaten  $x = (x_1, x_2)$  beschreiben, eine *konvexe* Menge  $M$  (nicht leer). *Konvex* bedeutet, daß jeweils die gesamte Verbindungsstrecke zweier Punkte  $x$  und  $y$  in  $M$  liegt, sofern  $x$  und  $y$  aus  $M$  sind (alle Punkte in einem Kreis, Rechteck, Dreieck, ... bilden konvexe Mengen). Die Menge  $M$  sei weiter *beschränkt* (sie enthalte keinen Strahl) und *abgeschlossen* (sie enthalte alle ihre 'Randpunkte'). Mit dem gewöhnlichen euklidischen Abstand  $d$  sind  $X$  und die Menge  $M$  metrische Räume. Damit sind stetige Funktionen  $f : M \rightarrow M$  definiert.

Völlig analog kann man vereinbaren, was *abgeschlossene, beschränkte konvexe Mengen und stetige Funktionen* im 3-dimensionalen euklidischen Raum und im  $n$ -dimensionalen sind.

### Theorem 2 (Brouwer's Fixpunktsatz)

Sei  $M \subset \mathbb{R}^n$  eine konvexe, abgeschlossene, beschränkte, nichtleere Menge im euklidischen Raum und  $f : M \rightarrow M$  eine stetige Funktion, so existiert mindestens ein Fixpunkt (nach Lutzien Egbertus Jan Brouwer).

## 6.1 Beweis nach Scarf und Hoang Tuy

Wir betrachten einen konstruktiven Beweis nach Scarf und Hoang Tuy. Zuerst betrachten wir, dass es sich bei  $M$  um ein Simplex handelt, den wir später verallgemeinern.

Zur Veranschaulichung benutzen wir Skizzen für  $n = 2$ .

**Lemma 3** *Jeder Punkt  $x \in M$  lässt sich eineindeutig als konvexe Kombination der Eckpunkte  $p_i$  des Simplex darstellen:*

$$x = \sum_{i=1}^{n+1} \alpha_i \cdot p_i \quad \sum_{i=1}^{n+1} \alpha_i = 1 \quad \forall_{i \in \mathbb{N}^*} i \leq n+1 : \alpha_i \geq 0$$

Der Beweis hierfür basiert auf der Lösbarkeit von linearen Gleichungssystemen.

Für unseren Spezialfall ( $n = 2$ ) ergibt sich also:

$$x = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3.$$

Da die Funktion  $f$  selbstabbildend ist, ergibt sich analog:

$$f(x) = \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3.$$

Da sich jedem  $x$  eineindeutig ein  $\alpha$ -Tupel und jedem  $f(x)$  ein  $\lambda$ -Tupel zuweisen lässt und darüber hinaus  $f$  stetig ist, ist auch die Abbildung der Alphas auf die Lambdas stetig.

### 6.1.1 Fixpunktbedingung

Die allgemeine Fixpunktbedingung  $f(x) = x$  ist bei unserem Fall äquivalent zu der Bedingung

$$\forall_{i \in \mathbb{N}^*} i \leq n+1 : \quad \alpha_i = \lambda_i.$$

Unter der Nebenbedingung, dass die Alphas und Lambdas Konvexkombinationen der Eckpunkte definieren, ist die folgende Bedingung ebenfalls äquivalent dazu:

$$\forall_{i \in \mathbb{N}^*} i \leq n+1 : \quad \alpha_i \leq \lambda_i.$$

Beweis: Die Aufsummierung von diesen  $n+1$  Ungleichung ergibt:

$$\underbrace{\alpha_1 + \alpha_2 + \dots + \alpha_{n+1}}_{=1} \leq \underbrace{\lambda_1 + \lambda_2 + \dots + \lambda_{n+1}}_{=1}.$$

Dies wiederum kann natürlich nur erfüllt werden, wenn für alle  $i$   $\alpha_i = \lambda_i$  gilt, da sonst die linke Seite echt kleiner wäre.

### 6.1.2 Unterteilung des Simplex

Das Ausgangssimplex lässt sich nun rekursiv zerteilen. Hierfür werden beispielsweise für den Fall  $n = 2$  (also das Dreieck in der Ebene) die Mittelpunkte der drei Seiten verbunden, wodurch die Figur in vier kongruente Teildreiecke zerfällt. Dies wird auch als erste Zerlegung bezeichnet.

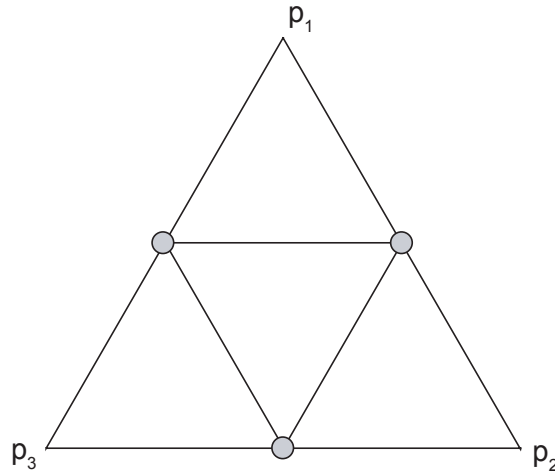


Abbildung 1: die erste Unterteilung des Simplex

Die zweite Unterteilung ergibt sich durch das Anwenden desselben Verfahrens auf die entstandenen Teildreiecke. Analog dazu ergibt sich die  $N$ -te Unterteilung. Die Eckpunkte aller so entstandenen Teilsimplizes bilden ein Raster.

All diesen Rasterpunkten kann nun mit einer geeigneten Indexfunktion ein Wert zugeordnet werden.

Die Menge  $I$  sei die Menge der gültigen Indexwerte und ist abhängig von der Stelle  $x$  (und damit von Alpha und Lambda) mit

$$I(x) = \{i | \alpha_i \leq \lambda_i\}.$$

Jede beliebige Indexfunktion  $i(x)$  heißt *zulässig*, wenn gilt:

1.  $i(x) \in I(x) \quad \forall x$
2.  $i(x) = j$ , falls  $\alpha_j = 0$  und  $x$  keine Ecke ist

Nun besitzt jedes Teilsimplex an seinen Eckpunkten bestimmte Indizes. Die nachfolgende Skizze zeigt die triviale Indizierung der Eckpunkte bei einem Dreieck mit der zweiten Unterteilung. Die Punkte auf den Kanten des Dreiecks besitzen in ihrer Konvexkombination keinen Anteil des gegenüberliegenden Punktes. Somit

müssen nach der zweiten Regel der Indexfunktionen alle Punkte auf der Kante (die dem Punkt  $p_j$  gegenüber liegt) mit  $j$  bezeichnet werden, da  $\alpha_j = 0$  ist. Die Eckpunkte werden nicht indiziert und in den nachfolgenden Betrachtungen ignoriert.

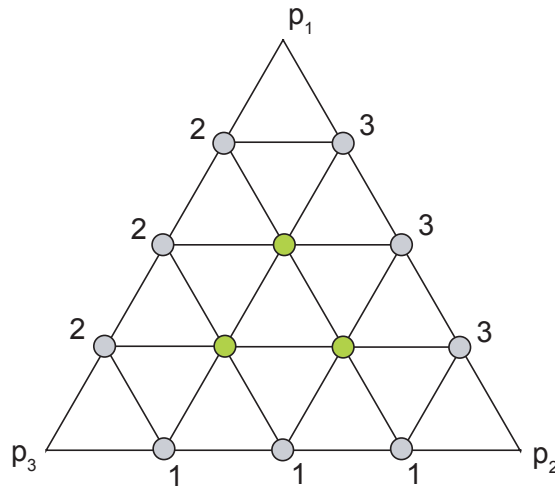


Abbildung 2: die  $N$ -te Unterteilung des Simplex (mit  $N = 2$ )

Wenn für ein Teilsimplex die Indizes aller Eckpunkte unterschiedlich sind, also alle Werte von 0 bis  $n + 1$  an den  $n + 1$  Ecken vorkommen, so bezeichnet man ein solches Teilsimplex als *normal*.

Im Fall  $n = 2$  handelt es sich um Dreiecke, während im Dreidimensionalen entsprechend Tetraeder mit 4 Ecken verwendet werden. Für höhere Dimensionen gilt entsprechendes.

Weiterhin werden solche Simplexseiten, die an Ecken mit genau den Nummern 1 bis  $n$  angrenzen, *ausgezeichnet* genannt. Im Zweidimensionalen sind diese Seiten die Kanten eines Dreiecks und im Dreidimensionalen entsprechend die Flächen eines Tetraeders.

Im Folgenden lässt sich nun ein Algorithmus beschreiben um ein solches normales Teilsimplex zu finden.

Wir starten im Eckpunkt  $p_{n+1}$ . Das dazugehörige Teilsimplex hat genau eine ausgezeichnete Seite, was sich aus der trivialen Indizierung der angrenzenden Simplexseiten ergibt. Nun wechselt man in das Teilsimplex, das ebenfalls an die ausgezeichnete Seite angrenzt. Nun ergeben sich zwei Fälle:

1. Das neue Teilsimplex ist normal.
2. Das neue Teilsimplex ist nicht normal. Somit existiert ein Index, der doppelt vergeben ist. Da es bei einem Simplex zu jeder Kombination von Eckpunk-

ten eine Seite gibt, lässt sich diejenige Ecke, deren Index zweimal vorkommt, nun austauschen. Dadurch entsteht zwingenderweise genau eine weitere ausgezeichnete Seite.

Im ersten Fall endet der Algorithmus hier und wir haben ein normales Teilsimplex gefunden. Im zweiten Fall verlassen wir einfach das Teilsimplex durch die zweite ausgezeichnete Seite und betreten ein weiteres Teilsimplex.

Dies wiederholt man solange bis ein normales Teilsimplex gefunden wurde. Da die Anzahl an Teilsimplizes endlich und somit beschränkt ist, muss auch der Algorithmus terminieren. Denn jedes Teilsimplex besitzt höchstens zwei ausgezeichnete Seiten, kann somit also nur einmal in genau einer Richtung “durchlaufen“ werden, weshalb keine Zyklen entstehen können. Dadurch muss immer irgendwann ein normales Teilsimplex erreicht werden.

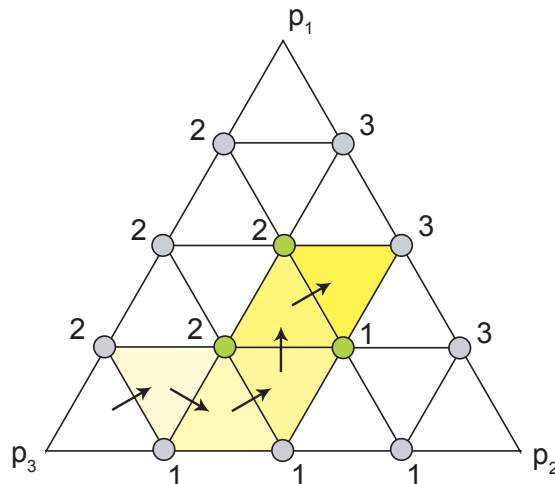


Abbildung 3: Weg durch den Simplex zu einem normalen Teilsimplex

Betrachten wir nun ein solches normales Teilsimplex näher. Da alle Indizes vorkommen, gibt es zu jedem  $i$  einen Eckpunkt  $h_i$  mit diesem Index  $i$ . Somit gilt:

$$\alpha_i(h_i) \leq \lambda_i(h_i) \quad \forall i$$

Dadurch werden *alle* Fixpunktgleichungen (allerdings nur in verschiedenen Punkten) im Teilsimplex erfüllt.

Aufgrund der gleichmäßigen Stetigkeit (stetig auf einer kompakten Menge) gilt:

$$\begin{aligned} \forall \varepsilon > 0 \quad \exists \delta > 0 & : \quad d(y, x) \leq \delta \quad \Rightarrow \quad d(\alpha_i(y), \alpha_i(x)) \leq \varepsilon \\ \forall \varepsilon > 0 \quad \exists \delta > 0 & : \quad d(y, x) \leq \delta \quad \Rightarrow \quad d(\lambda_i(y), \lambda_i(x)) \leq \varepsilon \end{aligned}$$

Somit lässt sich  $\alpha_i(x)$  beliebig genau durch  $\alpha_i(y)$  und  $\lambda_i(x)$  beliebig genau durch  $\lambda_i(y)$  annähern. Dies lässt sich nun auf die Fixpunktungleichungen anwenden:

$$\begin{aligned} & \alpha_i(x) \leq \lambda_i(x) & (6) \\ \Leftrightarrow \alpha_i(y) - \varepsilon & \leq \alpha_i(x) \leq \lambda_i(x) \leq \lambda_i(y) + \varepsilon \\ \Leftrightarrow & \alpha_i(y) - \varepsilon \leq \lambda_i(y) + \varepsilon \\ \Leftrightarrow & \alpha_i(y) \leq \lambda_i(y) + 2\varepsilon & (7) \end{aligned}$$

Dank der gleichmäßigen Stetigkeit gibt es nun immer ein Delta, sodass für ein vorgegebenes Epsilon die letzte Ungleichung erfüllt ist. Es muss also gelten:  $d(y, x) \leq \delta$ . Hierfür muss lediglich der Abstand zwischen zwei Punkten im Teilsimplex kleiner werden als Delta. Das kann erreicht werden, in dem man die Anzahl  $N$  der Unterteilungen genügend groß wählt, wodurch die Teilsimplizes beliebig klein gemacht werden.

Wenn man nun  $\varepsilon$  gegen null konvergieren lässt, so ergibt sich nach (7) folgende Ungleichung für alle  $i$  und  $j$ :

$$\alpha_i(h_j) \leq \lambda_i(h_j)$$

Somit sind alle Bedingungen für den Fixpunkt erfüllt, ein Fixpunkt existiert also.  $\square$

## 6.2 Erweiterung auf beliebige konvexe Mengen

Das oben vorgestellte Verfahren lässt sich nicht nur für Simplizes anwenden, sondern auch für beliebige konvexe Mengen. Hierzu wird die abgeschlossene, beschränkte konvexe Menge  $M$  im  $n$ -dimensionalen euklidischen Raum einem  $n$ -dimensionalen Simplex  $S$  einbeschrieben.

Nun muss die Funktion  $f$  auf ganz  $S$  erweitert werden – unter Beibehaltung der Stetigkeit. Hierfür bildet man alle Punkte aus  $S$  mit einer neuen Funktion  $g$  auf  $M$  ab. Auf die entstandenen Bildpunkte ist nun  $f$  anwendbar. Die Gesamtfunktion ist also  $\bar{f}(x) := f(g(x))$ .

Jedem Punkt  $x$  aus  $S$  wird solch ein  $x'$  aus  $M$  zugewiesen, sodass der euklidische Abstand zwischen  $x$  und  $x'$  minimal ist. Daraus erkennt man, dass allen Punkten außerhalb von  $M$  ein Randpunkt von  $M$  zugewiesen wird. Man erkennt auch, dass alle Punkte, die selbst schon aus  $M$  sind, wieder auf sich selbst abgebildet werden, da sie schließlich den kleinsten Abstand zu sich selbst haben. Die Eindeutigkeit dieser Zuordnung ergibt sich aus der Konvexität der Menge  $M$ . Denn angenommen es gäbe zwei unterschiedliche Punkte  $x'$  und  $x''$  aus  $M$  mit dem selben minimalen Abstand zu  $x$ , so hätte der Mittelpunkt zwischen den beiden

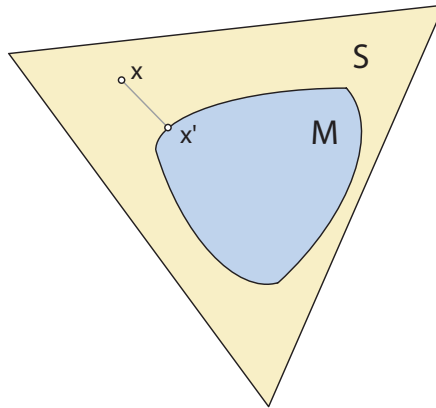


Abbildung 4: Konvexe Menge mit einhüllendem Simplex

(der wegen der Konvexität existiert) aufgrund von geometrischen Aspekten einen kleineren euklidischen Abstand zu  $x$ .

Der Beweis für die Stetigkeit erfolgt indirekt. Angenommen die Projektion  $f(g(x))$  ist unstetig in einem speziellen  $x$ , das auf  $x'$  abgebildet wird, so gibt es eine Folge  $x_n$  die gegen  $x$  konvergiert, deren Funktionswertefolge  $x'_n$  allerdings nicht gegen  $x'$  sondern  $x''$  konvergiert.

Damit gelten für die Abstände mit einer stetigen Abstandsfunktion  $d$  (der euklidische Abstand ist eine solche Funktion) folgende Relationen:

$$d(x, x') < d(x, x'') \quad (8)$$

$$d(x_n, x'_n) \leq d(x_n, x') \quad \forall n. \quad (9)$$

Nach der Stetigkeit von  $d$  folgt aus (9) für  $n \rightarrow \infty$ :

$$d(x, x'') \leq d(x, x'). \quad (10)$$

Dies ist aber ein Widerspruch zu (8), sodass  $x' = x''$  sein muss, also ist  $g$  stetig.

Nun ist der Brouwer'sche Fixpunktsatz auf das einhüllende Simplex mit der stetigen Funktion  $\bar{f}$  anwendbar. Demzufolge existiert ein  $\bar{x} \in S$  mit  $\bar{f}(\bar{x}) = \bar{x}$ . Da aber das Bild von  $\bar{f}$  in  $M$  liegt, muss auch  $\bar{x} \in M$  sein.

Da für alle  $x \in M$  stets  $f(x) = \bar{f}(x)$  gilt, ist  $\bar{x}$  auch ein Fixpunkt von  $f$  in  $M$ .

□



## 7 Sperners Lemma

### 7.1 Konvexe Hülle

Gegeben seien die Punkte  $p_1, \dots, p_m$ . Die Menge der Konvexkombinationen dieser Punkte bezeichnet man als konvexe Hülle

$$\text{conv} \{p_1, \dots, p_m\}.$$

### 7.2 Lemma

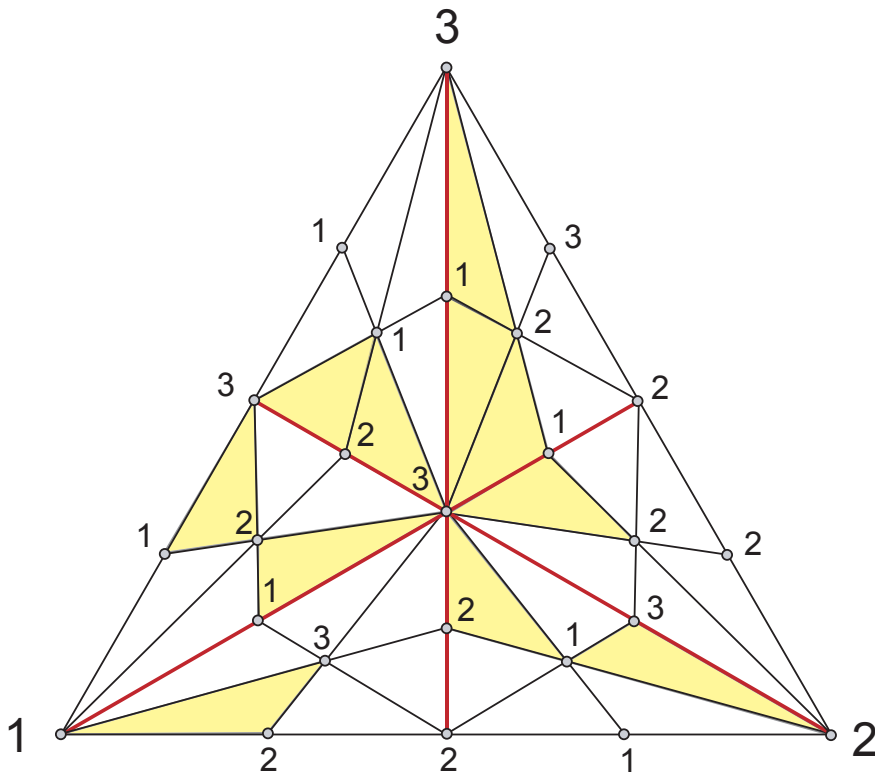


Abbildung 5: Veranschaulichung des Lemmas von Sperner

Ausgangspunkt ist das Verfahren von Brouwer. Jedoch wird ein Simplex mit den Ecken  $p_1, \dots, p_n$  über seinen Schwerpunkt  $r$  in Teilsimplizes unterteilt.

$$r = \frac{1}{n} \sum_{i=1}^n p_i.$$

Dabei wird  $r$  zu einer neuen Ecke und bildet mit jeweils einer anderen Ecke und dem Mittelpunkt der Seite dieser Ecke einen neuen Teilsimplex. Dieses Verfahren

kann nun wieder auf die neu entstandenen Teilsimplizes angewendet werden um eine beliebige ( $N$ -fache) Unterteilung vorzunehmen.

Die Seiten des Ausgangssimplex bilden eine konvexe Hülle, zu der  $x$  gehöre:

$$x \in \text{conv} \{p_1, \dots, p_n\}.$$

Sei nun die Indexmenge als

$$I(x) = \{i | \alpha_i \geq \lambda_i\}$$

definiert und sei

$$\forall x : i(x) \in I(x)$$

eine zulässige Indexfunktion mit dem Zusatz, dass dem Punkte  $p_{k \in \{1, \dots, n\}}$  jeweils  $k$  zugeordnet wird, so gibt es stets eine ungerade Anzahl normaler Teilsimplizes.

Beweis durch vollständige Induktion:

Induktionsanfang:  $n = 1$ :  $p_1$ : trivial

Es wird nun zwischen den folgenden zwei Sichtweisen auf das Simplex unterschieden:

### 7.2.1 Sehen

Man läuft durch alle Teilsimplizes  $T_k$  mit  $k = 1, \dots, N$  und zählt die gesehenen ausgezeichneten Seiten, deren Anzahl  $g(k)$  sei. Die Summe sei

$$s = \sum_{k=1}^N g(k).$$

Nun gibt es zwei Möglichkeiten:

1.  $T_k$  ist normal:  $g(k)$  ist somit 1, weil genau einmal die Eckennummern von 1 bis  $n$  vorkommen.
2.  $T_k$  ist nicht normal:  $g(k)$  ist somit 0 oder 2, weil mindestens eine Eckennummer doppelt vorkommt. Man benötigt nämlich die Nummern  $1, \dots, n-1$  für eine ausgezeichnete Seite. Man hat aber entweder die letzte Nummer (bei  $p_n$ ) doppelt, sodass jeweils mit einer der beiden doppelten Nummern eine ausgezeichnete Seite gebildet werden kann, oder man hat eine andere Nummer doppelt, sodass diese Nummer der ausgezeichneten Seite fehlt.

Man erkennt nun, dass die Summe  $s$  genau dann ungerade ist, wenn die Anzahl der normalen Teilsimplizes ungerade ist.

### 7.2.2 Gesehen werden

Als nächstes betrachtet man die Situation aus der Sicht der ausgezeichneten Seite  $\theta$  und zählt, wie oft diese gesehen wurde. Es ergeben sich wieder zwei Möglichkeiten:

1. Alle Punkte von  $\theta$  liegen in der Original-Simplexseite der Ecken  $p_1$  bis  $p_{n-1}$ :  $\theta$  wird genau einmal von innen aus gesehen. In den Original-Seiten mit den Ecken von  $p_{n-1}$  bis  $p_1$  gibt es keine ausgezeichneten Seiten, da hier eine Indexnummer fehlt.
2.  $\theta$  liegt nicht im Rand:  $\theta$  wird von den zwei benachbarten Teilsimplizes gesehen.

Somit ist  $s = a + 2b$ , wobei  $a$  die Anzahl der ausgezeichneten Seiten im Fall 1 ist und  $b$  im Fall 2. Man erkennt, dass  $s$  genau dann gerade ist, wenn auch  $a$  gerade ist.  $a$  ist aber auch gleichzeitig die Anzahl normaler Teilsimplizes in einem Simplex der Dimension  $n - 2$  und somit auch ungerade. Damit ist auch  $s$  ungerade.  $\square$

## 8 Minimax-Satz

Sei  $f(x, y) \in \mathbb{R}$  eine stetige Funktion, die in  $y$  konkav und in  $x$  konvex ist. Es gelte

$$\begin{aligned}x &\in X \subset \mathbb{R}^n, \\y &\in Y \subset \mathbb{R}^m \text{ und} \\n, m &\in \mathbb{N}^*,\end{aligned}$$

wobei  $X$  und  $Y$  konvexe, abgeschlossene und beschränkte Mengen seien.

Für den Sattelpunkt  $(x^*, y^*)$  gelte somit:

$$\forall x \in X, y \in Y : f(x, y^*) \leq f(x^*, y^*) \leq f(x^*, y)$$

Unter diesen Voraussetzungen existiert stets wenigstens ein Sattelpunkt.

### 8.1 Beispiel

Es werde Schere-Stein-Papier gespielt. Dabei sei  $x = (x_1, x_2, x_3)$  eine gemischte Strategie. Sei  $a_{ij}$  der Gewinn von Spieler 1 bei der Originalstrategie  $i$  (Stein) gegen die Originalstrategie  $j$  (Papier) des Gegners.

Bei  $x$  gegen die Strategie  $j$  beträgt der Erwartungswert des Gewinns

$$\sum_i x_i a_{ij}.$$

Insgesamt ergibt sich somit für  $x$  gegen  $y$

$$\sum_j \sum_i x_i a_{ij} y_j = x^T A y = f(x, y),$$

wobei  $A$  eine Matrix mit den Gewinneinträgen  $a_{ij}$  ist. Aus der Linearität in  $x$  und  $y$  folgt, dass  $f$  konvex-konkav ist.

## 8.2 Konvexität

Die Funktion  $g$  nennt man *konvex*, falls gilt:

$$\forall x, x' \in X \forall \lambda \in ]0, 1[: g(\lambda x + (1 - \lambda)x') \leq \lambda g(x) + (1 - \lambda)g(x')$$

## 8.3 Beweis des Minimax-Satzes

Sei  $f$  eine strikt konvex-konkave Funktion. Man betrachte nun irgendein Paar  $(x, y) \in M$ . Es seien

$$X^*(y) = \{\xi \in X | f(\xi, y) = \max_{x' \in X} f(x', y)\}$$

$$Y^*(x) = \{\xi \in Y | f(x, \xi) = \min_{y' \in Y} f(x, y')\}.$$

Als  $F$  nehme man nun  $F(x, y) = X^*(y) \times Y^*(x)$ , da in einer konvexen Menge das Maximum/Minimum genau einen Punkt beschreibt, falls  $f$  strikt konkav/konvex ist. Dieses  $F$  ist stetig, was später noch bewiesen wird.

Nun sei  $(x^*, y^*)$  der Fixpunkt.

Es gilt somit, dass  $x^*$  die beste Antwort auf  $y^*$  ist

$$\iff f(x^*, y^*) \geq f(x, y^*) \forall x \in X$$

und dass  $y^*$  die beste Antwort auf  $x^*$  ist

$$\iff f(x^*, y^*) \leq f(x^*, y) \forall y \in Y.$$

Daraus folgt die Behauptung.

## 8.4 Beweis der Stetigkeit von $X^*$ und $Y^*$

Sei  $(y_k)$  eine Folge, die gegen  $y$  konvergiert. Man betrachte nun das beste  $x_k$  zu  $y_k$ . Es konvergiere  $(x_k)$  gegen  $\bar{x}$ . Nun ist zu zeigen, dass  $\bar{x}$  optimal für  $y$  ist. Wäre dies nicht der Fall, dann würde gelten:

$$\exists \xi \in X : f(\xi, y) < f(\bar{x}, y) - \varepsilon, \quad \varepsilon > 0.$$

Daraus folgt, dass  $x_k$  nicht die beste Antwort zu  $y_k$  ist, da hier  $\xi$  besser ist und somit muss  $\bar{x}$  optimal sein.  $\square$

## 8.5 Allgemeiner Fall ohne Striktheit

$f$  sei geändert:

$$f_\varepsilon(x, y) = f(x, y) - \varepsilon \|x\|_2 + \varepsilon \|y\|_2$$
$$\|x\|_p = \sum_{i=1}^n x_i^p.$$

Diese neue Funktion ist strikt konvex-konkav und man kann den Minimax-Satz auf  $f_\varepsilon$  anwenden. Sei  $(x_\varepsilon, y_\varepsilon)$  der Sattelpunkt von  $f_\varepsilon$ . Für  $\varepsilon_k \rightarrow 0$  mit  $k \rightarrow \infty$  kann man nun die Folgen  $(x_{\varepsilon_k})$  und  $(y_{\varepsilon_k})$  betrachten. Nach Auswahl einer geeigneten Teilfolge konvergieren diese gegen einen Sattelpunkt von  $f$ . Der Beweis erfolgt analog zum Beweis der Stetigkeit von  $X^*$  und  $Y^*$ .

## 9 Retrakte

Sei  $B$  eine Einheitskugel. Eine Teilmenge  $C \subset B$  ist ein Retrakt, falls es eine stetige Abbildung  $f : B \rightarrow C$  gibt mit

$$\forall x \in C : f(x) = x. \quad (11)$$

### 9.1 Satz

Der Rand  $C = \partial B$  ist nie Retrakt von  $B$ .

### 9.2 Äquivalenz

Im Folgenden soll gezeigt werden, dass dieser Satz äquivalent zu Brouwers Fixpunktsatz ist. Dazu soll zum einen der Retraktsatz aus dem Brouwer-Satz geschlossen werden und umgekehrt.

### 9.3 Beweis über den Brouwer'schen Fixpunktsatz

Angenommen  $\partial B$  ist Retrakt von  $B$ . Es sei nun

$$g(x) = -f(x).$$

mit  $g : B \rightarrow B$ , wobei  $f$  die Retrakt-Abbildung ist.  $g$  besitzt den Fixpunkt  $x^* \in \partial B$ :

$$x^* = g(x^*) = -f(x^*).$$

Nach (11) gilt aber

$$x^* = f(x^*).$$

Wegen diesem Widerspruch kann der Rand der Einheitskugel nie deren Retrakt sein.  $\square$

### 9.4 Beweis von Brouwers Fixpunktsatz über den Retrakt-satz

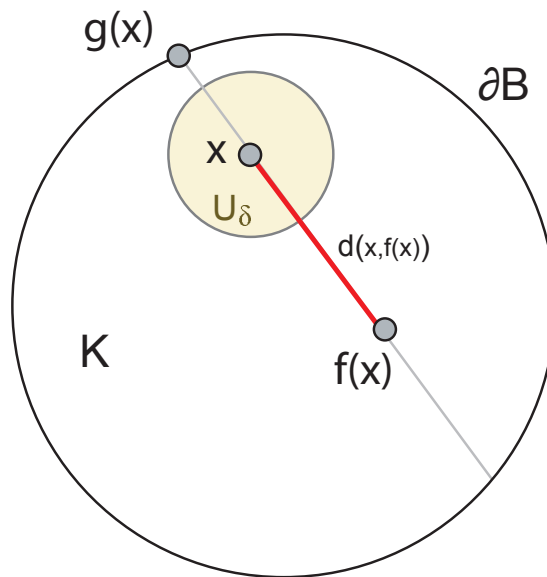


Abbildung 6: Einheitskreis

Angenommen Brouwers Satz sei falsch. Dann gibt es eine Kugel (analog zum Simplex), sodass mit einer geeigneten stetigen Abbildung  $f : K \rightarrow K$  der Fixpunktsatz nicht gilt.

O.B.d.A sei  $K$  die Einheitskugel  $B$ , dann gilt:

$$\forall x \in B : f(x) \neq x \tag{12}$$

$$\Rightarrow \inf d(x, f(x)) > 0 \tag{13}$$

Die Folgerung (13) muss gelten, denn sonst würde eine Folge  $(x_k)$  mit  $x_k \in B$  existieren, für die gilt:

$$\lim_{k \rightarrow \infty} d(x_k, f(x_k)) = 0.$$

Nun lässt sich eine geeignete Teilfolge  $(x_{k_t})$  auswählen, deren Grenzwert  $x^*$  sei. Wegen der Stetigkeit von  $f$  gilt nun

$$\lim_{t \rightarrow \infty} f(x_{k_t}) = f(x^*) \Rightarrow d(x^*, f(x^*)) = 0. \tag{14}$$

Aus (12) folgt somit

$$d(x, f(x)) \geq \delta > 0.$$

Nun definiert man  $g(x)$  als den Schnittpunkt der Geraden durch  $x$  und  $f(x)$  mit dem Rand von  $K$ . Dabei soll  $x$  zwischen  $g(x)$  und  $f(x)$  liegen. Diese neue Abbildung ist stetig (wegen der Stetigkeit von  $f$ ) und ein Retrakt, denn es gilt  $g(x) \in \partial B$  und mit  $x \in \partial B$  folgt per Konstruktion

$$g(x) = x.$$

Die Existenz dieses Fixpunktes steht im Widerspruch zu der Annahme (12) und somit war die Annahme falsch.  $\square$