Leopold-Franzens Universität Innsbruck

Bachelorarbeit

am Institut für Mathematik Innsbruck

Von Magischen Quadraten zu Magischen Würfeln

Sabina Malik

betreut von

Univ.-Prof. Dr. Tim Netzer

Abgabedatum: 10. Februar 2022
Matr.-Nr.: 11806741
Studienkennzahl: UC 033 201

universität
innsbruck

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Ich erkläre mich mit der Archivierung der vorliegenden Bachelorarbeit einverstanden.

11.02.22	Onalik
Datum	Unterschrift

Inhaltsverzeichnis

Einleitung

Zahlenrätsel begeistern MathematikerInnen ebenso wie Nicht-MathematikerInnen seit Anbeginn der Menschheit. Eines dieser Zahlenspiele ist das der magischen Quadrate. Die Aufgabe dabei ist es, ein $n \times n$ Feld mit den Zahlen von 1 bis n^2 zu füllen, sodass die Einträge in jeder Zeile und Spalte sowie der Diagonalen sich zur selben Zahl summieren. Seien es Entdeckungen des Autodidakten Srinivasa Ramanujan oder Kunstelemente in Kupferstichen von Albrecht Dürer (Abb. 1), berühmte Beispiele solcher Matrizen gibt es viele.

Im Folgenden passen wir diese Eigenschaften etwas an. Es muss nur erfüllt sein, dass sich Zeilen- und Spaltensummen gleichen, wobei diese jeweils 1 ergeben sollen. In dieser Arbeit ist der Begriff magischer Quadrate also gleichzusetzen mit dem der doppelt-stochastischen Matrizen. Ein Beispiel dafür ist die Einheitsmatrix und Permutationen dieser. Eine Eigenschaft dieser Matrizen erkennt man bereits an der Definition der Begriffe: Jede Konvexkombination aus Permutationsmatrizen ist doppelt-stochastisch. Die nicht-triviale Umkehrung des Satzes, dass jede doppelt-stochastische Matrix eine Konvexkombination aus Permutationsmatrizen ist, wurde von Birkhoff und von Neumann unabhängig voneinander bewiesen und stellt eine der wichtigsten Entdeckungen im Zusammenhang mit magischen Quadraten dar. Dieses Theorem wurde schon oft für allgemeinere Fälle als dem reellen Fall betrachtet, ein Beispiel hierfür ist [1], in welchem Einträge über $C^* - Algebren$ betrachtet wurden und der Satz für $n \geq 3$ widerlegt wurde. Auch in dieser Arbeit liegt der Fokus auf einer Verallgemeinerung des Satzes von Birkhoff und von Neumann: lässt sich dieser für stochastische Tensoren formulieren? Dabei legen wir einen Fokus auf $n \times n \times n$ Tensoren und weiten die bekannten Definitionen darauf aus. Können wir von den Eigenschaften magischer Quadrate auf magische Würfel schließen?

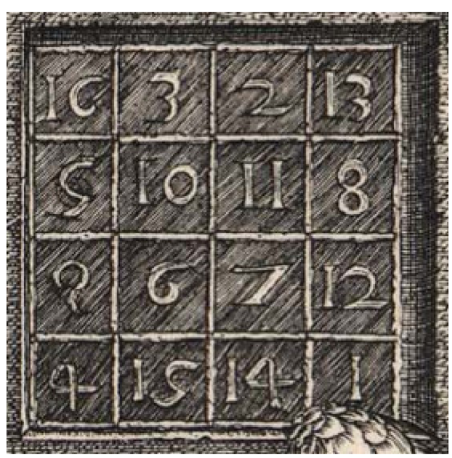


Abbildung 1: Magisches Quadrat in Albrecht Dürers „Melencolia“

1 Magische Quadrate

Bevor wir beginnen, uns mit magischen Würfeln auseinanderzusetzen, ist es zuerst wichtig, die Eigenschaften der einzelnen Schichten dieses Konstrukts zu verstehen. Deshalb wird als Einstieg in das Thema der zweidimensionale Fall betrachtet: Magische Quadrate. Im Folgenden sind die Einträge der Matrizen immer reelle Zahlen.

1.1 Wichtige Definitionen

Eine der bedeutendsten Eigenschaften von magischen Quadraten wurde von G. Birkhoff 1946 und J. von Neumann 1953 unabhängig voneinander entdeckt und im folgenden Satz festgehalten. [2]

Satz 1.1 (Birkhoff- von Neumann)**.** *Eine $n \times n$ Matrix ist genau dann doppelt-stochastisch, wenn sie eine Konvexkombination aus endlich vielen $n \times n$ Permutationsmatrizen ist.*

Um auf den Beweis des Satzes einzugehen, werden zuerst die für die Formulierung nötigen Definitionen erklärt.

Definition 1.1. Sei $n \in \mathbb{N}$ und $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ mit $a_{ij} \geq 0$

(i) A heißt *doppelt-stochastische Matrix*, wenn sich die Einträge jeder Zeile und jeder Spalte zu 1 summieren. Es gilt

$$\sum_{i=1}^{n} a_{ij} = 1 \qquad \forall 1 \leq j \leq n,$$
$$\sum_{j=1}^{n} a_{ij} = 1 \qquad \forall 1 \leq i \leq n.$$

(ii) Ist in jeder Zeile und Spalte genau ein Eintrag 1 und die restlichen 0, heißt A *Permutationsmatrix*. Jede Permutationsmatrix ist folglich eine doppelt-stochastische Matrix.

Beispiel 1.1.

(i)

$$\begin{pmatrix} \frac{1}{6} & \frac{5}{6} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{5}{6} & 0 & 0 & \frac{1}{6} \end{pmatrix}$$

(ii) Laut dem Satz von Birkhoff und von Neumann ist (i) eine Konvexkombination von Permutationsmatrizen

$$\begin{pmatrix} \frac{1}{6} & \frac{5}{6} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{5}{6} & 0 & 0 & \frac{1}{6} \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{5}{6} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Anhand der Definitionen erkennt man direkt, dass jede Konvexkombination aus Permutationsmatrizen eine doppelt-stochastische Matrix ist, wobei wir diese Tatsache noch genauer zeigen werden. Außerdem ergibt sich, dass jede ganzzahlige doppelt-stochastische Matrix eine Permutationsmatrix ist. Um die im Satz beschriebene Umkehrung zu beweisen, werden zusätzlich folgende Begriffe benötigt:

Definition 1.2. Sei V ein Vektorraum und $K \subseteq V$ eine konvexe Menge, dann heißt $x \in K$ *Extremalpunkt* von K, wenn für $a, b \in K$, $0 < \lambda < 1$ gilt

$$x = \lambda a + (1 - \lambda) b \quad \Rightarrow \quad x = a = b,$$

x lässt sich also nicht als Konvexkombination zweier verschiedener Punkte aus K darstellen.

Definition 1.3. Das konvexe Polytop

$$B_n = \{A = (a_{ij})_{1 \leq i,j \leq n} | A \text{ ist doppelt stochastisch}\} \subseteq \mathbb{R}^{n^2}$$

heißt *Birkhoff-Polytop*, benannt nach G. Birkhoff.

Die Punkte des Birkhoff-Polytops können als Elemente des $\mathbb{R}^{n^2}$ aufgefasst werden, da die Spaltenvektoren der $n \times n$ Matrizen aneinandergereiht werden können und wir somit einen Vektor mit n^2 Einträgen erhalten. Für besonderes Interesse sind für uns im Folgenden die Ecken dieser Menge, da die Ecken eines Polytops dessen Extremalpunkte sind.

1.2 Satz von Birkhoff und von Neumann

Zusammen mit den Definitionen des vorherigen Abschnitts lässt sich der Satz von Birkhoff und von Neumann umformulieren, die Aussage bleibt dabei erhalten.

Satz 1.2 (Birkhoff- von Neumann). *Die Extremalpunkte der Menge der doppelt-stochastischen Matrizen (Birkhoff Polytop B_n) sind genau die Permutationsmatrizen.*

Es gilt also zu zeigen, dass die Ecken des Birkhoff-Polytops genau die Permutationsmatrizen sind. Der Satz wurde auf viele verschiedene Arten bewiesen, im Folgenden werden wir uns an [2] und [3] orientieren, welche an der zweiten Formulierung des Satzes ansetzen. Hierfür benötigen wir einen Satz aus der linearen Optimierung, welcher uns hilft, die Ecken eines Polyeders zu charakterisieren.

Satz 1.3. *Sei $P \subset \mathbb{R}^d$ ein Polyeder*

$$P = \{u \in \mathbb{R}^d | \langle c_i, u\rangle \leq b_i \text{ für } i = 1, \ldots, n\} \quad c_i \in \mathbb{R}^d,\ b_i \in \mathbb{R}.$$

Es gilt $x \in P$ ist genau dann eine Ecke von P, wenn für die Indexmenge

$$I(x) = \{i \in \{1, \ldots, n\} | \langle c_i, x\rangle = b_i\}$$

die Menge der Vektoren $\{c_i | i \in I(x)\}$ den Vektorraum $\mathbb{R}^d$ linear aufspannen. In diesem Fall enthält die Menge $I(x)$ mindestens d Elemente.

Beweis. „$\Rightarrow$ " Wir nehmen an, dass $\operatorname{span}(\{c_i | i \in I(x)\}) \neq \mathbb{R}^d$ gilt. Wir können jetzt ein nicht-triviales $y \in \mathbb{R}^d$ finden, sodass $\langle y, c_i\rangle = 0$ für alle $i \in I(x)$ und setzen $x^+ = x + \epsilon y$ und $x^- = x - \epsilon y$ für ein $\epsilon > 0$. Dann gilt $x = \frac{x^+ + x^-}{2}$ und $x^+ \neq x^-$. Für $\epsilon > 0$ klein genug liegen x^+ und x^- im Polyeder P. Also haben wir gezeigt, dass x eine Konvexkombination zweier verschiedener Elemente in P ist und somit kein Extremalpunkt sein kann.

„$\Leftarrow$ " Sei nun $x \in P$ und die Vektoren c_i mit $i \in I(x)$ spannen den $\mathbb{R}^d$ auf. Wir nehmen an, dass x keine Ecke von P ist, also

$$x = \frac{y + z}{2} \quad \text{für } y, z \in P. \tag{1}$$

Es gilt

$$\langle c_i, y\rangle \leq b_i \text{ und } \langle c_i, z\rangle \leq b_i \text{ für } i = 1, \ldots n, \tag{2}$$
$$\langle c_i, x\rangle = b_i \text{ für } i \in I(x). \tag{3}$$

Wegen (1), (2) und (3) muss gelten, dass $\langle c_i, y\rangle = \langle c_i, z\rangle = b_i$ für $i \in I(x)$. Weiters wissen wir, dass das System linearer Gleichungen $\langle c_i, u\rangle = b_i$ mit $i \in I(x)$ eine eindeutige Lösung hat, also muss gelten $x = y = z$. Damit haben wir einen Widerspruch zur Annahme. □

Da dieser Satz für Polyeder gilt, können wir ihn insbesondere auf das Birkhoff-Polytop anwenden und erhalten zusätzliche Aussagen zu dessen Ecken, welche bei dieser Variante des Beweises des Birkhoff und von Neumanns Theorems benötigt werden.

Beweis Birkhoff - von Neumann. Die Menge B_n setzt sich aus Punkten $(a_{ij}) \in \mathbb{R}^{n^2}$ zusammen, welche folgende Bedingungen erfüllen müssen:

$$a_{ij} \geq 0 \quad \text{für } i, j = 1, \ldots, n \tag{4}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \quad \text{für } i = 1, \ldots, n \tag{5}$$

$$\sum_{i=1}^{n} a_{ij} = 1 \quad \text{für } j = 1, \ldots, n \tag{6}$$

Daraus erhalten wir, dass für alle $(a_{ij}) \in B_n$ gilt $0 \leq a_{ij} \leq 1$. Also ist B_n ein beschränkter konvexer Polyeder und somit ein Polytop. Wir betrachten nun beide Richtungen des Satzes, wobei die Rückrichtung, wie bereits erwähnt, einfach zu zeigen ist und deshalb bei der Ausführung des Beweises meist weggelassen wird.

„$\Leftarrow$" Sei $P = (p_{ij})$ eine $n \times n$ Permutationsmatrix. Angenommen P ist eine echte Konvexkombination aus $A = (a_{ij})$, $B = (b_{ij}) \in B_n$, also es gilt

$$P = \lambda A + (1 - \lambda) B \quad \text{für } \lambda \in (0, 1).$$

Falls $p_{ij} = 0$, muss wegen der Nichtnegativität der Einträge (4) gelten $a_{ij} = b_{ij} = 0$. Falls $p_{ij} = 1 = \lambda a_{ij} + (1 - \lambda) b_{ij}$ sind die restlichen Einträge der Zeile/Spalte gleich 0 und wegen (5) bzw. (6) muss gelten $a_{ij} = b_{ij} = 1$. Also gilt $P = A = B$, damit ist P keine Konvexkombination zweier verschiedener magischer Quadrate, ein Widerspruch zur Annahme.

„$\Rightarrow$" Für diese Richtung gehen wir über Induktion über n vor mit der Annahme, dass aus $A = (a_{ij})$ Extremalpunkt des Birkhoff-Polytops B_n folgt, dass A eine Permutationsmatrix sein muss. Der Induktionsanfang ist trivial, da es für $n = 1$ nur ein Element in B_1 gibt, $A = (1)$, welches dadurch offensichtlich Extremalpunkt ist und die Bedingung für Permutationsmatrizen erfüllt.
Wir nehmen $n \geq 2$ an und betrachten eine affine Teilmenge $N \subset \mathbb{R}^{n^2}$ bestehend aus den $n \times n$ Matrizen, welche (5) und (6) erfüllen. Eine solche Matrix A ist bereits durch die $(n-1)^2$ Einträge a_{ij} für $i, j = 1, \ldots n - 1$ eindeutig gegeben, da die restlichen durch diese bereits definiert sind:

$$a_{in} = 1 - \sum_{j=1}^{n-1} a_{ij} \quad \text{für } i = 1, \ldots, n - 1,$$

$$a_{nj} = 1 - \sum_{i=1}^{n-1} a_{ij} \quad \text{für } j = 1, \ldots, n - 1,$$

$$a_{nn} = (2 - n) + \sum_{i,j=1}^{n-1} a_{ij}.$$

Also gilt $\dim N = (n-1)^2$. In N können wir B_n durch die n^2 Ungleichungen $a_{ij} \geq 0$ definieren. Für einen Extremalpunkt gilt laut Satz 1.3, dass $(n-1)^2$ dieser Ungleichungen aktiv sein müssen, also $a_{ij} = 0$ für $(n-1)^2$ Einträge von A gilt.
Die $n \times n$ Matrizen in N können weder Nullzeilen noch -spalten haben, da sich die jeweiligen Einträge zu 1 summieren müssen, es muss also mindestens ein Eintrag positiv sein. Wir nehmen an, die Anzahl an positiven Einträgen in jeder Zeile ist mindestens 2. Damit bleiben nur noch maximal $n(n-2) < (n-1)^2$ Einträge gleich 0 übrig, ein Widerspruch zu Satz 1.3. Es muss also mindestens eine Zeile i_0 geben, sodass $a_{i_0 j} = 0$ für alle bis auf ein $j = j_0$. Um in N zu liegen, muss gelten $a_{i_0 j_0} = 1$. Streichen wir die i_0 Zeile und j_0 Spalte, erhalten wir eine $(n-1) \times (n-1)$ Untermatrix, welche Extremalpunkt von B_{n-1} ist. Deshalb können wir die Induktionsannahme anwenden und erhalten somit eine Permutationsmatrix. Wir erhalten also auch zusammen mit der i_0 Zeile und j_0 Spalte eine $n \times n$ Permutationsmatrix. Damit ist die Annahme gezeigt. $\square$

Wir haben gezeigt, dass die Extremalpunkte des Birkhoff-Polytops die Bedingungen an Permutationsmatrizen erfüllen. Da der Satz von Krein-Milman besagt, dass jedes Polytop die konvexe Hülle seiner Ecken ist, folgt daraus auch die ursprüngliche Formulierung des Satzes von Birkhoff - von Neumann. Ein Beweis für den Satz von Krein-Milman befindet sich ebenfalls in [2].

In [6] wird erwähnt, dass eine weitere beliebte Beweismethode folgendes Lemma über die positiven Diagonalen doppelt-stochastischer Matrizen verwendet. Als Diagonale bezeichnen wir in diesem Fall beliebige n Einträge einer $n \times n$ Matrix, welche alle auf verschiedenen Zeilen und Spalten liegen, da wir diese mithilfe von Permutationen der Spalten als Hauptdiagonale anordnen können. Da im selben Paper interessante Entdeckungen zu den Diagonalen magischer Würfel gemacht werden, schließen wir den Abschnitt zu magischen Quadraten mit diesem Lemma ab.

Lemma 1.1. *Sei $A = (a_{ij})_{i,j} \in B_n$ eine doppelt-stochastische Matrix. Dann lässt sich eine Permutation σ von $\{1, \dots, n\}$ finden, sodass $a_{1\sigma(1)}, \dots, a_{n\sigma(n)} > 0$. In anderen Worten: jede doppelt-stochastische Matrix hat eine positive Diagonale.*

Beweis. Siehe [5], Lemma 8.7.1 S.548. $\square$

Sei Δ_n die Menge der doppelt-stochastischen $n \times n$ Matrizen mit positiver Diagonale und Ω_n die konvexe Hülle der $n \times n$ Permutationsmatrizen. Jede Permutationsmatrix hat offensichtlich eine positive Diagonale. Es gilt also $\Omega_n \subseteq \Delta_n \subseteq B_n$ und laut Birkhoff - von Neumann besteht Gleichheit der Mengen. Doch gilt dies auch für stochastische Tensoren in höheren Dimensionen?

2 Magische Würfel

In diesem Kapitel gehen wir der Frage nach multidimensionalen stochastischen Tensoren und deren Eigenschaften nach. Wir wollen unsere Matrizen auf drei Indizes verallgemeinern und betrachten, wie sich diese Veränderung auf den Satz von Birkhoff und von Neumann auswirkt. Diese $n \times n \times n$ Tensoren haben eine Dimension mehr als die bereits betrachteten Quadrate, man kann sie sich also als Würfel vorstellen.

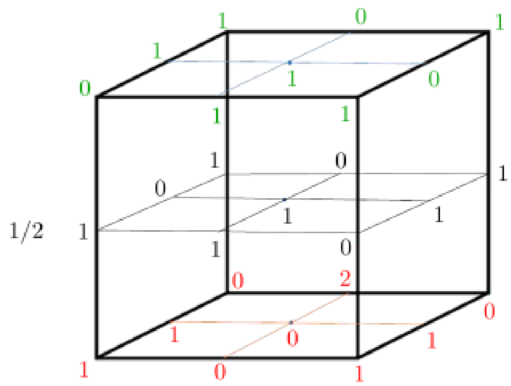


Abbildung 2: Beispiel eines Magischen Würfels, Quelle: [4] S.389

2.1 Verallgemeinerung der Definitionen

Um der Hauptfrage dieser Arbeit nachzugehen, müssen wir zuerst den Begriff doppelt-stochastischer Matrizen verallgemeinern, sodass die gewünschten Bedingungen immer noch erfüllt sind.
Zuerst gilt zu definieren, welche Einträge des Würfels sich zu 1 summieren sollen und auf welche Art diese angeordnet sind. Dafür führen wir, wie in [7], den Begriff der Hypermatrizen für beliebige Dimension m ein, fokussieren uns aber weiterhin auf $m = 3$. Weiters betrachten wir deren Linien, Diagonalen und Schichten, wie sie in [4] verwendet werden. Gleich wie bei den Quadraten verwenden wir auch hier die Begriffe stochastischer Würfel und magischer Würfel synonym.

Definition 2.1. Seien $n, m \in \mathbb{N}$

(i) Wir nennen die Abbildung $A : \underbrace{\{1, \dots, n\} \times \dots \times \{1, \dots, n\}}_{m-mal} \to \mathbb{R}$ *Hypermatrix* der Ordnung n und Dimension m und schreiben $A = (a_{i_1,\dots,i_m})$ mit $1 \leq i_1, \dots, i_m \leq n$.

(ii) Sei $A = (a_{ijk})$ eine Hypermatrix der Dimension 3 mit $a_{ijk} \geq 0$ und

$$\sum_{i=1}^{n} a_{ijk} = 1, \quad \forall j, k$$
$$\sum_{j=1}^{n} a_{ijk} = 1, \quad \forall i, k$$
$$\sum_{k=1}^{n} a_{ijk} = 1, \quad \forall i, j$$

Dann heißt A *stochastischer Würfel.*

Beispiel 2.1. Die folgenden Beispiele sind jeweils $3 \times 3 \times 3$ stochastische Würfel in Matrixschreibweise. In Abbildung 2 wird (i) anschaulich als Würfel dargestellt.

(i)

$$\frac{1}{2}\left(\begin{array}{ccc:ccc:ccc} 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 2 & 0 \end{array}\right)$$

(ii)

$$\left(\begin{array}{ccc:ccc:ccc} 0 & \frac{3}{5} & \frac{2}{5} & 1 & 0 & 0 & 0 & \frac{2}{5} & \frac{3}{5} \\ \frac{3}{5} & 0 & \frac{2}{5} & 0 & \frac{2}{5} & \frac{3}{5} & \frac{2}{5} & \frac{3}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{3}{5} & \frac{2}{5} & \frac{3}{5} & 0 & \frac{2}{5} \end{array}\right)$$

Definition 2.2. Sei $A = (a_{ijk})$ eine dreidimensionale Hypermatrix.

(i) Als eine *Schicht von A* wird der zweidimensionale Teil von A bezeichnet, welchen man durch Fixieren eines der drei Indizes erhält.

(ii) Der Schnitt zweier nicht-paralleler Schichten von A wird als *Linie* bezeichnet. Man erhält eine Linie auch durch Fixieren aller bis auf einen Index.

(iii) Eine *Diagonale von A* setzt sich aus n^2 Elementen zusammen, sodass für alle diese Elemente gilt, dass alle auf verschiedenen Linien liegen.

Eine Hypermatrix der Dimension 3 besteht aus insgesamt 9 verschiedenen Schichten, da wir 3 Möglichkeiten haben den Index zu fixieren und sich bei jeder Fixierung 3 Schichten ergeben. Im Falle eines stochastischen Würfels ist jede dieser Schichten wiederum eine doppelt-stochastische Matrix. In Beispiel 2.1, allen folgenden Beispielen und sonstigen Verwendungen dieses Begriffs setzen wir o.B.d.A. eine Fixierung des Index k des Würfels voraus und bezeichnen die Schichten als A_k für $1 \leq k \leq n$.

Definition 2.3. Sei $A = (a_{ijk})$ ein stochastischer Würfel. Gibt es $1 \leq k_1 \neq k_2 \leq n$, sodass der (i,j)-Eintrag der Schicht A_{k_1} sowohl als auch der Schicht A_{k_2} positiv ist, so heißen die beiden Schichten *untereinander zusammenhängend*. Sind alle Schichten A_k, $(k = 1, \ldots, n)$, untereinander zusammenhängend, so heißt A *zusammenhängender Tensor*.

Auch Permutationsmatrizen müssen umdefiniert werden, wobei diese im $n \times n$ Fall ja Permutationen der Einheitsmatrix I_n sind, was sich auf höhere Dimensionen nicht erweitern lässt. Wir können zwar die mögliche Wahl der Einträge so restringieren, dass die einzelnen Schichten Permutationsmatrizen im üblichen Sinn sind, finden jedoch keine Hauptdiagonale und haben somit keinen Einheitstensor definiert. Um die Verallgemeinerung von Quadraten auf Würfel zu verdeutlichen, werden in dieser Arbeit trotz fehlendem „Einheitswürfel“ die Elemente der im Folgenden definierten Menge als Permutationswürfel bezeichnet. Um die Menge der stochastischen Würfel, welche genau einen Eintrag gleich 1 auf jeder Linie haben, zu definieren, wird in [6] eine Verbindung zu lateinischen Quadraten hergestellt.

Definition 2.4. Sei $n \in \mathbb{N}$

(i) Eine $n \times n$ Matrix $L = (l_{ij})$ heißt *lateinisches Quadrat der Ordnung* n, wenn die Zahlen $1, \ldots, n$ jeweils genau einmal in jeder Zeile und Spalte vorkommen.

(ii) Eine dreidimensionale Hypermatrix $A = (a_{ijk})$ heißt *Permutationswürfel*, wenn sie ein stochastischer Würfel ist und auf jeder Linie genau ein Eintrag 1 entspricht. Jeder Permutationswürfel hat ein korrespondierendes lateinisches Quadrat, sodass gilt

$$a_{ijk} = 1 \Leftrightarrow l_{ij} = k.$$

Das bekannte Zahlenrätsel Sudoku ist üblicherweise ein lateinisches Quadrat der Ordnung 9. Auch hier summieren sich die Einträge der Zeilen und Spalten jeweils zur selben Zahl. Mit entsprechender Normierung erhalten wir also doppelt-stochastische Matrizen. Der Begriff der lateinischen Quadrate wird uns erneut bei der Frage nach der Anzahl an Ecken der Menge magischer Würfel unterkommen.

Laut ihrer Definition sind Permutationswürfel für $n \geq 2$ keine zusammenhängende Tensoren. Gilt für einen Eintrag in einer Schicht $(A_k)_{ij} = 1$, so muss für jede weitere Schicht $(A_{k_0})_{ij} = 0$ gelten, wobei $k_0 \neq k$, damit sich auf einer Linie nicht mehr als eine 1 befindet. Somit können keine der Schichten zusammenhängen.

Durch die Permutationswürfel lässt sich auch die Diagonale eines stochastischen Würfels $A = (a_{ijk})$ genauer charakterisieren. Nämlich ergibt sich die Menge dieser n^2 Einträge genau durch

$$D_P(A) = \{a_{ijk} | p_{ijk} = 1\}$$

für einen Permutationstensor $P = (p_{ijk})$. Damit lässt sich die Permanente einer Hypermatrix definieren. Wir bezeichnen mit Ω_n^3 die Menge aller Permutationswürfel.

Definition 2.5. Die Summe der Produkte der Einträge der Diagonalen einer Hypermatrix $A = (a_{ijk})$,

$$\operatorname{per}(A) := \sum_{P \in \Omega_n^3} \prod_{a_{ijk} \in D_P(A)} a_{ijk},$$

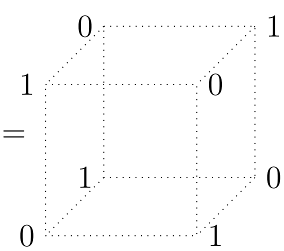
heißt *Permanente* des Tensors.

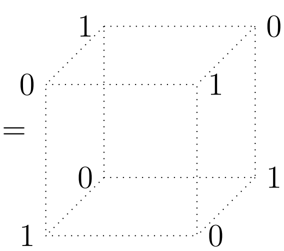
Für eine doppelt-stochastische Matrix A gilt laut Lemma 1.1 stets $\operatorname{per}(A) > 0$. Die Eigenschaft einer positiven Diagonale gilt jedoch nicht immer für stochastische Würfel, wie wir im nächsten Abschnitt feststellen werden.

2.2 Umformulierung Birkhoff und von Neumann

Um uns einen ersten Überblick zu verschaffen, wie sich die Menge der Extremalpunkte des Birkhoff-Polytops in Dimension 3 zusammensetzen könnte, wollen wir uns anhand von Beispielen orientieren. Dafür bezeichnen wir mit B_n^3 die Menge der $n \times n \times n$ stochastischen Würfel. Wir erhalten hier wieder eine konvexe und kompakte Menge, da sie sich als Schnitt endlich vieler abgeschlossener Halbräume ergibt und beschränkt ist. Durch die Beispiele fallen bereits erste Bedingungen auf, welche ausreichend sind, um die Ecken dieser Menge zu charakterisieren und welche wir später in diesem Abschnitt festhalten wollen. Dabei sind nicht die Eigenschaften von Permutationswürfeln vorausgesetzt, der Satz von Birkhoff- von Neumann lässt sich für magische Würfel also nicht allgemein umformulieren.

Betrachten wir den Fall $n = 2$, lässt sich die Gültigkeit von Birkhoff- von Neumann leicht zeigen. Es gibt nur die folgenden zwei Permutationswürfel in dieser Ordnung:

$P =$ 

$Q =$ 

Sei $0 \leq a \leq 1$. Alle weiteren stochastischen Würfel in dieser Menge haben diese Form:

$$A = \begin{pmatrix} a & 1-a & \vdots & 1-a & a \\ 1-a & a & \vdots & a & 1-a \end{pmatrix}$$

Man erkennt direkt, dass $A = aP + (1-a)Q$ gilt. Wir betrachten im Folgenden also die Fälle für $n \geq 3$.

2.2.1 Gegenbeispiele

Wir bezeichnen die konvexe Hülle der Permutationswürfel als Ω_n^3 und die Menge der stochastischen Würfel mit $\mathrm{per}(A) > 0$ als Δ_n^3. Für den Fall $n = 2$ lässt sich Lemma 1.1 auf den Würfel verallgemeinern, wie man an den obigen Überlegungen sieht. Somit gilt $\Omega_2^3 = \Delta_2^3 = B_2^3$. Für $n = 3$ haben wir mit Beispiel 2.1 (i) bereits ein Beispiel einer Hypermatrix ohne positiver Diagonale gefunden. Man findet keine n^2 Einträge in diesem Würfel, sodass keine von ihnen auf einer gleichen Linie liegen und keine von ihnen 0 ist. Außerdem lässt sich dieses Beispiel nicht als Konvexkombination von Permutationswürfeln schreiben. Das ergibt sich direkt daraus, dass die Permanente nicht positiv ist aber Permutationswürfel laut Definition immer eine Permanente gleich 1 besitzen. Um diese Eigenschaft allgemein zu betrachten, verwenden wir einen Satz aus [8], welcher notwendige Konditionen für die Darstellung magischer Würfel als Konvexkombination endlich vieler Permutationswürfel liefert. Dafür betrachten wir ein Lemma, welches Permutationswürfel anhand ihrer Schichten charakterisiert. Dabei bezeichnet $\circ$ das Hadamard-Produkt, welches die Einträge der Matrizen komponentenweise multipliziert. Es ergibt sich also für $A = (a_{ij}) \in \mathbb{R}^{n\times m}$ und $B = (b_{ij}) \in \mathbb{R}^{n\times m}$ das Hadamard-Produkt $A \circ B = (a_{ij}b_{ij}) \in \mathbb{R}^{n\times m}$.

Lemma 2.1. *Sei $P \in \mathbb{R}^{n\times n\times n}$ eine Hypermatrix der Ordnung n. Es gilt: P ist ein Permutationswürfel genau dann, wenn seine Schichten $P_k \in \mathbb{R}^{n\times n}$ Permutationsmatrizen sind für $1 \leq k \leq n$ und*

$$P_{k_1} \circ P_{k_2} = 0, \quad 1 \leq k_1 \neq k_2 \leq n.$$

Beweis. „$\Rightarrow$ “ Sei P ein Permutationswürfel. Laut Definition befindet sich auf jeder Linie genau ein Eintrag gleich 1. Dadurch gilt nach Fixierung eines Index, dass die Schicht P_k in jeder Zeile und Spalte genau einen 1-Eintrag hat und somit eine Permutationsmatrix ist. Wir nehmen an

$$(P_{k_1} \circ P_{k_2})_{ij} = p_{ijk_1}p_{ijk_2} = 1, \text{ für } k_1 \neq k_2.$$

Da die Einträge nur 0 oder 1 sein können, muss gelten $p_{ijk_1} = p_{ijk_2} = 1$, wodurch P zwei Einträge ungleich 0 auf einer Linie hat und somit kein Permutationswürfel ist.

„$\Leftarrow$ “ Seien nun P_k Permutationsmatrizen für $1 \leq k \leq n$, also $P = (P_1|P_2|\dots|P_n)$, und $P_{k_1} \circ P_{k_2} = 0$ für $1 \leq k_1 \neq k_2 \leq n$. Da die P_k genau n Einträge gleich 1 haben, hat P genau n^2 Einträge gleich 1. Laut Definition haben wir in unserem Würfel bei dieser Schreibweise den Index k fixiert. Fixieren wir zusätzlich den Index i oder j, erhalten wir Linien, welche auch Linien der Permutationsmatrizen sind und deshalb laut Definition jeweils genau einen Eintrag gleich 1 haben. Also gilt für den Fall $p_{ijk} = 1$, dass dieser Eintrag der einzige nichtnull Eintrag in den p_{-jk} Linien (für $1 \leq j, k \leq n$) und den p_{i-k} Linien (für $1 \leq i, k \leq n$) ist. Zu betrachten sind noch die Linien für Index i und Index j fixiert. Wir nehmen an, es gibt

i_0 und j_0, sodass

$$\sum_{k=1}^{n} p_{i_0 j_0 k} = 0, \tag{7}$$

also alle Einträge in dieser Linie gleich 0 sind. Da P_k Permutationsmatrizen sind, gilt für die i_0-Zeile in jeder Matrix, dass sich die Einträge zu 1 summieren. Damit erhalten wir insgesamt

$$\sum_{j=1}^{n} \sum_{k=1}^{n} p_{i_0 jk} = n.$$

und zusammen mit (7) ergibt sich, dass

$$\sum_{1 \le j \ne j_0 \le n} \sum_{k=1}^{n} p_{i_0 jk} = n.$$

Es müsste also ein $j \neq j_0$ geben, sodass

$$\sum_{k=1}^{n} p_{i_0 jk} \geq 2.$$

Damit hätten wir zwei Indizes $k_1 \neq k_2$ mit $(P_{k_1})_{i_0 j} = (P_{k_2})_{i_0 j} = 1$, also würde gelten $(P_{k_1} \circ P_{k_2})_{ij} \neq 0$, was der Annahme widerspricht. Daher muss gelten, dass es für gegebene Indizes i_0 und j_0 einen Index $1 \le k_0 \le n$ gibt, sodass $p_{i_0 j_0 k_0} = 1$, jede Linie von P hat also mindestens einen Eintrag gleich 1. Nehmen wir an, dass es einen weiteren Eintrag gleich 1 gibt, erhält man laut dem gleichen Argument für die Indizes i_0 und k_0 bzw. j_0 und k_0 denselben Widerspruch. Also gilt, P ist ein Permutationswürfel. □

Satz 2.1. *Sei $A = (a_{ijk}) \in \mathbb{R}^{n \times n \times n}$ ein stochastischer Würfel mit Schichten $A_k \in \mathbb{R}^{n \times n}$. Dann lässt A sich genau dann als Konvexkombination endlich vieler Permutationswürfel schreiben, wenn sich die A_k als Konvexkombinationen schreiben lassen, sodass*

$$A_k = \sum_{d=1}^{c} \alpha_k^{(d)} P_k^{(d)}, \quad 1 \le k \le n, \tag{8}$$

wobei $P_k^{(d)}$ Permutationsmatrizen sind und

$$P_{k_1}^{(d)} \circ P_{k_2}^{(d)} = 0, \quad \sum_{d=1}^{c} \alpha_k^{(d)} = 1, \quad \alpha_{k_1}^{(d)} = \alpha_{k_2}^{(d)} \geq 0,$$

für $1 \le k_1 \neq k_2 \le n$, $1 \le d \le c$.

Beweis. „$\Rightarrow$" Wir nehmen A als Konvexkombination von Permutationswürfeln $P^{(d)}$ an

$$A = \sum_{d=1}^{c} \alpha^{(d)} P^{(d)}, \sum_{d=1}^{c} \alpha^{(d)} = 1, 0 \leq \alpha^{(d)} \leq 1.$$

Wenn wir die einzelnen Schichten von A und $P^{(d)}$ betrachten, gilt für diese

$$A_k = \sum_{d=1}^{c} \alpha^{(d)} P_k^{(d)}, \quad 1 \leq k \leq n,$$

wobei $P_k^{(d)}$ Permutationsmatrizen sind und, da P ein Permutationswürfel ist, laut Lemma 2.1 $P_{k_1}^{(d)} \circ P_{k_2}^{(d)} = 0$ für $1 \leq k_1 \neq k_2 \leq n$ gilt. Damit ergibt sich bereits das gewünschte Resultat.

„$\Leftarrow$" Nun soll (8) gelten. Da damit die Konditionen von Lemma 2.1 für die $P_k^{(d)}$ gelten, können wir einen Permutationswürfel darstellen als $P^{(d)} = \left(P_1^{(d)} | P_2^{(d)} | \cdots | P_n^{(d)}\right)$. Laut Annahme können wir außerdem festlegen $\alpha_{k_1}^{(d)} = \alpha_{k_2}^{(d)} = \alpha^{(d)}$, wobei $\sum_{d=1}^{c} \alpha^{(d)} = 1$. Zusammen mit (8) gilt,

$$\sum_{d=1}^{c} \alpha^{(k)} P^{(k)} = \sum_{d=1}^{c} \alpha^{(d)} \left(P_1^{(d)} | \cdots | P_n^{(d)}\right) = \sum_{d=1}^{c} \left(\alpha^{(d)} P_1^{(d)} | \cdots | \alpha^{(d)} P_n^{(d)}\right) = (A_1 | \cdots | A_n) = A.$$

Damit ist A eine Konvexkombination endlich vieler Permutationswürfel. □

Bezeichne $A = (a_{ijk})$ den Würfel aus Beispiel 2.1 (i), die einzelnen Schichten A_k bestehen aus 3×3 Matrizen.

$$A_1 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \qquad A_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \qquad A_3 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

Betrachten wir jetzt alle 3×3 Permutationsmatrizen, können wir die einzelnen Schichten als Konvexkombinationen dieser schreiben, da für doppelt-stochastische Quadrate der Satz von Birkhoff- von Neumann weiterhin gültig ist.

$$P_{11} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad P_{21} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P_{22} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad P_{32} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Da jede Schicht genau zwei verschiedene Diagonalen hat, bekommen wir für $i = 1, 2, 3$ die eindeutige Darstellung

$$A_i = \sum_{j=1}^{n_{i1}} \alpha_{ij} P_{i1} + \sum_{j=1}^{n_{i2}} \beta_{ij} P_{i2},$$

wobei $0 \leq \alpha_{ij} \leq 1$ und $\sum_{j=1}^{n_{i1}} \alpha_{ij} = \sum_{j=1}^{n_{i2}} \beta_{ij} = \frac{1}{2}$. Es gilt außerdem

$$P_{11} \circ P_{21} \neq 0, \quad P_{21} \circ P_{31} \neq 0, \quad P_{22} \circ P_{31} \neq 0, \quad P_{21} \circ P_{12} \neq 0$$

Damit sind die Bedingungen aus Satz 2.1 nicht erfüllt und wir haben auch auf diese Art bestätigt, dass A keine Konvexkombination aus endlich vielen Permutationswürfeln sein kann.

Mit Beispiel 2.1 (ii) haben wir einen stochastischen Würfel mit positiver Diagonale, also ein Element aus Δ_n^3. Die Einträge können wie folgt gewählt werden.

$$\left(\begin{array}{ccc:ccc:ccc} & \frac{3}{5} & & 1 & & & & & \frac{3}{5} \\ \frac{3}{5} & & & & & \frac{3}{5} & & \frac{3}{5} & \\ & & \frac{1}{5} & & \frac{3}{5} & & \frac{3}{5} & & \end{array}\right)$$

Jedoch ergibt sich auch hier trotz positiver Diagonale, dass sich der Würfel nicht als Konvexkombination schreiben lässt. Die möglichen Permutationswürfel müssten folgende Form haben, um für die Konvexkombination in Frage zu kommen

$$\left(\begin{array}{ccc:ccc:ccc} 0 & * & * & 1 & 0 & 0 & 0 & * & * \\ * & 0 & * & 0 & * & * & * & * & 0 \\ * & * & * & 0 & * & * & * & 0 & * \end{array}\right),$$

wobei die freigelassenen Einträge 0 oder 1 sein können. Der einzig mögliche Permutationstensor, welcher Lemma 2.1 erfüllt und die obige Form hat, ist

$$(P_{12} | P_{31} | P_{11}) \,.$$

Damit ist auch dieses Beispiel nicht in Ω_n^3 enthalten. Wir haben gezeigt, dass die Teilmengenbeziehung $\Omega_n^3 \subset \Delta_n^3 \subset B_n^3$ echte Teilmengen sind.

2.2.2 Bedingungen an Ecken magischer Würfel

Bisher haben wir gezeigt, dass nicht alle Elemente der Menge der magischen Würfel als Konvexkombination von Permutationswürfeln dargestellt werden können, diese das Birkhoff-Polytop in Dimension 3 also nicht aufspannen. Wir können jedoch wieder leicht zeigen, dass die Permutationswürfel in der Menge der Extremalpunkte enthalten sind, auch wenn sie hier nur eine Teilmenge davon ausmachen.

Lemma 2.2. *Jeder Permutationswürfel $P = (p_{ijk})$ ist ein Extremalpunkt von B_n^3*

Beweis. Wir nehmen an, dass P kein Extremalpunkt ist. Dann muss es zwei stochastische Würfel $B \neq C \in B_n^3$ geben, sodass

$$P = \frac{1}{2}B + \frac{1}{2}C.$$

Da alle Einträge von P entweder 0 oder 1 sind, ergibt sich für die Einträge der anderen Tensoren $0 \leq b_{ijk}, c_{ijk} \leq 1$. Da B und C die Bedingungen stochastischer Würfel erfüllen müssen, gilt für $p_{ijk} = 0$, dass $b_{ijk} = c_{ijk} = 0$ und für $p_{ijk} = 1$, dass $b_{ijk} = c_{ijk} = 1$. Das widerspricht der Annahme, dass $B \neq C$. □

Es gilt die Umkehrung, nämlich der Satz von Birkhoff und von Neumann, in dieser Dimension eindeutig nicht. Unter bestimmten Bedingungen gilt zwar, dass ein stochastischer Würfel auch in Ω_n^3 liegt, wie man anhand von Satz 2.1 sehen kann, doch, wie unsere Beispiele zeigen, liegen auch viele außerhalb. Es liegt also nahe, dass die Menge der Extremalpunkte von B_n^3 mehr als nur die Permutationstensoren enthält. An den Beispielen wurde auch gezeigt, dass es, anders als bei magischen Quadraten, stochastische Würfel A gibt mit $\operatorname{per}(A) = 0$. Bevor wir also betrachten, welche weiteren Hypermatrizen weitere Ecken sein könnten, schauen wir uns ein Resultat aus [4] an, welches der Frage nach stochastischen Würfeln mit positiver Diagonale nachgeht.

Satz 2.2. *Sei A ein innerer Punkt von B_n^3, dann gilt $\operatorname{per}(A) > 0$. Folglich gilt, hat ein stochastischer Würfel keine positive Diagonale, gehört er zum Rand ∂B_n^3 von B_n^3.*

Beweis. Sei A aus dem Inneren von B_n^3. Dann existiert eine offene Kugel $\mathcal{B}(A)$ um A, welche auch in B_n^3 liegt. Wir betrachten einen beliebigen Permutationswürfel und können eine Verbindungsstrecke $c := tP + (1-t)A$ für ein $0 < t < 1$ finden, da $\Omega_n^3 \subseteq B_n^3$ und B_n^3 ein Polytop ist. Da Permutationswürfel immer eine positive Diagonale haben, liegt diese Verbindung in Δ_n^3. Den Schnitt zwischen c und dem Rand des Abschlusses der Umgebung von A, $\partial\,\mathrm{cl}(\mathcal{B}(A))$, bezeichnen wir mit C. Jetzt können wir die Antipode von C in Bezug auf das

Zentrum A finden und bezeichnen diese mit C'. Damit liegt A zwischen P und C' und wir können schreiben

$$A = sP + (1-s)C' \quad \text{für ein } 0 < s < 1.$$

Da P ein Permutationswürfel ist, muss auch gelten $A \in \Delta_n^3$, also $\mathrm{per}(A) > 0$. $\square$

Dieser Satz legt nahe, die Extremalpunkte anhand der Eigenschaften der Diagonalen der Würfel zu untersuchen. Dafür betrachten wir zwei Sätze aus [8]. Zuerst wird untersucht, unter welchen Bedingungen ein Extremalpunkt eine positive Diagonale haben kann. Dann werden zusätzliche Bedingungen formuliert, welche einen stochastischen Würfel zum Extremalpunkt machen.

Satz 2.3. *Sei A ein stochastischer Würfel mit $\mathrm{per}(A) > 0$. Es gilt, A ist genau dann ein Extremalpunkt von B_n^3, wenn A ein Permutationswürfel ist.*

Beweis. Die Rückrichtung ergibt sich aus Lemma 2.2. Um die andere Richtung zu zeigen, nehmen wir an, dass A kein Permutationswürfel ist. Da $\mathrm{per}(A) > 0$ gilt, gibt es einen zur positiven Diagonale $D_p(A)$ korrespondierenden Permutationswürfel. Wir setzen $a = \min_{a_{ijk}}\{a_{ijk} | p_{ijk} = 1\}$, also erhalten wir mit $A - aP$ einen nichtnegativen Tensor ungleich 0. Jetzt erhalten wir einen stochastischen Würfel mit $Q = \frac{A-aP}{1-a}$, wodurch A als Konvexkombination zweier verschiedener Punkte aus B_n^3 dargestellt wird. Damit kann A kein Extremalpunkt sein. $\square$

Satz 2.4. *Sei A ein zusammenhängender stochastischer Würfel mit $\mathrm{per}(A) = 0$. Hat jede Linie von A maximal zwei Einträge, welche positiv sind, und hat jede Schicht A_k genau zwei positive Diagonalen, sodass sich die Einträge der Diagonalen ungleich 1 nicht überschneiden, dann ist A ein Extremalpunkt von B_n^3*

Für den Beweis des Satzes wird noch ein Lemma benötigt, dessen Beweis in [8] nachzulesen ist. Das Lemma spricht eine Eigenschaft für magische Quadrate an, welche analoge Bedingungen zu den Schichten der magischen Würfel im obigen Satz erfüllen.

Lemma 2.3. *Sei A eine doppelt-stochastische Matrix, deren Zeilen und Spalten maximal zwei Einträge ungleich 0 haben und welche genau zwei positive Diagonalen hat, sodass sich* *die Einträge der Diagonalen, welche nicht 1 sind, nicht überschneiden. Falls es eine nicht-null Matrix E gibt, sodass $B = A + E$ und $C = A - E$ doppelt-stochastische Matrizen sind, dann gilt für die Einträge*

$$E_{ij} = \begin{cases} = 0, & A_{ij} \in \{0, 1\} \\ \neq 0, & 0 < A_{ij} < 1. \end{cases}$$

Außerdem befinden sich die positiven und negativen Einträge von E in verschiedenen Diagonalen von A.

Beweis Satz 2.4. Wir nehmen an, dass A kein Extremalpunkt ist und können deswegen $B, C \in B_n^3$ finden, sodass $A \neq B \neq C$ und $A = \frac{B+C}{2}$. Wir können also $B = A + E$ und $C = A - E$ schreiben, wobei E eine Hypermatrix der Dimension 3 ist, dessen Einträge sich in jeder Linie zu 0 summieren und dessen Schichten wir mit E_k bezeichnen. Damit bleiben die Liniensummen von B und C erhalten. Wir wissen, dass E mindestens einen Eintrag ungleich 0 haben muss, damit $A \neq B$ erfüllt ist. Wir können also annehmen, dass in der Schicht E_{k_0} der (i_0, j_0)-Eintrag ungleich 0 ist. Die Schichten von A erfüllen die Bedingungen von Lemma 2.3, also gilt für $1 \leq i, j \leq n$

$$(E_{k_0})_{ij} = \begin{cases} = 0, & (A_{k_0})_{ij} \in \{0, 1\} \\ \neq 0, & 0 < (A_{k_0})_{ij} < 1. \end{cases}$$

Da A zusammenhängend ist, gibt es in einer weiteren Schicht $k_1 \neq k_0$ einen positiven Eintrag an derselben Stelle (i_0, j_0). Damit die Summe der Einträge der Linien in E gleich 0 ergeben, muss gelten $(E_{k_1})_{i_0 j_0} = -(E_{k_0})_{i_0 j_0}$, da die Linien von E maximal 2 Einträge ungleich 0 haben. Es gilt

$$(E_{k_1})_{ij} = \begin{cases} = 0, & (A_{k_1})_{ij} \in \{0, 1\} \\ \neq 0, & 0 < (A_{k_1})_{ij} < 1. \end{cases}$$

Dieses Argument lässt sich für die weiteren Schichten $E_{k_0}, \ldots, E_{k_{n-1}}$ analog anwenden. Damit kann man für $0 \leq i, j, k \leq n$ festhalten

$$(E_k)_{ij} = \begin{cases} = 0, & (A_k)_{ij} \in \{0, 1\} \\ \neq 0, & 0 < (A_k)_{ij} < 1. \end{cases}$$

Da jede Linie von A maximal 2 Einträge hat, kann man, wie in Lemma 2.3 erwähnt, die positiven und negativen Einträge von E_k auf die beiden Diagonalen von A_k aufteilen. Es ergeben sich also jeweils zwei Möglichkeiten, die Einträge in einer Schicht von E aufzuteilen. An den Stellen, wo sich die Diagonalen schneiden, muss für A_k laut Voraussetzung der Eintrag gleich 1 sein und für E_k somit 0, es ergibt sich in diesem Fall also eine Nullzeile bzw. -spalte in dieser Schicht. Weiters gilt für E, dass die Einträge nur ungleich 0 sind, wenn A an der selben Stelle einen positiven Eintrag ($\neq 1$) hat. Damit ergibt sich wegen $\mathrm{per}(A) = 0$, dass sich auch für E keine positive oder negative Diagonale finden lässt. Dadurch können wir darauf schließen, dass es mindestens eine Linie von E geben muss, welche zwei positive oder zwei negative Einträge beinhaltet, wodurch die Summierung dieser Linie eine Zahl ungleich 0 liefert. Dadurch verlieren B und C ihre stochastischen Eigenschaften und sind nicht mehr in B_n^3 enthalten, wir erhalten einen Widerspruch. Es folgt, dass A ein Extremalpunkt von B_n^3 ist. $\square$

Anhand dieser Bedingungen lässt sich feststellen, dass Beispiel 2.1 (i) einer der Extremalpunkte ist. Dass die Diagonale nicht positiv ist, haben wir bereits gezeigt. Die zusätzliche Eigenschaft der Diagonalen der Schichten erkennt man direkt, wenn man die einzelnen Schichten in Matrixschreibweise betrachtet. Der im Beweis verwendete Tensor E kann folgende Form haben

$$\left(\begin{array}{ccc:ccc:ccc} 0 & -a & a & -a & a & 0 & a & 0 & -a \\ -a & a & 0 & 0 & -a & a & -a & 0 & a \\ a & 0 & -a & a & 0 & -a & 0 & 0 & 0 \end{array}\right),$$

wobei $a \neq 0$. Die Darstellung ist nicht eindeutig, da man in jeder Schicht die positiven und negativen Einträge vertauschen kann, jedoch ergibt sich für jede Darstellung, dass es Linien gibt, welche sich nicht zu 0 summieren. Wir erhalten also denselben Widerspruch wie im Beweis.

Mit diesen Resultaten wird eine ausreichende Bedingung gegeben, um einen stochastischen Würfel als Ecke des Birkhoff-Polytops in Dimension 3 zu charakterisieren. Jedoch sind auch damit nicht alle Extremalpunkte abgedeckt; es lassen sich weitere Beispiele finden, welche die geforderten Eigenschaften nicht erfüllen. Die Frage nach der Menge aller Extremalpunkte bleibt weiterhin offen, jedoch lässt sich die Anzahl bereits abschätzen. Wie in Abschnitt 2.1 bereits erwähnt, ist die Anzahl der $n \times n \times n$ Permutationswürfel gleich der der lateinischen Quadrate der Ordnung n, da es zwischen diesen eine eins-zu-eins Beziehung gibt. Laut [6] und [10] kann auch eine untere Schranke der gesamten Anzahl an Extremalpunkten mithilfe der Anzahl lateinischer Quadrate L_n bestimmt werden. Mit dieser Abschätzung wird gezeigt, dass die Menge der Permutationswürfel nur einen sehr kleinen Anteil an der gesamten Menge der Extremalpunkte ausmacht.

2.3 Doppelt-stochastische Tensoren

Neben den stochastischen Würfeln gibt es noch eine weitere Verallgemeinerung magischer Quadrate auf $n \times n \times n$ Tensoren. In [9] wird der Begriff doppelt-stochastischer Tensoren eingeführt. Auch Permutationstensoren erhalten dadurch eine andere Bedeutung.

Definition 2.6. Sei $n \in \mathbb{N}$

(i) Sei $A = (a_{ijk})$ ein nichtnegativer $n \times n \times n$ Tensor. A heißt *doppelt-stochastischer*

Tensor, wenn gilt

$$\sum_{i=1}^{n} a_{ijk} \leq 1 \quad \text{für } 1 \leq j,k \leq n,$$
$$\sum_{j,k \in \{1,\dots,n\}} a_{ijk} = 1 \quad \text{für } 1 \leq i \leq n,$$

also alle Einträge vertikaler Linien sich zu maximal 1 summieren und die Einträge horizontaler Ebenen sich genau zu 1 summieren.

(ii) Wenn in jeder horizontalen Ebene genau ein Eintrag gleich 1 ist, in jeder vertikalen Linie maximal ein Eintrag gleich 1 ist und die restlichen Einträge folglich 0 sind, dann heißt der doppelt-stochastische Tensor $P = (p_{ijk})$ *stochastischer Tensor*.

Betrachtet man diese Definition in Dimension 2, erhält man eine zeilen-stochastische Matrix, also eine Matrix, für welche nur gefordert ist, dass alle Zeilensummen 1 ergeben. Damit ist es nicht die Generalisierung magischer Quadrate, welche wir uns erwünscht haben, da wir „dreifach"-stochastische statt doppelt-stochastische Tensoren betrachten wollen. Dennoch ist es interessant zu betrachten, wie sich die Konvexkombinationen der stochastischen Permutationstensoren in diesem Fall verhalten. Es lässt sich zeigen, dass auch die Menge der doppelt-stochastischen Tensoren D_n^3 ein Polytop ist und auch hier, wie in allen bisher betrachteten Fällen, jeder stochastische Permutationstensor ein Extremalpunkt von D_n^3 ist. Mit dieser Verallgemeinerung ins Dreidimensionale erhalten wir Tensoren, auf diese sich der Satz von Birkhoff und von Neumann umformulieren lässt.

Satz 2.5. *Sei A ein nichtnegativer n-Tensor der Dimension 3. Es gilt, A ist ein doppelt-stochastischer Tensor genau dann, wenn A eine Konvexkombination endlich vieler stochastischer Permutationstensoren P_k ist,*

$$A = \sum_{k=1}^{d} \alpha_k P_k, \quad \textit{mit } \alpha_k \geq 0, \sum_{k=1}^{d} \alpha_k = 1.$$

Wie bereits oft in dieser Arbeit verwendet, kann man einen stochastischen Würfel als $n \times n^2$-Matrix anhand ihrer Schichten schreiben. Es lässt sich zeigen, dass es in diesem Fall eine $n^2 \times n^2$ Matrix C gibt, deren ersten n Zeilen A entsprechen und für welche gilt

(i) C ist eine doppelt-stochastische Matrix, genau dann wenn A ein doppelt-stochastischer Tensor ist.

(ii) C ist eine Permutationsmatrix, genau dann wenn A ein stochastischer Permutationstensor ist.

Dies ergibt sich dank der Eigenschaft, dass die vertikalen Linien sich *maximal* zu 1 summieren müssen. Daher gilt offensichtlich, dass die ersten n Zeilen von C die Bedingungen an doppelt-stochastische Tensoren erfüllen. Umgekehrt lässt sich die gegebene $n \times n^2$ Matrix mit weiteren n Zeilen erweitern, sodass die resultierende Matrix doppelt-stochastisch ist. Diese Erweiterung ist nicht eindeutig.

Mithilfe dieser Eigenschaften wollen wir nun den Beweis grob beschreiben. Da A aus dem verallgemeinerten Satz von Birkhoff und von Neumann ein stochastischer Würfel ist, können wir also eine Matrix C finden, welche die Schichten von A beinhaltet und auf welche der ursprüngliche Satz von Birkhoff und von Neumann zutrifft. Wir wissen also, C lässt sich als Konvexkombination endlich vieler Permutationsmatrizen P_k schreiben. Betrachten wir die ersten n Zeilen der P_k, befindet sich in jeder Zeile genau eine 1 und in jeder Spalte maximal eine 1. Dadurch sind die Bedingungen an Permutationstensoren in Matrixschreibweise erfüllt und deren Konvexkombination ergeben genau A.

3 Stochastische Tensoren höherer Dimension

Nun wieder zurück zu unserer ursprünglichen Definition stochastischer Tensoren, in welcher sich die Summe aller Einträge, wenn jeder bis auf einen Index festgehalten wird, zu 1 summieren. Im vorherigen Kapitel haben wir bereits den Begriff der *Hypermatrix* für beliebige Dimension m und Ordnung n definiert. Für alle weiteren Überlegungen haben wir uns jedoch auf die dritte Dimension beschränkt. Wir wollen nun betrachten, ob die Ergebnisse aus den vorherigen Abschnitten auch für $m > 3$ gelten, also für multistochastische Tensoren. Dabei bezeichnen wir eine nichtnegative Hypermatrix $A = (a_{i_1,\dots,i_m})$ als *multistochastisch*, wenn für alle $k \in \{1, \dots, m\}$ gilt

$$\sum_{i_k=1}^{n} a_{i_1,\dots,i_k,\dots,i_m} = 1, \quad 1 \leq j \neq k \leq n, 1 \leq i_j \leq n$$

und bezeichnen die Menge dieser analog als B_n^m. Dabei ergibt sich wieder eine kompakte und konvexe Menge. Für den Fall, dass jeder Eintrag 0 oder 1 ist, sprechen wir hier von *Permutationstensoren*. Um die Schichten der Elemente dieser Menge zu betrachten, müssen alle bis auf zwei Indizes festgehalten werden, damit wir einzelne Matrizen erhalten. In [8] wird die i_k-Linie von A notiert mit $A(i_1, \dots, i_{k-1}, :, i_{k+1}, \dots, i_m)$. Diese Notation und die Entdeckungen zu multistochastischen Tensoren übernehmen wir aus diesem Paper. Da sich die Eigenschaften eines Tensors, egal für welche Fixierung, analog verhalten, können wir o.B.d.A. den Fall $A(:, i_2, \dots, i_{m-1}, :)$ betrachten, bei dem alle bis auf den ersten und letzten Index fixiert sind. Damit ergeben sich $m-2$ fixe Indizes, also insgesamt n^{m-2} Schichten und A lässt sich schreiben als $A = (A_1|\cdots|A_{n^{m-2}})$. Um die Notation einfach zu halten, bezeichnen wir die Schichten erneut als A_k, wobei jetzt $1 \leq k \leq n^{m-2}$ gilt. Auch der Begriff *zusammenhängend* lässt sich mit dieser Definition sinnvoll auf multistochastische Tensoren anwenden.

Da wieder das Verhalten von Permutationstensoren in B_n^m betrachtet wird, also wieder Überlegungen zum Satz von Birkhoff und von Neumann in höheren Dimensionen gemacht werden, muss zuerst klargestellt werden, unter welchen Bedingungen ein Tensor als solcher gilt.

Lemma 3.1. *Sei P ein m-dimensionaler Tensor der Ordnung n. Es gilt, P ist genau dann ein Permutationstensor, wenn P_k Permutationsmatrizen sind für $1 \leq k \leq n^{m-2}$ und*

$$P(:, i_2, \dots, i_{k-1}, k_1, i_{k+1}, i_{m-1}, :) \circ P(:, i_2, \dots, i_{k-1}, k_2, i_{k+1}, i_{m-1}, :) = P_{k_1} \circ P_{k_2} = 0$$

für $1 \leq k_1 \neq k_2 \leq n$, wobei $\circ$ das Hadamard-Produkt bezeichnet.

Der Beweis dieses Lemmas ist, ebenso wie seine Formulierung, analog zu Lemma 2.1. Auch die triviale Richtung von Birkhoff und von Neumann ergibt sich wieder direkt. Man kann also zeigen, dass ein Permutationstensor P ein Extremalpunkt von B_n^m ist. Mithilfe der Permutationstensoren lässt sich die Diagonale eines multistochastischen Tensors A charakterisieren und damit auch der Begriff der Permanente $\operatorname{per}(A)$.

Ähnlich wie im letzten Kapitel lassen sich Satz 2.1 und 2.4 formulieren. Wir betrachten die multistochastischen Tensoren als Konvexkombination endlich vieler Permutationstensoren und beschreiben eine ausreichende Bedingung für die Extremalpunkte der Menge.

Satz 3.1. *Sei $A \in \mathbb{R}^{n^m}$ ein multistochastischer Tensor der Dimension m und Ordnung n mit Schichten $A_k \in \mathbb{R}^{n^{m-2}}$. Dann lässt sich A genau dann als Konvexkombination endlich vieler Permutationstensoren schreiben, wenn A_k sich als Konvexkombination darstellen lässt, sodass*

$$A_k = \sum_{d=1}^{c} \alpha_k^{(d)} P_k^{(d)}, \quad 1 \leq k \leq n^{m-2}$$

wobei $P_k^{(d)}$ Permutationsmatrizen sind und

$$\begin{aligned}
&\alpha^{(d)}_{(i_2,\ldots,i_{k-1},i_k=k_1,i_{k+1},\ldots,i_{m-1})} = \alpha^{(d)}_{(i_2,\ldots,i_{k-1},i_k=k_2,i_{k+1},\ldots,i_{m-1})} \geq 0,\\
&\sum_{d=1}^{c} \alpha^{(d)}_{(i_2,\ldots,i_{m-1})} = 1,\\
&P_{k_1} \circ P_{k_2} = 0
\end{aligned}$$

für $1 \leq k_1 \neq k_2 \leq n, 1 \leq d \leq c$.

Satz 3.2. *Sei A ein zusammenhängender multistochastischer Tensor mit* $\operatorname{per}(A) = 0$*. Hat jede Linie von A maximal zwei Einträge, welche positiv sind, und hat jede Schicht A_k genau zwei positive Diagonalen, sodass sich die Einträge der Diagonalen ungleich 1 nicht überschneiden, dann ist A ein Extremalpunkt von B_n^m.*

Gleich wie beim Beweis zu Satz 2.4 wird hier mit einem Widerspruch zur Aussage $A = \frac{B+C}{2}$ für $A \neq B \neq C \in B_n^m$ argumentiert. Es ergibt sich erneut, dass $B = A + E$ und $C = A - E$ nicht aufgehen kann für einen m-dimensionalen Tensor der Ordnung n, da dieser folgende Darstellung für $1 \leq i_k \leq n, 1 \leq k \leq m$ hätte

$$E_{i_1,\ldots i_m} = \begin{cases} = 0, & A_{i_1,\ldots,i_m} \in \{0,1\} \\ \neq 0, & 0 < A_{i_1,\ldots,i_m} < 1. \end{cases}$$

Für jede Schicht von E gibt es 2 Möglichkeiten, die positiven und negativen Diagonalen anzuordnen. Damit ergeben sich $2^{n^{m-2}}$ Möglichkeiten für E, aber keine von ihnen mit Diagonale. Der multistochastische Tensor A ergibt sich nicht als Konvexkombination zweier weiterer Elemente aus der Menge.

Mit diesen Sätzen ergibt sich auch für multistochastische Tensoren eine ausreichende Bedingung für die Extremalpunkte von B_n^m, die Frage nach einer notwendigen Bedingung bleibt weiterhin offen.

4 Fazit

Die Motivation dieser Arbeit war es, eine Antwort auf die Frage zu finden, ob sich der bekannte Satz von Birkhoff und von Neumann für magische Quadrate auch auf höhere Dimensionen verallgemeinern lässt. Dies konnte bereits für Dimension $m = 3$ und Ordnung $n > 2$ mit Gegenbeispielen widerlegt werden. Die Menge an magischen Würfeln wird also nicht nur von Permutationswürfeln aufgespannt. Auch die Eigenschaft der positiven Diagonale, welche doppelt-stochastische Matrizen erfüllen, konnte auf Hypermatrizen nicht ausgeweitet werden. Jedoch ließ sich eine notwendige und ausreichende Bedingung für einen multistochastischen Tensor A finden, sodass er sich als Konvexkombination von Permutationstensoren schreiben lässt. Weiters konnten Bedingungen an A formuliert werden, sodass dieser ein Extremalpunkt von B_n^m ist, wobei diese Tensoren zusammen mit den Permutationstensoren noch nicht die gesamte Menge an Extremalpunkten abdecken. Durch einen anderen Ansatz auf die Verallgemeinerung stochastischer Matrizen, nämlich doppelt-stochastische Tensoren, ließ sich ein dreidimensionales Konstrukt finden, für welches der Satz von Birkhoff und von Neumann weiterhin gültig ist.

Die Menge der stochastischen Würfel lässt immer noch viele Fragen offen, unter anderem die vollständige Charakterisierung ihrer Ecken. Sobald man diese versteht, öffnen sich auch die Türen für weitere Generalisierungen, zum Beispiel für *quantum magic cubes*, wie sie in [1] für Matrizen gemacht wurden. Die Verallgemeinerung von magischen Quadraten zu magischen Würfeln stellt also weiterhin genügend Motivation für zukünftige Arbeiten zur Verfügung.

Literaturverzeichnis

Literatur

[1] Gemma De Las Cuevas, Tom Drescher, and Tim Netzer. *Quantum Magic Squares: Dilations and their Limitations*. J. Math. Phys. 61, 111704 (2020).

[2] Alexander Barvinok. *A Course in Convexity*. Graduate Studies in Mathematics v. 54 (2002).

[3] Roger J. Webster *Convexity*. Oxford University Press (1994).

[4] Haixia Chang, Vehbi E. Paksoy, and Fuzhen Zhang. *Polytopes of Stochastic Tensors*. Ann. Funct. Anal. 7 no. 3 (2016), 386–393.

[5] Roger A. Horn, Charles R. Johnson. *Matrix Analysis*. 2nd ed., Cambridge Univ. Press (2013).

[6] Nathan Linial, Zur Luria. *On the vertices of the d-dimensional Birkhoff Polytope*. arXiv:1208.4218 (2012).

[7] Lek-Heng Lim. *Tensors and Hypermatrices*. Handbook of Linear Algebra 2nd. ed. Chapter 15 (2014).

[8] Lu-Bin Cui, Wen Li, Michael K. NG. *Birkhoff-Von Neumann Theorem For Multistochastic Tensors*. Siam J. Matrix Anal. Appl. Vol 35, No. 3 (2014), 956-973.

[9] Haibin Chen, Liqun Qi, Louis Caccetta, Guanglu Zhou. *Birkhoff-Von Neumann Theorem and Decomposition for Doubly Stochastic Tensors*. Linear Algebra and its Applications 583 (2019), 119-133.

[10] Zhongshan Li, Fuzhen Zhang, Xiao-Dong Zhang. *On the Number of Vertices of the Stochastic Tensor Polytope*. Linear and Multilinear Algebra, 65:10, 2064-2075, DOI: 10.1080/03081087.2017.1310178 (2017).