

Ein Beitrag zur  
Quantenkryptographie  
in endlichdimensionalen Systemen  
nebst weiteren Ergebnissen aus dem Gebiet  
der Quanteninformationstheorie



Vom Fachbereich Physik  
der Technischen Universität Darmstadt  
zur Erlangung der Würde  
eines Doktors der Naturwissenschaften  
(Dr. rer. nat.)  
genehmigte

D I S S E R T A T I O N

von

**Dipl.-Phys. Kedar S. Ranade**

aus Berlin

Darmstädter Dissertation  
Darmstadt 2009  
D17

Referent: Prof. Dr. rer. nat. GERNOT ALBER  
Korreferent: Prof. Dr. rer. nat. NORBERT GREWE  
Tag der Einreichung: Donnerstag, 30. Oktober 2008  
Tag der mündlichen Prüfung: Mittwoch, 4. Februar 2009

**Ein Beitrag zur  
Quantenkryptographie in endlichdimensionalen Systemen  
nebst weiteren Ergebnissen aus dem Gebiet  
der Quanteninformationstheorie**

**Zusammenfassung**

Diese Dissertation befaßt sich mit quantenkryptographischen Protokollen, die im Gegensatz zu den vorwiegend verwendeten zweidimensionalen Quantensystemen (Qubits) allgemeine endlichdimensionale Quantensysteme (Qudits) als Träger der Information zulassen. Hauptgegenstand der Untersuchungen ist dabei die maximal tolerierbare Fehlerrate solcher Protokolle und ihr Verhalten in Abhängigkeit von der Dimension der Informationsträger. Zu diesem Zweck wird eine Reihe von Konzepten eingeführt, die die Behandlung dieser Fragestellung erlauben. Insbesondere werden konkrete Protokolle vorgestellt, die bis zu einer maximal tolerierbaren Fehlerrate verwendbar sind, und es wird gezeigt, daß eine große Klasse von Protokollen bei höheren Fehleraten unbrauchbar ist. Es stellt sich unter anderem heraus, daß die maximal tolerierbare Fehlerrate in Zwei-Basis-Protokollen mit steigender Dimension auf bis zu 50 % wächst.

Neben den bisher genannten Schwerpunkten der Dissertation werden einige Einzelergebnisse auf dem Gebiet der Quanteninformationstheorie aufgeführt, die im Laufe des Promotionsvorhabens erzielt wurden.



**A contribution to  
Quantum Cryptography in finite-dimensional systems  
including further results from the field  
of Quantum Information Theory**

**Abstract**

This PhD thesis deals with quantum-cryptographic protocols which allow general finite-dimensional quantum systems (qudits) as carriers of information in contrast to the predominantly used two-dimensional quantum systems (qubits). The main focus of investigations is the maximum tolerable error rate of such protocols and its behaviour as a function of the dimension of the information carriers. For this purpose, several concepts are introduced which allow the treatment of this problem. In particular, protocols are presented which work up to a maximum tolerate error rate, and it is shown that a wide class of protocols cannot be used for higher error rates. Among other things, it turns out that the maximum tolerable error rate for two-basis protocols increases up to 50 % for high dimensions.

Apart from the above-mentioned main subjects of this thesis, some other results from the field of quantum information theory are given, which were achieved during this PhD project.



# Vorwort

Die vorliegende Dissertation wurde in den Jahren 2005 bis 2008 in der Arbeitsgruppe *Theoretische Quantenphysik* von Prof. Dr. rer. nat. GERNOT ALBER am *Institut für Angewandte Physik* angefertigt. In dieser Dissertation fasse ich die wesentlichen Ergebnisse meiner wissenschaftlichen Forschungen der vergangenen Jahre zusammen.

Die in dieser Dissertation behandelten Themen gehören alle zum Gebiet der *Quanteninformationstheorie*, einem verhältnismäßig jungen Zweig der theoretischen Physik, der sich zur Aufgabe nimmt, aus den axiomatischen Grundlagen der Quantenmechanik experimentell umsetzbare Ergebnisse zu gewinnen. Die Quanteninformationstheorie, so wie ich sie verstehe, unterscheidet sich von der übrigen Physik durch ihre Abstraktheit: während in der Physik im allgemeinen ein konkretes physikalisches System mathematisch beschrieben werden soll, nimmt sich die Quanteninformationstheorie den mathematischen *Formalismus* der Quantenmechanik zum Ausgangspunkt, untersucht Eigenschaften, die sich aus dem Formalismus ergeben und versucht erst dann, die Ergebnisse in physikalischen Systemen zu verwirklichen. Wohl eines der bekanntesten Teilgebiete der Quanteninformationstheorie ist die *Quantenkryptographie*, die sich mit der sicheren Übertragung von Nachrichten befaßt. Meine Untersuchungen zu diesem Thema, die in gewisser Weise die Untersuchungen meiner Diplomarbeit fortsetzen, bilden den ersten Hauptteil der Dissertation. Der zweite Hauptteil enthält Ergebnisse, die im Laufe der Zeit angefallen sind, aber nicht unmittelbar mit dem eigentlichen Promotionsthema in Verbindung stehen. Schließlich wird in einem Anhang ausführlich auf einige mathematische, physikalische und auch informationstheoretische Sachverhalte eingegangen, die – zumindest mittelbar – mit der Quanteninformationstheorie zusammenhängen; es handelt sich im wesentlichen um eine Zusammenstellung mehr oder minder bekannter Ergebnisse, die zum Verständnis des Textes erforderlich sind oder die in der Literatur uneinheitlich gebraucht werden oder nur schwierig aufzufinden sind. Ich habe versucht, den Text so zu verfassen, daß der Leser mit Ausnahme der Grundlagen der Quantenmechanik und gewisser mathematischer Kenntnisse

keinerlei Vorwissen zum Verständnis der Arbeit benötigt. Meinen Vorlieben entsprechend verwende ich wie in allen meinen Texten auch in dieser Dissertation die „alte“ deutsche Rechtschreibung; dies sollte dem Leser aber keine Schwierigkeiten bereiten.<sup>1</sup>

Ich bedanke mich bei Herrn Professor GERNOT ALBER für die Betreuung der Dissertation sowie bei den übrigen ehemaligen und gegenwärtigen Mitgliedern der Arbeitsgruppe. Für das Korrekturlesen der Dissertation mit vielen Verbesserungsvorschlägen danke ich Dipl.-Phys. ULRICH SEYFARTH. Schließlich bedanke ich mich bei meinen Eltern für die Unterstützung während des gesamten Studiums.

Unter dem Arbeitstitel *Fehlerkorrektur in höherdimensionalen Quantensystemen und Anwendungen auf Quantenalgorithmen und in der Kryptographie* wurde dieses Promotionsvorhaben in der Zeit vom 1. September 2005 bis zum 31. August 2008 durch ein Promotionsstipendium der Technischen Universität Darmstadt gefördert.

Darmstadt, im Oktober 2008

Kedar S. Ranade

In dieser nach der Ablegung der mündlichen Doktorprüfung<sup>2</sup> überarbeiteten Fassung wurden einige Fehler berichtigt, die mir unter anderem von den Referenten und von Dipl.-Phys. OLIVER KERN mitgeteilt wurden.

Mittlerweile sind die den Überlegungen des Kapitels 6 zugrundeliegenden Ergebnisse in zwei Arbeiten von MYHR u. a. [122, 123] erschienen.

Darmstadt, den 11. Februar 2009

Kedar S. Ranade

---

<sup>1</sup>Vgl. hierzu zum Beispiel Duden, Band 1 (Rechtschreibung der deutschen Sprache), 20. Auflage (Mannheim 1991), Der Große Duden: Wörterbuch und Leitfaden der deutschen Rechtschreibung, 16. Auflage (Leipzig 1968) und LUTZ MACKENSEN, *Großes Deutsches Wörterbuch* (ca. 1977); vgl. ferner Duden, Band 4 (Grammatik der deutschen Gegenwartssprache), 4. Auflage (Mannheim 1984).

<sup>2</sup>Die Prüfer waren Profs. Dres. GERNOT ALBER, NORBERT GREWE, ROBERT ROTH und THOMAS WALTHER, den Vorsitz führte der Dekan Prof. Dr. NORBERT PIETRALLA.

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>15</b>
1.1	Ein kurzer geschichtlicher Abriß . . . . .	15
1.1.1	Quantenmechanik . . . . .	15
1.1.2	Klassische Kryptographie . . . . .	16
1.1.3	Quanteninformationstheorie und -kryptographie . . . . .	17
1.2	Ziele und Aufbau dieser Arbeit . . . . .	19
<b>I</b>	<b>Quantenkryptographie</b>	<b>21</b>
<b>2</b>	<b>Allgemeine Einführung</b>	<b>23</b>
2.1	Klassische Begriffsbildungen . . . . .	23
2.1.1	Einheitswurzeln . . . . .	24
2.1.2	Wahrscheinlichkeitsverteilungen . . . . .	24
2.2	Grundbegriffe der Quantenmechanik . . . . .	25
2.2.1	Beschreibung von Quantenzuständen . . . . .	25
2.2.2	Zeitliche Entwicklung von Quantenzuständen . . . . .	26
2.2.3	Quantenoperationen . . . . .	26
2.2.4	Grundannahmen dieser Arbeit . . . . .	27
2.3	Qubit-Systeme . . . . .	27
2.3.1	Die Pauli-Matrizen . . . . .	28
2.3.2	Bloch-Kugel und Bloch-Vektor . . . . .	28
2.4	Verschränkungstheorie . . . . .	29
2.4.1	Bestimmung der Verschränkung . . . . .	29
2.4.2	Quantifizierung der Verschränkung . . . . .	30
2.4.3	Verschränkungsreinigung . . . . .	30
2.5	Verallgemeinerte Bell-Zustände . . . . .	31
2.5.1	Verallgemeinerte Bell-Basen . . . . .	31
2.5.2	Mischungen verallgemeinerter Bell-Zustände . . . . .	32
2.5.3	Verallgemeinert-isotrope Zustände . . . . .	33
2.6	Komplementäre Basen . . . . .	34

2.7	Einige unitäre Transformationen . . . . .	35
2.7.1	Verallgemeinerte XOR-Transformationen . . . . .	36
2.7.2	Die diskrete Fouriertransformation . . . . .	36
2.7.3	Verallgemeinerte Pauli-Matrizen . . . . .	37
<b>3</b>	<b>Grundlagen der Quantenkryptographie</b>	<b>39</b>
3.1	Klassische Kryptographie . . . . .	39
3.1.1	Aufgaben kryptographischer Protokolle . . . . .	39
3.1.2	Das <i>One-time pad</i> . . . . .	40
3.2	Begriffe der Quantenkryptographie . . . . .	41
3.2.1	Aufgaben quantenkryptographischer Protokolle . . . . .	41
3.2.2	Das Modell . . . . .	42
3.2.3	Zur Bedeutung der Authentisierung . . . . .	43
3.3	Quantenkryptographische Protokolle . . . . .	43
3.3.1	Protokolle, die komplementäre Basen verwenden . . . . .	43
3.3.2	Abschätzung der Kanalfehlerrate . . . . .	45
3.3.3	Angriffe auf die Protokolle . . . . .	46
3.4	Verschränkungsbasierte Protokolle . . . . .	46
3.4.1	Das Standard-Protokoll . . . . .	47
3.4.2	Zum Begriff der Sicherheit . . . . .	48
3.4.3	Vergleich mit <i>Prepare-and-Measure</i> -Protokollen . . . . .	49
3.5	Sicherheitsbeweise für Protokolle . . . . .	49
3.5.1	Calderbank-Shor-Steane-Codes . . . . .	49
3.5.2	Reduktion auf CSS-basierte Protokolle . . . . .	51
3.5.3	Reduktion auf <i>Prepare-and-Measure</i> -Protokolle . . . . .	52
3.5.4	Protokolle mit Zweiweg-Kommunikation . . . . .	53
3.5.5	Reduktion auf bell-diagonale Zustände . . . . .	54
3.6	Leitgedanken der Untersuchungen . . . . .	55
3.6.1	Ergebnisse über tolerierbare Fehlerraten . . . . .	55
3.6.2	Untersuchungen der tolerierbaren Fehlerraten . . . . .	56
3.6.3	Grundannahmen für die folgenden Kapitel . . . . .	57
3.6.4	Symmetrierelationen . . . . .	57
3.6.5	Asymptotisch-exponentielle Gleichheit . . . . .	58
<b>4</b>	<b>Korrigierbarkeit von Quantenzuständen</b>	<b>59</b>
4.1	Asymptotische Korrigierbarkeit . . . . .	59
4.1.1	Die Quanten-Shannon-Schranke . . . . .	60
4.1.2	Der Hauptsatz dieses Kapitels . . . . .	60
4.1.3	Die von-Neumann-Entropie als Schranke . . . . .	64
4.2	Asymptotische $B_n^{(d)}$ -Korrigierbarkeit . . . . .	65
4.2.1	Der $B_n^{(d)}$ -Schritt . . . . .	65

4.2.2	Die Entwicklung unter $B_n^{(d)}$ -Schritten . . . . .	66
4.2.3	Die Entwicklung der Ditfehler . . . . .	67
4.2.4	Die Entwicklung der Phasenfehler . . . . .	67
4.2.5	Eine neue Transformation . . . . .	68
4.2.6	Ein modifiziertes Protokoll . . . . .	68
4.2.7	Exponentielles Verhalten . . . . .	70
4.3	Quantenkryptographische Anwendungen . . . . .	71
4.3.1	Verallgemeinert-isotrope Zustände . . . . .	71
4.3.2	Abschätzung von $r^{(d)}$ für Zwei-Basis-Protokolle . . . . .	72
4.3.3	Maximal tolerierbare Fehlerraten . . . . .	72
<b>5</b>	<b>Konkrete Protokolle</b>	<b>75</b>
5.1	Korrektur der Phasenfehler . . . . .	75
5.1.1	Der $P_n^{(d)}$ -Schritt . . . . .	76
5.1.2	Das Verhalten der Zustände unter $P_n^{(d)}$ -Schritten . . . . .	78
5.2	Berechnung tolerierbarer Fehlerraten . . . . .	80
5.2.1	Das Protokoll . . . . .	81
5.2.2	Verallgemeinert-isotrope Zustände . . . . .	82
<b>6</b>	<b>Symmetrische Erweiterbarkeit von Quantenzuständen</b>	<b>83</b>
6.1	Grundlagen . . . . .	84
6.1.1	Bedeutung in der Quantenkryptographie . . . . .	84
6.1.2	Elementare Eigenschaften . . . . .	85
6.2	Symmetrische Erweiterbarkeit . . . . .	86
6.2.1	Die kommutative unitäre Gruppe . . . . .	86
6.2.2	Die $\mathfrak{U}_2$ -invarianten Zustände . . . . .	87
6.2.3	Erweiterungen $\mathfrak{U}_2$ -invarianter Zustände . . . . .	88
6.2.4	Die Spurbedingung . . . . .	88
6.2.5	Die Reduktion von $B_k$ auf $B'_k$ . . . . .	89
6.2.6	Die Aufstellung der Matrizen $B'_k$ . . . . .	90
6.2.7	Umformulierung der Spurbedingungen . . . . .	91
6.2.8	Vollständige Analyse für Qubits . . . . .	92
6.3	Quantenkryptographische Anwendungen . . . . .	94
6.3.1	Spezielle Matrizen . . . . .	94
6.3.2	$\mathfrak{U}_2$ -invariante Zustände . . . . .	96
6.3.3	Der verallgemeinert-isotrope Fall . . . . .	98
6.3.4	Ergebnisse . . . . .	99
<b>7</b>	<b>Schlußbemerkungen</b>	<b>101</b>
7.1	Ausblick auf weitere Untersuchungen . . . . .	101
7.2	Multi-Qudit-Emission . . . . .	102

7.3	De-Finetti-Sätze . . . . .	105
7.3.1	Ein geschichtlicher Überblick . . . . .	105
7.3.2	De-Finetti-Sätze für Quantensysteme . . . . .	106
7.3.3	Anwendungen . . . . .	107
<b>II Quanteninformationstheorie</b>		<b>109</b>
<b>8</b>	<b>Vollständig positive Abbildungen und der Jamiołkowski-Isomorphismus</b>	<b>111</b>
8.1	Einführung . . . . .	112
8.2	Der Jamiołkowski-Isomorphismus . . . . .	113
8.3	Positive Abbildungen . . . . .	114
8.4	Das Jamiołkowski-Kriterium . . . . .	116
8.5	Schlußbemerkungen . . . . .	118
<b>9</b>	<b>Entropische Unschärferelationen</b>	<b>119</b>
9.1	Einführung . . . . .	119
9.2	Relationen für gemischte Zustände . . . . .	121
9.3	Ideen zu einem Beweis . . . . .	122
<b>III Anhang</b>		<b>127</b>
<b>A</b>	<b>Ergänzungen zum Haupttext</b>	<b>129</b>
A.1	Die Entwicklung unter $B_n^{(d)}$ -Schritten . . . . .	129
A.2	Eigenschaften der Fouriertransformation . . . . .	131
A.3	Verzeichnis der verwendeten Notationen . . . . .	131
A.3.1	Allgemeine Symbole und Mengensymbole . . . . .	132
A.3.2	Weitere Symbole und Konventionen . . . . .	133
<b>B</b>	<b>Mathematische Grundlagen</b>	<b>135</b>
B.1	Wahrscheinlichkeitsrechnung . . . . .	135
B.1.1	Fakultät und Stirling-Reihe . . . . .	135
B.1.2	Binomialkoeffizienten und Binomischer Lehrsatz . . . . .	136
B.1.3	Zufallsvariablen und ihre Eigenschaften . . . . .	136
B.1.4	Binomial- und Normalverteilung . . . . .	137
B.2	Grundbegriffe der Algebra . . . . .	138
B.2.1	Relationen . . . . .	138
B.2.2	Grundlagen . . . . .	138
B.2.3	Untergruppen und Normalteiler . . . . .	140
B.2.4	Symmetrische und alternierende Gruppen . . . . .	140

B.2.5	Gruppenwirkungen . . . . .	141
B.3	Ringe und Körper . . . . .	142
B.3.1	Mengen mit zwei Verknüpfungen . . . . .	142
B.3.2	Restklassenringe und Primzahlkörper . . . . .	143
B.3.3	Endliche Körper . . . . .	144
B.3.4	Körpererweiterungen . . . . .	145
B.4	Begriffe der linearen Algebra . . . . .	146
B.4.1	Vektorräume, Moduln und Algebren . . . . .	146
B.4.2	Basen . . . . .	146
B.4.3	Lineare Abbildungen, Operatoren und Matrizen . . . . .	147
B.4.4	Das charakteristische Polynom . . . . .	147
B.4.5	Die Jordansche Normalform . . . . .	147
B.4.6	Vandermonde-Matrix und -Determinante . . . . .	148
B.4.7	Hilberträume . . . . .	149
B.5	Positive Matrizen . . . . .	150
B.5.1	Positive Matrizen und Minoren . . . . .	150
B.5.2	Das Hurwitz-Kriterium . . . . .	151
B.6	Tensorprodukte . . . . .	154
B.6.1	Grundlegende Definitionen . . . . .	154
B.6.2	Abbildung in den Produktraum, Skalarprodukt . . . . .	155
B.6.3	Operationen auf einem Tensorfaktor . . . . .	156
B.6.4	Teilspur und reduzierte Operatoren . . . . .	156
B.6.5	Die Schmidt-Zerlegung . . . . .	157
B.7	Darstellungen endlicher Gruppen . . . . .	159
B.7.1	Summen und Produkte von Darstellungen . . . . .	160
B.7.2	Irreduzible Darstellungen . . . . .	160
B.7.3	Charaktere von Darstellungen . . . . .	162
B.8	Normierte Vektorräume und Algebren . . . . .	162
B.8.1	Normen auf allgemeinen Vektorräumen . . . . .	163
B.8.2	Normen für quadratische Matrizen . . . . .	164
B.8.3	Zugeordnete Matrixnormen . . . . .	164
B.8.4	Gemeinsame Wahrscheinlichkeitsverteilungen . . . . .	165
B.9	Vermischte Ergebnisse . . . . .	168
B.9.1	Taylor-Reihen und Restglieder . . . . .	168
B.9.2	Die Fouriertransformation . . . . .	169
B.9.3	Das Haarsche Maß . . . . .	169
B.9.4	Die Singulärwertzerlegung . . . . .	170
B.9.5	Einige Ungleichungen . . . . .	171

<b>C</b>	<b>Informationstheorie und Kodierung</b>	<b>173</b>
C.1	Klassische Informationstheorie . . . . .	173
C.1.1	Die Shannon-Entropie . . . . .	173
C.1.2	Die relative Entropie . . . . .	174
C.1.3	Bedingte Entropie und gemeinsame Information . . . . .	175
C.1.4	Markow-Ketten und starke Subadditivität . . . . .	176
C.1.5	Rényi-Entropien . . . . .	178
C.1.6	Typische Sequenzen . . . . .	180
C.2	Quanteninformationstheorie . . . . .	181
C.2.1	Die von-Neumann-Entropie . . . . .	181
C.2.2	Grundbegriffe der Quanteninformationstheorie . . . . .	182
C.2.3	Die starke Subadditivität . . . . .	185
C.2.4	Die Holevo-Schranke und das HSW-Theorem . . . . .	186
C.2.5	Die Neumark-Erweiterung . . . . .	188
C.2.6	Die Subadditivität der Spurnorm . . . . .	190
C.2.7	Die Uhlmann-Fidelity . . . . .	191
C.3	Mathematische Hilfsmittel . . . . .	192
C.3.1	Schranken für Binomialkoeffizienten . . . . .	192
C.3.2	Chernoff-Schranken . . . . .	193
C.4	Grundbegriffe der Kodierungstheorie . . . . .	198
C.4.1	Allgemeine Begriffe . . . . .	198
C.4.2	Lineare Codes . . . . .	199
C.4.3	Diskrete Kugeln . . . . .	200
C.4.4	Gilbert-Varshamov-Schranken . . . . .	201
	<b>Schluß</b>	<b>207</b>
	Bibliographie . . . . .	207
	Lebenslauf des Verfassers . . . . .	229

# Kapitel 1

## Einführung

In der Einführung wird zunächst auf die geschichtlichen Ursprünge der in dieser Dissertation behandelten Themen eingegangen. Anschließend wird in einem zweiten Abschnitt ein Überblick über die im Rahmen dieses Promotionsvorhabens erzielten Ergebnisse gegeben.

### 1.1 Ein kurzer geschichtlicher Abriss

In diesem Abschnitt wird kurz die historische Entwicklung der Quantenmechanik und -informationstheorie sowie der klassischen Kryptographie und der Quantenkryptographie geschildert. Der Leser kann diesen Abschnitt ohne Nachteil überspringen.

#### 1.1.1 Quantenmechanik

Die gesamte moderne Physik wird wesentlich von den zwei großen Umwälzungen zu Beginn des 20. Jahrhunderts geprägt: der *Relativitätstheorie* und der *Quantenmechanik*. Während die Relativitätstheorie vorwiegend mit dem Namen ALBERT EINSTEINS verbunden wird, waren an der Entstehung der Quantenmechanik eine ganze Reihe bedeutender Physiker beteiligt.

Als Gründungsjahr der Quantenmechanik gilt allgemein das Jahr 1900, in dem MAX PLANCK seine Quantenhypothese aufstellte, um die spektrale Verteilung der Schwarzkörperstrahlung zu erklären. Die Frühzeit der Quantenmechanik war durch Versuche geprägt, bis dahin unerklärliche Phänomene physikalisch zu verstehen: hierzu gehören etwa die Lichtquantenhypothese EINSTEINS von 1905 und das BOHRsche Atommodell von 1913. Neue Eigenschaften von Teilchen, zum Beispiel der Elektronenspin, wurden in Versuchen entdeckt und interpretiert (vgl. auch das STERN-GERLACH-Experiment).

Der bis heute verwendete Formalismus der Quantenmechanik entstand in den Jahren ab 1925 mit den Arbeiten SCHRÖDINGERS über die *Wellenmechanik* und denen HEISENBERGS über die *Matrizenmechanik*. Etwas später erschienen zusammenfassende Darstellungen der Quantenmechanik bei DIRAC, auf den die Bra-Ket-Notation zurückgeht, und bei VON NEUMANN [124], der die mathematischen Grundlagen darlegte und den Begriff des Hilbertraums in die Quantentheorie einführte. Seit dieser Zeit blieben die formalen Grundlagen der Theorie weitgehend unangefochten, und die gegenwärtige Physik baut auf ihnen auf.

### 1.1.2 Klassische Kryptographie

Schon immer haben Menschen versucht, bestimmtes Wissen – Information – vor dem Zugriff anderer zu schützen. Neben Verfahren, die dieses Ziel durch die Tarnung der Information erreichen wollen, wurden und werden Verfahren entwickelt, um Nachrichten zu *verschlüsseln*. Der unbefugte Leser erhält nur einen verschlüsselten Text zur Einsicht, aus dem er – falls das Verfahren wirkt – nichts entnehmen kann. In der Geschichte der Menschheit wurden viele Verschlüsselungsverfahren erfunden, die meisten aber über kurz oder lang gebrochen.

Die wohl einfachsten Verschlüsselungsverfahren ersetzen jedes Zeichen der Nachricht durch ein anderes festes Zeichen, oft auch Phantasiezeichen.<sup>1</sup> Eine systematische Variante dieser Idee ist nach CAESAR benannt: man ersetze den Buchstaben A durch D, B durch E, ..., Z durch C; verschiebt man nicht immer um drei Buchstaben sondern um einen beliebigen festen Wert, so erhält man insgesamt 26 Möglichkeiten, die man durch einen Schlüsselbuchstaben benennen kann, in der CAESAR-Verschlüsselung zum Beispiel D. Auf ALBERTI und VIGNÈRE geht die Idee zurück, statt eines Buchstabens ein Schlüsselwort zu wählen, das durch Wiederholung auf die Länge des zu verschlüsselnden Textes gebracht wird, so daß die Verschlüsselung eines Buchstabens von seiner Position im Text abhängig ist; hat das Schlüsselwort  $n \in \mathbb{N}$  Buchstaben, so gibt es  $26^n$  mögliche Schlüssel. Dieses Verfahren wird beweisbar sicher, wenn das Schlüsselwort zufällig gewählt wird und genau so lang ist wie die Nachricht; dies ist einer der Ansatzpunkte der Quantenkryptographie (vgl. Unterabschnitt 3.1.2).

Neben vielen weiteren klassischen Verschlüsselungsverfahren, die an dieser Stelle nicht angesprochen werden können, hat die moderne Kryptographie im Computerzeitalter eine ganze Reihe komplizierter Verfahren hervorgebracht.

---

<sup>1</sup>Diese Verfahren können oft durch Häufigkeitsanalysen der Sprachen gebrochen werden; das gilt auch für etwas komplexere Verfahren, die Homophone, d. h. verschiedene Zeichen für den gleichen Buchstaben, und Nullen, d. h. Zeichen ohne Bedeutung, verwenden.

Bekannt ist zum Beispiel der engl. *Advanced Encryption Standard* (AES, auch Rijndael), der gegenwärtig Verwendung findet. Von großer Bedeutung sind sog. *asymmetrische Verfahren* (*Public-Key-Kryptographie*) wie insbesondere das RSA-Verfahren von RIVEST, SHAMIR und ADLEMAN und der DIFFIE-HELLMAN-Schlüsselaustausch. Ihre Sicherheit beruht auf der vermuteten, aber unbewiesenen Schwierigkeit, große Zahlen im Sinne der Komplexitätstheorie effizient in ihre Primfaktoren zu zerlegen oder diskrete Logarithmen zu berechnen. Niemand kann gegenwärtig sagen, ob diese Probleme für alle Zeiten schwierig sein werden; im allgemeineren Zusammenhang handelt es sich um die ungelöste Frage, ob die Komplexitätstheoretische Gleichheit  $P = NP$  erfüllt ist oder nicht.

Zuletzt sei angemerkt, daß alle bisher beschriebenen Verfahren den in der klassischen Kryptographie als KERCKHOFFs *Maxime* bezeichneten Grundsatz befolgen, daß die Sicherheit eines Verschlüsselungsverfahrens allein von der Unkenntnis des Schlüssels herrühren muß. Ein potentieller Angreifer darf sämtliche Verfahren (Algorithmen) und Maschinen kennen, nicht aber einen geheimen Schlüssel; der Vorteil ist, daß die Geheimhaltung sich nur auf den Schlüssel bezieht, der ggf. auch schnell ausgetauscht werden kann.

Für verschiedene Gesichtspunkte der klassischen Kryptographie verweise ich auf die Literatur; populärwissenschaftlich sind das Buch von SINGH [168] oder auch ein Artikel von VON RANDOW [142]; detaillierter sind für klassische Verfahren das Buch von BAUER [10], für modernere das von BUCHMANN [27].

### 1.1.3 Quanteninformationstheorie und -kryptographie

Ihre scheinbare Unanschaulichkeit führte dazu, daß viele Physiker sich nicht mit der Quantenmechanik – ebensowenig wie mit der Relativitätstheorie – abfinden wollten.<sup>2</sup> Im Jahre 1935 veröffentlichten EINSTEIN, PODOLSKY UND ROSEN [48] ein Gedankenexperiment, das heute als *EPR-Paradoxon* bekannt ist und welches die Unvollständigkeit der Quantenmechanik aufzeigen sollte; die Verfasser hofften, daß sich die Quantenmechanik mittels *verborgener Parameter* zu einer umfassenden klassischen Theorie ergänzen lassen könnte.<sup>3</sup>

---

<sup>2</sup>Bekannt sind etwa der EINSTEIN zugeschriebene Satz „Gott würfelt nicht!“ und seine Ablehnung „spukhafter Fernwirkungen“ ebenso wie SCHRÖDINGERS Ausspruch „Wenn es doch bei dieser verdammten Quantenspringerei bleiben soll, so bedaure ich, mich mit der Quantentheorie überhaupt beschäftigt zu haben.“ 1926 in Kopenhagen zu NIELS BOHR (zitiert nach HEISENBERG [73], S. 291).

<sup>3</sup>Dies ist in gewissem Sinne durch die BOHMsche *Führungswellentheorie* tatsächlich möglich, jedoch ist diese Theorie durch ihren nicht-lokalen Charakter nicht weniger seltsam als die gewöhnliche Quantenmechanik.

Hierdurch veranlaßt veröffentlichte SCHRÖDINGER [161] noch im gleichen Jahr eine Arbeit, in der er den Begriff der *Verschränkung*<sup>4</sup> prägte, der heutzutage als eines der Hauptmerkmale der Quantenmechanik gilt; er illustrierte ihn mit SCHRÖDINGERS *Katze*, die hier zum ersten Mal Erwähnung fand.

Den Grundstein für das gesamte Gebiet der Quanteninformationstheorie legte BELL [13] im Jahre 1964 mit seiner Arbeit über das EPR-Paradoxon, in der er zeigte, daß die Quantenmechanik als formales Gedankengebäude und *Theorien verborgener lokaler Parameter* (lokal-realistische Theorien) unvereinbar sind.<sup>5</sup> Die *BELLSche Ungleichung* zeigt eine Möglichkeit auf, durch ein *Experiment* zu prüfen, ob sich die reale Welt quantenmechanisch oder lokal-realistisch verhält. Eine dem Experiment besser zugängliche Verallgemeinerung dieser Ungleichung, die Quantenmechanik und Theorien verborgener lokaler Parameter unterscheiden kann, stammt von CLAUSER, HORNE, SHIMONY UND HOLT [41] und ist als *CHSH-Ungleichung* bekannt.

Das erste weithin bekannte Experiment, welches diese Fragen untersuchte, veröffentlichten ASPECT, DALIBARD UND ROGER [4] im Jahre 1982. Bis auf zwei Schlupflöcher (Zufälligkeit der Auswahl der Messungen und Detektionseffizienz), die jedes für sich geschlossen wurden, hat noch jedes Experiment die Quantenmechanik bestätigt; so veröffentlichten WEIHS u. a. [183] 1998 eine Arbeit, in der die Meßrichtungen der Detektoren zufällig und raumartig voneinander getrennt ausgewählt werden. Bis heute sind diese Fragestellungen Gegenstand der Forschung. Zwar hat die Quantenmechanik bisher allen theoretischen und experimentellen Einwendungen standgehalten, man kann jedoch nicht daraus schließen, daß dies für alle Zeiten so bleiben wird. Gegenwärtig muß jedoch die Quantenmechanik als richtig angesehen werden.

Neben diesen Grundlagenfragen begannen sich einige Leute für die technischen Anwendungen der Quantenmechanik zu interessieren. Die Idee des Quantencomputers – Quantensysteme mit Quantensystemen zu simulieren – wird allgemein FEYNMAN [51] zugeschrieben. Einer weiteren Öffentlichkeit bekannt wurde sie erst, nachdem SHOR zwischen 1994 und 1997 Algorithmen für Quantencomputer entwickelte, die in der Lage sind, große Zahlen effizient zu faktorisieren und diskrete Logarithmen effizient zu berechnen. Sollte es gelingen, einen Quantencomputer zu bauen, so stellt er eine Bedrohung für die Sicherheit fast aller gegenwärtig genutzten Verfahren der klassischen Kryptographie dar. Die Realisierung von Quantenrechnern, die diese Algorithmen in großem Maßstab umsetzen können, steht bislang noch aus.

---

<sup>4</sup>Zum ersten Mal taucht der Begriff als „Verschränkung der Voraussagen“ auf; im gleichen Jahr erscheint eine englischsprachige Arbeit [162], in der er den Begriff *entanglement* verwendet, der aber keine wörtliche Übersetzung des Begriffs „Verschränkung“ ist.

<sup>5</sup>Genaugenommen betrachtete er eine vereinfachte Variante des EPR-Paradoxons unter Verwendung zweier Spin-1/2-Teilchen, die BOHM UND AHARONOV [20] angaben.

Im Gegensatz zur klassischen Kryptographie versucht die Quantenkryptographie, die Sicherheit aus Naturgesetzen abzuleiten. Im Rahmen der untersuchten Modelle sind eine Reihe von Verfahren beweisbar sicher, vorausgesetzt, daß die Quantenmechanik als physikalische Theorie richtig ist. Wäre dies nicht der Fall, so hätte dies vielfältige Auswirkungen auf weite Gebiete der Physik und der Naturwissenschaften im ganzen.

Als erster Schritt in Richtung Quantenkryptographie gilt eine um 1969/70 angestellte Überlegung WIESNERS [185], die mangels Interesse der Physiker-gemeinde aber erst im Jahre 1983 veröffentlicht werden konnte. Berühmt werden sollte das 1984 von BENNETT UND BRASSARD [17] vorgestellte Protokoll (*BB84-Protokoll*), das auch heute noch als das Standardbeispiel aller quantenkryptographischen Protokolle gilt. Die Grundidee dieses Protokolls ist leicht verständlich, der *Beweis* der Sicherheit unter allen im Rahmen der Quantenmechanik erlaubten Angriffen aber schwierig zu führen. Auf den ursprünglichen, sehr langen Beweis von MAYERS [120] folgten weitere, von denen insbesondere der von SHOR UND PRESKILL [167] von 2000 zu nennen ist, der erstmals eine Verbindung zwischen der Sicherheit quantenkryptographischer Protokolle und der Verschränkung ausnutzte, die EKERT [49] schon im Jahre 1991 aufgezeigt hatte.

Da jede reale Übertragung fehlerbehaftet ist, ist es wichtig zu wissen, wie hoch der hierdurch erzeugte Fehler höchstens sein darf, damit die Sicherheit des Protokolls gewährleistet bleibt; dies ist die Frage nach der *maximal tolerierbaren Fehlerrate*. SHOR UND PRESKILL geben für das BB84-Protokoll eine tolerierbare Fehlerrate von 11,0% an, und GOTTESMAN UND LO [64] und CHAU [30] erhöhten diesen Wert auf 20,0%. Es zeigt sich, daß die tolerierbare Fehlerrate stark vom verwendeten Protokoll abhängt, und dies ist einer der Ausgangspunkte dieser Dissertation.

## 1.2 Ziele und Aufbau dieser Arbeit

Das Ziel dieser Dissertation ist es, einen Beitrag zum theoretischen Verständnis der Quanteninformationstheorie im allgemeinen und der Quantenkryptographie im besonderen zu leisten. Das Hauptaugenmerk wird dabei auf die von quantenkryptographischen Protokollen tolerierbaren Fehlerraten gelegt. Es wird angestrebt, möglichst hohe tolerierbare Fehlerraten zu erreichen; kann aber für eine Rate kein sicheres Protokoll angegeben werden, so wird versucht, eine möglichst große Klasse von Protokollen auszuschließen. Im Rahmen des Promotionsvorhabens sind neben Ergebnissen auf dem Gebiet der Quantenkryptographie noch eine Reihe kleinerer Ergebnisse angefallen, die ebenfalls in die Dissertation aufgenommen wurden.

Diese Dissertation unterteilt sich in zwei Hauptteile und einen Anhang, diese wiederum in Kapitel, Abschnitte und Unterabschnitte. Die Gliederung der Arbeit ist die folgende:

1. Der erste Hauptteil enthält die auf dem Gebiet der Quantenkryptographie erzielten Ergebnisse:
  - (a) das Kapitel 2 enthält eine kurze Einführung in das Gebiet der Quanteninformationstheorie im Hinblick auf die in der Quantenkryptographie verwendeten Konzepte und Methoden;
  - (b) das Kapitel 3 faßt die Grundideen der Quantenkryptographie in endlichdimensionalen Systemen zusammen und zeigt die grundlegenden Protokolle und Sicherheitsüberlegungen;
  - (c) im Kapitel 4 wird die Korrigierbarkeit bell-diagonaler Quantenzustände mittels Zweiweg-Fehlerkorrektur sowie asymmetrischer CSS-Codes als Einweg-Fehlerkorrektur untersucht; es werden allgemeine Schranken für tolerierbare Fehlerraten hergeleitet;
  - (d) im Kapitel 5 wird ein konkretes Protokoll vorgestellt, welches die in Kapitel 4 genannten Fehlerraten erreichen kann;
  - (e) in Kapitel 6 wird das Problem der Einweg-Korrektur aus dem Blickwinkel des Konzeptes der symmetrischen Erweiterbarkeit von Quantenzuständen angegangen; hiermit werden einige Aussagen des Kapitels 4 verallgemeinert;
  - (f) das Kapitel 7 enthält einige Anmerkungen zur Quantenkryptographie und einen kurzen Ausblick.
2. Der zweite Hauptteil enthält einige kleinere Ergebnisse, die sich im Laufe des Promotionsvorhabens ergeben haben:
  - (a) das Kapitel 8 enthält einen vereinfachten Beweis einer Aussage aus der Theorie der vollständig positiven Abbildungen;
  - (b) das Kapitel 9 beschäftigt sich mit einigen vorläufigen Ergebnissen zu entropischen Unschärferelationen.
3. Der Anhang enthält eine Reihe vorwiegend mathematischer Ergebnisse, auf die aus den Hauptteilen verwiesen wird oder die als ergänzende Information dienen, darunter ein Symbolverzeichnis auf S. 132f.

Es wird angestrebt, alle wesentlichen in dieser Dissertation getroffenen Aussagen vollständig und mathematisch präzise zu formulieren und lückenlos zu beweisen. Hierzu werden wichtige Aussagen in Definitionen, Lemmata (Hilfssätzen), Sätzen, Hauptsätzen und Korollaren (Folgerungen) formuliert.

**Erster Hauptteil**

**Quantenkryptographie  
in endlichdimensionalen  
Systemen**



## Kapitel 2

# Allgemeine Einführung

In diesem Kapitel werden die Begriffe eingeführt, die für das Verständnis der Quantenkryptographie und damit für den gesamten ersten Hauptteil erforderlich sind. Insbesondere wichtig sind hierbei die Grundbegriffe für Qudit-Systeme, also für die Beschreibung  $d$ -dimensionaler Quantensysteme. Für Details vergleiche man die Anhänge.

### 2.1 Klassische Begriffsbildungen

In der Informatik verwendet man den Begriff *Bit* (Kurzform von engl. *binary digit*) für eine Binärziffer. Das quantenmechanische Analogon zu einem Bit ist das *Qubit*, das heißt, ein quantenmechanisches System, das durch einen zweidimensionalen Hilbertraum beschrieben wird. Die Standardbeispiele für solche Systeme sind zum einen Spin- $1/2$ -Teilchen wie das Elektron, zum anderen Photonen, deren Polarisationszustände betrachtet werden.

Verzichtet man auf die Einschränkung, daß der Hilbertraum des Systems zweidimensional ist, fordert aber (der mathematischen Einfachheit wegen) einen endlichdimensionalen Hilbertraum, so nennt man das abstrakt gefaßte System ein *Qudit*, wobei das  $d$  auf die Dimension  $d \in \mathbb{N} \setminus \{1\}$  hinweist.<sup>1</sup> Es bietet sich an, eine Basis des Qudit-Hilbertraums  $\mathcal{H} = \mathbb{C}^d$  mit den Zahlen  $\mathbb{Z}_d := \{0, \dots, d-1\}$  zu benennen. Bestimmte Operationen auf Qudits können dann leicht mit Hilfe modularer Arithmetik beschrieben werden; hierzu bezeichnen „ $\oplus$ “ und „ $\ominus$ “ die Addition und Subtraktion modulo  $d$ . In Analogie zu Bit kann man nun noch von einem klassischen *Dit* sprechen.<sup>2</sup>

---

<sup>1</sup>Manche Autoren nennen die Hilbertraumdimension  $n$  oder  $N$  und sprechen dementsprechend von *Qunit* bzw. *QuNit*.

<sup>2</sup>Aus sprachlicher Sicht wäre wohl das Wort *Digit* (Ziffer, Stelle) vorzuziehen.

Es ist zunächst einmal völlig unbedeutend, welches physikalische System dabei tatsächlich vorliegt und ob es überhaupt ein sinnvolles physikalisches System gibt, in dem man das Qudit realisieren kann.

### 2.1.1 Einheitswurzeln

Für eine Zahl  $n \in \mathbb{N}$  bezeichnet man jede Zahl  $z \in \mathbb{C}$ , für die  $z^n = 1$  gilt, als  $n$ -te *Einheitswurzel*. Die Menge der  $n$ -ten Einheitswurzeln ist

$$\left\{ \exp \left[ \frac{2\pi i k}{n} \right] \mid k \in \mathbb{Z}_n \right\}. \quad (2.1)$$

Die Einheitswurzeln mit  $\text{ggT}(k, n) = 1$  erzeugen die Gruppe aller Einheitswurzeln, und man nennt sie *primitiv*. Die Einheitswurzel mit  $k = 1$ , also dem kleinsten positiven Argument, ist stets primitiv und heißt *Hauptwert*.

Im gesamten Text bezeichne  $z := z_d := \exp(2\pi i/d)$  den Hauptwert der  $d$ -ten Einheitswurzel. Die Abbildung  $\mathbb{Z}/d\mathbb{Z} \rightarrow \mathbb{C}$ ,  $k \mapsto z^k$  ist eine treue Darstellung von  $\mathbb{Z}/d\mathbb{Z}$ . Viele Texte verwenden die Bezeichnung  $\omega := z_d$ .

### 2.1.2 Wahrscheinlichkeitsverteilungen

Die Menge der Wahrscheinlichkeitsverteilungen einer Zufallsvariablen mit der Ergebnismenge  $\{1, \dots, n\}$  – oder allgemeiner mit einer Ergebnismenge der Mächtigkeit  $n \in \mathbb{N}$  – kann mit der Menge der auf Eins normierten  $n$ -Tupel nicht-negativer reeller Zahlen identifiziert werden. Es bietet sich also an, die Menge all dieser Wahrscheinlichkeitsverteilungen mit

$$\mathcal{W}_n := \left\{ (p_1, \dots, p_n) \in \mathbb{R}^n \mid (\forall i \in \{1, \dots, n\})(p_i \geq 0), \sum_{i=1}^n p_i = 1 \right\} \subseteq \mathbb{R}^n$$

zu bezeichnen. Als ein Maß für die Unordnung einer solchen Verteilung (oder für die Unvorhersagbarkeit eines Ereignisses) dient die *SHANNON-Entropie*.

#### Definition 2.1 (Shannon-Entropie)

Für  $p = (p_0, \dots, p_{d-1}) \in \mathcal{W}_d$  nennt man die Funktion  $H_d : \mathcal{W}_d \rightarrow [0; 1]$  mit

$$H_d(p) := - \sum_{i=0}^{d-1} p_i \log_d p_i = -(\ln d)^{-1} \sum_{i=0}^{d-1} p_i \ln p_i$$

die SHANNON-Entropie. Die binäre SHANNON-Entropie  $H : [0; 1] \rightarrow [0; 1]$  wird durch  $H(x) := H_2[(x, 1-x)]$  definiert.

Hier und im gesamten Text wird  $0 \log_b 0 := \lim_{x \rightarrow 0^+} x \log_b x = 0$  gesetzt. Die Basis des Logarithmus definiert die Informationseinheit; die Verwendung der Basis  $d$  bedeutet also, daß die Information in Dits gemessen wird.

## 2.2 Grundbegriffe der Quantenmechanik

Es wird vorausgesetzt, daß der Leser mit den Begriffen der Quantenmechanik weitgehend vertraut ist (vgl. für einen klassischen Zugang z. B. das Lehrbuch von FICK [52] oder für einen modernen Zugang das von BALLENTINE [8]). Ziel dieses einleitenden Abschnittes ist es, ihn mit den Grundgedanken der Quanteninformationstheorie und -kryptographie vertraut zu machen, die sich zum Teil von denen der übrigen Physik unterscheiden.

### 2.2.1 Beschreibung von Quantenzuständen

In der Quantenmechanik wird jedem System ein (separabler) Hilbertraum  $\mathcal{H}$  zugeordnet.<sup>3</sup> Die physikalisch meßbaren Größen (Observablen) werden durch selbstadjungierte Operatoren auf diesem Raum dargestellt; in der Quanteninformationstheorie nimmt man in der Regel an, daß umgekehrt auch jedem selbstadjungierten Operator eine Meßgröße zugeordnet werden kann.<sup>4</sup>

Der *Zustand* eines Quantensystems ist dadurch festgelegt, daß jeder Meßgröße ein Erwartungswert zugeordnet wird. Dies führt dazu, daß ein Zustand durch eine *Dichtematrix*<sup>5</sup>  $\rho$  auf dem Hilbertraum  $\mathcal{H}$  beschrieben wird, also durch eine positiv semidefinite Matrix, deren Spur Eins ergibt; die Menge der Dichtematrizen auf  $\mathcal{H}$  werde mit  $\mathcal{S}(\mathcal{H})$  (für engl. *state*) bezeichnet.

Nach dem Spektralsatz läßt sich eine Dichtematrix auf  $\mathcal{H} \cong \mathbb{C}^n$  stets in der Form  $\rho = \sum_{i=1}^n \lambda_i |i\rangle\langle i|$  für  $(\lambda_1, \dots, \lambda_n) \in \mathcal{W}_n$  schreiben. Verschwinden mit einer Ausnahme alle  $\lambda_i$ , so gilt  $\rho = |\Psi\rangle\langle\Psi|$  für einen *Zustandsvektor*  $|\Psi\rangle \in \mathcal{H}$ , und man spricht von einem *reinen Zustand*; die Dichtematrix ist also eine Projektion mit dem Rang 1. Die übrigen Zustände sind Konvexkombinationen der reinen Zustände, weshalb man sie als *gemischte Zustände* bezeichnet. Im folgenden werden die Begriffe Zustandsvektor, Zustand und Dichtematrix nicht streng unterschieden, wenn sich die Bedeutung aus dem Zusammenhang ergibt.

---

<sup>3</sup>Einige Sachverhalte, etwa die Existenz *vollständiger Systeme verallgemeinerter Eigenfunktionen*, lassen sich nur im allgemeineren Rahmen GELFANDscher *Raumtripel* erklären. Man vergleiche den *Satz von GELFAND UND KOSTJUTSCHENKO* [55] sowie die Monographie von GELFAND UND WILENKIN [56]; ein etwas allgemeinerer Fall findet sich bei GOULD [67].

<sup>4</sup>Die Konvention, Operatoren durch ein Dach zu kennzeichnen (etwa  $\hat{A}$ ) verwende ich nicht. Der Hauptgrund, Operatoren zu verwenden ist, daß nicht gleichzeitig meßbare (inkommensurable) physikalische Größen (inkompatible Observablen) durch nicht-kommutierende Operatoren dargestellt werden können.

<sup>5</sup>Ich unterscheide in der Regel nicht zwischen einem Operator als linearer Abbildung und seiner Darstellung als Matrix bzgl. einer festen Basis. Infolgedessen verwende ich die Begriffe *Dichtematrix* und *Dichteoperator* synonym. Andere Begriffe hierfür sind *statistische Matrix* bzw. *statistischer Operator*.

## 2.2.2 Zeitliche Entwicklung von Quantenzuständen

Die zeitliche Entwicklung eines Zustands ohne eine Messung wird durch die SCHRÖDINGER-Gleichung  $i\hbar(d/dt)|\Psi\rangle_t = H|\Psi\rangle_t$  bestimmt, die formal durch  $|\Psi\rangle_t = e^{-iHt/\hbar}|\Psi\rangle_{t=0}$  gelöst wird. Die Zeitentwicklung  $U(t) = e^{-iHt/\hbar}$  bildet also eine unitäre Einparameter-Gruppe. Umgekehrt existiert für jede solche Gruppe ein selbstadjungierter Erzeuger  $H$ , der zumindest theoretisch als ein HAMILTON-Operator betrachtet werden kann.

Tritt hingegen eine Messung auf, so ist die Entwicklung nicht mehr unitär. Der einer Meßgröße zugeordnete selbstadjungierte Operator  $A$  besitzt eine eindeutig bestimmte Spektralzerlegung  $A = \sum_{\mu} a_{\mu} P_{\mu}$ , wobei die  $a_{\mu}$  paarweise verschieden und die  $P_{\mu}$  Orthogonalprojektionen auf paarweise orthogonale Unterräume von  $\mathcal{H}$  sind. Der Zustand  $\rho$  geht dann mit der Wahrscheinlichkeit  $p_{\mu} := \text{Spur}(\rho P_{\mu})$  in den Zustand  $P_{\mu}\rho P_{\mu}$  über. Ignoriert man das Meßergebnis, so wird der Zustand zu  $\sum_{\mu} P_{\mu}\rho P_{\mu}$ . Da die Zahlenwerte  $a_{\mu}$  für den Zustandsübergang unbedeutend sind, nennt man die *Zerlegung der Eins* in orthogonale Projektionen  $(P_{\mu})_{\mu}$  eine VON-NEUMANN-Messung.

## 2.2.3 Quantenoperationen

Wird nur ein Untersystem eines größeren Systems betrachtet, wirkt aber der HAMILTON-Operator auf das Gesamtsystem, so ist die zeitliche Entwicklung auf dem Untersystem nicht zwingend unitär. Schränkt man sich auf das Untersystem ein, so kann die zeitliche Entwicklung (einschließlich Messungen) auf dem Untersystem durch eine *Quantenoperation* beschrieben werden; dies ist eine lineare, vollständig positive Abbildung  $\mathcal{E} : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$ .<sup>6</sup>

Eine alternative Beschreibung einer Quantenoperation ist die KRAUS- oder *Operatorsummendarstellung*; für KRAUS-Operatoren  $(E_{\mu})_{\mu=1}^n$ , die die Nebenbedingung  $\sum_{\mu=1}^n E_{\mu}^{\dagger} E_{\mu} = \mathbb{I}$  erfüllen, ist<sup>7</sup>

$$\rho \mapsto \rho' := \mathcal{E}(\rho) = \sum_{\mu=1}^n E_{\mu} \rho E_{\mu}^{\dagger}. \quad (2.2)$$

Setzt man  $A_{\mu} := E_{\mu}^{\dagger} E_{\mu} \geq 0$  fest, so gilt  $\sum_{\mu=1}^n A_{\mu} = \mathbb{I}$ . Umgekehrt definiert jede Zerlegung der Eins  $(A_{\mu})_{\mu=1}^n$  in positive Operatoren eine *verallgemeinerte Messung*; die Wahrscheinlichkeit für das Ergebnis  $\mu$  ist  $p_{\mu} = \text{Spur}(\rho A_{\mu})$ .<sup>8</sup>

<sup>6</sup>Für den Zusammenhang zwischen Quantenoperationen (stochastischen Abbildungen) und unitären Abbildungen auf größeren Systemen vgl. den STINESPRINGSchen Dilatationsatz; für dies und den Begriff der vollständig positiven Abbildung vgl. Kapitel 8.

<sup>7</sup>Die drei Beschreibungen von Quantenoperationen sind äquivalent; für Details vgl. zum Beispiel NIELSEN UND CHUANG [127] oder auch PRESKILL [137] oder KEYL [90].

<sup>8</sup>Man spricht auch von einem *POVM* für engl. *positive operator-valued measure*, also

## 2.2.4 Grundannahmen dieser Arbeit

In diesem Unterabschnitt werden einige Grundannahmen aufgelistet, die für die gesamte Arbeit gültig sind. Durch diese Annahmen werden gewisse Teile der Physik ausgeschlossen, die in den untersuchten Zusammenhängen nicht erforderlich sind oder die die Untersuchungen bedeutend erschweren würden. Das folgende werde daher vorausgesetzt:

- Alle Quantensysteme werden durch ihren Hilbertraum spezifiziert, der – wenn nicht gesondert auf das Gegenteil hingewiesen wird – stets als endlichdimensional angenommen werde.
- Es wird ausschließlich der mathematische Formalismus der nicht-relativistischen Quantenmechanik verwendet.
- Alle zusammengesetzten Quantensysteme bestehen aus unterscheidbaren Einzelsystemen („Teilchen“); die für Bosonen und Fermionen erforderliche Symmetrisierung entfällt.
- Als HAMILTON-Operator dient jeder selbstadjungierte Operator auf einem Hilbertraum, unabhängig davon, ob eine Realisierung in physikalischen Systemen bekannt ist oder nicht.

Trotz dieser Einschränkungen wird in dieser Arbeit Rücksicht auf die Realisierbarkeit der untersuchten Quantensysteme genommen; dies wird allerdings nur sehr allgemein abgehandelt (vgl. zum Beispiel verschränkungs-basierte und nicht-verschränkungs-basierte Protokolle in der Quantenkryptographie).

## 2.3 Qubit-Systeme

Der einfachste und zugleich nicht-triviale Fall eines Quantensystems ist ein *Qubit*; ein Qubit ist hierbei irgendein System, das durch einen zweidimensionalen Hilbertraum  $\mathcal{H} \cong \mathbb{C}^2$  beschrieben werden kann, gleich, um welches System es sich dabei handelt. Beispiele für Qubit-Systeme sind ein Spin-1/2-Teilchen, zum Beispiel ein Elektron, dessen Ortsfreiheitsgrad vernachlässigt wird, ein einzelnes Photon, dessen Polarisationszustände betrachtet werden oder auch ein Zwei-Niveau-System, welches Teil eines größeren Systems ist.

Im folgenden wird nur der zweidimensionale Hilbertraum betrachtet, der physikalisch in irgendeiner Weise mit dem System identifiziert werden kann.

---

einem Maß, dessen Werte positive Operatoren auf  $\mathcal{H}$  sind. Dies ist eine Verallgemeinerung des Spektralmaßes, das man einer VON-NEUMANN-Messung zuordnen kann.

### 2.3.1 Die Pauli-Matrizen

In der Quantenmechanik finden zur Beschreibung von Qubit-Systemen die PAULI-Matrizen Anwendung; diese Matrizen sind spurfrei, hermitesch, unitär mit den Eigenwerten  $+1$  und  $-1$  und werden durch

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{und} \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{C}^{2 \times 2} \quad (2.3)$$

definiert. Sehr häufig schreibt man statt  $\sigma_x$ ,  $\sigma_y$  und  $\sigma_z$  auch  $\sigma_1$ ,  $\sigma_2$  und  $\sigma_3$ , in der Quanteninformationstheorie oft auch  $X$ ,  $Y$  und  $Z$ . Die PAULI-Matrizen erfüllen die Rechenregeln  $\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i \sum_k \varepsilon_{ijk} \sigma_k$ , woraus die Kommutatorrelationen  $[\sigma_i, \sigma_j] = 2i \sum_k \varepsilon_{ijk} \sigma_k$  folgen. Für Qutrit-Systeme ( $d = 3$ ) gibt es die GELL-MANN-Matrizen, die zum Teil ähnliche Eigenschaften besitzen.

Man faßt die PAULI-Matrizen oft in einer Art Vektor  $\vec{\sigma} := (\sigma_x, \sigma_y, \sigma_z)^t$  zusammen und schreibt mit einem Vektor  $\vec{s} = (s_x, s_y, s_z)^t \in \mathbb{C}^3$  in Anlehnung an das Skalarprodukt  $\vec{\sigma} \cdot \vec{s} := s_x \sigma_x + s_y \sigma_y + s_z \sigma_z$ . Es gilt dann die Gleichung  $(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b} + i \vec{\sigma} \cdot (\vec{a} \times \vec{b})$ .

Zusammen mit der zweidimensionalen Einheitsmatrix  $\mathbb{1}$  bilden die PAULI-Matrizen eine Basis des Raumes  $\mathbb{C}^{2 \times 2}$ , man kann also jede komplexe  $2 \times 2$ -Matrix in der Form  $s_0 \mathbb{1} + \vec{\sigma} \cdot \vec{s}$  schreiben. Das charakteristische Polynom einer solchen Matrix ergibt sich zu  $\chi = \lambda^2 - 2s_0 \lambda + (s_0^2 - \vec{s}^2)$ , ihre Eigenwerte sind also  $\lambda_{1,2} = s_0 \pm |\vec{s}|$ . Eine Matrix ist genau dann hermitesch, wenn die Koeffizienten in dieser Basis reell sind.

### 2.3.2 Bloch-Kugel und Bloch-Vektor

Mithilfe der PAULI-Matrizen kann der Zustand eines einzelnen Qubits durch einen einzigen BLOCH-Vektor<sup>9</sup>  $\vec{s} = (s_x, s_y, s_z)^t$  mit  $\|\vec{s}\|_2 \leq 1$  beschrieben werden; die zugehörige Dichtematrix ist

$$\rho = \frac{1}{2} \left[ \mathbb{1} + \sum_{k \in \{x, y, z\}} \sigma_k \cdot s_k \right] = \frac{1}{2} \left[ \mathbb{1} + \vec{\sigma} \cdot \vec{s} \right]. \quad (2.4)$$

Man spricht hierbei von der BLOCH-Kugel; ihre Oberfläche heißt BLOCH-Sphäre. Die Eigenwerte des Zustands  $\rho$  sind  $1/2 \pm |\vec{s}|/2$ , ein Zustand ist somit genau dann rein, wenn sein BLOCH-Vektor  $\vec{s}$  normiert ist, also auf der Kugeloberfläche liegt.<sup>10</sup>

<sup>9</sup>Die Begriffe BLOCH-Vektor und BLOCH-Kugel gehen auf die Beschreibung des Elektronenspins und seines magnetischen Momentes zurück; die gleiche Konstruktion zur Beschreibung der Polarisationszustände eines einzelnen Photons wird in der Quantenoptik auch als POINCARÉ-Vektor bezeichnet.

<sup>10</sup>In einer hiervon zu unterscheidenden Konstruktion schreibt man die reinen Zustände in der Form  $(1, z)^t / (1 + |z|^2)^{1/2} \in \mathbb{C}^2$ , wobei der Parameter  $z$  als Element der RIEMANNschen Zahlenkugel  $\mathbb{C} \cup \{\infty\}$  aufgefaßt wird.

## 2.4 Verschränkungstheorie

Der in der Quanteninformationstheorie wohl wichtigste Begriff ist der der *Verschränkung*.

### Definition 2.2 (Separable und verschränkte Zustände)

Es sei  $\mathcal{H} := \mathcal{H}_A \otimes \mathcal{H}_B$  das Tensorprodukt der Hilberträume  $\mathcal{H}_A$  und  $\mathcal{H}_B$ . Läßt sich ein Zustandsvektor  $|\Psi\rangle_{AB} \in \mathcal{H}$  für ein  $|\Psi_1\rangle_A \in \mathcal{H}_A$  und ein  $|\Psi_2\rangle_B \in \mathcal{H}_B$  in der Form  $|\Psi\rangle_{AB} = |\Psi_1\rangle_A \otimes |\Psi_2\rangle_B$  schreiben, so nennt man ihn separabel, andernfalls verschränkt. Die konvexe Hülle separabler reiner Zustände bildet die Menge der separablen Zustände, die übrigen nennt man verschränkt.

Ein Zustand ist also genau dann separabel, wenn er für geeignete Zustandsvektoren  $|a_i\rangle_A \in \mathcal{H}_A$  und  $|b_i\rangle_B \in \mathcal{H}_B$ ,  $i \in \{1, \dots, n\}$ , und eine Wahrscheinlichkeitsverteilung  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  in der Form

$$\rho = \sum_{i=1}^n p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i| \quad (2.5)$$

geschrieben werden kann.<sup>11</sup> In der Quanteninformationstheorie ist üblicherweise die Verschränkung zweier räumlich entfernter Teilchen bedeutsam. Die Verschränkung bewirkt dann Korrelationen zwischen lokalen Messungen der Einzelteilchen; vgl. das EPR-Paradoxon.

### 2.4.1 Bestimmung der Verschränkung

Ein reiner Zustand  $|\Psi\rangle_{AB}$  ist genau dann separabel, wenn die reduzierte Dichtematrix  $\rho_A$  (oder  $\rho_B$ ) rein ist; dies kann z. B. durch die VON-NEUMANN-Entropie überprüft werden. Im Falle gemischter Zustände ist ein allgemein anwendbares Kriterium nicht bekannt, insbesondere, da die Menge der separablen Zustände zwar konvex ist, aber kein Polytop bildet, sie also nicht als konvexe Hülle einer endlichen Anzahl von Zuständen darstellbar ist.

Man kann zeigen, daß ein Zustand genau dann verschränkt ist, wenn es eine positive (aber nicht vollständig positive) Abbildung  $\Lambda : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$  derart gibt, daß  $(\mathbb{1}_A \otimes \Lambda_B)(\rho_{AB})$  nicht positiv semidefinit ist. Ein Beispiel für ein solches  $\Lambda$  ist die Transposition auf einem Tensorfaktor. Ist auf einem der Hilberträume  $\mathbb{C}^2 \otimes \mathbb{C}^2$ ,  $\mathbb{C}^2 \otimes \mathbb{C}^3$  oder  $\mathbb{C}^3 \otimes \mathbb{C}^2$  diese Teiltransponierte positiv, so ist der Zustand schon separabel; diese Aussage wird als das PERES-HORODECKI-Kriterium [133, 80] bezeichnet.

---

<sup>11</sup>Der Begriff der Verschränkung kann auch auf Systeme mit mehr als zwei Untersystemen verallgemeinert werden. Zum einen kann man die Systeme in zwei Gruppen einteilen, so daß die ursprüngliche Definition wieder angewendet werden kann (Biseparabilität), zum anderen kann man Vektoren der Form  $|\Psi_1\rangle \otimes \dots \otimes |\Psi_n\rangle$  vollständig separabel nennen und ihre konvexe Hülle betrachten. Auch weitere Definitionen sind gebräuchlich.

Andere Beispiele für Verschränkungskriterien sind das *Reduktionskriterium* (siehe Lemma 2.3) oder die Verwendung sogenannter *Verschränkungszeugen* (engl. *entanglement witnesses*); zu letzterem vgl. den Satz von HAHN und BANACH für konvexe Mengen.

## 2.4.2 Quantifizierung der Verschränkung

Es ist möglich, den Grad der Verschränkung quantitativ zu erfassen; dies geschieht durch *Verschränkungsmaße*. Ein Verschränkungsmaß ist dabei eine Funktion  $E : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathbb{R}_0^+$ , die gewissen Axiomen genügen muß. Welche Bedingungen im einzelnen gefordert werden, variiert in der Literatur. In der Regel wird gefordert, daß für separable Zustände  $E(\rho) = 0$  gilt und daß die Verschränkung unter *LOCC-Operationen* (engl. *local operations and classical communications*, lokale Operationen und klassische Kommunikation) nicht wächst. Üblicherweise ist  $E$  unter lokal-unitären Transformationen der Form  $U_A \otimes U_B$  invariant; für Details vgl. z. B. KEYL [90] oder CHRISTANDL [35].

Es zeigt sich, daß für reine Zustände die VON-NEUMANN-Entropie der reduzierten Dichtematrix das im wesentlichen einzige Verschränkungsmaß ist. Eines der wichtigsten Verschränkungsmaße ist die zur Bildung eines Zustands erforderliche Verschränkung (engl. *entanglement of formation*): diese ist über  $E(\rho) := \min \sum_i p_i S(\rho_{A;i})$  definiert, wobei über alle Zerlegungen der der Form  $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$  minimiert wird und  $\rho_{A;i} := \text{Spur}_B |\Psi_i\rangle\langle\Psi_i|$  ist.

## 2.4.3 Verschränkungsreinigung

Ein Ziel der Quanteninformationsverarbeitung ist es, maximal verschränkte Zustände herzustellen.<sup>12</sup> Liegt nun eine große Anzahl schwach verschränkter Zustände vor, so sollen *Protokolle zur Verschränkungsreinigung* aus diesen stärker verschränkte Zustände gewinnen; vgl. BENNETT u. a. [16, 18].

Üblicherweise betrachtet man zwei entfernte Parteien, die sich für beliebig großes  $n \in \mathbb{N}$  einen Anfangszustand  $\rho^{\otimes n}$  teilen. Ziel ist es, durch LOCC-Operationen asymptotisch  $m \leq n$  maximal verschränkte Zustände zu erzeugen; je nachdem, ob man nur Kommunikation von einer Partei zur anderen oder auch umgekehrt zuläßt, spricht man von Protokollen mit *Einweg-* oder *Zweiweg-Kommunikation* und nennt das im Grenzfall  $n \rightarrow \infty$  bestmögliche  $m/n$  die *destillierbare Verschränkung*  $D_1(\rho)$  bzw.  $D_2(\rho)$  eines Zustands. Es ist  $D_1(\rho) \leq D_2(\rho) \leq E(\rho)$ ; ist  $D_2(\rho) = 0$ , der Zustand aber verschränkt, so nennt man ihn nennt man *gebunden verschränkt*, andernfalls *frei verschränkt*.

---

<sup>12</sup>Ein reiner Zustand  $|\Psi\rangle$  ist maximal verschränkt, wenn die reduzierten Dichtematrizen maximal gemischt, also Vielfache der Einheitsmatrix sind. Dies ist genau dann der Fall, wenn in der SCHMIDT-Zerlegung alle Koeffizienten gleich sind.

## 2.5 Verallgemeinerte Bell-Zustände

In diesem Abschnitt werden die sogenannten BELL-Zustände für Qubits auf höherdimensionale Systeme verallgemeinert. Im folgenden werden daher Quantensysteme einer fest gewählten, endlichen Dimension  $d \in \mathbb{N} \setminus \{1\}$  betrachtet; aus algebraischen Gründen erfolgen zum Teil Einschränkungen auf Dimensionen, die eine Primzahl oder die Potenz einer Primzahl sind.

### 2.5.1 Verallgemeinerte Bell-Basen

Für Zwei-Qubit-Systeme, also Systeme mit dem Hilbertraum  $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ , ist eine Basis maximal verschränkter Zustände, die BELL-Basis, definiert, die leicht auf  $d$ -dimensionale Systeme verallgemeinert werden kann. Eine Möglichkeit hierzu besteht darin, für Werte  $l, m \in \mathbb{Z}_d$  die Zustände

$$|\Psi_{lm}\rangle := d^{-1/2} \sum_{k=0}^{d-1} z_d^{lk} |k\rangle |k \ominus m\rangle \quad \text{mit } z_d = e^{2\pi i/d} \quad (2.6)$$

zu betrachten (vgl. z. B. ALBER u. a. [2]); die Gesamtheit dieser Zustände  $\{|\Psi_{lm}\rangle | l, m \in \mathbb{Z}_d\}$  bildet eine Orthonormalbasis des Hilbertraums  $\mathbb{C}^d \otimes \mathbb{C}^d$  aus maximal verschränkten Zuständen und wird als *verallgemeinerte BELL-Basis* bezeichnet.<sup>13</sup>

In den folgenden Überlegungen wird es nützlich sein, die Werte von  $l$  und  $m$  als eine Art relative Phase bzw. einen relativen Dittwert zwischen den beiden Teilsystemen zu interpretieren. Gelegentlich wird die Kurzbezeichnung  $(l, m) := |\Psi_{lm}\rangle \langle \Psi_{lm}|$  mit  $l, m \in \mathbb{Z}_d$  verwendet.

Ist die Dimension eine Primzahlpotenz, läßt sie sich also (eindeutig) in der Form  $d = p^n$  für eine Primzahl  $p$  und einen Exponenten  $n \in \mathbb{N}$  schreiben, so wird vielfach noch eine andere Definition der verallgemeinerten BELL-Basis verwendet. Die Ein-Qudit-Basisvektoren werden hierfür mit den Elementen des endlichen Körpers  $\mathbb{F}_{p^n}$  bezeichnet, und man setzt

$$|\Psi'_{lm}\rangle := d^{-1/2} \sum_{k \in \mathbb{F}_{p^n}} z_p^{\text{Spur}(lk)} |k\rangle |k \ominus m\rangle, \quad (2.7)$$

wobei die Spur hier im Sinne der Theorie der Körpererweiterungen zu verstehen ist (vgl. Unterabschnitt B.3.4). Diese Konstruktion wird in dieser Arbeit nur in Einzelfällen verwendet.

---

<sup>13</sup>Die BELL-Basis (im engeren Sinne) entsteht hieraus im Fall  $d = 2$ ; man setzt in der Regel  $|\Phi^+\rangle := |\Psi_{00}\rangle$ ,  $|\Phi^-\rangle := |\Psi_{10}\rangle$ ,  $|\Psi^+\rangle := |\Psi_{01}\rangle$  und  $|\Psi^-\rangle := |\Psi_{11}\rangle$ . Der Zustand  $|\Psi^-\rangle$  beschreibt das *Singulett*, beim Zustand  $|\Psi^+\rangle$  handelt es sich um den Triplet-0-Zustand.

Genaugenommen müßte man im folgenden stets von verallgemeinerter BELL-Basis, verallgemeinert-bell-diagonalen Zuständen usw. sprechen. Da diese Dissertation sich jedoch zum Großteil mit derartigen verallgemeinerten Zuständen beschäftigt, wird dies der Kürze wegen oft unterlassen werden.

## 2.5.2 Mischungen verallgemeinerter Bell-Zustände

Im ersten Teil der Arbeit werden sehr häufig klassische Gemische verallgemeinerter BELL-Zustände betrachtet; mathematisch gesprochen sind dies die Konvexkombinationen verallgemeinerter BELL-Zustände, also Zustände der Form

$$\rho = \sum_{l,m=0}^{d-1} A_{lm} |\Psi_{lm}\rangle \langle \Psi_{lm}| \quad (2.8)$$

mit einem System nicht-negativer Koeffizienten  $(A_{lm})_{l,m=0}^{d-1}$ , das die Normierungsbedingung  $\sum_{l,m=0}^{d-1} A_{lm} = 1$  erfüllt; man kann also  $(A_{lm})_{l,m=0}^{d-1} \in \mathcal{W}_{d \times d}$  schreiben. Die Dichtematrix dieser Zustände ist diagonal in der BELL-Basis, weshalb man von *bell-diagonalen Zuständen* spricht; die Menge all dieser Zustände werde von nun an mit  $\mathcal{S}_{\text{bd}}^{(d)}$  bezeichnet, für  $d = 2$  auch mit  $\mathcal{S}_{\text{bd}}$ .

Die bell-diagonalen Zustände können in naheliegender Weise mit ihrer Koeffizientenmatrix<sup>14</sup> identifiziert werden, was in der gesamten Dissertation erfolgen wird; man schreibt also  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  oder ausführlich

$$\rho = \begin{pmatrix} A_{00} & A_{01} & \dots & A_{0,d-1} \\ A_{10} & A_{11} & \dots & A_{1,d-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{d-1,0} & A_{d-1,1} & \dots & A_{d-1,d-1} \end{pmatrix} \begin{array}{l} \rightarrow A_{0*} \\ \rightarrow A_{1*} \\ \vdots \\ \rightarrow A_{d-1*} \end{array} \quad (2.9)$$

$$\begin{array}{cccc} \downarrow & \downarrow & & \downarrow \\ A_{*0} & A_{*1} & \dots & A_{*d-1} \end{array}$$

Im weiteren Verlauf dieser Arbeit wird es darum gehen, einen vorgegebenen Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  in einen Zustand  $\rho' = (A'_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  zu überführen. Als Ziel wählt man sich einen beliebigen Referenzzustand der BELL-Basis aus, gewöhnlich den Zustand  $|\Psi_{00}\rangle \langle \Psi_{00}|$ , dessen Koeffizienten  $A_{lm} = \delta_{l0} \delta_{m0}$  sind.<sup>15</sup> Dies ermöglicht es, die Werte  $l, m \in \mathbb{Z}_d$  als Phasen- bzw. Ditfehler zu interpretieren; ein Fehler liegt also dann vor, wenn  $l$  oder  $m$  von Null verschieden sind.

Ein Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  definiert über seine Koeffizienten eine gemeinsame Wahrscheinlichkeitsverteilung der Dit- und Phasenfehler. Die zugehörigen Randverteilungen sind die Ditfehlerverteilung  $(A_{*m})_{m=0}^{d-1} \in \mathcal{W}_d$  und die Phasenfehlerverteilung  $(A_{l*})_{l=0}^{d-1} \in \mathcal{W}_d$ , die durch

$$A_{*m} := \sum_{l=0}^{d-1} A_{lm} \quad \text{bzw.} \quad A_{l*} := \sum_{m=0}^{d-1} A_{lm} \quad (2.10)$$

definiert werden; der Stern deutet also die Summierung über die betreffenden Elemente an.

---

<sup>14</sup>Diese Matrix ist nicht im Sinne einer linearen Abbildung zu verstehen!

<sup>15</sup>In der Quanteninformationstheorie nennt man diesen Zustand oft auch Singulett, auch wenn dies im strengen Sinne falsch ist, da er nicht  $U \otimes U$ - sondern  $U \otimes U^*$ -invariant ist.

### 2.5.3 Verallgemeinert-isotrope Zustände

Eine einfache und wichtige Unterklasse der bell-diagonalen Zustände sind die isotropen Zustände, das heißt, diejenigen Zustände, die unter allen Transformationen der Form  $U \otimes U^*$  invariant sind<sup>16</sup>; man kann zeigen, daß dies diejenigen bell-diagonalen Zustände sind, bei denen  $A_{lm} = (1 - A_{00})/(d^2 - 1)$  für alle  $(l, m) \neq (0, 0)$  ist. Als *verallgemeinert-isotrope Zustände* sollen in dieser Arbeit die Zustände mit der Koeffizientenmatrix<sup>17</sup>

$$\rho = (\alpha, \beta, \gamma, \delta) := \begin{pmatrix} \alpha & \gamma & \dots & \gamma \\ \beta & \delta & \dots & \delta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \delta & \dots & \delta \end{pmatrix} \in \mathcal{S}_{\text{bd}}^{(d)} \quad (2.11)$$

bezeichnet werden. Man erkennt unmittelbar, daß im Falle von Qubits alle bell-diagonalen Zustände diese Form haben. Die isotropen Zustände sind also die Zustände, für die  $\beta = \gamma = \delta$  gilt. NIKOLOPOULOS u. a. [130] verwenden den Begriff abweichend für Zustände, von denen nur  $\beta = \gamma$  gefordert wird.

#### Lemma 2.3 (Bell-diagonale Zustände und Verschränkung)

Ein Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ist zumindest dann destillierbar und daher insbesondere verschränkt, wenn  $\max \{A_{lm} \mid l, m \in \mathbb{Z}_d\} > 1/d$  ist.

*Beweis:* Nach dem Reduktionskriterium der HORODECKIS [79] ist ein Zustand zumindest dann destillierbar, wenn die Ungleichung  $\rho_A \otimes \mathbb{1}_d - \rho_{AB} \geq 0$  verletzt ist. Im betrachteten Fall ist  $\rho_A = d^{-1} \mathbb{1}_d$ , da die Elemente der BELL-Basis maximal verschränkt sind; also genügt  $d^{-1} \mathbb{1}_{d^2} - \rho_{AB} \not\geq 0$ , q. e. d.

Im Vorgriff auf die noch zu behandelnden Korrekturschritte soll schon an dieser Stelle ein Protokoll zur Verschränkungsreinigung angegeben werden:

1. Führe eine lokal-unitäre Transformation durch, so daß  $A_{00} > 1/d$  wird, z. B.  $\mathbb{1} \otimes X^{-m} Z^{-l}$  (vgl. Unterabschnitt 2.7.3), wenn  $A_{lm} > 1/d$  ist.
2. Mische den Zustand über die Gruppe  $\{U \otimes U^* \mid U \in \mathbb{C}^{d \times d}, U^{-1} = U^\dagger\}$  (Integration über das HAARSche Maß), so daß der Zustand isotrop wird.
3. Wende einen  $B_2^{(d)}$ -Schritt (siehe Unterabschnitt 4.2.1) an.
4. Wiederhole die Schritte 2 und 3 solange, bis der Zustand hinreichend nahe am Idealzustand  $|\Psi_{00}\rangle\langle\Psi_{00}|$  liegt.

---

<sup>16</sup>Im Gegensatz dazu sind die  $U \otimes U$ -invarianten Zustände die WERNER-Zustände.

<sup>17</sup>In früheren Arbeiten verwendete ich die Notation  $(a, b, c, d)$  statt  $(\alpha, \beta, \gamma, \delta)$ , was aber mit der Bezeichnung  $d$  für die Dimension kollidiert.

Im ersten Schritt wird  $A_{00} > 1/d$  erzwungen; der zweite Schritt erzeugt aus diesem Zustand den verallgemeinert-isotropen Zustand  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{\text{bd}}^{(d)}$  mit unverändertem  $A_{00} =: \alpha$  und  $\beta := \gamma := \delta := (1 - \alpha)/(d^2 - 1)$ . Der dritte Schritt gefolgt vom zweiten Schritt überführt den isotropen Zustand in einen anderen, ebenfalls isotropen Zustand, der durch ein  $\alpha'$  beschrieben werden kann. Aus der Gleichung (4.33) folgt für diese Abbildung

$$\alpha \mapsto \alpha' = f(\alpha) := \frac{1 + \alpha[\alpha d(d^2 + d - 1) - 2]}{d[1 + d + \alpha(\alpha d^2 - 2)] - 1}, \quad (2.12)$$

und man ist am Verhalten  $\lim_{n \rightarrow \infty} f^n(\alpha)$  interessiert. Für einen Fixpunkt fordert man  $\alpha' = \alpha$ , was auf eine Gleichung dritten Grades führt, die mittels der *Cardanischen Formel* [57] analytisch gelöst werden kann. Alternativ kann man die drei Fixpunkte  $d^{-2}$ ,  $d^{-1}$  und 1 auch einfach erraten.

Man kann nun zeigen, daß der Fixpunkt bei  $d^{-1}$  instabil und die beiden anderen stabil sind. Somit konvergiert das Verfahren für  $n \rightarrow \infty$  und  $\alpha > 1/d$  gegen den gewünschten Zustand.<sup>18</sup>

## 2.6 Komplementäre Basen

Der Begriff der komplementären Basen wird in der Quanteninformationstheorie vielfach gebraucht, insbesondere auch in der Quantenkryptographie; vgl. z. B. KLAPPENECKER UND RÖTTELER [94] und die Originalarbeiten von WOOTTERS UND FIELDS [186] und von BANDOPADYAY u. a. [9].<sup>19</sup>

### Definition 2.4 (Komplementäre Basen)

Zwei Basen  $B_1 = \{|a_0\rangle, \dots, |a_{d-1}\rangle\}$  und  $B_2 = \{|b_0\rangle, \dots, |b_{d-1}\rangle\}$  des Hilbert-raums  $\mathcal{H} = \mathbb{C}^d$  nennt man komplementär, wenn für alle  $i, j \in \{0, \dots, d-1\}$  die Gleichung  $|\langle a_i | b_j \rangle|^2 = 1/d$  erfüllt ist.

Im Englischen verwendet man häufiger den Begriff *mutually unbiased bases* (abgekürzt MUB). Beispiele paarweise komplementärer Basen sind die  $z$ - und die  $x$ -Basis in Qubit-Systemen. Allgemeiner sind zwei durch ihre normierten BLOCH-Vektoren definierten Basen eines Qubit-Systems genau dann komplementär, wenn diese BLOCH-Vektoren senkrecht aufeinander stehen, und man

---

<sup>18</sup>Für bell-diagonale Zwei-Qubit-Zustände  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{\text{bd}}$  gibt es auch das DEUTSCH-Protokoll [45, 116]. Hier wird statt Nr. 2 für  $D = (\mathbb{1}_2 - i\sigma_x)/\sqrt{2}$  die Transformation  $D \otimes D^*$  angewendet, die einen Zustand  $(l, m)$  nach  $(l, m \oplus l)$  überführt.

<sup>19</sup>Man beachte, daß der Begriff zunächst nur für endlichdimensionale Quantensysteme definiert ist; in gewissem Sinne sind aber z. B. Orts- und Impulsbasis eines freien Teilchens komplementär.

erkennt hieran, daß man drei – aber nicht mehr – paarweise komplementäre Basen wählen kann. Eine wichtige, in wesentlichen Teilen ungelöste Fragestellung in der Quanteninformationstheorie ist die Bestimmung der Anzahl paarweise komplementärer Basen in einem  $d$ -dimensionalen Hilbertraum. Für solche Systeme paarweise komplementärer Basen gilt der folgende Satz.

**Satz 2.5 (Systeme komplementärer Basen)**

*Ein System komplementärer Basen des Hilbertraums  $\mathcal{H} = \mathbb{C}^d$  besteht aus nicht mehr als  $d + 1$  Elementen. Ist  $d$  eine Primzahlpotenz, gibt es also eine Primzahl  $p \in \mathbb{N}$  und einen Exponenten  $n \in \mathbb{N}$ , so daß  $d = p^n$  gilt, so gibt es Systeme komplementärer Basen, die aus genau  $d + 1$  Elementen bestehen.*

Ein Beweis kann hier nicht angegeben werden. Die Tatsache, daß die Lösung im Falle von Primzahlpotenzdimensionen bekannt ist, rührt daher, daß genau in diesen Fällen endliche Körper vorliegen (vgl. Satz B.11). In diesem Fall sind mehrere konstruktive Lösungen bekannt.<sup>20</sup>

Der Fall, daß  $d$  keine Primzahlpotenz ist, ist ungelöst; so ist selbst für den Fall  $d = 6 = 2 \cdot 3$  kein System komplementärer Basen mit mehr als drei Elementen bekannt. Eine untere Schranke für die Maximalzahl paarweise komplementärer Basen bestimmt man, indem man die Primfaktorzerlegung  $d = p_1^{k_1} \cdot \dots \cdot p_n^{k_n}$  betrachtet; es sind dann Systeme komplementärer Basen mit  $\min \{p_i^{k_i} + 1 \mid i \in \{1, \dots, n\}\}$  Elementen konstruierbar, indem man den Hilbertraum  $\mathbb{C}^d$  formal mit dem Tensorprodukt  $\mathbb{C}^{(p_1^{k_1})} \otimes \dots \otimes \mathbb{C}^{(p_n^{k_n})}$  identifiziert, für jeden Tensorfaktor Systeme komplementärer Basen bestimmt und deren Tensorprodukt bildet.

## 2.7 Einige unitäre Transformationen

Auf Qudit-Systemen gibt es eine Reihe von unitären Transformationen, die in der Quanteninformationstheorie von Bedeutung sind. Neben den PAULI-Matrizen, die auf Qudit-Systeme verallgemeinert werden können, gehören hierzu auch eine Reihe von Transformationen, die in der allgemeinen Quantenmechanik eher nachrangige Bedeutung besitzen.

---

<sup>20</sup>Schreibt man die Vektoren einer Basis nebeneinander, so erhält man eine quadratische Matrix. Wählt man für die erste Basis die Standardbasis, so ist eine weitere Basis zu dieser komplementär, wenn die zugehörige Matrix eine HADAMARD-Matrix ist, das heißt, wenn alle ihre Einträge den Betrag  $1/\sqrt{d}$  besitzen.

### 2.7.1 Verallgemeinerte XOR-Transformationen

Eine der wichtigsten Transformationen auf zwei Qubits ist die CNOT-Transformation (engl. *controlled NOT*, gesteuerte Inversion), die als quantenmechanisches Analogon zum klassischen XOR (exklusives Oder) dient. Sie kann für Qudits zu einer *verallgemeinerten XOR-Transformation* GXOR (engl. *generalised XOR*) erweitert werden; dies erfolgt durch<sup>21</sup>

$$|k\rangle|l\rangle \mapsto \text{GXOR } |k\rangle|l\rangle := |k\rangle|k \oplus l\rangle. \quad (2.13)$$

Man nennt das erste Qudit die *Steuerung* (engl. *control*), das zweite das *Ziel* (engl. *target*). Betrachtet man zwei reine BELL-Zustände  $|\Psi_1\rangle_{AB}$  und  $|\Psi_2\rangle_{CD}$  und wendet je eine GXOR-Transformation von  $A$  nach  $C$  und von  $B$  nach  $D$  an, so spricht man von einer *bilateralen GXOR-Transformation*. In der Notation  $(l, m) := |\Psi_{lm}\rangle\langle\Psi_{lm}|$  für BELL-Zustände bewirkt sie

$$\text{GBXOR}[(l_1, m_1) \otimes (l_2, m_2)] = (l_1 \oplus l_2, m_1) \otimes (l_2, m_1 \oplus m_2). \quad (2.14)$$

Werden zwei Qudit-Paare  $QP_1$  und  $QP_2$  betrachtet, so werde dieser Prozeß als  $\text{GBXOR}(QP_1, QP_2)$  geschrieben.

### 2.7.2 Die diskrete Fouriertransformation

Die sogenannte HADAMARD-Transformation  $H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} / \sqrt{2}$  vermittelt auf Qubit-Systemen eine Transformation, die die  $z$ - und die  $x$ -Basis ineinander überführt und die in Zwei-Basis-Protokollen wie dem BB84-Protokoll verwendet wird. Für bell-diagonale Zwei-Qubit-Zustände wirkt sie als Abbildung  $(l, m) \mapsto (m, l)$ . Für qudit-basierte Zwei-Basis-Protokolle wird nun eine Art verallgemeinerte HADAMARD-Transformation benötigt, die ähnliche Eigenschaften besitzt.

Es zeigt sich, daß die diskrete Fouriertransformation  $\mathcal{F} : \mathbb{C}^d \rightarrow \mathbb{C}^d$  die gewünschten Eigenschaften mehr oder weniger besitzt. Sie ist durch

$$\mathcal{F} := d^{-1/2} \sum_{i,j=0}^{d-1} z_d^{ij} |i\rangle\langle j| \quad (2.15)$$

definiert. Es ist  $\mathcal{F}^2|i\rangle = |\ominus i\rangle$ , woraus  $\text{GXOR}(\mathbb{1} \otimes \mathcal{F}^2)|k\rangle|l\rangle = |k\rangle|k \oplus l\rangle$  folgt. Nach Anhang A.2 gilt  $(\mathcal{F} \otimes \mathcal{F}^*)|\Psi_{lm}\rangle = z^{lm}|\Psi_{m,d-l}\rangle$ , und die Anwendung von  $\mathcal{F} \otimes \mathcal{F}^*$  auf einen bell-diagonalen Zustand liefert bis auf Phasen eine Permutation der BELL-Basis: es gilt  $(l, m) \mapsto (m, \ominus l)$ .

---

<sup>21</sup>Allgemeiner könnte man  $|k\rangle|l\rangle \mapsto \text{GXOR } |k\rangle|l\rangle := |k\rangle|f_1 k \oplus f_2 l\rangle$  für  $f_1, f_2 \in \{-1, +1\}$  fordern. Eine bilaterale GXOR-Transformation auf zwei Qudit-Paaren soll eine Transformation der Form  $(l_1, m_1) \otimes (l_2, m_2) \mapsto (g_1 l_1 \oplus g_2 l_2, g_0 m_1) \otimes (h_0 l_2, h_1 m_1 \oplus h_2 m_2)$  bewirken, wobei  $g_0, g_1, g_2, h_0, h_1, h_2 \in \{-1, +1\}$  gewählt werden können. Es folgen die Bedingungen  $g_0 = g_1 = 1$ ,  $g_2 = -f_1 f_2$ ,  $h_0 = h_2 = f_2$  und  $h_1 = f_1$ . Die Werte  $f_1$  und  $f_2$  sind beliebig und wurden hier auf  $f_1 := 1$  und  $f_2 := -1$  gesetzt.

### 2.7.3 Verallgemeinerte Pauli-Matrizen

Ist  $p \neq 2$  eine Primzahl<sup>22</sup>, so läßt sich eine Basis des Hilbertraums  $\mathcal{H} = \mathbb{C}^p$  mit den Elementen des endlichen Körpers  $\mathbb{F}_p$  bezeichnen. Auf diesem Hilbertraum kann man verallgemeinerte Pauli-Matrizen definieren: für  $a, b \in \mathbb{F}_p$  setzt man<sup>23</sup>

$$X^a := \sum_{k \in \mathbb{F}_p} |k - a\rangle\langle k| \quad \text{und} \quad Z^b := \sum_{k \in \mathbb{F}_p} z_p^{kb} |k\rangle\langle k|; \quad (2.16)$$

die Abbildungen  $a \mapsto X^a$  und  $b \mapsto Z^b$  sind dann treue Darstellungen der additiven Gruppe des endlichen Körpers  $\mathbb{F}_p$ .<sup>24</sup> Man berechnet<sup>25</sup>  $X^a Z^b = z_p^{ab} Z^b X^a$  und erhält in der Abbildung  $\mathbb{F}_p \times \mathbb{F}_p \rightarrow \mathbb{C}^{p \times p}$ ,  $(a, b) \mapsto E_{ab} := X^a Z^b$  eine Strahldarstellung der additiven Gruppe  $\mathbb{F}_p \times \mathbb{F}_p$ . Die Menge  $\{E_{ab} \mid a, b \in \mathbb{F}_p\}$  bildet eine Basis der Operatoren auf  $\mathbb{C}^p$ , und mit  $k := a \cdot d - b \cdot c$  gilt  $E_{ab} E_{cd} = E_{cd} E_{ab} z_p^k$ . Da  $p$  eine Primzahl ist, kann man  $X^a$  als die  $a$ -te Potenz von  $X := X^1$  auffassen; sinngemäß ist  $Z^b$  die  $b$ -te Potenz von  $Z := Z^1$ .<sup>26</sup>

Ein PAULI-Operator auf dem Tensorprodukt von  $n \in \mathbb{N}$  Qudits, also auf dem Hilbertraum  $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$  ist das Tensorprodukt von PAULI-Matrizen auf den einzelnen Qudits, besitzt also die Form  $E_{a_1 b_1} \otimes \dots \otimes E_{a_n b_n}$ ; sind  $a_1 = \dots = a_n = 0$ , so spricht man von einem Typ- $Z$ -, für  $b_1 = \dots = b_n = 0$  von einem Typ- $X$ -Operator.

---

<sup>22</sup>Der Fall  $p = 2$  – und allgemeiner der Fall eines Körpers der Charakteristik 2 – unterscheidet sich etwas von den übrigen Fällen. Unter anderem treten an die Stelle der zweiten Einheitswurzeln die vierten Einheitswurzeln.

<sup>23</sup>Im Grunde genommen kann man jedem  $X^a$  und jedem  $Z^b$  eine beliebige  $p$ -te Einheitswurzel als Phasenfaktor anfügen; dies gilt auch für die  $E_{ab}$ .

<sup>24</sup>Diese lassen sich durch die diskrete Fouriertransformation  $\mathcal{F} = \sum_{ij} z_p^{ij} |i\rangle\langle j|$  ineinander überführen; es ist dann  $\mathcal{F} X^a \mathcal{F}^\dagger = Z^{-a}$ , wenn man  $\sum_{k \in \mathbb{F}_p} z_p^{kx} = p \cdot \delta_{x0}$  berücksichtigt.

<sup>25</sup> $X^a Z^b (X^a)^{-1} (Z^b)^{-1}$  heißt auch *gruppentheoretischer Kommutator* von  $X^a$  und  $Z^b$ .

<sup>26</sup>Im Falle von Primzahlpotenzen ist dies jedoch nicht unmittelbar möglich, und die Gleichung (2.16) muß als Definition aufgefaßt werden. Da die Gruppe  $(X^a)_{a \in \mathbb{F}_q}$  für  $n > 1$  nicht zyklisch ist, muß man eine Basis des Vektorraums  $\mathbb{F}_q \cong \mathbb{F}_{p^n}$  über  $\mathbb{F}_p$  verwenden, was die Notation verkompliziert; entsprechendes gilt für  $(Z^b)_{b \in \mathbb{F}_q}$ .



## Kapitel 3

# Grundlagen der Quantenkryptographie

In diesem Kapitel werden die Grundlagen der Quantenkryptographie vorgestellt und die Voraussetzungen zum Verständnis der Kapitel 4 bis 7 geschaffen.

### 3.1 Klassische Kryptographie

In diesem Abschnitt werden die Grundlagen der klassischen Kryptographie geschildert und der Zusammenhang zwischen klassischer Kryptographie und Quantenkryptographie hergestellt.

#### 3.1.1 Aufgaben kryptographischer Protokolle

Ziel eines kryptographischen Protokolls ist es, zwei Parteien den Austausch von Nachrichten derart zu ermöglichen, daß der Inhalt dieser Nachrichten einem eventuellen Lauscher verborgen bleibt.<sup>1</sup> Man betrachtet üblicherweise einen Sender, der in der Regel *Alice* ( $A$ ) genannt wird, einen Empfänger, der meist als *Bob* ( $B$ ) bezeichnet wird, und einen Lauscher, der in der Regel den Namen *Eve* ( $E$ , von engl. *eavesdropper*) trägt.

Die Aufgabe ist nun, es Alice mittels geeigneter Verschlüsselungsverfahren zu ermöglichen, eine Nachricht an Bob zu senden, die Eve zwar abhören darf, über deren Inhalt sie aber durch das Abhören keine Kenntnis erlangen soll. Um die Sicherheit eines Verfahrens zu beurteilen, müssen Annahmen an Eves Fähigkeiten getroffen werden. Bekannt ist zum Beispiel die Forderung, daß Eve eine Nachricht mit in der Schlüssellänge polynomialem Zeitaufwand

---

<sup>1</sup>Es gibt noch eine Reihe anderer Ziele, die in der Kryptographie angestrebt werden, das genannte ist aber wohl das älteste und bekannteste unter ihnen.

entziffern muß; vgl. z. B. BUCHMANN [27]. Die Details solcher Fragestellungen behandelt die Komplexitätstheorie.

In der Quantenkryptographie wird eine viel weitreichendere Sicherheit angestrebt, die *perfekte Sicherheit*. Das bedeutet, daß im Rahmen des verwendeten klassischen Modells Eve durch das Belauschen der Übertragung keine Information – sieht man von der Feststellung ab, daß überhaupt etwas übertragen wurde – über eine Nachricht erlangt.

### 3.1.2 Das *One-time pad*

Ein im Grunde genommen sehr einfaches Verfahren, um Nachrichten sicher zu verschlüsseln, ist das *One-time pad*.<sup>2</sup> Hierzu nimmt man an, daß Alices Nachricht (der Klartext) durch eine Zeichenkette  $t \in \mathbb{Z}_d^n$  beschrieben wird. Gleichzeitig teilen sich Alice und Bob einen Schlüssel  $s \in \mathbb{Z}_d^n$ , der a priori gleichverteilt ist und über den Eve keinerlei Kenntnis besitzt. Das *One-time pad* wird wie folgt verwendet:<sup>3</sup>

1. Alice erzeugt mittels  $c := t \oplus s \in \mathbb{Z}_d^n$  einen Chiffretext.
2. Alice sendet den Chiffretext  $c$  an Bob; Eve darf den Chiffretext  $c$  ebenfalls erfahren.
3. Bob entschlüsselt den Klartext zu  $t = c \ominus s$ .

Man kann zeigen, daß das *One-time pad* perfekte Sicherheit garantiert und daß Eve nichts über den Inhalt der Nachricht erfährt; vgl. SHANNON [166] oder das Lehrbuch von BUCHMANN [27]: Die Nachricht  $t$  kann als die Realisierung einer Zufallsvariablen  $T$  betrachtet werden, deren Wahrscheinlichkeitsverteilung vom erwarteten Inhalt, der verwendeten Sprache usw. beeinflußt wird. Die Wahrscheinlichkeiten  $P(T = t)$  entsprechen demnach Eves Vorwissen über die Nachricht. Der Schlüssel  $s$  ist die Realisierung einer gleichverteilten Zufallsvariable  $S$  über dem Raum  $\mathbb{Z}_d^n$ . Durch Kenntnisnahme des Chiffrextes  $C = T \oplus S = c$  versucht Eve, ihre Verteilung zu  $P(T = t | C = c)$  zu verbessern: im Idealfall ist die Wahrscheinlichkeit für ein bestimmtes  $t_0$  gleich Eins und die aller anderen gleich Null. Beachtet man, daß bei Unkenntnis der Realisierung von  $S$  der Chiffretext  $C$  gleichverteilt ist und daß eine

---

<sup>2</sup>Das *One-time pad* entsteht aus dem VIGNÈRE-Verfahren (vgl. Abschnitt 1.1.2), wenn der Schlüssel zufällig ist und die gleiche Länge wie der zu übermittelnde Text besitzt; häufig wird das Verfahren VERNAM [179] zugeschrieben.

<sup>3</sup>Will man Ver- und Entschlüsselung symmetrisch gestalten, so kann man  $c := s \ominus t$  und  $t = s \oplus c$  verwenden. In allen Fällen sind Schlüssel und Chiffretext austauschbar: die Information steckt in der Korrelation von beiden, die Kenntnis nur eines Teils ist wertlos.

gleichverteilte Zufallsvariable von jeder anderen Zufallsvariablen statistisch unabhängig ist, so berechnet man

$$P(T = t|C = c) \stackrel{\text{Def.}}{=} \frac{P(T = t \wedge C = c)}{P(C = c)} = \frac{P(T = t) \cdot P(C = c)}{P(C = c)} = P(T = t).$$

Dies bedeutet, das Eves Kenntnis über  $T$  unverändert bleibt, auch wenn sie  $C = c$  erfährt; das Abhören des Chiffretextes bringt ihr keinen Gewinn. Man beachte, daß die Sicherheit des *One-time pads* sich auf die folgenden Annahmen gründet:

- Die Auswahl des Schlüssels erfolgt zufällig ( $S$  ist gleichverteilt), und Alice und Bob kennen gemeinsam eine einzige *Realisierung* von  $S$ .
- Eve kennt die *Verteilung* von  $S$ , nicht aber die Realisierung.

Auch darf ein Schlüssel nur ein einziges Mal verwendet werden; wird ein Schlüssel wiederverwendet, so führt dies auf die VIGNÈRE-Verschlüsselung zurück, die schon im 19. Jahrhundert gebrochen wurde.

Das Hauptproblem des *One-time pads* ist die Erzeugung eines gemeinsamen geheimen Schlüssels von Alice und Bob, weshalb dieses Verfahren nur selten und nur für sehr wichtige Informationen verwendet wird.

## 3.2 Begriffe der Quantenkryptographie

Aufbauend auf die im vorangegangenen Abschnitt eingeführten Begriffe werden in diesem Abschnitt die Grundbegriffe der Quantenkryptographie vorgestellt. Dies betrifft insbesondere die Aspekte, die in dieser Dissertation behandelt werden.

### 3.2.1 Aufgaben quantenkryptographischer Protokolle

Im letzten Abschnitt wurde festgestellt, daß durch die Verwendung des *One-time pads* die perfekt sichere Nachrichtenübertragung möglich ist, wenn nur Alice und Bob sich einen Schlüssel hinreichender Länge teilen. Der Schlüssel ist dabei eine Zufallszahl, die nicht von Alice zu Bob übertragen werden muß, sondern von Alice und Bob gemeinsam erzeugt werden kann. Dies allein ist die Aufgabe der Quantenkryptographie, und man spricht aus diesem Grund oft auch von *Quanten-Schlüsselaustausch* oder engl. *quantum key distribution* (abgekürzt *QKD*).<sup>4</sup>

---

<sup>4</sup>Dies ist die klassische Aufgabenstellung der Quantenkryptographie; natürlich kann der erzeugte Schlüssel auch anderweitig verwendet werden, z. B. in anderen Verschlüsselungsverfahren (häufig wird AES genannt).

### 3.2.2 Das Modell

Es soll gezeigt werden, daß bestimmte quantenkryptographische Protokolle sicher sind. Hierzu müssen zunächst Annahmen an die Fähigkeiten der beteiligten Parteien getroffen werden. Klar ist, daß Alice und Bob in der Lage sein müssen, Quantenzustände zu präparieren und zu verarbeiten; es ist aber sinnvoll, die Anforderungen an die technische Umsetzung möglichst gering zu halten. Für die *Sicherheit* eines quantenkryptographischen Protokolls geht ein einfaches, aber plausibles Modell von den folgenden Annahmen aus:

- Die in den Laboratorien von Alice und Bob ausgeführten Handlungen, etwa die Präparation und Messung von Quantenzuständen oder die Realisierung klassischer Zufallsvariablen, sind Eve nicht zugänglich.
- Alice und Bob können über einen *Quantenkanal* Quanteninformation austauschen. Auf diesem Kanal kann Eve beliebige Quantenoperationen ausführen; eventuelle Fehler des Kanals werden Eve zugeschrieben.
- Alice und Bob können über einen *klassischen authentisierten Kanal* Nachrichten austauschen; Eve kann die gesamte Übertragung mitlesen, sie aber nicht manipulieren. Der Kanal soll fehlerfrei sein, was durch die Verwendung geeigneter Kodierungsverfahren hinreichend gut gelingt.

Die erste Annahme, daß die Laboratorien unangreifbar sind, erscheint sinnvoll, und sie wird auch in der klassischen Kryptographie gefordert. Eventuelle *Seitenkanalangriffe*, bei denen Eve Informationen aus der Unzulänglichkeit der Implementierung erhält, werden ausgeschlossen; vgl. auch Abschnitt 7.2 für einen Spezialfall.

Da jeder wirkliche Quantenkanal zumindest geringe Fehler aufweist, muß ein Sicherheitsbeweis darauf Rücksicht nehmen (ansonsten wäre ein solcher Beweis gleichermaßen trivial und nutzlos). Man kann nun annehmen, daß Eve – als ein sehr mächtiger und technisch fortschrittlicher Angreifer – den realen Quantenkanal durch einen idealen, fehlerfreien Kanal ersetzt und ihren Angriff so durchführt, daß gerade der Fehler des ursprünglichen Kanals auftritt, sie sich also durch den realen Kanal tarnt. Zumindest wird Eve durch diese Annahme nicht unterschätzt, und die Annahme ist im Sicherheitsbeweis nicht störend.

Der klassische Kanal ist schließlich notwendig, damit Alice und Bob miteinander öffentlich kommunizieren können; ohnehin wird er für den Austausch der klassischen verschlüsselten Nachricht benötigt. Mit der Frage nach der Bedeutung der Authentisierung befaßt sich der nächste Unterabschnitt.

### 3.2.3 Zur Bedeutung der Authentisierung

Eine der wichtigsten Voraussetzungen, die für die Sicherheit der Quantenkryptographie (und auch klassischer Protokolle) erfüllt sein muß, ist, daß der klassische Kanal, der Alice und Bob verbindet, authentisiert ist. Eve darf es nicht möglich sein, die übertragenen Nachrichten zu verfälschen oder zu unterdrücken. Ihre klassischen Angriffsmöglichkeiten werden also von aktiven zu passiven Angriffen reduziert; dies wurde schon von BENNETT UND BRASSARD [17] und von MAYERS [120] gefordert.

Würde man auf die Authentisierung verzichten, so wäre es möglich, daß Eve sich Alice gegenüber als Bob und Bob gegenüber als Alice ausgibt und mit jeder der beiden Parteien getrennt kommuniziert. Dies wird als *Man-in-the-Middle-Angriff* (vereinzelt auch *Janus-Angriff*) bezeichnet.

Um die Sicherheit quantenkryptographischer Protokolle zu gewährleisten, müssen Authentisierungsverfahren verwendet werden, die das gleiche Sicherheitsniveau wie die Protokolle aufweisen, da mit dem Brechen der Authentisierung das ganze System zerstört wird. Es gibt Authentisierungsverfahren, die die in der Quantenkryptographie geforderte unbedingte Sicherheit erreichen, sofern Alice und Bob sich nur einen kurzen Schlüssel zu Beginn teilen (vgl. WEGMAN UND CARTER [181]), weshalb man gelegentlich auch von engl. *quantum key expansion* spricht.<sup>5</sup>

## 3.3 Quantenkryptographische Protokolle

In diesem Abschnitt werden diejenigen quantenkryptographischen Protokolle vorgestellt, die in dieser Arbeit untersucht werden. Nicht behandelt werden andere Klassen von Protokollen wie das B92-Protokoll von BENNETT [15], das SARG-Protokoll von SCARANI u. a. [157] oder auch Protokolle mit kontinuierlichen Variablen.

### 3.3.1 Protokolle, die komplementäre Basen verwenden

Die wohl bekannteste Klasse von Protokollen ist die folgende: Alice und Bob wählen ein endlichdimensionales Quantensystem der Dimension  $d \in \mathbb{N} \setminus \{1\}$  aus und legen ein System von  $s \in \mathbb{N} \setminus \{1\}$  paarweise komplementären Basen fest. Die Basisvektoren seien nun  $|b, i\rangle \in \mathcal{H} \cong \mathbb{C}^d$ , wobei  $b \in \{1, \dots, s\}$  die Basis und  $i \in \mathbb{Z}_d$  den Ditwert bezeichne.

---

<sup>5</sup>Die Länge dieses Anfangsschlüssels verhält sich logarithmisch zur Länge der Nachricht, die authentisiert werden soll, was zu der etwas paradox klingenden Aussage führt, daß die zur Authentisierung aufzuwendende Schlüsselrate Null ist.

Das Protokoll lautet nun wie folgt:

1. Alice wählt gleichverteilt Werte  $i \in \mathbb{Z}_d$  und  $b \in \{1, \dots, s\}$  aus, notiert sich diese und sendet ein im reinen Zustand  $|b, i\rangle$  präpariertes Qudit an Bob; dies geschieht  $s \cdot n$  mal für ein hinreichend großes  $n \in \mathbb{N}$ .<sup>6</sup>
2. Bob mißt die Qudits in einer der  $s$  Basen, die er für jedes Qudit gleichverteilt auswählt; er notiert sich die Basis und sein Meßergebnis.
3. Alice und Bob vergleichen über den klassischen Kanal für jedes Qudit die zur Präparation bzw. Messung verwendete Basis; falls sie nicht die gleiche Basis verwendet haben, so verwerfen sie das Qudit. Aus den Ditzwerten der verbleibenden Qudits erhalten sie eine Zeichenfolge über dem Alphabet  $\mathbb{Z}_d$ ; im Mittel hat diese Zeichenfolge die Länge  $n$ .
4. Anhand einer Auswahl von  $m$  Zeichen ermitteln sie einen Schätzwert für die Fehlerverteilung des Kanals, indem sie ihre Werte über den klassischen Kanal vergleichen. Die verglichenen Zeichen sind dem Lauscher bekannt und müssen verworfen werden.
5. Abhängig von der geschätzten Fehlerrate entscheiden Alice und Bob, ob sie das Protokoll abbrechen oder ob sie durch Fehlerkorrektur und durch engl. *Privacy Amplification* (deutsch manchmal Privatsphärenverstärkung) einen sicheren Schlüssel erhalten können; ggf. wenden sie solche Verfahren an und verbleiben mit einem sicheren Schlüssel.

Beschränkt man sich auf Qubits ( $d = 2$ ), so gibt es im wesentlichen zwei Protokolle in dieser Klasse:

1. das *BB84-Protokoll*, das 1984 von BENNETT UND BRASSARD [17] vorgeschlagen wurde und welches das älteste dieser Protokolle ist; es verwendet die  $z$ - und die  $x$ -Basis sowie  $m = n/2$ ;
2. das *Protokoll mit sechs Zuständen* (engl. *six-state protocol*), das Qubits mit der  $z$ -,  $x$ - und  $y$ -Basis verwendet und in einer Arbeit von BRUß [25] untersucht wurde.

In dieser Arbeit werden Verallgemeinerungen dieser Protokolle auf höherdimensionale Quantensysteme untersucht; besonderen Wert wird dabei auf ein Protokoll gelegt, das zwei Basen, die Standardbasis  $\{|i\rangle | i \in \mathbb{Z}_d\}$  und die Fouriertransformierte  $\{\mathcal{F}|i\rangle | i \in \mathbb{Z}_d\}$  der Standardbasis von  $\mathbb{C}^d$  verwendet.

---

<sup>6</sup>Es wurden auch Protokolle vorgeschlagen, in denen die Basen nicht gleichverteilt ausgewählt werden. Dies erfordert aber bei der Abschätzung der Fehlerraten einen etwas größeren Aufwand; vgl. z. B. LO, CHAU UND ARDEHALI [113].

### 3.3.2 Abschätzung der Kanalfehlerrate

Im Grenzfall  $n \rightarrow \infty$  ist es durch statistische Tests möglich, Fehlerraten mit beliebiger Genauigkeit abzuschätzen. Im Falle einer endlichen Zahl übertragener Qudits müssen jedoch Verfahren der *schließenden Statistik* angewendet werden, die hier kurz angesprochen werden sollen. Sendet Alice  $i \in \mathbb{Z}_d$ , mißt Bob aber  $j \in \mathbb{Z}_d$ , so liegt ein Fehler vom Typ  $j \ominus i$  vor.

Werden  $n \in \mathbb{N}$  Qudits übertragen und gemessen, so gibt es unbekannte *absolute Häufigkeiten*  $K = (k_0, \dots, k_{d-1}) \in \mathbb{N}_0^d$  der Fehlertypen, für die offensichtlich  $\sum_{i=0}^{d-1} k_i = n$  gilt. Anhand einer *Stichprobe* von  $m \in \{0, \dots, n\}$  zufällig ausgewählten Qudits ermitteln die Parteien eine Häufigkeitsverteilung  $L = (l_0, \dots, l_{d-1}) \in \mathbb{N}_0^d$  mit  $\sum_{i=0}^{d-1} l_i = m$ , von der sie auf  $K$  zurückschließen wollen. Im Falle großer  $n$  und  $m$  ist nach dem Gesetz der großen Zahlen  $K \approx (n/m) \cdot L$  eine naheliegende Schätzung; von Bedeutung ist der Wert von  $K - L$ . Die bedingte Wahrscheinlichkeit, bei vorgegebenen Häufigkeiten  $K$  eine Stichprobe  $L$  zu erhalten, ist

$$P(L|K) = \binom{n}{m}^{-1} \times \prod_{i=0}^{d-1} \binom{k_i}{l_i}. \quad (3.1)$$

Im Falle von Qubits  $d = 2$  ist dies die *hypergeometrische Verteilung*, die im Urnenmodell der Statistik dem „Ziehen ohne Zurücklegen und ohne Berücksichtigung der Anordnung“ entspricht.

Nach der *Maximum-Likelihood-Methode* (selten *Methode der maximalen Mutmaßlichkeit*) bestimmt man nun dasjenige  $K$ , für das die *Likelihood-Funktion* (selten *Mutmaßlichkeitsfunktion*)  $f_L(K) := P(L|K)$  maximal wird. Der *Satz von BAYES*  $P(K|L) = P(L)^{-1} \cdot P(L|K) \cdot P(K)$  erlaubt es, den Ausdruck „umzudrehen“, und mit dem *Satz über die totale Wahrscheinlichkeit* läßt sich  $P(L) = \sum_K P(L|K)P(K)$  schreiben, wofür allerdings die Verteilung von  $K$  bekannt sein muß, die aber durch Eve beeinflußt wird.

Ein anderes, eher praktikables Verfahren, die Verteilung abzuschätzen, besteht darin, anzunehmen, daß ein *Kanal* vorliegt, der gewisse Fehlerraten  $p = (p_0, \dots, p_{d-1}) \in \mathcal{W}_d$  besitzt, und daß eine Stichprobe der Länge  $m$  gezogen wird. Diese Parameterabschätzung einer Multinomialverteilung kann auf Parameterabschätzungen mehrerer Binomialverteilungen zurückgeführt werden, die wiederum (für  $n, m \rightarrow \infty$  exakte) quantitative Abschätzungen der hypergeometrischen Verteilung ermöglichen. Für ein festes  $i \in \mathbb{Z}_d$  kann man nun das zum *Schätzwert*  $p_i = l_i/m$  zugehörige *Konfidenzintervall* (auch *Vertrauensbereich*) zu einem Niveau  $1 - \alpha$  bestimmen, d. h. das Intervall  $[p_-, p_+]$ , in dem  $p_i$  mit einer Wahrscheinlichkeit von mindestens  $1 - \alpha$  liegt, wozu man die tabellierten Quantile der Standard-Normalverteilung verwendet. Für Einzelheiten verweise ich auf das Buch von KREYSZIG [103].

### 3.3.3 Angriffe auf die Protokolle

Der einfachste Angriff auf ein Protokoll der beschriebenen Klasse besteht darin, daß Eve zufällig in einer der  $s$  möglichen Basen mißt (die Basen sind Eve bekannt, nicht aber die Basis, die für die Präparation eines einzelnen Qudits verwendet wird) und den gemessenen Zustand an Bob weitersendet (für Photonen ein *Intercept-and-Resend*-Angriff).

Die Wahrscheinlichkeit, in der richtigen Basis zu messen, ist  $1/s$ ; in diesem Fall erhält sie den Präparations- bzw. Meßwert von Alice und Bob. Andernfalls erzeugt der Angriff mit Wahrscheinlichkeit  $1 - d^{-1}$  einen Fehler, und die Gesamtfehlerrate, die Alice und Bob bestimmen, ist

$$p_{\text{Fehler}} = \frac{1}{s} \cdot 0 + \frac{s-1}{s} \cdot \frac{d-1}{d}. \quad (3.2)$$

Im BB84-Protokoll ( $d = 2$  und  $s = 2$ ) und im Protokoll mit sechs Zuständen ( $d = 2$  und  $s = 3$ ) ergeben sich somit Fehlerraten von 25 % bzw.  $33,3\%$ . Auch Eves gemessene Werte weisen diese Fehlerraten auf. Im Falle von  $s = 2$  und  $d \rightarrow \infty$  wächst diese Fehlerrate bis auf 50 %.

Bei dem geschilderten Angriff handelt es sich um einen *individuellen Angriff*, da jedes Qubit einzeln angegriffen und gemessen wird. Allgemeiner muß man sogenannte *kohärente Angriffe* betrachten: hierbei fängt Eve die gesamten  $n$  Qudits ab, führt Quantenoperationen aus (koppelt sie an Hilfsysteme, führt unitäre Transformationen und Messungen durch usw.) und sendet einen  $n$ -Qudit-Zustand an Bob weiter, der in sich und mit Eves System nach Belieben verschränkt sein kann. Erst nachdem Alice und Bob ihr Protokoll abgeschlossen haben, wertet sie ihren zurückgehaltenen Zustand durch verallgemeinerte Messungen aus und versucht, den Schlüssel zu erschließen.

MAYERS [120] gelang auf etwa 50 Seiten ein direkter Beweis der Sicherheit des BB84-Protokolls unter kohärenten Angriffen. Im folgenden soll der Umweg über verschränkungs-basierte Protokolle gegangen werden, für die zumindest die Grundidee eines Sicherheitsbeweises nahezu trivial ist.

## 3.4 Verschränkungs-basierte Protokolle

Neben der im vorangegangenen Abschnitt eingeführten Klasse von Protokollen der Quantenkryptographie, den sogenannten *Prepare-and-Measure*-Protokollen, gibt es eine weitere Klasse von Protokollen, die auf EKERT [49] zurückgeht und die als *verschränkungs-basierte Protokolle* bezeichnet werden.

Die Grundidee dieser Protokolle bilden die folgenden Beobachtungen:

1. Teilen sich Alice und Bob den Zustand  $|\Psi_{00}\rangle\langle\Psi_{00}|$ , so erhalten sie durch Messung bzgl. der Standardbasis den gleichen Wert aus  $\mathbb{Z}_d$ , der im vorhinein unbestimmt und gleichverteilt ist.
2. Eve erfährt nichts über das Ergebnis der Messung, da der Gesamtzustand von Alice, Bob und Eve die Form  $\rho_{ABE} = |\Psi_{00}\rangle\langle\Psi_{00}| \otimes \rho_E$  besitzt und Eves Messungen ihr keine Information liefern.<sup>7</sup>

Gelingt es nun, Alice und Bob mit einer genügend großen Anzahl  $n \in \mathbb{N}$  der unter Nummer 1 genannten Zustände zu versorgen, so können sie durch lokale Messungen einen sicheren, zufälligen Schlüssel der Länge  $n$  erzeugen.

### 3.4.1 Das Standard-Protokoll

Für ein konkretes, verschränkungsbasiertes Protokoll, wählen sich Alice und Bob ein System  $s$  paarweise komplementärer Basen aus, von denen die erste die Standardbasis sei; das Protokoll lautet dann wie folgt:

1. Alice präpariert  $n \in \mathbb{N}$  Qudit-Paare im Zustand  $\rho_0 := |\Psi_{00}\rangle\langle\Psi_{00}|$ .
2. Für jedes Paar wählt Alice gleichverteilt eine der  $s$  komplementären Basen aus und transformiert das zweite Qudit des Paares von der Standardbasis in diese Basis.
3. Alice sendet das zweite Qudit jedes Paares über den Quantenkanal zu Bob; Eve führt ihren Angriff durch.
4. Nachdem Bob den Empfang der  $n$  Qudits bestätigt hat, teilt Alice ihm die verwendeten Basen mit; Bob invertiert Alices Transformation.
5. Alice und Bob messen  $m$  der  $n$  Paare, vergleichen diese und schätzen hierdurch die Fehlerverteilung ab.
6. Abhängig von der geschätzten Fehlerverteilung brechen Alice und Bob das Protokoll ab, oder sie verwenden Protokolle zur Verschränkungereinigung, um  $n' \leq n - m$  Zustände zu erhalten, die hinreichend nahe am Idealzustand  $\rho_0$  liegen.
7. Sie messen lokal die Ditwerte und erhalten hierdurch einen Schlüssel der Länge  $n'$  über  $\mathbb{Z}_d$ .

---

<sup>7</sup>Die Anmerkung zu Satz C.19 ergibt  $|S(\rho_{AB}) - S(\rho_E)| \leq S(\rho_{ABE}) \leq S(\rho_{AB}) + S(\rho_E)$ , und für den reinen Zustand  $\rho_{AB} = |\Psi_{00}\rangle\langle\Psi_{00}|$  gilt  $S(\rho_{AB}) = 0$ . Es folgen  $S(\rho_{ABE}) = S(\rho_E)$  und mit  $S(\rho_{ABE}) = S(\rho_{AB}) + S(\rho_E)$  die gewünschte Aussage.

In Schritt 5 müssen sie die Positionen der  $m$  Prüf-Qudits zufällig auswählen, da Eve sonst die Prüf-Qudits und die Schlüssel-Qudits unterscheiden könnte. Alternativ können Alice und Bob eine zufällige Permutation der  $n$  Qudit-Paare vornehmen; dies hat zur Folge, daß Eves Angriff symmetrisch über alle Qudits wird, da sie in Schritt 3 noch keine Kenntnis über die verwendete Permutation hat. Dies wäre ein Ansatzpunkt für die DE-FINETTI-Sätze, die hier aber nicht verwendet werden (vgl. Unterabschnitt 7.3).

### 3.4.2 Zum Begriff der Sicherheit

Die Protokolle der Verschränkungsreinigung ermöglichen es, sich mit dem von Alice und Bob geteilten Zustand dem Idealzustand  $\rho_0^{\otimes n'} = (|\Psi_{00}\rangle\langle\Psi_{00}|)^{\otimes n'}$  beliebig zu nähern, können ihn aber niemals mit Sicherheit erreichen. Es muß daher einen Weg geben, aus der zulässigen Abweichung vom Idealzustand auf die dem Angreifer zugängliche Information zurückzuschließen. Dies erfolgt oft durch Schranken an die gemeinsame Information von Alice und Bob mit Eve (vgl. LO UND CHAU [112] oder meine Diplomarbeit [138]). Meines Erachtens ist jedoch die folgende Aussage einfacher verständlich: zwei Dichtematrizen  $\rho$  und  $\rho_0$  können durch eine verallgemeinerte Messung nicht mit einer Wahrscheinlichkeit größer als  $\delta(\rho, \rho_0) := \frac{1}{2} \|\rho - \rho_0\|_1$  unterschieden werden. Dies bedeutet auch, daß man statt einer *idealen* Dichtematrix  $\rho_0$  eine *reale* Dichtematrix  $\rho$  verwenden kann, wenn man einen Fehler von  $\delta(\rho, \rho_0)$  zuläßt.<sup>8</sup>

Den Sachverhalt kann man wie folgt begründen: Alice und Bob führen eine Messung am *realen* Zustand  $\rho$  aus, würden aber lieber den *idealen* Zustand  $\rho_0$  messen. Die Messungen kann man als Realisierungen einer *realen* Zufallsvariablen  $X$  und einer *idealen* Zufallsvariablen  $X_0$  auffassen. Die Norm  $\|X - X_0\|_1$  ist nun nicht größer als die Spurnorm  $\|\rho - \rho_0\|_1$ .<sup>9</sup> Nimmt man also an, daß Alice und Bob die Zufallsvariablen  $X$  und  $X_0$  immer gleichzeitig auswerten, so besagt der Satz B.38, daß es eine gemeinsame Verteilung  $P_{X X_0}$  derart gibt, daß sie mit einer Wahrscheinlichkeit von mindestens  $1 - \varepsilon$  übereinstimmen. Sie sind jedoch gezwungen  $X$  anstelle von  $X_0$  zu verwenden, was ihnen einen Fehler von  $\varepsilon$  liefert.

---

<sup>8</sup>Bekannt wurde mir diese Interpretation einschließlich des Satzes B.38 im Verlauf von Diskussionen, die in Erlangen zwischen dem 10. bis 14. Oktober 2005 mit RENATO RENNER und anderen anläßlich des SECOQC-QIT-Workshops geführt wurden.

<sup>9</sup>Eine verallgemeinerte Messung auf  $\mathcal{H} \cong \mathbb{C}^n$  wird für ein  $r \leq n^2$  (Lemma C.27) durch eine Zerlegung der Eins  $(E_k)_{k=1}^r$  dargestellt. Das Ergebnis der Messung an  $\rho$  ist eine Wahrscheinlichkeitsverteilung  $p = (p_k)_{k=1}^r \in \mathcal{W}_d$  mit  $p_k := \text{Spur } \rho E_k$ ; sinngemäß erzeugt die Messung an  $\rho_0$  ein  $p_0$ . Man berechnet  $|p_k - p_{0,k}| = |\text{Spur}(\rho - \rho_0)E_k| \leq \text{Spur } |\rho - \rho_0| E_k$  und hieraus  $\|p - p_0\|_1 \leq \text{Spur } |\rho - \rho_0| \cdot \sum_{k=1}^r E_k = \|\rho - \rho_0\|_1$ . Setzt man  $\rho - \rho_0 = Q - S$  für  $Q, S \geq 0$ , so erhält man Gleichheit, wenn kein  $E_k$  gleichzeitig Anteile der Träger von  $Q$  und  $S$  beinhaltet; vgl. NIELSEN UND CHUANG [127], Theorem 9.1 auf S. 405.

Für  $\rho_0 = |\Psi_{00}\rangle\langle\Psi_{00}|$  und  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  zeigt das Lemma C.29, daß  $\frac{1}{2} \|\rho^{\otimes n} - \rho_0^{\otimes n}\|_1 \leq \frac{n}{2} \|\rho - \rho_0\|_1 = n \cdot (1 - A_{00})$  gilt; ganz allgemein gilt die Interpretation auch für andere Referenzzustände als  $\rho_0$ .

### 3.4.3 Vergleich mit *Prepare-and-Measure*-Protokollen

Vergleicht man die *Prepare-and-Measure*-Protokolle mit den verschränkungs-basierten Protokollen, so stellt man fest:

- *Prepare-and-Measure*-Protokolle sind experimentell einfacher zu verwirklichen als verschränkungs-basierte Protokolle;
- verschränkungs-basierte Protokolle sind theoretisch leichter zu untersuchen als *Prepare-and-Measure*-Protokolle.

Den Ideen der grundlegenden Arbeiten von SHOR UND PRESKILL [167] und von GOTTESMAN UND LO [64] für Qubits folgend wird nun versucht, die Sicherheit verschränkungs-basierter Protokolle zu beweisen und hieraus durch geschickte Umformulierungen zu zeigen, daß gewisse *Prepare-and-Measure*-Protokolle ebenfalls sicher sind.

## 3.5 Sicherheitsbeweise für Protokolle

Im Unterabschnitt 3.4.1 wurde ein verschränkungs-basiertes Protokoll vorgestellt, welches nun in mehreren Schritten auf ein *Prepare-and-Measure*-Protokoll zurückgeführt werden soll, ohne die Sicherheit zu beeinträchtigen. Dies ist die Idee des Beweises von SHOR UND PRESKILL [167]; man vergleiche auch das Buch von NIELSEN UND CHUANG [127], Kapitel 12.6 auf S. 582–603.

### 3.5.1 Calderbank-Shor-Steane-Codes

Die CALDERBANK-SHOR-STEANE-*Codes* [28, 171] oder kurz *CSS-Codes* sind eine wichtige Klasse fehlerkorrigierender Quantencodes, deren Eigenschaften es erlauben, sie in Sicherheitsbeweisen für quantenkryptographische Protokolle zu verwenden. Insbesondere gibt es einfach auszuwertende, wenn auch nicht-konstruktive Schranken für die Existenz von CSS-Codes (vgl. Satz 4.1).

Betrachtet man für eine Primzahl<sup>10</sup>  $p$  den Hilbertraum  $\mathcal{H} = \mathbb{C}^p$ , so kann die Standardbasis dieses Raums durch Elemente aus  $\mathbb{F}_p$  bezeichnet werden.

---

<sup>10</sup>Die Einschränkung auf Primzahlpotenzen ist notwendig, weil lineare Codes verwendet werden; hier werden nur Primzahlen verwendet, da dies die Notation erheblich vereinfacht.

Es seien nun  $C_2 \subseteq C_1 \subseteq \mathbb{F}_p^n$  klassische lineare  $[n, k_2]$ - bzw.  $[n, k_1]$ -Codes im Sinne von Anhang C.4.2. Können nun  $C_1$  und  $C_2^\perp$  jeweils  $t_1$  und  $t_2$  Fehler korrigieren, so wird der CSS-Code  $\text{CSS}(C_1, C_2)$  in der Lage sein,  $t_1$  Ditfehler und  $t_2$  Phasenfehler zu korrigieren. Für  $u \in C_2$ ,  $x \in C_1/C_2$  und  $v \in \mathbb{F}_p^n/C_1$  setzt man nun<sup>11</sup>

$$|x; u, v\rangle := |C_2|^{-1/2} \sum_{y \in C_2} z_p^{u \cdot y} |x \oplus y \oplus v\rangle. \quad (3.3)$$

Die Vektoren  $|x; u, v\rangle$  bilden dann eine Orthonormalbasis des Hilbertraums, und man erhält CSS-Codes  $\text{CSS}_{u,v}(C_1, C_2) := \{|x; u, v\rangle \mid x \in C_1/C_2\} \cong C_1/C_2$ , die die Vektoren  $x$  kodieren. Man kann zeigen, daß alle diese Codes in bezug auf ihre Fehlerkorrektoreigenschaften äquivalent sind, und im folgenden wird zunächst nur den Fall  $u = 0$  und  $v = 0$  betrachtet.

Ein kodierter Vektor  $|x; 0, 0\rangle$  geht durch einen Fehler in einen verfälschten Vektor

$$|x'\rangle := |C_2|^{-1/2} \sum_{y \in C_2} z_p^{(x+y) \cdot e_2} |x \oplus y \oplus e_1\rangle. \quad (3.4)$$

über. Hierbei ist  $e_1$  das Muster der Ditfehler und  $e_2$  das der Phasenfehler; im letzteren Fall ist die genannte Schreibweise möglich, wenn man festsetzt, daß  $|0\rangle$  keinen Phasenfehler aufweisen kann und die Phasenfehler der übrigen relativ hierzu mißt. Dies ist möglich, da  $z_p^k$  außer für  $k = 0$  die Gruppe der  $p$ -ten Einheitswurzeln erzeugt. Auch eventuelle Superpositionen von  $|x'\rangle$  können korrigiert werden, da die Korrektur kohärent erfolgen wird. Koppelt man den Zustand  $|0\rangle^{\otimes(n-k)}$  an  $|x'\rangle$  an, so kann  $|x'\rangle|0\rangle^{\otimes(n-k)}$  in

$$|C_2|^{-1/2} \sum_{y \in C_2} z_p^{(x+y) \cdot e_2} |x \oplus y \oplus e_1\rangle |He_1\rangle \quad (3.5)$$

überführt werden, wenn  $H$  die Kontrollmatrix des Codes  $C_1$  ist. Hierzu beachte man  $H(x \oplus y \oplus e_1) = He_1$  wegen  $x + y \in C_1$ . Für die  $i$ -te Zeile von  $H$  wird also  $|e_1, \dots, e_{k_2}\rangle|0\rangle \mapsto |e_1, \dots, e_{k_2}\rangle|h_{11}e_1 \oplus \dots \oplus h_{1k}e_k\rangle$  angestrebt. Durch Anwenden der GXOR-Transformation vom ersten zum Zusatz-Qudit erhält man  $|e_1, \dots\rangle|\ominus e_1\rangle$  durch Anwenden von  $\mathcal{F}^2$  wird dies  $|e_1, \dots\rangle|e_1\rangle$ . Die  $h_{11}$ -fache Wiederholung erzeugt  $|e_1, \dots\rangle|h_{11}e_1\rangle$ . Entsprechend kann man die übrigen Qudits verarbeiten, so daß das gewünschte Ergebnis folgt. Anhand von  $He_1$  kann man  $e_1$  erschließen, falls das Gewicht nicht zu hoch ist. Demzufolge können die Ditfehler durch Anwendung von  $\sum_{He_1} X^{-e_1} \otimes |He_1\rangle\langle He_1|$  zu  $|x''\rangle = \sum_{y \in C_2} z_p^{(x+y) \cdot e_2} |x \oplus y\rangle$  korrigiert werden.

---

<sup>11</sup>Ist  $C_2 = \text{span}\{e_1, \dots, e_{k_2}\}$  und  $C_1 = \text{span}\{e_1, \dots, e_{k_1}\}$ , so erhält man die Isomorphie  $C_1/C_2 \cong \text{span}\{e_{k_1}, \dots, e_{k_2}\}$ ; dies gilt sinngemäß für  $\mathbb{F}_p^n/C_1$ . Im folgenden werden solche Identifizierungen benutzt.

Die Korrektur der Phasenfehler  $e_2$  kann auf die Korrektur von Ditfehlern zurückgeführt werden. Hierzu wendet man die diskrete Fouriertransformation auf jedes Qudit einzeln an und erhält

$$\begin{aligned} \mathcal{F}^{\otimes n}|x''\rangle &= |C_2|^{-1/2} p^{-n/2} \sum_{i,j \in \mathbb{F}_p^n} z_p^{i \cdot j} |i\rangle \langle j| \cdot \sum_{y \in C_2} z_p^{(x+y) \cdot e_2} |x \oplus y\rangle \\ &= |C_2|^{-1/2} p^{-n/2} \sum_{i \in \mathbb{F}_p^n, y \in C_2} z_p^{(i+e_2) \cdot (x+y)} |i\rangle. \end{aligned} \quad (3.6)$$

Setzt man  $i' := i \oplus e_2$  und verwendet Lemma C.40, so folgt weiter

$$\begin{aligned} \mathcal{F}^{\otimes n}|x''\rangle &= |C_2|^{-1/2} p^{-n/2} \sum_{i' \in \mathbb{F}_p^n} z_p^{i' \cdot x} \left( \sum_{y \in C_2} z_p^{i' \cdot y} \right) |i \ominus e_2\rangle \\ &= |C_2|^{+1/2} p^{-n/2} \sum_{i' \in C_2^\perp} z_p^{i' \cdot x} |i \ominus e_2\rangle, \end{aligned} \quad (3.7)$$

und dies läßt sich analog zu den Ditfehlern korrigieren.

### 3.5.2 Reduktion auf CSS-basierte Protokolle

Betrachtet man das Standard-Protokoll aus Unterabschnitt 3.4.1, so stellt man fest, daß Alice die zur Prüfung des Kanals ausgewählten Qudit-Paare schon unmittelbar nach Präparation messen kann; mit anderen Worten kann sie gleich Ein-Qudit-Zustände präparieren. Im weiteren werde angenommen, daß zur Fehlerkorrektur im vorletzten Schritt ein CSS-Code verwendet wird. Man berechnet nun

$$|\Phi_{00}\rangle^{\otimes n} = p^{-n/2} \sum_{j \in \mathbb{F}_p^n} |j\rangle_A |j\rangle_B = p^{-n/2} \sum_{x; u, v} |x; u, v\rangle_A |x; u, v\rangle_B. \quad (3.8)$$

Man kann zeigen, daß die CSS-Korrektur zufällige Werte für  $u$  und  $v$  liefert und die abschließende Messung ebenfalls zufällige Werte für  $x$  ergibt. Für die Sicherheit des Protokolls ist es also äquivalent, wenn Alice die Werte  $x$ ,  $u$  und  $v$  zufällig auswählt und den bzgl.  $\text{CSS}_{u,v}(C_1, C_2)$  kodierten Vektor  $x$  an Bob sendet. Man erhält somit das folgende *CSS-Protokoll*:

1. Alice erzeugt  $m$  zufällige Prüfdits und kodiert diese in der Standardbasis. Sie wählt einen zufälligen Schlüssel  $x \in C_1/C_2 \cong \mathbb{F}_p^{n'}$  und kodiert diesen für zufällige Werte  $u \in C_2$  und  $v \in \mathbb{F}_p^{n-m}/C_1$  bzgl.  $\text{CSS}_{u,v}(C_1, C_2)$ ; insgesamt entstehen  $n$  Qudits.
2. Die Schritte 2–4 des ursprünglichen Protokolls bleiben unverändert; danach übermittelt Alice die Positionen der Prüfdits sowie  $u$  und  $v$ .
3. Bob mißt die  $m$  Prüfdits in der Standardbasis und veröffentlicht die Ergebnisse; Alice und Bob schätzen die Fehlerverteilung ab; kann sie der CSS-Code nicht korrigieren, so bricht das Protokoll ab.

4. Bob dekodiert die verbleibenden  $n - m$  Qudits bzgl.  $\text{CSS}_{u,v}(C_1, C_2)$ .
5. Bob mißt seine Qudits und erhält den geheimen Schlüssel  $x$ .

Dieses modifizierte Protokoll ist also explizit nicht verschränkungsbasiert, seine Sicherheit aber äquivalent zum ursprünglichen.

### 3.5.3 Reduktion auf *Prepare-and-Measure*-Protokolle

Das im letzten Unterabschnitt beschriebene Protokoll hat noch den Nachteil, daß Alice und Bob Zustände aus vielen Qudits präparieren bzw. messen müssen. In einem zweiten Schritt soll das Protokoll daher in ein *Prepare-and-Measure*-Protokoll der Form von Unterabschnitt 3.3.1 überführt werden. Die Grundidee besteht darin, daß bei CSS-Codes Dit- und Phasenfehlerkorrektur entkoppelt sind; Alice und Bob könnten die Phasenfehler korrigieren, müssen es aber nicht, da diese keinen Einfluß auf den Schlüssel haben.

Man stellt nun fest, daß die Phaseninformation  $u$  des CSS-Codes für Bob unbedeutend ist, da die Phasenwerte überhaupt nicht gemessen werden; Alice braucht sie daher auch nicht zu senden. Effektiv sendet sie also den über  $u$  gemittelten gemischten Zustand

$$\begin{aligned}
 \rho_{x,v} &= \frac{1}{p^n} \sum_{u \in \mathbb{F}_p^n} |x; u, v\rangle \langle x; u, v| \\
 &= \frac{1}{p^n |C_2|} \sum_{u \in \mathbb{F}_p^n} \sum_{w_1, w_2 \in C_2} z_p^{u \cdot (w_1 - w_2)} |x \oplus w_1 \oplus v\rangle \langle x \oplus w_2 \oplus v| \\
 &= \frac{1}{|C_2|} \sum_{w \in C_2} |x \oplus w \oplus v\rangle \langle x \oplus w \oplus v|. \tag{3.9}
 \end{aligned}$$

Dieser Zustand kann erzeugt werden, indem Alice willkürlich ein  $w \in C_2$  auswählt und  $|x \oplus w \oplus v\rangle$  präpariert; da nun die Werte  $w \in C_2$ ,  $x \in C_1/C_2$  und  $v \in \mathbb{F}_p^{n-m}/C_1$  alle zufällig sind, kann Alice stattdessen auch einfach ein  $v' \in \mathbb{F}_p^{n-m}$  zufällig wählen und den Zustand  $|v'\rangle$  präparieren. Bobs Messung liefert dann  $v' \oplus \varepsilon$ , und Alice veröffentlicht  $v' \ominus \varepsilon$ , so daß Bob  $x \oplus \varepsilon$  berechnen und zu  $x$  korrigieren kann.

Schließlich kann man auf die Basistransformationen der Schritte 2 und 4 des ursprünglichen Protokolls verzichten, indem Alice in einer der  $s$  Basen zufällig präpariert und Bob in einer der Basen zufällig mißt. Nach der Übertragung sprechen sie sich bzgl. der verwendeten Basen ab, was mit einem Anteil von etwa  $1/s$  zu einer Übereinstimmung führt. Wählt man zur Sicherheit einen Überschußanteil  $\delta \in \mathbb{R}^+$  für die Zahl der übertragenen Zustände, so erhält man das folgende Protokoll:

1. Alice erzeugt  $(s + \delta)n$  zufällige Dits, kodiert sie zufällig in einer der  $s$  Basen und sendet sie an Bob; ferner wählt sie zufällig ein  $x \in C_1$ .
2. Bob empfängt die Qudits und mißt sie zufällig in einer der  $s$  Basen; er bestätigt Alice den Empfang.
3. Alice und Bob tauschen sich über die Basen aus, verwerfen alle Werte, bei denen sie unterschiedliche Basen verwendet haben und verbleiben in der Regel mit mehr als  $n$  Dits.
4. Anhand von  $m$  Dits schätzen sie die Fehlerverteilung ab; ist diese nicht korrigierbar, so brechen sie das Protokoll ab, andernfalls verbleibt Alice mit einer Ditfolge  $v'$  und Bob mit einer Ditfolge  $v' \oplus \varepsilon$ .
5. Alice veröffentlicht  $v' \ominus x$ . Bob subtrahiert dies von seinem Ergebnis und erhält nach Korrektur mit  $C_1$  die Ditfolge  $x$ .
6. Alice und Bob berechnen die Nebenklasse  $x + C_2$ , und erhalten einen gemeinsamen geheimen Schlüssel.

Die Sicherheit ist gegenüber dem ursprünglichen Protokoll unverändert. Es handelt sich um das Protokoll aus Unterabschnitt 3.3.1, das um ein Fehlerkorrekturverfahren ergänzt wurde.

### 3.5.4 Protokolle mit Zweiweg-Kommunikation

Die bislang untersuchten Protokolle verwenden CSS-Codes, zu deren Dekodierung nur Einweg-Kommunikation von Alice zu Bob benötigt wird. Es hat sich aber gezeigt, daß Protokolle mit Zweiweg-Kommunikation wesentlich höhere Fehlerraten tolerieren können als solche mit Einweg-Kommunikation. Allgemeine Untersuchungen, welche Zweiweg-Protokolle sich ähnlich wie die CSS-Codes auf *Prepare-and-Measure*-Protokolle zurückführen lassen, führten GOTTESMAN UND LO [64] und LO [111] durch. Im folgenden werden die Grundideen aufgeführt, für Details sei auf die Originalarbeiten verwiesen.

In einem Protokoll zur Verschränkungsreinigung mit Zweiweg-Kommunikation messen nun Alice und Bob lokal ihre Zustände, tauschen sich über das Ergebnis aus, entscheiden über den nächsten Schritt und wiederholen diesen Prozeß nach Belieben. In einem *Prepare-and-Measure*-Protokoll tun sie im wesentlichen das gleiche, aber nur für die Dit- und nicht für die Phasenwerte. Um ein Protokoll zur Verschränkungsreinigung in ein *Prepare-and-Measure*-Protokoll der Quantenkryptographie umzuformen, muß man nun die Messung und Korrektur der Dit- und Phasenfehler voneinander trennen, ähnlich wie es bei den CSS-Codes erfolgte.

Man fordert nun, daß Alice und Bob nur jeweils gleiche Typ- $X$ - und Typ- $Z$ -Messungen an ihren Qudits ausführen (wenn Alice  $M_A \otimes \mathbb{1}_B$  mißt, dann mißt Bob  $\mathbb{1}_A \otimes M_B$ ) und daß nur die Parität beider Messungen verwendet wird. Der Fortgang eines Protokolls soll nicht von den Ergebnissen der Typ- $X$ -Messungen (Phasenmessungen) abhängen, da diese in der Reduktion auf ein *Prepare-and-Measure*-Protokoll fortfallen sollen. Die Phasenmessungen dürfen auch nicht für die Korrektur der Ditzfehler mittels Typ- $X$ -Operatoren verwendet werden.

Die Dit- und Phasenmessungen erfolgen über die Messungen von Typ- $Z$ - bzw. Typ- $X$ -Operatoren, die mittels GBXOR-Transformationen implementiert werden können. Eventuelle Phasenfehlerkorrekturen erfolgen dann über Typ- $Z$ -Transformationen, die mithilfe von  $\text{GXOR}(Z \otimes \mathbb{1}) = (Z \otimes \mathbb{1}) \text{GXOR}$  und  $\text{GXOR}(\mathbb{1} \otimes Z) = (Z \otimes Z^{-1}) \text{GXOR}$  ans Ende des Protokolls verlagert werden. Dort sind sie für den Schlüssel aber unbedeutend und entfallen.

Es zeigt sich, daß die Protokolle, die die hier aufgeführten Bedingungen erfüllen, im wesentlichen Varianten der  $B_n^{(d)}$ - und  $P_n^{(d)}$ -Schritte sind, die in den folgenden Kapiteln eingeführt werden.

### 3.5.5 Reduktion auf bell-diagonale Zustände

Im verschränkungsbasierten Protokoll teilen sich Alice und Bob vor Beginn der Verschränkungsreinigung einen Zustand  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$  für  $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$ . Ein solcher Zustand ist jedoch theoretisch nur schwer zu handhaben; einfacher wäre es, wenn der Zustand in der Form  $\rho^{\otimes n}$  für ein  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  geschrieben werden könnte. Dies gelingt durch eine einfache Überlegung:

1. Anstelle von Alice könnte auch Eve den Zustand  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$  präparieren, was ihr bestenfalls mehr Macht gibt.
2. Nach ihrem eigentlichen Angriff könnte Eve jedes Qudit-Paar einzeln in der BELL-Basis messen, wodurch  $\rho$  zu einem Tensorprodukt reiner BELL-Zustände würde.
3. Wenden Alice und Bob wie schon gefordert eine zufällige Permutation aus  $S_n$  an, so kann der Zustand für  $n \rightarrow \infty$  in der Form  $\rho^{\otimes n}$  für ein  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  geschrieben werden.<sup>12</sup>

Da Eve nicht zu einer Messung gezwungen werden kann, können Alice und Bob nach Abschluß des Protokolls zusammenkommen und jedes Paar in der BELL-Basis messen. Diese Messung kommutiert aber mit den Messungen

---

<sup>12</sup>Der Grenzübergang ist erforderlich und erfolgt in Analogie zum Übergang von der hypergeometrischen auf die Binomialverteilung.

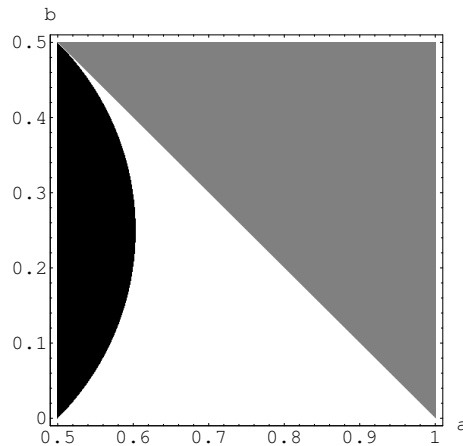


Abbildung 3.1: Korrigierbare Zustände und nicht-korrigierbare Zustände (weiß/schwarz) für  $\rho = (a, b, c, d) \in \mathcal{S}_{\text{bd}}$  mit  $a \geq 1/2$ ; grau: keine physikalischen Zustände.

$M_A \otimes M_B$ , die Alice und Bob anstelle ihrer lokalen Messungen ausführen könnten, so daß sie schon vor der eigentlichen Verschränkungsreinigung ausgeführt werden könnte. Dann könnte aber auch schon Eve die Messung in der BELL-Basis ausführen.

## 3.6 Leitgedanken der Untersuchungen

In diesem Abschnitt soll ein Einblick in die Untersuchungen der folgenden drei Kapitel gegeben werden. Hierbei stehen die Gedankengänge im Vordergrund, die die späteren Rechnungen leiten.

### 3.6.1 Ergebnisse über tolerierbare Fehlerraten

Schon in den ersten Sicherheitsbeweisen der Quantenkryptographie in den 90er Jahren wird die Frage untersucht, bis zu welcher Fehlerrate ein Quantenkanal genutzt werden kann, um einen sicheren Schlüssel zu erzeugen. Die Antwort auf diese Frage ist abhängig vom genauen Protokoll und seinen Parametern, die in einzelnen Realisierungen angepaßt werden.

Ein erstes Ergebnis erzielte MAYERS [120] in seinem Beweis der Sicherheit des BB84-Protokolls, mit dem er zeigte, daß das Protokoll bis zu einer Fehler-rate von 7,4% sicher gestaltet werden kann. SHOR UND PRESKILL [167] erhöhten diese Rate auf 11,0%, und für das Protokoll mit sechs Zuständen

ermittelte LO [110] eine Fehlerrate von 12,7%. Mittels komplizierterer Verfahren wie dem engl. *noisy preprocessing* können noch höhere Fehlerraten erzielt werden; so errechneten KERN UND RENES [89] tolerierbare Fehlerraten von 12,93% bzw. 14,59%. Die obere Schranke für Protokolle, die nur Einweg-Kommunikation verwenden, liegt bei unter 15%, der Fehlerrate, die entsteht, wenn der gesendete Zustand in zwei gleiche Teile (fehlerbehaftet) „geklont“ wird.

Durch die Verwendung von Zweiweg-Kommunikation erreichten GOTTESMAN UND LO [64] tolerierbare Fehlerraten von 18,9% und 26,4% für das BB84-Protokoll, die von CHAU [30] auf 20,0% und 27,6% verbessert wurden. In meiner Diplomarbeit [138] führte ich eine systematische Untersuchung dieser Protokolle durch und zeigte unter anderem, daß die von CHAU [30] gefundenen Raten nicht ohne weiteres erhöht werden können (vgl. Abb. 3.1). Ähnliche Ergebnisse fand auch die Gruppe um ACÍN [1, 6, 7]. Für Protokolle, die Zweiweg-Kommunikation erlauben, liegen obere Schranken der tolerierbaren Fehlerrate bei 25% bzw.  $33,3\%$  (Verschränkungsgrenze von NIKOLOPOULOS u. a. [129]).

In dieser Dissertation sollen ähnliche Fragestellungen untersucht werden, wenn statt Qubits beliebige endlichdimensionale Quantensysteme (Qudits) zugelassen werden. Für Zwei-Basis-Protokolle auf diesen Quantensystemen zeigten NIKOLOPOULOS UND ALBER [128], daß bei einer Fehlerrate unter  $(d-1)/(2d)$  die Zustände im verschränkungsbasierten Protokoll zwingend verschränkt sind, so daß die Gewinnung eines Schlüssels möglich erscheint.

### 3.6.2 Untersuchungen der tolerierbaren Fehlerraten

Eines der Ziele dieser Dissertation ist es, die maximal tolerierbaren Fehlerraten bestimmter quantenkryptographischer Protokolle mittels allgemeiner Methoden zu untersuchen. Zu diesen Methoden gehören insbesondere die Verschränkungsreinigung mittels Einweg- und Zweiweg-Protokollen. Als einziges echtes Zweiweg-Protokoll wird der  $B_n^{(d)}$ -Schritt verwendet, und in der Literatur finden sich keine Protokolle, die sich von den  $B_n^{(d)}$ -Schritten wesentlich unterscheiden.

Die Untersuchungen zu Einweg-Protokollen werden nur in sehr allgemeinem Rahmen geführt. Im Kapitel 4 wird eine allgemeine Schranke für die Existenz asymmetrischer CSS-Codes betrachtet, in Kapitel 5 wird der wohl einfachste denkbare Code, ein Wiederholungscode verwendet. Anschließend wird in Kapitel 6 eine Möglichkeit untersucht, wie man ausschließen kann, daß Einweg-Verfahren überhaupt existieren. Die Ergebnisse dieser drei Kapitel zeigen, daß die Wahl des genauen Einweg-Verfahrens ziemlich bedeutungslos ist, da im wesentlichen drei Mal die gleichen Schranken auftreten.

### 3.6.3 Grundannahmen für die folgenden Kapitel

Nach den Feststellungen in Unterabschnitt 3.5.5 genügt es, statt allgemeiner Zustände das Produkt identisch präparierter bell-diagonaler Zustände zu betrachten. Ferner werden in dieser Dissertation in erster Linie *tolerierbare Fehlerraten* untersucht, während *Schlüssel(erzeugungs)raten* keine Berücksichtigung finden.<sup>13</sup> Für die folgenden Kapitel gelte aus diesen Gründen die folgende

**Generalvoraussetzung:** Alice und Bob teilen sich eine unbegrenzte Anzahl bell-diagonaler Zwei-Qudit-Zustände, die sämtlich durch die gleiche reduzierte Dichtematrix beschrieben werden. Diese werde stets mit  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  bezeichnet.

In den folgenden Kapiteln wird nun die Wirkung der  $B_n^{(d)}$ -Schritte (vgl. Unterabschnitt 4.2.1) auf einen gegebenen Zustand untersucht; fast alle derzeit bekannten echten Zweiweg-Verfahren sind Varianten dieses  $B_n^{(d)}$ -Schritts. Setzt man die Gültigkeit von  $A_{*0} > \max \{A_{*m} \mid m \in \mathbb{Z}_d^*\}$  voraus, was einer Gesamtfehlerrate von unter 50 % entspricht, so konvergiert ein Zustand unter einem  $B_n^{(d)}$ -Schritt für  $n \rightarrow \infty$  im allgemeinen gegen den separablen Zustand  $\sigma := (d^{-1}\delta_{m0})_{lm} \in \mathcal{S}_{\text{bd}}^{(d)}$ , in dessen Umgebung sich aber verschränkte Zustände befinden.

Die Grundidee der untersuchten Protokolle besteht darin, den Zustand durch einen  $B_n^{(d)}$ -Schritt mit genügend großem  $n \in \mathbb{N}$  nahe an den Grenz-zustand  $\sigma$  zu bringen und anschließend zu prüfen, ob der Zustand mit anderen Protokollen, im allgemeinen Einweg-Protokollen, gereinigt werden kann. Für die Existenz oder Nicht-Existenz solcher Einweg-Protokolle werden in den einzelnen Kapiteln unterschiedliche Schranken verwendet.

### 3.6.4 Symmetrierelationen

In quantenkryptographischen Protokollen ist es allgemein nicht möglich, die Koeffizientenmatrix  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  zu bestimmen, wenn man sie auf *Prepare-and-Measure*-Protokolle zurückführen will. Zugänglich ist nur die Ditfehlerverteilung  $(A_{*m})_{m=0}^{d-1} \in \mathcal{W}_d$ , von der auf den Zustand zurückgeschlossen werden muß.

In den beiden betrachteten Klassen von Protokollen wendet Alice unitäre Transformationen  $\mathbb{I} \otimes U$  auf das zu übermittelnde Qudit an (oder präpariert gleich in einer anderen Basis), die dazu dienen, Eve über die verwendete Basis im Unklaren zu lassen. Eve kennt zwar die Menge der möglichen  $U$ ,

---

<sup>13</sup>In der Literatur wird der Begriff *Raten* in beiden Zusammenhängen gebraucht.

nicht aber das verwendete  $U$  im Einzelfall. Infolgedessen wirkt ihr Angriff so, als ob über alle möglichen  $U$  gemittelt würde. Im verschränkungsbasierten Protokoll kann Alice auch ihr erstes Qudit transformieren, so daß – mit einer kleinen Änderung der Notation – über  $U \otimes U^*$  gemittelt wird, der bell-diagonale Zustand  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  also diese Invarianz besitzt. Einige Beispiele für mögliche Transformationen in Protokollen sind

1. im BB84-Protokoll  $\{\mathbb{1} \otimes \mathbb{1}, H \otimes H^*\}$  für die sogenannte HADAMARD-Transformation  $H = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} / \sqrt{2} \in \mathbb{C}^{2 \times 2}$ , die für Qubits Zustände der Form  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{\text{bd}}$  mit  $\beta = \gamma$  liefern;
2. im Protokoll mit sechs Zuständen  $\{\mathbb{1} \otimes \mathbb{1}, T \otimes T^*, T^2 \otimes T^{*2}\}$  für die Transformation  $T := H \cdot \text{diag}(1, -i) \in \mathbb{C}^{2 \times 2}$ , die ähnlich Zustände mit  $\beta = \gamma = \delta$  liefern (isotrope Zwei-Qubit-Zustände);
3. in allgemeineren Zwei-Basis-Protokollen  $\{\mathbb{1} \otimes \mathbb{1}, \mathcal{F} \otimes \mathcal{F}^*\}$ , die auf die Symmetrierelationen  $A_{lm} = A_{d-m,l} = A_{m,d-l} = A_{d-l,d-m}$  führen, was letztlich aus  $\mathcal{F}^4 = \mathbb{1}$  folgt; vgl. NIKOLOPOULOS UND ALBER [128].

Mittels komplizierterer Verfahren lassen sich auch in höheren Dimensionen isotrope Zustände erzeugen; vgl. CHAU [31]. Diese Protokolle sind dann Verallgemeinerungen des Protokolls mit sechs Zuständen.

### 3.6.5 Asymptotisch-exponentielle Gleichheit

Ein Konzept, das an verschiedenen Stellen dieser Arbeit verwendet wird, betrifft das Verhalten einer Funktion  $f : D \rightarrow \mathbb{R}_0^+$ , deren Definitionsbereich  $D \subseteq \mathbb{R}$  nach oben unbeschränkt ist. Das exponentielle Verhalten einer solchen Funktion kann durch eine Größe  $z(f) := \lim_{n \rightarrow \infty} \sqrt[n]{f(n)}$  beschrieben werden, wobei im folgenden angenommen wird, daß der Grenzwert existiert.

Zwei Funktionen  $f, g : D \rightarrow \mathbb{R}$  haben das gleiche exponentielle Verhalten, wenn  $z(f) = z(g)$  gilt; in diesem Fall nenne ich die Funktionen *asymptotisch-exponentiell gleich* und schreibe  $f \stackrel{\text{aeg}}{=} g$ . Gilt  $z(f) = 1$ , so werde die Funktion  $f$  als *subexponentiell* bezeichnet.

Eine Funktion kann nun in einen exponentiellen und einen subexponentiellen Anteil zerlegt werden, das heißt, sie kann in der Form  $f(n) = c(n) z(f)^n$  geschrieben werden, wobei sich  $c : D \rightarrow \mathbb{R}_0^+$  subexponentiell verhält; dies folgt daraus, daß  $c(n) := f(n) \cdot z^{-n}$  wegen  $z(c) = \lim_{n \rightarrow \infty} \sqrt[n]{f(n) \cdot z^{-n}} = z^{-1} \cdot \lim_{n \rightarrow \infty} \sqrt[n]{f(n)} = 1$  subexponentiell ist.

Man kann recht einfach zeigen, daß das exponentielle Verhalten gewisse nützliche Eigenschaften besitzt: so sind  $z(\alpha f + \beta g) = \max\{z(f), z(g)\}$  und  $z(f \cdot g) = z(f) \cdot z(g)$ ; mittels  $z(f^k) = z(f)^k$  und  $z(cz^n) = z$  erkennt man, daß Potenzen subexponentieller Funktionen auch subexponentiell sind.

## Kapitel 4

# Korrigierbarkeit von Quantenzuständen

In diesem Kapitel werden einige der im letzten Kapitel eingeführten quantenkryptographischen Protokolle in allgemeinem Zusammenhang untersucht und hierdurch Ergebnisse früherer Arbeiten (vgl. RANADE [138] sowie RANADE UND ALBER [139]) von zweidimensionalen auf höherdimensionale Quantensysteme verallgemeinert. Die Hauptergebnisse dieses Kapitels, die in einer Arbeit von RANADE UND ALBER [140] veröffentlicht wurden, sind dabei

- das (Fast-)Kriterium für asymptotische Korrigierbarkeit, das sich in Hauptsatz 4.5 findet,
- die Untersuchung der asymptotischen  $B_n^{(d)}$ -Korrigierbarkeit sowie die Berechnung des charakteristischen Exponenten  $r^{(d)}$  in Gleichung (4.31) nebst seiner Interpretation in Satz 4.7 sowie
- die Abschätzung von  $r^{(d)}$  für Zwei-Basis-Protokolle und die Berechnung der tolerierbaren Fehlerraten im Abschnitt 4.3.

Im gesamten Kapitel bezeichne  $d \in \mathbb{N} \setminus \{1\}$  die Hilbertraumdimension des betrachteten Qudit-Systems. Die in den Entropieausdrücken auftretenden Logarithmen werden durchgängig zur Basis  $d$  genommen, das heißt, die Information wird in Dits gemessen.

### 4.1 Asymptotische Korrigierbarkeit

In diesem Abschnitt wird der Begriff der asymptotischen Korrigierbarkeit für qudit-basierte Protokolle der Quantenkryptographie eingeführt und untersucht.

### 4.1.1 Die Quanten-Shannon-Schranke

Der folgende Satz liefert eine hinreichende Bedingung für die Existenz fehlerkorrigierender asymmetrischer<sup>1</sup> CSS-Codes; er ist eine einfache Folgerung eines viel allgemeineren Satzes von HAMADA [68], dessen Beweis bei weitem zu kompliziert ist, um hier wiedergegeben werden; vgl. aber auch KERN [88].

#### Satz 4.1 (Quanten-Shannon-Schranke von Hamada)

Es seien  $d$  eine Primzahl und  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$ . Es sei weiterhin

$$\text{AsymCSS}[(A_{lm})_{l,m=0}^{d-1}] := 1 - H_d[(A_{*m})_{m=0}^{d-1}] - H_d[(A_{l*})_{l=0}^{d-1}],$$

wobei  $H_d$  die SHANNON-Entropie auf  $\mathcal{W}_d$  mit dem Logarithmus zur Basis  $d$  bezeichne. Ist  $\text{AsymCSS}(\rho) > 0$ , so existiert ein asymmetrischer CSS-Code, der den Zustand  $\rho$  korrigieren kann.

*Beweis:* In Theorem 2 (Anhang B, S. 8321) seiner Arbeit zeigt HAMADA [68], daß es CSS-Codes gibt, deren Fehlerrate durch  $d^{-2\nu E}$  beschränkt ist, wobei  $\nu$  die Codelänge bezeichnet; man kann zeigen, daß die dort genannte Funktion  $E$  genau dann positiv ist, wenn  $\text{AsymCSS}[(A_{lm})_{l,m=0}^{d-1}] > 0$  ist, q. e. d.

### 4.1.2 Der Hauptsatz dieses Kapitels

Ein Korrekturschritt  $S_n^{(d)}$  ist eine Quantenoperation, die als Eingang einen Zustand  $\rho^{\otimes n}$  erhält und mit einer von  $\rho$  und  $n$  abhängigen, im allgemeinen nicht-verschwindenden Wahrscheinlichkeit  $N_n$  am Ausgang einen Zustand  $\rho'^{\otimes n'}$  liefert, wobei hier stets  $\rho, \rho' \in \mathcal{S}_{\text{bd}}^{(d)}$  angenommen wird; mit Wahrscheinlichkeit  $1 - N_n$  erfolgt ein Fehlschlag, und alle  $n$  Qudit-Paare müssen verworfen werden. Im allgemeinen ist  $n' \leq n$ , und man geht davon aus, daß  $\rho'$  „stärker verschränkt“ ist als  $\rho$ . Der folgende Begriff ist grundlegend für alle weiteren Überlegungen.

#### Definition 4.2 (Asymptotische Korrigierbarkeit)

Es seien  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  und  $(S_n^{(d)})_{n \in \mathbb{N}}$  eine Folge möglicher Schritte eines Protokolls zur Verschränkungsreinigung. Der Zustand  $\rho$  heie asymptotisch korrigierbar, falls ein  $N_0 \in \mathbb{N}$  derart existiert, daß für alle  $n \in \mathbb{N}$ ,  $n \geq N_0$ , die Ungleichung  $\text{AsymCSS}[S_n^{(d)}(\rho)] > 0$  erfüllt ist. Gibt es umgekehrt ein  $N_0 \in \mathbb{N}$  derart, daß für alle  $n \in \mathbb{N}$ ,  $n \geq N_0$ , die Ungleichung verletzt ist, so heie der Zustand  $\rho$  asymptotisch nicht-korrigierbar.<sup>2</sup>

<sup>1</sup>Das heißt, daß sich die Verteilungen der Dit- und Phasenfehler unterscheiden dürfen.

<sup>2</sup>Trägt man  $f_\rho(n) := \text{AsymCSS}[S_n^{(d)}(\rho)]$  als (stetig interpolierte) Funktion von  $n \in \mathbb{N}$  auf, so ist in der Regel die konstante Nullfunktion die Asymptote von  $f_\rho$ . Liegt nun  $f_\rho(n)$  für  $n \rightarrow \infty$  über dieser Asymptote, so wird  $\rho$  asymptotisch korrigierbar genannt.

Entsprechend der Definition gilt die asymptotische Korrigierbarkeit eines Zustands nur in bezug auf eine festgelegte Folge  $(S_n^{(d)})_{n \in \mathbb{N}}$  von Korrekturschritten; um diese anzudeuten, werde kurz *asymptotisch  $S_n^{(d)}$ -korrigierbar* geschrieben.

In dieser Arbeit werden keine Schlüsselerzeugungsraten betrachtet, die Werte  $n$  und  $n'$  brauchen also nicht explizit berücksichtigt zu werden. Dies ermöglicht es,  $S_n^{(d)}$  als (im allgemeinen nicht-lineare) Abbildung auf  $\mathcal{S}_{\text{bd}}^{(d)}$  zu betrachten, die  $\rho = (A_{lm})_{l,m=0}^{d-1}$  auf  $\rho' = (A'_{lm})_{l,m=0}^{d-1}$  abbildet.

Die folgenden beiden Lemmata dienen als Vorbereitung des *Kriteriums für asymptotische Korrigierbarkeit*, das im unmittelbar darauffolgenden Satz formuliert und bewiesen wird.

**Lemma 4.3 (Schranken für die Shannon-Entropie)**

Es seien  $\xi = (\xi_0, \dots, \xi_{d-1}) \in \mathcal{W}_d$  und  $x_n := 1 - \xi_0 = \sum_{i=1}^{d-1} \xi_i$ . Es bezeichnen ferner  $\xi_{\min} := (\xi_0, x_n, 0, \dots, 0)$  und  $\xi_{\max} := (\xi_0, \frac{x_n}{d-1}, \dots, \frac{x_n}{d-1})$ . Dann sind

$$H_d(\xi_{\min}) = -(\ln d)^{-1} [\xi_0 \ln \xi_0 + x_n \ln x_n],$$

$$H_d(\xi_{\max}) = -(\ln d)^{-1} \left[ \xi_0 \ln \xi_0 + x_n \ln \frac{x_n}{d-1} \right],$$

und es gilt  $H_d(\xi_{\min}) \leq H_d(\xi) \leq H_d(\xi_{\max})$ .

Mit geeigneten Funktionen  $c : \mathcal{W}_d \rightarrow [\frac{1}{d-1}; 1]$  und  $c' : \mathcal{W}_d \rightarrow [0; 1]$  ist also

$$\begin{aligned} H_d(\xi) &= -(\ln d)^{-1} [\xi_0 \ln \xi_0 + x_n \ln c(\xi)x_n] \\ &= -(\ln d)^{-1} [\xi_0 \ln \xi_0 + x_n \ln x_n] - (\ln d)^{-1} x_n \ln c(\xi) \\ &= L \cdot H(x_n) - x_n \log_d c(\xi) = L \cdot H(x_n) + c'(\xi)x_n \log_d(d-1), \end{aligned} \quad (4.1)$$

wobei  $L := \ln 2 / \ln d = \log_d 2$  gesetzt wurde.

*Beweis:* Die SHANNON-Entropie der Verteilung  $\xi$  ist per Definition invariant unter Permutationen der  $\xi_i$ . Es läßt sich nun  $\xi$  als Konvexkombination von Permutationen von  $\xi_{\min}$  darstellen, und ebenso läßt sich  $\xi_{\max}$  als Konvexkombination von Permutationen von  $\xi$  darstellen. Aus der Konkavität der SHANNON-Entropie (vgl. etwa Satz C.22) folgt die behauptete Aussage, q. e. d.

Um die Taylorreihe der SHANNON-Entropie zu berechnen, benötigt man die Ableitungen der Funktion  $f : [0; 1] \rightarrow \mathbb{R}$ ,  $f(x) := -x \log_d x = -(\ln d)^{-1} x \ln x$ ; für diese berechnet man

$$f'(x) = -(\ln d)^{-1} (1 + \ln x), \quad (4.2)$$

$$f^{(n)}(x) = (\ln d)^{-1} \left[ \frac{(-1)^{n+1} (n-2)!}{x^{n-1}} \right] \text{ für } n \geq 2. \quad (4.3)$$

Die Taylorreihe von  $f$  an der Stelle  $p_1 = 1/d$  ergibt sich hiermit zu

$$(\ln d \cdot f)(1/d + \tilde{p}) = \frac{\ln d}{d} + \tilde{p}(-1 + \ln d) + \sum_{k=2}^{\infty} \frac{(-d)^{k-1} \tilde{p}^k}{k(k-1)}, \quad (4.4)$$

die im Punkt  $p_2 = 1$  zu

$$(\ln d \cdot f)(1 + \Delta p) = -\Delta p + \sum_{k=2}^{\infty} \frac{(-1)^{k+1} (\Delta p)^k}{k(k-1)}. \quad (4.5)$$

Insbesondere gilt  $(\ln d \cdot f)(1 + \Delta p) \leq -\Delta p$  für  $\Delta p \leq 0$ .

**Lemma 4.4 (Eine Taylor-Entwicklung der Shannon-Entropie)**

Es seien  $p = (p_0, \dots, p_{d-1}) \in \mathcal{W}_d$  eine beliebige Wahrscheinlichkeitsverteilung und  $g := (1/d, \dots, 1/d) \in \mathcal{W}_d$  die Gleichverteilung auf einer Menge mit  $d$  Elementen. Vorausgesetzt, daß es einen Faktor  $f > 0$  derart gibt, daß die Ungleichung  $p_i \geq f/d$  für alle  $i \in \mathbb{Z}_d$  gilt, so ist

$$H_d(p) = 1 - K \|g - p\|_2^2 + K' \varepsilon(p) \|g - p\|_3^3$$

für Zahlen  $K, K' > 0$  und eine beschränkte Funktion  $\varepsilon : \mathcal{W}_d \rightarrow [-1; 1]$ .

Allgemeiner läßt sich die  $p$ -te Ordnung der Taylorentwicklung durch die zugehörige  $p$ -Norm (vgl. Unterabschnitt B.8.1) abschätzen; im Falle geradzahlicher Ordnungen gilt sogar die Gleichheit.

*Beweis:* Eine Taylorentwicklung von  $H_d$  bis zur zweiten Ordnung um  $g$  ergibt

$$H_d(p) = 1 + \sum_{i=0}^{d-1} \left(1 - \frac{1}{\ln d}\right) (p_i - g_i) - \frac{d}{2 \ln d} \sum_{i=0}^{d-1} (g_i - p_i)^2 + R_2(p). \quad (4.6)$$

Die erste Summe verschwindet, weil  $p$  und  $g$  Wahrscheinlichkeitsverteilungen sind, und der Term der zweiten Ordnung kann mit  $K := d/(2 \ln d)$  in der im Lemma genannten Form geschrieben werden. Das Restglied ergibt sich in der LAGRANGESchen Form (vgl. Unterabschnitt B.9.1) zu

$$R_2(p) = \sum_{i=0}^{d-1} \frac{\tilde{p}_i^{-2}}{3! \cdot \ln d} (p_i - g_i)^3 \quad (4.7)$$

für geeignete Werte  $\tilde{p}_i$ , wobei  $p_i \leq \tilde{p}_i \leq 1/d$  oder  $1/d \leq \tilde{p}_i \leq p_i$  für jedes  $i \in \mathbb{Z}_d$  gilt. Nach Voraussetzung ist  $\tilde{p}_i \geq f/d$ , und dies führt auf

$$|R_2(p)| \leq \sum_{i=0}^{d-1} \frac{(f/d)^{-2}}{3! \cdot \ln d} |p_i - g_i|^3 = K' \|p - g\|_3^3, \quad (4.8)$$

wenn  $K' := d^2 \cdot (3! f^2 \cdot \ln d)^{-1}$  gesetzt wird, q. e. d.

### Hauptsatz 4.5 (Asymptotische Korrigierbarkeit)

Es sei  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ein bell-diagonaler Zwei-Qudit-Zustand, auf den für jedes  $n \in \mathbb{N}$  ein (fiktiver)  $S_n^{(d)}$ -Schritt angewendet werde; der entstehende Zustand sei  $\rho' = (A'_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$ , und es bezeichnen

- $x_n := \sum_{m=1}^{d-1} A'_{*m} = \sum_{m=1}^{d-1} \sum_{l=0}^{d-1} A'_{lm}$  die gesamte Ditfehlerrate und
- $y_n := \|g - p\|_2 / \sqrt{2}$  ein Maß für die Abweichung der Phasenfehlerrate  $p = (A'_{*l})_{l=0}^{d-1}$  von der Gleichverteilung  $g = (1/d, \dots, 1/d) \in \mathcal{W}_d$ .<sup>3</sup>

Ist die Folge  $(x_n)_{n \in \mathbb{N}}$  eine Nullfolge, so gelten die folgenden Aussagen:

1. Existiert ein  $r > 2$  derart, daß  $\sup \{x_n/y_n^r \mid n \in \mathbb{N}\} < \infty$  ist, so ist  $\rho$  asymptotisch  $S_n^{(d)}$ -korrigierbar.
2. Ist hingegen  $\inf \{x_n/y_n^2 \mid n \in \mathbb{N}\} > 0$ , so ist  $\rho$  asymptotisch  $S_n^{(d)}$ -nicht-korrigierbar.

Die Aussagen gelten entsprechend, wenn die Rollen von Dit- und Phasenfehlern vertauscht werden.

Die Grundidee des Satzes ist, daß  $x_n$  und  $y_n$  beide exponentiell gegen Null streben; das relative Verhalten dieser Größen wird durch einen Exponenten  $r$  beschrieben, für den  $x_n \stackrel{\text{aeg}}{\asymp} y_n^r$  gilt. Ist dieser Exponent größer als  $r_{\text{Grenz}} = 2$ , so ist der Zustand asymptotisch korrigierbar, ist er kleiner, dann nicht. Der Fall  $r = 2$  müßte gesondert betrachtet werden, ist aber praktisch ohne Belang.

*Beweis:* Es werden Taylorentwicklungen der Quanten-SHANNON-Schranke (Satz 4.1) von HAMADA betrachtet. Nach Lemma 4.3 gilt für die Verteilung  $\xi = (\xi_0, \dots, \xi_{d-1}) \in \mathcal{W}_d$  der Ditfehler

$$H_d(\xi) = L \cdot H(x_n) + c''(\xi) x_n \geq L \cdot H(x_n) \quad (4.9)$$

mit  $L = \ln 2 / \ln d = \log_d 2$  und  $c''(\xi) \in [0; \log_d(d-1)] \subseteq [0; 1]$ . Nach dem Lemma 4.4 gilt für die Verteilung  $p$  der Phasenfehler

$$\begin{aligned} H_d(p) &= 1 - K \cdot 2y_n^2 + K' \varepsilon(p) \|g - p\|_3^3 \\ &= 1 - K \cdot 2y_n^2 + K' \varepsilon'(p) \cdot (2y_n^2)^{3/2} \end{aligned} \quad (4.10)$$

mit  $K = d/(2 \ln d)$  und Funktionen  $\varepsilon, \varepsilon' : \mathcal{W}_d \rightarrow [-1; 1]$ , vorausgesetzt,  $y_n$  ist hinreichend klein; die zweite Gleichheit folgt aus  $(2y_n)^3 = \|g - p\|_2^3 \geq \|g - p\|_3^3$ . Mit  $\rho' := (A'_{lm})_{l,m=0}^{d-1}$  gilt somit

$$\begin{aligned} \text{AsymCSS}(\rho') &= 1 - H_d(\xi) - H_d(p) \\ &= -L \cdot H(x_n) - c''(\xi) x_n + 2K \cdot y_n^2 - 2\sqrt{2}K' \varepsilon'(p) \cdot y_n^3, \end{aligned} \quad (4.11)$$

---

<sup>3</sup>Der Faktor bei  $y_n$  dient ausschließlich der Konsistenz mit dem Qubit-Fall [138, 139].

und nach Division durch  $y_n^2$  erhält man

$$\text{AsymCSS}(\rho') > 0 \Leftrightarrow \frac{-L \cdot H(x_n)}{y_n^2} - c''(\xi) \frac{x_n}{y_n^2} + 2K - 2\sqrt{2}K' \varepsilon'(p) \cdot y_n > 0. \quad (4.12)$$

Die Regel von L'HÔPITAL zeigt, daß

$$\lim_{x \rightarrow 0^+} H(x)/x^s = \begin{cases} 0 & \text{für } s \in [0; 1) \\ +\infty & \text{für } s \in [1; \infty) \end{cases} \quad (4.13)$$

gilt. Die Bedingung aus Nr. 1 impliziert, daß  $x_n \leq cy_n^r$  für ein geeignetes  $c \geq 0$  ist, woraus wegen  $r > 2$  nun  $\lim_{n \rightarrow \infty} -L \cdot H(x_n)/y_n^2 \geq -L \cdot c^{2/r} H(x_n)/x^{2/r} = 0$  folgt. Dies bedeutet, daß in Formel (4.12) alle Terme mit Ausnahme von  $2K$  gegen Null streben. Für den Beweis der Nr. 2 stellt man fest, daß  $x_n \geq cy_n^2$  für ein geeignetes  $c > 0$  gilt. Ähnlich wie zuvor ergibt dies eine Ungleichung  $-L \cdot H(x_n)/y_n^2 \leq -L \cdot c \cdot H(x_n)/x \rightarrow -\infty$ . Der zweite Term ist negativ, alle anderen sind nach oben beschränkt, so daß im Fall  $n \rightarrow \infty$  die Quanten-SHANNON-Schranke nicht erfüllt wird, q. e. d.

### 4.1.3 Die von-Neumann-Entropie als Schranke

Gelegentlich wird für eine Dichtematrix  $\rho = (A_{lm})_{l,m=0}^{d-1}$  anstelle der Schranke aus Satz 4.1 der Ausdruck

$$1 - S(\rho) = 1 - S[(A_{lm})_{l,m=0}^{d-1}] = 1 - \left[ - \sum_{l,m=0}^{d-1} A_{lm} \log_d A_{lm} \right] \quad (4.14)$$

als Schranke für Codes verwendet; vgl. auch KERN [88].<sup>4</sup> Für eine Untersuchung bietet es sich an, die VON-NEUMANN-Entropie an der Stelle  $\sigma = (d^{-1}\delta_{m0})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  zu entwickeln; dies führt auf

$$\begin{aligned} 1 - S(\rho) &= 1 + \left[ \sum_{l=0}^{d-1} A_{l0} \log_d A_{l0} + \sum_{l=0}^{d-1} \sum_{m=1}^{d-1} A_{lm} \log_d A_{lm} \right] \\ &\approx \left( 1 - \frac{1}{\ln d} \right) (A_{*d} - 1) + \dots + x_n \log_d x_n, \end{aligned} \quad (4.15)$$

also im wesentlichen auf das gleiche Ergebnis wie zuvor. Das bedeutet, am Grenzexponenten  $r_{\text{Grenz}} = 2$  sollte sich nichts ändern. Im Kapitel 6 wird die Frage anderer Schranken von einem wesentlich allgemeineren Blickwinkel betrachtet, weswegen hier auf die genauen Ausführungen verzichtet wird.

---

<sup>4</sup>Faßt man eine bell-diagonale Dichtematrix als klassische gemeinsame Wahrscheinlichkeitsverteilung von Dit- und Phasenverteilungen auf, so ist die VON-NEUMANN-Entropie die zugehörige Verbundentropie. Die angegebene Schranke unterscheidet sich somit nur um die gemeinsame Information zwischen Dit- und Phasenverteilung von der aus Satz 4.1.

## 4.2 Asymptotische $B_n^{(d)}$ -Korrigierbarkeit

In diesem Abschnitt werden der  $B_n^{(d)}$ -Schritt und seine grundlegenden Eigenschaften eingeführt. Darauf aufbauend wird die asymptotische  $B_n^{(d)}$ -Korrigierbarkeit untersucht, die in Analogie zum Qubit-Fall [138, 139] durch einen *charakteristischen Exponenten*  $r^{(d)}$  beschrieben werden kann.<sup>5</sup>

### 4.2.1 Der $B_n^{(d)}$ -Schritt

Ein  $B_n^{(d)}$ -Schritt ist die Verallgemeinerung des  $B_n$ -Schrittes [138, 139] und wird für ein  $n \in \mathbb{N}$  wie folgt ausgeführt:<sup>6</sup>

1. Alice und Bob wählen zufällig  $n$  Qudit-Paare  $QP_1, \dots, QP_n$  aus.
2. Alice und Bob führen insgesamt  $n - 1$  GBXOR-Schritte der Form  $\text{GBXOR}(QP_1, QP_k)$  mit  $k \in \{2, \dots, n\}$  aus.
3. Alice und Bob messen an  $QP_2$  bis  $QP_n$  die Ditparität<sup>7</sup> und verwenden  $QP_1$  genau dann weiter, wenn alle Meßwerte gleich sind. Die Paare  $QP_2, \dots, QP_n$  werden verworfen.

Anders formuliert gilt für den zweiten Schritt

$$\bigotimes_{i=1}^n (l_i, m_i) \mapsto \left( \bigoplus_{i=1}^n l_i, m_1 \right) \otimes \left[ \bigotimes_{k=2}^n (l_k, m_1 \ominus m_k) \right]. \quad (4.16)$$

Falls nun  $m_1 \ominus m_k = 0$  für alle  $k \in \{2, \dots, n\}$  gilt, so wird das erste Paar  $QP_1$  weiterverwendet, sonst verworfen; dies führt dazu, daß verschiedene Spalten in der Koeffizientenmatrix  $(A_{lm})_{l,m=0}^{d-1}$  sich nicht mischen.<sup>8</sup>

Es gilt ferner  $B_n^{(d)} B_m^{(d)} = B_{nm}^{(d)}$ , und eine Sequenz von  $k$  hintereinander ausgeführten  $B_2^{(d)}$ -Schritten entspricht einem einzigen  $B_n^{(d)}$ -Schritt mit  $n = 2^k$ .

<sup>5</sup>Der  $B_n^{(d)}$ -Schritt beruht im wesentlichen auf dem von GOTTESMAN UND LO [64] eingeführten  $B$ -Schritt, der wiederum das quantenmechanische Analogon zur engl. *advantage distillation* von MAURER [119] ist. Ein verwandtes klassisches *effizientes* Zweiweg-Verfahren ist „Cascade“; vgl. BRASSARD UND SALVAIL [24] sowie HEID UND LÜTKENHAUS [71], für die Effizienz speziell die Tabellen 1 in beiden Texten; vgl. auch LO [111].

<sup>6</sup>Die etwas widersinnige Bezeichnung des  $B_n^{(d)}$ -Schrittes erklärt sich daraus, daß er eine Verallgemeinerung des  $B_n$ -Schrittes zur *Bitfehlerkorrektur* ist; in  $d$  Dimensionen könnte man auch von einem  $D_n$ -Schritt sprechen; vgl. z. B. NIKOLOPOULOS u. a. [130] für  $n = 2$ .

<sup>7</sup>Dies geschieht durch lokale Messungen von  $Z_A \otimes \mathbb{I}_B$  und  $\mathbb{I}_A \otimes Z_B$ ; durch Austausch ihrer Meßergebnisse (mit Zweiweg-Kommunikation) erschließen sie die Ditparität  $Z_A \otimes Z_B$ .

<sup>8</sup>Man kann durch Variationen der  $B_n^{(d)}$ -Schritte die Schlüsselrate wesentlich steigern, zum Beispiel, indem man die Messung in Schritt 3 am  $k$ -ten Paar unmittelbar nach der GBXOR-Transformation durchführt und das erste Paar bereits verwirft, wenn hier ein Fehler auftritt; dies soll hier aber nicht untersucht werden.

## 4.2.2 Die Entwicklung unter $B_n^{(d)}$ -Schritten

Die allgemeine Entwicklung bell-diagonaler Zustände unter  $B_n^{(d)}$ -Schritten wird durch den folgenden Satz bestimmt, dessen einfacher, aber langwieriger Beweis sich in Anhang A.1 (S. 129ff.) findet.

### Satz 4.6 (Allgemeine Entwicklung unter $B_n^{(d)}$ -Schritten)

Für jedes  $k \in \{1, \dots, n\}$  sei  $\rho^{(k)} = (A_{lm}^{(k)})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ein Zustand, auf deren Gesamtheit ein  $B_n^{(d)}$ -Schritt angewendet werde. Falls nicht alle Qudit-Paare verworfen werden, sei  $\rho' = (A'_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  der Zustand des verbleibenden Paares; seine Koeffizienten ergeben sich zu

$$A'_{lm} = (dN)^{-1} \sum_{i=0}^{d-1} \left[ z^{-il} \prod_{k=1}^n \left( \sum_{j=0}^{d-1} z^{ij} A_{jm}^{(k)} \right) \right].$$

Die Normierungskonstante  $N := \sum_{m=0}^{d-1} \left[ \prod_{k=1}^n \left( \sum_{l=0}^{d-1} A_{lm}^{(k)} \right) \right]$  ist hierbei die Wahrscheinlichkeit dafür, daß das erste Qudit-Paar weiterverwendet wird. Insbesondere ist der entstehende Zustand bell-diagonal und invariant unter Permutationen der Anfangszustände.

Die Idee hinter diesem Satz ist, daß sich die Dichtefehlerraten durch Potenzieren und Renormieren entwickeln; vgl. die Gleichung (4.18), die man auch direkt herleiten kann. Die Phasenfehlerentwicklung kann als eine Faltung interpretiert werden, die durch eine Sequenz aus Fouriertransformation, Multiplikation und inverser Fouriertransformation berechnet werden kann, was auf die genannte Entwicklung führt.

In den im folgenden betrachteten Fällen sind im allgemeinen alle Dichtematrizen gleich, also  $\rho^{(k)} =: \rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  für alle  $k \in \{1, \dots, n\}$ ; die Entwicklung vereinfacht sich dann zu

$$A'_{lm} = (dN)^{-1} \sum_{i=0}^{d-1} \left[ z^{-il} \left( \sum_{j=0}^{d-1} z^{ij} A_{jm} \right)^n \right]. \quad (4.17)$$

Für die folgenden Betrachtungen wird stets angenommen, daß die Bedingung  $A_{*0} > \max \{A_{*m} \mid m \in \mathbb{Z}_d^*\}$  erfüllt ist; andernfalls kann die Transformation  $\mathbb{1} \otimes X^{d-m}$  ausgeführt werden, wenn  $A_{*m} > \max \{A_{*j} \mid j \in \mathbb{Z}_d \setminus \{m\}\}$  ist. Es wird weiterhin angenommen, daß die Phasenfehler-Randverteilung unter  $B_n^{(d)}$ -Schritten für  $n \rightarrow \infty$  gegen die Gleichverteilung konvergiert.<sup>9</sup>

---

<sup>9</sup>Im Falle, daß der größte Wert  $A_{*m}$  nicht eindeutig ist, versagt das Korrekturverfahren. Die Aussage über die Phasenfehlerverteilung ist ähnlich: sie konvergiert gegen die Gleichverteilung, wenn es eine eindeutig bestimmte betragsgrößte Komponente einer Art „Spalten-Fouriertransformierten“ in der Koeffizientenmatrix gibt. Die nicht betrachteten Fälle sind kompliziert, und die betreffenden Zustände bilden eine Nullmenge in  $\mathcal{S}_{\text{bd}}^{(d)}$ .

### 4.2.3 Die Entwicklung der Ditfehler

Es sei  $\xi = (\xi_0, \dots, \xi_{d-1}) \in \mathcal{W}_d$  mit  $\xi_m := A_{*m}$ ,  $m \in \mathbb{Z}_d$ , die Verteilung der Ditfehler. Die Entwicklung  $B_n^{(d)} : \mathcal{W}_d \rightarrow \mathcal{W}_d$  mit  $\xi \mapsto \xi'$  bestimmt sich durch

$$\xi'_i = \frac{\xi_i^n}{\xi_0^n + \dots + \xi_{d-1}^n} \quad (4.18)$$

für alle  $i \in \mathbb{Z}_d$ ; vgl. Satz 4.6. Es gilt also insbesondere

$$x_n := 1 - \xi'_0 = \frac{\sum_{m=1}^{d-1} \xi_m^n}{\sum_{m=0}^{d-1} \xi_m^n} = \left[ \frac{\sum_{m=0}^{d-1} \xi_m^n}{\sum_{m=1}^{d-1} \xi_m^n} \right]^{-1} = \left[ 1 + \frac{\xi_0^n}{\sum_{m=1}^{d-1} \xi_m^n} \right]^{-1}. \quad (4.19)$$

Setzt man  $\xi_{\max} := \max \{\xi_m \mid m \in \mathbb{Z}_d^*\}$ , so gilt für den Nenner in der letzten Formel

$$\xi_{\max}^n \leq \sum_{m=1}^{d-1} \xi_m^n \leq (d-1)\xi_{\max}^n, \quad (4.20)$$

und man erkennt, daß der Fall  $\xi_1 = \xi_2 = \dots = \xi_{d-1}$  der *bestmögliche* Fall ist. Mit einer geeigneten Funktion  $h : \mathcal{W}_d \rightarrow [1; d-1]$  und  $\tilde{x} := \xi_0/\xi_{\max} > 1$  gilt somit

$$x_n = \left[ 1 + \frac{\xi_0^n}{h(\xi)\xi_{\max}^n} \right]^{-1} \stackrel{\text{aeg}}{=} \tilde{x}^{-n}, \quad (4.21)$$

die Entwicklung der Ditfehler ist also unabhängig von etwaigen Phasenfehlern, und es gilt  $\lim_{n \rightarrow \infty} \xi'_m = \delta_{m0}$ .

### 4.2.4 Die Entwicklung der Phasenfehler

Die Entwicklung der Phasenfehler unter  $B_n^{(d)}$ -Schritten ist wesentlich komplizierter als die der Ditfehler. Im Hauptsatz 4.5 wird jedoch nur der Parameter  $2y_n^2 = \|g - p\|_2^2$  benötigt, wobei  $p = (A_{0*}, \dots, A_{d-1,*})$  die Phasenfehlerverteilung und  $g = (1/d, \dots, 1/d)$  die Gleichverteilung bezeichnen; für diesen Parameter berechnet man

$$2y_n^2 = \|g - p\|_2^2 = \left\| \left( \frac{1}{d} - \frac{\sum_m \sum_i z^{-il} \left( \sum_j z^{ij} A_{jm} \right)^n}{dN} \right)_{l=0}^{d-1} \right\|_2^2. \quad (4.22)$$

Die euklidische Norm (2-Norm) ist unter der diskreten Fouriertransformation  $(x_i)_i \mapsto (d^{-1/2} \sum_i z^{ij} x_i)_j$  invariant, und wegen  $\sum_j z^{ij} = d \cdot \delta_{i0}$  folgt

$$2y_n^2 = \frac{1}{d} \left\| \left( \delta_{l0} - \frac{\sum_m \left( \sum_j z^{lj} A_{jm} \right)^n}{N} \right)_{l=0}^{d-1} \right\|_2^2. \quad (4.23)$$

Man stellt nun fest, daß sich das Kroneckersymbol  $\delta_{l0}$  gegen den Term für  $l = 0$  kürzt und erhält somit

$$2y_n^2 = \frac{1}{d} \left\| \left\| \left( \frac{\sum_m \left( \sum_j z^{lj} A_{jm} \right)^n}{N} \right)_{l=1}^{d-1} \right\|_2 \right\|^2. \quad (4.24)$$

Diese Größe ist im allgemeinen schwierig zu berechnen, man vermutet aber, daß für  $n \rightarrow \infty$  nur die Koeffizienten  $A_{lm}$  mit  $m = 0$  relevant sind, die die erste Spalte der Koeffizientenmatrix bilden. Dies wird im folgenden für bestimmte Dimensionen sehr elegant und explizit gezeigt werden.

### 4.2.5 Eine neue Transformation

Für ein Qudit-System seien  $U_1 := \sum_{x=0}^{d-1} z^{-x^2} |x\rangle\langle x|$  und

$$U := U_1 \otimes U_1^* = \sum_{x,y=0}^{d-1} z^{y^2-x^2} |x\rangle\langle x| \otimes |y\rangle\langle y|. \quad (4.25)$$

Die Anwendung dieser Transformation auf den BELL-Zustand  $|\Psi_{lm}\rangle$  liefert

$$\begin{aligned} U|\Psi_{lm}\rangle &= d^{-1/2} \sum_{x,y,k} z^{y^2-x^2} z^{lk} |x\rangle|y\rangle \langle x|k\rangle \langle y|k \ominus m\rangle \\ &= d^{-1/2} \sum_k z^{(k-m)^2-k^2+lk} |k\rangle|k \ominus m\rangle \\ &= d^{-1/2} z^{m^2} \sum_k z^{(l-2m)k} |k\rangle|k \ominus m\rangle = z^{m^2} |\Psi_{l \ominus 2m, m}\rangle, \end{aligned} \quad (4.26)$$

also für Dichtematrizen  $U : (l, m) \mapsto (l \ominus 2m, m)$ . Die Anwendung auf  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  führt daher in der Koeffizientenmatrix Permutationen innerhalb der Spalten aus. Entsprechend bewirkt eine Sequenz  $(\mathcal{F} \otimes \mathcal{F}^*) \rightarrow U \rightarrow (\mathcal{F} \otimes \mathcal{F}^*)^{-1}$  die Transformation  $(l, m) \mapsto (l, m \oplus 2l)$ , also Permutationen in den Zeilen.<sup>10</sup>

### 4.2.6 Ein modifiziertes Protokoll

Mithilfe der im letzten Unterabschnitt eingeführten Transformation  $U$  kann das Protokoll nun wie folgt modifiziert werden: bevor sie einen  $B_n^{(d)}$ -Schritt ausführen, wählen Alice und Bob zufällig ein  $n \in \mathbb{Z}_d$  aus und wenden  $U^n$  an. Dies bewirkt  $\rho \mapsto d^{-1} \sum_{n=0}^{d-1} U^n \rho (U^\dagger)^n$ , also eine Durchmischung innerhalb der Spalten der Koeffizientenmatrix, wobei die Spalte mit  $m = 0$  jedoch unverändert bleibt.

---

<sup>10</sup>Die Transformation  $U$  läßt sich ganz einfach auch auf die BELL-Basis für Primzahlpotenzen erweitern; hierzu setzt man  $U_1 := \sum_x z^{-\text{Spur}(x^2)} |x\rangle\langle x|$ , und vollkommen analog zur Rechnung in Formel (4.26) folgt  $U|\Psi'_{lm}\rangle = \sum_{x,y} z^{\text{Spur}(y^2-x^2)} = z^{\text{Spur}(m^2)} |\Psi'_{l \ominus 2m, m}\rangle$ . Es gilt auch hier  $U : (l, m) \mapsto (l \ominus 2m, m)$  für  $l, m \in \mathbb{F}_{p^n}$  mit  $d = p^n$ .

Betrachtet man genauer die Spalte zu einem Wert  $m \in \mathbb{Z}_d$ , so entstehen genau  $\text{ggT}(2m, d)$  verschiedene Zyklen in der entsprechenden Spalte. Im besten Fall ist also  $\text{ggT}(2m, d) = 1$  für alle  $m \in \mathbb{Z}_d^*$ , d. h. alle Spalten mit  $m \neq 0$  werden vollständig durchmischt<sup>11</sup>; dies ist genau für alle ungeraden Primzahlen der Fall.<sup>12</sup> Nach der Reduktion auf *Prepare-and-Measure*-Protokolle kann man die Mischoperation ganz weglassen, da sie die Ditwerte unverändert läßt. Für ein beliebiges  $l \in \mathbb{Z}_d$  berechnet man nun

$$\begin{aligned} \sum_{m=0}^{d-1} \left( \sum_{j=0}^{d-1} z^{lj} A_{jm} \right)^n &= \left( \sum_{j=0}^{d-1} z^{lj} A_{j0} \right)^n + \sum_{m=1}^{d-1} \left( \sum_{j=0}^{d-1} z^{lj} \frac{A_{*m}}{d} \right)^n \\ &= \left( \sum_{j=0}^{d-1} z^{lj} A_{j0} \right)^n + \sum_{m=1}^{d-1} \left( \frac{A_{*m}}{d} \right)^n \underbrace{\left( \sum_{j=0}^{d-1} z^{lj} \right)^n}_{=d^n \cdot \delta_{l0}}, \end{aligned} \quad (4.27)$$

und wegen  $l \neq 0$  folgt weiter (die Normierung  $N$  bleibt unverändert)

$$2y_n^2 \cdot dN^2 = \left\| \left( \left( \sum_j z^{lj} A_{j0} \right)^n \right)_{l=1}^{d-1} \right\|_2^2 = \left\| \left( \sum_j z^{lj} A_{j0} \right)_{l=1}^{d-1} \right\|_{2n}^{2n}. \quad (4.28)$$

Mit geeigneten Faktoren  $K(n) \in [1; d]$  ist also

$$2y_n^2 \cdot dN^2 = K(n) \left\| \left( \sum_j z^{lj} A_{j0} \right)_{l=1}^{d-1} \right\|_{\infty}^{2n}, \quad (4.29)$$

und für die Maximumsnorm gilt bekanntlich (vgl. Unterabschnitt B.8.2)

$$\sqrt{d} \left\| \left( \mathcal{F}(A_{j0})_j \right)_{l=1}^{d-1} \right\|_{\infty} = \left\| \left( \sum_j z^{lj} A_{j0} \right)_{l=1}^{d-1} \right\|_{\infty} = \max \left\{ \left| \sum_j z^{lj} A_{j0} \right| \mid l \in \mathbb{Z}_d^* \right\},$$

wobei  $\mathcal{F}$  hier die diskrete Fouriertransformation auf dem Raum  $\mathbb{C}^d$  bezeichnet. Das Problem entspricht der allgemeinen Suche nach dem betragsmäßig zweitgrößten Wert der Fouriertransformierten einer Wahrscheinlichkeitsverteilung. Die Fouriertransformierte einer Wahrscheinlichkeitsverteilung nennt man oft auch ihre *charakteristische Funktion*; vgl. z. B. Kapitel 12 bei KAMMLER [87]. Eine genauere Untersuchung erfolgt im nächsten Abschnitt.

---

<sup>11</sup>Dies bedeutet  $(A_{lm})_{l,m=0}^{d-1} \mapsto (A'_{lm})_{l,m=0}^{d-1}$  mit  $A'_{l0} = A_{l0}$  und  $A'_{lm} = A_{*m}/d$  für  $m \neq 0$ . Ist  $d$  eine Primzahlpotenz, so überträgt sich diese Überlegung allerdings *nicht* unmittelbar, selbst wenn man die BELL-Basis und die Mischoperation mithilfe endlicher Körper  $\mathbb{F}_{p^n}$  definiert.

<sup>12</sup>Wegen der Verwendung der Quanten-SHANNON-Schranke (Satz 4.1) und der abschließenden Behandlung für  $d = 2$  in früheren Arbeiten [138, 139] kommt dieser Einschränkung keine größere Bedeutung zu.

### 4.2.7 Exponentielles Verhalten

Es wurde bisher gezeigt, daß  $x_n \stackrel{\text{aeg}}{=} \tilde{x}^{-n}$  mit  $\tilde{x} = A_{*0} / \max \{A_{*m} \mid m \in \mathbb{Z}_d^*\}$  gilt. Für die Normierungskonstante des  $B_n^{(d)}$ -Schrittes ist also  $N_n = K'(n) A_{*0}^n$  für geeignete  $K'(n) \in [1; d]$ , und daher gilt auch

$$2y_n^2 = \frac{1}{d} \cdot \frac{K(n)}{K'(n)^2} \cdot \left( \frac{\max \left\{ \left| \sum_j z^{lj} A_{j0} \right| \mid l \in \mathbb{Z}_d^* \right\}}{A_{*0}} \right)^{2n} =: \frac{K(n)}{K'(n)^2} \cdot \frac{\tilde{y}^{2n}}{d}; \quad (4.30)$$

insbesondere ist also  $2y_n^2 \stackrel{\text{aeg}}{=} \tilde{y}^{2n}$ . Die Forderung  $x_n \stackrel{\text{aeg}}{=} y_n^{r^{(d)}}$  führt deswegen zu  $\tilde{x}^{-n} = \tilde{y}^{rn}$  bzw.  $-\ln \tilde{x} = r \ln \tilde{y}$  und hierdurch weiter auf<sup>13</sup>

$$r^{(d)} = -\frac{\ln \tilde{x}}{\ln \tilde{y}} = \frac{\ln [A_{*0} / \max \{A_{*m} \mid m \in \mathbb{Z}_d^*\}]}{\ln \left[ A_{*0} / \max \left\{ \left| \sum_j z^{lj} A_{j0} \right| \mid l \in \mathbb{Z}_d^* \right\} \right]}. \quad (4.31)$$

Im Falle von  $d = 2$  ist dies der Exponent  $r$  für Qubits [138, 139]. Der Vollständigkeit halber soll noch die Bedeutung des *charakteristischen Exponenten*  $r^{(d)}$  explizit formuliert und bewiesen werden.

#### Satz 4.7 (Asymptotische $B_n^{(d)}$ -Korrigierbarkeit)

Ein vorgegebener Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ist genau dann asymptotisch  $B_n^{(d)}$ -korrigierbar, wenn  $r^{(d)} > 2$  ist.

*Beweis:* Setzt man zur Vereinfachung  $r := r^{(d)}$ , so folgt unter Verwendung der Formeln (4.21) und (4.30) die Gleichung

$$\frac{x_n}{y_n^r} = u(n) \tilde{x}^{-1} \cdot \left( \tilde{y}^{2n} \frac{K(n)}{2dK'(n)} \right)^{-r/2} = \frac{u(n)}{(\tilde{x} \cdot \tilde{y}^r)^n} \left( \frac{K(n)}{2dK'(n)} \right)^{-r/2}. \quad (4.32)$$

Der charakteristische Exponent  $r^{(d)}$  ist derart gewählt, daß  $(\tilde{x} \cdot \tilde{y}^r)^n = 1$  und insbesondere auch  $x_n/y_n^r \stackrel{\text{aeg}}{=} 1$  gelten. Alle anderen Terme sind für  $n \in \mathbb{N}$  beschränkt, wobei die untere Schranke größer als Null ist. Aus Hauptsatz 4.5 folgt die Behauptung, q. e. d.

---

<sup>13</sup>Es ist wahrscheinlich, daß der Exponent auch ohne die vorgeschaltete Mischoperation die genannte Form hat, da die Werte der durchmischten Spalten unter  $B_n^{(d)}$ -Schritten für  $n \rightarrow \infty$  allesamt verschwinden. Aber selbst wenn dies so wäre, müßte man noch die Quanten-SHANNON-Schranke auf Nicht-Primzahldimensionen verallgemeinern, um allgemeinere Aussagen zu erhalten. Wegen der Verwendung von CSS-Codes erscheint eine Verallgemeinerung auf Primzahlpotenzen möglich, die auf allgemeinere Dimensionen aber nicht.

## 4.3 Quantenkryptographische Anwendungen

In quantenkryptographischen Protokollen ist  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  im allgemeinen nicht direkt zugänglich, sondern muß aus den Meßdaten, also der gemessenen Ditverteilung  $(A_{*m})_{m=0}^{d-1}$ , sowie den Symmetrien des Protokolls erschlossen werden. Diese Fragestellung wird in diesem Abschnitt untersucht; eine besondere Rolle spielen dabei die verallgemeinert-isotropen Zustände, die daher zunächst betrachtet werden sollen.

### 4.3.1 Verallgemeinert-isotrope Zustände

Es bietet sich an, statt der allgemeinen Form einer bell-diagonalen Dichtematrix zunächst speziellere Formen zu betrachten. Hierzu eignen sich die verallgemeinert-isotropen Zustände aus Unterabschnitt 2.5.2, da diese unter der Anwendung von  $B_n^{(d)}$ -Schritten auf wiederum verallgemeinert-isotrope Zustände abgebildet werden. Somit sind statt  $d^2$  Parametern nur noch vier Parameter zu berücksichtigen, so daß ein  $B_n^{(d)}$ -Schritt als Abbildung der Form  $B_n^{(d)} : (\alpha, \beta, \gamma, \delta) \mapsto (\alpha', \beta', \gamma', \delta')$  geschrieben werden kann. Aus dem Satz 4.6 folgt für die Koeffizienten

$$\alpha' = \{ [\alpha + (d-1)\beta]^n + (d-1)[\alpha - \beta]^n \} / dN, \quad (4.33)$$

$$\beta' = \{ [\alpha + (d-1)\beta]^n - [\alpha - \beta]^n \} / dN, \quad (4.34)$$

$$\gamma' = \{ [\gamma + (d-1)\delta]^n + (d-1)[\gamma - \delta]^n \} / dN, \quad (4.35)$$

$$\delta' = \{ [\gamma + (d-1)\delta]^n - [\gamma - \delta]^n \} / dN, \quad (4.36)$$

$$N = [\alpha + (d-1)\beta]^n + (d-1)[\gamma + (d-1)\delta]^n. \quad (4.37)$$

In diesem Spezialfall vereinfacht sich die Gleichung (4.31) zu

$$r^{(d)} = \left[ \ln \frac{\alpha + (d-1)\beta}{\gamma + (d-1)\delta} \right] / \ln \left[ \frac{\alpha + (d-1)\beta}{|\alpha - \beta|} \right], \quad (4.38)$$

und somit gilt  $r^{(d)} > 2 \Leftrightarrow \alpha^2 + (d-1)\beta^2 - (\alpha + (d-1)\beta)/d > 0$ , was die entsprechende Bedingung für Qubits [138, 139] verallgemeinert. Man beachte, daß diese Rechnung auch ohne die Mischoperation ausgeführt werden kann und auch in diesem Fall das gleiche Ergebnis liefert. Das von CHAU [31] untersuchte Protokoll erzeugt isotrope Zustände ( $\beta = \gamma = \delta$ ), und für Primzahlen liefert Gleichung (4.38) die gleichen tolerierbaren Fehlerraten (vgl. CHAU [31], erste Spalte in Tabelle 2). Dies trifft auch für Primzahlpotenzen zu, die Herleitung in diesem Kapitel deckt diese Fälle aber nicht ab.

### 4.3.2 Abschätzung von $r^{(d)}$ für Zwei-Basis-Protokolle

In dem von NIKOLOPOULOS UND ALBER [128] untersuchten Protokoll, das zwei durch eine Fouriertransformation verknüpfte komplementäre Basen verwendet, erhält man die Symmetrierelationen (vgl. Unterabschnitt 3.6.4)

$$A_{lm} = A_{d-m,l} = A_{m,d-l} = A_{d-l,d-m} \quad (4.39)$$

für jedes Tupel  $(l, m) \in \mathbb{Z}_d \times \mathbb{Z}_d$  und als Konsequenz hieraus  $A_{l*} = A_{*l}$  für jedes  $l \in \mathbb{Z}_d$ . Um in einem solchen Protokoll  $r^{(d)}$  zu bestimmen, muß zusätzlich zu den tatsächlich meßbaren Größen  $(A_{*m})_{m=0}^{d-1}$  nur noch der Wert von

$$M := \max \left\{ \left| \sum_{j=0}^{d-1} z^{lj} A_{j0} \right| \mid l \in \mathbb{Z}_d^* \right\} \quad (4.40)$$

bestimmt werden. Das Maximum einer Menge reeller Zahlen ist sicherlich größer als ihr Mittelwert, und offensichtlich ist  $|z| \geq \operatorname{Re} z$  für  $z \in \mathbb{C}$ . Aus dieser Überlegung folgt

$$\begin{aligned} M &\geq \frac{1}{d-1} \sum_{l=1}^{d-1} \left| \sum_{j=0}^{d-1} z^{lj} A_{j0} \right| \geq \frac{1}{d-1} \sum_{l=1}^{d-1} \operatorname{Re} \sum_{j=0}^{d-1} z^{lj} A_{j0} \\ &= \frac{1}{d-1} \operatorname{Re} \sum_{j=0}^{d-1} \left( \sum_{l=1}^{d-1} z^{lj} \right) A_{j0} = \frac{1}{d-1} \sum_{j=0}^{d-1} (d\delta_{j0} - 1) A_{j0} \\ &= \frac{dA_{00} - A_{*0}}{d-1} = \frac{(d-1)A_{00} + A_{00} - A_{*0}}{d-1} = A_{00} - \frac{\sum_{l=1}^{d-1} A_{l0}}{d-1}. \end{aligned} \quad (4.41)$$

Betrachtet man den verallgemeinert-isotropen Fall, so ist der letzte Ausdruck gerade  $\alpha - \beta$ , die Schranke also exakt. Dies besagt, daß in bezug auf den Wert von  $M$  der verallgemeinert-isotrope Fall der *schlechtestmögliche* Fall ist!

Man beachte aber auch, daß der verallgemeinert-isotrope Fall für die Dittfehlerkorrektur der *bestmögliche* Fall ist. Gilt jedoch  $A_{*m} = A_{*m'}$  für alle  $m, m' \in \mathbb{Z}_d^*$ , so ist der isotrope Kanal (mit  $\beta = \gamma$ ) der schlechtestmögliche in bezug auf die tolerierbare Fehlerrate. Im folgenden heiße ein Kanal, der  $A_{*m} = A_{*m'}$  für alle  $m, m' \in \mathbb{Z}_d^*$  erfüllt, *scheinbar-isotrop*, da ein böswilliger Lauscher hierdurch einen isotropen Kanal vortäuschen kann.

### 4.3.3 Maximal tolerierbare Fehlerraten

Es verbleibt nur noch, tolerierbaren Fehlerraten in der Quantenkryptographie zu berechnen. Verwendet man die Kurzbezeichnung  $x := A_{*m}$  und schreibt man  $\max \{A_{*m} \mid m \in \mathbb{Z}_d^*\} = f \cdot (1-x) \cdot (d-1)^{-1}$  für ein  $f \in [1; d-1]$ , so liest sich die Formel (4.31) als

$$r^{(d)} = \left( \ln \frac{x}{f \cdot \frac{1-x}{d-1}} \right) \cdot \left( \ln \frac{x}{M} \right)^{-1}. \quad (4.42)$$

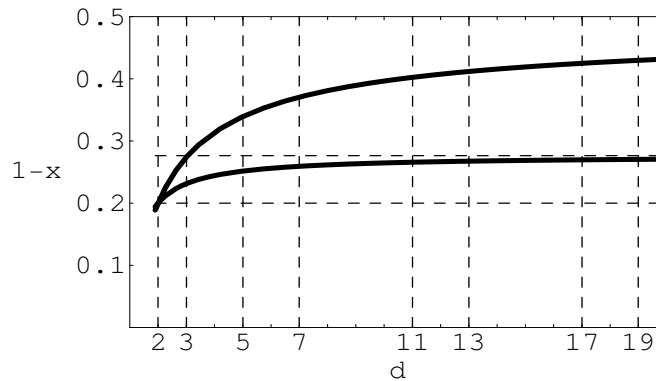


Abbildung 4.1: Untere Schranken für die maximal tolerierbare Fehlerrate ( $1 - x = 1 - A_{*0}$ ) als Funktion der Dimension  $d$ ; die obere Linie entspricht dem scheinbar-isotropen Fall  $f = 1$ , die untere dem Fall  $f = d - 1$ .

Der Fall  $f = 1$  ist der scheinbar-isotrope Kanal, während  $f = d - 1$  den Fall beschreibt, bei dem nur ein bestimmter Fehlertyp auftritt. Im Falle eines isotropen Kanals ist sicher  $f = 1$ , der Umkehrschluß gilt aber nicht.

Während  $x$  und  $f$  im Protokoll gemessen werden, muß für  $M$  die zuvor berechnete Abschätzung verwendet werden, weil eine explizite Lösung unbekannt ist. Aus der Formel für  $r^{(d)}$  erkennt man, daß untere Schranken für  $M$  auch untere Schranken für  $r^{(d)}$  und mithin für die maximal tolerierbare Fehlerrate liefern. Für das Protokoll ist also nur noch  $A_{00}$  abzuschätzen; dies geschieht mithilfe der Ungleichung

$$A_{*0} = A_{00} + \sum_{l=1}^{d-1} A_{l0} \leq A_{00} + \sum_{l=1}^{d-1} A_{l*} = A_{00} + \sum_{l=1}^{d-1} A_{*l} = A_{00} + (1 - A_{*0}), \quad (4.43)$$

aus der  $A_{00} \geq 2A_{*0} - 1$  folgt. Gleichheit liegt genau dann vor, wenn  $A_{lm} = 0$  für alle  $(l, m) \in \mathbb{Z}_d^* \times \mathbb{Z}_d^*$  gilt. Einsetzen in die Schranke für  $M$  liefert

$$M \geq x - d \cdot \frac{1 - x}{d - 1}. \quad (4.44)$$

Es wurde also gezeigt, daß unter allen scheinbar-isotropen Kanälen der isotrope Kanal der schlechtestmögliche Fall ist und unter diesen die Kanäle mit  $\delta = 0$  die schlechtesten Korrigierbarkeitseigenschaften haben. Je anisotroper die betrachteten Kanäle sind, desto schlechter ist die berechnete Schranke. Dies wird unmittelbar deutlich, indem man den Fall  $f = d - 1$  betrachtet, wo aufgrund der Symmetrien nur vier freie Parameter vorliegen; in diesem Fall können aber für  $d \rightarrow \infty$  beliebig hohe Fehlerraten toleriert werden.

Nichtsdestoweniger kann man in Abhängigkeit von  $f$  untere Schranken für  $x$ , also obere Schranken für die Fehlerrate  $1 - x$  berechnen; es ergibt sich, daß die Korrigierbarkeit nach Satz 4.7 zumindest dann gewährleistet ist, wenn

$$x > \frac{2d(2d-1) + (d-1)(f + \sqrt{(4d+f)f})}{2[(2d-1)^2 + (d-1)f]} \quad (4.45)$$

gilt, wobei die Verschränkungsbedingung  $x > (d+1)/(2d)$  von NIKOLOPOULOS UND ALBER [128] berücksichtigt wurde. Trägt man  $x$  für die Werte  $f = 1$  und  $f = d-1$  auf, so erhält man einen Bereich, in dem unteren Schranken der maximal tolerierbaren Fehlerraten liegen (vgl. Abbildung 4.1); beide Linien beginnen bei  $1 - x = 0,2$ . Für scheinbar isotrope Kanäle sind die unteren Schranken exakt. Beachtenswert ist vielleicht noch das asymptotische Verhalten dieser Kurven: betrachtet man  $x$  als Funktion von  $d$  und  $f$  und setzt  $f = k \cdot d$ , so folgt

$$\lim_{d \rightarrow \infty} x(d, k \cdot d) = \frac{4 + k + \sqrt{4 + k}}{2(4 + k)} = \frac{1}{2} \left[ 1 + \sqrt{\frac{k}{k + 4}} \right]. \quad (4.46)$$

Für den scheinbar-isotropen Fall ( $f = 1$ ) folgt hieraus  $x(d, 1) \rightarrow 0,5$ , für den Fall  $f = d-1$  hingegen  $x(d, d-1) \rightarrow 0,5 + (2\sqrt{5})^{-1}$ . Aus einem unbekanntem Grund entspricht die letztgenannte Zahl genau der Schranke, die CHAU [30] für das qubit-basierte Protokoll mit sechs Zuständen angibt.

## Kapitel 5

# Konkrete Protokolle

In diesem Kapitel soll ein konkretes Protokoll vorgestellt werden, das die Aufgabe der Einweg-Reinigung wahrnehmen kann. Historisch gesehen ist dieses Protokoll der Ausgangspunkt der Untersuchungen des Kapitels 4. Es wurde in Zusammenarbeit mit GEORGIOS NIKOLOPOULOS entwickelt und veröffentlicht [130]; es gründet sich im wesentlichen auf zwei Arbeiten von CHAU [30, 31].<sup>1</sup>

### 5.1 Korrektur der Phasenfehler

Bei dem im vorangehenden Kapitel vorgestellten  $B_n^{(d)}$ -Schritt wird die Ditfehlerrate verringert, die Phasenfehlerrate aber erhöht; es wurde untersucht, ob nach der Quanten-SHANNON-Schranke von HAMADA ein asymmetrischer CSS-Code existieren muß, der die Sicherheit des Protokolls gewährleistet, und der – wie jeder CSS-Code – die Sicherheit eines nicht verschränkungs-basierten Protokolls impliziert. Es wird allerdings kein unmittelbarer Weg aufgezeigt, wie dieser CSS-Code überhaupt zu finden ist.

In diesem Kapitel wird nun ein Korrekturschritt, der  $P_n^{(d)}$ -Schritt, vorgestellt, der in Umkehrung der Eigenschaften des  $B_n^{(d)}$ -Schrittes die Phasenfehlerrate verringert, dafür aber die Ditfehlerrate erhöht. Das Zusammenwirken dieser Schritte wird eine Reduktion sowohl der Dit- als auch der Phasenfehler ermöglichen.

---

<sup>1</sup>Der Gedanke, die CHAUSchen Protokolle zu verallgemeinern, stammt ausschließlich von GEORGIOS NIKOLOPOULOS, die weitere Ausarbeitung erfolgte gemeinsam. Die Darstellung in diesem Kapitel weicht in Formulierung und Notation von der veröffentlichten Arbeit ab, die wesentlichen Inhalte blieben aber unverändert. Die in der veröffentlichten Fassung angegebene Beschränkung auf Primzahldimensionen ist meines Erachtens überflüssig und wird daher weggelassen.

Das genaue Protokoll lautet nun wie folgt:

1. Alice und Bob teilen sich einen Zustand  $\rho^{\otimes N}$ , wobei  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  und  $N \in \mathbb{N}$  hinreichend groß ist; vgl. Unterabschnitt 3.6.3.
2. Sie wenden einen  $B_{n'}^{(d)}$ -Schritt an, wobei  $n'$  vom Ausgangszustand  $\rho$  abhängt.
3. Sie wenden einen  $P_n^{(d)}$ -Schritt an, wobei  $n$  von  $n'$  und dem Ausgangszustand  $\rho$  abhängt.

Fordert man daß der Zustand  $\rho'$  nach Ausführung des Protokolls nahe am Idealzustand  $\rho_0 = |\Psi_{00}\rangle\langle\Psi_{00}|$  liegt, daß also  $\|\rho' - \rho_0\|_1 \leq 2\varepsilon$  gilt, so müssen die Parameter  $n'$  und  $n$  des Protokolls in Abhängigkeit vom Ausgangszustand  $\rho = (A_{lm})_{l,m=0}^{d-1}$  – oder von den meßbaren Fehlerraten  $(A_{*m})_{m=0}^{d-1}$  – und der erlaubten Quantenfehlerrate  $\varepsilon$  bestimmt werden. Das Protokoll gewährleistet die Sicherheit, wenn solche Parameter für jedes  $\varepsilon > 0$  angegeben werden können; vgl. hierzu die Interpretation aus Unterabschnitt 3.4.2.

In der Praxis kann es angebracht sein, die Fehlerrate durch die genannten Verfahren nur unter eine bestimmte Schwelle zu bringen (vielleicht 5%) und anschließend effizientere Korrekturverfahren zu verwenden; hierdurch kann die Schlüsselrate erhöht werden. Hier geht es aber nur darum, überhaupt ein *konkretes* Protokoll anzugeben, das die Sicherheit garantiert.

### 5.1.1 Der $P_n^{(d)}$ -Schritt

Einen Korrekturschritt, der die Phasenfehlerrate vermindert, die Dittfehlerrate aber erhöht, ist leicht zu konstruieren: man wende auf alle beteiligten Qudit-Paare eine bilaterale Fouriertransformation  $\mathcal{F} \otimes \mathcal{F}^*$  an, führe einen  $B_n^{(d)}$ -Schritt aus und invertiere die Fouriertransformation. Der Nachteil dieses Korrekturschritts ist, daß er sich nicht auf ein *Prepare-and-Measure*-Protokoll reduzieren läßt, was sich schon im Falle von  $d = 2$  zeigt (vgl. GOTTESMAN UND LO [64] oder auch meine Diplomarbeit [138]).

Analog zum Qubit-Fall läßt sich aber ein Korrekturschritt definieren, der im folgenden als  $P_n^{(d)}$ -Schritt bezeichnet werde; ein  $P_n^{(d)}$ -Schritt verallgemeinert den  $P_n$ -Schritt für Qubit-Zustände [138, 139]. Die Phasenfehler  $p_l := A_{l*}$  definieren eine Verteilung  $p = (p_0, \dots, p_{d-1}) \in \mathcal{W}_d$ , und wählt man aus  $n$  Realisierungen die Mehrheit aller Werte aus, so funktioniert das Verfahren für  $n \rightarrow \infty$  nun immer, vorausgesetzt, daß  $p_0 > p_m := \max\{p_i \mid i \in \mathbb{Z}_d^*\}$  und  $n \in \mathbb{N}$  hinreichend groß ist.<sup>2</sup>

---

<sup>2</sup>Nach dem Gesetz der großen Zahlen ist dies offensichtlich, und für die Quantenkryptographie genügt es auch, wenn überhaupt ein eindeutiger größter Wert existiert; in diesem Falle wird ggf. bezüglich eines anderen, aber gleichermaßen sicheren Zustands gereinigt.

Quantenmechanisch wird ein  $P_n^{(d)}$ -Schritt für  $n \in \mathbb{N}$  wie folgt ausgeführt:<sup>3</sup>

1. Alice und Bob wählen  $n$  Qudit-Paare  $QP_1, \dots, QP_n$  aus.
2. Alice und Bob führen an allen Paaren eine bilaterale Fouriertransformation  $\mathcal{F} \otimes \mathcal{F}^*$  aus.<sup>4</sup>
3. Alice und Bob führen GBXOR( $QP_1, QP_k$ ) für  $k \in \{2, \dots, n\}$  aus.
4. Alice und Bob messen den Ditwert an  $QP_2, \dots, QP_n$  und stellen eine Häufigkeitsverteilung  $(H_s)_{s=0}^{d-1}$  mit  $\sum_{s=0}^{d-1} H_s = n - 1$  auf; anschließend erhöhen sie  $H_0$  um eins.
5. Aus der Menge  $\{s \in \mathbb{Z}_d \mid (\forall i \in \mathbb{Z}_d)(H_s \geq H_i)\}$  wählt Bob gleichverteilt ein Element  $s'$  aus und transformiert  $QP_1$  mittels  $l_1 \mapsto l_1 \ominus s'$ .
6. Alice und Bob führen an  $QP_1$  eine bilaterale Fouriertransformation aus und verwerfen alle anderen Paare.

Der Prozeß erklärt sich wie folgt: die bilaterale Fouriertransformation überführt den Bell-Zustand  $(l, m)$  in  $(m, \ominus l)$ . Somit liefern die ersten drei Schritte

$$\bigotimes_{i=1}^n (l_i, m_i) \mapsto \left( \bigoplus_{i=1}^n m_i, \ominus l_1 \right) \otimes \left[ \bigotimes_{k=2}^n (m_k, l_k \ominus l_1) \right]. \quad (5.1)$$

Im vierten Schritt messen Alice und Bob die Werte  $l_k \ominus l_1$  für  $k \in \{2, \dots, n\}$ . Fügen sie den erschlossenen Wert  $l_1 \ominus l_1 = 0$  hinzu, so erhalten die die absoluten Häufigkeiten der Abweichungen vom Wert  $l_1$  des ersten Paares. Im fünften Schritt wählt Bob einen (von evtl. mehreren) Maximalwerten aus und stellt durch eine lokale Operation  $l_1$  auf diesen Maximalwert ein, wendet also  $\mathbb{I} \otimes X^{-s'}$  an.<sup>5</sup>

Der  $P_n^{(d)}$ -Schritt, so wie er formuliert wurde, verwendet zwar klassische Zweiweg-Kommunikation, im zugeordneten *Prepare-and-Measure*-Protokoll wird er aber zum Einweg-Verfahren. Da Alice und Bob nur an den Ditwerten interessiert sind, bilden sie einfach die Parität  $\bigoplus_{i=1}^n m_i$  der klassischen Dits.

---

<sup>3</sup>Der  $P_n^{(d)}$ -Schritt entspricht einer klassischen Fehlerkorrektur mittels eines linearen Codes mit Generatormatrix  $G = (1, \dots, 1)^t \in \mathbb{Z}_d^{1 \times n}$ ; formal gesehen ist dies nur im Falle von Primzahlpotenzen  $d$  möglich, aufgrund seiner Einfachheit funktioniert das Verfahren aber auch allgemeiner.

<sup>4</sup>Statt der Fouriertransformation können sie auch eine andere Transformation verwenden, die Dit- und Phasenfehlerverteilungen austauscht.

<sup>5</sup>Im Falle von Qubits ( $d = 2$ ) war es möglich, durch die Forderung, daß  $n$  eine ungerade Zahl ist, Mehrdeutigkeiten bei der Wahl von  $s'$  zu vermeiden. Hier ist das nicht möglich, dies erschwert aber nur die genaue Formulierung, nicht die weitere Berechnung.

**Lemma 5.1 (Phasenfehlerraten nach  $B_n^{(d)}$ -Schritten)**

Erfüllt eine Dichtematrix  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  die Symmetriebedingungen  $A_{lm} = A_{m,d-l} = A_{d-l,d-m} = A_{d-m,l}$  und ist  $A_{0*} > \max\{A_{l*} | l \in \mathbb{Z}_d^*\}$ , so gilt  $A'_{0*} > \max\{A'_{l*} | l \in \mathbb{Z}_d^*\}$  nach der Ausführung eines  $B_n^{(d)}$ -Schrittes für geradzahliges  $n \in \mathbb{N}$ .

*Beweis:* Unter Verwendung des Satzes 4.6 erhält man für jeden Wert  $l \in \mathbb{Z}_d^*$  die Äquivalenz

$$\begin{aligned} A_{0*} > A_{l*} &\Leftrightarrow \sum_{m,i} \left[ \left( \sum_j z^{ij} A_{jm} \right)^n \right] > \sum_{m,i} \left[ z^{-il} \left( \sum_j z^{ij} A_{jm} \right)^n \right] \quad (5.2) \\ &\Leftrightarrow \sum_i \left[ \sum_m \left( \sum_j z^{ij} A_{jm} \right)^n \right] > \sum_i \left[ z^{-il} \sum_m \left( \sum_j z^{ij} A_{jm} \right)^n \right] \end{aligned}$$

für die Phasenfehlerraten nach einem  $B_n^{(d)}$ -Schritt. Da beide Seiten stets reell sind, ist diese Ungleichung zumindest dann erfüllt, wenn der Ausdruck  $\sum_m \left( \sum_j z^{ij} A_{jm} \right)^n$  für alle  $i \in \mathbb{Z}_d$  reell ist. Um dies zu zeigen, spaltet man die Summe in mehrere Teile auf: für den Summanden mit  $m = 0$  ergibt sich

$$\begin{aligned} \sum_{j=0}^{d-1} z^{ij} A_{j0} &= \frac{\sum_{j=0}^{d-1} z^{ij} A_{j0} + \sum_{j=0}^{d-1} z^{i(d-j)} A_{d-j,0}}{2} \\ &= \sum_{j=0}^{d-1} A_{j0} \frac{z^{ij} + z^{i(d-j)}}{2} = \sum_{j=0}^{d-1} A_{j0} \cos(2\pi i/d \cdot ij). \quad (5.3) \end{aligned}$$

Ist  $d$  eine gerade Zahl, so gilt diese Überlegung im wesentlichen auch für den Fall  $m = d/2$ . Für die übrigen Summanden soll gezeigt werden, daß die Summanden zu  $m$  und der zu  $d - m$  zueinander konjugiert komplex sind, ihre Summe also reell ist. Es genügt dafür zu zeigen, daß dies auf die Ausdrücke  $\sum_j z^{ij} A_{jm}$  zutrifft, da die Potenzierung mit  $n$  die fragliche Eigenschaft nicht ändert. Im Summanden zu einem  $m$  tritt der Summand  $z^{ij} A_{jm}$  auf, dem im Summanden zu  $d - m$  genau ein Summand  $z^{(d-i)(d-j)} A_{d-j,d-m}$  zugeordnet werden kann. Nach Voraussetzung ist  $A_{d-j,d-m} = A_{jm}$  reell, so daß mit  $z^{(d-i)(d-j)} = z^{-ij}$  die gewünschte Eigenschaft folgt, q. e. d.

Für allgemeine bell-diagonale Zustände gilt die Aussage nicht: sind zum Beispiel  $A_{00} = 1 - p$  und  $A_{10} = p$ , so wird nach einem  $B_n^{(d)}$ -Schritt das erste Paar stets weiterverwendet, aber für  $d \geq n$  ist zum Beispiel  $A'_{l*} = B(n; p; l)$  binomialverteilt. Selbst für  $p < 1/2$  ist  $A'_{l*}$  nicht zwingend der größte Wert.

**5.1.2 Das Verhalten der Zustände unter  $P_n^{(d)}$ -Schritten**

Das genaue Verhalten eines Zustands unter Anwendung des  $P_n^{(d)}$ -Schrittes ist in geschlossener Form sehr schwierig zu bestimmen. Für die Zwecke dieses

Kapitels genügt es aber, sowohl die Gesamt-Ditfehlerrate  $R_D := 1 - A_{*0}$ , als auch die Gesamt-Phasenfehlerrate  $R_P := 1 - A_{0*}$  nach oben abzuschätzen. Das folgende Lemma zeigt, daß die Ditfehlerrate höchstens linear, das heißt, proportional zu  $n$  wächst, die Phasenfehlerrate aber exponentiell in  $n$  abfällt.

**Lemma 5.2 (Schranken der Fehlerraten unter  $P_n^{(d)}$ -Schritten)**

Ist  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ein Zustand mit Ditfehlerrate  $R_D := 1 - A_{*0}$  und Phasenfehlerrate  $R_P := 1 - A_{0*}$ , so gelten für die Raten  $R'_D$  und  $R'_P$  nach Anwendung eines  $P_n^{(d)}$ -Schrittes die Ungleichungen

$$R'_D \leq nR_D \quad \text{und} \quad R'_P \leq \sum_{l=1}^{d-1} \left[ 1 - \left( \sqrt{A_{0*}} - \sqrt{A_{l*}} \right)^2 \right]^n,$$

vorausgesetzt, es gilt  $A_{0*} > A_{l*}$  für alle  $l \in \mathbb{Z}_d^*$ .

*Beweis:* Der Ditwert des ersten Paares nach einem  $P_n^{(d)}$ -Schritt ergibt sich durch modulare Addition der Ditwerte  $m_i$  der  $n$  Anfangspaare zu  $\oplus_i m_i$ . Diese Summe ist zumindest dann Null, wenn alle  $m_i = 0$  sind; die Wahrscheinlichkeit hierfür ist  $(1 - R_D)^n$ . Nach der BERNOULLISchen Ungleichung (vgl. bei HEUSER [74], Teil I, S. 61) gilt  $(1 + x)^n \geq 1 + nx$  für  $x \in [-1; +\infty)$  und  $n \in \mathbb{N}_0$ , so daß  $1 - R'_D \geq (1 - R_D)^n \geq 1 - nR_D$  und hieraus die Abschätzung für den Ditfehler folgt.<sup>6</sup>

Wie gezeigt wurde, nimmt der Phasenwert  $l_1$  des ersten Qudit-Paares nach Ausführung eines  $P_n^{(d)}$ -Schrittes den unter den  $n$  Qudit-Paaren am häufigsten auftretenden Wert an. Die Auswahl der  $n$  anfänglichen Qudit-Paare liefert die Realisierung einer multinomial verteilten Zufallsvariablen  $X$ , deren Wahrscheinlichkeitsverteilung  $p = (p_0, \dots, p_{d-1}) \in \mathcal{W}_d$  durch die Anfangsverteilung der Phasenfehler gegeben ist; es ist somit  $p_l = A_{l*}$  für  $l \in \mathbb{Z}_d$ . Die Realisierung liefere  $\eta_i$ -mal den Wert  $i \in \mathbb{Z}_d$ ; es ist  $\eta_i \in \mathbb{N}_0$  und  $\sum_{i=0}^{d-1} \eta_i = n$ . Ein Fehler tritt niemals auf, wenn alle  $\eta_i$  kleiner als  $\eta_0$  sind; umgekehrt gilt

$$R'_P \leq P \left[ \bigvee_{i=1}^{d-1} (\eta_i \geq \eta_0) \right] \leq \sum_{i=1}^{d-1} P(\eta_i \geq \eta_0). \quad (5.4)$$

Um die Summanden zu berechnen, genügt es, statt der allgemeinen Multinomialverteilung eine Trinomialverteilung für  $\eta_0$ ,  $\eta_i$  und  $\eta_{\text{Rest}} := n - \eta_0 - \eta_i$  zu betrachten. Setzt man nun  $p_{\text{Rest}} := 1 - p_0 - p_i$  und  $r_i := n - \eta_{\text{Rest}}$  und

---

<sup>6</sup>Im Grunde genommen kann man die Ditfehlerrate bei einem  $P_n^{(d)}$ -Schritt genauso behandeln wie die Phasenfehlerrate bei einem  $B_n^{(d)}$ -Schritt, das heißt, durch Fouriertransformation, Multiplikation und inverse Fouriertransformation. Nimmt man an, daß die anfängliche Ditfehlerrate klein ist, was zum Beispiel nach einem  $B_n^{(d)}$ -Schritt für hinreichend großes  $n \in \mathbb{N}$  der Fall ist, so ist die Wahrscheinlichkeit für mehrfache, sich aufhebende Fehler gering und die Abschätzung recht gut.

definiert die renormierten Wahrscheinlichkeiten durch  $\tilde{p}_0 := p_0/(p_0 + p_i)$  und  $\tilde{p}_i := p_i/(p_0 + p_i)$ , so führt dies auf

$$\begin{aligned} P(\eta_0, \eta_i, \eta_{\text{Rest}}) &= \sum_{\eta_{\text{Rest}}=0}^n \binom{n}{\eta_{\text{Rest}}} p_{\text{Rest}}^{\eta_{\text{Rest}}} \left[ \sum_{\eta_i=0}^{r_i} \binom{r_i}{\eta_i} p_0^{\eta_0} p_i^{\eta_i} \right] \\ &= \sum_{\eta_{\text{Rest}}=0}^n \binom{n}{\eta_{\text{Rest}}} p_{\text{Rest}}^{\eta_{\text{Rest}}} (p_0 + p_i)^{\eta_0 + \eta_i} \cdot \left[ \sum_{\eta_i=0}^{r_i} \binom{r_i}{\eta_i} \tilde{p}_0^{\eta_0} \tilde{p}_i^{\eta_i} \right]. \end{aligned} \quad (5.5)$$

Der Ausdruck in der zweiten eckigen Klammer läßt sich für jedes feste  $\eta_{\text{Rest}}$  wegen  $p_0 > p_i$  mithilfe der CHERNOFF-Schranke (Satz C.34) abschätzen, was

$$P(\eta_i \geq \eta_0 | \eta_{\text{Rest}}) = \sum_{\eta_i=n/2}^n \binom{n}{\eta_i} \tilde{p}_0^{\eta_0} \tilde{p}_i^{\eta_i} \leq (4\tilde{p}_0\tilde{p}_i)^{n/2} = \left( \frac{4p_0p_i}{(p_0 + p_i)^2} \right)^{n/2}$$

liefert. Setzt man dies in die Gesamtwahrscheinlichkeit ein, so erhält man

$$\begin{aligned} P(\eta_i \geq \eta_0) &= \sum_{\eta_{\text{Rest}}=0}^n P(\eta_{\text{Rest}}) \cdot P(\eta_i \geq \eta_0 | \eta_{\text{Rest}}) \\ &\leq \sum_{\eta_{\text{Rest}}=0}^n \binom{n}{\eta_{\text{Rest}}} p_{\text{Rest}}^{\eta_{\text{Rest}}} (1 - p_{\text{Rest}})^{1-\eta_{\text{Rest}}} \left( \frac{4p_0p_i}{(p_0 + p_i)^2} \right)^{n/2}. \end{aligned} \quad (5.6)$$

Formt man die rechte Seite mithilfe des binomischen Lehrsatzes weiter um, so ergibt sich schließlich

$$\left[ p_{\text{Rest}} + (1 - p_{\text{Rest}}) \frac{2\sqrt{p_0p_i}}{p_0 + p_i} \right]^n = (p_{\text{Rest}} + 2\sqrt{p_0p_i})^n = [1 - (\sqrt{p_0} - \sqrt{p_i})^2]^n,$$

woraus die Behauptung folgt, q. e. d.

Der Ausdruck des Lemmas läßt sich weiter vereinfachen, wenn man  $p_0 := A_{0*}$  und  $p_{\text{max}} := \max \{A_{l*} | l \in \mathbb{Z}_d^*\}$  betrachtet; nach Logarithmieren folgt aus Lemma B.43 die Ungleichung  $(1 - x)^n \leq e^{-nx}$  für  $x < 1$  und hiermit

$$\begin{aligned} R'_P &\leq (d - 1) [1 - (\sqrt{p_0} - \sqrt{p_{\text{max}}})^2]^n \\ &\leq (d - 1) \exp [-n(\sqrt{p_0} - \sqrt{p_{\text{max}}})^2]. \end{aligned} \quad (5.7)$$

## 5.2 Berechnung tolerierbarer Fehlerraten

In diesem Abschnitt wird ausgehend von den bisherigen Überlegungen ein Protokoll bestimmt, welches bei gegebener Fehlerrate und gewählten Sicherheitsparametern einen sicheren Schlüssel erzeugt.

### 5.2.1 Das Protokoll

Gegeben sei nun ein bell-diagonaler Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$ . Durch Anwendung des beschriebenen Protokolls soll der ideale (fehlerfreie) Zustand  $\rho_0 := |\Psi_{00}\rangle\langle\Psi_{00}| = (\delta_{l0}\delta_{m0})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  destilliert werden, was aber nur mit einem gewissen Fehler geschehen kann. Fordert man, daß  $A_{00} \geq 1 - \varepsilon$  oder  $Q := 1 - A_{00} \leq \varepsilon$  gilt, so berechnet man für den Abstand zwischen dem realen und dem idealen Zustand in der Spurnorm  $\|\rho - \rho_0\|_1 \leq 2\varepsilon$ , so daß die Interpretation aus Unterabschnitt 3.4.2 angewendet werden kann.

Um das konkrete Protokoll festzulegen, müssen also nur noch die Werte von  $n'$  und  $n$  (vgl. die Einleitung zum Abschnitt 5.1) in Abhängigkeit von Zustand und Sicherheitsparameter  $\varepsilon$  bestimmt werden. Nach Lemma 5.2 wächst die Ditfehlerrate unter einem  $P_n^{(d)}$ -Schritt auf höchstens ihr  $n$ -faches an. Bezeichnet also  $1 - A'_{*0}$  die Ditfehlerrate nach dem  $B_{n'}^{(d)}$ -Schritt, so ist sie nach dem darauffolgenden  $P_n^{(d)}$ -Schritt nicht größer als  $n \cdot (1 - A'_{*0})$ . Wählt man also

$$n := \left\lceil \frac{\varepsilon}{2} \cdot \frac{1}{(1 - A'_{*0})} \right\rceil, \quad (5.8)$$

so gilt für die Ditfehlerrate nach dem anschließenden  $P_n^{(d)}$ -Schritt die Abschätzung  $R''_D \leq n \cdot (1 - A'_{*0}) \approx \varepsilon/2$ . Mit Formel (5.7) folgt weiter

$$Q \leq R''_D + R''_P \leq \frac{\varepsilon}{2} + (d-1) \exp \left[ -n (\sqrt{p_0} - \sqrt{p_{\max}})^2 \right], \quad (5.9)$$

wenn  $p_0 := A'_{0*}$  und  $p_{\max} := \max \{A'_{l*} | l \in \mathbb{Z}_d^*\}$  die Phasenfehlerraten nach dem  $B_{n'}^{(d)}$ -Schritt bezeichnen. Es ist also nur noch ein  $n' \in \mathbb{N}$  derart zu finden, daß der zweite Term nicht größer als  $\varepsilon/2$  ist; umgeschrieben lautet diese Bedingung

$$-n (\sqrt{p_0} - \sqrt{p_{\max}})^2 \leq \ln \left( \frac{\varepsilon}{2(d-1)} \right) \quad (5.10)$$

oder, wenn man das  $n$  aus Gleichung (5.8) einsetzt und die GAUßklammer ignoriert,

$$\frac{(\sqrt{p_0} - \sqrt{p_{\max}})^2}{1 - A'_{*0}} \geq \frac{2}{\varepsilon} \ln \left( \frac{2(d-1)}{\varepsilon} \right). \quad (5.11)$$

Für  $\varepsilon \rightarrow 0$  strebt die rechte Seite gegen  $+\infty$ , die Korrigierbarkeit ist also genau dann gegeben, wenn die linke Seite, die die Parameter nach einem  $B_{n'}^{(d)}$ -Schritt enthält, für  $n' \rightarrow \infty$  gleichermaßen über alle Grenzen wächst.

Für einen vorgegebenen Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  und einen Sicherheitsparameter  $\varepsilon > 0$  kann diese Bedingung sehr leicht numerisch überprüft werden, eine allgemeine Aussage ist allerdings nur schwer zu treffen. Im nächsten Unterabschnitt wird diese Aufgabe für die verallgemeinert-isotropen Zustände vollständig gelöst werden.

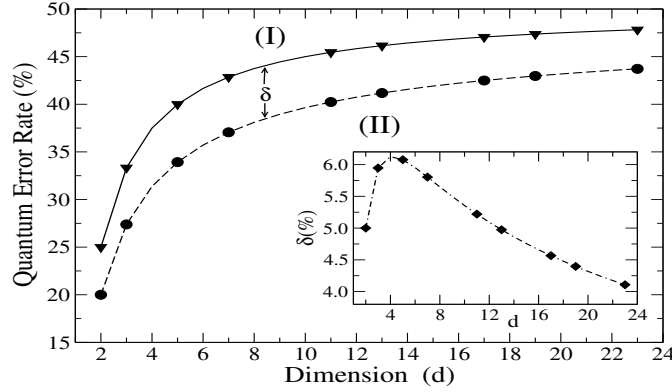


Abbildung 5.1: Tolerierbare Fehlerraten (untere Linie: für  $\beta = \gamma$  und  $\delta = 0$ ) und Verschränkungsgrenzen nach NIKOLOPOULOS u. a. [128] (obere Linie)

## 5.2.2 Verallgemeinert-isotrope Zustände

Im Falle verallgemeinert-isotroper Zustände  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{\text{bd}}^{(d)}$  kann der Bereich der Korrigierbarkeit exakt bestimmt werden. Schreibt man für die Normierungskonstante  $N := [\alpha + (d-1)\beta]^n + (d-1)[\gamma + (d-1)\delta]^n$ , so folgen nach einem  $B_n^{(d)}$ -Schritt aus den Formeln (4.33) bis (4.36) die Werte

$$\begin{aligned} p_0 &= \alpha' + (d-1)\gamma' = \frac{1}{d} \left( 1 + \frac{(d-1)(\alpha - \beta)^n + (d-1)^2(\gamma - \delta)^n}{N} \right), \\ p_{\max} &= \beta' + (d-1)\delta' = \frac{1}{d} \left( 1 - \frac{(\alpha - \beta)^n + (d-1)(\gamma - \delta)^n}{N} \right). \end{aligned} \quad (5.12)$$

Entwickelt man den Ausdruck  $\sqrt{p_0} - \sqrt{p_{\max}}$  mittels  $\sqrt{1+x} = 1 + x/2 + \dots$ , so erhält man

$$\sqrt{p_0} - \sqrt{p_{\max}} \approx \frac{1}{\sqrt{d}} \cdot \frac{d}{2} \cdot \frac{(\alpha - \beta)^n + (d-1)(\gamma - \delta)^n}{N} \quad (5.13)$$

in erster Ordnung. Mit  $A'_{*0} = [\alpha + (d-1)\beta]^n/N$  folgt

$$\frac{(\sqrt{p_0} - \sqrt{p_{\max}})^2}{1 - A'_{*0}} \approx \frac{d}{4} \cdot \frac{[(\alpha - \beta)^n + (d-1)(\gamma - \delta)^n]^2}{N \cdot (d-1)[\gamma + (d-1)\delta]^n}, \quad (5.14)$$

und da in den relevanten Fällen  $\alpha - \beta \geq \gamma - \delta$  ist, wird dies genau im Falle von  $\alpha - \beta > [\alpha + (d-1)\beta] \cdot [\gamma + (d-1)\delta]$  für  $n \rightarrow \infty$  beliebig groß. Diese Bedingung folgte aber im Kapitel 4 aus Gleichung (4.38) in Verbindung mit Lemma 4.7.

## Kapitel 6

# Symmetrische Erweiterbarkeit von Quantenzuständen

In den Kapiteln 4 und 5 wurde die Sicherheit einiger Protokolle der Quantenkryptographie unter Verwendung eines  $B_n^{(d)}$ -Schritts als Zweiweg-Verfahren und hinreichender Bedingungen für die Korrigierbarkeit eines Zustands durch Verfahren mit Einweg-Kommunikation betrachtet: im Kapitel 4 wurde die Quanten-SHANNON-Schranke verwendet, im Kapitel 5 ein festgelegter Code. In diesem Kapitel sollen ähnliche Untersuchungen anhand einer notwendigen Bedingung durchgeführt werden. Diese notwendige Bedingung besagt, daß der Zustand keine *symmetrische Erweiterung* besitzen darf; ist ein Zustand *symmetrisch erweiterbar*, so gibt es überhaupt kein Einweg-Verfahren, das den Zustand reinigen kann.<sup>1</sup>

Ein Großteil dieses Kapitels wird sich mit der Frage beschäftigen, welche Zustände einer Form, die zum Anfang des Kapitels eingeführt wird, symmetrisch erweiterbar sind; ein wichtiges Hilfsmittel der Rechnungen ist dabei das HURWITZ-Kriterium (Satz B.21). Mithilfe eines Kriteriums für symmetrische Erweiterbarkeit in der untersuchten Klasse (Hauptsatz 6.6) werden zum Ende des Kapitels obere Schranken für die maximal tolerierbaren Fehlerraten berechnet. Es wird sich zeigen, daß zumindest für die wichtigsten Fälle sich die gleichen Raten wie in den vorausgegangenen Kapiteln ergeben, die dort bestimmten Raten also nicht ohne weiteres verbessert werden können.

---

<sup>1</sup>Der Begriff der symmetrischen Erweiterbarkeit und seine unmittelbaren Folgen wurden mir während zweier SECOQC-QIT-Workshops in Wien (im April und im Dezember 2007) durch NORBERT LÜTKENHAUS und GEIR OVE MYHR bekannt. Eine Vorabfassung der Arbeit *Notes on symmetry properties for state extensions of Bell-diagonal states* von DOHERTY UND RENES erhielt ich von JOE RENES. Während eines eingeladenen Aufenthaltes in Cambridge führte ich mit MATTHIAS CHRISTANDL Gespräche zum Thema der symmetrischen Erweiterbarkeit.

## 6.1 Grundlagen

Zunächst wird der Begriff der *symmetrischen Erweiterbarkeit* von Quantenzuständen definiert, der verschiedentlich in der Literatur erscheint; vgl. zum Beispiel TERHAL, DOHERTY UND SCHWAB [175].

### Definition 6.1 (Symmetrische Erweiterbarkeit)

Für  $s, t \in \mathbb{N}$  nennt man einen Zustand  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  symmetrisch  $(s, t)$ -erweiterbar, wenn es einen Zustand  $\rho_{A_1, \dots, A_s, B_1, \dots, B_t} \in \mathcal{S}(\mathcal{H}_A^{\otimes s} \otimes \mathcal{H}_B^{\otimes t})$  gibt, der den folgenden Bedingungen genügt:

1. Für alle Permutationen  $\pi_A \in S_s$  und  $\pi_B \in S_t$  auf den beiden Untersystemen gilt  $\rho_{A_{\pi_A(1)}, \dots, A_{\pi_A(s)}, B_{\pi_B(1)}, \dots, B_{\pi_B(t)}} = \rho_{A_1, \dots, A_s, B_1, \dots, B_t}$ ;
2. es gilt  $\rho_{AB} = \text{Spur}_{A_2, \dots, A_s, B_2, \dots, B_t} \rho_{A_1, \dots, A_s, B_1, \dots, B_t}$ .

Man nennt einen Zustand  $\rho_{AB}$  symmetrisch erweiterbar (im engeren Sinne), wenn er symmetrisch  $(1, 2)$ -erweiterbar ist.

Unmittelbar aus der Definition folgt, daß separable Zustände für alle  $s, t \in \mathbb{N}$  symmetrisch erweiterbar sind; eine Erweiterung von  $\rho_{AB} = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$  ist zum Beispiel  $\rho_{AB} = \sum_i p_i \rho_{A,i}^{\otimes s} \otimes \rho_{B,i}^{\otimes t}$ . Umgekehrt sind reine verschränkte Zustände (außer für  $s = t = 1$ ) nicht symmetrisch erweiterbar.

Ein Kriterium für die symmetrische Erweiterbarkeit bell-diagonaler Zwei-Qubit-Zustände gaben DOHERTY UND RENES an, eine Vermutung für allgemeine Zwei-Qubit-Zustände stellte MYHR auf.

### 6.1.1 Bedeutung in der Quantenkryptographie

Für die Quantenkryptographie ist nun folgende einfache Überlegung wichtig: Teilen sich Alice und Bob einen symmetrisch  $(1, 2)$ -erweiterbaren Quantenzustand, so ist es unmöglich, ihn mittels Einweg-Kommunikation von Alice zu Bob zu reinigen, denn Eve könnte dieselben Operationen wie Bob ausführen, was ihr ermöglichen würde, denselben Zustand wie Bob zu erhalten. Ist dies der Fall, so kann der (reine) Gesamtzustand von Alice und Bob nicht maximal verschränkt sein, oder Eve erhielte dasselbe Ergebnis wie Bob.

Im folgenden wird ausschließlich die symmetrische  $(1, 2)$ -Erweiterbarkeit untersucht; bezeichnet  $\pi$  die Vertauschung der Teilsysteme  $A$  und  $B$ , so wird sich mit  $\pi|\Psi_{lm}\rangle = z^{lm}|\Psi_{l,d\oplus m}\rangle$  zeigen, daß in der Klasse der  $\mathfrak{U}_2$ -invarianten bell-diagonalen Zustände (siehe Unterabschnitt 6.2.2) die symmetrisch  $(1, 2)$ -erweiterbaren Zustände mit den symmetrisch  $(2, 1)$ -erweiterbaren Zuständen zusammenfallen, weshalb auch die Einweg-Kommunikation von Bob zu Alice nicht helfen würde.

### 6.1.2 Elementare Eigenschaften

Gegenwärtig ist kein allgemeines Kriterium bekannt, um festzustellen, ob ein Quantenzustand symmetrisch erweiterbar ist oder nicht. In diesem Abschnitt werden einige elementare Aussagen gezeigt, die diese Frage in bestimmten Situationen zu lösen helfen.

**Lemma 6.2 (Invariante Quantenzustände)**

*Ein Zustand  $\rho$  ist genau dann invariant bezüglich einer unitären Transformation  $U$ , wenn er mit  $U$  kommutiert.*

*Beweis:* Es gilt  $U\rho U^\dagger = \rho \Leftrightarrow U\rho = \rho U$ , da  $U$  unitär ist, q. e. d.

Das folgende Lemma wird die Herleitung eines Kriteriums für die symmetrische Erweiterbarkeit grundlegend vereinfachen. Der Satz und sein Beweis sollten sich auch auf beliebige  $(s, t)$ -Erweiterungen ausdehnen lassen, was hier aber nicht benötigt wird.

**Lemma 6.3 (Invarianz und symmetrische Erweiterbarkeit)**

*Ein  $U_A \otimes U_B$ -invarianter symmetrisch erweiterbarer Zustand  $\rho_{AB}$  besitzt eine  $U_A \otimes U_B \otimes U_B$ -invariante symmetrische Erweiterung.*

*Beweis:* Es sei  $\mathfrak{U}_2$  die durch  $U_A \otimes U_B$  erzeugte abgeschlossene Untergruppe der Gruppe der lokal-unitären Transformationen auf  $\mathcal{H}_A \otimes \mathcal{H}_B$ . Da die unitäre Gruppe auf dem Hilbertraum  $\mathcal{H}_A \otimes \mathcal{H}_B$  kompakt ist, ist somit auch  $\mathfrak{U}_2$  als abgeschlossene Untergruppe kompakt. Entsprechendes gilt für die durch  $U_A \otimes U_B \otimes U_B$  erzeugte abgeschlossene Gruppe  $\mathfrak{U}_3$ . Nach Voraussetzung ist also  $\rho_{AB}$  invariant bezüglich  $\mathfrak{U}_2$ , und es wird gezeigt, daß eine  $\mathfrak{U}_3$ -invariante symmetrische Erweiterung existiert.

Es sei  $\rho_{AB} = \sum_{ijpq} a_{ij,pq} |ij\rangle\langle pq|$  in der separablen Basis, und eine symmetrische Erweiterung dieses Zustands sei  $\rho_{ABB'} = \sum_{ijkpqr} a_{ijk,pqr} |ijk\rangle\langle pqr|$ . Schreibt man die unitären Abbildungen  $U_A$  und  $U_B$  in ihren Eigenbasen, also  $U_A = \sum_x \alpha_x |x\rangle\langle x|$  und  $U_B = \sum_y \beta_y |y\rangle\langle y|$ , so ist die reduzierte Dichtematrix des mit  $U_A \otimes U_B \otimes U_B$  transformierten Operators

$$\begin{aligned} \rho'_{AB} &:= \text{Spur}_{B'} [(U_A \otimes U_B \otimes U_B) \rho (U_A \otimes U_B \otimes U_B)^\dagger] \\ &= \sum_t {}_{B'} \langle t | (U_A \otimes U_B \otimes U_B) \rho (U_A \otimes U_B \otimes U_B)^\dagger | t \rangle_{B'}, \end{aligned} \quad (6.1)$$

wobei es sich anbietet, für die Basis  $(|t\rangle)_{t \in \mathbb{Z}_d}$  von  $\mathcal{H}_{B'}$  die Eigenbasis von  $U_B$  zu wählen. Kommt man überein, daß über alle Indizes summiert wird, wenn nichts anderes angegeben ist, so erhält man

$$\begin{aligned} \rho'_{AB} &= \sum \alpha_x \beta_y \beta_z a_{ijk,pqr} \alpha_u^* \beta_v^* \beta_w^* |xy\rangle \langle xy| ij\rangle \langle pq| uv\rangle \langle uv| \\ &\quad \times \langle t|z\rangle \langle z|k\rangle \langle r|w\rangle \langle w|t\rangle. \end{aligned} \quad (6.2)$$

Führt man die Summe über  $t$  aus, so wird  $\langle t|z\rangle \langle z|k\rangle \langle r|w\rangle \langle w|t\rangle$  zu  $\langle z|k\rangle \langle r|z\rangle$ , und die anschließende Summierung über  $z$  liefert mit  $\sum_z \beta_z \beta_z^* = 1$  nun

$$\begin{aligned} \rho'_{AB} &= \sum \alpha_x \beta_y a_{ijk,pqr} \alpha_u^* \beta_v^* |xy\rangle \langle xy|ij\rangle \langle pq|uv\rangle \langle uv| \cdot \langle r|k\rangle \\ &= \sum_{xyuvijpq} \alpha_x \beta_y \sum_k a_{ijk,pqk} \alpha_u^* \beta_v^* |xy\rangle \langle xy|ij\rangle \langle pq|uv\rangle \langle uv| \\ &= (U_A \otimes U_B) \rho_{AB} (U_A \otimes U_B)^\dagger = \rho_{AB}. \end{aligned} \quad (6.3)$$

Somit ist für jedes beliebige Element  $U \in \mathfrak{U}_3$  auch  $U \rho_{AB} U^\dagger$  eine symmetrische Erweiterung des Zustands  $\rho_{AB}$ . Integriert man über das normierte HAAR-Maß (vgl. Satz B.40) auf  $\mathfrak{U}_3$ , so entsteht der  $\mathfrak{U}_3$ -invariante Zustand  $\int_{U \in \mathfrak{U}_3} U \rho U^\dagger dU$ , der gleichzeitig selbst eine symmetrische Erweiterung des Zustands  $\rho_{AB}$  ist, q. e. d.

Ist  $\mathfrak{A} \subseteq M_n(\mathbb{C})$  eine Unteralgebra der Algebra der  $n \times n$ -Matrizen, so nennt man die Unteralgebra  $\mathfrak{A}' := \{B \in M_n(\mathbb{C}) \mid (\forall A \in M_n(\mathbb{C})) (AB = BA)\}$  ihre *Kommutante*. Nach dem Lemma findet man zu einer  $\mathfrak{U}_2$ -invarianten Dichtematrix, wenn überhaupt, immer auch eine symmetrische Erweiterung in der Kommutanten von  $\mathfrak{U}_3$ .

## 6.2 Symmetrische Erweiterbarkeit

Wie bereits festgehalten wurde, ist es selbst im Falle zweier Qubits schwierig, ein allgemeines Kriterium für die symmetrische Erweiterbarkeit anzugeben. Für die Zwecke dieser Dissertation genügt jedoch ein Kriterium für bell-diagonale Zustände oder noch spezieller, ein Kriterium für die Zustände, die im Unterabschnitt 4.2.6 betrachtet wurden.

In diesem Abschnitt wird die genannte Klasse von Zuständen durch die Invarianz bzgl. einer Gruppe  $\mathfrak{U}_2$  beschrieben, so daß Lemma 6.3 angewendet werden kann. Durch eine Reihe von Rechnungen wird schließlich der Hauptsatz 6.6 hergeleitet, der ein Kriterium für die symmetrische Erweiterbarkeit der untersuchten Zustände beinhaltet. Die Anwendungen dieses Kriteriums in der Quantenkryptographie sind Gegenstand der folgenden Abschnitte.

### 6.2.1 Die kommutative unitäre Gruppe

Es bezeichne  $\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\}$  den komplexen Einheitskreis, und für  $w = (w_0, \dots, w_{d-1}) \in \mathbb{T}^d$  sei  $U_w := \sum_{x=0}^{d-1} w_x |x\rangle \langle x|$  eine in der Standardbasis diagonale unitäre Abbildung. Die Menge aller solchen Abbildungen  $\mathfrak{U}_1 := \{U_w \mid w \in \mathbb{T}^d\}$  möge in dieser Arbeit als *kommutative unitäre Gruppe*

bezeichnet werden; sie ist maximal in dem Sinne, daß es keine größere kommutative Untergruppe der unitären Gruppe auf  $\mathbb{C}^d$  gibt, die sie umfaßt. Im folgenden Text werden die Bezeichnungen  $\mathfrak{U}_2 := \{U_w \otimes U_w^* \mid w \in \mathbb{T}^d\}$  und  $\mathfrak{U}_3 := \{U_w \otimes U_w^* \otimes U_w^* \mid w \in \mathbb{T}^d\}$  verwendet und die  $\mathfrak{U}_2$ - und  $\mathfrak{U}_3$ -invarianten Zustände untersucht.

### 6.2.2 Die $\mathfrak{U}_2$ -invarianten Zustände

Setzt man  $\rho_{AB} = \sum_{ijpq} a_{ij,pq} |ij\rangle\langle pq|$ , so berechnet sich

$$\begin{aligned} (U_w \otimes U_w^*)\rho_{AB} &= \sum_{xyijkl} w_x w_y^* a_{ijkl} |xy\rangle\langle xy| ij\rangle\langle kl| = \sum_{ijkl} w_i w_j^* a_{ijkl} |ij\rangle\langle kl|, \\ \rho_{AB}(U_w \otimes U_w^*) &= \sum_{xyijkl} w_x w_y^* a_{ijkl} |ij\rangle\langle kl| xy\rangle\langle xy| = \sum_{ijkl} w_k w_l^* a_{ijkl} |ij\rangle\langle kl|. \end{aligned}$$

Die Invarianz von  $\rho_{AB}$  bzgl.  $\mathfrak{U}_2$  liegt also genau dann vor, wenn für alle  $i, j, k, l$  die Gleichung  $w_i w_j^* a_{ij,kl} = w_k w_l^* a_{ij,kl}$  erfüllt ist. Ein  $a_{ij,kl}$  darf daher nur dann von Null verschieden sein, wenn  $w_i w_j^* = w_k w_l^*$  oder  $w_i w_l = w_k w_j$  ist, und dies trifft für alle unitären Abbildungen in  $\mathfrak{U}_2$  genau dann zu, wenn  $(i, l) = (j, k)$  oder  $(i, l) = (k, j)$  ist. Nach dem gesagten bleiben also nur Koeffizienten  $a_{ii,kk}$  und  $a_{ij,ij}$  übrig, so daß  $\rho_{AB}$  in Einer- und Zweierblöcke zerfällt.

#### Lemma 6.4 (Bell-diagonale $\mathfrak{U}_2$ -invariante Zustände)

Ein Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ist genau dann  $\mathfrak{U}_2$ -invariant, wenn für alle  $(l, m) \in \mathbb{Z}_d \times \mathbb{Z}_d^*$  die Gleichheit  $A_{lm} = A_{*m}/d$  gilt.

Im Bild der Koeffizientenmatrix sind somit die Einträge in der ersten Spalte beliebig, während sie in jeder der übrigen Spalten alle gleich sein müssen.

*Beweis:* Für den Zustand  $\rho_{AB} = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  berechnet man zunächst

$$\rho_{AB} = d^{-1} \sum_{lmkk'} A_{lm} z^{l(k-k')} |k\rangle\langle k| \otimes |m\rangle\langle k'| \otimes |m\rangle, \quad (6.4)$$

woraus  $a_{kk \otimes m, k'k' \otimes m} = d^{-1} \sum_l A_{lm} z^{l(k-k')}$  folgt. Ist nun  $m = 0$ , so sind alle Koeffizienten  $a_{kk, k'k'}$  möglich. Für  $m \neq 0$  gilt dies nur für  $k = k'$ , andernfalls muß der Koeffizient verschwinden. Die Fouriertransformierte aller Terme außer für  $k = k'$  verschwindet aber nur für konstante Funktionen, also wenn  $A_{lm} = A_{*m}/d$  für alle  $l \in \mathbb{Z}_d$  ist, q. e. d.

Die Zustände, die hier beschrieben wurden, sind also genau dieselben, die im Unterabschnitt 4.2.6 für Primzahlen  $d$  durch die Anwendung der diskreten Mischoperation erzeugt wurden

### 6.2.3 Erweiterungen $\mathfrak{U}_2$ -invarianter Zustände

Die  $\mathfrak{U}_2$ -invarianten bell-diagonalen Zustände besitzen nach Lemma 6.3, wenn überhaupt, eine  $\mathfrak{U}_3$ -invariante symmetrische Erweiterung. Für den Operator  $U := U_w \otimes U_w^* \otimes U_w^*$  berechnet man  $U|ijk\rangle = U_w \otimes U_w^* \otimes U_w^* |ijk\rangle = w_i w_j^* w_k^* |ijk\rangle$ , und für einen Zustand  $\rho = \sum_{ijk,pqr} a_{ijk,pqr} |ijk\rangle \langle pqr|$  ist die  $\mathfrak{U}_3$ -Invarianz zur Forderung äquivalent, daß ein Koeffizient  $a_{ijk,pqr}$  nur dann ungleich Null sein kann, wenn  $w_i w_j^* w_k^* = w_p w_q^* w_r^*$  oder  $w_i w_q w_r = w_p w_j w_k$  gilt.

Da diese Gleichheit für alle  $w \in \mathbb{T}^d$  gefordert wird, ist dies genau dann möglich, wenn  $(i, q, r)$  und  $(p, j, k)$  durch Permutation auseinander hervorgehen; dies definiert eine Äquivalenzrelation auf den Vektoren  $|ijk\rangle$ , die eine Blockmatrixstruktur liefert. Es zeigt sich nun, daß es drei Grundtypen von Blöcken gibt:

1. Ist entweder  $i = j$  oder  $i = k$  oder beides, so entstehen insgesamt  $d$  Blöcke  $B_k$  für  $k \in \mathbb{Z}_d$  der Größe  $2d - 1$ , die durch die Basisvektoren  $|ppk\rangle$  und  $|pkp\rangle$  für  $p \in \mathbb{Z}_d$  erzeugt werden.
2. Sind  $i, j$  und  $k$  allesamt verschieden, so gehören nur  $|ijk\rangle$  und  $|ikj\rangle$  derselben Äquivalenzklasse an, und es entstehen  $d(d-1)(d-2)/2$  Zweierblöcke  $C_{ijk}$ .
3. Ist schließlich  $i \neq j = k$ , so sind die Äquivalenzklassen einelementig, und es entstehen  $d(d-1)$  Einerblöcke  $D_{ij}$ .

Um sich zu vergewissern, daß dies wirklich alle Fälle sind, kann man die Dimensionen der Blöcke überprüfen, was auf das korrekte

$$[d(d-1)(d-2)/2] \cdot 2 + d(d-1) \cdot 1 + d \cdot (2d-1) = d^3 \quad (6.5)$$

führt. Im folgenden soll untersucht werden, ob eine symmetrische Erweiterung mit dieser Blockstruktur gefunden werden kann, was nach Lemma 6.3 stets möglich ist, wenn überhaupt eine solche Erweiterung existiert.

### 6.2.4 Die Spurbedingung

Die Spurbedingung betrifft alle diejenigen Elemente  $a_{ijk,pqr}$ , für die  $k = r$  ist; die anderen Einträge der Matrizen sind frei wählbar. Die Spurbedingung lautet

$$\sum_k a_{ijk,pqk} \stackrel{!}{=} a_{ij,pq} = \begin{cases} d^{-1} \tilde{A}_{ip} := d^{-1} \sum_l A_{l0} z^{l(i-p)} & \text{für } i = j \text{ und } p = q, \\ d^{-1} A_{*m}, m := i \ominus j, & \text{für } i = p \text{ und } j = q, \\ 0 & \text{in allen anderen Fällen.} \end{cases}$$

Falls sowohl der erste als auch der zweite Fall zutrifft ( $i = j = p = q$ ), so liefern sie beide die gleiche Bedingung. Aufgrund der Einschränkung  $k = r$  und der Blockstruktur ist die dritte Spurbedingung bedeutungslos.

Um festzustellen, in welchen Blöcken die Koeffizienten  $a_{ijk,pqk}$  für  $k = r$  liegen, berücksichtigt man, daß  $(i, q, r)$  und  $(p, j, k)$  durch Permutation auseinander hervorgehen müssen; die Forderung  $k = r$  reduziert dies auf die zwei Fälle  $i = j$  und  $p = q$  sowie  $i = p$  und  $j = q$ :

1. Der Koeffizient  $a_{iik,ppk}$  liegt im Block  $B_k$ , und der Ausdruck  $\sum_k a_{iik,ppk}$  beinhaltet je einen Summanden aus jedem der  $d$  Blöcke  $B_k$ .
2. Die Koeffizienten  $a_{ijk,ijk}$  sind genau die Diagonalelemente der Gesamtmatrix und befinden sich daher auf den Diagonalen aller Blöcke.

An dieser Stelle kann man nun festhalten, daß die Außerdiagonalelemente der Blöcke  $C_{ijk}$  auf Null gesetzt werden können, da sie (nach dem HURWITZ-Kriterium) andernfalls nur positivitätsschädigend wirken. Für das Vorliegen einer symmetrischen Erweiterung ist somit die gemeinsame Positivität der Matrizen  $B_k$  entscheidend, wenn nur die Diagonalelemente der Matrizen  $C_{ijk}$  und  $D_{ij}$  sämtlich nicht-negativ sind.

### 6.2.5 Die Reduktion von $B_k$ auf $B'_k$

Der Block  $B_k$  besitzt die Basiselemente  $|ppk\rangle$  und  $|pkp\rangle$  für  $k \neq p$  sowie das Ausnahmeelement  $|kkk\rangle$ . Die Symmetrie verlangt, daß  $a_{iik,ppk} = a_{iki,pkp}$  und  $a_{iik,pkp} = a_{iki,ppk}$  gelten; die erstgenannten Elemente sind spurrelevant, die übrigen, die nur für  $k \notin \{i, p\}$  von Belang sind, nicht.

Setzt man nun willkürlich  $a_{iik,pkp} := a_{iik,ppk}$ , so erhält man eine Matrix mit einer Art Viererblockstruktur und dem Ausnahmeelement  $|kkk\rangle$ . Diese willkürliche Festsetzung ist aber keine Einschränkung: um dies zu zeigen, betrachte man die aus den Basisvektoren  $|ppk\rangle$  (hier auch für  $k = p$ ) gebildete Untermatrix  $B'_k := (a_{iik,ppk})_{i,p=0}^{d-1} \in \mathbb{C}^{d \times d}$ .

#### Lemma 6.5 (Positivität der Matrizen)

*Die Matrizen  $B_k$  und  $B'_k$  sind gleichzeitig positiv semidefinit.*

*Beweis:* Nach dem HURWITZ-Kriterium ist mit  $B_k$  auch die Untermatrix  $B'_k$  positiv semidefinit. Für den Umkehrschluß betrachte man einen Minor von  $B_k$ ; enthält die Menge der Basiselemente, zu denen der Minor gebildet wird, ein Paar  $|ppk\rangle$  und  $|pkp\rangle$ , so verschwindet der Minor, da die zugehörige Untermatrix keinen vollen Rang besitzt. Andernfalls kann man in der Auswahl der Basiselemente alle  $|pkp\rangle$  durch  $|ppk\rangle$  ersetzen, ohne daß sich die zugehörige Untermatrix ändert, und dies ist eine Untermatrix von  $B'_k$ , q. e. d.

In der Matrix  $B'_k$  tauchen die zuvor willkürlich gewählten Elemente  $a_{iik,pkp}$  überhaupt nicht auf, und diese Untermatrix muß nach HURWITZ ohnehin positiv semidefinit sein. Jede andere Wahl der Elemente  $a_{iik,pkp}$  wirkt also höchstens positivitätsschädigend, und dies rechtfertigt die getroffene Wahl.

Indem man sich auf die Matrizen  $B'_k$  beschränkt, sind alle Symmetriebedingungen erschöpfend abgehandelt; die Spurbedingung läßt sich nun für die  $B'_k$  ebenfalls sehr leicht formulieren (vgl. Unterabschnitt 6.2.7).

### 6.2.6 Die Aufstellung der Matrizen $B'_k$

Relevant sind im folgenden nur die Matrizen  $B'_k$ , die übrigen sind positiv semidefinit, sofern nur ihre Diagonaleinträge nicht-negativ sind. Zur Verkürzung der Notation sei  $\lambda_{ijk} := a_{ijk,ijk}$ ; aufgrund der Symmetrie gilt dann  $\lambda_{ijk} = \lambda_{ikj}$  und nach der Spurbedingung  $\sum_k \lambda_{ijk} = d^{-1}A_{*,i\oplus j}$ . Trägt man dies auf, so ergibt sich für jedes feste  $i \in \mathbb{Z}_d$  ein Schema

$j/k$	0	1	...	$i$	...	$d-1$	Zeilen- summe
0	$\lambda_{i00}$	$\lambda_{i01}$	...	$\lambda_{i0i}$	...	$\lambda_{i,0,d-1}$	$d^{-1}A_{*,i}$
1	$\lambda_{i10}$	$\lambda_{i11}$	...	$\lambda_{i1i}$	...	$\lambda_{i,1,d-1}$	$d^{-1}A_{*,i\oplus 1}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$	
$i$	$\lambda_{ii0}$	$\lambda_{ii1}$	...	$\lambda_{iii}$	...	$\lambda_{i,i,d-1}$	$d^{-1}A_{*0}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	
$d-1$	$\lambda_{i,d-1,0}$	$\lambda_{i,d-1,1}$	...	$\lambda_{i,d-1,i}$	...	$\lambda_{i,d-1,d-1}$	$d^{-1}A_{*,i\oplus 1}$
Spalten- summe	$d^{-1}A_{*,i}$	$d^{-1}A_{*,i\oplus 1}$		$d^{-1}A_{*0}$		$d^{-1}A_{*,i\oplus 1}$	$d^{-1}$

mit  $j$  und  $k$  als Zeilen- bzw. Spaltenindex. Dieses Schema ist symmetrisch und enthält nur nicht-negative reelle Einträge. Das durch  $i = j$  oder  $i = k$  definierte „Kreuz“ enthält die Elemente der Matrizen  $B_k$ , die übrigen Diagonalelemente finden sich in den Matrizen  $D_{ij}$ , der Rest in den Matrizen  $C_{ijk}$ . Die Zeilen- und die Spaltensummen ergeben sich aus der zweiten Spurbedingung, die Summe über das gesamte Koeffizientenschema ist also gleich  $1/d$ .

Gibt es nun überhaupt ein Koeffizientenschema, daß die Bedingungen erfüllt und positiv semidefinite Matrizen  $B'_k$  liefert, so gibt es eines, in dem die Matrizen  $C_{ijk}$  verschwinden. Ist zum Beispiel in einem solchen  $C_{ijk}$  ein Eintrag  $\lambda_{ijk} =: x \geq 0$ , so ist wegen der Symmetrie auch  $\lambda_{ikj} = x$ . Setzt man nun  $\lambda'_{ijj} := \lambda_{ijj} + x$ ,  $\lambda'_{ikk} := \lambda_{ikk} + x$  und  $\lambda'_{ijk} := \lambda'_{ikj} := 0$ , so sind nach wie vor alle Einträge nicht-negativ, die Normierungsbedingungen erfüllt, aber  $C_{ijk} = 0$ . Führt man diesen Prozeß für alle Blöcke  $C_{ijk}$  durch, so sieht

man, daß sie allesamt als Null gewählt werden können, während die Matrizen  $D_{ij}$ , das heißt, die Elemente  $\lambda_{ijj} := d^{-1}A_{*,i\ominus j} - \lambda_{ijj}$  für die Einhaltung der Normierung sorgen; die Vergrößerung der Elemente auf dem Kreuz, also der Diagonalelemente der Matrizen  $B'_k$ , schadet niemals der Positivität.

Die Elemente auf dem Kreuz sind also unter Einhaltung der Bedingungen  $\lambda_{iik} \leq d^{-1}A_{*,i\ominus k}$  für alle  $i, k \in \mathbb{Z}_d$  sowie der Normierung  $\sum_{p \in \mathbb{Z}_d} \lambda_{ppk} = d^{-1}x$  für alle  $k \in \mathbb{Z}_d$  möglichst groß zu wählen. Die Wahl  $\lambda_{ijj} := d^{-1}A_{*,i\ominus j} - \lambda_{ijj}$  garantiert dann, daß die Einträge im Koeffizientenschema allesamt nicht-negativ sind und daß die Normierungsbedingungen erfüllt werden.

### 6.2.7 Umformulierung der Spurbedingungen

Die Matrix  $B'_k$  hat die Basiselemente  $|ppk\rangle$  für  $p \in \{0, \dots, d-1\}$ , wobei jetzt auch die Reihenfolge zu berücksichtigen ist. Summiert man nun die Matrizen  $B'_k$ , so lautet die erste Spurbedingung aus Unterabschnitt 6.2.4 jetzt

$$\sum_{k=0}^{d-1} B'_k = \left( \sum_k a_{iik,ppk} \right)_{ip} \stackrel{!}{=} d^{-1}(\tilde{A}_{ip})_{ip} =: \tilde{B}. \quad (6.6)$$

Zusammen mit den bisherigen Ergebnissen zeigt dies den folgenden Satz.

#### Hauptsatz 6.6 (Symmetrisch erweiterbare Zustände)

*Ein  $\mathfrak{U}_2$ -invarianter bell-diagonaler Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  ist genau dann symmetrisch erweiterbar, wenn die Matrix  $\tilde{B} = d^{-1}(\tilde{A}_{ip})_{ip} \in \mathbb{C}^{d \times d}$  mit  $\tilde{A}_{ip} := \sum_l A_{l0} z^{l(i-p)}$ ,  $i, p \in \mathbb{Z}_d$ , als Summe von  $d$  positiv semidefiniten Matrizen  $B'_k = (a_{iik,ppk})_{ip} \in \mathbb{C}^{d \times d}$  für  $k \in \mathbb{Z}_d$  geschrieben werden kann, deren sämtliche Diagonalelemente für alle Werte  $i, k \in \mathbb{Z}_d$  den Bedingungen  $\lambda_{iik} := a_{iik,iik} \leq d^{-1}A_{*,i\ominus k}$  genügen.*

Als erstes einfaches Beispiel für diesen Satz betrachte man einen Zustand, für den  $A_{*0} = 1$  gilt; dieser ist nur dann symmetrisch erweiterbar, wenn  $\tilde{A}_{ip} = 0$  ist, wenn also  $A_{lm} = d^{-1}\delta_{m,0}$  gilt und es sich um den separablen Zustand  $\sigma = d^{-1} \sum_{l=0}^{d-1} |\Psi_{l0}\rangle \langle \Psi_{l0}|$  handelt. Eine mögliche symmetrische Erweiterung ist der GHZ-Zustand  $|\Psi\rangle = d^{-1/2} \sum_k |kkk\rangle$ .

Notwendig dafür, daß ein  $\mathfrak{U}_2$ -invarianter bell-diagonaler Zustand symmetrisch erweiterbar ist, ist, daß zumindest die Matrix  $\tilde{B}$  positiv semidefinit ist. Man stellt fest, daß sie *zirkulant* ist, ihre Einträge also nur von der modularen Differenz zwischen Zeilen- und Spaltenindex abhängen. Die Matrix  $\tilde{B}$  ist nun immer positiv semidefinit, denn für die offensichtlich positiv semidefinite Matrix  $M = \sum_{jq} A_{0j} \delta_{jq} |j\rangle \langle q|$  und die unitäre diskrete Fouriertransformation

$\mathcal{F} = d^{-1/2} \sum_{il} z^{il} |i\rangle \langle l|$  ergibt sich

$$\begin{aligned} \mathcal{F} M \mathcal{F}^{-1} &= d^{-1} \sum_{iljqkp} z^{il} A_{0j} \delta_{jq} z^{-kp} |i\rangle \langle l|j\rangle \langle q|k\rangle \langle p| \\ &= d^{-1} \sum_{ilp} z^{l(i-p)} A_{0l} |i\rangle \langle p| = \tilde{B}. \end{aligned} \quad (6.7)$$

Die Eigenwerte bleiben aber unter solchen Transformationen unverändert, das heißt, auch  $\tilde{B}$  ist stets positiv semidefinit.

Eine positiv semidefinite Matrix bleibt positiv semidefinit, wenn ihre Diagonalelemente vergrößert werden. Am besten wäre es daher, wenn man die Diagonalelemente der  $B'_k$  alle auf ihren maximalen Wert setzen könnte; dies ist aber wegen der Normierung außer für  $A_{*0} = 1$  unmöglich. Man gewinnt aber aus dieser Überlegung eine notwendige Bedingung für die symmetrische Erweiterbarkeit.

**Korollar 6.7 (Bedingung für symmetrische Erweiterbarkeit)**

Ist ein  $\mathfrak{U}_2$ -invarianter bell-diagonaler Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  symmetrisch erweiterbar, so gilt für alle  $i, p \in \mathbb{Z}_d$  mit  $m := i \ominus p$  die Ungleichung

$$|\tilde{A}_{m0}| \leq \sum_{k=0}^{d-1} \sqrt{A_{*k} A_{*,k \oplus m}}.$$

*Beweis:* Ist der Zustand  $\rho$  symmetrisch erweiterbar, so findet man positiv semidefinite Matrizen  $B'_k$  mit den in Hauptsatz 6.6 beschriebenen Eigenschaften. Setzt man die Diagonalelemente dieser Matrizen auf ihren Maximalwert, setzt man also  $\lambda_{ikk} := d^{-1} A_{*,i \ominus k}$ , so bleiben sie positiv semidefinit, und nach dem HURWITZ-Kriterium sind alle Minoren zweiter Ordnung nicht-negativ, q. e. d.

Es ist möglich, aus dem Hauptsatz 6.6 eine Reihe weiterer Bedingungen für die symmetrische Erweiterbarkeit spezieller Zustände herzuleiten, es ist mir aber nicht gelungen, die Frage nach der symmetrischen Erweiterbarkeit  $\mathfrak{U}_2$ -invarianter bell-diagonale Zustände vollständig zu beantworten. Es wird sich aber zeigen, daß dies für die Anwendungen in dieser Dissertation gar nicht erforderlich ist.

### 6.2.8 Vollständige Analyse für Qubits

In diesem Unterabschnitt wird jetzt der Fall zweier Qubits betrachtet, der aufgrund seiner Einfachheit vollständig gelöst werden kann. Die Ergebnisse dienen ausschließlich der Veranschaulichung und werden für den Abschnitt über die Quantenkryptographie nicht benötigt; der Leser kann diesen Unterabschnitt ohne Nachteil überspringen.

Ein bell-diagonaler Zwei-Qubit-Zustand ist verallgemeinert-isotrop und hat die Form  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{\text{bd}}$ ; wegen der  $\mathfrak{U}_2$ -Invarianz gilt die Gleichheit  $\gamma = \delta = (1 - \alpha - \beta)/2$ . Setzt man  $x := \alpha + \beta > 1/2$  (andernfalls ist der Zustand separabel), so sucht man nach Hauptsatz 6.6 zwei positiv semidefinite Matrizen

$$B'_0 = \frac{1}{2} \begin{pmatrix} (1 - \sigma)x & r_0^* \\ r_0 & \tau x \end{pmatrix} \quad \text{und} \quad B'_1 = \frac{1}{2} \begin{pmatrix} \sigma x & r_1^* \\ r_1 & (1 - \tau)x \end{pmatrix}, \quad (6.8)$$

deren Summe

$$\tilde{B} = B'_0 + B'_1 = \frac{1}{2} \begin{pmatrix} x & \alpha - \beta \\ \alpha - \beta & x \end{pmatrix} \quad (6.9)$$

vorgegeben ist. Die zusätzlichen Bedingungen an die Diagonalelemente lauten  $(1 - \sigma)x \leq x$ ,  $\tau x \leq 1 - x$ ,  $\sigma x \leq 1 - x$  und  $(1 - \tau)x \leq x$ , zusammengefaßt also  $\sigma, \tau \in [0; (1 - x)/x]$ . Wegen  $x \geq 1/2$  garantiert dies auch die Nicht-Negativität aller Diagonalelemente, so daß  $B'_0$  und  $B'_1$  genau dann positiv semidefinit sind, wenn ihre Determinanten nicht-negativ sind; dies führt auf die Einschränkungen  $|r_0|^2 \leq (1 - \sigma)\tau x^2$  und  $|r_1|^2 \leq \sigma(1 - \tau)x^2$ . Da die Summe  $r_0 + r_1 = \alpha - \beta$  reell ist, können auch  $r_0$  und  $r_1$  selbst reell gewählt werden, da ein nicht-verschwindender Imaginärteil nur schaden würde. Somit ist die Bedingung

$$\alpha - \beta = r_0 + r_1 \leq x \left[ \sqrt{(1 - \sigma)\tau} + \sqrt{\sigma(1 - \tau)} \right] =: x \cdot f(\sigma, \tau) \quad (6.10)$$

notwendig und hinreichend für die Existenz einer symmetrischen Erweiterung. Man maximiert nun den Ausdruck  $f(\sigma, \tau)$  im erlaubten Bereich; hierzu berechnet man

$$\frac{d}{d\sigma} f(\sigma, \tau) = \frac{1}{2} \left[ \sqrt{(1 - \tau) \cdot \sigma^{-1}} - \sqrt{\tau \cdot (1 - \sigma)^{-1}} \right], \quad (6.11)$$

und für die Ableitung nach  $\tau$  ergibt sich der gleiche Ausdruck, wenn man  $\sigma$  und  $\tau$  vertauscht. Beide Ableitungen sind nicht-negativ für  $1 - \sigma - \tau \geq 0$ . Ist nun  $\sigma + \tau = 1$ , so ist  $f(\sigma, \tau) = 1$  und der Zustand wegen  $\alpha - \beta \leq x$  immer symmetrisch erweiterbar, vorausgesetzt, daß diese Wahl für  $\sigma$  und  $\tau$  überhaupt möglich ist. Dies ist der Fall für  $(1 - x)/x \geq 1/2$  oder  $x \leq 2/3$ .

Andernfalls sind  $\sigma$  und  $\tau$  größtmöglich, also gleich  $(1 - x)/x$  zu wählen, was schließlich auf die Ungleichung

$$\alpha - \beta \leq 2\sqrt{(1 - x)(2x - 1)} \quad (6.12)$$

führt; ausmultipliziert lautet dies  $-9\alpha^2 - 14\alpha\beta - 9\beta^2 + 12\alpha + 12\beta - 4 \geq 0$ , was mit den Ergebnissen von DOHERTY UND RENES übereinstimmt.

## 6.3 Quantenkryptographische Anwendungen

In diesem Abschnitt werden die Auswirkungen der bislang erzielten Ergebnisse im Hinblick auf tolerierbare Fehlerraten in der Quantenkryptographie untersucht. Der Hauptsatz 6.6 vereinfacht die Beantwortung der Frage, ob ein Zustand symmetrisch erweiterbar ist oder nicht, zwar wesentlich, ist aber immer noch zu kompliziert für praktische Anwendungen.

In diesem Abschnitt werden spezielle *Ansätze* für die Matrizen  $B'_k$  untersucht, mit deren Hilfe dann *hinreichende* Kriterien für die symmetrische Erweiterbarkeit hergeleitet werden können. Dies genügt, um die tolerierbaren Fehlerraten zu berechnen.

Wie man den Ausführungen des Kapitels 4 entnehmen kann, konvergiert ein Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  für  $A_{*0} > A_{*m}$ ,  $m \in \mathbb{Z}_d^*$ , im allgemeinen gegen den Zustand  $\sigma = (d^{-1}\delta_{m0})_{l,m=0}^{d-1} = d^{-1} \sum_{l=0}^{d-1} |\Psi_{l0}\rangle\langle\Psi_{l0}| \in \mathcal{S}_{\text{bd}}^{(d)}$ . Es wird daher ein Kriterium gesucht, welches zumindest in einer Umgebung dieses separablen Zustands „gut“ ist.

### 6.3.1 Spezielle Matrizen

In der Nähe des separablen Zustands  $\sigma$  sind alle  $A_{*m}$  für  $m \neq 0$  klein und somit für  $i \neq k$  auch alle  $\lambda_{ik}$  in einer Matrix  $B'_k$ . Die Ausnahmeelemente  $\lambda_{kkk}$  sind groß, d. h., sie liegen nahe bei Eins.

#### Lemma 6.8 (Positiv semidefinite Matrizen)

Sind  $\alpha, \beta_1, \dots, \beta_{d-1} > 0$ , so ist eine Matrix der Form

$$M = \begin{pmatrix} \alpha & \eta_1^* & \eta_2^* & \cdots & \eta_{d-1}^* \\ \eta_1 & \beta_1 & 0 & \cdots & 0 \\ \eta_2 & 0 & \beta_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \eta_{d-1} & 0 & \cdots & 0 & \beta_{d-1} \end{pmatrix} \in \mathbb{C}^{n \times n}$$

genau dann positiv semidefinit, wenn ihre Determinante nicht-negativ ist.

Identifiziert man  $\alpha$  mit dem großen Ausnahmeelement  $\lambda_{kkk}$  und die  $\beta_i$  mit den kleinen Elementen, so erkennt man, daß die Elemente, die in  $M$  den Wert Null haben, nach dem HURWITZ-Kriterium ohnehin kleiner als  $\beta_i\beta_j$ , also „quadratisch klein in  $\beta^{1/2}$ “ sein müssen, während die  $\eta_i$  nur „linear klein in  $\beta^{1/2}$ “ sind. Das Lemma vereinfacht jedoch die Feststellung der positiven Semidefinitheit wesentlich.

*Beweis:* Ist  $M$  positiv semidefinit, so gilt  $\det M \geq 0$  nach dem HURWITZ-Kriterium. Umgekehrt ist zu zeigen, daß alle Minoren der Matrix nicht-negativ sind. Für die Minoren erster Ordnung gilt dies nach Voraussetzung, und diejenigen, die ohne den ersten Basisvektor gebildet werden, sind als Determinanten von Diagonalmatrizen mit nicht-negativen Einträgen offensichtlich positiv semidefinit.

Entwickelt man die Determinante nach der ersten Spalte, so ergibt sich  $\det M = \alpha \operatorname{diag}(\beta_1, \dots, \beta_{d-1}) + \sum_{k=1}^{d-1} (-1)^k \eta_k \det M_{k,0}$ , wobei  $M_{k,0}$  diejenige Matrix bezeichnet, die aus  $M$  entsteht, indem man die  $(k+1)$ -te Zeile und die erste Spalte streicht; es ist also

$$M_{k,0} = \begin{pmatrix} \eta_1^* & \cdots & \eta_{k-1}^* & \eta_k^* & \eta_{k+1}^* & \cdots & \eta_{d-1}^* \\ \beta_1 & 0 & & & & & \\ & \ddots & \ddots & & & & \\ & & \beta_{k-1} & 0 & & & \\ & & & 0 & \beta_{k+1} & & \\ & & & & \ddots & \ddots & \\ & & & & & 0 & \beta_{d-1} \end{pmatrix}. \quad (6.13)$$

Vertauscht man in diesen Matrizen die  $k$ -te Spalte sukzessive mit der jeweils links von ihr stehenden Spalte, so erhält man nach  $k-1$  Vertauschungen eine obere Dreiecksmatrix, deren Determinante bis auf den Faktor  $(-1)^{k-1}$  mit der von  $M_{k,0}$  übereinstimmt. Die Determinante dieser Matrix ist das Produkt ihrer Diagonalelemente, also  $\eta_k \prod_{l \in \mathbb{Z}_d \setminus \{k\}} \beta_l$ ; somit ergibt sich

$$\det M = \alpha \cdot \prod_{l \in \mathbb{Z}_d} \beta_l - \sum_{k=1}^{d-1} \left[ |\eta_k|^2 \cdot \prod_{l \in \mathbb{Z}_d \setminus \{k\}} \beta_l \right], \quad (6.14)$$

was sich zu

$$\det M = \left\{ \alpha \cdot \prod_{l \in \mathbb{Z}_{d-1}} \beta_l - \sum_{k=1}^{d-2} \left[ |\eta_k|^2 \cdot \prod_{l \in \mathbb{Z}_{d-1} \setminus \{k\}} \beta_l \right] \right\} \beta_{d-1} - \left[ |\eta_{d-1}|^2 \cdot \prod_{l \in \mathbb{Z}_d \setminus \{d-1\}} \beta_l \right] \quad (6.15)$$

umschreiben läßt. Der Ausdruck in der geschweiften Klammer ist die Determinante derjenigen Untermatrix von  $M$ , bei der die letzte Zeile und die letzte Spalte gestrichen werden, also ein Minor. Er ist nicht-negativ, denn andernfalls wäre, da  $\beta_{d-1}$  und der letzte Term in eckigen Klammern nicht-negativ sind, auch  $\det M$  negativ, was ausgeschlossen wurde. Entsprechend kann auch jede andere  $(d-1) \times (d-1)$ -Untermatrix betrachtet werden, und setzt man dies absteigend fort, so muß jeder Minor nicht-negativ sein, q. e. d.

Da alle  $\beta_k$  positiv sind, kann die Determinante wie folgt umgeschrieben werden:

$$\det M = \left( \prod_{l \in \mathbb{Z}_d} \beta_l \right) \times \left( \alpha - \sum_{k=1}^{d-1} \frac{|\eta_k|^2}{\beta_k} \right). \quad (6.16)$$

Die rechte Seite kann als eine gewichtete 2-Norm auf  $\mathbb{C}^{d-1}$  betrachtet werden, das heißt, die zulässigen  $\eta = (\eta_1, \dots, \eta_{d-1}) \in \mathbb{C}^{d-1}$  definieren geometrisch einen Ellipsoiden in  $\mathbb{C}^{d-1}$  (oder  $\mathbb{R}^{d-1}$ , wenn man die Beträge  $|\eta_k|$  betrachtet), dessen Achsenlängen  $\sqrt{\alpha \cdot \beta_k}$  sind.

Läßt man  $\beta_i = 0$  zu, so müssen  $\eta_i$  und  $\eta_i^*$  nach den Bedingungen an die Minoren zweiter Ordnung verschwinden, und nach Streichen der zugehörigen Zeilen und Spalten kann man das Lemma 6.8 auf  $(d-1) \times (d-1)$ -Matrizen anwenden; im unbedeutenden Falle  $\alpha = 0$  müssen alle  $\eta_k$  verschwinden.

### 6.3.2 $\mathfrak{U}_2$ -invariante Zustände

Es wird nun ein hinreichendes Kriterium für symmetrische Erweiterbarkeit in der Umgebung des separablen Zustand  $\sigma = (d^{-1}\delta_{m0})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  gesucht. In einer solchen Umgebung sind die Werte  $A_{*m}$  für  $m \neq 0$  klein, das heißt, in einer Matrixzerlegung nach dem Hauptsatz 6.6 müssen die Werte  $\lambda_{iik}$  für  $i \neq k$  ebenfalls klein, die Ausnahmewerte  $\lambda_{kkk}$  also groß sein. Es bietet sich daher an, für die Matrix  $B'_0$  die Form des Lemmas 6.8 anzusetzen; für die Matrizen  $B'_1, B'_2$  usw. wandert das Ausnahmeelement  $\alpha$  (mitsamt den Werten  $\eta_k$ ) die Diagonale entlang, was aber ohne Bedeutung ist.

Es erscheint nun sinnvoll, die kleinen Elemente  $\beta_k$  auf ihren jeweiligen Maximalwert zu setzen und das Ausnahmeelement  $\alpha$  der Normierung anzupassen, denn eine Verkleinerung der ohnehin kleinen Elemente  $\beta_k$  erscheint aufgrund der Bedingungen an die Minoren zweiter Ordnung nicht wünschenswert. Man setzt also die Diagonalelemente der Matrizen  $B'_k$  für  $m \in \mathbb{Z}_d^*$  einschränkend auf  $\beta_m := d^{-1}A_{*m}$ , so daß die Normierung

$$d \cdot \alpha = x - \sum_{m \in \mathbb{Z}_d^*} A_{*m} = x - (1 - x) = 2x - 1 \quad (6.17)$$

erzwingt, wenn  $x := A_{*0}$  gesetzt wird. Alle Diagonalelemente sind in diesem Fall nicht-negativ, und mit der Gleichung (6.16) folgt als Kriterium dafür, daß die Matrix  $M$  positiv semidefinit ist, die Gültigkeit der Ungleichung

$$\sum_{m=1}^{d-1} \frac{|\eta_m|^2}{d^{-1}A_{*m}} \leq \frac{2x-1}{d}. \quad (6.18)$$

Falls ein  $A_{*m}$  verschwinden sollte, so ist die Division durch Null so zu interpretieren, daß das zugehörige  $\eta_m$  verschwinden muß.

Dafür, daß ein vorgegebener Zustand symmetrisch erweiterbar ist, muß nun für alle  $m \in \mathbb{Z}_d^*$  zusätzlich  $\eta_m + \eta_{d \ominus m}^* = d^{-1} \tilde{A}_{m0}$  gelten; dies ist wegen  $\tilde{A}_{ip} = \tilde{A}_{pi}^*$  grundsätzlich erfüllbar. Da in der Ungleichung (6.18) nur die Beträge der  $\eta_m$  auftauchen und diese möglichst klein gewählt werden sollen, können die Argumente (Phasen) von  $\eta_m$  und  $\eta_{d \ominus m}^*$  gleichgesetzt werden: in der GAUßschen Zahlenebene werden die Pfeile kollinear gewählt. Im folgenden werden daher die Argumente der  $\eta_m$  ignoriert und nur noch die Gültigkeit von  $|\eta_m| + |\eta_{d \ominus m}| = d^{-1} |\tilde{A}_{m0}|$  gefordert. Setzt man nun  $|\eta_m| =: \chi_m d^{-1} |\tilde{A}_{m0}|$ , so liest sich die Ungleichung als

$$\sum_{m=1}^{d-1} \chi_m^2 \frac{|\tilde{A}_{m0}|^2}{A_{*m}} \leq 2x - 1 = 1 - 2(1 - x) \quad (6.19)$$

und ist unter der Nebenbedingung  $\chi_m + \chi_{d \ominus m} = 1$  zu erfüllen. Weiter umgeformt lautet sie

$$\sum_{m=1}^{d-1} \left[ \chi_m^2 \frac{|\tilde{A}_{m0}|^2}{A_{*m}} + 2A_{*m} \right] \leq 1. \quad (6.20)$$

Wie schon mehrfach festgestellt wurde, konvergieren die hier betrachteten Zustände unter Anwendung eines  $B_n^{(d)}$ -Schrittes für  $n \rightarrow \infty$  gegen den separablen Zustand  $\sigma$ . Nach Satz 4.6 gilt für die relevanten Parameter nach Anwendung eines  $B_n^{(d)}$ -Schrittes  $A'_{*m} = A_{*m}^n / N$  und  $\tilde{A}'_{m0} = \tilde{A}_{m0}^n / N$  mit der Normierung  $N = \sum_m A_{*m}^n$ , so daß die Ungleichung

$$\sum_{m=1}^{d-1} \left[ \chi_m^2 \frac{1}{N} \left( \frac{|\tilde{A}_{m0}|^2}{A_{*m}} \right)^n + 2 \frac{A_{*m}^n}{N} \right] \leq 1 \quad (6.21)$$

im Grenzfall  $n \rightarrow \infty$  zu untersuchen ist. In diesem Fall strebt der zweite Term der linken Seite,  $2A_{*m}^n / N$ , immer gegen Null und wird daher im folgenden ignoriert werden.

Für den Ausdruck im ersten Term kann man die Abschätzung  $N \geq A_{*0}^n$  verwenden, die im Grenzfall großer  $n \in \mathbb{N}$  sehr gut wird; die fragliche Größe lautet also  $|\tilde{A}_{m0}|^2 / (A_{*0} A_{*m})$ . Ist diese Größe kleiner als Eins, so werden ihre Potenzen für  $n \rightarrow \infty$  verschwinden, ist sie größer als Eins, so muß sie durch ein entsprechend klein gewähltes  $\chi_m$  unterdrückt werden; der Fall, daß sie genau Eins ist, bleibt hier unberücksichtigt.

Ist von den beiden Ausdrücken  $|\tilde{A}_{m0}|^2 / (A_{*0} A_{*m})$  und  $|\tilde{A}_{m0}|^2 / (A_{*0} A_{*, d \ominus m})$  nur einer größer als Eins, so kann  $\chi_m$  bzw.  $\chi_{d \ominus m}$  auf Null gesetzt werden, so daß die Bedingung erfüllt werden kann. Andernfalls können nicht gleichzeitig  $\chi_m$  und  $\chi_{d \ominus m}$  so klein gewählt werden, daß die Bedingungen erfüllt werden, da sie gleichzeitig gegen Null streben und ihre Summe Eins ergeben müßte; dies beweist den folgenden Satz.

**Satz 6.9 (Symmetrische Erweiterbarkeit)**

Im Grenzfall  $n \rightarrow \infty$  ist ein bell-diagonaler  $\mathfrak{U}_2$ -invarianter Zustand nach Anwendung eines  $B_n^{(d)}$ -Schritts zumindest dann symmetrisch erweiterbar und mithin nicht mittels Einweg-Verfahren korrigierbar, wenn für jedes  $m \in \mathbb{Z}_d^*$  die Ungleichung  $|\tilde{A}_{m0}|^2 < A_{*0} \cdot \max\{A_{*m}, A_{*,d \ominus m}\}$  erfüllt wird.

Betrachtet man den scheinbar-isotropen Fall, in dem die  $A_{l0}$  beliebige Werte annehmen können, aber  $A_{*m} = (1-x)/(d-1)$  für alle  $m \neq 0$  ist, so lautet die Bedingung

$$M^2 < x \cdot \frac{1-x}{d-1}, \tag{6.22}$$

wenn man  $M := \max_{l=1}^{d-1} |\tilde{A}_{m0}|$  setzt. Dies ist aber bis auf den Fall der Gleichheit beider Seiten genau komplementär zu der in Kapitel 4 hergeleiteten Bedingung  $M^2 > A_{*0} \max_{m=1}^{d-1} A_{*m}$ , die sich dort aus Gleichung (4.31) und der Forderung  $r^{(d)} > 2$  ergab.

**6.3.3 Der verallgemeinert-isotrope Fall**

Zum Abschluß soll noch der Fall eines verallgemeinert-isotropen Zustands  $\rho = (\alpha, \beta, \gamma, \delta) \in \mathcal{S}_{bd}^{(d)}$  untersucht werden. Die Forderung, daß der Zustand bezüglich der Gruppe  $\mathfrak{U}_2$  invariant ist, führt dazu, daß es nur noch zwei freie nicht-negative Parameter  $\alpha$  und  $\beta$  gibt, die der Bedingung  $\alpha + (d-1)\beta \leq 1$  genügen müssen, und es gilt

$$A_{lm} = \begin{cases} \alpha, & \text{falls } l = m = 0, \\ \beta, & \text{falls } l \neq m = 0, \\ \frac{1-\alpha-(d-1)\beta}{d(d-1)} & \text{falls } m \neq 0 \text{ ist.} \end{cases} \tag{6.23}$$

Setzt man  $x := \alpha + (d-1)\beta$ , so folgt hieraus

$$\tilde{A}_{ip} = \begin{cases} x, & \text{falls } i = p, \\ \alpha - \beta, & \text{falls } i \neq p \text{ ist,} \end{cases} \quad \text{und} \quad A_{*m} = \begin{cases} x, & \text{falls } m = 0, \\ \frac{1-x}{d-1}, & \text{falls } m \neq 0 \text{ ist.} \end{cases}$$

Man erhält unmittelbar  $M = \alpha - \beta$ , also als hinreichende Bedingung für symmetrische Erweiterbarkeit  $(\alpha - \beta)^2(d-1) < x \cdot (1-x)$ . Nach Ausmultiplizieren unter Verwendung von  $x = \alpha + (d-1)\beta$  folgt schließlich

$$\begin{aligned} & (d-1)\alpha^2 - 2(d-1)\alpha\beta + (d-1)\beta^2 \\ & < \alpha + (d-1)\beta - \alpha^2 - 2(d-1)\alpha\beta - (d-1)^2\beta^2 \\ \Leftrightarrow & d\alpha^2 + [(d-1) + (d-1)^2]\beta^2 - [\alpha + (d-1)\beta] < 0 \\ \Leftrightarrow & \alpha^2 + (d-1)\beta^2 - \frac{\alpha + (d-1)\beta}{d} < 0. \end{aligned} \tag{6.24}$$

Die letztgenannte Bedingung ist bis auf den irrelevanten Fall der Gleichheit beider Seiten komplementär zu dem in der Zeile unter Gleichung (4.38) genannten Kriterium.

### 6.3.4 Ergebnisse

In diesem Abschnitt wurde gezeigt, dass die Verwendung anderer Einweg-Korrekturverfahren als die, die in den vorangehenden Kapiteln besprochen wurden, zumindest für die scheinbar-isotropen Zustände keinen Vorteil in bezug auf die maximal tolerierbare Fehlerrate bringt, wenn man den Zustand nach Anwendung eines  $B_n^{(d)}$ -Schritts im Grenzfall  $n \rightarrow \infty$  untersucht.

Zusammenfassend erhält man also im wesentlichen dieselben Ergebnisse wie in den vorangegangenen Kapiteln, aber unter einer weitaus größeren Klasse von Protokollen. Es erübrigt sich, die in den Kapiteln 4 und 5 gezeigten Graphiken noch einmal wiederzugeben; vgl. hierzu die Seiten 73 und 82.



# Kapitel 7

## Schlußbemerkungen

Dieses Kapitel beginnt mit einem Ausblick auf offene Fragen im Zusammenhang mit den erzielten Ergebnissen. In zwei weiteren Abschnitten erfolgt die Zusammenstellung einiger Anmerkungen zur Multi-Qudit-Emission und zu den DE-FINETTI-Sätzen; diese stehen im Zusammenhang zur Quantenkryptographie, ihre Bedeutung für diese Dissertation ist aber gering.

### 7.1 Ausblick auf weitere Untersuchungen

In den Kapiteln 4 bis 6 wurden Protokolle der Quantenkryptographie in endlichdimensionalen Systemen untersucht, die Verschränkungsreinigung mit klassischer Einweg- und Zweiweg-Kommunikation verwenden. Es zeigte sich, daß die dabei auftretenden maximal tolerierbaren Fehlerraten sehr robust gegenüber Modifikationen der Protokolle sind. Dies wirft folgende Fragen auf:

1. Gibt es Zweiweg-Korrekturschritte, mit deren Hilfe höhere Fehlerraten toleriert werden können? (Die bekannten Verfahren sind allesamt mit den  $B_n^{(d)}$ -Schritten verwandt und sollten daher keine höheren Fehlerraten tolerieren können.)
2. Sofern es vom  $B_n^{(d)}$ -Schritt wesentlich verschiedene Zweiweg-Korrekturschritte gibt: ist die Aneinanderreihung von  $B_n^{(d)}$ - und diesen anderen Schritten möglich und bringt sie eventuell Vorteile?
3. Haben die Ergebnisse über engl. *noisy preprocessing* (vgl. z. B. RENNER u. a. [148], KRAUS u. a. [102] und RENES UND SMITH [146]) und engl. *twisted states* (vgl. HORODECKI u. a. [81, 82]) einen Einfluß auf die tolerierbaren Fehlerraten? (Vermutlich nicht, ein wirklicher Beweis steht aber noch aus.)

Es ist weiterhin zu untersuchen, in welchen experimentellen Systemen die Ergebnisse tatsächlich angewendet werden können:

1. Welches physikalische System eignet sich als Qudit und welche Fehler-raten erwartet man bei einer Transmission (theoretisch/experimentell)? Eignen sich beispielsweise Multi-Qubits, die durch zwei parallele Kanäle geschickt werden?
2. Tauchen hierbei weitere Probleme (wie das der Multi-Qudit-Emission oder auch andere) auf?
3. Gibt es Möglichkeiten, die Ergebnisse dieser Untersuchungen in indirekten Sicherheitsbeweisen zu verwenden? (Vgl. z. B. BEAUDRY u. a. [11].)

Hat man eine Anwendung der Ergebnisse gefunden, so müssen für die Implementierung noch weitere Aufgaben gelöst werden:

1. Berechnung und Optimierung der erzielbaren Schlüsselraten, so daß eine Nutzung der Protokolle mit vertretbarem Aufwand möglich wird;
2. Überprüfung, inwieweit die untersuchten idealisierten Modelle tatsächlich eine experimentelle physikalische Situation beschreiben (Abwehr eventueller Seitenkanalangriffe).

## 7.2 Multi-Qudit-Emission

Verwendet man die von einem Laser erzeugten Photonen als Qubit, so hat man in der Quantenkryptographie das Problem, daß die Photonenzahl eines Lasers poisson-verteilt ist; wird mehr als nur ein Photon (im gleichen Polarisationszustand) emittiert, so kann der Lauscher eines der überzähligen Photonen entnehmen und es später, nachdem Alice und Bob sich über die Basen ausgetauscht haben, in der richtigen Basis messen. Dies ermöglicht dem Lauscher vollständige Kenntnis des vorgeblich geheimen Bits; vgl. hierzu LÜTKENHAUS [114], KHALIQUE u. a. [93] und KHALIQUE [92].<sup>1</sup>

Um die gleiche Fragestellung – ohne eine physikalische Erklärung – für Qudits zu untersuchen, betrachte man im verschränkungs-basierten Protokoll den Zustandsvektor  $|\Psi_{00}\rangle$  mit einem zusätzlichen Qudit, also den GHZ-Zustand (benannt nach GREENBERGER, HORN UND ZEILINGER 1989)

$$|\Psi'\rangle = d^{-1/2} \sum_{k=0}^{d-1} |k\rangle_A |k\rangle_B |k\rangle_E.$$

---

<sup>1</sup>Es gibt mehrere Möglichkeiten, diesen Effekt zu bekämpfen: experimentell werden Einphotonenquellen untersucht, und theoretisch kann man versuchen, den Lauscher zu behindern, indem man zwei Laser mit unterschiedlicher Verteilung verwendet und aus der Statistik zurückschließt, ob gelauscht wurde oder nicht (sog. *Decoy*-Protokolle).

Die Dimension von Eves Hilbertraum kann dabei auch größer als  $d$  sein, was hier ohne Bedeutung ist. Bildet man die Teilspur über Eves System, so liefert dies für Alice und Bob den separablen Zustand

$$\sigma := \text{Spur}_E |\Psi'\rangle\langle\Psi'| = d^{-1} \sum_{k=0}^{d-1} |k\rangle_A \langle k| \otimes |k\rangle_B \langle k| = (d^{-1} \delta_{m0})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)},$$

der keine DITfehler besitzt, dessen Phasenwerte aber gleichverteilt sind; dies ist auch der allgemeine Grenzzustand unter  $B_n^{(d)}$ -Schritten für  $n \rightarrow \infty$ .

Im weiteren werde angenommen, daß eine bekannte Verteilung  $P(X = n)$  für die Emission von Multi-Qudits vorliegt und daß  $\Delta \in [0; 1]$  die Wahrscheinlichkeit für solch eine Emission sei. Wird nur ein Qudit emittiert, so wird dies durch den Zustand  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  modelliert.

**Lemma 7.1 ( $B_n^{(d)}$ -Schritte mit Beimischungen)**

*Wird unter den Zuständen  $\rho^{(k)} \in \mathcal{S}_{\text{bd}}^{(d)}$  mindestens ein Zustand durch  $\sigma$  beschrieben, so verbleibt nach einem  $B_n^{(d)}$ -Schritt das erste Qudit-Paar, sofern es überhaupt weiterverwendet wird, im Zustand  $\sigma$ .*

Es ist nicht von vorneherein klar, daß dies auch bedeutet, daß der Lauscher alle Information über ein solches DIT hat, dies kann jedoch durch genauere Untersuchungen gezeigt werden.

*Beweis:* Zum Beweis genügt es, den Satz 4.6 anzuwenden; wegen der Permutationsinvarianz sei daher  $\rho^{(1)} = \sigma$ . Man berechnet

$$\left( \sum_{j=0}^{d-1} z^{ij} A_{jm}^{(0)} \right) = \left( \sum_{j=0}^{d-1} z^{ij} \frac{\delta_{m0}}{d} \right) = \delta_{m0} \cdot \delta_{i0}, \quad (7.1)$$

und hieraus folgt weiter

$$A'_{lm} = \frac{\delta_{m0}}{d \cdot N} \sum_{i=0}^{d-1} \left[ z^{-il} \delta_{i0} \cdot \prod_{k=2}^n \left( \sum_{j=0}^{d-1} z^{ij} A_{jm}^{(k)} \right) \right] = \frac{\delta_{m0}}{d \cdot N} \left[ \prod_{k=2}^n \left( \sum_{j=0}^{d-1} A_{jm}^{(k)} \right) \right].$$

Dies beschreibt den Zustand  $\sigma$ , wenn man

$$N = \sum_{m=0}^{d-1} \left[ \left( \sum_{l=0}^{d-1} \frac{\delta_{m0}}{d} \right) \cdot \prod_{k=2}^n \left( \sum_{l=0}^{d-1} A_{lm}^{(k)} \right) \right] = \prod_{k=2}^n \left( \sum_{l=0}^{d-1} A_{l0}^{(k)} \right) \quad (7.2)$$

berücksichtigt, q. e. d.

Im folgenden bezeichne  $(\rho, \Delta) \in \mathcal{S}_{\text{bd}}^{(d)} \times [0; 1]$  eine Gesamtheit, die zu einem Anteil  $(1 - \Delta)$  aus Zuständen  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$  besteht, denen mit einem Anteil  $\Delta$

Zustände  $\sigma$  beigemischt seien. Nach Lemma 4.6 definiert der  $B_n^{(d)}$ -Schritt eine Abbildung  $B_n^{(d)} : \mathcal{S}_{\text{bd}}^{(d)} \times [0; 1] \rightarrow \mathcal{S}_{\text{bd}}^{(d)} \times [0; 1]$  der Form

$$(\rho, \Delta) \mapsto (\rho', \Delta') := B_n[(\rho, \Delta)] = (B_n(\rho), \Delta'). \quad (7.3)$$

Als Schranke für die Existenz geeigneter CSS-Codes kann für Zustände  $(\rho, \Delta)$  statt der Quanten-SHANNON-Schranke (Satz 4.1) eine Art GLLP-Schranke für Qudits benutzt werden. Ihre Begründung erfolgt sinngemäß zu der von GOTTESMAN u. a. [65]; siehe auch KHALIQUE u. a. [93].

**Lemma 7.2 (Qudit-GLLP-Schranke)**

Es seien  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  und  $\Delta \in [0; 1]$ . Der Zustand  $(\rho, \Delta)$  ist mittels Einweg-Verfahren korrigierbar, falls der Ausdruck

$$1 - H_d[(1 - \Delta) \cdot (A_{*m})_{m=0}^{d-1} + \Delta \cdot (\delta_{m0})_{m=0}^{d-1}] - [\Delta + (1 - \Delta) \cdot H_d((A_{l*})_{l=0}^{d-1})]$$

positiv ist. Dies bedeutet, daß die Dittfehler auf  $(1 - \Delta)\rho + \delta\sigma$  zu korrigieren sind, die Phasenfehler auf  $\rho$  und  $\sigma$  aber getrennt korrigiert werden können.

Der Begriff der asymptotischen Korrigierbarkeit (vgl. Definition 4.2) kann auf  $(\rho, \Delta)$ -Zustände erweitert werden, indem man die Quanten-SHANNON-Schranke durch die Qudit-GLLP-Schranke ersetzt, was auf den folgenden Satz führt.

**Satz 7.3 (Asymptotische Korrigierbarkeit)**

Es seien  $\rho = (A_{lm})_{l,m=0}^{d-1} \in \mathcal{S}_{\text{bd}}^{(d)}$  und  $\Delta < 1$ . Ferner sei  $(S_n)_{n \in \mathbb{N}}$  ein Folge möglicher Korrekturschritte eines Protokolls zur Verschränkungsreinigung. Bezeichnet  $\Delta_n$  den Anteil der Zustände der Form  $\sigma$  nach Anwendung eines  $S_n^{(d)}$ -Schrittes, und verhält sich  $(1 - \Delta_n)$  subexponentiell in  $n \in \mathbb{N}$ , so ist  $(\rho, \Delta)$  genau dann asymptotisch  $S_n^{(d)}$ -korrigierbar, wenn  $\rho$  selbst es ist.

*Beweis:* Verwendet man die Qudit-GLLP-Schranke, so zeigt man analog zu Formel (4.12), daß  $\text{AsymCSS}(\rho') > 0$  genau dann gilt, wenn

$$\frac{-L \cdot H[(1 - \Delta_n)x_n]}{(1 - \Delta_n)y_n^2} - c''(\xi) \frac{x_n}{y_n^2} + 2K - 2\sqrt{2}K'\varepsilon'(p) \cdot y_n > 0$$

mit den dort verwendeten Bezeichnungen ist. Setzt man  $x'_n := (1 - \Delta_n)x_n$  und  $y'_n := \sqrt{1 - \Delta_n}y_n$ , so erhält man die gleiche Form wie in Formel (4.12), wenn man von dem zusätzlichen subexponentiellen Faktor im letzten Term absieht. Die Aussagen sind also die gleichen, q. e. d.

## 7.3 De-Finetti-Sätze

Die Argumentation von GOTTESMAN UND LO in Unterabschnitt 3.5.5 zeigt, daß man anstelle allgemeiner Zustände nur bell-diagonale Zustände untersuchen muß.

Ein anderer Zugang zur Vereinfachung der in quantenkryptographischen Protokollen auftretenden Zustände ist die Verwendung von DE-FINETTI-Sätzen. Die im Kontext von DE-FINETTI-Sätzen wichtigsten Begriffe sind *Symmetrie* und *Austauschbarkeit*.

### Definition 7.4 (Symmetrische und austauschbare Zustände)

Es sei  $\mathcal{H}$  ein Hilbertraum,  $k \in \mathbb{N}$  und  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})$ . Weiterhin bezeichne  $S_n$  die Gruppe der Permutationen (symmetrische Gruppe) von  $n \in \mathbb{N}$  Elementen, und für  $\pi \in S_n$  sei  $U_\pi$  die unitäre Darstellungsmatrix, die die Tensorfaktoren auf  $\mathcal{H}^{\otimes n}$  gemäß  $\pi$  vertauscht. Der Zustand  $\rho$  heißt

1. symmetrisch, falls  $U_\pi \rho U_\pi^\dagger = \rho$  für alle  $\pi \in S_k$  gilt;
2.  $n$ -austauschbar, falls  $n \geq k$  ist und eine symmetrische Dichtematrix  $\rho'$  auf  $\mathcal{H}^{\otimes n}$  derart existiert, daß  $\rho = \text{Spur}_{n-k} \rho'$  ist;<sup>2</sup>
3. (unendlich) austauschbar, falls  $\rho$  für alle Werte  $n \in \mathbb{N} \setminus \{1, \dots, k-1\}$   $n$ -austauschbar ist.

Für Wahrscheinlichkeitsverteilungen erfolgt die Definition sinngemäß, die Teilspur wird durch die Bildung von Randverteilungen ersetzt. Der englische Begriff für „austauschbar“ ist *exchangeable*.

### 7.3.1 Ein geschichtlicher Überblick

Der ursprüngliche Satz von DE FINETTI für austauschbare Wahrscheinlichkeitsverteilungen findet sich auf S. 37–38 seiner Arbeit [53], im ersten Absatz des Kapitels IV, insbesondere Gleichung (19), sowie in früheren italienischsprachigen Arbeiten. DIACONIS UND FREEDMAN [47] geben eine Erweiterung dieses Satzes auf endlich austauschbare Wahrscheinlichkeitsverteilungen an: sie zeigen eine Schranke, die von der endlichen Mächtigkeit des zugrundeliegenden Wahrscheinlichkeitsraums abhängt (Theorem 3 auf S. 746) und eine hiervon unabhängige (Theorem 13 auf S. 749).

Der erste DE-FINETTI-Satz, der auf Quantenzustände anwendbar ist, stammt von STØRMER [173], Theorem 3.1 auf S. 60, der allerdings den

---

<sup>2</sup>Wegen der Symmetrie auf  $\mathcal{H}^{\otimes n}$  ist es unbedeutend, welche der  $n - k$  Tensorfaktoren ausgespart werden.

Namen DE FINETTIS nicht erwähnt und der wegen seiner mathematischen Ausführungen schwierig zu lesen ist; vgl. hierzu auch den Satz von KREIN UND MILMAN, z. B. bei WERNER [184], Theorem VIII.4.4 auf S. 418. Ein einfacherer Beweis dieses Satzes stammt von HUDSON UND MOODY [83], Proposition 1 auf den S. 344–346.<sup>3</sup> Für endlichdimensionale Quantensysteme, also für alle Fälle mit Ausnahme des unendlichdimensionalen separablen Hilbert-raums, findet sich ein Beweis auch bei CAVES, FUCHS UND SCHACK [29], die zum ersten Mal explizit erwähnen, daß ein unendlich-austauschbarer Zustand stets als Konvexkombination von Produktzuständen geschrieben werden kann, er also die Form

$$\rho = \int_{\sigma \in \mathcal{S}(\mathcal{H})} \sigma^{\otimes n} dP(\sigma) \quad (7.4)$$

besitzt, wobei  $P$  ein geeignetes, vom Zustand  $\rho$  abhängiges Wahrscheinlichkeitsmaß auf der Menge  $\mathcal{S}(\mathcal{H})$  der Dichtematrizen auf  $\mathcal{H}$  ist.

### 7.3.2 De-Finetti-Sätze für Quantensysteme

Der DE-FINETTI-Satz für endlich austauschbare Quantenzustände findet sich bei CHRISTANDL u. a. [37], Theorem III.2 auf S. 4. Ein Vorläufer dieses Satzes stammt von KÖNIG UND RENNER [98], Korollar VI.2 auf S.19 in Verbindung mit Korollar V.2 auf S.17. CHRISTANDL u. a. zeigten, daß es für einen  $n$ -austauschbaren Quantenzustand  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})$  einen Zustand  $\rho'$  in der konvexen Hülle von Produktzuständen der Form  $\sigma^{\otimes k}$  derart gibt, daß  $\|\rho - \rho'\|_1 \leq (2d)^2 \cdot k/n$  ist. Diese Schranke ist explizit dimensionsabhängig; die Verfasser zeigten ferner, daß im Gegensatz zu den Ergebnissen von DIACONIS UND FREEDMAN [47] für Wahrscheinlichkeitsverteilungen keine dimensions-unabhängige Form möglich ist (Beispiel II.9 auf S. 481).

Eine hierzu verwandte Schranke, die aber keine Verallgemeinerung und kein Spezialfall der bisher genannten Sätze ist, zeigte RENNER [147] in den Abschnitten 4.1–4.3 auf S. 56–67 seiner Dissertation. Durch die Verwendung approximativer Produktzustände erzielte er eine in  $k/n$  exponentielle Schranke für die Abweichung vom Idealfall. Er betrachtet den symmetrischen Unterraum  $\text{Sym}(\mathcal{H}^{\otimes n}) := \{|\Psi\rangle \in \mathcal{H} \mid (\forall \pi \in S_n)(U_\pi |\Psi\rangle = |\Psi\rangle)\}$  von  $\mathcal{H}^{\otimes n}$  und zeigt, daß eine symmetrische Dichtematrix auf  $\mathcal{H}^{\otimes n}$  eine Purifizierung in  $\text{Sym}((\mathcal{H} \otimes \mathcal{H})^{\otimes n})$  besitzt. Er definiert die Menge  $\text{Sym}(\mathcal{H}^{\otimes n}, |\vartheta\rangle^{\otimes(n-r)}) := \text{Sym}(\mathcal{H}^{\otimes n}) \cap V$  für  $V := \text{span} \{U_\pi(|\vartheta\rangle^{\otimes(n-r)} \otimes |\Psi\rangle) \mid \pi \in S_n, |\Psi\rangle \in \mathcal{H}^{\otimes r}\}$ .

<sup>3</sup>Die Aussage, die auf S. 348 als Erweiterung des Satzes von DE FINETTI für Maße auf orthokomplementären Verbänden (engl. *lattices*) abgeschlossener Unterräume von Hilbert-räumen bezeichnet wird, ist *nicht* die hier als Verallgemeinerung betrachtete Aussage.

Dann betrachtet er einen reinen Zustand  $\rho_{n+k} \in \mathcal{S}(\mathcal{H}^{\otimes(n+k)})$  und eine Zahl  $r \in \{0, \dots, n\}$  und konstruiert ein Wahrscheinlichkeitsmaß<sup>4</sup>  $\nu$  auf der Menge  $\partial B_1(\mathcal{H})$  der normierten Vektoren auf  $\mathcal{H}$  und für jeden Vektor  $|\vartheta\rangle \in \partial B_1(\mathcal{H})$  eine reine Dichtematrix  $\bar{\rho}_n^{|\vartheta\rangle}$  auf  $\text{Sym}(\mathcal{H}^{\otimes n}, |\vartheta\rangle^{\otimes(n-r)})$ . Schließlich zeigt er die Ungleichung

$$\left\| \text{Spur}_k(\rho_{n+k}) - \int_{|\vartheta\rangle \in \partial B_1(\mathcal{H})} \bar{\rho}_n^{|\vartheta\rangle} d\nu(|\vartheta\rangle) \right\|_1 \leq 2 e^{-\frac{k(r+1)}{2(n+k)} + \frac{1}{2} \dim(\mathcal{H}) \ln k}, \quad (7.5)$$

die die genannte exponentielle Schranke darstellt. Man erhält somit das Ergebnis, daß die  $(n+k)$ -austauschbaren Zustände exponentiell nahe an gestörten symmetrischen Zuständen liegen.

### 7.3.3 Anwendungen

Es ist möglich, die DE-FINETTI-Sätze von CHRISTANDL u. a. und RENNER heranzuziehen, um die Reduktion des allgemeinen Zustands  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$  zu bewerkstelligen. In beiden Fällen ist die Aussage aufgrund ihrer Allgemeinheit aber deutlich schwächer als die Argumentation von GOTTESMAN UND LO. Zum einen erhält man keinen Zustand  $\rho^{\otimes n}$  für  $\rho \in \mathcal{S}_{\text{bd}}^{(d)}$ , sondern approximiert Zustände der in Gleichung (7.4) genannten Form. Insbesondere ist zunächst nicht klar, wie sich die Approximation unter der Verschränkungsreinigung verhält, ob zum Beispiel der Abstand vom Zustand der genannten Form unter  $B_n^{(d)}$ -Schritten (in der Spurnorm) nicht wächst.

---

<sup>4</sup>Mir ist aus seinem Beweis allerdings nicht klar, ob das Maß  $\nu$  wirklich ein Wahrscheinlichkeitsmaß ist, daß also  $\nu(\partial B_1(\mathcal{H})) = 1$  gilt.



# Zweiter Hauptteil

## Ergebnisse aus dem Gebiet der Quanteninformationstheorie



## Kapitel 8

# Vollständig positive Abbildungen und der Jamiołkowski-Isomorphismus

Vollständig positive Abbildungen sind in der Quanteninformationstheorie von großem Interesse, da sie in der Beschreibung der Eigenschaften und der zeitlichen Entwicklung von Quantenzuständen in vielfältiger Weise Verwendung finden. Eingeführt wurden sie in rein mathematischem Zusammenhang von STINESPRING [172], in der Physik erlangten sie Bedeutung durch Arbeiten von CHOI [33, 34], KRAUS [100, 101] und LINDBLAD [108].<sup>1</sup>

Völlig unabhängig davon untersuchte DE PILLIS [135] lineare Abbildungen zwischen Matrixalgebren, die die Hermitezität oder Positivität der Matrizen erhalten; seine in einem Punkt lückenhaften Ergebnisse wurden später durch JAMIOŁKOWSKI [85] vervollständigt. An verschiedenen Stellen wurde nun unter Bezugnahme auf diese Arbeit ein „JAMIOŁKOWSKI-Isomorphismus“ eingeführt und ein „JAMIOŁKOWSKI-Kriterium“ für vollständige Positivität verwendet, das in der Originalarbeit aber gar nicht auftaucht.<sup>2</sup>

Ungeachtet dieser Feststellungen ist es sehr wohl möglich, aus der Arbeit JAMIOŁKOWSKIS [85] das erwähnte Kriterium für vollständige Positivität zu gewinnen; tatsächlich liefert eine solche Betrachtung einen einfachen Beweis einer noch allgemeineren Aussage. Zur meiner Verwunderung fand sich dieser Beweis nicht in der veröffentlichten Literatur, weshalb dies in einer eigenständigen Arbeit [141] geschah.

---

<sup>1</sup>Für den *Satz von STINESPRING* vgl. dessen Originalarbeit [172], Theorem 1 auf S. 212, und PAULSEN [132], Kapitel 4 auf S. 43–46; eine etwas allgemeinere Fassung findet sich bei TAKESAKI [174], Theorem 3.6 in Kapitel IV auf S. 194–199.

<sup>2</sup>So findet sich dies zum Beispiel bei BRUß [26] auf S. 4241 über Gleichung (10), bei KEYL UND WERNER [91] in Abschnitt 5.3 auf den S. 76–78 und bei SALGADO u. a. [155] in Definition 2 und Theorem 3 auf S. 57.

## 8.1 Einführung

Da dieses Kapitel vorwiegend mathematischer Natur ist, wird – wie in der Mehrzahl der zuvor genannten Arbeiten und im Gegensatz zum restlichen Text der Dissertation – die in der Mathematik übliche Notation verwendet: die zu einer Zahl  $z \in \mathbb{C}$  konjugiert komplexe Zahl wird also mit  $\bar{z}$  (statt  $z^*$ ) bezeichnet, und die Bezeichnung überträgt sich auf Matrizen. Transponierte und Adjungierte einer Matrix  $A$  werden mit  $A^t$  bzw.  $A^*$  bezeichnet. Skalarprodukte werden in der Form  $\langle \cdot, \cdot \rangle$  geschrieben und sind im linken Argument linear, im rechten antilinear; zur Verdeutlichung wird ggf. als Index der Raum angegeben, auf dem sie wirken.

Es seien  $\mathcal{H}_A := \mathbb{C}^n$  und  $\mathcal{H}_B := \mathbb{C}^m$  zwei endlichdimensionale Hilberträume, denen ihre jeweiligen Matrixalgebren  $\mathfrak{A} := M_n(\mathbb{C})$  und  $\mathfrak{B} := M_m(\mathbb{C})$  zugeordnet seien. Diese Algebren bestehen aus allen Operatoren, die auf  $\mathcal{H}_A$  bzw.  $\mathcal{H}_B$  wirken und bilden bzgl. des HILBERT-SCHMIDT-Skalarproduktes  $\langle A_1, A_2 \rangle = \text{Spur } A_2^* A_1$  für  $A_1, A_2 \in \mathfrak{A}$  und sinngemäß für  $\mathfrak{B}$  selbst Hilberträume (vgl. hierzu auch Unterabschnitt B.8.2).

Der Raum der linearen Abbildungen von  $\mathfrak{A}$  nach  $\mathfrak{B}$  werde mit  $L(\mathfrak{A}, \mathfrak{B})$  bezeichnet. Eine Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  nennt man *hermitesitätserhaltend*, falls sie hermitesche  $A \in \mathfrak{A}$  auf hermitesche  $T(A) \in \mathfrak{B}$  abbildet; man nennt sie *positiv* (genauer wäre *positivitätserhaltend*), falls sie positive  $A$  auf positive  $T(A)$  abbildet. Für jedes  $k \in \mathbb{N}$  erzeugt  $T \in L(\mathfrak{A}, \mathfrak{B})$  eine Abbildung  $T_k := \mathbb{1}_k \otimes T \in L(M_k(\mathfrak{A}), M_k(\mathfrak{B}))$ , wobei  $M_k(\mathfrak{A})$  und  $M_k(\mathfrak{B})$  die Algebren der  $k \times k$ -Matrizen über  $\mathfrak{A}$  bzw.  $\mathfrak{B}$  sind; für eine Matrix  $A = (a_{ij})_{i,j=1}^k$  mit  $a_{ij} \in \mathfrak{A}$  definiert man hierzu

$$T_k \left[ \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} \right] := \begin{pmatrix} T(a_{11}) & \dots & T(a_{1k}) \\ \vdots & \ddots & \vdots \\ T(a_{k1}) & \dots & T(a_{kk}) \end{pmatrix}. \quad (8.1)$$

Die Abbildung  $T$  nennt man nun *k-positiv*, wenn  $T_k$  positiv ist; ist dies für alle  $k \in \mathbb{N}$  der Fall, so bezeichnet man sie als *vollständig positiv*. Allgemeine Untersuchungen dieser Begriffe finden sich bei CHOI [33].

In physikalischer Sprechweise kann  $T_k$  als Kopplung des Systems an ein Hilfssystem (eine *Ancilla*) der Dimension  $k$  verstanden werden, wobei aber keinerlei Operation auf dem Hilfssystem ausgeführt wird. Ist die Verwendung eines solchen Hilfssystems erlaubt, so muß eine quantenmechanisch zulässige Operation *k-positiv* sein. Läßt man beliebig große Hilfssysteme zu, so muß jede Quantenoperation vollständig positiv sein.

## 8.2 Der Jamiołkowski-Isomorphismus

In diesem Abschnitt werden zwei Abbildungen  $\mathcal{J}_1$  und  $\mathcal{J}_2$  eingeführt, die in der Literatur beide unter dem Namen JAMIOŁKOWSKI-*Isomorphismus* bekannt zu sein scheinen. Im folgenden heiÙe  $\mathcal{J}_1$  der JAMIOŁKOWSKI-*Isomorphismus* und  $\mathcal{J}_2$  der *modifizierte* JAMIOŁKOWSKI-*Isomorphismus*.

DE PILLIS [135] untersuchte die Abbildung  $\mathcal{J}_1 : L(\mathfrak{A}, \mathfrak{B}) \rightarrow \mathfrak{A} \otimes \mathfrak{B}$ , die durch die Forderung, daÙ  $\langle \mathcal{J}_1(T), A^* \otimes B \rangle_{\mathfrak{A} \otimes \mathfrak{B}} = \langle T(A), B \rangle_{\mathfrak{B}}$  für alle  $T \in L(\mathfrak{A}, \mathfrak{B})$ ,  $A \in \mathfrak{A}$  und  $B \in \mathfrak{B}$  gelte, definiert ist, und zeigte folgende Eigenschaften dieser Abbildung.

### Lemma 8.1 (Eigenschaften von $\mathcal{J}_1$ )

Die Abbildung  $\mathcal{J}_1$  ist wohldefiniert, und für jede Orthonormalbasis  $(E_i)_{i \in I}$  von  $\mathfrak{A}$  und jeden Operator  $T \in \mathfrak{A}$  gilt die Gleichung  $\mathcal{J}_1(T) = \sum_{i \in I} E_i^* \otimes T(E_i)$ . Ferner ist  $\mathcal{J}_1$  ein isometrischer Isomorphismus der Hilberträume  $L(\mathfrak{A}, \mathfrak{B})$  und  $\mathfrak{A} \otimes \mathfrak{B}$ .

Der modifizierte JAMIOŁKOWSKI-Isomorphismus wird wie folgt konstruiert: Man betrachtet die Basis  $(E_{ij})_{i,j=1}^n$  von  $\mathfrak{A}$ , die aus den Matrizen  $E_{ij}$  besteht, deren Eintrag in der  $j$ -ten Spalte der  $i$ -ten Zeile Eins ist, während alle anderen Einträge verschwinden (manchmal als WEYL-Basis bezeichnet). Für alle  $i, j \in \{1, \dots, n\}$  gilt dann  $E_{ij} = \overline{E_{ij}} = E_{ji}^* = E_{ji}^t$ . Der modifizierte JAMIOŁKOWSKI-Isomorphismus  $\mathcal{J}_2 : L(\mathfrak{A}, \mathfrak{B}) \rightarrow \mathfrak{A} \otimes \mathfrak{B}$  wird nunmehr durch die Festsetzung<sup>3</sup>

$$\mathcal{J}_2(T) := \sum_{i,j=1}^n E_{ij} \otimes T(E_{ij}) \quad (8.2)$$

definiert. Der Unterschied zum Isomorphismus  $\mathcal{J}_1$  ist, daÙ im ersten Tensorfaktor die Adjunktion fehlt. Für  $T_1, T_2 \in L(\mathfrak{A}, \mathfrak{B})$  berechnet man mithilfe von  $\langle E_{ij}, E_{kl} \rangle = \delta_{ik} \delta_{jl}$ , daÙ

$$\begin{aligned} \langle \mathcal{J}_1(T_1), \mathcal{J}_1(T_2) \rangle &= \left\langle \sum_{ij} E_{ij}^* \otimes T_1(E_{ij}), \sum_{kl} E_{kl}^* \otimes T_2(E_{kl}) \right\rangle \\ &= \sum_{ijkl} \langle E_{ij}^*, E_{kl}^* \rangle \langle T_1(E_{ij}), T_2(E_{kl}) \rangle \\ &= \sum_{ij} \langle T_1(E_{ij}), T_2(E_{ij}) \rangle \end{aligned} \quad (8.3)$$

gilt; es folgt  $\langle \mathcal{J}_1(T_1), \mathcal{J}_1(T_2) \rangle = \langle \mathcal{J}_2(T_1), \mathcal{J}_2(T_2) \rangle$  mithilfe einer ähnlichen Rechnung für  $\mathcal{J}_2$ , und somit ist  $\mathcal{J}_2$  in der Tat ein Isomorphismus.

---

<sup>3</sup>Während JAMIOŁKOWSKI [85] ausschließlich die Abbildung  $\mathcal{J}_1$  verwendet, findet sich  $\mathcal{J}_2$  implizit bei CHOI [34], zwischen Theorem 1 und Theorem 2 auf S. 286, ferner bei KEYL UND WERNER [91], S. 76–77, und bei SALGADO u. a. [156], Definition 2 auf S. 57. Einige Arbeiten verwenden auch die Bezeichnung CHOI-JAMIOŁKOWSKI-*Isomorphismus*.

Ein wesentlicher Unterschied zwischen den beiden Isomorphismen ist, daß  $\mathcal{J}_1$  basisunabhängig definiert wurde, während  $\mathcal{J}_2$  ausdrücklichen Bezug auf die Basis nimmt. Betrachtet man zum Beispiel eine weitere Basis  $(F_{ij})_{i,j=1}^n$  von  $\mathfrak{A}$ , so können ihre Elemente in der Form  $F_{ij} = \sum_{kl} \langle F_{ij}, E_{kl} \rangle_{\mathfrak{A}} E_{kl}$  ausgedrückt werden. Verwendet man diese Basis, um einen weiteren modifizierten JAMIOŁKOWSKI-Isomorphismus  $\mathcal{J}'_2$  zu definieren, so folgt

$$\mathcal{J}'_2(T) = \sum_{klpq} \left[ \sum_{ij} \langle F_{ij}, E_{kl} \rangle \langle F_{ij}, E_{pq} \rangle \right] E_{kl} \otimes T(E_{pq}), \quad (8.4)$$

und dies ist genau dann für alle  $T$  gleich  $\mathcal{J}_2(T)$ , wenn die innere Klammer  $\delta_{kp}\delta_{lq}$  ergibt. Als Beispiel betrachte man die kanonische Basis  $(e_i)_{i=1}^n$  von  $\mathcal{H}_A$  und einen unitären Operator  $U$  auf  $\mathcal{H}_A$ ; die Vektoren  $f_i := Ue_i$  bilden dann ebenfalls eine Basis von  $\mathcal{H}_A$ . Setzt man  $F_{ij} := \langle \cdot, f_j \rangle f_i$ , so kann gezeigt werden, daß  $\mathcal{J}_2 = \mathcal{J}'_2$  genau dann gilt, wenn  $U$  in dem Sinne *orthogonal* ist, daß  $U^t U = \mathbb{I}$  ist. Es gibt jedoch unitäre, aber nicht orthogonale Matrizen, und somit ist  $\mathcal{J}_2$  nicht für alle Basen gleich.

### 8.3 Positive Abbildungen

In diesem Abschnitt werden die Ergebnisse von DE PILLIS und JAMIOŁKOWSKI für die beiden Isomorphismen gezeigt. Ausgehend von diesen Ergebnissen wird im folgenden Abschnitt der Hauptsatz dieses Kapitels gezeigt werden. Das folgende Lemma charakterisiert die hermitezitätserhaltenden Abbildungen und wurde schon von DE PILLIS verwendet.

**Lemma 8.2 (Hermitezitätserhaltende Abbildungen)**

*Eine lineare Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  erhält die Hermitezität genau dann, wenn  $T(A^*) = T(A)^*$  für alle  $A \in \mathfrak{A}$  gilt. Ist  $(E_i)_{i \in I}$  eine Basis von  $\mathfrak{A}$ , so ist dies genau dann der Fall, wenn  $T(E_i^*) = T(E_i)^*$  für alle  $i \in I$  gilt.*

*Beweis:* Erhält  $T$  die Hermitezität, so gilt  $T(A)^* = T(A) = T(A^*)$  für hermitesches  $A$ ; allgemein läßt sich ein beliebiges  $A \in \mathfrak{A}$  als  $A = A_1 + iA_2$  zerlegen, wobei  $A_1, A_2 \in \mathfrak{A}$  hermitesch sind, so daß

$$\begin{aligned} T(A^*) &= T(A_1^* - iA_2^*) = T(A_1^*) - iT(A_2^*) \\ &= T(A_1)^* - iT(A_2)^* = [T(A_1) + iT(A_2)]^* = T(A)^* \end{aligned} \quad (8.5)$$

folgt. Umgekehrt berechnet man  $T(A) = T(A^*) = T(A)^*$  für hermitesches  $A \in \mathfrak{A}$ , was die erste Behauptung zeigt. Für die zweite Behauptung verwendet man die Zerlegung  $A = \sum_i a_i E_i$  und erhält hiermit  $T(A^*) = \sum_{i \in I} \bar{a}_i T(E_i^*) = \sum_{i \in I} \bar{a}_i T(E_i)^* = T(A)^*$ , wobei die zweite Gleichheit nach Voraussetzung gilt. Der Umkehrschluß ist offensichtlich, q. e. d.

Unter Zuhilfenahme dieses Lemmas kann der folgende Satz bewiesen werden.

**Satz 8.3 (Hermitesitätserhaltende und positive Abbildungen)**

*Steht  $\mathcal{J}$  entweder für  $\mathcal{J}_1$  oder  $\mathcal{J}_2$ , so ist eine Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  genau dann hermitesitätserhaltend, wenn  $\mathcal{J}(T)$  hermitesch ist; sie ist genau dann positiv, wenn  $\langle \mathcal{J}(T)x \otimes y, x \otimes y \rangle \geq 0$  für alle  $x \in \mathcal{H}_A$  und alle  $y \in \mathcal{H}_B$  ist.*

Für  $\mathcal{J}_1$  wurde der erste Teil von DE PILLIS, [135], Proposition 1.2 auf S. 133, gezeigt, der zweite von JAMIOLKOWSKI [85], Theorem 1 auf S. 276. Im folgenden werden daher nur die Beweise für  $\mathcal{J}_2$  aufgeführt, die aber leicht auf  $\mathcal{J}_1$  übertragen werden können.

*Beweis:* Zum Beweis der ersten Aussage berechnet man

$$\begin{aligned} \mathcal{J}_2(T)^* &= \left( \sum_{ij} E_{ij} \otimes T(E_{ij}) \right)^* = \sum_{ij} E_{ij}^* \otimes T(E_{ij})^* \\ &\stackrel{?}{=} \sum_{ij} E_{ji} \otimes T(E_{ij}^*) = \sum_{ij} E_{ji} \otimes T(E_{ji}) = \mathcal{J}_2(T) \end{aligned} \quad (8.6)$$

und stellt fest, daß nach dem zweiten Teil des Lemmas 8.2 die fragliche Gleichheit genau dann gilt, wenn  $T$  die Hermitesität erhält.

Der Beweis des zweiten Teils ist dem ursprünglichen Beweis JAMIOLKOWSKIS sehr ähnlich und verläuft wie folgt: Da jeder positive Operator spektral in eine positive reelle Linearkombination eindimensionaler Projektionsoperatoren zerlegt werden kann, genügt es, die Behauptung für ebendiese zu zeigen. Für den durch einen Einheitsvektor  $x \in \mathcal{H}$  definierten Projektionsoperator  $P_x = \langle \cdot, x \rangle x$  ist zu zeigen, daß  $T(P_x)$  positiv ist. Ist  $(f_p)_{p=1}^n$  eine Orthonormalbasis von  $\mathcal{H}_A$ , so berechnet man

$$\begin{aligned} T(P_x) &= \sum_{ij} \langle P_x, E_{ij} \rangle_{\mathfrak{A}} T(E_{ij}) = \sum_{ij} \text{Spur}(E_{ij}^* P_x) T(E_{ij}) \\ &= \sum_{ijp} \langle E_{ji} P_x f_p, f_p \rangle_{\mathcal{H}_A} T(E_{ij}) = \sum_{ijp} \langle E_{ji} x, f_p \rangle_{\mathcal{H}_A} \langle f_p, x \rangle_{\mathcal{H}_A} T(E_{ij}) \\ &= \sum_{ij} \langle E_{ji} x, x \rangle_{\mathcal{H}_A} T(E_{ij}). \end{aligned} \quad (8.7)$$

Somit ist  $T(P_x)$  genau dann positiv, wenn  $\sum_{ij} \langle E_{ji} x, x \rangle_{\mathcal{H}_A} \langle T(E_{ij}) y, y \rangle_{\mathcal{H}_B} \geq 0$  für alle  $x \in \mathcal{H}_A$  und alle  $y \in \mathcal{H}_B$  ist. Schreibt man  $x = (x_1, \dots, x_n)^t \in \mathcal{H}_A$  und bezeichnet mit  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^t \in \mathcal{H}_A$  den Vektor der konjugiert komplexen Einträge, so ist  $\langle E_{ji} x, x \rangle = x_i \cdot \bar{x}_j = \langle E_{ij} \bar{x}, \bar{x} \rangle$ , und hieraus folgt

$$\begin{aligned} \sum_{ij} \langle E_{ji} x, x \rangle_{\mathcal{H}_A} \langle T(E_{ij}) y, y \rangle_{\mathcal{H}_B} &= \sum_{ij} \langle E_{ij} \bar{x}, \bar{x} \rangle_{\mathcal{H}_A} \langle T(E_{ij}) y, y \rangle_{\mathcal{H}_B} \\ &= \langle \mathcal{J}_2(T) \bar{x} \otimes y, \bar{x} \otimes y \rangle, \end{aligned} \quad (8.8)$$

was für alle  $x \in \mathcal{H}_A$  und alle  $y \in \mathcal{H}_B$  gelten muß. Da mit  $x \in \mathcal{H}_A$  auch  $\bar{x} \in \mathcal{H}_A$  (wegen  $\mathcal{H}_A = \mathbb{C}^n$ ) ist und umgekehrt, ist der Satz gezeigt, q. e. d.

## 8.4 Das Jamiołkowski-Kriterium

Das Material aus dem vorangehenden Abschnitt erlaubt es, den Hauptsatz dieses Kapitels zu beweisen, der, wie CLARISSE [39] feststellt, implizit in einer Arbeit von TERHAL UND HORODECKI [176] enthalten ist.

Für Zustandsvektoren auf dem Tensorprodukt  $\mathcal{H}_A \otimes \mathcal{H}_B$  zweier endlich-dimensionaler Hilberträume ist die SCHMIDT-Zahl definiert (Satz B.24); die konvexe Hülle der den Vektoren einer SCHMIDT-Zahl nicht größer als  $k \in \mathbb{N}$  zugeordneten Dichtematrizen bildet die SCHMIDT-Klasse  $S_k$ . Somit ist  $S_1$  die Menge der separablen Zustände, und für  $k = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$  ist  $S_k$  die Gesamtheit aller Zustände. Allgemeiner kann man die Zugehörigkeit zu einer Klasse  $S_k \setminus S_{k-1}$  als ein Verschränkungsmaß ansehen.

Die zweite Aussage des Satzes 8.3 kann wie folgt interpretiert werden: *Eine Abbildung  $T$  ist genau dann positiv, wenn  $\mathcal{J}_2(T)$  positiv auf separablen Vektoren ist.* Die separablen Vektoren sind die Vektoren der SCHMIDT-Zahl Eins. Der folgende Satz verallgemeinert diese Beobachtung.

### Hauptsatz 8.4 (Jamiołkowski-Kriterium)

*Eine lineare Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  ist genau dann  $k$ -positiv, wenn die Ungleichung  $\langle \mathcal{J}_2(T)v, v \rangle \geq 0$  für alle Vektoren  $v \in \mathcal{H}_A \otimes \mathcal{H}_B$  gilt, die eine SCHMIDT-Zahl nicht größer als  $k$  besitzen.*

Die Grundidee des Beweises ist sehr einfach: nach Definition ist  $T$  genau dann  $k$ -positiv, wenn der Operator  $T_k := \mathbb{1}_k \otimes T \in L(M_k(\mathfrak{A}), M_k(\mathfrak{B}))$  positiv ist. Diese Eigenschaft kann über das modifizierte JAMIOŁKOWSKI-Kriterium aus Satz 8.3 überprüft werden, was nun geschehen soll.

*Beweis:* Es bezeichnen in diesem Beweis  $\mathcal{H}_{A;k} := \mathbb{C}^k \otimes \mathcal{H}_A \cong \bigoplus_{\alpha=1}^k \mathcal{H}_A$  und  $\mathcal{H}_{B;k} := \mathbb{C}^k \otimes \mathcal{H}_B \cong \bigoplus_{\beta=1}^k \mathcal{H}_B$  zwei Hilberträume mit ihren zugeordneten Matrixalgebren  $\mathfrak{A}_k := M_k(\mathfrak{A})$  und  $\mathfrak{B}_k := M_k(\mathfrak{B})$ . Der Operator  $T_k$  ist nach Satz 8.3 positiv, wenn für alle  $x \in \mathcal{H}_{A;k}$  und  $y \in \mathcal{H}_{B;k}$  die Ungleichung  $\langle \mathcal{J}_{2;k}(T_k)x \otimes y, x \otimes y \rangle \geq 0$  gilt, wobei  $\mathcal{J}_{2;k} : L(\mathfrak{A}_k, \mathfrak{B}_k) \rightarrow \mathfrak{A}_k \otimes \mathfrak{B}_k$  der modifizierte JAMIOŁKOWSKI-Isomorphismus auf den jeweiligen Räumen sei.

Für je eine Orthonormalbasis  $(E_{ij})_{i,j=1}^n$  von  $\mathfrak{A}$  und  $(e_{\alpha\beta})_{\alpha,\beta=1}^k$  von  $M_k(\mathbb{C})$  ist  $(e_{\alpha\beta} \otimes E_{ij})_{i,j=1}^n, \alpha,\beta=1}^k$  eine Orthonormalbasis von  $M_k(\mathfrak{A})$ ; man berechnet

$$\begin{aligned} \mathcal{J}_{2;k}(\mathbb{1}_k \otimes T) &= \sum_{i,j=1}^n \sum_{\alpha,\beta=1}^k e_{\alpha\beta} \otimes E_{ij} \otimes [\mathbb{1}_k \otimes T](e_{\alpha\beta} \otimes E_{ij}) \\ &= \sum_{i,j=1}^n \sum_{\alpha,\beta=1}^k e_{\alpha\beta} \otimes E_{ij} \otimes e_{\alpha\beta} \otimes T(E_{ij}). \end{aligned} \quad (8.9)$$

Bezeichnet  $(f_p)_{p=1}^k$  die kanonische Basis des Raums  $\mathbb{C}^k$ , so kann jeder Vektor  $x \in \mathcal{H}_{A;k}$  in der Form  $x = \sum_{p=1}^k f_p \otimes x_p$  mit geeigneten Elementen  $x_p \in \mathcal{H}_A$  geschrieben werden; entsprechend schreibt man  $y = \sum_{q=1}^k f_q \otimes y_q \in \mathcal{H}_{B;k}$  für Elemente  $y_q \in \mathcal{H}_B$ . Nach Satz 8.3 sind Bedingungen dafür zu finden, daß die Ungleichung

$$\langle \mathcal{J}_{2;k}(\mathbb{1}_k \otimes T)x \otimes y, x \otimes y \rangle \geq 0 \quad (8.10)$$

für alle  $x \in \mathcal{H}_{A;k}$  und alle  $y \in \mathcal{H}_{B;k}$  gilt. Setzt man  $x \otimes y = \sum_{pq} f_p \otimes x_p \otimes f_q \otimes y_q = \sum_{rs} f_r \otimes x_r \otimes f_s \otimes y_s$  und unterdrückt die Summierungsindizes, so schreibt sich die linke Seite dieser Bedingung als

$$\begin{aligned} & \sum \langle e_{\alpha\beta} f_p \otimes E_{ij} x_p \otimes e_{\alpha\beta} f_q \otimes T(E_{ij}) y_q, f_r \otimes x_r \otimes f_s \otimes y_s \rangle_{\mathbb{C}^k \otimes \mathcal{H}_A \otimes \mathbb{C}^k \otimes \mathcal{H}_B} \\ &= \sum \langle e_{\alpha\beta} f_p, f_r \rangle_{\mathbb{C}^k} \cdot \langle E_{ij} x_p, x_r \rangle_{\mathcal{H}_A} \cdot \langle e_{\alpha\beta} f_q, f_s \rangle_{\mathbb{C}^k} \cdot \langle T(E_{ij}) y_q, y_s \rangle_{\mathcal{H}_B} \quad (8.11) \\ &= \sum_{pqrs} \left[ \sum_{\alpha\beta} \langle e_{\alpha\beta} f_p, f_r \rangle_{\mathbb{C}^k} \langle e_{\alpha\beta} f_q, f_s \rangle_{\mathbb{C}^k} \right] \langle \mathcal{J}_2(T) x_p \otimes y_q, x_r \otimes y_s \rangle_{\mathcal{H}_A \otimes \mathcal{H}_B}. \end{aligned}$$

Mit  $\langle e_{\alpha\beta} f_p, f_r \rangle_{\mathbb{C}^k} = \delta_{\alpha r} \delta_{\beta p}$  vereinfacht sich die eckige Klammer zu

$$\sum_{\alpha\beta} \langle e_{\alpha\beta} f_p, f_r \rangle_{\mathbb{C}^k} \langle e_{\alpha\beta} f_q, f_s \rangle_{\mathbb{C}^k} = \left( \sum_{\beta} \delta_{\beta p} \delta_{\beta q} \right) \left( \sum_{\alpha} \delta_{\alpha r} \delta_{\alpha s} \right) = \delta_{pq} \delta_{rs},$$

woraus sich mit  $v := \sum_{p=1}^k x_p \otimes y_p$  schließlich

$$\sum_{p,r=1}^k \langle \mathcal{J}_2(T) x_p \otimes y_p, x_r \otimes y_r \rangle_{\mathcal{H}_A \otimes \mathcal{H}_B} = \langle \mathcal{J}_2(T) v, v \rangle_{\mathcal{H}_A \otimes \mathcal{H}_B} \quad (8.12)$$

ergibt. Da die Hilberträume, die durch die Vektorsysteme  $(x_p)_{p=1}^k$  und  $(y_p)_{p=1}^k$  aufgespannt werden, höchstens die Dimension  $k$  besitzen, ist die SCHMIDT-Zahl von  $v$  ebenfalls nicht größer als  $k$ , q. e. d.

Man beachte, daß dieser Satz falsch ist, wenn man  $\mathcal{J}_1$  statt  $\mathcal{J}_2$  benutzen wollte: die eckige Klammer in der Gleichung (8.11) lautete in diesem Fall nämlich  $\langle e_{\alpha\beta}^* f_p, f_r \rangle_{\mathbb{C}^k} = \delta_{\alpha p} \delta_{\beta q}$ , und dies führte auf ein anderes Ergebnis. Als Beispiel, daß das Ergebnis für  $\mathcal{J}_1$  falsch ist, betrachte man  $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$  mit  $\mathfrak{A} = \mathfrak{B} = \mathbb{C}^{2 \times 2}$ ; es sei  $T : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$  die Identität, die offensichtlich vollständig positiv ist. Der Vektor  $v := e_1 \otimes e_2 - e_2 \otimes e_1 \in \mathbb{C}^2$  ist aber ein Eigenvektor von  $\mathcal{J}_1(T)$  zum Eigenwert  $-1$ , so daß  $\langle \mathcal{J}_1(T)v, v \rangle$  negativ ist. Er hat die SCHMIDT-Zahl 2, und daher ist das Analogon des Hauptsatzes 8.4 mit  $\mathcal{J}_1$  statt  $\mathcal{J}_2$  falsch.

Schließlich seien noch zwei Folgerungen aus dem Hauptsatz genannt: die erste Folgerung ist ein Spezialfall viel allgemeinerer Sätze von CHOI [33], Theoreme 5 bis 8 auf S. 525–529, die zweite ist als das sogenannte JAMIÓŁKOWSKI-Kriterium für vollständige Positivität bekannt (vgl. z. B. CHOI [34], Theorem 2 auf S. 286, KEYL UND WERNER [91], Theorem 5.2 auf S. 78, und SALGADO u. a. [156], Theorem 3 auf S. 57).

**Korollar 8.5 (Vollständig positive Abbildungen)**

Eine  $\min\{n, m\}$ -positive Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  ist vollständig positiv.

*Beweis:* Nach dem Hauptsatz 8.4 ist die Abbildung  $T$  genau dann  $k$ -positiv, wenn sie positiv auf allen Vektoren einer SCHMIDT-Zahl nicht größer als  $k$  ist. Da aber jeder Vektor aus  $\mathcal{H}_A \otimes \mathcal{H}_B$  eine SCHMIDT-Zahl nicht größer als  $\min\{n, m\}$  besitzt, ist jede  $\min\{n, m\}$ -positive Abbildung vollständig positiv, q. e. d.

**Korollar 8.6 (Jamiołkowski-Kriterium im engeren Sinne)**

Eine lineare Abbildung  $T \in L(\mathfrak{A}, \mathfrak{B})$  ist genau dann vollständig positiv, wenn  $\mathcal{J}_2(T)$  positiv semidefinit ist.

*Beweis:* Ist die Abbildung  $T$  vollständig positiv, so ist sie nach dem Hauptsatz 8.4 positiv auf den Vektoren beliebiger SCHMIDT-Zahl. Da jeder Vektor eine wohldefinierte endliche SCHMIDT-Zahl besitzt, ist  $T$  positiv semidefinit. Der Umkehrschluß ist offensichtlich, q. e. d.

## 8.5 Schlußbemerkungen

Abschließend soll kurz die Frage erörtert werden, inwieweit die Ergebnisse des vorherigen Abschnitts auf beliebige Hilberträume verallgemeinert werden können; daß solche Verallgemeinerungen möglich sind, zeigten SALGADO UND SÁNCHEZ-GÓMEZ [155].

Hierzu kann man nach Unterabschnitt B.4.7 die Hilberträume  $\mathcal{H}_A = \ell^2(I)$  und  $\mathcal{H}_B = \ell^2(J)$  für Indexmengen  $I$  und  $J$  betrachten, auf denen man nicht die Menge aller Operatoren, sondern nur die Algebra der HILBERT-SCHMIDT-Operatoren betrachtet; dies ist die Algebra derjenigen Operatoren, deren HILBERT-SCHMIDT-Norm endlich ist. Ein Vektor  $v \in \mathcal{H}_A \otimes \mathcal{H}_B \cong \ell^2(I \times J)$  besitzt die Form  $v = \sum_{ij} v_{ij} e_i \otimes f_j$ , wobei  $I_0 := \{i \in I \mid (\exists j \in J)(v_{ij} \neq 0)\}$  und  $J_0 := \{j \in J \mid (\exists i \in I)(v_{ij} \neq 0)\}$  höchstens abzählbar sind; man kann den Vektor  $v$  somit als Element im separablen Hilbertraum  $\ell^2(I_0 \times J_0)$  auffassen, so daß man die SCHMIDT-Zerlegung anwenden kann.

Die genaue (mathematisch präzise) Ausarbeitung dieser Feststellungen wäre jedoch aufwendig und lieferte keine grundlegend neuen Erkenntnisse, weshalb hier darauf verzichtet wurde.

## Kapitel 9

# Entropische Unschärferelationen

In diesem Kapitel werden entropische Unschärferelationen angesprochen und eine neue Unschärferelation hergeleitet. Des weiteren werden einige Vermutungen hinsichtlich solcher Relationen geäußert.

### 9.1 Einführung

Aus den Vorlesungen und Lehrbüchern über die Quantenmechanik ist die HEISENBERG'sche *Unschärferelation* [72] für Ort und Impuls eines einzelnen Teilchens bekannt, die für andere Observable von ROBERTSON [152] und von SCHRÖDINGER [160] verallgemeinert wurde. Es ist bekanntlich

$$\Delta A \cdot \Delta B \geq |\langle \Psi | [A, B] | \Psi \rangle| / 2, \quad (9.1)$$

wobei  $\Delta A$  und  $\Delta B$  die Standardabweichungen bei Messungen der Observablen  $A$  und  $B$  sind und  $[A, B] := A \cdot B - B \cdot A$  ihren Kommutator bezeichnet. Für  $A = x$  und  $B = p$  erhält man hieraus wieder  $\Delta x \Delta p \geq \hbar/2$ .

Ein gelegentlich geäußerter Kritikpunkt an dieser Ungleichung ist, daß sie vom Zustandsvektor  $|\Psi\rangle$  des Systems abhängt. Ferner hängt die Standardabweichung von den Eigenwerten und ihrer Verteilung und nicht nur von der Spektralzerlegung (dem Spektralmaß) der Operatoren ab.

Im Gegensatz zur obenstehenden Relation drücken entropische Unschärferelationen die Unsicherheit bei quantenmechanischen Messungen nicht durch Standardabweichungen sondern durch Entropien aus. Die wohl bekannteste dieser Unschärferelationen stammt von MAASSEN UND UFFINK [115], deren Beweis nach einigen Vorbemerkungen angegeben werden soll.

Für eine Wahrscheinlichkeitsverteilung  $p = (p_1, p_2 \dots, p_n) \in \mathcal{W}_n$  und einen Parameter  $r \in (-1; \infty) \setminus \{0\}$  bezeichne<sup>1</sup>

$$M_r(p) := 2^{-R_{r+1}(p)} = \left( \sum_{i=1}^n p_i^{1+r} \right)^{1/r}, \quad (9.2)$$

und für  $r \in \{-1, 0, \infty\}$  betrachte man die stetige Fortsetzung. Für den Beweis der Unschärferelation wird das folgende Lemma benötigt.

**Lemma 9.1 (Satz von Riesz)**

Für eine unitäre Matrix  $T = (t_{ij})_{i,j=1}^n \in \mathbb{C}^{n \times n}$ , eine reelle Zahl  $p \in [1; 2]$  und einen Spaltenvektor  $x = (x_k)_{k=1}^n \in \mathbb{C}^n$  gilt

$$c^{1/q} \cdot \|Tx\|_q \leq c^{1/p} \cdot \|x\|_p,$$

wenn  $c := \max_{i,j=1}^n |t_{ij}|$  und  $q := p \cdot (p-1)^{-1} \in [2; \infty]$  gesetzt werden.

Dieses Lemma wurde ursprünglich von RIESZ [151], Nr. 11 auf S. 473, gezeigt. Für  $p = q = 2$  ist der Satz (mit Gleichheit beider Seiten) trivial, fehlten die Ausdrücke in  $c$ , so folgte er aus der Unitarität von  $T$  und der Tatsache, daß  $\|T\|_p$  monoton in  $p \in [1; \infty]$  fällt (vgl. Lemma B.33).

*Beweis:* Die Unitarität von  $T$  impliziert  $\|T\|_{2 \rightarrow 2} = 1$ , und nach Lemma B.35 gilt  $\|T\|_{1 \rightarrow \infty} = c$ . Aus dem Satz von RIESZ und THORIN (Satz B.36) mit den Werten  $(p_0, q_0) := (1, \infty)$ ,  $(p_1, q_1) := (2, 2)$  und  $\vartheta := 2 \cdot q^{-1}$  folgt nun

$$\|T\|_{p \rightarrow q} \leq \|T\|_{1 \rightarrow \infty}^{1-\vartheta} \cdot \|T\|_{2 \rightarrow 2}^{\vartheta} = c^{1-\vartheta} \cdot 1^{\vartheta}, \quad (9.3)$$

und man berechnet  $1 - \vartheta = (p^{-1} + q^{-1}) - 2 \cdot q^{-1} = p^{-1} - q^{-1}$ , q. e. d.

Besitzen zwei selbstadjungierte Operatoren bzw. Observablen einen gemeinsamen Eigenvektor  $|\Psi\rangle$ , so liefert die Messung jeder dieser Observablen mit Sicherheit den zugehörigen Eigenwert; die Entropie der Messungen ist also Null. Gibt es aber keinen gemeinsamen Eigenvektor, so ist mindestens eine der Entropien positiv. Zwei nicht-entartete Observablen definieren zwei Orthonormalbasen, bezüglich derer gemessen werden kann. Abhängig vom größten Überlapp zweier Basisvektoren kann nun eine untere Schranke an die Summe der zwei Entropien angegeben werden. Dies ist die Aussage des folgenden Satzes von MAASSEN UND UFFINK [115]; ein Teil dieses Satzes wurde von KRISHNA UND PARTHASARATHY [104] für eventuell entartete Operatoren verallgemeinert.

<sup>1</sup>Der Ausdruck  $R_\alpha$  bezeichnet die für  $\alpha \in [0; \infty]$  definierte RÉNYI- $\alpha$ -Entropie.

**Satz 9.2 (Entropische Unschärferelationen)**

Es seien  $(|a_i\rangle)_{i=1}^n$  und  $(|b_j\rangle)_{j=1}^n$  zwei Orthonormalbasen des Hilbertraums  $\mathbb{C}^n$ , deren größter Überlapp mit  $c := \max \{|\langle a_i|b_j\rangle| \mid i, j \in \{1, \dots, n\}\}$  bezeichnet werde. Ein normierter Vektor  $|\Psi\rangle \in \mathbb{C}^n$  definiert mittels  $p_i := |\langle a_i|\Psi\rangle|^2$  und  $q_j := |\langle b_j|\Psi\rangle|^2$  zwei Wahrscheinlichkeitsverteilungen  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  und  $q = (q_1, \dots, q_n) \in \mathcal{W}_n$ , die die Ungleichung

$$M_r(p) \cdot M_s(q) \leq c^2$$

erfüllen, sofern man  $s \in [0; \infty]$  und  $r := -s/(2s + 1) \in [-1; 0]$  wählt.

*Beweis:* Setzt man in Lemma 9.1 die Werte  $x_j := \langle a_j|\Psi\rangle$  und  $t_{ij} := \langle b_j|a_i\rangle$ , so ist  $T$  unitär, und es gilt  $(Tx)_i = \sum_{j=1}^n t_{ij}x_j = \langle b_i|\Psi\rangle$ . Mit  $a = 2(1 + r)$  berechnet man nun

$$\begin{aligned} \|x\|_a &= \left(\sum_{i=1}^n |x_i|^a\right)^{1/a} = \left(\sum_{i=1}^n |\langle a_i|\Psi\rangle|^{2(1+r)}\right)^{\frac{1}{2(1+r)}} \\ &= \left(\sum_{i=1}^n p_i^{1+r}\right)^{\frac{1}{2(1+r)}} = M_r(p)^{\frac{r}{2(1+r)}} \end{aligned} \tag{9.4}$$

und für  $a' = 2(1 + s)$  entsprechend  $\|Tx\|_{a'} = M_s(q)^{\frac{s}{2(1+s)}}$ . Beachtet man, daß die Bedingungen an  $r$  und  $s$  mit der Forderung  $a^{-1} + a'^{-1} = 1$  übereinstimmen, so folgt mit dem genannten Lemma also

$$c^{\frac{1}{2(1+s)}} \cdot M_s(q)^{\frac{s}{2(1+s)}} \leq c^{\frac{1}{2(1+r)}} \cdot M_r(p)^{\frac{r}{2(1+r)}}. \tag{9.5}$$

Unter Verwendung von  $r/(1 + r) = -s/(1 + s)$  läßt sich dies als

$$M_r(p)^{\frac{s}{2(1+s)}} \cdot M_s(q)^{\frac{s}{2(1+s)}} \leq c^{1 - \frac{1}{2(1+s)}} \cdot c^{-\frac{1}{2(1+s)}} = c^{1 - \frac{1}{1+s}} = c^{\frac{s}{1+s}} \tag{9.6}$$

schreiben, und das Potenzieren mit  $2(1 + s)/s$  liefert die Behauptung, q. e. d.

Wegen  $M_0(p) = 2^{-H(p)}$  ergibt sich im Spezialfall  $r = s = 0$  die Ungleichung  $2^{-H(p)} \cdot 2^{-H(q)} \leq c^2$  und nach Logarithmieren weiter  $H(p) + H(q) \geq -2 \log_2 c$ .

## 9.2 Relationen für gemischte Zustände

Es ist nach der Konkavität der SHANNON-Entropie offensichtlich, daß die obengenannte Unschärferelation  $H(p) + H(q) \geq -2 \log_2 c$  auch für gemischte Zustände gilt. Man kann sich aber die Frage stellen, ob es für gemischte Zustände eine *bessere* Relation gibt. Die folgende Vermutung benennt eine naheliegende Verallgemeinerung dieser Ungleichung.

**Vermutung 9.3 (Entropische Unschärferelation)**

Es seien  $\mathcal{H} \cong \mathbb{C}^d$  ein endlichdimensionaler Hilbertraum und  $\rho \in \mathcal{S}(\mathcal{H})$  ein Zustand. Sind  $B_1 = (|a_i\rangle)_{i=0}^{d-1}$  und  $B_2 = (|b_j\rangle)_{j=0}^{d-1}$  zwei Orthonormalbasen von  $\mathcal{H}$  und  $p$  und  $q$  die Wahrscheinlichkeitsverteilungen, die beim Messen von  $\rho$  bzgl.  $B_1$  bzw.  $B_2$  entstehen, so gilt die Ungleichung

$$H(p) + H(q) \geq -\log_2 \max_{ij} |\langle a_i | b_j \rangle|^2 + S(\rho).$$

Es handelt sich also um die entropische Unschärferelation für die SHANNON-Entropien, deren rechte Seite um die VON-NEUMANN-Entropie der Dichtematrix  $\rho$  ergänzt wurde. Sie ist ein Spezialfall einer noch allgemeineren Vermutung, die in Zusammenhang mit der Sicherheit quantenkryptographischer Protokolle steht; vgl. hierzu RENES UND BOILEAU [145].<sup>2</sup>

**Vermutung 9.4 (Vermutung über Holevo-Schranken)**

Alice, Bob und Eve teilen sich einen reinen Zustand  $|\Psi\rangle_{ABE}$ , und Alice kann bzgl. zwei Basen  $B_1$  und  $B_2$  messen. Dann gilt

$$\chi_B(B_1) + \chi_E(B_2) \leq H(p) + H(q) + \log_2 \max_{ij} |\langle a_i | b_j \rangle|^2.$$

Der Ausdruck  $\chi_B(B_1)$  bezeichnet die HOLEVO-Größe für Bob, wenn Alice bzgl.  $B_1$  mißt, und sinngemäß ist  $\chi_E(B_2)$  definiert.

Die entropische Unschärferelation folgt hieraus, indem man annimmt, daß Eves System eindimensional ist, d. h. in Wirklichkeit gar nicht betrachtet zu werden braucht.

## 9.3 Ideen zu einem Beweis

Es ist mir nicht gelungen, die Vermutung 9.3 allgemein analytisch zu zeigen, allerdings können – neben numerischen Untersuchungen – einige Spezialfälle behandelt werden.

**Satz 9.5 (Qubits und komplementäre Basen)**

Die Vermutung 9.3 gilt für Qubits, wenn die Basen komplementär sind.

*Beweis:* Im Falle eines Qubit-Systems können die beiden Basen durch je einen normierten BLOCH-Vektor bestimmt werden; der Zustand selbst kann ebenfalls durch einen BLOCH-Vektor beschrieben werden, der aber nicht unbedingt normiert sein muß (vgl. Unterabschnitt 2.3.2). Es bietet sich an, die Vektoren in Kugelkoordinaten anzugeben, für die hier die Konvention  $\vec{r} = (x, y, z)^t = (r \sin \vartheta \cos \varphi, r \sin \vartheta \sin \varphi, r \cos \vartheta)^t \in \mathbb{R}^3$  mit Zahlenwerten  $(r, \vartheta, \varphi) \in \mathbb{R}_0^+ \times [0; \pi] \times [0; 2\pi)$  verwendet werde.

---

<sup>2</sup>Die Vermutung wurde mir Ende Oktober 2007 durch JOE RENES persönlich mitgeteilt.

Da die relevanten Eigenschaften (speziell Skalarprodukte) unter räumlichen Drehungen auf der BLOCH-Kugel invariant sind, kann angenommen werden, daß die erste Basis den BLOCH-Vektor  $\vec{e}_z = (0, 0, 1) \in \mathbb{R}^3$  besitzt, es sich also um die  $z$ -Basis handelt. Der BLOCH-Vektor der zweiten Basis kann so gedreht werden, daß er in Kugelkoordinaten den Azimutalwinkel  $\varphi_0 = 0$  besitzt; sein Polarwinkel  $\vartheta_0$  kann aus dem Intervall  $[0; \pi/2]$  gewählt werden, denn andernfalls könnte man den BLOCH-Vektor durch sein Negatives ersetzen, was eine Umbenennung der Basiselemente bewirkt. Für den maximalen quadrierten Überlapp zweier Basisvektoren folgt also  $c^2 = \cos^2(\vartheta_0/2) = (1 + \cos \vartheta_0)/2$ .

Ein beliebiger Qubit-Zustand ist nun durch seinen BLOCH-Vektor festgelegt, der für einen Einheitsvektor  $\vec{e}$  in der Form  $\vec{s} = r\vec{e}$  geschrieben werden kann. Die Länge  $r \in [0; 1]$  legt die VON-NEUMANN-Entropie des Zustands und somit die rechte Seite der zu beweisenden Relation fest; es genügt also, die linke Seite über alle Einheitsvektoren  $\vec{e}$  zu minimieren. Hierzu stellt man fest, daß die Entropie einer Messung  $H[(1+r|\cos \alpha|)/2]$  beträgt, wenn bzgl. einer durch den Einheitsvektor  $\vec{n}$  definierten Basis gemessen wird und  $\alpha := \angle(\vec{e}, \vec{n})$  den eingeschlossenen Winkel zwischen  $\vec{e}$  und  $\vec{n}$  bezeichnet.

Geometrisch erkennt man aus diesen Überlegungen, daß das gesuchte Minimum für einen BLOCH-Vektor angenommen wird, der komplanar zu den die Basen definierenden Vektoren, also in der  $x$ - $z$ -Ebene liegt. Es kann daher für den BLOCH-Vektor des Zustands  $\varphi = 0$  und  $\vartheta \in [0; \vartheta_0]$  angenommen werden; zu zeigen ist für alle  $r \in [0; 1]$ ,  $\vartheta_0 \in [0; \pi/2]$  und  $\vartheta \in [0; \vartheta_0]$ , daß die Ungleichung

$$H\left(\frac{1+r\cos\vartheta}{2}\right) + H\left(\frac{1+r\cos(\vartheta_0-\vartheta)}{2}\right) \geq U + H\left(\frac{1+r}{2}\right) \quad (9.7)$$

mit  $U := -\log_2[(1 + \cos \vartheta_0)/2]$  erfüllt ist. Hierzu bietet es sich an, die binäre SHANNON-Entropie an der Stelle  $1/2$  zu entwickeln, was auf

$$H(1/2 \pm y) = 1 - (\ln 2)^{-1} \sum_{l \in 2\mathbb{N}} \frac{(2y)^l}{l(l-1)} \quad (9.8)$$

führt; man beachte, daß nur über gerade Zahlen  $l$  zu summieren ist. Aus der Ungleichung (9.7) wird hiermit

$$\sum_{l \in 2\mathbb{N}} \frac{r^l}{l(l-1)} [\cos^l \vartheta + \cos^l(\vartheta_0 - \vartheta) - 1] \leq \ln 2 \cdot \left( \log_2 \frac{1 + \cos \vartheta_0}{2} + 1 \right).$$

Im Falle komplementärer Basen ist notwendigerweise  $\vartheta_0 = \pi/2$ , so daß wegen  $\cos(\pi/2 - \vartheta) = \sin \vartheta$  nur noch

$$\sum_{l \in 2\mathbb{N}} \frac{r^l}{l(l-1)} [\cos^l \vartheta + \sin^l \vartheta - 1] \leq \ln 2 \cdot \left( \log_2 \frac{1+0}{2} + 1 \right) = 0 \quad (9.9)$$

für alle  $r \in [0; 1]$  und  $\vartheta \in [0; \pi/2]$  zu zeigen ist. Hierzu genügt es, wenn jeder Summand auf der linken Seite nicht-positiv ist, wenn also  $\cos^l \vartheta + \sin^l \vartheta \leq 1$  für jede gerade Zahl  $l \in 2\mathbb{N}$  gilt. Für  $l = 2$  liegt bekanntlich Gleichheit vor, für größere  $l$  ist die Ungleichung erfüllt, weil die Beträge der Cosinus- und Sinusfunktion nicht größer als Eins werden können, q. e. d.

Eine Verallgemeinerung des Beweises auf allgemeine Qubit-Basen erscheint möglich, insbesondere, weil sämtliche numerischen Ergebnisse für die Richtigkeit der Aussage sprechen. In höherdimensionalen Systemen konnte noch keine Lösung gefunden werden, es erscheint mir aber naheliegend, daß der Zustand, der die linke Seite der Ungleichung minimiert, auf einer Bahn liegen muß, die die eine Basis in die andere überführt.

In höherdimensionalen Fällen konnten nur einige sehr einfache Spezialfälle analytisch bewiesen werden, die das folgende Lemma zusammenfaßt.

**Lemma 9.6 (Einige Spezialfälle der Relation)**

*Die Vermutung 9.3 gilt, wenn (1) eine der Basen eine Eigenbasis der Dichtematrix ist, (2) eine der Basen komplementär zu einer Eigenbasis der Dichtematrix ist oder (3) die Dichtematrix maximal gemischt ist.*

*Beweis:* Es sei  $U := -\log_2 \max_{i,j} |\langle a_i | b_j \rangle|^2$  der logarithmierte Überlapp der Vektoren, die die zwei Basen  $B_1$  und  $B_2$  definieren.

Ist nun zum Beispiel  $B_1$  eine Eigenbasis der Dichtematrix, so ist wegen  $H(p) = S(\rho)$  nur noch  $H(q) \geq U$  zu zeigen. Ist  $\rho = \sum_i p_i |a_i\rangle\langle a_i|$ , so gilt  $q_j = \sum_i p_i |\langle a_i | b_j \rangle|^2$ , und man berechnet

$$\begin{aligned} H(q) &= \sum_{j=0}^{d-1} \left( \sum_{i=0}^{d-1} p_i |\langle a_i | b_j \rangle|^2 \right) \left[ -\log_2 \left( \sum_{k=0}^{d-1} p_k |\langle a_k | b_j \rangle|^2 \right) \right] \\ &\stackrel{?}{\geq} U \cdot \sum_{j=0}^{d-1} \sum_{i=0}^{d-1} p_i |\langle a_i | b_j \rangle|^2 = U \cdot \sum_{i=0}^{d-1} p_i = U. \end{aligned} \quad (9.10)$$

Die fragliche Ungleichung gilt wegen  $|\langle a_j | b_k \rangle|^2 \leq \max_{j,k} |\langle a_k | b_j \rangle|^2$ , woraus  $\sum_k p_k |\langle a_k | b_j \rangle|^2 \leq \sum_k p_k \max_{j,k} |\langle a_k | b_j \rangle|^2 = \max_{j,k} |\langle a_k | b_j \rangle|^2$  folgt; schließlich fällt der negative Logarithmus monoton, was die erste Aussage zeigt.

Ist im zweiten Fall zum Beispiel  $B_2$  komplementär zur Eigenbasis von  $\rho$ , so ist  $H(q) = 1$ ; somit gilt wegen  $H(p) \geq S(\rho)$  und  $1 \geq U$  die Ungleichung. Im dritten Fall ist die Dichtematrix maximal gemischt, das heißt, es gilt  $H(p) = H(q) = S(\rho) = 1$  und  $U \in [0; 1]$ , q. e. d.

Eine weitere Möglichkeit, die Vermutung 9.3 zu beweisen, erhält man, indem man eine Purifizierung der Dichtematrix  $\rho_A = \sum_i \lambda_i^2 |i\rangle\langle i|$  betrachtet und auf beiden Untersystemen lokale Messungen ausführt. Eine Purifizierung des Zustands lautet  $|\Psi\rangle_{AB} = \sum_i \lambda_i |i\rangle_A |i\rangle_B$ . Für zwei Basen  $B'_1$  und  $B'_2$  auf dem

Gesamtsystem erscheint im Satz 9.2 der Ausdruck

$$c_{\text{ges}} := \max \{ |\langle v|w \rangle|^2 \mid v \in B'_1, w \in B'_2 \}. \quad (9.11)$$

Als Beispiel sei  $B'_1 := B_1 \otimes B_1 := \{|a_i\rangle|a_k\rangle \mid i, k \in \{1, \dots, n\}\}$  das Tensorprodukt der Basis  $B_1$  mit sich selbst; analog sei  $B'_2 := B_2 \otimes B_2$ . Man erhält  $c_{\text{ges}} = c^2$ , wenn  $c$  den maximalen Überlapp von  $B_1$  und  $B_2$  bezeichnet. Für die Wahrscheinlichkeit der Messung des Zustands  $|a_i, a_k\rangle := |a_i\rangle|a_k\rangle$  erhält man

$$p_{ik} = |\langle a_i, a_k | \Psi \rangle|^2 = \left( \sum_{\nu} \lambda_{\nu} \langle a_i | \nu \rangle \langle a_i | \nu \rangle \right)^2, \quad (9.12)$$

und für die Messung von  $|b_j, b_l\rangle := |b_j\rangle|b_l\rangle$  gilt entsprechend

$$q_{jl} = |\langle b_j, b_l | \Psi \rangle|^2 = \left( \sum_{\nu} \lambda_{\nu} \langle b_j | \nu \rangle \langle b_l | \nu \rangle \right)^2. \quad (9.13)$$

Man erhält somit  $H[(p_{ik})_{ik}] + H[(q_{jl})_{jl}] \geq -4 \log_2 c$ . Faßt man die Messung als die Realisierung einer aus zwei Zufallsvariablen  $P_i$  und  $P_k$  mit der gleichen Verteilung  $P$  zusammengesetzten Zufallsvariablen  $P_{ik}$  auf, so erhält man  $H(P_{ik}) = H(P_i) + H(P_k) - H(P_i : P_k) = 2H(P) - H(P_i : P_k)$ . Das gleiche gilt sinngemäß für die andere Messung, und dies führt auf die Ungleichung

$$H(P) + H(Q) \geq -\log_2 c^2 + \frac{H(P_i : P_k) + H(Q_j : Q_l)}{2}. \quad (9.14)$$

Eine andere naheliegende Wahl für die zwei Basen des zusammengesetzten Systems ist  $B''_1 := B_1 \otimes B_2$  und  $B''_2 := B_2 \otimes B_1$ . Auch in diesem Fall ist  $c_{\text{ges}} = c^2$ , und ähnlich wie im ersten Fall folgt

$$w_{il} = \left( \sum_{\nu} \lambda_{\nu} \langle a_i | \nu \rangle \langle b_l | \nu \rangle \right)^2 \quad (9.15)$$

für die Messung von  $|a_i\rangle|b_l\rangle$ ; dies führt schließlich zu

$$H(P) + H(Q) \geq -\log_2 c^2 + H(P : Q). \quad (9.16)$$

Beide Rechnungen zeigen, daß für gemischte Zustände bessere Relationen als  $H(p) + H(q) \geq -2 \log_2 c$  existieren müssen. Auf der anderen Seite zeigen sie mit der HOLEVO-Schranke an die gemeinsame Information aber auch, daß keine bessere Relation als die in der Vermutung 9.3 aufgestellte möglich ist, wenn man sie als Korollar der Ursprungsrelation erhalten will.



# Anhang



# Anhang A

## Ergänzungen zum Haupttext

Dieses Kapitel beinhaltet einen längeren Beweis, der aus dem Hauptteil ausgelagert wurde, um den Lesefluß nicht übermäßig zu stören, und eine Eigenschaft der Fouriertransformation. Des weiteren findet sich hier ein Verzeichnis der verwendeten Notationen.

### A.1 Die Entwicklung unter $B_n^{(d)}$ -Schritten

In diesem Abschnitt wird der Satz 4.6 von Seite 66 bewiesen, was durch vollständige Induktion unter Verwendung der Ideen von MARTÍN-DELGADO UND NAVASCUÉS [118], Anhang 1 auf S. 179, erfolgt. Der Fall  $n = 1$  ist trivial und der Fall  $n = 2$  dient als Induktionsbasis. Hierzu seien  $(A_{lm})_{lm}, (B_{st})_{st} \in \mathcal{S}_{\text{bd}}^{(d)}$  zwei bell-diagonale Quantenzustände; der Anfangszustand der zwei Qudit-Paare lautet somit<sup>1</sup>

$$\rho = \sum_{lm} A_{lm}(l, m) \otimes \sum_{st} B_{st}(s, t) = \sum_{lmst} A_{lm} B_{st}(l, m) \otimes (s, t), \quad (\text{A.1})$$

und die Anwendung einer GBXOR-Operation vom ersten zum zweiten Paar macht daraus

$$\rho = \sum_{lmst} A_{lm} B_{st}(l \oplus s, m) \otimes (s, m \ominus t). \quad (\text{A.2})$$

Fordert man, daß die Messung der Ditparität auf dem zweiten Paar Übereinstimmung ergibt, daß also  $m \ominus t = 0$  gilt, und verwirft alle anderen Fälle, so schreibt sich der Gesamtzustand als

$$\rho = N_2^{-1} \sum_{lms} A_{lm} B_{sm}(l \oplus s, m) \otimes (s, 0), \quad (\text{A.3})$$

---

<sup>1</sup>Man beachte hierbei die Notation  $(l, m) := |\Psi_{lm}\rangle\langle\Psi_{lm}|$ ; sämtliche Vektor-, Matrix- und Summationsindizes laufen über  $\mathbb{Z}_d = \{0, \dots, d-1\}$  und werden ggf. weggelassen.

wobei  $N_2 := \sum_m A_{*m} B_{*m} = \sum_m [(\sum_l A_{lm})(\sum_l B_{lm})]$  für die Normierungskonstante steht. Die Bildung der Spur über das zweite, nach den lokalen Messungen nunmehr unbrauchbare Qudit-Paar liefert weiter

$$\begin{aligned} \rho &= N_2^{-1} \sum_{lms} A_{lm} B_{sm}(l \oplus s, m) = N_2^{-1} \sum_{lm'l'} A_{lm} B_{l' \ominus l, m}(l', m) \\ &= \sum_{lm} \left[ N_2^{-1} \sum_{l'} A_{l'm} B_{l' \ominus l, m} \right] (l, m), \end{aligned} \quad (\text{A.4})$$

was die Induktionsbasis zeigt. Man erkennt, daß  $N_2$  gleichzeitig die Wahrscheinlichkeit für die Übereinstimmung beider Ditzwerte ist; somit ist  $N_2$  auch die Wahrscheinlichkeit dafür, daß das erste Paar weiterverwendet wird.

Als Induktionsvoraussetzung werde nun angenommen, daß der Satz für alle natürlichen Zahlen bis zu einem festen Wert  $n \in \mathbb{N}$  gezeigt worden sei. Im nun folgenden Induktionsschritt ist er unter dieser Annahme für  $n+1$  zu zeigen. Hierzu betrachte man bell-diagonale Zustände  $\rho^{(k)} = (A_{lm}^{(k)})_{lm} \in \mathcal{S}_{\text{bd}}^{(d)}$  für  $k \in \{1, \dots, n+1\}$ . Die Anwendung eines  $B_n^{(d)}$ -Schrittes auf die ersten  $n$  Zustände liefere  $\rho' = (A'_{lm}) \in \mathcal{S}_{\text{bd}}^{(d)}$ , die Anwendung eines  $B_{n+1}^{(d)}$ -Schrittes auf alle  $n+1$  Zustände  $\rho'' = (A''_{lm}) \in \mathcal{S}_{\text{bd}}^{(d)}$ ; die zugehörigen Normierungskonstanten seien  $N_n$  und  $N_{n+1}$ .

Der  $B_{n+1}^{(d)}$ -Schritt kann ausgeführt werden, indem zuerst ein  $B_n^{(d)}$ -Schritt auf die ersten  $n$  Zustände angewendet wird und anschließend ein  $B_2^{(d)}$ -Schritt des verbleibenden ersten Paares mit dem  $(n+1)$ -ten Zustand ausgeführt wird. Wurde das erste Paar schon nach dem  $B_n^{(d)}$ -Schritt verworfen, so wird das  $(n+1)$ -te Paar ebenfalls verworfen. Diese Überlegung liefert mithilfe der Induktionsvoraussetzung für  $n=2$  der Ausdruck

$$A''_{lm} = \frac{1}{dN_2} \sum_i z^{-il} \left[ \left( \sum_j z^{ij} A'_{jm} \right) \left( \sum_{j'} z^{ij'} A_{j'm}^{(n+1)} \right) \right]. \quad (\text{A.5})$$

Mithilfe der Induktionsvoraussetzung für  $n$  schreibt sich der Ausdruck in der ersten runden Klammer als

$$\begin{aligned} \sum_j z^{ij} A'_{jm} &= \sum_j z^{ij} \left[ \frac{1}{dN_n} \sum_{i'} z^{-i'j} \prod_{k=1}^n \left( \sum_{j'} z^{i'j'} A_{j'm}^{(k)} \right) \right] \\ &= \frac{1}{dN_n} \sum_{i'} z^{(i-i')j} \left[ \prod_{k=1}^n \left( \sum_{j'} z^{i'j'} A_{j'm}^{(k)} \right) \right] \\ &= \frac{1}{N_n} \prod_{k=1}^n \left( \sum_{j'} z^{ij'} A_{j'm}^{(k)} \right), \end{aligned} \quad (\text{A.6})$$

wozu  $\sum_j z^{(i-i')j} = d\delta_{ii'}$  verwendet wurde. Es fällt nun auf, daß man den Term für das  $(n+1)$ -te Paar in das Produkt ziehen kann, so daß sich

$$A''_{lm} = \frac{1}{dN_2 N_n} \sum_i z^{-il} \prod_{k=1}^{n+1} \left( \sum_j z^{ij} A_{jm}^{(k)} \right) \quad (\text{A.7})$$

ergibt, wenn der Summationsindex  $j'$  durch  $j$  ersetzt wird. Es verbleibt nur noch zu zeigen, daß  $N_2 \cdot N_n = N_{n+1}$  ist. Unter Verwendung der bekannten Abkürzungen  $A_{*m}^{(k)} = \sum_l A_{lm}^{(k)}$  und  $A'_{*m} = \sum_l A'_{lm}$  berechnet man nach der Induktionsvoraussetzung  $N_n = \sum_m \left( \prod_{k=1}^n A_{*m}^{(k)} \right)$  und hieraus weiter

$$\begin{aligned} N_2 \cdot N_n &= \left( \sum_m A'_{*m} \cdot A_{*m}^{(n+1)} \right) \cdot N_n = \sum_m \underbrace{N_n^{-1} \prod_{k=1}^n A_{*m}^{(k)}}_{=A'_{*m}} \cdot A_{*m}^{(n+1)} \cdot N_n \\ &= \sum_m \prod_{k=1}^{n+1} A_{*m}^{(k)} = N_{n+1}. \end{aligned} \quad (\text{A.8})$$

Damit ist die Behauptung für  $n+1$  und mithin für alle  $n \in \mathbb{N}$  gezeigt, q. e. d.

## A.2 Eigenschaften der Fouriertransformation

Die Fouriertransformation  $\mathcal{F}$  verallgemeinert die aus dem BB84-Protokoll bekannte HADAMARD-Transformation; es ist daher sinnvoll, die Wirkung der Zwei-Qudit-Transformation  $\mathcal{F} \otimes \mathcal{F}^*$  auf einen reinen verallgemeinerten BELL-Zustand zu berechnen. Mit der Konvention, daß alle Summen über  $\mathbb{Z}_d = \{0, \dots, d-1\}$  laufen, notiert man

$$\mathcal{F} \otimes \mathcal{F}^* = d^{-1/2} \sum_{ij} z^{ij} |i\rangle \langle j| \otimes d^{-1/2} \sum_{pq} z^{-pq} |p\rangle \langle q|, \quad (\text{A.9})$$

und mit  $|\Psi_{lm}\rangle = d^{-1/2} \sum_k z^{lk} |k\rangle |k \ominus m\rangle$  ergibt sich

$$\begin{aligned} (\mathcal{F} \otimes \mathcal{F}^*) |\Psi_{lm}\rangle &= d^{-3/2} \sum_{ijpqk} z^{ij-pq+lk} |i\rangle \langle j|k\rangle |p\rangle \langle q|k \ominus m\rangle \\ &= d^{-3/2} \sum_{ipk} z^{ik-p(k-m)+lk} |i\rangle |p\rangle. \end{aligned} \quad (\text{A.10})$$

Verwendet man  $\sum_k z^{(i-p+l)k} = d\delta_{p,i \oplus l}$ , so folgt schließlich

$$\begin{aligned} (\mathcal{F} \otimes \mathcal{F}^*) |\Psi_{lm}\rangle &= d^{-1/2} \sum_{ip} \delta_{p,i \oplus l} z^{pm} |i\rangle |p\rangle = d^{-1/2} \sum_i z^{(i+l)m} |i\rangle |i \oplus l\rangle \\ &= z^{lm} d^{-1/2} \sum_i z^{im} |i\rangle |i \ominus (d \ominus l)\rangle = z^{lm} |\Psi_{m,d-l}\rangle. \end{aligned} \quad (\text{A.11})$$

## A.3 Verzeichnis der verwendeten Notationen

Im gesamten Text wird die in der Mathematik sowie in den Natur- und Ingenieurwissenschaften gebräuchliche Notation verwendet. Es werden im folgenden einige bekannte und weniger bekannte Schreibweisen aufgelistet.

### A.3.1 Allgemeine Symbole und Mengensymbole

Die folgende Liste enthält die wichtigsten im Text verwendeten Mengensymbole; weitere Mengensymbole werden sinngemäß gebildet.

$\mathbb{N}$	Menge der natürlichen Zahlen $\{1, 2, 3, 4, \dots\}$
$\mathbb{N}_0$	Menge der natürlichen Zahlen und Null $\mathbb{N} \cup \{0\}$
$\mathbb{Z}$	Integritätsbereich der ganzen Zahlen $\{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathbb{Q}$	Körper der rationalen Zahlen
$\mathbb{R}$	Körper der reellen Zahlen $(-\infty; \infty)$
$\mathbb{R}^+$	– positive reelle Zahlen $\{x \in \mathbb{R} \mid x > 0\}$
$\mathbb{R}_0^+$	– nicht-negative reelle Zahlen $\{x \in \mathbb{R} \mid x \geq 0\}$
$[a; b]$	– abgeschlossenes Intervall $\{x \in \mathbb{R} \mid a \leq x \leq b\}$
$(a; b)$	– offenes Intervall $\{x \in \mathbb{R} \mid a < x < b\}$
$\mathbb{C}$	Körper der komplexen Zahlen $\{a + bi \mid a, b \in \mathbb{R}\}$
$K^{n \times m}$	Menge der $n \times m$ -Matrizen über $K$ ( $n$ Zeilen, $m$ Spalten)
$M_n(K)$	Algebra der $n \times n$ -Matrizen über $K$

Neben den Mengensymbolen treten gelegentlich die in der folgenden Liste aufgeführten Bezeichnungen auf.

$\Leftrightarrow$	Äquivalenz („genau dann, wenn“)
$\Rightarrow$	Implikation („daraus folgt“)
$\forall$	Allquantor („für alle ...“, „für jedes ...“)
$\exists$	Existenzquantor („es existiert [mindestens] ein ...“)
$\exists_1, \exists!$	Existenzquantor („es existiert genau ein ...“)
$\wedge; \vee$	Konjunktion (logisches Und) bzw. Disjunktion (logisches Oder)
$:=$	definitionsgemäß gleich (Doppelpunkt auf der definierten Seite)
$\cong$	isomorph (strukturgleich)
q. e. d.	Ende eines Beweises: „quod erat demonstrandum“ („was zu beweisen war“, w. z. b. w.)
$ \cdot $	Betrag einer Zahl oder Mächtigkeit (Kardinalität) einer Menge
kgV	kleinstes gemeinsames Vielfaches
ggT	größter gemeinsamer Teiler
diag; $\mathbb{I}$	diagonale oder blockdiagonale Matrix; Einheitsmatrix
span	Lineare Hülle (Erzeugnis, Aufspann) einer Menge von Vektoren
$\log_b; \ln$	Logarithmus zur Basis $b$ ; natürlicher Logarithmus ( $b = e$ )
$[\cdot]$	GAUßsche Klammerfunktion <sup>2</sup> $[\cdot] : \mathbb{R} \rightarrow \mathbb{Z}$ , $[x] := \max \{n \in \mathbb{Z} \mid n \leq x\}$
sup	Supremum (kleinste obere Schranke, obere Grenze)
inf	Infimum (größte untere Schranke, untere Grenze)

Das KRONECKER-Symbol  $\delta_{ik}$  nimmt den Wert 1 an, falls  $i = k$  ist, sonst 0. Das LEVI-CIVITA-Symbol (der total antisymmetrischer Einheitstensor dritter Stufe)  $\varepsilon_{ijk}$  ist  $\pm 1$ , falls die Werte  $i, j, k$  zyklisch bzw. antizyklisch angeordnet sind, und verschwindet andernfalls.

### A.3.2 Weitere Symbole und Konventionen

Wenn nichts anderes gesagt wird, stehen die Variablen  $n$  und  $m$  für natürliche Zahlen; Laufvariablen sind üblicherweise  $i, j, k$  und  $l$ , und manchmal auch  $p, q, r, s$  usw. Die *imaginäre Einheit*  $i = \sqrt{-1}$  und die *EULERSche Zahl*  $e \approx 2,718\dots$  werden immer in aufrechter Schrift gesetzt. Die Buchstaben  $p$  und  $q$  stehen vielfach für Wahrscheinlichkeitsverteilungen. Die Konvention, über alle mehrfach auftretenden Indizes eines Ausdrucks zu summieren (EINSTEINSche Summenkonvention), wird nicht verwendet.

Im Zusammenhang mit quantenmechanischen Systemen, insbesondere in der Quantenkryptographie, werden sehr oft die folgenden Symbole verwendet:

$\mathcal{H}$	Hilbertraum
$\mathcal{S}(\mathcal{H})$	Dichtematrizen (Zustände) auf dem Hilbertraum $\mathcal{H}$
$\partial B_1(\mathcal{H})$	Einheitssphäre (Menge der normierten Vektoren) von $\mathcal{H}$
$d$	Dimension des Quantensystems: $\mathcal{H} \cong \mathbb{C}^d$
$\mathbb{Z}/d\mathbb{Z}$	Restklassenring modulo $d$
$\mathbb{Z}_d$	Vertretersystem $\{0, \dots, d-1\}$ des Restklassenrings $\mathbb{Z}/d\mathbb{Z}$
$\oplus, \ominus$	modulare Addition bzw. Subtraktion auf $\mathbb{Z}_d$
$\mathbb{Z}_d^*$	Kurzschreibweise für $\mathbb{Z}_d \setminus \{0\}$ (im allgemeinen nicht die Einheitengruppe des Rings $\mathbb{Z}/d\mathbb{Z}$ )
$\mathcal{S}_{\text{bd}}^{(d)}, \mathcal{S}_{\text{bd}}$	Menge der bell-diagonalen Zustände in $d$ bzw. 2 Dimensionen
$z, z_d$	Hauptwert der $d$ -ten Einheitswurzel $z_d = \exp(2\pi i/d)$
$B_n^{(d)}, P_n^{(d)}$	verschiedene Korrekturschritte
$\mathbb{F}_p, \mathbb{F}_q$	Endlicher Körper mit $p$ bzw. $q = p^n$ Elementen (vgl. Satz B.11)
$\mathcal{W}_n$	Wahrscheinlichkeitsverteilungen auf $n$ -elementigen Mengen
$S_n$	Symmetrische Gruppe (Permutationsgruppe) auf $n$ Elementen
$\stackrel{\text{aeg}}{=}$	asymptotisch-exponentiell gleich (vgl. Unterabschnitt 3.6.5)

Bell-diagonale Zustände werden durchgängig mit ihrer Koeffizientenmatrix  $(A_{lm})_{l,m=0}^{d-1}$  identifiziert, deren Randverteilungen der Phasen- und Ditwerte mit  $(A_{l*})_{l=0}^{d-1}$  bzw.  $(A_{*m})_{m=0}^{d-1}$  bezeichnet werden (vgl. Unterabschnitt 2.5.2).

---

<sup>2</sup>In der Informatik ist die Schreibweise  $\lfloor x \rfloor$  statt  $[x]$  üblich, und sinngemäß setzt man  $\lfloor x \rfloor := \min \{n \in \mathbb{Z} \mid x \leq n\}$ . Diese Schreibweisen werden hier *nicht* verwendet!



# Anhang B

## Mathematische Grundlagen

In diesem Kapitel sollen die notwendigen mathematischen Grundbegriffe für diese Dissertation zusammengestellt werden. Ziel dieses Kapitels ist es, einheitliche Begriffsbildungen zu treffen. Die aufgeführten Ergebnisse werden so allgemein gefaßt, daß sie in den entsprechenden mathematischen Zusammenhang eingeordnet werden können. Für die Beweise der Aussagen vergleiche man die Literatur.

### B.1 Wahrscheinlichkeitsrechnung

In diesem Abschnitt werden einige zum Verständnis der Dissertation nützliche Grundlagen der Wahrscheinlichkeitstheorie und der mathematischen Statistik abgehandelt. Die hierauf aufbauenden Begriffe der Informationstheorie (Entropien usw.) finden sich in Anhang C. Für Einzelheiten zur Wahrscheinlichkeitsrechnung verweise ich auf die Literatur, insbesondere auf die Bücher von KREYSZIG [103], LEHN UND WEGMANN [106], STANGE [170] und GNEDENKO [63].

#### B.1.1 Fakultät und Stirling-Reihe

Die *Fakultät* einer Zahl  $n \in \mathbb{N}_0$  ist durch die Festsetzungen  $0! := 1$  und  $(n+1)! := (n+1)n!$  rekursiv definiert.<sup>1</sup> Die Fakultät kann allgemein durch die STIRLING-Formel  $n! \approx \sqrt{2\pi n} n^n e^{-n}$  abgeschätzt werden. Eine quantitative Abschätzung erfolgt durch die sogenannte STIRLING-Reihe.

---

<sup>1</sup>Allgemeiner kann man die EULERSche *Gammafunktion*  $\Gamma$  betrachten, die eine meromorphe Funktion mit Polen in  $\mathbb{Z}_0^-$  ist und die die Funktionalgleichung  $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$  erfüllt; es gilt dann  $\Gamma(\alpha+1) = \alpha!$  für alle  $\alpha \in \mathbb{N}_0$ .

**Satz B.1 (Stirling-Reihe für die Abschätzung der Fakultät)**

Für  $n \in \mathbb{N}$  und  $m \in \mathbb{N}$  existiert ein  $\vartheta_m \in [0; 1]$  derart, daß

$$\frac{n!}{\sqrt{2\pi n} n^n e^{-n}} = \exp \left[ \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)n^{2k-1}} + \vartheta_m \cdot \frac{B_{2m+2}}{(2m+2)(2m+1)n^{2m+1}} \right]$$

gilt, wobei die BERNOULLI-Zahlen durch  $B_n := \lim_{x \rightarrow 0} (d^n/dx^n)[x/(e^x - 1)]$  definiert sind.<sup>2</sup>

Ein Beweis des Satzes verwendet die EULERSche *Summenformel* nebst der zugehörigen Restgliedabschätzung und findet sich ausführlich bei KNOPP [95] im XIV. Kapitel; Einzelheiten über die BERNOULLI-Zahlen findet man bei HEUSER [74], Teil I, Abschnitt 71 auf S. 410–413.

**B.1.2 Binomialkoeffizienten und Binomischer Lehrsatz**

Für  $n \in \mathbb{N}_0$  und  $k \in \{0, \dots, n\}$  ist der *Binomialkoeffizient* durch

$$\binom{n}{k} := \frac{n!}{(n-k)!k!} \quad (\text{B.1})$$

definiert. Er bestimmt die Anzahl der Möglichkeiten, aus einer  $n$ -elementigen Menge  $k \in \{0, \dots, n\}$  Elemente auszuwählen.<sup>3</sup> Für eine natürliche Zahl  $n \in \mathbb{N}_0$  und Zahlen  $a, b \in \mathbb{C}$  oder allgemeiner für kommutierende Elemente eines Rings mit Eins gilt der *Binomische Lehrsatz* [57]

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k. \quad (\text{B.2})$$

**B.1.3 Zufallsvariablen und ihre Eigenschaften**

*Zufallsvariablen* sind reellwertige meßbare Funktionen; in diesem Text seien sie stets *diskret*, d. h. ihre Bilder seien höchstens abzählbar. Für eine Zufallsvariable  $X$  definiert man *Erwartungswert* und *Varianz* mittels

$$\mathcal{E}(X) := \sum_{x \in \mathbb{R}} x \cdot P(X = x), \quad (\text{B.3})$$

$$\text{Var}(X) := \mathcal{E}((X - \mathcal{E}(X))^2) = \mathcal{E}(X^2) - \mathcal{E}(X)^2. \quad (\text{B.4})$$

<sup>2</sup>Alternativ können die BERNOULLI-Zahlen auch durch  $B_0 := 1$ ,  $B_1 := -1/2$  und  $\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0$  für  $k \geq 2$  rekursiv definiert werden; es gilt  $B_{2\lambda+1} = 0$  für alle  $\lambda \in \mathbb{N}$ . Es ist anzumerken, daß es sich bei STIRLING-Reihe (für  $m \rightarrow \infty$ ) nicht um eine konvergente sondern nur um eine *asymptotische* Reihe handelt,

<sup>3</sup>Im Urnenmodell der Kombinatorik nennt man dies das „Ziehen ohne Zurücklegen ohne Berücksichtigung der Anordnung“.

Der Erwartungswert ist der Wert, den man für den Mittelwert erwartet, wenn man einen Versuch sehr oft wiederholt. Häufig schreibt man  $\mu := \mathcal{E}(X)$  und  $\sigma^2 := \text{Var}(X)$ ; die *Standardabweichung* von  $X$  ist  $\sigma := \sqrt{\sigma^2}$ . Man bezeichnet häufig  $M_X(t) := \mathcal{E}(e^{tX})$  als die *momenterzeugende Funktion* von  $X$ , da  $\mathcal{E}(X^k) = M_X^{(k)}(0)$  gilt.

Für zwei Zufallsvariablen  $X$  und  $Y$  sowie  $a, b \in \mathbb{R}$  gilt die Linearität  $\mathcal{E}(aX + bY) = a\mathcal{E}(X) + b\mathcal{E}(Y)$  und, falls sie statistisch unabhängig sind, die Multiplikativität  $\mathcal{E}(X \cdot Y) = \mathcal{E}(X)\mathcal{E}(Y)$ . Für die Varianz gilt Translationsinvarianz und quadratische Homogenität, d. h.  $\text{Var}(aX + b) = a^2 \text{Var}(X)$ , im Falle der statistischen Unabhängigkeit von  $X$  und  $Y$  auch die Additivität  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ .

### B.1.4 Binomial- und Normalverteilung

Eine Zufallsvariable  $X$  genügt der *Binomialverteilung* mit der Erfolgswahrscheinlichkeit  $p$  (und  $q := 1 - p$ ), falls ihre Ergebnismenge  $\Omega = \{0, \dots, n\}$  ist und<sup>4</sup>

$$P(X = k) = B(n; p; k) := \binom{n}{k} p^k q^{n-k} \quad \text{für alle } k \in \{0, \dots, n\} \quad (\text{B.5})$$

gilt. Man nennt  $X$  in diesem Fall verkürzt  $(n, p)$ -binomialverteilt oder kurz  $B(n, p)$ -verteilt. Für diese Zufallsvariablen gilt  $\mathcal{E}(X) = np$ ,  $\text{Var}(X) = npq$  und ferner  $\mathcal{E}(e^{tX}) = (q + pe^t)^n = [1 + p(e^t - 1)]^n$ . Sind zwei Zufallsvariablen  $X_1$  und  $X_2$  statistisch unabhängig und  $B(n_1, p)$ - bzw.  $B(n_2, p)$ -verteilt, so genügt deren Summe  $X := X_1 + X_2$  einer  $B(n_1 + n_2, p)$ -Verteilung.

Eine kontinuierliche Zufallsvariable  $X$  genügt der Normalverteilung mit Mittelwert  $\mu$  und Standardabweichung  $\sigma$ , der  $N(\mu, \sigma)$ -Verteilung, wenn

$$P(a \leq X \leq b) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x=a}^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (\text{B.6})$$

gilt. Für  $\mu = 0$  und  $\sigma = 1$  spricht man von der *Standard-Normalverteilung*. In diesem Fall nennt man den Integranden einschließlich dem Vorfaktor die *GAUßsche Glockenkurve*  $\varphi(x)$ , die Funktion  $\Phi(z) := \int_{x=-\infty}^z \varphi(x) dx$  nennt man ihre *Verteilungs-* oder seltener auch *Summenfunktion*. Beide Funktionen sind tabelliert; vgl. z. B. KREYSZIG [103] oder STANGE [170].

Es ist  $P(a \leq X \leq b) \approx \Phi[(npq)^{-1/2}(b - np + 1/2)] - \Phi[(npq)^{-1/2}(a - np - 1/2)]$  für eine binomialverteiltes  $X$ ; die Abschätzung gilt als gut, falls  $npq > 9$  ist.

---

<sup>4</sup>Die Größe  $B(n; p; k)$  ist also die Wahrscheinlichkeit, bei  $n$  unabhängigen Versuchen genau  $k$  Erfolge zu erzielen, sofern  $p$  die Erfolgswahrscheinlichkeit eines Einzelversuchs ist.

## B.2 Grundbegriffe der Algebra

In diesem Abschnitt werden einige Begriffe der Algebra definiert; die aufgestellten Betrachtungen sind weitaus allgemeiner als zum Verständnis der Dissertation erforderlich, gestatten hierdurch aber einen besseren Überblick über bestimmte Zusammenhänge.

### B.2.1 Relationen

Ist  $X$  eine beliebige Menge, so nennt man eine Teilmenge  $R \subseteq M \times M$  des kartesischen Produktes eine *Relation*. Wichtige Eigenschaften, die solche Relationen besitzen können, sind (1) *Reflexivität*:  $(\forall x \in M)((x, x) \in R)$ ; (2) *Symmetrie*:  $(\forall x, y \in M)((x, y) \in R \Leftrightarrow (y, x) \in R)$ ; (3) *Antisymmetrie*:  $(\forall x, y \in M)((x, y) \in R \wedge (y, x) \in R \Rightarrow x = y)$  und (4) *Transitivität*:  $(\forall x, y, z \in M)((x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R)$ .

Ein (*Halb-*)*Ordnung* ist eine reflexive, antisymmetrische und transitive Relation (z. B.  $\leq$  auf  $\mathbb{R}$ ); gilt ferner  $(x, y) \in R \vee (y, x) \in R$ , so spricht man von einer *Totalordnung*. Eine *Äquivalenzrelation* ist eine reflexive, symmetrische und transitive Relation; sie teilt die Menge in paarweise disjunkte *Äquivalenzklassen* auf. Äquivalenzrelationen schreibt man häufig  $x \sim y$ .

### B.2.2 Grundlagen

Ist auf einer nicht-leeren Menge  $S$  eine Abbildung  $S^2 \rightarrow S$  definiert, so nennt man diese Abbildung eine *zweistellige Verknüpfung* (oder *binäre Operation*) auf  $S$ ; man nennt  $S$  *abgeschlossen* unter dieser Verknüpfung.

Eine *algebraische Struktur* ist ein Tupel, bestehend aus einer Menge  $S \neq \emptyset$  und einer oder mehreren Verknüpfungen auf ihr. Die Mächtigkeit  $|S|$  der Menge nennt man die *Ordnung* der algebraischen Struktur; man spricht von endlichen und unendlichen algebraischen Strukturen. Ist die Verknüpfung geläufig, so unterdrückt man in der Regel die Angabe der Verknüpfung. Ist eine Teilmenge einer Struktur bezüglich derselben Verknüpfungen selbst eine Struktur des gleichen Typs, so spricht man von einer *Unter-* oder *Teilstruktur*.

Die verschiedenen algebraischen Strukturen unterscheiden sich in den Eigenschaften, die die Verknüpfungen besitzen. Die einfachste denkbare algebraische Struktur ist ein *Gruppoid* oder *Magma*  $(G, \oplus)$ , bestehend aus einer Menge  $G$  und einer zweistelligen Verknüpfung  $\oplus : G \times G \rightarrow G$ ; für die Verknüpfung zweier Elemente  $a, b \in G$  schreibt man hier wie im allgemeinen  $a \oplus b := \oplus(a, b)$ .

Die wichtigsten grundlegenden Eigenschaften, die eine Verknüpfung erfüllen kann, sind die folgenden:

1. Assoziativgesetz:  $(\forall a, b, c \in G)((a \oplus b) \oplus c = a \oplus (b \oplus c))$
2. Existenz eines neutralen Elements:  $(\exists e \in G)(\forall a \in G)(a \oplus e = e \oplus a = a)$
3. Existenz inverser Elemente:  $(\forall a \in G)(\exists a' \in G)(a \oplus a' = a' \oplus a = e)$
4. Kommutativgesetz:  $(\forall a, b \in G)(a \oplus b = b \oplus a)$

Je nachdem, welche dieser Eigenschaften erfüllt werden, benennt man die algebraischen Strukturen.

**Definition B.2 (Gruppen und verwandte Strukturen)**

Eine Gruppe ist ein Gruppoid, in dem die Verknüpfung die Bedingungen 1 bis 3 erfüllt. Gilt schwächer nur die Bedingung 1, so spricht man von einer Halbgruppe; ein Monoid ist eine Halbgruppe, in der zusätzlich die Bedingung 2 erfüllt ist. Halbgruppen, Monoide und Gruppen nennt man kommutativ oder häufiger auch abelsch, wenn die Bedingung 4 gilt.

Die Einheitsgruppe  $M^\times$  eines Monoids  $M$  besteht aus seinen invertierbaren Elementen. Für eine Verknüpfung auf einer algebraischen Struktur wird in der Regel eine von zwei Schreibweisen verwendet, die als *additiv* bzw. als *multiplikativ* bezeichnet und in der folgenden Tabelle erklärt werden:

Notation	Verknüpfung	Neutrales Element	Inverse von $a$
additiv	+	0	$-a$
multiplikativ	$\cdot$	1	$a^{-1}$ .

Wie in der Arithmetik wird in multiplikativer Notation das Verknüpfungszeichen oft weggelassen, und man verwendet die üblichen Rechenregeln wie „Punkt vor Strich“ usw. Die additive Notation wird vorwiegend für kommutative Verknüpfungen verwendet.

Eine „strukturerhaltende“ Abbildung zweier gleichartiger algebraischer Strukturen nennt man einen *Homomorphismus*; zum Beispiel sind die Vektorraum-Homomorphismen die linearen Abbildungen.

**Definition B.3 (Eigenschaften von Homomorphismen)**

Ist ein Homomorphismus injektiv, surjektiv oder bijektiv, so spricht man von einem Monomorphismus, einem Epimorphismus bzw. einem Isomorphismus. Einen Homomorphismus einer Struktur auf sich nennt man Endomorphismus, ist dieser bijektiv einen Automorphismus.

Man nennt zwei Strukturen *isomorph*, wenn überhaupt ein Isomorphismus zwischen ihnen existiert; in diesem Falle sind sie in bezug auf die betrachtete Struktur völlig gleichwertig.

### B.2.3 Untergruppen und Normalteiler

Eine Teilmenge  $U$  einer Gruppe  $G$ , die hinsichtlich derselben Verknüpfung selbst eine Gruppe bildet, nennt man eine *Untergruppe* von  $G$ . Ist  $H \subseteq G$  eine Untergruppe, so bezeichnet man die Mengen der Form  $gH$  mit  $g \in G$  als *Linksnebenklassen* (engl. *left cosets*) von  $G$  bzgl.  $H$ ; die Zugehörigkeit zu einer Linksnebenklasse ist eine Äquivalenzrelation. Rechtsnebenklassen definiert man analog.

Fallen die Links- und Rechtsnebenklassen zusammen, gilt also  $gH = Hg$  für alle  $g \in G$ , so nennt man  $H$  einen *Normalteiler*<sup>5</sup> in  $G$ . Jede Gruppe  $G$  besitzt die *trivialen Normalteiler*  $\{e\}$  und  $G$ ; sind dies die einzigen Normalteiler, so nennt man die Gruppe *einfach*.

Ist  $N \subseteq G$  Normalteiler einer Gruppe  $G$ , so bildet die Menge der Nebenklassen selbst eine Gruppe, wenn man  $g_1N \cdot g_2N := g_1g_2N$  fordert; diese Gruppe nennt man *Quotienten-* oder *Faktorgruppe*  $G/N$ , ihre Mächtigkeit  $[G : N] = |G| / |N|$  nennt man den *Index*. Die Abbildung  $G \rightarrow G/N, g \mapsto gN$  ist ein Homomorphismus mit Kern  $N$ , und umgekehrt ist jeder Kern eines Homomorphismus von Gruppen ein Normalteiler.

#### Definition B.4 (Normalisator und Zentralisator)

Ist  $G$  eine Gruppe und  $H$  eine Teilmenge von  $G$ , so nennt man die Mengen  $N_G(H) := \{g \in G \mid gH = Hg\}$  und  $Z_G(H) := \{g \in G \mid (\forall h \in H)(gh = hg)\}$  den Normalisator bzw. Zentralisator von  $H$  in  $G$ . Den Zentralisator von  $G$  nennt man auch das Zentrum von  $G$ .

Die Normalisatoren und Zentralisatoren sind immer Untergruppen. Ist  $H$  eine Untergruppe, so ist  $N_G(H)$  die größte Gruppe, in der  $H$  ein Normalteiler ist; per Definition ist  $H$  genau für  $N_G(H) = G$  ein Normalteiler in  $G$ .

### B.2.4 Symmetrische und alternierende Gruppen

Eines der wichtigsten Beispiele endlicher Gruppen bilden die symmetrischen Gruppen, da nach dem *Satz von CAYLEY* jede endliche Gruppe der Ordnung  $n \in \mathbb{N}$  zu einer Untergruppe der symmetrischen Gruppe  $S_n$  isomorph ist.<sup>6</sup>

Für  $n \in \mathbb{N}$  ist die *symmetrische Gruppe*  $S_n$  die Gruppe aller Permutationen der Zahlen  $\{1, \dots, n\}$ . Die Elemente von  $S_n$  können in der *Zykelzerlegung* geschrieben werden, also als Produkt von  $k$ -Zyklen  $(p_0p_1 \dots p_{k-1})$  ( $p_i$  geht über in  $p_{(i+1) \bmod k}$ ), wobei keines der  $p_i$  in mehr als einem Zyklus vorkommt.

---

<sup>5</sup>Andere Begriffe sind *normale* oder *invariante Untergruppe* (engl. *normal subgroup*).

<sup>6</sup>Vgl. zu diesem Unterabschnitt zum Beispiel die Bücher von VAN DER WAERDEN [180], SPEISER [169] und KOCHENDOERFFER [96].

Andererseits kann jede Permutation als Produkt von Transpositionen ( $p_0p_1$ ) geschrieben werden. Die Untergruppe derjenigen Permutationen in  $S_n$ , die als Produkt einer geraden Zahl von Transpositionen geschrieben werden können, nennt man die alternierende Gruppe  $A_n$ ; sie wird zum Beispiel von den Dreierzyklen der Form  $(12k)$  für  $k \in \{3, \dots, n\}$  erzeugt. Zwei Permutationen in  $S_n$  oder  $A_n$  sind genau dann konjugiert, wenn in ihren Zykelzerlegungen die Längen aller Zyklen übereinstimmen.

Mit Ausnahme von  $n = 4$  sind die alternierenden Gruppen  $A_n$  einfach und bilden die einzigen nicht-trivialen Normalteiler der Gruppen  $S_n$ . Die Gruppen  $S_4$  und  $A_4$  besitzen als zusätzlichen Normalteiler die KLEINSche Vierergruppe  $V = \{e, (12)(34), (13)(24), (14)(23)\} \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ .

## B.2.5 Gruppenwirkungen

In der Theorie der CSS-Codes werden einige Begriffe aus der Theorie der *Gruppenwirkungen* (auch *-operationen* oder seltener *-aktionen*) verwendet, die aus diesem Grund hier kurz eingeführt werden sollen.

### Definition B.5 (Gruppenwirkungen)

Es seien  $G$  eine Gruppe und  $X$  eine Menge. Eine Abbildung  $\sigma : G \times X \rightarrow X$  bezeichnet man als Gruppenwirkung, falls

$$\sigma(1, x) = x \quad \text{und} \quad \sigma(g_1, \sigma(g_2, x)) = \sigma(g_1g_2, x)$$

für alle  $g_1, g_2 \in G$  und  $x \in X$  gelten. Man setzt auch  $gx := \sigma(g, x)$ .

Die Standardbeispiele für Wirkungen einer Gruppe  $G$  auf sich selbst sind die *Linksmultiplikation*  $(g, h) \mapsto gh$ , die *Rechtsmultiplikation*  $(g, h) \mapsto hg^{-1}$  sowie die *Konjugation*  $(g, h) \mapsto ghg^{-1}$ ; ebenso wirkt die Algebra  $M_n(K)$  auf den Vektorraum  $K^n$ .

Die Menge  $Gx$  heißt *Bahn* (oder *Orbit*) von  $x$  unter  $G$ ; existiert nur eine Bahn, so nennt man die Wirkung *transitiv*. Die Bahnen der Konjugationswirkung nennt man die *Konjugiertenklassen* der Gruppe. Ganz allgemein definieren die Bahnen einer Wirkung eine Äquivalenzrelation auf einer Gruppe.

Ist  $X$  eine Menge und  $S_X$  die Gruppe der Permutationen auf  $X$ , so entspricht jeder Gruppenwirkung  $\sigma : G \times X \rightarrow X$  genau ein Gruppen-Homomorphismus  $\varphi : G \rightarrow S_X$  und umgekehrt, wenn man die Gleichheit  $\sigma(g, x) = [\varphi(g)](x)$  für alle  $g \in G$  und  $x \in X$  fordert.

### Definition B.6 (Stabilisator)

Wirkt eine Gruppe  $G$  auf eine Menge  $X$ , und ist  $A \subseteq X$ , so nennt man die Menge  $G_A := \{g \in G \mid (\forall x \in A)(g \cdot x = x)\}$  den Stabilisator von  $A$  in  $G$ .

Andere Namen für Stabilisator sind *Standuntergruppe* und *Isotropiegruppe*.

## B.3 Ringe und Körper

In diesem Abschnitt werden algebraische Strukturen mit zwei Verknüpfungen, insbesondere Ringe und Körper, betrachtet. Ziel ist es unter anderem, einige Grundlagen aus der Theorie der endlichen Körper zusammenzustellen. Für Einzelheiten sei unter anderem auf die Bücher von BOSCH [22] und ARTIN [3], für endliche Körper insbesondere auf das von LIDL UND NIEDERREITER [107] verwiesen.

### B.3.1 Mengen mit zwei Verknüpfungen

Man betrachte ein Tripel  $(R, +, \cdot)$  derart, daß  $(R, +)$  und  $(R, \cdot)$  Gruppoide bilden. Die beiden Verknüpfungen (Addition und Multiplikation) können z. B. durch folgende *Distributivgesetze* in Verbindung gebracht werden:

1.  $(\forall a, b, c \in R)((a + b) \cdot c = a \cdot c + b \cdot c)$  und
2.  $(\forall a, b, c \in R)(c \cdot (a + b) = c \cdot a + c \cdot b)$ .

Ist die Multiplikation kommutativ, so fallen beide Bedingungen zusammen. In einem Ring bezeichne  $R^* := R \setminus \{0\}$ .<sup>7</sup>

#### Definition B.7 (Ringe und Körper)

Eine algebraische Struktur  $(R, +, \cdot)$  mit zwei Verknüpfungen derart, daß  $(R, +)$  eine abelsche Gruppe ist und die Distributivgesetze gelten, nennt man

1. einen Ring, falls  $(R, \cdot)$  eine Halbgruppe ist;
2. einen Ring mit Eins, falls  $(R, \cdot)$  ein Monoid ist und  $1 \neq 0$  gilt;<sup>8</sup>
3. einen Körper, falls zusätzlich  $(R \setminus \{0\}, \cdot)$  eine abelsche Gruppe ist.

Man spricht von einem kommutativen Ring (mit Eins), wenn die Multiplikation kommutativ ist; verzichtet man beim Körper auf die Forderung, daß die Multiplikation kommutativ ist, so spricht man von einem Schiefkörper.<sup>9</sup>

---

<sup>7</sup>Dieses Symbol wird häufig auch für die Einheitengruppe des Ringes gebraucht.

<sup>8</sup>Läßt man  $1 = 0$  zu, so erhält man eine einelementige Struktur: für ein Element  $a \in R$  gilt  $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$ , woraus sich nach beidseitigem Kürzen  $0 \cdot a = 0$  ergibt; aus der Voraussetzung folgt nun  $a = 1 \cdot a = 0 \cdot a = 0$ .

<sup>9</sup>Seltener wird für einen Ring schon die Existenz eines Einselementes gefordert. Statt Ring mit Eins sagt man auch *unitärer Ring*, statt Schiefkörper auch *Divisionsring*. Die englischen Begriffe für Körper und Schiefkörper sind *field* und *skew field*. Das bekannteste Beispiel eines Schiefkörpers bilden die HAMILTONschen *Quaternionen*, die für vier Basiselemente 1, i, j und k, die die Multiplikationsregeln  $i^2 = j^2 = k^2 = ijk = -1$  erfüllen, durch  $\mathbb{H} = \{x_0 \cdot 1 + x_1 \cdot i + x_2 \cdot j + x_3 \cdot k \mid x_0, x_1, x_2, x_3 \in \mathbb{R}\}$  definiert sind. Sie können realisiert werden, indem man 1, i, j und k mit den PAULI-Matrizen  $\mathbb{I}$ ,  $i\sigma_z$ ,  $i\sigma_y$  und  $i\sigma_x$  (in dieser Reihenfolge) identifiziert.

**Definition B.8 (Ideale)**

Eine nicht-leere additive Untergruppe  $I$  eines Ringes  $R$  nennt man Linksideal, falls  $RI = I$  ist, Rechtsideal, falls  $IR = I$  ist und (beidseitiges oder zweiseitiges) Ideal, falls beides zutrifft.

Ideale sind also Unterringe, und in kommutativen Ringen fallen alle drei Begriffe zusammen. Die Ideale  $I = \{0\}$  und  $I = R$  nennt man die *trivialen Ideale*; ein Ideal  $I \neq R$  nennt man *echtes Ideal*. Für  $a \in R$  nennt man  $Ra$  das *Haupt-Linksideal*,  $aR$  das *Haupt-Rechtsideal* und  $RaR$  das (*beidseitige*) *Hauptideal* zu  $a$ . In kommutativen Ringen fallen die Begriffe zusammen. Sind in einem Integritätsring alle Ideale Hauptideale, so spricht man auch von einem *Hauptidealring*.

Man kann einen Ring  $R$  nach einem Ideal  $I$  faktorisieren und erhält den Faktorring  $R/I$ ; ist  $R$  kommutativ, so auch  $R/I$ .

**Definition B.9 (Eigenschaften von Idealen)**

Ein Ideal  $I$  eines Ringes  $R$  heißt maximal, wenn es kein größeres echtes Ideal in  $R$  gibt; es heißt prim, falls  $(\forall a, b \in R)(ab \in I \Rightarrow a \in I \vee b \in I)$  gilt.

Ein Ideal  $I$  eines kommutativen Ringes mit Eins  $R$  ist genau dann maximal, wenn  $R/I$  ein Körper ist; es ist genau dann prim, wenn  $R/I$  ein Integritätsbereich ist; insbesondere ist jedes maximale Ideal prim. Ist  $R$  ein Integritätsbereich und  $c \in R$ , so ist  $R/cR$  genau dann ein Körper, wenn  $cR$  prim ist.

Zur Konstruktion endlicher Körper bedient man sich schließlich noch nullteilerfreier Ringe.

**Definition B.10 (Nullteiler und nullteilerfreie Ringe)**

Ist  $R$  ein Ring, so nennt man ein Element  $a \in R^*$  einen Links-Nullteiler, falls ein  $b \in R^*$  derart existiert, daß  $a \cdot b = 0$  gilt; entsprechend definiert man Rechts-Nullteiler sowie (beidseitige) Nullteiler<sup>10</sup>. Besitzt ein Ring  $R$  keine Nullteiler, so nennt man ihn nullteilerfrei. Ein kommutativer nullteilerfreier Ring mit Eins heißt Integritätsbereich (oder auch Integritätsring).

Die Nullteilerfreiheit eines Rings  $R$  wird also durch die Gültigkeit der Aussage  $(\forall a, b \in R)(a \cdot b = 0 \Rightarrow a = 0 \vee b = 0)$  charakterisiert. Der Ring der ganzen Zahlen  $(\mathbb{Z}, +, \cdot)$  ist das Standardbeispiel für einen Integritätsbereich.

**B.3.2 Restklassenringe und Primzahlkörper**

Faktorisiert man den Integritätsbereich  $\mathbb{Z}$  nach dem Hauptideal  $d\mathbb{Z}$  für ein  $d \in \mathbb{N} \setminus \{1\}$ , so erhält man  $\mathbb{Z}/d\mathbb{Z}$ , den *Restklassenring modulo  $d$* . Dieser Restklassenring wird in der Regel mit dem *Vertretersystem*  $\mathbb{Z}_d := \{0, \dots, d - 1\}$

---

<sup>10</sup>Es wird nicht gefordert, daß ein Element  $b \in R^*$  existiert, für das  $a \cdot b = b \cdot a = 0$  gilt!

identifiziert, wobei Addition und Subtraktion in dieser Arbeit durch „ $\oplus$ “ bzw. „ $\ominus$ “ bezeichnet werden; ferner werde  $\mathbb{Z}_d^* := \mathbb{Z}_d \setminus \{0\}$  gesetzt.

Der Restklassenring  $\mathbb{Z}/d\mathbb{Z}$  ist genau dann ein Körper, wenn  $d$  eine Primzahl ist. Andernfalls läßt sich beispielsweise  $d = a \cdot b$  für  $a, b \in \{2, \dots, d-1\}$  schreiben, so daß  $a$  und  $b$  Nullteiler sind. Die endlichen *Primzahlkörper* werden  $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$  geschrieben. Allgemein ist jeder endliche Integritätsbereich und nach dem *Satz von WEDDERBURN* auch jeder endliche Schiefkörper ein Körper.

### B.3.3 Endliche Körper

Ein *endlicher Körper* (engl. *finite field* oder *Galois field*) ist ein Körper  $K$  mit einer endlichen Ordnung  $|K|$ . Eine genaue Behandlung der endlichen Körper erfordert die Theorie der *Körpererweiterungen*, die hier nur gestreift werden kann. Für diese Dissertation bedeutsam ist der folgende Satz, der hier ohne Beweis aufgeführt wird.<sup>11</sup>

#### Satz B.11 (Klassifizierung der endlichen Körper)

*Ein endlicher Körper der Ordnung  $q \in \mathbb{N}$  existiert genau dann, wenn  $q$  eine Primzahlpotenz ist. Je zwei endliche Körper gleicher Ordnung sind isomorph.*

Im allgemeinen schreibt man  $\mathbb{F}_q$  für diesen bis auf Isomorphie eindeutigen Körper; eine andere häufig verwendete Schreibweise ist  $\text{GF}(q)$  für engl. *Galois field*. Will man die Primzahl und den Exponenten betonen,  $q = p^n$ , so schreibt man auch  $\mathbb{F}_{p^n}$  bzw.  $\text{GF}(p^n)$ .

#### Definition B.12 (Charakteristik eines Körpers)

*Ist  $K$  ein Körper, so nennt man die kleinste positive Anzahl von Einsen, die man addieren muß, um den Wert Null zu erhalten, die Charakteristik  $\text{char } K$  dieses Körpers; existiert keine solche Zahl, so setzt man  $\text{char } K := 0$ .*<sup>12</sup>

Unter den endlichen Körpern sind die genau die Körper  $\mathbb{F}_p$  für eine Primzahl  $p$  die *Primkörper*, das heißt, sie enthalten keinen echten Unterkörper. Den einzigen Primkörper der Charakteristik Null bilden die rationalen Zahlen  $\mathbb{Q}$ . Die Körper der Charakteristik 2 weisen wegen  $1+1 = 0$  einige Besonderheiten auf.

---

<sup>11</sup>Eine *Primzahl* ist bekanntlich eine Zahl  $n \in \mathbb{N}$ , die genau zwei Teiler besitzt; die ersten Primzahlen sind 2, 3, 5, 7 usw. (nicht aber 1); als *Primzahlpotenz* bezeichnet man eine Zahl, die sich als  $p^n$  für eine Primzahl  $p$  und ein  $n \in \mathbb{N}$  darstellen läßt. Nach dem *Fundamentalsatz der Zahlentheorie* (oder der *Arithmetik*) besitzt jede natürliche Zahl eine bis auf die Anordnung der Faktoren eindeutige Zerlegung im Primfaktoren.

<sup>12</sup>Die allgemeine Definition der Charakteristik in einem Integritätsbereich  $R$  lautet: ist  $\varphi : \mathbb{Z} \rightarrow R$  mit  $\varphi(n) := n \cdot 1$  der *kanonische Ringhomomorphismus*, so nennt man eine Zahl  $p \in \mathbb{N}_0$ , die das Hauptideal Kern  $\varphi$  erzeugt, die Charakteristik von  $R$ .

**Satz B.13 (Primzahlkörper)**

Der Restklassenring  $\mathbb{Z}/p\mathbb{Z}$  ist genau dann nullteilerfrei, wenn  $p$  eine Primzahl ist. Genau in diesem Fall ist er auch ein Körper, und man schreibt auch  $\mathbb{F}_p$ .

Für einen Körper  $K$  bezeichne  $K[X]$  den Ring der Polynome über  $K$  in der Variablen  $X$ .<sup>13</sup> Ein Polynom  $f \in K[X]$  nennt man *irreduzibel* (unzerlegbar), wenn es nicht in der Form  $f = gh$  für Polynome  $g, h \in K[X]$  mindestens ersten Grades geschrieben werden kann.

Man kann zeigen, daß es für  $n \in \mathbb{N}$  stets ein irreduzibles (normiertes) Polynom  $n$ -ten Grades gibt. Im folgenden Satz sei  $I(f) := \{fg \mid g \in K[X]\}$  das durch ein irreduzibles Polynom  $f \in K[X]$  erzeugte Hauptideal.

**Satz B.14 (Die Galois-Konstruktion)**

Es seien  $p$  eine Primzahl,  $n \in \mathbb{N}$  und  $f \in \mathbb{F}_p[X]$  ein irreduzibles Polynom  $n$ -ten Grades. Der Quotient  $\mathbb{F}_{p^n} := \mathbb{F}_p[X]/I(f)$  ist ein endlicher Körper, und es gilt  $|\mathbb{F}_{p^n}| = p^n$  sowie  $\text{char } \mathbb{F}_{p^n} = p$ .

Als Vertretersystem wählt man üblicherweise die Polynome aus  $K[X]$  mit Grad echt kleiner als  $n$ . Mittels der Koeffizienten dieser Polynome kann man also  $\mathbb{F}_{p^n}$  in naheliegender Weise mit  $\mathbb{Z}_p^n$  identifizieren, wenn nach der Multiplikation eine Polynomdivision mit  $f$  durchgeführt wird.

### B.3.4 Körpererweiterungen

Ist  $L$  ein Körper und  $K \subseteq L$  ein Unterkörper, so nennt man  $L/K$  eine *Körpererweiterung*. Eine Körpererweiterung  $L/K$  liefert immer auch einen Vektorraum  $L$  über dem Skalkörper  $K$ , dessen Dimension, der *Grad* von  $L/K$ , im folgenden als endlich angenommen werde (endliche Körpererweiterung).

Ein Element  $a \in L$  definiert eine  $K$ -lineare Abbildung (also eine Matrix)  $\varphi_a : L \rightarrow L$ ,  $\varphi_a(x) := ax$ , auf  $L$ . Bezeichnet  $\chi_a := \chi_{\varphi_a}$  das charakteristische Polynom, so definiert man *Spur* und *Norm*<sup>14</sup> von  $a$  mittels der Festsetzungen  $\text{Spur}(a) := \text{Spur } \varphi_a$  und  $N(a) := \det \varphi_a$ .

---

<sup>13</sup>Ein *Polynom* (oder auch *ganzrationale Funktion*) ist eine (endliche) Linearkombination von *Monomen*  $x^k$ ,  $k \in \mathbb{N}_0$ , über einem Ring  $R$ . Man schreibt  $f(x) = \sum_{i=0}^n a_i x^i$  für ein  $n \in \mathbb{N}_0$  und nennt  $n$ , falls  $a_n \neq 0$  ist, den *Grad* des Polynoms; der Grad des Nullpolynoms ist  $-\infty$ . Der Koeffizient  $a_n$  heißt *Leitkoeffizient* oder *führender Koeffizient*, ist er gleich Eins, so nennt man das Polynom *normiert*. Die Polynome bilden mit der üblichen Addition und Multiplikation einen Ring  $R[x]$ . Eine *Nullstelle* (oder älter *Wurzel*) eines Polynoms  $f$  ist ein  $\alpha \in R$ , für das  $f(\alpha) = 0$  gilt.

<sup>14</sup>Dieser Begriff hängt nicht mit dem gleichnamigen Begriff in der Funktionalanalysis zusammen.

**Lemma B.15 (Spur und Norm)**

Die Körpererweiterung  $\mathbb{F}_{p^n}/\mathbb{F}_p$  hat den endlichen Grad  $[\mathbb{F}_{p^n} : \mathbb{F}_p] = n$ , und für jedes Element  $a \in \mathbb{F}_{p^n}$  gilt

$$\chi_a(x) = \prod_{k=0}^{n-1} (x - a^{p^k}) = (x - a) \cdot (x - a^p) \cdot \dots \cdot (x - a^{p^{n-1}}),$$

insbesondere also  $\text{Spur}(a) = \sum_{k=0}^{n-1} a^{p^k}$  und  $N(a) = \prod_{k=0}^{n-1} a^{p^k}$ .

Die Spur ist ein lineares Funktional von  $\mathbb{F}_{p^n}$  nach  $\mathbb{F}_p$ , und alle linearen Funktionale von  $\mathbb{F}_{p^n}$  nach  $\mathbb{F}_p$  haben die Form  $\alpha \mapsto \text{Spur}(\beta\alpha)$  für ein  $\beta \in \mathbb{F}_{p^n}$ .

## B.4 Begriffe der linearen Algebra

In diesem Abschnitt werden einige (allgemein bekannte) Grundbegriffe der linearen Algebra zusammengestellt, die dem Leser vertraut sein sollten. Die HURWITZ-Kriterien zur Bestimmung der Definitheit und Semidefinitheit von Matrizen finden sich im Abschnitt B.5.

Für die geläufigen Begriffe der linearen Algebra (Rang, Determinante, Spur, Kern, Bild, selbstadjungierte, unitäre, normale Matrizen, Spektralsatz, Eigenwerttheorie) verweise ich auf die Lehrbuch-Literatur.

### B.4.1 Vektorräume, Moduln und Algebren

Ein Vektorraum ist ein Quadrupel  $(V, K, +, \cdot)$ , bestehend aus einer abelschen Gruppe  $(V, +)$  von *Vektoren*, einem Körper  $K$  und einer *Skalarmultiplikation*  $\cdot : K \times V \rightarrow V$ , die folgende Bedingungen erfüllt: (1)  $\lambda(x + y) = \lambda x + \lambda y$ , (2)  $(\lambda + \mu)x = \lambda x + \mu x$ , (3)  $\lambda(\mu x) = (\lambda\mu)x$  und (4)  $1 \cdot x = x$ , jeweils für alle  $x, y \in V$  und  $\lambda, \mu \in K$ . Man nennt  $K$  den *Skalkörper* des Vektorraums und spricht von einem *Vektorraum über  $K$* . Ist  $K \in \{\mathbb{R}, \mathbb{C}\}$ , so schreibt man sehr oft auch  $\mathbb{K}$  statt  $K$ . Läßt man zu, daß  $K$  nur ein Ring ist, so spricht man von einem *Modul* statt von einem Vektorraum. Ist  $V$  ein Ring und gilt  $\lambda(xy) = (\lambda x)y = x(\lambda y)$ , so handelt es sich um eine *Algebra*.

Die wichtigsten Vektorräume sind  $K^n$  für einen Körper  $K$  und eine natürliche Zahl  $n \in \mathbb{N}$ . In diesem Text seien alle Vektoren Spaltenvektoren. Die quadratischen  $n \times n$ -Matrizen  $M_n(K) := K^{n \times n}$  bilden eine Algebra.

### B.4.2 Basen

Jeder Vektorraum besitzt eine Basis. Zum Beispiel besitzen  $\mathbb{R}^n$  über  $\mathbb{R}$  und  $\mathbb{C}^n$  über  $\mathbb{C}$  die *kanonische Basis* (oder *Standardbasis*)  $\{e_1, \dots, e_n\}$ , wobei

der Vektor  $e_i := (0, \dots, 0, 1, 0, \dots, 0)^t$  (die  $i$ -te Komponente ist Eins, die übrigen allesamt Null) ist. Je nach Zusammenhang ist es sinnvoll, eine Basis als eine ungeordnete *Menge* zu betrachten oder eine Anordnung der Basisvektoren zu definieren.<sup>15</sup>

### B.4.3 Lineare Abbildungen, Operatoren und Matrizen

Eine *Matrix* bezeichnet ein „rechteckiges Schema von Zahlen“. Bezüglich fest gewählter Basen kann jede lineare Abbildung zwischen zwei Vektorräumen durch eine Matrix dargestellt werden.<sup>16</sup> Ein *Operator* ist eine *lineare* Abbildung (ein Homomorphismus) zwischen zwei Vektorräumen.

Eine *Diagonalmatrix*  $D = (d_{ij})_{i,j=1}^n$  hat nur Einträge auf ihrer Diagonale, das heißt, für  $i \neq j$  ist  $d_{ij} = 0$ ; hierfür werde  $D = \text{diag}(d_{11}, d_{22}, \dots, d_{nn})$  geschrieben. Die gleiche Schreibweise wird für *blockdiagonale Matrizen* verwendet, deren Diagonalelemente Blöcke von beliebig besetzten Matrizen sind.

### B.4.4 Das charakteristische Polynom

Ist  $V$  ein Vektorraum über dem Skalkörper  $K$  und  $A$  eine  $n \times n$ -Matrix über  $V$ , so nennt man  $\chi_A(\lambda) := \det(\lambda \mathbb{I} - A) \in K[\lambda]$  das *charakteristische Polynom* von  $A$ .<sup>17</sup> Es ist

$$\chi_A(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \quad a_i \in K, \quad (\text{B.7})$$

Die Koeffizienten können mithilfe des Lemmas B.20 berechnet werden. Nach dem Satz von HAMILTON und CAYLEY gilt  $\chi_A(A) = 0$  als Matrixgleichung. Die Nullstellen des charakteristischen Polynoms  $\chi_A$  sind die Eigenwerte der Matrix  $A$  in ihrer algebraischen Vielfachheit.

### B.4.5 Die Jordansche Normalform

Matrizen können, selbst wenn sie nicht normal und somit nicht diagonalisierbar sind, auf eine bestimmte Normalform gebracht werden.

---

<sup>15</sup>In der Funktionalanalysis unterscheidet man für Hilberträume zwischen *Vektorraum-basis* oder *HAMEL-Basis* einerseits und *Orthonormalbasis* oder *Hilbertraumbasis* andererseits. In den hier hauptsächlich betrachteten endlichdimensionalen Vektorräumen ist diese Unterscheidung unerheblich, im Zweifel werden Orthonormalbasen verwendet.

<sup>16</sup>Im allgemeinen unterscheidet man nicht zwischen einer linearen Abbildung und der Matrix, die sie (stets bezüglich einer festen Basis) repräsentiert. In dem Sinne steht Matrix immer für eine lineare Abbildung; die wichtigste Ausnahme hiervon ist die Koeffizientenmatrix für bell-diagonale Zustände im ersten Hauptteil der Dissertation.

<sup>17</sup>Auch üblich ist die Definition  $\chi_A(\lambda) := \det(A - \lambda \mathbb{I})$ ; die hier gewählte Form hat den Vorteil, daß der Leitkoeffizient stets Eins und das Polynom somit normiert ist.

Ein JORDAN-Kästchen oder JORDAN-Block der Größe  $n \in \mathbb{N}$  zum Eigenwert  $\lambda \in K$  ist eine Matrix  $J_\lambda^{(n)} = (j_{ik})_{i,k=1}^n$ , deren Einträge auf der Hauptdiagonalen allesamt  $\lambda$  sind und deren Einträge auf der ersten oberen Nebendiagonalen alle Eins sind; es ist also  $j_{ik} = \lambda\delta_{ik} + \delta_{i,k-1}$ . Das charakteristische Polynom zu solch einem Block ist  $\chi(\lambda) = \lambda^n$ , die algebraische Vielfachheit des Eigenwerts  $\lambda$  ist also  $n$ . Die Vektoren  $e_1, e_2, e_3$  usw. sind nun Eigenvektor, Hauptvektor 2. Stufe, Hauptvektor 3. Stufe usw.; die geometrische Vielfachheit des Eigenwerts ist also Eins.

**Satz B.16 (Satz über die Jordansche Normalform)**

Ist  $A \in K^{n \times n}$  eine Matrix über dem Körper  $K$  und zerfällt<sup>18</sup> das charakteristische Polynom  $\chi_A = \det(\lambda\mathbb{I} - A)$  dieser Matrix, so läßt sich  $A$  durch eine Ähnlichkeitstransformation in die JORDANSche Normalform überführen, d. h., es gibt eine invertierbare Matrix  $T \in K^{n \times n}$  und für  $i \in \{1, \dots, m\}$  JORDAN-Blöcke  $J_{\lambda_i}^{(n_i)} \in K^{n_i \times n_i}$  mit  $\sum_{i=1}^m n_i = n$  derart, daß

$$A' := T^{-1}AT = \text{diag} \left( J_{\lambda_1}^{(n_1)}, J_{\lambda_2}^{(n_2)}, \dots, J_{\lambda_m}^{(n_m)} \right)$$

ist. Die Matrix  $A'$  ist bis auf Permutationen der JORDAN-Blöcke eindeutig.

Ein Beweis findet sich zum Beispiel bei KOECHER [97], Nr. 9.3.6 auf S. 279, oder bei ZURMÜHL [187], § 19.1 auf S. 245–248. Ist der Körper  $K$  algebraisch abgeschlossen – was nach dem Fundamentalsatz der Algebra insbesondere für  $K = \mathbb{C}$  der Fall ist – so gilt die getroffene Aussage für alle quadratischen Matrizen über diesem Körper.

## B.4.6 Vandermonde-Matrix und -Determinante

Ist  $K$  ein Körper und sind  $x_1, \dots, x_n \in K$ , so nennt man  $V := V(x_1, \dots, x_n) := (x_i^{j-1})_{i,j=1}^n \in K^{n \times n}$  (oder manchmal auch  $V^t$ ) eine VANDERMONDE-Matrix. Für ihre Determinante, die VANDERMONDE-Determinante, gilt der folgende Satz.

**Satz B.17 (Vandermonde-Determinante)**

Es ist  $\det V(x_1, \dots, x_n) = \prod_{i>j} (x_i - x_j)$ , und  $V$  ist genau dann invertierbar, wenn die Werte  $x_i$  alle verschieden sind.

*Beweis:* Der Beweis erfolgt durch Induktion, wofür von  $n-1$  auf  $n$  geschlossen werden muß. Hierzu sei  $R := (\delta_{ij} - x_1\delta_{i,k-1})_{i,j=1}^n \in K^{n \times n}$  (ähnlich wie bei einem JORDAN-Kästchen). Es gilt  $\det R = 1$ , so daß mit  $V' := VR$  nach dem

<sup>18</sup>Ein Polynom  $f(x) = a_n x^n + \dots + a_1 x + a_0$  über einem Körper zerfällt in Linearfaktoren, wenn es in der Form  $f(x) = a_n (x - \lambda_1) \cdots (x - \lambda_n)$  geschrieben werden kann.

Multiplikationssatz für Determinanten  $\det V' = \det V$  folgt. Schreibt man nun  $V = (v_{ij})_{i,j=1}^n$  und  $V' = (v'_{ij})_{i,j=1}^n$ , so stimmen die ersten Spalten von  $V$  und  $V'$  überein, und für  $j \in \{2, \dots, n\}$  ergibt sich

$$v'_{ij} = v_{ij} - x_1 v_{i,j-1} = x_i^{j-1} - x_1 x_i^{j-2} = (x_i - x_1) x_i^{j-2}. \quad (\text{B.8})$$

Bei  $V'$  stehen in der ersten Spalte also lauter Einsen, in der ersten Zeile mit Ausnahme der ersten Komponente Nullen. Entwickelt man die Determinante nach der ersten Zeile, so zeigt sich, daß  $\det V = \det V'_{11}$  ist, wenn  $V'_{11}$  diejenige Untermatrix von  $V'$  bezeichnet, bei der die erste Zeile und die erste Spalte gestrichen werden. Zieht man aus jeder Zeile den Faktor  $(x_i - x_1)$  heraus, so ist die noch zu berechnende Determinante eine VANDERMONDE-Determinante der Größe  $n - 1$ . Somit folgt

$$\det V = \prod_{i=2}^n (x_i - x_1) \cdot \det V(x_2, \dots, x_n) \quad (\text{B.9})$$

und hieraus die Behauptung aus der Induktionsvoraussetzung, q. e. d.

### B.4.7 Hilberträume

Ein *Prähilbertraum* ist ein (hier stets komplexer) Vektorraum, auf dem ein *Skalarprodukt* definiert ist; dies ist in der Mathematik eine Funktion  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ , die folgende Eigenschaften erfüllt: sie ist (1) im linken Argument linear und im rechten antilinear, (2) *symmetrisch*, d. h.  $\langle w, v \rangle = \langle v, w \rangle^*$  und (3) *positiv definit*, d. h.  $\langle v, v \rangle > 0$  für  $v \neq 0$ . Die erste Eigenschaft definiert eine *Sesquilinearform*; Antilinearität bedeutet  $\langle v, \alpha w \rangle = \alpha^* \langle v, w \rangle$ . In der Physik wird das Skalarprodukt allgemein als im rechten Argument linear und im linken Argument antilinear definiert. Ein *Hilbertraum* ist ein Hilbertraum, der bezüglich der durch  $\|x\| := \sqrt{\langle x, x \rangle}$  induzierten Norm vollständig ist. Man kann direkte Summen und Tensorprodukte von Hilberträumen bilden; vgl. auch Abschnitt B.6.<sup>19</sup>

In der Quantenmechanik wird sehr oft die DIRAC- oder *Bra-Ket-Notation* verwendet, in der man einen Vektor als „halbes Skalarprodukt“  $|v\rangle$  schreibt (Ket-Vektor). Stetige lineare Funktionale auf einem Hilbertraum  $\mathcal{H}$  können nach dem Satz von RIESZ und FRECHET immer in der Form  $|v\rangle \mapsto \langle w|v\rangle$  für ein  $w \in \mathcal{H}$  geschrieben werden und definieren einen Vektor  $\langle w|$  (Bra-Vektor).

#### Satz B.18 (Klassifizierung der Hilberträume)

*Zwei Hilberträume sind genau dann isomorph, wenn ihre Dimension übereinstimmt.*

---

<sup>19</sup>Für alle Einzelheiten verweise ich auf die Lehrbücher, zum Beispiel von WERNER [184], von REED UND SIMON [143] und von CONWAY [42].

Um für jede mögliche Dimension (jede Kardinalzahl) einen Hilbertraum zu konstruieren, sei nun  $I$  eine beliebige Indexmenge und

$$\ell^2(I) := \left\{ (x_i)_{i \in I} \in \mathbb{C}^I \mid \sum_{i \in I} |x_i|^2 < \infty \right\}. \quad (\text{B.10})$$

Die Summe bedeutet, daß höchstens abzählbar viele der  $x_i$  nicht verschwinden und eine aus den übrigen Elementen gebildete Reihe absolut konvergiert.<sup>20</sup> Die Dimension von  $\ell^2(I)$  ist  $|I|$ , und das *kanonische* oder *Standard-Skalarprodukt* ist  $\langle x|y \rangle := \sum_{i \in I} x_i^* y_i$ . Auf den endlichdimensionalen Hilberträumen  $\mathbb{C}^n$  ( $\dim \mathbb{C}^n = n \in \mathbb{N}$ ) lautet dies  $\langle v|w \rangle := \sum_{i=1}^n v_i^* w_i$  für Vektoren  $v = (v_1, \dots, v_n)^t$  und  $w = (w_1, \dots, w_n)^t$ .

Die HILBERT-SCHMIDT-Operatoren auf einem Hilbertraum bilden bezüglich  $\langle A|B \rangle := \text{Spur } A^\dagger B$  selbst einen Hilbertraum; für Matrizen  $A = (a_{ij})_{i,j=1}^n$  und  $B = (b_{ij})_{i,j=1}^n$  ist dann  $\langle A|B \rangle = \sum_{i,j=1}^n a_{ij}^* b_{ij}$ .

## B.5 Positive Matrizen

In diesem Abschnitt werden einige Grundbegriffe aus der Theorie positiver Matrizen zusammengestellt, insbesondere das HURWITZ-Kriterium, das in Kapitel 6 an vielen Stellen verwendet wird.

### B.5.1 Positive Matrizen und Minoren

Eine Matrix  $A \in \mathbb{C}^{n \times n}$  nennt man *positiv semidefinit* oder kurz *positiv*, wenn für alle Vektoren  $v \in \mathbb{C}^n$  die Ungleichung  $\langle v|Av \rangle \geq 0$  gilt; man nennt sie *positiv definit*, wenn zusätzlich aus  $\langle v|Av \rangle = 0$  bereits  $v = 0$  folgt. Eine Matrix ist genau dann positiv definit oder semidefinit, wenn sie hermitesch ist und nur positive bzw. nicht-negative Eigenwerte besitzt.

Zu einer Matrix  $A \in \mathbb{C}^{n \times n}$  kann man durch Streichen von Zeilen und Spalten *Untermatrizen* bilden. Wählt man eine  $k$ -elementige nicht-leere Teilmenge  $M \subseteq \{1, \dots, n\}$  aus und streicht alle Zeilen und Spalten, deren Position nicht in  $M$  vorkommt, so erhält man eine  $M \times M$ -Untermatrix, deren Determinante man als einen *Minor* (oder eine *Unterdeterminante*)  $k$ -ter Ordnung bezeichnet. Liegt die spezielle Form  $M = \{1, \dots, k\}$  vor, so spricht man von einer *Hauptuntermatrix* und einem *Hauptminor*. Es gibt somit  $\binom{n}{k}$  Minoren  $k$ -ter Ordnung, von denen jeweils einer ein Hauptminor ist; insgesamt gibt es daher  $2^n - 1$  Minoren und  $n$  Hauptminoren.<sup>21</sup>

<sup>20</sup>Genaugenommen bedeutet dies, daß  $(x_i)_{i \in I}$  eine *summierbare Familie* ist; für den Begriff der *Summierbarkeit* vergleiche man BOURBAKI [23], § 5 (*Sommes infinies dans les groupes commutatifs*) auf S. III.36–45.

<sup>21</sup>Die Benennung ist nicht einheitlich: oft bezeichnet man die Determinanten jeder qua-

## B.5.2 Das Hurwitz-Kriterium

Um zu prüfen, ob eine Matrix positiv definit oder semidefinit ist, können anstelle der Definition die HURWITZ-Kriterien [84] verwendet werden, die sehr häufig auch nach SYLVESTER benannt werden. Da ihre Bedeutung für das Kapitel 6 außerordentlich wichtig ist, sie in den Lehrbüchern aber nur schwierig aufzufinden sind, werden sie hier mit ihrem Beweis angeführt.<sup>22</sup>

### Lemma B.19 (Kriterium für positiv definite Matrizen)

*Eine hermitesche Matrix ist genau dann positiv definit, wenn alle ihre Hauptminoren positiv sind.*

*Beweis:* Es sei  $M = (m_{ij})_{i,j=1}^n \in \mathbb{C}^{n \times n}$  eine hermitesche Matrix und für  $k \in \{1, \dots, n\}$  bezeichne  $M_k := (m_{ij})_{i,j=1}^k \in \mathbb{C}^{k \times k}$  ihre Hauptuntermatrizen.

Ist nun die Matrix  $M$  positiv definit, so gilt dies auch für jedes  $M_k$ , denn zu einem Vektor  $x' = (x_1, \dots, x_k)^t \in \mathbb{C}^k$  kann man durch Anfügen von  $n - k$  Nullen den Vektor  $x = (x_1, \dots, x_k, 0, \dots, 0)^t \in \mathbb{C}^n$  bilden, und man berechnet  $\langle x' | M_k x' \rangle = \langle x | M x \rangle \geq 0$ ; aus  $\langle x' | M_k x' \rangle = 0$  folgt  $\langle x | M x \rangle = 0$ , so daß  $x$  und damit  $x'$  verschwinden,  $M_k$  also positiv definit ist. Somit sind alle Eigenwerte der Matrix  $M_k$  und mit ihnen die Determinante als ihr Produkt positiv.

Für den Umkehrschluß seien nun alle Hauptminoren von  $M$  positiv. Der Beweis erfolgt mittels Induktion über  $k \in \{1, \dots, n\}$ . Die erste Untermatrix  $M_1 = (m_{11})$  ist wegen  $\det M_1 = m_{11} > 0$  offensichtlich positiv definit. Vorausgesetzt werde nun, daß die  $(n - 1)$ -te Untermatrix positiv definit ist.

Da die Matrix  $M_{n-1}$  positiv definit ist, gibt es eine invertierbare Matrix  $A \in \mathbb{C}^{(n-1) \times (n-1)}$ , für die  $M_{n-1} = A^\dagger A$  gilt. Im folgenden werden alle  $n \times n$ -Matrizen in Blockmatrixform mit einem Block der Größe  $n - 1$  und einem der Größe Eins geschrieben. Setzt man also  $B := \text{diag}(A, 1) \in \mathbb{C}^{n \times n}$  und verwendet die Abkürzungen  $v := (m_{1n}, \dots, m_{n-1,n})^t \in \mathbb{C}^{n-1}$ ,  $x := m_{nn} \in \mathbb{R}$  und  $c = (c_1, \dots, c_{n-1})^t := (A^\dagger)^{-1}v \in \mathbb{C}^{n-1}$ , so berechnet sich

$$\begin{aligned} C &:= (B^\dagger)^{-1} M_n B^{-1} = \begin{pmatrix} (A^\dagger)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_{n-1} & v \\ v^\dagger & x \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} (A^\dagger)^{-1} M_{n-1} A^{-1} & (A^\dagger)^{-1} v \\ v^\dagger A^{-1} & x \end{pmatrix} = \begin{pmatrix} \mathbb{I}_{n-1} & c \\ c^\dagger & x \end{pmatrix} \in \mathbb{C}^{n \times n}. \end{aligned} \quad (\text{B.11})$$

---

dratischen Untermatrix, als Minoren; im Englischen hat *minor* stets diese Bedeutung. Minoren in dem in dieser Arbeit verwendeten Sinne heißen im Englischen *principal minors*, die Hauptminoren *leading principal minors*.

<sup>22</sup>Der hier angegebene Beweis des Lemmas B.19 folgt dem unveröffentlichten *Skript zur Vorlesung „Lineare Algebra für Physiker“* von Prof. HELMUT MÄURER (TH Darmstadt).

Subtrahiert man von der letzten Zeile der Matrix  $C$  nacheinander das  $c_i^*$ -fache der  $i$ -ten Zeile für alle  $i \in \{1, \dots, n-1\}$ , so entsteht die obere Dreiecksmatrix

$$C' = \begin{pmatrix} \mathbb{1}_{n-1} & c \\ 0 & m \end{pmatrix} \quad (\text{B.12})$$

mit  $m := x - \sum_{i=1}^{n-1} |c_i|^2$ , die die gleiche Determinante wie  $C$  besitzt. Es ist daher  $m = \det C' = \det C = \det M_n > 0$ . Setzt man nun

$$F := \begin{pmatrix} \mathbb{1}_{n-1} & c \\ 0 & \sqrt{m} \end{pmatrix}, \quad (\text{B.13})$$

so ist  $F$  invertierbar und  $C = F^\dagger F$  damit ebenfalls, also positiv definit. Mit der Invertierbarkeit von  $B$  folgt, daß auch  $FB$  invertierbar, die Matrix  $M_n = B^\dagger C B = B^\dagger F^\dagger F B = (FB)^\dagger (FB)$  mithin positiv definit ist, q. e. d.

Man könnte nun vermuten, daß eine Matrix, deren sämtliche Hauptminoren nicht-negativ sind, positiv semidefinit ist. Dies ist aber falsch, was man zum Beispiel an der Matrix  $M := \text{diag}(0, -1) \in \mathbb{C}^{2 \times 2}$  erkennt. Es zeigt sich aber, daß eine Matrix positiv semidefinit ist, wenn *alle* Minoren nicht-negativ sind. Als Hilfsmittel zum Beweis dient das folgende Lemma.

**Lemma B.20 (Koeffizienten des charakteristischen Polynoms)**

Das charakteristische Polynom  $\chi_A = \det(\lambda \mathbb{1} - A)$  einer Matrix  $A \in \mathbb{C}^{n \times n}$  besitzt die Form  $\chi_A(\lambda) = \sum_{k=0}^n a_k \lambda^k$  mit dem Leitkoeffizienten  $a_n = 1$ . Der Koeffizient  $a_k$  ist das  $(-1)^{n-k}$ -fache der Summe über alle Minoren  $(n-k)$ -ter Ordnung; insbesondere sind  $a_{n-1} = -\text{Spur } A$  und  $a_0 = (-1)^n \det A$ .

*Beweis:* Es sei  $A = (a_1, \dots, a_n)$ , wobei  $a_i$  den  $i$ -ten Spaltenvektor bezeichne und  $e_i$  der  $i$ -te Einheitsvektor sei; das charakteristische Polynom ist dann  $\det(\lambda e_1 - a_1, \dots, \lambda e_n - a_n)$ . Für eine Teilmenge  $M \subseteq \{1, \dots, n\}$  bezeichne nun  $m_M$  die Determinante derjenigen Matrix, in deren  $i$ -ter Spalte  $a_i$  steht, wenn  $i \in M$  ist, und andernfalls  $e_i$ ; wertet man das charakteristische Polynom aus, indem man verwendet, daß die Determinante in jeder Spalte linear ist, so erhält man

$$\det(\lambda \mathbb{1} - A) = \sum_{M \subseteq \{1, \dots, n\}} (-1)^{|M|} \cdot \lambda^{n-|M|} \cdot m_M. \quad (\text{B.14})$$

Für jedes  $k \in \{0, \dots, n\}$  ist der Koeffizient vor  $\lambda^k$  also das  $(-1)^{n-k}$ -fache der Summe aller  $m_M$ , wenn über alle  $(n-k)$ -elementigen Teilmengen  $M$  von  $\{1, \dots, n\}$  summiert wird. Es genügt also zu zeigen, daß  $m_M$  der zur Menge  $M$  gehörende Minor ist.

Betrachtet man zunächst  $M = \{1, \dots, k\}$ , so folgt die Behauptung, indem man die Determinante  $m_M$  nacheinander nach der letzten, die übrigbleibende

Determinante nach der vorletzten usw. bis zur  $k + 1$ -ten Spalte entwickelt. Die übrigen  $m_M$  ändern ihren Wert nicht, wenn man die zugrundeliegende Matrix durch gemeinsames Vertauschen von Zeilen und Spalten auf die vorgenannte Form bringt, so daß die  $m_M$  Minoren der Matrix sind, q. e. d.

Es wird nun das HURWITZ-Kriterium für positiv semidefinite Matrizen formuliert und mithilfe des soeben gezeigten Lemmas auf das HURWITZ-Kriterium für positiv definite Matrizen zurückgeführt.

**Satz B.21 (Kriterium für positiv semidefinite Matrizen)**

*Eine hermitesche Matrix ist genau dann positiv semidefinit, wenn alle ihre Minoren nicht-negativ sind.*

*Beweis:* Ist eine Matrix positiv semidefinit, so ergibt sich die Behauptung entsprechend wie im Beweis des Lemmas B.19 aus der Tatsache, daß ein gemeinsames Umsortieren der Zeilen- und Spaltenanordnung die Semidefinitheit unverändert läßt.

Für die Umkehrung beachte man, daß eine Matrix  $A \in \mathbb{C}^{n \times n}$  genau dann positiv semidefinit ist, wenn die Matrizen  $A + \varepsilon \mathbb{I}$  für jedes  $\varepsilon > 0$  positiv definit sind; diese Eigenschaft kann mithilfe des Lemmas B.19 überprüft werden. Bezeichnet man mit  $S_k$  die Summe aller Minoren  $k$ -ter Ordnung von  $A$ , so berechnet man mithilfe des Lemmas B.20 nun

$$\begin{aligned} \det(A + \varepsilon \mathbb{I}) &= (-1)^n \det(-\varepsilon \mathbb{I} - A) \\ &= (-1)^n \sum_{k=0}^n (-1)^{n-k} S_{n-k} (-\varepsilon)^k = \sum_{k=0}^n S_{n-k} \varepsilon^k. \end{aligned} \quad (\text{B.15})$$

Nach Voraussetzung sind alle Minoren nicht-negativ, ihre Summen  $S_k$  also auch. Verschwindet also ein Hauptminor, ist zum Beispiel  $\det A = S_n = 0$ , so ist  $\det(A + \varepsilon \mathbb{I})$  für jedes noch so kleine  $\varepsilon > 0$  positiv. Sinngemäß gilt dies für alle Hauptminoren, die Matrix  $A + \varepsilon \mathbb{I}$  ist also positiv definit, q. e. d.

Aus den beiden HURWITZ-Kriterien erhält man leicht zwei weitere Ergebnisse, von denen im Kapitel 6 Gebrauch gemacht wird.

**Korollar B.22 (Blockmatrizen)**

*Ist eine aus Blockmatrizen bestehende Matrix  $M = \begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix}$  positiv definit oder semidefinit, so auch  $M' = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \cong A \oplus C$ .*

*Beweis:* Dies folgt aus den beiden HURWITZ-Kriterien, da mit  $M$  auch die Untermatrizen  $A$  und  $C$  positiv definit oder semidefinit sind und sich die Determinanten faktorisieren lassen, q. e. d.

**Korollar B.23 (Rang und Minoren einer Matrix)**

*Besitzt eine Matrix  $A \in \mathbb{C}^{n \times n}$  den Rang  $r \in \{0, \dots, n\}$ , so verschwinden alle Minoren  $(r + 1)$ -ter und höherer Ordnung.*

*Beweis:* Nach Voraussetzung gibt es keine  $r + 1$  Spalten von  $A$ , die linear unabhängig sind. Entsprechend sind auch die Spalten der Untermatrix zu einem Minor  $(r + 1)$ -ter oder höherer Ordnung linear abhängig, so daß ihre Determinante verschwindet, q. e. d.

## B.6 Tensorprodukte

In diesem Abschnitt wird das Tensorprodukt von Hilberträumen erläutert, das für die Beschreibung zusammengesetzter Quantensysteme und somit der Verschränkung von herausragender Wichtigkeit ist. Im Gegensatz dazu ist die *direkte Summe*  $\mathcal{H}_1 \oplus \mathcal{H}_2$  zweier Hilberträume  $\mathcal{H}_1$  und  $\mathcal{H}_2$  unbedeutend.

Das *Tensorprodukt* (auch *direktes Produkt* genannt) ist eine Möglichkeit, aus zwei Hilberträumen  $\mathcal{H}_1$  und  $\mathcal{H}_2$  einen größeren Hilbertraum  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$  zu konstruieren. Während das Tensorprodukt (von Hilberträumen und auch von anderen algebraischen Strukturen) in der Mathematik in der Regel durch eine sogenannte *universelle Eigenschaft* eingeführt wird, erweist sich für die Zwecke der Physik eine direktere Definition als einfacher verständlich; diese ist auch unter dem Namen *KRONECKER-Produkt* bekannt.

Im folgenden wird nur das Tensorprodukt zweier endlichdimensionaler komplexer Hilberträume behandelt, die als  $\mathcal{H}_A := \mathbb{C}^n$  und  $\mathcal{H}_B := \mathbb{C}^m$  für  $n, m \in \mathbb{N}$  angenommen werden können und die mit dem kanonischen Skalarprodukt versehen seien. Die Verallgemeinerung der Konstruktion auf beliebige Hilberträume ist ohne konzeptionelle Probleme möglich, ebenso die Bildung von Tensorprodukten aus mehr als zwei Tensorfaktoren, da die Produktbildung als Abbildung auf der Menge der Hilberträume (bis auf Isomorphie) assoziativ und kommutativ ist.

### B.6.1 Grundlegende Definitionen

In diesem Abschnitt sei eine Basis stets eine geordnete Menge von Vektoren; alle Vektoren und Matrizen werden bezüglich dieser Anordnung gebildet. Für die kanonische Basis  $\{e_1, e_2, \dots, e_n\}$  des Hilbertraums  $\mathbb{C}^n$  werde in DIRAC-Notation  $|i\rangle := e_{i+1}$  für  $i \in \{0, \dots, n - 1\}$  gesetzt, um Operationen auf dem Restklassenring  $\mathbb{Z}/n\mathbb{Z}$  (also Operationen modulo  $n$ ) formulieren zu können.

Die betrachteten Hilberträume  $\mathbb{C}^n$  und  $\mathbb{C}^m$  besitzen also die kanonischen Basen  $\{|0\rangle_A, \dots, |n - 1\rangle_A\}$  und  $\{|0\rangle_B, \dots, |m - 1\rangle_B\}$ , wobei die Räume durch tiefgestellte Indizes angedeutet werden. Das Tensorprodukt  $\mathcal{H}_A \otimes \mathcal{H}_B$  dieser beiden Räume ist nun definiert als ein Hilbertraum, der eine Orthonormalbasis

$$\{|i\rangle_A \otimes |j\rangle_B \mid i \in \{0, \dots, n - 1\}, j \in \{0, \dots, m - 1\}\} \quad (\text{B.16})$$

besitzen möge; da sie aus separablen Zuständen besteht, werde sie im weiteren die *separable Basis* genannt. Der Tensorproduktraum besitzt also die Dimension  $n \cdot m$ , ist also zum Raum  $\mathbb{C}^{nm}$  isomorph. Für die Basisvektoren des Produktraums ist eine Vielzahl von Schreibweisen gebräuchlich, die allesamt selbsterklärend sind, zum Beispiel

$$|ij\rangle_{AB} := |i, j\rangle_{AB} := |i\rangle_A |j\rangle_B := |i\rangle_A \otimes |j\rangle_B. \quad (\text{B.17})$$

Legt man die Reihenfolge der Faktoren ein für allemal fest, so kann man auch die Indizes  $A$  und  $B$  unterdrücken.

Für die *explizite Matrixdarstellung von Tensorprodukten* muß man noch die Reihenfolge festlegen, in der man die Produktvektoren  $|i, j\rangle$  mit der kanonischen Basis des  $\mathbb{C}^{nm}$  identifizieren will. Es bietet sich an, sie gewissermaßen „numerisch aufsteigend“ anzuordnen, also in der Reihenfolge

$$|0, 0\rangle, |0, 1\rangle, \dots, |0, m-1\rangle, |1, 0\rangle, |1, 1\rangle, \dots, |1, m-1\rangle, \dots \quad (\text{B.18})$$

Im Falle von mehr als zwei Tensorfaktoren kann diese Sortierung sinngemäß verallgemeinert werden. Man erkennt hieran, daß die Anordnung der Tensorfaktoren unbedeutend ist; die Bildung des Tensorproduktes ist also bis auf Isomorphie assoziativ und kommutativ.

## B.6.2 Abbildung in den Produktraum, Skalarprodukt

Je einem Paar von Vektoren aus  $\mathbb{C}^n$  und  $\mathbb{C}^m$  wird mittels einer Abbildung  $\otimes : \mathbb{C}^n \times \mathbb{C}^m \rightarrow \mathbb{C}^n \otimes \mathbb{C}^m$  ein Vektor im Tensorproduktraum zugeordnet; sie ist durch

$$\left( \sum_{i=0}^{n-1} v_i |i\rangle_A \right) \otimes \left( \sum_{j=0}^{m-1} w_j |j\rangle_B \right) \mapsto \sum_{i,j} v_i w_j |i\rangle_A \otimes |j\rangle_B \quad (\text{B.19})$$

definiert. Aus der Linearität des Skalarproduktes und der Forderung, daß die separable Basis eine Orthonormalbasis ist, erhält man für das Skalarprodukt zweier Vektoren des Tensorproduktraums

$$\left( \sum_{ij} \alpha_{ij} \langle ij| \right) \left( \sum_{kl} \beta_{kl} |kl\rangle \right) = \sum_{i,j,k,l} \alpha_{ij}^* \beta_{kl} \underbrace{\langle ij|kl\rangle}_{=\delta_{ik}\delta_{jl}} = \sum_{i,j} \alpha_{ij}^* \beta_{ij},$$

mithin also  $\langle \psi_1 \otimes \varphi_1 | \psi_2 \otimes \varphi_2 \rangle = \langle \psi_1 | \psi_2 \rangle \cdot \langle \varphi_1 | \varphi_2 \rangle$  für beliebige Vektoren  $|\psi_1\rangle, |\psi_2\rangle \in \mathbb{C}^n$  und  $|\varphi_1\rangle, |\varphi_2\rangle \in \mathbb{C}^m$ .

Eine Matrix auf dem Tensorproduktraum kann durch  $C = (c_{ij,kl}) \in \mathbb{C}^{nm \times nm}$  dargestellt werden, die dann die Form

$$C = \begin{pmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{pmatrix} \quad \text{mit } C_{ik} = \begin{pmatrix} c_{i1,k1} & \cdots & c_{i1,km} \\ \vdots & \ddots & \vdots \\ c_{im,k1} & \cdots & c_{im,km} \end{pmatrix} \in M_m(\mathbb{C}) \quad (\text{B.20})$$

hat. Ähnlich wie für Vektoren, nur in „zwei Richtungen“ bildet man das Tensorprodukt zweier Matrizen  $A \in \mathbb{C}^{n \times n}$  und  $B \in \mathbb{C}^{m \times m}$  mittels

$$A \otimes B = (a_{ik})_{ik} \otimes (b_{jl})_{jl} := (a_{ik} \cdot b_{jl})_{ijkl}, \quad (\text{B.21})$$

das heißt, es ist  $C_{ik} = a_{ik}B$ . Es folgt dann unmittelbar  $(A \otimes B)(|\psi\rangle \otimes |\varphi\rangle) = (A|\psi\rangle) \otimes (B|\varphi\rangle)$  für alle  $|\psi\rangle \in \mathbb{C}^n$  und alle  $|\varphi\rangle \in \mathbb{C}^m$  und  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$  für alle  $C \in \mathbb{C}^{n \times n}$  und alle  $D \in \mathbb{C}^{m \times m}$ .

### B.6.3 Operationen auf einem Tensorfaktor

Eine wichtige Transformation für Operatoren auf dem Tensorproduktraum ist die *Teiltransposition* oder *partielle Transposition*. Sie ist für eine Matrix  $C = (c_{ij,kl}) \in \mathbb{C}^{nm \times nm}$ , die die in Formel B.20 beschriebene Struktur habe, durch  $C^{TA} := (c_{kj,il})_{ijkl}$  (Transposition bezüglich des ersten Tensorfaktors  $\mathcal{H}_A = \mathbb{C}^n$ ) bzw. durch  $C^{TB} := (c_{il,kj})_{ijkl}$  (Transposition bezüglich des zweiten Tensorfaktors  $\mathcal{H}_B = \mathbb{C}^m$ ) definiert; im ersten Fall wird bildlich gesprochen, die äußere Matrix transponiert, das heißt,  $C^{TA} = (C_{ki})_{ik}$ , im zweiten Fall alle inneren Matrizen, also  $C^{TB} = (C_{ik}^t)_{ik}$ . Ausgeschrieben lautet dies

$$C^{TA} = \begin{pmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{pmatrix} \quad \text{und} \quad C^{TB} = \begin{pmatrix} C_{11}^t & \cdots & C_{1n}^t \\ \vdots & \ddots & \vdots \\ C_{n1}^t & \cdots & C_{nn}^t \end{pmatrix}, \quad (\text{B.22})$$

und offensichtlich gilt  $C^t = C^{TA}T_B = C^{TB}T_A$ .

### B.6.4 Teilspur und reduzierte Operatoren

Ist man statt am Gesamtsystem nur an einem der Untersysteme interessiert, so verwendet man statt eines auf dem Gesamtsystem definierten Operators dessen *Teilspur* oder *partielle Spur*. Für  $C = (c_{ij,kl}) \in M_{nm}(\mathbb{C})$  bestimmen sich die *reduzierten Operatoren* zu  $C_A := \text{Spur}_B C$  und  $C_B := \text{Spur}_A C$  mit

$$\text{Spur}_A C := \sum_{i=1}^n C_{ii} \in \mathbb{C}^{m \times m}, \quad (\text{B.23})$$

$$(\text{Spur}_B C)_{ik} := \sum_{j=1}^m c_{ij,kj} = (\text{Spur } C_{ik})_{ik} \in \mathbb{C}^{n \times n}. \quad (\text{B.24})$$

Es gilt  $\text{Spur } C = \text{Spur}_A \text{ Spur}_B C = \text{Spur}_B \text{ Spur}_A C$ . Man kann ähnlich auch *partielle Skalarprodukte* definieren.

### B.6.5 Die Schmidt-Zerlegung

Ein in der Quanteninformationstheorie allgemein sehr bedeutendes Hilfsmittel ist die *SCHMIDT-Zerlegung*.<sup>23</sup> In der Physik wird zumeist die endlichdimensionale Fassung verwendet, aber schon die Originalarbeit enthält eine unendlichdimensionale Fassung. Einen einfachen Beweis findet man etwa bei NIELSEN UND CHUANG [127].

#### Satz B.24 (Schmidt-Zerlegung)

Für jeden nicht-verschwindenden Vektor  $v \in \mathcal{H}_A \otimes \mathcal{H}_B$  existieren eine Zahl  $s \in \{1, \dots, \min\{n, m\}\}$ , Zahlen  $\lambda_1, \dots, \lambda_s \in \mathbb{R}^+$  und Orthonormalsysteme  $(e_i^A)_{i=1}^s$  und  $(e_i^B)_{i=1}^s$  in  $\mathcal{H}_A$  bzw.  $\mathcal{H}_B$  derart, daß  $v = \sum_{i=1}^s \lambda_i e_i^A \otimes e_i^B$  ist.

Man nennt  $s$  die *SCHMIDT-Zahl* des Vektors  $v$ ; daß sie eindeutig bestimmt ist, ergibt sich aus der im Beweis verwendeten Singulärwertzerlegung. Eine Verallgemeinerung der SCHMIDT-Zerlegung auf Systeme mit mehr als zwei Teilsystemen ist nicht allgemein möglich.

*Beweis:* Wählt man Orthonormalbasen  $(f_j^A)_{j=1}^n$  und  $(f_k^B)_{k=1}^m$  von  $\mathbb{C}^n$  und  $\mathbb{C}^m$  aus, so läßt sich  $v = \sum_{jk} a_{jk} f_j^A \otimes f_k^B$  schreiben. Nach der Singulärwertzerlegung (Korollar B.42) läßt sich die Koeffizientenmatrix  $A = (a_{jk}) \in \mathbb{C}^{n \times m}$  in der Form  $A = UDV$  schreiben, wobei die Matrizen  $U = (u_{ji})_{ji} \in \mathbb{C}^{n \times n}$  und  $V = (v_{ik})_{ik} \in \mathbb{C}^{m \times m}$  unitär sind und  $D = (\lambda_i \delta_{ii'})_{ii'} \in \mathbb{C}^{n \times m}$  gilt. Aus  $a_{jk} = \sum_i u_{ji} \lambda_i v_{ik}$  folgt daher

$$v = \sum_{ijk} u_{ji} \lambda_i v_{ik} f_j^A \otimes f_k^B = \sum_i \lambda_i \left( \sum_j u_{ji} f_j^A \right) \otimes \left( \sum_k v_{ik} f_k^B \right). \quad (\text{B.25})$$

Ordnet man die Komponenten so, daß genau die Werte  $\lambda_i$  für  $i \in \{1, \dots, s\}$  nicht verschwinden, so erhält man die gesuchte Zerlegung, indem man für diese Indizes  $e_i^A := \sum_j u_{ji} f_j^A$  und  $e_i^B := \sum_k v_{ik} f_k^B$  setzt, q. e. d.

Aus der Singulärwertzerlegung folgt, daß die Zahl  $s$  und die Werte  $\lambda_1, \dots, \lambda_s$  mit ihrer Vielfachheit eindeutig bestimmt sind. Gruppiert man die SCHMIDT-Zerlegung nach unterschiedlichen Zahlenwerten der  $\lambda_i$ , so hat sie die Form

$$v = \sum_{i=1}^n \lambda_i \left( \sum_{j=1}^{m(i)} e_{i,j}^A \otimes e_{i,j}^B \right), \quad (\text{B.26})$$

---

<sup>23</sup>Vergleiche die Originalarbeit von SCHMIDT [159], Gleichung (34) auf S. 466, sowie das Buch von COURANT UND HILBERT [43], Band I, § 10.8 auf S. 137. Im physikalischen Kontext findet sie sich unabhängig bei SCHRÖDINGER [162].

wobei  $\sum_{i=1}^n m(i) = s$  ist die  $\lambda_i$  nun paarweise verschieden sind. Damit der Vektor  $v$  sich nicht ändert, müssen also  $\sum_{j=1}^{m(i)} e_{i,j}^A \otimes e_{i,j}^B$  für  $i \in \{1, \dots, n\}$  invariant sein, was bedeutet, daß unitäre Transformationen, die dies tun, die direkte Summe von unitären Transformationen  $U \otimes U^*$  auf den bezeichneten Unterräumen sind; vgl. das folgende Lemma.

**Lemma B.25 (Invarianz bestimmter Zustände)**

Der Zustand  $|\Psi_{00}\rangle = d^{-1/2} \sum_{k=0}^{d-1} |k\rangle_A |k\rangle_B \in \mathbb{C}^d \otimes \mathbb{C}^d$  ist unter einem Produkt unitärer Transformationen der Form  $U \otimes V$  genau dann invariant, wenn  $V = U^*$  ist.

*Beweis:* Es seien  $U = \sum_{ij} u_{ij} |i\rangle\langle j| \in \mathbb{C}^{d \times d}$  und  $V = \sum_{pq} v_{pq} |p\rangle\langle q| \in \mathbb{C}^{d \times d}$  unitäre Matrizen; man berechnet hiermit

$$\begin{aligned} (U \otimes V)|\Psi_{00}\rangle &= d^{-1/2} \sum_{ijpqk} u_{ij} v_{pq} |i\rangle_A |j\rangle_B \otimes |p\rangle_B |q\rangle_A \\ &= d^{-1/2} \sum_{ip} \left[ \sum_k u_{ik} v_{pk} \right] |i\rangle_A \otimes |p\rangle_B. \end{aligned} \quad (\text{B.27})$$

Der Zustand  $|\Psi_{00}\rangle$  ist also genau dann invariant, wenn  $\sum_k u_{ik} v_{pk} = \delta_{ip}$  gilt. Der  $p$ -te Zeilenvektor von  $V$  ist also orthogonal zu allen Zeilenvektoren von  $U^*$  mit Ausnahme des  $p$ -ten; da die unitäre Matrix  $U^*$  vollen Rang besitzt und ihre Zeilenvektoren orthogonal aufeinander stehen, ist also der  $p$ -te Zeilenvektor von  $V$  ein Vielfaches des  $p$ -ten Spaltenvektors von  $U^*$ . Als Zeilenvektor einer unitären Matrix ist er normiert und seine Phase ist durch die Orthogonalitätsbedingung festgelegt; dies gilt für alle Zeilenvektoren der Matrix  $V$ , und somit ist  $V = U^*$ , q. e. d.

Mithilfe der SCHMIDT-Zerlegung kann man leicht zeigen, daß zu jedem gemischten Zustand eine *Purifizierung* existiert.<sup>24</sup>

**Satz B.26 (Purifizierung von Zuständen)**

Für jede Dichtematrix  $\rho_A$  auf einem Hilbertraum  $\mathcal{H}_A$  existiert ein Hilbertraum  $\mathcal{H}_B$  sowie ein reiner Zustand  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  derart, daß  $\rho_A = \text{Spur}_B |\Psi\rangle\langle\Psi|$  ist; dieser ist bis auf isometrische Transformationen des Hilfssystems eindeutig bestimmt.

*Beweis:* Ist  $\rho_A = \sum_{i=1}^s \lambda_i^2 |i\rangle_A \langle i|$  eine Spektralzerlegung von  $\rho_A$ , wobei die  $\lambda_i$  nicht-negativ gewählt seien, und ist  $(|i\rangle_B)_{i=1}^s$  ein Orthonormalsystem in einem Hilbertraum  $\mathcal{H}_B$ , so ist

$$|\Psi\rangle = \sum_{i=1}^s \lambda_i |i\rangle_A |i\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B. \quad (\text{B.28})$$

---

<sup>24</sup>Diese Purifizierung ist nicht mit dem Begriff der *entanglement purification* verwandt.

eine mögliche Purifizierung, die in ihrer SCHMIDT-Zerlegung angegeben ist. Ist nun  $|\Psi'\rangle = \sum_{j=1}^t \mu_j |j'\rangle_A |j'\rangle_{B'} \in \mathcal{H}_A \otimes \mathcal{H}_{B'}$  eine weitere Purifizierung von  $\rho_A$ , so folgt aus der Eindeutigkeit der Spektralzerlegung, daß wegen  $s = \text{Rang } \rho_A$  auch  $s = t$  sein muß und bis auf Umsortierung alle  $\lambda_i$  und  $\mu_j$  zusammenfallen. Das Orthonormalsystem von  $\mathcal{H}_B$  kann isometrisch auf eines von  $\mathcal{H}_{B'}$  abgebildet werden, q. e. d.

## B.7 Darstellungen endlicher Gruppen

Im folgenden sei  $V$  stets ein endlichdimensionaler komplexer Vektorraum, der mit  $\mathbb{C}^n$  für  $n = \dim V$  identifiziert werde. Seine Automorphismengruppe ist die *allgemeine lineare Gruppe*  $\text{GL}(V)$  (für engl. *general linear group*), die mit den invertierbaren Matrizen identifiziert werden soll; es ist also

$$\text{GL}(V) := \{A \in \mathbb{C}^{n \times n} \mid \det A \neq 0\}. \quad (\text{B.29})$$

Eine Darstellung wird nun wie folgt definiert.<sup>25</sup>

### Definition B.27 (Darstellungen endlicher Gruppen)

Als Darstellung einer endlichen Gruppe  $G$  bezeichnet man einen Gruppenhomomorphismus  $\rho : G \rightarrow \text{GL}(V)$ ; es gilt also  $\rho(gh) = \rho(g)\rho(h)$  für alle  $g, h \in G$ . Man nennt  $V$  den Darstellungsraum und seine Dimension den Grad der Darstellung; die Darstellung heißt *treu*, falls sie injektiv ist.<sup>26</sup>

Eine verwandter Begriff ist der der *Strahl-* oder *projektiven Darstellung*, der einen Homomorphismus auf  $\text{GL}(V)/\mathbb{C}^\times$  bezeichnet; hierbei ist  $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$  die Einheitengruppe der komplexen Zahlen. Es handelt sich also um eine Darstellung „bis auf einen komplexen Faktor“, der oft den Betrag Eins hat.

Die triviale Darstellung einer Gruppe  $G$  ist durch  $\rho : G \rightarrow \mathbb{C} \setminus \{0\}$  mittels  $\rho(g) = 1$  für alle  $g \in G$  definiert. Für die (*links-*)*reguläre Darstellung* einer Gruppe  $G$  sei  $V = \mathbb{C}^{|G|}$ , und mit einer Orthonormalbasis  $(e_g)_{g \in G}$  fordert man  $\rho(g)e_h := e_{gh}$  für alle  $g, h \in G$ . Weiterhin ist  $\rho(g^{-1}) = \rho(g)^{-1}$  für alle  $g \in G$ .

---

<sup>25</sup>Einzelheiten zur Darstellungstheorie finden sich zum Beispiel bei BOERNER [19], bei FULTON UND HARRIS [54] oder auch bei SERRE [164]. Ein großer Teil der Ergebnisse läßt sich mit geringen Anpassungen auf unendliche, aber kompakte Gruppen erweitern. Eine Darstellung ist dann ein *stetiger* Homomorphismus, der Raum  $V$  unverändert endlichdimensional, und man ersetzt den Ausdruck  $|G|^{-1} \sum_{g \in G} \dots$  durch die Integration über das normierte HAAR-Maß.

<sup>26</sup>Die englischen Begriffe sind *representation*, *representation space*, *degree* und *faithful*.

**Definition B.28 (Äquivalente Darstellungen)**

Sind  $\rho_1 : G \rightarrow \text{GL}(V)$  und  $\rho_W : G \rightarrow \text{GL}(W)$  Darstellungen, so nennt man eine Funktion  $\varphi : V \rightarrow W$  einen Homomorphismus von Darstellungen, wenn  $\varphi \circ \rho_V(g) = \rho_W(g) \circ \varphi : V \rightarrow W$  für alle  $g \in G$  gilt. Falls ein solcher Homomorphismus ungleich Null existiert, so nennt man die Darstellungen äquivalent (auch ähnlich oder isomorph).

Dieser Begriff ist stärker als der des Vektorraum-Homomorphismus, der nur die linearen Abbildungen von  $V$  nach  $W$  bezeichnet. Man nennt eine Darstellung *unitär*, falls für alle  $g \in G$  die Abbildung  $\rho(g)$  als Matrix unitär ist, und kann zeigen, daß jede Darstellung einer endlichen Gruppe auf einem Hilbertraum zu einer unitären Darstellung äquivalent ist.<sup>27</sup>

**B.7.1 Summen und Produkte von Darstellungen**

Es ist möglich, aus zwei Darstellungen einer Gruppe neue Darstellungen zu bilden; dies geschieht mittels direkter Summen und Tensorprodukten. Liegen hierzu zwei Darstellungen  $\rho_1 : G \rightarrow \text{GL}(V_1)$  und  $\rho_2 : G \rightarrow \text{GL}(V_2)$  derselben Gruppe  $G$  vor, so definieren die *direkte Summe*  $\rho_1 \oplus \rho_2 : G \rightarrow \text{GL}(V_1 \times V_2)$ ,  $\rho(g) := \rho_1(g) \oplus \rho_2(g)$  und das *Tensorprodukt*<sup>28</sup>  $\rho_1 \otimes \rho_2 : G \rightarrow \text{GL}(V_1 \otimes V_2)$ ,  $\rho(g) := \rho_1(g) \otimes \rho_2(g)$  Darstellungen von  $G$ .

**B.7.2 Irreduzible Darstellungen**

Betrachtet man einen Vektorraum  $V$  und zwei lineare Teilräume  $U_1, U_2 \subseteq V$ , so nennt man  $V$  die (*innere*) *direkte Summe* von  $U_1$  und  $U_2$ ,  $V = U_1 \oplus U_2$ , falls sowohl  $U_1 \cap U_2 = \{0\}$  als auch  $U_1 + U_2 := \{x + y \mid x \in U_1, y \in U_2\} = V$  gilt. Dies bedeutet, daß jedes Element aus  $V$  eindeutig in Elemente aus  $U_1$  und  $U_2$  zerlegt werden kann; es gilt somit  $V \cong U_1 \times U_2$ .

**Definition B.29 (Irreduzible Darstellungen)**

Einen linearen Teilraum  $U \subseteq V$  nennt man invariant unter einer Darstellung  $\rho : G \rightarrow \text{GL}(V)$ , falls  $\rho(G)U \subseteq U$  gilt. In diesem Fall erhält man mittels  $\rho_U : G \rightarrow \text{GL}(U)$ ,  $\rho_U(g) := \rho(g)|_U$  eine Unterdarstellung von  $\rho$ . Eine Darstellung  $\rho : G \rightarrow \text{GL}(V)$  nennt man irreduzibel, wenn  $\{0\}$  und  $V$  ihre einzigen invarianten Teilräume sind; andernfalls nennt man sie reduzibel.

<sup>27</sup>Ist  $G$  eine endliche Gruppe der Ordnung  $m$ , so gilt nach dem Satz von EULER-FERMAT ( $\forall g \in G)(g^m = 1)$ . Bildet man die Potenzen eines JORDAN-Blocks (vgl. Unterabschnitt B.4.5), so erkennt man, daß  $\varphi(g)$  notwendig normal und stärker noch unitär sein muß.

<sup>28</sup>Hiervon zu unterscheiden ist die folgende Konstruktion: Sind  $\rho_1 : G_1 \rightarrow \text{GL}(V_1)$  und  $\rho_2 : G_2 \rightarrow \text{GL}(V_2)$  Darstellungen zweier Gruppen  $G_1$  und  $G_2$ , so definiert  $\rho_1 \otimes \rho_2 : G_1 \times G_2 \rightarrow \text{GL}(V_1 \otimes V_2)$ ,  $(\rho_1 \otimes \rho_2)(g_1, g_2) := \rho_1(g_1) \otimes \rho_2(g_2)$  eine Darstellung von  $G_1 \times G_2$ , die auch als das *Tensorprodukt* bezeichnet wird. Sind  $\rho_1$  und  $\rho_2$  irreduzibel, so auch  $\rho_1 \otimes \rho_2$ , und jede irreduzible Darstellung von  $G_1 \times G_2$  ist von dieser Form.

Nach dem *Satz von MASCHKE* gibt es zu einem invarianten Teilraum  $U \subseteq V$  stets ein Vektorraumkomplement  $U'$  – d. h.,  $U \oplus U' = V$  und  $U \cap U' = \emptyset$  –, das selbst invariant ist. Dies bedeutet, daß die Darstellung in die direkte Summe zweier Darstellungen *zerfällt*; anschaulich bedeutet dies, daß die Darstellungsmatrizen durch eine Basistransformation in eine blockdiagonale Form gebracht werden können.<sup>29</sup>

Ist  $V$  ein Hilbertraum mit einem unter  $G$  invarianten Skalarprodukt, also  $(\forall g \in G, \forall x, y \in V)(\langle \rho(g)x | \rho(g)y \rangle = \langle x | y \rangle)$ , so gilt für dieses Skalarprodukt  $\langle x | y \rangle = |G|^{-1} \sum_{g \in G} \langle \rho(g)x | \rho(g)y \rangle$ , und das orthogonale Komplement  $U^\perp$  ist ein geeignetes Komplement. Diese Invarianz des Skalarproduktes ist für unitäre Darstellungen gegeben.

**Lemma B.30 (Invariante Unterräume)**

*Ist  $\varphi : V \rightarrow W$  ein Homomorphismus von Darstellungen, so sind Kern  $\varphi \subseteq V$  und Bild  $\varphi \subseteq W$  invariante Unterräume der jeweiligen Darstellungen.*<sup>30</sup>

*Beweis:* Für jedes  $g \in G$  gilt  $[\varphi \circ \rho_V(g)](\text{Kern } \varphi) = [\rho_W(g) \circ \varphi](\text{Kern } \varphi) = \{0\} \subseteq W$ , also  $\rho_V(g)(\text{Kern } \varphi) \subseteq \text{Kern } \varphi$ ; entsprechend gilt  $\rho_W(g)[\text{Bild } \varphi] = \rho_W(g)[\varphi(V)] = \varphi[\rho_V(g)(V)] = \varphi(V) = \text{Bild } \varphi$ , q. e. d.

**Satz B.31 (Schursches Lemma)**

*Es seien  $\rho_V : G \rightarrow \text{GL}(V)$  und  $\rho_W : G \rightarrow \text{GL}(W)$  irreduzible Darstellungen. Ist  $\varphi : V \rightarrow W$  ein Homomorphismus dieser Darstellungen, so gilt:*

1. *Entweder ist  $\varphi$  ein Isomorphismus oder verschwindet.*
2. *Ist  $V = W$ , so ist  $\varphi = \lambda \cdot \mathbb{1}$  für ein  $\lambda \in \mathbb{C}$ .*<sup>31</sup>

Den Darstellungsraum bezeichnet man daher oft auch kurz als *Darstellung*.

*Beweis:* Ist Kern  $\varphi = V$  oder Bild  $\varphi = \{0\}$ , so verschwindet  $\varphi$  identisch. Ist dies nicht der Fall, so gilt wegen der Irreduzibilität von  $V$  und  $W$  nun Kern  $\varphi = \{0\}$  und Bild  $\varphi = W$ ; die Abbildung  $\varphi$  ist also bijektiv und somit ein Homomorphismus.

<sup>29</sup>Definiert man Darstellungen allgemeiner über einem beliebigen Körper  $K$ , so gilt der Satz nicht für *modulare Darstellungen*, d. h.  $\text{char } K \neq 0$ , wenn die Charakteristik des Körpers die Gruppenordnung teilt.

<sup>30</sup>Für eine Funktion  $f : A \rightarrow B$  zwischen zwei Vektorräumen sind Kern  $f := f^{-1}(\{0\}) = \{x \in A | f(x) = 0\} \subseteq A$  und Bild  $f := f(A) = \{f(x) | x \in A\} \subseteq B$ .

<sup>31</sup>Der zweite Teil des SCHURschen Lemmas gilt nur über algebraisch abgeschlossenen Körpern: als ein Gegenbeispiel betrachte man die Darstellung von  $(\mathbb{Z}/4\mathbb{Z}, \oplus) \rightarrow \text{GL}(\mathbb{R}^2)$ , die durch  $1 + 4\mathbb{Z} \mapsto i\sigma_y \in \text{GL}(\mathbb{R}^2)$  definiert wird. Diese Darstellung ist irreduzibel, denn ein echter invarianter Teilraum wäre eindimensional, also von der Form  $\mathbb{R} \cdot (x, y)^t$  für einen Eigenvektor (!)  $(x, y)^t \in \mathbb{R}^2$  von  $i\sigma$ . Somit wäre  $i\sigma(x, y)^t = (y, -x)^t = \lambda(x, y)^t$  für ein  $\lambda \in \mathbb{R}$ , woraus aber  $x = y = 0$  folgte. Die Matrizen  $A = (a_{ij})_{i,j=1}^2$ , die mit  $i\sigma_y$  und somit mit allen vier Darstellungsmatrizen kommutieren, sind durch  $a_{11} = a_{22}$  und  $a_{12} = -a_{21}$  charakterisiert; dies sind jedoch nicht alles Vielfache der Einheitsmatrix.

Im zweiten Fall ist  $\varphi$  mit allen Darstellungsmatrizen vertauschbar, und dies gilt mit  $\lambda \in \mathbb{C}$  auch für  $\varphi - \lambda\mathbb{I}$ . Da dies nun ein Homomorphismus von Darstellungen ist, ist entweder  $\varphi = \lambda\mathbb{I}$  oder  $\det(\lambda\mathbb{I} - \varphi) \neq 0$ ; der zweite Fall ist nicht für alle  $\lambda \in \mathbb{C}$  möglich, da  $\mathbb{C}$  algebraisch abgeschlossen ist und das nicht-konstante Polynom  $\chi_\lambda = \det(\lambda\mathbb{I} - \varphi)$  mindestens eine Nullstelle haben muß, q. e. d.

### B.7.3 Charaktere von Darstellungen

Als *Charakter* einer Darstellung  $\rho : G \rightarrow \text{GL}(V)$  bezeichnet man die Funktion  $\chi_\rho$  (oder  $\chi_V$ ) :  $G \rightarrow \mathbb{C}$ ,  $\chi_\rho(g) := \text{Spur}[\rho(g)]$ . Der Charakter ist eine *Klassenfunktion* auf  $G$ , das heißt, eine Funktion, die auf den Konjugiertenklassen von  $G$  konstant ist. Man kann zeigen, daß zwei Darstellungen genau dann äquivalent sind, wenn sie denselben Charakter besitzen.

Definiert man auf der Menge der Funktionen  $G \rightarrow \mathbb{C}$  mittels der Festsetzung  $(\varphi, \psi) := |G|^{-1} \sum_{g \in G} \varphi(g)^* \psi(g)$  ein Skalarprodukt, so kann man zeigen, daß die *irreduziblen Charaktere*, d. h., die Charaktere aller nicht äquivalenten irreduziblen Darstellungen, eine Orthonormalbasis auf dem Raum der Klassenfunktionen auf  $G$  bilden; es gibt somit genausoviele nicht-äquivalente irreduzible Charaktere wie Konjugiertenklassen.

Bezeichnet  $[g] := \{h \in G \mid (\exists t \in G)(g = tht^{-1})\}$  die Konjugiertenklasse von  $g \in G$ , so gilt  $\sum \chi_i(g)^* \chi_i(h) = (|G| / |[g]|) \delta_{[g],[h]}$ , wenn über die inäquivalenten irreduziblen Charaktere summiert wird.<sup>32</sup>

#### Satz B.32 (Kanonische Zerlegung von Darstellungen)

Jede Darstellung  $\rho : G \rightarrow \text{GL}(V)$  zerfällt in die direkte Summe irreduzible Darstellungen  $\rho = m_1 \rho_1 \oplus \dots \oplus m_l \rho_l$ . Hierbei ist  $m_i = \langle \rho | \rho_i \rangle$ , und die einzelnen  $\rho_i$  sind inäquivalent. Bezeichnet  $n_i$  den Grad einer Darstellung  $m_i \rho_i$ , so ist  $p_i : V \rightarrow m_i V_i$ ,  $p_i = n_i |G|^{-1} \sum_{g \in G} \chi_{m_i \rho_i}(g)^* \rho(g)$  die zugehörige Projektion.<sup>33</sup>

## B.8 Normierte Vektorräume und Algebren

In der Funktionalanalysis ist eine *Norm* bekanntlich eine Funktion  $\|\cdot\| : V \rightarrow \mathbb{R}_0^+$  von einem Vektorraum  $V$  über einem Skalarkörper  $K$  in die Menge

<sup>32</sup>Eine endliche Gruppe ist genau dann abelsch, wenn all ihre irreduziblen Darstellungen den Grad Eins haben. Ist  $A$  eine abelsche Untergruppe einer endlichen Gruppe  $G$ , so hat jede irreduzible Darstellung von  $G$  höchstens den Grad  $|G|/|A|$ .

<sup>33</sup>Betrachtet man das Tensorprodukt einer Darstellung  $\rho : G \rightarrow \text{GL}(V)$  mit sich selbst, also  $\rho \otimes \rho : G \rightarrow \text{GL}(V \otimes V)$ , und den Automorphismus  $\vartheta(x \otimes y) := y \otimes x$  auf  $V \otimes V$ , so zerfällt  $V \otimes V$  in zwei invariante Teilräume, den *symmetrischen Teil*  $\text{Sym}(V)$  und den *alternierenden Teil*  $\text{Alt}(V)$ ; diese sind durch  $\{z \in V \otimes V \mid \vartheta(z) = \pm z\}$  definiert, besitzen die Dimensionen  $n(n \pm 1)/2$ , und ihre Charaktere sind  $\chi(g) = (\chi(g)^2 \pm \chi(g^2))/2$ .

der nicht-negativen reellen Zahlen, die den folgenden Bedingungen genügt: (1) sie ist *definit*, d. h.  $\|x\| = 0$  impliziert  $x = 0$ , (2) sie ist *homogen*, d. h.  $\|\lambda x\| = |\lambda| \|x\|$  für alle  $\lambda \in K$ , und (3) sie erfüllt die *Dreiecksungleichung*, d. h.  $\|x + y\| \leq \|x\| + \|y\|$ . Verzichtet man auf die erste Bedingung, so spricht man von einer *Halbnorm*.

In den folgenden Unterabschnitten werden die sogenannten  $p$ -Normen auf endlichdimensionalen Vektorräumen eingeführt; sie können mit geringen Anpassungen auf unendlichdimensionale diskrete oder kontinuierliche, endliche oder unendliche Maßräume erweitert werden. Statt Matrizen kann man allgemeiner Elemente einer Algebra betrachten. Hierzu sei auf die Literatur zur Funktionalanalysis verwiesen; vgl. z. B. WERNER [184].

### B.8.1 Normen auf allgemeinen Vektorräumen

Auf den Räumen  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  ist die  $p$ -Norm für  $x = (x_1, \dots, x_n) \in \mathbb{C}^n$  durch

$$\|x\|_p = \begin{cases} (\sum_{i=1}^n |x_i|^p)^{1/p} & \text{für } p \in [1; \infty) \\ \max \{|x_i| \mid i \in \{1, \dots, n\}\} & \text{für } p = \infty \end{cases} \quad (\text{B.30})$$

definiert; für  $p < 1$  gilt eine Art umgekehrte Dreiecksungleichung, es handelt sich also nicht um eine Norm. Für den folgenden Satz vgl. HARDY, LITTLEWOOD UND PÓLYA [70], Theorem 19 auf Seite 28, und HEUSER [74], Teil 1, Aufgabe 59.6 auf S. 351.<sup>34</sup>

#### Satz B.33 (Jensensche Ungleichung)

Für alle  $p, q \in [1; \infty]$  mit  $p < q$  und  $x = (x_1, \dots, x_n) \in \mathbb{C}^n$  gilt  $\|x\|_p \geq \|x\|_q$ ; Gleichheit gilt genau dann, wenn alle  $x_i$  mit höchstens einer Ausnahme verschwinden; ferner ist  $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$ .

*Beweis:* Vom trivialen Fall  $x = 0$  abgesehen kann man sich durch Umskalieren auf den Fall  $\|x\|_p = 1$  beschränken. Es sind dann alle  $|x_i| \leq 1$  und daher  $|x_i|^q \leq |x_i|^p$ , mithin also  $\|x\|_q \leq 1$ . Damit  $\|x\|_q = 1$  ist, muß  $|x_i|^q = |x_i|^p$  für alle  $i$  gelten, also  $x_i \in \{0, 1\}$  sein; wegen der Forderung  $\|x\|_p = 1$  kann nur ein  $x_i \neq 0$  sein.

Für die Aussage über den Grenzwert seien die Einträge ohne Beschränkung der Allgemeinheit reell, positiv und absteigend angeordnet, es gelte also  $x_1 = \dots = x_k > x_{k+1} \geq \dots \geq x_n$ ; man berechnet für die Norm dann  $\|x\|_p = x_1 [k + (x_{k+1}/x_1)^p + \dots + (x_n/x_1)^p]^{1/p}$ . Für  $p \rightarrow \infty$  strebt der Ausdruck in der eckigen Klammer gegen  $k$ , und  $k^{1/p}$  strebt gegen Eins, q. e. d.

---

<sup>34</sup>Die Aussage und ihr Beweis gelten allgemeiner auch dann, wenn  $p$  und  $q$  nur positiv sind, also keine Norm vorliegt.

Umgekehrt zeigt man unmittelbar  $\|x\|_p \leq n^{1/p} \|x\|_\infty$ , und Gleichheit liegt vor, wenn alle  $x_i$  den gleichen Betrag haben.

### B.8.2 Normen für quadratische Matrizen

Identifiziert man die Matrixalgebra  $\mathbb{C}^{n \times n}$  mit dem Vektorraum  $\mathbb{C}^{n^2}$ , so kann man auch hierauf die üblichen  $p$ -Normen definieren. In der Regel werden jedoch andere Normen betrachtet. Der *Betrag* einer Matrix  $T \in \mathbb{C}^{n \times n}$  ist durch  $|T| := \sqrt{T^\dagger T}$  definiert. Die *Operatornorm*  $\|T\|_\infty$  einer Matrix  $T$  ist nun der größte Eigenwert von  $|T|$ , und für  $p \in [1; \infty)$  setzt man

$$\|T\|_p := [\text{Spur } |T|^p]^{1/p} = [\text{Spur}(T^\dagger T)^{p/2}]^{1/p}. \quad (\text{B.31})$$

Ist die Matrix  $T$  normal, also diagonalisierbar, so ist dies die  $p$ -Norm der in ihrer Vielfachheit gezählten Eigenwerte. Für  $p = 1$  nennt man sie die *Spurnorm* oder *nukleare Norm*, für  $p = 2$  die *HILBERT-SCHMIDT-Norm* (auch *FROBENIUS-Norm*), die durch das *HILBERT-SCHMIDT-Skalarprodukt*  $\langle T|S \rangle := \text{Spur}(T^\dagger S)$  induziert wird, so daß die *CAUCHY-SCHWARTZsche Ungleichung* gilt.<sup>35</sup>

#### Lemma B.34 (Berechnung des Hilbert-Schmidt-Skalarproduktes)

Es seien  $A, B \in \mathbb{C}^{n \times n}$  und  $(|i\rangle)_{i=1}^n$  eine Orthonormalbasis von  $\mathbb{C}^n$ . Setzt man  $|m\rangle := \sum_{i=1}^n |i\rangle_A \langle i|_B \in \mathbb{C}^n \otimes \mathbb{C}^n$ , so gilt  $\langle A|B \rangle \stackrel{\text{Def.}}{=} \text{Spur } A^\dagger B = \langle m|A \otimes B|m\rangle$ .

*Beweis:* Man berechnet  $\langle m|A \otimes B|m\rangle = \sum_{i,j=1}^n {}_A \langle i|_B \langle i|(A \otimes B)|j\rangle_A |j\rangle_B = \sum_{i,j=1}^n \langle i|A|j\rangle \cdot \langle i|B|j\rangle$ . Mit der Gleichheit  $\langle i|A|j\rangle = \langle j|A^\dagger|i\rangle$  lautet dies  $\sum_{i,j=1}^n \langle j|A^\dagger|i\rangle \cdot \langle i|B|j\rangle = \sum_{j=1}^n \langle j|A^\dagger (\sum_{i=1}^n |i\rangle \langle i|) B|j\rangle = \text{Spur}(A^\dagger B)$ , q. e. d.

### B.8.3 Zugeordnete Matrixnormen

Sind  $(E, \|\cdot\|_E)$  und  $(F, \|\cdot\|_F)$  normierte Vektorräume, so definiert man für eine lineare Abbildung  $T : E \rightarrow F$  die Operatornorm bzgl.  $E$  und  $F$  durch

$$\|T\|_{E \rightarrow F} := \sup_{x \in E \setminus \{0\}} [\|x\|_E^{-1} \cdot \|Tx\|_F]. \quad (\text{B.32})$$

Betrachtet man spezieller die schon genannten  $p$ - und  $q$ -Normen, so werde in dieser Arbeit  $\|\cdot\|_{p \rightarrow q}$  geschrieben.<sup>36</sup> Für eine Matrix  $A = (a_{ij})_{i,j=1}^n$  nennt

---

<sup>35</sup>Allgemeiner gilt die *HÖLDERSche Ungleichung*: Erfüllen  $p, q, r \in [0; \infty]$  die Bedingung  $r^{-1} = p^{-1} + q^{-1}$ , so gilt  $\|S \cdot T\|_r \leq \|S\|_p \cdot \|T\|_q$  für beschränkte lineare Operatoren  $S$  und  $T$  auf einem Hilbertraum. Vergleiche hierzu PISIER UND XU [136], S. 1464, und MCCARTHY [121], Theorem 2.3 auf S. 254, und die Literatur zu den *SCHATTEN-Klassen*.

<sup>36</sup>Vorwiegend in der numerischen Mathematik wird der Fall  $p = q$  betrachtet und hierfür  $\|\cdot\|_p$  geschrieben. Wegen der Verwechslungsmöglichkeit mit den Normen des vorstehen-

man die Norm  $\|A\|_{\infty \rightarrow \infty} = \max_{i=1}^n \sum_{j=1}^n |a_{ij}|$  die *Zeilensummennorm* und  $\|A\|_{1 \rightarrow 1} = \|A^t\|_{\infty \rightarrow \infty}$  die *Spaltensummennorm*; ferner ist  $\|\cdot\|_{2 \rightarrow 2} = \|\cdot\|_{\infty}$ .

Im Hinblick auf entropische Unschärferelationen wird das folgende Lemma gezeigt, welches in Verbindung mit dem darauffolgenden *Satz von RIESZ und THORIN* Verwendung findet.

**Lemma B.35 (Die 1- $\infty$ -Norm)**

Für eine Matrix  $T = (t_{jk})_{j,k=1}^n \in \mathbb{C}^{n \times n}$  ist  $\|T\|_{1 \rightarrow \infty} = \max_{i,j \in \{1, \dots, n\}} |t_{ij}|$ .

*Beweis:* In Komponentenschreibweise gilt  $(Tx)_i = \sum_{k=1}^n t_{ik}x_k$ , und mit der Abkürzung  $c := \max_{i,j \in \{1, \dots, n\}} |t_{ij}|$  berechnet man

$$|(Tx)_i| = \left| \sum_{k=1}^n t_{ik}x_k \right| \leq \sum_{k=1}^n |t_{ik}| |x_k| \leq c \cdot \sum_{k=1}^n |x_k| = c \cdot \|x\|_1 \quad (\text{B.33})$$

für jedes  $i \in \{1, \dots, n\}$ . Somit ist  $\|Tx\|_{\infty} = \max_{i=1}^n |(Tx)_i| \leq c \cdot \|x\|_1$ , also  $\|T\|_{1 \rightarrow \infty} \leq c$ . Nun sei  $t_{ik}$  eine Komponente, für die  $c = |t_{ik}|$  gilt. Der Vektor  $x = (\delta_{jk})_{j=1}^n \in \mathbb{C}^n$  erfüllt  $\|x\|_1 = 1$ , und es gilt  $(Tx)_i = \sum_k t_{ik}x_k = t_{ik}$ , mithin auch  $\|Tx\|_{\infty} = \max_{j=1}^n |(Tx)_j| \geq |(Tx)_i| = |t_{ik}| = c$ , q. e. d.

**Satz B.36 (Interpolationssatz von Riesz und Thorin)**

Es seien  $T \in \mathbb{C}^{n \times n}$  sowie Werte  $p_0, p_1, q_0, q_1 \in [1; \infty]$  und  $\vartheta \in [0; 1]$  gegeben. Setzt man  $p^{-1} = (1 - \vartheta) \cdot p_0^{-1} + \vartheta \cdot p_1^{-1}$  und  $q^{-1} = (1 - \vartheta) \cdot q_0^{-1} + \vartheta \cdot q_1^{-1}$ , so gilt die Ungleichung

$$\|T\|_{p \rightarrow q} \leq \|T\|_{p_0 \rightarrow q_0}^{1-\vartheta} \cdot \|T\|_{p_1 \rightarrow q_1}^{\vartheta}.$$

Anders ausgedrückt ist die Abbildung  $f_T(x, y) := \ln \|T\|_{1/x \rightarrow 1/y}$  für jede quadratische Matrix  $T$  konvex.

Ein Beweis findet sich bei WERNER [184], Satz II.4.2 auf S. 73–77, oder auch bei REED UND SIMON [143], Band II auf S. 27–28, wo er als Korollar aus dem STEINschen *Interpolationssatz* (S. 40) folgt. Ein hierzu verwandter Satz findet sich auch bei HARDY, LITTLEWOOD UND PÓLYA [70], Anhang II in Verbindung mit den S. 214–220.

## B.8.4 Gemeinsame Wahrscheinlichkeitsverteilungen

Im diesem Unterabschnitt wird ein Satz bewiesen, der für das Verständnis des Sicherheitsbegriffs in der Quantenkryptographie (vgl. Unterabschnitt 3.4.2)

---

den Unterabschnitts wird die Schreibweise hier nicht verwendet. Es handelt sich hierbei um (*zugeordnete*) *Matrixnormen* (oder allgemeiner *Algebrennormen*) in dem Sinne, daß zusätzlich zu den Normaxiomen die Ungleichung  $\|AB\| \leq \|A\| \cdot \|B\|$  gilt.

bedeutsam ist. Er verbindet die Spurnorm zweier Dichtematrizen und ihre Unterscheidbarkeit mittels Messungen.

Im folgenden seien  $p = (p_i)_{i=1}^n \in \mathcal{W}_n$  und  $q = (q_j)_{j=1}^{n'} \in \mathcal{W}_{n'}$  gewöhnlich Wahrscheinlichkeitsverteilungen auf  $n$  bzw.  $n'$  Elementen. Im Falle von  $n = n'$  ist der *Variationsabstand* von  $p$  und  $q$  als die Hälfte der Spurnorm ihrer Differenz definiert, das heißt

$$\delta(p, q) := \frac{1}{2} \|p - q\|_1 = \frac{1}{2} \sum_{k=1}^n \left[ \max\{p_k, q_k\} - \min\{p_k, q_k\} \right]. \quad (\text{B.34})$$

Der Grund für die Definition des Variationsabstands ist der weiter unten stehende Satz B.38, dessen Kern bereits in folgendem Lemma steckt.

**Lemma B.37 (Gemeinsame Wahrscheinlichkeitsverteilungen)**

Zu  $p \in \mathcal{W}_n$  und  $q \in \mathcal{W}_{n'}$  findet man eine Matrix  $M = (m_{ij})_{i=1, j=1}^{n, n'} \in (\mathbb{R}_0^+)^{n \times n'}$  mit nicht-negativen Einträgen derart, daß die Gleichungen

$$p_i = \sum_{j=1}^{n'} m_{ij} \quad \text{und} \quad q_j = \sum_{i=1}^n m_{ij}$$

für alle Werte  $i \in \{1, \dots, n\}$  bzw.  $j \in \{1, \dots, n'\}$  erfüllt sind.

Die Matrix  $M$  kann als gemeinsame Wahrscheinlichkeitsverteilung aufgefaßt werden, deren Randverteilungen  $p$  und  $q$  sind; dies wird durch folgendes Diagramm veranschaulicht:

$$M = \begin{array}{ccc} \left( \begin{array}{ccc} m_{11} & \dots & m_{1n'} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{nn'} \end{array} \right) & \begin{array}{l} \rightarrow p_1 \\ \vdots \\ \rightarrow p_n \end{array} \\ \begin{array}{ccc} \downarrow & & \downarrow \\ q_1 & \dots & q_{n'} \end{array} \end{array}$$

Man erkennt, daß die Zeilensummen die Einträge aus  $p$  und die Spaltensummen diejenigen aus  $q$  ergeben.

*Beweis:* Es genügt, den Fall  $n = n'$  zu betrachten, da andernfalls das kürzere Tupel mit Nullen auf die Länge des längeren Tupels aufgefüllt werden kann. In der Matrix  $M$  stehen in den zusätzlichen Zeilen bzw. Spalten lauter Nullen, die anschließend wieder weggelassen werden können.

Der weitere Beweis erfolgt durch vollständige Induktion; hierbei ist der Induktionsanfang ( $n = 1$ ) offensichtlich, und als Induktionsvoraussetzung nehme man an, daß die Behauptung für alle Zahlen  $1, \dots, n - 1$  gezeigt sei.

Für den Induktionsschritt nehme man an, die Komponenten der Tupel  $p$  und  $q$  seien so angeordnet, daß für einen geeigneten Wert  $m \in \{0, \dots, n\}$  die

Ungleichungen  $p_i \geq q_i$  für  $i \in \{1, \dots, m\}$  und  $p_i \leq q_i$  für  $i \in \{m+1, \dots, n\}$  gelten; dies kann durch eine Permutation der  $i \in \{1, \dots, n\}$  geschehen, die anschließend wieder rückgängig gemacht wird. Es kann nun stets  $m \notin \{0, n\}$  gewählt werden, da andernfalls entweder alle Komponenten der Tupel gleich sind oder die Summe ihrer Komponenten verschieden sein muß, was aber ausgeschlossen wurde. In der Matrix  $M$  setze man nun

$$m_{kk} := \min \{p_k, q_k\} = \begin{cases} q_k & \text{für } k \in \{1, \dots, m\} \\ p_k & \text{für } k \in \{m+1, \dots, n\} \end{cases} . \quad (\text{B.35})$$

Die Anforderungen an die Spalten- und Zeilensummen besagen folglich, daß  $m_{ij} = 0$  im Falle von  $i \neq j$  ist, sofern  $i \geq m+1$  oder  $j \leq m$  gilt. Die gesuchte Matrix hat daher die Form

$$M = \left( \begin{array}{ccc|ccc} q_1 & & & & & \\ & \ddots & & & & \\ & & q_m & & * & \\ \hline & & & p_{m+1} & & \\ & 0 & & & \ddots & \\ & & & & & p_n \end{array} \right) . \quad (\text{B.36})$$

Bis auf den oberen rechten Block besteht die Matrix also nur aus Diagonalelementen. Es ist somit nur noch die Teilmatrix  $M' = (m_{ij})_{i=1, j=m+1}^m$  mit Zahlen zu besetzen, und zwar so, daß

$$\sum_{j=1}^m m_{ij} = p_i - q_i \quad \text{und} \quad \sum_{i=m+1}^n m_{ij} = q_j - p_j \quad (\text{B.37})$$

für  $i \in \{1, \dots, m\}$  bzw.  $j \in \{m+1, \dots, n\}$  gelten. Ferner rechnet man nach, daß  $\sum_{i=1}^m p_i - q_i = \sum_{i=m+1}^n q_i - p_i \geq 0$  gilt. Wie zuvor festgehalten wurde, kann  $m \notin \{0, n\}$  gewählt werden, so daß  $\max \{m, n-m\} < n$  ist. Dies ist nach Induktionsvoraussetzung lösbar, da dies bis auf eine Umskalierung die gleiche Fragestellung für den Wert  $\max \{m, n-m\} < n$  ist, so daß ein  $M'$  existiert, q. e. d.

**Satz B.38 (Wahrscheinlichkeitsverteilungen und Spurnorm)**

*Sind  $X$  und  $Y$  diskrete Zufallsvariablen auf einer Menge  $\{1, \dots, n\}$  mit zugehörigen Wahrscheinlichkeitsverteilungen  $p, q \in \mathcal{W}_n$ , so existiert eine gemeinsame Wahrscheinlichkeitsverteilung auf der Menge  $\{1, \dots, n\} \times \{1, \dots, n\}$  derart, daß sich  $p$  und  $q$  als deren Randverteilungen ergeben, das heißt, es können für jedes Tupel  $(x, y) \in \{1, \dots, n\} \times \{1, \dots, n\}$  nicht-negative Werte*

$P(X = x, Y = y)$  derart angegeben werden, daß

$$P(X = x) = \sum_{y=1}^n P(X = x, Y = y)$$

und  $P(Y = y) = \sum_{x=1}^n P(X = x, Y = y)$

für alle  $x, y \in \{1, \dots, n\}$  gelten und daß  $P(X \neq Y) \leq \delta(p, q)$  ist.

Die Aussage scheint zwar allgemein bekannt zu sein, ihr Beweis findet sich aber nicht in der Literatur.<sup>37</sup>

*Beweis:* Eine gemeinsame Wahrscheinlichkeitsverteilung für  $X$  und  $Y$  kann mit einer quadratischen Matrix nicht-negativer Zahlen  $A = (a_{ij})_{i,j=1}^n$  identifiziert werden, deren Komponentensumme Eins ergibt: hierzu verende man  $P(X = x, Y = y) = a_{xy}$ . Nach Lemma B.37 existiert nun eine solche Matrix, die die vorgegebenen Randverteilungen  $p$  und  $q$  erzeugt. Wegen der Gleichung  $P(X \neq Y) = 1 - \text{Spur } A$  ist nur noch zu zeigen, daß  $1 - \text{Spur } A \leq \delta(p, q)$  oder

$$1 - \sum_{k=1}^n \min \{p_k, q_k\} \leq \sum_{k=1}^n \frac{\max \{p_k, q_k\} - \min \{p_k, q_k\}}{2} \quad (\text{B.38})$$

gilt. Bringt man die linke Summe auf die rechte Seite und beachtet, daß  $\max \{p_k, q_k\} + \min \{p_k, q_k\} = p_k + q_k$  für jeden Summanden  $k \in \{1, \dots, n\}$  gilt, so folgt mittels  $\sum_{k=1}^n p_k = \sum_{k=1}^n q_k = 1$  sogar die Gleichheit beider Seiten, q. e. d.

## B.9 Vermischte Ergebnisse

In diesem Abschnitt werden einige Ergebnisse aus der Mathematik zusammengetragen, die im Verlauf der Arbeit verwendet wurden, sich aber nicht in die bisherige Systematik einordnen lassen.

### B.9.1 Taylor-Reihen und Restglieder

In der Physik werden sehr oft Taylorreihen verwendet, bei denen in der Regel nur die niedrigste nicht-verschwindende Ordnung betrachtet wird. Um die Fehler solcher Betrachtungen abzuschätzen, werden Restglieddarstellungen verwendet; der nachfolgende Satz liefert eine allgemeine Darstellung.

---

<sup>37</sup>Ohne Beweis wird sie unter anderem von CHRISTANDL, RENNER UND EKERT [38], Lemma 3.1 auf S. 9, von RENNER [147], Proposition 2.1.1 auf S. 24, sowie von KÖNIG UND RENNER [149], Lemma 2.1 auf S. 4 erwähnt.

**Satz B.39 (Taylor-Entwicklung und Schlömilchsches Restglied)**

Eine Funktion  $f : [x_0; x] \rightarrow \mathbb{R}$  sei auf  $[x_0; x]$   $n$ -mal stetig differenzierbar, und  $f^{(n+1)}$  existiere zumindest im Intervall  $(x_0; x)$ . Wählt man  $p \in \mathbb{N}$ , so ist

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}}{k!} (x - x_0)^k + R_n,$$

wobei ein (im allgemeinen von  $p$  abhängiges)  $\vartheta \in (0; 1)$  derart existiert, daß

$$R_n(x) = \frac{f^{(n+1)}(x_0 + \vartheta \cdot (x - x_0))}{n! p} (1 - \vartheta)^{n+1-p} (x - x_0)^{n+1}$$

gilt.

Den Ausdruck  $R_n$  nennt man das SCHLÖMILCHsche Restglied; für  $p = n + 1$  geht dieses in das Restglied von LAGRANGE, für  $p = 1$  in das von CAUCHY über; vergleiche zum Satz z. B. HEUSER [74], Teil I, S. 355–357 und S. 613.

**B.9.2 Die Fouriertransformation**

Es seien  $\alpha, C \in \mathbb{R}^+$  und  $k \in \{+1, -1\}$  vorgegeben. Zu einer integrierbaren Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  definiert man ihre *Fouriertransformierte*  $\mathfrak{F}\{f\}$  und ihre *inverse Fouriertransformierte*  $\mathfrak{F}^{-1}\{f\}$  mittels

$$\mathfrak{F}\{f\}(\xi) := C \left(\frac{\alpha}{2\pi}\right)^{n/2} \int_{x \in \mathbb{R}^n} f(x) e^{-i\alpha k \langle x, \xi \rangle} d^n x, \tag{B.39}$$

$$\mathfrak{F}^{-1}\{f\}(x) := C^{-1} \left(\frac{\alpha}{2\pi}\right)^{n/2} \int_{\xi \in \mathbb{R}^n} f(\xi) e^{+i\alpha k \langle x, \xi \rangle} d^n \xi, \tag{B.40}$$

wobei  $\langle x, \xi \rangle = \sum_{i=1}^n x_i \xi_i$  ist und über das LEBESGUE-Maß auf  $\mathbb{R}^n$  integriert wird. Sowohl  $\mathfrak{F}\{f\}$  als auch  $\mathfrak{F}^{-1}\{f\}$  sind beschränkt, stetig und fallen für  $\|x\| \rightarrow \infty$  auf Null. Für  $C = 1$  ist die Fouriertransformation symmetrisch.

**B.9.3 Das Haarsche Maß**

Aus der Theorie der topologischen Gruppen ist das HAARSche Maß bekannt, das auf jeder lokalkompakten Gruppe existiert. Hier wird nur der wesentlich einfachere Spezialfall kompakter Gruppen benötigt, der bei Mischoperationen (engl. *twirling*) zum Tragen kommt.

**Lemma B.40 (Satz über das Haarsche Maß)**

Auf einer kompakten (topologischen HAUSDORFFschen) Gruppe  $G$  existiert genau ein normiertes linksinvariantes Maß; dieses Maß ist gleichzeitig rechtsinvariant.

Ein Beweis dieses Satzes findet sich in den Lehrbüchern zur Maßtheorie, z. B. bei ELSTRODT [50], Kap. VIII, § 3, insbesondere Satz VIII.3.12 auf S. 362. Eine endliche Gruppe  $G$  bildet mit der diskreten Topologie versehen eine kompakte Gruppe, und das normierte HAAR-Maß stimmt mit dem auf Eins normierten Zählmaß überein, das heißt, für eine Menge  $A \subseteq G$  gilt  $\mu(A) = |A| / |G|$ .

Das HAAR-Maß ist also das eindeutig bestimmte invariante Maß über einer Gruppe  $G$ . Wirkt diese Gruppe auf einen Raum  $X$ , so ist eine Teilmenge  $A \subseteq X$  im allgemeinen nicht invariant unter  $G$ , kann aber durch den Übergang zu  $A' := \int_g gA dg$  invariant gemacht werden. Diese *Integration über das HAAR-Maß* kann wird in Mischoperationen (engl. *twirling*) verwendet.

## B.9.4 Die Singulärwertzerlegung

Die Singulärwertzerlegung ist eine Zerlegung, die für alle kompakten Operatoren – insbesondere für alle Matrizen – möglich ist; mit ihrer Hilfe kann die SCHMIDT-Zerlegung gezeigt werden.

### Satz B.41 (Singulärwertzerlegung kompakter Operatoren)

Zu einem kompakten linearen Operator  $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  existieren Orthonomalsysteme  $(e_i)_{i \in I}$  von  $\mathcal{H}_1$  und  $(f_i)_{i \in I}$  von  $\mathcal{H}_2$  und ein System positiver Zahlen  $(\lambda_i)_{i \in I} \in (\mathbb{R}^+)^I$ , so daß  $T = \sum_{i \in I} \lambda_i \langle \cdot, e_i \rangle f_i$  gilt.<sup>38</sup> Für die Mächtigkeit der Indexmenge  $I$  gilt  $|I| \leq \min \{ \dim \mathcal{H}_1, \dim \mathcal{H}_2 \}$ , und sie ist höchstens abzählbar unendlich.

Beweise finden sich z. B. bei WERNER [184], Satz VI.3.6 auf S. 271, oder bei REED UND SIMON [143], Band I, Theorem VI.17 auf S. 203–204. Im Falle endlichdimensionaler Hilberträume folgt aus diesem Satz die bekanntere Singulärwertzerlegung für allgemeine, nicht notwendigerweise quadratische Matrizen.

### Korollar B.42 (Singulärwertzerlegung von Matrizen)

Zu jeder Matrix  $T \in \mathbb{C}^{n \times m}$  gibt es unitäre Matrizen  $U \in \mathbb{C}^{n \times n}$  und  $V \in \mathbb{C}^{m \times m}$  derart, daß die Gleichung  $T = U \Sigma V$  gilt, wobei  $\Sigma \in \mathbb{C}^{n \times m}$  eine Matrix ist, bei der nur die Diagonalelemente nicht verschwinden und diese zusätzlich nicht-negativ sind.

Die *Singulärwerte* sind die  $\lambda_i$  bzw. die Diagonaleinträge der Matrix  $\Sigma$ .

---

<sup>38</sup>In Bra-Ket-Notation lautet diese Gleichheit  $T = \sum_{i \in I} \lambda_i |f_i\rangle \langle e_i|$ .

### B.9.5 Einige Ungleichungen

In diesem Unterabschnitt werden zwei elementare Ungleichungen gezeigt.

#### Lemma B.43 (Abschätzung des Logarithmus)

Für  $x \in \mathbb{R}^+$  gilt  $x^{-1}(x-1) \leq \ln x \leq x-1$ ; Gleichheit gilt nur für  $x = 1$ .

*Beweis:* Der Mittelwertsatz besagt, daß für eine differenzierbare Funktion  $f : [a; b] \rightarrow \mathbb{R}$  sich stets eine geeignete Stelle  $\xi \in (a; b)$  derart findet, daß  $f(b) - f(a) = f'(\xi)(b - a)$  gilt, es also einen inneren Punkt gibt, dessen Tangentensteigung der Sekantensteigung über  $[a; b]$  entspricht. Wählt man nun zunächst  $a = 1$  und  $b = x > 1$ , so folgt wegen  $\ln 1 = 0$  aus dem genannten, daß  $\ln x = (x-1)/\xi$  für ein  $\xi \in (1; x)$  ist, woraus die Behauptung für  $x > 1$  folgt. Im umgekehrten Falle setze man  $a = x$  und  $b = 1$ , woraufhin der Mittelwertsatz entsprechend angewendet werden kann, q. e. d.

Das folgende Lemma findet sich als Lemma 1.14 bei OHYA UND PETZ [131] auf S. 27.

#### Lemma B.44 (Eine Ungleichung)

Für  $u, v \in [0; 1]$  mit  $u \leq v$  gilt

$$2(u-v)^2 \leq u \ln \frac{u}{v} + (1-u) \ln \frac{1-u}{1-v}.$$

*Beweis:* Es soll gezeigt werden, daß die Funktion  $f(u, v) := u \ln u - u \ln v + (1-u) \ln(1-u) - (1-u) \ln(1-v) - 2(u-v)^2$  im betrachteten Bereich nicht-negativ ist. Im Falle von  $u = v$  erhält man durch Einsetzen  $f(u, v) = 0$ . Die partielle Ableitung nach  $v$  ergibt sich zu

$$\frac{\partial f}{\partial v} = -u \cdot \frac{1}{v} - (1-u) \cdot \frac{-1}{1-v} + 4 \cdot (u-v). \quad (\text{B.41})$$

Sie ist also nicht-negativ, falls

$$\frac{(1-u)v - (1-v)u}{v \cdot (1-v)} = \frac{v-u}{v \cdot (1-v)} \geq 4 \cdot (v-u) \quad (\text{B.42})$$

gilt; ferner ist  $v \cdot (1-v) \leq 1/4$ , q. e. d.



# Anhang C

## Informationstheorie und Kodierung

In diesem Kapitel werden einige Ergebnisse der Informationstheorie und Kodierung zusammengestellt. Im weiteren Sinne gehören sie sicherlich zur Mathematik, sind aber für diese eher von nachrangiger Bedeutung. Hingegen haben sie viele Anwendungen in der Daten- und Informationsübertragung und werden daher auch in der Quantenkryptographie verwendet.

### C.1 Klassische Informationstheorie

In diesem Abschnitt werden die Grundzüge der klassischen Informationstheorie behandelt. Hierzu gehören insbesondere die Definition der Entropie und die Untersuchung ihrer Eigenschaften; vgl. NIELSEN UND CHUANG [127].

Im gesamten Abschnitt bezeichnen  $X, Y, Z, \dots$  Zufallsvariablen, deren Ergebnismenge mit  $\{1, \dots, n\}$  identifiziert werden soll; einige Eigenschaften können aber auch für unendliche oder sogar kontinuierliche Zufallsvariablen verallgemeinert werden. Die den Zufallsvariablen zugeordneten Wahrscheinlichkeitsverteilungen seien  $P_X, P_Y$  usw., gemeinsame Verteilungen  $P_{XY}$  usw. Es werden Kurzschreibweisen wie  $p_i := p(x_i) := P(X_i = x_i)$  verwendet, wenn sie aus dem Zusammenhang ersichtlich sind.

#### C.1.1 Die Shannon-Entropie

Die SHANNON-*Entropie* ist der wohl wichtigste Begriff der Informationstheorie. In den nun folgenden Definitionen tauchen Logarithmenausdrücke auf. Je nach Zusammenhang können dabei für die Basis verschiedene Werte verwendet werden; in diesem Kapitel verwende ich vorwiegend den Logarith-

mus zur Basis 2, so daß die Information in Bits gemessen wird, während ich im ersten Hauptteil der Dissertation die Basis  $d$  bevorzugt, also Dits als Informationseinheit verwendet habe. In Rechnungen ist oft der natürliche Logarithmus ( $\ln$ ) zur Basis  $e$  angenehmer. Die Ausdrücke unterscheiden sich wegen der Rechenregel  $\log_b x = \ln x / \ln b$  nur um einen Faktor und werden austauschbar verwendet.

**Definition C.1 (Shannon-Entropie)**

Für eine Wahrscheinlichkeitsverteilung  $p = (p_1, p_2, p_3, \dots, p_n) \in \mathcal{W}_n$  ist die SHANNON-Entropie durch

$$H(p) := - \sum_{i=1}^n p_i \log_2 p_i = -(\ln 2)^{-1} \sum_{i=1}^n p_i \ln p_i \quad (\text{C.1})$$

definiert. Die SHANNON-Entropie einer Zufallsvariablen ist die SHANNON-Entropie der Wahrscheinlichkeitsverteilung, die diese Zufallsvariable besitzt.

Ein wichtiger Spezialfall der SHANNON-Entropie ergibt sich, wenn  $p_1 = 1 - x$  und  $p_i = (n-1)^{-1}x$  für alle  $i \in \{2, \dots, n\}$  und ein  $x \in [0; 1]$  ist; man schreibt dann verkürzt

$$H_n(x) := H(p_1, \dots, p_n) = -(1-x) \log_2(1-x) - x \log_2 \frac{x}{n-1}; \quad (\text{C.2})$$

im Falle von  $n = 2$  erhält man  $H(x) = -x \log_2 x - (1-x) \log_2(1-x)$  und nennt dies die *binäre SHANNON-Entropie*.

Betrachtet man zwei Zufallsvariablen  $X$  und  $Y$  mit einer gemeinsamen Verteilung  $P_{XY} = (p_{ij})_{i=1}^n_{j=1}^m$  für  $p_{ij} := P(X = i \wedge Y = j)$ , so definiert man die *Verbundentropie* oder *verbundene Entropie* von  $X$  und  $Y$  mittels

$$H(X, Y) := H(P_{XY}) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log_2 p_{ij}. \quad (\text{C.3})$$

**C.1.2 Die relative Entropie**

Sind den Zufallsvariablen  $X$  und  $Y$  die Verteilungen  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  und  $q = (q_1, \dots, q_n) \in \mathcal{W}_n$  zugeordnet, so wird ihre *relative Entropie* oder *KULLBACK-LEIBLER-Information* durch

$$H(X||Y) := H(p||q) := \sum_{i=1}^n p_i \log_2 \frac{p_i}{q_i} = -H(X) - \sum_{i=1}^n p_i \log_2 q_i \quad (\text{C.4})$$

definiert. Die relative Entropie ist nicht symmetrisch in  $p$  und  $q$  und somit auch keine Metrik. Ist die Zufallsvariable  $Y$  gleichverteilt, gilt also  $q_i = 1/n$  für alle  $i \in \{1, \dots, n\}$ , so erhält man  $H(X||Y) = -H(X) + \log_2 n$ .

**Lemma C.2 (Definitheit der relativen Entropie)**

Die relative Entropie von  $X$  und  $Y$  ist nicht-negativ und verschwindet genau für  $P_X = P_Y$ .

*Beweis:* Unter Verwendung des Lemmas B.43 berechnet man  $-\ln 2 \cdot H(X||Y) = \sum_{i=1}^n p_i \ln(q_i/p_i) \leq \sum_{i=1}^n p_i [(q_i/p_i) - 1] = \sum_{i=1}^n (q_i - p_i) = 0$ , und es gilt Gleichheit, wenn  $p_i = q_i$  für alle  $i \in \{1, \dots, n\}$  ist, q. e. d.

Sind  $P_X$  und  $P_Y$  die Randverteilungen einer gemeinsamen Verteilung  $P_{XY}$ , so sind  $X$  und  $Y$  genau dann statistisch unabhängig, wenn  $P_{XY} = P_X \cdot P_Y$  gilt, wobei  $(P_X \cdot P_Y)(X = x \wedge Y = y) := P_X(X = x) \cdot P_Y(Y = y)$  ist.

**Lemma C.3 (Subadditivität der Entropie)**

Für zwei Zufallsvariablen  $X$  und  $Y$  gilt die Gleichung

$$H(P_{XY}||P_X \cdot P_Y) = H(X) + H(Y) - H(X, Y),$$

mithin also  $H(X) + H(Y) \geq H(X, Y)$  mit Gleichheit genau dann, wenn  $X$  und  $Y$  statistisch unabhängig sind.

*Beweis:* Schreibt man  $p_{i*} = \sum_j p_{ij}$  und  $p_{*j} = \sum_i p_{ij}$  für die beiden Randverteilungen, so berechnet man  $-H(X, Y) + H(X) + H(Y) = \sum_{ij} p_{ij} \log_2 p_{ij} - \sum_i p_{i*} \log_2 p_{i*} - \sum_j p_{*j} \log_2 p_{*j} = \sum_{ij} p_{ij} [\log_2 p_{ij} - \log_2 p_{i*} - \log_2 p_{*j}] = \sum_{ij} p_{ij} [\log_2 \frac{p_{ij}}{p_{i*} \cdot p_{*j}}] = H(P_{XY}||P_X \cdot P_Y)$  und verwendet Lemma C.2, q. e. d.

**C.1.3 Bedingte Entropie und gemeinsame Information**

Ein weiterer Begriff der Informationstheorie ist die *bedingte Entropie* zweier Zufallsvariablen  $X$  und  $Y$ , die durch  $H(X|Y) = H(X, Y) - H(Y)$  bzw. durch  $H(Y|X) = H(X, Y) - H(X)$  definiert wird. Schließlich ist die *gemeinsame Information* von  $X$  und  $Y$  mittels  $H(X : Y) := H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$  festgelegt; oft schreibt man auch  $I(X : Y)$ . Der folgende Satz enthält einige Eigenschaften dieser Größen.

**Satz C.4 (Eigenschaften der Entropien)**

Für zwei Zufallsvariablen  $X$  und  $Y$  gilt

1.  $H(X, Y) = H(Y, X)$  und  $H(X : Y) = H(Y : X)$ ;
2.  $H(Y|X) \geq 0$ ,  $H(X : Y) \leq H(Y)$  und  $H(X) \leq H(X, Y)$  mit Gleichheit genau dann, wenn  $Y = f(X)$  für eine geeignete Funktion  $f$  ist; und
3.  $H(X, Y) \leq H(X) + H(Y)$ ,  $H(Y|X) \leq H(Y)$  und  $H(X : Y) \geq 0$  mit Gleichheit genau dann, wenn  $X$  und  $Y$  statistisch unabhängig sind.

*Beweis:* Die erste Aussage ist trivial, die dritte folgt aus Lemma C.3. Für die zweite Behauptung berechnet man  $H(X, Y) = -\sum_{xy} p(x, y) \log_2 p(x)p(y|x) = -\sum_x p(x) \log_2 p(x) - \sum_{xy} p(x, y) \log_2 p(y|x)$ . Der erste Term ist  $H(X)$ , die zweite Summe wegen  $p(y|x) \leq 1$  nicht-positiv; sie verschwindet, wenn  $p(x, y) > 0 \Rightarrow p(y|x) = 1$  gilt, wenn also  $Y$  aus  $X$  mit Sicherheit vorhergesagt werden kann, q. e. d.

### C.1.4 Markow-Ketten und starke Subadditivität

Im folgenden werden einige wichtige, aber nicht ganz triviale Eigenschaften der klassischen Entropie gezeigt. Im Vergleich zu ihren quantenmechanischen Verallgemeinerungen sind sie aber leichter zu beweisen.

#### Definition C.5 (Markow-Ketten)

Eine (endliche oder unendliche) Folge von Zufallsvariablen  $X_1, X_2, \dots$  heißt MARKOW-Kette oder MARKOW-Prozeß, wenn für jedes geeignete  $n \in \mathbb{N}$  die Gleichung

$$P(X_{n+1} = x_{n+1} | X_n = x_n \wedge \dots \wedge X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n)$$

gilt. Man schreibt  $X_1 \rightarrow X_2 \rightarrow \dots$ .

#### Lemma C.6 (Markow-Ketten)

Die Folge  $X \rightarrow Y \rightarrow Z$  ist genau dann eine MARKOW-Kette, wenn auch die Folge  $Z \rightarrow Y \rightarrow X$  eine MARKOW-Kette ist.

*Beweis:* Sieht man von zwei irrelevanten Fällen ab, so kann man  $p(x, y) > 0$  und  $p(y, z) > 0$  annehmen. In diesem Fall erhält man  $p(z|x, y) = p(z|y) \Leftrightarrow \frac{p(x, y, z)}{p(x, y)} = \frac{p(y, z)}{p(y)} \Leftrightarrow \frac{p(x, y, z)}{p(y, z)} = \frac{p(x, y)}{p(y)} \Leftrightarrow p(x|y, z) = p(x|y)$ , q. e. d.

#### Satz C.7 (Starke Subadditivität)

Für drei Zufallsvariablen gilt  $H(X, Y, Z) + H(Y) \leq H(X, Y) + H(Y, Z)$ ; Gleichheit gilt genau dann, wenn  $Z \rightarrow Y \rightarrow X$  eine MARKOW-Kette ist. Ferner ist  $H(X|Y, Z) \leq H(X|Y)$ .

*Beweis:* Die zweite Aussage folgt über  $H(X|Y, Z) = H(X, Y, Z) - H(Y, Z) \leq H(X, Y) - H(Y) = H(X|Y)$  aus der ersten. Für diese wiederum ist  $H(X, Y, Z) - H(X, Y) \leq H(Y, Z) - H(Y)$  zu zeigen; man formt dies mittels  $p(x, y, z) = p(z|x, y)p(y|x)p(x)$  und  $p(y, z) = p(y|x)p(x)$  und anschließendem Kürzen zu

$$-\sum_{xyz} p(x, y, z) \log_2 p(z|x, y) \leq -\sum_{yz} p(y, z) \log_2 p(z|y) \quad (\text{C.5})$$

um. Dies ist äquivalent zu  $\sum_{xyz} p(x, y, z) \log_2 \frac{p(z|x,y)}{p(z|y)} \geq 0$ , und man erkennt, daß Gleichheit genau für MARKOW-Ketten gilt. Für den Ausdruck im Logarithmus berechnet man weiter

$$p(z|x, y) = \frac{p(x, y, z)}{p(x, y)} = \frac{p(x, z|y) \cdot p(y)}{p(x|y) \cdot p(y)} = \frac{p(x, z|y)}{p(x|y)}, \quad (\text{C.6})$$

so daß nur noch  $\sum_y p(y) \sum_{xz} p(x, z|y) \log_2 \frac{p(x, z|y)}{p(x|y)p(z|y)} \geq 0$  zu zeigen ist. Dies ist aber eine über  $y$  gemittelte relative Entropie, also nicht-negativ, q. e. d.

**Satz C.8 (Kettenregel für die bedingte Entropie)**

Für beliebige Zufallsvariablen  $X_1, \dots, X_n$  und  $Y$  gilt

$$H(X_1, \dots, X_n|Y) = \sum_{i=1}^n H(X_i|X_1, \dots, X_{i-1}, Y).$$

*Beweis:* Der Fall  $n = 1$  ist trivial, und für  $n = 2$  erhält man  $H(X_1, X_2|Y) = H(X_1, X_2, Y) - H(Y) = H(X_1, X_2, Y) - H(X_1, Y) + H(X_1, Y) - H(Y) = H(X_2|X_1, Y) + H(X_1|Y)$  als Induktionsanfang. Hieraus folgt weiter

$$H(X_1, \dots, X_{n+1}|Y) = H(X_{n+1}|X_1, \dots, X_n, Y) + H(X_1, \dots, X_n|Y)$$

und, nachdem man die Induktionsvoraussetzung für den rechten Term eingesetzt hat, auch die Behauptung, q. e. d.

Die gemeinsame Information ist weder allgemein sub- noch superadditiv. Hierzu betrachte man  $B(1, 1/2)$ -verteilte Zufallsvariablen:

1.  $X$  und  $Y$  seien statistisch unabhängig und  $Z := (X + Y) \bmod 2$ ; es gilt dann  $1 = H(X, Y : Z) > H(X : Z) + H(Y : Z) = 0$ , da je zwei Zufallsvariablen statistisch unabhängig sind, man aber aus zweien die dritte eindeutig erschließen kann;
2.  $X_1 = X_2 = Y_1 = Y_2$ , wobei hier die Ergebnisse und nicht nur die Verteilungen gemeint sind; da jede Zufallsvariable eine Funktion jeder anderen ist, gilt  $2 = H(X_1 : Y_1) + H(X_2 : Y_2) > H(X_1, X_2 : Y_1, Y_2) = 1$ .

**Korollar C.9 (Datenverarbeitungsungleichung)**

Ist  $X \rightarrow Y \rightarrow Z$  eine MARKOW-Kette, so gilt

$$H(X : Y) \geq H(X : Z) \quad \text{und} \quad H(Z : Y) \geq H(Z : X).$$

*Beweis:* Aufgrund von Lemma C.6 muß nur die erste Ungleichung bewiesen werden, für die nur noch  $H(X|Z) \geq H(X|Y)$  zu zeigen ist. Nach dem gleichen Lemma gilt  $H(X|Z) \geq H(X|Y) = H(X|Y, Z)$ , und Satz C.7 liefert  $H(X|Z) \geq H(X|Y, Z)$ , q. e. d.

### C.1.5 Rényi-Entropien

Die RÉNYI- $\alpha$ -Entropien sind Verallgemeinerungen der SHANNON-Entropie. Im folgenden werden die grundlegenden Definitionen und Sätze aufgeführt; vgl. zum Beispiel den Anhang bei RÉNYI [150] oder das Buch von BECK UND SCHLÖGEL [12]. Die RÉNYI-Entropien finden Anwendungen in entropischen Unschärferelationen (z. B. in Satz 9.2).

#### Definition C.10 (Rényi- $\alpha$ -Entropien)

Es sei  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  eine Wahrscheinlichkeitsverteilung. Für ein  $\alpha \in \mathbb{R}^+ \setminus \{1\}$  definiert man die RÉNYI- $\alpha$ -Entropie mittels

$$R_\alpha(p) := \frac{1}{1 - \alpha} \log_2 \sum_{i=1}^n p_i^\alpha.$$

Für  $\alpha \in \{0, 1, \infty\}$  setzt man  $R_\alpha(p) := \lim_{\beta \rightarrow \alpha} R_\beta(p)$ .<sup>1</sup>

Mittels der Festsetzung  $R_\alpha(p||q) := \alpha(\alpha - 1)^{-1} \log_2 \sum_{i=1}^n p_i^\alpha / q_i^{\alpha-1}$  kann man, wie analog schon bei der SHANNON-Entropie, *relative Rényi-Entropien* definieren. Oft wird auch  $H_\alpha$  für die Rényi-Entropien geschrieben, was hier, um Verwechslungen zu vermeiden, aber nicht geschieht. Andere Verallgemeinerungen der SHANNON-Entropie sind auch bekannt, z. B. die TSALLIS-Entropien  $T_\alpha(p) := (\alpha - 1)^{-1} (1 - \sum_{i=1}^n p_i^\alpha)$ .

#### Lemma C.11 (Spezielle Entropien)

Für  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  ist

- $R_0(p) = \log_2 |\{i \in \{1, \dots, n\} \mid p_i \neq 0\}|$ , das heißt,  $R_0$  ist der Logarithmus der Mächtigkeit des Trägers der Funktion  $p : \{1, \dots, n\} \rightarrow \mathbb{R}$ ;
- $R_1(p)$  die SHANNON-Entropie der Verteilung  $p$ ;
- $R_2(p) = -\log_2 P(X = Y)$ , wenn  $X$  und  $Y$  statistisch unabhängig sind und beide die Verteilung  $p$  besitzen; man spricht von der Kollisionsentropie;
- $R_\infty(p) = -\log_2 \sup \{p_i \mid i \in \{1, \dots, n\}\}$ .

*Beweis:* Die Aussage für  $R_2$  folgt aus der Definition. Für die übrigen Aussagen verwendet man die Regel von DE L'HÔPITAL und berechnet

$$\lim_{\beta \rightarrow \alpha} R_\beta(p) \stackrel{\text{v.H.}}{=} - \lim_{\beta \rightarrow \alpha} \frac{\sum_{i=1}^n p_i^\beta \log_2 p_i}{\left| \sum_{i=1}^n p_i^\beta \right|} = - \lim_{\beta \rightarrow \alpha} \sum_{i=1}^n q_i(\beta) \log_2 p_i, \quad (\text{C.7})$$

<sup>1</sup>Man beachte, daß im Falle von  $\alpha \in \mathbb{R}^+ \setminus \{1\}$  die RÉNYI-Entropien auch in der Form  $R_\alpha(p) = \alpha^{-1} (1 - \alpha)^{-1} \log_2 \|p\|_\alpha$  geschrieben werden können, so daß einige Eigenschaften der  $p$ -Normen (vgl. Unterabschnitt B.8.1) sich auf diese übertragen.

wenn man  $q_i(\beta) := p_i^\beta / \sum_{j=1}^n p_j^\beta$  setzt. Für  $\alpha = 1$  ist  $q_i = p_i$ , so daß die SHANNON-Entropie vorliegt. Es gilt  $\lim_{\beta \rightarrow 0} q_i(\beta) = 1$ , wenn  $p_i \neq 0$  ist, sonst verschwindet der Grenzwert; dies erklärt den Ausdruck für  $R_0$ . Für  $\alpha = \infty$  beachte man, daß  $\lim_{\beta \rightarrow \infty} q_i(\beta) = 0$  ist, wenn  $p_i < \max \{p_j | j \in \{1, \dots, n\}\}$  ist; andernfalls ist der Grenzwert  $1/m$ , wenn  $m$  die Anzahl der maximalen Elemente ist. Einsetzen in die Formel liefert die Behauptung, q. e. d.

**Satz C.12 (Monotonieeigenschaften der Rényi-Entropien)**

Für jedes  $p \in \mathcal{W}_n$  fällt  $R_\alpha(p)$  monoton für  $\alpha \in [0; \infty]$ . Umgekehrt steigt  $(\alpha - 1) \cdot \alpha^{-1} \cdot R_\alpha(p)$  monoton für  $\alpha \in [1; \infty]$ .

*Beweis:* Mittels der Quotientenregel  $(\frac{u}{v})' = \frac{u'v - uv'}{v^2}$  sowie  $\log_2 x = \ln x / \ln 2$  berechnet man zunächst

$$(1 - \alpha)^2 \cdot \ln 2 \cdot \frac{d}{d\alpha} R_\alpha(p) = \frac{\sum_{i=1}^n p_i^\alpha \ln p_i}{\sum_{j=1}^n p_j^\alpha} \cdot (1 - \alpha) - \ln \sum_{i=1}^n p_i^\alpha \cdot (-1). \quad (\text{C.8})$$

Setzt man  $q_i := p_i^\alpha / \sum_{j=1}^n p_j^\alpha$ , so gilt mit  $q = (q_1, \dots, q_n) \in \mathcal{W}_n$  weiter

$$\begin{aligned} (1 - \alpha)^2 \cdot \ln 2 \cdot \frac{d}{d\alpha} R_\alpha(p) &= \sum_{i=1}^n q_i \ln p_i \cdot (1 - \alpha) + \sum_{i=1}^n q_i \ln \sum_{k=1}^n p_k^\alpha \\ &= \sum_{i=1}^n q_i \left[ \ln p_i + \ln p_i^{-\alpha} + \ln \sum_{k=1}^n p_k^\alpha \right] \\ &= \sum_{i=1}^n q_i \left[ \ln p_i - \ln p_i^\alpha \left( \sum_{k=1}^n p_k^\alpha \right)^{-1} \right] \\ &= \sum_{i=1}^n q_i \left( \ln \frac{p_i}{q_i} \right). \end{aligned} \quad (\text{C.9})$$

Somit gilt  $dR_\alpha(p)/d\alpha = -(1 - \alpha)^{-2} H(q||p)$ , und dies ist nach Lemma C.2 nicht-positiv, so daß  $R_\alpha(p)$  monoton fällt. Für den zweiten Teil berechnet man  $(\alpha - 1) \cdot \alpha^{-1} \cdot R_\alpha(p) = -\alpha^{-1} \log_2 \sum_{i=1}^n p_i^\alpha$  für  $\alpha > 1$  und erhält hiermit

$$\begin{aligned} \frac{d}{d\alpha} \left[ \frac{-1}{\alpha} \log_2 \sum_{i=1}^n p_i^\alpha \right] &= \frac{-1}{\alpha^2} \left[ \frac{\sum_{i=1}^n p_i^\alpha \log_2 p_i}{\sum_{j=1}^n p_j^\alpha} \cdot \alpha - \log_2 \sum_{k=1}^n p_k^\alpha \right] \\ &= \frac{-1}{\alpha^2} \left[ \sum_{i=1}^n q_i \left( \log_2 p_i^\alpha - \log_2 \sum_{k=1}^n p_k^\alpha \right) \right] \\ &= \frac{-1}{\alpha^2} \left[ \sum_{i=1}^n q_i \log_2 q_i \right], \end{aligned} \quad (\text{C.10})$$

also  $(d/d\alpha)(\alpha - 1)\alpha^{-1}R_\alpha(p) = \alpha^{-2}H(q) \geq 0$ , q. e. d.

Mithilfe des Lemmas C.11 folgt nun  $0 \leq R_\infty(p) \leq R_\alpha(p) \leq R_0(p) \leq \log_2 n$ , insbesondere also  $R_\alpha(p) \in [0; \log_2 n]$ . Des weiteren gilt für  $\beta \geq \alpha > 1$  die Ungleichung  $R_\alpha \geq R_\beta \geq \frac{\beta}{\beta-1} \frac{\alpha-1}{\alpha} R_\alpha$ , woraus für  $\alpha \rightarrow 2$  und  $\beta \rightarrow \infty$  auch  $R_2 \geq R_\infty \geq \frac{1}{2}R_2$  folgt.

## C.1.6 Typische Sequenzen

Es sei  $(X_i)_{i \in \mathbb{N}}$  eine Folge identischer und unabhängig verteilter Kopien einer Zufallsvariablen  $X$ . Ist  $f$  eine beliebige Funktion, so sind auch je endlich viele der  $f(X_i)$  statistisch unabhängig; hierzu sei  $A \subseteq \mathbb{N}$  eine endliche Menge, und man berechnet

$$\begin{aligned} P \left[ \bigwedge_{i \in A} f(X_i) = x_i \right] &= P \left[ \bigwedge_{i \in A} X_i \in f^{-1}(x_i) \right] \\ &= \prod_{i \in A} P [X_i \in f^{-1}(x_i)] = \prod_{i \in A} P [f(X_i) = x_i]. \end{aligned} \quad (\text{C.11})$$

### Definition C.13 (Typische Sequenzen)

Für  $\varepsilon > 0$  nennt man eine Sequenz  $x = (x_1, \dots, x_n) \in \Omega^n$   $\varepsilon$ -typisch, wenn die Ungleichung  $2^{-n[H(X)+\varepsilon]} \leq p(x_1, \dots, x_n) \leq 2^{-n[H(X)-\varepsilon]}$  erfüllt ist. Die Menge aller  $\varepsilon$ -typischen Sequenzen einer Länge  $n \in \mathbb{N}$  werde mit  $T(n, \varepsilon)$  bezeichnet.

Bezeichnen  $H(k) := |\{i \in \{1, \dots, n\} \mid x_i = k\}|$  und  $h(k) := n^{-1}H(k)$  die absolute und die relative Häufigkeit von  $k$  unter den  $x_i$ , so gilt nach der statistischen Unabhängigkeit  $p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i)$ ; man erhält weiter  $-n^{-1} \log_2 [\prod_i p(x_i)] = -n^{-1} \log_2 \prod_k p(k)^{H(k)} = -\sum_k h(k) \log_2 p(k)$ . Eine Sequenz  $x$  ist also genau dann  $\varepsilon$ -typisch, wenn die Bedingung

$$\left| \sum_{k=1}^d [h(k) - p(k)] \log_2 p(k) \right| \leq \varepsilon \quad (\text{C.12})$$

erfüllt ist. Diese Bedingung mißt den Abstand zwischen der Verteilung und der relativen Häufigkeit gewichtet mit der logarithmierten Verteilung. Ist die Wahrscheinlichkeit  $p(k)$  sehr klein, dann werden selbst kleine Ausreißer stark gewichtet, ist sie hingegen groß, so fallen sie nicht so stark ins Gewicht.<sup>2</sup>

### Satz C.14 (Typische Sequenzen)

Es gelten die folgenden zwei Aussagen.<sup>3</sup>

1. Für  $\varepsilon > 0$  und  $\delta > 0$  und hinreichend großes  $n \in \mathbb{N}$  ist die Wahrscheinlichkeit, daß eine Sequenz  $\varepsilon$ -typisch ist, mindestens  $1 - \delta$ .
2. Für  $\varepsilon > 0$  und  $\delta > 0$  und hinreichend großes  $n \in \mathbb{N}$  gilt die Ungleichung  $(1 - \delta)2^{n[H(X)-\varepsilon]} \leq |T(n, \varepsilon)| \leq 2^{n[H(X)+\varepsilon]}$ .

<sup>2</sup>Interessant ist die Analogie zur relativen Entropie: dort wird der Quotient im Logarithmus betrachtet, hier die Differenz der Verteilungen.

<sup>3</sup>In Quantorschreibweise liest sich die erste Aussage als  $(\forall \varepsilon, \delta > 0)(\exists N_0 \in \mathbb{N}_0)(\forall n \in \mathbb{N})(n \geq N_0 \Rightarrow P[(X_1, \dots, X_n) \in T(n, \varepsilon)] \geq 1 - \delta)$ , für die zweite wird die letzte der vier runden Klammern durch  $(n \geq N_0 \Rightarrow (1 - \delta)2^{n[H(X)-\varepsilon]} \leq |T(n, \varepsilon)| \leq 2^{n[H(X)+\varepsilon]})$  ersetzt.

*Beweis:* Nach der Vorbemerkung zu typischen Sequenzen sind die  $-\log_2 X_i$  statistisch unabhängig und besitzen den Erwartungswert  $H(X)$ . Nach dem schwachen Gesetz der großen Zahlen (vgl. LEHN UND WEGMANN [106], S. 88) gilt nun  $\lim_{n \rightarrow \infty} P(|-n^{-1} \sum_{i=1}^n \log_2 X_i - H(X)| \leq \varepsilon) = 1$  für jedes  $\varepsilon > 0$ . Die zweite Aussage gilt wegen  $1 \geq \sum_{x \in T(n, \varepsilon)} p(x) \geq \sum_{x \in T(n, \varepsilon)} 2^{-n[H(X)+\varepsilon]} = |T(n, \varepsilon)| 2^{-n[H(X)+\varepsilon]}$  und  $1 - \delta \leq \sum_{x \in T(n, \varepsilon)} p(x) \leq \sum_{x \in T(n, \varepsilon)} 2^{-n[H(X)-\varepsilon]} = |T(n, \varepsilon)| 2^{-n[H(X)-\varepsilon]}$ , q. e. d.

Gibt man für jedes  $n \in \mathbb{N}$  und ein  $R < H(X)$  eine Menge  $S_n$  von höchstens  $2^{nR}$  Sequenzen der Länge  $n$  vor, so kann man zeigen, daß für jedes  $\delta > 0$  und hinreichend großes  $n$  die Ungleichung  $\sum_{x \in S_n} p(x) \leq \delta$  erfüllt ist.

Hiermit kann der *erste SHANNONSche Hauptsatz*, der *Hauptsatz der Quellenkodierung* gezeigt werden; dieser besagt, daß für eine Folge  $(X_i)_{i \in \mathbb{N}}$  unabhängiger und identisch verteilter Zufallsvariablen  $X$  ein zuverlässiges Kompressionsverfahren der Rate  $R$  existiert, wenn  $R > H(X)$  ist, und das dies für  $R < H(X)$  nicht der Fall ist; für Details vgl. z. B. NIELSEN UND CHUANG [127], S. 539–542.

## C.2 Quanteninformationstheorie

In diesem Abschnitt werden verschiedene Begriffe der Quanteninformationstheorie zusammengestellt. Die ersten Unterabschnitte behandeln die Eigenschaften der VON-NEUMANN-Entropie, zum Ende des Abschnitts wird auf die NEUMARK-Erweiterung und die UHLMANN-Fidelity eingegangen.

### C.2.1 Die von-Neumann-Entropie

Die VON-NEUMANN-Entropie ist in gewisser Weise das quantenmechanische Analogon zur SHANNON-Entropie, wurde aber von VON NEUMANN [124] schon 1932 eingeführt, also vor SHANNONS Arbeit [165] von 1948.

#### Definition C.15 (Von-Neumann-Entropie)

Es sei  $\rho \in \mathcal{S}(\mathcal{H})$  eine Dichtematrix, und  $\rho = \sum_{i=1}^n p_i |i\rangle\langle i|$  eine Spektralzerlegung dieser Matrix. Die VON-NEUMANN-Entropie ist dann durch

$$S(\rho) := -\text{Spur}(\rho \log_2 \rho) = -\sum_{i=1}^n p_i \log_2 p_i,$$

definiert, also als die SHANNON-Entropie der in ihrer Vielfachheit gezählten Eigenwerte der Dichtematrix.<sup>4</sup>

---

<sup>4</sup>Man kann auch Quanten-RÉNYI- $\alpha$ -Entropien über  $R_\alpha(\rho) = (1-\alpha)^{-1} \log_2 \text{Spur}(\rho^\alpha)$  als Entropien der Eigenwerte einer Dichtematrix definieren. Die relativen Entropien werden durch  $R_\alpha(\rho||\sigma) = (\alpha-1)^{-1} \log_2 \text{Spur}(\rho^\alpha \sigma^{1-\alpha})$  definiert; vgl. OHYA UND PETZ [131].

Die VON-NEUMANN-Entropie verschwindet genau für die reinen Zustände und nimmt ihren Maximalwert nur für  $\rho = d^{-1}\mathbb{1}_d$  an. Ist  $\rho_{AB}$  ein reiner Zustand, so folgt aus der SCHMIDT-Zerlegung, daß  $S(\rho_A) = S(\rho_B)$  für die reduzierten Zustände  $\rho_A$  und  $\rho_B$  ist. Der folgende Satz besagt, daß die VON-NEUMANN-Entropie stetig ist; vgl. AUDENAERT [5].

**Satz C.16 (Stetigkeit der von-Neumann-Entropie)**

Für Dichtematrizen  $\rho$  und  $\sigma$  auf  $\mathcal{H} = \mathbb{C}^d$  bezeichne  $T := \frac{1}{2} \|\rho - \sigma\|_1$  ihren Variationsabstand; es gilt dann stets  $|S(\rho) - S(\sigma)| \leq T \log_2(d-1) + H(T)$ , und es gibt keine nur von  $T$  und  $d$  abhängige bessere Abschätzung.

Wählt man  $\rho = \text{diag}(1-T, T \cdot (d-1)^{-1}, \dots, T \cdot (d-1)^{-1}) \in \mathbb{C}^{d \times d}$  und  $\sigma = \text{diag}(1, 0, \dots, 0) \in \mathbb{C}^{d \times d}$ , so erkennt man, daß die Ungleichung scharf ist: man berechnet dann  $\frac{1}{2} \|\rho - \sigma\|_1 = T$  und  $S(\sigma) = 0$ , und es folgt weiter  $S(\rho) = -(1-T) \log_2(1-T) - T \log_2 T + T \log_2(d-1) = H(T) + T \log_2(d-1)$ .

Im folgenden werden ähnliche Kurzschreibweisen wie in Abschnitt C.1 verwendet; bezeichnet etwa  $\rho_{12}$  eine Dichtematrix auf  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , so sind  $\rho_1$  und  $\rho_2$  reduzierte Dichtematrizen. Auch Abkürzungen wie  $S_{12} := S(\rho_{12})$ ,  $S_1 := S(\rho_1)$  usw. können auftreten.

**Satz C.17 (Eigenschaften der von-Neumann-Entropie)**

Für Dichtematrizen  $\rho_1, \dots, \rho_n \in \mathcal{S}(\mathbb{C}^d)$  und  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  bezeichne  $\rho = \sum_{i=1}^n p_i \rho_i$ . Es gilt  $S(\rho) = H(p) + \sum_i S(\rho_i)$ , sofern die Träger der  $\rho_i$  paarweise orthogonal sind; allgemein gilt  $S(\sum_i p_i |i\rangle\langle i| \otimes \rho_i) = H(p) + \sum_i S(\rho_i)$ , wenn  $(|i\rangle)_{i=1}^n$  ein Orthonormalsystem ist.

*Beweis:* Man betrachte die Spektralzerlegungen  $\rho_i = \sum_j \lambda_{ij} |e_{ij}\rangle\langle e_{ij}|$ . Da nach Voraussetzung die Träger der  $\rho_i$  orthogonal sind, sind  $p_i \lambda_{ij}$  die Eigenwerte zu den Eigenvektoren  $|e_{ij}\rangle$  von  $\rho$ , so daß

$$\begin{aligned} S(\rho) &= - \sum_{ij} p_i \lambda_{ij} \log_2 p_i \lambda_{ij} = - \sum_{ij} p_i \lambda_{ij} \log_2 p_i - \sum_{ij} p_i \lambda_{ij} \log_2 \lambda_{ij} \\ &= - \sum_i p_i \log_2 p_i - \sum_i p_i \sum_j \lambda_{ij} \log_2 \lambda_{ij} = H(p) + \sum_i p_i S(\rho_i) \end{aligned}$$

gilt. Ferner ist  $S(|i\rangle\langle i| \otimes \rho_i) = S(|i\rangle\langle i|) + S(\rho_i) = S(\rho_i)$ , q. e. d.

## C.2.2 Grundbegriffe der Quanteninformationstheorie

Analog zum klassischen Fall definiert man die *bedingte* VON-NEUMANN-Entropie über  $S(A|B) := S(\rho_{AB}) - S(\rho_B)$  und die *gemeinsame Information* mittels  $S(A : B) := S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ ; für die *relative* VON-NEUMANN-Entropie setzt man  $S(\rho||\sigma) := \text{Spur } \rho \log_2 \rho - \text{Spur } \rho \log_2 \sigma$ . Aus dem nächsten Satz folgt eine Reihe von Eigenschaften der Entropie.

**Satz C.18 (Kleinsche Ungleichung)**

Es gilt  $S(\rho\|\sigma) \geq 0$ , und Gleichheit gilt genau für  $\rho = \sigma$ .

*Beweis:* Für Spektralzerlegungen  $\rho = \sum_i p_i |i\rangle\langle i|$  und  $\sigma = \sum_j q_j |j\rangle\langle j|$  ist die Matrix  $P = (P_{ij})_{ij} \in \mathbb{R}^{n \times n}$ ,  $P_{ij} := |\langle i|j\rangle|^2$ , doppelt-stochastisch, das heißt, es ist  $P_{ij} \geq 0$  und  $\sum_i P_{ij} = \sum_j P_{ij} = 1$  für alle  $i$  bzw.  $j$ . Es folgt

$$\sum_i \langle i|\log_2 \sigma|i\rangle = \sum_i \langle i|(\log_2 q_j |j\rangle\langle j|)|i\rangle = \sum_{ij} P_{ij} \log_2 q_j, \quad (\text{C.13})$$

was mit  $\langle i|\rho = p_i \langle i|$  auf die Gleichung  $S(\rho\|\sigma) = \sum_i p_i \log_2 p_i - \sum_i \langle i|\rho \log_2 \sigma|i\rangle = \sum_i p_i (\log_2 p_i - \sum_j P_{ij} \log_2 q_j)$  führt. Der Logarithmus ist streng konkav, so daß  $\sum_j P_{ij} \log_2 q_j \leq \log_2 \sum_j P_{ij}$  für jedes  $i$  gilt; Gleichheit gilt genau dann, wenn genau ein  $P_{ij} = 1$  ist und alle anderen verschwinden. Setzt man  $r_i := \sum_j P_{ij} q_j$ , so folgt hieraus  $S(\rho\|\sigma) \geq \sum_i p_i \log_2 (p_i/r_i)$  mit Gleichheit genau dann, wenn  $P$  eine Permutationsmatrix ist. Dieser Ausdruck hat die Form einer klassischen relativen Entropie, so daß die rechte Seite nicht-negativ ist und genau für  $p = r$  verschwindet. Dies zeigt  $S(\rho\|\sigma) \geq 0$ , und Gleichheit gilt genau dann, wenn die (bis auf Permutationen eindeutigen) Spektralzerlegungen gleich sind, q. e. d.

Man kann für Zustände  $\omega$  und  $\varphi$  auf einer endlichdimensionalen  $C^*$ -Algebra sogar die stärkere Ungleichung  $S(\omega\|\varphi) \geq \|\omega - \varphi\|_1^2/2$  zeigen; vgl. OHYA UND PETZ [131], Theorem 1.15 auf S. 27, in Verbindung mit Lemma B.44.

**Satz C.19 (Relative von-Neumann-Entropie und Subadditivität)**

Für  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  gilt  $S(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}\|\rho_A \otimes \rho_B)$ .

Insbesondere gilt die Subadditivität  $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ ; Gleichheit gilt nach Satz C.18 genau für  $\rho_{AB} = \rho_A \otimes \rho_B$ , selbst klassische Korrelationen sind nicht zugelassen. Umgekehrt folgt aus der Subadditivität auch die *Dreiecks- oder ARAKI-LIEB-Ungleichung*  $|S(\rho_A) - S(\rho_B)| \leq S(\rho_{AB})$ : purifiziert man  $\rho_{AB}$  zu  $\rho_{ABR}$ , so gilt  $S(\rho_{AR}) \leq S(\rho_A) + S(\rho_R)$ , und weil  $\rho_{ABR}$  rein ist, sind ferner  $S(\rho_{AR}) = S(\rho_B)$  und  $S(\rho_R) = S(\rho_{AB})$ ; die Gleichheitsbedingungen sind hier aber schwieriger zu formulieren.

*Beweis:* Mit der Spektralzerlegung  $\rho_{AB} = \sum_{ij} p_{ij} |i\rangle\langle i| \otimes |j\rangle\langle j|$  und der Größe<sup>5</sup>  $R := \text{Spur } \rho_{AB} \log_2 \rho_A \otimes \rho_B = \text{Spur}(\rho_{AB} \log_2 \rho_A \otimes \mathbb{1}) + \text{Spur}(\rho_{AB} \log_2 \mathbb{1} \otimes \rho_B)$  berechnet man  $S(\rho_{AB}\|\rho_A \otimes \rho_B) = \text{Spur } \rho_{AB} \log_2 \rho_{AB} - \text{Spur } \rho_{AB} \log_2 \rho_A \otimes \rho_B = -S(\rho_{AB}) - R$ .

---

<sup>5</sup>Beachte hierzu  $\log_2(\rho \otimes \sigma) = \log_2(\rho \otimes \mathbb{1}) + \log_2(\mathbb{1} \otimes \sigma) = (\log_2 \rho) \otimes \mathbb{1} + (\log_2 \mathbb{1}) \otimes \sigma$ : für  $\rho = \sum_i \lambda_i |i\rangle\langle i|$  und  $\sigma = \sum_j \mu_j |j\rangle\langle j|$  erhält man  $\log_2(\rho \otimes \sigma) = \log_2 \sum_{ij} \lambda_i \mu_j |i\rangle\langle i| \otimes |j\rangle\langle j| = \sum_{ij} (\log_2 \lambda_i + \log_2 \mu_j) |i\rangle\langle i| \otimes |j\rangle\langle j|$  nach dem Funktionalkalkül; weiterhin gilt zum Beispiel für den ersten Summanden  $\sum_{ij} \log_2 \lambda_i |i\rangle\langle i| \otimes |j\rangle\langle j| = \sum_i \log_2 \lambda_i |i\rangle\langle i| \otimes \sum_j |j\rangle\langle j| = \log_2 \sum_i \lambda_i |i\rangle\langle i| \otimes \mathbb{1} = \log_2 \rho \otimes \mathbb{1}$ .

Es gilt nun weiter  $\rho_{AB} \log_2 \rho_A \otimes \mathbb{1} = \sum_{ij} p_{ij} |i\rangle\langle i| \otimes |j\rangle\langle j| \cdot \log_2 \sum_k p_{k*} |k\rangle\langle k| \otimes \mathbb{1}$   
 $= \sum_{ij} p_{ij} \log_2 p_{i*} |i\rangle\langle i| \otimes |j\rangle\langle j| = \sum_i p_{i*} \log_2 p_{i*} |i\rangle\langle i| \otimes \sigma_i$  für  $\sigma_i := \sum_j p_{j|i} |j\rangle\langle j|$ .  
 Man erhält  $\text{Spur}(\rho_{AB} \log_2 \rho_A \otimes \mathbb{1}) = \text{Spur}(\sum_i p_{i*} \log_2 p_{i*} |i\rangle\langle i|)$  wegen der  
 Multiplikativität der Spur unter Tensorprodukten, und somit gilt

$$\begin{aligned} R &= \text{Spur} \left( \sum_i p_{i*} \log_2 p_{i*} |i\rangle\langle i| \right) + \text{Spur} \left( \sum_j p_{*j} \log_2 p_{*j} |j\rangle\langle j| \right) \\ &= \text{Spur} \rho_A \log_2 \rho_A + \text{Spur} \rho_B \log_2 \rho_B = -S(\rho_A) - S(\rho_B), \end{aligned} \quad (\text{C.14})$$

was die Behauptung zeigt, q. e. d.

**Lemma C.20 (Bedingte Entropie und Verschränkung)**

Ein Zustand  $|\Psi\rangle_{AB}$  ist genau dann verschränkt, wenn  $S(\rho_B|\rho_A) < 0$  ist.

*Beweis:* Man betrachte die SCHMIDT-Zerlegung  $|\Psi\rangle_{AB} = \sum_{i=1}^n \lambda_i |i\rangle_A \otimes |i\rangle_B$ ,  
 aus der  $\rho_A = \sum_{i=1}^n \lambda_i^2 |i\rangle\langle i|$  folgt. Somit gilt  $S(B|A) = S(A, B) - S(A)$   
 $= 0 - H(\lambda_1, \dots, \lambda_n)$ , und die SHANNON-Entropie verschwindet genau für  
 die SCHMIDT-Zahl 1, also wenn  $|\Psi\rangle_{AB}$  separabel ist, q. e. d.

**Satz C.21 (Entropieveränderung unter Messungen)**

Sind  $\rho$  ein Quantenzustand und  $(P_\mu)_{\mu=1}^n$  eine VON-NEUMANN-Messung und  
 bezeichnet  $\rho' := \sum_{\mu=1}^n P_\mu \rho P_\mu$  den Zustand nach Ausführung der Messung  
 ohne Bekanntwerden des Ergebnisses, so gilt stets

$$S(\rho') \geq S(\rho),$$

wobei Gleichheit beider Seiten genau dann vorliegt, wenn  $\rho = \rho'$  ist.

Die entsprechenden Aussagen für verallgemeinerte Messungen sind nicht all-  
 gemein gültig.

*Beweis:* Zum Beweis verwendet man die KLEINSche Ungleichung (Satz C.18),  
 aus der  $0 \leq S(\rho'|\rho) = -S(\rho) - \text{Spur}(\rho \log_2 \rho') = -S(\rho) + S(\rho')$  folgt, wobei  
 die letzte Gleichheit noch zu begründen ist. Hierzu berechnet man

$$\begin{aligned} \text{Spur}(\rho \log_2 \rho') &= \text{Spur} \left( \sum_{\mu=1}^n P_\mu \rho \log_2 \rho' \right) = \text{Spur} \left( \sum_{\mu=1}^n P_\mu^2 \rho \log_2 \rho' \right) \\ &= \text{Spur} \left( \sum_{\mu=1}^n P_\mu \rho \log_2 \rho' P_\mu \right) = \text{Spur} \left( \sum_{\mu=1}^n P_\mu \rho P_\mu \log_2 \rho' \right) \\ &= \text{Spur}(\rho' \log_2 \rho') = -S(\rho'), \end{aligned} \quad (\text{C.15})$$

wozu man die Eigenschaften  $\sum_{\mu=1}^n P_\mu = \mathbb{1}$  und  $P_\mu = P_\mu^2$ , die Invarianz der  
 Spur unter zyklischen Vertauschungen und die Tatsache verwendet, daß auf-  
 grund von  $P_\mu P_{\mu'} = \delta_{\mu\mu'} P_\mu$  und

$$P_\mu \rho' = P_\mu \cdot \left( \sum_{\mu'=1}^n P_{\mu'} \rho P_{\mu'} \right) = P_\mu \rho P_\mu = \left( \sum_{\mu'=1}^n P_{\mu'} \rho P_{\mu'} \right) \cdot P_\mu = \rho' P_\mu$$

auch  $\log_2 \rho'$  und  $P_\mu$  kommutieren, q. e. d.

**Satz C.22 (Strenge Konkavität der Entropie)**

Sind  $\rho_1, \dots, \rho_n$  Zustände und  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$ , so gilt

$$\sum_{i=1}^n p_i S(\rho_i) \leq S\left(\sum_{i=1}^n p_i \rho_i\right);$$

beide Seiten sind genau dann gleich, wenn alle  $\rho_i$  übereinstimmen, die zu den Werten  $p_i > 0$  gehören.

Umgekehrt gilt  $S(\sum_{i=1}^n p_i \rho_i) \leq H(p) + \sum_{i=1}^n p_i S(\rho_i)$  mit Gleichheit genau dann, wenn die Träger der  $\rho_i$  paarweise orthogonal sind; vgl. NIELSEN UND CHUANG [127], S. 518–519.

*Beweis:* Zum Beweis führe man ein Orthonormalsystem  $(|i\rangle)_{i=1}^n$  ein und definiere  $\rho_{AB} = \sum_{i=1}^n p_i \rho_i \otimes |i\rangle\langle i|$ . Es folgt dann  $S(\rho_{AB}) = H(p) + \sum_{i=1}^n p_i S(\rho_i)$  aus Satz C.17 sowie

$$S(\rho_A) = S\left(\sum_{i=1}^n p_i \rho_i\right) \quad \text{und} \quad S(\rho_B) = S\left(\sum_{i=1}^n p_i |i\rangle\langle i|\right) = H(p).$$

Die Subadditivität aus Satz C.19 liefert also  $S(\rho_A) \geq S(\rho_{AB}) - S(\rho_B) = \sum_{i=1}^n p_i S(\rho_i)$ , wobei die Gleichheit genau dann gilt, wenn  $\rho_{AB} = \rho_A \otimes \rho_B$  ist, was der Fall ist, wenn alle  $\rho_i$  zu einem  $p_i > 0$  gleich sind, q. e. d.

### C.2.3 Die starke Subadditivität

Die starke Subadditivität der VON-NEUMANN-Entropie ist eine der wichtigsten Eigenschaften der Quanteninformationstheorie, der Beweis ist allerdings – im Gegensatz zum klassischen Fall in Satz C.7 – zu schwierig, um an dieser Stelle wiedergegeben zu werden; vgl. RUSKAI [154] und die dort zitierten Arbeiten. Es folgt ein Satz, aus dem für den Spezialfall, daß  $\Phi$  die Bildung der Teilspur ist, eine Reihe weiterer Ungleichungen folgen, die anschließend angegeben werden.

**Satz C.23 (Eine Monotonieeigenschaft der relativen Entropie)**

Für jede Quantenoperation  $\Phi$  und jedes Paar positiv semidefiniter Matrizen  $\rho$  und  $\sigma$  gilt die Ungleichung  $S(\Phi(\rho) \|\Phi(\sigma)) \leq S(\rho \|\sigma)$ ; Gleichheit beider Seiten liegt genau dann vor, wenn  $\ln \rho - \ln \sigma = \Phi^\dagger[\ln \Phi(\rho) - \ln \Phi(\sigma)]$  ist.

Zu beachten ist hierbei, daß  $\Phi$  ein Operator auf Operatoren ist und die Adjunktion entsprechend zu interpretieren ist. Die Aussage des Satzes ist, daß die relative Entropie zweier Dichtematrizen unter Quantenoperationen kontraktiv ist.

Multipliziert man beide Seiten der Gleichheitsbedingung mit  $\rho$  und bildet die Spur, wendet man also  $\langle \rho | \cdot \rangle$  auf sie an, so folgt

$$\begin{aligned} S(\rho||\sigma) &= (\ln 2)^{-1} \langle \rho | \Phi^\dagger [\ln \Phi(\rho) - \ln \Phi(\sigma)] \rangle \\ &= (\ln 2)^{-1} \langle \Phi(\rho) | \ln \Phi(\rho) - \ln \Phi(\sigma) \rangle = S(\Phi(\rho)||\Phi(\sigma)), \end{aligned} \quad (\text{C.16})$$

die genannte Bedingung ist also zumindest hinreichend für die Gleichheit.

Äquivalent zum vorstehenden Satz sind unter anderem (1) die *starke Subadditivität*  $S_{123} + S_2 \leq S_{12} + S_{23}$ , (2) die Ungleichung  $S_1 + S_3 \leq S_{12} + S_{23}$ , (3) die Ungleichung  $S(\rho_{12}||\rho_2) \leq S(\rho_{123}||\rho_{23})$  für die relative Entropie, (4) die gemeinsame Konvexität<sup>6</sup> der relativen Entropie  $S(\rho||\sigma)$ , (5) die Konvexität der Abbildung  $\rho_{12} \mapsto S_1 - S_{12}$  und (6) die Nicht-Negativität der bedingten gemeinsamen Information  $H(X : Y|Z) := H(X|Z) + H(Y|Z) - H(X, Y|Z)$ .

Zum Beweis der Äquivalenz der ersten beiden Aussagen wähle man eine Purifizierung  $\rho_{1234}$  des Gesamtzustands  $\rho_{123}$ ; die SCHMIDT-Zerlegung liefert dann  $S_4 = S_{123}$  sowie  $S_{13} = S_{24}$ . Die Behauptung ist nun aus der Äquivalenz  $S_4 + S_3 \leq S_{24} + S_{23} \Leftrightarrow S_{123} + S_3 \leq S_{13} + S_{23}$  ersichtlich. Für die Äquivalenz der ersten und dritten Aussage berechnet man für  $s := \text{Spur}(\rho_{12} \log_2 \rho_2)$  nun

$$\begin{aligned} s &= \text{Spur} \left[ \sum_{ij} p_{ij} |i\rangle\langle i| \otimes |j\rangle\langle j| \cdot \log_2 \sum_{i'j'} p_{*j'} |i'\rangle\langle i'| \otimes |j'\rangle\langle j'| \right] \quad (\text{C.17}) \\ &= \text{Spur} \left[ \sum_{ij} p_{ij} \log_2 p_{*j} |i\rangle\langle i| \otimes |j\rangle\langle j| \right] = \left[ \sum_{ij} p_{ij} \log_2 p_{*j} \right] = -S(\rho_2), \end{aligned}$$

woraus sich  $S(\rho_{12}||\rho_2) = \text{Spur} \rho_{12} \log_2 \rho_{12} - \text{Spur} \rho_{12} \log_2 \rho_2 = -S(\rho_{12}) + S(\rho_2)$  ergibt; analog folgt  $S(\rho_{123}||\rho_{23}) = -S(\rho_{123}) + S(\rho_{23})$ .

## C.2.4 Die Holevo-Schranke und das HSW-Theorem

Man betrachte die folgende Fragestellung: eine Quelle  $A$  erzeugt entsprechend der Realisierung einer Zufallsvariablen  $X$  mit Wertebereich  $\{1, \dots, n\}$  und einer vorgegebenen Wahrscheinlichkeitsverteilung  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  die Quantenzustände  $\rho_1, \dots, \rho_n$ . Während die Verteilung und die Quantenzustände bekannt sind, ist die Realisierung verborgen.

Ein Proband hat nun die Aufgabe, unter Zuhilfenahme der Ergebnisse einer von ihm gewählten quantenmechanisch möglichen (verallgemeinerten)

---

<sup>6</sup>Eine Funktion  $f : X \times Y \rightarrow \mathbb{R}$  in zwei Veränderlichen nennt man *gemeinsam konvex*, falls die Ungleichung  $f(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) \leq \lambda f(x_1, y_1) + (1 - \lambda)f(x_2, y_2)$  für  $x_1, x_2 \in X$ ,  $y_1, y_2 \in Y$  und  $\lambda \in [0, 1]$  stets erfüllt ist. Setzt man  $y_1 = y_2$  oder  $x_1 = x_2$ , so stellt man fest, daß eine gemeinsam konvexe Funktion in jeder einzelnen Veränderlichen konvex ist; daß die Umkehrung nicht gilt, erkennt man an der Funktion  $f(x, y) := (x^2 + y^2) - 4xy$ , deren zweite Ableitungen beide positiv sind, die aber auf der Diagonalen  $x = y$  nicht konvex ist.

Messung des erzeugten Zustands eine Zufallsvariable  $Y$  derart zu konstruieren, daß  $H(X : Y)$  möglichst groß wird.

Zu diesem Zweck definiert der Proband positive Operatoren  $A_1, \dots, A_n$ , die er für eine Messung verwendet, wobei auch  $A_0 := \mathbb{I} - \sum_{i=1}^n A_i$  positiv sein muß. Wird  $A_i$  gemessen, so möge  $Y := i$  ausgegeben werden. Die Messung von  $A_0$  wird als Fehler betrachtet.

Die Quantenmechanik liefert nun  $P(Y = j | X = i) = \text{Spur } \rho_i A_j$  für  $j \neq 0$ , die Wahrscheinlichkeit eines Fehlers in diesem Fall ist somit  $P(Y \neq i | X = i) = \text{Spur } [\rho_i (\mathbb{I} - A_i)] = 1 - \text{Spur } \rho_i A_i$ .

**Definition C.24 (Holevo-Größe)**

Die Funktion  $\chi : \bigcup_{n \in \mathbb{N}} \mathcal{W}_n \times \mathcal{S}(\mathcal{H})^n \rightarrow \mathbb{R}_0^+$ , welche durch

$$\chi := \chi \left[ (p_1, \dots, p_n), (\rho_1, \dots, \rho_n) \right] := S \left( \sum_{i=1}^n p_i \rho_i \right) - \sum_{i=1}^n p_i S(\rho_i)$$

definiert ist, wird als HOLEVO-Größe bezeichnet.

Eine andere Formulierung der HOLEVO-Größe liefert die Rechnung

$$\begin{aligned} \chi &= S(\rho) - \sum_{i=1}^n p_i S(\rho_i) = - \sum_{i=1}^n p_i S(\rho_i) - \text{Spur} \left( \sum_{i=1}^n p_i \rho_i \right) \log_2 \rho \\ &= \sum_{i=1}^n p_i [\text{Spur } \rho_i \log_2 \rho_i] - \sum_{i=1}^n p_i [\text{Spur } \rho_i \log_2 \rho] \\ &= \sum_{i=1}^n p_i [\text{Spur } \rho_i \log_2 \rho_i - \text{Spur } \rho_i \log_2 \rho] = \sum_{i=1}^n p_i S(\rho_i || \rho). \end{aligned} \quad (\text{C.18})$$

Wählt man in der Entropie den Logarithmus zur Basis 2 und betrachtet ein festes  $n \in \mathbb{N}$ , so nimmt die HOLEVO-Größe Werte in  $[0; \log_2 n]$  an. Die untere Grenze folgt aus der Konkavität der Entropie, die obere Grenze wird angenommen, wenn der Minuend maximal, also gleich  $\log_2 n$  ist, und der Subtrahend minimal ist, also verschwindet. Dies ist genau dann der Fall, wenn alle  $\rho_i$  rein sind und  $\sum_{i=1}^n p_i \rho_i = \mathbb{I}/d$  ist, wenn also die  $\rho_i$  Projektionen auf eine Orthonormalbasis und  $p_i = 1/n$  für alle  $i \in \{1, \dots, n\}$  sind. HOLEVO [77] zeigte den folgenden Satz.

**Satz C.25 (Holevo-Schranke)**

Eine Quelle präpariere gemäß der unbekanntem Realisierung einer klassischen Zufallsvariable  $X$  mit Wahrscheinlichkeitsverteilung  $p = (p_1, \dots, p_n) \in \mathcal{W}_n$  die Zustände  $\rho_1, \dots, \rho_n$ . Bestimmt man durch verallgemeinerte Messung des Zustands eine Zufallsvariable  $Y$ , so gilt

$$H(X : Y) \leq \chi \left[ (p_1, \dots, p_n), (\rho_1, \dots, \rho_n) \right],$$

und die Gleichheit beider Seiten kann nur dann erzielt werden, wenn alle Zustände  $\rho_1, \dots, \rho_n$  kommutieren.

Der ursprüngliche Beweis von HOLEVO [77] ist recht schwierig; ein einfacher Beweis unter Verwendung der starken Subadditivität der VON-NEUMANN-Entropie – allerdings ohne die Bedingungen für die Gleichheit – findet sich bei NIELSEN UND CHUANG [127], S. 531–534.

Wie oben erwähnt wurde, ist die HOLEVO-Schranke im allgemeinen nicht erreichbar. Es zeigt sich aber, daß unter bestimmten Umständen eine etwas schwächere Aussage getroffen werden kann. Nach den fast gleichzeitig erschienenen Arbeiten von HOLEVO [78] und von SCHUMACHER UND WESTMORELAND [163] wird sie als das *HSW-Theorem* bezeichnet; vgl. zum Beweis von HOLEVO auch NIELSEN UND CHUANG [127], Theorem 12.8 auf S. 555–560.

Was man leicht erkennt, ist, daß die HOLEVO-Größe einer aus zwei unabhängigen Quellen zusammengesetzten Quelle gleich der Summe der einzelnen HOLEVO-Größen ist: es gilt  $\chi_{AB} = S(\sum_{ij} p_i q_j \rho_i \otimes \sigma_j) - \sum_{ij} p_i q_j S(\rho_i \otimes \sigma_j) = S(\sum_i p_i \rho_i \otimes \sum_j q_j \sigma_j) - \sum_{ij} p_i q_j (S(\rho_i) + S(\sigma_j)) = S(\sum_i p_i \rho_i) + S(\sum_j q_j \sigma_j) - \sum_i p_i S(\rho_i) - \sum_j q_j S(\sigma_j) = \chi_A + \chi_B$ ; insbesondere ist  $\chi_n = n\chi_1$ , wenn man  $n \in \mathbb{N}$  unabhängige Realisierungen einer Quelle hat. Das HSW-Theorem besagt, daß diese Schranke auch erreicht werden kann, wenn man den Sender geeignete Produktzustände  $\rho_1 \otimes \rho_2 \otimes \dots$  präparieren läßt (klassische Kapazität eines Quantenkanals).

## C.2.5 Die Neumark-Erweiterung

Die NEUMARK-Erweiterung [125] besagt, daß jede verallgemeinerte Messung als VON-NEUMANN-Messung auf einem höherdimensionalen Raum interpretiert werden kann. Ist  $(P_\mu)_{\mu=1}^n$  eine Zerlegung des Einheitsoperators, die eine verallgemeinerte Messung bildet, so kann jedes  $P_\mu$  spektral zerlegt werden, so daß man sich auf  $P_\mu^{(i)} = |v_\mu^{(i)}\rangle\langle v_\mu^{(i)}|$  für nicht notwendig normierte Vektoren  $|v_\mu^{(i)}\rangle$  beschränken kann.

### Lemma C.26 (Neumark-Erweiterung)

*Erfüllen Vektoren  $|v_1\rangle, \dots, |v_n\rangle \in \mathbb{C}^d$  die Bedingung  $\sum_{\alpha=1}^n |v_\alpha\rangle\langle v_\alpha| = \mathbb{I}_d$ , so gibt es eine lineare Abbildung  $\varphi: \mathbb{C}^d \rightarrow \mathbb{C}^n$  sowie eine Orthonormalbasis  $\{|\bar{v}_1\rangle, \dots, |\bar{v}_n\rangle\} \subseteq \mathbb{C}^n$ , so daß  $P|\bar{v}_k\rangle = \varphi(|v_k\rangle)$  für alle  $k \in \{1, \dots, n\}$  erfüllt ist, wenn  $P$  die Orthogonalprojektion von  $\mathbb{C}^n$  auf  $\varphi(\mathbb{C}^d)$  bezeichnet.*

*Beweis:* Im folgenden seien  $(|e_j\rangle)_{j=1}^d$  und  $(|\bar{e}_\alpha\rangle)_{\alpha=1}^n$  Orthonormalbasen von  $\mathbb{C}^d$  und  $\mathbb{C}^n$  derart, daß  $\varphi|e_j\rangle = |\bar{e}_j\rangle$  für alle  $j \in \{1, \dots, d\}$  gilt. Die Entwicklung der Vektoren  $|v_\alpha\rangle$  in  $\mathbb{C}^d$  für  $\alpha \in \{1, \dots, n\}$  ergibt sich hiermit zu

$$|v_\alpha\rangle = \sum_{j=1}^d c_{\alpha j} |e_j\rangle \quad \text{mit} \quad c_{\alpha j} := \langle e_j | v_\alpha \rangle. \quad (\text{C.19})$$

Hieraus folgt

$$\sum_{\alpha=1}^n c_{\alpha j} c_{\alpha k}^* = \sum_{\alpha=1}^n \langle e_j | v_\alpha \rangle \langle v_\alpha | e_k \rangle = \langle e_j | \left[ \sum_{\alpha=1}^n |v_\alpha\rangle \langle v_\alpha| \right] | e_k \rangle = \delta_{jk}.$$

Betrachtet man eine Matrix der Form

$$C = \left( \begin{array}{cccc|ccc} c_{11} & c_{12} & \cdots & c_{1d} & c_{1,d+1} & \cdots & c_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nd} & c_{n,d+1} & \cdots & c_{nn} \end{array} \right) \in \mathbb{C}^{n \times n}, \quad (\text{C.20})$$

so bedeutet dies, daß der linke Teil als ein Orthonormalsystem auf  $\mathbb{C}^n$  interpretiert werden kann, während der rechte Teil durch das GRAM-SCHMIDT-Verfahren so ergänzt werden kann, daß eine Orthonormalbasis vorliegt, das heißt, daß  $C$  unitär ist. Setzt man nun

$$|\bar{v}_\alpha\rangle := C|\bar{e}_\alpha\rangle = \sum_{\beta=1}^n c_{\alpha\beta} |\bar{e}_\beta\rangle \text{ für } \alpha \in \{1, \dots, n\}, \quad (\text{C.21})$$

so bilden diese Vektoren ebenfalls eine Orthonormalbasis, und wie gefordert gilt  $P|\bar{v}_\alpha\rangle = \sum_{\beta=1}^d c_{\alpha\beta} |\bar{e}_\beta\rangle =: \varphi(|v_\alpha\rangle)$  für alle  $\alpha \in \{1, \dots, n\}$ , q. e. d.

**Lemma C.27 (Maximale Anzahl der Kraus-Operatoren)**

Jede Quantenoperation  $\mathcal{E}$  auf  $\mathcal{H} = \mathbb{C}^d$  kann durch höchstens  $d^2$  KRAUS-Operatoren dargestellt werden, also  $\mathcal{E}(\rho) = \sum_{k=1}^M E_k \rho E_k^\dagger$  für  $M \in \{1, \dots, d^2\}$ .

*Beweis:* Zunächst gibt es zu jeder Quantenoperation einen Satz geeigneter Kraus-Operatoren  $(E_i)_{i=1}^n$  mit Matrixdarstellungen  $E_j = (e_{ab}^{(j)})_{a,b=1}^d$ . Man definiert nun eine weitere Matrix  $W = (W_{jk})_{j,k=1}^n \in \mathbb{C}^{n \times n}$  mittels der Festsetzungen  $W_{jk} = \text{Spur}(E_j^\dagger E_k) = \sum_{a,b=1}^d e_{ab}^{(j)*} e_{ab}^{(k)}$  und erhält somit

$$W = \sum_{a,b} \underbrace{\begin{pmatrix} e_{ab}^{(1)*} e_{ab}^{(1)} & \cdots & e_{ab}^{(1)*} e_{ab}^{(n)} \\ \vdots & & \vdots \\ e_{ab}^{(n)*} e_{ab}^{(1)} & \cdots & e_{ab}^{(n)*} e_{ab}^{(n)} \end{pmatrix}}_{\text{dyadisches Produkt}}. \quad (\text{C.22})$$

Die Summanden der rechten Seite sind dyadische Produkte. Setzt man also  $|e_{ab}\rangle := (e_{ab}^{(1)}, \dots, e_{ab}^{(n)})^t$ , so erhält man die Zerlegung  $W = \sum_{a,b=1}^d |e_{ab}\rangle \langle e_{ab}|$ . Jeder Summand hierbei ist ein Vielfaches einer Rang-1-Projektion, und da die Summe  $d^2$  Elemente hat, ist  $\text{Rang } W \leq d^2$ , q. e. d.

**Korollar C.28 (Allgemeine Messungen auf Quantensystemen)**

Jede verallgemeinerte Messung auf einem  $\mathcal{H} \cong \mathbb{C}^d$  kann als VON-NEUMANN-MESSUNG auf  $\mathcal{H}' \cong \mathbb{C}^{d^2}$  dargestellt werden.

*Beweis:* Nach Lemma C.27 kann jede verallgemeinerte Messung aus nicht mehr als  $d^2$  Operatoren gebildet werden, und nach Lemma C.26 werden hierfür  $d^2 - d = d(d - 1)$  zusätzliche Dimensionen gebraucht, q. e. d.

## C.2.6 Die Subadditivität der Spurnorm

Der quantenmechanische Variationsabstand ist durch  $\delta(\rho, \sigma) := \frac{1}{2} \|\rho - \sigma\|_1$  als die Hälfte der Spurnorm der Differenz zweier Dichtematrizen definiert; vgl. Unterabschnitt B.8.4 für das klassische Analogon.

### Lemma C.29 (Subadditivität der Spurnorm)

Sind  $\mathcal{H}$  und  $\mathcal{H}'$  Hilberträume, so gilt für beliebige Zustände  $\rho, \sigma \in \mathcal{S}(\mathcal{H})$  und  $\rho', \sigma' \in \mathcal{S}(\mathcal{H}')$  die Ungleichung

$$\delta(\rho \otimes \rho', \sigma \otimes \sigma') \leq \delta(\rho, \sigma) + \delta(\rho', \sigma'),$$

und Gleichheit gilt genau dann, wenn  $\delta(\rho, \sigma) = 0$  oder  $\delta(\rho', \sigma') = 0$  ist.

*Beweis:* Es wird zunächst  $\delta(P \times P', Q \times Q') \leq \delta(P, Q) + \delta(P', Q')$  für klassische Wahrscheinlichkeitsverteilungen gezeigt; es seien hierzu  $P = (p_i)_{i=1}^n \in \mathcal{W}_n$ ,  $Q = (q_i)_{i=1}^n \in \mathcal{W}_n$ ,  $P' = (p'_j)_{j=1}^m \in \mathcal{W}_m$  und  $Q' = (q'_j)_{j=1}^m \in \mathcal{W}_m$ , wobei man  $n := \dim \mathcal{H}$  und  $m := \dim \mathcal{H}'$  setzt. Die beiden gemeinsamen Verteilungen als klassische Versionen des Tensorproduktes sind  $P \times P' = (p_i p'_j)_{i=1, j=1}^n, m \in \mathcal{W}_{n \cdot m}$  und  $Q \times Q' = (q_i q'_j)_{i=1, j=1}^n, m \in \mathcal{W}_{n \cdot m}$ . Die Rechnung

$$\begin{aligned} \sum_{ij} |p_i p'_j - q_i q'_j| &= \sum_{ij} |p_i p'_j - q_i p'_j + q_i p'_j - q_i q'_j| \\ &\leq \sum_{ij} p'_j |p_i - q_i| + \sum_{ij} q_i |p'_j - q'_j| \\ &= \sum_i |p_i - q_i| + \sum_j |p'_j - q'_j| \end{aligned} \quad (\text{C.23})$$

zeigt, daß die Behauptung für klassische Wahrscheinlichkeitsverteilungen gilt. Mittels  $\delta(\rho \otimes \rho', \sigma \otimes \sigma') = \max \delta(P \times P', Q \times Q') \leq \max \delta(P, Q) + \max \delta(P', Q') = \delta(\rho, \sigma) + \delta(\rho', \sigma')$  folgt die Behauptung, wenn die Bedeutung der jeweiligen Maxima geklärt wird. Im ersten Fall wird über verallgemeinerte Messungen auf  $\mathcal{H} \otimes \mathcal{H}'$  maximiert, im zweiten über  $\mathcal{H}$  und  $\mathcal{H}'$  getrennt; vgl. NIELSEN UND CHUANG [127], Theorem 9.1 auf S. 405. Letztere ergeben sich als Teilspuren der ersteren, so daß die Maximierungen übereinstimmen, q. e. d.

Für Zustände, die keine Produktform besitzen, gilt die Aussage jedoch nicht: hierzu betrachte man  $P = (p_{ij})_{i=1, j=1}^n, m \in \mathcal{W}_{n \cdot m}$  und  $Q = (q_{ij})_{i=1, j=1}^n, m \in \mathcal{W}_{n \cdot m}$  mit Werten<sup>7</sup>  $p_{ij} := (nm)^{-1} [1 + \frac{1}{2}(-1)^{i+j}]$  und  $q_{ij} := (nm)^{-1} [1 - \frac{1}{2}(-1)^{i+j}]$ . Zu falsifizieren ist dann  $\sum_{ij} |p_{ij} - q_{ij}| \leq \sum_i |\sum_j (p_{ij} - q_{ij})| + \sum_j |\sum_i (p_{ij} - q_{ij})|$ . Die linke Seite ist Eins, für die rechte gilt  $(nm)^{-1} \cdot n \cdot g(m) + (nm)^{-1} \cdot m \cdot g(n) \leq (nm)^{-1} \cdot (n + m) = m^{-1} + n^{-1}$ , wenn  $g(n) = 1$  für ungerade Zahlen ist und sonst verschwindet. In den sinnvollen Fällen  $n, m \geq 2$  zeigt dies, daß die Ungleichung verletzt ist.

---

<sup>7</sup>Sind sowohl  $n$  als auch  $m$  ungerade, so sind addiere bzw. man den Term  $(2n^2m^2)^{-1}$  zu  $p_{ij}$  und  $q_{ij}$ , da andernfalls keine Wahrscheinlichkeitsverteilungen vorliegen.

### C.2.7 Die Uhlmann-Fidelity

Ein in der Quanteninformationstheorie sehr häufig verwendetes Abstandsmaß ist die von UHLMANN [177] eingeführte *Fidelity*; für  $\rho, \sigma \in \mathcal{S}(\mathcal{H})$  ist sie durch  $f(\rho, \sigma) := \text{Spur} \sqrt{\rho^{1/2} \sigma \rho^{1/2}}$  definiert. Die Fidelity liegt im Intervall  $[0; 1]$  und ist symmetrisch in  $\rho$  und  $\sigma$ ; es gilt genau dann  $f(\rho, \sigma) = 1$ , wenn  $\rho = \sigma$  ist. Ist z. B.  $\sigma = |\Psi\rangle\langle\Psi|$ , so vereinfacht sich der Ausdruck zu  $f(\rho, \sigma) = \text{Spur} \sqrt{\langle\Psi|\rho|\Psi\rangle}$ ; oft wird statt der Fidelity auch die quadrierte Fidelity  $F(\rho, \sigma) := f(\rho, \sigma)^2$  verwendet.

Die Fidelity ist selbst keine Metrik, jedoch sind die abgeleiteten Größen *Winkel*  $\arccos f(\rho, \sigma)$  und die *BURES-Metrik*  $\sqrt{\text{Spur} \rho + \text{Spur} \sigma - 2f(\rho, \sigma)}$  Metriken. Für den folgenden Satz vgl. NIELSEN UND CHUANG [127], S. 410f.

#### Satz C.30 (Satz von Uhlmann)

Für zwei Dichtematrizen  $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A)$  ist  $f(\rho, \sigma) = \max_{|\Psi\rangle, |\Phi\rangle} |\langle\Psi|\Phi\rangle|$ , wobei  $|\Psi\rangle, |\Phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  Purifizierungen von  $\rho$  bzw.  $\sigma$  sind.

Für das folgende betrachte man zwei reine Zustände  $\rho = |a\rangle\langle a|$  und  $\sigma = |b\rangle\langle b|$ , wobei für die Zustandsvektoren  $\langle a|b\rangle \in \mathbb{R}$  angenommen werden kann. Das GRAM-SCHMIDT-Verfahren erzeugt eine Orthonormalbasis  $\{|0\rangle, |1\rangle\}$  des von  $|a\rangle$  und  $|b\rangle$  aufgespannten Unterraums von  $\mathcal{H}$ ; es ist

$$|0\rangle := |a\rangle \quad \text{und} \quad |1\rangle := \frac{(\mathbb{I} - |0\rangle\langle 0|)|b\rangle}{\sqrt{\langle b|(\mathbb{I} - |0\rangle\langle 0|)^2|b\rangle}} = \frac{|b\rangle - \langle a|b\rangle|a\rangle}{\sqrt{1 - |\langle a|b\rangle|^2}}. \quad (\text{C.24})$$

Mit  $\vartheta = \arccos \langle a|b\rangle \in [0; \pi]$  schreibt sich also  $|b\rangle = \cos \vartheta |0\rangle + \sin \vartheta |1\rangle$ ; es folgt  $f(\rho, \sigma) = |\cos \vartheta|$ . Man berechnet weiter

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad \sigma = \begin{pmatrix} \cos^2 \vartheta & \sin \vartheta \cos \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix}, \quad (\text{C.25})$$

und mit dem Variationsabstand  $\delta(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1$  gilt

$$\delta(\rho, \sigma) = \frac{1}{2} \left\| \begin{pmatrix} 1 - \cos^2 \vartheta & -\sin \vartheta \cos \vartheta \\ -\sin \vartheta \cos \vartheta & -\sin^2 \vartheta \end{pmatrix} \right\|_1. \quad (\text{C.26})$$

Das charakteristische Polynom der Matrix ist  $\chi(x) = x^2 - \sin^2 \vartheta$ , seine Nullstellen also  $\lambda_{1;2} = \pm \sin \vartheta$ . Es folgt  $\delta(\rho, \sigma) = (1/2) \cdot 2 |\sin \vartheta| = \sqrt{1 - f(\rho, \sigma)^2}$ , also  $\delta(\rho, \sigma)^2 + f(\rho, \sigma)^2 = 1$  für reine Zustände.

Für allgemeine Zustände gilt  $1 - f(\rho, \sigma) \leq \delta(\rho, \sigma) \leq \sqrt{1 - f(\rho, \sigma)^2}$ , wofür auf NIELSEN UND CHUANG [127], S. 415–416, verwiesen sei.

## C.3 Mathematische Hilfsmittel

In diesem Abschnitt werden mathematische Aussagen zusammengestellt, die in der Theorie klassischer und quantenmechanischer Codes verwendet werden. Dies sind in einem ersten Unterabschnitt einige Schranken für Binomialkoeffizienten, in einem zweiten Unterabschnitt die sogenannten CHERNOFF-Schranken; obgleich letztere der mathematischen Statistik zuzuordnen sind, sind sie im allgemeinen nur schwierig aufzufinden.

### C.3.1 Schranken für Binomialkoeffizienten

In der Statistik und in der Kodierungstheorie werden sehr häufig Binomialkoeffizienten gebraucht. In den folgenden zwei Lemmata werden diese durch einfachere Ausdrücke abgeschätzt.

**Lemma C.31 (Ein Ausdruck für Binomialkoeffizienten)**

Ist  $n \in \mathbb{N}$  und wird  $\lambda \in [0; 1]$  so gewählt, daß  $\lambda n \in \{1, \dots, n-1\}$  ist, so existiert eine Funktion  $g: \mathbb{N} \rightarrow \mathbb{R}$  derart, daß mit der Abkürzung  $\mu := 1 - \lambda$  die Gleichung

$$\binom{n}{\lambda n} = \frac{g(n)}{g(\lambda n)g(\mu n)} \cdot \frac{2^{nH(\lambda)}}{\sqrt{2\pi\lambda\mu n}}$$

erfüllt ist und für die Funktionswerte  $g(n) \in [e^{1/12n-1/360n^3}; e^{1/12n}]$  gilt.

*Beweis:* Die Stirling-Formel (Satz B.1) besagt, daß  $n! = n^n e^{-n} \sqrt{2\pi n} \cdot g(n)$  für eine Funktion mit den gewünschten Eigenschaften ist. Hieraus berechnet man

$$\begin{aligned} \binom{n}{\lambda n} &= \frac{g(n)}{g(\lambda n)g(\mu n)} \cdot \frac{n^n}{(\lambda n)^{\lambda n} (\mu n)^{\mu n}} \cdot \frac{e^{-n}}{e^{-\lambda n} e^{-\mu n}} \cdot \frac{\sqrt{2\pi n}}{\sqrt{2\pi\lambda n} \cdot \sqrt{2\pi\mu n}} \\ &= \frac{g(n)}{g(\lambda n)g(\mu n)} \cdot \left( \frac{n}{(\lambda n)^\lambda (\mu n)^\mu} \right)^n \cdot 1 \cdot \frac{1}{\sqrt{2\pi\lambda\mu n}}, \end{aligned} \quad (\text{C.27})$$

und mithilfe von

$$\frac{n}{(\lambda n)^\lambda (\mu n)^\mu} = \frac{n}{\lambda^\lambda n^\lambda \cdot \mu^\mu n^\mu} = \lambda^{-\lambda} \mu^{-\mu} = 2^{-\lambda \log_2 \lambda - \mu \log_2 \mu} = 2^{H(\lambda)} \quad (\text{C.28})$$

folgt schließlich der gewünschte Ausdruck, q. e. d.

**Lemma C.32 (Abschätzung von Binomialkoeffizienten)**

Wählt man  $n \in \mathbb{N}$  und  $\lambda \in [0; 1]$  derart, daß  $\lambda n \in \{1, \dots, n-1\}$  ist, so gilt mit der Abkürzung  $\mu := 1 - \lambda$  die Abschätzung

$$\frac{2^{nH(\lambda)}}{\sqrt{8n\lambda\mu}} \leq \binom{n}{\lambda n} \leq \frac{2^{nH(\lambda)}}{\sqrt{2\pi\lambda\mu}}.$$

*Beweis:* Im Fall  $n = 1$  ist kein  $\lambda$  möglich und daher nichts zu zeigen; die Fälle  $n \in \{2, 3\}$ , also die Binomialkoeffizienten  $\binom{2}{1}$ ,  $\binom{3}{1}$  und  $\binom{3}{2}$ , können durch Nachrechnen überprüft werden. Unter Verwendung des Lemmas C.31 genügt es zu zeigen, daß

$$\frac{\sqrt{\pi}}{2} \leq \frac{g(n)}{g(\lambda n)g(\mu n)} \leq \sqrt{n} \quad (\text{C.29})$$

gilt, wenn  $g(n) \in [1; e^{1/12n}]$  ist. Die obere Schranke folgt aus der Ungleichung  $e^{1/12n} \leq \sqrt{n}$ , die für  $n \geq 2$  gilt. Mittels  $\lambda^{-1} + \mu^{-1} = (\lambda\mu)^{-1}$  berechnet man für die untere Schranke

$$\frac{g(n)}{g(\lambda n)g(\mu n)} \geq \frac{1}{e^{1/12\lambda n} \cdot e^{1/12\mu n}} = e^{-(1/\lambda+1/\mu)/12n} = e^{-1/12\lambda\mu n} \quad (\text{C.30})$$

und maximiert über die rechte Seite, minimiert also  $\lambda\mu n$ . Für festes  $n \in \mathbb{N}$  ist  $\lambda\mu$  für  $\lambda = n^{-1}$  minimal und  $\lambda\mu n = n^{-1} \cdot (n-1)n^{-1} \cdot n = (n-1)n^{-1}$ ; dieser Ausdruck wächst monoton mit steigendem  $n$ . Für  $n \geq 4$  folgt nun

$$\frac{g(n)}{g(\lambda n)g(\mu n)} \geq e^{-n/12(n-1)} \geq e^{-1/9} \approx 0,894839 \geq 0,886227 \approx \frac{\sqrt{\pi}}{2}, \quad (\text{C.31})$$

was die gewünschte Aussage liefert, q. e. d.

### C.3.2 Chernoff-Schranken

Die *CHERNOFF-Schranken* sind in den Lehrbüchern der mathematischen Statistik selten zu finden, aber in der Kodierungstheorie ein sehr häufig verwendetes Hilfsmittel.<sup>8</sup> Aus diesem Grund sollen in diesem Abschnitt drei eng verwandte, aber inäquivalente *CHERNOFF-Schranken* kurz besprochen werden, von denen im wesentlichen nur die erste bedeutsam ist. Die folgenden Ausführungen stützen sich ausschließlich auf Sekundärliteratur und wurden durch mich geringfügig erweitert.

Das zentrale Hilfsmittel im Beweis aller *CHERNOFF-Schranken* ist das folgende Lemma.

**Lemma C.33 (Markowsche Ungleichung)**

Für eine Zufallsvariable  $X \geq 0$  und  $a > 0$  gilt  $P(X \geq a) \leq a^{-1} \cdot \mathcal{E}(X)$ .

*Beweis:* Es gilt  $\mathcal{E}(X) = \sum_{x \in \mathbb{R}_0^+} x \cdot p_x \geq \sum_{x \geq a} x \cdot p_x \geq a \cdot \sum_{x \geq a} p_x$ , wobei  $p_x := P(X = x)$  ist; aus  $\sum_{x \geq a} p_x = P(X \geq a)$  folgt die Behauptung, q. e. d.

---

<sup>8</sup>Oft werden sie nach den Originalarbeiten von *CHERNOFF* [32] und *HOEFFDING* [75] auch als *CHERNOFF-HOEFFDING-Schranken* bezeichnet.

In den drei im folgenden vorgestellten Varianten der CHERNOFF-Schranke wird zunächst die MARKOWSche Ungleichung auf die momenterzeugende Funktion angewendet, anschließend aber, der Allgemeinheit der Aussagen angepaßt, unterschiedlich abgeschätzt.

### Chernoff-Schranke für Binomialverteilungen

Die erste und gleichzeitig wichtigste Ungleichung liefert der folgende Satz; der Beweis wurde dem Buch von ROMAN [153], S. 25–26, entnommen. Im Falle von  $p = 1/2$  nennt man die Ungleichung auch engl. *tail inequality*, deren Beweis sich auch bei NIELSEN UND CHUANG [127] auf S. 154 findet.

#### Satz C.34 (Chernoff-Schranke für Binomialverteilungen)

Sind  $n \in \mathbb{N}$  und  $p, \lambda \in [0; 1]$ , so gilt

$$\sum_{k=0}^{\lambda n} \binom{n}{k} p^k (1-p)^{n-k} \leq \left(\frac{p}{\lambda}\right)^{\lambda n} \left(\frac{1-p}{1-\lambda}\right)^{(1-\lambda)n} \quad \text{für } \lambda < p,$$

$$\sum_{k=\lambda n}^n \binom{n}{k} p^k (1-p)^{n-k} \leq \left(\frac{p}{\lambda}\right)^{\lambda n} \left(\frac{1-p}{1-\lambda}\right)^{(1-\lambda)n} \quad \text{für } \lambda > p.$$

Stehen  $\lambda$  und  $p$  für  $B(1, \lambda)$ - bzw.  $B(1, p)$ -verteilte Zufallsvariablen, so läßt sich die rechte Seite beider Ungleichungen zu  $2^{-nH(\lambda|p)}$  umschreiben.

*Beweis:* Wählt man eine Zahl  $t \in \mathbb{R}$  und definiert die Zufallsvariable  $X = e^{tY}$ , die nur nicht-negative Werte annimmt, und einen Wert  $a = e^{tb}$ , so liefert die MARKOWSche Ungleichung<sup>9</sup>

$$P(Y \leq b) \stackrel{t \leq 0}{\leq} P(e^{tY} \geq e^{tb}) \leq e^{-tb} \mathcal{E}(e^{tY}). \quad (\text{C.32})$$

Ist  $Y$  nun  $(n, p)$ -binomialverteilt und  $q := 1 - p$ , so gilt  $\mathcal{E}(e^{tY}) = (q + pe^t)^n$ , und für  $b \in \mathbb{R}$  ist

$$P(Y \leq b) = \sum_{k=0}^b \binom{n}{k} p^k q^{n-k} \leq e^{-tb} (q + pe^t)^n. \quad (\text{C.33})$$

Mit  $\lambda = b/n \in (0; 1)$  folgt nach Minimierung über  $t \in \mathbb{R}^-$  bzw.  $e^t \in (0; 1)$  die Bedingung  $e^t = \lambda \cdot (1 - \lambda)^{-1} \cdot q \cdot p^{-1}$ ; der Ausdruck auf der rechten Seite wird dann mit  $\mu := 1 - \lambda$  zu  $(\lambda \cdot q \cdot \mu^{-1} \cdot p^{-1})^{-\lambda n} (q \cdot \mu^{-1})^n = \lambda^{-\lambda n} \mu^{-\mu n} p^{\lambda n} q^{\mu n}$ , woraus die erste Ungleichung folgt.

Der zweite Fall geht aus dem ersten durch die Substitutionen  $p \rightarrow 1 - p$  und  $k \rightarrow n - k$  (man ersetze die Summationsgrenzen, nicht die Laufvariable) hervor, indem man das entstehende  $(1 - \lambda)$  durch  $\lambda$  ersetzt, q. e. d.

---

<sup>9</sup>ROMAN [153] nennt fälschlich die TSCHEBYSCHEFFSche Ungleichung.

Als Funktion von  $n \in \mathbb{N}$  betrachtet, besitzt die Schranke die Form  $a_0^n$  mit der Basis  $a_0 = \left(\frac{p}{\lambda}\right)^\lambda \left(\frac{1-p}{1-\lambda}\right)^{1-\lambda}$ ; man kann nun zeigen, daß man keine kleinere Basis als diese wählen kann. Hierzu schätzt man den Ausdruck auf der linken Seite durch den größten Summanden ab, setzt also  $k = \lambda n$  und verbleibt dann mit  $\binom{n}{\lambda n} \leq 2^{nH(\lambda)}$ . Nach Lemma C.32 ist  $\binom{n}{\lambda n} \geq (2n)^{-1/2} \cdot 2^{nH(\lambda)}$ , übersteigt also für  $n \rightarrow \infty$  jede Exponentialfunktion mit einer Basis kleiner als  $2^{H(\lambda)}$ .<sup>10</sup>

Das folgende Korollar findet sich als Lemma 3 in MAYERS' Arbeit [120] auf S. 358; der dort zitierte Text enthält allerdings keinen Beweis.

**Korollar C.35 (Chernoff-Schranke nach Mayers)**

Für eine  $(n, p)$ -binomialverteilte Zufallsvariable  $S$  und  $\Delta p \in [0; 1]$  gilt

$$\begin{aligned} P(S \geq n(p + \Delta p)) &\leq \exp[-2n(\Delta p)^2], \\ P(S \leq n(p - \Delta p)) &\leq \exp[-2n(\Delta p)^2]. \end{aligned}$$

*Beweis:* Man setze  $\lambda := p - \Delta p < p$  und betrachte die zweite Ungleichung. Nach der ersten Ungleichung aus Satz C.34 ist diese zumindest dann erfüllt, wenn  $-\lambda \ln \lambda - (1 - \lambda) \ln(1 - \lambda) + \lambda \ln p + (1 - \lambda) \ln(1 - p) \leq -2(p - \lambda)^2$  gilt. Für  $p = \lambda$  liegt Gleichheit vor; beidseitiges Ableiten nach  $p$  liefert die hinreichende Bedingung (beachte  $p - \lambda \geq 0$ )

$$\frac{\lambda}{p} - \frac{1 - \lambda}{1 - p} = \frac{-(p - \lambda)}{p(1 - p)} \leq -4(p - \lambda), \tag{C.34}$$

die wegen  $p(1 - p) \leq 1/4$  stets erfüllt ist. Der Beweis der ersten Ungleichung erfolgt analog, q. e. d.

Mithilfe des Satzes C.34 kann im Exponenten kein größerer Faktor als 2 gewählt werden, der nicht von  $p$  abhängt; man kann aber die Schranke für bekanntes  $p$  anpassen.

**Summen binomialverteilter Zufallsvariablen**

Die in Satz C.37 aufgeführte Variante der CHERNOFF-Schranke ist zum Großteil dem Buch von SCHICKINGER UND STEGER [158], S. 65–70, entnommen; zu ihrem Beweis wird das folgende Lemma benötigt.

**Lemma C.36 (Abschätzungen reellwertiger Funktionen)**

*Es gilt*

$$\begin{aligned} (1 + \delta) \ln(1 + \delta) &\geq +\delta + \min\{\delta, \delta^2\} / 3 && \text{für } \delta \in \mathbb{R}_0^+, \\ (1 - \delta) \ln(1 - \delta) &\geq -\delta + \delta^2 / 2 && \text{für } \delta \in [0; 1]. \end{aligned}$$

---

<sup>10</sup>Faßt man die linke Seite der Gleichungen im Satz C.34 als eine Funktion  $f(n)$  auf, so besagen der Satz und die Nachbemerkingen, daß  $z(f) = \left(\frac{p}{\lambda}\right)^\lambda \left(\frac{1-p}{1-\lambda}\right)^{1-\lambda}$  im Sinne des Unterabschnitts 3.6.5 gilt.

*Beweis:* Für  $\delta = 0$  liegt in beiden Fällen Gleichheit vor, und die Produktregel  $f = uv \Rightarrow f' = u'v + uv'$  liefert für die linken Seiten

$$\frac{d}{d\delta} [(1 \pm \delta) \ln(1 \pm \delta)] = \pm [\ln(1 \pm \delta) + 1] \stackrel{|\delta| < 1}{=} \pm \left[ 1 - \sum_{k=1}^{\infty} (\mp 1)^k \frac{\delta^k}{k} \right]. \quad (\text{C.35})$$

Für die rechte Seite der zweiten Ungleichung gilt  $(d/d\delta) [-\delta + \delta^2/2] = -1 + \delta$ , und die Aussage folgt wegen  $-1 + (\delta + \delta^2/2 + \dots) \geq -1 + \delta$ . Für den Beweis der ersten Ungleichung im Falle  $\delta \leq 1$  berechnet man

$$\begin{aligned} (1 + \delta) \ln(1 + \delta) &= \delta + \sum_{k=2}^{\infty} (-1)^k \frac{\delta^k}{k(k-1)} \\ &= \delta + \frac{\delta^2}{2} - \frac{\delta^3}{6} + \sum_{k=2}^{\infty} \left( \frac{\delta^{2k}}{2k} - \frac{\delta^{2k+1}}{2k+1} \right). \end{aligned} \quad (\text{C.36})$$

Alle Summanden der letzten Summe sind nun für  $\delta \in [0; 1]$  nicht-negativ, und somit genügt es zu zeigen, daß  $\delta + \delta^2/2 - \delta^3/6 \geq \delta + \delta^2/3$  oder  $(\delta^2 - \delta^3)/6 \geq 0$  gilt, was im betrachteten Intervall der Fall ist.

Da wegen  $2 \ln 2 \approx 1,386 \geq 1, \bar{3} = 4/3$  die erste Ungleichung an der Stelle  $\delta = 1$  erfüllt ist, genügt es nun, für den Fall  $\delta \geq 1$  festzuhalten, daß  $1 + \ln(1 + \delta) \geq 1 + \ln 2 \approx 1,693 \geq 4/3$  ist, q. e. d.

**Satz C.37 (Chernoff-Schranke für zweiwertige Verteilungen)**

Für  $i \in \{1, \dots, n\}$  seien  $X_i$  statistisch unabhängige  $B(1, p_i)$ -verteilte Zufallsvariablen mit (ggf. unterschiedlichen) Erfolgswahrscheinlichkeiten  $p_i \in [0; 1]$ . Die Summe  $X := \sum_{i=1}^n X_i$  hat den Erwartungswert  $\mu := \mathcal{E}(X) = \sum_{i=1}^n p_i$ , und es gilt

$$\begin{aligned} P(X \geq (1 + \delta)\mu) &\leq \left[ \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right]^\mu \leq e^{-\mu \min\{\delta, \delta^2\}/3} && \text{für } \delta \in \mathbb{R}_0^+, \\ P(X \leq (1 - \delta)\mu) &\leq \left[ \frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right]^\mu \leq e^{-\mu \delta^2/2} && \text{für } \delta \in [0; 1]. \end{aligned}$$

*Beweis:* Nach Lemma C.33 gilt zunächst  $P(X \geq (1 + \delta)\mu) \leq e^{-t(1+\delta)\mu} \mathcal{E}(e^{tX})$ . Mittels  $\mathcal{E}(e^{tX}) = \prod_{i=1}^n \mathcal{E}(e^{tX_i})$  und der Ungleichung  $1 + x \leq e^x$  für Werte  $x \geq 0$  berechnet man weiter  $P(X \geq (1 + \delta)\mu) \leq e^{-t(1+\delta)\mu} \prod_{i=1}^n [1 + p_i(e^t - 1)] \leq \exp[(e^t - 1)\mu - t(1 + \delta)\mu]$ . Durch Minimieren erhält man  $t = \ln(1 + \delta)$  und hieraus die erste Behauptung. Der zweite Fall folgt analog, und die gröberen Abschätzungen folgen unmittelbar aus dem vorstehenden Lemma, q. e. d.

### Allgemeine Verteilungen

Die folgende Schranke gilt auch für nicht binomialverteilte Zufallsvariablen.<sup>11</sup> Zunächst wird eine Eigenschaft der momenterzeugenden Funktion gezeigt, unmittelbar darauf die CHERNOFF-Schranke für allgemeine Verteilungen.

#### Lemma C.38 (Schranke für die momenterzeugende Funktion)

Es sei  $Z$  eine Zufallsvariable, die  $|Z| \leq 1$  und  $\mathcal{E}(Z) = 0$  erfülle. Für jedes  $t \in [-1; 1]$  gilt  $\mathcal{E}(e^{tZ}) \leq 1 + t^2(e - 2) \text{Var}(Z) \leq 1 + t^2 \text{Var}(Z)$ .

*Beweis:* Mithilfe von  $e^x = \sum_{k=0}^{\infty} x^k/k!$  berechnet man

$$\begin{aligned} \mathcal{E}(e^{tZ}) &= \sum_j p_j e^{tz_j} = \sum_j p_j \left( 1 + tz_j + \frac{(tz_j)^2}{2!} + \frac{(tz_j)^3}{3!} + \dots \right) \\ &= \underbrace{\sum_j p_j}_{=:A} + t \underbrace{\sum_j p_j z_j}_{=:B} + \underbrace{\sum_j p_j \left( \frac{(tz_j)^2}{2!} + \frac{(tz_j)^3}{3!} + \dots \right)}_{=:C}. \end{aligned} \quad (\text{C.37})$$

Die Normierung der Wahrscheinlichkeit liefert  $A = 1$ , nach Voraussetzung ist  $B = \mathcal{E}(Z) = 0$ , und mit  $|tz_j| \leq 1$  berechnet man

$$C \leq \sum_j p_j (tz_j)^2 \left( \frac{1}{2!} + \frac{1}{3!} + \dots \right) = t^2(e - 2) \sum_j p_j z_j^2. \quad (\text{C.38})$$

Die Summe im letzten Term ist die Varianz von  $Z$ , so daß  $C \leq t^2(e - 2) \text{Var}(Z)$  gezeigt wurde; schließlich ist noch  $e - 2 \approx 0,718 \leq 1$ , q. e. d.

#### Satz C.39 (Chernoff-Schranke für allgemeine Verteilungen)

Es seien  $X_1, \dots, X_n$  diskrete, statistisch unabhängige Zufallsvariablen derart, daß  $\mathcal{E}(X_i) = 0$  und  $|X_i| \leq 1$  für alle  $i \in \{1, \dots, n\}$  gelten möge. Ferner sei  $X := \sum_{i=1}^n X_i$  die Summe dieser Zufallsvariablen mit Varianz  $\sigma^2 := \text{Var}(X)$ . Es gilt dann

$$P(X \geq \lambda\sigma) \leq e^{-\lambda^2/(4(e-2))} \leq e^{-\lambda^2/4}$$

für jeden Wert  $\lambda \in [0; 2\sigma]$ ; ferner ist  $P(|X| \geq \lambda\sigma) \leq 2e^{-\lambda^2/4}$ .

*Beweis:* Für  $t \in [0; 1]$  ergibt sich mithilfe der MARKOWSchen Ungleichung zunächst  $P(X \geq \lambda\sigma) = P(tX \geq t\lambda\sigma) = P(e^{tX} \geq e^{t\lambda\sigma}) \leq e^{-t\lambda\sigma} \mathcal{E}(e^{tX})$ . Ferner ist

$$\mathcal{E}(e^{tX}) = \mathcal{E}\left(e^{t\sum_{i=1}^n X_i}\right) = \mathcal{E}\left(\prod_{i=1}^n e^{tX_i}\right) = \prod_{i=1}^n \mathcal{E}(e^{tX_i}), \quad (\text{C.39})$$

---

<sup>11</sup>Diese Formulierung folgt im wesentlichen den unveröffentlichten Notizen mit dem Titel „Chernoff Bound“ von KIRILL LEVCHENKO; siehe <http://www.cs.ucsd.edu/~klevchen/techniques/chernoff.pdf>.

wobei die letzte Gleichheit wegen der statistischen Unabhängigkeit der Zufallsvariablen gilt. Mit Lemma C.38 und  $1 + \alpha \leq e^\alpha$  für  $\alpha \geq 0$  folgt weiter

$$\mathcal{E}(e^{tX}) \leq \prod_{i=1}^n e^{(e-2)t^2 \text{Var}(X_i)} = e^{(e-2)t^2 \sum_{i=1}^n \text{Var}(X_i)} = e^{(e-2)t^2 \sigma^2}, \quad (\text{C.40})$$

wobei  $\sum_i \text{Var}(X_i) = \sigma^2$  wiederum aus der statistischen Unabhängigkeit folgt. Es gilt somit  $P(X \geq \lambda\sigma) \leq e^{t\sigma(t\sigma(e-2)-\lambda)}$ . Minimieren über die rechte Seite liefert  $t = \lambda/(2\sigma(e-2))$  und somit  $P(X \geq \lambda\sigma) \leq e^{-\lambda^2/(4(e-2))}$ . Der zweite Teil folgt aus  $P(|X| \geq \lambda\sigma) \leq P(X \geq \lambda\sigma) + P(-X \geq \lambda\sigma)$ , q. e. d.

## C.4 Grundbegriffe der Kodierungstheorie

Bei der Übertragung von Nachrichten über eine Leitung (einen Kanal) muß immer damit gerechnet werden, daß der Kanal Störungen auf der Nachricht hervorruft. Um eine Nachricht fehlerfrei übertragen zu können, kann man nun Kodierungen verwenden. Man vergleiche die Bücher von VAN LINT [109] und von MACWILLIAMS UND SLOANE [117].

### C.4.1 Allgemeine Begriffe

Ein *Alphabet* ist eine endliche, im allgemeinen mindestens zweielementige Menge  $\Omega$ ; bekannt sind das lateinische Alphabet  $\Omega = \{A, \dots, Z\}$  mit 26 Elementen (ohne Kleinbuchstaben, Umlaute, Sonderzeichen usw.) und die Binärzahlen (Bits)  $\Omega = \{0, 1\}$ . Für informationstheoretische Fragestellungen ist allein die Mächtigkeit  $q := |\Omega|$  des Alphabets entscheidend. Im folgenden sei daher  $\Omega = \mathbb{Z}/q\mathbb{Z} \cong \mathbb{Z}_q$  für ein vorgegebenes  $q \in \mathbb{N} \setminus \{1\}$ .

Eine *Nachricht* ist eine endliche Zeichenfolge von Elementen aus einem vorgegebenen Alphabet, also ein Element aus der Menge  $\bigcup_{k=1}^{\infty} \Omega^k$ . Für jedes  $n \in \mathbb{N}$  kann man auf  $\Omega^n$  das *HAMMING-Gewicht*  $\text{Gew} : \Omega^n \rightarrow \{0, \dots, n\}$  definieren; es ist<sup>12</sup>

$$\text{Gew}(v) := |\{k \in \{1, \dots, n\} \mid v_k \neq 0\}| \quad \text{für } v = (v_1, \dots, v_n) \in \Omega^n. \quad (\text{C.41})$$

Der *HAMMING-Abstand*  $d(v, w) := \text{Gew}(v \ominus w)$  zweier Vektoren  $v, w \in \Omega^n$  definiert eine Metrik auf  $\Omega^n$ .

Ein *Code*  $C$  ist ein  $M$ -Tupel  $(x_1, \dots, x_M)$  über  $\Omega^n$ . Enthält eine Nachricht das Zeichen  $i \in \{1, \dots, M\}$ , so wird dieses durch den Sender mittels  $i \mapsto x_i$  *kodiert* und an den Empfänger übertragen. Dieser erhält eine durch

---

<sup>12</sup>Ist  $q$  eine Primzahl und betrachtet man  $\Omega^n = \mathbb{F}_q^n$  als Vektorraum über  $\mathbb{F}_q$ , so ist das HAMMING-Gewicht eine Norm und der HAMMING-Abstand die zugeordnete Metrik.

den Kanal verfälschte Nachricht  $y \in \mathbb{Z}_q^n$  und versucht, den wahrscheinlichsten Wert für  $i$  zu erschließen, indem er dasjenige  $x_i$  wählt, das den kleinsten HAMMING-Abstand zu  $y$  besitzt und  $y$  zu  $i$  *dekodiert*. Meist nennt man nur die Menge  $\{x_1, \dots, x_M\}$  den Code und verzichtet auf die Erwähnung der Kodierungsabbildung. Die *Länge* des Codes  $C$  ist die Anzahl  $n = \log_q |\Omega^n|$  der *physikalischen Zeichen* in einem Block. Man faßt  $\log_q M$  als die Anzahl der *logischen Zeichen* auf und bezeichnet den Quotienten  $R := (\log_q M)/n$  als die *Informationsrate* oder kurz *Rate* des Codes  $C$ . Die *Minimaldistanz* des Codes ist  $d(C) := \min \{d(x_1, x_2) \mid x_1, x_2 \in C\}$ .

Man kann z. B. für  $q = 2$  zeigen, daß im Falle eines *binären symmetrischen Kanals*, der jedes übertragene Bit unabhängig mit einer Wahrscheinlichkeit  $p \in [0; 1]$  verfälscht, für  $R \leq 1 - H(p)$  ein Code mit dieser Rate existiert; diese Aussage bezeichnet man als die *SHANNON-Schranke* [165] (oder als den *zweiten SHANNONScher Hauptsatz* oder den *Satz über die Kanalkodierung*). Für einen Beweis vgl. das Buch von VAN LINT [109], § 2.2 auf S. 27–29.

In dieser Dissertation wird eine bei weitem allgemeinere Fassung der SHANNON-Schranke benötigt: sie gilt für CSS-Codes (Unterabschnitt 3.5.1), die das quantenmechanische Analogon klassischer linearer Codes bilden und die Beschränkung auf  $q = 2$  wird aufgegeben (vgl. hierzu Satz 4.1).

## C.4.2 Lineare Codes

Von besonderer Bedeutung sind Alphabete, deren Mächtigkeit eine Primzahlpotenz  $q = p^n$  ist, da dann das Alphabet mit der Struktur eines endlichen Körpers versehen werden kann; in diesem Fall sei  $\Omega = \mathbb{F}_q$ .

Da  $\mathbb{F}_q^n$  einen Vektorraum der Dimension  $n$  über  $\mathbb{F}_q$  bildet, können die Methoden der linearen Algebra angewendet werden; so setzt man z. B. für Vektoren  $x = (x_1, \dots, x_n)^t \in \mathbb{F}_q^n$  und  $y = (y_1, \dots, y_n)^t \in \mathbb{F}_q^n$  ein Produkt  $x \cdot y := \sum_{i=1}^n x_i y_i$ . Wählt man den Coderaum  $C$  als einen linearen Teilraum von  $\mathbb{F}_q^n$  und ist die Kodierungsabbildung  $G : \mathbb{F}_q^k \rightarrow \mathbb{F}_q^n$  linear, so spricht man von einem *linearen Code* der Länge  $n$ , der  $k$  Zeichen kodiert, und schreibt hierfür  $[n, k]_q$ -Code oder  $[n, k]$ -Code; ist seine Minimaldistanz  $d = \min \{\text{Gew}(x) \mid x \in C \setminus \{0\}\}$  bekannt, so schreibt man  $[n, k, d]_q$ -Code.

Stellt man  $G$  durch eine Matrix dar, so nennt man  $G$  die *Generatormatrix* des Codes  $C$ ; es ist dann  $C = \text{Bild } G$ . Alternativ kann der Code (als Menge) durch eine *Kontrollmatrix*  $H$  beschrieben werden; es ist dann  $C = \text{Kern } H$ .<sup>13</sup>

Der zu  $C$  *duale Code* ist  $C^\perp := \{x \in \mathbb{F}_q^n \mid (\forall y \in C)(x \cdot y = 0)\}$ . Seine Generatormatrix ist  $H^t$ , seine Kontrollmatrix  $G^t$ , es handelt sich also um einen  $[n, n - k]$ -Code; ferner ist  $(C^\perp)^\perp = C$ .

---

<sup>13</sup>Dies ist in etwa analog zur HESSESchen Normalform einer Ebene im  $\mathbb{R}^3$ .

**Lemma C.40 (Duale Codes)**

Es gilt  $\sum_{y \in C} z_p^{x \cdot y} = |C|$ , falls  $x \in C^\perp$  ist; andernfalls ist  $\sum_{y \in C} z_p^{x \cdot y} = 0$ .

*Beweis:* Ist  $x \in C^\perp$ , so gilt  $\sum_{y \in C} z_p^{x \cdot y} = \sum_{y \in C} 1 = |C|$ . Andernfalls schreibt man  $y = \sum_{i=1}^k y_i e_i$  für eine Basis  $\{e_1, \dots, e_k\}$  von  $C$  und erhält

$$\sum_{y \in C} z_p^{x \cdot y} = \sum_{y_1, \dots, y_k=0}^{p-1} z_p^{x \cdot (y_1 e_1 + \dots + y_k e_k)} = \prod_{i=1}^k \left( \sum_{y_i=0}^{p-1} (z_p^{x \cdot e_i})^{y_i} \right).$$

Da  $x \notin C^\perp$  angenommen wurde, verschwinden nicht alle Ausdrücke  $z_p^{x \cdot e_i}$ , und die Summierung über  $y_i$  ergibt Null; mit einem Faktor verschwindet dann auch das Produkt, q. e. d.

**C.4.3 Diskrete Kugeln**

Ähnlich wie in kontinuierlichen Vektorräumen  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  kann man auch über  $\Omega^n$  Kugeln betrachten, selbst dann, wenn  $q$  keine Primzahlpotenz ist. Betrachtet man  $n$ -Tupel über  $\Omega$ , ggf. also ein Element des Vektorraums  $\mathbb{F}_q^n$ , so bezeichnet  $B_r(x) := \{y \in \mathbb{F}_q^n \mid d(x, y) \leq r\}$  die Kugel mit einem Radius  $r$  um den Punkt  $x$ , wobei hier stets die Kugelhülle eingeschlossen sei.

Das Volumen dieser Kugel, also die Anzahl ihrer Elemente ist nicht von  $x$  abhängig und bestimmt sich zu

$$V_q(n, r) := |B_r(0)| = \sum_{k=0}^r \binom{n}{k} (q-1)^k. \tag{C.42}$$

Mit der Notation aus Gleichung (C.2) läßt sich das Volumen diskreter Kugeln wie folgt abschätzen.

**Lemma C.41 (Volumen diskreter Kugeln)**

Im Falle von  $r \leq n/q$  gilt  $V_q(n, r) \leq n \cdot q^{nH_q(r/n)-1}$ .

*Beweis:* Für  $n, r \in \mathbb{N}_0$  mit  $r \leq n$  gilt unter Verwendung des Binomischen Lehrsatzes die Ungleichung  $n^n = [(n-r) + r]^n = \sum_{k=0}^n \binom{n}{k} (n-r)^{n-k} \cdot r^k \geq \binom{n}{r} (n-r)^{n-r} r^r$ , also  $\binom{n}{r} \leq n^n \cdot [(n-r)^{n-r} \cdot r^r]^{-1}$ . Hieraus folgt

$$V_q(n, r) = \sum_{k=0}^r \binom{n}{k} (q-1)^k \leq \frac{n}{q} \binom{n}{r} (q-1)^r \leq \frac{n}{q} \cdot \frac{n^n}{(n-r)^{n-r} \cdot r^r} \cdot (q-1)^r,$$

und der letzte Ausdruck läßt sich zu  $n \cdot q^{nH_q(r/n)-1}$ , umschreiben, q. e. d.

### C.4.4 Gilbert-Varshamov-Schranken

Aus dem Volumen diskreter Kugeln ergibt sich durch „Ausschöpfung“ des Raumes fast unmittelbar die folgende Schranke, die für den Fall  $q = 2$  von GILBERT [60], Theorem 1 auf S. 507, stammt.

**Lemma C.42 (Gilbert-Schranke)**

Es sei  $A_q(n, d)$  die maximale Größe eines Codes der Länge  $n$  und minimalem HAMMING-Abstand  $d$  über einem Alphabet  $\Omega$  der Mächtigkeit  $q$ . Es gilt

$$A_q(n, d) \geq \frac{q^n}{\sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k}. \quad (\text{C.43})$$

*Beweis:* Ist  $C$  ein maximaler Code, so ist  $\Omega^n \subseteq \bigcup_{c \in C} B_{d-1}(c)$ , denn andernfalls gäbe es ein  $x \in \Omega^n$ , welches in keiner dieser Kugeln liegt. Somit wäre  $C \cup \{x\}$  ein Code der Minimaldistanz  $d$ ,  $C$  also entgegen der Annahme nicht maximal. Es folgt nun

$$q^n = |\Omega^n| \leq \left| \bigcup_{c \in C} B_{d-1}(c) \right| = |C| \cdot \sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k \quad (\text{C.44})$$

und  $A_q(n, d) \geq q^n / \sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k$ , q. e. d.

Für lineare Codes gibt es eine verwandte Schranke, die für den Fall  $q = 2$  von VARSHAMOV [178] herrührt. Bevor diese bewiesen wird, soll noch das folgende Lemma gezeigt werden.

**Lemma C.43 (Lineare Codes und ihre Minimaldistanz)**

Ist  $H \in \mathbb{F}_q^{(n-k) \times n}$  die Kontrollmatrix eines linearen Codes  $C$  der Länge  $n$ , dann ist die Minimaldistanz von  $C$  gleich dem Rang von  $H$ , in Zeichen  $d(C) = \text{Rang } H$ . Mit anderen Worten ist jede Auswahl von  $d-1$  Spalten linear unabhängig, und es gibt eine linear abhängige Auswahl aus  $d$  Spalten.

*Beweis:* Gibt es  $d$  linear abhängige Spalten, so findet man einen Spaltenvektor  $x \in \mathbb{F}_q^n \setminus \{0\}$  mit  $\text{Gew } x = d$ , so daß  $Hx = 0$  und daher  $x \in C$  gilt. Der Umkehrschluß gilt gleichermaßen, und die Behauptung folgt aus der Feststellung, daß die Minimaldistanz eines linearen Codes gleich dem minimalen Gewicht eines vom Nullvektor verschiedenen Codeworts ist, q. e. d.

**Lemma C.44 (Varshamov-Schranke)**

Ein  $[n, k, d]$ -Code über dem Körper  $\mathbb{F}_q$  existiert zumindest dann, wenn die Ungleichung  $V_q(n-1, d-2) < q^{n-k}$  erfüllt ist.

*Beweis:* Durch vollständige Induktion wird eine Kontrollmatrix  $H \in \mathbb{F}_q^{(n-k) \times n}$  erzeugt, die den Bedingungen des Lemmas C.43 genügt. Als ersten Spaltenvektor wähle man einen beliebigen nicht-verschwindenden Vektor der Länge  $n-k$ . Nimmt man nun an, daß  $i$  Spaltenvektoren derart ausgewählt sind, daß je  $d-1$  von diesen linear unabhängig sind, dann kann man einen weiteren Spaltenvektor finden, ohne daß diese Eigenschaft verändert wird, wenn

$$\sum_{j=1}^{d-2} \binom{i}{j} (q-1)^j < q^{n-k} - 1 \quad (\text{C.45})$$

gilt. Der Ausdruck  $\binom{i}{j}(q-1)^j$  ist nämlich die Anzahl der Linearkombinationen der Spalten, bei denen genau  $j$  Koeffizienten nicht verschwinden. Die linke Seite ist also die Mächtigkeit der Menge aller Linearkombinationen aus höchstens  $d-2$  Elementen, die rechte Seite hingegen die Anzahl der von Null verschiedenen Spaltenvektoren der Länge  $n-k$ .

Gilt nun die genannte Ungleichung, so existiert zumindest ein weiterer Spaltenvektor der Länge  $n-k$ , der nicht als solche Linearkombination dargestellt werden kann; dieser wird als  $i+1$ -te Spalte zu  $H$  hinzugefügt. Damit man die  $n$ -te Spalte wählen kann, muß die Bedingung für  $i = n-1$  erfüllt sein. Beachtet man, daß stets  $\binom{n+1}{k} = \frac{n! \cdot (n+1)}{(n+1-k)! \cdot k!} = \frac{n+1}{n+1-k} \cdot \binom{n}{k} \geq \binom{n}{k}$  gilt, so sieht man, daß diese auch die Gültigkeit der Ungleichung für  $i \leq n-2$  impliziert. Addiert man auf beiden Seiten die Zahl Eins, so ergibt sich die Behauptung, q. e. d.

Eine verwandte, aber schlechtere Schranke bewies HAMMING [69], ebenfalls für den Fall  $q = 2$ ; es handelt sich aber nicht um die obere Schranke, die üblicherweise seinen Namen trägt. Es ist möglich, sie aus der VARSHAMOV-Schranke herzuleiten, indem man die Gleichung  $V_q(n, d-1) - qV_q(n-1, d-2) = \binom{n-1}{d-1} (q-1)^{d-1} \geq 0$  zeigt, was durch Induktion über  $d$  erfolgen kann. Einfacher verständlich ist aber der direkte Beweis.

**Lemma C.45 (Hamming-Schranke)**

*Gilt  $V_q(n, d-1) < q^{n-k+1}$ , dann existiert ein  $[n, k, d]$ -Code.*

*Beweis:* Der Beweis erfolgt durch Induktion; der Fall  $k = 0$  ist dabei trivial. Es sei  $C_{k-1}$  ein  $[n, k-1, d]$ -Code, der wegen  $|C_{k-1}| \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i < q^n$  mit  $|C_{k-1}| = q^{k-1}$  nicht maximal ist (Satz C.42). Es existiert also ein  $x \in \mathbb{F}_q^n$  mit einem HAMMING-Abstand nicht kleiner als  $d$  zu jedem Codewort in  $C_{k-1}$ , und man betrachte den von  $C_{k-1}$  und  $\{x\}$  aufgespannten Code  $C_k$ . Für  $z = \alpha x + y$  mit  $\alpha \in \mathbb{F}_q \setminus \{0\}$  und  $y \in C_{k-1}$  gilt nun

$$\text{Gew}(z) = \text{Gew}(a^{-1}z) = \text{Gew}(x + a^{-1}y) = d(x, -a^{-1}y) \geq d, \quad (\text{C.46})$$

d. h.  $C_k$  hat das Minimalgewicht  $d$  und ist somit ein  $[n, k, d]$ -Code, q. e. d.

Die folgende Schranke ist nun die asymptotische Form der vorangehenden Schranken; in ihr tritt die SHANNON-Entropie auf.

**Satz C.46 (Asymptotische Gilbert-Varshamov-Schranke)**

Wählt man ein  $\lambda \in [0; (q-1)/q)$ , so existiert für hinreichend großes  $n \in \mathbb{N}$  ein fehlerkorrigierender  $[n, nR, \lambda n]$ -Code mit  $R \geq 1 - H_q(\lambda)$ .

*Beweis:* Aus der Definition von  $V_q(n, \lambda n)$  und der Nachbemerkung zu Satz C.34 folgt für  $\lambda \leq (q-1)/q$  der Ausdruck

$$\begin{aligned} \frac{V_q(n, \lambda n)}{q^n} &= \sum_{k=0}^{\lambda n} \binom{n}{k} \left(\frac{1}{q}\right)^{n-k} \left(\frac{q-1}{q}\right)^k \stackrel{\text{aeg}}{=} \left[ \left(\frac{q-1}{q} \cdot \frac{1}{\lambda}\right)^\lambda \left(\frac{1}{q} \cdot \frac{1}{1-\lambda}\right)^{1-\lambda} \right]^n \\ &= \frac{1}{q^n} \left[ \left(\frac{q-1}{\lambda}\right)^\lambda \left(\frac{1}{1-\lambda}\right)^{1-\lambda} \right]^n = \frac{q^{nH_q(\lambda)}}{q^n}, \end{aligned} \tag{C.47}$$

also  $V_q(n, \lambda n) \stackrel{\text{aeg}}{=} q^{nH_q(\lambda)}$  und hieraus  $\lim_{n \rightarrow \infty} n^{-1} \log_q V_q(n, \lambda n) = H_q(\lambda)$ . Mithilfe der CHERNOFF-Schranke aus Satz C.34 liefert dieses Argument auch  $V_q(n, \lambda n) \leq q^{nH_q(\lambda)}$ .

Für die Rate eines  $[n, k, d]$ -Codes gilt nun  $R = k/n$ , wobei mit Lemma C.42 für die Anzahl der effektiven Stellen die Ungleichung

$$k = \log_q A_q(n, \lambda n) \geq \log_q \frac{q^n}{V_q(n, \lambda n)} \geq \log_q \frac{q^n}{q^{nH_q(\lambda)}} = n[1 - H_q(\lambda)]$$

erfüllt ist, d. h. es existiert zumindest ein allgemeiner Code der gegebenen Rate und dem gegebenen Minimalabstand. Aus Lemma C.45 und der o.g. CHERNOFF-Schranke folgt wegen

$$nH_q(\lambda) \leq n - k + 1 \Leftrightarrow H_q(\lambda) \leq 1 - R + 1/n \Leftrightarrow R \leq 1 - H_q(\lambda) + 1/n,$$

daß dieser Code linear gewählt werden kann, q. e. d.

Das folgende Lemma für Codes über  $\mathbb{F}_2$  bezeichneten CALDERBANK UND SHOR [28] als engl. *simple greedy argument*.

**Lemma C.47 (Existenz klassischer linearer Codes)**

Es sei  $(\Phi_i)_{i \in \mathbb{N}}$  eine Folge von Mengen linearer  $[n_i, k_i]$ -Codes derart, daß  $n_i$  für  $i \rightarrow \infty$  unbeschränkt wachse,  $k_i/n_i > R$  für eine Rate  $R > 0$  und alle  $i \in \mathbb{N}$  gelte sowie folgende Bedingung erfüllt sei:

- Die Anzahl  $N_i := |\{C \in \Phi_i \mid v \in C\}|$  der Codes in  $\Phi_i$ , die einen Vektor  $v \in \mathbb{F}_2^{n_i} \setminus \{0\}$  enthalten, ist unabhängig vom gewählten Vektor.

Dann existieren Codes in  $(\Phi_i)_{i \in \mathbb{N}}$  derart, daß  $R \geq 1 - H(d/n)$  für alle  $n$  ist.

*Beweis:* Die Anzahl der von Null verschiedenen Codewörter in allen Codes eines Folgenglieds  $\Phi_i$  ist einerseits  $|\mathbb{F}_2^{n_i} \setminus \{0\}| \cdot N_i$ , andererseits ist sie auch die Anzahl der von Null verschiedenen Urbilder multipliziert mit der Anzahl der Codes in  $\Phi_i$ ; es gilt somit  $(2^{n_i} - 1)N_i = (2^{k_i} - 1)|\Phi_i|$ .

Die Anzahl der Vektoren mit einem HAMMING-Gewicht kleiner als  $d$  ist  $A := \sum_{j=0}^{d-1} \binom{n_i}{j}$ . Die Anzahl aller Codes in  $\Phi_i$  mit einem HAMMING-Gewicht kleiner als  $d$  ist also durch  $N_i A$  beschränkt, da die Bedingung weitere Codewörter in dieser Kugel nicht gestattet. Ist jedoch  $N_i A < |\Phi_i|$ , so muß es noch weitere Codes in  $\Phi_i$  geben, die daher ein HAMMING-Gewicht von mindestens  $d$  besitzen. Man berechnet

$$\sum_{j=1}^{d-1} \binom{n_i}{j} < \frac{2^{n_i} - 1}{2^{k_i} - 1} \xrightarrow{n \rightarrow \infty} 2^{n_i - k_i} = 2^{\log_2 2^{n_i/k_i}} \leq 2^{\log_2 2^R} = 2^R. \quad (\text{C.48})$$

Mithilfe der CHERNOFF-Schranke (Satz C.34) folgt  $A < \sum_{j=0}^d \binom{n_i}{j} \leq 2^{n_i H(d/n)}$ , also genügt  $H(d/n) \leq (n_i - k_i)/n = 1 - k_i/n_i \leq 1 - R$  im Grenzfall  $n \rightarrow \infty$ , q. e. d.

# Schluß



# Bibliographie

Die Bibliographie verzeichnet alle Arbeiten (Aufsätze in wissenschaftlichen Zeitschriften und Sammelbänden, Lehrbücher und Monographien usw.), auf die aus dem laufenden Text verwiesen wird oder die während der Anfertigung dieser Dissertation verwendet wurden. Nicht zitierfähige Verweise finden sich nicht in der Bibliographie, sondern werden in Fußnoten angegeben; in diesem Fall werden die Inhalte in einer in sich geschlossenen Form wiedergegeben.

Angegeben werden in der Bibliographie in der Regel die vollständigen Namen der Verfasser (soweit feststellbar), selbst wenn sie sich nicht auf den genannten Arbeiten finden, der vollständige Titel des Werks sowie die Fundstelle, bei Zeitschriften der vollständige Name, der Band (in Fettdruck), die Heft-/Ausgabennummer mit Datum sowie die Seitenzahl oder die Artikelnummer mit Seitenlänge, bei Büchern der Verlag und die Auflage, die eingesehen wurde.

Die meisten neueren Arbeiten finden sich als (zum Teil gegenüber der gedruckten Fassung erweiterte oder berichtigte) sog. *Preprint*-Fassung unter <http://www.arxiv.org> (gelegentlich wesentlich früher und auch unter anderem Titel). In der Regel wird nicht gesondert hierauf verwiesen; Ausnahmen erfolgen, wenn es sich um eine bedeutende Langfassung handelt, die Arbeit nicht in einer wissenschaftlichen Zeitschrift veröffentlicht wurde oder sonst nur schwer zugänglich ist.

Ein Großteil der aufgeführten Literatur ist im Internet verfügbar, z. B. kostenfrei beim *Göttinger Digitalisierungszentrum* (GDZ) oder auch bei *DigiZeitschriften*, *BNF Gallica*, *NUMDAM* usw. oder kostenpflichtig über *SpringerLink* oder *JSTOR*, bekannte Arbeiten zum Teil auch an anderer Stelle. Von vielen Arbeiten finden sich Referate oder Zusammenfassungen in den Referateorganen *Mathematical Reviews* (Math. Rev., MR) und *Zentralblatt für Mathematik und ihre Grenzgebiete* (Zentralblatt MATH, Zbl.). Arbeiten älterer Autoren finden sich meist auch in ihren *Gesammelten Werken*, bedeutende Arbeiten sind auch oft in thematischen Sammelbänden abgedruckt.

- [1] ANTONIO ACÍN, JOONWOO BAE, EMILI BAGAN, MARIAÀ BAIG, LLUIS MASANES UND RAMON MUÑOZ-TAPIA:  
„*Secrecy properties of quantum channels*“  
Physical Review A **73** (Nr. 1, Januar 2006), Nr. 012327 (5 Seiten)
  
- [2] GERNOT ALBER, ALDO DELGADO, NICOLAS GISIN UND IGOR JEX:  
„*Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces*“  
Journal of Physics A **34** (Nr. 42, 26. Oktober 2001), S. 8821–8833

- [3] MICHAEL ARTIN: „*Algebra*“  
Birkhäuser, Basel 1998
- [4] ALAIN ASPECT, JEAN DALIBARD UND GÉRARD ROGER:  
„*Experimental Test of Bell's Inequalities using time-varying Analyzers*“  
Physical Review Letters **49** (Nr. 25, 20. Dez. 1982), S. 1804–1807
- [5] KOENRAAD M. R. AUDENAERT:  
„*A sharp continuity estimate for the von Neumann entropy*“  
Journal of Physics A **40** (Nr. 28, 13. Juli 2007), S. 8127–8136
- [6] JOONWOO BAE: „*Entanglement and Quantum Cryptography*“  
(Dissertation, Barcelona)  
<http://icfo.es/images/publications/junu.pdf> (Januar 2007)
- [7] JOONWOO BAE UND ANTONIO ACÍN:  
„*Key distillation from quantum channels using two-way communication protocols*“  
Physical Review A **75** (Nr. 1, Januar 2007), Nr. 012334 (20 Seiten)
- [8] LESLIE E. BALLENTINE:  
„*Quantum mechanics: A modern development*“  
World Scientific 2003
- [9] SOMSHUBHRO BANDYOPADHYAY, P. OSCAR BOYKIN, VWANI ROYCHOWDHURY UND FARROKH VATAN:  
„*A New Proof for the Existence of Mutually Unbiased Bases*“  
Algorithmica **34** (Nr. 4, November 2002), S. 512–528
- [10] FRIEDRICH LUDWIG BAUER:  
„*Entzifferte Geheimnisse: Methoden und Maximen der Kryptologie*“  
Springer-Verlag, 3. Auflage 2000
- [11] NORMAND BEAUDRY, TOBIAS MORODER UND NORBERT LÜTKENHAUS:  
„*Squashing Models for Optical Measurements in Quantum Communication*“  
Preprint arXiv:0804.3082v4 [quant-ph] 17 Sep 2008
- [12] CHRISTIAN BECK UND FRIEDRICH SCHLÖGEL:  
„*Thermodynamics of chaotic systems: an introduction*“  
Cambridge University Press 1993

- [13] JOHN STEWART BELL: „*On the Einstein-Podolsky-Rosen-Paradox*“  
Physics **1** (Nr. 3, 1964), S. 195–200
- [14] INGEMAR BENGTSSON UND KAROL ŻYCZKOWSKI:  
„*Geometry of Quantum States: An Introduction to Quantum Entanglement*“  
Cambridge University Press 2006
- [15] CHARLES H. BENNETT:  
„*Quantum Cryptography Using Any Two Nonorthogonal States*“  
Physical Review Letters **68** (Nr. 21, 25. Mai 1992), S. 3121–3124
- [16] CHARLES H. BENNETT, HERBERT J. BERNSTEIN, SANDU POPESCU UND BENJAMIN SCHUMACHER:  
„*Concentrating partial entanglement by local operations*“  
Physical Review A **53** (Nr. 4, April 1996), S. 2046–2052
- [17] CHARLES H. BENNETT UND GILLES BRASSARD:  
„*Quantum cryptography: Public key distribution and coin tossing*“  
Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing (Bangalore/Indien, Dezember 1984)  
IEEE Computer Society Press, Los Alamitos (Kalifornien/USA),  
S. 175–179
- [18] CHARLES H. BENNETT, DAVID P. DiVICENZO, JOHN A. SMOLIN  
UND WILLIAM K. WOOTTERS:  
„*Mixed state entanglement and quantum error correction*“  
Physical Review A **54** (Nr. 5, November 1996), S. 3824–3851
- [19] HERMANN BOERNER:  
„*Darstellungen von Gruppen. Mit besonderer Berücksichtigung der Bedürfnisse der modernen Physik*“  
Springer-Verlag, 2. Auflage 1967
- [20] DAVID BOHM UND YAKIR AHARONOV:  
„*Discussion of Experimental Proof for the Paradox of Einstein, Rosen and Podolsky*“  
Physical Review **108** (Nr. 4, 15. November 1957), S. 1070–1076
- [21] H. BOMBIN UND MIGUEL A. MARTIN-DELGADO:  
„*Entanglement distillation protocols and number theory*“  
Physical Review A **72** (Nr. 3, September 2005), Nr. 032313 (17 Seiten)

- [22] SIEGFRIED BOSCH: „*Algebra*“  
Springer-Verlag, 6. Auflage 2006
- [23] NICOLAS BOURBAKI: „*Éléments de Mathématique*“ (III. Buch),  
„*Topologie générale: Chapitres 1 à 4*“  
Hermann, Paris 1971
- [24] GILLES BRASSARD UND LOUIS SALVAIL:  
„*Secret-Key Reconciliation by Public Discussion*“  
Lecture Notes in Computer Science, Band 765, Springer-Verlag 1994,  
S. 410–423
- [25] DAGMAR BRUß:  
„*Optimal eavesdropping in quantum cryptography with six states*“  
Physical Review Letters **81** (Nr. 14, 5. Oktober 1998), S. 3018–3021
- [26] DAGMAR BRUß: „*Characterizing entanglement*“  
Journal of Mathematical Physics **43** (Nr. 9, September 2002),  
S. 4237–4251
- [27] JOHANNES BUCHMANN: „*Einführung in die Kryptographie*“  
Springer-Verlag, 3. Auflage 2003
- [28] A. ROBERT CALDERBANK UND PETER W. SHOR:  
„*Good quantum error-correcting codes exist*“  
Physical Review A **54** (Nr. 2, August 1996), S. 1098–1105
- [29] CARLTON M. CAVES, CHRISTOPHER A. FUCHS UND RÜDIGER  
SCHACK:  
„*Unknown quantum states: The quantum de Finetti representation*“  
Journal of Mathematical Physics **43** (Nr. 9, Sept. 2002), S. 4537–4559
- [30] HOI-FUNG CHAU:  
„*Practical scheme to share a secret key through a quantum channel  
with a 27.6% bit error rate*“  
Physical Review A **66** (Nr. 6, Dez. 2002), Nr. 060302(R) (4 Seiten)
- [31] HOI-FUNG CHAU:  
„*Unconditionally Secure Key Distribution in Higher Dimensions by  
Depolarization*“  
IEEE Transactions on Information Theory **51** (Nr. 4, April 2005),  
S. 1451–1468

- [32] HERMAN CHERNOFF:  
„*A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*“  
Annals of Mathematical Statistics **23** (Nr. 4, Dez. 1952), S. 493–507
- [33] MAN-DUEN CHOI: „*Positive linear maps on  $C^*$ -algebras*“  
Canadian Journal of Mathematics **24** (Nr. 3, 1972), S. 520–529
- [34] MAN-DUEN CHOI:  
„*Completely Positive Linear Maps on Complex Matrices*“  
Linear Algebra and its Applications **10** (1975), S. 285–290
- [35] MATTHIAS CHRISTANDL:  
„*The Structure of Bipartite Quantum States – Insights from Group Theory and Cryptography*“ (Dissertation, Cambridge)  
Preprint arXiv:quant-ph/0604183v1 25 Apr 2006
- [36] MATTHIAS CHRISTANDL, ARTUR EKERT, MICHAŁ UND PAWEŁ HORODECKI, JONATHAN OPPENHEIM UND RENATO RENNER:  
„*Unifying Classical and Quantum Key Distillation*“,  
erschienen bei SALIL P. VADHAN (Hrsg.): „*Theory of Cryptography*“,  
Lecture Notes in Computer Science, Band 4392, Springer-Verlag 2007,  
S. 456–477; auch Preprint arXiv:quant-ph/0608199v3 28 Feb 2007
- [37] MATTHIAS CHRISTANDL, ROBERT KÖNIG, GRAEME MITCHISON  
UND RENATO RENNER:  
„*One-and-a-Half Quantum de Finetti Theorems*“  
Communications in Mathematical Physics **273** (Nr. 2, Juli 2007),  
S. 473–498
- [38] MATTHIAS CHRISTANDL, RENATO RENNER UND ARTUR EKERT:  
„*A Generic Security Proof for Quantum Key Distribution*“  
Preprint arXiv:quant-ph/0402131v2 4 Mar 2004
- [39] LIEVEN CLARISSE:  
„*Characterization of distillability of entanglement in terms of positive maps*“  
Physical Review A **71** (Nr. 3, März 2005), Nr. 032332
- [40] LIEVEN CLARISSE:  
„*Entanglement Distillation: A Discourse on Bound Entanglement in Quantum Information Theory*“ (Dissertation, York)  
Preprint arXiv:quant-ph/0612072v1 10 Dec 2006

- [41] JOHN F. CLAUSER, MICHAEL A. HORNE, ABNER SHIMONY UND RICHARD A. HOLT:  
„*Proposed experiment to test local hidden-variable theories*“  
Physical Review Letters **23** (Nr. 15, 13. Oktober 1969), S. 880–884
- [42] JOHN BLIGH CONWAY: „*Functional analysis*“  
Springer-Verlag, 2. Auflage 1990
- [43] RICHARD COURANT UND DAVID HILBERT:  
„*Methoden der Mathematischen Physik*“ (zwei Bände)  
Springer-Verlag 1924 und 1937
- [44] IMRE CSISZÁR UND JÁNOS KÖRNER:  
„*Broadcast Channels with Confidential Messages*“  
IEEE Transactions on Information Theory **24** (Nr. 3, Mai 1978),  
S. 339–348
- [45] DAVID DEUTSCH, ARTUR EKERT, RICHARD JOSZA,  
CHIARA MACCHIAVELLO, SANDU POPESCU UND ANNA SANPERA:  
„*Quantum Privacy Amplification and the Security of Quantum  
Cryptography over Noisy Channels*“  
Physical Review Letters **77** (Nr. 13, 23. Sept. 1996), S. 2818–2821
- [46] IGOR DEVETAK UND ANDREAS WINTER:  
„*Distillation of secret key and entanglement from quantum states*“  
Proceedings of the Royal Society A **461** (Nr. 2053, 8. Januar 2005),  
S. 207–235
- [47] PERSI W. DIACONIS UND DAVID A. FREEDMAN:  
„*Finite exchangeable sequences*“  
Annals of Probability **8** (Nr. 4, August 1980), S. 745–764
- [48] ALBERT EINSTEIN, BORIS PODOLSKY UND NATHAN ROSEN:  
„*Can quantum-mechanical description of physical reality be considered  
complete?*“  
Physical Review **47** (Nr. 10, 15. März 1935), S. 777–780
- [49] ARTUR K. EKERT:  
„*Quantum cryptography based on Bell’s Theorem*“  
Physical Review Letters **67** (Nr. 6, 5. August 1991), S. 661–663
- [50] JÜRGEN ELSTRODT: „*Maß- und Integrationstheorie*“  
Springer-Verlag, 4. Auflage 2005

- [51] RICHARD PHILIPS FEYNMAN: „*Simulating physics with computers*“  
International Journal of Theoretical Physics **21** (Nr. 6–7, Juni 1982),  
S. 467–488
- [52] EUGEN FICK:  
„*Einführung in die Grundlagen der Quantentheorie*“  
Akademische Verlagsgesellschaft, Frankfurt am Main 1968
- [53] BRUNO DE FINETTI:  
„*La prévision: les lois logiques, ses sources subjectives*“  
Annales de l’Institut Henri Poincaré **7** (Nr. 1, 1937), S. 1–68
- [54] WILLIAM FULTON UND JOE HARRIS:  
„*Representation Theory – A First Course*“  
Springer-Verlag 1991
- [55] ИСРАЕЛ МОИСЕВИЧ ГЕЛЬФАНД (I. M. GELFAND) UND  
АНАТОЛИЙ ГОРДЕЕВИЧ КОСТЮЧЕНКО  
(A. G. KOSTJUTSCHENKO):  
„*Разложение по собственным функциям дифференциальных и  
других операторов*“ (Über die Entwicklung von Differential- und  
anderen Operatoren nach Eigenfunktionen)  
Доклады Академии Наук СССР **103** (Nr. 3, 21. Juli 1955),  
S. 349–352  
Englische Übersetzung:  
„*Eigenfunction expansions for differential and other operators*“ in  
IZRAIL M. GELFAND: „*Collected papers*“, Band I (Springer 1987),  
S. 505–509
- [56] ISRAEL MOJSEJEWITSCH GELFAND UND NAUM JAKOWLEWITSCH  
WILENKIN:  
„*Verallgemeinerte Funktionen (Distributionen)*“,  
Band IV: „*Einige Anwendungen der harmonischen Analyse.  
Gelfandsche Raumtripel*“  
VEB Deutscher Verlag der Wissenschaften, Berlin 1964
- [57] WALTER GELLERT, HERBERT KÜSTNER, MANFRED HELLWICH  
UND HERBERT KÄSTNER (Hrsg.):  
„*Kleine Enzyklopädie – Mathematik*“  
VEB Verlag Enzyklopädie Leipzig 1965, 2. Auflage 1967 (240.–265.  
Tausend der Gesamtauflage)

- [58] WALTER GELLERT, HERBERT KÜSTNER, WERNER SEIDEL UND KONRAD SENGLAUB (Hrsg.):  
„*Kleine Enzyklopädie – Natur*“  
VEB Bibliographisches Institut Leipzig, 16., überarb. Auflage 1966  
(1276.–1325. Tausend)
- [59] CHRISTIAN GERTHSEN: „*Physik*“  
Springer-Verlag, 20. Auflage 1999
- [60] EDGAR NELSON GILBERT: „*A Comparison of Signaling Alphabets*“  
Bell System Technical Journal **31** (Nr. 3, Mai 1952), S. 504–522
- [61] NICOLAS GISIN, GRÉGOIRE RIBORDY, WOLFGANG TITTEL UND HUGO ZBINDEN: „*Quantum cryptography*“  
Reviews of Modern Physics **71** (Nr. 1, Januar 2002), S. 145–195
- [62] ANDREW M. GLEASON:  
„*Measures on the Closed Subspaces of a Hilbert Space*“  
Journal of Mathematics and Mechanics **6** (Nr. 6, 1957), S. 885–893
- [63] BORIS VLADIMIROVICH GNEDENKO: „*The Theory of Probability*“  
Mir Publishers, Moskau 1969
- [64] DANIEL GOTTESMAN UND HOI-KWONG LO:  
„*Proof of Security of Quantum Key Distribution with two-way classical communications*“  
IEEE Transactions on Information Theory **49** (Nr. 2, Februar 2003),  
S. 457–475
- [65] DANIEL GOTTESMAN, HOI-KWONG LO, NORBERT LÜTKENHAUS UND JOHN PRESKILL:  
„*Security of quantum key distribution with imperfect devices*“  
Quantum Information and Computation **4**  
(Nr. 5, 8. September 2004), S. 325–360
- [66] DANIEL GOTTESMAN UND JOHN PRESKILL:  
„*Secure quantum key distribution using squeezed states*“  
Physical Review A **63** (Nr. 2, Februar 2001), Nr. 022309 (18 Seiten)
- [67] GERALD G. GOULD:  
„*The spectral representation of normal operators on a rigged hilbert space*“  
Journal of the London Mathematical Society **43** (1968), S. 745–754

- [68] MITSURU HAMADA:  
*„Reliability of Calderbank-Shor-Steane codes and security of quantum key distribution“*  
Journal of Physics A **37** (Nr. 34, 27. Aug. 2004), S. 8303–8328
- [69] RICHARD WESLEY HAMMING:  
*„Error Detecting and Error Correcting Codes“*  
Bell System Technical Journal **29** (Nr. 2, April 1950), S. 147–160
- [70] GODFREY HAROLD HARDY, JOHN EDENSOR LITTLEWOOD UND  
GEORGE PÓLYA: *„Inequalities“*  
Cambridge University Press, 2. Auflage 1952 und Nachdrucke
- [71] MATTHIAS HEID UND NORBERT LÜTKENHAUS:  
*„Efficiency of coherent-state quantum cryptography in the presence of loss: Influence of realistic error correction“*  
Physical Review A **73** (Nr. 5, Mai 2006), Nr. 052316 (7 Seiten)
- [72] WERNER HEISENBERG:  
*„Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik“*  
Zeitschrift für Physik **43** (Nr. 3–4, März 1927), S. 172–198
- [73] WERNER HEISENBERG:  
*„Die Entwicklung der Deutung der Quantentheorie“*  
Physikalische Blätter **12** (Nr. 7, Juli 1956), S. 289–304
- [74] HARRO HEUSER: *„Lehrbuch der Analysis“*  
B. G. Teubner: Teil 1 (15. Auflage, 2003), Teil 2 (13. Auflage, 2004)
- [75] WASSILY HOEFFDING:  
*„Probability Inequalities for Sums of Bounded Random Variables“*  
Journal of the American Statistical Association **58**  
(Nr. 301, März 1963), S. 13–30
- [76] OSKAR HÖFLING:  
*„Physik: Lehrbuch für Unterricht und Selbststudium“*  
Dümmler, Bonn, 14. Auflage 1985
- [77] АЛЕКСАНДР СЕМЁНОВИЧ ХОЛЕВО (A. S. HOLEVO/KHOLEVO):  
*„Информационные аспекты квантового измерения“*  
(Information Theory Aspects of Quantum Measurements)  
Проблемы передачи информации **9** (Nr. 2, April–Juni 1973),  
S. 31–42

„Некоторые оценки для количества информации передаваемого квантовым каналом связи“ (Some Estimates of Information Transmitted through Quantum Communication Channel)  
a. a. O. (Nr. 3, Juli–September 1973), S. 3–11

Englische Übersetzungen:

„Information-theoretical aspects of quantum measurement“  
Problems of Information Transmission **9** (1973), S. 110–118

„Bounds for the quantity of information transmitted by a quantum communication channel“  
a. a. O., S. 177–183

[78] ALEXANDER S. HOLEVO:

„The Capacity of the Quantum Channel with General Signal States“  
IEEE Transactions on Information Theory **44** (Nr. 1, Januar 1998),  
S. 269–273

[79] MICHAŁ UND PAWEŁ HOROECKI:

„Reduction criterion of separability and limits for a class of distillation protocols“  
Physical Review Letters **59** (Nr. 6, Juni 1999), S. 4206–4216

[80] MICHAŁ, PAWEŁ UND RYSZARD HOROECKI:

„Separability of mixed states: necessary and sufficient conditions“  
Physics Letters A **223** (November 1996), S. 1–8

[81] KAROL, MICHAŁ UND PAWEŁ HORODECKI UND JONATHAN OPPENHEIM:

„Secure Key from Bound Entanglement“  
Physical Review Letters **94** (Nr. 16, 29. April 2005), Nr. 160502

[82] KAROL, MICHAŁ UND PAWEŁ HORODECKI UND JONATHAN OPPENHEIM:

„General paradigm for distilling classical key from quantum states“  
Preprint arXiv:quant-ph/0506189v1 22 Jun 2005

[83] ROBIN LYTH HUDSON UND GRAHAM R. MOODY:

„Locally Normal Symmetric States and an Analogue of de Finetti's Theorem“  
Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **33**  
(1976), S. 343–351

- [84] ADOLF HURWITZ:  
*„Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt“*  
Mathematische Annalen **46** (Nr. 2, 1895), S. 273–284
- [85] ANDRZEJ JAMIOŁKOWSKI:  
*„Linear transformations which preserve trace and positive semidefiniteness of operators“*  
Reports on Mathematical Physics **3** (Nr. 4, Dez. 1972), S. 275–278
- [86] RICHARD JOZSA, DANIEL ROBB UND WILLIAM K. WOOTTERS:  
*„Lower bound for accessible information in quantum mechanics“*  
Physical Review A **49** (Nr. 2, Februar 1994), S. 668–677
- [87] DAVID A. KAMMLER: *„A first course in Fourier analysis“*  
Cambridge University Press 2000
- [88] OLIVER KERN:  
*„Randomized dynamical decoupling strategies and improved one-way key rates for quantum cryptography“* (Dissertation, TU Darmstadt)
- [89] OLIVER KERN UND JOSEPH M. RENES:  
*„Improved one-way rates for BB84 and 6-state protocols“*  
Quantum Information and Computation **8** (Nr. 8/9, September 2008), S. 756–772
- [90] MICHAEL KEYL:  
*„Fundamentals of Quantum Information Theory“*  
Physics Reports **369** (Nr. 5, Oktober 2002), S. 431–548
- [91] MICHAEL KEYL UND REINHARD F. WERNER: *„Channels and Maps“*, erschienen als Kapitel 5 bei DAGMAR BRUß UND GERD LEUCHS (Hrsg.): *„Lectures on Quantum Information“*, Wiley-VCH, November 2006, S. 73–86
- [92] AEYSHA KHALIQUE:  
*„Robustness bounds and practical limitations of quantum key distribution“* (Dissertation, TU Darmstadt, 2008)
- [93] AEYSHA KHALIQUE, GEORGIOS M. NIKOLOPOULOS UND GERNOT ALBER:  
*„Postponement of dark-count effects in practical quantum key-distribution by two-way post-processing“*  
European Physical Journal D **40** (Nr. 3, Dezember 2006), S. 453–464

- [94] ANDREAS KLAPPENECKER UND MARTIN RÖTTELER:  
„*Constructions of Mutually Unbiased Bases*“  
Preprint arXiv:quant-ph/0309120v1 15 Sep 2003
- [95] KONRAD KNOPP: „*Theorie und Anwendung der unendlichen Reihen*“  
Springer-Verlag, 4. Auflage 1947
- [96] RUDOLF KOCHENDÖRFFER:  
„*Lehrbuch der Gruppentheorie unter besonderer Berücksichtigung der  
endlichen Gruppen*“  
Akademische Verlagsgesellschaft, Leipzig 1966
- [97] MAX KOECHER: „*Lineare Algebra und analytische Geometrie*“  
Springer-Verlag, 4. Auflage 1997
- [98] ROBERT KÖNIG UND RENATO RENNER:  
„*A de Finetti representation for finite symmetric quantum states*“  
Journal of Mathematical Physics **46** (2005), Nr. 122108 (23 Seiten)
- [99] EMANUEL KNILL UND RAYMOND LAFLAMME:  
„*Theory of quantum error-correcting codes*“  
Physical Review A **55** (Nr. 2, Februar 1997), S. 900–911
- [100] KARL KRAUS: „*General state changes in quantum theory*“  
Annals of Physics **64** (Nr. 2, Juni 1971), S. 311–335
- [101] KARL KRAUS: „*States, effects, and operations*“  
Springer-Verlag 1983
- [102] BARBARA KRAUS, NICOLAS Gisin UND RENATO RENNER:  
„*Lower and Upper Bounds on the Secret-Key Rate for Quantum Key  
Distribution Protocols Using One-Way Classical Communication*“  
Physical Review Letters **95** (Nr. 8, 19. August 2005), Nr. 080501
- [103] ERWIN KREYSZIG: „*Statistische Methoden und ihre Anwendungen*“  
Vandenhoeck & Ruprecht/Göttingen, 5. Auflage 1975
- [104] MADDALY KRISHNA UND KALYANAPURAM RANGACHARI  
PARTHASARATHY:  
„*An entropic uncertainty relation principle for quantum  
measurements*“  
Sankhyā: The Indian Journal of Statistics A **64** (Nr. 3, 2002),  
S. 842–851

- [105] GERD LAßNER:  
„*Mathematische Beschreibung von Observablen-Zustandssystemen*“  
Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig,  
Mathematisch-Naturwissenschaftliche Reihe **22** (Nr. 2, 1973),  
S. 103–138
- [106] JÜRGEN LEHN UND HELMUT WEGMANN:  
„*Einführung in die Statistik*“  
B. G. Teubner, 4. Auflage 2004
- [107] RUDOLF LIDL UND HARALD NIEDERREITER: „*Finite fields*“  
Cambridge University Press 1984 (1. Nachdruck 1987)
- [108] GÖRAN LINDBLAD:  
„*Completely Positive Maps and Entropy Inequalities*“  
Communications in Mathematical Physics **40** (Nr. 2, Juni 1975),  
S. 147–151
- [109] JACOBUS HENDRICUS VAN LINT: „*Introduction to Coding Theory*“  
Springer-Verlag, 3. Auflage 1999
- [110] HOI-KWONG LO:  
„*Proof of unconditional security of six-state quantum key distribution scheme*“  
Quantum Information and Computation **1** (Nr. 2, August 2001),  
S. 81–94
- [111] HOI-KWONG LO:  
„*Method for decoupling error correction from privacy amplification*“  
New Journal of Physics **5** (2003), Nr. 36 (24 Seiten)
- [112] HOI-KWONG LO UND HOI FUNG CHAU:  
„*Unconditional Security of Quantum Key Distribution over arbitrarily long distances*“  
Science **283** (1999), S. 2050–2056 nebst den Anhängen in der  
Preprint-Fassung arXiv:quant-ph/9803006v5 6 Dez 1999
- [113] HOI-KWONG LO, HOI FUNG CHAU UND MOHAMMED ARDEHALI:  
„*Efficient quantum key distribution scheme and proof of its unconditional security*“  
Preprint arXiv:quant-ph/0011056v2 30 Nov 2001

- [114] NORBERT LÜTKENHAUS:  
„*Security against individual attacks for realistic quantum key distribution*“  
Physical Review A **61** (Nr. 5, Mai 2000), Nr. 052304 (10 Seiten)
- [115] HANS MAASSEN UND JOS B. M. UFFINK:  
„*Generalized Entropic Uncertainty Relations*“  
Physical Review Letters **60** (Nr. 12, 21. März 1988), S. 1103–1106
- [116] CHIARA MACCHIAVELLO:  
„*On the analytical convergence of the QPA procedure*“  
Physics Letters A **246** (September 1998), S. 385–388
- [117] FLORENCE J. MACWILLIAMS UND NEIL J. A. SLOANE:  
„*The Theory of error-correcting Codes*“  
North Holland Mathematical Library, 11. Druck 2003
- [118] MIGUEL ANGEL MARTÍN-DELGADO UND MIGUEL NAVASCUÉS:  
„*Distillation protocols for mixed states of multilevel qubits and the quantum renormalization group*“  
European Physical Journal D **27** (Nr. 2, November 2003), S. 169–180
- [119] UELI M. MAURER:  
„*Secret Key Agreement by Public Discussion from Common Information*“  
IEEE Transactions on Information Theory **39** (Nr. 3, Mai 1993),  
S. 733–742
- [120] DOMINIC MAYERS:  
„*Unconditional Security in Quantum Cryptography*“  
Journal of the Association for Computing Machinery **48**  
(Nr. 3, Mai 2001), S. 351–406
- [121] CHARLES A. MCCARTHY: „ $c_p$ “  
Israel Journal of Mathematics **5** (1967), S. 249–271
- [122] GEIR OVE MYHR UND NORBERT LÜETKENHAUS:  
„*Spectrum conditions for symmetric extendible states*“  
Preprint arXiv:0812.3667v1 [quant-ph] 19 Dec 2008
- [123] GEIR OVE MYHR, JOSEPH M. RENES, ANDREW C. DOHERTY  
UND NORBERT LÜETKENHAUS:  
„*Symmetric extension in two-way quantum key distribution*“  
Preprint arXiv:0812.3607v1 [quant-ph] 18 Dec 2008

- [124] JOHANN VON NEUMANN:  
„*Mathematische Grundlagen der Quantenmechanik*“  
Springer-Verlag 1932 (Nachdrucke 1968, 1996)
- [125] МАРК АРОНОВИЧ НАЙМАРК (M. A. NEUMARK/NAIMARK):  
„*О самосопряженных расширениях второго рода симметрического оператора*“ mit englischer Zusammenfassung  
„*Self-adjoint extensions of the second kind of a symmetric operator*“  
Известия Академии Наук СССР – Серия математическая – 4 (Nr. 1, 1940)<sup>14</sup>, S. 53–104  
„*Спектральные функции симметрического оператора*“ mit engl. Zusammenfassung „*Spectral functions of a symmetric operator*“ a. a. O. (Nr. 3, 1940), S. 277–318
- [126] SARAH R. NICHOLS UND WILLIAM K. WOOTTERS:  
„*Between entropy and subentropy*“  
Preprint arXiv:quant-ph/0207010v1 2 Jul 2002
- [127] MICHAEL A. NIELSEN UND ISSAC L. CHUANG:  
„*Quantum Computation and Quantum Information*“  
Cambridge University Press, 5. Druck 2000
- [128] GEORGIOS M. NIKOLOPOULOS UND GERNOT ALBER:  
„*Security bound of two-basis quantum-key-distribution protocols using qudits*“  
Physical Review A **72** (Nr. 3, September 2005), Nr. 032320 (10 Seiten)
- [129] GEORGIOS M. NIKOLOPOULOS, AEYSHA KHALIQUE UND GERNOT ALBER:  
„*Provable entanglement and information cost for qubit-based quantum key-distribution protocols*“  
European Physical Journal D **37** (Nr. 3, März 2006), S. 441–450
- [130] GEORGIOS M. NIKOLOPOULOS, KEDAR S. RANADE UND GERNOT ALBER:  
„*Error tolerance of two-basis quantum-key-distribution protocols using qudits and two-way classical communication*“  
Physical Review A **73** (Nr. 3, März 2006), Nr. 032325 (9 Seiten)

---

<sup>14</sup>Nebentitel *Bulletin de l'Académie de Sciences de l'URSS – Série mathématique* –

- [131] MASANORI OHYA UND DÉNEZ PETZ:  
„*Quantum Entropy and its Use*“  
Springer-Verlag, 2. korrigierter Druck 2004
- [132] VERN IVAL PAULSEN:  
„*Completely bounded maps and operator algebras*“  
Cambridge University Press 2003
- [133] ASHER PERES: „*Separability criterion for density matrices*“  
Physical Review Letters **77** (Nr. 8, 19. August 1996), S. 1413–1415
- [134] DÉNEZ PETZ:  
„*Quantum Information Theory and Quantum Statistics*“  
Springer-Verlag 2008
- [135] JOHN DE PILLIS:  
„*Linear transformations which preserve hermitian and positive semidefinite operators*“  
Pacific Journal of Mathematics **23** (Nr. 1, Oktober 1967), S. 129–137
- [136] GILLES PISIER UND QUANHUA XU: „*Non-Commutative  $L^p$ -spaces*“, erschienen als Kapitel 34 bei WILLIAM B. JOHNSON UND JORAM LINDENSTRAUSS (Hrsg.): „*Handbook of the Geometry of Banach spaces*“, Band 2, North Holland, März 2001, S. 1459–1517
- [137] JOHN PRESKILL: „*Quantum Information and Computation*“  
Vorlesungsskript (ca. 1998), verfügbar unter  
<http://www.theory.caltech.edu/people/preskill/>
- [138] KEDAR S. RANADE:  
„*Quantenkryptographie und Verschränkung – Zur maximal tolerierbaren Fehlerrate der Quantenkryptographie mit Zweiweg-Kommunikation im Schema von Gottesman und Lo*“  
(Diplomarbeit, TU Darmstadt, 28. April 2005)
- [139] KEDAR S. RANADE UND GERNOT ALBER:  
„*Asymptotic correctability of Bell-diagonal quantum states and maximum tolerable bit error rates*“  
Journal of Physics A **39** (Nr. 7, 17. Februar 2006), S. 1701–1716
- [140] KEDAR S. RANADE UND GERNOT ALBER:  
„*Asymptotic correctability of Bell-diagonal qudit states and lower bounds on tolerable error probabilities in quantum cryptography*“  
Journal of Physics A **40** (Nr. 1, 5. Januar 2007), S. 139–153

- [141] KEDAR S. RANADE UND MAZHAR ALI:  
„*The Jamiolkowski Isomorphism and a Simplified Proof for the Correspondence Between Vectors Having Schmidt Number  $k$  and  $k$ -Positive Maps*“  
Open Systems and Information Dynamics **14** (Nr. 4, Dezember 2007), S. 371–378
- [142] THOMAS VON RANDOW: „*Verschlüsselte Botschaften*“  
Bild der Wissenschaft (Nr. 3, März 1994), S. 32–36
- [143] MICHAEL REED UND BARRY SIMON:  
„*Methods of modern mathematical physics*“  
Band I: Functional Analysis (Revised and enlarged edition 1980)  
Band II: Fourier Analysis, Self-Adjointness (1975)
- [144] JOSEPH M. RENES:  
„*Spherical-code key-distribution protocols for qubits*“  
Physical Review A **70** (Nr. 5, November 2004), Nr. 052314 (4 Seiten)
- [145] JOSEPH M. RENES UND JEAN-CHRISTIAN BOILEAU:  
„*Strong Complementary Information Tradeoff*“  
Preprint arXiv:0806.3984v1 [quant-ph] 24 Jun 2008
- [146] JOSEPH M. RENES UND GRAEME SMITH:  
„*Noisy Processing and Distillation of Private Quantum States*“  
Physical Review Letters **98** (Nr. 2, 12. Januar 2007), Nr. 020502
- [147] RENATO RENNER:  
„*Security of Quantum Key Distribution*“ (Dissertation, ETH Zürich)  
Preprint arXiv:quant-ph/0512258v1 30 Dec 2005
- [148] RENATO RENNER, NICOLAS GISIN UND BARBARA KRAUS:  
„*Information-theoretic security proof for quantum-key-distribution protocols*“  
Physical Review A **72** (Nr. 1, Juli 2005), Nr. 012332
- [149] RENATO RENNER UND ROBERT KÖNIG:  
„*Universally Composable Privacy Amplification Against Quantum Adversaries*“  
Preprint arXiv:quant-ph/0403133v2 15 Apr 2004
- [150] ALFRED RÉNYI:  
„*Wahrscheinlichkeitsrechnung (mit einem Anhang über Informationstheorie)*“  
VEB Deutscher Verlag der Wissenschaften, Berlin 1962

- [151] MARCEL RIESZ:  
„*Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires*“  
Acta Mathematica (Uppsala) **49** (Nr. 3–4, 1926–27), S. 465–497
- [152] HOWARD PERCY ROBERTSON: „*The Uncertainty Principle*“  
Physical Review **34** (Nr. 1, 1. Juli 1929), S. 163–164
- [153] STEVEN ROMAN: „*Coding and Information Theory*“  
Springer-Verlag 1992
- [154] MARY BETH RUSKAI:  
„*Inequalities for quantum entropy: A review with conditions for equality*“ nebst Erratum  
Journal of Mathematical Physics **43** (Nr. 9, September 2002),  
S. 4358–4375 und **46** (Nr. 1, Januar 2005), Nr. 019901 (1 Seite)
- [155] DAVID SALGADO UND JOSÉ L. SÁNCHEZ-GÓMEZ:  
„*A simple proof of the Jamiolkowski criterion for complete positivity of linear maps of algebras of Hilbert-Schmidt operators*“  
Preprint math-ph/0411074v1 24 Nov 2004
- [156] DAVID SALGADO, JOSÉ L. SÁNCHEZ-GÓMEZ UND MIGUEL A. FERRERO:  
„*A Simple Proof of the Jamiolkowski Criterion for Complete Positivity of Linear Maps*“  
Open Systems and Information Dynamics **12** (Nr. 1, März 2005),  
S. 55–64
- [157] VALERIO SCARANI, ANTONIO ACÍN, GRÉGOIRE RIBORDY UND NICOLAS GISIN:  
„*Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations*“  
Physical Review Letters **92** (Nr. 5, 6. Februar 2004),  
Nr. 057901 (4 Seiten)
- [158] THOMAS SCHICKINGER UND ANGELIKA STEGER:  
„*Diskrete Strukturen*“,  
Band 2: Wahrscheinlichkeitstheorie und Statistik  
Springer-Verlag 2001

- [159] ERHARD SCHMIDT:  
*„Zur Theorie der linearen und nichtlinearen Integralgleichungen.  
I. Teil: Entwicklung willkürlicher Funktionen nach Systemen  
vorgeschriebener.“*  
Mathematische Annalen **63** (Nr. 4, 28. März 1907), S. 433–476
- [160] ERWIN SCHRÖDINGER: *„Zum HEISENBERG'schen Unschärfeprinzip“*  
Sitzungsberichte der Preußischen Akademie der Wissenschaften,  
Physikalisch-mathematische Klasse (Nr. 19, 19. Juni 1930),  
S. 296–303
- [161] ERWIN SCHRÖDINGER:  
*„Die gegenwärtige Situation in der Quantenmechanik“*  
Die Naturwissenschaften **23** (Nr. 48, 29. November 1935), S. 807–812;  
a. a. O. (Nr. 49, 6. Dezember 1935), S. 823–828;  
a. a. O. (Nr. 50, 13. Dezember 1935), S. 844–849
- [162] ERWIN SCHRÖDINGER:  
*„Discussion of probability relations between separated systems“*  
Proceedings of the Cambridge Philosophical Society **31**  
(Nr. 4, Oktober 1935), S. 555–563;  
*„Probability relations between separated systems“*  
a. a. O. **32** (Nr. 3, Juli 1936), S. 446–452
- [163] BENJAMIN SCHUMACHER UND MICHAEL D. WESTMORELAND:  
*„Sending classical information via noisy quantum channels“*  
Physical Review A **56** (Nr. 1, Juli 1997), S. 131–138
- [164] JEAN-PIERRE SERRE: *„Linear Representations of finite Groups“*  
Springer-Verlag 1977
- [165] CLAUDE ELWOOD SHANNON:  
*„A mathematical theory of communications“*  
Bell System Technical Journal **27** (Nr. 3, Juli 1948), S. 379–423, und  
a. a. O. (Nr. 4, Oktober 1948), S. 623–656
- [166] CLAUDE ELWOOD SHANNON:  
*„Communication Theory of Secrecy Systems“*  
Bell System Technical Journal **28** (Nr. 4, Oktober 1949), S. 656–715
- [167] PETER W. SHOR UND JOHN PRESKILL:  
*„Simple Proof of Security of the BB84 Quantum Key Distribution  
Protocol“*  
Physical Review Letters **85** (Nr. 2, 10. Juli 2000), S. 441–444

- [168] SIMON SINGH: „*The Code Book*“  
Archer Press 1999
- [169] ANDREAS SPEISER:  
„*Die Theorie der Gruppen von endlicher Ordnung*“  
Birkhäuser, Basel, 4. Auflage 1956
- [170] KURT STANGE: „*Angewandte Statistik*“ (zwei Bände),  
Springer-Verlag 1970 und 1971
- [171] ANDREW M. STEANE:  
„*Multiple Particle Interference and Quantum Error Correction*“  
Proceedings of the Royal Society A **452**  
(Nr. 1954, 8. November 1996), S. 2551–2577
- [172] WILLIAM FORREST STINESPRING:  
„*Positive functions on  $C^*$ -algebras*“  
Proceedings of the American Mathematical Society **6**  
(Nr. 2, April 1955), S. 211–216
- [173] ERLING STØRMER:  
„*Symmetric States of Infinite Tensor Products of  $C^*$ -algebras*“  
Journal of Functional Analysis **3** (1969), S. 48–68
- [174] MASAMICHI TAKESAKI: „*Theory of Operator Algebras I*“  
Springer-Verlag 1979, Nachdruck 2002
- [175] BARBARA M. TERHAL, ANDREW C. DOHERTY UND DAVID  
SCHWAB:  
„*Symmetric Extensions of Quantum States and Local Hidden Variable  
Theories*“  
Physical Review Letters **90** (Nr. 15, 18. April 2003), Nr. 157903  
(4 Seiten), Langfassung arXiv:quant-ph/0210053v2 15 Oct 2002
- [176] BARBARA M. TERHAL UND PAWEŁ HORODECKI:  
„*Schmidt number for density matrices*“  
Physical Review A **61** (Nr. 4, April 2000), Nr. 040301(R)
- [177] ARMIN UHLMANN: „*Sätze über Dichtematrizen*“  
Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig,  
Mathematisch-Naturwissenschaftliche Reihe **20** (Nr. 4/5, 1971),  
S. 633–637  
  
„*Endlich-dimensionale Dichtematrizen I*“ und „*II*“  
a. a. O. **21** (Nr. 4, 1972), S. 421–452; **22** (Nr. 2, 1973), S. 139–177

- [178] РОМ РУБЕНОВИЧ ВАРШАМОВ (R. R. VARSHAMOV):  
„Оценка числа сигналов в кодах с коррекцией ошибок“  
(The evaluation of signals in codes with correction of errors)  
Доклады Академии Наук СССР **117** (Nr. 4, 11. Dezember 1957),  
S. 739–741  
  
Englische Übersetzung:  
„Estimate of the number of signals in error correcting codes“  
in IAN F. BLAKE (Hrsg.): „Algebraic Coding Theory: History and  
Development“ (Dowden, Hutchinson & Ross 1973), S. 68–71
- [179] GILBERT SANDFORD VERNAM:  
„Cipher Printing Telegraph Systems For Secret Wire and Radio  
Telegraphic Communications“  
Journal of the American Institute of Electrical Engineers **55** (1926),  
S. 109–115
- [180] BARTEL LEENDERT VAN DER WAERDEN: „Algebra“ (Teil I)  
Springer-Verlag, 6. Auflage 1964
- [181] MARK N. WEGMAN UND J. LAWRENCE CARTER:  
„New hash functions and their use in authentication and set equality“  
Journal of Computer and System Sciences **22** (Nr. 3, Juni 1981),  
S. 265–279
- [182] ALFRED WEHRL: „General properties of entropy“  
Reviews of Modern Physics **50** (Nr. 2, April 1978), S. 221–260
- [183] GREGOR WEIHS, THOMAS JENNEWEIN, CHRISTOPH SIMON,  
HARALD WEINFURTER UND ANTON ZEILINGER:  
„Violation of Bell’s Inequality under Strict Einstein Locality  
Conditions“  
Physical Review Letters **81** (Nr. 23, 7. Dezember 1998), S. 5039–5043
- [184] DIRK WERNER: „Funktionalanalysis“  
Springer-Verlag, 5., erweiterte Auflage 2005
- [185] STEPHEN WIESNER: „Conjugate coding“  
ACM SIGACT News **15** (Nr. 1, Januar 1983), S. 78–88
- [186] WILLIAM K. WOOTTERS UND BRIAN D. FIELDS:  
„Optimal State-Determination by Mutually Unbiased Measurements“  
Annals of Physics **191** (Nr. 2, 1. Mai 1989), S. 363–381

- [187] RUDOLF ZURMÜHL: „*Matrizen und ihre technischen Anwendungen*“  
Springer-Verlag, 4. Auflage 1966

Kedar S. Ranade  
Hamburger Str. 12  
65760 Eschborn/Ts.  
Deutschland/Germany  
E-Mail: kedar.ranade@gmx.de



## Lebenslauf

### Angaben zur Person

Name und Vorname: RANADE, KEDAR SHRIKANT  
Geburtsdatum und -ort: 25. Juli 1979 in Berlin (West)  
Staatsangehörigkeit: deutsch  
Eltern: Dr.-Ing. SHRIKANT TRIMBAK RANADE und  
Dipl.-Ing. LATA RANADE, geborene LELE,  
Mag. tex. nauk (Moskau)

### Ausbildung und Beruf

08/1986 – 07/1990 Grundschule Süd-West in Eschborn am Taunus  
08/1990 – 06/1999 Albert-Einstein-Schule<sup>15</sup> (Gymnasium)  
in Schwalbach am Taunus  
23. Juni 1999 Abitur an der Albert-Einstein-Schule (Note 1,6)  
mit den Leistungskursen Chemie und Mathematik  
09/1999 – 06/2000 Wehrdienst beim Fernmeldeaufklärungsregiment 940  
in Daun (Fm-Aufklärer Tastfunk, Hauptgefreiter)  
07/2000 – 09/2000 Tätigkeit bei der Deutschen Bank in Eschborn  
10/2000 – 06/2005 Studium der Physik an der Technischen Universität  
Darmstadt  
22. Juni 2005 Diplom in Physik an der Technischen Universität  
Darmstadt (mit Auszeichnung bestanden)  
seit 09/2005 Promotionsstudium der Physik an der Technischen  
Universität Darmstadt  
seit 09/2008 Wissenschaftlicher Mitarbeiter im Fachbereich  
Physik der Technischen Universität Darmstadt

---

<sup>15</sup>Bis zum März 1992 trug die Schule den Namen *Eichwald-Gymnasium*.

## Wissenschaftliche Tätigkeit

### Diplomarbeit

Meine Diplomarbeit verfaßte ich am Fachbereich Physik der Technischen Universität Darmstadt in der Arbeitsgruppe „Theoretische Quantenphysik“ von Prof. Dr. GERNOT ALBER im Institut für Angewandte Physik. Ihr Titel lautet

#### Quantenkryptographie und Verschränkung

Zur maximal tolerierbaren Fehlerrate der Quantenkryptographie mit Zweiweg-Kommunikation im Schema von Gottesman und Lo

(fertiggestellt am 26. April 2005, eingereicht am 28. April 2005).

Die Hauptergebnisse der Diplomarbeit wurden als eigenständige Arbeit in einer begutachteten wissenschaftlichen Zeitschrift veröffentlicht [1].

### Veröffentlichungen

in begutachteten wissenschaftlichen Zeitschriften

In der folgenden Liste sind neben den Zeitschriftenveröffentlichungen mit ihren DOI-Angaben (Digital Object Identifier) auch die *Preprint*- Fassungen und eventuelle Referate in den Mathematical Reviews (MR) und im Zentralblatt für Mathematik und ihre Grenzgebiete (Zbl) angegeben.

Alle genannten Veröffentlichungen finden sich gegenwärtig unter der Adresse

<http://prp0.prp.physik.tu-darmstadt.de/~ranade/>

auf den Internet-Seiten des Verfassers.

- [1] KEDAR S. RANADE UND GERNOT ALBER:<sup>16</sup>  
„*Asymptotic correctability of Bell-diagonal quantum states and maximum tolerable bit error rates*“  
Journal of Physics A: Mathematical and General **39**  
(Nr. 7, 17. Februar 2006), S. 1701–1716  
DOI: <http://dx.doi.org/10.1088/0305-4470/39/7/014>  
auch Preprint [arXiv:quant-ph/0510041v1](http://arxiv.org/abs/quant-ph/0510041v1) 6 Oct 2005  
Referate „MR2210177 (2006m:81076)“ und „Zbl 1085.81033“

---

<sup>16</sup>Die Zusammenfassung findet sich auch in den „Verhandlungen der Deutschen Physikalischen Gesellschaft 7/2006 (AMOP-Frühjahrstagung 2006)“, Q 76.1 (S. 184).

- [2] GEORGIOS M. NIKOLOPOULOS, KEDAR S. RANADE UND GERNOT ALBER:  
*„Error tolerance of two-basis quantum-key-distribution protocols using qudits and two-way classical communication“*  
Physical Review A **73** (Nr. 3, März 2006), Nr. 032325 (9 Seiten)  
DOI: <http://dx.doi.org/10.1103/PhysRevA.73.032325>  
auch Virtual Journal of Quantum Information **6** (Nr. 4, April 2006)  
auch Preprint arXiv:quant-ph/0602008v1 1 Feb 2006
- [3] KEDAR S. RANADE UND GERNOT ALBER:<sup>17</sup>  
*„Asymptotic correctability of Bell-diagonal qudit states and lower bounds on tolerable error probabilities in quantum cryptography“*  
Journal of Physics A: Mathematical and Theoretical **40**  
(Nr. 1, 5. Januar 2007), S. 139–153  
DOI: <http://dx.doi.org/10.1088/1751-8113/40/1/008>  
auch Preprint arXiv:quant-ph/0609196v1 26 Sep 2006  
Referate „MR2304937 (2008b:94080)“ und „Zbl 1105.81020“
- [4] KEDAR S. RANADE UND MAZHAR ALI:  
*„The Jamiołkowski Isomorphism and a Simplified Proof for the Correspondence Between Vectors Having Schmidt Number  $k$  and  $k$ -Positive Maps“*  
Open Systems and Information Dynamics **14**  
(Nr. 4, Dezember 2007), S. 371–378  
DOI: <http://dx.doi.org/10.1007/s11080-007-9062-2>  
auch Preprint arXiv:quant-ph/0702255v1 27 Feb 2007  
(unter dem Titel *„The Jamiołkowski isomorphism and a conceptionally simple proof for the correspondence between vectors having Schmidt number  $k$  and  $k$ -positive maps“*)  
Referat „Zbl 1134.81342“

## Sonstige Veröffentlichungen

- [5] MAZHAR ALI, A. RAVI P. RAU UND KEDAR S. RANADE:  
*„Disentanglement in qubit-qudit systems“*  
Preprint arXiv:0710.2238v1 [quant-ph] 11 Oct 2007

---

<sup>17</sup>Die Zusammenfassung findet sich auch in den „Verhandlungen der Deutschen Physikalischen Gesellschaft 3/2007 (AMOP-Frühjahrstagung 2007)“, Q 35.6 (S. 188).

## Veröffentlichungen in Vorbereitung

- [6] KEDAR S. RANADE UND GERNOT ALBER:  
*„Symmetric extendibility for a class of qudit states“*
- [7] KEDAR S. RANADE UND GERNOT ALBER:  
*„Symmetric extendibility for qudits and tolerable error rates in quantum cryptography“*

## Teilnahme an Konferenzen

An den folgenden Konferenzen nahm ich teil und stellte – mit Ausnahme der SECOQC-QIT-Treffen und den Konferenzen in Hirschegg und Cuernavaca – jeweils ein Poster vor:

- Frühjahrstagung der Deutschen Physikalischen Gesellschaft (DPG),  
Sektion Atome, Moleküle, Optik und Plasma (AMOP):
  - Frankfurt am Main, 13.–17. März 2006
  - Düsseldorf, 19.–23. März 2007
  - Darmstadt, 17.–21. März 2008
- Treffen (Workshop) der SECOQC-QIT-Gruppe:
  - Erlangen, 21.–25. Februar 2005
  - Erlangen, 10.–14. Oktober 2005
  - Wien, 6.–10. März 2006
  - Wien, 23.–27. April 2007
  - Wien, 17.–20. Dezember 2007
  - Wien, 28.–31. Juli 2008
- International DFG Workshop „Quantum Entanglement – From Error Correction to Secure Key Distribution“, Hirschegg, 30. März – 2. April 2004
- International Workshop „Quantum Information“, Darmstadt, 14.–16. Dezember 2005
- International DFG Workshop „Quantum Information Processing“, Cochem an der Mosel, 28.–30. März 2007

- Informal Quantum Information Gathering (IQING) 5, Innsbruck, 11.–14. April 2007
- Workshop on modern trends in quantum optics and quantum information, Prag, 1.–4. Mai 2008
- Complexity of classical simulations of many body quantum dynamics, Cuernavaca/Mexiko, 1.–10. August 2008

## Vorträge

Die folgenden (zumindest teilweise öffentlichen) Vorträge wurden von mir gehalten:

- „*Error tolerance of two-basis QKD*“ (Wien, 8. März 2006)
- „*Quantum cryptography with qudits and two-way classical communication*“ (Wien, 25. April 2007)
- „*Quantum cryptography with qudits – Two-way error correction and tolerable error rates*“ (University of Cambridge, 21. Juni 2007)
- „*Verborgene Parameter, Kontextualität und das KOCHEN-SPECKER-Theorem*“ (TU Darmstadt, 24. Januar 2008)
- „*Quantum cryptography with qudits – Two-way error correction and tolerable error rates*“ (TU Prag, 29. April 2008)
- „*Quantum cryptography with qudits – Two-way error correction and tolerable error rates*“ (Cuernavaca, 5. August 2008)
- „*Quantenkryptographie in endlichdimensionalen Systemen*“ (TU Darmstadt, 16. Dezember 2008)

## Lehre

Übungsbetreuung der folgenden Lehrveranstaltungen (Vorlesungen):

- Theoretische Physik II: Quantenmechanik (SS 2005)
- Theoretische Quantenoptik (SS 2007)
- Theoretische Physik II: Quantenmechanik (SS 2008)

Zweitbegutachtung von Diplomarbeiten:

- ULRICH SEYFARTH: *Quantenkryptographie mit kontinuierlichen Variablen – Diskussion verschiedener Realisierungen mit Qudits* (Mai 2008)

## Sonstiges

- Sprachkenntnisse: Deutsch (muttersprachlich), Englisch (fließend), Französisch (Schulunterricht 1992–97), Latein (Latinum 1999), Marāṭhī (Grundkenntnisse), Russisch (UNICert, Stufe I)
- Wissenschaftliche Interessen: Physik (insbes. Grundlagen der Quantenmechanik und Quanteninformationstheorie), Mathematik (Funktionalanalysis), Kryptographie
- Mitgliedschaften: Deutsche Physikalische Gesellschaft
- Auszeichnungen: Buchpreis der Deutschen Physikalischen Gesellschaft für hervorragende Leistungen im Fach Physik 1999
- Gutachter für das European Physical Journal D

**Erklärung gemäß § 9 Abs. 1 Satz 5 und gemäß § 8 Abs. 1(c)**

der „Allgemeinen Bestimmungen der Promotionsordnung  
der Technischen Universität Darmstadt“  
vom 12. Januar 1990 (ABl. 1990, S. 658)  
in der Fassung der VI. Änderung vom 15. Februar 2006

Hiermit versichere ich an Eides Statt, daß ich die vorliegende Dissertation selbständig, nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfaßt habe. Ich erkläre ferner, daß ich bisher keinen Versuch unternommen habe, an einer anderen Hochschule das Promotionsverfahren einzuleiten.

*Kedar Ranade*

Darmstadt, den 28. Oktober 2008

Kedar S. Ranade

Satz und Gestaltung mit  $\text{\LaTeX}$  2 $\epsilon$  und Kile unter openSUSE 10.2  
durch den Verfasser; PDF-Erzeugung mit dvips und ps2pdf

Schriftart: Latin Modern 12pt auf DIN A4 (210mm  $\times$  297mm)