

Fachbereich Informatik und Mathematik  
ISMI - Institut für Stochastik  
& Mathematische Informatik

# **Geometrie für Anfänger**

**WS 2009/10**

H. Dinges

24. April 2010



---

## Einleitung und Vorwort

Die Geometrie ist eine ehrwürdige mathematische Disziplin mit vielfältigen Wurzeln. Sie nahm ihren Anfang bei den Techniken der Landvermessung im alten Orient. (Der Satz von Pythagoras ist eine Erkenntnis, die in diese Zeit zurückreicht.) Etwas später entwarfen die philosophisch motivierten Griechen eine mathematische Theorie der Geometrie. Diese Geometrie nach der Methode des Euklid wurde über Jahrhunderte systematisch gelehrt; neben den Fächern Arithmetik, Astronomie und Musik war sie ein Fach im ‘quadrivium’.

Die von den Griechen geforderte *inhaltliche Trennung* der Geometrie von der Arithmetik hatte lange Zeit Geltung; sie ist erst durch die sog. Arithmetisierung der Analysis im 19. Jahrhundert obsolet geworden. Auch die *methodische Unterscheidung*, welche synthetischer Geometrie gegen analytische Geometrie setzt, hat heute ihre Bedeutung verloren. Die Geometrie ist aber durch diese Zusammenführungen nicht in der Arithmetik oder Algebra aufgegangen. Trotz aller technischen Verwobenheit mit anderen Disziplinen gilt das *geometrische Denken* auch heute als ein mathematisches Grundprinzip eigener Art, welches es im Mathematikunterricht zu entwickeln gilt. Bei den verschiedensten Gelegenheiten stellen die Mathematiker fest, dass eine ‘geometrische Betrachtungsweise’ das ‘intuitive’ Verständnis mathematischer Aussagen entscheidend konsolidiert und das Weiterspinnen mathematischer Gedanken unterstützt.

Es ist anscheinend schwierig, einen ‘Kern’ der Geometrie dingfest zu machen und daraus einen Leitfaden für eine zeitgemäße Lehre herzuleiten. Lehrbücher der Geometrie, die sich um festgefügte Theorien bemühen, mögen manchen Puristen gefallen; sie erscheinen aber leicht ein wenig steril. Andere, lebendigere Geometriebücher erscheinen andererseits leicht als Sammlungen von (oft begeisternd eleganten) Kunstgriffen. Liebhaber gibt es für beide Typen von Herangehensweisen. Dagegen vertreten wir die Ansicht, dass für ein zeitgemäßes Geometrieverständnis die Systematik ebensowenig ausschlaggebend ist wie die unmittelbare Anschauung. Die Geometrie sollte man u. E. weder als eine stringente mathematische Methode verstehen, noch als einen festumrissenen Gegenstandsbereich. Geometrisches Denken verstehen wir lieber als einen sehr allgemeinen Zug mathematischen Denkens. Diese Auffassung gibt nun aber leider keine Richtschnur für eine Grundvorlesung Geometrie. Man hat (mit Recht) behauptet, dass man in (fast) jedem mathematischen Arbeitsgebiet eine geistige Kraft findet, die nicht einem Kalkül eingeordnet werden kann, und die, wenn sie auch nicht mit der traditionellen räumlichen Anschauung arbeitet, doch der räumlichen Anschauung eng verwandt ist. Ganz gleich, in welche Art von Mathematik man sich einarbeitet, in jedem Fall sollte die Befähigung zum (allgemein und abstrakt verstandenen) geometrischen Denken auf weitere Gelegenheiten das eigentliche Ziel der Bemühung sein. Diese Befähigung gilt es in (fast) allen mathematischen Anfängerveranstaltungen zu entwickeln. In unserer Anfängervorlesung wollen wir demgemäß gar nicht erst versuchen, ein Panorama der klassischen Geometrien zu entwerfen. Stattdessen entwickeln wir einige für viele Zusammenhänge bedeutsame mathematische Ideen, bei welchen die ‘geometrische Anschauung’ (nach unserer Einschätzung) eine überzeugende Rolle spielt.

Durch unser Vorgehen und unsere Stoffauswahl werden möglicherweise Erwartungen enttäuscht, die sich bei manchen Hörern aufgrund der Schulgeometrie herausgebildet haben könnten, die aber in der höheren Mathematik eine andere Wendung nehmen müssen. Die folgenden allgemeinen Gedanken zum Phänomen Geometrie versuchen diesen Hörern Anhaltspunkte für eine gewisse Orientierung zu geben. Ausgehend von den Ideen der traditionellen euklidischen Geometrie wollen wir vorab, auf untechnische Weise, die Wurzeln einiger heute wichtiger geometrischer Ideen andeuten.

**Figuren:** Die Schulgeometrie beginnt mit Figuren in der Ebene wie Geraden, Winkeln, Dreiecken, Kreisscheiben usw.. Das Ziel ist ein belastbares Verständnis für den uns umgebenden Raum. So kann die Mathematik heute nicht mehr ansetzen. Sie kann die Ebene und den Anschauungsraum nicht mehr als gegeben betrachten. Stattdessen konstruiert die Mathematik Räume (in mehr oder weniger loser Anlehnung an die traditionellen Vorstellungen), und sie wählt in diesen Räumen Klassen von Teilmengen aus, die man als Figuren versteht. Die Untersuchung konkreter Figuren in konkreten Räumen ist bis heute ein Thema der Geometrie. Auch wir werden unten solche Untersuchungen unternehmen. Wir werden insbesondere die Kreise und Geraden in der komplexen Ebene eingehend studieren, um daran Überlegungen zur sog. nichteuklidischen Geometrie anzuschließen. Wir werden auch erste Schritte in die sog. konvexe Geometrie hinein machen; die zugrundeliegenden Räume sind da reelle Vektorräume, die Figuren sind konvexe Mengen.

**Bewegungen:** Vom Standpunkt der euklidischen Schulgeometrie sind Figuren äquivalent, wenn sie kongruent sind, d. h., wenn die eine aus der anderen durch eine starre Bewegung hervorgeht. (In der sog. Ähnlichkeitsgeometrie sind auch noch Streckungen zugelassen.) Die an den Figuren abzulesenden ‘Größen’ bleiben bei den Bewegungen invariant. Die neueren Geometrien gehen in der Regel den umgekehrten Weg; sie beginnen mit der Festlegung einer Transformationsgruppe. Die Größen der betreffenden Geometrie sind diejenigen, die bei allen Transformationen unverändert bleiben. Im Fall der euklidischen Geometrie sieht das so aus: die Bewegungsgruppe wird von den Translationen und den Drehungen erzeugt; man stellt dann fest, dass die Länge die einzige (!) Distanzfunktion ist, welche gegenüber Translationen und Drehungen invariant ist. Man argumentiert weiter: Für die Länge gilt der Satz von Pythagoras (oder besser die sog. Parallelogrammgleichung). Und daraus leitet man den Begriff des inneren Produkts zweier Verschiebungen her sowie die Idee des Winkels.

Diese Sicht suggeriert Ansätze, welche über euklidische Geometrie hinausgehen: Gibt es bei den uns vorgegebenen Figuren noch tiefergehende Eigenschaften, die selbst bei drastischeren Transformationen als den starren Bewegungen nicht zerstört werden? Eine typische Wendung dieser Art ist die Topologie, die wir hier natürlich nicht weiter verfolgen können. Wir werden die affine und die projektive Geometrie kennenlernen sowie die dazugehörigen Bewegungsgruppen. Des weiteren werden wir uns auch in diesem Sinn näher beschäftigen mit den ‘Bewegungen’ der klassischen nichteuklidischen Geometrie, sowie mit der Gruppe der eigentlichen Lorentz-Transformationen (des Raum-Zeit-Kontinuums).

**Koordinaten:** Die ‘Methode der Koordinaten’ von Descartes ist eine Entdeckung von fundamentaler Bedeutung für die Geometrie. Auf die Koordinaten gründet sich seit Descartes die ‘analytische Geometrie’. Aus moderner Sicht sind Koordinaten (auf einem Grundraum) Funktionen auf diesem Raum mit der Eigenschaft, dass die Punkte des Raums durch die Spezifikation der Werte dieser Funktionen festgelegt sind. Die interessierenden Figuren werden durch Relationen zwischen den Koordinaten beschrieben. Die Flexibilität und die Kraft der auf die Koordinaten gegründeten ‘analytischen’ Herangehensweise ergibt sich daraus, dass man Funktionen addieren und multiplizieren kann. Im Fall reellwertiger Koordinaten kommen auch noch die Maximumsbildung und allerlei Grenzprozesse hinzu.

**Die klassischen Matrixgruppen;** Für die klassischen Geometrien gibt es ausgezeichnete Koordinatensysteme. Der Übergang von einem ausgezeichneten zu einem anderen ausgezeichneten Koordinatensystem wird durch eine Matrix beschrieben. Die dabei auftretenden Matrizen bilden eine Gruppe, und zwar eine Gruppe die man auch als eine Gruppe von Bewegungen verstehen kann. Matrizengruppen dieser Art sind heutzutage ein wichtiger Gegenstand der geometrischen Theorie.

**Synthetische Geometrien:** Die Geometrie von Euklid geht axiomatisch vor: die Lehrsätze werden durch logische Schlüsse aus einem Anfangssystem von Aussagen abgeleitet, die man Axiome oder Postulate nennt. Das Vorgehen von Euklid hat Mathematiker wie Philosophen aus vielen Jahrhunderten fasziniert, und es gewann eine immense Bedeutung für die Entwicklung der abendländischen Wissenschaft.

Die ersten Axiome von Euklid beziehen sich auf Punkte und Verbindungsgeraden in der Ebene, das dritte Axiom fordert, dass man zu jedem Mittelpunkt und jedem Radius den Kreis zeichnen kann. Das soll hier als ein Hinweis genügen; wir wollen nicht alle euklidischen Axiome auflisten oder gar diskutieren. Euklids Axiomensystem hatte Schwächen. Eine Aufbesserung, die allen modernen Anforderungen genügt, gelang erst D. Hilbert in seinen ‘Grundlagen der Geometrie’ (1899). Bei Hilbert wie bei Euklid handelt es sich um eine sog. synthetische Herangehensweise, es wird nicht (wie bei der sog. analytischen Herangehensweise) auf irgendwelche Zahlensysteme rekurriert.

Man hat im 20. Jahrhundert auf synthetischem Weg eine ganze Reihe von (mehr oder weniger schwachen) Geometrien etabliert. Hier geht es nicht mehr unbedingt um das bis ins 19. Jahrhundert gültige Anliegen, dass die Geometrie den uns umgebenden Raum beschreiben soll. Die Geraden haben auch nicht mehr unbedingt die wohlvertrauten Eigenschaften der sog. Zahlengeraden  $\mathbb{R}$ .

Das folgende (den modernen Anforderungen an eine axiomatisch gegebene geometrische Struktur entsprechende) Beispiel, (welches auf den ersten Blick zu Euklids Vorstellungen zu passen scheint), soll die Idee des synthetischen Aufbaus einer Geometrie verdeutlichen.

### Die affine Ebene:

Unter einer affinen Ebene versteht man eine Menge  $X$  (ihre Elemente heißen Punkte) zusammen mit einer Menge von Teilmengen von  $X$  (diese Teilmengen heißen Geraden), so dass die folgenden drei Axiome erfüllt sind:

- Zu verschiedenen Punkten  $p, q \in X$  gibt es genau eine Gerade  $Y \subset X$  mit  $p, q \in Y$ ,
- Ist  $Y$  eine Gerade und  $p$  ein Punkt, so gibt es durch  $p$  genau eine zu  $Y$  parallele Gerade  $Y'$ , die Geraden  $Y$  und  $Y'$  heißen parallel, wenn  $Y = Y'$  oder  $Y \cap Y' = \emptyset$ ,
- Es gibt drei Punkte, die nicht auf einer Geraden liegen.

Der Leser sollte sich an dieser Stelle daran erinnern, wie in Vorlesung über Lineare Algebra der Begriff des zweidimensionalen  $\mathbb{K}$ -affinen Raums eingeführt wird. ( $\mathbb{K}$  ist ein Körper mit  $1 + 1 \neq 0$ .) Der Weg von der synthetischen zur analytischen Begriffsbestimmung ist ein Thema in mathematischen Disziplinen, die im 20. Jahrhundert geblüht haben: ‘Grundlagen der Geometrie’ und ‘Geometrische Algebra’. Wir werden diesen Weg hier nicht gehen.

**Axiomatik:** Ein wesentlicher Grund für die Attraktivität der synthetischen Geometrien bis in die Mitte des 20. Jahrhunderts hinein war die Tatsache, dass die verschiedenen vor- oder nichteuklidischen Theorien zunächst einmal die einzigen Disziplinen waren, die konsequent von den Axiomen ausgingen. Der Idee des axiomatischen Aufbaus, die vielen mathematisch Interessierten als fundamental galt, hat nun aber seine Bindung an die Geometrie verloren; das axiomatische Vorgehen beherrscht heute zahllose Gebiete. Gleichzeitig hat sich das Verständnis von Axiomatik gewandelt. Axiomensysteme dienen nicht mehr (wie zu Euklids Zeiten) dazu, vermeintlich evidente Tatsachen zu identifizieren, um darauf auf logischen Weg ein Gebäude von Wahrheiten zu gründen. Sie dienen in der heutigen Mathematik dazu, axiomatisch abgegrenzte mathematische Objekte und Beziehungen, deren Tragweite noch auszuloten ist, sauber zu scheiden von dem, was man zu kennen glaubt.

Die synthetische Geometrie hat ihre Wertschätzung verloren. Wenn heute jemand Geometrie (fern von allen analytischen Elementen) synthetisch betreiben will, dann gilt das den meisten Mathematikern als Sektierertum. Das (möglicherweise für die allererste Schulgeometrie gültige) Argument, dass ein die Zahlen vermeidender synthetischer Zugang der Anschaulichkeit dient, kann heute nicht mehr überzeugen. Die moderne Auffassung von geometrischem Denken hat die Fiktion einer ‘natürlichen Anschauung des wirklichen Raums’ hinter sich gelassen. Die moderne Idee von Anschauung fordert, dass die Wissenschaftler auf der Grundlage fortgeschrittener analytischer Konstruktionen eine passende ‘Anschauung’ entwickeln, die für sie erkenntnisleitend ist nach dem Muster der traditionellen geometrische Anschauung.

**Die Vielfalt geometrischer Disziplinen:** Die Menge der mathematischer Systeme, die man heute zur Geometrie rechnet, ist kaum zu überschauen. Prominente umfassende

---

Gebiete sind die Topologie, die Differentialgeometrie und die Algebraische Geometrie. Der Gegenstand dieser geometrischen Disziplinen sind Klassen von Räumen  $M$ , auf welchen Funktionenmengen  $\mathfrak{F}$  mit besonderen Eigenschaften ausgezeichnet sind. Wenn  $(M_1, \mathfrak{F}_1)$  zwei Objekte sind, dann studiert man Abbildungen  $M_1 \rightarrow M_2$ , für welche der Pullback die Funktionen aus  $\mathfrak{F}_2$  in Funktionen aus  $\mathfrak{F}_1$  abbildet.

Nicht alle modernen Ansätze zu geometrischem Denken sind von der Analysis bestimmt: Bei den analysisfernen Geometrien handelt es sich vorwiegend um Lineare Geometrien. Die Räume sind da  $\mathbb{K}$ -Vektorräume oder daraus abgeleitete Räume wie  $\mathbb{K}$ -affine oder  $\mathbb{K}$ -projektive Räume. (Affine Räume sind Vektorräume, die der Nullpunkt ‘vergessen haben’; projektive Räume gewinnt man aus affinen Räumen durch die Hinzunahme unendlich ferner Punkte.) Die ausgezeichneten Funktionen sind in erster Linie die linearen Funktionen und in zweiter Linie die quadratischen Funktionen; (im Falle  $\mathbb{K} = \mathbb{C}$  die hermitesche Formen.)

Die klassischen linearen Geometrien werden üblicherweise in den Anfängerveranstaltungen zur Linearen Algebra angesprochen, leider oft nur als ein Anhängsel zur Vektorraumtheorie. Die früher hochgeschätzten Veranstaltungen und Lehrtexte mit dem Titel ‘Lineare Algebra und analytische Geometrie’ werden in neuester Zeit immer mehr um den Geometrieanteil gekürzt. Und man kann auch sonst feststellen, dass die Lehre der klassischen Geometrien in den vergangenen Jahrzehnten immer mehr eingeschränkt worden ist.

Die für die Lehre Verantwortlichen glauben offenbar davon ausgehen zu können, dass das geometrische Denken (dessen Bedeutung nicht bestritten wird) auch im Rahmen von aktuelleren Disziplinen vermittelt werden kann, unter Umgehung der Tradition von Euklid und in Distanz zu den festgefügt klassischen Geometrien. Die Kontroverse ist nicht entschieden. Einerseits ist es wohl wahr, dass man in (fast) jeder mathematischen Disziplin Anlässe für (im weiteren Sinn) geometrische Betrachtungen findet. Andererseits wird aber nicht nur von den Traditionalisten bestritten, dass das ein ausreichender Grund sei, auf die Bekanntschaft mit den traditionellen Lehrstoffen der Geometrie zu verzichten. Experimentieren ist angesagt.

### Vorwort. Die aktuelle Herausforderung

Der Frankfurter Fachbereich hat die gewohnte vierstündige Fortführung der Pflichtveranstaltung ‘Lineare Algebra’ (im zweiten Fachsemester) für das Bachelorstudium abgeschafft, und an ihre Stelle eine 2+1-stündige Pflichtveranstaltung ‘Geometrie’ gesetzt, (welche die Veranstaltung Lineare Algebra zum Modul ‘Grundstrukturen’ komplettiert). Es ist klar, dass es in dieser Veranstaltung nicht darum gehen kann, irgendwelche klassischen Geometrien von Grund auf zu entwickeln. Die hier dokumentierte Veranstaltung, die ausserhalb des Taktes im Wintersemester stattfand, und daher eine recht überschaubare Anzahl von Teilnehmern hatte, war auch Studierenden des Lehramts empfohlen. Sie ist als Pilotprojekt zu verstehen.

Nicht nur das Studienziel der Teilnehmer war uneinheitlich, auch die Vorbildung, denn nicht alle hatten dieselbe Pflichtveranstaltung ‘Lineare Algebra’ besucht. Mein Plan für die zweistündige Veranstaltung war der, ohne strikten systematischen Anspruch solche geometrische Gedanken zu vermitteln, die einen wichtigen Platz im breiten Feld der Elementarmathematik haben. Es schien mir vordringlich, geometrisches Argumentieren im Umgang mit solchen Gegenständen einzuüben, die im weiteren Studium einen hohen Stellenwert haben. Hervorragend geeignet erschienen mir die komplexen Zahlen, die in den neuen Studienplänen leider sehr in den Schatten getreten sind. Ich sehe weitere für das Einüben geometrischer Argumentation gut geeignete Themenbereiche, die aber vielleicht nicht alle in einer zweistündigen Veranstaltung Platz finden. Hervorheben möchte ich die konvexe Geometrie. Ich hätte sie in meiner Veranstaltung gerne weiter verfolgt; in diesem Bereich waren jedoch die Vorkenntnisse allzu unterschiedlich. Die Behandlung der projektiven Räume (über beliebigen Körpern) auf der anderen Seite würde ich nicht in den Kanon aufnehmen. Einen fertigen Plan für eine (2- oder 4-stündige ‘Geometrie’ im zweiten Fachsemester) habe ich nicht anzubieten; es wird wohl noch einiger didaktischer Experimente bedürfen, bis die Fachbereiche überzeugende Lehrpläne zur Geometrie gefunden haben. Vorerst wird es gut sein, sich mit lockeren Vorgaben zu begnügen. Es muss sich erst noch zeigen, wohin man die ‘Modulbeschreibungen’ entwickeln kann und will.

Die für das im Aufbau befindliche Frankfurter Bachelorstudium ist (noch sehr vorläufig!) folgendermaßen formuliert:

#### Die aktuelle Modulbeschreibung ‘Grundstrukturen’:

‘LINEARE ALGEBRA’ (4-stündig): *Vektorräume und affine Räume, euklidische Räume, Lineare Abbildungen und Matrizen, Determinanten und Eigenwerte, Lineare (Un-)Gleichungen, Konvexität.*

‘GEOMETRIE’ (2-stündig): *2- und 3-dimensionale Geometrie, (darstellende) euklidische Geometrie, projektive und hyperbolische Geometrie, endliche Konzepte.*

**Weiterführende LITERATUR :**

- [1] DUBROVIN, B. A., FOMENKO, A. T.; NOVIKOV S. P. : ‘Modern Geometrie - Methods and Applications’ Springer 1984.
- [2] BERGER, M. : ‘GEOMETRY I, II’ Springer 1980.
- [3] COURANT, R. u. ROBBINS, H.: ‘Was ist Mathematik ?’ 5. Aufl. Springer 1991.
- [4] COXETER, H. S. M. : ‘Reelle Projektive Geometrie der Ebene’ Oldenbourg 1955.
- [5] COXETER, H. S. M. : ‘Unvergängliche Geometrie’ Birkhäuser 1963.
- [6] DINGES, H.: ‘Einführung in die Mathematik I’ im Netz WS 2007/08.
- [7] KLINGENBERG, W. : ‘Lineare Algebra und Geometrie’ 3. Aufl. Springer 1992.
- [8] KOSTRIKIN, A. I. u. MANIN, Yu. I.: ‘Lineare Algebra und Geometrie’ Gordon and Breach 1988.
- [9] STRUIK, D. J. : ‘Abriss der Geschichte der Mathematik’ VEB Deutscher Verlag der Wissenschaften 1963. (Kapitel 16-18 zum 19. Jahrh.)
- [10] VALENTINE, F. A.: ‘Convex Sets’ McGraw-Hill Book Company 1964.

Man beachte: Die hier angegebenen Referenzen haben sehr verschiedenartigen Charakter. Alle geben Antworten auf die Frage, was es mit der Geometrie auf sich hat. Sie stellen sich aber nicht in der gleichen Weise und nicht in gleichem Maße der Tatsache, dass die Geometrie ein kulturelles Phänomen besonderer Art ist; immerhin war die Geometrie seit der Zeit der alten Griechen ein Lehrfach, welches das abendländische Denken immer wieder in besonderer Weise befruchtet hat.

Andererseits tragen auch nicht alle Texte der Tatsache Rechnung, dass heute in allen möglichen Studiengängen Prüfungen im Fach Geometrie gefordert werden. Nicht alle angegebenen Lehrbücher taugen für die Vorbereitung auf die Prüfungen. Ein auffälliger Grund sind insbesondere die deutlichen Unterschiede in der Nähe zur (linearen) Algebra und in ihrer Haltung zur geometrischen Anschauung.

## Inhaltsverzeichnis

<b>1</b>	<b>Komplexe <math>2 \times 2</math>-Matrizen und nichteuklidische Geometrie.</b>	<b>1</b>
1.1	Geometrische Aspekte des Systems der komplexen Zahlen . . . . .	1
1.2	Kreisverwandtschaften und die Riemann'sche Zahlenkugel. . . . .	5
1.3	Modelle für die klassischen nichteuklidischen Ebenen. . . . .	12
1.4	Lorentztransformationen. Quaternionen für die Drehungen. . . . .	21
1.5	Der hyperbolische Raum und seine Tangentialräume . . . . .	35
1.6	Die Bewegungen des hyperbolischen Raums . . . . .	49
1.7	Der Raum der Kreise . . . . .	63
1.8	Anhang: Gruppenwirkungen, insbesondere Konjugation . . . . .	68
<b>2</b>	<b>Die Geometrie des Hilbertraums</b>	<b>82</b>
2.1	Einordnung in die Welt der normierten Vektorräume . . . . .	83
2.2	Die Struktur des Hilbertraums. Orthogonalität . . . . .	90
2.3	Orthogonale Projektionen im Matrizenkalkül. Kleinste Quadrate. . . . .	98
2.4	Exkurs: Quadratische Formen im reellaffinen Raum . . . . .	106
2.5	Hermitische Formen im Hilbertraum. Spektralzerlegung. . . . .	117
2.6	Diagonalisierbare Endomorphismen . . . . .	124
2.7	Kontinuierliche Drehungen. Kommutatoren. . . . .	130
2.8	Selbstadjungierte Operatoren als Observable. . . . .	139

# 1 Komplexe $2 \times 2$ -Matrizen und nichteuklidische Geometrie.

Wenn man in den heute üblichen Anfängerveranstaltungen die reellen und die komplexen Zahlen behandelt, dann stehen in der Analysis die Ideen des Kontinuierlichen derart im Vordergrund, dass die algebraischen und geometrischen Aspekte kaum wahrgenommen werden. Auf der anderen Seite hat sich die zweite Grundvorlesung, die lange Jahre 'Lineare Algebra und analytische Geometrie' hiess, vielerorts zu einer Einführung in die Theorie der Vektorräume über beliebigen Körpern deformiert. Die besonderen Möglichkeiten einer geometrisch eingefärbten Argumentation, die durch die reellen und die komplexen Zahlen eröffnet werden, sind weit zurückgedrängt worden. In Opposition zu dieser Entwicklung wollen im ersten Kapitel dieser Vorlesung versuchen, für die komplexen Zahlen einen guten Platz in der Anfängerausbildung zu erkämpfen. Wir kommen dann auch sehr schnell zu den komplexen  $2 \times 2$ -Matrizen, und wir werden sehen, dass diese in vielen Bereichen der Geometrie einen zentralen Platz einnehmen.

## 1.1 Geometrische Aspekte des Systems der komplexen Zahlen

Das System  $\mathbb{C}$  der komplexen Zahlen wird üblicherweise (u. U. schon im Schulunterricht, sicher aber in der ersten mathematischen Anfängervorlesung) durch Zahlbereichserweiterung konstruiert. Im Vordergrund steht da meistens das Herausarbeiten der Körperstruktur.

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Das Systeme  $\mathbb{R}$  und  $\mathbb{C}$  tragen aber auch noch eine ganze Reihe von Strukturen, auf die wir hier noch einmal hinweisen wollen.

1) Die Menge  $\mathbb{R}$  der reellen Zahlen ist vollständig angeordnet, und jede nach oben beschränkte Teilmenge besitzt in  $\mathbb{R}$  eine kleinste obere Schranke, das sog. Supremum dieser Zahlenmenge.

2) In  $\mathbb{C}$  hat man die 'komplexe Konjugation'  $z \mapsto \bar{z}$  (man notiert auch  $z^*$  statt  $\bar{z}$ ). Die komplexe Konjugation lässt die reellen Zahlen unverändert. Für  $z \in \mathbb{C}$  sind

$$\Re z = \frac{1}{2}(z + z^*) \quad \text{und} \quad \Im z = \frac{1}{2i}(z - z^*)$$

reelle Zahlen. Wenn  $z$  und  $w$  komplexe Zahlen sind, dann gilt

$$\begin{aligned} \overline{(z + w)} &= \bar{z} + \bar{w} && (= (z + w)^*) \\ \overline{(z \cdot w)} &= \bar{w} \cdot \bar{z} && (= (z \cdot w)^*) \\ |z + w| &\leq |z| + |w| \\ |z \cdot w| &= |z| \cdot |w| \end{aligned}$$

Zu jeder komplexen Zahl  $\neq 0$  gibt es eine multiplikative Inverse; man nennt sie auch die reziproke Zahl;  $z^{-1} = \frac{\bar{z}}{|z|^2}$ .

3) Das System  $\mathbb{C}$  hat auch die Struktur eines zweidimensionalen reellen Vektorraums. Man spricht von der Gauss'schen Zahlenebene. Diese ist aber nicht nur einfach ein reellaffiner Raum, man hat darin ausgezeichnete Punkte, nämlich  $0, 1, i$ . Die Norm  $\sqrt{z \cdot \bar{z}}$  macht ihn zu einem euklidischen Raum. Für die komplexe Zahl  $z = a + ib$  mit  $a, b \in \mathbb{R}$  gilt

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z \cdot \bar{z}}; \quad \text{nur die Null hat die Norm} = 0.$$

4) Polarkoordinaten: Die Punkte der euklidischen Ebene kann man auch durch ihre Polarkoordinaten beschreiben:

$$(x, y) = (r \cdot \cos \varphi, r \cdot \sin \varphi).$$

Für  $r = \sqrt{r^2 + y^2} \neq 0$  ist der Winkel  $\varphi$  bis auf ein ganzzahliges Vielfaches von  $2\pi$  eindeutig bestimmt. (Man sagt, der Winkel  $\varphi$  ist ein wohlbestimmter Punkt in  $\mathbb{R}/2\pi$ )

Man notiert

$$z = |z| \cdot (\cos \varphi + i \sin \varphi) = |z| \cdot e^{i\varphi} \quad \text{mit} \quad \varphi \in \mathbb{R}/2\pi,$$

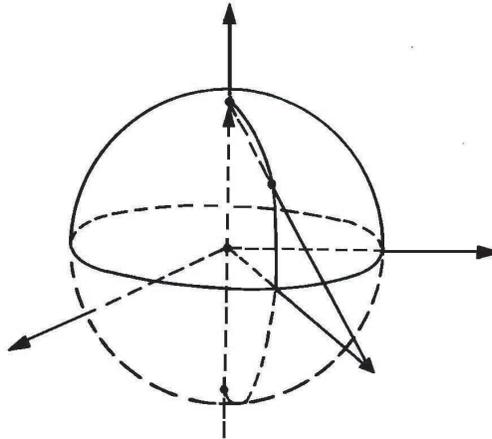
5) Wir fügen der komplexen Zahlenebene  $\mathbb{C}$  einen zusätzlichen Punkt hinzu, den „unendlich fernen“ Punkt  $\infty$ . So erhalten wir die erweiterte Zahlenebene  $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ . Anschaulich kann man die Menge  $\bar{\mathbb{C}}$  mit der sog. Riemann'schen Zahlenkugel identifizieren. Man darf sich hier wirklich eine Einheitssphäre im 3-dimensionalen Anschauungsraum vorstellen; und die sog. stereographische Projektion leistet die Identifikation mit der erweiterten Zahlenebene; und diese konkrete Entsprechung hat viele bemerkenswerte anschauliche Züge. Die Entsprechung ist z. B. winkeltreu, und die Kreise auf der Riemann'schen Zahlenkugel entsprechen den Kreisen (oder Geraden) in  $\bar{\mathbb{C}}$ . Die übliche geometrische Anschauung stellt hier kaum Fallen, wie wir sehen werden.

Schwierigkeiten macht nur die Interpretation der komplexen  $2 \times 2$ -Matrizen, welche zur Darstellung der 'Kreisverwandtschaften' herangezogen werden. Man darf sie nämlich ('anschaulich') nicht identifizieren mit den invertierbaren linearen Abbildungen des komplexen Vektorraums  $\mathbb{C}^2$  in sich. Wir werden nach und nach verstehen, dass die komplexen  $2 \times 2$ -Matrizen im Schnittpunkt mehrerer 'geometrischer' Theorien stehen, und dort jeweils verschiedene 'anschauliche' Interpretationen erfordern.

6) Die Riemann'sche Zahlenkugel: Es sei  $S^2$  die der Einheitssphäre, eingebettet in den euklidischen  $\mathbb{R}^3$ . Die stereographische Projektion setzt die Punkte der Sphäre  $S^2$  in Beziehung zu den Punkten der Äquatorebene. Das geschieht dadurch, dass jede Gerade durch den Nordpol, die nicht zur Äquatorebene parallel ist, sowohl die Einheitssphäre als auch die Äquatorebene in genau einem Punkt trifft. (Siehe Bild!) Wenn wir den Nordpol mit dem unendlich fernen Punkt von  $\bar{\mathbb{C}}$  in Beziehung setzen, dann haben wir eine Bijektion  $\bar{\mathbb{C}} \leftrightarrow S^2$ . In diesem Zusammenhang nennt man die zweidimensionale Einheitssphäre  $S^2$

die Riemann'sche Zahlenkugel. In kartesischen Koordinaten lautet die Bijektion

$$\mathbb{C} \ni z = x + iy \quad \longleftrightarrow \quad (tx, ty, 1 - t) \in S^2, \quad \text{wobei \u00fcrbrigens } t = \frac{2}{1 + |z|^2} .$$



Wir beweisen nun den

**Satz.** Die stereographische Projektion bildet die Kreise auf der Riemann'schen Zahlenkugel auf die Kreise und Geraden in  $\bar{\mathbb{C}}$  ab.

**Beweis.** Im euklidischen  $\mathbb{R}^3 = \{(x, y, u)\}$  sei  $K$  der Kegel mit dem Scheitel  $N = (0, 0, 1)$ , welche die  $(x, y)$ -Ebene im Kreis  $(x - x_0)^2 + (y - y_0)^2 = r^2$  schneidet.

Wir zeigen: Die Schnittpunkte dieses Kegels mit der Einheitssph\u00e4re  $\{x^2 + y^2 + u^2 = 1\}$  liegen in einer Ebene. Das Schnittgebilde ist daher ein Kreis auf der Sph\u00e4re.

Betrachten wir n\u00e4mlich die Gleichung des Kegels

$$(x - x_0(1 - u))^2 + (y - y_0(1 - u))^2 = r^2 \cdot (1 - u)^2 .$$

und subtrahieren wir die Gleichung der Einheitssph\u00e4re  $x^2 + y^2 + u^2 = 1 - u^2$ . Wir erhalten eine Gleichung, welche im Schnittgebilde  $K \cap S^2$  g\u00fcltig ist, n\u00e4mlich

$$-2x_0 \cdot x + x_0^2(1 - u) - 2y_0 \cdot y + y_0^2(1 - u) = r^2(1 - u) - (1 + u) .$$

Wenn man den Faktor  $1 - u$  ausklammert, erh\u00e4lt man die Gleichung einer Ebene. Das Schnittgebilde dieser Ebene mit der Einheitssph\u00e4re ist der gesuchte Kreis auf der Sph\u00e4re.

Eine wichtige Eigenschaft der stereographischen Projektion ist ihre Winkeltreue

**Satz.** Kreise in  $\bar{\mathbb{C}}$  schneiden sich in  $\bar{\mathbb{C}}$  im gleichen Winkel wie ihre stereographischen Bilder.

**Beweis.** Es gen\u00fcgt, Geraden zu studieren, die sich im Punkt  $z \in \bar{\mathbb{C}}$  im Winkel  $\alpha$  schneiden. Die stereographischen Bilder sind Kreise, die sich im Nordpol und im Bildpunkt von  $z$  schneiden. Da sie sich im Nordpol offensichtlich im Winkel  $\alpha$  schneiden, schneiden sie sich auch im Bildpunkt im Winkel  $\alpha$ .

### Einige Thesen:

Zu einer geometrischen Theorie (im traditionellen Verständnis) gehört zunächst einmal eine Klasse von ‘Figuren’, denen man besondere Aufmerksamkeit schenken will. Noch wichtiger allerdings ist die Gruppe der für diese Theorie charakteristischen ‘Bewegungen’. Genauer gesagt: Das zentrale Charakteristikum einer geometrischen Theorie ist eine Gruppe, die sich in allerlei Transformationsgruppen manifestiert.

In der traditionellen 3-dimensionalen euklidischen Geometrie beispielsweise geht es zunächst um Punkte, Geraden und Ebenen, (wobei nach dem modernen auf Hilberts ‘Grundlagen der Geometrie’ zurückgehenden Verständnis nicht gesagt werden muss, was nun unter einem Punkt, einer Geraden und einer Ebene zu verstehen ist; es kommt nur auf die durch die Axiome festgelegten Beziehungen an.) Für die ‘Kongruenz-Geometrie’ ist (neben der Gruppe der Translationen) die entscheidende Gruppe die orthogonale Gruppe  $O(3, \mathbb{R})$ . Die Elemente dieser Gruppe wirken als Bijektionen sowohl auf der Menge der Punkte als auch auf Menge der Geraden und auf der Menge der Ebenen, (sowie bei Bedarf auch noch auf weiteren Klassen von Figuren.)

In der ‘Geometrie zu  $\mathbb{C}$ ’, die wir in diesem Kapitel I entwickeln wollen, sind für uns neben den Punkten (zunächst einmal) die Kreise die besonderen Figuren (die Geraden werden zu den Kreisen gezählt). Die besonderen Bewegungen zu dieser Geometrie sind die sog. Kreisverwandtschaften, die wir in der nächsten Vorlesungsstunde studieren werden. Dahinter steckt die abstrakte Gruppe  $SL(2, \mathbb{C})$ . Diese Gruppe und einige ihrer Untergruppen (wie z. B. die spezielle unitäre Gruppe  $SU(2)$ ) werden wir im Laufe der Vorlesung immer wieder antreffen; sie spielt nämlich nicht nur in der ‘Geometrie zu  $\mathbb{C}$ ’ eine wichtige Rolle.

### Hinweis :

Die Gruppe  $SL(2, \mathbb{C})$  und die aus ihr abgeleiteten Gruppen, die hier diskutiert werden, sind glatte reelle Mannigfaltigkeiten im Sinne der Topologie und haben als solche eine **Dimension**. Der Dimensionsbegriff für (reelle) Mannigfaltigkeiten ist intuitiv nicht schwer zu erfassen; eine mathematisch solide Definition ist aber leider nicht so einfach wie beim Dimensionsbegriff für Vektorräume. Es sollte intuitiv verständlich sein, wenn wir berichten:

Die Gruppen  $SL(2, \mathbb{C})$  und  $\mathfrak{M}$  sind reell sechsdimensional: man braucht zunächst einmal vier komplexe, d.h. 8 reelle Zahlen, um eine komplexe  $2 \times 2$ -Matrix  $A$  festzulegen; durch die Bedingung  $\det A = 1$  wird die Freiheit eingeschränkt; es bleiben 6 Dimensionen. Die Gruppen  $SU(2)$  und  $SU(1, 1)$  sind reell dreidimensional. Um ein Element festzulegen, braucht man komplexe Zahlen  $(a, b)$  mit  $|a|^2 + |b|^2 = 1$  bzw.  $|a|^2 - |b|^2 = 1$ . Die Gruppen  $O(3, \mathbb{R})$  und  $SO(3, \mathbb{R})$  sind natürlich ebenfalls dreidimensional. Die sechsdimensionale spezielle lineare Gruppe  $SL(2, \mathbb{C})$  wird uns wieder begegnen, wenn wir uns mit der Raum-Zeit im Sinne der Relativitätstheorie befassen.

## 1.2 Kreisverwandtschaften und die Riemann'sche Zahlenkugel.

In dieser Vorlesung wollen wir einen speziellen Typ von Bijektionen der erweiterten Gauss'schen Zahlenebene  $\bar{\mathbb{C}}$  studieren. Diese Bijektionen heißen Möbiustransformationen oder auch Kreisverwandtschaften, weil sie Kreise in Kreise überführen. (Die Geraden gelten in diesem Zusammenhang als die Kreise durch den unendlich fernen Punkt).

**Sprechweise.** Die Operationen mit komplexen Zahlen  $a, b$  können dazu verwendet werden, spezielle Bijektionen von  $\bar{\mathbb{C}}$  zu definieren.

- Die Abbildung  $\tau_a : z \mapsto z + a$  heisst die Translation um  $a \in \mathbb{C}$ .
- Die Abbildung  $\sigma_b : z \mapsto b \cdot z$  heisst die Drehstreckung zu  $b \neq 0$ .
- Die Abbildung  $\rho : z \mapsto \frac{1}{z}$  heisst die Reziprokenabbildung.

Die Translationen und Drehstreckungen bilden (definitionsgemäss!) den unendlichfernen Punkt auf sich ab. Für die Reziprokenabbildung ist festgelegt:  $\rho(\infty) = 0$ ,  $\rho(0) = \infty$ .

**Definition.** Die Gruppe der Möbiustransformationen  $\mathfrak{M}$  ist die Gruppe der Bijektionen von  $\bar{\mathbb{C}}$ , die von den Translationen, den Drehstreckungen und der Reziprokenabbildung erzeugt wird.

Die Elemente von  $\mathfrak{M}$  nennt man auch die gebrochen linearen Abbildungen von  $\bar{\mathbb{C}}$ .

**Satz.**

- Jede invertierbare komplexe  $2 \times 2$ -Matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  liefert eine Möbiustransformation  $\varphi_A(\cdot) : z \mapsto \frac{az+b}{cz+d}$
- Zu jeder Möbiustransformation  $\varphi(\cdot)$  existieren genau zwei darstellende Matrizen  $\pm A$  mit  $\det A = 1$ .

**Beweis.** Wir zeigen zunächst, dass es zu den elementaren Möbiustransformationen darstellende Matrizen mit der Determinante = 1 gibt

- $\tau_a(\cdot) = \varphi_A(\cdot)$  mit  $A = \pm \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$
- $\sigma_b(\cdot) = \varphi_B(\cdot)$  mit  $B = \frac{1}{\pm\sqrt{b}} \cdot \begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix}$
- $\rho(\cdot) = \varphi_C(\cdot)$  mit  $C = \pm i \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Wir zeigen  $\varphi_B(\varphi_A(\cdot)) = \varphi_{B \cdot A}(\cdot)$ . Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \text{also} \quad B \cdot A = \begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + \delta d \end{pmatrix} = \begin{pmatrix} s & t \\ u & v \end{pmatrix}$$

Wir haben dann

$$\varphi_B(\varphi_A(z)) = \frac{\alpha \cdot \frac{az+b}{cz+d} + \beta}{\gamma \cdot \frac{az+b}{cz+d} + \delta} = \frac{\alpha(az+b) + \beta(cz+d)}{\gamma(az+b) + \delta(cz+d)} = \frac{sz+t}{uz+v}.$$

*Insbesondere gilt: Wenn die Matrix  $A$  die Abbildung  $\varphi(\cdot)$  darstellt, dann stellt die inverse Matrix  $A^{-1}$  die Umkehrabbildung  $\varphi^{-1}(\cdot)$  dar.*

*Wenn zwei Matrizen  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , und  $B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  dieselbe gebrochene lineare Abbildung  $\varphi(\cdot)$  darstellen, dann sind die entsprechenden Zeilen bis auf einen Faktor gleich; denn die Linearfaktoren des quadratischen Polynoms  $(a \cdot z + b) \cdot (\gamma \cdot z + \delta) = q(z) = (\alpha \cdot z + \beta) \cdot (c \cdot z + d)$  sind bis auf einen Faktor gleich. Die Forderung  $\det A = 1 = \det B$  legt die Faktoren bis auf das Vorzeichen fest.*

*Wir zeigen noch, dass man jede lineargebrochene Abbildung  $\varphi_A$  durch Hintereinanderschalten von vier elementaren Möbiustransformationen gewinnen kann. Im Falle  $c = 0$ , das ist der Fall, wo der unendlichferne Punkt Fixpunkt ist, genügt sogar schon eine Drehstreckung gefolgt von einer Translation. Im Falle  $c \neq 0$ , (o. B. d.  $A \cdot c = 1$ ) haben wir*

$$\frac{az + b}{z + d} = \frac{a(z + d) + b - ad}{z + d} = a + \frac{b - ad}{z + d}.$$

*Wir verschieben also um  $d$ , wenden die Reziprokenabbildung an, multiplizieren mit  $(b - ad)$  und verschieben um  $a$ . Vier elementare Möbiustransformationen reichen aus.*

## Die Bilder von Kreisen

Es ist offensichtlich, dass Translationen und Drehstreckungen Kreise in Kreise abbilden und Gerade in Gerade. Wir zeigen, dass auch die Reziprokenabbildung Kreise in Kreise abbildet, wenn man auch die Geraden als Kreise gelten lässt, nämlich als Kreise durch den unendlichfernen Punkt. Wenn wir das gezeigt haben, dann ist klar, dass jede Möbiustransformation Kreise in Kreise abbildet.

**Satz.** *Wenn  $K = \{z : z \in K\}$  ein Kreis in  $\bar{\mathbb{C}}$  ist, dann ist auch  $K' = \{\frac{1}{z} : z \in K\}$  ein Kreis.*

**Beweis.** *Die Kreislinie mit dem Mittelpunkt  $z_0 \in \mathbb{C}$  und dem Radius  $r > 0$  ist die Lösungsmenge der Gleichung  $|z - z_0|^2 = r^2$ . Die Gleichung kann auch auf folgende Weisen geschrieben werden:*

$$(z - z_0)(\bar{z} - \bar{z}_0) - r^2 = 0 \quad \text{oder} \quad z \cdot \bar{z} - z_0 \cdot \bar{z} - \bar{z}_0 \cdot z + |z_0|^2 - r^2 = 0, \quad \text{oder auch} \\ a \cdot z\bar{z} + \bar{\alpha}z + \alpha\bar{z} + b = 0, \quad \text{mit} \quad a, b \in \mathbb{R}, \alpha \in \mathbb{C}, D := |\alpha|^2 - ab > 0.$$

*Wir bemerken: Die letzte Form passt auch für die 'Kreise durch den Punkt  $\infty$ '; für  $a = 0$  erhalten wir nämlich eine Geradengleichung*

$$\bar{\alpha}z + \alpha\bar{z} + b = 0 \iff (\bar{\alpha} + \alpha)x + (\bar{\alpha} - \alpha)iy + b = 0 \iff \\ (2\Re\alpha) \cdot x + (2\Im\alpha)y + b = 0.$$

*Wenn  $z \in \mathbb{C} \setminus \{0, \infty\}$  die obige Kreisgleichung erfüllt, dann erfüllt  $w = \frac{1}{z}$  die Gleichung*

$$a \cdot \frac{1}{w\bar{w}} + \bar{\alpha} \cdot \frac{1}{w} + \alpha \frac{1}{\bar{w}} + b = 0, \quad \text{oder} \quad b \cdot w\bar{w} + \alpha w + \bar{\alpha}\bar{w} + a = 0, \\ \text{mit} \quad a, b \in \mathbb{R}, \alpha \in \mathbb{C}, D := |\alpha|^2 - ab > 0.$$

Das ist die gewünschte Kreisgleichung für die Bildmenge. Die Gleichungen passen (richtig verstanden) auch auf die Fälle, wo die Kreise durch 0 oder  $\infty$  gehen.

Wir führen noch einen weiteren Beweis, einen Beweis, welcher auf die übliche Parameterdarstellung der Kreise Bezug nimmt. Es genügt, für jeden Kreis um den Mittelpunkt 1 zu zeigen, dass das Bild bzgl. der Reziprokenabbildung ein Kreis ist. Der Kreis mit dem Radius  $r$  um den Punkt 1 schneidet die reelle Achse in den Punkten  $1+r$  und  $1-r$ . Die Reziprokenabbildung macht aus diesem Kreis eine Kurve, welche die reelle Achse in den Punkten  $\frac{1}{1+r}$  und  $\frac{1}{1-r}$  schneidet. Im Fall  $r=1$  ist das Bild des Kreises die Gerade  $\{w : \Re w = \frac{1}{2}\}$ ; denn

$$\frac{1}{1+e^{it}} = \frac{1}{2} + \frac{i}{2} \cdot \tan(t/2)$$

Im Fall  $r \neq 1$  ist zu zeigen: das Bild ist der Kreis mit dem Mittelpunkt  $M = \frac{1}{2}(\frac{1}{1+r} + \frac{1}{1-r}) = \frac{1}{1-r^2}$  und dem Radius  $R = \frac{1}{2}|\frac{1}{1+r} - \frac{1}{1-r}| = |\frac{r}{1-r^2}|$ . In Formeln

$$\{\frac{1}{z} : z = 1 + r \cdot e^{it} \quad \text{mit } t \in \mathbb{R}\} = \{w : w = M + R \cdot e^{is} \quad \text{mit } s \in \mathbb{R}\}$$

Der Nachweis ist eine leichte Rechnung.

$$\begin{aligned} \frac{1}{1+r \cdot e^{it}} - M &= \frac{1}{1+re^{it}} - \frac{1}{1-r^2} = R \cdot g(t) \quad \text{mit} \\ g(t) &= \frac{1-r^2}{r} \left[ \frac{1}{1+re^{it}} - \frac{1}{1-r^2} \right] = \frac{1}{r(1+re^{it})} [(1-r^2) - (1+re^{it})] = (-1)e^{it} \frac{1+r \cdot e^{-it}}{1+r \cdot e^{it}} \end{aligned}$$

$g(t)$  hat offenbar den Betrag 1 für alle reellen  $t$ .

### Übertragung auf die Zahlenkugel

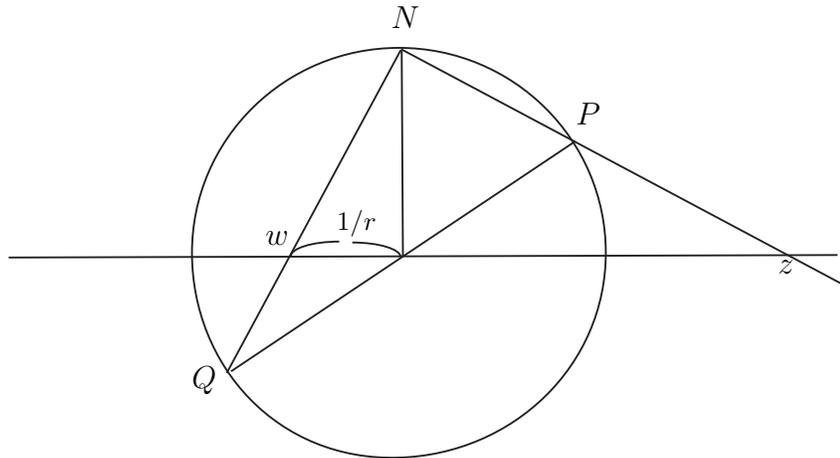
Die eben betrachteten Kreisverwandtschaften über der Grundmenge  $\bar{\mathbb{C}}$  liefern (vermittelt durch die stereographische Projektion) Kreisverwandtschaften über der Einheitskugel  $S^2$ . Die Beschreibung in Koordinaten mag für manche Abbildungen komplizierter sein; dafür kann aber die anschauliche Beschreibung manchmal überzeugender sein. Manche Möbius-Transformationen liefern einfach Drehungen der Einheitskugel, wie wir sehen werden. Ausserdem verliert der unendlichferne Punkt von  $\bar{\mathbb{C}}$  (bzw. der Nordpol der Zahlenkugel) seinen scheinbaren Sonderstatus; die Zahlenkugel ist als ein kompakter Raum zu verstehen. Die Wirkung der Reziprokenabbildung ergibt sich aus der folgenden elementargeometrischen Überlegung über die Spiegelung am Einheitskreis. (Die komplexe Konjugation entspricht der Spiegelung an der  $xz$ -Ebene.)

**Satz.** Die Spiegelung am Einheitskreis entspricht im stereographischen Bild der Spiegelung an der Äquatorebene.

**Beweis.** Zwei komplexe Zahlen  $z, w$  gehören zu Diametralpunkten  $P, Q$  auf der Riemann'schen Zahlenkugel, wenn gilt  $z \cdot \bar{w} = -1$ . Für  $z = r \cdot e^{i\varphi}$  bedeutet das nämlich

$w = -\frac{1}{r}e^{i\varphi}$ , und die Skizze zeigt die Behauptung, wenn man sich an den elementargeometrischen Satz erinnert, der besagt:

Die Höhe eines rechtwinkligen Dreiecks teilt die Hypotenuse so, dass das Produkt der Abschnitte das Quadrat der Höhe ist.



Als Konsequenz ergibt sich übrigens ein dritter Beweis, dass die Reziprokenabbildung Kreise in Kreise abbildet.

**Doppelverhältnisse** Die Möbiustransformationen sind als Bijektionen von  $\bar{\mathbb{C}}$  definiert. Sie liefern aber natürlich auch injektive Abbildungen der Menge aller geordneten Punktpaare, der Menge aller geordneten Punkttripel usw. Wir wollen hier die Möbiustransformationen noch als Abbildungen von Punkttripeln studieren.

**Satz.** (*‘Transitivitätstheorem’*) Die Gruppe  $\mathfrak{M}$  aller Möbiustransformationen wirkt einfach transitiv auf der Menge  $\Omega$  aller geordneten Punkttripel.

Genauer gesagt: Wenn  $z_1, z_2, z_3$  und  $w_1, w_2, w_3$  in  $\bar{\mathbb{C}}$  paarweise verschieden sind, dann gibt es genau eine Möbiustransformation  $\varphi(\cdot)$  mit

$$\varphi(z_1) = w_1, \varphi(z_2) = w_2, \varphi(z_3) = w_3 .$$

Der Beweis bedarf einiger Vorbereitung

**Definition.** („Doppelverhältnis“)

Für vier paarweise verschiedene komplexe Zahlen definiert man das Doppelverhältnis

$$DV(z_1, z_2, z_3, z_4) = \frac{z_3 - z_1}{z_3 - z_2} \cdot \frac{z_4 - z_2}{z_4 - z_1} .$$

**Bemerke :**

Man definiert das Doppelverhältnis auch (in naheliegender Weise), wenn eines der Argumente der unendlich ferne Punkt ist, und auch, wenn ein Paar der Argumente gleich ist. Es gibt anscheinend keine allgemein akzeptierte Konvention, in welcher die Argumente aufzuzählen sind. Das Doppelverhältnis bleibt ungeändert, wenn man die Argumente gemäß der sog. Klein'schen Vierergruppe permutiert. Demgemäß können für die permutierten Argumente höchstens 6 verschiedene Zahlenwerte auftreten. Das sind die Werte

$$\lambda, \quad \frac{1}{\lambda}, \quad 1 - \lambda, \quad \frac{1}{1 - \lambda}, \quad \frac{\lambda}{\lambda - 1}, \quad \frac{\lambda - 1}{\lambda}.$$

**Hinweis :** In der projektiven Geometrie ordnet man jedem Quadrupel von Punkten auf einer Geraden ein Doppelverhältnis zu. Wir wollen die Parallele hier nicht diskutieren.

**Satz.**

Werden die Zahlen  $z_1, z_2, z_3, z_4 \in \bar{\mathbb{C}}$  durch eine Möbiustransformation in die Zahlen  $w_1, w_2, w_3, w_4$  abgebildet, so gilt

$$DV(z_1, z_2, z_3, z_4) = DV(w_1, w_2, w_3, w_4).$$

(„Das Doppelverhältnis ist eine Invariante gegenüber allen Kreisverwandtschaften.“)

**Beweis.** Es genügt, die Invarianz für die elementaren Möbiustransformationen nachzuweisen. Für die Translationen und die Drehstreckungen ist die Behauptung trivial. Für die Reziprokenabbildung  $z \mapsto \frac{1}{z}$  ist sie leicht nachzurechnen.

**Beweis des Transitivitätstheorems**

Wenn  $\varphi(\cdot)$  die Punkte  $z_1, z_2, z_3$  in die Punkte  $w_1, w_2, w_3$  abbildet, dann liegt das Bild  $w = \varphi(z)$  für jedes  $z$  fest. Die Gleichung  $DV(z_1, z_2, z_3, z) = DV(w_1, w_2, w_3, w)$  kann man nämlich nach  $w$  auflösen,  $w$  ergibt sich aus  $z$  durch eine lineare gebrochene Abbildung  $w = \frac{az+b}{cz+d}$ .

Wenn wir z. B.  $(z_1, z_2, z_3)$  durch die Möbiustransformation  $\varphi(\cdot)$  nach  $(0, \infty, 1)$  abbilden möchten, dann ergibt sich  $w = \varphi(z)$  aus der Gleichung

$$DV(0, \infty, w, 1) = w = \frac{z - z_1}{z - z_2} \cdot \frac{z_3 - z_2}{z_3 - z_1} = DV(z_1, z_2, z, z_3)$$

Da wir jedes Tripel paarweise verschiedener Punkte in umkehrbarer Weise in das Tripel  $(0\infty, 1)$  abbilden können, können wir es auch in jedes andere Tripel paarweise verschiedener Punkte abbilden. Es gibt genau eine Möbiustransformation, die das leistet.

### Corollar

Die Punkte  $z_1, z_2, z_3, z_4$  liegen genau dann einem Kreis, wenn ihr Doppelverhältnis reell ist.

**Beweis.** Für Punkte auf der reellen Achse ist das Doppelverhältnis trivialerweise reell. Jeder Kreis kann durch eine Möbiustransformation auf die reelle Achse abgebildet werden.

### Hinweise:

1) Im berühmten Buch ‘Unvergängliche Geometrie’ von Coxeter trägt das sechste Kapitel die Überschrift ‘Kreise und Kugeln’ (Seiten 104 - 127). Man kann ‘cum grano salis’ sagen, dass dort die Inhalte der hier entwickelten ‘Geometrie zu  $\bar{\mathbb{C}}$ ’ auf synthetischem anschaulichem Weg entwickelt werden.

Der methodische Unterschied zu unserem Ansatz ist aber gravierend. In Coxeters Buch wird der Kalkül der komplexen Zahlen nicht einbezogen; und die Algebra kommt auch nicht mit dem Begriff einer Matrix (oder einer Matrizen­gruppe) ins Bild. Die Anschauung bleibt ganz auf dem Boden der klassischen Euklidischen Geometrie.

Hier und in vielen weiteren Kapiteln der ‘Unvergänglichen Geometrie’ kann man erfahren, dass die Geometrie als eine brillante, aber museal abgeschlossene Geistesübung praktiziert werden kann. Es scheint mir nicht ausgemacht, dass sich der große didaktische Erfolg dieser Herangehensweise ins 21. Jahrhundert fortsetzen wird.

2) Unsere Herangehensweise stellt insofern eine Abkehr von den alten Idealen des Geometrieunterrichts dar, als wir nicht primär darauf achten wollen, dass neue mathematische Aussagen ‘methodenrein’ durch stringente Ableitungen auf einer möglichst schmalen axiomatischen Basis bewiesen werden. Wir legen mehr Wert darauf, dass die Zusammenhänge (mit den gerade passenden Mitteln) einsichtig gemacht werden. Der Philosoph B. Spinoza (1632 – 1677), dem es bekanntlich wichtig war, seine Ethik ‘more geometrico’ darzustellen, hätte das sicher als einen Verrat am Geist der Geometrie empfunden.

3) (Für diejenigen, die in der ‘Linearen Algebra’ bereits einiges über projektive Geometrie erfahren haben.) Eine wichtige Methode, einen sog.  $\mathbb{K}$ -projektiven Raum der Dimension  $n$  zu konstruieren, besteht bekanntlich darin, einen  $n$ -dimensionalen  $\mathbb{K}$ -affinen Raum um die sog. unendlichfernen Punkte zu erweitern. Besonders einfach erscheint der Fall  $n = 1$ , der Fall der  $\mathbb{K}$ -projektiven Geraden also. Man sollte sich aber bzgl der Anschauung nicht allzu sehr in Sicherheit wiegen; schon der Fall  $\mathbb{K} = \mathbb{C}$  stellt eine Herausforderung für die Anschauung dar, wie die eben entwickelte ‘Geometrie zu  $\bar{\mathbb{C}}$ ’ zeigt.

4) Eine alternative Konstruktionsmöglichkeit eines  $n$ -dimensionalen  $\mathbb{K}$ -projektiven Raum bietet bekanntlich die Identifikation der Punkte mit den eindimensionalen linearen Teilräumen eines  $(n + 1)$ -dimensionalen  $\mathbb{K}$ -Vektorraums. Die erweiterte Gauss’sche Zahlenebene  $\bar{\mathbb{C}}$  kann als der eindimensionale komplexe projektive Raum verstanden werden, indem man ihre Punkte mit den eindimensionalen (komplexen !) Teilvektorräume im

Raum  $\mathbb{C}^2$  identifiziert.

$$\left\{ a \cdot \begin{pmatrix} z_1 \\ z_0 \end{pmatrix} : a \in \mathbb{C} \right\} \longleftrightarrow z = \frac{z_1}{z_0} \quad \text{für } z_0 \neq 0; \quad \text{und } \longleftrightarrow \infty, \quad \text{falls } z_0 = 0, z_1 \neq 0.$$

Die 'projektiven Abbildungen' werden durch Äquivalenzklassen invertierbarer linearer Abbildungen eines zweidimensionalen komplexen Vektorraums beschrieben. Wenn man den Vektorraum als den Raum der komplexen 2-Spalten präsentiert, dann entsprechen den projektiven Abbildungen die Äquivalenzklassen invertierbarer komplexer  $2 \times 2$ -Matrizen (wobei Matrizen, die sich nur um einen Faktor unterscheiden als äquivalent zu betrachten sind). Jede invertierbare komplexe  $2 \times 2$ -Matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  liefert eine invertierbare lineare Abbildung des Vektorraums  $\mathbb{C}^2$  auf sich. Und die Äquivalenzklasse liefert eine lineargebrochene Abbildung der erweiterten Zahlenebene  $\bar{\mathbb{C}}$  auf sich.

$$\begin{pmatrix} z_1 \\ z_0 \end{pmatrix} \longmapsto A \cdot \begin{pmatrix} z_1 \\ z_0 \end{pmatrix}; \quad z = \frac{z_1}{z_0} \longmapsto \frac{az + b}{cz + d}.$$

Die Gruppe der Projektivitäten entspricht also der Gruppe der Möbiustransformationen. Diesen Gesichtspunkt wollen wir hier aber nicht weiter verfolgen.

### 1.3 Modelle für die klassischen nichteuklidischen Ebenen.

Die Geometrie von Euklid handelt von Punkten, Geraden, usw. .

Euklid gibt Definitionen dieser Objekte; er sagt z. B.

- Ein Punkt ist, was keine Teile hat.
- Eine Linie ist eine breitenlose Länge.
- Eine Gerade ist eine Linie, die bezüglich der Punkte auf ihr stets gleich liegt.

Euklid formuliert für die Geometrie fünf Postulate (man spricht heute von Axiomen). Gefordert wird,

1. dass man von jedem Punkt nach jedem Punkt die Strecke ziehen könne
2. dass man eine begrenzte gerade Linie gerade verlängern könne
3. dass man mit jedem Mittelpunkt und Abstand einen Kreis zeichnen könne
4. dass alle rechten Winkel einander gleich seien
5. dass zu einer Geraden in einer Ebene durch einen Punkt ausserhalb höchstens eine diese nicht schneidende Gerade existiere. ('Parallelenaxiom')

Zweck der Definitionen (soweit sie die Grundbegriffe betreffen) ist es bei Euklid, den Bezug zu vertrauten geometrischen Erfahrungen herzustellen und die Postulate zu motivieren. Es ist nicht zu übersehen, dass diese 'Definitionen' für den Aufbau von Euklids Theorie ohne Bedeutung sind. Hilbert, der 1899 in seinen 'Grundlagen der Geometrie' das euklidische System auf unanfechtbare Beine gestellt hat, sagt sehr pointiert, es sei für die Theorie irrelevant, was die Punkte, Geraden und Ebenen mit der Wirklichkeit zu tun haben. Die Bedeutung der Grundbegriffe sei allein dadurch bestimmt, dass diese die Axiome erfüllen.

Andererseits erklärt Hilbert in der Einleitung seiner Arbeit: *Die vorliegende Untersuchung ist ein neuer Versuch, für die Geometrie ein vollständiges und möglichst einfaches System von Axiomen aufzustellen.* Offenkundig geht es also Hilbert nicht um einen beliebigen Formalismus, sondern um eine Präzisierung dessen, was Euklid mit 'Geometrie' gemeint hat, und was wir alle als die Eigenschaften des uns umgebenden Raums kennen oder zu kennen glauben.

Schon im Altertum ist die Frage aufgeworfen worden, ob das Parallelenaxiom unabhängig ist oder nicht. Wenn es eine logische Folge der übrigen Axiome ist, dann müsste es möglich sein, dass man es aus der Liste der Axiome streicht und mittels der anderen Axiome beweist. Verschiedene Autoren versuchten das Parallelenaxiom durch die indirekte Methode zu beweisen, d. h. das Gegenteil anzunehmen und daraus unsinnige Konsequenzen abzuleiten. Dies gelang nicht, und es konnte nicht gelingen, weil, wie sich später herausgestellt hat, das Parallelaxiom tatsächlich logisch unabhängig ist. Der Russe N. I. Lobatchevski (1793 – 1856) erarbeitete eine Theorie auf der Grundlage des alternativen

Axioms, dass durch einen Punkt ausserhalb einer Geraden mindestens zwei Geraden existieren, die mit dieser in einer Ebene liegen und sie nicht schneiden. Ihm und unabhängig von ihm auch dem Ungarn Janos Bolyai (1802 – 1860) gelang es, auf der Grundlage des alternativen Axioms in allen Einzelheiten eine in sich stimmige Theorie aufzubauen. Gauss, der unbestrittene ‘princeps mathematicorum’, hatte, wie er dem jungen Bolyai mitteilte, die Ergebnisse schon früher gefunden aber nicht veröffentlicht, weil er aufsehenerregende Publizität scheute. Immerhin war es eine der grundlegenden Annahmen der Kantschen Philosophie (so, wie sie damals weithin verstanden wurde), dass die euklidischen Axiome als unumstößliche Wahrheiten dem Reich der reinen Anschauung angehören.

Um die Widerspruchsfreiheit der neuen Geometrie zu zeigen, genügt es natürlich nicht, eine große Anzahl nichtabsurder nichteuklidischer Sätze abzuleiten, wie das Bolyai und Lobatschewski taten. Endgültige Überzeugungskraft haben erst ‘Modelle’, die alle Axiome Euklids mit Ausnahme des Parallelaxioms befriedigen. Ein besonders einfaches derartiges Modell wurde auf der Grundlage von Ideen des englischen Geometers Arthur Cayley (1821 – 1895) von Felix Klein (1849 – 1925) angegeben.

Modelle der gewünschten Art werden konstruiert, indem man zuerst Objekte der gewöhnlichen euklidischen Geometrie betrachtet, und dann einige dieser Objekte und die Beziehungen zwischen ihnen so umbenennt, dass eine nichteuklidische Geometrie entsteht. Bei der Entwicklung der Theorie ist es natürlich unerwünscht, das intuitive Verständnis an ein spezielles Modell zu knüpfen; die ‘Anschauung’ sollte sich allein von dem nähren, was in den Axiomen steckt, sodass auch neue Interpretationen oder ‘Anwendungen’ ins Bild passen. Ein zentrales (vom Modell unabhängiges) Element der neuen Theorie ist das System der ‘nichteuklidischen Bewegungen’, welche in der neuen Theorie die Idee der Kongruenz begründet.

Intuitive Ansätze für allgemeinere nichteuklidische Geometrien ergeben sich mit Hilfe der (aus der Physik stammenden) Vorstellung, dass die Geraden den Lichtstrahlen in einem Medium entsprechen, in welchem die Lichtgeschwindigkeit vom Ort abhängt. Die allgemeinere Idee gehört in die sog. Riemann’sche Geometrie, mit der wir uns hier nicht befassen können. Elementare Hinweise zum Spezialfall der hyperbolischen Geometrie im Sinne von Lobatschewski findet man z. B. im Buch von Courant und Robbins.

Wir werden Im Folgenden auf der Grundlage der oben entwickelten ‘Geometrie zu  $\bar{\mathbb{C}}$ ’ einige Blicke werfen auf die beiden klassischen nichteuklidischen Geometrien, die die elliptische bzw. die hyperbolischen Ebene genannt werden. Bei Coxeter finden wir (auf Seite 126) die Anmerkung, dass die Namen nicht in unmittelbarer Beziehung stehen zu den bekannten Kurven Ellipse und Hyperbel; man habe sich bei der Bezeichnung leiten lassen von der folgenden Analogie: In der hyperbolischen Geometrie hat jede Gerade zwei unendlichferne Punkte, so wie die Hyperbel zwei Asymptoten hat. Die Ellipse hat keine Asymptoten, und in der elliptischen Geometrie gibt es keine unendlichfernen Punkte; zwei Gerade haben da immer einen Schnittpunkt.

Die beiden klassischen nichteuklidischen Geometrien stehen in enger Beziehung zu zwei Gruppen, die uns (wie auch eine Reihe verwandter Matrizen Gruppen) auch noch anderswo interessieren werden. Es scheint an der Zeit, etwas weiter auszuholen.

**Definition.** (Klassische Gruppen)

a) Die allgemeine lineare Gruppe  $GL(n, \mathbb{K})$  besteht aus den nichtsingulären  $n \times n$ -Matrizen mit Einträgen aus  $\mathbb{K}$ .

b) Die spezielle lineare Gruppe  $SL(n, \mathbb{K})$  besteht aus den nichtsingulären  $n \times n$ -Matrizen mit der Determinante  $= 1$ .

c) Die orthogonale Gruppe  $O(n, \mathbb{K})$  besteht aus den  $n \times n$ -Matrizen  $A$  mit  $A^T A = E_n$ . (Das Produkt mit der transponierten Matrix ist die Einheitsmatrix.)

d) Die spezielle orthogonale Gruppe  $SO(n, \mathbb{K})$  besteht aus den orthogonalen  $n \times n$ -Matrizen  $A$  mit  $\det A = 1$ .

e) Wenn  $\mathbb{K} = \mathbb{R}$ , dann notiert man auch  $O(n)$  statt  $O(n, \mathbb{R})$ , sowie  $SO(n)$  statt  $SO(n, \mathbb{R})$ .

f) Wenn  $\mathbb{K} = \mathbb{C}$ , dann definiert man die unitäre Gruppe  $U(n)$  oder  $U(n, \mathbb{C})$  als die Gruppe der komplexen  $n \times n$ -Matrizen mit  $A^* A = E_n$ . (Das Produkt mit der hermitisch konjugierten Matrix ist die Einheitsmatrix.)

g) Die unitären Matrizen mit der Determinante  $= 1$  bilden die Gruppe  $SU(n)$  (oder  $SU(n, \mathbb{C})$ ). (Man bemerke, dass man die unitären Matrizen aus den speziellen unitären Matrizen gewinnt, in dem man sie mit einer Zahl vom Betrag 1 multipliziert.)

Wir werden uns zunächst einmal besonders für die Gruppen  $SL(2; \mathbb{C})$  und  $SU(2, \mathbb{C})$  interessieren. Wir bemerken

$$A \in SU(2) \iff A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 + |b|^2 = 1.$$

Ausserdem werden wir es zu tun haben mit der Gruppe  $SU(1, 1)$  aller derjenigen komplexen  $2 \times 2$ -Matrizen  $A$  mit Determinante  $= 1$ , für die gilt

$$A^* \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{oder äquivalent dazu} \quad A = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 - |b|^2 = 1.$$

Wir erinnern daran, dass es zu jeder Möbiustransformation  $\varphi(\cdot)$  genau zwei darstellende Matrizen  $\pm A \in SL(2; \mathbb{C})$  gibt; und wir haben eine Isomorphie  $\mathfrak{M} \cong SL(2; \mathbb{C})/\{\pm E_2\}$ . Wir werden im Laufe der Vorlesung zwei Untergruppen genauer untersuchen, nämlich

$$SU(2; \mathbb{C})/\{\pm E_2\} \quad \text{und} \quad SU(1, 1)/\{\pm E_2\}.$$

Die erste spielt eine Rolle in der sog. elliptischen nichteuklidischen Geometrie, und übrigens auch in der Geometrie des dreidimensionalen euklidischen Raums. Die zweite spielt eine Rolle in der sog. hyperbolischen nichteuklidischen Geometrie.

### Hinweis auf die reellprojektive Ebene.

Einen zweidimensionalen reellprojektiven Raum  $\Omega = P_2(\mathbb{R})$  gewinnt man üblicherweise dadurch, dass man die Punkte  $p \in \Omega$  als Geraden durch den Nullpunkt eines dreidimensionalen reellen Vektorraums  $V$  deutet. Es ist manchmal praktisch,  $V$  mit dem euklidischen  $\mathbb{R}^3$  zu identifizieren. Die ‘Richtungen’ im  $\mathbb{R}^3$  kann man durch die Punkte der Einheitssphäre veranschaulichen; die Punkte  $p \in \Omega$  entsprechen dann den Äquivalenzklassen, wenn man Antipoden identifiziert.

Die ‘Geraden’ in  $\Omega$  sind zunächst durch die zweidimensionalen Teilvektorräume des  $\mathbb{R}^3$  gegeben; wir wollen sie lieber durch die Grosskreise mit identifizierten Antipoden beschreiben. Zu zwei Punkten  $p, q \in \Omega$  gibt es genau eine ‘Verbindungsgerade’. Zwei ‘Geraden’ schneiden sich immer in einem Antipodenpaar, also einem Punkt  $p \in \Omega$ . Es bezeichne nun  $\Gamma$  die Menge der ‘Geraden’. Das Paar  $(\Omega, \Gamma)$  hat viele Eigenschaften, die analog sind zu den Punkten und Geraden in der synthetischen euklidischen Geometrie – mit der Ausnahme des Parallelenaxioms.

Man kann  $(\Omega, \Gamma)$  als Grundgerüst einer nichteuklidischen Geometrie verstehen. (Man nennt sie die klassische elliptische Geometrie.) Was genauer zu studieren ist, ist die Gruppe der Bewegungen, welche die Kongruenz in unserer elliptischen Geometrie  $(\Omega, \Gamma)$  definiert. Da bietet sich die Gruppe  $\mathfrak{G}$  der Drehungen an, wobei Drehungen der Einheitssphäre hier als Bijektionen von  $\Omega$  verstanden werden. Diese Gruppe von Bijektionen wirkt transitiv auf  $\Omega$ . Die Fixgruppen der Punkte sind isomorph zu Kreislinien. Man kann von Winkeln sprechen; die Bewegungen sind winkelerhaltend. Interessant ist der Zusammenhang zwischen der Winkelsumme in einem Dreieck und der Fläche des Dreiecks. Für die Definition eines bewegungsinvarianten Abstands zweier Punkte bieten sich mehrere Möglichkeiten an, die wir hier aber nicht verfolgen wollen. – Der Kalkül der Längen Winkel und Flächen (von Dreiecken) trifft sich in wesentlichen Zügen mit der sog. sphärischen Trigonometrie.

Wir werden uns im Folgenden in mehreren Ansätzen mit der Gruppe der Bewegungen der nichteuklidischen Ebene auseinandersetzen. Zunächst einmal stützen wir uns auf den Kalkül der komplexen Zahlen. (Später werden wir auch hermitesche  $2 \times 2$ -Matrizen ins Spiel bringen.)

### Der Effekt der stereographischen Projektion.

Vermöge der stereographischen Projektion können wir Bijektionen der 2-Sphäre durch Bijektionen der erweiterten komplexen Ebene  $\bar{\mathbb{C}}$  beschreiben und umgekehrt. Wir haben oben zwei Zahlen  $z_1, z_2 \in \bar{\mathbb{C}}$  zueinander diametral genannt, wenn  $z_1 \cdot \bar{z}_2 = -1$  gilt; und wir haben gesehen, dass die Diametralität zum Ausdruck bringt, dass die stereographischen Bilder Antipoden sind. Bijektionen von  $\bar{\mathbb{C}}$ , welche die Diametralität respektieren, (Diametralpunkte gehen in Diametralpunkte) können als Bijektionen von  $\Omega$  verstanden werden. Hier können nun (es mag als ein Wunder erscheinen!) die Möbiustransformationen ins Spiel gebracht werden. Die Gruppe derjenigen Möbiustransformationen, die die Diametralität respektieren, ist nämlich recht reichhaltig: zu jedem Paar  $z, w$  gibt es eine Möbiustransformation dieser Art, die  $z$  in  $w$  überführt.

**Satz.** Die Möbiustransformation zu einer beliebigen unitären Matrix  $U$  (mit Determinante = 1) bildet Diametralpunkte in Diametralpunkte ab. Diese Möbiustransformationen können daher als Bijektionen von  $\Omega$  interpretiert werden.

Kreise in  $\bar{\mathbb{C}}$ , welche ein Paar von Diametralpunkten enthalten, werden durch die stereographische Projektion zu Großkreisen; und können deswegen als Elemente von  $\Gamma$ , also als ‘Gerade’ in  $\Omega$  interpretiert werden.

**Beweis.** Seien  $z_1$  und  $z_2$  zueinander diametral; d.h.  $\frac{z_1 \bar{z}_2 + 1}{\bar{z}_1 z_2 + 1} = 0$ .  
 Durch eine direkte Rechnung bestätigen wir  $\varphi_U(z_1) \cdot \varphi_U(z_2) + 1 = 0$ :

$$\begin{aligned} & \frac{az_1 + b}{-\bar{b}z_1 + \bar{a}} \cdot \frac{\bar{a}\bar{z}_2 + \bar{b}}{-bz_2 + a} + 1 = \\ &= \frac{1}{\text{Nenner}} [a\bar{a}z_1\bar{z}_2 + b\bar{a}\bar{z}_2 + a\bar{b}z_1 + |b|^2 + |b|^2z_1\bar{z}_2 - \bar{b}az_1 - \bar{a}b\bar{z}_2 + |a|^2] \\ &= \frac{1}{\text{Nenner}} [(|a|^2 + |b|^2)(z_1 \cdot \bar{z}_2 + 1)] = 0 \end{aligned}$$

**Satz.** Jede Möbiustransformation zu einem unitären  $U$ , welche nicht die Identität ist, besitzt genau zwei Fixpunkte; und diese Fixpunkte sind zueinander diametral.

**Beweis.** Wir suchen die Fixpunkte zunächst in speziellen Fällen:

Für  $|a| = 1$ , also  $b = 0$  haben wir die Fixpunkte  $0$  und  $\infty$ .

Für  $a \neq 1$  reell und  $b = |b| \cdot e^{i\phi}$  haben wir Fixpunkte auf dem Einheitskreis, nämlich  $z_{1,2} = \pm i \cdot e^{i\phi}$ . Im Fall  $b \neq 0$  gilt in der Tat allgemein :  $z$  ist Fixpunkt genau dann, wenn  $\bar{b} \cdot z$  die folgende quadratische Gleichung löst,

$$z = \varphi_U(z) = \frac{az + b}{-\bar{b}z + \bar{a}} \Leftrightarrow -\bar{b}z^2 + \bar{a}z = az + b \Leftrightarrow (\bar{b}z)^2 + (a - \bar{a})\bar{b}z + |b|^2 = 0$$

Da die Abbildung die Diametralität respektiert, ist klar, dass die beiden Lösungen  $z_1, z_2$  der quadratischen Gleichung zueinander diametral sind. Man kann das aber natürlich auch durch eine direkte Rechnung bestätigen.

**Hinweis:**

Es stellt sich heraus, dass die  $U$  als Drehungen der Zahlenkugel wirken, ebenso wie die speziellen orthogonalen Matrizen auf die 3-Spalten wirken.  $SU(2)/\{\pm E_2\} \cong SO(3, \mathbb{R})$ . Die Codierung der eindimensionalen Teilvektorräume durch die Paare von Diametralpunkte in  $\bar{\mathbb{C}}$  und die Berechnung der Fixpunkte der Bewegungen aus den Einträgen von  $U$  (als Lösungen quadratischen Gleichungen) sind aber reichlich unübersichtlich. Das ist einer der Gründe dafür, dass wir später (ohne Bezugnahme auf die Möbiustransformationen) nochmals eine effektive Wirkung der Gruppe  $SU(2)/\{\pm E_2\}$  auf ein Modell der elliptischen Ebene konstruieren werden. (Ein zweiter Grund ist übrigens der, dass wir damit die sog. Spinordarstellung der Lorentzgruppe vorbereiten.)

### Die Gruppe der hyperbolischen Bewegungen.

In unserem ersten Ansatz zur sog. hyperbolischen Geometrie wollen wir die Punkte der Geometrie durch Punkte in der komplexen Zahlenebene modellieren und die Geraden als spezielle Kreise. Die Grundlage für die Theorie der Bewegungen liefert der

#### Satz.

Für jedes  $A \in SU(1,1)$  bildet die Möbiustransformation  $\varphi_A(\cdot)$  das Innere des Einheitskreises auf das Innere des Einheitskreises ab.

Jeder Kreis, der den Einheitskreis senkrecht schneidet, wird durch  $\varphi_A(\cdot)$  auf einen ebensolchen Kreis abgebildet.

**Beweis.** Es seien  $a, b$  komplexe Zahlen mit  $|a|^2 - |b|^2 = 1$  und  $\psi(z) = \frac{az+b}{bz+\bar{a}}$ . Es gilt

$$\begin{aligned} 1 - \psi(z)\psi(\bar{z}) &= 1 - \frac{az+b}{bz+\bar{a}} \cdot \frac{\bar{a}\bar{z}+\bar{b}}{b\bar{z}+a} = \\ &= \frac{1}{\text{Nenner}} \cdot [(\bar{b}z + \bar{a})(b\bar{z} + a) - (az + b)(\bar{a}\bar{z} + \bar{b})] = \\ &= \frac{1}{\text{Nenner}} \cdot [(|a|^2 - |b|^2)(1 - z\bar{z})] \quad \text{mit Nenner} > 0. \end{aligned}$$

Man kann auch einfacher argumentieren: der Einheitskreis wird auf sich abgebildet und der Punkt  $\infty$  geht in einen Punkt mit Betrag  $> 1$

$$|\psi(e^{it})| = \left| \frac{ae^{it} + b}{\bar{b}e^{it} + \bar{a}} \right| = \left| \frac{ae^{it} + b}{\bar{a}e^{-it} + \bar{b}} \right| = 1; \quad \psi(\infty) = \frac{a}{b}.$$

Eine Kreislinie bleibt genau dann bei der Spiegelung am Einheitskreis ( $z \mapsto w = \frac{1}{\bar{z}}$ ) unverändert, welche sie den Einheitskreis senkrecht schneidet. Daraus ergibt sich eine Die Gleichung der Kreislinie  $K$  sei

$$a \cdot z\bar{z} + \bar{a}z + \alpha\bar{z} + b = 0, \quad \text{mit } a, b \in \mathbb{R}, \alpha \in \mathbb{C}, D := |\alpha|^2 - ab > 0.$$

Durch die Spiegelung am Einheitskreis kommen wir zur Lösungsmenge der Gleichung

$$b \cdot w\bar{w} + \bar{a}w + \alpha\bar{w} + a = 0.$$

Die Geraden durch den Nullpunkt ( $a = 0 = b$ ) bleiben unverändert. Auch für die übrigen Kreise ist  $a = b$  notwendig und hinreichend für die Invarianz. Möbiustransformationen sind winkelerhaltend. Die auf dem Einheitskreis senkrechten Kreislinien werden durch jede Möbiustransformation  $\varphi_A(\cdot)$  mit  $A \in SU(1,1)/\{\pm E_2\}$  in eine auf dem Einheitskreis senkrechte Kreislinie abgebildet.

**Modelle der hyperbolischen Geometrie** In den Modellen der hyperbolischen Geometrie wirkt die Gruppe  $SU(1, 1)/\{\pm E_2\}$  auf den Punkten eines Grundraums  $\Omega$  und auf den ‘Geraden’ des Modells, wobei die Geraden spezielle Teilmengen von  $\Omega$  sind. Wie die Wirkung in den einzelnen Modellen festgelegt ist, werden wir sehen; zunächst beschreiben wir nur  $\Omega$  und die Menge  $\Gamma$  der Geraden.

- Im Modell von H. Poincaré (1854 –1912) sind die Punkte  $\omega$  die inneren Punkte des Einheitskreises  $\in \mathbb{C}$ ; die Geraden sind die auf dem Einheitskreis senkrechten Kreisbögen.
- Im Modell von F. Klein (1849– 1925) entsprechen ebenfalls die Punkte im Einheitskreis den Punkten der hyperbolischen Ebene. Die Geraden sind hier aber die Sehnen durch den Einheitskreis. (Bei Coxeter heisst das Klein’sche Modell das Modell von Beltrami (1835 – 1900).)
- In einem weiteren sehr beliebten Modell sind die Punkte die Punkte der oberen Halbebene, und die Geraden die Halbkreise mit dem Mittelpunkt auf der reellen Achse (einschliesslich der auf der reellen Achse senkrechten Geraden).

Wir sollten bemerken, dass alle diese Modelle in einer Beziehung täuschen. In den Modellen von Poincaré und Klein scheint (gemäß euklidischer Anschaulichkeit) der Mittelpunkt ausgezeichnet zu sein; im dritten Modell scheinen die senkrechten Geraden eine besondere Rolle zu spielen. In der hyperbolischen Geometrie sind alle Punkte und alle Geraden gleichberechtigt. In einer reifen Anschauung zur (streng axiomatisch aufgebauten) hyperbolischen Geometrie gibt es keine ausgezeichneten Punkte.

**Zusammenhänge** Wenn man das Modell von Poincaré mit der stereographischen Projektion auf die untere Halbkugel abbildet, dann sind die Geraden die auf dem Äquator senkrechten Halbkreise. Wenn man jetzt senkrecht auf die Äquatorebene projiziert, dann erhält man das Klein’sche Modell.

Wenn man den Einheitskreis mit einer Möbiustransformation auf die obere Halbene abbildet (etwa mit der inversen Cayley-Transformation), dann wird aus dem Modell von Poincaré das dritte Modell.

Die technischen und die die Anschauung betreffenden Vor- und Nachteile der verschiedenen Modelle zeigen sich erst beim eingehenden Studium ausgewählter Phänomene. Im Klein’schen Modell sieht man besonders leicht, dass es zu jedem Punktepaar genau eine Verbindungsgerade gibt. Man darf sich vorstellen, dass es zu jeder hyperbolischen Geraden zwei ‘unendlichferne’ Punkte gibt, die man in den Modellen von Poincaré und Klein durch die Punkte auf der Peripherie des Einheitskreises visualisieren kann. Zu einem Punkt  $p$  ausserhalb deiner Geraden gibt es eine ‘rechtseitige Parallele’ und eine ‘linksseitige Parallele’; dabei versteht man hier unter einer ‘Parallelen’ eine Gerade, welche die gegebene Gerade in einem ihrer unendlichfernen Punkte ‘trifft’.

In der hyperbolischen Geometrie gibt es eine **Winkelmessung**. Die Bewegungen sind winkelerhaltend. Das Modell von Poincaré und das dritte Modell sind winkeltreu; man nennt sie daher konforme Modelle. Die Bewegungen sind spezielle Möbiustransformationen. Bei Poincaré sind es diejenigen, die den Einheitskreis auf sich abbilden, im dritten Modell diejenigen, die die obere Halbebene auf sich abbilden.

Die Bewegungen im Klein'schen Modell sind spezielle projektive Abbildungen der reellprojektiven Ebene, diejenigen nämlich, die den Einheitskreis auf sich abbilden. Man nennt das Klein'sche Modell daher ein projektives Modell.

In der hyperbolischen Geometrie gibt es eine (bis auf einen Skalenfaktor eindeutig bestimmte) bewegungsinvariante **Metrik**. In allen unseren Modellen berechnet sich der Abstand zweier Punkte als der Logarithmus eines Doppelverhältnisses. Wir wollen das hier nicht vertiefen, Es sei aber bemerkt, dass es sich im Modell um ein Doppelverhältnis im Sinne der projektiven Geometrie handelt; auf der Verbindungsgeraden von  $p$  und  $q$  liegen auch zwei unendlichferne Punkte; und zu diesem Quadrupel gibt es Doppelverhältnis. In den beiden konformen Modellen interessiert das Doppelverhältnis von vier komplexen Zahlen, und dieses Doppelverhältnis ist eine reelle Zahl, weil die vier Punkte auf einem Kreis liegen. (Den Zusammenhang zwischen den Doppelverhältnissen in den verschiedenen Modellen erörtern wir in einer Übungsaufgabe.)

## Schlussbemerkungen

1. Wenn man die These mathematisch untermauern will, dass der einzig wesentliche Unterschied zwischen der euklidischen Geometrie und der hyperbolischen Geometrie im Sinne von Lobatchewski und Bolyai im Parallelenaxiom beschlossen sei, dann setzt das voraus, dass man die übrigen Axiome, die in beiden Geometrien gelten sollen, wirklich präzise fasst. Eine Geometrie, die sich nur auf solche Axiome stützt, nennt man absolute Geometrie. Solche absoluten Geometrien, (auf deren Ausformung die Geometer des 20. Jahrhunderts in der Tat viel Mühe verwandt haben), wollen wir hier nicht diskutieren. Lassen wir es bei der Feststellung, dass es neben der euklidischen Geometrie eine weitere Geometrie gibt, die im Großen ebenso viele Symmetrien in sich trägt wie die euklidische, und im Kleinen nicht von ihr zu unterscheiden ist.
2. Coxeter zitiert F. Klein mit den Worten: *Jeder Lehrer muß notwendig etwas von der Nichteuklidischen Geometrie kennen; denn sie gehört nun einmal zu den wenigen Teilen der Mathematik, die zumindest in einzelnen Schlagworten in weiteren Kreisen bekannt geworden ist; nach ihr kann daher jeder Lehrer jeden Moment gefragt werden.*

Das Argument von Klein ist nicht mehr zeitgemäß. Die Aufregung um die Entdeckung der nichteuklidischen Theorie im 19. Jahrhundert kam daher, dass hier die

Vorstellung zerstört wurde, dass die Axiome Euklids einen unveränderlichen Rahmen für unsere Raumvorstellung bilden. Die Geometrie von Lobatchevski hat heute den Status einer mathematischen Spezialität. Man muss heute weiter ausholen, wenn man versuchen will, der gebildeten Öffentlichkeit einen Eindruck davon zu vermitteln, was die Mathematiker unter Geometrie verstehen, und was die Aussagen der Geometrie möglicherweise mit dem physikalischen Raum zu tun haben.

3. Das Modell von Poincaré erlaubt eine merkwürdige Interpretation, welche aufzeigt, welche weitreichenden Möglichkeiten der Verallgemeinerung sich für die Geometrie bieten, wenn man nicht die Geraden an den Anfang stellt, sondern eine (lokale!) Metrik: Stellen wir uns vor, dass die Kreisscheibe von einem Medium erfüllt ist, in welchem die Lichtgeschwindigkeit proportional ist zum Abstand vom Rand. Es stellt sich heraus, dass die Lichtstrahlen, d. h. die Kurven kürzesten optischen Wegs, Kreisbögen sind, die auf der Peripherie senkrecht stehen, geradeso, wie es im Modell von Poincaré sein soll. Diese Interpretation lässt die weitreichenden Symmetrien der hyperbolischen Ebene (zunächst einmal) wie ein Wunder erscheinen.

Wenn man andere Annahmen über die Lichtgeschwindigkeit (auf einer glatten Mannigfaltigkeit) macht, dann bekommt man allgemeinere Kurven kürzesten optischen Wegs. Einfache globale Aussagen über diese Strecken, (ähnlich denen in der euklidischen Geometrie oder in der Geometrie von Lobatchevski) kann man i. Allg. nicht erwarten. Ebensowenig kann man eine große Gruppe globaler Symmetrien erwarten. Die lokal vorgegebenen Metriken sind das Thema der sog. Riemann'schen Geometrie. Insbesondere in der allgemeinen Relativitätstheorie zeigt sich, dass der lokale Ansatz von größtem physikalischen Interesse ist. Wir werden ihn hier nicht verfolgen.

Wir wenden uns im nächsten Kapitel einer linearen Geometrie mit hoher Symmetrie zu, die (nicht zuletzt wegen der vielfältigen in ihr beschlossenen Symmetrien) grundsätzliche Bedeutung für die Quantenphysik besitzt.

## 1.4 Lorentztransformationen. Quaternionen für die Drehungen.

Eine für die Geometrie besonders wichtige Matrizen­gruppe ist die Gruppe der unimodularen komplexen  $2 \times 2$ -Matrizen. (Matrizen mit der Determinante = 1 heissen unimodulare Matrizen.) Wir wollen hier diese Gruppe, die spezielle lineare Gruppe  $SL(2, \mathbb{C})$  in Verbindung bringen zur Gruppe der Isometrien des sog. Minkowskiraums. Die Untergruppe  $SU(2, \mathbb{C})$  wird sich als nützlich erweisen für das eingehende Studium der Gruppe  $SO(3, \mathbb{R})$  der Drehungen des dreidimensionalen euklidischen Raum.

Der Minkowskiraum modelliert den vierdimensionalen reellen Vektorraum zum affinen Raum der ‘Ereignisse’ im Sinne der speziellen Relativitätstheorie. Die Mathematiker definieren: Ein vierdimensionaler reeller Vektorraum  $\mathbb{M}$  wird zu einem Minkowskiraum, indem man eine nichtausgeartete quadratische Form mit der Signatur  $(1, 3)$  auszeichnet. Den Begriff der Signatur einer reellen quadratischen Form  $q(\cdot)$  werden wir später beim Satz 2.4.4 auf Seite 114 allgemein diskutieren. Hier genügt es zunächst, zu wissen, dass es im Minkowskiraum Koordinatensysteme  $(x^0, x^1, x^2, x^3)$  gibt, sodass gilt  $q = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$ . Ein Koordinatensystem dieser Art heisst ein Inertialsystem. Die symmetrische Bilinearform zu dieser quadratischen Form notieren wir  $(\cdot \parallel \cdot)$ . Wenn  $(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$  die zu  $(x^0, x^1, x^2, x^3)$  duale Basis ist, dann gilt

$$\left( \sum_0^3 \mathbf{u}_i x^i \parallel \sum_0^3 \mathbf{u}_i y^i \right) = x^0 y^0 - \sum_1^3 x^i y^i.$$

Insbesondere gilt  $(\mathbf{u}_i, \parallel \mathbf{u}_j) = d_{ij}$ , wo  $D = D_{\mathcal{L}}$  die Diagonalmatrix ist, welche den ersten Diagonaleintrag +1 und sonst Diagonaleinträge -1 hat. Und dies charakterisiert die ‘Orthonormalbasen’ im Minkowskiraum  $(\mathbb{M}, (\cdot \parallel \cdot))$ . (oder in  $(\mathbb{M}, q(\cdot))$ )

**Sprechweise 1.4.1.** Die Vektoren  $\mathbf{v}$  mit  $q(\mathbf{v}) > 0$  heissen die zeitartigen Vektoren. Zwei Ereignisse, die durch einen zeitartigen Vektor verbunden sind, erscheinen in geeigneten Inertialsystemen als Ereignisse am gleichen Ort. Zwei Ereignisse, die durch einen raumartigen Vektor  $\mathbf{w}$  verbunden sind ( $q(\mathbf{w}) < 0$ ) erscheinen in geeigneten Inertialsystemen als Ereignisse, die gleichzeitig stattfinden. Die Vektoren  $\mathbf{u}$  mit  $q(\mathbf{u}) = 0$  heissen die Vektoren auf dem Lichtkegel oder auch die isotropen Vektoren.

Die Menge der zeitartigen Vektoren zerfällt in zwei Zusammenhangskomponenten; jede dieser Zusammenhangskomponenten ist ein konvexer Kegel. Indem man die eine Zusammenhangskomponente den Vorwärtskegel, die andere den Rückwärtskegel nennt, zeichnet man eine Zeitrichtung aus. Auch der Lichtkegel  $\{\mathbf{v} : q(\mathbf{v}) = 0\}$  (ohne den Nullpunkt) zerfällt in zwei Komponenten. Jeder der Teile besteht aus einer Familie von Halbgeraden (‘Strahlen’). Auf einem rückwärts gerichteten Strahl liegen diejenigen Ereignisse, die zum Zeitpunkt in einer bestimmten ‘Richtung’ wahrgenommen werden können; der Ort dieser Ereignisse ist nämlich genau so weit entfernt, wie das Licht braucht, um die Entfernung zu überwinden. Es sollte suggestiv sein wenn man notiert

$$\mathbb{M} \setminus \{0\} = \mathbb{M}_{++} \cup \mathbb{M}_{--} \cup \mathbb{M}_{-+} \cup \mathbb{M}_{0+} \cup \mathbb{M}_{0-}.$$

(Die Menge  $\mathbb{M}_{+-}$  ist die Menge der raumartigen Vektoren. Bei Bedarf wird der Nullpunkt zu der einen oder anderen dieser Vektorenmengen dazugenommen; ansonsten sind die Mengen paarweise disjunkt.)

Die Beobachter in verschiedenen Inertialsystemen messen Zeit und Entfernung verschieden. Sie sind sich aber einig, was ein vom gegebenen Nullpunkt nach rückwärts weisender Strahl auf dem Lichtkegel ist. Die Strahlen entsprechen den Punkten auf dem Himmelsgewölbe ('Firmament'). Das Himmelsgewölbe hat die topologische Struktur einer 2-Sphäre. Die Beobachter in verschiedenen Inertialsystemen (mit demselben Nullpunkt) sehen diese Sphäre verschieden: der Wechsel von einem Inertialsystem zu einem anderen erscheint als eine Transformation der Punkte des Himmelsgewölbes. Wie ein Beobachter mit Hilfe der Raumkoordinaten  $(u, v, w)$  in seinem Inertialsystem die Strahlen parametrisiert, ist eine andere Frage. Er beobachtet die Richtung in seiner 'Horizontalebene', der  $(u, v)$ -Ebene und den Winkel gegen die Horizontalebene, und muss daraus einen Parameterwert gewinnen, mit dem sich leicht rechnen lässt, wenn es gilt, zu einem anderen Inertialsystem überzugehen. Wir werden unten eine Parametrisierung konstruieren, welche die Strahlen mit den Punkten  $z \in \bar{\mathbb{C}}$  identifiziert, und damit (vermöge der stereographischen Projektion) mit den Punkten der Riemannschen Zahlenkugel. Es wird sich zeigen, dass die spezielle Parametrisierung dazu führt, dass der Übergang zu einem anderen Inertialsystem durch eine Möbiustransformation beschrieben wird. Im Spezialfall von Beobachtern, die gegeneinander in Ruhe sind, bewirkt jeder Übergang zu einem anderen Inertialsystem eine Drehung der Riemann'schen Zahlenkugel.

### Lorentzmatrizen

**Definition 1.1.** Die Matrix eines Koordinatenwechsels, welcher die Gestalt der Bilinearform reproduziert, welche also ein Inertialsystem in ein Inertialsystem überführt, heisst eine Lorentzmatrix.

#### Satz 1.4.1.

Die reelle  $4 \times 4$ -Matrix  $M$  ist genau dann eine Lorentzmatrix, wenn  $M^T D_{\mathcal{L}} M = D_{\mathcal{L}}$ .

**Beweis.** Genau dann ist  $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  die Basis zu einem Inertialsystem, wenn gilt

$$(\mathbf{v}_k || \mathbf{v}_l) = \left( \sum \mathbf{u}_i m^i_k || \sum \mathbf{u}_j m^j_l \right) = \sum m^i_k d_{ij} m^j_l = d_{kl}.$$

Die Menge  $\mathcal{M}$  der Lorentzmatrizen ist offenbar eine Gruppe. Die Determinante einer Lorentzmatrix kann die Werte  $+1$  und  $-1$  haben. Wir werden später sehen, dass die Menge der Lorentzmatrizen eine sechsdimensionale Mannigfaltigkeit ist, die in vier Zusammenhangskomponenten zerfällt. Die Zusammenhangskomponente der Einheitsmatrix heisst die Gruppe der orthochronen eigentlichen (oder positiv orientierten) Lorentzmatrizen. Wir bezeichnen sie mit  $\mathcal{M}_+^\uparrow$ . Eine einprägsame Notation für die Zerlegung ist

$$\mathcal{M} = \mathcal{M}_+^\uparrow \cup \mathcal{M}_-^\uparrow \cup \mathcal{M}_+^\downarrow \cup \mathcal{M}_-^\downarrow.$$

$$E \in \mathcal{M}_+^\uparrow, \quad D_{\mathcal{L}} \in \mathcal{M}_-^\uparrow, \quad -E \in \mathcal{M}_+^\downarrow, \quad -D_{\mathcal{L}} \in \mathcal{M}_-^\downarrow.$$

Wir werden sehen, dass die Gruppe  $\mathcal{M}_+^\uparrow$  eng verwandt ist mit der speziellen linearen Gruppe  $SL(2, \mathbb{C})$ . Wir konstruieren in der Tat einen Gruppenisomorphismus

$$\Phi : SL(2, \mathbb{C}) / \{\pm E_2\} \longrightarrow \mathcal{M}_+^\uparrow$$

Dieser macht die (zunächst möglicherweise etwas undurchsichtige) eigentliche Lorentzgruppe transparent und leicht handhabbar. Unsere Konstruktion wird auch die Isomorphie der eigentlichen Lorentzgruppe mit der Gruppe aller Möbiustransformationen durchsichtig machen.

### Lorentz-Matrizen und Lorentz-Transformationen

Wenn man in einem  $n$ -dimensionalen Vektorraum eine Basis auszeichnet, dann entsprechen bekanntlich die nichtsingulären  $n \times n$ -Matrizen in eindeutiger Weise den Automorphismen des Vektorraums. Diejenigen positiv gerichteten Lorentztransformationen, die den Vorwärtskegel (und damit auch den Rückwärtskegel) auf sich abbilden, nennen wir die orthochronen eigentlichen Lorentztransformationen des Minkowskiraums  $(\mathbb{M}, q(\cdot))$ . Wir bezeichnen die Gruppe dieser Isometrien mit  $\mathcal{L}_\mathbb{M}^{\uparrow+}$ . Die Abbildung  $\varphi$  mit  $\varphi(\mathbf{e}_i) = \mathbf{e}'_i = \sum_j \mathbf{e}_j \cdot l^j_i$  ist also genau dann eine (orthochrone eigentliche) Lorentztransformation, wenn  $L$  eine (orthochrone eigentliche) Lorentzmatrix ist.

Die Eigenschaft von  $\varphi : (\mathbb{M}, q(\cdot)) \longrightarrow (\mathbb{M}, q(\cdot))$ , Lorentztransformation zu sein, bedeutet  $q(\varphi(\mathbf{v})) = q(\mathbf{v})$  für alle  $\mathbf{v} \in \mathbb{M}$ , oder ausführlicher:  $(\varphi(\mathbf{v}) \parallel \varphi(\mathbf{w})) = (\mathbf{v} \parallel \mathbf{w})$  für alle  $\mathbf{v}, \mathbf{w} \in \mathbb{M}$ . Die Menge der Lorentztransformationen von  $(\mathbb{M}, q(\cdot))$  bezeichnen wir mit  $\mathcal{L}$  oder ausführlich mit  $\mathcal{L}_\mathbb{M}$ . Sie zerfällt wie die Menge der Lorentzmatrizen in vier Komponenten

$$\mathcal{L} = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow \cup \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\downarrow.$$

Die Orthochronie von  $\varphi$  bedeutet, dass für ein zeitartiges  $\mathbf{v}$  das Bild  $\varphi(\mathbf{v})$  in denselben Kegel zeigt wie  $\mathbf{v}$ . (Für die darstellende Matrix  $L$  bedeutet das  $l^0_0 > 0$ .) Wenn zusätzlich  $\det L > 0$ , dann ist  $\varphi$  eine orthochrone eigentliche Lorentztransformation. Wir schreiben  $\varphi \in \mathcal{L}_+^\uparrow$ . Man überlegt sich leicht, dass die Charakterisierung der Eigenschaft  $\varphi \in \mathcal{L}_+^\uparrow$  nicht von der Basiswahl abhängt. Wir werden gelegentlich eine ON-Basis  $(\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  im Minkowski-Raum  $(\mathbb{M}, q(\cdot))$  wählen. Mit einer solchen Wahl zeichnen wir eine Zeitrichtung aus, weil wir generell fordern, dass  $\mathbf{e}_0$  in den Vorwärtskegel zeigen soll; ausserdem zeichnen wir damit eine Orientierung aus, wenn wir fordern, dass die gewählte ONB positiv gerichtet ist.

Für die Matrizen der Koordinatentransformation  $L \in \mathcal{M}_+^\uparrow$  war die Schreibweise und die Interpretation ganz traditionell festgelegt; es ging da um reelle  $4 \times 4$ -Matrizen  $L$ , die durch Multiplikation von links auf 4-Spalten wirken:  $\tilde{\mathbf{x}} \longmapsto L \cdot \tilde{\mathbf{x}}$ , oder auch als Multiplikation von rechts, wenn es um die (als Zeile aufgelisteten) Basen ging:

$$(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \longmapsto (\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \cdot L.$$

Bei den Elementen der Transformationsgruppe  $\mathcal{L}_+^\uparrow$  sind wir freier in Schreibweise und Interpretation. Endomorphismen  $\chi(\cdot)$  eines Minkowskiraums  $\mathbb{M}$  müssen nicht notwendigerweise durch Matrizen beschrieben werden, die durch Multiplikation von links auf Spalten wirken; denn  $\mathbb{M}$  muss nicht notwendigerweise als der Raum der reellen 4-Spalten präsentiert werden.

### Der Raum der hermiteschen $2 \times 2$ -Matrizen als Minkowskiraum

Die Menge  $\mathbb{H}$  der hermiteschen  $2 \times 2$ -Matrizen ist ein vierdimensionaler reeller Vektorraum. Seine Elemente  $H$  können wir beispielsweise in der folgenden Form präsentieren

$$H = a_0 \cdot \sigma_0 + a_1 \cdot \sigma_x + a_2 \cdot \sigma_y + a_3 \cdot \sigma_3 = a_0 \cdot \sigma_0 + \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix},$$

wo  $\sigma_0 = e = E$  die Einheitsmatrix ist, und  $\sigma_1 = \sigma_x$ ,  $\sigma_2 = \sigma_y$ ,  $\sigma_3 = \sigma_z$  die bekannten Pauli-Spinmatrizen.

Den Raum  $\mathbb{H}$  machen wir zu einem Minkowskiraum durch die quadratische Form

$$\begin{aligned} q(H) &= a_0^2 - a_1^2 - a_2^2 - a_3^2 = (a_0 + a_3)(a_0 - a_3) - (a_1 - ia_2)(a_1 + ia_2) \\ &= \det H = \frac{1}{2}(\text{trace}H)^2 - \frac{1}{2}(\text{trace}H^2). \end{aligned}$$

Zum Nachweis der letzten Gleichung bemerken wir  $\text{trace} H = 2a_0$  und

$$\begin{aligned} \sigma_x^2 &= \sigma_y^2 = \sigma_z^2 = \sigma_0; & \sigma_i \sigma_j + \sigma_j \sigma_i &= 0 \quad \text{für } i \neq j, \quad i, j \geq 1 \\ (a_1 \cdot \sigma_1 + a_2 \cdot \sigma_2 + a_3 \cdot \sigma_3) \cdot (a_1 \cdot \sigma_1 + a_2 \cdot \sigma_2 + a_3 \cdot \sigma_3) &= (a_1^2 + a_2^2 + a_3^2) \sigma_0, \\ \frac{1}{2}(\text{trace}H^2) &= a_0^2 + a_1^2 + a_2^2 + a_3^2, \\ \frac{1}{2}(\text{trace}H)^2 - \frac{1}{2}(\text{trace}H^2) &= \frac{1}{2}(2a_0)^2 - (a_0^2 + a_1^2 + a_2^2 + a_3^2). \end{aligned}$$

Aus der letzten Gleichung ergibt sich auch sofort eine kurze Formel für die symmetrische Bilinearform, nämlich

$$(K \parallel H) = \frac{1}{2}(\text{trace}K)(\text{trace}H) - \frac{1}{2}\text{trace}(K \cdot H)$$

Die positivdefiniten  $H$  betrachten wir als die Vektoren im Vorwärtskegel; die negativdefiniten sind die Vektoren des Rückwärtskegels. Das Quadrupel  $(\sigma_0, \sigma_1, \sigma_2, \sigma_3)$  betrachten wir als eine orthochrone positiv gerichtete Orthonormalbasis von  $(\mathbb{H}, \det(\cdot))$ . Das passt, weil die Matrix mit den Einträgen  $(\sigma_i \parallel \sigma_j)$  in der Tat die Lorentzmatrix  $D_{\mathcal{L}}$  ist. Die Pauli-Matrizen haben die Spur 0, und wegen  $\sigma_i \sigma_j + \sigma_j \sigma_i = 0$  auch ihre Produkte. Die Quadrate  $\sigma_j^2$  sind die Einheitsmatrix sind, und haben somit die Spur = 2.

Zu jeder unimodularen Matrix  $A \in SL(2, \mathbb{C})$  assoziieren wir einen Vektorraumautomorphismus des speziellen Minkowskiraums  $(\mathbb{H}, \det(\cdot))$ :

$$\varphi_A : \mathbb{H} \longrightarrow \mathbb{H}; \quad \varphi_A(H) = AHA^*.$$

Die  $\varphi_A$  erhalten die Minkowski-Form; denn wegen  $\det A = 1$  gilt für alle  $H$

$$q(\varphi_A(H)) = \det AHA^* = \det H = q(H).$$

Jede Abbildung  $\varphi_A$  bildet den Vorwärtskegel auf den Vorwärtskegel ab. Die Menge  $SL(2, \mathbb{C})$  ist zusammenhängend; daher haben alle Automorphismen  $\varphi_A$  dieselbe Orientierung, und zwar (wie die Identitätsabbildung) positive Orientierung.

**Satz 1.4.2.** Die Zuordnung  $A \mapsto \varphi_A(\cdot) = A(\cdot)A^*$  liefert einen injektiven Gruppenhomomorphismus der speziellen linearen Gruppe in die Gruppe  $\mathcal{L}_+^\uparrow$  der orthochronen positiv-orientierten linearen Isometrien des speziellen Minkowskiraums  $(\mathbb{H}, \det(\cdot))$ :

$$\Phi : SL(2, \mathbb{C}) \ni A \mapsto \varphi_A \in \mathcal{L}_+^\uparrow.$$

**Beweis 1.4.1.** Die Einheitsmatrix wird auf die Identitätsabbildung  $id_{\mathbb{H}}$  abgebildet, und es gilt

$$\varphi_B(\varphi_A(H)) = BAHA^*B^* = (BA)H(BA)^* = \varphi_{AB}(H).$$

Wir bestimmen den Kern von  $\Phi$ , d.h. die Menge derjenigen Matrizen  $A$ , für welche gilt  $\varphi_A = id_{\mathbb{H}}$ . Wenn  $\varphi_A(\sigma_0) = \sigma_0$ , dann heisst das, dass  $A$  unitär ist:  $AA^* = E$ ,  $A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$  mit  $|a|^2 + |b|^2 = 1$ . Aus  $\varphi_A(\sigma_z) = \sigma_z$  ergibt sich  $|a|^2 = 1$ ,  $b = 0$ . Der Kern ist also  $\{\pm E\}$ .

Wenn wir zeigen, dass  $\Phi$  surjektiv ist, dann haben wir gezeigt, dass  $\Phi$  der gewünschte Isomorphismus ist.— Wir stellen diese Aufgabe noch etwas zurück.

**Die Wirkung der  $\varphi_A$  auf den Lichtkegel des Raums  $(\mathbb{H}, \det(\cdot))$ .**

Für jedes reelle  $d$  gilt  $\varphi_A(d \cdot H) = d \cdot \varphi_A(H)$ . Wir wollen im Folgenden zwei hermitesche Matrizen als äquivalent betrachten, wenn sie sich nur um einen reellen Faktor  $d \neq 0$  unterscheiden; und wir wollen die  $\varphi_A$  als Bijektionen des Raums der Äquivalenzklassen verstehen.

**Satz 1.4.3.**

Es sei  $H$  eine hermitesche Matrix vom Rang 1, die nicht von der Form  $d \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  ist.

Dann existieren  $d \in \mathbb{R} \setminus \{0\}$  und  $z \in \bar{\mathbb{C}}$ , sodass

$$H = d \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} \cdot (\bar{z}, 1) = d \cdot \begin{pmatrix} |z|^2 & z \\ \bar{z} & 1 \end{pmatrix}$$

Wenn  $A$  irgendeine nichtsinguläre Matrix ist, dann existiert  $e \in \mathbb{R} \setminus \{0\}$ , sodass

$$AHA^* = e \cdot \begin{pmatrix} w \\ 1 \end{pmatrix} \cdot (\bar{w}, 1) \quad \text{mit} \quad w = \mu_A(z) = \frac{a_{11}z + a_{12}}{a_{21}z + a_{22}}.$$

**Beweis.** Nehmen wir an, dass  $H$  in der rechten unteren Ecke den Eintrag  $d \neq 0$  hat. Diesen Faktor ziehen wir vor, sodass jetzt in der rechten unteren Ecke die 1 steht. In den Positionen ausserhalb der Diagonale stehen  $z$  und  $\bar{z}$ . Aus  $\det H = 0$  ergibt sich der Eintrag  $|z|^2$  in der linken oberen Ecke.

Wir können den Ausnahmefall integrieren, indem wir zu Äquivalenzklassen übergehen und beachten

$$H \sim \begin{pmatrix} z \\ 1 \end{pmatrix} \cdot (\bar{z}, 1) \sim \begin{pmatrix} 1 \\ \frac{1}{z} \end{pmatrix} \cdot (1, \frac{1}{\bar{z}}).$$

Der Ausnahmefall entspricht dann dem unendlichfernen Punkt  $z = \infty \in \bar{\mathbb{C}}$ .

Wir setzen  $\mu(\infty) = \frac{a_{11}}{a_{21}}$  und  $\mu_A(\tilde{z}) = \infty$ , wenn  $a_{21}\tilde{z} + a_{22} = 0$ .

So haben wir eine Bijektion  $\mu_A : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ . Und es gilt  $\mu_B(\mu_A(\cdot)) = \mu_{B \cdot A}(\cdot)$ .

Die  $\mu_A$  sind die Möbiustransformationen, mit denen wir uns früher ausführlich beschäftigt haben. Wie wir hier sehen, treten sie auf, wenn man die Automorphismen  $\varphi_A$  auf die Menge  $\Omega$  der Äquivalenzklassen der hermiteschen Matrizen vom Rang 1 ‘einschränkt’. — Wir werden übrigens  $\Omega$  später als die Menge der ‘unendlichfernen’ Punkte des hyperbolischen Raums deuten.

Es ist klar, und wir haben das auch schon früher diskutiert, dass die Abbildung  $\mu_A$  eine Matrix  $A \in SL(2, \mathbb{C})$  bis auf den Faktor  $\pm 1$  festlegt.

Wir haben hier eine Bijektion  $\Omega \longleftrightarrow \bar{\mathbb{C}}$  konstruiert. Die  $\omega$  entsprechen den Strahlen im Lichtkegel. Und das bedeutet in Verbindung mit der (noch nicht bewiesenen) Surjektivität von  $\Phi$ : Die Wirkung der eigentlichen Lorentztransformationen ist die Wirkung der Möbiustransformationen und die Gruppe der eigentlichen Lorentztransformationen ist durch ihre Wirkung auf der Menge  $\Omega$  der Lichtstrahlen eindeutig bestimmt ist.

Man bedenke: Bei der üblichen Weise, Gerade  $\omega$  im Raum der 4-Spalten (d. h. Koordinaten in einem Inertialsystem) zu beschreiben, ist die Wirkung der durch  $4 \times 4$ -Matrizen präsentierten Lorentztransformationen weniger transparent: Die Standardbasis in  $\mathbb{H}$  verbindet nämlich die Zahl  $z = u + iv$  mit einer 4-Spalte, wie folgt

$$\begin{pmatrix} z \\ 1 \end{pmatrix} (\bar{z}, 1) = \begin{pmatrix} |z|^2 & z \\ \bar{z} & 1 \end{pmatrix} = \begin{pmatrix} u^2 + v^2 & u + iv \\ u - iv & 1 \end{pmatrix} \longleftrightarrow \frac{1}{2} \begin{pmatrix} u^2 + v^2 + 1 \\ 2u \\ -2v \\ u^2 + v^2 - 1 \end{pmatrix}.$$

Zur Befestigung der eben gewonnenen Einsichten wollen wir die Eigenschaften des speziellen Minkowskiraums  $(\mathbb{H}, (\cdot \| \cdot))$  noch einmal von Grund auf entwickeln.

**Sprechweise 1.4.2** (Standard-Minkowski-Raum).

Den vierdimensionalen reellen Vektorraum  $\mathbb{H}$  der hermiteschen  $2 \times 2$ -Matrizen mit der Bilinearform  $(H \| K) = \frac{1}{2}(\text{trace}H)(\text{trace}K) - \frac{1}{2}\text{trace}(HK)$  nennen wir den Standard-Minkowski-Raum, wenn wie oben die Menge der positivdefiniten Matrizen als der Vorwärtskegel ausgezeichnet ist. Das Quadrupel  $(E, \sigma_x, \sigma_y, \sigma_z)$  heisst die Standardbasis von  $\mathbb{H}$ . Das Tripel der Pauli-Matrizen  $(\sigma_x, \sigma_y, \sigma_z)$  nennen wir die Standardbasis des Raums  $\mathbb{H}_0$  der spurlosen hermiteschen Matrizen.

Ein Quadrupel  $(H_0, H_1, H_2, H_3)$  heisst eine orthochrone Orthonormalbasis, wenn  $H_0$  positiv definit ist und  $(H_i \parallel H_j)_{ij} = D_{\mathcal{L}}$  ('Lorentzmatrix').

Die 'Standardbasis' ist in der Tat eine orthochrone Orthonormalbasis. Es gilt nämlich  $\text{trace} E = 2$  und daher  $(E \parallel E) = \frac{1}{2} \cdot 2^2 - \frac{1}{2} \cdot 2 = 1$ . Ein  $H$  ist genau dann orthogonal zu  $E$  (d. h.  $(E \parallel H) = 0$ ), wenn  $\text{trace} H = 0$ . Die Pauli-Matrizen sind orthogonal zu  $E$ . Sie haben die 'Norm'  $-1$  (d. h.  $(\sigma_i \parallel \sigma_i) = -1$ ), und sie sind paarweise zueinander orthogonal.

**Satz 1.4.4** (Die Determinante als quadratische Form auf  $\mathbb{H}$ ).

Zu jedem  $H \in \mathbb{H}$  existiert eine unitäre Matrix  $U$ , sodass  $U^* H U$  eine Diagonalmatrix ist. Die Diagonaleinträge sind die Eigenwerte  $\lambda, \mu$ . Es gilt

$$\begin{aligned} \text{trace } H &= \text{trace } U^* H U = \lambda + \mu; & \text{trace } H^2 &= \text{trace } (U^* H U)^2 = \lambda^2 + \mu^2; \\ (H \parallel H) &= \frac{1}{2}(\lambda + \mu)^2 - \frac{1}{2}(\lambda^2 + \mu^2) = \lambda \cdot \mu = \det H. \end{aligned}$$

Für jede Matrix  $A \in SL(2, \mathbb{C})$  gilt

$$(A H A^* \parallel A H A^*) = \det(A H A^*) = \det H = (H \parallel H) \quad \text{für alle } H \in \mathbb{H},$$

und wegen der Bilinearität der Form auch  $(A H A^* \parallel A K A^*) = (H \parallel K)$  für alle  $H, K$ .

**Satz 1.4.5** ('Zeitartige' Vektoren in  $\mathbb{H}$ ).

Ein  $H$  mit  $\det H > 0$  (man spricht von einem zeitartigen Vektor) ist entweder positivdefinit oder negativdefinit. Wenn  $H$  positiv definit ist, dann gibt es genau eine positivdefinite Matrix  $\tilde{A}$  mit  $(\tilde{A})^2 = \tilde{A} \tilde{A}^* = H$ . Ist  $A = \tilde{A} U$  mit  $U$  unitär, dann gilt  $A A^* = H$ . Umgekehrt: Wenn  $A$  eine Matrix ist mit  $A A^* = H$ , dann ist  $\tilde{A}^{-1} A$  unitär.

**Satz 1.4.6** (Orthonormaler Koordinatenwechsel im Standard- Minkowskiraum  $(\mathbb{H}, \det(\cdot))$ ).

Ist  $(H_0, H_1, H_2, H_3)$  eine orthochrone Orthonormalbasis, dann existiert eine Matrix  $A$ , sodass  $(A H_0 A^*, A H_1 A^*, A H_2 A^*, A H_3 A^*)$  entweder die Standardbasis ist oder aber die Basis  $(\sigma_0, \sigma_x, -\sigma_y, \sigma_z)$ . Die transformierende Matrix  $A$  ist bis auf das Vorzeichen eindeutig bestimmt.

(Im erstgenannten Fall nennt man  $(H_0, H_1, H_2, H_3)$  eine positivorientierte orthochrone Orthonormalbasis, im zweiten Fall eine negativ orientierte orthochrone Orthonormalbasis)

**Beweis.** Wählen wir ein  $A_0$  sodass  $A_0 H_0 A_0^* = E$ . Die Matrizen  $A_0 H_i A_0^*, i = 1, 2, 3$ , sind dann spurlose Matrizen, die paarweise zueinander orthogonal sind. Sie haben allesamt die Eigenwerte  $+1$  und  $-1$ . Wählen wir eine unitäre Matrix sodass  $U(A_0 H_3 A_0^*) U^* = \sigma_z$ . Mit  $A_1 = U A_0$  ist dann  $(K_0, K_1, K_2, K_3) = (A_1 H_0 A_1^*, A_1 H_1 A_1^*, A_1 H_2 A_1^*, A_1 H_3 A_1^*)$  eine orthochrone Orthonormalbasis mit  $K_0 = E, K_3 = \sigma_z$ . Die Elemente  $K_1, K_2$  sind

sowohl zu  $E$  als auch zu  $\sigma_z$  senkrecht. Das impliziert, dass ihre Diagonaleinträge verschwinden; denn  $\text{trace}(K\sigma_z)$  ist die Differenz der Diagonaleinträge von  $K$ . Ausserdem gilt  $0 = (K_1 \parallel K_2) = -\frac{1}{2}\text{trace}(K_1K_2)$ .

$$K_1 = \begin{pmatrix} 0 & e^{i\alpha} \\ e^{-i\alpha} & 0 \end{pmatrix} \quad K_2 = \begin{pmatrix} 0 & e^{i\beta} \\ e^{-i\beta} & 0 \end{pmatrix}, \quad 0 = \text{trace} \begin{pmatrix} e^{i(\alpha-\beta)} & 0 \\ 0 & e^{-i(\alpha-\beta)} \end{pmatrix}$$

Aus  $0 = e^{i(\alpha-\beta)} + e^{-i(\alpha-\beta)} = 2\cos(\alpha-\beta)$  gewinnen wir  $\alpha-\beta = \pm\frac{\pi}{2} + k \cdot \pi$ ;  $e^{i\alpha} = \pm i \cdot e^{i\beta}$ . Wir betrachten die orthochrone Orthonormalbasis  $(K'_0, K'_1, K'_2, K'_3)$  mit  $K'_i = A_\alpha K_i A_\alpha^*$  mit der Diagonalmatrix  $A_\alpha$  mit den Diagonaleinträgen  $e^{-i\alpha/2}$  und  $e^{i\alpha/2}$ . Offenbar gilt  $K'_0 = K_0 = E$ ,  $K'_3 = K_3 = \sigma_z$ , sowie im ersten Fall

$$K'_1 = \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \cdot \begin{pmatrix} 0 & e^{i\alpha} \\ e^{-i\alpha} & 0 \end{pmatrix} \cdot \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x.$$

$$K'_2 = \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \cdot e^{i\alpha} \\ i \cdot e^{-i\alpha} & 0 \end{pmatrix} \cdot \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y.$$

Im zweiten Fall erhalten wir die orthochrone Orthonormalbasis  $(E, \sigma_x, -\sigma_y, \sigma_z)$ . In diesem Fall gibt es keine Matrix  $A$  mit  $AA^* = E$ ,  $A\sigma_z A^* = \sigma_z$ , und  $A\sigma_y A^* = \sigma_x$ ,  $A\sigma_x A^* = \sigma_y$ . Die Annahme einer die Positionen vertauschenden Matrix  $A$  führt zum Widerspruch: Wegen  $\sigma_x\sigma_y = i \cdot \sigma_z = -\sigma_y\sigma_x$  haben wir nämlich

$$i \cdot \sigma_z = A(i \cdot \sigma_z)A^* = A\sigma_x\sigma_y A^* = (A\sigma_x A^*)(A\sigma_y A^*) \neq \sigma_y\sigma_x = -i \cdot \sigma_z.$$

Das Ergebnis soll nun noch in einer etwas anderen Sprache formuliert werden. Zuerst erinnern wir an die klassische Beschreibung von Basiswechseln; es geht hier um reelle  $4 \times 4$ -Matrizen:

Zu jedem Wechsel von einer (orthochronen) ONB zu einer (orthochronen) ONB gibt genau eine  $4 \times 4$ -Matrix  $L = L^\varphi$ , sodass  $L^T \cdot D_{\mathcal{L}} \cdot L = D_{\mathcal{L}}$  und

$$(H_0, H_1, H_2, H_3) \xrightarrow{\varphi} (K_0, K_1, K_2, K_3) = (H_0, H_1, H_2, H_3) \cdot L^\varphi; \quad K_j = \sum_i H_i \cdot l^i_j.$$

Es gibt zweierlei Basiswechsel  $\varphi$  von einer orthochronen ONB eines Minkowskiraums  $(\mathbb{M}, (\|\cdot\|))$  zu einer anderen orthochronen ONB, die eigentlichen und die uneigentlichen. Die eigentlichen orthochronen Basiswechsel sind gekennzeichnet durch  $l^0_0 > 0$ ,  $\det L > 0$ .

In dem, was wir bewiesen haben, geht es um komplexe  $2 \times 2$ -Matrizen:

**Satz 1.4.7.** *Zu jedem eigentlichen Basiswechsel im Standard-Minkowski-Raum  $(\mathbb{H}, (\|\cdot\|))$  der hermiteschen  $2 \times 2$ -Matrizen gibt es genau zwei unimodulare Matrizen  $\pm A$ , sodass*

$$(H_0, H_1, H_2, H_3) \xrightarrow{\varphi^A} (K_0, K_1, K_2, K_3), \quad K_i = AH_i A^*.$$

### Unbewegte Beobachter

Wenn zwei orthochrone ONB eines Minkowskiraums  $(\mathbb{M}, (\cdot\|\cdot))$  denselben zeitartigen Vektor enthalten, dann bedeutet das, dass sich die entsprechenden Beobachter einig sind, was gleichzeitige Ereignisse sind. Die gleichzeitigen Ereignisse bilden einen dreidimensionalen euklidischen Raum, in welchem für beide Beobachter die zeitartigen Basisvektoren eine ONB bilden. Der Koordinatenwechsel  $\varphi$  ist (nach der klassischen Auffassung) durch eine Matrix  $M^\varphi \in SO(3, \mathbb{R})$  gegeben. Diese  $3 \times 3$ -Matrix erscheint als rechte untere Teilmatrix in der Lorentzmatrix  $L^\varphi$ .

Im speziellen Raum  $(\mathbb{H}, (\cdot\|\cdot))$  wollen wir annehmen, dass der für beide Beobachter gemeinsame zeitartige Vektor die Einheitsmatrix ist. Und wir wollen die obige Konstruktion gegen die klassische Auffassung vom Koordinatenwechsel in einem dreidimensionalen euklidischen Raum stellen. Da der Vektor  $E$  festbleiben soll,  $AEA^* = E$  ist unsere Koordinatentransformation  $H \mapsto AHA^*$  durch eine unimodulare unitäre Matrix  $A = U$  gegeben.

$$K_i = UH_iU^*, \quad \text{und andererseits} \quad (K_1, K_2, K_3) = (H_1, H_2, H_3) \cdot M^\varphi.$$

Beide Tripel sind ONB im euklidischen Raum  $(\mathbb{H}_0, \sqrt{-\det(\cdot)})$ . Der Übergang von  $U = U^\varphi = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$  (mit  $|a|^2 + |b|^2 = 1$ ) zur reellen  $3 \times 3$ -Matrix  $M = M^\varphi$  heisst in der Literatur die Cayley-Parametrisierung der Gruppe  $SO(3, \mathbb{R})$ .  $\pm U^\varphi \leftrightarrow L^\varphi$ .

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{pmatrix} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} = \sigma_x y^1 + \sigma_y y^2 + \sigma_z y^3,$$

$$\text{und andererseits gilt mit} \quad L = L^U \quad \begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = L \cdot \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

Die Sachlage ist besonders übersichtlich, wenn sich die Beobachter auch noch über den dritten Basisvektor (die 'z-Achse') einig sind. Der Basiswechsel entspricht dann einer Drehung um eben diese 'z-Achse'. Für eine unitäre Matrix bedeutet nämlich  $U\sigma_zU^* = \sigma_z$ , oder  $\sigma_zU\sigma_z = U$ , dass  $U$  eine Diagonalmatrix ist:  $U = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}$  Für diese Diagonalmatrix erhalten wir mit

$$U\sigma_xU^* = \sigma_x \cos \alpha - \sigma_y \sin \alpha, \quad U\sigma_yU^* = \sigma_x \sin \alpha + \sigma_y \cos \alpha$$

$$U\left(\sum \sigma_i x^i\right)U^* = \sum \sigma_i y^i \quad \text{mit} \quad \begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Unsere Matrix  $U$  liefert also die Drehung um den Winkel  $\alpha$ .

#### Satz 1.4.8.

Zu jedem positiv orientierten Basiswechsel im euklidischen Raum  $(\mathbb{H}_0, \sqrt{-\det(\cdot)})$  der spurlosen hermiteschen  $2 \times 2$ -Matrizen gibt es genau zwei unitäre Matrizen  $\pm U$ , sodass

$$(h_1, h_2, h_3) \xrightarrow{\varphi_U} (k_1, k_2, k_3), \quad k_i = Uh_iU^*.$$

**Einschub: Quaternionen für den Kalkül der Drehungen**

Eine Lorentztransformation heisst eine Drehung, wenn sie einen zeitartigen Vektor festlässt. Für die für die Drehungen relevanten Matrizen wollen wir jetzt einen bequemen Kalkül entwickeln. Die Grundlage für den Kalkül ist die Erkenntnis, dass die Matrizen

$$\mathbf{j} = \frac{1}{i}\sigma_x, \quad \mathbf{k} = \frac{1}{i}\sigma_y, \quad \mathbf{l} = \frac{1}{i}\sigma_z$$

sowohl schiefhermitisch als auch unitär sind. Es zeigt sich, dass wir sie in einer ökonomischen Weise zusammenbringen kann. Beide Typen von Matrizen passen in den vierdimensionalen reellen Vektorraum der reellen Linearkombinationen

$$\mathbb{G}^\sigma = \left\{ \mathbf{q} = a_0\mathbf{e} + a_1\mathbf{j} + a_2\mathbf{k} + a_3\mathbf{l} = a_0 \cdot \sigma_0 + a_1 \cdot \frac{1}{i}\sigma_x + a_2 \cdot \frac{1}{i}\sigma_y + a_3 \cdot \frac{1}{i}\sigma_z \right\}.$$

Die Elemente sind die Matrizen der Form  $\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$ ; und das sind gerade die reellen Vielfachen der unitären Matrizen. Man schreibt manchmal

$$\mathbb{G}^\sigma = [\sigma_0] \oplus \frac{1}{i} \cdot \mathbb{H}_0 \quad \text{im Gegensatz zu} \quad \mathbb{H} = [\sigma_0] \oplus \mathbb{H}_0.$$

Wir holen weiter aus: Bei  $\mathbb{G}^\sigma$  handelt es sich um ein (im Matrizenkalkül konkretisiertes) Modell des Schiefkörpers der Hamilton'schen Quaternionen. Wie man diesen Schiefkörper in der Algebra abstrakt definiert, werden wir gleich im Anschluss sehen. Zuerst bringen wir Ordnung in die Bezeichnungen. Einselemente treten im Folgenden in verschiedenen Umgebungen auf, und sie sollen möglichst zur Umgebung passend bezeichnet werden:

**Notation.**

- $E = E_2$  oder  $e$  bezeichnet die Gruppeneins in der Gruppe  $SL(2, \mathbb{C})$  oder auch in der Untergruppe  $SU(2, \mathbb{C})$ .  $E$  wird später auch die Eins im Ring aller komplexen  $2 \times 2$ -Matrizen bezeichnen.
- $\sigma_0$  ist der ausgezeichnete zeitartige Vektor im Standard-Minkowski-Raum  $(\mathbb{H}, (\cdot\|\cdot))$ .  $\sigma_0$  ist hier also wirklich die Einheitsmatrix.
- $\mathbf{e}$  bezeichnet das Einselement im Schiefkörper der Quaternionen  $(\mathbb{G}, +, \cdot)$ . Nur für die Standarddarstellung  $\mathbb{G}^\sigma$  könnte man mit gutem Recht identifizieren:  $\mathbf{e} = \sigma_0 = E$ .

**Definition 1.2** (Der Raum der Quaternionen  $(\mathbb{G}, +, \cdot, *, \|\cdot\|)$ ).

Es sei  $\mathbb{G}$  ein vierdimensionaler reeller Vektorraum mit ausgezeichneter Basis  $(\mathbf{e}, \mathbf{j}, \mathbf{k}, \mathbf{l})$ . Es sei eine den Distributivgesetzen genügende Multiplikation definiert, sodass  $\mathbf{e}$  das Einselement ist und für die übrigen Basiselemente gilt

$$\mathbf{j}^2 = \mathbf{k}^2 = \mathbf{l}^2 = -\mathbf{e},$$

$$\mathbf{j} \cdot \mathbf{k} = \mathbf{l} = -\mathbf{k} \cdot \mathbf{j}, \quad \mathbf{k} \cdot \mathbf{l} = \mathbf{j} = -\mathbf{l} \cdot \mathbf{k}, \quad \mathbf{l} \cdot \mathbf{j} = \mathbf{k} = -\mathbf{j} \cdot \mathbf{l}.$$

Für ein Quaternion  $\mathbf{q} = a_0\mathbf{e} + a_1\mathbf{j} + a_2\mathbf{k} + a_3\mathbf{l}$  sei definiert

$$\mathbf{q}^* = a_0\mathbf{e} - a_1\mathbf{j} - a_2\mathbf{k} - a_3\mathbf{l} \quad (\text{das konjugierte Quaternion}) \quad \text{und}$$

$$\|\mathbf{q}\| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2} \quad (\text{die Norm}).$$

Das System heisst dann der Schiefkörper der Quaternionen.

Es zeigt sich, dass das System der Quaternionen alle Eigenschaften eines Körpers mit Ausnahme der Kommutativität der Multiplikation besitzt. Insbesondere besitzt jedes Element  $\neq 0$  eine multiplikative Inverse. Die Konstruktion von  $\mathbb{G}^\sigma$  liefert einen schnellen Beweis für das Assoziativgesetz.

**Satz 1.4.9.** *Der Vektorraumisomorphismus*

$$\sigma : (\mathbb{G}, +, \cdot, *) \longrightarrow (\mathbb{G}^\sigma, +, \cdot, *) \quad a_0\mathbf{e} + a_1\mathbf{j} + a_2\mathbf{k} + a_3\mathbf{l} \longmapsto a_0 \cdot \sigma_0 + a_1 \cdot \frac{1}{i}\sigma_x + a_2 \cdot \frac{1}{i}\sigma_y + a_3 \cdot \frac{1}{i}\sigma_z.$$

hat die Eigenschaften

1. Für das zu  $\mathbf{q}$  konjugierte Quaternion gilt  $\sigma(\mathbf{q}^*) = \sigma(\mathbf{q})^*$  (hermitesche Konjugation),
2. für das Produkt gilt  $\sigma(\mathbf{q} \cdot \mathbf{r}) = \sigma(\mathbf{q}) \cdot \sigma(\mathbf{r})$  (Matrizenmultiplikation).

Dies zeigt, dass die Multiplikation in  $\mathbb{G}$  assoziativ ist und dass die Konjugation die Reihenfolge der Faktoren umkehrt:  $(\mathbf{q} \cdot \mathbf{r})^* = \mathbf{r}^* \cdot \mathbf{q}^*$ . Wir müssen uns lediglich mit den Normen und den Bilinearformen in den verschiedenen Räumen befassen. Die Bilinearform zur Hilbertraumnorm bezeichnen wir mit  $\langle \cdot | \cdot \rangle$ , zur Unterscheidung von der Bilinearform  $(\cdot | \cdot)$  im Minkowskiraum.

### Sprechweisen und einfache Bemerkungen

1. Die  $\mathbf{q} \in \mathbb{G}$  mit  $\|\mathbf{q}\| = 1$  nennen wir die Einheitsquaternionen oder auch die normierten Quaternionen. Für jedes  $\mathbf{q}$  gilt  $\mathbf{q} \cdot \mathbf{q}^* = \|\mathbf{q}\|^2 \cdot \mathbf{e}$ ;  $\frac{\mathbf{q}}{\|\mathbf{q}\|}$  ist also ein normiertes Quaternion. Wir benützen auch den Ausdruck unitäres Quaternion wegen

$$\|\mathbf{q}\| = 1 \iff \sigma(\mathbf{q}) \in SU(2) \iff \sigma(\mathbf{q}) = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad \text{mit } \|a\|^2 + \|b\|^2 = 1.$$

Die Standardbasis  $(\mathbf{e}, \mathbf{j}, \mathbf{k}, \mathbf{l})$  ist eine Orthonormalbasis des vierdimensionalen euklidischen Raums  $(\mathbb{G}, \|\cdot\|)$ .

2. Die  $\mathbf{q}$  mit  $\mathbf{q} = -\mathbf{q}^*$  nennen wir die reinen Quaternionen oder auch die Quaternionen ohne reellen Anteil.  $(\mathbb{G}_0, \|\cdot\|)$  bezeichnet den dreidimensionalen euklidischen Raum der reinen Quaternionen.

$$\mathbf{q} \in \mathbb{G}_0 \iff \frac{1}{i}\sigma(\mathbf{q}) \in \mathbb{H}_0 \quad (\text{hermitisch spurlos}) \iff \det \sigma(\mathbf{q}) = -\|\mathbf{q}\|^2.$$

3. Für  $\mathbf{q}, \mathbf{r} \in \mathbb{G}_0$  gilt  $\|\mathbf{q} \cdot \mathbf{r}\| = \|\mathbf{q}\| \cdot \|\mathbf{r}\|$ , sowie  $\mathbf{q} \perp \mathbf{r} \iff \mathbf{q} \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{q} = 0$ .

In der Tat gilt wegen  $\mathbf{q}^2 = \|\mathbf{q}\|^2 \cdot (-\mathbf{e})$  und  $2 \cdot \langle \mathbf{q} | \mathbf{r} \rangle = \|\mathbf{q} + \mathbf{r}\|^2 - \|\mathbf{q}\|^2 - \|\mathbf{r}\|^2$ .

$$\begin{aligned} 2 \cdot \langle \mathbf{q} | \mathbf{r} \rangle (-\mathbf{e}) &= (\mathbf{q} + \mathbf{r})^2 - \mathbf{q}^2 - \mathbf{r}^2 = \mathbf{q} \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{q} \\ \|\mathbf{q}\| \cdot \|\mathbf{r}\|^2 \cdot \mathbf{e} &= (\mathbf{q} \cdot \mathbf{r})(\mathbf{q} \cdot \mathbf{r})^* = \mathbf{q} \cdot \mathbf{r} \cdot \mathbf{r}^* \cdot \mathbf{q}^* = \|\mathbf{q}\|^2 \|\mathbf{r}\|^2 \cdot \mathbf{e} \end{aligned}$$

4. Jedes  $\mathbf{q} \in \mathbb{G}$  besitzt eine eindeutige Zerlegung  $\mathbf{q} = \tau(\mathbf{q}) \cdot \mathbf{e} + \mathbf{s}$  mit  $\mathbf{s} \in \mathbb{G}_0$ .

Hierbei ist  $\tau(\mathbf{q}) = \frac{1}{2} \text{trace } \sigma(\mathbf{q})$ ,  $\|\mathbf{s}\|^2 = \|\mathbf{q}\|^2 - |\tau(\mathbf{q})|^2$ .  $\tau(\mathbf{q}) \cdot \mathbf{e} = \frac{1}{2}(\mathbf{q} + \mathbf{q}^*)$ .

5. Es seien  $\mathbf{q}, \mathbf{r} \in \mathbb{G}_0$  mit  $\|\mathbf{q}\| = 1 = \|\mathbf{r}\|$ . Es existieren dann in eindeutiger Weise  $\alpha \in [0, \pi)$  und  $\mathbf{s} \in \mathbb{G}_0$  mit  $\|\mathbf{s}\| = 1$ , sodass  $\mathbf{q} \cdot \mathbf{r} = \cos \alpha \cdot (-\mathbf{e}) + \sin \alpha \cdot \mathbf{s}$ .

$$\begin{aligned} (a_1 \mathbf{j} + a_2 \mathbf{k} + a_3 \mathbf{l}) \cdot (b_1 \mathbf{j} + b_2 \mathbf{k} + b_3 \mathbf{l}) &= \\ = (a_1 b_1 + a_2 b_2 + a_3 b_3)(-\mathbf{e}) &+ (a_2 b_3 - a_3 b_2) \cdot \mathbf{j} + (a_3 b_1 - a_1 b_3) \cdot \mathbf{k} + (a_1 b_2 - a_2 b_1) \cdot \mathbf{l}. \end{aligned}$$

Bei einem sorgfältigen Verständnis der Orientierung ist die Zahl  $\alpha$  als der Winkel zwischen  $\mathbf{q}$  und  $\mathbf{r}$  zu interpretieren. Dabei ist für  $\cos \alpha$  die Reihenfolge der Faktoren irrelevant.

$$\mathbf{q} \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{q} = \mathbf{q} \cdot \mathbf{r} + (\mathbf{q} \cdot \mathbf{r})^* = 2 \cdot \cos \alpha (-\mathbf{e}).$$

Bei der Vertauschung der Faktoren kehrt sich das Vorzeichen von  $\sin \alpha \cdot \mathbf{s}$  um.

$$\mathbf{q} \cdot \mathbf{r} - \mathbf{r} \cdot \mathbf{q} = \mathbf{q} \cdot \mathbf{r} - (\mathbf{q} \cdot \mathbf{r})^* = 2 \cdot \sin \alpha \cdot \mathbf{s}.$$

Diese Formel zeigt nun schnell, dass  $\mathbf{s}$  zu  $\mathbf{q}$  und  $\mathbf{r}$  orthogonal ist; denn  $\mathbf{q}^2$  ist ein Vielfaches der Identität und

$$\mathbf{q} \cdot \mathbf{s} + \mathbf{s} \cdot \mathbf{q} = 0 \quad \text{wegen} \quad \mathbf{q} \cdot (\mathbf{q} \cdot \mathbf{r} - \mathbf{r} \cdot \mathbf{q}) + (\mathbf{q} \cdot \mathbf{r} - \mathbf{r} \cdot \mathbf{q}) \cdot \mathbf{q} = 0.$$

Ebenso folgt  $\mathbf{s} \perp \mathbf{r}$ .

Hinweis: In der (veralteten!) Vektoranalysis definierte man das **Vektorprodukt** für Vektoren im orientierten dreidimensionalen ‘Anschauungsraum’. In dieser Tradition hätte man notiert  $\mathbf{q} \times \mathbf{r} = |\sin \alpha| \cdot \mathbf{s}$ , und man hätte gesagt, dass  $(\mathbf{q}, \mathbf{r}, \mathbf{s})$  ein positiv orientiertes Dreibein mit dem Volumen  $|\sin \alpha|$  ist. Weiter hätte man gesagt: „Man gewinnt  $\mathbf{r}$  aus  $\mathbf{q}$ , indem man um die Achse  $\mathbf{s}$  im positiven Sinn um den Winkel  $\alpha$  weiterdreht.“

In dieser Sprechweise hätte man zum Beispiel gesagt: Da  $\mathbf{j} \cdot \mathbf{k} = \mathbf{l}$ ; oder ausführlich  $\mathbf{j} \cdot \mathbf{k} = \mathbf{l} = \sin \alpha \cdot \mathbf{l}$  mit  $\alpha = \frac{\pi}{2}$ . kann man sagen: „man gewinnt  $\mathbf{k}$ , indem man  $\mathbf{j}$  um den Winkel  $90^\circ$  im positiven Sinn um die z-Achse dreht“.

Diese Interpretation der (reinen!) Quaternionen ist aber, wie gesagt, veraltet.

Der bessere Kalkül der Drehungen, den wir hier jetzt entwickeln wollen, ist der in letzter Zeit (vor allem im Zusammenhang mit der sog. Robotik) sehr populär geworden.

**Notation 1.1.** Es sei  $\mathbf{s} \in \mathbb{G}_0$  mit  $\|\mathbf{s}\| = 1$  und  $\alpha \in \mathbb{R}/(0, 2\pi)$ .

Zum normierten Quaternion  $\mathbf{s}_{\alpha/2} = \cos(\alpha/2)\mathbf{e} + \sin(\alpha/2) \cdot \mathbf{s}$  definiert man die Abbildung

$$\chi_{\mathbf{s}_\alpha} : (\mathbb{G}_0, \|\cdot\|) \longrightarrow (\mathbb{G}_0, \|\cdot\|); \quad \mathbf{q} \longmapsto \mathbf{s}_{\alpha/2} \cdot \mathbf{q} \cdot \mathbf{s}_{\alpha/2}^*$$

und man nennt sie die Drehung um den Winkel  $\alpha$  um die gerichtete Achse  $\mathbf{s}$  in Quaternionenschreibweise.

Wenn man die Achse umorientiert,  $\mathbf{s} \mapsto -\mathbf{s}$ , und gleichzeitig das Vorzeichen des Winkels ändert,  $\alpha \mapsto -\alpha$ , dann ergibt sich dieselbe Drehung. Man beachte auch  $\mathbf{s}_\pi = -\mathbf{e}$ ; und  $\chi_{\mathbf{s}_\alpha}$  ist genau dann die Identitätsabbildung, wenn  $\alpha$  ein ganzzahliges Vielfaches von  $2\pi$  ist. Zur Identitätsabbildung gibt es keine wohlbestimmte Achse; die Identitätsabbildung ist der Grenzfall  $\chi_{\mathbf{e}}$ . Jede andere Drehung bestimmt die Achse  $\mathbf{s}$  bis aufs Vorzeichen.

Zu den Interpretationen von  $\mathbf{s} \in \mathbb{G}_0$  als Drehachse und  $\alpha$  als Drehwinkel sagen wir gleich mehr. Zunächst stellen wir fest, dass jedes Einheitsquaternion  $\mathbf{p}$  eine orientierungserhaltende Isometrie  $\chi$  liefert:

$$\chi(\mathbf{q}) = \chi_{\mathbf{p}}(\mathbf{q}) = \mathbf{p} \cdot \mathbf{q} \cdot \mathbf{p}^*.$$

Es gilt nämlich  $\chi(\mathbf{q})^* = \chi(\mathbf{q}^*) = -\chi(\mathbf{q})$  und  $\chi(\mathbf{q} \cdot \mathbf{r}) = \chi(\mathbf{q}) \cdot \chi(\mathbf{r})$ , insbesondere

$$\begin{aligned} \mathbf{j}' = \chi(\mathbf{j}), \quad \mathbf{k}' = \chi(\mathbf{k}), \quad \mathbf{l}' = \chi(\mathbf{l}) &\implies \mathbf{j}' \cdot \mathbf{k}' = \mathbf{l}'. \\ \|\chi(\mathbf{q})\|^2 \cdot \mathbf{e} = \chi(\mathbf{q})\chi(\mathbf{q})^* = \mathbf{p} \cdot \mathbf{q} \cdot \mathbf{p}^* \cdot \mathbf{p} \cdot \mathbf{q}^* \cdot \mathbf{p}^* &= \chi(\mathbf{q} \cdot \mathbf{q}^*) = \|\mathbf{q}\|^2 \cdot \mathbf{e} \end{aligned}$$

Sei  $\mathbf{p} = \cos(\alpha/2) \cdot \mathbf{e} + \sin(\alpha/2) \cdot \mathbf{s}$ .

- Der von  $\mathbf{s}$  aufgespannte Teilraum  $V \subset \mathbb{G}_0$ , bleibt fest unter der Abbildung  $\chi_{\mathbf{p}}(\cdot)$ , weil  $\mathbf{s}$  mit  $\mathbf{p}$  kommutiert.
- Der zweidimensionale Orthogonalraum  $V^\perp$  wird um den Winkel  $\alpha$  gedreht.

Für  $\mathbf{q} \in V^\perp$  gilt nämlich wegen  $\mathbf{q} \cdot \mathbf{s} + \mathbf{s} \cdot \mathbf{q} = 0$

$$\begin{aligned} \mathbf{p} \cdot \mathbf{q} \cdot \mathbf{p}^* &= (\cos(\alpha/2)\mathbf{e} + \sin(\alpha/2)\mathbf{s}) \cdot \mathbf{q} \cdot (\cos(\alpha/2)\mathbf{e} - \sin(\alpha/2)\mathbf{s}) \\ &= (\cos(\alpha/2)\mathbf{e} + \sin(\alpha/2)\mathbf{s}) \cdot (\cos(\alpha/2)\mathbf{e} + \sin(\alpha/2)\mathbf{s}) \cdot \mathbf{q} \\ &= \left( (\cos^2(\alpha/2) - \sin^2(\alpha/2)\mathbf{s}) \cdot \mathbf{e} + 2 \cdot \cos(\alpha/2) \cdot \sin(\alpha/2) \cdot \mathbf{s} \right) \cdot \mathbf{q} \\ &= \cos \alpha \cdot \mathbf{q} + \sin \alpha \cdot \mathbf{r}, \end{aligned}$$

wo  $\mathbf{r} = \mathbf{s} \cdot \mathbf{q}$  der Vektor ist, welcher  $(\mathbf{q}, \mathbf{r}, \mathbf{s})$  zu einer positiv gerichteten ONB macht. Dieselbe Rechnung ergibt für das Bild des Vektors  $\mathbf{r} = \mathbf{s} \cdot \mathbf{q}$  ( $\mathbf{r} \perp \mathbf{s}$ ) wegen  $\mathbf{sr} = \mathbf{ssq} = -\mathbf{q}$

$$\mathbf{p} \cdot \mathbf{r} \cdot \mathbf{p}^* = (\cos \alpha \cdot \mathbf{e} + \sin \alpha \cdot \mathbf{s}) \cdot \mathbf{r} = \cos \alpha \cdot \mathbf{r} - \sin \alpha \cdot \mathbf{q}.$$

Die hier konstruierte Darstellung der Drehungen durch die normierten Quaternionen ist deswegen bei den Anwendern sehr beliebt, weil einerseits das Hintereinanderschalten leicht zu rechnen ist und andererseits Drehachsen und Drehwinkel am Resultat direkt abzulesen sind.

**Beispiel** (Drehungen des Einheitswürfels):

Wir machen zuerst eine  $90^\circ$ -Drehung um die z-Achse ( $\chi_1 = \chi_{\mathbf{p}_1}$  mit  $\mathbf{p}_1 = \frac{1}{\sqrt{2}}\mathbf{e} + \frac{1}{\sqrt{2}}\mathbf{l}$ ) (Es gilt  $\cos 45^\circ = \frac{1}{\sqrt{2}} = \sin 45^\circ$ ). Darauf machen wir eine  $90^\circ$ -Drehung um die y-Achse ( $\chi_2 = \chi_{\mathbf{p}_2}$  mit  $\mathbf{p}_2 = \frac{1}{\sqrt{2}}\mathbf{e} + \frac{1}{\sqrt{2}}\mathbf{k}$ ).

Es entsteht eine Drehung  $\chi(\cdot) = \chi_2(\chi_1(\cdot)) = \chi_{\mathbf{p}_2 \cdot \mathbf{p}_1}(\cdot)$ . Das resultierende Quaternion ist  $\mathbf{p} = \frac{1}{2}(\mathbf{e} + \mathbf{k})(\mathbf{e} + \mathbf{l}) = \frac{1}{2}\mathbf{e} + \frac{\sqrt{3}}{2}\mathbf{s}$  mit  $\mathbf{s} = \frac{1}{\sqrt{3}}(\mathbf{j} + \mathbf{k} + \mathbf{l})$ , und das ist wegen  $\cos 60^\circ = \frac{1}{2}$  eine  $120^\circ$ -Grad-Drehung um die Raumdiagonale im  $(+++)$ -Oktanten.

Schliessen wir noch eine  $180^\circ$ -Drehung um eine Seitendiagonale an. ( $\chi_3 = \chi_{\mathbf{p}_3}$  mit  $\mathbf{p}_3 = \frac{1}{\sqrt{2}}(\mathbf{j} + \mathbf{k})$ ). Es ergibt sich die  $270^\circ$ -Drehung um die x-Achse; denn  $\chi_3 \circ \chi_2 \circ \chi_1(\cdot) = \chi_{\mathbf{r}}(\cdot)$  mit  $\mathbf{r} = \mathbf{p}_3 \cdot \frac{1}{2}(\mathbf{e} + \mathbf{j} + \mathbf{k} + \mathbf{l}) = \frac{1}{2\sqrt{2}}(-2\mathbf{e} + 2\mathbf{j})$ .

Die **Exponentialabbildung** spielt bekanntlich eine wichtige Rolle im Matrizenkalkül. Man kann sie auch bei den Quaternionen ins Spiel bringen: Für ein reines Einheitsquaternion  $\mathbf{s}$  ( $\mathbf{s}^* = -\mathbf{s}$ ,  $\mathbf{s}^2 = -\mathbf{e}$ ) erhalten wir

$$\begin{aligned} \exp(\beta\mathbf{s}) &= \mathbf{e} + \beta\mathbf{s} + \frac{1}{2!}(\beta\mathbf{s})^2 + \frac{1}{3!}(\beta\mathbf{s})^3 + \dots \\ &= \mathbf{e} \cdot \left(1 - \frac{1}{2!}\beta^2 + \frac{1}{4!}\beta^4 - \dots\right) + \mathbf{s} \cdot \left(\beta - \frac{1}{3!}\beta^3 + \frac{1}{5!}\beta^5 - \dots\right) \\ &= \cos \beta \cdot \mathbf{e} + \sin \beta \cdot \mathbf{s}. \end{aligned}$$

Man kann also die Drehungen in der Form  $\mathbf{s}_{\alpha/2}(\cdot)\mathbf{s}_{\alpha/2}^* = \exp(\alpha/2 \cdot \mathbf{s})(\cdot)\exp(\alpha/2 \cdot \mathbf{s}^*)$  schreiben.

In dieser Notation wird deutlich, dass es sich bei festem  $\mathbf{s}$  um die ‘einparametrische Gruppe’ der gleichförmigen Drehungen um die ‘Achse’  $\mathbf{s}$  handelt. Man kann  $\alpha$  als Zeitparameter deuten. Für  $\|\mathbf{s}\| = 1$  erhalten wir die gleichförmigen Drehung um die Achse  $\mathbf{s}$ , mit Winkelgeschwindigkeit = 1 und positiv gerichtet im Bezug auf die Achse. Fixiert bleiben nämlich die Vektoren auf der von  $\mathbf{s}$  erzeugten Achse; und es sind genau die  $\alpha = k \cdot 2\pi$ , für welche die Transformation  $\mathbf{s}_\alpha(\cdot)\mathbf{s}_\alpha^*$  die Identität ergibt.

Die gleichförmigen Drehbewegungen bilden bekanntlich einen dreidimensionalen Vektorraum. Die ‘Generatoren’ oder ‘Winkelgeschwindigkeiten’ oder ‘Dralls’ sind (in dem hier entwickelten Kalkül) die reinen Quaternionen  $\mathbf{s}$ . Linearkombinationen der reinen Quaternionen bedeuten Überlagerungen der gleichförmigen Drehbewegungen.— Wir werden in allgemeinerem Rahmen mehr über die Exponentialabbildung zu sagen haben.

Für die effektive Drehung eines reinen Quaternion  $\mathbf{q} \in \mathbb{G}_0$  in der Zeitspanne  $t$  gilt

$$\begin{aligned} D_t(\mathbf{q}) &= \exp(t/2 \cdot \mathbf{s}) \cdot \mathbf{q} \cdot \exp(-t/2 \cdot \mathbf{s}) = \\ &= \left(\cos(t/2\|\mathbf{s}\|) \cdot \mathbf{e} + \sin(t/2\|\mathbf{s}\|) \cdot \frac{\mathbf{s}}{\|\mathbf{s}\|}\right) \cdot \mathbf{q} \cdot \left(\cos(t/2\|\mathbf{s}\|) \cdot \mathbf{e} - \sin(t/2\|\mathbf{s}\|) \cdot \frac{\mathbf{s}}{\|\mathbf{s}\|}\right) \end{aligned}$$

## 1.5 Der hyperbolische Raum und seine Tangentialräume

Wir gehen hier nochmals von einem Minkowskiraum  $(\mathbb{M}, (\cdot\|\cdot))$  aus; und wir gelangen dann durch die Konstruktion einer Semimetrik zum hyperbolischen Raum  $(\mathcal{Hyp}_{\mathbb{M}}, d(\cdot, \cdot))$ .

Ein Beobachter  $Q$ , der sich in seinem Inertialsystem in Ruhe befindet, definiert im Minkowskiraum eine zeitartige Gerade, nämlich die Gerade, die von dem zeitartigen Einheitsvektor  $\mathbf{u}_0$  in seiner Orthogonalbasis aufgespannt wird. Man sagt manchmal, dass der Beobachter sich entlang dieser Geraden bewegt. (Wir können die Eigenzeit als eine natürliche Parametrisierung dieser Geraden ansehen.) Zu einem zweiten Beobachter  $P$  gehört ein zweiter zeitartiger Einheitsvektor. Wir wollen uns jetzt mit Paaren zeitartiger Vektoren und den von ihnen aufgespannten Ebenen  $L_0$  beschäftigen. Eine solche ‘Ebene’ ist als die ‘Verbindungsgerade’ der Punkte  $P, Q$  im hyperbolischen Raum zu verstehen.

**Satz 1.5.1** (Umgekehrte Schwarz’sche Ungleichung).

Es seien  $\mathbf{u}_0, \mathbf{v}_0$  Vektoren im Vorwärtskegel des Minkowskiraums  $(\mathbb{M}, (\cdot\|\cdot))$ . Es gilt dann

$$(\mathbf{u}_0\|\mathbf{v}_0) \geq \sqrt{(\mathbf{u}_0\|\mathbf{u}_0)}\sqrt{(\mathbf{v}_0\|\mathbf{v}_0)}.$$

Gleichheit gilt nur, wenn die Vektoren zueinander proportional sind.

**Beweis.** Die Ungleichung gilt auch, wenn die Vektoren beide im Rückwärtskegel liegen. Wenn sie in entgegengesetzten Kegeln liegen, muss das Vorzeichen korrigiert werden. Wenn die Vektoren nicht proportional sind, dann spannen sie eine Ebene  $L_0$  auf, welche auch raumartige Vektoren enthält. Die quadratische Funktion

$$(\mathbf{u}_0 + t \cdot \mathbf{v}_0\|\mathbf{u}_0 + t \cdot \mathbf{v}_0) = (\mathbf{u}_0\|\mathbf{u}_0) + 2t \cdot (\mathbf{u}_0\|\mathbf{v}_0) + t^2 \cdot (\mathbf{v}_0\|\mathbf{v}_0) = a + 2tb + t^2c \quad \text{für } t \in \mathbb{R}$$

nimmt beide Vorzeichen an, und das bedeutet  $b^2 - ac > 0$ . Man bemerke: Eine merkwürdige Konsequenz ist die ‘verkehrte Dreiecksungleichung’  $\|\mathbf{u}_0 + \mathbf{v}_0\| \geq \|\mathbf{u}_0\| + \|\mathbf{v}_0\|$ .

**Definition 1.3** (Hyperbolischer Abstand).

Es seien  $\mathbf{u}, \mathbf{v}$  Vektoren im Vorwärtskegel und  $P, Q$  die von ihnen aufgespannten eindimensionalen Vektorräume. Wir schreiben

$$(\mathbf{u}\|\mathbf{v}) = \sqrt{(\mathbf{u}\|\mathbf{u})} \cdot \sqrt{(\mathbf{v}\|\mathbf{v})} \cdot \cosh \omega,$$

und nennen  $d(P, Q) = |\omega|$  den (hyperbolischen) Abstand von  $P$  und  $Q$ .

**Notation 1.2.**

Mit  $\mathbb{M}_{++}$  bezeichnen wir den Vorwärtskegel als konvexe Teilmenge des Minkowskiraums  $\mathbb{M}$ . Mit  $(\mathcal{Hyp}_{\mathbb{M}}, d(\cdot, \cdot))$  bezeichnen wir den metrischen Raum der zeitartigen eindimensionalen Teilvektorräume.

Die Definition findet ihre Berechtigung darin, dass  $d(\cdot, \cdot)$  im Raum  $\mathcal{Hyp}_{\mathbb{M}}$  der zeitartigen eindimensionalen Teilvektorräume tatsächlich die Eigenschaften einer Metrik hat. Die kritische Eigenschaft ist die Dreiecksungleichung, die wir sofort beweisen werden. Es wird bequem sein, mit den Vektoren im Minkowskiraum zu rechnen. Wir definieren

$$\cosh d(\mathbf{u}, \mathbf{v}) = \frac{(\mathbf{u} \parallel \mathbf{v})}{\sqrt{(\mathbf{u} \parallel \mathbf{u})} \cdot \sqrt{(\mathbf{v} \parallel \mathbf{v})}} \quad \text{für Vektoren im Vorwärtskegel } \mathbb{M}_{++}$$

Für andere Paare von Vektoren bleibt die ‘hyperbolische Distanz’ undefiniert.

Wir werden sehen: Die hyperbolische Distanz ist eine Semi-Metrik auf dem Vorwärtskegel  $\mathbb{M}_{++}$ , die zu einer Metrik wird, wenn man zu Äquivalenzklassen im Sinne der projektiven Geometrie übergeht.

### Hyperbolische Dreieckslehre

**Satz 1.5.2** (Dreiecksungleichung). *Es seien  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  Vektoren im Vorwärtskegel und  $P, Q, R$  die aufgespannten Vektorräume. Es gilt dann*

$$d(P, R) \leq d(P, Q) + d(Q, R).$$

*Gleichheit gilt genau dann wenn für nichtnegative  $a, b$  gilt  $\mathbf{v} = a \cdot \mathbf{u} + b \cdot \mathbf{w}$ . (Wir sagen in diesem Fall, dass  $Q$  auf der Verbindungsstrecke von  $P$  und  $R$  liegt.)*

**Beweis 1.5.1.** *Wir können annehmen, dass die Vektoren normiert sind. Wir bezeichnen*

$$\omega_1 = d(Q, P) = d(\mathbf{v}, \mathbf{u}), \quad \omega_2 = d(Q, R) = d(\mathbf{v}, \mathbf{w}).$$

*In dem von  $\mathbf{u}, \mathbf{v}$  aufgespannten Vektorraum  $L_0$  gibt es genau zwei auf  $\mathbf{v}$  ‘orthogonale’ Einheitsvektoren,  $\pm \mathbf{u}_1$ . Es handelt sich um raumartige Einheitsvektoren:  $(\mathbf{u}_1 \parallel \mathbf{u}_1) = -1$ . Wir beweisen  $\mathbf{u} = \cosh \omega_1 \cdot \mathbf{v} \pm \sinh \omega_1 \cdot \mathbf{u}_1$ . Da  $\mathbf{u}$  ein normierter Vektor im Raum  $L_0$  ist, gilt*

$$\begin{aligned} \mathbf{u} &= a \cdot \mathbf{v} \pm b \cdot \mathbf{u}_1 \quad \text{mit} \quad 1 = (\mathbf{u} \parallel \mathbf{u}) = a^2 - b^2, \quad \text{und} \\ -b^2 &= (b \cdot \mathbf{u}_1 \parallel b \cdot \mathbf{u}_1) = (\mathbf{u} - a \cdot \mathbf{v} \parallel \mathbf{u} - a \cdot \mathbf{v}) = 1 + a^2 - 2a \cdot (\mathbf{u} \parallel \mathbf{v}); \quad 2a \cdot (\mathbf{u} \parallel \mathbf{v}) = 2a^2. \end{aligned}$$

*Wir wählen das  $\mathbf{u}_1$  so, dass das  $+$ -Zeichen gilt. In der gleichen Weise konstruieren wir  $\mathbf{w}_1$  in der von  $\mathbf{v}, \mathbf{w}$  aufgespannten Ebene  $L_1$ ; und wir erhalten*

$$\mathbf{w} = \cosh \omega_2 \cdot \mathbf{v} + \sinh \omega_2 \cdot \mathbf{w}_1$$

*Die Vektoren  $\mathbf{u}_1, \mathbf{w}_1$  liegen im Orthogonalraum von  $\mathbf{v}$ . Und der trägt eine euklidische Metrik. Der eingeschlossene Winkel sei  $\alpha$ :  $\langle \mathbf{u}_1 \mid \mathbf{w}_1 \rangle = -(\mathbf{u}_1 \parallel \mathbf{w}_1) = \cos \alpha$ . Wir beweisen*

$$(\mathbf{u} \parallel \mathbf{w}) = \cosh \omega_1 \cdot \cosh \omega_2 - \sinh \omega_1 \cdot \sinh \omega_2 \cdot \cos \alpha$$

*In der Tat gilt  $(\mathbf{u} \parallel \mathbf{w}) = (\cosh \omega_1 \cdot \mathbf{v} + \sinh \omega_1 \cdot \mathbf{u}_1 \parallel \cosh \omega_2 \cdot \mathbf{v} + \sinh \omega_2 \cdot \mathbf{w}_1)$*

Der hyperbolische Cosinus ist auf der positiven Halbachse ansteigend. Die Dreiecksungleichung  $\omega_1 + \omega_2 \geq d(P, R)$  ist daher äquivalent mit der Aussage  $\cosh(\omega_1 + \omega_2) \geq \cosh d(P, R) = (\mathbf{u} \parallel \mathbf{w})$ . Diese Ungleichung ergibt sich mit Hilfe des Additionstheorems für den hyperbolischen Cosinus:  $\cosh(\omega_1 + \omega_2) = \cosh \omega_1 \cdot \cosh \omega_2 + \sinh \omega_1 \cdot \sinh \omega_2$ . Im Extremfall  $\alpha = 0, \cos \alpha = 1$  gilt  $\mathbf{u}_1 = \mathbf{w}_1$ ; und wir erhalten  $d(P, R) = |d(P, Q) - d(Q, R)|$ . Entsprechend ergibt sich  $d(P, R) = d(P, Q) + d(Q, R)$  für  $\alpha = \pi, \cos \alpha = -1$ .

Die Dreiecksungleichung wird also genau dann zu einer Gleichung, wenn  $P, Q, R$  kollinear sind mit  $Q$  zwischen  $P$  und  $R$ . Die Metrik kann somit als Kriterium für Kollinearität dienen. Die kürzeste Verbindung zweier Punkte ist die Verbindungsstrecke. Entlang der Verbindungsstrecke addieren sich die Abstände.

**Sprechweise 1.5.1** (Die Richtungen im Punkt  $Q$ ).

Es seien  $Q, P \in \mathcal{Hyp}_{\mathbb{M}}$ , dargestellt durch Einheitsvektoren  $\mathbf{v}$  bzw.  $\mathbf{u}$  im Vorwärtskegel. Es sei  $d(P, Q) = \omega$ . Es gibt genau einen zu  $\mathbf{v}$  orthogonalen Einheitsvektor  $\mathbf{u}_1$  ( $(\mathbf{v} \parallel \mathbf{u}_1) = 0, (\mathbf{u}_1 \parallel \mathbf{u}_1) = -1$ ), sodass  $\mathbf{u} = \mathbf{v} \cdot \cosh \omega + \mathbf{u}_1 \cdot \sinh \omega$ .

Wir sagen in diesem Fall, dass  $P$  im Abstand  $\omega$  von  $Q$  in der Richtung  $\mathbf{u}_1$  liegt.

Bemerke: Die Menge der möglichen von  $Q$  (oder  $\mathbf{u}$ ) ausgehenden Richtungen ist die Einheitskugel im Orthogonalraum von  $\mathbf{v}$ . Im Raum der von  $Q$  ausgehenden Richtungen haben wir eine euklidische Winkelmessung.

Was wir bewiesen haben, ist der

**Satz 1.5.3** (Hyperbolischer Cosinussatz). Seien  $\Delta(P, Q, R)$  ein Dreieck mit dem Winkel  $\alpha$  bei  $Q$ . Für die Seitenlängen  $\omega_1 = d(P, Q)$ ,  $\omega_2 = d(R, Q)$ ,  $\omega = d(P, R)$  gilt dann

$$c = c_1 \cdot c_2 - s_1 \cdot s_2 \cdot \cos \alpha.$$

wobei  $c = \cosh \omega$ ,  $c_1 = \cosh \omega_1$ ,  $c_2 = \cosh \omega_2$ ,  $s_1 = \sinh \omega_1$ ,  $s_2 = \sinh \omega_2$ .

Ihm entspricht in der euklidischen Geometrie der Satz:

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos \alpha.$$

Und dieser vertraute Satz approximiert unsere Formel für kleine Seitenlängen:

$$1 + \frac{1}{2}\omega^2 \sim \left(1 + \frac{1}{2}\omega_1^2\right) \cdot \left(1 + \frac{1}{2}\omega_2^2\right) - \omega_1 \cdot \omega_2 \cdot \cos \alpha,$$

Wir wollen die Aussage noch etwas weiter verfolgen. Auf der durch  $Q$  und  $P$  bestimmten Geraden gibt es noch einen weiteren Punkt  $P'$  im Abstand  $\omega_1$  von  $Q$ . Die beiden Punkte sind repräsentiert durch die Einheitsvektoren

$$\begin{aligned} \mathbf{u} &= \cosh \omega_1 \cdot \mathbf{v} + \sinh \omega_1 \cdot \mathbf{u}_1 \\ \mathbf{u}' &= \cosh \omega_1 \cdot \mathbf{v} - \sinh \omega_1 \cdot \mathbf{u}_1 \end{aligned}$$

Entsprechend sei  $R'$  definiert. Das ‘Parallelogramm’ mit den Ecken  $P, R, P', R'$  hat die Seitenlängen  $|\omega| = d(P, R)$  und  $|\omega'| = d(R, P')$ . Die Zahlen  $2|\omega_1| = d(P, P')$  und  $2|\omega_2| = d(R, R')$  können wir als die Längen der Diagonalen verstehen.

Die Abstände der ‘Ecken’  $P, P'$  vom weiteren Punkt  $R$  haben wir oben berechnet.

$$\begin{aligned} \cosh d(P, R) &= \cosh \omega_1 \cdot \cosh \omega_2 + \sinh \omega_1 \cdot \sinh \omega_2 \cdot \cos \alpha \\ \cosh d(P', R) &= \cosh \omega_1 \cdot \cosh \omega_2 - \sinh \omega_1 \cdot \sinh \omega_2 \cdot \cos \alpha \end{aligned}$$

Addition ergibt den

**Satz 1.5.4** (Hyperbolische Parallelogrammgleichung).

Mit den Bezeichnungen von oben gilt

$$\begin{aligned} \cosh \omega + \cosh \omega' &= 2 \cdot \cosh \omega_1 \cdot \cosh \omega_2. \\ \cosh d(P, R) + \cosh d(P', R) &= 2 \cdot \cosh d(Q, P) \cdot \cosh d(Q, R). \end{aligned}$$

Für kleine  $\omega$  gilt  $\cosh \omega = 1 + \frac{1}{2}\omega^2 + O(\omega^4)$ . Die ‘hyperbolische Parallelogrammgleichung’ geht über in die euklidische Parallelogrammgleichung; denn

$$(1 + \frac{1}{2}\omega^2) + (1 + \frac{1}{2}\omega'^2) \sim 2 \cdot (1 + \frac{1}{2}\omega_1^2) \cdot (1 + \frac{1}{2}\omega_2^2), \quad 2\omega^2 + 2\omega'^2 \sim d^2(P, P') + d^2(R, R').$$

**Sprechweise 1.5.2.** Die Punkte  $P, Q, R \in \mathcal{Hyp}_{\mathbb{M}}$  seien repräsentiert durch die Einheitsvektoren  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{M}_{++}$ . Die raumartigen Einheitsvektoren  $\mathbf{u}_1, \mathbf{w}_1$  im Orthogonalraum von  $\mathbf{v}$  seien wie oben konstruiert. Wir nennen dann  $\mathbf{u}_1, \mathbf{w}_1$  die Richtungen der Schenkel des Dreiecks  $\Delta(P, Q, R)$  im Eckpunkt  $Q$ . Die Zahl  $\cos \gamma = \langle \mathbf{u}_1, \mathbf{w}_1 \rangle$  heisst der Cosinus des Winkels bei  $Q$ . Wenn  $\cos \gamma = 0$ , dann sagt man, dass das Dreieck in  $Q$  einen rechten Winkel hat.

**Satz 1.5.5.** Im Fall eines bei  $Q$  rechtwinkligen Dreiecks  $\Delta(P, Q, R)$  gilt der hyperbolische Satz des Pythagoras

$$\cosh d(P, R) = \cosh d(Q, P) \cdot \cosh d(Q, R).$$

Daraus folgt offenbar für ‘kleine’ Dreiecke  $d(P, R)^2 \sim d(Q, P)^2 + d(Q, R)^2$ .

Es gibt weitere bemerkenswerte hyperbolische Analoga zu vertrauten Konstruktionen der euklidischen Geometrie: Die Differenz der obigen Abstandsformeln ergibt  $\cosh d(P, R) - \cosh d(P', R) = -2 \cos \alpha \cdot \sinh \omega_1 \cdot \sinh \omega_2$ , also

$$\cos \alpha = \frac{\cosh d(P', R) - \cosh d(P, R)}{2 \sinh d(Q, P) \cdot \sinh d(Q, R)}.$$

Dies entspricht der Polarisierungskonstruktion für das euklidische Skalarprodukt

$$\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4\langle \mathbf{x} | \mathbf{y} \rangle = 4 \cdot \|\mathbf{x}\| \cdot \|\mathbf{y}\|^2 \cdot \cos \alpha.$$

Bekanntlich ist in einem rechtwinkligen euklidischen Dreieck die Summe der beiden spitzen Winkel ein rechter:  $\alpha + \beta = \pi/2$ ,  $\cos \beta = \sin \alpha$ ,  $\cos^2 \beta + \cos^2 \alpha = 1$ .

In einem hyperbolischen Dreieck ist die Winkelsumme echt kleiner als zwei rechte Winkel. Wir studieren sie für ein rechtwinkliges Dreieck. Wenn die Längen der Katheten nach  $\infty$  streben, dann streben die Winkel nach 0.

**Satz 1.5.6.** *Es sei  $\Delta(P, Q, R)$  ein Dreieck mit einem rechten Winkel bei  $Q$  und den beiden spitzen Winkeln  $\alpha$ ,  $\beta$ . Mit  $s = \sinh d(P, R)$ ,  $s_1 = \sinh d(P, Q)$ ,  $s_2 = \sinh d(R, Q)$  gilt dann*

$$\cos^2 \beta + \cos^2 \alpha - 1 = \frac{s_1^2 \cdot s_2^2}{s^2} = \frac{s_1^2 \cdot s_2^2}{s_1^2 + s_2^2 + s_1^2 \cdot s_2^2}.$$

**Beweis.** *Wir benützen die Abkürzungen  $c = \cosh d(P, R)$ ,  $c_1 = \cosh d(P, Q)$ ,  $c_2 = \cosh d(R, Q)$ . Der hyperbolische Cosinussatz sagt uns*

$$\begin{aligned} c &= c_1 \cdot c_2, & c_1 &= c_2 \cdot c - s_2 \cdot s \cdot \cos \alpha, & c_2 &= c_1 \cdot c - s_1 \cdot s \cdot \cos \beta, \\ s \cdot \cos \alpha &= \frac{1}{s_2} (c_2 \cdot c - c_1) = c_1 \cdot s_2, & s \cdot \cos \beta &= c_2 \cdot s_1, \\ s^2 \cdot (\cos^2 \alpha + \cos^2 \beta) &= (c_1 \cdot s_2)^2 + (c_2 \cdot s_1)^2 \\ s^2 \cdot (\cos^2 \alpha + \cos^2 \beta - 1) &= c_1^2 \cdot (1 - c_2)^2 + c_2^2 \cdot (1 - c_1)^2 - (c_1 \cdot c_2)^2 + 1 = \\ &= c_1^2 \cdot c_2^2 - c_1^2 - c_2^2 + 1 = s_1^2 \cdot s_2^2. \end{aligned}$$

Bekanntlich ist in einem rechtwinkligen euklidischen Dreieck der Cosinus eines Winkels der Quotient 'Länge der anliegenden Kathete, geteilt durch die Länge der Hypotenuse' ( $a^2 + b^2 = c^2$ ;  $\cos \alpha = \frac{b}{c}$ ,  $\sin \alpha = \frac{a}{c}$ .)

Eine entsprechende Beziehung gibt es in der hyperbolischen Geometrie; an die Stelle der Seitenlängen  $\omega_1, \omega_2, \omega$  tritt hier aber der hyperbolische Tangens bzw. der hyperbolische Sinus dieser Seitenlängen.

**Satz 1.5.7** (Kathete und Hypotenuse).

*Es sei  $\Delta(P, Q, R)$  ein Dreieck mit einem rechten Winkel bei  $Q$ , dem Winkel  $\alpha$  bei  $R$  und den Seitenlängen  $\omega = d(P, R)$ ,  $\omega_1 = d(P, Q)$ ,  $\omega_2 = d(R, Q)$ . Es gilt dann*

$$\cos \alpha = \frac{\tanh \omega_2}{\tanh \omega}, \quad \sin \alpha = \frac{\sinh \omega_1}{\sinh \omega}.$$

**Beweis.** *Die Gleichung für den Cosinus ergibt sich aus der eben hergeleiteten Beziehung  $s \cdot \cos \alpha = c_1 \cdot s_2$ ,  $\cos \alpha = \frac{s_2 \cdot c}{c_2 \cdot s}$ . Daraus folgt*

$$\sin^2 \alpha = 1 - \cos^2 \alpha = \frac{c_2^2 \cdot s^2 - s_2^2 \cdot c^2}{c_2^2 \cdot s^2} = \frac{c_2^2 \cdot (c^2 - 1) - s_2^2 \cdot c^2}{c_2^2 \cdot s^2} = \frac{c_1^2 - 1}{s^2} = \frac{\sinh^2 \omega_1}{\sinh^2 \omega}.$$

**Satz 1.5.8** (Lot und Parallelwinkel).

*Es sei  $L_0$  eine Gerade in  $\mathcal{H}_{\mathbb{M}}$  und  $R$  ein Punkt ausserhalb.*

1. Es existiert dann genau ein  $Q \in L_0$ , sodass

$$d(R, Q) = \min\{d(R, \mathbf{v}) : \mathbf{v} \in L_0\}.$$

2. Wenn  $R$  durch  $\mathbf{w}$  dargestellt wird und

$$\mathbf{w} = \mathbf{v}_0 + \mathbf{w}' \quad \text{mit} \quad \mathbf{v}_0 \in L_0, \quad \mathbf{w}' \perp L_0,$$

dann stellt  $\mathbf{v}_0$  den Punkt  $Q$  dar, den ‘Fußpunkt des Lots’ von  $R$  auf  $L_0$ . Es gilt

$$\sinh d(R, L_0) = \sinh d(R, Q) = \frac{\|\mathbf{w}'\|}{\|\mathbf{w}\|}.$$

3. Für jedes  $P \in L_0$  ist  $\Delta(R, Q, P)$  ein bei  $Q$  rechtwinkliges Dreieck, für dessen Winkel bei  $R$  gilt

$$\cos \alpha < \tanh d(R, Q).$$

**Beweis.** Sei  $R$  sei durch den Einheitsvektor  $\mathbf{w}$  repräsentiert, und  $\mathbf{w} = \mathbf{v}_0 + \mathbf{w}'$  mit  $\mathbf{v}_0 \in L_0$ ,  $\mathbf{w}' \perp L_0$ . Sei  $\mathbf{v}_1$  der (bis auf das Vorzeichen eindeutig bestimmte) Einheitsvektor  $\in L_0$ , der orthogonal zu  $\mathbf{v}_0$  ist. Jedes  $P \in L_0$  besitzt eine Darstellung durch einen Einheitsvektor der Form  $\mathbf{v}_\omega = \frac{1}{\|\mathbf{v}_0\|} \cdot \mathbf{v}_0 \cdot \cosh \omega + \mathbf{v}_1 \cdot \sinh \omega$ .

Die Distanz  $d(P_\omega, R)$  ist minimal für  $\omega = 0$ ; denn  $\cosh d(\mathbf{v}_\omega, \mathbf{w}) = (\mathbf{v}_\omega | \mathbf{v}_0 + \mathbf{w}') = \cosh \omega \cdot \|\mathbf{v}_0\|$ . Wegen  $1 = \|\mathbf{w}\|^2 = \|\mathbf{v}_0\|^2 - \|\mathbf{w}'\|^2$  bedeutet das (für die orthogonale Zerlegung eines beliebigen darstellenden  $\mathbf{w}$ )

$$\mathbf{w} = \mathbf{v}_0 + \mathbf{w}' \quad \text{mit} \quad \mathbf{v}_0 \in L_0, \quad \mathbf{w}' \perp L_0 \quad \implies \quad \sinh d(\mathbf{w}, L_0) = \frac{\|\mathbf{w}'\|}{\|\mathbf{w}\|}.$$

Der Einheitsvektor zu  $R$  ist

$$\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{\mathbf{v}_0}{\|\mathbf{v}_0\|} \cdot \cosh \eta + \frac{\mathbf{w}'}{\|\mathbf{w}'\|} \cdot \sinh \eta \quad \text{mit} \quad \eta = d(R, L_0).$$

Dass der Winkel im Fußpunkt  $P_0$  ein rechter ist, ergibt sich aus dem hyperbolischen Satz von Pythagoras. Die obige Formel für  $d(R, P_\omega)$  liefert nämlich

$$\cosh d(R, P_\omega) = \cosh \omega \cdot \frac{\|\mathbf{v}_0\|}{\|\mathbf{w}\|} = \cosh d(P_0, P_\omega) \cosh d(R, P_0).$$

Wir betrachten für  $\omega \rightarrow \pm\infty$  den Winkel  $\alpha_\omega$  bei  $R$  im rechtwinkligen Dreieck  $\Delta(R, P_0, P_\omega)$ . Der dritte Winkel  $\beta_\omega$  strebt nach 0 und aus der obigen Formel für  $\cos^2 \alpha_\omega + \cos^2 \beta_\omega - 1$  ergibt sich

$$\cos^2 \alpha_\omega \longrightarrow \frac{\sinh^2 \eta}{1 + \sinh^2 \eta} = \tanh^2 \eta.$$

**Sprechweise 1.5.3** (Parallele durch einen Punkt).

Nach Lobatchevski heisst der Winkel  $\tilde{\alpha}$  mit  $\cos \tilde{\alpha} = \tanh d(R, L_0)$  der Parallelenwinkel in  $R$  bezüglich der hyperbolischen Geraden  $L_0$ . Wenn  $N$  eine der beiden Geraden in der durch  $L_0$  und  $R$  (ausserhalb  $L_0$ !) bestimmten Ebene ist, die mit dem Lot auf  $L_0$  den Winkel  $\pm \tilde{\alpha}$  hat, dann ‘schneidet’ sie  $L_0$  in einem ‘unendlichfernen’ Punkt. Die beiden (Halb)-Geraden  $N_{\pm}$  heissen die rechtsseitige und linksseitige Parallele zu  $L_0$  durch  $R$ .

In der euklidischen Geometrie haben bekanntlich alle Punkte einer zu  $L_0$  parallelen Geraden  $N$  denselben Abstand von der Geraden  $L_0$ . In der hyperbolischen Geometrie ist das ganz anders.

**Satz 1.5.9.** *Es sei  $N$  eine hyperbolische Gerade durch den Punkt  $R_0$ , die zur Geraden  $L_0$  parallel ist.  $R_{\omega}$  sei der Punkt auf der Geraden  $N$  im Abstand  $\omega$ . Es gilt dann*

$$\sinh d(R_{\omega}, L_0) = e^{-\omega} \cdot \sinh d(R_0, L_0).$$

**Beweis.** Die Punkte  $R$  der Halbgeraden  $N$  durch  $R_0$  mit der Richtung  $\mathbf{w}_1$  haben wir bisher vorzugsweise durch Einheitsvektoren repräsentiert:

$$R_0 \leftrightarrow \mathbf{w}_0, \quad R \leftrightarrow \mathbf{w}_0 \cdot \cosh \omega + \mathbf{w}_1 \cdot \sinh \omega, \quad \text{mit } \omega = d(R_0, R) \in [0, \infty).$$

Im Folgenden bevorzugen wir die Parametrisierung

$$R = R_{\beta} \leftrightarrow \mathbf{w}_{\beta} = \mathbf{w}_0 \cdot (1 - \beta) + (\mathbf{w}_1 + \mathbf{w}_0) \cdot \beta = \mathbf{w}_0 + \mathbf{w}_1 \cdot \beta \quad \text{mit } \beta = \tanh d(R_0, R) \in [0, 1).$$

Wir bemerken  $\|\mathbf{w}_{\beta}\| = \sqrt{1 - \beta^2} = \frac{1}{\cosh \omega}$ .  $\sqrt{\frac{1 - \beta}{1 + \beta}} = e^{-\omega}$ .

$L_0$  sei die Gerade durch  $P_0$ , den Fußpunkt des Lots von  $R$  auf  $L_0$  (repräsentiert durch den Einheitsvektor  $\mathbf{v}_0$ ) mit der Richtung  $\mathbf{v}_1$ . ( $(\mathbf{v}_0 \parallel \mathbf{v}_1) = 0$ ,  $(\mathbf{v}_1 \parallel \mathbf{v}_1) = -1$ ). Die Annahme, dass die Halbgeraden  $N$  und  $L_0$  sich im Unendlichen schneiden, bedeutet  $\mathbf{w}_1 + \mathbf{w}_0 = c \cdot (\mathbf{v}_1 + \mathbf{v}_0)$  mit  $c > 0$ . Wir gewinnen  $c = \frac{1}{\cosh \eta}$  aus  $\mathbf{w}_0 = \mathbf{v}_0 \cdot \cosh \eta + \mathbf{w}'$ , mit  $\mathbf{w}' \perp L_0$ ; denn  $-1 = (\mathbf{w}_1 \parallel \mathbf{w}_1) = (c \cdot (\mathbf{v}_1 + \mathbf{v}_0) - \mathbf{w}_0 \parallel c \cdot (\mathbf{v}_1 + \mathbf{v}_0) - \mathbf{w}_0) = -2c \cdot (\mathbf{v}_1 + \mathbf{v}_0 \parallel \mathbf{w}_0) + 1$ . Die orthogonale Zerlegung von  $\mathbf{w}_{\beta}$  bezüglich  $L_0 \oplus L_0^{\perp}$  ergibt

$$\mathbf{w}_{\beta} = \mathbf{w}_0 + \mathbf{w}_1 \cdot \beta = (\mathbf{w}_0 + \mathbf{w}_1) \cdot \beta + \mathbf{w}_0 \cdot (1 - \beta) = \mathbf{u} + \mathbf{w}' \cdot (1 - \beta) \quad \text{mit } \mathbf{u} \in L_0.$$

Die Formel für den Abstand von  $L_0$  ergibt

$$\sinh d(\mathbf{w}_{\beta}, L_0) = \frac{(1 - \beta) \|\mathbf{w}'\|}{\|\mathbf{w}_{\beta}\|} = \frac{1 - \beta}{\sqrt{1 - \beta^2}} \sinh \eta = e^{-\omega} \cdot \sinh d(R_0, L_0).$$

Mit den Richtungen in den verschiedenen Punkten  $\tilde{P} \in \mathcal{Hyp}_{\mathbb{M}}$  und mit den ‘tangentialen’ Vektoren müssen wir uns nun etwas eingehender befassen.

## Die Orthogonalräume als Tangentialräume

**Sprechweise 1.5.4.** Ist der Punkt  $\tilde{P} \in \mathcal{Hyp}_{\mathbb{M}}$  durch den zeitartigen Vektor  $\tilde{\mathbf{v}}$  repräsentiert, so nennt man den Orthogonalraum zu  $\tilde{\mathbf{v}}$  den Tangentialraum im Punkt  $\tilde{P}$ .

$$\mathcal{T}_{\tilde{P}} = \{\mathbf{v} : (\tilde{\mathbf{v}} \parallel \mathbf{v}) = 0\}.$$

Die Tangentialräume sind euklidische Räume:  $(\mathcal{T}_{\tilde{P}}, \langle \cdot | \cdot \rangle)$  mit  $\langle \cdot | \cdot \rangle = -(\|\cdot\|)$ .

Hinweis: Tangentialräume spielen in der Theorie der glatten Mannigfaltigkeiten eine fundamentale Rolle. In der allgemeinen Theorie repräsentieren die Elemente  $\mathbf{v}$  des Tangentialraums im Punkt  $\tilde{P}$  die Momentangeschwindigkeiten, in welchen parametrisierte glatte Kurven durch den Punkt  $\tilde{P}$  hindurchgehen können. Bei einer  $n$ -dimensionalen Mannigfaltigkeit ergibt das für jeden Punkt  $P$  einen  $n$ -dimensionalen Vektorraum  $\mathcal{T}_P$ . — Die Riemannschen Mannigfaltigkeiten sind dadurch ausgezeichnet, dass die Tangentialräume eine euklidische Struktur haben; man kann hier von der Länge eines Tangentialvektors und vom Cosinus des Winkels zwischen Tangentialvektoren sprechen.

Die Tangentialräume zu verschiedenen Fusspunkten  $P$  haben in der allgemeinen Theorie (abgesehen von den Glattheitseigenschaften) nicht viel miteinander zu tun. In unserem Fall scheint das anders zu sein: die Tangentialräume zu den verschiedenen  $P \in \mathcal{Hyp}_{\mathbb{M}}$  werden in der obigen Definition als dreidimensionale Teilräume des vierdimensionalen Minkowskiraums  $\mathbb{M}$  eingeführt. Auch bei der Festlegung der euklidischen Struktur in den verschiedenen Tangentialräumen wird Bezug genommen auf die metrische Struktur von  $\mathbb{M}$ . Wir werden sehen, wie das alles in unserem sehr besonderen Fall zusammenpasst.

**Didaktischer Hinweis:** Die obige Definitionsweise erinnert an die Vorstellungsweise der elementaren Analysis, wo man glatte Funktionen studiert, die auf einem  $N$ -dimensionalen affinen Raum  $\mathcal{A}$  definiert sind. Die ‘Tangentialebene’ zu einer glatten Funktion  $f(\mathbf{x})$  im Punkt  $\tilde{\mathbf{x}}$  versteht man (im Sinne der Schulmathematik) als die Hyperebene  $\mathcal{H}_{\tilde{\mathbf{x}}} \subset \mathcal{A}$ , welche folgendermaßen bestimmt ist:

$$\mathcal{H}_{\tilde{\mathbf{x}}} \ni \tilde{\mathbf{x}} + \mathbf{v} \iff f(\tilde{\mathbf{x}} + t \cdot \mathbf{v}) - f(\tilde{\mathbf{x}}) = o(t) \quad \text{für } t \rightarrow 0.$$

Die Menge der ‘Verschiebungsvektoren’  $\mathbf{v}$  der affinen ‘Tangentialebene’  $\mathcal{H}_{\tilde{\mathbf{x}}}$  nennt man die Menge der Tangentialvektoren zur Funktion  $f(\cdot)$  im Punkt  $\tilde{\mathbf{x}}$ . Man sagt dann,  $\mathbf{v}$  sei ein Tangentialvektor in  $\tilde{\mathbf{x}}$ , wenn  $\mathbf{v}$  im Nullraum des Gradienten von  $f$  im Punkt  $\tilde{\mathbf{x}}$  liegt.

Betrachten wir in diesem Sinne die Funktion  $f(\mathbf{x}) = x_0^2 - x_1^2 - x_2^2 - x_3^2$  im affinen Raum der reellen 4-Spalten  $(x_0, x_1, x_2, x_3)^T$ . Der Gradient im Punkt  $\tilde{\mathbf{x}}$  ist die Zeile  $(2\tilde{x}_0, -2\tilde{x}_1, -2\tilde{x}_2, -2\tilde{x}_3)$ . Der Nullraum ist die Menge der Spalten  $(v_0, v_1, v_2, v_3)^T$  mit  $2\tilde{x}_0 v_0 - 2\tilde{x}_1 v_1 - 2\tilde{x}_2 v_2 - 2\tilde{x}_3 v_3 = 0$ . Diese Bedingung kann man nun bequem in der Notation des Minkowskiraums schreiben:  $(\tilde{\mathbf{x}} \parallel \mathbf{v}) = 0$ . Wichtig ist nun die Bemerkung, dass es sich um eine Bedingung an den ‘Vektor’  $\mathbf{v}$  handelt, die für alle Repräsentanten  $\tilde{\mathbf{u}}$  des Punkts  $\tilde{P} \in \mathcal{Hyp}_{\mathbb{M}}$  dieselbe ist. Wenn man den Orthogonalraum zum zeitartigen Vektor  $\tilde{\mathbf{u}}$ , einen Teilvektorraum des Minkowskiraums also, ‘gleichsetzen’ will mit dem Tangentialraum  $\mathcal{T}_{\tilde{P}}$  im Sinne des hyperbolischen Raums als Riemannscher Mannigfaltigkeit,

dann muss man sich vergewissern, dass die Riemann'sche Metriken in Tangentialräumen tatsächlich die sind, die sich aus den Metriken in den zeitartigen Teilräumen des Minkowskiraums ergeben. Die folgende Rechnung zeigt, dass das tatsächlich der Fall ist.

**Satz 1.5.10.** *Der Punkt  $\tilde{P}$  sei durch den zeitartigen Einheitsvektor  $\tilde{\mathbf{v}}$  repräsentiert. Wir nähern uns auf der Kurve  $\{P_t : t \rightarrow 0\}$ , wo  $P_t$  durch den Vektor  $\tilde{\mathbf{v}} + t \cdot \mathbf{v}$  repräsentiert wird, mit  $(\tilde{\mathbf{v}} \parallel \mathbf{v}) = 0$ . Für den hyperbolische Abstand gilt dann  $d(t) = d(\tilde{P}, P_t) = \|\mathbf{v}\| \cdot |t| + O(t^2)$ .*

**Beweis.** *Wir nehmen an  $(\mathbf{v} \parallel \mathbf{v}) = -1$  und beweisen  $\cosh d(t) = 1 + \frac{1}{2}t^2 + O(t^4)$ .*

$$\begin{aligned} \|\tilde{\mathbf{v}}\| &= 1, & (\tilde{\mathbf{v}} + t \cdot \mathbf{v} \parallel \tilde{\mathbf{v}}) &= 1 \\ \implies \cosh d(t) &= (\tilde{\mathbf{v}} + t \cdot \mathbf{v} \parallel \tilde{\mathbf{v}} + t \cdot \mathbf{v})^{-1/2} = (1 + t^2(-1))^{-1/2} = 1 + \frac{1}{2}t^2 + O(t^4). \end{aligned}$$

### Der 3-dimensionale hyperbolische Raum, Begriffsbestimmung

Mathematische Laien fragen gelegentlich, was denn der euklidische Raum sei. Studierenden der Linearen Algebra geben sich mit der folgenden Begriffsbestimmung zufrieden: *Ein euklidischer Raum ist ein (orientierter dreidimensionaler) reellaffiner Raum, in dessen Vektorraum eine positive quadratische Form ausgezeichnet ist.* Eine ähnlich knappe Begriffsbestimmung für den hyperbolischen Raum wird den Studierenden selten angeboten. Es kommen mehrere interessante Zugänge in Frage.

In den Anfängen (bei Bolyai, Lobatschevski, Gauss) orientierte man sich an den auf Euklid zurückgehenden Begriffsbestimmungen für den euklidischen Raum. Es gibt Punkte, Geraden (und im dreidimensionalen Fall) auch Ebenen. Die Inzidenzstruktur und die Winkelmessung stehen am Anfang der Begriffsbestimmung. Diese Art von Zugang wurde um 1900 von Hilbert auf sichere Füße gestellt, und die Pioniere der Disziplin 'Grundlagen der Geometrie' haben das im gleichen Sinn für andere Räume, insbesondere für den hyperbolischen Raum geleistet. Im Falle der hyperbolischen Geometrie konnte man sehr viel übernehmen. Es entstand auch eine sog. absolute Geometrie, in welcher man solche Sätze ableitete, welche weder auf das Parallelenaxiom noch auf seine Negation Bezug nehmen.

Nun kann aber alternativ die hyperbolische Geometrie als ein besonders schönes Beispiel innerhalb der Riemann'schen Geometrie studiert werden. Und es spricht viel für einen solchen Zugang, wenn man bedenkt, welche zentrale Stellung die Begriffe der Differentialgeometrie in der heutigen Mathematik einnehmen: Glatte Mannigfaltigkeiten, lokale Koordinatensysteme, Tangentialräume und Cotangentialräume, Differentialformen, Riemann'sche Metrik, Geodäten usw. .

Historisch gesehen wurde die Riemann'sche Geometrie im Anschluss an Arbeiten von Gauss und Riemann entwickelt. Während Gauss die Theorie der gekrümmten zweidimensionalen Flächen im dreidimensionalen euklidischen Raum entwickelte, löste sich Riemann vom Gedanken der eingebetteten Fläche und konzentrierte sich auf die mathematische Beschreibung der inneren Eigenschaften der '*n*-fach ausgedehnten Grössen'. Riemann beginnt seine berühmte Untersuchung 'Über die Hypothesen, die der Geometrie zu Grunde liegen' (1854) mit den Worten:

Bekanntlich setzt die Geometrie sowohl den Begriff des Raumes, als die ersten Grundbegriffe für die Constructionen im Raume als etwas Gegebenes voraus. Sie giebt von ihnen nur Nominaldefinitionen, während die wesentlichen Bestimmungen in Form von Axiomen auftreten. Das Verhältnis der Voraussetzungen bleibt dabei im Dunkeln; man sieht weder ein, ob und in wie weit ihre Verbindung nothwendig, noch a priori, ob sie möglich ist.

Der Ansatz von Riemann hat die Vorstellungen von Geometrie von Grund auf verändert. Eine Verbindung der alten euklidischen Vorstellungen von Geometrie mit den Herangehensweisen im Sinne von Riemann, Einstein und anderen ist eine didaktische Herausforderung.

**Beispiele für Mannigfaltigkeiten** Wir haben das Wort Mannigfaltigkeit im Sinne der Riemann'schen Geometrie schon öfters benützt. Nachdem wir nun schon einige interessante Beispiele kennengelernt haben, scheint es an der Zeit, ein wenig spezifischer zu werden.

Man sagt: Eine  $n$ -dimensionale Mannigfaltigkeit sieht lokal genau so aus, wie eine offene Menge im  $n$ -dimensionalen reellaffinen Raum. Was das genau heisst, lernt man in der Analysis. Wir wollen die Idee hier lieber durch einige Beispiele verdeutlichen.

Man ist gelegentlich froh für eine Veranschaulichung einer Mannigfaltigkeit durch eine Teilmenge eines höherdimensionalen affinen Raums. Ein interessantes Beispiel einer dreidimensionalen Mannigfaltigkeit ist  $SU(2, \mathbb{C})$ , die Menge aller unimodularen unitären Matrizen  $U$ . Topologisch gesehen handelt es sich um eine 3-Sphäre.

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 + |b|^2 = 1 \quad \iff \quad (a_1, a_2, b_1, b_2) \in \mathbb{R}^4 \quad \text{mit} \quad a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1.$$

Damit haben wir  $SU(2, \mathbb{C})$  als die Einheitssphäre in den euklidischen  $\mathbb{R}^4$  eingebettet. In der Mannigfaltigkeit gibt es eine ausgezeichnete Riemann'sche Metrik, es gibt nämlich eine (bis auf einen 'Skalierungsfaktor' eindeutig bestimmte) translationsinvariante Riemann'sche Metrik. Die geodätischen Linien sind die Großkreise, d.h. eine kürzeste Kurve von  $P$  nach  $Q$  verläuft immer entlang dem Großkreis, den man bei der obigen Einbettung dadurch gewinnt, dass man die Ebene durch  $P, Q$  und den Koordinatenursprung mit der Sphäre zum Schnitt bringt.

Neben der Einbettung ist der Übergang zu Äquivalenzklassen ein wichtiges Konstruktionsprinzip. Beispielsweise gewinnt man die Punkte des  $n$ -dimensionalen projektiven Raums als die eindimensionalen Teilvektorräume eines  $(n+1)$ -dimensionalen Vektorraums. Man betrachtet zwei  $(n+1)$ -Tupel  $(x_0, x_1, \dots, x_n)$  als äquivalent, wenn sie durch Multiplikation mit einem Faktor  $\neq 0$  auseinander hervorgehen. Wenn man die Äquivalenzklassen (naheliegenderweise) durch normierte Tupel repräsentieren will, ( $x_0^2 + x_1^2 + \dots + x_n^2 = 1$ ), dann muss man beachten, dass zwei gegenüberliegende Punkte stets zu identifizieren sind. Es stellt sich heraus, dass (für  $n > 1$ ) der  $n$ -dimensionale projektive Raum  $\mathbb{P}^n$  global ganz anders aussieht als die  $n$ -Sphäre. So gibt es z. B. in  $\mathbb{P}^n$  geschlossene Kurven, die man nicht auf einen Punkt zusammenziehen kann.

Wir haben oben die dreidimensionale Mannigfaltigkeit  $\mathcal{Hyp}_{\mathbb{M}}$  als eine Teilmenge des vierdimensionalen projektiven Raums verstanden: seine Punkte sind die zeitartigen Geraden. Hier bieten sich mehrere Methoden an, aus den Äquivalenzklassen Repräsentanten auszuwählen. Man erhält das sog. Hyperboloidmodell, wenn man den (wohlbestimmten!) normierten Einheitsvektor  $\tilde{\mathbf{v}}$  im Vorwärtskegel als Repräsentanten der zeitartigen Geraden  $\tilde{P} \in \mathcal{Hyp}_{\mathbb{M}}$  wählt. Das Hyperboloidmodell hat den Vorzug, dass die euklidischen Metriken in den Tangentialhyperebenen durch die (auf die dazugehörigen Vektorräume eingeschränkte) Minkowskimetrik beschrieben werden. Eine Basis des Tangentialraums im Punkt  $\tilde{\mathbf{v}}$  ist genau dann eine ONB, wenn sie zusammen mit  $\tilde{\mathbf{v}}$  eine ONB des Minkowskiraums ist.

### Das Ballmodell des hyperbolischen Raums

Eine andere Auswahl von Repräsentanten gewinnt man, wenn man den Vorwärtskegel mit einer affinen Hyperebene schneidet, die parallel zu einem raumartigen Teilvektorraum ist. Die Menge der Repräsentanten erscheint in dieser Repräsentation als das Innere einer Kugel, oder vielmehr (in der hier angemessenen Sprache der affinen Geometrie korrekt gesagt) als das Innere eines Ellipsoids. (Wenn wir vom Ball-Modell sprechen, dann soll das nicht das Missverständnis hervorrufen, dass hier irgendwo eine euklidische Metrik im Spiel ist. Eine euklidische Metrik hat aber auf dem Ellipsoid ('Ball') nichts verloren. Euklidische Metriken gibt es nur auf den Tangentialräumen) Ist der schneidende affine Raum parallel zum Tangentialraum in  $\tilde{P}$ , so beschreibt die auf das Ellipsoid eingeschränkte Minkowskinorm die infinitesimalen Größenverhältnisse in der Nähe von  $\tilde{P}$ ; in anderen Punkte erscheinen die Winkel verzerrt. Der anschauliche Vorzug des 'Ballmodells' ist der, dass man hier die Geraden (d.h. die Geodäten des hyperbolischen Raums) als Geradenstücke ('Sehnen') im Ellipsoid erkennt. Die Ebenen im hyperbolischen Raum erscheinen als ebene Schnitte durch das Ellipsoid. Die beiden unendlichfernen Punkte 'auf' einer Geraden sieht man als die Schnittpunkte mit dem Rand des Ellipsoids. Die Menge der unendlichfernen Punkte 'auf' einer Ebene erscheint als ein Kreis (besser eine Ellipse) auf dem Rand des Ellipsoids. Eine anderes ist es nun aber, in diesem Bild Abstände und Winkel zu bestimmen. Darüber gibt der folgende Satz Auskunft.

**Satz 1.5.11** (Distanz und Doppelverhältnis).

*Es seien  $\mathbf{u}_0, \mathbf{v}_0$  Vektoren im Vorwärtskegel des Minkowskiraums  $(\mathbb{M}, (\cdot\|\cdot))$  mit  $(\mathbf{u}_0\|\mathbf{v}_0) = \cosh \omega \cdot \|\mathbf{u}_0\| \cdot \|\mathbf{v}_0\|$ . Die Gerade durch diese Punkte schneidet den Lichtkegel in zwei Punkten  $\mathbf{u}', \mathbf{u}''$ . Für das Doppelverhältnis der vier Punkte auf der Geraden gilt*

$$DV(\mathbf{u}', \mathbf{u}''; \mathbf{u}_0, \mathbf{v}_0) = e^{\pm 2\omega}.$$

**Beweis.** *Das Doppelverhältnis ist positiv, weil  $\mathbf{u}_0, \mathbf{v}_0$  zwischen den Punkten  $\mathbf{u}', \mathbf{u}''$  liegen. Das Doppelverhältnis ist genau dann  $< 1$ , wenn  $\mathbf{u}_0$  näher an  $\mathbf{u}'$  liegt als  $\mathbf{v}_0$ . Die folgende Version der Behauptung vermeidet die Unsicherheit mit den Vorzeichen:*

$$|\omega| = \frac{1}{2} \cdot |\ln DV(\mathbf{u}', \mathbf{u}''; \mathbf{u}_0, \mathbf{v}_0)|.$$

Die Gerade durch die Punkte  $\mathbf{u}_0, \mathbf{v}_0$  ist die Punktmenge  $\{\mathbf{u} = (1 - a) \cdot \mathbf{u}_0 + a \cdot \mathbf{v}_0\}$ . Bekanntlich gilt

$$DV(P, Q; R, S) = DV(R, S; P, Q) = \frac{1-a}{a} \cdot \frac{b}{1-b}, \quad \text{wenn}$$

$$P = (1 - a)R + aS, \quad Q = (1 - b)R + bS.$$

Wir können annehmen, dass die Vektoren normiert sind. Man prüft leicht die folgenden Formeln für die Schnittpunkte mit dem Lichtkegel.

$$\mathbf{u}' = \frac{2}{\sinh \omega/2} \cdot (e^{\omega/2} \mathbf{u}_0 - e^{-\omega/2} \mathbf{v}_0); \quad \mathbf{u}'' = \frac{2}{\sinh \omega/2} \cdot (e^{\omega/2} \mathbf{v}_0 - e^{-\omega/2} \mathbf{u}_0).$$

$(\mathbf{u}' \parallel \mathbf{u}') = \text{const} \cdot (e^\omega - 2(\mathbf{u}_0 \parallel \mathbf{v}_0) + e^{-\omega}) = 0$ . Einsetzen ergibt das Resultat  $e^{2\omega}$ .

Wir kehren zum Abschluss des Abschnitts noch einmal zurück in die Vorstellungswelt des Minkowskiraums. Die Möglichkeit, von der Vorstellungswelt des hyperbolischen Raums in diese Vorstellungswelt (und zurück) zu wechseln, erweist sich immer wieder einmal als hilfreich.

**Das sog. Zwillingparadox** Wir stellen uns zwei Beobachter vor, die sich zu einem gewissen Zeitpunkt am gleichen Ort  $O$  befinden. Der eine bleibt (im Sinne seines Inertialsystems) in Ruhe, und kommt damit in der Zeitspanne  $t$  hin zu einem Ereignis  $O + \mathbf{w} \in \mathbb{M}$ . Der andere ‘bewegt’ sich mit einer gleichförmigen Geschwindigkeit zum Ereignis  $O + \mathbf{u}_0$  und braucht dazu (in seinem Inertialsystem!) die Zeit  $\|\mathbf{u}_0\|$ ; er ‘reist’ dann mit gleichförmiger Geschwindigkeit weiter zum Ereignis  $O + \mathbf{u}_0 + \mathbf{v}_0 = O + \mathbf{w}$ , um dort den Zwilling zu treffen, für den (nach seiner Zeitrechnung) die längere Zeit  $\|\mathbf{w}\| = \|\mathbf{u}_0 + \mathbf{v}_0\|$  vergangen ist.

**Der Lorentzfaktor** Es seien  $P$  und  $Q$  Beobachter mit verschiedenen Inertialsystemen zur Zeit  $0$  am gleichen Ort. Jeder befinde sich aus seiner Sicht in Ruhe. Sie ‘bewegen’ sich also auf Geraden im Minkowskiraum, auf den Geraden, die durch die zeitartigen Vektoren  $\mathbf{u}_0$  bzw.  $\mathbf{v}_0$  in ihren Inertialsystemen gegeben sind. Der Beobachter  $P$  stellt fest, dass sich der Beobachter  $Q$  nach der Zeit  $1$  (nach seiner Zeitrechnung) im Punkt  $O + \mathbf{w}$  befindet, der durch den Schnittpunkt der Hyperebene  $L = \{x_0 = 1\}$  mit der Weltlinie von  $Q$  bestimmt ist. Dieses Ereignis erreicht  $Q$  (in seiner Zeitrechnung) nach der Zeitspanne  $s$ , die wir aus dem hyperbolischen Abstand der Vektoren  $\mathbf{u}_0, \mathbf{v}_0$  bestimmen. Dieser hyperbolischen Abstand lässt uns auch bestimmen, mit welcher Geschwindigkeit sich  $Q$  entfernt (aus der Sicht von  $P$ ).

$$s \cdot \mathbf{v}_0 = \mathbf{w} = 1 \cdot \mathbf{u}_0 + \mathbf{v} \quad \text{mit} \quad (\mathbf{v} \parallel \mathbf{u}_0) = 0.$$

Das ‘innere Produkt’ mit  $\mathbf{u}_0$  liefert  $s \cdot \cosh \omega = 1$ . Für die Länge von  $\mathbf{v}$  im euklidischen Orthogonalraum von  $\mathbf{u}_0$  ergibt sich

$$\langle \mathbf{v} \parallel \mathbf{v} \rangle = -(\mathbf{u}_0 - s \cdot \mathbf{v}_0 \parallel \mathbf{u}_0 - s \cdot \mathbf{v}_0) = 1 - 2s \cdot (\mathbf{u}_0 \parallel \mathbf{v}_0) + s^2 = -1 + s^2 = -1 + \frac{1}{(\mathbf{u}_0 \parallel \mathbf{v}_0)^2}.$$

Dies erlaubt eine Interpretation des 'hyperbolischen Abstands' durch die Relativgeschwindigkeit  $|v| = \|\mathbf{v}\|$ , nämlich

$$(\mathbf{u}_0 \| \mathbf{v}_0) = \cosh \omega = \frac{1}{\sqrt{1-v^2}}; \quad v = \tanh \omega.$$

Wenn der hyperbolischen Abstand  $|\omega|$  sehr groß wird, dann nähert sich die Relativgeschwindigkeit  $|v| = \|\mathbf{v}\|$  dem Wert 1; in physikalischen Dimensionen ist das natürlich die Lichtgeschwindigkeit. Die Wahl eines Vorzeichens für  $\omega$  impliziert eine Wahl des Vorzeichens für  $v$ . Das Vorzeichen ist wichtig für das folgende 'Additionstheorem' für Relativgeschwindigkeiten in einer festen Ebene des Minkowskiraums.

**Satz 1.5.12.** *Es seien  $\mathbf{u}_0, \mathbf{v}_0$  Einheitsvektoren im Vorwärtskegel. Dazwischen, in der von ihnen aufgespannten Ebene  $L_0$  sei  $\mathbf{w}_0$  ein weiterer zeitartiger Einheitsvektor*

$$(\mathbf{u}_0 \| \mathbf{w}_0) = \cosh \omega_1 = \frac{1}{\sqrt{1-v_1^2}} \quad (\mathbf{w}_0 \| \mathbf{v}_0) = \cosh \omega_2 = \frac{1}{\sqrt{1-v_2^2}}$$

*Für die Geschwindigkeit  $v$ , mit der sich die Beobachter zu  $\mathbf{u}_0$  und  $\mathbf{v}_0$  voneinander entfernen, gilt dann*

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}; \quad \frac{1}{\sqrt{1-v^2}} = \cosh (\omega_1 + \omega_2) = (\mathbf{u}_0 \| \mathbf{v}_0)$$

**Beweis.** *Wir wollen  $c := (\mathbf{u}_0 \| \mathbf{v}_0) = \cosh \omega$  berechnen. Wir zeigen*

$$(\mathbf{v}_0 \| \mathbf{w}_0) = \cosh \omega \cdot \cosh \omega_1 - \sinh \omega \cdot \sinh \omega_1 = \cosh (\omega - \omega_1) \quad \text{also} \quad \omega - \omega_1 = \omega_2.$$

*In der Ebene  $L_0$  gibt es einen (raumartigen) Einheitsvektor  $\mathbf{u}_1$  mit  $(\mathbf{u}_0 \| \mathbf{u}_1) = 0$ . Für diesen (bis auf das Vorzeichen eindeutig bestimmten) Vektor gilt*

$$\begin{aligned} \mathbf{u}_1 &= a\mathbf{u}_0 + b\mathbf{u}_1; & 0 &= (\mathbf{u}_1 \| \mathbf{u}_0) = a + b \cdot c, \\ \mathbf{v}_0 &= c \cdot \mathbf{u}_0 + \frac{1}{b} \cdot \mathbf{u}_1; & 1 &= (\mathbf{v}_0 \| \mathbf{v}_0) = c^2 - \frac{1}{b^2}; & \frac{1}{b^2} &= (\sinh \omega)^2. \\ \mathbf{v}_0 &= \mathbf{u}_0 \cosh \omega \pm \mathbf{u}_1 \sinh \omega, & \mathbf{w}_0 &= \mathbf{u}_0 \cosh \omega_1 \pm \mathbf{u}_1 \sinh \omega_1, \\ (\mathbf{v}_0 \| \mathbf{w}_0) &= \cosh \omega \cosh \omega_1 - \sinh \omega \sinh \omega_1 = \cosh (\omega - \omega_1). \end{aligned}$$

*Die Formel für die Geschwindigkeiten ergibt sich aus dem Additionstheorem für den hyperbolischen Tangens*

$$\tanh (\omega_1 + \omega_2) = \frac{\tanh \omega_1 + \tanh \omega_2}{1 + \tanh \omega_1 \cdot \tanh \omega_2}.$$

Im Minkowskiraum  $\mathbb{M}$  sei eine zeitartige Gerade  $\tilde{P}$  als 'Zentrum' ausgezeichnet. Ein zentrierter Minkowskiraum ist die geeignete mathematische Struktur um die Gesamtheit aller gleichförmigen Geschwindigkeiten aus der Sicht eines Beobachters, der sich in seinem

Inertialsystem, dem sog. Laborsystem, in Ruhe befindet, zu studieren. Unser Beobachter sieht andere Beobachter, die sich in ihrem Inertialsystem in Ruhe befinden mit gleichförmiger Geschwindigkeit von sich wegfliegen, in einer bestimmten Richtung mit einer gewissen Geschwindigkeit  $v$ . (Wir stellen uns vor, dass sie sich alle zu einem Zeitpunkt 0 allesamt am gleichen Ort befanden.) Die möglichen Richtungen werden durch die Einheitsvektoren im Tangentialraum  $\mathcal{T}_{\tilde{P}}$  beschrieben. Bei der Messung der Relativgeschwindigkeit müssen wir sorgfältig argumentieren. Wenn sich ein Objekt (z. B. in der Richtung der ersten Koordinatenachse) mit der Geschwindigkeit  $v$  vom zentralen Beobachter  $\tilde{P}$  entfernt, dann läuft es im Minkowskiraum  $\mathbb{R} \times \tilde{\mathbb{M}}$  auf der Geraden  $G_v = \{(t, v \cdot t, 0, 0) : t \in \mathbb{R}\}$ . Diese Gerade wollen wir umparametrisieren;

$$t = \frac{1}{\sqrt{1-v^2}} \cdot s, \quad G_v = \left\{ \left( \frac{1}{\sqrt{1-v^2}} \cdot s, \frac{v}{\sqrt{1-v^2}} \cdot s, 0, 0 \right) : s \in \mathbb{R} \right\}$$

Wenn wir jetzt auch noch die Relativgeschwindigkeit mit dem hyperbolischen Tangens umparametrisieren,  $v = \tanh \omega$ , erhalten wir

$$\frac{1}{\sqrt{1-v^2}} = \cosh \omega, \quad \frac{v}{\sqrt{1-v^2}} = \sinh \omega, \\ G_v = \{ (s \cdot \cosh \omega, s \cdot \sinh \omega, 0, 0) : s \in \mathbb{R} \}.$$

Diese Gerade  $G_v$  wird durch einen Punkt  $P_v$  im hyperbolischen Abstand  $|\omega|$  in x-Richtung repräsentiert. Die beobachtete Relativgeschwindigkeit ist  $\tanh \omega$ .

Stellen wir uns ein Objekt vor, welcher mit (aus seiner Sicht) gleichbleibender Beschleunigung in eine bestimmte Richtung von  $\tilde{P}$  wegfliegt, beginnend mit der Relativgeschwindigkeit 0 im Ausgangspunkt. Der Punkt  $P_t$  im hyperbolischen Raum der Geschwindigkeiten 'wandert' auf der betreffenden Geraden. Aus der Sicht der hyperbolischen Geometrie würde man sagen, dass er die Gerade mit gleichförmiger Geschwindigkeit durchläuft. Aus der Sicht des Minkowskiraums bedeutet das, dass das Objekt zur Zeit  $\omega$  die Relativgeschwindigkeit  $v_\omega = \tanh(c \cdot \omega)$  erreicht. Der Ort steht hier nicht zur Debatte.

## 1.6 Die Bewegungen des hyperbolischen Raums

Jede eigentliche Lorentztransformationen  $\varphi(\cdot) \in \mathcal{L}_{\mathbb{M}}^{\uparrow+}$  liefert eine positivorientierte isometrische Abbildung des hyperbolischen Raums  $\mathcal{Hyp}_{\mathbb{M}}$ . Wir nennen sie die durch  $\varphi(\cdot)$  bestimmte Bewegung. Offensichtlich ist  $\varphi(\cdot) \in \mathcal{L}_{\mathbb{M}}^{\uparrow+}$  durch diese Bijektion von  $\mathcal{Hyp}_{\mathbb{M}}$  eindeutig bestimmt; wir bezeichnen sie ebenfalls mit  $\varphi(\cdot)$ . Die Gruppe dieser Bewegungen nennen wir die Bewegungsgruppe  $\mathcal{L}_{\mathcal{Hyp}}^+$  des hyperbolischen Raums  $\mathcal{Hyp}$ .

Hinweis: Es gibt axiomatische Beschreibungen des hyperbolischen Raums und seiner Bewegungsgruppe, die sich nicht auf die hier ausgearbeitete Darstellung in einem Minkowskiraum beziehen. Darauf wollen wir aber nicht eingehen. Für uns ist es nur eine Umdeutung, wenn wir einmal von einer eigentlichen Lorentztransformation  $\varphi(\cdot)$ , und das andere Mal von einer Bewegung  $\varphi(\cdot)$  des hyperbolischen Raums sprechen. Die Umdeutungen sollen von Fall zu Fall die 'Anschauung' unterstützen.

**Sprechweise 1.6.1** (Drehungen). Wenn eine Bewegung des hyperbolischen Raums  $\mathcal{Hyp}_{\mathbb{M}}$  einen Punkt  $\tilde{P}$  festlässt, dann heisst sie eine (hyperbolische) Drehung um den Punkt  $\tilde{P}$ . Die dazugehörige Lorentztransformation des Minkowskiraum  $\mathbb{M}$  wird gleichfalls eine (hyperbolische) Drehung genannt, eine Drehung um die  $\tilde{P}$  repräsentierende zeitartige Gerade. Die dreidimensionale Gruppe  $\mathcal{U}_{\tilde{P}}$  der Drehungen um den fest gewählten Punkt  $\tilde{P}$  (als Untergruppe der Bewegungsgruppe  $\subset \mathcal{L}_{\mathcal{Hyp}}^+$ ) heisst die Fixgruppe von  $\tilde{P}$ .

Mit der Struktur solcher Fixgruppen  $\mathcal{U}_{\tilde{P}}$  haben wir uns früher ausführlich beschäftigt, als wir die Inertialsysteme unbewegter Beobachter studierten. Die Elemente von  $\mathcal{U}_{\tilde{P}}$  entsprechen in offensichtlicher Weise den Drehungen des Tangentialraums im Punkt  $\tilde{P}$ . Man kann sagen: Zu jeder euklidischen Drehung des Tangentialraums im Punkt  $\tilde{P}$  gibt es genau eine Bewegung des hyperbolischen Raums, welche den Tangentialraum in  $\tilde{P}$  so dreht. Die Struktur der Fixgruppen ist somit klar. Wir haben mit Hilfe der Quaternionen einen bequemen Kalkül für das Rechnen in einer solchen Gruppe entwickelt.

Im Folgenden wollen wir nun die volle Gruppe aller hyperbolischen Bewegungen in entsprechender Weise gründlich verstehen. Dabei wird es meistens bequemer sein, an die eigentliche Lorentzgruppe zu denken, die Gruppe der orthochronen positivorientierten Isometrien des Minkowskiraums. So, wie wir die dreidimensionale Gruppe  $SU(2, \mathbb{C})\{\pm E_2\}$  zur Durchdringung der Fixgruppen  $\mathcal{U}_{\tilde{P}}$  herangezogen haben, werden wir jetzt die sechsdimensionale Gruppe  $SL(2, \mathbb{C})/\{\pm E_2\}$  ins Spiel bringen.

Zunächst diskutieren wir einen speziellen Typ von Bewegungen, welcher gegeneinander bewegte Beobachter betrifft.

### Satz 1.6.1.

*Es seien  $\mathbf{u}_0, \mathbf{v}_0$  Einheitsvektoren im Vorwärtskegel des Minkowskiraums  $\mathbb{M}$ .  $L_0$  sei die von ihnen aufgespannte Ebene und  $L_0^\perp$  das orthogonale Komplement. Es gibt dann genau eine eigentliche Lorentztransformation  $\varphi(\cdot) \in \mathcal{L}_+^\uparrow$ , welche  $\mathbf{u}_0$  in  $\mathbf{v}_0$  überführt und den Orthogonalraum  $L_0^\perp$  punktweise festlässt.*

**Beweis.** Es sei  $(\mathbf{u}_2, \mathbf{u}_3)$  eine ONB von  $L_0^\perp$ . Es existiert genau ein zu  $\mathbf{u}_0$  orthogonaler Einheitsvektor  $\mathbf{u}_1$ , sodass mit  $\omega = d(\mathbf{u}_0, \mathbf{v}_0)$  gilt  $\mathbf{v}_0 = \mathbf{u}_0 \cdot \cosh \omega + \mathbf{u}_1 \cdot \sinh \omega$ .

Die folgende Abbildung (und keine andere) leistet nun das Verlangte

$$(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \xrightarrow{\varphi} (\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \begin{pmatrix} \cosh \omega & \sinh \omega & 0 & 0 \\ \sinh \omega & \cosh \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

denn  $\mathbf{v}_1 = \mathbf{u}_0 \cdot \sinh \omega + \mathbf{u}_1 \cdot \cosh \omega$  ist der eindeutig bestimmte Vektor, welcher  $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{u}_2, \mathbf{u}_3)$  zu einer gleichgerichteten ONB des Minkowskiraums macht.

**Sprechweise 1.6.2** (Boosts).

Eine eigentliche Lorentztransformation des Minkowskiraums  $(\mathbb{M}, (\cdot\|\cdot))$ , welche eine raumartige Ebene  $L_0^\perp$  punktweise festlässt, heisst bei den Physikern ein boost; ein deutsches Wort scheint nicht eingeführt zu sein. Die ‘Ebene’  $L_0 \subset \mathbb{M}$  oder die dazugehörige Gerade im hyperbolischen Raum  $\mathcal{Hyp}$  nennen wir die Leitlinie des boost. Ist  $\varphi(\cdot)$  ein boost mit der Leitlinie  $L_0$ , dann nennen wir seine Wirkung auf  $\mathcal{Hyp}$  eine hyperbolische Translation mit der Leitlinie  $L_0$ . Die Identität zählen wir auch zu den boosts; für sie kann jede Gerade die Rolle der Leitlinie spielen.

Die Gruppe aller boosts zur Leitlinie  $L_0$  (einschliesslich der Identität) bezeichnen wir mit  $\mathcal{B}_{L_0}$ . Wir nennen die Zahl  $|\omega|$  die Länge des boost, wenn dieser die Punkte auf der Leitlinie um die Distanz  $|\omega|$  verschiebt, wenn also für einen (und damit für jeden) Einheitsvektor  $\mathbf{u} \in L_0$  gilt  $\mathbf{u}\|\varphi(\mathbf{u}) = \cosh \omega$ .

Bemerke: Zur ‘Achse’  $L_0$  und der Länge  $|\omega|$  gibt es zwei boosts; der eine ist invers zum anderen. Es seien  $\varphi_{\omega_1}(\cdot)$  und  $\varphi_{\omega_2}(\cdot)$  boosts mit den Längen  $|\omega_1|$  bzw.  $|\omega_2|$ , die denselben Raum  $L_0^\perp$  festlassen. Die zusammengesetzte Isometrie ist wieder ein boost, welcher  $L_0^\perp$  festlässt, und zwar entweder einer der boosts mit der Länge  $|\omega_1 + \omega_2|$  oder einer der boosts mit der Länge  $|\omega_1 - \omega_2|$ .

Mit diesen Bezeichnungen können wir nun den Satz von oben folgendermaßen aussprechen:

**Satz 1.6.2.** Zu jedem Paar von Punkten  $P \neq Q$  im hyperbolischen Raum gibt es genau einen boost, welcher  $P$  in  $Q$  überführt:  $\varphi_Q^P(\cdot)$ . Seine Leitlinie ist die durch  $P$  und  $Q$  bestimmte Gerade.

Der boost  $\varphi_Q^P(\cdot)$  ist dadurch ausgezeichnet vor anderen Bewegungen, die  $P$  in  $Q$  überführen, dass die Ebenen, welche die Gerade durch  $P$  und  $Q$  enthalten, in sich abgebildet werden. (Bildlich könnte man sagen, dass sie keine Schraubung bewirken.) Ein boost der Länge  $|\omega|$  verschiebt jeden Punkt der Leitlinie in einen Bildpunkt in der Distanz  $|\omega|$ . Die Punkte, die nicht auf der Leitlinie liegen, erfahren eine größere Versetzung, wie der folgende Satz zeigt.

**Satz 1.6.3.** Sei  $\varphi(\cdot)$  ein boost der Länge  $|\omega|$  und  $R$  ein Punkt in der Entfernung  $d = d(R, L_0)$  von der Leitlinie. Dann gilt

$$\cosh d(R, \varphi(R)) = \cosh \omega + (\cosh \omega - 1) \cdot \sinh^2 d(R, L_0).$$

**Beweis.** Wir repräsentieren  $R$  wie im Satz auf Seite 39 durch den Einheitsvektor  $\mathbf{w} = \mathbf{v}_0 + \mathbf{w}'$  mit  $\mathbf{v}_0 \in L_0$ ,  $\mathbf{w}' \perp L_0$ . Wir haben gesehen  $\sinh d(R, L_0) = \|\mathbf{w}'\|$ . Da  $\mathbf{w}$  und  $\varphi(\mathbf{w})$  Einheitsvektoren sind, haben wir

$$\begin{aligned} \cosh d(\mathbf{w}, \varphi(\mathbf{w})) &= (\mathbf{w} \parallel \varphi(\mathbf{w})) = (\mathbf{v}_0 + \mathbf{w}' \parallel \varphi(\mathbf{v}_0) + \mathbf{w}') \\ &= (\mathbf{v}_0 \parallel \varphi(\mathbf{v}_0)) - \|\mathbf{w}'\|^2 = \|\mathbf{v}_0\|^2 \cosh \omega - \|\mathbf{w}'\|^2 \\ &= \cosh \omega \cdot (1 + \|\mathbf{w}'\|^2) - \|\mathbf{w}'\|^2 \\ &= \cosh \omega + (\cosh \omega - 1) \cdot \sinh^2 d. \end{aligned}$$

Bemerkung: Für einen Punkt  $R \notin L_0$  ist die Bahn unter der einparametrischen Gruppe der boosts mit der Leitlinie  $L_0$  keine Gerade.

**Sprechweise.** Wenn  $\chi, \psi \in \mathcal{L}_{\text{Hyp}}^+$  und  $\tilde{\psi} = \chi \circ \psi \circ \chi^{-1}$ , dann heisst  $\tilde{\chi}$  die mittels  $\chi$  zu  $\psi$  konjugierte Bewegung.

**Satz 1.6.4.** Die zu einem boost konjugierten Bewegungen sind boosts. Alle boosts derselben Länge sind zueinander konjugiert.

**Beweis.** Sei  $L_0$  die Leitlinie zum boost  $\varphi_Q^P(\cdot)$  und  $\chi(\cdot)$  eine irgendeine hyperbolische Bewegung. Es sei  $P' = \chi(P)$ ,  $Q' = \chi(Q)$ ,  $L'_0 = \chi(L_0)$ . Die Bewegung  $\chi \circ \varphi_Q^P \circ \chi^{-1}(\cdot)$  ist der boost  $\varphi_{Q'}^{P'}$ ; denn sie bildet  $P'$  auf  $Q'$  ab, und sie lässt die  $\mathbf{w} \perp L'_0$  fest. Für  $\mathbf{w} \in L'_0$  gilt nämlich  $\mathbf{v} = \chi^{-1}(\mathbf{w}) \perp L_0$ ,  $\varphi_Q^P(\mathbf{v}) = \mathbf{v}$ ,  $\chi \circ \varphi_Q^P \circ \chi^{-1}(\mathbf{w}) = \chi(\mathbf{v}) = \mathbf{w}$ .

Auf der anderen Seite: Zu jedem Paar gleichgerichteter ONB  $\mathcal{B} = (\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  und  $\mathcal{B}' = (\mathbf{e}'_0, \mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$  gibt es genau eine Bewegung  $\chi(\cdot)$ , welche  $\mathcal{B}$  in  $\mathcal{B}'$  überführt. Es seien nun  $P, P' \in \mathcal{H}_{\text{Hyp}}$  und  $\mathbf{e}_0, \mathbf{e}'_0$  die repräsentierenden Einheitsvektoren im Vorwärtskegel.  $\mathbf{v}$  und  $\mathbf{v}'$  seien die  $Q$  bzw.  $Q'$  repräsentierenden Einheitsvektoren im Vorwärtskegel. Es existiert genau ein  $\mathbf{e}_1 \perp \mathbf{e}_0$  und genau ein  $\mathbf{e}'_1 \perp \mathbf{e}'_0$ , sodass mit  $\omega = d(P, Q) = d(P', Q')$

$$\mathbf{v} = \cosh \omega \cdot \mathbf{e}_0 + \sinh \omega \cdot \mathbf{e}_1, \quad \mathbf{v}' = \cosh \omega \cdot \mathbf{e}'_0 + \sinh \omega \cdot \mathbf{e}'_1.$$

Wir komplettieren  $(\mathbf{e}_0, \mathbf{e}_1)$  zu einer ONB  $\mathcal{B}$  und  $(\mathbf{e}'_0, \mathbf{e}'_1)$  zu einer gleichgerichteten ONB  $\mathcal{B}'$ . Die Bewegung  $\chi(\cdot)$ , die  $\mathcal{B}$  in  $\mathcal{B}'$  überführt, leistet das Verlangte:  $\chi \circ \varphi_Q^P \circ \chi^{-1}(\cdot) = \varphi_{Q'}^{P'}(\cdot)$ .

**Einparametrische Gruppen von Boosts** Zu jeder ‘Leitlinie’  $L_0$  gehört eine Gruppe von boosts  $\mathcal{B}_{L_0}$ . Eine Parametrisierung  $\mathcal{B}_{L_0} = \{\beta_t : t \in \mathbb{R}\}$  kann dadurch festgelegt werden, dass man von zwei Punkten  $P_0, P_1$  auf der Leitlinie verlangt, dass  $P_0$  in der vorgeschriebenen ‘Zeit’  $\eta$  in  $P_1$  übergeführt werden soll.  $\beta_\eta(P_0) = P_1$ . Wenn die Zeit  $\eta$  für Punkte im Abstand  $|\eta|$  gefordert wird, dann sagen wir, dass die Parametrisierung

normiert ist. Die Punkte auf der Leitlinie werden von einer solchermaßen parametrisierten Gruppe mit der ‘Geschwindigkeit’ = 1 verschoben. In jedem Punkt  $\tilde{P}$  der Leitlinie liefert die Bahn  $\{\beta_t(\tilde{P}) : t \in \mathbb{R}\}$  einen normierten Tangentialvektor. Umgekehrt gibt es zu jedem normierten Tangentialvektor im Punkt  $\tilde{P}$  genau eine normierte parametrisierte Gruppe von boosts mit einer Leitlinie  $L_0 \ni \tilde{P}$ .

Eine normierte einparametrische Gruppe von boosts kann auch festgelegt werden durch ein Paar unendlichferner Punkte  $(P_1^\infty, P_2^\infty)$ . (Man verlangt, dass für  $\omega \rightarrow +\infty$  der zweite Punkt  $P_2^\infty$  angesteuert wird.  $DV((P_1^\infty, P_2^\infty); P, \varphi(P)) = e^{2\omega}$ . Wenn man die unendlichfernen Punkte vertauscht und das Vorzeichen von  $\omega$  umkehrt, dann erhält man denselben boost.

Wenn man einen boost als eine Lorentztransformation versteht, dann kann man bekanntlich von der dazugehörigen Bijektion der Menge  $\Omega$  der unendlichfernen Punkte sprechen. Bei einem boost ( $\neq id$ ) bleiben genau zwei unendlichferne Punkte fest. Wir wissen bereits, dass die Lorentztransformationen durch ihre Wirkung auf  $\Omega$  (das ‘Firmament’) eindeutig bestimmt sind. Die Menge aller Möbiustransformationen  $\{\tilde{\mathbb{C}} \ni z \mapsto e^\omega \cdot z : \omega \in \mathbb{R}\}$  liefert ein Beispiel für eine normierte einparametrische Gruppe von boosts. Und das ist das typische Beispiel; denn alle normierten einparametrischen Gruppen von boosts sind dazu konjugiert.

### Auszeichnung eines Zentrums

Wir wollen im Minkowskiraum  $(\mathbb{M}, (\cdot|\cdot))$  einen Einheitsvektor  $\mathbf{e}_0$  im Vorwärtskegel, und damit einen Punkt  $Z \in \mathcal{Hyp}_{\mathbb{M}}$  als ‘Zentrum’ auszeichnen. Der Orthogonalraum zu  $\mathbf{e}_0$  ist mit dem Tangentialraum  $(\mathcal{T}_Z, \langle \cdot | \cdot \rangle)$  zu identifizieren.

**Sprechweise 1.6.3.** Einen boost  $\varphi(\cdot)$  eines zentrierten Minkowskiraums  $\mathbb{M}^Z$  nennen wir einen zentralen boost, wenn seine Leitlinie  $L_0$  das Zentrum enthält. Eine Drehung von  $\mathbb{M}^Z$ , die das Zentrum festlässt, nennen wir eine zentrale Drehung.

Bemerke: Eine Bewegung ( $\neq id$ ) ist genau dann ein zentraler boost, wenn die Menge der Fixpunkte  $L_0^\perp$  eine Ebene ist, die orthogonal zu  $\mathbf{e}_0$  ist.

Die dreidimensionale Mannigfaltigkeit aller zentralen boosts  $\mathcal{B}_Z$  wir jetzt zusammen mit der dreidimensionalen Gruppe  $\mathcal{U}_{\tilde{P}}$  aller zentralen Drehungen betrachten. Zusammen ergeben sie als eine Art semidirektes Produkt die Gruppe aller eigentlichen Lorentztransformation (oder

Zu jedem  $P \neq Z$  gibt es genau einen zentralen boost, welcher das Zentrum nach  $P$  abbildet. Eine solche Eigenschaft hat in der euklidischen Theorie die Menge aller Parallelverschiebungen. Es wird sich zeigen, dass die zentralen boosts für den hyperbolischen Raum eine ähnliche Rolle spielen wie die Parallelverschiebungen für den euklidischen Raum mit einem ausgezeichneten Zentrum. Die Sache ist u. a. deswegen ertwas komplizierter, weil  $\mathcal{B}_Z$  keine Gruppe ist, geschweige denn ein Vektorraum. Zwar ist die Umkehrabbildung eines zentralen boost ein zentraler boost; das Hintereinanderschalten zweier zentraler boost ist aber nur dann ein zentraler boost, wenn die Leitlinien identisch sind.

Wir bemerken: Ist  $\varphi(\cdot)$  ein zentraler boost und  $\delta(\cdot)$  eine zentrale Drehung, dann ist  $\delta \circ \varphi \circ \delta^{-1}$  ein zentraler boost. Dieser ist genau dann gleich  $\varphi(\cdot)$ , wenn  $\delta(\cdot)$  eine Drehung um die Leitlinie des boosts ist. Allgemein gilt  $\delta \circ \varphi_P^Z \circ \delta^{-1} = \varphi_{\delta(P)}^Z$ .

**Satz 1.6.5.** *Es sei  $\mathbb{M}^Z$  ein Minkowskiraum mit einem ausgezeichneten Zentrum  $Z$ . Zu jeder Bewegung  $\chi(\cdot) \in \mathcal{L}_{\mathbb{M}}^{\uparrow+}$  existieren genau ein zentraler boost und eine zentrale Drehung, sodass gilt*

$$\chi(\cdot) = \delta \circ \varphi(\cdot) \quad \text{mit} \quad \varphi(\cdot) \in \mathcal{B}_Z, \quad \delta(\cdot) \in \mathcal{U}_Z$$

**Beweis.** *Es sei  $P = \chi^{-1}(Z)$  und  $\varphi = \varphi_P^Z$ . Es gilt dann  $\chi \circ \varphi_P^Z(Z) = Z$ . Also gehört  $\delta(\cdot) = \chi \circ \varphi_P^Z(\cdot)$  zur Fixgruppe von  $Z$ , und es gilt  $\chi = \delta \circ \varphi_P^Z = (\delta \circ \varphi_P^Z \circ \delta^{-1}) \circ \delta$ .*

*Die Eindeutigkeit der Faktorisierung liegt auf der Hand.*

**Satz 1.6.6.** *Sind  $\varphi_1, \varphi_2$  zentrale boosts, so gibt es genau eine zentrale Drehung  $\delta$  sodass  $\varphi = \varphi_2 \circ \delta \circ \varphi_1$  ein zentraler boost ist.*

**Beweis.** *Nehmen wir an, wir hätten ein Paar  $\delta, \varphi$ , welches das Verlangte leistet. Es gilt dann  $\delta^{-1} \circ \varphi_2^{-1} \circ \varphi \circ \varphi_1 = (\varphi_1 \circ \varphi_2)$ . Der mit  $\varphi_2^{-1}$  konjugierte zentrale boost  $\psi = \varphi_2^{-1} \circ \varphi \circ \varphi_1$  ergibt sich nach dem vorigen Satz in eindeutiger Weise aus der Bewegung  $\varphi_1 \circ \varphi_2$ .*

*Wir gewinnen also mit dem vorigen Satz  $\delta^{-1}, \psi$  zur Bewegung  $\varphi_1 \circ \varphi_2$ . Das Paar  $\delta, \varphi = \varphi_2 \circ \psi \circ \varphi_1^{-1}$ , und nur dieses Paar leistet das Verlangte. Wir können das gesuchte Paar aber auch auf einem anderen Weg beschreiben. Mit demselben Argument finden wir  $\varphi_2^{-1} \circ \varphi \circ \varphi_1 = \psi = \varphi_1 \circ \varphi \circ \varphi_1^{-1}$ . Für den gesuchten zentralen boost  $\varphi = \varphi_R^Z$  gilt*

$$\varphi_R^Z = \varphi_2 \circ \varphi_1 \circ \varphi_R^Z \circ \varphi_1^{-1} \circ \varphi_2^{-1} = (\varphi_2 \circ \varphi_1) \circ \varphi_R^Z \circ (\varphi_2 \circ \varphi_1)^{-1} = \varphi_{\chi(R)}^{\chi(Z)}, \quad \text{mit} \quad \chi = \varphi_2 \circ \varphi_1.$$

*Dies gibt eine Auskunft über die Leitlinie des gesuchten zentralen boosts  $\varphi = \varphi_R^Z$ :*

*$R$  liegt auf derselben Geraden durchs Zentrum wie der Punkt  $\varphi_2 \circ \varphi_1(Z)$ . Die Drehachse von delta steht senkrecht auf den Leitlinien von  $\varphi_1$  und  $\varphi_2$ .*

### Hinweis auf die Bewegungen eines zentrierten euklidischen Raums

Unsere Faktorisierung der hyperbolischen Bewegungen  $\chi = \delta \circ \varphi$  mit  $\delta \in \mathcal{U}_Z, \varphi \in \mathcal{B}_Z$  entspricht der Faktorisierung der euklidischen Bewegungen eines euklidischen Raums mit einem ausgezeichneten Zentrum. Jede Bewegung kann man dadurch erhalten, dass man zunächst parallel verschiebt und dann um das Zentrum dreht; die Verschiebung  $\mathbf{a}$  und die Drehung  $D$  sind eindeutig bestimmt.

$$\chi(\cdot) \leftrightarrow (D, \mathbf{a}) : \quad \mathbb{E} \ni \mathbf{v} \mapsto D \cdot (\mathbf{v} + \mathbf{a}).$$

Im euklidischen Fall ist die Menge der Verschiebungen ebenso wie die Menge der zentralen Drehungen eine Gruppe. Die Gruppe der euklidischen Bewegungen ist ein sog. semidirektes Produkt.

$$\left. \begin{array}{l} \chi_1(\cdot) \leftrightarrow (D_1, \mathbf{a}_1) \\ \chi_2(\cdot) \leftrightarrow (D_2, \mathbf{a}_2) \end{array} \right\} \quad \chi_2 \circ \chi_1(\cdot) = \chi(\cdot) \leftrightarrow (D_2 \cdot D_1, \mathbf{a}) \quad \text{mit einem wohlbestimmten} \quad \mathbf{a}.$$

Wenn man an den Raum  $\mathbb{R}^3$  der 3-Spalten wie üblich als einen zentrierten euklidischen Raum versteht, dann erscheinen die Drehungen  $D$  als Matrizen aus  $SO(3, \mathbb{R})$  und die Verschiebungen (ebenso wie die Punkte, die ‘Ortsvektoren’) als 3-Spalten. Die Formel für die Komposition ist

$$(D_2, \mathbf{a}_2) \circ (D_1, \mathbf{a}_1) = (D_2 D_1, \mathbf{a}_2^{D_1} + \mathbf{a}_1) \quad \text{mit} \quad \mathbf{a}_2^{D_1} = D_1^{-1} \mathbf{a}_2.$$

Man muss nicht an Spalten und orthogonale Matrizen denken. Man kann auch an den zentrierten euklidischen Raum der spurlosen hermiteschen Matrizen  $(\mathbb{H}_0, -\det(\cdot))$  denken. Hier erscheinen die Drehungen als (bis aufs Vorzeichen eindeutig bestimmte) unimodulare unitäre Matrizen  $\pm U: D: H \mapsto U H U^*$ . Eine Translation erscheint auch hier als eine Addition, und zwar als die Addition eines  $h \in \mathbb{H}_0: (D, h): H \mapsto U(H + h)U^*$ . Die Formel für die Komposition lautet also

$$(U_2, h_2) \circ (U_1, h_1) = (U, h) \quad \text{mit} \quad U = U_2 \cdot U_1, \quad h = h_2^{U_1} + h_1 = U_1^* h_2 U_1 + h_1.$$

Wir werden im Folgenden die entsprechenden Rechengesetze für die eigentlichen Lorentztransformationen des Standard-Minkowski-Raums  $(\mathbb{H}, \det(\cdot))$  ableiten.

### Bewegungen des Standard-Minkowski-Raums

Im Standard-Minkowski-Raum  $(\mathbb{H}, \det(\cdot))$ . wollen wir die Einheitsmatrix (oder genauer die von ihr erzeugte zeitartige Gerade als Zentrum auszeichnen. Gelegentlich denken wir uns den Raum dadurch orientiert, dass wir festlegen, dass  $(E, \sigma_x, \sigma_y, \sigma_z)$  eine positiv orientierte Orthonormalbasis ist.

Die Lorentztransformationen  $\chi(\cdot)$  werden wir nicht durch reelle  $4 \times 4$ -Matrizen ausdrücken, sondern durch die unimodularen  $2 \times 2$ -Matrizen, gemäss dem bereits erwähnten Isomorphismus  $SL(2, \mathbb{C})/\{\pm E_2\} \longleftrightarrow \mathcal{L}_+^\uparrow$ . Der Kalkül stellt sich als eine Erweiterung des Kalküls der Quaternionen dar.

Zuerst erinnern wir an einige **Bezeichnungen und Fakten**, mit denen wir uns schon beim ersten Studium des Standard-Minkowski-Raums befasst haben.

1. Der Orthogonalraum  $[E]^\perp$  ist der Raum der spurlosen hermiteschen Matrizen  $\mathbb{H}_0$ . Diesen Raum verstehen wir als den euklidischen Raum der zentralen Richtungen im hyperbolischen Raum  $\mathcal{Hyp}_{\mathbb{H}}$ .
2. Die  $H \in \mathbb{H}$  sind Matrizen, die man multiplizieren kann; solche Produkte liegen aber nur unter sehr speziellen Umständen in einem der uns interessierenden Räumen.
3. Die spurlosen hermiteschen Matrizen und ganz besonders diejenigen mit der Norm 1 sind nicht nur spezielle Elemente des ‘Grundraums’  $\mathbb{H}$ ; sie spielen auch eine fundamentale Rolle bei der Beschreibung der Bewegungen. (Die Wörter Lorentztransformation und hyperbolische Bewegung benutzen wir synonym.) In diesem Zusammenhang benutzen wir dann auch alternative Bezeichnungen, z. B.  $\Sigma_0 = \mathbb{H}_0$ .

4.  $\Sigma_0^1$  bezeichne die Menge der normierten spurlosen hermiteschen Matrizen. Das sind genau die hermiteschen Matrizen mit den Eigenwerten  $\pm 1$ . Sie haben die Gestalt

$$\tau = a_1 \cdot \sigma_x + a_2 \cdot \sigma_y + a_3 \cdot \sigma_3 = \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}, \quad \text{mit } a_1^2 + a_2^2 + a_3^2 = 1.$$

Sie können, wie bereits bemerkt, als die zentralen Richtungen interpretiert werden.

5. Die spurlosen schieferhermiteschen Matrizen haben die Gestalt  $\frac{1}{i}\sigma$  mit  $\sigma \in \mathbb{H}_0$ . Wenn  $\sigma \in \Sigma_0^1$ , dann kann  $\frac{1}{i}\sigma$  als die Achse einer zentralen Drehung interpretiert werden. Eine solche schieferhermitesche Matrizen entspricht nämlich einem normierten reinen Quaternion.
6. Die Matrizen  $\tau \in \Sigma_0^1$  und die Matrizen  $\frac{1}{i}\sigma \in \frac{1}{i}\Sigma_0^1$  spielen eine wichtige Rolle, wenn wir die Exponentialabbildung ins Spiel bringen.

$$\begin{aligned} \det H = 1, \quad H \text{ pos. def.} &\iff H = \exp(\omega \cdot \tau) \quad \text{mit } \omega \in \mathbb{R} \\ \det U = 1, \quad UU^* = E &\iff U = \exp(\alpha \cdot \frac{1}{i}\sigma). \quad \text{mit } \alpha \in \mathbb{R}/2\pi. \end{aligned}$$

Wir bemerken  $\exp(\omega \cdot \tau) = \cosh \omega \cdot E + \sinh \omega \cdot \tau$ ;  $\exp(\alpha \cdot \frac{1}{i}\sigma) = \cos \alpha \cdot E + \sin \alpha \cdot (\frac{1}{i}\sigma)$ . Hierbei sind  $\omega \cdot \tau$  und  $\alpha \cdot (\frac{1}{i}\sigma)$  eindeutig bestimmt.

7. Wenn  $\tau_1, \tau_2 \in \Sigma_0^1$  mit  $\langle \tau_1 | \tau_2 \rangle = \cos \beta \neq \pm 1$ , dann existiert  $\sigma \in \Sigma_0^1$ , sodass

$$\begin{aligned} \tau_1 \cdot \tau_2 &= \cos \beta \cdot E + \sin \beta \cdot (\frac{1}{i}\sigma), & \tau_2 \cdot \tau_1 &= \cos \beta \cdot E + \sin \beta \cdot (-\frac{1}{i}\sigma), \\ \frac{1}{2}(\tau_1 \cdot \tau_2 + \tau_2 \cdot \tau_1) &= \cos \beta \cdot E, & \frac{1}{2}(\tau_1 \cdot \tau_2 - \tau_2 \cdot \tau_1) &= \sin \beta \cdot (\frac{1}{i}\sigma). \end{aligned}$$

Die spurlosen hermiteschen Matrizen  $H, K \in (\mathbb{H}_0, \langle \cdot | \cdot \rangle)$  sind genau dann orthogonal, wenn sie antikommutieren:  $H \cdot K + K \cdot H = 0$ .

8. Für  $\tau_1, \tau_2, \sigma \in \Sigma_0^1$  wie oben gilt  $\sigma \perp \tau_1, \tau_2$ ; denn wegen  $\tau_1^2 = E$

$$\sin \beta \cdot (\frac{1}{i}\sigma) \cdot \tau_1 = \frac{1}{2}\tau_1 \cdot \tau_2 \cdot \tau_1 - \frac{1}{2}\tau_2 = -\tau_1 \cdot \sin \beta \cdot (\frac{1}{i}\sigma).$$

Ebenso für  $\tau_2$ .

Wir haben oben gesehen, dass jede positivorientierte orthochrone Lorentztransformation des Standard-Minkowski-Raums  $\mathbb{H}$  durch eine (bis auf das Vorzeichen eindeutig bestimmte) unimodulare Matrix  $A \in SL(2, \mathbb{C})$  dargestellt werden kann:  $\chi(\cdot) = A \cdot (\cdot) \cdot A^*$ . Dabei werden die zentralen Drehungen  $\delta(\cdot)$  durch die unitären Matrizen  $U \in SU(2)$  dargestellt.

**Sprechweise 1.6.4.** Die zentrale Drehung

$$\delta(\cdot) = U \cdot (\cdot) \cdot U^* = \exp(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma) \cdot (\cdot) \cdot \exp(-\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma)$$

nennen wir die (zentrale) Drehung mit dem Winkel  $\alpha$  um die Achse  $\sigma \in \Sigma_0^1$ .

Zur Begründung wiederholen wir ein Argument aus dem Abschnitt über Quaternionen: Die Bewegung  $\delta(\cdot) = \exp(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma) \cdot (\cdot) \cdot \exp(-\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma)$  hat die Fixpunkte  $E$  und  $\sigma$ . Für  $\tau \in \Sigma_0^1$  mit  $\tau \perp \sigma$  haben wir  $\tau\sigma\tau^{-1} = -\sigma$  und wir gewinnen  $\rho = \frac{1}{i}\sigma \cdot \tau \in \Sigma_0^1$  mit  $\rho \perp \sigma, \tau$ . Wegen  $-\tau = \frac{1}{i}\sigma \cdot \rho$  gilt

$$\begin{aligned} \tau \cdot \exp(-\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma) &= (\tau \cdot \exp(-\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma) \cdot \tau^{-1}) \cdot \tau = \exp(-\frac{1}{2}\alpha \cdot \tau(\frac{1}{i}\sigma)\tau^{-1}) \cdot \tau = \exp(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma) \cdot \tau. \\ \delta(\tau) &= \exp(\alpha \cdot \frac{1}{i}\sigma) \cdot \tau = \cos \alpha \cdot \tau + \sin \alpha \cdot \rho \\ \delta(\rho) &= \exp(\alpha \cdot \frac{1}{i}\sigma) \cdot \rho = \cos \alpha \cdot \rho - \sin \alpha \cdot \tau. \end{aligned}$$

Und das sind die vertrauten Formeln für eine Drehung in der zu  $\sigma$  orthogonalen Ebene.

**Satz 1.6.7.** *Der zentrale boost, welcher das Zentrum  $E$  in die positivdefinite Matrix  $H = \exp(\omega \cdot \tau)$  ( $\tau \in \Sigma_0^1$ ) überführt, ist*

$$\varphi_H^E(\cdot) = A \cdot (\cdot) \cdot A^* = \exp(\frac{1}{2}\omega \cdot \tau) \cdot (\cdot) \cdot \exp(\frac{1}{2}\omega \cdot \tau)$$

**Beweis.** Wenn  $\sigma \in \Sigma_0^1$  zu  $\tau$  orthogonal ist,  $\sigma \cdot \tau + \tau \cdot \sigma = 0$ ,  $\sigma \cdot \tau \cdot \sigma^{-1} = -\tau$ , dann gilt

$$\sigma \cdot A = \sigma \cdot \exp(\frac{1}{2}\omega \cdot \tau) \cdot \sigma^{-1} \cdot \sigma = \exp(\frac{1}{2}\omega \sigma \cdot \tau \cdot \sigma^{-1}) \cdot \sigma = A^{-1} \cdot \sigma, \quad A \cdot \sigma \cdot A = \sigma.$$

Die Menge der Fixpunkte ist also der zweidimensionale Vektorraum  $\{E, \tau\}^\perp$ .

Die Matrix  $A = A^* = \exp(\frac{1}{2}\omega \cdot \tau)$  kommutiert mit  $E$  und  $K = H^{-1}$ . Es gilt daher  $A \cdot E \cdot A^* = H$ ,  $A \cdot K \cdot A^* = E$ . Der eindeutig bestimmte zentrale boost  $\varphi_H^E(\cdot) = \varphi_E^K(\cdot)$  ist  $A \cdot (\cdot) \cdot A^*$ , wo  $A$  die positivdefinite Quadratwurzel aus  $H$  ist:  $A^2 = AA^* = H$ . Es handelt sich um einen boost der Länge  $|\omega|$ .

Bekanntlich transformiert die Exponentialabbildung jede spurlose Matrix in eine unimodulare Matrix. Für die spurlosen hermiteschen Matrizen gilt darüber hinaus

$$H \in \mathbb{H}_0, \quad H^* = H, \quad t \in \mathbb{R} \implies \exp(\frac{1}{i}tH) \text{ unitär, } \exp(tH) \text{ positivdefinit.}$$

Wir erinnern bei dieser Gelegenheit, dass  $U = \exp(\frac{1}{i}H) = \exp(\frac{\alpha}{2} \cdot \frac{1}{i}\sigma)$  mit  $\frac{1}{2}\alpha = \|H\|$  die Matrix zur zentralen Drehung um die Achse zum Winkel  $\alpha$  ist.

$$SU(2) = \{U : U = \exp(\frac{\alpha}{2} \cdot \frac{1}{i}\sigma) \text{ mit } \alpha \in [0, 2\pi) \text{ und } \sigma \in \Sigma_0^1\}.$$

Der Exponentialausdruck interessiert übrigens für beliebige reelle  $\alpha$ ; man beachte aber  $\exp(\frac{1}{2}(\alpha + 2\pi) \cdot \frac{1}{i}\sigma) = -U$ .

In derselben Weise interessieren wir uns für die Matrizen zu den zentralen boosts.. Wir bezeichnen

$$\mathfrak{B} = \{A : A = \exp(H) \text{ mit } H \in \mathbb{H}_0\} = \{A : A = \exp(\frac{\omega}{2} \cdot \tau) \text{ mit } \omega \geq 0 \text{ und } \tau \in \Sigma_0^1\}.$$

Dies sind, wie wir eben gesehen haben, in der Tat die Matrizen  $A \in \mathfrak{B}$  zu den zentralen boosts. ( $H \mapsto AHA^*$ .) Für  $A = \exp(H) = \exp(\frac{\omega}{2} \cdot \tau)$  ist  $\omega = 2 \cdot \|H\|$  die Länge des zentralen boosts und  $\tau$  seine Richtung.

**Faktorisierung** Zu jeder nichtsingulären Matrix  $B$  existiert bekanntlich eine unitäre Matrix  $U$  und eine positivdefinite Matrix  $A$ , sodass  $B = U \cdot A$ ;  $A$  ist die positivdefinite Quadratwurzel aus  $B^* \cdot B$ . Wenn  $\det B = 1$ , dann  $\det A = 1 = \det U$ . Also

$$B = U \cdot A = \exp\left(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma\right) \cdot \exp\left(\frac{1}{2}\omega \cdot \tau\right)$$

Für die Lorentztransformation  $\chi(\cdot) = B \cdot (\cdot) \cdot B^* = UA \cdot (\cdot) \cdot A^*U^*$  liefert das die Faktorisierung  $\chi(\cdot) = \delta \circ \varphi(\cdot)$ , wo  $\delta(\cdot)$  die zentrale Drehung zu  $U$  ist und  $\varphi(\cdot)$  der zentrale boost zu  $A$ .

Die eindeutige ‘Polarfaktorisierung’  $B = UA$  mit  $U$  unitär,  $A$  positiv definit gibt es für alle nichtsingulären komplexen  $n \times n$ -Matrizen  $B$ . Die Berechnung ist aber im Allgemeinen recht aufwendig. Für den Kalkül der hyperbolischen Bewegungen reicht uns die Berechnung der Faktorisierung in einem übersichtlichen Spezialfall, wie wir sofort sehen werden.

**Satz 1.6.8** (Eine Quaternionendarstellung).

a) Seien  $\sigma, \tau \in \Sigma_0^1$  mit  $\sigma \perp \tau$ ,  $b_0, b_1, b_2 \in \mathbb{R}$  und  $B = b_0E + b_1\frac{1}{i}\sigma + b_2\tau$ .  
Wenn  $b_0^2 + b_1^2 - b_2^2 = 1$ , dann gilt  $\det B = 1$ .

b) Wenn  $B \in SL(2, \mathbb{C})$  mit  $b_0 = \frac{1}{2}\text{trace}B > 0$ , dann existieren eindeutig bestimmte  $b_1, b_2 \geq 0$  und (im Falle  $b_1, b_2 > 0$ ) eindeutig bestimmte  $\sigma, \tau \in \Sigma_0^1$ , sodass  $B = b_0E + b_1\frac{1}{i}\sigma + b_2\tau$ . Es gilt  $\sigma \perp \tau$  und  $b_0^2 + b_1^2 - b_2^2 = 1$ .

c) Wenn  $B = b_0E + b_1\frac{1}{i}\sigma + b_2\tau$  wie oben und  $\tan \frac{1}{2}\alpha = \frac{b_1}{b_0}$ , dann gilt

$$B = \left(\cos \frac{1}{2}\alpha \cdot E + \sin \frac{1}{2}\alpha \cdot \frac{1}{i}\sigma\right) \cdot A, \quad \text{mit } A \text{ positiv definit.}$$

**Beweis.** Wir betrachten hier die Determinante als quadratische Form auf dem komplexen Vektorraum aller  $2 \times 2$ -Matrizen. Die dazugehörige Bilinearform wird ebenso wie im Standard-Minkowski-Raum bezeichnet. Wir erinnern an die Formel  $(B' \| B'') = \frac{1}{2}(\text{trace}B')(\text{trace}B'') - \frac{1}{2}\text{trace}(B' \cdot B'')$ . Für  $\sigma, \tau \in \Sigma_0^1$  gilt  $\sigma \cdot \tau = \cos \beta \cdot E + \sin \beta \cdot \frac{1}{i}\rho$  mit  $\rho \in \Sigma_0^1$ ,  $(\sigma \| \tau) = -\cos \beta$ . Daher

$$\det(b_0E + b_1\frac{1}{i}\sigma + b_2\tau) = (b_0E + b_1\frac{1}{i}\sigma + b_2\tau \| b_0E + b_1\frac{1}{i}\sigma + b_2\tau) = b_0^2 + b_1^2 - b_2^2 + 2b_1b_2\frac{1}{i}(\sigma \| \tau).$$

wegen  $E \perp \sigma$  für alle spurlosen  $B$  und  $(\sigma \| \sigma) = -1 = (\tau \| \tau)$ ,  $(\sigma \| \tau)$  reell für alle  $\sigma, \tau \in \Sigma_0^1$ . Die Behauptung a) folgt.

Sei  $B$  eine Matrix mit  $b_0 = \frac{1}{2}\text{trace}B > 0$ . Für die spurlose Matrix  $N = B - b_0E$  sind  $\frac{i}{2}(N - N^*)$  und  $\frac{1}{2}(N + N^*)$  spurlose hermitesche Matrizen. Sie haben Eigenwerte  $\pm b_1$  bzw.  $\pm b_2$ , also die euklidische Norm  $|b_1|$  bzw.  $|b_2|$ .

$$B = b_0E + \frac{1}{2}(N - N^*) + \frac{1}{2}(N + N^*) = b_0E + b_1\frac{1}{i}\sigma + b_2\tau \quad \text{mit wohlbestimmten } \sigma, \tau \in \Sigma_0^1.$$

Aus  $\det B = 1$  ergibt sich wie oben die Behauptungen  $b_0^2 + b_1^2 - b_2^2 = 1$  und  $\sigma \perp \tau$ .

Zu  $B = b_0E + b_1\frac{1}{i}\sigma + b_2\tau$  konstruieren wir die unimodulare unitäre Matrix

$$U = \frac{b_0}{\sqrt{b_0^2+b_1^2}} \cdot E + \frac{b_1}{\sqrt{b_0^2+b_1^2}} \cdot \frac{1}{i}\sigma = \cos \frac{1}{2}\alpha \cdot E + \sin \frac{1}{2}\alpha \cdot \frac{1}{i}\sigma = \exp\left(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma\right).$$

Wenn wir  $B$  von links mit  $U^*$  multiplizieren, erhalten wir eine positivdefinite Matrix  $A$ .

$$\begin{aligned} U^*B &= \left(\frac{b_0}{\sqrt{b_0^2+b_1^2}} \cdot E - \frac{b_1}{\sqrt{b_0^2+b_1^2}} \cdot \frac{1}{i}\sigma\right) \cdot (b_0E + b_1\frac{1}{i}\sigma + b_2\tau) \\ &= \frac{b_0^2+b_1^2}{\sqrt{b_0^2+b_1^2}} \cdot E + \left(\frac{b_0}{\sqrt{b_0^2+b_1^2}} \cdot E - \frac{b_1}{\sqrt{b_0^2+b_1^2}} \cdot \frac{1}{i}\sigma\right) \cdot b_2\tau \\ &= \sqrt{b_0^2+b_1^2} \cdot E + \frac{b_2}{\sqrt{b_0^2+b_1^2}} \cdot (b_0\tau - b_1\frac{1}{i}\sigma\tau) \\ &= \sqrt{1+b_2^2} \cdot E + b_2 \cdot (\cos \frac{1}{2}\alpha \cdot \tau + \sin \frac{1}{2}\alpha \cdot \rho) \in \mathfrak{P}. \end{aligned}$$

denn wegen  $\sigma \perp \tau$  ist  $\rho = -\frac{1}{i}\sigma\tau = \frac{1}{i}\tau\sigma$  hermitisch und es gilt  $U^*B = \exp\left(\frac{\omega}{2} \cdot \tau'\right)$ .

$$B = U \cdot A = \exp\left(\frac{1}{2}\alpha \cdot \frac{1}{i}\sigma\right) \cdot \exp\left(\frac{\omega}{2} \cdot \tau'\right)$$

$$\text{mit } \sinh \frac{\omega}{2} = b_2, \quad \tau' = \cos \frac{1}{2}\alpha \cdot \tau + \sin \frac{1}{2}\alpha \cdot \rho.$$

Die hier gelungene Berechnung der Faktorisierung funktioniert zwar (in dieser Direktheit) nur für unimodulare Matrizen mit reeller Spur. Dies genügt aber für unsere Zwecke, wie wir sehen werden. (siehe aber auch den Nachtrag zur Faktorisierung). Wir wiederholen zunächst in etwas veränderter Notation die Methode der Quaternionen für den

### Kalkül der zentralen Drehungen

Eine Matrix  $U \in SU(2)$  hat Eigenwerte vom Betrag 1 mit dem Produkt = 1, sagen wir  $\exp(\pm i\alpha/2)$ ; die Spur ist  $2b_0 = 2 \cos(\alpha/2)$ .  $U - b_0E$  ist eine schieferhermitische spurlose Matrix mit den Eigenwerten  $b_1 = \pm i \sin(\alpha/2)$ . Wir haben  $U = b_0E + b_1 \cdot \frac{1}{i}\sigma$  mit  $\sigma \in \Sigma_0^1$ , als Spezialfall des obigen Satzes mit  $b_2 = 0$ . Es ergibt sich also  $U = \exp(\alpha/2 \cdot \frac{1}{i}\sigma)$ , und das ist die darstellende Matrix der zentralen Drehung mit dem Winkel  $\alpha$  um die Achse  $\sigma$ .

Wir berechnen das Produkt zweier solcher Matrizen  $U_j \in SU(2)$ :

$$U_j = \exp\left(\frac{\alpha_j}{2} \cdot \frac{1}{i}\sigma_j\right) = c_j \cdot E + s_j \cdot \frac{1}{i}\sigma_j = c_j \cdot \left(E + t_j \cdot \frac{1}{i}\sigma_j\right) \quad \text{mit } t_j = \tan \frac{\alpha_j}{2}.$$

Bekantlich gilt  $\sigma_2 \cdot \sigma_1 = \cos \beta \cdot E + \sin \beta \cdot \frac{1}{i}\rho$ ,  $\sigma_1, \sigma_2 \perp \rho \in \Sigma_0^1$ , wo  $\beta$  als der Winkel zwischen den beiden Achsen zu deuten ist

$$\begin{aligned} U_2 \cdot U_1 &= c_2c_1 \cdot \left(E + t_2 \cdot \frac{1}{i}\sigma_2\right) \cdot \left(E + t_1 \cdot \frac{1}{i}\sigma_1\right) \\ &= c_2c_1 \cdot \left((1 - t_1t_2 \cos \beta) \cdot E + \frac{1}{i}(t_1\sigma_1 + t_2\sigma_2 + t_1t_2 \sin \beta\rho)\right) = \exp\left(\frac{\alpha}{2} \cdot \frac{1}{i}\sigma\right). \end{aligned}$$

Aus dem Koeffizienten von  $E$  ist der Drehwinkel  $\alpha$  abzulesen:

$$\cos \alpha/2 = c_2c_1 \cdot (1 - t_1t_2 \cos \beta) = c_2c_1 - s_1s_2 \cos \beta.$$

Die Linearkombination  $t_1\sigma_1 + t_2\sigma_2 + t_1t_2 \sin \beta\rho$  ist bis auf den Normierungsfaktor die Drehachse der resultierenden Drehung.

### Die Kombination der zentralen Boosts

Zu jeder Bewegung  $\chi(\cdot)$  des hyperbolischen Raums  $\mathcal{Hyp}_{\mathbb{H}}$  gibt es eine unimodulare Matrix  $B$  (bis auf das Vorzeichen eindeutig), sodass  $\chi(H) = BHB^*$ . Das Hintereinanderschalten spiegelt sich in der Matrizenmultiplikation  $\chi_2 \circ \chi_1 \Leftrightarrow B_2B_1$ . Nun besitzt jedes  $B \in SL(2, \mathbb{C})$  eine eindeutige Faktorisierung  $B = UA$  mit  $U \in SU(2)$ ,  $A$  positiv definit; und das bedeutet, dass man jede Bewegung gewinnen kann, indem man auf eine wohlbestimmten zentralen boost eine wohlbestimmte zentrale Drehung folgen lässt; denn zu den zentralen boosts gehören die unimodularen (positiv)-definiten Matrizen und zu den zentralen Drehungen die speziellen unitären Matrizen.

Bei den Bewegungen eines euklidischen Raums (mit einem ausgezeichneten Zentrum) hatten wir eine ähnliche Situation. Man bekommt jede Bewegung dadurch, dass man eine wohlbestimmte zentrale Drehung auf eine wohlbestimmte Translation folgen lässt.

Die Situation bei den hyperbolischen Bewegungen ist deswegen etwas komplizierter, weil die Menge der zentralen boosts keine Untergruppe der Bewegungsgruppe ist. Das Produkt zweier positivdefiniter Matrizen ist nicht notwendigerweise hermitisch. Es gibt aber genau eine positivdefinite Matrix  $A = \Pi(A_2A_1)$ , sodass  $A_2 \cdot A_2 = B = U \cdot A$  mit  $U \in SU(2)$ . Wir werden Formeln für  $U$  und  $A = \Pi(A_2A_1)$  herleiten, von denen man sagen könnte, dass sie sich an den Quaternionenkalkül anschliessen. Man kann jedenfalls aus ihnen leicht die Richtung und die Länge des zugehörigen boost ablesen, und auch Achse und Drehwinkel der ‘korrigierenden Drehung’  $U$ . Die Formeln zeigen auch, dass für boosts kleiner Länge das Hintereinanderschalten in der Nähe des Zentrums approximiert werden kann durch das (der Vektorrechnung folgende) Hintereinanderschalten von Translationen.

Für die unimodularen positivdefiniten  $2 \times 2$ -Matrizen gibt es mehrere bemerkenswerte (eindeutig bestimmte!) Darstellungsweisen, z. B.

$$A = \cosh \frac{\omega}{2} \cdot E + \sinh \frac{\omega}{2} \cdot \tau \quad \text{mit } \omega \in \mathbb{R}^+ \text{ und } \tau \in \Sigma_0^1.$$

Für die Matrizen  $A_1, A_2$ , für die wir jetzt  $\Pi(A_2A_1)$  berechnen, benützen wir Abkürzungen

$$A_j = c_j \cdot E + s_j \cdot \tau_j = c_j \cdot (E + t_j \cdot \tau_j) \quad \text{mit } t_j = \tanh \frac{\omega_j}{2}.$$

Es gilt  $\tau_2 \cdot \tau_1 = \cos \beta \cdot E + \sin \beta \cdot \frac{1}{i}\sigma$  mit wohlbestimmten  $\sin \beta \geq 0, \sigma \in \Sigma_0^1$ , und  $\tau_1 \cdot \tau_2 = \cos \beta \cdot E - \sin \beta \cdot \frac{1}{i}\sigma$ . Dies ergibt

$$\begin{aligned} A_2 \cdot A_1 &= (c_2 \cdot E + s_2 \cdot \tau_2) \cdot (c_1 \cdot E + s_1 \cdot \tau_1) \\ &= (c_2c_1 + s_2s_1 \cos \beta) \cdot E + s_2s_1 \cdot \frac{1}{i}\sigma + (c_2s_1\tau_1 + c_2s_1\tau_1) \\ &= b_0 \cdot E + b_1 \cdot \frac{1}{i}\sigma + c_1c_2 \cdot (t_1\tau_1 + t_2\tau_2) \\ &= b_0 \cdot E + b_1 \cdot \frac{1}{i}\sigma + b_2 \cdot \tau, \end{aligned}$$

wo  $\tau$  die Richtung des ‘resultierenden’ Tangentialvektors  $t_1\tau_1 + t_2\tau_2$  ist und

$$b_2 = c_1c_2 \cdot \|t_1\tau_1 + t_2\tau_2\| = c_1c_2 \cdot \sqrt{t_1^2 + t_2^2 + 2t_1t_2 \cos \beta}.$$

Die positive Zahl  $b_0$  kennen wir schon aus dem hyperbolischen Cosinussatz. Da auch  $b_1$  und  $b_2$  nichtnegativ sind, sind wir in der Situation des obigen Satzes. Es gilt mit einer wohlbestimmten positivdefiniten Matrix  $A = \Pi(A_2 \cdot A_1)$

$$A_2 \cdot A_2 = \left( \frac{b_0}{\sqrt{b_0^2 + b_1^2}} \cdot E + \frac{b_1}{\sqrt{b_0^2 + b_1^2}} \cdot \frac{1}{i} \sigma \right) \cdot A = \exp\left(\frac{1}{2} \alpha \cdot \frac{1}{i} \sigma\right) \cdot A \quad \text{mit} \quad \tan \frac{\alpha}{2} = \frac{b_1}{b_0} = \frac{t_1 t_2 \sin \beta}{1 + t_1 t_2 \cos \beta}.$$

Die Rechnung von oben liefert uns auch die gewünschte Darstellung

$$A = \Pi(A_2 \cdot A_1) = \cosh \frac{\omega}{2} \cdot E + \sinh \frac{\omega}{2} \cdot \tau',$$

$$\text{wo} \quad \sinh \frac{\omega}{2} = c_1 c_2 \sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos \beta} \quad \text{und} \quad \tau' = \cos \frac{\alpha}{2} \cdot \tau + \sin \frac{\alpha}{2} \cdot \tau \cdot \sigma.$$

Wir können nun auch irgendwelche Bewegungen hintereinanderschalten, deren Matrizen uns in faktorisierten Form gegeben sind:  $B_2 = U_2 A_2$   $B_1 = U_1 A_1$ . Das Produkt schreiben wir in der Form  $B_2 \cdot B_1 = U_2 A_2 \cdot U_1 A_1 = U_2 U_1 \cdot A_2^{U_1} \cdot A_1$  mit der positivdefiniten konjugierten Matrix  $A_2^{U_1} = U_1^{-1} A_2 U_1$ . Die Operation  $\Pi(\cdot)$  liefert uns die Faktorisierung  $A_2^{U_1} A_1 = U \cdot A$  und damit die Polarfaktorisierung  $B_2 \cdot B_1 = U_2 U_1 U \cdot A$ . — Wir stoßen hier also auf einen etwas verallgemeinerten Begriff eines semidirekten Produkts.

Näherung für zentrale boosts kleiner Länge: Wenn die Längen  $\omega_j$  und damit die Zahlen  $s_j, t_j$  von erster Ordnung klein sind, dann ist  $\tan \frac{\alpha}{2}$  von zweiter Ordnung klein. Die Korrektur  $\exp(\frac{1}{2} \alpha \cdot \frac{1}{i} \sigma)$  fällt nicht ins Gewicht. Ausserdem gilt in erster Näherung  $c_j \sim 1$ .

$$\begin{aligned} \Pi(A_2 \cdot A_1) &\sim A_2 \cdot A_1 \sim (E + t_2 \tau_2) \cdot (E + t_1 \tau_1) \sim E + t \tau \\ \text{mit} \quad \tau &= t_2 \tau_2 + t_1 \tau_1, \quad t = \sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos \beta}. \end{aligned}$$

Die zentralen boosts werden in erster Näherung wie die Vektoren im zentralen Tangentialraum kombiniert: die Richtungen und die Längen werden wie im euklidischen Vektorraum kombiniert.

### Die Algebra des Raums $\mathbb{C}^{2 \times 2}$

Bisher hatten sich unsere Rechnungen mit den komplexen  $2 \times 2$ -Matrizen an geometrischen Themen orientiert. Man kann aber natürlich die Sache auch rein algebraisch betrachten. Es kommen hier mehrere Strukturen zusammen, die dann in der höheren Algebra in ganz verschiedenen Richtungen verallgemeinert werden. Wir beginnen mit der Vektorraum.

$\mathbb{C}^{2 \times 2}$  ist ein vierdimensionaler komplexer und damit ein achtdimensionaler reeller Vektorraum. Eine erste interessante Projektion ist die Abbildung

$$B \mapsto b_0 \cdot E \quad \text{mit} \quad b_0 = b_0(B) = \frac{1}{2} \text{trace} B.$$

Der Kern dieser Projektion ist der Raum der spurlosen Matrizen.

Eine zweite interessante Projektion des reellen Vektorraums  $\mathbb{C}^{2 \times 2}$  ergibt sich aus der hermiteschen Konjugation, (welche eine Involution des reellen Vektorraums ist)

$$B \mapsto \Re(B) = \frac{1}{2}(B + B^*)$$

Sie ordnet der Matrix  $B$  den hermiteschen Anteil zu. Das Bild wollen wir wie oben mit  $\mathbb{H}$  bezeichnen. Die komplementäre Projektion ordnet der Matrix  $B$  ihren schieferhermiteschen Anteil zu; das Bild bezeichnen wir mit  $i \cdot \mathbb{H}$ . Die zu diesen Projektionen gehörende direkte Zerlegung ist also

$$\mathbb{C}^{2 \times 2} = \mathbb{H} \oplus i \cdot \mathbb{H}; \quad B = \Re(B) + i\Re(\frac{1}{i}B) = \frac{1}{2}(B + B^*) + \frac{1}{2}(B - B^*).$$

Die beiden Projektionen sind verträglich, liefern also zusammen eine direkte Zerlegung in zwei eindimensionale und zwei dreidimensionale Teilvektorräume

$$\mathbb{C}^{2 \times 2} = [E] \oplus [iE] \oplus \mathbb{H}_0 \oplus i \cdot \mathbb{H}_0.$$

Die Determinante  $\det(\cdot)$  ist eine quadratische Form

$$\det B = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = \frac{1}{2}(a+d)^2 - \frac{1}{2}(a^2 + bc + bc + d^2) = \frac{1}{2}(\text{trace} B)^2 - \frac{1}{2}(\text{trace} B^2).$$

Die Einschränkung auf den vierdimensionalen Vektorraum  $\mathbb{H}$  der hermiteschen Matrizen macht diesen zu einem Minkowskiraum  $(\mathbb{H}, (\cdot|\cdot))$ .

$$(H_1|H_2) = \frac{1}{2}(\text{trace} H_1)(\text{trace} H_2) - \frac{1}{2}\text{trace}(H_1 H_2).$$

Daraus ergibt sich eine euklidische Struktur auf dem dreidimensionalen Vektorraum  $\mathbb{H}_0$  der spurlosen hermiteschen Matrizen:  $\langle \cdot | \cdot \rangle = -(\cdot|\cdot)$ .

$$\|H\|^2 = -\det \begin{pmatrix} \alpha & b \\ \bar{b} & -\alpha \end{pmatrix} = \alpha^2 + |b|^2, \quad \text{wobei } \alpha \in \mathbb{R}, b \in \mathbb{C},$$

Wir kommen jetzt zur Matrizenmultiplikation: Die Determinante ist bekanntlich ein multiplikatives Funktional mit  $\det B^* = (\det B)^*$ . Entscheidend ist nun die Tatsache, dass  $\mathbb{G} = [E] \oplus i \cdot \mathbb{H}_0$  ein multiplikativ abgeschlossener Teilraum ist. (— Es handelt sich um eine Matrix-Realisierung des Raums der Quaternionen.) Die Beweis ergibt sich aus der Feststellung, dass für jedes  $H \in \mathbb{H}_0$  das Quadrat ein Vielfaches der Einheitsmatrix ist:  $H^2 = \|H\|^2 \cdot E$ . Eine spurlose hermitesche Matrix hat nämlich reelle Eigenwerte  $\pm\lambda$  (und daher die Determinante  $-\lambda^2$ .) Das ergibt  $H^2 = \lambda^2 \cdot E = \langle H|H \rangle \cdot E$ . Es folgt

$$\frac{1}{2}(HK + KH) = \frac{1}{2}((H + K)^2 - H^2 - K^2) = \langle H|K \rangle \cdot E \quad \text{für } H, K \in \mathbb{H}_0.$$

Die Matrizen  $H, K \in \mathbb{H}_0$  sind genau dann orthogonal  $H \perp K$ , wenn sie antikommutieren. Für die spurlose hermitesche Matrix  $L = \frac{1}{2i}(HK - KH) \in \mathbb{H}_0$  gilt  $L \perp H$ , ( und ebenso  $\perp K$ ); denn  $2i(LH + HL) = (HK - KH)H - H(HK - KH) = 0$ . Weiter gilt  $\|L\|^2 \cdot E = LL^* = (HK - \langle H|K \rangle \cdot E) \cdot (KH - \langle H|K \rangle \cdot E) = HKKH - (HK + KH)\langle H|K \rangle + \langle H|K \rangle^2 \cdot E$ . und das ergibt wegen  $HK = \frac{1}{2}(HK + KH) + \frac{1}{2}(HK - KH)$  für alle  $H, K \in \mathbb{H}_0$

$$HK = \langle H|K \rangle \cdot E + i \cdot L \quad \text{mit } L \perp H, K, \quad \|L\|^2 = \|H\|^2 \|K\|^2 - \langle H|K \rangle^2.$$

Dies ist die moderne Version der Lehre von sog. Vektorprodukt im euklidischen Raum.

Manchmal ist es angezeigt, im Hilbertraum  $\mathbb{G} = [E] \oplus i \cdot \mathbb{H}_0$  eine ONB einzuführen wie z. B.  $\{E, \frac{1}{i}\sigma_x, \frac{1}{i}\sigma_y, \frac{1}{i}\sigma_z\}$ .

Als ein Beispiel für das Rechnen im System  $\mathbb{C}^{2 \times 2}$  berechnen wir die (eindeutig bestimmte!) Polarzerlegung  $B = U \cdot A$  einer beliebigen nichtsingulären Matrix. Es genügt offenbar, den Fall  $\det B = 1$  zu betrachten—dies ist der Fall, wo die unitäre Matrix die Determinante = 1 hat. Wir wollen hier aber nicht so wie oben voraussetzen, dass  $\text{trace} B = r + is$  reell ist.

**Satz 1.6.9.**

Gegeben sei  $B \in SL(2, \mathbb{C})$ . Es existieren dann  $a_0 > 0$ , eine spurlose hermitische Matrix  $K \in \mathbb{H}_0$  und eine unitäre Matrix  $U$ , sodass gilt  $B = U \cdot (a_0 \cdot E + K)$ .

**Beweis.** Wenn wir  $B$  mit  $\det B = 1$  gemäß  $\mathbb{C}^{2 \times 2} = \mathbb{H} \oplus i \cdot \mathbb{H}$  zerlegen, dann erhalten wir

$$B = (r + is)E + iB_1 + B_2 \quad \text{mit} \quad B_1, B_2 \in \mathbb{H}_0, \quad \langle B_1 | B_2 \rangle = r \cdot s.$$

In der Tat gilt wegen  $(E || E) = 1, \quad (E || B_1) = 0 = (E || B_2)$

$$\begin{aligned} 1 = \det B &= (B || B) = (r + is)^2 - (B_1 || B_1) + (B_2 || B_2) + i(B_1 || B_2) + i(B_2 || B_1) \\ &= r^2 - s^2 + 2irs + \|B_1\|^2 - \|B_2\|^2 - 2i \cdot \langle B_1 | B_2 \rangle \\ &= r^2 - s^2 + b_1^2 - b_2^2 + 2irs - 2i \cdot \langle B_1 | B_2 \rangle. \end{aligned}$$

Wir definieren  $a_0^2 = r^2 + b_1^2 = 1 + s^2 + b_2^2$  und machen den Ansatz  $A = a_0 \cdot E + K$ . Die Matrix  $K \in \mathbb{H}_0$  gewinnen wir aus dem Ansatz  $B^* \cdot B = A^* \cdot A = a_0^2 \cdot E + \|K\|^2 \cdot E + 2a_0 \cdot K$ .

$$\begin{aligned} B^* \cdot B &= ((r - is)E - iB_1 + B_2) \cdot ((r + is)E + iB_1 + B_2) \\ &= (r^2 + s^2)E + (-i(r - is) + i(r + is))B_1 + (r + r)B_2 - i \cdot (B_1 B_2 - B_2 B_1) \\ &= (r^2 + s^2 + b_1^2 + b_2^2) \cdot E + 2s \cdot B_1 + 2r \cdot B_2 + 2D = (2a_0^2 - 1) \cdot E + 2a_0 \cdot K. \end{aligned}$$

wo  $D \in \mathbb{H}_0$  mit  $D \perp B_1, B_2$  bestimmt ist aus  $B_1 \cdot B_2 = \langle B_1 | B_2 \rangle \cdot E + i \cdot D$ .

Die Matrix  $U = \frac{r}{a_0} \cdot E + \frac{i}{a_0} \cdot B_1$  ist unitär; denn  $U \cdot U^* = \left(\frac{r}{a_0}\right)^2 \cdot E + \left(\frac{b_1}{a_0}\right)^2 \cdot E = E$ .

Wir rechnen nun nach:

$$\begin{aligned} U^* \cdot B &= \left(\frac{r}{a_0} \cdot E - \frac{i}{a_0} \cdot B_1\right) \cdot ((r + is) \cdot E + i \cdot B_1 + B_2) \\ &= \frac{1}{a_0} \cdot \left(r(r + is)E - i(r + is)B_1 + irB_1 + rB_2 + B_1^2 - iB_1B_2\right) \\ &= \frac{1}{a_0} \cdot (r(r + is) + b_1^2 - irs) \cdot E + \frac{1}{a_0} \cdot (sB_1 + rB_2 + D) \\ &= a_0 \cdot E + K = A. \end{aligned}$$

## 1.7 Der Raum der Kreise

Im Folgenden benützen wir einige weitergehende Begriffsbildungen, die wir im abschließenden Anhang erläutern werden. Die Bahn einer Gruppenwirkung ist ein solcher Begriff. Ein anderes wichtiges Konzept ist das Konzept der Pullback-Abbildung.

Die Möbiustransformationen bilden Kreise in  $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  in Kreise ab, (wo die Geraden als Kreise durch den unendlichfernen Punkt zu betrachten sind). Der Raum der Kreise ist eine drei-dimensionale Mannigfaltigkeit, auf der die Gruppe  $SL(2, \mathbb{C})/\mathbb{R}$  wirkt. Er weist gewisse Analogien mit dem hyperbolischen Raum auf. Er gilt allerdings aus synthetischer Sicht als etwas unübersichtlicher, weil die Analogien zur wohlbekannteren Geometrie des euklidischen Raums nicht auf der Hand liegen.

Wenn man Kreise als Nullstellengebilde von Funktionen beschreibt, dann kommen hermitesche Formen ins Spiel. Die hermiteschen Formen spielen eine zentrale Rolle in der Hilbertraumgeometrie, wie wir später sehen werden. Hier geht es nun aber um anderweitig geometrisch inspirierte Fragen, und zwar im besonders übersichtlichen komplex-zweidimensionalen Fall. Es wird nochmals der Minkowskiraum ins Spiel kommen.

Wir schicken einige algebraische Vorbemerkungen voraus.

### Lemma.

Es sei  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$  eine Basis eines komplexen Vektorraums  $V$ . Jede hermitesche Matrix  $H = (h_{jk})$ ,  $(H^* = H)$  definiert eine hermitesche Form  $\mathbf{h}(\cdot)$  auf  $V$ :

$$\mathbf{h}(v) = \mathbf{h}\left(\sum_j \mathbf{u}_j \cdot a^j\right) = \sum_{jk} \bar{a}^j h_{jk} a^k = (\bar{a}^1, \dots, \bar{a}^n) \cdot H \cdot \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$$

Jede hermitesche Form  $\mathbf{h}(\cdot)$  wird durch genau eine hermitesche Matrix  $H$  beschrieben.

**Notation.** Wenn wir in  $V$  eine Basis eingeführt haben, dann repräsentieren wir die Vektoren durch Spalten. Die Linearformen repräsentieren wir (der dualen Basis entsprechend) durch Zeilen. Nicht nur die hermiteschen Formen auf  $V$ , sondern auch die auf dem Dualraum  $V^*$  repräsentieren wir durch hermitesche Matrizen.

Den  $n^2$ -dimensionalen reellen Vektorraum der hermiteschen Formen auf  $V$  bezeichnen wir mit  $\mathbb{M}_V^*$ . Der  $n^2$ -dimensionalen reellen Vektorraum der hermiteschen Formen auf  $V^*$  kann in natürlicher Weise als sein Dualraum verstanden werden. Wenn er im Folgenden (eher selten) vorkommen wird, dann bezeichnen wir ihn mit  $\mathbb{M}_V$ . In dieser natürlichen Auffassung ergeben die darstellenden Matrizen  $L \in \mathbb{M}$ ,  $H \in \mathbb{M}^*$  den Skalar  $(L, H) = \text{trace}(L \cdot H)$ .

Wenn  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$  eine Basis von  $V$  ist, dann verstehen wir die Elemente der dualen Basis  $(\mathbf{z}^1, \dots, \mathbf{z}^n)$  auch als komplexwertige Funktionen auf  $V$ . Die (reellwertigen!) hermiteschen Formen  $\mathbf{h} \in \mathbb{M}^*$  notieren wir dementsprechend

$$\mathbf{h}(\cdot) = \sum_{jk} \bar{\mathbf{z}}^j(\cdot) \cdot h_{jk} \cdot \mathbf{z}^k(\cdot). \quad \text{als Funktion auf } V.$$

Wohlbekannt aus der Linearen Algebra ist der Begriff des Eigenwerts eines Endomorphismus und auch der Begriff des Eigenwerts einer hermiteschen Form in einem unitären Raum. Wir erinnern an den berühmten Spektralsatz.

**Lemma.** *Es sei  $(V, \|\cdot\|)$  ein  $n$ -dimensionaler unitärer Raum. Zu jeder hermiteschen Form  $\mathbf{h}(\cdot)$  existieren reelle Zahlen  $\lambda_1 \geq \dots \geq \lambda_n$  und eine ONB  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ , sodass*

$$\mathbf{h}(\cdot) = \lambda_1 \cdot |\mathbf{z}^1|^2 + \dots + \lambda_n \cdot |\mathbf{z}^n|^2.$$

*Die absteigende Folge der ‘Eigenwerte’ von  $\mathbf{h}(\cdot)$  ist eindeutig bestimmt. Wenn man zu einer anderen Hilbertraumnorm übergeht, dann wird sich i. Allg. die Folge der ‘Eigenwerte’ ändern. Auch das Produkt der Eigenwerte wird sich i. Allg. ändern; es ändert sich aber für alle  $\mathbf{h}(\cdot)$  um denselben Faktor  $> 0$ .*

**Beweis.** *Das Produkt der Eigenwerte ist die Determinante der darstellenden Matrix  $H$ . Wenn man zu einer anderen Basis übergeht, (z. B. zu einer ONB bzgl. der neuen Hilbertraum-Norm), dann wird aus der darstellenden Matrix  $H$  die Matrix  $B^*HB$ , wo  $B$  die Matrix der Basistransformation ist. Wir bemerken, dass dabei positivdefinite Matrizen zu positivdefiniten Matrizen werden, positiv semidefinite Matrizen werden zu positivsemidefinten Matrizen, usw.*

Wir bemerken

1. Im Fall  $n = 2$  zerfällt der Raum der hermiteschen Formen

$$\mathbb{M}_V^* = (M_{++}^* \cup M_{+0}^*) \cup M_{+-}^* \cup (M_{0-}^* \cup M_{--}^*).$$

Und diese disjunkte Zerlegung ist unabhängig von der Wahl einer Basis. Uns wird im Folgenden hauptsächlich der Raum  $M_{+-}^*$  der indefiniten Formen interessieren.

2. Die darstellende Matrix für  $\mathbf{h}(\cdot) \in M_{+-}^*$  hat bei jeder Basiswahl eine negative Determinante; die Eigenwerte haben verschiedenes Vorzeichen.
3. Wenn wir im Folgenden nur Basistransformationen mit der Determinante 1 in Betracht ziehen, dann kommt jeder hermiteschen Form auf  $V$  eine ‘Determinante’ zu. Sie ist da offenbar eine quadratische Form, die  $\mathbb{M}^*$  zu einem Minkowskiraum macht.
4. Wenn wir in  $V$  eine Hilbertraum  $\|\cdot\|$  auszeichnen, dann bedeutet das, dass wir im Minkowskiraum  $\mathbb{M}_{(V, \|\cdot\|)}^*$  einen zeitartigen Einheitsvektor (nämlich  $\|\cdot\|^2$ ) als ‘Zentrum’ auszeichnen. Die darstellende Matrix ist die Einheitsmatrix  $E$ .
5. Die Menge der spurlosen hermiteschen Matrizen ist der Orthogonalraum (im Sinne der Minkowski-Struktur von  $\mathbb{M}^*$ ). Zu jedem  $H$  existiert genau eine Zahl  $a = \frac{1}{2}\text{trace}H$ , sodass  $H - a \cdot E$  spurlos ist.  $\mathbf{h}(\cdot) - a \cdot \|\cdot\|^2$  hat Eigenwerte  $\pm\lambda$ , wo  $-\lambda^2 = \det(H - a \cdot E)$ ,  $q(\mathbf{h}) = \det H = (a - \lambda)(a + \lambda)$ .

**Satz.** Wenn im zweidimensionalen Raum  $V$  eine positivdefinite hermitesche Form, also eine Hilbertraum ausgezeichnet ist, dann ergeben sich bequeme Formeln für die quadratische Form  $q(\cdot)$  auf  $\mathbb{M}_V^*$  und die daraus abgeleitete symmetrische Bilinearform  $(\cdot \parallel \cdot)$ , nämlich

$$\begin{aligned} (H \parallel K) &= \frac{1}{2} \text{trace} H \cdot \text{trace} K - \frac{1}{2} \text{trace}(H \cdot K) \\ &= ab + (H_0 \parallel K_0) = ab - \langle H_0 | K_0 \rangle \\ &= \frac{1}{2} \text{trace}((2a \cdot E - H) \cdot K), \quad \text{wenn} \\ a &= \frac{1}{2} \text{trace} H, \quad b = \frac{1}{2} \text{trace} K, \quad H = a \cdot E + H_0, \quad K = b \cdot E + K_0. \end{aligned}$$

**Beweis.** Die Beweise sind einfache Rechnungen, die wir hier nicht vorführen wollen. Wir bemerken aber

$$H = \begin{pmatrix} h & \bar{c} \\ c & k \end{pmatrix} \implies 2a \cdot E - H = \begin{pmatrix} k & -\bar{c} \\ -c & h \end{pmatrix}.$$

**Lemma.** Es seien  $V$  und  $W$  komplexe Vektorräume und  $\beta(\cdot) : V \rightarrow W$  eine lineare Abbildung. Der Pullback  $\beta^*(\cdot)$  bildet dann hermitesche Form auf  $W$  in hermitesche Formen auf  $V$  ab.

**Beweis.** Wir erinnern noch einmal an einige Eigenschaften der hermiteschen Formen auf einem  $n$ -dimensionalen komplexen Vektorraum. Die einfachsten hermiteschen Formen  $\mathbf{h}(\cdot)$  ergeben sich als Absolutquadrat einer Linearform  $\mathbf{h}(\cdot) = |\ell(\cdot)|^2$  – man nennt sie die Formen vom Rang 1. Diese spannen den  $n^2$ -dimensionalen reellen Vektorraum der hermiteschen Formen auf. Die Matrixdarstellung (bezüglich jeder beliebigen Basis von  $V$  bzw.  $W$ ) ist ein Produkt ‘Spalte  $\times$  Zeile’:  $\ell^* \cdot \ell$ . Der Pullback macht daraus wieder ein Produkt ‘Spalte  $\times$  Zeile’. Die Spalte ist jeweils hermitisch konjugiert zur Zeile, die für die Linearform  $\ell(\cdot)$  steht.

Eine für unseren Zweck bequeme Definition des Begriffs der hermiteschen Form ist die auf der Grundlage der Parallelogrammgleichung (die wir übrigens in der Hilbertraum-Geometrie ausführlich studieren werden)  $\mathbf{h}(\mathbf{v} + \mathbf{w}) + \mathbf{h}(\mathbf{v} - \mathbf{w}) = 2 \cdot \mathbf{h}(\mathbf{v}) + 2 \cdot \mathbf{h}(\mathbf{w})$ . Auf beiden Wegen ergibt sich sehr einfach die Behauptung.

Wir bemerken zur Übung: Ist  $H$  zur Darstellung von  $\mathbf{h} \in \mathbb{M}^*$  vom Rang 1,  $H = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} (\bar{z}^1, \bar{z}^2)$  mit  $\text{trace} H = |z^1|^2 + |z^2|^2$ ; und ist  $L$  ebenfalls vom Rang 1, dann haben wir

$$\begin{aligned} 2 \cdot \text{trace} H \cdot E - H &= \begin{pmatrix} z^2 \bar{z}^2 & -z^1 \bar{z}^2 \\ -z^2 \bar{z}^1 & z^1 \bar{z}^1 \end{pmatrix} = \begin{pmatrix} z^2 \\ -z^1 \end{pmatrix} (\bar{z}^2, -\bar{z}^1), \\ \frac{1}{2} \text{trace}(H \cdot L) &= \frac{1}{2} \text{trace} \left( \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} (\bar{z}^1, \bar{z}^2) \begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix} (a_1, a_2) \right) = |a_1 \cdot z^1 + a_2 \cdot z^2|^2, \\ (H \parallel L) &= \frac{1}{2} \text{trace} \left( \begin{pmatrix} z^2 \\ -z^1 \end{pmatrix} (\bar{z}^2, -\bar{z}^1) \begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix} (a_1, a_2) \right) = |a_1 \cdot z^2 - a_2 \cdot z^1|^2. \end{aligned}$$

### Kreise in der komplexen Ebene

Wir haben früher die Kreise  $K \subset \bar{\mathbb{C}}$  durch Gleichungen beschrieben:

$$\mathfrak{k} = \left\{ z : a \cdot z\bar{z} + \bar{\alpha}z + \alpha\bar{z} + b = 0 \right\} = \left\{ z : a + \bar{\alpha}\frac{1}{z} + \alpha\frac{1}{\bar{z}} + b\frac{1}{z\bar{z}} = 0 \right\}.$$

wobei  $a, b \in \mathbb{R}$ ,  $\alpha \in \mathbb{C}$  mit  $ab - |\alpha|^2 < 0$ . Oder in Matrixschreibweise:

$$z = \frac{z_1}{z_2} \in \mathfrak{k} \iff (\bar{z}_1, \bar{z}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0 = (\bar{z}_1, \bar{z}_2) \cdot H \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

wo die indefinite hermitische Matrix  $H$  durch den Kreis bis auf einen Faktor eindeutig bestimmt ist.

**Hinweis:** Die Pullback-Abbildung  $\mu_A^*(\cdot)$  bildet Funktionen auf  $\bar{\mathbb{C}}$  in Funktionen auf  $\bar{\mathbb{C}}$  ab. Das Nullstellengebilde  $K_\Psi$  der Funktion  $\Psi(w) = a \cdot w\bar{w} + \bar{\alpha}w + \alpha\bar{w} + b$  wird in das Nullstellengebilde der zurückgenommenen Funktion  $\mu_A^*(\Psi)$  abgebildet. Man muss aber feststellen, dass die inhomogene hermitische Funktionen  $\Psi$  vom Pullback  $\mu_A^*$  nicht in eine hermitische Funktion abgebildet wird. Die folgende Rechnung zeigt einen Faktor, welcher als störend empfunden werden muss, das Nullstellengebilde aber nicht betrifft.

Wenn  $z = \frac{z^1}{z^2}$ ,  $w = \mu_A(z)$ , dann gilt  $w = \frac{w^1}{w^2}$  mit  $\begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$ , und

$$\begin{aligned} \Psi(w) &= |w^2|^{-2} \cdot (\bar{w}_1, \bar{w}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \\ \mu_A^*(\Psi).(z) &= \Psi \left( \frac{a_{11}z^1 + a_{12}z^2}{a_{21}z^1 + a_{22}z^2} \right) = |a_{21}z^1 + a_{22}z^2|^{-2} \cdot (\bar{z}^1, \bar{z}^2) A^* \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} A \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}. \end{aligned}$$

Für die Nullstellengebilde der Funktionen  $\Psi$  und  $\mu^*(\Psi)$  spielen Faktoren wie hier offenbar keine Rolle. Eine glattere Theorie ergibt sich aber, wenn man die Nullstellengebilde mit Hilfe von homogenen hermitischen Funktionen beschreibt.

Wir haben früher bewiesen, dass die Möbiustransformationen Kreise in Kreise transformieren: die Translationen und die Drehstreckungen tun das trivialerweise. Die Reziprokenabbildung tut es, und somit alle aus diesen elementaren Transformationen zusammengesetzten Abbildungen—and das sind die Möbiustransformationen. Wir wollen die Situation hier weiter untersuchen.

Es sei  $\mu_A(\cdot)$  die Möbiustransformation zur unimodularen Matrix; man könnte auch genauer  $\mu_{\pm A}$  schreiben.  $\{\mu_A : A \in SL(2, \mathbb{C}) / \{\pm E\}\}$  ist zunächst einmal eine Gruppenwirkung auf  $\bar{\mathbb{C}}$  (oder auf der Riemannschen Zahlenkugel.) Sie ist dann aber auch eine Gruppenwirkung auf der Menge aller Kreise  $\mathcal{K}$ . Wenn es der Klarheit dient, bezeichnen wir diese Bijektion von  $\mathcal{K}$  mit  $\tilde{\mu}_A(\cdot)$ .

$$\mathcal{K} \ni \mathfrak{k} \longmapsto \tilde{\mu}_A(\mathfrak{k}) = \{\mu_A(z) : z \in \mathfrak{k}\} = \{w : \mu_A^{-1}(w) \in \mathfrak{k}\}.$$

Es gilt offenbar  $\tilde{\mu}_E = id$ ,  $\tilde{\mu}_B(\tilde{\mu}_A(K)) = \tilde{\mu}_{B \cdot A}(K)$ , also  $\tilde{\mu}_B \circ \tilde{\mu}_A(\cdot) = \tilde{\mu}_{B \cdot A}(\cdot)$  usw.

**Definition.** Eine homogene Geradengleichung ist eine Gleichung der Form

$$a \cdot z^1 \bar{z}^1 + \bar{\alpha} z^1 \bar{z}^2 + \alpha \bar{z}^1 \bar{z}^2 + b \cdot z^2 \bar{z}^2 = 0, \quad \text{oder} \quad (\bar{z}_1, \bar{z}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0.$$

wo die hermitesche Matrix  $H = \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix}$  indefinit ist,  $ab - |\alpha|^2 < 0$ .

Wir haben also eine Bijektion der Menge  $\mathcal{K}$  auf die Menge  $\mathbb{H}_{+-}/\mathbb{R}$  der raumartigen Geraden im Standard-Minkowski-Raum.

Die Gruppenwirkung auf der Menge  $\mathcal{K}$ , die durch die Familie der Möbiustransformationen gegeben ist, kann also auch als Gruppenwirkung auf der Menge der raumartigen Geraden im Minkowskiraum verstanden werden. So, wie die Menge der raumartigen Geraden der hyperbolische Raum 'ist, so'ist' die Menge der raumartigen Geraden die Menge der Kreise in  $\bar{\mathbb{C}}$ . Wir haben schon früher gesehen, dass die Menge der Geraden im Lichtkegel die erweiterte Zahlenebene  $\bar{\mathbb{C}}$  ist oder (vermöge der stereographischen Projektion) die Riemann'sche Zahlenkugel.

Fazit: Die Gruppe  $SL(2, \mathbb{C})/\{\pm E\}$  wirkt treu auf dem Minkowskiraum (oder besser: auf der Menge der Geraden im Minkowskiraum). Sie hat drei Bahnen. Die Wirkung auf dem hyperbolischen Raum und die Wirkung auf der Riemann'schen Zahlenkugel haben wir bereits ausführlich studiert. Nun wollen wir uns auch noch der Wirkung auf der Menge der Kreise  $\mathcal{K}$  zuwenden.

Die Pullback-Abbildung  $\mu_A^*(\cdot)$  bildet Funktionen auf  $\bar{\mathbb{C}}$  in Funktionen auf  $\bar{\mathbb{C}}$  ab. Er bildet somit das Nullstellengebilde  $K_\Psi$  der Funktion  $\Psi(w) = a \cdot w\bar{w} + \bar{\alpha}w + \alpha\bar{w} + b$  in das Nullstellengebilde der zurückgenommenen Funktion  $\mu_A^*(\Psi)$  ab. Er bildet aber nicht die inhomogene hermitesche Funktionen  $\Psi$  in hermitesche Funktion ab. Die folgende Rechnung ergibt einen Faktor, welcher als störend empfunden werden kann, das Nullstellengebilde aber nicht betrifft.

Wenn  $z = \frac{z^1}{z^2}$ ,  $w = \mu_A(z)$ , dann gilt  $w = \frac{w^1}{w^2}$  mit  $\begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$ , und

$$\begin{aligned} \Psi(w) &= |w^2|^{-2} \cdot (\bar{w}_1, \bar{w}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \\ \mu_A^*(\Psi).(z) &= \Psi \left( \frac{a_{11}z^1 + a_{12}z^2}{a_{21}z^1 + a_{22}z^2} \right) = |a_{21}z^1 + a_{22}z^2|^{-2} \cdot (\bar{z}^1, \bar{z}^2) \cdot A^* \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}. \end{aligned}$$

Für die Nullstellengebilde der Funktionen  $\Psi$  und  $\mu^*(\Psi)$  spielen Faktoren wie hier offenbar keine Rolle. Es liegt daher nahe, die Nullstellengebilde mit Hilfe der dazugehörigen homogenen hermiteschen Funktionen zu beschreiben.

$$\begin{aligned} K'' &= \{w = \frac{w^1}{w^2} : \Psi(w) = 0\} = \{w = \frac{w^1}{w^2} : (\bar{w}^1, \bar{w}^2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = 0\}, \\ K' &= \{z = \frac{z^1}{z^2} : (\bar{z}^1, \bar{z}^2) \cdot A^* \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = 0\} \end{aligned}$$

Damit erhalten wir eine Bijektion der Menge  $\mathcal{K}$  auf die Menge Äquivalenzklassen  $\mathbb{M}_{+-}^*/\mathbb{R}$  von indefiniten hermiteschen Formen auf dem Vektorraum  $V$ . Dies wollen wir noch genauer diskutieren. Bisher war von Matrixdarstellungen der hermiteschen Formen die Rede, nachdem wir eine Hilbertraumnorm in  $V$  ausgezeichnet hatten; wir brauchten das, da wir in den Konstruktionen auf die Spur Bezug nehmen wollten. Wir kommen jetzt aber nochmal auf die Bemerkung zurück, dass wir keine ausgezeichnete Norm auf  $V$  (d.h. kein ausgezeichnetes positivdefinites Element in  $\mathbb{M}^*$ ) ins Spiel bringen müssen. Wir können vom Produkt der Eigenwerte einer hermiteschen Form auf  $V$  auch dann reden, wenn wir uns darauf verständigen, dass wir nur solche Basen in Betracht ziehen, die durch einen Basiswechsel mit der Determinante 1 auseinander hervorgehen. Zu einer hermiteschen Form  $\mathbf{h}(\cdot)$  und jeder Basis  $(\mathbf{z}^1(\cdot), \mathbf{z}^2(\cdot))$  von  $V^*$  existiert eine hermitesche Matrix  $H$ , sodass

$$\mathbf{h}(\cdot) = (\bar{\mathbf{z}}^1(\cdot), \bar{\mathbf{z}}^2(\cdot)) \cdot H \cdot \begin{pmatrix} \mathbf{z}^1(\cdot) \\ \mathbf{z}^2(\cdot) \end{pmatrix} \quad \text{auf } V.$$

Wenn  $\begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = B \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}$  ( mit  $B \in SL(2, \mathbb{C})/\pm E$ ), dann

$$\mathbf{h}(\cdot) = (\bar{\mathbf{w}}^1(\cdot), \bar{\mathbf{w}}^2(\cdot)) \cdot B^* H B \cdot \begin{pmatrix} \mathbf{w}^1(\cdot) \\ \mathbf{w}^2(\cdot) \end{pmatrix} \quad \text{auf } V \quad \text{mit} \quad \det B^* H B = \det H.$$

Zwei indefinite hermitesche Formen auf  $V$  beschreiben genau dann denselben Kreis auf der Riemann'schen Zahlenkugel, wenn sie sich um einen Faktor unterscheiden. Dies ist die gesuchte Bijektion  $\mathcal{K} \longleftrightarrow \mathbb{M}_{+-}^*$ , die sich nicht auf Matrixdarstellungen stützt, und doch die Determinante (das Produkt der Eigenwerte) im Spiel lässt. Es ist wichtig zu bemerken, dass man in  $\mathcal{K}$  zunächst einmal keine ausgezeichnete Funktion  $\det(\cdot)$  im Blick hat. Erst die Bijektion bringt diese ausgezeichnete Funktion in den Vordergrund. Bei den Möbiustransformationen der Kreise wie bei den eigentlichen Lorentztransformationen des Minkowskiraums bleibt diese Funktion invariant.

## 1.8 Anhang: Gruppenwirkungen, insbesondere Konjugation

In den obigen Überlegungen zur Drehgruppe und zur Lorentzgruppe spielt die Konjugationsabbildung eine entscheidende Rolle.  $((g, \omega) \mapsto g\omega g^{-1})$ . Diese Konstruktion wollen wir hier in einem allgemeineren Kontext etwas genauer studieren.

**Definition 1.4** (Transformationsgruppe).

Es sei  $\Omega$  eine Menge. Eine Menge  $\mathcal{T}$  von bijektiven Abbildungen von  $\Omega$  heisst eine Transformationsgruppe über  $\Omega$ , wenn gilt

- (i) die Identitätsabbildung  $I_\Omega$  gehört zu  $\mathcal{T}$
- (ii) wenn  $\pi(\cdot)$  zu  $\mathcal{T}$  gehört, dann auch die Umkehrabbildung  $\pi^{-1}(\cdot)$
- (iii) wenn  $\pi'(\cdot)$  und  $\pi''(\cdot)$  zu  $\mathcal{T}$  gehören, dann auch die zusammengesetzte Abbildung  $\pi' \circ \pi''(\cdot)$ .

**Sprechweise 1.8.1.** Wenn  $\Omega$  eine endliche Menge ist, dann spricht man auch von einer Permutationsgruppe ('von' oder 'über')  $\Omega$ . Die Gruppe aller Bijektionen der endlichen Menge  $\Omega$  heisst die symmetrische Gruppe über  $\Omega$ . Man bezeichnet sie üblicherweise mit  $\mathcal{S}_\Omega$ . Die Menge aller Permutationen der Menge  $\{1, 2, \dots, n\}$  nennt man die symmetrische Gruppe von  $n$  Elementen  $\mathcal{S}_n$ .

Bemerkung: Eine Aufzählung der  $n$ -punktigen Menge  $\Omega$  liefert einen Isomorphismus  $\mathcal{S}_\Omega \longleftrightarrow \mathcal{S}_n$ . Die Ordnung der symmetrischen Gruppe über  $n$  Elementen ist  $|\mathcal{S}_n| = n!$ .

Die Gruppen, die in der Geometrie auftreten, sind in der Regel Transformationsgruppen. Es ist aber unbedingt nötig, auch die Definition einer abstrakten Gruppe im Auge zu behalten. Hier gibt es dann keine Interpretation der Gruppenelemente; die Gruppenelemente sind nichts weiter als die Elemente einer Menge mit einer bestimmten Struktur.

**Definition 1.5** (Gruppe).

Sei  $(G, e, \circ)$  eine abstrakte Menge mit einem ausgezeichneten Element  $e$  und einer Verknüpfung  $\circ : G \times G \rightarrow G$ . Das Tripel  $(G, e, \circ)$  heisst eine Gruppe, wenn gilt

- (i)  $g \circ e = e \circ g = g$  für alle  $g \in G$  ('Gruppeneins')
- (ii)  $\forall g \exists g' : g \circ g' = e = g' \circ g$  ('Gruppeninverse')
- (iii)  $\forall g, h, k \quad g \circ (h \circ k) = (g \circ h) \circ k$  ('Assoziativgesetz')

**Definition 1.6** (Gruppenisomorphismus).

Zwei Gruppen  $(G', e', \circ)$  und  $(G'', e'', \circ)$  heissen *isomorph*, wenn es eine bijektive Abbildung gibt, welche das Einselement in das Einselement überführt und Produkte in Produkte.

$$i(e') = e'' \quad \text{und} \quad i(g \circ h) = i(g) \circ i(h) \quad \text{für alle } g, h.$$

Jede Bijektion  $i(\cdot)$ , die das leistet, heisst ein *Gruppenisomorphismus* von  $G'$  auf  $G''$ . Ein Gruppenisomorphismus von  $G$  auf  $G$  heisst ein *Automorphismus* der Gruppe  $G$ .

**Definition 1.7** (Gruppenhomomorphismus).

Seien  $(G', e', \circ)$  und  $(G'', e'', \circ)$  Gruppen. Eine Abbildung  $\varphi : G' \rightarrow G''$  heisst ein Gruppenhomomorphismus, wenn gilt  $\varphi(e') = e''$ ;  $\varphi(g' \circ h') = \varphi g' \circ \varphi h'$  für alle  $g', h' \in G'$ .

**Definition 1.8** (Untergruppe).

Eine Teilmenge  $U$  einer Gruppe  $G$  heisst eine Untergruppe von  $(G, e, \circ)$ , wenn  $U$  mit der auf  $U$  eingeschränkten Verknüpfung selbst eine Gruppe ist, wenn also mit  $g$  auch  $g^{-1}$  gehört, und  $g, h \in U \implies g \circ h \in U$ .

**Sprechweise 1.8.2** (Nebenklasse). Wenn  $U$  eine Untergruppe ist und  $a \in G$ , dann heisst die Menge  $a \circ U = \{a \circ u : u \in U\}$  die Linksnebenklasse zu  $a$ . Entsprechend heisst  $U \circ a$  die Rechtsnebenklasse zu  $a$ .

Bemerke: Die Linksnebenklassen der Untergruppe  $U$  zu  $a$  und  $b$  sind entweder gleich oder disjunkt. (Sie sind genau dann gleich, wenn  $a^{-1} \circ b \in U$ .) Die Linksnebenklassen zu  $U$  liefern also eine Partition der Gruppe  $G$ . Dasselbe gilt für das System der Rechtsnebenklassen.

**Definition 1.9** (Konjugationsabbildungen).

Zu jedem Gruppenelement  $h$  definiert man die Konjugation mit  $h$  als die Bijektion

$$\psi_h : G \rightarrow G; \quad G \in g \mapsto h \circ g \circ h^{-1}.$$

Diese Abbildung  $\psi_h$  heisst auch der von  $h$  erzeugte innere Automorphismus (von  $G$ ).

Wenn  $U$  eine Untergruppe ist, dann heisst die Untergruppe  $\psi_h U = \{h \circ u \circ h^{-1} : u \in U\}$  die vermöge  $h$  konjugierte Untergruppe.

**Definition 1.10.** Eine Untergruppe  $H$  heisst ein Normalteiler der Gruppe  $G$  ('normal subgroup' im Englischen), wenn  $\psi_h U = U$  für alle  $h \in G$ .

**Satz 1.8.1** (Homomorphiesatz).

*Eine Untergruppe  $H$  ist genau dann Normalteiler, wenn jede Linksnebenklasse auch Rechtsnebenklasse ist. Die Gesamtheit der Nebenklassen wird in diesem Falle mit  $G/H$  bezeichnet. Die Nebenklasse eines Produkts  $a \circ b$  hängt in diesem Falle nur von den von  $a$  bzw.  $b$  erzeugten Nebenklassen ab. Die Multiplikation macht  $G/H$  zu einer Gruppe. (Man nennt sie die Faktorgruppe.)*

Wenn  $\varphi : G' \rightarrow G''$  ein Gruppenhomomorphismus ist, dann ist der Kern, d.h. die Menge  $\ker \varphi = \{h' : \varphi(h') = e''\}$  ein Normalteiler, und  $G'/\ker \varphi$  ist in natürlicher Weise isomorph zum Bild im  $\varphi = \{g'' : g'' = \varphi(g')\}$ .

## Gruppenwirkungen

### Definition 1.11 (Gruppenwirkung).

Sei  $(G, e, \circ)$  eine abstrakte Gruppe und  $\Omega$  eine Menge. Jedem  $g \in G$  sei eine Bijektion  $\varphi_g(\cdot)$  zugeordnet, sodass gilt

$$\varphi_e(\cdot) = I_\Omega(\cdot); \quad \varphi_h(\varphi_g(\cdot)) = \varphi_{h \circ g}(\cdot),$$

Die Familie  $\Phi = \{\varphi_g : g \in G\}$  heisst dann eine Wirkung von  $G$  auf  $\Omega$  ('group-action' im Englischen). Sie heisst eine treue Gruppenwirkung, wenn allein  $\varphi_e$  die Identitätsabbildung ist. ('faithful' Im Englischen.)

Bemerkte: Wenn die Gruppenwirkung  $\{\varphi_g : g \in G\}$  nicht treu ist, wenn also  $H = \{h : \varphi_h = id_\Omega\} \neq \{e\}$ , dann kann man die Familie der Bijektionen als eine treue Wirkung der Faktorgruppe  $G/H$  verstehen.

### Definition 1.12 (Bahnen und Fixgruppen).

Es sei  $\{\varphi_g : g \in G\}$  eine Wirkung der Gruppe  $G$  auf der Grundmenge  $\Omega$ .

Für ein  $\omega^* \in \Omega$  heisst die Punktmenge  $\Omega_{\omega^*} = \{\omega : \omega = \varphi_g(\omega^*) \text{ mit } g \in G\} \subseteq \Omega$  die Bahn des Punkts  $\omega^*$  ('orbit' im Englischen).

Die Menge  $G_{\omega^*} = \{g : \varphi_g(\omega^*) = \omega^*\} \subseteq G$  heisst die Fixgruppe des Punkts  $\omega^*$ .

**Satz 1.8.2.** *Die Bahnen zweier Punkte  $\omega^*$  und  $\omega^{**}$  sind entweder disjunkt oder identisch. Wenn  $\varphi(\omega^*) = \omega^{**}$ , dann sind die Fixgruppen von  $\omega^*$  und  $\omega^{**}$  zueinander konjugiert.*

$$\varphi_g(\omega^{**}) = \omega^{**} \iff \varphi_g(\varphi_h(\omega^*)) = \varphi_h(\omega^*) \iff \varphi_{h^{-1} \circ g \circ h}(\omega^*) = \omega^*.$$

**Sprechweise 1.8.3.** Man sagt, eine Gruppe wirkt transitiv auf  $\Omega$ , wenn jeder Punkt  $\omega_1$  durch eine geeignete Transformation in einen beliebig vorgegebenen Punkt  $\omega_2$  übergeführt werden kann.  $\forall \omega_1, \omega_2 \exists g : \varphi(\omega_1) = \varphi(\omega_2)$ .

Man sagt, sie wirkt einfach transitiv, wenn es genau ein solches  $g$  gibt.

Ein (nahezu triviales) Beispiel ist das folgende: Die Gruppe  $G$  wirkt vermöge der Multiplikation von links einfach transitiv auf  $\Omega = G$ . ( $\varphi_g(\omega) = g \circ \omega$ .)

Bei dieser Gelegenheit erinnern wir an die Definition eines  $\mathbb{K}$ -affinen Raums:

*Ein  $\mathbb{K}$ -affiner Raum ist eine Menge, auf welcher ein  $\mathbb{K}$ -Vektorraum einfach transitiv operiert. Die Operationen heissen die Parallelverschiebungen.*

Betrachten wir die Begriffsbildungen etwas näher im Falle von endlichen Gruppen. Die Anzahl der Elemente der Gruppe  $G$  heisst die Gruppenordnung und wird mit  $|G|$  bezeichnet. Wenn eine endliche Gruppe auf einer Menge  $\Omega$  operiert, dann sind natürlich alle Bahnen  $\Omega_{\omega^*}$  endlich.

### Satz 1.8.3 (Gruppenordnungen).

*Die endliche Gruppe  $G$  operiere auf der Menge  $\Omega$ . Für die Fixgruppe eines beliebigen Punkts  $\omega \in \Omega$  und seine Bahn gilt dann  $|G_\omega| \cdot |\Omega_\omega| = |G|$ .*

**Beweis.** Betrachten wir die Linksnebenklassen  $g \circ G_\omega$ . Sie partitionieren  $G$  und haben alle dieselbe Mächtigkeit. Ihre Anzahl ist also  $m = \frac{|G|}{|G_\omega|}$ . Sie entsprechen andererseits in eindeutiger Weise den Punkten der Bahn von  $\omega$ :  $g \circ G_\omega \rightleftharpoons \varphi_g(\omega)$ . Es gilt nämlich  $h \in g \circ G_\omega \Leftrightarrow \varphi_h(\omega) = \varphi_g(\omega)$ .

*Beispiel 1.8.1* (Konjugation für die symmetrische Gruppe). Die Konjugationsabbildung kann man als eine Wirkung der Gruppe  $G$  auf der Menge  $\Omega = G$  verstehen.

$$\psi_g(\omega) = g \circ \omega \circ g^{-1}; \quad \text{es gilt} \quad \psi_h(\psi_g(\cdot)) = \psi_{h \circ g}(\cdot)$$

Die Bahnen dieser Gruppenwirkung können sehr verschieden aussehen. So ist z. B. die Bahn des Einheitselements einpunktig. Einpunktig sind auch die Bahnen derjenigen Elemente, die mit allen  $g \in G$  kommutieren. (Die Gruppentheoretiker nennen diese Menge das Zentrum der Gruppe; es handelt sich offenbar um einen Normalteiler.)

Betrachten wir konkret die symmetrische Gruppe über  $n$  Elementen,  $\mathcal{S}_n$ . Die Punkte  $\omega$  sollen in der Zykelschreibweise dargestellt sein. Den konjugierten Punkt  $g \circ \omega \circ g^{-1}$  kann man dadurch finden, dass man die Ziffern in den Zyklen der Permutation  $g$  unterwirft. Wir sagen,  $\omega$  sei vom Typ  $(m_1, m_2, \dots, m_n)$ , wenn  $\omega$  in der Schreibweise mit disjunkten Zyklen  $m_j$  Zyklen der Länge  $j$  aufweist ( $m_1$  ist beispielsweise die Anzahl der Fixpunkte). Mögliche Typen sind die Tupel  $(m_1, m_2, \dots, m_n)$  mit  $1 \cdot m_1 + 2 \cdot m_2 + \dots + n \cdot m_n = n$ .

Die Konjugationsabbildung macht aus jeder zyklischen Permutation der Länge  $j$  eine zyklische Permutation der Länge  $j$ , und man kann jede zyklische Permutation in jede zyklische Permutation derselben Länge überführen. Die Gruppe  $\mathcal{S}_n$  wirkt also transitiv auf der Menge  $\Omega_j$  aller zyklischen  $\omega$  der Länge  $j$ . Die Fixgruppe eines Zyklus der Länge  $n$  hat offenbar die Mächtigkeit  $n$ . Die Anzahl der verschiedenen Zykeln der Länge  $n$  ist  $(n-1)! = \frac{1}{n}n!$ . Man kann sich unschwer vorstellen, wie die Fixgruppe eines Zyklus der Länge  $j$  aussieht.

Die Bahn eines  $\omega$  vom Typ  $(m_1, m_2, \dots, m_n)$  besteht aus allen  $\omega'$  vom gleichen Typ. Es ist aber eine etwas mühsame Arbeit, die Anzahl der verschiedenen Konjugationsklassen (oder Bahnen) auszurechnen. Es geht um die zahlentheoretische Frage: auf wieviele Weisen kann man die Zahl  $n$  zerlegen in der Weise

$$n = 1 \cdot m_1 + 2 \cdot m_2 + \dots + n \cdot m_n?$$

Im Fall  $n = 4$  sind die Konjugationsklassen durch die folgenden Zerlegungen gegeben:

1.  $4 = 4 \cdot 1$  (es gibt sechs Viererzykeln)
2.  $4 = 1 \cdot 1 + 3 \cdot 1$  (es gibt acht Dreierzykeln bei einem Fixpunkt)
3.  $4 = 2 \cdot 2$  (es gibt drei Produkte zweier disjunkter Zweierzykeln)
4.  $4 = 2 \cdot 1 + 1 \cdot 2$  (es gibt sechs Transpositionen)
5.  $4 = 4 \cdot 1$  (die Identität)

Es gibt also fünf Bahnen der Konjugationsabbildung zur symmetrischen Gruppe  $\mathcal{S}_4$ .

*Beispiel 1.8.2.* Die Oktaedergruppe ist die Gruppe aller derjenigen Drehungen des dreidimensionalen Raums, die den Einheitswürfel auf sich abbilden. Man kann sich dazu mehrere Gruppenwirkungen ausdenken:

Die  $g \in G$  wirken als Permutationen der Menge  $\Omega_F$  der Seitenflächen. Sie wirken andererseits auf der Menge  $\Omega_E$  der Ecken des Würfels, oder auch auf der Menge  $\Omega_K$  der Kanten. Es handelt sich in jedem Falle um eine treue Gruppenwirkung, welche transitiv auf den Grundraum  $\Omega$  wirkt.

Studieren wir etwas genauer die Wirkung auf der sechspunktigen Menge  $\Omega_F$ . Jede Bahn hat die Größe  $|\Omega_F| = 6$ . Die Fixgruppe einer Fläche  $\omega$  ist die Gruppe  $G_\omega$  der Drehungen um die Achse, die auf der Fläche  $\omega$  senkrecht steht. Man kann um  $90^\circ$ ,  $180^\circ$  und  $270^\circ$  drehen; dazu kommt die Identität.  $|G_\omega| \cdot |\Omega_F| = 4 \cdot 6 = 24 = |G|$ .

Bei der transitiven Wirkung auf die achtpunktige Grundmenge  $\Omega_E$  ist die Fixgruppe jeder Ecke die Menge der (einschliesslich der Identität drei) Drehungen um die entsprechende Raumdiagonale.  $3 \cdot 8 = 24 = |G|$ .

Bei der transitiven Wirkung auf die zwölfpunktige Grundmenge  $\Omega_K$  besteht die Fixgruppe jeder Kante aus der Identität und der Klappung um diese Kante.  $2 \cdot 12 = 24 = |G|$ .

Besonders ‘sparsam’ ist die Wirkung von  $G$  auf die vierpunktige Menge  $\Omega_D$  der Raumdiagonalen des Würfels. Die Fixgruppe jeder Raumdiagonalen ist trivial. Die Größe einer Bahn ist also 24; die Wirkung ist transitiv; zu jeder Permutation der vier Raumdiagonalen gibt es genau ein Drehung. Damit erhalten wir einen Isomorphismus  $G \longleftrightarrow \mathcal{S}_4$ . Und wir erkennen die geometrische Realisierung der verschiedenen oben aufgelisteten Konjugationsklassen der symmetrischen Gruppe: Vom Typ 1 sind die Drehungen um  $90^\circ$  bzw.  $270^\circ$  um eine flächensenkrechte Achse, vom Typ 2 sind die Drehungen um eine Raumdiagonale, usw.

*Beispiel 1.8.3 (Pullbackwirkungen).* Es sei  $G$  eine endliche Gruppe und  $\Omega$  eine genügend große Menge von Funktionen auf  $G$ . Für  $h \in G$  und  $f(\cdot) \in \Omega$  sei  $g(\cdot) = \psi_h f(\cdot)$  die Funktion  $g(\omega) = f(h^{-1} \circ \omega)$ . Wir schreiben auch  $g = \varphi_{h^{-1}}^* f$ . Für  $h_1, h_2 \in G$  gilt  $\psi_{h_1}(\psi_{h_2} f) = \psi_{h_1 \circ h_2} f$ ; denn mit  $g(\cdot) = \psi_{h_2} f(\cdot)$  haben wir

$$\psi_{h_1} g(\omega) = g(h_1^{-1} \circ \omega) = f(h_2^{-1} \circ h_1^{-1} \circ \omega) = f((h_2 \circ h_1)^{-1} \circ \omega) = \psi_{h_1 \circ h_2} f(\omega).$$

Wenn die Funktionenmenge  $\Omega$  mit  $f$  auch alle  $\varphi_{h^{-1}}^* f$  enthält, dann haben wir eine Wirkung von  $G$  auf  $\Omega$ . Denken wir an komplexwertige Funktionen auf  $G$ ; wir verstehen sie als komplexe  $G$ -Spalten. Sei nun  $\Omega = \mathbb{C}_{Sp}^G$  der Vektorraum aller komplexen  $G$ -Spalten. Die Abbildungen  $\psi_h = \varphi_{h^{-1}}^*$  sind durch Permutationsmatrizen  $P_h$  zu beschreiben:  $\psi_g f = P_h \cdot f$  (Matrizenprodukt). Wir haben dafür gesorgt, dass gilt  $P_{h_1 \circ h_2} = P_{h_1} \cdot P_{h_2}$  (Matrizenprodukt).

Die Konstruktionsmethode hat in der elementaren Mengenlehre einen (mittlerweile auch im Deutschen gern gebrauchten) Namen:

**Sprechweise 1.8.4** (Pullback, zurückgenommene Funktionen).

Gegeben sei eine Abbildung  $\varphi : \Omega_1 \longrightarrow \Omega_2$ . Für jede Funktion  $f_2(\cdot)$  auf  $\Omega_2$  definiert man

die ‘zurückgenommene’ Funktion  $f_1 = \varphi^* f_2$ ; dies ist die Funktion auf  $\Omega_1$  mit  $f_1(\omega_1) = f_2(\varphi\omega_1)$ . Die Abbildung  $\varphi^* : f_1 \mapsto f_2 = \varphi^* f_1$  nennt man die Pullback-Abbildung. Genauer: Um wirklich eine Abbildung  $\varphi^*$  zu gewinnen, sind Funktionenräume  $\mathfrak{F}_2$  und  $\mathfrak{F}_1$  zu spezifizieren, sodass  $\varphi^*(f_2) \in \mathfrak{F}_1$  für alle  $f_2 \in \mathfrak{F}_2$ . Man wird nicht immer die Menge aller (komplexwertigen) auf  $\Omega_2$  bzw.  $\Omega_1$  in Betracht ziehen. In vielen Anwendungen sind  $\mathfrak{F}_2$  und  $\mathfrak{F}_1$  passende komplexe Vektorräume. Man bemerke  $\varphi^*(\alpha f + \beta g) = \alpha \varphi^* f + \beta \varphi^* g$ .

Beim Übergang von den Punktabbildungen  $\varphi(\cdot), \psi(\cdot)$  zu den Pullback-Abbildungen ist zu beachten, dass sich beim Hintereinanderschalten die Reihenfolge umkehrt:

$$(\varphi \circ \psi)^*(\cdot) = \psi^*(\varphi^*(\cdot)) \quad \Omega_1 \xrightarrow{\varphi} \Omega_2 \xrightarrow{\psi} \Omega_3 \quad \mathfrak{F}_1 \xleftarrow{\varphi^*} \mathfrak{F}_2 \xleftarrow{\psi^*} \mathfrak{F}_3.$$

Im Beispiel oben betrachteten wir Pullbackabbildungen  $\psi_g(\cdot) = \varphi_{g^{-1}}^*(\cdot)$  auf einer Menge  $\mathfrak{F}$  von Funktionen über  $\Omega$ . Die Familie  $\{\psi_g : g \in G\}$  ist eine Gruppenwirkung wegen

$$\psi_{h \circ g} = \varphi_{(h \circ g)^{-1}}^* = \varphi_{g^{-1} \circ h^{-1}}^* = \varphi_{h^{-1}}^* \circ \varphi_{g^{-1}}^* = \psi_h \circ \psi_g.$$

Hinweis: Im Augenblick denken wir hauptsächlich an endliche Gruppen. Ein Blick auf einen wichtigen Fall mit einer kontinuierliche Gruppe mag aber erhellend sein. Es sei  $G = SL(n, \mathbb{C})$  die spezielle lineare Gruppe und  $\Omega$  der Hilbertraum  $(\mathbb{C}_{Sp}^n, \langle \cdot | \cdot \rangle)$  der komplexen  $n$ -Spalten  $\mathbf{v}$ . Die Matrizenmultiplikation von links,  $\varphi_A(\mathbf{v}) = A \cdot \mathbf{v}$ , liefert über  $\Omega$  eine Darstellung  $\Phi = \{\varphi_A : A \in G\}$ .

Die hermiteschen Formen  $h(\cdot)$  auf  $\mathbb{C}_{Sp}^n$  sind durch die hermiteschen  $n \times n$ -Matrizen  $H$  gegeben,  $h(\mathbf{v}) = \mathbf{v}^* H \mathbf{v}$ . Es bezeichne  $\mathcal{F}$  den  $n^2$ -dimensionalen reellen Vektorraum dieser hermiteschen Formen. Die Pullbackabbildungen  $\varphi_A^* : h \mapsto \varphi_A^*(h)$  sind Bijektionen dieses Funktionenraums. Für  $h$  zur Matrix  $H$  ist  $k = \varphi_A^*(h)$  die Form zur Matrix  $K = A^* H A$ ;  $h(\varphi_A \cdot \mathbf{v}) = \mathbf{v}^* A^* H A \mathbf{v}$ . Die Familie  $\{\varphi_{A^{-1}}^* : A \in GL(n, \mathbb{C})\}$  ist eine Gruppenwirkung auf dem Funktionenraum  $\mathcal{F}$ , die man auch als eine Gruppenwirkung auf der Menge  $\mathcal{H}$  der hermiteschen Matrizen verstehen kann:  $\rho_A(H) = (A^{-1})^* H A^{-1}$ . für  $H \in \mathcal{H}$ . Offenbar gilt  $\rho_B(\rho_A(H)) = \rho_{BA}(H)$ .

Eine sehr ähnliche Gruppenwirkung wird uns später beschäftigen, nämlich die Familie der Abbildungen  $\sigma_A : H \mapsto A H A^*$  für  $H \in \mathcal{H}, A \in SL(n, \mathbb{C})$ .

### Gruppen von Endomorphismen eines Vektorraums.

Eine wichtiges Gebiet der linearen Algebra ist die Theorie der Gruppendarstellungen. Man kann sie als einen Spezialfall der Theorie der Gruppenwirkungen entwickeln. Man kann aber auch etwas anders beginnen, etwa mit der

**Definition.** Es sei  $(G, e, \cdot)$  eine abstrakte Gruppe. Jedem  $a \in G$  sei eine komplexe  $n \times n$ -Matrix  $A_a$  zugeordnet, sodass  $A_{a \cdot b} = A_a \cdot A_b$  (Matrizenprodukt). Die Familie  $\{A_a : a \in G\}$  heisst dann eine Matrixdarstellung von  $G$ .

Aus  $A_a \cdot A_{a^{-1}} = A_e$  und  $A_e \cdot A_a = A_e = A_a \cdot A_e$  ergibt sich, dass alle Matrizen  $A_a$  den gleichen Rang  $r$  haben. Alle  $A_a$  liefern in der Tat (durch die Multiplikation von links)

Automorphismen eines  $r$ -dimensionalen Raums von Spalten. Wenn  $A_e$  die Einheitsmatrix ist, dann sprechen wir von einer Darstellung von vollem Rang. Im allgemeinen Fall können wir  $A_e$  als eine Projektion auf einen Teilvektorraum  $U$  des Spaltenraums verstehen, den wir den Träger der Matrixdarstellung nennen. Bemerkte: Wenn wir im ‘Träger’ einer Matrixdarstellung eine Basis einführen, dann liefern uns die  $A_a$  eine Matrixdarstellung durch nichtsinguläre  $r \times r$ -Matrizen. Auf das Thema Basenwahl kommen wir später. Hier stellen wir fest.

**Lemma.** *Es sei  $\Phi$  eine Matrixdarstellung der Gruppe  $G$  über  $V$  mit dem Träger  $U$ , und es sei  $P$  eine Projektion von  $V$ , die mit  $\Phi$  verträglich ist in dem Sinn, dass  $P$  mit allen  $A_a$  kommutiert. ( $\forall a \in G : A_a \cdot P = P \cdot A_a$ .) Dann ist  $\Psi = \{\psi_a = A_a \cdot P : a \in G\}$  eine Matrixdarstellung mit dem Träger  $P(U)$ .*

Die geschickte Wahl von Basen ist das zentrale Anliegen der Darstellungstheorie. Um da nichts vorwegzunehmen, beginnt die Theorie besser ganz ohne Basen mit der

**Definition 1.13.** Es sei  $(G, e, \cdot)$  eine abstrakte Gruppe und  $V$  ein komplexer Vektorraum.  $\mathfrak{A}_V$  sei die Menge der Automorphismen des Vektorraums  $V$ . Eine Darstellung der Gruppe  $G$  von vollem Rang über  $V$  ist eine Abbildung  $\Phi : G \longrightarrow \mathfrak{A}_V$  mit  $\Phi(e) = id_V$  und  $\Phi(a \cdot b) = \Phi(a) \circ \Phi(b)$  für alle  $a, b \in G$ .

**Notation.** Statt  $\Phi(a)$  oder  $\Phi(a)(\cdot)$  notieren wir auch manchmal  $\varphi_a(\cdot)$ . Für den zusammengesetzten Automorphismus  $\Phi(a) \circ \Phi(b)$  notieren wir auch  $\Phi(a) \cdot \Phi(b)$ , wie das bei Operatoren üblich ist.

Im Anschluss an die oben eingeführten Sprechweisen können wir sagen: Eine Gruppenwirkung von vollem Rang ist eine Gruppenwirkung über einem komplexen Vektorraum, wo alle Bijektionen Vektorraumautomorphismen sind.

**Satz 1.8.4** (Satz von der vollständigen Reduzierbarkeit).

*Es sei  $\Phi = \{\varphi_a : a \in G\}$  eine Darstellung einer endlichen Gruppe  $G$  über dem Vektorraum, und es sei  $U \subset V$  ein  $\Phi$ -invarianter Teilraum. ( $\forall a : \varphi_a(U) \subseteq U$ .) Dann existiert eine Projektion auf  $U$ , welche mit  $\Phi$  verträglich ist.*

**Beweis 1.8.1.** *Wir zeigen zunächst, dass für jeden Endomorphismus  $\sigma$  der ‘symmetrisierte’ Endomorphismus mit allen  $\varphi_g$  kommutiert:*

$$\bar{\sigma} = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ \sigma \circ \varphi_a \implies \forall g \in G : \bar{\sigma} \circ \varphi_g = \varphi_g \circ \bar{\sigma}$$

In der Tat liefert der Übergang vom Summationsindex  $a$  zum Summationsindex  $b = ag$

$$\bar{\sigma} \circ \varphi_g = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \sigma \varphi_{ag} = \varphi_g \circ \frac{1}{|G|} \sum_{a \in G} \varphi_{(ag)^{-1}} \sigma \varphi_{ag} = \varphi_g \circ \frac{1}{|G|} \sum_{b \in G} \varphi_{b^{-1}} \sigma \varphi_b = \varphi_g \circ \bar{\sigma}.$$

Es sei nun  $\tau$  irgendeine Projektion auf  $U$  und  $\bar{\tau} = \pi$  der symmetrisierte Endomorphismus. Wegen der Invarianz von  $U$  gilt

$$\begin{aligned} \tau \circ \varphi_a \circ \tau &= \varphi_a \circ \tau \quad \text{für alle } a \in G \\ \tau \circ \pi &= \frac{1}{|G|} \sum_{a \in G} (\tau \circ \varphi_{a^{-1}} \circ \tau) \circ \varphi_a = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ \tau \circ \varphi_a = \pi; \\ \pi \circ \tau &= \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ (\tau \circ \varphi_a \circ \tau) = \frac{1}{|G|} \sum_{a \in G} (\varphi_{a^{-1}} \circ \varphi_a) \circ \tau = \tau. \end{aligned}$$

Das zeigt zunächst einmal, dass der symmetrisierte Endomorphismus  $\bar{\tau} = \pi$  denselben Rang hat wie  $\tau$ . Er hat Werte in  $U$  und ist in der Tat eine Projektion auf  $U$ ; denn  $\pi \circ \pi = \pi \circ \tau \circ \pi = \tau \circ \pi = \pi$ .

**Sprechweise 1.8.5.** Eine Gruppendarstellung  $\Phi = \{\varphi_a : a \in G\}$  über einem Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$  heisst eine unitäre Darstellung, wenn alle  $\varphi_a$  Isometrien des Hilbertraums sind.

**Satz 1.8.5.** Es sei  $\Phi = \{\varphi_a : a \in G\}$  eine unitäre Darstellung einer Gruppe über dem Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ . Wenn der Teilraum  $U$   $\Phi$ -invariant ist, dann ist auch das orthogonale Komplement  $W = U^\perp$   $\Phi$ -invariant.

**Beweis.**  $w \perp U \iff \forall u \in U : \langle w | u \rangle = 0 \iff \forall a \in G \forall u \in U : \langle w | \varphi_{a^{-1}} u \rangle = 0 \iff \forall a \in G \forall u \in U : \langle \varphi_a w | u \rangle = 0 \iff \forall a \in G : \varphi_a w \perp U$ .

Bei den Darstellungen endlicher Gruppen ist keine Einschränkung, wenn man fordert, dass die Automorphismen Isometrien sind. Mit der Technik der Symmetrisierung beweisen wir nun nämlich den

**Satz 1.8.6.** Es sei  $\Phi = \{\varphi_a : a \in G\}$  eine Darstellung der endlichen Gruppe  $G$  über dem komplexen Vektorraum  $V$ . Es existiert dann eine Hilbertraummetrik, sodass alle  $\varphi_a$  Isometrien sind.

**Beweis 1.8.2.** Hier ist es bequem, in  $V$  eine Basis  $\{\mathbf{v}_j : j \in J\}$  zu wählen und die Automorphismen  $\varphi_a$  durch die  $J \times J$ -Matrizen  $B_a$  auszudrücken, die auf dem Raum der  $J$ -Spalten durch Multiplikation von links wirken.

Wählen wir nun eine positivdefinite Matrix  $Q$ . Für die ‘hermitisch symmetrisierte’ positivdefinite Matrix

$$\begin{aligned} H &:= \frac{1}{|G|} \sum_{a \in G} B_a^* \cdot Q \cdot B_a. \quad \text{gilt dann} \\ B_g^* \cdot H \cdot B_g &= H \quad \text{für alle } g \in G. \text{ In der Tat} \end{aligned}$$

In der Tat gilt wegen  $B_a \cdot B_g = B_{ag}$ , und  $B_g^* \cdot B_a^* = (B_{ag})^*$

$$B_g^* \cdot H \cdot B_g = \frac{1}{|G|} \sum_a B_g^* B_a^* Q B_a B_g = \frac{1}{|G|} \sum_h B_h^* Q B_h = H, \quad ,$$

denn die Summe geht über alle  $h = ag$  ( $a \in G$ ) und das bedeutet nur eine andere Aufzählung aller Elemente von  $G$ .

$H$  definiert eine Hilbertraumstruktur und alle  $B_g$  sind Isometrien für diese Struktur

$$\langle B_g y | B_g y \rangle = y^* B_g^* H B_g y = y^* H y = \langle y | y \rangle$$

Diese Hilbertraum auf dem Raum der  $J$ -Spalten liefert eine Hilbertraumnorm auf dem Träger  $V$  der gegebenen Darstellung  $\Phi$ . Wenn wir dort nun eine ONB  $\{\mathbf{u}_i : i \in I\}$  einführen, dann sind die darstellenden Matrizen unitäre  $I \times I$ -Matrizen.

Die Eigenschaft einer Darstellung, dass zu einem invarianten Teilraum ein invariantes Komplement existiert, heisst die Eigenschaft der vollständigen Reduzierbarkeit der Darstellung. Sie hat weitreichende Konsequenzen. Wir haben zu jedem invarianten Teilraum  $U$  eine mit der Darstellung verträgliche direkte Zerlegung  $V = U \oplus W$ . Die Darstellung zerfällt somit in Darstellungen mit den Trägern  $U$  und  $W$ . Die Zerlegung kann weitergetrieben werden: Wenn  $U'$  ein weiterer  $\Phi$ -invarianter Teilraum ist und  $W'$  sein invariantes Komplement, dann sind auch die Teilräume  $U' \cap U$ ,  $U' \cap W$ ,  $W' \cap U$ ,  $W' \cap W$   $\Phi$ -invariante Teilräume. Man kann daher (im endlichdimensionalen Fall) die verträglichen direkten Zerlegungen so lange verfeinern, bis die Summanden unzerlegbar sind in dem Sinne, dass sie keine nichttrivialen invarianten Teilräume besitzen.

Hinweis auf die Reduktionstheorie der Darstellungen einer endlichen Gruppe:

Die Träger der Darstellungen sind endlichdimensionale komplexe Vektorräume. Die Darstellungen mit einem unzerlegbarem Träger nennt man irreduzible Darstellungen. Zwei irreduzible Darstellungen heissen äquivalent,  $\Phi \sim \Psi$ , wenn ein Isomorphismus der Träger  $\sigma(\cdot)$  existiert, sodass  $\sigma \circ \varphi_a \circ \sigma^{-1} = \psi_a$  für alle  $a \in G$ . In der Darstellungstheorie für endliche Gruppen zeigt man, dass es nur endlich viele nichtäquivalente irreduzible Darstellungen gibt, und man entwickelt Verfahren, um diese explizit zu bestimmen.— Wir können diese faszinierende Theorie im gegenwärtigen Rahmen nicht weiter verfolgen.

**Nochmals Möbiustransformationen und der Standard-Minkowski-Raum**

Im Folgenden benützen wir einige weitergehende Begriffsbildungen, die wir im abschließenden Anhang erläutern. Die Bahn einer Gruppenwirkung ist ein solcher Begriff. Ein anderes wichtiges Konzept ist das Konzept der Pullback-Abbildung.

Die Möbiustransformationen bilden Kreise in  $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  in Kreise ab, (wo die Geraden als Kreise durch den unendlichfernen Punkt zu betrachten sind). Wenn man Kreise als Nullstellengebilde von Funktionen beschreibt, dann kommen hermitische Formen ins Spiel. Wir werden sehen: Die Möbiustransformationen  $\mu_B(\cdot)$  begründen interessante Pullbackabbildungen für allerlei Funktionenmengen. In Verbindung mit solchen Pullbacks interessieren wir uns in besonderem Maße für eine Gruppenwirkung  $\{\beta(\cdot) : B \in SL(2, \mathbb{C})\}$  auf dem Raum  $\mathbb{M}^*$  der hermitischen Formen auf einem zweidimensionalen komplexen Vektorraum  $V$ . Zuerst einige Vorbemerkungen.

**Lemma.** *Es sei  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$  eine Basis eines komplexen Vektorraums  $V$ . Jede hermitische Matrix  $H = (h_{jk})$ , ( $H^* = H$ ) definiert eine hermitische Form  $\mathbf{h}(\cdot)$  auf  $V$ :*

$$\mathbf{h}(v) = \mathbf{h}\left(\sum_j \mathbf{u}_j \cdot a^j\right) = \sum_{jk} \bar{a}^j h_{jk} a^k = (\bar{a}^1, \dots, \bar{a}^n) \cdot H \cdot \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$$

*Jede hermitische Form  $\mathbf{h}(\cdot)$  wird durch genau eine hermitische Matrix  $H$  beschrieben.*

**Notation.** Wenn wir in  $V$  eine Basis eingeführt haben, dann repräsentieren wir die Vektoren durch Spalten. Die Linearformen repräsentieren wir (der dualen Basis entsprechend) durch Zeilen.

Auch die hermitischen Formen auf dem Dualraum repräsentieren wir durch hermitische Matrizen. Sie können in natürlicher Weise als die Elemente des Dualraums zu  $\mathbb{M}^*$  verstanden werden; diesen Dualraum bezeichnen wir mit  $\mathbb{M}$ . In dieser natürlichen Auffassung ergeben  $L \in \mathbb{M}$ ,  $H \in \mathbb{M}^*$  den Skalar  $(L, H) = \text{trace}(L \cdot H)$ .

Die duale Basis  $(\mathbf{z}^1, \dots, \mathbf{z}^n)$  verstehen wir auch als ein  $n$ -Tupel von Funktionen auf  $V$ , und wir notieren dementsprechend

$$\mathbf{h}(\cdot) = \sum_{jk} \bar{\mathbf{z}}^j(\cdot) \cdot h_{jk} \cdot \mathbf{z}^k(\cdot).$$

**Lemma.** *Es sei  $(V, \|\cdot\|)$  ein  $n$ -dimensionaler unitärer Raum. Zu jeder hermitischen Form  $\mathbf{h}(\cdot)$  existieren reelle Zahlen  $\lambda_1 \geq \dots \geq \lambda_n$  und eine ONB  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ , sodass*

$$\mathbf{h}(\cdot) = \lambda_1 \cdot |\mathbf{z}^1|^2 + \dots + \lambda_n \cdot |\mathbf{z}^n|^2.$$

*Die absteigende Folge der ‘Eigenwerte’ von  $\mathbf{h}(\cdot)$  ist eindeutig bestimmt, wenn die ausgezeichnete Form  $\|\cdot\|^2$  gegeben ist. Das Produkt dieser Eigenwerte ändert sich für alle  $\mathbf{h}(\cdot)$  um denselben Faktor, wenn man eine andere Hilbertraum-Norm auszeichnet.*

**Beweis.** Das Produkt der Eigenwerte ist die Determinante der darstellenden Matrix  $H$ . Wenn man zu einer anderen Basis übergeht, (z. B. zu einer ONB bzgl. der neuen Hilbertraum-Norm), dann wird aus der darstellenden Matrix  $H$  die Matrix  $B^*HB$ , wo  $B$  die Matrix der Basistransformation ist.

Da wir im Folgenden nur Basistransformationen mit der Determinante 1 in Betracht ziehen, kommt jeder hermiteschen Form eine ‘Determinante’ zu. Diese Funktion  $q(\cdot)$  auf  $\mathbb{M}^*$  soll uns im Fall  $n = 2$  näher beschäftigen. Sie ist da nämlich eine quadratische Form, die  $\mathbb{M}^*$  zu einem Minkowskiraum macht, zu einem reell vierdimensionalen Raum mit einer quadratischen Form mit der Signatur  $(1, 3)$ .

Wenn wir einen zeitartigen Einheitsvektor (nämlich  $\|\cdot\|^2$ ) als das Zentrum unseres Minkowskiraums  $\mathbb{M}^*$  auszeichnen, dann hat das die Konsequenz, dass die Spur von  $(h)$  ein wohldefinierte Zahl ist. Die darstellende Matrix von  $\mathbf{h}$  hat in jeder ONB (bzgl. der ausgezeichneten Norm) dieselbe Summe der Diagonalelemente,  $\text{trace}(H) = \text{trace}(\mathbf{h})$ , nämlich die Summe der Eigenwerte. Wenn  $a = \frac{1}{2}\text{trace}H$ , dann hat  $\mathbf{h}(\cdot) - a \cdot \|\cdot\|^2$  Eigenwerte  $\pm\lambda$ , wo  $-\lambda^2 = \det(H - a \cdot E)$ ,  $q(\mathbf{h}) = \det H = (a - \lambda)(a + \lambda)$ . In diesem Fall haben wir bequeme Formeln für die quadratische Form  $q(\cdot)$  und die daraus abgeleitete symmetrische Bilinearform  $(\cdot \| \cdot)$ , nämlich

$$\begin{aligned} (H \| K) &= \frac{1}{2}\text{trace}H \cdot \text{trace}K - \frac{1}{2}\text{trace}(H \cdot K) \\ &= ab + (H_0 \| K_0) = ab - \langle H_0 | K_0 \rangle \\ &= \frac{1}{2}\text{trace}((2a \cdot E - H) \cdot K), \quad \text{wenn} \\ a &= \frac{1}{2}\text{trace}H, \quad b = \frac{1}{2}\text{trace}K, \quad H = a \cdot E + H_0, \quad K = b \cdot E + K_0. \end{aligned}$$

**Lemma.** Es seien  $V$  und  $W$  komplexe Vektorräume und  $\beta(\cdot) : V \rightarrow W$  eine lineare Abbildung. Der Pullback  $\beta^*(\cdot)$  bildet dann hermitesche Form auf  $W$  in hermitesche Formen auf  $V$  ab.

**Beweis.** Wir erinnern noch einmal an einige Eigenschaften der hermiteschen Formen auf einem  $n$ -dimensionalen komplexen Vektorraum. Die einfachsten hermiteschen Formen  $\mathbf{h}(\cdot)$  ergeben sich als Absolutquadrat einer Linearform.  $\mathbf{h}(\cdot) = |\ell(\cdot)|^2$ . Diese spannen den  $n^2$ -dimensionalen reellen Vektorraum der hermiteschen Formen auf. Die Matrixdarstellung ist ein Produkt ‘Spalte  $\times$  Zeile’  $\ell^* \cdot \ell$ . Der Pullback macht daraus wieder ein Produkt ‘Spalte  $\times$  Zeile’. Die Spalte ist jeweils hermitisch konjugiert zur Zeile, die für die Linearform  $\ell(\cdot)$  steht.

Die für unseren Zweck bequemste Definition ist die auf der Grundlage der Parallelogrammgleichung  $\mathbf{h}(\mathbf{v} + \mathbf{w}) + \mathbf{h}(\mathbf{v} - \mathbf{w}) = 2 \cdot \mathbf{h}(\mathbf{v}) + 2 \cdot \mathbf{h}(\mathbf{w})$ . Auf beiden Wegen ergibt sich sehr einfach die Behauptung.

Wir bemerken zur Übung: Ist  $H$  zur Darstellung von  $\mathbf{h} \in \mathbb{M}^*$  vom Rang 1,  $H =$

$\begin{pmatrix} z^1 \\ z^2 \end{pmatrix}(\bar{z}^1, \bar{z}^2)$  mit  $\text{trace}H = |z^1|^2 + |z^2|^2$ ; und ist  $L$  ebenfalls vom Rang 1, dann haben wir

$$2 \cdot \text{trace}H \cdot E - H = \begin{pmatrix} z^2 \bar{z}^2 & -z^1 \bar{z}^2 \\ -z^2 \bar{z}^1 & z^1 \bar{z}^1 \end{pmatrix} = \begin{pmatrix} z^2 \\ -z^1 \end{pmatrix}(\bar{z}^2, -\bar{z}^1),$$

$$\frac{1}{2} \text{trace}(H \cdot L) = \frac{1}{2} \text{trace}\left(\begin{pmatrix} z^1 \\ z^2 \end{pmatrix}(\bar{z}^1, \bar{z}^2) \begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix}(a_1, a_2)\right) = |a_1 \cdot z^1 + a_2 \cdot z^2|^2,$$

$$(H \parallel L) = \frac{1}{2} \text{trace}\left(\begin{pmatrix} z^2 \\ -z^1 \end{pmatrix}(\bar{z}^2, -\bar{z}^1) \begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix}(a_1, a_2)\right) = |a_1 \cdot z^2 - a_2 \cdot z^1|^2.$$

### Kreise in der komplexen Ebene

Wir haben früher die Kreise  $K \subset \bar{\mathbb{C}}$  durch Gleichungen beschrieben:

$$K = \{z : a \cdot z\bar{z} + \bar{\alpha}z + \alpha\bar{z} + b = 0\} = \{z : a + \bar{\alpha}\frac{1}{z} + \alpha\frac{1}{\bar{z}} + b\frac{1}{z\bar{z}} = 0\}.$$

wobei  $a, b \in \mathbb{R}$ ,  $\alpha \in \mathbb{C}$  mit  $ab - |\alpha|^2 < 0$ .

In Matrixschreibweise

$$z = \frac{z_1}{z_2} \in K \iff (\bar{z}_1, \bar{z}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0 = (\bar{z}_1, \bar{z}_2) \cdot L \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

Die hermitesche Matrix  $L$  ist durch den Kreis bis auf einen Faktor eindeutig bestimmt. Wir haben also eine Bijektion der Menge  $\mathcal{K}$  auf die Menge  $\mathbb{H}_{+-}/\mathbb{R}$  der raumartigen Geraden im Standard-Minkowski-Raum.

Die Gruppenwirkung auf der Menge  $\mathcal{K}$ , die durch die Familie der Möbiustransformationen gegeben ist, kann also auch als Gruppenwirkung auf der Menge der raumartigen Geraden im Minkowskiraum verstanden werden. So, wie die Menge der raumartigen Geraden der hyperbolische Raum ‘ist, so ist’ die Menge der raumartigen Geraden die Menge der Kreise in  $\bar{\mathbb{C}}$ . Wir haben schon früher gesehen, dass die Menge der Geraden im Lichtkegel die erweiterte Zahlenebene  $\bar{\mathbb{C}}$  ist oder (vermöge der stereographischen Projektion) die Riemann’sche Zahlenkugel.

Fazit: Die Gruppe  $SL(2, \mathbb{C})/\{\pm E\}$  wirkt treu auf dem Minkowskiraum (oder besser: auf der Menge der Geraden im Minkowskiraum). Sie hat drei Bahnen. Die Wirkung auf dem hyperbolischen Raum und die Wirkung auf der Riemann’schen Zahlenkugel haben wir bereits ausführlich studiert. Nun wollen wir uns auch noch der Wirkung auf der Menge der Kreise  $\mathcal{K}$  zuwenden.

Es sei  $\mu_A(\cdot)$  die Möbiustransformation zur unimodularen Matrix; man könnte auch genauer schreiben  $\mu_{\pm A}$ .  $\{\mu_A : A \in SL(2, \mathbb{C})/\{\pm E\}\}$  ist zunächst einmal eine Gruppenwirkung auf  $\bar{\mathbb{C}}$  (oder auf der Riemann’schen Zahlenkugel.) Sie ist dann aber auch eine Gruppenwirkung auf der Menge  $\mathcal{K}$ . Wenn es der Klarheit dient, bezeichnen wir diese Bijektion von  $\mathcal{K}$  mit  $\tilde{\mu}_A(\cdot)$ .

$$\mathcal{K} \ni K \longmapsto \tilde{\mu}_A(K) = \{\mu_A(z) : z \in K\} = \{w : \mu_A^{-1}(w) \in K\}.$$

Es gilt offenbar  $\tilde{\mu}_E = id$ ,  $\tilde{\mu}_B(\tilde{\mu}_A(K)) = \tilde{\mu}_{B \cdot A}(K)$ , also  $\tilde{\mu}_B \circ \tilde{\mu}_A(\cdot) = \tilde{\mu}_{B \cdot A}(\cdot)$  usw.

Die Pullback-Abbildung  $\mu_A^*(\cdot)$  bildet Funktionen auf  $\bar{\mathbb{C}}$  in Funktionen auf  $\bar{\mathbb{C}}$  ab. Er bildet somit das Nullstellengebilde  $K_\Psi$  der Funktion  $\Psi(w) = a \cdot w\bar{w} + \bar{\alpha}w + \alpha\bar{w} + b$  in das Nullstellengebilde der zurückgenommenen Funktion  $\mu_A^*(\Psi)$  ab. Er bildet aber nicht die inhomogene hermitische Funktionen  $\Psi$  in hermitische Funktion ab. Die folgende Rechnung ergibt einen Faktor, welcher als störend empfunden werden kann, das Nullstellengebilde aber nicht betrifft.

Wenn  $z = \frac{z^1}{z^2}$ ,  $w = \mu_A(z)$ , dann gilt  $w = \frac{w^1}{w^2}$  mit  $\begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$ , und

$$\Psi(w) = |w^2|^{-2} \cdot (\bar{w}_1, \bar{w}_2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

$$\mu_A^*(\Psi)(z) = \Psi\left(\frac{a_{11}z^1 + a_{12}z^2}{a_{21}z^1 + a_{22}z^2}\right) = |a_{21}z^1 + a_{22}z^2|^{-2} \cdot (\bar{z}^1, \bar{z}^2) \cdot A^* \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}.$$

Für die Nullstellengebilde der Funktionen  $\Psi$  und  $\mu^*(\Psi)$  spielen Faktoren wie hier offenbar keine Rolle. Es liegt daher nahe, die Nullstellengebilde mit Hilfe der dazugehörigen homogenen hermitischen Funktionen zu beschreiben.

$$K'' = \left\{ w = \frac{w^1}{w^2} : \Psi(w) = 0 \right\} = \left\{ w = \frac{w^1}{w^2} : (\bar{w}^1, \bar{w}^2) \cdot \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = 0 \right\},$$

$$K' = \left\{ z = \frac{z^1}{z^2} : (\bar{z}^1, \bar{z}^2) \cdot A^* \begin{pmatrix} a & \bar{\alpha} \\ \alpha & b \end{pmatrix} A \cdot \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = 0 \right\}$$

Damit erhalten wir eine Bijektion der Menge  $\mathcal{K}$  auf die Menge Äquivalenzklassen  $\mathbb{M}_{+-}^*/\mathbb{R}$  von indefiniten hermitischen Formen auf dem Vektorraum  $V$ . Dies wollen wir noch genauer diskutieren. Bisher war von Matrixdarstellungen der hermitischen Formen die Rede, nachdem wir eine Hilbertraumnorm in  $V$  ausgezeichnet hatten; wir brauchten das, da wir in den Konstruktionen auf die Spur Bezug nehmen wollten. Wir kommen jetzt aber nochmal auf die Bemerkung zurück, dass wir keine ausgezeichnete Norm auf  $V$  (d.h. kein ausgezeichnetes positivdefinites Element in  $\mathbb{M}^*$ ) ins Spiel bringen müssen. Wir können vom Produkt der Eigenwerte einer hermitischen Form auf  $V$  auch dann reden, wenn wir uns darauf verständigen, dass wir nur solche Basen in Betracht ziehen, die durch einen Basiswechsel mit der Determinante 1 auseinander hervorgehen. Zu einer hermitischen Form  $\mathbf{h}(\cdot)$  und jeder Basis  $(\mathbf{z}^1(\cdot), \mathbf{z}^2(\cdot))$  von  $V^*$  existiert eine hermitische Matrix  $H$ , sodass

$$\mathbf{h}(\cdot) = (\bar{\mathbf{z}}^1(\cdot), \bar{\mathbf{z}}^2(\cdot)) \cdot H \cdot \begin{pmatrix} \mathbf{z}^1(\cdot) \\ \mathbf{z}^2(\cdot) \end{pmatrix} \quad \text{auf } V.$$

Wenn  $\begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = B \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}$  ( mit  $B \in SL(2, \mathbb{C})/\pm E$ ), dann

$$\mathbf{h}(\cdot) = (\bar{\mathbf{w}}^1(\cdot), \bar{\mathbf{w}}^2(\cdot)) \cdot B^* H B \cdot \begin{pmatrix} \mathbf{w}^1(\cdot) \\ \mathbf{w}^2(\cdot) \end{pmatrix} \quad \text{auf } V \quad \text{mit} \quad \det B^* H B = \det H.$$

Zwei indefinite hermitesche Formen auf  $V$  beschreiben genau dann denselben Kreis auf der Riemann'schen Zahlenkugel, wenn sie sich um einen Faktor unterscheiden. Dies ist die gesuchte Bijektion  $\mathcal{K} \longleftrightarrow \mathbb{M}_{+-}^*$ , die sich nicht auf Matrizendarstellungen stützt, und doch die Determinante (das Produkt der Eigenwerte) im Spiel lässt. Es ist wichtig zu bemerken, dass man in  $\mathcal{K}$  zunächst einmal keine ausgezeichnete Funktion  $\det(\cdot)$  im Blick hat. Erst die Bijektion bringt diese ausgezeichnete Funktion in den Vordergrund. Bei den Möbiustransformationen der Kreise wie bei den eigentlichen Lorentztransformationen des Minkowskiraums bleibt diese Funktion invariant.

## 2 Die Geometrie des Hilbertraums

In der Grundvorlesung ‘Lineare Algebra’ geht es zunächst einmal um Vektoren und Linearformen. Sehr bald kommen aber auch quadratische Funktionen und quadratische Formen ins Bild. Und das auf mehrere Weisen, wie wir sehen werden. Zum ersten lernt man die

**Definition 2.1.** Ein endlichdimensionaler reeller Vektorraum wird zu einem euklidischen Raum, indem man eine positivdefinite quadratische Form auszeichnet; die dazugehörige symmetrische Bilinearform  $\langle \cdot | \cdot \rangle$  heisst das Skalarprodukt des euklidischen Raums.

Es geht hier nicht wirklich um eine quadratische Funktion, wie man sie in der elementaren Analysis kennenlernt, (etwa als eine Näherung zweiter Ordnung für eine glatte Funktion). Es geht vielmehr um eine **quadratische Form** auf einem (möglicherweise unendlichdimensionalen) reellen Vektorraum. Was das ist, werden wir ausführlich erörtern; die hervorstechende Eigenschaft ist die sog. Parallelogrammgleichung.

Wir wollen hier nicht die reellen euklidischen Räume in den Vordergrund stellen sondern ihre komplexen Verwandten, die Prä-Hilberträume, die man in der Linearen Algebra (im endlichdimensionalen Fall) unitäre Räume nennt. Die Struktur kommt hier nicht von einer quadratischen Form sondern von einer **hermitischen Form**, dem Analogon in komplexen Vektorräumen.

**Definition 2.2.** Ein komplexer Vektorraum wird zu einem unitären Raum, indem man eine positiv definite hermitische Form auszeichnet; die dazugehörige Sesquilinearform  $\langle \cdot | \cdot \rangle$  heisst das Skalarprodukt im unitären Raum.

Wir werden uns hier den euklidischen und unitären Räumen von einer Seite nähern, die in der Linearen Algebra nicht üblich ist. Wir wollen unsere Räume zunächst als normierte Vektorräume betrachten. (Man spricht auch von einem Prä-Banachraum, wenn in einem reellen oder komplexen Vektorraum eine Norm ausgezeichnet worden ist.) Das Skalarprodukt, welches man im Spezialfall des euklidischen Raums oder des Prä-Hilbertraums aus der Norm gewinnen kann, ist aus dieser Sicht eine abgeleitete Struktur. Wir wollen nicht durchgehend annehmen, dass unsere normierten Vektorräume endlichdimensional sind. Im unendlichdimensionalen Fall nehmen wir bei Bedarf an, dass die Räume vervollständigt sind, dass wir es also mit veritablen Banach- bzw. Hilberträumen zu tun haben. Der Anfänger muss sich keine Sorgen wegen möglicherweise bescheidener topologischer Vorbildung machen. Bei unserem Vorgehen ist der Bedarf an topologischen Begriffen sehr gering, der Nutzen im Sinne einer Vorbereitung auf höhere Analysis ist aber beträchtlich. Wir werden uns nicht in die Analysis hinein verwickeln lassen: wenn irgendetwas Konstruktion ernsthaft Analysis benötigen, dann werden wir allenfalls die Resultate formulieren. Bei den Beweisen werden wir uns stets auf den endlichdimensionalen Fall konzentrieren.

*Beispiel 2.0.4* (Standard-Hilbertraum). Das einfachste und wichtigste Beispiel eines  $n$ -dimensionalen Hilbertraums ist der Raum der komplexen  $n$ -Spalten  $\mathbf{x}$  mit der Norm  $\|\mathbf{x}\| = \sqrt{\sum |x_k|^2}$ . Wir nennen ihn den Standard-Hilbertraum der Dimension  $n$  und bezeichnen

ihn mit  $\mathbb{C}_{St}^n$ . Das dazugehörige innere Produkt lässt sich bequem in der Notation der Matrizenrechnung ausdrücken:

$$\langle \mathbf{y} | \mathbf{x} \rangle = \mathbf{y}^* \cdot \mathbf{x} = \sum \bar{y}_k \cdot x_k. \quad (\mathbf{y}^* \text{ ist die Zeile mit den Einträgen } \bar{y}_k.)$$

*Beispiel 2.0.5* (Trigonometrische Polynome). Ein unendlichdimensionales Beispiel ist der Vektorraum der trigonometrischen Polynome, welchen man zu einem Prä-Hilbertraum  $(\mathcal{T}, \langle \cdot | \cdot \rangle)$  macht mit dem Skalarprodukt

$$\langle g | f \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{g}(t) \cdot f(t) dt = \sum \bar{d}_n \cdot c_n \quad \text{für} \quad f(t) = \sum c_n e^{it}, \quad g(t) = \sum d_n e^{it}.$$

Die analytischen Fragen, die dieser Raum aufwirft, etwa die konkrete Vervollständigung und die Konvergenz von Fourier-Reihen, werden uns hier nicht beschäftigen.

**Figuren und Bewegungen:** Inwiefern geht es im Folgenden um Geometrie?

Wir haben oben gesagt, dass zu einer Geometrie zunächst einmal die besonderen Figuren und Bewegungen gehören. In einer analytischen Geometrie kommen noch die besonderen Funktionen dazu. Wir stellen uns hier in unserer elementaren 'Geometrie des Hilbertraums' auf den Standpunkt, dass die interessierenden Figuren die abgeschlossenen Teilvektorräume und die Orthonormalsysteme sind. Die Bewegungen sind die bijektiven linearen Isometrien. (Im Standardbeispiel  $\mathbb{C}_{St}^n$  sind das die unitären  $n \times n$ -Matrizen  $U$ , die in der üblichen Weise auf Spalten wirken.) Die besonderen Funktionen sind einerseits die linearen Funktionen (bzw. die beschränkten Linearformen) und andererseits die hermiteschen Formen (bzw. die selbstadjungierten endlichdimensionalen) Operatoren. Eine besondere Rolle spielen auch die quadratischen Funktionen im euklidischen Raum.

## 2.1 Einordnung in die Welt der normierten Vektorräume

Der Begriff der Linearform und die Idee der Vektorraumdualität sollten aus der Vorlesung 'Lineare Algebra' bekannt sein, und die Begriffe Metrik und Norm aus der Vorlesung 'Analysis I'. Wir bringen sie in einer Weise zusammen, die uns nicht nur hier in der Geometrie des Hilbert-Raums nützlich sein wird, sondern auch in der sog. konvexen Geometrie, die wir im nächsten Kapitel ansprechen wollen.

**Definition 2.3.** (Norm)

Ein reeller oder komplexer Vektorraum  $V$  wird zu einem normierten Vektorraum  $(V, \|\cdot\|)$ , indem man eine Norm auszeichnet.

Eine Norm ist eine nichtnegative Funktion  $\|\cdot\|$  mit den Eigenschaften

- (i)  $\|a \cdot v\| = |a| \cdot \|v\|$  für alle Skalare  $a$  ('Absolute Homogenität')
- (ii)  $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$  ('Subadditivität')
- (iii)  $\|v\| = 0 \iff v = 0$ .

**Satz 2.1.1.** (*Metrik und Vervollständigung:*)

Eine Norm auf  $V$  liefert eine translationsinvariante Metrik auf den affinen Räumen mit dem Tangentialraum  $V$ , wenn man den Abstand  $d(P, Q)$  als die Norm des Verschiebungsvektors  $\overrightarrow{PQ}$  definiert.

Da man jeden metrischen Raum  $(E, d(\cdot, \cdot))$  vervollständigen kann, kann man insbesondere jeden normierten Vektorraum vervollständigen. Die Operation der Linearkombination überträgt sich in stetiger Weise auf die Vervollständigung.

**Sprechweise.** (Banach-Raum) Ein vollständiger normierter (reeller oder komplexer) Vektorraum heisst ein Banach-Raum.

**Definition 2.4.** (Stetige Linearformen und Duale Norm )

$(V, \|\cdot\|)$  sei ein normierter Vektorraum. Eine Linearform  $\ell(\cdot)$  auf  $V$  (auch notiert  $\langle \ell, \cdot \rangle$ ) heisst eine stetige Linearform, wenn sie auf der Einheitskugel beschränkt ist. Ihr Supremum auf der Einheitskugel heisst die Norm der Linearform und wird mit  $\|\ell\|$  bezeichnet.

$$\|\ell\| = \sup\{|\langle \ell, v \rangle| : \|v\| \leq 1\} = \sup\left\{\frac{|\langle \ell, v \rangle|}{\|v\|} : v \in V\right\}.$$

**Bemerke :** Die ursprüngliche Norm  $\|\cdot\|$  und die duale Norm  $\|\cdot\|$  leben auf ganz verschiedenen Räumen; es besteht keine Verwechslungsgefahr, wenn man sie mit demselben Symbol  $\|\cdot\|$  bezeichnet.

Eine stetige Linearform auf einem Prä-Banachraum ist gleichmäßig stetig; sie lässt sich daher (auf eindeutige Weise) zu einer stetigen Linearform auf der Vervollständigung fortsetzen.

In der Theorie der Banachräume beweist man, dass es auf einem normierten Vektorraum viele stetige Linearformen gibt. Eine Variante des berühmten Satzes von Hahn-Banach (die wir später in der konvexen Geometrie beweisen werden) besagt:

**Satz 2.1.2** (Satz von Hahn-Banach). Sei  $(V, \|\cdot\|)$  ein normierter Vektorraum und  $\hat{v}$  ein Element mit  $\|\hat{v}\| = 1$ . Dann existiert eine stetige Linearform  $\hat{\ell}$  mit

$$\hat{\ell}(\hat{v}) = 1 \quad \text{und} \quad \hat{\ell}(v) \leq \|v\| \quad \text{für alle } v \in V.$$

**Corollar:** Es gilt also nicht nur  $|\langle \ell, v \rangle| \leq \|\ell\| \cdot \|v\|$  für alle  $v$ , sondern sogar

$$\|v\| = \sup\left\{\frac{|\langle \ell, v \rangle|}{\|\ell\|} : \ell \text{ stetige Linearform}\right\} \quad \text{für alle } v \in V.$$

**Notation.** Der normierte Raum aller stetigen Linearformen auf dem normierten Raum  $(V, \|\cdot\|)$  heisst der (topologische) Dualraum; er wird mit  $(V^*, \|\cdot\|)$  bezeichnet.

Hinweis: Eine vollständigere Version des Satzes von Hahn-Banach, die wir hier aber nicht benötigen, sagt folgendes

**Satz.** *Wenn auf einem Teilraum  $\hat{V}$  von  $(V, \|\cdot\|)$  eine Linearform gegeben ist, die auf der Einheitskugel durch 1 beschränkt ist, dann gibt es eine Fortsetzung auf den ganzen Raum, welche gleichfalls auf der Einheitskugel durch 1 beschränkt ist.*

Bei manchen normierten Vektorräumen ist es schwierig, konkretes Wissen über die Menge aller stetigen Linearformen zu gewinnen. Für manche Belange ist es dann bequemer, mit einem genügend großen Teilraum des Dualraums zu argumentieren.

**Sprechweise.** (Dualitätspaar) Es seien  $(V, \|\cdot\|)$  und  $(F, \|\cdot\|)$  normierte Vektorräume. Eine Bilinearform  $(\cdot, \cdot)$  auf  $F \times V$  macht  $(F, V)$  zu einem Dualitätspaar, wenn gilt

$$\begin{aligned}\|f\| &= \sup\{|(f, v)| : \|v\| \leq 1\} \quad \text{für alle } f \in F, \\ \|v\| &= \sup\{|(f, v)| : \|f\| \leq 1\} \quad \text{für alle } v \in V.\end{aligned}$$

*Beispiel 2.1.1.* Es sei  $V = \ell^\infty$  der Raum aller beschränkten reellen (oder komplexen) Zahlenfolgen mit der Supremumsnorm, (die man bekanntlich auch die Norm der gleichmässigen Konvergenz nennt). Die summablen Folgen  $\mathbf{p} = (p_j)_j$  mit der Norm  $\|\mathbf{p}\| = \sum |p_j| < \infty$  sind geeignet, die  $v \in V$  auszuwerten:  $(\mathbf{p}, v) = \sum p_j \cdot x_j$ . Wir haben ein Dualitätspaar, wie man leicht sieht. Man kann zeigen, dass es stetige Linearformen auf  $\ell^\infty$  gibt, die nicht durch ein  $\mathbf{p} \in \ell^1$  gegeben sind. Diese müssen uns aber nicht interessieren.

*Beispiel 2.1.2.* Berühmte Beispiele von übersichtlicher Art liefern die Räume  $\ell^p$  der  $p$ -summablen Folgen für  $p \in (1, \infty)$ . Die Norm ist hier  $\|v\|_p = (\sum |x_j|^p)^{1/p}$ . Zu  $p$  gehört  $q$ , sodass  $1/p + 1/q = 1$ .

**Satz 2.1.3** (Hölders Ungleichung).

a) *Wenn man den Raum der finiten Folgen einmal mit der  $p$ -Norm ausstattet und das andere Mal mit der  $q$ -Norm, dann erhält man ein Dualitätspaar. Für alle endlichen Folgen  $v$  und  $w$  gilt*

$$\begin{aligned}\sup\{|\sum w_j \cdot v_j| : \sum |v_j|^p \leq 1\} &= \|w\|_q, \\ \sup\{|\sum w_j \cdot v_j| : \sum |w_j|^q \leq 1\} &= \|v\|_p.\end{aligned}$$

b) *Der Dualraum zu  $(V = \ell^p, \|\cdot\|_p)$  ist  $(V^* = \ell^q, \|\cdot\|_q)$ . Das bedeutet: jede stetige Linearform auf dem Raum der  $p$ -summablen Folgen ist durch eine  $q$ -summable Folge gegeben. Und es gilt*

$$|\sum w_j \cdot v_j| \leq \left(\sum |w_j|^q\right)^{1/q} \cdot \left(\sum |v_j|^p\right)^{1/p}.$$

Die Beweise werden in den Anfängervorlesungen geführt.

Der Spezialfall  $p = 2 = q$  (mit der Ungleichung von Cauchy – Schwarz – Bunyakovski) ist relevant für die Hilbertraum-Geometrie. Er enthält nämlich die Aussage, dass der Raum der quadratsummablen Folgen vollständig ist, weil der Dualraum eines normierten Raums in jedem Falle vollständig ist.

## Sesquilinearformen und Skalarprodukt

**Definition 2.5.** Eine hermitesche Sesquilinearform ist eine komplexwertige Funktion  $\Phi(\cdot, \cdot)$  auf  $V \times V$  mit den Eigenschaften

- (i)  $\Phi(v, w) = \overline{\Phi(w, v)}$
- (ii)  $\Phi(v, \alpha w) = \alpha \cdot \Phi(v, w)$
- (iii)  $\Phi(v, w_1 + w_2) = \Phi(v, w_1) + \Phi(v, w_2)$ .

Eine reellwertige Funktion  $\Psi(\cdot)$  auf einem komplexen Vektorraum  $V$  heisst eine hermitesche Form, wenn es eine hermitesche Sesquilinearform  $\Phi(\cdot, \cdot)$  gibt, sodass  $\Phi(v, v) = \Psi(v)$ .

Eine Sesquilinearform heisst positivdefinit, wenn  $\Psi(v) = \Phi(v, v) > 0$  für alle  $v \neq 0$ .

**Sprechweise.** Eine positivdefinite Sesquilinearform ist nach allgemeinem Sprachgebrauch immer auch hermitisch, so wie eine positivdefinite Matrix immer auch hermitisch ist.

**Definition 2.6.** (Skalarprodukt und Hilbertraumnorm)

Ein komplexer Vektorraum  $V$  wird zu einem unitären Raum (oder Prähilbertraum), indem eine positiv definite hermitesche Form ausgezeichnet wird; die dazugehörige Sesquilinearform  $\langle \cdot | \cdot \rangle$  heisst das Skalarprodukt im unitären Raum. Die Quadratwurzel aus der hermiteschen Form heisst die dazugehörige Hilbertraumnorm.

$$\|v\| = \sqrt{h(v)} = \langle v | v \rangle^{1/2}.$$

Ein vollständiger unitärer Raum heisst ein Hilbertraum. Ein Hilbertraum, in welchem es eine überall dichte Teilmenge gibt, heisst ein separabler Hilbertraum.

Wir sind in der Tat berechtigt, von einer Norm zu sprechen.  $\sqrt{h(\cdot)}$  ist offenbar abso-luthomogen und auch subadditiv, wie die folgende Überlegung zeigt:

**Satz 2.1.4.**

Für ein Skalarprodukt und die dazugehörige Hilbertraumnorm  $\|\cdot\| = \sqrt{h(\cdot)}$  gilt

$$\begin{aligned} |\langle w | v \rangle| &\leq \|w\| \cdot \|v\|, && \text{('Cauchy - Schwarz - Ungl.')} \\ \|w + v\| &\leq \|w\| + \|v\|, && \text{('Dreiecksungleichung')} \\ \|w + v\|^2 + \|w - v\|^2 &= 2 \cdot \|w\|^2 + 2 \cdot \|v\|^2. && \text{('Parallelogrammgl.')} \end{aligned}$$

**Beweis.** Wir können uns auf den Fall beschränken, wo  $\langle w | v \rangle$  positiv ist. Für alle reellen Zahlen  $a$  gilt  $0 \leq \langle v + aw | v + aw \rangle = \|v\|^2 + 2a \cdot \langle w | v \rangle + a^2 \cdot \|w\|^2$ . Eine quadratische Funktion  $A + 2aB + a^2C$  ist nur dann auf der reellen Achse nichtnegativ, wenn  $B^2 \leq AC$ . Das beweist die Schwarz'sche Ungleichung. Die Dreiecksungleichung folgt; denn

$$(\|w\| + \|v\|)^2 = \|w\|^2 + \|v\|^2 + 2 \cdot \|w\| \cdot \|v\| \geq \|w\|^2 + \|v\|^2 + 2 \cdot |\langle w | v \rangle| \geq \langle w+v | w+v \rangle.$$

Die Parallelogrammgleichung ergibt sich durch eine triviale Rechnung.

Die Parallelogrammgleichung charakterisiert die Hilbertraumnormen. Es gilt der

**Satz 2.1.5.** (Satz von Jordan und von Neumann)

Wenn eine Norm in einem komplexen Vektorraum die Parallelogrammgleichung erfüllt, dann existiert genau eine positivdefinite Sesquilinearform  $\langle \cdot | \cdot \rangle$ , sodass  $\langle v | v \rangle = \|v\|^2$ .

Auf den Beweis kommen wir an anderer Stelle zurück.

Die Frage nach den stetigen Linearformen beantwortet sich für unitäre Räume erstaunlich einfach. Die stetigen Linearformen sind gleichmässig stetig und lassen sich daher (in eindeutiger Weise!) zu stetigen Linearformen auf der Vervollständigung fortsetzen; und diese Linearformen stehen in eineindeutiger Beziehung zu den Elementen des vervollständigten Raums.

**Satz 2.1.6.** (Satz von F. Riesz)

Sei  $(\hat{V}, \langle \cdot | \cdot \rangle)$  ein Prä-Hilbertraum und  $(V, \langle \cdot | \cdot \rangle)$  seine Vervollständigung. Zu jeder stetigen Linearform  $\ell(\cdot)$  auf  $V$  gibt es genau ein  $w$  in der Vervollständigung  $V$  von  $\hat{V}$ , sodass  $\ell(\cdot) = \langle w | \cdot \rangle$ .

**Beweis.** Es sei  $\ell(\cdot)$  eine stetige Linearform, nicht identisch 0, und  $\hat{v}$  ein Vektor ausserhalb des abgeschlossenen Teilraums  $L = \{v : \ell(v) = 0\}$ . Es gilt  $d = \inf\{\|\hat{v} - v\| : v \in L\} > 0$ . Wir finden eine Folge  $(v_n)_n$  in  $L$  mit  $\lim\|\hat{v} - v_n\| = d$ , und wir zeigen, dass es sich um eine Cauchy-Folge handelt, eine Folge also, die einen Limes  $\tilde{v} \in L$  besitzt. Die Cauchy-Eigenschaft ergibt sich mit Hilfe der Parallelogrammgleichung folgendermaßen: Es seien  $v_n, v_m \in L$  mit  $\|\hat{v} - v_n\|^2 < (d + \varepsilon)^2$ ,  $\|\hat{v} - v_m\|^2 < (d + \varepsilon)^2$ . Wegen  $\frac{1}{2}(v_n + v_m) \in L$  gilt

$$\begin{aligned} 4 \cdot d^2 &\leq 4 \cdot \left\| \hat{v} - \frac{1}{2}(v_n + v_m) \right\|^2 = \|\hat{v} - v_n + \hat{v} - v_m\|^2 \\ &= 2\|\hat{v} - v_n\|^2 + 2\|\hat{v} - v_m\|^2 - \|(\hat{v} - v_n) - (\hat{v} - v_m)\|^2 \leq 4(d + \varepsilon)^2 - \|v_n - v_m\|^2. \end{aligned}$$

also  $\frac{1}{4}\|v_n - v_m\|^2 \leq (d + \varepsilon)^2 - d^2$ .

Den Grenzwert  $\tilde{v} \in L$  der Cauchy-Folge nennen wir den Fusspunkt des Lots von  $\hat{v}$  auf den Teilraum  $L$ , die Differenz  $\hat{w} = \hat{v} - \tilde{v}$  nennen wir das Lot. Für jedes  $v \in L$  gilt  $\langle \hat{w} | v \rangle = 0$ ; denn es gilt für alle Skalare  $a$

$$d^2 \geq \|\hat{v} - (\tilde{v} + av)\|^2 = \langle \hat{w} - av | \hat{w} - av \rangle = d^2 - 2\Re(a\langle \hat{w} | v \rangle) + |a|^2 \cdot \|v\|^2,$$

und das impliziert  $0 = \langle \hat{w} | v \rangle$ . Für jedes  $v \in V$  gilt offenbar  $\ell(\hat{w})v - \ell(v)\hat{w} \in L$ , also

$$\langle \hat{w} | \ell(\hat{w})v - \ell(v)\hat{w} \rangle = 0; \quad \ell(v) = \frac{\ell(\hat{w})}{\langle \hat{w} | \hat{w} \rangle} \cdot \langle \hat{w} | v \rangle.$$

Und das liefert die gesuchte Darstellung der Linearform  $\ell(\cdot) = \langle c\hat{w} | \cdot \rangle$  mit  $\bar{c} = \frac{\ell(\hat{w})}{\langle \hat{w} | \hat{w} \rangle}$ .

In der höheren Theorie der Hilberträume interessiert man sich durchaus auch für gewisse unbeschränkte Linearformen.

*Beispiel 2.1.3* (Dirichlet-Kerne). Es sei  $(\mathcal{T}, \langle \cdot | \cdot \rangle)$  der Prä-Hilbertraum der trigonometrischen Polynome. Die Auswertung der Polynome in einem festen Punkt  $t_0 \in \mathbb{R}/2\pi$  ist eine unbeschränkte Linearform  $\ell_{t_0}$ . Es gibt eine Folge von trigonometrischen Polynomen  $v_N$  mit  $\|v_N\| \leq 1$  und  $\ell_{t_0}(v_N) \rightarrow \infty$ . Ein Beispiel liefern die sog. Dirichlet-Kerne  $D_N(t)$  und die Fejér-Kerne, die uns auch noch bei anderen Gelegenheiten (z. B. auf Seite ??) begegnen werden.

$$\begin{aligned} D_N(t) &= \sum_{n=-N}^N e^{int} = 1 + 2 \cdot \cos t + 2 \cdot \cos 2t + \cdots + 2 \cdot \cos Nt = \\ &= (e^{it/2} - e^{-it/2})^{-1} \cdot (e^{i(N+1/2)t} - e^{-i(N+1/2)t}) = \frac{\sin(N+1/2)t}{\sin t/2}. \\ F_N(t) &= \frac{1}{N} (D_0(t) + D_1(t) + \cdots + D_{N-1}(t)) = \frac{1}{N} \left( \frac{\sin Nt/2}{\sin t/2} \right)^2 \\ &= 1 + 2(1 - \frac{1}{N}) \cos t + 2(1 - \frac{2}{N}) \cos 2t + \cdots + 2(1 - \frac{N-1}{N}) \cos(N-1)t \end{aligned}$$

Es gilt  $\|D_N\| = \sqrt{2N+1}$ . Die Auswertung im Nullpunkt wächst schneller mit  $N \rightarrow \infty$ , nämlich  $D_N(0) = 2N+1$ .

### Ergänzungen zum Beispiel

1) Im Hinblick auf spätere Beispiele ist es geboten, die  $2\pi$ -periodischen Funktionen  $D_N(t)$  und  $F_N(t)$  näher anzuschauen. Man bemerke z. B.

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) dt = 1; \quad \text{und daraus} \quad \frac{\pi}{2} = \int_0^{2\pi} \frac{\sin(2N+1)s}{\sin s} ds.$$

2) Interessant ist auch die  $2\pi$ -periodische Funktion  $s_N(t)$ , die im Intervall  $(0, 2\pi)$  eine Stammfunktion ist für die Funktion  $\frac{1}{2}(D_N(t) - 1) = \cos t + \cos 2t + \cdots + \cos Nt$ .

$$s_N(t) = \int_{-\pi}^t (\cos s + \cos 2s + \cdots + \cos Ns) ds = \sin t + \frac{1}{2} \sin 2t + \cdots + \frac{1}{N} \sin Nt.$$

Der Limes wird üblicherweise die Euler'sche Sägezahn genannt:

$$s(t) = \frac{1}{2}(\pi - t) = \sin t + \frac{1}{2} \sin 2t + \frac{1}{3} \sin 3t + \cdots \quad \text{für } t \in (0, 2\pi).$$

Die unstetige Funktion  $s(t)$  springt von  $-\frac{\pi}{2}$  nach  $\frac{\pi}{2}$  an den Stellen  $t = 2\pi \cdot k$ , ( $k \in \mathbb{Z}$ ).

3) Die Fejér-Kerne sind positiv mit  $\frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(t) dt = 1$ . Man bemerke auch

$$F_{2N+1}(t) = \frac{1}{2N+1} \left( \frac{\sin(2N+1)t/2}{\sin t/2} \right)^2 = \frac{1}{2N+1} (D_N(t))^2.$$

Die Varianz der zentrierten Wahrscheinlichkeitsdichten  $\frac{1}{2\pi} \cdot F_N(t) dt$  auf dem Intervall  $(-\pi, \pi)$  verhält sich wie  $C/N$  für  $N \rightarrow \infty$ , denn

$$N \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2 \cdot F_N(t) dt = N \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \frac{t}{\sin(t/2)} \right)^2 \cdot \sin^2(Nt/2) dt \approx 2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \frac{t/2}{\sin(t/2)} \right)^2 dt,$$

4) Die Fejér-Kerne wirken als ‘Glättungskerne’ für die integrierbaren  $2\pi$ -periodischen Funktionen vermöge

$$f(\cdot) \longmapsto f_N(\cdot) = f \star F_N(\cdot), \quad f_N(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-s) \cdot F_N(s) ds.$$

Wenn speziell  $f(\cdot)$  ein trigonometrisches Polynom ist, dann ist auch  $f_N(\cdot)$  ein trigonometrisches Polynom, und zwar ein trigonometrisches Polynom vom Grad  $< N$ .

$$f(t) = \sum c_n \cdot e^{int} \implies f \star F_N(t) = \sum \left(1 - \frac{|n|}{N}\right)^+ \cdot c_n \cdot e^{int}.$$

5) Wir überlassen es der Veranstaltung zur Analysis, zu untersuchen, was man über die Konvergenz der Funktionenfolge  $f \star F_N$  sagen kann, wenn  $f$  eine nicht sehr reguläre  $2\pi$ -periodische Funktion ist. Dirichlet hat darüber einen eindrucksvollen Satz bewiesen. Das Thema erscheint auch in das Umfeld eines berühmten Satzes von Herglotz über positivdefinite Folgen.

## 2.2 Die Struktur des Hilbertraums. Orthogonalität

Wir erinnern an eine Begriffsbildung aus der Welt der  $\mathbb{K}$ -Vektorräume, die wir dann mutatis mutandis auf die Welt der Hilberträume übertragen wollen.

**Sprechweise 2.2.1.** (Duale Paarung, Annihilatoren)

Es seien  $V$  und  $F$   $\mathbb{K}$ -Vektorräume und  $(\cdot, \cdot)$  eine nichtausgeartete Bilinearform, d. h. eine  $\mathbb{K}$ -wertige Funktion auf  $F \times V$  mit

$$\forall f \in F : (f, v) = 0 \implies v = 0, \quad \forall v \in V : (f, v) = 0 \implies f = 0$$

Man spricht dann von einer dualen Paarung oder einem Dualitätssystem. Und man definiert zu jeder Teilmenge  $W$  von  $V$  sowie zu jeder Teilmenge  $G$  von  $F$  den Annihilator

$$\begin{aligned} W^\dagger &= \{f : (f, w) = 0 \text{ für alle } w \in W\} && \subseteq F \\ G^\dagger &= \{v : (g, v) = 0 \text{ für alle } g \in G\} && \subseteq V \end{aligned}$$

Ein Annihilator ist stets ein Teilvektorraum. In der Theorie der linearen Gleichungssysteme beweist man den

**Satz 2.2.1** (Durchschnitt und Verbindung in einem Dualitätssystem).

*Es sei  $(F, V, (\cdot, \cdot))$  ein  $n$ -dimensionales Dualitätssystem.*

*Wenn  $W \subseteq V$  (bzw.  $G \subseteq F$ ) einen  $r$ -dimensionalen Teilraum aufspannt, dann ist  $W^\dagger$  (bzw.  $G^\dagger$ ) ein  $n - r$ -dimensionaler Teilraum im 'dualen' Raum.*

*Für Teilvektorräume  $W_1, W_2 \subseteq V$  gilt*

$$\begin{aligned} (1) \quad & (W_1 + W_2)^\dagger = W_1^\dagger \cap W_2^\dagger, \\ (2) \quad & (W_1 \cap W_2)^\dagger = W_1^\dagger + W_2^\dagger, \\ (3) \quad & W^{\dagger\dagger} = W. \end{aligned}$$

*Die entsprechenden Gleichheiten gelten für Teilvektorräume  $G_1, G_2 \subseteq F$ .*

Hinweise: Wenn man es mit einer symmetrischen nichtausgearteten Bilinearform auf  $V \times V$  zu tun haben, dann identifiziert man  $W^\dagger$  mit dem sog. orthogonalen Komplement, und man benützt die Bezeichnung  $W^\perp$ . Der bekannteste Fall ist der Fall eines endlichdimensionalen euklidischen Raums.

In der symplektischen Geometrie (und in der klassischen Mechanik) trifft man auf antisymmetrische nichtausgeartete Bilinearformen (auf  $2n$ -dimensionalen reellen Vektorräumen). Die Bezeichnungen für den Annihilator scheinen da nicht einheitlich zu sein. (Bei V.I. Arnold findet sich das paradoxe Wort skew-orthogonal complement.)

Wir werden uns damit nicht aufhalten. Unser Thema sind hier allein die unitären Räume mit einer ausgezeichneten nichtausgearteten Sesquilinearform, die wir mit  $\langle \cdot | \cdot \rangle$  bezeichnen, (wobei Linearität im zweiten Argument verlangen).

Wir formulieren noch einmal den oben bewiesenen Satz von F. Riesz über den Dualraum eines Prä-Hilbertraums:

**Satz 2.2.2.** Sei  $(\bar{V}, \|\cdot\|)$  ein Prä-Hilbertraum und  $(V, \|\cdot\|)$  seine Vervollständigung. Für jedes  $w \in V$  ist  $\langle w|\cdot\rangle$  eine stetige Linearform. Und zu jeder stetigen Linearform  $\ell(\cdot)$  auf  $\bar{V}$  gibt es genau ein  $w \in V$  sodass  $\ell(\cdot) = \langle w|\cdot\rangle$ .

**Notation. (Dirac's Notation)** Nach einem weithin hochgeschätzten Vorschlag von P. Dirac notiert man  $\langle w| = \langle w|\cdot\rangle$ ; und man nennt  $\langle w|$  den bra-Vektor zum Vektor  $w$ . Die Vektoren  $v$  nennt man auch die ket-Vektoren; und man notiert  $|v\rangle = v$ . Ein 'bra' zusammen mit einem 'ket' liefert eine Klammer ('bracket'), und das ist der Skalar  $\langle w|v\rangle$ .

**Notation 2.1.** Wenn  $U$  eine Menge von Vektoren in einem Prä-Hilbertraum ist und  $w$  ein Vektor mit  $\langle w|u\rangle = 0$  für alle  $u \in U$ , wenn also  $\langle w| \in U^\dagger$ , dann sagt man, dass  $w$  auf  $U$  senkrecht steht. Man notiert  $w \in U^\perp$  oder auch  $w \perp U$ .

**Sprechweise 2.2.2.** Für einen abgeschlossenen linearen Raum  $W$  in einem Hilbertraum heisst  $W^\perp$  das orthogonale Komplement. — Komplement heisst bekanntlich Ergänzungstück. Der Name orthogonales Komplement wird durch die folgenden Sätze legitimiert.

**Satz 2.2.3.** (Die Isometrie  $V \longleftrightarrow V^*$ )

Wenn  $(V, \|\cdot\|)$  ein Hilbertraum ist, dann ist die Abbildung  $V \ni w \longleftrightarrow \langle w| \in V^*$  (Dualraum) eine bijektive Isometrie zweier Hilberträume, welche konjugiert-linear ist in dem Sinn

$$\langle w_1 + w_2| = \langle w_1| + \langle w_2|; \quad \langle w_1 \cdot a| = \bar{a} \cdot \langle w_1|.$$

(Eine Abbildung eines metrischen Raums in einen metrischen Raum heisst eine Isometrie, wenn die Bilder von  $P$  und  $Q$  denselben Abstand haben wie  $P$  und  $Q$ .)

Die bijektive Isometrie gemäß dem Satz von F. Riesz liefert zum Annihilator einer Menge von Vektoren, den sog. Orthogonalraum der Menge.

**Satz 2.2.4.** (Orthogonalraum)

Es sei  $(V, \|\cdot\|)$  ein Hilbertraum; d. h. ein vollständiger normierter Raum, dessen Norm die Parallelogrammgleichung erfüllt.  $\langle \cdot|\cdot\rangle$  sei das Skalarprodukt.

Der Annihilator  $W^\dagger$  einer Teilmenge  $W \subset V$  ist ein abgeschlossener Teilraum des Dualraums  $(V^*, \|\cdot\|)$  (mit der dualen Norm).

Dazu gehört als ein abgeschlossener Teilraum von  $V$  der sog. Orthogonalraum zu  $W$

$$W^\perp = \{u : \langle u|w\rangle = 0 \text{ für alle } w \in W\} = \{u : \langle u| \in W^\dagger\}$$

**Satz 2.2.5.** Im Hilbertraum  $(V, \|\cdot\|)$  sei  $W$  ein abgeschlossener linearer Teilraum. Jedes  $v \in V$  besitzt dann genau eine Zerlegung

$$v = v_0 + v_1 \quad \text{mit} \quad v_0 \in W, \quad v_1 \in W^\perp.$$

Die Abbildung  $\pi_W : v \rightarrow v_0$  ist eine lineare Projektion mit

$$\langle u|\pi_W v\rangle = \langle \pi_W u|v\rangle \quad \text{für alle } u, v \in V.$$

Die Eigenschaft, welche die orthogonalen Projektionen vor den übrigen idempotenten Endomorphismen auszeichnet, heisst die Selbstadjungiertheit.

**Definition 2.7.** Ein Endomorphismus  $\varphi$  eines Hilbertraums  $V$  heisst selbstadjungiert, wenn gilt

$$\langle u | \varphi(v) \rangle = \langle \varphi(v) | v \rangle \quad \text{für alle } u, v \in V.$$

**Beweis 2.2.1.** Der Beweis des Satzes funktioniert wie beim Satz von Riesz 2.1.6 auf Seite 87. Es gibt auf der abgeschlossenen Menge  $W$  genau einen Punkt  $v_0 = \pi_W(v)$  mit minimalem Abstand zu  $v$ , den ‘Fußpunkt des Lots auf  $W$ ’; das ‘Lot’  $v_1 = v - v_0$  ist dadurch gekennzeichnet, dass es auf  $W$  senkrecht steht. Die Zuordnung  $\pi_W$  ist linear; denn wenn zwei Vektoren auf  $W$  senkrecht stehen, dann auch ihre Summe.

$$v - v_0 \perp W, \quad w - w_0 \perp W \implies (v + w) - (v_0 + w_0) \perp W.$$

Es gilt  $V = W + W^\perp$ ; und es handelt sich hierbei nicht nur um eine ‘direkte Zerlegung’ im Sinne der allgemeinen Theorie der Vektorräume. Wir wissen nicht nur  $W \cap W^\perp = \{0\}$ ; wir wissen auch, dass jeder Vektor in dem einen Summanden senkrecht steht auf jedem Vektor im anderen Summanden. Man spricht von einer orthogonalen Zerlegung.

Wenn in diesem Sinne  $u = u_0 + u_1$ ;  $v = v_0 + v_1$ . dann gilt  $\langle u | v_0 \rangle = \langle u_0 | v_0 \rangle = \langle u_0 | v \rangle$ .

**Sprechweise 2.2.3.** Der kleinste abgeschlossene lineare Teilraum, welcher die Menge  $W$  umfasst, heisst die abgeschlossene lineare Hülle von  $W$ . Wir bezeichnen ihn mit  $\overline{\text{span}} W$ .

Bemerkung: Die abgeschlossene lineare Hülle ist der Durchschnitt aller abgeschlossenen linearen Teilräume, die  $W$  umfassen. Sie kann aber auch charakterisiert werden als die abgeschlossene Hülle des von  $W$  aufgespannten Teilvektorraums. Die Begriffsbildung ist in allen Banachräumen vernünftig.

Im Falle der Hilberträume werden wir eine weitere interessante Charakterisierung der abgeschlossenen linearen Hülle einer Menge von Vektoren kennenlernen. Wir werden nämlich zeigen: Es gilt  $\overline{\text{span}} W = W^{\perp\perp}$  (für eine beliebige Menge  $W$ .)

**Sprechweise 2.2.4.** Seien  $W_1$  und  $W_2$  Teilmengen eines Hilbertraums mit  $\langle w_1 | w_2 \rangle = 0$  für alle  $w_1 \in W_1$ ,  $w_2 \in W_2$ . Man sagt dann, dass  $W_1$  und  $W_2$  aufeinander senkrecht stehen, und man notiert  $W_1 \perp W_2$ . Man bemerke  $W_1 \perp W_2 \iff \overline{\text{span}} W_1 \perp \overline{\text{span}} W_2$ .

**Satz 2.2.6.** Für jede Teilmenge  $W$  eines Hilbertraums ist  $W^\perp$  ein abgeschlossener linearer Teilraum. Es gilt  $W_1 \subseteq W_2 \implies W_1^\perp \supseteq W_2^\perp$ ,  $W^{\perp\perp} = \overline{\text{span}} W$ .

**Beweis 2.2.2.** Die beiden ersten Aussagen sind trivial.  $W^{\perp\perp}$  ist eine abgeschlossene lineare Obermenge von  $W$ ; sie umfasst daher  $\overline{\text{span}} W$ . Wir zeigen für einen abgeschlossenen linearen Teilraum  $W$ , dass  $W^{\perp\perp}$  nicht echt größer sein kann. Wir zeigen: Wenn  $u \notin W$ , dann  $u \notin W^{\perp\perp}$ . Sei also  $u$  ein Vektor, der nicht zum abgeschlossenen linearen Teilraum  $W$  gehört, und sei  $u_0 \in W$  der Fußpunkt des Lots. Das nichtverschwindende Lot  $u - u_0$  steht auf allen Elementen von  $W$  senkrecht.  $0 \neq u - u_0 \in W^\perp$ .  $u$  steht nicht auf allen  $v \in W^\perp$  senkrecht,  $u$  steht nämlich nicht auf  $u - u_0$  senkrecht wegen  $\langle u - u_0 | u \rangle = \langle u - u_0 | u - u_0 \rangle \neq 0$ . Also gilt  $u \notin W^{\perp\perp}$ .

Wir haben zu einem beliebigen abgeschlossenen linearen Teilraum  $W$  ( $W^{\perp\perp} = W$ ) eine orthogonale Zerlegung gefunden,  $V = W \oplus W^\perp$  gefunden. Für das weitere werden wir feinere orthogonale Zerlegungen benötigen. Wir erinnern zuerst an einige Begriffsbildungen aus der Linearen Algebra, die nichts mit der Hilbertraumstruktur zu tun haben.

**Definition 2.8.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Seien  $V_1, V_2, \dots, V_m$  Teilvektorräume, sodass jedes  $v \in V$  genau eine Darstellung besitzt  $v = v_1 + v_2 + \dots + v_m$  mit  $v_j \in V_j$ . Man sagt dann, dass die  $V_j$  eine direkte Zerlegung von  $V$  liefern, oder dass  $V$  die direkte Summe der  $V_j$  ist. Man notiert  $V = V_1 \oplus V_2 \oplus \dots \oplus V_m = \bigoplus_1^m V_j$ .

**Sprechweise 2.2.5.** Ein Endomorphismus  $\varphi(\cdot)$  eines Vektorraums heisst eine Projektion, wenn gilt  $\varphi \circ \varphi(\cdot) = \varphi(\cdot)$ . Wenn  $V_0$  der Nullraum und  $V_1$  der Bildraum ist, dann heisst er die Projektion auf  $V_1$  entlang  $V_0$ . Die direkte Zerlegung  $V = V_0 \oplus V_1$  heisst die direkte Zerlegung zur Projektion.

Ist umgekehrt eine direkte Zerlegung  $V = V_0 \oplus V_1$  gegeben, dann heisst der Endomorphismus, welcher  $v = v_0 + v_1$  den Summanden  $v_1$  zuordnet, die Projektion auf  $V_1$  entlang  $V_0$ . Wenn  $\pi$  eine Projektion ist, dann heisst die Projektion  $id - \pi$  die komplementäre Projektion. Zwei Projektionen  $\pi_1$  und  $\pi_2$  heissen disjunkt, wenn  $\pi_2 \circ \pi_1 = \pi_1 \circ \pi_2 = 0$ .

**Satz 2.2.7.**

*Wenn  $\pi_1$  und  $\pi_2$  kommutierende Projektionen sind, dann sind  $\pi_1 \circ \pi_2$  und  $\pi_1 - \pi_1 \circ \pi_2$  disjunkte Projektionen.*

*Wenn  $\pi_1$  und  $\pi_2$  disjunkte Projektionen sind, dann sind  $\pi_1 + \pi_2$  und  $\pi_3 = id - (\pi_1 + \pi_2)$  Projektionen, und  $\pi_1, \pi_2, \pi_3$  sind paarweise disjunkt.*

*Sind  $\pi_1, \pi_2, \dots, \pi_m$  paarweise disjunkte Projektionen mit  $\pi_1 + \pi_2 + \dots + \pi_m = id$ , dann gilt  $V = V_1 \oplus V_2 \oplus \dots \oplus V_m$  mit  $V_j = \pi_j(V)$ .*

*Ist umgekehrt eine direkte Zerlegung  $V = V_1 \oplus V_2 \oplus \dots \oplus V_m$  gegeben, haben wir also für jedes  $v \in V$  eine eindeutige Summendarstellung  $v = v_1 + v_2 + \dots + v_m$  mit  $v_j \in V_j$ , dann liefern die Zuordnungen  $\pi_j : v \mapsto v_j$  ein  $m$ -Tupel paarweise disjunkter Projektionen, dessen Summe die Identitätsabbildung ist.*

Wir befassen uns jetzt mit speziellen Projektionen eines Hilbertraums:

**Definition 2.9.** Eine orthogonale Projektion eines Hilbertraums ist eine Projektion  $\pi$  auf einen abgeschlossenen Teilraum  $W$  entlang des orthogonalen Komplements  $W^\perp$ .

Wir haben gesehen, dass es zu jedem abgeschlossenen linearen Teilraum  $W$  eine orthogonale Projektion  $\pi$  auf  $W$  gibt. Wenn  $\pi$  diese orthogonale Projektion ist, dann ist  $id - \pi$  ist die orthogonale Projektion auf  $W^\perp$ . Wenn zwei orthogonale Projektionen zueinander disjunkt sind, dann sagen wir auch, sie seien zueinander orthogonal.

**Satz 2.2.8.** *Wenn zwei orthogonale Projektionen  $\pi_1$  und  $\pi_2$  kommutieren, dann liefert  $\pi_1 \circ \pi_2$  die orthogonale Projektion auf den Durchschnitt der Bildräume. In diesem Fall sind  $\pi_1 \circ \pi_2, \pi_1 - \pi_1 \circ \pi_2, \pi_2 - \pi_1 \circ \pi_2$  paarweise disjunkte orthogonale Projektionen; und  $\pi_3 = \pi_1 - \pi_1 \circ \pi_2 + \pi_2$  ist die orthogonale Projektion auf die Summe der Bildräume.*

**Beweis 2.2.3.**

Die Endomorphismen  $\varphi_1 = \pi_1 \circ \pi_2$ ,  $\varphi_2 = \pi_1 - \pi_1 \circ \pi_2$ ,  $\varphi_3 = \pi_2 - \pi_1 \circ \pi_2$ ,  $\varphi_4 = id - \pi_3$  sind jedenfalls paarweise disjunkte Projektionen, die sich zur Identität aufsummieren. Sei  $V = V_1 \oplus V_2 \oplus V_3 \oplus V_4$  die entsprechende direkte Zerlegung. Der Bildraum  $V_1$  von  $\pi_1 \circ \pi_2 = \pi_2 \circ \pi_1$  ist in den Bildräumen  $W_1$  und  $W_2$  enthalten, und jedes Element des Durchschnitts wird von  $\pi_1 \circ \pi_2$  auf sich abgebildet. Somit gilt  $V_1 = W_1 \cap W_2$ .

Bis hierher war nicht von Orthogonalität die Rede. In derselben Weise folgern wir nun aber auch  $V_2 = W_1 \cap W_2^\perp$  aus  $\pi_1 \circ (id - \pi_2) = (id - \pi_2) \circ \pi_1$ . Ebenso ergibt sich  $V_3 = W_1^\perp \cap W_2$  und  $V_4 = W_1^\perp \cap W_2^\perp$ . Die Vektorräume stehen paarweise auf einander senkrecht, und ihre Summe ist der gesamte Raum.

**Notation.** Der Deutlichkeit halber schreiben wir manchmal auch  $V = W \overset{\perp}{\oplus} W^\perp$ ; und für feinere orthogonale Zerlegungen  $V = V_1 \overset{\perp}{\oplus} V_2 \overset{\perp}{\oplus} \dots$

Im Satz haben wir zu einem Paar kommutierender orthogonaler Projektionen eine orthogonale Zerlegung in vier Teilräume gewonnen

$$V = V_1 \overset{\perp}{\oplus} V_2 \overset{\perp}{\oplus} V_3 \overset{\perp}{\oplus} V_4. \quad \text{Wir notieren auch} \quad id = \varphi_1 \overset{\perp}{\oplus} \varphi_2 \overset{\perp}{\oplus} \varphi_3 \overset{\perp}{\oplus} \varphi_4.$$

**Satz 2.2.9.** (Unendliche Summen paarweise orthogonaler Summanden)

Es seien  $V_1, V_2, \dots$  paarweise orthogonale Teilvektorräume des Hilbertraums  $V$ .

Zu jedem  $\tilde{v} \in \tilde{V} = \overline{\text{span}}(V_1 \oplus V_2 \oplus \dots)$  existiert dann eine eindeutig bestimmte Folge  $v_1, v_2, \dots$  mit  $v_j \in V_j$ , sodass  $\|(\tilde{v} - (v_1 + v_2 + \dots + v_n))\|^2 = \sum_{n+1}^\infty \|v_j\|^2$ .

Andererseits existiert für jede Folge  $(v_j)_j$  mit  $\sum_{n+1}^\infty \|v_j\|^2 < \infty$  ein  $\tilde{v} \in \tilde{V}$ , sodass

$$\|(\tilde{v} - (v_1 + v_2 + \dots + v_n))\|^2 = \sum_{n+1}^\infty \|v_j\|^2.$$

Der Beweis gehört in die Analysis. Das Kernproblem ist die Vollständigkeit des Raums der quadratsummablen Folgen. Das entscheidende Hilfsmittel ist das sog. Lemma von Fatou (oder der Satz von der dominierten Konvergenz).

**Notation 2.2.** Für ein  $v \in V$  sei  $v_0$  der Fusspunkt des Lots auf  $V_0 = \tilde{V}^\perp$ . Es gilt

$$\|v - (v_0 \oplus v_1 \oplus v_2 \oplus \dots \oplus v_n)\|^2 \rightarrow 0. \quad \text{Man notiert} \quad v = v_0 \overset{\perp}{\oplus} v_1 \overset{\perp}{\oplus} v_2 \overset{\perp}{\oplus} \dots = \overset{\perp}{\oplus}_j v_j.$$

Man notiert auch  $V = V_0 \overset{\perp}{\oplus} V_1 \overset{\perp}{\oplus} V_2 \overset{\perp}{\oplus} \dots = \overset{\perp}{\oplus}_j V_j$ .

Man sagt auch: Die paarweise orthogonalen(!) Projektionen  $\pi_j$  summieren sich zur Identität, und man notiert  $id = \pi_0 + \pi_1 + \pi_2 + \dots$ , wobei aber zu beachten ist, dass die Folge der orthogonalen Projektionen  $\rho_{n+1} = id - (\pi_0 + \pi_1 + \pi_2 + \dots + \pi_n)$  nicht in der Operatornorm nach 0 strebt. (Eine Projektion, die nicht identisch = 0 ist, hat immer die Operatornorm = 1.)

*Beispiel.* (Eindimensionale direkte Summanden)

Besonders übersichtlich ist der Fall, wo die paarweise orthogonalen Teilvektorräume eindimensional sind. Sei  $V_j$  aufgespannt vom Einheitsvektor  $\tilde{w}_j$ ; und  $V = \overset{\perp}{\oplus}(\text{span } \tilde{w}_j)$ . Die

orthogonalen Projektionen sind  $\pi_j(\cdot) = \langle \tilde{w}_j | \cdot \rangle$ . Zu jedem  $v$  gehört eine Folge von Koeffizienten  $c_j = \langle \tilde{w}_j | v \rangle$  mit  $\sum \|c_j\|^2 = \|v\|^2$ ; und es gilt  $v = \sum_j c_j \cdot \tilde{w}_j$ . Umgekehrt liefert jede quadratsummierbare Folge  $(c_j)$  einen Vektor  $v$ .

Dieser Fall liegt konkret vor bei dem Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ , den wir durch Vervollständigung des Prä-Hilbertraums der trigonometrischen Polynome erhalten. Die Einheitvektoren  $\tilde{w}_n = e^{int}$  (für  $n = \dots, -1, 0, 1, 2, \dots$ ) sind paarweise orthogonal. In unserer formalen Schreibweise erscheinen die Vektoren  $v \in V$  als unendliche Summen  $v = \sum_{-\infty}^{\infty} c_n \cdot e^{int}$  mit  $\sum \|c_j\|^2 < \infty$ . Und die Koeffizienten ergeben sich als die Skalarprodukte  $c_n = \langle \tilde{w}_n | v \rangle$ . Damit ist es gelungen, die Elemente der Vervollständigung  $V$  mit den quadratsummablen Folgen zu identifizieren.

**Hinweis auf die Theorie der Fourier-Reihen:** In der Analysis bringt man die Elemente der Vervollständigung in Zusammenhang mit den  $2\pi$ -periodischen Funktionen. Man geht davon aus, dass man ein trigonometrisches Polynom  $v$  nicht nur als eine finite Folge verstehen kann sondern auch als eine  $2\pi$ -periodische Funktion.

$$f(t) = \sum c_n \cdot e^{int} \quad \longleftrightarrow \quad v \quad \longleftrightarrow \quad \mathbf{c} = (\dots, c_1, c_0, c_1, c_2, \dots),$$

Die Koeffizienten ergeben sich durch eine Integration:  $c_n = \langle \tilde{w}_n | v \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-int} \cdot f(t) dt$ . Auch das Skalarprodukt im Prä-Hilbertraum kann man durch eine Integration gewinnen. Für  $w = g(t) = \sum d_n \cdot e^{int}$  gilt  $\langle w | v \rangle = \sum \bar{d}_n \cdot c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{g(t)} \cdot f(t) dt$ . Es stellen sich

### Fragen

1. Kann man auch die Elemente der Vervollständigung (bzw. die stetigen Linearformen) in Verbindung bringen mit geeigneten Typen  $2\pi$ -periodischer Funktionen? Wenn ja, welche 'Funktionen' sind das?
2. Kann man die Gleichung  $h(t) = \sum_{-\infty}^{\infty} a_n \cdot e^{int}$  für quadratsummable Koeffizientenfolgen erklären mit den Summationstechniken für Funktionenfolgen? Es funktioniert weder mit dem Begriff der gleichmäßigen Konvergenz, noch mit dem Begriff der punktweisen Konvergenz.
3. Angenommen, wir hätten 'Funktionen'  $f(t)$  und  $g(t)$ , die wir mit den unendlichen Summen  $v = \sum c_n \cdot e^{int}$  bzw.  $w = \sum d_n \cdot e^{int}$  in Verbindung gebracht haben. Kann man dann das Skalarprodukt durch eine Integration gewinnen?

$$\langle \tilde{w} | v \rangle = \sum \bar{d}_n \cdot c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{g(t)} \cdot f(t) dt.$$

Offenbar braucht man für die Beantwortung dieser Fragen eine leistungsfähige Theorie der Summation von Funktionenreihen und eine leistungsfähige Theorie der Integration.

Im 19. Jahrhundert haben Fourier, Dirichlet, Riemann, Fejér und andere Mathematiker tiefgründige Untersuchungen dazu angestellt. Die definitiven Antworten sind im Jahre 1906 von E. Fischer und F. Riesz auf der Grundlage der Lebesgue'schen Integrationstheorie (von 1901) gegeben worden.

### Orthonormalsysteme

Sei  $(V, \|\cdot\|)$  ein separabler Hilbertraum. Ist  $(v_m)_m$  eine überall dichte Folge, dann ist  $V$  auch die abgeschlossene Hülle der Vereinigung der Teilvektorräume  $\text{span}(\{v_1, v_2, \dots, v_m\})$ . Umgekehrt gilt aber auch: Wenn es eine aufsteigende Folge endlichdimensionaler Teilräume gibt, die in  $V$  dicht liegt, dann ist  $V$  separabel.

#### Sprechweise 2.2.6. (ONS und ONB)

Man nennt eine Folge  $(w_n)_n$  total bezüglich des abgeschlossenen linearen Teilraums  $W$ , wenn gilt  $\overline{\text{span}}(\{w_1, w_2, \dots\}) = W$ .

Eine Menge von Einheitsvektoren, die paarweise zueinander orthogonal sind, heisst ein Orthogonalsystem (ONS). Ein ONS, welches total ist bzgl.  $W$ , heisst eine Orthonormalbasis von  $W$ .

Eine bzgl.  $W$  totale Folge gibt es offenbar für jeden abgeschlossenen linearen Teilraum  $W$  eines separablen Hilbertraums; und man kann die Folge linear unabhängig wählen, so also, dass  $\text{span}(\{w_1, w_2, \dots, w_n\})$  die Dimension  $n$  hat für  $n = 1, 2, \dots$ . Der nächste Satz garantiert die Existenz von Orthonormalbasen.

#### Satz 2.2.10. ('Orthonormalisierungsverfahren nach Gram-Schmidt')

Die Folge  $(v'_n)_n$  sei linear unabhängig und total für den separablen Hilbertraum  $V$ . Es existieren paarweise orthogonale Vektoren mit der Norm 1, sodass

$$\text{span}\{v_1, v_2, \dots, v_n\} = \text{span}\{v'_1, v'_2, \dots, v'_n\} \quad \text{für } n = 1, 2, \dots$$

**Beweis 2.2.4.** Wir konstruieren die  $v_n$  induktiv.  $v_1$  ist notwendigerweise ein Vielfaches von  $v'_1$ . Um die Konstruktion eindeutig zu machen, wählen wir  $c > 0$ , sodass  $\|c \cdot v'_1\| = 1$ . Es seien  $v_1, \dots, v_n$  bereits konstruiert.  $v_{n+1}$  ist nicht Linearkombination dieser  $n$  Vektoren. Wir finden Koeffizienten  $c > 0$  und  $a_1, \dots, a_n \in \mathbb{C}$  sodass  $v_{n+1} = c \cdot v'_{n+1} - \sum_1^n a_j \cdot v_j$  ein Einheitsvektor ist, der zu  $V_n = \text{span}\{v_1, v_2, \dots, v_n\} = \text{span}\{v'_1, v'_2, \dots, v'_n\}$  orthogonal ist. Das ist zu erreichen mit  $a_j = c \cdot \langle v_j \mid v'_{n+1} \rangle$ ; denn mit diesen Koeffizienten gilt  $\langle v_k \mid cv'_{n+1} - \sum a_j v_j \rangle = 0$  für  $k = 1, \dots, n$ .

Mit dem Verfahren werden uns unten näher befassen. Wir werden sehen, dass die Matrix des Basiswechsels  $(v_1, v_2, \dots, v_n) = (v'_1, v'_2, \dots, v'_n) \cdot A$  mit den wohlbekanntten Mitteln der linearen Algebra recht einfach zu berechnen ist.

Zunächst einige Konsequenzen für die Struktur der Hilberträume. Wir zeigen, dass es zwischen Hilberträumen derselben Dimension normerhaltende Isomorphismen gibt. Hilberträume derselben Dimension unterscheiden sich also nur durch die Art, wie die Elemente präsentiert sind, nicht in ihrer inneren Struktur.

**Satz 2.2.11.** Es sei  $u_1, u_2, \dots$  eine ONB für den unendlichdimensionale Hilbertraum  $U$  und  $v_1, v_2, \dots$  eine ONB für den Hilbertraum  $V$ . Es existiert dann genau eine lineare Abbildung  $\varphi : U \rightarrow V$  mit  $v_j = \varphi(u_j)$  für alle  $j$ , und diese ist eine Isometrie.

Das entsprechende gilt, wenn  $U$  und  $V$   $n$ -dimensionale Hilberträume sind.

Der Beweis liegt auf der Hand.

**Satz.** *Jede surjektive isometrische Abbildung eines Hilbertraums  $U$  auf einen Hilbertraum  $V$  ist eine lineare Abbildung, also ein isometrischer Isomorphismus.*

**Beweis.** *Die Isometrie garantiert die Injektivität. Wenn  $\|\varphi(u)\| = \|u\|$  für alle  $u$ , dann liefert die Polarisierung  $\langle \varphi(u_1) | \varphi(u_2) \rangle = \langle u_1 | u_2 \rangle$  für alle  $u_1, u_2$ . Die isometrische Bijektion bildet jede ONB von  $U$  auf eine ONB von  $V$  ab. Sei  $(u_j)_j$  eine ONB und  $v_j = \varphi(u_j)$ . Für (endliche) Linearkombinationen  $u = \sum u_j \cdot a_j$  gilt  $\varphi(u) = \sum v_j \cdot a_j$ ; denn für alle  $k$  gilt  $\langle v_k | \varphi(u) \rangle = \langle u_k | \sum u_j a_j \rangle = a_k$ .*

**Satz 2.2.12.** *Es sei  $u_1, u_2, \dots, u_n$  eine ONB für den  $n$ -dimensionalen Hilbertraum  $U$ .*

a) *Das  $n$ -Tupel  $v_1, v_2, \dots, v_n$  ist genau dann eine ONB, wenn die Matrix des Basiswechsels eine unitäre Matrix ist.*

b) *Sei  $\varphi(\cdot)$  der Endomorphismus, welcher durch  $\varphi(u_j) = \sum_l u_l \cdot a_{lj}$  festgelegt ist.  $\varphi(\cdot)$  ist genau dann eine Isometrie, wenn  $A$  eine unitäre Matrix ist.  $A^* \cdot A = E_n$ ;  $A \in SU(n, \mathbb{C})$ .*

**Beweis 2.2.5.** *Wir beweisen nur die Aussage b):  $v_j = \sum_m u_m \cdot a_{mj}$*

$$\langle v_k | v_j \rangle = \langle \sum_l u_l \cdot a_{lk} | \sum_m u_m \cdot a_{mj} \rangle = \sum_l \bar{a}_{lk} \cdot a_{lj} = (A^* \cdot A)_{kj}$$

*Dies liefert genau dann das Kroneckersymbol  $\delta_{kj}$ , wenn  $A$  eine unitäre Matrix ist.*

**Fazit:** Die für die Geometrie des  $n$ -dimensionalen Hilbertraums charakteristischen ‘Figuren’ sind die Orthonormalsysteme. Die ‘Bewegungen’ in der Hilbertraumgeometrie sind durch die unitären Matrizen zu beschreiben.

## 2.3 Orthogonale Projektionen im Matrizenkalkül. Kleinste Quadrate.

Wir wollen jetzt den Matrix-Kalkül nützen, um gewisse orthogonale Projektionen konkret zu berechnen.

Es wird sich als günstig für die Notation erweisen, dass wir sorgfältig umgehen mit der Idee der dualen Hilberträume. Wir haben bereits gesehen, dass der Dualraum eines Hilbertraums (mit der dualen Norm ausgestattet) ein Hilbertraum ist; wir können daher vom inneren Produkt zweier Linearformen sprechen, insbesondere von der Orthogonalität von Linearformen.

Wenn  $(u_1, u_2, \dots, u_k)$  ein ONS im Hilbertraum  $V$  ist, dann ist das Tupel der bra-Vektoren  $(\langle u_1 |, \langle u_2 |, \dots, \langle u_k |)$  ein ONS im Dualraum  $V^*$ . Ein Vektor  $w \in V$  hat dieselbe Norm wie der bra-Vektor  $\langle w |$  in  $V^*$ ; denn

$$\| \langle w | \| = \sup \left\{ \frac{|\langle w | v \rangle|}{\|v\|} \right\} = \|w\|.$$

Bei der Konstruktion des Skalarprodukts  $s(\ell^1, \ell^2)$  empfiehlt es sich (im Interesse einer übersichtlichen Notation), darauf zu achten, dass das Skalarprodukt in  $V^*$  im ersten Argument linear und im zweiten antilinear wird. Die bra-Vektoren  $\langle w_1 |$  und  $\langle w_2 |$  haben dann in  $V^*$  dasselbe innere Produkt wie die Vektoren  $w_1$  und  $w_2$  in  $V$ .

**Definition 2.10** (Gram-Matrix).

Sind  $v'_1, v'_2, \dots, v'_n$  Vektoren in einem Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ , so nennt man die positiv semidefinite Matrix  $G$  mit den Einträgen  $g_{jk} = \langle v'_j | v'_k \rangle$  die Gram-Matrix des  $n$ -Tupels. Wenn  $v'_1, v'_2, \dots, v'_n$  eine Basis von  $V$  ist, dann nennt man  $G$  auch die Gram-Matrix des Skalarprodukts bzgl. dieser Basis.

**Satz 2.3.1** (Transformation der Gram-Matrizen). *Es sei  $G_v$  die Gram-Matrix des  $n$ -Tupels  $(v'_1, v'_2, \dots, v'_n)$ . Die Vektoren  $w'_1, w'_2, \dots, w'_m$  seien Linearkombinationen der  $v'$*

$$(v'_1, v'_2, \dots, v'_n) \cdot A = (w'_1, w'_2, \dots, w'_m), \quad \sum_i v'_i \cdot a^i_j = w'_j.$$

Für die Gram-Matrix  $G_w$  der  $w'$  gilt dann  $G_w = A^* \cdot G_v \cdot A$ .

**Beweis.**  $\langle w'_i | w'_j \rangle = \langle \sum v'_k a_{ki} | \sum v'_l a_{lj} \rangle = \sum_{kl} \overline{a_{ki}} g_{kl} a_{lj}$ .

Wir werden nun eine Anknüpfung an die Gauss'sche Eliminationsmethode herleiten.

**Satz 2.3.2** (Orthogonalisierung durch LU-Faktorisierung).

Seien  $v'_1, v'_2, \dots, v'_n$  linear unabhängig mit der Gram-Matrix  $G$ . Die Matrix des Orthonormierungsverfahrens ist dann charakterisiert durch die Eigenschaften:

$A$  ist eine obere Dreiecksmatrix mit positiven Diagonaleinträgen und  $A^* \cdot G \cdot A = E_n$ .

Hinweis: Wir schreiben die kritische Gleichung  $A^* \cdot G \cdot A = E_n$  in der Form  $A^* \cdot G = A^{-1}$  und gewinnen daraus einen Hinweis, wie man die untere Dreiecksmatrix  $A^*$  und die obere Dreiecksmatrix  $A^{-1}$  durch das sog. Gauss'sche Eliminationsverfahren gewinnen kann. Mit  $A$  gewinnen wir leicht die inverse Gram-Matrix  $G^{-1} = A \cdot A^*$ .— Die Behauptungen werden sich aus dem Beweis des folgenden, etwas technischer formulierten Satzes ergeben.

**Zur Erinnerung:** Bekanntlich bemüht man sich in der elementaren linearen Algebra, durch sog. elementare Zeilenoperationen eine gegebene Matrix  $M$  in eine obere Dreiecksmatrix umzuwandeln. Wenn die Zeilen und Spalten von  $M$  passend angeordnet sind, dann gelingt das auch. Zu fordern ist, dass jede linke obere Teilmatrix  $M_{k \times k}$  nichtsingulär ist, für  $k \leq \text{Rang}$ . Diese Bedingung ist bei den Gram-Matrizen  $G$  erfüllt, weil alle  $G_{k \times k}$  positiv definit sind.

Das Hintereinanderschalten der elementaren Zeilentransformation bewirkt bekanntlich die Multiplikation (von links!) mit einer unteren Dreiecksmatrix  $L$ , die normiert ist in dem Sinn, dass die Diagonalelemente allesamt = 1 sind. Es ist eine Konsequenz der positiven Definitheit, dass die entstehende obere Dreiecksmatrix  $U = L \cdot G$  in der Diagonalen positive Einträge  $d_1^2, d_2^2, \dots$  hat; das Produkt der ersten  $k$  Diagonalelemente ist die Determinante der Teilmatrix  $G_{k \times k}$ . Ist nun  $D$  die Diagonalmatrix mit den positiven Diagonaleinträgen  $d_1, d_2, \dots$ , dann haben wir  $L \cdot G = D^2 \cdot U$  mit einer normierten oberen Dreiecksmatrix. Wir formulieren das Resultat als

**Satz 2.3.3.** *Es sei  $G$  eine positiv definite  $n \times n$ -Matrix. Für die linken oberen  $k \times k$ -Matrizen gelte  $\det G_{k \times k} = d_1^2 \cdot d_2^2 \cdot \dots \cdot d_k^2$ . Es sei  $D$  die Diagonalmatrix mit den positiven Diagonaleinträgen  $d_1, d_2, \dots, d_n$ . Es gibt genau eine normierte untere Dreiecksmatrix  $L$ , sodass  $L \cdot G$  eine obere Dreiecksmatrix  $D^2 \cdot U$  ist. Für die Matrix  $A$  des Diagonalisierungsverfahrens ( $A^*GA = E$ ) gilt dann*

$$A^* = D^{-1} \cdot L; \quad A^{-1} = D \cdot U \quad \text{mit} \quad L^* = U^{-1}.$$

**Beweis 2.3.1.** *Die Existenz der normierten Dreiecksmatrizen haben wir eben gezeigt. Wir beweisen die Eindeutigkeit:*

$$\begin{aligned} L \cdot G = D^2 \cdot U, \quad \tilde{L} \cdot G = \tilde{D}^2 \cdot \tilde{U} &\implies L^{-1}D^2U = G = \tilde{L}^{-1}\tilde{D}^2\tilde{U} \\ \implies \tilde{L}L^{-1}D^2U = \tilde{D}^2\tilde{U} &\implies \tilde{L}L^{-1} = \tilde{D}^2\tilde{U}U^{-1}D^{-2}. \end{aligned}$$

Wenn eine untere Dreiecksmatrix gleich einer oberen Dreiecksmatrix ist, dann ist sie eine Diagonalmatrix. Daraus folgt  $\tilde{L} = L$ ,  $\tilde{U} = U$ .

Für die obere Dreiecksmatrix  $A$  des Orthonormalisierungsverfahrens leistet die Umrechnung der gegebenen Gram-Matrix in die Einheitsmatrix,  $A^*GA = E$ . Die Dreiecksmatrizen  $A^*$  und  $A$  haben dieselben positiven Diagonalelemente. Es gibt also genau eine Diagonalmatrix  $D$ , sodass  $L = DA^*$  und  $U = D^{-1}A^{-1}$  normierte Dreiecksmatrizen sind.  $LG = D^2U$ ;  $DA^*G = DA^{-1}$  Da die normierten Dreiecksmatrizen durch diese Faktorisierung eindeutig bestimmt sind, haben wir  $A^*$  und  $A^{-1}$  identifiziert.

**Satz 2.3.4.** *Es sei  $(u_1, u_2, \dots, u_n)$  eine Orthonormalbasis des Hilbertraums  $(U, \langle \cdot | \cdot \rangle)$ , und  $(w_1, w_2, \dots, w_p)$  linear unabhängig mit der linearen Hülle  $W$  und der Gram-Matrix  $G = (\langle w_j | w_k \rangle)_{jk}$ .*

$$(u_1, u_2, \dots, u_n) \cdot X = (w_1, w_2, \dots, w_p), \quad \sum_i u_i \cdot x^i_j = w_j.$$

Es gilt dann  $G = X^*EX$ , und die Fußpunkte der Lote der Basisvektoren auf  $W$  sind

$$(\pi(u_1), \pi(u_2), \dots, \pi(u_n)) = (w_1, w_2, \dots, w_p) \cdot F = (u_1, u_2, \dots, u_n) \cdot P.$$

mit den Matrizen  $F = G^{-1}X^*$ ,  $P = XG^{-1}X^* = X \cdot F$ .

**Beweis 2.3.2.** *Es genügt, die Matrix  $P$  zu berechnen;  $F$  ist durch  $X \cdot F$  eindeutig bestimmt, weil die Spalten von  $P$  Linearkombinationen der linear unabhängigen Spalten von  $X$  sind, und deshalb die Koeffizienten eindeutig bestimmt sind. Die Matrix  $P = XG^{-1}X^*$  ist eine hermitesche Projektionsmatrix:  $P \cdot P = P = P^*$ . Aus  $PX = X$  ergibt sich, dass  $P$  die Spalten von  $X$  fest lässt. Wir studieren den durch  $P$  definierten Endomorphismus*

$$(\varphi(u_1), \varphi(u_2), \dots, \varphi(u_n)) = (u_1, u_2, \dots, u_n) \cdot P, \quad \varphi(u_j) = \sum_i u_i \cdot p^i_j$$

Es gilt  $\varphi \circ \varphi = \varphi$  und  $\varphi$  ist selbstadjungiert wegen

$$\langle \varphi(u_j) | u_k \rangle = \langle \sum_i u_i \cdot p^i_j | u_k \rangle = \bar{p}^k_j \quad \langle u_j | \varphi(u_k) \rangle = \langle u_j | \sum_i u_i \cdot p^i_k \rangle = p^j_k.$$

Die Abbildung  $\varphi$  ist idempotent und selbstadjungiert mit dem Bild  $W$ ; also gilt  $\varphi = \pi$ .

Die Projektionsmatrix hat im vorliegenden Fall deswegen eine so einfache Gestalt,  $P = XG^{-1}X^*$  (wo  $G = X^*X$ ), weil die Vektoren, die den Bildraum  $W$  der gesuchten orthogonalen Projektion aufspannen, als Linearkombinationen in einer ONB gegeben waren. Wir wollen die orthogonale Projektion auf  $W$  auch in Fällen studieren, wo  $W$  in anderer Weise als ein span oder als das Nullstellengebilde eines homogenen linearen Gleichungssystems vorgegeben ist. Dazu wollen zunächst einmal die hier gestellte Aufgabe und ihre Lösung in die Welt des Standard-Hilbertraums  $\mathbb{C}_{St}^n$  (mit dem üblichen Skalarprodukt) übersetzen. Die Sache wird übersichtlich, wenn wir auch noch die Faktorisierung der Gram-Matrix ins Feld führen.  $E_{p \times p} = A^*GA$ ;  $G^{-1} = A \cdot A^*$ .

Ist  $\mathbf{c}$  die Spalte zum Vektor  $\sum_i u_i \cdot b^i$  und  $\mathbf{d}$  die Spalte zum Fußpunkt  $\pi(u) = \sum_i \pi(u_i) \cdot d^i = \sum_j u_j \cdot \sum_i p^j_i d^i$ , so sagt uns der Satz

$$\mathbf{d} = P \cdot \mathbf{c} \quad (\text{mit}) \quad P = XG^{-1}X^* = (XA) \cdot (XA)^*.$$

Die Matrix  $B = XA$  (vom Format  $n \times p$ ) hat orthonormierte Spalten,  $B^*B = E_{p \times p}$ ; denn  $B^*B = A^*X^*XA = E_{p \times p}$ . (Diese Rechnung hätten wir nicht gebraucht, wenn wir uns

daran erinnert hätten, dass  $A$  zur Orthogonalisierung des Spaltenraums von  $X$  konstruiert wurde.) Die Matrix  $BB^* = P$  ist die gesuchte hermitesche Projektionsmatrix auf den Spaltenraum der Matrix  $X$ ; denn es gilt  $P \cdot P = P = P^*$ ,  $P \cdot X = X$ .

Wir können das Resultat nun auch mit der Notation von Dirac formulieren. Die orthogonale Projektion auf den eindimensionalen, vom Einheitsvektor  $v$  aufgespannten Teilraum hat offenbar die Form

$$\pi(\cdot) = v \cdot \langle v | \cdot \rangle, \quad \text{oder auch} \quad \pi = |v\rangle \cdot \langle v| \quad .$$

Wenn  $v_1, \dots, v_d$  eine ONBasis von  $W$  ist, dann ist  $\pi_W = \sum |v_j\rangle \cdot \langle v_j|$  die orthogonale Projektion auf  $W$ .

### Kleinste Quadrate:

Bekanntlich ist ein inhomogenes lineares Gleichungssystem  $B \cdot \mathbf{x} = \mathbf{y}$  nur dann lösbar, wenn  $\mathbf{y}$  im Spaltenraum der Matrix  $B$  (vom Format  $n \times d$ ) liegt. Wenn das nicht der Fall ist, spricht man von einem überbestimmten Gleichungssystem. Überbestimmte Gleichungssysteme treten z. B. dort auf, wo Daten  $\mathbf{y}$  gemessen werden, die nicht mit dem vorgegebenen linearen Modell vereinbar sind. Man kann manchmal davon ausgehen, dass die Nichtübereinstimmung durch Messfehler verursacht sind, (die die Messungen durchaus nicht uninformativ machen). Man will dann doch davon ausgehen, dass eine  $d$ -Spalte  $\mathbf{s}$  existiert, sodass gilt  $\mathbf{y} = B \cdot \mathbf{s} + \varepsilon$  mit kleinen Fehlern  $\varepsilon_i$ . In der ‘Ausgleichsrechnung’ bestimmt man  $\hat{\mathbf{s}}$  so, dass der ‘Fehlervektor’ minimal ist,— wobei man sich natürlich noch verständigen muss, in welcher Norm man den Fehler messen will:  $\|\varepsilon\| = \min$ .

Die Suche nach dem besten  $d$ -Tupel  $\hat{\mathbf{s}}$  ist besonders einfach, wenn  $\|\cdot\|$  eine Hilbertraum (im Raum der  $n$ -Spalten) ist, wenn also eine positivdefinite  $n \times n$ -Matrix  $Q$  gegeben ist, sodass  $\|\varepsilon\|^2 = \varepsilon^* Q \varepsilon$ . (Wenn alle Größen reell sind, könnte man natürlich statt  $\varepsilon^*$  auch  $\varepsilon^T$  schreiben.) Die Aufgabe besteht dann offenbar darin, die  $n$ -Spalte  $\mathbf{y}$  orthogonal auf den Spaltenraum der Matrix  $B$  zu projizieren.

**Satz 2.3.5** (Kleinste Quadrate). *Es sei  $B$  eine  $n \times d$ -Matrix mit dem  $d$ -dimensionalen Spaltenraum  $W$ ; und  $Q$  eine positivdefinite  $n \times n$ -Matrix. Zu jeder  $n$ -Spalte  $\mathbf{y}$  existiert genau ein  $\hat{\mathbf{y}} = B \cdot \hat{\mathbf{s}} \in W$ , sodass*

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = (\mathbf{y} - \hat{\mathbf{y}})^* Q (\mathbf{y} - \hat{\mathbf{y}}) = \min .$$

Es gilt  $\hat{\mathbf{y}} = P \mathbf{y}$  mit  $P = B(B^*QB)^{-1}B^*Q$ .

**Beweis.** Der Vektor  $\mathbf{y} - \hat{\mathbf{y}}$  ist das Lot von  $\mathbf{y}$  auf  $W$ . Die Bedingung  $\mathbf{y} - \hat{\mathbf{y}} \perp W$  bedeutet für  $\hat{\mathbf{y}} = B \cdot \hat{\mathbf{s}}$ :  $\forall \mathbf{s} : (\mathbf{y} - \hat{\mathbf{y}})^* Q B \mathbf{s} = 0 \iff B^* Q (\mathbf{y} - B \hat{\mathbf{s}}) = 0 \iff B^* Q B \hat{\mathbf{s}} = B^* Q \mathbf{y}$ .

Man bemerke: Wenn in unserem Hilbertraum  $(C_{S_p}^n, Q)$  ein Endomorphismus  $\varphi$  durch die Matrix  $S$  beschrieben wird, dann wird der adjungierte Endomorphismus  $\varphi^*$  durch die Matrix  $T$  beschrieben mit  $T^*Q = QS$ , also  $T = Q^{-1}S^*Q$ . In der Tat haben wir

$$\langle z | \varphi y \rangle = \langle \varphi^* z | y \rangle \iff z^* Q S y = (T z)^* Q y = z^* T^* Q y.$$

Die oben berechnete Matrix  $P$  liefert also in der Tat einen selbstadjungierten Endomorphismus unseres Hilbertraums  $(\mathbb{C}_{Sp}^n, Q)$ .

Bemerke: Wenn es um die Berechnung der orthogonalen Projektion auf  $W$  geht, dann hat die Matrix  $B$  nur eine Hilfsrolle; jede  $n \times p$ -Matrix mit dem Spaltenraum  $W$  liefert dieselbe Projektion. In der Tat liefert  $A = BS$  (mit einer invertierbaren  $p \times p$ -Matrix  $S$ ) dieselbe Matrix  $P = BS(S^*B^*QBS)^{-1}S^*B^*Q = B(B^*QB)^{-1}B^*Q$ . Besonders übersichtlich wird die Formel, wenn wir  $S$  mit Hilfe des Orthogonalisierungsverfahrens so wählen, dass  $S^*(B^*QB)S = E_{p \times p}$ . Dann haben wir nämlich  $P = AA^*Q$  mit  $A = BS$ .

### Erwartungstreue Schätzung mit minimaler Varianz:

Es seien  $Y_1, \dots, Y_n$  komplexwertige Zufallsgrößen mit der bekannten (positivdefiniten!) Kovarianzmatrix  $C$ . Vom Erwartungswert  $\mathfrak{E}Y$  der zufälligen  $n$ -Spalte  $Y$  sei nur bekannt, dass er in einem gewissen linearen Teilraum  $W$  liegt. Der Erwartungswert einer Linearkombination der Komponenten,  $\xi \cdot Y$  soll durch eine Linearkombination  $\hat{\xi} \cdot Y$  geschätzt werden, sodass gilt

$$(\xi - \hat{\xi}) \cdot w = 0 \quad \text{für alle } w \in W, \quad \text{var}(\hat{\xi} \cdot Y) = \min.$$

Wir suchen lieber die Zeile  $\hat{\eta} = \xi - \hat{\xi} \in W^\dagger$  mit  $\text{var}((\xi - \hat{\eta}) \cdot Y) = \min$ ; und wir übersetzen die Aufgabe in die Sprache der Hilberträume.  $V^*$  sei der Hilbertraum der komplexen  $n$ -Zeilen  $\xi$  mit der Norm  $\|\xi\| = \sqrt{\xi C \xi^*}$ . Damit assoziieren wir die Sesquilinearform  $s(\xi, \eta) = \text{cov}(\xi Y, \eta Y) = \xi C \eta^*$ . (Es entspricht der Tradition der linearen Statistik, dass  $s(\cdot, \cdot)$  im ersten Argument linear und im zweiten Argument antilinear ist.) Gesucht ist der Punkt  $\hat{\eta}$  im Zeilenraum  $W^\dagger$ , in welchem die quadratische Funktion  $W^\dagger \ni \eta \mapsto \|\xi - \eta\|^2$  ihr Minimum annimmt. Gesucht ist also der Fußpunkt  $\hat{\eta}$  des Lots von  $\xi$  auf den Raum  $W^\dagger$ .  $(\xi - \hat{\eta}) \cdot Y$  ist dann der beste lineare Schätzer für  $\xi \cdot \mathfrak{E}Y$ .

Wenn  $W$  die Dimension  $d$  hat, dann hat  $W^\dagger$  die Dimension  $n - d$ ; es existiert eine  $(n - d) \times n$ -Matrix  $M^*$ , deren Zeilenraum  $W^\dagger$  ist.  $W^\dagger = \{\eta : \eta = t \cdot M^*\}$ . Wenn  $W$  der Spaltenraum der  $d \times n$ -Matrix  $B$  ist, dann ist  $M^*$  gekennzeichnet durch  $M^*B = 0$ .

$M^*$  und  $B$  stehen in der Beziehung, die von den homogenen linearen Gleichungssystemen wohl vertraut ist; man könnte  $M^*$  einen Annihilator für  $B$  nennen, und  $B^*$  einen Annihilator von  $M$ . Wir suchen hier eine Projektion auf den Zeilenraum von  $M^*$ , so wie wir oben eine Projektion auf den Spaltenraum von  $B$  gesucht haben.

Nun die konkrete Berechnung: Die Bedingung  $(\xi - \hat{\eta}) \perp W^\dagger$  bedeutet für  $\hat{\eta} = \hat{t}M^*$ :  $\forall t : (\xi - \hat{\eta})CMt^* = 0$ . Wegen  $(\xi - \hat{t}M^*)CA = 0$  oder  $\hat{t}(M^*CM) = \xi CM$  liefert das

$$\hat{\eta} = \xi R \quad \text{mit} \quad R = CM(M^*CM)^{-1}M^*.$$

Man bemerke. Die Abbildung  $\xi \mapsto \xi R$  mit  $R = CM(M^*CM)^{-1}M^*$  ist in der Tat eine selbstadjungierte Projektion unseres Hilbertraums  $(V^*, C) = (\mathbb{C}_{\mathbb{Z}}^n, C)$ ; denn  $CR^* = RC$ .

Es kommt nur auf den Spaltenraum von  $M$  an.  $MT$  mit einem invertierbaren  $T$  hat denselben Spaltenraum.  $W^\dagger = \{\eta : \eta = s \cdot N^* = s \cdot T^*M^*\}$ .

Die Rechnungen laufen hier sehr ähnlich wie oben bei der Methode der kleinsten Quadrate. Man kann die Aufgabe aber auch noch etwas anders anpacken, indem man nicht von  $M^*$  (oder  $N^*$ ) ausgeht, sondern von einem Annihilator  $B$ .

**Satz 2.3.6.** *Es sei  $C$  eine positivdefinite  $n \times n$ -Matrix;  $C = Q^{-1}$ .  $W$  sei der Spaltenraum einer  $n \times p$ -Matrix  $B$  mit linear unabhängigen Spalten.  $W^\dagger$  bezeichnet seinen Annihilator;  $W^\dagger = \{\eta : \eta B = 0\}$ .  $(\mathbb{C}_z^n, C)$  sei der Hilbertraum mit dem inneren Produkt  $\langle \xi | \eta \rangle = \xi C \eta^*$ . Das Lot von  $\xi$  auf den Raum  $W^\dagger$  ist  $\hat{\xi} = \xi P$  mit  $P = B(B^*QB)^{-1}B^*Q$ .*

**Beweis 2.3.3.** *Für ein vorgegebenes  $\xi$  sei  $\xi = \hat{\xi} + \hat{\eta}$  die orthogonale Zerlegung mit  $\hat{\eta} \in W^\dagger$  und  $\hat{\xi} \in (W^\dagger)^\perp$ . Zunächst stellen wir fest:  $\hat{\eta} \in W^\dagger \implies \xi - \hat{\xi} \in W^\dagger \implies (\xi - \hat{\xi})B = 0 \implies \hat{\xi}B = \xi B$ . Andererseits haben wir*

$$\begin{aligned} \hat{\xi} \in (W^\dagger)^\perp &\iff \forall \eta \in W^\dagger : \hat{\xi} C \eta^* = 0 \iff (\eta B = 0 \implies \eta(\hat{\xi} C)^* = 0) \\ &\iff \exists s : (\hat{\xi} C)^* = B \cdot s; \quad \hat{\xi} = s^* B^* Q \implies \exists s : s^*(B^*QB) = \hat{\xi} B = \xi B. \\ &\implies s^* = \xi B (B^*QB)^{-1} \implies \hat{\xi} = s^* B^* Q = \xi B (B^*QB)^{-1} B^* Q. \end{aligned}$$

Der Satz löst das Problem des besten linearen erwartungstreuen Schätzung eines Erwartungswerts. (BLUE ‘Best Linear Unbiased Estimator’). Der Vergleich mit dem Satz 2.3.5 zeigt die Verbindung zur Methode der kleinsten Quadrate. Man entnimmt den Formeln den

**Satz 2.3.7** (Satz von Gauss-Markov). *Wenn man eine Linearkombination der Komponenten eines Zufallsvektors  $Y$  erwartungstreu mit minimaler Varianz schätzen will, dann nehme man die betreffende Linearkombination des ‘ausgeglichenen’ Datenvektors  $\hat{Y}$ . Es gilt nämlich*

$$\hat{\xi} Y = \xi \cdot P Y = \xi \hat{Y}.$$

Wir vergleichen das Resultat mit dem im vorigen Satz:  $\hat{\xi} = \xi - \hat{\eta} = \xi(I - R)$ . Wir haben hier disjunkte Projektionen  $P$  und  $R$ ,  $P$  vom Rang  $p$ ,  $R$  vom Rang  $n - p$ .

$$P = B(B^*QB)^{-1}B^*Q; \quad R = CM(M^*CM)^{-1}M.$$

Wegen  $M^*B = 0, CQ = E$  ergibt sich durch direkte Rechnung  $R \circ P = 0 = P \circ R$ .

### Dualität bei endlichdimensionalen Hilberträumen

Die Formeln sollten an Transparenz gewinnen, wenn wir den Gedanken der Vektorraum-dualität im konkreten Fall ein wenig vertiefen.

**Notation 2.3.** Es seien  $Q$  und  $C$  zueinander inverse positivdefinite  $n \times n$ -Matrizen.

$V_{Sp} = (\mathbb{C}_{Sp}^n, Q)$  ist der Hilbertraum der  $n$ -Spalten  $y$  mit der Norm  $\|y\| = \sqrt{y^*Qy}$ .

$V_Z = (\mathbb{C}_Z^n, C)$  bezeichnet den Hilbertraum der  $n$ -Zeilen  $\eta$  mit der Norm  $\|\eta\| = \sqrt{\eta C \eta^*}$ .

Für  $x, y \in V_{Sp}$  ist  $\langle x | y \rangle_{Sp} = x^* \cdot Q \cdot y$ .

Für  $\xi, \eta \in V_Z$  ist  $\langle \xi | \eta \rangle_Z = \xi \cdot C \cdot \eta^*$ .

Wir stellen einige bekannte (oder leicht zu beweisende) Tatsachen zusammen

**Satz 2.3.8** (Duale Hilberträume). 1. Die Räume  $V_{S_p}$  und  $V_Z$  bilden als komplexe Vektorräume vermöge des Matrizenprodukts (Zeile  $\times$  Spalte = Skalar) ein Dualitätspaar.

2. Die hier definierten Normen sind zueinander duale Normen

$$\begin{aligned} \sup\{ |\xi \cdot x| : x^* Q x \leq 1 \} &= \sqrt{\xi C \xi^*} \quad \text{für } \xi \in V_Z, \\ \sup\{ |\xi \cdot x| : \xi C \xi^* \leq 1 \} &= \sqrt{x^* Q x} \quad \text{für } x \in V_{S_p}. \end{aligned}$$

3. Jeder Linearform entspricht in antilinearer Weise ein Vektor

$$\begin{aligned} \text{Der Spalte } x \text{ entspricht die Zeile} & \quad x^* Q \text{ oder } \langle x | \cdot \rangle_{S_p} \\ \text{Der Zeile } \xi \text{ entspricht die Spalte} & \quad C \xi^* \text{ oder } \langle \cdot | \xi \rangle_Z \end{aligned}$$

4. Die Abbildung ist eine Isometrie.

Wenn nämlich  $\xi = x^* C$ ,  $\eta = y^* C$ , also  $x = Q \xi^*$ ,  $y = Q \eta^*$ , dann gilt

$$\langle x | y \rangle_{S_p} = x^* \cdot Q \cdot y = (x^* Q) \cdot C \cdot (Q y) = \xi C \eta^* = \langle \xi | \eta \rangle_Z.$$

5. Wenn  $A$  eine  $n \times n$ -Matrix ist, dann ergibt das Endomorphismen in den beiden Räumen und dazu adjungierte Endomorphismen

$$\begin{aligned} \langle y | Ax \rangle_{S_p} &= y^* Q A x = (B y)^* Q x = \langle B y | x \rangle_{S_p} \quad \text{mit } B^* Q = Q A, \\ \langle \xi A | \eta \rangle_Z &= \xi A C \eta^* = \xi C (B \eta)^* = \langle \xi | \eta B \rangle_Z \quad \text{mit } A C = C B^*. \end{aligned}$$

6. Die Matrix  $A$  liefert genau dann für den Spaltenraum einen selbstadjungierten Endomorphismus, wenn sie einen selbstadjungierten Endomorphismus des Zeilenraums liefert.

Als wir oben unsere orthogonalen Projektionen konstruiert haben, haben wir einmal im Spaltenraum argumentiert und das andere Mal vom Zeilenraum her. Bei der Methode der kleinsten Quadrate suchten wir die Projektion auf den Spaltenraum der Matrix  $B$ . Bei der Konstruktion des BLUE suchten wir die Projektion auf den Zeilenraum von  $M^*$ . Setzen wir nun  $N^* = B^* Q$  und betrachten wir die Formel für die Projektionsmatrix  $R$ : Die zu invertierende Gram-Matrix wird  $N^* C N = B^* Q C Q B = B^* Q B$ ; und die Projektionsmatrix wird  $C N (B^* Q B)^{-1} N^* = B (B^* Q B)^{-1} B^* Q = P$ . Dies liefert den

**Satz 2.3.9** (Projektionsmatrizen).

Sei  $B$  eine  $n \times p$ -Matrix mit linear unabhängigen Spalten und  $V_1$  ihr Spaltenraum. Die Matrix  $P = B (B^* Q B)^{-1} B^* Q$  ist dann die Matrix der orthogonalen Projektion von  $V_{S_p} = (\mathbb{C}_{S_p}^n, Q)$  auf  $V_1$ .  $P$  ist gleichzeitig die Matrix der orthogonalen Projektion von  $V_Z =$

$(\mathbb{C}_Z^n, C)$  auf den Zeilenraum von  $N^* = B^*Q$ .

Sei  $A$  eine  $d \times n$ -Matrix mit linear unabhängigen Zeilen und  $V_0 = \{x : Ax = 0\}$ . Die orthogonale Projektion von  $V_{Sp} = (\mathbb{C}_{Sp}^n, Q)$  entlang  $V_0$  wird beschrieben durch die Matrix  $S = CA^*(ACA^*)^{-1}A$ .  $S$  ist auch die Matrix der orthogonalen Projektion des Zeilenraums entlang  $\{\xi : \xi CA^* = 0\}$ .

### Schlussbemerkung zur Matrix-Methode

Die Sache mit den dualen Hilberträumen, bestehend aus Spalten bzw. Zeilen wird schlank, wenn  $Q$  und  $C = Q^{-1}$  die Einheitsmatrix sind, wenn sich also alle Matrizendarstellungen auf eine ONB stützen.

Unpassende Basen werden wir in Zukunft nach Möglichkeit vermeiden. In der linearen Statistik hat man leider keine Wahl.

Erfreulicherweise scheint es, dass man sich bei den Anwendungen in der Quantenmechanik, wo man gelegentlich Matrizendarstellungen ins Spiel bringen will, wirklich immer auf Orthonormalbasen beschränken kann. In diesem Fall werden dann die Koordinatenwechsel durch unitäre Matrizen beschrieben; und die hermitesche Konjugation beschreibt wirklich den Übergang zum adjungierten Endomorphismus. Das Orthogonalisierungsverfahren nach Gram-Schmidt spielt dort keine Rolle. Ein großes Thema ist stattdessen die unitäre Diagonalisierung der hermiteschen Matrizen (oder der selbstadjungierten Operatoren). Damit werden wir uns gründlich zu befassen haben.

## 2.4 Exkurs: Quadratische Formen im reellaffinen Raum

Der Umgang mit quadratischen Kurven, Flächen und Funktionen hat lange Traditionen.

Die Griechen zur Zeit von Euklid beschäftigten sich nicht nur mit Geraden, Kreisen, Ebenen und Kugeln; sie befassten sich auch mit den sog. Kegelschnitten, den Ellipsen, Parabeln und Hyperbeln. Diese ‘Kurven zweiter Ordnung’ kann man nicht nur (wie der Name suggerieren könnte) als Schnittkurven eines Kegels mit einer Ebene erhalten; es gibt viele Möglichkeiten der Definition oder auch der Konstruktion mit elementaren Mitteln. Diesen wurde zu manchen Zeiten ein Platz in der Schulmathematik eingeräumt.

Kepler entdeckte beim Studium astronomischer Daten, dass die Planeten sich auf Ellipsenbahnen bewegen, und Newton gab einen mathematischen Beweis auf der Grundlage des Gesetzes, welches die Massenanziehung regelt. Bei den Beweisen, die man heute den Studierenden vorträgt, erfährt man Interessantes, sowohl über die Kegelschnitte als geometrische Objekte als auch über die Beschreibung der Kegelschnitte durch Funktionen.

Quadratische Gleichungen in einer reellen Veränderlichen sind ein aktuelles Thema der Schulmathematik. Manchmal werden in diesem Umfeld auch noch die quadratischen Gleichungen in zwei Variablen behandelt und die Kegelschnitte als ihre Lösungsmengen.

In den heute üblichen Vorlesungen zur Analysis werden neben den linearen auch die quadratischen Näherungen von glatten Funktionen mehrerer Veränderlicher diskutiert

$$\begin{aligned} f(x, y) &= a_0 + a_1x + a_2y + \frac{1}{2}a_{11}x^2 + a_{12}xy + \frac{1}{2}a_{22}y^2 + \text{Rest} \\ &= a_0 + (a_1, a_2) \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x, y) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \text{Rest} \end{aligned}$$

Hier sind  $x$  und  $y$  sog. lokale Koordinaten. Im approximierenden Polynom ist der Anteil vom Grad 1 durch den Gradienten bestimmt, der Anteil vom Grad 2 durch die symmetrische Hesse-Matrix. Wenn der Nullpunkt ein stationärer Punkt ist, wenn also der Gradient im Nullpunkt verschwindet, dann gibt (im nichtausgearteten Fall) die quadratische Form zur Hesse-Matrix Auskunft, ob ein lokales Extremum vorliegt oder ein sog. Sattelpunkt. Fragen dieser Art wurden in der zweiten Hälfte des 20. Jahrhunderts in den mathematischen Anfängervorlesungen mehr oder weniger gründlich behandelt.

Für Mathematiker geschah das üblicherweise in der Veranstaltung, in welcher auch die Matrizenrechnung gelehrt wurde. Das waren Veranstaltungen mit Namen wie ‘Lineare Algebra und Analytische Geometrie’ oder ‘Lineare Algebra und Matrix-Theorie’. Die Gewichte haben sich nach und nach verschoben. Die Veranstaltungen zur Linearen Algebra haben sich mehr und mehr von den klassischen geometrischen Themen abgewandt. (Die Entwicklungen liefen übrigens sehr ähnlich im englischsprachigen und im russischsprachigen Universitätsbetrieb.) In der linearen Algebra neueren Zuschnitts diskutiert man allenfalls am Rande die Kurven und Flächen zweiter Ordnung. Man diskutiert stattdessen die symmetrischen Bilinearformen. Bevor wir uns dahin wenden, werfen wir einen (reichlich ungeometrischen) Blick auf die quadratischen Funktionen auf einem affinen Raum.

**Vorbemerkung: Additive Funktionen und ‘lineare’ Stetigkeit**

Eine reellwertige Funktion  $f(x)$  auf der reellen Achse  $\mathbb{R}$  nennt man eine additive Funktion, wenn gilt  $f(x + y) = f(x) + f(y)$  für alle  $x, y \in \mathbb{R}$ . Für eine solche Funktion gilt offenbar  $f(r) = r \cdot f(1)$  für alle rationalen  $r$ . Die stetigen additiven Funktionen sind die linearen Funktionen. Die Mathematiker haben keine Schwierigkeit zu akzeptieren, dass es auch unstetige additive Funktionen gibt, obwohl man die Möglichkeit ausschliessen kann, dass man eine solche unstetige additive Funktion jemals zu Gesicht bekommt.

**Sprechweise.** Sei  $f(v)$  eine reelle oder komplexe Funktion auf einem reellen oder komplexen Vektorraum  $V$ . Wir sagen,  $f$  sei linear stetig, wenn für jedes Paar von Vektoren  $v, w$  die Funktion  $F(t) = f(v + tw)$  stetig ist auf  $\mathbb{R}$  bzw. auf  $\mathbb{C}$ .

**Was ist eine quadratische Funktion?** Es sei  $L$  ein affiner Raum zum reellen Vektorraum  $V$ . Affine und quadratische Funktionen sind hier reellwertige Funktionen auf  $L$ ; Linearformen und quadratische Formen sind reellwertige Funktionen auf  $V$ . Bilinearformen sind reellwertige Funktionen auf  $V \times V$ . Was affine Funktionen, Linearformen und Bilinearformen sind, sollte bekannt sein. Was quadratische Funktionen und quadratische Formen (hier im reellen Fall!) sind, werden wir gründlich diskutieren.

Eine Funktion  $a(\cdot)$  auf  $L$  heisst bekanntlich eine affine Funktion, wenn eine Linearform  $\ell(\cdot)$  existiert, sodass für alle  $P \in L$  gilt  $a(P + t \cdot v) - a(P) = t \cdot \ell(v)$ .

In Worten: “Für jedes feste  $v \in V$  ist der Zuwachs beim Schritt  $v$  unabhängig vom Ort.” Die Abhängigkeit vom Schritt  $v \in V$  ist notwendigerweise eine Linearform ist; denn

$$a(P + su + tv) - a(P + su) = t \cdot \ell(v); \quad a(P + su) - a(P) = s \cdot \ell(u)$$

$$\ell(su + tv) = a(P + su + tv) - a(P) = t \cdot \ell(v) + a(P + su) - a(P) = s \cdot \ell(u) + t \cdot \ell(v).$$

Eine übliche Definition des Begriffs der quadratischen Funktion auf  $L$  ist die folgende

**Definition 2.11.** Die Funktion  $q(P)$  heisst eine quadratische Funktion, wenn ein Punkt  $P_0$ , eine Linearform  $\ell_0(\cdot)$  und eine symmetrische Bilinearform  $b_0(\cdot, \cdot)$  existieren, sodass

$$(\text{quadr}) \quad q(P) = q(P_0 + v) = q(P_0) + \ell_0(v) + \frac{1}{2}b_0(v, v).$$

Bemerke: Wenn  $P_1 = P_0 + u$  irgendein Punkt ist, dann gilt

$$q(P) = q(P_1 + w) = q(P_0 + u + w) = q(P_0) + \ell_0(u + w) + \frac{1}{2}b_0(u + w, u + w)$$

$$= [q(P_0) + \ell_0(u) + \frac{1}{2}b_0(u, u)] + [\ell_0(w) + b_0(u, w)] + \frac{1}{2}b_0(w, w)$$

$$= q(P_1) + \ell_1(v) + \frac{1}{2}b_0(w, w).$$

Der Punkt  $P_0$  spielt also keine ausgezeichnete Rolle, die Bilinearform ist ortsunabhängig, und die Linearform (der ‘Gradient’), als Funktion mit Werten in  $V^*$ , hängt in affiner Weise vom Ort ab; genauer  $\ell(P_0 + u, \cdot) - \ell(P_0, \cdot) = b(u, \cdot)$  für jeden Punkt  $P_0$ .

Hinweis: Die Formel (quadr) sollte daran erinnern, wie für eine glatte Funktion  $f(\cdot)$  die Taylorapproximation zweiter Ordnung aussieht

$$f(P_0 + t \cdot v) = f(P_0) + t \cdot f'(P_0) \cdot v + \frac{1}{2}t^2 \cdot v^T \cdot f''(P_0) \cdot v + o(t^2) \quad \text{für } t \rightarrow 0.$$

Die Linearform  $f'(P_0)$  (hier als Zeile notiert, während  $v$  als Spalte notiert ist,) heisst der Gradient im Punkt  $P_0$ , die symmetrische Matrix  $f''(P_0)$  heisst die Hesse-Matrix im Punkt  $P_0$ . Wenn  $f(\cdot)$  keine affine Funktion ist, dann ist der Gradient ortsabhängig; wenn  $f(\cdot)$  keine quadratische Funktion ist, dann ist die Hesse-Matrix ortsabhängig.

Definitionen (als Begriffsbestimmungen), welche Existenzbedingungen enthalten, gelten als weniger schön als solche, welche direkt nachrechenbare Forderungen stellen. Insofern könnte man mit der obigen Definition nicht voll zufrieden sein. Der folgende Satz behebt diesen Schönheitsfehler. Die Idee ist aus der Analysis vertraut: die quadratischen Funktionen sind diejenigen, für welche alle Richtungsableitungen affine Funktionen sind.

**Satz 2.4.1.** *Eine linearstetige Funktion  $q(\cdot)$  auf einem reellaffinen Raum ist genau dann eine quadratische Funktion, wenn für jedes  $v$  der Zuwachs  $q(P + v) - q(P)$  eine affine Funktion  $a_v(P)$  ist.*

**Beweis.** *Wir haben bereits gesehen, dass die Eigenschaft (quadr) die hier geforderte Eigenschaft impliziert.*

*Wenn für festes  $v$  der Zuwachs  $a_v(\cdot)$  eine affine Funktion ist, dann existiert eine Linearform  $B(v, \cdot)$ , sodass  $a_v(P + tw) - a_v(P) = t \cdot B(v, w)$ , unabhängig von  $P$ .*

*$B(\cdot, \cdot)$  ist symmetrisch; denn es gilt (für beliebiges  $P$ )*

$$B(v, w) = a_v(P + w) - a_v(P) = q(P + w + v) - q(P + w) - q(P + v) + q(P) = B(w, v).$$

*Es gilt  $B(v + tw, v + tw) = B(v, v) + t^2 B(w, w) + 2t \cdot B(v, w)$ .*

*Die Funktion  $\ell_0(v) = q(P_0 + v) - q(P_0) - \frac{1}{2}B(v, v)$  ist additiv für jedes  $P_0$ ; denn*

$$\ell_0(v + tw) - \ell_0(v) - \ell_0(tw) = q(P_0 + v + tw) - q(P_0 + v) - q(P_0 + tw) + q(P_0) - B(v, tw) = 0.$$

*Durch vollständige Induktion finden wir  $\ell_0(v + nw) = n \cdot \ell_0(w)$  für alle natürlichen Zahlen  $n$ ; dies ergibt auch  $\ell_0(v + rw) = r \cdot \ell_0(w)$  für alle rationalen Zahlen. Die angenommene lineare Stetigkeit von  $q(\cdot)$  impliziert die Linearität von  $\ell_0(\cdot)$ . Wir haben also für jedes  $P_0$  eine Linearform  $\ell_0$ , sodass gilt*

$$q(P_0 + v) = q(P_0) + \ell_0(v) + \frac{1}{2}B(v, v)$$

Im gleiche Sinn wollen wir nun auch diskutieren, was eine quadratische Form auf einem reellen Vektorraum ist.

**Definition 2.12.** Eine Funktion  $\Psi(\cdot)$  auf einem Vektorraum heisst eine quadratische Form, wenn es eine symmetrische Bilinearform  $\Phi(\cdot, \cdot)$  gibt, sodass  $\Phi(v, v) = \Psi(v)$ .

Es zeigt sich, dass die symmetrische Bilinearform durch  $\Psi(\cdot)$  eindeutig bestimmt ist, und wir werden in der sog. Polarisierungskonstruktion eine explizite Formel für  $\Phi(\cdot, \cdot)$  finden. Man könnte somit auch definieren: Die Funktion  $\Psi(\cdot)$  ist eine quadratische Form genau dann, wenn die Polarisierungskonstruktion auf eine symmetrische Bilinearform führt.

**Satz 2.4.2.** Jede quadratische Form ist linearstetig. Eine linearstetige Funktion  $\Psi(\cdot)$  ist genau dann eine quadratische Form, wenn sie die Parallelogrammgleichung erfüllt.

$$\Psi(v+w) + \Psi(v-w) = 2 \cdot \Psi(v) + 2 \cdot \Psi(w) \quad \text{für alle } v, w \in V.$$

**Lemma 2.4.1.** Wenn die Funktion  $\Psi(v)$  die Parallelogrammgleichung erfüllt, dann gilt  $\Psi(r \cdot v) = r^2 \cdot \Psi(v)$  für alle rationalen  $r$ .

**Beweis.** Für  $v = w = 0$  ergibt sich  $\Psi(0) = 0$ ; wenn wir  $v = 0$  setzen, ergibt sich  $\Psi(-w) = \Psi(w)$ ; für  $v = w$  erhalten wir  $\Psi(2v) = 4 \cdot \Psi(v)$ . Wir beweisen durch vollständige Induktion  $\Psi(nv) = n^2 \Psi(v)$  für alle  $n \in \mathbb{N}$ . In der Tat gilt

$$\begin{aligned} \Psi((n+1)v) &= 2 \cdot \Psi(nv) + 2 \cdot \Psi(v) - \Psi((n-1)v) = (2n^2 + 2 - (n-1)^2) \Psi(v) = (n^2 + 2n + 1) \Psi(v). \\ \Psi(v) &= \Psi(n \cdot \frac{1}{n}v) = n^2 \cdot \Psi(\frac{1}{n}v); & \Psi(\frac{1}{n}v) &= \frac{1}{n^2} \Psi(v); & \Psi(\frac{m}{n}v) &= (\frac{m}{n})^2 \Psi(v). \end{aligned}$$

**Lemma 2.4.2** (Polarisierungskonstruktion).

Zur Funktion  $\Psi(\cdot)$ , welcher die Parallelogrammgleichung erfüllt, sei definiert

$$\Phi(v, w) = \frac{1}{4} [\Psi(v+w) - \Psi(v-w)] = \Phi(w, v).$$

Für alle  $u, v, w$  gilt dann  $\Phi(u, v+w) = \Phi(u, v) + \Phi(u, w)$ .

$\Phi(\cdot, \cdot)$  ist die einzige symmetrische biadditive Funktion mit  $\Phi(v, v) = \Psi(v)$ .

Wenn  $\Psi(\cdot)$  linearstetig ist, dann ist  $\Phi(\cdot, \cdot)$  eine Bilinearform.

**Beweis.** Wegen  $u+v = u + \frac{v+w}{2} + \frac{v-w}{2}$ ,  $u+w = u + \frac{v+w}{2} - \frac{v-w}{2}$  und den entsprechenden Gleichungen für  $-v, -w$  haben wir

$$\begin{aligned} \Psi(u+v) + \Psi(u+w) &= 2 \cdot \Psi(u + \frac{v+w}{2}) + 2 \cdot \Psi(\frac{v-w}{2}) \\ \Psi(u-v) + \Psi(u-w) &= 2 \cdot \Psi(u - \frac{v+w}{2}) + 2 \cdot \Psi(\frac{v-w}{2}) \end{aligned}$$

Subtraktion ergibt

$$4 \cdot \Phi(u, v) + 4 \cdot \Phi(u, w) = 8 \cdot \Phi(u, \frac{v+w}{2}) = 4 \cdot \Phi(u, v+w)$$

wobei sich die letzte Gleichung aus dem Spezialfall Fall  $v = w$  ergibt:  $\Phi(u, 2v) = 2 \cdot \Phi(u, v)$ . Durch vollständige Induktion erhält man  $\Phi(u, nv) = n \cdot \Phi(u, v)$  für alle  $n \in \mathbb{N}$ . und damit  $\Phi(u, rv) = r \cdot \Phi(u, v)$  für alle rationalen  $r$ . Wenn  $\Psi(\cdot)$  linearstetig ist, dann sind  $\Psi(u+tv)$  und  $\Psi(u-tv)$  stetige Funktionen von  $t$ , und  $\Phi$  ist bilinear. Umgekehrt gilt: Wenn die Polarisierungskonstruktion für  $\Psi(\cdot)$  auf eine Bilinearform führt, dann ist  $\Psi(\cdot)$  linearstetig; denn  $\Psi(u+tv) = \Psi(u) + 2t\Phi(u, v) + t^2\Psi(v)$ . Es gilt  $\Phi(v, v) = \Psi(v)$ , und für jede biadditive symmetrische  $\Theta(\cdot, \cdot)$  mit dieser Eigenschaft gilt

$$4 \cdot \Theta(v, w) = \Theta(v+w, v+w) - \Theta(v-w, v-w) = \Psi(v+w) - \Psi(v-w) = 4 \cdot \Phi(v, w).$$

### Traditionelle Ansätze der Begriffsbestimmung

Die folgende Definition eines Kegelschnitts stammt aus dem alten Griechenland (vgl. Coxeter[5] S. 149)

**Definition.** Ein Kegelschnitt ist der geometrische Ort eines Punkts  $P$ , dessen Abstand von einem festen Punkt  $O$  gleich  $\epsilon$  mal seinem Abstand von einer festen Geraden  $g$  ist, wo  $\epsilon$  eine positive Konstante ist. Für  $\epsilon < 1$  heisst der Kegelschnitt eine Ellipse, für  $\epsilon = 1$  eine Parabel, und für  $\epsilon > 1$  eine Hyperbel.  $\epsilon$  heisst die Exzentrizität des Kegelschnitts. Der Punkt  $O$  heisst der Brennpunkt, die Gerade  $g$  die Leitlinie.

Es ist eine lohnende Aufgabe in der Schulgeometrie, die geometrische Beschreibung in eine Beschreibung mit Polarkordinaten und in eine Beschreibung mit cartesischen Koordinaten überzuführen.

Die Definition der Griechen platziert die Kegelschnitte in die **euklidische Ebene**. Das Wort Kegelschnitt suggeriert eine etwas allgemeinere Klasse von Objekten. Wenn man einen Kegel (im reellaffinen dreidimensionalen Raum) mit einer Ebene schneidet, dann erhält man nämlich nicht nur Ellipsen, Parabeln und Hyperbeln; man erhält auch (für die Ebenen durch den Scheitel) Paare sich schneidender Geraden oder auch einen isolierten Punkt.

Die Klassifikation aller Kegelschnitte in der **affinen Ebene** war bis in die Mitte des 20. Jahrhunderts hinein ein beliebtes Thema der Lehrbücher über ‘Analytische Geometrie’. Metrische Begriffe wie Brennpunkte, Leitlinien, Hauptachsen usw. kamen da mangels einer euklidischen Struktur nicht ins Bild. Andere Ideen wie die von Pol und Polare (bei nichtausgeartete Kegelschnitten) traten in den Vordergrund. Manche Lehrbücher machten sich die Mühe, alle Quadriken im dreidimensionalen **reellaffinen Raum** zu klassifizieren. Bei dieser Gelegenheit lernte man u. U. auch einiges über spezielle Quadriken wie Ellipsoide oder Hyperboloide. Man kann fragen, welche dieser Themen überholt sind. Können oder sollen die oben entwickelten Betrachtungen über Bilinearformen in der Lehre an ihre Stelle treten.

Hinweis: In den vergangenen Jahrzehnten wurden in den Vorlesungen über ‘Linearen Algebra’ Quadriken auch in **höherdimensionalen nicht notwendigerweise reellen affinen Räumen** definiert und studiert; (man nennt diese Punktmenge auch Hyperflächen zweiter Ordnung.) Der Begriff der Quadrik setzt natürlich voraus, dass der Begriff der quadratischen Funktion in einem affinen Raum geklärt ist; denn eine Quadrik im affinen Raum ist definiert als das Nullstellengebilde einer **quadratischen Funktion**. Wir überlassen es der Algebra, quadratischen Funktionen auf einem  $\mathbb{K}$ -affinen Raum zu untersuchen.

Man kann die Sache etwas übersichtlicher machen, indem man zu homogenen Koordinaten übergeht. Man hat dann die quadratischen Formen zu studieren. Eine quadratische Form ist die Einschränkung einer (eindeutig bestimmten!) symmetrischen Bilinearform auf die ‘Diagonale’  $\Psi(v) = \Phi(v, v)$ . Es ist Sache der Tensoralgebra, die Bilinearformen und allgemeiner die Multilinearformen (insbesondere die symmetrischen und die alternierenden Multilinearformen) zu studieren. Besonders die alternierenden Multilinearformen

haben große Bedeutung für die Geometrie; das kommt allerdings üblicherweise erst in der Vorlesung über Differentialgeometrie zum Vorschein. Die Frage stellt sich: Ist es möglich oder gar wünschenswert, diese Begriffsbildungen bereits in den Anfängerveranstaltungen vorzubereiten, oder sollen wir uns Zeit nehmen für die Quadriken, wie das im 20. Jahrhundert üblich war?

### Didaktische Anmerkungen zu den Definitionen und der ‘Anschauung’.

Die quadratischen Funktionen sind ein passender Anlass, über die Rolle der Definitionen in der Mathematik zu sinnieren. Man kann immer wieder einmal die Frage hören, ob die heute üblichen formalen Definitionen die anvisierten Begriffe wirklich treffen?

Es sollte zu denken geben, dass die quadratischen Objekte (Kurven, Flächen, Funktionen, Formen) über Jahrhunderte einen prominenten Platz in der Mathematik und in der Lehre der Mathematik eingenommen haben. Die Definitionen und die Art des Definierens haben sich aber grundlegend verändert; die neueren quadratischen Objekte sind daher nicht mehr ganz die alten. In den Debatten um den mathematischen Lehrstoff sollte die Frage erörtert werden, inwieweit die quadratischen Objekte geeignet sind oder es verdienen, weiterhin einen besonderen Platz in unserem Mathematikverständnis einzunehmen. Manchmal wird dem widersprochen; man sagt, es drehe sich doch heute alles um Linearität und Linearisierung. Als ein spezifisches Argument für die Behandlung der quadratischen Objekte hat lange die Anschaulichkeit gedient; man kann und sollte aber fragen, in welchem Sinn dieses Argument noch aktuell ist.

Hinter den alten griechischen Definitionen der Kegelschnitte stand zweifellos eine Anschauung. Man sieht die Kegelschnitte, wenn man einen Lichtkegel erzeugt und den ausgeleuchteten Bereich auf einem Schirm beobachtet. Im geometrischen Schulunterricht schätzt man es, wenn die Schüler die Gegenstände, die mathematisch behandelt werden sollen, irgendwie auch sinnlich erfahren erfahren können. Andererseits gehört zu einem reifen Verständnis von moderner Mathematik die Einsicht, dass die mathematischen Gegenstände nicht durch die Anschauung vorgegeben werden, sondern durch Definitionen; die Mathematik behandelt keine empirisch gegebenen Objekte oder Zusammenhänge. Das ist den Erkenntnistheoretikern wichtig. Immanuel Kant sagt ausdrücklich, dass Definitionen an der Spitze eines Axiomensystems stehen müssen, sodass das Verständnis seines begrifflichen Inhalts vollständig von den Definitionen abhängt. *“Dagegen haben wir in der Mathematik gar keinen Begriff vor der Definition, als durch welche der Begriff allererst gegeben wird, sie muss also und kann auch jederzeit davon anfangen.”* Das behagt beileibe nicht allen, die mit Mathematik zu tun haben. Und ergibt sich eine didaktische Herausforderung.

Durch mathematische Definitionen wird abgegrenzt, welche Objekte oder Zusammenhänge die Mathematik genauer studieren möchte. Bei der Abgrenzung geht es in der Regel nicht um singuläre Objekte oder Typen von Objekten, die man bereits gut zu kennen glaubt, wie etwa die Kegelschnitte. Weitergefasste Definitionen oder Axiomensysteme **schaffen** Objektklassen, die sich durch gemeinsame Eigenschaften auszeichnen.

Wenn der Mathematiker eine neue Definition formuliert hat, weiss er nicht von vorneherein, welche (möglicherweise recht verschiedenartigen, ihm mehr oder weniger bekannten) Objekte oder Systeme damit erfasst werden. Mit der Definition entstehen Probleme der Klassifikation. Es ist durchaus erwünscht, dass der Mathematiker bei ihrer Ausarbeitung manchmal auf Objekte oder Systeme stößt, an die vorher niemand gedacht hat. Manche solche Fundstücke verdienen sehr wohl eine nähere Untersuchung. — Ein Beispiel wäre etwa die nichteuklidische Geometrie nach Bolyai und Lobatchevski.

Nach der reinen Lehre muss man von mathematische Definitionen nur verlangen, dass sie in sich widerspruchsfrei sind. Die Mathematiker können sich (etwa durch Analogiebetrachtungen) allerlei Definitionen und Axiomensysteme einfallen lassen; und auf der Grundlage formaler Phantasie ist auch schon viel interessante Mathematik entstanden. Bei manchen Definitionen der sog. Strukturmathematik kann allerdings der Eindruck entstehen, dass sie ausschliesslich die Funktion haben, irgendwelche Klassifikationsbemühungen auf den Weg zu bringen. Bei Definitionen ohne ‘anschaulichen’ Hintergrund besteht jedoch die Gefahr, dass sich die Untersuchungen in wenig fruchtbaren logischen Schlüssen verlieren. Es gibt Studierende der Mathematik, die das stringente logische Schliessen so sehr lieben, dass sie gern auf jedweden anschaulichen Hintergrund verzichten; sie wollen sich auch keine ‘Vorstellungen’ von den Gegenständen zu machen, die durch Definitionen innerhalb formaler Systeme geschaffen werden. Ebenso wenig fragen sie, ob die ‘Gegenstände’ auch ausserhalb des formalen Systems Bestand haben, ob man auf alternativen Wegen möglicherweise zu den ‘gleichen’ Gegenständen gelangt, und wie sich die auf anderen Wegen gewonnenen Einsichten zusammenfügen. (Ein Prüfungssystem, welches das Gelernte in kleinsten Happen abprüft, befördert übrigens solche geistige Abstinenz.)

Dieser Typ eines Formalisten stellt aber doch eine Minderheit dar. In der Regel sind die an der Mathematik Interessierten froh, wenn die formalen Untersuchungen immer wieder einmal konvergieren und sich schliesslich auch zusammenreimen in einer adäquaten Anschauung. Die meisten an der Mathematik Interessierten sind beispielsweise doch sehr beeindruckt von der Einsicht, dass die Kegelschnitte die Planetenbahnen beschreiben. Die analytische Beschreibung ergänzt hier auf schier wunderbare Weise die synthetische; und es entsteht in überzeugender Weise ein umfassenderes ‘Bild’ von den ‘quadratischen Objekten’.

Wir werden im nächsten Unterabschnitt die selbstadjungierten Operatoren studieren. Wir werden sie zunächst kennenlernen als eine bequeme Präsentationsform der hermiteschen Formen in einem Hilbertraum; man könnte sie daher als quadratische Objekte verstehen. In der Tat sind die Spektralwerte eines selbstadjungierten Operators eng verwandt mit den Hauptachsen eines Kegelschnitts. Es ist aber eine offene Frage, auf welchen Wegen diese ‘formalen’ Zusammenhänge mit einer adäquaten Anschauung unterlegt werden können. Eine Anschaulichkeit neuer Art wäre jedenfalls dringend zu wünschen angesichts der zentralen Bedeutung, welche diese Gegenstände in der Quantenmechanik einnehmen.

### Ergänzungen zu traditionellen Themen

Bevor wir uns wieder dem Hauptanliegen unserer Veranstaltung zuwenden, der Geometrie des Hilbertraums, machen wir noch (ohne ausführliche Beweise) einige elementare Anmerkungen zum klassischen Thema der quadratischen Funktionen in einem  $n$ -dimensionalen affinen Raum. Der zugrundeliegende reellaffine Raum hat hier keine euklidische Struktur; es ist also keine positivdefinite quadratische Form ausgezeichnet, es gibt kein Skalarprodukt und keine Orthogonalität. Fast alle Überlegungen lassen sich problemlos übertragen auf  $\mathbb{K}$ -wertige Funktionen auf endlichdimensionalen  $\mathbb{K}$ -Vektorräumen.

Die Menge der quadratischen Funktionen ist ein Vektorraum, und zwar im Falle eines  $n$ -dimensionalen affinen Raums  $A$  ein Vektorraum der Dimension  $1 + n + \frac{n(n+1)}{2}$ . Das punktweise Produkt zweier affiner Funktionen ist eine quadratische Funktion; im endlichdimensionalen Fall bilden diese Produkte ein aufspannendes System. Bei den Formen gilt Entsprechendes. Man kann genauer sagen: Wenn  $x^1, x^2, \dots, x^n$  eine Basis des Raums  $V^*$  der Linearformen ist, dann bilden die Produkte  $x^j \cdot x^k$  mit  $j \leq k$  eine Basis des Vektorraums der quadratischen Formen. Die Beweise sind dem Leser überlassen.

### Die Idee der quadratischen Ergänzung

Aus dem Schulunterricht sollte die Konstruktion der quadratischen Ergänzung bekannt sein. Wir wollen an einigen Zahlenbeispielen zeigen, wie die Idee funktioniert. Wir werden (ohne strenge Beweise) verstehen, dass jede quadratische Funktion auf einem endlichdimensionalen  $\mathbb{K}$ -Vektorraum ( $1 + 1 \neq 0$ ) als Linearkombination von Quadraten und einer affinen Funktion gewonnen werden kann.

*Beispiel.* Es seien  $x = x(P)$ ,  $z = z(P)$  und  $y = y(P)$  affine Funktionen. Für die folgenden (exemplarischen) Funktionen  $f$ ,  $g$ ,  $h$  gilt

$$\begin{aligned}
 f = f(P) &= 11x^2 + 19y^2 + 6x - 38y + 15 \\
 &= 11 \cdot (x^2 + 2 \cdot \frac{3}{11}x) + 19 \cdot (y^2 - 2y) + 15 \\
 &= 11 \cdot (x + \frac{3}{11})^2 + 19 \cdot (y - 1)^2 - \frac{9}{11} - 19 + 15. \\
 g = g(P) &= x^2 + 4xy + 5y^2 + 10z^2 + 2xz + 10yz - 2z - 2 \\
 &= x^2 + 2x(2y + z) + 5y^2 + 10yz - 2z - 2 \\
 &= (x + 2y + z)^2 - 4y^2 - 4yz - z^2 + 5y^2 + 10z^2 + 10z - 2 \\
 &= (x + 2y + z)^2 + y^2 + 6yz + 9z^2 - 2z - 2 \\
 &= (x + 2y + z)^2 + (y + 3z)^2 - 2(z + 1). \\
 h = h(P) &= xy + z^2 - 1 \\
 &= \frac{1}{4}(x + y)^2 - \frac{1}{4}(x - y)^2 + z^2 - 1 \\
 k = k(P) &= x^2 + 11y^2 + 10\sqrt{3}xy = (x + 5\sqrt{3}y)^2 - 64y^2 \\
 &= 16 \cdot \left(\frac{1}{2}x + \frac{1}{2}\sqrt{3}y\right)^2 - 4 \cdot \left(-\frac{1}{2}\sqrt{3}x + \frac{1}{2}y\right)^2.
 \end{aligned}$$

Es zeigt sich, dass  $n$ -dimensionalen Raum nicht mehr als  $n$  Quadrate benötigt werden. Wir wollen das im (etwas übersichtlicheren) Fall der quadratischen Formen etwas genauer diskutieren.

**Sprechweise.** Wir sagen, dass der Vektor  $v$  zum Nullraum  $V_0$  der quadratischen Form  $\Psi(\cdot)$  gehört, wenn für alle  $w$  gilt  $\Psi(v+w) - \Psi(v-w) = 0$ , wenn also  $\Phi(v, \cdot)$  die Nullform ist. Wenn der Nullraum  $V_0$  die Dimension  $n - r$  hat, dann heisst  $r$  der Rang von  $\Psi$ .

Die Idee der quadratischen Ergänzung liefert (hier ohne Beweis) den

**Satz 2.4.3.** *Wenn  $\Psi(\cdot)$  eine quadratische Form vom Rang  $r$  ist, dann existieren linear unabhängige Linearformen  $\ell^1, \dots, \ell^r$  und Koeffizienten  $d_1, \dots, d_r \neq 0$ , sodass gilt*

$$\Psi(\cdot) = d_1 \cdot (\ell^1(\cdot))^2 + \dots + d_r \cdot (\ell^r(\cdot))^2.$$

Alle  $\ell^k$  verschwinden auf dem Nullraum von  $\Psi$ .

Die Konstruktion lässt viel Freiheit. Im reellen Fall gibt es aber eine wichtige Invariante, die sich auf die Vorzeichen der Koeffizienten bezieht.

**Sprechweise.** Es sei  $\Psi(\cdot)$  eine quadratische Form auf dem  $n$ -dimensionalen reellen Vektorraum  $V$ . Wir sagen,  $\Psi(\cdot)$  sei indefinit, wenn  $\Psi(\cdot)$  sowohl positive als auch negative Werte annimmt. Wenn die Einschränkung von  $\Psi(\cdot)$  auf den Teilvektorraum  $W$  positivdefinit ist, dann sagen wir  $W$  sei totalpositiv für  $\Psi(\cdot)$ . Entsprechend sagen wir,  $W$  sei totalnegativ für  $\Psi(\cdot)$ , wenn gilt  $\Psi(w) < 0$  für alle  $w \in W \setminus \{0\}$ . Man nennt  $w$  einen isotropen Vektor, wenn  $\Psi(w) = 0$ . Die Menge aller isotropen Vektoren heisst der Nullkegel der quadratischen Form.

**Satz 2.4.4** (Sylvesters Trägheitsgesetz). *Es sei  $\Psi(\cdot)$  eine indefinite quadratische Form vom Rang  $r$  auf einem  $n$ -dimensionalen reellen Vektorraum  $V$ .  $V_0$  sei der Nullraum (der Dimension  $n_0 = n - r$ ). Es existiert dann eine direkte Zerlegung  $V = V_0 \oplus V_+ \oplus V_-$  mit totalpositivem  $V_+$  und totalnegativen  $V_-$ .*

*Für die entsprechend zerlegten  $v = v_0 + v_+ + v_-$  gilt  $\Psi(v) = \Psi(v_+) + \Psi(v_-)$ .*

*Die Dimensionen  $n_+ = \dim V_+$ , und  $n_- = \dim V_-$  sind durch  $\Psi$  eindeutig bestimmt.*

**Sprechweise 2.4.1.** Eine reelle quadratische mit  $n_0 = 0$  heisst eine nichtausgeartete quadratische Form der Signatur  $(n_+, n_-)$ .

**Beweis.** *Da  $\Psi$  indefinit ist, gibt es einen Vektor  $w_1$  so, dass der von ihm aufgespannte eindimensionale Vektorraum totalnegativ ist. Es sei  $W$  ein totalnegativer Teilraum maximaler Dimension  $n_-$ , und es sei*

$$W^\Psi = \{v : \Phi(v, \cdot) = 0 \text{ für alle } w \in W\}.$$

*$W^\Psi$  hat die Dimension  $n - n_-$ . Wenn nämlich  $w_1, \dots, w_{n_-}$  eine Basis von  $W$  sind, dann sind die dazugehörige Linearformen  $\Phi(w_j, \cdot)$  linear unabhängig.  $\sum a_j \cdot \Phi(w_j, \cdot)$  verschwindet nicht im Punkt  $\sum a_j w_j$ . Wir zeigen, dass  $\Psi(\cdot)$  auf diesem Vektorraum positiv semi-definit ist. Für  $\tilde{v} \in W^\Psi$  und  $w \in W$  gilt nämlich  $\Psi(w + a\tilde{v}) = \Psi(w) + a^2\Psi(\tilde{w})$ . Wäre*

$\Psi(\tilde{w}) < 0$ , dann wäre der vergrößerte Raum  $\text{span}\{W, \tilde{w}\}$  totalnegativ. Wenn  $V_+$  ein Komplement von  $V_0$  in  $W^\Psi$  ist, dann ist  $V_+$  totalpositiv mit der Dimension  $n_+ = n - r - n_-$ . Ist  $V = V_0 \oplus V'_+ \oplus V'_-$  eine weitere direkte Zerlegung mit  $V'_+$  totalpositiv,  $V'_-$  totalnegativ, so folgt aus  $V'_+ \cap (V_0 \oplus V_-) = \{0\}$  zunächst  $n'_+ = \dim V'_+ \leq n - (r + n_-) = n_+$ . Aus Symmetriegründen ergibt sich  $n'_+ = n_+$ ,  $n'_- = n_-$ .

Man bemerke: Die Konstruktion lässt viel Freiheit, wie man daraus sieht, dass es zu jedem  $w_1$  mit  $\Psi(w_1) < 0$  eine direkte Zerlegung gibt, sodass  $w_1 \in V_-$ .

Eine ähnliche Argumentation führt auch für quadratische Formen auf einem  $\mathbb{K}$ -Vektorraum zu interessanten Aussagen. Der Bequemlichkeit halber befassen wir uns nur mit quadratischen Formen vom vollen Rang  $n$ . Andernfalls kann man den (wohlbestimmten!) Nullraum abspalten und die Einschränkung von  $\Psi$  auf irgendeinen Komplementärraum von  $V_0$  studieren.

### Sprechweise 2.4.2 (Verträgliche direkte Zerlegungen).

Es sei  $\Psi$  eine quadratische Form von vollem Rang auf dem  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraum  $V$ . (Man spricht auch von einer nichtausgearteten quadratischen Form.)

Man sagt von einer direkten Zerlegung  $V = V_1 \oplus \cdots \oplus V_k$ , sie sei verträglich mit  $\Psi$ , wenn für die ihr entsprechend zerlegten Vektoren  $v = v_1 + \cdots + v_k$  gilt  $\Psi(v) = \Psi(v_1) + \cdots + \Psi(v_k)$ .

**Lemma.** Sei  $\Psi$  eine nichtausgeartete quadratische Form auf  $V$  ( $\dim V = n$ ) und  $W$  ein  $m$ -dimensionaler Teilraum, dann ist  $W^\Psi = \{v : \Phi(v, \cdot) = 0 \text{ für alle } w \in W\}$  ein  $(n - m)$ -dimensionaler Teilraum mit  $W \cap W^\Psi = \{0\}$ . (Man nennt  $W^\Psi$  das  $\Psi$ -Komplement von  $W$ .) Es gilt

1.  $(W_1 + W_2)^\Psi = W_1^\Psi \cap W_2^\Psi \quad (W^\Psi)^\Psi = W.$
2.  $V = V_1 \oplus \cdots \oplus V_k$  ist genau dann verträglich mit  $\Psi$ , wenn  $V_i \subseteq V_j^\Psi$  für alle  $i \neq j$ .

Der Beweis ist derselbe wie bei den orthogonalen Komplementen. Die Einsicht, die uns die Methode der quadratischen Ergänzung erbracht hat, kann man jetzt folgendermaßen formulieren:

**Satz 2.4.5.** Zu jeder quadratischen Form existiert eine verträgliche direkte Zerlegung in eindimensionale Teilräume.

Seien nämlich  $\ell^1, \ell^2, \dots, \ell^n$  wie bei der Konstruktion der quadratischen Ergänzungen linear unabhängige Linearformen und  $d_1, d_2, \dots, d_n$  Skalare, sodass

$$\Psi(\cdot) = d_1 \cdot (\ell^1(\cdot))^2 + \cdots + d_n \cdot (\ell^n(\cdot))^2.$$

Sei  $(v_1, v_2, \dots, v_n)$  die duale Basis, und sei  $V_i$  der von  $v_i$  aufgespannte Teilraum. Wir haben dann für die Vektoren  $v = \sum v_i \cdot y^i$  die verträgliche Zerlegung  $\Psi(v) = \sum d_i \cdot (y^i)^2$ .

### Rückkehr zur Geometrie des Hilbertraums

Wir haben immer wieder darauf hingewiesen, dass die Methode der quadratischen Ergänzung große Freiheiten lässt, wenn es darum geht, eine quadratische Form als Linearkombination von quadrierten Linearformen darzustellen. Im euklidischen Raum gibt es nun aber ganz besondere verträgliche Zerlegungen, solche nämlich die eindeutig durch die Forderung bestimmt sind, dass auf jedem direkten Summanden die quadratische Form ein Vielfaches der euklidischen Norm ist. Diesen ganz besonderen Zerlegungen, den sog. Spektralzerlegungen wollen wir uns im nächsten Unterabschnitt zuwenden, (und zwar nicht nur im euklidischen Fall, sondern auch im komplexen Fall, wo es um die Zerlegungen der hermiteschen Formen geht.)

Ein Zahlenbeispiel haben wir oben bereits kennengelernt,

$$\begin{aligned}\Psi(v) &= x^2(v) + 11y^2(v) + 10\sqrt{3}xy(v) = (x + 5\sqrt{3}y)^2 - 64y^2 \\ &= 16 \cdot \left(\frac{1}{2}x + \frac{1}{2}\sqrt{3}y\right)^2 - 4 \cdot \left(-\frac{1}{2}\sqrt{3}x + \frac{1}{2}y\right)^2.\end{aligned}$$

Wenn wir davon ausgehen, dass die Linearformen  $x, y$  ein orthonormales Koordinatensystem auf dem Vektorraum  $V$  sind, dann wird man mit der ersten Darstellung unserer quadratischen Form  $\Psi$  Form als Linearkombination von Quadraten nicht in gleichem Maße zufrieden sein wie mit der zweiten, die ja ebenfalls orthonormierte Koordinaten benützt. Die Linearformen  $\ell^1 = \frac{1}{2}x + \frac{1}{2}\sqrt{3}y$ ,  $\ell^2 = -\frac{1}{2}\sqrt{3}x + \frac{1}{2}y$  bilden in der Tat ein Orthonormalsystem. ('Cartesische Koordinaten')

Hier kann man sich wirklich erinnert fühlen an die Hyperbelgleichung aus dem Schulunterricht:  $\frac{1}{a^2}x^2 - \frac{1}{b^2}y^2 = 1$ . Man würde in der Schule möglicherweise sagen: Für die Hyperbel  $\{P : k(P) = 1\}$  sind  $a = 1/4$ ;  $b = 1/2$  die Längen der Hauptachsen, und die sind um  $60^\circ$  verdreht.

In der Linearen Algebra würde man das Resultat der Rechnung so fassen

$$A^T \cdot \begin{pmatrix} 1 & 5\sqrt{3} \\ 5\sqrt{3} & 11 \end{pmatrix} \cdot A = \begin{pmatrix} 16 & 0 \\ 0 & -4 \end{pmatrix} \quad \text{mit} \quad A = \begin{pmatrix} 1/2 & -1/2\sqrt{3} \\ 1/2\sqrt{3} & 1/2 \end{pmatrix} \quad A^T \cdot A = E.$$

In diesem Kontext versteht man die Rechnung als ein Zahlenbeispiel zu dem

**Satz.** *Zu jeder reellen symmetrischen Matrix  $B$  existiert eine orthogonale Matrix  $A$ , so dass  $A^T B A = D$  eine Diagonalmatrix ist.*

In mehr geometrischer Sprache handelt es sich um ein Beispiel für den angekündigten

**Satz 2.4.6.** *Ist  $k(\cdot)$  eine quadratische Form in einem  $n$ -dimensionalen euklidischen Raum  $(V, \|\cdot\|)$ , dann existieren reelle Zahlen  $d_1, \dots, d_n$  und ein orthonormales System von Linearformen  $\ell^1, \dots, \ell^n$ , sodass gilt*

$$k(\cdot) = d_1 \cdot (\ell^1(\cdot))^2 + \dots + d_n \cdot (\ell^n(\cdot))^2.$$

Diesen und den entsprechenden Satz im Hilbertraum werden wir im nächsten Unterabschnitt beweisen.

## 2.5 Hermitische Formen im Hilbertraum. Spektralzerlegung.

**Definition 2.13** (Hermitische Form, hermitische Sesquilinearform).

Eine reellwertige linear stetige Funktion  $\Psi(\cdot)$  auf einem komplexen Vektorraum  $V$  nennt man eine hermitische Form, wenn sie die Parallogrammgleichung erfüllt und  $\Psi(iv) = \Psi(v)$ .

Die komplexwertige Funktion  $\Phi(\cdot, \cdot)$  auf  $V \times V$  heisst eine hermitische Sesquilinearform, wenn gilt

- (i)  $\Phi(v, w) = \overline{\Phi(w, v)}$
- (ii)  $\Phi(v, \alpha w) = \alpha \cdot \Phi(v, w)$  für alle  $\alpha \in \mathbb{C}$
- (iii)  $\Phi(v, w_1 + w_2) = \Phi(v, w_1) + \Phi(v, w_2)$ .

**Satz 2.5.1** (Polarisierungskonstruktion). *Zu jeder hermitischen Form  $\Psi(\cdot)$  existiert genau eine hermitische Sesquilinearform  $\Phi(\cdot, \cdot)$  mit  $\Phi(v, v) = \Psi(v)$ . Es gilt*

$$\Phi(v, w) = \frac{1}{4} [\Psi(v + w) - \Psi(v - w) + i \cdot \Psi(v - iw) - i \cdot \Psi(v + iw)]$$

**Beweis 2.5.1.**

1) *Zunächst zeigen wir die Eindeutigkeit: Ist  $\Theta(\cdot, \cdot)$  eine hermitische Sesquilinearform mit  $\Theta(v, v) = \Psi(v)$  so gilt*

$$\begin{aligned} \Theta(v + w, v + w) - \Theta(v - w, v - w) &= 2\Theta(v, w) + 2\Theta(w, v), \\ i\Theta(v - iw, v - iw) - i\Theta(v + iw, v + iw) &= 2i\Theta(v, -iw) + 2i\Theta(-iw, v) \\ &= 2\Theta(v, w) - 2\Theta(w, v). \end{aligned}$$

*Die Summe ergibt  $\Psi(v + w) - \Psi(v - w) + i \cdot \Psi(v - iw) - i \cdot \Psi(v + iw) = 4\Theta(v, w)$ .*

2) *Die Rechnung im vorigen Abschnitt zeigt, dass in der Polarisierungskonstruktion sowohl der Realteil als auch der Imaginärteil biadditiv ist. Insbesondere gilt, zunächst für alle rationalen  $r$ , dann aber (wegen der linearen Stetigkeit) auch für alle reellen  $r$   $\Phi(v, rw) = r\Phi(v, w)$ .*

3) *Die Formel zeigt sofort  $\Phi(v, iw) = i\Phi(v, w)$ . Wegen  $\Psi(iv) = \Psi(v)$  haben wir weiter*

$$\begin{aligned} \Psi(iv - iw) &= \Psi(v - w), & \Psi(iv + iw) &= \Psi(v + w) \\ (-4i)\Phi(iv, w) &= -i\Psi(iv + w) - i\Psi(iv - w) - \Psi(v - w) + \Psi(v + w) = 4\Phi(w, v). \end{aligned}$$

Der Satz liefert den Beweis des Satzes 2.1.5 von Jordan und von Neumann auf Seite 87. Wir brauchen nur zu bemerken, dass aus der positiven Definitheit die lineare Stetigkeit der Funktion  $\Psi(\cdot) = \|\cdot\|^2$  folgt. In der Tat gilt wegen der Dreiecksungleichung (die sich aus der Schwarz'schen Ungleichung ergibt)

$$\begin{aligned} \Psi(v + t_1 w) - \Psi(v + t_0 w) &= \|(v + t_0 w) + (t_1 - t_0) \cdot w\|^2 - \|v + t_0 w\|^2 \\ &\leq (\|v + t_0 w\| + |t_1 - t_0| \cdot \|w\|)^2 - \|v + t_0 w\|^2 = 2c \cdot |t_1 - t_0| + d \cdot |t_1 - t_0|^2. \end{aligned}$$

*Beispiel* (Hermitische Formen vom Rang 1). Wenn  $\ell(\cdot)$  eine Linearform ist, dann ist  $\Psi(\cdot) = |\ell(\cdot)|^2$  eine hermitische Form; denn es gilt  $|\ell(iv)|^2 = |\ell(v)|^2$  und

$$\begin{aligned} |\ell(v+w)|^2 + |\ell(v-w)|^2 &= \ell(v+w)\overline{\ell(v+w)} + \ell(v-w)\overline{\ell(v-w)} = \\ &= 2\ell(v)\overline{\ell(v)} + 2\ell(w)\overline{\ell(w)} = 2|\ell(v)|^2 + 2|\ell(w)|^2 \end{aligned}$$

Die hermitische Sesquilinearform dazu ist  $\Phi(v, w) = \overline{\ell(v)}\ell(w)$ , wie man leicht nachrechnet.

In der Notation von Dirac (für einen Hilbertraum!) liefert die Linearform  $\ell = \langle u |$  die hermitische Form  $\Psi(v) = \|\ell(v)\|^2 = \langle v|u\rangle\langle u|v\rangle$  und die hermitische Sesquilinearform  $\Phi(v, w) = \langle v|u\rangle\langle u|w\rangle$ , kurz geschrieben  $\Phi(\cdot, \cdot) = |u\rangle\langle u|$ .

Wir erwähnen einen (wenig geometrisch anmutenden) Satz, der aus der Linearen Algebra bekannt sein dürfte:

**Satz.** Wenn  $V$  ein  $n$ -dimensionaler komplexer Vektorraum ist, dann ist der Raum  $\mathcal{H}$  aller hermitischen Sesquilinearformen ein  $n^2$ -dimensionaler reeller Vektorraum. Wenn  $(u_1, \dots, u_n)$  eine Basis von  $V$  ist, dann wird ein  $h(\cdot, \cdot) \in \mathcal{H}$  durch eine hermitische Matrix  $H$  beschrieben

$$h\left(\sum u_i y^i, \sum u_j x^j\right) = \sum_{ij} \overline{y^i} h_{ij} x^j = \mathbf{y}^* \cdot H \cdot \mathbf{x}.$$

Die Diagonaleinträge von  $H$  sind reelle Zahlen, die Einträge über der Diagonalen können irgendwelche komplexen Zahlen sein, die Einträge unter der Diagonalen ergeben sich aus der Hermitizität. So ergibt sich die reelle Dimension  $\dim \mathcal{H} = n + 2 \cdot \frac{1}{2}n(n-1) = n^2$ . Wir bemerken, dass der Rang der Matrix  $H$  in der Tat der Rang der Sesquilinearform ist. Insbesondere sind die Formen vom Rang 1 durch Matrizen der Gestalt  $h_{ij} = \bar{a}_i \cdot a_j$  gegeben.

Unsere Aufmerksamkeit gilt im Folgenden den hermitischen Formen auf einem  $n$ -dimensionalen Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ . Wenn wir eine Basis einführen, dann wird es immer eine ONB sein. Ein Basiswechsel wird daher immer durch eine unitäre Matrix  $U$  bewerkstelligt. ( $U^*U = E$ ). Die darstellende Matrix  $H$  einer hermitischen Form transformiert sich bei einem Basiswechsel nach der Regel  $H \rightsquigarrow U^*HU = U^{-1}HU$ . Wir erinnern daran, dass auch der Dualraum von  $V$  ein Hilbertraum ist. Wenn  $(w_1, \dots, w_n)$  eine Orthonormalbasis von  $V$  ist, dann ist das Tupel der bra-Vektoren  $(\langle w_1 |, \dots, \langle w_n |)$  eine Orthonormalbasis  $(\ell^1(\cdot), \dots, \ell^n(\cdot))$  des Dualraums  $V^*$ .

Der Hauptsatz über hermitische Formen auf einem Hilbertraum hat mehrere nützliche Formulierungen und viele Namen. ‘Satz von der Hauptachsendarstellung’ und ‘Spektralsatz’ sind zwei solche Namen.

**Satz 2.5.2** (Satz von den Hauptachsen).

Wenn  $\Psi(\cdot)$  eine hermitische Form auf einem  $n$ -dimensionalen Hilbertraum  $V$  ist, dann existieren reelle Zahlen  $d_1 \geq d_2 \geq \dots \geq d_n$  und eine ONB  $(\ell^1(\cdot), \dots, \ell^n(\cdot))$  von  $V^*$ , sodass

$$\Psi(\cdot) = d_1 \cdot |\ell^1(\cdot)|^2 + \dots + d_n \cdot |\ell^n(\cdot)|^2.$$

Die nichtsteigende Folge der ‘Spektralwerte’  $d_1 \geq d_2 \geq \dots \geq d_n$  ist eindeutig bestimmt.

Wir werden einen dazu äquivalenten Satz beweisen; zuvor wollen wir aber noch einige Begriffe einführen, die für die Einsicht und den Beweis nützlich sind.

**Sprechweise 2.5.1** (Eigenvektor, Eigenwert, Eigenraum).

Ist  $\varphi : V \rightarrow V$  ein Endomorphismus und  $\tilde{v} \neq 0$  ein Vektor mit  $\varphi(\tilde{v}) = d \cdot \tilde{v}$ , dann heisst  $\tilde{v}$  ein Eigenvektor für  $\varphi(\cdot)$  zu Eigenwert  $d$ .

Der Teilraum  $V_d = \{v : \varphi(v) = d \cdot v\}$  heisst der Eigenraum zum Eigenwert  $d$ .

Bemerke: Der Eigenraum  $V_d$  ist der Nullraum für den Endomorphismus  $\varphi - d \cdot id$ . Einen Nullraum gibt es für jede lineare Abbildung; Eigenräume zu  $d$ , (die für die meisten  $d$  trivial sind,) gibt es aber nur für Endomorphismen.

**Sprechweise 2.5.2.** Ein Endomorphismus eines endlichdimensionalen Hilbertraums wird ein selbstadjungierter Operator genannt, wenn für alle  $w, v$  gilt  $\langle w | \varphi(v) \rangle = \langle \varphi(w) | v \rangle$ .

**Satz 2.5.3.** Sei  $\Psi(\cdot)$  eine hermitesche Form auf einem Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ , und sei  $\Phi(\cdot, \cdot)$  die dazugehörige hermitesche Sesquilinearform. Es existiert dann genau ein Endomorphismus  $\varphi : V \rightarrow V$ , sodass gilt

$$\Phi(w, v) = \langle w | \varphi(v) \rangle = \langle \varphi(w) | v \rangle.$$

Umgekehrt liefert jeder selbstadjungierte Operator  $\varphi(\cdot)$  eine hermitesche Sesquilinearform  $\Phi(w, v) = \langle \varphi(w) | v \rangle$ .

**Beweis.** Wenn  $\Phi(\cdot, \cdot)$  eine Sesquilinearform ist, dann existiert (nach dem Satz von Riesz) für jedes  $w$  ein bra-Vektor  $\langle \varphi(w) |$ , sodass  $\Phi(w, \cdot) = \langle \varphi(w) | \cdot \rangle$ . Die Zuordnung  $w \mapsto \varphi(w)$  ist ein Endomorphismus. Ihre Selbstadjungiertheit ergibt sich aus

$$\langle \varphi(w) | v \rangle = \Phi(w, v) = \overline{\Phi(v, w)} = \overline{\langle \varphi(v) | w \rangle} = \langle w | \varphi(v) \rangle.$$

Die zweite Behauptung liegt auf der Hand.

Wir haben oben bereits die folgende Notation vorbereitet:

**Notation.** Wenn  $u$  ein Einheitsvektor ist, und  $\Psi(\cdot) = |\langle u | \cdot \rangle|^2$ , dann ist die von dieser hermiteschen Form erzeugte hermitesche Sesquilinearform

$$\Phi(\cdot, \cdot) = |u\rangle\langle u|, \quad \text{genauer} \quad \Phi(w, v) = \langle w | u \rangle \langle u | v \rangle \text{ für alle } w, v,$$

und der dazugehörige selbstadjungierte Operator ist die orthogonale Projektion auf den von  $u$  aufgespannten Vektorraum.

Wenn  $(u_1, \dots, u_k)$  ein ONBasis des Teilraums  $U$  ist, dann ist der selbstadjungierte Operator zu  $\Phi(\cdot, \cdot) = \sum_j |u_j\rangle\langle u_j|$  die orthogonale Projektion auf  $U$ .

In dieser Notation können wir den Satz 2.5.2 folgendermaßen umformen

**Satz 2.5.4** (Existenz einer ONBasis aus Eigenvektoren).

Zu jeder hermitischen Sesquilinearform  $\Phi(\cdot, \cdot)$  existiert eine absteigende Folge reeller Zahlen  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  und eine ONB  $(u_1, \dots, u_n)$ , sodass

$$\Phi(\cdot, \cdot) = \lambda_1 \cdot |u_1\rangle\langle u_1| + \dots + \lambda_n \cdot |u_n\rangle\langle u_n|.$$

Die Basisvektoren  $u_j$  sind Eigenvektoren für den dazugehörigen selbstadjungierten Operator  $\varphi(\cdot) = \sum \lambda_j \cdot |u_j\rangle\langle u_j| = \sum \lambda_j \cdot \pi_j(\cdot)$ .  $\varphi(u_j) = \lambda_j \cdot u_j$ .

Die  $u_j$  nennt man auch Hauptachsenvektoren (zu den Eigenwerten  $\lambda_j$ ). Die  $\lambda_j$  sind offenbar gerade die Zahlen, für welche  $\varphi(\cdot)$  einen nichttrivialen Eigenraum besitzt. Die Menge der Zahlen  $\lambda_j$  mit ihren Vielfachheiten wird das **Spektrum des Operators** genannt.

Wir können dem Hauptsatz auch noch eine Form geben, in welcher alle auftretenden Objekte durch die hermitische Form  $\Psi(\cdot)$  (oder die hermitische Sesquilinearform  $\Phi(\cdot, \cdot)$ , oder den selbstadjungierten Operator  $\varphi(\cdot)$ ) **eindeutig bestimmt** sind. Dieses Ziel erreichen wir, indem wir in der obigen Summendarstellung die Summanden mit dem gleichen Eigenwert zusammenzufassen. Es bezeichne also  $n_{(\lambda)}$  die Vielfachheit, mit welcher  $\lambda$  in der absteigenden Folge auftritt und  $\pi_{(\lambda)}$  die orthogonale Projektion auf den  $n_{(\lambda)}$ -dimensionalen Vektorraum  $V_{(\lambda)}$ , der von den betreffenden  $u_j$  aufgespannt wird. Dies führt zu der folgenden Version des Hauptsatzes über hermitische Formen in einem  $n$ -dimensionalen Hilbertraum:

**Satz 2.5.5** (Spektralzerlegung).

Ist  $\varphi(\cdot)$  ein selbstadjungierter Operator in einem endlichdimensionalen Hilbertraum, so existieren paarweise disjunkte orthogonale Projektionen  $\pi_{(\lambda)}$ , sodass

$$\varphi(\cdot) = \sum \lambda \cdot \pi_{(\lambda)}(\cdot), \quad \text{wo die Summe über die Menge der Eigenwerte zu erstrecken ist.}$$

Zur Eingewöhnung in die Konstruktionen und Notationen bemerken wir:

Anstelle der hermitischen Form  $\Psi(\cdot)$  untersucht man manchmal lieber die hermitische Form  $\Psi_c(\cdot) = \Psi(\cdot) + c \cdot \|\cdot\|^2$  (mit irgendeinem  $c \in \mathbb{R}$ ).

Wenn  $\Phi(\cdot, \cdot)$  die hermitische Sesquilinearform und  $\varphi(\cdot)$  der selbstadjungierte Operator zu  $\Psi(\cdot)$  sind, dann ergeben sich jetzt  $\Phi_c(\cdot, \cdot) + c\langle \cdot, \cdot \rangle$  und  $\varphi_c(\cdot) + c \cdot id$ ; denn für alle  $w$  gilt

$$\langle \varphi(w) + cw | \cdot \rangle = \Phi(w, \cdot) + c\langle w | \cdot \rangle = \Phi_c(w, \cdot) = \langle \varphi_c(w) | \cdot \rangle.$$

Die Eigenwerte  $\lambda_j$  sind durch  $\lambda_j + c$  zu ersetzen. Unverändert können bleiben die zu konstruierenden ONBasen und die paarweise disjunkten orthogonalen Projektionen. Ist nämlich  $(\ell^1(\cdot), \dots, \ell^n(\cdot)) = (\langle u_1 | \cdot \rangle, \dots, \langle u_n | \cdot \rangle)$  eine Orthonormalbasis von  $V^*$ , so gilt

$$\begin{aligned} \sum |\ell^j(\cdot)|^2 &= \|\cdot\|^2, & \langle w | v \rangle &= \sum \bar{\ell}^j(w) \cdot \ell^j(v) = \sum \langle w | u_j \rangle \langle u_j | v \rangle, \\ v &= \langle \cdot | v \rangle = \sum |u_j\rangle \langle u_j | v \rangle; & \langle \cdot, \cdot \rangle &= \sum |u_j\rangle \langle u_j|. \end{aligned}$$

Bei den Beweisen können wir uns jedenfalls auf den Fall einer positivdefiniten hermiteschen Form beschränken. Die Menge  $\{v; \Psi(v) \leq 1\}$  können wir uns dann als ein Ellipsoid vorstellen, dessen ‘Hauptachsen’ wir bestimmen wollen. Nach all den Umformungen der Behauptungen ist es jetzt Zeit für die Beweise. Wir stützen uns auf ein berühmtes Argument von R. Courant, welches nicht auf Operatoren und Eigenwerte gerichtet ist, sondern direkt (und effektiv) an der hermiteschen Form und der dazugehörigen hermiteschen Sesquilinearform angreift.

**Beweis 2.5.2.** *Irgendwo auf der Einheitssphäre unseres Hilbertraums  $S = \{v; \|v\|^2 = 1\}$  nimmt die hermitesche Form ihr Maximum  $\lambda_1$  an, sagen wir im Einheitsvektor  $u_1$ .*

$$\lambda_1 = \sup\{\Psi(u) : \|u\| = 1\} = \Psi(u_1) \quad \text{mit} \quad \|u_1\| = 1.$$

Wir werden sehen, dass die Menge aller  $u$  mit  $\Psi(u) = \lambda_1 \cdot \|u\|^2$  ein Vektorraum  $V_1$  ist. Der Durchschnitt mit der Einheitssphäre  $V_1 \cap S$  ist die  $n_1$ -dimensionale Sphäre  $S_1$ , in welcher das Ellipsoid  $\{w : \Psi(w) \leq \lambda_1\}$  die Einheitssphäre trifft. Um das zu sehen, betrachten wir für ein beliebiges  $v \in V$  die folgenden Funktionen von  $t \in \mathbb{C}$ :

$$\begin{aligned} f(t) &= \Psi(u_1 + tv) - \Psi(u_1) = 2\Re\Phi(u_1, tv) + |t|^2\Psi(v), \\ g(t) &= \|u_1 + tv\|^2 - \|u_1\|^2 = 2\Re\langle u_1 | tv \rangle + |t|^2\|v\|^2 \\ f(t) - \lambda_1 g(t) &= 2\Re(\Phi(u_1, tv) - \lambda_1\langle u_1 | tv \rangle) + |t|^2(\Psi(v) - \lambda_1\|v\|^2) \leq 0 \quad \text{für alle } t. \end{aligned}$$

Für alle  $v \perp u_1$  folgt  $\Phi(u_1, v) = 0$  aus der Ungleichung.

Für  $v \perp u_1$  mit  $\Psi(v) = \lambda_1\|v\|^2$  ergibt sich zusätzlich zunächst  $\Psi(u_1 + tv) - \lambda_1\|u_1 + tv\|^2 = 0$  für alle  $t$  und dann auch  $\Psi(au_1 + bv_1) - \lambda_1\|au_1 + bv_1\|^2 = 0$  für alle  $a, b$ .

Jedes  $w \in V$  besitzt eine Darstellung  $w = au_1 + v$  mit  $v \perp u_1$ . Es gilt daher

$$\forall w \in V : \Phi(u_1, w) - \lambda_1\langle u_1 | w \rangle = 0 \iff \Psi(u_1) - \lambda_1\|u_1\|^2 = 0.$$

So ist also  $\{u : \Psi(u) = \lambda_1\|u\|^2\} = \{u : \Phi(u, \cdot) = \lambda_1\langle u | \cdot \rangle\} = \{u : \varphi(u) = \lambda_1 \cdot u\}$  der Eigenraum  $V_1$  zum größten Eigenwert  $\lambda_1$ .

Die hermitesche Form  $\Psi(\cdot)$  ist verträglich mit der orthogonalen Zerlegung  $V = V_1 \oplus V_1^\perp$  in dem Sinne, dass für die entsprechend zerlegten Vektoren gilt

$$w = u_1 + v \implies \Psi(w) = \Psi(u_1 + v) = \Psi(u_1) + \Psi(v).$$

Den Eigenraum zum zweitgrößten Eigenwert finden wir, indem wir unser Verfahren auf die Einschränkung von  $\Psi(\cdot)$  auf den Raum  $V_1^\perp$  anwenden. — Und so gewinnen wir induktiv die orthogonale Zerlegung  $V = V_1 \overset{\perp}{\oplus} V_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} V_K$ , wo  $V_k$  der Eigenraum zum Eigenwert  $\lambda_k$  ist, mit der strikt abfallenden Folge  $\lambda_1 > \lambda_2 > \dots > \lambda_K$ .

Wenn wir nun in jedem der Eigenräume ( $V_\lambda$  von der Dimension  $n_\lambda$ ) eine ONB einführen, dann liefert uns das ein ONB  $u_1, u_2, \dots, u_n$  des Gesamtraums  $V$ , welche aus Eigenvektoren zu den reellen Eigenwerten  $d_1 \geq d_2 \geq \dots \geq d_n$  besteht. Die duale Basis  $(\langle u_1 |, \langle u_2 |, \dots, \langle u_n |) = (\ell^1(\cdot), \ell^2(\cdot), \dots, \ell^n(\cdot))$  ist ein  $n$ -tupel von paarweise orthogonalen Linearformen der Norm 1 mit der Eigenschaft  $\Psi = \sum d_j \cdot |\ell^j(\cdot)|^2$ .

**Satz 2.5.6** (Satz von Fischer und Courant).

Es seien  $\Psi(v)$  und  $\Psi'(v)$  hermitische Formen auf einem  $n$ -dimensionalen Hilbertraum mit  $\Psi(\cdot) \geq \Psi'(\cdot)$ . Für die Spektralwerte  $d_1 \geq d_2 \geq \dots \geq d_n$  und  $d'_1 \geq d'_2 \geq \dots \geq d'_n$  gilt dann  $d_j \geq d'_j$  für alle  $j$ .

**Beweis 2.5.3.** Es sei  $\Psi(\cdot) = d_1 \cdot |\ell^1(\cdot)|^2 + d_2 \cdot |\ell^2(\cdot)|^2 + \dots + d_n \cdot |\ell^n(\cdot)|^2$  mit ortho-normierten  $\ell^j$ . Die dazu duale ONB sei  $u_1, u_2, \dots, u_n$ . Die ersten  $j$  dieser Basisvektoren spannen einen Vektorraum  $W_j = V_1 \oplus V_2 \oplus \dots \oplus V_j$  auf; sein orthogonales Komplement ist  $W_j^\perp = V_{j+1} \oplus \dots \oplus V_n$ . Es gilt

$$\begin{array}{ll} v \in V_1 \implies \Psi(v) = d_1 \|v\|^2, & v \in V_1^\perp \implies \Psi(v) \leq d_1 \|v\|^2, \\ v \in W_2 \implies \Psi(v) \geq d_2 \|v\|^2, & v \in W_2^\perp \implies \Psi(v) \leq d_3 \|v\|^2, \\ & \dots \quad \dots \\ v \in W_j \implies \Psi(v) \geq d_j \|v\|^2, & v \in W_j^\perp \implies \Psi(v) \leq d_{j+1} \|v\|^2, \\ & \dots \quad \dots \end{array}$$

Es sei  $U$  eine Hyperebene. Da  $\dim U + \dim V_2 = n+1$  und daher  $U \cap V_2$  nicht der Nullraum ist, haben wir

$$d_1 \geq \sup \left\{ \frac{\Psi(v)}{\|v\|^2} : v \in U \right\} \geq d_2.$$

Wir betrachten jetzt allgemeiner ein Nullstellengebilde  $U = \{v : L(v) = 0\}$  eines homogenen Gleichungssystems  $L$  vom Rang  $j-1$ . Da  $\dim U + \dim W_j = n+1$ , gibt es einen Vektor im Durchschnitt. Es gilt

$$\sup \left\{ \frac{\Psi(v)}{\|v\|^2} : L(v) = 0 \right\} \geq d_j \quad \text{für alle } L \text{ vom Rang } j-1$$

Andererseits ist auf dem Raum  $W_{j-1}^T = \{v : \ell^1(v) = 0, \dots, \ell^{j-1}(v) = 0\}$  der Quotient  $\leq d_j$ . Das ergibt Courants Charakterisierung des  $j$ -ten Spektralwerts:

$$d_j = \inf \left\{ \sup \left\{ \frac{\Psi(v)}{\|v\|^2} : L(v) = 0 \right\} : L \text{ vom Rang } j-1 \right\}.$$

Wenn wir die hermitische Form verkleinern, dann kann der  $j$ -te Spektralwert nicht größer werden.

Die Charakterisierung leistet noch etwas mehr.

**Satz 2.5.7.** Es sei  $\Psi(v)$  eine hermitische Formen auf einem  $n$ -dimensionalen Hilbertraum mit den Spektralwerten  $d_1 \geq d_2 \geq \dots \geq d_n$ . Wenn man die hermitische Form auf eine Hyperebene  $U$  einschränkt, dann erhält man eine hermitische Form, für deren Spektralwerte  $\mu_j$  gilt  $d_1 \geq \mu_1 \geq d_2 \geq \mu_2 \geq d_3 \geq \dots \geq d_{n-1} \geq \mu_{n-1} \geq d_n$ .

**Beweis.** Die Hyperebene sei durch die Linearform  $\tilde{\ell}$  gegeben.

$$\mu_j = \inf \left\{ \sup \left\{ \frac{\Psi(v)}{\|v\|^2} : \tilde{\ell}(v) = 0, L(v) = 0 \right\} : L \text{ vom Rang } j - 1 \right\} \leq d_{j+1},$$

weil hier das Infimum über eine eingeschränkte Klasse von Systemen vom Rang  $j$  erstreckt wird. Das Argument auf die hermitesche Form  $-\Psi$  angewendet, ergibt  $d_j \geq \mu_j$ .

### Eine didaktische These

Nach einer pragmatischen Auffassung vom Matrizenkalkül gibt es da zwei zentrale Rechenaufgaben. Das ‘Lösen’ linearer Gleichungssysteme, insbesondere das Invertieren einer Matrix ist die eine. Die andere, anspruchsvollere, ist das Auffinden von Eigenvektoren und Eigenwerten.

In der Frage der Eigenwerte gibt es nun aber deutliche Unterschiede zwischen der algebraischen und der numerisch orientierten Sichtweise. Für die Algebraiker spielt das charakteristische Polynom eine zentrale Rolle, und damit kommen die Determinanten ins Spiel, sowie der sog. Fundamentalsatz der Algebra, wenn man nicht gleich bei den allgemeinen  $\mathbb{K}$ -Vektorräumen und der abstrakten Algebra bleiben will. Die Numeriker versuchen aus guten Gründen die Berechnung von charakteristischen Polynomen und ihren Nullstellen zu vermeiden. Sie beschränken sich lieber darauf, die Hauptachsen einer quadratischen Form (und simultan ihre Spektralwerte) zu bestimmen. Verfahren im Sinne des Satzes von Fischer und Courant haben dabei große Bedeutung.

Wenn es um die mathematische Grundbildung geht, wird man den Berechnungsproblemen natürlich nicht den höchsten Stellenwert einräumen. Algebraische Begrifflichkeit und geometrische Anschauung sind die wichtigeren Aspekte in der Anfängerausbildung. Aber auch unter diesen Gesichtspunkten scheinen mir die Überlegungen zum Hauptachsenproblem durchaus Vorteile zu bieten gegenüber der heute üblichen algebraischen Herangehensweise an die charakteristischen Polynome. Damit plädiere ich nicht für Themen wie die Klassifikation der Quadriken im alten Stil. Ich meine aber doch, dass die sog. Jordan’sche Normalform einer Matrix nicht die Aufmerksamkeit verdient, die ihr heute in den Lehrtexten der Linearen Algebra entgegengebracht wird. Eigenwerte sind vor allem da wichtig, wo sie als Spektralwerte in Erscheinung treten. Und da eröffnen sich auch Türen des Verständnisses.

## 2.6 Diagonalisierbare Endomorphismen

**Sprechweise 2.6.1.** Es sei  $\varphi$  ein Endomorphismus eines  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraums  $V$ . Wir sagen von einer direkten Zerlegung  $V = \bigoplus V_j$ , dass sie mit  $\varphi$  verträglich ist, wenn  $\varphi$  jedes  $V_j$  in sich abbildet.

**Lemma.** *Es seien  $\pi_j$  die Projektionen zur direkten Zerlegung  $V = \bigoplus V_j$ . Die folgenden Bedingungen sind notwendig und hinreichend für die Verträglichkeit von  $\varphi$  mit  $\pi_j$*

- $\forall j : \pi_j \circ \varphi \circ \pi_j = \varphi \circ \pi_j$ ,
- $\forall j \neq k : \pi_k \circ \varphi \circ \pi_j = 0$ ,
- $\forall j : \pi_j \circ \varphi = \varphi \circ \pi_j$ .

**Sprechweise 2.6.2.** Wir sagen von einem Endomorphismus  $\varphi$ , er sei diagonalisierbar, wenn es eine direkte Zerlegung  $V = \bigoplus V_j$  gibt, sodass für jedes  $j$  die Einschränkung von  $\varphi$  auf  $V_j$  ein Vielfaches der Identität ist,

**Lemma.** *Der Endomorphismus  $\varphi$  ist genau dann diagonalisierbar, wenn eine direkte Zerlegung  $id = \sum_j \pi_j$  und Koeffizienten  $\lambda_j$  existieren, sodass gilt  $\varphi = \sum_j \lambda_j \pi_j$*

Bemerkung: Wenn man die  $\pi$  zu gleichen Eigenwerten zusammenfasst, dann erhält man die direkte Zerlegung in die Eigenräume; wir nennen sie die Grobzerlegung des diagonalisierbaren Endomorphismus.

Wenn man andererseits in den direkten Summanden  $V_j = \pi_j(V)$  Basen einführt, dann ergibt das für  $V$  eine Basis, die aus Eigenvektoren besteht. In dieser Basis wird  $\varphi$  durch eine Diagonalmatrix dargestellt. Wir sprechen hier von einer Feinzerlegung des diagonalisierbaren Endomorphismus  $\varphi$ .

**Satz 2.6.1** (Simultan diagonalisierte kommutierende Endomorphismen). *Wenn zwei diagonalisierbare Endomorphismen eines  $\mathbb{K}$ -Vektorraums kommutieren,  $\varphi \circ \psi = \psi \circ \varphi$ , dann sind sie simultan diagonalisierbar, d. h. es gibt paarweise disjunkte Projektionen  $\{\pi_{jk} : j \in J, k \in K\}$  und Koeffizienten  $a_j$  und  $b_k$ , sodass gilt*

$$\varphi(\cdot) = \sum_{jk} a_j \cdot \pi_{jk}, \quad \psi(\cdot) = \sum_{jk} b_k \cdot \pi_{jk}.$$

**Beweis 2.6.1.** *Betrachten wir die Grobzerlegungen zu den kommutierenden Endomorphismen*

$$\varphi(\cdot) = \sum_j a_j \cdot \pi_j(\cdot), \quad \psi(\cdot) = \sum_k b_k \cdot \rho_k(\cdot).$$

Der Beweis des Satzes ist erbracht, wenn wir zeigen, dass die Projektionen  $\pi$  und  $\rho$  kommutieren, d. h.  $\pi_j \circ \rho_k = \rho_k \circ \pi_j$  für alle  $j, k$ .

Wir zeigen in einem ersten Schritt  $\psi \circ \pi_l = \pi_l \circ \psi$  für alle  $l \in J$ ; Da  $\bigoplus V_j$  die größte mit  $\varphi$  verträgliche Zerlegung ist, haben wir  $v \in V_j \Leftrightarrow \varphi(v) = a_j v$ . Daraus folgt

$v \in V_j \Rightarrow \psi(v) \in V_j$ ; denn  $\varphi(\psi(v)) = \psi(\varphi(v)) = a_j\psi(v)$ . Dieses Ergebnis kann man auch so schreiben:

$$\pi_l \circ \psi \circ \pi_j = \begin{cases} 0 & \text{falls } l \neq j \\ \psi \circ \pi_l(v) & \text{falls } l = j \end{cases}.$$

Summieren wir über alle  $j$ , so ergibt sich  $\pi_l \circ \psi = \psi \circ \pi_l$  für alle  $l$  (wegen  $\sum_j \pi_j = id$ ).

Wir fassen zusammen: Wenn  $\psi$  mit dem grobzerlegten  $\varphi$  kommutiert, dann kommutiert  $\psi$  auch mit den einzelnen Projektionen.

Das gleiche Argument zeigt: Da (für jedes  $l$ ) der Endomorphismus  $\pi_l$  mit  $\psi$  kommutiert, kommutiert er auch mit jeder der in der Grobzerlegung auftretenden Projektion.

Das Ergebnis kann auch folgendermaßen formuliert werden:

**Satz** (Simultane Diagonalisierbarkeit). *Es sei  $\{\varphi^{(\alpha)} : \alpha \in I\}$  eine Schar von diagonalisierbaren Endomorphismen, die paarweise kommutieren. Es existiert dann eine Basis, in welcher alle  $\varphi^{(\alpha)}$  durch eine Diagonalmatrix dargestellt werden.*

Oder in der Sprache der Matrizen.

**Satz.** *Sei  $\{A^{(\alpha)} : \alpha \in I\}$  eine Schar von diagonalisierbaren Matrizen, die paarweise kommutieren. Es existiert dann eine Matrix  $S$ , sodass  $S^{-1}A^{(\alpha)}S = D^{(\alpha)}$  (Diagonalmatrix).*

Hinweis: Wenn man sich eingehender mit der Diagonalisierbarkeit befassen will, dann benötigt man den Begriff des Minimalpolynoms eines Endomorphismus. Man kann zeigen: Ein Endomorphismus eines  $\mathbb{K}$ -Vektorraums ist genau dann diagonalisierbar, wenn sein Minimalpolynom in Linearfaktoren der Vielfachheit 1 zerfällt.

Dies wollen wir nicht weiter verfolgen. Wir wenden uns wieder den Hilberträumen zu.

## Unitäre Diagonalisierbarkeit

Wir haben aus hermiteschen Formen auf einem  $n$ -dimensionalen Hilbertraum selbstadjungierte Endomorphismen abgeleitet. Jetzt wollen bei den weiteren Konstruktionen von den Endomorphismen ausgehen. Entscheidende Bedeutung hat dabei der Begriff des adjungierten Operators. Wir formulieren die Definition für den unendlichdimensionalen Fall, obwohl wir uns dann sehr bald wieder auf die endlichdimensionalen Hilberträume beschränken wollen.

**Definition 2.14** (Adjungierte Operatoren).

Es sei  $(\tilde{V}, \langle \cdot | \cdot \rangle)$  ein Prähilbertraum und  $(V, \langle \cdot | \cdot \rangle)$  seine Vervollständigung.  $\varphi : \tilde{V} \rightarrow V$  sei eine lineare Abbildung. Jedes  $w \in V$  liefert eine Linearform  $\ell(\cdot) = \langle w | \varphi(\cdot) \rangle$ . Wenn es sich um eine beschränkte Linearform handelt, dann gibt es genau ein  $\tilde{w} \in V$ , sodass  $\ell(\cdot) = \langle \tilde{w} | \cdot \rangle$ . Wir schreiben in diesem Fall  $w \in \mathcal{D}$  und  $\tilde{w} = \varphi^*(w)$ . Die somit auf dem ‘Definitionsbereich’  $\mathcal{D}$  definierte lineare Abbildung  $\varphi^* : \mathcal{D} \rightarrow V$  heisst der zu  $\varphi$  adjungierte Operator.

Hinweis: Der Definitionsbereich ist ein Vektorraum, der im unendlichdimensionalen Fall nicht notwendigerweise abgeschlossen ist. — Dies macht Probleme, die in der Funktionalanalysis gründlich bearbeitet werden. Bevor wir uns wieder dem endlichdimensionalen Fall zuwenden, formulieren wir ein Beispiel (vgl. Satz 2.8.1 auf Seite 142):

*Beispiel.* Es sei  $(\tilde{V}, \langle \cdot | \cdot \rangle)$  der Raum der trigonometrischen Polynome. Die Vektoren  $v \in \tilde{V}$  kann man bekanntlich einerseits durch finite Folgen darstellen und andererseits durch  $2\pi$ -periodische Funktionen

$$f(t) = \sum c_n \cdot e^{int} \quad \longleftrightarrow \quad v \quad \longleftrightarrow \quad \mathbf{c} = (\dots, c_1, c_0, c_1, c_2, \dots),$$

Wir definieren eine lineare Abbildung  $\varphi(\cdot)$  in beiden Darstellungen

$$\frac{1}{i} \frac{\partial}{\partial t} f \quad \longleftrightarrow \quad \varphi(v) \quad \longleftrightarrow \quad \varphi(\mathbf{c}) = (\dots, (-1)c_1, 0, 1c_1, 2c_2, \dots),$$

Die Funktion  $e^{int}$  ist also Eigenvektor zum Eigenwert  $n$  (für alle  $n \in \mathbb{Z}$ ). Der adjungierte Operator  $\varphi^*$  ist eine Fortsetzung von  $\varphi$  auf einen Teilraum  $\mathcal{D} \subset V$ , welcher in  $V$  dicht liegt, aber nicht der ganze Hilbertraum ist.

In der Funktionalanalysis gibt man sich große Mühe, durch Wahl der passenden Definitionsbereiche Operatoren dieser Art zu ‘wesentlich’ selbstadjungierten Operatoren zu machen. Für Operatoren dieser Art gibt es nämlich eine wohlausgearbeitete Theorie mit ausserordentlich wichtigen Anwendungen.

### Adjungierte im Matrizenkalkül

Im Raum der Spalten  $V = \mathbb{C}_{Sp}^n$  ist jeder Endomorphismus  $\varphi$  durch eine Matrix  $A$  beschrieben. Wenn wir durch eine positivdefinite Matrix  $Q$  eine Hilbertraumstruktur einführen, dann ergibt sich für den adjungierten Operator die Matrix  $A_Q^* = Q^{-1}A^*Q$ . In der Tat gilt für alle  $\mathbf{x}, \mathbf{y}$

$$\langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle = \mathbf{y}^* Q \mathbf{A} \mathbf{x} = \mathbf{y}^* Q \mathbf{A} Q^{-1} Q \mathbf{x} = (Q^{-1} A^* Q \mathbf{y})^* Q \mathbf{x} = \langle \mathbf{A}_Q^* \mathbf{y} | \mathbf{x} \rangle.$$

Bemerke: Das Adjungieren ist eine antilineare Abbildung der Menge aller Endomorphismen in sich:  $(a \cdot \varphi_1 + b \cdot \varphi_2)^* = \bar{a} \cdot \varphi_1^* + \bar{b} \cdot \varphi_2^*$ . Und es gilt  $(\varphi^*)^* = \varphi$ ; denn

$$\langle w | \varphi(v) \rangle = \langle \varphi^*(w) | v \rangle = \overline{\langle v | \varphi^*(w) \rangle} = \overline{\langle \varphi^{**}(v) | w \rangle} = \langle w | \varphi^{**}(v) \rangle.$$

Im Falle des Standard-Skalarprodukts ist die adjungierte Abbildung einfach durch die hermitisch konjugierte Matrix gegeben. Die selbstadjungierten Operatoren entsprechen den hermitischen Matrizen. Man kann das auch so ausdrücken: Wenn wir in einem  $n$ -dimensionalen Hilbertraum eine Orthonormalbasis auszeichnen, dann ist jeder selbstadjungierte Operator  $\varphi$  durch eine hermitische Matrix gegeben. (Auf der anderen Seite ist übrigens jeder isometrische Endomorphismus durch eine unitäre Matrix gegeben.)

Den Spektralsatz können wir auch folgendermaßen formulieren: Wenn wir eine zu  $\varphi$  passende ONB einführen, dann ist die darstellende Matrix eine Diagonalmatrix mit reellen Diagonaleinträgen. In der Sprache der Matrizenrechnung ist das der

**Satz 2.6.2** (Simultan diagonalisierte hermitische Matrizen).

Gegeben sei eine positivdefinite Matrix  $C$  und eine hermitische Matrix  $H$ , beide vom Format  $n \times n$ . Es existieren dann reelle Zahlen  $d_1 \geq d_2 \geq \dots \geq d_n$  und eine Basis  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  des Spaltenraums  $\mathbb{C}_{Sp}^n$ , sodass gilt

$$\mathbf{a}_j^* \cdot C \cdot \mathbf{a}_k = \delta_{jk}, \quad \mathbf{a}_j^* \cdot H \cdot \mathbf{a}_k = d_j \cdot \delta_{jk}.$$

Es existiert also eine Matrix  $A$  und eine reelle Diagonalmatrix  $D$ , sodass

$$A^*CA = E, \quad A^*HA = D.$$

Die Menge der  $d_j$  mit ihren Vielfachheiten  $n_j$  ist eindeutig bestimmt.

**Beweis.** Wir machen den Spaltenraum mit  $C$  zu einem Hilbertraum  $(\mathbb{C}^2, C)$ ,  $\|\mathbf{a}\|^2 = \mathbf{a}^* \cdot C \cdot \mathbf{a}$ . Die Matrix  $H$  liefert eine hermitische Form  $\Psi(\mathbf{a}) = \mathbf{a}^* \cdot H \cdot \mathbf{a}$ . Nach dem Spektralsatz existiert eine ONB mit  $\Phi(\mathbf{a}_j, \mathbf{a}_k) = d_j \cdot \delta_{jk}$ . Die  $d_j$  sind die Eigenwerte mit ihren Vielfachheiten der Form  $\Psi$  in diesem Hilbertraum.

**Satz 2.6.3** (Satz von der unitären Diagonalisierbarkeit einer hermitischen Matrix).

Zu jeder hermitischen Matrix  $H$  existiert eine unitäre Matrix  $U$ , sodass

$$U^*HU = D \text{ (Diagonalmatrix)} \quad U^*U = E.$$

Wenn  $H$  reell ist, dann kann man  $U$  reell wählen. Zu jeder symmetrischen Matrix  $S$  existiert also eine reelle Matrix  $O$ , sodass  $O^T S O = D$ ,  $O^T O = E$ .

**Beweis.** Die Existenz ist der Spezialfall des obigen Satzes für  $C = E$ . Es gilt  $HU = UD$ . Die  $j$ -te Spalte der Matrix  $UD$  ist  $d_j$  mal die  $j$ -te Spalte der Matrix  $U$ , sie ist also Eigenvektor des Endomorphismus zur Matrix  $H$ , und zwar zum Eigenwert  $d_j$ . Wir bemerken, dass Eigenvektoren zu verschiedenen Eigenwerten aufeinander senkrecht stehen; denn  $\varphi(u) = du, \varphi(v) = ev \implies d\langle u|v \rangle = \langle \varphi(u)|v \rangle = \langle u|\varphi(v) \rangle = e\langle u|v \rangle$ ; und das ist im Falle  $d \neq e$  nur möglich, wenn  $u \perp v$ .

Wenn  $H$  reelle Einträge hat, dann ist für jeden Eigenvektor  $\mathbf{u}_j$  zum Eigenwert  $d_j$  auch der Real- und der Imaginärteil ein Eigenvektor. Sie stehen auch senkrecht auf den Eigenvektoren zu den übrigen Eigenwerten. Die Real- und Imaginärteile der Spalten  $\mathbf{u}_j$  zum Eigenwert  $d_j$  sind ein aufspannendes System des  $n_j$ -dimensionalen Vektorraums der reellen Eigenvektoren zum Eigenwert  $d_j$ . Jede ONBasis dieses reellen Vektorraums kann die Rolle der Spalten  $\mathbf{u}_j$  übernehmen.

**Satz 2.6.4** (Die positive Quadratwurzel).

Ist  $C$  positiv definit, so existiert genau eine positivdefinite Matrix  $B$  mit  $B \cdot B = C$ .

**Beweis.** Es sei  $U$  unitär, sodass  $U^*CU = D$  (Diagonalmatrix), also  $C = UDU^*$ . Die Diagonalmatrix hat positive Diagonaleinträge. Es sei  $D^{1/2}$  die Diagonalmatrix mit den Diagonaleinträgen  $\sqrt{d_j}$ . Es ist die einzige positivdefinite Matrix, deren Quadrat  $D$  ist. Die Matrix  $B = UD^{1/2}U^*$  ist positivdefinit und es gilt  $B^2 = UD^{1/2}D^{1/2}U^* = C$ . Ist  $F$  eine weitere positivdefinite Matrix mit  $F^2 = C$ , so gilt  $(U^*FU)(U^*FU) = U^*CU = D$ , daher  $U^*FU = D^{1/2}$ ,  $F = B$ .

**Satz 2.6.5** (Polarfaktorisierung). *Zu jeder nichtsingulären Matrix  $A$  existiert eine positivdefinite Matrix  $R$  und eine unitäre Matrix  $V$ , sodass gilt  $A = R \cdot V$ . Die Faktoren  $R$  und  $V$  sind eindeutig bestimmt.*

**Beweis.** *Die Matrix  $AA^*$  ist positiv definit. Sei  $R = (AA^*)^{1/2}$ ; dann ist  $V = R^{-1}A$  unitär; denn  $V^*V = A^*R^{-1}R^{-1}A = E$ . Wenn  $A = R' \cdot V'$  mit  $R'$  positivdefinit und  $V'$  unitär. Dann gilt  $AA^* = R' \cdot V'V'^* \cdot R' = R'^2$ , und nach dem vorigen Satz  $R' = R$ .*

**Sprechweise 2.6.3** (Normale Endomorphismen und normale Matrizen).

Ein Endomorphismus  $\varphi$  eines Hilbertraums heisst normal, wenn er mit seinem adjungierten kommutiert.  $\varphi^* \circ \varphi = \varphi \circ \varphi^*$ .

Eine komplexe  $n \times n$ -Matrix heisst eine normale Matrix, wenn die Matrizen  $N$  und  $N^*$  kommutieren, wenn also gilt  $NN^* = N^*N$ .

Die wichtigsten Beispiele normaler Endomorphismen sind die selbstadjungierten Endomorphismen und die isometrischen Selbstabbildungen des Hilbertraums.

Die wichtigsten Beispiele normaler Matrizen sind die hermiteschen Matrizen und die unitären Matrizen.

**Satz 2.6.6.** *Die Matrix  $N$  ist genau dann normal, wenn die hermiteschen Matrizen  $S = \frac{1}{2}(N + N^*)$  und  $T = \frac{1}{2i}(N - N^*)$  kommutieren. Die nichtsinguläre Matrix  $N$  ist genau dann normal, wenn die Faktoren in der Polarfaktorisierung kommutieren.*

**Beweis.** *Es gilt  $N = S + iT$ ,  $N^* = S - iT$ . (Man nennt  $S$  den hermiteschen und  $iT$  den antihermiteschen Anteil der Matrix  $N$ .)  $NN^* = S^2 + T^2 + i(TS - ST)$ ,  $N^*N = S^2 + T^2 + i(-TS + ST)$ . Diese beiden Matrizen sind genau dann gleich, wenn  $TS - ST = 0$ . Wenn  $N = RV$  die Polarfaktorisierung ist, dann gilt  $N^* = V^*R$  und  $NN^* = R^2$ ,  $N^*N = V^*R^2V$ . Die beiden nichtsingulären Matrizen sind genau dann gleich, wenn  $R^2 = V^*R^2V = (V^*RV)(V^*RV)$ . Da aber die positivdefinite Matrix  $R^2$  nur eine Darstellung als Quadrat einer positivdefiniten Matrix besitzt, haben wir  $R = V^*RV$ ,  $VR = RV$ .*

Wir haben oben gesehen, dass ein selbstadjungierter Operator diagonalisierbar ist; nach dem Spektralsatz 2.5.5 lässt er sich als reelle Linearkombination von paarweise orthogonalen Projektionen dargestellt.  $\varphi = \sum \lambda_j \pi_j$ . Wenn  $\pi = \sum z_j \pi_j$  eine komplexe Linearkombination paarweise orthogonaler Projektionen ist, dann gilt  $\pi^* = \sum \bar{z}_j \pi_j$  wegen der Selbstadjungiertheit der  $\pi_j$ . Da nun für einen normalen Operator  $\varphi$  die Operatoren  $\frac{1}{2}(\varphi + \varphi^*)$  und  $\frac{1}{2}(\varphi - \varphi^*)$  kommutierende Endomorphismen sind, haben wir die Verhältnisse von oben; es existiert eine simultane Diagonalisierung. Wir werden folgern, dass die normalen Operatoren gerade die komplexen Linearkombination paarweise orthogonaler Projektionen sind. Zunächst bemerken wir Wenn ein Endomorphismus als komplexe Linearkombination paarweise orthogonaler Projektionen dargestellt werden kann, dann kommutieren  $\varphi$  und  $\varphi^*$ , denn  $\varphi^* = \sum \bar{\lambda}_j \pi_j$ . Wir werden zeigen, dass die komplexen Linearkombinationen von paarweise orthogonalen Projektionen gerade die normalen Endomorphismen sind.  $\varphi$  und  $\varphi^*$  sind daher diagonalisierbare Operatoren im Sinne der obigen allgemeineren Überlegungen. Zunächst bemerken wir

**Satz 2.6.7.** Wenn  $\varphi$  ein normaler Endomorphismus ist, dann haben  $\varphi$  und  $\varphi^*$  dieselben Eigenvektoren.  $\varphi(w) = \lambda \cdot w \iff \varphi^*(w) = \bar{\lambda} \cdot w$ .

**Beweis.** Wenn  $\varphi^* \circ \varphi = \varphi \circ \varphi^*$ , dann haben  $\varphi$  und  $\varphi^*$  denselben Nullraum; denn  $\|\varphi(w)\|^2 = \langle \varphi(w) | \varphi(w) \rangle = \langle \varphi^* \circ \varphi(w) | w \rangle = \langle \varphi \circ \varphi^*(w) | w \rangle = \langle \varphi^*(w) | \varphi^*(w) \rangle = \|\varphi^*(w)\|^2$ . Der adjungierte Operator zu  $\varphi - \lambda \cdot id$  ist  $\varphi^* - \bar{\lambda} \cdot id$ .

Wir wissen bereits, dass die gemeinsamen Eigenvektoren  $u$  der kommutierenden selbstadjungierten Operatoren  $\sigma = \frac{1}{2}(\varphi + \varphi^*)$  und  $\tau = \frac{1}{2i}(\varphi - \varphi^*)$  den ganzen Raum aufspannen. Diese gemeinsamen Eigenvektoren sind auch gemeinsame Eigenvektoren der Operatoren  $\varphi$  und  $\varphi^*$ .  $\sigma(u) = au, \tau(u) = bu \iff \varphi(u) = (a + ib)u, \varphi^*(u) = (a - ib)u$ . Die Eigenvektoren eines normalen Operators  $\varphi$  spannen den ganzen Raum auf.

Wir haben also bewiesen.

**Satz 2.6.8** (Satz von der unitären Diagonalisierbarkeit eines normalen Operators).

Ein Endomorphismus  $\varphi$  eines Hilbertraums  $V$  ist genau dann ein normaler Operator ( $\varphi \circ \varphi^* = \varphi^* \circ \varphi$ ), wenn er unitär diagonalisierbar ist, d. h.  $\varphi = \sum \lambda_j \pi_j$  mit paarweise orthogonalen Projektionen  $\pi_j$ . Dies ist genau dann der Fall, wenn es eine ONB gibt, die aus Eigenvektoren besteht.

Bemerke: Sei  $u$  ein Eigenvektor des invertierbaren normalen Operators  $\varphi = \rho \circ \chi$  (mit  $\rho$  positivdefinit,  $\chi$  unitär),  $\varphi(u) = (a + ib)u = r \cdot e^{i\phi}u$ .

Dann ist  $u$  auch Eigenvektor für  $\rho$  und  $\chi$ .  $\rho u = ru, \chi(u) = e^{i\phi}u$ .

Die unitären Operatoren sind diejenigen normalen Operatoren, deren Eigenwerte den Betrag 1 haben.

**Der reelle Fall** Eine hermitesche Matrix mit reellen Einträgen heisst auch eine symmetrische reelle Matrix.  $A = A^T = A^*$ . Als hermitesche Matrix hat sie reelle Eigenwerte  $d_1, \dots, d_n$ . Man kann eine ONB des Spaltenraums wählen, die aus reellen Eigenspalten besteht. Wenn  $S$  die Matrix mit diesen Spalten ist, dann gilt  $A \cdot S = S \cdot D$ , wo  $D$  die Diagonalmatrix mit den Diagonaleinträgen  $d_j$  ist. Wir bemerken: reelle Matrizen mit  $A^T A = E$  (man nennt sie orthogonale Matrizen) haben nicht notwendigerweise lauter reelle Eigenwerte; die unitären Matrizen  $S$ , die sie diagonalisieren ( $S^{-1} A S = D$ ) man kann sie i. Allg. nicht reell wählen. Die Sache ist hier ein wenig komplizierter. Ein Beispiel einer orthogonalen Matrix ohne reelle Eigenwerte ist  $A = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$ . Wir halten fest:

**Satz 2.6.9** (Orthogonale Diagonalisieren einer symmetrischen reellen Matrix).

Wenn  $A$  eine symmetrische reelle Matrix ist.  $A = A^T$ . Es existiert dann eine orthogonale Matrix  $S \in O(n, \mathbb{R})$ , ( $S^T S = E$ ), sodass  $S^T A S = D$  (Diagonalmatrix).

Diesen Satz haben wir bereits angekündigt, als es darum ging, eine reelle quadratische Form  $\Psi(v)$  in einem  $n$ -dimensionalen euklidischen Raum als Linearkombination von  $n$  quadrierten Linearformen darzustellen. Man kann die Linearformen  $\ell^j(v)$  so wählen, dass sie eine ONB bilden:  $\Psi(v) = \sum_{j=1}^n d_j \cdot |\ell^j(v)|^2$ . Die Koeffizienten mit ihren Vielfachheiten sind eindeutig bestimmt: Wenn sie paarweise verschieden sind, dann sind die  $\ell^j$  bis auf den Faktor  $\pm 1$  eindeutig bestimmt.

## 2.7 Kontinuierliche Drehungen. Kommutatoren.

Für eine beliebige komplexe  $n \times n$ -Matrix  $A$  und  $t \in \mathbb{C}$  definiert man bekanntlich

$$e^{tA} = \exp(tA) = E + tA + \frac{1}{2!}t^2A^2 + \frac{1}{3!}t^3A^3 + \dots$$

(Die Reihe konvergiert gleichmässig, wenn  $tA$  in einem kompakten Bereich liegt.)

Man beweist: Wenn  $A$  und  $B$  kommutieren ( $AB = BA$ ), dann gilt  $e^{A+B} = e^A e^B$ .

Entsprechend definiert man für Endomorphismen  $\alpha$  eines reellen oder komplexen Vektorraums und  $t \in \mathbb{C}$  die Endomorphismen  $\exp(t\alpha) = id + t\alpha + \frac{1}{2!}t^2\alpha^2 + \dots$ , und man beweist:  $\alpha \circ \beta = \beta \circ \alpha \implies \exp(\alpha + \beta) = \exp(\alpha) \circ \exp(\beta)$ .

**Satz 2.7.1.** Sei  $\alpha$  ein selbstadjungierter Operator im Hilbertraum  $(V, \langle \cdot | \cdot \rangle)$ . Dann sind die Operatoren  $\varphi_t = \exp(-it\alpha)$  unitär für alle  $t \in \mathbb{R}$  und es gilt  $\varphi_s \circ \varphi_t = \varphi_{s+t}$ .

Man nennt in diesem Falle die Schar  $\{\varphi_t = \exp(-it\alpha) : t \in \mathbb{R}\}$  eine einparametrische Gruppe von Isometrien. Der Operator  $-i\alpha$  wird manchmal ihr infinitesimaler Generator genannt; man schreibt  $-i\alpha = \lim_{s \rightarrow 0} \frac{1}{s}(\varphi_s - id)$ . Im unendlichdimensionalen Fall muss man sich mit der Frage auseinandersetzen, in welchem Sinne der Limes verstanden werden soll. Im unendlichdimensionalen Fall gibt es auch einparametrische Gruppen von Isometrien, deren infinitesimaler Operator ein unbeschränkter mit  $-i$  multiplizierter selbstadjungierter Operator ist.

**Satz 2.7.2.** Ist  $A$  eine hermitesche Matrix, und  $V$  eine unitäre Matrix, sodass  $V^*AV = D$  die Diagonalmatrix ist mit den Diagonaleinträgen  $d_1, d_2, \dots$ , dann ist

$$V^* \exp(-itA)V = \exp(-itV^*AV)$$

die Diagonalmatrix mit den Einträgen  $e^{-itd_1}, e^{-itd_2}, \dots$

Zu jedem unitären  $\varphi$  existieren selbstadjungierte Operatoren  $\alpha$ , sodass  $\varphi = \exp(-i\alpha)$ .

**Beweis.** Die Behauptungen über Operatoren werden zu den entsprechenden Behauptungen über Matrizen, wenn wir die Operatoren in einer ONB darstellen.

Zum Beweis der letzten Behauptung wählen wir zur unitären Matrix  $U$  eine unitäre Matrix  $V$  so, dass  $VUV^*$  eine Diagonalmatrix  $B$  ist. Ihre Diagonaleinträge sind komplexe Zahlen mit dem Betrag 1, sagen wir  $e^{-id_1}, e^{-id_2}, \dots$ .  $B = \exp(-iD)$ . (Beachte, dass  $D$  nicht eindeutig bestimmt ist.) Die hermitesche Matrix  $A = V^*DV$  leistet das Verlangte; denn  $VUV^* = \exp(-iD) \implies U = V^* \exp(-iD)V = \exp(-iV^*DV) = \exp(-iA)$ .

### Beispiele im unendlichdimensionalen Fall

Es sei wie oben  $(\mathcal{T}, \langle \cdot | \cdot \rangle)$  der Prähilbertraum der trigonometrischen Polynome und  $(V, \langle \cdot | \cdot \rangle)$  seine Vervollständigung. Wir wollen die  $f(\cdot) \in \mathcal{T}$  punktweise mit der  $2\pi$ -periodischen Funktion  $h(\cdot)$  multiplizieren. Für genügend schöne Funktionen ist die Multiplikation mit  $h$  ein Endomorphismus  $M_h : V \longrightarrow V$ . Man muß sich hier nicht auf stetige

$h$  beschränken. Nur im einfachsten Fall, wo  $h$  selbst ein trigonometrisches Polynom ist,  $h(t) = \sum_k h_k \cdot e^{ikt}$  können wir das Bild  $M_h(\mathbf{v})$  für alle  $f(t) \longleftrightarrow \mathbf{v} \longleftrightarrow \sum c_n \cdot e^{int} \in \mathcal{T}$  in den beiden Darstellungsformen bequem hinschreiben.

$$h(t) \cdot f(t) \longleftrightarrow M_h(\mathbf{v}) \longleftrightarrow \mathbf{k} = (\dots, k_{-1}, k_0, k_1, k_2, \dots)$$

wobei sich die Folge  $\mathbf{k}$  als das Faltungsprodukt ergibt:  $\mathbf{k} = \mathbf{h} \star \mathbf{c}$ .

Der adjungierte Operator ist die Multiplikation mit  $\bar{h}(\cdot)$ ; das bedeutet im Spezialfall für die Folgen die Faltung mit der Koeffizientenfolge von  $\bar{h}(t) = \sum \bar{h}_{-k} e^{ikt}$ .

$$\langle g | M_h | f \rangle = \frac{1}{2\pi} \int \bar{g}(t) \cdot h(t) \cdot f(t) dt.$$

Der Operator  $M_h$  ist also genau dann selbstadjungiert, wenn  $h(\cdot)$  reellwertig ist,  $h_{-k} = \bar{h}_k$ . Wenn das der Fall ist, dann ist  $U_s = \exp(isM_h)$  ein unitärer Operator für alle  $s \in \mathbb{R}$ .  $U_s$  multipliziert  $f(\cdot)$  mit einer Funktion vom Betrag = 1. In der Tat  $U_s f = \exp(ish(\cdot))f(\cdot)$ . Für ‘gute’  $h$  erscheint  $U_s$  in der Reihendarstellung als Faltung der Koeffizienten. mit einer gewissen unendlichen Koeffizientenfolge  $(u_n(s))_n$ . In weniger guten Fällen, (wie etwa bei dem unten untersuchten Multiplikationsoperatoren  $M_e$  und  $\exp(isM_e)$ ) erweist sich die Faltung als eine unpassende Konstruktion.

Für die Operatoren  $M_h$  und  $U_s$  gibt es nur dann Eigenvektoren, wenn  $h$  Stellen der Konstanz aufweist. Wenn  $f \in V$  (im Intervall  $(-\pi, +\pi)$ ) ausserhalb einer kleinen Umgebung von  $t_0$  verschwindet, dann ist  $f$  ‘näherungsweise’ ein Eigenvektor  $M_h f \approx h(t_0) \cdot f$ ,  $U_s f \approx e^{ish(t_0)} \cdot f$ . Man kann die Vorstellung präzisieren, wenn man die Idee der Spektralzerlegung verallgemeinert zur Idee des Spektralintegrals.

An die Stelle von  $id = \sum \pi_\lambda$ ,  $\varphi = \sum \lambda \cdot \pi_\lambda$  tritt

$$id = \int_{-\pi}^{\pi} d\Pi_t, \quad M_h = \int_{-\pi}^{\pi} h(t) d\Pi_t, \quad U_s = \int_{-\pi}^{\pi} e^{ish(t)} d\Pi_t.$$

Die Integration bzgl. einer Schar von verträglichen orthogonalen Projektionen wird in der Theorie der Operatoren im Hilbertraum mathematisch exakt gemacht. Wir können hier die Idee nur ‘intuitiv’ benutzen.

Betrachten wir noch ein weiteres Beispiel einer einparametrischen Gruppe von Isometrien auf unserem Hilbertraum  $\mathcal{T}$ . Die Verschiebungen  $U_s(f)(\cdot) = f(\cdot - s)$  sogar den Prä-Hilbertraum der trigonometrischen Polynom isometrisch auf sich ab. Es gibt keinerlei Probleme mit den Eigenvektoren für die unitären Operatoren  $U_s$ . Die  $2\pi$ -periodische Funktion  $e_n(t) = e^{int}$  ist für  $U_s$  Eigenvektor zum Eigenwert  $e^{-ins}$ , und diese Eigenvektoren bilden eine Orthogonalbasis. Technische Probleme bereitet hier der zur unitären Gruppe gehörige ‘infinitesimale’ Operator. Er ist ein unbeschränkter Operator, nämlich der ‘wesentlich’ selbstadjungierte Operator  $A = \frac{1}{i} \frac{\partial}{\partial t}$ . Man schreibt symbolisch  $U_s = \exp(-isA)$ . Für jede stetig differenzierbare  $2\pi$ -periodische  $f$  gilt in der Tat  $\frac{1}{s}(U_s f - f) \rightarrow -f' = -iA(f)$  im Sinne der gleichmäßigen Konvergenz. (Für weniger reguläre Funktionen  $f$  im Definitionsbereich des Operators  $A$  muss man auf einen anderen Konvergenzbegriff rekurren.)

Die ‘wesentliche Selbstadjungiertheit’ von  $A$  wird dadurch begründet, dass sich für ‘gute’  $2\pi$ -periodische Funktionen  $f, g$  mit partieller Integration ergibt

$$\langle g|Af \rangle = \frac{1}{2\pi} \int \bar{g}(t) \cdot \frac{1}{i} \frac{\partial}{\partial t} f(t) dt = \frac{1}{2\pi} \int \left(-\frac{1}{i}\right) \frac{\partial}{\partial t} \bar{g}(t) \cdot f(t) dt = \langle Ag|f \rangle.$$

Jede ‘reine Sinusschwingung’  $e_n(\cdot) = e^{in(\cdot)}$  ist Eigenvektor für den Operator  $A$ , und zwar zum ganzzahligen Eigenwert  $n$ , und diese ‘reinen Sinusschwingungen’ bilden eine ONBasis. Eine Schwierigkeit ergibt sich hier nun aber daraus, dass die Menge der Eigenwerte (nach oben und nach unten) unbeschränkt ist.

Wir werden uns mit den Operatoren  $A$  und  $M_h$  weiter beschäftigen, nachdem wir uns etwas näher mit den infinitesimalen Operatoren im endlichdimensionalen Fall auseinandergesetzt haben.

### Matrizen als neutrale Rechengrößen.

Als es um Endomorphismen eines Hilbertraums  $(V, \langle \cdot | \cdot \rangle)$  ging, benützten wir auch (weitgehend synonym) das Wort Operator. Beim Begriff des Endomorphismus  $\varphi(\cdot)$ ,  $\psi(\cdot)$  denken wir (neben der Linearkombination) vor allem an das Hintereinanderschalten  $\varphi \circ \psi(\cdot)$ . Beim Begriff des Operatoren  $A, B$  denken wir eher an das Multiplizieren nach dem Muster der Matrizenmultiplikation  $A \cdot B$ . Die Entsprechung wird geleistet, wenn man im Vektorraum eine Basis auszeichnet, und die Koeffizienten der Vektoren als Spaltenvektoren notiert. Dieses aus der Anfängervorlesung bekannte Verfahren sollte nun aber nicht vergessen machen, dass die Matrizen nicht nur zur Darstellung von Endomorphismen des Spaltenraums gebraucht werden; gewisse Matrizen leisten insbesondere auch dann gute Dienste, wenn es um die Darstellung von quadratischen oder hermiteschen Formen geht,— und es gibt durchaus noch weitere wichtige Interpretationen.

Bei den folgenden Definitionen und Konstruktionen verstehen wir die Matrizen zunächst nur als Rechengrößen. Neben den bekannten Operationen der Addition und Multiplikation von Matrizen werden wir uns auch noch für das sog. Lie-Produkt zweier Matrizen interessieren. — Die gelegentlich Interpretation der Matrizen als Darstellungsform für Endomorphismen oder Formen kommt allenfalls als technische Hilfsüberlegung vor.

### Definition 2.15 (Die klassischen Gruppen).

1. Die allgemeine lineare Gruppe  $GL(n, \mathbb{K})$  (‘general linear group’) besteht aus den nichtsingulären  $n \times n$ -Matrizen mit Einträgen aus dem Körper  $\mathbb{K}$ .
2. Die spezielle lineare Gruppe  $SL(n, \mathbb{K})$  (‘special linear group’) besteht aus den  $n \times n$ -Matrizen über  $\mathbb{K}$  mit Determinante = 1.
3. Die orthogonale Gruppe  $O(n, \mathbb{K})$  besteht aus den Matrizen  $A$  mit  $A^T A = E_n$ . Die spezielle orthogonale Gruppe  $SO(n, \mathbb{K})$  ist die Untergruppe, die aus den orthogonalen Matrizen mit Determinante = 1 besteht.

4. Die unitäre Gruppe  $U(n) = U(n, \mathbb{C})$  ('unitary group') besteht aus den komplexen Matrizen  $U$  mit  $U^*U = E_n$ . Die spezielle unitäre Gruppe  $SU(n)$  ist die Untergruppe, die aus den unitären Matrizen mit Determinante = 1 besteht.

**Definition 2.16** (Das Lie-Produkt für Matrizen). Ein Vektorraum  $\mathcal{A}$  von  $n \times n$ -Matrizen heisst eine Matrix-Lie-Algebra, wenn gilt

$$A, B \in \mathcal{A} \implies [A, B] = A \cdot B - B \cdot A \in \mathcal{A}$$

Die Matrix  $[A, B]$  heisst der Kommutator von  $A$  und  $B$ .

Bemerke: Die Verknüpfung  $A, B \longrightarrow [A, B]$  ist bilinear. Es gilt  $[B, A] = -[A, B]$ . Die Verknüpfung ist nicht assoziativ; eine Art Ersatz für die Assoziativität ist

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (\text{'Jacobi's Identität'})$$

Den Nachweis der Identität überlassen wir dem Leser.

*Beispiel 2.7.1* (Schiefsymmetrische Matrizen). Die Menge aller schiefsymmetrischen Matrizen  $\mathcal{S} = \{A : A + A^T = 0\}$  ist eine Lie-Algebra; denn

$$A^T = -A, B^T = -B \implies [A, B]^T = [B^T, A^T] = [-B, -A] = -[A, B].$$

Diese Lie-Algebra wird (im Falle  $1 + 1 \neq 0$ ) mit  $\mathfrak{o}(n, \mathbb{K})$  bezeichnet. Es verschwinden die Diagonaleinträge, insbesondere also die Spur.  $\mathfrak{o}(n, \mathbb{K})$  ist ein  $\mathbb{K}$ -Vektorraum der Dimension  $\frac{1}{2}n(n-1)$ .

*Beispiel 2.7.2* (Schiefhermitesche Matrizen). Die Lie-Algebra  $\mathfrak{u}(n) = \mathfrak{u}(n, \mathbb{C})$  ist die Menge der schiefhermiteschen Matrizen, d. h.  $\mathfrak{u}(n) = \{A : A + A^* = 0\}$ . Wenn  $A$  und  $B$  hermitisch sind, dann ist der Kommutator schiefhermitisch.

Bemerke:  $A$  ist genau dann schiefhermitisch, wenn  $\frac{1}{i}A$  hermitisch ist, ( $a_{jk} = -\bar{a}_{kj}$ ). Die Menge der schiefhermiteschen Matrizen ist ein reeller Vektorraum der Dimension  $n^2$ ; sie ist kein komplexer Vektorraum. Jede komplexe  $n$ -Matrix besitzt eine eindeutige Darstellung als Summe einer hermiteschen und einer schiefhermiteschen Matrix:  $A = S + iT$  mit  $S = \frac{1}{2}(A + A^*)$ ,  $T = \frac{1}{2i}(A - A^*)$ .

**Satz 2.7.3.** Wenn  $B$  eine schiefsymmetrische reelle Matrix ist, dann ist  $\exp(tB)$  eine orthogonale Matrix für alle  $t \in \mathbb{R}$ .

Wenn  $B$  schiefhermitisch ist, dann ist  $\exp(tB)$  unitär für alle reellen  $t$ . Wenn  $H$  hermitisch ist, dann ist  $\{\exp(-itH) : t \in \mathbb{R}\}$  eine Gruppe unitärer Matrizen.

Wir wollen diese Konstruktionen noch etwas genauer studieren, nachdem wir uns zunächst über die Spur bzw. die Determinante der hier auftretenden Matrizen befasst haben. Wir bemerken, dass die Spur eines Lie-Produkts immer verschwindet.

**Lemma.** Für jede komplexe Matrix  $B$  gilt  $\det \exp(t \cdot B) = \exp(t \cdot \text{trace} B)$ .

**Beweis.**  $g(t) = \det \exp(t \cdot B)$  eine multiplikative Funktion:  $g(s+t) = g(s) \cdot g(t)$ . Wenn man sie für  $t \rightarrow 0$  betrachtet,  $g(t) = \det(E + tB + \dots)$ , dann sieht man  $g(t) = \exp(t \cdot \text{trace} B)$ .

Bemerkung: Wir haben oben mit Hilfe des Satzes von der unitären Diagonalisierbarkeit der hermiteschen Matrix  $H$  bewiesen, dass  $U_t = \exp(-itH)$  unitär ist für alle reellen  $t$ . Man kann das auch direkt sehen. Aus der Reihendarstellung sieht man  $U_{-t}^* = (\exp(itH))^* = \exp(-itH)$ . Andererseits  $U_{-t}U_t = E$  und das zeigt, dass  $U_t$  unitär ist. Wir formulieren das Ergebnis nochmals als

**Satz 2.7.4.** Die Exponentialabbildung bildet die Lie-Algebra  $u(n)$  surjektiv auf die unitäre Gruppe  $U(n)$  ab. Dabei wird die Lie-Algebra der spurlosen schiefhermiteschen Matrizen surjektiv auf die spezielle unitäre Gruppe  $SU(n)$  abgebildet.

Bemerkung: Wenn  $B$  schiefsymmetrisch ist, dann ist  $A_t = \exp(t \cdot B)$  eine spezielle orthogonale Matrix für alle reellen  $t$ . Es gilt nämlich

$$A_t^T = (e^{tB})^T = (E + tB + \frac{1}{2!}B^2 + \dots)^T = E - tB + \frac{1}{2!}B^2 + \dots = e^{-tB} = A_{-t} = A_t^{-1}.$$

also  $(e^{tB})^T \cdot e^{tB} = E$ . Da nun jede schiefsymmetrische reelle Matrix  $B$  die Spur 0 hat, hat  $\exp(t \cdot B)$  die Determinante = 1 für alle reellen  $t$ . Wir werden sehen, dass die Exponentialabbildung die Lie-Algebra  $o(n, \mathbb{R})$  surjektiv auf die spezielle orthogonale Gruppe  $SO(n, \mathbb{R})$  abbildet.

### Die gleichförmigen Drehungen des 3-dimensionalen euklidischen Raums

Die Elemente der Lie-Algebra  $o(3, \mathbb{R})$  sind die Matrizen der Gestalt

$$J_\omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} = \omega_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + \omega_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + \omega_3 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$J_\omega = \omega_1 \cdot J_x + \omega_2 \cdot J_y + \omega_3 \cdot J_z.$$

Man berechnet leicht die Kommutatoren

$$[J_x, J_y] = J_z, \quad [J_y, J_z] = J_x, \quad [J_z, J_x] = J_y.$$

Daraus ergibt sich wegen der Bilinearität des Lie-Produkts für beliebige Tripel  $\mathbf{a}, \mathbf{b}$

$$[J_{\mathbf{a}}, J_{\mathbf{b}}] = J_{\mathbf{c}} = (a_2 \cdot b_3 - a_3 \cdot b_2)J_x + (a_3 \cdot b_1 - a_1 \cdot b_3)J_y + (a_1 \cdot b_3 - a_3 \cdot b_1)J_z.$$

Hinweis: Die Formel für das Tripel  $\mathbf{c}$  dürfte manchen Leser an die Berechnung des sog. Kreuzprodukt (oder 'Vektorprodukt')  $\mathbf{a} \times \mathbf{b} = \mathbf{c}$  aus der Vektoranalysis erinnern. Die Formel ergibt sich dort aber aus Überlegungen, die nicht in den Matrizenkalkül passen.

Die Exponentialabbildung liefert zu jedem  $J_\omega$  eine einparametrische Gruppe orthogonaler Matrizen  $O_t = O_\omega(t) = \exp(t \cdot J_\omega)$ . Man nennt sie die Gruppe der gleichförmigen Drehungen mit der Winkelgeschwindigkeit  $\sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2}$  um die durch das Tripel  $\omega$  bestimmte Achse. Die Spalte  $\omega$  ist in der Tat Fixpunkt ist für die Abbildungen  $O_\omega(t) = \exp(t \cdot J_\omega)$ . Eine leichte Rechnung ergibt nämlich

$$(\omega_1 J_x + \omega_2 J_y + \omega_3 J_z) \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Eine einfache explizite Formel für die Einträge der Matrizen  $O_t$  findet man für die Drehungen um die Koordinatenachsen. Wir bemerken dazu; Wenn  $j = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix}$  dann gilt  $j^2 = -e$ ,  $j^3 = -j$ ,  $j^4 = e, \dots$  und daher

$$\begin{aligned} \exp(t \cdot j) &= e + tj + \frac{1}{2!}t^2(-e) + \frac{1}{3!}t^3(-j) + \frac{1}{4!}t^4(+e) + \dots \\ &= e \cdot (1 - \frac{1}{2!}t^2 + \frac{1}{4!}t^4 - \dots) + j \cdot (t - \frac{1}{3!}t^3 + \frac{1}{5!}t^5 + \dots) = e \cdot \cos t + j \cdot \sin t. \end{aligned}$$

Dies ergibt  $\exp(t \cdot J_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}$ . Wenn  $t$  ein ganzzahliges Vielfaches von  $2\pi$  ist, haben wir eine Volldrehung.

Die Gruppen  $\{O_\omega(t) : t \in \mathbb{R}\}$  mit dem gleichen Wert  $\sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2}$  sind zueinander konjugiert. Konkret sind die Matrizen  $J_x, J_y, J_z$  zueinander konjugiert vermittelt der zyklischen Permutationen

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{bzw.} \quad P^2 = P^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

$$P^{-1} \cdot J_x \cdot P = J_y, \quad P^{-1} \cdot J_y \cdot P = J_z, \quad P^{-1} \cdot J_z \cdot P = J_x.$$

Wenn  $\sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2} = 1$  und  $A$  eine spezielle orthogonale Matrix ist, die den Strahl durch  $\omega$  in die  $z$ -Achse überführt, dann gilt

$$\exp(t \cdot J_\omega) = A^{-1} \cdot \exp(t \cdot J_z) \cdot A.$$

Nach diesem Exkurs in die Welt der reellen Matrizen wenden wir uns wieder den unitären Matrizen zu. Wir werden sehen, dass die Gruppe  $SU(2)$  einerseits sehr ähnlich ist zur Gruppe  $SO(3, \mathbb{R})$ .

### Die Struktur von $u(2)$ und $SU(2)$

Die Lie-Algebra  $u(2)$  der schieferhermiteschen  $2 \times 2$ -Matrizen ist ein vierdimensionaler reeller Vektorraum. Die Matrix  $\frac{1}{i}e$  gehört dazu; sie hat die Spur  $\frac{2}{i}$ . Eine bequeme Basis des dreidimensionalen Vektorraums der spurlosen schieferhermiteschen ergibt sich aus bekannten Paulimatrizen  $\sigma_x, \sigma_y, \sigma_z$

$$\mathbf{j} = \frac{1}{i}\sigma_x = \frac{1}{i} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \mathbf{k} = \frac{1}{i}\sigma_y = \frac{1}{i} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \mathbf{l} = \frac{1}{i}\sigma_z = \frac{1}{i} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Wir bemerken für später:

1. Die Menge  $\{\pm e, \pm \mathbf{j}, \mathbf{k}, \mathbf{l}\}$  ist eine Gruppe. Ihre Multiplikationstabelle ist bestimmt durch

$$\mathbf{j}^2 = \mathbf{k}^2 = \mathbf{l}^2 = -e, \\ \mathbf{j} \cdot \mathbf{k} = \mathbf{l} = -\mathbf{k} \cdot \mathbf{j}; \quad \mathbf{k} \cdot \mathbf{l} = \mathbf{j} = -\mathbf{l} \cdot \mathbf{k}; \quad \mathbf{l} \cdot \mathbf{j} = \mathbf{k} = -\mathbf{j} \cdot \mathbf{l};$$

2. Man nennt die Matrizen  $\mathbf{j}, \mathbf{k}, \mathbf{l}$  auch die Standard-Einheitsquaternionen. Die Matrizen  $\frac{1}{2}\mathbf{j}, \frac{1}{2}\mathbf{k}, \frac{1}{2}\mathbf{l}$  haben dieselben Kommutatoren wie die schiefssymmetrischen  $3 \times 3$ -Matrizen  $J_x, J_y, J_z$ , die wir oben als Basis des Raums  $\mathfrak{o}(3, \mathbb{R})$  gewählt haben

$$\left[ \frac{1}{2}\mathbf{j}, \frac{1}{2}\mathbf{k} \right] = \frac{1}{2}\mathbf{l}, \quad \left[ \frac{1}{2}\mathbf{k}, \frac{1}{2}\mathbf{l} \right] = \frac{1}{2}\mathbf{j}, \quad \left[ \frac{1}{2}\mathbf{l}, \frac{1}{2}\mathbf{j} \right] = \frac{1}{2}\mathbf{k}.$$

3. Für die spurlose hermitische Matrix  $\mathbf{h} = b_x \cdot \sigma_x + b_y \cdot \sigma_y + b_z \cdot \sigma_z$  mit  $b_x^2 + b_y^2 + b_z^2 = 1$  gilt  $\mathbf{h}^2 = e$ ,  $\exp(-i\pi\mathbf{h}) = -e$  und für alle  $t \in \mathbb{R}$

$$\exp(-it\mathbf{h}) = e + (-it)\mathbf{h} + \frac{1}{2!}(-it)^2 e + \frac{1}{3!}(-it)^3 \mathbf{h} + \dots = \cos t \cdot e - i \sin t \cdot \mathbf{h}.$$

4. Insbesondere gilt

$$\exp\left(\frac{1}{2}i\psi\sigma_z\right) = \begin{pmatrix} e^{i\frac{\psi}{2}} & 0 \\ 0 & e^{-i\frac{\psi}{2}} \end{pmatrix}, \quad \exp\left(\frac{1}{2}i\theta\sigma_x\right) = \begin{pmatrix} \cos \frac{\theta}{2} & i \sin \frac{\theta}{2} \\ i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix}.$$

Die Elemente der speziellen unitären Gruppe  $SU(2)$  sind die Matrizen

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 + |b|^2 = 1.$$

Wenn man die Real- und Imaginärteile der Zahlen  $a$  und  $b$  als Koordinaten im vierdimensionalen euklidischen Raum auffasst, dann erscheint die Menge  $SU(2)$  als die Einheitskugel:  $(\Re a)^2 + (\Im a)^2 + (\Re b)^2 + (\Im b)^2 = 1$ . Wir haben oben den nichtsingulären  $2 \times 2$ -Matrizen Transformationen der Riemann'schen Zahlenkugel zugeordnet. Die Möbiustransformationen zu den (speziellen) unitären  $U$  bilden Punkte, die einander gegenüberliegen in gegenüberliegende Punkte ab. Diese Transformationen kann man als Drehungen der Riemann'schen Zahlenkugel verstehen. Die Drehungen um die  $z$ -Achse beispielsweise entsprechen den Matrizen  $\exp\left(i\frac{1}{2}\phi \cdot \sigma_z\right)$ . Im stereographischen Bild der Transformationen von  $\bar{\mathbb{C}}$  sind das nämlich die Abbildungen

$$\bar{\mathbb{C}} \ni z \longrightarrow \frac{e^{i\phi/2} \cdot z + 0}{0 + e^{-i\phi/2}} = e^{i\phi} \cdot z.$$

Entsprechendes gilt für die Wirkung der Matrizen  $\exp\left(i\frac{1}{2}\phi \cdot \mathbf{h}\right)$ , wo  $\mathbf{h}$  eine spurlose hermitische Matrix mit  $\det \mathbf{h} = -1$  ist:

$$\mathbf{h} = b_x \cdot \sigma_x + b_y \cdot \sigma_y + b_z \cdot \sigma_z; \quad \det \begin{pmatrix} b_z & b_x - ib_y \\ b_x + ib_y & -b_z \end{pmatrix} = -b_z^2 - b_x^2 - b_y^2 = -1.$$

Man die Drehungen aus Drehungen um Achsen zusammensetzen, wobei man sogar (aufgrund der Nichtkommutativität) auf die Drehungen um die  $y$ -Achse verzichten kann.

**Satz 2.7.5** (Euler'sche Winkel).

Jedes  $U \in SU(2)$  besitzt eine Produktdarstellung

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} = \exp\left(\frac{1}{2}i\phi\sigma_z\right) \cdot \exp\left(\frac{1}{2}i\theta\sigma_x\right) \cdot \exp\left(\frac{1}{2}i\psi\sigma_z\right).$$

wobei (im Falle  $b \neq 0$ )  $\phi \in [0, 2\pi)$ ,  $\theta \in (0, \pi)$ ,  $\psi \in [-2\pi, 2\pi)$ .

**Beweis.** Das Produkt ist die Matrix

$$\begin{pmatrix} \cos \frac{\theta}{2} \cdot e^{i\alpha} & i \sin \frac{\theta}{2} \cdot e^{i\beta} \\ i \sin \frac{\theta}{2} \cdot e^{-i\beta} & \cos \frac{\theta}{2} \cdot e^{-i\alpha} \end{pmatrix}$$

mit  $\alpha = \frac{1}{2}(\phi + \psi)$ ,  $\beta = \frac{1}{2}(\phi - \psi)$ . Wir finden also  $a$  und  $\alpha$  aus  $a = \cos \frac{\theta}{2} \cdot e^{i\alpha}$ ; schliesslich ergibt sich  $\beta$  aus dem Argument von  $b$ .

Die Gruppe  $SU(2)$  ist (nach dem Urteil der meisten Mathematiker) einfacher zu verstehen als die Gruppe  $SO(3, \mathbb{R})$ , obwohl wir doch bekanntlich nicht in einem 2-dimensionalen Hilbertraum leben, sondern in einem dreidimensionalen euklidischen Raum. Wir zitieren dazu R. Feynman: (Lecture Notes Kapitel 6. Spin 1/2): *Die Tatsachen der Verknüpfungen von Drehungen und was daraus entsteht, sind anschaulich schwierig zu erfassen. Es ist einigermaßen seltsam, wo wir doch in drei Dimensionen leben, aber es ist schwer für uns einzusehen, was geschieht, wenn wir uns einmal so herum und dann so herum drehen. Wenn wir Fische oder Vögel wären und ein wirkliches Verständnis von dem hätten, was geschieht, wenn wir im Raum Purzelbäume schlagen, könnten wir diese Dinge vielleicht leichter erfassen.*

**Anhang: Hamiltons Quaternionen** Eine bemerkenswerte algebraische Struktur hat die Gesamtheit der reellen Linearkombinationen der Matrizen  $e, \mathbf{j}, \mathbf{k}, \mathbf{l}$ . Man nennt sie den Schiefkörper  $\mathbb{H}$  der Quaternionen (in der Standarddarstellung).

Hamilton (1805 -1865) führte die Quaternionen als Zahlenquadrupel ein, nach demselben Schema, wie man die komplexen Zahlen als Zahlenpaare konstruiert. In leicht modernisierter Sprache definiert man die Quaternionen als formale Ausdrücke

$$a_0 \cdot e + a_1 \cdot \mathbf{j} + a_2 \cdot \mathbf{k} + a_3 \cdot \mathbf{l} \quad \text{mit} \quad a_0, a_1, a_2, a_3 \in \mathbb{R}.$$

Quaternionen werden komponentenweise addiert und distributiv ausmultipliziert ausgehend von den oben genannten Multiplikationsregeln. Wie für die komplexen Zahlen definiert man zu einem Quaternion  $\mathbf{q}$  das konjugierte Quaternion  $\mathbf{q}^*$  und die Norm  $\|\mathbf{q}\|$ :

$$\begin{aligned} \mathbf{q}^* &= a_0 \cdot e - a_1 \cdot \mathbf{j} - a_2 \cdot \mathbf{k} - a_3 \cdot \mathbf{l}. \\ \|\mathbf{q}\| &= \sqrt{\mathbf{q} \cdot \mathbf{q}^*} = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2} \end{aligned}$$

Man kann direkt nachrechnen, dass das System der Quaternionen  $(\mathbb{H}, +, \cdot, *)$  eine assoziative  $\mathbb{R}$ -Algebra ist mit  $(\mathbf{q} \cdot \mathbf{r})^* = \mathbf{r}^* \cdot \mathbf{q}^*$ . Man kann sich die Arbeit aber sparen, wenn man sich überzeugt hat, dass man die Quaternionen durch komplexe  $2 \times 2$ -Matrizen darstellen kann, wo die Addition und die Multiplikation die Verknüpfungen des Matrizenkalküls sind und die Konjugation durch die hermitesche Konjugation gegeben ist. Eine naheliegende Übersetzung ist die Übersetzung der ‘Basisquaternionen’ in die Pauli-Matrizen, die wir oben beschrieben haben:  $\mathbf{j} \longleftrightarrow \frac{1}{i}\sigma_x$ ,  $\mathbf{k} \longleftrightarrow \frac{1}{i}\sigma_y$ ,  $\mathbf{l} \longleftrightarrow \frac{1}{i}\sigma_z$ .

Das System der Quaternionen hat die bemerkenswerte Eigenschaft, dass jedes Element ausser dem Nullelement eine Inverse besitzt, nämlich  $\mathbf{q}^{-1} = \frac{1}{\|\mathbf{q}\|^2} \mathbf{q}^*$ . Die Quaternionen bilden einen sog. Schiefkörper.

Ein Quaternion mit der Norm 1 nennt man ein Einheitsquaternion. Bei unserer Darstellung im Matrizenkalkül entsprechen die Einheitsquaternionen den unitären Matrizen mit Determinante 1. Ein Quaternion  $\mathbf{m}$  nennt ein Reines Quaternion, wenn  $\mathbf{m}^* = -\mathbf{m}$ . Die reinen Quaternionen  $\mathbf{m}$  entsprechen also den spurlosen schiefhermiteschen Matrizen. Wir haben uns damit im Abschnitt zum Minkowski-Raum beschäftigt. Das Thema dieses Kapitels soll weiterhin die Geometrie des Hilbertraums sein.

## 2.8 Selbstadjungierte Operatoren als Observable.

In der Quantenmechanik beschreibt man die Zustände eines Systems durch Vektoren in einem Hilbertraum ('ket-Vektoren' in der Sprechweise von Dirac). Bei der 'Überlagerung' (Superposition) von Zuständen gelten die Regeln des Hilbertraums. Vektoren beschreiben genau dann denselben Zustand, wenn sie sich nur um einen Faktor unterscheiden; es ist daher manchmal angemessen, nur Einheitsvektoren zur Beschreibung zu heranzuziehen; der Einheitsvektor zu einem Zustand ist bis auf einen Faktor vom Betrag 1 bestimmt.

Meistens erscheint der Hilbertraum als ein unendlichdimensionaler Raum komplexwertiger Funktionen (oder als ein Raum von Äquivalenzklassen komplexwertiger Funktionen) auf einem Teilbereich eines  $\mathbb{R}^d$ . Für den Einstieg in die Gedankenwelt der Zustände und Observablen kann man aber einige bedeutsame Gedanken schon im endlichdimensionalen Fall explizieren. Der endlichdimensionale Fall kann durchaus physikalische Realitäten beschreiben; in den Feynman Lectures werden allerlei Zweizustandssysteme ausführlich diskutiert.

Wenn man in einem endlichdimensionalen Hilbertraum eine ONB  $\{\mathbf{u}_j : j \in J\}$  gewählt hat, dann erscheinen die Elemente  $\mathbf{v}$  zunächst einmal als  $J$ -Spalten; diese kann man dann natürlich auch als komplexwertige Funktionen auf der Indexmenge  $J$  verstehen.  $f_{\mathbf{v}} = \{f_{\mathbf{v}}(j) : j \in J\} = \{\langle \mathbf{u}_j | \mathbf{v} \rangle : j \in J\}$ . Funktionen, die sich nur um einen Faktor unterscheiden, beschreiben denselben Zustand.

(Hinweis: Häufig ergibt sich die Wahl der Basis durch die Auszeichnung eines selbstadjungierten Operators, dessen Eigenvektoren man dann zur Basis macht. Wenn der Operator nur einfache Eigenwerte hat, dann sind die Basisvektoren bis auf Faktoren vom Betrag 1 eindeutig bestimmt. Es ist in diesem Falle naheliegend, das Spektrum  $Sp$  als Indexmenge der Basis zu betrachten; die ket-Vektoren sind dann also komplexwertige Funktionen auf der Teilmenge  $Sp$  der reellen Achse. Im Hinblick auf Verallgemeinerungen versteht man diese Funktionen auf  $Sp$  manchmal auch als Äquivalenzklassen von Funktionen auf  $\mathbb{R}$ , wobei Funktionen als äquivalent gelten, wenn sie auf  $Sp$  übereinstimmen. Diese Art von Repräsentation der ket-Vektoren durch Äquivalenzklassen von Funktionen auf der reellen Achse eignet sich auch für den Fall, wo der ausgezeichnete Operator ein kontinuierliches Spektrum hat. Wir werden Beispiele kennenlernen.)

Die zeitliche Veränderung des Zustands wird in der Quantenmechanik durch eine einparametrische Gruppe unitärer Transformationen beschrieben, gemäß der Schrödinger-Gleichung. Bei  $n$ -Zustandssystemen hat diese im zeitlich stationären Fall die Gestalt

$$i\hbar \frac{d}{dt} f(j) = \sum_j H_{ij} \cdot f(j); \quad i\hbar \frac{d}{dt} \langle \mathbf{u}_j | \mathbf{v} \rangle = \sum_j \langle \mathbf{u}_j | \hat{H} | \mathbf{u}_j \rangle \cdot \langle \mathbf{u}_j | \mathbf{v} \rangle,$$

wo  $H$  die Matrix zum Hamiltonoperator  $\hat{H}$  ist. (In den Feynman Lecture Notes Bd.III Glchg (16.49)) Der Zustandsvektor  $\mathbf{v}$ , der (in den jeweiligen ONSystemen) auch die komplexe Amplitude heisst, verändert sich in der Zeit gemäß  $\mathbf{v}_t = \exp(-it\frac{1}{\hbar}\hat{H}) \cdot \mathbf{v}_0$ .

Im nichtstationären Fall tritt eine sog. verallgemeinerte Gruppe unitärer Operatoren an die Stelle der einparametrischen Gruppe  $\{U_t : U_t = \exp(-it\hat{H}) : t \in \mathbb{R}\}$ . Dabei handelt es sich um eine zweiparametrische Familie unitärer Operatoren  $\{U(t_1, t_2) : t_1, t_2 \in \mathbb{R}\}$  mit  $U(t_3, t_1) = U(t_3, t_2) \cdot U(t_2, t_1)$ . Diese ‘verallgemeinerte Gruppe’ ist bestimmt durch eine Schar selbstadjungierter Operatoren  $\{H_t : t \in \mathbb{R}\}$  vermöge

$$\lim_{s \rightarrow 0} \frac{1}{s} (U(t+s, t) - U(t, t)) = -\frac{i}{\hbar} H_t. \quad \text{für alle } t$$

Ein Zustandsvektor verändert sich in der Zeit gemäß  $\mathbf{v}_{t_2} = U(t_2, t_1)\mathbf{v}_1$  gemäß der infinitesimalen Veränderung  $\lim_{s \rightarrow 0} \frac{1}{s} (\mathbf{v}_{t+s} - \mathbf{v}_t) = -i\frac{1}{\hbar} H_t \cdot \mathbf{v}_t$ .

Die Planck-Konstante  $\hbar$  muss uns als Mathematiker nicht interessieren. Wir dürfen auch annehmen, dass  $\mathbf{v}_0$  und damit alle  $\mathbf{v}_t$  Einheitsvektoren sind.

Wenn man im stationären Fall eine ONB einführt, die aus Eigenvektoren  $\mathbf{u}_j$  zum Operator  $\hat{H}$  besteht, dann wird der Hamiltonoperator  $H$  durch reelle Diagonalmatrix dargestellt. Die Diagonaleinträge  $d_j$  sind zu interpretieren als die möglichen Energiewerte des Systems. Die  $\mathbf{u}_j$  stellen die zeitlich invarianten Zustände zu diesen Energiewerten dar. Wenn  $\mathbf{v}_0 = \sum \mathbf{u}_j \cdot \beta_j$  der Anfangszustand des Systems ist, dann ist  $\mathbf{v}_t = \sum \mathbf{u}_j \cdot e^{itd_j} \beta_j$  der Zustand zur Zeit  $t$ . Die Wahrscheinlichkeit, dass die Messung der Energie den Wert  $d_j$  erbringt, ist  $p_j = |\langle \mathbf{u}_j | \mathbf{v}_t \rangle|^2 = |\beta_j|^2$ .

**Observable:** Der Zustand eines quantenmechanischen Systems kann nicht in allen Aspekten simultan beobachtet werden. Für die Beobachtung einer physikalischen Größe muss das System nämlich mit einem entsprechenden Mechanismus wechselwirken, und der Zustand wird durch die Wechselwirkung gestört, sodass weitere Beobachtungen nicht mehr den ursprünglichen Zustand des Systems betreffen. Man braucht einen ‘Ofen’, der immer wieder denselben Zustand produziert. Man kann den Zustand in mehrfachen Ansätzen in allen Aspekten beobachten und dann die zufälligen (!) Ergebnisse statistisch auswerten.

Die Beobachtung eines Aspekts ist durch einen selbstadjungierten Operator (genannt ‘Observable’) gekennzeichnet. Seine Eigenwerte sind die möglichen Werte, die die betreffende physikalische Größe annehmen kann. Der Zustand  $\mathbf{v}$  des Systems bestimmt, mit welchen Wahrscheinlichkeiten diese Werte beobachtet werden. Ist  $A = \sum_{\lambda} \lambda \cdot \pi_{\lambda} = \sum_{\lambda} \lambda \cdot |\mathbf{u}_{\lambda}\rangle\langle \mathbf{u}_{\lambda}|$  die Spektralzerlegung der Observablen und  $\mathbf{v} = \sum_{\lambda} \mathbf{u}_{\lambda} \cdot \langle \mathbf{u}_{\lambda} | \mathbf{v} \rangle$  der entsprechend zerlegte Zustand ( $\|\mathbf{v}\|^2 = 1$ ), dann ist  $p_{\lambda} = |\langle \mathbf{u}_{\lambda} | \mathbf{v} \rangle|^2$  die Wahrscheinlichkeit, dass die Beobachtung den Wert  $\lambda$  liefert. Die reelle Zahl  $\mathfrak{E}_{\mathbf{v}} A = \sum_{\lambda} \lambda \cdot p_{\lambda} = \langle \mathbf{v} | A | \mathbf{v} \rangle$  ist der Erwartungswert der Observablen  $A$  im Zustand  $\mathbf{v}$ . Nur in dem Falle, wo  $\mathbf{v}$  ein Eigenvektor des Operators  $A$  ist, wird mit Sicherheit ein bestimmter Wert  $\lambda$  beobachtet. Andernfalls hat die Beobachtung positive Varianz. Es ist aber zu sagen, dass für die Theorie die von einem Zustand des Systems abgeleitete Verteilung auf der Menge der möglichen Werte der Observablen  $A$  nicht dieselbe Rolle spielt wie die Verteilung einer Zufallsgröße in der klassischen (‘kommutativen’) Stochastik.

In der nichtkommutativen Theorie der Observablen bestimmen die Zustände des Systems (nicht direkt beobachtbare!) sog. Wahrscheinlichkeitsamplituden, die auf die Wahl einer ONB  $\{\mathbf{u}_j : j \in J\}$  bezogen sind. Das sind komplexwertige Größen, aus deren Absolutquadrat sich die (statistisch beobachtbare!) Wahrscheinlichkeitsverteilung gewisser Observabler ergibt.

Für die mathematische Theorie wird ein Zustand  $\mathbf{v}$  vollständig beschrieben durch einen zerlegten Einheitsvektor  $\mathbf{v} = \sum_{j \in J} \mathbf{u}_j \cdot \langle \mathbf{u}_j | \mathbf{v} \rangle$ ,  $\|\mathbf{v}\|^2 = 1$ . Die Familie komplexer Zahlen  $\{a_j : a_j = \langle \mathbf{u}_j | \mathbf{v} \rangle : j \in J\}$ , als eine komplexwertige Funktion  $a(\cdot)$  auf der Indexmenge  $J$  verstanden, heisst die ‘Amplitudenfunktion’ des Zustands bzgl. der gewählten Basis. Diese ‘Funktion’  $|a(\cdot)|^2$  ergibt eine Wahrscheinlichkeitsgewichtung  $\{p_j : j \in J\}$ , aus der sich die Verteilung derjenigen Observablen  $A$  ablesen lässt, die mit der orthogonalen Zerlegung  $id = \sum_{j \in J} |\mathbf{u}_j\rangle\langle \mathbf{u}_j|$  ‘verträglich’ sind. Das sind diejenigen  $A$ , für welche alle  $\mathbf{u}_j$  Eigenvektoren sind. Wenn  $A$  ein verträglicher selbstadjungierter Operator ist mit  $A\mathbf{u}_j = \lambda_j \cdot \mathbf{u}_j$  für alle  $j$ , dann kann man die Wirkung des Operators  $A$  als die Wirkung eines Multiplikationsoperators auf der Amplitudenfunktion verstehen.  $A \rightsquigarrow M_\lambda : a(\cdot) \mapsto \lambda(\cdot) \cdot a(\cdot)$ . Es gilt sodann

$$\begin{aligned} \text{Ws}_{\mathbf{v}}(A \text{ hat den Wert } \lambda) &= \sum_{\{j:\lambda_j=\lambda\}} p_j = \sum_{\{j:\lambda_j=\lambda\}} |a_{\mathbf{v}}(j)|^2 \\ \mathfrak{E}_{\mathbf{v}}A &= \sum_j \lambda_j \cdot p_j = \langle \mathbf{v} | A | \mathbf{v} \rangle = \tilde{\lambda} \\ \text{var}_{\mathbf{v}}A &= \sum_j \lambda_j^2 p_j - \tilde{\lambda}^2 = \sum_j (\lambda_j - \tilde{\lambda})^2 p_j = \langle \mathbf{v} | (A - \tilde{\lambda} \cdot id)^2 | \mathbf{v} \rangle \end{aligned}$$

Bemerkung: Es ist nicht wirklich wichtig, dass der Hilbertraum  $V$  in paarweise orthogonale eindimensionale Teilräume  $V_j = \Pi_j(V)$  zerlegt wird. Es genügt, dass die orthogonale Zerlegung so fein ist, dass die zu untersuchende Observable  $A$  auf jedem  $V_j$  ein Vielfaches der Identität ist.  $A \circ \pi_j = \lambda_j \cdot \pi_j$ . Entsprechendes gilt, wenn (im unendlichdimensionalen Fall) für manche Observable die orthogonale Summe  $id = \sum_j \pi_j$  durch ein Integral zu ersetzen ist, z. B.  $id = \int d\Pi_t$ . Die orthogonale Zerlegung muss hinreichend fein für  $A$  sein.

**‘Unbestimmtheit’:** In der klassischen ‘kommutativen’ Wahrscheinlichkeitstheorie konstruiert man zum Begriff der Varianz (einer reellwertigen Zufallsgrößen) den Begriff der Kovarianz zweier Zufallsgrößen. Die Konstruktion benötigt das Produkt der beiden Zufallsgrößen. Eine entsprechende Konstruktion funktioniert nicht bei den Observablen in der Quantenmechanik, weil Produkte  $AB$  i. Allg. nicht selbstadjungiert sind. Hier ist folgendes festzustellen: Wenn  $A$  und  $B$  selbstadjungierte Operatoren sind, dann ist auch der ‘Kommutator’  $\frac{1}{i}(AB - BA)$  selbstadjungiert. Es stellt sich heraus, dass in dem Falle, dass sein Erwartungswert im Zustand  $\mathbf{v}$  einen Betrag  $\neq 0$  hat, die beiden Observable nicht simultan ‘scharf konzentrierte’ Werte haben können.

**Satz 2.8.1** (Heisenbergs Unschärferelation). *Seien  $A$  und  $B$  selbstadjungierte Operatoren und  $C = \frac{1}{i}(AB - BA)$ . Es gilt dann für alle Zustände  $\mathbf{v}$*   $\sqrt{\text{var}_{\mathbf{v}}A} \cdot \sqrt{\text{var}_{\mathbf{v}}B} \geq \frac{1}{2} \cdot |\mathfrak{E}_{\mathbf{v}}C|$ .

**Beweis.** *Wir können uns auf den Fall  $\mathfrak{E}_{\mathbf{v}}A = \mathfrak{E}_{\mathbf{v}}B = 0$  beschränken. Für jedes reelle  $t$  ist  $(A + tiB)^*(A + tiB)$  ein positiv semidefiniter Operator. Er hat einen nichtnegativen Erwartungswert in jedem Zustand  $\mathbf{v}$ .*

$$\begin{aligned} 0 \leq \mathfrak{E}_{\mathbf{v}}((A + tiB)^*(A + tiB)) &= \mathfrak{E}_{\mathbf{v}}(A^*A) + t^2 \cdot \mathfrak{E}_{\mathbf{v}}(B^*B) + ti \cdot \mathfrak{E}_{\mathbf{v}}((A^*B - B^*A)) \\ &= \text{var}_{\mathbf{v}}A + t^2 \cdot \text{var}_{\mathbf{v}}B - t \cdot \mathfrak{E}_{\mathbf{v}}C \end{aligned}$$

*Eine quadratische Funktion  $a + t^2b - tc$  ist nur dann nichtnegativ, wenn  $c^2 \leq 4ab$ .*

### Zur Vorstellungswelt der Quantenmechanik

Die Pioniere der Quantenmechanik haben sich tiefgründige Gedanken zum Begriff der Observablen gemacht. Man bemüht sich nunmehr seit fast hundert Jahren darum, die intuitiven Vorstellungen zusammenzubringen mit den Möglichkeiten der technischen Bearbeitung. Nicht nur die Physiker sondern auch die Mathematiker sind gefordert. Und es hat sich gezeigt, dass die Geometrie der Hilberträume ein entscheidender Ansatzpunkt ist. Für ein quantenmechanisches System mit einem unendlichdimensionalen Zustandsraum wird geometrisches Verständnis etwas anderer Art gefordert als man das in den endlichdimensionalen Geometrien gewöhnt ist — von der euklidischen, relativistischen, nichteuklidischen, symplektischen Geometrie bis zur Differentialgeometrie.

Die Notwendigkeit einer neuartigen Geometrie zeigte sich vor etwa 100 Jahren. (Manche sprechen auch von der Notwendigkeit einer neuen Logik.) Die Ansammlung von Informationen über Phänomene im atomaren und mikroskopischen Bereich während des ersten Viertels des 20. Jahrhunderts hatte eine wachsende Verwirrung hervorgerufen, die schliesslich durch die Formulierung der Prinzipien der Quantenmechanik in Form gebracht werden konnte. Fundamental waren die Idee der Superposition von Zuständen, (die zu partieller Auslöschung führen kann), und die Idee der Unbestimmtheit. Sie erschienen höchst sonderbar; sie trugen aber Wesentliches bei zu einem (mehr oder weniger) intuitiven Verständnis für so fundamentale Fragen wie die nach der ‘inneren Struktur’ der Photonen. Es ist aber noch immer so, dass die Anschauung hinterherhinken hinter dem technischen Verständnis. In den Einschätzungen beziehe ich mich hauptsächlich auf Feynman’s Lecture Notes und auf das klassische Buch ‘The Principles of Quantum Mechanics’ von P. Dirac. Nun kann man sagen, dass es nicht das hauptsächliche Anliegen der wissenschaftlichen Physik sei, einsichtige Bilder bereitzustellen; es gehe vielmehr darum, die Gesetze zu formulieren, welche die Phänomene beherrschen und diese Gesetze für die Entdeckung neuer Phänomene zu nutzen. Wenn ein Bild existiert, dann ist das umso besser. Aber bei den Phänomenen der Atome kann man wohl kein Bild im üblichen Sinn des Wortes ‘Bild’ erwarten. Man darf sich von einem Bild kein Modell erhoffen, welches im Wesentlichen entlang der traditionellen Denkweisen funktioniert. Man sollte, so schreibt Dirac, die Bedeutung des Wortes ‘Bild’ dahingehend ausweiten, dass damit jede

Betrachtungsweise der fundamentalen Gesetze erfasst ist, welche ihre innere Konsistenz offensichtlich erscheinen lässt. In diesem Sinne sollte es möglich sein, meint Dirac, im Laufe der Zeit ein Bild von den atomaren Phänomenen zu erwerben, indem man sich mit den Gesetzen der Quantenmechanik vertraut macht. In Feynman's Lecture Notes (Band III.1) lesen wir: *Weil das Verhalten der Atome so ganz ausserhalb unserer normalen Erfahrung liegt, ist es sehr schwierig, sich daran zu gewöhnen, und es erscheint beiden, dem Neuling wie dem erfahrenen Physiker, seltsam und geheimnisvoll. Selbst die Experten verstehen es nicht so, wie sie es gerne möchten, und das ist ganz klar, denn jede menschliche Erfahrung bezieht sich auf große Objekte. Wir wissen, wie sich große Objekte verhalten werden, aber die kleinen verhalten sich nicht so. Darum müssen wir unsere Erfahrungen durch eine Art von Abstraktion oder Imagination sammeln und nicht durch Anschluss an unsere direkten Erfahrungen.*

Beim Begriff der Observablen nimmt Dirac als intuitiven Ausgangspunkt die Vorstellung von der 'reellwertigen dynamischen Variablen'. Er sagt dann aber, dass offenbar nicht alle 'dynamischen Variablen' ausreichen, um ein vollständiges System zu bilden. Solche 'dynamischen Variablen' können nicht gemessen werden. Er kommt daher zu der Festlegung, dass nur diejenigen dynamischen Variablen als Observable gelten können, wenn sie ein vollständiges System von Eigenvektoren besitzen. Damit das funktioniert, muss Dirac den Begriff des Eigenvektors weiter fassen als die Funktionalanalysis. (Auch den Begriff der Funktion und die Technik der Integration musste Dirac erweitern, um die Orthogonalität von 'Eigenvektoren', die durch Funktionen gegeben sind auszudrücken.) Die Mathematiker haben die meisten dieser von Dirac erwünschten Begriffe auf sichere Grundlagen gestellt. In einer Anfängerveranstaltung kann aber bestenfalls angedeutet werden, welche Konstruktionen und Vorgehensweisen sich als tragfähig erwiesen haben.

Im endlichdimensionalen Fall gibt es mehrere zueinander äquivalente Definitionen einer Observablen zu einem Hilbertraum  $V$ . Man definiert z. B.: Eine Observable ist ein selbstadjungierter Endomorphismus. Man denkt da also an eine Abbildung. Man kann die Observable auch als eine hermitesche Form  $\langle \mathbf{v} | A | \mathbf{v} \rangle$ . In beiden Vorstellungen kommen die Vektoren (oder Zustände des Systems) ins Bild. Man kann die Erwähnung der Argumente  $\mathbf{v}$  vermeiden, wenn man das folgende Muster der Definition einer reellwertigen Funktion  $f$  (auf einer Grundmenge  $S$ ) variiert. Eine Funktion  $f$  ist bekanntlich dadurch bestimmt, dass jedem Intervall  $(a, b] \subseteq \mathbb{R}$  eine Teilmenge  $f^{-1}((a, b]) \subseteq S$  zugeordnet ist. (Diese Teilmenge heisst das volle Urbild des Intervalls.) Andererseits beweist man leicht:

**Lemma 2.8.1.** *Wenn  $P(\cdot)$  jedem Intervall  $(-\infty, t]$  eine Teilmenge von  $S$  zuordnet, sodass*

1.  $s \leq t \implies P(s) \subseteq P(t)$ ,
2.  $\bigcap_{\varepsilon > 0} P(t + \varepsilon) = P(t)$ ,
3.  $\bigcap_t P(t) = \emptyset; \quad \bigcup_t P(t) = S$ .

*dann existiert genau eine Funktion  $f(\cdot)$  auf  $S$  sodass  $f^{-1}((a, b]) = P(b) \setminus P(a)$ .*

An diese Bemerkung knüpft die folgende Definition an:

**Definition 2.17.** Eine Spektralschar auf einem Hilbertraum ist eine Familie orthogonaler Projektionen  $\{\Pi_t : t \in \mathbb{R}\}$  mit den Eigenschaften

1.  $\Pi_s \circ \Pi_t = \Pi_{\min\{s,t\}},$
2.  $\lim_{\varepsilon \searrow 0} \Pi_{t+\varepsilon} = \Pi_t,$
3.  $\bigcap_t \Pi_t = 0; \quad \bigcup_t \Pi_t = id.$

Der Satz von der unitären Diagonalisierbarkeit eines selbstadjungierten Operators wird im unendlichdimensionalen Fall zum Spektralsatz, welcher (untechnisch ausgedrückt) besagt, dass für jeden ‘gutartigen’ symmetrischen Operator  $A$  ( $\langle \mathbf{w} | A \mathbf{v} \rangle = \langle A \mathbf{w} | \mathbf{v} \rangle$ ) genau eine Spektralschar existiert, sodass gilt  $\int t d\Pi_t = A$ . Wenn  $h(\cdot)$  eine ‘gute’ (etwa beschränkte) Funktion ist, dann gilt  $h(A) = \int h(t) d\Pi_t$ .— Was das alle bedeutet, werden wir im Beispiel unten sehen.

Diese in der Funktionalanalysis präzisierte Aussage gibt eine (nach allgemeiner Meinung gelungene) Auskunft auf Dirac’s Frage, welche mathematischen Anforderungen an eine Observable zu stellen sind. Es geht also nicht um die Existenz von Eigenvektoren im strengen Sinn. Die Formulierung mit der Spektralschar hat nicht nur im unendlichdimensionalen Fall einen Vorteil gegenüber der Formulierung im Stil von Dirac, welche die Existenz einer aus Eigenvektoren bestehenden ONB fordert. In der Wahl einer ONB steckt nämlich eine unerfreuliche Willkür; selbst in dem Fall, wo alle Eigenwerte einfach sind, kann man die Elemente der ONB noch mit Faktoren vom Betrag 1 multiplizieren, und das ändert dann die ‘Amplitudenfunktionen’ der Zustände um einen nichtkonstanten Faktor.

Wir kommen nochmals zurück auf die zeitliche Entwicklung eines quantenmechanischen Systems. Das oben vorgestellte Bild heisst das Bild von Schrödinger: jeder einzelne ket-vektor  $|\mathbf{v}\rangle$  wandelt sich in der Zeitspanne von  $t_1$  bis  $t_2$  in den ket-Vektor  $U(t_2, t_1)|\mathbf{v}\rangle$ , wo  $\{U(t_1, t_2) : t_1, t_2 \in \mathbb{R}\}$  die verallgemeinerte Gruppe unitärer Operatoren zur Schar  $\{H_t : t \in \mathbb{R}\}$  der selbstadjungierten Hamiltonoperatoren ist, welche zu den verschiedenen Zeiten die infinitesimale Entwicklung des Systems bestimmen. Im zeitlich stationären Fall entwickelt sich der Zustand  $|\mathbf{v}_0\rangle$  in  $|\mathbf{v}_t\rangle = \exp(-itH)|\mathbf{v}_0\rangle$ .

Man kann nun aber die Zustände aus dem Spiel lassen, und direkt nach der Veränderung der Observablen fragen? — Man spricht hier von Heisenbergs Bild. Im zeitlich stationären Fall bringt die zeitliche Entwicklung in der Zeitspanne  $t$  den Übergang

$$\begin{aligned} A &\rightsquigarrow A_t = \exp(-itH)A \exp(-itH) && \text{gemäß} \\ \langle \mathbf{v}_t | A | \mathbf{v}_t \rangle &= \langle \mathbf{v}_0 | \exp(itH)A \exp(-itH) | \mathbf{v}_0 \rangle \end{aligned}$$

Die infinitesimale Entwicklung wird (auch im instationären Fall, gegeben durch die Schar  $\{H_t : t \in \mathbb{R}\}$ ) durch eine lineare Differentialgleichung beschrieben

$$\frac{d}{dt}A_t = \lim_{\frac{1}{t}} \frac{1}{t} ((E + itH_t)A(E - itH_t) - A) = \frac{1}{i} (H_t A - A H_t) = \frac{1}{i} [A, H_t].$$

**Fazit:** Die grundsätzliche Idee der quantenmechanischen Beschreibung, die unser geometrisches Vorstellungsvermögen sehr ernsthaft herausfordert, können wir kurz folgendermaßen beschreiben: Die Beobachtung einer 'dynamischen Variablen' liefert keinen bestimmten sondern einen zufälligen Wert, wobei die Wahrscheinlichkeitsverteilung in quadratischer Weise vom Zustandsvektor des Systems abhängt. Die zeitliche Veränderung des Zustandsvektors ist durch eine lineare Differentialgleichung in einem Hilbertraum bestimmt. (Die Linearität korrespondiert zur Idee der Superposition von Zuständen.)

**Hinweis:** Die Bewegungsgleichung für die Observablen in Heisenbergs Bild erinnert merkwürdigerweise auch irgendwie an die Bewegungsgleichung in der klassischen Hamilton'schen Mechanik. Wenn  $\{q_1, \dots, q_n; p_1, \dots, p_n\}$  kanonische Koordinaten sind, und  $H$  die Hamilton-Funktion ist, dann gilt

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r} \quad \frac{dp_r}{dt} = -\frac{\partial H}{\partial q_r}$$

Für eine Funktion  $v$  der Koordinaten, welche die Zeit nicht explizit enthält, gilt demnach

$$\begin{aligned} \frac{dv}{dt} &= \sum_r \left\{ \frac{\partial v}{\partial q_r} \frac{dq_r}{dt} + \frac{\partial v}{\partial p_r} \frac{dp_r}{dt} \right\} \\ &= \sum_r \left\{ \frac{\partial v}{\partial q_r} \frac{\partial H}{\partial p_r} - \frac{\partial v}{\partial p_r} \frac{\partial H}{\partial q_r} \right\} \\ &= [v, H]. \quad (\text{'Poisson-Klammer'}) \end{aligned}$$

Die Fachleute sind seit jeher überzeugt, dass es sich nicht um eine oberflächliche Analogie handelt. Die Zusammenhänge sind aber bis heute nicht wirklich verstanden, weder mathematisch noch physikalisch.—

### Einige Rechenübungen im unendlichdimensionalen Fall.

Mit unseren Bemerkungen zur Quantenmechanik wollten wir darauf verweisen, dass die infinitesimalen Operatoren und ihre Kommutatoren in der Vorstellungswelt der Quantentheorie eine zentrale Rolle spielen. Wir kehren zu einer rein mathematischen Betrachtung zurück. Anhand unseres Standardbeispiels für einen unendlichdimensionalen Hilbertraums, des Raums der quadratintegrierbaren  $2\pi$ -periodischen Funktionen wollen wir einen Einblick geben, was bei den Operatoren und ihren Kommutatoren zu beachten ist, wenn auch unbeschränkte selbstadjungierte Operatoren (mit diskretem oder kontinuierlichem Spektrum) ins Bild kommen, Für eine mathematisch saubere Behandlung fehlen uns natürlich die technischen Mittel.

In der Vorstellungswelt der Quantenmechanik sind die möglichen Zustände des Systems abstrakt durch 'ket-Vektoren' gegeben. Diese werden mit Hilfe von Spektralscharen, die üblicherweise aus selbstadjungierten Operatoren hergeleitet werden, 'zerlegt', was dann u. U. zu einer Darstellung durch eine Amplitudenfunktion führt.

In unserem konkreten Fall haben wir von vorneherein zwei Darstellungen der abstrakten normierten ket-Vektoren durch Amplitudenfunktionen, nämlich die Darstellung durch Folgen  $\mathbf{c} = (c_n)_{n \in \mathbb{Z}}$  mit  $\sum_n |c_n|^2 = 1$  und die Darstellung durch  $2\pi$ -periodische Funktionen  $f(t)$  mit  $\int_0^{2\pi} |f(t)|^2 dt = 2\pi$ . (Wer die Lebesgue'sche Integrationstheorie studiert hat, weiss natürlich, dass es sich bei dieser zweiten Darstellung in Wirklichkeit um Äquivalenzklassen borel-messbarer  $2\pi$ -periodischer Funktionen handelt.)

Wir werden spezielle selbstadjungierte Operatoren kennenlernen, die in der einen oder in der anderen Darstellung als Multiplikationsoperatoren erscheinen. Und wir werden uns mit den Kommutatoren einiger solcher Operatoren befassen.

Wir haben bereits darauf hingewiesen, dass der Operator  $A = \frac{1}{i} \frac{\partial}{\partial t}$ , der zunächst nur auf dem Prä-Hilbertraum der trigonometrischen Polynome definiert wird, in eindeutiger Weise zu einem (unbeschränkten!) selbstadjungierten Operator fortgesetzt werden kann. Die Eigenvektoren des Operators sind die Vielfachen der 'reinen Sinusschwingungen'  $e_n$ ,  $n \in \mathbb{Z}$ . ( $e_n(t) = e^{int}$  für  $t \in \mathbb{R}/2\pi$ ). Die  $e_n$  bilden eine ONB.

Für einen Einheitsvektor  $\mathbf{v} = f(t) = \sum_n c_n \cdot e^{int}$  können wir die Folge  $(c_n)_{n \in \mathbb{Z}}$  als die Amplitudenfunktion eines Zustands verstehen. Wenn sich das System in diesem Zustand befindet, dann nimmt die Observable  $A$  den Wert  $n$  mit der Wahrscheinlichkeit  $|c_n|^2$  an. In Zuständen mit  $\sum |n| \cdot |c_n|^2 < \infty$  existiert der Erwartungswert  $\tilde{a} = \sum n \cdot |c_n|^2 = \mathfrak{E}_{\mathbf{v}} A = \langle \mathbf{v} | A | \mathbf{v} \rangle$ ; und wir haben  $\text{var}_{\mathbf{v}} A = \sum_n (n - \tilde{a})^2 |c_n|^2$ . Für die 'reinen' Zustände  $\mathbf{e}_n$  verschwindet die Varianz.

Interessante Zustände, in denen die Werte von  $A$  weit streuen, sind z. B. die folgenden:

$$\mathbf{v}_N(t) = \frac{1}{\sqrt{2N+1}} \cdot D_N(t-s) = \frac{1}{\sqrt{2N+1}} \sum_{-N}^N e^{-ins} \cdot e^{int}.$$

Diese Funktionen haben wir oben die (ins Zentrum  $s$  verschobenen) normierten Dirichlet-Kerne genannt. (Vgl. das Beispiel 2.1.3 auf Seite 88.) Offenbar hat der Operator  $A$  in einem solchen Zustand eine Varianz von der Größenordnung  $\frac{1}{3}N^2 + O(N)$  für  $N \rightarrow \infty$ . Wir werden andere Operatoren kennenlernen, die in diesen Zuständen eine recht kleine Varianz  $\approx \frac{c}{N}$  haben.

Ein studierenswerter Typ von Operatoren sind die sog. kontinuierlichen Faltungsoperatoren. Für ein trigonometrisches Polynom  $k(t) = \sum k_m \cdot e^{imt}$  definieren wir

$$N_k : V \ni f(\cdot) \mapsto N_k f(\cdot) = (k \star f)(\cdot) = \frac{1}{2\pi} \int_0^{2\pi} k(s) \cdot f(\cdot - s) ds.$$

Im Folgenbild unseres Hilbertraums erscheint  $N_k$  als ein Multiplikationsoperator

$$N_k : V \ni f(\cdot) = \sum_n c_n \cdot e^{in(\cdot)} \mapsto N_k f(\cdot) = \sum_n (k_n c_n) \cdot e^{in(\cdot)}.$$

Es gilt nämlich  $\frac{1}{2\pi} \int_0^{2\pi} e_m(s) \cdot e_n(\cdot - s) ds = e_n(\cdot) \cdot \delta_{mn}$ , also  $e_m \star e_n = e_n \cdot \delta_{mn}$ , und daher  $k \star \sum c_n \cdot e_n = \sum (k_n c_n) \cdot e_n$ . Der Operator  $N_k$  ist genau dann selbstadjungiert, wenn  $k(\cdot)$  reellwertig ist, wenn also  $k_{-m} = \bar{k}_m$ .

Wir müssen nicht fordern, dass die Folge  $(k_m)_m$  eine finite Folge ist. Jede beschränkte Folge definiert einen beschränkten Operator  $N_k$ ; denn für beschränkte Folgen gilt  $\sum |c_n|^2 < \infty \implies \sum |k_n c_n|^2 < \infty$ . Es handelt sich dann aber nicht notwendigerweise um einen Faltungsoperator. Der Operator  $N_k$  ist genau dann eine Isometrie, wenn alle  $k_m$  den Betrag haben. Ein Beispiel ist der Identitätsoperator, der zur konstanten Folge  $(k_m) = 1$  gehört. Die Folge  $k_m = e^{-ims}$  liefert den Operator  $U_s$  der Verschiebung um  $s$ , denn  $e^{-ims} \cdot e_n(\cdot) = e_n(\cdot - s)$ . Die unbeschränkte Folge  $(k_m) = m$  liefert den unbeschränkten Operator  $A = \frac{1}{i} \frac{\partial}{\partial t}$ .

Wir wenden uns einem anderen Typ von Operatoren zu. Wenn  $h(\cdot)$  eine reellwertige  $2\pi$ -periodische Funktion ist, dann ist  $M_h : f(\cdot) \longrightarrow (h \cdot f)(\cdot)$  ein selbstadjungierter Operator, und zwar ein beschränkter Operator, wenn die Funktion  $h$  beschränkt ist. Wenn  $h = \sum h_k e^{ikt}$  ein trigonometrisches Polynom ist, dann erscheint der Operator im Folgenbild als ein diskreter Faltungsoperator:

$$M_h : V \ni f(\cdot) = \sum c_n e^{in(\cdot)} \longmapsto M_h f(\cdot) = (h f)(\cdot) = \sum d_n e^{in(\cdot)} \quad \text{mit } d_n = \sum h_k c_{n-k}.$$

Für das Folgende finden wir es bequemer, die Funktionen  $f \in V$  und die Funktionen  $h$  als Funktionen auf dem Intervall  $(0, 2\pi]$  zu verstehen, (die dann gegebenenfalls  $2\pi$ -periodisch fortgesetzt werden.) Interessant ist z. B die Multiplikation mit der  $2\pi$ -periodischen Funktion  $e(t)$ , die auf  $(0, 2\pi]$  mit der Koordinatenvariablen übereinstimmt.  $e(t) = \pi + 2s(t)$ , wo  $s(\cdot)$  die Euler'sche Sägezahnfunktion ist. Im Beispiel auf Seite 88 haben wir gesehen  $s(t) = \lim s_K(t)$  mit

$$s_K(t) = \sin t + \frac{1}{2} \sin 2t + \cdots + \frac{1}{K} \sin Kt,$$

$$s'_K(t) = \cos t + \cos 2t + \cdots + \cos Kt = \frac{1}{2}(D_K(t) - 1).$$

Die Multiplikationsoperatoren haben keine Eigenvektoren. Man könnte unscharf sagen: Funktionen, die ausserhalb einer kleinen Umgebung von  $s$  sehr klein sind, sind approximativ Eigenvektoren zum Eigenwert  $h(s)$ . Solche Einheitsvektoren wie die oben beschriebenen  $\mathbf{v}_N(t) = \frac{1}{\sqrt{2N+1}} \cdot D_N(t-s)$  könnte man (für große  $N$ ) als approximative Eigenvektoren für  $M_h$  ansehen  $M_h v_N \approx h(s) \cdot v_N$ . Dirac hatte seinerzeit vorgeschlagen, den Limes wirklich zu konstruieren; er kam zu den sog. Diracfunktionen, die fern von  $s$  verschwinden, im Punkt  $s$  aber so groß sind, dass die ein Integral  $\neq 0$  haben. Die Schwierigkeiten einer solchen Konstruktion kann man vermeiden, wenn man sich auf die Idee der kontinuierlichen orthogonalen Zerlegung stützt: Es sei  $\Pi_t$  für  $t \in (0, 2\pi)$  die orthogonale Projektion, die der Funktion  $f(\cdot)$  die Funktion  $\Pi_t f(\cdot) = (f \cdot \mathbf{1}_{(0,t)})(\cdot)$  (im Intervall  $(0, 2\pi]$ ) zuordnet. Es gilt dann in einem offensichtlichen Sinn  $id = \int_0^{2\pi} d\Pi_t$ , und die kontinuierliche orthogonale Zerlegung ist mit allen Multiplikationsoperatoren  $M_h$  verträglich.  $M_h f = \int_0^{2\pi} h(t) d\Pi_t f$ . Die kontinuierliche orthogonale Zerlegung  $\{\Pi_t : t \in (0, 2\pi)\}$  erscheint als die Spektralzerlegung des Operators  $M_e$ ; denn es gilt  $M_e = \int_0^{2\pi} t \cdot d\Pi_t$ .

Bzgl. unserer kontinuierlichen orthogonalen Zerlegung dient die Funktion  $f(t)$  selbst als eine Amplitudenfunktion des Zustands  $\mathbf{v} \leftrightarrow f(t)$ . Wir haben die Entsprechung  $\mathbf{w}_n \leftrightarrow e^{int}$ . Die Normiertheit eines ket-Vektors  $\mathbf{v}$  bedeutet  $1 = \|\mathbf{v}\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt$ . Die Spektralwerte des Operators  $M_e$  sind die Punkte  $t \in (0, 2\pi]$ .  $M_e f(t) = t \cdot f(t)$ . Für alle Intervalle  $(a, b] \subset (0, 2\pi]$  gilt

$$\text{Ws}_f(M_e \text{ hat einen Wert in } (a, b]) = \frac{1}{2\pi} \int_a^b |f(t)|^2 dt.$$

$$\text{Ws}_f(M_h \text{ hat einen Wert } \in (a, b]) = \frac{1}{2\pi} \int \mathbf{1}_{\{t: a < h(t) \leq b\}} |f(t)|^2 dt.$$

Das kann man auch in der Vorstellungswelt der Zustände interpretieren: In jedem beliebigen Zustand des Systems  $\mathbf{v}$  ist die Verteilung der Observablen  $M_h$  das  $h$ -Bild der Verteilung von  $M_e$ . Man rechnet also wie in der 'kommutativen' Wahrscheinlichkeitstheorie, solange man es nur mit den Observablen  $M_h$  zu tun hat. Man kann sagen, die Observable  $M_h$  sei das  $h$ -Bild der Observablen  $M_e$  und man könnte symbolisch schreiben  $h(M_e) = M_h$ . Es gilt jedenfalls

$$\begin{aligned} \mathfrak{E}_f M_h &= \frac{1}{2\pi} \int_0^{2\pi} h(t) \cdot |f(t)|^2 dt = \langle f | M_h | f \rangle = \tilde{h} \\ \text{var}_f M_h &= \frac{1}{2\pi} \int_0^{2\pi} (h(t) - \tilde{h})^2 \cdot |f(t)|^2 dt = \langle f | (M_h - \tilde{h} \cdot id)^2 | f \rangle. \end{aligned}$$

Für den speziellen Zustand  $\mathbf{v}_N = f_N(t) = \frac{1}{\sqrt{2N+1}} \cdot D_N(t-s) = \frac{1}{\sqrt{2N+1}} \left( \sum_{-N}^N e^{-ins} e^{int} \right)$ . haben wir früher (auf Seite 88) gesehen  $|f_N|^2(t) = F_{2N+1}(t-s)$ . Die Observable  $M_e$  hat also in diesem Zustand die Wahrscheinlichkeitsdichte  $\frac{1}{2\pi} F_{2N+1}(t-s) dt$  im Intervall  $(0, 2\pi]$ . Genauer gesagt

$$\text{Ws}_{f_N}(M_e \text{ hat einen Wert } \in (a, b]) = \frac{1}{2\pi} \int_a^b F_{2N+1}(t-s) dt, \quad \text{für } 0 < a < b < 2\pi.$$

Im Falle  $N = 0$  ist das die uniforme Verteilung im Intervall  $(0, 2\pi]$ ; für große  $N$  ist die Verteilung der Observablen  $M_e$  auf eine kleine Umgebung von  $s$  konzentriert. Der Erwartungswert des Multiplikationsoperators  $M_h$  im speziellen Zustand  $\mathbf{v}_N$  ist

$$\tilde{h}_N = \mathfrak{E}_{\mathbf{v}_N} M_h = \langle \mathbf{v}_N | M_h | \mathbf{v}_N \rangle = \frac{1}{2\pi} \int_0^{2\pi} h(t) \cdot F_{2N+1}(t-s) dt \approx h(s).$$

Die Varianz der Observablen  $M_e$  im Zustand  $\mathbf{v}_N$  ist  $\text{var}_{\mathbf{v}_N} M_e = \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2 \cdot F_{2N+1}(t) dt \approx \frac{C}{N}$ . Die Varianz der Observablen  $M_h$  ist  $\approx (h'(s))^2 \cdot \frac{C}{N}$ , wenn  $h(\cdot)$  glatt ist.

**Allerlei Kommutatoren** Kommen wir zu den Kommutatoren der speziellen Observablen  $A = \frac{1}{i} \frac{\partial}{\partial t}$  mit den ‘Multiplikations- Observablen’  $M_h$ .

**Lemma.** Wenn  $h$  eine stetig differenzierbare  $2\pi$ -periodische Funktion ist, dann ist der Kommutator des Multiplikationsoperators  $M_h$  mit  $A = \frac{1}{i} \frac{\partial}{\partial t}$  der Multiplikationsoperator  $M_{h'}$ .  $\frac{1}{i}[M_h, A] = M_{h'}$ . In den Eigenzuständen des Operators  $A$ ,  $(\mathbf{w}_n \longleftrightarrow e^{int})$ , verschwindet der Erwartungswert dieses Kommutators  $M_{h'}$ .

**Beweis.**

$$\begin{aligned} \langle g | \frac{1}{i}(M_h A - A M_h) | f \rangle &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{g}(t) \cdot \left( -h(t) f'(t) + \frac{\partial}{\partial t}(h(t) \cdot f(t)) \right) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{g}(t) \cdot h'(t) \cdot f(t) dt = \langle g | M_{h'} | f \rangle. \end{aligned}$$

Im Zustand  $\mathbf{v} = f(t) = \sum c_n \cdot e^{int}$ , ( $\|\mathbf{v}\|^2 = 1$ ) liefert die Unschärferelation

$$\sqrt{\text{var}_{\mathbf{v}} M_h} \cdot \sqrt{\text{var}_{\mathbf{v}} A} \geq \frac{1}{2} |\langle \mathbf{v} | M_{h'} | \mathbf{v} \rangle| = \frac{1}{4\pi} \left| \int_0^{2\pi} h'(s) \cdot |f(s)|^2 ds \right|.$$

Die Ungleichung zeigt wegen  $\text{var}_{\mathbf{w}_n} A = 0$ , dass der Erwartungswert des Kommutators  $M_{h'}$  in diesen Zuständen verschwindet. Man kann aber auch direkt nachrechnen

$$\langle \mathbf{w}_n | \frac{1}{i}[M_h, A] | \mathbf{w}_n \rangle = \langle \mathbf{w}_n | M_{h'} | \mathbf{w}_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} h'(s) ds = 0.$$

Der Erwartungswert des Multiplikationsoperator  $M_h$  im Zustand  $\mathbf{w}_n$  ist

$$\mathfrak{E}_{\mathbf{w}_n} M_h = \langle \mathbf{w}_n | M_h | \mathbf{w}_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} h(t) dt = \tilde{h}.$$

Die Heisenberg’sche Unschärferelation liefert nur die triviale Abschätzung der Varianz; man kann hier die Varianz aber direkt ausrechnen:  $\text{var}_{\mathbf{w}_n} M_h = \frac{1}{2\pi} \int_0^{2\pi} (h(t) - \tilde{h})^2 dt$ ,

Betrachten wir nun die Observable  $M_e$ . Ihre Verteilung im Zustand  $\mathbf{w}_n$  ist die Gleichverteilung auf dem Intervall  $(0, 2\pi]$ . Da  $e(\cdot)$  unstetig ist, können wir den Kommutator mit  $A$  nicht wie oben berechnen. Insbesondere ist die Formel  $\langle \mathbf{w}_n | \frac{1}{i}[M_h, A] | \mathbf{w}_n \rangle = 0$  nicht anwendbar. Wir stellen nämlich fest, dass  $e'$  im Inneren des Intervalls konstant = 1. Der Multiplikationsoperator  $M_{e'}$  ist die Identität. Wir können hier den Begriff des Kommutators für unbeschränkte Operatoren nicht klären. Zur Unterstützung der Behauptung  $\frac{1}{i}[M_e, A] = id$  können wir aber immerhin ein relevantes Integral ausrechnen. Für ‘gute’ Zustände  $\mathbf{v} = f(t)$ ,  $\mathbf{w} = g(t)$  gilt:

$$\begin{aligned} \langle \mathbf{w} | \frac{1}{i}(M_e A - A M_e) | \mathbf{v} \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \bar{g}(t) \left( -e(t) \frac{\partial}{\partial t} f(t) + \frac{\partial}{\partial t}(e(t) \cdot f(t)) \right) dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \bar{g}(t) \left( -t \cdot f'(t) + \frac{\partial}{\partial t}(t \cdot f(t)) \right) dt = \langle \mathbf{w} | \mathbf{v} \rangle. \end{aligned}$$