

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

Einführung in die Mathematik I
(Elementare Algebra und Lineare Geometrie)

WS 2007/08

H. Dinges

Vorwort

Die Mathematik ist in der Öffentlichkeit präsent, und sie gilt als wichtig. Als Lerngebiet hat sie zur Zeit aber leider nicht die *Attraktivität*, die sie haben könnte oder sollte. Es gelingt nicht so leicht, auf den Wegen der Mathematik die Sehnsucht nach den Weiten der Wissenschaft zu evozieren. Man kann vielleicht darüber streiten, ob unser Bildungssystem genügend viele reine Mathematiker hervorbringt, es steht aber ausser Zweifel, dass unsere Gesellschaft mehr mathematisch Gebildete gut brauchen könnte, Leute, die nicht nur die mathematischen Techniken des 19. Jahrhunderts (im Sinne der Schulmathematik) leidlich beherrschen, sondern auch mit den abstrakteren mathematischen Denkweisen des 20. Jahrhunderts flüssig umgehen können, und sich dabei von einem positiven und offenen Verhältnis zur Mathematik leiten lassen.

Die gegenwärtige *Umstellung aller Studiengänge* auf das Bachelor-Master-Schema sind ein dringender Anlass neu nachzudenken, inwieweit die mathematischen Anfängervorlesungen dem Anspruch einer in die Breite wirkenden Einführung ins mathematische Denken genügen können. Nach meiner Meinung ist ein Umdenken erforderlich. Die mathematischen Fachbereiche sollten es als einen gesellschaftlichen Auftrag auffassen, für die einschlägig begabten Studierenden dem (wohl zutreffenden aber oft billig gebrauchten) Schlagwort von der zunehmenden Mathematisierung aller Wissenschaften fassbare Konturen zu geben. Es wäre wünschenswert, dass die Anfänger mit mathematiknahen Studienzielen (insbesondere im Lehramt, in der Physik und in der Informatik) in den mathematischen Anfängervorlesungen etwas vom frischen Geist der Mathematik spürten. Das kann weder durch ein Training in den gebräuchlichsten Anwendungen geleistet werden noch durch eine Einweisung in die Routinen einer abgehobenen mathematischen Forschung. Es wäre eine sträfliche Abkehr von den (von den Bildungstheoretikern) zurecht hochgeschätzten Prinzipien des akademischen Studierens, wenn die mathematischen Fachbereiche die Pflege des mathematischen Weltverstehens von Anfang an organisatorisch und inhaltlich separierten von den durch das Bachelor-System nahegelegten schematischen Pflichtmoduln ‘Mathematik’ ohne großen geistigen Anspruch. Es sollte den Anfängervorlesungen darauf ankommen, möglichst vielen Studienanfängern die für die *Mathematisierung* charakteristischen Denkweisen der Abstraktion und Verallgemeinerung nahezubringen.

Der hier vorgelegte Text ‘Einführung in die Mathematik I’ macht Vorschläge, wie Themen aus *Elementarer Algebra und Linearer Geometrie* im Anschluss an die Schulmathematik so weiterentwickelt werden können, dass ein modernes Mathematikverständnis sichtbar wird. Mein Text sucht in den Abschnitten I – IV immer wieder Anschluss an die zentralen Gegenstände der Schulmathematik: Rechengrößen, Abbildungen, Ungleichungen und euklidische analytische Geometrie. Erst der Abschnitt V zeigt den Charakter einer mathematischen Theorie. Er ordnet mit vollständigen Beweisen die vorher adhoc entwickelten Techniken der Vektor- und Matrizenrechnung.

Bei meiner Abkehr von den üblichen Einführungen war ich durch große Vorbilder ermutigt. Vor allem möchte ich Feynman’s ‘Lectures on physics’ nennen und den ‘Course in mathematics for students of physics’ von P. Bamberg und S. Sternberg (erschieden 1988 bei Cambridge University Press). Durch Meisterwerke dieser Art kann man eine Idee bekommen, wie tragfähige didaktische Modularisierungen aussehen können, und wie man bei den hochmotivierten Studierenden Begeisterung nähren kann, ohne die weniger leistungsfähigen abzuhängen. Wenn diese Vorbilder den Alltag der Ausbildung noch nicht in größerem Umfang erreicht haben, so haben sie doch die Landschaft verändert, in welcher die jüngeren Dozenten den Vorlesungsstil suchen können, welcher den lokalen Erfordernissen gerecht wird.

Das primäre Anliegen unserer Einführung sind die *mathematischen Sachverhalte*, das Ideal des kunstgerechten Beweisens stellen wir zunächst einmal zurück. Eine klare Unterscheidung zwischen Aussagen über Tatbestände und Beweisen erwarten wir natürlich vom Dozenten und (mit Abstrichen) von den Tutoren. Aufwendigere Beweise werden in unserem Text ausgespart, wenn sie keinen wesentlichen Beitrag leisten, die Einsicht in den Sachverhalt zu befördern, oder wenn sie durch ihre Methodik herausführen aus der aktuellen Lernumgebung. Gute Themen sind für uns diejenigen, bei denen die Studierenden jenseits der Beweise Erkenntnisse gewinnen, nach dem Muster ‘Ach so verhält sich die Sache’. In den Übungen ermutigen wir sie zu einsichtigem Argumentieren; wir versuchen zu vermeiden, dass das (für die forschende Mathematik unabdingbare) penible mathematische Schliessen die inhaltlich orientierten Anfänger lähmt. Unser Ansatz sollte die Heranbildung ‘echter’ Mathematiker nicht gefährden; die primär am Beweisen interessierten Studierenden werden mit Sicherheit in späteren Mathematikvorlesungen ausreichend bedient.

Die Anfänger machen sich erfahrungsgemäß kaum Gedanken über den *Aufbau eines Vorlesungszyklus*. Wir wollen nichts dagegen einwenden, wenn sie zunächst einmal jeden einzelnen Unterabschnitt so verstehen, dass hier jeweils zwei oder drei wichtige mathematische Ideen skizziert werden, die dann später bei Bedarf in mathematisch streng aufgebaute Theorien einzubauen sind. Für eine Einführung, die in die Breite wirken soll, passt u. E. weder das Ideal eines strengen Aufbaus noch das Ideal der Methodenreinheit. Viele der wichtigen elementaren Sachverhalte enthalten im Kern Denkmuster aus Linearer Algebra und Analysis; wir halten nichts von einer Trennung dieser Bereiche in der Anfängerausbildung.

Die meisten Unterabschnitte sind im Rahmen von Anfängervorlesungen für Mathematiker, Physiker und Lehramtskandidaten mehrfach erprobt worden. Die Erfahrung hat gezeigt, dass man in einer vierstündigen Vorlesung etwas mehr als 20 Unterabschnitte durchnehmen kann. Grundsätzlich ist jeder Unterabschnitt als das Thema einer Doppelstunde konzipiert, wobei aber nicht intendiert ist, dass das Thema in der Doppelstunde ‘erledigt’ wird. Das auf der betreffenden Ebene Verbindliche wird jeweils gleich zu Beginn des Unterabschnitts vorgestellt. Gegen Ende des Unterabschnitts bietet der Text dann aber mehr als das, was in der Kürze der Zeit durchgenommen werden kann. Es wird vorgeschlagen, dass die Studierenden im ersten Durchlauf (parallel zu den Übungen) die ersten Seiten jedes Unterabschnitts gründlich bearbeiten; der (mehr oder weniger umfangreiche) Rest kann den Studierenden zum freiwilligen Lesen empfohlen werden. Manche Unterabschnitte haben in der Tat gegen Ende eher den Charakter einer Erzählung zum Zweck einer untechnischen Allgemeinbildung als den Charakter eines Lehrtexts.

Es liegt u. E. in der Natur der wichtigen mathematischen Ideen, dass sie bei einer ersten Begegnung nicht vollständig erfasst und abgearbeitet werden können. Viele wichtige Ideen haben einen recht elementaren Kern; und dieser sollte aus verschiedenen Blickwinkeln betrachtet werden, bevor man die einzelnen Aspekte in Systeme presst und dort endgültig fixiert. Gemäß dieser Überzeugung werden in unserem Text die großen Ideen in einem *spiraligen Sinn* öfters berührt; sie erscheinen nach und nach in weiter gespannten Horizonten, und sie werden mit immer höher entwickelten Methoden bearbeitet, (wobei wir penibel darauf achten, nichts zu sagen, was später revidiert werden muss). Beim spiraligen Vorgehen muss der Anfänger keine bleibenden Wissenslücken befürchten, wenn ihm ein wichtiger mathematischer Gegenstand bei der ersten Konfrontation dunkel oder gar abschreckend erscheint. Das spiralige Vorgehen sorgt für weitere Gelegenheiten, eine positive Beziehung zu den Gegenständen aufzubauen, wenn sie

anderswo aus einem anderen Blickwinkel beleuchtet werden; nicht selten erscheint dann auch der ursprüngliche Zugang auf einmal licht und einsichtig.

Die *wiederholte Zuwendung* zu wichtigen mathematischen Gegenständen wirkt erfahrungsgemäß auf die Studierenden nicht langweilend, sondern im Gegenteil motivierend. Die Erfahrung zeigt andererseits, dass das verständnisarme Auswendiglernen für die Prüfung durch den spiraligen Aufbau erschwert wird. Der objektive Vorteil des spiraligen Vorgehens gegenüber dem linearen ist im gelungenen Fall eine bessere Vernetzung und Flexibilisierung des Wissens, was bekanntlich ein vielbeschworenes aber nicht unbedingt bequemes Anliegen eines akademischen Studiums ist.

Wir skizzieren noch einige Leitlinien der Gliederung unserer Einführung:

- I Im Abschnitt I erweitern wir die *Arithmetik*; das ‘Buchstabenrechnen’ wird über die in der Schule betrachteten Zahlbereiche hinausgeführt. Dabei folgen wir nicht dem Weg der abstrakten Algebra, auf dem die Studierenden mit den Axiomensystemen für Gruppen, Ringe und Vektorräume überfallen werden. Wir bevorzugen den ‘synthetischen’ Zugang gegenüber dem ‘analytischen’; und das heisst hier, dass die neuen Typen von *Rechengrößen* zunächst einmal konkret in Erscheinung treten. Es dürfte deutlich werden, dass es im Folgenden darum gehen muss, die konkreten Rechengrößen, insbesondere die Matrizen (und die Matrizenmultiplikation) von den ‘Anwendungen’ her gründlich zu verstehen.
- II Die Idee der *Abbildungen*, die nicht an eine bestimmte Darstellungsform gebunden sind, ist erfahrungsgemäß für die Anfänger schwerer zu begreifen als die Idee der Rechengrößen und die Idee der in der Schule behandelten Funktionen auf einem Intervall. Die Abbildungen in abstrakter Form und die Gruppen von Abbildungen signalisieren den Anfang einer modernen Auffassung von *Geometrie*.
- III Wir befassen uns mit *Ungleichungen*, insbesondere mit der Dreiecksungleichung und mit der Sublinearität. Die topologischen Grundbegriffe für metrische Räume werden vorgestellt. Wir behandeln normierte Vektorräume und Trennungssätze für konvexe Mengen sowie (im endlichdimensionalen Fall) die Legendre-Dualität. Im Begriff der konvexen Funktion ergeben sich auch Beziehungen zur elementaren Analysis, die aber nicht vertieft werden. In einem Anhang werden die in unserer Einführung betrachteten speziellen Funktionentypen zusammengestellt.
- IV Der Abschnitt IV ist lang geraten zugunsten der Studierenden der Physik. Hier wird die intuitive Geometrie des euklidischen Anschauungsraums weiterentwickelt zur axiomatisch aufgebauten *Geometrie des Hilbertraums*. Es sollte klar werden, dass die Räume, in denen die Physik operiert, nicht als vorgegeben zu betrachten sind; sie müssen mathematisch konstruiert werden.
- V Der Abschnitt V enthält die übliche *Lineare Algebra I* ohne die Determinanten. Neben den Matrizen studieren wir die Tableaus; dabei wird einerseits die fundamentale Idee der Vektorraumdualität deutlicher als üblich herausgestellt und andererseits die Technik der linearen Programmierung vorbereitet.

Jedem einzelnen Abschnitt ist eine **Themenübersicht** vorangestellt, damit man ihn für spezielle Lehrzwecke leichter aus dem Gesamtzusammenhang herauslösen kann. Ein ausführlicher **Index** ganz am Ende des ermöglicht einen schnellen Überblick über die Themenverwandtschaften quer durch unsere ‘Einführung’.

Ich hoffe, dass die Themenauswahl und der Rhythmus des hier vorgelegten Ansatzes einige Kollegen ermutigen werden, aus den überholten Bahnen der Anfängerausbildung auszubrechen und nach Wegen zu suchen, die Mathematik des 20. Jahrhunderts bei den verschiedenen potentiellen Nutzniessern populärer zu machen. Ich wage die These, dass eine spiralig aufgebaute Sammlung von Vorlesungen wie die hier vorgelegte dem Ansehen der Mathematik in einer breiteren Öffentlichkeit besser dienen kann als eine linear aufgebaute mathematische Theorie.

Für den Übungsbetrieb erwarte ich, dass sich aus den Einzelbetrachtungen vielseitige Anregungen für attraktive (und doch nicht zu schwierige) Aufgaben ergeben. Gute Aufgaben sind für uns solche, bei welchen der Anfänger spürt, dass er (unter guter Anleitung) selbständig vorankommen kann beim Studium mathematischer Gegenstände. Die Anfängerveranstaltungen sollten sich nicht aufhalten mit solchen Aufgaben, die den Studierenden, die noch recht wenig von mathematischen Inhalten wissen, das Beweisen von mathematischen Kleinigkeiten zumuten.

Schliesslich hoffe ich, dass unsere Texte auch fortgeschrittene Studierende ansprechen. Ich habe jedenfalls gehört und beobachtet, dass die Tutoren, die die Anfängerübungen betreuen, die Texte schätzen; sie erarbeiten sich durch diese Texte eine reifere Einstellung zu den Sachverhalten, die sie als Anfänger zwar irgendwie aufgenommen, aber in ihrer Bedeutung und ihrer Vernetzung nicht einschätzen konnten.

Frankfurt, 19. März 2008

Inhaltsverzeichnis

I	Rechengrößen	1
I.1	Von den Zahlen zu den Matrizen	1
I.2	Komplexe Zahlen; die Euler'sche Formel	9
I.3	Polynome als Rechengrößen	19
I.4	Allerlei komplexe $n \times n$ -Matrizen. Faktorisierungen	31
I.5	Vektoren sind Elemente eines Vektorraums. Basiswechsel	41
I.6	Minimalpolynome. Verallgemeinerte Eigenvektoren.	51
I.7	Das Faltungsprodukt über \mathbb{Z} . Trigonometrische Polynome	59
I.8	Newton's Binomialreihe. Formale Potenzreihen	71
II	Abbildungen	83
II.1	Abbildungen, Permutationen, Gruppen	83
II.2	Die Riemann'sche Zahlenkugel; Kreisverwandtschaften	91
II.3	Einige spezielle Transformationsgruppen	99
II.4	Affine Räume, affine Funktionen, affine Abbildungen	111
II.5	Quadratische Funktionen	125
II.6	Polynome als Abbildungen von \mathbb{C} in sich	137
III	Metrik, Norm, Konvexität	151
III.1	Dreiecksungleichung und Subadditivität	151
III.2	Normierte Vektorräume und stetige Linearformen	161
III.3	Konvexe Mengen und konvexe Funktionen	169
IV	Inneres Produkt. Orthogonalität	189
IV.1	Parallelogrammgleichung und Polarisierung	189
IV.2	Hauptachsen, Orthogonale Zerlegungen, Spektralsatz.	199
IV.3	Einparametrische Matrixgruppen	209
IV.4	Gekoppelte Oszillatoren	217
IV.5	Die schwingende Saite	231
IV.6	Diskrete Fourier-Transformation	239
IV.7	Drehungen im Minkowskiraum	247
IV.8	Zerlegung einer Gruppendarstellung	259
IV.9	Spurfunktionen. Die Charaktertafel einer endlichen Gruppe G	271
IV.10	Unitäre und hermitesche Matrizen	283
V	Lineare Algebra	295
V.1	Vektorräume: Basis und Dimension	295
V.2	Lineare Gleichungssysteme	301
V.3	Tableaus. Vollständiger Austausch	309
V.4	Eliminationsmethode; LU-Faktorisierung	319
V.5	Matrizen vom Rang r , Singulärwertzerlegung	327
V.6	Kleinste Quadrate und lineare Schätzung	337
A	Anhang : Mengen und Relationen	348

Erläuterungen zum Inhalt

I. Rechengrößen

Wir beginnen mit algebraischen Themen, die zum Teil schon im Schulunterricht vorgezeichnet sind. Die Sprache und die Notation der naiven Mengenlehre, die wir am konkreten Stoff benützen, ist im Anhang zusammengestellt.

I.1 Von den Zahlen zu den Matrizen

Zahlbereichserweiterungen $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$. Die Assoziativgesetze und die Distributivgesetze gelten auch für Matrizen. Was versteht man unter der Inversen einer $I \times J$ -Matrix? Die Determinante einer 2×2 -Matrix; die Multiplikativität dieser Determinantenfunktion $\det(\cdot)$.

I.2 Komplexe Zahlen; die Euler'sche Formel

Komplexe Zahlen als Punkte in der Gauss'schen Zahlenebene. Darstellung durch spezielle reelle 2×2 -Matrizen. Konjugation. Der Betrag einer komplexen Zahl. Polarkoordinaten und die Euler'sche Formel $z = |z| \cdot (\cos \varphi + i \sin \varphi) = |z| \cdot e^{i\varphi}$. Wechselströme. Die Exponentialreihe im Reellen und im Komplexen.

$$e^z = \lim \left(1 + \frac{z}{n} \right)^n = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \dots$$

Nicht nur für Zahlen, sondern auch für kommutierende Matrizen gilt: $\exp(A + B) = \exp(A) \cdot \exp(B)$

I.3 Polynome als Rechengrößen

Der Polynomring $\mathbb{C}[z]$ ist ein kommutativer Ring mit Einselement. Der Grad eines Polynoms. Horner-Schema. Umzentrieren. Teilen mit Rest. Euklidischer Algorithmus. Gebrochenrationale Funktionen. Partialbruchzerlegung.

I.4 Allerlei komplexe $n \times n$ -Matrizen. Faktorisierungen.

Dieser Abschnitt dient der Eingewöhnung in Rechentechniken, die später immer wieder gebraucht werden. Das Rechnen in \mathbb{C} wird (mit der nötigen Vorsicht und den nötigen Änderungen) auf einen allgemeineren Objektbereich, nämlich die Algebra der komplexen $n \times n$ -Matrizen ausgedehnt: Addition, Multiplikation und die (hermitesche) Konjugation sind die Rechenoperationen.

Hinweis zum Aufbau: Die Abbildungen, die man oftmals mit den Matrizen verbindet, interessieren uns hier noch nicht. Wir interessieren uns auch noch nicht für die (reellen symmetrischen) Matrizen als Darstellungsform für quadratische Funktionen oder für die (hermiteschen) Matrizen als Darstellungsform für (hermitesche) Sesquilinear-Formen. Wir formulieren hier schon einmal ohne Beweise einige Faktorisierungssätze für nichtsinguläre Matrizen; die üblichen Verallgemeinerungen und die dazugehörigen Algorithmen sind Thema von Abschnitt V.

I.5 Vektoren sind Elemente eines Vektorraums. Basiswechsel.

Das Axiomensystem für einen \mathbb{K} -Vektorraum. Mengen von Matrizen als Beispiele. Verschiebungsvektoren und „Ortsvektoren“ im Anschauungsraum. Didaktisches: Es ist zu eng, wenn man sagt: „Alle Größen, die wie ein Schritt im Raum eine Richtung besitzen, werden Vektoren genannt“. Die Matrizen eines Basiswechsels.

I.6 Minimalpolynome. Verallgemeinerte Eigenvektoren.

Dieser Unterabschnitt kombiniert die Rechenregeln für Polynome mit den Rechenregeln für quadratische Matrizen. Ausserdem kommt der Vektorraum aller n -Spalten ins Spiel. In einem leichten Vorgriff auf später (‘Abbildungen’) kommen neben den Matrizen eines Basiswechsels hier auch die quadratischen Matrizen zur Darstellung eines Endomorphismus ins Bild. Zum ersten Mal treffen wir auch auf direkte Zerlegungen. Die hauptsächlichen Anwendungen der verallgemeinerten Eigenvektoren liegen in Bereichen, die wir erst viel später in den Blick nehmen (nämlich bei den einparametrischen Matrixgruppen oder bei den linearen Differentialgleichungen mit konstanten Koeffizienten). Der Unterabschnitt I.6 kann daher zurückgestellt werden.

Hinweis zum Aufbau: Die beiden folgenden Unterabschnitte I.7 und I.8 enthalten Brücken zur elementaren Analysis und zur Stochastik (im Sinne der Schulmathematik). Dabei bieten sich Aufgaben mit MAPLE an.

Von der Technik her ist ein Sprung von I.5 (Axiomatik der Vektorräume) direkt in den Abschnitt V möglich; dort dringen wir tiefer ein in die Welt der \mathbb{K} -Vektorräume. Einen solchen Sprung wollen wir aber nicht empfehlen, weil reine ‘Lineare Algebra’ den Horizont einer Einführung in die Mathematik doch allzusehr einschränkt. Dem ausschliesslich algebraisch orientierten Leser wird empfohlen, sofort in den Abschnitt II (‘Abbildungen’) zu springen.

I.7 Das Faltungsprodukt über \mathbb{Z} . Trigonometrische Polynome.

Hier studieren wir einige konkrete \mathbb{C} -Vektorräume mit zusätzlicher Struktur. Gewisse Räume von komplexen Gewichtungen auf \mathbb{Z} tragen (vermöge des Faltungsprodukts) eine ähnliche algebraische Struktur wie die oben besprochene (assoziative und kommutative) Algebra $\mathbb{C}[z]$ der Polynome. Das Interesse geht allerdings in eine andere Richtung; bei den trigonometrischen Polynomen und bei den formalen Potenzreihen gibt es kein Teilen mit Rest und kein Umzentrieren; andere Strukturen treten in den Vordergrund.

Das Falten von Gewichtungen $\mathbf{a} = (a_n)_{n \in \mathbb{Z}}$. „Schriftliches Multiplizieren“ ist Falten mit Übertrag. Charakteristische Funktionen. Die geometrische Reihe und die geometrischen Gewichtungen werden ausführlich behandelt.

Trigonometrische Polynome und Reihen.

I.8 Newton’s Binomialreihe. Formale Potenzreihen

Der Begriff der erzeugenden Funktion ist eine Variante des Begriffs der charakteristischen Funktion. Die erzeugenden Funktionen leisten manchmal gute Dienste, wenn es um ‘Gewichtungen’ auf \mathbb{Z}_+ geht. Newton’s Binomialreihe bringt die Binomialkoeffizienten in den Blick. Der im Weiteren entwickelte Kalkül der formalen Potenzreihen kann auch als eine Einstimmung auf die komplexe Funktionentheorie verstanden werden.

Im Vordergrund aber steht noch einmal das rein algebraische Anliegen von Teil I: Rechengesetze. Bei den formalen Potenzreihen hat man nicht nur Addition und Multiplikation mit den Assoziativ- und die Distributivgesetzen. Man hat auch noch die Reziprokenbildung für gewisse Objekte, die Verkettung für gewisse andere Objekte und die (formale!) Ableitung.

II. Abbildungen

Das geometrische Denken ist eine wichtige Komponente aller Mathematik; der Dreh- und Angelpunkt ist dabei der Begriff der Abbildung. Spezifischer geht es in der Geometrie um Gruppen von Abbildungen. Man hat sogar gesagt, dass Geometrie schlussendlich nichts anderes ist als das Studium von Symmetriegruppen; damit ist man wieder bei der Algebra. Die andere Komponente der Geometrie ist die Anschaulichkeit, wobei man aber natürlich nicht bei den allzu einfachen Vorstellungen von einem uns umgebenden Raum stehen bleibt. Wir geben hier eine erste Einführung in die Gedankenwelt der Transformationsgruppen. Besonders schöne ‘anschauliche’ Beispiele liefern uns hier die erweiterte komplexe Ebene und die 2×2 -Matrizen. Wir beginnen jedoch sehr allgemein im Sinne der abstrakten Mengenlehre.

II.1 Abbildungen, Permutationen, Gruppen

Definitionsbereich, Zielbereich und Bildbereich einer Abbildung. Injektive, surjektive, bijektive Abbildungen. Das Hintereinanderschalten von Abbildungen. Die Darstellung von Abbildungen endlicher Mengen durch Graphen. Permutationen. Ihre Zyklendarstellung. Transformationsgruppen, speziell Permutationsgruppen. Die abstrakte Definition einer Gruppe. Die Oktaedergruppe als Beispiel. Isomorphie von Gruppen. Untergruppen. Die alternierende Gruppe als Untergruppe der symmetrischen Gruppe. Volles Urbild. Pullback.

II.2 Riemann’sche Zahlenkugel; Kreisverwandtschaften

Die Spiegelung am Einheitskreis bildet Kreise in Kreise ab. Drei Beweise. Die allgemeine Kreisgleichung in $\bar{\mathbb{C}}$. Die Gruppe der lineargebrochenen Abbildungen (Möbiustransformationen) ist erzeugt von den Translationen, den Drehstreckungen und der Reziprokenabbildung. Die Riemann’sche Zahlenkugel, stereographische Projektion. Kreise in $\bar{\mathbb{C}}$ entsprechen Kreisen auf der Zahlenkugel. Die stereographische Projektion ist winkeltreu. Hinweise auf nichteuklidische Geometrie.

II.3 Gruppenwirkung; spezielle Transformationsgruppen

Gruppenwirkung (‘group action’), allgemeine Definition. Transitivität. Fixgruppen. Die Wirkung von $\mathrm{SU}(2)$ auf der Riemann’schen Zahlenkugel. Die Wirkung von $\mathrm{SU}(1, 1)$ auf dem Innern des Einheitskreises und auf der oberen Halbebene. Cayley-Transformation. Die Affinitäten eines Spaltenraums $\mathbb{K}_{\mathrm{Sp}}^I$, dargestellt mit der Hilfe von Matrizen. Mehr über die lineargebrochenen Abbildungen der Riemann’schen Zahlenkugel: Die Gruppe aller Möbiustransformationen wirkt einfach transitiv auf der Menge aller Zahlentripel. Das Doppelverhältnis von Quadrupeln $(z_1 : z_2 : z_3 : z_4)$ bleibt erhalten.

II.4 Affine Räume, affine Funktionen, affine Abbildungen

Ein affiner Raum ist eine Punktmenge L , auf welcher ein Vektorraum V einfach transitiv wirkt. Der Kalkül der Ortsvektoren; Schwerpunktskoordinaten. Anschluss an die Schulmathematik. Affine Funktionen und ihre Linearformen. Der Dualraum V^* . Duale Basen: $\{v_j : j \in J\}$ in V und $\{\ell^j(\cdot) : j \in J\}$ in V^* .

Affine Koordinatensysteme. Affine Abbildungen von affinen Räumen. Eine Abbildung Φ eines affinen Raums in einen affinen Raum ist genau dann eine affine Abbildung, wenn der Pullback Φ^* den affinen Funktionen affine Funktionen zuordnet.

II.5 Quadratische Funktionen

Koordinatenfreier Zugang zum Begriff der quadratischen Funktion. Die Bilinearform zu einer quadratischen Funktion ('Polarisierung'). Satz: Sei $q(\cdot)$ eine quadratische Form und $Q(\cdot)$ die dazugehörige quadratische Form. Zu jedem Punkt P existiert dann eine Linearform $\ell_P(\cdot)$, sodass gilt

$$q(P + v) = q(P) + \ell_P(v) + \frac{1}{2}Q(v); \quad \text{oder in der Sprache der Analysis}$$

$$q(x) = q(x_0) + q'(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)^\top \cdot H \cdot (x - x_0).$$

Bemerkenswert ist, dass der „Gradient“ $\ell_P(\cdot)$ bzw. $q'(x_0)$ in affiner Weise von P bzw. x_0 abhängt.

Koordinatenwechsel. Hinweise auf Quadriken. Die quadratische Form der Minkowski-Geometrie (Relativitätstheorie).

Hinweis zum Aufbau: Die quadratischen Funktionen fallen etwas heraus aus unserem Hauptthema 'Abbildungen'; dabei geht es nämlich um Funktionen, die linear kombiniert werden und nicht um Abbildungen, die hintereinandergeschaltet werden. Der Gegenstand sollte aber u. E. bald angesprochen werden, jedenfalls vor dem Unterabschnitt IV.2. Ein Grund, bald von den quadratischen Funktionen zu reden, ist auch der, dass sie schon zu Beginn des Studiums als ein Thema der elementaren Analysis aufkommen, z. B. bei der Approximation zweiter Ordnung.

II.6 Polynome als Abbildungen von \mathbb{C} in sich

Der Fundamentalsatz der Algebra besagt: Für ein Polynom $p(\cdot)$ vom Grad n hat die Gleichung $p(z) = w$ für jedes w genau n Lösungen (Vielfachheiten berücksichtigt). Erste Beweisskizze: Umlaufzahlen. Zweite Beweisskizze: Polynomiale Abbildungen sind offene Abbildungen und die stetige Funktion $|p(z)|$ nimmt ihr Minimum an. Gauss zu einem weiteren Beweis. Didaktische Thesen zur Beziehung von Algebra und Geometrie.

Historischer Anhang : Irrationalzahlen und Geometrie

Warum erscheinen Arithmetik und Geometrie traditionell (seit den alten Griechen) getrennt?

Geometrisierung des Zahlbereichs \mathbb{C} bei Gauss.

Arithmetisierung der Analysis im späten 19. Jahrhundert (Dedekind, Cantor).

III. Metrik, Norm, Konvexität

Eine wichtige Methode, einer abstrakten Menge S Struktur zu geben, besteht darin, dass man den Punktpaaren einen Abstand $d(P, Q)$ zuordnet. Eine Abstandsfunktion $d(\cdot, \cdot)$ ist eine nichtnegative Funktion $d(\cdot, \cdot)$, welche die sog. Dreiecksungleichung erfüllt. Auf reellen und komplexen affinen Räumen interessiert man sich vor allem für solche Abstandsfunktionen, die durch eine Längenmessung (oder spezieller durch eine Norm) im Raum der Verschiebungsvektoren gegeben sind: $d(P, Q) = \|PQ\|$.

Metrik und Norm stehen am Anfang der Analysis. In unserer elementargeometrisch orientierten Anfängerveranstaltungen soll jedoch die Infinitesimalrechnung nicht ganz ausgeklammert bleiben. Die reellen Zahlen und die Ungleichungen werden gebraucht. Im Folgenden bedienen uns auch (ohne Beweis) der Ordnungsvollständigkeit von (\mathbb{R}, \leq) : Jede nach oben beschränkte Teilmenge von \mathbb{R} besitzt in \mathbb{R} eine kleinste obere Schranke („Supremum“).

Damit ist das Feld offen für den Begriff der konvexen Funktion. Das punktweise Supremum von affinen Funktionen ist eine konvexe Funktion. Wir befassen uns allgemein mit den konvexen Funktionen auf endlichdimensionalen reellaffinen Räumen sowie mit Normen auf einem endlichdimensionalen (reellen oder komplexen) Vektorräumen. Bei der Legendre-Dualität tritt die Vektorraumdualität eindrucksvoll in Erscheinung.

III.1 Dreiecksungleichung und Subadditivität

Semimetriken. Der Begriff ist auch in der diskreten Mathematik bedeutungsvoll. Ein Beispiel ist die Hamming-Distanz auf $\{0, 1\}^N$.

Jede monotone subadditive Funktion auf dem \mathbb{R}^n macht aus einem n -Tupel von Semimetriken eine Semimetrik. Eine Vorform der Minkowski-Ungleichung liefert ein Beispiel.

Metrische Räume sind ein passender Kontext für die elementaren topologischen Begriffe: Konvergenz, offene und abgeschlossene Mengen, stetige und unterhalbstetige Funktionen.

III.2 Normierte Vektorräume und stetige Linearformen

Seminormen auf (möglicherweise unendlichdimensionalen) Vektorräumen. Die 2-Norm auf dem Raum der trigonometrischen Polynome. Stetige Linearformen; die duale Norm. Die p -Norm auf dem Raum der finiten Folgen. Die q -Norm ist die duale Norm.

III.3 Konvexe Mengen und konvexe Funktionen

Konvexe Mengen in einem endlichdimensionalen reellaffinen Raum. Distanzfunktionen. (‘Minkowski-Funktionale’). Trennende Hyperebenen. Die Stützfunktion einer konvexen Menge. Abgeschlossene konvexe Mengen und unterhalbstetige konvexe Funktionen. Der Epigraph einer konvexen Funktion ist genau dann eine abgeschlossene konvexe Menge, wenn die Funktion unterhalbstetig ist. Jede unterhalbstetige konvexe Funktion das Supremum ihrer affinen Minoranten ist. Legendre-Transformation. Zueinander duale konvexe Kegel. Die Jensen’sche Ungleichung.

IV. Inneres Produkt, Orthogonalität

Wir befassen uns mit normierten Vektorräumen, bei welchen die Norm auf ein inneres Produkt gegründet ist; man nennt sie Prä-Hilberträume oder im endlichdimensionalen reellen Fall euklidische Räume. Die Konstruktion, die aus einem Prä-Hilbertraum einen Hilbertraum macht ('Vervollständigung'), überlassen wir der Einführung in die Analysis. Hier diskutieren wir nur solche Aspekte des inneren Produkts, die keinen wesentlichen Gebrauch von der Vollständigkeit machen.

IV.1 Parallelogrammgleichung und Polarisierung

Mit der sog. Polarisierungskonstruktion beweisen wir den berühmten Satz von Jordan und v. Neumann: Ein normierter Vektorraum ist genau dann ein Prä-Hilbertraum, wenn die Norm die sog. Parallelogrammgleichung erfüllt.

$$\|v + w\|^2 + \|v - w\|^2 = 2 \cdot \|v\|^2 + 2 \cdot \|w\|^2.$$

In solchen Räumen gibt es den Begriff des Orthonormalsystems und die Orthogonalisierungstechnik nach Gram-Schmidt. Wir skizzieren die Notation von P. Dirac für die bra- und ket-Vektoren.

Hinweis zum Aufbau: Das Instrumentarium, welches wir bis hierher entwickelt haben, liefert eine axiomatische algebraische Begründung für den (in der Schulmathematik naiv hingenommenen) Begriff des euklidischen Anschauungsraums und gleichzeitig eine axiomatische Begründung für den Zustandsraum eines quantenmechanischen Systems.

Die beiden folgenden Unterabschnitte verbinden die geometrische Struktur des (Prä-) Hilbertraums mit der Matrizenrechnung. Die weiteren Unterabschnitte skizzieren 'Anwendungsfälle' dieser geometrischen Struktur.

IV.2 Hauptachsen, Orthogonale Zerlegungen, Spektralsatz

Auf Vektorräumen mit Hilbertraumstruktur gibt der Begriff der Sesquilinearform Anlass zu einer höchst bedeutsamen Theorie; es gilt das Theorem von den Hauptachsen. Wir skizzieren einen Beweis mit dem Maximierungsargument nach R. Courant. Wir folgern die simultane Diagonalisierbarkeit zweier hermitescher Formen, wo die eine positiv definit ist; und wir folgern die unitäre Diagonalisierbarkeit einer hermiteschen Matrix.

Eine angemessene Beschreibung der Resultate benötigt den Begriff der verträglichen direkten Zerlegung eines \mathbb{K} -Vektorraums und den Begriff der orthogonalen Projektion eines Hilbertraums.

Wir bemerken auch die Entsprechung zwischen den hermiteschen Sesquilinearformen und den selbstadjungierten Operatoren. Der Satz von den Hauptachsen entspricht dem Satz von der Spektralzerlegung eines selbstadjungierten Operators. Eine Matrix ist genau dann unitär diagonalisierbar, wenn sie 'normal' ist in dem Sinn, dass sie mit der adjungierten Matrix kommutiert.

IV.3 Einparametrische Matrixgruppen

Der infinitesimale Operator einer einparametrischen Matrixgruppe. Gleichförmige Drehun-

gen im Anschauungsraum. Einparametrische Gruppen von 2×2 -Matrizen. Einfache und doppelte Nullstellen des charakteristischen Polynoms. In einem Anhang stellen wir zusammen, wie die Matrizen recht verschiedene Rollen spielen können.

IV.4 Gekoppelte Oszillatoren

Der einfache harmonische Oszillator.

$$\begin{aligned} m\ddot{x} + r \cdot \dot{x} + k \cdot x &= f(t) \\ L \cdot \ddot{Q} + R\dot{Q} + \frac{1}{C} \cdot Q &= U(t) . \end{aligned}$$

Die allgemeine Lösung der homogenen Gleichung (im schwach gedämpften Fall). Zwei schwach gekoppelte Oszillatoren ohne Reibung und ohne anregende Kraft. Die *normal modes* im symmetrischen Fall. Die Lösungen sind fastperiodische Funktionen. ‘Resonanz’. Drei Oszillatoren in symmetrischen Anordnung. Die *normal modes* für n identische Oszillatoren in ringförmiger Anordnung. Stehende und laufende Wellen auf einem Ring.

IV.5 Die schwingende Saite

Die Gleichung der schwingenden Saite und die Gleichung der schwingenden Membrane. Separationsansatz für die Wellengleichung. Ebene Wellen. Intuitives zur Wellenbewegung: Wellenflächen, Ausbreitungsgeschwindigkeit, Wellenzahl, Frequenz, Phasengeschwindigkeit. Eigenschwingungen. Historisches zur schwingenden Saite und zur Fourier-Analyse.

IV.6 Diskrete Fourier-Transformation

Die DFT als Übergang zu neuen orthonormalen Koordinaten. Die DFT als Approximation der üblichen kontinuierlichen Fourier-Transformationen. Die Schnelle Fourier-Transformation FFT. Periodogramme.

IV.7 Drehungen im Minkowskiraum

Wir befassen uns eingehend mit dem euklidischen Anschauungsraum und mit dem Minkowskiraum der speziellen Relativitätstheorie. Wir entwickeln die sog. Spinordarstellung. Die Elemente w des vierdimensionalen Minkowskiraums werden durch die hermiteschen 2×2 -Matrizen dargestellt. Die ‘Drehungen’ (das sind die eigentlichen Lorentztransformationen) werden dann durch die Elemente A der speziellen linearen Gruppe $SL(2, \mathbb{C})$ beschrieben, $w \mapsto AwA^*$. Die Gruppe $SL(2, \mathbb{C})$ ist eine zweifache Überlagerung der eigentlichen Lorentzgruppe; diese ist damit in überschaubarer Weise isomorph zur Gruppe aller Möbiustransformationen.

Die Quaternionen werden herangezogen, um die Drehungen des 3-dimensionalen euklidischen Anschauungsraums zu studieren (Drehachse und Drehwinkel).

IV.8 Zerlegung einer Gruppendarstellung

Die Idee der verträglichen direkten Zerlegung eines Vektorraums, die wir bereits beim Spektralsatz kennengelernt haben, wird in einem weiteren bedeutsamen Fall ausgeführt: Jede Darstellung einer endlichen Gruppe durch Endomorphismen eines Vektorraums lässt sich in irreduzible Darstellungen zerlegen. Die etwas gröbere ‘kanonische’ direkte Zerlegung ist eindeutig bestimmt. Die Darstellungen sind durch ihre Spurfunktionen eindeutig bestimmt.

IV.9 Spurfunktionen, die Charakertafel einer endlichen Gruppe G

Die Theorie der Spurfunktionen auf einer endlichen Gruppe hat Ähnlichkeiten mit der Theorie der Fourier-Transformation. Die Analysis ist einfacher, weil es um Funktionen auf einer endlichen Menge geht. Die Algebra jedoch wird etwas komplizierter, sobald man auch nichtkommutative Gruppen in Betracht zieht; die Partition der Gruppe in Konjugationsklassen spielt eine wesentliche Rolle. Für einige spezielle Gruppen G gewinnen wir einen vollständigen Überblick über die Schar der irreduziblen Darstellungen ('Charaktertafel').

IV.10 Unitäre und hermitische Matrizen

Wir stellen nochmals zusammen, wo uns die unitären Matrizen schon früher begegnet sind. Die unitäre Diagonalisierung einer hermitischen 2×2 -Matrix wird explizit durchgeführt. Darauf aufbauend entwickeln wir Jacobi's Iterationsverfahren zur Diagonalisierung einer hermitischen $n \times n$ -Matrix. Gute Dienste leistet dabei die Hilbert-Schmidt-Norm.

V. Lineare Algebra, Gleichungssysteme, Dualität

Die 'Matrizenrechnung' hat für viele Fächer einen Platz in den Anfängervorlesungen; denn lineare Gleichungssysteme fallen in allen quantitativ arbeitenden Wissenschaften an. Andererseits wird die 'lineare Algebra' von den Mathematikern als der erste Schritt der grundlagenorientierten Mathematikausbildung geschätzt; für den Mathematiker ist nämlich der Begriff des Vektorraums zusammen mit dem Begriff der Gruppe der Grundbaustein aller Algebra und Geometrie.

Die Ziele der mathematischen Grundlegung sind selbstverständlich verschieden von den Zielen des Rechnens mit Matrizen. Nach all den Ansätzen in den Teilen I bis IV dürfen wir hoffen, dass sich hier im Teil V ein grundsätzliches Verständnis für Vektorräume verbindet mit einem Einblick in die Mechanismen der 'Matrizenrechnung'; ein solides Verständnis für das Prinzip der Linearität braucht u. E. ein gewisses Maß von Erfahrungen mit anschaulichen mathematischen Objekten.

Ein wesentliches Anliegen bei unserer Behandlung der Theorie der linearen Gleichungssysteme (oder allgemeiner gesagt, zur Theorie der linearen Abhängigkeit) ist die Vektorraumdualität. Ein lehrreiches elementares Vehikel scheint uns die Idee des vollständig ausgetauschten Tableaus zu sein; wir sehen darin nicht nur eine Alternative zu der (in Schulbüchern favorisierten) Eliminationsmethode zur 'Lösung' linearer Gleichungssysteme. Die Verbindung zwischen der Tableau-Methode mit der Matrizenrechnung wird (möglicherweise übertrieben ausführlich) dargelegt; der Leser mag sie überspringen. Die Unterabschnitte V.5 und IV.6 zeigen die Dualität in Anwendungen, die auf eine Hilbertraumnorm Bezug nehmen.

V.1 Vektorräume: Basis und Dimension

Vektorraumaxiome. Teilvektorräume. Die lineare Hülle einer Schar von Vektoren $\text{span}\{v_\alpha : \alpha \in I\}$. Lineare Unabhängigkeit. Eine Basis ist ein linear unabhängiges System maximaler Länge. Eine Basis ist ein aufspannendes System minimaler Länge. Beweis des Austauschlemmas von E. Steinitz. Linearformen. Der Dualraum. Duale Basen. Die

natürliche Paarung.

V.2 Lineare Gleichungssysteme

Ein lineares Gleichungssystem führt auf Vektorräume und auf affine Räume. Die allgemeine Lösung des inhomogenen Systems $Ax = b$ ergibt sich, wenn man zu einer speziellen Lösung die allgemeine Lösung der homogenen Gleichung dazuaddiert.

Ein lineares Gleichungssystem definiert eine lineare Abbildung $\varphi : x \mapsto Ax$. Wenn A den Spaltenrang r hat, dann hat das Bild im φ die Dimension r . Wenn A den Zeilenrang r hat, dann hat der Kern $\ker \varphi = \{x : Ax = 0\}$ die Dimension $n - r$. Zeilenrang = Spaltenrang. Wenn A den Rang r hat, dann gibt es eine nichtsinguläre Teilmatrix von der Größe $r \times r$.

V.3 Tableaus; Vollständiger Austausch

Zeilentableaus zur „Lösung“ linearer Gleichungssysteme. Spaltentableaus zur Konstruktion einer Basis des Spaltenraums. (Austauschlemma von Steinitz). Vollständiger Austausch: der Kernblock, der Bildblock und der Inversblock. Weitere Anwendungen des vollständigen Austausches.

V.4 Eliminationsmethode; LU-Faktorisierung

Das Gauss'sche Eliminationsverfahren. Zwei Zahlenbeispiele. Zeilenoperationen sind Multiplikationen von links. Untere und obere Dreiecksmatrizen. Die LU-Zerlegung $A = L^{-1} \cdot D \cdot U$ für eine nichtsinguläre Matrix A mit geeignet angeordneten Zeilen und Spalten. Inhomogene Systeme. Systeme mit unbestimmter rechter Seite. Zahlenbeispiel. Hinweis auf die Minimierung einer quadratischen Funktion.

V.5 Matrizen vom Rang r , Singulärwertzerlegung

Eine $m \times n$ Matrix A vom Rang r kann man als Summe von r Matrizen vom Rang 1 darstellen. Anders gesagt: Es gibt Faktorisierungen $A = B \cdot C$, wo B das Format $m \times r$ hat und C das Format $r \times n$; die Spalten von B bilden eine Basis des Spaltenraums von A , die Zeilen von C bilden eine Basis des Zeilenraums von A . Wir diskutieren insbesondere die Faktorisierung, die zum vollständigen Austausch gehört.

Wir betrachten dann eine komplexe $m \times n$ -Matrix A als Abbildung des Hilbertraums der n -Spalten in den Hilbertraum der m -Spalten. Wenn A den Rang r hat, dann gibt es Zahlen $\lambda_1, \dots, \lambda_r$, ein orthonormiertes r -Tupel von n -Spalten v_1, \dots, v_r und ein orthonormiertes r -Tupel von m -Spalten w_1, \dots, w_r , sodass

$$A \cdot v_k = \lambda_k \cdot w_k \quad \text{für } k = 1, \dots, r; \quad A \cdot v = 0, \quad \text{wenn } v \perp \{v_k\}.$$

Wenn wir die m -Spalten w_k zu einer Matrix Q zusammenstellen und die n -Spalten v_k zu einer Matrix P^* , dann erhalten wir die Faktorisierung

$$A = Q \cdot \Lambda \cdot P \quad \text{mit } \Lambda \text{ Diagonalmatrix und} \quad Q^* Q = I_r = P P^*.$$

Der Beweis benützt wieder Courant's Maximierungsargument. Die unitäre Diagonalisierung einer normalen Matrix A ergibt sich als Corollar.

V.6 Kleinste Quadrate und lineare Schätzung

Die Gleichung $B \cdot x = y$ besitzt eine Lösung nur für die y im Spaltenraum von B .

Es stellt sich die Aufgabe: Man finde für ein beliebiges y solche Spalten \hat{y} und \hat{x} , dass

$$B \cdot \hat{x} = \hat{y} , \quad \|y - \hat{y}\| = \min .$$

Die Methode der kleinsten Quadrate löst diese Aufgabe in dem Fall, wo $\|\cdot\|$ eine Hilbertraumnorm im Spaltenraum ist, $\|y\|^2 = y^* Q y$ mit Q positiv definit. Der Schlüssel ist eine Projektionsmatrix P , die aus einer ‘Normalgleichung’ gewonnen wird. Als Beispiel diskutieren wir die Bestimmung einer Ausgleichsgeraden.

Schliesslich entwickeln wir in aller Ausführlichkeit die duale Fragestellung als ein Thema der linearen Statistik.

Anhang : Mengen und Relationen

Die Mengendefinition von G. Cantor (1895). Mengenalgebra und Boole’sche Algebra. Strukturierte Mengen allgemein. Präordnungen. Äquivalenzrelationen. Die Partition nach Äquivalenzklassen. Anwendung auf Gruppen: Nebenklassen. Gruppenhomomorphismus. Faktorgruppen. Quotientenräume zu einem Vektorraum.

I Rechengrößen

I.1 Von den Zahlen zu den Matrizen

In der Schule gibt man sich (zurecht) viel Mühe mit den Zahlbereichserweiterungen

$$\mathbb{N} \subseteq \frac{\mathbb{Q}_+}{\mathbb{Z}} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Wir werden die Ideen, die den Erweiterungsschritten zugrunde liegen, bei passenden Gelegenheiten kommentieren. Die Erweiterung $\mathbb{Z} \subseteq \mathbb{Q}$ heißt das Bruchrechnen – davon wird in I.3. die Rede sein. Die Erweiterung $\mathbb{Q} \subseteq \mathbb{R}$ heißt Vervollständigung; wir überlassen sie der Vorlesung über Analysis. Die Erweiterung $\mathbb{R} \subseteq \mathbb{C}$ ist das Thema der zweiten Vorlesung I.2. Es sollte schnell eine Sache der Gewohnheit werden, mit den komplexen Zahlen genauso sicher umzugehen wie mit den (in der Schule eingeübten) reellen Zahlen. In I.4 wird dieses Rechnen (mit der nötigen Vorsicht) auf die Matrizen übertragen.

Zahlen sind ein sehr spezieller Typ von Rechengrößen. Manche der bekannten Rechenregeln gelten auch in allgemeineren Bereichen von Rechengrößen. Wir denken insbesondere an die

Assoziativgesetze

$$\begin{aligned}(a + b) + c &= a + (b + c) \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c)\end{aligned}$$

Distributivgesetze

$$\begin{aligned}(a + b) \cdot c &= a \cdot c + b \cdot c \\ a \cdot (b + c) &= a \cdot b + a \cdot c\end{aligned}$$

Kommutativgesetze

$$\begin{aligned}a + b &= b + a \\ a \cdot b &= b \cdot a\end{aligned}$$

Nach den neutralen Elementen (0 für die Addition und 1 für die Multiplikation) werden wir später fragen. Auch die Fragen nach der additiven Inversen ($-a$ zu a) und nach der multiplikativen Inversen (a^{-1} zu a) werden wir erst später aufnehmen.

Eine konkrete und äußerst wichtige Art von Rechengrößen sind die Matrizen, ganz besonders die reellen und die komplexen Matrizen.

Definition

I und J seien endliche Indexmengen; $|I| = m$, $|J| = n$.

Eine $I \times J$ -Matrix A ist dadurch gegeben, dass jedem Paar (i, j) eine Zahl a_{ij} (oder a_j^i) zugeordnet ist. a_{ij} heißt der Eintrag in der i -ten Zeile und der j -ten Spalte.

Meistens denkt man sich die Indexmengen durchnummeriert. Die Matrix A erscheint dann als ein rechteckiges Zahlenschema vom Format $m \times n$.

Operationen

Summe von Matrizen Zwei Matrizen vom gleichen Format $I \times J$ kann man addieren. Es werden die entsprechenden Einträge addiert:

$$A + B = C \quad \text{mit} \quad a_{ij} + b_{ij} = c_{ij} \quad \text{für alle } (i, j) \in I \times J .$$

Matrizenprodukt Wenn A eine $I \times J$ -Matrix ist und B eine $J \times K$ -Matrix, dann kann man das Matrizenprodukt bilden. $C = A \cdot B$ ist die $I \times K$ -Matrix mit den Einträgen

$$c_{ik} = \sum_j a_{ij} \cdot b_{jk} \quad \text{oder} \quad c_k^i = \sum_j a_j^i \cdot b_k^j \quad \text{für alle } (i, k) \in I \times K .$$

Skalare Multiplikation Wenn α eine Zahl ist, dann bezeichnet $\alpha \cdot A$ die $I \times J$ -Matrix mit den Einträgen $\alpha \cdot a_{ij}$. Die $I \times J$ -Matrizen kann man somit **linear kombinieren**: Wenn α und β Zahlen sind und A, B $I \times J$ -Matrizen, dann ist $\alpha \cdot A + \beta \cdot B$ eine wohldefinierte $I \times J$ -Matrix.

Die Definition des Matrizenprodukts mag dem Anfänger merkwürdig erscheinen. Die Sinnhaftigkeit ergibt sich aus den zahlreichen „Anwendungen“, die wir nach und nach kennen lernen werden. Hier interessieren uns zunächst nur die Rechengesetze.

Satz

Für Matrizen gelten die Assoziativgesetze und die Distributivgesetze. Das Kommutativgesetz der Multiplikation ist nicht gültig. Es gilt also

$$\begin{aligned} (A + B) + C &= A + (B + C) \quad , \quad (A \cdot B) \cdot C &= A \cdot (B \cdot C) \\ A \cdot (B + C) &= A \cdot B + A \cdot C \quad ; \quad (A + B) \cdot C &= A \cdot C + B \cdot C \quad , \end{aligned}$$

falls die Formate stimmen.

Die Gleichungen sind so zu verstehen: Wenn die linke (oder die rechte) Seite definiert ist, dann ist auch die rechte (bzw. die linke) Seite eine wohldefinierte Matrix und das Ergebnis ist auf beiden Seiten dasselbe. Die Beweise sind einfach: Das erste Distributivgesetz besagt

$$\sum_j a_{ij} \cdot (b_{jk} + c_{jk}) = \sum_j a_{ij} \cdot b_{jk} + \sum_j a_{ij} \cdot c_{jk} \quad \text{für alle } (i, k) .$$

Das Assoziativgesetz der Multiplikation besagt, dass bei der Berechnung der Einträge der Matrix $D = A \cdot B \cdot C$ nicht auf die Klammerung ankommt. Jedes $d_{i\ell}$ ergibt sich als Doppelsumme.

$$d_{i\ell} = \sum_{j,k} a_{ij} \cdot b_{jk} \cdot c_{k\ell} \quad \text{für alle } (i, \ell) \in I \times L .$$

Quadratische Matrizen

Eine $I \times I$ -Matrix heißt eine quadratische Matrix.

Wenn A und B $I \times I$ -Matrizen sind, dann sind auch $A + B$ und $A \cdot B$ wohldefinierte $I \times I$ -Matrizen. Im Bereich der $I \times I$ -Matrizen sind also die Linearkombination und die Multiplikation unbeschränkt definiert.

Die $I \times I$ -Matrix, deren Einträge allesamt $= 0$ sind, heißt die **$I \times I$ -Nullmatrix**. Die Nullmatrix ist das neutrale Element für die Addition. Die $I \times I$ -Matrix mit Einsen in der Diagonale und Nullen außerhalb der Diagonale heißt die **$I \times I$ -Einheitsmatrix**; sie wird hier mit $E_{I \times I}$ bezeichnet.

(Die gebräuchlichere Bezeichnung ist I oder Id für ‘Identitäts(matrix)’; hier benötigen wir aber (vorerst) die Buchstaben I, J, K, \dots als Namen für Indexmengen für Scharen oder ‘Familien’; das können Familien von Spalten sein, oder Familien von Zeilen sein, später auch für Scharen von Vektoren oder Scharen von Linearformen.)

Die Einheitsmatrix ist das neutrale Element für die Multiplikation. Genauer gesagt:

Satz

$E = E_{I \times I}$ ist die einzige Matrix mit $E \cdot A = A$ für alle $I \times I$ -Matrizen A .
Und $E_{I \times I}$ ist auch die einzige Matrix mit $A \cdot E = A$ für alle A .

Der Beweis ergibt sich leicht aus der folgenden allgemeineren Betrachtung über die Wirkung einer Matrix von links (oder rechts).

Zeilen- und Spaltenoperationen

Eine $I \times J$ -Matrix kann man als ein I -Tupel von J -Zeilen auffassen oder als ein J -Tupel von I -Spalten. Multiplikation von links bedeutet Zeilenoperation; Multiplikation von rechts bedeutet Spaltenoperation.— Genauer:

Wenn man die $I \times J$ -Matrix A von links multipliziert, dann sind die Zeilen von $B \cdot A$ Linearkombinationen der Zeilen von A . Die Spalten von $A \cdot C$ sind Linearkombinationen der Spalten von A .

Beispiel Zuerst werden die Zeilen von A linear kombiniert, dann die Spalten

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} a_{11} + a_{21} & a_{12} + a_{22} & a_{13} + a_{23} \\ a_{11} + 2a_{21} & a_{12} + 2a_{22} & a_{13} + 2a_{23} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{13} & a_{11} - a_{12} \\ a_{21} + a_{23} & a_{21} - a_{22} \end{pmatrix}$$

Diagonalmatrizen

Eine $I \times I$ -Matrix heißt eine Diagonalmatrix, wenn die Einträge außerhalb der Diagonalen Nullen sind. Mit den Einträgen d_i in der Diagonalen gilt

$D \cdot A$ entsteht aus A indem man die Zeilen von A mit den Zahlen d_i multipliziert;

$A \cdot D$ entsteht aus A , indem man die Spalten von A mit den Zahlen d_i multipliziert.

Zwei zentrale Sätze der Matrizenrechnung

Theorem 1

Seien I und J Indexmengen derselben Mächtigkeit $n : |I| = n = |J|$.

Zur $I \times J$ -Matrix A gibt es genau dann eine Linksinverse, wenn es eine Rechtsinverse gibt, und es handelt sich um dieselbe Matrix B vom Format $J \times I$.

$$A \cdot B = E_{I \times I}, \quad B \cdot A = E_{J \times J}.$$

Kurz gesagt: Eine Linksinverse ist automatisch eine Rechtsinverse (und umgekehrt).

Die Schwierigkeit des Beweises liegt bei der Existenz; die endliche Dimension spielt eine wesentliche Rolle. Einfach ist der Beweis der Eindeutigkeit.

$$A \cdot B = E_{I \times I}, \quad C \cdot A = E_{J \times J} \implies B = (C \cdot A) \cdot B = C \cdot (A \cdot B) = C.$$

Hinweis

Die Theorie der linearen Operatoren (die wir im 3. Semester skizzieren werden) ist eine Verallgemeinerung der Matrizenrechnung. In dieser Theorie, im sog. unendlichdimensionalen Fall gibt es Operatoren A , welche zwar eine Linksinverse besitzen, aber keine Rechtsinverse (oder umgekehrt).

Definition

Eine $I \times J$ -Matrix A heißt **nichtsingulär**, oder invertierbar, wenn sie eine Inverse B existiert. Die (eindeutig bestimmte !) **Inverse** der nichtsingulären Matrix A wird mit A^{-1} bezeichnet.

Satz Sei $C = A \cdot B$ mit A und B invertierbar. Dann ist auch C invertierbar, und es gilt

$$C^{-1} = (A \cdot B)^{-1} = B^{-1} A^{-1}.$$

Der Beweis ist eine einfache Anwendung des Assoziativgesetzes:

$$C \cdot (B^{-1} A^{-1}) = A \cdot (B B^{-1}) \cdot A^{-1} = E; \quad (B^{-1} A^{-1}) \cdot C = B^{-1} \cdot (A^{-1} A) \cdot B = E.$$

Theorem 2

Auf der Menge aller $I \times I$ -Matrizen existiert eine multiplikative Funktion, welche genau dann $\neq 0$ ist, wenn die Matrix nichtsingulär ist. Diese Funktion heisst die **Determinantenfunktion**; sie wird mit $\det(\cdot)$ bezeichnet.

$$\begin{aligned} \det(A \cdot B) &= (\det A) \cdot (\det B) \\ \det(A) \neq 0 &\Leftrightarrow A \text{ ist nichtsingulär.} \end{aligned}$$

Die Beweise der Theoreme verschieben wir auf später.

Komplexe 2×2 -Matrizen

Wir werden die Aussagen der Theoreme vorerst nur für die 2×2 -Matrizen verwenden; und für diesen Fall werden sie durch die folgenden einfachen Rechnungen bestätigt.

Definition

Für 2×2 -Matrizen $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ definiert man

$$\det A = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc .$$

Satz

Wenn $\det A \neq 0$, dann gilt

$$A^{-1} = \frac{1}{\det A} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Hinweis :

Es gilt als mathematisch unsauber, wenn man sagt: Die Determinante einer 2×2 -Matrix verschwindet genau dann, wenn die Zeilen (und damit auch die Spalten) proportional sind. Beim Reden über Proportionalität spielen die Nullen eine etwas unglückliche Rolle. Beispiele:

$$\det \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} = 0 , \quad \det \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} = 0 .$$

Wer würde im ersten Fall sagen, dass die Spalten „zueinander proportional“ sind? Die korrekte Ausdrucksweise, die sich auch für $n \times n$ -Matrizen für $n > 2$ bewährt, ist die:

Die Determinante verschwindet genau dann, wenn die Zeilen **linear abhängig** sind, d.h. wenn es eine Zeile gibt, welche sich als Linearkombination der übrigen Zeilen darstellen läßt. In diesem Fall gibt es auch eine Spalte, die sich als Linearkombination der übrigen Spalten darstellen läßt. (Dies ist ein mathematischer Satz, den wir hier nur für 2×2 -Matrizen bestätigen können.) Im Falle, wo alle Einträge $\neq 0$ sind, haben wir in der Tat

$$ad - bc = 0 \iff \frac{a}{b} = \frac{c}{d} \iff \frac{a}{c} = \frac{b}{d} .$$

Es ist allgemein richtig und sauber ausgedrückt, wenn man sagt: Die Determinante einer $n \times n$ -Matrix verschwindet genau dann, wenn die Zeilen (oder auch die Spalten) linear abhängig sind, wenn es also eine Zeile (bzw. Spalte) gibt, die sich als Linearkombination der übrigen Zeilen (bzw. Spalten) darstellen lässt.

$$\det A = 0 \iff \exists \quad x \neq 0 : \quad A \cdot x = 0 \iff \exists \quad \xi \neq 0 : \quad \xi \cdot A = 0 .$$

Betrachten wir die Matrix $A - \lambda \cdot \text{Id}$, dann ergibt sich der

Satz

Genau dann existiert eine Spalte $x \neq 0$ mit $A \cdot x = \lambda \cdot x$, wenn λ die quadratische Gleichung $(a - \lambda)(d - \lambda) - bc = 0$ löst.

Eine solche Spalte x heisst eine Eigenspalte zum Eigenwert λ .

Multiplikativität der Determinantenfunktion für $n=2$

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \quad B = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \quad A \cdot B = \begin{pmatrix} a\alpha + c\beta & a\gamma + c\delta \\ b\alpha + d\beta & b\gamma + d\delta \end{pmatrix}$$

Durch Nachrechnen bestätigt man

$$(a\alpha + c\beta)(b\gamma + d\delta) - (a\gamma + c\delta)(b\alpha + d\beta) = (ad - bc) \cdot (\alpha\delta - \beta\gamma) .$$

Neben der Determinante $\det(A)$ ist auch noch die **Spur** $\text{trace}(A)$ eine besonders wichtige Kenngröße einer quadratischen Matrix A . Die Spur ist die Summe der Diagonalelemente.

Im Falle $n = 2$ ist zu bemerken

$$\det(A - \lambda \text{Id}) = \det \begin{pmatrix} a - \lambda & c \\ b & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - \text{trace}(A) \cdot \lambda + \det(A) .$$

Ein merkwürdiger Satz

Sei A eine 2×2 -Matrix mit der Spur t und der Determinante d . Es gilt dann

$$A^2 - t \cdot A + d \cdot E = 0 .$$

Beispiel

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} , \quad t = 5 , \quad d = -2$$

$$A^2 - 5 \cdot A - 2E = \begin{pmatrix} 7 & 15 \\ 10 & 12 \end{pmatrix} - \begin{pmatrix} 5 & 15 \\ 10 & 20 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} .$$

Der Satz kann leicht durch eine direkte Rechnung bestätigt werden. Seine Verallgemeinerung auf $n \times n$ -Matrizen heißt der Satz von Cayley-Hamilton. Wir werden diesen Satz beweisen, wenn wir später das charakteristische Polynom einer quadratischen Matrix studieren.

Aufgaben zu I.1

Aufgabe I.1.1 :

Eine quadratische Matrix P heisst eine Permutationsmatrix, wenn sie nur Einträge 1 und 0 hat, wobei in jeder Zeile und in jeder Spalte genau eine 1 steht.

1. Machen Sie eine vollständige Liste aller Permutationsmatrizen vom Format 3×3 .
2. Zeigen Sie: Wenn P_1 und P_2 Permutationsmatrizen (sagen wir vom Format $J \times J$) sind, dann ist auch das Produkt eine Permutationsmatrix.

Aufgabe I.1.2 :

Eine $I \times J$ -Matrix P heisst eine stochastische Matrix, wenn sie nichtnegative Einträge hat und alle Zeilensummen $= 1$ sind

$$p_{ij} \geq 0 \quad \text{für alle } i, j, \quad \sum_j p_{ij} = 1 \quad \text{für alle } i.$$

Sei Q eine stochastische $J \times K$ -Matrix und sei $R = P \cdot Q$. ($r_{ik} = \sum_j p_{ij} q_{jk}$).
Zeigen Sie, dass R eine stochastische Matrix ist.

Aufgabe I.1.3 :

- a) Berechnen Sie die Potenzen P^2, P^3, P^4, \dots der folgenden Matrix

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

- b) Berechnen Sie die Inverse $B = A^{-1}$ für die Matrix

$$A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \frac{1}{2}(P + P^2).$$

Hinweis zu b) : Zeigen Sie $2A^2 = A + E$ und multiplizieren Sie mit B .

Aufgabe I.1.4 : (Spur einer Matrix)

Die Spur (trace) einer quadratischen Matrix M ist die Summe der Diagonalelemente $\text{tr}(M)$. Seien A eine $I \times J$ -matrix und B eine $J \times I$ -Matrix.

Zeigen Sie

$$\text{tr}(A \cdot B) = \text{tr}(B \cdot A).$$

Aufgabe I.1.5 : (Transponierte Matrix)

Sei A eine reelle $I \times J$ -Matrix. $A = (a_{ij})$. Man definiert dazu die transponierte Matrix A^T . Dies ist die $J \times I$ -Matrix B mit den Einträgen $b_{ji} = a_{ij}$ (für $i \in I, j \in J$).

Seien M und N Matrizen, für welche $M \cdot N$ definiert ist (wo also die Formate passen).

Zeigen Sie

$$(M \cdot N)^T = N^T \cdot M^T.$$

Aufgabe I.1.6 : (Diagonalmatrizen)

A sei irgendeine 3×3 -Matrix.

1. Für die Diagonalmatrix D mit den Diagonalelementen $1, 3, -1$ gelte $A \cdot D = D \cdot A$. Zeigen Sie, dass A tatsächlich eine Diagonalmatrix ist!
2. Für die Diagonalmatrix \tilde{D} mit den Diagonalelementen $1, 3, 1$ gelte $A \cdot \tilde{D} = \tilde{D} \cdot A$. Was können Sie über die Matrix A sagen?
3. Es seien D und E Diagonalmatrizen mit nichtverschwindenden Diagonalelementen d_j bzw. e_j .
Zeigen Sie: Wenn eine nichtsinguläre Matrix W existiert, sodass $D \cdot W = W \cdot E$, dann gilt $E = D$. In diesem Falle zerfällt W in Blöcke gemäss den verschiedenen Diagonaleinträgen d_j .

I.2 Komplexe Zahlen; die Euler'sche Formel

Die komplexen Zahlen gewinnt man aus Paaren reeller Zahlen: $z = a + ib$ mit $a, b \in \mathbb{R}$. Seit Gauss veranschaulicht man die komplexen Zahlen als Punkte in der Gauss'schen Zahlenebene \mathbb{C} . Die Koordinaten heißen der Realteil und der Imaginärteil von z .

$$z = a + ib, \quad a = \Re z, \quad b = \Im z.$$

Verknüpfungen

Komplexe Zahlen kann man addieren und multiplizieren. Für $z = a + ib$, $w = c + id$ definiert man

$$\begin{aligned} z + w &= (a + c) + i(b + d) \\ z \cdot w &= (ac - bd) + i(ad + bc). \end{aligned}$$

Die Menge der Punkte mit Imaginärteil $= 0$ heißt die reelle Achse \mathbb{R} . Wenn man die eben definierten Verknüpfungen auf die Teilmenge \mathbb{R} einschränkt, erhält man die bekannte Addition (bzw. Multiplikation) reeller Zahlen; insofern besteht keine Verwechslungsgefahr, wenn man die Zeichen $+$ und \cdot auch für die Verknüpfungen in \mathbb{C} benützt. Die Null, das neutrale Element für die Addition, und die Eins, das neutrale Element für die Multiplikation, liegen auf der reellen Achse. Statt $0 + i \cdot 0$ schreibt man einfach 0, statt $1 + i \cdot 0$ schreibt man 1. Allgemeiner schreibt man $a + i \cdot 0 = a$.

$(\mathbb{C}, 0, 1, +, \cdot)$ ist somit eine **Erweiterung** von $(\mathbb{R}, 0, 1, +, \cdot)$.

Assoziativität und Distributivität

Die Addition und die Multiplikation komplexer Zahlen sind offenbar kommutative Verknüpfungen. Außerdem gelten die Assoziativ- und die Distributivgesetze

$$\begin{aligned} (z_1 + z_2) + z_3 &= z_1 + (z_2 + z_3); & (z_1 \cdot z_2) \cdot z_3 &= z_1 \cdot (z_2 \cdot z_3) \\ (z_1 + z_2) \cdot w &= z_1 \cdot w + z_2 \cdot w, & z \cdot (w_1 + w_2) &= z \cdot w_1 + z \cdot w_2. \end{aligned}$$

Dies kann man direkt nachrechnen. Es ist aber auch lehrreich, die Beweise mit Hilfe der Matrizenrechnung zu führen. Dies geht folgendermaßen: Wir identifizieren die komplexen Zahlen $z = a + ib$, $w = c + id$ mit speziellen reellen 2×2 -Matrizen, nämlich

$$z \longleftrightarrow \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad w \longleftrightarrow \begin{pmatrix} c & -d \\ d & c \end{pmatrix}.$$

Summe und Produkt sind wieder Matrizen dieser speziellen Gestalt; es gilt in der Tat:

$$z + w \longleftrightarrow \begin{pmatrix} a + c & -(b + d) \\ b + d & a + c \end{pmatrix}; \quad z \cdot w \longleftrightarrow \begin{pmatrix} ac - bd & -(bc + ad) \\ bc + ad & (ac - bd) \end{pmatrix}.$$

Wir wissen bereits, dass für die Addition und die Multiplikation quadratischer Matrizen die Assoziativ- und die Distributivgesetze gültig sind.

Konjugation und Normbildung

Wenn $z = a + ib$ ($a, b \in \mathbb{R}$) eine komplexe Zahl ist, dann heißt $a - ib$ die (komplex) konjugierte Zahl; sie wird mit \bar{z} (oder auch z^*) bezeichnet. Die Abbildung $z \mapsto \bar{z}$ heißt die **komplexe Konjugation** (in \mathbb{C}).

Die **Norm** der komplexen Zahl $z = a + ib$ ist die positive Zahl

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

.

Satz

Wenn z und w komplexe Zahlen sind, dann gilt

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

Beweis Die Norm macht die Gauss'sche Zahlenebene \mathbb{C} zu einem euklidischen Raum.

Die dritte Aussage ist die „Dreiecksungleichung“ in der euklidischen Ebene.

Weiter gilt $|z \cdot w|^2 = (z \cdot w) \cdot \overline{(z \cdot w)} = (z \cdot \bar{z}) \cdot (w \cdot \bar{w}) = |z|^2 \cdot |w|^2$.

Bemerke auch: $|z|^2$ ist die Determinante der (wie oben) zugeordneten Matrix.

Multiplikative Inverse

Zu jeder komplexen Zahl $\neq 0$ gibt es eine multiplikative Inverse; man nennt sie auch die reziproke Zahl und notiert sie

$$\frac{1}{z} \quad \text{oder} \quad z^{-1}.$$

Beweis

$$z = a + ib \Rightarrow \frac{1}{z} = \frac{a}{a^2 + b^2} + i \frac{-b}{a^2 + b^2} = \frac{1}{|z|^2} \cdot \bar{z}.$$

In der Tat gilt $z \cdot \frac{1}{|z|^2} \bar{z} = 1$.

Bemerke

Die Reziproke der konjugierten Zahl ist die Konjugierte der reziproken Zahl.

Die Norm der reziproken Zahl ist die Reziproke der Norm. In Formeln:

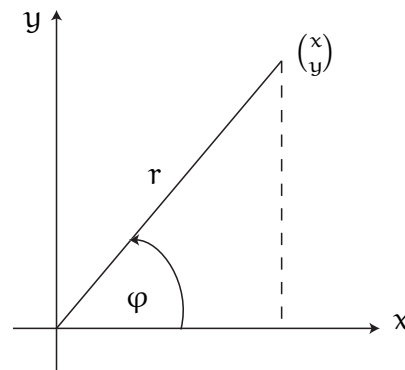
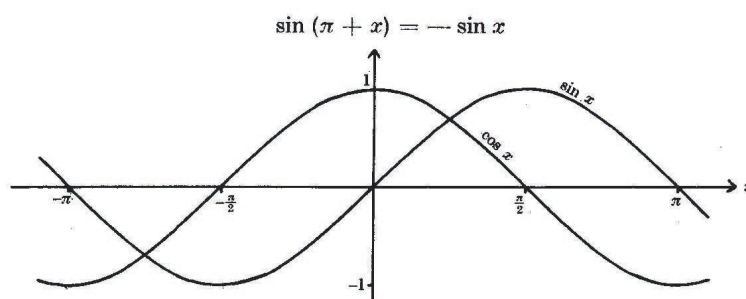
$$\frac{1}{\bar{z}} = \left(\frac{1}{z} \right)^* ; \quad \left| \frac{1}{z} \right| = \frac{1}{|z|}.$$

Polarkoordinaten

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

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

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



Satz

Für jede komplexe Zahl $z = a + ib \neq 0$ gibt es genau ein $\varphi \in \mathbb{R}/2\pi$, sodass

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

Die Rechtfertigung der Notation $e^{i\varphi}$ wird sich unten ergeben, wo wir die Exponentialfunktion im Komplexen diskutieren. Hinter der Notation steckt in Wirklichkeit eine tiefe mathematische Einsicht.

Hinweis

Feynman, Bd. I, Kap. 22 : „Algebra“ schließt mit dem Satz:

„Wir fassen zusammen: Die bemerkenswerteste Formel der Mathematik ist $e^{i\vartheta} = \cos \vartheta + i \sin \vartheta$. Das ist unser Juwel.“

Satz

Seien z und w komplexe Zahlen $\neq 0$

$$z = |z| \cdot e^{i\varphi} , \quad w = |w| \cdot e^{i\psi} .$$

Dann gilt

$$z \cdot w = |z \cdot w| \cdot e^{i(\varphi+\psi)} .$$

Beweis

$$\begin{aligned}
z \cdot w &= |z| \cdot (\cos \varphi + i \sin \varphi) \cdot |w| \cdot (\cos \psi + i \sin \psi) \\
&= |z \cdot w| [(\cos \varphi \cdot \cos \psi - \sin \varphi \cdot \sin \psi) + i(\sin \varphi \cos \psi + \cos \varphi \cdot \sin \psi)] \\
&= |z \cdot w| \cdot [\cos(\varphi + \psi) + i \sin(\varphi + \psi)] = |z \cdot w| \cdot e^{i(\varphi + \psi)}
\end{aligned}$$

(gemäß den Additionstheoremen für die trigonometrischen Funktionen)

Bemerke:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \cdot \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} = \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix}$$

Komplexe Amplituden in der Theorie der Wechselströme

Seien $\omega \in \mathbb{R}$, $t_0 \in \mathbb{R}$ und $|A| \in \mathbb{R}_+$. Die reellwertige Funktion

$$f(t) = |A| \cdot \cos(\omega(t - t_0)) \quad t \in \mathbb{R}$$

nennt man in der Theorie der Wechselspannungen und Wechselströme eine „reine Sinusschwingung“ mit der maximalen Amplitude $|A|$, der Kreisfrequenz ω und der Phasenverzögerung t_0 . (Wir bemerken: $\nu = \frac{\omega}{2\pi}$ heißt die Frequenz, $T = \frac{1}{\nu} = \frac{2\pi}{\omega}$ heißt die Periode.)

Wenn man reine Sinusschwingungen zur Kreisfrequenz ω überlagert, bekommt man reine Sinusschwingungen. Unser Beweis wird auch eine einfache Formel ergeben, wie die Phasenverzögerung und die maximale Amplitude der resultierenden Schwingung zu berechnen ist.

Wir arbeiten mit komplexen Zahlen. Es gilt

$$\begin{aligned}
f(t) &= \Re(|A| \cdot e^{i\omega(t-t_0)}) \\
&= \Re(A \cdot e^{i\omega t}) \quad \text{mit } A = |A|e^{-i\omega t_0}.
\end{aligned}$$

Seien $f_1(t)$ und $f_2(t)$ zwei reine Sinusschwingungen

$$f_1(t) = \Re(A_1 \cdot e^{i\omega t}), \quad f_2(t) = \Re(A_2 \cdot e^{i\omega t}).$$

Wir haben dann

$$f_1(t) + f_2(t) = \Re((A_1 + A_2)e^{i\omega t}).$$

Die komplexen Amplituden A_1 und A_2 werden einfach addiert.

Die Exponentialfunktion im Reellen

Im Schulunterricht wird die Exponentialfunktion gerne mit der Vorstellung von der kontinuierlichen Verzinsung motiviert.

Wenn sich ein Kapital K mit dem jährlichen Zinssatz α (z.B. $\alpha = 5\% = 0,05$) verzinst, dann hat man nach einem Jahr das Kapital $K \cdot (1 + \alpha)$. Wenn halbjährlich verzinst wird mit dem halbjährlichen Zinssatz $\frac{\alpha}{2}$, dann ergibt sich nach einem Jahr

$$K \cdot \left(1 + \frac{\alpha}{2}\right) \left(1 + \frac{\alpha}{2}\right).$$

Bei monatlicher Verzinsung zum monatlichen Zinssatz $\frac{\alpha}{12}$ hat man nach einem Jahr

$$K \cdot \left(1 + \frac{\alpha}{12}\right)^{12}.$$

Wenn das Jahr in n Einheiten eingeteilt wird mit dem Zinssatz $\frac{\alpha}{n}$ für jede Periode, dann hat man nach einem Jahr

$$K \cdot \left(1 + \frac{\alpha}{n}\right)^n.$$

Im Grenzwert $n \rightarrow \infty$ ergibt sich $K \cdot e^{\alpha}$ nach einem Jahr und $K \cdot e^{\alpha \cdot t}$ nach t Jahren.

Fazit : Das Kapital wächst exponentiell in der Zeit

$$K(t) = K(0) \cdot e^{\alpha \cdot t}.$$

Bemerke : Exponentielles Wachstum tritt bei vielen Gelegenheiten auf, immer dann nämlich, wenn der momentane Zuwachs proportional zum Bestand ist.

$$\begin{aligned} f'(t) &= \lim_{n \rightarrow \infty} n \left[f\left(t + \frac{1}{n}\right) - f(t) \right] = \alpha \cdot f(t) \\ f(t) &= f(0) \cdot e^{\alpha \cdot t}. \end{aligned}$$

Cauchy hat in der Tat den Satz **bewiesen**, dass jede differenzierbare Funktion $f(t)$ mit $f' = \alpha \cdot f$ ein Vielfaches der Funktion $e^{\alpha \cdot t}$ ist.

Die Exponentialfunktion auf der imaginären Achse

Wir studieren für reelles φ die Folge

$$z_n = \left(1 + \frac{1}{n}(i\varphi)\right)^n$$

z_n ist die n -te Potenz einer komplexen Zahl mit einem Betrag sehr nahe bei 1 und einem Winkel von der Größenordnung $\varphi_n \approx \frac{1}{n}\varphi$.

$$1 + \frac{1}{n}(i\varphi) = \sqrt{1 + \frac{1}{n^2}\varphi^2} \cdot e^{i\varphi_n},$$

wobei $\tan \varphi_n = \frac{1}{n}\varphi$, also $\varphi_n = \arctan\left(\frac{1}{n}\varphi\right)$.

$$z_n = \left(1 + \frac{1}{n^2}\varphi^2\right)^{n/2} \cdot e^{in \cdot \varphi_n}.$$

Der erste Faktor konvergiert nach 1. Außerdem gilt $n\varphi_n \rightarrow \varphi$; denn die Arcustangensfunktion hat im Nullpunkt die Ableitung = 1 (Bild!) Es gilt also

$$\cos(n\varphi_n) + i \sin(n\varphi_n) \longrightarrow \cos \varphi + i \sin \varphi.$$

Somit haben wir mit Schulkenntnissen „bewiesen“

$$\left(1 + \frac{1}{n}(i\varphi)\right)^n \longrightarrow \cos \varphi + i \sin \varphi = e^{i\varphi}.$$

Hinweis : In der Anfängervorlesung für Mathematiker betrachtet man die Grenzübergänge $n \rightarrow \infty$ noch etwas genauer. Für Physik-Studierende wollen wir das auf später verschieben. Hier bei den klassischen speziellen Funktionen dürfen wir so sorglos argumentieren wie die Leute im 18. Jahrhundert.

Die Exponentialfunktion über \mathbb{C}

Satz Für jedes komplexe z existiert der Limes

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n =: e^z.$$

Es gilt

$$e^z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \dots$$

Beweisskizze (im Stile des Entdeckers L. Euler, 1748)

Nach der Binomialformel gilt:

$$\left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} \left(\frac{z}{n}\right)^k \quad \text{mit} \quad \binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!}.$$

Für festes k und $n \rightarrow \infty$ haben wir

$$\binom{n}{k} \cdot \left(\frac{1}{n}\right)^k = \frac{1}{k!} \cdot \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right) \rightarrow \frac{1}{k!}.$$

Wir haben also

$$\left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!} z^k \cdot a_k^{(n)}$$

mit Koeffizienten $a_k^{(n)}$ zwischen 0 und 1, welche für $n \rightarrow \infty$ nach 1 konvergieren. Der Grenzwert für $n \rightarrow \infty$ ist die Exponentialreihe. Wenn man einen strengen Beweis wünscht, dann muss man nachweisen, dass man den Grenzübergang gliedweise durchführen kann. Man zeigt in der Tat den

Satz Die Konvergenz

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!} z^k = e^z$$

ist gleichmäßig in jedem beschränkten Teil von \mathbb{C} .

Hinweis

Die Rechnung läßt sich ausdehnen.

Sei A eine quadratische Matrix. Es gilt dann

$$\lim_{n \rightarrow \infty} \left(I + \frac{1}{n}A\right)^n = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots =: e^A = \exp(A).$$

So wirkt die „Exponentialabbildung“ also auch auf quadratische Matrizen. In der Quantenmechanik wirkt die Exponentialabbildung auf Operatoren : $U_t = \exp(itH)$, wo H ein Hamilton-Operator ist, $t \in \mathbb{R}$. Wir werden das im 3. Semester erörtern.

Multiplikativität

In der Schule lernt man, dass die Exponentialabbildung Summen (in \mathbb{R}) zu Produkten (in \mathbb{R}_+) macht

$$\exp(s + t) = \exp(s) \cdot \exp(t) \quad \text{für } s, t \in \mathbb{R} .$$

Noch bekannter ist diese „Funktionalgleichung“, wenn man sie für die Umkehrfunktion hinschreibt. Die Exponentialabbildung besitzt in der Tat auf \mathbb{R}_+ eine Umkehrabbildung. Sie heißt der natürliche Logarithmus. Es gilt

$$\ln(x \cdot y) = \ln x + \ln y \quad \text{für } x, y \in \mathbb{R}_+ .$$

Diese „Funktionalgleichung“ ist bekanntlich die Grundlage für das Rechnen mit dem Rechenschieber sowie für das Rechnen mit den Logarithmentafeln.

Es ist nun zu beachten, dass die Exponentialfunktion im Komplexen keine Umkehrabbildung besitzt. Zu jedem $w \in \mathbb{C}$ ($w \neq 0$) gibt es viele z mit $e^z = w$. (Die Menge der Lösungen ist allerdings überschaubar, die Lösungen unterscheiden sich nur um ein ganzzahliges Vielfaches von $2\pi i$; denn $\exp(2\pi i k) = 1$ für alle $k \in \mathbb{Z}$). Wir rechnen jedenfalls lieber mit der Exponentialfunktion als mit dem Logarithmus.

Satz

Wenn z und w komplexe Zahlen sind, dann gilt

$$\exp(z + w) = \exp(z) \cdot \exp(w) .$$

1. Beweisskizze

$$\left(1 + \frac{z}{n}\right)^n \cdot \left(1 + \frac{w}{n}\right)^n = \left[1 + \frac{z+w}{n} + \frac{zw}{n^2}\right]^n \approx \left(1 + \frac{z+w}{n}\right)^n .$$

2. Beweisskizze

Hier betrachten wir gleich den allgemeineren Fall der kommutierenden Matrizen.

Satz

Seien A und B kommutierende Matrizen, $A \cdot B = B \cdot A$.

Es gilt dann für alle $t \in \mathbb{R}$

$$\exp(t(A + B)) = \exp(t \cdot A) \cdot \exp(t \cdot B) .$$

Beweisskizze

Wir sammeln die Summanden mit demselben Faktor t^n im Produkt.

$$\begin{aligned}
 & \exp(t \cdot A) \cdot \exp(t \cdot B) = \\
 & = (I + t \cdot A + \frac{1}{2!}t^2A^2 + \frac{1}{3!}t^3A^3 + \dots)(I + t \cdot B + \frac{1}{2!}t^2B^2 + \frac{1}{3!}t^3B^3 + \dots) \\
 & = I + t(A + B) + \frac{1}{2!}t^2(A^2 + 2AB + B^2) + \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \dots .
 \end{aligned}$$

Im Falle kommutierender Matrizen haben wir die kombinatorisch zu beweisenden Binomialformeln

$$\begin{aligned}
 A^2 + 2AB + B^2 &= (A + B)^2 \\
 A^3 + 3A^2B + 3AB^2 + B^3 &= (A + B)^3 \\
 &\dots\dots\dots
 \end{aligned}$$

Daraus ergibt sich die Behauptung

$$\exp(t \cdot A) \cdot \exp(t \cdot A) = I + t(A + B) + \frac{1}{2!}t^2(A + B)^2 + \dots = \exp(t(A + B)) .$$

Aufgaben zu I.2

Aufgabe I.2.1 :

Berechnen Sie das Quadrat und die dritte Potenz von

$$z_0 = \frac{1}{2} \cdot \sqrt{3} + \frac{1}{2}i = r \cdot e^{i\varphi} \quad (\text{Zeichnung})$$

Was sind r und φ ?

Aufgabe I.2.2 : (Fünfte Einheitswurzeln)

- Die Gleichung $z^5 - 1 = 0$ hat 5 Lösungen w^0, w^1, w^2, w^3, w^4 (Zeichnung!).
- Es gibt reelle Zahlen $a < A$, sodass gilt

$$z^5 - 1 = (z - 1)(z^4 + z^3 + z^2 + z + 1) = (z - 1)(z^2 + az + 1)(z^2 + Az + 1)$$

Bestimmen Sie diese Zahlen a, A .

- Wir gewinnen daraus die 5-ten Einheitswurzeln w^j und die Zahlen $\cos \frac{2\pi}{5} = \cos 72^\circ$, $\sin \frac{2\pi}{5} = \sin 72^\circ$. Betrachten Sie nämlich

$$w_1 = -\frac{a}{2} + i\sqrt{1 - \frac{a^2}{4}}; \quad w_2 = -\frac{A}{2} + i\sqrt{1 - \frac{A^2}{4}}.$$

Zeigen Sie $w_2 = w_1^2$ und berechnen Sie $w_3 = w_1^3$; $w_4 = w_1^4$. Für jedes der w_k gilt $w_k^5 = 1$.

- Zeigen Sie (Bild!)

$$\begin{aligned} \cos 72^\circ &= -\frac{a}{2} = \frac{1}{4}(\sqrt{5} - 1); & \sin 72^\circ &= \frac{1}{4}\sqrt{10 + 2\sqrt{5}} \\ \cos 36^\circ &= +\frac{A}{2} = \frac{1}{4}(\sqrt{5} + 1); & \sin 36^\circ &= \frac{1}{4}\sqrt{10 - 2\sqrt{5}}. \end{aligned}$$

Aufgabe I.2.3 :

Wir multiplizieren komplexe 2×2 -Matrizen von einer speziellen Gestalt

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} e & f \\ ? & ? \end{pmatrix}$$

Berechnen Sie die Zahlen e und f sowie die beiden fehlenden Einträge in der Produktmatrix!

Aufgabe I.2.4 :

Bekanntlich definiert man e^z auch für komplexe Argumente z . Es gilt

$$e^{z+w} = e^z \cdot e^w \quad \text{für alle } z, w \in \mathbb{C} .$$

Man definiert dann für komplexes z

$$\cos z = \frac{1}{2} (e^{iz} + e^{-iz}) ; \quad \sin z = \frac{1}{2i} (e^{iz} - e^{-iz}) .$$

Zeigen Sie die Additionstheoreme

$$\begin{aligned} \cos(z + w) &= \cos z \cdot \cos w - \sin z \cdot \sin w \\ \sin(z + w) &= \sin z \cdot \cos w + \cos z \cdot \sin w . \end{aligned}$$

Aufgabe I.2.5 :

Es sei $t \in \mathbb{R}$; E sei die 4×4 -Einheitsmatrix und $N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Berechnen Sie $e^{t(-E+N)}$.

Hinweis: Es dürfte die Aufgabe erleichtern und zu einem übersichtlichen Resultat führen, wenn Sie benützen, dass für kommutierende Matrizen A, B gilt $e^{A+B} = e^A \cdot e^B$.

I.3 Polynome als Rechengrößen

Ein Polynom mit komplexen Koeffizienten ist ein Ausdruck der Form

$$a_0 + a_1 \cdot z + a_2 \cdot z^2 + \dots + a_n \cdot z^n \quad \text{mit} \quad a_j \in \mathbb{C}$$

(z heißt die Unbestimmte des Polynoms).

Über Polynome ist sowohl Theoretisches als auch Praktisches zu sagen. Einiges sollte schon von der Schule her geläufig sein. Wir behandeln

I. Das Horner-Schema; Umzentrieren

II. Teilen mit Rest; euklidischer Algorithmus

III. Gebrochenrationale Funktionen; Partialbruchzerlegung

Das Lösen polynomialer Gleichungen $p(z) = 0$ hat theoretische und praktische Aspekte. Quadratische Gleichungen sind von der Schule her bekannt. Alles, was darüber hinaus geht, ist delikat. Dieses Thema werden wir erst nach und nach in den Griff bekommen.

$\mathbb{C}[z]$ als kommutativer Ring mit Eins

Polynome kann man addieren und multiplizieren. Es gibt ein neutrales Element für die Addition, das „Nullpolynom“, sowie ein neutrales Element für die Multiplikation, das Polynom $\equiv 1$.

Die Menge aller Polynome bezeichnet man mit $\mathbb{C}[z]$. In

$$(\mathbb{C}[z], 0, 1, +, \cdot)$$

gelten die Assoziativ-, die Distributiv- und die Kommutativgesetze.

Bemerkung

Eine multiplikative Inverse gibt es nur für diejenigen Polynome, die ein skalares Vielfaches des Einspolynoms sind (aber nicht das Nullpolynom!). Die konstanten Polynome, auch Polynome vom Grad 0 genannt, heißen die Einheiten von $\mathbb{C}[z]$.

Definition (Grad eines Polynoms)

Man sagt von einem Polynom $p(z) = \sum a_j \cdot z^j$, dass es einen Grad $< n$ hat, wenn $a_j = 0$ für $j = n, n+1, \dots$. Die Einheiten haben den Grad 0; dem Nullpolynom ordnet man üblicherweise den Grad $-\infty$ zu. Mit dieser Konvention gilt für den Grad $\deg(\cdot)$ („degree“)

$$\deg(p \cdot q) = \deg(p) + \deg(q) \quad \text{für alle} \quad p, q \in \mathbb{C}[z].$$

Sprechweise

Ein Polynom mit führendem Koeffizienten $= 1$ nennt man ein normiertes Polynom.

Bemerkung: Das Produkt normierter Polynome ist normiert.

Satz 1

Wenn ein Polynom $p(\cdot)$ im Punkt $z_0 \in \mathbb{C}$ verschwindet, dann besitzt es den Linearfaktor $z - z_0$. Es existiert ein Polynom $p^{(1)}(z)$, sodass gilt

$$p(z) = (z - z_0) \cdot p^{(1)}(z) .$$

Beweis

$p(z) - p(z_0) = \sum a_j(z^j - z_0^j) = (z - z_0) \cdot \sum a_j^{(1)} \cdot z^j$; denn aus jedem Summanden kann man den Faktor $z - z_0$ herausziehen

$$z^j - z_0^j = (z - z_0)(z^{j-1} + z^{j-2} \cdot z_0 + \dots + z \cdot z_0^{j-2} + z_0^{j-1}) .$$

Wenn man mehrere Nullstellen eines Polynoms kennt, kann man die entsprechenden Linearfaktoren (nach und nach) herausziehen; beispielsweise

$$\begin{aligned} z^3 - 1 &= (z - 1)(z^2 + z + 1) = (z - 1) \left(z + \frac{1}{2} + \frac{i}{2} \cdot \sqrt{3} \right) \left(z + \frac{1}{2} - \frac{i}{2} \sqrt{3} \right) \\ z^4 - 1 &= (z^2 - 1)(z^2 + 1) = (z - 1)(z + 1)(z - i)(z + i) . \end{aligned}$$

Corollar

Ein Polynom vom Grad $\leq n$, welches in mindestens $n + 1$ Punkten verschwindet, ist das Nullpolynom. Wenn zwei Polynome vom Grad $\leq n$ in $n + 1$ Punkten übereinstimmen, dann sind sie gleich.

Das Auffinden von Nullstellen eines Polynoms ist ein Problem mit langer Tradition. Schon im 17. Jahrhundert (A. Girard, 1629) wurde vermutet, dass jedes nichtkonstante Polynom eine Nullstelle besitzt. Den ersten Beweis gab C. F. Gauss in seiner Dissertation (1799). Gauss gab später noch weitere Beweise (siehe II.6).

Wenn man die Vielfachheiten der Nullstellen berücksichtigt, dann kann man sagen:

Satz 2 („Fundamentalsatz der Algebra“)

Jedes Polynom vom Grad n besitzt n Nullstellen. Es zerfällt in Linearfaktoren

$$p(z) = \text{const} \cdot (z - z_1)^{\ell_1} \cdot (z - z_2)^{\ell_2} \cdot (z - z_n)^{\ell_n}$$

mit z_j paarweise verschieden und $\ell_1 + \dots + \ell_n = n$.

Über die Techniken, mit welchen die Numeriker die Nullstellen eines Polynoms finden, können wir hier nichts sagen. Ein nützliches Hilfsmittel ist das Horner-Schema.

Ia) Horner-Schema

Es gibt viele Wege, den Wert eines Polynoms $p(z)$ vom Grad n einem Punkt z_0 zu evaluieren; man braucht mehr oder weniger viele Additionen und Multiplikationen. Das geschickteste Verfahren benötigt n Multiplikationen mit z_0 und n Additionen

$$\begin{aligned} p(z_0) &= a_n \cdot z_0^n + a_{n-1} \cdot z_0^{n-1} + \dots + a_1 \cdot z_0 + a_0 = \\ &= \left\{ [(a_n z_0 + a_{n-1}) z_0 + a_{n-2}] z_0 + \dots + a_1 \right\} z_0 + a_0 . \end{aligned}$$

Man berechnet nach und nach die Zahlen

$$\begin{aligned} a_{n-1}^{(1)} &= a_n \cdot z_0 + a_{n-1} , & a_{n-2}^{(1)} &= a_{n-1}^{(1)} \cdot z_0 + a_{n-2} \dots \\ \dots a_1^{(1)} &= a_2^{(1)} \cdot z_0 + a_1 , & p(z_0) &= a_1^{(1)} \cdot z_0 + a_0 . \quad \square \end{aligned}$$

Wir wollen die Idee des Horner-Schemas noch etwas weiter treiben. Als Vorbereitung stellen wir fest: Die Zwischenergebnisse $a_j^{(1)}$ liefern uns ein Polynom $p^{(1)}(z)$ mit

$$p(z) = p^{(1)}(z) \cdot (z - z_0) + p(z_0) .$$

Dies muss man nachrechnen (im Fall $n = 3$)

$$\begin{aligned} p(z) &= [(a_3 \cdot z + a_2)z + a_1] \cdot z + a_0 \\ &= [(a_3 z + a_2^{(1)})z + a_1^{(1)}] \cdot (z - z_0) + p(z_0) \\ \text{denn} \quad &a_j^{(1)} - a_{j+1}^{(1)} \cdot z_0 = a_j \quad \text{für } j = n, n-1, \dots . \end{aligned}$$

Ib) Umzentrieren eines Polynoms

Satz 3

Sei $p(z)$ ein Polynom und z_0 eine Zahl.

Es existieren dann Koeffizienten b_0, b_1, \dots, b_n , sodass

$$\begin{aligned} p(z) &= a_n z^n + a_{n-1} \cdot z^{n-1} + \dots + a_1 z + a_0 \\ &= a_n \cdot (z - z_0)^n + b_{n-1} \cdot (z - z_0)^{n-1} + \dots + b_1 \cdot (z - z_0) + b_0 . \end{aligned}$$

Beweis

1. Dass das Umzentrieren möglich ist, ist klar. Nach dem binomischen Lehrsatz haben wir für jedes n

$$z^n = ((z - z_0) + z_0)^n = \sum \binom{n}{k} z_0^{n-k} \cdot (z - z_0)^k .$$

Die rechte Seite ist eine Linearkombination von

$$(z - z_0)^n , (z - z_0)^{n-1} , \dots , (z - z_0) , 1 .$$

Also ist auch $\sum a_j z^j$ eine Linearkombination dieser Potenzen.

2. Es kommt uns jetzt darauf an, die Koeffizienten b_0, b_1, \dots mit möglichst wenig Aufwand zu berechnen. Dazu dient uns das Horner-Schema. Wie oben finden wir $p^{(1)}(z)$, sodass

$$p(z) = p^{(1)}(z) \cdot (z - z_0) + b_0 .$$

Nach demselben Verfahren

$$p^{(1)}(z) = p^{(2)}(z)(z - z_0) + b_1 .$$

3. Die Berechnungen liefern die Tabelle

$$\begin{array}{ccccccc}
 a_n & a_{n-1} & a_{n-2} & \dots & a_2 & a_1 & a_0 \\
 a_n & a_{n-1}^{(1)} & a_{n-2}^{(1)} & \dots & a_2^{(1)} & a_1^{(1)} & \boxed{b_0} \\
 a_n & a_{n-1}^{(2)} & a_{n-2}^{(2)} & \dots & a_2^{(2)} & \boxed{b_1} & \\
 \vdots & & \vdots & \vdots & & & \\
 a_n & a_{n-1}^{(n-2)} & a_{n-2}^{(n-2)} & \boxed{b_{n-3}} & & & \\
 a_n & a_{n-1}^{(n-1)} & \boxed{b_{n-2}} & & & & \\
 a_n & \boxed{b_{n-1}} & & & & &
 \end{array}$$

Wir berechnen Zeile für Zeile von links

$$a_j^{(k+1)} = a_{j+1}^{(k+1)} \cdot z_0 + a_j^{(k)} \quad \text{für } j = n, n-1, \dots, k+1$$

und schließlich

$$b_k = a_k^{(k+1)} = a_{k+1}^{(k+1)} \cdot z_0 + a_k^{(k)}.$$

Beispiel

$$p(z) = z^3 - 3z^2 + 2z + 1 \quad \text{und} \quad z_0 = 1.$$

Wir erhalten

$$p(z) = (z-1)^3 + 0 \cdot (z-1)^2 + (-1)(z-1) + 1.$$

Das Schema liefert nämlich

$$\begin{array}{cccc}
 1 & -3 & 2 & 1 \\
 1 & -2 & 0 & \boxed{1} \\
 1 & -1 & \boxed{-1} & \\
 1 & \boxed{0} & &
 \end{array}$$

IIa) Teilen mit Rest in \mathbb{Z} In der Schule lernt man das Teilen mit Rest in \mathbb{Z} :

$$\begin{array}{r}
 36271 : 17 = 2133 \text{ Rest } 10 \\
 \underline{- 34} \\
 22 \\
 \underline{- 17} \\
 57 \\
 \underline{- 51} \\
 61 \\
 \underline{- 51} \\
 10
 \end{array}$$

Grundlage des Verfahrens ist die Tatsache: Sind a, b positive ganze Zahlen mit $a > b$, dann gibt es $m \in \mathbb{N}$ und $r < a$, sodass $a = m \cdot b + r$.

Das Teilen mit Rest liefert auch den Schlüssel zur Bestimmung des größten gemeinsamen Teilers zweier natürlicher Zahlen a und b ($a > b$). Offenbar gilt $\text{ggT}(a, b) = \text{ggT}(a - b, b)$ und daher $\text{ggT}(a, b) = \text{ggT}(r, b) = \text{ggT}(b, r)$ ($b > r$). Iteriertes Anwenden liefert den berühmten

Satz 4 (Euklidischer Algorithmus)

Den größten gemeinsamen Teiler $\text{ggT}(\mathbf{a}, \mathbf{b})$ findet man durch iteriertes Teilen mit Rest

$$\begin{aligned} \mathbf{a} &= \mathbf{m}_1 \cdot \mathbf{b} + \mathbf{r}_2 \\ \mathbf{b} &= \mathbf{m}_2 \cdot \mathbf{r}_2 + \mathbf{r}_3 \\ \mathbf{r}_2 &= \mathbf{m}_3 \cdot \mathbf{r}_3 + \mathbf{r}_4 \\ &\dots \\ \mathbf{r}_{k-2} &= \mathbf{m}_{k-1} \cdot \mathbf{r}_{k-1} + \mathbf{r}_k \\ \mathbf{r}_{k-1} &= \mathbf{m}_k \cdot \mathbf{r}_k + \mathbf{0} . \end{aligned}$$

Der größte gemeinsame Teiler ist \mathbf{r}_k . Es existieren $\mathbf{s}, \mathbf{t} \in \mathbb{Z}$, sodass

$$\text{ggT}(\mathbf{a}, \mathbf{b}) = \mathbf{s} \cdot \mathbf{a} + \mathbf{t} \cdot \mathbf{b} .$$

Beweis

1. \mathbf{r}_k teilt alle \mathbf{r}_j und insbesondere \mathbf{a} und \mathbf{b} .
2. Wenn eine Zahl \mathbf{c} sowohl \mathbf{a} als auch \mathbf{b} teilt, dann teilt sie jedes \mathbf{r}_j , insbesondere \mathbf{r}_k .
3. Es gilt

$$\begin{aligned} \mathbf{r}_k &= \mathbf{r}_{k-2} - \mathbf{m}_{k-1} \cdot \mathbf{r}_{k-1} = \mathbf{r}_{k-2} - \mathbf{m}_{k-1}(\mathbf{r}_{k-3} - \mathbf{m}_{k-2} \cdot \mathbf{r}_{k-2}) = \\ &= \tilde{\mathbf{s}} \cdot \mathbf{r}_{k-2} + \tilde{\mathbf{t}} \cdot \mathbf{r}_{k-3} = \dots = \mathbf{s} \cdot \mathbf{a} + \mathbf{t} \cdot \mathbf{b} . \end{aligned}$$

IIb) Teilen mit Rest in $\mathbb{C}([z])$ **Satz 5**

Seien $\mathbf{p}(z)$ und $\mathbf{q}(z)$ Polynome mit $\deg(\mathbf{p}) > \deg(\mathbf{q})$. Es existieren dann Polynome $\mathbf{m}(z)$ und $\mathbf{r}(z)$ mit

$$\mathbf{p}(z) = \mathbf{m}(z) \cdot \mathbf{q}(z) + \mathbf{r}(z) , \quad \deg(\mathbf{r}) < \deg(\mathbf{q}) .$$

Beispiel

$$\begin{aligned} \mathbf{p}(z) &= -4z^5 + 2z^4 - 14z^3 + 6z^2 - 14z + 10 \\ \mathbf{q}(z) &= 2z^3 + 3z - 1 . \end{aligned}$$

Wir zeigen

$$\mathbf{p}(z) = \mathbf{q}(z) \cdot (-2z^2 + z - 4) + (z^2 - z + 6) .$$

Das geht genauso wie für Zahlen:

$$\begin{array}{r} (-4z^5 + 2z^4 - 14z^3 + 6z^2 - 14z + 10) : (2z^3 + 3z - 1) = (-2z^2 + z - 4) + \text{Rest} \\ \underline{-4z^5 + 0 - 6z^3 + 2z^2} \\ 2z^4 - 8z^3 + 4z^2 - 14z \\ \underline{ 2z^4 + 0 + 3z^2 - z} \\ -8z^3 + z^2 - 13z + 10 \\ \underline{ -8z^3 + 0 - 12z + 4} \\ \text{Rest :} z^2 - z + 6 \end{array}$$

Beweis des Satzes (durch Induktion nach $n = \deg(p)$)

$q(z)$ sei fest. Wir nehmen an, der Satz gelte für alle $\tilde{p}(z)$ mit $\deg(\tilde{p}) < n$. $p(z)$ sei ein Polynom vom Grad n .

$$p(z) = a_n z^n + \dots \quad ; \quad q(z) = b_m z^m + \dots$$

$$p(z) - q(z) \cdot \left(\frac{a_n}{b_m} \cdot z^{n-m} \right) = \tilde{p}(z) \text{ ist eine Polynom vom Grad } < n .$$

Satz 6

Seien p und q Polynome.

Es existiert dann ein normiertes Polynom n , sodass gilt

1. n teilt p und n teilt q
2. Jedes Polynom, welches p und q teilt, teilt auch n .

Dieses Polynom ist eindeutig bestimmt. Es heißt der normierte größte gemeinsame Teiler: $n = \text{ggT}(p, q)$.

Es existieren Polynome s und t , sodass $n = s \cdot p + t \cdot q$.

Beweis (Euklidischer Algorithmus)

Wir machen immer wieder Teilen mit Rest.

Zur Vereinfachung der Notation $p(z) = r_0$, $q(z) = r_1$

$$\begin{aligned} r_0 &= m_1 \cdot r_1 + r_2 \\ r_1 &= m_2 \cdot r_2 + r_3 \\ &\dots \\ r_{k-2} &= m_{k-1} \cdot r_{k-1} + r_k \\ r_{k-1} &= m_k \cdot r_k + 0 . \end{aligned}$$

Irgendwann geht die Division auf; denn $\deg(r_1) > \deg(r_2) > \dots$. Wie bei den Zahlen gewinnt man r_k als Summe $r_k = s \cdot p + t \cdot q$. Dies ist bis auf einen Faktor der normierte $\text{ggT}(r_0, r_1)$.

Beispiel

$$\begin{aligned} \text{ggT}(z^3 - 1, z^2 - 1) &= z - 1 \\ (z - 1) &= s \cdot p(z) + t \cdot q(z) \quad \text{mit } s = 1 \text{ und } t = z \\ z - 1 &= (z^3 - 1) - z \cdot (z^2 - 1) . \end{aligned}$$

III a) Gebrochenrationale Funktionen Wir betrachten Brüche von Polynomen. Wir haben Gleichheit

$$\frac{p(z)}{q(z)} = \frac{\tilde{p}(z)}{\tilde{q}(z)} ,$$

wenn die Brüche durch Kürzen und Erweitern auseinander hervorgehen. (Der Nenner darf nicht das Nullpolynom sein.) $p(z) \cdot \tilde{q}(z) = q(z) \cdot \tilde{p}(z)$.

In der Schule lernt man die Bruchrechnung für ganze Zahlen. Für Polynome geht alles genauso. Das Multiplizieren ist leicht. Beim Addieren dagegen muss man bekanntlich aufpassen; man muss die Summanden auf einen gemeinsamen Nenner bringen!

$$\begin{aligned}\left(\frac{p_1}{q_1}\right)\left(\frac{p_2}{q_2}\right) &= \frac{p_1 \cdot p_2}{q_1 \cdot q_2} \\ \left(\frac{p_1}{q_1}\right) + \left(\frac{p_2}{q_2}\right) &= \frac{1}{q_1 \cdot q_2} \cdot (p_1 q_2 + p_2 \cdot q_1) .\end{aligned}$$

Die Brüche von Polynomen nennt man manchmal die **gebrochenrationalen** Funktionen, während man die Polynome **ganzrationale** Funktionen nennt. Die ganzrationalen Funktionen kann man als spezielle gebrochenrationale Funktionen (mit dem Nenner =1) auffassen. Wir haben eine Erweiterung von $\mathbb{C}([z])$ nach demselben Schema wie die Erweiterung $\mathbb{Z} \subseteq \mathbb{Q}$. Alle Brüche mit Ausnahme der Null haben eine multiplikative Inverse. Die Menge der gebrochenrationalen Funktionen ist ein Körper. Den Beweis überlassen wir der Vorlesung über Elementare Algebra. Das Stichwort dort lautet: Erweiterung eines Integritätsbereichs zu einem Körper.

Beispiel für das Addieren:

$$\frac{1}{z-1} + \frac{1}{z+1} - \frac{2}{z} = \frac{2z}{(z-1)(z+1)} - \frac{2}{z} = \frac{1}{z(z^2-1)}((2z^2 - 2(z^2 - 1))) = \frac{2}{z^3 - z} .$$

IIIb) Partialbruchzerlegung Im Folgenden nennen wir eine gebrochenrationale Funktion $\frac{p(z)}{q(z)}$ einen echten Bruch, wenn $\deg(p) < \deg(q)$. Die Funktionen der Form $\frac{1}{(z-\bar{z})^\ell}$ ($\ell \in \mathbb{N}$), nennen wir (in diesem Abschnitt) reine Potenzen.

Satz 7

Jeder echte Bruch lässt sich als Linearkombination reiner Potenzen darstellen. Sei

$$\begin{aligned}p(z) &= a_0 + a_1 \cdot z + \dots + a_{n-1} \cdot z^{n-1} \\ q(z) &= (z - z_1)^{\ell_1} \dots (z - z_m)^{\ell_m} \text{ mit } z_j \text{ paarweise verschieden.}\end{aligned}$$

Der Bruch $\frac{p(z)}{q(z)}$ ist dann darstellbar als Linearkombination der „reinen Potenzen“

$$\left(\frac{1}{z - z_1}\right)^{\ell_1}, \left(\frac{1}{z - z_1}\right)^{\ell_1-1}, \dots, \left(\frac{1}{z - z_1}\right), \left(\frac{1}{z - z_2}\right)^{\ell_2}, \dots, \left(\frac{1}{z - z_m}\right).$$

Beispiele

$$1. \frac{2}{z \cdot (z-1)(z+1)} = \frac{1}{z-1} + \frac{1}{z+1} - \frac{2}{z} \text{ (siehe oben)}$$

$$2. \frac{a_0 + a_1 z + a_2 z^2}{(z-1)^2 \cdot (z+1)} = \frac{\alpha_1}{(z-1)^2} + \frac{\alpha_2}{z-1} + \frac{\alpha_3}{z+1}$$

Wir diskutieren in der nächsten Vorlesung, wie man in solchen Fällen die n -Tupel (a_0, \dots, a_{n-1}) und $(\alpha_1, \dots, \alpha_n)$ ineinander umrechnet.

Beweis des Satzes

Wir beweisen durch Induktion nach dem Grad n des Nennerpolynoms.

Für $\deg(q) = 1$ ist nichts zu beweisen.

Sei \tilde{z} eine ℓ -fache Nullstelle von $q(z)$

$$q(z) = (z - \tilde{z})^\ell \cdot Q(z) \quad \text{mit} \quad Q(\tilde{z}) \neq 0.$$

Wenn wir die Zahl α geeignet wählen, dann gilt

$$\begin{aligned} (p - \alpha Q)(\tilde{z}) &= 0 \\ \frac{p(z)}{q(z)} - \alpha \frac{1}{(z - \tilde{z})}^\ell &= \frac{p(z) - \alpha Q(z)}{q(z)}. \end{aligned}$$

Wir können durch $z - \tilde{z}$ kürzen und haben

$$\frac{p(z)}{q(z)} = \frac{\alpha}{(z - \tilde{z})^\ell} + \frac{p_1(z)}{q_1(z)},$$

wo $\deg(q_1) < \deg(q)$ und $\deg(p_1) < \deg(q_1)$.

Die Induktionsvoraussetzung greift.

Beispiele

1. $\frac{4}{(z-1)^2(z+1)} = \frac{\alpha}{(z-1)^2} + \frac{\beta}{(z-1)} + \frac{\gamma}{(z+1)}$
 $\frac{4}{(z-1)^2(z+1)} - \frac{2}{(z-1)^2} = \frac{4-2(z+1)}{(z-1)^2(z+1)} = \frac{(-2)(z-1)}{(z-1)^2(z+1)} = \frac{-2}{(z-1)(z+2)}$
2. $\frac{4}{z^4-1} = \frac{4}{(z-1)(z+1)(z-i)(z+i)} = \frac{1}{z-1} - \frac{1}{z+1} + \frac{i}{z-i} - \frac{i}{z+i}.$

Anhang I.3: Die Stammfunktionen der gebrochenrationalen Funktionen

In der Schule studiert man die Stammfunktionen der gebrochenrationalen Funktionen im Reellen

$$f(x) = P(x) + \frac{r(x)}{q(x)} \quad \text{für } x \in \text{Intervall} .$$

Die Stammfunktion des Polynoms $P(x)$ macht keine Schwierigkeiten; den ‘echten’ Bruch mit Zählergrad kleiner als Nennergrad behandelt man mit der Partialbruchzerlegung. Interessant sind die Beispiele:

$$\begin{aligned} f(x) &= \frac{\alpha}{x - x_1} \quad \text{für } x > x_1 & \text{oder speziell} & \quad f(x) = \frac{1}{x} \quad \text{für } x > 0 , \\ f(x) &= \frac{\alpha + \beta x}{(x - x_1)^2 + c^2} & \text{oder speziell} & \quad f(x) = \frac{1}{1 + x^2} . \end{aligned}$$

In der Schule lernt man

$$\text{StF}\left(\frac{1}{x}\right) = \ln x + \text{const} ; \quad \text{StF}\left(\frac{1}{1 + x^2}\right) = \arctan x + \text{const} .$$

Und man beweist das, indem man die Ableitung der Logarithmusfunktion bzw. der Arcustangensfunktion (als Umkehrabbildungen definiert) mittels der Kettenregel berechnet.

$$\begin{aligned} \frac{d}{dx} e^x &= e^x ; & \frac{d}{dx} \frac{\sin x}{\cos x} &= 1 + \tan^2 x . \\ e^{\ln x} = x &\implies e^{\ln x} \cdot \frac{d}{dx} \ln x = 1 &\implies \frac{d}{dx} \ln x &= \frac{1}{x} . \\ \tan(\arctan x) = x &\implies (1 + \tan^2)(\arctan x) \cdot \frac{d}{dx} \arctan x = 1 &\implies \frac{d}{dx} \arctan x &= \frac{1}{1 + x^2} . \end{aligned}$$

Im Leistungskurs oder in der Infinitesimalrechnung für Mathematiker lernt man, dass das alles nicht so einfach ist. Es müssen Fragen geklärt werden wie z. B.

1. Wo sind unsere Funktionen definiert, und welche Werte können sie annehmen?
2. Existiert eine Umkehrfunktion; gibt es zu jedem Bildpunkt einen eindeutig bestimmten Urbildpunkt?
3. In welchen Bereichen gibt es eine Stammfunktion?
4. Sind die Voraussetzungen für die Anwendung der Kettenregel erfüllt?

In unserem Fall muss man sich klarmachen, dass man den Tangens auf das Intervall $(-\pi/2, \pi/2)$ einschränken muss, wenn man eine Umkehrfunktion haben will, die im Nullpunkt verschwindet. Der Logarithmus ist in der reellen Theorie nur für positive Argumente definiert. In den richtig gewählten Bereichen geht dann aber alles glatt mit Integrabilität und Differenzierbarkeit.

Wir wollen uns hier noch mit den Stammfunktionen der gebrochen rationalen Funktionen mit komplexen Koeffizienten und **komplexem Argument** z befassen.

Die Exponentialfunktion ist in der gesamten komplexen Ebene wohldefiniert; die Tangensfunktion bleibt nur in den Punkten $(k + 1/2)\pi$, d.h. in den Nullstellen der Cosinusfunktion $\cos z = \frac{1}{2}(e^{iz} + e^{-iz})$, undefiniert.

Die Exponentialfunktion bildet den Streifen $\{z : |\Im z| < \pi\}$ bijektiv auf die ‘geschlitzte komplexe Ebene’ ab. Für jedes w in der geschlitzten Ebene gibt es einen eindeutig definierten Winkel im Intervall $(-\pi, \pi)$, den man hier den Arcus von w oder das ‘Argument’ $\arg(w)$ nennt. Wir haben daher tatsächlich eine Umkehrfunktion in der geschlitzten Ebene (die man den Hauptwert des Logarithmus nennt und wie im Reellen mit \ln bezeichnet).

$$w = e^{x+iy} = e^x \cdot (\cos y + i \sin y) \implies x = \ln |w|, \quad y = \arg(w) \implies \ln w = \ln |w| + i \cdot \arg(w).$$

Der Logarithmus im Komplexen ist in der Tat eine Stammfunktion der Funktion $\frac{1}{z}$ in dem Sinne, dass die Ableitung diese Funktion ergibt. Dies sieht man so:

$$\begin{aligned} \ln(w_1 \cdot w_2) &= \ln w_1 + \ln w_2 + k \cdot 2\pi i \quad \text{mit einem geeigneten } k \in \mathbb{Z} \\ \lim_{h \rightarrow 0} \frac{1}{h} (\ln(w+h) - \ln w) &= \lim_{h \rightarrow 0} \frac{1}{h} \ln\left(\frac{w+h}{w}\right) = \frac{1}{w} \lim_{h \rightarrow 0} \frac{w}{h} \ln\left(1 + \frac{h}{w}\right) = \frac{1}{w}. \end{aligned}$$

Wenn man die Arcustangensfunktion im Komplexen studieren will, dann sollte man sich wohl am besten auf den Streifen $\{z : |\Re z| < \pi/2\}$ konzentrieren. Als Stammfunktion der Funktion $\frac{1}{1+w^2} = \frac{1}{2i} \left(\frac{1}{w-i} - \frac{1}{w+i} \right)$, die im Ursprung verschwindet, kommt in Betracht

$$A(w) = \frac{1}{2i} \ln \frac{1+iw}{1-iw} \quad (= \frac{1}{2i} \ln \frac{w-i}{w+i} + \text{const}).$$

Und diese Funktion kann in der Tat (in einer gewissen Umgebung des Ursprungs) als die Umkehrfunktion zum Tangens gelten. Wir haben nämlich

$$\begin{aligned} \tan z &= \frac{\sin z}{\cos z} = \frac{2 e^{iz} - e^{-iz}}{2i e^{iz} + e^{-iz}} = \frac{1 e^{2iz} - 1}{i e^{2iz} + 1} \quad \text{und daher gilt für} \\ z &= \frac{1}{2i} \ln \frac{1+iw}{1-iw} \quad e^{2iz} = \frac{1+iw}{1-iw} \quad \text{schliesslich} \quad \tan z = \frac{1(1+iw) - (1-iw)}{i(1+iw) - (1-iw)} = w. \end{aligned}$$

Beachte Unsere (etwas unscharfen) Andeutungen zum Thema der gebrochenrationalen Funktionen im Komplexen gehören eigentlich nicht in den gegenwärtigen Abschnitt I: ‘Rechengrößen’. Die Fragen der Definitionsbereiche von Abbildungen werden im Abschnitt II grundsätzlich studiert. Am Ende des Abschnitts II werden übrigens dann auch Beweise des ‘Fundamentalsatzes der Algebra’ angedeutet; eine strenge Durchführung dieser Beweise müssen wir allerdings der Analysis überlassen.

Der didaktische Grund für das Abweichen von der grossen Linie ist der, dass wir uns die Gelegenheit nicht entgehen lassen wollten, einerseits an Schulkenntnisse zu erinnern und andererseits weitergehende geometrische und analytische Überlegungen anzudeuten.

Aufgaben zu I.3

Aufgabe I.3.1 :

Wir haben gesehen (Aufgabe I.2.2)

$$z^5 - 1 = (z - 1)(z^4 + z^3 + z^2 + z + 1) = (z - 1)(z^2 + az + 1)(z^2 + Az + 1)$$

und wir haben die Zahlen a, A bestimmt.

Finden Sie Zahlen a_0, a_1, a_2, a_3, a_4 , sodass gilt

$$\frac{1}{z^5 - 1} = \frac{a_0}{z - 1} + \frac{a_1 + a_2 z}{z^2 + az + 1} + \frac{a_3 + a_4 z}{z^2 + Az + 1}.$$

Aufgabe I.3.2 : (Achte Einheitswurzeln)

Eine komplexe Zahl w heisst eine primitive achte Einheitswurzel, wenn gilt $w^4 = -1$.

1. Aus einem Bild können Sie Real- und Imaginärteile der primitiven achten Einheitswurzeln z_1, z_2, z_3, z_4 ablesen (z. B. $z_1 = w, z_2 = \bar{w}, z_3 = -w, z_4 = -\bar{w}$). Geben Sie diese an!
2. Zeigen Sie:

$$z^4 + 1 = p(z) \cdot q(z) \quad \text{wenn} \quad p(z) = (z - w)(z - \bar{w}), \quad q(z) = (z + w)(z + \bar{w}).$$

3. Finden Sie Zahlen a_1, a_2, a_3, a_4 , sodass gilt

$$\frac{1}{z^4 + 1} = \frac{a_1}{z - z_1} + \frac{a_2}{z - z_2} + \frac{a_3}{z - z_3} + \frac{a_4}{z - z_4}.$$

4. Zeigen Sie dass man $\frac{a_1}{z - z_1} + \frac{a_2}{z - z_2}$ und $\frac{a_3}{z - z_3} + \frac{a_4}{z - z_4}$ als Quotienten reeller Polynome darstellen kann.

I.4 Allerlei komplexe $n \times n$ -Matrizen. Faktorisierungen

Wenn A und B komplexe $n \times n$ -Matrizen sind, dann sind auch $A + B$ und $A \cdot B$ komplexe $n \times n$ -Matrizen. Man kann mit den komplexen $n \times n$ -Matrizen fast genauso rechnen wie mit den komplexen Zahlen; (\mathbb{C} ist in der Tat der Spezialfall $n=1$). Gegenüber den reellen Zahlen hat man noch die sog. hermitesche Konjugation als eine weitere Operation.

Zwei wichtige Unterschiede gegenüber dem Rechnen mit Zahlen sind für $n \geq 2$ bedeutungsvoll: $B \cdot A$ ist nicht immer gleich $A \cdot B$, und es ist nicht nur das Nullelement, welches keine multiplikative Inverse besitzt (in der Tat gibt es Matrizen $\neq O$, deren Produkt die Nullmatrix ist).

In der Fachsprache sagt man: Der Ring (oder die Algebra) der komplexen $n \times n$ -Matrizen ist nicht kommutativ und er ist nicht nullteilerfrei.

Wir stellen die **Rechenregeln** zusammen

$$(A + B) + C = A + (B + C)$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

$$A + B = B + A$$

$$A \cdot (B + C) = A \cdot B + A \cdot C$$

$$(B + C) \cdot A = B \cdot A + C \cdot A.$$

Es gibt ein neutrales Element für die Addition ('Nullelement') und ein neutrales Element für die Multiplikation ('Einselement')

$$A + O = A = O + A$$

$$A \cdot I = A = I \cdot A.$$

Bemerkung: 'Axiome' Wenn wir zu den hier aufgelisteten Regeln noch die Existenz der additiven Inversen $-A$ dazunehmen, dann haben wir ein Axiomensystem für einen *Ring mit Einselement*.

Bei unseren $n \times n$ -Matrizen haben wir außerdem eine **Multiplikation mit Skalaren** (d. h. hier mit komplexen Zahlen $\alpha, \beta, \gamma, \dots$); und dafür gelten die Regeln

$$\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$$

$$(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$$

$$\alpha \cdot (A \cdot B) = (\alpha \cdot A) \cdot B$$

$$(\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$$

$$1 \cdot A = A.$$

Damit haben wir die Axiome einer *assoziativen Algebra*. Die Axiome gelten z. B. auch für die Algebra der reellen $n \times n$ -Matrizen (mit reellen Skalaren). Weitere wichtige Beispiele für assoziative Algebren sind die Menge der Polynome mit komplexen (bzw. reellen) Koeffizienten. Diese Algebren sind im Unterschied zu unserer Algebra der $n \times n$ -Matrizen kommutativ und nullteilerfrei.

Hermitische Konjugation Bei den komplexen $n \times n$ -Matrizen interessiert noch eine weitere Struktur: die (*hermitische*) *Konjugation*. Wir haben eine Zuordnung $*$: für jede Matrix A ist A^* eine wohldefinierte Matrix. Die Einträge von $B = A^*$ ergeben sich, indem man die Zeilen mit den Spalten vertauscht ('Transposition') und gleichzeitig die Einträge durch die konjugiertkomplexen Einträge ersetzt.

$$b_{kl} = \bar{a}_{lk}$$

Diese (offensichtlich involutorische) Abbildung hat die Eigenschaften

$$\begin{aligned}(A + B)^* &= A^* + B^* \\ (A \cdot B)^* &= B^* \cdot A^* \\ (A^*)^* &= A \\ (\alpha \cdot A)^* &= \alpha^* \cdot A^*.\end{aligned}$$

(α^* ist dasselbe wie die konjugiertkomplexe Zahl $\bar{\alpha}$). (Man beachte die Reihenfolge der Faktoren: sie wird bei der hermitischen Konjugation umgekehrt.)

Invertierbare Matrizen Eine komplexe $n \times n$ -Matrix A heißt *nichtsingulär* oder auch *invertierbar*, wenn es eine Matrix B gibt, sodass gilt

$$A \cdot B = I = B \cdot A.$$

Man zeigt leicht: Wenn es eine solche Matrix zu A gibt, dann ist sie eindeutig bestimmt; man nennt sie die zu A *inverse Matrix* und bezeichnet sie mit A^{-1} .

Man bemerke: Wenn A und B invertierbare Matrizen sind, dann ist auch die Produktmatrix $A \cdot B$ invertierbar, und es gilt

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

Satz Eine komplexe $n \times n$ -Matrix A ist genau dann invertierbar, wenn gilt

$$A \cdot x \neq 0 \quad \text{für alle Spalten } x \neq 0.$$

Diese wichtige Einsicht werden wir uns später im Rahmen der Vektorraumtheorie gründlich erarbeiten. Wenn wir die n -Spalten als Vektoren verstehen, dann besagt der Satz:

Eine $n \times n$ -Matrix ist genau dann invertierbar, wenn ihre Spalten linear unabhängig sind.

Wir werden zwei wichtige Algorithmen kennenlernen, welche bei einer Matrix feststellen, ob diese invertierbar ist und gegebenenfalls die Inverse berechnen. (Der eine ist ein Spezialfall des Gauss'schen Eliminationsverfahrens, der andere ist ein Spezialfall des 'vollständigen Austauschs' für ein Tableau.)

Für Matrizen vom Format $n > 4$ ist der Rechenaufwand für die Invertierung beträchtlich. Im Falle $n = 2$ gibt es eine einfache Formel:

Die Matrizen

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{und} \quad \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

sind zueinander invers.

Hermitische und unitäre Matrizen

- Eine komplexe $n \times n$ -Matrix H heißt *hermitische* Matrix, wenn sie gleich ihrer hermitisch Konjugierten ist: $H = H^*$.
- Eine komplexe $n \times n$ -Matrix U heißt *unitäre* Matrix, wenn gilt $U^* \cdot U = I = U \cdot U^*$.
- Eine hermitische Matrix S mit lauter reellen Einträgen heißt auch eine (reelle) *symmetrische* Matrix: $S = S^T$.
- Eine unitäre Matrix T mit lauter reellen Einträgen heißt eine *orthogonale* Matrix: $T \cdot T^T = I = T \cdot T^T$.

Bemerkung Eine Diagonalmatrix D ist genau dann hermitisch, wenn die Diagonalelemente reell sind. Sie ist genau dann unitär, wenn die Diagonalelemente den Betrag 1 haben.

Satz Wenn H hermitisch ist, dann ist $U_t = \exp(itH)$ unitär für jedes reelle t .

$$\begin{aligned} U_t &= I + itH + \frac{1}{2!}(it)^2 H^2 + \dots, \\ U_t^* &= I + (-it)H + \frac{1}{2!}(-it)^2 H^2 + \dots = U_{-t}. \end{aligned}$$

Die Summe zweier hermitischer Matrizen ist hermitisch; wenn man eine hermitische Matrix mit einer reellen(!) Zahl multipliziert, dann erhält man wieder eine hermitische Matrix. Die Gesamtheit aller hermitischen $n \times n$ Matrizen ist also gegenüber reellen Linearkombinationen stabil. In der Fachsprache drückt man das so aus:

Die Gesamtheit aller hermitischen $n \times n$ Matrizen ist ein reeller Vektorraum.

(Seine Dimension ist übrigens n^2 .)

Das Produkt zweier unitärer Matrizen ist unitär; die Inverse einer unitären Matrix ist unitär. In der Fachsprache drückt man das so aus: *Die Gesamtheit aller unitären $n \times n$ -Matrizen ist eine Gruppe.* (Der allgemein übliche Name ist übrigens $U(n)$.)

Eine interessante Klasse von Matrizen, welche sowohl die hermitischen als auch die unitären Matrizen umfasst, ist die Klasse der *normalen* Matrizen. Eine $n \times n$ -Matrix N heißt eine *normale* Matrix, wenn sie mit ihrer Konjugierten kommutiert: $N^* \cdot N = N \cdot N^*$.

Jede komplexe $n \times n$ -Matrix A besitzt genau eine Darstellung $A = S + iT$ mit hermitischen S und T . Die Matrix ist genau dann normal, wenn S und T kommutieren.

$$S = \frac{1}{2}(A + A^*), \quad T = \frac{1}{2i}(A - A^*).$$

Positiv definite Matrizen Eine hermitesche Matrix H heisst *positiv definite* Matrix, wenn gilt

$$x^* \cdot H \cdot x > 0 \quad \text{für alle Spalten } x \neq 0.$$

Sie heisst *positiv semidefinit*, wenn gilt

$$x^* \cdot H \cdot x \geq 0 \quad \text{für alle Spalten } x.$$

Unitär diagonalisierbare Matrizen Eine Matrix A heisst *unitär diagonalisierbar*, wenn eine unitäre Matrix U existiert und eine Diagonalmatrix D , sodass

$$U^* \cdot A \cdot U = D, \quad \text{d. h.} \quad A \cdot U = U \cdot D \quad \text{mit } U \text{ unitär.}$$

Satz Eine Matrix A ist genau dann unitär diagonalisierbar, wenn sie normal ist. Insbesondere ist also jede hermitesche Matrix unitär diagonalisierbar.

Den Beweis dieses extrem wichtigen Satzes verschieben wir auf später. Der Satz wird manchmal der Satz von der **Hauptachsentransformation** genannt.

Diagonalisierbare Matrizen Eine Matrix A heisst *diagonalisierbar*, wenn eine nicht-singuläre Matrix B existiert und eine Diagonalmatrix D , sodass

$$B^{-1} \cdot A \cdot B = D, \quad \text{d. h.} \quad A \cdot B = B \cdot D \quad \text{mit } B \text{ invertierbar.}$$

Didaktischer Hinweis Die Theorie der diagonalisierbaren Matrizen weist Überschneidungen auf mit der Theorie der Hauptachsentransformation. Es gibt z. B. den Satz:

Eine hinreichende Bedingung für die Diagonalisierbarkeit von A ist die Bedingung, dass das charakteristische Polynom nur einfache Nullstellen besitzt. Die Diagonalelemente sind die Eigenwerte, die Spalten von B sind die Eigenvektoren (der zu A gehörenden linearen Abbildung der Spaltenvektoren). Wenn A eine normale Matrix ist, dann stehen die Eigenvektoren paarweise aufeinander senkrecht.

Leider führen die Überschneidungen der Theorie der Hauptachsen mit der Theorie der diagonalisierbaren Matrizen bei den Anfängern manchmal zu Kurzschlüssen. Wir wollen die Thematik der (nicht notwendigerweise) diagonalisierbaren Matrizen der fortgeschrittenen linearen Algebra überlassen. Auch die allgemeine Theorie der charakteristischen Polynome und ihrer Nullstellen (Eigenwerte eines Endomorphismus oder seiner darstellenden Matrix) stellen wir zurück. Hier mag der Hinweis genügen, dass der Satz von der Jordan'schen Normalform als die Krönung dieser Theorie gilt.

Die Theorie der Hauptachsen dagegen wird uns immer wieder beschäftigen, z. B. beim Studium der hermiteschen Formen und bei der Singulärwertzerlegung einer rechteckigen Matrix.

Der folgende Satz (den wir hier noch nicht beweisen) stellt eine matrizentheoretische Analogie her zu der Darstellung einer komplexen Zahl in Polarkoordinaten.

Satz

Jede nichtsinguläre Matrix A besitzt genau eine Darstellung $A = R \cdot U$ mit positiv definitem R und unitärem U ; sie ist genau dann normal, wenn R und U kommutieren.

Der entscheidende Schritt beim Beweis dieses Satzes ist das folgende Lemma von der positiv definiten Quadratwurzel zu einer positiv definiten Matrix:

Lemma

Zu jeder positiv definiten Matrix H (wie z. B. $H = A \cdot A^$) gibt es genau eine positiv definite Matrix R mit $R^2 = R^*R = H$.*

Diese Sätze sind natürlich trivial, wenn man den Satz von der Hauptachsentransformation schon hat.

Weitere Faktorisierungen einer nichtsingulären Matrix

Eine $n \times n$ -Matrix L heisst untere (bzw. obere) *Dreiecksmatrix*, wenn die Einträge oberhalb (bzw. unterhalb) der Diagonalen allesamt verschwinden; wir sprechen von einer *normierten Dreiecksmatrix*, wenn zusätzlich die Diagonalelemente allesamt gleich 1 sind.

Man bemerke, dass das Produkt zweier normierter unterer Dreiecksmatrizen eine normierte untere Dreiecksmatrix ist, und dass die Inverse einer normierten unteren Dreiecksmatrix eine normierte untere Dreiecksmatrix ist.

Die Inverse einer Matrix kann man offenbar spaltenweise berechnen; für jede Spalte gilt es, das lineare Gleichungssystem für die entsprechende Einheitsspalte zu lösen.

Bei den (normierten) Dreiecksmatrizen ist die Lösung eines beliebigen linearen Gleichungssystems sehr einfach. Das folgende Beispiel zeigt, warum man im Falle der oberen Dreiecksmatrizen von backward substitution (oder ‘von hinten Einsetzen’) spricht.

Beispiel Wir berechnen die Lösung $(x, y, z)^T$ des Gleichungssystems

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Die dritte Gleichung ergibt sofort z ; wenn wir z schon haben, dann liefert die zweite Zeile y ; wenn wir y, z schon haben, dann liefert die erste Zeile x .

Satz

Wenn man die Spalten einer nichtsingulären Matrix A geeignet anordnet, dann gewinnt man eine Matrix \tilde{A} , welche eine Faktorisierung erlaubt

$$\tilde{A} = L \cdot D \cdot M,$$

wo L eine normierte untere Dreiecksmatrix ist, D eine Diagonalmatrix und M eine normierte obere Dreiecksmatrix. Die Faktoren sind durch \tilde{A} eindeutig bestimmt.

Eine Konstruktion von M (und auch von $L \cdot D$) erhält man mittels des Gauss'schen Eliminationsverfahren, (welches möglicherweise schon von der Schule her bekannt ist). Dieses Verfahren wird für allgemeine rechteckige Matrizen in V.4 ausführlich besprochen. Die Eindeutigkeit der Faktorisierung im Spezialfall der nichtsingulären Matrizen ergibt sich folgendermaßen:

$$L_1 \cdot D_1 \cdot M_1 = \tilde{A} = L_2 \cdot D_2 \cdot M_2 \implies$$

$$(L_2^{-1} \cdot L_1) \cdot D_1 = D_2 \cdot (M_2 \cdot M_1^{-1})$$

Die linke Seite ist eine untere Dreiecksmatrix, die rechte ist eine obere Dreiecksmatrix; da sie gleich sind, sind sie in Wirklichkeit Diagonalmatrizen mit denselben Diagonalelementen. Die Diagonalmatrizen D_1 und D_2 sind gleich und

$$L_2^{-1} \cdot L_1 = I = M_2 \cdot M_1^{-1}.$$

Diese sog. LU-Zerlegung ist übrigens ein probates Mittel zur Berechnung der Determinante $\det \tilde{A}$. Die Determinante einer normierten Dreiecksmatrix ist nämlich $= 1$ und die Determinante einer Diagonalmatrix ist das Produkt der Diagonalelemente.

Permutationsmatrizen Das Umstellen der Spalten einer Matrix A kann man dadurch erreichen, dass man die Matrix mit einer Permutationsmatrix P von rechts multipliziert. $\tilde{A} = A \cdot P$. Eine $n \times n$ -Matrix P heißt eine Permutationsmatrix, wenn sie in jeder Zeile und in jeder Spalte genau einen Eintrag 1 besitzt und sonst nur Nulleinträge.

GS-Zerlegung Das berühmte Orthogonalisierungsverfahren nach Gram-Schmidt liefert eine weitere interessante eindeutige Faktorisierung einer beliebigen nichtsingulären Matrix A : $A = U \cdot M$, wo M eine obere Dreiecksmatrix und U eine unitäre Matrix ist. Darüber wird in der Theorie der Hilberträume zu sprechen sein. (Abschnitt IV).

Hinweis Die Bedeutung der hier genannten Faktorisierungen wird sich erschliessen, wenn wir erst einmal die Matrizen als Darstellungsformen von Abbildungen (Abschnitt II) bzw. als Darstellungsformen von Bilinearformen verstehen. (Abschnitt IV). Man bleibt an der Oberfläche, wenn man, wie wir hier die Matrizen nur als Rechengrößen behandelt.

Aufgaben zu I.4**Aufgabe I.4.1 :** (Eigenwerte hermitescher 2×2 -Matrizen)

Eine komplexe Zahl λ heisst ein Eigenwert der komplexen 2×2 -Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$, wenn gilt

$$\det(A - \lambda I) = 0 = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}.$$

Zeigen Sie, dass für hermitesche Matrizen die Eigenwerte reell sind.

Aufgabe I.4.2 : (Produkte unitärer Matrizen)

Eine komplexe $I \times J$ -Matrix U heisst eine unitäre Matrix, wenn gilt

$$U \cdot U^* = E_{I \times I} \quad \text{und} \quad U^* \cdot U = E_{J \times J}.$$

- a) Zeigen Sie: Wenn U, V unitäre Matrizen sind, sodass das Produkt $U \cdot V$ definiert ist, dann ist $U \cdot V$ eine unitäre Matrix.
- b) Zeigen Sie, dass jede unitäre Matrix invertierbar ist, und dass die Inverse unitär ist.

Hinweis: Aus dem hier bewiesenen Sachverhalt ergibt sich (in einer Sprache, die später ausführlich entwickelt wird): Die Menge $\mathcal{U}(J)$ aller unitären $J \times J$ -Matrizen ist eine Gruppe.

Aufgabe I.4.3 : (Pauli-Matrizen, Multiplikationstafel)

Die folgenden komplexen 2×2 -Matrizen heissen die Pauli-Matrizen („Spin-Matrizen“)

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Vergewissern Sie sich, dass sie allesamt unitäre Matrizen sind, und berechnen Sie die Produkte $\sigma_i \cdot \sigma_j$ ($i, j = 1, 2, 3$). (‘Multiplikationstafel’)

Aufgabe I.4.4 : (Die Gruppe $\mathcal{SU}(1, 1)$)

Sei V eine komplexe 2×2 -Matrix mit

$$V \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} V^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \det V = 1.$$

Zeigen Sie : Es gibt komplexe Zahlen a, b mit

$$V = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}, \quad |a|^2 - |b|^2 = 1.$$

Zeigen Sie weiter: Die Menge dieser Matrizen, (die man üblicherweise mit $\mathcal{SU}(1, 1)$ bezeichnet) hat die Eigenschaften:

Das Produkt von Matrizen aus $\mathcal{SU}(1, 1)$ liegt in $\mathcal{SU}(1, 1)$;

jede Matrix in $\mathcal{SU}(1, 1)$ ist invertierbar und die Inverse gehört zu $\mathcal{SU}(1, 1)$.

Aufgabe I.4.5 :

J bezeichne die spezielle Matrix

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Zeigen Sie :

- a) Für jede komplexe 2×2 -Matrix A mit $\det A = 1$ gilt

$$A^T \cdot J \cdot A = J .$$

- b) Wenn für eine 2×2 -Matrix A gilt

$$\det A = 1 , \quad A^* \cdot J \cdot A = J$$

dann ist A eine reelle Matrix.

- c) Betrachten Sie

$$C = \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \quad C^{-1} = \frac{1}{2i} \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$$

Zeigen Sie :

Wenn A eine reelle Matrix ist, dann gibt es komplexe Zahlen a, b , sodass

$$C \cdot A \cdot C^{-1} = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}$$

Aufgabe I.4.6 : (Eine Faktorisierung)

Eine komplexe $n \times n$ -Matrix C heisst (bekanntlich) eine positiv semidefinite Matrix, wenn gilt

$$C = C^* , \quad x^* \cdot C \cdot x \geq 0 \quad \text{für alle Spalten } x .$$

Wenn C ausserdem nichtsingulär ist (d. h. $C \cdot x \neq 0$ für alle $x \neq 0$), dann heisst C eine positivdefinite Matrix.

1. Zeigen Sie: für jede beliebige Matrix A ist $C = A^* \cdot A$ positiv semidefinit; wenn $A \cdot x \neq 0$ für alle $x \neq 0$, dann ist C positiv definit.
2. Sei A eine $n \times n$ -Matrix mit $A^* \cdot A$ positiv definit; und sei R eine positivdefinite Matrix mit $R^2 = A^* \cdot A = (R^* \cdot R)$. Zeigen Sie:
Es existiert genau eine unitäre Matrix U , sodass gilt $A = U \cdot R$.
3. Seien A, R und U wie in 2). Zeigen Sie, dass R und U genau dann kommutieren, wenn A und A^* kommutieren.

Aufgabe I.4.6 :

Vorbemerkung: Für jede komplexe n -Spalte x definieren wir $\|x\| = \sqrt{x^*x}$.

Bemerge: Die $n \times n$ -Matrix A ist genau dann nichtsingulär, wenn A^*A positiv definit ist, also genau dann, wenn $\|Ax\|^2 > 0$ für alle $x \neq 0$.

Zeigen Sie

- a) Für jede hermitesche $n \times n$ -Matrix H ist $(H + iE)$ invertierbar. (E ist die Einheitsmatrix)

(**Hinweis:** $\|(H + iE)x\|^2 \geq \|x\|^2$)

- b) Wenn H hermitisch ist, dann ist

$$U := (H - iE)(H + iE)^{-1}$$

eine unitäre Matrix. Was bedeutet diese Aussage für $n = 1$?

- c) Klären Sie: Für welche unitären U gibt es eine Matrix H mit

$$U = (H - iE)(H + iE)^{-1} \quad ?$$

Ist H durch U eindeutig bestimmt?

(Klären Sie zuerst den Fall $n = 1$).

I.5 Vektoren sind Elemente eines Vektorraums. Basiswechsel

Vektoren sind Rechengrößen, die man **linear kombinieren** kann: Mit v und w ist auch $\alpha v + \beta w$ ein Vektor (für beliebige Zahlen α, β).

Vektoren kann man **nicht multiplizieren**.

In der Physik braucht man vor allem die reellen und die komplexen Vektorräume. Die Algebra betrachtet Vektorräume zu beliebigen Koeffizientenkörpern \mathbb{K} . Wir sprechen im Folgenden ebenfalls von \mathbb{K} -Vektorräumen, denken aber vor allem an \mathbb{R} -Vektorräume oder \mathbb{C} -Vektorräume. Die Rechenregeln in \mathbb{R} und in \mathbb{C} sind bekannt. Wir ersparen es uns, die Rechenregeln aufzulisten, die einen allgemeinen Körper \mathbb{K} auszeichnen. (Kurz gesagt: Ein Körper ist ein kommutativer Ring mit Einselement, in welchem jedes Element ausser der Null ein Inverses besitzt.) Unser Thema ist hier die **axiomatische Kennzeichnung** der \mathbb{K} -Vektorräume.

Definition

Ein \mathbb{K} -Vektorraum ist ein Bereich von Rechengrößen mit einem ausgezeichneten Element 0 , in welchem man addieren und „skalar“ mit Elementen von \mathbb{K} multiplizieren kann

$$(V, 0, +, \cdot).$$

Dabei gelten die folgenden Regeln der Addition

- (i) $u + 0 = u$ für alle $u \in V$
- (ii) $u + v = v + u$ für alle $u, v \in V$
- (iii) $(u + v) + w = u + (v + w)$ für alle $u, v, w \in V$
- (iv) Zu jedem $u \in V$ existiert ein Vektor $(-u)$ mit $u + (-u) = 0$

sowie die folgenden Regeln der skalaren Multiplikation

- (v) $\alpha(u + v) = \alpha u + \alpha v$ für $\alpha \in \mathbb{K}, u, v \in V$
- (vi) $(\alpha + \beta) \cdot u = \alpha u + \beta u$ für $\alpha, \beta \in \mathbb{K}, u \in V$
- (vii) $\alpha \cdot (\beta u) = (\alpha \cdot \beta) \cdot u$ für $\alpha, \beta \in \mathbb{K}, u \in V$
- (viii) $1 \cdot u = u$ für alle $u \in V$.

Bei der Kurzbeschreibung der folgenden Beispiele verwenden wir bereits Begriffe wie Basis und Dimension, die wir erst in V.1 von den Axiomen her begründen werden.

Beispiel 1

Sei I eine endliche Indexmenge, $|I| = m$. Die Menge der reellen I -Spalten ist ein \mathbb{R} -Vektorraum. Man bezeichnet ihn mit \mathbb{R}^I oder \mathbb{R}_{Sp}^I .

Addition $u + v$ erfolgt durch Addition der entsprechenden Einträge, die skalare Multiplikation $\alpha \cdot u$ besteht darin, dass man alle Einträge von u mit derselben reellen Zahl α multipliziert.

Andererseits ist auch die Menge der I -Zeilen ein Vektorraum. Wir bezeichnen ihn mit \mathbb{R}_Z^I . Beide Vektorräume \mathbb{R}^I sind Vektorräume der reellen Dimension $|I| = m$.

Beispiel 2

Die Menge $\mathbb{C}([z])$ aller Polynome mit komplexen Koeffizienten ist ein unendlichdimensionaler \mathbb{C} -Vektorraum.

Die Menge V aller komplexen Polynome vom Grad $< n$ ist ein \mathbb{C} -Vektorraum der komplexen Dimension n .

Beispiel 3

Eine komplexe $I \times I$ -Matrix A heißt eine hermitesche Matrix, wenn gilt

$$a_{ij} = a_{ji}^* \text{ für alle } (i, j) \in I \times I.$$

Die Menge dieser Matrizen ist ein reeller Vektorraum, aber kein komplexer Vektorraum. Es liegt auf der Hand, dass die (reelle!) Dimension $= m^2$ ist.

Beispiel 4

Eine reelle $I \times I$ -Matrix A heißt eine symmetrische Matrix, wenn gilt

$$a_{ij} = a_{ji} \text{ für alle } (i, j) \in I \times I.$$

Die Menge V dieser Matrizen ist ein reeller Vektorraum der Dimension $\frac{m(m+1)}{2}$.

Bemerke

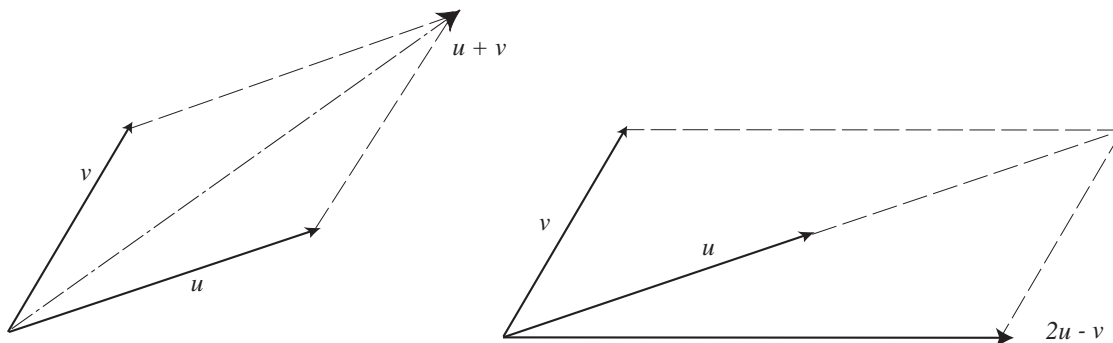
Einen beliebigen \mathbb{C} -Vektorraum kann man zu einem \mathbb{R} -Vektorraum machen, indem man die skalare Multiplikation auf \mathbb{R} einschränkt. Die reelle Dimension ist zweimal die komplexe Dimension.

Beispiel 5

Die Menge der Translationen des dreidimensionalen Anschauungsraums ist ein dreidimensionaler \mathbb{R} -Vektorraum. Die Elemente nennt man in der Schule üblicherweise „Pfeilklassen“. Die Addition nennt man die Vektoraddition. Mit diesem (möglicherweise aus dem Schulunterricht bekannten) Beispiel wollen wir uns etwas näher beschäftigen. Wir machen hier einen ersten Schritt in die

Affine Geometrie

1. Eine Translation des zweidimensionalen Anschauungsraums nennen wir einen Verschiebungsvektor. Verschiebungsvektoren kann man linear kombinieren.



2. Wir zeichnen einen Punkt P_0 im Anschauungsraum aus. Den Punkt P , welcher durch die Verschiebung w entsteht, bezeichnen wir

$$P = P_0 + w .$$

Wir lesen die Formel so: w angewandt auf P_0 liefert P .

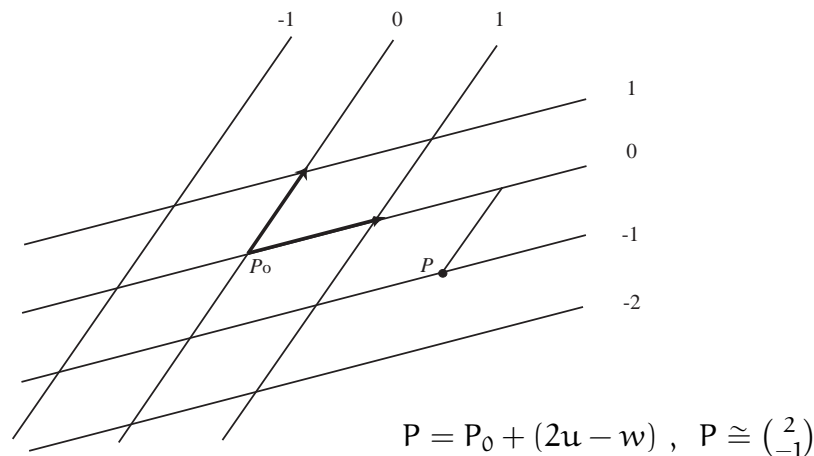
Wir bemerken :

$$(P_0 + w_1) + w_2 = P_0 + (w_1 + w_2) = P_0 + (w_2 + w_1) = (P_0 + w_2) + w_1 .$$

Die Verschiebungen werden addiert.

3. Sei V der Vektorraum der Verschiebungsvektoren. Jedes Paar linear unabhängiger Vektoren $\{u, v\}$ bildet eine Basis. Das soll heißen: In jedem $w \in V$ gibt es genau ein Paar von Zahlen α, β , sodass $w = \alpha u + \beta v$.
4. Sei L die Menge der Punkte P des Anschauungsraums. Die Punkte nennt man auch Ortsvektoren. Bemerke: Ortsvektoren kann man nicht addieren. Ortsvektoren (d.h. Punkte) sind dazu da, verschoben zu werden.
Wenn ein Punkt $P_0 \in L$ und eine Basis von V ausgezeichnet ist, dann entspricht jedem $P \in L$ ein Zahlenpaar. Es gibt genau ein Zahlenpaar α, β , sodass gilt

$$P = P_0 + (\alpha u + \beta v) .$$



Dieses Paar nennt man die Koordinaten von P bzgl. des von P_0 und $\{u, v\}$ erzeugten **affinen Koordinatensystems**.

5. Im dreidimensionalen Anschauungsraum ist jedes linear unabhängige Tripel von Verschiebungsvektoren eine Basis des Vektorraums aller Translationen. Der Anschauungsraum selber ist eine Punktmenge, welche der Menge aller Zahlentripel entspricht, wenn man eine Punkt P_0 und eine Basis von V ausgezeichnet hat.

Didaktische Hinweise

Der Begriff des Vektors ist vielen schon im Schulunterricht begegnet, allerdings vermutlich in derart verkürzter Form, dass wir darauf nicht aufbauen können. In schulnaher Sprache schreibt Feynman in Band I, Kapitel 11 :

Alle Größen, die wie ein Schritt im Raum eine Richtung besitzen, werden Vektoren genannt. Ein Vektor besteht aus drei Zahlen. Um einen Schritt im Raum darzustellen, sagen wir vom Anfangspunkt zu einem besonderen Punkt P, benötigen wir wirklich drei Zahlen [...].

Dieser Auffassung von dem, was ein Vektor ist, können wir uns nicht anschließen. Wir haben eine Reihe von Gründen vorzubringen.

Feynman unterschlägt hier, dass es neben den Vektoren (im Sinne eines Schritts) auch noch die Covektoren gibt. Kräfte und Impulse sind Covektoren. Die Unterscheidung zwischen Vektoren und Covektoren (und die Idee der dualen Vektorräume) lernt der Physiker in der höheren Mechanik z.B. beim Übergang von den Lagrange-Gleichungen zu den Hamilton-Gleichungen.

Feynman unterschlägt, dass man erst einmal eine Basis wählen muss, wenn man die Verschiebungsvektoren („Schritte“) mit Zahlentripeln identifizieren will. In der Newton'schen Mechanik nimmt man in der Regel orthonormierte Basen. Diese spezielle Wahl korrespondiert mit der Symmetrie der Newton'schen Gesetze.

Im Zitat bindet Feynman den Vektorbegriff an den dreidimensionalen Anschauungsraum. Entsprechend hat man in der speziellen Relativitätstheorie die „Vierervektoren“. Im vierdimensionalen Raum-Zeitkontinuum bevorzugt man dann natürlich solche Basen, die zur relativistischen Invarianz der physikalischen Gesetze passen.

Feynman ist übrigens nicht wirklich der Meinung, dass man den Physik-Anfängern den Standpunkt der axiomatischen Mathematik nicht zumuten könne. Im Gegenteil; schon im Kapitel 12.1 „Was ist eine Kraft“ lesen wir:

In der Tat ist es das Herrliche der Mathematik, dass wir nicht sagen müssen, wovon wir reden. Das Herrliche liegt darin, dass die Gesetze, die Argumente und die Logik unabhängig davon sind, was „es“ ist. Wenn wir irgendeinen anderen Satz von Objekten haben, der das gleiche System von Axiomen befolgt wie die Euklidische Geometrie und wenn wir dann neue Definitionen machen und diese mit der korrekten Logik verfolgen, so werden die Folgerungen korrekt sein, und es macht keinen Unterschied, was der Gegenstand war.

Die algebraische Vektorraumtheorie hat einen sehr weiten Horizont, der sich für die Physikstudierenden erst nach und nach mit interessanten Objekten füllt. Die Vektorraumstruktur spielt insbesondere dort eine wesentliche Rolle, wo es ein Superpositionsprinzip gibt, also z.B. in der Theorie der Wellen und in der Quantenmechanik. Die quantenmechanischen Zustände werden ket-Vektoren genannt. (Dual dazu gibt es die bra-Vektoren). Und diese Vektoren haben gewiß nichts mit Schritten im Raum zu tun. Sie lassen sich auch nur in seltenen Fällen durch Zahlentripel darstellen.

Abstrakte Vektorräume.

Mit dem aktuellen Kapitel betreten wir den Bereich der axiomatisch aufgebauten Mathematik. Auf die Frage: Was ist ein Vektor? muss die Antwort lauten: Ein Vektor ist ein Element eines Vektorraums. Ein Vektorraum ist ein Bereich von Rechengrößen, in welchem nach Maßgabe der Vektorraum-Axiome linear kombiniert wird.

In unseren Beispielen haben wir den Begriff der Dimension eines \mathbb{K} -Vektorraums in intuitiver Weise gebraucht: Wieviele Skalare (d.h. Elemente von \mathbb{K}) braucht man um einen Vektor zu identifizieren? Ein Mathematiker, der den Horizont für alle Anwendungen offen halten will, kann mit dem intuitiven Gebrauch des Dimensionsbegriffs nicht zufrieden sein. Wir benötigen einen gesicherten Platz für den Begriff der Dimension. Einen solchen sicheren Platz liefert der folgende Satz. Man könnte ihn als einen ersten Hauptsatz der Vektorraumtheorie bezeichnen. Wir werden ihn an geeigneter Stelle beweisen (siehe Abschnitt V.1 ‘Basis und Dimension’). Wir wollen aber zunächst Erfahrungen mit der Anwendung dieses Satzes sammeln.

Satz (“Vektorräume haben eine Dimension”)

Jeder \mathbb{K} -Vektorraum V besitzt eine Dimension; sie kann endlich oder unendlich sein. In einem endlichdimensionalen \mathbb{K} -Vektorraum gibt es Vektorraumbasen. Eine Basis ist ein Tupel von Vektoren $\{v_j : j \in J\}$, sodass jedes $v \in V$ auf genau eine Weise aus den v_j linear kombiniert werden kann.

$$v = \sum_{j \in J} v_j \cdot \alpha^j.$$

Alle Basen von V haben dieselbe Länge; diese Länge $|J|$ heißt die Dimension von V . Jedes linear unabhängige System $\{u_i : i \in I\}$ der Länge $|I| = |J| = \dim V$ ist eine Basis.

Zur Erinnerung : ein System von Vektoren $\{u_i : i \in I\}$ heißt linear unabhängig, wenn keines der u_i als Linearkombination der übrigen dargestellt werden kann.

Beispiel 1

Die Menge V der Polynome vom Grad $< n$ ist ein n -dimensionaler \mathbb{C} -Vektorraum. Für jedes $z_0 \in \mathbb{C}$ ist $\{1, z - z_0, (z - z_0)^2, \dots, (z - z_0)^{n-1}\}$ eine Basis.

Beispiel 2

Sei $q(z)$ ein Polynom vom Grade n . Die Menge aller „echten“ Brüche $\frac{p(z)}{q(z)}$ ($\deg(p) < n$) ist ein n -dimensionaler \mathbb{C} -Vektorraum.

Wenn $q(z)$ n verschiedene Nullstellen hat

$$q(z) = (z - z_1)(z - z_2) \dots (z - z_n) \quad (z_j \text{ paarweise verschieden}),$$

dann ist $\left\{ \frac{1}{z - z_1}, \frac{1}{z - z_2}, \dots, \frac{1}{z - z_n} \right\}$ eine Basis. Im allgemeinen Fall

$$q(z) = (z - z_1)^{\ell_1} \cdot (z - z_2)^{\ell_2} \cdot \dots \cdot (z - z_m)^{\ell_m} \quad (z_j \text{ paarweise verschieden})$$

ist das folgende System eine Basis

$$\left\{ \left(\frac{1}{z - z_1} \right)^{k_1} : 1 \leq k_1 \leq \ell_1 \right\} \cup \dots \cup \left\{ \left(\frac{1}{z - z_m} \right)^{k_m} : 1 \leq k_m \leq \ell_m \right\}$$

Die Matrizen eines Basiswechsels

Im n -dimensionalen \mathbb{C} -Vektorraum V seien zwei Basen gegeben :

$$\{\mathbf{u}_i : i \in I\} \quad \text{und} \quad \{\mathbf{v}_j : j \in J\}.$$

Der Übersichtlichkeit halber bezeichnen wir hier die Vektoren mit fetten Buchstaben.

Basiswechsel Wir können die Basisvektoren \mathbf{u}_i in eindeutigerweise als Linearkombinationen der Basisvektoren \mathbf{v}_j darstellen, und umgekehrt

$$\sum_j \mathbf{v}_j \cdot b_i^j = \mathbf{u}_i \quad \text{für alle } i; \quad \sum_i \mathbf{u}_i \cdot a_j^i = \mathbf{v}_j \quad \text{für alle } j.$$

In Matrizenschreibweise

$$(\mathbf{u}_1, \dots, \mathbf{u}_n) \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix} = (\mathbf{v}_1, \dots, \mathbf{v}_n); \quad (\mathbf{v}_1, \dots, \mathbf{v}_n) \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix} = (\mathbf{u}_1, \dots, \mathbf{u}_n).$$

Wir zeigen, dass die $I \times J$ -Matrix A und die $J \times I$ -Matrix B zueinander invers sind.

$$\begin{aligned} (\mathbf{u}_1, \dots, \mathbf{u}_n) A \cdot B &= (\mathbf{v}_1, \dots, \mathbf{v}_n) B = (\mathbf{u}_1, \dots, \mathbf{u}_n); \\ (\mathbf{v}_1, \dots, \mathbf{v}_n) B \cdot A &= (\mathbf{u}_1, \dots, \mathbf{u}_n) A = (\mathbf{v}_1, \dots, \mathbf{v}_n). \end{aligned}$$

Da jedes \mathbf{u}_i nur in trivialer Weise in der \mathbf{u} -Basis dargestellt werden kann, und jedes \mathbf{v}_j nur in trivialer Weise in der \mathbf{v} -Basis, ergibt sich

$$A \cdot B = E_{I \times I}; \quad B \cdot A = E_{J \times J}.$$

Koordinatenwechsel Jeder Vektor \mathbf{w} ist eine Linearkombination der Basisvektoren mit eindeutig bestimmten Koeffizienten

$$\sum_i \mathbf{u}_i \cdot y^i = \mathbf{w} = \sum_j \mathbf{v}_j \cdot x^j.$$

Die Koeffizienten y^i ergeben sich in linearer Weise aus den Koeffizienten x^j und zwar vermittelt der $I \times J$ -Matrix A . Entsprechendes gilt für die Rückabbildung.

$$\begin{aligned} \sum_j \mathbf{v}_j \cdot x^j = \mathbf{w} &= \sum_i \sum_j \mathbf{v}_j \cdot b_i^j \implies \sum_i b_i^j y^i = x^j; \\ \sum_i \mathbf{u}_i \cdot y^i = \mathbf{w} &= \sum_j \sum_i \mathbf{u}_i \cdot a_j^i \implies \sum_j a_j^i x^j = y^i. \end{aligned}$$

In Matrizennotation

$$\begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix} \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} ; \quad \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} .$$

Wir fassen die **u**-‘Koordinaten’ des Vektors **w** zur I-Spalte **y** zusammen, und seine **v**-‘Koordinaten’ zur J-Spalte **x**, und haben dann

$$A \cdot x = y ; \quad B \cdot y = x .$$

Die Matrizen **A** und **B** = **A**⁻¹ sind somit die Matrizen der Koordinatenwechsels, ebenso wie sie die Matrizen des Basiswechsels sind.

Tableau-Schreibweise Die Gleichung **A** · **x** = **y** kann man so lesen: Die Spalte **y** ergibt sich als Linearkombination der Spalten von **A** mit den Koeffizienten **x**¹, ..., **x**ⁿ. Das notiert man manchmal in einem sog. Spaltentableau

$$\begin{array}{ccc} x^1 & \dots & x^n \\ \boxed{\begin{array}{ccc} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{array}} & = & \begin{array}{c} y^1 \\ \vdots \\ y^n \end{array} \end{array}$$

Die Gleichungen **v**_j = ∑_i **u**_i · **a**_jⁱ notiert man manchmal in einem sog. Zeilentableau.

$$\begin{array}{ccc} \mathbf{u}_1 & \boxed{\begin{array}{ccc} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{array}} & \\ \vdots & & \\ \mathbf{u}_n & & \\ & = \mathbf{v}_1 & \dots = \mathbf{v}_n \end{array}$$

Die Tableauschreibweise werden wir in den Unterabschnitten V.2 und V.3 verallgemeinern und weiterentwickeln.

Beispiele für Basiswechsel

Im 3-dimensionalen **C**-Vektorraum **V** der quadratischen Polynome betrachten wir die Basen

$$\{1, z, z^2\} \quad \text{und} \quad \{1, (z-1), (z-1)^2\} .$$

Jedes quadratische Polynom **p**(**z**) hat Darstellungen

$$a_0 + a_1 z + a_2 z^2 = p(z) = b_0 + b_1(z-1) + b_2(z-1)^2$$

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} ; \quad \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

Es gilt nämlich

$$(1, z, z^2) \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = (1, z-1, (z-1)^2) ; \quad (1, z-1, (z-1)^2) \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = (1, z, z^2) .$$

Hinweis : Die Matrizen eines Koordinatenwechsels sind ein wichtiges Thema in der Differentialrechnung mehrerer Veränderlicher. Diese Matrizen heißen dort die Jacobi-Matrizen eines Koordinatenwechsels.

Didaktischer Hinweis

Wer konkrete Polynome in eine neue Basis umrechnen will, der wird i. d. Regel nicht zuerst die Matrizen A und $B = A^{-1}$ des Basiswechsels ausrechnen und dann die Matrizen auf das konkrete Koeffizienten-Tupel anwenden. Wir haben beispielsweise in I.3 gesehen, wie man gemäß dem Hornerschema ein Polynom mit wenig Rechenoperationen umzentriert. Das iterativ arbeitende Horner-Schema verarbeitet konkrete Koeffizienten-Tupel; es dient einem anderen Zweck als das matrizentheoretische Schema des Basiswechsels. Ein Rechenschema ist hier wie auch häufig sonst etwas anderes als eine allgemeine Formel (in die man potentiell etwas ‘einsetzen’ kann).

Warnung : Vektoren kann man linear kombinieren, aber nicht multiplizieren. Wenn man es dennoch tut, dann nicht aufgrund der Vektorraumstruktur. In der Regel ist das Produkt dann auch ein Objekt ausserhalb des Vektorraums. Man kann z.B. affine Funktionen (punktweise) multiplizieren; das Resultat ist aber keine affine sondern eine quadratische Funktion. Wenn man Vektoren vermöge des Tensorprodukts miteinander multipliziert, dann erhält man keinen Vektor sondern einen Tensor zweiter Stufe. Wir halten es nicht für geschickt, wenn man den Anfängern in der sog. Vektoranalysis sagt, man könne Vektoren im dreidimensionalen euklidischen Raum mittels des äusseren Produkts multiplizieren und das Ergebnis sei ein Vektor. Irgendwann später muss man nämlich doch einräumen, dass das „äussere Produkt“ aus einem Paar von „polaren“ Vektoren einen „axialen“ Vektor macht. Das „äussere Produkt“ ist eine Konstruktion, die für gewisse spezielle Rechnungen nützlich ist, dann aber mathematisch nicht weiterführt. Wir überlassen seine Behandlung der Vorlesung über Mechanik (wenn sie dort angebracht erscheinen sollte).

Aufgaben zu I.5

Aufgabe I.5.1 : (Pauli-Matrizen, Spur und Determinante)

Wir betrachten die Menge aller komplexen 2×2 -Matrizen als einen vierdimensionalen komplexen Vektorraum.

- a) Zeigen Sie, dass die Pauli-Matrizen $\sigma_1, \sigma_2, \sigma_3$ zusammen mit der Einheitsmatrix σ_0 eine Basis bilden

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- b) Sei $A = a_0\sigma_0 + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$.

Zeigen Sie :

- (i) $A^* = \bar{a}_0 \cdot \sigma_0 + \bar{a}_1\sigma_1 + \bar{a}_2\sigma_2 + \bar{a}_3\sigma_3$
- (ii) $\text{trace}A = 2a_0$
- (iii) $\det A = a_0^2 - (a_1^2 + a_2^2 + a_3^2)$

- c) Die reellen Linearkombinationen

$$\alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3$$

bilden einen dreidimensionalen reellen Vektorraum \mathcal{H} . Die Zuordnung

$$\mathcal{H} \ni A \longmapsto \|A\| = (-\det A)^{1/2}$$

liefert die euklidische Norm auf \mathcal{H} .

Zeigen Sie : Für jedes ($U \in \text{SU}(2)$) (d.h. $U^*U = E$, $\det U = 1$) gilt

$$\|U^*AU\| = \|A\| \quad \text{für alle } A \in \mathcal{H}.$$

(*Hinweis* : Beachten Sie $\text{trace}(A \cdot B) = \text{trace}(B \cdot A)$.)

Aufgabe I.5.2 : (Pauli-Matrizen, Eigenwerte)

Die Menge \mathcal{M} aller hermiteschen 2×2 -Matrizen ist ein vierdimensionaler reeller Vektorraum.

Zeigen Sie :

- a) Die Pauli-Matrizen $\sigma_1, \sigma_2, \sigma_3$ zusammen mit der Einheitsmatrix σ_0 bilden eine Basis.
- b) Die Pauli-Matrizen haben die Eigenschaften, dass sie die Eigenwerte $+1$ und -1 haben

$$\det(\sigma_j) = -1, \quad \text{trace}(\sigma_j) = 0 \quad \text{für } j = 1, 2, 3.$$

c) Finden Sie weitere (oder alle) hermitesche 2×2 -Matrizen

$$A = \alpha_0 \sigma_0 + \alpha_1 \sigma_1 + \alpha_2 \sigma_2 + \alpha_3 \sigma_3 \quad ,$$

welche die Eigenwerte $+1$ und -1 besitzen.

Zur Erinnerung :

$$1. \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

2. Die Eigenwerte der 2×2 -Matrix A sind die Nullstellen des Polynoms

$$\det(A - z \cdot e) = z^2 - z \cdot \text{trace}(A) + \det A \quad .$$

Aufgabe I.5.3 : (Quaternionen)

Seien $\sigma_1, \sigma_2, \sigma_3$ die Pauli-Matrizen und

$$j = \frac{1}{i} \sigma_1 = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} ; \quad k = \frac{1}{i} \sigma_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} ; \quad \ell = \frac{1}{i} \sigma_3 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} .$$

e sei die Einheitsmatrix.

a) Zeigen Sie : Jede Matrix q von der Gestalt

$$q = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

läßt sich als reelle Linearkombination dieser Matrizen darstellen.

$$q = \alpha_0 \cdot e + \alpha_1 \cdot j + \alpha_2 \cdot k + \alpha_3 \cdot \ell .$$

Die Menge dieser q bezeichnen wir mit \mathbb{H} („Hamiltons Quaternionen“).

b) Jede Matrix dieser Art mit Ausnahme der Nullmatrix besitzt eine Inverse. Zum Beweis zeigen Sie :

$$q^* \cdot q = (\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2) \cdot e .$$

Man schreibt

$$q^* \cdot q = \|q\|^2 \cdot e .$$

c) Die Linearkombinationen von j, k, ℓ sind dadurch ausgezeichnet, dass sie die Spur $= 0$ haben. Wir nennen sie die reinen Quaternionen. Der dreidimensionale Vektorraum der reinen Quaternionen wird mit \mathbb{H}_0 bezeichnet.

$$\begin{aligned} \text{Sei } q &= \cos\left(\frac{\varphi}{2}\right) \cdot e + \sin\left(\frac{\varphi}{2}\right) \cdot (\alpha_1 j + \alpha_2 k + \alpha_3 \ell) \\ \text{mit } \alpha_1^2 + \alpha_2^2 + \alpha_3^2 &= 1 . \end{aligned}$$

Zeigen Sie : Für $x \in \mathbb{H}_0$ gilt $q^* \cdot x \cdot q \in \mathbb{H}_0$ und

$$q^* \cdot x \cdot q = x \Leftrightarrow x = \text{const} \cdot (\alpha_1 j + \alpha_2 k + \alpha_3 \ell) .$$

I.6 Minimalpolynome. Verallgemeinerte Eigenvektoren.

Dieser Unterabschnitt kombiniert die Rechenregeln für Polynome mit den Rechenregeln für quadratische Matrizen. Ausserdem kommt der Vektorraum aller n -Spalten ins Spiel.

Definition

1. Zu einer $n \times n$ -Matrix A betrachten wir die Potenzen $A^0 = I, A^1 = A, A^2, A^3, \dots$. Für jedes Polynom $p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ definieren wir

$$p(A) = a_n A^n + a_{n-1} A^{n-1} + \dots + a_1 A + a_0 I.$$

2. Es sei $m = m(A)$ die kleinste Zahl, für welche die Matrizen A^0, A^1, \dots, A^m linear abhängig sind.

$$A^m + a_{m-1} \cdot A^{m-1} + \dots + a_1 \cdot A + a_0 \cdot I = 0;$$

Das normierte Polynom

$$p(z) = p_A(z) = z^m + a_{m-1} \cdot z^{m-1} + \dots + a_1 \cdot z + a_0$$

nennen wir das **Minimalpolynom zur Matrix A** .

3. Es sei v eine n -Spalte, und sei $k = k(A, v)$ die kleinste Zahl, für welche die ‘Vektoren’ v, Av, A^2v, \dots, A^kv linear abhängig sind

$$A^k v + b_{k-1} \cdot A^{k-1} v + \dots + b_1 \cdot Av + b_0 \cdot v = 0.$$

Das normierte Polynom

$$p_v(z) = p_{A,v}(z) = z^k + b_{k-1} \cdot z^{k-1} + \dots + b_1 \cdot z + b_0$$

nennen wir das **Minimalpolynom zu v** . (A ist fest.)

Satz

Es sei p_A das Minimalpolynom zur Matrix A und p_v das Minimalpolynom zum Vektor v . Es gilt dann

1. Jedes Polynom \tilde{p} mit $\tilde{p}(A) = 0$ (Nullmatrix) ist ein Vielfaches von p_A .
2. Jedes Polynom \tilde{p} mit $\tilde{p}(A)v = 0$ (Nullvektor) ist ein Vielfaches von p_v .
3. Das Minimalpolynom p_A ist das kleinste gemeinsame Vielfache der p_v .

Vorbemerkung über Teilbarkeit

Für normierte Polynome mit Koeffizienten in einem beliebigen Körper \mathbb{K} gibt es ‘größte gemeinsame Teiler’ (‘greatest common divisor’)(gcd) und ‘kleinste gemeinsame Vielfache’ (‘least common multiple’)(lcm). Zu jedem Paar von normierten Polynomen p und q gibt es normierte Polynome $a = \gcd(p, q)$ und $b = \text{lcm}(p, q)$, charakterisiert durch die Eigenschaften

$$a = \gcd(p, q) \iff \left\{ \begin{array}{l} a \mid p, \quad a \mid q \\ \forall c \quad (c \mid p, c \mid q) \Rightarrow c \mid a \end{array} \right\}$$

$$b = \text{lcm}(p, q) \iff \left\{ \begin{array}{l} p \mid b, \quad q \mid b \\ \forall c \quad (p \mid c, q \mid c) \Rightarrow b \mid c \end{array} \right\}$$

Mit dem Euklidischen Algorithmus findet man Polynome m, n , sodass $a = m \cdot p + n \cdot q$.

Beweis des Satzes

1. p_A ist das Polynom vom kleinsten Grad, welches A annulliert. Wenn $\tilde{p}(A) = 0$, dann auch $a(A) = \gcd(\tilde{p}, p_A)(A) = 0$ mit $\deg a \leq \deg p_A$. Also $a = p_A$.
2. p_v ist das Polynom vom kleinsten Grad mit $p(A)v = 0$. Wenn $\tilde{p}(A)v = 0$, dann auch $a(A)v = \gcd(\tilde{p}, p_v)(A)v = 0$ mit $\deg a \leq \deg p_v$. Also $a = p_v$.
3. Sei \tilde{p} ein Polynom, welches Vielfaches von jedem p_v ist. Es gilt dann $\tilde{p}(A)v = 0$ für alle v , also $\tilde{p}(A) = 0$ und daher $p_A \mid \tilde{p}$.

Eine unmittelbare Folgerung ist der

Satz

Wenn das Minimalpolynom zum Vektor v den Grad m hat, dann sind die Vektoren $v, Av, A^2v, \dots, A^{m-1}v$ linear unabhängig. Sei W_v der von ihnen aufgespannte Vektorraum. Seine Elemente sind von der Gestalt $w = \sum_0^{m-1} c_j A^j v = q(A)v$ mit einem eindeutig bestimmten Polynom $q(\cdot)$ vom Grad $< m$. Jedes $q(A)$ bildet den Vektorraum W_v in sich ab; es gilt also $q(A)W_v \subseteq W_v$.

Die Matrizen $q(A)$ als Endomorphismen

Für das Folgende ist es manchmal bequem, wenn man die Matrizen $q(A)$ abstrakt als Endomorphismen eines n -dimensionalen \mathbb{K} -Vektorraums versteht. (Unten vorzugsweise $\mathbb{K} = \mathbb{C}$.)

1. Jede $J \times J$ -Matrix entspricht einer linearen Abbildung des Spaltenraums \mathbb{K}_{Sp}^J in sich

$$A \rightsquigarrow \varphi_A(\cdot) \quad \varphi_A : \mathbb{K}_{Sp}^J \longrightarrow \mathbb{K}_{Sp}^J, \quad \varphi_A(v) = A \cdot v.$$

2. Eine lineare Abbildung φ eines Vektorraums V in sich nennt man einen Endomorphismus von V . Wählt man in V eine Basis $\{v_j : j \in J\}$, so wird der Endomorphismus durch eine $J \times J$ -Matrix dargestellt.

$$(\varphi(v_1), \varphi(v_2), \dots, \varphi(v_n)) = (v_1, v_2, \dots, v_n) \cdot A; \quad \varphi(v_j) = \sum_i v_i \cdot a_j^i$$

$$v = \sum v_j \cdot x^j \implies \varphi(v) = \sum v_j \cdot y^j \quad \text{mit} \quad A \cdot x = y, \quad \sum a_j^i \cdot x^j = y^i.$$

3. Wechselt man zu einer anderen Basis $\{u_i : i \in I\}$ und ist B die $J \times I$ -Matrix des Basiswechsels, so wird der Endomorphismus in der neuen Basis durch die $I \times I$ -Matrix $\tilde{A} = B^{-1}AB$ dargestellt.

$$u_i = \sum v_j \cdot b_i^j; \quad v_j = \sum u_i \cdot \tilde{b}_j^i \quad \text{mit} \quad \tilde{B} = B^{-1};$$

$$\varphi(u_i) = \varphi\left(\sum v_j \cdot b_i^j\right) = \sum \varphi(v_j) \cdot b_i^j = \sum_{jk} v_k \cdot a_j^k b_i^j = \sum_{ijk} u_i \cdot \tilde{b}_k^l a_j^k b_i^j = \sum u_l \cdot \tilde{a}_i^l.$$

4. Zwei Matrizen A und \tilde{A} heissen ähnlich, wenn es eine invertierbare Matrix B gibt, sodass $\tilde{A} = B^{-1}AB$. Ähnliche Matrizen haben offenbar dasselbe Minimalpolynom. Ähnliche Matrizen kann man auffassen als Darstellungen desselben Endomorphismus in verschiedenen Basen.

Satz

Das Minimalpolynom $p_A(z)$ einer Matrix A sei als Produkt zweier teilerfremder Polynome gegeben.

$$p_A(z) = p_1(z) \cdot p_2(z) \quad \text{mit} \quad \gcd(p_1, p_2) = 1.$$

Es sei $\varphi_1 = p_1(A)$, $\varphi_2 = p_2(A)$ und für $i = 1, 2$

$$U_i = \{v : \varphi_i(v) = 0\} = \ker \varphi_i; \quad W_i = \{w : w = \varphi_i(v)\} = \text{im } \varphi_i.$$

Es gilt dann

$$W_1 = U_2, \quad W_2 = U_1. \quad W_1 \oplus W_2 = V.$$

Für jedes $\psi = q(A)$ gilt $\psi(W_i) \subseteq W_i$.

Beweis

1. Für jedes $\psi = q(A)$ gilt $\psi \circ \varphi_i = \varphi_i \circ \psi$. Daraus ergibt sich sofort

$$\psi(U_i) \supseteq U_i, \quad \psi(W_i) \subseteq W_i.$$

Aus $\varphi_1 \circ \varphi_2 = p_1(A) \cdot p_2(A) = 0$ ergibt sich $W_1 \subseteq U_2$ und $W_2 \subseteq U_1$.

2. Wir zeigen $U_1 \cap U_2 = \{0\}$. Jedes $u \neq 0$ liefert ein Minimalpolynom $p_u(z)$, (welches $p_1(z) \cdot p_2(z)$ teilt). Wenn $\varphi_i(u) = 0$, also $u \in U_i$, dann teilt $p_u(z)$ das Polynom $p_i(z)$. Da p_1 und p_2 teilerfremd sind, kann p_u nicht beide Polynome teilen. Es existiert also kein $u \neq 0$ in $U_1 \cap U_2$.

3. Wir zeigen $W_1 + W_2 = V$, m. a. Worten: jedes v besitzt mindestens eine Darstellung $v = \varphi_1(v_1) + \varphi_2(v_2)$. Da die Polynome p_i teilerfremd sind, gibt es Polynome r_i sodass $r_1(z)p_1(z) + r_2(z)p_2(z) = 1$. Daraus ergibt sich $r_1(A) \circ \varphi_1 + r_2(A) \circ \varphi_2 = \text{Id}$; und daher für alle v

$$v = r_1(A)\varphi_1(v) + r_2(A)\varphi_2(v) = \varphi_1 r_1(A)v + \varphi_2 r_2(A)v = \varphi_1(v_1) + \varphi_2(v_2) = w_1 + w_2.$$

Dies ist eine Zerlegung $v = u_2 + u_1$ mit $u_i \in U_i$ und zwar wegen 2) die einzige solche Zerlegung.

Die Eigenwerte eines Endomorphismus

Es sei A eine quadratische Matrix. Eine Spalte w , die nicht die Nullspalte ist, heisst ein Eigenvektor von A zum Eigenwert λ , wenn $A \cdot w = \lambda w$ gilt. Wir werden sehen, dass die Nullstellen des Minimalpolynoms p_A die Eigenwerte der Matrix A sind. Wir werden auch sehen, wie die Vielfachheit einer Nullstelle des Minimalpolynoms p_A zu interpretieren ist.

Definition (Verallgemeinerte Eigenvektoren zum Eigenwert λ)

Eine Spalte w nennen wir einen verallgemeinerten Eigenvektor der Höhe h zum Eigenwert λ , wenn gilt

$$(A - \lambda I)^h w = 0, \quad (A - \lambda I)^{h-1} w \neq 0.$$

Bemerkungen

Es sei A eine $n \times n$ -Matrix. Ein Skalar λ ist genau dann ein Eigenwert, wenn er eine Nullstelle des Minimalpolynoms ist, wenn also gilt $p_A(z) = (z - \lambda)^{h_\lambda} \cdot q_\lambda(z)$ mit einem $h_\lambda \geq 1$ (und o. B. d. A. $q_\lambda(\lambda) \neq 0$).

1. Die Menge V_λ aller verallgemeinerten Eigenvektoren zum Eigenwert λ ist ein Vektorraum. Seine Dimension $l_\lambda = \dim V_\lambda$ heisst die algebraische Vielfachheit von λ . Die Dimension $l_\lambda^{(1)}$ des Vektorraums aller Eigenvektoren heisst üblicherweise die geometrische Vielfachheit des Eigenwerts λ . Wir werden unten an einem Beispiel sehen, dass man weder die Dimension l_λ noch die Dimension $l_\lambda^{(1)}$ am Minimalpolynom ablesen kann.
2. Ist v ein verallgemeinerter Eigenvektor der Höhe h zum Eigenwert λ , so ist sein Minimalpolynom von der Form $p_v = (z - \lambda)^h$. Wegen $(A - \lambda I)^h v = 0$ ist nämlich p_v ein Teiler von $(z - \lambda)^h$; wegen $(A - \lambda I)^{h-1} v \neq 0$ ist p_v kein echter Teiler.
3. Wenn das Minimalpolynom p_A im Punkt λ von der Ordnung h verschwindet, $p_A(z) = (z - \lambda)^h \cdot q_\lambda(z)$ mit $q_\lambda(\lambda) \neq 0$, dann gibt es einen verallgemeinerten Eigenvektor von der Höhe h , (aber offenbar keinen von der Höhe $h + 1$). Es existiert nämlich ein Vektor w mit $0 \neq (A - \lambda I)^{h-1} q_\lambda(A)w = (A - \lambda I)^{h-1} v$, und $(A - \lambda I)^h v = 0$. Der Vektor v ist also ein verallgemeinerter Eigenvektor der Höhe h . Die Vektoren $v, Av, \dots, A^{h-1}v$ sind linear unabhängige Vektoren in V_λ . Also gilt $h \leq \dim V_\lambda$.

4. Wie wir gesehen haben ist die Vielfachheit h der Nullstelle λ im Minimalpolynom $p_A(z)$ die maximale Höhe eines verallgemeinerten Eigenvektors zum Eigenwert λ . Für $i = 1, 2, \dots, h$ sei

$$V_\lambda^{(i)} = \{v : (A - \lambda I)^i v = 0\}; \quad l_\lambda^{(i)} = \dim V_\lambda^{(i)}.$$

$l_\lambda^{(i)}$ ist also die Dimension des Vektorraums aller verallgemeinerten Eigenvektoren mit Höhe $\leq i$. Es gilt dann

$$V_\lambda = V_\lambda^{(h)} \supsetneq V_\lambda^{(h-1)} \supsetneq V_\lambda^{(h-2)} \supsetneq \dots \supsetneq V_\lambda^{(1)} \supsetneq \{0\};$$

$$\dim V_\lambda = l_\lambda = l_\lambda^{(h)} > l_\lambda^{(h-1)} > \dots > l_\lambda^{(1)} > 0.$$

Nach Annahme gibt es nämlich ein \tilde{w} mit $(A - \lambda I)^h \tilde{w} = 0$ und $(A - \lambda I)^{h-1} \tilde{w} \neq 0$. Für $i = 0, 1, \dots, h-1$ sei $\tilde{w}_i = (A - \lambda I)^i \tilde{w}$. Es gilt dann $(A - \lambda I)^{h-i} \tilde{w}_i = 0$ und $(A - \lambda I)^{h-i-1} \tilde{w}_i \neq 0$.

5. Man kann weiter zeigen: Es existieren verallgemeinerte Eigenvektoren v_1, v_2, \dots, v_m mit den Höhen h_1, h_2, \dots, h_m , sodass die Vektoren

$$v_1, (A - \lambda I)v_1, \dots, (A - \lambda I)^{h_1-1}v_1, v_2, (A - \lambda I)v_2, \dots, v_m, \dots, (A - \lambda I)^{h_m-1}v_m$$

eine Basis von V_λ bilden; $h_1 + h_2 + \dots + h_m = l_\lambda = \dim V_\lambda$. Wenn man den Endomorphismus von V_λ in dieser Basis darstellt, dann zerfällt die darstellende Matrix in Blöcke der Größe h_1, h_2, \dots, h_m . (Ohne Beweis!)

(Hinweis: Die Blöcke sind sog. Jordan-Blöcke wie im folgenden Beispiel 2.)

Beispiele

1. Für eine 2×2 -Matrix kann das Minimalpolynom nur linear oder quadratisch sein. Es ist genau dann linear, wenn A ein Vielfaches der Einheitsmatrix ist. Andernfalls haben wir $p_A(z) = \det(A - zI) = z^2 - \text{trace} A \cdot z + \det A$. (Wir haben das bereits in I.1 nachgerechnet.)
2. Betrachten wir die 5×5 -Matrix A mit dem einzigen Eigenwert λ

$$A = \begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}; \quad p_A(z) = (z - \lambda)^3.$$

Die Matrix zerfällt in zwei Blöcke. Die erste und die dritte ‘Einheitsspalte’ sind Eigenvektoren. Die zweite ‘Einheitsspalte’ hat die Höhe 2, die fünfte ist ein verallgemeinerter Eigenvektor der Höhe 3. Die Summe dieser Höhen ist $2 + 3 = 5 = \dim V_\lambda$. (Wir bemerken: das sog. charakteristische Polynom ist $\det(A - zI) = (\lambda - z)^5$.)

Satz

Sei A eine $n \times n$ -Matrix, deren Minimalpolynom in Linearfaktoren zerfällt

$$p_A(z) = (z - \lambda_1)^{h_1} \cdot (z - \lambda_2)^{h_2} \cdot \dots \cdot (z - \lambda_m)^{h_m} .$$

Dann lässt sich jede Spalte v in eindeutiger Weise als Summe von verallgemeinerten Eigenwerten darstellen.

$$v = \sum v_j \quad \text{mit} \quad v_j \text{ so, dass } (A - \lambda_j I)^{h_j} v_j = 0 .$$

Beweis Wir schreiben für ein festes j

$$p_A(z) = (z - \lambda_j)^{h_j} \cdot q_j(z) .$$

Der Satz von oben liefert für jedes j einen Vektorraum

$$V_j = \ker(A - \lambda_j I)^{h_j} = \text{im}(q_j(A)) .$$

Die $q_j(z)$ sind teilerfremd. Es existieren daher Polynome $r_j(z)$, sodass $\sum r_j(z) \cdot q_j(z) = 1$.

Für die Endomorphismen $\varphi_j = q_j(A)$ und $\rho_j = r_j(A)$ bedeutet das $\sum \rho_j \circ \varphi_j = \text{Id}$.

Für ein beliebiges v betrachten wir $w_j = \varphi_j(v)$, $v_j = \rho_j(w_j)$. Wir haben die Zerlegung

$$v = \sum \rho_j \circ \varphi_j(v) = \sum \rho_j(w_j) = \sum v_j .$$

Es gilt $v_j \in V_j$, weil $w_j \in \text{im} \varphi_j = V_j$, und weil ρ_j diesen Vektorraum $V_j = \ker(A - \lambda_j I)^{h_j}$ in sich abbildet. Somit haben wir eine Darstellung von v als Summe von verallgemeinerten Eigenvektoren.

Die Eindeutigkeit ergibt sich, wenn wir beweisen, dass der Nullvektor nur in trivialer Weise als Summe von verallgemeinerten Eigenvektoren zu den verschiedenen Eigenwerten λ_j dargestellt werden kann.

$$0 = \sum v_j \quad \text{mit} \quad v_j \in V_j \implies \forall j \quad v_j = 0 .$$

Das sieht man so : $v_j \in V_j$, $\varphi_j(v_j) = 0 \implies v_j = 0$; denn das Minimalpolynom zu $v_j \neq 0$ ist eine Potenz von $(z - \lambda_j)$. Andererseits gilt $q_j(A)v_i = 0$ für jeden verallgemeinerten Eigenvektor zu einem Eigenwert $\lambda_i \neq \lambda_j$. Somit gilt für jedes j

$$0 = \sum v_j \implies \forall j \quad 0 = q_j(0) = q_j(v_j) \implies \forall j \quad v_j = 0 .$$

Satz

Sei A eine komplexe $n \times n$ -Matrix und $p_{A,v}(z)$ das Minimalpolynom zur Spalte v

$$p_{A,v}(z) = \prod_j (z - \lambda_j)^{h_j} = (z - \lambda_k)^{h_k} \cdot r_k(z) .$$

Es gilt dann (für jedes k): Die Spalte $w_k = r_k(A)v$ ist ein verallgemeinerter Eigenvektor der Höhe h_k zum Eigenwert λ_k .

Der Beweis liegt auf der Hand. (Der Satz kann nützlich sein, wenn man Eigenwerte und dazugehörige Eigenvektoren sucht. Wie man das Minimalpolynom $p_{A,v}(z)$ findet, werden wir sehen, wenn wir uns in Abschnitt V. mit dem Basis-Austauschverfahren befassen.)

Hinweise

1. Wenn wir uns später mit Determinanten beschäftigen, werden wir sehen, dass das charakteristische Polynom $\tilde{p}(z) := \det(A - zI)$ die Matrix A annulliert, $\tilde{p}(A) = 0$. Dies ist der berühmte Satz von Cayley-Hamilton.. Das charakteristische Polynom ist also ein Vielfaches des Minimalpolynoms der Matrix A .
2. Bei komplexen Matrizen braucht man sich nur für die Nullstellen zu interessieren. Im charakteristischen Polynom sind die Vielfachheiten der Nullstellen die algebraischen Vielfachheiten $\lambda_j = \dim V_j$, während sie im Minimalpolynom (wie wir gesehen haben) die maximalen Höhen sind.
3. Bei manchen komplexen Matrizen (z. B. bei den Matrizen A mit $A^*A = AA^*$) ist von vorneherein klar, dass jeder verallgemeinerte Eigenvektor ein veritabler Eigenvektor ist. Hier hat dann also das Minimalpolynom nur einfache Nullstellen; und die Vielfachheiten der Nullstellen im charakteristischen Polynom sind die Dimensionen der Eigenräume.
4. Es ist i. Allg. eine anspruchsvolle Aufgabe für die Numerik, wenn man von einer gegebenen $n \times n$ -Matrix A mit $n > 4$ die (verallgemeinerten) Eigenvektoren berechnen möchte. Es ist i. Allg. nicht empfehlenswert, das charakteristische Polynom zu berechnen; die Berechnung der Koeffizienten ist mühevoll.. Es dürfte weniger aufwendig sein, zunächst einmal das Minimalpolynom eines (geschickt gewählten) Vektors v zu berechnen, und dessen Nullstellen zu bestimmen. Es ist als eine willkommene Reduktion der Aufgabe zu begrüßen, wenn $p_v(z)$ für das gewählte v ein Polynom geringeren Grades ist; bei Polynomen hohen Grades ist nämlich bekanntlich die Zerlegung in Linearfaktoren i. Allg. ein aufwendiges numerisches Problem.

I.7 Das Faltungsprodukt über \mathbb{Z} . Trigonometrische Polynome

Die rein algebraische Betrachtung der \mathbb{K} -Vektorräume werden wir in Abschnitt V weiterführen. Dort werden wir allerdings auch auf die fundamentale Idee der Abbildung bauen, so wie sie in Abschnitt II entwickelt wird.

Hier studieren wir noch einige konkrete \mathbb{C} -Vektorräume mit zusätzlicher Struktur. Gewisse Räume von komplexen Gewichtungen auf \mathbb{Z} tragen (vermöge des Faltungsprodukts) eine ähnliche algebraische Struktur wie die oben besprochene (assoziative und kommutative) Algebra $\mathbb{C}[z]$ der Polynome. Das Interesse geht allerdings in eine andere Richtung; bei den trigonometrischen Polynomen und bei den formalen Potenzreihen gibt es kein Teilen mit Rest und kein Umzentrieren; andere Strukturen treten in den Vordergrund.

Es ergeben sich Ausblicke auf Gegenstände, die später in der Analysis (oder in der Stochastik) in einem vertieften Sinn zu studieren sind. Der eilige Leser mag das überspringen und sich sofort den Abbildungen im Sinne von Abschnitt II zuwenden.

Sprechweisen

Eine komplexwertige Funktion auf \mathbb{Z}_+ kann man als eine „einseitige Folge“ bezeichnen

$$\mathbf{a} = (a_0, a_1, a_2, \dots) .$$

Eine komplexwertige Funktion auf \mathbb{Z} nennt man manchmal eine „zweiseitige Folge“

$$\mathbf{a} = (\dots, a_{-1}, a_0, a_1, a_2, \dots) .$$

Wir sprechen von einer finiten Folge, wenn nur endlich viele der a_n ungleich 0 sind. Im Folgenden nennen wir diese finiten Folgen **finite Gewichtungen auf \mathbb{Z}** .

Wir sprechen von einer linksseitig (bzw. rechtsseitig) finiten Gewichtung, wenn es ein $N \in \mathbb{Z}$ gibt, sodass $a_n = 0$ für alle $n < N$ (bzw. für alle $n > N$).

Faltung (Definition)

Seien \mathbf{a} und \mathbf{b} finite Gewichtungen auf \mathbb{Z} . Wir definieren dann ihr **Faltungsprodukt** $\mathbf{c} = \mathbf{a} * \mathbf{b}$ als die Gewichtung \mathbf{c} mit den Gewichten

$$c_n = \sum_k a_k \cdot b_{n-k} = \sum_\ell a_{n-\ell} \cdot b_\ell \quad \text{für alle } n \in \mathbb{Z} .$$

Die Faltung heisst im Englischen ‘convolution’. In der älteren deutschen Literatur findet man auch den Namen ‘Cauchy-Produkt’.

Bemerke :

1. Man kann auch Gewichtungen, die nur linksseitig bzw. rechtsseitig finit sind, falten; und man erhält als Faltungsprodukt eine linksseitig (bzw. rechtsseitig) finite Gewichtung.

2. Wenn man Gewichtungen falten möchte, die nicht einseitig finit sind, dann treten u.U. Probleme mit der Konvergenz unendlicher Reihen auf. Diese Fragen der Konvergenz sind aber harmlos, wenn die Gewichtungen ein endliches Gesamtgewicht haben. Wir betrachten vorwiegend solche Gewichtungen, also $(a_n)_{n \in \mathbb{Z}}$ mit $\sum |a_n| < \infty$. Es gilt dann für $c = a * b$. $\sum |c_n| \leq \sum |a_k| \cdot \sum |b_\ell|$.
3. Eine Gewichtung $(p_n)_{n \in \mathbb{Z}}$ heißt eine Wahrscheinlichkeitsgewichtung auf \mathbb{Z} , wenn gilt

$$p_n \geq 0 \text{ für alle } n, \quad \sum p_n = 1.$$

Das Faltungsprodukt zweier Wahrscheinlichkeitsgewichtungen ist offenbar eine Wahrscheinlichkeitsgewichtung.

4. Sei $\sum a_n z^n$ ein Polynom mit komplexen Koeffizienten. Die Koeffizientenfolge kann man als eine finite Gewichtung verstehen. Die Produktbildung im Ring $\mathbb{C}[z]$ entspricht der Faltung der Koeffizientenfolgen.

Satz

Für die Faltung von finiten Gewichtungen auf \mathbb{Z} gelten die Assoziativgesetze und die Distributivgesetze. Die Faltung ist kommutativ.

$$\begin{aligned} a * (b * c) &= (a * b) * c \\ a * (b + c) &= a * b + a * c \\ a * b &= b * a. \end{aligned}$$

Diese Rechenregeln gelten auch für rechtsseitig finite Gewichtungen und für Gewichtungen mit endlichem Gesamtgewicht.

Beweis des Assoziativgesetzes

Sei $d = a * (b * c)$, d.h.

$$d_n = \sum_k a_k \cdot \left(\sum_{\ell+m=n-k} b_\ell c_m \right) = \sum_{k+\ell+m=n} a_k \cdot b_\ell \cdot c_m.$$

Dieselbe Dreifachsumme ergibt sich für das Faltungsprodukt $(a * b) * c$.

Der Beweis des Distributivgesetzes ist ebenso einfach. Die Kommutativität ist trivial.

Bemerkung („Schriftliches Multiplizieren“ ist „Faltung + Übertrag“)

$\begin{array}{r} 3441 \\ \text{mal } 211 \\ \hline 3441 \\ 3441 \\ 6882 \\ \hline 726051 \end{array}$	$\begin{aligned} 3441 \cdot 211 &= (3 \cdot 10^3 + 4 \cdot 10^2 + 4 \cdot 10 + 1)(2 \cdot 10^2 + 1 \cdot 10 + 1) = \\ &= c_5 \cdot 10^5 + c_4 \cdot 10^4 + c_3 \cdot 10^3 + c_2 \cdot 10^2 + c_1 \cdot 10 + c_0 \\ \text{mit } c_0 &= a_0 b_0, \quad c_1 = a_1 b_0 + a_0 b_1, \quad c_2 = a_2 b_0 + a_1 b_1 + a_0 b_2, \dots \end{aligned}$
--	---

Die c_n sind zunächst noch nicht die Ziffern im (dezimal dargestellten) Produkt 726051;

durch den „Übertrag“ erreicht man

$$= d_5 \cdot 10^5 + d_4 \cdot 10^4 + d_3 \cdot 10^3 + d_2 \cdot 10^2 + d_1 \cdot 10 + d_0 \quad \text{mit} \quad d_j = \{0, 1, 2, \dots, 9\}.$$

Das Falten von finiten Gewichtungen (im Sinne der Definition) ist insofern einfacher als das „schriftliche Multiplizieren“ (im Sinne der Schulmathematik), als hier keine Überträge zu beachten sind.

Falten in der Wahrscheinlichkeitstheorie

Seien X und Y unabhängige ganzzahlige Zufallsgrößen mit

$$\begin{aligned} \text{Ws}(\{X = k\}) &= a_k \quad \text{für } k \in \mathbb{Z} \\ \text{Ws}(\{Y = \ell\}) &= b_\ell \quad \text{für } \ell \in \mathbb{Z}. \end{aligned}$$

Die Wahrscheinlichkeitsgewichtung für die Summe $Z = X + Y$ ergibt sich durch Faltung der Gewichtungen a und b .

$$\text{Ws}(X + Y = n) = c_n = \sum_{k+\ell=n} a_k \cdot b_\ell \quad \text{für } n \in \mathbb{Z}.$$

Es gilt nämlich

$$\text{Ws}(Z = n) = \sum_{k+\ell=n} \text{Ws}(X = k, Y = \ell) = \sum_{k+\ell=n} \text{Ws}(X = k) \cdot \text{Ws}(Y = \ell).$$

Beispiele

1. Zwei faire Würfel werden unabhängig geworfen. Die Augensumme $Z = X + Y$ hat eine Verteilung, die auf die Zahlenmenge $\{2, 3, \dots, 11, 12\}$ konzentriert ist.

$$\begin{aligned} \text{Ws}(\{X = k\}) &= \frac{1}{6} \quad \text{für } k = 1, 2, \dots, 6; & \text{Ws}(\{Y = \ell\}) &= \frac{1}{6} \quad \text{für } \ell = 1, 2, \dots, 6. \\ \text{Ws}(\{Z = 7\}) &= \frac{1}{6}, & \text{Ws}(\{Z = 6\}) &= \text{Ws}(\{Z = 8\}) = \frac{5}{36}, & \text{Ws}(\{Z = 5\}) &= \frac{4}{36}, \dots \end{aligned}$$

2. Seien X und Y unabhängig und gleichmäßig verteilt auf der Punktmenge $\{-N, -N+1, \dots, -1, 0, +1, \dots, N\}$. Die Gewichte von $Z = X + Y$ liefern eine ‘Dreiecksverteilung’:

$$\text{Ws}(\{X + Y = n\}) = \frac{1}{2N+1} \cdot \left(1 - \frac{|n|}{2N+1}\right)^+.$$

3. Ein Versuch mit der Erfolgswahrscheinlichkeit p ist n -mal unabhängig durchgeführt worden. Die Anzahl S der Erfolge ist die Summe der Zufallsgrößen X_j mit dem Wert 1, wenn der j -te Versuch ein Erfolg war und dem Wert 0, wenn er ein Misserfolg war. $S = X_1 + X_2 + \dots + X_n$. Die Verteilung von S erhält man, wenn man die Zweipunktgewichtung $(\dots, 0, (1-p), p, 0, \dots)$ n -mal mit sich selber faltet. Wir erhalten die Binomialgewichtung

$$\text{Ws}(\{S = k\}) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{für } k \in \mathbb{Z}.$$

4. Ein Versuch mit der Erfolgswahrscheinlichkeit p wird unabhängig wiederholt. T_1 sei die Wartezeit bis zum ersten Erfolg, $T_1 + T_2$ die Wartezeit bis zum zweiten Erfolg, ..., $T_1 + T_2 + \dots + T_n$ sei die Wartezeit bis zum n -ten Erfolg. Wir bestimmen

$$\text{Ws}(\{T_1 + \dots + T_n = n + k\}) = a_k^{(n)} \quad \text{für } k = 0, 1, 2, \dots$$

Offenbar gilt

$$\text{Ws}(\{T_1 = k\}) = (1 - p)^{k-1} \cdot p \quad \text{für } k = 1, 2, \dots$$

Auf $(k-1)$ Mißerfolge folgt ein erster Erfolg. Wir setzen $(1-p) = c$. Die Gewichtung für T_1 ist also

$$(1 - c) \cdot (0, 1, c, c^2, \dots)$$

Die Gewichtung für $T_1 + T_2$ ist das Faltungsprodukt

$$(1 - c)^2 (0, 1, c, c^2, \dots) * (0, 1, c, c^2, \dots)$$

Die Gewichtung für $T_1 + T_2 + \dots + T_n$ ergibt sich durch n -faches Falten. Die Rechnung werden wir unten mit Newton's Binomialreihe durchführen. Es ergibt sich

$$a_k^{(n)} = (1 - c)^n \cdot \binom{-n}{k} (-c)^k = \binom{n+k-1}{n-1} p^{n-1} \cdot (1-p)^k \cdot p$$

Hier verwenden wir die Notation der verallgemeinerten **Binomialkoeffizienten**.

Für $\alpha \in \mathbb{R}$, $k = 0, 1, 2, \dots$ definiert man

$$\binom{\alpha}{k} = \frac{1}{k!} \alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)$$

Auch die folgenden Bezeichnungen aus der Kombinatorik erweisen sich als nützlich.

$$\begin{aligned} (\alpha)_{(k)} &= \alpha(\alpha-1) \dots (\alpha-k+1) && \text{ („}\alpha \text{ untere Faktorielle } k\text{“)} \\ (\alpha)^{(m)} &= \alpha \cdot (\alpha+1) \cdot \dots \cdot (\alpha+m-1) && \text{ („}\alpha \text{ obere Faktorielle } m\text{“)} \\ (k+1)^{(n-1)} &= (k+1) \cdot (k+2) \cdot \dots \cdot (k+n-1) && = (n+k-1)_{(n-1)} \end{aligned}$$

Hinweise

Man kann das Resultat auch stochastisch begründen:

$a_k^{(n)}$ ist die Wahrscheinlichkeit dass zur Zeit $n+k$ der n -te Erfolg eintritt. Der $(n+k)$ -te Versuch muss ein Erfolg sein, nachdem in den vorausgehenden $n+k-1$ Versuchen genau $(n-1)$ Erfolge eingetreten sind. In der Tat ist

$$\binom{n+k-1}{n-1} p^{n-1} \cdot (1-p)^k$$

die Wahrscheinlichkeit, dass in $(n+k-1)$ Versuchen genau $(n-1)$ Erfolge eintreten.

Für den Stochastiker überzeugend ist übrigens auch die Formel

$$a_k^{(n)} = p^n \cdot \frac{1}{(n-1)!} (k+1)^{(n-1)} \cdot \exp(k \cdot \ln(1-p)) \quad \text{für } k = 0, 1, 2, \dots$$

Sie erinnert ihn nämlich an die berühmten Gamma-Dichten

$$\frac{1}{\Gamma(n)} \cdot x^{n-1} \cdot e^{-x} dx \quad (\text{für } x > 0).$$

Man kann sagen: So wie die geometrische Gewichtung das diskrete Gegenstück zur Exponentialdichte ist, so sind die „negativen Binomialverteilungen“ die diskreten Analoga zu den Gamma-Dichten.

Definition (trigonometrische Polynome und charakteristische Funktionen)

- Ein trigonometrisches Polynom ist eine 2π -periodische komplexwertige Funktion von der Form

$$f(t) = \sum a_n \cdot e^{int}, \quad t \in \mathbb{R}/2\pi \quad (a_n)_n \text{ finit.}$$

Man nennt $f(t)$ auch die charakteristische Funktion zur finiten Koeffizientenfolge $(a)_n$. Wenn $a_n = 0$ für alle n mit $|n| > N$, dann spricht man von einem trigonometrischen Polynom vom Grad $\leq N$.

- Allgemeiner definiert man die charakteristische Funktion zur Gewichtung $a = (\dots, a_{-1}, a_0, a_1, \dots)$ mit endlichem Gesamtgewicht $\sum |a_n| < \infty$.

$$f(t) = \sum a_n e^{int}, \quad t \in \mathbb{R}/2\pi.$$

Satz

Das punktweise Produkt trigonometrischer Polynome ist ein trigonometrisches Polynom

$$h(t) = f(t) \cdot g(t) \quad \text{für } t \in \mathbb{R}/2\pi.$$

*Sind a und b Gewichtungen und $c = a * b$, dann ist die charakteristische Funktion des Faltungsprodukts das punktweise Produkt der charakteristischen Funktionen.*

Der Beweis ergibt sich durch distributives Ausmultiplizieren. Wegen $e^{ikt} \cdot e^{ilt} = e^{i(k+l)t}$ haben wir

$$\begin{aligned} f(t) \cdot g(t) &= \left(\sum_k a_k e^{ikt} \right) \cdot \left(\sum_\ell b_\ell \cdot e^{i\ell t} \right) = \sum_{k,\ell} a_k b_\ell \cdot e^{i(k+\ell)t} \\ &= \sum_n e^{int} \left(\sum_{k+\ell=n} a_k \cdot b_\ell \right) = \sum_n c_n e^{int} = h(t). \end{aligned}$$

Beispiele Die folgenden Beispiele beziehen sich in der einen oder anderen Weise auf die (von der Schule her bekannte) Summationsformel für die **geometrische Reihe**. Wir leiten einige Formeln her, die sich immer wieder als nützlich erweisen, sowohl in der Analysis als auch ganz besonders in der Stochastik.

1. Die Dirichlet-Dichte der Ordnung n ist das reellwertige trigonometrische Polynom

$$D_n(t) = \sum_{-n}^n e^{ikt} = (e^{it/2} - e^{-it/2})^{-1} \cdot \left(e^{i(n+\frac{1}{2})t} - (e^{-i(n+\frac{1}{2})t}) \right) = \frac{1}{\sin t/2} \cdot \sin\left((n+\frac{1}{2})t\right).$$

Die Funktion $f(t) = \frac{1}{2N+1} D_N(t)$ ist offenbar die charakteristische Funktion der Gleichverteilung auf $[-N, N]$.

2. Die Fejér-Dichte der Ordnung n ist das nichtnegative trigonometrische Polynom

$$F_n(t) = \frac{1}{n} \cdot (D_0(t) + D_1(t) + \cdots + D_{n-1}(t)).$$

Es zeigt sich, dass $h(t) = \frac{1}{2N+1} F_{2N+1}(t)$ die charakteristische Funktion der oben betrachteten ‘Dreiecksverteilung’ ist. $F_{2N+1}(t) = (2N+1)^{-1} \cdot (D_N(t))^2$.

Die Summationsformel für die endliche geometrische Reihe liefert in der Tat

$$\begin{aligned} n \cdot F_n(t) &= \frac{1}{\sin t/2} \cdot \frac{1}{2i} \sum_0^{n-1} (e^{i(k+\frac{1}{2})t} - e^{-i(k+\frac{1}{2})t}) \\ &= \frac{1}{\sin t/2} \cdot \frac{1}{2i} (e^{it/2}(e^{it} - 1)^{-1}(e^{int} - 1) - e^{-it/2}(e^{-it} - 1)^{-1}(e^{-int} - 1)) \\ &= \left(\frac{1}{(2i) \cdot \sin t/2} \right)^2 \cdot (e^{int} - 1 + e^{-int} - 1) \\ &= \left(\frac{\sin nt/2}{\sin t/2} \right)^2 \end{aligned}$$

Bemerke:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(t) dt &= 1; & \frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(t) dt &= 1. \\ D_n(0) &= 2n+1, & F_n(0) &= n. \end{aligned}$$

Bemerkung Die charakteristische Funktion einer Wahrscheinlichkeitsgewichtung hat im Nullpunkt den Wert $= 1$; sie ist überall dem Betrag nach ≤ 1 .

Für spätere Gelegenheiten bemerken wir: Für jedes $t_0 \in \mathbb{R}/2\pi$ liefert $\frac{1}{2\pi} F_n(t - t_0) dt$ eine Wahrscheinlichkeitsdichte auf $\mathbb{R}/2\pi$; diese Schar von Wahrscheinlichkeitsdichten nennt man den Fejér-Kern der Ordnung n auf $\mathbb{R}/2\pi$.

Beispiel (‘Doppeltgeometrische’ Gewichtung) Wir betrachten nochmals geometrische Gewichtungen. Die Differenz zweier unabhängiger geometrisch verteilter Zufallsgrößen mit demselben Erwartungswert hat eine Verteilung, die man doppelte geometrische (oder ‘beidseitig geometrische’) Gewichtung nennt. Man kann das sowohl von der Faltungsformel her leicht bestätigen als auch mit Hilfe der charakteristischen Funktionen.

Seien X und Y unabhängige Zufallsgrößen mit

$$W_S(X = k) = W_S(Y = k) = \frac{1-c}{c} \cdot c^k \quad \text{für } k = 1, 2, 3, \dots;$$

Dann gilt

$$W_S(X - Y = n) = \text{const} \cdot c^{|n|} \quad \text{für } n = 0, \pm 1, \pm 2, \pm 3, \dots;$$

Beweis

Die Gewichtung ist offensichtlich symmetrisch. Für $n = 0, 1, 2, \dots$ gilt

$$W_S(X - Y = n) = \sum_{k=1}^{\infty} W_S(X = n + k) \cdot W_S(Y = k) = \left(\frac{1-c}{c}\right)^2 \cdot c^n \sum_{k=1}^{\infty} c^{2k} = \text{const} \cdot c^{|n|}.$$

Betrachten wir andererseits die charakteristischen Funktionen

$$\begin{aligned} f(t) &= \frac{c}{1-c} \cdot \mathcal{E} e^{itX} = \sum_{k=1}^{\infty} c^k e^{itk} = \frac{ce^{it}}{1-ce^{it}} \\ |f(t)|^2 &= \left(\frac{c}{1-c}\right)^2 \cdot \mathcal{E} e^{it(X-Y)} = \frac{c^2}{(1-ce^{it})(1-ce^{-it})} \end{aligned}$$

und vergleichen wir damit

$$\begin{aligned} 1 + f(t) + \overline{f(t)} &= \frac{(1-ce^{it})(1-ce^{-it}) + ce^{it}(1-ce^{-it}) + ce^{-it}(1-ce^{it})}{(1-ce^{it})(1-ce^{-it})} \\ &= \frac{1-c^2}{(1-ce^{it})(1-ce^{-it})} = \frac{1-c^2}{c^2} \cdot |f(t)|^2. \end{aligned}$$

Die linke Seite zeigt, dass die Koeffizienten von $|f(t)|^2$ nach beiden Seiten geometrisch abfallen. Die 2π -periodische Funktion $|f(t)|^2$ können wir auch folgendermaßen schreiben:

$$|f(t)|^2 = \frac{c^2}{1+c^2+2c \cdot \cos t} = \frac{c^2}{(1-c)^2 + 4c \cdot \sin^2(t/2)}.$$

Bei dieser Gelegenheit sei bemerkt: Wenn X und Y unabhängige exponentiell verteilte Zufallsgrößen mit dem Erwartungswert $1/\lambda$ sind, dann hat ihre Differenz die doppelte exponentielle Dichte $\frac{\lambda}{2} \cdot e^{-\lambda|x|} dx$. Die charakteristische Funktion ist

$$\mathcal{E} e^{it(X-Y)} = \frac{\lambda}{2} \cdot \int_{-\infty}^{\infty} e^{itx} e^{-\lambda|x|} dx = \frac{\lambda}{2} \cdot \left(\frac{1}{\lambda - it} + \frac{1}{\lambda + it} \right) = \frac{\lambda^2}{\lambda^2 + t^2}.$$

Satz

Die Koeffizienten eines trigonometrischen Polynoms $f(t)$ sind eindeutig bestimmt. Es gilt

$$f(t) = \sum a_n e^{int} \iff \forall n \in \mathbb{Z} \quad a_n = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \cdot f(t) dt .$$

Der Beweis ergibt sich aus der sog.

Orthogonalitätsbeziehung :

Für $n, m \in \mathbb{Z}$ gilt

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \cdot e^{imt} dt = \begin{cases} 0 & \text{falls } n \neq m \\ 1 & \text{falls } n = m \end{cases}$$

Beweis der Orthogonalitätsrelation

Der Fall $n = m$ ist trivial. Für $m - n = k \neq 0$ gilt

$$\int_{-\pi}^{+\pi} e^{ikt} dt = \int_{-\pi}^{+\pi} \cos(kt) dt + i \cdot \int_{-\pi}^{+\pi} \sin(kt) dt = 0 .$$

Es wird hier nämlich über volle Perioden integriert.

Beweis des Satzes :

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \cdot \sum a_m e^{imt} dt = a_n \quad \text{für alle } n \in \mathbb{Z} .$$

Inneres Produkt

Eine interessante Folgerung aus der Orthogonalitätsrelation ist der

Satz

Für jedes trigonometrische Polynom $f(t) = \sum a_n e^{int}$ gilt

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} |f(t)|^2 dt = \sum |a_n|^2 .$$

Wenn $f(t) = \sum a_n e^{int}$ und $g(t) = \sum b_n e^{int}$ trigonometrische Polynome sind, dann gilt

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} \bar{g}(t) f(t) dt = \sum \bar{b}_n a_n .$$

Beweis :

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} \bar{g}(t) \cdot f(t) dt = \frac{1}{2\pi} \int \left(\sum \bar{b}_n e^{-int} \right) \cdot \left(\sum a_m e^{imt} \right) dt = \sum \bar{b}_n a_n .$$

Ausblick

Diese das ‘innere Produkt’ erhaltende Beziehung zwischen einem Vektorraum 2π -periodischer Funktionen einerseits und dem Vektorraum der finiten Folgen (a_n) andererseits lässt sich fortsetzen. Zunächst einmal funktioniert das ohne jedes Problem für die Folgen mit endlichem Gesamtgewicht und die entsprechenden ‘absolut konvergenten’ trigonometrischen Reihen. (Unendliche Reihen der Form $\sum_{-\infty}^{+\infty} a_n e^{int}$ heißen trigonometrische Reihen.)

In der Analysis geht man dann auch noch weiter; mit einer Vervollständigungskonstruktion (analog zur Vervollständigung $\mathbb{Q} \subset \mathbb{R}$) gelangt man zu einem sog. Hilbertraum. Es ist aber zu bemerken, dass im vervollständigten Bereich keine Faltung der Folgen definiert werden kann; die Vektorraumstruktur verbindet sich hier mit einer multiplikativen Struktur, sondern mit einer interessanten topologischen Struktur. Diese topologische Struktur werden wir in der Analysis diskutieren.

Hinweis zur Geschichte der trigonometrischen Reihen

Von B. Riemann (1826 – 1866) stammt die Idee, einer ‘beliebigen integrierbaren’ 2π -periodischen Funktion $f(t)$ eine Folge von Koeffizienten $a = (\dots, a_{-1}, a_0, a_1, a_2, \dots)$ zuzuordnen, nämlich

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \cdot f(t) dt \quad \text{für alle } n \in \mathbb{Z} .$$

Zur Funktion $f(t)$ betrachtete man die sog. formale Fourier-Reihe $f(\cdot) \mapsto \sim \sum_n a_n e^{int}$. Es wurde gefragt, unter welchen Umständen die formale Fourier-Reihe zu einer Funktion $f(\cdot)$ ein sinnvolles Objekt im Sinne der Analysis ist. Weiter wurde gefragt, ob oder gegebenenfalls wie man mit diesen formalen Objekten rechnen kann.

Schon L. Euler (1707 – 1783) hatte trigonometrische Reihen diskutiert, die zu speziellen unstetigen Funktionen gehören und daher ein sonderbares Konvergenzverhalten der Partialsummen zeigen: $\sum_{-N}^N a_n e^{int} \rightarrow ?$

Ein Beispiel ist die ‘Euler’sche Sägezahnfunktion’ $f(t)$:

$$f(\cdot) \text{ ist } 2\pi\text{-periodisch mit } f(t) = \frac{1}{2}(\pi - t) \quad \text{für } t \in (0, 2\pi),$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \frac{1}{2}(\pi - t) dt = \frac{1}{2in} \quad \text{für } n \neq 0 ,$$

$$f^{(N)}(t) = c_{-N} e^{-iNt} + \dots + c_N e^{iNt} = \sin t + \frac{1}{2} \sin 2t + \dots + \frac{1}{N} \sin Nt .$$

(Es sollte lehrreich sein, diese Funktionenfolge auf dem Bildschirm zu realisieren!) Riemann entwickelte 1854 einen Vorschlag, was für seine Theorie Integrabilität bedeuten sollte. Seine Theorie der Integrabilität hatte aber Mängel. Erst 1901, durch die Integrationstheorie von H. Lebesgue (1875 – 1941) wurde der Weg frei für eine wirklich befriedigende Einsicht. Fischer und Riesz haben 1906 gezeigt, dass man die Elemente des vervollständigten Raums sowohl durch die quadratintegrablen 2π -periodische Funktionen repräsentieren kann als auch durch die quadratsummablen Folgen $(a_n)_n$: $\sum |a_n|^2 < \infty$.

Aufgaben zu I.7**Aufgabe I.7.1 :**

Ein trigonometrisches Polynom lässt sich auch als eine Linearkombination der Funktionen $\cos kt$ und $\sin kt$ darstellen

$$\sum_{-N}^N c_n e^{int} = f(t) = \frac{1}{2} a_0 + \sum_{k=1}^N (a_k \cos kt + b_k \sin kt).$$

Gewinnen Sie die Koeffizienten a_k , b_k aus den Koeffizienten c_n und umgekehrt. Zeigen Sie, dass die Funktion $f(t)$ auf $\mathbb{R}/2\pi$ genau dann reellwertig ist, wenn $c_{-n} = \bar{c}_n$ für alle n . Dies ist genau dann der Fall, wenn alle a_k und b_k reell sind.

Aufgabe I.7.2 : Berechnen Sie für $k, l \in \mathbb{Z}$ die Zahlen

$$\begin{aligned} c_{kl} &= \frac{1}{\pi} \int_{-\pi}^{+\pi} \cos kt \cdot \cos lt \, dt, \\ g_{kl} &= \frac{1}{\pi} \int_{-\pi}^{+\pi} \sin kt \cdot \cos lt \, dt, \\ s_{kl} &= \frac{1}{\pi} \int_{-\pi}^{+\pi} \sin kt \cdot \sin lt \, dt. \end{aligned}$$

Hinweis:

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} \cdot e^{imt} \, dt = \begin{cases} 0 & \text{für } m \neq n \\ 1 & \text{für } m = n. \end{cases}$$

Aufgabe I.7.2 :

Es sei $f(t) = \frac{1}{2} a_0 + \sum_{k=1}^N (a_k \cos kt + b_k \sin kt)$.

Zeigen Sie

$$a_k = \frac{1}{\pi} \int_{-\pi}^{+\pi} \cos kt \cdot f(t) \, dt; \quad b_k = \frac{1}{\pi} \int_{-\pi}^{+\pi} \sin kt \cdot f(t) \, dt.$$

Aufgabe I.7.3 : Sei r eine positive Zahl < 1 und

$$f(t) = \sum_{n=1}^{\infty} r^n \cdot e^{int} \quad \text{für } t \in \mathbb{R}/2\pi.$$

Es gilt dann mit gewissen Koeffizienten c_n bzw. a_k

$$|f(t)|^2 = \overline{f(t)} \cdot f(t) = \sum_{-\infty}^{+\infty} c_n \cdot e^{int} = \frac{1}{2} a_0 + \sum_1^{\infty} a_k \cdot \cos(kt) .$$

a) Berechnen Sie diese Koeffizienten

$$c_0, \quad c_1 = \bar{c}_{-1}, \quad c_2 = \bar{c}_{-2}, \quad \dots$$

b) Finden Sie eine einfache Formel für $|f(t)|^2$. (Bild!)

Hinweis : Verstehen Sie $f(t)$ als eine geometrische Reihe.

c) (Für diejenigen, die schon einmal mit MAPLE gespielt haben.) Plotten Sie die Funktion $|f(t)|^2$ für verschiedene r , $r \nearrow 1$.

I.8 Newton's Binomialreihe. Formale Potenzreihen

Der Begriff der erzeugenden Funktion ist eine Variante des Begriffs der charakteristischen Funktion. Die erzeugenden Funktionen leisten manchmal gute Dienste, wenn es um 'Gewichtungen' auf \mathbb{Z}_+ geht.

Sei $\mathbf{a} = (a_0, a_1, a_2, \dots)$ eine Gewichtung auf \mathbb{Z}_+ mit $\sum |a_n| < \infty$. Dann definiert man dazu die erzeugende Funktion

$$h(z) = \sum_0^{\infty} a_n \cdot z^n \quad \text{für } |z| \leq 1.$$

Hinweis Dieselbe Idee liegt auch dem Begriff der Z-Transformierten eines Signals $\{f(n) : n \in \{0, 1, 2, \dots\}\}$ zugrunde, welcher im Gebiet der Nachrichtenübertragung bekannt ist. Man definiert dort $F(z) = \sum_0^{\infty} f(n)z^{-n}$, wobei man davon ausgeht, dass das Signal so dimensioniert ist, dass die Summe für $|z| > 1$ konvergiert.

Bemerke :

1. Die charakteristische Funktion ist als eine komplexwertige Funktion auf dem abgeschlossenen Einheitskreis zu verstehen. Manchmal betrachtet man sie auch nur als Funktion auf dem abgeschlossenen Einheitsintervall $[-1, +1]$.
2. Für finite Gewichtungen auf \mathbb{Z}_+ ist die erzeugende Funktion ein Polynom.
3. Die charakteristische Funktion $f(\cdot)$ der Gewichtung \mathbf{a} ergibt sich als die Einschränkung der erzeugenden Funktion $h(\cdot)$ auf die Peripherie des Einheitskreises $z = e^{it}$.

$$f(t) = h(e^{it}).$$

4. Für eine Wahrscheinlichkeitsgewichtung $a_k = \text{Ws}(X = k)$ notiert man

$$f(t) = \mathcal{E} \exp(itX); \quad h(z) = \mathcal{E} z^X.$$

Es gilt $h'(z) = \sum \text{Ws}(X = k) k z^{k-1}$, insbesondere $h'(1) = \mathcal{E}X$, sowie $h''(1) = \sum \text{Ws}(X = k) k(k-1) = \mathcal{E}X(X-1) = \text{var}X + \mathcal{E}X(\mathcal{E}X - 1)$.

Beispiele Ein Experiment mit der Erfolgswahrscheinlichkeit p wird unabhängig wiederholt. $N^{(n)}$ sei die Anzahl der Erfolge nach n Versuchen. $T^{(n)}$ sei die Wartezeit bis zum n -ten Erfolg. Diese beiden Zufallsgrößen sind Summen von n unabhängigen identisch verteilten Zufallsgrößen.

$$\mathcal{E} z^{N^{(n)}} = \sum_k \binom{n}{k} p^k (1-p)^{n-k} z^k = ((1-p) + pz)^n = (1-p)^n \left(1 + \frac{p}{1-p} z\right)^n;$$

$$\mathcal{E} z^{T^{(n)}-n} = \left(\sum_{l=0}^{\infty} \text{Ws}(\{T^{(1)} = 1+l\}) z^l\right)^n = \left(\frac{p}{1-(1-p)z}\right)^n = p^n (1-(1-p)z)^{-n}.$$

Die Koeffizienten $a_k^{(n)}$ dieser Reihe haben wir oben mit einem stochastischen Argument bestimmt. Ein zweiter Zugang ergibt sich durch die sog. binomische Reihe. Diese Reihe ist eine Entdeckung von Newton, die im 18. Jahrhundert als ein Höhepunkt der „kombinatorischen Analysis“ gefeiert wurde, im 19. Jahrhundert dann aber in eine allgemeine Theorie der im Einheitskreis holomorphen Funktionen eingereiht wurde.

Satz (Newton's Binomialreihe)

Für alle $\alpha \in \mathbb{C}$ und alle z mit $|z| < 1$ gilt

$$(1+z)^\alpha = 1 + \alpha \cdot z + \binom{\alpha}{2} z^2 + \binom{\alpha}{3} z^3 + \dots = \sum_0^\infty \binom{\alpha}{k} z^k,$$

Wir betrachten zuerst Spezialfälle, die möglicherweise von der Schule her bekannt sind, nämlich die sog. Binomialformeln sowie die Abkömmlinge der geometrischen Reihe, die bei den ‘negativen Binomialverteilungen’ auftreten.

1. Für $\alpha = 1, 2, \dots$ ist die binomische Reihe eine endliche Summe. Die Aussage ist der binomische Lehrsatz

$$\begin{aligned} (1+z)^2 &= 1 + 2z + z^2 \\ (1+z)^3 &= 1 + 3z + 3z^2 + z^3 \\ (1+z)^4 &= 1 + 4z + 6z^2 + 4z^3 + z^4 \\ &\dots \quad \dots \end{aligned}$$

2. Für $\alpha = -1$ ist die binomische Reihe nichts anders als die geometrische Reihe, denn $\binom{-1}{k} = \frac{1}{k!}(-1)(-2)\dots(-k) = (-1)^k$. Durch Differentiation gewinnt man die Reihe für $(1+z)^{-n}$

$$\begin{aligned} (1-z)^{-1} &= 1 + z + z^2 + z^3 + \dots, \\ (1-z)^{-2} &= 1 + 2z + 3z^2 + 4z^3 + \dots \\ 2 \cdot (1-z)^{-3} &= 2 \cdot 1 + (3 \cdot 2)z + (4 \cdot 3)z^2 + (5 \cdot 4)z^3 + \dots \\ &\dots \\ (1-z)^{-n} &= \frac{1}{(n-1)!} \sum_0^\infty (k+1)^{(n-1)} (-z)^k, . \end{aligned}$$

Diese Reihe ist die Binomialreihe wegen

$$\frac{(k+1)^{(n-1)}}{(n-1)!} \cdot \frac{k!}{k!} = \binom{n+k-1}{k} = \frac{(n)^{(k)}}{k!} = (-1)^k \frac{(-n)_{(k)}}{k!} = (-1)^n \binom{-n}{k}.$$

Anmerkung: Die Ableitung einer Funktion, die durch eine Reihe gegeben ist, kann man bekanntlich nicht immer dadurch finden, dass man die Reihe gliedweise differenziert. Probleme dieser Art werden in der Analysis abgehandelt. Sie sind hier nicht unser Thema.

Wir akzeptieren die folgenden Aussagen aus der elementaren Analysis:

1. Der Hauptwert der Logarithmusfunktion lässt sich in einer Umgebung der Eins durch eine Potenzreihe darstellen

$$\ln(1+z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4 + \cdots \quad \text{für } |z| < 1.$$

2. Auch der Hauptwert der Funktion $(1+z)^\alpha = \exp(\alpha \cdot \ln(1+z))$ lässt sich dort durch eine Potenzreihe darstellen

$$(1+z)^\alpha = 1 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots \quad \text{für } |z| < 1.$$

Wir wollen die Koeffizienten diskutieren. Dies ist eine algebraisch-kombinatorische Aufgabe, die man loslösen kann von den Fragen der Konvergenz. Newton hat für solche Aufgaben (insbesondere für das 'Lösen' von Differentialgleichungen) die **Methode des Koeffizientenvergleichs** entwickelt. Diese Methode wollen wir im Folgenden im Rahmen der algebraischen Theorie der formalen Potenzreihen entwickeln.

Formale Potenzreihen

Eine formale Potenzreihe ist ein formaler Ausdruck der Form

$$A(z) = \sum_0^\infty a_n \cdot z^n, \quad \text{oder} \quad A(z) = a_0 + a_1 z + a_2 z^2 + \cdots.$$

Die Koeffizienten sollen bei uns hier komplexe Zahlen sein; es kommen hier aber auch allgemeinere Systeme von Koeffizienten in Frage. Für die beiden führenden Koeffizienten notieren wir auch $a_0 = A(0)$; $a_1 = A'(0)$.

Man beachte: In der Theorie der **formalen** Potenzreihen ist nicht vorgesehen, dass man für die Unbestimmte z irgendwelche 'Werte' einsetzt.

Operationen

1. Formale Potenzreihen kann man linear kombinieren; man hat einen unendlichdimensionalen komplexen Vektorraum.
2. Formale Potenzreihen kann man miteinander 'multiplizieren'; die Koeffizientenfolgen werden gefaltet. Man hat eine assoziative und kommutative Algebra.
3. Formale Potenzreihen kann man 'ableiten'. Die Ableitung (auch Derivation genannt) ist eine lineare Abbildung, welche die Produktregel erfüllt.

$$D : A(z) \mapsto A'(z) . \quad \sum_0^\infty a_n z^n \mapsto \sum_1^\infty a_n n z^{n-1} ;$$

$$D(A \cdot B) = (DA) \cdot B + A \cdot (DB) ; \quad (AB)' = A'B + AB' .$$

4. Eine formale Potenzreihe $F(z)$ mit $F(0) = 0$ kann man in eine formale Potenzreihe $A(z)$ 'einsetzen'. Man erhält

$$B(z) = A(F(z)) = a_0 + a_1 F(z) + a_2 F^2(z) + a_3 F^3(z) + \cdots .$$

Satz (Umkehrung)

Wenn F und G einsetzbare formale Potenzreihen sind, $F(0) = 0$; $G(0) = 0$, dann ist $H(z) = G(F(z))$ einsetzbar, und es gilt für alle A

$$A(H(z)) = A(G(F(z))) = B(F(z)) \quad \text{mit} \quad B(z) = A(G(z)).$$

Jedes einsetzbare $F(z)$ mit $F'(0) \neq 0$ besitzt eine 'Umkehrung' $G(z)$

$$F(G(z)) = 0 + 1 \cdot z + 0 \cdot z^2 + 0 \cdot z^3 + \dots = G(F(z)).$$

Satz (Reziproke)

Eine formale Potenzreihe $A(z)$ besitzt genau dann eine multiplikative Inverse $B(z)$, wenn $A(0) \neq 0$.

$$A(z) \cdot B(z) = 1 + 0 \cdot z + 0 \cdot z^2 + \dots$$

Man nennt B die reziproke formale Potenzreihe und notiert auch $B(z) = \frac{1}{A}(z)$.

Die Konstruktionen und die Beweise führt man mit der Methode des Koeffizientenvergleichs; die Annahmen garantieren, dass immer nur endliche Summen auftreten, wenn es nach und nach um die Bestimmung der gesuchten formalen Potenzreihe geht. Die Beweise überlassen wir dem Leser.

Beispiele

1. Es gibt genau eine formale Potenzreihe $E(z)$ mit $E'(z) = E(z)$ und $E(0) = 1$.

$$E(z) = 1 + a_1 z + a_2 z^2 + a_3 z^3 + \dots = a_1 + 2a_2 z + 3a_3 z^2 + 4a_4 z^3 + \dots = E'(z)$$

impliziert nämlich $a_1 = 1$, $2a_2 = a_1$, $a_2 = \frac{1}{2!}$, $3a_3 = a_2$, $a_3 = \frac{1}{3!}$, \dots , $a_n = \frac{1}{n!}$.
So gewinnen wir also die Exponentialreihe

$$E(z) = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \dots = \sum \frac{1}{n!}z^n.$$

Mit der Faltungsformel erhalten wir für $\alpha, \beta \in \mathbb{C}$

$$\begin{aligned} E(\alpha z) \cdot E(\beta z) &= (1 + \alpha z + \frac{1}{2!}\alpha^2 z^2 + \dots) \cdot (1 + \beta z + \frac{1}{2!}\beta^2 z^2 + \dots) \\ &= 1 + (\alpha + \beta)z + \frac{1}{2!}(\alpha^2 + 2\alpha\beta + \beta^2)z^2 + \dots = E((\alpha + \beta)z). \end{aligned}$$

2. Die formale Potenzreihe $E(z) - 1$ besitzt eine Umkehrung $L(z)$. Die Differentiation von $E(L(z)) - 1 = z$ ergibt

$$\begin{aligned} 1 &= E(L(z)) \cdot L'(z) = (z + 1) \cdot L'(z); \quad L'(z) = \frac{1}{1+z} = 1 - z + z^2 - z^3 + \dots \\ L(z) &= z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4 + \dots \end{aligned}$$

Diese formale Potenzreihe $L(z)$ hat dieselben Koeffizienten wie die wohlbekannte Potenzreihenentwicklung von $\ln(1+z)$.

3. Die Menge aller $A(z)$ mit $A''(z) = -A(z)$ ist ein zweidimensionaler Vektorraum. Eine bequeme Basis sind

$$\begin{aligned} C(z) &= 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \frac{1}{6!}z^6 + \dots &= \frac{1}{2} (E(iz) + E(-iz)) ; \\ S(z) &= z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \frac{1}{7!}z^7 + \dots &= \frac{1}{2i} (E(iz) - E(-iz)) ; \end{aligned}$$

4. Für jede komplexe Zahl α ist $B_\alpha(z) = E(\alpha \cdot L(z))$ eine wohldefinierte formale Potenzreihe. Ihre Koeffizienten gewinnen wir aus der Differentialgleichung

$$B'_\alpha(z) = E(\alpha \cdot L(z)) \cdot \alpha L'(z) = B_\alpha(z) \cdot \alpha \cdot \frac{1}{1+z} .$$

Der Koeffizientenvergleich zu $(1+z) \cdot B'_\alpha(z) = \alpha \cdot B_\alpha(z)$ liefert wegen

$$\begin{aligned} B_\alpha(z) &= 1 + b_1z + b_2z^2 + b_3z^3 + \dots \\ (1+z) \cdot B'_\alpha(z) &= (1+z)(b_1 + 2b_2z + 3b_3z^2 + \dots) \\ &= b_1 + (b_1 + 2b_2)z + (2b_2 + 3b_3)z^2 + \dots \end{aligned}$$

$$b_1 = \alpha, \quad b_1 + 2b_2 = \alpha b_1, \quad 2b_2 + 3b_3 = \alpha b_2, \quad 3b_3 + 4b_4 = \alpha b_3, \quad \dots$$

$$b_1 = \alpha; \quad b_2 = \frac{1}{2}(\alpha - 1)b_1 = \binom{\alpha}{2}; \quad b_3 = \frac{1}{3}(\alpha - 2)b_2 = \binom{\alpha}{3}; \quad \dots$$

Somit ergibt sich die Newton'sche Binomialreihe

$$B_\alpha(z) = 1 + \binom{\alpha}{1}z + \binom{\alpha}{2}z^2 + \binom{\alpha}{3}z^3 + \dots$$

Eine Anwendung in der kombinatorischen Geometrie

Wir betrachten ein konvexes n -Eck in der Ebene mit den Ecken P_0, P_1, \dots, P_{n-1} (im Gegenuhrzeigersinn). Mit Diagonalen (die sich nicht überschneiden sollen!) kann man dieses n -Eck auf allerlei Weisen in Dreiecke zerlegen; a_n sei die Anzahl der Möglichkeiten. $A(z) = \sum_n a_n z^n$ sei die erzeugende Funktion. Es wird sich nützlich erweisen, wenn wir die ersten Koeffizienten (durch Konvention!) folgendermassen festlegen: $a_0 = 0 = a_1$, $a_2 = 1 = a_3$. Wegen $a_4 = 2$ und $a_5 = 5$ haben wir also $A(z) = z^2 + z^3 + 2z^4 + 5z^5 + \dots$. Für $n = 6, 7, \dots$ gilt die Rekursionsformel

$$a_n = \sum_{k=2}^{n-1} a_k a_{n-k+1}.$$

Sei nämlich für eine Zerlegung des n -Ecks k der kleinste Index, für welchen die Diagonale $\overline{P_0 P_k}$ in der Zerlegung vorkommt (der Fall $k = n - 1$ entspreche dem Fall, dass keine Diagonale von P_0 ausgeht). Die Zerlegungen zu den Fällen $\{k = 2\}$ und $\{k = n - 2\}$

entsprechen den Zerlegungen eines $(n-1)$ -Ecks. Für die k dazwischen ergibt sich jede Zerlegung durch eine (eindeutig bestimmte) Zerlegung der Teilfiguren mit den Ecken P_1, P_2, \dots, P_k und P_0, P_k, \dots, P_{n-1} ; hier gibt es $a_k \cdot a_{n-k+1}$ Möglichkeiten.

Die Rekursion gilt auch für $n = 3$. Somit haben wir

$$z \cdot (A(z) - z^2) = A^2(z); \quad (A - z/2)^2 = \frac{1}{4}z^2(1 - 4z); \quad 2A(z) = z - z \cdot \sqrt{1 - 4z}.$$

Die binomische Reihe ergibt

$$2 \cdot A(z) = (-z) \sum_{n=1} \binom{1/2}{n} (-4z)^n; \quad 2a_n = (-1)^n \binom{1/2}{n-1} 4^{n-1}$$

$$a_n = \frac{2 \cdot 6 \cdot 10 \cdot \dots \cdot (4n-10)}{2 \cdot 3 \cdot 4 \cdot 5 \cdot \dots \cdot (n-1)} \quad \text{für } n = 3, 4, 5, \dots$$

Didaktische Anmerkung

Der Kalkül der **formalen Potenzreihen** erscheint heutzutage selten im Programm der Mathematikerausbildung. Die analytische Theorie der **lokal konvergenten Potenzreihen** (das ist die ‘Funktionentheorie’ im Sinne von Weierstraß (1815 -1897)) hat diesen algebraischen Kalkül fast völlig verdrängt. Gegen die Feststellung, dass der Koeffizientenvergleich in vielen Anwendungsgebieten sehr beliebt ist, machen die Mathematiker geltend, dass die Aussagen des Kalküls keine Aussagen über Größen sind; man kann für die Unbestimmte nichts ‘einsetzen’ (keine Zahl, keine Matrix, keine Funktion). Obwohl die Exponentialreihe im Kalkül sehr wohl vorkommt, liegen Aussagen wie beispielsweise $\exp(i\pi) = -1$ ausserhalb der Reichweite des Kalküls. Man schüttet manchmal den Anfängern gegenüber das Kind mit dem Bade aus, wenn man suggeriert, dass das Rechnen mit formalen Potenzreihen und insbesondere die Methode des Koeffizientenvergleichs mathematisch wertlos ist, indem die Rechnungen zwar möglicherweise Plausibilitätsargumente liefern, aber doch keine Beweiskraft besitzen.

Wir halten dagegen: Die formalen Potenzreihen leben in einem (in sich widerspruchsfreien!) formalen System. Gleichheitsbeziehungen zwischen formalen Potenzreihen können sehr wohl interessante Aussagen sein, mögen die Potenzreihen auch den Konvergenzradius 0 haben. Man muss sie eben als kombinatorische Aussagen interpretieren, so wie wir das oben im Beispiel aus der kombinatorischen Geometrie gesehen haben.

Aufgaben zu I.8

Aufgabe I.8.1 :

Analytische Vorbereitung:

- (i) Wenn $B = b_0 + b_1x + b_2x^2 + \dots$, dann ergeben sich die Koeffizienten c_n der formalen Potenzreihe $C = B^2$ folgendermaßen

$$c_0 = b_0^2, \quad c_1 = b_0b_1 + b_1b_0, \quad c_2 = b_0b_2 + b_1b_1 + b_2b_0, \quad \dots$$

- (ii) Zur formalen Potenzreihe $C = 1 - x$ gibt es genau ein B mit

$$B^2 = C = 1 - x; \quad \text{man schreibt } B = (1 - x)^{\frac{1}{2}}.$$

Die Koeffizienten drückt man mit den sog. Binomialkoeffizienten aus:

$$\binom{\alpha}{n} := \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}, \quad \alpha \in \mathbb{R}, n = 0, 1, 2, \dots$$

Man findet

$$(1-x)^{\frac{1}{2}} = B = \sum \binom{\frac{1}{2}}{n} (-x)^n = 1 - \frac{1}{2}x - \frac{1}{8}x^2 \pm \dots$$

Hier ist kein Beweis erwünscht! Sie werden die Tatsache in der folgenden geometrischen Anzahlbestimmung (Teil c) brauchen.

Geometrische Vorbereitung:

- (iii) Ein konvexes Viereck kann man auf zwei verschiedene Weisen durch das Einzeichnen einer Diagonalen in Dreiecke zerlegen.

Ein konvexes Fünfeck kann man auf a_5 verschiedene Weisen durch das Einzeichnen von sich nicht schneidenden Diagonalen in Dreiecke zerlegen. $a_5 = ?$

Entsprechend sei a_n definiert. Berechnen Sie auch a_6 .

- (iv) Konvention: $a_0 = 0, a_1 = 0, a_2 = 1, a_3 = 1$. Die Formeln werden damit übersichtlich.

- a) Zeigen Sie für alle $n \geq 5$

$$a_n = a_2a_{n-1} + a_3a_{n-2} + \dots + a_{n-1}a_2 \quad \text{Summe der Länge } n-2.$$

Anleitung: Die Ecken seien P_0, P_1, \dots, P_{n-1} (im Gegenuhrzeigersinn numeriert). Diejenigen Zerlegungen, wo keine Diagonale durch P_0 geht, entsprechen den Zerlegungen des $(n-1)$ -Ecks mit den Ecken P_1, P_2, \dots, P_{n-1} (Zeichnung!). Bei den anderen gibt es ein kleinstes k , so daß P_0P_k in der Zerlegung vorkommt; es kommt dann auch P_1P_k in der Zerlegung vor. Unsere Zerlegung ergibt sich durch Kombination einer Zerlegung der Teilfiguren mit den Ecken P_1, \dots, P_k und $P_0, P_k, P_{k+1}, \dots, P_{n-1}$.

b) Zeigen Sie für die formale Potenzreihe $A = a_0 + a_1x + a_2x^2 + \dots$

$$(*) \quad x(A - x^2) = A^2.$$

c) Zeigen Sie, daß für alle $n \geq 3$

$$a_n = \frac{2 \cdot 6 \cdot 10 \cdot 14 \cdot \dots \cdot (4n - 10)}{2 \cdot 3 \cdot 4 \cdot 5 \cdot \dots \cdot (n - 1)}.$$

II Abbildungen

Themenübersicht

Es geht hier um Geometrie.

Von der Schulmathematik her ist nicht leicht zu erschließen, was Geometrie für einen Physik-Studierenden bedeuten kann. Geometrie sollte anschaulich sein; Geometrie kann aber sicherlich nicht bei den naturwüchsigen Vorstellungen vom Anschauungsraum stehen bleiben.

Ein wichtiger Mann für die Mathematische Physik (in Deutschland und in den USA) war Richard Courant (1888–1972). Wir entnehmen einige Sätze aus dem Kapitel über Geometrie in einem Buch, welches allen Mathematik- oder Physik-Anfängern wärmstens zu empfehlen ist:

R. Courant u. H. Robbins: „*Was ist Mathematik?*“, Springer-Verlag, 5. Auflage, 2001. Dort heißt es:

Wenn irgendeine bestimmte Klasse von Transformationen einer Figur gegeben ist (wie etwa die Klasse aller starren Bewegungen oder Kompressionen oder Inversionen an Kreisen u.s.w.), so können wir fragen, welche Eigenschaften der Figur bei dieser Klasse von Transformationen ungeändert bleiben. Das System der Sätze über diese Eigenschaften ist dann die Geometrie, die dieser Klasse von Transformationen zugeordnet ist.“

Feynman äußert sich in seinen Vorlesungen, Bd I, immer wieder zur Symmetrie physikalischer Gesetze. Er übernimmt eine Definition von Hermann Weyl: Ein Ding ist symmetrisch, wenn man es einer bestimmten Operation aussetzen kann und es danach als genau das gleiche erscheint wie vor der Operation. (Kapitel 11). In Kapitel 52.3 benennt Feynman einen sehr tiefgründigen und schönen Sachverhalt in der Quantenmechanik: Für jedes Symmetriegesetz gibt es ein entsprechendes Erhaltungsgesetz. Feynman deutet dort auch einige Symmetrien in der Quantenmechanik an, welche kein klassisches Analogon besitzen und für die es keine Beschreibungsform in der klassischen Physik gibt.

Im Folgenden geben wir eine erste Einführung in die Gedankenwelt der Transformationsgruppen, wobei wir uns sehr um Anschaulichkeit im Sinne der Schulgeometrie bemühen. Die komplexen Zahlen und die 2×2 -Matrizen spielen dabei eine zentrale Rolle. Wir beginnen allerdings sehr allgemein im Sinne der abstrakten Mengenlehre.

II.1 Abbildungen, Permutationen, Gruppen

Definitionsbereich, Zielbereich und Bildbereich einer Abbildung. Injektive, surjektive, bijektive Abbildungen. Das Hintereinanderschalten von Abbildungen. Die Darstellung von Abbildungen endlicher Mengen durch Graphen. Permutationen. Ihre Zyklendarstellung. Transformationsgruppen, speziell Permutationsgruppen. Die abstrakte Definition einer Gruppe. Gruppenwirkung. Die Oktaedergruppe als Beispiel. Isomorphie von Gruppen. Untergruppen. Die alternierende Gruppe als Untergruppe der symmetrischen Gruppe. Volles Urbild. Pullback.

II.2 Die Riemann'sche Zahlenkugel; Kreisverwandtschaften

Die Spiegelung am Einheitskreis bildet Kreise in Kreise ab. Vier Beweise. Lineargebrochene Abbildungen (Möbiustransformationen) sind erzeugt von den Translationen, den Drehstreckungen und der Reziprokenabbildung. Die allgemeine Kreisgleichung in \mathbb{C} . Die Riemann'sche Zahlenkugel, stereographische Projektion. Kreise in \mathbb{C} entsprechen Kreisen auf der Zahlenkugel. Die stereographische Projektion ist winkeltreu. Spezielle Kreise.

II.3 Einige spezielle Transformationsgruppen

Gruppenwirkung, allgemeine Definition. Transitivität. Fixgruppen. Die Wirkung von $SU(2)$ auf der Riemann'schen Zahlenkugel. Die Wirkung von $SU(1, 1)$ auf dem Innern des Einheitskreises und auf der oberen Halbebene (Cayley-Transformation). Die Möbiustransformationen, die den Punkt ∞ festlassen, bilden die affine Gruppe in einer komplexen Dimension. Man kann sie auch auffassen als die orientierungserhaltenden Ähnlichkeitsabbildungen der euklidischen Ebene. Verallgemeinerung: Die Affinitäten eines Spaltenraums \mathbb{K}^J . Die Gruppe aller Möbiustransformationen wirkt einfach transitiv auf der Menge aller Zahlentripel. Das Doppelverhältnis von Quadrupeln (z_1, z_2, z_3, z_4) bleibt unter Möbiustransformationen erhalten.

II.4 Affine Räume, affine Funktionen, affine Abbildungen

Ein affiner Raum ist eine Punktmenge L , auf welcher ein Vektorraum V einfach transitiv wirkt. Der Kalkül der Ortsvektoren; Schwerpunktskoordinaten. Affine Teilräume und ihre Tangentialräume. Anschluss an die Schulmathematik: Affine und quadratische Funktionen auf dem affinen Raum aller Zahlenpaare. Amplituden ebener Wellen. Koordinatenfreier Zugang zum Begriff der affinen Funktion. Linearformen auf dem Vektorraum V ; der Dualraum V^* . Duale Basen: $\{v_j : j \in J\}$ in V und $\{\ell^j(\cdot) : j \in J\}$ in V^* . Affine Koordinatensysteme. Abstrakte Definitionen des Begriffs der affinen Abbildung eines affinen Raums. Schluss: Eine Abbildung Φ eines affinen Raums in einen affinen Raum ist genau dann eine affine Abbildung, wenn der Pullback Φ^* den affinen Funktionen auf dem Bildraum affine Funktionen auf dem Urbildraum zuordnet.

II.5 Quadratische Funktionen

Koordinatenfreier Zugang zum Begriff der quadratischen Funktion auf einem affinen Raum. Gemischte Differenzen, die Bilinearform $B(\cdot, \cdot)$ zu einer quadratischen Funktion.

$$q(P + v) = q(P) + \ell_P(v) + \frac{1}{2}Q(v) .$$

$Q(\cdot)$ ist die quadratische Form zu $B(\cdot, \cdot)$; der „Gradient“ $\ell_P(\cdot)$ hängt in affiner Weise vom Fußpunkt P ab. Quadratische Funktionen in affinen Koordinaten.

$$q(x) = q(x_0) + q'(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)^\top \cdot H \cdot (x - x_0) .$$

Koordinatenwechsel. Hinweise auf Quadriken. Die quadratische Form der Minkowski-Geometrie (spezielle Relativitätstheorie).

II.6 Polynome als Abbildungen von \mathbb{C} in sich

Der Fundamentalsatz der Algebra besagt: Für ein Polynom $p(\cdot)$ vom Grad n hat die Gleichung $p(z) = w$ für jedes w genau n Lösungen (Vielfachheiten berücksichtigt).

Erste Beweisskizze: Umlaufzahlen. Zweite Beweisskizze: Polynomiale Abbildungen sind offene Abbildungen und die stetige Funktion $|p(z)|$ nimmt ihr Minimum an. Topologische Anmerkungen von Gauss zu einem weiteren Beweis des Fundamentalsatzes des komplexen Zahlensystems. Geometrisches zur Zhukowski-Abbildung. Didaktische Thesen zur Beziehung von Algebra und Geometrie.

Historischer Anhang : Irrationalzahlen und Geometrie

Irrationalzahlen und Geometrie. Warum erscheinen Arithmetik und Geometrie traditionell (seit den alten Griechen) getrennt? Geometrisierung des Zahlbereichs \mathbb{C} bei Gauss. Arithmetisierung der Analysis im späten 19. Jahrhundert (Dedekind, Cantor).

II Abbildungen

II.1 Abbildungen, Permutationen, Gruppen

Seien M und N irgendwelche nichtleere Mengen. Wenn jedem $x \in M$ durch irgendeine Vorschrift ein Element $y \in N$ zugeordnet ist, dann spricht man von einer **Abbildung**

$$\varphi : M \rightarrow N ; \quad M \ni x \mapsto \varphi(x) \in N .$$

M heißt der **Definitionsbereich** von φ ; N heißt der **Zielbereich**. Die Menge derjenigen y , die wirklich als Bild auftreten können, heißt der **Bildbereich**

$$\text{im}\varphi = \{y : y = \varphi(x) \text{ mit } x \in M\} = \varphi(M) .$$

Spezielle Typen von Abbildungen

Von einer Abbildung $\varphi : M \rightarrow N$ sagt man

- a) φ ist **injektiv**, wenn verschiedene Urbilder verschiedene Bilder haben

$$x_1 \neq x_2 \implies \varphi(x_1) \neq \varphi(x_2)$$

- b) φ ist **surjektiv**, wenn $\text{im}\varphi = N$

$$\forall y \in N \exists x \in M : \varphi(x) = y .$$

- c) φ ist **bijektiv**, wenn φ sowohl injektiv als auch surjektiv ist.

Bei bijektiven Abbildungen existiert also für jedes $y \in N$ genau ein $x \in M$ mit $\varphi(x) = y$. Eine bijektive Abbildung nennt man auch eine Bijektion von M auf N . Zu einer Bijektion $\varphi : M \rightarrow N$ gibt es eine wohlbestimmte **Umkehrabbildung**. Es gibt also eine Abbildung $\psi : N \rightarrow M$, sodass gilt

$$\begin{aligned} \varphi(\psi(y)) &= y \text{ für alle } y \in N \\ \psi(\varphi(x)) &= x \text{ für alle } x \in M . \end{aligned}$$

Man schreibt $\psi(y) = \varphi^{-1}(y)$. Wir werden sehen, dass das Symbol $\varphi^{-1}(\cdot)$ (leider!) auch noch für andere Konstruktionen benutzt wird, z. B. für die ‘Volle-Urbild-Abbildung’. (Siehe unten)

Das Hintereinanderschalten von Abbildungen

Seien $\varphi : M_1 \rightarrow M_2$ und $\psi : M_2 \rightarrow M_3$ Abbildungen.

Durch Hintereinanderschalten gewinnt man eine Abbildung $\chi : M_1 \rightarrow M_3$.

$$\begin{array}{ccccc} M_1 & \xrightarrow{\varphi} & M_2 & \xrightarrow{\psi} & M_3 \\ & \searrow \chi \nearrow & & & \end{array}$$

Man notiert $\chi(\cdot) = \psi(\varphi(\cdot)) = \psi \circ \varphi(\cdot)$. Beachte die Reihenfolge!

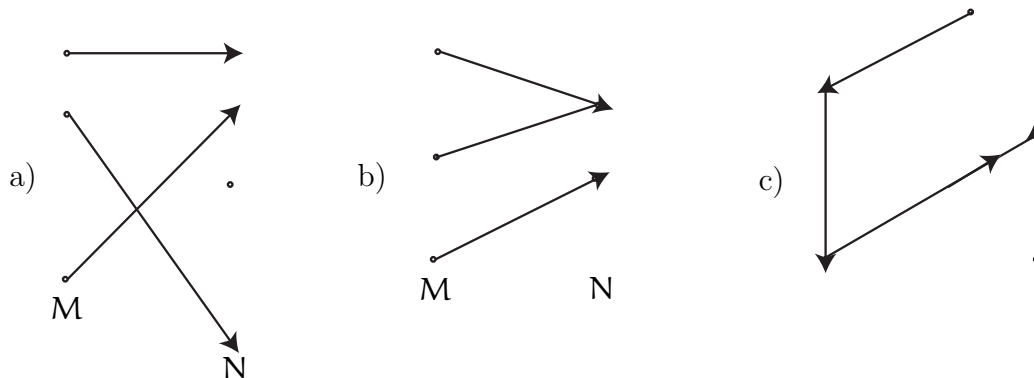
Graphen zur Visualisierung von Abbildungen endlicher Mengen

Seien S und K endliche Mengen. Jedem $k \in K$ sei ein Punktepaar $(\alpha(k), \beta(k)) \in S \times S$ zugeordnet. Man spricht dann von einem (gerichteten) Graphen mit der Scheitelmengemenge S und der Kantenmenge K .

Spezialfälle:

- Wenn $S = M + N$ (disjunkte Vereinigung) und alle Kanten von M nach N gehen, spricht man von einem bipartiten Graphen.
- Jede Abbildung $\varphi : M \rightarrow N$ (M, N endlich) kann man durch einen bipartiten Graphen visualisieren; die Kanten führen von $x \in M$ nach $\varphi(x) \in N$. Ein bipartiter Graph über $S = M + N$ gehört genau dann zu einer Abbildung von M nach N , wenn von jedem $x \in M$ genau eine Kante ausgeht und diese nach einem Scheitel in N zeigt.
- Eine Abbildung von S in sich, $\varphi : S \rightarrow S$, wird durch den Graphen über S visualisiert, dessen Kanten von x nach $\varphi(x)$ weisen (für alle $x \in S$). Ein Graph über S gehört genau dann zu einer Abbildung von S in sich, wenn von jedem Scheitel genau eine Kante ausgeht.
- Eine Bijektion von S erscheint als ein Graph, bei welchem jeder Scheitel genau einmal als Anfangspunkt und genau einmal als Endpunkte einer Kante auftritt.

Beispiele

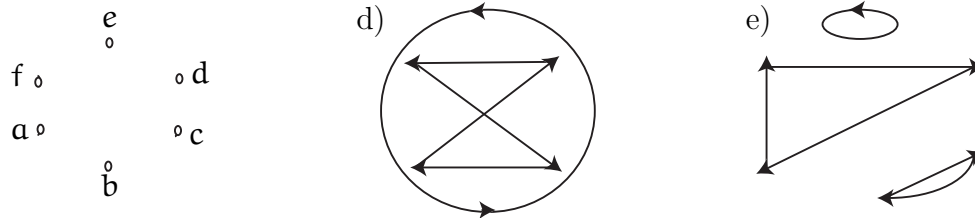


a) und b) stellen Abbildungen von M nach N dar. Die Abbildung zu a) ist injektiv, die zu b) ist surjektiv; c) stellt eine Abbildung einer fünfpunktigen Menge in sich dar, die weder injektiv noch surjektiv ist.

Permutationen

Eine bijektive Abbildung einer endlichen Menge S auf sich heißt eine **Permutation** auf S (oder eine Permutation der Elemente von S).

Die Graphen d) und e) zeigen Permutationen über der sechspunktigen Scheitelmenge $S = \{a, b, c, d, e, f\}$



Notation Für Permutationen sind verschiedene algebraische Darstellungen im Gebrauch, für Beispiel d)

$$\begin{pmatrix} a & b & c & d & e & f \\ d & e & a & f & b & c \end{pmatrix} \quad \text{oder} \quad (a \ d \ f \ c) (b \ e)$$

(„Zyklendarstellung“)

für Beispiel e)

$$\begin{pmatrix} a & b & c & d & e & f \\ f & c & b & a & e & d \end{pmatrix} \quad \text{oder} \quad (a \ f \ d) (b \ c) (e)$$

(„Zyklendarstellung“)

Sprechweisen

Die Menge aller Permutationen einer Grundmenge der Mächtigkeit n heisst die symmetrische Gruppe über n Objekten. Sie wird mit \mathcal{S}_n bezeichnet und hat $n!$ Elemente.

Die Menge aller geraden Permutationen heisst die alternierende Gruppe über n Objekten. Sie wird mit \mathcal{A}_n bezeichnet und hat $n!/2$ Elemente.

(Was eine gerade Permutation ist, werden wir unten diskutieren.)

Bemerke :

Das Wort Permutation wird (z.B. im Schulunterricht) auch für eine Anordnung (von n Objekten) gebraucht. Wir wollen aber den dynamischen Aspekt betonen: Permutationen sind Bijektionen, die man hintereinanderschalten kann.

Definition (Permutationsgruppe)

Eine Menge \mathcal{P} von bijektiven Abbildungen einer endlichen Menge Ω heisst eine Permutationsgruppe über Ω , wenn gilt

- (i) die Identitätsabbildung I_Ω gehört zu \mathcal{P}
- (ii) wenn $\pi(\cdot)$ zu \mathcal{P} gehört, dann auch die Umkehrabbildung $\pi^{-1}(\cdot)$
- (iii) wenn $\pi'(\cdot)$ und $\pi''(\cdot)$ zu \mathcal{P} gehören, dann auch die zusammengesetzte Abbildung $\pi' \circ \pi''(\cdot)$.

Allgemeinere Gruppen

Definition (Transformationsgruppe)

Eine Menge G von bijektiven Abbildungen einer Grundmenge Ω heißt eine Transformationsgruppe, wenn gilt

- (i) die Identitätsabbildung I_Ω gehört zu G
- (ii) wenn $\varphi(\cdot)$ zu G gehört, dann auch die Umkehrabbildung $\varphi^{-1}(\cdot)$
- (iii) wenn $\varphi(\cdot)$ und $\psi(\cdot)$ zu G gehören, dann auch die zusammengesetzte Abbildung $\varphi \circ \psi(\cdot)$.

Definition (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 (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.

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

Jede Bijektion $i(\cdot)$, die das leistet, heißt ein *Gruppenisomorphismus* (von G' auf G'').

Definition (Untergruppe)

Eine Teilmenge U einer Gruppe G heißt eine Untergruppe von (G, e, \circ) , wenn gilt

- (i) $e \in U$
- (ii) $g \in U \Rightarrow g^{-1} \in U$
- (iii) $g, h \in U \Rightarrow g \circ h \in U$

Ergänzend definieren wir noch den Begriff der Gruppenwirkung

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

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

Man sagt dann, dass die Gruppe G (von links) auf Ω wirkt.

Beispiel („Oktaedergruppe“)

Sei G die Menge aller derjenigen Drehungen des dreidimensionalen Anschauungsraums, die den Einheitswürfel in sich überführen.

Leute mit guter geometrischer Anschauung können sehen, dass G 24 Elemente hat.

Um jede der drei Koordinatenachsen kann man um 90° , 180° und 270° drehen. (Die Identität lassen wir vorerst aus dem Spiel.)

Um jede der vier Raumdiagonalen kann man um 120° und 240° drehen.

Um jede der 12 Kanten kann man „klappen“; gegenüberliegende Kanten führen jedoch zur gleichen Drehung. Schließlich kommt die Identität dazu

$$3 \cdot 3 + 4 \cdot 2 + 6 + 1 = 24 .$$

Zu unserer Gruppe G gehört eine „Permutationsgruppe“ auf der Menge F aller Seitenflächen ($|F| = 6$); nennen wir sie G_F . Zu G gehört auch eine Permutationsgruppe über E , der Eckenmenge mit $|E| = 8$ Elementen; nennen wir sie G_E . Entsprechend können wir eine Permutationsgruppe G_K auf der Kantenmenge K gewinnen. Die Gruppen sind offenbar allesamt isomorph. Da sie uns andererseits als Untergruppen verschiedener symmetrischer Gruppen \mathcal{S}_n gegeben sind, handelt es sich um verschiedene Transformationsgruppen.

Hinweis

Im Folgenden werden wir uns nicht nur mit endlichen Gruppen befassen, sondern auch mit kontinuierlichen Gruppen, insbesondere mit kontinuierlichen Untergruppen der allgemeinen linearen Gruppe $GL(2, \mathbb{C})$. (Die allgemeine lineare Gruppe $GL(n, \mathbb{K})$ ist die Gruppe aller invertierbaren $n \times n$ -Matrizen mit Einträgen aus dem Körper \mathbb{K} . Die Untergruppe der Matrizen mit der Determinante $= 1$ heisst die spezielle lineare Gruppe $SL(n, \mathbb{K})$.)

Zuerst wollen wir aber noch einmal zu den Permutationen zurückkommen.

Gerade und ungerade Permutationen

Die einfachsten Permutationen (neben der Identität) sind die Transpositionen. Sie vertauschen zwei Elemente der Grundmenge und lassen alle anderen fest. In der Zyklennotation werden sie durch einen einzigen Zweierzyklus beschrieben. Statt $(a, b)(c)(d)(e)(f)$ schreibt man übrigens einfach (a, b) ; die ‘Einerzyklen’ sind die Identität.

Satz

Jede Permutation kann man durch das Hintereinanderschalten von Transpositionen gewinnen.

Anders gesagt: Aus jeder Anordnung von n Objekten kann man jede andere Anordnung dadurch gewinnen, dass man Transpositionen hintereinander schaltet.

Der Beweis liegt auf der Hand.

Definition

Eine Permutation heißt eine ungerade (bzw gerade) Permutation, wenn sie als das Produkt einer ungeraden (bzw. geraden) Anzahl von Transpositionen gewonnen werden kann.

Um die Sinnhaftigkeit dieser Definition zu erweisen, müssen wir einen Satz beweisen.

Satz

Das Produkt einer ungeraden Anzahl von Transpositionen läßt sich nicht als Produkt einer geraden Anzahl von Transpositionen gewinnen.

Beweis

Wir bringen die Elemente der n -punktigen Menge Ω in eine Reihenfolge. Anders gesagt: Wir identifizieren Ω mit $\{1, 2, \dots, n\}$. Eine Permutation π denken wir uns dadurch gegeben, dass durchnummerierte Objekte auf durchnummerierte Plätze verteilt werden.

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & & \pi(n) \end{pmatrix}.$$

Wir sagen, dass das Zahlenpaar $i < j$ eine Fehlstellung verursacht, wenn $\pi(i) > \pi(j)$. $N(\pi)$ sei die Anzahl der Fehlstellungen für die Permutation π . Die Identität hat keine Fehlstellungen; die inverse Permutation π^{-1} hat dieselbe Anzahl von Fehlstellungen wie π . Offenbar gilt: Ist τ eine Transposition benachbarter Objekte, so gilt $N(\tau \circ \pi) = N(\pi) \pm 1$. Entweder haben die Objekte $\pi(i), \pi(i+1)$ eine Fehlstellung geliefert, die durch τ aufgehoben wird, oder aber τ verursacht eine zusätzliche Fehlstellung. Die Gleichung

$$(-1)^{N(\tau \circ \pi)} = (-1)^{N(\pi)} \cdot (-1)^{N(\tau)}$$

gilt nicht nur für die Transpositionen benachbarter Elemente, sondern allgemein für alle π und τ ; denn jede Permutation τ kann man durch das Hintereinanderschalten von Transpositionen benachbarter Elemente erzeugen.

Die alternierende Gruppe ist die Menge der Permutationen, die durch eine gerade Anzahl von Transpositionen gewonnen werden können.

Zum Abschluss dieser Vorlesung wollen wir noch einmal sehr allgemein über Abbildungen beliebiger Mengen sprechen.

Volles Urbild und Pullback

Sei $\varphi : M \rightarrow N$ eine Abbildung. Für eine Teilmenge $N' \subseteq N$ nennt man die Menge M' aller derjenigen $x \in M$, deren Bild in N' liegt, das **volle Urbild**. Man notiert

$$M' = \{x : \varphi(x) \in N'\} = \varphi^{-1}(N').$$

Die „Volle-Urbild-Abbildung“ $\varphi^{-1}(\cdot)$ bildet Teilmengen von N in Teilmengen von M ab. Dabei gilt

$$\begin{aligned} \varphi^{-1}(N) &= M, \quad \varphi^{-1}(\emptyset) = \emptyset \\ \varphi^{-1}(N \setminus B) &= M \setminus \varphi^{-1}(B) \text{ für alle } B \subseteq N \\ \varphi^{-1}(B_1 \cup B_2) &= \varphi^{-1}(B_1) \cup \varphi^{-1}(B_2). \end{aligned}$$

Allgemeiner: Für jede Schar von Teilmengen $\{B_j : j \in J\}$ gilt

$$\varphi^{-1}\left(\bigcup_j B_j\right) = \bigcup_j \varphi^{-1}(B_j), \quad \varphi^{-1}\left(\bigcap_j B_j\right) = \bigcap_j \varphi^{-1}(B_j).$$

Bemerke : Wenn $\varphi : M \rightarrow N$ bijektiv ist, dann ist das volle Urbild jeder einpunktigen Menge eine einpunktige Menge

$$\varphi^{-1}(\{y\}) = \{x\} \Leftrightarrow y = \varphi(x) .$$

Hier könnte für den Anfänger die Gefahr bestehen, dass er die Volle-Urbild-Abbildung, die eine Mengenabbildung ist, verwechselt mit der Umkehrabbildung, die eine Punktabbildung ist.

Sprechweisen

Sei $\varphi : M \rightarrow N$ eine beliebige Abbildung. Man ordnet jeder (reellwertigen) Funktion g auf N eine (reellwertige) Funktion $f = \varphi^*(g)$ zu

$$f(x) = g(\varphi(x)) \text{ für alle } x \in M .$$

f heißt die **zurückgenommene Funktion**; φ^* heißt die **Pullback-Abbildung** (für reellwertige Funktionen).

Bemerke : Der Pullback φ^* ordnet der punktweisen Summe zweier Funktionen die punktweise Summe der Bildfunktionen zu. Ebenso steht es um punktweise Produkte.

$$\begin{aligned} \varphi^*(\alpha g_1 + \beta g_2) &= \alpha \cdot \varphi^*(g_1) + \beta \varphi^*(g_2) \\ \varphi^*(g_1 \cdot g_2) &= \varphi^*(g_1) \cdot \varphi^*(g_2) . \end{aligned}$$

In derselben Weise gewinnt man die Pullback-Abbildung für \mathbb{C} -wertige Funktionen und allgemein die Pullback-Abbildung für Funktionen mit irgend einem Wertebereich. Man benützt immer dasselbe Symbol φ^* .

Fazit : Der Pullback φ^* bildet Funktionen auf dem Zielraum in wohlbestimmte Funktionen auf dem Urbildraum ab.

Spezialfall : Jede $\{0, 1\}$ -wertige Funktion auf dem Urbildraum kann man mit einer Teilmenge von Ω identifizieren.

$$1_A(\omega) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A \end{cases}$$

Die Funktion 1_A heißt die **Indikatorfunktion** der Menge A . Der Pullback bildet Indikatorfunktionen in Indikatorfunktionen ab.

$$\varphi^*(1_{A_2}) = 1_A \Leftrightarrow A_1 = \varphi^{-1}(A_2) .$$

Man kann somit den Pullback φ^* als eine Verallgemeinerung der Volle-Urbild-Abbildung auffassen.

Warnung Es erweist sich nicht als förderlich für die intuitive Vorstellung vom Pullback $\varphi^*(\cdot)$, wenn man dabei stehen bleibt, dass jede zurückgenommene Funktion $\varphi^*(f)$ durch eine Verkettung (d. h. Hintereinanderschaltung) von ‘Abbildungen’ entsteht. Man sollte nicht an die einzelnen Funktionen $\varphi^*(f)$ denken, sondern vielmehr an die Abbildung $\varphi^*(\cdot)$ des betreffenden Funktionenraums.

Aufgaben zu II.1

Aufgabe II.1.1 :

Die Elemente der Gruppe \mathcal{S}_6 wollen wir als Bijektionen der Menge $\Omega = \{a, b, c, d, e, f\}$ verstehen und als Produkte von elementfremden Zyklen schreiben.

- a) Berechnen Sie (d. h. finden Sie die disjunkte Zykeldarstellung für) die Produkte:

$$\begin{aligned}\pi' &= (a, b) (b, c, d, e, f) (a, b) = & ? \\ \pi'' &= (b, c) (b, c, d, e, f) (b, c) = & ? \\ \pi''' &= (b, e) (b, c, d, e, f) (b, e) = & ? \\ \pi'''' &= (a, d) (a, b, c) (d, e, f) (a, d) = & ?\end{aligned}$$

- b) Zeigen Sie, dass zwei beliebige k -Zyklen in der symmetrischen Gruppe zueinander konjugiert sind. M. a. W. : Wenn π' und π'' Zyklen der Länge k sind, dann existiert eine Permutation g , sodass $g \pi' g^{-1} = \pi''$.

Aufgabe II.1.2

Seien φ, ψ Abbildungen und χ die zusammengesetzte Abbildung.

Zeigen Sie

- a) Wenn φ und ψ injektiv sind, dann auch χ .
b) Wenn φ und ψ surjektiv sind, dann auch χ .

Es seien M und N irgendwelche (endliche) Mengen. Mit \mathbb{R}^M bezeichnen wir die Menge aller reellwertigen Funktionen auf M ; ebenso ist \mathbb{R}^N definiert. Wir definieren eine Abbildung ('Pullbackabbildung')

$$\varphi^* : \mathbb{R}^M \longleftarrow \mathbb{R}^N ; \quad g(y) \longmapsto f(x) = \varphi^*(g)(x) = g(\varphi(x)) .$$

Zeigen Sie

- c) $\varphi^*(\cdot)$ ist genau dann injektiv, wenn $\varphi(\cdot)$ surjektiv ist
d) $\varphi^*(\cdot)$ ist genau dann surjektiv, wenn $\varphi(\cdot)$ injektiv ist.

II.2 Die Riemann'sche Zahlenkugel; Kreisverwandtschaften

Wir fügen der komplexen Zahlenebene \mathbb{C} einen zusätzlichen Punkt hinzu, den „unendlich fernen“ Punkt ∞ . So erhalten wir die erweiterte Zahlenebene $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Die Menge $\bar{\mathbb{C}}$ werden wir später mit der Riemann'schen Zahlenkugel identifizieren.

In dieser Vorlesung wollen wir einen speziellen Typ von Bijektionen der Menge $\bar{\mathbb{C}}$ studieren. Sie heißen die Kreisverwandtschaften, weil sie Kreise in Kreise überführen. (Die Geraden gelten in diesem Zusammenhang als die Kreise durch den unendlich fernen Punkt). Ein wichtiges Beispiel einer Kreisverwandtschaft ist die „Reziprokenabbildung“: $z \mapsto \frac{1}{z}$. (Die Null geht in den Punkt ∞ über, und ∞ geht nach 0). Wir werden mehrere Beweise führen für die Tatsache, dass die Reziprokenabbildung **Rez** Kreise in Kreise überführt.

Erster Beweis

1. Vorbemerkung: Wenn man ein Figur K dreht, dann ergibt die Reziprokenabbildung ein gedrehtes Bild

$$\text{Rez}(e^{i\varphi} \cdot K) = e^{-i\varphi} \cdot (\text{Rez}K) .$$

Wenn man die Figur mit einem reellen Faktor streckt, dann wird das Bild um diesen Faktor gestaucht

$$\text{Rez}(\alpha \cdot K) = \frac{1}{\alpha} \cdot \text{Rez}(K) .$$

Es genügt also, wenn wir zeigen, dass die Kreise mit dem Mittelpunkt $z_0 = 1$ in Kreise übergeführt werden.

2. Der Kreis mit dem Mittelpunkt $z_0 = 1$ und dem Radius $r = 1$ ist ein Sonderfall; er geht durch den Nullpunkt. Wir müssen uns überlegen, dass er durch die Reziprokenabbildung in eine Gerade übergeführt wird. Es ist auch schon klar, welche Gerade in Frage kommt. Da der Kreis durch den Punkt $\tilde{z} = 2$ geht, muss das Bild durch den Punkt $\tilde{w} = \frac{1}{2}$ gehen.

In der Tat ist das Bild des Kreises

$$K = \{z : z = 1 + e^{it}, \quad t \in \mathbb{R}\}$$

die Menge der Punkte

$$\frac{1}{1 + e^{it}} = \frac{1}{2} + \frac{1}{2} \frac{1 - e^{it}}{1 + e^{it}} = \frac{1}{2} - \frac{1}{2i} \frac{\sin t/2}{\cos t/2} = \frac{1}{2} + \frac{i}{2} \tan(t/2) .$$

Wenn t von $-\pi$ bis $+\pi$ läuft, dann läuft $\tan t/2$ von $-\infty$ bis $+\infty$.

3. Wenn wir schon wüssten, dass das Bild des Kreises

$$K = \{z : z = 1 + r \cdot e^{it}\}$$

ein Kreis ist, dann wäre klar, dass es der Kreis durch die Punkte $\frac{1}{1-r}$ und $\frac{1}{1+r}$ sein muss. Der Mittelpunkt w_0 des Bildkreises müsste also $w_0 = \frac{1}{2} \left(\frac{1}{1-r} + \frac{1}{1+r} \right)$ sein und

der Radius $R = \frac{1}{2} \left(\frac{1}{1-r} - \frac{1}{1+r} \right)$. In der Tat gilt nun

$$\begin{aligned} \frac{1}{1+re^{it}} - w_0 &= \frac{1}{1+re^{it}} - \frac{1}{1-r^2} = R \cdot g(t) \quad \text{mit} \\ g(t) &= \frac{1-r^2}{r} \left[\frac{1}{1+re^{it}} - \frac{1}{1-r^2} \right] = \\ &= \frac{1}{r(1+re^{it})} [(1-r^2) - (1+re^{it})] = (-1)e^{it} \frac{1+r \cdot e^{-it}}{1+r \cdot e^{it}}. \end{aligned}$$

Offenbar gilt $|g(t)| = 1$; und damit ist alles bewiesen.

Bemerke Die Abbildung $z \mapsto \frac{1}{\bar{z}}$ heißt die Spiegelung am Einheitskreis. Die Punkte des Einheitskreises sind Fixpunkte für diese Abbildung. Die Strahlen durch den Ursprung werden auf sich abgebildet.

$$z = r \cdot e^{i\varphi} \Rightarrow \frac{1}{\bar{z}} = \frac{1}{r} \cdot e^{i\varphi}.$$

Bemerke : Das Innere des Einheitskreises wird auf das Äussere abgebildet. Bei der Spiegelung am Einheitskreis wird die Orientierung geändert; bei der Reziprokenabbildung bleibt die Orientierung unverändert.

Definition

Wenn A eine nichtsinguläre komplexe 2×2 -Matrix ist,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{mit} \quad ad - bc \neq 0,$$

dann heißt die Abbildung

$$\varphi_A(\cdot) : z \mapsto \frac{az + b}{cz + d}$$

die Möbiustransformation zu A .

Bemerke :

Zwei Matrizen, die sich nur um einen Faktor unterscheiden, liefern dieselbe Möbiustransformation. Jede Möbiustransformation $\psi(\cdot)$ kann also mit einer Matrix mit der Determinante 1 realisiert werden. A und $-A$ liefern dieselbe Möbiustransformation. Matrizen, die nicht proportional sind, liefern verschiedene Transformationen. wie man leicht sieht, z.B. mit Hilfe des folgenden Satzes.

Satz 1

a) Jede Möbiustransformation $\varphi_A(\cdot)$ ist eine bijektive Abbildung von $\bar{\mathbb{C}}$ auf sich mit

$$\varphi_A(\infty) = \frac{a}{c}, \quad \varphi_A\left(-\frac{d}{c}\right) = \infty, \quad \varphi_A(0) = \frac{b}{d}.$$

b) Die Umkehrabbildung ist eine Möbiustransformation.

c) $\varphi_B(\varphi_A(\cdot)) = \varphi_{B \cdot A}(\cdot)$ für alle A, B .

Beweis von b)

$$\begin{aligned}
 w &= \frac{az + b}{cz + d} \Leftrightarrow (cz + d)w = az + b \Leftrightarrow \\
 &\Leftrightarrow z(cw - a) = -dw + b \Leftrightarrow z = \frac{dw - b}{-cw + a} .
 \end{aligned}$$

Beweis von c)

Sei $B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, also $B \cdot A = \begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + \delta a \end{pmatrix}$.

Wir haben

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

Elementare Möbiustransformationen

1. Die Abbildung $z \mapsto z + b$ ist eine Translation um b .
2. Die Abbildung $z \mapsto az$ ist eine Drehstreckung.
3. Die Möbiustransformation zu $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ oder zu $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ liefert die Reziprokenabbildung.

Satz 2

Jede Möbiustransformation kann man durch Hintereinanderschalten elementarer Möbiustransformationen gewinnen.

Beweis

$$\psi(z) = \frac{az + b}{cz + d} .$$

Im Falle $c = 0$ reicht eine Drehstreckung gefolgt von einer Translation (oder auch in umgekehrter Reihenfolge). Das ist der Fall, wo der Punkt ∞ fest bleibt.

Im Falle $c \neq 0$ können wir $c = 1$ annehmen.

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

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

Satz 3 (Kreisverwandtschaften)

Jede Möbiustransformation bildet Kreise auf Kreise ab.

Beweis

Die elementaren Möbiustransformationen haben diese Eigenschaft. Für die Translationen und die Drehstreckungen ist das klar. Für die Reziprokenabbildung haben wir oben einen ersten Beweis geliefert. Unser Beweis hat sich auf eine Parametrisierung der Kreise gestützt. Man kann Kreise auch durch Gleichungen beschreiben.

$$\{z : z = z_0 + r \cdot e^{it}\} = K = \{z : |z - z_0|^2 = r^2\}.$$

Die Kreisgleichung müssen wir noch etwas umformen, damit sie auch die Kreise durch den unendlich fernen Punkt erfasst. Für die gewöhnlichen Kreise haben wir

$$(z - z_0)(\bar{z} - \bar{z}_0) - r^2 = 0 \quad \text{oder} \quad z \cdot \bar{z} - z_0 \cdot \bar{z} - \bar{z}_0 \cdot z + |z_0|^2 - r^2 = 0$$

oder auch

$$a \cdot z\bar{z} + \bar{\alpha} \cdot z + \alpha \bar{z} + b = 0,$$

wobei zu fordern ist $a, b \in \mathbb{R}$, $\alpha \in \mathbb{C}$, $D := |\alpha|^2 - ab > 0$.

Der Fall $a = 0$ liefert die Geradengleichung. In der Tat

$$\begin{aligned} \bar{\alpha}z + \alpha\bar{z} + b = 0 &\Leftrightarrow (\bar{\alpha} + \alpha)x + (\bar{\alpha} - \alpha)iy + b = 0 \Leftrightarrow \\ &\Leftrightarrow (2\Re\alpha) \cdot x + (2\Im\alpha)y + b = 0. \end{aligned}$$

Satz 4 (Zweiter Beweis für die Kreisverwandtschaft)

Es genüge z der Kreisgleichung

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

Dann genügt $w = \frac{1}{z}$ ebenfalls einer Kreisgleichung.

Beweis

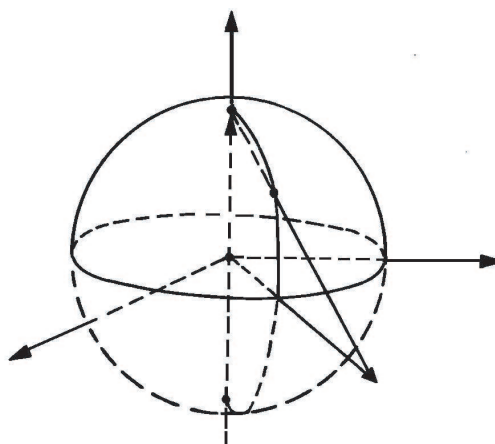
$$0 = a + \bar{\alpha} \cdot \frac{1}{\bar{z}} + \alpha \cdot \frac{1}{z} + b \cdot \frac{1}{|z|^2} = a + \bar{\alpha} \cdot \bar{w} + \alpha w + b \cdot |w|^2.$$

Die Bedingung $|\alpha|^2 > ab$ ändert ihre Gestalt nicht.

Die Riemann'sche Zahlenkugel

Die stereographische Projektion setzt die Punkte der Einheitssphäre 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. Wenn wir den Nordpol mit dem unendlich fernen Punkt von \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

$$z = x + iy \leftrightarrow (tx, ty, 1 - t) \quad \text{mit} \quad t = \frac{2}{1 + |z|^2}.$$

**Lemma**

Im euklidischen $\mathbb{R}^3 = \{(x, y, u)\}$ sei K der Kegel mit dem Scheitel $N = (0, 0, 1)$, welche die (x, y) -Ebene in einem Kreis $(x - x_0)^2 + (y - y_0)^2 = r^2$ schneidet. Die Schnittpunkte dieses Kegels mit der Einheitssphäre $\{x^2 + y^2 + u^2 = 1\}$ liegen in einer Ebene. Das Schnittgebilde ist ein Kreis auf der Sphäre.

Beweis

Die Gleichung des Kegels ist

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

Wir subtrahieren die Gleichung

$$x^2 + y^2 = 1 - u^2$$

und erhalten eine Gleichung, aus welcher man den Faktor $1 - u$ ausklammern kann.

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

Dies ist eine lineare Gleichung

$$ax + by + cu = d. \quad \text{q.e.d.}$$

Wir haben also den

Satz 5

Die stereographische Projektion bildet die Kreise auf der Riemann'schen Zahlenkugel auf die Kreise in $\bar{\mathbb{C}}$ ab (und umgekehrt).

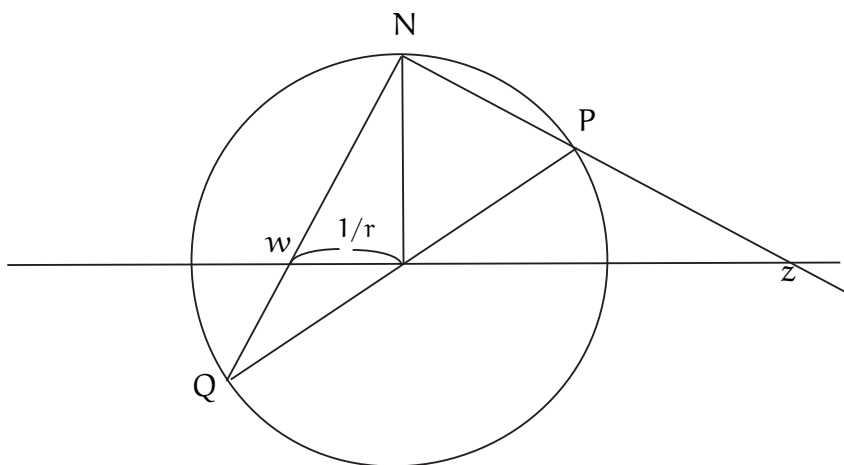
Corollar :

Wenn man die Möbiustransformationen als Bijektionen der Riemann'schen Zahlenkugel auffasst, dann bilden sie Kreise in Kreise ab. Wir werden sehen, dass der Nordpol keine ausgezeichnete Rolle spielt. Alle Drehungen der Riemann'schen Zahlenkugel lassen sich durch spezielle Möbiustransformationen realisieren.

Die Spiegelung am Einheitskreis ist im stereographischen Bild die Spiegelung an der Äquatorebene.

Zwei komplexe Zahlen z, w gehören zu Diametralpunkten P, Q auf der Riemann'schen Zahlenkugel, wenn

$$z \cdot \bar{w} = -1 \text{ , } \text{ d.h. } z = r \cdot e^{i\varphi}, w = -\frac{1}{r}e^{i\varphi} \text{ .}$$



Satz 6

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

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

Spezielle Kreise

- Ein Kreis schneidet den Einheitskreis genau dann in einem rechten Winkel, wenn er durch die Spiegelung am Einheitskreis in sich übergeführt wird. Das bedeutet $\mathbf{a} - \mathbf{b} = 0$ in der Kreisgleichung.

$$a \cdot |z|^2 + 2 \cdot \Re(\bar{\alpha}z) + b = 0 \quad \text{mit} \quad |\alpha|^2 - ab > 0.$$

- Ein Großkreis auf der Zahlenkugel ist dadurch gekennzeichnet, dass er den Äquator in diametralen Punkten schneidet. Dies bedeutet $\mathbf{a} + \mathbf{b} = 0$.

Hinweise auf nichteuklidische Geometrie Die Riemann'sche Zahlenkugel wird gern herangezogen, wenn es gilt, Modelle der nichteuklidischen Geometrien zu konstruieren.

Im Klein'schen Modell der hyperbolischen Ebene werden die Punkte auf der südlichen Halbkugel als die Punkte interpretiert, und die auf dem Äquator senkrechten Kreise als die Geraden. Zu je zwei Punkten gibt es genau eine sie verbindende 'Gerade'. Zwei 'Gerade' müssen sich nicht schneiden; durch einen Punkt ausserhalb einer Geraden gibt es eine rechtsseitige und eine linksseitige 'Parallele', das sind Geraden, welche die gegebene Gerade im 'Unendlichen' schneiden.

Die südliche Halbkugel kann entweder durch die orthogonale oder durch die stereographische Projektion bijektiv auf den Einheitskreis abgebildet werden. Im ersten Fall erscheinen die 'Geraden' als Kreissehnen, im zweiten Fall als Kreise, die auf dem Einheitskreis senkrecht stehen. Manchmal bildet man dann noch den Einheitskreis mit Hilfe der Cayley-Transformation (siehe unten) winkelerhaltend auf die obere Halbebene ab. In diesem Modell entsprechen dann Punkte der hyperbolischen Ebene den Punkten der oberen Halbebene, während die 'Geraden' den auf der reellen Achse senkrechten Kreisen entsprechen. Die Winkel der hyperbolischen Ebene sind durch dieses Modell getreu wiedergegeben; die Winkelsumme in einem 'Dreieck' ist stets kleiner als π . Darauf und auf die Abstandsbestimmung wollen wir hier aber nicht eingehen.

In der sphärischen Geometrie spielen die Großkreise die Rolle der Geraden. Durch je zwei Punkte, die nicht zueinander diametral liegen, gibt es genau einen Großkreis. Zwei Großkreise schneiden sich immer, und zwar in zwei Punkten, die zueinander diametral liegen. Man betrachtet in der sphärischen Geometrie wie in der euklidischen Geometrie 'Dreiecke'. Hier stellt man aber fest, dass die Winkelsumme keineswegs gleich zwei rechten Winkeln ist; sie hat vielmehr etwas mit der Fläche des Dreiecks zu tun hat.– Auch die 'Sphärische Geometrie' soll uns hier nicht weiter beschäftigen.

Eine sehr überzeugende untechnische Einführung in die Gedankenwelt der nichteuklidischen Geometrie und speziell der hyperbolischen Ebene findet man in dem bereits erwähnten Buch

R. Courant und H. Robbins: „Was ist Mathematik?“ Springer-Verlag, 5. Auflage 2001, S.130–179.

Aufgaben zu II.2**Aufgabe II.2.1 :**

Wir sagen von zwei Punkten z_0, z_1 in \mathbb{C} , dass sie zueinander diametral liegen, wenn gilt $z_0 \cdot \bar{z}_1 = -1$. ($z_0 = re^{i\phi} \Rightarrow z_1 = -\frac{1}{r}e^{i\phi}$).

Ausserdem wird verabredet, dass der Nullpunkt und der unendlich ferne Punkt zueinander diametral sind.

(Man bemerke: Die Bildpunkte bzgl. der stereographischen Abbildung sind Diametralpunkte auf der Riemann'schen Zahlenkugel.)

a) Es sei $\varphi(\cdot)$ eine Möbiustransformation der Form

$$z \mapsto \varphi(z) = \frac{az + b}{-\bar{b}z + \bar{a}}.$$

Zeigen sie, dass zueinander diametrale Punkte in zueinander diametrale Punkte abgebildet werden.

b) Gibt es noch weitere Möbiustransformationen, welche jedes Paar von diametralen Punkten in ein Paar von diametralen Punkten abbilden?

II.3 Einige spezielle Transformationsgruppen

Definition („Group action“)

(G, e, \circ) sei eine abstrakte Gruppe, Ω sei eine Menge, jedem $g \in G$ sei eine Bijektion von Ω zugeordnet. Das Tripel

$$(G, \Omega, \{\varphi_g(\cdot) : g \in G\})$$

heißt eine Gruppenwirkung, wenn gilt

$$\begin{aligned}\varphi_e &= \text{Id}_\Omega \\ \varphi_h(\varphi_g(\cdot)) &= \varphi_{h \circ g}(\cdot) \quad \text{für alle } g, h \in G.\end{aligned}$$

Man sagt in diesem Fall, dass G auf Ω (von links) wirkt („ G acts on Ω from the left“).

Bemerkung : Wenn nur $\varphi_e(\cdot)$ als die identische Transformation (die alle Punkte ω festläßt) wirkt, spricht man von einer *effektiven* Gruppenwirkung. In diesem Fall ist g durch $\varphi_g(\cdot)$ eindeutig bestimmt (für alle $g \in G$). Man kann G mit der Transformationsgruppe $\{\varphi_g(\cdot) : g \in G\}$ „identifizieren“; wir haben einen Gruppenisomorphismus zwischen G und der Menge aller $\varphi_g(\cdot)$ (mit dem Hintereinanderschalten von Abbildungen).

Beispiel 1 (Möbiustransformationen)

Die spezielle lineare Gruppe $SL(2, \mathbb{C})$ ist die Gruppe der komplexen 2×2 -Matrizen mit Determinante = 1. Wir betrachten sie hier als eine abstrakte Gruppe G .

Vermöge der Möbius-Konstruktion wirkt sie auf $\bar{\mathbb{C}}$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \varphi_A(z) = \frac{az + b}{cz + d} \quad \text{für } z \in \bar{\mathbb{C}}.$$

Diese Gruppenwirkung ist nicht effektiv, weil $\varphi_A(\cdot)$ die Matrix A nur bis auf das Vorzeichen bestimmt. Sie ist unter Geometern beliebt, weil sie in gewissem Sinne anschaulicher ist als die effektive Wirkung auf der Menge aller komplexen 2-Spalten.

Zunächst wollen wir eingehend zwei Untergruppen von $SL(2, \mathbb{C})$ studieren, Untergruppen, die für die Geometrie und für die Quantenphysik gleichermassen wichtig sind:

- Die spezielle unitäre Gruppe $SU(2)$ beherrscht die Quantentheorie von Systemen mit zwei Zuständen. Der Physikstudent sollte dazu Feynman lesen: ‘Vorlesungen über Physik’, Teil III, Kapitel 11: Weitere Zweizustandssysteme. Ihre Wirkung auf der Riemann’sche Zahlenkugel entspricht der Gruppe aller Drehungen des dreidimensionalen Anschauungsraums, wie wir sehen werden.
- Die Gruppe $SU(1, 1)$ begegnet den Physikern unter anderem bei der Beschreibung des Tunneleffekts. Der Name $SU(1, 1)$ wird einsichtig, wenn man auch symplektische Matrizen studiert (siehe auch die Aufgabe am Ende). Für den Geometer ist zu sagen: Ihre Wirkung auf $\bar{\mathbb{C}}$, oder genauer auf dem Einheitskreis, entspricht der Gruppe aller Bewegungen der hyperbolischen Ebene.

Beispiel 2 (Sphärische Bewegungen)

Die spezielle unitäre Gruppe $SU(2)$ besteht aus den komplexen Matrizen der Gestalt

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad |a|^2 + |b|^2 = 1.$$

Es handelt sich wirklich um eine Gruppe: Produkte und Inverse haben dieselbe Gestalt.

Satz

Die Abbildungen $\varphi_U(\cdot)$ zu $U \in SU(2)$ respektieren die Diametralität; sie bilden diametrale Punkte in diametrale Punkte ab.

Jedes $\varphi_U(\cdot) \neq \text{Id}$ hat genau zwei Fixpunkte; und die sind zueinander diametral.

Beweis :

1. Seien z_1 und z_2 zueinander diametral; d.h. $z_1 \bar{z}_2 + 1 = 0$.

Wir müssen zeigen: $\varphi_U(z_1) \cdot \overline{\varphi_U(z_2)} + 1 = 0$. Das ist eine einfache Rechnung :

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

2. z ist Fixpunkt genau dann, wenn z eine gewisse quadratische Gleichung löst,

$$\begin{aligned} z = \frac{az + b}{-\bar{b}z + \bar{a}} & \Leftrightarrow -\bar{b}z^2 + \bar{a}z = az + b \Leftrightarrow (\bar{b}z)^2 + (a - \bar{a})\bar{b}z + |b|^2 = 0 \\ & \Leftrightarrow \left(\bar{b}z + \frac{a - \bar{a}}{2}\right)^2 = \left(\frac{a - \bar{a}}{2}\right)^2 - |b|^2. \end{aligned}$$

Im Falle $b = 0$ haben wir die Fixpunkte 0 und ∞ , für a und b reell haben wir die Fixpunkte $\pm i$, für a reell und b rein imaginär haben wir die Fixpunkte ± 1 . Diese drei Fälle werden wir unten genauer betrachten.

3. Wir zeigen allgemein, dass die beiden Lösungen z_1, z_2 der quadratischen Gleichung zueinander diametral sind:

$$\begin{aligned} \bar{b} \cdot z_{1,2} &= -\frac{a - \bar{a}}{2} \pm \alpha \quad \text{mit} \quad \alpha^2 = \left(\frac{a - \bar{a}}{2}\right)^2 - |b|^2; \\ (\bar{b}z_1) \cdot (b\bar{z}_2) &= \left(\frac{a - \bar{a}}{2}\right)^2 - \alpha^2 = -|b|^2; \quad \text{also} \quad z_1 \bar{z}_2 = -1. \end{aligned}$$

Satz

Die $\varphi_{\mathbf{U}}(\cdot)$ zu den $\mathbf{U} \in \mathrm{SU}(2)$ wirken als die Drehungen der Riemann'schen Zahlenkugel. (Man nennt sie auch die sphärischen Bewegungen von $\bar{\mathbb{C}}$.)

Wir wollen das hier nicht beweisen, wir betrachten einige

Beispiele sphärischer Bewegungen

1. Die Matrizen $\mathbf{U}_z(\varphi) = \begin{pmatrix} e^{\frac{i\varphi}{2}} & 0 \\ 0 & e^{-\frac{i\varphi}{2}} \end{pmatrix}$ liefern Drehungen um den Nordpol,
2. die Matrizen $\mathbf{U}_x(\varphi) = \begin{pmatrix} \cos \frac{\varphi}{2} & i \sin \frac{\varphi}{2} \\ i \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{pmatrix}$ liefern Drehungen um die x -Achse, und
3. die Matrizen $\mathbf{U}_y(\varphi) = \begin{pmatrix} \cos \frac{\varphi}{2} & \sin \frac{\varphi}{2} \\ -\sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{pmatrix}$ liefern Drehungen um die y -Achse.

Es handelt sich jeweils um Drehungen um den Winkel φ .

Man bemerke auch $\mathbf{U}(\varphi + 2\pi) = -\mathbf{U}(\varphi)$.

Unsere Matrizen lassen sich mit Hilfe der berühmten Pauli-Spinmatrizen schreiben; es gilt

$$\mathbf{U}_z(\varphi) = \exp(i\varphi/2 \cdot \sigma_z) \quad \mathbf{U}_x(\varphi) = \exp(i\varphi/2 \cdot \sigma_x) \quad \mathbf{U}_y(\varphi) = \exp(i\varphi/2 \cdot \sigma_y)$$

mit den Paulimatrizen

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Hinweise

1. Die Pauli-Matrizen sind hermitesche Matrizen mit der Spur $= 0$. Sie bilden in der Tat eine Basis des reellen Vektorraums aller spurlosen hermiteschen 2×2 -Matrizen. Wenn σ eine reelle Linearkombination der Paulimatrizen ist, dann ist die Schar $\{\mathbf{U}_t = \exp(it\sigma) : t \in \mathbb{R}\}$ eine sog. einparametrische Untergruppe der speziellen unitären Gruppe $\mathrm{SU}(2)$.
2. Die Gruppe der Drehungen des euklidischen \mathbb{R}^3 wird üblicherweise mit $\mathrm{SO}(3, \mathbb{R})$ identifiziert. $\mathrm{SO}(3, \mathbb{R})$ besteht aus den reellen 3×3 -Matrizen \mathbf{O} mit $\mathbf{O}^T \mathbf{O} = \mathbf{E}$ und $\det \mathbf{O} = 1$. $\mathrm{SO}(3, \mathbb{R})$ heißt die reelle spezielle orthogonale Gruppe. Man kann eine Isomorphie zeigen

$$\mathrm{SO}(3, \mathbb{R}) \cong \mathrm{SU}(2) / \{+1, -1\}.$$

Dies ist in der Tat nur noch einmal in anderen Worten die obige Behauptung: Die Gruppe der Möbiustransformationen zu den Matrizen aus $\mathrm{SU}(2)$ ist isomorph zur Gruppe der Drehungen $\mathrm{SO}(3, \mathbb{R})$. (Ohne Beweis).

3. In den Übungen werden wir eine Wirkung von $\mathrm{SU}(2)$ auf dem \mathbb{R}^3 kennenlernen. Der Raum \mathbb{R}^3 wird dort mit der Menge der reinen Quaternionen identifiziert, die durch Konjugation transformiert werden. Hier bleiben wir lieber bei der Wirkung auf $\bar{\mathbb{C}}$; wir verstehen also die Möbiustransformationen $\varphi_U(\cdot)$ als Drehungen im stereographischen Bild.

Beispiel 3

Die Gruppe $\mathrm{SU}(1, 1)$ besteht aus den Matrizen der Form

$$V = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{mit } |a|^2 - |b|^2 = 1.$$

Es handelt sich in der Tat um eine Gruppe; Produkte und Inverse haben dieselbe Gestalt.

Satz

Jede Möbiustransformation von der Gestalt

$$\psi(z) = \frac{az + b}{\bar{b}z + \bar{a}}, \quad |a|^2 - |b|^2 = 1$$

bildet das Innere des Einheitskreises auf sich ab.

Beweis Wir zeigen: Aus $1 - z\bar{z} > 0$ folgt $1 - \psi(z)\bar{\psi}(z) > 0$.

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

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

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

Bemerkung : Man kann leicht beweisen, dass jede Möbiustransformation, die das Innere des Einheitskreises auf sich abbildet, von der angegebenen Gestalt ist. Diese Möbiustransformationen heißen die hyperbolischen Bewegungen des Einheitskreises.

Beispiel 4 (Reelle Möbiustransformationen)

Die Möbiustransformationen zu den Matrizen mit reellen Einträgen bilden offenbar eine Transformationsgruppe. Diese Transformationen bilden die obere Halbebene auf sich ab. Wir bringen Sie (mit Hilfe der sog. Cayley-Transformation) in Beziehung zur Gruppe der hyperbolischen Bewegungen oder auch zur Gruppe derjenigen Möbiustransformationen, die das Innere des Einheitskreises auf sich abbilden.

1. Die folgende spezielle Möbiustransformation heißt die **Cayley-Transformation**

$$z \mapsto w = \chi(z) = \frac{z - i}{z - (-i)}.$$

Sie bildet die obere Halbebene auf das Innere des Einheitskreises ab; denn es sind genau die z in der oberen Halbebene, die von i einen kleineren Abstand haben als von $-i$.

2. Die Umkehrung $\chi^{-1}(\cdot)$ der Cayley-Transformation ist

$$w \mapsto z = \chi^{-1}(w) = -i \cdot \frac{w + 1}{w - 1} = \frac{w + 1}{iw - i}.$$

Wenn $\varphi_g(\cdot)$ das Innere des Einheitskreises auf sich abbildet, dann bildet $\chi^{-1} \circ \varphi_g \circ \chi(\cdot)$ die obere Halbebene auf sich ab.

3. Die Gruppe der Möbiustransformationen, die die obere Halbebene auf sich abbilden, kann andererseits beschreiben als die Menge der Möbiustransformationen mit reellen Koeffizienten

$$z \mapsto \psi(z) = \frac{az + b}{cz + d} \quad \text{mit } a, b, c, d \text{ reell } ad - bc = 1$$

Es gilt nämlich

$$\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} a + ic & b + id \\ b - id & a - ic \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \begin{pmatrix} 2(a + b) & 2(c - d) \\ -2(c + d) & 2(a - b) \end{pmatrix}$$

Fazit : Die Gruppe der hyperbolischen Bewegungen ist konjugiert zur Gruppe der reellen Möbiustransformationen im Sinne der

Definition Sei G eine abstrakte Gruppe. Zwei Untergruppen G_1 und G_2 heißen konjugiert (in G), wenn ein $h \in G$ existiert, sodass $G_2 = hG_1h^{-1}$.

Hinweis :

Die Gruppe aller reellen 2×2 -Matrizen mit Determinante $= 1$ hat eine interessante Untergruppe, nämlich die Gruppe aller ganzzahligen reellen 2×2 -Matrizen mit der Determinante $= 1$. Diese wird in der Gittertheorie studiert, aber auch in der Theorie der Modulfunktionen.

Transitivität und Fixgruppen

Definition

Man sagt von einer Gruppenwirkung $(G, \Omega, \{\varphi_g(\cdot) : g \in G\})$, dass sie auf Ω **transitiv** ist, wenn es zu jedem Paar ω_1, ω_2 mindestens ein $\varphi_h(\cdot)$ gibt mit $\varphi_h(\omega_1) = \omega_2$.

In diesem Fall sagt man auch, dass G transitiv auf Ω wirkt. Wenn es genau ein h gibt, dann sagt man, dass die Gruppe G **einfach transitiv** auf Ω wirkt.

Definition

Sei $(G, \Omega, \{\varphi_g(\cdot) : g \in G\})$ eine Gruppenwirkung; und sei $\tilde{\omega} \in \Omega$.

Die Fixgruppe von $\tilde{\omega}$ ist dann die Gruppe $G_{\tilde{\omega}} := \{g : \varphi(g\tilde{\omega}) = \tilde{\omega}\}$.

Bemerkung

Seien ω_1, ω_2 Punkte und $h \in G$, sodass $\varphi_h(\omega_1) = \omega_2$. Für die Fixgruppen G_{ω_1} und G_{ω_2} gilt dann

$$G_{\omega_2} = hG_{\omega_1}h^{-1}, \quad G_{\omega_1} = h^{-1}G_{\omega_2}h.$$

Für jedes $g \in G_{\omega_1}$ gilt nämlich $hgh^{-1} \in G_{\omega_2}$ wegen

$$\omega_1 = \varphi_{gh^{-1}}(\omega_2); \quad \varphi_{hgh^{-1}}(\omega_2) = \omega_2.$$

Beispiel 5 (Oktaedergruppe)

Sei G die Gruppe aller derjenigen Drehungen des Anschauungsraums, die den Einheitswürfel in sich überführen. (Wir haben oben gesehen, dass es sich um eine Gruppe mit 24 Elementen handelt; man nennt sie die Oktaedergruppe.) Wir diskutieren jetzt drei effektive Gruppenwirkungen der Oktaedergruppe.

- G wirkt transitiv auf der Menge F der Flächen. Die Fixgruppe jeder Fläche hat 4 Elemente.
- G wirkt transitiv auf der Menge E aller Ecken. Die Fixgruppe jeder Ecke hat 3 Elemente.
- G wirkt transitiv auf der Menge K aller Kanten. Die Fixgruppe jeder Kante hat 2 Elemente.

Bemerke : $24 = |F| \cdot 4 = |E| \cdot 3 = |K| \cdot 2$; denn es gibt 6 Flächen, 8 Ecken und 12 Kanten. Dieser Beobachtung liegt ein allgemeines Prinzip zugrunde. Um es zu beweisen müssten wir den Begriff der Nebenklassen (cosets) einer Untergruppe in einer Gruppe G einführen. Wir wollen das auf den Anhang verschieben.

Beispiel 6

Die Gruppe \mathfrak{M} aller Möbiustransformationen wirkt bekanntlich transitiv auf $\bar{\mathbb{C}}$.

Sei \mathfrak{M}_{∞} die Fixgruppe des Punkts ∞ . Die Elemente von \mathfrak{M}_{∞} sind die Abbildungen

$$z \longmapsto az + b \quad \text{mit } a \neq 0.$$

Diese Gruppe, aufgefasst als Transformationsgruppe auf \mathbb{C} , heißt die *affine Gruppe in einer komplexen Dimension*.

Bemerke :

1. Die zugrundeliegende abstrakte Gruppe \mathfrak{A} kann man natürlicherweise auch als Transformationsgruppe des zweidimensionalen Anschauungsraums auffassen; dies liefert uns die Gruppe der orientierungserhaltenden Ähnlichkeitstransformationen.
2. Wir können \mathfrak{A} auch als Transformationsgruppe auf der Menge Ω aller Kreise auffassen. Sie wirkt transitiv; denn zu jedem Paar von Kreisen gibt es eine Transformation in \mathfrak{A} , welche den einen in den anderen überführt. Die Fixgruppe eines Kreises ist die Menge der Drehungen um den Kreismittelpunkt.
3. Wir können die Gruppe \mathfrak{A} auch abstrakt beschreiben.

Jedes Element $\varphi(\cdot)$ von \mathfrak{A} entspricht einem Paar komplexer Zahlen (a, b) mit $a \neq 0$. Die Verknüpfungsregel lautet

$$(c, d) \circ (a, b) = (ca, d + cb) \quad \text{denn} \\ c(az + b) + d = caz + (d + cb) \quad \text{für alle } z \in \mathbb{C} .$$

Die Gruppeneins ist das Paar $(1, 0)$. Das inverse Element zu (a, b) ist $(a^{-1}, -a^{-1} \cdot b)$. Bemerke: Die Menge aller Paare $(1, b)$ ist eine kommutative Untergruppe, die Menge aller Translationen. Eine weitere kommutative Untergruppe ist die Menge aller Drehstreckungen um einen festen Punkt ω^* . Wenn dieser Punkt der Nullpunkt ist, dann besteht diese Gruppe aus der Menge aller Paare $(a, 0)$, $a \in \mathbb{C}$.

Beispiel 7 (Affinitäten des affinen Raums \mathbb{K}_{Sp}^J)

Wenn man in einem n -dimensionalen \mathbb{K} -affinen Raum ein affines Koordinatensystem wählt, dann entsprechen die Punkte („Ortsvektoren“) den Spalten der Länge n . (Hauptsächlich interessieren uns $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$, hier kann aber \mathbb{K} ein beliebiger Körper sein).

Sei $\Omega = \mathbb{K}_{\text{Sp}}^J$ die Menge aller J -Spalten mit Einträgen aus dem Körper \mathbb{K} . Wir konstruieren eine Bijektion von Ω für jedes Paar (A, b) , wo A eine nichtsinguläre $J \times J$ -Matrix ist und b eine J -Spalte.

$$x \longmapsto Ax + b \quad \text{für jede } J\text{-Spalte } x .$$

Die Gruppe dieser Bijektionen heißt die **affine Gruppe** auf \mathbb{K}_{Sp}^J . Die Elemente heißen die **Affinitäten**.

Das Paar (A, b) ist durch die Abbildung eindeutig bestimmt. Die Verknüpfungsregel lautet

$$(C, d) \circ (A, b) = (CA, d + Cb) .$$

Die Inverse zu (A, b) ist $(A^{-1}, -A^{-1}b)$. Wir werden den Begriff der affinen Abbildung in der nächsten Vorlesung ausführlich diskutieren.

Beispiel 8 (Cartesische Koordinatenwechsel im \mathbb{R}^3)

Die üblichen cartesischen Koordinaten im euklidischen Anschauungsraum mit dem Nullpunkt O bezeichnen wir mit $x(\cdot), y(\cdot), z(\cdot)$. Es handelt sich um lineare Funktionen mit

$$x^2(P) + y^2(P) + z^2(P) = \|\vec{OP}\|^2 \quad \text{für alle } P.$$

Definition :

Ein Tripel linearer Funktionen $x'(\cdot), y'(\cdot), z'(\cdot)$ im \mathbb{R}^3 heißt ein cartesisches Koordinatensystem, wenn gilt

$$x'^2(P) + y'^2(P) + z'^2(P) = \|\vec{OP}\|^2 \quad \text{für alle } P.$$

Die neuen cartesischen Koordinaten ergeben sich aus den alten mittels einer Matrix A

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}; \quad (x', y', z') \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (x, y, z) \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Die Normbedingung bedeutet für die Transformationsmatrix A , dass gilt $A^T \cdot A = E$. Wir wiederholen die schon früher erwähnte

Definition

- a) Eine reelle 3×3 -Matrix A heißt eine **orthogonale Matrix**, wenn $A^T \cdot A = E$. Die Gruppe aller reellen 3×3 -Matrizen mit $A^T \cdot A = E$ wird mit $O(3, \mathbb{R})$ bezeichnet („Orthogonale Gruppe“).
- b) Die reellen 3×3 -Matrizen A mit

$$A^T A = E, \quad \det A = 1$$

bilden die Gruppe $SO(3, \mathbb{R})$ („Special orthogonal group“).

Hinweis : Wir haben eben gesehen, dass man die $A \in O(3, \mathbb{R})$ als die Matrizen der cartesischen Koordinatentransformationen deuten kann. ($\det A = 1$ bedeutet, dass die Orientierung erhalten bleibt.) Man kann andererseits die Matrizen $A \in SO(3, \mathbb{R})$ auch mit den Drehungen des euklidischen Anschauungsraums „identifizieren“. Und damit haben wir wieder den mehrfach erwähnten Anschluss an die Gruppe $SU(2)/\{+1, -1\}$. Die klassischen Gruppen von $n \times n$ -Matrizen mit $n > 2$ werden wir an anderer Stelle zu studieren haben.

Beispiel 9

Eine Punkteabbildung eines Raums Ω liefert auch eine Abbildung der Punktepaare, der Punktetripel usw.; sie liefert auch Abbildungen von ‘Figuren’ (etwa im Sinne von Punktemengen). Eine lange Tradition in der Schulmathematik hat z. B. das Abbilden von Dreiecken (in der Anschauungsebene). Wir wollen hier noch einmal die Möbiustransformationen studieren, jetzt aber als Abbildungen von Punktetripeln.

Theorem

Die Gruppe \mathfrak{M} aller Möbiustransformationen wirkt einfach transitiv auf der Menge Ω aller Punktetripel. Wenn $z_1, z_2, z_3 \in \bar{\mathbb{C}}$ paarweise verschieden sind und w_1, w_2, w_3 paarweise verschieden sind, dann gibt es genau eine Möbiustransformation $\varphi(\cdot)$ mit

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

Zur Vorbereitung des Beweises studieren wir sog. Doppelverhältnisse

Definition („Doppelverhältnis“)

Für vier paarweise verschiedene komplexe Zahlen definiert man das Doppelverhältnis

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

Bemerkung :

Es kommt auf die Reihenfolge an. Wenn man die 24 möglichen Permutationen durchgeht, dann können aber nicht 24 verschiedene Werte auftreten, sondern höchstens 6, nämlich

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

Hinweis :

In der projektiven Geometrie ordnet man jedem Quadrupel von Punkten auf einer Geraden ein Doppelverhältnis zu. Die Formel ist ähnlich, die Sache ist aber doch nicht ganz dieselbe.

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.

Satz

Wenn die Punkte z_1, z_2, z_3, z_4 auf einem Kreis liegen, dann ist das Doppelverhältnis reell.

Beweis

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

Beweis des Theorems

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 von jedem weiteren 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 lineargebrochene Abbildung

$$w = \frac{az + b}{cz + d};$$

und das war zu zeigen.

Hinweis :

Die klassischen Gruppen, die wir diskutiert haben sind glatte 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 nämlich zunächst einmal vier komplexe, d.h. 8 reelle Zahlen, um eine komplexe 2×2 -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 affine Gruppe in einer komplexen Dimension hat (als reelle Mannigfaltigkeit) die Dimension vier. 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.

Aufgaben zu II.3

Aufgabe II.3.1 :

Es sei (G, e, \cdot) eine endliche Gruppe. Zu $\tilde{g} \in G$ sei $K_{\tilde{g}}$ die Konjugationsklasse:

$$K_{\tilde{g}} = \{g : g = h\tilde{g}h^{-1} \quad \text{mit} \quad h \in G\}.$$

Zeigen Sie, dass die Mächtigkeit von $K_{\tilde{g}}$ die Gruppenordnung teilt.

(Hinweis: Die Aufgabe betrifft einen Spezialfall der folgenden Aufgabe.)

Aufgabe II.3.2 :

Es sei (G, e, \cdot) eine endliche Gruppe. $\{\varphi_g : g \in G\}$ sei eine Gruppenwirkung auf irgendeiner Menge Ω . Für ein $\tilde{\omega} \in \Omega$ sei $G_{\tilde{\omega}}$ die Fixgruppe und $\Omega_{\tilde{\omega}}$ die ‘Bahn’ (der ‘Orbit’).

$$G_{\tilde{\omega}} = \{g : \varphi_g(\tilde{\omega}) = \tilde{\omega}\}; \quad \Omega_{\tilde{\omega}} = \{\omega : \omega = \varphi_g(\tilde{\omega}) \quad \text{mit} \quad g \in G\}.$$

Zeigen Sie für die Mächtigkeiten dieser Mengen $|G| = |\Omega_{\tilde{\omega}}| \cdot |G_{\tilde{\omega}}|$.

Aufgabe II.3.3 :

a) Machen Sie eine Liste aller Konjugationsklassen der symmetrischen Gruppe S_3 .

b) Wieviele verschiedene Konjugationsklassen gibt es in den symmetrischen Gruppen S_4 und S_5 ?

c) Man unterscheidet die Permutationen in S_n nach ihrem ‘Typ’:
Eine Permutation π heisst Permutation vom Typ (m_1, m_2, m_3, \dots) , wenn ihre disjunkte Zyklendarstellung m_k Zyklen der Länge k enthält. (m_1 ist die Anzahl der Punkte, die unter π festbleiben, m_2 ist die Anzahl der Transpositionen, Man bemerke $1 \cdot m_1 + 2 \cdot m_2 + 3 \cdot m_3 + \dots = n$.)

Zeigen Sie:

Die Anzahl der Permutationen vom Typ $(m_1, m_2, m_3, \dots, m_n)$ ist

$$N(m_1, m_2, \dots, m_n) = \frac{n!}{m_1! \cdot 2^{m_2} m_2! \cdot 3^{m_3} m_3! \dots} \quad \text{falls} \quad m_1 + 2m_2 + \dots = n,$$

und andernfalls gleich 0.

Aufgabe II.3.4 (Zum Begriff der symplektischen Matrix)

a) Sei $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$

Zeigen Sie : Für jede reelle 2×2 -Matrix A gilt

$$A^T \cdot J \cdot A = (\det A) \cdot J,$$

b) Sei R eine komplexe 2×2 -Matrix mit

$$R^* \cdot J \cdot R = J \quad \text{und} \quad \det R = 1 .$$

Zeigen Sie, dass R in Wirklichkeit eine reelle Matrix ist. Die entsprechenden Möbiustransformationen bilden die obere Halbebene auf sich ab.

c) Sei R eine reelle 2×2 -Matrix mit $\det R = 1$, d.h. $R \in \mathrm{SL}(2, \mathbb{R})$.
Zeigen Sie

$$R^T \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

d) Sei $R \in \mathrm{SL}(2, \mathbb{R})$ und C die spezielle Matrix

$$C = \begin{pmatrix} 1 & -i \\ 1 & +i \end{pmatrix}, \quad C^{-1} = \frac{1}{2i} \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$$

Zeigen Sie

$$A := CRC^{-1} \in \mathrm{SU}(1, 1)$$

d.h.

$$A = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 - |b|^2 = 1 .$$

Diesen Sachverhalt kann man folgendermaßen ausdrücken:

Die Gruppen $\mathrm{SL}(2, \mathbb{R})$ und $\mathrm{SU}(1, 1)$ sind zueinander konjugiert. Die dazugehörigen Möbiustransformationen sind vermöge der Cayley-Transformation zueinander konjugiert.

Hinweis :

Zu jeder geraden Dimension $2n$ gibt es symplektische Matrizen. Eine reelle $(2n) \times (2n)$ -Matrix A heißt eine symplektische Matrix, wenn gilt

$$A^T \cdot \begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix} \cdot A = \begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}$$

Die symplektischen Matrizen spielen eine wichtige Rolle in der Hamilton'schen Mechanik. (V. I. Arnold : *Mathematical Methods of Classical Mechanics*, Springer-Verlag 1988)

.

II.4 Affine Räume, affine Funktionen, affine Abbildungen

Wir beginnen mit einer mathematischen Definition des Typs von Raum, dessen einfachste Ausprägung man im Schulunterricht den Anschauungsraum nennt. Im Weiteren befassen wir uns dann mit wichtigen Typen von Funktionen auf solchen Räumen. Abbildungen werden nur kursorisch behandelt; die gründliche Behandlung der linearen Abbildungen ist Sache der Linearen Algebra (Abschnitt V).

Definition

Ein **affiner Raum** der Dimension n ist eine Punktmenge L , auf welcher ein n -dimensionaler Vektorraum V einfach transitiv wirkt. Zu jedem Punktepaar $P, Q \in L$ gibt es genau einen „Verschiebungsvektor“ $v \in V$, welcher P nach Q verschiebt.

Man notiert $P + v = Q$ oder auch $v = \overrightarrow{PQ}$. Der Raum aller Verschiebungsvektoren V heisst der **Tangentialraum** des affinen Raums L .

Der Anfänger ist gut beraten, wenn er hier zunächst an den zweidimensionalen oder dreidimensionalen Anschauungsraum denkt. Weiter sollte er dann an reellaffine Räume denken. Für komplexaffine Räume gibt es nämlich Spezialitäten, die uns in Abschnitt IV beschäftigen werden (Wir denken an die hermiteschen Formen als Gegenstücke zu den quadratischen Formen). Die Überlegungen gelten aber sehr wohl auch für allgemeine endlichdimensionale \mathbb{K} -affine Räume. (Gelegentlich benötigen wir für unsere Konstruktionen, dass $\frac{1}{2}$ ein wohldefinierter Skalar ist, dass also gilt $1 + 1 \neq 0$ in \mathbb{K} ; wir werden darauf hinweisen.)

Der Kalkül der Ortsvektoren

Die Punkte des affinen Raums L bezeichnen wir mit P, Q, R, \dots ; die Verschiebungsvektoren bezeichnen wir mit v, w, \dots . Wenn Q aus P durch Verschiebung mit v hervorgeht, schreiben wir auch $v = \overrightarrow{PQ} = Q - P$.

Sei $P \in L$, $v \in V$, $v \neq 0$. Die Punktmenge

$$\{R : R = P + \lambda v, \lambda \in \mathbb{K}\}$$

nennt man die **Gerade** durch P mit der Richtung von v . Bemerke: Durch jedes Paar von Punkten $P \neq Q$ gibt es genau eine Gerade, nämlich

$$\{R : R = (1 - \lambda)P + \lambda Q, \quad \lambda \in \mathbb{K}\}.$$

Die Geometer nennen die Zahl $\frac{\lambda}{\lambda - 1} = V(P, Q, R)$ das Teilungsverhältnis, in welchem R die Strecke PQ teilt.

Die Notation $R = (1 - \lambda)P + \lambda Q$ ist praktischer als etwa die Notationen $R = P + \lambda \overrightarrow{PQ} = P + \lambda(Q - P) = Q + (1 - \lambda)\overrightarrow{QP}$.

Diese Notation soll jetzt verallgemeinert werden. Wir definieren $P = \sum \lambda_k \cdot P_k$, wenn $\sum \lambda_k = 1$.

Satz

Seien P_0, P_1, \dots, P_m Punkte in einem affinen Raum L , und $\lambda_0, \lambda_1, \dots, \lambda_m$ Skalare mit $\sum \lambda_k = 1$. Dann ist $\sum \lambda_k P_k$ ein wohlbestimmter Punkt $\in L$.

Beweis

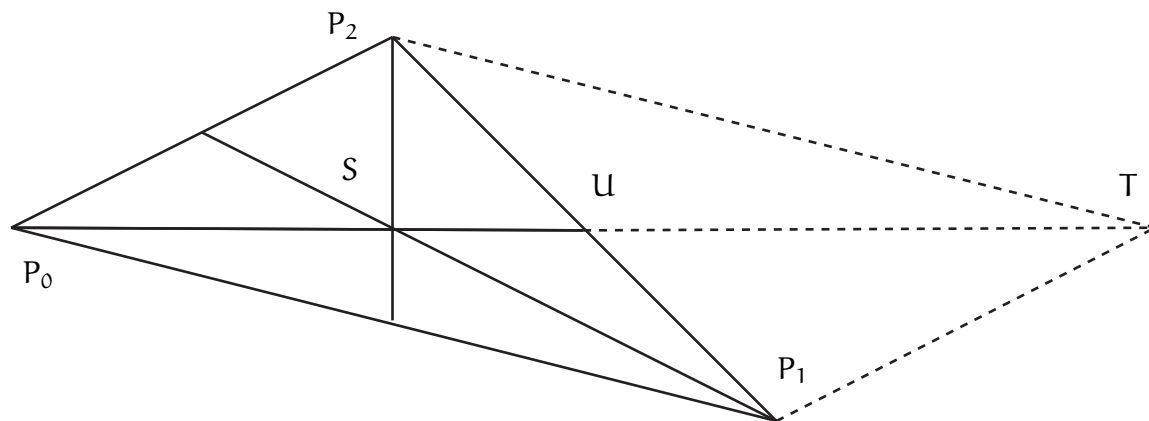
Sei P_* ein beliebig gewählter Punkt und $v_k = \vec{P_* P_k} = P_k - P_*$ für $k = 0, 1, \dots, m$. Den Punkt $P = P_* + \sum \lambda_k v_k$ notieren wir auch

$$P = P_* + \sum \lambda_k \cdot \vec{P_* P_k} = P_* + \sum \lambda_k (P_k - P_*) .$$

Wegen $\sum \lambda_k = 1$ ist das Ergebnis unabhängig von P_* .

Sei nämlich $P_* = Q_* + v_*$, und $w_k = \vec{Q_* P_k} = v_* + v_k$. Dann haben wir

$$Q_* + \sum \lambda_k w_k = Q_* + \sum \lambda_k (v_* + v_k) = P_* + \sum \lambda_k v_k = P .$$

Beispiel :

$$S = \frac{1}{3}P_0 + \frac{1}{3}P_1 + \frac{1}{3}P_2 = \frac{1}{3}P_0 + \frac{2}{3}U = \frac{1}{3}P_0 + \frac{2}{3} \left(\frac{1}{2}P_1 + \frac{1}{2}P_2 \right) .$$

Der Schwerpunkt von gleichen Massen in den Ecken eines Dreiecks ist der Schnittpunkt der „Schwerlinien“. Der Schwerpunkt S teilt jede Schwerlinie im Verhältnis $\frac{1}{3} : \frac{2}{3}$.

Wenn man positive Massen λ_k in die Ecken legt (wir nehmen an, dass sie sich zu 1 aufaddieren, $\sum \lambda_k = 1$), dann ist ihr Schwerpunkt ein Punkt $R = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2$ im Innern des Dreiecks. Man nennt $(\lambda_0, \lambda_1, \lambda_2)$ die Schwerpunktskoordinaten von R .

Mit reellen Tripeln $(\lambda_0, \lambda_1, \lambda_2)$ mit $\sum \lambda_k = 1$ kann man auch all die übrigen Punkte R in der Ebene durch P_0, P_1, P_2 gewinnen.

Beispielsweise hat der Punkt T im Bild die ‘Schwerpunktskoordinaten’ $(-1, +1, +1)$

$$T = P_2 + \vec{P_0 P_1} = P_1 + \vec{P_0 P_2} = -P_0 + P_1 + P_2 .$$

Der Schwerpunkt S liegt auf der Verbindungsstrecke von P_0 nach T , $S = \frac{2}{3}P_0 + \frac{1}{3}T$.

Die Punkte eines affinen Raums nennt man auch Ortsvektoren. Wenn P_0, P_1, \dots, P_m Ortsvektoren sind, dann ist $\sum \lambda_k \cdot P_k$ im Falle $\sum \lambda_k = 1$ ein Ortsvektor, und im Falle $\sum \lambda_k = 0$ ein Verschiebungsvektor; wenn die Summe der λ weder 0 noch 1 ist, bleibt der Ausdruck undefiniert.

Definition

Eine Teilmenge M' eines affinen Raums (L, V) heißt ein **affiner Teilraum**, wenn es einen Teilvektorraum $V' \subseteq V$ gibt, sodass (für ein $P_0 \in M'$)

$$M' = \{R : R = P_0 + v' \quad \text{mit } v' \in V'\} .$$

V' heißt der **Tangentia Raum** von M' . Die Dimension von V' heißt auch die Dimension von M' .

Bemerkung : P_0 kann offenbar beliebig in M' gewählt werden.

Sei v_1, v_2, \dots, v_m eine Basis von V' , und sei $P_k = P_0 + v_k$. Dann gilt für jeden Punkt $P \in M'$

$$P = P_0 + \sum_1^m \lambda_k v_k = \lambda_0 P_0 + \lambda_1 P_1 + \dots + \lambda_m P_m \quad \text{mit wohlbestimmten } \lambda_k \quad \left(\sum_0^m \lambda_k = 1 \right) .$$

$(\lambda_0, \lambda_1, \dots, \lambda_m)$ nennt man die **Schwerpunktskoordinaten** von P bzgl. $\{P_0, P_1, \dots, P_m\}$.

Satz Sei \mathbb{K} ein Körper mit $1 + 1 \neq 0$.

Eine Teilmenge M' des \mathbb{K} -affinen Raums (L, V) ist genau dann ein affiner Teilraum, wenn gilt

$$\forall P, Q \in M' \quad \forall \lambda \in \mathbb{K} \quad (1 - \lambda)P + \lambda Q \in M' .$$

(„Mit P und Q gehört auch die gesamte Verbindungsgerade zu M' “.)

Beweis

Die Bedingung ist offenbar notwendig; wenn wir einen Teilvektorraum haben, welcher M' liefert, dann ist die Aussage richtig. Um zu zeigen, dass sie hinreichend ist, wählen wir $P_0 \in M'$ und zeigen, dass $V' = \{P_0 P \text{ mit } P \in M'\}$ ein Vektorraum ist, d. h.

$$(i) \quad v \in V' , \quad \lambda \in \mathbb{K} \Rightarrow \lambda v \in V'$$

$$(ii) \quad v, w \in V' \Rightarrow \frac{1}{2}(v + w) \in V' .$$

$$\text{zu (i) : } (1 - \lambda)P_0 + \lambda \cdot (P_0 + v) \in M' \Rightarrow P_0 + \lambda v \in M' \Rightarrow \lambda v \in V'$$

$$\text{zu (ii) : } \frac{1}{2}(P_0 + v) + \frac{1}{2}(P_0 + w) \in M' \Rightarrow P_0 + \frac{1}{2}(v + w) \in M' \Rightarrow \frac{1}{2}(v + w) \in V' .$$

Wir haben also in der Tat

$$v, w \in V' , \alpha, \beta \in \mathbb{K} \Rightarrow \alpha v + \beta w \in V' .$$

„Räume sind Strukturierte Mengen“

Nach allgemeinem Sprachgebrauch der Mathematiker ist ein Raum eine strukturierte Menge. Topologische Räume, messbare Räume und viele andere Typen von Räumen erhalten ihre Struktur durch ein ausgezeichnetes Teilmengensystem. Ein affiner Raum bekommt seine Struktur durch die algebraische Struktur des Vektorraums seiner Verschiebungen. Bei vielen strukturierten Mengen Ω erweisen sich gewisse Klassen von Funktionen f auf Ω als noch wichtiger als die Punkte des Raums.

Wir müssen uns genauer mit den affinen Funktionen auf einem affinen Raum befassen, weil die affinen Funktionen in der Tat den fundamentalsten Typ von Funktionen auf einem affinen Raum darstellen. Wir werden nach und nach sehen, wie die affinen Funktionen als Bausteine dienen für weitere Typen von Funktionen (wie z. B. die quadratischen Funktionen, die polynomialen Funktionen oder auch solche Funktionen wie die Amplitudenfunktionen zu ebenen Wellen).

Anschluss an die Schulmathematik

In der Schule definiert man: Eine reellwertige Funktion $f(\cdot)$ auf dem zweidimensionalen Anschauungsraum heißt eine affine (oder „lineare“) Funktion, wenn ihr ‘Funktionsgraph’ eine Ebene ist. Die Niveaumengen $\{P : f(P) = \text{const}\}$ sind zueinander parallele Geraden. Eine affine Funktion $f(\cdot)$ auf dem dreidimensionalen Anschauungsraum veranschaulicht man von vornherein durch das System der Niveaumengen; hier hat man parallele Ebenen. (Der ‘Funktionsgraph’ wäre eine Hyperebene im vierdimensionalen ‘Anschauungsraum’.)

Wir wollen uns hier mit den affinen Funktionen auf einem n -dimensionalen affinen Raum beschäftigen. Die Niveaumengen sind zueinander parallele Hyperebenen. Hyperebenen sind $(n - 1)$ -dimensionale affine Teilräume, ebenso wie Gerade eindimensionale und Ebenen zweidimensionale affine Teilräume sind.

Affine und quadratische Funktionen im Sinne der Schulmathematik

Sei L der affine Raum der Zahlenpaare $\begin{pmatrix} x \\ y \end{pmatrix}$. Die affinen Funktionen $f(\cdot)$ auf L sind die Funktionen der Gestalt

$$f(\cdot) = a_0 + a_1x + a_2y = a_0 + (a_1, a_2) \begin{pmatrix} x \\ y \end{pmatrix}.$$

Um eine affine Funktion festzulegen, braucht man eine Konstante a_0 und eine Zeile (a_1, a_2) .

Die quadratischen Funktionen sind die Funktionen der Gestalt

$$\begin{aligned} q(\cdot) &= a_0 + a_1x + a_2y + \frac{1}{2}a_{11}x^2 + \frac{1}{2}(a_{12} + a_{21})xy + \frac{1}{2}a_{22}y^2 \\ &= a_0 + (a_1, a_2) \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x, y) \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Um eine quadratische Funktion festzulegen, braucht man eine Zahl a_0 , eine Zeile (a_1, a_2) und eine symmetrische Matrix.

Bemerkung Die Menge aller affinen Funktionen auf dem zweidimensionalen Anschauungsraum ist ein dreidimensionaler Vektorraum (bzgl. der punktweisen Linearkombination); die Menge der quadratischen Funktionen ist ein sechsdimensionaler Vektorraum. Wir werden allgemein sehen: Ist L ein n -dimensionaler \mathbb{K} -affiner Raum, dann ist die Menge $\text{Aff}(L)$ aller affinen Funktionen ein $(1+n)$ -dimensionaler Vektorraum; die Menge $\text{Quad}(L)$ aller quadratischen Funktionen ist ein \mathbb{K} -Vektorraum der Dimension $1 + n + \frac{1}{2}n(n+1)$.

Weitere Funktionen auf dem Anschauungsraum : Amplituden ebener Wellen

Das Raum-Zeitkontinuum der speziellen Relativitätstheorie ist ein vierdimensionaler affiner Raum. Es sei durch die affinen Funktionen t, x^1, x^2, x^3 koordinatisiert.

(Der Physiker nimmt nicht irgendein linear unabhängiges Quadrupel von Koordinatenfunktionen, sondern ein solches, welches gut zum Lichtkegel passt. Die Bezeichnungen weisen darauf hin, dass t als eine Zeitkoordinate verstanden werden soll und die x^j als Raumkoordinaten. Davon aber später.)

Die affinen Funktionen sind die Grundlage für den Begriff der ebenen Welle. Das ist folgendermassen zu verstehen: Jede affine Funktion hat die Gestalt

$$\text{const} + \omega t - k_1 x^1 - k_2 x^2 - k_3 x^3 .$$

Die komplexe Amplitudenfunktion der dazugehörigen ebenen Welle hat die Gestalt

$$\begin{aligned} A(t, \vec{x}) &= A \cdot \exp \left(-i[\omega t - k_1 x^1 - k_2 x^2 - k_3 x^3] \right) \\ \text{oder auch} \quad &= A \cdot \exp \left(-\frac{i}{\hbar}[Et - p_1 x^1 - p_2 x^2 - p_3 x^3] \right) \end{aligned}$$

$$E = \hbar\omega, \quad p_1 = \hbar k_1, \quad p_2 = \hbar k_2, \quad p_3 = \hbar k_3$$

ω heisst die Kreisfrequenz, die k_j heissen die Wellenzahlen, A ist die komplexe Amplitude. E steht für Energie, die p_j deuten auf einen Impuls hin.

Wir bemerken

- In einem festen Punkt im Raum \vec{x} ist die Amplitude $A(\cdot, \vec{x})$ eine reine Sinusschwingung mit der maximalen Amplitude $|A|$.
- Betrachten wir zu einem festen Zeitpunkt t_0 die Menge der Punkte im Raum, in welchen der Realteil der Amplitude maximal ist.

$$M_{t_0} = \left\{ \vec{x} : k\vec{x} = \omega t_0 + \text{const} + 2\pi m, \quad m \in \mathbb{Z} \right\} .$$

Die Wellengipfel zur Zeit t_0 bilden eine Schar äquidistanter Ebenen.

- Wenn t_0 wächst, dann „bewegen“ sich die Ebenen

$$M_t^{(m)} = \left\{ \vec{x} : k\vec{x} = \omega t + \text{const} + 2\pi m \right\}$$

mit einer sogen. Phasengeschwindigkeit \vec{v} . Es gilt

$$\mathbf{k} \cdot (\vec{x}_t - \vec{v} \cdot t) = \text{const} \quad \text{für} \quad \vec{x}_t \in M_t^{(m)} .$$

Der „Vektor“ \vec{v} ist allerdings nicht eindeutig bestimmt, weil man ja nicht sagen kann, welcher Punkt auf einem Wellengipfel zur Zeit t welchem Punkt auf dem Wellengipfel zur Zeit $t + \Delta t$ entspricht. Es gibt keine materielle Bewegung und daher keine wirkliche Geschwindigkeit; nur $\omega = \mathbf{k} \cdot \vec{v}$ ist eindeutig bestimmt.

- d) Wenn zwei gegeneinander laufende ebene Wellen mit derselben maximalen Amplitude $|A|$ überlagert werden, dann erhält man eine sog. stehende Welle

$$\begin{aligned} & |A|e^{i\alpha} \cdot \exp(-i[\omega t - \mathbf{k}\vec{x}]) + |A|e^{i\beta} \cdot \exp(-i[\omega t + \mathbf{k}\vec{x}]) \\ &= |A| \cdot e^{-i\omega(t-t_0)} \cdot 2 \cos(\mathbf{k}\vec{x} + \delta) \\ \text{mit } i\omega t_0 &= \frac{\alpha + \beta}{2} \quad \text{und} \quad \delta = \frac{\alpha - \beta}{2} . \end{aligned}$$

In allen Positionen haben wir reine Sinusschwingungen, mit derselben Phase. Die Amplitude hängt vom Ort ab. Die Knoten liegen auf den parallelen Ebenen

$$\{\vec{x} : \mathbf{k}\vec{x} + \delta = m\pi \text{ mit } m \in \mathbb{Z}\} .$$

- e) In einem Medium, in welchem ebene Wellen mit allen möglichen Wellenzahlen und den dazugehörigen Kreisfrequenzen betrachtet werden, heißt $\omega = \Omega(\mathbf{k})$ die Dispersionsrelation des Mediums.

Hinweis : Im Endeffekt interessieren den Physiker nicht die einzelnen ebenen Wellen, sondern die Überlagerungen solcher Wellen. Schreiben wir kurz $f_{\mathbf{k}}(\cdot)$ für die Funktion

$$f_{\mathbf{k}}(t, \vec{x}) = \exp(-i[\Omega(\mathbf{k})t - \mathbf{k}\vec{x}]) ,$$

wobei $\Omega(\mathbf{k})$ eine gewisse Funktion ist („Dispersionsrelation“). Interessant sind insbesondere diejenigen Überlagerungen ebener Wellen, die ein „Wellenpaket“ beschreiben

$$g(t, \vec{x}) = \int f_{\mathbf{k}}(t, \vec{x}) \cdot Q(\mathbf{k}) d\mathbf{k} .$$

Dabei ist $Q(\mathbf{k})d\mathbf{k}$ eine „geeignete“ komplexe Gewichtung auf dem dreidimensionalen Raum der Wellenzahlen. Für jede feste Zeit kann man $g(t, \cdot)$ als ein sog. Fourier-Integral verstehen

$$g(t, \vec{x}) = \int \exp(i\mathbf{k}\vec{x}) \exp(-i\Omega(\mathbf{k})t) Q(\mathbf{k}) d\mathbf{k} .$$

Mit solchen Fourier-Integralen hat der Physiker bei vielen Gelegenheiten zu arbeiten. Die Amplitudenfunktionen ebener Wellen erscheinen als Bausteine.

Koordinatenfreier Zugang

In der Schule und in den traditionellen Lehrbüchern für Anwender geht man davon aus, dass ein affiner Raum ein bestimmtes Koordinatensystem ‘mitbringt’. In unserer Definition eines \mathbb{K} -affinen Raums ist davon jedoch ebenso wenig die Rede wie in der Definition eines Vektorraums von einer ausgezeichneten Basis die Rede ist. Der koordinatenfreie Zugang bietet die Möglichkeit, in jedem Fall ein zum Problem passendes Koordinatensystem zu wählen (ohne sich mit Koordinatenwechsel abzumühen). Die Koordinatensysteme werden den Gegebenheiten angepasst; wenn man z. B. im Anschauungsraum auch den euklidischen Abstandsbegriff zur Geltung bringen will, dann wird man sog. orthonormale Koordinatensysteme bevorzugen; im relativistischen Raum-Zeit-Kontinuum wird man Koordinatensysteme bevorzugen, in welchen der Lichtkegel eine übersichtliche Darstellung besitzt. Bevor wir uns speziellen Typen von affinen Koordinaten (auf speziellen Räumen) zuwenden, wollen wir zuerst die Theorie der affinen Koordinatensysteme auf einem endlichdimensionalen \mathbb{K} -affinen Raum in voller Allgemeinheit entwickeln. Zu diesem Zweck erinnern wir uns an die (in I.5 entwickelten) elementaren Begriffe der Vektorraumtheorie.

Linearformen, Dualraum, duale Basen

Definition (Linearform)

Sei V ein n -dimensionaler \mathbb{K} -Vektorraum. Eine \mathbb{K} -wertige Funktion $\ell(\cdot)$ auf V heißt eine **Linearform** (oder eine \mathbb{K} -lineare Funktion), wenn gilt

- (i) $\ell(\alpha v) = \alpha \cdot \ell(v)$ für alle $\alpha \in \mathbb{K}$, $v \in V$,
- (ii) $\ell(v + w) = \ell(v) + \ell(w)$ für alle $v, w \in V$.

Man kann die beiden Bedingungen offenbar auch durch eine einzige Bedingung ersetzen.

- (iii) $\ell(\alpha v + \beta w) = \alpha \cdot \ell(v) + \beta \cdot \ell(w)$ für alle $\alpha, \beta \in \mathbb{K}$, $v, w \in V$.

Definition (Dualraum)

Den \mathbb{K} -Vektorraum aller \mathbb{K} -Linearformen auf V nennt man den Dualraum von V ; man bezeichnet ihn mit V^* .

Notation

Statt $\ell(v)$ schreibt man auch $\langle \ell, v \rangle$. Die Linearform $\ell(\cdot)$ notiert man auch $\langle \ell, \cdot \rangle$.

Beispiel: Spalten und Zeilen

Sei V der \mathbb{K} -Vektorraum aller J -Spalten mit Einträgen aus \mathbb{K} . $V = \mathbb{K}_{\text{Sp}}^J$. Jede J -Zeile ξ liefert vermöge des Matrizenprodukts eine Linearform

$$\langle \xi, v \rangle = \xi \cdot v \quad (\text{Zeile} \times \text{Spalte} = \text{Skalar}) .$$

Jede Linearform $\ell(\cdot)$ ist von dieser Gestalt.

Die Einträge x^j von v und die Einträge ξ_j von ξ sind die Koeffizienten, wenn man v und ξ mit Hilfe der ‘Einheitsspalten’ v_j bzw. mit Hilfe der ‘Einheitszeilen’ darstellt.

$$v = \sum v_j \cdot x^j \quad \text{und mit} \quad \xi_j = \langle \ell, v_j \rangle$$

$$\langle \ell, v \rangle = \langle \ell, \sum v_j \cdot x^j \rangle = \sum \xi_j \cdot x^j.$$

(Die Hoch- und Tiefstellung der Indizes entspricht der Tradition. Da die Indizes j bei uns keine Zahlen sind, sondern Elemente der abstrakten Indexmenge J , ist eine Verwechslung mit Potenzen eines Skalars nicht zu befürchten.)

Satz (Duale Basen)

Der Dualraum V^* des Vektorraums V hat dieselbe Dimension wie V .

Zu jeder Basis $\{v_j : j \in J\}$ von V existiert eine Basis $\{\ell^j(\cdot) : j \in J\}$ von V^* mit der Eigenschaft

$$\langle \ell^k, v_j \rangle = \begin{cases} 1 & \text{falls } k = j \\ 0 & \text{falls } k \neq j \end{cases}$$

Beweis

Nach dem ersten Hauptsatz der Theorie endlich-dimensionaler \mathbb{K} -Vektorräume gibt es Basen.

1. Sei $\{v_j : j \in J\}$ irgendeine Basis von V . Jedes $v \in V$ besitzt eine Darstellung

$$v = \sum v_j \cdot \alpha^j \quad \text{mit wohlbestimmten } \alpha^j.$$

Der j -te Koeffizient α^j hängt in linearer Weise von v ab. $\alpha^j = \ell^j(v)$.

2. Wir zeigen, dass $\{\ell^j(\cdot) : j \in J\}$ eine Basis von V^* ist.

Sei $\ell(\cdot)$ irgendeine Linearform, $\beta_j = \ell(v_j)$. Es gilt dann

$$\ell(v) = \ell\left(\sum \alpha^j v_j\right) = \sum \alpha^j \cdot \ell(v_j) = \sum \alpha^j \cdot \beta_j = \sum \ell^j(v) \cdot \beta_j,$$

$$\text{also} \quad \ell(\cdot) = \sum \beta_j \cdot \ell^j(\cdot)$$

Die $\ell^j(\cdot)$ sind linear unabhängig. Die Koeffizienten in der Darstellung $\ell(\cdot) = \sum \beta_j \ell^j(\cdot)$ sind für jedes $\ell(\cdot)$ eindeutig bestimmt.

$\{\ell^j(\cdot) : j \in J\}$ ist die duale Basis zu $\{v_j : j \in J\}$.

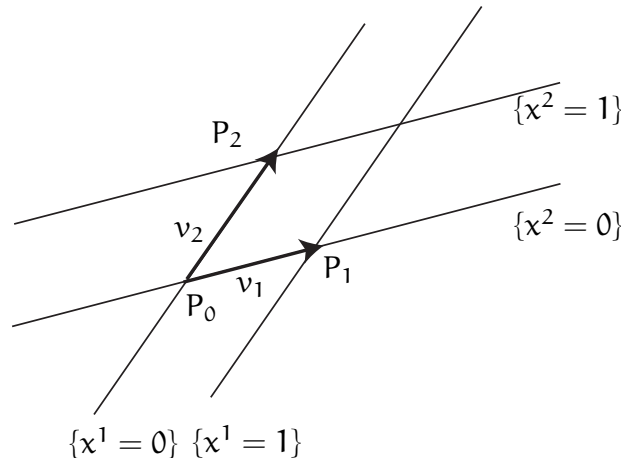
Satz

Die Linearformen auf dem Dualraum entsprechen den Elementen von V ; $V^{**} \longleftrightarrow V$. Die duale Basis zur dualen Basis ist die ursprüngliche Basis.

Diese Tatsache wird offensichtlich, wenn wir den Vektor v mit der ‘Funktion’ $m(\cdot) = \langle \cdot, v \rangle$ identifizieren; — ebenso wie wir die Linearform ℓ mit der ‘Funktion’ $\langle \ell, \cdot \rangle$ identifiziert haben.

Bildliche Darstellung

Im zweidimensionalen reellen Anschauungsraum veranschaulichen wir affine Funktionen x^1, x^2 und die dazugehörigen Linearformen ℓ^1, ℓ^2 durch Paare von Niveaulinien.



Die zu $\{\ell^1(\cdot), \ell^2(\cdot)\}$ duale Basis ist das Paar von Vektoren $v_1 = \vec{P_0 P_1}, v_2 = \vec{P_0 P_2}$.

Affine Funktionen**Definition**

Eine \mathbb{K} -wertige Funktion $f(\cdot)$ auf einem \mathbb{K} -affinen Raum (L, V) heißt eine **\mathbb{K} -affine Funktion**, wenn es eine Linearform $\ell(\cdot)$ auf V gibt, sodass gilt

$$f(P + v) - f(P) = \ell(v) \quad \text{für alle } P \in L.$$

Satz

Sei \mathbb{K} ein Körper mit $1 + 1 \neq 0$. (L, V) sei ein \mathbb{K} -affiner Raum. Eine \mathbb{K} -wertige Funktion $f(\cdot)$ ist genau dann eine \mathbb{K} -affine Funktion, wenn gilt

$$\forall P, Q \in L \quad \forall \lambda \in \mathbb{K} \quad f((1 - \lambda)P + \lambda Q) = (1 - \lambda)f(P) + \lambda \cdot f(Q).$$

Beweis :

Die Bedingung ist offenbar notwendig. Um zu zeigen, dass sie hinreichend ist, wählen wir einen Punkt P_0 und beweisen, dass $\ell(v) = f(P_0 + v) - f(P_0)$ eine Linearform $\ell(\cdot)$ ist, indem wir zeigen

$$(i) \quad \ell(\lambda v) = \lambda \cdot \ell(v) \quad \text{für alle } \lambda \in \mathbb{K}, v \in V$$

$$(ii) \quad \ell\left(\frac{1}{2}(v + w)\right) = \frac{1}{2}\ell(v) + \frac{1}{2}\ell(w) \quad \text{für alle } v, w \in V.$$

$$\begin{aligned} \text{zu (i)} : f(P_0) + \ell(\lambda v) &= f(P_0 + \lambda v) = f((1 - \lambda)P_0 + \lambda(P_0 + v)) \\ &= (1 - \lambda)f(P_0) + \lambda \cdot f(P_0 + v) = f(P_0) + \lambda \cdot \ell(v). \end{aligned}$$

$$\begin{aligned} \text{zu (ii)} : f(P_0) + \ell\left(\frac{1}{2}(v + w)\right) &= f\left(P_0 + \frac{1}{2}(v + w)\right) = \\ &= f\left(\frac{1}{2}(P_0 + v) + \frac{1}{2}(P_0 + w)\right) = \frac{1}{2}f(P_0 + v) + \frac{1}{2}f(P_0 + w) = f(P_0) + \frac{1}{2}(\ell(v) + \ell(w)). \end{aligned}$$

Corollar

Eine \mathbb{K} -wertige Funktion ist genau dann \mathbb{K} -affin, wenn ihre Einschränkung auf jede Gerade \mathbb{K} -affin ist. (Hier ist $1 + 1 \neq 0$ vorauszusetzen.)

Bemerke : In der Schule lernt man, dass eine affine Funktion dadurch gekennzeichnet ist, dass die Ableitung überall dieselbe ist, (woraus sich ergibt, dass der Zuwachs beim Fortschreiten um den Verschiebungsvektor v überall derselbe ist).

Eine etwas kompliziertere aber doch analoge Aussage gilt auch im Fall eines beliebigen Körpers. Zunächst stellen wir fest: Wenn $g(\cdot)$ eine \mathbb{K} -wertige Funktion ist, sodass der Zuwachs $g(P + v) - g(P)$ unabhängig von P ist für jedes feste $v \in L$, dann ist dieser Zuwachs $m(\cdot) = g(P + \cdot) - g(P)$ eine additive Funktion. Es gilt

$$\begin{aligned} m(v + w) &= m(v) + m(w) \quad \text{für alle } v, w; \quad \text{denn} \\ m(v + w) - m(w) &= g(P + v + w) - g(P) - (g(P + w) - g(P)) = m(v). \end{aligned}$$

Im allgemeinen Fall ist aber zu beachten, dass eine additive Funktion nicht notwendigerweise \mathbb{K} -linear ist. ($m(\lambda v) = \lambda m(v)$?)

Der Körper \mathbb{Q} ist ein Sonderfall. Es gilt: Jede additive Funktion auf einem \mathbb{Q} -affinen Raum ist \mathbb{Q} -linear. Der Beweis liegt auf der Hand.

Für den Körper $\mathbb{K} = \mathbb{R}$ braucht man etwas Zusätzliches, um für eine Funktion mit positionsunabhängigen Zuwächsen zu garantieren, dass sie eine affine Funktion ist: Jede additive Funktion auf einem reellen Vektorraum, welche auf jeder Geraden stetig ist, ist eine \mathbb{R} -lineare Funktion. Unstetige additive Funktionen auf der reellen Achse kann man sich zwar kaum vorstellen; die Analytiker bestehen aber darauf, dass es solche Funktionen gibt.

Affine Koordinaten

Sei L ein n -dimensionaler \mathbb{K} -affiner Raum mit dem Tangentialraum V .

Ein n -Tupel von affinen Funktionen $\{a^j(\cdot) : j \in J\}$ heißt **affines Koordinatensystem** auf L , wenn es zusammen mit der konstanten Funktion $\equiv 1$ eine Basis des Raums aller affinen Funktionen bilden, d. h. wenn die zugehörigen Linearformen $\{\ell^j(\cdot) : j \in J\}$ eine Basis von V^* bilden.

Für einen Punkt P heissen die ‘Funktionswerte’ $a^j(P)$ die affinen Koordinaten von P bezgl. des betreffenden Koordinatensystems. Zu jedem J -Tupel von Skalaren $\{x^j : j \in J\}$ gibt es genau einen Punkt mit den Koordinaten x^j .

Lineare und affine Abbildungen**Definition**

- a) V und W seien \mathbb{K} -Vektorräume. Eine Abbildung $\varphi : V \rightarrow W$ heißt eine **\mathbb{K} -lineare Abbildung**, wenn gilt

$$\forall \alpha, \beta \in \mathbb{K}, v_1, v_2 \in V \quad \varphi(\alpha v_1 + \beta v_2) = \alpha \cdot \varphi(v_1) + \beta \cdot \varphi(v_2) .$$

- b) (L, V) und (M, W) seien \mathbb{K} -affine Räume. Eine Abbildung $\Phi : L \rightarrow M$ heißt eine **\mathbb{K} -affine Abbildung**, wenn eine lineare Abbildung $\varphi : V \rightarrow W$ existiert, sodass

$$\Phi(P_0 + v) = \Phi(P_0) + \varphi(v) \quad .$$

(Der Punkt P_0 kann offenbar beliebig gewählt werden.)

Beispiele

1. V sei der \mathbb{K} -Vektorraum aller J -Spalten;
 W sei der \mathbb{K} -Vektorraum aller I -Spalten.
 Wenn A eine $I \times J$ -Matrix ist, dann ist

$$\varphi : \mathbb{R}^J \ni x \longmapsto Ax \in \mathbb{R}^I$$

eine lineare Abbildung von V nach W . Jede lineare Abbildung von \mathbb{R}^J nach \mathbb{R}^I ist durch eine $I \times J$ -Matrix A gegeben; A ist durch die Abbildung eindeutig bestimmt.

2. Jede affine Abbildung Φ des affinen Raums \mathbb{R}^J in den affinen Raum \mathbb{R}^I hat die Gestalt

$$x \longmapsto b + Ax$$

mit einer wohlbestimmten I -Spalte b und einer wohlbestimmten $I \times J$ -Matrix A .

Satz

Sei \mathbb{K} ein Körper mit $1+1 \neq 0$. Seien (L, V) und (M, W) \mathbb{K} -affine Räume. Eine Abbildung

$$\Phi : L \rightarrow M$$

ist genau dann eine affine Abbildung, wenn gilt

$$\forall P, Q \in L \quad \forall \lambda \in \mathbb{K} \quad \Phi((1-\lambda)P + \lambda Q) = (1-\lambda)\Phi(P) + \lambda \cdot \Phi(Q) \quad .$$

Beweis

Die Bedingung ist offenbar notwendig; sie folgt aus den Eigenschaften der linearen Abbildung zu Φ . Um zu zeigen, dass sie auch hinreichend ist, wählen wir einen Punkt P_0 und definieren

$$\varphi(v) = \Phi(P_0 + v) - \Phi(P_0) \quad .$$

Nach dem Schema von oben zeigen wir leicht, dass $\varphi(\cdot)$ eine lineare Abbildung ist.

$$(i) \quad \varphi(\lambda v) = \lambda \cdot \varphi(v) \quad \text{für alle } \lambda \in \mathbb{K}, v \in V$$

$$(ii) \quad \varphi\left(\frac{1}{2}(v+w)\right) = \frac{1}{2}(\varphi(v) + \varphi(w)) \quad .$$

Bemerke: Es gibt einen Körper mit nur zwei Elementen; in diesem Körper $\mathbb{K} = \mathbb{F}_2$ gibt es nur das Nullelement und das Einselement. Für Abbildungen eines \mathbb{F}_2 -Vektorraums ist die Aussage

$$\forall P, Q \in L \quad \forall \lambda \in \mathbb{K} \quad \Phi((1-\lambda)P + \lambda Q) = (1-\lambda)\Phi(P) + \lambda \cdot \Phi(Q) \quad .$$

leer; sie ist für alle Funktionen erfüllt und nicht nur für die \mathbb{F}_2 -affinen Funktionen.

Im Falle $1+1 \neq 0$ besagt unser Satz in der Sprache der Geometer:

Satz

Eine Abbildung $\Phi : L \rightarrow M$ ist genau dann eine affine Abbildung, wenn sie Gerade so in Gerade abbildet, dass die Teilungsverhältnisse erhalten bleiben.

Bemerke :

Wenn P_k , $k = 1, \dots, m$ irgendwelche Punkte sind und λ_k Skalare mit $\sum \lambda_k = 1$, dann gilt für jede affine Abbildung Φ

$$\Phi\left(\sum \lambda_k P_k\right) = \sum \lambda_k \cdot \Phi(P_k) \quad ,$$

Satz :

Seien P_0, P_1, \dots, P_n Punkte in einem n -dimensionalen affinen Raum L , die nicht in einer Hyperebene liegen. Seien Q_0, Q_1, \dots, Q_n Punkte in einem affinen Raum M . Es existiert dann genau eine affine Abbildung $\Phi : L \rightarrow M$ mit $\Phi(P_k) = Q_k$ für $k = 0, 1, \dots, n$.

Der Beweis liegt auf der Hand.

Pullback und duale Abbildung

Wenn $\Psi : \Omega_1 \rightarrow \Omega_2$ irgendeine Abbildung ist, dann ordnet bekanntlich der Pullback Ψ^* jeder \mathbb{K} -wertigen Funktion auf Ω_2 eine Funktion auf Ω_1 zu. Der Pullback bildet die Summe von Funktionen in die Summe der Bildfunktionen ab und das punktweise Produkt von \mathbb{K} -wertigen Funktionen in das punktweise Produkt der Bildfunktionen.

Das besondere bei den affinen Abbildungen

$$\Phi : L \rightarrow M$$

besteht darin, dass der Pullback die affinen Funktionen auf M nicht in irgendwelche Funktionen auf L abbildet, sondern in affine Funktionen. Wir werden das sofort beweisen. (In der nächsten Vorlesung werden wir übrigens sehen, dass der Pullback einer affinen Abbildung die quadratischen Funktionen in quadratische Funktionen überführt, und allgemeiner, dass polynomiale Funktionen vom Grad $\leq m$ in polynomiale Funktionen vom Grad $\leq m$ abgebildet werden.)

Satz :

V und W seien \mathbb{K} -Vektorräume. Eine Abbildung $\varphi : V \rightarrow W$ ist genau dann eine lineare Abbildung, wenn der Pullback die Linearformen auf W in Linearformen auf V abbildet.

(Die Abbildung $\varphi^* : W^* \leftarrow V^*$ heißt in diesem Fall die zu φ **duale Abbildung**.)

Beweis des Satzes

1. Sei φ eine lineare Abbildung und $m(\cdot) \in W^*$. $\ell(\cdot)$ sei die zurückgenommene Funktion

$$\ell(v) := m(\varphi(v)) \quad .$$

$\ell(\cdot)$ ist Linearform; denn

$$\begin{aligned}\ell(\alpha v_1 + \beta v_2) &= m(\varphi(\alpha v_1 + \beta v_2)) = m(\alpha \varphi(v_1) + \beta \varphi(v_2)) \\ &= \alpha \cdot m(\varphi(v_1)) + \beta \cdot m(\varphi(v_2)) = \alpha \cdot \ell(v_1) + \beta \cdot \ell(v_2) \quad .\end{aligned}$$

2. Sei $\{w_i : i \in I\}$ eine Basis von W und sei $\{m^i(\cdot) : i \in I\}$ die duale Basis. $\varphi(\cdot)$ sei eine Abbildung, $\ell^i(\cdot)$ sei das Bild von $m^i(\cdot)$ bzgl. des Pullback; $\ell^i(v) = m^i(\varphi(\cdot))$. Wenn die $\ell^i(\cdot)$ linearformen sind, dann gilt für jeden Vektor v

$$\varphi(v) = w = \sum \alpha^i w_i = \sum m^i(w) \cdot w_i = \sum m^i(\varphi(v)) \cdot w_i = \sum \ell^i(v) \cdot w_i \quad .$$

Also gilt

$$\begin{aligned}\varphi(\alpha v) &= \sum \ell^i(\alpha v) \cdot w_i = \alpha \cdot \sum \ell^i(v) w_i = \alpha \cdot \varphi(v) \\ \varphi(v_1 + v_2) &= \sum \ell^i(v_1 + v_2) \cdot w_i = \varphi(v_1) + \varphi(v_2) \quad .\end{aligned}$$

Den Beweis des folgenden Satzes überlassen wir dem Leser.

Satz :

Seien L und M affine Räume. Eine Abbildung

$$\Phi : L \rightarrow M$$

ist genau dann eine affine Abbildung, wenn der Pullback jeder affinen Funktion auf M eine affine Funktion auf L zuordnet.

Hinweis

Wir werden noch mehrere Gelegenheiten finden, wo ein spezieller Typ von Abbildungen $\Psi : \Omega_1 \rightarrow \Omega_2$ dadurch gekennzeichnet ist, dass der Pullback Ψ^* Funktionen einer gewissen Art (auf Ω_2) Funktionen einer gewissen Art (auf Ω_1) zuordnet.

Aufgaben zu II.4

Aufgabe II.4.1 :

Eine Teilmenge K eines affinen Raums heißt bekanntlich eine konvexe Menge, wenn gilt

$$\forall P_0, P_1 \in K \quad \forall \lambda \in [0, 1] \quad (1 - \lambda)P_0 + \lambda P_1 \in K.$$

Zu jeder Teilmenge eines reellaffinen Raums gibt es offenbar eine kleinste sie umfassende konvexe Menge. (Man nennt sie die konvexe Hülle.) Sie ist der Durchschnitt aller die gegebene Menge umfassenden konvexen Mengen.

Sei $M = \{P_\alpha : \alpha \in I(\text{Indexmenge})\}$ eine beliebige Familie von Punkten in einem affinen Raum.

Zeigen Sie: Die konvexe Hülle ist die Menge aller Punkte der Gestalt

$$P = \lambda_0 P_0 + \lambda_1 P_1 + \dots + \lambda_m P_m, \quad \text{mit den } P_j \in M, \quad \lambda_j \geq 0, \quad \sum \lambda_j = 1.$$

Aufgabe II.4.2 :

Man nennt ein $(m+1)$ -Tupel von Punkten P_0, P_1, \dots, P_m in einem n -dimensionalen affinen Raum ein Tupel in allgemeiner Lage, wenn die Verschiebungsvektoren $\vec{P_0 P_1}, \vec{P_0 P_2}, \dots, \vec{P_0 P_m}$ linear unabhängig sind.

Zeigen Sie: Wenn P_0, P_1, \dots, P_m ein m -Tupel in allgemeiner Lage ist, dann besitzt jedes P in der affinen Hülle genau eine Darstellung

$$P = \lambda_0 P_0 + \lambda_1 P_1 + \dots + \lambda_m P_m, \quad \text{mit } \sum \lambda_j = 1.$$

(Die Zahlen λ_j nennt man manchmal die Schwerpunktskoordinaten des Punkts P .)

II.5 Quadratische Funktionen

Bei unserer Diskussion der quadratischen Funktionen auf einem endlichdimensionalen affinen Raum bemühen wir uns um eine Annäherung der Standpunkte von Algebra, Geometrie und Analysis. Wir denken vor allem an den reellen Fall, obschon fast alle Argumente auch für jeden Grundkörper \mathbb{K} mit $1 + 1 \neq 0$ funktionieren. Wir stellen daher vorerst nur die Forderung $1 + 1 \neq 0$, und wir werden (koordinatenfrei!) den Begriff der quadratischen Funktion auf einem \mathbb{K} -affinen Raum entwickeln.

Definition

Eine \mathbb{K} -wertige Funktion $q(\cdot)$ auf einem \mathbb{K} -affinen Raum heißt **quadratische Funktion**, wenn für jedes $v \in V$ der Zuwachs $q_v(P) = q(P + v) - q(P)$ eine affine Funktion $q_v(\cdot)$ ist.

Eine **quadratische Form** ist eine quadratische Funktion $Q(v)$ auf einem Vektorraum mit $Q(\alpha v) = \alpha^2 Q(v)$ für alle Skalare α . Eine quadratische Form nennt man auch eine homogene quadratische Funktion.

Satz :

Wenn $f(\cdot)$ und $g(\cdot)$ affine Funktionen sind, dann ist das punktweise Produkt $h = f \cdot g$ eine quadratische Funktion.

Beweis Wir haben Linearformen $\ell(\cdot)$ und $m(\cdot)$ mit

$$f(P + v) = f(P) + \ell(v), \quad g(P + v) = g(P) + m(v).$$

Für jedes feste v gilt

$$\begin{aligned} h(P + v) - h(P) &= f(P + v) \cdot g(P + v) - f(P) \cdot g(P) = \\ &= [f(P + v) - f(P)] \cdot g(P + v) + f(P) \cdot [g(P + v) - g(P)] \\ &= \ell(v) \cdot [g(P) + m(v)] + f(P) \cdot m(v) \\ &= c_0 + c_1 \cdot f(P) + c_2 \cdot g(P) \quad . \end{aligned}$$

Der Zuwachs von $h(\cdot)$ als Funktion von P ist also eine Linearkombination der affinen Funktionen $f(\cdot)$ und $g(\cdot)$ plus eine Konstante. Er ist somit eine affine Funktion, und zwar für jedes v . (Die Koeffizienten der Linearkombination hängen natürlich von v ab.)

Definition

Eine quadratische Funktion wird auch eine polynomiale Funktion vom Grad ≤ 2 genannt. Eine \mathbb{K} -wertige Funktion $g(\cdot)$ heißt eine **polynomiale Funktion** vom Grad $\leq m$, wenn für jedes v der Zuwachs $g_v(P) = g(P + v) - g(P)$ eine polynomiale Funktion vom Grad $\leq m - 1$ ist. ($m = 2, 3, \dots$).

Durch vollständige Induktion beweist man leicht den

Satz :

Wenn $f(\cdot)$ polynomial vom Grad $\leq \ell$ ist und $g(\cdot)$ polynomial vom Grad $\leq m$, dann ist das punktweise Produkt $h(\cdot) = (f \cdot g)(\cdot)$ polynomial vom Grad $\leq \ell + m$.

Hinweis Bei der Analyse der polynomialen Funktionen auf einem affinen Raum arbeitet man üblicherweise mit symmetrischen Multilinearformen und Tensoren. Die zu diesen Begriffen gehörige Theorie heißt die **Multilineare Algebra**. In diese Theorie gehören übrigens auch die alternierenden Multilinearformen und speziell die Determinanten. In dieser Anfängervorlesung wollen wir auf all das noch nicht eingehen. Wir wollen hier nur die quadratischen Funktionen diskutieren, und dafür benötigen wir nur einen sehr speziellen Begriff aus dieser Theorie, nämlich den Begriff der symmetrischen Bilinearform. — Für unsere bescheidenen Zwecke können wir wieder auf die Matrizen zurückgreifen. Man beachte aber, dass die hier auftretenden symmetrischen Matrizen keine Gemeinsamkeiten mit den Matrizen eines Basiswechsels haben und sehr wenige Gemeinsamkeiten mit den Matrizen zur Darstellung einer linearen Abbildung. (Man nennt sie manchmal Gram-Matrizen).

Definition : V sei ein \mathbb{K} -Vektorraum.

Eine \mathbb{K} -wertige Funktion $B(\cdot, \cdot)$ auf $V \times V$ heißt **symmetrische Bilinearform**, wenn gilt

- (i) $B(v, w) = B(w, v)$
- (ii) $B(v, \alpha w) = \alpha \cdot B(v, w)$
- (iii) $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$.

Die Funktion $Q(v) = B(v, v)$ heisst die von $B(\cdot, \cdot)$ erzeugte quadratische Form.

Die Bedingung (iii) heißt die Additivität im zweiten Argument; wegen (i) haben wir auch Additivität im ersten Argument. Zusammen mit (ii) erhalten wir die Linearität in beiden Argumenten, also die Bilinearität von $B(\cdot, \cdot)$.

Bemerke: In einem \mathbb{Q} -Vektorraum ist (ii) eine Konsequenz von (iii).

Satz

Sei $Q(\cdot)$ die quadratische Form zur Bilinearform $B(\cdot, \cdot)$. $Q(\cdot)$ ist dann in der Tat eine homogene quadratische Funktion auf V , und es gilt

- 1) $Q(v + w) = B(v + w, v + w) = Q(v) + Q(w) + 2 \cdot B(v, w)$
- 2) $Q(v + w) + Q(v - w) = 2 \cdot Q(v) + 2 \cdot Q(w)$
- 3) $B(v, w) = \frac{1}{4}Q(v + w) - \frac{1}{4}Q(v - w) = \frac{1}{2}[Q(v + w) - Q(v) - Q(w)]$.

Der Beweis ergibt sich aus der Bilinearität. Die Gleichung 2) heisst die Parallelogrammgleichung; Die Gleichung 3) zeigt, wie man eine Bilinearform aus der von ihr erzeugten quadratischen Form zurückgewinnen kann.

Hinweis Die Parallelogrammgleichung wird uns auch bei den sog. Sesquilinearformen in Abschnitt IV begegnen. Der folgende Satz zeigt, wie man zu einer beliebigen quadratischen Form eine erzeugenden Bilinearform gewinnt. Mit etwas mehr Mühe werden wir in IV sehen, wie man von einer ‘guten’ hermiteschen Form zu einer erzeugenden Sesquilinearform gelangt.

Satz

Jede quadratische Form ist von einer symmetrischen Bilinearform erzeugt.

Wenn $q(\cdot)$ eine quadratische Funktion auf dem \mathbb{K} -affinen Raum (L, V) ist, dann gibt es eine quadratische Form $Q(\cdot)$ und (zu jedem $P \in L$) eine Linearform $\ell_P(\cdot)$, sodass gilt

$$q(P + v) - q(P) = \ell_P(v) + \frac{1}{2}Q(v) .$$

Beweis

Wir zeigen dass die sog. **gemischte Differenz** zweimal die gesuchte Bilinearform liefert.

$$B(v, w) := q(P + v + w) - q(P + w) - q(P + v) + q(P) .$$

Wir wissen, dass der Zuwachs in Richtung v eine affine Funktion $q_v(\cdot)$ ist. Zu dieser affinen Funktion gehört eine Linearform; wir nennen sie $B(v, \cdot)$. Es gilt

$$\begin{aligned} q_v(P + \cdot) - q_v(P) &= B(v, \cdot) \\ q(P + v + \cdot) - q(P + \cdot) - q(P + v) + q(P) &= B(v, \cdot) . \end{aligned}$$

Wenn wir w einsetzen, erhalten wir die gemischte Differenz, eine symmetrische Funktion der Argumente, welche linear im zweiten Argument ist. $B(\cdot, \cdot)$ ist also bilinear.

Wenn wir für eine quadratische Form $Q(\cdot)$ die gemischte Differenz auf der Diagonale auswerten, dann erhalten wir $B(v, v) = Q(v + v) - Q(v) - Q(v) = 2Q(v)$. Die gegebene quadratische Form ist also von der halben gemischten Differenz erzeugt.

Wir betrachten jetzt eine beliebige quadratische Funktion $q(\cdot)$. Für festes P ist die Differenz $\ell(v) = q(P + v) - q(P) - \frac{1}{2}Q(v)$ eine Linearform, denn für jedes w gilt

$$\begin{aligned} \ell(v + w) - \ell(w) &= q(P + v + w) - q(P + w) - \left(B(v, w) + \frac{1}{2}Q(w) \right) \\ &= q(P + v) - q(P) - \frac{1}{2}Q(w) = \ell(v) . \end{aligned}$$

Fazit: Wir haben somit den Begriff der quadratischen Funktion in ähnlicher Weise auf den algebraischen Begriff der symmetrischen Bilinearform zurückgeführt wie wir oben den Begriff der affinen Funktion auf den algebraischen Begriff der Linearform zurückgeführt haben.

Bemerkungen:

1. Wir haben früher gesehen, dass eine Funktion $f(\cdot)$ genau dann eine affine Funktion ist, wenn die Einschränkung auf jede Gerade eine affine Funktion ist.

$$f((1 - \lambda)P + \lambda Q) = (1 - \lambda) \cdot f(P) + \lambda \cdot f(Q) .$$

Analog gilt : Eine Funktion $q(\cdot)$ ist genau dann eine quadratische Funktion, wenn die Einschränkung auf jede Ebene eine quadratische Funktion ist. Wenn nämlich (für festes v) die Funktion $q_v(\cdot)$ auf jeder Geraden $\{R : R = P + \lambda v\}$ affin ist, dann ist sie auf dem ganzen Raum affin.

2. Eine Bilinearform $B(\cdot, \cdot)$ kann man als lineare Abbildung $\varphi : V \rightarrow V^*$ auffassen. Jedem $v \in V$ wird eine Linearform $B(v, \cdot)$ zugeordnet und die Zuordnung ist linear.

$$B(\alpha v_1 + \beta v_2, \cdot) = \alpha \cdot B(v_1, \cdot) + \beta \cdot B(v_2, \cdot) .$$

In der projektiven Geometrie benützt man die nichtausgearteten Bilinearformen, um den ‘Punkten’ Hyperebenen zuzuordnen; man stellt so die sog. Beziehung zwischen Pol und Polare her. Die projektive Geometrie ist hier aber nicht unser Thema.

3. In der reellen Theorie würde man $\ell_P(\cdot)$ den Gradienten von $q(\cdot)$ im Punkt P nennen; der Gradient ist eine Linearform auf dem Tangentialraum im Fußpunkt. Bei einer Funktion, die nicht affin ist, hängt der Gradient vom Fußpunkt P ab; bei quadratischen Funktionen in affiner Weise. Bei uns sind alle Tangentialräume identisch. Unsere Bilinearform macht die Veränderung des Gradienten explizit. Die Verschiebung des Fußpunkts um w bewirkt die Addition der Linearform $B(w, \cdot)$:

$$\ell_{P+w}(\cdot) = \ell_P(\cdot) + B(w, \cdot).$$

In der Tat haben wir $q(P+v) - q(P) = \ell_P(v) + \frac{1}{2}Q(v)$. Für $R = P + w$ ergibt sich

$$\begin{aligned} \ell_R(v) + \frac{1}{2}Q(v) &= q(R+v) - q(R) = q(P+w+v) - q(P+w) \\ &= \ell_P(w+v) - \ell_P(w) + \frac{1}{2}Q(v+w) - \frac{1}{2}Q(w) = \ell_P(v) + B(w, v) + \frac{1}{2}Q(v) . \end{aligned}$$

Koordinatendarstellung

Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und $\{v_j : j \in J\}$ eine Basis. Wie üblich werden die Vektoren v durch **J-Spalten** x dargestellt; die Darstellung $v = \sum v_j \cdot x^j$ liefert die Einträge x^j . Jede Linearform $\ell(\cdot)$ wird durch eine **J-Zeile** ξ dargestellt. Die Einträge ξ_j ergeben sich aus der Darstellung $\ell(\cdot) = \sum \xi_j \cdot \ell^j(\cdot)$ mittels der dualen Basis $\{\ell^j(\cdot) : j \in J\}$.

Satz (Darstellende Matrix)

Zu jeder symmetrischen Bilinearform $B(\cdot, \cdot)$ gibt es genau eine symmetrische $J \times J$ -Matrix B , sodass gilt

$$B(v, w) = x^T \cdot B \cdot y \quad \text{für} \quad v = \sum v_j \cdot x^j, \quad w = \sum v_j \cdot y^j .$$

Beweis Die Matrix mit den Einträgen $b_{jk} = B(v_j, v_k)$ und keine andere leistet das Verlangte. (Man nennt sie manchmal die Gram-Matrix). Es gilt in der Tat

$$B(v, w) = B\left(\sum v_j x^j, \sum v_k y^k\right) = \sum x^j \cdot y^k \cdot B(v_j, v_k) = \sum x^j b_{jk} y^k ;$$

und in Matrizenschreibweise

$$\begin{aligned} B(v, w) &= x^T \cdot B \cdot y = \\ &= (x^1, \dots, x^n) \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} . \end{aligned}$$

Wenn wir $v = w$ setzen, dann erhalten wir die quadratische Form $Q(v) = x^T \cdot B \cdot x$.

Satz (Basiswechsel)

Sei $Q(\cdot)$ eine quadratische Form auf einem \mathbb{K} -Vektorraum V .

Seien $\{v_j : j \in J\}$ und $\{u_i : i \in I\}$ Basen von V , A sei die $I \times J$ -Matrix des Basiswechsels

$$(v_1, \dots, v_n) = (u_1, \dots, u_n) \cdot A, \quad \text{d. h.} \quad v_j = \sum u_i \cdot a_j^i \quad \text{für alle } j \in J.$$

Für die symmetrischen Matrizen, welche $Q(\cdot)$ bzgl. dieser Basen darstellen, gilt dann

$$B_{J \times J} = A^\top \cdot B_{I \times I} \cdot A.$$

Beweis Bemerke zuerst: A ist eine $I \times J$ -Matrix, A^\top ist eine $J \times I$ -Matrix; das Produkt $B_{J \times J} = A^\top \cdot B_{I \times I} \cdot A$ ist also tatsächlich eine wohldefinierte symmetrische $J \times J$ -Matrix.

Die Einträge der darstellenden Matrizen erhalten wir, indem wir die Bilinearform in allen Paaren von Basisvektoren auswerten:

$$B(v_j, v_k) = B\left(\sum u_i \cdot a_j^i, \sum u_l \cdot a_k^l\right) = \sum_{i,l} a_j^i B(u_i, u_l) a_k^l.$$

Die Koordinatendarstellung einer quadratischen Form auf V kann man auch folgendermaßen zum Ausdruck bringen: Sei $Q(\cdot)$ eine quadratische Form zur symmetrischen Bilinearform $B(\cdot, \cdot)$. Sei $\{v_j : j \in J\}$ eine Basis von V und sei $\{\ell^j(\cdot) : j \in J\}$ die duale Basis. Es gilt dann

$$Q(v) = \sum_{j,k} b_{jk} \cdot \ell^j(v) \cdot \ell^k(v) \quad \text{mit} \quad b_{jk} = B(v_j, v_k).$$

Die quadratischen Formen nennt man auch homogene quadratische Funktionen (auf einem Vektorraum). Für die ‘nichthomogenen’ quadratischen Funktionen auf einem \mathbb{K} -affinen Raum L ($1 + 1 \neq 0$) haben wir in einem affinen Koordinatensystem $\{x^j : j \in J\}$: die Darstellung

$$q(\cdot) = b_0 + \sum_j b_j \cdot x^j + \frac{1}{2} \sum_{j,k} b_{jk} \cdot x^j \cdot x^k.$$

In der Analysis würde man sagen: Der Skalar b_0 ist der Wert und die Zeile (b_1, \dots, b_n) der Gradient im Koordinatenursprung; und die symmetrische Matrix mit den Einträgen b_{jk} ist die Hesse-Matrix der zweiten Ableitungen.

Einen Weg zur affinen Klassifikation weist die Methode der quadratischen Ergänzung, die wir hier nur an einigen zweidimensionalen Beispielen erläutern wollen.

Beispiele zur quadratischen Ergänzung

Eine homogene quadratische Funktion auf dem Vektorraum V aller 2-Spalten $\begin{pmatrix} x \\ y \end{pmatrix}$ ist eine Funktion der Form

$$q(v) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = (x, y) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Eine solche Funktion kann man auf verschiedene Weisen als signierte Summe von zwei Quadraten schreiben

$$q(v) = \pm x'(v)^2 \pm y'(v)^2 \quad \text{mit Linearformen} \quad x'(\cdot), y'(\cdot).$$

Beispiele

$$\begin{aligned}
-7x^2 + 12xy + 2y^2 &= -7\left(x - \frac{6}{7}y\right)^2 + \frac{50}{7}y^2 \\
&= -25x^2 + 2(y + 3x)^2 \\
&= (x + 2y)^2 - 2(-2x + y)^2 \\
x^2 + 2xy + 3y^2 &= (x + y)^2 + 2y^2 \\
&= \frac{2}{3}x^2 + 3\left(y + \frac{1}{3}x\right)^2 \\
&= \frac{1}{2}\left(x + (1 + \sqrt{2})y\right)^2 + \frac{1}{2}\left(x + (1 - \sqrt{2})y\right)^2 \\
x^2 + y^2 &= \left(\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y\right)^2 + \left(\frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y\right)^2 \\
&= \left(\frac{1}{2}x + \frac{1}{2}\sqrt{3}y\right)^2 + \left(-\frac{1}{2}\sqrt{3}x + \frac{1}{2}y\right)^2
\end{aligned}$$

Wir machen den Raum V zu einem **euklidischen Raum** mit der üblichen Norm $\|v\| = (x^2 + y^2)^{\frac{1}{2}}$. Die Linearformen $\alpha x + \beta y$ und $\gamma x + \delta y$ stehen somit genau dann aufeinander senkrecht, wenn gilt $\alpha \cdot \gamma + \beta \cdot \delta = 0$. In den ersten beiden Beispielen stehen die Linearformen aus der letzten Zeile aufeinander senkrecht. Für die ersten Zeilen haben wir uns weniger Mühe gemacht; wir haben die (aus der Schule wohlbekannte) Methode der quadratischen Ergänzung angewandt.

Die Zerlegung mit orthogonalen Linearformen ist in den ersten beiden Beispielen eindeutig bestimmt (wie wir sehen werden); im dritten Beispiel gibt es unendlich viele ‘orthogonale Zerlegungen’ (der offizielle Name ist ‘Hauptachsenzerlegung’).

Hinweis: Quadriken

Das Nullstellengebilde einer quadratischen Funktion in einem (n -dimensionalen) reellaffinen Raum L nennt man eine Hyperfläche zweiter Ordnung oder auch eine Quadrik.

(Im Falle $n = 3$ sollte man wohl besser von einer ‘Fläche’ und im Falle $n = 2$ von einer ‘Kurve’ sprechen. In jedem Falle ist aber zu beachten, dass das Nullstellengebilde ‘ausgeartet’ sein kann; es kann z. B. leer sein oder nur aus einem einzigen Punkt bestehen, und es gibt noch viele weitere Möglichkeiten, wo die Rede von einer Hyperfläche unangebracht erscheint.)

In vielen traditionellen Lehrbüchern der ‘Analytischen Geometrie’ findet man für den reellen und für den komplexen Fall eine vollständige Klassifikation aller Quadriken in den Dimensionen $n = 2$ und $n = 3$. Darauf wollen wir hier verzichten.

Kegelschnitte Wichtige ‘Kurven zweiter Ordnung’ sind die natürlich die Ellipsen, die Hyperbeln und die Parabeln, die ja (wie die Namen verraten) schon von den alten Griechen studiert wurden. Diese nichtausgearteten ‘Kegelschnitte’ erscheinen bekanntlich auch in der Kepler’schen Theorie der Bahnen von Himmelskörpern. Allerdings werden sie in diesen klassisch geometrischen Zusammenhängen normalerweise nicht in affinen Koordinaten dargestellt. Die Wahl von passenden affinen Koordinatensystemen und der Matrizenkalkül sind natürlich kräftige Hilfsmittel bei der Untersuchung der ‘Kurven zweiter Ordnung’.

Die folgende Konstruktion erklärt den traditionellen Namen ‘Kegelschnitt’: Ein Doppelkegel im dreidimensionalen Anschauungsraum ist in passenden affinen Koordinaten das Nullstellengebilde der homogenen quadratischen Funktion $z^2 - x^2 - y^2$; die Durchschnitte mit Hyperebenen heissen Kegelschnitte. Wenn nun die quadratische Funktion $q(\cdot)$ in einem bestimmten affinen Koordinatensystem gegeben ist

$$q(P) = b_0 + \sum_{j=1}^n b_j \cdot x^j(P) + \frac{1}{2} \sum_{j,k=1}^n b_{jk} \cdot x^j(P) x^k(P),$$

dann geht man bequemerweise zu einer quadratischen Form auf dem \mathbb{R}^{n+1} über

$$Q(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \cdot \mathbf{Q} \cdot \mathbf{x} = \frac{1}{2} \sum_{j,k=0}^n q_{jk} \cdot x^j x^k \quad \text{für die Spalten } \mathbf{x} = (x^0, \dots, x^n)^\top, \quad \text{wo}$$

$$q_{00} = 2b_0, \quad q_{0j} = q_{j0} = b_j \text{ für } j = 1, \dots, n \text{ und } q_{jk} = b_{jk} \text{ für } j, k = 1, \dots, n.$$

Nun wird die Funktion $q(\cdot)$ durch die Einschränkung der quadratischen Form auf die Hyperebene $\{x^0 = 1\}$ beschrieben.

Es zeigt sich, dass die interessantesten Typen quadratischer Funktionen durch die Wahl von passenden affinen Koordinaten $x(\cdot), y(\cdot)$ in die folgende Gestalt gebracht werden können:

$$\begin{aligned} q(P) &= x^2(P) + y^2(P) + \text{const} \\ \text{bzw.} &= x^2(P) - y^2(P) + \text{const} \\ \text{bzw.} &= x^2(P) - y(P) + \text{const} \\ \text{bzw.} &= x^2(P) + \text{const} . \end{aligned}$$

Im ersten Fall sind die Niveaumengen leer oder Ellipsen, im zweiten Fall sind sie Hyperbeln oder ein Paar sich schneidender Geraden, und im dritten Fall sind sie Parabeln. Das Nullstellengebilde der Funktion $x^2(P) + \text{const}$ kann ein Geradenpaar sein oder eine ‘Doppelgerade’, wenn es nicht überhaupt leer ist.

Exkurs: Eine spezielle quadratische Form im ‘Anschauungsraum’

In der speziellen Relativitätstheorie hat man es mit einem vierdimensionalen reellaffinen Raum zu tun. Die Punkte heißen die Weltpunkte. Man stellt die Verschiebungsvektoren (in einem „Inertialsystem“) durch ein Quadrupel von Zahlen dar.

$$(x_0, x_1, x_2, x_3) \quad \text{oder} \quad (ct, x, y, z) \quad .$$

Jedem Verschiebungsvektor wird eine Zahl zugeordnet.

$$Q(\vec{PR}) = x_0^2 - (x_1^2 + x_2^2 + x_3^2) = (ct)^2 - \|x\|^2 \quad .$$

Diese reellwertige Funktion ist die quadratische Form zu einer Bilinearform. Man sagt von zwei Weltpunkten P und R , dass sie **zeitartig** zueinander liegen, wenn $Q(\vec{PR}) > 0$. Man nennt den Verschiebungsvektor **raumartig**, wenn $Q(\vec{PR}) < 0$. Die Verschiebungsvektoren mit dem Q -Wert $= 0$ sind die Vektoren auf dem **Lichtkegel**. Man zeigt: Wenn $Q(\vec{PR}) > 0$, dann kann man ein Inertialsystem einführen, in welchem die Weltpunkte P und R denselben Ort zu verschiedenen Zeiten einnehmen. $\sqrt{Q(\vec{PR})}$ gibt an, welche „Eigenzeit“ zwischen den Ereignissen liegt. Wenn $Q(\vec{PR}) < 0$, dann kann man ein Inertialsystem einführen, in welchem P und R als gleichzeitig erscheinen. $\sqrt{-Q(\vec{PR})}$ ist der euklidische Abstand. Die zu $Q(\cdot)$ gehörige Bilinearform $B(\cdot, \cdot)$ ist für den Physiker ebenfalls wichtig; sie ist aber für den Anfänger schwerer zu interpretieren.

Eine quadratische Form derselben Gestalt wird auch für die sog. Covektoren benutzt. Ein Covektor ist z.B. der sog. **Vierervektorimpuls** eines bewegten Teilchens. In jedem Inertialsystem wird er durch ein Quadrupel (E, p_x, p_y, p_z) beschrieben; E ist die Energie, p ist der Impuls bzgl. des gewählten Koordinatensystems. Die Zahl

$$m_0 := \frac{1}{c} \sqrt{\left(\frac{E}{c}\right)^2 - (p_x^2 + p_y^2 + p_z^2)} \geq 0$$

ist die **Ruhemasse** des bewegten Teilchens.

Bemerke : Interessant ist z. B. der Vierervektorimpuls der Photonen. Ein Photon besitzt Energie $E > 0$ ($E = \hbar \cdot \omega$), obwohl es keine Ruhemasse hat. Die Energie besitzt es durch immerwährende Bewegung mit Lichtgeschwindigkeit (siehe Feynman, Bd I, Kapitel 17.5).

Hinweis Koordinatensysteme

Die affinen Koordinaten des Anschauungsraums verallgemeinern die von der Schule her bekannten cartesischen Koordinaten. Sie sind sehr gut geeignet für die Darstellung der polynomialen Funktionen und vieler anderer wichtiger Funktionen.

In der höheren Theorie der Funktionen auf dem Anschauungsraums betrachtet man dann aber auch allgemeinere Koordinatensysteme, sog. „krummlinige“ Koordinaten (auf Teilbereichen des Anschauungsraums oder auch auf sog. Mannigfaltigkeiten). Das bekannteste Beispiel sind die Polarkoordinaten auf der geschlitzten Ebene. Die Wahl eines geschickten Koordinatensystems ist ein Thema, welches uns immer wieder einmal beschäftigen wird.

Bei den krummlinigen Koordinaten und in der Theorie der glatten Mannigfaltigkeiten ist zu beachten, dass jeder Punkt seinen eigenen Tangentialraum hat. Zu jeder glatten Funktion f gewinnt man eine Linearform $\langle df, \cdot \rangle$ auf jedem dieser Tangentialräume. Der ‘Gradient’ von $\langle df, \cdot \rangle$ beschreibt die Änderungsgeschwindigkeit von f beim Fortschreiten entlang eines Tangentialvektors.

Wenn man ein ‘lokales Koordinatensystem’ bei P gewählt hat, dann kann man eine glatte Funktionen bei P ‘in erster Näherung’ durch eine affine Funktion und ‘in zweiter Näherung’ durch eine quadratische (Funktion in diesen Koordinaten) approximieren. Auch deswegen sind die affinen und die quadratischen Funktionen beliebte Funktionen.

Für die Approximationen ‘höherer Ordnung’ benützt man dann in der ‘Taylor-Formel’ polynomiale Funktionen höheren Grades; um sie zu beschreiben, müssten wir über den Matrizenkalkül hinausgehen.

Aufgaben zu II.5

Aufgabe II.5.1 :

Eine Ellipse im euklidischen \mathbb{R}^2 kann man durch eine quadratische Gleichung beschreiben. Wenn der Mittelpunkt im Koordinatenursprung liegt, handelt es sich um eine homogene quadratische Gleichung, die durch eine symmetrische 2×2 -Matrix vorgegeben werden kann.

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : a_{11}x^2 + (a_{12} + a_{21})xy + a_{22}y^2 = 1 \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : (x, y)A \begin{pmatrix} x \\ y \end{pmatrix} = 1 \right\}$$

(Die Teile a) und b) sind eine Anregung für diejenigen, die schon einmal mit MAPLE gespielt haben, die übrigen Studierenden können die Anregung einfach übergehen und sofort mit dem Rechnen beginnen.)

- a) MAPLE kann Ihnen helfen, Ellipsen zu zeichnen. Finden Sie heraus, wie dies mit dem MAPLE Befehl `implicitplot` funktioniert. Zeichnen Sie

$$E_1 : \frac{1}{9}x^2 + \frac{1}{4}y^2 = 1$$

$$E_2 : \frac{1}{9} \left(\frac{x+y}{\sqrt{2}} \right)^2 + \frac{1}{4} \left(\frac{x-y}{\sqrt{2}} \right)^2 = 1$$

$$E_3 : 9 \cdot x^2 + 4xy + 6y^2 = 1$$

$$E_4 : 10 \cdot \left(\frac{1}{\sqrt{5}}(2x + y) \right)^2 + 5 \cdot \left(\frac{1}{\sqrt{5}}(-x + 2y) \right)^2 = 1 .$$

Betrachten Sie die Bilder der Ellipsen unter Variation des Bildbereichs im Befehl `implicitplot`.

- b) Ergänzung zu E_3 & E_4 : Verifizieren sie

$$\begin{aligned} (x, y) \begin{pmatrix} 9 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \\ = (x', y') \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}, &\text{ wenn} \\ \begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}; \quad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \end{aligned}$$

- c) Auf jedem Kreis um den Nullpunkt gibt es einen Punkt, in welchem $9x^2 + 4xy + 6y^2$ maximal (bzw. minimal) ist.
Die Maximalpunkte (bzw. Minimalpunkte) liegen auf einer Geraden durch den Nullpunkt. Diese Geraden („Hauptachsen“) sollen bestimmt werden.

Hinweis : Wenn man einen Punkt P^* mit den kartesischen Koordinaten x^*, y^* um den Winkel ϕ dreht, erhält man den Punkt P_ϕ mit den kartesischen Koordinaten

$$\begin{aligned}a(\phi) &= \cos \phi \cdot x^* + \sin \phi \cdot y^* \\b(\phi) &= -\sin \phi \cdot x^* + \cos \phi \cdot y^*\end{aligned}$$

Die Funktion $F(\phi) = 9a^2(\phi) + 4a(\phi) \cdot b(\phi) + 6b^2(\phi)$ soll maximiert (bzw. minimiert) werden. Die Bedingung $\frac{d}{d\phi}F(\phi) = 0$ führt zum Ziel, wenn man beachtet

$$a'(\phi) = b(\phi), b'(\phi) = -a(\phi).$$

Man erhält eine quadratische Gleichung für $\frac{b}{a}(\tilde{\phi})$.

II.6 Polynome als Abbildungen von \mathbb{C} in sich

In dieser Vorlesung studieren wir keine Vektorräume von Funktionen und keine Gruppen von Abbildungen. Wir diskutieren spezielle Abbildungen von \mathbb{C} in sich; wir interessieren uns vor allem für die polynomialen Abbildungen

$$z \longmapsto p(z) = a_0 + a_1 z + \dots + a_n z^n.$$

Die Polynome werden nicht als Elemente des Polynomrings $\mathbb{C}[z]$ betrachtet, sondern als Abbildungen von \mathbb{C} auf \mathbb{C} . Der Fundamentalsatz der Algebra besagt, dass ein nichtkonstantes Polynom $p(z)$ jeden Wert annimmt; die Gleichung $p(z) = w$ hat Lösungen für jedes w . (Wenn man Vielfachheiten berücksichtigt, gibt es genau n Lösungen für jedes Polynom vom Grad n).

Im Schulunterricht werden quadratische Gleichungen behandelt. Die Formeln, die man dort für reelle Polynome lernt, gelten unverändert für komplexe Polynome

$$az^2 + 2bz + c = 0 \qquad z_{1,2} = \frac{1}{a} \left(-b \pm \sqrt{b^2 - ac} \right).$$

(Die Quadratwurzel ist im Komplexen bis auf das Vorzeichen definiert.)

Im 16. Jahrhundert hat man gefunden, dass man die Gleichungen dritten und vierten Grades mit ähnlichen Mitteln behandeln kann. Weithin bekannt ist die Cardanische Formel. Der damals eingeschlagene Weg erwies sich aber als eine Sackgasse. Er taugt nicht für Polynome vom Grad > 4 .

C. F. Gauss hat schließlich das Problem mit geometrischen Mitteln gelöst. Er gab mehrere Beweise des Fundamentalsatzes, den ersten in seiner Dissertation, 1799.

Die Algebra hat weiterhin viel zum Thema zu sagen. Durch Arbeiten von N.H. Abel (1802-1829) und E. Galois (1811-1832) wurde klar, warum Gleichungen vom Grad > 4 prinzipiell anders behandelt werden müssen als die Gleichungen vom Grad ≤ 4 . Galois entwickelte zu diesem Zweck den Begriff der Permutationsgruppe; er erkannte die grundlegenden Eigenschaften der Transformationsgruppen, die zu den Wurzeln einer polynomialen Gleichung gehören. Galois schaffte es aber nicht, seine Entdeckungen zu publizieren. Am Vorabend seines Todes in einem Duell hatte er einen Freund gebeten: „Du wirst Jacobi oder Gauss bitten, ihre Meinung nicht über die Wahrheit, sondern über Bedeutung der Sätze zu sagen. Danach wird es, so hoffe ich, einige Leute geben, die es für vorteilhaft erachten, diesen ganzen Wirrwarr zu entziffern“. Der Brief und die Entdeckungen wurden, soweit wir wissen, Jacobi und Gauss nicht vorgelegt; sie wurden der mathematischen Öffentlichkeit erst 1846 bekannt, zu einer Zeit, als Cauchy begonnen hatte, über Gruppentheorie zu arbeiten. Volles Verständnis für den Ansatz von Galois wurde erst um 1870 geschaffen (C. Jordan, F. Klein, S. Lie).

Erste Beweisskizze

Wir wollen zeigen, dass ein nichtkonstantes Polynom $p(\cdot)$ Nullstellen besitzt. Wir betrachten die p -Bilder \mathfrak{C}_r von Kreisen um den Nullpunkt $\{z_t = r \cdot e^{it}, t \in [0, 2\pi]\}$. Für kleine $r > 0$ liegen die Bildpunkte $w_t = p(z_t)$ nahe bei a_0 . Für große r überwiegt der letzte Term; er ist von der Größenordnung r^n .

$$w_t = a_0 + a_1 \cdot z_t + \dots + a_n \cdot (z_t)^n = a_n \cdot r^n [e^{int} + \text{Rest}] .$$

Der Faktor a_n ist eine Drehstreckung. $\{e^{int} : t \in [0, 2\pi]\}$ durchläuft n -mal den Einheitskreis $\left\{\frac{1}{a_n} w_t : t \in [0, 2\pi]\right\}$ ist also eine Kurve nahe am n -mal durchlaufenen Kreis mit dem großen Radius r^n . Die Kurven $\mathfrak{C}_r = \{p(re^{it}) : t \in [0, 2\pi]\}$ müssen für gewisse $r_1 \leq r_2 \leq \dots \leq r_n$ durch den Nullpunkt gehen, da die Anzahl der Umschlingungen des Nullpunkts von 0 auf n anwächst. (Hier wird ein Argument benutzt, welches heute unter dem Namen „Zwischenwertsatz“ diskutiert wird. Der erste, der es klar ausgesprochen hat, war Bolzano. Cauchy machte es um 1821 weithin bekannt.)

Das Argument mit den Umlaufzahlen zeigt: Wenn die Kurve $\mathfrak{C}_r = \{p(r \cdot e^{it}) : t \in [0, 2\pi]\}$ den Nullpunkt k -mal umschlingt, dann liegen k Nullstellen z_1, \dots, z_k in der Kreisscheibe $\{z : |z| < r\}$. Die Nullstellen müssen allerdings nicht paarweise verschieden sein; es können auch Mehrfach-Nullstellen auftreten. Es ist eindrucksvoll, wenn man die Bildkurven \mathfrak{C}_r für verschiedene r auf dem Bildschirm verfolgt.

MAPLE erweist sich als ein überzeugendes Hilfsmittel.

Zweite Beweisskizze

Der Absolutbetrag $|p(z)|$ ist groß, wenn $|z|$ groß ist. Es existiert ein R und eine Konstante sodass $|p(z)| \geq \text{const} \cdot |R|^n$ für $|z| \geq R$. Der Minimalwert von $|p(z)|$ kann also nicht weit draußen angenommen werden.

Sei \tilde{z} ein Punkt, in welchem $|p(z)|$ seinen Minimalwert annimmt. Wenn wir zeigen können, dass $p(\tilde{z}) \neq 0$ unmöglich ist, dann wissen wir, dass \tilde{z} eine Nullstelle ist.

Die Überlegung bringt einen Begriff ins Spiel, der uns in allgemeineren Zusammenhängen immer wieder begegnen wird, den Begriff der offenen stetigen Abbildung. Eine stetige Abbildung $\varphi(\cdot)$ heißt eine offene Abbildung, wenn sie jede offene Menge U auf eine offene Menge $\varphi(U)$ abbildet. Um das zu verstehen, muss man wissen, was eine offene Menge ist. Def. : Eine Menge U heißt offen, wenn sie mit jedem $P \in U$ auch eine volle Umgebung von P enthält.

Beispiel : Offene Teilmengen der reellen Achse \mathbb{R} sind leicht zu charakterisieren: Jede offene Menge U ist eine Vereinigung von offenen Intervallen, $U = \bigcup (a_i, b_i)$. Eine stetige reellwertige Funktion auf \mathbb{R} wollen wir hier als eine Abbildung verstehen. Sie ist genau dann eine offene Abbildung, wenn sie strikt monoton ist (isoton oder antiton). Wenn sie in einer Umgebung eines Punkts \tilde{x} nicht strikt monoton ist, dann ist das Bild einer kleinen Umgebung nicht offen.

Die nicht konstanten polynomialen Abbildungen von \mathbb{C} in sich sind stets offene Abbildungen. Das $p(\cdot)$ -Bild jeder offenen Kreisscheibe $\{z : |z - z_0| < r\}$ umfasst eine volle Umgebung von $p(z_0)$.

In der Tat sehen die Bilder kleiner Kreisscheiben fast wie Kreisscheiben aus

$$p(z) = p(z_0) + b_1 \cdot (z - z_0) + \dots + b_n \cdot (z - z_0)^n .$$

In den Punkten mit $b_1 \neq 0$ geht $p(z) - p(z_0)$ aus $(z - z_0)$ näherungsweise durch eine Drehstreckung hervor

$$p(z) - p(z_0) = b_1 \cdot (z - z_0) [1 + \text{Rest}(z_0, z)]$$

mit einem kleinen Rest für kleines $|z - z_0|$. Sei z_0 ein Punkt mit $0 = b_1 = \dots = b_{k-1}, b_k \neq 0$. Es gilt dann für $z \sim z_0$.

$$p(z) - p(z_0) = b_k(z - z_0)^k [1 + \text{Rest}(z_0, z)] .$$

Die Abbildung $z \mapsto z^k$ vergrößert die Winkelabstände um den Faktor k . Der Kreis mit dem Radius r wird in k -facher Umdeckung auf den Kreis mit dem Radius r^k abgebildet.

Corollar : Wenn $p(\cdot)$ in z_0 nicht verschwindet, dann gibt es in jeder Umgebung Punkte z_1 , die näher am Nullpunkt liegen.

Wenn $|p(\cdot)|$ in z_0 ein lokales Minimum besitzt, dann ist z_0 eine Nullstelle des Polynoms.

Bemerke : Aus der Beweisskizze wird ein Beweis, wenn wir wissen, dass die Funktion $|p(z)|$ ihr Minimum annimmt.

Wir werden lernen: Jede stetige reellwertige Funktion auf einem kompakten Raum nimmt ihr Minimum an. Der Schlüsselbegriff ist der Begriff der Kompaktheit. Dieser stand Gauss nicht zur Verfügung. (Er geht auf Weierstraß zurück. Bei Weierstraß findet sich auch ein Beweis(!), dass jede beschränkte abgeschlossene Teilmenge von \mathbb{C} kompakt ist.) Gauss verband mit der Idee der komplexen Zahlenebene ohne Weiteres die Überzeugung, dass jede Folge in der abgeschlossenen Kreisscheibe $\{|z| \leq R\}$ eine konvergente Teilfolge besitzt.

Bemerkung : Es genügt, den Fundamentalsatz der Algebra für diejenigen Polynome $p(z)$ zu beweisen, die keine Mehrfach-Nullstellen haben; denn man kann zu jedem Polynom $p(z)$ sehr leicht ein Polynom $q(z)$ konstruieren, welches dieselben Nullstellen hat, aber alle nur einfach. Sei nämlich

$$\begin{aligned} p(z) &= a_n \cdot z^n + a_{n-1} \cdot z^{n-1} + \dots + a_1 \cdot z + a_0 \\ p'(z) &= (n a_n) \cdot z^{n-1} + (n-1) a_{n-1} \cdot z^{n-2} + \dots + 2 a_2 z + a_1 \\ r(z) &= \text{ggT}(p, p') , \quad q(z) = \frac{p(z)}{r(z)} . \end{aligned}$$

Dann hat $q(\cdot)$ dieselben Nullstellen wie $p(\cdot)$, aber jede nur mit der Vielfachheit 1. Dies sieht man so: z_0 sei eine ℓ -fache Nullstelle von $p(\cdot)$ mit $\ell > 1$. $p(z) = (z - z_0)^\ell \cdot p_0(z)$, $p_0(z_0) \neq 0$. Nach der Produktregel (die ganz ohne Grenzübergänge trivial zu beweisen ist), gilt

$$\begin{aligned} p'(z) &= \ell \cdot (z - z_0)^{\ell-1} \cdot p_0(z) + (z - z_0)^\ell \cdot p'_0(z) = (z - z_0)^{\ell-1} \cdot p_1(z) \\ \text{ggT}(p, p') &= (z - z_0)^{\ell-1} \cdot r_1(z) \end{aligned}$$

$q(\cdot)$ hat also z_0 als einfache Nullstelle

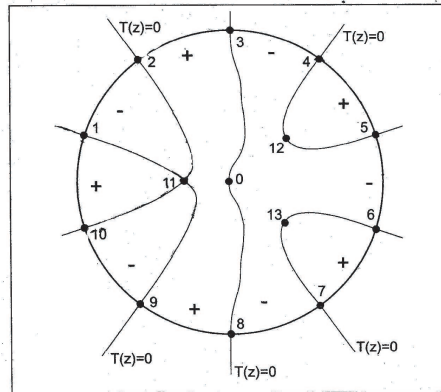
Dritte Beweisskizze

Gauss betrachtet die reellwertige Funktion $T(z) = \Re \left(\frac{1}{a_n \cdot R^n} p(z) \right)$ auf einem großen Kreis $\{z : |z| = R\}$. $T(\cdot)$ verhält sich bei einem Umlauf $z_t = R \cdot e^{it}$ wie $\cos(n \cdot t)$. $T(R \cdot e^{it})$ wechselt bei einem Umlauf n -mal das Vorzeichen von $+$ nach $-$ und n -mal von $-$ nach $+$.

Gauss verfolgt die Trennlinien ins Innere des Kreises. (Sie schneiden sich übrigens nicht, wenn $p(\cdot)$ keine Mehrfach-Nullstellen hat.) Er betrachtet andererseits die Funktion

$$S(z) = \Im \left(\frac{1}{a_n \cdot R^n} p(z) \right)$$

und ihre Trennlinien der Positivitätsbereiche. Dort wo sie sich mit den Trennlinien für $T(z)$ schneiden, haben wir Nullstellen von $p(\cdot)$.



Beispiel für eine Funktion 5. Grades aus Gaussens viertem Beweis.

Es ist visionär zu nennen, was Gauss (1849) zu seinem Argument schreibt:

Ich werde die Beweisführung in einer der Geometrie der Lage entnommenen Einkleidung darstellen, weil jene dadurch die größte Anschaulichkeit und Einfachheit gewinnt. Im Grunde gehört aber der eigentliche Inhalt der ganzen Argumentation einem höheren von Räumlichkeiten unabhängigen Gebiete der allgemeinen abstrakten Größenlehre an, dessen Gegenstand die nach der Stetigkeit zusammenhängenden Größenkombinationen sind, einem Gebiete, welches zur Zeit noch wenig ausgebaut ist, und in welchem man sich auch nicht bewegen kann ohne eine von räumlichen Bildern entlehnte Sprache.

(Zitiert nach „Die vier Gauss’schen Beweise für die Zerlegung ganzer algebraischen Funktionen . . .“, Verlag von Wilhelm Engelmann, Leipzig und Berlin, 1913.)

Exkurs : Auch für den Physiker gibt es gute Gründe, die komplexwertigen Funktionen einer komplexen Variablen als Abbildungen aufzufassen. Hier ein Beispiel:

Die Zhukowski-Abbildung

Wir diskutieren eine spezielle nichtpolynomiale Abbildung, die Zhukowski-Abbildung

$$\varphi : z \mapsto \frac{1}{2} \left(z + \frac{1}{z} \right) .$$

Sie galt früher als wichtig für die Strömungslehre („Zhukowski-Profil“). Wir behandeln nur einige mathematische Aspekte.

1. Die Zhukowski-Abbildung $\varphi(\cdot)$ bildet den Einheitskreis auf die („doppeltbelegte“) Strecke S zwischen -1 und $+1$ ab.

$$\frac{1}{2}(e^{it} + e^{-it}) = \cos t \quad .$$

2. Es gilt $\varphi\left(\frac{1}{z}\right) = \varphi(z)$. Sowohl das Äußere des Einheitskreises als auch das Innere des punktierten Einheitskreises $\{z : 0 < |z| < 1\}$ wird bijektiv auf $\mathbb{C} \setminus S$ abgebildet. Die Gleichung

$$w = \frac{1}{2} \left(z + \frac{1}{z} \right)$$

hat nämlich für jedes $w \neq S$ genau zwei Lösungen

$$\begin{aligned} 2zw &= z^2 + 1, & (z-w)^2 &= w^2 - 1 \\ z_{1,2} &= w \pm \sqrt{w^2 - 1} \quad . \end{aligned}$$

Eine der Lösungen hat Betrag > 1 ; die andere ist die reziproke Zahl mit Betrag < 1 .

3. Die Kreise um den Nullpunkt werden in konfokale Ellipsen abgebildet

$$\varphi(r \cdot e^{i\varphi}) = \frac{1}{2} \left(r + \frac{1}{r} \right) \cos \varphi + \frac{i}{2} \left(r - \frac{1}{r} \right) \sin \varphi = u + iv$$

Wir haben
$$\frac{u^2}{a^2} + \frac{v^2}{b^2} = 1, \quad a^2 - b^2 = 1$$

$$\text{mit } a = \frac{1}{2} \left(r + \frac{1}{r} \right) \quad \text{und} \quad b = \frac{1}{2} \left(r - \frac{1}{r} \right)$$

4. Die Strahlen vom Nullpunkt aus werden auf konfokale Hyperbelbögen abgebildet. Mit $c = \cos \varphi$, $s = \sin \varphi$

$$\text{haben wir} \quad \frac{u^2}{c^2} - \frac{v^2}{s^2} = 1, \quad c^2 + s^2 = 1 \quad .$$

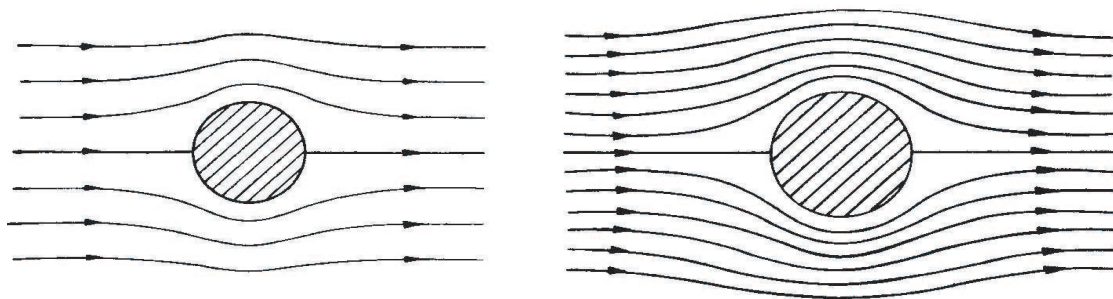
5. Wir studieren das volle Urbild der waagerechten Geraden $\{z : \Im \varphi(z) = v\}$. Der Wert $v = 0$ ist ein Sonderfall. Für $v \neq 0$ zerfällt das Urbild in zwei Zusammenhangskomponenten. Die eine liegt außerhalb des Einheitskreises; die andere geht daraus durch Spiegelung am Einheitskreis hervor. Die Komponente außerhalb des Einheitskreises ist der Graph G einer Funktion $y = Y_v(x)$, nämlich

$$\begin{aligned} G &= \{(x, y) : x^2 + y^2 > 1, \quad \Im \left(x + iy + \frac{1}{x+iy} \right) = 2v\} \quad \text{oder} \\ G &= \{(x, y) : x^2 + y^2 > 1, \quad y - \frac{y}{x^2+y^2} = 2v\} \quad . \end{aligned}$$

Für große x^2 haben wir $Y_v(x) \approx 2v$. Für $v \rightarrow 0$ erhalten wir Kurven, die sich im Bereich $|x| < 1$ an den Einheitskreis anschmiegen.

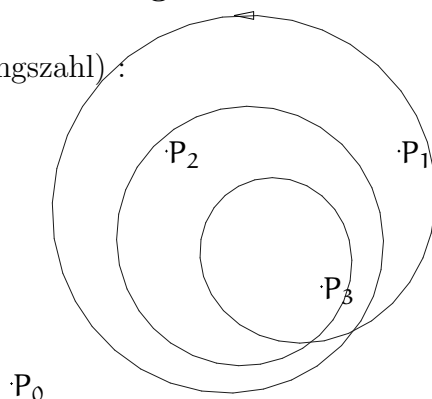
Mit MAPLE findet man ein Bild. Dieses Bild findet man auch in Feynman Bd II, Kapitel 40: „Die Strömung von trockenem Wasser“ (Figur 40.11a).

Ein ähnliches Bild ist Fig. 41.5 im Kapitel „Die Strömung von nassem Wasser“. Die Unterschrift zu 41.5 lautet: Viskose Strömung (niedrige Geschwindigkeiten) um einen kreisförmigen Zylinder.



Aufgaben zu II.6

Aufgabe II.6.1 (Windungszahl):



Vorbereitungen :

A) Das Bild veranschaulicht den Begriff „Windungszahl“ $w(P, \mathfrak{C})$ einer orientierten geschlossenen glatten Kurve \mathfrak{C} um einen Punkt P .

Im Bild : $w(P_j, \mathfrak{C}) = j$ (für $j = 0, 1, 2, 3$).

B) Für verschiedene Kurven \mathfrak{C}_r in der komplexen w -Ebene sollen die Umlaufszahlen $w_r := w(0, \mathfrak{C}_r)$ um den Nullpunkt studiert werden. \mathfrak{C}_r sei das Bild des (im positiven Sinn einfach durchlaufenen) Kreises $K_r := \{z \in \mathbb{C} : |z - i| = r\}$ bzgl. einer polynomialen Abbildung

$$z \mapsto w = p(z)$$

Aufgabe :

a) Benützen Sie MAPLE, um Bilder für den konkreten Fall

$$p(z) := \left(z - \left(\frac{1}{2} - i\right)\right)(z - (1 + i)) = \left(z - \left(-\frac{1}{2} + \frac{i}{2}\right)\right)$$

für Radien r mit $0,4 \leq r \leq 2,5$ in einer Animation darzustellen. Lesen Sie die Umlaufszahlen ab!

- b) Wenn man w_r (als Funktion von $r > 0$) betrachtet, dann erhält man Sprünge an zwei Stellen r^* und r^{**} . Berechnen Sie diese.
- c) Interpretieren Sie, was es für einen Kreis K_r bedeutet, dass sein Bild einen „Scheitelpunkt“ besitzt.
Berechnen Sie die Abstände der Nullstellen von $p'(z)$ zum Kreismittelpunkt und betrachten Sie die Bilder der Kreise mit diesen Radien.

Thesen zum Verhältnis von Arithmetik und Geometrie

Die Wendung von den Polynomen als Rechengrößen (im Sinne der Algebra) zu den Polynomen als Abbildungen (im Sinne der Geometrie) verdient Kommentare. Man hat zurecht gesagt, dass es sich beim berühmten „Fundamentalsatz der Algebra“ überhaupt nicht um einen Satz der Algebra handelt; man sollte lieber vom „Fundamentalsatz des komplexen Zahlensystems“ sprechen. Die Beweise des Satzes brauchen nämlich geometrische (oder genauer: „topologische“) Argumente.

Arithmetik und Geometrie galten in der Tradition der griechischen Philosophie bis hinein ins 19. Jahrhundert als Wissensgebiete, die man aus logischen Gründen voneinander getrennt sehen muss. Die Unterscheidung war allerdings spätestens im 17. Jahrhundert durch die praktische Entwicklung in Frage gestellt worden. Bei der Infinitesimalrechnung von Newton und Leibniz war die Einordnung völlig unklar; diese Theorie handelte von *geometrischen Variablen*, von Größen, die explizit mit geometrischen Kurven assoziiert wurden; die Theorie orientierte sich recht weitgehend an intuitiven geometrischen Begriffen. Das irritierte die Mathematiker des 18. und 19. Jahrhunderts. Leute wie Euler, Lagrange und Cauchy versuchten in ihren Begründungen der Analysis die geometrische Intuition durch Prinzipien der Arithmetik zu ersetzen; konsequenterweise erscheint in ihren Büchern über Infinitesimalrechnung kein einziges geometrisches Diagramm.

Man darf sagen, dass diese Versuche der Arithmetisierung letztlich nicht erfolgreich waren. Bolzano (1815) gründete seinen Beweis der ‘Zwischenwertsatzes’ auf eine Kombination von Vorstellungen über die reellen Zahlen mit Vorstellungen über die reelle Achse. Ein entscheidender Schritt der Annäherung von Geometrie und Arithmetik war dann die von Gauss initiierte Geometrisierung des Systems der komplexen Zahlen. Eine wirklich tragfähige Brücke ergab sich aber erst gegen Ende des 19. Jahrhunderts. Da gelang Dedekind, Cantor und Weierstrass die sog. Arithmetisierung der Analysis. Es gelang, den Irrationalzahlen einen legitimen Platz in der Arithmetik zu verschaffen, und so die Idee der reellen Geraden mit der Arithmetik zu verbinden.

Im 20. Jahrhundert haben sich Auffassungen durchgesetzt, die man in Thesen wie den folgenden zuspitzen kann:

- Geometrie kommt nicht weit ohne Algebra. Auch die Geometrie braucht Formeln und formales Rechnen. Abstrakte Vektorräume und Gruppen sind nicht nur Strukturen der Algebra, sie sind auch das Rückgrat der Geometrie.
- Die Algebra kann das geometrische Denken nicht ersetzen. Die Objekte der Mathematik sind nicht identisch mit ihrer formalen Beschreibung. Das Denken in Bildern und Diagrammen ist ein wesentliches Moment aller Mathematik. Zu jeder Art von Mathematik gehört eine gewisse Anschaulichkeit, die der geometrischen Anschauung analog ist.
- Die zu neuen Erkenntnissen passende ‘Anschauung’ lässt sich nicht immer kurzerhand mit der gewohnten geometrischen Anschauung zur Deckung bringen. Versuche der Reduktion können u. U. zu Fehlvorstellungen führen, wie man z. B. in der Quantenmechanik sehen kann. Der erweiterte Schatz an Rechengrößen sollte dazu nützlich sein, die erweiterte Vorstellungswelt zu festigen.

Historischer Anhang : Irrationalzahlen und Geometrie

(Entlang von D.J. Struik „Abriss der Geschichte der Mathematik“, Berlin 1963, VEB Deutscher Verlag der Wissenschaften.)

In ihrer Suche nach den ewigen Gesetzen des Kosmos studierten die Pythagoreer Geometrie, Arithmetik, Astronomie und Musik (das „Quadrivium“ wurde dieser Fächerkanon in der mittelalterlichen Universität genannt). Anhand von inkommensurablen Strecken entdeckten die Pythagoreer das Irrationale. Das Verhältnis von Seite und Diagonale eines Quadrats kann nicht durch „Zahlen“ ausgedrückt werden - d.h. durch positive ganze Zahlen und ihre Verhältnisse, die in der Pythagoreischen Arithmetik allein zugelassenen Begriffe.

Dabei blieben die Griechen nicht stehen. Das Fünfte Buch von Euklid stellt die Theorie inkommensurabler Größen von Eudoxus in rein geometrischer Form dar. Die Theorie der Proportionen von Eudoxus beseitigte die arithmetische Theorie der Pythagoreer, die nur für kommensurable Größen gültig war. Typisch ist die Definition 5 von Buch V der Euklidischen Elemente:

„Man sagt, dass Größen in demselben Verhältnis stehen, die erste zur zweiten wie die dritte zur vierten, wenn bei beliebiger Vervielfältigung die Gleichvielfachen der ersten und der dritten den Gleichvielfachen der zweiten und vierten gegenüber, paarweise entsprechend genommen, entweder zugleich größer oder zugleich kleiner sind.“

Euklid hat der neuen Theorie der inkommensurablen Größen ein besonderes Gewicht beigemessen. Im zehnten Buch, welches meist als das schwierigste angesehen wird, entwickelt Euklid eine geometrische Klassifizierung quadratischer Irrationalitäten und von Quadratwurzeln aus solchen. Wir würden sagen, dass es um Zahlen von der Form $\sqrt{a + \sqrt{b}}$ geht. Bei den Griechen wurden die Schwierigkeiten mit den Irrationalitäten nicht, wie im Orient oder in Europa im Zeitalter der Renaissance, durch eine Verallgemeinerung des Zahlbegriffs aufgelöst, sondern (unter Ablehnung einer auf Zahlen beruhenden Theorie) durch den Versuch einer geometrischen Synthese.

Die Harmonie zwischen Arithmetik und Geometrie erschien den Griechen auch noch durch eine andere Schwierigkeit gestört, die aus den Erörterungen über die Realität der Veränderungen entstanden war. Zeno von Elea (um 450 v.u.Z.) bezweifelte die Zulässigkeit von einigen älteren intuitiven Begriffen bezüglich des unendlich Kleinen und des unendlich Großen. Die vier Paradoxien des Zeno erregten großes Aufsehen; und es wird bis heute über sie geredet (besonders über „Achilles und die Schildkröte“ und den „fliegenden Pfeil“). Die Überlegungen von Zeno machten klar, dass eine endliche Strecke in unendlich viele kleine Strecken zerlegt werden kann, von denen jede eine endliche Länge besitzt. Sie zeigten außerdem, dass es schwierig zu erklären ist, was man eigentlich damit meint, wenn man sagt, dass eine Linie aus Punkten zusammengesetzt ist.

Die Trennung von Arithmetik und Geometrie bestimmte über 2000 Jahre das Bild von Mathematik. Formeln gab es nur in der Arithmetik; man hatte keine algebraischen Zeichen für die geometrischen Größen. Als im 15. Jahrhundert die Trigonometrie als eine von der Astronomie unabhängige Wissenschaft entstand (ein wichtiger Pionier war Regiomontanus, 1464), wurden noch alle Sätze (wie z.B. der Sinussatz für sphärische Dreiecke) in Worten ausgedrückt.

Auf der anderen Seite hat sich im späten Mittelalter die Einstellung zu den Zahlen gewandelt. Mit der Ausdehnung des Handels hat sich das Interesse an Mathematik langsam in den Städten des Nordens ausgebreitet. Zunächst war das Interesse hauptsächlich ein praktisches. Führer in der Liebe zur praktischen Mathematik waren die Rechenmeister, die auch Buchführung und Navigation lehrten. Ihnen stand nur selten ein Mann mit Universitätsbildung zur Seite, der sie aufgrund seines Studiums der Astronomie hätte befähigen können, die Bedeutung von verbesserten Rechenverfahren zu verstehen.

Die theoretische Mathematik war zwar während des Mittelalters nicht völlig untergegangen; sie wurde aber nicht bei den Praktikern, sondern bei den scholastischen Philosophen kultiviert. Hier führte das Studium von Plato und Aristoteles im Verein mit Grübeleien über die Natur der Gottheit zu spitzfindigen Spekulationen über die Natur der Bewegung, des Kontinuums und der Unendlichkeit. Das änderte sich in der Renaissance.

Mittelpunkte eines neuen geistigen Lebens waren im 15. Jahrhundert die italienischen Städte und in enger Verbindung die mitteleuropäischen Städte Nürnberg, Wien und Prag. Der Fall von Konstantinopel im Jahre 1453, der das Byzantinische Reich beendete, führte viele griechische Gelehrte in die Städte des Westens. Das Interesse an den griechischen Originaltexten wuchs. Noch blieben aber die Klassiker das Nonplusultra der Wissenschaft. Es bedeutete eine ungeheure und freudige Überraschung, als italienische Mathematiker am Anfang des 16. Jahrhunderts tatsächlich zeigten, dass es möglich war, eine neue mathematische Theorie zu entwickeln, die bei den Alten und den Arabern nicht vorhanden war.

Diese Theorie, die zur allgemeinen Lösung der kubischen Gleichung führte, wurde von Scipio de Ferro und seinen Schülern an der Universität von Bologna entdeckt. Diese Universität war damals eine der größten und berühmtesten in Europa. Allein ihre astronomische Fakultät hatte zeitweise sechzehn Lektoren. Aus allen Teilen Europas strömten die Studenten herbei, um die Vorlesungen zu hören - und zu den öffentlichen Disputationen, die ebenfalls die Aufmerksamkeit großer, sportlich eingestellter Hörermassen auf sich lenkten. Unter diesen Studenten befanden sich zeitweise Pacioli, Albrecht Dürer und Kopernikus. Charakteristisch für dieses neue Zeitalter war der Wunsch, nicht nur das klassische Erbe aufzunehmen, sondern zugleich Neues zu schaffen und über die von den Klassikern abgesteckten Grenzen hinaus vorzudringen. Die Erfindung der Buchdruckerkunst und die Entdeckung Amerikas waren Beispiele für solche Möglichkeiten.

War es möglich, in der Mathematik Neues zu finden? Griechen und Orientalen hatten ihren Scharfsinn an der Lösung der kubischen Gleichung versucht, aber sie hatten nur einige Spezialfälle numerisch lösen können. Die Mathematiker in Bologna versuchten nun, die allgemeine Lösung zu finden. Die kubischen Gleichungen konnten auf drei Fälle reduziert werden

$$x^3 + px = q \quad , \quad x^3 = px + q \quad , \quad x^3 + q = px \quad ,$$

wobei p und q positive Zahlen waren. Die Lösung wird heute als Cardanische Formel bezeichnet; sie hat im Falle $x^3 + px = q$ die Form

$$x = \sqrt[3]{\sqrt{\frac{p^3}{27} + \frac{q^2}{4}} + \frac{q}{2}} - \sqrt[3]{\sqrt{\frac{p^3}{27} + \frac{q^2}{4}} - \frac{q}{2}} \quad .$$

Es gab heftige Auseinandersetzung um Prioritäten (siehe Struik, Seite 98) .

Wie man sieht, führte diese Lösung Größen der Form $\sqrt[3]{a + \sqrt{b}}$ ein, die von den Euklidischen $\sqrt{a + \sqrt{b}}$ verschieden waren. Cardano betrachtete auch negative Zahlen, die er „fiktiv“ nannte, aber er wusste nichts anzufangen mit dem sog. „casus irreducibilis“ der kubischen Gleichung, in dem drei reelle Lösungen vorhanden sind, die als Summe oder Differenz von solchen Zahlen erschienen, die wir heute komplexe Zahlen nennen.

Die Schwierigkeit wurde von dem letzten der großen Bologneser Mathematiker des 16. Jahrhunderts Raffael Bombelli, behoben, dessen „Algebra“ 1572 erschien. Er schrieb $3i$ als $\sqrt{0-9}$ (wörtlich so: $R[0m.9]$, R für radix, m für meno). Der Kalkül ergab Formeln wie z.B.

$$\sqrt[3]{52 + \sqrt{0-2209}} = 4 + \sqrt{0-1} \quad .$$

Der Kalkül erlaubte den „casus irreducibilis“ zu behandeln. Bombellis Buch wurde viel gelesen, insbesondere auch von Leibniz und Euler. Von da an verloren die komplexen Zahlen etwas von ihrem übernatürlichen Charakter, obwohl sich ihre uneingeschränkte Anerkennung als Zahlen erst im 19. Jahrhundert durchsetzte.

In seiner Arbeit von 1831 beseitigte Gauss ein für allemal das Geheimnis, das die komplexen Zahlen immer noch umgeben hatte, durch ihre Darstellung als Punkte in der Ebene. Hier wird also die Geometrie herangezogen, um die Zahlen zu verstehen. Parallel dazu hat man auch verstanden, warum das Wurzelziehen nicht geeignet ist, um polynomiale Gleichungen vom Grad > 4 zu lösen. (Bei den einschlägigen Arbeiten von Abel und Galois liegen übrigens auch wesentliche Anstöße für die Theorie der endlichen Gruppen.)

Im Laufe des 19. Jahrhunderts wurde klar, dass man für einen strengen Beweis des Fundamentalsatzes der Algebra die Vollständigkeit der komplexen Ebene benötigt. Man braucht eine Theorie der Irrationalzahlen. Die heutige Theorie der Irrationalzahlen, wie sie von Dedekind und Weierstraß gegen Ende des 19. Jahrhunderts entwickelt wurde, folgt fast wörtlich dem Gedankengang von Eudoxus. In zwei kleinen Bändchen „Stetigkeit und irrationale Zahlen“ (1872) und „Was sind und was sollen die Zahlen“ (1882), vollendete Dedekind das für die moderne Mathematik, was Eudoxus für die griechische Mathematik geleistet hatte. Es besteht eine große Ähnlichkeit zwischen dem „Dedekind’schen Schnitt“ und der alten Theorie von Eudoxus. Cantor und Weierstraß gaben arithmetische Definitionen der irrationalen Zahlen, die auf ähnlichen Betrachtungen beruhen. Cantor ging dann mit seiner Mengenlehre noch einen wesentlichen Schritt weiter in der „Arithmetisierung der Mathematik“.

Die Theorien der Irrationalzahlen haben durch die Verwendung moderner arithmetischer Methoden weitere Perspektiven eröffnet. Die zunächst umstrittenen Ideen zum Aktual-Unendlichen von G. Cantor haben schließlich doch vollständige Anerkennung gefunden, als die außerordentliche Bedeutung dieser Denkweise für die Begründung der Theorie der reellen Funktionen und der Topologie immer deutlicher wurde.

Endgültig klar wurde die Bedeutung der abstrakten Konstruktionen, nachdem Lebesgue im Jahre 1901 die Mengenlehre durch seine Maß- und Integrationstheorie bereichert hatte.

III Metrik, Norm, Konvexität

Themenübersicht III

Eine wichtige Methode, einer abstrakten Menge S Struktur zu geben, besteht darin, dass man den Punktpaaren einen Abstand $d(P, Q)$ zuordnet. Die Dreiecksungleichung bietet einen Einstieg in die Begriffswelt der metrischen Räume.

Auf reellen und komplexen affinen Räumen interessiert man sich vor allem für solche Abstandsfunktionen, die durch eine Längenmessung (Norm) im Raum der Verschiebungsvektoren gegeben sind: $d(P, Q) = \|\overrightarrow{PQ}\|$.

Auf diesem Weg kommen wir zum Begriff der Konvexität und zu Themen der Dualität in reellaffinen Vektorräumen.

Bei allen Überlegungen spielt die Ordnung auf der reellen Achse \mathbb{R} eine wichtige Rolle. Wir werden wesentlich ausnützen, dass jede nach oben beschränkte Teilmenge von \mathbb{R} in \mathbb{R} eine kleinste obere Schranke („Supremum“) besitzt. Eine eingehende Analyse dieser Vollständigkeitseigenschaft von (\mathbb{R}, \leq) überlassen wir aber der Vorlesung über Analysis.

III.1 Dreiecksungleichung und Subadditivität

Semimetriken. Beispiele: Abstand auf einem zusammenhängenden Graphen. Hamming-Distanz auf $\{0, 1\}^N$. Subadditive Funktionen auf dem positiven Oktanten. Eine monotone subadditive Funktion macht aus einem Tupel von Semimetriken eine Semimetrik. Vorformen der Hölder’schen und der Minkowski’schen Ungleichung.

Metrische Räume für die Analysis. Konvergente Folgen in einem metrischen Raum, abgeschlossene und offene Mengen. Stetige und unterhalbstetige Funktionen auf einem metrischen Raum. Eine Funktion ist genau dann unterhalbstetig, wenn ihr Epigraph abgeschlossen ist.

III.2 Normierte Vektorräume und stetige Linearformen.

Seminormen auf einem reellen oder komplexen Vektorraum. Die 2-Norm auf dem Raum der trigonometrischen Polynome. Die Schwarz’sche Ungleichung. Stetige Linearformen auf einem normierten Vektorraum; die duale Norm. Die p -Norm auf dem Raum der finiten Folgen. Die q -Norm ist die duale Norm.

III.3 Konvexe Mengen und konvexe Funktionen.

Wir beschränken uns hier wieder auf den endlichdimensionalen Fall. Die Distanzfunktion zu einer konvexen Nullumgebung. Stützfunktionen. Fortsetzung einer Linearform unter einer positiv homogenen subadditiven Funktion. Die Trennung disjunkter konvexer Mengen. Unterhalbstetige konvexe Funktionen. Beispiele aus der elementaren Analysis. Stützhyperebenen. Die Legendre-Transformation. Duale konvexe Kegel.

Nachtrag: Die Jensen’sche Ungleichung.

Rückblick: Die bisher ins Auge gefassten Typen von Funktionen im Überblick.

III Metrik, Norm, Konvexität

III.1 Dreiecksungleichung und Subadditivität

Definition

Eine abstrakte Punktmenge S wird dadurch zu einem **metrischen Raum**, dass man eine positive Funktion $d(\cdot, \cdot)$ auf $S \times S$ auszeichnet, welche den Forderungen an eine Metrik genügt.

Definition (Metrik)

Eine Funktion $d(\cdot, \cdot)$ heißt eine **Semi-Metrik**, wenn gilt

- (i) $d(P, Q) \geq 0$ für alle P, Q , $d(P, P) = 0$ für alle P .
- (ii) $d(P, Q) = d(Q, P)$ („Symmetrie“)
- (iii) $d(P, R) \leq d(P, Q) + d(Q, R)$ („Dreiecksungleichung“)

Die Semi-Metrik ist eine Metrik, wenn gilt

- (iv) $d(P, Q) = 0 \implies P = Q$.

Beispiele :

- 1) Die reelle Achse wird im Schulunterricht von vorneherein als ein metrischer Raum eingeführt. $d(x, y) = |y - x|$.
- 2) Die komplexe Ebene wird meistens mit der Metrik $d(z, w) = |w - z| = \sqrt{(w - z)(\bar{w} - \bar{z})}$ ausgestattet. Manchmal ist aber der Abstand auf der Riemann'schen Zahlkugel vorteilhafter; denn für diese Metrik bietet sich eine Erweiterung zu einer Metrik auf \mathbb{C} an.
- 3) Der Anschauungsraum wird meistens mit der euklidischen Metrik ausgestattet. Wir werden aber auch andere Metriken schätzen lernen.
- 4) Man stelle sich ein Straßennetz vor, wo jeder Straße k eine nichtnegative Zahl $e(k)$ zugeordnet ist. In der Fachsprache: S sei die Scheitelmengende eines ungerichteten Graphen; jeder Kante k sei eine Zahl $e(k) \geq 0$ zugeordnet. Nehmen wir an, dass der Graph zusammenhängend ist. Für jeden Weg von P nach Q betrachten wir die Summe der $e(\cdot)$ -Werte auf den Kanten. $d(P, Q)$ sei das Infimum. $d(\cdot, \cdot)$ ist dann eine Semi-Metrik auf S . Wenn es keine Straßen gibt, die ohne Kosten zu begehen sind, dann ist $d(\cdot, \cdot)$ eine Metrik.
- 5) Sei $S = \{0, 1\}^N$ die Menge aller Null-Eins-Folgen der Länge N . Eine beliebige Metrik ist die Hamming-Distanz. $d(P, Q)$ gibt an, in wievielen Punkten die beiden Null-Eins-Folgen verschieden sind. Man deutet S als die Menge der Ecken eines N -dimensionalen Würfels, dessen Kanten k mit $e(k) = 1$ belegt sind. Der Hamming-Abstand ist dann die Metrik im Sinne des Beispiels 4.

Gewisse Sprechweisen, die für den Anschauungsraum geläufig sind, kann man mit Vorteil auf allgemeine metrische Räume übertragen. Wir beginnen mit dem Begriff der Kugel (auch Vollkugel, im Englischen ‘ball’). Für $\tilde{P} \in (S, \text{dist}(\cdot, \cdot))$ und $r > 0$ heissen die Mengen

$$B_r(\tilde{P}) = \{Q : d(Q, \tilde{P}) < r\} \quad \text{und} \quad B_{\leq r}(\tilde{P}) = \{Q : d(Q, \tilde{P}) \leq r\}$$

die (offene) bzw. die abgeschlossene Kugel vom Radius r mit dem Mittelpunkt \tilde{P} .

Was es allgemein mit den offenen und abgeschlossenen Mengen in einem metrischen Raum auf sich hat, werden wir später sehen. Hier wollen wir zunächst einmal die Kugeln im Raum der Null-Eins-Folgen diskutieren.

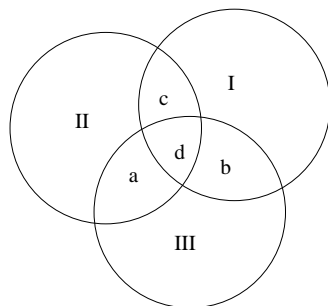
Ein fehlerkorrigierender Code

Jeder Punkt $x^* \in \{0, 1\}^n$ (Beispiel 5) hat genau n unmittelbare Nachbarn und $\binom{n}{2}$ Nachbarn im Abstand $= 2$. Die abgeschlossene Kugel $B(x^*, \leq 1)$ mit dem Mittelpunkt x^* und dem Radius 1 enthält offenbar genau $n + 1$ Punkte; und die Anzahl der Punkte in der abgeschlossenen Kugel mit dem Radius 2 ist $|B_{\leq 2}(x^*)| = 1 + n + \binom{n}{2}$.

Ein merkwürdiger Raum ist der Raum $S = \{0, 1\}^7$. In diesem Raum mit $2^7 = 2^4 \cdot 2^3$ Punkten gibt es $2^4 = 16$ paarweise disjunkte Kugeln mit dem Radius 1. Man benützt dieses Faktum zur Konstruktion eines „fehlerkorrigierenden“ Code.

Fehlererdeckende und fehlerkorrigierende Code dienen dazu, die Übertragung von Bit-Folgen durch gestörte Kanäle sicherer zu machen. Man sendet z.B. statt jedes Quadrupels von Bits ein besonderes 7-Tupel, wo die zusätzlichen drei Bits als Kontrollbits fungieren. Das folgende Diagramm veranschaulicht einen beliebigen Code dieser Art (Er heißt Hamming’s (7,4)-Code). Der Empfänger ist in der Lage, das gemeinte Quadrupel $abcd$ auch dann zu rekonstruieren, wenn die Übertragung (höchstens) eines der 7 gesendeten Bits $abcd$ I II III verändert hat. — Man beachte: Wenn mit allzugroßer Wahrscheinlichkeit mehr als eines der sieben Bits falsch übertragen wird, dann ist Hamming’s (7,4)-Code nicht hilfreich.

Das Diagramm stellt das „Codebuch“ auf übersichtliche Weise dar.



Die Kontrollbits I, II und III werden so gewählt, daß die Summe in jedem Kreis gerade ist. Beispielsweise wird das Quadrupel $(a b c d) = (1 0 0 1)$ in das 7-Tupel $(a b c d, I, II, III) = 1 0 0 1 1 1 0$ kodiert.

Wenn nun der Empfänger ein 7-Tupel empfängt, welches der Forderung nicht genügt, dann kann er es durch Abänderung in genau einer Position passend machen.

Zum Beweis müssen wir Fälle unterscheiden: Wenn ein empfangenes 7-Tupel alle drei Gleichungen verletzt, dann ist \mathbf{d} zu korrigieren. Wenn genau eine der Gleichungen verletzt ist, dann ist das entsprechende Kontrollbit zu korrigieren. Wenn genau zwei Gleichungen verletzt sind, ist a bzw. b bzw. c zu korrigieren. Das Decodieren ist also genauso einfach wie das Codieren.

Konstruktionen für Semimetriken

Eine interessante Konstruktion, die aus einer (Semi-)Metrik eine weitere (Semi-)Metrik macht, ist die folgende: Wenn $\mathbf{d}(\cdot, \cdot)$ eine (Semi-)Metrik auf der Menge S ist, dann ist auch

$$e(\cdot, \cdot) = \frac{\mathbf{d}(\cdot, \cdot)}{1 + \mathbf{d}(\cdot, \cdot)}$$

eine (Semi-)Metrik. Wir werden das unten beweisen. Man bemerke, dass die Kugeln für $e(\cdot, \cdot)$ dieselben sind wie die Kugeln für $\mathbf{d}(\cdot, \cdot)$.

Wenn $\mathbf{d}_1(\cdot, \cdot)$ und $\mathbf{d}_2(\cdot, \cdot)$ Semimetriken sind, dann ist offenbar auch die Summe $\mathbf{d}_1 + \mathbf{d}_2$ eine Semimetrik, möglicherweise sogar eine Metrik. Es sei z. B. $\tilde{\mathbf{d}}_1$ eine Semimetrik auf der Menge S_1 und $\tilde{\mathbf{d}}_2$ eine Semimetrik auf der Menge S_2 . Wir können daraus eine Semimetrik auf dem cartesischen Produkt $S = S_1 \times S_2$ gewinnen, indem wir konstruieren

$$\mathbf{d}(x, y) = \mathbf{d}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \tilde{\mathbf{d}}_1(x_1, y_1) + \tilde{\mathbf{d}}_2(x_2, y_2).$$

Die beiden Summanden können als Semimetriken auf dem Produktraum aufgefasst werden. Wenn $\tilde{\mathbf{d}}_1$ und $\tilde{\mathbf{d}}_2$ Metriken sind, dann ist $\mathbf{d}(\cdot, \cdot)$ offenbar eine Metrik auf dem Produktraum.

Eine weitere (Semi-)Metrik ist, wie wir sehen werden

$$e(x, y) = e\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \sqrt{\tilde{\mathbf{d}}_1^2(x_1, y_1) + \tilde{\mathbf{d}}_2^2(x_2, y_2)}.$$

In der Tat erhalten wir für jedes $p \geq 1$ eine (Semi-)Metrik

$$e_p(x, y) = e_p\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (\tilde{\mathbf{d}}_1^p(x_1, y_1) + \tilde{\mathbf{d}}_2^p(x_2, y_2))^{1/p}.$$

Um das zu beweisen, holen wir weiter aus.

Eine nichtnegative Funktion $F(s) = F(s_1, \dots, s_n)$ auf dem 'positiven Oktanten' \mathbb{R}_+^n nennen wir **monoton**, wenn gilt

$$s \leq t \quad (\text{in dem Sinne } s_j \leq t_j \text{ für alle } j) \implies F(s) \leq F(t),$$

wir nennen sie **subadditiv**, wenn gilt

$$F(s + t) \leq F(s) + F(t).$$

Satz

Seien $d_1(\cdot, \cdot), \dots, d_n(\cdot, \cdot)$ Semimetriken auf S und $F(\cdot)$ eine monotone subadditive Funktion auf \mathbb{R}_+^n mit $F(0) = 0$. Dann ist

$$e(\cdot, \cdot) = F(d_1(\cdot, \cdot), \dots, d_n(\cdot, \cdot))$$

eine Semimetrik.

Beweis :

Die Symmetrie von $e(\cdot, \cdot)$ ist trivial; wir müssen die Dreiecksungleichung nachweisen. Und in der Tat ergeben sich aus der Monotonie und der Subadditivität die Ungleichungen

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

Beispiele :

- Die Funktion $F(r) = \frac{r}{1+r} = 1 - \frac{1}{1+r}$ ist monoton auf \mathbb{R}_+ ; und für $s, t > 0$ gilt

$$F(r+s) = \frac{r+s}{1+r+s} = \frac{r}{1+r+s} + \frac{s}{1+r+s} \leq F(r) + F(s).$$

- Eine weitere monotone subadditive Funktion auf \mathbb{R}_+ ist $F(r) = r \wedge 1 = \min\{r, 1\}$. Die Kugeln zur beschränkten Metrik $e(\cdot, \cdot) = F(d(\cdot, \cdot))$ sind hier für kleine Radien dieselben wie die Kugeln zur Ausgangsmetrik $d(\cdot, \cdot)$.
- Eine wichtige monotone subadditive Funktion ist die ‘Quadratwurzel aus der Quadratsumme’

$$F(s_1, \dots, s_n) = \sqrt{s_1^2 + \dots + s_n^2} \text{ auf } \mathbb{R}_+^n.$$

Als Vorbereitung für den Beweis der Subadditivität beweisen wir die Ungleichung

$$\sum s_j \cdot t_j \leq \sqrt{\sum s_j^2} \cdot \sqrt{\sum t_j^2} \quad \text{für } s, t \in \mathbb{R}_+^n.$$

Wenn wir alle s_j oder alle t_j mit derselben positiven Zahl multiplizieren, dann multiplizieren sich beide Seiten der behaupteten Ungleichung mit eben dieser Zahl. Es genügt daher, die Ungleichung im Spezialfall $\sum s_j^2 = 1 = \sum t_j^2$ zu beweisen.

Bekanntlich ist das geometrische Mittel zweier positiver Zahlen nicht größer als das arithmetische Mittel: $\sqrt{a \cdot b} \leq \frac{1}{2}(a+b)$. Wir benützen diese Ungleichung für $a_j = s_j^2$, $b_j = t_j^2$ und erhalten $s_j \cdot t_j \leq \frac{1}{2}s_j^2 + \frac{1}{2}t_j^2$. Summation über alle j ergibt die Behauptung

$$\sum s_j^2 = 1, \quad \sum t_j^2 = 1 \quad \implies \sum s_j \cdot t_j \leq 1.$$

Die Subadditivität der Wurzel aus der Quadratsumme ist eine einfache Konsequenz:

$$\begin{aligned} F^2(s+t) &= \sum (s_j + t_j)^2 = \\ &= \sum s_j^2 + \sum t_j^2 + 2 \sum s_j \cdot t_j \leq \sum s_j^2 + \sum t_j^2 + 2 \cdot \sqrt{\sum s_j^2} \cdot \sqrt{\sum t_j^2} = \\ &= (F(s) + F(t))^2. \end{aligned}$$

Interessant ist auch die

Verallgemeinerung Sei $p \geq 1$ und $\frac{1}{p} + \frac{1}{q} = 1$, und $F_p(t) = (\sum t_j^p)^{1/p}$ für $t \in \mathbb{R}_+^n$. Dann gilt für $s, t \in \mathbb{R}_+^n$

$$\begin{aligned} \sum s_j \cdot t_j &\leq F_p(s) \cdot F_q(t) = \left(\sum s_j^p\right)^{1/p} \cdot \left(\sum t_j^q\right)^{1/q} \\ F_p(s+t) &\leq F_p(s) + F_p(t) \end{aligned}$$

Beweisskizze

1. Wenn $p \geq 1$ und $\frac{1}{p} + \frac{1}{q} = 1$, dann gilt wegen der Konkavität der Logarithmusfunktion für alle $a, b \in \mathbb{R}_+$

$$\ln\left(\frac{1}{p}a + \frac{1}{q}b\right) \geq \frac{1}{p} \ln a + \frac{1}{q} \ln b, \quad a^{1/p} \cdot b^{1/q} \leq \frac{1}{p} \cdot a + \frac{1}{q} \cdot b.$$

Gleichheit gilt genau dann, wenn $a = b$.

2. Wir wenden die Ungleichung an auf $a_j = s_j^p$, $b_j = t_j^q$. Summation über j liefert $\sum s_j \cdot t_j \leq 1$ für n -Tupel s, t mit $F_p(s) = 1 = F_q(t)$, und daher $\sum s_j \cdot t_j \leq F_p(s) \cdot F_q(t)$ für alle s, t .
3. Zu jedem s gibt es ein t , sodass Gleichheit gilt ($t_j^q = s_j^p$). Somit haben wir

$$\begin{aligned} \sup \left\{ \sum s_j \cdot t_j : F_q(t) \leq 1 \right\} &= F_p(s) \quad \text{für jedes } s \in \mathbb{R}_+^n \\ F_p(s' + s'') &= \sup \left\{ \sum (s'_j + s''_j) \cdot t_j : F_q(t) \leq 1 \right\} \leq \\ &\leq \sup \left\{ \sum s'_j \cdot t_j : F_q(t) \leq 1 \right\} + \sup \left\{ \sum s''_j \cdot t_j : F_q(t) \leq 1 \right\} \\ &= F_p(s') + F_p(s''). \end{aligned}$$

Anhang: Metrische Räume für die Analysis, Sprechweisen

Viele Begriffe, die man in der elementaren Analysis zunächst für reelle Folgen bzw. für Teilmengen der reellen Achse einführt, besitzen eine direkte Entsprechung im Kontext der metrischen Räume. Es sei also $(S, \text{dist}(\cdot, \cdot))$ ein beliebiger metrischer Raum.

1. (Konvergente Folgen) Man sagt von einer Punktfolge $(P_n)_n$ in S , dass sie gegen \tilde{P} **konvergiert** und notiert $\lim_{n \rightarrow \infty} P_n = \tilde{P}$, wenn gilt

$$\forall \varepsilon > 0 \quad \exists N \quad \forall n \geq N \quad \text{dist}(P_n, \tilde{P}) < \varepsilon.$$

Entsprechend definiert man den Begriff der Cauchy-Folge. Ein triviales Beispiel einer konvergenten Folge ist eine Folge, die schließlich konstant ist. In manchen metrischen Räumen gibt es keine anderen konvergenten Folgen.

2. (Beschränkte Mengen) Für die Menge $M \subseteq S$ heisst $\sup\{\text{dist}(P', P'') : P', P'' \in M\}$ der Durchmesser; die Menge heisst **beschränkt**, wenn der Durchmesser endlich ist.

3. (Abgeschlossene Mengen) Eine Teilmenge A eines metrischen Raums $(S, \text{dist}(\cdot, \cdot))$ heißt eine (in S) **abgeschlossene Menge**, wenn gilt: Immer wenn eine A -Folge (in S) konvergiert, dann gehört der Limespunkt zu A .
4. (Offene Mengen) Eine Teilmenge U eines metrischen Raums $(S, \text{dist}(\cdot, \cdot))$ heißt eine (in S) **offene Menge**, wenn gilt

$$\forall P_0 \in U \quad \exists r > 0 : \{P : \text{dist}(P_0, P) < r\} \subseteq U .$$

Satz: Eine Menge U ist genau dann (in S) offen, wenn ihr Komplement $A = S \setminus U$ (in S) abgeschlossen ist.

Beweisskizze: Die Punkte im Komplement der abgeschlossenen Menge A sind diejenigen, die man nicht als Limespunkte einer A -Folge erhalten kann. Das sind diejenigen Punkte Q , für die es in einer genügend kleinen Kugel $B_r(Q)$ keine Punkte von A gibt.

Hinweis: Viele Lehrbücher benützen den Satz zur Definition des Begriffs der abgeschlossenen Menge. Diese Lehrbücher sprechen zuerst über die Klasse \mathcal{U} aller offenen Mengen und sagen dann: Die abgeschlossenen Mengen sind die Komplemente der offenen Mengen.

Die Anfänger müssen hier vor einem Denkfehler gewarnt werden: Es ist nicht so, dass eine Menge M , die nicht offen ist, notwendigerweise abgeschlossen ist. Typischerweise ist eine Teilmenge M eines metrischen Raums $(S, \text{dist}(\cdot, \cdot))$ weder offen noch abgeschlossen. Die leere Menge und der Gesamtraum S sind sowohl offen als auch abgeschlossen.

5. (Abgeschlossene Hülle und offener Kern) Der Durchschnitt von (beliebig vielen) abgeschlossenen Mengen ist abgeschlossen; die Vereinigung von (beliebig vielen) offenen Mengen ist offen.

Die grösste offene Teilmenge von M heisst das Innere von M oder der **offene Kern**; die Elemente dieser Teilmenge heissen die **inneren Punkte** von M . Wenn ein Punkt \tilde{P} im Inneren der Menge M liegt, dann sagt man, M sei eine **Umgebung** von \tilde{P} .

Die kleinste abgeschlossene Obermenge von M heisst die **abgeschlossene Hülle** von M ; ihre Punkte heissen die **Berührungspunkte** von M .

Die Punkte, die zur abgeschlossenen Hülle von M , aber nicht zum offenen Kern gehören, heissen die Randpunkte von M . Sie sind dadurch gekennzeichnet, dass in einer beliebigen Umgebung sowohl Punkte von M liegen, als auch Punkte, die nicht zu M gehören. Die Menge aller Randpunkte heisst der topologische Rand von M .

6. Es sei M eine beliebige Menge. Der Abstand eines Punktes P von M ist die Zahl $\text{dist}(P, M) = \inf\{\text{dist}(P, Q) : Q \in M\}$. Für $\varepsilon > 0$ heisst die Menge $M^\varepsilon = \{P : \text{dist}(P, M) < \varepsilon\}$ die ε -Umgebung von M . Es handelt sich offenbar um eine offene Obermenge von M .
7. Wenn M und N beschränkte Mengen sind, dann nennt man die Zahl $\inf\{\varepsilon : M \subseteq N^\varepsilon, N \subseteq M^\varepsilon\}$ den Hausdorff-Abstand der Mengen M und N . Der Hausdorff-Abstand ist offenbar eine Semi-Metrik auf der Menge aller beschränkten Teilmengen des metrischen Raums $(S, \text{dist}(\cdot, \cdot))$.

8. Eine reellwertige Funktion $f(\cdot)$ auf einem metrischen Raum heisst stetig im Punkt \tilde{P} , wenn $\{P : |f(P) - f(\tilde{P})| < \varepsilon\}$ eine Umgebung von \tilde{P} ist für jedes $\varepsilon > 0$, wenn also

$$\forall \varepsilon > 0 \exists \delta > 0 \quad \{P : |f(P) - f(\tilde{P})| < \varepsilon\} \supseteq B_\delta(\tilde{P}).$$

oder etwas anders ausgedrückt

$$f(\cdot) \text{ stetig in } \tilde{P} \iff \forall \varepsilon > 0 \exists \delta > 0 \forall P \quad \text{dist}(P, \tilde{P}) < \delta \Rightarrow |f(P) - f(\tilde{P})| < \varepsilon.$$

9. (Unterhalbstetigkeit) Eine Funktion $f(\cdot)$ auf einem metrischen Raum mit Werten in $\mathbb{R} \cup \{+\infty\}$ heisst unterhalbstetig im Punkt \tilde{P} , wenn $\{P : f(P) > f(\tilde{P}) - \varepsilon\}$ eine Umgebung von \tilde{P} ist für jedes $\varepsilon > 0$. (Unterhalbstetigkeit in einem Punkt \tilde{P} mit $f(\tilde{P}) = +\infty$ bedeutet, dass für jede gegen \tilde{P} konvergierende Folge die Funktionswerte gegen $+\infty$ konvergieren.) Man bemerke: Wenn $f_1(\cdot), f_2(\cdot), \dots$ im Punkt \tilde{P} unterhalbstetig sind, dann ist auch das punktweise Supremum $f = \bigvee f_j$ im Punkt \tilde{P} unterhalbstetig.
10. Eine unterhalbstetige Funktion auf einem metrischen Raum $(S, \text{dist}(\cdot, \cdot))$ ist eine $\mathbb{R} \cup \{+\infty\}$ -wertige Funktion, die in jedem Punkt unterhalbstetig ist. Das ist genau dann der Fall, wenn $\{P : f(P) > b\}$ offen ist für jedes $b \in \mathbb{R}$.
11. Zu einer Funktion $f(\cdot)$ mit Werten in $\mathbb{R} \cup \{+\infty\}$ auf dem Raum S definiert man den ‘Funktionsgraphen’ Γ_f und den ‘Epigraphen’ $\Gamma_{\geq f}$. Dabei handelt es sich um Teilmengen des Produktraums $\mathbb{R} \times S$, und zwar

$$\Gamma_f = \{(f(P), P) : P \in S\}, \quad \Gamma_{\geq f} = \{(y, P) : P \in S, y \geq f(P)\}.$$

Satz

Der Epigraph einer Funktion f auf dem metrischen Raum $(S, \text{dist}(\cdot, \cdot))$ ist genau dann abgeschlossen, wenn f unterhalbstetig ist.

Beweisskizze: f ist genau dann unterhalbstetig, wenn für jede konvergente Folge $(P_n)_n$ in S gilt

$$\lim_n P_n = \tilde{P} \implies \liminf f(P_n) \geq f(\tilde{P}).$$

Es sei $(y_n, P_n)_n$ sei eine konvergente Folge im Epigraphen. Wenn f unterhalbstetig ist, dann liegt der Grenzwert (\tilde{y}, \tilde{P}) im Epigraphen, denn

$$\tilde{y} = \lim y_n \geq \liminf f(P_n) \geq f(\tilde{P}).$$

Der Epigraph ist also abgeschlossen.

Wenn umgekehrt der Epigraph abgeschlossen ist, dann gilt für jede gegen \tilde{P} konvergierende Folge (P_n) , dass für jede Teilfolge, entlang welcher $f(P_n)$ gegen einen Wert \tilde{y} konvergiert, $\tilde{y} \geq f(\tilde{P})$. Und das bedeutet $\liminf f(P_n) \geq f(\tilde{P})$.

Aufgaben zu III.1

Aufgabe III.1.1 :

Zeigen Sie: In jedem metrischen Raum sind die Mengen $B_r(\tilde{P}) = \{P : \text{dist}(P, \tilde{P})\}$ für jedes $r > 0$ offene Mengen. (Sie heissen die offenen Kugeln mit dem Radius r .)

Zeigen Sie weiter: Die Mengen $B_{\leq r}(\tilde{P}) = \{P : \text{dist}(P, \tilde{P}) \leq r\}$ sind abgeschlossen. (Sie heissen die abgeschlossenen Kugeln mit dem Radius r .)

Aufgabe III.1.2 :

Zeigen Sie: Die Indikatorfunktion einer offenen Menge ist unterhalbstetig.

Zeigen Sie weiter: Zur Menge A sei

$$g_A(P) = \begin{cases} 0 & \text{wenn } P \in A \\ +\infty & \text{sonst} \end{cases}.$$

Zeigen Sie, dass $g_A(\cdot)$ genau dann unterhalbstetig ist, wenn A abgeschlossen ist.

Aufgabe III.1.3 :

Es seien $(S_1, d_1(\cdot, \cdot)), (S_2, d_2(\cdot, \cdot)), \dots$ metrische Räume. Für Punktepaaire im cartesischen Produkt $S = S_1 \times S_2 \times \dots = \{x = (x_1, x_2, \dots) : x_j \in S_j\}$ definieren wir

$$d(x, y) = \sum_j \frac{1}{2^j} \cdot (d_j(x_j, y_j) \wedge 1).$$

Zeigen Sie, dass $d(\cdot, \cdot)$ eine Metrik ist.

Zeigen Sie weiter: Eine S -Folge $x^{(1)}, x^{(2)}, \dots$ ist genau dann Cauchy-Folge bzgl. der Metrik $d(\cdot, \cdot)$, wenn für alle j die j -te Komponente $x_j^{(1)}, x_j^{(2)}, \dots$ im metrischen Raum $(S_j, d_j(\cdot, \cdot))$ eine Cauchyfolge ist.

Aufgabe III.1.4 :

Für ein festes d -Tupel positiver Zahlen $z = (z_1, \dots, z_d)$ betrachten wir die p -Norm als Funktion von $p \in [1, +\infty]$. $p \longrightarrow \|z\|_p = (\sum |z_j|^p)^{1/p}$. Zeigen Sie

1. Die Funktion $\ln \|z\|_p$ ist monoton fallend; die Funktion $f(p) = p \cdot \ln \|z\|_p$ ist monoton steigend.
2. Die Funktion $f(p) = p \cdot \ln \|z\|_p$ ist konvex.
3. Für $1 \leq r \leq s$ gilt $d^{-1/r} \|z\|_r \leq d^{-1/s} \|z\|_s$.
4. Finden Sie für das Paar $r \leq s$ eine Konstante $C = C(r, s)$ sodass gilt $\|z\|_s \leq \|z\|_r \leq C \cdot \|z\|_s$.

Hinweise:

zu 1): Beachten Sie, dass für $p \leq q$ die Einheitskugel zur p -Norm enthalten ist in der Einheitskugel zur q -Norm.

zu 2): Zeigen Sie für $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ und $r_1, r_2 \geq 1$

$$f\left(\frac{1}{p}r_1 + \frac{1}{q}r_2\right) \leq \frac{1}{p}f(r_1) + \frac{1}{q}f(r_2)$$

indem Sie die Höldersche Ungleichung anwenden auf $x_i := |z_i|^{\frac{1}{p}r_1}$ und $y_i := |z_i|^{\frac{1}{q}r_2}$.

zu 3): Zeigen Sie zuerst mit Hilfe der Hölderschen Ungleichung $\|z\|_1 \leq d^{1-1/p} \cdot \|z\|_p$ und wenden Sie dann nochmals die Höldersche Ungleichung an mit $s = rp$.

III.2 Normierte Vektorräume und stetige Linearformen

Hier sind auch unendlichdimensionale (reelle oder komplexe) Vektorräume zugelassen. Man denke z. B. an den Vektorraum der trigonometrischen Polynome oder an den Raum der stetigen Funktionen auf dem Einheitsintervall.

Definition

Ein reeller oder komplexer Vektorraum V wird dadurch zu einem **normierten Vektorraum**, dass man eine Funktion auf V auszeichnet, welche den Forderungen an eine Norm genügt.

Definition (Seminorm)

Eine nichtnegative Funktion $p(\cdot)$ auf einem reellen oder komplexen Vektorraum V heißt eine **Seminorm**, wenn gilt

- (i) $p(\alpha \cdot v) = |\alpha| \cdot p(v)$ für alle Skalare α ('Absolute Homogenität')
- (ii) $p(v_1 + v_2) \leq p(v_1) + p(v_2)$ ('Subadditivität')

Man nennt $p(\cdot)$ eine Norm, wenn zusätzlich gilt

- (iv) $p(v) = 0 \implies v = 0$.

Die Metrik zu einer Norm Normen benützt man zur Konstruktion spezieller translationsinvarianter Metriken auf affinen Räumen; man definiert den Abstand $d(P, Q)$ als die Norm des Verschiebungsvektors \overrightarrow{PQ} . Die Eigenschaften der Norm garantieren offenbar, dass $d(\cdot, \cdot)$ den Forderungen an eine Metrik genügt.

Die L^2 -Norm und die l^2 -Norm

V sei der komplexe Vektorraum aller trigonometrischen Polynome. Für $f(\cdot) \in V$ sei

$$\|f\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{+\pi} |f(t)|^2 dt \right)^{1/2} \quad (,,2\text{-Norm}").$$

Um die Subadditivität zu beweisen, müssen wir zeigen

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2 \quad \text{oder} \quad \|f\|^2 + \|g\|^2 + 2 \cdot \|f\| \cdot \|g\| \geq \|f + g\|^2 .$$

Wir zeigen, wie sich das aus der folgenden ebenfalls wichtigen Ungleichung ergibt, die wir unten beweisen werden

$$\frac{1}{2\pi} \int |f(t) \cdot g(t)| dt \leq \|f\|_2 \cdot \|g\|_2 .$$

Man beachte: Diese Ungleichung kann als ein kontinuierliches Analogon verstanden werden zu einer oben bewiesenen Ungleichung für die Wurzel aus der Quadratsumme. Es

handelt sich um Versionen der ‘Ungleichung von Cauchy-Schwarz-Buniakowski’, (kurz: ‘Schwarz’sche Ungleichung’), die uns noch öfters beschäftigen wird. Die Subadditivität der ‘2-Norm’ $\|\cdot\|_2$ folgt wegen

$$\|f + g\|^2 = \frac{1}{2\pi} \int |f + g|^2 dt = \frac{1}{2\pi} \int (|f|^2 + |g|^2 + 2\Re(\bar{f} \cdot g)) dt .$$

Beweis der Schwarz’schen Ungleichung

Wir betrachten für ein festes Paar f, g die folgende nichtnegative Funktion von λ auf der reellen Achse

$$\begin{aligned} q(\lambda) &\leq \| |f| + \lambda |g| \|^2 = \|f\|^2 + |\lambda|^2 \cdot \|g\|^2 + \frac{1}{2\pi} \int 2 \cdot \lambda |f \cdot g|(t) dt \\ &= a + \lambda^2 \cdot c + 2b\lambda \quad \text{mit } a = \|f\|^2, \quad c = \|g\|^2 . \end{aligned}$$

Man lernt bereits in der Schulmathematik, dass $b^2 \leq a \cdot c$ gilt für jede quadratische Funktion $q(\lambda)$, die keine reelle Nullstelle besitzt.

Man kann das Resultat auch folgendermaßen formulieren

Satz

Für trigonometrische Polynome sei definiert

$$\langle f | g \rangle = \frac{1}{2\pi} \int_0^{2\pi} (\bar{f}(t) g(t)) dt .$$

Es gilt dann

$$|\langle f | g \rangle| \leq \|f\|_2 \cdot \|g\|_2 .$$

Bemerkung :

In der letzteren Form ist die Schwarz’sche Ungleichung für trigonometrische Polynome identisch mit der Schwarz’schen Ungleichung für Zahlenfolgen. Wie wir in der Vorlesung I.6 gesehen haben, kann man nämlich die Norm eines trigonometrischen Polynoms sehr einfach aus seinen Koeffizienten berechnen:

$$\|f\|_2 = \left(\sum |a_n|^2 \right)^{1/2} \quad \text{falls} \quad f(t) = \sum a_n e^{int} .$$

Die Schwarz’sche Ungleichung besagt in diesem Fall wegen

$$\langle f | g \rangle = \frac{1}{2\pi} \int (\bar{f} \cdot g)(t) dt = \sum \bar{a}_n \cdot b_n$$

nichts anderes als die Ungleichung

$$\sum |a_n b_n| \leq \left(\sum |a_n|^2 \right)^{1/2} \cdot \left(\sum |b_n|^2 \right)^{1/2} .$$

für finite Zahlenfolgen $(a_n)_n$ und $(b_n)_n$.

Stetige Linearformen

Der Raum aller Linearformen auf einem Vektorraum V ist in jedem Fall ein Vektorraum. Wenn V ein unendlichdimensionaler normierter Vektorraum ist, dann nennt man den Raum aller Linearformen den algebraischen Dualraum V^* – zur Unterscheidung vom (viel interessanteren) Raum der stetigen Linearformen, den man üblicherweise mit V' bezeichnet und einfach den **Dualraum** von $(V, \|\cdot\|)$ nennt.

Definition (Stetige Linearform)

$(V, \|\cdot\|)$ sei ein normierter Vektorraum. Eine Linearform $\ell(\cdot)$ auf V heißt eine **stetige Linearform**, wenn sie auf der Einheitskugel beschränkt ist. Ihr Supremum auf der Einheitskugel wird mit $\|\ell\|$ bezeichnet. $\|\ell\|$ heißt die Norm der Linearform.

Die Zuordnung

$$V' \ni \ell \longmapsto \|\ell\| = \sup\{|\langle \ell, v \rangle| : \|v\| \leq 1\}$$

heißt die **duale Norm**.

Bemerke : Die ursprüngliche Norm $\|\cdot\|$ auf V und die duale Norm $\|\cdot\|$ auf V' leben auf ganz verschiedenen Räumen; es besteht keine Verwechslungsgefahr, wenn man sie mit demselben Symbol $\|\cdot\|$ bezeichnet.

Die p-Normen auf dem Raum der finiten Folgen

Satz (Minkowski-Ungleichung)

Sei V der komplexe Vektorraum der finiten Folgen

$$a = (\dots, a_{-1}, a_0, a_1, a_2, \dots) .$$

Für ein festes $p \geq 1$ definiert man

$$\|a\|_p = \left(\sum |a_j|^p \right)^{1/p} .$$

Es gilt dann

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p \quad \text{für alle } a, b \in V .$$

Der Satz besagt im Wesentlichen, dass $\|\cdot\|_p$ eine Norm ist. Der Beweis der Subadditivität erfordert Vorbereitungen. Wir werden ihn so führen, dass wir auch verstehen, wie die stetigen Linearformen auf V aussehen und wie man ihre Norm berechnet.

Beispiele stetiger Linearformen liefern uns die finiten Folgen ξ , wenn wir definieren

$$\langle \xi, a \rangle = \sum \xi_j \cdot a_j .$$

(Man sollte sich die $a \in V$ als \mathbb{Z} -Spalten vorstellen und die ξ als \mathbb{Z} -Zeilen.)

Es wird uns mit einem Trick gelingen, die Norm dieser Linearformen ξ zu bestimmen.

Satz :

Sei $(V, \|\cdot\|_p)$ der normierte Vektorraum aller p -summablen Folgen \mathbf{a} , $\|\mathbf{a}\|_p = (\sum |\mathbf{a}_j|^p)^{1/p}$. Die beschränkten Linearformen entsprechen dann genau den q -summablen Folgen ξ , wobei $1/p + 1/q = 1$.

Die duale Norm zu $(V, \|\cdot\|_p)$ ist die q -Norm

$$\sup\{|\langle \xi, \mathbf{a} \rangle| : \sum |\mathbf{a}_j|^p \leq 1\} = \|\xi\|_q = (\sum |\xi_j|^q)^{1/q}.$$

Beweis Es ist klar, dass jede Linearform auf dem Vektorraum aller finiten Folgen \mathbf{a} durch eine (i. Allg. nicht finite) Folge ξ gegeben ist. Wir wollen für eine beliebige Folge ξ mit $\sum |\xi_j|^q < \infty$ beweisen

1. Für jede finite Folge \mathbf{a} mit $\sum |\mathbf{a}_j|^p < 1$ gilt $\sum_j |\xi_j \cdot \mathbf{a}_j| \leq (\sum |\xi_j|^q)^{1/q}$
2. Zu jedem $\varepsilon > 0$ existiert ein finites $\mathbf{a}^{(\varepsilon)}$ mit

$$\sum |\mathbf{a}_j^{(\varepsilon)}|^p \leq 1 \quad \text{und} \quad \sum |\xi_j \mathbf{a}_j^{(\varepsilon)}| \geq (1 - \varepsilon) \cdot \left(\sum |\xi_j|^q\right)^{1/q}.$$

Und dafür genügt es offenbar, für finite Folgen \mathbf{a} und ξ das Folgende nachzuweisen:

Satz (Hölder'sche Ungleichung)

Seien ξ und \mathbf{a} finite Folgen; und sei (für $1/p + 1/q = 1$)

$$\|\mathbf{a}\|_p = (\sum |\mathbf{a}_j|^p)^{1/p}; \quad \|\xi\|_q = (\sum |\xi_j|^q)^{1/q}.$$

Dann gilt

$$\sum |\xi_j \mathbf{a}_j| \leq \|\xi\|_q \cdot \|\mathbf{a}\|_p.$$

Gleichheit gilt genau dann, wenn

$$|\mathbf{a}_j|^p = |\xi_j|^q \quad \text{für alle } j.$$

Beweis

1. Der Logarithmus $\ln(\cdot)$ ist eine konkave Funktion. Es gilt daher

$$\ln\left(\frac{1}{p}\mathbf{a} + \frac{1}{q}\mathbf{b}\right) \geq \frac{1}{p}\ln \mathbf{a} + \frac{1}{q}\ln \mathbf{b} \quad \text{für alle } \mathbf{a}, \mathbf{b} > 0.$$

Gleichheit gilt genau dann, wenn $\mathbf{a} = \mathbf{b}$.

Die Aussage kann man umformulieren

$$\mathbf{a}^{1/p} \cdot \mathbf{b}^{1/q} \leq \frac{1}{p}\mathbf{a} + \frac{1}{q}\mathbf{b}.$$

Bemerke: Im Falle $p = q = 2$ ergibt sich die bekannte Ungleichung $\sqrt{\mathbf{a} \cdot \mathbf{b}} \leq \frac{1}{2}(\mathbf{a} + \mathbf{b})$. In Worten: „Das geometrische Mittel positiver Zahlen ist nicht größer als das arithmetische Mittel“.

2. Es genügt offenbar, nichtnegative Folgen ξ, \mathbf{a} zu untersuchen; außerdem können wir annehmen

$$\|\xi\|_q = 1 = \|\mathbf{a}\|_p.$$

Wir müssen dann zeigen

$$\sum \xi_j a_j \leq 1.$$

In der Tat gilt

$$a_j \xi_j = (a_j^p)^{1/p} \cdot (\xi_j^q)^{1/q} \leq \frac{1}{p} a_j^p + \frac{1}{q} \xi_j^q.$$

Summation über j liefert die gewünschte Ungleichung.

3. Gleichheit haben wir genau dann, wenn

$$|a_j|^p = |\xi_j|^q \quad \text{für alle } j. \quad \text{q.e.d.}$$

Corollar Wenn wir die Rollen von p und q vertauschen, dann haben wir

$$\|\mathbf{a}\|_p = \sup\{|\langle \xi, \mathbf{a} \rangle| : \sum |\xi_j|^q \leq 1\}.$$

Die Funktion $\|\cdot\|_p$ ist das Supremum der Seminormen $|\langle \xi, \cdot \rangle|$, erstreckt über eine gewisse Menge von ξ . Sie ist daher sublinear und die Minkowski-Ungleichung ist bewiesen.

Hinweise

1. Der Raum der absolutsummablen Folgen \mathbf{a} mit der Norm $\|\mathbf{a}\|_1 = \sum |a_j|$ ist ein Grenzfall, der Aufmerksamkeit verdient. Offenbar entsprechen die beschränkten Linearformen den beschränkten Folgen ξ und es gilt

$$\sup\{|\langle \xi, \mathbf{a} \rangle| : \sum |a_j| \leq 1\} = \sup_j |\xi_j|.$$

Die duale Norm ist also die Supremumsnorm auf dem Raum aller beschränkten Folgen ξ .

2. Für jede feste finite Folge ξ ist die Funktion $q \mapsto \|\xi\|_q$ antiton auf dem Intervall $(1, \infty)$. Es gilt

$$p < q, \quad \|\xi\|_p \leq 1 \implies \|\xi\|_q \leq 1$$

$$\|\xi\|_\infty = \lim_{q \rightarrow \infty} \searrow \|\xi\|_q.$$

3. In der Funktionalanalysis beweist man, dass es auf dem Raum $\ell^\infty(\mathbb{Z})$ der beschränkten Folgen mit der Supremumsnorm neben den absolutsummablen Folgen \mathbf{a} (man schreibt $\mathbf{a} \in \ell^1(\mathbb{Z})$) noch weitere beschränkte Linearformen gibt. Der Beweis ist jedoch nichtkonstruktiv, man kann keine solchen exotischen beschränkten Linearformen angeben (Stichwort: Banach-Limiten). Dennoch: Der Dualraum des Dualraums von ℓ^1 ist hier echt größer als der Raum ℓ^1 der summablen Folgen. In der Fachsprache sagt man: Der Folgenräume ℓ^p mit $1 < p < \infty$ sind reflexive Banachräume; der Raum ℓ^1 jedoch ist nicht reflexiv.

Der bekannteste Fall ist der Fall $p = 2 = q$. Hier können wir die Resultate auch in der Sprache der trigonometrischen Reihen formulieren. Der normierte Vektorraum $(V_{\text{fin}} \|\cdot\|_2)$ kann nämlich mit dem Raum der trigonometrischen Polynome identifiziert werden, und die 2-Norm erhält dabei eine vertraute Gestalt. Die 2-Norm des trigonometrischen Polynoms $f = \sum c_n \cdot e^{int}$ ist

$$\|f\|_2 = \left(\sum |c_n|^2 \right)^{1/2} = \left(\frac{1}{2\pi} \int_{-\pi}^{+\pi} |f(t)|^2 dt \right)^{1/2}.$$

In der Funktionalanalysis zeigt man mit Hilfe der Lebesgue'schen Integrationstheorie, dass es zu jeder quadratsummablen Folge $(\xi_n)_n$ eine 2π -periodische Funktion $h(t)$ gibt, sodass

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} \overline{h(t)} \cdot f(t) dt = \sum \overline{\xi_n} \cdot c_n \quad \text{für alle } f = \sum c_n \cdot e^{int}.$$

Dies sind also die stetigen Linearformen auf dem Raum der trigonometrischen Polynome. Beispiele für nichtstetige Linearformen sind die Auswertungen der trigonometrischen Polynome im Nullpunkt (oder in irgendeinem anderen Punkt t_0). Die Unstetigkeit ersieht man daran, dass es trigonometrische Polynome mit Norm ≤ 1 gibt, für welche die Auswertung einen beliebig vorgegebenen Wert überschreitet. Für das auf die Norm = 1 normierte N -te Dirichlet-Polynom $f(t)$ beispielsweise liefert die Auswertung im Nullpunkt den Wert $\frac{1}{2}\sqrt{2N+1}$, denn

$$f(t) = \frac{1}{\sqrt{2N+1}} \sum_{-N}^{+N} e^{int} = \frac{1}{\sqrt{2N+1}} \cdot D_N(t) = \frac{1}{\sqrt{2N+1}} \cdot \frac{1}{\sin t} \cdot \sin \left(N + \frac{1}{2} \right) t.$$

Hinweis: Punktetrennende Paarungen

Man zeigt in der Funktionalanalysis, dass es auf jedem normierten Raum 'viele' beschränkte Linearformen gibt. (Das Stichwort ist der Satz von Hahn-Banach.) Manchmal ist es aber schwer, den Dualraum in seiner Gänze zu überblicken. Deswegen bevorzugt man es bei manchen Konstruktionen, nicht alle beschränkten Linearformen ins Spiel zu bringen. Beispielsweise kann man sich bei den meisten interessanten Konstruktionen im Raum der beschränkten Folgen mit denjenigen Linearformen begnügen, die durch eine summable Folge gegeben sind. Man definiert:

Definition

Seien $(V, \|\cdot\|)$ und $(W, \|\cdot\|)$ normierte Vektorräume und $B(\cdot, \cdot)$ eine Bilinearform auf $W \times V$ mit

- (i) $|B(w, v)| \leq c \cdot \|w\| \cdot \|v\|$
- (i) $B(w, \cdot) = 0 \Rightarrow w = 0$
- (i) $B(\cdot, v) = 0 \Rightarrow v = 0$.

Man nennt $B(\cdot, \cdot)$ dann eine punktetrennende Paarung, die bzgl. des gegebenen Paares von Normen beschränkt ist.

Beispiel

Sei V der unendlichdimensionale Raum der finiten Folgen.

$$B(y, x) = \sum y_j \cdot x_j \quad \text{für } y \in V, x \in V.$$

$B(\cdot, \cdot)$ ist eine punktgetrennende Paarung auf $V \times V$, welche beschränkt ist bzgl. dem Paar von Normen $\|\cdot\|_p$ und $\|\cdot\|_q$ (wobei $\frac{1}{p} + \frac{1}{q} = 1$). Auch im Fall $p = 1, q = \infty$ gilt

$$\begin{aligned} \|y\|_q &= \sup\{|B(y, x)| : \|x\|_p \leq 1\} \\ \|x\|_p &= \sup\{|B(y, x)| : \|y\|_q \leq 1\}. \end{aligned}$$

Die wichtigsten Paarungen sind zweifellos die Paarungen eines normierten Vektorraums $(V, \|\cdot\|)$ mit einem (genügend großen) Teilraum seines Dualraums $(V', \|\cdot\|)$.

Die Bilinearform auf $V' \times V$

$$B(\ell, v) = \ell(v) = \langle \ell, v \rangle \quad \text{für } \ell \in V', v \in V.$$

heißt die natürliche Paarung zu $(V, \|\cdot\|)$.

Hinweis Auch in der Integrationstheorie gibt es die Hölder'sche Ungleichung und die Minkowski'sche Ungleichung:

Wenn f eine p -integrierte Funktion ist und g eine q -integrierte Funktion mit $1/p + 1/q = 1$, dann ist das punktweise Produkt $h = f \cdot g$ integrierbar, und es gilt

$$\int |f \cdot g| \leq \|f\|_p \cdot \|g\|_q = \left(\int |f|^p \right)^{1/p} \cdot \left(\int |g|^q \right)^{1/q}.$$

Gleichheit gilt genau dann, wenn $|f|^p = |g|^q$ (fast überall).

Aufgaben zu III.2

Aufgabe III.2.1 :

Zeigen Sie: Eine positivhomogene Funktion (auf einem reellen Vektorraum) ist genau dann konvex, wenn sie subadditiv ist.

Aufgabe III.2.2 :

Der Dualraum des Dualraums eines endlichdimensionalen Vektorraums V kann bekanntlich ('in kanonischer Weise') mit V identifiziert werden: $V^{**} = V$.

Sei nun $n(\cdot)$ eine Metrik auf dem endlichdimensionalen (reellen oder komplexen) Vektorraum V und $n^*(\cdot)$ die duale Norm. Zeigen Sie, dass die dazu duale Norm gleich $n(\cdot)$ ist: $n^{**} = n$.

Aufgabe III.2.3 :

Es sei $p \geq 1$, $1/p + 1/q = 1$ und $a, b > 0$. Zeigen Sie mit den Mitteln der Schulmathematik

$$\sup_{0 < t < 1} (a \cdot t^{1/p} + b \cdot (1-t)^{1/p}) = (a^q + b^q)^{1/q}.$$

Empfehlung: Behandeln Sie zuerst den Fall $p = 2 = q$.

Aufgabe III.2.4 : (Hilbert-Schmidt-Norm)

Für $n \times n$ -Matrizen mit komplexen Einträgen a_{kl} definiert man

$$\|A\| = \|A\|_{\text{HS}} := \left(\sum_{k,l} |a_{kl}|^2 \right)^{1/2}.$$

Dies ist eine Norm auf einem n^2 -dimensionalen Vektorraum. Man kann hier aber auch multiplizieren. Zeigen Sie

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|$$

Hinweis: Wenden Sie die Schwarzsche Ungleichung an auf jedes $c_{kn} = \sum_l a_{kl} b_{ln}$.

Aufgabe III.2.5 : (Norm zur Faltung)

Für trigonometrische Polynome

$$f(t) = \sum_k a_k \cdot e^{ikt}$$

definieren wir hier einmal

$$\|f\| = \sum_k |a_k|$$

Dies ist eine Norm auf einem ∞ -dimensionalen Vektorraum. In diesem Raum kann man auch multiplizieren: $f \cdot g = h$ ist ein trigonometrisches Polynom (wo übrigens die Koeffizienten von h durch die Faltungsformel gegeben sind).

Zeigen Sie $\|f \cdot g\| \leq \|f\| \cdot \|g\|$.

III.3 Konvexe Mengen und konvexe Funktionen

In einem n -dimensionalen reellen Vektorraum bieten sich Verallgemeinerungen des Begriffs der Semi-Norm an. Wir betrachten die Distanz-Funktionen und die Stützfunktionen.

Definition

Eine Funktion $F(\cdot)$ auf einem reellen Vektorraum V heißt **positiv homogen**, wenn gilt

$$F(\alpha v) = \alpha \cdot F(v) \quad \text{für alle } \alpha \geq 0, v \in V.$$

Eine positiv homogene Funktion auf einem n -dimensionalen reellen Vektorraum, die nicht-negativ und subadditiv ist, nennt man ein **Minkowski-Funktional** oder auch kurz eine **M-Norm**.

Bemerkung: Eine positiv homogene Funktion ist genau dann subadditiv, wenn sie konvex ist. Eine M-Norm ist genau dann eine Seminorm wenn sie symmetrisch ist.

Konstruktion (Distanzfunktion zu K)

Sei V ein reeller Vektorraum. Zu einer beliebigen konvexen Nullumgebung $K \subseteq V$ definiert man die Distanzfunktion

$$m(v) = m_K(v) = \inf \left\{ \alpha : \alpha > 0, \frac{v}{\alpha} \in K \right\} \quad \text{für } v \in V.$$

Diese ist ein Minkowski-Funktional; sie ist also nichtnegativ, positiv homogen und subadditiv.

Der offene Kern K° und die abgeschlossene Hülle K^a liefern dieselbe Distanzfunktion $m(\cdot)$; und es gilt

$$K^\circ = \{v : m(v) < 1\}, \quad K^a = \{v : m(v) \leq 1\}.$$

Beweis :

Wir beweisen nur die Subadditivität der Distanzfunktion.

Seien $v, w \in V$ mit $m(v) = \alpha$, $m(w) = \beta$. Es gilt dann $\frac{v}{\alpha}, \frac{w}{\beta} \in K^a$ und daher

$$(1 - \lambda) \frac{v}{\alpha} + \lambda \frac{w}{\beta} \in K^a; \quad m\left((1 - \lambda) \frac{v}{\alpha} + \lambda \frac{w}{\beta}\right) \leq 1 \quad \text{für alle } \lambda \in [0, 1].$$

Speziell für $1 - \lambda = \frac{\alpha}{\alpha + \beta}$, $\lambda = \frac{\beta}{\alpha + \beta}$ ergibt sich $m(v + w) \leq \alpha + \beta$.

Konstruktion (Stützfunktion zu K)

Zur Menge $K \subseteq V$ definiert man die ‘Stützfunktion’ auf dem Dualraum V^*

$$\varphi(\theta) = \varphi_K(\theta) = \sup \{ \langle \theta, v \rangle : v \in K \} \quad \text{für } \theta \in V^*.$$

Bemerkungen:

1. Jede Stützfunktion ist subadditiv und positiv homogen mit Werten in $\mathbb{R}_+ \cup \{+\infty\}$. Die abgeschlossene konvexe Hülle von K liefert offenbar dieselbe Stützfunktion.
2. Die Stützfunktion von K ist genau dann endlichwertig, wenn K beschränkt ist. In jedem Falle ist die Stützfunktion unterhalbstetig als das punktweise Supremum einer Familie von Linearformen $\{\langle \cdot, v \rangle : v \in K\}$.
3. Wenn K die Einheitskugel einer Norm auf V ist, dann ist die Stützfunktion $\varphi_K(\cdot)$ die duale Norm.
4. Wenn K eine konvexe Nullumgebung, $m_K(v)$ ihre Distanzfunktion und $\varphi_K(\theta)$ ihre Stützfunktion ist, dann sind die folgenden Bedingungen an ein $\theta \in V^*$ äquivalent

$$\varphi_K(\theta) \leq 1 \iff \langle \theta, v \rangle \leq 1 \text{ für alle } v \in K \iff \langle \theta, \cdot \rangle \leq m_K(\cdot).$$

Die Menge $M = \{\theta : \varphi_K(\theta) \leq 1\} \subseteq V^*$ ist hier also die Menge aller linearen Minoranten der Distanzfunktion $m_K(\cdot)$. Wir werden beweisen, dass das punktweise Supremum dieser Minoranten die Distanzfunktion ist.

$$m_K(v) = \sup\{\langle \theta, v \rangle : \theta \in M\} \text{ für alle } v \in V.$$

Da M beschränkt und abgeschlossen ist, wird das Supremum angenommen. Es wird also darum gehen, die folgende Aussage zu beweisen:

$$\forall \tilde{v} \in V \quad \exists \theta : \left(\langle \theta, \cdot \rangle \leq m_K(\cdot) \right) \wedge \left(\langle \theta, \tilde{v} \rangle = m_K(\tilde{v}) \right).$$

Satz : (Fortsetzung einer Linearform)

Gegeben ist eine positiv homogene subadditive Funktion $q(\cdot)$ auf dem n -dimensionalen reellen Vektorraum V , sowie eine Linearform $\tilde{l}(\cdot)$ auf einem Teilraum \tilde{V} , die von der Einschränkung von $q(\cdot)$ majorisiert wird. ($\tilde{l}(w) \leq q(w)$ für alle $w \in \tilde{V}$). Dann existiert eine Linearform $l(\cdot) \in V^*$ mit

$$\forall w \in \tilde{V} \quad l(w) = \tilde{l}(w), \quad \forall v \in V \quad l(v) \leq q(v).$$

(In Worten: Jede auf einem Teilraum gegebene lineare Minorante von $q(\cdot)$ kann zu einer linearen Minoranten auf dem ganzen Raum fortgesetzt werden. – Für unendlichdimensionale Vektorräume trägt dieser Satz den Namen ‘Satz von Hahn-Banach’.)

Beweis :

Sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V so, dass $\{v_1, v_2, \dots, v_k\}$ eine Basis von $V_k = \tilde{V}$ ist. $\{l^1(\cdot), l^2(\cdot), \dots, l^n(\cdot)\}$ sei die duale Basis. Es gibt genau ein k -Tupel von Koeffizienten $(\beta_1, \dots, \beta_k)$ so, dass $\tilde{l}(\cdot)$ die Einschränkung von $\sum^k \beta_j \cdot l^j(\cdot)$ ist. Wir suchen $\beta = \beta_{k+1}$ sodass $\sum^{k+1} \beta_j \cdot l^j(\cdot)$ auf V_{k+1} von $q(\cdot)$ majorisiert wird. Jeder Vektor in $V_{k+1} \setminus V_k$ ist

entweder ein positives Vielfaches von $v_{k+1} + w'$ oder ein positives Vielfaches von $-v_{k+1} + w''$ mit $w', w'' \in \tilde{V}$. Die Forderung an β lautet also

$$\forall w', w'' \in \tilde{V} \quad \beta + \tilde{l}(w') \leq q(v_{k+1} + w'), \quad -\beta + \tilde{l}(w'') \leq q(-v_{k+1} + w'') \\ \sup\{\tilde{l}(w'') - q(-v_{k+1} + w'')\} \leq \beta \leq \inf\{-\tilde{l}(w') + q(v_{k+1} + w')\}.$$

Die Forderung ist erfüllbar, weil gilt

$$\forall w', w'' \in \tilde{V} \quad l(w') + l(w'') = l(w' + w'') \leq q(w' + w'') \leq q(v_{k+1} + w') + q(-v_{k+1} + w'').$$

Wir wählen β_{k+1} und fahren so fort. Nach $n - k$ Schritten haben wir $(\beta_{k+1}, \dots, \beta_n)$ so, dass $l(\cdot) = \sum^n \beta_j \cdot l^j(\cdot)$ eine von $q(\cdot)$ majorisierte Fortsetzung von $\tilde{l}(\cdot)$ ist.

Es ist noch etwas zum Start ($k=0$) unserer schrittweisen Konstruktion zu sagen: Die möglichen Werte $\beta = \beta_1$ für den ersten Vektor v_1 sind die Zahlen im Intervall $[-q(-v_1), q(v_1)]$. Die Wahl $\beta_1 = q(v_1)$ führt auf eine lineare Minorante $l(\cdot)$ mit $l(v_1) = q(v_1)$. Da $v_1 \neq 0$ beliebig gewählt werden kann, folgt: Jede positiv homogene sublineare Funktion $q(\cdot)$ ist das punktweise Maximum aller ihrer linearen Minoranten

$$q(\cdot) = \sup\{l(\cdot) : l \text{ linear} \leq q\}.$$

Man kann das Resultat auch als einen Trennungssatz formulieren

Satz (Stützebenen in Randpunkten offener konvexer Mengen)

Es sei K_0 eine offene konvexe Menge in einem affinen Raum, und P_1 ein Punkt ausserhalb. Es existiert dann eine affine Funktion $\alpha(\cdot)$, sodass

$$\alpha(P_1) \geq 0 \quad \text{und} \quad \{P : \alpha(P) < 0\} \supseteq K_0.$$

Beweis : Wir wählen einen inneren Punkt $P_0 \in K_0$ und identifizieren die Punkte P mit den Verschiebungsvektoren $v = \overrightarrow{P_0 P}$. Die Distanzfunktion von K_0 relativ zu P_0 sei $q(\cdot)$. Da P_1 nicht in K_0 liegt, gilt $q(\overrightarrow{P_0 P_1}) \geq 1$. Sei $l(\cdot)$ eine lineare Minorante von $q(\cdot)$ mit $l(\overrightarrow{P_0 P_1}) = q(\overrightarrow{P_0 P_1}) \geq 1$. Wegen $\forall P \in K_0 \quad l(\overrightarrow{P_0 P}) < 1$ leistet die affine Funktion $\alpha(P) = l(\overrightarrow{P_0 P}) - 1$ das Verlangte.

Sprechweise: Wenn P_1 auf dem Rand von K_0 liegt, dann nennt man die Hyperebene $H = \{P : \alpha(P) = 0\}$ eine Stützhyperebene oder auch Stützebene im Punkt P_1 . Wenn es nur eine Stützhyperebene im Punkt P_1 gibt, dann nennt man sie die *Tangentialebene* im Punkt P_1 .

Satz (Trennungssatz für konvexe Mengen)

Es sei K_0 eine offene konvexe Menge in einem affinen Raum, und K_1 eine beliebige dazu disjunkte konvexe Menge. Es existiert dann eine affine Funktion $\alpha(\cdot)$, sodass

$$\{Q : \alpha(Q) \geq 0\} \supseteq K_1 \quad \text{und} \quad \{P : \alpha(P) < 0\} \supseteq K_0.$$

Beweis : Sei L die Menge aller Verschiebungsvektoren von K_0 nach K_1 : $L = K_1 - K_0 = \{\vec{v} = \overrightarrow{P_0 P_1} : P_0 \in K_0; P_1 \in K_1\}$. L ist offenbar eine offene konvexe Menge, die den Nullpunkt nicht enthält. Es existiert eine Linearform \tilde{l} , die auf L echt positiv ist. Es sei $\alpha(\cdot)$ eine affine Funktion von der Gestalt $\alpha(P) = s + \tilde{l}(\overrightarrow{P_* P})$. Es gilt $\alpha(P_0) - \alpha(P_1) = -\tilde{l}(\overrightarrow{P_0 P_1}) > 0$ für alle $P_0 \in K_0, P_1 \in K_1$. Wir finden

$$s := \sup\{\alpha(P_0) : P_0 \in K_0\} \leq \inf\{\alpha(P_1) : P_1 \in K_1\}$$

Es gilt also $\alpha(P_1) - s \geq 0$ für alle $P_1 \in K_0$, und $\alpha(P_0) - s < 0$ für alle $P_0 \in K_0$. (Die strikte Ungleichung < 0 ergibt sich aus der Annahme, dass K_0 offen ist.)

Abgeschlossene konvexe Mengen und unterhalbstetige konvexe Funktionen

Bisher ging es hauptsächlich um offene konvexe Mengen und positivhomogene konvexe Funktionen auf einem reellen Vektorraum. Hier beschäftigen wir uns jetzt mit abgeschlossenen konvexen Mengen und unterhalbstetigen konvexen Funktionen auf einem reellaffinen Raum $(L; V)$. Manchmal wird es bequem sein, einen 'Nullpunkt' P_0 auszuzeichnen und die Punkte P des affinen Raums mit den Verschiebungsvektoren $\vec{v} = \overrightarrow{P_0 P}$ zu identifizieren. Man denkt gern an den Raum $V = \mathbb{R}_{\text{Sp}}^n$ der reellen n -Spalten, die mit $\mathbf{x}, \mathbf{y}, \dots$ bezeichnet werden; (Koordinatentransformationen spielen nämlich hier keine wichtige Rolle). Der Dualraum V^* ist mit dem Raum aller n -Zeilen θ zu identifizieren. Die Linearform θ liefert im Vektor \mathbf{x} den Wert $\theta(\mathbf{x}) = \langle \theta, \mathbf{x} \rangle = \theta \cdot \mathbf{x}$ mit dem Punkt im Sinne des Matrizenprodukts.

Rekapitulieren zuerst den Begriff der Konvexität ohne topologisches Beiwerk.

Definition (Konvexe Menge)

Eine Teilmenge K eines affinen Raums heißt eine konvexe Menge, wenn gilt

$$\forall P_0, P_1 \in K \quad \forall \lambda \in [0, 1] \quad (1 - \lambda)P_0 + \lambda P_1 \in K.$$

Satz (Konvexe Hülle)

Zu jeder Teilmenge eines affinen Raums gibt es eine kleinste sie umfassende konvexe Menge.

(Man nennt sie die konvexe Hülle.)

Beweis : Der Durchschnitt einer beliebigen Familie von konvexen Mengen ist eine konvexe Menge, (wenn er nicht leer ist). Der Durchschnitt aller die gegebene Menge umfassenden konvexen Mengen ist eine konvexe Obermenge und zwar offenbar die kleinste. Hinweis: Eine andere, 'konkretere' Kennzeichnung der konvexen Hülle findet man in den Übungen.

Definition (Konvexe Funktionen)

Eine Funktion $k(\cdot)$ auf einem affinen Raum heißt eine **konvexe Funktion**, wenn gilt

$$\forall P_0, P_1 \quad \forall \lambda \in [0, 1] \quad k((1 - \lambda)P_0 + \lambda P_1) \leq (1 - \lambda) \cdot k(P_0) + \lambda \cdot k(P_1).$$

Verabredung In manchen Lehrbüchern sind konvexe Funktionen reellwertige Funktionen, die auf einer konvexen Teilmenge eines affinen Raums definiert sind und dort die angegebene Eigenschaft besitzen. (‘Die Funktionswerte liegen unterhalb der Sehnen’.) Es ist bequemer, von den konvexen Funktionen zu verlangen, dass sie auf dem gesamten Raum definiert sind, wobei aber auch zugelassen ist, dass sie den Wert $+\infty$ annehmen. Man bemerke: Aus der Definition der Konvexität folgt, dass der ‘Endlichkeitsbereich’ $\{P : k(P) < \infty\}$ eine konvexe Menge K ist. Wenn auf einer konvexen Menge eine (endlichwertige) konvexe Funktion gegeben ist, dann gewinnt man dazu eine konvexe Funktion auf dem gesamten affinen Raum, indem man ausserhalb den Wert $+\infty$ festlegt.

Satz Eine Funktion ist genau dann konvex, wenn ihre Einschränkung auf jede Gerade konvex ist

(Der Beweis ist trivial.)

Den Beweis des folgenden Satzes überlassen wir dem Leser.

Satz

Wenn für eine konvexe Funktion $f(\cdot)$ gilt, dass auch $-f(\cdot)$ konvex ist, dann ist $f(\cdot)$ eine **affine** Funktion. Es gilt für alle reellen λ (und nicht nur für diejenigen im Einheitsintervall)

$$f((1-\lambda)P + \lambda Q) = (1-\lambda) \cdot f(P) + \lambda \cdot f(Q).$$

(Eine Funktion $h(\cdot)$ heißt eine **konkave** Funktion, wenn $-h(\cdot)$ konvex ist.)

Bemerke : Die Gesamtheit aller affinen Funktionen auf einem affinen Raum ist bekanntlich ein Vektorraum. Die Menge der konvexen Funktionen (über einem affinen Raum S) ist kein Vektorraum. Man kann konvexe Funktionen addieren und mit positiven Skalaren multiplizieren; man kann aber nicht subtrahieren. Man sagt, dass die Menge der konvexen Funktionen über S ein **konvexer Kegel** ist. Dieser Funktionenkegel hat die höchst bemerkenswerte Eigenschaft, dass er gegenüber punktweiser Maximumsbildung abgeschlossen ist.

Es sei nämlich $\{k_i : i \in I\}$ eine Familie konvexer Funktionen und k das punktweise Supremum. Für festes P, Q , $\lambda \in [0, 1]$ und alle i gilt

$$k_i((1-\lambda)P + \lambda Q) \leq (1-\lambda) \cdot k_i(P) + \lambda \cdot k_i(Q) \leq (1-\lambda) \cdot k(P) + \lambda \cdot k(Q) .$$

Es folgt $k((1-\lambda)P + \lambda Q) \leq (1-\lambda) \cdot k(P) + \lambda \cdot k(Q)$.

Satz

Zu jeder Funktion $g(\cdot)$ auf S , die mindestens eine konvexe Minorante besitzt, gibt es eine größte konvexe Minorante.

Beweis :

Das punktweise Maximum aller konvexen Minoranten ist konvex. Es ist offenbar die größte konvexe Minorante.

Einen Anschluß an die Schulmathematik liefert der

Satz

Eine $\mathbb{R} \cup \{+\infty\}$ -wertige Funktion auf der reellen Achse, welche auf einem offenen Intervall endlich ist, ist genau dann konvex, wenn die rechtsseitige (oder die linksseitige) Ableitung isoton ist.

(Der Beweis ist eine leichte Übungsaufgabe)

Beispiele von konvexen Funktionen über \mathbb{R}

$$1) \quad k_1(x) = a_0 + a_1x + a_2x^2 \quad \text{mit } a_2 \geq 0$$

$$2) \quad k_2(x) = \begin{cases} -\ln x & \text{für } x > 0 \\ +\infty & \text{für } x \leq 0 \end{cases}$$

$$3) \quad k_3(x) = \begin{cases} x \cdot \ln x & \text{für } x \geq 0 \\ +\infty & \text{für } x < 0 \end{cases}$$

$$4) \quad k_4(x) = |x| = x^+ + x^- \quad (x^+ = \max\{0, x\}, \quad x^- = \max\{0, -x\})$$

$$5) \quad k_5(x) = \sum p_j |x - x_j| \quad \text{mit } p_j \geq 0, \quad x_j \text{ beliebig.}$$

Die Ableitung von $k_5(\cdot)$ ist konstant in jedem Intervall, in welchem kein x_j liegt. Sie springt im Punkt x_j um den Betrag $2p_j$. Man skizziere z.B.

$$k(x) = \frac{1}{2}|x - 1| + \frac{1}{2}|x + 1| = \max\{|x|, 1\}.$$

Satz

Wenn K eine konvexe Menge ist, dann ist auch die abgeschlossene Hülle K^a konvex.

Beweis

Die Punkte $P, Q \in K^a$ sind die Limiten von K -Folgen. Wenn $P = \lim P_n$, $Q = \lim Q_n$, dann gilt für jedes $\lambda \in [0, 1]$ $(1 - \lambda)P + \lambda Q = \lim (1 - \lambda)P_n + \lambda Q_n \in K^a$.

Die kleinste abgeschlossene konvexe Obermenge einer Menge M nennt man die abgeschlossene konvexe Hülle. Offenbar ist eine Funktion f auf dem affinen Raum L genau dann konvex, wenn der Epigraph eine konvexe Teilmenge von $\mathbb{R} \times L$ ist. Wir haben oben in III.1 gesehen, dass der Epigraph genau dann abgeschlossen ist, wenn die Funktion f unterhalbstetig ist.

Satz

Wenn $k(\cdot)$ irgendeine konvexe Funktion ist, dann ist die abgeschlossene Hülle des Epigraphen der Epigraph einer unterhalbstetigen Funktion $k^{**}(\cdot)$, die sich von $k(\cdot)$ allenfalls in den Randpunkten des Endlichkeitsbereichs unterscheidet. (Ohne Beweis!)

Das punktweise Supremum von **affinen Funktionen** ist offenbar eine unterhalbstetige konvexe Funktion. Höchst bemerkenswert ist nun die Tatsache, dass man jede unterhalbstetige konvexe Funktion als das Supremum von affinen Funktionen gewinnen kann.

Theorem (Satz vom Supremum der affinen Minoranten)

Jede unterhalbstetige konvexe Funktion auf einem endlichdimensionalen reellaffinen Raum ist das Supremum ihrer affinen Minoranten.

Beweis

Wir betrachten Funktionen auf dem Vektorraum der n -Spalten $V = \mathbb{R}_{\text{Sp}}^n$. Die affinen Funktionen haben die Gestalt $a(x) = \theta x - c$, wo c eine Zahl und θ eine n -Zeile ist. Ein solches $a(\cdot)$ ist genau dann eine Minorante von $k(\cdot)$, wenn $a(x) - y \leq 0$ für alle $y \geq k(x)$, wenn also

$$-y + \theta x - c \leq 0 \quad \text{für alle } (y, x) \text{ im Epigraphen } \Gamma_{\geq k}.$$

Wenn $k(\cdot)$ im Punkt \tilde{x} unterhalbstetig ist, dann existiert zu jedem $M < k(\tilde{x})$ eine Kugel $B_\delta(\tilde{x})$, sodass $k(x) > M$ gilt für alle $x \in B_\delta(\tilde{x})$. Die offene konvexe Menge $U = (-\infty, M) \times B_\delta(\tilde{x}) \subset \mathbb{R} \times V$ ist also disjunkt zum Epigraphen $\Gamma_{\geq k}$.

Wenn $k(\cdot)$ auch noch konvex ist, dann existiert in $\mathbb{R} \times V$ eine trennende Hyperebene.

$$\tilde{b}y + \tilde{\theta}x - \tilde{c} \leq 0 \quad \text{für alle } (y, x) \in \Gamma_{\geq k}, \quad \tilde{b}y + \tilde{\theta}x - \tilde{c} > 0 \quad \text{für alle } (y, x) \in U.$$

Der Koeffizient \tilde{b} kann nicht positiv sein. Wir unterscheiden die Fälle $\tilde{b} = 0$ und $\tilde{b} < 0$. Im zweiten Fall können wir $\tilde{b} = -1$ annehmen. Die Funktion $\tilde{a}(x) = \tilde{\theta}x - \tilde{c}$ ist eine affine Minorante mit $\tilde{a}(x) = \tilde{\theta}x - \tilde{c} \geq M$ für alle $x \in B_\delta(\tilde{x})$. Wir haben also eine affine Minorante gefunden, welche im Punkt \tilde{x} den vorgegebenen Wert $M < k(\tilde{x})$ übersteigt.

Der Koeffizient \tilde{b} in der trennenden Hyperebene kann nur dann verschwinden, wenn $k(\tilde{x}) = \infty$. Wir haben in diesem Fall $\tilde{\theta}x - \tilde{c} \leq 0$ für alle x im Endlichkeitsbereich von k , und $\tilde{\theta}x - \tilde{c} > 0$ für alle $x \in B_\delta(\tilde{x})$. Wir müssen zu jedem M eine affine Minorante finden, die in \tilde{x} größer als M ist. Wenn $\bar{a}(x)$ irgendeine Minorante ist, dann erfüllt $\bar{a}(x) + C \cdot (\tilde{\theta}x - \tilde{c})$ für genügend großes C die Forderung.

Legendre-Transformierte

Eine nicht konstante affine Funktion $a(x) = \theta x - c$ (auf dem Vektorraum V der n -Spalten) nennen wir eine affine Funktion mit der Richtung θ . Eine solche Funktion ist Minorante der Funktion $f(\cdot)$ genau dann, wenn

$$\forall x \in V \quad \theta x - c \leq f(x), \quad \text{d. h. wenn} \quad c \geq \sup\{\theta x - f(x) : x \in V\}.$$

Wir definieren auf dem Dualraum V^* die Funktion

$$f^*(\theta) = \sup\{\theta x - f(x) : x \in V\} \quad \text{für } \theta \in V^*.$$

Wenn $f(\cdot)$ mindestens eine affine Minorante besitzt, dann ist $f^*(\cdot)$ nicht identisch $+\infty$. Wir interessieren uns nur für solche $f(\cdot)$.

Sprechweise: Es sei $k(\cdot)$ eine unterhalbstetige konvexe Funktion. Für jedes θ mit $k^*(\theta) < +\infty$ nennt man $\{(y, x) : y = \theta x - k^*(\theta)\}$ eine Stützebene an den Epigraphen $\Gamma_{\geq k}$, wobei aber zu beachten ist, dass eine Stützebene dieser Art nicht notwendig einen nichtleeren Durchschnitt mit dem Epigraphen hat. Es ist diejenige Hyperebene mit der Richtung θ , die Epigraphen am nächsten kommt.

Beispiel (Stützgeraden)

Betrachten wir die konvexe Funktion $k(x) = \sqrt{1+x^2}$. Zu jeder Richtung θ mit $|\theta| \leq 1$ existiert eine Stützgerade, nämlich der Graph von $y = \theta \cdot x - k^*(\theta) = \theta \cdot x + \sqrt{1-\theta^2}$. Für $|\theta| < 1$ entspricht die Stützgerade einer Tangente, für $\theta = \pm 1$ erhalten wir die Asymptoten $y = x$ bzw. $y = -x$.

Definition (Legendre-Transformierte)

Wenn $k(\cdot)$ auf dem Vektorraum V unterhalbstetig und konvex ist, und

$$k^*(\theta) = \sup\{\theta x - k(x) : x \in V\} \quad \text{für } \theta \in V^*,$$

dann heisst die Funktion $k^*(\cdot)$ die Legendre-Transformierte von $k(\cdot)$.

Wir bemerken, dass die Legendre-Transformierte als das Supremum einer Schar von affinen Funktionen $\{\langle \cdot, x \rangle - k(x)\}$ eine unterhalbstetige konvexe Funktion ist. Wir können daher die Legendre-Transformation auf diese Funktion $k^*(\cdot)$ anwenden und erhalten eine Funktion $k^{**}(\cdot)$ auf dem Vektorraum $V^{**} = V$. Der Satz vom Supremum der affinen Minoranten liefert $k^{**} = k$, denn für jedes $x \in V$ gilt

$$k^{**}(x) = \sup\{\langle \theta, x \rangle - k^*(\theta) : \theta \in V^*\} = \sup\{a(x) : a(\cdot) \text{ ist affine Minorante von } k(\cdot)\}.$$

Wir betrachten einige speziellere Situationen. Insbesondere erläutern wir die Beziehung der Legendre-Transformierten zum Begriff der Stützfunktion einer konvexen Menge:

Wenn K eine abgeschlossene konvexe Menge ist, und $g_K(x) = \begin{cases} 0 & \text{für } x \in K \\ +\infty & \text{für } x \notin K \end{cases}$

dann ist g_K eine unterhalbstetige konvexe Funktion und ihre Legendre-Transformierte ist die Stützfunktion von K . In der Tat gilt

$$g_K^*(\theta) = \sup\{\langle \theta, x \rangle - g_K(x) : x \in V\} = \sup\{\langle \theta, x \rangle : x \in K\}.$$

Hier wird aus einer Funktion, die nur die Werte 0 und $+\infty$ annimmt, eine positiv definite Funktion. Dies und die Umkehrung bringt der folgende Satz zum Ausdruck.

Satz

Wenn $f(\cdot)$ nur die Werte 0 und $+\infty$ annimmt, dann ist $f^*(\cdot)$ positiv homogen.

Wenn $f(\cdot)$ positiv homogen ist, dann nimmt $f^*(\cdot)$ nur die Werte 0 und $+\infty$ an.

Beispiele

1. Sei K das achsenparallele Einheitsquadrat

$$K = \left\{ \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} : |x^1| \leq 1, |x^2| \leq 1 \right\} = \left\{ \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} : \max_i |x^i| \leq 1 \right\}.$$

Die Stützfunktion liefert für die 2-Zeile $\xi = (\xi_1, \xi_2)$ den Wert $\varphi(\xi) = |\xi_1| + |\xi_2|$.

Die Legendre-Transformierte dieser positiv homogenen Funktion $\varphi(\xi)$ hat den Wert 0 auf K und den Wert $+\infty$ auf dem Komplement,

$$\sup\{(\xi_1 x^1 + \xi_2 x^2) - (|\xi_1| + |\xi_2|) : \xi \in V^*\} = \begin{cases} 0 & \text{für } x \in K \\ +\infty & \text{für } x \notin K \end{cases}.$$

2. Sei K der Einheitskreis

$$K_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x^2 + y^2 \leq 1 \right\}.$$

Die Stützfunktion ist

$$k_K^*(\xi, \eta) = \sup\{\xi x + \eta y : x \in K\} = \sqrt{\xi^2 + \eta^2}.$$

Ihre Legendre-Transformierte ist

$$\sup\{\xi x + \eta y - \sqrt{\xi^2 + \eta^2}\} = \begin{cases} 0 & \text{für } x \in K \\ \infty & \text{für } x \notin K \end{cases}.$$

3. Sei $p \geq 1$ und

$$K_p = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : |x|^p + |y|^p \leq 1 \right\}.$$

Die Stützfunktion ergibt sich aus den Überlegungen zu den p -Normen, die wir in III.1 angestellt haben

$$k^*(\xi, \eta) = (|\xi|^q + |\eta|^q)^{1/q} \quad \text{mit} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Eine Teilmenge K eines Vektorraums V heisst ein Kegel (mit der Spitze im Nullpunkt), wenn gilt $\forall v \in V, \alpha > 0 \quad \alpha \cdot v \in K$. Ein Kegel K ist offenbar genau dann ein konvexer Kegel, wenn gilt $\forall v, w \in K \quad v + w \in K$. Man definiert den zum Kegel $K \subseteq V$ dualen Kegel $K^* = \{\theta : \langle \theta, v \rangle \leq 0 \text{ für alle } v \in K\}$.

Corollar (Duale Kegel)

Wenn $k(\cdot)$ auf dem Kegel K verschwindet und sonst den Wert $+\infty$ hat, dann verschwindet die Legendre-Transformierte $k^*(\cdot)$ auf dem dualen Kegel K^* und hat sonst den Wert $+\infty$.

Beweis:

Wenn $\langle \theta, v \rangle > 0$ für ein $v \in K$, dann gilt $k^*(\theta) = \sup\{\langle \theta, v \rangle : v \in K\} = +\infty$.

Wenn $\langle \theta, v \rangle \leq 0$ für alle $v \in K$, dann haben wir $k^*(\theta) = 0$.

Wir haben gesehen: Für jeden Kegel K in einem endlichdimensionalen Vektorraum V ist der Kegel K^{**} die abgeschlossene konvexe Hülle. Diesen Sachverhalt kann man auch folgendermaßen ausdrücken:

Satz

Sei K ein konvexer Kegel im n -dimensionalen Vektorraum. Ein Vektor w liegt genau dann in der abgeschlossenen konvexen Hülle von K , wenn gilt

$$\langle \xi, w \rangle \leq 0 \quad \text{für alle } \xi \text{ mit } \forall v \in K \quad \langle \xi, v \rangle \leq 0.$$

Es sei K der kleinste konvexe Kegel, welcher die Vektoren v_1, v_2, \dots, v_M enthält. Ein Vektor w gehört genau dann zu K , wenn er eine positive Linearkombination der v_j ist; dies ist genau dann der Fall wenn $\langle \xi, w \rangle \leq 0$ gilt für jede Linearform ξ mit $\forall j \quad \langle \xi, v_j \rangle \leq 0$.

Nachtrag : Die Jensen'sche Ungleichung

Satz : Wenn $k(\cdot)$ konvex ist und P der Schwerpunkt einer normierten Massenverteilung

$$P = \sum p_i P_i \quad \text{mit} \quad p_i \geq 0, \quad \sum p_i = 1 \quad .$$

dann gilt

$$k(P) \leq \sum p_i \cdot k(P_i) \quad .$$

(„Der k -Wert im Schwerpunkt ist kleiner oder gleich dem gewichteten Mittel der k -Werte“.)

Bemerke: Wenn die Gewichte auf zwei Punkte konzentriert sind, $P = (1 - \lambda)P_0 + \lambda P_1$, dann ist das die definierende Eigenschaft der Konvexität. Für Gewichtungen, die auf endlich viele Punkte konzentriert sind, kann man die Ungleichung mittels vollständiger Induktion nach der Anzahl der gewichteten Punkte beweisen.

Die Jensen'sche Ungleichung gilt auch in allgemeineren Fällen. Für jede Wahrscheinlichkeitsgewichtung $\mu(\cdot)$ auf L , die einen Schwerpunkt $\text{bar}(\mu)$ besitzt, gilt

$$k(\text{bar}(\mu)) \leq \int k(x) d\mu(x) \quad .$$

Der „Schwerpunkt“ $\text{bar}(\mu)$ kann übrigens als derjenige Punkt P charakterisiert werden, für welchen gilt

$$f(P) = f(\text{bar}(\mu)) = \int f(x) d\mu(x) \quad \text{für alle affinen } f(\cdot) \quad .$$

Eine noch bequemere Formulierung der Jensen'schen Ungleichung gewinnt man in der Sprache der Wahrscheinlichkeitstheorie: Die Stochastiker assoziieren mit einer Wahrscheinlichkeitsgewichtung $\mu(\cdot)$ auf L eine L -wertige Zufallsgröße X mit der Verteilung μ , d. h.

$$\Pr(X \in B) = \mu(B) \quad \text{für alle } B \subseteq L \quad .$$

Die Jensen'sche Ungleichung erhält die Gestalt:

Satz : (Jensen'sche Ungleichung)

Sei $k(\cdot)$ auf L konvex und X eine L -wertige Zufallsgröße, die einen Erwartungswert $\mathcal{E}X$ besitzt. Dann gilt

$$k(\mathcal{E}X) \leq \mathcal{E}(k(X)) \quad .$$

Der Beweis für allgemeine $\mu(\cdot)$ kann (für unterhalbstetige konvexe Funktionen) sehr einfach mit dem Fundamentalsatz $k(\cdot)$ geführt werden: da nach Definition des Schwerpunkts für jede affine Minorante $f(\cdot)$ gilt

$$f(\text{bar}\mu) = \int f(x) d\mu \leq \int k(x) d\mu(x) \quad .$$

haben wir auch für das Supremum der affinen Minoranten $k(\text{bar}\mu) \leq \int k(x) d\mu(x)$. ■

Aufgaben zu III.3

Aufgabe III.3.1 :

Es seien p_j und q_j nichtnegative Zahlen für $j = 1, 2, \dots, n$ mit $\sum p_j = 1 = \sum q_j$.

Zeigen Sie $\sum p_j \cdot \ln \frac{p_j}{q_j} \geq 0$.

Hinweis: Beachten Sie, dass die Funktion $k(x) = x \cdot \ln x$ konvex ist.

Aufgabe III.3.2 :

Es sei $M = \{P_i : i \in I\}$ eine Familie von Punkten in einem n -dimensionalen reellaffinen Raum (L, V) . Der Durchschnitt aller konvexen bzw. affinen Obermengen von M ist die kleinste konvexe bzw. affine Obermenge von M . Man nennt sie die konvexe bzw. affine Hülle von M und bezeichnet sie üblicherweise mit M^{conv} bzw. mit M^{aff} . Zeigen Sie

1. M^{aff} besteht aus den Punkten der Gestalt $P = \sum_{\text{endl}} \lambda_i \cdot P_i$ mit $\lambda_i \in \mathbb{R}$, $\sum \lambda_i = 1$.
2. M^{conv} besteht aus den Punkten der Gestalt $P = \sum_{\text{endl}} \lambda_i \cdot P_i$ mit $\lambda_i \geq 0$, $\sum \lambda_i = 1$.
3. Es existiert eine Teilfamilie $M' = \{P_i : i \in I'\}$ mit $|I'| = n + 1$, sodass gilt

$$\forall P \in L \exists \{\lambda_i : i \in I'\} : \left(\lambda_i \in \mathbb{R} \right) \wedge \left(\sum_{I'} \lambda_i = 1 \right) \wedge \left(\sum \lambda_i \cdot P_i = P \right).$$

Wenn $M^{\text{aff}} = L$, und I' der Forderung gemäß gewählt ist (man sagt, dass sich die Punkte aus M' in allgemeiner Lage befinden), dann sind Koeffizienten für jeden Punkt P eindeutig bestimmt.

4. ('Satz von Caratheodory')
Zu jedem $Q \in M^{\text{conv}}$ existiert eine Teilfamilie $M' = \{P_i : i \in I'\}$ mit $|I'| = n + 1$, sodass gilt

$$\exists \{\lambda_i : i \in I'\} : \left(\lambda_i \geq 0 \right) \wedge \left(\sum_{I'} \lambda_i = 1 \right) \wedge \left(\sum \lambda_i \cdot P_i = Q \right).$$

5. Wenn $|I| \geq n + 2$, dann existiert eine Partition der Indexmenge $I = I_+ + I_-$ und ein Punkt \tilde{P} , der sowohl in $M_+^{\text{conv}} = \{P_i : i \in I_+\}$ als auch in $M_-^{\text{conv}} = \{P_i : i \in I_-\}$ liegt.
6. Diskutieren Sie diese Sachverhalte für ein Quadrupel von Punkten in der Ebene.

Hinweis zu 5.): Betrachte zu den Punkten P_0, P_1, \dots, P_{n+1} die linear abhängigen Vektoren $v_i = \overrightarrow{P_0 P_i}$, $i = 1, \dots, n + 1$. Bei der linearen Abhängigkeit $\sum \alpha_i \cdot v_i = 0$ können wir annehmen $\sum \alpha_i \geq 0$, $\sum \alpha_i^+ = 1$. Wir haben dann

$$\sum \alpha_i^+ \cdot \overrightarrow{P_0 P_i} = \sum \alpha_i^- \cdot \overrightarrow{P_0 P_i}, \quad \sum \alpha_i^+ \cdot P_i = \tilde{P} = \sum \alpha_i^- \cdot P_i + \left(\sum \alpha_i \right) \cdot P_0.$$

Aufgabe III.3.3 :

Berechnen Sie mit den Mitteln der Schulmathematik die Legendre-Transformierten $\varphi_i(\theta) = k_i^*(\theta)$ für die folgenden Funktionen $k_i(x)$:

- $k_1(x) = c + bx + \frac{1}{2}ax^2$ mit $a > 0$,
- $k_2(x) = x - 1 - \ln x$ für $x > 0$ und $= +\infty$ für $x \leq 0$,
- $k_3(x) = x \cdot \ln x$ für $x \geq 0$ und $= +\infty$ für $x < 0$,
- $k_4(x) = |x|$,
- $k_5(x) = \frac{1}{3}(|x-1| + |x| + |x+1|)$

Hinweis: Es mag nützlich sein, zuerst die zweite Aussage der nächsten Aufgabe (über runde konvexe Funktionen) für $n = 1$ zu durchdenken.

Aufgabe III.3.4 :

Es sei $k(x)$ eine konvexe Funktion auf dem Vektorraum V aller n -Spalten x . U sei eine offene konvexe Teilmenge des Endlichkeitsbereichs. Wir sagen, k sei rund auf U , wenn k auf U zweimal stetig differenzierbar ist mit einer positiv definiten Hesse-Matrix $k''(x)$ für alle $x \in U$.

Zeigen Sie: Wenn die konvexe Funktion $k(x)$ auf U rund ist, dann ist die Legendre-Transformierte $\varphi(\theta) = k^*(\theta)$ rund auf der Menge U^* , wo U^* das Bild von U bzgl. der Abbildung $U \ni x \longrightarrow k'(x)$ ('Gradientenabbildung') ist.

Zeigen Sie weiter, dass $U^* \ni \theta \longrightarrow \varphi'(\theta)$ die Umkehrung von k' ist.

Rückblick : Verschiedene Typen von Funktionen

Durch unsere Betrachtungen zur Konvexität mag sich mancher Leser an die Begriffsbildungen der elementaren Analysis erinnert fühlen. Wir wollen hier aber die analytischen Aspekte nicht weiter verfolgen.

Wir sind auf spezielle Typen von Funktionen gestossen, und erinnern zum Abschluss dieses Abschnitts III daran, wie die jeweiligen Definitionsbereiche aussehen und welche Werte die bisher studierten Funktionen annehmen.

- a) Eine Metrik (auf S) ist eine \mathbb{R}_+ -wertige Funktion auf $S \times S$.
- b) Eine Norm (auf V) ist eine \mathbb{R}_+ -wertige Funktion auf V .
- c) Eine konvexe Funktion $k(\cdot)$ ist eine $\mathbb{R} \cup \{+\infty\}$ -wertige Funktion auf einem affinen Raum (L, V) .
- d) Ihre Legendre-Transformierte $k^*(\cdot)$ ist eine $\mathbb{R} \cup \{+\infty\}$ -wertige Funktion auf V^* . (Bei der Konstruktion wird ein Punkt $\in L$ als der Nullpunkt ausgezeichnet.)

Früher schon haben wir gesehen:

- a) Eine affine Funktion ist eine \mathbb{K} -wertige Funktion auf einem \mathbb{K} -affinen Raum.
- b) Eine Linearform ist eine \mathbb{K} -wertige Funktion auf einem \mathbb{K} -Vektorraum.
- c) Eine quadratische Form ist eine \mathbb{K} -wertige Funktion auf einem \mathbb{K} -Vektorraum. Zu ihr gehört ($1 + 1 \neq 0$ in \mathbb{K}) eine symmetrische Bilinearform auf $V \times V$.

Dazu werden wir noch lernen: Eine hermitesche Form (siehe IV.2) ist eine reellwertige Funktion auf einem \mathbb{C} -Vektorraum V . Zu ihr gehört eine hermitesche Sesquilinearform auf $V \times V$. Sie nimmt komplexe Werte an.

Noch früher haben wir gesehen:

- a) Ein Polynom mit komplexen Koeffizienten ist zunächst ein formaler Ausdruck, eine Rechengröße. Man kann mit ihr aber auch eine „ganz rationale“ Funktion assoziieren (welche auf \mathbb{C} komplexe Werte annimmt). Ihr Absolutbetrag ist eine nichtnegative Funktion auf \mathbb{C} .
- b) Eine gebrochenrationale Funktion $\frac{p(z)}{q(z)}$ kann man als eine $\bar{\mathbb{C}}$ -wertige Funktion auf $\bar{\mathbb{C}}$ verstehen.
- c) Bei den lineargebrochenen „Funktionen“ denken wir an Abbildungen von $\bar{\mathbb{C}}$ auf sich (Möbiustransformationen).

Ein Typ von Funktionen, welcher dem Funktionsbegriff der Schulmathematik nahe steht, sind die trigonometrischen Polynome. Diese Funktionen leben auf $\mathbb{R}/2\pi$ und nehmen komplexe Werte an. Trigonometrische Polynome kann man auch durch ihre Koeffizientenfolgen beschreiben; dabei handelt es sich sozusagen um komplexwertige Funktionen auf \mathbb{Z} . In den p -summablen Folgen haben wir weitere komplexwertige Funktionen auf \mathbb{Z} .

kennengelernt.

Der Anfänger sollte sich nicht dadurch verwirren lassen, dass wir uns manchmal für individuelle Funktionen (oder Abbildungen) interessiert haben, und manchmal eher für Funktionenräume (meistens Vektorräume von Funktionen) oder für Mengen von Abbildungen (insbesondere Gruppen von Abbildungen).

IV Inneres Produkt. Orthogonalität

Themenübersicht IV

In der euklidischen Geometrie kann man nicht nur von der Länge einer Strecke sprechen, d.h. von der Norm eines Verschiebungsvektors; man hat auch den Begriff des Winkels zwischen Vektoren. Der Cosinus des Winkels ergibt sich bekanntlich aus dem inneren Produkt

$$\langle \mathbf{v} | \mathbf{w} \rangle = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos \angle .$$

Auch in gewissen normierten Vektorräumen unendlicher Dimension kann man ein inneres Produkt definieren. Dies ist dann möglich, wenn die Norm die sog. Parallelogrammgleichung erfüllt.

$$\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2 \cdot \|\mathbf{v}\|^2 + 2 \cdot \|\mathbf{w}\|^2 .$$

Man nennt einen solchen normierten Raum einen Prä-Hilbertraum oder im endlichdimensionalen reellen Fall eine euklidischen Vektorraum.

Einen Hilbertraum gewinnt man aus einem Prä-Hilbertraum durch einen Vervollständigungsprozess, den wir später behandeln werden. Hier diskutieren wir diejenigen Aspekte des inneren Produkts, die keinen wesentlichen Gebrauch von der Vollständigkeit machen.

IV.1 Parallelogrammgleichung und Polarisierung

Beispiele zur Parallelogrammgleichung: die euklidische Norm im Anschauungsraum, die 2-Norm im Raum der trigonometrischen Polynome. Die Norm, die zu einer positiv definiten Matrix $J \times J$ -Matrix gehört. Die Ungleichung von Cauchy-Schwarz. Polarisierungskonstruktion. Der Satz von Jordan und von Neumann. Das Orthogonalisierungsverfahren nach Gram-Schmidt. ONBasen und orthonormale Koordinaten. bra- und ket-Vektoren. Didaktische Anmerkungen zum euklidischen Anschauungsraum.

IV.2 Hauptachsen, Orthogonale Zerlegungen, Spektralsatz

Das Theorem von den Hauptachsen einer hermiteschen Form. Beweis mit Courants Maximierungsargument. Orthogonale Projektionen. Simultane Diagonalisierung zweier hermitescher Formen, wo eine positiv definit ist. Auf einem endlichdimensionalen Hilbertraum gibt es eine natürliche Entsprechung zwischen den hermiteschen Formen und den selbstadjungierten Operatoren. Der Satz von den Hauptachsen entspricht dem Spektralsatz für selbstadjungierte Operatoren. In orthonormalen Koordinaten werden diese Objekte durch die hermiteschen Matrizen beschrieben. Das Theorem von den Hauptachsen einer hermiteschen Form entspricht dem Satz von der unitären Diagonalisierbarkeit einer hermiteschen Matrix.

Nicht nur die hermiteschen (und die unitären) Matrizen sind unitär diagonalisierbar. Genau dann gibt es zur komplexen Matrix N eine unitäre Matrix U mit $U^* N U = D$ (Diagonalmatrix), wenn N ‘normal’ ist in dem Sinne $N^* N = N N^*$.

IV.3 Einparametrische Matrixgruppen

Der infinitesimale Operator einer einparametrischen Matrixgruppe. Gleichförmige Drehungen im Anschauungsraum. Einparametrische Gruppen von 2×2 -Matrizen. Einfache und doppelte Nullstellen des charakteristischen Polynoms. In einem Anhang stellen wir zusammen, wie die Matrizen recht verschiedene Rollen spielen können. Quaternionen.

IV.4 Gekoppelte Oszillatoren

Der einfache harmonische Oszillator.

$$\begin{aligned} m\ddot{x} + r \cdot \dot{x} + k \cdot x &= f(t) \\ L \cdot \ddot{Q} + R\dot{Q} + \frac{1}{C} \cdot Q &= U(t) . \end{aligned}$$

Die allgemeine Lösung der homogenen Gleichung (im schwach gedämpften Fall). Zwei schwach gekoppelte Oszillatoren ohne Reibung und ohne anregende Kraft. Die *normal modes* im symmetrischen Fall. Die Lösungen sind fastperiodische Funktionen. „Resonanz“. Drei Oszillatoren in symmetrischer Anordnung. Die *normal modes* für n identische Oszillatoren in ringförmiger Anordnung. Stehende und laufende Wellen auf einem Ring.

IV.5 Die schwingende Saite

Die Gleichung der schwingenden Saite und die Gleichung der schwingenden Membrane. Separationsansatz für die Wellengleichung. Ebene Wellen. Intuitives zur Wellenbewegung: Wellenflächen. Ausbreitungsgeschwindigkeit. Wellenzahl, Frequenz, Phasengeschwindigkeit. Die Eigenschwingungen. Historisches zur schwingenden Saite und zur Fourier-Analyse.

IV.6 Diskrete Fourier-Transformation

Die diskrete Fourier-Transformation DFT als Übergang zu einem neuen orthonormierten Koordinatensystem. Die unitäre Matrix mit den Einträgen

$$u_k^j = \frac{1}{\sqrt{n}} \cdot \exp(i \frac{2\pi}{n} kj) = \frac{1}{\sqrt{n}} \cdot \zeta_{(n)}^j ,$$

wo die $\zeta_{(n)}$ die n -ten Einheitswurzeln sind.

Die DFT als Approximation zu den üblichen kontinuierlichen Fourier-Transformationen. Shift-Operator und Differenzenquotient. Die Schnelle Fourier-Transformation FFT. Die Funktion $f(t)$ werde im Zeitbereich $(-T, T)$ in Zeitabständen $\frac{2T}{n}$ abgetastet; die DFT $y(\omega)$ ist dann 2Ω -periodisch, $\Omega = \frac{\pi}{T} \cdot \frac{n}{2}$. Das Periodogramm zeigt, wie die ‘Energie’ auf das Intervall $(-\Omega, \Omega)$ verteilt ist. Der Effekt des Herausschneidens ist abzumildern.

IV.7 Drehungen im Minkowskiraum

Der dreidimensionale Vektorraum \mathbb{P} der spurlosen hermiteschen 2×2 -Matrizen erweist sich als ein nützliches Modell des dreidimensionalen euklidischen Anschauungsraums; der vierdimensionale Vektorraum \mathbb{M} aller hermiteschen 2×2 -Matrizen liefert ein Modell für

den Minkowskiraum der speziellen Relativitätstheorie. Das ‘Normquadrat’ und das ‘innere Produkt’ kann mit Spur- und Determinantenbildung zum Ausdruck gebracht werden.

Wir verschaffen uns in diesen Modellen eine Übersicht über die Drehungen des Anschauungsraums und die eigentlichen Lorentztransformationen. Die Gruppe $SU(2)$ wirkt auf \mathbb{P} ; die Gruppe $SL(2, \mathbb{C})$ wirkt auf \mathbb{M} . Dabei erweist sich $SU(2)$ als eine zweifache Überlagerung der Drehungsgruppe $SO(3, \mathbb{R})$ und $SL(2, \mathbb{C})$ als eine zweifache Überlagerung der eigentlichen Lorentzgruppe (die man auch als die Gruppe der Drehungen des Minkowskiraums bezeichnet).

Wir betrachten auch die Quaternionen und ihre Darstellung durch 2×2 -Matrizen.

IV.8 Zerlegung einer Gruppendarstellung

Eine Gruppendarstellung (group representation) ist eine spezielle Art von Gruppenwirkung (group action): jedem Gruppenelement wird ein Endomorphismus eines Vektorraums zugeordnet. (Bei uns geht es um endliche Gruppen und um endlichdimensionale komplexe Vektorräume). Jede Darstellung kann man als direkte Summe von irreduziblen Darstellungen gewinnen, wobei die Vielfachheiten der verschiedenen irreduziblen Teildarstellungen eindeutig bestimmt sind. In einer passenden Basis betrachtet erscheinen also die Endomorphismen als Blockmatrizen. Der Begriff der direkten Zerlegung eines \mathbb{K} -Vektorraums und der Begriff der orthogonalen Zerlegung eines endlichdimensionalen Hilbertraums wird nochmals von Grunde auf entwickelt.

IV.9 Spurfunktionen, die Charaktertafel der Gruppe G .

Eine endliche Gruppe G besitzt nur endlichviele nichtisomorphe Darstellungen; die Anzahl ist gleich der Anzahl der Konjugationsklassen der Gruppe. Bei kommutativen Gruppen sind alle irreduziblen Darstellungen eindimensional; sie entsprechen den Gruppencharakteren. Wir bemerken einige Analogien zur Fourier-Transformation für die (unendlichen!) Gruppen \mathbb{Z} , $\mathbb{R}/2\pi$ und \mathbb{R} .

Bei einer beliebigen endlichen Gruppe G bilden die Spurfunktionen der irreduziblen Darstellungen eine Orthonormalbasis in einem interessanten Vektorraum komplexwertiger Funktionen auf G .

IV Inneres Produkt. Orthogonalität

IV.1 Parallelogrammgleichung und Polarisierung

Einen normierten (reellen oder komplexen) Vektorraum $(V, \|\cdot\|)$ nennt man auch einen Prä-Banachraum. Man nennt ihn einen **Prä-Hilbertraum**, wenn das Normquadrat die sog. Parallelogrammgleichung erfüllt

$$\|v + w\|^2 + \|v - w\|^2 = 2 \cdot \|v\|^2 + 2\|w\|^2 \quad (\text{Parallelogrammgleichung})$$

Eine Norm, welche die Parallelogrammgleichung erfüllt, heisst eine **Hilbertraumnorm**.

Hinweis : In der Funktionalanalysis braucht man Banachräume und speziell Hilberträume. Diese normierten Vektorräume entstehen durch Vervollständigung aus Prä-Banachräumen bzw. Prä-Hilberträumen. Im endlichdimensionalen Fall braucht man keine Vervollständigung, die endlichdimensionalen Prä-Banachräume sind von vorneherein vollständig.

Satz (1. Beispiel)

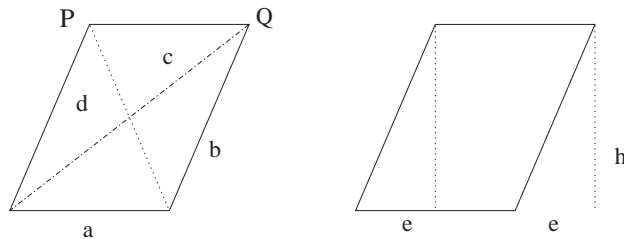
Die euklidische Norm im Anschauungsraum erfüllt die Parallelogrammgleichung.

1. Beweis

Der erste Beweis baut elementargeometrisch auf dem Satz von Pythagoras auf.

Betrachten wir ein Parallelogramm mit den Seitenlängen a, b . Die Längen der Diagonalen seien c und d . Wir zeigen

$$c^2 + d^2 = 2a^2 + 2b^2$$



Wir fällen das Lot von P und Q auf die Grundlinie und haben nach dem Satz von Pythagoras

$$c^2 = (a + e)^2 + h^2, \quad d^2 = (a - e)^2 + h^2, \quad b^2 = e^2 + h^2.$$

Addition der beiden ersten Gleichungen ergibt

$$c^2 + d^2 = 2a^2 + 2e^2 + 2h^2 = 2a^2 + 2b^2.$$

2. Beweis

Wir deuten die Strecken als komplexe Zahlen. Zu zeigen ist

$$|z + w|^2 + |z - w|^2 = 2|z|^2 + 2|w|^2.$$

Das ergibt sich sofort, wenn man die komplex konjugierten Zahlen ins Spiel bringt

$$(z + w)(\bar{z} + \bar{w}) + (z - w)(\bar{z} - \bar{w}) = z\bar{z} + w\bar{w} + z\bar{z} + w\bar{w}.$$

Die gemischten Produkte fallen weg. Die folgende Rechnung werden wir später brauchen; sie ist der einfachste Fall der sog. Polarisierungskonstruktion (im Komplexen).

Sind z, w komplexe Zahlen, so gilt

$$\begin{aligned} |z + w|^2 - |z - w|^2 &= 2z\bar{w} + 2\bar{z}w = 4 \cdot \Re(\bar{z}w) \\ |z + w|^2 - |z - w|^2 + i|z - iw|^2 - i|z + iw|^2 &= 4 \cdot \bar{z}w. \end{aligned}$$

Ein zweites Beispiel für den Begriff des Prä-Hilbertraums ist der Vektorraum der trigonometrischen Polynome mit der 2-Norm.

Satz (2. Beispiel: trigonometrische Polynome)

Betrachte trigonometrische Polynome

$$f(t) = \sum c_n e^{int}, \quad g(t) = \sum d_n e^{int}$$

mit den Normen

$$\begin{aligned} \|f\| &= \left(\frac{1}{2\pi} \int |f(t)|^2 dt \right)^{1/2} = \left(\sum |c_n|^2 \right)^{1/2} \text{ bzw.} \\ \|g\| &= \left(\frac{1}{2\pi} \int |g(t)|^2 dt \right)^{1/2} = \left(\sum |d_n|^2 \right)^{1/2}. \end{aligned}$$

Es gilt die Parallelogrammgleichung

$$\begin{aligned} \|f + g\|^2 + \|f - g\|^2 &= 2\|f\|^2 + 2\|g\|^2, \\ \sum |c_n + d_n|^2 + \sum |c_n - d_n|^2 &= 2 \sum |c_n|^2 + 2 \sum |d_n|^2. \end{aligned}$$

Der Beweis ist trivial. Auf der Grundlage der Feststellung, dass die komplexen Zahlen $f(t), g(t)$ (für alle t) und die komplexen Zahlen c_n, d_n (für alle n) die entsprechenden Gleichungen erfüllen, findet man auch

$$\begin{aligned} \|f + g\|^2 - \|f - g\|^2 + i \cdot \|f - ig\|^2 - i \cdot \|f + ig\|^2 &= \\ = 4 \cdot \frac{1}{2\pi} \int (\bar{f} \cdot g)(t) dt &= 4 \cdot \sum \bar{c}_n d_n. \end{aligned}$$

Das Ergebnis wird üblicherweise mit $4 \cdot \langle f | g \rangle$ bezeichnet. Man nennt es das **innere Produkt** von f und g (bzgl. der Hilbertraumnorm im Raum der trigonometrischen Polynome).

Es ist darauf hinzuweisen, dass die trigonometrischen Polynome keinen Hilbertraum bilden; die Vervollständigung dieses Raums wird in der Analysis ausführlich diskutiert.

Hermitische Sesquilinearformen

Definition

Eine komplexwertige Funktion $\Phi(\cdot, \cdot)$ auf $V \times V$ heißt eine **hermitische Sesquilinearform** (auf V), wenn gilt

- (i) $\Phi(v, w) = \overline{\Phi(w, v)}$
- (ii) $\Phi(v, \alpha w) = \alpha \cdot \Phi(v, w)$
- (iii) $\Phi(v, w_1 + w_2) = \Phi(v, w_1) + \Phi(v, w_2)$.

Die reellwertige Funktion $\Psi(v) = \Phi(v, v)$ heißt die **hermitische Form** zu $\Phi(\cdot, \cdot)$.

Wenn $\Psi(v) > 0$ für alle $v \neq 0$, dann heisst Φ eine **positiv definite Sesquilinearform** oder auch ein **inneres Produkt**.

Hilfssatz (Ungleichung von Cauchy-Schwarz)

Wenn $\Phi(\cdot, \cdot)$ eine positiv definite Sesquilinearform ist und $\Psi(\cdot)$ die hermitische Form dazu, dann gilt

$$|\Phi(v, w)|^2 \leq \Psi(v) \cdot \Psi(w) \quad \text{für alle } v, w.$$

Die Quadratwurzel von Ψ liefert eine Hilbertraumnorm.

$$\begin{aligned} \sqrt{\Psi(\alpha v)} &= |\alpha| \sqrt{\Psi(v)}, & \sqrt{\Psi(v+w)} &\leq \sqrt{\Psi(v)} + \sqrt{\Psi(w)}; \\ \Psi(v+w) + \Psi(v-w) &= 2\Psi(v) + 2\Psi(w). \end{aligned}$$

Beweis

Wir können o. B. d. A. annehmen, dass $\Phi(v, w)$ reell ist; denn die beteiligten Grössen ändern sich nicht, wenn man v durch $v \cdot e^{i\phi}$ ersetzt. Für alle reellen λ haben wir

$$0 \leq \Phi(v + \lambda w, v + \lambda w) = \Psi(v) + 2\lambda b + \lambda^2 \Psi(w) \quad \text{mit } b = \Phi(v, w) = \Phi(w, v).$$

Schon auf der Schule lernt man, dass $b^2 \leq ac$ gilt für eine quadratische Funktion $q(\lambda) = a\lambda^2 + 2b\lambda + c$, die nirgends auf \mathbb{R} negativ ist. Dies ergibt die Schwarz'sche Ungleichung.

$$\begin{aligned} \Psi(v+w) &= \Psi(v) + \Psi(w) + 2\Re \Phi(v, w), \\ \Psi(v+w) &\leq \Psi(v) + \Psi(w) + 2\sqrt{\Psi(v)}\sqrt{\Psi(w)} = (\sqrt{\Psi(v)} + \sqrt{\Psi(w)})^2. \end{aligned}$$

Satz

Eine hermitische Sesquilinearform ist durch die von ihr erzeugte hermitische Form eindeutig bestimmt.

Beweis : Die Polarisierung im hermitischen Fall ist ein wenig komplizierter als die Polarisierung bei quadratischen Formen. Es gilt

$$\begin{aligned} \Psi(v + \alpha w) - \Psi(v - \alpha w) &= 4\Re(\alpha b) \quad \text{mit } b = \Phi(v, w) \\ \frac{1}{4} \left(\Psi(v + \alpha w) - \Psi(v - \alpha w) - i\Psi(v + i\alpha w) + \Psi(v - i\alpha w) \right) &= \alpha b = \Phi(v, \alpha w). \end{aligned}$$

Satz Eine hermitesche $J \times J$ -Matrix H liefert eine hermitesche Sesquilinearform auf dem Vektorraum aller J -Spalten, wenn man definiert

$$\Phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^* H \mathbf{y} .$$

Sei $\Phi(\cdot, \cdot)$ eine hermitesche Sesquilinearform auf einem Vektorraum V mit einer ausgezeichneten Basis $\{\mathbf{v}_j : j \in J\}$. Dann gibt es genau eine hermitesche Matrix H , sodass gilt

$$\Phi\left(\sum_k \mathbf{y}^k \mathbf{v}_k, \sum_j \mathbf{x}^j \mathbf{v}_j\right) = \sum_{k,j} \bar{\mathbf{y}}^k h_{kj} \mathbf{x}^j .$$

Sei $\{\mathbf{u}_i : i \in I\}$ eine weitere Basis und A die $I \times J$ -Matrix des Basiswechsels

$$\mathbf{v}_j = \sum_i \mathbf{u}_i a_j^i .$$

Für die dazugehörige hermitesche $I \times I$ -Matrix K gilt dann

$$H = A^* \cdot K \cdot A .$$

Der Beweis liegt auf der Hand. Die $J \times J$ -Matrix H mit den Einträgen $h_{kj} = \Phi(\mathbf{v}_k, \mathbf{v}_j)$ heisst die **Gram-Matrix** von Φ bezüglich der betreffenden Basis.

Orthonormalbasen

Die Prä-Hilberträume sind die komplexen Vektorräume mit einer ausgezeichneten positiv definiten Sesquilinearform, (die man das innere Produkt nennt und mit $\langle \cdot | \cdot \rangle$ bezeichnet.) $\|\mathbf{v}\|^2 = \langle \mathbf{v} | \mathbf{v} \rangle$ Das innere Produkt eröffnet die Möglichkeit, in einem unendlichdimensionalen (Prä)-Hilbertraum Basen (oder Koordinatensysteme) einzuführen, mit welchen man fast genau so bequem operieren kann wie mit den Basen in endlichdimensionalen Vektorräumen. Diese besonderen Basen heissen Orthonormalbasen.

Definition

Eine Familie von Vektoren in einem (Prä-)Hilbertraum $(\mathbf{e}_n)_{n \in I}$ heisst ein **Orthonormalsystem** (kurz ein ONS), wenn gilt

$$\langle \mathbf{e}_n | \mathbf{e}_m \rangle = \begin{cases} 1 & \text{falls } n = m \\ 0 & \text{falls } n \neq m \end{cases}$$

Man spricht von einer Orthonormalbasis ONB im Prä-Hilbertraum V , wenn die Vervollständigung des aufgespannten Vektorraums gleich der Vervollständigung von V ist.

Satz (Orthonormalisierungsverfahren nach Gram-Schmidt)

Bezeichne $\langle \cdot | \cdot \rangle$ das innere Produkt in einem Prä-Hilbertraum V .

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$ seien linear unabhängige Vektoren. Es existiert dann ein ONS $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots$, sodass $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ denselben Teilraum aufspannt wie $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ (für jedes m).

Beweis

Es gibt ein Vielfaches von v_1 , welches die Länge 1 hat. e_1 sei ein solches. Wir konstruieren die e_m rekursiv. $\{e_1, \dots, e_m\}$ seien bereits konstruiert. v_{m+1} ist nicht Linearkombination der v_j mit $j \leq m$ und daher auch nicht Linearkombination der e_j mit $j \leq m$.

Sei $b_j = \langle e_j | v_{m+1} \rangle$ für $j = 1, \dots, m$ und $\tilde{e}_{m+1} = v_{m+1} - \sum_{j=1}^m b_j e_j$.

Es gilt $\langle e_k | \tilde{e}_{m+1} \rangle = \langle e_k | v_{m+1} \rangle - \sum_{j=1}^m b_j \langle e_k | e_j \rangle = \langle e_k | v_{m+1} \rangle - b_k = 0$ für $k = 1, \dots, m$.

Man wähle nun e_{m+1} als ein Vielfaches von \tilde{e}_{m+1} mit der Norm 1.

— Und so geht das weiter.

Bemerkungen

- Die orthonormalen Vektoren e_1, e_2, \dots sind bis auf Faktoren vom Betrag 1 festgelegt. Da e_n eine Linearkombination der ‘früheren’ v_k ($k \leq n$) ist, liefert uns das Schmidt’sche Orthonormalisierungsverfahren eine obere Dreiecksmatrix M , sodass

$$(e_1, e_2, e_3, \dots) = (v_1, v_2, v_3, \dots) \cdot M.$$

- Der Raum \mathbb{C}_{sp}^J aller komplexen J -Spalten wird bekanntlich zu einem Hilbertraum, indem man eine positiv definite $J \times J$ -Matrix H auszeichnet und damit das innere Produkt definiert. Eine ONBasis ist ein n -tupel von J -Spalten, welches, zu einer Matrix B zusammengefasst, liefert $B^*HB = I$ (Einheitsmatrix).
- Wir haben eben gesehen, dass es zu jeder positivdefiniten Matrix H und zu jeder nichtsingulären Matrix A eine obere Dreiecksmatrix M gibt, sodass für $B = AM$ gilt $B^*HB = I$. Zu jeder positivdefiniten Matrix H gibt es eine normierte obere Dreiecksmatrix N sodass gilt $N^*HN = D$ (Diagonalmatrix). Eine Verallgemeinerung dieses Satzes wird uns beim gaus’schen Eliminationsverfahren begegnen. Er hat natürlich nichts zu tun mit dem ‘Satz von der Hauptachsentransformation’, nach welchem zu jeder hermiteschen Matrix H eine unitäre Matrix U existiert, sodass $U^*HU = D$ (Diagonalmatrix).

‘Stetige’ hermitische Formen

Aus einem Grund, der sofort ersichtlich werden wird, nennen wir die hermiteschen Formen, die von einer hermiteschen Sesquilinearform erzeugt werden, stetige hermitesche Formen. Es darf vermutet werden, dass es ‘unstetige’ hermitesche Formen nirgendwo bedeutungsvoll sind. Die Grundlagentheoretiker bestehen darauf, dass es sie gibt, wenn man sie auch nicht wirklich konstruieren kann. Wir brauchen die ‘Stetigkeit’, um lückenlose Beweise zu führen.

Definition Sei V ein komplexer Vektorraum. Eine stetige reellwertige Funktion $\Psi(\cdot)$ auf V heißt eine **hermitesche Form**, wenn gilt

- (i) $\Psi(\alpha v) = |\alpha|^2 \cdot \Psi(v)$ für alle $\alpha \in \mathbb{C}$, $v \in V$
- (ii) $\Psi(v + w) + \Psi(v - w) = 2 \cdot \Psi(v) + 2 \cdot \Psi(w)$ für alle v, w .

Die hermitesche Form $\Psi(\cdot)$ heisst stetig, wenn $\Psi(v + \alpha w)$ eine stetige Funktion von $\alpha \in \mathbb{C}$ ist, für jedes Paar v, w .

Bemerke: Jede hermitesche Form, die von einer hermiteschen Sesquilinearform erzeugt wird, ist stetig; wir haben nämlich

$$\Psi(v + \alpha w) = a + |\alpha|^2 c + 2\Re(\alpha b) \quad \text{mit Konstanten } a, b, c.$$

Das Quadrat einer Hilbertraumnorm ist eine stetige hermitesche Form. Die Dreiecksungleichung zusammen mit $\|\alpha w\| = |\alpha| \|w\|$ impliziert die Stetigkeit.

Satz

Zu jeder stetigen hermiteschen Form $\Psi(\cdot)$ gibt es genau eine Sesquilinearform $\Phi(\cdot)$ mit $\Phi(v, v) = \Psi(v)$.

Beweis

Zu $\Psi(\cdot)$ besorgen wir uns mit der Polarisierungskonstruktion eine Funktion $\Phi(\cdot, \cdot)$ mit $\Phi(v, v) = \Psi(v)$. Es ist nachzuweisen, dass $\Phi(\cdot, \cdot)$ eine Sesquilinearform ist.

$$\Phi(v, w) = \frac{1}{4} (\Psi(v + w) - \Psi(v - w) + i \cdot \Psi(v - iw) - i \cdot \Psi(v + iw)) .$$

1. Wir zeigen $\Phi(v, w) = \overline{\Phi(w, v)}$. Dazu müssen wir nur bemerken

$$\Psi(v - iw) = \Psi(i(v - iw)) = \Psi(w + iv) , \quad \Psi(v + iw) = \Psi(w - iv) .$$

Diese Beobachtung liefert auch $\Phi(v, iw) = i \cdot \Phi(v, w)$.

2. Wir zeigen

$$\Phi(u, v) + \Phi(u, w) = 2 \cdot \Phi\left(u, \frac{1}{2}(v + w)\right) .$$

Setze $m = u + \frac{1}{2}(v + w)$, $d = \frac{1}{2}(v - w)$. Wegen $u + v = m + d$, $u + w = m - d$ und der Parallelogrammgleichung ergibt sich

$$\begin{aligned} \Psi(u + v) + \Psi(u + w) &= 2\Psi(m) + 2\Psi(d) \\ \Psi(u - v) + \Psi(u - w) &= 2 \cdot \Psi\left(u - \frac{1}{2}(v + w)\right) + 2\Psi(d) . \end{aligned}$$

Die Differenz liefert

$$\begin{aligned} &\left[\Psi(u + v) - \Psi(u - v) \right] + \left[\Psi(u + w) - \Psi(u - w) \right] = \\ &= 2 \cdot \Psi\left(u + \frac{1}{2}(v + w)\right) - 2 \cdot \Psi\left(u - \frac{1}{2}(v + w)\right) . \end{aligned}$$

Wir addieren die entsprechende Formel mit iv, iw und erhalten

$$\left[4 \cdot \Phi(u, v) \right] + \left[4 \cdot \Phi(u, w) \right] = 8 \cdot \Phi\left(u, \frac{1}{2}(v + w)\right) .$$

Die Rechnung ergibt also

$$\Phi(u, v) = 2 \cdot \Phi\left(u, \frac{1}{2}v\right) \quad \text{und} \quad \Phi(u, v + w) = \Phi(u, v) + \Phi(u, w) .$$

3. Aus der Beobachtung, dass man i aus dem zweiten Argument herausziehen kann, ergibt sich

$$\Phi(u, \alpha v) = \alpha \cdot \Phi(u, v)$$

für alle α von der Gestalt $\alpha = \frac{1}{2^n}(k + i\ell)$, $n, k, \ell \in \mathbb{Z}$. Für jedes feste Paar u, v ist also $\alpha \mapsto \Phi(u, \alpha v)$ eine additive Funktion $g(\alpha)$, welche auf der Menge $\{\alpha = \frac{1}{2^n}(k + i\ell)\}$ mit einer linearen Funktion $c \cdot \alpha$ übereinstimmt. Die Analytiker bestehen darauf, dass es auf \mathbb{R} (oder \mathbb{C}) additive Funktionen gibt, die nicht linear sind; wir brauchen eine Stetigkeitsannahme (oder etwas ähnliches), um zu garantieren, dass $\Phi(u, \lambda v) = \lambda \Phi(u, v)$ gilt für alle komplexen Zahlen λ .

Zusammen mit der obigen Bemerkung über Hilbertraumnormen haben wir nun also einen vollständigen Beweis für das berühmte

Theorem (Jordan und v. Neumann)

Sei $(V, \|\cdot\|)$ ein Prä-Hilbertraum. Für $v, w \in V$ definiert man

$$\langle v | w \rangle := \frac{1}{4}\|v + w\|^2 - \frac{1}{4}\|v - w\|^2 + \frac{i}{4}\|v - iw\|^2 - \frac{i}{4}\|v + iw\|^2.$$

Es gilt dann

- (i) $\langle v | w \rangle = \overline{\langle w | v \rangle}$
- (ii) $\langle u | v + w \rangle = \langle u | v \rangle + \langle u | w \rangle$
- (iii) $\langle v | \alpha w \rangle = \alpha \langle v | w \rangle$ für alle $\alpha \in \mathbb{C}$
- (iv) $\langle v | v \rangle = \|v\|^2 > 0$ für alle $v \neq 0$

Man nennt $\langle \cdot | \cdot \rangle$ das **innere Produkt** zur Prä-Hilbertraum-Norm $\|\cdot\|$.

Linearformen mit endlicher Norm

Es sei $(V, \langle \cdot | \cdot \rangle)$ ein Prä-Hilbertraum. Eine Linearform ℓ heisst eine beschränkte Linearform, wenn $\ell(\cdot)$ auf der Einheitskugel beschränkt ist; und man definiert

$$\|\ell\| = \sup\{|\ell(v)| : \|v\| \leq 1\} = \sup\left\{\frac{|\ell(v)|}{\|v\|} : v \in V\right\}.$$

Auf dem Prä-Hilbertraum der trigonometrischen Polynome gibt es viele interessante Beispiele für Linearformen mit unendlicher Norm. Beispielweise ist die Auswertung in einem Punkt ein unbeschränktes lineares Funktional. Man betrachte etwa die Dirichlet-Funktionen $D_N = \sum_{-N}^N e^{int}$. Die Norm ist $\sqrt{2N+1}$; die Auswertung im Nullpunkt liefert die (mit $N \rightarrow \infty$) viel schneller wachsenden Werte $2N+1$.

Satz

Es sei $(V, \langle \cdot | \cdot \rangle)$ ein Prä-Hilbertraum. Für jedes $v \in V$ ist dann $\langle v | \cdot \rangle$ eine Linearform ℓ mit $\|\ell\| = \|v\|$.

Der Beweis ergibt sich in trivialer Weise aus der Schwarz'schen Ungleichung.

Satz

Die Menge aller beschränkten Linearformen auf einem (Prä-)Hilbertraum V ist ein Hilbertraum. Jede beschränkte Linearform entspricht einem Element in der Vervollständigung von V .

Uns fehlen hier die Mittel für einen Beweis im unendlichdimensionalen Fall.

Hinweis Wir wollen die Situation in dem uns wohlbekannten Beispiel der trigonometrischen Polynome schildern. (Das Beispiel ist durchaus typisch!)

Sei $(c_n)_{n \in \mathbb{Z}}$ eine quadratsummable Zahlenfolge, $C = \sqrt{\sum |c_n|^2} < \infty$. Für die trigonometrischen Polynome $f(t) = \sum a_n e^{int}$ definieren wir $\ell(f) = \sum \bar{c}_n a_n$. Nach der Schwarz'schen Ungleichung gilt $|\ell(f)| = |\sum \bar{c}_n a_n| \leq C \|f\|$. Die Folge $(c_n)_{n \in \mathbb{Z}}$ liefert also eine beschränkte Linearform ℓ mit Norm $\leq C$. Die Norm ist in der Tat gleich C , wie man sich leicht überlegt. Man kann sich auch leicht überlegen, dass es keine weiteren beschränkten Linearformen auf dem Prä-Hilbertraum V der trigonometrischen Polynome gibt.

Für die finiten Folgen ist C die Hilbertraumnorm des trigonometrischen Polynoms $g(t) = \sum c_n e^{int}$ und man kann die Auswertung des linearen Funktionalis auch durch ein Integral beschreiben:

$$\ell(f) = \langle g | f \rangle = \frac{1}{2\pi} \int \bar{g}(t) f(t) dt \quad \text{für alle } f \in V.$$

In der Lebesgue'schen Integrationstheorie wird geklärt, dass es für beliebige quadratintegrierbare Folgen $(c_n)_{n \in \mathbb{Z}}$ so etwas wie eine Funktion (oder vielmehr eine Äquivalenzklasse von Funktionen) gibt, sodass $\ell(f)$ durch ein Integral $\frac{1}{2\pi} \int \bar{g} \cdot f$ ausgedrückt werden kann. Die Art von Funktionen g , die man so zu einer quadratsummablen Folge gewinnt, heißen die 2π -periodischen quadratintegrierbaren Funktionen. (In Wirklichkeit handelt es sich gar nicht um Funktionen, sondern um Äquivalenzklassen von Funktionen, solche 'Funktionen' g haben i. Allg. keinen wohlbestimmten Wert in einem Punkt.) Dieselbe Art von 'Funktionen' ist auch geeignet, um die Elemente f der Vervollständigung unseres Prä-Hilbertraums zu beschreiben; eine alternative Beschreibung ergibt sich durch die quadratsummablen Folgen $(a_n)_{n \in \mathbb{Z}}$. Die auf die Vervollständigung fortgesetzte Linearform kann als Integral oder als unendliche Summe geschrieben werden:

$$\ell(f) = \sum_{n=-\infty}^{\infty} \bar{c}_n a_n = \frac{1}{2\pi} \int \bar{g}(t) f(t) dt = \langle g | f \rangle,$$

(wobei die Art der Konvergenz zu thematisieren ist.)

Bemerkungen zum endlichdimensionalen Fall

- Sei H eine positiv definite $J \times J$ -Matrix und V der Hilbertraum dazu.

$$\langle \mathbf{y} | \mathbf{x} \rangle = \mathbf{y}^* \cdot H \cdot \mathbf{x} \quad \text{für alle } J\text{-Spalten } \mathbf{y}, \mathbf{x}.$$

Zu jeder J-Spalte $\mathbf{z} \in V$ gehört eine Linearform, nämlich die J-Zeile $\xi = \mathbf{z}^* \cdot H$. Die Menge aller Linearformen ist der Hilbertraum aller J-Zeilen ξ mit der Norm zur Matrix H^{-1} : $\|\xi\|^2 = \xi \cdot H^{-1} \cdot \xi^*$.

- Es sei $(V, \langle \cdot | \cdot \rangle)$ ein endlichdimensionaler Hilbertraum und $\{v_j : j \in J\}$ eine ONBasis. Die duale Basis ist dann eine ONBasis für den Hilbertraum der Linearformen. Ein solches orthonormiertes System von Linearformen nennt man ein **orthonormiertes Koordinatensystem**.
- Die Abbildung $v \mapsto \ell(\cdot) = \langle v | \cdot \rangle$ ist eine antilineare Isometrie des Hilbertraums V in den Raum der Linearformen (Mit dem Adjektiv ‘antilinear’ machen wir darauf aufmerksam, dass zwar Summen in Summen abgebildet werden, aber nicht Linearkombinationen in Linearkombinationen, wie das bei einer linearen Abbildung gefordert ist; wenn v in ℓ abgebildet wird, dann wird αv in $\bar{\alpha} \ell$ abgebildet.)

Didaktische Anmerkungen

- 1) In der Schulgeometrie geht man davon aus, dass der Begriff des rechten Winkels (Orthogonalität) zusätzlich zur Längenmessung intuitiv gegeben ist. Die Vektoren v und w sind zueinander orthogonal, wenn die Vektoren $v + w$ und $v - w$ dieselbe Länge haben,
Wenn man den Satz von Pythagoras zum sog. Cosinussatz verallgemeinern will, dann stößt man auf den Begriff des inneren Produkts. Um dieses einzuführen, stützt man sich üblicherweise auf kartesische Koordinaten. Es gilt als intuitiv klar, dass es kartesische Koordinaten gibt.
- 2) Wir haben hier gesehen, dass es auch anders geht. Man braucht nur die Parallelogrammgleichung (als Axiom) und die Idee der Polarisierung.
Im reellen Fall reduziert sich die Polarisierungskonstruktion auf

$$\langle v | w \rangle = \frac{1}{4} \|v + w\|^2 - \frac{1}{4} \|v - w\|^2.$$

Man schreibt $\langle v | w \rangle = \|v\| \|w\| \cos \phi$ und erhält in Verallgemeinerung des Satzes von Pythagoras den sog. Cosinussatz

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2 \|v\| \|w\| \cos \phi.$$

- 3) Das didaktische Problem mit dem Skalarprodukt besteht darin, dass es die Unterscheidung zwischen Vektoren und Linearformen überflüssig zu machen scheint; es gibt einen natürlichen Antiisomorphismus, (welcher im reellen Fall ein Vektorraum-Isomorphismus ist). In orthonormierten Koordinaten stellt sich diese Bijektion als der Übergang von der Spalte y zur Zeile y^* dar. Der Anfänger neigt erfahrungsgemäss dazu, Spalten ohne langes Nachdenken auch einmal als Zeilen zu schreiben. (In elementaren Lehrbüchern wird man dazu manchmal sogar ermuntert: Zeilen seien doch nichts anderes als ‘umgefallene’ Spalten, im einen wie im anderen Fall gehe es einfach um n -Tupel von Zahlen.)

- 4) P.A. Dirac und viele Physiker nach ihm benützen die Bezeichnung $|\mathbf{v}\rangle$ für die Zustände eines quantenmechanischen Systems und sie nennen $|\mathbf{v}\rangle$ einen ket-Vektor. Die Linearformen $\langle \mathbf{w} | \cdot \rangle$ nennen sie die bra-Vektoren. Hier wird also ganz klar unterschieden zwischen den Vektoren und den Linearformen. Die Zahl $\langle \mathbf{w} | \mathbf{v} \rangle$, das innere Produkt von \mathbf{w} und \mathbf{v} , ist die bracket („Klammer“). Für einen Einheitsvektor \mathbf{v} ist der Ausdruck $|\mathbf{v}\rangle\langle \mathbf{v} | \cdot \rangle$ als die orthogonale Projektion auf den von \mathbf{v} aufgespannten Vektorraum zu deuten. Über Dirac's Notation, die auch aus mathematischer Sicht große Vorzüge hat, werden wir später noch mehr sagen.

Hinweise

Wenn man den gegenwärtigen Unterabschnitt IV.1 mit dem Unterabschnitt II.5 vergleicht, könnte man vielleicht der Vorstellung verfallen, dass die ‘stetigen’ hermiteschen Formen für die komplexen Vektorräume einfach nur das sind, was die quadratischen Formen für die reellen Vektorräume sind. Gegenüber diesem flüchtigen Eindruck ist jedoch zu betonen, dass die Theorie der hermiteschen Formen ihre Tiefe und Bedeutung vor allem dadurch gewinnt, dass man die hermiteschen Formen auf Hilberträumen betrachtet, also auf Räumen mit einer sehr speziellen metrischen Struktur.

Wir werden in der nächsten Vorlesung sehen, dass man in der Hilbertraumtheorie die hermiteschen Sesquilinearformen als ‘symmetrische’ (oder ‘hermitische’ oder ‘selbstadjungierte’) Operatoren deuten kann. In unendlichdimensionalen Hilberträumen muss man genauer hinsehen, um gewisse hermitesche Sesquilinearen als ‘selbstadjungierte’ oder ‘wesentlich selbstadjungierte’ Operatoren zu verstehen. Und es sind diese selbstadjungierten Operatoren, die man in der Quantenphysik sehr wesentlich benötigt; sie repräsentieren nämlich die ‘Observablen’. Der Spektralsatz ist ein Satz von fundamentaler Bedeutung.

Es ist leider so, dass man für die unendlichdimensionale Theorie tiefliegende Techniken aus der Analysis benötigt. Viele physikalisch bedeutsame Ideen kommen aber auch schon in der endlichdimensionalen Theorie zur Geltung, wie Feynman in seinen berühmten elementaren Vorlesungen zeigt; die mathematische Behandlung von Quantensystemen mit endlich vielen Zuständen braucht weniger Analysis als Geometrie und Algebra, und die erforderlichen Techniken können bereits im ersten Studienjahr erarbeitet werden. Dazu wollen wir hier einen Beitrag leisten.

IV.2 Hauptachsen, Orthogonale Zerlegungen, Spektralsatz.

Ein endlichdimensionaler komplexer Vektorraum bekommt die Struktur eines Hilbertraums, wenn man eine Norm auszeichnet, welche die Parallelogrammgleichung erfüllt. Man kann auch sagen: Ein endlichdimensionaler Hilbertraum ist ein Vektorraum mit einem ausgezeichneten inneren Produkt: $(V, \langle \cdot | \cdot \rangle)$. Die Dimension wird im Folgenden n sein, wenn nichts anderes gesagt ist.

Der Satz von den Hauptachsen einer hermiteschen Form

Sei $(V, \langle \cdot | \cdot \rangle)$ ein n -dimensionaler Hilbertraum und $\Psi(\cdot)$ eine stetige hermitesche Form. Dann existieren reelle Zahlen $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ und ein orthonormales Koordinatensystem $\{\ell^1, \dots, \ell^n\}$, sodass gilt

$$\Psi(v) = \sum \lambda_j \cdot |\ell^j(v)|^2.$$

Mit anderen Worten: Es existiert eine ONBasis $\{e_j\}$, sodass gilt

$$\Psi(\cdot) = \lambda_1 \cdot |\langle e_1 | \cdot \rangle|^2 + \lambda_2 \cdot |\langle e_2 | \cdot \rangle|^2 + \dots + \lambda_n \cdot |\langle e_n | \cdot \rangle|^2.$$

Die von den Einheitsvektoren e_j aufgespannten eindimensionalen Vektorräume heißen Hauptachsen der hermiteschen Form; die λ_j heißen charakteristische Werte,— oder manchmal auch Eigenwerte (aus einem Grund, der verständlich wird, wenn wir uns unten mit den selbstadjungierten Operatoren befassen).

Beweis (mit Courant's Maximierungsargument)

Wenn wir zur reellwertigen Funktion Ψ ein reelles Vielfaches des Normquadrats dazuaddieren, dann erhalten wir wieder eine stetige hermitesche Form. Es genügt daher, den Satz für positive Ψ zu beweisen. $\Phi(\cdot, \cdot)$ sei die dazugehörige hermitesche Sesquilinearform.

$$\Psi(v+w) = \Phi(v+w, v+w) = \Psi(v) + \Psi(w) + 2 \cdot \Re \Phi(v, w).$$

Unser positives Ψ nimmt irgendwo auf der Einheitskugel unseres endlichdimensionalen Hilbertraums das Maximum $\lambda_1 > 0$ an, sagen wir in e_1 . Sei v orthogonal zu e_1 , $\langle e_1 | v \rangle = 0$. Wir zeigen $\Phi(e_1, \beta v) = 0$ für alle β

$$\Psi(e_1 + \beta v) = \Psi(e_1) + |\beta|^2 \Psi(v) + 2 \cdot \Re(\beta a). \quad \text{mit } a = \Phi(e_1, v).$$

Für den auf Länge 1 normierten Vektor gilt

$$\Psi(e_1) \geq \Psi\left(\frac{e_1 + \beta v}{\|e_1 + \beta v\|}\right) = \frac{1}{1 + |\beta|^2} \Psi(e_1 + \beta v).$$

So ergibt sich $2 \cdot \Re(\beta a) \leq (1 + |\beta|^2) \Psi(e_1) - \Psi(e_1) - |\beta|^2 \Psi(v) = |\beta|^2 (\Psi(e_1) - \Psi(v))$. Die rechte Seite verschwindet von zweiter Ordnung für $\beta \rightarrow 0$; also gilt $a = \Phi(e_1, v) = 0$.

Mit V_1^\perp bezeichnen wir den $(n-1)$ -dimensionalen Vektorraum aller auf e_1 orthogonalen Vektoren: $V_1^\perp = \{w : \langle e_1 | w \rangle = 0\}$. Die Einschränkung von Ψ auf V_1^\perp erfüllt die Bedingungen des Satzes. Der Beweis des Satzes ergibt sich durch vollständige Induktion nach

der Dimension. Wenn $\{e_2, \dots, e_n\}$ eine ONBasis von V_1^\perp ist, welche das Verlangte für die Einschränkung von Ψ leistet, dann ist $\{e_1, e_2, \dots, e_n\}$ eine ONBasis von V mit

$$\Psi\left(\sum \alpha^j e_j\right) = \sum \lambda_j |\alpha^j|^2.$$

Konsequenzen für die hermitesche Sesquilinearformen: Die hermiteschen Sesquilinearform zur hermiteschen Form $|\ell(v)|^2$ ist offenbar $\Phi(w, v) = \overline{\ell(w)} \cdot \ell(v)$.

Der Satz liefert für die zu Ψ gehörende Sesquilinearform

$$\begin{aligned} \Phi(w, \cdot) &= \sum \lambda_j \overline{\ell^j(w)} \cdot \ell^j(\cdot) = \sum \lambda_j \overline{\langle e_j | w \rangle} \cdot \langle e_j | \cdot \rangle = \sum \lambda_j \langle w | e_j \rangle \cdot \langle e_j | \cdot \rangle, \\ \text{oder kurz} \quad \Phi(\cdot, \cdot) &= \sum \lambda_j \langle \cdot | e_j \rangle \cdot \langle e_j | \cdot \rangle. \end{aligned}$$

Wir werden sehen, dass man die rechte Seite gemäss der Dirac'schen Notation auch als eine selbstadjungierte Abbildung $\varphi : V \rightarrow V$ interpretieren kann. Zunächst ziehen wir Konsequenzen für den Matrizenkalkül.

Satz (Simultane Diagonalisierung)

Sei T eine positiv definite $I \times I$ -Matrix, ($|I| = n$), und sei H eine beliebige hermitesche $I \times I$ -Matrix. Dann existiert eine invertierbare Matrix A , sodass

$$A^* \cdot T \cdot A = E \quad \text{und} \quad A^* \cdot H \cdot A = \Lambda \quad (\text{Diagonalmatrix}).$$

(Die Einheitsmatrix wird hier vorübergehend mit $E = E_{I \times I}$ bezeichnet, um den Konflikt mit dem Namen der Indexmenge I zu vermeiden.)

Beweis

Mit der Matrix T machen wir den Vektorraum \mathbb{C}_{Sp}^I zu einem Hilbertraum

$$\|y\|_T = (y^* \cdot T \cdot y)^{1/2}.$$

Ein n -Tupel von I -Spalten $\{w_k : k \in K\}$ ist genau dann eine Orthonormalbasis, wenn

$$w_k^* \cdot T \cdot w_\ell = \begin{cases} 1 & \text{für } k = \ell \\ 0 & \text{für } k \neq \ell \end{cases}$$

also genau dann, wenn die Matrix A mit den Spalten w_k die Gleichung $A^* T A = E$ erfüllt. Wir haben jetzt noch die Freiheit, welche ONBasis wir wählen wollen. Wir wählen die Orthonormalbasis so, dass die hermitesche Form $\Psi(y) = y^* H y$ diagonalisiert ist, d.h. $\{w_k : k \in K\}$ mit

$$\begin{aligned} w_k^* H w_k &= \lambda_k \quad \text{für } k \in K \\ w_k^* H w_\ell &= 0 \quad \text{für } k \neq \ell. \end{aligned}$$

Für die Matrix A mit den Spalten w_k bedeutet das $A^* H A = \Lambda$, wo Λ die Diagonalmatrix mit den Diagonalelementen λ_k ist. Dies ist die Behauptung.

Als Spezialfall dieses Satzes erhalten wir den vielbenützten

Satz (Unitäre Diagonalisierung einer hermiteschen Matrix)

Zu jeder hermiteschen Matrix H gibt es eine unitäre Matrix U , sodass $U^ H U$ eine Diagonalmatrix ist.*

Mit anderen Worten: Es gibt eine Orthonormalbasis des Spaltenraums, die aus ‘Eigenvektoren’ besteht.

Hinweis Die Hauptachsendarstellung in der Form

$$\Psi(v) = \sum \lambda_j \cdot |\ell^j(v)|^2. \quad \{\ell^j : j = 1, \dots, n\} \text{ orthonormal}$$

hat insofern einen Schönheitsfehler, dass die ONBasis nur dann eindeutig bestimmt ist, wenn die λ_j paarweise verschieden sind. Die natürliche Betrachtung, die dann auch zu den Verallgemeinerungen gut passt, ergibt sich, wenn man die e_j , die zum gleichen ‘Eigenwert’ λ gehören, zusammenfasst. Sei V_λ der von diesen e_j aufgespannte Vektorraum, und sei π_λ die orthogonale Projektion auf V_λ (was eine orthogonale Projektion ist, werden wir unten erläutern), dann liefert der Satz:

Es gibt $\lambda_1 > \dots > \lambda_m$ und paarweise disjunkte orthogonale Projektionen π_1, \dots, π_m , sodass $\Psi(v) = \sum \lambda_k \cdot \|\pi_k(v)\|^2$. Die λ_k und die π_k sind durch Ψ eindeutig bestimmt.

Projektionen eines \mathbb{K} -Vektorraums

Eine lineare Abbildung eines Vektorraums V in sich heißt ein **Endomorphismus** von V . Ein Endomorphismus φ mit $\varphi \circ \varphi = \varphi$ heißt eine **Projektion**.

Wenn V_1 das Bild bezgl. einer Projektion π ist, und V_0 der Nullraum, $V_1 = \pi(V)$, $V_0 = \{v : \pi(v) = 0\}$, dann sagt man, π sei **Projektion auf V_1 entlang V_0** .

Zwei Projektionen π_1, π_2 heißen **disjunkte Projektionen**, wenn $\pi_1 \circ \pi_2 = 0 = \pi_2 \circ \pi_1$.

Bemerkungen:

1. Wenn π die Projektion auf V_1 entlang V_0 ist, dann ist $\text{id} - \pi$ die Projektion auf V_0 entlang V_1 . π und $\text{id} - \pi$ heißen zueinander komplementäre Projektionen.
2. Ein Teilvektorraum V_0 heisst ein **Komplementärraum** zu V_1 (in V), wenn gilt $V_0 \cap V_1 = \{0\}$ und $\dim V_0 + \dim V_1 = \dim V$. Es sind genau diese Paare, zu welchen es eine Projektion gibt.
3. Wenn π_1, π_2 disjunkte Projektionen sind, dann sind $\pi_1 + \pi_2$ und $\pi_3 = \text{id} - (\pi_1 + \pi_2)$ Projektionen und π_1, π_2, π_3 sind paarweise disjunkt.
4. Es seien $\pi_1, \pi_2, \dots, \pi_m$ paarweise disjunkte Projektionen mit $\pi_1 + \pi_2 + \dots + \pi_m = \text{id}$; und es bezeichne V_k den Wertebereich von π_k . Jedes $v \in V$ besitzt dann genau eine Darstellung $v = v_1 + v_2 + \dots + v_m$ mit $v_k \in V_k$.

5. Umgekehrt: Seien V_1, V_2, \dots, V_m Teilvektorräume von V , sodass zu jedem v genau eine Darstellung existiert $v = v_1 + v_2 + \dots + v_m$ mit $v_k \in V_k$. Dann sind die Abbildungen $\pi_k : v \mapsto v_k$ paarweise disjunkte Projektionen, die sich zur Identität aufsummieren.

Notation Man schreibt in diesem Fall

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_m = \bigoplus_{k=1}^m V_k.$$

und sagt, V sei die direkte Summe der V_k . Man spricht auch von einer **direkten Zerlegung** des Vektorraums V .

Bemerkung: Jede Basis von V liefert eine direkte Zerlegung in eindimensionale Teilvektorräume. Sei $V = V_0 \oplus V_1$ und v_1, \dots, v_l eine Basis von V_0 , v_{l+1}, \dots, v_m eine Basis von V_1 ; dann ist v_1, \dots, v_m eine Basis von V .

Verträglichkeit: Man sagt von einer quadratischen Form $Q(v)$, sie sei verträglich mit der direkten Zerlegung $V = \bigoplus_{k=1}^m V_k$, wenn es quadratische Formen Q_k auf V_k gibt, sodass gilt

$$Q(v) = Q(\pi_1(v) + \dots + \pi_m(v)) = Q_1(\pi_1(v)) + \dots + Q_m(\pi_m(v)).$$

Es ist klar, was es heisst, dass eine stetige hermitesche Form auf einem komplexen Vektorraum verträglich ist mit einer vorgegebenen direkten Zerlegung. Wir werden unten auch erklären, was es bedeutet, dass eine Gruppendarstellung über V verträglich ist mit einer gewissen direkten Zerlegung. Bei den hermiteschen Formen und den Gruppendarstellungen interessiert uns allerdings vor allem die Verträglichkeit mit sog. orthogonalen Zerlegungen eines Hilbertraums.

Definition

1. Zu jeder Menge M in einem Hilbertraum V definiert man den **Orthogonalraum** zu M .

$$M^\perp = \{w : \langle w | v' \rangle = 0\} \quad \text{für alle } v' \in M \subseteq V.$$

2. Man sagt von zwei Teilvektorräumen V_1, V_2 eines Hilbertraums, dass sie **zueinander orthogonal** sind, wenn gilt

$$\langle v_1 | v_2 \rangle = 0 \quad \text{für alle } v_1 \in V_1, v_2 \in V_2.$$

Man notiert $V_1 \perp V_2$ oder $V_1 \subseteq V_2^\perp$ oder $V_2 \subseteq V_1^\perp$.

3. Man nennt eine Projektion auf V_1 die **orthogonale Projektion** auf V_1 , wenn ihr Nullraum der Orthogonalraum V_1^\perp ist.
4. Eine direkte Zerlegung eines Hilbertraums $(V, \langle \cdot | \cdot \rangle)$ heisst eine **orthogonale Zerlegung**, wenn die V_k paarweise orthogonal sind. Man notiert dann

$$V = V_1 \boxplus V_2 \boxplus \dots \boxplus V_m,$$

und man sagt, V sei die **orthogonale Summe** der V_k .

Hinweis Die orthogonalen Projektionen spielen auch in der Theorie der unendlichdimensionalen Hilberträume eine wichtige Rolle. Man beweist da, dass es zu jedem ‘abgeschlossenen’ Teilraum V_1 die orthogonale Projektion auf V_1 gibt; jedes $v \in V$ besitzt eine eindeutige Zerlegung $v = v_1 + v_0$ mit $v_1 \in V_1$ und $v_0 \in V_1^\perp$; v_0 heisst das Lot auf V_1 , und v_1 heisst der Fußpunkt des Lots. v_1 ist derjenige Punkt in V_1 , der von v minimalen Abstand hat.

Man beachte: Um zu garantieren, dass es einen Punkt mit minimalen Abstand gibt, braucht man die Vollständigkeit des Raums V_1 . Im endlichdimensionalen Fall gibt es keine Probleme: jeder endlichdimensionale Teilvektorraum ist vollständig.

Den Fußpunkt v_1 des Lots von v auf einen endlichdimensionalen Teilraum V_1 kann man auch ohne Minimierungskonstruktion direkt angeben: Sei $\{w_k : k \in K\}$ eine ONBasis des Teilvektorraums. Es gilt dann

$$v_1 = \sum_k \alpha_k w_k \quad \text{mit} \quad \alpha_k = \langle v | w_k \rangle .$$

Der Vektor $v_0 = v - v_1$ steht nämlich senkrecht auf allen w_k und damit auf V_1 .

Selbstadjungierte Operatoren

Ein Endomorphismus eines endlichdimensionalen Hilbertraums $\varphi : V \rightarrow V$ heisst ein selbstadjungierter Operator, wenn gilt

$$\langle w | \varphi(v) \rangle = \langle \varphi(w) | v \rangle \quad \text{für alle} \quad v, w \in V .$$

Hilfssatz

Eine Projektion eines Hilbertraums ist genau dann eine orthogonale Projektion, wenn sie selbstadjungiert ist.

Beweis

Sei V_1 der Wertebereich der Projektion π und V_0 der Wertebereich von $\text{id} - \pi$. Es gilt

$$V_0 \perp V_1 \iff \forall v, w \quad \langle \pi w | (\text{id} - \pi)v \rangle = 0 \iff \forall v, w \quad \langle \pi w | v \rangle = \langle \pi w | \pi v \rangle$$

und wegen der Symmetrie der rechten Seite

$$\begin{aligned} &\implies \forall v, w \quad \langle \pi w | v \rangle = \langle w | \pi v \rangle \implies \pi \text{ selbstadjungiert.} \\ \pi \text{ selbstadjungiert} &\iff \forall v, w \quad \langle \pi w | v \rangle = \langle w | \pi v \rangle = \langle w | \pi(\pi v) \rangle = \langle \pi w | \pi v \rangle \\ &\implies \forall v, w \quad \langle \pi w | (\text{id} - \pi)v \rangle = 0 \implies V_0 \perp V_1 . \end{aligned}$$

Satz (Natürliche Bijektion)

In einem endlichdimensionalen Hilbertraum gibt es eine natürliche Bijektion zwischen den hermiteschen Sesquilinearformen und den selbstadjungierten Operatoren.

Man beachte: Es handelt sich um einen Isomorphismus zwischen reellen Vektorräumen der Dimension n^2 . Eine hermitesche Sesquilinearform erscheint bezüglich einer Basis als

eine hermitesche $n \times n$ -Matrix; in der Diagonalen stehen irgendwelche reellen Zahlen, oberhalb der Diagonalen kann man die Real- und Imaginärteile der Einträge beliebig wählen. $n + 2\frac{n(n-1)}{2} = n^2$.

Konstruktion Wenn $\Phi(\cdot, \cdot)$ eine hermitesche Sesquilinearform auf dem n -dimensionalen Hilbertraum V ist, dann haben wir zwei Abbildungen S und T , die V auf den Raum V^* der Linearformen abbilden:

$$S : w \mapsto S(w) = \Phi(w, \cdot) ; \quad T : w \mapsto T(w) = \langle w | \cdot \rangle .$$

Beide sind antilinear; T ist eine Bijektion. Die zusammengesetzte Abbildung

$$\varphi : v \mapsto \varphi(v) = T^{-1} \circ S(v)$$

ist ein Endomorphismus von V , und zwar ein selbstadjungierter Endomorphismus.

Beweis: Für einen Vektor u gilt $u = \varphi(v) = T^{-1} \circ S(v)$ genau dann, wenn für alle w gilt $\langle u | w \rangle = \Phi(v, w)$. Wir haben also

$$\begin{aligned} \Phi(v, w) &= \langle \varphi(v) | w \rangle , & \Phi(w, v) &= \langle \varphi(w) | v \rangle ; \\ \langle \varphi(w) | v \rangle &= \overline{\Phi(v, w)} = \langle w | \varphi(v) \rangle . \end{aligned}$$

Darstellung in orthonormierten Koordinaten

Sei $\{e_j : j \in J\}$ eine ONBasis, und sei H die Gram-Matrix für $\Phi(\cdot, \cdot)$. Die Vektoren werden durch J -Spalten dargestellt: für $w \leftrightarrow y$, $v \leftrightarrow x$ gilt

$$\langle w | v \rangle = y^* \cdot x , \quad \Phi(w, v) = y^* \cdot H \cdot x = y^* \cdot (Hx) .$$

Sei $\varphi(\cdot)$ die lineare Abbildung zu $x \mapsto Hx$. Es gilt dann

$$\langle w | \varphi(v) \rangle = y^* \cdot (Hx) = \Phi(w, v) .$$

Jede hermitesche Matrix kann einerseits als ein selbstadjungierter Operator φ auf dem Raum der Spalten (mit der ‘natürlichen’ Hilbertraumnorm) verstanden werden und andererseits als die Gram-Matrix einer hermiteschen Form Ψ . Die resultierende Bijektion der beiden n^2 -dimensionalen reellen Vektorräume hängt nicht von der Wahl der ONBasis ab; die darstellende Matrix H wird in $U^* H U$ transformiert, wenn U die unitäre Matrix der Wechsels von einer ONBasis zu einer anderen ONBasis ist.

Beispiel: Sei $\pi(\cdot)$ eine orthogonale Projektion. Da sie ein selbstadjungierter Operator ist, entspricht ihr eine hermitesche Sesquilinearform oder auch eine stetige hermitesche Form $\Psi(\cdot)$, und diese ist $\Psi(v) = \|\pi(v)\|^2$. In der Tat gilt

$$\Psi(v) = \Phi(v, v) = \langle v | \pi(v) \rangle = \langle v | \pi(\pi v) \rangle = \|\pi(v)\|^2 .$$

Der Satz von der Hauptachsenzerlegung kann nun auch als ein Satz über selbstadjungierte Operatoren formuliert werden

Satz (Spektralzerlegung selbstadjungierter Operatoren)

Zu jeder stetigen hermiteschen Form auf einem endlichdimensionalen Hilbertraum gibt es Zahlen $\lambda_1 > \dots > \lambda_m$ und eine orthogonale Zerlegung $V = V_1 \boxplus V_2 \boxplus \dots \boxplus V_m$, sodass Ψ auf V_k gleich λ_k mal das Normquadrat ist.

$$\Psi(v_1 + \dots + v_m) = \lambda_1 \|v_1\|^2 + \dots + \lambda_m \|v_m\|^2 .$$

Der zu Ψ gehörende selbstadjungierte Operator hat die Gestalt

$$\phi(\cdot) = \lambda_1 \pi_1 + \dots + \lambda_m \pi_m .$$

Normale Operatoren

Zu jedem Endomorphismus ϵ eines endlichdimensionalen Hilbertraums V gibt es einen eindeutig bestimmten adjungierten Operator ϵ^*

$$\langle w | \epsilon(v) \rangle = \langle \epsilon^*(w) | v \rangle .$$

Wenn ϵ in einer ONBasis durch die Matrix M dargestellt wird, dann wird ϵ^* in dieser ONBasis durch die hermitisch konjugierte Matrix M^* dargestellt.

$$y^* \cdot Mx = (M^*y)^* \cdot x .$$

Für hintereinandergeschaltete Endomorphismen gilt $(\epsilon \circ \varphi)^* = \varphi^* \circ \epsilon^*$; denn

$$\langle (\epsilon \circ \varphi)^* w | v \rangle = \langle w | (\epsilon \circ \varphi)(v) \rangle = \langle \epsilon^* w | \varphi v \rangle = \langle \varphi^*(\epsilon^* w) | v \rangle .$$

Jeder Endomorphismus besitzt eine eindeutige Darstellung

$$\epsilon = \sigma + i\tau, \quad \text{mit } \sigma, \tau \text{ selbstadjungiert; } \quad \sigma = \frac{1}{2}(\epsilon + \epsilon^*) \quad \tau = \frac{1}{2i}(\epsilon - \epsilon^*) .$$

Der Endomorphismus heisst ein **normaler Endomorphismus**, wenn gilt $\epsilon\epsilon^* = \epsilon^*\epsilon$. Dies ist genau dann der Fall, wenn der ‘Realteil’ σ und der ‘Imaginärteil’ τ kommutieren.

$$\epsilon\epsilon^* - \epsilon^*\epsilon = (\sigma + i\tau) \cdot (\sigma - i\tau) - (\sigma - i\tau) \cdot (\sigma + i\tau) = 2i(\tau\sigma - \sigma\tau) .$$

Satz

Wenn ϵ ein normaler Endomorphismus ist, dann gibt es paarweise verschiedene komplexe Zahlen $\alpha_l + i\beta_l$ und paarweise disjunkte orthogonale Projektionen π_l , sodass gilt

$$\epsilon = \sum (\alpha_l + i\beta_l) \pi_l .$$

Die Darstellung ist eindeutig bestimmt.

Beweisskizze

Wir zerlegen zuerst den selbstadjungierten Operator in der bekannten Weise

$$\epsilon\epsilon^* = \epsilon^*\epsilon = \sum \lambda_i \pi_i \quad \lambda_i \text{ paarweise verschieden.}$$

Die Eigenräume $V_i = \{v : \epsilon\epsilon^*(v) = \lambda_i v\}$ werden sowohl durch ϵ als auch durch ϵ^* in sich abgebildet.

$$v \in V_i \implies \epsilon\epsilon^*(\epsilon(v)) = \epsilon(\epsilon^*\epsilon(v)) = \epsilon(\lambda_i v) = \lambda_i \epsilon(v).$$

Ebenso für ϵ^* . Die Einschränkung von $\lambda_i^{-1/2}\epsilon$ auf V_i ist ein unitärer Operator.

Es genügt daher, den Satz für unitäre Operatoren zu beweisen:

Wenn U ein unitärer Operator ist, dann existieren paarweise verschiedene Zahlen vom Betrag 1 und paarweise disjunkte orthogonale Projektionen π_l , sodass $U = \sum e^{i\phi_l} \pi_l$.

Real- und Imaginärteil des unitären Operators U sind kommutierende selbstadjungierte Operatoren, Wir zerlegen in der bekannten Weise

$$\sigma = \sum \alpha_j \rho_j, \quad \tau = \sum \beta_k \rho_k.$$

Wegen $UU^* = I$ haben wir $\sigma^2 + \tau^2 = I$ und $\sigma\tau - \tau\sigma = 0$.

$$I = \sum \alpha_j^2 \rho_j + \sum \beta_k^2 \rho_k = \sum (\alpha_j^2 \rho_j \rho_k + \beta_k^2 \rho_k \rho_j) \quad (H)$$

Wenn wir diese Gleichung von rechts oder links mit ρ_j oder ρ_k multiplizieren, ergibt sich

$$\rho_j = \alpha_j^2 \rho_j + \sum \beta_k^2 \rho_j \rho_k = \alpha_j^2 \rho_j + \sum \beta_k^2 \rho_k \rho_j \quad (Re)$$

$$\rho_k = \sum \alpha_j^2 \rho_k \rho_j + \beta_k^2 \rho_k = \sum \alpha_j^2 \rho_j \rho_k + \beta_k^2 \rho_k \quad (Im)$$

$$\rho_j \rho_k = \alpha_j^2 \rho_j \rho_k + \beta_k^2 \rho_j \rho_k; \quad \rho_k \rho_j = \alpha_j^2 \rho_k \rho_j + \beta_k^2 \rho_k \rho_j.$$

Wir können annehmen, dass σ und τ keine Eigenwerte vom Betrag 1 haben. (Andernfalls studieren wir den unitären Operator $U \cdot e^{i\phi}$). Die letzte Gleichung liefert $\rho_j \rho_k \neq 0 \implies \alpha_j^2 + \beta_k^2 = 1$. Die Summen in den Gleichungen (Re) und (Im) haben nur zwei Summanden für jedes j bzw. für jedes k . Wenn $\alpha_j^2 + \beta_k^2 = 1$, dann passen zu $\alpha = \alpha_j$ die Werte β_k und $-\beta_k$. Zu $\beta = \beta_k$ passen die Werte $\pm\alpha$. Die dazugehörigen Projektionen seien ρ_+, ρ_- (für das gegebene α) bzw. ρ_+, ρ_- für das gegebene β . Die Gleichungen (Re) und (Im) liefern eine Projektion π mit

$$\rho_+ + \rho_- = \pi = \rho_+ + \rho_-.$$

Der einfachste Fall ist der, wo für alle α, β nur eine einzige Vorzeichenkombination (nämlich eben die zu α, β) eine nichtverschwindende Projektion $\rho\rho = \pi_l$ liefert. In diesem Fall können wir die verträglichen Paare mit einem Index $l \in L$ indizieren und die dazugehörigen Koeffizienten mit α_l, β_l bezeichnen.

$$\begin{aligned} \sigma &= \sum \alpha_j \rho_j = \sum \alpha_l \pi_l, & \tau &= \sum \beta_k \rho_k = \sum \beta_l \pi_l; \\ U &= \sigma + i\tau = \sum (\alpha_l + i\beta_l) \pi_l. \end{aligned}$$

Intuitiv wird man vermuten, dass man den allgemeinen Fall durch eine ‘Drehung’ $U \mapsto U \cdot e^{i\Phi}$ auf den einfachen Fall zurückführen kann. Eine alternative Beweisidee ist die, dass man aus $\sigma\tau - \tau\sigma = 0$ schliesst, dass die Projektionen ρ_j und ρ_k kommutieren; denn dann liefern die ‘Produkte’ $\rho_j\rho_k$ tatsächlich Projektionen (die trivialerweise orthogonal und paarweise disjunkt sind). Man kann das Problem leicht zurückführen auf das einfache

Hilfssatz *Seien ρ_+ und ρ_- Projektionen und ρ_+, ρ_- die komplementären Projektionen. Wenn die Operatoren $\rho_+ - \rho_-$ und $\rho_+ + \rho_-$ kommutieren, dann kommutieren die ρ mit den ρ und die ‘Produkte’ sind Projektionen.*

Übersetzung in den Matrizenkalkül Wenn man im Hilbertraum eine ONBasis einführt, dann sind die normalen Operatoren durch normale Matrizen N dargestellt, $N^*N = NN^*$. $H = NN^*$ ist eine positiv semidefinite Matrix mit gewissen Eigenwerten $\lambda_i \geq 0$. Der Eigenraum zum Eigenwert 0 muss uns nicht weiter interessieren. Die Einschränkung von N auf den Eigenraum zu $\lambda > 0$ ist bis auf den Faktor $\sqrt{\lambda}$ ein unitärer Operator. Auf jedem solchen Eigenraum für H können wir eine ONBasis so einführen, dass die Basisvektoren gleichzeitig Eigenwerte sind für den selbstadjungierten Realteil und den selbstadjungierten Imaginärteil des normalen Operators $N = S + iT$.— Das ist der entscheidende Schritt! Wenn wir die Spalten zu dieser ONBasis zu einer unitären Matrix V zusammenstellen, dann haben wir $NV = V \cdot D$ wo D die Diagonalmatrix ist, deren Diagonalelemente die zusammengehörigen Eigenwerte der Operatoren S und iT sind. Damit ist bewiesen

Satz (Unitäre Diagonalisierung)

Zu jeder normalen Matrix N , ($NN^ = N^*N$) existiert eine unitäre Matrix V , sodass V^*NV eine Diagonalmatrix ist. Eine Matrix, die nicht normal ist kann nicht unitär diagonalisiert werden.*

IV.3 Einparametrische Matrixgruppen

Wir haben schon früher bemerkt, dass man auf quadratische Matrizen die Exponentialabbildung anwenden kann. Für kommutierende Matrizen macht die Exponentialabbildung aus der Summe ein Produkt. Die einparametrischen Matrixgruppen sind ein einfacher aber wichtiger Anwendungsfall.

Definition Für jedes $t \in \mathbb{R}$ sei $X(t)$ eine komplexe $n \times n$ -Matrix, sodass gilt

$$\lim_{t \rightarrow 0} X(t) = I \quad \text{und} \quad X(s+t) = X(s) \cdot X(t) \quad \text{für alle } s, t.$$

Die Schar $\{X(t) : t \in \mathbb{R}\}$ heisst dann eine **einparametrische Gruppe** von $n \times n$ -Matrizen.

Satz

Für jede einparametrische Gruppe $\{X(t) : t \in \mathbb{R}\}$ existiert der Grenzwert

$$A = \lim_{h \rightarrow 0} \frac{1}{h} (X(h) - I) \quad \text{und es gilt} \quad X(t) = \exp(t \cdot A) \quad \text{für alle } t.$$

Die Matrix A heisst der **infinitesimale Generator** oder der infinitesimale Operator der einparametrischen Matrixengruppe $\{X(t) : t \in \mathbb{R}\}$.

Bemerkung (Spurlose infinitesimale Operatoren): Für jede einparametrische Matrixgruppe ist die Determinante $d(t) = \det(X(t))$ eine multiplikative Funktion, $d(s+t) = d(s) \cdot d(t)$, also von der Form $d(t) = e^{ct}$; aus dem Verhalten für $t \rightarrow 0$ ersieht man $c = \text{trace}(A)$. Genau dann, wenn der infinitesimale Generator die Spur 0 hat, besteht die Gruppe aus Matrizen mit der Determinante 1.

Hinweis Die Charakterisierung der einparametrischen Gruppen von Matrizen durch ihre infinitesimalen Generatoren gehört zur Theorie der linearen Differentialgleichungen mit konstanten Koeffizienten. Wir betonen hier diejenigen Aspekte, die in der Theorie der einparametrischen Gruppen von Isometrien eines Hilbertraums ihre natürliche Fortsetzung finden; die infinitesimalen Operatoren sind dann selbstadjungierte Operatoren. Zunächst wollen wir aber doch noch allgemeiner bleiben.

Speziellere einparametrische Gruppen Eine einparametrische Gruppe von $n \times n$ -Matrizen nennt man auch eine einparametrische Untergruppe der Allgemeinen Linearen Gruppe $GL(n, \mathbb{C})$. Jede $n \times n$ -Matrix A kommt als infinitesimaler Generator in Betracht.

- Jede Matrix mit der Spur 0 ('spurlose Matrix') liefert eine einparametrische Untergruppe der Speziellen Linearen Gruppe $SL(n, \mathbb{C})$.
- Einparametrische Gruppen orthogonaler Matrizen sind dadurch gekennzeichnet, dass der infinitesimale Generator schiefssymmetrisch ist.

$$X(-t) = X^T(t) \quad \text{für alle } t \quad \Longleftrightarrow \quad A^T = -A.$$

- Eine einparametrische Gruppe von $n \times n$ -Matrizen besteht genau dann aus unitären Matrizen, wenn $U(t) = e^{-itH}$ mit hermischem H .
(Eine komplexe Matrix A mit $A^* = -A$, also $H = iA$ hermitisch, nennt man eine schiefhermitische Matrix.)
- Die spurlosen antihermitischen Matrizen sind die infinitesimalen Generatoren der einparametrischen Untergruppen der Speziellen Unitären Gruppe $SU(n)$.

$$U(-t) = U(t)^* \quad \text{für alle } t \text{ und } \det = 1 \quad \Longleftrightarrow \quad A^* = -A \quad \text{und} \quad \text{trace} = 0.$$

Beispiel: Gleichförmige Drehungen im Anschauungsraum

Die einparametrischen Untergruppen der orthogonalen Gruppe $O(3, \mathbb{R})$ liegen allesamt in der Speziellen Orthogonalen Gruppe $SO(3, \mathbb{R})$. Man interpretiert sie als gleichförmige Drehungen. Der infinitesimale Operator $A = \begin{pmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{pmatrix}$ gibt Auskunft über die Drehachse und die Drehgeschwindigkeit. In der traditionellen Vektoranalysis nennt man das Tripel (a, b, c) den zur gleichförmigen Drehbewegung gehörenden ‘axialen Vektor’.

Bekanntlich gibt die Überlagerung gleichförmiger Drehbewegungen wieder eine gleichförmige Drehbewegung. Die Überlagerung spiegelt sich in der Addition der infinitesimalen Generatoren, die Multiplikation mit einem Skalar ändert die Drehgeschwindigkeit.

Satz

Sei $\{X(t) : t \in \mathbb{R}\}$ eine einparametrische Gruppe mit dem infinitesimalen Generator A . Wenn B eine invertierbare Matrix ist, dann bilden die Matrizen $Y(t) = B^{-1} \cdot X(t) \cdot B$ eine einparametrische Gruppe mit dem infinitesimalen Generator $B^{-1} \cdot A \cdot B$.

Beweis

$$\begin{aligned} \exp(t \cdot B^{-1}AB) &= I + t \cdot B^{-1}AB + \frac{1}{2!}t^2(B^{-1}AB)^2 + \dots \\ &= B^{-1}(I + t \cdot A + \frac{1}{2!}t^2 \cdot A^2 + \dots)B = Y(t). \end{aligned}$$

Korollar

Wenn $B^{-1}AB$ die Diagonalmatrix mit den Diagonalelementen λ_j ist, dann ist $B^{-1}X(t)B$ die Diagonalmatrix mit den Diagonalelementen $e^{t\lambda_j}$.

Einparametrische Gruppen von komplexen 2×2 -Matrizen

Satz

Wenn der infinitesimale Operator A einer einparametrischen Gruppe von 2×2 -Matrizen diagonalisierbar ist, $B^{-1}AB = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, dann gilt mit dieser Matrix B und den Zahlen λ

$$X(t) = B \cdot \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \cdot B^{-1}.$$

Wenn der infinitesimale Operator A nicht diagonalisierbar ist, dann existiert eine Zahl $\tilde{\lambda}$ und eine invertierbare Matrix B , sodass gilt

$$B^{-1}AB = \begin{pmatrix} \tilde{\lambda} & 1 \\ 0 & \tilde{\lambda} \end{pmatrix} = \tilde{\lambda} \cdot I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad B^{-1}X(t)B = e^{\tilde{\lambda}t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Beweis

1. Die Matrix A ist genau dann diagonalisierbar, wenn es zwei linear unabhängige Eigenvektoren gibt. Seien \mathbf{x} und \mathbf{y} linear unabhängige Spalten mit

$$A \cdot \mathbf{x} = \lambda_1 \mathbf{x}, \quad A \cdot \mathbf{y} = \lambda_2 \mathbf{y}.$$

Wir fassen diese Spalten zur 2×2 -Matrix $B = (\mathbf{x}, \mathbf{y})$ zusammen und erhalten $A \cdot B = B \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Das Übrige zum diagonalisierbaren Fall erledigt das Korollar. (Eine hinreichende Bedingung für die Diagonalisierbarkeit ist übrigens die Bedingung, dass das charakteristische Polynom von A zwei verschiedene Nullstellen hat.)

2. Nehmen wir an, dass das charakteristische Polynom die doppelte Nullstelle $\tilde{\lambda}$ hat.

$$\det(A - \lambda I) = \lambda^2 - \lambda \cdot \text{trace}(A) + \det A = (\lambda - \tilde{\lambda})^2.$$

Nach dem Satz von Cayley-Hamilton (siehe I.1) gilt $(A - \tilde{\lambda}I)^2 = 0$.

Wenn $A - \tilde{\lambda}I = 0$, dann ist nichts weiter zu zeigen. Im anderen Fall gibt es eine Spalte \mathbf{y} mit $(A - \tilde{\lambda}I)\mathbf{y} = \mathbf{x} \neq 0$. Wir fassen die linear unabhängigen Spalten zur Matrix $B = (\mathbf{x}, \mathbf{y})$ zusammen. Es gilt dann

$$(A - \tilde{\lambda}I) \cdot B = (A - \tilde{\lambda}I)(\mathbf{x}, \mathbf{y}) = (0, \mathbf{x}).$$

Aus der Invertierbarkeit von B ergibt sich

$$I = B^{-1}(\mathbf{x}, \mathbf{y}); \quad B^{-1}\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad B^{-1}\mathbf{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{also} \\ B^{-1}AB = \tilde{\lambda} \cdot I + B^{-1}(0, \mathbf{x}) = \tilde{\lambda} \cdot I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \tilde{\lambda} & 1 \\ 0 & \tilde{\lambda} \end{pmatrix}.$$

3. Schliesslich haben wir

$$\exp(t \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = I + t \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Hinweis In der höheren Linearen Algebra zeigt man, dass man jede $n \times n$ -Matrix A auf Dreiecksgestalt bringen kann $B^{-1}AB = \Delta$ (Dreiecksmatrix); die Matrizen $B^{-1} \exp(tA)B$ haben in der Diagonalen die Einträge $e^{\lambda_j t}$.

Reelle infinitesimale Generatoren

Bei reellen einparametrischen Gruppen von 2×2 -Matrizen kann man mehrere Fälle unterscheiden. Wenn das charakteristische Polynom des infinitesimalen Generators eine doppelte Nullstelle besitzt, dann ist diese notwendigerweise reell, und auch die standardisierende Matrix B kann reell gewählt werden. Hier, wie auch im Falle von zwei verschiedenen reellen Eigenwerten, kann die Bestimmung einer standardisierenden Matrix B ganz im Reellen behandelt werden.

Wenn das charakteristische Polynom keine reellen Nullstellen hat, dann sind die Nullstellen notwendigerweise konjugiert komplex. Dieser Fall (mit negativer Spur) wird uns in der nächsten Vorlesung beim schwach gedämpften harmonischen Oszillator begegnen.

Bemerkung: Konkrete Diagonalisierung

Die Diagonalisierung einer vorgegebenen $n \times n$ -Matrix A ist eine numerische Herausforderung, wenn die Dimension größer als 2 ist.

Wir haben gesehen, dass die explizite Rechnung für $n = 2$ sehr einfach ist: Man berechnet die Spur und die Determinante und damit das charakteristische Polynom

$$p_A(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc);$$

die Nullstellen dieses quadratischen Polynoms sind die Eigenwerte der Matrix A ; wenn man die Eigenwerte hat, dann findet man die Spalten einer ‘standardisierenden Matrix’ B durch das Lösen von linearen Gleichungssystemen

$$A \cdot B = B \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{bzw.} \quad A \cdot B = B \cdot \begin{pmatrix} \tilde{\lambda} & 1 \\ 0 & \tilde{\lambda} \end{pmatrix}.$$

Wenn man nun denselben Weg für größeres n begehen wollte, dann wäre schon das Berechnen des charakteristischen Polynoms eine aufwendige Arbeit. Sodann müsste man die Nullstellen eines Polynoms vom Grade n (i. Allg. durch ein Iterationsverfahren) lösen. Schließlich sind im diagonalisierbaren Fall die Spalten einer Matrix B als Lösungen von linearen Gleichungssystemen zu bestimmen. Im Fall mehrfacher Nullstellen gibt es zusätzliche Arbeit; man wird sich übrigens zunächst damit zufrieden geben, die vorgegebene Matrix auf Dreiecksgestalt zu bringen.

Für unitär diagonalisierbare Matrizen (insbesondere für hermitesche Matrizen) ist die Lage ungleich günstiger. Es gibt direktere Iterationsverfahren zur Bestimmung einer diagonalisierenden unitären Matrix B , welche simultan die Eigenwerte liefern. Die Maximierungsmethode (nach R. Courant) haben wir bereits in IV.2 skizziert; wir werden sie in V.5 (bei der Singulärwertzerlegung) wieder aufnehmen.

Anhang: Die verschiedenen Rollen der Matrizen

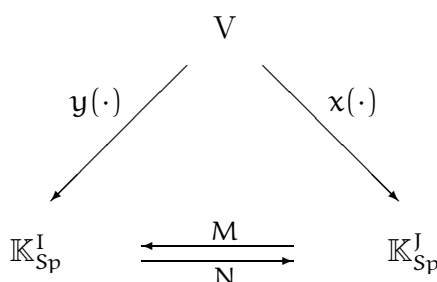
Die Matrizen können (wie die Zahlen) recht verschiedenen Zwecken dienen.

1. **Basiswechsel** Es sei V ein n -dimensionaler Vektorraum.

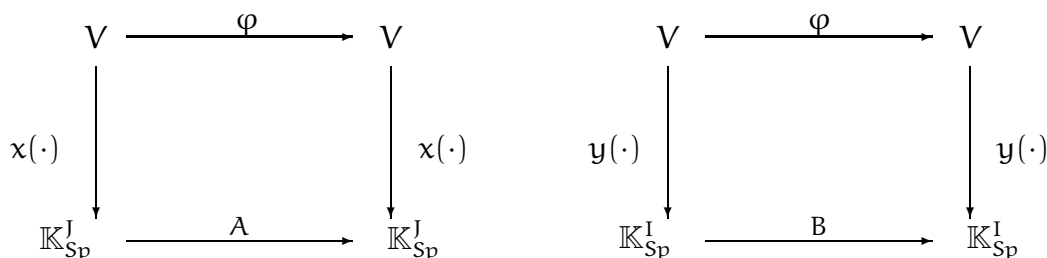
Wenn $\{u_i : i \in I\}$ und $\{v_j : j \in J\}$ Basen von V sind, dann beschreibt man den Basiswechsel durch eine $I \times J$ -Matrix M und die dazu inverse $J \times I$ -Matrix $N = M^{-1}$ folgendermaßen

$$\begin{aligned}(v_1, \dots, v_n) &= (u_1, \dots, u_n) \cdot M ; & v_j &= \sum u_i \cdot m_j^i \\ (u_1, \dots, u_n) &= (v_1, \dots, v_n) \cdot N ; & u_i &= \sum v_j \cdot n_i^j .\end{aligned}$$

Einen Vektor w kann man einerseits durch eine J -Spalte x darstellen und andererseits durch eine I -Spalte y . $\sum v_j \cdot x^j = w = \sum u_i \cdot y^i$. $Mx = y \iff x = Ny$.



2. **Endomorphismen** Sei φ eine lineare Abbildung von V in sich; $\varphi : V \rightarrow V$. Einen solchen „Endomorphismus“ kann man in „ x -Koordinaten“ durch eine $J \times J$ -Matrix A oder in „ y -Koordinaten“ durch eine $I \times I$ -Matrix B beschreiben.



$$\begin{aligned}A \circ x(\cdot) &= x \circ \varphi(\cdot) , & B \circ y(\cdot) &= y \circ \varphi(\cdot) \\ \text{Wegen } x &= Ny = M^{-1}y , & y &= Mx = N^{-1}x \text{ haben wir} \\ A \circ Ny(\cdot) &= x \circ \varphi(\cdot) = Ny \circ \varphi(\cdot) ; \\ N^{-1}ANy(\cdot) &= y \circ \varphi(\cdot) = By(\cdot) .\end{aligned}$$

Somit ergibt sich für die den Endomorphismus darstellenden Matrizen in den verschiedenen Koordinatensystemen

$$\begin{aligned}B &= N^{-1}AN = MAM^{-1} \\ A &= M^{-1}BM = NBN^{-1}\end{aligned}$$

3. **Homomorphismen** Eine (nicht notwendigerweise invertierbare) $I \times J$ -Matrix kann dazu dienen, eine lineare Abbildung $\varphi : V \rightarrow U$ zu beschreiben, wenn in V eine Basis $\{v_j : j \in J\}$ und im Zielraum U eine Basis $\{u_i : i \in I\}$ ausgezeichnet ist.

$$(\varphi(v_1), \varphi(v_2), \dots, \varphi(v_n)) = (u_1, u_2, \dots, u_m) \cdot A ;$$

$$\varphi\left(\sum v_j \cdot x^j\right) = \sum u_i \cdot \left(\sum a_j^i x^j\right) .$$

Wenn man in den beteiligten Vektorräumen zu anderen Basen übergeht, dann ändert sich die beschreibende Matrix gemäss $A \mapsto B^{-1}AC$.

(Ein Spezialfall sind die quadratischen (nicht notwendigerweise invertierbaren) Matrizen, die die Endomorphismen φ, ψ, \dots eines Vektorraums mit einem ausgezeichneten Koordinatensystem darstellen; siehe oben.)

Hintereinanderschalten

Seien φ und ψ Homomorphismen : $\varphi : U \rightarrow V$; $\psi : V \rightarrow W$.

Es seien $x(\cdot)$, $y(\cdot)$, $z(\cdot)$ Koordinatenabbildungen für U bzw. V bzw. W . Man notiert

$$\begin{array}{ccccc} U & \xrightarrow{\varphi} & V & \xrightarrow{\psi} & W \\ \downarrow x & & \downarrow y & & \downarrow z \\ \mathbb{C}_{Sp}^I & \xrightarrow{A} & \mathbb{C}_{Sp}^J & \xrightarrow{B} & \mathbb{C}_{Sp}^K \end{array}$$

Die Produktmatrix $C = B \cdot A$ ist die darstellende Matrix für die zusammengesetzte Abbildung $\psi(\varphi(\cdot))$. Man beachte die Reihenfolge $x \xrightarrow{\varphi} Ax \xrightarrow{\psi} B(Ax) = (BA)x$.

Die Matrizen tauchen auch noch in anderen Zusammenhängen auf. (Wir denken hier an die reellen und die komplexen Matrizen.)

1. Symmetrische Matrizen erscheinen als Koeffizientensysteme in einer quadratischen Form. Hermitesche Matrizen können dazu dienen, stetige hermitesche Formen (oder auch hermitesche Sesquilinearformen) (in einem gegebenen Koordinatensystem) zu beschreiben. Sie heissen dann **Gram-Matrizen**. In einem Hilbertraum mit einer gegebenen ONBasis können die hermiteschen Matrizen genausogut als Darstellung der selbstadjungierten Operatoren verstanden werden. In diesen Zusammenhängen spielt die Matrixmultiplikation keine Rolle, wohl aber die Transformation auf andere Koordinaten. $H \mapsto B^*HB$.
2. Die quadratischen (nicht notwendigerweise invertierbaren) Matrizen können auch als **infinitesimale Operatoren** von Matrixgruppen in Erscheinung treten. Wir haben am Beispiel der uniformen Drehungen angedeutet, dass in diesem Kontext

die Vektorraumstruktur bedeutungsvoll ist. Wir werden später sehen, dass daneben noch eine zweite algebraische Operation bedeutungsvoll ist, der sog. **Kommutator**. Der Kommutator ist die bilineare Operation, welche einem Paar von infinitesimalen Operatoren A, B die Matrix $AB - BA$ zuordnet; man notiert auch $[A, B]$ und nennt die Operation $[\cdot, \cdot]$ die Lie-Klammer. —Wegen dieser Struktur sollte man die infinitesimalen Operatoren nicht in einen Topf mit den Endomorphismen werfen.

Wir bemerken: Für beliebige A, B ist der Kommutator eine spurlose Matrix. Wenn A und B schiefhermitisch sind, dann ist auch der Kommutator schiefhermitisch.

3. Die Matrizen treten manchmal auch in Kontexten auf, wo keine der hier genannten Interpretationen sinnvoll oder nützlich ist. Wir werden z. B. unten (in einer ad-hoc Konstruktion) die Elemente des vierdimensionalen Vektorraums aller hermiteschen 2×2 -Matrizen als die Verschiebungsvektoren des relativistischen Raum-Zeitkontinuums interpretieren. Bei der Beschreibung der sog. Lorentzgruppe bewähren sich die formalen Möglichkeiten, die mit der Multiplikation von 2×2 -Matrizen gegeben sind, in überraschender Weise.

Hinweis

Bei den quadratischen Matrizen, welche Endomorphismen darstellen, haben sowohl die Linearkombinationen als auch die Matrizenprodukte eine vernünftige Interpretation. Wenn ϵ ein Endomorphismus ist, dann sind auch $\epsilon^2, \epsilon^3, \dots$ und allgemeiner

$$\varphi = a_0 \cdot I + a_1 \cdot \epsilon + \dots + a_n \cdot \epsilon^n$$

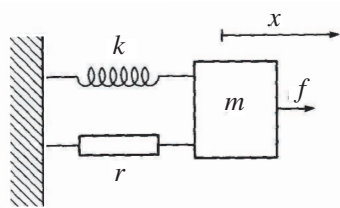
wohldefinierte Endomorphismen. Die Gesamtheit aller so zu einem einzelnen ϵ gewonnenen Endomorphismen ist eine assoziative Algebra. Diese Algebra hat eine äusserst einfache Struktur, wenn ϵ diagonalisierbar ist. Mit dem allgemeinen Fall beschäftigt man sich in der höheren linearen Algebra unter dem Stichwort ‘Jordan’sche Normalform’.

IV.4 Gekoppelte Oszillatoren

Der einfache harmonische Oszillator

Ein wichtiges Element der mathematischen Physik ist der harmonische Oszillator. Schwingungen mit vielen Freiheitsgraden kann man nämlich manchmal als eine Überlagerung von harmonischen Schwingungen begreifen. In seiner einfachsten Form erscheint der harmonische Oszillator sowohl in der elementaren Mechanik als auch in der elementaren Elektrizitätslehre.

In der Mechanik stellt man sich eine Masse m vor, die an eine Feder schwingt, deren rücktreibende Kraft proportional zur Auslenkung ist. Außerdem wirkt eine Reibungskraft proportional zur Geschwindigkeit.



Wegen Kraft = Masse \times Beschleunigung haben wir für die Auslenkung $x(t)$ als Funktion der Zeit die lineare Differentialgleichung

$$m \cdot \ddot{x} + r \cdot \dot{x} + k \cdot x = f(t) \quad ,$$

wobei $f(\cdot)$ die anregende Kraft ist.

In einem mechanischen Oszillator steckt zu jedem Zeitpunkt t ein gewisser Teil der Energie in der kinetischen Energie der bewegten Masse und ein gewisser Teil in der potentiellen Energie der gespannten Feder. Die Summe

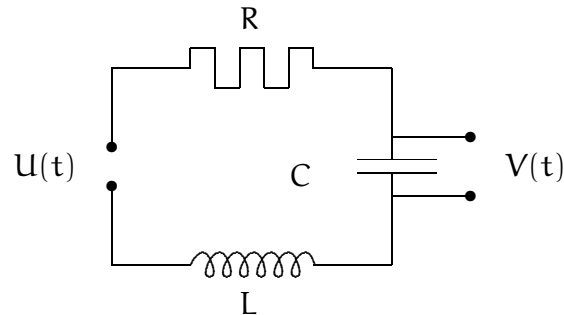
$$\frac{1}{2}m \cdot \dot{x}^2(t) + \frac{1}{2}k \cdot x^2(t)$$

ist zeitlich konstant, wenn keine äußeren Kräfte angreifen. Im allgemeineren Fall gilt für den angeregten harmonischen Oszillator

$$\frac{d}{dt} \left(\frac{1}{2}m \cdot \dot{x}^2 + \frac{1}{2}k \cdot x^2 \right) = \dot{x}[m\ddot{x} + kx] = \dot{x}[-r\dot{x} + f(t)] \quad .$$

Ein Teil der mechanischen Energie geht durch Reibung verloren; („Energiedissipation“); außerdem leistet die anregende Kraft Arbeit (Arbeit = Kraft \times Weg). Diese Leistung ist $f(t) \cdot \dot{x}(t)$.

In der Elektrizitätslehre denkt man an einen elementaren Schwingkreis. Es gibt da eine Spule (Induktivität L), einen Kondensator (Kapazität C) und einen Ohm'schen Widerstand R .



Der Spannungsabfall entlang der Schaltung ist gleich der antreibenden Spannung $U(t)$ („Elektromotorische Kraft“). Für die Ladung $Q(t)$ auf dem Kondensator gilt

$$L \cdot \ddot{Q} + R \cdot \dot{Q} + \frac{1}{C} \cdot Q = U(t) \quad .$$

Für eine einheitliche Bezeichnung setzen wir

$$\gamma = \frac{r}{m} \text{ bzw. } = \frac{R}{L} ; \quad \omega_0^2 = \frac{k}{m} \text{ bzw. } = \frac{1}{LC}$$

$$\tilde{\omega} = \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}$$

(Wir denken an den schwach gedämpften Fall $\omega_0 > \frac{1}{2}\gamma$.)

Die homogene Gleichung

$$\ddot{x} + \gamma \cdot \dot{x} + \omega_0^2 \cdot x = 0$$

löst man mit dem Ansatz $x(t) = \tilde{x} \cdot e^{i\omega t}$. Diese „reine Sinusschwingung“ ist eine Lösung, wenn ω die folgende quadratische Gleichung löst

$$(i\omega)^2 + \gamma \cdot (i\omega) + \omega_0^2 = 0 \quad .$$

Im schwach gedämpften Fall hat die quadratische Gleichung komplexe Lösungen, nämlich

$$i\omega_{+,-} = -\frac{1}{2}\gamma \pm i\sqrt{\omega_0^2 - \frac{1}{4}\gamma^2} = -\frac{1}{2}\gamma \pm i\tilde{\omega} \quad .$$

Damit erhalten wir zwei linear unabhängige Lösungen der homogenen Gleichung

$$\begin{aligned} x_+(t) &= \exp\left(-\frac{\gamma}{2} \cdot t\right) \cdot \exp(i\tilde{\omega}t) \\ x_-(t) &= \exp\left(-\frac{\gamma}{2} \cdot t\right) \cdot \exp(-i\tilde{\omega}t) \quad . \end{aligned}$$

Hinweis : In der Theorie der Differentialgleichungen zeigt man (unter sehr allgemeinen Bedingungen), dass der Lösungsraum einer skalaren Differentialgleichung zweiter Ordnung zweidimensional ist. Man zeigt nämlich, dass der Lösungsverlauf (lokal) eindeutig bestimmt ist, wenn zu irgendeiner Zeit t_0 die Auslenkung $x(t_0)$ und die Geschwindigkeit $\dot{x}(t_0)$ bekannt ist.

Das liefert für unsere spezielle Situation den

Satz

Die Lösungskurven $x(\cdot)$ der Gleichung

$$\ddot{x} + \gamma \cdot \dot{x} + \omega_0^2 \cdot x = 0$$

sind die Funktionen der Form

$$\begin{aligned} x(t) &= \exp\left(-\frac{\gamma}{2} \cdot t\right) [A \cdot e^{i\tilde{\omega}t} + B \cdot e^{-i\tilde{\omega}t}] \\ \text{wo } \tilde{\omega} &= \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2} \text{ und } A, B \text{ beliebig } \in \mathbb{C}. \end{aligned}$$

Bemerkung

Die reellen Lösungen kann man auch folgendermaßen schreiben

$$\begin{aligned} x(t) &= |A| \cdot \exp\left(-\frac{\gamma}{2}t\right) \cdot \cos(\tilde{\omega}(t - t_0)) \\ \text{mit } |A|, t_0 &\in \mathbb{R} \text{ beliebig.} \end{aligned}$$

Hinweis :

Mit der inhomogenen Gleichung werden wir uns eingehend beschäftigen, wenn wir Integrationstheorie und insbesondere Fourier-Integrale studieren.

Im Folgenden wollen wir den Satz von der Hauptachsentransformation anwenden, um gekoppelte harmonische Oszillatoren zu studieren. Wir beschränken uns auf den ungedämpften Fall.

Zwei schwach gekoppelte Oszillatoren

Wir denken an ein mechanisches System, welches aus zwei ungedämpften Oszillatoren besteht, die durch eine schwache Feder miteinander verbunden sind. Die Bewegungsgleichungen lauten:

$$\begin{aligned} m_1 \cdot \ddot{x}_1 &= -k_1 \cdot x_1 - h(x_1 - x_2) \\ m_2 \cdot \ddot{x}_2 &= -k_2 \cdot x_2 - h(x_2 - x_1). \end{aligned}$$

Mit Matrizen geschrieben

$$T \cdot \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = -H \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

wobei

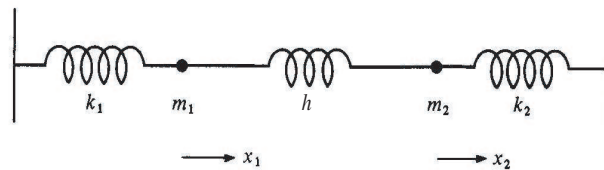
$$T = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}; \quad H = \begin{pmatrix} k_1 + h & -h \\ -h & k_2 + h \end{pmatrix}.$$

Wir betrachten h als einen Parameter.

Ohne Koppelung ($h = 0$)



Mit Koppelung ($h > 0$)



Der symmetrische Fall

Wenn die beiden Oszillatoren identisch sind ($m_1 = m_2 = m$ und $k_1 = k_2 = k$), dann kann man leicht einige spezielle Lösungen erraten. Die Oszillatoren können beispielsweise miteinander („synchron“) oder auch gegeneinander schwingen. Im zweiten Fall wird die Kreisfrequenz größer sein als im ersten $\omega_2 > \omega_1$.

Die Lösungen der homogenen Schwingungsgleichungen sind die komplexen Linearkombinationen der vier speziellen Lösungen

$$e^{\pm i\omega_1 t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad e^{\pm i\omega_2 t} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Wir müssen nur noch die Kreisfrequenzen $\omega_1 < \omega_2$ berechnen. Sie hängen von $\omega_0^2 := \frac{k}{m}$ und $s := \frac{h}{m}$ ab.

$$\begin{aligned} \ddot{x}_1 &= -(\omega_0^2 + s)x_1 + s \cdot x_2 \\ \ddot{x}_2 &= sx_1 - (\omega_0^2 + s) \cdot x_2 \\ (i\omega_1)^2 &= -(\omega_0^2 + s) + s = -\omega_0^2 \\ (i\omega_2)^2 &= -(\omega_0^2 + s) - s = -(\omega_0^2 + 2s) \end{aligned}$$

Also $\omega_1 = \omega_0$ und $\omega_2 = \sqrt{\omega_0^2 + 2s} \sim \omega_0 + \frac{s}{\omega_0}$.

Satz :

Die **reellen Lösungen** haben die Gestalt

$$\begin{aligned} x_1(t) &= \rho_1 \cdot \cos(\omega_1 t + \alpha_1) + \rho_2 \cdot \cos(\omega_2 t + \alpha_2) \\ x_2(t) &= \rho_1 \cdot \cos(\omega_1 t + \alpha_1) - \rho_2 \cdot \cos(\omega_2 t + \alpha_2). \end{aligned}$$

Beweis

$$A_+ e^{i\omega t} + A_- e^{-i\omega t} = (A_+ + A_-) \cos \omega t + i \cdot (A_+ - A_-) \sin \omega t .$$

Der Realteil ergibt sich aus

$$\begin{aligned} a &= \Re(A_+ + A_-) \quad \text{und} \quad b = \Im(A_+ - A_-) \\ a \cdot \cos \omega t - b \sin \omega t &= \sqrt{a^2 + b^2} \cdot [\cos \alpha \cdot \cos \omega t - \sin \alpha \cdot \sin \omega t] \\ &= \sqrt{a^2 + b^2} \cdot \cos(\omega t + \alpha) \end{aligned}$$

Diese Rechnung machen wir für ω_1 und für ω_2 . Der Satz ist bewiesen.

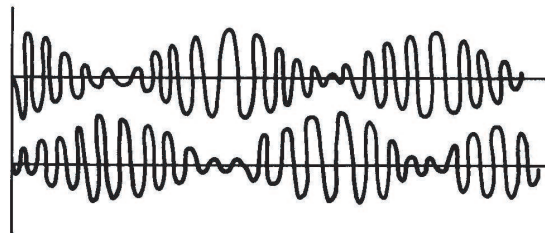
Bemerkung

1. Wenn $\rho_1 = 0$ oder $\rho_2 = 0$, dann haben wir periodische Lösungen; im Allgemeinen sind die Lösungen sog. fastperiodische Funktionen der Zeit.
2. Die reellen Zahlen $\rho_1, \rho_2 \geq 0$, $\alpha_1, \alpha_2 \in \mathbb{R}/2\pi$ sind durch die Anfangsbedingungen bestimmt. Wir können für jeden der Oszillatoren die anfängliche Auslenkung und die Anfangsgeschwindigkeit vorgeben.

Beispiel : (ähnlich wie im Bild)

$$x_1(0) = 1, \quad \dot{x}_1(0) = 0, \quad x_2(0) = 0, \quad \dot{x}_2(0) = 0 .$$

Der erste Oszillator wird zur Zeit 0 ausgelenkt, der zweite ist in der Nulllage; beide sind anfangs in Ruhe. Wir zeigen, dass der zweite Oszillator nach und nach die Energie übernimmt, die wir in die erste Feder gesteckt haben. Der erste Oszillator wird gebremst, bis er zu einem gewissen Zeitpunkt zur Ruhe kommt. Danach wird er wieder vom schwingenden zweiten Oszillator angeregt. Das Phänomen der Energieübertragung von einem schwingfähigen System auf ein anderes heißt **Resonanz**.



In unserem symmetrischen Beispiel haben wir den eher seltenen Fall, dass alle Energie des ursprünglich angeregten Oszillators auf den zweiten Oszillator hinüber und dann wieder herüber geht. In allgemeineren Fällen (z.B. $m_1 \neq m_2$ oder $k_1 \neq k_2$) sinkt die Energie im ursprünglich angeregten Oszillator nur bis zu einem Wert > 0 ; dann kommt sie fastperiodisch zurück.

Hier lässt sich die Rechnung wirklich explizit durchführen.
Die Anfangsbedingungen liefern

$$\begin{aligned}x_1(t) &= \frac{1}{2} \cos(\omega_1 t) + \frac{1}{2} \cos(\omega_2 t) \\x_2(t) &= \frac{1}{2} \cos(\omega_1 t) - \frac{1}{2} \cos(\omega_2 t) .\end{aligned}$$

Die Formeln werden übersichtlich, wenn wir die Additionstheoreme für die trigonometrischen Funktionen benutzen

$$\begin{aligned}\cos \alpha + \cos \beta &= 2 \cdot \cos \left(\frac{\alpha - \beta}{2} \right) \cdot \cos \left(\frac{\alpha + \beta}{2} \right) \\ \cos \alpha - \cos \beta &= -2 \cdot \sin \left(\frac{\alpha - \beta}{2} \right) \cdot \sin \left(\frac{\alpha + \beta}{2} \right) .\end{aligned}$$

Mit den Konstanten

$$\begin{aligned}\tilde{\omega} &= \frac{1}{2}(\omega_1 + \omega_2) \approx \sqrt{\frac{k}{m}} + \frac{s}{2\omega_0} \\ \delta &= \frac{1}{2}(\omega_1 - \omega_2) \approx \frac{s}{2\omega_2} = \frac{h}{2k} \sqrt{\frac{k}{m}} \quad (\text{für } \delta \text{ klein})\end{aligned}$$

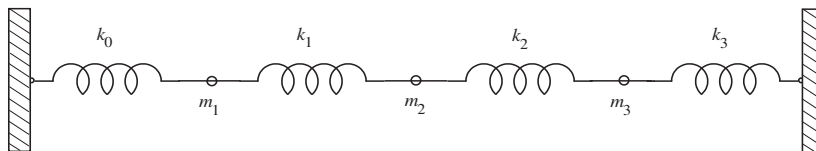
erhalten wir die „amplitudenmodulierten“ Schwingungen

$$\begin{aligned}x_1(t) &= \cos(\delta t) \cdot \cos(\tilde{\omega} t) \\ x_2(t) &= \sin(\delta t) \cdot \sin(\tilde{\omega} t)\end{aligned}$$

Diese sieht man im Bild.

Drei Oszillatoren in symmetrischer Anordnung

Wir denken jetzt an drei Massen m_1, m_2, m_3 , die untereinander und mit festen Verankerungen durch Federn verbunden sind, wie das Bild zeigt:



Wenn wir die Massen in vertikaler Richtung aus der Ruhelage heraus auslenken oder anstoßen und dann das System sich selbst überlassen, dann entstehen Schwingungen $x^1(t), x^2(t), x^3(t)$ in vertikaler Richtung („Transversalschwingungen“) Bei kleinen Auslenkungen können wir die Kräfte in longitudinaler Richtung vernachlässigen; die resultierende rücktreibende Kraft wirkt nämlich im Wesentlichen in vertikaler Richtung. Diese Kräfte erzeugen Beschleunigungen

$$m_1 \cdot \ddot{x}^1 = -k_0 \cdot x^1 - k_1(x^1 - x^2)$$

$$m_2 \cdot \ddot{x}^2 = -k_1 \cdot (x^2 - x^1) - k_2(x^2 - x^3)$$

$$m_3 \cdot \ddot{x}^3 = -k_2 \cdot (x^3 - x^2) - k_3 \cdot x^3$$

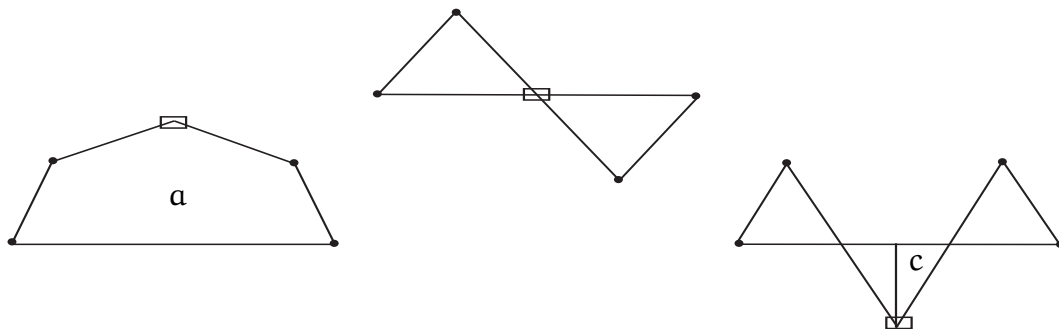
$$\begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix} \begin{pmatrix} \ddot{x}^1 \\ \ddot{x}^2 \\ \ddot{x}^3 \end{pmatrix} = \begin{pmatrix} k_0 + k_1 & -k_1 & 0 \\ -k_1 & k_1 + k_2 & -k_2 \\ 0 & -k_2 & k_2 + k_3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

$$T \cdot \ddot{x} = -Hx$$

Hier ist T eine sehr einfache positiv definite Matrix; die linke Seite ist die Änderung des Impulses. H ist ebenfalls positiv definit; die rechte Seite ist der „Gradient“ der potentiellen Energie in den Federn. Wir diskutieren den symmetrischen Fall

$$m_1 = m_3 = 1, \quad k_0 = k_3 = k, \quad k_1 = k_2 = h.$$

Es bleiben drei Parameter: die Masse in der Mitte sowie die Federkonstanten h, k . Die Form der Eigenschwingungen können wir im Wesentlichen erraten.



Die zweite Eigenschwingung ist klar. Es kommt darauf hinaus, dass wir in der ersten und der dritten Position einen harmonischen Oszillator mit der Federkonstanten $h + k$ haben. $\omega_2 = \sqrt{h + k}$.

Für die Grundschiwingung mit der noch unbekannten Kreisfrequenz ω_1 machen wir den Ansatz

$$(i\omega_1)^2 \begin{pmatrix} 1 \\ ma \\ 1 \end{pmatrix} = - \begin{pmatrix} h+k & -h & 0 \\ -h & 2h & -h \\ 0 & -h & h+k \end{pmatrix} \begin{pmatrix} 1 \\ a \\ 1 \end{pmatrix}$$

Dies ergibt die Gleichungen

$$\omega_1^2 = (k + h) - ha, \quad \omega_1^2 \cdot ma = -2h + 2ha.$$

Eliminieren wir ω_1^2 , so erhalten wir eine quadratische Gleichung für \mathbf{a}

$$2h(\mathbf{a} - 1) = m \cdot \mathbf{a} \cdot [k - h(\mathbf{a} - 1)] .$$

Die eine Lösung ist das gesuchte \mathbf{a} , die andere ist für die Grundschiwingung offenbar die Zahl $c < 0$ zur Frequenz ω_3 . Diese Zahlen liefern die gesuchten Frequenzen

$$\omega_1 = \sqrt{k + h - h\mathbf{a}} , \quad \omega_2 = \sqrt{k + h} , \quad \omega_3 = \sqrt{k + h - hc} .$$

Bezeichnen wir

$$\mathbf{n}_1 = \begin{pmatrix} 1 \\ \mathbf{a} \\ 1 \end{pmatrix} ; \mathbf{n}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} ; \mathbf{n}_3 = \begin{pmatrix} 1 \\ c \\ 1 \end{pmatrix} ;$$

Die Lösungen unserer Schwingungsgleichung sind

$$\mathbf{x}(t) = (\mathbf{A}_+ \cdot e^{i\omega_1 t} + \mathbf{A}_- e^{-i\omega_1 t}) \mathbf{n}_1 + (\mathbf{B}_+ \cdot e^{i\omega_2 t} + \mathbf{B}_- e^{-i\omega_2 t}) \mathbf{n}_2 + (\mathbf{C}_+ \cdot e^{i\omega_3 t} + \mathbf{C}_- e^{-i\omega_3 t}) \mathbf{n}_3$$

mit beliebigen komplexen Koeffizienten $\mathbf{A}_+, \mathbf{A}_-, \mathbf{B}_+, \mathbf{B}_-, \mathbf{C}_+, \mathbf{C}_-$. Wie oben können wir die reellen Lösungen auch folgendermaßen schreiben: Wir haben die zeitabhängige 3-Spalte

$$\mathbf{x}(t) = \rho_1 \cdot \cos(\omega_1 t + \alpha_1) \cdot \mathbf{n}_1 + \rho_2 \cdot \cos(\omega_2 t + \alpha_2) \cdot \mathbf{n}_2 + \rho_3 \cdot \cos(\omega_3 t + \alpha_3) \cdot \mathbf{n}_3 .$$

Betrachten wir speziell die Lösung zu der Anfangssituation, wo die erste Masse ausgelenkt, aber anfangs in Ruhe ist, während sich die beiden anderen in der Ruhelage befinden. Diese Anfangsbedingung liefert uns Zahlen ρ_1, ρ_2, ρ_3 und die Lösung

$$\mathbf{x}(t) = \rho_1 \cdot \cos(\omega_1 t) \cdot \mathbf{n}_1 + \rho_2 \cdot \cos(\omega_2 t) \cdot \mathbf{n}_2 + \rho_3 \cdot \cos(\omega_3 t) \cdot \mathbf{n}_3 .$$

Auch hier wird man sich überzeugen, dass die Energie zunächst nach und nach vom ersten Oszillator weggeht und dann wieder zurück kommt. Wenn die Zahlen $\omega_1, \omega_2, \omega_3$ kommensurabel sind, dann sind die Lösung periodisch, andernfalls fastperiodisch.

Hinweis :

Anschauungsmaterial findet man z.B. unter <http://www.mapleapps.com/Info.asp?ArticleID=134>.

Die „Normal modes“ im allgemeinen Fall

Wir denken (der Übersichtlichkeit halber) an Massenpunkte $\mathbf{m}_1, \dots, \mathbf{m}_n$, die sich in der Ruhelage in Punkten $\mathbf{P}_1, \dots, \mathbf{P}_n$ der Ebene befinden. Sie seien mit Federn untereinander und mit irgendwelchen festen Punkten so verbunden, dass für kleine vertikale Auslenkungen die resultierenden rücktreibenden Kräfte im Wesentlichen vertikal gerichtet sind. (Wir können auch daran denken, dass die Massen \mathbf{m}_i auf vertikalen Schienen gleiten.) Jede Auslenkung \mathbf{x} (I-Tupel) liefert eine erhöhte potentielle Energie im System der Federn. Diese potentielle Energie ist eine quadratische Funktion der Auslenkung: $Q(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x}$. Die rücktreibenden Kräfte sind daher lineare Funktionen

$$\mathbf{k}_i = - \sum_j \mathbf{a}_{ij} \cdot \mathbf{x}^j \quad \text{für } i \in I .$$

Newton's Gesetz liefert

$$m_i \cdot \ddot{x}_i(t) = k_i(t) = - \sum_j a_{ij} \cdot x^j(t) .$$

Somit haben wir die **Gleichung für die ungedämpfte Schwingung**

$$T \cdot \ddot{x} = -A \cdot x$$

wo T die Diagonalmatrix mit den Diagonalelementen m_i ist und A eine symmetrische Matrix.

Lösung der Schwingungsgleichung

Nach dem Satz von der simultanen Diagonalisierung gibt es eine $I \times J$ -Matrix, sodass

$$N^* T N = E , \quad N^* A N = \Lambda \quad (\text{Diagonalmatrix})$$

Wir schreiben $x(t) = N y(t)$ und erhalten eine Lösung $x(\cdot)$ von $T \ddot{x} = -A \cdot x$, wenn

$$\ddot{y} = -\Lambda y , \text{ d.h. } y^j(t) = \exp(\pm i \omega_j t) \quad \text{mit } \omega_j = \sqrt{\lambda_j}, j = 1, \dots, n .$$

In der Tat haben wir

$$\ddot{y} = N^* T N \ddot{y} = N^* T \ddot{x} = -N^* A \cdot x = -N^* A N y .$$

Seien $\{n_j : j \in J\}$ die Spalten der Matrix N . Sie heißen die **normal modes** zu den Kreisfrequenzen ω_j .

Die Lösungen von $T \ddot{x} = -A x$ sind die I-Spalten von der Form

$$x(t) = \sum_{j \in J} \left(A_j^{(+)} e^{i \omega_j t} + A_j^{(-)} e^{-i \omega_j t} \right) \cdot n_j .$$

Die Spalte n_j erfüllt die Gleichung

$$A n_j = T n_j \cdot \omega_j^2 .$$

Wenn wir also n_j (bis auf eine Konstante) erraten können, dann berechnet sich leicht die dazugehörige Kreisfrequenz. In der Tat ergibt sich $AN = TN \cdot \Lambda$ aus der Formel für die simultane Diagonalisierung

$$\begin{aligned} N^* A N &= \Lambda , \quad N^* T N = E \\ A N &= (N^*)^{-1} \cdot \Lambda = T N \cdot \Lambda . \end{aligned}$$

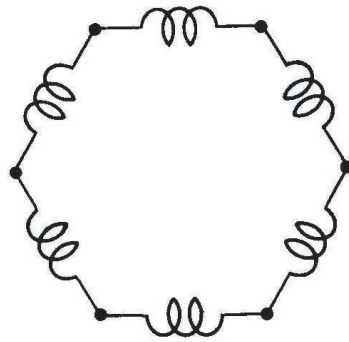
Für die Bestimmung der Lösungsmenge unserer Gleichung $T \ddot{x} = -A x$ brauchen wir also die normal modes nur bis auf eine Konstante; die Normierung $n_j^* T n_j = 1$ ist irrelevant.

Verallgemeinerung

Denken wir an eine Folge von n Massenpunkten, die entlang einer Geraden mit den nächsten Nachbarn verbunden sind und an den beiden Enden mit festen Verankerungen. Wir haben n normal modes. Der mit der geringsten Frequenz ω_1 hat einen Bauch, der zu ω_2 hat zwei Bäuche, usw. Bei dem „normal mode“ zur höchsten Frequenz ω_{n-1} haben die Auslenkungen alternierendes Vorzeichen. Interessant ist nun, wie die Frequenz von der Anzahl der Bäuche abhängt („Dispersionsrelation“). Es zeigt sich, dass die Kreisfrequenz $\Omega(k)$ für kleine k ungefähr linear anwächst. Für sehr große n erinnert die Anordnung an eine Saite, die an den Enden eingespannt ist. Die Frequenzen der Eigenschwingungen hängen im Grenzfall $n \rightarrow \infty$ in linearer Weise von der Anzahl der Knoten ab.

Wellen über einem Kreisring

Wir diskutieren jetzt einen so schön symmetrischen Fall, dass wir alle normal modes und alle Frequenzen explizit ausrechnen können. Bei diesem Beispiel gibt es nicht nur stehende Wellen mit eins, zwei, \dots , $n-1$ Bergen bzw. Tälern, die ihre Knoten irgendwo haben können, es gibt auch laufende Wellen, die rechts herum oder links herum mit Kreisfrequenzen $\omega_1, \omega_2, \dots, \omega_{[\frac{n}{2}]}$ laufen können. Was es mit den laufenden Wellen auf sich hat, werden wir dann im kontinuierlichen Fall (Abschnitt IV.5) noch deutlicher sehen. n Massenpunkte seien in einem Kreis angeordnet. Sie sind mit ihren Nachbarn durch Federn verbunden sind. Alle Massen seien m , alle Federkonstanten seien k .



$$m \cdot \ddot{x}^j = -k(x^j - x^{j+1}) - k(x^j - x^{j-1}) \quad \text{für } j = 0, 1, 2, \dots$$

Dabei ist $x^{n+j}(t) = x^j(t)$ zu setzen. Mit $\omega^2 = \frac{k}{m}$ haben wir also

$$\begin{aligned} \ddot{x}^j &= -\omega^2 \cdot [2x^j - x^{j-1} - x^{j+1}] \quad \text{für } j \in \mathbb{Z}/n\mathbb{Z} \\ \ddot{\mathbf{x}} &= -\omega^2 \cdot \mathbf{A} \mathbf{x}, \end{aligned}$$

wobei \mathbf{A} die Matrix ist, die in der Diagonalen den Wert 2 hat, in den beiden Nebendiagonalen den Wert -1 und noch zusätzlich den Eintrag -1 in der rechten oberen und in der linken unteren Ecke. Bezeichnet \mathbf{P} die Permutationsmatrix, welche die Einheitsspalten zyklisch vertauscht, so gilt

$$\mathbf{A} = 2 \cdot \mathbf{E} - \mathbf{P} - \mathbf{P}^{-1}.$$

Diese Matrix ist singulär. Wenn $\tilde{\mathbf{x}}_0$ die Spalte mit lauter Einsen ist, dann gilt $\mathbf{A}\tilde{\mathbf{x}}_0 = 0$. Zu diesem $\tilde{\mathbf{x}}_0$ gehören zwei linear unabhängige Lösungen der Schwingungsgleichung, nämlich

$$\mathbf{x}(t) = \text{const} \cdot \tilde{\mathbf{x}}_0 \quad \text{und} \quad \mathbf{x}(t) = \text{const} \cdot (t\tilde{\mathbf{x}}_0).$$

Von Schwingungen kann man hier nicht reden; uns interessieren die übrigen $(n-1)$ normal modes und ihre Frequenzen $\omega_1, \omega_2, \dots, \omega_{n-1}$.

Die normal modes können wir leicht erraten. Für jede n -te Einheitswurzel $\tau = \exp\left(i k \cdot \frac{2\pi}{n}\right)$ ist $\tilde{\mathbf{x}} = (\tau, \tau^2, \dots, \tau^{n-1}, 1)^T$ ein Eigenvektor; denn $\mathbf{P}\tilde{\mathbf{x}} = \tau\tilde{\mathbf{x}}$, $\mathbf{P}^{-1} = \frac{1}{\tau}\tilde{\mathbf{x}} = \bar{\tau} \cdot \tilde{\mathbf{x}}$

$$\mathbf{A}\tilde{\mathbf{x}} = (2 - \tau - \bar{\tau})\tilde{\mathbf{x}} = \left(2 - 2\cos\left(k\frac{2\pi}{n}\right)\right) \cdot \tilde{\mathbf{x}} = 4\sin^2\left(k \cdot \frac{\pi}{n}\right) \cdot \tilde{\mathbf{x}}.$$

Somit haben wir die Eigenfrequenzen

$$\omega_k = \omega \cdot \sin\left(k \cdot \frac{\pi}{n}\right) \quad k = 0, 1, 2, \dots, n-1$$

Die dazugehörigen Eigenvektoren sind paarweise orthogonal. Für $k \neq \ell$ gilt nämlich

$$\begin{aligned} \tilde{x}_\ell^* \cdot \tilde{x}_k &= \sum_{j=1}^n \exp\left(-j \cdot 2\pi i \frac{\ell}{n}\right) \cdot \exp\left(+j \cdot 2\pi i \frac{k}{n}\right) \\ &= \sum_{j=1}^n \exp\left(j \cdot \frac{2\pi i(k-\ell)}{n}\right) = \sum_{j=1}^n \tau^j = 0. \end{aligned}$$

Bemerke : Diese Orthogonalität ist analog zu der Orthogonalitätsrelation

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(-i\ell t) \cdot \exp(ikt) dt = \begin{cases} 0 & \text{falls } k = \ell \\ 1 & \text{falls } k \neq \ell \end{cases}$$

Wir haben nun alle Lösungen der Gleichung

$$\ddot{x} = -\omega^2 \cdot Ax.$$

Es sind die Linearkombinationen der folgenden Spalten

$$\exp(\pm i\omega_1 t) \cdot \tilde{x}_1, \dots, \exp(\pm i\omega_{n-1} t) \cdot \tilde{x}_{n-1},$$

sowie \tilde{x}_0 und $t \cdot \tilde{x}_0$, wo \tilde{x}_0 die Spalte mit lauter Einsen ist.

Die Einträge in \tilde{x}_k sind die Potenzen der Einheitswurzel $\exp\left(ik \cdot \frac{2\pi}{n}\right)$.

Die Realteile und die Imaginärteile der normal modes des \tilde{x}_k entsprechen den Sinuskurven auf dem Kreisring. Es gilt

1. \tilde{x}_{n-k} und \tilde{x}_k sind komplex konjugiert; denn

$$\begin{aligned} \exp\left(i(n-k) \cdot \frac{2\pi}{n}\right) &= \exp\left(-ik \cdot \frac{2\pi}{n}\right) \\ \frac{1}{2}(\tilde{x}_k + \tilde{x}_{n-k}) &\quad \text{und} \quad \frac{1}{2i}(\tilde{x}_k - \tilde{x}_{n-k}) \end{aligned}$$

sind reelle Spalten.

2. Die Punkte P_j , über denen die Massen m_j schwingen, können wir mit Punkten des Einheitskreises identifizieren

$$P_j \leftrightarrow \exp\left(ij \cdot \frac{2\pi}{n}\right) \quad j \in \mathbb{Z}.$$

Ein J -Tupel $(y^j)_{j=1,\dots,n}$ entspricht einer 2π -periodischen Funktion $y(\varphi)$, die nur in den Punkten $\varphi_j = j \cdot \frac{2\pi}{n}$ ($j \in \mathbb{Z}$) beobachtet wird. Beispielsweise

$$\begin{aligned} \tilde{x}_k &\longleftrightarrow \exp(ik\varphi) \quad \text{für } \varphi = j \cdot \frac{2\pi}{n}, \quad j \in \mathbb{Z} \\ \tilde{x}_{n-k} &\longleftrightarrow \exp(-ik\varphi) \quad \text{für } \varphi = j \cdot \frac{2\pi}{n}, \quad j \in \mathbb{Z} \\ \Re \tilde{x}_k &= \Re \tilde{x}_{n-k} \longleftrightarrow \cos(k\varphi) \\ \Im \tilde{x}_k &= -\Im \tilde{x}_{n-k} \longleftrightarrow \sin(k\varphi). \end{aligned}$$

3. Wir interessieren uns für die Lösungen mit den Eigenfrequenzen $\pm\omega_k = \pm\omega \sin\left(\frac{k\pi}{n}\right)$, $k = 1, 2, \dots$. Hier müssen wir zwei Fälle unterscheiden. Wenn n ungerade ist, dann erhalten wir für $k = 1, 2, \dots, \frac{n-1}{2}$ je vier linear unabhängige Lösungen

$$\begin{aligned} & \exp(i\omega_k t) \cdot \cos(k\varphi), \quad \exp(i\omega_k t) \cdot \sin(k\varphi) \\ & \exp(-i\omega_k t) \cdot \cos(k\varphi), \quad \exp(-i\omega_k t) \cdot \sin(k\varphi). \end{aligned}$$

Wenn n gerade ist, dann liefern $k = 1, \dots, \frac{n}{2}-1$ je vier linear unabhängige Lösungen, während für $k = \frac{n}{2}$ nur zwei linear unabhängige Lösungen existieren; denn

$$\begin{aligned} \sin\left(\frac{n}{2}\varphi\right) &= \sin\left(\frac{n}{2} \cdot j \cdot \frac{2\pi}{n}\right) = \sin(j \cdot \pi) \equiv 0 \\ \cos\left(\frac{n}{2}\varphi_j\right) &= \cos(j\pi) \text{ ist alternierend } +1 \text{ und } -1. \end{aligned}$$

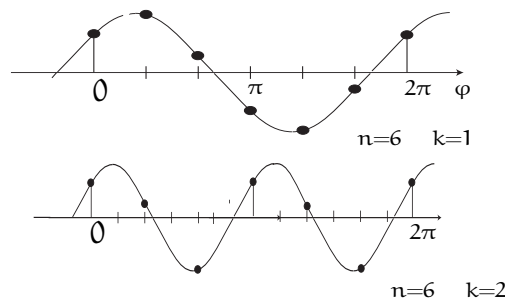
Es sind uns bei diesen Umrechnungen keine Lösungen verloren gegangen. Im geraden Fall haben wir (neben den Nichtschwingungen \tilde{x}_0 und $t \cdot \tilde{x}_0$) $4\left(\frac{n}{2}-1\right) = 2(n-1)$ linear unabhängige Lösungen, im ungeraden Fall ebenfalls $4\frac{n-1}{2} = 2(n-1)$.

Fazit : Die normal modes sind sinusförmig. Für $k = 1$ haben wir genau einen Nulldurchgang von positiv nach negativ. Für $k = 2$ haben wir genau zwei solche Nulldurchgänge. Im geraden Fall sind für $k = \frac{n}{2}$ die Funktionswerte dieses normal mode alternierend positiv und negativ.

4. Betrachten wir für $1 \leq k \leq \frac{n}{2}$ die reelle Linearkombination

$$\begin{aligned} x(t, \varphi) &= a \cdot e^{i\omega_k t} \cos(k\varphi) + b \cdot e^{i\omega_k t} \cdot \sin(k\varphi) \\ &= \sqrt{a^2 + b^2} \cdot \exp(i\omega_k t) \cdot \cos(k\varphi + \alpha) \end{aligned}$$

Es handelt sich um eine **stehende Welle** mit $2k$ Knoten (k heißt die Wellenzahl)



Die Bilder zeigen eine Momentaufnahme des Realteils oder des Imaginärteils von $x(t, \varphi)$.

5. Ebenso interessant sind die **laufenden Wellen** zu einem festen k , $1 \leq k \leq \frac{n}{2}$

$$\begin{aligned} & \exp(i\omega_k t) \cdot \left[A \cdot (\cos k\varphi + i \sin k\varphi) + B(\cos k\varphi - i \sin k\varphi) \right] \\ &= A \cdot \exp(i[\omega_k t + k\varphi]) + B \cdot \exp(i[\omega_k t - k\varphi]). \end{aligned}$$

Die Summanden sind gegeneinander laufende Wellen. Betrachten wir den zweiten Summanden

$$\exp(i\omega_k t - ik\varphi) = \exp\left(-ik\left[\varphi - \frac{\omega_k}{k}t\right]\right) .$$

Es handelt sich um eine laufende Welle mit der Wellenzahl k und der Phasengeschwindigkeit $\frac{\omega_k}{k}$. Der Realteil $\cos(k[\varphi - vt])$ hat seinen Maximalwert dort wo $\varphi - vt$ in der Nähe eines ganzzahligen Vielfachen von 2π hat. Der Wellenkamm bewegt sich mit der Geschwindigkeit v .

Wenn man von k zu $-k$ (oder zu $n - k$) über geht, dann erhält man die in der Gegenrichtung laufende Welle.

Bemerke : Im geraden Fall verschwindet für $k = \frac{n}{2}$ der Unterschied zwischen einer laufenden und einer stehenden Welle (weil wir nur in diskreten Positionen beobachten).

6. Wenn man laufende Wellen zu verschiedenen Wellenzahlen k in geeigneter Weise überlagert, dann erhält man laufende „Wellenpakete“. Solche Wellenpakete werden wir im kontinuierlichen Fall an geeigneter Stelle diskutieren (Teil „Asymptotisches Verhalten“). Wir werden sehen, dass die „Dispersionsrelation“ eine entscheidende Rolle spielt; die Dispersionsrelation $\Omega(k)$ gibt an, wie die Kreisfrequenz von der Wellenzahl abhängt. In unserem Fall hier haben wir $\Omega(k) = \omega \sin\left(k \cdot \frac{\pi}{n}\right)$, $k = 1, 2, \dots$. Für $n \rightarrow \infty$ und entsprechende Dimensionierung der Massen und der Federkonstanten erhalten wir für $k = 1, 2, \dots$ einen linearen Anstieg der Kreisfrequenzen als Funktion von k . Die Phasengeschwindigkeit $\frac{1}{k}\omega_k$ ist für die kleinen k nahezu unabhängig von der Wellenzahl.

IV.5 Die schwingende Saite

Stellen wir uns eine elastische Saite vor, die an den Stellen $x = 0$ und $x = L$ eingespannt ist. Sie sei mit einer Massendichte $m(x)dx$ belegt und habe in jedem x eine Elastizität $E(x)$, welche angibt, welche Kraft der Dehnung entgegensteht, wobei angenommen wird, dass diese Kraft proportional zur Dehnung ist.

In dem üblichen idealisierenden Modell (für kleine Auslenkungen) hat man nach Newton's Formel

$$m(x) \cdot \frac{\partial^2}{\partial t^2} u(t, x) = E(x) \frac{\partial^2}{\partial x^2} u(t, x) \quad .$$

(Masse \times Beschleunigung = rücktreibende Kraft)?

Dass die Dehnung mit der zweiten Ableitung zu tun hat, liegt im Falle der Longitudinalwellen auf der Hand. Bei den transversalen Schwingungen erscheint die rücktreibende Kraft als eine Resultierende der entlang der Saite angreifenden Kräfte.

Uns soll hier nur der Fall interessieren, wo $m(\cdot)$ und $E(\cdot)$ entlang der Saite konstant sind.

Mit $c = \sqrt{\frac{E}{m}}$ haben wir dann die

Gleichung der schwingenden Saite

$$\left(\frac{1}{c^2} \cdot \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) u(t, x) = 0$$

$$u(t, 0) = 0 = u(t, L) \quad \text{für alle } t \quad .$$

Die Lösung ist durch die Auslenkung und die Geschwindigkeit zur Zeit eindeutig bestimmt. Seien also $u(0, x)$ und $\dot{u}(0, x)$ gegeben.

Hinweise :

- a) Die analoge Gleichung in zwei Raumdimensionen heißt die

Gleichung der schwingenden Membrane

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \right) u(t, x) = 0$$

$$u(t, (x, y)) = 0 \quad \text{für } (x, y) \text{ am Rand der Membran.}$$

Für die Anfangsbedingungen zur Zeit 0 gilt dasselbe wie für die schwingende Saite.

b) In drei Raumdimensionen wird daraus die

Wellengleichung

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) u(t, x, y, z) = 0$$

wobei $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ „Laplace Operator“

c) Lassen wir alle Randbedingungen außer Acht, so können wir schon einmal Lösungen eines interessanten Typs hinschreiben, die sogenannten ebenen Wellen.

$$u(t, x, y, z) = \exp\left(i(\omega t - (k_1 x + k_2 y + k_3 z))\right)$$

mit einer beliebigen „Wellenzahl“ (k_1, k_2, k_3) und der dazu passenden „Kreisfrequenz“

$$\omega = \|k\| \cdot c \quad .$$

Man hat in der Tat

$$\Delta u(t, x, y, z) = (-k_1^2 - k_2^2 - k_3^2) \cdot u = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} u(t, x, y, z) \quad .$$

Die ebenen Wellen sind ein besonders wichtiger Fall von Lösungen, die man faktorisieren kann.

d) Machen wir nun den allgemeinen „Separationsansatz“ :

$$u(t, x, y, z) = g(t) \cdot h(x, y, z) \quad .$$

Die Wellengleichung fordert für Lösungen dieser Gestalt

$$\frac{1}{c^2} g''(t) \cdot h(x) = g(t) \cdot \Delta h(x)$$

und das bedeutet, dass eine Konstante existiert, sodass

$$\frac{1}{c^2} \frac{g''(t)}{g(t)} = \text{const} = \frac{\Delta h(x)}{h(x)} \quad .$$

Die Lösungen der Gleichung

$$g''(t) = -\omega^2 \cdot g(t)$$

sind natürlich nichts anderes als die Linearkombinationen der Funktionen

$$g(t) = e^{i\omega t} \quad \text{und} \quad g(t) = e^{-i\omega t}$$

oder $g(t) = \cos(\omega t)$ und $g(t) = \sin(\omega t)$.

Sei auf der anderen Seite $h(\cdot)$ eine Lösung der Gleichung

$$\Delta h = -k^2 \cdot h \quad .$$

Dann ist mit $\omega^2 = k^2 \cdot c^2$

$$\exp(i\omega t) \cdot h(x)$$

eine Lösung der Wellengleichung.

Besonders einfach ist natürlich der Fall einer einzigen Raumdimension, d.h. der Fall der schwingenden Saite. Bevor wir uns diesem Fall (mit der angegebenen Randbedingung) zuwenden, skizzieren wir kurz den allgemeinen

Hintergrund

Die Theorie der Wellenbewegung ist eine unendliche Geschichte für Mathematiker wie für Physiker. Sie ist auch eine Herausforderung an die Lehre. In Mathematikbüchern wird man schwerlich eine umfassende Definition von Wellenbewegung finden; und in Physikbüchern baut man ohnehin lieber auf konkrete Assoziationen als auf allgemeine Definitionen. Im berühmten Lehrbuch „Physik“ von Gerthsen findet sich (auf Seite 102) die folgende (nach meiner Meinung sehr gelungene) „Definition und Beschreibung“

*„Wenn im Innern eines deformierbaren Mediums eine Verschiebung aus der Ruhelage (Deformation) bewirkt oder „erregt“ wird, so bleibt diese nicht auf das Erregungszentrum beschränkt, sondern sie teilt sich den Nachbargebieten mit, die (zeitlich verzögert) ebenfalls deformiert werden. Eine Erregung pflanzt sich nach allen Richtungen mit einer charakteristischen **Ausbreitungsgeschwindigkeit** fort. Wir nennen diesen zeitlich **und** räumlich veränderlichen Zustand eine **Welle**“ ...*

*„Flächen im Medium, deren Punkte mit gleicher Phase schwingen, bezeichnen wir als **Wellenflächen**. Sie umschließen das Erregungszentrum. Ist dieses punktförmig, und ist die Ausbreitungsgeschwindigkeit unabhängig von der Richtung und überall konstant, dann sind die Wellenflächen Kugelflächen (**Kugelwellen**). Liegt das Erregungszentrum im Unendlichen oder mindestens sehr weit entfernt, oder geht die Welle von einer überall mit gleicher Phase schwingenden Ebene aus, dann sind die Wellenflächen Ebenen (**ebene Wellen**),*

*Linien, die vom Erregungszentrum ausgehend, die Wellenflächen überall senkrecht durchsetzen, bezeichnen wir als **Strahlen**.*

*Wenn die Deformation im Erregungszentrum eine harmonische Schwingung der Teilchen um ihre Ruhelage mit der Schwingungsdauer $T = \frac{1}{\nu}$ ist, so setzt sich diese Schwingung durch den ganzen Körper hindurch fort. Benachbarte Teilchen schwingen in der Phase gegeneinander versetzt. In regelmäßigen Abständen folgen aber Teilchen, die in der Schwingungsphase miteinander übereinstimmen. Wir nennen diesen Abstand innerhalb einer Welle die **Wellenlänge***

Während das Erregungszentrum eine volle Schwingung vollführt hat, ist die Erregung bis zu einem Punkt vorgedrungen, in dem die Schwingung nun mit der des Zentrums gleichphasig ist. Sein Abstand ist nach obiger Definition die Wellenlänge. Er ist gleich der Geschwindigkeit, genauer der Phasengeschwindigkeit c der Wellenausbreitung mul-

tipliziert mit der Schwingungsdauer T .

$$c = \frac{\lambda}{T} = v \cdot \lambda$$

(c ist nicht die Geschwindigkeit eines Körpers, sondern die eines Zustandes.)

Die Gleichung stellt eine Beziehung zwischen den oben definierten Größen c, λ und T dar; sie sagt aber nichts aus über die Abhängigkeit der Phasengeschwindigkeit von der Frequenz“.

Der letzte Satz aus Gerthsens Buch macht deutlich, dass die hier entwickelten Intuitionen wesentlich über die Theorie der Wellengleichung hinausgehen. In der Wellengleichung gibt es nämlich nur eine einzige Konstante c , während bei allgemeineren Wellenphänomenen die „Dispersion“ eine fundamentale Rolle spielt. Die Dispersionsrelation beschreibt, wie die Kreisfrequenz ω von der Wellenzahl (k_1, k_2, k_3) abhängt. Man könnte auch sagen, dass die Dispersionsrelation angibt, wie die Phasengeschwindigkeit von der Wellenzahl abhängt.

$$c = \frac{\Omega(k_1, k_2, k_3)}{\|k\|} = \omega \cdot (2\pi\lambda) \quad .$$

Bemerke : Der Betrag der Wellenzahl hat die Dimension einer reziproken Länge, $\|k\| = \frac{1}{2\pi\lambda}$. In isotropen Medien hängt die Phasengeschwindigkeit nur von der Wellenlänge ab; bei der klassischen Wellengleichung ist sie von der Wellenzahl (k_1, k_2, k_3) gänzlich unabhängig.

Elektromagnetische Wellen breiten sich im materiefreien Raum mit Lichtgeschwindigkeit aus. In einem Medium mit dem Brechungsindex n ist die Phasengeschwindigkeit aber nicht c , sondern $\frac{c}{n}$. Der Brechungsindex ist abhängig von der Frequenz.

Man lese Feynman, *Vorlesungen über Physik*, Band 1, Kap. 31, „Der Ursprung des Brechungsindex“.

Der Begriff der Phasengeschwindigkeit ist am Bild der ebenen Wellen entwickelt. Eine mathematische Begründung dafür, dass er für alle Wellenbewegungen fundamental ist, ergibt sich aus der Tatsache, dass man sehr allgemeine Funktionen $u(t, x, y, z)$ aus Funktionen der Form

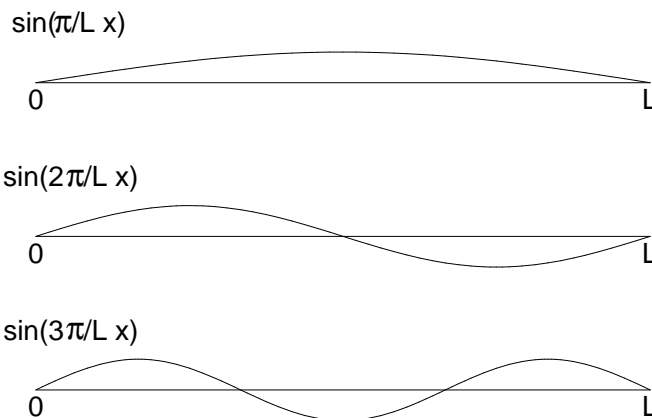
$$\exp \left(i(\omega t - (k_1 x + k_2 y + k_3 z)) \right)$$

im Sinne einer Integration linear kombinieren kann.

Dieses mathematische Phänomen ist der Gegenstand der Theorie der Fourier-Integrale in höheren Dimensionen.

Eigenschwingungen der Saite

- 1) Für die in $x = 0$ und $x = L$ eingespannte Saite gibt es eine „Grundschiwingung“ und eine Folge von Oberschwingungen. Die Wellenzahlen, die hier in Frage kommen, sind die ganzzahligen Vielfachen von $\frac{\pi}{L}$. Die Auslenkungen haben die Form



- 2) Für $n \in \mathbb{N}$ sei $k = \pm \frac{n\pi}{L}$ und $\omega = \pm k \cdot c$. Die Funktion

$$u(t, x) := \sin kx \cdot e^{i\omega t}$$

erfüllt die Gleichungen

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} u(t, x) = -k^2 \cdot u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x) \quad .$$

- 3) Das Produkt $u(t, x)$ kann man auch als Summe schreiben

$$i \cdot u(t, x) = \frac{1}{2} \exp(i(\omega t + kx)) - \frac{1}{2} \exp(i(\omega t - kx)) \quad .$$

Die Summanden sind eindimensionale ebene Wellen; man nennt sie laufende Wellen mit den Wellenzahlen $\pm k$ und der Kreisfrequenz ω . Die Phasengeschwindigkeit ist $c = \frac{\omega}{|k|}$.

$u(t, x)$ kann man sich also so zustandekommen denken, dass sich zwei gegeneinander laufende ebene Wellen überlagern, wobei sich die Beiträge in den Positionen $x = 0, \pm L, \pm 2L, \dots$ zu allen Zeiten auslöschen.

- 4) Im \mathbb{C} -Vektorraum der komplexen Lösungen läßt sich bequemer rechnen als im \mathbb{R} -Vektorraum der reellen Lösungen. Die beobachtete Auslenkung kann als der Realteil einer komplexen Lösung verstanden werden. Sie hat die Form

$$v(t, x) = |a| \cdot \sin(kx) \cdot \cos(\omega(t - t_0)) \quad .$$

Die maximale Amplitude $|a| \cdot \sin(kx)$ wird zu den Zeitpunkten $t_0 + \frac{2\pi}{\omega} m = t_0 + m \cdot T$ ($m \in \mathbb{Z}$) angenommen; $T = \frac{2\pi}{\omega}$ ist die Schwingungsdauer $\nu := \frac{1}{T} = \frac{\omega}{2\pi}$ ist die Frequenz.

Die Geschwindigkeit, mit welcher die Saite durch die Nulllage geht, ist $\pm \omega \cdot |a| \cdot \sin(kx)$. Die kinetische Energie ergibt sich durch Integration des Quadrats.

$$\text{kin.Energie} = \left(\frac{1}{4} M \cdot L \right) \cdot |a|^2 \cdot |\omega|^2 \quad (M = \text{Gesamtmasse})$$

Zu den Zeitpunkten maximaler Auslenkung ist diese Energie in potentielle Energie (Deformationsenergie) umgewandelt.

- 5) **Die Überlagerung von Oberschwingungen** Die Gesamtheit der Lösungen der Gleichung

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) u(t, x) = 0$$

$$u(t, 0) = 0 = u(t, L)$$

ist ein Vektorraum. Jede Linearkombination der oben studierten Lösungen ist eine Lösung. Das gilt nicht nur für endliche Linearkombinationen, sondern auch für manche unendliche Überlagerungen

$$u(t, x) = \sum_{-\infty}^{+\infty} (i c_n) \sin\left(\frac{n\pi x}{L}\right) \cdot \exp\left(\frac{i n \pi c t}{L}\right) \quad .$$

Es handelt sich um eine reelle Lösung, wenn

$$c_{-n} = \bar{c}_n \quad \text{für alle } n \quad .$$

Mit der obigen Rechnung läßt sich das umschreiben

$$u(t, x) = \frac{1}{2} \sum c_n \exp\left(i \frac{n\pi}{L} (x + ct)\right) - \frac{1}{2} \sum c_n \cdot \exp\left(i \frac{n\pi}{L} (-x + ct)\right)$$

$$= \frac{1}{2} g(x + ct) - \frac{1}{2} g(-x + ct) \quad .$$

Bemerke :

- a) Die Funktion

$$g(y) = \sum c_n \exp\left(i \frac{n\pi}{L} y\right)$$

ist genau dann reell, wenn $c_{-n} = \bar{c}_n$ für alle n .

- b) Zur Zeit 0 sind Auslenkung und Geschwindigkeit

$$u(0, x) = \frac{1}{2} g(x) - \frac{1}{2} g(-x)$$

$$\frac{1}{c} \frac{\partial}{\partial t} u(0, x) = \frac{1}{2} g'(x) - \frac{1}{2} g'(-x) \quad .$$

Wenn die Funktion $g(\cdot)$ ungerade ist, dann verschwindet die Anfangsgeschwindigkeit.

In diesem Falle hat man $u(t, x) = \frac{1}{2} g(x + ct) + \frac{1}{2} g(x - ct)$.

6) Es stellen sich nun zwei Fragen :

I. Wie allgemein ist die durch Überlagerung gewonnene Lösung der Gleichung

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) u(t, x) = 0$$

$$u(t, 0) = 0 = u(t, L) \quad \text{für alle } t \quad .$$

II. Welche Koeffizientenfolgen $(c_n)_n$ kommen in Betracht? Unter welchen Umständen liefert die unendliche Summe tatsächlich eine Lösung?

Diese Fragen haben eine lange und interessante Geschichte.

Historisches zur schwingenden Saite

Der berühmte Mathematiker und Enzyklopädist J.B.d'Alembert hat 1747 das folgende Ergebnis publiziert: Wenn man einer elastischen Saite der Länge L , die in den Punkten $x = 0$ und $x = L$ eingespannt ist, in den Punkten x die Auslenkung $g(x)$ gibt und dann losläßt, so wird ihre Auslenkung zu allen Zeiten t gegeben durch

$$u(t, x) = \frac{1}{2}g(x + ct) + \frac{1}{2}g(x - ct) \quad .$$

D'Alembert dachte vermutlich an Funktionen $g(\cdot)$, die durch einen analytischen Ausdruck gegeben sind, welcher die erforderlichen Differentiationen ermöglicht. L.Euler propagierte zu dieser Zeit schon allgemeinere Auffassungen, was man unter einer Funktion im Sinne der Analysis zu verstehen hat und welche Funktionen $g(\cdot)$ daher in Betracht gezogen werden können. In seinem Lehrbuch von 1755 gibt Euler die folgende Definition:

„Sind nun Größen auf die Art voneinander abhängig, dass keine davon eine Veränderung erfahren kann, ohne zugleich eine Veränderung der anderen zu bewirken, so nennt man diejenige, deren Veränderung man als die Wirkung von der Veränderung der anderen betrachtet, eine Funktion von dieser; eine Benennung, die sich so weit erstreckt, dass sie alle Arten, wie eine Größe durch eine andere bestimmt werden kann, unter sich begreift.“

Wenn $g(\cdot)$ eine (in Eulers Sinn beliebige) Funktion ist, die zweimal differenzierbar ist, dann genügt

$$u(t, x) = \frac{1}{2}g(x + ct) + \frac{1}{2}g(x - ct)$$

offenbar der Gleichung $\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) u(t, x) = 0$.

D. Bernoulli ging 1753 bei seiner Lösung des Problems der schwingenden Saite von der physikalischen These aus, dass jeder Ton durch Überlagerung von Grund- und Obertönen entsteht. Bernoulli schloß daraus, dass die Auslenkung immer in der Form

$$u(t, x) = \sum_{n=1}^{\infty} c_n \cdot \sin \frac{n\pi x}{L} \cdot \cos \frac{n\pi ct}{L}$$

dargestellt werden kann. Bernoulli meinte, dass man jede mögliche Anfangsauslenkung durch eine Reihe darstellen kann.

Euler, d'Alembert und später auch Lagrange verwarfen die Behauptung Bernoullis. Die

völlige Aufklärung dieser Frage mußte bis 1824 vertagt werden, als Fourier die Zweifel an der Gültigkeit der Darstellung einer „beliebigen“ Funktion durch trigonometrische Reihen beseitigte.

Wir können die Frage hier noch nicht sofort klären. Wir müssen vorher einiges über konvergente Folgen und konvergente Reihen lernen.

Bemerkung : Die Absolutquadrate der Amplituden c_n geben Auskunft, wie sich die Gesamtenergie auf die Obertöne verteilt.

$$E_n = \text{const} \cdot n^2 \cdot |c_n|^2 \quad .$$

Da die Gesamtenergie endlich ist, ist

$$g'(y) = \frac{\pi}{L} \sum c_n(i n) \exp \left(i \frac{n\pi}{L} y \right)$$

durch eine quadratsummierbare Reihe gegeben.

Die Theorie läßt also Auslenkungen $u(t, x)$ zu, die nicht stetig differenzierbar sind.

IV.6 Diskrete Fourier-Transformation

Sei $\{t_j : j \in J\}$ eine Familie von Punkten auf der reellen Achse, $|J| = n$. Wenn wir eine komplexwertige Funktion $f(t)$ in den Punkten t_j auswerten, dann erhalten wir Zahlen $f(t_j)$, die wir die Abtastwerte der Funktion nennen. Wir schreiben sie in eine J -Spalte f . Wir interessieren uns im Folgenden für 2π -periodische Funktionen und ihre Abtast-Werte in den äquidistanten Punkten $t_j = \frac{2\pi}{n} \cdot j$, $j = 1 \dots, n$. Den Vektorraum $\mathcal{H} = \mathcal{H}^J$ aller J -Spalten v machen wir zu einem Hilbertraum durch das innere Produkt $\langle w|v \rangle = \frac{1}{n} \sum \bar{w}^j \cdot v^j$. Bemerkte: Wenn n gross ist, dann gilt für glatte Funktionen $f(t)$, $g(t)$ und die ihnen zugeordneten J -Spalten f, g

$$\langle g|f \rangle = \frac{1}{n} \sum g(\bar{t}_j) f(t_j) \approx \frac{1}{2\pi} \int_0^{2\pi} g(\bar{t}) f(t) dt.$$

Die J -Spalte e_k zur Funktion $e_k(t) = \exp(ikt)$ nennen wir die J -Spalte zur Wellenzahl k (für ganzzahliges k). Offenbar gilt $e_k = e_l$, wenn $k-l$ durch n teilbar ist. Wir interessieren uns nur für die e_k mit $k \in K = \{1, \dots, n\}$; diese fassen wir zur $J \times K$ -Matrix E zusammen, und wir setzen $U = \frac{1}{\sqrt{n}} \cdot E$.

Lemma Die Familie $\{e_k : k \in K\}$ ist eine Orthonormalbasis des Hilbertraums \mathcal{H} . Die Matrix U ist eine unitäre $J \times K$ -Matrix.

Beweis Alle Einträge der Spalte e_k haben den Betrag 1; somit gilt $\|e_k\| = 1$. Wenn k nicht durch n teilbar ist, dann sind die Einträge von e_k die Potenzen einer Einheitswurzel $\zeta \neq 1$; die Summe der Einträge ist daher $= 0$.

$$\zeta = \exp(i \frac{2\pi}{n} \cdot k); \quad 1 + \zeta + \zeta^2 + \dots + \zeta^{n-1} = \frac{1}{1 - \zeta} \cdot (1 - \zeta^n) = 0.$$

Wenn $k-l$ nicht durch n teilbar ist, dann gilt

$$n \cdot \langle e_k | e_l \rangle = \sum \bar{e}_k^j \cdot e_l^j = \sum_j \exp(i \frac{2\pi}{n} (l-k) \cdot j) = 0.$$

Daraus ergibt sich $E^* \cdot E = n \cdot I_{K \times K}$. U ist unitär.

Beispiel, $n = 4$: Die folgende Matrix U ist unitär

$$U = \frac{1}{\sqrt{4}} \cdot \begin{pmatrix} +i & -1 & -i & +1 \\ -1 & +1 & -1 & +1 \\ -i & -1 & +i & +1 \\ +1 & +1 & +1 & +1 \end{pmatrix}; \quad u_k^j = \frac{1}{\sqrt{4}} \cdot \exp(i \frac{\pi}{2} \cdot k \cdot j) \quad j, k = 1, \dots, 4.$$

Wir können die unitäre Matrix U als Matrix eines **Basiswechsels** deuten:
 Sei g_j die J -Spalte mit dem Eintrag \sqrt{n} in der j -ten Position und Nullen sonst.
 Die Familie $\{g_j : j \in J\}$ ist dann eine Orthonormalbasis. Es gilt

$$e_k = \sum_j g_j \cdot \frac{1}{\sqrt{n}} \cdot e_k^j \quad \text{für alle } k. \quad (1)$$

$$(e_1, \dots, e_n) = (g_1, \dots, g_n) \cdot \frac{1}{\sqrt{n}} \cdot E = (g_1, \dots, g_n) \cdot U. \quad (2)$$

Die Umkehrung ist durch die $K \times J$ -Matrix $U^{-1} = U^*$ gegeben.
 Diese hat die Einträge $\bar{u}_j^k = \frac{1}{\sqrt{n}} \cdot \exp(-i \frac{2\pi}{n} \cdot k \cdot j)$. Es gilt

$$g_j = \sum_k e_k \cdot \frac{1}{\sqrt{n}} \cdot \exp(-i \frac{2\pi}{n} \cdot k \cdot j) \quad \text{für alle } j.$$

Satz Für eine J -Spalte v mit den Einträgen v^j haben wir

$$v = \sum_j g_j \cdot \frac{1}{\sqrt{n}} v^j = \sum_k e_k \cdot y^k \quad \text{mit} \quad y^k = \frac{1}{n} \cdot \sum_j \exp(-i \frac{2\pi}{n} \cdot k \cdot j) \cdot v^j.$$

Für eine 2π -periodische Funktion $f(t)$ haben wir in allen Abtaststellen $t_j = \frac{2\pi}{n}j$

$$f(\cdot) = \sum_k \exp(ik\cdot) \cdot y^k \quad \text{mit} \quad y^k = \frac{1}{n} \cdot \sum_j \exp(-i \frac{2\pi}{n} \cdot k \cdot j) \cdot f(\frac{2\pi}{n}j).$$

Der Beweis ergibt sich, wie üblich, indem man in die erste Summe die Formel für die g_j einbringt. Bemerge: Für grosse n und glatte $f(\cdot)$ kann man die Summe für den k -ten Koeffizienten durch ein Integral approximieren

$$y^k \approx Y^k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-ikt) \cdot f(t) dt.$$

Sprechweisen

- Der Übergang von den Abtastwerten $\{f(t_j) : j \in J\}$; $t_j = \frac{2\pi}{n}j$ zum K -Tupel der Koeffizienten $\{Y^k : k \in K\}$ heisst die diskrete Fourier-Transformation zur Maschenweite $\frac{2\pi}{n}$.
- Der Übergang von der 2π -periodischen Funktion $f(t)$ zur Koeffizientenfolge $\{Y^k : k \in \mathbb{Z}\}$ heisst die Fourier-Transformation von 2π -periodischen Funktionen. Y^k heisst der k -te Fourierkoeffizient von $f(\cdot)$.
- Der Übergang von der (quadrat)-integrablen Funktion $f(t)$ auf \mathbb{R} zur Funktion $Y(\omega) = \frac{1}{2\pi} \cdot \int \exp(-i\omega t) f(t) dt$ heisst die Fourier-Transformation auf \mathbb{R} .

Über die ‘Definitionsbereiche’ der verschiedenen Typen von Fourier-Transformationen wird noch ausführlich zu sprechen sein. Nur für die diskrete Fourier-Transformation funktioniert alles ohne Konvergenzargumente und höhere Integrationstheorie.

Wir halten fest: Die Koeffizienten y^k ergeben sich als Linearkombinationen der Abtastwerte $f(t_j)$; und das trigonometrische Polynom $\sum_1^n y^k \cdot e^{ikt}$ hat an den Abtastpunkten die vorgegebenen Werte $f(t_j)$.

Die Theorie der **Fourier-Reihen** befasst sich mit dem Hilbertraum der quadratisch integrierbaren 2π -periodischen Funktionen. Es geht um die Funktionen mit einer Darstellung

$$f(t) = \sum_{-\infty}^{\infty} c_k \cdot e^{ikt} \quad \text{wobei} \quad \sum |c_k|^2 < \infty.$$

Wichtige Operatoren auf diesem Hilbertraum $\mathcal{L}^2(\mathbb{R}/2\pi)$ sind die Shift-Operatoren B_ε und die Differenzen-Operatoren D_ε

$$(B_\varepsilon f)(t) = f(t - \varepsilon); \quad D_\varepsilon = \frac{1}{\varepsilon} \cdot (I - B_\varepsilon) \quad \text{also} \quad (D_\varepsilon f)(t) = \frac{1}{\varepsilon} [f(t) - f(t - \varepsilon)].$$

Unser Hilbertraum \mathcal{H}^J ist eine endlichdimensionale Karikatur des Raums $\mathcal{L}^2(\mathbb{R}/2\pi)$. Auch auf ihm gibt es Shift-Operatoren und Differenzenoperatoren, allerdings nur zu ε , die ein ganzzahliges Vielfaches von $\frac{2\pi}{n}$ sind; und für diese gilt $B_\varepsilon^n = I$. Der Shift um $\frac{2\pi}{n}$ ist die lineare Abbildung S , welche die Einträge zyklisch permutiert: Der j -te Eintrag von $S\mathbf{w}$ ist der $(j-1)$ -te Eintrag von \mathbf{w} . $(S\mathbf{w})^j = w^{j-1}$. Es handelt sich um eine Isometrie; sie wird in jeder ONBasis durch eine unitäre Matrix beschrieben. In unserem Fall können wir leicht eine ONBasis angeben, in welcher die darstellende Matrix eine Diagonalmatrix ist. Die J -Spalte zur Wellenzahl k , also \mathbf{e}_k mit den Einträgen $e_k^j = \zeta^j$ (wobei $\zeta = \exp(i\frac{2\pi}{n}k)$) ist normierter Eigenvektor zum Eigenwert ζ^{-1} . Die \mathbf{e}_k , entstanden aus $\mathbf{e}_k(t) = e^{ikt}$, sind auch die Eigenvektoren des normalen Operators D_ε ; die Eigenwerte sind hier $\lambda_k = \frac{n}{2\pi} \cdot (1 - e^{-i\frac{2\pi}{n}k})$. Für grosses n und $k = 0, \pm 1, \pm 2, \dots$ gilt $\lambda_k \approx ik$. Der Operator $\frac{1}{i}D_\varepsilon$ auf dem unendlichdimensionalen Hilbertraum $\mathcal{L}^2(\mathbb{R}/2\pi)$ approximiert (in einem gewissen Sinn) den selbstadjungierten Operator $T = \frac{1}{i}\frac{\partial}{\partial t}$ mit den Eigenwerten $k = 0, \pm 1, \pm 2, \dots$ und den dazugehörigen Eigenvektoren \mathbf{e}^{ikt} .— Der Differentialoperator T ist ein unbeschränkter selbstadjungierter Operator. Eine mathematisch solide Theorie dieser Art von Operatoren wird (nach umfangreichen Vorbereitungen) in der Funktionalanalysis entwickelt.

Schnelle Fourier-Transformation, (FFT = Fast Fourier Transformation)

In der Datenanalyse kennt man das Problem, für sehr viele und sehr lange (numerisch gegebene) Folgen $\mathbf{x} = \{x^j : j = 0, 1, \dots, n-1\}$ die diskrete Fourier-Transformation $\mathbf{y} = \{y^k : k = 0, 1, \dots, n-1\}$ numerisch zu berechnen.

$$\mathbf{y} = \frac{1}{\sqrt{n}} \mathbf{E}^* \cdot \mathbf{x}, \quad y^k = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} x^j \cdot e^{-i\frac{2\pi}{n}jk}$$

Wenn man eine n -Spalte \mathbf{x} mit einer $n \times n$ -Matrix \mathbf{A} multipliziert, dann braucht man (im Allgemeinen) n Multiplikationen für jeden Eintrag von $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}$; zur Berechnung

der gesamten n -Spalte y braucht man also n^2 Multiplikationen. Bei der Berechnung der diskreten Fourier-Transformierten einer n -Spalte x (oder besser: einer n -periodischen Folge) kann man sich zunutze machen, dass die unitäre Matrix $\frac{1}{\sqrt{n}}E^*$ eine sehr spezielle Struktur hat. Es erweist sich als günstig, wenn n das Produkt vieler kleiner Primzahlen ist; am günstigsten ist der Fall $n = 2^m$. Wir werden sehen, dass in diesem Fall $2n \log_2 n$ Multiplikationen ausreichen.

Wir behandeln die Produktzerlegung $n = R \cdot S$:

Da jedes $j \in \{0, 1, \dots, n-1\}$ eine eindeutige Darstellung besitzt

$$j = u \cdot R + v, \quad \text{mit } u \in \{0, \dots, S-1\}, \quad v \in \{0, \dots, R-1\},$$

können wir schreiben

$$y^k = \frac{1}{\sqrt{R}} \sum_{v=0}^{R-1} e^{-i \frac{2\pi k}{n} v} \cdot \frac{1}{\sqrt{S}} \sum_{u=0}^{S-1} x^{v+uR} e^{-i \frac{2\pi k}{S} u} = \frac{1}{\sqrt{R}} \sum_{v=0}^{R-1} e^{-i \frac{2\pi k}{n} v} \cdot z_v^{(k)}$$

Die äussere Summe erfordert R Multiplikationen für jedes k . Für jedes v ergibt sich die S -periodische Folge $\{z_v^{(0)}, z_v^{(1)}, \dots, z_v^{(S-1)}, \dots\}$ als das Ergebnis einer diskreten Fourier-Transformation.

Sei $M(S)$ die Anzahl der Multiplikationen, mit welcher man die diskrete Fourier-Transformation einer S -periodischen Folge bewältigen kann. (Die Division durch \sqrt{S} wird nicht mitgezählt.) Unsere Rechnung zeigt

$$M(S \cdot R) \leq R \cdot M(S) + R \cdot RS \quad \text{also} \quad \frac{1}{SR} M(SR) \leq \frac{1}{S} M(S) + R.$$

Speziell $n = 2^m$: Die diskrete Fourier-Transformation einer 2-periodischen Folge kann man mit 4 Multiplikationen bewältigen. Wegen $\frac{1}{2^m} M(2^m) \leq \frac{1}{2^{m-1}} M(2^{m-1}) + 2$ haben wir

$$\frac{1}{2^m} M(2^m) \leq 2m; \quad M(n) \leq 2n \log_2 n \quad \text{für } n = 2^m.$$

Diskrete Fourier-Analyse

Eine Funktion $f(t)$ wird auf dem Intervall $[-T, +T]$ in n äquidistanten Punkten t_j abgetastet. Die diskrete Fourier-Transformation liefert eine n -periodische Folge $\{y^k : k \in \mathbb{Z}/n\}$. Von diskreter Fourier-Analyse spricht man, wenn man aufgrund dieser Folge y Eigenschaften der Funktion $f(\cdot)$ zu erschliessen versucht.

Beispiel: Wir stellen uns ein System von gekoppelten Oszillatoren vor. Im ungedämpften Fall ist der Schwingungszustand bekanntlich durch eine Überlagerung der *normal modes* gegeben. Für jeden einzelnen Oszillator finden wir eine fastperiodische Schwingung $f(t) = \sum_{\alpha} c_{\alpha} \cdot e^{i\omega_{\alpha} t}$. Die diskrete Fourier-Transformation ist ein bewährtes Mittel, um einige der Kreisfrequenzen ω_{α} der *normal modes* zu entdecken.

Die niederfrequenten Schwingungen können wir allerdings nur dann entdecken, wenn T gross ist; wenn nämlich die Wellenlänge gross im Vergleich zu $2T$ ist, ($\omega \ll \pi/T$), dann

sieht die Einschränkung auf das Intervall $(-T + T)$ wie eine Konstante aus. Die hochfrequenten Schwingungen können wir nur dann entdecken, wenn die Abtastdichte gross ist. Auf dem Gitter mit der Maschenweite $2T/n$ sind nämlich die reinen Sinusschwingungen $e^{i\tilde{\omega}t}$ und $e^{i(\tilde{\omega}+n\pi/T)t}$ nicht zu unterscheiden. Wir müssen uns also darauf beschränken, nach den Beiträgen mit $|\tilde{\omega}| < \Omega = \frac{\pi}{T} \cdot \frac{n}{2}$ zu suchen. (Ω wird manchmal die Nyquist-Frequenz genannt)

Machen wir die Suche nach den Beiträgen $\exp(i\tilde{\omega}t)$ etwas konkreter:

Die Abtastpunkte sind $t_j = 2T/n \cdot j : [-n/2] \leq j < [n/2]$.

Die Kreisfrequenz mit der Wellenzahl k ist $\omega_k = \frac{\pi}{T} \cdot k : k \in \mathbb{Z}/n$.

Die Funktion $f(t)$, oder vielmehr das n -tupel $f = \{f(t_j)\}$ erhält das Normquadrat

$$\|f\|^2 = \frac{1}{n} \sum_j |f(t_j)|^2 = \frac{1}{2T} \int_{-T}^T \left| \sum_k y^k \cdot e^{i\omega_k t} \right|^2 dt = \sum_k |y^k|^2 \quad (3)$$

$$\text{mit} \quad y^k = \frac{1}{n} \cdot \sum_j \exp(-i \frac{2\pi k}{n} \cdot j) \cdot f(t_j) \quad \text{für} \quad k \in \mathbb{Z}/n. \quad (4)$$

Wir definieren nun zu $f(\cdot)$ das sog. **Periodogramm**. Das ist die 2Ω -periodische positive Treppenfunktion $p(\cdot)$ mit dem Wert $p(\omega_k) = \frac{T}{\pi} |y^k|^2$ im Intervall $(\omega_k - \frac{1}{2} \frac{\pi}{T}, \omega_k + \frac{1}{2} \frac{\pi}{T})$.

Man bemerke, dass das Integral des Periodogramms über die Periode das Normquadrat liefert:

$$\int_{-\Omega}^{\Omega} p(\omega) d\omega = \sum_k |y^k|^2 = \|f\|^2 = \frac{1}{n} \sum_j |f(t_j)|^2.$$

In physikalischer Analogie darf man sagen, dass das Periodogramm die Energie auf die Frequenzen verteilt, also sozusagen eine Energiedichte beschreibt.

Beispiele für Periodogramme:

Für die konstante Funktion $f_0(\cdot) \equiv 1$ hat das Periodogramm den Wert $\frac{T}{\pi}$ im Abstand $< \frac{1}{2} \frac{\pi}{T}$ vom Nullpunkt; sonst im Intervall $|\omega| \leq \Omega = \frac{\pi}{T} \cdot \frac{n}{2}$ hat dieses Periodogramm den Wert 0. Allgemeiner: Für jedes $\tilde{\omega} = \omega_k = \frac{\pi}{T} \cdot k$ hat das Periodogramm der Funktion $e^{i\tilde{\omega}t}$ den Wert $\frac{T}{\pi}$ im Intervall $\tilde{\omega} \pm \frac{1}{2} \frac{\pi}{T}$; sonst im Intervall $|\omega| \leq \Omega$ hat das Periodogramm $p_{\tilde{\omega}}(\cdot)$ den Wert 0.

Berechnen wir jetzt das Periodogramm $p_{\tilde{\omega}}(\cdot)$ für die $2T$ -periodische Funktion $f_{\tilde{\omega}}(\cdot)$, die im Intervall $[-T, T]$ die Werte $e^{i\tilde{\omega}t}$ hat. (Es kommt natürlich nur auf die Werte in den Abtastpunkten t_j an; und wir können uns daher auf $\tilde{\omega}$ mit $|\tilde{\omega}| \leq \Omega = \frac{\pi}{T} \cdot \frac{n}{2}$ beschränken.) Zur Darstellung der Werte $p_{\tilde{\omega}}(\omega_k)$ benützen wir den wohlbekannten n -ten Fejér-Kern $K_n(\cdot)$. Zur Erinnerung:

Lemma Die Funktion

$$K_n(\phi) = \frac{1}{2\pi n} \cdot \frac{\sin^2(\frac{n}{2}\phi)}{\sin^2(\frac{1}{2}\phi)}$$

ist nichtnegativ und 2π -periodisch mit $\int_{-\pi}^{\pi} K_n(\phi) d\phi = 1$. Es gilt

$$K_n(0) = \frac{n}{2\pi} \quad \text{und} \quad K_n(\phi) < \frac{1}{2\pi n \sin^2(\delta/2)} \quad \text{für} \quad \delta < |\phi| < \pi.$$

Berechnung Die diskrete Fouriertransformierte von $e^{i\tilde{\omega}t}$ im Intervall $(-T, +T)$ kann elementar berechnet werden

$$\begin{aligned} n \cdot y^k &= \sum_{-[n/2] \leq j < [n/2]} e^{-i\frac{2\pi}{n}kj} \cdot f_{\tilde{\omega}}\left(\frac{2T}{n} \cdot j\right) \\ &= \sum_{-[n/2] \leq j < [n/2]} e^{-i\frac{2\pi}{n}kj} \cdot e^{i\tilde{\omega}\frac{2T}{n}j} = \sum_j \zeta_{(k)}^j = \frac{1}{1 - \zeta_{(k)}} \cdot \zeta_{(k)}^{-[n/2]} \cdot (1 - \zeta_{(k)}^n) \end{aligned}$$

mit $\zeta_{(k)} = \exp\left(i\frac{2T}{n}(\tilde{\omega} - \omega_k)\right) = \exp(i\phi_k)$.

Bemerkung: Für eine Zahl $\zeta = \exp(i\phi)$ vom Betrag 1 gilt

$$|1 - \zeta|^2 = |1 - e^{i\phi}|^2 = |e^{i\phi/2} - e^{-i\phi/2}|^2 = 4 \cdot \sin^2(\phi/2).$$

Somit haben wir die Werte des Periodogramms

$$p_{\tilde{\omega}}(\omega_k) = \frac{T}{\pi} \cdot |y_k|^2 = \frac{T}{\pi n^2} \cdot \frac{\sin^2(n/2 \cdot \phi_k)}{\sin^2(1/2 \cdot \phi_k)} = \frac{2T}{n} \cdot K_n(\phi_k) = \frac{2T}{n} \cdot K_n\left(\frac{2T}{n}(\omega_k - \tilde{\omega})\right).$$

Bemerkung: Das Integral der Funktion $g_{\tilde{\omega}}(\omega) = \frac{2T}{n} \cdot K_n\left(\frac{2T}{n}(\omega - \tilde{\omega})\right)$, erstreckt über das Intervall $(-\Omega, \Omega)$ ergibt den Wert 1. Das Integral = 1 erhalten wir auch für unsere Treppenfunktion $p_{\tilde{\omega}}(\omega)$. Für $\tilde{\omega} \neq \frac{\pi}{T} \cdot k$ haben die Periodogramme der reinen Sinusschwingungen $\exp(i\tilde{\omega}t)$ jedoch Nebenmaxima, die den Praktiker stören.

Plausibilitätsbetrachtungen

1. Die diskrete Fourier Transformation ist eine lineare Abbildung: Die DFT von $(f + g)(t)$ ist $y + z$, wenn y die DFT von $f(t)$ und z die DFT von $g(t)$ ist. Das Periodogramm von $(f + g)(t)$ ist approximativ die Summe der Periodogramme, wenn die Produkte der Fourierkoeffizienten allesamt ‘klein’ sind.

$$p_{(f+g)}(\omega_k) = p_f(\omega_k) + p_g(\omega_k) + y^k \cdot z^k.$$

Wenn $f(t)$ und $g(t)$ reine Sinusschwingungen sind, deren Kreisfrequenzen nicht nahe aneinander liegen, dann ist das annähernd so; die Bereiche, wo p_f und p_g ‘gross’ sind, überlappen sich nicht. Allgemeiner kann man sagen, dass das Periodogramm von $f(t) = \sum_{\alpha} c_{\alpha} \cdot e^{i\omega_{\alpha}t}$ (wenn die Kreisfrequenzen paarweise einen substantiellen Abstand haben) als eine Funktion erscheint, welche an den Punkten ω_{α} Gipfel der Höhe $\approx |c_{\alpha}|^2$ besitzt.

Hinweis: Funktionen der Frequenz (oder der Energie) mit diesem Erscheinungsbild trifft man sehr häufig in der messenden Physik. Häufig werden sie als ‘Resonanzkurven’ verstanden. Manchmal kann man sie aber auch als Periodogramme interpretieren.

2. Die diskrete Fourier Transformation liefert zu den n Abtastwerten $\{f(t_j) : t_j = \frac{2T}{n}j\}$ eine interpolierende $2T$ -periodische Funktion \tilde{f}

$$\tilde{f}(t) = \sum_{-[\frac{n}{2}] \leq k < [\frac{n}{2}]} y^k \exp(ik\frac{\pi}{T}t) \quad \text{mit} \quad y^k = \frac{1}{n} \sum f(t_j) \exp(-ik\frac{\pi}{T}t_j)$$

Wenn es darum geht, irgendwelche fastperiodische Funktionen zu analysieren, dann erscheint es grob und unsensibel, die Abtastwerte wie die einer $2T$ -periodischen Funktion zu behandeln; die periodische Funktion \tilde{f} versucht einen drastischen Umbruch an den Intervallenden zu bewältigen, welcher mit der zu analysierenden Funktion überhaupt nichts zu tun hat. Die Praktiker haben sich Tricks der Datenaufbereitung ausgedacht, die den Abschnideffekt abmildern ('tempern' oder englisch 'tempering' heisst der Fachausdruck). Die Praktiker bestimmen das Periodogramm zu Werten $g(t_j)$, die im Innern des Intervalls mit den $f(t_j)$ übereinstimmen und am Rande für einen guten Anschluss sorgen. Damit können sie in der Tat lästige Nebenmaxima zu den 'Spitzen' in den interessanten Punkten ω_α abflachen.

3. Das Tempern ist auch in anderen Situationen beliebt. Nehmen wir an, wir sollen aus langen Realisierungen x^1, x^2, \dots, x^n einer stationären Zeitreihe X^1, X^2, \dots auf die sog. Spektraldichte schliessen. (Eine stationäre Zeitreihe ist das, was die Nachrichteningenieure ein farbiges Rauschen in diskreter Zeit nennen.) Sollten wir die Abschnitte nicht lieber auf die Länge $\tilde{n} = 2^m$ kürzen und/oder den Bruch an den Enden abmildern, indem wir beispielsweise die Werte auf die Nullfolge hin abflachen? Die FFT funktioniert dann bekanntlich besonders gut. Ein schnell zu berechnendes Periodogramm von gefälliger Gestalt wird möglicherweise den Analysanten auf interessante Phänomene aufmerksam machen. Und wenn das der Effekt des Temporns ist, dann wird man auf eine mathematische Fundierung der Tricks gern verzichten; die Willkür beim Abschneiden ist ohnehin nicht zu vermeiden.

Zur Geschichte Die diskrete Fouriertransformation hat eine lange Geschichte. Gauss hat sie benützt, um die Bahnen von Himmelskörpern zu approximieren. Gauss hat für seine Rechnungen auch schon die Schnelle Fourier Transformation benützt. Diese Technik war dann aber weitgehend vergessen, bis sie 1966 von J. Tukey und J. W. Cooley wieder aufgegriffen wurde. Heute wird die 'Cooley–Tukey FFT' in grossem Stil angewendet (u. a. bei der Analyse von seismologischen Daten und in der Bildverarbeitung.)

Gauss studierte die Matrix mit den Einträgen $\exp(i\frac{2\pi}{n}jk)$ auch aus der Sicht der Zahlentheorie; (sie entspricht der sog. Charaktertafel der zyklischen Gruppe mit n Elementen). Einen kurzen Überblick über aktuelle Weiterungen im Sinne der Reinen Mathematik geben Maslen und Rockmore im Artikel:

"The Cooley–Tukey FFT and Group Theory", NOTICES of the American Mathematical Society. Vol 48 Nr.10 (Nov 2001).

IV.7 Drehungen im Minkowskiraum

In der speziellen Relativitätstheorie hat man einen vierdimensionalen reellen Vektorraum mit einer ausgezeichneten quadratischen Form $(W, Q(\cdot))$. Die quadratische Form hat die Signatur $(1, 3)$, sie verschwindet auf dem Lichtkegel. Obwohl $Q(\cdot)$ nicht das Quadrat einer Norm im Sinne einer Metrik ist, notiert man häufig $Q(w) = \|w\|^2$ und man nennt

$$\langle w \| v \rangle = \frac{1}{2}(Q(w + v) - Q(w) - Q(v)) \quad \text{das innere Produkt.}$$

Der Raum $(W, Q(\cdot))$ heisst der **Minkowskiraum**. Er besitzt eine (durch die Physik gegebene) Orientierung; die Menge der Vektoren mit $Q(w) > 0$ zerfällt in den ‘Vorwärtskegel’ und den ‘Rückwärtskegel’.

Die Gruppe derjenigen linearen Abbildungen, welche den Minkowskiraum isometrisch, orthochron und orientierungstreu auf sich abbilden, heisst **eigentliche Lorentzgruppe**. Es handelt sich um eine zusammenhängende sechsdimensionale Gruppe von linearen ‘Isometrien’ des Minkowskiraums $(W, Q(\cdot))$.

Wenn man in W eine Basis einführt und die Koeffizienten durch 4-Spalten darstellt, dann wird aus der eigentlichen Lorentzgruppe eine Gruppe von 4×4 -Matrizen. Es empfiehlt sich, als Basis eine positiv orientierte Orthonormalbasis zu wählen, d.h. ein Quartupel von Vektoren $\{w_0, w_1, w_2, w_3\}$, sodass

$$\langle w_j \| w_k \rangle = \begin{cases} 1 & \text{wenn } j = k = 0 \\ -1 & \text{wenn } j = k \neq 0 \\ 0 & \text{wenn } j \neq k, \end{cases}$$

wo w_0 in den Vorwärtskegel deutet. (Man spricht von einem **Inertialsystem**.)

Der algebraische Zugang zur Lorentzgruppe

Es bezeichne D die Diagonalmatrix mit den Diagonalelementen $+1, -1, -1, -1$. Mit $O((1, 3), \mathbb{R})$ bezeichnet man die Gruppe \tilde{D} aller 4×4 -Matrizen L mit

$$L^T \cdot D \cdot L = D.$$

Diese Gruppe ist eine 6-dimensionale Mannigfaltigkeit, welche in vier Zusammenhangskomponenten zerfällt.

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

Die Zusammenhangskomponente des Einselements \mathcal{D}_+^\uparrow besteht aus den Matrizen, welche die eigentlichen Lorentztransformationen bzgl. eines Inertialsystems darstellen. Sie haben die Determinante $+1$; und sie haben einen positiven Eintrag in der linken oberen Ecke. Die Elemente der übrigen Zusammenhangskomponenten gewinnt man, indem man diese Matrizen mit $-I$ oder D oder $-D$ multipliziert.

Ein alternatives Modell für den Minkowskiraum Es ist in keiner Weise zwingend, die Vektoren w des Minkowskiraums durch 4-Spalten darzustellen. Wir wollen im Folgenden die Vektoren w lieber durch die hermiteschen 2×2 -Matrizen repräsentieren

$$w = \alpha_0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \alpha_3 & \alpha_1 - i\alpha_2 \\ \alpha_1 + i\alpha_2 & -\alpha_3 \end{pmatrix} \quad \text{mit reellen } \alpha_0, \alpha_1, \alpha_2, \alpha_3.$$

Eine bequeme Basis liefern die Pauli-Matrizen

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

zusammen mit der Einheitsmatrix, die in diesem Zusammenhang auch als σ_0 notiert wird.

Bemerkenswert ist die Formel für die Determinante und die für die Spur des Quadrats

$$\det(w) = \alpha_0^2 - (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) = \frac{1}{2}(\text{trace } w)^2 - \frac{1}{2}\text{trace}(w^2).$$

Es handelt sich um eine quadratische Form, welche den vierdimensionalen Raum \mathbb{M} aller hermiteschen 2×2 -Matrizen $w = \alpha_0\sigma_0 + \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3$ in der Tat zu einem Modell $(\mathbb{M}, \det(\cdot))$ des Minkowskiraums macht.

Polarisierung ergibt für $v = \begin{pmatrix} \beta_0 + \beta_3 & \beta_1 - i\beta_2 \\ \beta_1 + i\beta_2 & \beta_0 - \beta_3 \end{pmatrix}$

$$\langle w \| v \rangle = \alpha_0\beta_0 - (\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3) = \frac{1}{2}(\text{trace } w)(\text{trace } v) - \frac{1}{2}\text{trace}(w \cdot v).$$

Hinweis Wir werden sehen, dass es Rechenbequemlichkeiten mit sich bringt, dass wir die Koeffizienten $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ nicht als Einträge in einer 4-Spalte notieren, sondern eben als die Real- oder Imaginärteile von Einträgen in der dazugehörigen hermiteschen 2×2 -Matrix. Der Matrizenkalkül erlaubt uns Darstellungen der (orthochronen orientierungserhaltenden) Isometrien des Raums $(\mathbb{M}, \det(\cdot))$, die in mancherlei Hinsicht intuitiver sind als die Darstellungen durch 4×4 -Matrizen. Wir werden z. B. aus der alternativen Darstellung leicht ersehen, dass die eigentliche Lorentzgruppe isomorph ist zur Gruppe der Möbiustransformationen. Für die Gruppe der Möbiustransformationen wie für die eigentliche Lorentzgruppe liefert die Spezielle Lineare Gruppe $SL(2, \mathbb{C})$ eine leicht zu erfassende zweifache Überlagerung. Die topologisch sehr einfache Gruppe $SL(2, \mathbb{C})$ ist in der Tat die sog. universelle Überlagerung.

Bevor wir uns mit dem vierdimensionalen Minkowski-Raum beschäftigen, studieren wir ein Modell des orientierten dreidimensionalen euklidischen Raum. Die Vektoren repräsentieren wir durch die spurlosen hermiteschen 2×2 -Matrizen σ . Der dreidimensionale reelle Vektorraum \mathbb{P} der spurlosen hermiteschen 2×2 -Matrizen wird zu einem Modell des **orientierten euklidischen Raums** gemacht, indem wir festlegen, dass die Pauli-Matrizen $\{\sigma_1, \sigma_2, \sigma_3\}$ eine positiv orientierte Orthonormalbasis bilden.

$$\sigma = \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3 = \begin{pmatrix} \alpha_3 & \alpha_1 - i\alpha_2 \\ \alpha_1 + i\alpha_2 & -\alpha_3 \end{pmatrix}.$$

Der folgende Satz zeigt, wie man die Norm $\|\sigma\|$ und das innere Produkt $\langle\sigma|\tau\rangle$ in \mathbb{P} durch Spur- und Determinantenbildung gewinnen kann. (Zur Bezeichnung: Das Symbol \mathbb{P} wird in weiten Bereichen der Mathematik für ‘projektive Räume’ verwendet. Damit haben wir an dieser Stelle nichts zu tun; das Symbol soll hier nur an die Pauli-Matrizen erinnern.)

Satz (Das innere Produkt in \mathbb{P})

1. Das innere Produkt im dreidimensionalen euklidischen Raum \mathbb{P} ergibt sich mit Hilfe des Spuoperators

$$\|\sigma\|^2 = \frac{1}{2}\text{trace}(\sigma^2) ; \quad \langle\sigma|\tau\rangle = \frac{1}{4}\text{trace}(\sigma\tau + \tau\sigma).$$

2. Die Vektoren σ, τ sind genau dann zueinander orthogonal, wenn die entsprechenden Matrizen antikommutieren

$$\sigma \perp \tau \iff \sigma\tau + \tau\sigma = 0 .$$

Zum Beweis benötigen wir den

Hilfssatz:

- Für die Pauli-Matrizen gelten die Multiplikationsregeln

$$\begin{aligned} \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 = I \quad (= \sigma_0) \\ \sigma_1\sigma_2 &= i \cdot \sigma_3 = -\sigma_2\sigma_1; \quad \sigma_2\sigma_3 = i \cdot \sigma_1 = -\sigma_3\sigma_2; \quad \sigma_3\sigma_1 = i \cdot \sigma_2 = -\sigma_1 \cdot \sigma_3 . \end{aligned}$$

- Für jede spurlose hermitesche Matrix $\sigma \in \mathbb{P}$ gilt

$$\sigma^2 = (\alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3)^2 = (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) \cdot \sigma_0 = \|\sigma\|^2 \cdot I .$$

Beweis des Satzes

1. Die erste Formel ergibt sich trivialerweise aus dem Hilfssatz. Wir bemerken, dass die Norm von σ in \mathbb{P} bis auf den Faktor $\sqrt{2}$ die Hilbert-Schmidt-Norm ist..

$$\|\sigma\|_{\text{HS}}^2 = \text{trace}(\sigma\sigma^*) = \text{trace}(\sigma^2) = \|\sigma\|^2 \cdot \text{trace}(I) = 2 \cdot \|\sigma\|^2 .$$

2. Die Polarisierungsgleichung ergibt für $\langle\sigma|\tau\rangle$

$$\frac{1}{2}(\|\sigma+\tau\|^2 - \|\sigma\|^2 - \|\tau\|^2) = \frac{1}{4}(\text{trace}(\sigma+\tau)^2 - \text{trace}(\sigma^2) - \text{trace}(\tau^2)) = \frac{1}{4}\text{trace}(\sigma\tau + \tau\sigma)$$

3. Das Quadrat jeder spurlosen 2×2 -Matrix ist ein Vielfaches der Einheitsmatrix. Die Matrix $(\sigma + \tau)^2 - \sigma^2 - \tau^2$ verschwindet genau dann, wenn ihre Spur verschwindet.
4. Man bemerke, dass man die Norm im euklidischen Raum \mathbb{P} auch aus der Determinante gewinnen kann:

$$\|\sigma\|^2 = -\det(\sigma) .$$

Drehungen des 3-dimensionalen Anschauungsraums

Zu einer Drehung des Anschauungsraums $\neq \text{id}$ gehört nach landläufigem Verständnis eine Drehachse und ein Drehwinkel; wenn man verlangt, dass der Drehwinkel zwischen 0° und 180° ist, dann sind Drehachse und Drehwinkel eindeutig bestimmt.

In den mathematischen Vorlesungen beginnt man meistens an anderer Stelle. Man sagt: Wenn man einen euklidischen dreidimensionalen Raum mit einer Orthonormalbasis ausstattet und die Vektoren durch 3-Spalten darstellt, dann sind die Drehungen durch die orthogonalen Matrizen mit der Determinante 1 gegeben. $M^T M = I$; $\det M = 1$. Die Gruppe aller Drehungen ist somit mit der Matrixgruppe $SO(3, \mathbb{R})$ zu identifizieren.

Die geometrische Struktur von $SO(3, \mathbb{R})$ ist nicht so einfach zu begreifen, wie man vielleicht denken mag.

Feynman schreibt in seinen berühmten elementaren Vorlesungen

It is rather strange, because we live in three dimensions, but it is hard for us to appreciate what happens if we turn this way and then that way. Perhaps if we were fish or birds and had a real appreciation of what happens when we turn somersaults in space, we could more easily appreciate such things.

Wir werden sehen, dass die Gruppe der Drehungen $SO(3, \mathbb{R})$ eng verwandt ist mit der Gruppe $SU(2)$; es gibt aber einen Unterschied, und wer den nicht beachtet, wird gewisse Phänomene der Quantenmechanik (wie z. B. den Spin) nicht richtig verstehen. Wir werden auch sehen: Als Mannigfaltigkeit betrachtet, hat $SU(2)$ die übersichtliche Struktur einer dreidimensionalen Sphäre; der Raum der Drehungen hingegen hat die weniger übersichtliche Struktur des dreidimensionalen reellprojektiven Raums.

Zweizustandssysteme Die spurlosen hermiteschen 2×2 -Matrizen findet man in der Quantenmechanik als die Hamiltonoperatoren für Zweizustandssysteme. Das bekannteste Zweizustandssystem trifft man beim Studium des Spins eines Elektrons.

(Feynman: Vorlesungen über Physik Teil III, Kap. 14: Weitere Zweizustandssysteme)

Wenn man sich im Zustandsraum eines Zweizustandssystems für eine Orthonormalbasis entschieden hat, dann kann man diese ‘reinen’ Zustände ‘Spin-up’ und ‘Spin-down’ (bezgl. der gewählten z-Richtung) nennen. Die übrigen möglichen Zustände werden durch komplexe Linearkombinationen dieser ‘reinen’ Zustände ‘dargestellt, als 2-Spalten notiert. Wenn man das Elektron in ein Magnetfeld mit den Komponenten B_x , B_y , B_z bringt, dann ändert sich der Zustand gemäß der dazugehörigen einparametrischen Gruppe

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} (t) = \exp\left(-t \cdot \frac{i}{\hbar} H\right) \cdot \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} (0), \quad i\hbar \cdot \frac{d}{dt} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} (t) = H \cdot \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} (t)$$

mit $H = -\mu \cdot [B_x \cdot \sigma_1 + B_y \cdot \sigma_2 + B_z \cdot \sigma_3] = -\mu \cdot \sigma$.

Die Eigenvektoren von H repräsentieren diejenigen Spinzustände, die (bis auf die nicht-beobachtbare Phase) zeitlich unverändert bleiben; die gemischten Zustände werden im Laufe der Zeit durch unitäre Matrizen transformiert. Entscheidend ist der Spin-Operator H , welche die Veränderung in infinitesimaler Zeit beschreibt.

Betrachten wir nun (im Sinne der reinen Mathematik) die spurlosen hermiteschen 2×2 -Matrizen σ als Endomorphismen des Raums V der komplexen 2-Spalten. Es gilt der

Satz

Ein $\sigma \in \mathbb{P}$ mit $\|\sigma\| = \lambda > 0$ hat die Eigenwerte $\pm\lambda$ und liefert damit eine orthogonale Zerlegung in die Eigenräume $V = V_+^\sigma \oplus V_-^\sigma$.

Zu jeder orthogonalen Zerlegung $V = V_+ \oplus V_-$ existiert genau ein σ mit $\|\sigma\| = 1$.

Bemerke: Wenn man σ durch ein positives Vielfaches ersetzt, dann bleibt die orthogonale Zerlegung dieselbe. Wenn man das Vorzeichen ändert, dann vertauschen V_+ und V_- die Plätze. Nehmen wir einen Einheitsvektor v , welcher V_+ aufspannt (er ist bis auf einen Faktor e^{ia} vom Betrag 1 eindeutig bestimmt); das dazugehörige σ ist bis auf einen positiven(!) Faktor eindeutig bestimmt. Zu $-\sigma$ gehört der dazu orthogonale Einheitsvektor. Die in alten Zeiten gerne bemühte Analogie von Spin und einer Art ‘internes Drehmoment’ ist irreführend. (Siehe Feynman, Bd. III Kap. 14 oder auch A. I. Kostrikin und Yu. I. Manin: Linear Algebra and Geometry, Gordon and Breach Science Publishers, 1989. § 11: ‘Three-Dimensional Euclidian Space’.)

Satz (‘Achsendarstellung’ für $U \in SU(2)$)

Sei $\sigma \in \mathbb{P}$ mit $\|\sigma\| = 1$. Für jedes reelle t gilt dann

$$U(t) = \exp(it\sigma) = \cos t \cdot I + i \sin t \cdot \sigma \in SU(2).$$

Umgekehrt: Zu jedem $U \in SU(2)$, $U \neq \pm I$, gibt es genau ein $t \in (0, \pi)$ und genau ein σ mit $\|\sigma\| = 1$, sodass gilt

$$U = \cos t \cdot I + i \sin t \cdot \sigma.$$

Beweis: Wegen $\sigma^2 = -I$ haben wir

$$\exp(it\sigma) = I + it \cdot \sigma + \frac{1}{2!}(it)^2 \cdot I + \frac{1}{3!}(it)^3 \cdot \sigma + \dots = \cos t \cdot I + i \sin t \cdot \sigma.$$

Dass $U(t)$ die Determinante 1 hat, ergibt sich aus unseren allgemeinen Überlegungen über die Determinante in einer einparametrischen Matrizenengruppe.

Umgekehrt: Die speziellen unitären Matrizen sind die Matrizen der Gestalt

$$U = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \quad \text{mit} \quad |a|^2 + |b|^2 = 1.$$

Wenn $\Re a = \frac{1}{2} \text{trace} U = \cos t$, dann ist $U - \cos t \cdot I$ eine spurlose schiefhermitesche Matrix

$$U - \cos t \cdot I = \begin{pmatrix} i\alpha_3 & -\bar{b} \\ b & -i\alpha_3 \end{pmatrix}.$$

Die Determinante ist $\det(U - \cos t \cdot I) = \alpha_3^2 + |b|^2 = 1 - \cos^2 t = \sin^2 t$.

Für $U \neq \pm I$ ist $t \in (0, \pi)$ mit $\sin t > 0$ eindeutig bestimmt. Es gilt

$$\sigma = \frac{1}{i \sin t} (U - \cos t \cdot I) \in \mathbb{P} \quad \text{und} \quad \|\sigma\|^2 = -\det \sigma = 1.$$

Jede Matrix $U \neq \pm I$ in der Speziellen Unitären Gruppe $SU(2)$ besitzt also genau eine Darstellung

$$U = c \cdot I + is \cdot \sigma \quad \text{mit} \quad c \in \mathbb{R}, \quad s > 0, \quad \sigma \in \mathbb{P}, \quad \|\sigma\| = 1.$$

Man könnte sagen, dass die Gruppe mit dem infinitesimalen Generator σ dieses U zu den Zeiten $t + 2\pi k$ ($k \in \mathbb{Z}$) erreicht, wobei $\cos t = c$, $\sin t = s$.

Quaternionen

W. R. Hamilton (1805–1865) hatte bereits seine bedeutenden Beiträge zur Analytischen Mechanik, zur Optik und zur Variationsrechnung gemacht, als er 1843 die Quaternionen entdeckte. Von diesen Rechenobjekten war er so begeistert, dass er ihnen den ganzen weiteren Teil seines wissenschaftlichen Lebens widmete. Hamilton 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

$$a_0 \cdot 1 + a_1 \cdot j + a_2 \cdot k + a_3 \cdot l \quad \text{mit} \quad a_0, a_1, a_2, a_3 \in \mathbb{R}.$$

Quaternionen werden komponentenweise addiert und distributiv ausmultipliziert ausgehend von den folgenden Multiplikationsregeln:

$$\begin{aligned} j^2 = k^2 = l^2 = -1; \quad 1 \text{ ist das Einselement;} \\ j \cdot k = l = -k \cdot j, \quad k \cdot l = j = -l \cdot k, \quad l \cdot j = k = -j \cdot l. \end{aligned}$$

Man definiert zu einem Quaternion q das konjugierte Quaternion q^* und die Norm $\|q\|$:

$$\begin{aligned} q &= a_0 \cdot 1 + a_1 \cdot j + a_2 \cdot k + a_3 \cdot l \\ \longmapsto \quad q^* &= a_0 \cdot 1 - a_1 \cdot j - a_2 \cdot k - a_3 \cdot l. \\ \|q\| &= \sqrt{q \cdot q^*} = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2} \end{aligned}$$

Man kann direkt nachrechnen, dass das System der Quaternionen $(\mathbb{H}, +, \cdot, *)$ eine assoziative \mathbb{R} -Algebra ist mit $(q \cdot r)^* = r^* \cdot q^*$. Man kann sich die Arbeit aber sparen, wenn man sich überzeugt hat, dass man die Quaternionen durch komplexe 2×2 -Matrizen darstellen kann, wo die Addition und die Multiplikation die Verknüpfungen des Matrizenkalküls sind und die Konjugation durch die hermitesche Konjugation gegeben ist. Eine mögliche Übersetzung ist die Übersetzung der ‘Basisquaternionen’ in die Pauli-Matrizen:

$$1 \longleftrightarrow \sigma_0, \quad j \longleftrightarrow \frac{1}{i} \sigma_1, \quad k \longleftrightarrow \frac{1}{i} \sigma_2, \quad l \longleftrightarrow \frac{1}{i} \sigma_3.$$

Das System der Quaternionen hat eine sehr bemerkenswerte Eigenschaft: Jedes Element ausser dem Nullelement besitzt eine Inverse, nämlich $q^{-1} = \frac{1}{\|q\|^2} q^*$. Man spricht vom **Schiefkörper der Quaternionen**.

Ein Quaternion mit der Norm 1 nennt man ein **Einheitsquaternion**. Bei unserer Übersetzung entsprechen die Einheitsquaternionen den unitären Matrizen mit Determinante 1:

$$\begin{aligned} \mathbf{q} &= a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{j} + a_2 \cdot \mathbf{k} + a_3 \cdot \mathbf{l} \quad \text{mit} \quad a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1 \\ \mathbf{U} &= a_0 \cdot \sigma_0 - ia_1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - ia_2 \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - ia_3 \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} a_0 - ia_3 & -a_2 - ia_1 \\ a_2 - ia_1 & a_0 + ia_3 \end{pmatrix} = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \quad \text{mit} \quad \det \mathbf{U} = |a|^2 + |b|^2 = 1. \end{aligned}$$

Als eine Mannigfaltigkeit betrachtet ist die Menge der Einheitsquaternionen eine dreidimensionale Sphäre.

Wir nennen ein Quaternion \mathbf{m} ein **Reines Quaternion**, wenn $\mathbf{m}^* = -\mathbf{m}$. Jedes Reine Quaternion \mathbf{m} entspricht also einer spurlosen **schiefhermitischen Matrix**.

Hinweis

Im achtdimensionalen reellen Vektorraum aller komplexen 2×2 -Matrizen bilden die Reinen Quaternionen (oder genauer: die ihnen zugeordneten Matrizen) einen dreidimensionalen reellen Teilraum; die Menge \mathbb{H} aller Quaternionen bildet einen vierdimensionalen Teilvektorraum.

Wenn \mathbb{M} , wie oben, den vierdimensionalen Minkowski-Raum aller hermiteschen Matrizen bezeichnet, dann könnte man gegenüberstellen:

$$\mathbb{H} \cong \mathbb{R} \cdot \sigma_0 \oplus \frac{1}{i} \cdot \mathbb{P}; \quad \mathbb{M} \cong \mathbb{R} \cdot \sigma_0 \oplus \mathbb{P}.$$

Wir wollen die obigen Konstruktionen zum Anschauungsraum noch etwas weiterführen. Wir können einige Aussagen (über spurlose hermitesche Matrizen) dadurch durchsichtiger machen, dass wir sie als Aussagen über schiefhermitesche Matrizen (oder reine Quaternionen) formulieren.

Kommutatoren und das sog. äussere Produkt

Der Vektorraum \mathbb{E} der Reinen Quaternionen \mathbf{m} ($\mathbf{m}^* = -\mathbf{m}$) mit der Norm $\|\mathbf{m}\| = \sqrt{\mathbf{m} \cdot \mathbf{m}^*}$ ist ein euklidischer Vektorraum. Wir geben \mathbb{E} eine Orientierung, indem wir festlegen, dass $\{\mathbf{j}, \mathbf{k}, \mathbf{l}\}$ eine positiv orientierte Orthonormalbasis ist.

Satz ('Kommutatoren')

- Wenn \mathbf{m} und \mathbf{n} Reine Quaternionen sind, dann ist auch $\mathbf{q} = \frac{1}{2}(\mathbf{m} \cdot \mathbf{n} - \mathbf{n} \cdot \mathbf{m})$ ein Reines Quaternion, und es gilt

$$\begin{aligned} \mathbf{q} &\perp \mathbf{m} \quad \text{und} \quad \mathbf{q} \perp \mathbf{n} \\ \|\mathbf{q}\|^2 &= \|\mathbf{m}\|^2 \cdot \|\mathbf{n}\|^2 - \|\langle \mathbf{m} | \mathbf{n} \rangle\|^2. \end{aligned}$$

- Wenn \mathbf{m} und \mathbf{n} orthogonale Einheitsvektoren sind, dann ist $\{\mathbf{m}, \mathbf{n}, \mathbf{m} \cdot \mathbf{n}\}$ eine positiv orientierte Orthonormalbasis.

Beweis Die ersten Aussagen bestätigt man durch triviales Rechnen. Die letzte Aussage ergibt sich aus einem Stetigkeitsargument: Das Produkt der zueinander orthogonalen Einheitsvektoren ist die Hälfte des Kommutators, hat also die Norm =1. Man kann das Paar in das Paar \mathbf{j}, \mathbf{k} hineindreihen; da am Ende \mathbf{l} herauskommt, hatten wir auf dem ganzen Weg ein positiv orientiertes Orthonormalsystem $\mathbf{m}, \mathbf{n}, \mathbf{m} \cdot \mathbf{n}$.

Für die Berechnung der Norm des Kommutators bedienen wir uns der Matrixdarstellung. Wir haben oben gezeigt, dass für spurlose hermitesche Matrizen gilt

$$\langle \sigma | \tau \rangle = \frac{1}{4} \text{trace}(\sigma \tau + \tau \sigma); \quad \sigma \tau + \tau \sigma = 2c \cdot \sigma_0 \quad \text{mit} \quad c = \langle \sigma | \tau \rangle.$$

Daraus ergibt sich

$$\begin{aligned} (\sigma \tau - \tau \sigma)^2 &= (-2\tau \sigma + 2c\sigma_0)(-2\tau \sigma + 2c\sigma_0) = 4\tau \sigma \tau \sigma - 8c\tau \sigma + 4c^2\sigma_0 \\ &= 4\tau(-\tau \sigma + 2c\sigma_0)\sigma - 8c\tau \sigma + 4c^2\sigma_0 = (-4\|\tau\|^2\|\sigma\|^2 + 4\|\langle \sigma | \tau \rangle\|^2)\sigma_0. \end{aligned}$$

Die halbe Spur liefert das Quadrat der euklidischen Norm von $\mathbf{m} \cdot \mathbf{n} - \mathbf{n} \cdot \mathbf{m} = 2\mathbf{q}$.

Die Formel dürfte manchen Leser an den Formalismus des ‘äusseren Produkts’ erinnern.

$$\langle \mathbf{m} | \mathbf{n} \rangle = \|\mathbf{m}\| \|\mathbf{n}\| \cos \phi \quad \implies \quad \|\mathbf{q}\| = \|\mathbf{m}\| \|\mathbf{n}\| \sin \phi.$$

Eine zweite Gruppenwirkung der $SU(2)$

Im Abschnitt II.3 haben wir schon einmal eine interessante Aktion der Gruppe $SU(2)$ gesehen. Jeder unitären Matrix U haben wir damals eine Bijektion der erweiterten komplexen Ebene $\bar{\mathbb{C}}$ zugeordnet

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad \longmapsto \quad \varphi_U(z) = \frac{a + bz}{-\bar{b} + \bar{a}}.$$

Wenn man $\bar{\mathbb{C}}$ durch die stereographische Projektion mit der Riemann’schen Zahlenkugel identifiziert, dann liefern diese φ_U die Drehungen. (Jede Drehung kann so gewonnen werden, und U ist bis auf das Vorzeichen eindeutig bestimmt.) Wir diskutieren jetzt eine Wirkung der Gruppe $SU(2)$ auf dem Raum \mathbb{P} (oder auf dem Raum der reinen Quaternionen).

Satz

Die 3-dimensionale Gruppe $SU(2)$ wirkt auf dem 3-dimensionalen euklidischen Raum \mathbb{P} mit

$$\varphi_U(\sigma) = U\sigma U^* \quad \text{für} \quad U \in SU(2), \sigma \in \mathbb{P}.$$

Jedes $\varphi_U(\cdot)$ ist eine Drehung. Jede Drehung ist ein $\varphi_U(\cdot)$, und U ist durch φ_U bis auf das Vorzeichen eindeutig bestimmt.

Beweis

Wenn σ spurlos hermitisch ist, dann ist auch $\varphi_U(\sigma)$ spurlos hermitisch. Es gilt offenbar $\varphi_U \circ \varphi_V(\cdot) = \varphi_{UV}(\cdot)$. Jedes $\varphi_U(\cdot)$ ist eine Isometrie wegen $\det U\sigma U^* = \det \sigma$ und $\|\sigma\|^2 = -\det \sigma$.

Bemerke: Wenn $\{U(t) : t \in \mathbb{R}\}$ eine einparametrische Untergruppe von $SU(2)$ ist, dann ist $\{\varphi_{U(t)} : t \in \mathbb{R}\}$ eine einparametrische Gruppe von Drehungen, die sich in der Koordinatisierung durch $\sigma_1, \sigma_2, \sigma_3$ als eine einparametrische Gruppe von orthogonalen Matrizen $\{M(t) : t \in \mathbb{R}\}$ darstellt, (aus Stetigkeitsgründen mit $\det M(t) = 1$). Die Gruppenwirkung ist nicht effektiv; nicht nur die Gruppeneins $\sigma_0 \in SU(2)$ sondern auch $-\sigma_0$ wirkt als die Identitätsabbildung. Wir werden jetzt durch eine konkrete Rechnung zeigen, dass es keine weiteren U gibt, die die Identitätsabbildung liefern; und wir werden durch diese Rechnung auch sehen, dass die Abbildung von $SU(2)$ in $SO(3, \mathbb{R})$ surjektiv ist. Die Konstruktion liefert also einen konkreten Isomorphismus

$$SU(2)/\{\pm\sigma_0\} \longleftrightarrow SO(3, \mathbb{R}).$$

Fortsetzung des Beweises: Sei τ eine spurlose hermitesche 2×2 -Matrix mit $\|\tau\| = 1$. Sei $t \in (0, \pi)$ und $U = U(t) = \exp(-it/2 \cdot \tau) = \cos t/2 \cdot I - i \sin t/2 \cdot \tau$.

$$\varphi_U(\sigma) = U\sigma U^* = \exp(-it/2 \cdot \tau) \cdot \sigma \cdot \exp(it/2 \cdot \tau) \quad \text{für alle} \quad \sigma \in \mathbb{P}.$$

Wenn σ mit τ kommutiert, also ein Vielfaches von τ ist, dann gilt $\varphi_U(\sigma) = \sigma$.

Wenn $\sigma \perp \tau$ mit $\|\sigma\| = 1$, dann gilt $\sigma\tau + \tau\sigma = 0$ und $\rho = \frac{i}{2}(-\sigma\tau + \tau\sigma)$ ist die spurlose hermitesche Matrix, welche $\{\sigma, \tau, \rho\}$ zu einem positiv orientierten Orthonormalsystem macht.

$$\begin{aligned} \varphi_U(\sigma) &= (\cos t/2 \cdot I - i \sin t/2 \cdot \tau) \cdot \sigma \cdot (\cos t/2 \cdot I + i \sin t/2 \cdot \tau) \\ &= \cos^2 t/2 \cdot \sigma + \sin^2 t/2 \cdot \tau\sigma\tau + i(\cos t/2 \cdot \sin t/2) \cdot (\sigma\tau - \tau\sigma) \\ &= \cos t \cdot \sigma + i \sin t \cdot \frac{1}{2}(\sigma\tau - \tau\sigma) = \cos t \cdot \sigma - \sin t \cdot \rho. \end{aligned}$$

Ebenso ergibt sich $\varphi_U(\rho) = \cos t \cdot \rho + \sin t \cdot \sigma$. Die Abbildung $\varphi_{U(t)}$ ist also die Drehung um den Winkel t mit der durch τ gegebenen Drehachse. Wir haben oben gesehen, dass jedes $U \in SU(2)$ mit $U \neq \pm\sigma_0$ genau eine Darstellung der angegebenen Art besitzt. Andererseits gibt es zu jeder Drehung $\neq \text{id}$ einen Winkel $< \pi$ und eine eindeutig bestimmte orientierte Drehachse. Die Drehungen des Anschauungsraums entsprechen (in einer beliebigen gewählten Orthonormalbasis) den 3×3 -Matrizen $\in SO(3, \mathbb{R})$. Somit ist der Satz bewiesen.

Bemerkung zur geometrischen Struktur Wenn wir die Gruppenstruktur einmal ausblenden und nur an die Mannigfaltigkeitsstruktur denken, dann können wir sagen, dass die 3-dimensionale Sphäre $SU(2)$ eine zweifache Überlagerung der Mannigfaltigkeit $SO(3, \mathbb{R})$ ist; man gewinnt die Punkte von $SO(3, \mathbb{R})$, indem man gegenüberliegende Punkte der Sphäre ‘identifiziert’. Die so gewonnene Mannigfaltigkeit ist nicht einfach zusammenhängend; es gibt geschlossene Kurven, die man in $SO(3, \mathbb{R})$ nicht auf einen Punkt ‘zusammenziehen’ Eine zweifach durchlaufene geschlossene Kurve kann man allerdings immer auf einen Punkt zusammenziehen, weil man sie nämlich auf die 3-Sphäre liften kann und die n -Sphäre für $n > 1$ einfach zusammenhängend ist.

Unsere Sphäre ist in einen vierdimensionalen Raum eingebettet und es bietet sich an, die Punkte von $SO(3, \mathbb{R})$ mit den eindimensionalen Teilräumen dieses Raums identifizieren. In diesem Sinne ‘ist’ also $SO(3, \mathbb{R})$ der dreidimensionale reellprojektive Raum.

Nachdem wir jetzt den 3-dimensionalen Anschauungsraum und seine Drehungen mit Hilfe der spurlosen 2×2 -Matrizen (bzw. mit Hilfe der Reinen Quaternionen) beleuchtet haben, machen wir ganz ähnliche Überlegungen zum **Minkowski-Raum** und den eigentlichen **Lorentz-Transformationen**. Hier brauchen wir jetzt den vierdimensionalen Raum \mathbb{M} aller hermiteschen 2×2 -Matrizen $w = \alpha_0 \cdot \sigma_0 + \sigma$ mit $\sigma \in \mathbb{P}$. Die ausgezeichnete quadratische Form haben wir schon oben durch Determinanten und Spuren ausgedrückt.

$$\begin{aligned}\|w\|^2 &= \det(w) = \frac{1}{2}(\text{trace } w)^2 - \frac{1}{2}\text{trace}(w^2) \\ \langle w \| v \rangle &= \frac{1}{2}(\text{trace } w)(\text{trace } v) - \frac{1}{2}\text{trace}(w \cdot v) .\end{aligned}$$

Die eigentlichen Lorentztransformationen kann man mit einigem Recht die Drehungen des Minkowskiraums nennen. Durch den folgenden Satz wird diese Auffassung unterstrichen.

Satz (Die sog. Spinordarstellung der eigentlichen Lorentz-Gruppe)

Die 6-dimensionale Gruppe $SL(2, \mathbb{C})$ wirkt als zusammenhängende Gruppe von Isometrien auf dem 4-dimensionalen Minkowskiraum $(\mathbb{M}, \det(\cdot))$ vermöge

$$\varphi_A(w) = AwA^* \quad \text{für } A \in SL(2, \mathbb{C}), \quad w \in \mathbb{M} .$$

Jede orthochrone, orientierungstreue lineare Isometrie von \mathbb{M} ist durch ein φ_A gegeben, und A ist bis auf das Vorzeichen eindeutig bestimmt.

Beweis

Jedes φ_A ist eine lineare Isometrie; denn wegen $\det A = 1 = \det A^*$ und der Multiplikativität der Determinantenfunktion gilt

$$\|\varphi_A(w)\|^2 = \det(AwA^*) = \det w = \|w\|^2 .$$

Die infinitesimalen Operatoren der einparametrischen Gruppen in $SL(2)$ bilden einen 6-dimensionalen Vektorraum (er wird häufig mit $\mathfrak{sl}(2)$ bezeichnet). Seine Elemente γ sind die spurlosen Matrizen, eine bequeme Basis bilden die Pauli-Matrizen σ_j zusammen mit den Matrizen $i\sigma_j$. Wir schreiben $\gamma = \sigma + i\tau$ mit $\sigma, \tau \in \mathbb{P}$.

Sei $\tau \in \mathbb{P}$ mit $\|\tau\|^2 = 1$ und sei $\{\sigma_0, \tau, \rho, \sigma\}$ eine positiv orientierte Orthonormalbasis. Wir haben bereits berechnet

$$\begin{aligned}U(t) &= \exp(it\tau) = \cos t \cdot I + i \sin t \cdot \tau ; \\ \varphi_{U(-t/2)}(\sigma_0) &= \sigma_0 , \quad \varphi_{U(-t/2)}(\tau) = \tau , \\ \varphi_{U(-t/2)} \begin{pmatrix} \sigma \\ \rho \end{pmatrix} &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} \sigma \\ \rho \end{pmatrix}\end{aligned}$$

In derselben Weise gewinnen wir

$$\begin{aligned} A(t) &= \exp(t\tau) = \cosh t \cdot I + i \sinh t \cdot \tau; \\ \varphi_{A(-t/2)}(\sigma) &= \sigma, \quad \varphi_{A(-t/2)}(\rho) = \rho, \\ \varphi_{A(-t/2)} \begin{pmatrix} \sigma_0 \\ \tau \end{pmatrix} &= \begin{pmatrix} \cosh t & -\sinh t \\ \sinh t & \cosh t \end{pmatrix} \begin{pmatrix} \sigma_0 \\ \tau \end{pmatrix} \end{aligned}$$

Eine ‘Drehung’ des Minkowskiraums ist offenbar dadurch bestimmt, dass eine gewisse orthochrone positiv orientierte Orthonormalbasis $\{\sigma'_0, \sigma'_1, \sigma'_2, \sigma'_3\}$ in die Standardnormalbasis $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ hineingedreht wird. Bezeichne \mathcal{D} die Gruppe aller dieser ‘Drehungen’. Wir zeigen, dass die Zuordnung

$$SL(2, \mathbb{C}) \ni A \mapsto \varphi_A \in \mathcal{D}$$

surjektiv ist: Es gibt zunächst einmal ein $\tau \in \mathbb{P}$ in der von σ'_0 und σ_0 aufgespannten Ebene, sodass für $A = \exp(\tau)$ gilt $\varphi_A(\sigma'_0) = \sigma_0$. Dieses φ_A bildet den Orthogonalraum von σ'_0 , das heisst den von $\{\sigma'_1, \sigma'_2, \sigma'_3\}$ aufgespannten Raum auf den Orthogonalraum von σ_0 , das heisst auf den Raum \mathbb{P} ab. Dabei werden $\sigma'_1, \sigma'_2, \sigma'_3$ in eine Orthonormalbasis von \mathbb{P} überführt. Es gibt sodann ein unitäres $U = \exp(i\sigma)$, sodass φ_U diese Basisvektoren in die die Standardbasisvektoren $\sigma_1, \sigma_2, \sigma_3$ abbildet.

Es gilt $\varphi_{-\sigma_0} = \text{id} = \varphi_{\sigma_0}$; und für keine andere Matrix $A \in SU(2)$ ist φ_A die Identitätsabbildung. Wir haben also einen Gruppenisomorphismus

$$SL(2, \mathbb{C}) / \{\pm \sigma_0\} \longleftrightarrow \mathcal{D}.$$

Hier ist nochmals eine gute Gelegenheit, an den Abschnitt II zu erinnern.

Satz

Die 6-dimensionale Gruppe $SL(2, \mathbb{C})$ wirkt als eine Gruppe (nichtlinearer!) Bijektionen der Riemann’schen Zahlenkugel \mathbb{C} vermöge

$$\chi_A(z) = \frac{az + b}{cz + d} \quad \text{für} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ad - bc = 1, \quad z \in \mathbb{C} \cup \{\infty\}.$$

Jede dieser ‘Möbiustransformationen’ χ_A erhält Doppelverhältnisse. A ist durch χ_A bis auf das Vorzeichen eindeutig bestimmt.

Bezeichne \mathcal{M} die Gruppe aller lineargebrochenen Abbildungen der Riemann’schen Zahlenkugel. Die Zuordnung

$$SL(2, \mathbb{C}) \ni A \mapsto \chi_A \in \mathcal{M}$$

liefert einen Gruppenisomorphismus

$$SL(2, \mathbb{C}) / \{\pm \sigma_0\} \longleftrightarrow \mathcal{M}.$$

Folgerung: Die Gruppe aller Möbiustransformationen ist isomorph zur eigentlichen Lorentzgruppe.

IV.8 Zerlegung einer Gruppendarstellung

Die Idee der orthogonalen Zerlegung ist nicht nur für einzelne selbstadjungierte Operatoren bedeutungsvoll, sondern auch für die Darstellungen von (endlichen) Gruppen über (endlichdimensionalen) komplexen Vektorräumen. Der Anfänger kann das Thema zurückstellen.

Wenn wir im Folgenden von einer abstrakten endlichen Gruppe (G, e, \circ) sprechen, dann denken wir z. B. an die Oktaedergruppe mit 24 Elementen, die wir schon früher betrachtet haben. Die symmetrische Gruppe \mathcal{S}_3 werden wir in allen Einzelheiten behandeln. In der nächsten Vorlesung werden wir auch manchmal auf Analogien bei den unendlichen Gruppen \mathbb{Z} , $\mathbb{R}/2\pi$ und \mathbb{R} verweisen.

Definition

Es sei (G, e, \circ) eine abstrakte endliche Gruppe und V ein komplexer Vektorraum. Jedem $g \in G$ sei eine lineare Abbildung φ_g von V in sich zugeordnet, sodass gilt

$$\varphi_h(\varphi_g(\cdot)) = \varphi_{h \circ g}(\cdot) \quad \text{für alle } g, h \in G$$

Dann heißt die Familie $\Phi = \{\varphi_g : g \in G\}$ eine **Darstellung** von G über V .

Bemerkungen:

- Der Endomorphismus φ_e ist in jedem Fall eine Projektion, sagen wir die Projektion von V auf V_1 entlang V_0 ; und alle φ_g sind Surjektionen von V auf V_1 , welche die Elemente von V_0 in den Nullvektor abbilden. Die Dimension von V_1 heisst der **Rang** (oder auch die Dimension) der Darstellung. Wenn φ_e die Identitätsabbildung ist, dann spricht man von einer Darstellung von vollem Rang. In jedem Fall bilden die Einschränkungen der φ_g auf V_1 eine Darstellung über V_1 von vollem Rang.
- Wenn eine Gruppendarstellung dem Einselement e und keinem anderen g die Identitätsabbildung zuordnet, dann ist jedes Gruppenelement g durch die Abbildung $\varphi_g(\cdot)$ eindeutig bestimmt und man spricht dann von einer **treuen Darstellung** (im Englischen: faithful group representation of G over V).

Hinweis: In II.3 haben wir definiert, was eine Gruppenwirkung (group action) ist. Die Gruppendarstellungen von vollem Rang ($\varphi_e = \text{id}_V$) kann man als Spezialfälle dieses Begriffs auffassen. Bei den Gruppendarstellungen von vollem Rang geht es um lineare Bijektionen eines komplexen Vektorraums, während man bei allgemeinen Gruppenwirkungen beliebige Bijektionen einer beliebigen Menge Ω zulässt. In der Theorie der Gruppendarstellungen nutzt man die Möglichkeiten des Basiswechsels, um die Gruppenaktion transparenter zu machen.

Bemerke: Wenn eine Gruppe G auf einer endlichen Menge Ω wirkt, dann kann man daraus eine Gruppendarstellung machen, indem man jeder Permutation die entsprechende Permutationsmatrix zuordnet. Die Permutationsmatrizen, die die Einheitsspalten permutieren, werden dann als lineare Abbildungen des Spaltenraums $V = \mathbb{C}_{\mathcal{S}_p}^\Omega$ interpretiert.

Ein Spezialfall ist die ‘Standarddarstellung’ oder ‘reguläre Darstellung’; hier geht man davon aus, dass die Gruppe G durch Multiplikation von links auf die Menge G wirkt.

Satz (Die direkte Summe von Darstellungen)

Sei $V = V_1 \oplus \cdots \oplus V_m$, und sei $\Phi^{(i)} = \{\varphi_g^{(i)} : g \in G\}$ eine Darstellung von G über V_i , ($i = 1, \dots, m$). Dann liefern die Endomorphismen

$$\varphi_g : v = v_1 \oplus \cdots \oplus v_m \mapsto \varphi_g(v) = \varphi_g^{(1)}(v_1) \oplus \cdots \oplus \varphi_g^{(m)}(v_m)$$

eine Darstellung Φ von G über V . Man schreibt $\Phi = \Phi^{(1)} \oplus \cdots \oplus \Phi^{(m)}$.
Der Beweis ist trivial.

Hilfssatz 1 Sei π die Projektion von V auf V_1 entlang V_0 . Für Endomorphismen φ gilt

$$\begin{aligned} \varphi(V_1) \subseteq V_1 &\iff \varphi \circ \pi = \pi \circ \varphi \circ \pi \\ \varphi(V_1) \subseteq V_1 \text{ und } \varphi(V_0) \subseteq V_0 &\iff \varphi \circ \pi = \pi \circ \varphi. \end{aligned}$$

Beweis Wenn für alle $\pi(v)$ das φ -Bild in V_1 liegt, dann bedeutet das $\pi(\varphi(\pi(v))) = \varphi(\pi(v))$ für alle v . Damit ist die erste Behauptung bewiesen.

Sei $\varphi \circ \pi = \pi \circ \varphi \circ \pi$. Es gilt dann

$$\begin{aligned} (\text{id} - \pi)\varphi(\text{id} - \pi) &= \varphi(\text{id} - \pi) \iff \varphi - \varphi \circ \pi - \pi \circ \varphi + \pi \circ \varphi \circ \pi = \varphi - \varphi \circ \pi \\ &\iff \pi \circ \varphi = \pi \circ \varphi \circ \pi = \varphi \circ \pi. \end{aligned}$$

Die Einschränkungen der φ_g auf V_1 liefern in diesem Fall eine Darstellung $\Phi^{(1)}$ über V_1 ; die Einschränkungen auf V_0 liefern eine Darstellung $\Phi^{(0)}$, und es gilt $\Phi = \Phi^{(0)} \oplus \Phi^{(1)}$.

Sprechweise Die Projektion π heisst **verträglich** mit der Darstellung Φ , wenn gilt

$$\pi \circ \varphi_g = \varphi_g \circ \pi \text{ für alle } g \in G.$$

Wir haben eben gesehen: wenn π eine mit Φ verträgliche Projektion ist, dann auch $\text{id} - \pi$. Wenn π_1 und π_2 kommutierende Projektionen sind, die mit Φ verträglich sind, dann ist auch $\pi_1 \circ \pi_2$ eine mit Φ verträgliche Projektion.

Satz (Direkte Zerlegung einer Darstellung)

Sei Φ eine Darstellung von G über V und seien π_1, \dots, π_m paarweise disjunkte Projektionen, die mit Φ verträglich sind.

$$\pi_i \circ \varphi_g = \varphi_g \circ \pi_i \text{ für alle } g \in G, \quad i = 1, \dots, m$$

Dann ist auch $\pi = \pi_1 + \cdots + \pi_m$ eine mit Φ verträgliche Projektion.

Die Einschränkung von Φ auf $V_i = \pi_i(V)$ ist eine Darstellung $\Phi^{(i)}$. Mit $\pi_0 = \text{id} - \pi$ gilt

$$\begin{aligned} V &= V_0 \oplus V_1 \oplus \cdots \oplus V_m & \Phi &= \Phi^{(0)} \oplus \Phi^{(1)} \oplus \cdots \oplus \Phi^{(m)}; \\ \varphi_g(v) &= \varphi_g(v_0 + v_1 + \cdots + v_m) = \varphi_g^{(0)}(v_0) + \varphi_g^{(1)}(v_1) + \cdots + \varphi_g^{(m)}(v_m). \end{aligned}$$

Man sagt in diesem Fall, die Darstellung Φ über V und die direkte Zerlegung $V = V_0 \oplus \cdots \oplus V_m$ sind miteinander **verträglich**.

Bemerke ('Blockmatrizen'): In der Situation des Satzes sei $\{v_j : j \in I_i\}$ eine Basis von V_i für $i = 0, \dots, m$ und somit $\{v_j : j \in I_0 + \cdots + I_m\}$ eine Basis von $V = V_0 \oplus \cdots \oplus V_m$. Die φ_g darstellende Matrix A_g vom Format $(I_0 + \cdots + I_m) \times (I_0 + \cdots + I_m)$ zerfällt dann in Blöcke vom Format $I_i \times I_i$; das soll heissen: die Einträge a_{jk} von A_g können nur dann von Null verschieden sein, wenn j und k zum gleichen I_i gehören.

Es sollte plausibel sein, dass eine verträgliche direkte Zerlegung einen informativen Einblick in die Struktur der Darstellung liefert. Wenn es beispielsweise für eine Darstellung eine direkte Zerlegung in eindimensionale Darstellungen gibt, dann bedeutet das, dass man alle φ_g diagonalisieren kann. Wir werden sehen, dass es eine solche Zerlegung genau dann gibt, wenn die Gruppe der Transformationen φ_g kommutativ ist.

Definition

Die Darstellung $\{\varphi_g(\cdot) : g \in G\}$ über V heisst **reduzibel**, wenn es einen echten Teilvektorraum gibt, der von allen $\varphi_g(\cdot)$ in sich abgebildet wird. Andernfalls heisst sie **irreduzibel**.

Hilfssatz 2 (Satz vom invarianten Komplement)

Sei Φ eine Darstellung über V von vollem Rang; und sei V_1 ein Teilraum, der von allen φ_g in sich abgebildet wird, $\varphi_g(V_1) \subseteq V_1$ für alle φ_g . Dann gibt es einen Komplementärraum V_0 , welcher ebenfalls von allen φ_g in sich abgebildet wird.

Mit anderen Worten: Zu jedem invarianten Teilraum gibt es eine Projektion auf V_1 , welche mit der Darstellung verträglich ist. $\pi\varphi_g = \varphi_g\pi$ für alle $g \in G$. Es kann mehrere invariante Komplemente geben.

Wir beweisen zuerst einen Hilfssatz, den wir immer wieder brauchen werden.

Hilfssatz 3 (Symmetrisierung)

Für jeden Endomorphismus σ kommutieren alle φ_g mit dem sog. symmetrisierten Endomorphismus $\bar{\sigma}$.

$$\bar{\sigma} = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ \sigma \circ \varphi_a; \quad \bar{\sigma} \circ \varphi_g = \varphi_g \circ \bar{\sigma} \quad \text{für alle } g.$$

Beweis

$$\bar{\sigma} \circ \varphi_g = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \sigma \varphi_{ag} = \varphi_g \circ \frac{1}{|G|} \sum_{a \in G} \varphi_{(ag)^{-1}} \sigma \varphi_{ag} = \varphi_g \circ \frac{1}{|G|} \sum_{b \in G} \varphi_{b^{-1}} \sigma \varphi_b = \varphi_g \circ \bar{\sigma};$$

denn die $b = ag$ durchlaufen die gesamte Gruppe.

Beweis von Hilfssatz 2 Es sei τ irgendeine Projektion auf den invarianten Teilraum V_1 . Es gelte also

$$\tau \circ \varphi_a \circ \tau = \varphi_a \circ \tau \quad \text{für alle } a \in G.$$

Wir zeigen, dass der symmetrisierte Endomorphismus $\bar{\tau} = \pi$ eine Projektion auf V_1 ist. Zunächst einmal haben wir

$$\begin{aligned}\tau \circ \pi &= \frac{1}{|G|} \sum_{a \in G} (\tau \circ \varphi_{a^{-1}} \circ \tau) \circ \varphi_a = \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ \tau \circ \varphi_a = \pi; \\ \pi \circ \tau &= \frac{1}{|G|} \sum_{a \in G} \varphi_{a^{-1}} \circ (\tau \circ \varphi_a \circ \tau) = \frac{1}{|G|} \sum_{a \in G} (\varphi_{a^{-1}} \circ \varphi_a) \circ \tau = \tau.\end{aligned}$$

Daraus folgt $\pi \circ \pi = \pi \circ \tau \circ \pi = \tau \circ \pi = \pi$. $\pi = \tau \circ \pi$ ist also eine Projektion auf $V_1 = \tau(V)$. Wie jeder symmetrisierte Endomorphismus kommutiert π mit allen φ_g . Hilfssatz 1 liefert die Behauptung.

Sei $V = V_1 \oplus \dots \oplus V_m$ irgend eine Φ -verträgliche direkte Zerlegung des Darstellungsraums. Wenn man die einzelnen V_j weiter in Φ -verträglicher Weise zerlegt, dann ergibt das eine verfeinerte Φ -verträgliche direkte Zerlegung von V . Man kann das so weit treiben bis alle Summanden irreduzibel sind. Das ergibt den

Satz

Jede Darstellung einer endlichen Gruppe besitzt eine Zerlegung in irreduzible Darstellungen $\Phi = \Phi^{(1)} \oplus \dots \oplus \Phi^{(m)}$.

Hinweis Die Zerlegung in irreduzible Darstellungen ist nicht eindeutig bestimmt. Wir werden eine etwas gröbere Zerlegung kennenlernen, die eindeutig bestimmt ist; man nennt sie manchmal die kanonische Zerlegung der Gruppendarstellung.

Definitionen

Seien V und W komplexe Vektorräume. Sei $\Phi = \{\varphi_g(\cdot) : g \in G\}$ eine Darstellung über V , und sei $\Psi = \{\psi_g(\cdot) : g \in G\}$ eine Darstellung über W .

- Eine lineare Abbildung $\chi : V \rightarrow W$ heisst ein **Verflechtungsoperator** (im Englischen ‘intertwining operator’), wenn gilt

$$\chi \circ \varphi_g(\cdot) = \psi_g \circ \chi(\cdot) \quad \text{für alle } g \in G.$$

- Die Darstellungen heissen **äquivalente Darstellungen**, wenn es einen invertierbaren Verflechtungsoperator gibt. Mit anderen Worten, wenn es einen Vektorraum-Isomorphismus χ gibt, sodass

$$\chi \circ \varphi_g \circ \chi^{-1} = \psi_g \quad \text{für alle } g \in G.$$

- Die Darstellung Ψ heisst **disjunkt** zur Darstellung Φ , wenn der Nulloperator der einzige Verflechtungsoperator ist.

Hilfssatz 4

Es seien Φ und Ψ Darstellungen von vollem Rang der endlichen Gruppe G , Φ über V und Ψ über W . Es sei $\chi : V \rightarrow W$ ein Verflechtungsoperator und nicht der Nulloperator. $V_0 = \{v : \chi v = 0\}$ sei der Kern der Abbildung χ , und $W_1 = \chi(V)$ sei das Bild. Es gilt dann

- V_0 ist ein Φ -invarianter Teilraum von V , und W_1 ist ein Ψ -invarianter Teilraum von W .
- Sei V_1 das Φ -invariante Komplement von V_0 , und sei $\Phi^{(1)}$ die Einschränkung von Φ auf V_1 . $\Phi^{(1)}$ ist dann äquivalent zur Einschränkung von Ψ auf W_1 .

Beweis: Für jedes g sind φ_g und ψ_g bijektiv. Der Kern V_0 ist Φ -invariant; denn

$$v \in V_0 \iff \chi v = 0 \iff \psi_g \chi v = 0 \iff \chi \varphi_g v = 0 \iff \varphi_g v \in V_0.$$

Die Einschränkung des Verflechtungsoperators χ auf V_1 ist eine bijektive Abbildung $\chi_1 : V_1 \rightarrow W_1$. Für die auf V_1 eingeschränkten φ und die auf W_1 eingeschränkten ψ gilt

$$\chi_1 \varphi_g^{(1)} = \psi_g^{(1)} \chi_1, \quad \text{also} \quad \chi_1 \varphi_g^{(1)} \chi_1^{-1} = \psi_g^{(1)} \quad \text{für alle } g \in G.$$

Satz ('Lemma von I. Schur')

Es seien Φ und Ψ irreduzible Darstellungen der endlichen Gruppe G , Φ über V und Ψ über W .

- a) Wenn $\chi : V \rightarrow W$ ein Verflechtungsoperator ist, dann gilt entweder $\chi = 0$, oder aber χ ist nichtsingulär und die beiden irreduziblen Darstellungen sind äquivalent.
- b) Wenn ρ ein Endomorphismus ist mit $\rho \varphi_g = \varphi_g \rho$ für alle g , dann ist ρ ein Vielfaches der Identität.

Der Beweis von a) ergibt sich aus dem Hilfssatz: Der Kern ist entweder der ganze Raum oder der Nullraum; ebenso steht es um das Bild. Der Beweis von b) ergibt sich aus der Beobachtung, dass $\rho - \lambda I$ mit allen φ_g kommutiert.

Corollare Wenn ein φ_h in einer irreduziblen Darstellung mit allen φ_g kommutiert, dann ist φ_h ein Vielfaches des Identitätsoperators.

Alle irreduziblen Darstellungen einer kommutativen Gruppe sind eindimensional.

Bemerkung Ein Endomorphismus eines eindimensionalen Vektorraums ist durch einen Skalar gegeben: jeder Vektor v wird mit dieser Zahl multipliziert. Diese Zahl ist die Spur des Endomorphismus.

Sprechweisen

Es sei $\Phi = \{\varphi_g : g \in G\}$ eine (nicht notwendigerweise irreduzible) Darstellung über V . Die Algebra aller Endomorphismen ρ mit $\rho \varphi_g = \varphi_g \rho$ für alle g , heisst die **Vertauschungsalgebra**.

Es seien Φ über V und Ψ über W Darstellungen. Der Vektorraum aller Verflechtungsoperatoren $\chi : V \rightarrow W$ bezeichnet man oft mit $\mathcal{R}(\Phi, \Psi)$.

Konstruktion

Es seien Φ und Ψ Darstellungen der endlichen Gruppe G von vollem Rang, Φ über V und Ψ über W . Zu jedem Vektorraum-Homomorphismus $\chi : V \rightarrow W$ gewinnt man einen Verflechtungsoperator $\bar{\chi} \in \mathcal{R}(\Phi, \Psi)$, wenn man konstruiert

$$\bar{\chi}(v) = \sum_{g \in G} \psi_{g^{-1}} \chi \varphi_g(v) \quad \text{für alle } v \in V.$$

Wie oben beim Beweis des Hifssatzes beweist man $\bar{\chi} \varphi_h = \psi_h \bar{\chi}$ für alle $h \in G$.

Unitäre Darstellungen

Definition (Isometrie)

Es sei $(V, \langle \cdot | \cdot \rangle)$ ein Hilbertraum. Eine invertierbare lineare Abbildung $\chi : V \rightarrow V$ heißt eine **Isometrie**, wenn gilt

$$\|\chi(v)\| = \|v\| \quad \text{für alle } v \in V.$$

Mit der Polarisierungskonstruktion zeigt man $\langle \chi(v) | \chi(w) \rangle = \langle v | w \rangle$ für alle v, w .

Definition

Eine Darstellung über den Hilbertraum $(V, \langle \cdot | \cdot \rangle)$ heißt eine **unitäre Darstellung**, wenn alle $\varphi_g(\cdot)$ Isometrien sind.

(Der Name ist nicht besonders glücklich, man sollte vielleicht lieber von einer Darstellung durch Isometrien sprechen. Die Idee, die hinter dem Namen steckt, ist die: Wenn man auf V eine Orthonormalbasis einführt, dann entspricht jeder Isometrie eine unitäre Matrix.)

Darstellung mit unitären Matrizen

Eine Isometrie des Spaltenraums \mathbb{C}_{Sp}^I mit der „natürlichen“ Hilbertraumstruktur $\|y\| = \sqrt{y^* y}$ ist durch eine unitäre $I \times I$ -Matrix gegeben. Jede unitäre Matrix U liefert andererseits eine Isometrie $x \mapsto Uy$. In der Tat: Sei U eine $I \times I$ -Matrix mit $\|Uy\|^2 = \|y\|^2$ für alle y . Dann gilt auch

$$z^* \cdot U^* Uy = z^* y \quad \text{für alle } y, z \in \mathbb{C}_{\text{Sp}}^I.$$

Daraus ergibt sich $U^* U = E$, U ist eine unitäre Matrix.

Eine unitäre Darstellung ist hier also eine Schar unitärer Matrizen $\{A_a : a \in G\}$ mit $A_a A_b = A_{ab}$ für alle $a, b \in G$.

Definition

Seien $(V, \langle \cdot | \cdot \rangle)$ und $(W, \langle \cdot | \cdot \rangle)$ n -dimensionale Hilberträume. Sei $\{\varphi_g(\cdot) : g \in G\}$ eine unitäre Darstellung über V und $\{\psi_g(\cdot) : g \in G\}$ eine unitäre Darstellung über W . Man sagt, die Darstellungen seien **unitär äquivalent**, wenn eine Isometrie

$$\chi : (V, \|\cdot\|) \rightarrow (W, \|\cdot\|), \quad (\|\chi(v)\| = \|v\| \text{ für alle } v)$$

existiert, sodass

$$\chi \circ \varphi_g(\cdot) = \psi_g \circ \chi(\cdot) \quad \text{für alle } g \in G.$$

Theorem

- a) Zu jeder Darstellung $\{\varphi_g(\cdot) : g \in G\}$ auf V existiert eine Hilbertraumnorm, sodass die Darstellung eine unitäre Darstellung ist.
- b) Wenn $\{\varphi_g : g \in G\}$ über V und $\{\psi_g : g \in G\}$ über W äquivalente Darstellungen sind, dann gibt es Hilbertraumnormen auf V und W und eine isometrische Abbildung $\chi : (V, \|\cdot\|) \rightarrow (W, \|\cdot\|)$, welche die unitäre Äquivalenz etabliert.

Beweis von a) :

Wir wählen auf V irgendeine Basis $\{u_i : i \in I\}$. Wir haben dann eine Darstellung durch $I \times I$ -Matrizen $\{B_a : a \in G\}$.

1. Wählen wir irgendeine positiv definite $I \times I$ -Matrix Q und betrachten wir dazu die positiv definite Matrix

$$H := \frac{1}{|G|} \sum_{a \in G} B_a^* \cdot Q \cdot B_a .$$

H heißt die **symmetrisierte Matrix**. Sie hat die bemerkenswerte Eigenschaft

$$B_g^* \cdot H \cdot B_g = H \quad \text{für alle } g \in G .$$

2. In der Tat gilt wegen $B_a \cdot B_g = B_{ag}$, und $B_g^* \cdot B_a^* = (B_{ag})^*$

$$B_g^* \cdot H \cdot B_g = \frac{1}{|G|} \sum_a B_g^* B_a^* Q B_a B_g = \frac{1}{|G|} \sum_h B_h^* Q B_h = H ,$$

denn die Summe geht über alle $h = ag$ ($a \in G$) und das bedeutet nur eine andere Aufzählung aller Elemente von G .

3. H definiert eine Hilbertraumstruktur und alle B_g sind Isometrien für diese Struktur

$$\|y\|_H^2 = y^* H y , \quad \|B_g y\|_H^2 = y^* B_g^* H B_g y = y^* H y = \|y\|_H^2 .$$

4. Unsere Hilbertraumnorm $\|\cdot\|_H$ liefert eine Hilbertraumnorm auf dem ursprünglichen Darstellungsraum V . Die $\varphi_g(\cdot)$ sind nun Isometrien. Wenn wir eine Orthonormalbasis einführen $\{v_j : j \in J\}$, dann werden die $\varphi_g(\cdot)$ durch unitäre $J \times J$ -Matrizen repräsentiert.

Beweis von b):

Es sei $\Phi = \{A_g\}$ eine Darstellung mit unitären $J \times J$ -Matrizen und $\Psi = \{B_g\}$ eine Darstellung mit unitären $I \times I$ -Matrizen. $\chi : \mathbb{C}_{Sp}^J \rightarrow \mathbb{C}_{Sp}^I W$ sei ein invertierbarer Verflechtungsoperator, dargestellt durch die invertierbare $I \times J$ -Matrix T ; $TA_g = B_g T$.

Wie wir wissen, besitzt jede invertierbare Matrix eine Faktorisierung $T = US$, wo U unitär

und S positivdefinit ist. Wir haben $SA_g = (U^*B_gU)S$. Die positivdefinite Matrix S besitzt eine unitäre Diagonalisierung $S = V^*DV$. So ergibt sich

$$V^*DVA_g = U^*B_gV^*DV \quad \text{also} \quad D(VA_gV^*) = (W^*B_gW)D$$

mit der unitären Matrix $W = UV^*$.

Daraus folgt $VA_gV^* = W^*B_gW$ für alle $g \in G$, wie wir sofort beweisen werden mit Hilfe eines allgemeinen Hilfssatzes über unitäre Matrizen, die mit einer positivdefiniten Diagonalmatrix multipliziert werden. Wenn dieser Hilfssatz bewiesen ist, dann haben wir die unitäre Äquivalenz der beiden unitären Darstellungen.

Hilfssatz 5 *Sei D eine positivdefinite Diagonalmatrix; und seien A und B unitäre Matrizen mit $DA = BD$. Es gilt dann $A = B$, und A zerfällt in Blöcke gemäss den verschiedenen Eigenwerten von D .*

Beweis

$$DA = BD \iff A^{-1}D^{-1} = D^{-1}B^{-1} \iff DA^* = B^*D.$$

In den Positionen (k, l) bzw. (l, k) haben wir die Einträge

$$\begin{aligned} d_k a_{kl} &= b_{kl} d_l & d_l a_{lk} &= b_{lk} d_k \\ d_k \bar{a}_{lk} &= \bar{b}_{lk} d_l & d_l \bar{a}_{kl} &= \bar{b}_{kl} d_k. \end{aligned}$$

Für jedes Paar mit $d_k = d_l$ ergibt sich $a_{kl} = b_{kl}$. Wenn $d_k \neq d_l$, dann ergibt sich $a_{kl} = 0 = b_{kl}$, $a_{lk} = 0 = b_{lk}$. Die Matrizen A und B zerfallen also in Blöcke gemäss den verschiedenen Werten d . Und es gilt $A = B$.

Satz

Es sei Φ eine mit der direkten Zerlegung $V = V_1 \oplus \cdots \oplus V_m$ verträgliche Darstellung. Wenn V ein Hilbertraum ist und Φ eine unitäre Darstellung, dann handelt es sich um eine orthogonale Zerlegung $V = V_1 \boxplus \cdots \boxplus V_m$. (Man notiert $\Phi = \Phi^{(1)} \boxplus \cdots \boxplus \Phi^{(m)}$.)

Beweis Es genügt zu zeigen: Das verträgliche Komplement zu einem Φ -invarianten Teilraum V_1 ist das orthogonale Komplement V_1^\perp .

1. Wenn φ eine Isometrie ist, dann gilt

$$\langle w | v \rangle = \langle \varphi(w) | \varphi(v) \rangle = \langle \varphi^* \varphi(w) | v \rangle \quad \text{also} \quad \varphi^* = \varphi^{-1}$$

Bei einer unitären Darstellung gilt also $\varphi_g^* = \varphi_{g^{-1}}$ für alle $g \in G$.

2. Wenn $v_1 \in V_1$, dann gilt (für alle g) $\varphi_g^*(v_1) \in V_1$ und für jedes $v_0 \in V_0 = V_1^\perp$

$$\langle \varphi_g(v_0) | v_1 \rangle = \langle v_0 | \varphi_g^*(v_1) \rangle = 0 \quad \text{also} \quad \varphi_g(v_0) \in V_0.$$

Der Satz von der Zerlegbarkeit einer Darstellung in irreduzible Darstellungen lautet also im Falle der Darstellungen durch Isometrien:

Satz

Eine unitäre Darstellung Φ kann man verträglich zerlegen, sodass die Summanden zueinander orthogonale irreduzible unitäre Darstellungen sind.

Wir haben gesehen: Die Theorie der Darstellungen einer endlichen Gruppe über komplexen Vektorräumen wird nicht eingeschränkt, wenn man sich auf unitäre Darstellungen über Hilberträumen beschränkt. Die Hilbertraumstruktur kann u. U. manche Konstruktionen durchsichtiger machen.

Hinweise Einen Schönheitsfehler der Zerlegung in irreduzible Komponenten mag man darin sehen, dass die Zerlegung nur dann eindeutig bestimmt ist, wenn es keine höheren Vielfachheiten gibt ($c_r = 0$ oder 1 für alle r). Man kennt das Phänomen von der Hauptachsenzerlegung; die Zerlegung ist nur dann eindeutig, wenn alle Eigenwerte die Vielfachheit ≤ 1 besitzen.

Wir werden sehen: Eindeutig ist diejenige verträgliche Zerlegung von Φ , die man aus einer Zerlegung in irreduzible Summanden dadurch erhält, dass man alle zueinander äquivalenten Summanden zusammenfasst. Man nennt sie manchmal die kanonische Zerlegung. Diese verträgliche Zerlegung ist zu charakterisieren als **Zerlegung in paarweise disjunkte Darstellungen**. Jeder direkte Summand ist die disjunkte Summe einer gewissen Anzahl von äquivalenten Darstellungen.

Wir werden sehen, dass es für jede endliche Gruppe (bis auf Äquivalenz!) nur endlich viele irreduzible Darstellungen gibt. Sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste von paarweise nichtäquivalenten irreduziblen Darstellungen. Wir werden weiter sehen: Für jede Darstellung Φ sind die Vielfachheiten c_r , mit welchen die einzelnen $\Phi^{(r)}$ in Φ vorkommen, eindeutig bestimmt. Man notiert die Zerlegung in disjunkte Darstellungen

$$\Phi = c_1 \Phi^{(1)} \oplus \cdots \oplus c_m \Phi^{(m)} = \bigoplus_{r \in L} c_r \Phi^{(r)} .$$

Die Zerlegung einer Darstellung in irreduzible Summanden sieht man gerne in Analogie zur Faktorisierung einer natürlichen Zahl in Primfaktoren; die Zerlegung in disjunkte Darstellungen entspricht in dieser Analogie der Faktorisierung in paarweise teilerfremde Primzahlpotenzen.

Die Standarddarstellung oder ‘Reguläre Darstellung’

Definition Es bezeichne V den Hilbertraum aller komplexwertigen Funktionen $f(\cdot)$ auf der endlichen Gruppe G . G operiert auf V , wenn wir (für $b \in G$) definieren

$$\tilde{\varphi}_b : V \longrightarrow V ; \quad (\tilde{\varphi}_b(f))(a) = f(b^{-1}a) \quad \text{für } a \in G .$$

In der Tat gilt $\tilde{\varphi}_e(f) = f$, $\tilde{\varphi}_b(\tilde{\varphi}_c(f)) = \tilde{\varphi}_{bc}(f)$ für alle $f \in V$. Die $\tilde{\varphi}_b$ sind Isometrien. Die unitäre Darstellung $\tilde{\Phi} = \{\tilde{\varphi}_b : b \in G\}$ heisst die Standarddarstellung von G . Die Funktion $\chi_a(\cdot)$, die im Punkt d den Wert 1 und sonst überall den Wert 0 hat, wird durch die Abbildung $\tilde{\varphi}_b$ in die Funktion $\chi_{ba}(\cdot)$ abgebildet.

Bemerkungen

1. Man kann die $\tilde{\varphi}_b$ als Pullback-Abbildungen deuten. $\tilde{\varphi}_b$ ist offenbar der Pullback zur Abbildung $G \ni a \mapsto b^{-1}a \in G$.
2. Man kann V mit dem Raum der G -Spalten identifizieren. Die $\tilde{\varphi}_b$ sind dann durch Permutationsmatrizen $P_b(\cdot, \cdot)$ dargestellt.
3. Eine weitere bequeme Beschreibung der $f \in V$ liefern die sog. formalen Linearkombinationen

$$f = \sum_{a \in G} f(a) \cdot a; \quad \tilde{\varphi}_b f = \sum_{a \in G} f(a) \cdot ba = \sum_{a \in G} f(b^{-1}a) \cdot a.$$

(Man könnte sagen, dass das Symbol a der G -Spalte entspricht, die in der Position a den Eintrag 1, und sonst überall den Eintrag 0 hat.)

Satz

Es sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste irreduzibler Darstellungen von G . n_r bezeichne die Dimension von $\Phi^{(r)}$.

Wenn man die Standarddarstellung $\tilde{\Phi}$ in irreduzible Bestandteile zerlegt, dann tritt jedes $\Phi^{(r)}$ mit der Vielfachheit $c_r = n_r$ auf. Es gilt $\sum n_r^2 = |G|$.

Der Beweis führen wir, nachdem wir uns im nächsten Abschnitt genauer mit den irreduziblen Darstellungen und ihren Spurfunktionen auseinandergesetzt haben.

Beispiel 1 (Die symmetrische Gruppe \mathcal{S}_3)

Die symmetrische Gruppe \mathcal{S}_3 hat 6 Elemente: drei Transpositionen, zwei Dreierzyklen und die Gruppeneins e . (Dies ist die Unterscheidung nach den drei Konjugationsklassen.) Neben der trivialen Darstellung gibt es eine weitere eindimensionale Darstellung; ihre Spurfunktion ist die ‘Signatur’; die Signatur hat den Wert -1 in den Transpositionen und sonst den Wert $+1$. Es gibt noch genau eine weitere irreduzible Darstellung, und zwar eine zweidimensionale. Wir erraten leicht eine Realisation durch unitäre 2×2 -Matrizen: Wir identifizieren die drei Punkte, die es zu permutieren gilt, mit den dritten Einheitswurzeln in der komplexen Ebene \mathbb{C} , und fragen nach den linearen Transformationen der Ebene, welche das von ihnen gebildete gleichseitige Dreieck in sich überführen. In komplexer Notation sind das (mit $\theta = 2\pi/3$)

$$z \mapsto z; \quad z \mapsto e^{i\theta}z; \quad z \mapsto e^{-i\theta}z; \quad z \mapsto \bar{z}; \quad z \mapsto e^{-i\theta}\bar{z}; \quad z \mapsto e^{i\theta}\bar{z}.$$

In Real- und Imaginärteil ausgedrückt leisten das die unitären Matrizen

$$\begin{pmatrix} +1 & 0 \\ 0 & +1 \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ +\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & +\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}; \\ \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & +\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & +\frac{1}{2} \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}.$$

Die Spurfunktion dieser zweidimensionalen Darstellung hat den Wert 2 im Einselement, den Wert -1 in den zyklischen Permutationen, und den Wert 0 in den Transpositionen.

Beispiel 2 (Die zyklische Gruppe)

Sei N eine natürliche Zahl. Die additive Gruppe der Reste modulo N heisst bekanntlich die zyklische Gruppe der Ordnung N . Als kommutative Gruppe besitzt sie nur eindimensionale irreduzible Darstellungen. Unser Satz besagt, dass es genau $N = |G|$ irreduzible Darstellungen gibt; sie sind durch multiplikative Funktionen gegeben, die Spurfunktionen $\chi^{(r)}(\cdot)$. Und diese kann man hier natürlich ohne grosse Theorie sofort angeben. Die irreduziblen Spurfunktionen sind die Funktionen

$$\chi^{(r)}(k) = \exp\left(\frac{2\pi r}{N} \cdot k\right) \quad \text{für } k \in G = \mathbb{Z}/N\mathbb{Z}.$$

In der Standarddarstellung sind die Endomorphismen die Verschiebungen der N -periodischen Folgen $(a_k)_{k \in G}$. Die Zerlegung in irreduzible Summanden ist die orthogonale Zerlegung der N -periodischen Folgen in Linearkombinationen der Eigenfunktionen $\exp(\frac{2\pi r}{N} \cdot k)$. (Diese Zerlegung ist bekanntlich das Thema der diskreten Fourier-Transformation.)

Hinweis

Für gewisse Klassen von unendlichen Gruppen gibt es eine Darstellungstheorie, die in mancherlei Hinsicht ähnlich ist zur Darstellungstheorie für endliche Gruppen. Man beschränkt sich da allerdings von vorneherein auf unitäre Darstellungen auf Hilberträumen (was für endliche Gruppen keine Einschränkung bedeutet, wie wir gesehen haben). Wenn es um die irreduziblen Darstellungen kompakter Gruppen geht, dann werden nur endlichdimensionale Hilberträume benötigt; denn nach dem berühmten ‘Satz von Peter und Weyl’ ist jede irreduzible Darstellung einer kompakten Gruppe endlichdimensional. Es gibt hier abzählbar unendlich viele irreduzible Darstellungen. Die Standarddarstellung zerlegt sich in eine abzählbare direkte Summe irreduzibler unitärer Darstellungen. –Man muss hier etwas tiefer in die Analysis einsteigen.

Ein Beispiel ist die kommutative kompakte Gruppe $G = \mathbb{R}/2\pi$. Die eindimensionalen irreduziblen Darstellungen sind durch die stetigen multiplikativen 2π -periodischen Funktionen gegeben; das sind die Funktionen $\chi^{(r)}(t) = e^{irt}$, $r \in \mathbb{L} = \mathbb{Z}$. Der Darstellungsraum V für die Standarddarstellung ist der Hilbertraum aller quadratintegrierbaren 2π -periodischen Funktionen; das ist die Vervollständigung des Raums der trigonometrischen Polynome. $V = \mathcal{L}^2(\mathbb{R}/2\pi)$. Der Operator $\tilde{\varphi}_b$ zu $b \in G$ ist die Abbildung $f(t) \mapsto f(t - b)$. Die Funktionen e^{irt} sind Eigenfunktionen für alle diese Verschiebungoperatoren. Die Verschiebungoperatoren bilden eine einparametrische unitäre Gruppe. Es genügt, wenn wir uns mit dem infinitesimalen Operator befassen. Der infinitesimale Operator $-iH$ ($U_t = \exp(-itH)$) bildet die guten Funktionen aus V in Funktionen aus V ab:

$$\begin{aligned} f &\mapsto \lim_{h \rightarrow 0} \frac{1}{h} (U_h f - f) = -iHf, \\ -iHf(t) &= \lim_{h \rightarrow 0} \frac{1}{h} (f(t - h) - f(t)) = -\frac{\partial}{\partial t} f(t); \quad \text{kurz} \quad H = \frac{1}{i} \frac{\partial}{\partial t}. \end{aligned}$$

Der infinitesimale Operator ist also (bis auf den Faktor $-i$) die Differentiation.

In Analogie zu dem, was wir für kommutative endliche Gruppen bewiesen haben, gilt: Alle irreduziblen Darstellungen sind eindimensional und die Standarddarstellung besitzt

eine eindeutige (hier allerdings unendliche) Zerlegung in irreduzible Komponenten. Die orthogonale Zerlegung der Standarddarstellung in die irreduziblen Bestandteile (jeder mit der Vielfachheit 1) bedeutet für die $v \in V = L^2(\mathbb{R}/2\pi)$ die **Fourier-Zerlegung** .

Didaktischer Hinweis

Unser ‘Beispiel’ $\mathbb{R}/2\pi$ ist von der Algebra her nicht sehr interessant; mit der (durchaus interessanten) Analysis werden wir uns in einer späteren Vorlesung auseinandersetzen.

Von der Analysis her nicht schwieriger, algebraisch jedoch hochinteressant (und für die Anwendungen äusserst wichtig) ist die Theorie der irreduziblen Darstellungen der uns bereits gut bekannten nichtkommutativen kompakten Gruppen SO_3 oder SU_2 . Die Chemiker würden sie gerne im Grundstudium behandelt wissen. Die Mathematiker können jedoch zu diesem Zeitpunkt allenfalls mit Analogien aus der Theorie der Darstellungen endlicher Gruppen zu Diensten sein.

IV.9 Spurfunktionen. Die Charaktertafel einer endlichen Gruppe G .

Zu jeder Darstellung Φ einer endlichen Gruppe G gehört eine Spurfunktion. Das ist die komplexwertige Funktion

$$\chi^\Phi(g) = \text{trace } \varphi_g \quad \text{für } g \in G.$$

Äquivalente Darstellungen haben offenbar dieselbe Spurfunktion. Wir werden sehen, dass die Spurfunktion die Darstellung bis auf Äquivalenz eindeutig bestimmt.

Die Spurfunktion einer direkten Summe $\Phi = \Phi_1 \oplus \Phi_2$ ist die Summe der Spurfunktionen. $\chi^\Phi(e) = n$ liefert den Rang ('Dimension') der Darstellung. Es gilt stets $\chi^\Phi(g^{-1}) = \bar{\chi}^\Phi(g)$; denn jede Darstellung ist zu einer unitären Darstellung äquivalent. Sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste der irreduziblen Darstellungen der Gruppe G , sei $\chi^{(r)}(\cdot)$ die Spurfunktion von $\Phi^{(r)}$, und sei $\chi^{(r)}(e) = n_r$ ihre Dimension. Jede Spurfunktion ist von der Form $\chi^\Phi(\cdot) = \sum_{r \in L} c_r \chi^{(r)}(\cdot)$, wo die 'Vielfachheiten' c_r natürliche Zahlen sind. Ein interessantes Beispiel ist die Standarddarstellung: Die Vielfachheiten c_r sind hier die Dimensionen $c_r = n_r$ (was wir noch beweisen müssen); somit gilt $\sum_r n_r^2 = |G|$.

Wir werden sehen, dass die irreduziblen Spurfunktionen $\chi^{(r)}(\cdot)$ zueinander orthogonal sind, und dass der von ihnen aufgespannte Vektorraum der Raum der Klassenfunktionen ist; somit ist die Länge der Liste $|L|$ gleich der Anzahl $|K|$ der Konjugationsklassen in der Gruppe.

Wir wollen das noch etwas genauer sagen: Die Gruppenelemente a_1 und a_2 werden konjugierte Gruppenelemente genannt, wenn es ein $b \in G$ gibt, sodass $a_2 = b^{-1}a_1b$. Die Gruppe G zerfällt in Konjugationsklassen. Mit K bezeichnen wir die Menge der Konjugationsklassen. Die Mächtigkeit der Konjugationsklasse k wird mit h_k bezeichnet. Man bemerke: die Konjugationsklasse von $b \in G$ hat genau dann die Mächtigkeit 1, wenn b mit allen Gruppenelementen kommutiert; dies trifft beispielsweise für die Gruppeneins zu. Jede Spurfunktion ist auf jeder Konjugationsklasse konstant. Eine Funktion auf G , die auf jeder Konjugationsklasse konstant ist, nennen wir eine K -messbare Funktion oder eine Klassenfunktion. (Eine K -messbare Funktionen $f(\cdot)$ kann man natürlich auch als eine Funktion auf der Menge K auffassen.)

Die Menge aller komplexwertigen Funktionen $f(\cdot), g(\cdot), \dots$ auf G machen wir zu einem Hilbertraum $\mathcal{H}(G)$ durch das innere Produkt

$$\langle f | g \rangle = \frac{1}{|G|} \sum_{a \in G} \bar{f}(a) g(a).$$

Den Teilraum aller K -messbaren Funktionen bezeichnen wir mit $\mathcal{K}(G)$. Für seine Elemente $f(\cdot)$ gilt

$$\|f\|^2 = \frac{1}{|G|} \sum_{a \in G} |f(a)|^2 = \frac{1}{|G|} \sum_{k \in K} h_k |f(k)|^2.$$

Wir werden sehen: Die Charaktere der irreduziblen Darstellungen $\Phi^{(r)}$ bilden eine Orthonormalbasis $\{\chi^{(r)} : r \in L\}$ von $\mathcal{K}(G)$. Wir beweisen hier schon einmal

Satz

Die einzige Klassenfunktion, die auf allen $\chi^{(r)}(\cdot)$ senkrecht steht, ist die Nullfunktion.

Beweis Sei $\Phi = \{\varphi_g : g \in G\}$ eine beliebige Darstellung (über dem Vektorraum V), und sei $f(\cdot)$ eine Klassenfunktion. Der Endomorphismus $\varphi_f = \frac{1}{|G|} \sum_{g \in G} f^*(g) \varphi_g$ kommutiert dann mit jedem φ_a . In der Tat gilt

$$\varphi_a^{-1} \circ \varphi_f \circ \varphi_a = \frac{1}{|G|} \sum_{g \in G} f^*(g) \varphi_{a^{-1}ga} = \frac{1}{|G|} \sum_{h \in G} f^*(aha^{-1}) \varphi_h = \varphi_f$$

Für jede irreduzible Darstellung $\Phi^{(r)}$ ist dieser Endomorphismus (nach dem Schur'schen Lemma) ein Vielfaches der Identität $\varphi_f^{(r)} = d_f^{(r)} I$. Die Konstante bestimmen wir aus der Spur

$$d_f^{(r)} \cdot n_r = \frac{1}{|G|} \sum_{g \in G} f^*(g) \text{trace}(\varphi_g^{(r)}) = \frac{1}{|G|} \sum_{g \in G} f^*(g) \chi^{(r)}(g) = \langle f | \chi^{(r)} \rangle.$$

Wenn die Klassenfunktion $f(\cdot)$ auf allen $\chi^{(r)}$ senkrecht steht, dann ist φ_f der Nulloperator für alle Darstellungen Φ . Insbesondere gilt dann $\sum f^*(g) \tilde{\varphi}_g = 0$ für die Standarddarstellung $\tilde{\Phi}$, die bekanntlich dadurch gekennzeichnet ist, dass es im Darstellungsraum eine Basis $\{v_g : g \in G\}$ gibt mit $\tilde{\varphi}_h(v_g) = v_{hg}$ für alle $g, h \in G$. Die Anwendung dieses Nullooperators auf den Basisvektor v_e liefert den Nullvektor $0 = |G| \tilde{\varphi}_f(v_e) = \sum_{g \in G} f^*(g) \tilde{\varphi}_g(v_e) = \sum_{g \in G} f^*(g) v_g$. Die lineare Unabhängigkeit der v_g ergibt $f^*(\cdot) = 0$.

Exkurs: Multiplikative Funktionen

Beschäftigen wir uns zuerst mit den **multiplikativen Funktionen** auf der Gruppe G , das heisst mit den Spurfunktionen $f(\cdot)$ der eindimensionalen irreduziblen Darstellungen. ($f(gh) = f(g)f(h)$ für alle $g, h \in G$) Man nennt diese Funktionen $f(\cdot)$ auch (eindimensionale) Gruppencharaktere.

Hilfssatz Eine komplexwertige multiplikative Funktion $f(\cdot)$ auf der endlichen Gruppe G kann nur Werte vom Betrag 1 annehmen. Weiter gilt

$$f(e) = 1, \quad f(g^{-1}) = \bar{f}(g), \quad \sum_{g \in G} f(g) = 0, \quad \text{wenn } f \text{ nicht identisch } 1 \text{ ist.}$$

Beweis Nur die Aussage über die Quersumme bedarf eines Beweises.

Sei $s = \sum_{g \in G} f(g)$. Für jedes h gilt

$$s \cdot f(h) = \left(\sum_{g \in G} f(g) \right) f(h) = \sum_{g \in G} f(gh) = s;$$

wenn nicht alle $f(h)$ gleich 1 sind, dann folgt $s = 0$.

Satz

Die Menge aller multiplikativen Funktionen auf einer endlichen Gruppe ist eine Gruppe bzgl. der punktweisen Multiplikation.

Beweis Die Funktion $\equiv 1$, der ‘Einheitscharakter’, ist das Einselement. Mit $f(\cdot)$ ist auch $f^{-1}(\cdot)$ eine multiplikative Funktion. Wenn $f_1(\cdot)$ und $f_2(\cdot)$ multiplikative Funktionen sind, dann auch das punktweise Produkt $f_1 f_2(\cdot)$.

Eine Teilaussage des oben angekündigten Satzes können wir schnell beweisen:

Satz

Die multiplikativen Funktionen auf einer endlichen Gruppe bilden ein Orthonormalsystem im Hilbertraum $\mathcal{K}(G)$.

Beweis Der Normierungsfaktor sorgt dafür, dass die Funktion $\equiv 1$ und überhaupt jeder Gruppencharakter die Norm 1 hat. Wenn f_1 und f_2 verschiedene multiplikative Funktionen sind, dann ist $\bar{f}_1 \cdot f_2$ eine multiplikative Funktion, die nicht identisch 1 ist; ihre Quersumme verschwindet.

Das einfachste Beispiel ist die (additive) Gruppe der ganzen Zahlen modulo n . Die Gruppencharaktere bilden eine Orthonormalbasis. Die Darstellung der n -periodischen Funktionen in dieser Orthonormalbasis ist die diskrete Fourier-Transformation DFT.

Hinweis (Die einfachsten unendlichen Gruppen)

Manche der oben gemachten Aussagen über die eindimensionalen Darstellungen einer endlichen Gruppen haben ein Analogon bei nichtendlichen Gruppen. Unter einem Gruppencharakter auf der topologischen Gruppe G versteht man eine multiplikative Funktion, die ausserdem **beschränkt und stetig** ist. Im Hinblick auf die **Fourieranalyse** sind die folgenden ‘Beispiele’ bemerkenswert.

1. Die Gruppencharaktere für die Gruppe $\mathbb{R}/2\pi = \{t : t \in \mathbb{R}/2\pi\}$ sind die Funktionen $\dots e^{-2it}, e^{-it}, 1, e^{it}, e^{2it}, e^{3it}, \dots$. Es gibt keine weiteren multiplikativen 2π -periodischen Funktionen, die stetig und beschränkt sind.
2. Die Gruppencharaktere auf der Gruppe $\mathbb{Z} = \{n : n \in \mathbb{Z}\}$ sind die Folgen $e^{i\omega n}$ mit $\omega \in \mathbb{R}/2\pi$. Es gibt hier also eine kontinuierliche Schar von Gruppencharakteren.
3. Die Gruppencharaktere für die Gruppe $\mathbb{R} = \{t : t \in \mathbb{R}\}$ sind die Funktionen $e^{i\omega t}$ mit $\omega \in \mathbb{R}$. Es gibt keine weiteren stetigen beschränkten multiplikativen Funktionen auf der reellen Achse.

Interessant sind die **Analogien zum Satz von der Quersumme** und dem daraus abgeleiteten Satz von der Orthogonalität der Gruppencharaktere.

Im Beispiel 1 ist die Hilbertraumnorm durch das Integral des Absolutbetrags gegeben, und es gilt in der Tat

$$\frac{1}{2\pi} \int e^{-ikt} e^{ilt} dt = \begin{cases} 1 & \text{wenn } k = l \\ 0 & \text{wenn } k \neq l \end{cases}$$

Im Beispiel 2 ist die Quersumme keine konvergente Reihe. Es ist nicht mathematisch korrekt, wenn man (z. B. in Physikerkreisen) gern einmal schreibt: $\sum_{-\infty}^{\infty} e^{i\omega n} = 0$, falls

$\omega \neq 0$, oder auch

$$\sum_{n=-\infty}^{\infty} e^{-i\omega_1 n} e^{i\omega_2 n} = 0 \quad \text{für } \omega_1 \neq \omega_2.$$

Die unendliche Summe ist nicht wohldefiniert. Man bemerke aber: Wenn man die Mittelwerte über endliche Abschnitte (von $-N$ bis $+N$), dann findet die man die altbekannten Dirichlet-Kerne

$$\frac{1}{2N+1} (e^{-i\omega N} + \dots + e^{i\omega N}) = \frac{1}{2N+1} \frac{\sin((N + \frac{1}{2})\omega)}{\sin \frac{1}{2}\omega},$$

und die werden klein für $N \rightarrow \infty$, wenn nicht ω ein ganzzahliges Vielfaches von 2π ist.

Auf das Beispiel 3 bezieht sich ein (bei Physikern beliebtes) heuristisches Prinzip, welches sagt

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega_1 t} e^{i\omega_2 t} dt = \delta(\omega_1 - \omega_2).$$

Was damit zum Ausdruck gebracht werden soll, lernt man in der Theorie der Fourier-Integrale.

Weitere Analogien: Eine endlichdimensionale unitäre Darstellung der Gruppe \mathbb{R} ist offenbar nichts anderes als eine **einparametrische Gruppe** unitärer Matrizen $U_t = \exp(itH)$. Die Eigenvektoren der hermiteschen Matrix H spannen invariante Teilräume auf. Es existiert eine aus Eigenvektoren bestehende Orthonormalbasis. Dies ist der Satz von der Hauptachsentransformation. Jede endlichdimensionale Darstellung der Gruppe \mathbb{R} ist die direkte Summe von eindimensionalen Darstellungen.

Die Standarddarstellung einer unendlichen Gruppe ist eine unendlichdimensionale unitäre Darstellung dieser Gruppe. Ihre Zerlegung in irreduzible Darstellungen erfordert eine unendliche direkte Summe oder aber (im nichtkompakten Fall) sog. direktes Integral. Direkte Integrale zur Zerlegung unendlichdimensionaler Darstellungen sind der Gegenstand der Theorie der Spektralzerlegungen. Im Falle $G = \mathbb{R}$ geht es um die Spektralzerlegung der selbstadjungierten infinitesimalen Operatoren.

Die Spurfunktionen aller irreduziblen Darstellungen

Es sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste paarweise nichtäquivalenter irreduzibler Darstellungen. Wir studieren die dazugehörigen Spurfunktionen $\chi^{(r)}(\cdot)$ als Elemente des $|K|$ -dimensionalen Funktionenraums $\mathcal{K}(G)$. Die Werte der $\chi^{(r)}$ notieren wir als Spalten in einer $K \times L$ -Matrix; und diese Matrix $\mathcal{C}(G)$ nennen wir die **Charaktertafel** der endlichen Gruppe G .

(Warnung: In manchen Lehrbüchern wird die transponierte Matrix die Charaktertafel genannt; die Werte der verschiedenen irreduziblen Spurfunktionen stehen dann als Zeilen in der Charaktertafel; bei uns jedoch erscheinen sie, wie generell die Elemente des Funktionenraums $\mathcal{K}(G)$ als Spaltenvektoren.)

Wir haben gesehen, dass die Spurfunktionen, die zu den eindimensionalen irreduziblen Darstellungen gehören, die multiplikativen Funktionen sind. Auch für die irreduziblen Spurfunktionen der Dimension n_r gibt es eine Multiplikationsformel, nämlich

Multiplikationsformel

Sei $\Phi^{(r)} = \{\varphi_a^{(r)} : a \in G\}$ eine irreduzible Darstellung der endlichen Gruppe G über dem n_r -dimensionalen Vektorraum V_r . Für die Spurfunktion $\chi^{(r)}(\cdot)$ gilt dann

$$\frac{\chi^{(r)}(b)\chi^{(r)}(a)}{n_r} = \frac{1}{|G|} \sum_{g \in G} \chi^{(r)}(bgag^{-1}) \quad \text{für alle } a, b \in G.$$

Beweis

Für jeden Endomorphismus M von V_r ist der symmetrisierte Endomorphismus \bar{M} ein Vertauschungsoperator und daher (nach dem Lemma von Schur) ein Vielfaches der Identität. Da er dieselbe Spur hat wie M , haben wir

$$\bar{M} = \frac{1}{|G|} \sum_{g \in G} \varphi_{g^{-1}}^{(r)} M \varphi_g^{(r)} = \frac{1}{n_r} \text{trace } M \cdot I_{V_r}.$$

Für $M = \varphi_a^{(r)}$ erhalten wir

$$\frac{1}{|G|} \sum_{g \in G} \varphi_{g^{-1}}^{(r)} \varphi_a^{(r)} \varphi_g^{(r)} = \frac{1}{|G|} \sum_{g \in G} \varphi_{g^{-1}ag}^{(r)} = \frac{\chi^{(r)}(a)}{n_r} \cdot I_{V_r}.$$

Wenn wir mit $\varphi_b^{(r)}$ multiplizieren und die Spur bilden, erhalten wir die Multiplikationsformel.

Die explizite Konstruktion der kanonischen Zerlegung

Wir haben bereits darauf hingewiesen, dass sich jede Darstellung auf genau eine Weise als direkte Summe von paarweise disjunkten Darstellungen schreiben lässt. ('kanonische Zerlegung'). Wir zeigen, wie man die Projektionen auf die Summanden der kanonischen Zerlegung einer gegebenen Darstellung erhält, wenn man die entsprechenden irreduziblen Spurfunktionen kennt.

Satz (Disjunkte Projektionen mit Hilfe der $\chi^{(r)}$)

Es sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste irreduzibler Darstellungen der Gruppe G .

Zu festem $r \in L$ bezeichne $\chi^{(r)}(\cdot)$ die Spurfunktion und $n_r = \chi^{(r)}(e)$ den Rang, d.h. die Dimension des Darstellungsraums V_r .

Sei $\Psi = \{\psi_a : a \in G\}$ eine Darstellung über W , zerlegt in disjunkte Darstellungen

$$\Psi = \bigoplus_{s \in L} \Psi_s, \quad \text{wobei } \Psi_s \text{ auf } W_s \text{ wirkt :} \quad W = \bigoplus_{s \in L} W_s.$$

Für jedes $r \in L$ ist dann

$$\pi^{(r)} = \frac{n_r}{|G|} \sum_{a \in G} \chi^{(r)}(a^{-1}) \cdot \psi_a$$

die Projektion auf W_r entlang $W'_r = \bigoplus_{s \neq r} W_s$.

Vorbemerkungen zum Beweis

- Wenn man von Ψ zu einer äquivalenten Darstellung $S^{-1}\Psi S$ übergeht, dann transformiert sich der Endomorphismus $\pi^{(r)}$ in der richtigen Weise

$$\frac{n_r}{|G|} \sum_{a \in G} \chi^{(r)}(a^{-1}) \cdot S^{-1} \psi_a S = S^{-1} \pi^{(r)} S.$$

Es genügt also, unitäre Darstellungen Ψ zu studieren; der Satz besagt in diesem Fall, dass die $\pi^{(r)}$ die orthogonalen Projektionen auf die Summanden der kanonischen Zerlegung sind.

- Der direkte Summand Ψ_r ist äquivalent mit der direkten Summe von c_r Kopien der irreduziblen Darstellung $\Phi^{(r)}$. Wenn $\Phi^{(r)}$ durch unitäre $n_r \times n_r$ -Matrizen $U_g^{(r)}$ realisiert wird, dann wird Ψ_r (bei passender Basiswahl in W_r) durch Blockmatrizen mit c_r identischen Blöcken $U_g^{(r)}$ realisiert.
- Es genügt also zu zeigen: Wenn man die Blöcke $U_g^{(s)}$ mit den Gewichten $\frac{n_r}{|G|} \bar{\chi}^{(r)}(g)$ linear kombiniert, dann erhält man die Nullmatrix für jedes $s \neq r$; und im Fall $s = r$ ist

$$\frac{n_r}{|G|} \sum_g \bar{\chi}^{(r)}(g) \cdot U_g^{(r)} \quad \text{die Einheitsmatrix.}$$

Notation

Für jedes r aus der Liste L sei $\Phi^{(r)} = \{U_a^{(r)} : a \in G\}$ eine irreduzible Darstellung durch unitäre $J_r \times J_r$ -Matrizen ($|J_r| = n_r$). Den Eintrag in der Position (j, k) bezeichnen wir mit $u_{jk}^{(r)}(\cdot)$. Wir verstehen die Funktionen $u_{jk}^{(r)}(\cdot)$ als Elemente im Hilbertraum $\mathcal{H}(G)$. Die Spurfunktionen verstehen wir auch als Elemente im Teilraum $\mathcal{K}(G)$ der K -messbaren Funktionen.

Satz

Die Eintragsfunktionen $u_{jk}^{(r)}(\cdot)$ sind paarweise orthogonal; und es gilt für jedes r

$$\|u_{jk}^{(r)}(\cdot)\|^2 = \begin{cases} 0 & \text{wenn } j \neq k \\ \frac{1}{n_r} & \text{wenn } j = k \end{cases}$$

Bevor wir den Satz beweisen, bemerken wir einige Konsequenzen:

Der Satz sagt insbesondere, dass der von den Eintragsfunktionen $u_{jk}^{(r)}$ (zu einem festen r) aufgespannte Vektorraum \mathcal{H}_r die Dimension n_r^2 besitzt, und dass diese Vektorräume \mathcal{H}_r zueinander orthogonal sind. Daraus folgt $\sum_r n_r^2 \leq |G|$ sowie $|L| \leq |K|$. Da wir andererseits wissen, dass die $\chi^{(r)}$ den Raum der Klassenfunktionen aufspannen, können wir folgern, dass die irreduziblen Spurfunktionen eine Orthonormalbasis des Raums $\mathcal{K}(G)$ sind; es gilt $|L| = |K|$.

Vorbemerkungen zum Matrizenkalkül

- Die Einträge einer komplexen $J \times K$ -Matrix $A = (a_{jk})_{jk}$ kann man als Matrizenprodukte schreiben. Im Vektorraum $V = \mathbb{R}_{Sp}^J$ der J -Spalten x bezeichne x_j die j -te Einheitsspalte, (für $j \in J$); y_k sei die k -te Einheitsspalte im Vektorraum W der K -Spalten y . Wenn A eine $J \times K$ -Matrix ist, dann ist $a_{jk} = x_j^* A y_k$ der Eintrag in der Position (j, k) .
- Andererseits verstehen wir das Matrizenprodukt $M = y \cdot x^*$ als eine $K \times J$ -Matrix, die einen Homomorphismus $\varphi : V \rightarrow W$ darstellt. (Er bildet alle J -Spalten z in Vielfache von y ab; man nennt ihn einen Vektorraumhomomorphismus vom Rang 1).
Im Fall $K = J$ ist $y \cdot x^*$ eine quadratische Matrix mit der Spur $x^* y$. Wenn x ein Einheitsvektor ist, dann ist $x \cdot x^*$ die orthogonale Projektion auf den von x aufgespannten Vektorraum. (Man vergleiche auch Dirac's Notation.)

Beweis des Satzes

Seien nun $\{U_a^{(r)} : a \in G\}$ und $\{U_a^{(s)} : a \in G\}$ nichtäquivalente unitäre Darstellungen der Gruppe G . (Bemerke: die $U_a^{(r)}$ sind unitäre $J_r \times J_r$ -Matrizen mit $U_{a^{-1}}^{(r)} = U_a^{(r)*}$).
Wir betrachten für allgemeine J_r -Spalten x und allgemeine J_s -Spalten y die Funktionen

$$f_x^{(r)}(a) = x^* U_a^{(r)} x \quad \text{und} \quad f_y^{(s)}(a) = y^* U_a^{(s)} y .$$

sowie die Funktionen von der Bauart $g_{x,z}(a) = x^* U_a z$, die man als Linearkombinationen der Funktionen $f_{x+z}, f_{x-z}, f_{x+iz}, f_{x-iz}$ gewinnt. ('Polarisierung')

1) Nach dem Lemma von Schur verschwindet für jede $J_s \times J_r$ -Matrix M die symmetrisierte Matrix

$$\bar{M} = \frac{1}{|G|} \sum_{a \in G} U_{a^{-1}}^{(s)} M U_a^{(r)}$$

Für beliebige J_s -Spalten w und beliebige J_r -Spalten v gilt also

$$0 = w^* \cdot \bar{M} \cdot v = \frac{1}{|G|} \sum_{a \in G} w^* U_{a^{-1}}^{(s)} M U_a^{(r)} v$$

$$\text{und für } M = y \cdot x^*$$

$$0 = \frac{1}{|G|} \sum_{a \in G} w^* U_{a^{-1}}^{(s)} y \cdot x^* U_a^{(r)} v = \frac{1}{|G|} \sum_{a \in G} g_{w,y}^{(s)}(a^{-1}) \cdot g_{x,v}^{(r)}(a) = \frac{1}{|G|} \sum_{a \in G} \bar{g}_{y,w}^{(s)}(a) \cdot g_{x,v}^{(r)}(a) .$$

2) Im Fall $r = s$ ist (nach dem Lemma von Schur) der symmetrisierte Endomorphismus zu M ein Vielfaches der Einheitsmatrix. Da \bar{M} dieselbe Spur hat wie M , gilt

$$\bar{M} = \frac{1}{|G|} \sum_{a \in G} U_{a^{-1}}^{(r)} M U_a^{(r)} = \frac{1}{n_r} \text{trace } M \cdot I_{n_r} .$$

Für beliebige J_r -Spalten w, v gilt

$$w^* \cdot \bar{M} \cdot v = \frac{1}{|G|} \sum_{a \in G} w^* u_{(a^{-1})}^{(r)} M u_a^{(r)} v = \frac{1}{n_r} \text{trace } M \cdot (w^* v)$$

und für $M = z \cdot x^*$ mit der Spur $x^* z$,

$$\frac{1}{|G|} \sum_{a \in G} w^* u_{a^{-1}}^{(r)} z \cdot x^* u_a^{(r)} v = \frac{1}{|G|} \sum_{a \in G} \bar{g}_{z,w}^{(r)}(a) \cdot g_{x,v}^{(r)}(a) = \frac{1}{n_r} \cdot (x^* z)(w^* v) .$$

3) Die Orthonormalbasis $\{x_j : j \in J_r\}$ des n_r -dimensionalen Spaltenraums $V_r = \mathbb{C}_{Sp}^{J_r}$ liefert uns die n_r^2 Eintragsfunktionen $u_{jk}^{(r)}(\cdot) = g_{x_j, x_k}^{(r)}(\cdot)$. Diese sind paarweise orthogonal mit der Norm $(n_r)^{-1/2}$. Sie sind auch orthogonal zu den Eintragsfunktionen für die nichtäquivalenten unitären Darstellungen $\{u_a^{(s)} : a \in G\}$. Die Spurfunktionen $\chi^{(r)}(\cdot) = \text{trace } U^{(r)}(\cdot) = \sum_j u_{jj}^{(r)}(\cdot)$ für die verschiedenen r aus der Liste bilden ein Orthonormalsystem im Raum $\mathcal{K}(G)$.

Corollar

Sei $\chi^{(r)}(\cdot)$ die Spurfunktion der irreduziblen Darstellung $\Phi^{(r)}$, und sei $\chi(\cdot)$ die Spurfunktion einer Darstellung Ψ mit der irreduziblen Darstellung $\Psi = \bigoplus_{s \in L} c_s \Phi^{(s)}$. Es gilt dann

$$\langle \chi^{(r)} | \chi \rangle = \frac{1}{|G|} \sum_{a \in G} \bar{\chi}^{(r)}(a) \cdot \chi(a) = c_r .$$

Für die Standarddarstellung $\Psi = \tilde{\Phi}$ (auch die reguläre Darstellung genannt) ergeben sich die Vielfachheiten $\tilde{c}_r = n_r$ für alle r aus der Liste L .

Beweis

Die Spurfunktion der Standarddarstellung hat nur in der Gruppeneins einen von 0 verschiedenen Wert, und zwar den Wert $|G|$; andererseits ist $\chi^{(r)}(e)$ die Dimension der Darstellung $\Phi^{(r)}$. Es gilt

$$\chi(\cdot) = \sum_{s \in L} c_s \chi^{(s)}(\cdot) \implies \langle \chi^{(r)} | \sum_{s \in L} c_s \chi^{(s)} \rangle = c_r .$$

$$\tilde{c}_r = \frac{1}{|G|} \sum_{a \in G} \bar{\chi}^{(r)}(a) \cdot \tilde{\chi}(a) = \frac{1}{|G|} \bar{\chi}^{(r)}(e) \cdot |G| = n_r .$$

Da die Dimension von $\bigoplus_{r \in L} n_r \Phi^{(r)}$ gleich $\sum_{r \in L} n_r^2$ ist, haben wir $\sum_{r \in L} n_r^2 = |G|$.

Konkrete Charaktertafeln

Wir sind oben davon ausgegangen, dass wir irreduzible Spurfunktionen zur gegebenen Gruppe (oder gar die ganze Charaktertafel) kennen. Wenn wir sie nicht kennen, wissen wir wenigstens, nach was wir suchen müssen. Für die folgenden recht kleinen Gruppen können wir die Charaktertafel mit Hilfe der bisherigen Einsichten erraten. Wenn man

ansonsten die Charaktertafel einer speziellen Gruppe wirklich berechnen will, dann kommt es sehr wesentlich darauf an, wie einem die Gruppe gegeben ist, und was man so alles weiss über die Gruppe. Wir werden unten sehen, wie man die irreduziblen Spurfunktionen als die Lösungen von Eigenwertproblemen gewinnen kann, wenn man die Gruppentafel gut kennt.

Beispiel 1 (Tetraedergruppe)

Die Menge aller Drehungen, welche ein reguläres Tetraeder im dreidimensionalen Anschauungsraum in sich überführen, ist eine Gruppe G mit 12 Elementen. In der Tat: Um jede Höhe kann man um 120° im positiven oder im negativen Sinn drehen. Um jede Kante kann man klappen; dabei führen aber gegenüberliegende Kanten zu gleichen Klappung. Darüber hinaus gibt es nur noch die Identität.

Da es vier Höhen gibt, haben wir $4 \cdot 2 = 8$ Drehungen; da es 6 Kanten gibt haben wir 3 Klappungen. Alle Klappungen sind zueinander konjugiert; die positiven Drehungen sind zueinander konjugiert, ebenso die negativen; die positiven Drehungen sind aber nicht konjugiert zu den negativen Drehungen. Es gibt daher 4 Konjugationsklassen. Wenn wir $|G| - 1 = 11$ als Summe von drei Quadraten darstellen wollen, dann gibt es nur die Möglichkeit $11 = 1^2 + 1^2 + 3^2$. Unsere Gruppe G hat also neben der trivialen Darstellung noch zwei weitere eindimensionale irreduzible Darstellungen und eine dreidimensionale Darstellung. Die eindimensionalen Darstellungen sind durch multiplikative Funktionen gegeben; dazu benötigen wir eine dritte Einheitswurzel ϵ eine. Die Charaktertafel hat die Gestalt

$$\begin{array}{l} e(1) \\ +\text{Dr}(4) \\ -\text{Dr}(4) \\ \text{Kl}(3) \end{array} \left| \begin{array}{cccc} 1 & 1 & 1 & 3 \\ 1 & \epsilon & \epsilon^2 & 0 \\ 1 & \epsilon^2 & \epsilon & 0 \\ 1 & 1 & 1 & -1 \end{array} \right.$$

(Die letzte Spalte ergibt sich aus der Orthonormalität der Spalten mit Gewichtungen) Noch einfacher ergibt sie sich aus der Orthonormalität der Zeilen der Charaktertafel, die wir unten beweisen werden.

Beispiel 2 (Die symmetrische Gruppe \mathcal{S}_3)

Die symmetrische Gruppe \mathcal{S}_3 hat 6 Elemente, drei Transpositionen (nennen wir sie $\mathbf{a}, \mathbf{b}, \mathbf{c}$), zwei Dreierzyklen (nennen wir sie \mathbf{l}, \mathbf{m}) und die Gruppeneins \mathbf{e} . (Dies ist die Unterscheidung nach den drei Konjugationsklassen.) Neben dem trivialen Charakter $\equiv 1$ haben wir noch die ‘Signatur’ als eine multiplikative Funktion; die Signatur hat den Wert -1 in den Transpositionen und sonst den Wert $+1$. Die Formel $|G| = \sum n_i^2$ zeigt, dass es genau eine weitere irreduzible Darstellung gibt, und zwar eine zweidimensionale. Die Charaktertafel hat die Gestalt

$$\begin{array}{l} e(1) \\ zy(2) \\ tr(3) \end{array} \left| \begin{array}{ccc} 1 & 1 & 2 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \end{array} \right.$$

Wir haben oben eine Realisierung der zweidimensionalen Darstellung durch unitäre Matrizen konkret angegeben. Die Einträge $\sqrt{2}f_{11}, \sqrt{2}f_{12}, \sqrt{2}f_{21}, \sqrt{2}f_{22}$ bilden ein Orthonormalsystem und zusammen mit der Signatur und der Einsfunktion eine Orthonormalbasis im Hilbertraum Raum aller Funktionen auf \mathcal{S}_3 .

Faltung und die reduzierte Gruppenalgebra

Wir betrachten jetzt Gewichtungen auf der Gruppe G . Die Bausteine sind die sog. δ -Maße δ_g . Wir benützen hier aber auch die Schreibweise der formalen Linearkombinationen.

Definition

- Die Elemente der Gruppenalgebra sind die formalen Linearkombinationen

$$\mu = \sum_{a \in G} p(a) \cdot a = \sum_{a \in G} p(a) \cdot \delta_a$$

- Die **Multiplikation** oder ‘**Faltung**’ von μ mit $\nu = \sum_{a \in G} q(a) \cdot a$ liefert

$$\mu * \nu = \sum_{a, b \in G} p(a) q(b) \cdot ab = \sum_{a \in G} r(a) \cdot a \quad \text{mit} \quad r(a) = \sum_{a_1 a_2 = a} p(a_1) q(a_2) = \sum_{g \in G} p(g) q(g^{-1}a)$$

- Die Elemente der **reduzierten Gruppenalgebra** sind diejenigen μ , die mit allen a kommutieren.

Bemerkung: Der Vektorraum aller Gewichtungen sollte als der Dualraum zum Vektorraum der Funktionen auf der Gruppe gesehen werden. Wenn k eine Konjugationsklasse in G ist, $k = \{b : b \in k\}$, dann ist das arithmetische Mittel $\rho_k = \frac{1}{h(k)} \sum_{b \in k} b$ ein Element der reduzierten Gruppenalgebra. Jedes Element der reduzierten Gruppenalgebra ist als eine Linearkombination der ausgeglichenen Gewichtungen ρ_k zu gewinnen; die Elemente der reduzierten Gruppenalgebra sind diejenigen Gewichtungen, welche konjugierte Gruppenelemente mit demselben Gewicht belegen. Die Mittelung über Konjugationsklassen ist eine surjektive Projektion. Die Faltung macht den Vektorraum der Gewichtungen zu einer Algebra. Die Faltung zweier Elemente der reduzierten Gruppenalgebra liefert ein Element der Gruppenalgebra, welches mit allen a kommutiert. Die reduzierte Gruppenalgebra ist also eine Teilalgebra; die Projektion auf die reduzierte Gruppenalgebra ist mit dem Faltungsprodukt verträglich.

Die Strukturkoeffizienten

Wenn i und j Konjugationsklassen sind, dann gilt

$$\rho_i * \rho_j = \sum_{k \in K} c_{ij}^k \rho_k$$

mit gewissen Koeffizienten c_{ij}^k , die man die **Strukturkoeffizienten** der reduzierten Gruppenalgebra nennt.

Beispiel

Es ist eine einfache Übungsaufgabe, die Strukturkoeffizienten für die symmetrische Gruppe \mathcal{S}_3 zu berechnen. Man bemerke z. B. (Wir schreiben die Permutationen in Zykelschreibweise)

$$\begin{aligned} \left(\frac{1}{3}(\mathbf{a}, \mathbf{b}) + \frac{1}{3}(\mathbf{b}, \mathbf{c}) + \frac{1}{3}(\mathbf{c}, \mathbf{a})\right) * \left(\frac{1}{3}(\mathbf{a}, \mathbf{b}) + \frac{1}{3}(\mathbf{b}, \mathbf{c}) + \frac{1}{3}(\mathbf{c}, \mathbf{a})\right) &= \frac{1}{3}\mathbf{e} + \frac{1}{3}(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \frac{1}{3}(\mathbf{a}, \mathbf{c}, \mathbf{b}) \\ \left(\frac{1}{3}(\mathbf{a}, \mathbf{b}) + \frac{1}{3}(\mathbf{b}, \mathbf{c}) + \frac{1}{3}(\mathbf{c}, \mathbf{a})\right) * \left(\frac{1}{2}(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \frac{1}{2}(\mathbf{a}, \mathbf{c}, \mathbf{b})\right) &= \frac{1}{3}(\mathbf{a}, \mathbf{b}) + \frac{1}{3}(\mathbf{a}, \mathbf{b}) + \frac{1}{3}(\mathbf{a}, \mathbf{b}) \\ \left(\frac{1}{2}(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \frac{1}{2}(\mathbf{a}, \mathbf{c}, \mathbf{b})\right) * \left(\frac{1}{2}(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \frac{1}{2}(\mathbf{a}, \mathbf{c}, \mathbf{b})\right) &= \frac{1}{2}\mathbf{e} + \frac{1}{4}(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \frac{1}{4}(\mathbf{a}, \mathbf{c}, \mathbf{b}) \end{aligned}$$

Eine besondere Rolle spielt die Konjugationsklasse der Gruppeneins, $k_1 = \{\mathbf{e}\}$. Das dazugehörige ρ bezeichnen wir mit \mathbf{e} oder auch mit ρ_1 ; es ist das Einselement der Gruppenalgebra. Die Strukturkoeffizienten c_{ij}^1 können wir leicht berechnen.

$$\rho_i * \rho_j = c_{ij}^1 \mathbf{e} + \sum_{k \neq k_1} c_{ij}^k \rho_k \quad \text{mit} \quad c_{ij}^1 = \begin{cases} \frac{1}{h(i)} & \text{wenn } j = \tilde{i} \\ 0 & \text{wenn } j \neq \tilde{i} \end{cases}$$

Zum Beweis müssen wir zwei Fälle unterscheiden: Wenn $i = \{\mathbf{b} : \mathbf{b} \in i\}$ eine Konjugationsklasse ist, dann ist die Menge $\tilde{i} = \{\mathbf{b}^{-1} : \mathbf{b} \in i\}$ ebenfalls eine Konjugationsklasse; die übrigen Konjugationsklassen j enthalten kein Gruppenelement, welches zu einem der $\mathbf{b} \in i$ invers ist. Das ergibt

Satz

Es sei $\{\Phi^{(r)} : r \in L\}$ eine vollständige Liste der irreduziblen Darstellungen von G. Der Wert des r -ten Charakters in der k -ten Konjugationsklasse (mit der Mächtigkeit $h(k)$) sei $\chi^r(k)$. Es gilt dann

$$\frac{1}{|G|} \sum_r \bar{\chi}^r(l) \chi^r(k) = \begin{cases} 0 & \text{wenn } k \neq l \\ \frac{1}{h(k)} & \text{wenn } k = l \end{cases}.$$

Beweis

1. Wir mitteln die Endomorphismen φ_b^r über die \mathbf{b} in der Konjugationsklasse k und erhalten einen Operator ψ_k^r , der mit allen φ_a^r kommutiert und daher ein Vielfaches der Identität I^r ist

$$\begin{aligned} \psi_k^r &= \frac{1}{h(k)} \sum_b \varphi_b^r = \varphi_a \psi_k^r \varphi_a^{-1} = z(k, r) \cdot I^r \\ \text{mit } z(k, r) &= \frac{1}{n_r} \cdot \chi^{(r)}(k) . \end{aligned}$$

Die Konstanten ergeben sich, wenn wir die Spur bilden, da nämlich alle φ_b^r dieselbe Spur $\chi^r(k)$ haben (und die Identität I_r die Spur n_r hat).

2. Die Darstellung $\Phi^{(r)}$ liefert eine mit Addition und Multiplikation ('Faltung' bzw. Hintereinanderschalten) verträgliche Abbildung der Gruppenalgebra in die Algebra der Endomorphismen von V_r . Der Endomorphismus ψ_k^r ist das Bild von ρ_k . Daher gilt

$$\begin{aligned}\psi_i^r \psi_j^r &= \sum_{k \in K} c_{ij}^k \psi_k^r; & z(i, r) z(j, r) &= \sum_{k \in K} c_{ij}^k z(k, r); \\ \chi^{(r)}(i) \cdot \chi^{(r)}(j) &= n_r \sum_{k \in K} c_{ij}^k \chi^{(r)}(k) \quad \text{für alle } i, j \in K, \quad r \in L\end{aligned}$$

3. Wir summieren über alle r .

$\sum_r n_r \chi^{(r)}(\cdot)$ ist die Spurfunktion der Standarddarstellung (mit dem Wert $|G|$ im Einselement und $= 0$ sonst). Also ergibt sich

$$\sum_{r \in L} \chi^{(r)}(i) \cdot \chi^{(r)}(j) = c_{ij}^1 |G|$$

Die obige Formel für den Koeffizienten c_{ij}^1 liefert die Behauptung; denn für alle $b \in G$ gilt $\chi^{(r)}(b^{-1}) = \bar{\chi}^{(r)}(b)$.

Die Berechnung von $\chi^{(r)}(\cdot)$ als Lösung eines Eigenwertproblems

Wenn man die Strukturkoeffizienten c_{ij}^k der reduzierten Gruppenalgebra aus der 'Gruppentafel' abgelesen hat, dann kann man die Berechnung der irreduziblen Spurfunktionen als ein Eigenwertproblem formulieren.

Es sei $\chi^{(r)}(\cdot)$ die Spurfunktion einer irreduziblen Darstellung vom Rang n_r und es sei $z = z^{(r)}$ die K -Spalte mit den Einträgen $z(k, r) = \frac{1}{n_r} \cdot \chi^{(r)}(k)$.

Für ein festes i, r sei $\lambda = z(i, r)$ und A die $K \times K$ -Matrix mit den Einträgen $a_{jk} = c_{ij}^k$. Wir haben dann

$$A \cdot z = \lambda z \quad \text{wegen} \quad z(i, r) z(j, r) = \sum_{k \in K} c_{ij}^k z(k, r);$$

Für jedes r ist also $z^{(r)}$ ein Eigenvektor der bekannten Matrix A , wobei sich seine 'Normierung' durch den Wert λ in der Position i ergibt.

IV.10 Unitäre und hermitische Matrizen

Die unitären $n \times n$ -Matrizen sind uns immer wieder begegnet. Um den Überblick zu sichern, rekapitulieren wir:

- a) Seien $\{u_i : i \in I\}$ und $\{v_j : j \in J\}$ ONBasen in einem n -dimensionalen Hilbertraum $(V, \|\cdot\|)$. Die Matrix des Basiswechsels ist dann eine unitäre Matrix.

$$v_j = \sum u_i \cdot a_j^i, \quad A^* A = E_{J \times J}.$$

Die Rücktransformation wird von der $J \times J$ -Matrix $A^* = A^{-1}$ geleistet. $AA^* = E_{I \times I}$.

- b) Sei $\{v_j : j \in J\}$ eine ONBasis von $(V, \|\cdot\|)$ und $\{w_k : k \in K\}$ eine ONBasis von $(W, \|\cdot\|)$. Eine lineare Abbildung

$$\varphi : (V, \|\cdot\|) \rightarrow (W, \|\cdot\|)$$

heißt eine **Isometrie**, oder eine isometrische Abbildung, wenn gilt

$$\begin{aligned} \|\varphi(v)\| &= \|v\| \quad \text{für alle } v \in V, \quad \text{oder (damit äquivalent)} \\ \langle \varphi(v_1) | \varphi(v_2) \rangle &= \langle v_1 | v_2 \rangle \quad \text{für alle } v_1, v_2 \in V. \end{aligned}$$

(Für den Beweis der Invarianz des Skalarprodukts benütze man die Polarisierungsgleichung.)

Wenn man die Isometrie φ mit den Basen darstellt

$$\varphi(v_j) = \sum w_k \cdot b_j^k,$$

dann gilt für die $K \times J$ -Matrix $B : B^* \cdot B = E_{J \times J}$.

Wenn V und W dieselbe Dimension haben, dann nennt man φ einen **unitären Operator**. Die darstellende Matrix B ist unitär.

- c) Betrachten wir den Spezialfall $V = W$.

Die Menge der Isometrien ist eine Transformationsgruppe. Wenn man eine Basis $\{u_i : i \in I\}$ auszeichnet, dann entsprechen die Isometrien den unitären $I \times I$ -Matrizen. Sei $x(\cdot)$ die Koordinatenabbildung zu $\{u_i : i \in I\}$.

$$\begin{array}{ccccc} u & \xrightarrow{\varphi} & u & \xrightarrow{\psi} & u \\ \downarrow x & & \downarrow x & & \downarrow x \\ \mathbb{C}_{Sp}^I & \xrightarrow{A} & \mathbb{C}_{Sp}^I & \xrightarrow{B} & \mathbb{C}_{Sp}^I \end{array}$$

Die Produktmatrix $C = B \cdot A$ ist die darstellende Matrix für die zusammengesetzte Abbildung $\psi(\varphi(\cdot))$. (Beachte die Reihenfolge $x \xrightarrow{\varphi} Ax \xrightarrow{\psi} B(Ax) = (BA)x$.)

Hinweise :

Die Menge der unitären $n \times n$ -Matrizen $U(n, \mathbb{C})$ ist einerseits eine Gruppe und andererseits eine reelle Mannigfaltigkeit. Was eine Mannigfaltigkeit ist, werden wir im zweiten Semester lernen. Im vorliegenden Fall ist der Mannigfaltigkeitscharakter aber nicht schwer zu verstehen. Wir können $U(n, \mathbb{C})$ als eine n^2 -dimensionale „Fläche“ im $2n^2$ -dimensionalen Raum \mathbb{R}^{2n^2} begreifen. $U^*U = E_{n \times n}$ liefert offenbar n^2 Bedingungen an die Real- und Imaginärteile einer komplexen $n \times n$ -Matrix. Jede der Bedingungen sondert eine Hyperfläche aus dem \mathbb{R}^{2n^2} aus. Diese Hyperflächen schneiden sich in einer regulären Weise; der Durchschnitt hat die Dimension $2n^2 - n^2 = n^2$. Um zu sehen, dass alle Schnittpunkte in der Tat reguläre Punkte sind, muss man sich vergewissern, dass die Gradienten der definierenden Funktionen linear unabhängig sind. Darüber mehr im zweiten Semester.

Es gibt weitere interessante Plausibilitätsbetrachtungen, die auf die Dimension n^2 führen. Man kann z.B. Beziehungen zwischen den unitären und den hermiteschen Matrizen herstellen, die lokal umkehrbar sind. Die Cayley-Transformation (die wir in den Übungen studiert haben) ist eine interessante Möglichkeit. Ihre unendlichdimensionale Variante war in der Geschichte der Funktionalanalysis bedeutsam.

Wir skizzieren noch einen anderen Zusammenhang, der in der Theorie der Lie-Gruppen und Lie-Algebren ausführlich behandelt wird. Man zeigt da: Die Menge der schiefhermiteschen $n \times n$ -Matrizen ist die Lie-Algebra zur Lie-Gruppe $U(n, \mathbb{C})$. Inwiefern man in der Menge der schiefhermiteschen Matrizen eine „Multiplikation“ hat, werden wir später sehen.

Die Exponentialabbildung**Konstruktion :**

Sei A eine komplexe $n \times n$ -Matrix. Dann ist die Schar der Matrizen

$$M_t = \exp(t \cdot A)$$

eine (einparametrische) Matrixgruppe. Es gilt nämlich $M_0 = E$, $M_s \cdot M_t = M_{s+t}$ für alle s, t .

a) Wenn $\text{trace}(A) = 0$, dann gilt $\det(M_t) = 1$ für alle t .

b) Wenn H hermitisch ist, dann sind die Matrizen

$$U_t = \exp(itH)$$

eine einparametrische Gruppe von unitären Matrizen.

Beweis

$$1. \quad M_t = E + t \cdot A + \frac{1}{2!}t^2A^2 + \dots, \quad M_s \cdot M_t = M_{s+t}$$

$$\frac{1}{h}(M_h - E) \xrightarrow{h \rightarrow 0} A$$

Man nennt A den infinitesimalen Operator der Matrixgruppe $\{M_t : t \in \mathbb{R}\}$.

2. Die Determinante ist multiplikativ. Für die stetige komplexwertige Funktion $d(t) = \det(M_t)$ gilt also $d(s+t) = d(s) \cdot d(t)$ mit $d(0) = 1$.

Es gibt nicht viele stetige Funktionen mit dieser Eigenschaft. Es existiert eine Zahl α , sodass $d(t) = \exp(\alpha \cdot t)$. Wir identifizieren α durch den Grenzwert für $t \rightarrow 0$.

$$\begin{aligned}\det M_t &\approx \det(E + tA) \approx 1 + t \cdot \text{trace} A \quad \text{für } t \rightarrow 0 ; \\ \det M_t &= \exp(t \cdot \text{trace} A) .\end{aligned}$$

$$3. M_t^* = E + t \cdot A^* + \frac{1}{2!}t^2 \cdot (A^*)^2 + \dots .$$

Wenn A antihermitisch ist, $A^* = -A$, dann haben wir

$$M_t^* = M_{-t} \quad \text{für alle } t \in \mathbb{R} . \quad M_t \cdot M_t^* = E . \quad \text{q.e.d.}$$

Beispiel

Sei $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ die zweite Pauli-Matrix.

σ_2 ist eine spurlose hermitische Matrix. Die Matrizen

$$U_t = \exp(it\sigma_2) , \quad t \in \mathbb{R}$$

bilden eine Gruppe spezieller unitärer Matrizen. Es gilt $\sigma_2^2 = \sigma_0$ und

$$\begin{aligned}U_t &= \sigma_0 + it \cdot \sigma_2 + \frac{1}{2!}(it)^2 \cdot \sigma_2^2 + \frac{1}{3!}(it)^3 \cdot \sigma_2 + \dots \\ &= \sigma_0 \left[1 - \frac{1}{2!}t^2 + \frac{1}{4!}t^4 - \dots \right] + i\sigma_2 \cdot \left[t - \frac{1}{3!}t^3 + \frac{1}{5!}t^5 - \dots \right] \\ &= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} .\end{aligned}$$

Für die erste und die dritte Pauli-Matrix $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ erhalten wir

$$U_t = \begin{pmatrix} \cos t & i \sin t \\ i \sin t & \cos t \end{pmatrix} \quad \text{bzw.} \quad U_t = \begin{pmatrix} \exp(it) & 0 \\ 0 & \exp(-it) \end{pmatrix} .$$

Weitere einparametrische Untergruppen der $SU(2, \mathbb{C})$ erhalten wir für beliebige spurlose hermitische Matrizen $h = \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3$, (α_j reell).

Wenn $\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$, dann gilt $h^2 = \sigma_0$ und daher

$$\begin{aligned}U_t &= \sigma_0 + it h + \frac{1}{2!}(it)^2 \cdot \sigma_0 + \frac{1}{3!}(it)^3 \cdot h + \dots \\ &= \sigma_0 \cdot \cos t + i h \cdot \sin t .\end{aligned}$$

Man kann hier auch direkt nachrechnen

$$\begin{aligned}U_t \cdot U_s &= (\sigma_0 \cdot \cos t + i h \cdot \sin t) \cdot (\sigma_0 \cdot \cos s + i h \cdot \sin s) \\ &= \sigma_0 [\cos t \cdot \cos s - \sin t \sin s] + i h [\sin t \cdot \cos s + \cos t \cdot \sin s] \\ &= \sigma_0 \cdot \cos(t+s) + i h \sin(t+s) = U_{t+s} .\end{aligned}$$

e unitäre Diagonalisierung der hermiteschen 2×2 -Matrizen

Wir haben in V.2 gezeigt: Zu jeder hermiteschen $n \times n$ -Matrix H gibt es eine unitäre Matrix U , sodass $U^* \cdot H \cdot U = \Lambda$ (Diagonalmatrix). Die Einträge der Diagonalmatrix sind die Eigenwerte λ_j ; die Spalten von U sind die normierten Eigenvektoren.

$$HU = U \cdot \Lambda, \quad U^* \cdot U = E.$$

Wir werden ein Iterationsverfahren zur approximativen Bestimmung von U kennen lernen. Es stammt von Jacobi (1804-1851). Zuvor wollen wir den Fall $n = 2$ explizit lösen.

Wenn die unitäre 2×2 -Matrix U die Diagonalisierung der hermiteschen 2×2 -Matrix h leistet, dann leistet sie auch die Diagonalisierung von $\alpha \cdot \sigma_0 + \beta h$ für α, β beliebig reell. Es genügt also die h mit $\text{trace}(h) = 0$ und $\det h = -1$ zu untersuchen. Das sind die hermiteschen 2×2 -Matrizen h mit den Eigenwerten ± 1 . Für jedes solche h gilt $h^2 = \sigma_0$; denn h^2 ist eine hermitesche Matrix mit dem doppelten Eigenwert $+1$. (Aus der unitären Diagonalisierung $U^* \cdot h^2 \cdot U = \sigma_0$ schließen wir $h^2 = \sigma_0$.)

Wir betrachten den Raum aller spurlosen hermiteschen 2×2 -Matrizen als einen dreidimensionalen euklidischen Raum $(\mathbb{M}_0, \|\cdot\|)$, indem wir die Pauli-Matrizen $\sigma_1, \sigma_2, \sigma_3$ als eine Orthogonalbasis auszeichnen.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$h = \alpha_1 \sigma_1 + \alpha_2 \sigma_2 + \alpha_3 \sigma_3 = \begin{pmatrix} \alpha_3 & b \\ \bar{b} & -\alpha_3 \end{pmatrix} \quad \text{mit } b = \alpha_1 - i\alpha_2$$

Es gilt

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = -\det h = -[-\alpha_3^2 - |b|^2] = \alpha_3^2 + |b|^2.$$

Hinweis

Die Norm im euklidischen Raum $(\mathbb{M}_0, \|\cdot\|)$ kann man auf verschiedene Weisen ausdrücken

$$\|h\|^2 = -\det h, \quad \|h\|^2 = \frac{1}{2} \text{trace}(h^2), \quad \|h\|^2 \cdot \sigma_0 = h^2.$$

Die erste Formel wird uns interessieren, wenn wir uns später mit dem Minkowski-Raum befassen. Die dritte Formel werden wir vertieft verstehen, wenn wir uns mit dem Lie-Produkt befassen. Hier interessiert uns vor allem die mittlere Formel. Die „Hilbert-Schmidt-Norm“, die den Raum aller $n \times n$ -Matrizen zu einem Hilbertraum macht, kann als Verallgemeinerung gesehen werden, wie wir unten sehen werden. $\|A\|^2 := \text{trace}(A^*A)$.

Jetzt wieder zum Fall $n = 2$.

Satz :

Sei $\mathbf{h} \in \mathbb{M}_0$, sei also \mathbf{h} eine spurlose hermitische 2×2 -Matrix

$$\mathbf{h} = \begin{pmatrix} \alpha & \mathbf{b} \\ \bar{\mathbf{b}} & -\alpha \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R}, \quad \mathbf{b} \in \mathbb{C}.$$

Die Eigenwerte sind $\pm\lambda = \pm\sqrt{\alpha^2 + |\mathbf{b}|^2}$. Eigenvektoren sind die Spalten der Matrix

$$\mathbf{N} = \begin{pmatrix} -\mathbf{b} & \mathbf{b} \\ \lambda + \alpha & \lambda - \alpha \end{pmatrix} \quad (\text{“normal modes“})$$

Wenn man diese Spalten auf die Länge 1 normiert, dann erhält man die unitäre Matrix \mathbf{U} , welche \mathbf{h} diagonalisiert

$$\mathbf{U}^* \cdot \mathbf{h} \cdot \mathbf{U} = \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

Beweis

Wir rechnen einfach nach

$$\begin{aligned} \mathbf{h} \cdot \begin{pmatrix} -\mathbf{b} \\ \lambda + \alpha \end{pmatrix} &= \begin{pmatrix} \alpha & \mathbf{b} \\ \bar{\mathbf{b}} & -\alpha \end{pmatrix} \begin{pmatrix} -\mathbf{b} \\ \lambda + \alpha \end{pmatrix} = \begin{pmatrix} \lambda \cdot \mathbf{b} \\ -|\mathbf{b}|^2 - \alpha^2 - \alpha\lambda \end{pmatrix} = -\lambda \begin{pmatrix} -\mathbf{b} \\ \lambda + \alpha \end{pmatrix} \\ \mathbf{h} \cdot \begin{pmatrix} +\mathbf{b} \\ \lambda - \alpha \end{pmatrix} &= \begin{pmatrix} \alpha & \mathbf{b} \\ \bar{\mathbf{b}} & -\alpha \end{pmatrix} \begin{pmatrix} +\mathbf{b} \\ \lambda - \alpha \end{pmatrix} = \lambda \begin{pmatrix} +\mathbf{b} \\ \lambda - \alpha \end{pmatrix} \end{aligned}$$

Diese Eigenvektoren stehen aufeinander senkrecht

$$\mathbf{N}^* \cdot \mathbf{N} = \mathbf{D}^2 \quad (\text{Diagonalmatrix})$$

Für $\mathbf{U} := \mathbf{N} \cdot \mathbf{D}^{-1}$ haben wir $\mathbf{U}^* \mathbf{U} = \mathbf{E}$

$$\begin{aligned} \mathbf{h} \cdot \mathbf{N} &= \mathbf{N} \cdot \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix} \\ \mathbf{h} \cdot \mathbf{N} \mathbf{D}^{-1} &= \mathbf{N} \cdot \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix} \mathbf{D}^{-1} = \mathbf{N} \mathbf{D}^{-1} \cdot \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix} \\ \mathbf{U}^{-1} \cdot \mathbf{h} \cdot \mathbf{U} &= \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad \text{q.e.d.} \end{aligned}$$

Die Elemente der Diagonalmatrix \mathbf{D}^{-1} sind $[|\mathbf{b}|^2 + (\lambda + \alpha)^2]^{1/2}$ und $[|\mathbf{b}|^2 + (\lambda - \alpha)^2]^{1/2}$. Man kann sie etwas knapper schreiben; denn

$$\begin{aligned} |\mathbf{b}|^2 + (\lambda + \alpha)^2 &= |\mathbf{b}|^2 + \alpha^2 + \lambda^2 + 2\alpha\lambda = 2\lambda(\lambda + \alpha) \\ |\mathbf{b}|^2 + (\lambda - \alpha)^2 &= 2\lambda \cdot (\lambda - \alpha). \end{aligned}$$

Spezialfall $\alpha = 0$:

$$\mathbf{h} = \begin{pmatrix} 0 & \mathbf{b} \\ \bar{\mathbf{b}} & 0 \end{pmatrix}, \quad \lambda = |\mathbf{b}|$$

$$\mathbf{N} = \begin{pmatrix} -\mathbf{b} & \mathbf{b} \\ |\mathbf{b}| & |\mathbf{b}| \end{pmatrix} \quad D^2 = \mathbf{N}^* \cdot \mathbf{N} = 2|\mathbf{b}|^2 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Mit $\mathbf{b} = |\mathbf{b}| \cdot e^{i\varphi}$ haben wir

$$\mathbf{u} = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\varphi} & e^{i\varphi} \\ 1 & 1 \end{pmatrix} \quad \mathbf{u}^* \cdot \mathbf{h} \cdot \mathbf{u} = \begin{pmatrix} -|\mathbf{b}| & 0 \\ 0 & |\mathbf{b}| \end{pmatrix}.$$

Zahlenbeispiel

$$\mathbf{H} = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} \quad \lambda = \sqrt{3^2 + 4^2} = 5$$

$$\mathbf{N} = \begin{pmatrix} -4 & 4 \\ 8 & 2 \end{pmatrix}$$

$$\mathbf{u} = \begin{pmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}$$

$$\frac{1}{5} \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -5 & 0 \\ 0 & 5 \end{pmatrix}$$

Die Hilbert-Schmidt-Norm

Satz :

Der n^2 -dimensionaler Vektorraum aller komplexen $n \times n$ -Matrizen \mathbf{A} wird zu einem Hilbertraum, wenn man definiert

$$\|\mathbf{A}\|^2 = \text{trace}(\mathbf{A}^* \cdot \mathbf{A}) \quad .$$

Beweis

Seien \mathbf{a}_{ij} die Einträge der Matrix \mathbf{A} . Die Einträge von $\mathbf{A}^* \cdot \mathbf{A}$ sind $\mathbf{b}_{ik} = \sum_j \bar{\mathbf{a}}_{ij} \cdot \mathbf{a}_{jk}$.

$$\text{trace}(\mathbf{A}^* \cdot \mathbf{A}) = \sum_i \mathbf{b}_{ii} = \sum_{i,j} \bar{\mathbf{a}}_{ii} \cdot \mathbf{a}_{ij} = \sum_{i,j} |\mathbf{a}_{ij}|^2.$$

Das Quadrat der Hilbert-Schmidt-Norm von \mathbf{A} ist also die Quadratsumme der Einträge. Das innere Produkt ist

$$\langle \mathbf{A} | \mathbf{B} \rangle = \text{trace}(\mathbf{A}^* \cdot \mathbf{B}) \quad .$$

Das ergibt sich aus der Polarisierungsidentität.

Bemerke :

Man könnte sagen, dass $\|A\|^2$ die Summe der Normquadrate der Spalten von A ist. Diese Beobachtung wird vertieft durch den

Satz :

Sei $\varphi : V \rightarrow V$ eine lineare Abbildung eines Hilbertraums V in sich. Seien $\{v_j : j \in J\}$ und $\{w_k : k \in K\}$ ON Basen und A bzw. B die φ darstellenden Matrizen.

Es gilt dann

$$\sum \|\varphi(v_j)\|^2 = \text{trace}(A^*A) = \text{trace}(B^* \cdot B) = \sum \|\varphi(w_k)\|^2 .$$

(Diese Zahl bezeichnet man mit $\|\varphi\|^2$ und man nennt $\|\varphi\|$ die Hilbert-Schmidt-Norm der linearen Abbildung $\varphi : (V, \|\cdot\|) \rightarrow (V, \|\cdot\|)$.)

Beweis

Sei U die $J \times K$ -Matrix des Basiswechsels. Es gilt

$$\begin{aligned} B &= U^{-1} \cdot A \cdot U = U^* \cdot A \cdot U \\ \varphi(v_j) &= \sum v_i \cdot a_j^i, \quad \varphi(w_k) = \sum w_\ell \cdot b_k^\ell \\ \sum_j \|\varphi(v_j)\|^2 &= \langle \sum_{i_1} v_{i_1} \cdot a_j^{i_1} \mid \sum_{i_2} v_{i_2} \cdot a_j^{i_2} \rangle = \\ &= \sum_{j,i} \bar{a}_j^i \cdot a_j^i = \text{trace}(A^*A) \\ \text{Ebenso } \sum_k \|\varphi(w_k)\|^2 &= \text{trace}(B^* \cdot B) \\ \text{trace}(B^* \cdot B) &= \text{trace}(U^* A^* U U^* A U) = \text{trace}(A^* A) . \end{aligned}$$

Sei H die hermitische Matrix mit den Einträgen h_{ij} . Die Zahlen $|h_{ij}|^2$ nennen wir die Gewichte in den Positionen (i, j) . Die Summe über alle Gewichte ergibt das Quadrat der Hilbert-Schmidt-Norm. Es gilt $\|H\|^2 = \|U^* H U\|^2$ für jede unitäre Matrix. Wir definieren

$$\Delta(H) = \sum_{i \neq j} |h_{ij}|^2 = \|H\|^2 - \sum_i |h_{ii}|^2 .$$

Jacobi's Iterationsverfahren

Sei H eine hermitische $n \times n$ -Matrix. Für jedes unitäre B gilt $\|B^* H B\|^2 = \|H\|^2$.

Man finde unitäre B_1, B_2, \dots , sodass die transformierten Matrizen $H_{k+1} = B_k^* \cdot H_k \cdot B_k$ immer mehr Gewicht auf die Diagonale legen.

Mit dem Verfahren welches wir im Folgenden beschreiben, erreichen wir in jedem Falle

$$\Delta(H_k) \leq q^k \cdot \Delta(H_0)$$

mit einer Zahl $q < 1$, die nur von der Dimension abhängt. $q = 1 - \frac{2}{n(n-1)}$ ist leider nahe bei 1 für großes n . In den praktisch wichtigen Fällen darf man aber erwarten, dass die Konvergenz viel schneller ist.

Konstruktion

Der Schritt von $H = H_k$ zu $K = H_{k+1}$ erfolgt folgendermaßen. Man wählt ein Paar von Zeilenidizes $i \neq j$, wo $|h_{ij}|^2$ groß ist. h sei die Teilmatrix von H mit dem Format $\{i, j\} \times \{i, j\}$. b sei die 2×2 -Matrix, welche h auf Diagonalgestalt bringt. Für die Diagonalelemente μ_i, μ_j von b^*hb gilt $\mu_i^2 + \mu_j^2 = h_{ii}^2 + h_{jj}^2 + 2 \cdot |h_{ij}|^2$.

Mit der Matrix b konstruieren wir eine $n \times n$ -Matrix B mit den Blöcken b und E_{n-2} (Einheitsmatrix). Nehmen wir der Übersichtlichkeit halber $\{i, j\} = \{1, 2\}$ an.

$$B = \begin{pmatrix} b & \vdots & 0 \\ \dots\dots\dots & & \\ 0 & \vdots & E_{n-2} \end{pmatrix} ; \quad B^*HB = B^* \cdot \begin{pmatrix} H_{|} & \vdots & H_{||} \\ \dots\dots\dots & & \\ H_{|}^{\parallel} & \vdots & H_{||}^{\parallel} \end{pmatrix} B = \begin{pmatrix} b^*hb & \vdots & b^* \cdot H_{||} \\ \dots\dots\dots & & \\ H_{|}^{\parallel} \cdot b & \vdots & H_{||}^{\parallel} \end{pmatrix}$$

Das Gewicht der Diagonale wird um den Betrag $2 \cdot |h_{12}|^2$ erhöht

$$\Delta(B^* \cdot H \cdot B) = \Delta(H) - 2|h_{ij}|^2 .$$

Wenn wir $\{i, j\}$ so wählen, dass $|h_{ij}|^2$ maximal ist, dann $(n^2 - n)|h_{ij}|^2 \geq \Delta(H)$

$$\Delta(B^* \cdot H \cdot B) \leq \left(1 - \frac{2}{n^2 - n}\right) \cdot \Delta(H) .$$

Bemerke :

Jeder Iterationsschritt erfordert die Berechnung einer 2×2 -Matrix b , wie wir sie oben beschrieben haben. Ausserdem sind die beiden Spalten von $H_{|}^{\parallel} \cdot b$ aus den Spalten von $H_{|}^{\parallel}$ linear zu kombinieren.

V Lineare Algebra, Gleichungssysteme, Dualität

Didaktische Vorrede

Die lineare Algebra hat einen seit 50 Jahren etablierten Platz am Anfang jeder grundlagenorientierten Mathematikausbildung. Der Begriff des Vektorraums ist zusammen mit dem Begriff der Gruppe der Grundbaustein aller Algebra und Geometrie. Kurse über lineare Algebra müssen andererseits auch diejenigen bestehen, die sich zunächst einmal nicht für mathematische Grundstrukturen interessieren; denn lineare Gleichungssysteme nehmen in allen quantitativ arbeitenden Wissenschaften einen breiten Raum ein.

Für die Studierenden der Physik empfiehlt sich ein mittlerer Weg. Da man am Ende des ersten Semesters ein gewisses geometrisches Grundwissen voraussetzen kann, kann man sich der Aufgabe widmen, ein abstraktes Verständnis für Vektorräume zu verbinden mit einem ersten Einblick in den praktischen Umgang mit linearen Gleichungen.

In unserer Darstellung der linearen Algebra wird das bereits angesammelte Wissen über Vektorräume ausgerichtet auf den abstrakten Begriff der Dualität. Die Methode der Tableaus hat aus diesem Grund bei uns den Vorrang gegenüber der in Schulbüchern favorisierten Eliminationsmethode.

Die Dualitätstheorie für Vektorräume birgt keine technischen Schwierigkeiten; es ist ihre Abstraktheit, die sie für Anfänger schwierig macht. Didaktisch gutgemeinte Versuche, die Begriffe „konkret“ zu machen, führen leider sehr leicht in Sackgassen. Als Beispiele didaktischer Fallen möchte ich nennen: Es ist gefährlich, Bilinearformen und lineare Abbildungen in einen Topf zu rühren, indem man beide Objekttypen stets durch „konkrete“ Darstellungen an Matrizen bindet. Es ist gefährlich, Nullräume mit Orthogonalräumen zu verwechseln, was nahe liegt, wenn man die Vorstellungen stets an die euklidischen Räume knüpft. Es ist gefährlich, duale Abbildungen mit adjungierten Operatoren zu verwechseln, was nahe liegt, wenn man allzu schnell an euklidische (oder unitäre) Räume denkt.

Wir hoffen, dass wir durch unser vorsichtiges Sprechen über Linearformen und Bilinearformen in den Abschnitten I und IV die Studierenden solcher Verwechslungsgefahren enthoben haben. Die Studierenden sollten verstanden haben, dass man sich aus der allgemeinen Theorie herausbegibt, wenn man darauf besteht, Vektoren zu multiplizieren (‘inneres Produkt’, ‘Vektorprodukt’ u.dgl.). Die Disziplin beim mathematischen Operieren sollte sich positiv auswirken, wenn es später um Tensorprodukte und um Dachprodukte geht. Wir versuchen, in der abstrakten Theorie der Vektorräume, der Forderung nach geometrischer Anschaulichkeit auf einem höheren Niveau zu entsprechen. Weitere geometrische Aspekte der Dualität in Sinne der ‘Vektorraumgeometrie’ verschieben wir in die ‘Multilineare Algebra’, die sich natürlicherweise anschliessen würde an die hier abgehandelte erste Einführung in die lineare Geometrie.

Themenübersicht

V.1 Vektorräume: Basis und Dimension

Vektorraumaxiome. Teilvektorräume. Die lineare Hülle einer Schar von Vektoren $\text{span}\{v_\alpha : \alpha \in I\}$. Lineare Unabhängigkeit. Eine Basis ist ein linear unabhängiges System maximaler Länge. Eine Basis ist ein aufspannendes System minimaler Länge. Beweis des Austauschlemmas von E. Steinitz. Linearformen. Der Dualraum. Duale Basen. Die natürliche Paarung.

V.2 Lineare Gleichungssysteme

Ein lineares Gleichungssystem führt auf Vektorräume und auf affine Räume. Die allgemeine Lösung des inhomogenen Systems $Ax = b$ ergibt sich, wenn man zu einer speziellen Lösung die allgemeine Lösung der homogenen Gleichung dazuaddiert. Ein lineares Gleichungssystem definiert eine lineare Abbildung $\varphi : x \mapsto Ax$. Wenn A den Spaltenrang r hat, dann hat das Bild $\text{im } \varphi$ die Dimension r . Wenn A den Zeilenrang r hat, dann hat der Kern $\ker \varphi = \{x : Ax = 0\}$ die Dimension $n - r$. Zeilenrang = Spaltenrang. Wenn A den Rang r hat, dann gibt es eine nichtsinguläre Teilmatrix von der Größe $r \times r$.

V.3 Tableaus. Vollständiger Austausch

Zeilentableaus zur „Lösung“ linearer Gleichungssysteme. Spaltentableaus zur Konstruktion einer Basis des Spaltenraums. (Austauschlemma von Steinitz). Vollständiger Austausch: der Kernblock, der Bildblock und der Inversblock. Weitere Anwendungen des vollständigen Austausches.

V.4 Eliminationsmethode; LU-Faktorisierung

Das Gauss'sche Eliminationsverfahren. Zwei Zahlenbeispiele. Zeilenoperationen sind Multiplikationen von links. Untere und obere Dreiecksmatrizen. Die LU-Zerlegung $A = L^{-1} \cdot D \cdot U$ für eine nichtsinguläre Matrix A mit geeignet angeordneten Zeilen und Spalten. Inhomogene Systeme. Systeme mit unbestimmter rechter Seite. Zahlenbeispiel. Hinweis auf die Minimierung einer quadratischen Funktion.

?? Matrizen vom Rang r , Singulärwertzerlegung

Eine $m \times n$ Matrix A vom Rang r kann man als Summe von r Matrizen vom Rang 1 darstellen. Anders gesagt: Es gibt Faktorisierungen $A = B \cdot C$, wo B das Format $m \times r$ hat und C das Format $r \times n$; die Spalten von B bilden eine Basis des Spaltenraums von A , die Zeilen von C bilden eine Basis des Zeilenraums von A . Wir diskutieren insbesondere die Faktorisierung, die zum vollständigen Austausch gehört.

Dann beschäftigen wir nochmals ausführlich mit den komplexen Matrizen und den unitären Diagonalisierungen. Wir konstruieren (mit der Maximierungsmethode von Courant) die

sog. Singulärwertzerlegung einer beliebigen komplexen $m \times n$ Matrix A .

$$A = Q \cdot \Lambda \cdot P \quad \text{mit } \Lambda \text{ Diagonalmatrix und } \quad Q^*Q = I_r = PP^* .$$

Wir greifen zunächst die Rechnungen aus I.3 nochmals auf, und konstruieren dann auf einem geometrischen Weg (ähnlich wie in IV.2) die (nicht immer eindeutige) Singulärwertzerlegung. Die unitäre Diagonalisierung einer normalen Matrix ergibt sich als Corollar.

V.6 Kleinste Quadrate und lineare Schätzung

Die Gleichung $B \cdot x = y$ besitzt eine Lösung nur für die y im Spaltenraum von B . Für ein beliebig vorgegebenes y sucht man nun eine Spalte \hat{y} und eine Spalte \hat{x} , sodass gilt

$$B \cdot \hat{x} = \hat{y} , \quad \|y - \hat{y}\| = \min .$$

Diese Aufgabe wird in dem Fall, wo $\|\cdot\|$ eine Hilbertraumnorm im Spaltenraum ist, durch die Methode der kleinsten Quadrate gelöst.

Als Beispiel diskutieren wir ausführlich die Bestimmung einer Ausgleichsgeraden.

In der Linearen Statistik löst man ein verwandtes Problem: Das vorgegebene n -Tupel y wird als eine Realisation eines Zufallsvektors Y verstanden. Man nimmt an, dass man zwar die Covarianzmatrix $C = \text{cov}(Y, Y)$ kennt, vom Erwartungswert aber nur weiss, dass er im Spaltenraum von B liegt.

$$Y = B \cdot s + U \quad \text{mit} \quad \mathbb{E}U = 0 , \quad \text{cov}(U, U) = C , \quad s \text{ unbekannt} .$$

Gesucht ist ein linearer Schätzer $N \cdot Y$ mit

$$\mathbb{E}(N \cdot Y) = s \quad \text{falls} \quad \mathbb{E}Y = s \quad \text{und} \quad \text{cov}(NY, NY) = \min .$$

Aus der Sicht der Linearen Algebra geht es auch hier um eine orthogonale Projektion, die durch die Lösung einer ‘Normalgleichung’ bestimmt wird. Wir zeigen, dass für $Q = C^{-1}$ die beiden Vorgehensweisen zur gleichen Projektionsmatrix führen. -Das ist im Wesentlichen der sog. Satz von Gauss - Markov.

V Lineare Algebra

V.1 Vektorräume: Basis und Dimension

Wir haben bereits mehrere \mathbb{R} -Vektorräume und \mathbb{C} -Vektorräume kennengelernt. Für die abstrakte Theorie der (endlichdimensionalen) Vektorräume, die wir jetzt entwickeln wollen, brauchen wir keine speziellen Eigenschaften des Skalarenkörpers. Sei also \mathbb{K} irgendein Körper.

Definition (\mathbb{K} -Vektorraum)

Ein \mathbb{K} -Vektorraum ist ein Tupel

$$(V, 0, +, \cdot) .$$

Hierbei ist V eine Menge mit einem ausgezeichneten Element 0 und einer Verknüpfung $+$, welche $(V, 0, +)$ zu einer kommutativen Gruppe macht, wo 0 das neutrale Element ist.

Die skalare Multiplikation \cdot ist eine Abbildung

$$\mathbb{K} \times V \rightarrow V, (\alpha, v) \mapsto \alpha v$$

mit den Eigenschaften

$$\begin{aligned} \alpha(v_1 + v_2) &= \alpha v_1 + \alpha v_2, \\ (\alpha + \beta)v &= \alpha v + \beta v, \\ \alpha(\beta v) &= (\alpha \cdot \beta)v, \\ 1 \cdot v &= v. \end{aligned}$$

Definitionen

- a) Eine Teilmenge eines Vektorraums, die gegenüber Addition und skalarer Multiplikation abgeschlossen ist, heißt ein Teilvektorraum.
- b) Der kleinste Teilvektorraum, welcher eine gegebene Familie $\{v_\alpha : \alpha \in I\}$ enthält, heißt die lineare Hülle dieser Familie oder der von der Familie erzeugte Vektorraum. Man notiert $\text{span}\{v_\alpha : \alpha \in I\}$. Die Familie heißt ein aufspannendes System für diesen Vektorraum.
- c) Ein Teilvektorraum V heißt endlichdimensional, wenn er ein endliches aufspannendes System besitzt.
- d) Ein endliches aufspannendes System minimaler Länge heißt eine Basis des Teilvektorraums. Die Länge der Basis heißt die Dimension des Teilvektorraums.
- e) Eine Familie $\{v_\alpha : \alpha \in I\}$ heißt linear unabhängig, wenn man den Nullvektor nur auf die triviale Weise als Linearkombination der v_α darstellen kann.
(In unseren \mathbb{K} -Vektorräumen gibt es nur endliche Linearkombinationen.)

$$\{v_\alpha : \alpha \in I\} \text{ linear unabhängig} \Leftrightarrow \left(\sum \lambda_\alpha v_\alpha = 0 \Rightarrow \forall \alpha : \lambda_\alpha = 0 \right) .$$

Satz :

Die Dimension eines endlichdimensionalen Vektorraums V ist die maximale Länge eines linear unabhängigen Tupels.

Wenn V n -dimensional ist, dann ist jedes linear unabhängige n -Tupel eine Basis.

Der Beweis wird sich leicht ergeben, wenn wir zuerst einige wichtige Konstruktionen in Vektorräumen diskutieren. Wir können den Satz noch etwas anders formulieren.

Satz (Charakterisierung einer Basis)

Sei V ein n -dimensionaler Vektorraum. Dann gilt

- a) Jedes linear unabhängige System der Länge n ist ein aufspannendes System.
- b) Jedes aufspannende System der Länge n ist ein linear unabhängiges System.

Der Satz weist mehrere Wege, Basen zu konstruieren:

a) Wenn ein linear unabhängiges System noch nicht den ganzen Vektorraum V aufspannt, dann kann man irgendeinen Vektor, der nicht in der linearen Hülle liegt, hinzunehmen, ohne die lineare Unabhängigkeit zu gefährden.

b) Wenn ein aufspannendes System nicht linear unabhängig ist, dann kann man gewisse Vektoren weglassen ohne die lineare Hülle zu verkleinern.

c) Sei $\{v_1, \dots, v_n\}$ eine Basis von V und $w = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$.

Wenn $\alpha_k \neq 0$, dann kann man v_k durch w ersetzen und erhält eine neue Basis („Austauschlemma“).

Beweis des Austauschlemma

o.B.d.A. nehmen wir an $\alpha_1 \neq 0$. Wir können dann v_1 durch w ersetzen; denn

$$v_1 = \frac{1}{\alpha_1} w - \sum_{k=2}^n \frac{\alpha_k}{\alpha_1} v_k$$

Jedes $\tilde{w} = \beta_1 v_1 + \sum_{k=2}^n \beta_k v_k$ kann auch mit Hilfe des Systems $\{w, v_2, \dots, v_n\}$ ausgedrückt werden

$$\tilde{w} = \frac{\beta_1}{\alpha_1} w + \sum_{k=2}^n \left(\beta_k - \frac{\alpha_k \cdot \beta_1}{\alpha_1} \right) v_k .$$

Den Austausch notiert man als Pivottransformation eines sog. Spaltentableau.

v_1	α_1	β_1		w	$\frac{1}{\alpha_1}$	$\frac{\beta_1}{\alpha_1}$
v_2	α_2	β_2		v_2	$-\frac{\alpha_2}{\alpha_1}$	$\beta_2 - \frac{\alpha_2}{\alpha_1} \beta_1$
	\vdots				\vdots	
v_n	α_n	β_n		v_n	$-\frac{\alpha_n}{\alpha_1}$	$\beta_n - \frac{\alpha_n}{\alpha_1} \beta_1$
	$= w$	$= \tilde{w}$			$= v_1$	$= \tilde{w}$

Diesen Gedanken werden wir später gründlich verfolgen.

Linearformen, duale Basis

Definition V sei ein \mathbb{K} -Vektorraum.

- a) Eine \mathbb{K} -wertige Funktion $\ell(\cdot)$ auf V heißt eine **Linearform**, wenn gilt

$$\ell(\alpha v_1 + \beta v_2) = \alpha \cdot \ell(v_1) + \beta \cdot \ell(v_2) \quad \text{für alle } \alpha, \beta \in \mathbb{K}, v_1, v_2 \in V.$$

- b) Der Vektorraum aller Linearformen heißt der **Dualraum** von V und wird mit V^* bezeichnet.

- c) Wenn $\{v_1, \dots, v_n\}$ eine Basis von V ist und $\ell^{(1)}, \dots, \ell^{(n)}$ die Linearformen mit
- $$\ell^{(i)}(v_j) = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$
- dann heißt $\{\ell^{(1)}, \dots, \ell^{(n)}\}$ die **duale Basis**.

Die Bezeichnung „duale Basis“ braucht eine Rechtfertigung. Wir müssen uns überzeugen.

- i) Die $\ell^{(i)}(\cdot)$ sind linear unabhängig.
- ii) Die $\ell^{(i)}(\cdot)$ erzeugen den ganzen Dualraum V^* .

Dies ergibt sich aus den folgenden Bemerkungen:

1. Eine Linearform $\ell(\cdot)$ ist durch ihre Werte in den Basisvektoren v_j eindeutig bestimmt.

$$\ell(v_j) = \beta_j \Rightarrow \ell\left(\sum \alpha^j v_j\right) = \sum \beta_j \cdot \alpha^j.$$

Es gilt

$$\ell(\cdot) = \sum \beta_j \cdot \ell^{(j)}(\cdot).$$

2. Da jedes $v \in V$ eine eindeutige Darstellung $v = \sum \alpha^j v_j$ besitzt, sind die Abbildungen $v \mapsto \alpha^j$ Linearformen. Sie sind in der Tat die Elemente der dualen Basis. Es gilt

$$v = \sum \ell^{(j)}(v) \cdot v_j.$$

Notation : Statt $\ell(v)$ notiert man $\langle \ell, v \rangle$. Die Abbildung

$$\langle \cdot, \cdot \rangle : V^* \times V \rightarrow \mathbb{K}$$

heißt die **natürliche Paarung** für V .

Bemerkungen

1. Die natürliche Paarung ist bilinear

$$\begin{aligned} \langle \alpha \ell' + \beta \ell'', \cdot \rangle &= \alpha \cdot \langle \ell', \cdot \rangle + \beta \cdot \langle \ell'', \cdot \rangle \\ \langle \cdot, \alpha w_1 + \beta w_2 \rangle &= \alpha \cdot \langle \cdot, w_1 \rangle + \beta \cdot \langle \cdot, w_2 \rangle. \end{aligned}$$

2. Jedes $v \in V$ liefert eine Linearform auf V^* und jede Linearform auf V^* kommt von einem v . Man sagt kurz: V^{**} ist in natürlicher Weise isomorph zu V .
3. Die duale Basis zu $\{\ell^{(1)}, \dots, \ell^{(n)}\}$ ist die ursprüngliche Basis $\{v_1, \dots, v_n\}$ (wenn man V^{**} mit V identifiziert).

Konvention

Wenn man in einem Vektorraum V eine Basis $\{v_j : j \in J\}$ ausgezeichnet hat, dann beschreibt man den Vektor $v = \sum \alpha^j \cdot v_j$ durch die **J-Spalte** mit den Einträgen α^j . Die Linearformen $\ell(\cdot)$, die man auch **Covektoren** nennt, beschreibt man durch **J-Zeilen**. $\ell(\cdot) = \sum \beta_j \cdot \ell^{(j)}(\cdot)$ ist die Zeile mit den Einträgen β_j .

$\langle \ell, v \rangle$ ist das „Matrizenprodukt“ $\sum \beta_j \cdot \alpha^j$ (Zeile \times Spalte).

Achtung!

1. Die natürliche Paarung nennt man manchmal das innere Produkt. Man darf dieses innere Produkt aber nicht mit dem inneren Produkt („Skalarprodukt“) in euklidischen Räumen verwechseln.

Das Skalarprodukt im euklidischen (bzw. hilbertschen) Raum ist eine Abbildung

$$\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{R} \text{ (bzw. } \mathbb{C} \text{)}$$

die bilinear (bzw. sesquilinear) ist.

2. Man beachte auch: Aus einem Paar von Spaltenvektoren kann man nicht ohne weiteres einen Skalar gewinnen; man muss den einen Spaltenvektor zuerst zu einem Zeilenvektor machen, z.B. mit Transposition oder durch hermitesche Konjugation.

$$(x, y) \mapsto x^T \cdot y \text{ oder } (x, y) \mapsto x^* \cdot y .$$

Hinweis Bei den d -dimensionalen Mannigfaltigkeiten hat man in jedem Punkt $P \in M$ den Tangentialraum T_P . Dies ist der d -dimensionale \mathbb{R} -Vektorraum aller Tangentialvektoren. Sein Dualraum T_P^* heißt der Cotangentialraum. Das Differential df einer glatten Funktion („der Gradient“) liefert in jedem Punkt P einen Covektor.

Die natürliche Paarung macht aus dem Gradienten und dem Tangentialvektor der Kurve $\gamma(\cdot)$ zu jedem Zeitpunkt t eine reelle Zahl

$$\langle df, \dot{\gamma}(t) \rangle = \frac{d}{dt} f(\gamma(t)) .$$

Man sagt : Die infinitesimale Änderung der Funktion f entlang der Kurve ist das „innere Produkt“ des Gradienten mit dem Tangentialvektor.

Man sollte bei dieser Aussage nicht denken, dass man eine euklidische Struktur in T_P braucht. Das „innere Produkt“ ist die natürliche Paarung.

Satz

Seien V_1 und V_2 Teilvektorräume von V . Die lineare Hülle $V_1 + V_2$ und der Durchschnitt sind dann ebenfalls Teilräume. Es gilt

$$\dim(V_1 \cap V_2) + \dim(V_1 + V_2) = \dim V_1 + \dim V_2 .$$

Beweis

Sei u_1, \dots, u_d eine Basis von $V_1 \cap V_2$. Wir können sie mit geeigneten v_1, \dots, v_ℓ zu einer Basis $\{u_1, \dots, u_d, v_1, \dots, v_\ell\}$ von V_1 fortsetzen; und mit geeigneten w_1, \dots, w_k zu einer Basis $\{u_1, \dots, u_d, w_1, \dots, w_k\}$ von V_2 . Alle die $d + \ell + k$ Vektoren zusammen liefern eine Basis von $V_1 + V_2$. Dass sie $V_1 + V_2$ aufspannen ist trivial. Die lineare Unabhängigkeit ergibt sich folgendermaßen. Sei $u + v + w = 0$ mit $u \in V_1 \cap V_2$, $v \in V' = \text{span}\{v_1, \dots, v_\ell\}$ und $w \in V'' = \text{span}\{w_1, \dots, w_k\}$. Dann gilt $-w = u + v \in V_1 \cap V_2$, also $v = 0$ und $w = 0$.

Bemerkung :

Jeder Vektor $\tilde{v} \in V_1 + V_2$ hat eine eindeutige Darstellung $\tilde{v} = u + v + w$ mit $u \in V_1 \cap V_2$, $v \in V'$, $w \in V''$. Wir haben eine direkte Zerlegung

$$V_1 + V_2 = (V_1 \cap V_2) \oplus V' \oplus V'' .$$

(vgl. IV.2)

Man spricht hier von einer „inneren“ direkten Zerlegung, weil ein umfassender Vektorraum von vornherein gegeben ist. Bei einer „äußeren“ direkten Summe denkt man an die folgende

Konstruktion

Seien V und W \mathbb{K} -Vektorräume. Das kartesische Produkt $V \times W$ (d.h. die Gesamtheit aller Paare (v, w)) besitzt dann in natürlicher Weise eine Vektorraumstruktur

$$\begin{aligned} \alpha \cdot (v, w) &= (\alpha v, \alpha w) \\ (v_1, w_1) + (v_2, w_2) &= (v_1 + v_2, w_1 + w_2) . \end{aligned}$$

Man nennt diesen Vektorraum der Dimension $n + m$ die äussere direkte Summe und bezeichnet ihn mit $V \oplus W$. Zu jedem Paar $(v, w) \in V \times W$ erhält man einen Vektor $v \oplus w \in V \oplus W$.

Sei (v_1, \dots, v_n) eine Basis von V , und sei (w_1, \dots, w_m) eine Basis von W . Wir betrachten dazu in der ‘äusseren’ direkten Summe die Vektoren $\tilde{v}_j = v_j \oplus 0$ und die Vektoren $\tilde{w}_k = 0 \oplus w_k$. $(\tilde{v}_1, \dots, \tilde{v}_n, \tilde{w}_1, \dots, \tilde{w}_m)$ ist dann eine Basis von $V \oplus W$; jedes $u \in V \oplus W$ besitzt genau eine Darstellung $u = x^1 \cdot \tilde{v}_1 + \dots + x^n \cdot \tilde{v}_n + y^1 \cdot \tilde{w}_1 + \dots + y^m \cdot \tilde{w}_m$.

Den Dualraum von $V \oplus W$ kann man mit der direkten Summe $V^* \oplus W^*$ identifizieren.

Hinweis

Wir werden später zu V, W den Vektorraum $V \otimes W$, das Tensorprodukt konstruieren. Man darf das Zeichen \otimes nicht mit dem Zeichen \times für das mengentheoretische cartesische Produkt verwechseln. Das Tensorprodukt ist ein Vektorraum der Dimension $(\dim V) \cdot (\dim W)$; die Elemente sind Objekte einer neuen Art. Die Physiker kennen Trägheitstensoren, Spannungstensoren und dergleichen mehr. Die abstrakte Konstruktion der Tensorprodukte beginnen wir im Abschnitt VII.

V.2 Lineare Gleichungssysteme

Ein **inhomogenes** lineares Gleichungssystem ist ein Gleichungssystem der Form

$$\sum_j a_{ij} \cdot x_j = b_i \quad \text{für } i \in I.$$

Dabei ist die $I \times J$ -Matrix A und die I -Spalte b gegeben. Man schreibt auch

$$\sum_j a_j^i \cdot x^j = b^i \quad \text{für } i \in I \quad \text{oder kurz} \quad Ax = b.$$

Wenn man die rechte Seite b gleich 0 setzt, erhält man das dazugehörige **homogene** Gleichungssystem $Ax = 0$.

Wir betrachten auch lineare Gleichungssysteme mit **unbestimmter rechter Seite**.

Zahlenbeispiel

$$\begin{aligned} 1 \cdot x^1 + 2 \cdot x^2 + 4 \cdot x^3 + 1 \cdot x^4 &= -y^1 \\ 1 \cdot x^1 + 1 \cdot x^2 + 1 \cdot x^3 + 0 \cdot x^4 &= -y^2 \\ -1 \cdot x^1 + 0 \cdot x^2 + 2 \cdot x^3 + 1 \cdot x^4 &= -y^3. \end{aligned}$$

In Matrizenschreibweise

$$\begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ x^4 \end{pmatrix} = \begin{pmatrix} -y_1 \\ -y_2 \\ -y_3 \end{pmatrix}$$

Die „Unbekannten“ müssen nicht mit x^1, x^2, x^3, x^4 durchnummeriert sein. Wenn wir unser Zahlenbeispiel genauer analysieren, werden wir die Bezeichnung x^1, x^2, z^1, z^2 bevorzugen. In Tableau-Schreibweise sieht man am schönsten, dass es darauf ankommt, die Spalte $-y$ als Linearkombination der Spalten unserer Matrix darzustellen.

$$\begin{array}{cccc|l} x^1 & x^2 & z^1 & z^2 & \\ \hline 1 & 2 & 4 & 1 & = -y^1 \\ 1 & 1 & 1 & 0 & = -y^2 \\ -1 & 0 & 2 & 1 & = -y^3 \end{array}$$

Es stellen sich drei Probleme

- I. Für welche y existiert eine Lösung?
- II. Man finde eine spezielle Lösung zu jedem guten y .
- III. Finde eine übersichtliche Darstellung der Menge aller Lösungen.

Die linearen Abbildungen zu einer Matrix

Sei A eine $I \times J$ -Matrix, $|I| = m$, $|J| = n$. Die Abbildung

$$\varphi : x \mapsto Ax, \quad \varphi : V \longrightarrow W$$

ordnet jeder J -Spalte eine I -Spalte zu. Man definiert

$$\begin{aligned} \ker \varphi &= \{v : \varphi(v) = 0\} && \text{„Kern von } \varphi\text{“} \\ \operatorname{im} \varphi &= \{w : w = \varphi(v)\} && \text{„Bild von } V \text{ bzgl. } \varphi\text{“} \end{aligned}$$

Die drei Probleme lauten in dieser Sprache

- I. Charakterisiere das Bild $\operatorname{im} \varphi$.
- II. Finde ein spezielles Urbild zu jedem guten w .
- III. Finde eine übersichtliche Beschreibung von $\ker \varphi$.

Die zu φ duale Abbildung φ^* bildet die Linearformen auf W , (d.h. die I -Zeilen) in Linearformen auf V ab. Das φ^* -Bild von W^* ist ein Raum von Linearformen.

$$\varphi^* : \xi \mapsto \xi A; \quad \varphi^* : V^* \leftarrow W^* .$$

Es gilt in der Tat für alle $\xi \in W^*$, $v \in V$ mit $\eta = \xi A = \varphi^*(\xi)$ und $w = \varphi(v)$

$$\begin{aligned} \eta \cdot v &= (\xi A) \cdot v = \xi \cdot Av = \xi \cdot w \\ \langle \varphi^*(\xi), v \rangle &= \langle \xi, \varphi(v) \rangle . \end{aligned}$$

Interessant sind nun auch die Vektorräume

$$\begin{aligned} \ker \varphi^* &= \{\xi : \xi A = 0\} \\ \operatorname{im} \varphi^* &= \{\eta : \eta = \xi A\} \end{aligned}$$

Wir schreiben in Tableauform

$$\begin{array}{c} x^1, \dots, x^n \\ \boxed{} \end{array} \begin{array}{l} = -y^1 \\ \vdots \\ = -y^m \end{array} \quad \begin{array}{c} \xi_1 \\ \vdots \\ \xi_m \\ \boxed{} \end{array} \begin{array}{l} = \eta_1, \dots, = \eta_n \end{array}$$

Antworten für das Zahlenbeispiel

Anhand unseres Zahlenbeispiels können wir nun klarmachen, wie Antworten auf die drei Fragen aussehen können.

Das Gleichungssystem ist

$$A \cdot \begin{pmatrix} x \\ z \end{pmatrix} = -y \quad \text{mit} \quad A = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 2 & 1 \end{pmatrix}$$

Wir werden eine Methode kennenlernen, wie man das entsprechende Tableau umformen kann in ein äquivalentes Tableau. In unserem Zahlenbeispiel gelangt man zu einem äquivalenten Tableau, welches in der rechten unteren Ecke bemerkenswerte Nullen aufweist.

$$\begin{array}{cccc|l}
 x^1 & x^2 & z^1 & z^2 & \\
 \hline
 1 & 2 & 4 & 1 & = -y^1 \\
 1 & 1 & 1 & 0 & = -y^2 \\
 -1 & 0 & 2 & 1 & = -y^3
 \end{array}
 \qquad
 \begin{array}{cccc|l}
 y^1 & y^2 & z^1 & z^2 & \\
 \hline
 -1 & 2 & -2 & -1 & = -x^1 \\
 1 & -1 & 3 & 1 & = -x^2 \\
 -1 & 2 & 0 & 0 & = -y^3
 \end{array}$$

Aus dem gemäß der Methode des vollständigen Austauschs umgeformten Tableau kann man nun in der Tat alles ablesen, was uns interessiert:

I) Eine Lösung existiert genau dann, wenn

$$-y^1 + 2y^2 = -y^3.$$

II) Zu jedem guten y gibt es eine spezielle Lösung von der Form $(x^1, x^2, 0, 0)^T$, nämlich

$$\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -x^1 \\ -x^2 \end{pmatrix}.$$

III) Die Lösungsmenge der homogenen Gleichung $A \cdot \begin{pmatrix} x \\ z \end{pmatrix} = 0$ ist zweidimensional.

Wir können z.B. $z = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$ frei wählen. Die Lösung dazu $\tilde{x} = (\tilde{x}^1, \tilde{x}^2, z^1, z^2)^T$ berechnet sich mit der Matrix rechts oben.

$$\begin{pmatrix} -2 & -1 \\ +3 & +1 \end{pmatrix} \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = \begin{pmatrix} -\tilde{x}^1 \\ -\tilde{x}^2 \end{pmatrix}.$$

Bevor wir uns dem Verfahren der Berechnung zuwenden, diskutieren wir

Geometrische Aspekte

Satz 1 (Affine Lösungsräume)

Die allgemeine Lösung der inhomogenen Gleichung $Ax = b$ ergibt sich dadurch, dass man zu einer speziellen Lösung \tilde{x} die allgemeine Lösung v der homogenen Gleichung dazu addiert.

Beweis

(Wir benützen den Beweis, um an Begriffe aus früheren Abschnitten zu erinnern.)

1. Die Menge $\{x : Ax = b\}$ ist (wenn sie nicht leer ist) ein affiner Raum.

$$Ax_1 = v, \quad Ax_2 = b \implies \forall \lambda \in \mathbb{K} \quad A \cdot ((1 - \lambda)x_1 + \lambda x_2) = b.$$

2. Für jeden Vektor \mathbf{v} mit $A\mathbf{v} = \mathbf{0}$ gilt

$$A\mathbf{x} = \mathbf{b} \implies A(\mathbf{x} + \mathbf{v}) = \mathbf{b} .$$

Das Lösungsgebilde der homogenen Gleichung ist der Tangentialraum für jeden der affinen Räume $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$.

Satz 2 (Zeilenrang)

Sei $\{g^i(\cdot) : i \in I\}$ eine Familie von Linearformen auf dem \mathbf{n} -dimensionalen \mathbb{K} -Vektorraum V . Wenn die $g^i(\cdot)$ einen \mathbf{r} -dimensionalen Raum aufspannen, dann ist das gemeinsame Nullstellengebilde $\{\mathbf{v} : g^i(\mathbf{v}) = 0 \text{ für alle } i\}$ ein $(\mathbf{n} - \mathbf{r})$ -dimensionaler Teilvektorraum.

Sei A eine $I \times J$ -Matrix, $|I| = \mathbf{m}$, $|J| = \mathbf{n}$. Wenn die Zeilen einen \mathbf{r} -dimensionalen Raum aufspannen, dann ist der Lösungsraum der homogenen Gleichung $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ ein $(\mathbf{n} - \mathbf{r})$ -dimensionaler Teilraum des Raums aller J -Spalten.

Kurz gesagt: $\dim(\ker \varphi) = \dim V - \dim(\operatorname{im} \varphi^*)$.

Beweis

o.B.d.A. nehmen wir an, dass die $g^i(\cdot)$ linear unabhängig sind. Wir haben also ein linear unabhängiges \mathbf{r} -Tupel $g^1(\cdot), g^2(\cdot), \dots, g^r(\cdot)$. Wir setzen es fort zu einer Basis von V^* , dem Raum aller Linearformen auf V . $\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$ sei die duale Basis. Ein Vektor $\mathbf{v} = \sum \mathbf{v}_j \mathbf{x}^j$ gehört genau dann zu dem gemeinsamen Nullstellengebilde der $g^i(\cdot)$, wenn $\mathbf{x}^j = 0$ für $j = 1, \dots, \mathbf{r}$. Die Menge dieser \mathbf{v} bildet einen $(\mathbf{n} - \mathbf{r})$ -dimensionalen Vektorraum.

Sprechweise :

Wenn die Menge der Zeilen einer Matrix A einen \mathbf{r} -dimensionalen Raum aufspannt, dann sagt man A hat den **Zeilenrang** \mathbf{r} . (Technisch ausgedrückt: $\dim(\operatorname{im} \varphi^*) = \mathbf{r}$.)

Satz 3 (Spaltenrang = Zeilenrang)

Für jede lineare Abbildung $\varphi : V \rightarrow W$ gilt $\dim(\ker \varphi) + \dim(\operatorname{im} \varphi) = \dim V$.

Für jede $I \times J$ -Matrix gilt Spaltenrang = Zeilenrang.

Beweis

Der Kern $\ker \varphi$ sei $(\mathbf{n} - \mathbf{r})$ -dimensional. Wir zeigen, dass das Bild $\operatorname{im} \varphi$ \mathbf{r} -dimensional ist. Sei $\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$ eine Basis von $\ker \varphi$. Wir setzen das linear unabhängige $(\mathbf{n} - \mathbf{r})$ -Tupel zu einer Basis von V fort: $\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$. Die Bilder $\mathbf{w}_j = \varphi(\mathbf{v}_j)$, $j = 1, \dots, \mathbf{r}$ spannen $\operatorname{im} \varphi$ auf. Wir zeigen, dass sie linear unabhängig sind. In der Tat

$$\sum \lambda^i \mathbf{w}_j = \mathbf{0} \implies \varphi\left(\sum \lambda^i \mathbf{v}_j\right) = \mathbf{0} \implies \sum \lambda^i \mathbf{v}_j \in \ker \varphi .$$

Dies impliziert $0 = \lambda^1 = \dots = \lambda^r$.

Sprechweise :

Wenn die Spalten einer Matrix A einen r -dimensionalen Vektorraum aufspannen, dann sagt man, A hat den **Spaltenrang** r . (Technisch ausgedrückt: $\dim(\operatorname{im} \varphi) = r$.)

Fortsetzung des Beweises: Wir haben eben gesehen

$$\dim(\ker \varphi) = n - \dim(\operatorname{im} \varphi) .$$

Im Satz 2 haben wir andererseits gesehen

$$\dim(\ker \varphi) = n - \dim(\operatorname{im} \varphi^*) .$$

Die Abbildungen φ und φ^* haben denselben Rang. Der Spaltenrang ist gleich dem Zeilenrang.

Satz 4 (Reguläre Fenster einer Matrix)

Wenn die $I \times J$ -Matrix A den Rang r hat, dann gibt es eine Teilmatrix A' , die eine invertierbare $r \times r$ -Matrix ist.

Wir wollen den Sachverhalt genauer verstehen und führen dazu einige Sprechweisen ein: Eine Teilmenge $I' \subseteq I$ nennen wir eine „gute“ Zeilenmenge (von A), wenn die Zeilen in den Positionen $\in I'$ linear unabhängig sind. Eine Teilmenge $J' \subseteq J$ nennen wir eine „gute“ Spaltenmenge, wenn die Spalten $\{w_j : j \in J'\}$ linear unabhängig sind. $I' \times J'$ nennen wir ein „gutes“ Fenster (für die Matrix A), wenn die Teilmatrix vom Format $I' \times J'$ invertierbar ist. Der Satz vom Rang sagt, dass es zu jedem guten I' ein J' gibt, sodass $I' \times J'$ ein „gutes“ Fenster ist. Wir können Schritt für Schritt „Pivots“ $(i_1, j_1), (i_2, j_2), \dots, (i_r, j_r)$ wählen, sodass für jedes k die Teilmatrix vom Format $\{i_1, \dots, i_k\} \times \{j_1, \dots, j_k\}$ invertierbar ist. Satz 4 ist bewiesen.

Ausblick : Wir werden sehen, wie man „gute“ Fenster der Größe $r \times r$ schrittweise findet. Die Methode löst die folgende Aufgabe: Wähle aus der Menge der Spalten (und Zeilen) einer Matrix A eine Basis für den Spaltenraum (und Zeilenraum) aus. Die Aufgabe kann als eine Konkretisierung des Problems I angesehen werden: Charakterisiere das Bild $\operatorname{im} \varphi$, finde eine Basis bestehend aus den Spalten der gegebenen Matrix.

Zweites Zahlenbeispiel

Wir studieren

$$Ax = -y \quad \text{mit der Matrix}$$

$$A = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 2 \end{pmatrix} \quad \text{statt} \quad \begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 2 & 1 \end{pmatrix} .$$

Der Vergleich mit dem ersten Zahlenbeispiel, wo die Matrix den Zeilenrang 2 hat, zeigt, dass wir es jetzt mit einer Matrix vom Rang 3 zu tun haben. $Ax + y = 0$ ist für jedes y

lösbar.

Die drei ersten Spalten bilden ein gutes Fenster $A|$. Es existiert eine Matrix $B|$, sodass

$$B| \cdot A| = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{nämlich } B| = \begin{pmatrix} 1 & -2 & -2 \\ -2 & 5 & 3 \\ 1 & -2 & -1 \end{pmatrix}.$$

Wir werden sehen, wie man B_1 schrittweise berechnet. Wir werden natürlich auch $\ker \varphi$ bestimmen. Eine konzise Darstellung aller Rechenergebnisse findet man im „vollständig ausgetauschten Tableau“.

$$\begin{array}{c}
 \begin{array}{cc} x & z \end{array} \\
 \boxed{\begin{array}{ccc|c} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 2 \end{array}} = -y \quad \longleftrightarrow \quad \boxed{\begin{array}{cc} y & z \end{array}} \\
 \boxed{\begin{array}{ccc|c} 1 & -2 & -2 & -3 \\ -2 & 5 & 3 & 4 \\ 1 & -2 & -1 & -1 \end{array}} = -x .
 \end{array}$$

Hier kann man alles ablesen, was in den Fragen I, II, III gefragt wurde, z.B.: Das Lösungsgebilde des homogenen Systems ist eindimensional.; z kann man beliebig wählen.

V.3 Tableaus. Vollständiger Austausch

Gegeben seien zwei lineare Beziehungen

$$\begin{aligned}\alpha \cdot x^1 + a_{12} \cdot x^2 + a_{13} \cdot x^3 &= -y^1 \\ a_{21} \cdot x^1 + a_{22} \cdot x^2 + a_{23} \cdot x^3 &= -y^2 \quad .\end{aligned}$$

Wenn $\alpha \neq 0$, dann können wir x^1 und y^1 „austauschen“

$$\begin{aligned}\frac{1}{\alpha} \cdot y^1 + \frac{1}{\alpha} a_{12} \cdot x^2 + \frac{1}{\alpha} a_{13} \cdot x^3 &= -x^1 \\ -\frac{1}{\alpha} a_{21} \cdot y^1 + b_{22} \cdot x^2 + b_{23} \cdot x^3 &= -y^2\end{aligned}$$

wobei

$$b_{22} = a_{22} - a_{21} \cdot \frac{1}{\alpha} \cdot a_{12} \quad , \quad b_{23} = a_{23} - a_{21} \cdot \frac{1}{\alpha} \cdot a_{13} \quad .$$

Ebenso geht das für Gleichungssysteme mit m Gleichungen. Wir schreiben die Beziehungen als Zeilentableau:

x^1	x^{\parallel}		y^1	x^{\parallel}	
α	$a_2^1 \quad \dots \quad a_n^1$	$= -y^1$	$\frac{1}{\alpha}$	$\frac{1}{\alpha} a_2^1 \quad \dots \quad \frac{1}{\alpha} a_n^1$	$= -x^1$
a_1^2	\vdots		$-\frac{1}{\alpha} a_1^2$	\vdots	
\vdots	\vdots		\vdots	\vdots	
\vdots	A_2^2	$= -y^{\parallel}$	\vdots	B_2^2	$= -y^{\parallel}$
\vdots	\vdots		\vdots	\vdots	
a_1^m	\vdots		$-\frac{1}{\alpha} a_1^m$	\vdots	

Der Übergang heißt der Austausch zur Pivot-Position (1,1). Einen solchen Übergang gibt es offenbar zu jeder Pivot-Position, in welcher ein Eintrag $\neq 0$ steht. Die Unbestimmten an den Rändern werden ausgetauscht.

Bemerke : Wenn man zweimal um dieselbe Position austauscht, kommt man zum Ausgangstableau zurück. (Pivot ist das englische Wort für Dreh- und Angelpunkt.)

Wir wollen möglichst viele x -Variablen gegen y -Variablen austauschen. Die Sache kommt zu einem Ende, wenn im nichtausgetauschten Bereich nur noch Nullen stehen. Ein solches Tableau nennen wir ein vollständig ausgetauschtes Tableau. Nach einer geeigneten Umnummerierung der Indexmengen hat ein vollständig ausgetauschtes Zeilentableau die Gestalt

y^{\parallel}	x^{\parallel}	
$B_{\parallel}^{\parallel}$	$B_{\parallel}^{\parallel}$	$= -x^{\parallel}$
\vdots	\vdots	
$B_{\parallel}^{\parallel}$	$B_{\parallel}^{\parallel}$	$= -y^{\parallel}$

mit $B_{\parallel}^{\parallel} = 0$.

Erstes Zahlenbeispiel

Wir pivotieren zuerst um die Position (1, 1) und dann um die Position (2, 2)

$$\begin{array}{cccc|l}
 x^1 & x^2 & z^1 & z^2 & \\
 \hline
 1^* & 2 & 4 & 1 & = -y^1 \\
 1 & 1 & 1 & 0 & = -y^2 \\
 -1 & 0 & 2 & 1 & = -y^3
 \end{array}
 \qquad
 \begin{array}{cccc|l}
 y^1 & x^2 & z^1 & z^2 & \\
 \hline
 1 & 2 & 4 & 1 & = -x^1 \\
 -1 & -1^* & -3 & -1 & = -y^2 \\
 1 & 2 & 6 & 2 & = -y^3
 \end{array}$$

$$\begin{array}{cccc|l}
 y^1 & y^2 & z^1 & z^2 & \\
 \hline
 -1 & 2 & -2 & -1 & = -x^1 \\
 1 & -1 & 3 & 1 & = -x^2 \\
 -1 & 2 & 0 & 0 & = -y^3
 \end{array}
 \qquad
 \begin{array}{cc|cc|l}
 y^1 & & & z & \\
 \hline
 B_I^I & \vdots & B_{II}^I & & = -x^1 \\
 \dots\dots\dots & & & & \\
 -1 & 2 & \vdots & 0 & 0 & = -y^3
 \end{array}$$

Aus dem vollständig ausgetauschten Tableau können wir schnell ablesen, was wir über das ursprüngliche Gleichungssystem wissen wollten, nämlich

$$A \begin{pmatrix} x^1 \\ z \end{pmatrix} = -y^1.$$

- I. Lösungen existieren, wenn y die letzte Zeile erfüllt $(-1) \cdot y^1 + 2 \cdot y^2 = -y^3$.
- II. Eine spezielle Lösung $\begin{pmatrix} x^1 \\ 0 \end{pmatrix}$ gibt es zu jedem guten y^1 , nämlich $-x = B_I^I y^1$.
- III. Die allgemeine Lösung $\begin{pmatrix} x^1 \\ z \end{pmatrix}$ ist $-x = B_I^I y^1 + B_{II}^I \cdot z$.

Der vollständige Austausch wird zu einem Algorithmus, wenn eine Regel für die Pivot-Wahl gegeben ist. Wir wollen uns da aber nicht festlegen.

Anmerkung zur Pivot-Wahl

A sei eine $I \times J$ -Matrix vom Rang r . Eine Folge $(i_1, j_1), (i_2, j_2), \dots, (i_r, j_r)$ nennen wir eine günstige Pivot-Wahl, wenn jede der Teilmatrizen vom Format

$$\{i_1, \dots, i_k\} \times \{j_1, \dots, j_k\} \quad , \quad k = 1, \dots, r$$

„bequem“ invertierbar ist.

Bemerke

Für jemanden, der exakt rechnen kann, ist jede Folge $(i_1, j_1), \dots, (i_r, j_r)$, für welche die entsprechenden $k \times k$ -Matrizen (für $k = 1, \dots, r$) invertierbar sind, eine mögliche Pivot-Wahl. Solche Folgen gibt es nach Satz 4 in V.2.

Die Numeriker wissen, dass es „schlecht konditionierte“ Matrizen gibt; das sind Matrizen, bei welchen eine einigermaßen zuverlässige Inversion hohen Rechenaufwand erfordert. Mit

dieser Problematik wollen wir uns nicht befassen. In unseren Zahlenbeispielen konnten wir die Pivots entlang der Diagonale wählen. Wir würden die Wahl anderer Pivot-Position bevorzugen, wenn wir nach $(k-1)$ Austauschschritten in der Position (k, k) einen sehr kleinen Eintrag antreffen, wie im folgenden Zahlenbeispiel:

Drittes Zahlenbeispiel

Wir betrachten das Gleichungssystem zur gestörten Matrix

$$A_\varepsilon = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 + \varepsilon & 2 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Die inverse Matrix A_ε^{-1} sollte nahe bei A_0^{-1} liegen. Durch schlechte Pivot-Wahl würden aber problematische Zwischenresultate auftreten; die linke obere 2×2 -Matrix ist nämlich nahe an einer singulären Matrix. Wir wählen $(1, 1)$, $(3, 3)$, $(2, 2)$ als unsere Pivots.

$$\begin{array}{ccc|l} x^1 & x^2 & x^3 & \\ \hline 1^* & 1 & 0 & = -y^1 \\ 1 & 1 + \varepsilon & 2 & = -y^2 \\ 0 & \frac{1}{2} & \frac{1}{2} & = -y^3 \end{array} \qquad \begin{array}{ccc|l} y^1 & x^2 & x^3 & \\ \hline 1 & 1 & 0 & = -x^1 \\ -1 & \varepsilon & 2 & = -y^2 \\ 0 & \frac{1}{2} & \frac{1}{2}^* & = -y^3 \end{array}$$

$$\begin{array}{ccc|l} y^1 & x^2 & y^3 & \\ \hline 1 & 1 & 0 & = -x^1 \\ -1 & (\varepsilon - 2)^* & -4 & = -y^2 \\ 0 & 1 & 2 & = -x^3 \end{array} \quad \text{und für} \quad \begin{array}{ccc|l} y^1 & y^2 & y^3 & \\ \hline \frac{1}{2} & \frac{1}{2} & -2 & = -x^1 \\ \frac{1}{2} & -\frac{1}{2} & 2 & = -x^2 \\ -\frac{1}{2} & \frac{1}{2} & 0 & = -x^3 \end{array}$$

$\varepsilon = 0$ weiter

Die Inverse von A_ε ergibt sich ganz harmlos als eine Matrix in der Nähe von

$$A_0^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -4 \\ 1 & -1 & 4 \\ -1 & 1 & 0 \end{pmatrix}$$

Spaltentableau Wir haben oben ein Gleichungssystem durch ein Zeilentableau ausgedrückt. Ebenso interessant ist das Spaltentableau. Diesem wollen wir hier aber eine andere Interpretation geben.

Gegeben sei ein System von I-Spalten $\{w_j : j \in J\}$. Der i -te Eintrag in w_j sei a_j^i . Wenn wir mit u_i die i -te Einheitsspalte bezeichnen, dann haben wir

$$w_j = \sum u_i \cdot a_j^i \quad \text{für } j \in J.$$

Dies schreiben wir als ein Spaltentableau

$$\begin{array}{c} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{array} \boxed{\begin{array}{c} A \end{array}} \\ = \mathbf{w}_1 \quad \dots \quad = \mathbf{w}_n$$

Wenn $\mathbf{a}_1^1 \neq 0$, dann ist $\{\mathbf{w}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ eine Basis und die weiteren \mathbf{w}_j werden mittels des ausgetauschten Tableaus durch die neue Basis ausgedrückt. Das haben wir bereits in V1 gesehen. (Steinitz'sches Austauschlemma). Wir wollen möglichst viele \mathbf{u}_i gegen passende \mathbf{w}_j austauschen. Damit gelangen wir zum vollständig ausgetauschten Tableau.

$$\begin{array}{c} \mathbf{w}_1 \\ \dots \\ \mathbf{u}_\parallel \end{array} \boxed{\begin{array}{ccc} B_\parallel^1 & \vdots & B_\parallel^1 \\ \dots & \dots & \dots \\ B_\parallel^\parallel & \vdots & 0 \end{array}} \\ = \mathbf{u}_\parallel \quad \dots \quad = \mathbf{w}_\parallel$$

Das r -Tupel \mathbf{w}_\parallel ist eine Basis von $\text{span}\{\mathbf{w}_j : j \in J\}$. Die Matrix B_\parallel^1 zeigt, wie die übrigen \mathbf{w}_j aus den Elementen dieser Basis linear kombiniert werden:

$$\mathbf{w}_\parallel \cdot B_\parallel^1 = \mathbf{w}_\parallel \quad .$$

Auch die übrigen Blöcke haben interessante Interpretationen. Bevor wir das ausführen, wollen wir einen matrizentheoretischen Satz über alle möglichen Zwischenergebnisse beweisen.

Satz

Durch eine Serie von Pivot-Transformationen sei ein äquivalentes Tableau entstanden. Aus der $(I^1 + I^\parallel) \times (J^1 + J^\parallel)$ -Matrix A sei die $(I^1 + I^\parallel) \times (I^1 + J^\parallel)$ -Matrix B entstanden. Es gilt dann für die Blöcke

$$(i) \quad B_\parallel^1 = (A_\parallel^1)^{-1}$$

$$(ii) \quad B_\parallel^1 = (A_\parallel^1)^{-1} \cdot A_\parallel^\parallel \quad B_\parallel^\parallel = -A_\parallel^\parallel \cdot (A_\parallel^1)^{-1}$$

$$(iii) \quad B_\parallel^\parallel = A_\parallel^\parallel - A_\parallel^\parallel \cdot (A_\parallel^1)^{-1} \cdot A_\parallel^1$$

Beweis

- Wir bemerken, dass die Regeln dieselben sind wie die beim einmaligen Austausch eines 2×2 -Tableaus; man muss allerdings auf die Formate achten und auf die Reihenfolge der Faktoren

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \rightsquigarrow \begin{pmatrix} \frac{1}{\alpha} & \frac{1}{\alpha} \cdot \beta \\ -\gamma \cdot \frac{1}{\alpha} & \delta - \gamma \cdot \frac{1}{\alpha} \cdot \beta \end{pmatrix}$$

2. Wir orientieren uns am Spaltentableau. Aus der Basis $(u_{\perp}, u_{\parallel})$ des Raums der I-Spalten ist die Basis $(w_{\perp}, u_{\parallel})$ hervorgegangen.

$$\begin{array}{c}
 u_{\perp} \\
 \vdots \\
 u_{\parallel}
 \end{array}
 \begin{array}{|c|c|c|}
 \hline
 A_{\perp}^{\perp} & \vdots & A_{\parallel}^{\perp} \\
 \hline
 \dots\dots\dots & & \\
 \hline
 A_{\perp}^{\parallel} & \vdots & A_{\parallel}^{\parallel} \\
 \hline
 \end{array}
 \begin{array}{c}
 w_{\perp} \\
 \vdots \\
 u_{\parallel}
 \end{array}
 \begin{array}{|c|c|c|}
 \hline
 B_{\perp}^{\perp} & \vdots & B_{\parallel}^{\perp} \\
 \hline
 \dots\dots\dots & & \\
 \hline
 B_{\perp}^{\parallel} & \vdots & B_{\parallel}^{\parallel} \\
 \hline
 \end{array}$$

$$\begin{array}{cc}
 = w_{\perp} & = w_{\parallel}
 \end{array}
 \qquad
 \begin{array}{cc}
 = u_{\perp} & = w_{\parallel}
 \end{array}$$

3. $u_{\perp} A_{\perp}^{\perp} + u_{\parallel} \cdot A_{\perp}^{\parallel} = w_{\perp}$; $w_{\perp} \cdot B_{\perp}^{\perp} + u_{\parallel} \cdot B_{\perp}^{\parallel} = u_{\perp}$.

Wir multiplizieren mit der nichtsingulären Matrix B_{\perp}^{\perp}

$$u_{\perp} A_{\perp}^{\perp} B_{\perp}^{\perp} + u_{\parallel} \cdot A_{\perp}^{\parallel} B_{\perp}^{\perp} = w_{\perp} B_{\perp}^{\perp} = u_{\perp} - u_{\parallel} \cdot B_{\perp}^{\parallel}.$$

Da $(u_{\perp}, u_{\parallel})$ eine Basis ist, haben wir für die Matrizen

$$A_{\perp}^{\perp} \cdot B_{\perp}^{\perp} = E \text{ und } A_{\perp}^{\parallel} B_{\perp}^{\perp} = -B_{\perp}^{\parallel}.$$

Damit sind (i) und (ii) bewiesen.

4. Die späteren Spalten liefern den Beweis von (iii)

$$\begin{aligned}
 u_{\perp} \cdot A_{\parallel}^{\perp} + u_{\parallel} \cdot A_{\parallel}^{\parallel} &= w_{\parallel} = w_{\perp} \cdot B_{\parallel}^{\perp} + u_{\parallel} \cdot B_{\parallel}^{\parallel} = \\
 &= (u_{\perp} \cdot A_{\perp}^{\perp} + u_{\parallel} \cdot A_{\perp}^{\parallel}) \cdot B_{\parallel}^{\perp} + u_{\parallel} \cdot B_{\parallel}^{\parallel} \\
 &= u_{\perp} \cdot A_{\perp}^{\perp} B_{\parallel}^{\perp} + u_{\parallel} \cdot (A_{\perp}^{\parallel} B_{\parallel}^{\perp} + B_{\parallel}^{\parallel}) .
 \end{aligned}$$

Der vollständige Austausch löst mehrere Probleme. A sei eine $I \times J$ -Matrix.

- Indem wir Pivots wählen $(i_1, j_1) \dots, (i_r, j_r)$ finden wir invertierbare Teilmatrizen. Die Teilmatrix vom Format $\{i_1, \dots, i_k\}$ ist für jedes $k \leq r$ invertierbar. Im Laufe der Rechnung wird für jedes A_k die inverse Matrix bestimmt. Aus der Sicht der Anwender ist, wie gesagt, die Pivot-Wahl eine heikle Sache. Man möchte ja nicht irgendwelche invertierbare Teilmatrizen identifizieren und invertieren, sondern nach Möglichkeit solche, welche überzeugende Einsichten vermitteln. Für die Numerik ist es wichtig, dass die Rundungsfehler nicht explodieren.
- Nehmen wir an, wir hätten Partitionen der Indexmengen gewählt, $I = I^{\perp} + I^{\parallel}$, $J = J^{\perp} + J^{\parallel}$ mit $|I^{\perp}| = r = |J^{\perp}|$. Wir haben dann zum Anfangstableau vom Format $(I^{\perp} + I^{\parallel}) \times (J^{\perp} + J^{\parallel})$ ein vollständig ausgetauschtes Tableau vom Format $(J^{\perp} + I^{\parallel}) \times (I^{\perp} + J^{\parallel})$.

$$\begin{array}{c}
 x^{\perp} \quad x^{\parallel} \\
 \xi_{\perp} \begin{array}{|c|c|c|}
 \hline
 A_{\perp}^{\perp} & \vdots & A_{\parallel}^{\perp} \\
 \hline
 \dots\dots\dots & & \\
 \hline
 A_{\perp}^{\parallel} & \vdots & A_{\parallel}^{\parallel} \\
 \hline
 \end{array} = -y^{\perp} \\
 \xi_{\parallel} \begin{array}{|c|c|c|}
 \hline
 A_{\perp}^{\perp} & \vdots & A_{\parallel}^{\perp} \\
 \hline
 \dots\dots\dots & & \\
 \hline
 A_{\perp}^{\parallel} & \vdots & A_{\parallel}^{\parallel} \\
 \hline
 \end{array} = -y^{\parallel} \\
 = \eta_{\perp} \quad = \eta_{\parallel}
 \end{array}
 \qquad
 \begin{array}{c}
 y^{\perp} \quad x^{\parallel} \\
 y_{\perp} \begin{array}{|c|c|c|}
 \hline
 B_{\perp}^{\perp} & \vdots & B_{\parallel}^{\perp} \\
 \hline
 \dots\dots\dots & & \\
 \hline
 B_{\perp}^{\parallel} & \vdots & 0 \\
 \hline
 \end{array} = -x^{\perp} \\
 \xi_{\parallel} \begin{array}{|c|c|c|}
 \hline
 B_{\perp}^{\perp} & \vdots & 0 \\
 \hline
 \dots\dots\dots & & \\
 \hline
 B_{\perp}^{\parallel} & \vdots & 0 \\
 \hline
 \end{array} = -y^{\parallel} \\
 = \xi_{\perp} \quad = \eta^{\parallel}
 \end{array}$$

3. Man denkt meistens an das Zeilentableau. Von daher erklären sich die folgenden Namen: Die $J^I \times J^I$ -Matrix $K = B_{||}^I$ heißt der **Kernblock**; . Die $(I^I \times I^I)$ -Matrix $B_{|}^I$, nennt man den **Bildblock**. $B_{|}^I = (A_{|}^I)^{-1}$ heißt der **Inversblock** für den gegebenen vollständigen Austausch.

Wir haben gesehen

$$B_{||}^I = (A_{|}^I)^{-1} \cdot A_{||}^I, \quad B_{|}^I = -A_{|}^I \cdot (A_{|}^I)^{-1}$$

$$B_{||}^I = 0 = A_{||}^I - A_{|}^I \cdot (A_{|}^I)^{-1} \cdot A_{||}^I.$$

Damit ist der Satz bewiesen.

Satz

Zur Matrix A betrachten wir die Abbildungen

$$\varphi : x \mapsto Ax; \quad \varphi^* : \xi \mapsto \xi A.$$

Die Blöcke, die sich bei einem vollständigen Austausch ergeben, können dann folgendermaßen interpretiert werden:

- a) Die Spalten der $J \times J^I$ -Matrix $\begin{pmatrix} B_{||}^I \\ -E \end{pmatrix}$ sind eine Basis von $\ker \varphi$.
- b) Die Spalten der $I \times I^I$ -Matrix $\begin{pmatrix} -E \\ B_{|}^I \end{pmatrix}$ sind eine Basis von $\operatorname{im} \varphi$.
- c) Die Zeilen der $I^I \times I$ -Matrix $(B_{|}^I : E)$ sind linear unabhängige Linearformen, die auf $\operatorname{im} \varphi$ verschwinden.
- d) Die Zeilen der $J^I \times J$ -Matrix $(E : B_{||}^I)$ sind linear unabhängige Linearformen, die auf $\ker \varphi$ verschwinden.

Beweis

Man kann die Behauptungen aus dem Tableau ablesen; man kann aber auch innerhalb des Matrizenkalküls argumentieren.

1. Wir beweisen a) und d) zusammen. Die Spalten der Matrix in a) sind linear unabhängig; die Matrix hat den Rang $n - r$. Die Matrix $(E : B_{||}^I)$ aus d) hat den Rang r . Es gilt trivialerweise

$$(E : B_{||}^I) \begin{pmatrix} B_{||}^I \\ -E \end{pmatrix} = 0.$$

Wir zeigen, dass $(E:B_{||}^I)$ die Elemente aus $\ker \varphi$ annulliert; damit ist dann alles bewiesen. Die J -Spalte $x = \begin{pmatrix} x^I \\ x_{||}^I \end{pmatrix}$ wird genau dann von A annulliert, wenn sie von den ersten r Zeilen annulliert wird. Es gilt

$$\begin{aligned} (\mathbf{A}_{\parallel}^{\dagger} \mathbf{A}_{\parallel}^{\dagger})_{(\mathbf{x}_{\parallel}^{\dagger})} &= 0 \Leftrightarrow (\mathbf{A}_{\parallel}^{\dagger})^{-1} (\mathbf{A}_{\parallel}^{\dagger} \mathbf{A}_{\parallel}^{\dagger})_{(\mathbf{x}_{\parallel}^{\dagger})} = 0 \\ &\Leftrightarrow (\mathbf{E}, \mathbf{B}_{\parallel}^{\dagger})_{(\mathbf{x}_{\parallel}^{\dagger})} = 0 \end{aligned}$$

2. In derselben Weise beweist man b) und c). Dort haben wir es mit I-Zeilen in $(B \parallel E)$ zu tun und mit I-Spalten in der Matrix aus b).

$$(\mathbf{B}_\parallel^{\parallel} \vdash \mathbf{E}) \cdot \begin{pmatrix} -\mathbf{E} \\ \mathbf{B}_\parallel^{\parallel} \end{pmatrix} = 0 \quad .$$

Das Bild $\text{im } \varphi$ ist der Spaltenraum von A , aber auch der Spaltenraum von

$$\begin{pmatrix} \mathbf{A}_{||} \\ \mathbf{A}_{\perp} \end{pmatrix} \text{ und von } \begin{pmatrix} \mathbf{A}_{||} \\ \mathbf{A}_{\perp} \end{pmatrix} \cdot (\mathbf{A}_{\perp})^{-1} = \begin{pmatrix} \mathbf{E} \\ -\mathbf{B}_{||} \end{pmatrix} .$$

Damit ist b) bewiesen, und auch c).

Erstes Zahlenbeispiel

$$\begin{array}{ccccc} & x^1 & & z & \\ \hline 1 & 2 & \vdots & 4 & 1 \\ 1 & 1 & \vdots & 1 & 0 \\ \hline & & & & \\ -1 & 0 & \vdots & 2 & 1 \\ \hline \end{array} = -y^1 \quad \begin{array}{ccccc} & y^1 & & z & \\ \hline -1 & 2 & \vdots & -2 & -1 \\ 1 & -1 & \vdots & 3 & 1 \\ \hline & & & & \\ -1 & 2 & \vdots & 0 & 0 \\ \hline \end{array} = -x^1$$

a) $\ker \varphi$ ist der Spaltenraum von $\begin{pmatrix} -2 & -1 \\ 3 & 1 \\ \dots\dots\dots \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$;

b) $\text{im}\varphi$ ist der Spaltenraum von $\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ \dots\dots\dots \\ -1 & +2 \end{pmatrix}$;

- c) $\text{im } \phi$ kann durch eine einzige Gleichung beschrieben werden:

$$y \in \operatorname{im} \varphi \Leftrightarrow (-1)y^1 + 2y^2 + y^3 = 0, \quad \text{Zeile } (-1, 2 \vdots 1).$$

d) Für die Beschreibung von $\ker \varphi$ braucht man zwei Gleichungen:

$$x \in \ker \varphi \Leftrightarrow \begin{pmatrix} 1 & 0 & \vdots & -2 & -1 \\ 0 & 1 & \vdots & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = 0 \quad .$$

Konsequenz

Gegeben seien irgendwelche I-Spalten. W sei der von ihnen aufgespannte Raum. Wir möchten eine Basis von W und ein linear unabhängiges System von Linearformen, dessen Nullstellengebilde W ist. Was ist zu tun? Wir müssen jedenfalls die Dimension von W bestimmen. Dazu bietet sich an, Pivots $(i_1, j_1), (i_2, j_2), \dots$ zu wählen, bis man bei einem vollständig ausgetauschten Tableau angelangt ist. Wir haben oben schon gesehen, dass die $\{w_j : j \in J'\}$ eine Basis von W bilden. Andererseits bilden die $\{g^i(\cdot) : i \in I'\}$ eine Basis des Nullraums W^\diamond . Wenn wir nun aber schon rechnen mussten, um den Rang der Matrix zu bestimmen, können wir auch bequemere Basen angeben: W ist der Spaltenraum von $\begin{pmatrix} B_{||}^| \\ -E \end{pmatrix}$, W^\diamond ist der Zeilenraum von $(E; B_{||}^|)$.

Fazit : Der vollständige Austausch ist geeignet, bequeme Basen des Spaltenraums (und des Zeilenraums) einer $I \times J$ -Matrix konkret zu bestimmen. Er ist auch geeignet, Basen für die Nullräume des Spaltenraums (und des Zeilenraums) zu bestimmen. Den Schlüssel liefern die Matrizen $B_{||}^|$ und $B_{|}^||$.

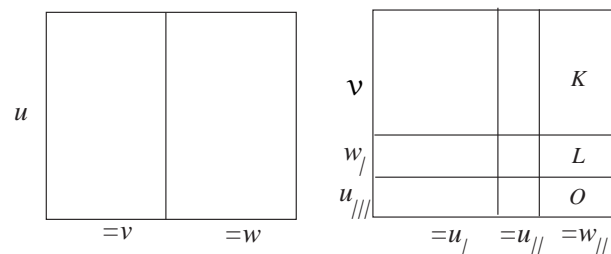
Der vollständige Austausch kann auch noch für andere Probleme verwendet werden, z.B.

Aufgabe : Seien $\{v_j : j \in J\}$ und $\{w_k : k \in K\}$ linear unabhängige Systeme von I-Spalten. Gesucht ist eine Basis des Durchschnitts

$$W_0 = \text{span}\{v_j\} \cap \text{span}\{w_k\} \quad .$$

Konstruktion

Wir schreiben die v und w in ein Spaltentableau, tauschen alle v_j aus und dann noch möglichst viele der w_k . Der Kernblock zeigt uns eine Basis des Durchschnitts



Die ausgetauschten w seien w_1, \dots, w_p , die restlichen seien w_{p+1}, \dots, w_{p+q} . Für $k = p+1, \dots, p+q$ haben wir

$$w_k = \sum v_j \cdot k_k^j + \sum_{m=1}^p w_m \cdot \ell_k^m \quad .$$

Die Vektoren

$$\tilde{v}_k = \sum v_j \cdot k_k^j = w_k - \sum_1^p w_m \ell_k^m \quad k = p+1, \dots, p+q$$

sind eine Basis des Durchschnitts.

Zweites Zahlenbeispiel

v_1, v_2, w_1, w_2 seien die Spalten der Matrix, die wir oben als zweites Zahlenbeispiel eingeführt haben. Die aufgespannten Räume sind Ebenen im \mathbb{R}^3 , die sich in einer Gerade schneiden (da sie nicht parallel sind). Diese Gerade kann durch einen aufspannenden Vektor beschrieben werden.

$$\begin{array}{c} u \\ \begin{array}{|c|c|c|c|} \hline 1 & 2 & 4 & 1 \\ \hline 1 & 1 & 1 & 0 \\ \hline -1 & 0 & 1 & 2 \\ \hline \end{array} \\ = v_1 \quad = v_2 \quad = w_1 \quad = w_2 \end{array} \qquad \begin{array}{c} v_1 \\ v_2 \\ w_1 \\ \begin{array}{|c|c|c|c|} \hline 1 & -2 & -2 & -3 \\ \hline -2 & 5 & 3 & 4 \\ \hline 1 & -2 & -1 & -1 \\ \hline \end{array} \\ = u_1 \quad = u_2 \quad = u_3 \quad = w_2 \end{array}$$

Die nichtausgetauschte Spalte sagt

$$w_2 = -3v_1 + 4v_2 - w_1 \quad .$$

Der Vektor

$$\tilde{v} = -3v_1 + 4v_2 = w_2 + w_1 = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$$

spannt den Durchschnitt auf.

Verkürzte Rechnung

Wenn man ein Gleichungssystem $Ax = b$ für ein festes b lösen will, dann muss man nicht den ganzen vollständigen Austausch leisten. Zunächst einmal verwandelt man das Problem in ein homogenes Tableau.

$$\begin{array}{c} x \qquad -1 \\ \begin{array}{|c|c|c|} \hline \vdots & & \\ \hline A & \vdots & b \\ \hline \vdots & & \\ \hline \end{array} = 0 \end{array}$$

Wir beginnen mit einer Pivot-Wahl (i_1, j_1) . Alle $m \cdot (n+1)$ Einträge erfahren eine Neuberechnung („update“, Aktualisierung). Wenn die (erweiterte) Matrix den Rang r hat, dann erfordert der vollständige Austausch $m(n+1) \cdot r$ updates. (Die zusätzliche Spalte kommt natürlich nicht als Pivot-Spalte in Betracht). So viele updates werden aber nicht wirklich gebraucht.

Nehmen wir an, dass wir in den ersten $k-1$ Schritten $\{i_1, \dots, i_{k-1}\}$ gegen $\{j_1, \dots, j_{k-1}\}$

ausgetauscht haben, und jetzt um die Pivot-Position (i_k, j_k) „drehen“ wollen. Um den Eintrag a_{mn} auf den neuen Stand zu bringen, brauchen wir gewisse Einträge in Zeile i_k und in der Spalte j_k . Wenn wir nur die späteren Spalten aktualisieren wollen, dann brauchen wir keine Information über die Einträge in den früheren $k-1$ Spalten. Das Aktualisieren, welches am Ende nur dazu dienen soll, die letzte Spalte des vollständig ausgetauschten Tableaus zu berechnen, wird immer weniger aufwendig.

Zweites Zahlenbeispiel

$$\begin{pmatrix} 1 & 2 & 4 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} ; \quad \begin{array}{cccc|l} x^1 & x^2 & x^3 & -1 & \\ \hline 1^* & 2 & 4 & 1 & = 0 \\ 1 & 1 & 1 & 0 & = 0 \\ -1 & 0 & 1 & 2 & = 0 \end{array}$$

Das Aktualisieren ergibt :

$$\begin{array}{cccc|l} 0 & x^2 & x^3 & -1 & \\ \hline \vdots & 2 & 4 & 1 & = -x^1 \\ \vdots & -1^* & -3 & -1 & = 0 \\ \vdots & 2 & 5 & 3 & = 0 \end{array} \quad \begin{array}{cccc|l} 0 & 0 & x^3 & -1 & \\ \hline \vdots & \vdots & -2 & -1 & = -x^1 \\ \vdots & \vdots & 3 & 1 & = -x^2 \\ \vdots & \vdots & -1^* & 1 & = 0 \end{array} \quad \begin{array}{cccc|l} 0 & 0 & 0 & -1 & \\ \hline \vdots & \vdots & \vdots & -3 & = -x^1 \\ \vdots & \vdots & \vdots & 4 & = -x^2 \\ \vdots & \vdots & \vdots & -1 & = -x^3 \end{array}$$

Somit haben wir die Lösung $x = \begin{pmatrix} -3 \\ 4 \\ 1 \end{pmatrix}$ mit $9 + 6 + 3 = 18$ Aktualisierungen gefunden.

V.4 Eliminationsmethode; LU-Faktorisierung

Wir diskutieren ein weiteres Verfahren zur „Lösung“ linearer Gleichungssysteme, das sog. Gauss'sche Eliminationsverfahren. Es gibt davon mehrere Varianten; wir betrachten diejenige, die auf die sog. LU-Faktorisierung führt. (LU steht für „lower-upper“, dabei denkt man an untere und obere Dreiecksmatrizen.)

Beim Gauss'schen Verfahren geht es um ein I -Tupel von homogenen linearen Gleichungen auf dem Raum \mathbb{K}_{Sp}^I . Die $g^i(\cdot)$ sind durch J -Zeilen gegeben; sie werden zu einer $I \times J$ -Matrix A zusammengefasst. Man sucht die Lösungsmenge $\{x : Ax = 0\}$. Bei der Eliminationsmethode geht man vom System $\{g^i(\cdot) : i \in I\}$ durch sog. elementare Zeilentransformationen zu äquivalenten Systemen $\{h^k(\cdot) : k \in K\}$ über, welche denselben Lösungsraum haben, für welche aber die Lösungsmenge leichter überschaubar ist. Das Verfahren geht davon aus, dass die Zeilen und die Spalten angeordnet sind. Im Laufe der Rechnung unternimmt man dann aber u.U. Vertauschungen von Zeilen oder Spalten. Man kann dieses Anordnen offenbar als eine Folge von Pivot-Wahlen sehen $(i_1, j_1), (i_2, j_2), \dots$ deuten. Eindeutige Regeln für die Pivot-Wahl werden zum Gauss-Verfahren in der Regel nicht angegeben; insofern ist die Gauss'sche Eliminationsmethode ebenso wenig wie der vollständige Austausch ein Algorithmus (im strengen Sinn). Am übersichtlichsten ist die Folge der Operationen, wenn I und J von vorneherein so angeordnet sind, dass die Pivot-Wahl $(1, 1), (2, 2), \dots, (r, r)$ möglich (oder sogar empfehlenswert) ist. In diesem Falle funktioniert das Verfahren so:

Die erste Zeile der $m \times n$ -Matrix A bleibt unangetastet. (Man geht davon aus, dass der Eintrag in der Position $(1, 1)$ eine Zahl $\alpha_1 \neq 0$ ist.) Von den weiteren Zeilen wird ein Vielfaches der ersten Zeile subtrahiert, sodass alle die entstehenden Zeilen an erster Stelle den Eintrag $= 0$ haben. Die zweite Zeile bleibt von nun an unangetastet. (Man geht davon aus, dass sie in der Diagonalposition $(2, 2)$ einen Eintrag $\alpha_2 \neq 0$ hat.) Von den späteren Zeilen wird nun ein Vielfaches dieser neuen zweiten Zeile subtrahiert, sodass die entstehenden Zeilen auch in der zweiten Position den Eintrag 0 haben. Die dritte Zeile bleibt von nun an unangetastet usw.

Wenn A den Rang r hat und alle Teilmatrizen vom Format $\{1, 2, \dots, k\} \times \{1, 2, \dots, k\}$ für $1 \leq k \leq r$ nichtsingulär sind, dann entsteht nach r Schritten eine Matrix der Gestalt

$$\begin{pmatrix} \tilde{A} & M \\ 0 & 0 \end{pmatrix},$$

wo \tilde{A} eine obere Dreiecksmatrix vom Format $r \times r$ ist.

Erstes Zahlenbeispiel

$$\begin{aligned}
A &= \begin{pmatrix} 1^* & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 2 & 1 \end{pmatrix} \\
A^{(1)} &= \begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & -1 & -3 & -1 \\ 0 & 2 & 6 & 2 \end{pmatrix} \\
A^{(2)} &= \begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & -1 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \tilde{A} & M \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Für eine 4-Spalte x gilt

$$Ax = 0 \Leftrightarrow A^{(1)} \cdot x = 0 \Leftrightarrow A^{(2)} \cdot x = 0 .$$

Schreiben wir $x = \begin{pmatrix} x^1 \\ z \end{pmatrix}$, so bedeutet das

$$\tilde{A}x^1 + Mz = 0 \quad \text{oder} \quad x^1 = -\tilde{A}^{-1} \cdot Mz .$$

Die 2-Spalte z kann beliebig gewählt werden; dazu gibt es dann genau ein x^1 , sodass

$$A \begin{pmatrix} x^1 \\ z \end{pmatrix} = 0 .$$

Da \tilde{A} eine obere Dreiecksmatrix ist, kann man die Einträge x^r, x^{r-1}, \dots, x^1 rekursiv zu den Einträgen von Mz bestimmen („Back-Substitution“).

Man kann alternativ das Eliminationsverfahren von unten her neu beginnen. Manche Bücher sprechen in diesem Fall vom Gauss-Jordan-Verfahren.

Beispiel a) $z = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, Mz = \begin{pmatrix} -4 \\ +3 \end{pmatrix}$

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} + \begin{pmatrix} -4 \\ +3 \end{pmatrix} = 0$$

also $x^2 = +3$ und $+4 = x^1 + 2x^2 = x^1 + 6, x^1 = -2$.

Beispiel b) $z = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, Mz = \begin{pmatrix} -1 \\ +1 \end{pmatrix}$

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} + \begin{pmatrix} -1 \\ +1 \end{pmatrix} = 0$$

also $x^2 = +1$ und $+1 = x^1 + 2x^2 = x^1 + 2$; $x^1 = -1$.

Damit haben wir eine Basis der Lösungsmenge, nämlich

$$\{(-2, +3, -1, 0)^T, \quad (-1, +1, 0, -1)^T\}.$$

Zweites Zahlenbeispiel

$$A = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 2 \end{pmatrix}$$

$$A^{(2)} = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & -1 & -3 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} = (\tilde{A}; M) \text{ mit } M = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Zu jeder Zahl z gibt es genau eine 3-Spalte x^l mit $\begin{pmatrix} x^l \\ z \end{pmatrix}$ aus der Lösungsmenge.

Bestimmen wir sie zu $z = -1$.

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & -3 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

also $x^3 = -1$, $-1 = -x^2 - 3x^3 = -x^2 + 3$; $x^2 = 4$.

$x^1 + 8 - 4 = 1$; $x^1 = -3$.

Damit haben wir eine Lösung des homogenen Systems, nämlich

$$(-3, 4, -1, -1)^T.$$

Dreiecksmatrizen

Zeilenoperationen bedeuten bekanntlich die Multiplikation von links. Bei unseren elementaren Zeilenoperationen wird mit unteren Dreiecksmatrizen multipliziert, die in der Diagonalen Einsen haben. Wir nennen sie normierte untere Dreiecksmatrizen.

Satz

Sei A eine $m \times n$ -Matrix vom Rang r , deren Teilmatrizen am oberen linken Ende bis zur Größe $r \times r$ invertierbar sind.

Es existiert dann eine normierte untere Dreiecksmatrix L , sodass

$$L \cdot A = \begin{pmatrix} \tilde{A} & M \\ 0 & 0 \end{pmatrix},$$

wo \tilde{A} eine nichtsinguläre obere Dreiecksmatrix von der Größe $r \times r$ ist.

Die linke obere Ecke von L ist eindeutig bestimmt.

Beweis :

Die Existenz ergibt sich aus dem angegebenen Eliminationsverfahren. Die Menge der normierten unteren Dreiecksmatrizen ist eine Gruppe.

Die Eindeutigkeitsaussage beweisen wir zunächst für $r = m$.

$$\begin{aligned} L \cdot A &= (\tilde{A}, M), & K \cdot A &= (\tilde{B}, N) \\ L^{-1} \cdot \tilde{A} &= A = K^{-1} \cdot \tilde{B} \\ K \cdot L^{-1} &= \tilde{B} \cdot \tilde{A}^{-1} \end{aligned}$$

Die linke Seite ist eine normierte untere, die rechte Seite ist eine obere Dreiecksmatrix. Also

$$K \cdot L^{-1} = E = \tilde{B} \cdot \tilde{A}^{-1}, \quad K = L, \quad \tilde{A} = \tilde{B}.$$

Im allgemeinen Fall haben wir

$$\begin{pmatrix} L_{\perp} & 0 \\ L_{\parallel} & L_{\parallel} \end{pmatrix} \begin{pmatrix} A_{\perp} & A_{\parallel} \\ A_{\perp} & A_{\parallel} \end{pmatrix} = \begin{pmatrix} \tilde{A} & M \\ 0 & 0 \end{pmatrix}$$

also $L_{\perp} \cdot A_{\perp} = \tilde{A}$, und daraus ist L_{\perp} eindeutig zu bestimmen.

Die Kurzfassung des Satzes heißt der

Satz von der eindeutigen LU-Faktorisierung

Die Zeilen und die Spalten einer nichtsingulären Matrix kann man so anordnen, dass die so angeordnete Matrix A in eindeutiger Weise faktorisiert werden kann.

$$A = L^{-1} \cdot D \cdot U,$$

wo D eine Diagonalmatrix, L eine spezielle untere und U eine normierte obere Dreiecksmatrix ist.

Inhomogene Systeme

Das inhomogene Gleichungssystem $Ax = b$ verwandeln wir wie oben (Ende V.3) in ein Tableau mit der erweiterten Matrix $(A:b)$. die Lösungsmenge ist dieselbe wie die zum folgenden Tableau

$$\begin{array}{ccc|ccc} & x_{\perp} & x_{\parallel} & & & -1 \\ \hline \tilde{A} & \vdots & M & \vdots & \tilde{b} & \\ \dots & \dots & \dots & \dots & \dots & \\ 0 & \vdots & 0 & \vdots & c & \end{array} = 0 \quad \text{wobei} \quad \begin{pmatrix} \tilde{b} \\ c \end{pmatrix} = Lb.$$

$$A \begin{pmatrix} x_{\perp} \\ x_{\parallel} \end{pmatrix} - b = 0 \Leftrightarrow LA \begin{pmatrix} x_{\perp} \\ x_{\parallel} \end{pmatrix} - Lb = 0.$$

- I) Das System $Ax = b$ ist genau dann lösbar, wenn $c = 0$, wenn also Lb von der Stelle $r + 1$ an Einträge $= 0$ hat, d.h. wenn $(A:b)$ denselben Rang hat wie A .
- II) Für ein gutes b bekommen wir eine spezielle Lösung aus der impliziten Gleichung $\tilde{A}x^I = \tilde{b}$.
- III) Zu vorgegebenem x^I bestimmt sich die Lösung $\begin{pmatrix} x^I \\ x^{II} \end{pmatrix}$ aus $\tilde{x}^I + Mx^{II} - \tilde{b} = 0$.
Hier kann man wieder mit Rück-Substitution arbeiten.

Unbestimmte rechte Seite

Statt $Ax = -y$ schreiben wir

$$\begin{array}{cc} x & y \\ \boxed{A \quad \vdots \quad E} = 0 \end{array}$$

Äquivalent ist $L(Ax + y) = 0$, also

$$\begin{array}{ccc} x^I & x^{II} & y \\ \boxed{\begin{array}{ccc} \tilde{A} & : & M & : \\ \dots\dots\dots & & & : & L \\ 0 & : & 0 & : \end{array}} \begin{array}{l} = 0 \\ \\ = 0 \end{array} \end{array}$$

Die Lösbarkeitsbedingung lautet $Ly = \begin{pmatrix} \tilde{y} \\ 0 \end{pmatrix}$.

Zu einem guten y ergibt sich für jedes x^{II} eine Lösung $\begin{pmatrix} x^I \\ x^{II} \end{pmatrix}$ aus der durch Rückwärtssubstitution leicht lösbaren Gleichung

$$\tilde{A}x^I + Mx^{II} + \tilde{y} = 0 .$$

Im Spezialfall linear unabhängiger Zeilen von A :

$$\boxed{\tilde{A}x^I + Mx^{II} + Ly = 0 .}$$

Bemerke : Man kann weiter mit \tilde{A}^{-1} multiplizieren und erhält die „explizite“ Lösung

$$x^I + (\tilde{A}^{-1}M)x^{II} + (A_1^I)^{-1} \cdot y = 0 .$$

Zweites Zahlenbeispiel

$$\begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x^I \\ z \end{pmatrix} = \begin{pmatrix} -y^1 \\ -y^2 \\ -y^3 \end{pmatrix}$$

Wir haben elementare Zeilenumformungen vorgenommen.

$$L = \begin{pmatrix} 1 & \cdot & \cdot \\ 0 & 1 & \cdot \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdot & \cdot \\ -1 & 1 & \cdot \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \cdot & \cdot \\ -1 & 1 & \cdot \\ -1 & 2 & 1 \end{pmatrix}$$

um zur oberen Dreiecksmatrix \tilde{A} zu gelangen

$$(\tilde{A}:\mathbf{M}) = L \cdot A = \begin{pmatrix} 1 & 2 & 4 & \vdots & 1 \\ 0 & -1 & -3 & \vdots & -1 \\ 0 & 0 & -1 & \vdots & 1 \end{pmatrix}$$

Die Gleichung

$$\tilde{A} \cdot \mathbf{x}^I + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \cdot z + L \cdot \mathbf{y} = 0$$

kann durch „Rückwärtssubstitution“ nach \mathbf{x}^I aufgelöst werden.

Man kann die Gleichung aber auch mit \tilde{A}^{-1} multiplizieren; d.h. man kann durch Zeilenoperationen den Anfang der langen Matrix $(\tilde{A}:\mathbf{M}:L)$ auf Diagonalgestalt bringen. In unserem Fall ergibt sich

$$\mathbf{x}^I + \begin{pmatrix} -3 \\ 4 \\ -1 \end{pmatrix} \cdot z + \begin{pmatrix} 1 & -2 & -2 \\ -2 & 5 & 3 \\ 1 & -2 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \\ \mathbf{y}^3 \end{pmatrix} = 0 \quad .$$

Die Matrix am Ende ist $\tilde{A}^{-1} \cdot L = (A^I)^{-1}$. Das ist eine Matrix, die wir oben durch vollständigen Austausch berechnet haben, um ein Durchschnittsproblem zu lösen. Als Zeilentableau stellt sich die „Lösung“ des Gleichungssystems so dar:

	y	z	
1	-2	-2	\vdots
-2	5	3	\vdots
1	-2	-1	\vdots

-3	4	-1
----	---	----

=

$-\mathbf{x}^1$

=

$-\mathbf{x}^2$

=

$-\mathbf{x}^3$

Bemerkungen zum praktischen Lösen von Gleichungssystemen Der vollständige Austausch erfordert im Beispiel $3 \cdot 12 = 36$ Aktualisierungen. Beim abgekürzten Verfahren sind es 18 Aktualisierungen. Es ist nicht viel nutzloses Rechnen im Spiel. Beim Eliminationsverfahren braucht man (wenn man nicht so genau hinsieht) ungefähr ebenso viele. Ein großer Vorteil der Tableaus scheint die Übersichtlichkeit zu sein; in jeder Etappe der Rechnung sind alle Informationen in knapper Form auf dem Tisch. (Man

denke z.B. an die Zeilen- oder Spaltenumordnungen). Der Grund dafür, dass die elementaren Lehrbücher die Eliminationsmethode bevorzugen, ist eigentlich unverständlich. Eine löbliche Ausnahme ist

Gewirtz, A.; Sitomer, H.; Tucker, A.W.: *Constructive Linear Algebra*, Prentice-Hall Inc., N. Jersey, 1974

Eine andere Sache sind die Erfahrungen der Numeriker. Bei speziellen Typen von Matrizen empfehlen sie mit guten Gründen spezielle Methoden der Matrizeninversion. Dabei ist in vielen Fällen aus der Natur- und Ingenieur-Wissenschaften überhaupt kein finites Verfahren angezeigt; man nähert sich mit geeigneten Iterationsverfahren der Lösung einer Gleichung $Ax = b$. Bei Problemen wie dem folgenden würde ein Praktiker wohl kaum mit dem Eliminationsverfahren arbeiten.

Aufgabe

Sei $q(\cdot)$ eine quadratische Funktion auf einem hochdimensionalen \mathbb{R}^n .

$$q(x) = a_0 + a \cdot x + \frac{1}{2}x^T \cdot A \cdot x$$

mit einer positiv definiten Matrix A und einer vorgegebenen Zeile a .

Gesucht ist die (eindeutig bestimmte!) Minimalstelle \tilde{x} , d.h. dasjenige \tilde{x} für welches gilt

$$q(x) = \text{const} + \frac{1}{2}(x - \tilde{x})^T \cdot A \cdot (x - \tilde{x}) .$$

Das Problem kann als ein lineares Gleichungssystem verstanden werden. Wegen

$$\frac{1}{2}(x - \tilde{x})^T \cdot A \cdot (x - \tilde{x}) = \frac{1}{2}x^T \cdot A \cdot x - (\tilde{x}^T \cdot A)x + \frac{1}{2}\tilde{x}^T \cdot A \cdot \tilde{x}$$

kommt es nämlich darauf an, \tilde{x} so zu bestimmen, dass $-\tilde{x}^T \cdot A = a$, also $A\tilde{x} = -a^T$. Die Tatsache, dass A positiv definit ist, erlaubt viel effektivere Herangehensweisen als die oben betrachteten Eliminationsmethoden.

V.5 Matrizen vom Rang r , Singulärwertzerlegung

Im Abschnitt I.4 ‘Die Algebra der komplexen $n \times n$ Matrizen’ haben wir uns hauptsächlich mit nichtsingulären Matrizen befasst. Es wurde gesagt, dass es für nichtsinguläre Matrizen eine Reihe von übersichtlichen eindeutigen Faktorisierungen gibt. Wir werden nach und nach sehen, wie sich die Aussagen (entsprechend verkompliziert) auf nichtquadratische Matrizen übertragen lassen. Die Dualität erweist sich dabei als ein entscheidendes Moment.

Hier wollen wir zunächst eine *additive* Zerlegung einer $m \times n$ Matrix vom Rang r diskutieren. Wir werden beweisen: Jede $m \times n$ Matrix A vom Rang r lässt sich als Summe von r Matrizen vom Rang 1 schreiben. Es wird sich dann bald zeigen, dass es auch da wieder um Faktorisierungen geht.

Eine **Matrix A vom Rang 1** hat (grob gesagt) lauter zueinander proportionale Zeilen und lauter zueinander proportionale Spalten. Wenn in der Position (i_0, j_0) ein Eintrag $\alpha \neq 0$ steht, $\mathbf{b} = (b^i)$ die j_0 -te Spalte und $\mathbf{c} = (c_j)$ die i_0 -te Zeile ist, dann sind die Einträge von A gleich

$$a_j^i = \frac{1}{\alpha} \cdot b^i \cdot c_j. \quad (\text{Matrix von der Form } \textit{Spalte} \times \textit{Zeile}).$$

Im Matrizenkalkül versteht man gewöhnlich die Spalten als Vektoren und die Zeilen als Linearformen (auch Covektoren genannt). Das Produkt in dem hier vorliegenden Sinn nennt man das Tensorprodukt (des Vektors \mathbf{b} mit dem Covektor \mathbf{c}). Unsere Matrix A vom Rang 1 ist in dieser Sprache ein contra-covarianter Tensor vom Rang 1.

Wir bemerken, dass eine Summe solcher Matrizen vom Rang 1 als ein Matrizenprodukt geschrieben werden kann:

$$\sum_k b_{(k)}^i \cdot c_j^{(k)} = (B \cdot C)_j^i,$$

wo B die $m \times r$ -Matrix mit den Spalten $\mathbf{b}_{(k)}$ und C die $r \times n$ -Matrix mit den Zeilen $\mathbf{c}^{(k)}$ ist.

Dualität Wir vermeiden hier die übliche Bezeichnung V^* für den Dualraum von V und die Bezeichnung φ^* für die duale Abbildung, weil $*$ bereits allzusehr strapaziert wurde, z. B. bei der hermiteschen Konjugation. Statt des hochgestellten Stern $*$ benützen wir den hochgestellten Dolch † , also V^\dagger bzw. φ^\dagger . In der Matrizenrechnung haben wir

$$\begin{aligned} V &= \mathbb{K}_{\text{Sp}}^n, & V^\dagger &= \mathbb{K}_{\text{Z}}^n, & W &= \mathbb{K}_{\text{Sp}}^m, & W^\dagger &= \mathbb{K}_{\text{Z}}^m; \\ \varphi(x) &= Ax, & \varphi^\dagger(\xi) &= \xi A. \\ \langle \xi, \varphi(x) \rangle &= \xi Ax = \langle \varphi^\dagger(\xi), x \rangle. \end{aligned}$$

Verschiedene Bezeichnungsweisen Sei $\varphi: V \rightarrow W$ eine lineare Abbildung vom Rang 1

- Im Tensorkalkül schreibt man die Abbildung φ in der Form

$$\varphi = \frac{1}{\alpha} \cdot w \otimes n \quad \text{mit } w \in W \text{ und } n \in V^\dagger,$$

und man liest die Formel: $\frac{1}{\alpha}$ ‘mal’ w ‘tensor’ n ‘ergibt’ φ .

- Im Matrizenkalkül versteht man

$$A = \frac{1}{\alpha} \cdot b \cdot c \quad \text{in dem Sinn: } m\text{-Spalte ‘mal’ } n\text{-Zeile ‘ergibt’ } m \times n\text{-Matrix.}$$

- Wir bevorzugen für das Folgende eine Notation, welche der beliebten Dirac-Notation in der Hilbertraumtheorie nachempfunden ist (siehe unten):

$$\varphi = \frac{1}{\alpha} \cdot \langle \cdot, w \rangle \langle n, \cdot \rangle.$$

Die Abbildung $\varphi: V \rightarrow W$ bildet den Vektor $v \in V$ ab in den Vektor

$$\varphi(v) = \frac{1}{\alpha} w \cdot \langle n, v \rangle.$$

Die duale Abbildung $\varphi^\dagger: W^\dagger \rightarrow V^\dagger$ bildet den Covektor m ab in den Covektor

$$\varphi^\dagger(m) = \frac{1}{\alpha} \langle m, w \rangle \cdot n.$$

In der Notation mit den Spitzklammern hat unser Zerlegungstheorem die Gestalt

Theorem

Sei A eine $m \times n$ -Matrix vom Rang r . Es existieren dann m -Spalten $b_{(1)}, b_{(2)}, \dots, b_{(r)}$ sowie n -Zeilen $c^{(1)}, c^{(2)}, \dots, c^{(r)}$, sodass gilt

$$A = \sum_{k=1}^r \langle \cdot, b_{(k)} \rangle \langle c^{(k)}, \cdot \rangle.$$

Die Zeilen $c^{(k)}$ sind durch A und die Kollektion der Spalten $b_{(k)}$ eindeutig bestimmt.

Das Zerlegungstheorem lässt sich folgendermassen verschärfen

Sei A eine $m \times n$ -Matrix vom Rang r ; und sei B eine $m \times r$ -Matrix mit demselben Spaltenraum. Es existiert dann genau eine Matrix C , sodass $A = B \cdot C$.

Beweis Seien $\mathbf{b}_{(k)}$ die Spalten der Matrix B . Jeder Bildvektor $A\mathbf{v}$ besitzt eine eindeutige Darstellung

$$A\mathbf{v} = \sum \mathbf{b}_{(k)} \cdot \mathbf{y}^k.$$

Die Koeffizienten \mathbf{y}^k hängen in linearer Weise von \mathbf{v} ab. Wir haben also Linearformen $\mathbf{c}^{(k)}$, sodass

$$A\mathbf{v} = \sum \mathbf{b}_{(k)} \cdot \langle \mathbf{c}^{(k)}, \mathbf{v} \rangle = \sum \langle \cdot, \mathbf{b}_{(k)} \rangle \langle \mathbf{c}^{(k)}, \mathbf{v} \rangle.$$

Genauso ergibt sich der duale Satz: Man kann sich eine $r \times n$ -Matrix C mit demselben Zeilenraum wie A vorgeben; und findet genau eine $m \times r$ -Matrix B , sodass $A = B \cdot C$. Die symmetrische Formulierung lautet

Theorem

Sei A eine $m \times n$ -Matrix vom Rang r ; sei M eine $m \times r$ -Matrix mit demselben Spaltenraum und N eine $r \times n$ -Matrix mit demselben Zeilenraum. Es existiert dann genau eine $r \times r$ -Matrix E , sodass gilt $A = M \cdot E \cdot N$.

Spezialfälle In den Anwendungen wird man sich natürlich nicht irgendein r -Tupel von Spalten oder irgendein r -Tupel von Zeilen vorgeben.

1. Sei z.B. φ eine diagonalisierbare Abbildung ($\varphi: V \rightarrow V$) mit den Eigenwerten λ_k und den Eigenvektoren \mathbf{v}_k . Sei $\{ \mathbf{n}^k \}$ die duale Basis. Es gilt dann

$$\varphi = \sum \lambda_k \langle \cdot, \mathbf{v}_k \rangle \langle \mathbf{n}^k, \cdot \rangle.$$

Zum Beweis müssen wir nur zeigen, dass die Eigenvektoren \mathbf{v}_k die richtigen Bilder haben, nämlich $\lambda_k \cdot \mathbf{v}_k$. Das aber liegt auf der Hand.

2. Besonders wichtig ist der Fall, wo die vorgegebene $n \times n$ Matrix H unitär diagonalisierbar ist. Sei etwa H positiv semidefinit mit orthonormierten Eigenvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$ zu den Eigenwerten $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$. Es gilt dann

$$H = \sum \lambda_k \cdot \mathbf{u}_k \cdot \mathbf{u}_k^* \quad (\text{in Matrixschreibweise}).$$

Wenn man die Summe nur bis r erstreckt, dann erhält man eine Approximation von H durch eine Matrix vom Rang $\leq r$. (Dies ist die Idee der sog. Hauptkomponentenanalyse für eine Kovarianzmatrix in der statistischen Datenanalyse.)

Bemerkung zur Dualität

Die Faktorisierungen gemäss dem Theorem (oder seinem Dual) $A = B \cdot C$ bedeuten für die zu A gehörige Spaltenabbildung $\varphi(\cdot)$ ($\varphi(\mathbf{v}) = A\mathbf{v}$) eine Faktorisierung

$\varphi(\cdot) = \mu(\nu(\cdot))$, wobei gilt: ν ist surjektiv, μ ist injektiv,

Für die duale Abbildung φ^\dagger ($\varphi^\dagger(\xi) = \xi A$) haben wir die Faktorisierung

$\varphi^\dagger(\cdot) = \nu^\dagger(\mu^\dagger(\cdot))$, wobei gilt: ν^\dagger ist injektiv, μ^\dagger ist surjektiv.

Kernblock, Bildblock und Inversblock Sei A eine $m \times n$ Matrix vom Rang r , deren Zeilen und Spalten so angeordnet sind, dass die ersten r Spalten eine $m \times r$ Matrix M vom Rang r und die ersten r Zeilen eine $r \times n$ Matrix N vom Rang r bilden. Die invertierbare $r \times r$ Matrix in der linken oberen Ecke sei mit A_r bezeichnet. Das Faktorisierungstheorem und sein Dual sind anwendbar: Es existieren Matrizen K und L , sodass gilt

$$A = M \cdot (I_r, K) = \begin{pmatrix} I_r \\ L \end{pmatrix} A_r (I_r, K) = \begin{pmatrix} I_r \\ L \end{pmatrix} N = A ,$$

oder auch

$$A = M \cdot A_r^{-1} \cdot N .$$

Bemerke: Diese Darstellung einer Matrix A haben wir im Falle $r = 1$ oben bereits diskutiert. Wenn der Eintrag α in der linken oberen Ecke nicht verschwindet, dann gilt

$$A = \mathbf{b} \cdot \frac{1}{\alpha} \cdot \mathbf{c} , \quad \text{wo } \mathbf{b} \text{ die erste Spalte und } \mathbf{c} \text{ die erste Zeile ist.}$$

Die Matrizen A_r^{-1} , K , und L sind uns beim vollständigen Austausch bereits begegnet. Die folgenden Tableaus sind äquivalent:

$$\begin{array}{ccc} & x^\perp & x^\parallel \\ \xi_\perp & \boxed{A_r} & * \\ \xi_\parallel & * & * \\ & = \eta_\perp & = \eta_\parallel \end{array} \quad \begin{array}{ccc} & y^\perp & x^\parallel \\ \eta_\perp & \boxed{A_r^{-1}} & K \\ \xi_\parallel & -L & 0 \\ & = \xi_\perp & = \eta_\parallel \end{array} \quad \begin{array}{l} = -y^\perp \\ = -y^\parallel \end{array}$$

Der Kernblock K beschreibt den Kern der ‘Spaltenabbildung’ $\varphi: x \mapsto Ax$. In der Tat gilt

$$A \begin{pmatrix} x^\perp \\ x^\parallel \end{pmatrix} = 0 \iff N \begin{pmatrix} x^\perp \\ x^\parallel \end{pmatrix} = 0 \iff (I, K) \begin{pmatrix} x^\perp \\ x^\parallel \end{pmatrix} = 0 \iff -x^\perp = Kx^\parallel .$$

Die letzte Gleichung ist auch die Beschreibung des Kerns $\ker(\varphi)$ in der Tableaudarstellung.

Die Matrix L beschreibt einerseits das Bild $\text{im}(\varphi)$; sie beschreibt aber gleichzeitig auch den Kern der ‘Zeilenabbildung’ $\xi \mapsto \xi A$, d.h. den Kern der dualen Abbildung $\ker(\varphi^\dagger)$. In der Tat gilt

$$(\xi_\perp, \xi_\parallel)A = 0 \iff (\xi_\perp, \xi_\parallel)M = 0 \iff (\xi_\perp, \xi_\parallel) \begin{pmatrix} I_r \\ L \end{pmatrix} = 0 \iff \xi_\perp = -\xi_\parallel L .$$

Die letzte Gleichung ist auch die Beschreibung von $\ker(\varphi^\dagger)$ in der Tableaudarstellung.

Hinweis zur Numerik Es sei A eine (reelle oder komplexe) $m \times n$ Matrix, welche (möglicherweise) ‘nah’ an einer Matrix von einem kleinen Rang r liegt. Man möchte A durch eine Matrix vom Rang r approximieren. Dazu kann man das Austauschverfahren heranziehen. Man wählt nach und nach die Pivot-Positionen $(i_1, j_1), (i_2, j_2), \dots$ so, dass der Eintrag in der Pivot-Position jeweils dem Betrag nach maximal ist. Wenn nach r Austauschschritten die $(m-r) \times (n-r)$ Matrix R der verbleibenden Einträge dem Betrag nach ‘klein’ sind, hören wir auf. (Die Zeilen und Spalten denken wir uns aus typographischer Bequemlichkeit so angeordnet, dass die Pivotierungen entlang der Diagonalen anfallen.) Es gilt

$$A = M \cdot A_r^{-1} \cdot N + \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix} = \begin{pmatrix} I_r \\ L \end{pmatrix} \cdot A_r \cdot (I_r, K) + \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$$

Es bleibt natürlich die Frage, welche Restmatrizen R als ‘klein’ gelten können.

Ausserdem stellt sich die Frage, ob es geschickt ist, darauf zu bestehen, dass die approximierende Matrix vom Rang r einfach gewisse Zeilen und Spalten aus der gegebenen Matrix A übernehmen soll.

Hinweis: Bei der sogenannten Singulärwertzerlegung einer komplexen Matrix A geht man ganz anders vor; man verlangt, dass in der Zerlegung $A = Q \cdot \Lambda \cdot P$ die Matrix Λ eine Diagonalmatrix mit positiven Einträgen ist, und dass Q orthonormierte Spalten, P orthonormierte Zeilen besitzt; $Q^*Q = I_r$, $PP^* = I_r$.

Dass es eine solche Zerlegung gibt, werden wir unten einerseits algebraisch beweisen und andererseits geometrisch. Bevor wir uns diesen Beweisen zuwenden, wollen wir aber eine Notation besprechen, welche die Verhältnisse in Hilberträumen sehr übersichtlich macht.

Dirac's Notation

Sei $(V, \langle \cdot | \cdot \rangle)$ ein Hilbertraum für die Quantenmechanik. Die Elemente v liefern einerseits die sog. ket-Vektoren $|v\rangle$, welche für die möglichen Zustände des Systems stehen; andererseits liefern sie die sog. bra-Vektoren $\langle v|$, welche den Linearformen entsprechen.

Die Linearform $\langle w|$ angewandt auf $|v\rangle$ liefert das innere Produkt $\langle w|v\rangle$, die ‘bracket’. (Nach einem berühmten Satz von F. Riesz ist jede stetige Linearform auf einem Hilbertraum durch einen bra-Vektor gegeben.)

Die Entsprechung $|v\rangle \longleftrightarrow \langle v|$ zwischen dem gegebenen Hilbertraum und seinem Dualraum ist eine Isometrie, die allerdings kein Isomorphismus komplexer Vektorräume ist; sie ist nämlich nicht linear sondern $|\alpha v\rangle = \alpha \cdot |v\rangle \longleftrightarrow \langle v| \cdot \bar{\alpha} = \langle \alpha v|$.

In der Dirac'schen Notation bezeichnet $A = |w\rangle\langle u|$ einen eindimensionalen Operator; dieser ordnet jedem ket-Vektor $|v\rangle$ ein Vielfaches von $|w\rangle$ zu, nämlich $|w\rangle \cdot c$ mit $c = \langle u|v\rangle$. Andererseits ordnet dieses A jedem bra-Vektor $\langle v|$ ein Vielfaches von $\langle u|$ zu, nämlich $c \cdot \langle u|$ mit $c = \langle v|w\rangle$. Speziell ist für ein w mit $\|w\| = 1$ der Operator $|w\rangle\langle w|$ die orthogonale Projektion auf den von w aufgespannten eindimensionalen Vektorraum. Genauer gesagt:

$$|w\rangle\langle w|: \quad |v\rangle \longmapsto |w\rangle \cdot \langle w|v\rangle, \quad \langle u| \longmapsto \langle u|w\rangle \cdot \langle w|.$$

Beispiel: Seien w_1, w_2, \dots, w_n orthonormale Vektoren. Dann ist

$$A = |w_1\rangle\langle w_1| + \dots + |w_n\rangle\langle w_n| = \sum_{k=1}^n |w_k\rangle\langle w_k|$$

die orthogonale Projektion auf den von den w_k aufgespannten Teilvektorraum.

Operatoren und ihre Adjungierten

Sei $A: V \rightarrow W$ eine lineare Abbildung von einem Hilbertraum $(V, \langle \cdot | \cdot \rangle)$ in einen Hilbertraum $(W, \langle \cdot | \cdot \rangle)$. Man spricht von einem (beschränkten) linearen Operator.

Für $v \in V$ bezeichnet $A|v\rangle = |Av\rangle$ den ket-Vektor zum Bild von v ; auf der anderen Seite bezeichnet (für $w \in W$) $\langle w|A$ die Linearform mit

$$\langle w|A \text{ 'angewandt auf' } |v\rangle = \langle w|A|v\rangle = \langle w| \text{ 'angewandt auf' } A|v\rangle.$$

Definition Der adjungierte Operator ist die eindeutig bestimmte Abbildung

$$A^*: W \rightarrow V \quad \text{mit} \quad \langle A^*w| = \langle w|A.$$

Das bedeutet

$$\forall w \in W, v \in V \quad \langle A^*w|v\rangle = \langle w|A|v\rangle = \langle w|Av\rangle.$$

Eine lineare Abbildung eines Hilbertraums in sich mit $A = A^*$ heisst ein selbstadjungierter Operator.

Hinweis In der Funktionalanalysis betrachtet man auch unbeschränkte Operatoren, insbesondere unbeschränkte selbstadjungierte Operatoren. Dieser Typ von Operatoren spielt eine zentrale Rolle in der Theorie der quantenmechanischen Observablen.

Matrizen als Abbildungen von Hilberträumen Mittels einer beliebigen positivdefiniten $n \times n$ -Matrix Q können wir den Spaltenraum \mathbb{C}_{Sp}^n zu einem Hilbertraum machen, indem wir definieren: Für $x, y \in \mathbb{C}_{Sp}^n$ sei

$$\langle y|x\rangle = y^* \cdot Q \cdot x.$$

Betrachten wir als einfachsten Fall den Fall $Q = I$. In diesem Fall haben wir

$$\langle y|x\rangle = y^* \cdot x.$$

Jede $m \times n$ -Matrix A wirkt als lineare Abbildung auf die Spalten. $A: x \mapsto Ax$. Der adjungierte Operator ist durch die **hermitisch konjugierte** Matrix A^* gegeben; denn

$$y^* \cdot (Ax) = (A^* \cdot y)^* \cdot x = \langle y|A|x\rangle.$$

Die hermiteschen Matrizen liefern also die selbstadjungierten Operatoren (wohlgemerkt im Fall $Q = I$). Eine $n \times n$ -Matrix ist unitär, wenn ihre Spalten eine Orthonormalbasis bilden.

$U^* \cdot U = I$. Unser Beispiel von oben bedeutet für die Matrizenrechnung: Wenn $\{u_1, \dots, u_n\}$ eine Orthonormalbasis des Spaltenraums ist, dann ist $\sum_1^n u_k \cdot u_k^*$ die Einheitsmatrix. Die Summe $\sum_1^m u_k \cdot u_k^*$ ist die orthogonale Projektion auf den von $\{u_1, \dots, u_m\}$ aufgespannten Teilraum.

Allgemeiner gilt: Für beliebige komplexe Zahlen d_1, d_2, \dots ist

$$N = d_1 \cdot u_1 \cdot u_1^* + \dots + d_n \cdot u_n \cdot u_n^* = \sum |u_k\rangle d_k \langle u_k|$$

eine normale Matrix; und jede normale Matrix kann in dieser Form dargestellt werden. Die Spalten u_i liefern uns die unitäre Matrix, welche die unitäre Diagonalisierung von N leistet. Wenn nämlich D die Diagonalmatrix mit den Diagonalelementen d_i ist, dann haben wir

$$N = U \cdot D \cdot U^* \quad \text{und daher} \quad U^* \cdot N \cdot U = D.$$

Auf die unitäre Diagonalisierung einer beliebigen normalen Matrix N werden wir unten nochmals zu sprechen kommen. Zuerst beweisen wir das höchst bemerkenswerte

Theorem (Der Satz von der Singulärwertzerlegung)

Sei $A : V \rightarrow W$ eine lineare Abbildung von einem Hilbertraum in einen Hilbertraum, welche den Rang r besitzt. Dann gibt es Zahlen $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$, orthonormierte w_1, \dots, w_r in W und orthonormierte v_1, \dots, v_r in V , sodass gilt

$$A = |w_1\rangle \lambda_1 \langle v_1| + \dots + |w_r\rangle \lambda_r \langle v_r|.$$

Die λ_k sind eindeutig bestimmt. Im Falle, wo sie paarweise verschieden sind, sind auch die Einheitsvektoren w_k und v_k bis auf einen gemeinsamen Faktor vom Betrag 1 eindeutig bestimmt.

Algebraischer Beweis Im Spezialfall, wo A eine nichtsinguläre $n \times n$ Matrix ist, können wir den Beweis leicht mit Hilfe des Satzes von der unitären Diagonalisierbarkeit einer positivdefiniten Matrix führen. Der Satz von der positivdefiniten Quadratwurzel (siehe I.4) liefert zunächst einmal die Faktorisierung $A = RU$, wo U unitär und R die positivdefinite Quadratwurzel aus AA^* ist. ($U = R^{-1}A$ ist offensichtlich eine unitäre Matrix.) Wenn wir jetzt R unitär diagonalisieren, $V^*RV = \Lambda$, dann haben wir

$$R = V\Lambda V^*, \quad A = V\Lambda V^*U = Q \cdot \Lambda \cdot P,$$

mit unitären Matrizen Q , P und einer positivdefiniten Diagonalmatrix Λ .

Mit Hilfe des Zerlegungssatzes von oben erledigen wir den allgemeinen Fall: Wir wählen orthonormierte m -Spalten, welche eine Basis des Spaltenraums von A bilden und fassen sie zu einer $m \times r$ Matrix M zusammen; wir wählen orthonormierte n -Zeilen, welche den Zeilenraum von A aufspannen und fassen sie zu einer $r \times n$ Matrix N zusammen. Es gibt dann genau eine Matrix E , sodass $A = M \cdot E \cdot N$. Die nichtsinguläre Matrix E

schreiben wir als Produkt $E = U_1 \Lambda U_2$ mit unitären $r \times r$ -Matrizen U_1, U_2 und einer Diagonalmatrix Λ mit dem Ergebnis

$$A = MEN = MU_1 \Lambda U_2 N = Q \Lambda P \quad \text{mit} \quad Q^* Q = I_r = P P^* .$$

Wir bemerken: Die Spalten von Q sind Eigenvektoren für die positiv semidefinite $m \times m$ Matrix AA^* ; die Zeilen von P sind Eigenvektoren für die positiv semidefinite $n \times n$ Matrix A^*A . Triviales Rechnen ergibt in der Tat

$$(A \cdot A^*)Q = Q \cdot \Lambda^2, \quad P(AA^*) = \Lambda^2 \cdot P .$$

Die Tatsache weist übrigens auch einen Weg der Berechnung, wenn man die unitäre Diagonalisierung der Matrizen AA^* und A^*A beherrscht, (etwa mit Jacobi's Iterationsverfahren, siehe IV.6.) Hier wird auch klar: Wenn der Eigenwert λ_k^2 einfach ist, dann ist die entsprechende Spalte in Q sowie die entsprechende Zeile in P bis auf einen Faktor (vom Betrag 1) eindeutig bestimmt. Im allgemeinen Fall hat man darauf zu achten, dass man die Spalte und die Zeile im betreffenden Eigenraum zueinander passend wählt. Wir werden sehen, was das bedeutet, wenn wir unten das sog. Maximierungsargument von R. Courant in Gang setzen.

Der hier eben vorgestellte algebraische Existenzbeweis für die Singulärwertzerlegung von A benützt die unitäre Diagonalisierbarkeit von positiv semidefiniten Matrizen. Der folgende geometrische Beweis des Satzes von der Singulärwertzerlegung hat diese Schwäche nicht; er erfasst den Satz von der unitären Diagonalisierbarkeit einer normalen Matrix als einen Spezialfall. Die wesentliche Idee des Beweises steckt im folgenden Maximierungsargument (nach R. Courant). Wir brauchen es hier für lineare Abbildungen eines n -dimensionalen Hilbertraums V in einen m -dimensionalen Hilbertraum W ; es gilt in gleicher Weise für kompakte Operatoren in unendlichdimensionalen Hilberträumen.

Maximierungsargument

Sei $\varphi : V \rightarrow W$ eine lineare Abbildung eines Hilbertraums in einen Hilbertraum; und sei (w_1, v_1) ein Paar von Einheitsvektoren mit

$$\|\varphi\| = \langle w_1 | \varphi | v_1 \rangle, \quad \|w_1\| = 1 = \|v_1\| .$$

Es gilt dann

1. $\varphi(v_1)$ ist ein Vielfaches von w_1 , $\varphi(v_1) = \|\varphi\| \cdot w_1$.
2. Wenn v auf v_1 senkrecht steht, dann steht $\varphi(v)$ senkrecht auf w_1 .

Beweis Wir bemerken

$$\begin{aligned} \varphi(v) &= \sup\{ |\langle w | \varphi(v) \rangle| : \|w\| \leq 1 \}, \\ \|\varphi\| &= \sup\{ \|\varphi(v)\| : \|v\| \leq 1 \} = \sup\{ |\langle w | \varphi(v) \rangle| : \|w\| \leq 1, \|v\| \leq 1 \} \end{aligned}$$

Sei (w_1, v_1) ein maximierendes Paar und $w = w_1 + \alpha \tilde{w}$ mit $\|\tilde{w}\| = 1$ und $\tilde{w} \perp w_1$. Der folgende Bruch ist dem Betrage nach $\leq \|\varphi\|$ für alle α :

$$\frac{\langle w | \varphi | v_1 \rangle}{\|w\|} = \frac{\|\varphi\| + \bar{\alpha} \cdot \langle \tilde{w} | \varphi(v_1) \rangle}{\sqrt{1 + |\alpha|^2}} .$$

Daher gilt $\tilde{w} \perp \varphi(v_1)$; jedes auf w_1 senkrechte \tilde{w} ist senkrecht auf $\varphi(v_1)$; w_1 ist proportional zu $\varphi(v_1)$, und zwar

$$\varphi(v_1) = \lambda_1 \cdot w_1 = \|\varphi\| \cdot w_1.$$

Es bezeichne V_1 den von v_1 aufgespannten Vektorraum und W_1 den von w_1 aufgespannten Vektorraum. Wir zeigen, dass φ den Orthogonalraum V_1^\perp in den Orthogonalraum W_1^\perp abbildet:

Sei $v = v_1 + \alpha \tilde{v}$ mit $\|\tilde{v}\| = 1$ und $\tilde{v} \perp v_1$.

Der folgende Bruch ist dem Betrage nach $\leq \|\varphi\|$ für alle α :

$$\frac{\langle w_1 | \varphi | v \rangle}{\|v\|} = \frac{\|\varphi\| + \alpha \cdot \langle w_1 | \varphi | \tilde{v} \rangle}{\sqrt{1 + |\alpha|^2}}.$$

Daraus folgt $\varphi(\tilde{v}) \perp w_1$. q.e.d.

Konstruktion der Singulärwertzerlegung

Sei wie oben

$$\lambda_1 = \|\varphi\| = \langle w_1 | \varphi | v_1 \rangle \quad \text{mit} \quad \|w_1\| = 1 = \|v_1\|,$$

und sei

$$\varphi_2 = \varphi - |w_1\rangle \lambda_1 \langle v_1|.$$

Die Abbildung φ_2 bildet V_1 in die Null ab und stimmt auf V_1^\perp mit φ überein. Es sei

$$\lambda_2 = \|\varphi_2\| = \langle w_2 | \varphi | v_2 \rangle \quad \text{mit} \quad \|w_2\| = 1 = \|v_2\|,$$

Die Abbildung

$$\varphi_3 = \varphi_2 - |w_2\rangle \lambda_2 \langle v_2| = \varphi - \sum_{k=1}^2 |w_k\rangle \lambda_k \langle v_k|$$

bildet den von $\{v_1, v_2\}$ aufgespannten Vektorraum in die Null ab; auf dem orthogonalen Komplement stimmt sie mit φ überein.

Wenn wir in dieser Weise weiterkonstruieren, dann landen wir nach r Schritten bei der Nullabbildung φ_r , und es ergibt sich

$$\varphi = \sum_{k=1}^r |w_k\rangle \lambda_k \langle v_k|.$$

Normale Matrizen Sei N eine normale Matrix, $N^*N = NN^*$. Die Singulärwertzerlegung liefert zunächst einmal

$$N = Q \cdot \Lambda \cdot P \quad \text{mit } \Lambda \text{ Diagonalmatrix und} \quad Q^*Q = I_r = PP^*$$

(Die nichtnegativen Diagonalelemente λ sind absteigend angeordnet).

Da N normal ist, liefert eine triviale Rechnung

$$NN^* = Q\Lambda^2Q^* = N^*N = P^*\Lambda^2P.$$

Daraus folgt insbesondere, dass P^* denselben Spaltenraum hat wie Q .

Wir überzeugen uns, dass sich die entsprechenden Spalten nur um Faktoren vom Betrag 1 unterscheiden: Die maximierenden Paare $(w_1, v_1), (w_2, v_2), \dots$ wurden schrittweise nach demselben Prinzip gewählt; es genügt daher zu zeigen, dass w_1 ein Vielfaches von v_1 ist, dass also v_1 ein Eigenvektor ist. Wie wir eben gesehen haben, sind v_1 und w_1 Eigenvektoren der Abbildung $\varphi \circ \varphi^*$, und zwar zum selben Eigenwert $\lambda_1^2 = \|\varphi\|^2$. Wenn dieser Eigenwert einfach ist, dann ist klar, dass sich w_1 und v_1 nur um einen Faktor unterscheiden. Im allgemeinen Fall führt ein Stetigkeitsargument zum Ziel, -welches wir hier aber nicht weiter ausführen wollen.

Wir haben somit gesehen: es existiert eine unitäre Diagonalmatrix E , sodass $P^* = Q \cdot E$. Mit der Diagonalmatrix $M = \Lambda \cdot E^*$ haben wir dann also

Theorem

Wenn N eine normale $n \times n$ Matrix vom Rang r ist, dann existiert eine komplexe Diagonalmatrix M und eine $n \times r$ Matrix Q , sodass

$$N = Q \cdot M \cdot Q^* \quad \text{und} \quad Q^* \cdot Q = I_r.$$

Übersetzen wir die Resultate nun in Dirac's Notation:

Die vorgegebene $m \times n$ Matrix A betrachten wir als eine Abbildung φ des Raums der n -Spalten in den Raum der m -Spalten, beide Räume mit dem üblichen Skalarprodukt ausgestattet. Die Spalten von P^* können wir als orthonormierte Vektoren v_k verstehen, die Spalten von Q als orthonormierte Vektoren w_k .

$$A = Q \cdot \Lambda \cdot P \quad \text{bedeutet dann} \quad \varphi = \sum |v_k\rangle \lambda_k \langle w_k|.$$

Für eine normale Matrix N ergibt sich $P^* = Q \cdot E$ und wir haben mit der komplexen Diagonalmatrix $M = \Lambda \cdot E^*$

$$N = Q \cdot M \cdot Q^* \quad \text{bzw.} \quad \psi = \sum |v_k\rangle \mu_k \langle v_k|.$$

Dieser Satz verallgemeinert offenbar den Satz von der Hauptachsentransformation einer hermiteschen Form. (siehe IV.2) Die unitäre Diagonalisierung der normalen Matrix N ergibt sich folgendermaßen: Das orthonormierte r -Tupel der Spalten von Q wird zu einer Orthonormalbasis ergänzt, sodass wir als eine unitäre Matrix U gewinnen, deren vorderer Abschnitt Q ist. Die Diagonalmatrix M wird zu einer Diagonalmatrix D vom Format $n \times n$ gemacht, indem wir sie mit Nullen auffüllen. Es gilt

$$N = Q \cdot M \cdot Q^* = U \cdot D \cdot U^* \quad \text{und} \quad U^* \cdot U = I_n = U \cdot U^*,$$

und daher $U^* \cdot N \cdot U = D$ (Diagonalmatrix).

V.6 Kleinste Quadrate und lineare Schätzung

Ein Eingangsthema der linearen Algebra ist bekanntlich das ‘Lösen’ linearer Gleichungssysteme. Wenn für ein \mathbf{y} die Gleichung $\mathbf{A} \cdot \mathbf{x} = \mathbf{y}$ keine Lösung besitzt, dann spricht man von einem überbestimmten Gleichungssystem. Man kann sich die Aufgabe stellen, ein $\hat{\mathbf{x}}$ zu finden, sodass gilt

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{y} + \text{Rest} \quad \text{mit} \quad \|\text{Rest}\| = \min.$$

In dem Falle, wo $\|\cdot\|$ eine Hilbertraum-Norm ist, wird diese Aufgabe durch die sog. Methode der kleinsten Quadrate gelöst. Wir wollen im Folgenden mit einer veränderten Notation arbeiten. Wir beweisen den

Satz

Gegeben sei eine positivdefinite $\mathbf{n} \times \mathbf{n}$ Matrix \mathbf{Q} und eine $\mathbf{n} \times \mathbf{p}$ Matrix \mathbf{B} vom Rang \mathbf{p} .

- Zu jeder \mathbf{n} -Spalte \mathbf{y} existiert dann genau eine \mathbf{p} -Spalte $\hat{\mathbf{s}}$, sodass gilt

$$\mathbf{y} = \mathbf{B} \cdot \hat{\mathbf{s}} + \mathbf{u} \quad \text{und} \quad \mathbf{u}^* \cdot \mathbf{Q} \cdot \mathbf{u} = \min,$$

- Die Abbildung $\mathbf{y} \mapsto \hat{\mathbf{y}} = \mathbf{B} \cdot \hat{\mathbf{s}}$ ist die orthogonale Projektion auf den Spaltenraum V_1 der Matrix \mathbf{B} , (orthogonal bezüglich der durch \mathbf{Q} gegebenen Norm).

Beweis

1. Wir betrachten auf dem Raum aller \mathbf{p} -Spalten \mathbf{s} die quadratische Funktion

$$q(\mathbf{s}) := \|\mathbf{y} - \mathbf{B} \cdot \mathbf{s}\|^2.$$

Sie besitzt eine Minimalstelle $\hat{\mathbf{s}}$. Mit der Bezeichnung $\hat{\mathbf{y}} = \mathbf{B} \cdot \hat{\mathbf{s}}$ haben wir für alle \mathbf{s}

$$\begin{aligned} q(\hat{\mathbf{s}}) &\leq q(\hat{\mathbf{s}} + \mathbf{s}) = \|\mathbf{y} - \mathbf{B} \cdot (\hat{\mathbf{s}} + \mathbf{s})\|^2 = \|(\mathbf{y} - \hat{\mathbf{y}}) - \mathbf{B} \cdot \mathbf{s}\|^2 \\ &= \|\mathbf{y} - \hat{\mathbf{y}}\|^2 - \langle \mathbf{y} - \hat{\mathbf{y}} | \mathbf{B} \cdot \mathbf{s} \rangle - \langle \mathbf{B} \cdot \mathbf{s} | \mathbf{y} - \hat{\mathbf{y}} \rangle + \|\mathbf{B} \cdot \mathbf{s}\|^2 \\ &= \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\mathbf{B} \cdot \mathbf{s}\|^2 - 2 \cdot \Re(\mathbf{s}^* \mathbf{B}^* \cdot \mathbf{Q} \cdot (\mathbf{y} - \hat{\mathbf{y}})). \end{aligned}$$

Dies impliziert $\mathbf{B}^* \mathbf{Q} (\mathbf{y} - \hat{\mathbf{y}}) = 0$, also

$$(\mathbf{B}^* \mathbf{Q} \mathbf{B}) \cdot \hat{\mathbf{s}} = \mathbf{B}^* \mathbf{Q} \cdot \mathbf{y}.$$

2. Da bei uns hier die $\mathbf{p} \times \mathbf{p}$ Matrix $\mathbf{B}^* \mathbf{Q} \mathbf{B}$ positiv definit ist, hängt die eindeutig bestimmte Lösung der Gleichung $\hat{\mathbf{s}}$ in linearer Weise von \mathbf{y} ab. Auch $\hat{\mathbf{y}} = \mathbf{B} \cdot \hat{\mathbf{s}}$ ist eine lineare Funktion von \mathbf{y} . Die lineare Abbildung $\pi : \mathbf{y} \mapsto \hat{\mathbf{y}}$ wird durch eine $\mathbf{n} \times \mathbf{n}$ Matrix \mathbf{P} dargestellt.

$$\mathbf{P} = \mathbf{B} \cdot (\mathbf{B}^* \mathbf{Q} \mathbf{B})^{-1} \cdot \mathbf{B}^* \mathbf{Q}.$$

Dass die lineare Abbildung $\pi(\cdot)$ eine Projektion ist, ist aus der Konstruktion klar; man kann aber $\pi \circ \pi = \pi$ natürlich auch im Matrizenkalkül verifizieren.

3. Die Orthogonalität (bzgl. der Q -Norm) bedeutet für die Projektion $\pi = \pi^* \circ \pi = \pi^*$ und für die darstellende Matrix $P^*QP = QP = P^*Q$. In der Tat

$$\begin{aligned} \forall w, v \quad \langle (\text{id} - \pi)w | \pi(v) \rangle &= 0 & \text{d. h.} \quad \pi^*(\text{id} - \pi) &= 0, \\ \forall w, v \quad w^*(I - P)^* QPv &= 0 & \text{d. h.} \quad (I - P^*) \cdot Q \cdot P &= 0. \end{aligned}$$

Hinweis: Eine Gleichung der Form

$$(B^*QB) \cdot \hat{s} = B^*Q \cdot y.$$

mit einer positiv semidefiniten Matrix Q und einer beliebigen Matrix B heisst (seit der Zeit von Gauss) eine Normalgleichung. Sie hat für jedes y eine Lösung, weil, wie man leicht sieht, die Matrizen B^*QB und B^*Q denselben Rang haben

Adjungierte Matrizen Sei Q eine positivdefinite $n \times n$ Matrix und R eine positiv definite $m \times m$ Matrix. Wir staten die Spaltenräume mit den dazugehörigen Hilbertraum-Strukturen aus.

$$(V, \langle \cdot | \cdot \rangle) = (\mathbb{C}_{\text{Sp}}^n, Q); \quad (W, \langle \cdot | \cdot \rangle) = (\mathbb{C}_{\text{Sp}}^m, R).$$

Die lineare Abbildung $\alpha: V \rightarrow W$ sei durch die $m \times n$ Matrix A dargestellt: $\alpha: v \mapsto w = Av$. Der adjungierte Operator α^* ist dann natürlich i. Allg. nicht durch die hermitisch konjugierte Matrix A^* dargestellt. Es gilt vielmehr

$$\alpha^*: w \mapsto v = Q^{-1}A^*Rw.$$

In der Tat gilt

$$\langle \alpha^*w | v \rangle = (Q^{-1}A^*Rw)^*Qv = w^*RAQ^{-1}Qv = w^*RAv = \langle w | Av \rangle.$$

Spezialfall: Sei α ein Endomorphismus des Hilbertraums $(\mathbb{C}_{\text{Sp}}^n, Q)$ und A die darstellende Matrix. α ist genau dann ein selbstadjungierter Operator, wenn QA eine hermitesche Matrix ist. Genau in diesem Fall gilt nämlich

$$Q^{-1}A^*Q = A, \quad \text{also} \quad A^*Q = QA = (QA)^*.$$

Beispiel (Ausgleichsgerade)

Messungen haben n Punkte im \mathbb{R}^2 ergeben,

$$(t_1, y_1), (t_2, y_2), \dots, (t_n, y_n),$$

die (mit Abweichungen in der y -Richtung) im Wesentlichen auf einer Geraden $y = \alpha + \beta t$ zu liegen scheinen. Welche ist die 'richtige', die 'wahre' Gerade? Wir wollen α und β , oder auch für gewisse \tilde{t} den Wert der Ausgleichsgeraden im Punkt \tilde{t} , also $\alpha + \beta \tilde{t}$ schätzen. Gemäss der Methode der kleinsten Quadrate wird verlangt

$$y_i = \alpha + \beta t_i + u_i, \quad \|u\| = \min.$$

Wir wählen die Norm $\|\mathbf{u}\|^2 = \sum q_i |\mathbf{u}_i|^2$, wo die q_i positive Zahlen sind (die zum Ausdruck bringen, wie unterschiedlich wir die Genauigkeit in den verschiedenen Punkten t_i veranschlagen). Ohne Beschränkung der Allgemeinheit können wir $\sum q_i = 1$ annehmen; zur Abkürzung setzen wir

$$\bar{t} = \sum q_i t_i, \quad \tau^2 = \sum q_i (t_i - \bar{t})^2 = \sum q_i t_i^2 - \bar{t}^2, \quad \bar{y} = \sum q_i y_i.$$

In Matrizenschreibweise haben wir

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}; \quad \sum q_i \cdot |u_i|^2 = \min.$$

In unserem Fall kann man die Gleichung für das optimale $\hat{\mathbf{s}} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ und ihre Lösung einfach hinschreiben:

$$(B^*QB) \cdot \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = B^*Qy = \begin{pmatrix} q_1 & \cdots & q_n \\ q_1 t_1 & \cdots & q_n t_n \end{pmatrix} \cdot y$$

$$B^*QB = \begin{pmatrix} 1 & \bar{t} \\ \bar{t} & \bar{t}^2 + \tau^2 \end{pmatrix}; \quad (B^*QB)^{-1} = \frac{1}{\tau^2} \cdot \begin{pmatrix} \bar{t}^2 + \tau^2 & -\bar{t} \\ -\bar{t} & 1 \end{pmatrix}$$

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \frac{1}{\tau^2} \cdot \begin{pmatrix} \bar{t}^2 + \tau^2 & -\bar{t} \\ -\bar{t} & 1 \end{pmatrix} \cdot \begin{pmatrix} q_1 & \cdots & q_n \\ q_1 t_1 & \cdots & q_n t_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Die zweite Zeile ergibt die geschätzte Steigung

$$\hat{\beta} = \frac{1}{\tau^2} \cdot (-\bar{t}q_1 + q_1 t_1, \dots, -\bar{t}q_n + q_n t_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \frac{1}{\tau^2} \cdot \sum q_i (t_i - \bar{t}) y_i.$$

Ferner entnehmen wir der expliziten Formel, dass die geschätzte Gerade im Punkt \bar{t} den Wert \bar{y} annimmt. In der Tat

$$\hat{\alpha} + \hat{\beta} \bar{t} = \frac{1}{\tau^2} \cdot (\tau^2, 0) \cdot B^*Qy = \bar{y}.$$

Fazit: Die Ausgleichsgerade geht durch den Schwerpunkt der Punktwolke ('Streudiagramm', 'scatterplot') mit dem 'empirischen Regressionskoeffizienten' als Steigung.

Hinweis: Regressionskoeffizienten sind ebenso wie auch Kovarianzen Termini aus der linearen Statistik; auf sie kommen wir weiter unten zu sprechen.

Vorausschau: Die Methode der kleinsten Quadrate in der hier vorgestellten Form liefert uns zu dem Zahlen- n -tupel \mathbf{y} das p -tupel $\hat{\mathbf{s}} = (\hat{\alpha}^1, \dots, \hat{\alpha}^p)^\top$. Die Zuordnung ist linear, die $\hat{\alpha}^j$ sind also Linearkombinationen der \mathbf{y}_i : $\hat{\alpha}^j = \hat{\xi}^{(j)} \cdot \mathbf{y}$ für $j = 1, \dots, p$. Den von den Zeilen $\hat{\xi}^{(j)}$ aufgespannten Vektorraum wollen wir im Folgenden noch aus einem veränderten Blickwinkel charakterisieren. Indem wir uns in die Gedankenwelt der BLUE-Schätzung (Best Linear Unbiased Estimators) begeben, gelangen wir zu einer dualen Betrachtung.

Der statistische Zugang

Während man in der Datenanalyse die Beobachtungswerte $\mathbf{y}_1, \dots, \mathbf{y}_n$ einfach als Zahlen versteht, interpretiert man sie in der Mathematischen Statistik als die Realisierungen von Zufallsgrößen Y_1, \dots, Y_n . Man stellt sich vor, dass diesen Zufallsgrößen eine (allerdings nicht vollständig bekannte) gemeinsame Verteilung zukommt. Ein stochastisches Modell wird durch die Annahme etabliert, dass die unbekannte gemeinsame Verteilung zu einer Klasse (oder Familie) von möglichen (oder besser ‘nominellen’) Verteilungen gehört. Mit den Methoden der mathematischen Statistik sucht man nun unter diesen nominellen Verteilungen diejenigen, welche am besten zu den beobachteten Realisierungen passen. Das letztendliche Ziel sind möglichst genaue Vorhersagen für (noch) nicht durchgeführte Messungen.

Beispiel (Wiederholte Messungen) Wir haben ein Werkstück mehrmals unabhängig mit gut justierten Präzisionswaagen gewogen und dabei Messwerte $\mathbf{y}_1, \dots, \mathbf{y}_n$ erhalten. Wir nehmen an, dass es sich um Realisierungen von unabhängigen normalverteilten Zufallsgrößen mit dem allen gemeinsamen unbekannten Mittelwert μ und den bekannten Varianzen σ_i^2 handelt. Welches Ergebnis erwarten wir, wenn wir eine weitere unabhängige Messung mit einer gut justierten Waage durchführen?

Gauss hat bewiesen, dass das gewichtete Mittel $\sum q_i \mathbf{y}_i$ mit $q_i = \text{const} \cdot \sigma_i^{-2}$ die beste Schätzung des wahren μ ist.

Hinweis: Gauss hat übrigens auch gezeigt, dass gewichtete Mittel keine optimalen Schätzer sind, wenn die zugrunde Fehlerverteilung nicht normal ist. Wenn beispielsweise die Y_i unabhängige doppeltexponentiell verteilte Zufallsgrößen mit derselben Varianz sind, dann ist der Stichprobenmedian der optimale Schätzer, und nicht etwa das Stichprobenmittel. (Das Stichprobenmittel ist aber natürlich ein besserer Schätzer als jedes andere lineare Schätzverfahren).

Die Sätze von Gauss können wir hier nicht beweisen; wir müssten zuerst einmal herausarbeiten, was es (im gegebenen Umfeld) heissen soll, dass ein Verfahren besser ist als ein anderes. Das jedoch überlassen wir der Mathematischen Statistik.

Im Folgenden werden wir nur ganz spezielle Schätzverfahren miteinander vergleichen, die sog. erwartungstreuen linearen Schätzer. Wir werden sehen, dass der Vergleich ein Anwendungsfall der Linearen Algebra ist. Konkret werden wir wieder auf die orthogonalen Projektionen geführt, die wir schon bei der Methode der kleinsten Quadrate gefunden haben.

Wir wollen aber zunächst etwas weiter ausholen. Wir benötigen den Begriff eines (reellen oder komplexen) Zufallsvektors. Als eine Konkretisierung werden wir auch den Begriff eines gauss’schen Zufallsvektors skizzieren.

Covarianzmatrizen

Ein n -tupel von reellwertigen oder komplexwertigen Zufallsgrößen Z_1, \dots, Z_n schreibt man in der Linearen Statistik als eine n -Spalte Z , und man spricht von einem Zufallsvektor der Länge n . In der Linearen Statistik geht man immer davon aus, dass die Zufallsgrößen endliche Varianz haben.

$$\sigma_i^2 = \text{var} Z_i = \mathbb{E}|Z_i|^2 - |\mathbb{E} Z_i|^2 = \mathbb{E}|Z_i - \mathbb{E} Z_i|^2 < \infty.$$

Die Kovarianzen σ_{ij} fasst man in der Covarianzmatrix zusammen. So ist also

$$\text{cov}(Z, Z) = \mathbb{E}(Z \cdot Z^*) - \mathbb{E} Z \cdot \mathbb{E} Z^*$$

die hermitesche $n \times n$ Matrix mit den Einträgen $\sigma_{ij} = \text{cov}(Z_i, Z_j)$. Allgemeiner definiert man für Zufallsvektoren W und Z der Längen m bzw. n die $m \times n$ -Covarianzmatrix

$$\text{cov}(W, Z) = \mathbb{E}(W \cdot Z^*) - \mathbb{E} W \cdot \mathbb{E} Z^*.$$

Wenn $\text{cov}(W, Z) = 0$, dann sagt man, die Zufallsvektoren seien unkorreliert.

Beachte: Das Funktional $\text{cov}(\cdot, \cdot)$ ist in beiden Argumenten additiv, es ist in dem Sinn sesquilinear, dass es im ersten Argument linear und im zweiten antilinear ist – eine bewährte Konvention bei komplexwertigen Zufallsvektoren. Aus der Sesquilinearität ergibt sich der

Satz

Sei $C = \text{cov}(W, Z)$ eine Covarianzmatrix vom Format $m \times n$, und seien A und B komplexwertige Matrizen vom Format $p \times m$ bzw. $q \times n$. Dann gilt

$$\text{cov}(A \cdot W, B \cdot Z) = A \cdot C \cdot B^*.$$

Semidefinitheit Die Covarianzmatrix eines Zufallsvektors Z ist immer positiv semidefinit; sie ist genau dann positiv definit, wenn die Verteilung auf keinen echten linearen Teilraum konzentriert ist. In der Tat, wenn C die Covarianzmatrix ist und ξ irgendeine n -Zeile, dann gilt

$$\xi \cdot C \cdot \xi^* = \sum_{i,j} \xi_i \cdot c_{ij} \cdot \xi_j^* = \text{cov}(\xi \cdot Z, \xi \cdot Z) = \text{var}(\xi \cdot Z) \geq 0,$$

und > 0 für alle ξ , für welche die Zufallsgrösse $\xi \cdot Z$ nicht konstant ist. Die Verteilung von $Z - \mathbb{E} Z$ ist auf den Spaltenraum von C konzentriert. Für eine Zeile ξ gilt $\xi \cdot C = 0$ genau dann, wenn $\text{var}(\xi \cdot Z) = 0$.

Satz (Regressionsmatrix)

Sei X ein Zufallsvektor der Länge n und $Y = M \cdot X$ mit einer $m \times n$ -Matrix M . Es existiert dann eine $n \times m$ -Matrix N , sodass $X - N \cdot Y$ und Y unkorreliert sind.

$X = N \cdot Y + (X - N \cdot Y)$ ist also eine Zerlegung des gegebenen Zufallsvektors in unkorrelierte Summanden.

Beweis Wir suchen N mit

$$\text{cov}(X - N \cdot Y, Y) = 0 \quad \text{das heisst} \quad (I - NM)CM^* = 0.$$

Die Zeilen von N bestimmen sich als Lösungen von Normalgleichungen.

$$N \cdot MCM^* = CM^*.$$

Im Falle, wo $MCM^* = \text{cov}(Y, Y)$ positiv definit ist, gibt es nur eine Regressionsmatrix, nämlich $N = CM^*(MCM^*)^{-1}$.

Beispiel, $n=2, m=1$ Sei X der Zufallsvektor, welcher die Werte (t_i, y_i) mit den Wahrscheinlichkeiten q_i annimmt; $q_i \geq 0$, $\sum q_i = 1$.

Die Regression von Y auf den ersten Eintrag T ist gegeben durch

$$0 = \text{cov}(Y - \beta \cdot T, T) \quad \beta = \text{cov}(Y, T) \cdot (\text{var} T)^{-1}.$$

Dieser Koeffizient β ist uns bereits bei der Berechnung der Ausgleichsgeraden begegnet

$$\text{var} T = \sum q_i (t_i - \bar{t})^2 = \tau^2 \quad \text{cov}(Y, T) = \sum q_i y_i (t_i - \bar{t}).$$

Die Steigung der Ausgleichsgeraden für die gewichtete Punktwolke kann also als Regressionskoeffizient gedeutet werden.

Gauss'sche Vektoren

Eine reellwertige Zufallsgrösse X heisst eine gauss'sche Zufallsgrösse mit dem Erwartungswert μ und der Varianz σ^2 , oder auch eine $N(\mu, \sigma^2)$ -Variable, wenn gilt

$$Ws(X \in (x, x + dx)) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx \quad \text{für alle } x \in \mathbb{R}.$$

Eine reellwertige $N(0, 1)$ -Variable heisst auch eine (reelle) standardnormalverteilte Zufallsgrösse. Eine komplexwertige Zufallsgrösse, deren Real- und Imaginärteil unabhängige $N(0, 1/2)$ -Variablen sind, heisst eine komplexe standardnormalverteilte Zufallsgrösse.

Wenn Z_1, \dots, Z_n unabhängige (reelle bzw. komplexe) standardnormalverteilte Zufallsgrössen sind, dann heisst der entsprechende Zufallsvektor ein (reeller bzw. komplexer) gauss'scher Einheitsvektor. Das Bild eines gauss'schen Einheitsvektors bezüglich einer linearen Abbildung in einen affinen Raum L heisst ein L -wertiger gauss'scher Vektor. Bei L denken wir hier vor allem an die Spaltenräume \mathbb{R}_{Sp}^m bzw. \mathbb{C}_{Sp}^m . Wir sprechen dann von einem gauss'schen Vektor der Länge m .

Satz Die Verteilung eines gauss'schen Vektors der Länge m ist durch den Erwartungswert und die Covarianzmatrix eindeutig bestimmt.

Wenn man diesen bemerkenswerten Satz beweisen will, empfiehlt es sich, zuerst den folgenden Spezialfall zu beweisen:

Ist X ein gauss'scher Vektor mit $\mathbb{E}X = 0$, $\text{cov}(X, X) = I_m$, dann sind die Einträge unabhängige standardnormalverteilte Zufallsgrössen. *M. a. Worten: Wenn die Einträge*

eines gauss'schen Vektors paarweise uncorreliert sind, dann sind sie unabhängig.

Diese Aussagen ergeben sich im Fall reeller Zufallsvektoren sehr leicht aus der expliziten Formel für die Dichte eines gauss'schen Vektors mit dem Erwartungswert μ und der nichtsingulären Covarianzmatrix C .

Wenn Q die inverse Matrix bezeichnet, $Q = C^{-1}$, dann haben wir die Dichte

$$\left(\frac{1}{\sqrt{2\pi}}\right)^m \cdot (\det Q)^{1/2} \cdot \exp\left(-\frac{1}{2} \cdot (x - \mu)^\top Q (x - \mu)\right) dx$$

Es gibt noch weitere interessante Kennzeichnungen der gauss'schen Vektoren, z. B.

Ein Zufallsvektor X mit $\mathbb{E}X = 0$ ist genau dann ein gauss'scher Vektor mit der Covarianzmatrix C , wenn für jede Zeile ξ mit $\xi Q \xi^* = 1$ die Zufallsgrösse $\xi \cdot X$ standardnormalverteilt ist.

Die Beweise überlassen wir der Elementaren Stochastik. Wir betrachten die gauss'schen Vektoren als das Standardbeispiel für die Konstruktionen der Linearen Statistik.

Lineare Statistik

Die Lineare Statistik handelt von Zufallsvektoren der Länge n mit unbekanntem Erwartungswert, deren Covarianzmatrix (bis auf einen Faktor σ^2) bekannt ist. Es wird angenommen, dass der Erwartungswert in einem gegebenen p -dimensionalen Teilraum liegt; der wahre Erwartungswert ist durch eine Linearkombination der Messwerte zu schätzen, und zwar so, dass die Varianz des erwartungstreuen Schätzers minimal ist.

In statistischer Sprache formuliert man das Problem auch folgendermassen:

Lineare Einflussgrössen Gegeben sind die n -Spalten b_1, \dots, b_p , die uns angeben, wie gewisse 'Einflussgrössen' auf die Erwartungswerte der Beobachtungen Y_1, \dots, Y_n wirken. Was durch Linearkombination der b_j nicht erklärt werden kann, wird durch Zufallsfehler U_i mit Erwartungswert $= 0$ erklärt.

$$Y_i = \sum b_j \cdot \alpha^j + U_i.$$

Man nimmt an, dass die Covarianzmatrix (bis auf eine multiplikative Konstante σ^2) bekannt ist. Gesucht sind die 'Gewichte' α^j in der Linearkombination. In der Linearen Statistik bestimmt man (für $j = 1, \dots, p$) diejenigen Linearkombinationen der Y_i , $\hat{\alpha}^1, \dots, \hat{\alpha}^p$, welche 'im quadratischen Mittel' möglichst nahe bei den α^j liegen.

$$\mathbb{E}|\hat{\alpha}^j - \alpha^j|^2 = \min \quad \text{oder auch} \quad \mathbb{E}(\hat{\alpha}^j - \alpha^j) = 0 \text{ und } \text{var}(\hat{\alpha}^j - \alpha^j) = \min.$$

Beispiel (Die Schätzung einer Regressionsgeraden)

Die Festigkeit eines Materials wird bei verschiedenen Temperaturen (mehrmals) gemessen. Die (von der Temperatur abhängige) Messgenauigkeit führt zu unabhängigen (oder unkorrelierten) Zufallsgrössen Y_i , $i = 1, \dots, n$ mit gewissen Varianzen σ_i^2 . Vom Erwartungswert wollen wir annehmen, dass er ('in erster Näherung') linear von der Temperatur t abhängt. $\mathbb{E}Y_i = \alpha + \beta \cdot t_i$. Die unbekannten Werte α und β sollen aus den Y_i ermittelt

werden. Wenn wir für den ‘Temperaturgang’ der Festigkeit eine quadratische Funktion ansetzen wollten, dann müssten wir der ‘Design-Matrix’ B eine Spalte hinzufügen, etwa mit den Einträgen t_i^2 oder den Einträgen $(t_i - \bar{t})^2$ und wir müssten demgemäss auch noch eine weitere Einflussgrösse γ schätzen.

Bei den Schätzverfahren beschränken wir uns auf lineare Schätzer, also

$$\hat{\alpha} = \hat{\xi}^{(1)} \cdot Y, \quad \hat{\beta} = \hat{\xi}^{(2)} \cdot Y$$

mit geeigneten n -Zeilen $\hat{\xi}^{(1)}$ und $\hat{\xi}^{(2)}$. Wir werden sehen

$$\hat{\beta} = \sum q_i \cdot Y_i \cdot (t_i - \bar{t}), \quad \hat{\alpha} + \hat{\beta} \cdot \bar{t} = \sum q_i \cdot Y_i,$$

wobei $q_i = \text{const} \cdot \frac{1}{\sigma_i^2}$ mit $\sum q_i = 1$ zu wählen ist und die gewichteten Mittelwerte \bar{t} , \bar{y} , wie oben im Abschnitt über die Ausgleichsgerade definiert sind.

Es zeigt sich also, dass die BLUE-Schätzung auf dasselbe Ergebnis führt wie die Bestimmung der Ausgleichsgeraden. – Dieses Phänomen werden wir im allgemeinen Fall wiederfinden. In den Lehrbüchern der mathematischen Statistik findet man den Sachverhalt unter dem Stichwort ‘Satz von Gauss - Markov’.

In Matrizenschreibweise lautet unsere **Aufgabe** :

Gegeben ist eine $n \times p$ - Matrix B mit linear unabhängigen Spalten. Von einem Zufallsvektor Y mit bekannter positivdefiniter Covarianzmatrix C sei bekannt, dass er die Form hat

$$Y = B \cdot s + U \quad \text{mit} \quad \mathfrak{E}(U) = 0.$$

Gesucht ist die $p \times n$ - Matrix N mit

$$\mathfrak{E}(N \cdot Y) = s \quad \text{falls} \quad \mathfrak{E}Y = B \cdot s \quad \text{und} \quad \text{cov}(NY, NY) = \min.$$

Die erste Bedingung erweist $N \cdot Y$ als einen erwartungstreuen Schätzer der unbekannten p -Spalte s . Die zweite Bedingung macht Gebrauch von der

Sprechweise Von zwei positiv semidefiniten Matrizen sagt man, dass C_1 kleiner oder gleich C_2 ist, $C_1 \preceq C_2$, wenn die Differenz positiv semidefinit ist, d. h. wenn

$$\eta \cdot C_1 \cdot \eta^* \leq \eta \cdot C_2 \cdot \eta^* \quad \text{für alle Zeilen } \eta.$$

Wir werden sehen, dass die Aufgabe tatsächlich eine eindeutig bestimmte Lösung besitzt; diese Lösung N heisst die BLUE- Matrix für den unbekannten Parameter s .

Die Berechnung der BLUE-Matrix zum Paar B, C wollen wir vorbereiten durch einen Satz, der auch für sich interessant ist:

Satz

Gegeben sei eine positivdefinite $n \times n$ Matrix Q und eine $n \times p$ Matrix B vom Rang p .

- Zu jeder n -Zeile ξ existiert dann genau eine n -Zeile $\hat{\xi}$, sodass gilt

$$\mathfrak{E}(\hat{\xi} \cdot Y) = \mathfrak{E}(\xi \cdot Y), \quad \text{d. h.} \quad (\xi - \hat{\xi}) \cdot B = 0 \quad (5)$$

$$\text{var}(\hat{\xi} \cdot Y) \leq \text{var}((\xi - \eta) \cdot Y) \quad \text{für alle } \eta \text{ mit } \eta \cdot B = 0. \quad (6)$$

- Die Abbildung $\phi : \xi \mapsto \hat{\xi}$ ist eine orthogonale Projektion des Zeilenraums. $\mathcal{W} = (\mathbb{C}_Z^n, \mathbb{C})$.

Beweis

1. Für jedes feste ξ nimmt die nichtnegative Funktion $q(\eta) = \text{var}((\xi - \eta) \cdot Y)$ auf dem $(n - p)$ -dimensionalen Vektorraum $\mathcal{W}_0 := \{\eta : \eta \cdot B = 0\}$ ihr Minimum in genau einem Punkt $\hat{\xi} = \xi - \hat{\eta}$ an. Die Abbildung $\phi : \xi \mapsto \hat{\xi}$ ist idempotent in dem Sinn $\phi \circ \phi = \phi$.
2. Das Bild $\mathcal{W}_1 := \text{im}(\phi)$ ist ein Vektorraum. Eine n -Zeile ζ liegt nämlich genau dann in diesem Bildraum, wenn gilt

$$\text{var}(\zeta \cdot Y) \leq \text{var}((\zeta + \eta) \cdot Y) \quad \text{für alle } \eta \text{ mit } \eta \cdot B = 0$$

Das bedeutet $\text{cov}(\eta \cdot Y, \zeta \cdot Y) = 0$ für alle diese η , d. h. in Matrizenschreibweise

$$\eta \cdot C \cdot \zeta^* = 0 \quad \text{für alle } \eta \text{ mit } \eta \cdot B = 0.$$

ζ liegt also genau dann im Bild, wenn $C\zeta^*$ ein gewisses lineares Gleichungssystem löst; dessen Lösungssystem ist gerade der Spaltenraum von B .

$$\mathcal{W}_1 = \{\zeta : C\zeta^* = Br \quad \text{für eine } p\text{-Spalte } r\}$$

3. Jede n -Zeile ξ besitzt genau eine Zerlegung $\xi = \zeta + \eta$ mit $\zeta \in \mathcal{W}_1$ und $\eta \in \mathcal{W}_0$. Der Zeilenraum ist eine direkte Summe und in der Tat eine orthogonale Summe

$$\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_0 \quad \text{mit} \quad \mathcal{W}_1 \perp \mathcal{W}_0;$$

denn $\eta \cdot C \cdot \zeta^* = \eta \cdot (Bs) = 0$ für alle $\eta \in \mathcal{W}_0$ und alle s .

4. Die darstellende Matrix der orthogonalen Projektion ϕ auf \mathcal{W}_1 nennen wir P :
 $\phi: \quad \xi \mapsto \hat{\xi} = \xi \cdot P$. Die Inverse zur Covarianzmatrix C nennen wir $Q = C^{-1}$.
 $\hat{\xi}$ ist dasjenige $\zeta = r^* B^* Q$, für welches gilt $(\xi - r^* B^* Q)B = 0$. Aus dieser Bedingung an den Erwartungswert $\mathfrak{E}(\xi - \hat{\xi}) \cdot Y = 0$ erhalten wir

$$r^* = \xi B (B^* Q B)^{-1} \quad \text{und} \quad \hat{\xi} = r^* B^* Q = \xi B (B^* Q B)^{-1} B^* Q = \xi \cdot P.$$

Bemerkung Die Matrix $P = B(B^*QB)^{-1}B^*Q$ ist uns bereits bei der Methode der kleinsten Quadrate begegnet. Dort diente sie für eine Abbildung π des Spaltenraums $\mathcal{V} = (\mathbb{C}_{Sp}^n, Q)$: $\pi : y \mapsto \hat{y} = P \cdot y$. Wir zeigten, dass es sich um eine orthogonale Projektion handelt. $P \cdot P = P$ und $(Q \cdot P)$ ist eine hermitesche Matrix.

Hier nun brauchen wir P zur Darstellung einer orthogonalen Projektion des Zeilenraums $\mathcal{W} = (\mathbb{C}_Z^n, Q)$: $\phi : \xi \mapsto \hat{\xi} = \xi \cdot P$. $P \cdot P = P$ und $(P \cdot C)$ ist eine hermitesche Matrix.

Wir kehren zur Konstruktion der BLUE-Matrix zurück. Wir rekapitulieren: Es wurde eine n -Spalte Y beobachtet, über welche unser stochastisches Modell besagt:

$$Y = B \cdot s + U \quad \text{mit} \quad \mathbb{E}U = 0, \quad \text{cov}(U, U) = C.$$

Die Matrizen B und C sind bekannt; das unbekannte s soll (möglichst zuverlässig) erwartungstreu linear geschätzt werden. Gesucht ist also eine Matrix N , sodass

$$\mathbb{E}(N \cdot Y) = s \quad \text{falls} \quad \mathbb{E}Y = s \quad \text{und} \quad N \cdot C \cdot N^* = \text{cov}(NY, NY) = \min.$$

Wenn $S = N \cdot Y$ der beste erwartungstreuer Schätzer für die unbekannte p -Spalte s ist, dann gilt

$$\text{cov}(NY, NY) \preccurlyeq \text{cov}((N + E)Y, (N + E)Y) \quad \text{für jede } p \times n\text{-Matrix } E \text{ mit } E \cdot B = 0.$$

Für die Matrix N bedeutet das $E \cdot C \cdot N^* = 0$ für alle E mit $E \cdot B = 0$. Die Spalten von $C \cdot N^*$ lassen sich daher als Linearkombinationen der Spalten von E schreiben, und zwar auf eindeutige Weise, weil die Matrix B nach Voraussetzung linear unabhängige Spalten hat. Andererseits gilt wegen der Erwartungstreue $N \cdot B = I_p$. Somit gilt

$$\exists R : \quad C \cdot N^* = B \cdot R; \quad N^* = QBR; \quad B^*QB \cdot R = I_p.$$

R ist die Inverse der positiv definiten Matrix B^*QB und

$$N = R^*B^*Q = (B^*QB)^{-1}B^*Q$$

ist die BLUE-Matrix für die p -Spalte s . Die BLUE-Matrix für die n -Spalte $\mathbb{E}Y = Bs$ erhalten wir durch Multiplikation mit B . $BN \cdot Y = \hat{Y}$. Die BLUE-Zeile für den Erwartungswert der Zufallsgrößen $\xi \cdot Y$ ist ξBN .

Wir stellen fest, dass $B \cdot N$ die Projektionsmatrix P von oben ist. Die Gleichung

$$\hat{\xi} \cdot Y = \xi PY = \xi \cdot \hat{Y}$$

zeigt, dass die Methode der BLUE-Schätzung auf dasselbe Ergebnis führt wie die Methode der kleinsten Quadrate, wenn die Norm-Matrix Q die Inverse der Covarianzmatrix C ist.– Dies ist im Wesentlichen der sog. Satz von Gauss-Markov, den wir für den Spezialfall der einfachen Regression schon oben angetroffen haben.

Hinweis: Zum Gauss-Markov Theorem in seiner endgültigen Ausprägung gehört auch noch eine Aussage über die beste quadratische Schätzung der Konstanten σ^2 in der Covarianzmatrix $\text{cov}(Y, Y) = \sigma^2 \cdot C$. Wir wollten hier aber bei der Linearen Algebra im engeren Sinn bleiben.

Bemerkung zur Dualität Sei Y ein Zufallsvektor der Länge n mit $\mathbb{E}Y = 0$ und $C = \text{cov}(Y, Y)$ positiv definit. Der n -dimensionale Vektorraum aller Linearkombinationen der Einträge trägt eine Hilbertraumstruktur; für jeden n -Zeile ξ ist $\text{var}(\xi \cdot Y) = \xi C \xi^*$ das Quadrat der Norm. Es liegt nahe, diesen Hilbertraum von Zufallsgrößen mit dem Zeilenraum (\mathbb{C}_Z^n, C) zu identifizieren. Allerdings ist es in der elementaren Linearen Algebra weithin üblich, die Vektoren durch Spalten darzustellen und die Linearformen als Zeilen. Bei uns hier sind nun aber die Linearformen die Spalten; den Wert der ‘Linearform’ x im Vektor ξ erhält man, wie üblich, durch das ‘Matrizenprodukt’ $\langle \xi, x \rangle = \xi \cdot x$. Der Dualraum ist mit der dualen Norm auszustatten; das ist in unserem Fall die Hilbertraumnorm zur Matrix $Q = C^{-1}$. Die Hilberträume

$$(\mathbb{C}_{\text{Sp}}^n, Q) \quad \text{und} \quad (\mathbb{C}_Z^n, C)$$

sind zueinander dual. Bekanntlich gibt es eine natürliche antilineare Isometrie zwischen einem Hilbertraum und seinem Dualraum (in Dirac’s Notation ist es die Entsprechung $|\nu\rangle \longleftrightarrow \langle \nu|$). Hier in der Matrixalgebra wird daraus

$$x \mapsto \xi = (Qx)^* \quad \xi \mapsto x = (\xi C)^* .$$

Für die Spalte $x \in (\mathbb{C}_{\text{Sp}}^n, Q)$ und ihr Bild $\xi \in (\mathbb{C}_Z^n, C)$ gilt

$$\|\xi\|^2 = \xi C \xi^* = (x^* Q) C (Qx) = x^* Q x = \|x\|^2 .$$

Die Identifikation $\xi \leftrightarrow \xi \cdot Y$ ist intuitiv naheliegend. $\langle \xi | \eta \rangle = \text{cov}(\xi Y, \eta Y)$ Es lohnt sich, die Zeilen als die primär gegebenen ‘Vektoren’ anzusehen. Allerdings ist das innere Produkt im ersten Argument linear und im zweiten antilinear, entgegen den Konventionen der Funktionalanalysis.

A Anhang : Mengen und Relationen

Mengen

G. Cantor führte 1895 die folgende Sprechweise ein: „Unter einer Menge M verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die „Elemente“ von M genannt werden) zu einem Ganzen“.

Nach der heute allgemein akzeptierten (anfangs aber aus philosophischen Gründen heftig bekämpften) Auffassung von Cantor sind beliebige Mengen legitime Objekte mathematischen Betrachtung.

Eine Menge M gilt als wohldefiniert, wenn festgelegt ist, welche Elemente zu ihr gehören und was es heißt, dass zwei Elemente von M als gleich zu betrachten sind.

Die Konstruktion einer Menge erfolgt in der Regel so, dass man aus den Elementen bekannter Mengen neue Objekte zusammenbaut und dann festlegt, welche dieser Konstrukte als gleich (oder äquivalent) zu betrachten sind.

Ein Beispiel ist aus der Schule bekannt:

Die Menge \mathbb{Q} der rationalen Zahlen wird dort so konstruiert: Man betrachtet alle Zahlenpaare (m, n) mit $m \in \mathbb{Z}$, $n \in \mathbb{N}$ und definiert

$$(m_1, n_1) \sim (m_2, n_2) \Leftrightarrow m_1 \cdot n_2 = m_2 \cdot n_1 \quad .$$

Ähnlich konstruiert man bekanntlich die Menge der gebrochenrationalen Funktionen.

Bevor wir uns den Problemen um Gleichheit und Äquivalenz zuwenden, skizzieren wir einen Kalkül der Mengen, der schon von der Schule her bekannt sein sollte.

Mengenalgebra

Betrachten wir Teilmengen A, B, C, \dots einer festen Grundmenge Ω . Zwei solche Teilmengen A, B kann man vereinigen oder auch durchschneiden; man erhält $A \cup B$ bzw. $A \cap B$; und das sind wohldefinierte Teilmengen von Ω . Das Komplement von A (in Ω) bezeichnet man

$$\Omega \setminus A \quad \text{oder} \quad \neg A \quad \text{oder} \quad A^c \quad .$$

Das „Komplement von A in B “ bezeichnet man

$$B \setminus A \quad \text{oder} \quad B \cap (\neg A) \quad \text{oder} \quad B \cap A^c \quad .$$

Die Gesamtheit aller Teilmengen von Ω heißt die Potenzmenge $\mathfrak{P}(\Omega)$. Die Komplementbildung ist eine einstellige Operation, Vereinigung und Durchschnittsbildung sind zweistellige Operationen (auch zweistellige Verknüpfungen genannt). Die volle Menge ist ein ausgezeichnetes Element, die leere Menge \emptyset ist ihr Komplement.

Die Eigenschaften der strukturierten Menge

$$(\mathfrak{P}(\Omega), \emptyset, \neg, \cup, \cap)$$

kann man in ähnlicher Weise fassen, wie man das bei den Zahlen (oder den Polynomen) gewohnt ist. Die wichtigsten Rechenregeln sind

- (i) $A \cup B = B \cup A$ $A \cap B = B \cap A$
- (ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- (iii) $A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$
- (iv) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- (v) $\neg(\neg A) = A$ für alle A
- (vi) $\neg(A \cup B) = (\neg A) \cap (\neg B)$

Wenn man diese Eigenschaften als ein Axiomensystem einer strukturierten Menge $(\mathfrak{A}, \emptyset, \neg, \cup, \cap)$ auffasst, dann hat man den Begriff der **Boole'schen Algebra**. Man kann sich jetzt daran machen, ein möglichst kleines System von elementaren Rechenregeln („Axiomen“) ausfindig zu machen, aus welchem man alle in einer Boole'schen Algebra gültigen Regeln ableiten kann. Man kann auch Systeme $(\mathfrak{A}, \cup, \cap)$ untersuchen, in welchen nur die Eigenschaften (i), (ii), (iii), bzw. (i), (ii), (iii), (iv) gefordert werden. Das führt zu den (distributiven) Verbänden, einem Thema, welches einen Physiker zunächst einmal nicht interessieren muss.

Hinweis : Ein bekannter Prüfstein für die Vollständigkeit eines von irgendjemanden vorgeschlagenen Axiomensystems ist die Regel

$$(vii) \quad (A \cup B) \cap (B \cup C) \cap (C \cup A) = (A \cap B) \cup (B \cap C) \cup (C \cap A) .$$

Es sei dem Leser überlassen, sie aus (i), (ii), (iii), (iv) herzuleiten.

Aussagenkalkül

Die axiomatische Herangehensweise geht auf George Boole (1815-1864) zurück. In seiner Untersuchung „The Laws of Thought“ (1854) legt Boole dar, wie die Gesetze der formalen Logik, die von Aristoteles aufgestellt worden waren, zum Gegenstand eines Kalküls gemacht werden können. Der Beweis, dass das axiomatische System von Boole äquivalent ist mit der Mengenalgebra, hat M. Stone 1936 gezeigt. (Der Beweis passt nicht in eine Anfängervorlesung).

Im Boole'schen Aussagenkalkül notiert man

die Vereinigung („Disjunktion“) $A \vee B$ („A oder B“) und

den Durchschnitt („Konjunktion“) $A \wedge B$ („A und B“).

$\neg A$ liest man „nicht A“.

Hinweise zur Notation :

- a) Die Logiker haben auch noch spezielle Zeichen für gewisse weitere Verknüpfungen. Man schreibt z.B. $A \curvearrowright B$ oder $A \rightarrow B$ statt $(\neg A) \cup B$ und liest „Wenn A dann B“. Das Zeichen \curvearrowright heißt der Subjunktionspfeil, manchmal auch der Implikationspfeil, obwohl er für den strengen Logiker nicht genau dasselbe bedeutet wie das Kürzel \Rightarrow , welches die Mathematiker als den Implikationspfeil bezeichnen.

- b) Die Informatiker benützen auch gerne das „exklusive Oder“ $A \triangle B = (A \vee B) \setminus (A \wedge B)$, sowie die Verknüpfungen NAND und NOR.
- c) Die Mathematiker benützen die Symbole der Logik in einer naiven („halbformalen“) Weise. Sie verlassen sich darauf, dass aus dem Kontext hervorgeht, was die Formeln meinen; die Kürzel werden als Elemente einer Kurzschrift benützt, nicht als die Grundlage von Beweisen.
- d) Wir werden die Kurzschrift üben, wenn wir Analysis treiben. Da kommen dann allerdings noch die Quantoren hinzu: der Existenzquantor \exists und der Allquantor \forall . Den Allquantor schreibt man manchmal auch \bigwedge , um anzudeuten, dass es um eine Konjunktion einer (i.Allg. unendlichen) Familie von Aussagen geht.

Strukturierte Mengen

Die grundsätzliche Bedeutung des Mengenbegriffs für die Mathematik des 20. Jahrhunderts kann man nicht erraten, wenn man nur an die Mengenalgebra denkt. (Konkret betriebene Mengenalgebra ist ein Stoff für die elementare Schulmathematik, ebenso wie die elementare Arithmetik). Die philosophische Bedeutung des Mengenbegriffs ergibt sich daraus, dass man in der heutigen reinen Mathematik alle Objekte der mathematischen Betrachtung als Elemente einer strukturierten Menge versteht. Das heißt: Man fragt in der reinen Mathematik nicht nach der „wahren Natur“ der Gegenstände, mit welchen operiert wird; diese sind (mit Cantor gesprochen) irgendwelche wohlunterschiedene Objekte unserer Anschauung oder unseres Denkens. Es genügt, wenn feststeht, welche Elemente zur Menge gehören und außerdem bei je zwei Elementen feststeht, ob sie gleich oder verschieden sind. Bei den Strukturen, die man einer abstrakten Menge aufprägen kann, kann man grob gesagt drei Typen unterscheiden: algebraische, topologische und Ordnungsstrukturen. Beispiele für algebraische Strukturen haben wir bereits gesehen, allen voran die Vektorraumstruktur und die Gruppenstruktur.

In Abschnitt III haben wir gesagt: Aus einer Menge $(S, =)$ wird ein metrischer Raum $(S, d(\cdot, \cdot))$, wenn man eine Metrik auszeichnet. Eine Metrik liefert einen wichtigen Typ einer topologischen Struktur.

Was eine Ordnungsstruktur ist, ergibt sich aus der folgenden Definition.

Definition

Eine Menge S wird zu einer (partiell) **geordneten Menge**, wenn man eine (partielle) Ordnung auszeichnet.

Eine zweistellige Relation R heißt eine **Präordnung**, wenn gilt

- (i) xRx für alle $x \in S$ (Reflexivität)
- (ii) $xRy, yRz \Rightarrow xRz$ (Transitivität)

Eine Präordnung R heißt eine (partielle) Ordnung, wenn zusätzlich gilt

- (iii) $xRy, yRx \Rightarrow x = y$ (Antisymmetrie)

Beispiel 1

Die übliche Ordnung „ \leq “ in \mathbb{R} (oder in \mathbb{Q}) ist eine Ordnungsrelation im Sinne der Definition. Für diese Ordnung gilt zusätzlich

$$(iv) \quad \forall x, y : xRy \text{ oder } yRx .$$

Man spricht von einer **totalen Ordnung**.

Beispiel 2

Auf \mathbb{N} ist die Teilbarkeitsrelation eine partielle Ordnung. Man notiert

$$a|b, \text{ wenn ein } m \in \mathbb{N} \text{ existiert, so dass } am = b \quad („a \text{ teilt } b“)$$

Beispiel 3

Die Inklusion \subseteq ist eine partielle Ordnung auf der Potenzmenge $\mathfrak{P}(\Omega)$ (zu einer beliebigen Grundmenge Ω).

In allen drei Beispielen hat die Ordnung die bemerkenswerte Eigenschaft, dass es zu jedem Paar ein Maximum und ein Minimum gibt, d.h. eine kleinste Majorante und eine größte Minorante.

Im Beispiel 2 heißt das Minimum der größte gemeinsame Teiler, das Maximum das kleinste gemeinsame Vielfache. Im Beispiel 3 ist das Minimum die Schnittmenge, das Maximum die Vereinigungsmenge.

Definition

(M, \leq) sei eine partiellgeordnete Menge.

Man sagt, c sei das Minimum von a und b und notiert $c = \min\{a, b\}$ oder auch $c = a \wedge b$, wenn gilt

- (i) $c \leq a, c \leq b$
- (ii) $\forall d : (d \leq a, d \leq b) \Rightarrow d \leq c$.

Man sagt, e sei das Maximum von a, b und notiert $e = \max\{a, b\}$ oder auch $e = a \vee b$, wenn gilt

- (i) $a \leq e, b \leq e$
- (ii) $\forall d : (a \leq d, b \leq d) \Rightarrow e \leq d$.

Bemerkung :

Wenn zu einem Paar a, b das Minimum (bzw. das Maximum) existiert, dann ist es eindeutig bestimmt. Es können nicht zwei verschiedene Elemente als Minimum (Maximum) qualifizieren. Größte untere Schranken nennt man auch Infima, kleinste obere Schranken heißen Suprema.

Definition

Eine partiell geordnete Menge (M, \leq) heißt ein **Verband** (engl. „lattice“), wenn zu jedem Paar a, b das Minimum und das Maximum existiert.

Beispiel 4

$(\mathfrak{K}, \subseteq)$ sei die Menge aller konvexen Teilmengen des d -dimensionalen Anschauungsraums, geordnet durch die Inklusion. $(\mathfrak{K}, \subseteq)$ ist ein Verband. Das Minimum ist der mengentheoretische Durchschnitt; das Maximum ist die konvexe Hülle der Vereinigungsmenge.

Beispiel 5

Die Menge aller stetigen reellwertigen Funktionen über dem Einheitsintervall wird mit $C([0, 1])$ bezeichnet. Sie ist ein Verband bzgl. der punktweisen Ordnung.

Bemerkung : $C([0, 1])$ ist außerdem ein reeller Vektorraum bzgl. der punktweisen Addition. Dieser „Vektorverband“ wird uns in der Integrationstheorie beschäftigen.

Äquivalenzrelationen**Definition**

Eine zweistellige Relation über einer Grundmenge Ω heißt eine **Äquivalenzrelation**, wenn gilt

- (i) $a \sim a$ für alle $a \in \Omega$ (Reflexivität)
- (ii) $a \sim b, b \sim c \Rightarrow a \sim c$ (Transitivität)
- (ii) $a \sim b \Rightarrow b \sim a$ (Symmetrie)

Beispiel 1

Sei $\Omega = \mathbb{N}$ (Menge der natürlichen Zahlen), $m \in \mathbb{N}$ fest. Man sagt, a und b seien äquivalent modulo m und notiert

$$a \equiv b \pmod{m}, \text{ wenn } m | (a - b)$$

(„ m teilt die Differenz“).

Beispiel 2

Sei $d(\cdot, \cdot)$ eine Pseudometrik auf der Menge Ω . Wir erhalten eine Äquivalenzrelation wenn wir definieren

$$P \sim Q \stackrel{\text{def}}{\iff} d(P, Q) = 0.$$

(Statt Pseudometrik sagt man auch Semimetrik oder Prämetrik.)

Beispiel 3

Sei \leq eine Präordnung auf einer Menge S . Wir erhalten eine Äquivalenzrelation auf S , wenn wir definieren

$$x \sim y \iff x \leq y \text{ und } y \leq x.$$

Bemerkung : Eine Präordnung ist also genau dann eine Ordnung, wenn die dazugehörige Äquivalenzrelation die Gleichheit ist.

Satz

Sei \sim eine Äquivalenzrelation über Ω .

Für ω_1, ω_2 sei $\tilde{\omega}_1 = \{\omega : \omega \sim \omega_1\}$, $\tilde{\omega}_2 = \{\omega : \omega \sim \omega_2\}$. Die Mengen $\tilde{\omega}_1, \tilde{\omega}_2$ sind entweder disjunkt oder gleich.

Es gilt $\tilde{\omega}_1 = \tilde{\omega}_2 \iff \omega_1 \sim \omega_2$.

Die Gesamtheit aller Äquivalenzklassen ist eine Menge im Sinne von

Cantors Definition (1895) : „Unter einer Menge M verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die „Elemente“ von M genannt werden) zu einem Ganzen“.

Notation Sei $(M, =)$ eine Menge und sei \sim eine Äquivalenzrelation auf M . (M, \sim) bezeichnet dann die Menge aller Äquivalenzklassen.

Beispiel

Die Gesamtheit aller Restklassen modulo m ist eine Menge M mit $|M| = m$ Elementen. Die Menge der Äquivalenzklassen modulo m ist eine Menge der Mächtigkeit m . Sie ist außerdem ein Ring und für Primzahlen m sogar ein Körper.

Beispiel

Seien M und N Mengen und $\varphi : M \rightarrow N$ eine Abbildung. Wir setzen

$$m_1 \sim m_2 \stackrel{\text{def}}{\iff} \varphi(m_1) = \varphi(m_2) .$$

Die Äquivalenzklassen stehen in eindeutiger Beziehung zu den Punkten $n \in \text{im } \varphi$

$$\tilde{m}_1 = \{m : \varphi(m) = \varphi(m_1)\} .$$

Definition

Man sagt von einer Abbildung

$$\varphi : M \rightarrow N$$

dass sie mit der Äquivalenzrelation \sim auf M verträglich ist, wenn gilt

$$m_1 \sim m_2 \Rightarrow \varphi(m_1) = \varphi(m_2) .$$

Man kann in diesem Fall sagen, dass φ eine Abbildung von (M, \sim) nach N liefert.

Bemerke

Der Begriff der Verträglichkeit passt auch auf Verknüpfungen. Sei $\varphi : M_1 \times M_2 \rightarrow N$ eine Abbildung. Auf M_1 und auf M_2 sei eine Äquivalenzrelation gegeben. Man sagt, φ sei mit diesen Äquivalenzen verträglich, wenn gilt

$$a_1 \sim b_1, a_2 \sim b_2 \Rightarrow \varphi(a_1, a_2) = \varphi(b_1, b_2) .$$

Beispiel :

Die Addition und die Multiplikation in \mathbb{Z} ist mit der Äquivalenz modulo m verträglich. Man kann Äquivalenzklassen modulo m addieren und multiplizieren.

Didaktischer Hinweis

Der Anfänger versucht erfahrungsgemäß, bei den Objekten der Mathematik in Stufen zu denken. Die Zahlen erscheinen da als Objekte auf einer ersten Stufe; Mengen von Zahlen oder Funktionen erscheinen als Objekte auf einer höheren Stufe; Funktionenräume oder Mengensysteme erscheinen auf einer noch höheren Stufe. In der Umgangssprache schlägt sich dies z.B. so nieder, dass man eine Vorschrift, welche jedem Element eines Funktionenraums eine Zahl zuordnet, gerne ein „Funktional“ nennt, um einen Unterschied zu machen zu den „Funktionen“, die hier als Elemente des Definitionsbereichs des Funktionalen auftreten. Das Integral ist in diesem Sinne ein Funktional auf einer Menge von Funktionen (auf einem meßbaren Grundraum).

Die Unterscheidung von Stufen (bei den Objekten der Mathematik, oder bei den Aussagen) ist in der heutigen Mathematik offiziell nicht vorgesehen. In manchen Überlegungen zur Grundlegung der Mathematik wird die Stufung aber sehr wohl diskutiert.

Im Anschluss an Cantors Ideen zur Mengenlehre waren nämlich logische Paradoxien konstruiert worden (z.B. die Paradoxie von den „Mengen, die sich nicht als Element enthalten“ von B. Russell), denen man mit einer „Stufenlogik“ beizukommen versuchte. Davon ist man abgekommen. Das folgende Beispiel exemplifiziert die heute übliche Einebnung der Stufen. Übrigens hat auch das Ineinssetzen des Subjunktionspfeils mit dem Implikationspfeil etwas mit der Einebnung der Stufen zu tun.

Exkurs : Partitionen und Äquivalenzrelationen**Definition**

M sei eine Menge („Grundmenge“). Eine Familie paarweise disjunkter Teilmengen M_j ($j \in J$) heißt eine Partition von M , wenn die Vereinigung gleich M ist.

Man notiert $M = \sum_{\{j: j \in J\}} M_j$ oder $M = M_1 + M_2 + \dots$, wenn es sich um eine abzählbare Familie handelt. Die nichtleeren M_j heißen die Atome der Partition.

Wir machen die Gesamtheit aller Partitionen von M zu einer geordneten Menge (\mathfrak{M}, \preceq) und wir zeigen, dass man diese Menge mit der Menge aller Äquivalenzrelationen über M identifizieren kann.

Definition

Man sagt von einer Partition $M = \sum M'_j$, dass sie feiner ist als eine Partition $M = \sum M''_k$, wenn jedes M'_j in einem M''_k enthalten ist. Man notiert

$$(M = \sum M'_j) \preceq (M = \sum M''_k) \quad .$$

Satz

Die Relation \preceq ist eine (Prä)-Ordnung auf der Gesamtheit \mathfrak{M} aller Partitionen.

Es ist nun klar, wann zwei Partitionen von M als gleich zu betrachten sind; \mathfrak{M} ist eine Menge.

Jeder Partition von M entspricht in umkehrbar eindeutiger Weise eine Äquivalenzrelation über M . Zwei Elemente $m, n \in M$ sind bzgl. der entsprechenden Äquivalenzrelation äquivalent, wenn sie im gleichen Atom der Partition liegen.

Bemerke : Die geordnete Menge (\mathfrak{M}, \preceq) ist ein Verband und zwar ein sog. vollständiger Verband. Zu jeder Familie von Partitionen gibt es eine grösste gemeinsame Verfeinerung und eine feinste gemeinsame Vergröberung.

Allgemeinere Relationen

Äquivalenzrelationen und Präordnungen sind besonders wichtige Typen von zweistelligen Relationen. In der „universellen Algebra“ definiert man allgemein, was eine n -stellige Relation auf einer Menge M ist. Eine n -stellige Relation R ist nicht anderes als eine Teilmenge des cartesischen Produkts $M \times M \times \dots \times M$; ein n -Tupel (a_1, \dots, a_n) erfüllt die Relation oder es erfüllt sie nicht. Eine Verknüpfung \square wie die Multiplikation in einer Gruppe (oder die Addition in einem Vektorraum) kann als eine Abbildung aufgefasst werden.

$$\square : M \times M \rightarrow M ; (a, b) \mapsto c = a \square b$$

oder aber als eine dreistellige Relation $R \subseteq M \times M \times M$.

$$(a, b, c) \in R \iff a \square b = c .$$

Es ist bemerkenswert, dass Begriffe wie Abbildung, Verknüpfung, Relation usw. logisch auf den Begriff der Teilmenge (eines cartesischen Produkts) zurückgeführt werden können. Ein Beispiel ist aus der Schule bekannt: Von einer reellwertigen Funktion $f(\cdot)$ auf einem Intervall $[a, b]$ sagt man zwar, dass es sich um eine Zuordnung handelt; man zeichnet dann aber den „Funktionsgraphen“, und das ist eine Teilmenge des Produktraums $[a, b] \times \mathbb{R}$.

Nebenklassen, Faktorgruppen, Quotientenräume

Sei (G, e, \cdot) eine Gruppe. U sei eine Untergruppe („subgroup“), d.h. eine Teilmenge von G , welche multiplikativ abgeschlossen ist und mit g auch g^{-1} enthält.

Zu jedem $a \in G$ definiert man die „Linksnebenklasse“ (left coset)

$$aU = \{b : b = a \cdot u \text{ mit } u \in U\} = \{b : a^{-1}b \in U\}$$

und die „Rechtsnebenklasse“ (right coset)

$$U \cdot a = \{c : c = u \cdot a \text{ mit } u \in U\} = \{c : c \cdot a^{-1} \in U\} .$$

Wir befassen uns hier mit der Menge Ω aller Linksnebenklassen.

Satz

Zwei Linksnebenklassen sind entweder identisch oder disjunkt.

$$a_1U = a_2U \iff a_2^{-1} \cdot a_1 \in U \iff a_1^{-1} \cdot a_2 \in U .$$

Der Beweis ist trivial.

Corollar

Die Menge der verschiedenen Linksnebenklassen ist eine Partition der Gruppe. Wir erhalten eine Äquivalenzrelation auf G , wenn wir definieren

$$a_1 \sim a_2 \pmod{U_\ell} \iff a_1^{-1} \cdot a_2 \in U \iff a_2 \in a_1 \cdot U .$$

Bemerke : Dasselbe gilt für die Partition in Rechtsnebenklassen. Dort gilt

$$a_1 \sim a_2 \pmod{U_r} \iff a_2 \cdot a_1^{-1} \in U \iff a_2 \in U \cdot a_1 . .$$

Definition

Eine Untergruppe H von (G, e, \cdot) heißt ein Normalteiler („normal subgroup“), wenn gilt

$$a \cdot H \cdot a^{-1} = H \text{ für alle } a \in G .$$

Bemerke :

H ist genau dann ein **Normalteiler**, wenn jede Linksnebenklasse auch eine Rechtsnebenklasse ist.

Der Beweis ist trivial.

Beispiel

G sei die Oktaedergruppe; das ist die Gruppe aller derjenigen Drehungen des Anschauungsraums, welche den Einheitswürfel (oder das Einheitsoktaeder) in sich überführen.

U sei die Menge derjenigen Drehungen, welche die obere (und damit auch die untere) Fläche des Würfels in sich überführen. Es gilt $|G| = 24$, $|U| = 4$. Es gibt $6 = \frac{1}{4} \cdot 24$ verschiedene Linksnebenklassen; wenn a die Deckfläche in die rechte Seitenfläche überführt, dann besteht $a \cdot U$ aus all denjenigen Drehungen, welche die Deckfläche in eben diese Seitenfläche überführen. U ist kein Normalteiler.

Satz

Sei G eine endliche Gruppe und U eine Untergruppe. Dann ist $|U|$ ein Teiler von $|G|$ und der Quotient ist die Anzahl der verschiedenen Linksnebenklassen (oder Rechtsnebenklassen).

Definition

Seien G_1 und G_2 Gruppen. Eine Abbildung

$$\varphi : G_1 \rightarrow G_2$$

heißt **Gruppenhomomorphismus**, wenn gilt

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) \quad \text{für alle } a, b \in G_1.$$

Bemerkung

Ein Gruppenhomomorphismus bildet das Einselement in das Einselement ab; und es gilt

$$\varphi(a^{-1}) = (\varphi(a))^{-1} \quad \text{für alle } a.$$

Satz (Homomorphie-Satz)

Wenn $\varphi : G_1 \rightarrow G_2$ ein Gruppenhomomorphismus ist, dann ist das „Bild“ eine Untergruppe von G_2 ; der „Kern“ $K = \{g_1 : \varphi(g_1) = e_2\}$ ist ein Normalteiler in G_1 . Die Bildpunkte entsprechen in eindeutiger Weise den Nebenklassen zum Kern.

$$\varphi(a) = \varphi(b) \iff ab^{-1} \in K \iff ba^{-1} \in K.$$

Der Beweis liegt auf der Hand.

Satz (Faktorgruppen)

Sei H ein Normalteiler in G . Es gilt dann

$$\begin{aligned} a \sim b(\text{mod} H) &\Rightarrow a^{-1} \sim b^{-1}(\text{mod} H) \\ a \sim c, b \sim d(\text{mod} H) &\Rightarrow ab \sim c \cdot d(\text{mod} H) \end{aligned}$$

Die Gruppenoperationen sind also verträglich mit der Äquivalenzrelation $\text{mod} H$.

Die Gruppe der Äquivalenzklassen heißt die Faktorgruppe von G modulo H und wird mit G/H bezeichnet.

Bemerkung :

Die Abbildung

$$\varphi : G \rightarrow G/H$$

ist ein surjektiver Gruppenhomomorphismus mit dem Kern H .

Bemerke :

Sei G eine Gruppe, U eine Untergruppe und Ω die Menge aller Linksnebenklassen. G operiert dann transitiv auf Ω vermöge der „Linksmultiplikation“ $L_a(\cdot)$. Für $\omega = b \cdot U$ gilt

$$L_a(\omega) = a \cdot \omega = a \cdot (bU) = (a \cdot b)U \in \Omega .$$

Es gilt offenbar

$$L_e(\cdot) = \text{Id (auf } \Omega) \quad \text{und} \quad L_{a_1}(L_{a_2}(\cdot)) = L_{a_1 a_2}(\cdot) .$$

Für diese Aussage muss man nicht voraussetzen, dass U ein Normalteiler ist.

Definition

Sei V ein \mathbb{K} -Vektorraum und W ein Teilvektorraum. Man definiert für $v_1, v_2 \in V$

$$v_1 \sim v_2(\text{mod } W) \iff v_1 - v_2 \in W .$$

Die Menge der Äquivalenzklassen heißt der Quotientenraum von V modulo W und wird mit V/W bezeichnet.

Bemerke : Die Äquivalenzrelation ist mit den Vektorraumoperationen verträglich, der Quotientenraum ist ein Vektorraum.

Konjugationsklasse

Sei (G, e, \cdot) eine Gruppe. Man sagt von zwei Gruppenelementen g_1 und g_2 , dass sie konjugiert sind, wenn ein $h \in G$ existiert, sodass $hg_1h^{-1} = g_2$. Das Konjugiertsein ist offenbar eine Äquivalenzrelation. Die einzelnen Äquivalenzklassen können ganz verschiedene Mächtigkeiten haben; die Konjugationsklasse des Einheitselement beispielsweise ist immer einpunktig. Die Anzahl der verschiedenen Konjugationsklassen ist, wie wir in IV.8 sehen, für eine endliche Gruppe gleich der Anzahl der irreduziblen Gruppendarstellungen. Die Gruppe wirkt auf die Menge G vermöge der Operation

$$G \ni h \longmapsto \psi_h(\cdot) \quad \text{mit} \quad \psi_h(g) = hgh^{-1} \quad \text{für} \quad g \in G.$$

Dabei bilden die Konjugationen ψ_h jede Konjugationsklasse auf sich ab.

Beispiele

- Die symmetrische Gruppe \mathcal{S}_3 hat 6 Elemente; drei von ihnen sind Transpositionen, zwei sind zyklische Permutationen; zusammen mit dem Einheitslement hat man 3 Konjugationsklassen.
- Die Oktaedergruppe hat 24 Elemente, partitioniert in 5 Konjugationsklassen: es gibt 6 Drehungen um 90° (um die Achse durch einen Seitenmittelpunkt), es gibt 3 Drehungen um 180° um eine solche Achse, ausserdem 6 Drehungen um 180° um die Achse durch einen Kantenmittelpunkt, sowie 8 Drehungen (um 120° bzw. 240°) um eine Raumdiagonale.

Index

- Abbildung, 83
- abgeschlossene Menge, 156
- Adjungierte, 338
- adjungierter Operator, 332
- affine Funktion, 173
- affiner Teilraum, 113
- affines Koordinatensystem, 43, 120
- Affinität, 105
- antilinear, 331, 347
- Äquivalenzrelation, 352
- Assoziativgesetz, 1, 60
- aufspannendes System, 295
- Ausgleichsgerade, 338
- Austauschlemma, 296
- Axiome
 - assoziative Algebra, 31
 - Gruppe, 86
 - Gruppenwirkung, 99
 - Ring mit Einselement, 31
 - Transformationsgruppe, 86
 - Vektorraum, 41
- Basis, 45, 118, 296
- beschränkt, 155
- bijektiv, 83
- Bildbereich, 83
- Bildblock, 314, 330
- bilinear, 297
- Binomialreihe, 72
- Blockmatrix, 261, 266
- bra-Vektor, 198, 331
- Bruchrechnen, 1
- Cardanische Formel, 146
- Cayley-Transformation, 103
- charakteristische Funktion, 63
- Charaktertafel, 245, 271
- Cotangentialraum, 298
- Courant, R., 79, 97
- Covarianz, 341
- Covektor, 298, 327
- Darstellung
 - reguläre, 260
 - unitäre, 264
- Definitionsbereich, 83
- Determinante, 4
- Diagonalisierung, 225, 261
 - simultane, 200
 - unitäre, 201
- Diagonalmatrix, 3, 8
- Differentiation, 237, 241
- Dimension, 108
- Dirac's Notation, 198, 277, 328, 331
- direkte Summe
 - äußere, 299
- direkte Zerlegung, 299
- Dirichlet-Kern, 64, 166, 195, 274
- Dispersionsrelation, 116, 234
- Distributivgesetz, 1, 60
- Drehachse, 250
- Dreiecksmatrix, 35
- Dreiecksungleichung, 151
- Dreiecksverteilung, 61
- duale Abbildung, 122
- duale Basis, 297
- duale Norm, 163
- Dualraum, 163, 297
- ebene Wellen, 115, 233
- Eigenschwingung, 234
- Eigenvektor, 54, 282
- einfach zusammenhängend, 255
- Einheitsmatrix, 3
- einseitige Folge, 59
- Elektromotorische Kraft, 218
- Ellipse, 130, 141
- Endomorphismus, 52, 201
- Energiedissipation, 217
- Epigraph, 149
- erzeugende Funktion, viii, 71
- Euklidischer Algorithmus, 24
- Eulers Funktionsbegriff, 237
- Exponentialabbildung, 284
- Faktorgruppe, 357

- Faltungsprodukt, 59, 280
- fastperiodische Funktion, 221, 242
- Fejér-Kern, 64, 243
- Feynman, 44, 79, 250
- finite Folge, 59
- finite Gewichtung, 59
- Fourier-Analyse
 - diskrete, 242
- Fourier-Integral, 234
- Fourier-Transformation
 - diskrete,DFT, 239
 - schnelle,FFT, 241
- Fourier-Transformierte, 240
- Frequenz, 12
- Fundamentalsatz
 - der Algebra, x, 20, 137, 144
- ganzrationale Funktion, 25
- gauss'sche Zahlenebene, vii, 137
- gauss'scher Vektor, 342
- Gauss-Jordan-Verfahren, 320
- Gauss-Markov Theorem, 344, 346
- gebrochenrationale Funktion, 25
- Gewichtung, 59
- Gleichungssystem
 - homogenes, 301
 - inhomogenes, 301
- Gradient, 298
- Gram-Matrix, 192
- Graph, 151
 - gerichteter, 84
- Gruppe, 86
- Gruppen
 - $SL(2, \mathbb{C})$, 99, 108
 - $SL(2, \mathbb{R})$, 110
 - $SO(3, \mathbb{R})$, 106
 - $SU(1, 1)$, 99, 102
 - $SU(2)$, 99
 - $U(n)$, 33
 - einparametrische, 209
 - affine Gruppe, 105
- Gruppenalgebra, 280
- Gruppencharakter, 272
- Gruppendarstellung, 259
 - äquivalente, 262
 - disjunkte, 262, 267
 - irreduzible, 261
 - Standard- oder 'reguläre', 267
 - unitär äquivalente, 264
- Gruppenhomomorphismus, 357
- Gruppenisomorphismus, 86
- Gruppenwirkung
 - effektive, 99
 - transitive, 104
- Hölder'sche Ungleichung, 149
- Hülle
 - abgeschlossene, 156
 - abgeschlossene konvexe, 174
 - konvexe, 181
- Hamilton, 252
- Hamilton-Operator, 15, 250
- Hamming-Distanz, 151
- harmonischer Oszillator, 223
 - einfach, 217
 - gekoppelt, 219
- Hauptachsen, 199
- Hauptkomponenten, 329
- hermitisch, 32, 33
- hermitesche Form, 193
- Hilbert-Schmidt-Norm, 168, 288
- homogen
 - absolut, 161
 - positiv, 169
- Horner-Schema, 20
- Hülle
 - konvexe, 172
 - lineare, 295
- hyperbolische Bewegungen, 102
- Indikatorfunktion, 89
- Infimum, 351
- inhomogen, 322
- injektiv, 83
- inneres Produkt, 195
- inneres Produkt, Skalarprodukt, 191, 298
- Inversblock, 314
- inverse Matrix, 4
- Isometrie, 264, 283
- Jacobi's Iterationsverfahren, 289, 334

Jacobi-Matrix, 48
 Jordan-Block, 55

 \mathbb{K} -affine Abbildung, 121
 \mathbb{K} -affine Funktion, 119
 \mathbb{K} -lineare Abbildung, 120
 Kantenmenge, 84
 Kegel, 177
 dualer, 177
 Kernblock, 314, 330
 ket-Vektor, 198, 331
 kleinste Quadrate, 337
 Koeffizientenvergleich, 73
 Kommutator, 254
 Komplement, 348
 invariantes, 261
 Komplementärraum, 201
 komplexe Konjugation, 10
 Konjugationsklasse, 271
 Konjugationsklassen, 358
 Konvergenz, 155
 konvexer Kegel, 173, 177
 Kugel, 159

 Laplace Operator, 232
 laufende Wellen, 229
 Legendre-Transformation, 175
 Lichtkegel, 132, 247
 linear abhängig, 5
 linear unabhängig, 32, 43, 45
 Linearform, 117, 297
 Linearkombination, 295
 Linksnebenklasse, 356
 Lorentzgruppe, 247
 Lot, 203

 Möbiustransformation, 257
 Matrix
 ähnliche, 53
 Maximierungsargument, 199, 334
 Maximum, 351
 Metrik, 151
 Minimalpolynom, 51
 Minimum, 351
 Minkowski-Funktional, 169
 Minkowskiraum, 247

 Minorante (konvexe), 173
 Multilineare Algebra, 126
 multiplikative Inverse, 1, 10

 Nebenklasse, 104
 Newton, 72
 nichtsinguläre Matrix, 4, 32
 Norm, 10
 normal modes, 224
 Normalgleichung, 338
 Normalteiler, 356
 Nullmatrix, 3

 offene Menge, 138, 156
 Oktaedergruppe, 87, 104, 356, 358
 ONBasis, 283
 orthochron, 247
 orthogonale Gruppe, 106
 orthogonale Matrix, 33, 106
 orthogonale Projektion, 332, 333, 337, 345
 orthogonale Summe, 345
 Orthogonalitätsrelation, 66
 Orthogonalraum, 202
 Orthonormalbasis, 247
 Orthonormalisierung, 192
 Orthonormalsystem, 192
 Ortsvektor, 43, 113

 Paarung, natürliche, 297
 Parallelogrammgleichung, 189
 Partition einer Menge, 354
 Pauli-Spinmatrizen, 37, 49, 101, 248
 Periode, 12
 Periodogramm, 243
 Permutation, 36, 84, 87
 Phasengeschwindigkeit, 233
 Pivot-Wahl, 310
 Pivottransformation, 296
 Polarisierung, 197, 277
 Polynom, 19
 polynomiale Funktion, 125
 positiv definit, 34
 Potenzmenge, 348
 Prämetrik, 352
 Präordnung, 350
 Produkt

- äusseres, 48, 254
 - inneres, 190, 247
- Projektionen
 - disjunkte, 201
 - orthogonale, 202
- Pullback-Abbildung, 89, 122, 268
- quadratische Funktion, 125
- quadratische Matrix, 3
- Quaternion, 50, 252
- Quotientenraum, 358
- Rand, 156
- raumartig, 132
- Rechtsnebenklasse, 356
- reelle Achse, 9
- Regression, 339, 341
- Resonanz, 221
- Reziprokenabbildung, 91
- Riemann'sche Zahlenkugel, 94
- Ruhemasse, 132
- Satz
 - von Cayley-Hamilton, 6, 57, 211
 - von Hahn-Banach, 170
 - von Jordan und v. Neumann, 195
- Scheitelmenge, 84
- Schwerpunktskoordinaten, 112
- schwingende Saite, 231, 233, 236, 237
- selbstadjungiert, 203, 241, 274, 332, 338
- Seminorm, 161
- Sesquilinearform, 191
- Shift-Operator, 241
- Singulärwertzerlegung, 331, 333
- Skalarenkörper, 295
- Spaltenrang, 305
- sphärische Bewegungen, 101
- Spur einer Matrix, 6, 7
- Spurfunktion, 271
- Stützfunktion, 169
- Stützhyperebene, 171, 175
- Standarddarstellung, 260, 268
- stehende Welle, 229
- stetige Linearform, 163
- subadditiv, 153
- Summe
 - direkte, 202
 - orthogonale, 202
- Supremum, 351
- surjektiv, 83
- symmetrische Bilinearform, 126
- symmetrische Gruppe, 85
- symmetrische Matrix, 33
- Symmetrisierung, 261
- Tangentialebene, 171
- Tangentialraum, 113, 298
- Teilbarkeit, 52
- Teilungsverhältnis, 111
- Teilvektorraum, 295
- Tensorprodukt, 299, 327
- Transformationsgruppe, 86
- transitiv, 104
- transponierte Matrix, 8
- trigonometrische Reihen, 67, 238
- trigonometrisches Polynom, 63, 190
- Überlagerung von Grund-
 - und Oberschwingungen, 236
- Umgebung, 156
- Umkehrabbildung, 83
- Umzentrieren, 21
- unabhängig, linear, 295
- Ungleichung
 - Cauchy-Schwarz, 162
 - Hölder, 149, 164, 167
 - Jensen, 179
 - Minkowski, 163
- unitär, 33
- unitär diagonalisierbar, 34, 329, 333
- unitärer Operator, 264, 283
- unkorreliert, 341
- Untergruppe, 86, 356
- unterhalbstetig, 149, 157
- Vektor, 41, 44, 45
- Vektorraum
 - normiert, 161
- \mathbb{C} -Vektorraum, 42
- \mathbb{K} -Vektorraum, 41
- \mathbb{R} -Vektorraum, 41
- Verband, 351

- Verflechtungsoperator, 262
- Vertauschungsalgebra, 263
- Vervollständigung, 1
- Vielfachheit
 - eines Eigenwerts, 54
- Vierervektoren, 44
- Vierervektorimpuls, 132
- volles Urbild, 88
- vollständig, xi, xii, 147, 203

- Wahrscheinlichkeitsgewichtung auf \mathbb{Z} , 60
- Wellengleichung, 232, 234
- Wellenzahl, 115, 232, 235, 239
- Windungszahl, 142

- Z-Transformierte, 71
- Zeilenrang, 304
- zeitartig, 132
- Zeitreihe, 245
- Zerlegung
 - Φ -verträgliche, 261
 - direkte, 202
 - kanonische, 267, 275
 - orthogonale, 202
- Zielbereich, 83
- zurückgenommene Funktion, 89
- zweiseitige Folge, 59
- Zweizustandssystem, 250