

M1: WS 07/08

Robert Beig

9. März 2011

Inhaltsverzeichnis

1	Vorbemerkung	3
2	Lineare Algebra	5
2.1	Vektorräume	5
2.1.1	Vektoren und ihre Komponenten	5
2.1.2	Teilräume, direkte Summe	6
2.1.3	Matrizen und lineare Gleichungssysteme	7
2.1.4	Lineare Operatoren	8
2.1.5	Das Tensorprodukt $\mathbb{V} \otimes \mathbb{V}$	13
2.2	Inneres Produkt	14
2.2.1	Definition und wichtige Eigenschaften	14
2.2.2	Norm	15
2.2.3	Orthonormalsysteme	16
2.2.4	Orthogonalprojektionen	17
2.2.5	Lineare Funktionale und Adjungierte	19
2.2.6	Spezialfälle	20
2.2.7	Spektraltheorie linearer Operatoren	22
2.2.8	Dirac-Notation	25
2.2.9	Quantenmechanik	26
2.2.10	Symmetrische Operatoren, kleine Schwingungen	28
2.3	Anhang: Normalformen	29
3	ODE's	32
3.1	Lineare Systeme	32
3.1.1	Homogene autonome Systeme	32
3.1.2	Inhomogene Systeme	35
3.1.3	Lineare nicht-autonome Gleichungen	35
3.2	Nichtlineare Systeme	37
3.2.1	Autonome Systeme erster Ordnung	37
3.2.2	Der Existenz- und Eindeutigkeitsatz	42
3.2.3	Stabilität	44
3.2.4	Nichtlineare ODE's für $\mathbf{n} = \mathbf{1}$	45
3.2.5	$\mathbf{n} = \mathbf{2}$: planare dynamische Systeme	47

3.2.6	Nicht-autonome Systeme und Gleichungen höherer Ordnung	50
3.2.7	Exakte Differentialgleichungen, Differentialformen	55
4	Komplexe Analysis	58
4.1	Komplexe Zahlen	58
4.2	Potenzreihen	58
4.3	Exp und Log	59
4.4	Analytische Funktionen	60
4.5	Holomorphie	61
4.6	Residuensatz	65
4.7	Anwendungen	66

Kapitel 1

Vorbemerkung

Dieser Vorlesungsbehef soll ein Gerüst wichtiger Tatsachen liefern, die hoffentlich das Erlernen des Stoffs erleichtern. Mit Beweisen wird sehr unterschiedlich verfahren. Diese können gänzlich fehlen, z.B. weil sie als aus der Vorlesung "Lineare Algebra" bekannt vorausgesetzt werden, in der Vorlesung nachgeliefert oder eben weggelassen werden. (Umgekehrt wird aber fallweise in der Vorlesung auf das Skriptum verwiesen werden.) Weiters gibt es Beweisskizzen, die einem vollständigen Beweis nahekommen und von den ehrgeizigeren unter ihnen vielleicht vervollständigt werden können. Vollständige Beweise gibt es hier auch, aber nur fallweise. Das ist einer von vielen Gründen, warum das Skriptum den Besuch der Vorlesung keinesfalls ersetzen kann.

Zum Programm der Vorlesung will ich folgendes sagen: Die Rolle der Mathematik in der Physik ist vielschichtig. Sie ist eine Methode zum Lösen von Problemen, aber darüber hinaus die Sprache, in der die Naturgesetze und die darin vorkommenden Konzepte formuliert sind. Daher ist es hilfreich, wenn Physiker, auch die experimentell arbeitenden, ein Verständnis der Mathematik entwickeln, das ein wenig über deren Aufgaben als Rechenmethode hinausgeht.

Eine Schwierigkeit beim Lernen ist, dass in den physikalischen Theorien praktisch immer mehrere Theorien der Mathematik gleichzeitig zum Einsatz kommen und im Zuge eines Physikstudiums nicht die Zeit ist, all diese Theorien vorweg zu lernen. Anders als in der reinen Mathematik ist man also zuweilen auf "learning by doing" angewiesen. Das setzt aber voraus, dass es eine gewisse eiserne Reserve an mathematischen Konzepten und Techniken gibt, die man wirklich versteht, sodass man den Rest in gutem Glauben hinnehmen und anwenden kann, ohne den Halt zu verlieren. Daraus ergibt sich die Maxime: "Im Zweifel weniger, und das ordentlich".

Der erste Abschnitt über lineare Algebra ist z.T. Wiederholung von Bekanntem, z.T. eine Einführung in orthogonale und unitäre Vektorräume, die in vielen Gebieten der Physik von grosser Bedeutung sind (z.B. unitäre Vektorräume in der Quantenmechanik). Der zweite Abschnitt handelt von Systemen gewöhnlicher Differentialgleichungen (kurz: "ODE's" für "Ordinary Differential Equations").

Die traditionell beliebten speziellen Lösungsansätze für spezielle Gleichungstypen werden geopfert, im Zeitalter von Computeralgebrasystemen wie Mathematica oder Maple ohne schlechtes Gewissen. Betont wird, wegen seiner überragenden Rolle in der Physik und weit über sie hinaus, der Gesichtspunkt der gewöhnlichen Differentialgleichung (bzw. eines Systems von solchen) als dynamisches System. Die Funktionentheorie muss aus Zeitgründen im Eiltempo absolviert werden. Hier folgen wir im wesentlichen dem Buch von H.Cartan, "Elementare Theorie der analytischen Funktionen einer oder mehrerer komplexen Veränderlichen" (B.I-Hochschultaschenbücher 1966).

Kapitel 2

Lineare Algebra

2.1 Vektorräume

2.1.1 Vektoren und ihre Komponenten

Hier sind ein paar Grundelemente der linearen Algebra zusammengestellt. Die Vektorraum-Axiome werden hier nicht wiederholt (Vorschlag: schlagen Sie in Ihrer Mitschrift aus linearer Algebra nach). Skalare sind für uns reelle oder komplexe Zahlen. Wichtigstes Konzept ist das der linearen Unabhängigkeit von Vektoren. Zwei Vektoren in einem Vektorraum \mathbb{V} heißen linear unabhängig, wenn keiner proportional (ein skalares Vielfaches) des anderen ist. Insbesondere ist keiner der Vektoren der Nullvektor. Eine Kollektion \mathbf{e}_i ($i = 1, 2, \dots, k$) von k Vektoren heisst linear unabhängig (l.u.), wenn die Vektorgleichung für die Skalare α_i

$$\sum_{i=1}^k \alpha_i \mathbf{e}_i = \mathbf{0}, \quad (2.1)$$

wobei die rechte Seite der Nullvektor in \mathbb{V} ist, einzig die triviale Lösung $\alpha_1 = 0 = \alpha_2 = \dots = \alpha_k$ hat. Diese Bedingung ist äquivalent damit, dass keiner der Vektoren \mathbf{e}_i sich als Linearkombination der anderen schreiben lässt. Sie ist weiters äquivalent damit, dass für jeden Vektor, der eine Linearkombination der \mathbf{e}_i ist (man sagt auch: "für jeden Vektor in der linearen Hülle von $\{\mathbf{e}_i\} \dots$ "), diese Linearkombination eindeutig ist. Wenn eine Kollektion von Vektoren l.u. ist, dann auch jede Teilkollektion davon. Angenommen, wir haben eine Kollektion von k Vektoren, die l.u. sind. Dann kann man versuchen, einen weiteren Vektor hinzuzufügen, sodass die entstehende Menge noch immer l.u. ist. Wenn dieser Prozess bei einer Zahl n von Vektoren aufhört, heisst diese Kollektion eine Basis von \mathbb{V} . Jedes Element von \mathbb{V} lässt sich dann eindeutig als Linearkombination von Elementen dieser Basis darstellen. Man kann weiters zeigen, dass alle Basen, d.h. alle Kollektionen von Vektoren, die l.u. sind und deren lineare Hülle ganz \mathbb{V} ist, aus n Elementen besteht. Es macht also Sinn, von n als "der Dimension von \mathbb{V} "

zu sprechen.

Wenn \mathbf{x} ein beliebiger Vektor ist und \mathbf{e}_i eine Basis bilden, gibt es eindeutige Zahlen x_i , sodass $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ ist. Die Zahlen x_i heissen die Komponenten des Vektors \mathbf{x} bezüglich der Basis \mathbf{e}_i . In diesem Sinn ist jeder n -dimensionale Vektorraum \mathbb{V} im reellen Fall äquivalent dem \mathbb{R}^n , im komplexen Fall mit \mathbb{C}^n .

Zwischenfrage: Was sind die Komponenten eines Basisvektors, z.B. \mathbf{e}_i , bezüglich der Basis $\{\mathbf{e}_j, j = 1, \dots, n\}$? Antwort: δ_{ij} , wobei δ_{ij} - das Kronecker-Delta - gegeben ist durch: $\delta_{ij} = 1$, wenn $i = j$ ist, andernfalls ist δ_{ij} gleich Null.

Es ist manchmal wichtig, zwischen dem Vektor \mathbf{v} als Element von \mathbb{V} und seiner Komponentendarstellung in einer konkreten Basis zu unterscheiden. Angenommen \mathbf{e}'_i bilden eine andere Basis. Dann gibt es Zahlen t_{ij} , sodass $\mathbf{e}_i = \sum_{j=1}^n t_{ji} \mathbf{e}'_j$. Die Komponenten x'_i des Vektors \mathbf{x} bezüglich der Basis \mathbf{e}'_i sind daher

$$x'_i = \sum_{j=1}^n t_{ij} x_j \quad (2.2)$$

Bemerkung: Man kann auch \mathbf{e}'_i durch die \mathbf{e}_j - und daher x_i durch x'_i - ausdrücken, und zwar ist

$$x_i = \sum_{j=1}^n s_{ij} x'_j \quad (2.3)$$

wobei $s_{ij} = (t^{-1})_{ij}$ die zu t_{ij} inverse Matrix ist, also

$$\sum_{j=1}^n s_{ij} t_{jk} = \delta_{ik}. \quad (2.4)$$

Bemerkung: man vergegenwärtige sich bei den Formeln (2.1 - 2.4) die Bedeutung des Summenzeichens, insbesondere "freie Indizes vs. Summationsindizes". Zur Sicherheit sollte man diese Relationen in den Fällen $n = 1, 2, 3$ explizit, d.h. termweise aufschreiben.

2.1.2 Teilräume, direkte Summe

Eine Teilmenge \mathbb{T} des Vektorraums \mathbb{V} heisst Teilraum von \mathbb{V} , wenn er abgeschlossen bezüglich der Vektorraumoperationen - d.h. Addition und skalare Multiplikation - in \mathbb{V} ist. Wenn es zwei Teilräume \mathbb{S} und \mathbb{T} von \mathbb{V} gibt, sodass jeder Vektor $\mathbf{x} \in \mathbb{V}$ sich als Summe eines Vektors $\mathbf{u} \in \mathbb{S}$ und eines Vektors $\mathbf{v} \in \mathbb{T}$ schreiben lässt und diese Zerlegung eindeutig ist, heisst \mathbb{V} die direkte Summe von \mathbb{S} und \mathbb{T} , und wir schreiben $\mathbb{V} = \mathbb{S} \oplus \mathbb{T}$. Dies ist für zwei Teilräume genau dann der Fall, wenn die Summe ihrer Dimensionen $= n$ ist und sie nur den Vektor $\mathbf{0} \in \mathbb{V}$ gemeinsam haben.

2.1.3 Matrizen und lineare Gleichungssysteme

Eine $n \times m$ - Matrix \mathbf{A} ist eine Anordnung der Art $(\mathbf{A})_{ij} = a_{ij}$, wobei a_{ij} reelle oder komplexe Zahlen sind und $i = 1, \dots, n$ und $j = 1 \dots m$ ist. Man kann sich \mathbf{A} als rechteckiges Zahlenschema mit n Zeilen und m Spalten vorstellen. In diesem ist z.B. a_{32} die Zahl in der 3. Zeile der 2. Spalte. Solche Matrizen treten in erster Linie als Darstellungen linearer Abbildungen von einem m -dimensionalen Vektorraum in einen n -dimensionalen Vektorraum auf, wenn in beiden Vektorräumen eine Basis gewählt worden ist. Wir betrachten sie in diesem Abschnitt als Objekte "per se". Hier ein paar Definitionen. Sei \mathbf{A} eine $n \times m$ - Matrix:

1. transponierte Matrix \mathbf{A}^T : eine $m \times n$ - Matrix \mathbf{A}^T der Form $(\mathbf{A}^T)_{ji} = a_{ij}$
2. komplex konjugierte Matrix $\bar{\mathbf{A}}$: klar
3. adjungierte Matrix \mathbf{A}^\dagger : eine $m \times n$ - Matrix der Form $(\mathbf{A}^\dagger)_{ji} = \overline{a_{ij}}$
4. Addition und skalare Multiplikation: Matrizen desselben Typs kann man gliedweise addieren und mit Skalaren multiplizieren.
5. Multiplikation: eine $n \times m$ - Matrix \mathbf{A} und eine $m \times p$ - Matrix \mathbf{B} kann man multiplizieren. Das Ergebnis ist eine $n \times p$ - Matrix \mathbf{C} der Form $c_{ij} = \sum_{k=1}^m a_{ik}b_{kj}$.

Ein lineares Gleichungssystem bestehend aus n Gleichungen für m Unbekannte hat die Gestalt

$$\sum_{j=1}^m a_{ij}x_j = y_i. \quad (2.5)$$

Mithin ist $x_i \in \mathbb{R}^m(\mathbb{C}^m)$ und $y_i \in \mathbb{R}^n(\mathbb{C}^n)$. Wir bezeichnen mit $im(\mathbf{A})$ die Menge aller $y_i \in \mathbb{R}^n(\mathbb{C}^n)$, für die Glg. (2.5) (eine oder mehrere) Lösungen hat. Wir bezeichnen mit $ker(\mathbf{A})$ die Menge aller $x_i \in \mathbb{R}^m(\mathbb{C}^m)$, sodass Glg. (2.5) für $y_i = 0$ erfüllt ist. Analog ist $ker(\mathbf{A}^T)$ die Menge aller $y_i \in \mathbb{R}^n(\mathbb{C}^n)$, sodass $\sum_{i=1}^n y_i a_{ij} = 0$ ist. Diese drei Räume sind lineare Teilräume der jeweiligen Vektorräume, denen sie angehören. Die Dimension von $im(\mathbf{A})$ heisst Rang $\rho(\mathbf{A})$ des Gleichungssystems bzw. der Matrix \mathbf{A} . Dieser ist auch dasselbe wie die Dimension der linearen Hülle der Spaltenvektoren $\{a_{i1}, a_{i2}, \dots, a_{im}\}$. Sei \mathbb{T} ein Teilvektorraum von $\mathbb{R}^k(\mathbb{C}^k)$. Dann definieren wir den Annihilator $\mathbb{T}^\perp \subset \mathbb{R}^k(\mathbb{C}^k)$ von \mathbb{T} als die Menge aller z_i , $i = 1, \dots, k$, sodass $\sum_{i=1}^k z_i x_i = 0$ für alle $x_i \in \mathbb{T}$. Man kann zeigen, dass \mathbb{T}^\perp ein Teilvektorraum der Dimension $k - l$ ist, wenn $l \leq k$ die Dimension von \mathbb{T} ist. Fundamental sind nunmehr die folgenden Aussagen:

- **G1** ("Fredholm'sche Alternative"): $im(\mathbf{A}) = (ker(\mathbf{A}^T))^\perp$

Bemerkung: Die Aussage $im(\mathbf{A}) \subset (ker(\mathbf{A}^T))^\perp$ wäre offensichtlich gewesen. Warum?

- **G2** ("Spaltenrang = Zeilenrang"): $\rho(\mathbf{A}) = \rho(\mathbf{A}^T)$

Wenn man **G1** auf \mathbf{A}^T anwendet und **G2** benutzt sowie $(\mathbf{A}^T)^T = \mathbf{A}$, folgt weiters

- **G3**: $\dim(\ker(\mathbf{A})) + \rho(\mathbf{A}) = m$

Die Zahl $\dim(\ker(\mathbf{A}))$ gibt ein Mass für die Nicht-Eindeutigkeit der Lösungen von (2.5). Daraus ergibt sich folgender - hier etwas provokant formulierter - Sachverhalt:

- **G4**: Sei in Glg.(2.5) $m = n$. Dann ist (2.5) genau dann lösbar für alle $y_i \in \mathbb{R}^n(\mathbb{C}^n)$, wenn die Lösung eindeutig ist. In diesem Fall existiert die inverse Matrix \mathbf{A}^{-1} , die die Relationen $\sum_{j=1}^n (a^{-1})_{ij} a_{jk} = \delta_{ik} = \sum_{j=1}^n a_{ij} (a^{-1})_{jk}$ erfüllt. Solche Matrizen heissen regulär oder nicht-singulär.

2.1.4 Lineare Operatoren

Sei also nun \mathbf{A} eine lineare Abbildung von \mathbb{V} nach \mathbb{V} (Dimension von \mathbb{V} sei n), also

$$\mathbf{A}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{A} \mathbf{x}_1 + \alpha_2 \mathbf{A} \mathbf{x}_2 \quad (2.6)$$

für beliebige Vektoren $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{V}$ und Skalare α_1, α_2 . Dann folgt für $\mathbf{y} = \mathbf{A} \mathbf{x} = \sum_{i=1}^n y_i \mathbf{e}_i$, dass

$$y_i = \sum_{j=1}^n a_{ij} x_j, \quad (2.7)$$

wobei a_{ij} gegeben ist durch $\mathbf{A} \mathbf{e}_j = \sum_{i=1}^n a_{ij} \mathbf{e}_i$. Die lineare Abbildung \mathbf{A} ist also bezüglich der Basis \mathbf{e}_i gegeben durch die quadratische $(n \times n)$ -Matrix a_{ij} reeller oder komplexer Zahlen.

Zwischenbemerkung 1: Man beachte den begrifflichen Unterschied zwischen der "passiven" Transformation (2.2) und der "aktiven Transformation" (2.6). Insbesondere ist eine passive Transformation immer umkehrbar.

Zwischenbemerkung 2: Oft ist es im reellen Fall (a_{ij} reell) nützlich, im nachhinein komplexe Werte für die Skalare - und daher Koordinaten von Vektoren - zuzulassen. Der Grund ist (siehe später), dass Operatoren im komplexen Vektorraum immer Eigenwerte haben (die dann i.a. komplex sind), nicht aber solche im reellen Vektorraum.

Wichtiges Beispiel eines Operators ist die Dilatation $\mathbf{D} \mathbf{x} = d \mathbf{x}$. Wenn $d = 0$ ist, ist das die Null-Abbildung, wenn $d = 1$ ist, die identische Abbildung \mathbf{E} . Die Matrixdarstellung der obigen Dilatation ist $d_{ij} = d \delta_{ij}$, wobei δ_{ij} das Kronecker-Delta ist. Ist \mathbf{B} eine weitere lineare Abbildung mit Matrixdarstellung b_{ij} bezüglich der Basis \mathbf{e}_i , dann ist die zusammengesetzte Abbildung $\mathbf{C} = \mathbf{B} \mathbf{A}$ ebenfalls eine

lineare Abbildung, und es gilt

$$c_{ij} = \sum_{k=1}^n b_{ik} a_{kj} . \quad (2.8)$$

Wenn erste Indizes als Zeilenindizes, zweite Indizes als Spaltenindizes aufgefasst werden, entspricht (2.8) gerade der üblichen Matrixmultiplikation.

Beachte: der Operator $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$ ist im allgemeinen nicht die Null-Abbildung, dementsprechend ist das Matrix-Produkt (2.8) nicht-kommutativ, d.h. im allgemeinen ist

$$[a, b]_{ij} = \sum_{k=1}^n (a_{ik} b_{kj} - b_{ik} a_{kj}) \neq 0 \quad (2.9)$$

Beachte: die Reihenfolge der Faktoren a, b auf der r.S. von (2.9) spielt keine Rolle, d.h. zum Beispiel ist klarerweise $a_{ik} b_{kj} = b_{kj} a_{ik}$, wohl aber die Verteilung der Indizes auf a, b ! Sehr wohl kommutieren lineare Abbildungen für $n = 1$ und solche, für die es eine Basis gibt, in der ihre jeweiligen Matrixdarstellungen beide diagonal sind.

Unter Basisänderung (siehe Absatz vor (2.2)) verhält sich die Matrixdarstellung a_{ij} eines Operators \mathbf{A} wie folgt:

$$a'_{ij} = \sum_{k=1}^n \sum_{l=1}^n t_{ik} a_{kl} (t^{-1})_{lj} . \quad (2.10)$$

Gleichung (2.10) kann man auch in Matrix-Notation schreiben:

$$\mathbf{A}' = \mathbf{TAT}^{-1} . \quad (2.11)$$

Hier ist allerdings zu beachten, dass die Matrizen \mathbf{A} und \mathbf{A}' als verschiedene Darstellungen ein und desselben Objekts aufgefasst werden.

Bemerkung: Die Matrixkomponenten einer Dilatation sind in jeder Basis dieselben.

Invariante Teilräume

Ein Teilraum $\mathbb{T} \subset \mathbb{V}$ heisst invariant unter dem Operator \mathbf{A} , wenn $\mathbf{A}\mathbb{T} \subset \mathbb{T}$. Sei $\mathbb{S} \oplus \mathbb{T} = \mathbb{V}$, und beide Teilräume invariant unter \mathbf{A} . Seien $\mathbf{B} = \mathbf{A}_{\mathbb{S}}$ und $\mathbf{C} = \mathbf{A}_{\mathbb{T}}$ die Einschränkungen von \mathbf{A} auf die jeweiligen Teilräume. Dann hat, bezüglich einer Basis von \mathbb{V} , die nur aus Elementen besteht, die einem der beiden Teilräume angehören, die Gleichungsmatrix von \mathbf{A} die Blockgestalt

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}$$

Die Spur

Die Spur eines Operators \mathbf{A} ist definiert als die Summe der Diagonalelemente der Matrix a_{ij} . Dass die Bezeichnung "Spur der linearen Abbildung" gerechtfertigt ist, liegt an folgendem Umstand: die Spur $\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$ ändert sich nicht unter Basistransformationen (2.10) (anders ausgedrückt: "tr(\mathbf{A}) ist eine Invariante der linearen Abbildung \mathbf{A} "), denn:

$$\sum_{i=1}^n a'_{ii} = \sum_{i=1}^n \sum_{k=1}^n \sum_{l=1}^n t_{ik} a_{kl} (t^{-1})_{li} = \sum_{k=1}^n \sum_{l=1}^n a_{kl} \delta_{lk} = \sum_{k=1}^n a_{kk} \quad (2.12)$$

wobei wir im ersten Gleichheitszeichen (2.10), im zweiten (2.4) und im letzten die Definition des Kronecker-Deltas verwendet haben.

Bemerkung 1: Für die Spur der Einheitsabbildung gilt: $\text{tr}(\mathbf{E}) = n$.

Bemerkung 2: Es gilt $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$.

Weiters: wenn $\mathbf{C} = \mathbf{AB}$, also $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$, dann gilt auch

$$c'_{ij} = \sum_{k=1}^n a'_{ik} b'_{kj} \quad (2.13)$$

Wir können also weitere Invarianten von \mathbf{A} gewinnen, wenn wir \mathbf{A}^2 , \mathbf{A}^3 , etc. bilden und dann die Spur nehmen. Wir erwähnen nebenbei, dass es einen Zusammenhang gibt zwischen den so gebildeten Invarianten und einer anderen Invarianten von \mathbf{A} , nämlich der Determinante $\det(\mathbf{A})$.

Determinante

Wir erinnern an eine mögliche Definition der Determinante. Sei $\epsilon_{i_1 i_2 \dots i_n}$ das total antisymmetrische Symbol in n Dimensionen. D.h. $\epsilon_{i_1 i_2 \dots i_n}$ ändert das Vorzeichen, wenn zwei beliebige Indizes vertauscht werden und $\epsilon_{12 \dots n} = 1$. Dann ist

$$\det(\mathbf{A}) = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} a_{i_2 2} \dots a_{i_n n} \epsilon_{i_1 i_2 \dots i_n} \quad (2.14)$$

Daraus kann man viele Eigenschaften der Determinante herleiten. Wir erwähnen nur einige wenige:

- **D1:** $\det(\mathbf{E}) = 1$, $\det(-\mathbf{E}) = (-1)^n$
- **D2:** $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$
- **D3:** $\det(\mathbf{A}^T) = \det(\mathbf{A})$, wobei \mathbf{A}^T die zur transponierten Matrix gehörige lineare Abbildung ist.

- **D4:** Der Operator \mathbf{A} ist invertierbar genau dann, wenn $\det(\mathbf{A})$ nicht gleich Null ist.
- **D5:** Wenn \mathbf{A} Blockgestalt hat (siehe Abschn. invariante Teilräume), gilt:

$$\det(\mathbf{A}) = \det(\mathbf{B})\det(\mathbf{C}) \quad (2.15)$$

Genau genommen, haben wir die Determinante für quadratische Matrizen und nicht lineare Operatoren definiert. Aus **D2** folgt aber die Invarianz der Determinante, denn $\det(\mathbf{TAT}^{-1}) = \det(\mathbf{T}^{-1}\mathbf{TA}) = \det(\mathbf{A})$.

Wir kommen kurz auf den Zusammenhang mit der Spur zurück. Dieser wird immer komplizierter, je grösser n ist. Für $n = 2$ gilt (bitte nachrechnen!)

$$2 \det(\mathbf{A}) = (\operatorname{tr}(\mathbf{A}))^2 - \operatorname{tr}(\mathbf{A}^2). \quad (2.16)$$

Zwischenfrage: warum kann z.B. für $n = 2$ die folgende Formel niemals richtig sein: $2 \det(\mathbf{A}) = (\operatorname{tr}(\mathbf{A}))^2 - \operatorname{tr}(\mathbf{A})$??

Für $n = 3$ gilt (Mathematica ?)

$$6 \det(\mathbf{A}) = (\operatorname{tr}(\mathbf{A}))^3 - 3 \operatorname{tr}(\mathbf{A}^2) \operatorname{tr}(\mathbf{A}) + 2 \operatorname{tr}(\mathbf{A}^3) \quad (2.17)$$

Eigenwerte und Eigenvektoren

Ein Skalar λ heisst Eigenwert des Operators \mathbf{A} , wenn es einen vom Nullvektor verschiedenen Vektor \mathbf{x} im n -dimensionalen Vektorraum \mathbb{V} gibt, sodass

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x} \iff (\mathbf{A} - \lambda \mathbf{E}) \mathbf{x} = \mathbf{0} \iff \sum_{j=1}^n (a_{ij} - \lambda \delta_{ij}) x_j = 0 \quad (2.18)$$

ist. Wenn \mathbb{V} komplex ist, gibt es immer Eigenwerte, und von diesen können höchstens n voneinander verschieden sein. Der Grund ist folgender: λ ist genau dann Eigenwert (siehe **G4** und **D4**), wenn der Operator $\mathbf{A} - \lambda \mathbf{E}$ singular - d.h. nicht invertierbar - ist, was äquivalent ist mit

$$\det(\mathbf{A} - \lambda \mathbf{E}) = 0. \quad (2.19)$$

Die linke Seite von Glg.(2.19) hat die Form eines Polynom n -ten Grades ("charakteristisches Polynom")

$$P(\lambda) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0, \quad (2.20)$$

daher hat (2.19) nach dem Fundamentalsatz der Algebra immer Lösungen in \mathbb{C} , von denen maximal n voneinander verschieden sind. Die Menge der Eigenwerte heisst Spektrum der linearen Transformation. Wenn \mathbb{V} reeller Vektorraum ist,

kann man seine sog. Komplexifizierung $\mathbb{V}^{\mathbb{C}}$ betrachten. Am einfachsten erhält man diese so: man nimmt eine Basis $\{e_i\}$ und lässt einfach für die Komponenten von Vektoren bezüglich dieser Basis komplexe Werte zu. Es resultiert ein komplexer Vektorraum der selben Dimension. Die lineare Transformation \mathbf{A} wird sodann zu einer reellen Transformation in diesem komplexen Vektorraum, also mit reeller Matrixdarstellung in der beschriebenen Basis. Das resultierende Eigenwertproblem (2.18) führt zur selben Glg. $P(\lambda) = 0$ mit reellen Koeffizienten a_0, \dots, a_{n-1} , deren im allgemeinen komplexe Nullstellen aber als Eigenwerte Sinn machen.

Bemerkung: In einem reellen Vektorraum gibt es zu jedem komplexen Eigenwert auch den komplex konjugierten. Daraus folgt z.B., dass bei ungerader Dimension es mindestens einen reellen Eigenwert geben muss.

Die Eigenwerte sind, wie die Determinante oder die Spur, Invarianten der linearen Abbildung. Daher muss dies auch für die Koeffizienten des charakteristischen Polynoms gelten. Dies sieht man direkt aus $\mathbf{TAT}^{-1} - \lambda\mathbf{E} = \mathbf{T}(\mathbf{A} - \lambda\mathbf{E})\mathbf{T}^{-1}$ und der Invarianz der Determinante.

Beispiel: Im Fall $n = 2$ lautet das charakteristische Polynom

$$P(\lambda) = \lambda^2 - \text{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) \quad (2.21)$$

und im Fall $n = 3$

$$P(\lambda) = -\lambda^3 + \text{tr}(\mathbf{A})\lambda^2 - \frac{1}{2}[(\text{tr}(\mathbf{A}))^2 - \text{tr}(\mathbf{A}^2)]\lambda + \det(\mathbf{A}) \quad (2.22)$$

Folgendes ist ein allgemeines Faktum ("Satz von Cayley-Hamilton"), das wir nicht beweisen: Für jede Abbildung \mathbf{A} gilt, dass

$$P(\mathbf{A}) = \mathbf{0} \quad (2.23)$$

Dies ist so zu verstehen, dass z.B. λ^2 durch \mathbf{A}^2 zu ersetzen ist und λ^0 durch \mathbf{E} . Wenn wir (2.23) für $n = 3$ anwenden und die Spur bilden, erhalten wir unter Verwendung von (2.22) die Identität (2.17).

Sei nun λ ein Eigenwert von \mathbf{A} . Die Menge aller Vektoren \mathbf{x} , die Gleichung (2.18) lösen, bilden einen Teilraum \mathbb{T}_λ von \mathbb{V} . Dieser Teilraum ist invariant unter \mathbf{A} , d.h. $\mathbf{A}\mathbb{T} \subset \mathbb{T}$. (Wenn \mathbb{T}_λ ein-dimensional ist, heisst λ "nicht ausgeartet".) Man kann beweisen: Wenn $\{\mathbf{x}_i\}$ Eigenvektoren zu paarweise verschiedenen Eigenwerten sind, sind sie linear unabhängig. Angenommen nun \mathbf{A} hat n voneinander verschiedene Eigenwerte λ_i . Dann sind diese Eigenwerte nicht ausgeartet. Weiters spannen die zugehörigen "Eigenräume" \mathbb{T}_i ganz \mathbb{V} auf. M.a.W. gibt es eine Basis, die aus Eigenvektoren besteht. Das heisst weiters folgendes: Es gibt eine Basistransformation \mathbf{T} , sodass die transformierte Matrix $\mathbf{A}' = \mathbf{TAT}^{-1}$ eine Diagonalmatrix ist. Man sagt: " \mathbf{A} ist diagonalisierbar". Diagonalelemente sind gerade die Eigenwerte λ_i . Nicht jede lineare Abbildung (resp. quadratische Matrix) ist diagonalisierbar.

Beispiel 1: Sei $n = 2$ und $\mathbf{A}\mathbf{e}_1 = \mathbf{0}$, $\mathbf{A}\mathbf{e}_2 = \mathbf{e}_1$. Dann ist $\lambda = 0$ der einzige Eigenwert (nicht ausgeartet). Die Matrix \mathbf{A} ist aber nicht diagonalisierbar.

Beispiel 2: Die Matrix

$$\mathbf{A} = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$$

ist diagonalisierbar, denn $\det(\mathbf{A} - \lambda\mathbf{E}) = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2)$. Die Vektoren $\mathbf{f}_1 = (1, 1)$ resp. $\mathbf{f}_2 = (2, 1)$ sind Eigenvektoren zu $\lambda_1 = -1$ resp. $\lambda_2 = 2$. Aus (2.1.1) folgt, dass der Zusammenhang zwischen neuer und alter Basis gegeben ist durch $\mathbf{f}_i = \sum_j s_{ji}\mathbf{e}_j$. Die alten Basiselemente haben per definition die Komponenten $(\mathbf{e}_i)_j = \delta_{ij}$, daher ist $(\mathbf{f}_i)_j = s_{ji}$. Daher ist

$$\mathbf{S} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad \mathbf{S}^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Daher muss gelten, dass

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

ist. Das ergibt auch die explizite Rechnung.

2.1.5 Das Tensorprodukt $\mathbb{V} \otimes \mathbb{V}$

Dieses Konzept spielt eine grosse Rolle in der Allgemeinen Relativitätstheorie sowie in der Quantentheorie. Wenn $\{\mathbf{e}_i\}$, $i = 1, \dots, n$ eine Basis bilden, betrachten wir Paare $(\mathbf{e}_i, \mathbf{e}_j)$, denen wir das Symbol $\mathbf{e}_i \otimes \mathbf{e}_j$ zuordnen, und zu diesen den n^2 -dimensionalen Vektorraum $\mathbb{V} \otimes \mathbb{V}$ aller ihrer Linearkombinationen $\mathbf{r} = \sum_{i,j} r_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$. Das Tensorprodukt $\mathbf{e}_i \otimes \mathbf{e}_j$ der Basiselemente wird nun zu einer Abbildung erweitert, die jedem Paar \mathbf{x}, \mathbf{y} von Vektoren in \mathbb{V} ein Element $\mathbf{r} = \mathbf{x} \otimes \mathbf{y}$ aus $\mathbb{V} \otimes \mathbb{V}$ zuordnet, und zwar durch die Forderung der Linearität in beiden Faktoren, also

$$\mathbf{x} \otimes \mathbf{y} = \sum_{i,j} x_i y_j \mathbf{e}_i \otimes \mathbf{e}_j$$

Wir ersparen uns den Beweis, dass diese Konstruktion des Raums $\mathbb{V} \otimes \mathbb{V}$ und des Tensorprodukts von Vektoren in \mathbb{V} von der gewählten Basis unabhängig ist. Elemente von $\mathbb{V} \otimes \mathbb{V}$ heissen auch "kontravariante Tensoren 2-ter Stufe". Ihre Komponenten r_{ij} verhalten sich unter Basistransformationen folgendermassen

$$r'_{ij} = \sum_{k,l} t_{ik} t_{jl} r_{kl}$$

oder, in Matrixnotation,

$$\mathbf{R}' = \mathbf{T}\mathbf{R}\mathbf{T}^T \tag{2.24}$$

Natürlich ist nicht jedes Element in $\mathbf{r} \in \mathbb{V} \otimes \mathbb{V}$ zerlegbar, d.h. von der Form $\mathbf{r} = \mathbf{x} \otimes \mathbf{y}$, also $r_{ij} = x_i y_j$. Zum Beispiel für $n = 2$ ist das Kriterium dafür gerade, dass $\det(r_{ij}) = 0$ ist. Weiters ist im allgemeinen $\mathbf{x} \otimes \mathbf{y} \neq \mathbf{y} \otimes \mathbf{x}$.

Wegen der Anwendungen in der Quantenmechanik erwähnen wir noch folgende Konstruktion: Wenn \mathbf{A}, \mathbf{B} Abbildungen von \mathbb{V} sind, induziert das eine lineare Abbildung $\mathbf{A} \otimes \mathbf{B}$ von $\mathbb{V} \otimes \mathbb{V}$ gemäss

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{x} \otimes \mathbf{y}) = \mathbf{A}\mathbf{x} \otimes \mathbf{B}\mathbf{y} \quad (2.25)$$

2.2 Inneres Produkt

2.2.1 Definition und wichtige Eigenschaften

Ein inneres Produkt in einem reellen oder komplexen Vektorraum \mathbb{V} ist eine Abbildung $\langle | \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ (oder \mathbb{C}) mit den Eigenschaften:

- **S1:** $\langle \mathbf{x} | \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{x} | \mathbf{z} \rangle$
- **S2:** $\langle \mathbf{x} | \alpha \mathbf{y} \rangle = \alpha \langle \mathbf{x} | \mathbf{y} \rangle$
- **S3:** $\langle \mathbf{y} | \mathbf{x} \rangle = \overline{\langle \mathbf{x} | \mathbf{y} \rangle}$

Es folgt, dass $\langle | \rangle$ "konjugiert-linear" im ersten Faktor ist, also $\langle \alpha \mathbf{x} + \beta \mathbf{y} | \mathbf{z} \rangle = \bar{\alpha} \langle \mathbf{x} | \mathbf{z} \rangle + \bar{\beta} \langle \mathbf{y} | \mathbf{z} \rangle$. Ausserdem ist $\langle \mathbf{x} | \mathbf{x} \rangle$ reell.

- **S4:** $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$ und $\langle \mathbf{x} | \mathbf{x} \rangle = 0 \Rightarrow \mathbf{x} = \mathbf{0}$

Aus **S4** folgt insbesondere

- **S4':** Wenn

$$\langle \mathbf{y} | \mathbf{x} \rangle = 0 \quad \forall \mathbf{x} \in \mathbb{V} \implies \mathbf{y} = \mathbf{0} \quad (2.26)$$

Aus diesen Axiomen folgt für die Gestalt des Skalarprodukts in einer Basis, dass

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{ij} g_{ij} \bar{x}_i y_j, \quad (2.27)$$

wobei die Matrix $g_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle$ hermitisch ist, d.h. $\mathbf{G}^\dagger = \mathbf{G}$ oder

$$g_{ij} = \bar{g}_{ji} \quad (2.28)$$

Unter Basistransformationen wie in (2.1.1) verhält sich die Matrix \mathbf{G} folgendermassen

$$\mathbf{T}^\dagger \mathbf{G}' \mathbf{T} = \mathbf{G} \iff \mathbf{G}' = (\mathbf{T}^\dagger)^{-1} \mathbf{G} \mathbf{T}^{-1} = (\mathbf{T}^{-1})^\dagger \mathbf{G} \mathbf{T}^{-1} \quad (2.29)$$

Wir sehen also: "Nicht alles, was durch eine Matrix beschrieben wird, transformiert sich gleich".

Folgender Satz ist nur im komplexen Fall richtig:

Satz: Wenn im komplexen Vektorraum \mathbb{V} für einen linearen Operator \mathbf{A} gilt

$$\langle \mathbf{x} | \mathbf{A}\mathbf{x} \rangle = 0 \quad \forall \mathbf{x} \in \mathbb{V}, \quad (2.30)$$

ist \mathbf{A} die Nullabbildung.

Beweis: Wir haben

$$\langle \mathbf{x} | \mathbf{A}\mathbf{y} \rangle + \langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle = \langle \mathbf{x} + \mathbf{y} | \mathbf{A}(\mathbf{x} + \mathbf{y}) \rangle - \langle \mathbf{x} | \mathbf{A}\mathbf{x} \rangle - \langle \mathbf{y} | \mathbf{A}\mathbf{y} \rangle = 0 \quad (2.31)$$

wobei das erste Gleichheitszeichen sich durch Rechnung und das zweite durch Anwendung von (2.30) ergibt. Setzen wir in (2.31) auf der l.S. \mathbf{y} durch $i\mathbf{y}$, erhalten wir

$$i \langle \mathbf{x} | \mathbf{A}\mathbf{y} \rangle - i \langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle = 0. \quad (2.32)$$

Aufgrund (2.31, 2.32) ist folglich $\mathbf{A}\mathbf{x}$ auf alle Vektoren orthogonal und daher - wegen $\mathbf{S}4'$ - gleich null, w.z.b.w.

2.2.2 Norm

Die Norm $\|\mathbf{x}\|$ eines Vektors $\mathbf{x} \in \mathbb{V}$ ist definiert als

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} \quad (2.33)$$

Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ heissen orthogonal, wenn $\langle \mathbf{x} | \mathbf{y} \rangle = 0$ ist. In diesem Fall gilt der

Satz von Pythagoras: Wenn \mathbf{x}, \mathbf{y} orthogonal sind, ist

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \quad (2.34)$$

Im reellen Fall gilt Gleichheit in (2.34) genau dann, wenn \mathbf{x}, \mathbf{y} orthogonal sind.

Beweis: Rechnung!

Wenn $\mathbf{y} \neq \mathbf{0}$ ist, kann man jeden Vektor \mathbf{x} eindeutig in einen zu \mathbf{y} parallelen und einen orthogonalen Vektor zerlegen:

$$\mathbf{x} = \frac{\langle \mathbf{y} | \mathbf{x} \rangle}{\|\mathbf{y}\|^2} \mathbf{y} + \left(\mathbf{x} - \frac{\langle \mathbf{y} | \mathbf{x} \rangle}{\|\mathbf{y}\|^2} \mathbf{y} \right) \quad (2.35)$$

Fundamental sind die folgenden beiden Ungleichungen:

Cauchy-Schwarz'sche Ungleichung:

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \quad (2.36)$$

wobei Gleichheit in (2.36) genau dann besteht, wenn \mathbf{x} und \mathbf{y} linear abhängig sind, also einer ein skalares Vielfaches des jeweils anderen ist.

Beweis: Vorlesung.

Aus (2.36) folgt die ebenfalls fundamentale Dreiecksungleichung

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (2.37)$$

Wenn \mathbb{V} reell ist, definieren wir den von Vektoren $\neq \mathbf{0}$ eingeschlossenen Winkel ¹ $\sphericalangle(\mathbf{x}, \mathbf{y})$ durch

$$\cos(\sphericalangle(\mathbf{x}, \mathbf{y})) = \frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \quad (2.38)$$

2.2.3 Orthonormalsysteme

Es gibt eine Basis $\{\mathbf{e}_i\}$, die bezüglich $\langle \mathbf{x} | \mathbf{y} \rangle$ orthonormal ist d.h. in der gilt

$$g_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle = \delta_{ij} \quad (2.39)$$

Beweisskizze ("Gram-Schmidt-Verfahren"): Wenn \mathbb{V} die Dimension zwei hat, starten wir mit einer Basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ und ersetzen diese durch das VONS $\{\mathbf{f}_1, \mathbf{f}_2\}$, wobei (vgl. (2.35))

$$\mathbf{f}_1 = \frac{1}{\|\mathbf{e}_1\|} \mathbf{e}_1, \quad \mathbf{f}_2 = \frac{1}{\|\mathbf{e}_2 - \frac{\langle \mathbf{e}_1 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1\|} \left(\mathbf{e}_2 - \frac{\langle \mathbf{e}_1 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 \right) \quad (2.40)$$

Wenn \mathbb{V} dreidimensional ist und \mathbf{e}_3 ein dritter Basisvektor, definieren wir

$$\mathbf{f}_3 = \frac{1}{\|\mathbf{e}_3 - \langle \mathbf{f}_1 | \mathbf{e}_3 \rangle \mathbf{f}_1 - \langle \mathbf{f}_2 | \mathbf{e}_3 \rangle \mathbf{f}_2\|} \left(\mathbf{e}_3 - \langle \mathbf{f}_1 | \mathbf{e}_3 \rangle \mathbf{f}_1 - \langle \mathbf{f}_2 | \mathbf{e}_3 \rangle \mathbf{f}_2 \right), \quad (2.41)$$

usw. W.z.b.w.

Diesen Sachverhalt kann man mit Hinsicht auf (2.29) auch so ausdrücken: Zu jeder hermiteschen, positiv definiten Matrix \mathbf{G} gibt es eine nicht-singuläre Matrix \mathbf{T} , sodass $\mathbf{G}' = (\mathbf{T}^{-1})^\dagger \mathbf{G} \mathbf{T}^{-1}$ die Einheitsmatrix ist.

In einer ON-Basis hat das Skalarprodukt mithin die Form

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_i \bar{x}_i y_i \quad (2.42)$$

Die Komponenten α_i eines Vektors \mathbf{x} bezüglich einer ON-Basis \mathbf{e}_i lauten

$$\alpha_i = \langle \mathbf{e}_i | \mathbf{x} \rangle. \quad (2.43)$$

Zum Beweis schreiben wir

$$\mathbf{x} = \sum_j \alpha_j \mathbf{e}_j \quad (2.44)$$

¹Genauer ist das der nicht-orientierte Winkel², und er ist durch diese Formel eindeutig nur im Intervall $[0, \pi]$ definiert.

und multiplizieren Glg.(2.44) von links skalar mit \mathbf{e}_i . Weiters gilt ("Parseval'sche Gleichung"):

$$\sum_i |\alpha_i|^2 = \|\mathbf{x}\|^2, \quad (2.45)$$

wobei die Betragsstriche auf der linken Seite im Sinn der komplexen Zahlen zu verstehen sind: $|z|^2 = \bar{z}z$.

2.2.4 Orthogonalprojektionen

Wenn \mathbb{T} ein Teilraum von \mathbb{V} ist, definieren wir sein orthogonales Komplement \mathbb{T}^\perp folgendermassen

$$\mathbb{T}^\perp = \{\mathbf{x} \in \mathbb{V} \mid \langle \mathbf{x} | \mathbf{y} \rangle = 0 \quad \forall \mathbf{y} \in \mathbb{T}\} \quad (2.46)$$

Zu beachten ist, dass diese Definition mit der im Abschnitt **2.1.3** nur im reellen Fall übereinstimmt. Wegen der Positiv-Definitheit des Skalarprodukts ist offensichtlich

$$\mathbb{T} \cap \mathbb{T}^\perp = \{\mathbf{0}\} \quad (2.47)$$

Darüber hinaus gilt

Satz: $\mathbb{T} \oplus \mathbb{T}^\perp = \mathbb{V}$

Folglich gibt es einen linearen Operator $\mathbf{P}_\mathbb{T}$, der jedem $\mathbf{x} \in \mathbb{V}$ seinen "Anteil" (siehe **2.1.2**) in \mathbb{T} zuordnet. Hier sind wichtige Eigenschaften der Orthogonalprojektion $\mathbf{P}_\mathbb{T}$:

- **P1:** $\mathbf{P}_\mathbb{T}^2 = \mathbf{P}_\mathbb{T}$
- **P2:** $im(\mathbf{P}_\mathbb{T}) = ker(\mathbf{E} - \mathbf{P}_\mathbb{T}) = \mathbb{T}$, $ker(\mathbf{P}_\mathbb{T}) = im(\mathbf{E} - \mathbf{P}_\mathbb{T}) = \mathbb{T}^\perp$
- **P3:** $\langle \mathbf{P}_\mathbb{T}\mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | \mathbf{P}_\mathbb{T}\mathbf{y} \rangle$

Die Eigenschaft **P3** heisst, dass $\mathbf{P}_\mathbb{T}$ selbstadjungiert ist (siehe **2.2.5**).

Gegeben ein Teilraum $\mathbb{T} \in \mathbb{V}$, wie kann man eine explizite Form für $\mathbf{P}_\mathbb{T}$ finden? Eine Antwort ist: wir wählen eine Orthonormalbasis \mathbf{m}_i von \mathbb{T} . Dann ist $\mathbf{P}_\mathbb{T}\mathbf{x} = \sum_i c_i \mathbf{m}_i$ gegeben durch die Bedingung $\langle \mathbf{m}_j | \mathbf{x} - \mathbf{P}_\mathbb{T}\mathbf{x} \rangle = 0$ für alle j , also $c_j = \langle \mathbf{m}_j | \mathbf{x} \rangle$.

Betrachten wir nun das folgende Extremalproblem: Gegeben ein Teilraum $\mathbb{T} \subset \mathbb{V}$ und $\mathbf{x} \in \mathbb{V}$. Wir suchen $\mathbf{y} \in \mathbb{T}$, sodass $\|\mathbf{x} - \mathbf{y}\| = \min!$. Das folgende Resultat zeigt, dass dieses Problem eine eindeutige Lösung hat, die durch $\mathbf{y} = \mathbf{P}_\mathbb{T}\mathbf{x}$ gegeben ist.

Satz: Sei \mathbb{T} ein Teilraum von \mathbb{V} und $\mathbf{y} \in \mathbb{T}$ beliebig. Dann gilt

$$\|\mathbf{x} - \mathbf{P}_{\mathbb{T}}\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| \quad (2.48)$$

mit Gleichheit genau dann, wenn $\mathbf{y} = \mathbf{P}_{\mathbb{T}}\mathbf{x}$.

Beweis: Es ist

$$\|\mathbf{x} - \mathbf{P}_{\mathbb{T}}\mathbf{x}\|^2 \leq \|\mathbf{x} - \mathbf{P}_{\mathbb{T}}\mathbf{x}\|^2 + \|\mathbf{P}_{\mathbb{T}}\mathbf{x} - \mathbf{y}\|^2 \quad (2.49)$$

Die Vektoren $\mathbf{x} - \mathbf{P}_{\mathbb{T}}\mathbf{x}$ und $\mathbf{P}_{\mathbb{T}}\mathbf{x} - \mathbf{y}$ sind orthogonal. Die rechte Seite von (2.49) ist daher wegen Pythagoras gleich $\|\mathbf{x} - \mathbf{P}_{\mathbb{T}}\mathbf{x} + \mathbf{P}_{\mathbb{T}}\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$, w.z.b.w.

Dieser Satz kann folgendermassen benutzt werden: Wir wollen \mathbf{x} durch Elemente in \mathbb{T} approximieren. Eine "gute" Näherung \mathbf{u} sei eine solche, für die der Abstand zu \mathbf{x} minimal ist. Sie ist durch $\mathbf{u} = \mathbf{P}_{\mathbb{T}}\mathbf{x}$ gegeben.

Beispiel: Sei $\mathbb{V} = C[-\pi, \pi]$, der Vektorraum der reellwertigen stetigen Funktionen im Intervall $[-\pi, \pi]$, \mathbf{x} entspreche der Funktion $\sin t$ und \mathbb{T} werde aufgespannt durch die Funktionen $\{t, t^3, t^5\}$. Das Skalarprodukt zwischen Elementen $\mathbf{f} : t \in [-\pi, \pi] \mapsto f(t) \in \mathbb{R}$ und $\mathbf{g} : t \in [-\pi, \pi] \mapsto g(t) \in \mathbb{R}$ sei gegeben durch

$$\langle \mathbf{f} | \mathbf{g} \rangle = \int_{-\pi}^{\pi} f(t)g(t)dt \quad (2.50)$$

Die im obigen Sinn beste Approximation ("Approximation im Mittel") von $\sin t$ ist also eine Funktion $f(t)$ der Form $f(t) = \alpha_1 t + \alpha_2 t^3 + \alpha_3 t^5$, wo die Zahlen $\alpha_1, \alpha_2, \alpha_3$ so gewählt sind, dass

$$\int_{-\pi}^{\pi} |\sin t - f(t)|^2 dt = \min! \quad (2.51)$$

Indem man zuerst das Gram-Schmidt-Verfahren auf die Elemente $\{t, t^3, t^5\}$ und dann (2.43, 2.44) anwendet, erhält man für $\mathbf{f} = \mathbf{P}_{\mathbb{T}}\mathbf{x}$ - unter Benutzung einer Dezimalnäherung für π - angenähert (Mathematica?)

$$f(t) = 0.987862t - 0.155271t^3 + 0.00564312t^5 \quad (2.52)$$

Die - punktweise - Übereinstimmung mit $\sin t$ im Intervall $[-\pi, \pi]$ ist exzellent. Alternativ könnte man die partiellen Ableitungen des Ausdrucks (2.51) nach den α_i gleich Null setzen und das entstehende Gleichungssystem für $(\alpha_1, \alpha_2, \alpha_3)$ lösen. Zum Vergleich kann man die Taylorentwicklung 5-ter Ordnung von $\sin t$ um $t = 0$ betrachten: $\sin t = t - \frac{1}{6}t^3 + \frac{1}{120}t^5 + O(t^7)$. Diese ist gut in der Nähe des Ursprungs. Es stellt sich heraus, sie ist gut etwa für $t \in (-2, 2)$.

Bemerkung: Dieses Beispiel geht genau genommen über den bisherigen Rahmen hinaus. Die Funktion $\sin t$ gehört nicht dem Raum \mathbb{V} an, sondern z.B. dem (unendlich-dimensionalen) Raum der stetigen Funktionen in $[-\pi, \pi]$.

2.2.5 Lineare Funktionale und Adjungierte

Ein lineares Funktional auf \mathbb{V} ist eine lineare Abbildung $\mathbf{l} : \mathbb{V} \rightarrow \mathbb{R}(\mathbb{C})$. Zum Beispiel ist für irgendein $\mathbf{u} \in \mathbb{V}$ die Zuordnung

$$\mathbf{l}_{\mathbf{u}}(\mathbf{x}) = \langle \mathbf{u} | \mathbf{x} \rangle \quad (2.53)$$

ein lineares Funktional. Wir formulieren ohne Beweis den

Satz: Jedes lineare Funktional \mathbf{l} ist von der Form (2.53) für einen eindeutigen Vektor \mathbf{u} .

Sei \mathbf{A} ein Operator und $\mathbf{u} \in \mathbb{V}$. Wir betrachten das lineare Funktional $\mathbf{l}_{\mathbf{y}'}$, das gegeben ist durch

$$\mathbf{l}_{\mathbf{y}'}(\mathbf{x}) = \langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle \quad (2.54)$$

Dann gibt es wegen des vorangegangenen Satzes einen Vektor $\mathbf{u} \in \mathbb{V}$, sodass

$$\langle \mathbf{u} | \mathbf{x} \rangle = \langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle \quad (2.55)$$

Da dieses \mathbf{u} linear von \mathbf{y} abhängt, gibt es einen Operator, der \mathbf{y} in \mathbf{u} abbildet. Dieser erhält die Bezeichnung \mathbf{A}^\dagger und heisst Adjungierte von \mathbf{A} :

$$\langle \mathbf{A}^\dagger \mathbf{y} | \mathbf{x} \rangle = \langle \mathbf{y} | \mathbf{A}\mathbf{x} \rangle \quad (2.56)$$

Im reellen Fall schreiben wir zuweilen statt \mathbf{A}^\dagger das Symbol \mathbf{A}^T . Wenn nicht eigens erwähnt, gelten die folgenden Aussagen sowohl im reellen als im komplexen Fall. Hier sind einige Eigenschaften:

- **Ad1:** $(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger$
- **Ad2:** $(\alpha \mathbf{A})^\dagger = \bar{\alpha} \mathbf{A}^\dagger$
- **Ad3:** $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$
- **Ad4:** $(\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$

Abbildungen \mathbf{A} , sodass $\mathbf{A}^\dagger = \mathbf{A}$ heissen selbstadjungiert. Zum Beispiel ist $\mathbf{E}^\dagger = \mathbf{E}$. Generell gilt:

Satz: Ein Operator \mathbf{A} ist selbstadjungiert genau dann, wenn ihre Matrixdarstellung a_{ij} in einer ON-Basis bezüglich $\langle | \rangle$ die Eigenschaft hat

$$a_{ij} = \overline{a_{ji}} \quad (2.57)$$

Abbildungen \mathbf{U} , die das Skalarprodukt invariant lassen, heissen unitär (im reellen Fall: isometrisch). Die definierende Eigenschaft ist also

$$\langle \mathbf{U}\mathbf{x} | \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle \quad (2.58)$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. Folglich ist der Vektor $(\mathbf{U}^\dagger \mathbf{U} - \mathbf{E})\mathbf{x}$ auf alle Vektoren orthogonal, daher gleich Null (wegen **S4'**). Mithin gilt

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{E} = \mathbf{U} \mathbf{U}^\dagger \quad (2.59)$$

Es ist nicht sehr schwer, sich zu überlegen, dass ein Operator genau dann unitär ist, wenn er eine Orthonormalbasis wieder in eine solche überführt.

Unitäre (orthogonale) Operatoren bilden kontinuierliche Familien. Sei z.B. $\mathbf{U}(\epsilon)$ eine 1-parametrische Familie unitärer Operatoren mit $\mathbf{U}(0) = \mathbf{E}$. Dann folgt aus (2.59) für $\mathbf{I} = \mathbf{U}'(0)$, dass

$$\mathbf{I}^\dagger + \mathbf{I} = 0 \quad (2.60)$$

also \mathbf{I} anti-selbstadjungiert ist. Die Gesamtheit aller derartigen "infinitesimalen Operatoren" bildet im reellen Fall einen reellen Vektorraum der Dimension $\frac{n(n-1)}{2}$, im komplexen Fall einen reellen Vektorraum der Dimension n^2 . Wichtig sind weiters die unitären Operatoren mit Determinante 1 ("spezielle unitäre Operatoren"). Hier kommt bei den infinitesimalen Operatoren noch die Bedingung verschwindender Spur hinzu. Im reellen Fall schränkt das die Zahl der kontinuierlichen Parameter nicht ein, im komplexen Fall wird die Zahl der Parameter um 1 reduziert, sodass die infinitesimalen Operatoren einen reellen Vektorraum der Dimension $n^2 - 1$ aufspannen.

2.2.6 Spezialfälle

Orthogonaler \mathbb{R}^2

Physikalisch kann man sich diesen Raum als Konfigurationsraum eines (nichtrelativistischen) Punktteilchens in der Ebene vorstellen. In einer Orthogonalbasis $\{\mathbf{e}_1, \mathbf{e}_2\}$ ist $\langle \mathbf{x} | \mathbf{y} \rangle = x_1 y_1 + x_2 y_2$. Eine zuweilen nützliche Relation ist

$$\epsilon_{ij} \epsilon_{kl} = \delta_{ki} \delta_{jl} - \delta_{kj} \delta_{il} \quad (2.61)$$

wobei ϵ_{ij} der Spezialfall $n = 2$ des total antisymmetrischen Symbols in (2.14) ist. Orthogonale Operatoren, also solche, die

$$\mathbf{O}^T \mathbf{O} = \mathbf{E} \quad (2.62)$$

erfüllen, haben Determinante ± 1 . Jene mit Determinante 1 bilden die Drehgruppe $\mathbf{SO}(2)$, gegeben durch die 1-parametrische Familie

$$O_{ij}(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

Eine kurze Rechnung zeigt

$$\cos \angle(\mathbf{O}\mathbf{x}, \mathbf{x}) = \cos \phi \quad (2.63)$$

Orthogonaler \mathbb{R}^3

Die ist z.B. der Konfigurationsraum eines Punktteilchens im Raum. Die allgemeinste Formel, die das Skalarprodukt δ_{ij} mit dem Epsilon-Symbol verbindet, lautet hier

$$\epsilon_{ijk}\epsilon_{lmn} = \delta_{il}(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + \delta_{in}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) + \delta_{im}(\delta_{jn}\delta_{kl} - \delta_{jl}\delta_{kn}) \quad (2.64)$$

Die Standard-Identitäten für das Vektorprodukt $\mathbf{x} \times \mathbf{y}$, gegeben durch

$$(\mathbf{x} \times \mathbf{y})_i = \sum_{j,k} \epsilon_{ijk} x_j y_k, \quad (2.65)$$

kann man - z.B. mittels Spurbildung - aus (2.64) herleiten.

Die orthogonalen Operatoren mit positiver Determinante sind die Drehgruppe $\mathbf{SO}(3)$. Ihre Elemente sind durch 3 Parameter charakterisiert, als die man sich z.B. die sog. Euler'schen Winkel denken kann.

Unitärer \mathbb{C}^2

Wichtigste physikalische Realisierung ist der Zustandsraum eines Quantenteilchens mit Spin $\frac{1}{2}$, wobei die Translationsfreiheitsgrade unterdrückt sind. Die speziellen unitären Operatoren bilden die Gruppe $\mathbf{SU}(2)$. Eine mögliche Basis für die infinitesimalen Operatoren \mathbf{I}_j , $j = 1, 2, 3$ kann man folgendermassen wählen: man schreibt $\mathbf{I}_j = i\boldsymbol{\sigma}_j$, wobei die $\boldsymbol{\sigma}$'s selbstadjungiert und spurfrei sind, z.B.

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Diese heissen Pauli'sche Spinmatrizen. Sei nun $\mathbf{U} \in \mathbf{SU}(2)$. Dann sind die drei Matrizen $\mathbf{U}\boldsymbol{\sigma}_i\mathbf{U}^\dagger$ wieder selbstadjungiert und spurfrei, denn

$$(\mathbf{U}\boldsymbol{\sigma}_i\mathbf{U}^\dagger)^\dagger = (\mathbf{U}^\dagger)^\dagger \boldsymbol{\sigma}_i^\dagger \mathbf{U}^\dagger = \mathbf{U}\boldsymbol{\sigma}_i\mathbf{U}^\dagger \quad (2.66)$$

und

$$\text{tr}(\mathbf{U}\boldsymbol{\sigma}_i\mathbf{U}^\dagger) = \text{tr}(\mathbf{U}^\dagger\mathbf{U}\boldsymbol{\sigma}_i) = \text{tr}(\boldsymbol{\sigma}_i) = 0 \quad (2.67)$$

Folglich gibt es reelle Zahlen R_{ij} , sodass

$$\mathbf{U}\boldsymbol{\sigma}_i\mathbf{U}^\dagger = \sum_j \boldsymbol{\sigma}_j R_{ji} \quad (2.68)$$

Sei nun $\mathbf{x} \in \mathbb{R}^3$. Dann ist

$$-\det \left(\sum_i \boldsymbol{\sigma}_i x_i \right) = \|\mathbf{x}\|^2, \quad (2.69)$$

wobei auf der rechten Seite die Norm im \mathbf{R}^3 steht, und daher

$$\det \left(\mathbf{U} \left(\sum_i \sigma_i x_i \right) \mathbf{U}^\dagger \right) = \det(\mathbf{U}) \det(\mathbf{U}^\dagger) \sum_i \det(\sigma_i x_i) = -\|\mathbf{x}\|^2 \quad (2.70)$$

Unter Benutzung von (2.69) sehen wir nun, dass der Vektor $\mathbf{x}' = \mathbf{R}\mathbf{x}$, gegeben durch $x'_i = \sum_j R_{ij}x_j$ die Relation

$$\|\mathbf{x}'\| = \|\mathbf{x}\| \quad (2.71)$$

erfüllt. Wir stellen ohne Beweis fest, dass wegen $\det(\mathbf{U}) = 1$ auch $\det(\mathbf{R}) = 1$ ist. Es gehört also zu jedem Element in $\mathbf{SU}(2)$ genau ein Element in $\mathbf{SO}(3)$.

2.2.7 Spektraltheorie linearer Operatoren

Hier kehren wir zur Frage zurück, ob es zu einer linearen Abbildung eine Basis gibt, die aus Eigenvektoren dieses Operators besteht. Nun haben, wie wir gesehen haben, auch reelle Abbildungen im allgemeinen komplexe Eigenwerte. Es ist daher wichtig, den reellen Fall (orthogonaler Vektorraum) auf den komplexen Fall (unitärer Vektorraum) zurückführen zu können, wie wir im nächsten Absatz erklären.

Wir müssen das orthogonale Skalarprodukt zu einem unitären "erweitern". Wir schreiben $\langle \mathbf{x} | \mathbf{y} \rangle$ in einer ON-Basis, es gilt also $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_i x_i y_i$. Den komplexifizierten Vektorraum denken wir uns nun als den Raum der Vektoren, die bez. dieser Basis komplexe Komponenten haben, und definieren als unitäres Skalarprodukt in diesem Raum naheliegenderweise $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_i \bar{x}_i y_i$. Die Abbildung \mathbf{A} wird wie gehabt zu einer reellen Abbildung im komplexifizierten Vektorraum.

Im Rahmen eines Vektorraums mit Skalarprodukt ist die natürliche Version der Eingangs-Frage nach einer Basis aus Eigenvektoren die, ob es eine Orthonormalbasis gibt, die aus Eigenvektoren besteht. Aus Beispiel 1 in (2.1.4) folgt bereits, dass das nicht für jeden Operator möglich sein kann. Im unitären Vektorraum - auf den wir uns im Augenblick konzentrieren - kann man aber schnell eine notwendige Bedingung angeben, und zwar so: In einer Basis von Eigenvektoren ist die zur linearen Abbildung \mathbf{A} gehörige Matrix a_{ij} diagonal. Wenn diese Basis eine ON-Basis ist, ist die Matrix der adjungierten Abbildung \mathbf{A}^\dagger gegeben durch \bar{a}_{ji} . Diese ist dann ebenfalls diagonal und kommutiert daher - im Sinn des Matrix-Produkts - mit der Matrix a_{ij} . Also muss gelten, dass

$$[\mathbf{A}, \mathbf{A}^\dagger] = 0 \quad (2.72)$$

Derartige Operatoren heißen normal. Sie enthalten selbstadjungierte und unitäre Operatoren als Spezialfall. Es stellt sich heraus - wir werden das nicht beweisen - dass die Bedingung (2.72) auch hinreichend für die Diagonalisierbarkeit ist. Eine

andere notwendige Bedingung für Diagonalisierbarkeit ist diese: Wenn \mathbf{x} ein Eigenvektor von \mathbf{A} zum Eigenwert λ ist, so ist $\bar{\mathbf{x}}$ Eigenvektor von \mathbf{A}^\dagger zum Eigenwert $\bar{\lambda}$. Auch diese Bedingung stellt sich als hinreichend heraus.

Wir betrachten nun selbstadjungierte Operatoren. Die folgenden Aussagen sind fundamental. Ihre Bedeutung in der Physik kann gar nicht hoch genug eingeschätzt werden.

Satz: Sei $\mathbf{A} = \mathbf{A}^\dagger$. Dann sind alle Eigenwerte reell.

Beweis: Einsetzen von $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ in $\langle \mathbf{A}\mathbf{x} | \mathbf{x} \rangle = \langle \mathbf{x} | \mathbf{A}\mathbf{x} \rangle$ ergibt

$$\bar{\lambda} \|\mathbf{x}\|^2 = \lambda \|\mathbf{x}\|^2, \quad (2.73)$$

Bemerkung: Die Eigenwerte λ eines unitären Operators erfüllen $|\lambda| = 1$, also $\lambda = e^{i\omega}$ für $\omega \in \mathbb{R}$.

Weiters gilt der

Satz: Sei $\mathbf{A} = \mathbf{A}^\dagger$. Dann sind Eigenvektoren zu verschiedenen Eigenwerten zueinander orthogonal.

Beweis: Wir haben $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ und $\mathbf{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_2$. Dann gilt

$$\lambda_1 \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle = \langle \mathbf{A}\mathbf{x}_1 | \mathbf{x}_2 \rangle = \langle \mathbf{x}_1 | \mathbf{A}\mathbf{x}_2 \rangle = \langle \mathbf{x}_1 | \lambda_2 \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle, \quad (2.74)$$

also sind die beiden Eigenvektoren orthogonal, sofern $\lambda_1 \neq \lambda_2$, w.z.z.w.

Beispiel: Sei \mathbb{V} der Vektorraum unendlich oft differenzierbarer Funktionen $f : \mathbb{R}^1 \rightarrow \mathbb{C}$ mit der zusätzlichen Eigenschaft, dass $f(x) = f(x + 2\pi)$ ist und dem Skalarprodukt $\langle f | g \rangle = \int_0^{2\pi} \overline{f(x)} g(x) dx$. Dann ist der lineare Operator \mathbf{A} gegeben durch $(\mathbf{A}f)(x) = \frac{1}{i} f'(x)$ selbstadjungiert. Beweis: partielle Integration. Also hat die Eigenwertgleichung $f' = i\lambda f$ nur reelle Lösungen λ . Dies ist zunächst erstaunlich, da die Gleichung $f' = i\lambda f$ für beliebiges $\lambda \in \mathbb{C}$ durch $f(x) \sim e^{i\lambda x}$ gelöst wird. Diese Funktion ist aber nur dann periodisch, wenn ($n \in \mathbb{Z}$) $\lambda = \lambda_n = n$ ist. Der letzte Satz und der jetzt folgende sind allgemeiner für normale Operatoren richtig.

Satz: Sei $\mathbf{A} = \mathbf{A}^\dagger$ und λ ein Eigenwert von \mathbf{A} . Dann ist $(\mathbb{T}_\lambda)^\perp$ unter \mathbf{A} invariant.

Beweis: leicht.

Für den berühmten Spektralsatz fehlt noch die folgende Aussage.

Satz: Sei $\mathbf{A} = \mathbf{A}^\dagger$ und λ_1 ein Eigenwert von \mathbf{A} . Dann ist die Dimension des Teilraums \mathbb{T}_{λ_1} ("geometrische Vielfachheit von λ_1 ") gleich der Ordnung von λ_1 als Nullstelle des charakteristischen Polynoms ("algebraische Vielfachheit").

Beweis: Wegen des Satzes in (2.2.4) können wir eine Orthonormalbasis wählen, die aus Elementen von \mathbb{T}_{λ_1} und solchen in $(\mathbb{T}_{\lambda_1})^\perp$ besteht. In einer solchen Basis hat die zugehörige Matrix Blockgestalt, wobei der zu \mathbb{T}_{λ_1} gehörige Block das λ_1 -fache der Einheitsmatrix ist. Die zur Abbildung $\mathbf{A} - \lambda\mathbf{E}$ gehörige Matrix hat ebenfalls Blockgestalt. Das charakteristische Polynom lautet daher (siehe D5)

$$\det(\mathbf{A} - \lambda\mathbf{E}) = (\lambda_1 - \lambda)^{\dim \mathbb{T}_{\lambda_1}} Q(\lambda), \quad (2.75)$$

wobei λ_1 keine Nullstelle des Polynoms $Q(\lambda)$ ist, w.z.b.w.

Nun ist es leicht, den Spektralsatz zu beweisen.

Satz: Sei \mathbf{A} selbstadjungiert. Dann gibt es eine unitäre Matrix \mathbf{T} , sodass

$\mathbf{A}' = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} = \mathbf{T}\mathbf{A}\mathbf{T}^\dagger$ diagonal ist.

Hier ist eine äquivalente Formulierung:

Satz: Jede selbstadjungierte Abbildung hat die Form

$$\mathbf{A} = \sum_{\alpha} \lambda_{\alpha} \mathbf{P}_{\lambda_{\alpha}}, \quad (2.76)$$

wobei λ_{α} die verschiedenen (reellen) Eigenwerte von \mathbf{A} sind und $\mathbf{P}_{\lambda_{\alpha}}$ die zugehörigen Orthogonalprojektionen. Diese haben die Eigenschaften

$$\sum_{\alpha} \mathbf{P}_{\lambda_{\alpha}} = \mathbf{E}, \quad \mathbf{P}_{\lambda_{\alpha}} \mathbf{P}_{\lambda_{\beta}} = 0 \text{ sofern } \lambda_{\alpha} \neq \lambda_{\beta} \quad (2.77)$$

Jede diagonalisierbare lineare Abbildung hat die Gestalt (2.76), nur dass dann die Eigenwerte nicht notwendig reell sind. Z.B. für unitäre Abbildungen hat die Spektralzerlegung die Form ($\omega_{\alpha} \in \mathbb{R}$)

$$\mathbf{U} = \sum_{\alpha} e^{i\omega_{\alpha}} \mathbf{P}_{e^{i\omega_{\alpha}}}, \quad (2.78)$$

Mithilfe der Darstellung (2.76) kann man Funktionen der entsprechenden Operatoren definieren und effizient berechnen ("Funktionalkalkül"). Sei $f: \mathbb{C} \rightarrow \mathbb{C}$. Dann definieren wir die lineare Abbildung $f(\mathbf{A})$ gemäss

$$f(\mathbf{A}) = \sum_{\alpha} f(\lambda_{\alpha}) \mathbf{P}_{\lambda_{\alpha}} \quad (2.79)$$

Zum Beispiel folgt aus (2.79), dass

$$e^{\mathbf{0}} = \mathbf{E} \quad (2.80)$$

Wenn f analytisch ist, kann man Funktionen f allgemeiner Operatoren über die Potenzreihenentwicklung definieren. Man kann zeigen, dass die entsprechende Matrix-Reihe konvergiert. Es ist dann nicht schwer zu sehen, dass Matrix-Funktionen invariant definiert sind, d.h.

$$f(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}) = \mathbf{T}f(\mathbf{A})\mathbf{T}^{-1} \quad (2.81)$$

Lemma: Sei $\mathbf{A}^\dagger = -\mathbf{A}$, d.h. anti-selbstadjungiert. Dann ist $e^{\mathbf{A}}$ unitär (bzw. orthogonal im reellen Fall).

Beispiel (wichtig!): \mathbb{V} ist der unitäre \mathbb{C}^2 und \mathbf{A} die lineare Abbildung, die in einer ON-Basis die Form

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

hat. Diese ist anti-selbstadjungiert, aber die komplexe Abbildung $i\mathbf{A}$ ist selbstadjungiert, ihre Eigenwerte λ_{\pm} sind ± 1 . Zugehörige normierte Eigenvektoren sind $\mathbf{e}_{\pm} = \frac{1}{\sqrt{2}}(\mathbf{e}_1 \pm i\mathbf{e}_2)$, oder $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(\mathbf{e}_+ + \mathbf{e}_-)$ und $\mathbf{e}_2 = \frac{1}{\sqrt{2}i}(\mathbf{e}_+ - \mathbf{e}_-)$. Die Abbildungsmatrix \mathbf{T} (siehe die Diskussion vor (2.2)) ist daher

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Die Matrix \mathbf{T} ist, wie es sein muss, unitär. Weiters ist

$$\mathbf{T}(i\mathbf{A})\mathbf{T}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

ebenfalls wie es sein muss. Nun zur Darstellung (2.76). Wenn $e_{\pm,i}$ die Komponenten der Vektoren \mathbf{e}_{\pm} in der Basis $\mathbf{e}_1, \mathbf{e}_2$ sind, gilt für die Komponenten der entsprechenden Projektoren $p_{+,ij} = e_{+,i}\overline{e_{+,j}}$ und $p_{-,ij} = e_{-,i}\overline{e_{-,j}}$. Die Rechnung ergibt, dass

$$\mathbf{P}_{\pm} = \frac{1}{2} \begin{pmatrix} 1 & \mp i \\ \pm i & 1 \end{pmatrix}$$

und $i\mathbf{A} = 1 \cdot \mathbf{P}_+ + (-1) \cdot \mathbf{P}_-$, wie es sein muss. Wir berechnen nun die Transformation $\exp(\mathbf{A}t)$, folgend (2.79). Sie lautet

$$\exp(\mathbf{A}t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

(Ende des Beispiels.)

In der Quantenmechanik von Bedeutung ist der folgende

Satz: Seien \mathbf{A} und \mathbf{B} selbstadjungierte Operatoren die kommutieren, also $[\mathbf{A}, \mathbf{B}] = 0$. Dann gibt es eine unitäre Transformation \mathbf{U} , die beide diagonalisiert.

Beweis: Sei λ ein Eigenwert von \mathbf{A} und \mathbf{x} ein zugehöriger Eigenvektor. Dann ist

$$(\mathbf{A} - \lambda\mathbf{E})\mathbf{B}\mathbf{x} = \mathbf{B}\mathbf{A}\mathbf{x} - \lambda\mathbf{B}\mathbf{x} = \mathbf{B}\lambda\mathbf{x} - \lambda\mathbf{B}\mathbf{x} = 0 \quad (2.82)$$

Folglich ist \mathbb{T}_{λ} unter \mathbf{B} invariant. Falls λ nicht-ausartet ist, muss $\mathbf{B}\mathbf{x}$ also proportional zu \mathbf{x} , daher Eigenvektor auch von \mathbf{B} sein. Andernfalls suchen wir eine Basis von \mathbb{T}_{λ} , die aus Eigenvektoren von $\mathbf{B}_{\lambda} = \mathbf{B}|_{\mathbb{T}_{\lambda}}$ besteht. Indem wir diese Prozedur für alle Eigenwerte von \mathbf{A} durchführen, erhalten wir eine Orthonormalbasis, die \mathbf{A} und \mathbf{B} gleichzeitig diagonalisiert.

2.2.8 Dirac-Notation

Das Arena (genauer: der Zustandsraum) der Quantenmechanik ist ein Vektorraum mit unitären Skalarprodukt. Ausser wenn es um sog. innere Freiheitsgrade wie Spin geht, hat dieser Vektorraum unendliche Dimension (sog. Hilbertraum), was

unseren gegenwärtigen Rahmen sprengt und was wir in der Folge aber ignorieren. Vektoren $\mathbf{x} \in \mathbb{V}$ werden nach Dirac als "ket" (2-ter Teil von "bracket") geschrieben, also $|\mathbf{x}\rangle$. Der zu einem Vektor \mathbf{y} gehörige "bra", geschrieben als $\langle \mathbf{y}|$, ist das lineare Funktional \mathbf{I}_y im Sinn von (2.53). Wegen der komplexen Linearität des Skalarprodukts im ersten Faktor gilt also für den einem ket $|\mathbf{x}\rangle$ zugeordneten bra $\langle \mathbf{x}|$ die Relation $\langle c\mathbf{x}| = \bar{c}\langle \mathbf{x}|$.

Das Symbol \mathbf{A} gegeben durch

$$\mathbf{A} = |\mathbf{y}\rangle\langle \mathbf{z}| \quad (2.83)$$

ist folglich eine lineare Abbildung, die jedem Vektor $\mathbf{x} \in \mathbb{V}$ den Vektor $\langle \mathbf{z}|\mathbf{x}\rangle \mathbf{y}$ zuordnet. Wenn $\mathbf{y} = \mathbf{z}$ und $\|\mathbf{y}\| = 1$, ist der Operator \mathbf{A} gerade die Orthogonalprojektion auf den durch den Vektor \mathbf{y} aufgespannten Teilraum. Die Spektralzerlegung (2.76) schreibt sich z.B. im Fall, dass alle Eigenwerte λ_α nicht-ausgeartet sind, als

$$\mathbf{A} = \sum_{\alpha} \lambda_{\alpha} |\mathbf{y}_{\alpha}\rangle\langle \mathbf{y}_{\alpha}|, \quad (2.84)$$

wobei \mathbf{y}_{α} die zugehörigen normierten Eigenvektoren sind. Weiters gilt für eine beliebige Orthonormalbasis \mathbf{e}_i die sog. "Vollständigkeitsrelation"

$$\mathbf{E} = \sum_i |\mathbf{e}_i\rangle\langle \mathbf{e}_i|, \quad (2.85)$$

wie durch Anwendung von (2.76) auf den selbstadjungierten Operator $\mathbf{A} = \mathbf{E}$ ersichtlich ist. Oft schreibt man auch kurz $|i\rangle$ statt $|\mathbf{e}_i\rangle$.

2.2.9 Quantenmechanik

Zustände ψ eines quantenmechanischen Systems sind - auf $\|\psi\| = 1$ normierte - Vektoren in einem unitären Vektorraum. Physikalischen Observablen entsprechen selbstadjungierte Operatoren, deren Eigenwerten (die ja reell sind) mögliche Messwerte. Der Erwartungswert der Observablen \mathbf{A} im Zustand ψ ist definiert als

$$\mathcal{E}_{\psi}(\mathbf{A}) = \langle \psi | \mathbf{A} \psi \rangle \quad (2.86)$$

Dieser ist reell, denn

$$\overline{\langle \psi | \mathbf{A} \psi \rangle} = \langle \mathbf{A} \psi | \psi \rangle = \langle \psi | \mathbf{A} \psi \rangle, \quad (2.87)$$

wobei wir im 1.ten Schritt Axiom **S3** und im 2.ten Schritt die Selbstadjungiertheit von \mathbf{A} benutzt haben. Wenn ψ Eigenvektor von \mathbf{A} zum Eigenwert λ ist, stimmt der Erwartungswert von \mathbf{A} mit λ überein. Die Dispersion (oder Unschärfe) der Observablen \mathbf{A} im Zustand ψ ist definiert als

$$\Delta_{\psi}(\mathbf{A}) = [\mathcal{E}_{\psi}(\mathbf{A}^2) - (\mathcal{E}_{\psi}(\mathbf{A}))^2]^{\frac{1}{2}} \quad (2.88)$$

Dass das Argument unter der Wurzel in (2.88) nicht-negativ ist, sieht man so. Es gilt

$$\mathcal{E}_\psi([\mathbf{A} - \mathcal{E}_\psi(\mathbf{A}) \mathbf{E}]^2) = \mathcal{E}_\psi(\mathbf{A}^2) - 2(\mathcal{E}_\psi(\mathbf{A}))^2 + (\mathcal{E}_\psi(\mathbf{A}))^2 = \mathcal{E}_\psi(\mathbf{A}^2) - (\mathcal{E}_\psi(\mathbf{A}))^2 \quad (2.89)$$

Also ist das Argument der Wurzel der Erwartungswert des Quadrats eines selbstadjungierten Operators \mathbf{B} und $\langle \psi | \mathbf{B}^2 \psi \rangle = \langle \mathbf{B} \psi | \mathbf{B} \psi \rangle \geq 0$. Weiters zeigt diese Rechnung, dass die Unschärfe = 0 genau dann ist, wenn ψ Eigenvektor von \mathbf{A} ist. Wichtig ist der

Satz: Wenn es sich bei der Observablen um eine Orthogonalprojektion \mathbf{P} handelt, gilt

$$0 \leq \mathcal{E}_\psi(\mathbf{P}) \leq 1 \quad (2.90)$$

Beweis: Wir haben

$$0 \leq (\mathcal{E}_\psi(\mathbf{P}))^2 \leq \langle \mathbf{P} \psi | \mathbf{P} \psi \rangle = \langle \psi | \mathbf{P}^2 \psi \rangle = \langle \psi | \mathbf{P} \psi \rangle = \mathcal{E}_\psi(\mathbf{P}), \quad (2.91)$$

wobei wir im 1.ten Schritt Cauchy-Schwarz, im 2.ten $\mathbf{P}^\dagger = \mathbf{P}$ und im letzten $\mathbf{P}^2 = \mathbf{P}$ verwendet haben. w.z.b.w.

Sei $\mathbb{T} = im(\mathbf{P})$, sodass $\mathbf{P} = \mathbf{P}_\mathbb{T}$. Dann ist nach den Interpretationsregeln der Quantenmechanik die Grösse $\mathcal{E}_\psi(\mathbf{P})$ die Übergangswahrscheinlichkeit zwischen ψ und dem Raum der Zustände in \mathbb{T} .

Wenn \mathbb{V} der Zustandsraum z.B. eines Teilchens für ein bestimmtes physikalisches System ist: was ist im Rahmen dieses Systems der Raum der Zustände zweier Teilchen? (Im Fall der klassischen Mechanik wäre dies das cartesische Produkt des Konfigurationsraums für ein Teilchen mit sich selbst.) In der Quantenmechanik ist dies das Tensorprodukt $\mathbb{V} \otimes \mathbb{V}$ ². Dafür muss dieser auch zu einem unitären Raum erklärt werden. Das geht natürlicherweise so, dass, wenn $\{\mathbf{e}_i\}$ eine ON-Basis von \mathbb{V} sind, die Tensorprodukte $\mathbf{e}_i \otimes \mathbf{e}_j$ eine ON-Basis von $\mathbb{V} \otimes \mathbb{V}$ bilden sollen, d.h.

$$\langle \mathbf{e}_i \otimes \mathbf{e}_j | \mathbf{e}_k \otimes \mathbf{e}_l \rangle = \delta_{ik} \delta_{jl} \quad (2.92)$$

Daraus folgt

$$\langle \mathbf{x} \otimes \mathbf{y} | \mathbf{u} \otimes \mathbf{v} \rangle = \langle \mathbf{x} | \mathbf{u} \rangle \langle \mathbf{y} | \mathbf{v} \rangle \quad (2.93)$$

und, für allgemeine Tensoren $\mathbf{r} = \sum_{i,j} r_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$, $\mathbf{s} = \sum_{i,j} s_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$, dass

$$\langle \mathbf{r} | \mathbf{s} \rangle = \sum_{i,j} \overline{r_{ij}} s_{ij} \quad (2.94)$$

Sei nun \mathbb{V} der unitäre \mathbb{C}^2 . Wir schreiben $|\mathbf{e}_1\rangle = |\uparrow\rangle$ ("spin up") und $|\mathbf{e}_2\rangle = |\downarrow\rangle$ ("spin down") und betrachten den Zweiteilchenzustand $|\psi\rangle \in \mathbb{V} \otimes \mathbb{V}$, gegeben durch

$$|\psi\rangle = \frac{1}{\sqrt{2}} [|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle] \quad (2.95)$$

²Genauer gilt dies nur in dem Fall, wo die Teilchen unterscheidbar sind.

Aus der Bemerkung über Zerlegbarkeit in Abschn.(2.1.5) folgt, dass ψ nicht zerlegbar ist. (Denn hier ist $r_{01} = -r_{10} = \frac{1}{\sqrt{2}}$, $r_{00} = r_{11} = 0$, also $\det(r_{ij}) = \frac{1}{2}$) Nun ist eine weitere Interpretationsregel der Quantenmechanik, dass, wenn eine Messung an einem Zustand $|\psi\rangle$ einen scharfen Wert für eine Observable erbringt, dieser Zustand als Resultat dieser Messung in den Zustand

$$|\bar{\psi}\rangle = \frac{1}{\|\mathbf{P}\psi\|} \mathbf{P}|\psi\rangle \quad (2.96)$$

übergeht, wobei \mathbf{P} die Orthogonalprojektion auf den Raum der zu diesem Messwert gehörigen Eigenvektoren ist. Dieser Vorgang heisst "Kollaps (oder Reduktion) des Zustandsvektors".

Wenn nun eine Messung ergibt: "Teilchen A hat Spin up, was auch immer Teilchen B tut", ist der zugehörige Projektor \mathbf{P} von der Form

$$\mathbf{P} = |\uparrow\rangle\langle\uparrow| \otimes [|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|], \quad (2.97)$$

und folglich ist

$$|\bar{\psi}\rangle = |\uparrow\rangle \otimes |\downarrow\rangle \quad (2.98)$$

Die Wahrscheinlichkeit für den Eintritt dieses Messergebnisses ist $\mathcal{E}_{\bar{\psi}}(\mathbf{P}) = \frac{1}{2}$. Mit derselben Wahrscheinlichkeit ergibt die Messung, dass Teilchen A Spin down hat. Der kollabierte Zustandsvektor ist dann

$$|\bar{\psi}\rangle = |\downarrow\rangle \otimes |\uparrow\rangle \quad (2.99)$$

Im ersten Fall ergibt eine weitere Messung an Teilchen B, dass dieses mit Wahrscheinlichkeit 1 Spin down hat. Im zweiten Fall ergibt eine weitere Messung an B, dass es mit Wahrscheinlichkeit 1 Spin up hat. Dies ist das Phänomen der Verschränkung (die von Einstein-Podolsky-Rosen sogenannte "spukhafte Fernwirkung").

2.2.10 Symmetrische Operatoren, kleine Schwingungen

Wir betrachten hier den Fall eines reellen Vektorraums \mathbb{V} mit positiv definitem Skalarprodukt $\langle | \rangle$. Die reellen, selbstadjungierten Abbildungen \mathbf{A} heissen auch symmetrische Operatoren. In einer ON-Basis für $\langle | \rangle$ entsprechen ihnen reelle symmetrische Matrizen, also $\mathbf{A}^T = \mathbf{A}$. Der Spektralsatz lautet hier:

Satz: Zu jeder symmetrischen Matrix \mathbf{A} gibt es eine orthogonale Matrix \mathbf{O} , also $\mathbf{O}^T = \mathbf{O}^{-1}$, sodass $\mathbf{O}\mathbf{A}\mathbf{O}^{-1}$ eine Diagonalmatrix ist.

Diese Situation tritt in der Punktmechanik in folgender Form auf: Man betrachtet ein konservatives System mit N Freiheitsgraden in der Nähe einer Gleichgewichtslage. Die Newton'schen Bewegungsgleichungen für die infinitesimalen Abweichungen $\mathbf{Q} \in \mathbb{R}^N$ von der Ruhelage lauten dann ($\dot{}$ ist Ableitung nach der Zeit)

$$\sum_{j=1}^N (k_{ij}\ddot{Q}_j + u_{ij}Q_j) = 0, \quad (2.100)$$

wobei k_{ij} ("kinetische Energie") und u_{ij} ("potentielle Energie") symmetrische Matrizen sind und k_{ij} positiv definit ist, also $\sum_{i,j} k_{ij} Q_i Q_j > 0$ für alle $\mathbf{0} \neq \mathbf{Q} \in \mathbb{R}^N$. Der Exponentialansatz $Q_i(t) = q_i e^{i\omega t}$ ergibt

$$\sum_j (-\omega^2 k_{ij} + u_{ij}) q_j = 0 \quad (2.101)$$

Wir machen die folgende Beobachtung: die lineare Abbildung mit der Matrix $a_{ij} = \sum_l k_{il}^{-1} u_{lj}$ ($\mathbf{A} = \mathbf{K}^{-1}\mathbf{U}$) ist selbstadjungiert bezüglich des durch $\langle \mathbf{q} | \mathbf{q}' \rangle_{\mathbf{K}} = \langle \mathbf{q} | \mathbf{K} \mathbf{q}' \rangle$ definierten Skalarprodukts. Denn wir haben $\langle \mathbf{q} | \mathbf{A} \mathbf{q}' \rangle_{\mathbf{K}} = \langle \mathbf{q} | \mathbf{K} \mathbf{A} \mathbf{q}' \rangle = \langle \mathbf{q} | \mathbf{K} \mathbf{K}^{-1} \mathbf{U} \mathbf{q}' \rangle = \langle \mathbf{q} | \mathbf{U} \mathbf{q}' \rangle = \langle \mathbf{U} \mathbf{q} | \mathbf{q}' \rangle = \langle \mathbf{q}' | \mathbf{U} \mathbf{q} \rangle = \langle \mathbf{q}' | \mathbf{A} \mathbf{q} \rangle_{\mathbf{K}} = \langle \mathbf{A} \mathbf{q} | \mathbf{q}' \rangle_{\mathbf{K}}$. Wenn wir weiters Glg.(2.101) mit \mathbf{K}^{-1} multiplizieren, erhalten wir gerade die Eigenwertgleichung für a_{ij} mit $\lambda = \omega^2$. Folglich sind alle ω^2 , für die Gl.(2.101) nicht-triviale Lösungen hat, reell. Sei \mathbf{O} die bezüglich $\langle | \rangle_{\mathbf{K}}$ orthogonale Matrix, die \mathbf{A} diagonalisiert. Dann führen die Komponenten von $\mathbf{Q}' = \mathbf{O} \mathbf{Q}$ (unabhängige) Bewegungen gemäss $\ddot{Q}'_i + \omega_i^2 Q'_i = 0$ durch, sind also "entkoppelt". Die Natur dieser Bewegungen hängt vom Vorzeichen von $\lambda_i = \omega_i^2$ ab.

1. $\omega_i^2 > 0$: harmonische Schwingungen, d.h. $Q'_i(t) = c_1 \sin \omega_i t + c_2 \cos \omega_i t$
2. $\omega_i^2 < 0$: exponentielles Verhalten, d.h. $Q'_i(t) = c_1 e^{\omega_i t} + c_2 e^{-\omega_i t}$
3. $\omega_i^2 = 0$: $Q'_i(t) = c_1 + c_2 t$

2.3 Anhang: Normalformen

In diesem Anhang behandeln wir Normalformen reeller Operatoren für $n = 2$. Die Natur einer reellen 2×2 -Matrix hängt im wesentlichen vom Vorzeichen der Diskriminante

$$\Delta = (\text{tr}(\mathbf{A}))^2 - 4 \det(\mathbf{A}) \quad (2.102)$$

ihres charakteristischen Polynoms ab.

Typ 1, $\Delta > 0$: Es gibt 2 voneinander verschiedene reelle Eigenwerte, zugehörige Eigenvektoren bilden daher eine Basis. In Termen dieser Basis schreibt sich die zugehörige Matrix als ($\lambda_1 \neq \lambda_2$)

$$\mathbf{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Typ 2a), $\Delta = 0$, $\lambda_1 = \lambda_2 = \lambda$ und $\mathbf{A} = \lambda \mathbf{E}$: Hier gilt (diesmal in jeder Basis)

$$\mathbf{A} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

Typ 2b), $\Delta = 0$, $\lambda_1 = \lambda_2 = \lambda$, $\mathbf{A} - \lambda\mathbf{E} \neq \mathbf{0}$: Um diesen Fall zu analysieren, betrachten wir den Teilraum $\text{im}(\mathbf{A} - \lambda\mathbf{E})$. Dieser ist wegen **G3** 1-dimensional und unter \mathbf{A} invariant, folglich ein Eigenraum von \mathbf{A} . Wegen $\Delta = 0$ muss also $\text{im}(\mathbf{A} - \lambda\mathbf{E}) = \text{ker}(\mathbf{A} - \lambda\mathbf{E})$ sein (und daher $(\mathbf{A} - \lambda\mathbf{E})^2 = \mathbf{0}$).

Wir wählen nun Vektoren $(\mathbf{e}_1, \mathbf{e}_2)$ von \mathbb{V} , wobei \mathbf{e}_2 ein beliebiger vom Nullvektor verschiedener Vektor ist, der nicht in $\text{ker}(\mathbf{A} - \lambda\mathbf{E})$ liegt und $\mathbf{e}_1 = (\mathbf{A} - \lambda\mathbf{E})\mathbf{e}_2$. Folglich ist $\mathbf{e}_1 \neq \mathbf{0}$ und liegt in $\text{ker}(\mathbf{A} - \lambda\mathbf{E})$. Die beiden bilden also eine Basis. In dieser nimmt die Matrix \mathbf{A} die Form an

$$\mathbf{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Typ 3, $\Delta < 0$: Hier gibt es einen komplexen Vektor \mathbf{e} mit

$$\mathbf{A}\mathbf{e} = (a - ib)\mathbf{e}, \quad (2.103)$$

wobei $b \neq 0$. Mit $\mathbf{e}_1 = \text{Re}(\mathbf{e})$ und $\mathbf{e}_2 = \text{Im}(\mathbf{e})$ folgt aus (2.103) und der Realität von \mathbf{A} , dass

$$\mathbf{A}\mathbf{e}_1 = a\mathbf{e}_1 + b\mathbf{e}_2, \quad \mathbf{A}\mathbf{e}_2 = -b\mathbf{e}_1 + a\mathbf{e}_2. \quad (2.104)$$

Folglich hat in der Basis $\mathbf{e}_1, \mathbf{e}_2$ die Matrix \mathbf{A} die Form

$$\mathbf{A} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Wir haben also bewiesen, dass

Satz: Jede reelle 2×2 -Matrix ist ähnlich zu einer der 3 obigen Typen.

Wir wollen, als Vorbereitung für die Systeme linearer Differentialgleichungen, die Normalformen dazu benutzen, um die Exponentialfunktion einer reellen 2×2 -Matrix zu berechnen.

Als Vorbereitung benötigen wir zunächst:

Lemma: Wenn $[\mathbf{A}, \mathbf{B}] = 0$, dann ist

$$e^{\mathbf{A}}e^{\mathbf{B}} = e^{(\mathbf{A}+\mathbf{B})} \quad (2.105)$$

Der Beweis folgt formal dem aus der Analysis bekannten für die numerische Exponentialfunktion mithilfe der Potenzreihendarstellung. Man muss sich aber über die Konvergenz den Kopf zerbrechen.

Bemerkung: Daraus folgt insbesondere, dass jede Abbildung von der Form $e^{\mathbf{A}}$ nicht-singulär ist, und zwar gilt (siehe (2.80)), dass $(e^{\mathbf{A}})^{-1} = e^{-\mathbf{A}}$.

Wir kehren zu unseren reellen 2×2 -Matrizen zurück: Für Typ 1 und Typ 2a ist offenbar

$$e^{\mathbf{A}t} = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}$$

Für Typ 2b schreiben wir $\mathbf{A} = (\mathbf{A} - \lambda\mathbf{E}) + \lambda\mathbf{E}$. Wegen der Nilpotenz von $\mathbf{A} - \lambda\mathbf{E}$ gilt

$$e^{(\mathbf{A}-\lambda\mathbf{E})t} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

und daher wegen (2.105)

$$e^{\mathbf{A}t} = \begin{pmatrix} e^{\lambda t} & t e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

Analog ergibt sich unter Verwendung des Beispiels in (2.2.7) für den Typ 3 die Formel

$$e^{\mathbf{A}t} = e^{at} \begin{pmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{pmatrix}$$

Um die Exponentialfunktion einer beliebigen reellen 2×2 - Matrix zu berechnen, kann man folgenderweise vorgehen: Man findet zuerst die Transformation \mathbf{T} auf die jeweilige Normalform \mathbf{A}' , also $\mathbf{A}' = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}$ und schreibt unter Benutzung von (2.79)

$$e^{\mathbf{A}} = e^{\mathbf{T}^{-1}\mathbf{A}'\mathbf{T}} = \mathbf{T}^{-1}e^{\mathbf{A}'}\mathbf{T}. \quad (2.106)$$

Schliesslich entnimmt man die Gestalt von $e^{\mathbf{A}'}$ den obigen Formeln.

Kapitel 3

ODE's

3.1 Lineare Systeme

3.1.1 Homogene autonome Systeme

Die einfachste derartige Differentialgleichung lautet

$$\dot{x} = \alpha x \tag{3.1}$$

Hier ist $x : t \in \mathbb{R} \mapsto x(t) \in \mathbb{R}$, $\dot{}$ bezeichnet Ableitung nach t und $\alpha \in \mathbb{R}$. Gesucht ist $x(t)$, das (3.1) löst und mit der Anfangsbedingung (AB), dass $x(0) = x_0$. Die Funktion

$$x(t) = e^{\alpha t} x_0 \tag{3.2}$$

erfüllt dies Bedingungen. Sie ist auch die einzige, denn sei $y(t)$ eine andere Lösung. Dann erfüllt $\bar{x}(t) = e^{-\alpha t} x(t)$ die Gleichung $\dot{\bar{x}} = 0$. Also erfüllt die allgemeine Lösung $x(t) = \text{const } e^{\alpha t}$, und die Konstante wird durch die AB festgelegt.

Die r.S. von (3.1) hängt nicht explizit von t ab. Das bewirkt, dass mit $x(t)$ auch $x(t - t_0)$ Lösung ist, und zwar zur AB $x(t_0) = x_0$.

Sei nun $\mathbf{x} \in \mathbb{R}^n$. Wir suchen Kurven $t \in \mathbb{R} \mapsto \mathbf{x}(t) \in \mathbb{R}^n$, sodass

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x} , \tag{3.3}$$

wobei \mathbf{A} ein (t - unabhängiger) linearer Operator im \mathbb{R}^n ist. Wie im Spezialfall (3.1) ($n = 1$) sehen wir, dass die allgemeine Lösung von Glg.(3.3) von der Form $e^{\mathbf{A}t} \mathbf{y}$ ist, wobei \mathbf{y} ein t - unabhängiger ("konstanter") Vektor in \mathbb{R}^n ist. Die AB fixiert dann $\mathbf{y} = \mathbf{x}_0$. Die Lösung des Anfangswertproblems für Gleichung (3.3) lautet daher

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0 \tag{3.4}$$

Man kann $e^{\mathbf{A}t} \mathbf{x}$ als eine 1-parametrische Familie von Transformationen ansehen (den sog. Fluss von (3.3)). Deren Bildmenge (der Orbit oder die Bahn durch \mathbf{x}) wird durch den Fluss in sich selbst abgebildet, denn:

$$e^{\mathbf{A}s} e^{\mathbf{A}t} \mathbf{x} = e^{\mathbf{A}(s+t)} \mathbf{x} \tag{3.5}$$

Die Relation zwischen \mathbf{y} und \mathbf{x} : " \mathbf{y} liegt auf dem Orbit durch \mathbf{x} " ist also eine Äquivalenzrelation. Folglich liegt jeder Punkt auf genau einem Orbit. Der Orbit durch $\mathbf{0}$ besteht nur aus $\mathbf{0}$.

Wir bemerken, dass man die Relation (3.5) auch so ausdrücken kann: Sei \mathbf{g}_t die (lineare) Abbildung, die dem Anfangswert $\mathbf{x}(0)$ den Bahnpunkt $\mathbf{x}(t)$ zuordnet. Dann gilt das Gruppengesetz

$$\mathbf{g}_s \circ \mathbf{g}_t = \mathbf{g}_{s+t} \quad (3.6)$$

Fall $n = 1$ und $\alpha \neq 0$: Hier gibt es 3 Orbits: den Ursprung, den "rechten" und den "linken". Z.B. für $\alpha > 0$ laufen die Punkte auf dem rechten Orbit für $t \rightarrow \infty$ nach $+\infty$ und für $t \rightarrow -\infty$ in den Ursprung. Auf dem linken Orbit laufen die Punkte für $t \rightarrow -\infty$ in den Ursprung und für $t \rightarrow \infty$ nach $-\infty$. Der Ursprung ist also eine "Quelle". (Für $\alpha < 0$ ist der Ursprung eine Senke.) Wenn $\alpha = 0$, bewegt sich nichts, also jeder Punkt ist ein Orbit.

Fall $n = 2$ und \mathbf{A} nicht-singulär: Nach den Resultaten von Abschnitt 2.3 gibt es, nach Übergang $\mathbf{x} \mapsto \mathbf{T}\mathbf{x}$, folgende Möglichkeiten:

$$\begin{array}{ll} (I) \quad \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad (\lambda \neq 0) & (II) \quad \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad (\mu < \lambda < 0) \\ (III) \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad (\lambda \neq 0) & (IV) \quad \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad (\lambda < 0 < \mu) \\ (V) \quad \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \quad (\alpha, \beta \neq 0) & (VI) \quad \begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix} \quad (\beta \neq 0) \end{array}$$

Wir nehmen $\lambda < 0$ an. Im Fall (I) ist der Ursprung ein "eigentlicher Knotenpunkt" ("proper node"). Es kommen alle Bahnen für $t \rightarrow -\infty$ aus dem Unendlichen und münden für $t \rightarrow \infty$ in den Ursprung, jede unter einem anderen Winkel. Im Fall (II) ("uneigentlicher Knoten") münden alle Bahnen in einer der zwei entgegengesetzten Richtungen in den Ursprung, die zum Eigenraum \mathbb{T}_λ gehört - mit Ausnahme jener beiden Bahnen, die zum Eigenraum \mathbb{T}_μ gehören. Dies sieht man so: der Fluss ist gegeben durch $x_1(t) = e^{\lambda t}x_1(0)$, $x_2(t) = e^{\mu t}x_2(0)$. Wenn $x_1(0) = 0$ ist, läuft die Bahn für $t \rightarrow \infty$ - je nach dem Vorzeichen von $x_2(0)$ - entlang der positiven oder negativen x_2 - Achse in den Ursprung. Wenn $x_2(0) = 0$ ist läuft die Bahn je nach dem Vorzeichen von $x_1(0)$ entlang der positiven oder negativen x_1 - Achse in den Ursprung. Wenn $x_1(0) > 0$, $x_2(0) > 0$ ist, verbleibt die Bahn im ersten Quadranten, und zwar so, dass $\frac{x_2(t)}{x_1(t)}$ nach 0 geht, also münden diese Bahnen tangential zur x_1 - Achse in den Ursprung. Analoges gilt für Bahnen in den anderen Quadranten. Im Fall (III) für $\lambda < 0$ münden alle Bahnen in einem der zwei entgegengesetzten Richtungen in den Ursprung, die zum Eigenraum \mathbb{T}_λ gehören. Denn hier ist $x_1(t) = e^{\lambda t}(x_1(0) + tx_2(0))$, $x_2(t) = e^{\lambda t}x_2(0)$. Daher sind wieder die positive und negative x_1 - Achse Bahnen, die für $t \rightarrow \infty$ in den Ursprung münden. Wenn die Bahn in der oberen Halbebene startet, dann

bleibt sie in der oberen Halbebene. Wenn t hinreichend gross ist, ist $x_1(t)$ positiv, also enden diese Bahnen im ersten Quadranten, genauer: sie erreichen diesen Quadranten (von links) und verbleiben dann in diesem. Zur Zeit $T = -\frac{1}{\lambda}$ nach Erreichen dieses Quadranten erreicht x_1 ein Maximum und geht dann nach 0. Weiters geht $\frac{x_2(t)}{x_1(t)}$ nach 0. Mithin münden diese Bahnen tangential zur positiven x_1 -Achse in den Ursprung. In analoger Weise enden Bahnen, die in der unteren Halbebene starten, im dritten Quadranten und münden entlang der negativen x_1 -Achse in den Ursprung. Im Fall (IV) ist der Ursprung ein Sattelpunkt: die positive (negative) x_2 -Achse ist ein Orbit, der für $t \rightarrow \infty$ ins Unendliche läuft und für $t \rightarrow -\infty$ in den Ursprung. Bei der x_1 -Achse ist es umgekehrt. Alle anderen Bahnen bleiben in ihrem jeweiligen Quadranten und erfüllen dort $x_2(t) = \left|\frac{x_1(t)}{x_1(0)}\right|^{\frac{1}{\lambda}} x_2(0)$ nebst $x_1(t) = e^{\lambda t} x_1(0)$. Daher geht $x_2(t)$ für $t \rightarrow \infty$ nach Unendlich und für $t \rightarrow -\infty$ nach Null. Im Fall (V) ist der Ursprung ein Spiralpunkt: Es gilt (mit $r^2 = x_1^2 + x_2^2$), dass $r(t) = e^{\alpha t} r(0)$. Desweiteren gilt für $\phi = \tan^{-1} \frac{x_2}{x_1}$, dass $\tan(\phi(t)) = \frac{\tan \beta t + \tan(\phi(0))}{1 - \tan \beta t \tan(\phi(0))}$ oder $\phi(t) = \beta t + \phi(0)$. Im Fall (VI) sind alle Bahnen ausser dem Ursprung geschlossen, die Bahnen in (V,VI) laufen für $\beta > 0$ gegen den Uhrzeiger. Siehe hand-out!

Wir betrachten als Anwendung den harmonischen Oszillator in einer Dimension:

$$\ddot{q} + \omega^2 q = 0, \quad q \in \mathbb{R}^1, \quad \omega \neq 0 \quad (3.7)$$

Die Substitution $p = \dot{q}$ liefert für den Vektor $\mathbf{x} = (q, p)$ ein System der Form $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, wobei

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}$$

Der Operator \mathbf{A} hat die Eigenwerte $\lambda_{\pm} = \pm i\omega$. Wir haben also Typ 3 mit $a = 0$. Die Eigenvektoren können als $\mathbf{e}_{\pm} = (1, \pm i\omega)$ gewählt werden. Für die Matrix \mathbf{S} , sodass $\mathbf{A}' = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ die Normalform

$$\mathbf{A}' = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$$

hat, können daher als Spalten die Vektoren $\mathbf{e} = \operatorname{Re}(\mathbf{e}_+)$ und $\mathbf{f} = \operatorname{Im}(\mathbf{e}_+)$ gewählt werden, also

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & \omega \end{pmatrix}$$

Daher ist

$$e^{\mathbf{A}t} = \mathbf{S} e^{\mathbf{A}'t} \mathbf{S}^{-1} = \mathbf{S} \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} \mathbf{S}^{-1} = \begin{pmatrix} \cos \omega t & \frac{1}{\omega} \sin \omega t \\ -\omega \sin \omega t & \cos \omega t \end{pmatrix} \quad (3.8)$$

Also ist die Lösung des Anfangswertproblems für die Gleichung (3.7) gegeben durch

$$q(t) = \cos \omega t q(0) + \frac{\sin \omega t}{\omega} \dot{q}(0) \quad (3.9)$$

3.1.2 Inhomogene Systeme

Diese sind von der Form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}(t), \quad (3.10)$$

wobei $\mathbf{b}(t)$ ein vorgegebener, zeitabhängiger Vektor in \mathbb{R}^n ist. Wir machen den Ansatz

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{y}(t) \quad (3.11)$$

Folglich ist $\dot{\mathbf{x}} - \mathbf{A}\mathbf{x} = e^{\mathbf{A}t}\dot{\mathbf{y}}$. Man erhält daher (spezielle) Lösungen von (3.10) durch

$$\mathbf{x}(t) = e^{\mathbf{A}t} \int_{t_0}^t e^{-\mathbf{A}t'} \mathbf{b}(t') dt' \quad (3.12)$$

Wie immer bei linearen Gleichungen kann die allgemeine Lösung als Linearkombination von allgemeiner Lösung des homogenen Problems und einer speziellen Lösung des inhomogenen Problems geschrieben werden. Bei gegebenen Anfangswerten für $t = 0$ erhalten wir

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-t')} \mathbf{b}(t') dt' \quad (3.13)$$

Die Natur des 2-ten Terms der r.S. in (3.13), z.B. dessen asymptotisches Verhalten für $t \rightarrow \infty$, hängt sowohl vom Typ von \mathbf{A} als auch den Eigenschaften von \mathbf{b} ab (z.B. Stichwort: "Resonanz").

Wenn wir im Fall des harmonischen Oszillators von vorhin die äussere Kraft $f(t)$ vorgeben, also

$$\ddot{q} + \omega^2 q = f \quad (3.14)$$

führt das zu einem $\mathbf{b}(t)$ der Form $b_i = (0, f)$, und aufgrund von (3.8) lautet die q -Komponente des inhomogenen Terms in (3.13)

$$\int_0^t \frac{\sin \omega(t-t')}{\omega} f(t') dt' \quad (3.15)$$

3.1.3 Lineare nicht-autonome Gleichungen

Lineare, nicht-autonome Systeme sind von der Form (siehe (3.2.6))

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t) \quad (3.16)$$

Hier ist nur der Fall $n = 1$ einfach, da dann die \mathbf{A} 's für verschiedenes t trivialerweise miteinander kommutieren. Wir haben dann

$$\dot{x} = a(t)x + b(t) \quad (3.17)$$

Eine Lösung für $b = 0$ lautet

$$x(t) = e^{\int_0^t a(t') dt'} \quad (3.18)$$

und die allgemeine homogene Lösung ergibt sich durch Multiplikation von (3.18) mit einer Konstanten, die gleich x_0 ist. Die volle Lösung ist

$$x(t) = e^{\int_0^t a(t') dt'} x_0 + \int_0^t F(t, t') b(t') dt', \quad (3.19)$$

wobei

$$F(t, t') = e^{\int_{t'}^t a(t'') dt''} \quad (3.20)$$

Angenommen, $b(t)$ geht hinreichend schnell für $t \rightarrow -\infty$ nach 0, z.B. ist $\equiv 0$ für $t < -T$. Dann ist eine andere spezielle Lösung die folgende

$$x_I(t) = \int_{-\infty}^t F(t, t') b(t') dt' \quad (3.21)$$

Sei $\Theta(t)$ die Funktion mit $\Theta(t) = 1$ für $t > 0$ und $\Theta(t) = 0$ für $t < 0$. Die Funktion

$$G(t, t') = \Theta(t - t') F(t, t') \quad (3.22)$$

heißt (retardierte) Green-Funktion der Glg.(3.17). Die Lösung x_I schreibt sich dann als

$$x_I(t) = \int_{-\infty}^{\infty} G(t, t') b(t') dt' \quad (3.23)$$

Da G nicht einmal differenzierbar ist, ist es schon gar nicht erlaubt, in (3.21) unter dem Integralzeichen zu differenzieren. Wenn wir dies formal trotzdem tun, wäre die Schlussfolgerung: Anwendung des linearen Operators $\frac{d}{dt} - a(t)$ auf $G(t, t')$ ergibt eine "Funktion" $\delta(t, t') = \delta(t - t')$ mit der Eigenschaft, dass für alle Funktionen $b(t)$ die Relation $\int_{-\infty}^{\infty} \delta(t - t') b(t') dt' = b(t)$ gilt. Eine integrable Funktion mit dieser Eigenschaft kann es nicht geben, denn angenommen, es gäbe z.B. $\delta(t)$, sodass $\int_{-\infty}^{\infty} \delta(t) b(t) dt = b(0)$. Dann wählen wir für $b(t)$ die Funktion

$$c_\epsilon(t) = e^{-\frac{\epsilon^2}{\epsilon^2 - t^2}} \text{ wenn } |t| < \epsilon, \text{ ansonsten } = 0$$

Dann müsste gelten, dass $\int_{-\infty}^{\infty} \delta(t) c_\epsilon(t) dt = e^{-1}$. Andererseits müsste, wenn $\delta(t)$ eine integrable Funktion ist, gelten, dass $\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta(t) c_\epsilon(t) dt = 0$ - ein Widerspruch. Im Rahmen der Distributionstheorie erhalten Objekte $\delta(t - t')$ einen Sinn (als lineares Funktional in einem geeigneten Funktionenraum). Diese können beliebig oft differenziert werden. Zum Beispiel ist die Ableitung der Funktion $\Theta(t)$, definiert durch $\Theta(t) = 1$ für $t > 0$ und $\Theta(t) = 0$ für $t < 0$, **im Distributionssinn**, gleich $\delta(t)$, denn:

$$\int \Theta'(t) b(t) dt := - \int \Theta(t) b'(t) dt = b(0) \quad (3.24)$$

Man kann sich das Funktional $\delta(t)$ mit $\int_{-\infty}^{\infty} \delta(t) b(t) dt = b(0)$ als Grenzwert einer Funktionenfolge f_ϵ mit $\epsilon > 0$ vorstellen mit der Eigenschaft, dass $\lim_{\epsilon \rightarrow 0} f_\epsilon(t) = 0$ für $t \neq 0$ und $\int_{-\infty}^{\infty} f_\epsilon(t) dt = 1$ für alle $\epsilon > 0$.

Beispiel: Sei $f_\epsilon(t) = \frac{1}{\pi} \frac{\epsilon}{t^2 + \epsilon^2}$. Wir haben $\int_{-\infty}^{\infty} f_\epsilon(t) dt = \frac{1}{\pi} [\arctan \frac{\beta}{\epsilon} - \arctan \frac{\alpha}{\epsilon}]$. Folglich ist $\int_{-\infty}^{\infty} f_\epsilon(t) dt = 1$ und $\int_{-\infty}^t f_\epsilon(t') dt' = \Theta(t)$. Mithin ist $\lim_{\epsilon \rightarrow 0} f_\epsilon(t) = \delta(t)$.

3.2 Nichtlineare Systeme

3.2.1 Autonome Systeme erster Ordnung

Sei $\mathbf{x} \in \mathbb{R}^n$ und $v_i(\mathbf{x})$ ($i = 1, \dots, n$) Funktionen der reellen Variablen \mathbf{x} . Die v_i seien C^1 -Funktionen (stetig differenzierbar), d.h. die n^2 partiellen Ableitungen

$$\frac{\partial v_i(\mathbf{x})}{\partial x_j} =: \partial_j v_i(\mathbf{x}) \quad (3.25)$$

existieren und sind stetig. Ein autonomes System gewöhnlicher Differentialgleichungen 1.Ordnung (kurz: "autonomes System") ist ein Gleichungssystem der Form

$$\dot{x}_i = v_i(\mathbf{x}) \quad \text{oder} \quad \dot{x}_i = v_i(x_j) \quad \text{oder} \quad \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}) \quad (3.26)$$

Eine Lösung des Systems (3.26) ist per Definition eine C^1 -Kurve, also eine stetig differenzierbare Abbildung $\gamma : t \mapsto \mathbf{x}(t)$ von einem Intervall I des \mathbb{R} nach \mathbb{R}^n , sodass für alle $t \in I$ gilt:

$$\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}(t)) \quad (3.27)$$

Die Funktionen v_i hat man sich also als Komponenten eines Vektorfelds zu denken, und eine Lösung als Kurve, deren Geschwindigkeitsvektor an jeder Stelle der Bahn mit diesem Vektorfeld übereinstimmt. Eine solche Kurve heisst auch Integralkurve des Vektorfelds. Ein Beispiel für $n = 2$ wäre das Geschwindigkeitsfeld einer zeitunabhängigen, idealen Flüssigkeit in der ebenen Hydrodynamik. Einer Lösung der Gleichung (3.26), genauer gesagt: der zu einer Lösung gehörigen Bildmenge $\{\mathbf{x} \in \mathbb{R}^2 | x_i = \gamma_i(t), t \in I\}$, entspricht die Bahn eines einzelnen Flüssigkeitsteilchens.¹

Manchmal ist das Vektorfeld nur in einer offenen Teilmenge Ω des \mathbb{R}^n definiert. Der Lösungsbegriff modifiziert sich dann in der offensichtlichen Weise.

Angenommen, \mathbf{v} hat einen Punkt $\bar{\mathbf{x}}$ seines Definitionsbereichs als Nullstelle (z.B. der Ursprung bei linearen Vektorfeldern (siehe Abschn.(3.1.1)) Dann heisst $\bar{\mathbf{x}}$ singulärer oder Gleichgewichtspunkt. Die Bedeutung von $\bar{\mathbf{x}}$ ist, dass man sofort die folgende Lösung von (3.26) angeben kann: $\mathbf{x}(t) \equiv \bar{\mathbf{x}}$. Die zugehörige Bahn besteht also aus dem einzigen Punkt $\bar{\mathbf{x}}$.

Verhalten unter Transformationen, Vektorfelder

Unter einer Transformation wollen wir eine C^1 - Abbildung $\mathbf{f} : x_i \in \mathbb{R}^n \mapsto y_i = f_i(x_j) \in \mathbb{R}^n$ oder zwischen offenen Teilmengen Ω und $\Omega' = \mathbf{f}(\Omega)$, die umkehrbar ist mit Umkehrabbildung ebenfalls C^1 . Daraus folgt, dass die \mathbf{x} - abhängige

¹Allerdings ist die Grundaufgabe der Hydrodynamik nicht so sehr die uns hier interessierende Bestimmung solcher Bahnen bei vorgegebenem Geschwindigkeitsfeld, sondern die Bestimmung, unter gewissen physikalisch motivierten Bedingungen, des (i.a. zeitabhängigen) Geschwindigkeitsfelds selber - und diese unterliegt einem System partieller Differentialgleichungen erster Ordnung, den sog. Euler-Gleichungen, siehe T1

Matrix $\partial_i f_j(x_k)$ nirgendwo singularär ist. Die (i.a. nichtlineare) Abbildung \mathbf{f} kann man sich als aktive Transformation vorstellen, aber auch als "passive" Koordinatentransformation, z.B. im euklidischen \mathbb{R}^3 zwischen cartesischen Koordinaten und Zylinderkoordinaten. Wir wollen, dass die Gleichung "kovariant" bezüglich all dieser Transformationen ist. D.h. wir wollen ein Transformationsgesetz für \mathbf{v} , sodass, wenn $\mathbf{x}(t)$ Lösung von (3.26) ist, die Bildkurve $\mathbf{y}(t) = \mathbf{f}(\mathbf{x}(t))$ Lösung des transformierten Systems

$$\dot{y}_i = \bar{v}_i(\mathbf{y}) \quad (3.28)$$

ist. Die Kettenregel sagt, dass das der Fall ist, wenn

$$\bar{v}_i(\mathbf{y}) = \sum_j (\partial_j f_i)(\mathbf{x}) v_j(\mathbf{x}), \quad (3.29)$$

woraus folgt, dass

$$\bar{v}_i(\mathbf{y}) = \sum_j (\partial_j f_i)(\mathbf{f}^{-1}(\mathbf{y})) v_j(\mathbf{f}^{-1}(\mathbf{y})) \quad (3.30)$$

Das ist gerade das Transformationsverhalten eines Vektorfelds, wie es in der Vektoranalysis gefordert wird.

Wir verwenden fallweise eine Schreibweise, die Vektorfelder als partielle Differentialoperatoren erster Ordnung auffasst, also z.B. statt $\mathbf{v} = (v_1, v_2, \dots, v_n)$ die Notation

$$\mathbf{v} = \sum_i v_i \frac{\partial}{\partial x_i} = \sum_i v_i \partial_i \quad (3.31)$$

Diese hat vom Standpunkt der Differentialgeometrie eine tiefere Bedeutung. Für uns ist sie eine nützliche Buchhaltungsvorschrift. Zum Beispiel ergibt sich die Transformationsformel für Vektorfelder aus (3.31) formal so: wir schreiben die Kettenregel in der Form

$$\frac{\partial}{\partial x_i} = \sum_j (\partial_i f_j) \frac{\partial}{\partial y_j} \quad (3.32)$$

Nunmehr setzen wir (3.32) in die linke Seite von

$$\mathbf{v} = \sum_i v_i \frac{\partial}{\partial x_i} = \sum_i \bar{v}_i \frac{\partial}{\partial y_i} \quad (3.33)$$

ein und lesen die Transformationsregel (3.30) ab.

Beispiel: Für den Wechsel zwischen cartesischen und Polarkoordinaten in der Ebene gilt: $r = \sqrt{x_1^2 + x_2^2}$ und $\phi = \arctan \frac{x_2}{x_1}$. Daraus ergibt sich

$$\frac{\partial}{\partial x_1} = \cos\phi \frac{\partial}{\partial r} - \frac{\sin\phi}{r} \frac{\partial}{\partial \phi}, \quad \frac{\partial}{\partial x_2} = \sin\phi \frac{\partial}{\partial r} + \frac{\cos\phi}{r} \frac{\partial}{\partial \phi} \quad (3.34)$$

Man kann die Rechnung, die zu den Gleichungen (3.34) führt, folgendermassen auffassen. Wir betrachten die "Koordinatenvektorfelder" $\frac{\partial}{\partial x_1}$ und $\frac{\partial}{\partial x_2}$, d.h. die Vektorfelder mit den (konstanten) Komponenten $(1, 0)$ bzw. $(0, 1)$ im (x_1, x_2) -Koordinatensystem. Ihre Komponenten im (r, ϕ) -Koordinatensystem errechnen sich so: wir rechnen die partiellen Ableitungen $\frac{\partial}{\partial x_1}$ und $\frac{\partial}{\partial x_2}$ nach der Kettenregel um und drücken die entstehenden Komponenten, die zunächst in Termen von (x_1, x_2) aufscheinen, gemäss $x_1 = r \cos \phi$, $x_2 = r \sin \phi$ durch die neuen Koordinaten (r, ϕ) aus. Sei das uns interessierende Vektorfeld nun das lineare Vektorfeld, das zur Gleichung vom Typ (V) in Abschnitt (3.1.1) gehört, also

$$\mathbf{v} = (\alpha x_1 - \beta x_2) \frac{\partial}{\partial x_1} + (\beta x_1 + \alpha x_2) \frac{\partial}{\partial x_2}. \quad (3.35)$$

Nach Einsetzen von (3.34) ergibt eine kurze Rechnung, dass

$$\mathbf{v} = \alpha r \frac{\partial}{\partial r} + \beta \frac{\partial}{\partial \phi} \quad (3.36)$$

Das Gleichungssystem für die Integralkurven lautet folglich in den neuen Koordinaten

$$\dot{r} = \alpha r, \quad \dot{\phi} = \beta, \quad (3.37)$$

und diese integrieren sich zu

$$r(t) = e^{\alpha t} r(0), \quad \phi(t) = \beta t + \phi(0) \quad (3.38)$$

in Übereinstimmung mit Abschnitt (2.1.1).

Symmetrien

Eine Transformation \mathbf{f} heisst Symmetrie von \mathbf{v} , wenn $\bar{v}_i(\mathbf{y}) = v_i(\mathbf{f}(\mathbf{x}))$, also ²

$$\sum_j (\partial_j f_i)(\mathbf{x}) v_j(\mathbf{x}) = v_i(\mathbf{f}(\mathbf{x})) \quad (3.39)$$

In diesem Fall ist mit jeder Lösung $\mathbf{x}(t)$ des Systems (3.26), auch $\mathbf{f}(\mathbf{x}(t))$ eine Lösung von (3.26).

Beispiel: Sei \mathbf{v} das "lineare Vektorfeld" $\mathbf{v}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ und \mathbf{f} die lineare Abbildung $\mathbf{f}(\mathbf{x}) = \mathbf{B}\mathbf{x}$. Dann ist \mathbf{f} eine Symmetrie des Vektorfelds \mathbf{v} genau dann, wenn $[\mathbf{B}, \mathbf{A}] = 0$.

In der Physik kommen häufig kontinuierliche Familien von Symmetrien vor, d.h. Transformationen $\mathbf{f}(\epsilon)$, die differenzierbar von einem reellen Parameter ϵ abhängen und sodass $\mathbf{f}(0) = \mathbf{id}$ bzw. $\mathbf{f}(0, \mathbf{x}) = \mathbf{x}$. Sei $\mathbf{f}(\epsilon)$ eine solche kontinuierliche Symmetrie von \mathbf{v} . Dann folgt, mit der Definition

$$\mathbf{w} = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \mathbf{f}(\epsilon) \quad (3.40)$$

²Vergleiche: Eine Transformation \mathbf{f} ist Symmetrie eines Skalarfelds Φ , wenn $\Phi(\mathbf{x}) = \Phi(\mathbf{f}(\mathbf{x}))$.

aus (3.39), dass

$$[\mathbf{w}, \mathbf{v}]_i := \sum_j w_j (\partial_j v_i) - v_j (\partial_j w_i) = 0 \quad (3.41)$$

Das Vektorfeld $[\mathbf{w}, \mathbf{v}]$ heisst die Lie-Klammer der Vektorfelder \mathbf{w} und \mathbf{v} .

Beispiel: Sei \mathbf{v} von der Koordinate x_1 unabhängig. Dann sind die Transformationen $f_1(\epsilon, \mathbf{x}) = x_1 + \epsilon$, $f_2(\epsilon, \mathbf{x}) = x_2$, $f_3(\epsilon, \mathbf{x}) = x_3, \dots, f_n(\epsilon, \mathbf{x}) = x_n$ Symmetrien. Das Vektorfeld \mathbf{w} in diesem Fall hat die Komponenten $w_i = \delta_{i1}$.

Wir betrachten als wichtiges Beispiel ein Kraftfeld im orthogonalen \mathbb{R}^3 : $\mathbf{F} = \nabla U$, wobei U invariant unter einer Drehung ist, also $U(\mathbf{R}\mathbf{q}) = U(\mathbf{q})$ für ein $\mathbf{R} \in SO(3)$. Daraus folgt

$$\sum_j r_{ji} (\partial_j U)(\mathbf{R}\mathbf{q}) = (\partial_i U)(\mathbf{q}) \quad \text{oder} \quad (\mathbf{R}^T \nabla U)(\mathbf{R}\mathbf{q}) = (\nabla U)(\mathbf{q}) \quad (3.42)$$

und daher, da $\mathbf{R}^T \mathbf{R} = \mathbf{E}$ ist,

$$(\partial_i U)(\mathbf{R}\mathbf{q}) = \sum_j r_{ij} (\partial_j U)(\mathbf{q}) \quad (3.43)$$

Das heisst aber, dass das durch $\mathbf{F} = \nabla U$ gegebene Vektorfeld unter der Drehung $\mathbf{y} = \mathbf{R}\mathbf{x}$ invariant ist. Sei U nun unter allen Drehungen invariant, also $U(\mathbf{x}) = u(\|\mathbf{x}\|)$. Dann muss \mathbf{F} auch unter infinitesimalen Drehungen invariant sein. Das heisst, in Anbetracht von (2.60), dass alle Vektorfelder \mathbf{v} der Form $v_i = \sum_j I_{ij} x_j$, wobei $I_{ij} + I_{ji} = 0$ ist, mit \mathbf{F} verschwindende Lie-Klammer haben. Also muss gelten, dass

$$\sum_j [v_j (\partial_j \partial_i U) - (\partial_j U) (\partial_j v_i)] = 0 \quad (3.44)$$

ist. Das wollen wir explizit nachrechnen: Nun ist ja $U(\mathbf{x}) = u(\|\mathbf{x}\|)$ und daher

$$\partial_i U = \frac{x_i}{\|\mathbf{x}\|} u', \quad \partial_i \partial_j U = \frac{x_i x_j}{\|\mathbf{x}\|^2} u'' + \left(\frac{\delta_{ij}}{\|\mathbf{x}\|} - \frac{x_i x_j}{\|\mathbf{x}\|^3} \right) u'$$

und $\partial_j v_i = I_{ij}$. Indem wir all dies in (3.44) einsetzen, stellen wir fest, dass die $x_i x_j$ -Ausdrücke im ersten Term sich wegen der Antisymmetrie von I_{ij} zu Null summieren und der δ_{ij} -Term sich mit dem zweiten Ausdruck in (3.44) weghebt. Also ist (3.44) bewiesen.

Folgendes ist ein zentrales Beispiel für Symmetrie in der Physik:

Sei $n = 6$ mit $x_i = (\vec{q}, \vec{p}) \in \mathbb{R}^3 \times \mathbb{R}^3$. Unter \mathbb{R}^3 ist hier wieder der orthogonale \mathbb{R}^3 verstanden. Sei U eine Funktion von \mathbf{q} . Die Gleichungen

$$\dot{\mathbf{q}} = \frac{1}{m} \mathbf{p} \quad \dot{\mathbf{p}} = -\nabla U \quad (3.45)$$

sind äquivalent zu den Newton'schen Bewegungsgleichungen für ein Partikelchen der Masse m im äusseren Potential U (siehe nächster Abschnitt). Wenn nun U

wie vorhin invariant unter einer Drehung ist, dann ist \mathbf{v} unter der Transformation $\mathbf{f}(\mathbf{q}, \mathbf{p}) = (\mathbf{R}\mathbf{q}, \mathbf{R}\mathbf{p})$ invariant. Wenn U sphärisch symmetrisch ist, muss daher das Vektorfeld

$$\mathbf{w} = \sum_{i,j} I_{ji} \left(q_i \frac{\partial}{\partial q_j} + p_i \frac{\partial}{\partial p_j} \right) \quad (3.46)$$

wobei \mathbf{I} ein beliebiger antisymmetrischer Operator im \mathbb{R}^3 ist, eine kontinuierliche Symmetrie sein, d.h. es gilt $[\mathbf{w}, \mathbf{v}] = 0$. Wir lassen die Rechnung aus.

Erhaltungsgrößen

Eine Funktion $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ heisst Erhaltungsgröße für das System (3.26), wenn

$$\mathbf{v}(\Phi) = \sum_i v_i(\mathbf{x}) \partial_i \Phi(\mathbf{x}) = 0 \quad (3.47)$$

Beispiel: Sei \mathbf{v} das Vektorfeld im \mathbb{R}^3 gegeben durch $\mathbf{v} = -x_2 \partial_1 + x_1 \partial_2$. Dann ist die Funktion $\Phi(\mathbf{x}) = \|\mathbf{x}\|^2$ erhalten.

Erhaltungsgrößen sind "erhalten" in dem Sinn, dass sie - entlang jeder Lösung von (3.26) - konstant sind, d.h. $\frac{d}{dt} \Phi(\mathbf{x}(t)) = 0$, sofern $\mathbf{x}(t)$ die Glg. (3.26) erfüllt. Beweis: Kettenregel! Es ist leicht, Erhaltungsgrößen anzugeben, wenn man schon alle Lösungen von (3.26) kennt. Der interessante Fall ist, wenn man Erhaltungsgrößen schon vor dem Lösen von (3.26) kennt³. Man kann dann die Zahl der abhängigen Variablen in (3.26) verringern, und das geht so:

Sei \mathbf{x}_0 ein Punkt, sodass $\nabla \phi(\mathbf{x}_0) \neq \mathbf{0}$ ist, z.B. sei $\partial_n \phi(\mathbf{x}_0) \neq 0$. Dann kann man, wegen des Satzes über implizite Funktionen (Analysis II), in einer Umgebung von \mathbf{x}_0 die Gleichung $\Phi(x_1, x_2, \dots, x_n) = y_n$ nach x_n auflösen, d.h. es gibt eine Funktion $F(x_1, x_2, \dots, x_{n-1}, y_n)$, sodass $\Phi(x_1, x_2, \dots, F(x_1, x_2, \dots, y_n)) = y_n$ ist. Wir führen statt (x_1, x_2, \dots, x_n) die neuen Koordinaten $y_1 = x_1, y_2 = x_2, \dots, y_n = \Phi(x_1, x_2, \dots, x_n)$ ein. Die Umkehrabbildung lautet $x_1 = y_1, x_2 = y_2, \dots, x_n = F(y_1, \dots, y_n)$. Das Vektorfeld \mathbf{v} in den neuen Koordinaten lautet

$\bar{v}_1(y_1, \dots, y_n) = v_1(y_1, \dots, F(y_1, \dots, y_n)), \bar{v}_2(y_1, \dots, y_n) = v_2(y_1, \dots, F(y_1, \dots, y_n)), \dots$
und $\bar{v}_n(y_1, y_2, \dots, y_n) = \sum_j \partial_j \Phi(x_1, x_2, \dots, x_n) v_j(x_1, x_2, \dots, x_n) = 0$ wegen (3.47).

Geometrisch gesprochen, ist das Vektorfeld \mathbf{v} tangential zur Hyperfläche $\Phi = \Phi_0 = \text{const}$, und die neuen Koordinaten sind dieser Hyperfläche "angepasst". Ein Beispiel ist die Energiehyperfläche im Phasenraum der klassischen Mechanik. Wir können nunmehr einen fixen Wert $y_n = \Phi_0$ annehmen und das reduzierte Gleichungssystem ($i = 1, 2, \dots, n-1$)

$$\dot{y}_i = v_i(y_1, y_2, \dots, y_{n-1}, F(y_1, y_2, \dots, y_{n-1}, \Phi_0)) \quad (3.48)$$

integrieren.

Beispiel: $\mathbb{R}^2 = \{(q, p)\}$, $\mathbf{v} = \frac{1}{m} p \partial_q - U'(q) \partial_p$ und $\Phi(q, p) = \frac{1}{2m} p^2 + U(q)$ ist

³Dies hängt sehr oft mit dem Vorliegen von kontinuierlichen Symmetrien zusammen.

erhalten. Wir folgen der obigen Prozedur, indem wir die Variablen (q, p) durch $y_1 = q$ und $y_2 = \Phi(q, p) = \frac{1}{2m}p^2 + U(q)$ ersetzen. In der Nähe von Punkten (q, p) , für die $p \neq 0$ ist, z.B. $p > 0$, ist die Funktionalmatrix

$$\begin{pmatrix} \frac{\partial y_1}{\partial q} & \frac{\partial y_1}{\partial p} \\ \frac{\partial y_2}{\partial q} & \frac{\partial y_2}{\partial p} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ U'(q) & \frac{p}{m} \end{pmatrix}$$

invertierbar und die Umkehrabbildung existiert und ist durch $q = y_1$ und $p = [2m(y_2 - U(y_1))]^{\frac{1}{2}}$ gegeben. Nun gilt $\partial_q = \partial_{y_1} + U' \partial_{y_2}$ und $\partial_p = \frac{p}{m} \partial_{y_2}$, sodass

$$\mathbf{v} = \frac{p}{m} \partial_q - U' \partial_p = \frac{p}{m} (\partial_{y_1} + U') \partial_{y_2} - U' \frac{p}{m} \partial_{y_2} = \frac{[2m(y_2 - U(y_1))]^{\frac{1}{2}}}{m} \partial_{y_1} \quad (3.49)$$

In Übereinstimmung mit der allgemeinen Theorie verbleibt daher einzig die Gleichung

$$\dot{y}_1 = \left[\frac{2}{m} (\Phi_0 - U(y_1)) \right]^{\frac{1}{2}} \quad (3.50)$$

zu integrieren.

3.2.2 Der Existenz- und Eindeutigkeitsatz

Wir geben diesen Satz hier nur etwas informell wieder und verzichten auf einen Beweis.

Satz: Sei das Vektorfeld $\mathbf{v}(\mathbf{x})$ in einem offenen Gebiet Ω des \mathbb{R}^n differenzierbar und $\mathbf{x}_0 \in \Omega$. Dann gibt es in Ω eine eindeutige Kurve, die die Gleichung $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$ löst und die Anfangsbedingung $\mathbf{x}(0) = \mathbf{x}_0$ erfüllt. (Insbesondere können sich diese Kurven niemals kreuzen.)

Warnung: Der Satz sagt nicht, dass die Kurve $\mathbf{x}(t)$ für alle Zeiten t existiert: die Kurve könnte in endlicher Zeit den Bereich Ω verlassen. Ja, selbst wenn $\Omega = \mathbb{R}^n$ ist, kann es passieren, dass das Vektorfeld im Unendlichen so schnell anwächst, dass das Teilchen in endlicher Zeit ins Unendliche katapultiert wird.

Beispiel: Sei $n = 1$ und $\mathbf{v}(\mathbf{x}) = v(x) = x^2$. Dann erfüllt die Kurve $x(t) = \frac{x_0}{1-x_0 t}$ die Gleichung mit der Anfangsbedingung $x(0) = x_0 > 0$. Die Kurve ist nur für $t \in (-\infty, 1/x_0)$ definiert. Wir haben $\lim_{t \rightarrow -\infty} x(t) = 0$ und $\lim_{t \rightarrow 1/x_0} x(t) = \infty$. Ein für $t = 0$ auf der positiven Zahlengeraden bei x_0 startendes Teilchen fliegt also für $t \rightarrow 1/x_0$ nach $+\infty$.

Zur obigen - lokalen - Aussage über Existenz und Eindeutigkeit kommt eine über "lokale Stabilität", die sehr informell so lautet: Wenn wir die Lösungen $\mathbf{x}_\epsilon(t)$ der Gleichung (3.27) zu den Anfangsbedingungen $\mathbf{x}_\epsilon(0) = \mathbf{x}_\epsilon$ betrachten und \mathbf{x}_ϵ hängt stetig von ϵ ab, so hängen, wenn \bar{t} ein fester Punkt im Definitionsbereich aller \mathbf{x}_ϵ 's ist, auch die Punkte $\mathbf{x}_\epsilon(\bar{t})$ stetig von ϵ ab. Wir werden diese Sätze hier nicht beweisen. Ihre Gültigkeit ist aber aus folgenden Gründen sehr plausibel: Im linearen Fall haben wir die Sätze tatsächlich bewiesen, und sogar mehr: wenn

$\Omega = \mathbb{R}^n$, waren die Lösungen für alle $-\infty < t < \infty$ definiert.

Ein weiteres Argument ist die Existenz einer Version, in der ein diskreter Zeitschritt, z.B. Δt , vorgegeben wird, innerhalb dessen die "Lösung" als linear angenommen wird, und zwar so:

$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{v}(\mathbf{x}_0) \Delta t$, $\mathbf{x}_2 = \mathbf{x}_1 + \mathbf{v}(\mathbf{x}_1) \Delta t$, ..., $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{v}(\mathbf{x}_n) \Delta t$. Dies heisst "Euler-Diskretisierung". Unter gewissen Umständen konvergiert diese im Limes $\Delta t \rightarrow 0$ gegen eine Lösung der Gleichung $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$.

Als Beispiel nehmen wir die lineare Gleichung $\dot{x} = x$ in \mathbb{R} mit $x(0) = x_0$ als AB. Wir teilen das Intervall $[0, t]$ in n gleiche Teile, d.h. $\Delta t = \frac{t}{n}$. Dann ist $x_{n+1} = x_n + x_n \Delta t = (1 + \Delta t)x_n = (1 + \Delta t)^{n+1}x_0$ oder $x_n = (1 + \Delta t)^n x_0 = (1 + \frac{t}{n})^n x_0$. Also ist $\lim_{n \rightarrow \infty} x_n(t) = \lim_{n \rightarrow \infty} (1 + \frac{t}{n})^n x_0 = e^t x_0$, wie es sein muss.

Schliesslich könnte man sich die Lösung $\mathbf{x}(t)$ als Taylorreihe um $t = 0$ vorstellen. Man sieht dann, dass bei gegebenem \mathbf{x}_0 alle Taylorkoeffizienten rekursiv berechnet werden können, z.B. gilt $\dot{x}_i(0) = v_i(\mathbf{x}_0)$, $\ddot{x}_i(0) = \sum_j \dot{x}_j(0) (\partial_j v_i)(\mathbf{x}_0)$, $\ddot{\ddot{x}}_i(0) = \sum_j \ddot{x}_j(0) (\partial_j v_i)(\mathbf{x}_0) + \sum_{j,k} \dot{x}_j(0) \dot{x}_k(0) (\partial_k \partial_j v_i)(\mathbf{x}_0)$, usw.

Der obige fundamentale Satz gilt im wesentlichen auch für nicht-autonome Systeme, also solche von der Form

$$\dot{\mathbf{x}} = \mathbf{v}(t, \mathbf{x}) \quad (3.51)$$

(siehe auch Abschnitt (3.2.6)). Das heisst, gegeben ein Anfangswert \mathbf{x}_0 : dann existiert ein offenes Intervall I , das 0 enthält, sodass $\mathbf{x}(t)$ für alle $t \in I$ die Gleichung (3.51) erfüllt und $\mathbf{x}(0) = \mathbf{x}_0$ ist. Eine analoge Aussage gilt natürlich, wenn die "Anfangszeit" $t = 0$ durch einen anderen Zeitpunkt t_0 ersetzt wird. Wir erklären, wie man im Prinzip nicht-autonome Systeme auf autonome zurückführen kann⁴. Die Idee ist, die unabhängige Variable t auf der r.S. von (3.51) zu einer abhängigen zu machen. Und zwar ersetzt man (3.51) durch das $(n+1)$ -dimensionale System mit der AB $\mathbf{x}(0) = \mathbf{x}_0$ durch

$$\frac{d\mathbf{x}}{dt} \equiv \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, s), \quad \frac{ds}{dt} \equiv \dot{s} = 1 \quad (3.52)$$

mit der Anfangsbedingung $\mathbf{x}(0) = \mathbf{x}_0$, $s(0) = 0$. Wegen der 2. Gleichung in (3.52) gilt dann $s = t$, und wir haben eine Lösung des ursprünglichen Systems.

Was aber autonome Systeme unterscheidet, ist diese Eigenschaft: wenn $\mathbf{x}(t)$ eine Lösung ist, dann auch $\mathbf{y}(t) = \mathbf{x}(t+s)$ für alle s , sodass $t+s$ im Definitionsbereich von \mathbf{x} liegt. Dies hat zur Folge, dass das nicht-lineare Analogon der Familie \mathbf{g}_t von Abbildungen aus (3.1.1) ebenfalls die Gruppeneigenschaft

$$\mathbf{g}_s \circ \mathbf{g}_t = \mathbf{g}_{s+t} \quad (3.53)$$

Zur Erinnerung: $\mathbf{g}_s(\mathbf{x})$ ist der Bahnpunkt $\mathbf{y}(s)$ mit der Anfangsbedingung $\mathbf{y}(0) = \mathbf{x}$. Der Beweis geht so: wenn wir die linke Seite von (3.53) auf den Punkt \mathbf{x}

⁴Wir sagen "im Prinzip", weil das nicht unbedingt der beste Weg ist, in einem konkreten Fall eine nicht-autonome Gleichung zu lösen.

anwenden, liefert sie den Bahnpunkt, der für $s = 0$ durch $\mathbf{g}_t(\mathbf{x})$ geht. Nun beschreibt $\mathbf{g}_{s+t}(\mathbf{x})$ wegen der obigen Eigenschaft autonomer Systeme ebenfalls eine Lösungskurve des Systems. Da diese weiters für $s = 0$ die selbe Anfangsbedingung erfüllt, wie die linke Seite, müssen wegen des Eindeutigkeitsatzes beide Seiten übereinstimmen.

Bemerkung: Aus (3.53) folgt, dass \mathbf{g}_t invertierbar ist. Und zwar gilt $(\mathbf{g}_t)^{-1} = \mathbf{g}_{-t}$.

3.2.3 Stabilität

Wir haben früher erwähnt, dass in einem Fix- oder Gleichgewichtspunkt $\bar{\mathbf{x}}$ des Vektorfelds, also $\mathbf{v}(\bar{\mathbf{x}}) = 0$, die konstante Kurve $\mathbf{x}(t) \equiv \bar{\mathbf{x}}$ eine Integralkurve ist. Der Eindeutigkeits-Teil des obigen Satzes gestattet nun, von dieser als DER Lösung zum Anfangswert $\bar{\mathbf{x}}$ zu sprechen. Es stellt sich die Frage nach der Stabilität dieses Gleichgewichtszustands, genauer nach dem Verhalten von Bahnen, die in der Nähe von $\bar{\mathbf{x}}$ starten, für sehr späte Zeiten. (Für kurze Zeiten sagt das obige Resultat, dass die Bahn in der Nähe von $\bar{\mathbf{x}}$ verbleibt.) Für $t \rightarrow \infty$ könnte die Bahn sich z.B. von $\bar{\mathbf{x}}$ weit entfernen, in der Nähe bleiben, oder sogar asymptotisch gegen $\bar{\mathbf{x}}$ streben. Wichtiges Mittel dafür ist die Linearisierung von $\dot{\mathbf{x}} = \mathbf{v}$ an der Lösung $\mathbf{x}(t) \equiv \bar{\mathbf{x}}$. Diese ist folgendermassen definiert: wir denken uns eine 1-parametrische Familie $\mathbf{x}(\lambda; t)$ von Lösungen mit $\mathbf{x}(0; t) = \mathbf{x}(t) \equiv \bar{\mathbf{x}}$. Mit der Definition

$$\boldsymbol{\xi}(t) = \left. \frac{\partial}{\partial \lambda} \mathbf{x}(\lambda; t) \right|_{\lambda=0} \quad (3.54)$$

und der Matrix

$$(\mathbf{A})_{ij} = (\partial_j v_i)(\bar{\mathbf{x}}) \quad (3.55)$$

ergibt sich aus $\dot{\mathbf{x}} = \mathbf{v}$, dass

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} \quad (3.56)$$

Die allgemeine Stabilitätstheorie können wir hier nicht behandeln⁵. Aber z.B. sagt ein allgemeiner Satz, den wir nicht beweisen, dass, wenn alle Eigenwerte von \mathbf{A} negativen Realteil haben, das Verhalten für grosse positive Zeiten von Lösungen von $\dot{\mathbf{x}} = \mathbf{v}$, die nahe bei $\bar{\mathbf{x}}$ starten, dem der Lösungen von (3.56) entspricht. Insbesondere gilt, dass $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \bar{\mathbf{x}}$ ("asymptotische Stabilität")

Beispiel: Die Bewegungsgleichung des ebenen Pendels mit Dämpfung $\Gamma > 0$ lautet

$$\ddot{\theta} + 2\Gamma\dot{\theta} + \omega^2 \sin\theta = 0 \quad (3.57)$$

Mit der Bezeichnung $\mathbf{x} = (\theta, \psi)$ und $\psi = \dot{\theta}$ ist die Gleichung (3.57) äquivalent zu $\dot{\mathbf{x}} = \mathbf{v}$, wobei

$$\mathbf{v} = \psi \partial_\theta - (\omega^2 \sin\theta + 2\Gamma\psi) \partial_\psi \quad (3.58)$$

⁵Einfach ist nur der Fall $n = 1$, siehe nächster Abschnitt.

Für $(\theta, \psi) = (0, 0)$ hat \mathbf{v} einen Gleichgewichtspunkt. Die Matrix \mathbf{A} lautet

$$\begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\Gamma \end{pmatrix},$$

und daher ist die Gleichgewichtslage asymptotisch stabil. Anders verhält es sich mit der durch $(\theta, \psi) = (\pi, 0)$ gegebenen Gleichgewichtslage. Hier ist die entsprechende Matrix \mathbf{A} gleich

$$\begin{pmatrix} 0 & 1 \\ \omega^2 & -2\Gamma \end{pmatrix},$$

und das obige Stabilitätsresultat ist nicht anwendbar.

3.2.4 Nichtlineare ODE's für $n = 1$

Hier ist $\mathbf{x} = x \in \mathbb{R}$ und

$$\dot{x} = v(x), \quad (3.59)$$

wobei wir der Einfachheit halber annehmen, dass v in ganz \mathbb{R} definiert ist. In jedem Intervall, in dem $v \neq 0$ ist, ist die Umkehrfunktion von $t(x)$, gegeben durch

$$t = \int_{x_0}^x \frac{dy}{v(y)}, \quad (3.60)$$

eine (und daher: die) Lösung von (3.59) mit der Anfangsbedingung $x(0) = x_0$, sofern $v(x_0) \neq 0$ ist. Wenn $v(x_0) = 0$ ist, ist die Lösung natürlich $x(t) \equiv x_0$. Sei nun wieder $v(x_0) \neq 0$, etwa $v(x_0) > 0$. Dann ist $t(x)$ in (3.60) im grössten Intervall $[x_0, X)$ definiert, sodass $v(x) > 0$ ist (insbesondere könnte $X = \infty$ sein). Folglich existiert die Lösung $x(t)$ für t im Intervall $[0, T)$, wobei

$$T = \int_{x_0}^X \frac{dy}{v(y)} \quad (3.61)$$

Jetzt gibt es zwei Möglichkeiten.

a) $X = \infty$, d.h. v ist für alle $x > x_0$ positiv. Hier ist der Definitionsbereich von $x(t)$ Unendlich - d.h. die Lösung existiert für alle positiven Zeiten - genau dann, wenn das Integral (3.61) divergiert ($T = \infty$). Andernfalls⁶ hat dieses Integral einen endlichen Wert T , und die Lösung existiert nur für $0 \leq t < T$. Das ist der früher erwähnte Fall, wo das Teilchen in endlicher Zeit ins Unendliche befördert wird: dies setzt voraus, dass $v(x)$ für $x \rightarrow \infty$ hinreichend stark anwächst⁷.

b) $X < \infty$, also ist $v(X) = 0$: aus der Differenzierbarkeit von $v(x)$ folgt, dass $\lim_{y \rightarrow X} \frac{v(y)}{X-y}$ konvergiert, insbesondere endlich ist. Folglich gibt es eine positive

⁶Dass T nur divergieren oder endlich sein kann, liegt daran, dass es ein Integral über eine positive Funktion ist: eine nicht-abnehmende unendliche Folge reeller Zahlen geht entweder nach Unendlich oder hat einen Grenzwert.

⁷So ist z.B. für $x_0 > 0$ das Integral $\int_{x_0}^{\infty} \frac{dy}{y^k}$ genau dann endlich, wenn $k > 1$.

Konstante C , sodass im Intervall $x_0 \leq y < X$ die Ungleichung $\frac{C}{v(y)} > \frac{1}{X-y}$ gilt. Daher ist $T = \infty$, und die Lösung existiert für alle positiven Zeiten. Wir haben es hier mit einer Bahn zu tun, die sich dem Gleichgewichtspunkt X beliebig annähert. Dass das unendlich lange dauern muss, folgt schon aus dem Eindeutigkeits-Teil des fundamentalen Existenzsatzes.

Hinsichtlich der Abhängigkeit der Lösung $x(t)$ vom Anfangswert x_0 bemerken wir ohne Beweis, dass

$$\partial_{x_0} x(t) = \frac{v(x(t))}{v(x_0)} \quad (3.62)$$

ist, sofern $v(x_0) \neq 0$, andernfalls ist klarerweise $\partial_{x_0} x(t) \equiv 1$.

Wir betrachten jetzt ein instruktives

Beispiel: Die sog. logistische Gleichung ($x \geq 0$)

$$\dot{x} = x(1 - x), \quad (3.63)$$

Man kann $\dot{x} = x$ als das Reproduktionsgesetz einer Spezies auffassen und den Term $-x^2$ als den Effekt, dass bei starker Zunahme der Zahl der Individuen die Konkurrenz um Nahrung zunimmt, was die Reproduktionsrate absenkt. Die Gleichung hat die beiden Gleichgewichtslagen $x = 0$ und $x = 1$. Für $0 < x_0 < 1$ muss gelten, dass $\lim_{t \rightarrow \infty} x(t) = 1$. Die explizite Lösung lautet

$$x(t) = \frac{e^t}{\frac{1-x_0}{x_0} + e^t} \quad (3.64)$$

Wir betrachten noch die verwandte Gleichung

$$\dot{x} = x(1 - \epsilon x), \quad \epsilon > 0 \quad (3.65)$$

Diese kann durch die "Skalierung" $x = \frac{1}{\epsilon} y$ auf (3.64) zurückgeführt werden. Die Lösung zur Anfangsbedingung $x(0) = x_0$ mit $0 < x_0 < \frac{1}{\epsilon}$ ergibt sich zu

$$x(t) = \frac{e^t}{\frac{1-\epsilon x_0}{x_0} + \epsilon e^t} \quad (3.66)$$

Hier gilt natürlich, dass $\lim_{t \rightarrow \infty} x(t) = \frac{1}{\epsilon}$. Im Limes $\epsilon \rightarrow 0$ geht $x(t)$ gegen $\bar{x}(t) = e^t x_0$, wie es sein muss.

Zusammenfassend können wir sagen, dass die Bahnstruktur im Fall $n = 1$ einfach ist: Bahnen laufen in den in ihrer Richtung nächsten Gleichgewichtspunkt und, wenn es keinen gibt, ins Unendliche. Ersteres dauert unendlich lange. In höheren Dimensionen ist die Situation wesentlich komplizierter: z.B. kann eine Bahn asymptotisch gegen eine periodische Bahn streben (also insbesondere weder ins Unendliche, noch gegen einen Gleichgewichtspunkt laufen). Für $n = 2$ gibt es immerhin noch eine einigermaßen vollständige Theorie, für $n \geq 3$ auch das nicht mehr. Systeme mit $n \geq 3$ zeigen i.a. chaotisches Verhalten (was freilich nicht heisst, dass diese Systeme einer mathematischen Behandlung unzugänglich sind).

3.2.5 $n = 2$: planare dynamische Systeme

Wir behandeln hier nur den Fall sog. Hamilton'scher Systeme. In Anlehnung an die Mechanik schreiben wir $(x_1, x_2) = (q, p)$ und nehmen an, dass es eine Funktion $H(q, p)$ gibt, sodass $v_1 = \partial_p H$ und $v_2 = -\partial_q H$. Das Gleichungssystem lautet also

$$\dot{q} = (\partial_p H)(q, p), \quad \dot{p} = -(\partial_q H)(q, p) \quad (3.67)$$

Fundamental ist die folgende Beobachtung: H ist eine Erhaltungsgrösse für das System (3.67). Denn

$$\dot{q} \partial_q H + \dot{p} \partial_p H = (\partial_p H)(\partial_q H) - (\partial_q H)(\partial_p H) = 0 \quad (3.68)$$

Der folgende Einschub ist etwas schwieriger und daher optional.

Einschub: Formel und Satz von Liouville*

Eine weitere wichtige Eigenschaft des Systems (3.67) ist, dass \mathbf{v} divergenzfrei ist, denn

$$\partial_1 v_1 + \partial_2 v_2 = \partial_q \partial_p H - \partial_p \partial_q H = 0 \quad (3.69)$$

Die Bedeutung dieser Aussage beruht auf folgender Formel von Liouville. Sei Ω ein endlicher Bereich in \mathbb{R}^2 . Sei weiters $\Omega(t)$ das Bild dieses Bereichs unter dem Fluss eines Vektorfelds \mathbf{v} , m.a.W. $\Omega(t) = \mathbf{g}_t(\Omega)$, wobei $\mathbf{g}_t(\mathbf{x})$ (siehe 3.2.2) definiert ist durch $\dot{\mathbf{g}}_t = \mathbf{v}(\mathbf{g}_t)$ und $\mathbf{g}_0(\mathbf{x}) = \mathbf{x}$. Sei $V(t)$ definiert durch

$$V(t) = \int_{\Omega(t)} dq' dp' \quad (3.70)$$

Dann gilt

$$\dot{V}(t) = \int_{\Omega(t)} (\operatorname{div} \mathbf{v}) dq' dp' \quad (3.71)$$

Wir führen im Integral in (3.70) die neue Integrationsvariable (q, p) ein gemäss $\mathbf{x}' = (q', p') = \mathbf{g}_t(\mathbf{x}) = \mathbf{g}_t((q, p))$. Dann ergibt die Regel für das Verhalten von Integralen unter Transformationen, dass

$$V(t) = \int_{\Omega(t)} dq' dp' = \int_{\Omega(0)} \left| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right| dq dp \quad (3.72)$$

Wir wollen die rechte Seite von (3.72) unter dem Integralzeichen nach t differenzieren. Wir benutzen zunächst die Formel (siehe Übungen)

$$\partial_t \det \mathbf{A} = \det \mathbf{A} \operatorname{tr}(\mathbf{A}^{-1} \dot{\mathbf{A}}) \quad (3.73)$$

und $\dot{\mathbf{g}} = \mathbf{v}(\mathbf{g})$. Daraus folgt, dass

$$(\dot{\mathbf{A}})_{ij} = \frac{\partial \dot{g}_i}{\partial x_j} = \sum_k (\partial_j g_k)(\mathbf{x}) (\partial_k v_i)(\mathbf{g}(\mathbf{x})) \quad (3.74)$$

und somit

$$\operatorname{tr}(\mathbf{A}^{-1}\dot{\mathbf{A}}) = \sum_k (\partial_k v_k)(\mathbf{g}(\mathbf{x})) . \quad (3.75)$$

Daher ist

$$\dot{V}(t) = \int_{\Omega(0)} \sum_k (\partial_k v_k)(\mathbf{g}_t(\mathbf{x})) \left| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right| dq dp = \int_{\Omega(t)} \sum_k (\partial_k v_k)(q', p') dq' dp' \quad (3.76)$$

wobei wir im letzten Schritt wieder die Transformationsformel für Integrale benutzt haben. Das beendet den Beweis von (3.71).

Wir schliessen aus (3.72), dass im Fall der Hamilton'schen Gleichungen (3.67) das Volumen V konstant ist. Dies ist der für die statistische Mechanik fundamentale Satz von Liouville.

Eine kleine Anwendung im gegenwärtigen Zusammenhang ist, dass folgende Situation unmöglich ist: Es gibt einen Gleichgewichtspunkt \mathbf{x}_0 und dazu eine Umgebung Ω , sodass alle in Ω startenden Bahnen nach \mathbf{x}_0 streben. Denn dann müsste gelten, dass $\lim_{t \rightarrow \infty} V(t) = 0$, im Widerspruch zu $\dot{V}(t) = 0$.

Die für die Physik wichtigste Gruppe Hamilton'scher Systeme sind die von der Form

$$H = \frac{p^2}{2} + U(q) \quad (3.77)$$

(Teilchen der Masse 1 im äusseren Potential U). Bahnen liegen auf Niveaulinien der Funktion $H : (q, p) \in \mathbb{R}^2 \mapsto H(q, p) = \frac{p^2}{2} + U(q) \in \mathbb{R}$. Anders ausgedrückt: Jede Bahn liegt auf einer durch $H^{-1}(E)$ gegebenen Kurve. (Vergleiche: "Höhenschichtlinien".) E ist der konstante Wert von H entlang dieser Bahn.

Beispiel 1: Das Potential $U(q)$ sei "vom Typ einer Parabel", d.h. konvex, d.h. $U'' < 0$. Dann gibt es genau ein Minimum für $q = \bar{q}$ und die Gleichung $E = U(x)$ hat für $E > \bar{E} = U(\bar{q})$ genau 2 Lösungen $a(E)$ und $b(E)$, und es gelte $U'(a) < 0$ und $U'(b) > 0$. Dann sind die Mengen $H^{-1}(E)$ geschlossene Kurven: der einzelne Punkt $(q, p) = (\bar{q}, 0)$ für $E = \bar{E}$ und für zunehmendes $E > \bar{E}$ eine Folge ineinanderliegender, den Punkt $(\bar{q}, 0)$ enthaltender, Kurven vom Typ eines Kreises (keine Selbstüberschneidungen), gegeben durch $p = \pm \sqrt{2(E - U(q))}$. Die entsprechenden Lösungen sind Kurven $(q(t), p(t))$, die periodisch in t sind⁸. Die Länge der Periode ist

$$T(E) = \sqrt{2} \int_{a(E)}^{b(E)} \frac{dq}{\sqrt{E - U(q)}} \quad (3.78)$$

Der Integrand verhält sich an den Grenzen wie $(a - q)^{-\frac{1}{2}}$ bzw. $(b - q)^{-\frac{1}{2}}$, konvergiert also. Sei z.B. $U(q) = \frac{\omega^2 q^2}{2}$ mit $\omega > 0$. Dann gibt es ein eindeutiges Minimum bei $q = 0$ und Energie $E = 0$. Für $E > 0$ sind die Mengen $H^{-1}(E)$ die Ellipsen

⁸Hier ist streng genommen noch etwas zu beweisen, das wir später nachholen.

$p^2 + \omega^2 q^2 = 2E$. Für $T(E)$ erhalten wir die bekannte Formel

$$T(E) = \sqrt{2} \int_{-\frac{\sqrt{2E}}{\omega}}^{\frac{\sqrt{2E}}{\omega}} \frac{dq}{\sqrt{E - \frac{\omega^2 q^2}{2}}} = \frac{2}{\omega} \int_{-1}^1 \frac{dy}{\sqrt{1 - y^2}} = \frac{2\pi}{\omega} \quad (3.79)$$

In diesem speziellen Fall ist also die Schwingungsdauer von der Energie unabhängig.

Beispiel 2: Ebenes Pendel ohne Dämpfung ("Schaukel"). Hier ist (siehe **(3.2.3)**) $H = \frac{\psi^2}{2} - \omega^2 \cos\theta$. Den Phasenraum $\{(\theta, \psi)\}$ stellt man sich am besten als Zylinder $\mathbb{S}^1 \times \mathbb{R}$ vor, in dem jeder der Punkte $(-\pi, \theta)$ jeweils mit (π, θ) identifiziert ist. In diesem gibt es die 2 Gleichgewichtslagen $(\theta, \psi) = (0, 0)$ und $(\theta, \psi) = (\pi, 0)$. Diese haben die Energie $E = -\omega^2$ bzw. $E = \omega^2$. Die Struktur der Gesamtheit der Bahnen ist folgende:

- (i) $E < -\omega^2$: keine Bahnen
- (ii) $E = -\omega^2$: ein Punkt, nämlich die stabile Gleichgewichtslage $(0, 0)$
- (iii) $-\omega^2 < E < \omega^2$: $H^{-1}(E)$ ist eine im Uhrzeigersinn durchlaufene geschlossene Kurve mit $(0, 0)$ in ihrem Inneren: diese entsprechen Schwingungen um die stabile Gleichgewichtslage.
- (iv) $E = \omega^2$: bestehen aus 3 Komponenten, zunächst der instabilen Gleichgewichtslage $(\pi, 0)$. Weiters zwei offenen Kurve, die $(-\pi, 0)$ mit $(\pi, 0)$ (auf dem Zylinder derselbe Punkt) verbinden, und zwar eine für $\psi > 0$ von links nach rechts und eine für $\psi < 0$ von rechts nach links: dies sind die 2 Kriechbahnen, die für $t \rightarrow \infty$ und $t \rightarrow -\infty$ in den instabilen Gleichgewichtspunkt laufen.
- (v) $E > \omega^2$: $H^{-1}(E)$ besteht aus 2 Komponenten, beides auf dem Zylinder geschlossene Bahnen. Erstens nach rechts laufende für $\psi > 0$, zweitens nach links laufende für $\psi < 0$. Dies sind die 2 Klassen von Bahnen, bei denen die Schaukel "die volle Umdrehung macht".

Wie in **(3.2.1)** erklärt, reduziert sich das Problem auf die Integration der Gleichung ($H = H_0 = E$)

$$\dot{q} = \sqrt{2(E - U(q))} \quad (3.80)$$

(Ende des Beispiels 2) Wir haben Komponenten von $H^{-1}(E)$, die geschlossene Kurven waren, mit periodischen Bahnen des dynamischen Systems identifiziert. Hier ist aber noch etwas zu beweisen. Die zugehörige Formel, nämlich

$$t(q) = \int_{a(E)}^q \frac{dq'}{\sqrt{2(E - U(q'))}} \quad (3.81)$$

macht zunächst nur im Intervall $0 \leq t \leq \frac{1}{2} T(E)$ Sinn. Die Umkehrfunktion hat die Eigenschaft, dass

$$\lim_{t \rightarrow 0} \dot{q}(t) = \lim_{t \rightarrow 0} p(t) = \lim_{t \rightarrow \frac{T}{2}} \dot{q}(t) = \lim_{t \rightarrow \frac{T}{2}} p(t) = 0 \quad (3.82)$$

Wir setzen das im Intervall $[0, \frac{T}{2}]$ definierte $(q_1(t), p_1(t))$ in das Intervall $[\frac{T}{2}, T]$ symmetrisch fort, also: $(q_2(t), p_2(t)) = (q_1(T-t), -p_1(T-t))$. Wegen $p(\frac{T}{2}) = 0$, ist diese Fortsetzung stetig. Sie erfüllt weiters die Bewegungsgleichungen (warum?) und $(q_2(T), p_2(T)) = (q_1(0), p_1(0))$. Sie kann nunmehr periodisch fortgesetzt werden.

3.2.6 Nicht-autonome Systeme und Gleichungen höherer Ordnung

Es ist schon mehrfach angeklungen, dass Gleichungen höherer Ordnung auf ein System 1.Ordnung zurückgeführt werden können. Wir betrachten hier zunächst Gleichungen der Form

$$x^{(n)} + a_1(t)x^{(n-1)} + a_2(t)x^{(n-2)} + \dots + a_n(t)x = 0 \quad (3.83)$$

Dies ist eine einzelne ODE n -ter Ordnung. Sie ist nicht autonom, wenn die Koeffizienten a_i von t abhängen. Durch die Substitutionen $x_1 = x, x_2 = \dot{x}, \dots, x_n = x^{(n-1)}$ sowie $\mathbf{x} = (x_1, x_2, \dots, x_n)$ erhalten wir

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} \quad (3.84)$$

wobei

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \cdot \\ & & & & & & 0 & 1 \\ -a_n & \cdot & \cdot & \cdot & -a_2 & -a_1 & & \end{pmatrix} \quad (3.85)$$

Die nun folgenden Aussagen gelten, wenn nicht ausdrücklich auf (3.83) Bezug genommen wird, unabhängig davon, ob $\mathbf{A}(t)$ in (3.84) die Form (3.85) hat oder nicht. Wir stellen zunächst ohne Beweis fest, dass dieses Gleichungssystem für gegebenes $\mathbf{x}(0) = \mathbf{x}_0$ für alle $-\infty < t < \infty$ eine Lösung besitzt und diese eindeutig ist. Folglich hat auch die Gleichung (3.83) Lösungen für alle $-\infty < t < \infty$, die eindeutig sind, wenn $x(0), \dot{x}(0), \dots, x^{(n-1)}(0)$ vorgegeben werden. Beinahe evident ist folgender

Satz: Die Lösungen von (3.84) bilden einen n -dimensionalen reellen Vektorraum. Beweis: Die reellwertigen Funktionen $\mathbf{x} : t \in \mathbb{R} \mapsto \mathbf{x}(t) \in \mathbb{R}^n$, die Lösungen von (3.84) sind, bilden einen reellen Vektorraum \mathbb{V} . Wir betrachten den Operator $\mathbf{B} : \mathbb{V} \rightarrow \mathbb{R}^n$, der jedem $\mathbf{x} \in \mathbb{V}$ den Wert $\mathbf{x}(0)$ zuordnet. Diese Abbildung ist surjektiv: das folgt aus dem Existenzsatz, nämlich dass zu jeder Anfangs (End-) Bedingung \mathbf{x}_0 eine Lösung existiert. Sie ist aber auch injektiv, denn der Kern von \mathbf{B} besteht nur aus der Nulllösung von (3.84), wegen des Eindeutigkeitsatzes, qed. Eine Basis von \mathbb{V} heisst ein Fundamentalsystem für die Gleichung (3.84). Es gilt

der (offensichtliche)

Satz: Vektorwertige Funktionen $\mathbf{x}_i(t)$ mit $1 \leq i \leq n$ in \mathbb{V} bilden ein Fundamentalsystem genau dann, wenn die Funktion W , gegeben durch die Determinante

$$W = \det[\mathbf{x}_1(0), \mathbf{x}_2(0), \dots, \mathbf{x}_n(0)] \quad (3.86)$$

ungleich 0 ist. Ebenso klar ist, dass diese Bedingung äquivalent ist mit

$$W(t) = \det[\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t)] \equiv \det \mathbf{M}(t) \neq 0 \quad (3.87)$$

für irgendein t . Die Grösse W heisst Wronski-Determinante. Etwas expliziter können wir folgendes sagen. Zunächst gilt

$$[\dot{\mathbf{x}}_1(t), \dot{\mathbf{x}}_2(t), \dots, \dot{\mathbf{x}}_n(t)] = \mathbf{A}(t) [\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t)] \quad (3.88)$$

Folglich ist (siehe (3.73))

$$\dot{W}(t) = W(t) \operatorname{tr}(\dot{\mathbf{M}}(t)\mathbf{M}^{-1}(t)) = \operatorname{tr}(\mathbf{A}(t))W(t) \quad (3.89)$$

oder

$$W(t) = e^{\int_0^t \operatorname{tr}(\mathbf{A}(t')) dt'} W(0), \quad (3.90)$$

und der 1.Faktor auf der r.S. von (3.90) ist niemals = 0.

Diese Überlegungen übertragen sich natürlich sinngemäss auf die Gleichung (3.83). Beispiel: Seien $x(t), y(t)$ Lösungen der physikalisch wichtigen Gleichung

$$\ddot{x} + \omega^2(t)x = 0. \quad (3.91)$$

Dann ist $W(t) = xy - \dot{y}x$, und wegen des Fehlens eines Terms mit 1.Ableitung in (3.91) gilt, dass $\operatorname{tr}(\mathbf{A}(t)) = 0$. Folglich ist in diesem Fall $\dot{W}(t) = 0$, was man auch direkt nachrechnen kann.

Wenn wir im Besitz eines Fundamentalsystems $\{\mathbf{x}_i(t)\}$ sind, ist die Lösung des Anfangswertproblems für (3.84) leicht: Wir setzen

$$\mathbf{x}(t) = \sum_i \alpha_i \mathbf{x}_i(t), \quad (3.92)$$

und das Gleichungssystem $\mathbf{x}(0) = \mathbf{x}_0$ für die Konstanten α_i hat eine eindeutige Lösung.

Wie löst man die inhomogene Gleichung

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t) \quad (3.93)$$

mit der Anfangsbedingung $\mathbf{x}(0) = \mathbf{x}_0$? Es genügt, die spezielle Lösung $\mathbf{y}(t)$ von (3.93) zu finden, die die AB $\mathbf{y}(0) = 0$ erfüllt, und das geht so: wir suchen n

⁹Wir haben hier die Existenz von \mathbf{M}^{-1} , also das Nichtverschwinden von W benutzt. Man kann zeigen, dass die Formel (3.89) auch ohne diese Einschränkung gilt.

Lösungen $\mathbf{x}_i(t)$ von (3.84), die $(\mathbf{x}_i)_j(t_0) = \delta_{ij}$ erfüllen. Man könnte auch sagen wir betrachten die Operator-Differentialgleichung

$$\dot{\mathbf{G}} = \mathbf{A}\mathbf{G} \quad (3.94)$$

mit der Anfangsbedingung $\mathbf{G}(t_0) = \mathbf{E}$. Die Spalten der zu \mathbf{G} gehörigen Matrix bilden gerade die eben eingeführten Vektoren $\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)$. Der Operator \mathbf{G} ist eine Funktion $\mathbf{G} = \mathbf{G}(t; t_0)$ mit $\mathbf{G}(t_0, t_0) = \mathbf{E}$. (Im autonomen Fall würde gelten, dass $\mathbf{G}(t; t_0)$ nur von $t - t_0$ abhängt.) Wir bilden nun

$$\mathbf{y}(t) = \int_0^t \mathbf{G}(t; t') \mathbf{b}(t') dt' \quad (3.95)$$

und erhalten

$$\dot{\mathbf{y}}(t) = \int_0^t \mathbf{A}(t)\mathbf{G}(t; t') \mathbf{b}(t') dt' + \mathbf{G}(t, t)\mathbf{b}(t) = \mathbf{A}(t)\mathbf{y}(t) + \mathbf{b}(t), \quad (3.96)$$

wie gewünscht. Wir können diese Überlegung auf die inhomogene Gleichung

$$x^{(n)} + a_1(t)x^{(n-1)} + a_2(t)x^{(n-2)} + \dots + a_n(t)x = b(t) \quad (3.97)$$

übertragen, indem wir in (3.93) wie oben (3.85) benutzen und für \mathbf{b} den Vektor $\mathbf{b}(t) = (0, 0, \dots, b(t))$ einsetzen. Nun übersetzt sich die inhomogene Lösung (3.95) in der Sprache der ursprünglichen Gleichung (3.83) folgendermassen: Sei $G(t, t_0)$ die (bezüglich t) Lösung von (3.83) mit der Anfangsbedingung $G(t_0, t_0) = 0, \dot{G}(t_0, t_0) = 0, \dots, G^{(n-1)}(t_0, t_0) = 1$. Dann ist ¹⁰

$$y(t) = \int_0^t G(t, t')b(t')dt' \quad (3.98)$$

die Lösung von (3.97) mit den Anfangswerten $y(0) = 0, \dot{y}(0) = 0, \dots, y^{(n-1)}(0) = 0$. Natürlich lässt sich das auch direkt nachrechnen.

Bemerkung: Wenn die Koeffizienten a_i zeitunabhängig sind, hängt die "Green-Funktion" $G(t, t')$ nur von $t - t'$ ab.

Wir betrachten zum Abschluss noch diesen, d.h. den zeitunabhängigen Fall von (3.83), also

$$x^{(n)} + a_1 x^{(n-1)} + a_2 x^{(n-2)} + \dots + a_n x = 0 \quad (3.99)$$

Der schnellste Weg zum Auffinden von Lösungen ist der Exponentialansatz $x(t) = e^{\lambda t}$. Wir erhalten die Bedingung

$$Q(\lambda) \equiv \lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0 \quad (3.100)$$

¹⁰Die Funktion $G(t, t_0)$ ist nichts anderes als die $(1, n)$ -Komponente von $\mathbf{G}(t, t_0)$, also die erste Komponente des Vektors $\mathbf{x}_n(t)$.

Die Funktion $Q(\lambda)$ in (3.100) ist nichts anderes als

$$Q(\lambda) = (-1)^n P(\lambda) \quad (3.101)$$

wobei $(-1)^n P(\lambda) = \det(\lambda \mathbf{E} - \mathbf{A})$, d.h. $P(\lambda)$ das charakteristische Polynom des Operators \mathbf{A} in (3.85) ist (Beweis von (3.101) mindestens für $n = 2, 3$?).

Bemerkung: Es ist kein Zufall, dass eine Relation zwischen $Q(\lambda)$ und $P(\lambda)$ besteht. Denn der Ansatz $x(t) = e^{\lambda t}$ bewirkt für (3.99) den Ansatz $\mathbf{x}(t) = \mathbf{c} e^{\lambda t}$ für den konstanten Vektor $\mathbf{c} = (1, \lambda, \dots, \lambda^{n-1})$, und das führt zur Gleichung $(\lambda \mathbf{E} - \mathbf{A}) \mathbf{c} = 0$, also der Eigenwertbedingung für den Operator (3.85) im zeitunabhängigen Fall. Also gehört zu jeder exponentiellen Lösung von (3.99) mindestens ein Eigenvektor von \mathbf{A} . Umgekehrt ist jeder Eigenvektor von \mathbf{A} zum Eigenwert λ proportional zum Vektor \mathbf{c} , und jede Komponente dieses Vektors mal $e^{\lambda t}$ ist Lösung von (3.99).

Seien $\lambda_i, i = 1, \dots, r$ die verschiedenen Eigenwerte von \mathbf{A} . Dann gibt es natürliche Zahlen q_1, \dots, q_r , sodass

$$Q(\lambda) = (\lambda - \lambda_1)^{q_1} \cdots (\lambda - \lambda_r)^{q_r} \quad (3.102)$$

Es gilt, dass $q_1 + \dots + q_r = n$. Die Zahlen q_i sind die algebraischen Vielfachheiten der Eigenwerte λ_i . Da die Koeffizienten a_i reell sind, ist zu jedem komplexen Eigenwert auch die komplex Konjugierte ein Eigenwert. Das fundamentale Resultat ist dieses. Sei $0 \leq k \leq r_i - 1$. Wir betrachten folgende Funktionenfamilie. Für λ_i reell, definieren wir $e_{k,i}(t) = t^k e^{\lambda_i t}$. Wenn λ_i komplex ist, sei $f_{k,i}(t) = t^k \cos \lambda_i t$ und $g_{k,i}(t) = t^k \sin \lambda_i t$. Dann gilt der

Satz: Die allgemeine Lösung der Gleichung (3.99) ist eine (eindeutige) Linearkombination der n Funktionen $e_{k,i}(t), f_{k,i}(t), g_{k,i}(t)$.

Beweisskizze: Wir können die Gleichung (3.99) schreiben als

$$Q \left(\frac{d}{dt} \right) x = \left(\frac{d}{dt} - \lambda_1 \mathbf{E} \right)^{q_1} \cdots \left(\frac{d}{dt} - \lambda_r \mathbf{E} \right)^{q_r} x = 0, \quad (3.103)$$

wobei Q nunmehr als Operator aufgefasst wird, der Funktionen von \mathbb{R}^1 nach \mathbb{R}^1 ineinander abbildet. Wichtig ist, dass alle Faktoren auf der r.S. von (3.103) miteinander kommutieren. Entscheidend ist nun die Identität

$$\left(\frac{d}{dt} - \lambda \mathbf{E} \right)^{k+1} t^k e^{\lambda t} = 0, \quad k = 0, 1, \dots \quad (3.104)$$

Sei t der Multiplikationsoperator mit der Funktion $f(t) = t$. Dann gilt wegen der Leibniz-Regel, dass

$$\left[\frac{d}{dt} - \lambda \mathbf{E}, t \right] = \mathbf{E} \quad (3.105)$$

und daher

$$\left[\frac{d}{dt} - \lambda \mathbf{E}, t^k \right] = k t^{k-1}, \quad k = 1, 2, \dots \quad (3.106)$$

ist. Folglich ist

$$\begin{aligned} \left(\frac{d}{dt} - \lambda \mathbf{E}\right)^{k+1} t^k e^{\lambda t} &= \left(\frac{d}{dt} - \lambda \mathbf{E}\right)^k \left(\frac{d}{dt} - \lambda \mathbf{E}\right) t^k e^{\lambda t} \\ &= \left(\frac{d}{dt} - \lambda \mathbf{E}\right)^k \left(\left[t^k \left(\frac{d}{dt} - \lambda \mathbf{E}\right) + k t^{k-1} \right] e^{\lambda t} \right) \\ &= k \left(\frac{d}{dt} - \lambda \mathbf{E}\right)^k t^{k-1} e^{\lambda t} \end{aligned}$$

Daraus folgt (3.104) unmittelbar. Es bleibt zu zeigen, dass die Funktionen im obigen Satz linear unabhängig über \mathbb{R} sind, was dasselbe ist, wie dass die Funktionen $t^k e^{\lambda_i t}$ mit $k = 0, 1, \dots, r_i - 1$, wobei λ_i reell oder komplex ist, linear über \mathbb{C} sind. Sei nun Q_i jenes Polynom, das aus Q dadurch hervorgeht, dass der Faktor $(\lambda - \lambda_i)^{r_i}$ durch $(\lambda - \lambda_i)^{r_i-1}$ ersetzt wird und R_i jenes Polynom, das aus Q hervorgeht, indem der Faktor $(\lambda - \lambda_i)^{r_i}$ weggelassen wird. Dann gilt, dass

$$Q_i \left(\frac{d}{dt}\right) t^k e^{\lambda_j t} = \begin{cases} 0 & \text{wenn } i \neq j, 0 \leq k < r_i \\ 0 & \text{wenn } i = j, 0 \leq k < r_i - 1 \\ R_i(\lambda_j) k! e^{\lambda_j t} & \text{wenn } i = j, k = r_i - 1, \end{cases}$$

und die 3-te Zeile hierin ist ungleich Null. Daher folgt aus

$$\sum_i \sum_k^{r_i-1} c_{ik} t^k e^{\lambda_i t} = 0 \quad (3.107)$$

durch Anwendung von $Q_i \left(\frac{d}{dt}\right)$, dass $c_{i r_i-1} = 0$ ist. In ähnlicher Weise zeigt man, dass auch alle übrigen c_{ik} gleich Null sind. Ende der Beweisskizze.

Beispiel: Wir betrachten die Gleichung

$$x^{(4)} + 4x^{(3)} + 5x^{(2)} + 4\dot{x} + 4x = 0 \quad (3.108)$$

Das charakteristische Polynom hat die Gestalt $Q(\lambda) = \lambda^4 + 4\lambda^3 + 5\lambda^2 + 4\lambda + 4$ mit Nullstellen -2 , i und $-i$. Nun ist $Q'(-2) = 4(-2)^3 + 12(-2)^2 + 10(-2) + 4 = 0$. Also ist -2 eine doppelte Nullstelle und $Q(\lambda) = (\lambda + 2)^2(\lambda - i)(\lambda + i)$. Die allgemeine Lösung von (3.108) ist also

$$x(t) = A e^{-2t} + B t e^{-2t} + C \cos t + D \sin t \quad (3.109)$$

für Konstanten A, B, C, D . Die Lösung eines AW-Problems mit vorgegebenen $x(0), \dot{x}(0), x^{(2)}(0), x^{(3)}(0)$ ergibt ein lineares Gleichungssystem für A, B, C, D . Aus der obigen Theorie folgt, dass dieses Gleichungssystem eindeutig lösbar ist.

3.2.7 Exakte Differentialgleichungen, Differentialformen

Wir behandeln hier, wegen des geometrischen Interesses und als Vorbereitung auf die Funktionentheorie, ausnahmsweise einen speziellen Typ von nichtautonomen ODE's für $n = 1$, die durch Quadraturen, also explizite Integrale, gelöst werden können. Wir gehen aus von

$$\dot{x} = \frac{dx}{dt} = \frac{P(x, t)}{Q(x, t)} \quad (3.110)$$

Eine Differentialform 1-ten Grades (oder "1-Form" oder "Pfaff'sche Form") im \mathbb{R}^2 ist ein Ausdruck der Art

$$\omega = A(x, y) dx + B(x, y) dy \quad (3.111)$$

Sei γ eine Kurve im \mathbb{R}^2 , gegeben durch $x = x(s), y = y(s)$. Dann ist die auf γ eingeschränkte Differentialform $\omega|_\gamma$ gegeben durch

$$\omega|_\gamma = [A(x(s), y(s)) x' + B(x(s), y(s)) y'] ds \quad (3.112)$$

Man könnte sagen, der Ausdruck in eckigen Klammern in (3.112) ist das innere Produkt, entlang der Kurve γ , von ω mit dem Tangentenvektor $v = x' \partial_x + y' \partial_y$ an die Kurve γ . Das Integral von ω über eine Kurve γ ("Linienintegral"), die z.B. die Punkte p_1 und p_2 verbindet ist definiert als das Riemann-Integral

$$\int_\gamma \omega = \int_{s_1}^{s_2} \omega|_\gamma \quad (3.113)$$

wobei $\gamma(s_1) = p_1$ und $\gamma(s_2) = p_2$ ist. Man rechnet leicht nach, dass das Linienintegral einer 1-Form über eine Kurve von der Parametrisierung dieser Kurve unabhängig ist und bei Änderung der Orientierung das Vorzeichen wechselt. Insbesondere kann man Integrale über geschlossene Kurven betrachten und schreibt dann $\oint_\gamma \omega$. Wenn eine geschlossene Kurve γ einen Bereich \mathcal{D} umrandet, gilt (Satz von Stokes)

$$\oint_\gamma \omega = \oint_\gamma (A dx + B dy) = \int_{\mathcal{D}} (\partial_x B - \partial_y A) dx dy \quad (3.114)$$

Eine Differentialform ω heisst geschlossen, wenn der Integrand des Flächenintegrals auf der r.S. von (3.114) verschwindet, also

$$\partial_x B = \partial_y A \quad (3.115)$$

Eine 1-Form ω heisst exakt, wenn es eine skalare Funktion $F(x, y)$ gibt, sodass

$$A(x, y) = \partial_x F(x, y), \quad B(x, y) = \partial_y F(x, y) \quad (3.116)$$

Man schreibt (3.116) auch in der Form

$$\omega = dF = \partial_x F dx + \partial_y F dy \quad (3.117)$$

Die geometrische Bedeutung von (3.117) ist diese: für jede Kurve, die etwa Punkte p_1 und p_2 verbindet, hängt das Integral $\int_\gamma \omega$ nur von den Endpunkten ab, denn $\int_\gamma \omega = F(p_2) - F(p_1)$. Insbesondere ist $\oint \omega = 0$ für alle geschlossenen Kurven. Wenn ω auf ganz \mathbb{R}^2 definiert ist, folgt aus der Geschlossenheit die Exaktheit, und zwar gibt die folgende Formel eine explizite Konstruktion für F bei gegebenem ω . Sei ω im ganzen \mathbb{R}^2 geschlossen¹¹. Dann definieren wir

$$F(x, y) = \int_0^1 [A(tx, ty)x + B(tx, ty)y] dt \quad (3.118)$$

Beweis: Es gilt, dass

$$\partial_x F(x, y) = \int_0^1 A(tx, ty) dt + \int_0^1 t[x \partial_x A(tx, ty) + y \partial_x B(tx, ty)] dt \quad (3.119)$$

Unter Verwendung von (3.115) schreiben sich 2.ter und 3.ter Term in (3.119) als

$$\int_0^1 t[x \partial_x A(tx, ty) + y \partial_y A(tx, ty)] dt = \int_0^1 t \frac{d}{dt} A(tx, ty) dt = [tA(tx, ty)]_0^1 - \int_0^1 A(tx, ty) dt \quad (3.120)$$

Einsetzen von (3.120) in (3.119) ergibt die erste Gleichung in (3.116). Der Beweis der zweiten Gleichung in (3.116) ist analog.

Oft ist aber ω nur auf einer offenen, zusammenhängenden Menge $\mathcal{D} \subset \mathbb{R}^2$ definiert, und dann ist die Situation komplizierter. Zunächst kann man zeigen, dass die Bedingung, es möge $\int_\gamma \omega = 0$ für alle geschlossenen Kurven γ sein, auch hinreichend für die Exaktheit einer geschlossenen 1-Form ω ist. Wenn nun eine geschlossene Kurve γ einen offenen Bereich berandet, in dem ω geschlossen ist, ist das offenbar der Fall, wie der Satz von Stokes zeigt. Es kann aber sein, dass ω in einem Bereich \mathcal{D} definiert ist, in dem nicht alle geschlossenen Kurven ein Gebiet beranden. Es folgt ein typisches Beispiel für dieses Phänomen. Sei $\Omega = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ und $\omega = -\frac{y}{r^2} dx + \frac{x}{r^2} dy$, wobei $r^2 = x^2 + y^2$. Die 1-Form ω ist geschlossen, denn $\partial_y(-\frac{y}{r^2}) = \frac{-x^2 + y^2}{r^4} = \partial_x(\frac{x}{r^2})$. Aus dem Satz von Stokes folgt nun, dass $\oint_\gamma \omega$ gleich Null ist für jede geschlossene Kurve, die den Nullpunkt nicht umschließt. Sei γ aber ein Kreis um den Ursprung mit Radius R , der im Gegenuhrzeigersinn durchlaufen wird, also $x = R \cos s$, $y = R \sin s$ mit $0 \leq s \leq 2\pi$. Dann ist

$$\oint_\gamma \omega = \int_0^{2\pi} \frac{1}{R^2} (\cos^2 s + \sin^2 s) R^2 ds = 2\pi \quad (3.121)$$

Aus dem Satz von Stokes folgt, dass, nicht nur für den obigen Kreis, sondern für alle gleich orientierten geschlossenen Kurven, die den Nullpunkt einmal umschließen, das Integral $\oint_\gamma \omega$ den Wert 2π hat.

¹¹Das folgende Argument erfordert genaugenommen lediglich, dass Ω sternförmig bezüglich des Nullpunkts ist.

Schliesslich betrachten wir das Beispiel einer geschlossenen 1-Form ω , die ebenfalls nur in $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ definiert ist, aber trotzdem exakt, und zwar $\omega = \frac{1}{x^2+y^2}(x dx + y dy)$. Diese 1-Form erfüllt $\omega = dF$ mit $F = \frac{1}{2} \log(x^2 + y^2)$.

Wir kehren zur Gleichung (3.110) zurück. Angenommen, es ist

$$\partial_t Q + \partial_x P = 0 \quad (3.122)$$

Dann ist die 1-Form gegeben durch

$$\omega = Q dx - P dt \quad (3.123)$$

exakt. Nun gehört zu jeder Funktion $x(t)$, die die Gleichung (3.110) löst, eine Kurve γ im \mathbb{R}^2 mit $(x(s), t(s))$ mit $x'(s) = \dot{x}(t(s)) t'(s)$, sodass $\omega|_\gamma = 0$ ist. Diese Bedingung ist wiederum äquivalent mit der Forderung, dass γ eine Niveaulinie der Funktion F ist¹².

Die Funktionen P, Q in der Gleichung (3.110) sind natürlich nicht eindeutig. Man kann versuchen, beide mit derselben Funktion $R(x, t)$ ("integrierender Faktor") zu multiplizieren, sodass $P' = R P$ und $Q' = R Q$ die Bedingung (3.122) erfüllen. Es zeigt sich, dass das (in hinreichend kleinen Umgebungen im (x, t) - Raum) immer möglich ist. Dummerweise ist aber das Auffinden eines integrierenden Faktors im allgemeinen gleich schwer wie das ursprüngliche Problem, nämlich die Gleichung (3.110) zu lösen.

¹²Die Funktion $-F$ ist nichts anderes als eine Hamilton-Funktion des autonomen Systems $x' = P, t' = Q$

Kapitel 4

Komplexe Analysis

4.1 Komplexe Zahlen

Elemente $z \in \mathbb{C}$ sind Paare reeller Zahlen, geschrieben als $z = a + ib$, wobei $a, b \in \mathbb{R}$ sind. Wir schreiben $a = \Re z$ und $b = \Im z$. Komplexe Zahlen werden nach der Regel $z + z' = a + a' + i(b + b')$ addiert. Die Menge \mathbb{C} ist also als Vektorraum identisch mit \mathbb{R}^2 . Sie hat aber weiters die Möglichkeit einer Multiplikation, die durch die Regel $zz' = aa' - bb' + i(ab' + ba')$ erklärt ist. Multiplikation und Addition genügen den üblichen Rechenregeln. Weiters gilt, dass $(0+i)^2 = -1+i0$ ist, kurz: $i^2 = -1$. Jedes Element $z \neq 0$ hat genau ein Inverses z^{-1} , das durch $z^{-1}z = 1$ definiert ist. Explizit erhalten wir

$$z^{-1} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2} \quad (4.1)$$

Die komplex Konjugierte \bar{z} einer Zahl $z = a + ib$ ist definiert als $\bar{z} = a - ib$. Die Norm $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$ stimmt mit der euklidischen Norm im \mathbb{R}^2 überein. Insbesondere gilt, dass

$$|z + z'| \leq |z| + |z'| \quad (4.2)$$

Ausserdem ist

$$|zz'| = |z||z'| \quad (4.3)$$

4.2 Potenzreihen

Eine formale Potenzreihe ist ein Ausdruck $f(X)$ der Art $f(X) = \sum_{n \geq 0} a_n X^n$, wobei die Variable X und die Koeffizienten a_n komplexe oder reelle Zahlen sind. Potenzreihen können addiert und miteinander multipliziert werden. Z.B. sind die Koeffizienten c_n der Potenzreihe $(fg)(X)$, wobei $f(X) = \sum_{n \geq 0} a_n X^n$ und $g(X) = \sum_{n \geq 0} b_n X^n$ ist, gleich der endlichen Summe $c_n = \sum_{k \geq 0, l \geq 0, k+l=n} a_k b_l$. Man kann formale Potenzreihen auch ineinander einsetzen, also die Reihe $f(g(X))$ bilden. Wenn in $f(X) = \sum_{n \geq 0} a_n X^n$ die Koeffizienten $a_0 = 0$ und $a_1 \neq 0$

sind, kann man eine eindeutige formale Potenzreihe für $f^{-1}(X)$ definieren, wobei $f^{-1}(f(X)) = X$ ist. Die Ableitung $f'(X)$ einer formalen Potenzreihe ist natürlich gegeben durch $f'(x) = \sum_{n \geq 1} n a_n X^{n-1}$.

Für den Begriff der Konvergenz und seine Verschärfungen (absolute Konvergenz, gleichmäßige Konvergenz) müssen wir auf die Vorlesung Analysis 1 verweisen. Sei nun $\sum_{n \geq 0} a_n X^n$ eine Potenzreihe, die für $X_0 \neq 0$ konvergiert. Dann konvergiert $\sum_{n \geq 0} a_n X^n$ absolut für jedes x mit $|X| < |X_0|$. Es gibt eine reelle Zahl, genannt Konvergenzradius, $\rho \geq 0$, sodass $\sum_{n \geq 0} a_n X^n$ für $|X| < \rho$ absolut konvergiert und für $|X| > \rho$ divergiert.

Eine wichtige Formel für den Konvergenzradius ist

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} \quad (4.4)$$

Wir bemerken, dass ρ gleich Null oder gleich ∞ sein kann. Summe und Produkt zweier Potenzreihen mit Konvergenzradius $\geq \rho$ hat ebenfalls Konvergenzradius $\geq \rho$. Die Ableitung $f'(X)$ einer Potenzreihe $f(X)$ hat denselben Konvergenzradius wie die ursprüngliche Potenzreihe und es gilt, dass $\lim_{h \rightarrow 0} \frac{f(X+h) - f(X)}{h} = f'(X)$. Insbesondere ist eine Potenzreihe innerhalb ihres Konvergenzkreises (bzw. Intervalls) unendlich oft differenzierbar. Die Koeffizienten einer konvergenten Potenzreihe bestimmen sich durch die fundamentale Formel

$$a_n = \frac{1}{n!} f^{(n)}(0) \quad (4.5)$$

Es folgt, dass eine Funktion $f(X)$ höchstens eine formale Potenzreihe mit Konvergenzradius $\rho \neq 0$ hat. Für gegebene Funktion $f(X)$ ist die Reihe $\sum_{n \geq 0} a_n X^n$ daher die Taylorreihe bezüglich des Ursprungs.

Beispiele: Die Potenzreihe $\sum_{n \geq 0} X^n$ konvergiert für $|X| < 1$ und divergiert für $|X| \geq 1$. Es gilt $\frac{1}{1-X} = \sum_{n \geq 0} X^n$.

4.3 Exp und Log

Die komplexe Exponentialfunktion ist definiert durch die Reihenentwicklung

$$e^z = \sum_{n \geq 0} \frac{1}{n!} z^n \quad (4.6)$$

Sie hat Konvergenzradius $\rho = \infty$. Ihre Ableitung ist $\frac{d}{dz} e^z = e^z$. Sie erfüllt, auch im Komplexen, die fundamentale Relation

$$e^z e^{z'} = e^{z+z'} \quad (4.7)$$

Weiters gilt, wenn y reell ist, dass $\overline{e^{iy}} = e^{-iy}$ und daher, wegen (4.7),

$$|e^{iy}| = 1 \tag{4.8}$$

ist. Daraus folgt, dass $e^{x+iy} = e^x e^{iy}$ nirgends $= 0$ ist. Man sieht weiters, dass

$$e^{iy} = \cos y + i \sin y \tag{4.9}$$

und $e^{iy} = 1 \Leftrightarrow y = 2k\pi$ mit $k \in \mathbb{Z}$. Schliesslich zeigt man, dass jede Zahl $z \in \mathbb{C}$ mit $|z| = 1$ in der Form $z = e^{iy}$ geschrieben werden kann. Die reelle Zahl y ist eindeutig bis auf Addition eines ganzzahligen Vielfachen von 2π . Mithin kann jede komplexe Zahl in "Polarform" geschrieben werden, also

$$z = r e^{i\phi} . \tag{4.10}$$

Hier ist offensichtlich $r = |z|$. Der "Winkel" ϕ heisst Argument von z , also $\phi = \operatorname{arg} z$.

Der komplexe Logarithmus $\log z$ ist für $z \neq 0$ definiert durch

$$\log z = \log |z| + i \operatorname{arg} z \tag{4.11}$$

Der Logarithmus auf der rechten Seite von (4.11) ist der reelle Logarithmus. Offenbar ist $e^{\log z} = z$ und modulo 2π gilt, dass $\log z z' = \log z + \log z'$. Wegen der Mehrdeutigkeit von $\operatorname{arg} z$ macht (4.11) ohne weitere Einschränkungen keinen Sinn als Funktion von z . Eine Möglichkeit dies zu tun ist, die r.S. von (4.11) auf die rechte Halbebene $\Re z > 0$ und $\operatorname{arg} z = \phi$ auf $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$ einzuschränken. Die dadurch definierte stetige Funktion ist der sog. Hauptzweig von $\log z$. Man kann $\operatorname{arg} z$ auch stetig auf die entlang der negativen reellen Achse inkl. des Ursprungs "aufgeschlitzte" komplexe Ebene ausdehnen. Der Bildbereich von $\operatorname{arg} z$ ist dann $-\pi < \operatorname{arg} z < \pi$. Allgemein kann man eine beliebige radiale Halbgerade zum Winkel ψ , $0 \leq \psi < 2\pi$ inkl. des Ursprungs entfernen und auf dem verbleibenden Gebiet der Funktion $\operatorname{arg} z$ die Werte $\psi - 2\pi < \operatorname{arg} z < \psi$ (modulo 2π) geben. Jede so definierte stetige Funktion $\operatorname{arg} z$ definiert einen Zweig von $\operatorname{arg} z$ und daher von $\log z$.

Wenn z und α komplexe Zahlen sind, definieren wir

$$z^\alpha = e^{\alpha \log z} \tag{4.12}$$

Jeder Zweig von $\operatorname{arg} z$ definiert einen Zweig von z^α .

4.4 Analytische Funktionen

Eine in einer offenen Menge \mathcal{D} in \mathbb{R} oder \mathbb{C} definierte Funktion $f(x)$ heisst in \mathcal{D} analytisch, wenn ihre Taylorreihe bez. jeden Punkts von \mathcal{D} konvergiert. Genauer: es gibt für jeden Punkt $x_0 \in \mathcal{D}$ eine Potenzreihe $\sum_{n \geq 0} a_n X^n$ mit positivem

Konvergenzradius, sodass

$$f(x) = \sum_{n \geq 0} a_n (x - x_0)^n \quad \text{mit } |x - x_0| \text{ hinreichend klein} \quad (4.13)$$

ist. Fundamental ist der

Satz: Sei $f(x) = \sum_{n \geq 0} a_n x^n$ eine konvergente Potenzreihe mit $\rho > 0$. Dann ist $f(x)$ analytisch in der Kreisscheibe (im Intervall) $|x| < \rho$.

Die Pointe dieses Satzes ist, dass sich $f(x)$ bezüglich jeden Punktes im Konvergenzkreis in eine Taylorreihe entwickeln lässt. Weiters zeigt sich, dass der Konvergenzradius dieser Potenzreihe $\geq \rho - |x_0|$ ist, also so gross, dass die zugehörige Konvergenzscheibe mindestens in die Scheibe $|x| < \rho$ hineinpasst.

Analytische Funktionen sind unendlich oft differenzierbar. Folgende beiden Eigenschaften analytischer Funktionen gelten i.a. nicht für unendlich oft differenzierbare Funktionen.

Satz: Wenn zwei in einer offenen zusammenhängenden Menge analytische Funktionen in der Umgebung eines Punktes $\in \mathcal{D}$ übereinstimmen, dann stimmen sie in ganz \mathcal{D} überein. Weiters gilt der

Satz: Die Nullstellenmenge einer in einer offenen zusammenhängenden Menge \mathcal{D} analytischen Funktion f ist diskret, es sei denn f verschwindet identisch.

Physikalische Zwischenbemerkung: Nach dem vorletzten Satz ist es insbesondere für eine analytische Funktion unmöglich, ausserhalb einer kompakten Menge zu verschwinden, ohne identisch = 0 zu sein. Wenn wir physikalische Zustände grob in dynamische und Gleichgewichtszustände einteilen, ist plausibel, dass die ersteren i.a. nicht durch analytische Funktionen beschrieben werden können. Nun werden, von der Punktmechanik abgesehen, physikalische Zustände in der Physik in der Regel durch partielle Differentialgleichungen charakterisiert, und zwar dynamische in der Regel durch hyperbolische Gleichungen¹ und Gleichgewichtszustände oft durch elliptische Gleichungen. Und in der Tat kann man beweisen, dass Lösungen elliptischer Gleichungen (unter sehr allgemeinen Annahmen) analytisch sind, solche hyperbolischer Gleichungen i.a. nicht.

4.5 Holomorphie

Sei $f(z)$ eine Funktion der komplexen Variablen z und γ eine Kurve in \mathbb{C} . Wir können, das Kurvenintegral $\int_{\gamma} f(z) dz$ betrachten, wobei $dz = dx + i dy$ gesetzt ist. Wir können f als Funktion der reellen Variablen x und y auffassen, und diese wieder durch z und \bar{z} ausdrücken gemäss

$$dz = dx + i dy, \quad d\bar{z} = dx - i dy. \quad (4.14)$$

¹Die Schrödingergleichung der Quantenmechanik nimmt hier eine Sonderstellung ein.

Also gilt

$$dx = \frac{1}{2}(dz + d\bar{z}) \quad dy = \frac{1}{2i}(dz - d\bar{z}) \quad (4.15)$$

Folglich ist

$$df = \partial_x f dx + \partial_y f dy = \partial_z f dz + \partial_{\bar{z}} f d\bar{z}, \quad (4.16)$$

wobei

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y) \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y) \quad (4.17)$$

Wenn wir andererseits $f = P + iQ$ setzen, ist

$$\omega = f dz = P dx - Q dy + i(Q dx + P dy) \quad (4.18)$$

Die Geschlossenheit der Differentialform ω ist äquivalent mit

$$\partial_y P = -\partial_x Q \quad \partial_y Q = \partial_x P \quad (4.19)$$

und diese wieder mit

$$\partial_{\bar{z}} f = 0 \Leftrightarrow \partial_x f + i\partial_y f = 0 \quad (4.20)$$

Die Gleichungen (4.19) heißen Cauchy-Riemannsche Differentialgleichungen.

Wir erwähnen den wichtigen, leicht nachprüfbaren Umstand, dass die reellen Funktionen P und Q harmonisch sind, also der 2-dimensionalen Laplacegleichung genügen.

Satz: Sei $f(z)$ eine Funktion der komplexen Variablen z in einem offenen Gebiet $\mathcal{D} \subset \mathbb{C}$. Dann sind folgende zwei Bedingungen gleichwertig: (i) f ist eine differenzierbare Funktion der reellen Variablen x und y und erfüllt überdies die Bedingung (4.20). (ii) Die Funktion f ist "komplex differenzierbar" oder holomorph, d.h. der Limes

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(z+h) - f(z)}{h} \text{ existiert}$$

für alle $z \in \mathcal{D}$, wobei h eine variable komplexe Zahl ist. Wir sehen also, dass vom reellen Standpunkt die komplexe Differenzierbarkeit zusätzlich zur reellen Differenzierbarkeit die Integrabilitätsbedingungen (4.20), also die Geschlossenheit der Differentialform ω , beinhaltet. Wichtiges Beispiel einer überall reell differenzierbaren, aber nirgends holomorphen Funktion ist $F(z) = \bar{z}$.

Die Bedingungen (4.20) bewirken den

Satz: Sei $f(z)$ eine holomorphe Funktion in einem offenen Gebiet \mathcal{D} . Dann ist $\omega = f(z) dz$ geschlossen.

Umgekehrt gilt der

Satz: Sei $f(z)$ in \mathcal{D} stetig und $\omega = f(z) dz$ geschlossen. Dann ist $f(z)$ in \mathcal{D} holomorph.

Weiters folgt aus der Stokes'schen Formel der

Satz: Wenn $f(z)$ in \mathcal{D} holomorph ist, gilt, dass $\oint_{\gamma} f(z) dz = 0$ für jeden geschlossenen Weg γ , der sich stetig auf einen Punkt zusammenziehen lässt.

Wir haben gesehen, dass die Holomorphie von $f(z)$ äquivalent zur Geschlossenheit von $f(z)dz$ ist. Was ist die komplexe Version des Begriffs der Exaktheit? Es ist die Existenz einer Stammfunktion $F(z)$, also einer komplexen Funktion, sodass $f(z)dz = dF = F'(z)dz$. Diese Stammfunktion ist ebenfalls holomorph. Und die Forderung der Exaktheit von ω ist wieder äquivalent zur Bedingung, dass $\oint_{\gamma} f(z)dz = 0$ ist für jeden geschlossenen Weg im betrachteten Gebiet. In einfach zusammenhängenden Gebieten von \mathbb{C} , also in solchen, wo jeder geschlossene Weg eine offene Menge berandet, ist somit die Holomorphie äquivalent mit der Existenz einer Stammfunktion.

Beispiel 1: Die Funktionen $f(z) = z^m$, $m \geq 0$ sind holomorph in ganz \mathbb{C} . Da \mathbb{C} einfach zusammenhängend ist, besitzt sie auch eine Stammfunktion. Diese ist natürlich $F(z) = \frac{1}{m+1}z^{m+1}$.

Beispiel 2: Die Funktion $f(z) = \frac{1}{z}$ ist offenbar holomorph in $\mathbb{C} \setminus \{0\}$. Dann ist $\oint_{\gamma} f(z) dz = 0$ für jeden geschlossenen Weg, der den Ursprung nicht umschließt. Wenn andererseits γ den Ursprung einmal umrundet, gilt

$$\oint_{\gamma} \frac{dz}{z} = 2\pi i \quad (4.21)$$

Dies folgt entweder aus **(3.2.7)** oder durch $z = re^{i\phi}$ und $dz = ir e^{i\phi} d\phi$, sodass

$$\oint_{\gamma} \frac{dz}{z} = i \int_0^{2\pi} d\phi = 2\pi i \quad (4.22)$$

Beispiel 2': Die (eine) Stammfunktion des komplexen Logarithmus. Hier betrachten wir das Kurvenintegral

$$\int_{\gamma_{r,\phi}} \frac{dz}{z} \quad (4.23)$$

Der Weg $\gamma_{r,\phi}$ verbindet den Punkt $z = 1$ mit dem Punkt (r, ϕ) , wobei $-\pi < \phi < \pi$ und $r > 0$ ist. Und zwar geht der Weg zuerst entlang der positiven reellen Achse und dann entlang des Kreises mit Radius r (wenn $y > 0$ im Gegenuhrzeigersinn, ansonsten im Uhrzeigersinn). Das Ergebnis ist

$$\int_{\gamma_{r,\phi}} \frac{dz}{z} = \log r + i\phi \quad (4.24)$$

Dies ist gerade der Hauptzweig des Logarithmus.

Beispiel 3: Die Funktionen $f(z) = z^m$ für $m < -1$ sind holomorph in $\mathbb{C} \setminus \{0\}$, aber nicht in \mathbb{C} . Sie haben aber trotzdem eine Stammfunktion, und zwar $F(z) = \frac{1}{m+1}z^{m+1}$. Man kann auch explizit nachrechnen, dass $\oint_{\gamma} f(z) dz = 0$ ist für einen Kreis mit Radius R um den Ursprung, denn:

$$\int_{\gamma} z^m dz = iR^{m+1} \int_0^{2\pi} e^{i(m+1)\phi} d\phi = 0, \quad m \neq -1 \quad (4.25)$$

Für das Folgende nützlich ist dieser Sachverhalt: sei $f(z)$ holomorph in \mathcal{D} mit $0 \in \mathcal{D}$ und $f(0) = 0$. Dann ist die Funktion $g(z)$, definiert durch $g(z) = \frac{1}{z}f(z)$ für $z \neq 0$ und $g(0) = f'(0)$, holomorph in \mathcal{D} . Fundamental ist die Cauchy'sche Integralformel.

Satz: Sei $f(z)$ in \mathcal{D} holomorph und γ ein geschlossener Weg, der den Punkt $z_0 \in \mathcal{D}$ einfach umschliesst. Dann ist

$$\oint_{\gamma} \frac{f(z)}{z - z_0} dz = 2\pi i f(z_0) \quad (4.26)$$

Beweisskizze: Man schreibt $\frac{f(z)}{z - z_0} = \frac{f(z) - f(z_0)}{z - z_0} + \frac{f(z_0)}{z - z_0}$ und benutzt, dass der erste Summand auf der r.S. holomorph ist.

Anwendung: Wir beweisen die Formel ($a, b > 0$)

$$\int_0^{2\pi} \frac{dt}{a^2 \cos^2 t + b^2 \sin^2 t} = \frac{2\pi}{ab} \quad (4.27)$$

Dazu gehen wir aus von (4.21), wobei für γ der Weg mit einer Ellipse als Bild gewählt wird, mit grosser bzw. kleiner Halbachse a bzw. b und Mittelpunkt im Ursprung. Also ist γ gegeben durch $z(t) = a \cos t + i b \sin t$ mit $0 \leq t \leq 2\pi$ und

$$2\pi i = \oint_{\gamma} \frac{dz}{z} = \int_0^{2\pi} \frac{(-a^2 + b^2) \cos t \sin t + i ab (\sin^2 t + \cos^2 t)}{a^2 \cos^2 t + b^2 \sin^2 t} dt, \quad (4.28)$$

woraus sich (4.27) ergibt.

Eine wichtige Konsequenz der Cauchy'schen Integralformel ist der

Satz: Sei die Funktion $f(z)$ in der offenen Kreisscheibe $|z| < R$ holomorph. Dann hat sie in dieser Kreisscheibe eine konvergente Taylorentwicklung $f(z) = \sum_{n \geq 0} a_n z^n$ mit Konvergenzradius $\geq R$.

Man kann die Koeffizienten a_n von $f(z)$ ausser durch (4.5) auch folgendermassen darstellen. Im Ausdruck

$$f(z) = \frac{1}{2\pi i} \oint \frac{f(z') dz'}{z' - z} \quad \text{wenn } |z| < R \quad (4.29)$$

entwickeln wir, zunächst nur für $|z| < |z'|$, $\frac{1}{z' - z} = \frac{1}{z'} (1 + \frac{z}{z'} + \dots + \frac{z^n}{z'^{n+1}} + \dots)$. Daraus folgt, dass

$$a_n = \frac{1}{2\pi i} \oint \frac{f(z') dz'}{z'^{n+1}} \quad n \geq 0 \quad (4.30)$$

Wir bemerken, dass umgekehrt jede analytische Funktion komplex differenzierbar, also holomorph ist. Da aber nach dem vorangehenden Satz jede holomorphe Funktion in einer ganz im Holomorphiegebiet liegenden Kreisscheibe eine konvergente Taylorreihe bezüglich des Ursprungs hat, ergibt sich für unsere frühere Diskussion des Analytizitätsbegriffs folgende Verschärfung: Sei $f(z)$ in einem offenen Gebiet \mathcal{D} analytisch und $z_0 \in \mathcal{D}$. Dann konvergiert die Taylorreihe bezüglich

z_0 in jeder Kreisscheibe um z_0 , die ganz in \mathcal{D} liegt. Diese Aussage wäre im Reellen falsch. Dazu betrachte man die Funktion $f(x) = \frac{1}{1+x^2}$. Diese ist auf ganz \mathbb{R} analytisch. Aber der Konvergenzradius ihrer Taylorreihe um den Ursprung ist nur $= 1$. Der "Grund" hierfür ist, dass die "komplex fortgesetzte" Funktion $f(z) = \frac{1}{1+z^2}$ in $z = \pm i$ singularär ist.

4.6 Residuensatz

Wenn die offene Kreisscheibe im vorangehenden Satz durch einen Kreisring ersetzt wird, ist die Situation anders. Vorbereitend erwähnen wir folgenden Sachverhalt: Man betrachte die formale Potenzreihe $\sum_{n<0} a_n X^n = \sum_{n>0} a_{-n} X^{-n}$. Wenn diese Reihe den Konvergenzradius $\frac{1}{\rho}$ hat, definiert sie eine für $|z| > \frac{1}{\rho}$ holomorphe Funktion.

Satz: Sei $f(z)$ im Kreisring $R_1 < |z| < R_2$ holomorph, wobei die Fälle $R_2 = 0$ und $R_1 = \infty$ zugelassen sind. Dann lässt dort $f(z)$ sich in eine konvergente "Laurentreihe" entwickeln, also eine Reihe der Art

$$f(z) = \sum_{-\infty < n < \infty} a_n z^n, \quad (4.31)$$

wobei die Summe in (4.31) nunmehr über alle ganzen Zahlen n geht. Wieder gilt, dass

$$a_n = \frac{1}{2\pi i} \oint \frac{f(z) dz}{z^{n+1}}, \quad (4.32)$$

wobei n beliebig in \mathbb{Z} ist. Aus der Formel (4.32) folgt insbesondere, dass

$$\frac{1}{2\pi i} \oint \frac{z^m}{z^{n+1}} dz = \delta_{mn} \quad (4.33)$$

ist. Das wussten wir allerdings bereits (siehe (4.29,4.30)).

Jedenfalls folgt für jede im Kreisring holomorphe Funktion, dass

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz = a_{-1}, \quad (4.34)$$

wenn γ den Kreisring einmal positiv durchläuft. Diese Formel wird meist angewandt, wenn $R_1 = 0$ ist, also f in der punktierten Kreisscheibe definiert ist. In diesem Fall nennt man den Koeffizienten a_{-1} das Residuum der Funktion $f(z)$ im Ursprung, kurz: $\text{Res}(f, 0)$ und die Formel (4.34) die einfachste Version des Residuensatzes.

Wenn die Funktion $f(z)$ sich in die volle Kreisscheibe fortsetzen lässt - und genau dann - sind alle negativen Koeffizienten in der Taylorreihe (und daher das Residuum) gleich $= 0$. Wenn endlich viele negative Koeffizienten ungleich 0 sind,

spricht man vom Ursprung als einem Pol endlicher Ordnung k für die Funktion $f(z)$, wobei $k \in \mathbb{N}$ so ist, dass $g(z) = z^k f(z)$ sich holomorph in die volle Kreisscheibe mit $g(0) \neq 0$ fortsetzen lässt.

Oft ist es nützlich, statt einer offenen Kreisscheibe das Komplement einer solchen, also $|z| > R$ für hinreichend grosses R zu betrachten und in diesem, statt $f(z)$, die in der im Ursprung punktierten Kreisscheibe $|z'| < \frac{1}{R}$ definierte Funktion $-\frac{1}{z'^2} f(\frac{1}{z'})$. Wenn $f(z)$ in $|z| > R$ holomorph ist, hat es dort eine konvergente Laurent-Entwicklung $\sum a_n z^n$. Daher hat $-\frac{1}{z'^2} f(\frac{1}{z'})$ eine konvergente Laurent-Entwicklung in $|z'| < \frac{1}{R}$, und zwar mit Residuum gleich $-a_{-1}$. Dieses heisst Residuum von $f(z)$ im unendlich fernen Punkt $z = \infty$.

Bemerkung: Das Konzept des unendlich fernen Punktes ist mehr als nur formal. Man kann den Raum, der aus \mathbb{C} durch Hinzunahme des unendlich fernen Punktes entsteht, mit der Riemannschen Zahlenkugel \mathbb{S}^2 identifizieren (siehe Übungsaufgabe 51).

Angenommen, die Funktion $f(z)$ sein in einer offenen Teilmenge \mathcal{D} holomorph, ausgenommen endlich vieler Punkte, den singulären Stellen. Dann gilt der Residuensatz:

$$\oint_{\gamma} f(z) dz = 2\pi i \sum \text{Res}(f, z_i), \quad (4.35)$$

wobei γ der (orientierte) Rand von \mathcal{D} ist und z_i die innerhalb \mathcal{D} gelegenen singulären Punkte von f . Der Satz gilt gleichlautend, wenn \mathcal{D} den Punkt $z = \infty$ enthält.

4.7 Anwendungen

Sei $f(z) = \frac{g(z)}{(z-z_0)^k}$, wobei g holomorph mit $g(z_0) \neq 0$ ist. Dann ist $\text{Res}(f, z_0)$, das Residuum von $f(z)$ bei z_0 , gleich dem Koeffizienten von $(z-z_0)^{k-1}$ in der Taylorentwicklung von $g(z)$ im Punkt z_0 .

Sei nun $f(z) = \frac{P(z)}{Q(z)}$ und z_0 eine einfache Polstelle von Q mit $P(z_0) \neq 0$. Dann ist

$$\text{Res}(f, z_0) = \frac{P(z_0)}{Q'(z_0)} \quad (4.36)$$

Beispiel 1: Die Funktion $f(z) = \frac{e^{iz}}{z^2+1} = \frac{e^{iz}}{(z+i)(z-i)}$. Folglich ist $\text{Res}(f, i) = -\frac{i}{2e}$ und $\text{Res}(f, -i) = \frac{ie}{2}$.

Beispiel 2: Die Funktion $f(z) = \frac{e^{iz}}{z(z^2+1)^2}$. Offensichtlich ist $\text{Res}(f, 0) = 1$. Für den doppelten Pol z.B. in $z = i$ definieren wir $g(z) = \frac{e^{iz}}{z(z+i)^2}$. Wir finden $g'(i) = -\frac{3}{4e}$. Nun soll der Residuenkalkül auf die Berechnung einiger Typen bestimmter Integrale verwendet werden.

Typ 1. Integrale der Form

$$I = \int_0^{2\pi} R(\sin t, \cos t) dt, \quad (4.37)$$

wobei $R(x, y)$ eine rationale Funktion ist, die auf dem Kreis $x^2 + y^2 = 1$ keine Pole hat. Wir setzen $z = e^{it}$, sodass z den Einheitskreis beschreibt. Also ist I das $2\pi i$ -fache der Residuensumme der Funktion

$$\frac{1}{iz} R\left(\frac{1}{2i}\left(z - \frac{1}{z}\right), \frac{1}{2}\left(z + \frac{1}{z}\right)\right)$$

Beispiel:

$$I = \int_0^{2\pi} \frac{dt}{a + \sin t}, \quad (4.38)$$

wobei a reell und > 1 ist. Dann gilt

$$I = 2\pi i \sum \operatorname{Res} \frac{2}{z^2 + 2iaz - 1} \quad (4.39)$$

Die Funktion auf der r.S. von (4.39) hat zwei Pole im Endlichen, davon einen im Einheitskreis, und zwar bei $z_0 = -ia + i\sqrt{a^2 - 1}$. Das Residuum an dieser Stelle ist gleich $\frac{1}{i\sqrt{a^2 - 1}}$, also ist $I = \frac{2\pi}{\sqrt{a^2 - 1}}$.

Typ 2: Integrale der Form

$$I = \int_{-\infty}^{+\infty} R(x) dx \quad (4.40)$$

wobei R eine rationale Funktion ist, die keine reellen Pole hat und im Unendlichen mindestens wie $\frac{1}{x^2}$ nach 0 geht. Zunächst betrachten wir das Integral von $R(z)$ über einen in der oberen Halbebene gelegenen Halbkreis γ_S , dessen Radius S so gross gewählt ist, dass alle Polstellen von $R(z)$ innerhalb des (ganzen) Kreises liegen. Dann ist

$$\left| \int_{\gamma_S} f(z) dz \right| \leq \int_{\gamma_S} |f(z)| dz \leq \int_0^{2\pi} \frac{\operatorname{const}}{S^2} d\phi \rightarrow 0 \text{ mit } S \rightarrow \infty \quad (4.41)$$

Nun ist aber

$$\lim_{S \rightarrow \infty} \left[\int_{-S}^{+S} R(x) dx + \int_{\gamma_S} f(z) dz \right] = 2\pi i \sum \operatorname{Res}(R(z)), \quad (4.42)$$

wobei über alle Residuen in der oberen Halbebene summiert wird. Also ist

$$\int_{-\infty}^{+\infty} R(x) dx = 2\pi i \sum \operatorname{Res}(R(z)). \quad (4.43)$$

Analog ist

$$\int_{-\infty}^{+\infty} R(x) dx = -2\pi i \sum \operatorname{Res}(R(z)), \quad (4.44)$$

wobei über alle Residuen in der unteren Halbebene summiert wird.

Beispiel:

$$I = \int_0^{+\infty} \frac{dx}{1+x^6} dx \quad (4.45)$$

Die in der oberen Halbebene gelegenen Polstellen sind die Lösungen $e^{i\frac{\pi}{6}}, e^{3i\frac{\pi}{6}}, e^{5i\frac{\pi}{6}}$ der Gleichung $z^6 = e^{i\pi} = -1$. In jedem Pol ist das Residuum gleich $\frac{1}{6z^5} = -\frac{z}{6}$. Folglich ist

$$I = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{dx}{1+x^6} dx = -\frac{\pi i}{6} (e^{i\frac{\pi}{6}} + e^{i\frac{3\pi}{2}} + e^{5i\frac{\pi}{6}}) = \frac{\pi}{6} (2 \sin \frac{\pi}{6} + 1) = \frac{\pi}{3} \quad (4.46)$$

3. Typ: Integrale der Form

$$I = \int_{-\infty}^{+\infty} f(x)e^{ix} dx, \quad (4.47)$$

wobei $f(z)$ in der oberen Halbebene, ausser in höchstens endlich vielen Punkten, holomorph ist und auf der reellen Achse keine singulären Punkte hat. Man kann zeigen, dass das Integral (4.47) konvergiert, wenn die Funktion $f(x)$ reell ist und für $|x| \rightarrow \pm\infty$ monoton nach 0 strebt. Wir beweisen den

Satz: Wenn $\lim_{|z| \rightarrow \infty} f(z) = 0$ für $y \geq 0$, gilt

$$\lim_{S \rightarrow \infty} \int_{-S}^{+S} f(x)e^{ix} dx = 2\pi i \sum \operatorname{Res}(f(z)e^{iz}), \quad (4.48)$$

wobei die Residuensumme über die in der oberen Halbebene liegenden singulären Punkte von f geht.

Bemerkung: Wenn wir statt $\int_{-\infty}^{+\infty} f(x)e^{ix} dx$ das Integral $\int_{-\infty}^{+\infty} f(x)e^{-ix} dx$ betrachten, muss der komplexe Weg in der unteren Halbebene geschlossen werden. Entscheidend ist wieder das Verhalten von

$$I_S = \int_{\gamma_S} f(z)e^{iz} dz \quad (4.49)$$

für grosses S . Sei $M(S)$ die obere Grenze von $f(Se^{i\phi})$ für $0 \leq \phi \leq \pi$. Dann ist

$$I_S \leq M(S) \int_0^\pi e^{-S \sin \phi} S d\phi = 2M(S) \int_0^{\frac{\pi}{2}} e^{-S \sin \phi} S d\phi \quad (4.50)$$

Aber im Intervall $0 \leq \phi \leq \frac{\pi}{2}$ gilt, dass $\sin \phi \geq \frac{2\phi}{\pi}$ ist. Um das zu sehen, betrachten wir die Funktion

$$h(\phi) = \frac{2}{\pi} \phi - \sin \phi \quad (4.51)$$

Diese erfüllt $h''(\phi) = \sin \phi > 0$ für $0 < \phi < \frac{\pi}{2}$. Daher kann h im offenen Intervall $(0, \frac{\pi}{2})$ kein Maximum annehmen. Aber $h(0) = h(\frac{\pi}{2}) = 0$. Folglich ist im besagten Intervall $h \leq 0$. Also ist

$$\int_0^{\frac{\pi}{2}} e^{-S \sin \phi} S d\phi \leq \int_0^{\frac{\pi}{2}} e^{-\frac{2}{\pi} S \phi} S d\phi \leq \int_0^{\infty} e^{-\frac{2}{\pi} S \phi} S d\phi = \frac{\pi}{2}, \quad (4.52)$$

und I_S geht mit $S \rightarrow \infty$ nach 0, da $M(S)$ das tut, was den Beweis von (4.48) beendet.

Für das Integral (4.47) ist folgendes zu beachten. Die Formel (4.48) sagt insbesondere, dass die linke Seite endlich ist. Diese ist gleich dem Integral I , sofern letzteres konvergiert, was aber nicht automatisch der Fall ist. Dieses muss nämlich an beiden Grenzen unabhängig konvergieren. Dass ist z.B. der Fall, wenn $\int_{-\infty}^{+\infty} |f(x)| dx < \infty$ ist. Wenn aber wie im vorliegenden Fall der Integrand gleich $f(x)\psi(x)$ mit ψ eine oszillierende Funktion ist, kann das Integral (4.47) konvergieren, ohne absolut zu konvergieren. Ein klassisches Resultat, das wir nicht beweisen, ist

Lemma: Gehe $f(x)$ monoton nach Null für grosses x . Ausserdem sei $|\int_a^x \psi(x') dx'|$ mit $x \rightarrow \infty$ beschränkt. Dann konvergiert $\int_a^{+\infty} f(x')\psi(x') dx'$.

Ein weiteres nützliches Resultat ist folgendes: Habe $g(z)$ in $z = 0$ einen einfachen Pol. Dann ist das Halbkreis-Integral (im Gegenuhrzeigersinn)

$$\lim_{\epsilon \rightarrow 0} \int_{\gamma_\epsilon} g(z) dz = \pi i \operatorname{Res}(g, 0) \quad (4.53)$$

Denn g hat eine Laurententwicklung der Form $g(z) = \frac{a}{z} + h(z)$, wobei h im Ursprung holomorph ist. Daher geht $\int_{\gamma_\epsilon} g(z) dz$ im Limes $\epsilon \rightarrow 0$ nach Null. Das Integral $\int_{\gamma_\epsilon} \frac{a}{z} dz = a\pi i$.

Wir betrachten nun das Integral

$$I = \int_0^{+\infty} \frac{\sin x}{x} dx \quad (4.54)$$

Dieses konvergiert bei $x = 0$ wegen $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$ und wegen des obigen Lemmas bei $x = \infty$. Daher ist

$$I = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{\sin x}{x} dx = \frac{1}{2i} \lim_{\epsilon \rightarrow 0} \left[\int_{-\infty}^{-\epsilon} \frac{e^{ix}}{x} dx + \int_{+\epsilon}^{+\infty} \frac{e^{ix}}{x} dx \right] \quad (4.55)$$

Nun ist, da $\frac{e^{iz}}{z}$ keine komplexen Pole hat,

$$\oint_{\gamma} \frac{e^{iz}}{z} dz = \left(\lim_{\epsilon \rightarrow 0} \left[\int_{-\infty}^{-\epsilon} + \int_{+\epsilon}^{+\infty} - \int_{\gamma_\epsilon} \right] + \lim_{R \rightarrow \infty} \int_{\gamma_R} \right) \frac{e^{iz}}{z} dz = 0 \quad (4.56)$$

Das Minuszeichen im dritten Term ergibt sich durch die gegenüber (4.53) umgekehrte Orientierung. Es folgt

$$\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}. \quad (4.57)$$

4. Typ: Wir betrachten Integrale der Form

$$I = \int_0^{+\infty} \frac{R(x)}{x^\alpha} dx, \quad (4.58)$$

wobei $0 < \alpha < 1$ ist und R eine rationale Funktion ist, die für $x \geq 0$ keine Polstellen hat und im Unendlichen mindestens wie $\frac{1}{x}$ abfällt. Das Integral I konvergiert also sowohl bei $x = 0$, als auch bei $x = +\infty$. Für die komplexe Funktion z^α nehmen wir die entlang der positiven reellen Achse aufgeschnittene komplex Ebene, also den Zweig, in dem $0 < \arg z < 2\pi$ ist und als geschlossenen Weg in diesem Gebiet: den grossen positiv orientierten Kreis γ_S plus eine unterhalb des Schnitts zwischen $x = S$ und $x = \epsilon$ verlaufende Gerade plus einen negativ orientierten kleinen Kreisbogen γ_ϵ plus eine oberhalb des Schnitts verlaufende Gerade zwischen $x = \epsilon$ und $x = S$. Man sieht wieder, dass die Kreisbogen-Integrale im Limes $S \rightarrow \infty$ bzw. $\epsilon \rightarrow 0$ verschwinden.

Nun ist, wenn z gegen die positive reelle Achse geht, der Limes von $z^{-\alpha}$ gleich $|z|^{-\alpha}$, wenn man von oberhalb des Schnitts kommt. Wegen (4.12) ist aber der Limes, wenn man von unterhalb kommt, gleich $e^{-2\pi i \alpha}$. In Summe erhalten wir

$$(1 - e^{-2\pi i \alpha}) I = 2\pi i \sum \operatorname{Res} \left(\frac{R(z)}{z^\alpha} \right) \quad (4.59)$$

Beispiel: Das Integral

$$I = \int_0^{+\infty} \frac{dx}{x^\alpha(1+x)} \quad 0 < \alpha < 1 \quad (4.60)$$

Die Funktion $R(z) = \frac{1}{1+z}$ hat einen einfachen Pol bei $z = -1$. In dem gewählten Zweig hat $z^{-\alpha}$ den Wert $e^{-i\pi\alpha}$. Daher ist

$$I = \frac{2\pi i}{e^{i\pi\alpha}(1 - e^{-2\pi i \alpha})} = \frac{\pi}{\sin \pi \alpha} \quad (4.61)$$