

---

Mathematische Ergänzungen  
zur Theoretischen Physik 1

---

H. van Hees

28. Februar 2018



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>3</b>
<b>1 Lineare Algebra</b>	<b>9</b>
1.1 Geometrische Einführung von Euklidischen Vektoren	9
1.1.1 Definition von Vektoren als Verschiebungen	9
1.1.2 Vektoraddition	10
1.1.3 Länge (Norm) von Vektoren	11
1.1.4 Lineare Unabhängigkeit von Vektoren und Basen	12
1.1.5 Der Vektorraum $\mathbb{R}^3$	13
1.1.6 Basistransformationen	14
1.1.7 Das Skalarprodukt	16
1.1.8 Geometrische Anwendungen des Skalarprodukts	17
1.1.9 Das Skalarprodukt im $\mathbb{R}^3$	18
1.2 Axiomatische Begründung der linearen Algebra und Geometrie	19
1.3 Kartesische Basen und orthogonale Transformationen	23
1.4 Das Kreuzprodukt	25
1.5 Das Spatprodukt	31
1.6 Lineare Gleichungssysteme und Determinanten	31
1.6.1 Lineare Gleichungssysteme	32
1.6.2 Determinanten als Volumenform	36
1.6.3 Determinanten von Matrizen	37
1.6.4 Transformationsverhalten des Kreuzprodukts	41
1.7 Drehungen	43
1.7.1 Drehungen in der Ebene	43
1.7.2 Drehungen im Raum	45
1.7.3 Euler-Winkel	47
1.8 Tensoren und ko- und kontravariante Komponenten	49
<b>2 Vektoranalysis</b>	<b>53</b>
2.1 Kurven in Ebene und Raum	53
2.1.1 Ebene Kurven	53
2.1.2 Raumkurven und Fresnetsche Formeln	56
2.1.3 Anwendung auf die Bewegung eines Partikelchens	60
2.2 Skalare Felder, Gradient und Richtungsableitung	63

## INHALTSVERZEICHNIS

2.2.1	Der Gradient in beliebigen Basen . . . . .	66
2.3	Extremwertaufgaben . . . . .	67
2.4	Vektorfelder, Divergenz und Rotation . . . . .	69
2.4.1	Divergenz und Rotation in beliebigen Basen . . . . .	70
2.5	Potentialfelder . . . . .	71
2.6	Wegintegrale und Potentialfelder . . . . .	73
2.7	Flächenintegrale und der Stokessche Satz . . . . .	75
2.7.1	Orientierte Flächen im Raum . . . . .	75
2.7.2	Definition des Flächenintegrals . . . . .	77
2.7.3	Unabhängigkeit des Flächenintegrals von der Parametrisierung . . . . .	78
2.7.4	Koordinatenunabhängige Definition der Rotation . . . . .	79
2.7.5	Der Integralsatz von Stokes . . . . .	80
2.7.6	Der Greensche Satz in der Ebene . . . . .	81
2.8	Das Poincaré-Lemma . . . . .	81
2.8.1	Der Energieerhaltungssatz . . . . .	84
2.9	Volumenintegrale, Divergenz und Gaußscher Integralsatz . . . . .	86
2.9.1	Definition des Volumenintegrals . . . . .	86
2.9.2	Die koordinatenunabhängige Definition der Divergenz . . . . .	88
2.9.3	Der Gaußsche Integralsatz . . . . .	89
2.9.4	Die Greenschen Integralsätze im Raum . . . . .	89
2.10	Krummlinige Orthogonalkoordinaten . . . . .	90
2.10.1	Definition krummliniger Orthogonalkoordinaten . . . . .	90
2.10.2	Polarkoordinaten in der Ebene . . . . .	93
2.10.3	Die Differentialoperatoren grad, div, rot und $\Delta$ . . . . .	95
2.11	Solenoidfelder und Vektorpotentiale . . . . .	98
2.12	Die Poisson-Gleichung und Green-Funktionen . . . . .	100
2.13	Der Helmholtzsche Zerlegungssatz der Vektoranalysis . . . . .	104
2.13.1	Bestimmung des Potentialfeldanteils . . . . .	105
2.13.2	Bestimmung des Solenoidfeldanteils . . . . .	105
2.14	Allgemein kovariante Vektoranalysis . . . . .	107
2.14.1	Konsistenz mit dem Kalkül in krummlinigen Orthonormalkoordinaten . . . . .	113
<b>3</b>	<b>Tensoren und Tensorfelder</b> . . . . .	<b>115</b>
3.1	Multilinearformen . . . . .	115
3.2	Symmetrische Bilinearformen und Hauptachsentransformationen . . . . .	117
3.3	Kegelschnitte im $\mathbb{R}^2$ . . . . .	121
3.4	Elementare Kreiselltheorie . . . . .	125
3.5	Allgemein kovariante Tensorschreibweise . . . . .	134
<b>4</b>	<b>Komplexe Zahlen</b> . . . . .	<b>139</b>
4.1	Definition der komplexen Zahlen . . . . .	139
4.2	Potenzreihen . . . . .	141

<b>5</b>	<b>Gewöhnliche Differentialgleichungen</b>	<b>145</b>
5.1	Differentialgleichungen 1. Ordnung	146
5.1.1	Separierbare Differentialgleichungen	146
5.1.2	Homogene Differentialgleichungen 1. Ordnung	147
5.1.3	Exakte Differentialgleichung 1. Ordnung	148
5.1.4	Homogene lineare Differentialgleichung 1. Ordnung	148
5.1.5	Inhomogene lineare Differentialgleichung 1. Ordnung	148
5.2	Lineare Differentialgleichungen 2. Ordnung	149
5.3	Der ungedämpfte harmonische Oszillator	152
5.4	Der gedämpfte harmonische Oszillator	155
5.4.1	Schwingfall ( $\omega_0 > \gamma$ )	156
5.4.2	Kriechfall ( $\omega_0 < \gamma$ )	158
5.4.3	Aperiodischer Grenzfall ( $\omega_0 = \gamma$ )	159
5.4.4	Direkte Lösung im aperiodischen Grenzfall	159
5.5	Der getriebene gedämpfte Oszillator	160
5.5.1	Spezielle Lösung der inhomogenen Gleichung	161
5.5.2	Amplitudenresonanzfrequenz	162
5.5.3	Energieresonanz	163
5.5.4	Lösung des Anfangswertproblems	164
5.5.5	Resonant angetriebener ungedämpfter Oszillator	165
5.5.6	Allgemeine äußere Kräfte und die $\delta$ -Distribution	166
<b>A</b>	<b>Zusammenfassung der Analysis für reelle Funktionen einer Variablen</b>	<b>171</b>
A.1	Mengen und reelle Zahlen	171
A.2	Folgen und Grenzwerte	173
A.3	Satz vom Supremum und Infimum	180
A.4	Lineare und quadratische Gleichungen	180
A.5	Funktionen und Stetigkeit	182
A.6	Differentialrechnung für Funktionen einer reellen Veränderlichen	183
A.6.1	Definition der Ableitung einer Funktion	183
A.6.2	Formeln zur Ableitung	184
A.6.3	Kurvendiskussionen, Mittelwertsatz der Differentialrechnung	189
A.7	Integralrechnung	191
A.7.1	Definition des Riemann-Integrals	191
A.7.2	Der Mittelwertsatz der Integralrechnung	194
A.7.3	Der Hauptsatz der Analysis	194
A.7.4	Integrationstechniken	196
A.7.5	Funktionenfolgen und -reihen	197
A.7.6	Taylor-Entwicklung und Potenzreihen	200
A.8	Die strikte Definition der trigonometrischen Funktionen	205
	<b>Literaturverzeichnis</b>	<b>209</b>

*INHALTSVERZEICHNIS*

# Einleitung

## Allgemeine Hinweise zur Vorlesung

Dies ist das Manuskript zur Vorlesung „Mathematische Ergänzungen zur Theoretischen Physik 1“. Ziel dieser Vorlesung ist es, die in der Vorlesung „Theoretische Physik 1“ benötigten mathematischen Methoden zu vertiefen und vor allem auf konkrete physikalische Probleme anzuwenden. Der Schwerpunkt liegt entsprechend weniger auf formalen Beweisen als vielmehr auf der Vermittlung der Rechentechnik, die sehr wichtig für das Verständnis der theoretischen Physik ist.

Inhaltlich ist die Vorlesung durch die physikalischen Anwendungen in der klassischen Newtonschen Mechanik festgelegt. Entsprechend beginnen wir mit einer Einführung der wesentlichen Grundlagen der Vektorrechnung, wobei wir von anschaulichen geometrischen Definitionen in Ebene und Raum ausgehen. Dies wird naturgemäß für die Behandlung der Dynamik eines oder mehrerer Massenpunkte benötigt. Wir gelangen dann schnell über die Darstellung der Vektoren mittels reeller Komponenten bzgl. einer Basis zur quantitativen Darstellung von Bewegungsabläufen.

Deren weitere Analyse erfordert dann auch die Erweiterung der Methoden der so definierten **linearen Algebra** zu solchen der eigentlichen **Vektor-Analyse**. Entsprechend werden wir über die physikalischen Größen Ort, Geschwindigkeit und Beschleunigung die Ableitung von Vektorfunktionen nach äußeren Parametern (hier naturgemäß der Zeit) einführen und geometrisch deuten.

Ebenso werden die wesentlichen Grundbegriffe der **Feldtheorie** eingeführt, wie die Ableitungen **skalarer Felder** und **Vektorfelder** (grad, div und rot), die durch den „Nabla-Operator“  $\vec{\nabla}$  kompakt dargestellt werden. Ebenso besprechen wir auch ausführlich das Rechnen mit Komponenten, den sog. **Ricci-Kalkül**.

Den zweiten Schwerpunkt bilden Techniken zur Lösung von **gewöhnlichen Differentialgleichungen**, also die Integration der einfachsten Typen von Bewegungsgleichungen, wie sie in der klassischen Mechanik auftreten, insbesondere das wichtige Beispiel des **harmonischen Oszillators**. Dazu führen wir auch **komplexe Zahlen** ein und besprechen die wichtigsten **elementaren Funktionen** wie Polynome, die Exponentialfunktion, die trigonometrischen Funktionen und deren Umkehrungen.

Es sei betont, daß in der Vorlesung nicht notwendig alle Inhalte dieses Manuskripts abgearbeitet werden müssen. Der Inhalt der Vorlesung richtet sich nicht zuletzt auch nach den Bedürfnissen der Hörerinnen und Hörer der Hauptvorlesung und dient auch der Bereitstellung mathematischer Grundlagen und Methoden für die Vorlesung!

**Literaturempfehlungen:** Die Literatur zum Thema „Mathematik für Physiker“ ist nahezu unerschöpflich. Für die Zielsetzung dieser Vorlesung ist besonders [Gro05] zu empfehlen, das den mathematischen Teil des Vorlesungsstoffes vollständig abdeckt. Eine sehr anschauliche Behandlung der Vektoranalysis sind [BK88, SH99]. Auch viele Theorie-Lehrbücher bieten einen Überblick über die zur Anwendung kommende Mathematik. Als besonders gelungen empfinde ich die älteren Bücher [Joo89, Sau73, Som92]. Die Theorievorlesung selbst orientiert sich an dem Skript von D. Rischke [Ris09]. Standardlehrbücher sind noch [Gre03, Sch07, Nol13].

*INHALTSVERZEICHNIS*

## Kapitel 1

# Lineare Algebra

Die **Vektorrechnung** oder **lineare Algebra** und die **Vektoranalysis** stellen die wesentliche mathematische Grundlage zur Beschreibung physikalischer Systeme dar. In dieser Vorlesung behandeln wir die **Newtonsche Mechanik**, in der es vornehmlich um die Beschreibung der Bewegung von als punktförmig idealisierten Körpern aufgrund vorgegebener auf sie einwirkender Kräfte geht.

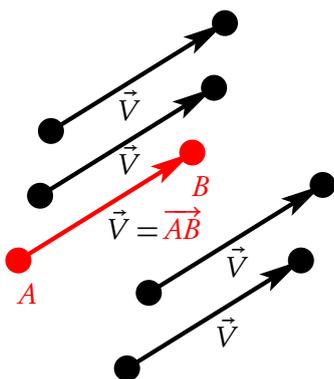
Wir gehen von der Euklidischen Geometrie des physikalischen dreidimensionalen Raumes aus, die wir als bekannt voraussetzen. Wir werden schrittweise anhand des Begriffes der **Verschiebung** im physikalischen Anschauungsraum die Algebra von Vektoren (Addition und Multiplikation mit reellen Zahlen) motivieren und dann abstrahieren.

### 1.1 Geometrische Einführung von Euklidischen Vektoren

Wir gehen davon aus, daß die Grundlagen der **Euklidischen Geometrie** von der Schulmathematik her bekannt sind und entwickeln deren Formulierung als **analytische Geometrie** mit Hilfe von Vektoren. Dies ist für die gesamte moderne Naturwissenschaft, wie sie von Galilei und Newton im 17. Jh. begründet worden ist, von entscheidender Bedeutung, denn die analytische Geometrie macht die Analyse von **Bewegungen** von Körpern im Raum der Analysis, also der **Differential- und Integralrechnung**, zugänglich.

Die Anwendungen der hier behandelten mathematischen Methoden in der Physik gehen aber weit über die klassische Mechanik hinaus. Sie stellen daher die unverzichtbare Grundlage für das gesamte Physikstudium dar. Insbesondere die Vektoranalysis wird in nahezu allen Teilbereichen der Physik benötigt. Außer in den Theorievorlesungen 1 und 2, die von der klassischen bzw. analytischen Newtonschen (und speziell relativistischen) Mechanik handeln, kommt sie vor allem in der Elektrodynamik (Theorie 3) zum Einsatz.

#### 1.1.1 Definition von Vektoren als Verschiebungen

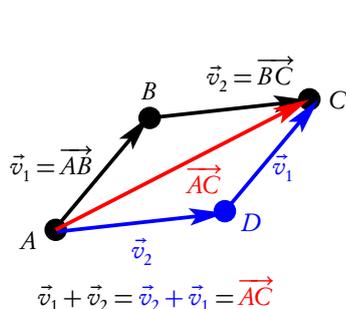


Wir führen den Begriff des **Vektors** anhand der Beschreibung von **Verschiebungen** ein. Seien also  $A$  und  $B$  zwei Punkte. Aufgrund der Axiome, die der Euklidischen Geometrie zugrundeliegen, können wir diese beiden Punkte durch eine **gerade Strecke** verbinden, die eine Länge besitzt. Wir markieren zugleich die Richtung mit einem Pfeil, legen also fest, daß wir die Verschiebung des Punktes  $A$  nach  $B$  entlang der geraden Strecke meinen. Dieses Objekt, das die Verschiebung von  $A$  nach  $B$  durch **Länge und Richtung** festlegt, bezeichnen wir als den Vektor  $\vec{AB}$ .

Nun interessieren wir uns oft nicht für die konkreten beiden Punkte, die wir ineinander verschieben, sondern lediglich für Länge und Richtung der Ver-

schiebung selbst. Daher identifizieren wir alle Pfeile, die aus  $\overrightarrow{AB}$  durch **Parallelverschiebung** hervorgehen und bezeichnen diese Klasse von Pfeilen als den Vektor  $\vec{v} = \overrightarrow{AB}$ . Wir bezeichnen die Menge aller Vektoren in einer Ebene mit  $E^2$  bzw.  $E^3$ , wobei das  $E$  für „Euklidischer Vektorraum“ und der hochgestellte Index 2 bzw. 3 die Dimensionszahl angibt.

### 1.1.2 Vektoraddition



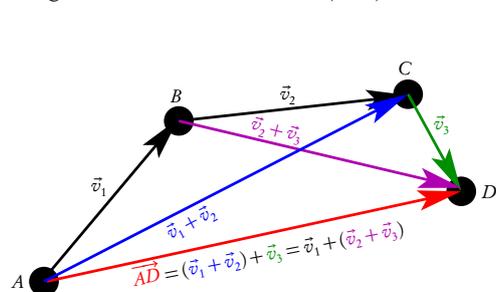
Geben wir zwei Vektoren  $\vec{v}_1 = \overrightarrow{AB}$  und  $\vec{v}_2 = \overrightarrow{BC}$  vor, so definieren wir die Hintereinanderausführung der Verschiebungen (s. nebenstehende Skizze), die direkt zur Verschiebung  $\overrightarrow{AC}$  führt, als die **Summe der Vektoren**:  $\vec{v}_1 + \vec{v}_2 = \overrightarrow{AC}$ .

Durch Parallelverschiebung von  $\vec{v}_2$ , so daß sein Anfangspunkt in  $A$  zu liegen kommt, ergibt den Punkt  $D$  vermöge  $\vec{v}_2 = \overrightarrow{AD}$ . Nach den Gesetzen der Euklidischen Geometrie ist dann  $\overrightarrow{DC} = \vec{v}_1$ . Entsprechend folgt  $\vec{v}_2 + \vec{v}_1 = \overrightarrow{AD} + \overrightarrow{DC} = \overrightarrow{AC} = \vec{v}_1 + \vec{v}_2$ . Das bedeutet, daß die **Vektoraddition kommutativ** ist, d.h. die Summe zweier Vektoren hängt nicht von der Reihenfolge der

Summation ab.

Nun führen wir noch den (nur scheinbar sinnlosen) **Nullvektor**  $\overrightarrow{AA} = \vec{0}$  ein. Es ist klar, daß das im Sinne von Verschiebungen bedeutet, daß gar keine Verschiebung ausgeführt wird. Im Sinne unserer Äquivalenzklassenbildung gilt für jeden anderen Punkt  $B$  ebenfalls, daß  $\overrightarrow{BB} = \vec{0}$  ist. Entsprechend folgt für die Addition  $\overrightarrow{AA} + \overrightarrow{AB} = \overrightarrow{AB}$  bzw.  $\vec{0} + \vec{v}_1 = \vec{v}_1 + \vec{0} = \vec{v}_1$ . Der Nullvektor ist also das **neutrale Element der Vektoraddition**.

Es ist auch klar, daß wir zu jedem Verschiebungsvektor  $\vec{v} = \overrightarrow{AB}$  den die umgekehrte Verschiebung kennzeichnenden Vektor  $(-\vec{v}) = \overrightarrow{BA}$  zuordnen können. Der Summe dieser beiden Vektoren entspricht gerade die Verschiebung von  $A$  nach  $B$  und dann wieder zurück zu  $A$ . Insgesamt haben wir also gar keine Verschiebung ausgeführt. Es ist also  $\vec{v} + (-\vec{v}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{0}$ .



Betrachten wir nun drei Vektoren  $\vec{v}_1 = \overrightarrow{AB}$ ,  $\vec{v}_2 = \overrightarrow{BC}$  und  $\vec{v}_3 = \overrightarrow{CD}$ . Dann ist

$$\vec{v}_1 + \vec{v}_2 = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}. \quad (1.1.1)$$

und folglich

$$(\vec{v}_1 + \vec{v}_2) + \vec{v}_3 = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD}. \quad (1.1.2)$$

Addieren wir jetzt diese drei Vektoren in einer etwas anderen Reihenfolge, und zwar bilden wir zuerst die Summe

$$\vec{v}_2 + \vec{v}_3 = \overrightarrow{BC} + \overrightarrow{CD} = \overrightarrow{BD}. \quad (1.1.3)$$

Dann folgt

$$\vec{v}_1 + (\vec{v}_2 + \vec{v}_3) = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}. \quad (1.1.4)$$

Vergleichen wir dies mit (1.1.2), ergibt sich das **Assoziativgesetz der Vektoraddition**

$$(\vec{v}_1 + \vec{v}_2) + \vec{v}_3 = \vec{v}_1 + (\vec{v}_2 + \vec{v}_3). \quad (1.1.5)$$

Dies zeigt, daß wir hinsichtlich der Addition mit Vektoren formal genauso wie mit reellen Zahlen rechnen können. Insbesondere können wir auch Gleichungen lösen. Seien z.B.  $\vec{a}$  und  $\vec{b}$  vorgegebene Vektoren. Wir

suchen nun einen Vektor  $\vec{x}$ , der die Gleichung  $\vec{a} + \vec{x} = \vec{b}$  erfüllt. Hätten wir Zahlen vorliegen, könnten wir einfach  $\vec{a}$  auf beiden Seiten der Gleichungen abziehen, um  $\vec{x}$  zu finden. Aufgrund der eben hergeleiteten Rechenregeln funktioniert das auch für Vektoren, denn es gilt

$$\vec{b} + (-\vec{a}) = (\vec{a} + \vec{x}) + (-\vec{a}) = (-\vec{a}) + (\vec{a} + \vec{x}) = [(-\vec{a}) + \vec{a}] + \vec{x} = \vec{0} + \vec{x} = \vec{x}. \quad (1.1.6)$$

Es ist eine gute *Übung* sich zu vergewissern, welche der oben hergeleiteten Rechenregeln bei den einzelnen Umformungsschritten verwendet wurden! Entsprechend definieren wir die **Subtraktion von Vektoren** in der naheliegenden Weise als  $\vec{b} - \vec{a} = \vec{b} + (-\vec{a})$ .

### 1.1.3 Länge (Norm) von Vektoren

Bisher haben wir nicht von der Eigenschaft von Vektoren Gebrauch gemacht, daß sie auch eine **Länge** besitzen. Die Länge des Vektors  $\vec{v} = \overrightarrow{AB}$  ist dabei natürlich einfach durch die Länge der Strecke  $|\vec{v}| = |\overrightarrow{AB}|$  im Sinne der Euklidischen Geometrie definiert. Man nennt  $|\vec{v}|$  auch die **Euklidische Norm** des Vektors  $\vec{v}$  oder (vor allem in der Physik) auch einfach den **Betrag** oder die **Länge** des Vektors  $\vec{v}$ . Der Betrag ist eine positive reelle Zahl. Daß für die Länge von Strecken *reelle* Zahlen benötigt werden und nicht etwa rationale Zahlen ausreichen, ist keinesfalls trivial. Erst D. Hilbert hat Ende des 19. Jh. bemerkt, daß die Manipulationen mit Lineal und Zirkel, wie sie Euklid im Altertum ausgeführt bzw. axiomatisch begründet hat, die reellen Zahlen erfordern, also auch irrationale Zahlen benötigt werden.

Die Euklidische Norm von Vektoren erbt nun naturgemäß einige Eigenschaften vom Längenbegriff der Euklidischen Geometrie. Z.B. ist die Länge des Nullvektors 0:  $|\vec{0}| = 0$ , denn ein Punkt besitzt definitionsgemäß keine Ausdehnung. Ist umgekehrt  $\vec{v}$  ein Vektor mit  $|\vec{v}| = 0$  ist offenbar  $\vec{v} = \vec{0}$ .

Weniger trivial ist die **Dreiecksungleichung**. Sind nämlich  $A, B$  und  $C$  drei beliebige nicht auf einer Gerade gelegene Punkte, dann gilt für die Seiten des von ihnen definierten Dreiecks stets  $|AB| + |BC| > |AC|$ . Seien also  $\vec{v}_1 = \overrightarrow{AB}$ ,  $\vec{v}_2 = \overrightarrow{BC}$ , so gilt

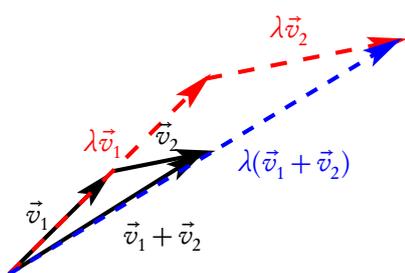
$$|\overrightarrow{AC}| = |\vec{v}_1 + \vec{v}_2| < |\vec{v}_1| + |\vec{v}_2|. \quad (1.1.7)$$

Liegen die drei Punkte auf einer Geraden und  $B$  zwischen  $A$  und  $C$ , so gilt offenbar  $|AB| + |BC| = |AC|$ . In diesem Fall gilt also  $|\vec{v}_1 + \vec{v}_2| = |\vec{v}_1| + |\vec{v}_2|$ . Es gilt also für alle Vektoren  $\vec{v}_1$  und  $\vec{v}_2$  immer die **Dreiecksungleichung**

$$|\vec{v}_1 + \vec{v}_2| \leq |\vec{v}_1| + |\vec{v}_2|. \quad (1.1.8)$$

Das Gleichheitszeichen gilt offenbar dann und nur dann, wenn die Vektoren  $\vec{v}_1$  und  $\vec{v}_2$  parallel zueinander sind.

Nun gibt es in der Euklidischen Geometrie zu zwei Punkten  $A$  und  $B$  und jeder reellen Zahl  $\lambda > 0$  einen Punkt auf der durch  $A$  und  $B$  eindeutig festgelegten Geraden einen Punkt  $C$ , so daß  $|AC| = \lambda|AB|$ , wobei wir festlegen, daß für  $\lambda < 1$  der Punkt  $C$  zwischen  $A$  und  $B$  und für  $\lambda > 1$  der Punkt  $B$  zwischen  $A$  und  $C$  liegen soll. Entsprechend definieren wir die Multiplikation des Vektors  $\vec{v} = \overrightarrow{AB}$  mit der reellen positiven Zahl  $\lambda$  durch  $\lambda\vec{v} = \overrightarrow{AC}$ . Anders ausgedrückt bedeutet die Verschiebung um den Vektor  $\lambda\vec{v}$  eine Verschiebung in die gleiche Richtung wie die durch  $\vec{v}$  vorgegebene Verschiebung, aber um eine um den Faktor  $\lambda$  verschiedene Länge. Wir wollen eine solche Multiplikation von Vektoren mit reellen Zahlen auch für  $\lambda < 0$  definieren. Wie wir gleich sehen werden, ist es sinnvoll, in diesem Fall  $\lambda\vec{v} = -(|\lambda|\vec{v})$ . Dies liegt nahe, denn für  $\lambda < 0$  ist  $\lambda = -|\lambda|$ . Wir verschieben in diesem Fall also um eine um den Faktor  $\lambda$  geänderte Strecke in entgegengesetzter Richtung zu  $\vec{v}$ . Schließlich definieren wir noch, daß  $0\vec{v} = \vec{0}$  sein soll. Man macht sich schnell klar, daß für zwei Zahlen  $\lambda_1, \lambda_2 \in \mathbb{R}$  das **Assoziativgesetz**  $\lambda_1(\lambda_2\vec{v}) = (\lambda_1\lambda_2)\vec{v}$  gilt.



Es ist unmittelbar einsichtig, daß  $(\lambda_1 + \lambda_2)\vec{v} = \lambda_1\vec{v} + \lambda_2\vec{v}$ . Es ergeben sich aus diesen Rechenregeln sofort die unmittelbar einleuchtende Formeln wie  $\vec{v} + \vec{v} = 2\vec{v}$ , d.h. führt man zweimal dieselbe Verschiebung hintereinander aus, erhält man eine Verschiebung in die gleiche Richtung aber um die doppelte Länge. Aus der nebenstehenden Skizze entnehmen wir, daß aufgrund des Strahlensatzes der Euklidischen Geometrie auch das **Distributivgesetz**  $\lambda(\vec{v}_1 + \vec{v}_2) = \lambda\vec{v}_1 + \lambda\vec{v}_2$  gilt.

### 1.1.4 Lineare Unabhängigkeit von Vektoren und Basen

Seien  $\vec{b}_1$  und  $\vec{b}_2$  zwei nichtparallele Vektoren. Das bedeutet, daß es keine reelle Zahl  $\lambda$  gibt, für die  $\lambda\vec{b}_1 = \vec{b}_2$  ist. Man nennt solche Vektoren **linear unabhängig**. Eine etwas allgemeinere Definition ist, daß zwei Vektoren linear unabhängig voneinander sind, genau dann wenn aus  $\lambda_1\vec{b}_1 + \lambda_2\vec{b}_2 = \vec{0}$  folgt, daß notwendig  $\lambda_1 = \lambda_2 = 0$  sein muß. Beide Definitionen sind offenbar äquivalent. Gilt nämlich  $\lambda\vec{b}_1 = \vec{b}_2$ , so ist  $\lambda\vec{b}_1 - \vec{b}_2 = \vec{0}$ . Es ist also zumindest  $\lambda_2 = -1 \neq 0$ , so daß die Vektoren linear abhängig sind. Ist umgekehrt  $\lambda_1\vec{b}_1 + \lambda_2\vec{b}_2 = \vec{0}$  und  $\lambda_2 \neq 0$ , so gilt  $\vec{b}_2 = -(\lambda_1/\lambda_2)\vec{b}_1$ , d.h. die Vektoren sind nach der ersten Definition linear abhängig.

Dies läßt sich nun auf beliebig viele Vektoren verallgemeinern. Wir nennen eine beliebige endliche Menge von Vektoren  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  **voneinander linear unabhängig**, genau dann wenn aus

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \dots + \lambda_n\vec{v}_n = \sum_{j=1}^n \lambda_j\vec{v}_j = \vec{0} \quad (1.1.9)$$

notwendig  $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$  folgt. Andernfalls heißen die Vektoren **voneinander linear abhängig**.

Betrachten wir nun als Beispiel Vektoren in einer Ebene. Seien  $\vec{b}_1$  und  $\vec{b}_2$  zwei beliebige voneinander linear unabhängige Vektoren. Dann können wir jeden beliebigen Vektor  $\vec{v}$  durch **Linearkombination** aus diesen **Basisvektoren** zusammensetzen. Wir schreiben die entsprechenden Zahlen, die **Komponenten** von  $\vec{v}$  als  $v^1$  und  $v^2$ , d.h. wir können stets Zahlen  $v^j$  ( $j \in \{1, 2\}$ ) finden, so daß

$$\vec{v} = v^1\vec{b}_1 + v^2\vec{b}_2 = \sum_{j=1}^2 v^j\vec{b}_j. \quad (1.1.10)$$

Es ist nun auch klar, daß diese Zahlen eindeutig sind. Die Vektoren  $\vec{b}_1$  und  $\vec{b}_2$  sind nämlich voneinander linear unabhängig, denn sie sind nicht parallel zueinander, weisen also in verschiedene Richtungen in der Ebene. Seien nun  $\lambda^1$  und  $\lambda^2$  irgendwelche Komponenten von  $\vec{v}$ , folgt nämlich

$$\vec{0} = \vec{v} - \vec{v} = (v^1 - \lambda^1)\vec{b}_1 + (v^2 - \lambda^2)\vec{b}_2. \quad (1.1.11)$$

Da  $\vec{b}_1$  und  $\vec{b}_2$  linear unabhängig sind, folgt daraus notwendig, daß  $v^1 - \lambda^1 = 0$ , also  $v^1 = \lambda^1$ , und  $v^2 - \lambda^2 = 0$ , also  $v^2 = \lambda^2$  sein muß.

Man nennt nun eine Menge von Vektoren  $\{\vec{b}_1, \dots, \vec{b}_n\}$  **vollständig**, wenn man jeden Vektor  $\vec{v}$  als Linearkombination dieser Vektoren darstellen kann. Falls diese Menge zusätzlich auch noch linear unabhängig ist, ist diese Linearkombination für jeden Vektor eindeutig, was man genauso beweist für unser Beispiel mit zwei Vektoren in der Ebene, und man nennt entsprechend jede vollständige Menge linear unabhängiger Vektoren **eine Basis** des Vektorraumes. Für  $E^2$  bestehen alle Basen offensichtlich aus genau zwei Vektoren.

<sup>1</sup>Man beachte, daß hier die hochgestellten Zahlen keine Potenzen bezeichnen sondern Indizes, die die Komponenten durchnummerieren. Der Sinn, die Indizes für die Komponenten oben und die für die Basisvektoren unten zu schreiben, wird im folgenden noch klar werden.

Genauso bestehen alle Basen im dreidimensionalen Raum offensichtlich aus beliebigen Mengen von genau drei linear unabhängigen Vektoren. Man nennt einen Vektorraum, der eine Basis aus endlich vielen Vektoren besitzt, einen **endlichdimensionalen Vektorraum**. Offenbar bilden die geometrischen Vektoren wie wir sie in diesem Abschnitt definiert haben, in einer Ebene einen zweidimensionalen bzw. im Raum einen dreidimensionalen Vektorraum.

Wir merken hier nur an, daß es Vektorräume beliebiger endlicher Dimension aber auch solche unendlicher Dimension gibt. In diesem Manuskript befassen wir uns nur mit endlichdimensionalen Vektorräumen, und zwar vornehmlich mit den in der Euklidischen Geometrie des physikalischen Raumes der Newtonschen Mechanik auftretenden zwei- und dreidimensionalen Vektorräumen. Beschränkt man sich auf Vektoren entlang einer Geraden, hat man es auch mit einem eindimensionalen Vektorraum zu tun.

### 1.1.5 Der Vektorraum $\mathbb{R}^3$

Wir haben im vorigen Abschnitt gesehen, daß wir durch Einführung einer Basis  $\{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$  jeden räumlichen Vektor  $\vec{v}$  durch seine drei Komponenten  $v^1$ ,  $v^2$  und  $v^3$  eindeutig als Linearkombination dieser Basisvektoren

$$\vec{v} = v^1 \vec{b}_1 + v^2 \vec{b}_2 + v^3 \vec{b}_3 = \sum_{j=1}^3 v^j \vec{b}_j \quad (1.1.12)$$

darstellen können. Da solche Summenbildungen im folgenden ständig auftreten, läßt man oft auch die Summenzeichen einfach weg. Diese Konvention geht auf Einstein zurück, der sie bei der Formulierung der Allgemeinen Relativitätstheorie eingeführt hat. Man spricht daher auch von der **Einsteinschen Summationskonvention**. Wir werden sie allerdings zu Anfang noch nicht benutzen, um klar daran zu erinnern, wann wir Größen summieren wollen. Es ist klar, daß umgekehrt auch durch beliebige drei Zahlen  $(v^j)$  ( $j \in \{1, 2, 3\}$ ) (1.1.12) ein Vektor  $\vec{v}$  definiert ist. Haben wir also, einmal eine Basis festgelegt, können wir genauso gut mit diesen **geordneten Zahlentripeln** arbeiten, und zwar ordnen wir diese Zahlentripel gewöhnlich in einer Spalte an

$$\vec{v} \mapsto \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \equiv (v^j) \equiv \bar{v}. \quad (1.1.13)$$

Wir bezeichnen die so definierten **Spaltenvektoren** mit demselben Symbol wie die geometrischen Vektoren, zur Unterscheidung aber mit einem Oberstrich anstelle eines Pfeilchens. Man muß sich dabei immer vergewissern, bzgl. welcher Basis eine solche Darstellung als Spaltenvektor gemeint ist.

Seien nun  $\vec{v}$  und  $\vec{w}$  beliebige räumliche Vektoren. Dann muß sich die Summe dieser Vektoren eindeutig durch die Spalten  $\bar{v}$  und  $\bar{w}$  darstellen lassen. Dies ist in der Tat einfach zu sehen, denn es gilt

$$\vec{v} + \vec{w} = \sum_{j=1}^3 v^j \vec{b}_j + \sum_{j=1}^3 w^j \vec{b}_j = \sum_{j=1}^3 (v^j \vec{b}_j + w^j \vec{b}_j) = \sum_{j=1}^3 (v^j + w^j) \vec{b}_j. \quad (1.1.14)$$

Es ist also der Summe der beiden Vektoren eindeutig der Spaltenvektor

$$\bar{v} + \bar{w} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} + \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} = \begin{pmatrix} v^1 + w^1 \\ v^2 + w^2 \\ v^3 + w^3 \end{pmatrix} \quad (1.1.15)$$

zugeordnet. Es werden also einfach die entsprechenden Komponenten addiert.

Genauso zeigt man (*Übung*), daß  $\lambda \vec{v}$  der Spaltenvektor

$$\lambda \bar{v} = \lambda \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} \lambda v^1 \\ \lambda v^2 \\ \lambda v^3 \end{pmatrix} \quad (1.1.16)$$

zugeordnet ist, d.h. es werden die Komponenten einfach mit der Zahl  $\lambda$  multipliziert.

Ebenso ist es leicht einzusehen (*Übung*), daß die in Spalten angeordneten Zahlentripel mit den Rechenregeln (1.1.15) und (1.1.16) genauso wie die geometrisch definierten Vektoren einen dreidimensionalen reellen Vektorraum bilden, für die bzgl. Addition von Vektoren und Multiplikation von Vektoren mit reellen Zahlen dieselben Rechenregeln gelten. Da die Vektorkomponenten reelle Zahlen sind, nennt man diesen Vektorraum mit den so definierten Rechenoperationen  $\mathbb{R}^3$ . Wir haben also eine umkehrbar eindeutige Abbildung zwischen dem geometrischen Vektorraum  $E^3$  und dem aus den Zahlentripeln des  $\mathbb{R}^3$  gebildeten Vektorraum gefunden. Die Zahlentripel  $\mathbb{R}^3$  bilden zudem die gleiche **algebraische Struktur** wie der geometrische Vektorraum. Man spricht bei solchen umkehrbar eindeutigen Abbildungen zwischen zwei solcherart gleichartigen algebraischen Strukturen, für die sich die Rechenoperationen zudem noch umkehrbar eindeutig entsprechen von **Homomorphismen**. Vom Standpunkt einer rein axiomatischen Definition eines Vektorraumes sind die durch einen Homomorphismus verknüpften algebraischen Strukturen nicht unterscheidbar. Sie sind vollständig zueinander **äquivalent**.

### 1.1.6 Basistransformationen

Es ist klar, daß wir die Betrachtungen des vorigen Abschnitts mit einer beliebigen Basis ausführen können. Es ist allerdings sehr wichtig, stets in Erinnerung zu behalten, daß diese Abbildung von der willkürlichen Wahl der **Basisvektoren** abhängig ist. In diesem Abschnitt befassen wir uns daher mit **Basistransformationen**, also der Frage, wie wir die Komponenten von Vektoren bzgl. einer Basis  $\{\vec{b}_j\}_{j \in \{1,2,3\}}$  in die Komponenten des gleichen Vektors bzgl. einer anderen Basis  $\{\vec{b}'_j\}_{j \in \{1,2,3\}}$  umrechnen können. Definitionsgemäß gilt

$$\vec{v} = \sum_{j=1}^3 v^j \vec{b}_j = \sum_{k=1}^3 v'^k \vec{b}'_k. \quad (1.1.17)$$

Da die Abbildung des Vektors  $\vec{V} \in E^3$  auf die  $\mathbb{R}^3$ -Vektoren  $\vec{v}$  bzw.  $\vec{v}'$  jeweils umkehrbar eindeutig sind, muß es entsprechend eine umkehrbar eindeutige Abbildung dieser beiden  $\mathbb{R}^3$ -Vektoren geben.

Dazu müssen wir nur bedenken, daß es offenbar neun eindeutige Zahlen  $T^k_j$  gibt, so daß

$$\vec{b}_j = \sum_{k=1}^3 T^k_j \vec{b}'_k \quad (1.1.18)$$

gilt. Das ist so, weil definitionsgemäß die Vektoren  $\vec{b}'_k$  eine Basis bilden und somit jeder Vektor eine eindeutige Linearkombination dieser drei Basisvektoren ist, insbesondere also auch  $\vec{b}_j$ . Setzen wir (1.1.18) in (1.1.17) ein, ergibt sich

$$\vec{v} = \sum_{j=1}^3 v^j \vec{b}_j = \sum_{j=1}^3 \sum_{k=1}^3 T^k_j v^j \vec{b}'_k = \sum_{k=1}^3 \left( \sum_{j=1}^3 T^k_j v^j \right) \vec{b}'_k = \sum_{k=1}^3 v'^k \vec{b}'_k. \quad (1.1.19)$$

Da die Linearkombination von  $\vec{v}$  bzgl. der Basis  $\vec{b}'_k$  eindeutig ist, folgt daraus, daß notwendig

$$v'^k = \sum_{j=1}^3 T^k_j v^j \quad (1.1.20)$$

sein muß. Kennen wir also die neun Zahlen  $T^k_j$ , können wir direkt die Komponenten bzgl. der alten Basis in diejenigen der neuen umrechnen.

Gewöhnlich ordnet man die  $T^k_j$  in ein rechteckiges  $3 \times 3$ -Zahlenschema, eine sog. **Matrix** an:

$$\hat{T} \equiv (T^k_j) = \begin{pmatrix} T^1_1 & T^1_2 & T^1_3 \\ T^2_1 & T^2_2 & T^2_3 \\ T^3_1 & T^3_2 & T^3_3 \end{pmatrix}. \quad (1.1.21)$$

Die Transformation (1.1.20) schreibt man dann kurz in der Form

$$\bar{v}' = \hat{T} \cdot \bar{v} \equiv \hat{T}\bar{v}. \quad (1.1.22)$$

Ausführlich geschrieben lautet die Rechenvorschrift

$$\begin{pmatrix} v'^1 \\ v'^2 \\ v'^3 \end{pmatrix} = \begin{pmatrix} T^1_1 & T^1_2 & T^1_3 \\ T^2_1 & T^2_2 & T^2_3 \\ T^3_1 & T^3_2 & T^3_3 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} T^1_1 v^1 + T^1_2 v^2 + T^1_3 v^3 \\ T^2_1 v^1 + T^2_2 v^2 + T^2_3 v^3 \\ T^3_1 v^1 + T^3_2 v^2 + T^3_3 v^3 \end{pmatrix}. \quad (1.1.23)$$

Die einzelnen Einträge haben also die Form „Zeile  $\times$  Spalte“, wobei man die Multiplikation jeweils als die angegebene Summe von Produkten anzusehen hat.

Die gleiche Argumentation führt auch umgekehrt zur Berechnung der Komponenten  $\bar{v}$  aus den Komponenten  $\bar{v}'$ . Dazu müssen wir nur die entsprechende Matrix via

$$\vec{b}'_k = \sum_{j=1}^3 U^j_k \vec{b}_j \quad (1.1.24)$$

eingeführen. Dann folgt

$$\bar{v} = \hat{U}\bar{v}'. \quad (1.1.25)$$

Kombinieren wir (1.1.25) mit (1.1.22) folgt für alle  $\bar{v} \in \mathbb{R}^3$

$$\bar{v} = \hat{U}\bar{v}' = \hat{U}\hat{T}\bar{v}. \quad (1.1.26)$$

Zum einen führt dies das **Produkt zweier Matrizen** ein;  $\hat{U}\hat{T}$  ist wieder eine  $\mathbb{R}^{3 \times 3}$ -Matrix, und man bildet sie wieder nach dem Schema „Zeile  $\times$  Spalte“. Zum anderen folgt aber aus (1.1.26), daß  $\hat{U}\hat{T}$  eine Matrix sein muß, die alle Vektoren  $\bar{v} \in \mathbb{R}^3$  auf sich selbst abbildet. Das kann nur die **Einheitsmatrix**

$$\mathbb{1}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv \text{diag}(1, 1, 1) \quad (1.1.27)$$

sein. Es gilt also notwendig

$$\hat{U}\hat{T} = \mathbb{1}_3. \quad (1.1.28)$$

Es ist also  $\hat{U}$  die **inverse Matrix** zu  $\hat{T}$  im Sinne der Matrizenmultiplikation. Man schreibt daher auch

$$\hat{U} = \hat{T}^{-1}. \quad (1.1.29)$$

Wir können natürlich auch umgekehrt von (1.1.22) ausgehen und dann mit (1.1.25) auf

$$\bar{v}' = \hat{T}\bar{v} = \hat{T}\hat{U}\bar{v} \quad (1.1.30)$$

schließen. Es gilt also mit denselben Argumenten wie eben auch

$$\hat{T}\hat{U} = \mathbb{1}_3 \Rightarrow \hat{T} = \hat{U}^{-1}. \quad (1.1.31)$$

Wir weisen bereits hier darauf hin, daß i.a. die *Matrizenmultiplikation nicht kommutativ ist*, d.h. für irgendwelche Matrizen  $\hat{A}, \hat{B} \in \mathbb{R}^{3 \times 3}$  gilt i.a.

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}! \tag{1.1.32}$$

Außerdem ist beileibe nicht jede Matrix invertierbar, d.h. haben wir irgendeine Matrix  $\hat{A}$  gegeben, braucht die inverse Matrix  $\hat{A}^{-1}$  nicht zu existieren. Das kennen wir schon von den reellen Zahlen: hier ist auch die 0 bzgl. der Multiplikation nicht invertierbar. Wir werden uns mit der Matrizenrechnung in Abschnitt 1.6 noch ausführlich beschäftigen und diese Fragen genauer studieren.

### 1.1.7 Das Skalarprodukt

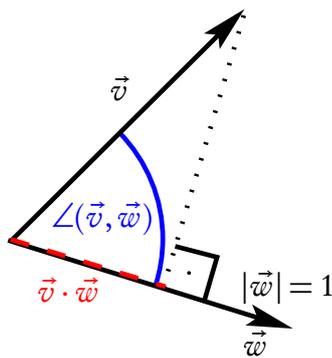
Wir haben nun zwar schon einen sehr beachtlichen Teil der Euklidischen Geometrie in die Sprache der Vektoren übersetzt und damit als „Analytische Geometrie“ für die Physik bequemer handhabbar gemacht. Offensichtlich fehlt aber noch die Behandlung von **Winkeln**. Dazu benötigen wir eine weitere Rechenoperation für zwei Vektoren  $\vec{v}$  und  $\vec{w}$ , das **Skalarprodukt**. In der modernen mathematischen Literatur spricht man auch von einem **inneren Produkt**. Wir geben einfach die Definition des Skalarproduktes an und untersuchen dann seine Eigenschaften:

$$\vec{v} \cdot \vec{w} = |\vec{v}||\vec{w}| \cos[\angle(\vec{v}, \vec{w})]. \tag{1.1.33}$$

Dabei ist der Winkel  $\angle(\vec{v}, \vec{w}) = \angle(\vec{w}, \vec{v}) \in [0, \pi]$  der Winkel zwischen den beiden Vektoren, wenn man sie so verschiebt, daß ihre Anfangspunkte aufeinanderfallen (s. die folgende Abbildung). Da der Cosinus im Intervall  $[0, \pi]$  streng monoton fallend ist, wird durch das Skalarprodukt und die Länge der Vektoren der Winkel eindeutig definiert:

$$\angle(\vec{v}, \vec{w}) = \arccos\left(\frac{\vec{v} \cdot \vec{w}}{|\vec{v}||\vec{w}|}\right), \quad \vec{v}, \vec{w} \neq \vec{0}.$$

Falls mindestens einer der beiden Vektoren im Skalarprodukt der Nullvektor ist, ist der Winkel zwischen diesen Vektoren zwar unbestimmt. Wir definieren daher noch zusätzlich, daß für alle Vektoren  $\vec{w}$  stets  $\vec{0} \cdot \vec{w} = \vec{w} \cdot \vec{0} = 0$  gelten soll. Insbesondere ist natürlich auch  $\vec{0} \cdot \vec{0} = 0$ .



Die geometrische Bedeutung des Skalarproduktes wird verständlich, wenn wir für  $\vec{w}$  einen **Einheitsvektor**, also einen Vektor der Länge 1 wählen. Sei also  $|\vec{w}| = 1$ . Dann ist

$$\vec{v} \cdot \vec{w} = |\vec{v}| \cos[\angle(\vec{v}, \vec{w})] \quad \text{falls } |\vec{w}| = 1. \tag{1.1.34}$$

Dies ist dem Betrag nach gerade die Länge der senkrechten Projektion des Vektors  $\vec{v}$  auf die Richtung von  $\vec{w}$  (vgl. Abbildung). Wegen des cos gilt hinsichtlich des Vorzeichens

$$\vec{v} \cdot \vec{w} \begin{cases} > 0 & \text{falls } \angle(\vec{v}, \vec{w}) \in [0, \pi/2), \\ = 0 & \text{falls } \angle(\vec{v}, \vec{w}) = \pi/2, \\ < 0 & \text{falls } \angle(\vec{v}, \vec{w}) \in (\pi/2, \pi]. \end{cases} \tag{1.1.35}$$

Das Skalarprodukt verschwindet also entweder, wenn  $\vec{v} = \vec{0}$  oder  $\vec{w} = \vec{0}$  ist oder wenn die Vektoren aufeinander senkrecht stehen (denn es ist  $\cos(\pi/2) = \cos 90^\circ = 0$ . Für  $\vec{v} \neq \vec{0}$  und  $\vec{w} \neq \vec{0}$  schreibt man dann  $\vec{v} \perp \vec{w}$  („ $\vec{v}$  steht senkrecht auf  $\vec{w}$ “).

Dem Skalarprodukt eines Vektors mit sich selbst kommt eine besondere Bedeutung zu. Wegen  $\cos 0 = 1$  folgt nämlich

$$\vec{v} \cdot \vec{v} \equiv \vec{v}^2 = |\vec{v}|^2. \tag{1.1.36}$$

Daraus folgt sofort

$$\vec{v} \cdot \vec{v} \geq 0, \quad \vec{v} \cdot \vec{v} = 0 \Leftrightarrow \vec{v} = \vec{0}. \quad (1.1.37)$$

Man sagt daher, daß das Skalarprodukt **positiv definit** ist.

Aus der Definition (1.1.33) ist unmittelbar klar, daß das Skalarprodukt **kommutativ** ist, d.h. es kommt auf die Reihenfolge der Multiplikation nicht an

$$\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}. \quad (1.1.38)$$

Weiter ist es auch linear in beiden Argumenten, d.h. es gilt

$$(\lambda \vec{v}) \cdot \vec{w} = |\lambda| |\vec{v}| |\vec{w}| \cos[\angle(\lambda \vec{v}, \vec{w})]. \quad (1.1.39)$$

Nun gilt aber gemäß der nebenstehenden Abbildung

$$\angle(\lambda \vec{v}, \vec{w}) = \begin{cases} \angle(\vec{v}, \vec{w}) & \text{falls } \lambda > 0, \\ \pi - \angle(\vec{v}, \vec{w}) & \text{falls } \lambda < 0 \end{cases} \quad (1.1.40)$$

ist. Wegen  $\cos(\pi - \alpha) = -\cos \alpha$  folgt also für  $\lambda \neq 0$  aus (1.1.39)

$$(\lambda \vec{v}) \cdot \vec{w} = \text{sign } \lambda |\lambda| |\vec{v}| |\vec{w}| \cos[\angle \vec{v}, \vec{w}] = \lambda (\vec{v} \cdot \vec{w}) \equiv \lambda \vec{v} \cdot \vec{w}. \quad (1.1.41)$$

Außerdem entnimmt man der nebenstehenden Abbildung, daß für  $|\vec{w}| = 1$  auch das Distributivgesetz, also

$$(\vec{v}_1 + \vec{v}_2) \cdot \vec{w} = \vec{v}_1 \cdot \vec{w} + \vec{v}_2 \cdot \vec{w} \quad (1.1.42)$$

gilt. Falls  $\vec{w} = \vec{0}$  ist, gilt diese Gleichung sicher. Für einen Vektor  $\vec{w} \neq \vec{0}$ , der kein Einheitsvektor ist, können wir stets  $\vec{w} = |\vec{w}| \vec{w}/|\vec{w}|$  schreiben. Nun ist  $\vec{w}/|\vec{w}|$  ein Einheitsvektor (*warum?*), und wegen (1.1.41) folgt

$$(\vec{v}_1 + \vec{v}_2) \cdot \vec{w} = |\vec{w}| (\vec{v}_1 + \vec{v}_2) \cdot \frac{\vec{w}}{|\vec{w}|} = |\vec{w}| \left( \vec{v}_1 \cdot \frac{\vec{w}}{|\vec{w}|} + \vec{v}_2 \cdot \frac{\vec{w}}{|\vec{w}|} \right) = \vec{v}_1 \cdot \vec{w} + \vec{v}_2 \cdot \vec{w}. \quad (1.1.43)$$

Wegen des Kommutativgesetzes gilt dies freilich auch, wenn die Summe im zweiten Argument steht.

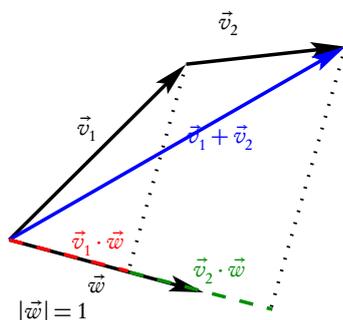
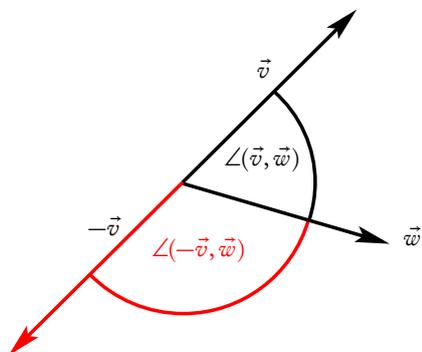
Das Skalarprodukt ist daher auch eine **symmetrische Bilinearform**. Symmetrisch heißt es deshalb, weil das Kommutativgesetz gilt und bilinear, weil es bzgl. beider Argumente eine Lineare Abbildung (in die reellen Zahlen) ist. Wir können nämlich (1.1.41) und (1.1.43) zusammenfassen zu

$$(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2) \cdot \vec{w} = \lambda_1 \vec{v}_1 \cdot \vec{w} + \lambda_2 \vec{v}_2 \cdot \vec{w}. \quad (1.1.44)$$

Mit der Definition des Skalarprodukts ist die Struktur des **Euklidischen Vektorraumes** nunmehr vollständig beschrieben. Ein Vektorraum heißt demnach euklidisch, wenn neben der Vektoralgebra mit den Operationen der Addition von Vektoren und der Multiplikation mit reellen Zahlen auch noch eine positiv definite Bilinearform definiert ist.

### 1.1.8 Geometrische Anwendungen des Skalarprodukts

Wir können nun das Skalarprodukt verwenden, um mit den Mitteln der analytischen Geometrie bekannte Sätze der Geometrie zu beweisen.



Als erstes beweisen wir den **Cosinus-Satz**. Seien drei Punkte  $A, B$  und  $C$  gegeben, die nicht auf einer Geraden liegen. Wir setzen dazu  $\vec{v}_1 = \overrightarrow{AB}$  und  $\vec{v}_2 = \overrightarrow{AC}$ . Dann ist offenbar  $\vec{v}_1 - \vec{v}_2 = \overrightarrow{AB} + \overrightarrow{CA} = \overrightarrow{CB} = \vec{v}_3$ . Für die Länge der Seite  $AC$  gilt demnach

$$|BC|^2 = \vec{v}_3 \cdot \vec{v}_3 = \vec{v}_3^2 = (\vec{v}_1 - \vec{v}_2)^2 = \vec{v}_1^2 + \vec{v}_2^2 - 2\vec{v}_1 \cdot \vec{v}_2 = |AB|^2 + |AC|^2 - 2|AB||AC|\cos\alpha, \quad (1.1.45)$$

wobei  $\alpha = \angle(\overrightarrow{AB}, \overrightarrow{AC})$  der von den Seiten  $\overline{AB}$  und  $\overline{AC}$  eingeschlossene Winkel ist, und das ist der Cosinus-Satz. Dabei haben wir ausgenutzt, daß wir mit dem Vektorprodukt formal wie mit Zahlen rechnen können und insbesondere durch Ausmultiplikation die gewohnten **binomischen Formeln** analog wie bei Zahlen gelten. Falls  $\alpha = \pi/2$ , liegt offenbar ein **rechtwinkliges Dreieck** vor, und dann wird (1.1.45) wegen  $\cos(\pi/2) = 0$

$$|BC|^2 = |AB|^2 + |AC|^2, \quad (1.1.46)$$

und das ist der **Satz des Pythagoras**.

Schließlich gilt wegen  $|\cos\alpha| \leq 1$  stets die **Cauchy-Schwarzsche Ungleichung**

$$|\vec{v}_1 \cdot \vec{v}_2| \leq |\vec{v}_1||\vec{v}_2|. \quad (1.1.47)$$

Das Gleichheitszeichen gilt nur falls  $\cos\alpha = 1$ , d.h.  $\alpha = 0$  (denn definitionsgemäß soll ja  $\alpha \in [0, \pi]$  liegen) oder falls  $\cos\alpha = -1$ , d.h.  $\alpha = \pi$  ist. MaW. das Gleichheitszeichen in (1.1.47) gilt genau dann, wenn  $\vec{v}_1 \parallel \vec{v}_2$  ist.

Wir können nun die **Dreiecksungleichung** aus der positiven Definitheit des Skalarproduktes beweisen, denn es gilt

$$\begin{aligned} |\vec{v}_1 + \vec{v}_2|^2 &= (\vec{v}_1 + \vec{v}_2)^2 \\ &= \vec{v}_1^2 + \vec{v}_2^2 + 2\vec{v}_1 \cdot \vec{v}_2 \\ &\leq \vec{v}_1^2 + \vec{v}_2^2 + 2|\vec{v}_1 \cdot \vec{v}_2| \\ &\leq \vec{v}_1^2 + \vec{v}_2^2 + 2|\vec{v}_1||\vec{v}_2| = (|\vec{v}_1| + |\vec{v}_2|)^2, \end{aligned} \quad (1.1.48)$$

bzw., weil immer nur positive Zahlen quadriert werden,

$$|\vec{v}_1 + \vec{v}_2| \leq |\vec{v}_1| + |\vec{v}_2|. \quad (1.1.49)$$

Umgekehrt folgt aus der positiven Definitheit des Skalarproduktes auch die Cauchy-Schwarzsche Ungleichung (1.1.47).

### 1.1.9 Das Skalarprodukt im $\mathbb{R}^3$

Wegen der Bilinearität können wir das Skalarprodukt zwischen zwei Vektoren auch mit Hilfe der Vektorkomponenten  $\vec{v}$  bzgl. einer beliebigen Basis ausdrücken. Dazu müssen wir allerdings offenbar die Skalarprodukte zwischen den Basisvektoren kennen. Dies ergibt wieder eine Matrix. Wir definieren also die **Metrikkomponenten** bzgl. der gegebenen Basis  $\vec{b}_j$  gemäß

$$g_{jk} = \vec{b}_j \cdot \vec{b}_k. \quad (1.1.50)$$

Für zwei beliebige Vektoren  $\vec{v}$  und  $\vec{w}$  mit den Komponenten  $\vec{v} = (v^j)$  und  $\vec{w} = (w^k)$  folgt dann wegen der Bilinearität des Skalarproduktes

$$\vec{v} \cdot \vec{w} = \left( \sum_{j=1}^3 v^j \vec{b}_j \right) \cdot \left( \sum_{k=1}^3 w^k \vec{b}_k \right) = \sum_{j=1}^3 \sum_{k=1}^3 v^j w^k \vec{b}_j \cdot \vec{b}_k = \sum_{j=1}^3 \sum_{k=1}^3 g_{jk} v^j w^k. \quad (1.1.51)$$

Wir schreiben nun  $(g_{jk})$  als Matrix, machen aber die Tatsache, daß es sich diesmal nicht um die Transformationsmatrix zwischen Basisvektoren sondern um die darstellende Matrix einer Bilinearform (in diesem Fall des Skalarproduktes) handelt, durch einen übergestellten Doppelpfeil kenntlich<sup>2</sup>:

$$\overleftrightarrow{g} = (g_{jk}) = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix}. \quad (1.1.52)$$

Man beachte auch, daß nun *beide* Indizes unten zu sehen kommen.

Wir können nun (1.1.51) mit Hilfe von Matrix-Vektorprodukten schreiben, wenn wir den linken Vektor als Zeile schreiben. Dazu definieren wir

$$\overline{v}^T = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix}^T = (v^1 \quad v^2 \quad v^3). \quad (1.1.53)$$

Man nennt diese Operation an Spaltenvektoren des  $\mathbb{R}^3$  **Transposition**. Diese Operation kann man auch mit Matrizen vornehmen. Man schreibt einfach alle Spalten als Zeilen:

$$\overleftrightarrow{g}^T = \begin{pmatrix} g_{11} & g_{21} & g_{31} \\ g_{12} & g_{22} & g_{32} \\ g_{13} & g_{23} & g_{33} \end{pmatrix} \Leftrightarrow (g^T)_{jk} = g_{kj}. \quad (1.1.54)$$

Da das Skalarprodukt kommutativ ist, gilt allerdings

$$g_{jk} = \vec{b}_j \cdot \vec{b}_k = \vec{b}_k \cdot \vec{b}_j = g_{kj} \Leftrightarrow \overleftrightarrow{g}^T = \overleftrightarrow{g}. \quad (1.1.55)$$

Mit diesen Definitionen folgt, daß

$$\vec{v} \cdot \vec{w} = \overline{v}^T \overleftrightarrow{g} \vec{w} = \overline{w}^T \overleftrightarrow{g} \vec{v}. \quad (1.1.56)$$

Es ist nun einfach, die Metrikkomponenten bzgl. einer beliebigen anderen Basis  $\vec{b}'_k$  durch die der Basis  $\vec{b}_j$  auszudrücken. Dazu benötigen wir lediglich (1.1.24). Wegen der Bilinearität des Skalarproduktes folgt dann

$$g'_{jk} = \vec{b}'_j \cdot \vec{b}'_k = \sum_{l,m=1}^3 U^l_j U^m_k \vec{b}_l \cdot \vec{b}_m = \sum_{l,m=1}^3 U^l_j U^m_k g_{lm}. \quad (1.1.57)$$

In Matrixschreibweise ergibt dies (*Übung!*)

$$\overleftrightarrow{g}' = \hat{U}^T \overleftrightarrow{g} \hat{U}. \quad (1.1.58)$$

## 1.2 Axiomatische Begründung der linearen Algebra und Geometrie

Bislang haben wir die diversen Rechenoperationen mit Vektoren geometrisch motiviert und uns dabei auf die dreidimensionale (räumliche) Euklidische Geometrie berufen. Es ist nun gerade für die theoretische Physik wichtig, diese Betrachtungen auf abstraktere Füße zu stellen. Bereits im nächsten Semester werden wir

<sup>2</sup>Es ist sehr wichtig, eine solche Unterscheidung zu treffen, denn wie wir gleich sehen werden, unterscheidet sich das Transformationsverhalten einer solchen Bilinearform (auch **Tensor zweiter Stufe** genannt) unter einem Basiswechsel entscheidend von dem der Vektorkomponenten. Eigentlich wäre es daher auch adäquater gewesen, eine Unterscheidung bereits bei der Bezeichnung der Komponenten zu treffen. Dies geschieht in der weiterführenden Literatur durch die Verwendung von hoch- und tiefgestellten Indizes. Dies habe ich in diesem Skript vermieden, um die Kompatibilität mit der in der Theorie-Vorlesung gewählten Schreibweise zu wahren.

z.B. die **Spezielle Relativitätstheorie** kennenlernen, die sich am elegantesten und leicht verständlichsten in einem vierdimensionalen Punktraum und entsprechenden vierdimensionalen Vektoren beschreiben läßt, der sog. **Minkowski-Raumzeit**. Die Algebra der Vektoren ist dabei nicht wesentlich komplizierter als die aus der soeben geometrischen Anschauung gewonnenen Rechenregeln für den Euklidischen dreidimensionalen Raum. Nicht zuletzt ist die folgende Zusammenstellung der Axiome auch eine gute Merkhilfe der wichtigsten Formeln.

Wir beginnen mit der axiomatischen Begründung des **Vektorraums**. Ein Vektorraum ist zunächst eine abstrakte Menge  $V$  von Objekten, für die zunächst eine Abbildung  $V \times V \rightarrow V$ , die **Addition** definiert wird. Diese mathematische Pfeilschreibweise für eine Abbildung bedeutet hierbei folgendes: auf der linken Seite steht der **Definitionsbereich** oder **Urbildbereich** der Abbildung. In unserem Fall sind das Paare von Elementen aus der Menge  $V$ . Hier steht  $V \times V$  nämlich für ein **Mengenprodukt**, welches einfach eine neue Menge ist, die aus **geordneten Paaren** von Elementen aus  $V$  besteht. Dabei ist ausdrücklich i.a.  $(\vec{v}_1, \vec{v}_2) \neq (\vec{v}_2, \vec{v}_1)$ , d.h. die Reihenfolge bei dem geordneten Paar soll ausdrücklich berücksichtigt werden. Rechts vom Pfeil steht die **Wertemenge** oder der **Bildbereich**. Für unser Fall besagt dies, daß jedem geordneten Paar von Vektoren aus  $V$  eindeutig ein Vektor, ebenfalls aus  $V$ , zugeordnet wird. Die Addition schreiben wir dabei mit einem Pluszeichen. Dies drückt man formal wie folgt aus  $(\vec{v}_1, \vec{v}_2) \mapsto \vec{v}_1 + \vec{v}_2$ .

Dies mag auf den ersten Blick ein wenig pedantisch und kompliziert erscheinen. Es ist jedoch oft von großem Nutzen, sich klar zu machen, welchen Definitionsbereich und Wertebereich Rechenoperationen eigentlich haben, und dafür ist diese pedantische Schreibweise der Mathematiker äußerst nützlich.

Die Rechenoperationen müssen nun noch weiter spezifiziert werden. Dabei ist es aus mathematischer Sicht unerheblich, um welche konkreten Objekte es sich bei dem, was wir „Vektoren“ nennen, es sich eigentlich handelt. Es werden einfach „Spielregeln“, die **Axiome**, aufgestellt, die formale Operationen festlegen. Für unseren Fall der Addition von Vektoren sind dies die folgenden Axiome:

1. Für beliebige drei Vektoren  $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in E^3$  gilt stets das **Assoziativgesetz**

$$(\vec{v}_1 + \vec{v}_2) + \vec{v}_3 = \vec{v}_1 + (\vec{v}_2 + \vec{v}_3). \quad (1.2.1)$$

2. Es existiert ein **neutrales Element der Addition**, das wir mit  $\vec{0} \in E^3$  bezeichnen. Es zeichnet sich dadurch aus, daß für alle Vektoren  $\vec{v} \in E^3$

$$\vec{v} + \vec{0} = \vec{v} \quad (1.2.2)$$

gilt.

3. Zu jedem Vektor  $\vec{v} \in E^3$  existiert ein bzgl. der Addition **inverses Element**  $(-\vec{v}) \in E^3$ , für das

$$\vec{v} + (-\vec{v}) = \vec{0}. \quad (1.2.3)$$

4. Für beliebige zwei Vektoren  $\vec{v}_1, \vec{v}_2 \in E^3$  gilt das **Kommutativgesetz**

$$\vec{v}_1 + \vec{v}_2 = \vec{v}_2 + \vec{v}_1. \quad (1.2.4)$$

Es führt für diese Physikvorlesung zu weit, den gesamten Apparat der linearen Algebra, die wir oben durch Berufung auf die geometrische Anschauung hergeleitet haben, aus diesen Axiomen zu beweisen. Dies ist Gegenstand der Mathematikvorlesung.

Als ein Beispiel möge der wichtige Satz dienen, daß das neutrale Element der Addition  $\vec{0}$  eindeutig ist. Nehmen wir dazu an, es sei auch  $\vec{0}'$  ein neutrales Element der Addition. Wegen der Eigenschaft (1.2.2) des neutralen Elements  $\vec{0}$  gilt

$$\vec{0}' = \vec{0}' + \vec{0} \stackrel{(1.2.4)}{=} \vec{0} + \vec{0}'. \quad (1.2.5)$$

Nun ist aber  $\vec{0}'$  voraussetzungsgemäß ebenfalls ein neutrales Element, und gemäß (1.2.2) gilt demnach

$$\vec{0}' = \vec{0} + \vec{0}' = \vec{0}, \quad (1.2.6)$$

d.h. daß notwendig die beiden neutralen Elemente der Addition gleich sind.

Wir bemerken noch, daß eine algebraische Struktur, die auf einer Menge wie die eben definierte Addition definiert ist und (1.2.1)-(1.2.3) genügt als **Gruppe** bezeichnet wird. So einfach diese Axiome auch anmuten mögen, so liefern sie doch eine sehr reichhaltige mathematische Struktur. Die Gruppentheorie ist ein riesiges Teilgebiet der Mathematik, und Teile der Gruppentheorie spielen übrigens für die Physik eine herausragende Rolle, die im Verlauf des Physikstudiums noch klar wird. Nimmt man zu diesen drei Axiomen noch das Kommutativgesetz (1.2.4) hinzu, spricht man von einer **Abelschen Gruppe**. Wir können also die Axiome der Vektoraddition auch kurz zusammenfassen, indem wir bemerken, daß die Vektoraddition eine Abelsche Gruppe auf der Menge der Vektoren  $V$  bildet.

Um die so definierte additive Gruppe zu einem Vektorraum zu erweitern, benötigen wir die Multiplikation von Vektoren mit einem Skalar. Dabei sind die Skalare Elemente eines **Zahlenkörpers**, der in unserem Fall stets der Körper der reellen Zahlen  $\mathbb{R}$  ist. Auch diese algebraische Struktur kann wiederum axiomatisch begründet werden. Wir wollen dies hier der Übersichtlichkeit halber aber unterlassen. Die Multiplikation mit Skalaren erfüllt die folgenden Axiome:

1. Für alle  $\lambda, \mu \in \mathbb{R}$  und alle  $\vec{v} \in V$  gilt  $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$  und  $\lambda(\mu\vec{v}) = (\lambda\mu)\vec{v}$ .
2. Für alle  $\vec{v} \in V$  ist  $1\vec{v} = \vec{v}$ .
3. für alle  $\vec{v}, \vec{w} \in V$  und alle  $\lambda \in \mathbb{R}$  gilt  $\lambda(\vec{v} + \vec{w}) = \lambda\vec{v} + \lambda\vec{w}$ .

Man weist leicht einige weitere einfache Rechenregeln nach (*Übung!*), z.B.

$$0\vec{v} = \vec{0}, \quad \lambda\vec{0} = \vec{0}, \quad (-1)\vec{v} = -\vec{v}, \dots \quad (1.2.7)$$

Mit diesen Rechenregeln ist der Begriff des Vektorraumes bereits vollständig bestimmt. Als wichtige Begriffe ergeben sich daraus, wie oben hergeleitet, die **Linearkombinationen** und **Basen** und im Zusammenhang damit der **Dimension**. Wir werden unten als weitere wichtige Strukturen die **linearen Abbildungen** verschiedener Art kennenlernen, deren Untersuchung weitere Eigenschaften der Vektorräume charakterisieren. Es ist übrigens ohnehin ein Merkmal der modernen axiomatischen Methode der Mathematik, Strukturen einzuführen und dann entsprechende diese Strukturen charakterisierende Abbildungen zwischen diesen Strukturen zu analysieren.

Für die Geometrie und damit auch für deren Anwendung in der Physik als Modell für den physikalischen Raum ist weiter die Einführung des Skalarprodukts wichtig. Es handelt sich um eine Abbildung  $V \times V \rightarrow \mathbb{R}$ ,  $\vec{v}, \vec{w} \mapsto \vec{v} \cdot \vec{w}$ , das die folgenden Axiome erfüllt

1. Für alle  $\vec{v}, \vec{w}, \vec{x} \in E^3$  gilt  $\vec{v} \cdot (\vec{w} + \vec{x}) = \vec{v} \cdot \vec{w} + \vec{v} \cdot \vec{x}$ .
2. Für alle  $\lambda \in \mathbb{R}$  und alle  $\vec{v}, \vec{w} \in E^3$  gilt  $(\lambda\vec{v}) \cdot \vec{w} = \lambda(\vec{v} \cdot \vec{w})$ .
3. Für alle  $\vec{v}, \vec{w} \in E^3$  gilt  $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$ .
4. Für alle  $\vec{v} \in E^3$  gilt  $\vec{v}^2 := \vec{v} \cdot \vec{v} \geq 0$ .
5. Es ist  $\vec{v}^2 = 0$  genau dann wenn,  $\vec{v} = \vec{0}$  ist.

Eine Abbildung  $V \times V \rightarrow V$ , welche die ersten drei Axiome erfüllt, heißt **symmetrische Bilinearform**. Gilt auch noch das vierte Axiom, heißt die Bilinearform **positiv semidefinit**. Ist schließlich auch noch das fünfte Axiom erfüllt, heißt die Bilinearform **positiv definit**.

Wir wollen als Beispiel mit dem Umgang des so axiomatisch definierten Skalarprodukts die **Schwarzsche Ungleichung** (1.1.47) beweisen, wobei der **Betrag eines Vektors** durch

$$|\vec{v}| = \sqrt{\vec{v}^2} \geq 0 \quad (1.2.8)$$

definiert ist. Wegen der positiven Definitheit des Skalarprodukts gilt für alle Vektoren  $\vec{v} \neq \vec{0}$  und  $\vec{w} \in E^3$

$$P(\lambda) = \frac{1}{\vec{v}^2} (\lambda \vec{v} + \vec{w})^2 \geq 0. \quad (1.2.9)$$

Dabei dürfen wir wegen der positiven Definitheit (5. Axiom des Skalarprodukts) durch  $\vec{v}^2 \neq 0$  dividieren. Wir können weiter das Skalarprodukt ausmultiplizieren (zur *Übung* mache man sich das anhand der obigen Axiome klar):

$$P(\lambda) = \lambda^2 + 2 \frac{\vec{v} \cdot \vec{w}}{\vec{v}^2} + \frac{\vec{w}^2}{\vec{v}^2} \lambda \geq 0. \quad (1.2.10)$$

Dieses quadratische Polynom können wir nun noch weiter umformen zu

$$P(\lambda) = \left( \lambda + \frac{\vec{v} \cdot \vec{w}}{\vec{v}^2} \right)^2 + \frac{\vec{w}^2}{\vec{v}^2} - \frac{(\vec{v} \cdot \vec{w})^2}{(\vec{v}^2)^2} \geq 0. \quad (1.2.11)$$

Da diese Ungleichung für alle  $\lambda \in \mathbb{R}$  gilt, muß notwendig

$$\frac{\vec{w}^2}{\vec{v}^2} - \frac{(\vec{v} \cdot \vec{w})^2}{(\vec{v}^2)^2} \geq 0 \quad (1.2.12)$$

gelten. Mit  $(\vec{v}^2)^2 > 0$  multipliziert ergibt sich wegen der Definition des Betrages von Vektoren (1.2.8) schließlich

$$(\vec{v} \cdot \vec{w})^2 \leq \vec{v}^2 \vec{w}^2 \Rightarrow |\vec{v} \cdot \vec{w}| \leq |\vec{v}| |\vec{w}|. \quad (1.2.13)$$

Aus dem Beweis folgt auch, daß das Gleichheitszeichen genau dann gilt, wenn es ein  $\lambda \in \mathbb{R}$  gibt, so daß  $P(\lambda) = 0$  ist, also wenn  $\vec{w} = -\lambda \vec{v}$  ist, d.h. wenn die Vektoren **kolinear** sind. Mit der Rechnung (1.1.48) haben wir gezeigt, daß für den via (1.2.8) definierten Betrag die **Dreiecksungleichung** gilt und damit der Betrag die Axiome einer Norm erfüllt, die wir in Abschnitt 1.1.3 besprochen haben. Wegen der Gültigkeit der Schwarzschen Ungleichung (1.2.13) ist es auch sinnvoll, den Winkel zwischen zwei nicht verschwindenden Vektoren gemäß

$$\cos[\angle(\vec{v}, \vec{w})] = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}| |\vec{w}|} \quad (1.2.14)$$

einzuführen und damit die Winkelmessung zu begründen.

Man kann zeigen, daß sich mit diesen Regeln alle Sätze der Euklidischen Geometrie ergeben. Dazu muß man nur noch die Punkte als Menge einführen und postulieren, daß es entsprechend den oben im umgekehrten Sinne aus den Eigenschaften des Euklidischen Punktraumes erschlossenen Regeln für Verschiebungsvektoren  $\vec{AB}$  gelten. Wir wollen auch diese Betrachtungen nicht im einzelnen hier ausführen. Wir bemerken nur, daß ein solches Konstrukt aus einem (reellen endlichdimensionalen) Euklidischen Vektorraum und einer Menge von Punkten als **Euklidischer affiner Punktraum** bezeichnet wird. Eine Verallgemeinerung dieser mathematisch-abstrakten Definition spielt später in der Begründung der **Speziellen Relativitätstheorie** noch eine wichtige Rolle, wo die Beschreibung der Strukturen von Raum und Zeit durch die sog. **Minkowski-Raumzeit** erfolgt. Die Minkowski-Raumzeit entpuppt sich dabei als vierdimensionaler affiner Punktraum, wobei im Vektorraum allerdings kein positiv definites Skalarprodukt sondern eine indefinite nicht ausgeartete Bilinearform definiert ist. Es handelt sich also um einen sog. **pseudo-Euklidischen affinen Punktraum**.

### 1.3 Kartesische Basen und orthogonale Transformationen

Mit der Einführung eines Skalarprodukts gibt es eine besonders bequeme Klasse von **Basen**, die **Orthonormalbasen** oder **Kartesische Basen**. Dazu wählt man als Basisvektoren beliebige drei paarweise zueinander senkrechte Einheitsvektoren, ein sog. **Dreibein**. Anschaulich ist unmittelbar klar, daß es beliebig viele solcher Dreibeine gibt und damit auch beliebig viele Orthonormalbasen. Wie wir gleich sehen werden, vereinfachen sie das Rechnen mit Vektorkomponenten und dem Matrixprodukt bzw. den dazugehörigen Metrikkomponenten sowie die Transformation von einem Orthonormalsystem zu einem anderen erheblich.

Es sei also  $\{\vec{e}_j\}_{j \in \{1,2,3\}}$  eine beliebige Orthonormalbasis. Voraussetzungsgemäß sind diese drei Vektoren auf 1 normiert und stehen paarweise aufeinander senkrecht. Es gilt also

$$g_{jk} = \vec{e}_j \cdot \vec{e}_k = \delta_{jk} = \begin{cases} 1 & \text{falls } j = k, \\ 0 & \text{falls } j \neq k. \end{cases} \quad (1.3.1)$$

Man nenne  $\delta_{jk}$  das **Kronecker-Symbol**. Es ist klar, daß es die Komponenten der Einheitsmatrix repräsentiert. Bzgl. einer kartesischen Basis wird die darstellende Matrix des Skalarproduktes also extrem bequem, denn es gilt gemäß (1.3.1)

$$\overleftrightarrow{g} = \overleftrightarrow{\delta} = \mathbb{1}_3. \quad (1.3.2)$$

Für kartesische Koordinaten gilt also gemäß (1.1.56)

$$\vec{v} \cdot \vec{w} = \bar{v} \cdot \bar{w} = \bar{v}^T \overleftrightarrow{g} \bar{w} = \bar{v}^T \bar{w} = \sum_{j,k=1}^3 \delta_{jk} v^j w^k = \sum_{j=1}^3 v^j w^j = v^1 w^1 + v^2 w^2 + v^3 w^3. \quad (1.3.3)$$

Wir betrachten nun wieder die Frage, wie sich Basiswechsel auswirken. Wir betrachten zunächst die Transformation zwischen den Basisvektoren einer kartesischen Basis  $\{\vec{e}_j\}_{j \in \{1,2,3\}}$  und einer anderen beliebigen Basis  $\{\vec{b}'_j\}_{j \in \{1,2,3\}}$ . Es gilt wieder wie für ganz allgemeine Basen (1.1.19) und (1.1.24):

$$\vec{e}_j = \sum_{k=1}^3 T^k_j \vec{b}'_k, \quad \vec{b}'_k = \sum_{j=1}^3 U^j_k \vec{e}_j, \quad (1.3.4)$$

wobei die Matrizen  $\hat{T}$  und  $\hat{U}$  zueinander invers sind:  $\hat{T}\hat{U} = \hat{U}\hat{T} = \mathbb{1}_3$  bzw.  $\hat{U} = \hat{T}^{-1}$ . Insbesondere sind diese Matrizen notwendig invertierbar, wenn sowohl die  $\vec{e}_j$  als auch die  $\vec{b}'_j$  jeweils eine Basis bilden, also vollständig und linear unabhängig sind.

Für die Vektorkomponenten gelten entsprechend (1.1.20) bzw. (1.1.21) und (1.1.26):

$$\bar{v}' = \hat{T} \bar{v}, \quad \bar{v} = \hat{U} \bar{v}' = \hat{T}^{-1} \bar{v}'. \quad (1.3.5)$$

Die Tatsache, daß die  $\vec{e}_j$  hierbei ein Orthonormalsystem bilden, führt in diesem Fall noch nicht zu irgendwelchen Vereinfachungen. Betrachten wir nun jedoch, wie sich die darstellenden Matrizen für das Skalarprodukt verhalten. Hier ist (1.3.2) entscheidend, denn für kartesische Basen ist demnach die darstellende Matrix einfach die Einheitsmatrix! Die Transformationsvorschrift (1.1.58) vereinfacht sich nämlich zu

$$\overleftrightarrow{g}' = \hat{U}^T \overleftrightarrow{g} \hat{U} = \hat{U}^T \mathbb{1}_3 \hat{U} = \hat{U}^T \hat{U}. \quad (1.3.6)$$

Daraus läßt sich die für spätere Überlegungen sehr wichtige Schlußfolgerung ziehen, daß die Metrikkomponenten bzgl. einer beliebigen Basis eine invertierbare Matrix bilden. Eine Bilinearform, deren Metrikkomponenten diese Eigenschaft besitzen, heißt **nicht ausgeartete Bilinearform**. Das **Skalarprodukt** ist also eine

**nicht ausgeartete positiv definite symmetrische Bilinearform.** Jede Bilinearform mit diesen Eigenschaften kann als Skalarprodukt verwendet werden und definiert dann auf dem jeweils betrachteten Vektorraum genau die geometrische Struktur, die diesen Vektorraum zu einem Euklidischen Vektorraum macht, d.h. sie induziert eine Norm, die die Länge von Vektoren festlegt, den Begriff des Winkels zwischen zwei Vektoren und insbesondere die Eigenschaft, daß Vektoren orthogonal zueinander sein können (also aufeinander senkrecht stehen).

Um nun also zu zeigen, daß

$$\overset{\leftrightarrow}{g}' = (\overset{\leftrightarrow}{g})^{-1} \quad (1.3.7)$$

existiert, müssen wir uns bereits ein wenig in die Matrizenrechnung einarbeiten. Wir wissen, daß aufgrund der Tatsache, daß  $\hat{U}$  und  $\hat{T}$  zueinander inverse Basistransformationen repräsentieren, daß  $\hat{U}$  eine invertierbare Matrix ist und daß  $\hat{U}^{-1} = \hat{T}$  ist. Wir wissen weiter, daß dies impliziert, daß  $\hat{T}\hat{U} = \hat{U}\hat{T} = \mathbb{1}_3$  ist. Dies ist keinesfalls selbstverständlich, weil i.a. die *Matrizenmultiplikation nicht kommutativ* ist. Wenn aber die Matrizen zueinander invers sind, ergibt ihr Produkt unabhängig von der Reihenfolge stets die Einheitsmatrix. Seien nun  $\hat{A}$  und  $\hat{B}$  beide invertierbar. Dann ist auch das Produkt  $\hat{A}\hat{B}$  eine invertierbare Matrix, und es gilt

$$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}. \quad (1.3.8)$$

Man muß hierbei strikt auf die *Reihenfolge* der Matrixmultiplikation achten! Zum Beweis rechnen wir

$$(\hat{A}\hat{B})(\hat{B}^{-1}\hat{A}^{-1}) = \hat{A}(\hat{B}\hat{B}^{-1})\hat{A}^{-1} = \hat{A}\mathbb{1}_3\hat{A}^{-1} = \hat{A}\hat{A}^{-1} = \mathbb{1}_3, \quad (1.3.9)$$

d.h.  $\hat{B}^{-1}\hat{A}^{-1}$  ist tatsächlich die inverse Matrix zu  $\hat{A}\hat{B}$ , wie in (1.3.8) behauptet.

Weiter behaupten wir, daß mit  $\hat{U}$  auch  $\hat{U}^T$  invertierbar ist, und daß

$$(\hat{U}^T)^{-1} = (\hat{U}^{-1})^T = \hat{T}^T \quad (1.3.10)$$

ist. Zum Beweis müssen wir nur das entsprechende Matrixprodukt in Komponenten ausschreiben:

$$(\hat{U}^T\hat{T}^T)_{jl} = \sum_{k=1}^3 (\hat{U}^T)_{jk} (\hat{T}^T)_{kl} = \sum_{k=1}^3 U_{kj} T_{lk} = (\hat{T}\hat{U})_{lj} = \delta_{lj} = \delta_{jl}. \quad (1.3.11)$$

Das bedeutet aber

$$\hat{U}^T\hat{T}^T = \mathbb{1}_3, \quad (1.3.12)$$

womit (1.3.10) bewiesen ist.

Damit ist wegen (1.3.6) aber  $\overset{\leftrightarrow}{g}$  invertierbar. Wenden wir nämlich (1.3.8) und (1.3.10) auf das entsprechende Produkt an, erhalten wir

$$\overset{\leftrightarrow}{g} = (\overset{\leftrightarrow}{g}')^{-1} = (\hat{U}^T\hat{U})^{-1} = \hat{U}^{-1}(\hat{U}^T)^{-1} = \hat{T}\hat{T}^T. \quad (1.3.13)$$

Nun betrachten wir den Fall, daß auch die zweite Basis eine Orthonormalbasis ist, d.h.  $\vec{b}'_j = \vec{e}'_j$  sei auch eine Orthonormalbasis, d.h. es gilt auch  $\overset{\leftrightarrow}{g}' = \mathbb{1}_3$ . Dann gilt aber gemäß (1.3.6)

$$\overset{\leftrightarrow}{g}' = \hat{U}^T\hat{U} = \mathbb{1}_3. \quad (1.3.14)$$

Damit gilt aber für diesen speziellen Fall, d.h. für die Matrix des Basiswechsels zwischen zwei orthonormalen Basen

$$\hat{T} = \hat{U}^{-1} = \hat{U}^T. \quad (1.3.15)$$

Man nennt Matrizen mit dieser Eigenschaft **orthogonale Matrizen**. Natürlich gilt auch für die Umkehrtransformation

$$\hat{T}^{-1} = \hat{U} = (\hat{U}^T)^T \stackrel{(1.3.15)}{=} \hat{T}^T. \quad (1.3.16)$$

Also ist auch  $\hat{T}$  eine orthogonale Matrix.

Es ist auch sehr einfach, diese Transformationsmatrizen zu bestimmen, wenn man die beiden Orthonormalbasen  $\vec{e}_j$  und  $\vec{e}'_k$  gegeben hat. Man muß sich natürlich stets vergewissern, daß beide Basen auch tatsächlich Orthonormalbasen sind! Dann ist nämlich

$$\vec{e}_j = \sum_{k=1}^3 T^k_j \vec{e}'_k. \quad (1.3.17)$$

Dann folgt wegen  $\vec{e}'_l \cdot \vec{e}'_k = \delta_{lk}$

$$\vec{e}'_l \cdot \vec{e}_j = \vec{e}'_l \cdot \sum_{k=1}^3 T^k_j \vec{e}'_k = \sum_{k=1}^3 T^k_j \vec{e}'_l \cdot \vec{e}'_k = \sum_{k=1}^3 T^k_j \delta_{lk} = T^l_j. \quad (1.3.18)$$

Wenden wir nun (1.3.16) an. Daraus folgt

$$U_{jk} = (\hat{T}^T)_{jk} = T^k_j = \vec{e}'_k \cdot \vec{e}_j = \vec{e}_j \cdot \vec{e}'_k. \quad (1.3.19)$$

Das folgt natürlich sofort auch aus

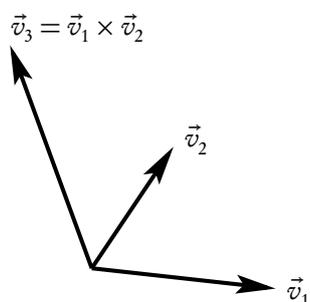
$$\vec{e}'_k = \sum_{j=1}^3 U_{jk} \vec{e}_j \quad (1.3.20)$$

und der Orthonormalität der  $\vec{e}_j$  (*nachrechnen!*).

Wir halten also fest: Die Matrizen für die Basiswechsel zwischen **Orthonormalbasen** sind **orthogonale Matrizen**.

## 1.4 Das Kreuzprodukt

Alle Betrachtungen, die wir bislang angestellt haben, kann man ohne Schwierigkeiten auf beliebige Euklidische Vektorräume verallgemeinern, also auch auf solche, die eine andere Dimension als 3 besitzen. Z.B. kann man sich auf Vektoren in einer Ebene beschränken, und alle Betrachtungen gelten für den entsprechenden Euklidischen Vektorraum in zwei Dimensionen.



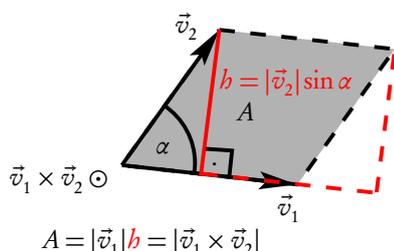
Nun führen wir ein Konzept ein, das es in dieser Form nur für Vektoren in dreidimensionalen Euklidischen Vektorräumen gibt, das **Kreuzprodukt**. Dabei werden zwei Vektoren miteinander Multipliziert, und das Ergebnis ist wieder ein Vektor. Seien also  $\vec{v}_1$  und  $\vec{v}_2$  zwei Vektoren. Dann soll der Betrag des Vektorprodukts  $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$  durch

$$|\vec{v}_1 \times \vec{v}_2| = |\vec{v}_1| |\vec{v}_2| \sin[\angle(\vec{v}_1, \vec{v}_2)] \quad (1.4.1)$$

gegeben sein. Dabei ist  $\angle(\vec{v}_1, \vec{v}_2) \in [0, \pi]$  so definiert wie beim Skalarprodukt. Da für  $\alpha \in [0, \pi]$  stets  $\sin \alpha \geq 0$  ist, ist (1.4.1) sinnvoll, weil die rechte Seite nie negativ wird. Seien nun  $\vec{v}_1$  und  $\vec{v}_2$  nicht parallel. Dann soll die Richtung des Vektorprodukts dadurch bestimmt sein, daß  $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$  senkrecht auf der durch  $\vec{v}_1$  und  $\vec{v}_2$  bestimmten Ebene steht und gemäß der **Rechte-Hand-Regel** orientiert ist, d.h. streckt man den Daumen der rechten Hand in Richtung von  $\vec{v}_1$  und den Zeigefinger in Richtung von  $\vec{v}_2$ , gibt der Mittelfinger die Richtung von  $\vec{v}_3$  an. Dazu

äquivalent ist die **Korkenzieherregel**: Dreht man einen Korkenzieher so, daß  $\vec{v}_1$  auf kürzestem Wege (also um einen Winkel im Intervall  $[0, \pi]$ ) in die Richtung von  $\vec{v}_2$  verdreht wird, bewegt sich der Korkenzieher<sup>3</sup> in Richtung von  $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$ . Wir halten gleich noch fest, daß das Kreuzprodukt mit dem Nullvektor stets wieder den Nullvektor ergeben soll:

$$\vec{v} \times \vec{0} = \vec{0} \times \vec{v} = \vec{0}. \quad (1.4.2)$$



Nun wollen wir uns die geometrische Bedeutung des Betrages (1.4.1) des Kreuzproduktes klar machen. Im nebenstehenden Bild zeichnen wir die beiden Vektoren  $\vec{v}_1$  und  $\vec{v}_2$  ein, wobei die Zeichenebene die durch diese beiden Vektoren definierte Ebene sein soll. Die Richtung des Kreuzproduktes deuten wir in solchen Darstellungen durch einen Kreis mit einem Punkt an, wenn das Kreuzprodukt  $\vec{v}_1 \times \vec{v}_2$  gemäß der Rechte-Hand-Regel aus der Zeichenebene heraus weist. Weist das Kreuzprodukt hingegen in die Zeichenebene hinein, zeichnen wir dies als kleinen Kreis mit einem Kreuz ein. Der Skizze entnehmen wir, daß der Betrag (1.4.1) gerade der

**Fläche** des grau schraffierten durch  $\vec{v}_1$  und  $\vec{v}_2$  aufgespannten Parallelogramms entspricht.

Nun wollen wir die für Produkte üblichen Rechenregeln überprüfen. Betrachten wir zunächst das Produkt in umgekehrter Reihenfolge. Da  $\angle(\vec{v}_2, \vec{v}_1) = \angle(\vec{v}_1, \vec{v}_2)$  ist, ist gemäß (1.4.1)  $|\vec{v}_2 \times \vec{v}_1| = |\vec{v}_1 \times \vec{v}_2|$ . Allerdings kehrt sich bei Anwendung der Rechte-Hand-Regel die Richtung des Kreuzproduktes um. Das Kreuzprodukt ist also **antikommutativ**, d.h. es gilt

$$\vec{v}_2 \times \vec{v}_1 = -\vec{v}_1 \times \vec{v}_2. \quad (1.4.3)$$

Daraus folgt insbesondere, daß das Vektorprodukt eines Vektors mit sich selbst stets verschwindet:

$$\vec{v} \times \vec{v} = -\vec{v} \times \vec{v} \Rightarrow \vec{v} \times \vec{v} = 0. \quad (1.4.4)$$

Es läßt sich auch wieder leicht zeigen (*Übung!*), daß

$$(\lambda \vec{v}_1) \times \vec{v}_2 = \lambda(\vec{v}_1 \times \vec{v}_2) \quad \text{und} \quad \vec{v}_1 \times (\lambda \vec{v}_2) = \lambda(\vec{v}_1 \times \vec{v}_2) \quad (1.4.5)$$

gilt.

Etwas komplizierter gestaltet sich der Beweis des **Distributivgesetzes**. Wir wollen zeigen, daß in der Tat

$$(\vec{v}_1 + \vec{v}_2) \times \vec{v}_3 = \vec{v}_1 \times \vec{v}_3 + \vec{v}_2 \times \vec{v}_3 \quad (1.4.6)$$

gilt.

Zuerst zeigen wir, daß wir jeden Vektor  $\vec{v}$  in einen Anteil parallel und einen senkrecht zu  $\vec{v}_3$  aufspalten können<sup>4</sup>. Zuerst führen wir den **Einheitsrichtungsvektor**

$$\hat{v}_3 = \frac{\vec{v}_3}{|\vec{v}_3|} \quad (1.4.7)$$

ein<sup>5</sup>. Von (1.1.34) wissen wir, daß die vorzeichenbehaftete Länge der Projektion des Vektors  $\vec{v}$  auf die Richtung von  $\vec{v}_3$  durch  $\vec{v} \cdot \hat{v}_3$  gegeben ist. Demnach definieren wir

$$\vec{v}_{||} = (\hat{v}_3 \cdot \vec{v}) \hat{v}_3. \quad (1.4.8)$$

<sup>3</sup>Dabei setze ich einen gewöhnlichen Korkenzieher voraus, der ein „Rechtsgewinde“ hat. Der Schraubensinn ist mathematisch genau durch diese Korkenzieherregel im Sinne des Vektorproduktes definiert. Es gibt natürlich Scherzartikelkorkenzieher mit Linksgewinde, die wir hier aber explizit nicht betrachten!

<sup>4</sup>wir nehmen im folgenden an, daß  $\vec{v}_3 \neq \vec{0}$  ist, denn in diesem Fall ist wegen (1.4.6) sicher richtig, da dann auf beiden Seiten der Nullvektor steht

<sup>5</sup>Hier bedeutet  $\hat{v}_3$  natürlich keine Matrix sondern einen Vektor der Länge 1 in Richtung von  $\vec{v}_3$ .

## 1.4 · Das Kreuzprodukt

Es ist nun klar, daß entsprechend

$$\vec{v}_\perp = \vec{v} - \vec{v}_\parallel \quad (1.4.9)$$

senkrecht auf  $\vec{v}_3$  steht. In der Tat gilt

$$\vec{v}_3 \cdot \vec{v}_\perp = \vec{v}_3 \cdot \vec{v} - \vec{v}_3 \cdot \vec{v}_\parallel = v^3(\hat{v}_3 \cdot \vec{v} - \hat{v}_3 \cdot \vec{v}_\parallel). \quad (1.4.10)$$

Nun ist aber

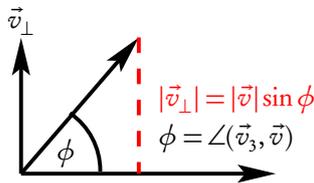
$$\hat{v}_3 \cdot \vec{v}_\parallel = (\hat{v}_3 \cdot \vec{v})\hat{v}_3 \cdot \hat{v}_3 = \hat{v}_3 \cdot \vec{v}, \quad (1.4.11)$$

und daraus folgt wegen (1.4.10), daß tatsächlich

$$\hat{v}_3 \cdot \vec{v}_\perp = 0 \Rightarrow \hat{v}_3 \perp \vec{v}_\perp \quad (1.4.12)$$

ist.

Aus der nebenstehenden Skizze folgt nun, daß



$$\begin{aligned} |\vec{v}_\perp \times \vec{v}_3| &= |v^\perp| |\vec{v}_3| \sin(\pi/2) = |\vec{v}_\perp| |\vec{v}_3| \\ &= |\vec{v}| |\vec{v}_3| \sin[\angle(\vec{v}, \vec{v}_3)] = |\vec{v} \times \vec{v}_3| \end{aligned} \quad (1.4.13)$$

ist. Aufgrund der Rechte-Hand-Regel stimmen auch die Richtungen der Kreuzprodukte überein, d.h. es gilt

$$\vec{v} \times \vec{v}_3 = \vec{v}_\perp \times \vec{v}_3. \quad (1.4.14)$$

Also können wir in (1.4.7) annehmen, daß  $\vec{v}_1 \perp \vec{v}_3$  und  $\vec{v}_2 \perp \vec{v}_3$ , denn es ist (*warum?*)

$$(\vec{v}_1 + \vec{v}_2)_\perp = \vec{v}_{1\perp} + \vec{v}_{2\perp} \quad (1.4.15)$$

und folglich

$$(\vec{v}_1 + \vec{v}_2) \times \vec{v}_3 = (\vec{v}_1 + \vec{v}_2)_\perp \times \vec{v}_3 = (\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \vec{v}_3. \quad (1.4.16)$$

Wir brauchen also das Distributivgesetz nur noch für den Fall zu zeigen, daß  $\vec{v}_1$  und  $\vec{v}_2$  senkrecht auf  $\vec{v}_3$  stehen.

Damit liegen aber die drei Vektoren  $\vec{v}_{1\perp} \times \hat{v}_3$ ,  $\vec{v}_{2\perp} \times \hat{v}_3$  und  $(\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \hat{v}_3$  alle in der zu  $\hat{v}_3$  senkrechten Ebene, die wir in der nebenstehenden Skizze zur Zeichenebene machen. Nun ist aber auch  $\vec{v}_{1\perp} \times \hat{v}_3 \perp \vec{v}_{1\perp}$ ,  $\vec{v}_{2\perp} \times \hat{v}_3 \perp \vec{v}_{2\perp}$  und  $(\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \hat{v}_3 \perp (\vec{v}_{1\perp} + \vec{v}_{2\perp})$ . Da die in diese Kreuzprodukte mit  $\hat{v}_3$  eingehenden Vektoren allesamt senkrecht auf  $\hat{v}_3$  stehen und  $|\hat{v}_3| = 1$  ist, gilt zudem

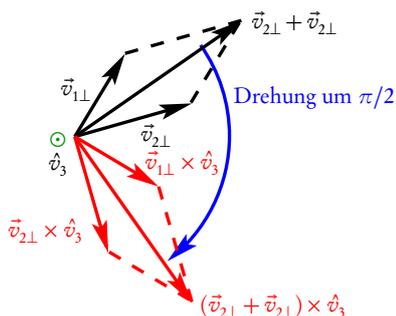
$$\begin{aligned} |\vec{v}_{1\perp} \times \hat{v}_3| &= |\vec{v}_{1\perp}| \sin(\pi/2) = |\vec{v}_{1\perp}|, \\ |\vec{v}_{2\perp} \times \hat{v}_3| &= |\vec{v}_{2\perp}| \sin(\pi/2) = |\vec{v}_{2\perp}|, \\ |(\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \hat{v}_3| &= |\vec{v}_{1\perp} + \vec{v}_{2\perp}| \sin(\pi/2) = |\vec{v}_{1\perp} + \vec{v}_{2\perp}|. \end{aligned} \quad (1.4.17)$$

Damit ist  $\vec{v}_{1\perp} \times \hat{v}_3$  gemäß der Rechte-Hand-Regel einfach der im Uhrzeigersinn um  $\pi/2$  gedrehte Vektor  $\vec{v}_{1\perp}$ . Analoges gilt für die anderen beiden Vektoren (vgl. nebenstehende Abbildung). Damit ist aber gezeigt, daß

$$(\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \hat{v}_3 = \vec{v}_{1\perp} \times \hat{v}_3 + \vec{v}_{2\perp} \times \hat{v}_3. \quad (1.4.18)$$

Wegen  $\vec{v}_3 = |\vec{v}_3| \hat{v}_3$  und (1.4.5) gilt dann aber auch

$$(\vec{v}_{1\perp} + \vec{v}_{2\perp}) \times \vec{v}_3 = \vec{v}_{1\perp} \times \vec{v}_3 + \vec{v}_{2\perp} \times \vec{v}_3. \quad (1.4.19)$$



Wegen (1.4.16) ist damit endlich das Assoziativgesetz (1.4.6) bewiesen.

Es ist wichtig zu bemerken, daß das Assoziativgesetz für das Vektorprodukt i.a. *nicht* gilt:

$$(\vec{v}_1 \times \vec{v}_2) \times \vec{v}_3 \neq \vec{v}_1 \times (\vec{v}_2 \times \vec{v}_3). \quad (1.4.20)$$

Das wird schon aus der Überlegung klar, daß der Vektor auf der linken Seite dieser Ungleichung in der von  $\vec{v}_1$  und  $\vec{v}_2$  aufgespannten Ebene liegt, während sich der auf der rechten Seite in der von  $\vec{v}_2$  und  $\vec{v}_3$  aufgespannten Ebene befindet, und diese beiden Ebenen sind i.a. bereits nicht gleich!

Nun betrachten wir die Frage, wie das Vektorprodukt mittels einer Basis durch die Komponenten der Vektoren, also im Vektorraum von Zahlentripeln  $\mathbb{R}^3$  darstellen läßt. Für allgemeine Basen ist dies i.a. recht schwierig. Für *kartesische Basen* wird dies hingegen recht einfach. Wir nehmen dazu zusätzlich an, daß die Orthonormalbasis  $\vec{e}_j$  rechtshändig sei. Die Numerierung der Vektoren sei also so, daß  $\vec{e}_1 \times \vec{e}_2 = +\vec{e}_3$  ist. Man macht sich anhand der Rechte-Hand-Regel leicht klar, daß dann die drei Beziehungen

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2 \quad (1.4.21)$$

gelten (vgl. die nebenstehende Abbildung). Es ist charakteristisch für das Skalarprodukt, daß die beiden letztgenannten Beziehungen aus der ersten durch **zyklischen Vertauschen der Indizes**  $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$  hervorgehen.

Um nun (1.4.21) symbolisch leichter handhaben zu können, führen wir das im folgenden sehr wichtige **Levi-Civita-Symbol**  $\epsilon_{ijk}$  ein. Es hat drei Indizes, die die Werte  $\{1, 2, 3\}$  annehmen können. Es gilt  $\epsilon_{123} = 1$  und ansonsten ist  $\epsilon_{ijk}$  total antisymmetrisch unter Vertauschungen irgendwelcher zwei Indizes. Damit sind die Werte für  $\epsilon$  eindeutig bestimmt: Sind zwei Indizes gleich, wird es notwendig 0, denn es ist ja wegen der Indexvertauschungsregel z.B.  $\epsilon_{112} = -\epsilon_{112}$ , und daraus folgt notwendig  $\epsilon_{112} = 0$ . Für die in (1.4.21) vorkommenden Indexreihenfolgen ist das Levi-Civita-Symbol immer positiv. Es ist ja z.B.  $\epsilon_{231} = -\epsilon_{213} = +\epsilon_{123} = +1$ . Ein weiteres Beispiel ist  $\epsilon_{321} = -\epsilon_{312} = +\epsilon_{132} = -\epsilon_{123} = -1$ . Die allgemeine Beziehung (1.4.21), die die Rechtshändigkeit des verwendeten Orthonormalsystems charakterisiert schreibt sich also mit Hilfe des Levi-Civita-Symbols als

$$\vec{e}_j \times \vec{e}_k = \epsilon_{jkl} \vec{e}_l = \sum_{l=1}^3 \epsilon_{jkl} \vec{e}_l. \quad (1.4.22)$$

Im letzten Schritt konnten wir die Summe bilden, da ohnehin nur für  $j \neq k$  etwas von 0 Verschiedenes herauskommt. Falls  $j = k$ , stimmt die obigen Gleichungskette ohnehin, da dann auch  $\epsilon_{jkl} = 0$  ist. Falls  $j \neq k$  trägt nur dasjenige  $l$  bei, für das  $l \notin \{j, k\}$  ist, und damit können wir einfach die Summe hinzufügen.

Jetzt ist es leicht, zu ermitteln, wie das Vektorprodukt sich im Sinne der Darstellung der Vektoren im  $\mathbb{R}^3$  bzgl. der rechtshändigen Orthonormalbasis  $\vec{e}_j$  verhält, denn wegen der obigen Rechenregeln des Vektorprodukts folgt

$$\begin{aligned} \vec{v} \times \vec{w} &= \left( \sum_{j=1}^3 v^j \vec{e}_j \right) \times \left( \sum_{k=1}^3 w^k \vec{e}_k \right) = \sum_{j,k=1}^3 v^j w^k (\vec{e}_j \times \vec{e}_k) = \sum_{j,k,l=1}^3 v^j w^k \epsilon_{jkl} \vec{e}_l \\ &= \sum_{l=1}^3 \left( \sum_{j,k=1}^3 v^j w^k \epsilon_{jkl} \right) \vec{e}_l. \end{aligned} \quad (1.4.23)$$

Damit ist aber klar, wie das Skalarprodukt in  $\mathbb{R}^3$  dargestellt wird:

$$(\vec{v} \times \vec{w})_l = \sum_{j,k=1}^3 \epsilon_{jkl} v^j w^k = \sum_{j,k=1}^3 \epsilon_{ljk} v^j w^k. \quad (1.4.24)$$

Wir schreiben auch für die entsprechenden Spaltenvektoren  $\bar{v} \times \bar{w}$ . Dann lautet (1.4.24) ausführlich geschrieben (*es ist sehr lehrreich, sich das ausführlich anhand von (1.4.24) herzuleiten!*)

$$\bar{v} \times \bar{w} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \times \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} = \begin{pmatrix} v^2 w^3 - v^3 w^2 \\ v^3 w^1 - v^1 w^3 \\ v^1 w^2 - v^2 w^1 \end{pmatrix}. \quad (1.4.25)$$

Man beachte, daß man sich nur die erste Zeile dieser Gleichung merken muß, denn die anderen beiden Zeilen folgen daraus durch die schon oben besprochene zyklische Vertauschung der Indizes  $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ . Man muß allerdings bei all dem beachten, daß (1.4.25) nur bzgl. *rechtshändiger Orthonormalbasen* gilt!

Auf das Transformationsverhalten von (1.4.25) können wir erst später in Abschnitt 1.6 eingehen. Dort wird sich zeigen, daß unter solchen orthogonalen Basistransformation, die auch *die Orientierung der Basen ungeändert rechtshändig* lassen, sich die Komponenten von  $\bar{v} \times \bar{w}$  wiederum **wie ein Vektor** transformieren.

Nun beschäftigen wir uns noch mit mehrfachen Vektorprodukten wie  $\vec{u} \times (\vec{v} \times \vec{w})$ . Diese Frage läßt sich geometrisch nicht so ohne weiteres klären. Allerdings wird es in kartesischen Komponenten a la (1.4.24) möglich, obwohl man auch dabei ein wenig mit den Levi-Civita-Symbolen und den Vorzeichen zu kämpfen hat. Wir wenden also (1.4.24) zweimal an:

$$[\vec{u} \times (\vec{v} \times \vec{w})]_j = \sum_{k,l=1}^3 u_k (\vec{v} \times \vec{w})_l \epsilon_{jkl} = \sum_{k,l,m,n=1}^3 u_k v^m w^n \epsilon_{jkl} \epsilon_{lmn}. \quad (1.4.26)$$

Um dies zu vereinfachen, benötigen wir die Summe über  $l$ , die die beiden miteinander multiplizierten Levi-Civita-Symbole betrifft. Wegen  $\epsilon_{lmn} = -\epsilon_{mln} = +\epsilon_{mnl}$  folgt

$$\sum_{l=1}^3 \epsilon_{jkl} \epsilon_{lmn} = \sum_{l=1}^3 \epsilon_{jkl} \epsilon_{mnl}. \quad (1.4.27)$$

Daraus wird klar, daß diese Summe von vornherein nur von 0 verschieden ist, wenn  $j \neq k$  und wenn  $\{j, k\} = \{m, n\}$  ist, weil ansonsten das Produkt aus Levi-Civita-Symbolen sicher verschwindet, weil sonst für alle  $l \in \{1, 2, 3\}$  stets in wenigstens einem dieser Levi-Civita-Symbole zwei gleiche Indizes vorkommen und diese damit verschwinden. Es verbleiben nun nur noch zwei Fälle zu unterscheiden, nämlich (a)  $j = m$  und  $k = n$  und (b)  $j = n$  und  $k = m$ . In Fall (a) trägt für  $j \neq k$  (und dann auch  $m \neq n$ ) genau ein Summand auf, nämlich der, für den  $l \notin \{j, k\} = \{m, n\}$  ist. In dem Fall sind dann aber die *geordneten* Indextripel  $(j, k, l) = (m, n, l)$  und damit ist das Produkt der beiden Levi-Civita-Symbole  $+1$ . Fall (a) trägt also zu der Summe (1.4.27) den Beitrag  $+\delta_{jm} \delta_{kn}$  bei. Ebenso überlegt man sich, daß Fall (b) den Beitrag  $-\delta_{jn} \delta_{km}$  beiträgt. Wir haben also schließlich

$$\sum_{l=1}^3 \epsilon_{jkl} \epsilon_{lmn} = \sum_{l=1}^3 \epsilon_{jkl} \epsilon_{mnl} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}. \quad (1.4.28)$$

Setzen wir dies nun in (1.4.26) ein, können wir sogleich die Summen über  $m$  und  $n$  ausführen:

$$[\vec{u} \times (\vec{v} \times \vec{w})]_j = \sum_{k,m,n=1}^3 u^k v^m w^n (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) = \sum_{k=1}^3 (u^k w^k v^j - u^k v^k w^j). \quad (1.4.29)$$

Nun ist aber gemäß (1.3.3)

$$\sum_{k=1}^3 u^k w^k = \vec{u}^T \vec{w} = \vec{u} \cdot \vec{w}. \quad (1.4.30)$$

Wir können also für (1.4.29)

$$[\vec{u} \times (\vec{v} \times \vec{w})]_j = (\vec{u} \cdot \vec{w}) v^j - (\vec{u} \cdot \vec{v}) w^j \quad (1.4.31)$$

schreiben. Da die Komponenten eines Vektors nun diesen Vektor eindeutig bestimmen, gilt also

$$\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}. \quad (1.4.32)$$

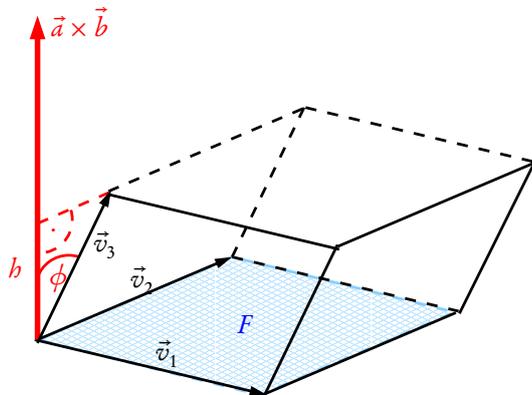
Diese Formel läßt sich wie folgt merken: Der Vektor auf der linken Seite muß in der von  $\vec{v}$  und  $\vec{w}$  aufgespannten Ebene liegen. Entsprechend stehen auf der rechten Seite außerhalb der Skalarprodukte diese Vektoren  $\vec{v}$  und  $\vec{w}$ . Dabei steht der im Tripelprodukt mittlere Vektor  $\vec{v}$  mit dem positiven Vorzeichen auf dieser rechten Seite der Gleichung und entsprechend der letzte Vektor  $\vec{w}$  mit dem negativen.

Wir können aus dieser Regel leicht auch das entsprechend anders geklammerte Vektorprodukt herleiten, denn es gilt

$$(\vec{u} \times \vec{v}) \times \vec{w} = -\vec{w} \times (\vec{u} \times \vec{v}) = -[(\vec{w} \cdot \vec{v})\vec{u} - (\vec{w} \cdot \vec{u})\vec{v}] = (\vec{w} \cdot \vec{u})\vec{v} - (\vec{w} \cdot \vec{v})\vec{u}. \quad (1.4.33)$$

Es gilt also hinsichtlich der Vorzeichen auf der rechten Seite wieder die analoge Merkregel wie für (1.4.32). Wir sehen nun explizit, daß in der Tat i.e. das Kreuzprodukt nicht assoziativ ist.

## 1.5 Das Spatprodukt



$$\text{vol}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = \pm hF$$

Es ist klar, daß wir nun auch kombinierte Produkte bilden können. Von besonderer Bedeutung ist das **Spatprodukt** aus drei Vektoren  $(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3$ . Seine geometrische Bedeutung wird aus der nebenstehenden Zeichnung deutlich. Dazu bemerken wir, daß

$$\text{vol}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = |\vec{v}_1 \times \vec{v}_2| |\vec{v}_3| \cos \phi \quad (1.5.1)$$

mit  $\cos \phi$ ,  $\phi = \angle(\vec{v}_1 \times \vec{v}_2, \vec{v}_3)$ . Aus der Zeichnung wird deutlich, daß es sich vom Betrag her um das Volumen des durch die drei Vektoren aufgespannten **Parallelepipeds** oder **Spats** ist. Das Spatprodukt ist offenbar positiv, wenn die drei Vektoren eine rechtshändige und negativ wenn sie eine linkshändige Basis bilden.

Falls die Vektoren linear abhängig sind, d.h. liegen die Vektoren alle in einer Ebene oder sind sogar alle parallel zueinander, verschwindet das Spatprodukt. Sind nämlich die drei Vektoren linear abhängig, gibt es Zahlen  $\lambda_1$  und  $\lambda_2$ , so daß  $\vec{v}_3 = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2$ . Nun ist das Vektorprodukt  $\vec{v}_1 \times \vec{v}_2$  ein sowohl zu  $\vec{v}_1$  als auch zu  $\vec{v}_2$  senkrechter Vektor, und das Skalarprodukt mit  $\vec{v}_3$  verschwindet demnach. Nehmen wir umgekehrt an, das Spatprodukt der drei Vektoren verschwindet, bedeutet dies, daß  $\vec{v}_3$  senkrecht auf  $\vec{v}_1 \times \vec{v}_2$  liegt, und das besagt, daß  $\vec{v}_3$  in der von  $\vec{v}_1$  und  $\vec{v}_2$  aufgespannten Ebene liegt und also wieder  $\vec{v}_3 = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2$  gilt. Daraus folgt, daß drei Vektoren dann und nur dann linear abhängig sind, wenn das Spatprodukt verschwindet.

Als nächstes zeigen wir, daß man im Spatprodukt „Kreuz und Punkt vertauschen“ dürfen, d.h. daß stets

$$(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = \vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) \quad (1.5.2)$$

gilt. Das ist aufgrund der geometrischen Bedeutung des Spatprodukts klar, denn das Volumen eines Parallelepipedes ist unabhängig davon, welche Fläche man als „Grundfläche“ auffaßt, und das Vorzeichen ist lediglich durch die Reihenfolge als Rechts- oder Linkssystem festgelegt. Wie wir oben gesehen haben verschwindet das Spatprodukt ohnehin, wenn die Vektoren linear abhängig sind.

Wir können (1.5.2) aber auch beweisen, indem wir die Komponentendarstellung von Skalar- und Kreuzprodukt bzgl. einer kartesischen Basis verwenden, also (1.3.3) und (1.4.24):

$$(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = \sum_{j,k,l=1}^3 v_1^j v_2^k v_3^l \epsilon_{jkl} = \sum_{j,k,l=1}^3 v_1^j v_2^k v_3^l \epsilon_{klj} = \vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3). \quad (1.5.3)$$

Durch Vertauschen der Vektoren im Kreuzprodukt unter Beachtung seiner Antisymmetrie können wir auch andere Kombinationen im Spatprodukt erreichen. Es ist klar, daß durch das Vertauschen der drei Vektoren sich nur das Vorzeichen ändert. Es ist eine gute *Übung*, sich klar zu machen, daß stets

$$(\vec{v}_j \times \vec{v}_k) \cdot \vec{v}_l = \epsilon_{jkl} (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 \quad (1.5.4)$$

gilt.

## 1.6 Lineare Gleichungssysteme und Determinanten

Wir verlassen für die nächsten Abschnitte den dreidimensionalen Euklidischen Vektorraum und betrachten das Problem, **lineare Gleichungssysteme** zu lösen, was uns auf die wichtigsten Begriffe der **Matrizenrechnung** führt.

### 1.6.1 Lineare Gleichungssysteme

Ein **lineares Gleichungssystem** stellt uns vor die Aufgabe,  $n$  Gleichungen der Form

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= y_1, \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= y_2, \\ &\vdots \\ A_{n1}x_1 + A_{n2}x_2 + \cdots + A_{nn}x_n &= y_n \end{aligned} \tag{1.6.1}$$

zu lösen. Dabei sind die  $A_{jk} \in \mathbb{R}$  und  $y_j \in \mathbb{R}$  gegeben und die  $x_j$  gesucht. Wir können diese Gleichung kurz mit Hilfe der Matrix-Vektor-Schreibweise formulieren:

$$\hat{A}\bar{x} = \bar{y}, \tag{1.6.2}$$

wobei  $\bar{x}, \bar{y} \in \mathbb{R}^n$  Spaltenvektoren mit  $n$  reellen Komponenten sind und  $\hat{A} \in \mathbb{R}^{n \times n}$  die **quadratische Matrix**

$$\hat{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \tag{1.6.3}$$

bezeichnen. Wir betrachten nun zunächst einige einfache nichttriviale Beispiele für den Fall  $n = 2$ . Beginnen wir mit dem Gleichungssystem

$$\begin{aligned} 3x_1 + x_2 &= 5, \\ x_1 - 3x_2 &= 7. \end{aligned} \tag{1.6.4}$$

Man kann in solch einem einfachen Fall auf vielerlei Arten vorgehen, um das Gleichungssystem zu lösen. Wir wollen hier als systematisches Verfahren den sog. **Gauß-Algorithmus** betrachten. Die grundlegende Strategie ist dabei, durch lineare Operationen an den einzelnen Gleichungen eine Art **Dreiecksform** des Gleichungssystems zu erzielen, so daß man leicht die unbestimmten Variablen nacheinander bestimmen kann. In dem hier vorliegenden zweidimensionalen Fall benötigen wir dazu nur zwei einfache Schritte. Zuerst multiplizieren wir die zweite Gleichung mit 3. Das ergibt

$$\begin{aligned} 3x_1 + x_2 &= 5, \\ 3x_1 - 9x_2 &= 21. \end{aligned} \tag{1.6.5}$$

Diese Strategie hat dazu geführt, daß der Koeffizient vor  $x_1$  in beiden Gleichungen gleich geworden ist. Wir behalten nun die erste Gleichung aus (1.6.4) bei und ersetzen die zweite Gleichung durch die Gleichung, die entsteht, wenn man von ihr die erste Gleichung in (1.6.5) abzieht. Dadurch eliminieren wir  $x_1$  aus dieser Gleichung. Es entsteht

$$\begin{aligned} 3x_1 + x_2 &= 5, \\ -10x_2 &= 16. \end{aligned} \tag{1.6.6}$$

Dividieren wir die letzte Gleichung durch  $(-10)$ , erhalten wir bereits die Lösung  $x_2 = -16/10 = -8/5 = -1,6$ .

brauchen diese Lösung nur in die erste Gleichung einzusetzen und können dann leicht nach  $x_1$  auflösen:

$$3x_1 - 8/5 = 5 \Rightarrow 3x_1 = 5 + 8/5 = 33/5 \Rightarrow x_1 = 11/5 = 2,2. \tag{1.6.7}$$

In dem Fall erhalten wir also eine **eindeutige Lösung**, nämlich  $x_1 = -8/5$ ,  $x_2 = 11/5$ .

Es ist nun offenbar unnötig, bei diesen Rechnungen beständig die unbekanntenen Variablen mitzuschreiben. Stattdessen können wir einfach die Matrix aus den Koeffizienten

$$\hat{A} = \begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix} \quad (1.6.8)$$

und den Spaltenvektor

$$\bar{y} = \begin{pmatrix} 5 \\ 7 \end{pmatrix} \quad (1.6.9)$$

zu der  $\mathbb{R}^{n \times n+1}$ -Matrix

$$\begin{pmatrix} 3 & 1 & 5 \\ 1 & -3 & 7 \end{pmatrix} \quad (1.6.10)$$

ergänzen und mittels linearer Manipulationen mit Zeilen auf die einfache Form bringen, so daß für die Teilmatrix  $\hat{A}$  zunächst nur eine **obere Dreiecksmatrix** übrig bleibt. Dies geschieht genau wie eben mit den Gleichungen demonstriert, indem man die obere Zeile von der mit 3 multiplizierten unteren Zeile abzieht. Daraus entsteht die Matrix

$$\begin{pmatrix} 3 & 1 & 5 \\ 0 & -10 & 16 \end{pmatrix}. \quad (1.6.11)$$

Wir können nun auch systematisch noch die verbliebenen Außerdiagonalargumente der Untermatrix  $\hat{A}$  eliminieren, indem wir geeignete Vielfache bilden und die weiter unten stehende Zeile geeignet abziehen. Hier müssen wir zum 10-fachen der ersten Zeile nur die zweite Zeile addieren und danach schließlich die erste Zeile noch durch 30 und die zweite durch  $-10$  teilen. Daraus entsteht nacheinander

$$\begin{pmatrix} 30 & 0 & 66 \\ 0 & -10 & 16 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2,2 \\ 0 & 1 & -1,6 \end{pmatrix}. \quad (1.6.12)$$

Diese abkürzende Schreibweise bedeutet nun einfach, daß das Gleichungssystem zu dem Gleichungssystem identisch ist, daß wir die beiden ersten Spalten als neue Koeffizientenmatrix  $\hat{A}' = \mathbb{1}_2$  auffassen und die verbliebene dritte Spalte als neuen Vektor  $\bar{y}'$ . Dann ist aber  $\hat{A}'\bar{x} = \bar{x} = \bar{y}'$ . Damit können wir das Ergebnis  $x_1 = -2,2$  und  $x_2 = -1,6$  sofort an dieser Matrix ablesen!

Wir können auf diese Art auch die **Umkehrmatrix** der vorgegebenen Koeffizientenmatrix  $A$  finden. Wir müssen nur die beiden Gleichungen  $\hat{A}\bar{x}_1 = \bar{e}_1 := (1,0)^T$  und  $\hat{A}\bar{x}_2 = \bar{e}_2 := (0,1)^T$  lösen. Dann bilden die Spalten  $\bar{x}_1$  und  $\bar{x}_2$  offensichtlich die Umkehrmatrix, denn dann ist

$$\hat{A}(\bar{x}_1, \bar{x}_2) = (\hat{A}\bar{x}_1, \hat{A}\bar{x}_2) = (\bar{e}_1, \bar{e}_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2. \quad (1.6.13)$$

Statt die beiden Gleichungssysteme einzeln zu lösen können, wir dies wieder simultan erledigen, indem wir zunächst neben die Koeffizientenmatrix  $\hat{A}$  die Einheitsmatrix  $\mathbb{1}_2$  schreiben. Daraus ergibt sich in unserem Beispiel die  $\mathbb{R}^{n \times 2n}$ -Matrix und ausgehend davon die folgende Rechensequenz (die Manipulationen sind wieder dieselben wie eben bei der Lösung des linearen Gleichungssystems

$$\begin{pmatrix} 3 & 1 & 1 & 0 \\ 1 & -3 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & -10 & -1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 30 & 0 & 9 & -3 \\ 0 & -10 & -1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 3/10 & 1/10 \\ 0 & 1 & 1/10 & -3/10 \end{pmatrix} \quad (1.6.14)$$

Die zweite Hälfte dieser Matrix ist demnach die gesuchte Umkehrmatrix

$$\hat{A}^{-1} = \begin{pmatrix} 3/10 & 1/10 \\ 1/10 & -3/10 \end{pmatrix}. \quad (1.6.15)$$

In der Tat rechnet man nach, daß

$$\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3/10 & 1/10 \\ 1/10 & -3/10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2, \quad (1.6.16)$$

wie erwartet.

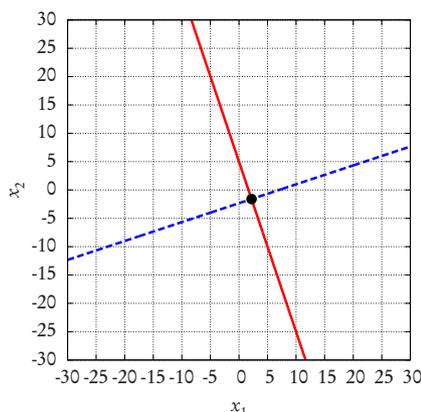
Betrachten wir nun nochmals das Gleichungssystem (1.6.5). Zunächst stellen wir fest, daß in dem gegebenen Fall das Gleichungssystem für jeden Vektor  $\bar{y}$  genau eine Lösung hat, weil nach unserer obigen Rechnung die Koeffizientenmatrix  $\hat{A}$  eine Inverse besitzt, denn es gilt ja

$$\hat{A}\bar{x} = \bar{y} \Leftrightarrow \bar{x} = \hat{A}^{-1}\bar{y}. \quad (1.6.17)$$

Jetzt wollen wir die Gleichungen geometrisch deuten. Dazu stellen wir uns vor, daß  $\bar{x}$  die kartesischen Komponenten eines geometrischen Vektors sind. Dann bedeutet die erste Gleichung

$$3x_1 + x_2 = 5 \Rightarrow x_2 = -3x_1 + 5, \quad (1.6.18)$$

daß die diese Gleichungen erfüllenden Vektoren auf einer **Geraden** liegen, die durch diese Gleichung bestimmt ist. Dabei bedeutet der Koeffizient 3 vor  $x_1$  auf der rechten Seite, daß die Gerade die Steigung 3 besitzt (also monoton von links unten nach rechts oben in der Zeichenebene ansteigt; vgl. die nebenstehende Skizze) und die Konstante 5, daß die Gerade durch den Punkt  $(0;5)$  verläuft, d.h. die  $x_2$ -Achse bei  $x_2 = 5$  schneidet. Ebenso läßt sich die zweite Gleichung als Gerade in der Ebene interpretieren:



$$x_1 - 3x_2 = 7 \Rightarrow x_2 = -\frac{1}{3}x_1 + \frac{7}{3}. \quad (1.6.19)$$

Offenbar bestimmt die Lösung des Gleichungssystems den **Schnittpunkt** dieser beiden Geraden. Nun können sich in der Tat zwei Geraden in der Ebene (und auch im Raum) maximal einmal schneiden, d.h. wenn überhaupt ein Schnittpunkt existiert, dann genau einer. Im vorliegenden Fall schneiden sich die beiden Geraden tatsächlich in dem einen Punkt mit den oben berechneten Koordinaten  $x_1 = 1,6$ ,  $x_2 = -2,2$ .

Nun kann es aber sein, daß die Geraden parallel verlaufen und sich nie schneiden. Dann ist die eine Zeile der Koeffizientenmatrix ein Vielfaches der ersten Zeile, denn genau dann ergibt sich für die durch sie im obigen Sinne definierten Geraden die gleiche Steigung. Es darf dann aber nicht  $\gamma_2$  dasselbe Vielfache von  $\bar{\gamma}_1$  sein. Dann besitzt das Gleichungssystem nämlich offenbar keine Lösung, und die Geraden sind Parallelen. In der Tat besitzen die entsprechenden Geraden dieselbe Steigung aber unterschiedliche Schnittpunkte mit der  $x_2$ -Achse. Ist hingegen die eine Gleichung einfach ein Vielfaches der anderen Gleichung, so definieren offensichtlich beide Gleichungen dieselbe Gerade, d.h. sie liegen aufeinander, und das lineare Gleichungssystem besitzt dann beliebig viele Lösungen, die graphisch gerade durch diese Gleichung bestimmt ist. Der Fall, daß die beiden Zeilenvektoren in der Koeffizientenmatrix linear abhängig sind, führt dazu, daß im Gauß-Algorithmus die zweite Zeile bereits im ersten Schritt identisch verschwindet, und wir daher keine Umkehrmatrix finden können, denn wir können ja nicht mehr mit Hilfe der unteren Zeile die das obere Außerdiagonalelement eliminieren. Für den Fall, daß wir ein konkretes Gleichungssystem lösen, ist entweder nach dem ersten Eliminationsschritt auch das Element in der 3. Spalte zu 0 (dann ergeben sich unendlich viele Lösungen, die im Graphen auf einer Geraden liegen) oder die letzte Spalte verschwindet oder das Element in der 3. Spalte ergibt sich zu  $\neq 0$  (dann gibt es keine Lösung für das Gleichungssystem, und die Geraden im Graphen liegen parallel zueinander und schneiden sich demnach nicht).

Dieses Lösungsverhalten verallgemeinert sich auch auf höhere Dimensionen. Wir können dies mit Hilfe des Gauß-Algorithmus stets im Einzelfall untersuchen, und wir wollen daher darauf nicht genauer eingehen.

Betrachten wir stattdessen noch ein dreidimensionales Beispiel zur Lösung eines linearen Gleichungssystems mit drei Unbekannten:

$$\begin{aligned}x_1 + 2x_2 + 5x_3 &= 12, \\2x_1 + 4x_2 + 2x_3 &= 16, \\3x_1 + 10x_2 + x_3 &= 30.\end{aligned}\tag{1.6.20}$$

Als erstes schreiben wir die erweiterte Matrix für dieses Gleichungssystem auf:

$$\begin{pmatrix}1 & 2 & 5 & 12 \\2 & 4 & 2 & 16 \\3 & 10 & 1 & 30\end{pmatrix}.\tag{1.6.21}$$

Wir können uns die Arbeit etwas erleichtern, wenn wir die zweite Zeile durch 2 dividieren:

$$\begin{pmatrix}1 & 2 & 5 & 12 \\1 & 2 & 1 & 8 \\3 & 10 & 1 & 30\end{pmatrix}.\tag{1.6.22}$$

Jetzt beginnen wir mit den Gauß-Eliminationen, um die Koeffizientenmatrix in obere Dreiecksgestalt zu bringen. Wir starten mit der Elimination der ersten Spalte der beiden letzten Zeilen, indem wir die erste Zeile bzw. die dreifache erste Zeile subtrahieren:

$$\begin{pmatrix}1 & 2 & 5 & 12 \\0 & 0 & -4 & -4 \\0 & 4 & -14 & -6\end{pmatrix}.\tag{1.6.23}$$

Beim Lösen linearer Gleichungen (und auch beim Invertieren von Matrizen) dürfen wir auch Zeilen beliebig vertauschen, weil dies nur der Vertauschung der entsprechenden Gleichungen innerhalb des Gleichungssystems entspricht. Nachdem wir die zweite Zeile durch  $-4$  dividiert haben, vertauschen wir das Resultat mit der letzten Zeile. Dann ergibt sich (nachdem wir auch die dann mittlere Zeile noch durch 2 dividiert haben

$$\begin{pmatrix}1 & 2 & 5 & 12 \\0 & 2 & -7 & -3 \\0 & 0 & 1 & 1\end{pmatrix}.\tag{1.6.24}$$

Damit ist die Koeffizientenmatrix bereits in oberer Dreiecksgestalt. Wir wollen nun auch die oberen Außerdiagonalelemente eliminieren. Dazu ziehen wir erst die zweite von der ersten Zeile ab:

$$\begin{pmatrix}1 & 0 & 12 & 15 \\0 & 2 & -7 & -3 \\0 & 0 & 1 & 1\end{pmatrix}.\tag{1.6.25}$$

Schließlich addieren wir das 7-fache der letzten Zeile zur zweiten und dividieren das Resultat durch 2. Zugleich subtrahieren das 12-fache der letzten Zeile von der ersten Zeile:

$$\begin{pmatrix}1 & 0 & 0 & 3 \\0 & 1 & 0 & 2 \\0 & 0 & 1 & 1\end{pmatrix}.\tag{1.6.26}$$

Damit ist der Lösungsvektor durch die letzte Spalte gegeben:  $\bar{x} = (3, 2, 1)^T$ . Einsetzen in das ursprüngliche Gleichungssystem bestätigt diese Lösung. Diese Probe sollte man zur Kontrolle am Ende der Rechnung stets ausführen!

### 1.6.2 Determinanten als Volumenform

In diesem Abschnitt führen wir **Determinanten** ein. Wir haben bereits Beispiele für Determinanten kennengelernt, ohne daß wir dies explizit erwähnt haben. So ist bereits das in Abschnitt 1.5 eingeführte Spatprodukt das Beispiel für eine Determinante. Gewöhnlich definiert man aber Determinanten nicht als **multilineare Abbildungen** von Vektoren, sondern als Funktionen von **quadratischen Matrizen**. Allerdings stellt das Beispiel des Spatprodukts eine schöne Motivation für Determinanten dar, und zwar gestatten sie die Verallgemeinerung des Begriffs von **orientierten Volumeninhalten** auf beliebige Dimensionen. Beim Spatprodukt haben wir gesehen, daß

$$\text{vol}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 \quad (1.6.27)$$

vom Betrag her das Volumen des von den drei Vektoren bestimmten Parallelepipeds ergibt, und das Vorzeichen bestimmt, ob die Vektoren in der angegebenen Reihenfolge eine rechts- oder linkshändiges System von Basisvektoren darstellen, falls sie linear unabhängig sind. Falls sie linear abhängig sind, verschwindet das Spatprodukt. Betrachten wir nun die Berechnung des Spatprodukts aus den Komponenten der Vektoren bzgl. einer *kartesischen Basis* (1.5.3)

$$(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = \sum_{j,k,l=1}^3 v_1^j v_2^k v_3^l \epsilon_{jkl}, \quad (1.6.28)$$

wird schon nahegelegt, wie dieser Begriff des orientierten Volumens auf beliebige Dimensionen zu verallgemeinern sein wird: Wir müssen für den  $n$ -dimensionalen Euklidischen Raum  $\mathbb{R}^n$  lediglich ein Levi-Civita-Symbol mit  $n$  Indizes einführen, das wie im dreidimensionalen Fall total antisymmetrisch unter Vertauschen dieser Indizes ist und für das

$$\epsilon_{12\dots n} = 1 \quad (1.6.29)$$

gilt. Dann können wir für  $n$  Vektoren die **Volumenform**

$$\text{vol}(\vec{v}_1, \dots, \vec{v}_n) = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 j_2 \dots j_n} v_1^{j_1} v_2^{j_2} \dots v_n^{j_n} \quad (1.6.30)$$

einführen. Dabei heißt dieses Konstrukt **Form**, weil es offensichtlich eine **lineare Abbildung**  $V^n \rightarrow \mathbb{R}$  ist, die vollständig **antisymmetrisch** beim Vertauschen zweier beliebiger Argumente ist (*warum?*). Diese speziellen linearen Abbildungen spielen eine wichtige Rolle. Wir werden uns in dieser Vorlesung aber nur mit der Volumenform beschäftigen.

Wir können weiter  $(\vec{v}_1, \dots, \vec{v}_n) = \hat{A}$  auch als  $\mathbb{R}^{n \times n}$ -Matrix auffassen. Dies führt dann auf die Definition der **Determinante einer Matrix**

$$\det \hat{A} = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 j_2 \dots j_n} A^{j_1}_1 A^{j_2}_2 \dots A^{j_n}_n. \quad (1.6.31)$$

Bleiben wir aber noch bei der geometrischen Bedeutung von (1.6.30). Betrachten wir zunächst den Spezialfall  $n = 2$  der Ebene. Dann ist

$$\text{vol}(\vec{v}_1, \vec{v}_2) = v_1^1 v_2^2 - v_1^2 v_2^1. \quad (1.6.32)$$

Daß dies vom Betrag her die **Fläche** des von  $\vec{v}_1$  und  $\vec{v}_2$  aufgespannten Parallelogramms in der Ebene ist, machen wir uns schnell klar, indem wir diese Betrachtung im dreidimensionalen Raum anstellen. Interpretieren wir also die zweidimensionalen Vektoren als kartesische Komponenten dreidimensionaler Vektoren in der 12-Ebene gemäß

$$\vec{w}_1 = \begin{pmatrix} v_1^1 \\ v_1^2 \\ v_1^3 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} v_2^1 \\ v_2^2 \\ v_2^3 \end{pmatrix}, \quad (1.6.33)$$

erhalten wir für das Kreuzprodukt

$$\vec{w}_1 \times \vec{w}_2 = \begin{pmatrix} 0 \\ 0 \\ v^{11}v^{22} - v^{12}v^{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \text{vol}(\vec{v}_1, \vec{v}_2) \end{pmatrix}. \quad (1.6.34)$$

Nun ist das Kreuzprodukt zweier dreidimensionaler Vektoren der Länge nach der Flächeninhalt des von diesen beiden Vektoren definierten Parallelogramms, und das Vorzeichen bestimmt sich nach der Rechte-Hand-Regel. Es ist offenbar positiv, wenn man den Vektor  $\vec{v}_1$  um den kleineren Drehwinkel entgegen dem Uhrzeigersinn in die Richtung von  $\vec{v}_2$  drehen kann und entsprechend negativ, wenn diese Drehung im Uhrzeigersinn erfolgt. Dies definiert eine Orientierung in der Ebene.

Im  $n$ -dimensionalen Raum mit  $n \geq 4$  ist es etwas schwierig, die Volumenform (1.6.30) geometrisch zu veranschaulichen. Wir können sie aber einfach als Definition des orientierten Volumens eines  $n$ -dimensionalen Parallelepipeds im  $n$ -dimensionalen Raum auffassen.

Wir befassen uns mit der Frage, wie man das Volumen eines  $n$ -dimensionalen Parallelepipeds berechnet, wenn man die entsprechenden Vektorkomponenten bzgl. einer beliebigen nicht-Kartesischen Basis gegeben hat, im nächsten Abschnitt, weil wir dazu die Eigenschaften der Determinanten von Matrizen benötigen.

### 1.6.3 Determinanten von Matrizen

Wir wollen nun einige Rechenregeln für die Determinante einer Matrix herleiten. Wir gehen dabei von der Definition (1.6.31) aus. Aus der Definition des Levi-Civita-Symbols ist sofort klar, daß die Determinante verschwindet, wenn die Spaltenvektoren der Matrix linear abhängig sind.

Man sieht dies für den Fall, daß zwei Spalten sogar gleich sind, sofort ein. Seien beispielsweise die ersten beiden Spalten gleich, d.h. gilt  $A_{j_1} = A_{j_2}$  für alle  $j \in \{1, 2, \dots, n\}$ , so wechselt einerseits wegen  $\epsilon_{j_1 j_2 j_3 \dots j_n} = -\epsilon_{j_2 j_1 j_3 \dots j_n}$  die Determinante ihr Vorzeichen. Da andererseits aber die ersten beiden Spalten gleich sind, ändert sich nichts, wenn man im ursprünglichen Ausdruck einfach die  $j_1$  und  $j_2$  bei den Matrixelementen vertauscht, d.h. die Determinante ändert andererseits ihr Vorzeichen bei diesem Tausch nicht. Es ist also dann  $\det \hat{A} = -\det \hat{A}$ , und daraus folgt notwendig  $\det \hat{A} = 0$ .

Seien nun die Spalten der Matrix linear abhängig. Dann läßt sich eine Spalte (sagen wir die erste) als Linearkombination aller übrigen Spalten darstellen, d.h. es gibt Zahlen  $\lambda_k$  mit  $k \in \{2, \dots, n\}$ , so daß

$$A_{j_1} = \sum_{k=2}^n \lambda_k A_{j_k}. \quad (1.6.35)$$

Dann folgt

$$\det \hat{A} = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} \left( \sum_{k=2}^n A_{j_1 k} \right) A_{j_2 2} \cdots A_{j_n n} = \sum_{k=2}^n \lambda_k \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} A_{j_1 k} A_{j_2 2} \cdots A_{j_n n}. \quad (1.6.36)$$

Im letzten Ausdruck ist aber für jedes  $k \in \{2, \dots, n\}$  die innere Summe die Determinante einer Matrix, mit zwei gleichen Spalten, d.h. jeder dieser Summanden verschwindet, und damit ist  $\det \hat{A} = 0$ , falls die Spalten von  $\hat{A}$  linear abhängig sind.

Nun betrachten wir die Determinante der **transponierten Matrix**. Es gilt ja  $(\hat{A}^T)_{j_k} = A_{kj}$ , und damit ist

$$\det \hat{A}^T = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} (\hat{A}^T)_{j_1 1} \cdots (\hat{A}^T)_{j_n 1} = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} A_{1j_1} \cdots A_{nj_n}. \quad (1.6.37)$$

Nun liefern gemäß der Definition des Levi-Civita-Symbols nur solche Terme einen von 0 verschiedenen Beitrag zur Determinante, für die  $(j_1, \dots, j_n)$  eine **Permutation** (d.h. Anordnung in geänderter Reihenfolge) von

$(1, 2, \dots, n)$  ist, und es wird über alle möglichen dieser Anordnungen summiert. Man kann demnach im letzten Ausdruck in (1.6.37) genauso gut die Produkte auch bzgl. der hinteren Indizes „sortieren“ und über alle Permutationen von  $(1, 2, \dots, n)$  der vorderen Indizes summieren. Das ist aber gerade wieder die Determinante der ursprünglichen Matrix. Wir erhalten also

$$\det \hat{A}^T = \det \hat{A}. \quad (1.6.38)$$

Betrachten wir nun die Determinante eines Produktes zweier  $\mathbb{R}^{n \times n}$ -Matrizen.

Es ergibt sich daraus übrigens sofort, daß sich die Determinante nicht ändert, wenn man beliebige Vielfache einer Spalte zu irgendeiner anderen Spalte oder beliebige Vielfache einer Zeile zu irgendeiner anderen Zeile addiert. Durch Manipulationen wie wir sie beim Gaußalgorithmus verwendet haben, ändert sich also die Determinante der Matrix nicht. Wenn wir allerdings einfach eine Zeile oder Spalte mit einer reellen Zahl multiplizieren, multipliziert sich auch die Determinante mit dieser Zahl. Wir werden weiter unten noch sehen, wie man dies ausnutzen kann, um Determinanten effizient zu berechnen.

Der Definition der Determinante gemäß gilt

$$\det(\hat{A}\hat{B}) = \sum_{j_1, \dots, j_n=1}^n \sum_{k_1, \dots, k_n=1}^n \epsilon_{j_1 \dots j_n} A_{j_1 k_1} B_{k_1 1} A_{j_2 k_2} B_{k_2 2} \cdots A_{j_n k_n} B_{k_n n}. \quad (1.6.39)$$

Wir können nun offenbar die Summenzeichen vertauschen, weil man Summanden in Summen beliebig vertauschen darf. Die Summation über die  $j_i$  betrifft nur Matrixkomponenten der Matrix  $\hat{A}$ . Wir können also zugleich die Matrixelemente der Matrix  $\hat{B}$  aus der Summe über die  $j_i$  ausklammern. Das führt zu

$$\det(\hat{A}\hat{B}) = \sum_{k_1, \dots, k_n=1}^n B_{k_1 1} B_{k_2 2} \cdots B_{k_n n} \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} A_{j_1 k_1} A_{j_2 k_2} \cdots A_{j_n k_n}. \quad (1.6.40)$$

Betrachten wir nun die inneren Summen über die  $j_i$ , sehen wir, daß es sich um Determinanten von Matrizen handelt, die aus den Spalten mit den Nummern  $k_1, k_2, \dots, k_n$  bestehen. Diese Determinanten verschwinden allesamt wenn zwei oder mehr  $k_l$ 's gleich sind, d.h. es ergibt sich nur ein von 0 verschiedener Wert, wenn unter den der  $j_i$ -Summe entsprechenden Determinanten lediglich Vertauschungen der Spalten der ursprünglichen Matrix vorgenommen wurden. Vertauscht man aber zwei Spalten in einer Matrix ändert deren Determinante nur das Vorzeichen. Insgesamt folgt aufgrund dieser Überlegung

$$\sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} A_{j_1 k_1} A_{j_2 k_2} \cdots A_{j_n k_n} = \epsilon_{k_1 k_2 \dots k_n} \det \hat{A}. \quad (1.6.41)$$

Setzen wir das in (1.6.40) ein, folgt

$$\det(\hat{A}\hat{B}) = \det \hat{A} \sum_{k_1, \dots, k_n=1}^n \epsilon_{k_1 k_2 \dots k_n} B_{k_1 1} B_{k_2 2} \cdots B_{k_n n} = (\det \hat{A})(\det \hat{B}). \quad (1.6.42)$$

Wir haben also den Satz bewiesen, daß die Determinante des Produktes zweier Matrizen das Produkt der Determinanten der einzelnen Matrizen ist.

Für **invertierbare** Matrizen folgt daraus sofort, daß

$$\det(\hat{A}\hat{A}^{-1}) = (\det \hat{A})(\det \hat{A}^{-1}) = \det \mathbb{1}_n = 1. \quad (1.6.43)$$

Daraus folgt aber, daß die Determinante einer invertierbaren Matrix stets von 0 verschieden ist und daß

$$\det(A^{-1}) = \frac{1}{\det \hat{A}} \quad (1.6.44)$$

gilt.

Jetzt wollen wir eine rekursive Methode zur Berechnung der Determinante herleiten, den sog. **Entwicklungssatz**, und zwar betrachten wir die Entwicklung nach der **ersten Spalte**. Dazu müssen wir in der Definition der Determinante nur die Summation über  $j_1$  „abseparieren“:

$$\det \hat{A} = \sum_{j_1, \dots, j_n=1}^n \epsilon_{j_1 \dots j_n} A_{j_1 1} A_{j_2 2} \cdots A_{j_n n} = \sum_{j_1=1}^n A_{j_1 1} \sum_{j_2, \dots, j_n=1}^n \epsilon_{j_1 j_2 \dots j_n} A_{j_2 2} A_{j_3 3} \cdots A_{j_n n}. \quad (1.6.45)$$

Betrachten wir nun die innere Klammer. Dies ist nun bis auf das Vorzeichen einfach die Determinante der Matrix, die aus der ursprünglichen Matrix entsteht, wenn man die  $j_1$ -te Zeile und die erste Zeile streicht. Um das Vorzeichen zu erhalten, betrachten wir die Summanden mit  $j_1 = 1$  und  $j_1 = 2$ . Für  $j_1 = 1$  kommt in der inneren Summe  $\epsilon_{1 j_2 \dots j_n}$  zu stehen. Wenn eines der  $j_k$  ( $k \in \{2, \dots, n\}$ ) den Wert 1 annimmt, verschwindet das Levi-Civita-Symbol, und der Summand trägt nichts zur Summe bei, was dem Streichen der  $j_1 = 1$ -sten Zeile entspricht. Andernfalls erhält man offenbar einfach die Determinante der Matrix  $\hat{A}'_{11}$ , also derjenigen  $\mathbb{R}^{(n-1) \times (n-1)}$ -Matrix, die aus der Matrix  $\hat{A}$  durch Streichen der ersten Spalte und der ersten Zeile entsteht. Für  $j_1 = 2$  steht in der inneren Summe das Levi-Civita-Symbol  $\epsilon_{2 j_2 j_3 \dots j_n}$ . Für  $j_2 = 1, j_3 = 3, \dots, j_n = n$ , erhält man den Faktor  $\epsilon_{213 \dots n} = -1$ . Man macht sich klar, daß wegen  $\epsilon_{2 j_2 \dots j_n} = -\epsilon_{j_2 2 j_3 \dots j_n}$  die innere Summe gerade  $-\det \hat{A}'_{12}$  entspricht. Dabei ist  $\hat{A}'_{12}$  die  $(n-1) \times (n-1)$ -Matrix, die aus  $\hat{A}$  durch Streichen der ersten Zeile und der zweiten Spalte entsteht. Für  $j_1 = 3$  trifft ergibt sich wegen  $\epsilon_{3 j_2 \dots j_n} = +\epsilon_{j_1 j_2 3 j_3 \dots j_n}$  hingegen wieder  $+\det \hat{A}'_{13}$  usw. Insgesamt folgt also

$$\det \hat{A} = \sum_{j_1=1}^n (-1)^{j_1+1} A_{j_1 1} \det \hat{A}'_{j_1 1}. \quad (1.6.46)$$

Die Berechnung einer  $\mathbb{R}^{n \times n}$ -Matrix ist damit auf die Berechnung einer  $\mathbb{R}^{(n-1) \times (n-1)}$ -Matrix zurückgeführt. Man überlegt sich auf dieselbe Weise wie eben, daß man auch nach einer beliebigen anderen Spalte entwickeln kann. Die sorgfältige Analyse der Vorzeichen liefert für die **Entwicklung nach der  $i$ -ten Spalte**

$$\det \hat{A} = \sum_{j=1}^n (-1)^{j+i} A_{j i} \det \hat{A}'_{j i}. \quad (1.6.47)$$

Dabei ist  $i \in \{1, 2, \dots, n\}$  beliebig aber während der Rechnung fest gewählt.

Wegen (1.6.38) folgt, daß man genauso gut nach einer beliebigen Zeile entwickeln kann, denn der Entwicklung der Determinante  $\det(\hat{A}^T)$  nach der  $i$ -ten Spalte entspricht einer Entwicklung von  $\det A = \det(\hat{A}^T)$  nach der  $i$ -ten Zeile:

$$\det \hat{A} = \sum_{j=1}^n (-1)^{j+i} A_{i j} \det \hat{A}'_{i j}. \quad (1.6.48)$$

Eine Wichtige Folgerung daraus ist, daß in dem Fall, daß  $\det \hat{A} \neq 0$  die Inverse  $\hat{B}$  der Matrix  $\hat{A}$  existiert, und deren Komponenten durch die Unterdeterminanten wie folgt bestimmt sind:

$$B_{j k} = (\hat{A}^{-1})_{j k} = \frac{(-1)^{j+k}}{\det \hat{A}} \det \hat{A}'_{k j}, \quad (1.6.49)$$

wobei auf die *Reihenfolge* der Indizes der Unterdeterminante zu achten ist: Auf der rechten Seite der Gleichung stehen die Indizes in der umgekehrten Reihenfolge wie auf der linken Seite!

Um diese Behauptung zu beweisen, rechnen wir dies einfach nach. Zunächst ist

$$(\hat{A} \hat{B})_{i k} = \sum_{j=1}^n A_{i j} B_{j k} = \frac{1}{\det \hat{A}} \sum_{j=1}^n (-1)^{j+k} A_{i j} \det \hat{A}'_{k j}. \quad (1.6.50)$$

Ist nun  $i = k$ , so ergibt sich wegen (1.6.48) auf der rechten Seite 1. Falls  $i \neq k$  ist, berechnet man die Determinante der Matrix, die aus  $\hat{A}$  entsteht, wenn man die  $i$ -te Zeile durch die  $k$ -te Zeile ersetzt. In dieser Matrix sind aber zwei Zeilen gleich, und die Determinante verschwindet daher. Damit ist in der Tat gezeigt, daß

$$(\hat{A}\hat{B})_{ik} = \delta_{ik} \Rightarrow \hat{A}\hat{B} = \mathbb{1}_n \quad (1.6.51)$$

ist, und das war zu zeigen.

Auch die Lösung des linearen Gleichungssystems  $\hat{A}\bar{x} = \bar{y}$  im Fall, daß  $\hat{A}$  invertierbar ist, d.h. eine und nur eine Lösung existiert, kann man nun mit Determinanten geschlossen angeben, denn es ist dann offenbar

$$\bar{x} = \hat{A}^{-1}\bar{y}. \quad (1.6.52)$$

Schreiben wir dies in Komponenten aus und verwenden (1.6.49), ergibt sich

$$x_j = \sum_{k=1}^n (\hat{A}^{-1})_{jk} y_k = \frac{1}{\det \hat{A}} \sum_{k=1}^n (-1)^{j+k} \det \hat{A}'_{kj} y_k. \quad (1.6.53)$$

Die Summe ist aber gerade die Determinante derjenigen Matrix, die aus der Koeffizientenmatrix  $\hat{A}$  entsteht, wenn man die  $j$ -te Spalte durch den Spaltenvektor  $\bar{y}$  ersetzt. Schreiben wir die Spaltenvektoren dieser Matrix als

$$\bar{a}_j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{nj} \end{pmatrix}, \quad (1.6.54)$$

so gilt also die **Cramersche Regel**

$$y_j = \frac{\text{vol}(\bar{a}_1, \dots, \bar{a}_{j-1}, \bar{y}, \bar{a}_{j+1}, \dots, \bar{a}_n)}{\det \hat{A}}, \quad (1.6.55)$$

wobei wir uns der Schreibweise (1.6.30) bedient haben. Es ist natürlich klar, daß es in der Praxis wesentlich ökonomischer ist, nicht die  $n+1$  Determinanten zu berechnen, die man benötigt um gemäß (1.6.55) ein lineares Gleichungssystem zu lösen sondern stattdessen das oben beschriebene Gaußsche Eliminationsverfahren zu benutzen.

Schließlich wollen wir noch das Volumen eines  $n$ -dimensionalen Parallelepipeds bzgl. beliebiger nicht notwendig kartesischer Koordinaten herleiten. Seien dazu  $\bar{x}_1, \dots, \bar{x}_n$  die  $n$  Spaltenvektoren aus *kartesischen* Koordinaten der Vektoren  $\vec{x}_1, \dots, \vec{x}_n$ , die das Parallelepiped aufspannen. Dann ist definitionsgemäß dessen Volumen

$$V = \text{vol}(\bar{x}_1, \dots, \bar{x}_n). \quad (1.6.56)$$

Nun seien wieder die Transformationsmatrizen  $\hat{T}$  und  $\hat{U}$  zwischen der kartesischen Basis  $\vec{e}_j$  und der beliebigen anderen Basis  $\vec{b}'_j$  wie in (1.3.4). Dann gilt gemäß (1.3.5)  $\bar{x}_j = \hat{U}\bar{x}'_j$ . Seien nun die  $\hat{X} = (\bar{x}_1, \dots, \bar{x}_n)$  und  $\hat{X}' = (\bar{x}'_1, \dots, \bar{x}'_n)$  die aus den Spaltenvektoren gebildeten Matrizen, so gilt  $\hat{X} = \hat{U}\hat{X}'$  und folglich wegen (1.6.42)

$$V = \det \hat{X} = \det(\hat{U}\hat{X}') = (\det \hat{U})(\det \hat{X}'). \quad (1.6.57)$$

Wegen (1.3.6) gilt nun

$$\det \overleftrightarrow{\hat{g}} = \det(\hat{U}^T \hat{U}) = (\det \hat{U}^T)(\det \hat{U}) = (\det \hat{U})^2. \quad (1.6.58)$$

Dabei haben wir (1.6.38) benutzt. Es gilt also

$$V = \text{sign}(\det \hat{U}) \sqrt{\det \overset{\leftrightarrow}{g'}} \det \hat{X}' = \text{sign}(\det \hat{U}) \sqrt{\det \overset{\leftrightarrow}{g'}} \text{vol}(\bar{x}'_1, \dots, \bar{x}'_n). \quad (1.6.59)$$

Das Vorzeichen der Determinante der Transformationsmatrix bestimmt also, ob die Volumenform bzgl. der neuen Vektorkomponenten das gleiche Vorzeichen besitzt wie die bzgl. der ursprünglichen kartesischen Basis. Man nennt daher Basiswechsel, für die  $\det \hat{U} = 1/\det \hat{T} > 0$  ist **orientierungserhaltende Basistransformationen**. Falls  $\det \hat{U} < 0$  ist, brauchen wir nur irgendwelche zwei Basisvektoren der Basis  $\bar{b}'_j$  miteinander zu vertauschen. Im folgenden nehmen wir stets an, daß alle Basen stets gleichartig zueinander **rechtshändig orientiert** sind.

Betrachten wir nun den für uns wichtigsten Spezialfall, daß die neue Basis  $\bar{b}'_j = \bar{e}'_j$  gleichfalls eine kartesische Basis ist. Dann ist gemäß Abschnitt 1.3 die Matrix des Basiswechsels eine orthogonale Matrix, d.h. es gilt

$$\hat{U}^T = \hat{U}^{-1} \quad (1.6.60)$$

und damit  $\overset{\leftrightarrow}{g}' = \mathbb{1}_n$ . Dann folgt aus (1.6.59), daß die Volumformen bzgl. der beiden Basen wie zu erwarten übereinstimmen. Außerdem ist für orthogonale Matrizen

$$(\det \hat{U})^2 = \det(\hat{U} \hat{U}^T) = \det \mathbb{1}_n = 1 \Rightarrow \det \hat{U} = \det \hat{T} = \pm 1. \quad (1.6.61)$$

Für **orientierungserhaltende orthogonale Matrizen** ist also  $\det \hat{T} = +1$ .

#### 1.6.4 Transformationsverhalten des Kreuzprodukts

Wir untersuchen nun, wie sich die Komponenten des Kreuzprodukts zweier Vektoren im dreidimensionalen Raum unter **orthogonalen Transformationen** verhalten. Von der Anschauung her erwarten wir, daß sich die Komponenten des Kreuzprodukts wie ein Vektor transformiert. Wir werden gleich sehen, daß dies mit einer kleinen Einschränkung tatsächlich zutrifft.

Gemäß (1.4.24) gilt für die Komponenten des Vektorprodukts  $\vec{x} = \vec{v} \times \vec{w}$  bzgl. einer beliebigen *kartesischen* Basis

$$x^j = \sum_{k,l=1}^3 \epsilon_{jkl} v^k w^l. \quad (1.6.62)$$

Für die Komponenten bzgl. einer anderen kartesischen Basis

$$\bar{e}'_a = \sum_{b=1}^3 U^b_a \bar{e}_b \quad (1.6.63)$$

gilt

$$\bar{v} = \hat{U} \bar{v}'. \quad (1.6.64)$$

Daraus folgt

$$x^j = \sum_{k,l=1}^3 \epsilon_{jkl} (\hat{U} \bar{v}')^k (\hat{U} \bar{w}')^l = \sum_{a,b,k,l=1}^3 \epsilon_{jkl} U^k_a U^l_b \bar{v}'^a \bar{w}'^b. \quad (1.6.65)$$

Nun ist aber

$$\sum_{k,l=1}^3 \epsilon_{jkl} U^k_a U^l_b = \sum_{c=1}^3 (-1)^{j+c} \epsilon_{abc} \det \hat{U}'_{jc}, \quad (1.6.66)$$

wobei  $\hat{U}'_{j_c} \in \mathbb{R}^{2 \times 2}$  wieder die durch Streichen der  $j$ -ten Zeile und  $c$ -ten Spalte aus  $\hat{U}$  hervorgehende Untermatrix bezeichnet. Dies macht man sich durch explizites Ausrechnen klar (*Übung*). Nun ist aber für eine orthogonale  $3 \times 3$ -Matrix gemäß (1.6.49)

$$\det \hat{U}'_{j_c} = (-1)^{j+c} \det \hat{U} \hat{U}^{-1)^c_j = (-1)^{j+c} \det \hat{U} (\hat{U}^T)^c_j = (-1)^{j+c} \det \hat{U} U^j_{j_c}. \quad (1.6.67)$$

Setzen wir dies in (1.6.66) ein, erhalten wir schließlich

$$\sum_{k,l=1}^3 \epsilon_{jkl} U^k_a U^l_b = \det \hat{U} \sum_{c=1}^3 \epsilon_{abc} U^j_c. \quad (1.6.68)$$

Verwenden wir dies nun in (1.6.65), folgt damit

$$x^j = \det \hat{U} \sum_{a,b,c=1}^3 \epsilon_{abc} U^j_c v'^a w'^b = \det \hat{U} \sum_{c=1}^3 U^j_c x'^c, \quad (1.6.69)$$

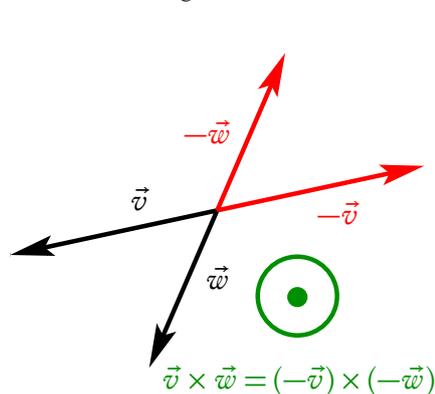
denn für die Komponenten des Vektorprodukts bzgl. der Basis  $(\vec{e}'_c)$  gilt

$$x'_c = \sum_{a,b=1}^3 \epsilon_{abc} v'^a w'^b. \quad (1.6.70)$$

In Matrix-Vektor-Produktschreibweise bedeutet nun aber (1.6.68)

$$\bar{x} = \det \hat{U} \hat{U} \bar{x}' \Rightarrow \bar{v} \times \bar{w} = (\hat{U} \bar{v}') \times (\hat{U} \bar{w}') = \det \hat{U} \hat{U} (\bar{v}' \times \bar{w}'). \quad (1.6.71)$$

Da  $\hat{U}$  eine orthogonale Matrix ist, ist  $\det U = \pm 1$ . Dann besagt (1.6.71), daß sich für **orthogonale Basistransformationen** die Komponenten des Vektorprodukts wie ein Vektor transformieren, wenn die Basistransformation **orientierungserhaltend** ist und andernfalls ein zusätzliches Vorzeichen auftritt. Man nennt solche Größen daher genauer auch **Pseudovektoren** oder **axiale Vektoren**. Im letzteren Zusammenhang bezeichnet man dann die gewöhnlichen Vektoren auch als **polare Vektoren**.



Diese Benennung ergibt Sinn, wenn man als Spezialfall die orthogonale Transformation, die durch  $\hat{U} = -\mathbb{1}$  gegeben ist. Dies bedeutet, daß die neue Orthonormalbasis  $(\vec{e}'_k)$  aus der alten  $(\vec{e}_j)$  hervorgeht, indem man sich die Vektoren in einem gemeinsamen Anfangspunkt befestigt denkt (dem Ursprung des kartesischen Koordinatensystems) und an diesem Punkt spiegelt. Es ist klar, daß dann aus einem rechtshändigen ein linkshändiges Orthonormalbasissystem wird. Irgendwelche zwei polare Vektoren  $\vec{v}$  und  $\vec{w}$  mit Anfangspunkt im Ursprung werden durch diese Spiegelung einfach mit  $(-1)$  multipliziert, d.h.  $\vec{v}' = -\vec{v}$  und  $\vec{w}' = -\vec{w}$ , aber ihr Vektorprodukt wird zu  $\vec{v}' \times \vec{w}' = (-\vec{v}) \times (-\vec{w}) = +\vec{v} \times \vec{w}$ . Dies folgt auch anschaulich aus der Rechten-Hand-Regel (s. die

nebenstehende Skizze). Während also polare Vektoren unter **Raumspiegelungen** ihr Vorzeichen wechseln ist dies für ihr Vektorprodukt nicht der Fall. Das Vektorprodukt zweier polarer Vektoren hat also geometrisch zumindest hinsichtlich ihrer Richtung eher mit einem Drehsinn als mit einer Parallelverschiebung zu tun. Wir werden weiter unten noch sehen, daß das Vektorprodukt eng mit der Beschreibung von Drehungen von Vektoren um eine beliebige Achse zu tun hat.

Es ist klar, daß man (1.6.71) auch in der Form

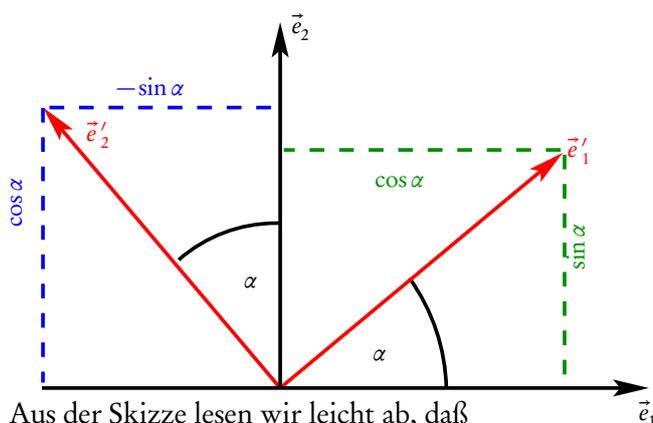
$$\bar{v}' \times \bar{w}' = (\hat{T} \bar{v}) \times (\hat{T} \bar{w}) = \det \hat{T} \hat{T} (\bar{v} \times \bar{w}) \quad (1.6.72)$$

schreiben kann, denn natürlich ist auch  $\hat{T} = \hat{U}^{-1} = \hat{U}^T$  eine orthogonale Matrix.

## 1.7 Drehungen

Wir haben in Abschnitt 1.3 die orthogonalen Transformationen als diejenigen Basistransformationen eingeführt, die eine beliebige Orthonormalbasis in beliebige andere Orthonormalbasen abbilden. Im vorigen Abschnitt haben wir die Transformationen, für deren Transformationsmatrix  $\det \hat{T} = \det \hat{U} = +1$  gilt, als orientierungserhaltend erkannt. Anschaulich ist klar, daß dies geometrisch einer **Drehung** des Koordinatensystems in der Ebene oder im Raum entspricht. In diesem Abschnitt betrachten wir die Drehungen etwas genauer.

### 1.7.1 Drehungen in der Ebene



Aus der Skizze lesen wir leicht ab, daß

$$\begin{aligned}\vec{e}'_1 &= \cos \alpha \vec{e}_1 + \sin \alpha \vec{e}_2, \\ \vec{e}'_2 &= -\sin \alpha \vec{e}_1 + \cos \alpha \vec{e}_2.\end{aligned}\quad (1.7.2)$$

Die Transformationsmatrix ist also

$$\hat{U} = \hat{D}(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.\quad (1.7.3)$$

Es ist nun leicht nachzuprüfen, daß dies in der Tat eine orthogonale Transformation in der Ebene ist, denn es gilt

$$\begin{aligned}\hat{U} \hat{U}^T &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \alpha + \sin^2 \alpha & \cos \alpha \sin \alpha - \sin \alpha \cos \alpha \\ -\cos \alpha \sin \alpha + \sin \alpha \cos \alpha & \sin^2 \alpha + \cos^2 \alpha \end{pmatrix} = \mathbb{1}_2,\end{aligned}\quad (1.7.4)$$

denn für alle  $\alpha$  gilt bekanntlich  $\cos^2 \alpha + \sin^2 \alpha = 1$ . Damit ist  $\hat{U}^T = \hat{U}^{-1}$  und folglich die Transformationsmatrix eine Orthogonalmatrix. Sie ist offensichtlich auch orientierungserhaltend, denn die Entwicklung der Determinante nach der ersten Spalte ergibt

$$\det \hat{U} = +\cos^2 \alpha - (-\sin^2 \alpha) = \cos^2 \alpha + \sin^2 \alpha = +1.\quad (1.7.5)$$

Als nächstes betrachten wir die Hintereinanderausführung zweier Drehungen. Es möge also zuerst die Basis  $(\vec{e}'_1, \vec{e}'_2)$  durch Drehung um den Winkel  $\alpha$  aus der Basis  $(\vec{e}_1, \vec{e}_2)$  und dann  $(\vec{e}''_1, \vec{e}''_2)$  durch Drehung um den Winkel  $\beta$  aus der Basis  $(\vec{e}'_1, \vec{e}'_2)$  hervorgehen. Mit Hilfe der entsprechenden Drehmatrizen geschrieben ergibt sich

$$\vec{e}''_l = \sum_{k=1}^2 D^k_l(\alpha) \vec{e}'_k = \sum_{j,k=1}^2 D^k_l(\alpha) D^j_k(\beta) \vec{e}_j.\quad (1.7.6)$$

Die Transformationsmatrix, die direkt die Transformation von den  $(\vec{e}_j)$  nach den  $(\vec{e}'_l)$  angibt, ist also durch

$$U^j_l = \sum_{k=1}^2 D^j_k(\beta) D^k_l(\alpha) \Rightarrow \hat{U} = \hat{D}(\beta) \hat{D}(\alpha) \quad (1.7.7)$$

gegeben. Anschaulich ist klar, daß  $\hat{U} = \hat{D}(\alpha + \beta)$  sein sollte (vgl. die nebenstehende Skizze). Das können wir nun auch einfach nachrechnen:

$$\begin{aligned} \hat{D}(\beta) \hat{D}(\alpha) &= \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos \beta \cos \alpha - \sin \beta \sin \alpha & -\cos \beta \sin \alpha - \sin \beta \cos \alpha \\ \sin \beta \cos \alpha + \cos \beta \sin \alpha & \cos \beta \cos \alpha - \sin \beta \sin \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} = \hat{D}(\alpha + \beta). \end{aligned} \quad (1.7.8)$$

Dabei haben wir die **Additionstheoreme** der trigonometrischen Funktionen

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta, \quad \sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta \quad (1.7.9)$$

verwendet.

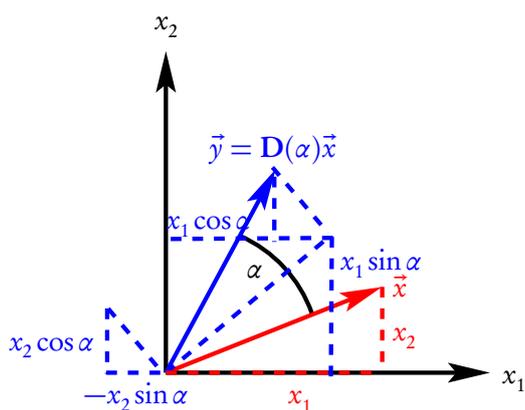
Aus (1.7.8) wird auch sofort klar, daß **Drehungen in der Ebene kommutieren**, d.h. es gilt

$$\hat{D}(\beta) \hat{D}(\alpha) = \hat{D}(\alpha) \hat{D}(\beta). \quad (1.7.10)$$

Aus (1.7.8) folgt auch, daß

$$\hat{D}(-\alpha) \hat{D}(\alpha) = \hat{D}(0) = \mathbb{1}_2 \quad (1.7.11)$$

ist. Es ist daher sinnvoll, negative Drehwinkel als Drehungen im Uhrzeigersinn (mathematisch negativer Drehsinn) zu betrachten. Andererseits ist klar, daß solche Drehungen auch als Drehungen im positiven Sinne um den Komplementärwinkel  $2\pi - |\alpha|$  angesehen werden können. Alle Drehungen in der Ebene werden also durch die Matrizen  $\hat{D}(\alpha)$  mit  $\alpha \in [0, 2\pi)$  erfaßt.



Wir betrachten nun noch Drehungen in einem etwas anderen Sinne, und zwar als **lineare Abbildung** von Vektoren. Lineare Abbildungen wollen wir mit einem aufrechten fett gedruckten Symbol bezeichnen. Dann soll die Drehung  $\mathbf{D}(\alpha)$  um den (positiven) Drehwinkel  $\alpha$  einer Drehung entgegen dem Uhrzeigersinn um diesen Winkel bedeuten. Setzen wir also

$$\vec{y} = \mathbf{D}(\alpha)\vec{x}, \quad (1.7.12)$$

so folgt aus der nebenstehenden Skizze, daß für die Komponenten in einem positiv orientierten kartesischen Koordinatensystem

$$\bar{y} = \hat{D}(\alpha)\bar{x} \quad (1.7.13)$$

gilt.

Manchmal bezeichnet man die Drehungen, die auftreten, wenn man die Transformation von **Vektorkomponenten** bei Drehungen einer kartesischen Basis in eine andere betrachtet als **passive Drehungen**, denn in dem Fall bleiben die Vektoren  $\vec{v}$  fest, und nur die Komponenten ändern sich, weil man diesen Vektor durch Komponenten bzgl. verschiedener Basen betrachtet. Es geschieht also bei dieser Auffassung von Drehungen

nichts mit dem Vektor als geometrischem Objekt sondern nur etwas hinsichtlich der Darstellung als  $\mathbb{R}^2$ -Spaltenvektor aufgrund der Änderung der Basis.

Dazu bemerken wir, daß diese beiden Arten von Transformationen aus Sicht der Physik durchaus unterschiedliche Bedeutung besitzen: Wie nämlich im Lauf der Theorie-Vorlesungen noch klar werden wird, spielen die **Symmetrien der Naturgesetze** eine große Rolle in der Physik. Dies wird bereits an dieser Stelle deutlich: Allein dadurch, daß wir den physikalischen Raum als euklidischen affinen Raum modellieren, implizieren wir, daß dieser Raum bestimmte Symmetrien besitzt. So läßt sich, rein mathematisch betrachtet, kein Ort vor irgendeinem anderen Ort auszeichnen. Wir können geometrische Figuren beliebig parallel zueinander verschieben, ohne daß sich deren wesentliche Eigenschaften, wie z.B. die Längen und Winkel eines Dreiecks, irgendwie verändern. Diese Symmetrie unter **Verschiebungen oder Translationen** bezeichnet man als **Homogenität des Raumes**. Daran ist bemerkenswert, daß zunächst einmal eine solche Transformation aufgrund der mathematischen Struktur überhaupt definiert ist und zum anderen auch bestimmte Eigenschaften von mathematischen Objekten bei einer solchen Transformation ungeändert bleiben. Genauso verändern sich Eigenschaften wie Längen und Winkel auch nicht bei **Drehungen**. Diese Symmetrie bezeichnen wir als **Isotropie des Raumes**. Betrachten wir nun über diese rein mathematischen Gegebenheiten hinausgehend auch **physikalische Gesetzmäßigkeiten**, wie z.B. hier in der Newtonschen Mechanik: Soll nun die Beschreibung des physikalischen Raumes als Euklidischer affiner Raum vollständig und streng korrekt sein, sollten auch die **Naturgesetze**, z.B. die Gleichungen, die die Bewegung von Partikeln beschreiben, unabhängig davon sein, wo man eine solche Bewegung beobachtet und wie man die Versuchsanordnung relativ zu irgendeinem willkürlich gewählten Bezugssystem orientiert. Anders betrachtet bedeutet dies, daß es unmöglich ist, aufgrund des Verhaltens irgendwelcher physikalischer Objekte irgendeinen Ort oder irgendeine Orientierung auszuzeichnen. Dies ist die **aktive Auffassung** von Symmetrien. Andererseits ist es von vornherein klar, daß die Wahl einer Orthonormalbasis und eines Koordinatenursprungs im Raum und die Zuordnung von Koordinaten zu physikalischen Vektoren nichts an den Aussagen der Naturgesetze ändern darf, damit diese Beschreibung überhaupt sinnvoll als Beschreibung eines Naturvorgangs gelten kann. Die Naturgesetze bleiben demnach auch ungeändert, wenn man ein neues Koordinatensystem einführt, das sich lediglich durch eine Verschiebung und/oder eine Drehung der Orthonormalbasis unterscheidet. Dies ist die **passive Auffassung** von Symmetrieprinzipien.

## 1.7.2 Drehungen im Raum

Wir betrachten nun Drehungen im Raum. Anschaulich ist klar, daß wir jede Drehung im Raum durch eine **Drehachse**, die wir durch einen Einheitsvektor  $\vec{n}$  festlegen, und einen **Drehwinkel**  $\varphi \in [0, \pi]$  im Raum eindeutig beschreiben können. Die Drehung ist dann wieder durch die **Rechte-Hand-Regel** festgelegt: Streckt man den Daumen in Richtung von  $\vec{n}$ , zeigen die gekrümmten Finger in die Richtung der Drehung. Hat man die Richtung so gewählt, daß man einen Drehwinkel  $> \pi$  benötigen würde, kann man offenbar einfach statt  $\vec{n}$  die Drehachse umorientieren, also stattdessen  $-\vec{n}$  als Drehachse wählen, und beschreibt dann die gleiche Drehung mit einem Drehwinkel  $< \pi$ . Die entsprechende lineare Abbildung bezeichnen wir mit  $D_{\vec{n}}(\varphi)$ .

Wir wollen uns nun überlegen, wie wir diese (aktive) Drehung von Vektoren formelmäßig für die kartesischen Koordinaten  $\vec{v}$  eines beliebigen Vektors  $\vec{v}$  bzgl. einer rechtshändigen Orthonormalbasis  $(\vec{e}_j)$  beschreiben können. Es ist klar, daß sich nur die Komponenten des Vektors, die senkrecht zu  $\vec{n}$  stehen, ändern, während die Komponente in Richtung von  $\vec{n}$  ungeändert bleibt. Wir können nun wie in (1.4.8) und (1.4.9) gezeigt, den Vektor  $\vec{v}$  in seinen Anteil parallel und senkrecht zu  $\vec{v}$  zerlegen:

$$\vec{v}_{\parallel} = (\vec{n} \cdot \vec{v})\vec{n}, \quad \vec{v}_{\perp} = \vec{v} - \vec{v}_{\parallel}. \quad (1.7.14)$$

Nun gilt

$$D_{\vec{n}}\vec{v}_{\parallel} = \vec{v}_{\parallel}. \quad (1.7.15)$$

Falls  $\vec{v} = \vec{v}_{\parallel}$  ist, sind wir dann schon fertig: Der Vektor zeigt dann in Richtung der Drehachse und ändert sich

daher unter der betrachteten Drehung gar nicht. Falls  $\vec{v}_{\text{perp}} \neq \vec{0}$  ist, so können wir eine neue rechtshändige Orthonormalbasis  $(\vec{e}'_k)$  einführen, so daß  $\vec{e}'_3 = \vec{n}$  ist. Den Vektor  $\vec{e}'_1$  wählen wir

$$\vec{e}'_2 = \frac{\vec{n} \times \vec{v}}{|\vec{n} \times \vec{v}|}. \quad (1.7.16)$$

Dies ist stets eine wohldefinierte Gleichung, denn  $\vec{n} \times \vec{v} \neq 0$ , weil  $\vec{v} \neq \lambda \vec{n}$  ist. Offenbar sind nun  $\vec{e}'_1$  und  $\vec{e}'_3 = \vec{n}$  aufeinander senkrecht stehende Vektoren, und wir müssen weiter

$$\vec{e}'_1 = \vec{e}'_2 \times \vec{e}'_3 \quad (1.7.17)$$

setzen, damit  $(\vec{e}'_j)$  eine rechtshändige Orthonormalbasis wird. Bzgl. dieser Basis kennen wir nun aber bereits das Resultat der Drehung, denn es handelt sich offenbar einfach um die Drehung von  $\vec{v}_{\perp}$  in der Ebene Senkrecht zu  $\vec{e}'_3$ , also in der  $1'2'$ -Ebene. Mit Hilfe von (1.4.32) erhalten wir

$$\vec{e}'_1 = \frac{(\vec{n} \times \vec{v}) \times \vec{n}}{|\vec{n} \times \vec{v}|} = \frac{\vec{v}(\vec{n} \cdot \vec{n}) - \vec{n}(\vec{n} \cdot \vec{v})}{|\vec{n} \times \vec{v}|} = \frac{\vec{v}_{\text{perp}}}{|\vec{n} \times \vec{v}|} = \frac{\vec{v}_{\perp}}{|\vec{v}_{\perp}|}. \quad (1.7.18)$$

Dabei haben wir benutzt, daß  $(\vec{n} \times \vec{v}) \perp \vec{n}$  und  $|\vec{n}| = 1$  ist, d.h. daß

$$|\vec{n} \times \vec{v}| = |\vec{v}_{\perp}| \quad (1.7.19)$$

ist. Weiter ist

$$v'^1 = \vec{e}'_1 \cdot \vec{v} = |\vec{v}_{\perp}|, \quad v'^2 = \vec{e}'_2 \cdot \vec{v} = 0, \quad v'^3 = \vec{e}'_3 \cdot \vec{v} = \vec{n} \cdot \vec{v}. \quad (1.7.20)$$

Sei nun  $\vec{w} = \mathbf{D}_{\vec{n}}(\varphi)\vec{v}$ . Dann gilt offenbar  $w'^3 = v'^3$ , und

$$\begin{pmatrix} w'^1 \\ w'^2 \end{pmatrix} = \hat{D}(\varphi) \begin{pmatrix} |\vec{v}_{\perp}| \\ 0 \end{pmatrix} = |\vec{v}_{\perp}| \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}. \quad (1.7.21)$$

Dabei haben wir uns der Drehungen in der  $1'2'$ -Ebene (1.7.13) bedient. Nun finden wir

$$\vec{w} = \mathbf{D}_{\vec{n}}(\varphi)\vec{v} = w'^1 \vec{e}'_1 + w'^2 \vec{e}'_2 + w'^3 \vec{e}'_3 = |\vec{v}_{\perp}| \left( \cos \varphi \frac{\vec{v}_{\perp}}{|\vec{v}_{\perp}|} + \sin \varphi \frac{\vec{n} \times \vec{v}}{|\vec{v}_{\perp}|} \right) + (\vec{n} \cdot \vec{v})\vec{n}. \quad (1.7.22)$$

Wegen (1.7.18) ist aber

$$\vec{v}_{\perp} = |\vec{v}_{\perp}| \vec{e}'_1 = (\vec{n} \times \vec{v}) \times \vec{n}, \quad (1.7.23)$$

und wir erhalten schließlich die geschlossene Formel für die (aktive) Drehung eines Vektors um die durch  $\vec{n}$  gegebene Achse um den Drehwinkel  $\varphi$ :

$$\begin{aligned} \vec{w} &= \mathbf{D}_{\vec{n}}(\varphi)\vec{v} = (\vec{n} \cdot \vec{v})\vec{n} + \cos \varphi (\vec{n} \times \vec{v}) \times \vec{n} + \sin \varphi \vec{n} \times \vec{v} \\ &= (\vec{n} \cdot \vec{v})(1 - \cos \varphi)\vec{n} + \cos \varphi \vec{v} + \sin \varphi \vec{n} \times \vec{v}. \end{aligned} \quad (1.7.24)$$

Wir wollen dies nun noch durch die Komponenten bzgl. einer beliebigen rechtshändigen Orthonormalbasis  $(\vec{e}_j)$  ausdrücken. Dazu müssen wir nur die einzelnen Terme ausschreiben. Zunächst gilt

$$[(\vec{n} \cdot \vec{v})\vec{n}]_j = \sum_{k=1}^3 n_j n_k v^k. \quad (1.7.25)$$

Für die entsprechende Matrix, die die Projektion von  $\vec{v}$  auf die Richtung des Einheitsvektors  $\vec{n}$  beschreibt, schreiben wir

$$\hat{P}_{\parallel}(\vec{n}) = \vec{n} \otimes \vec{n} = (n_j n_k) = \begin{pmatrix} n_1^2 & n_1 n_2 & n_1 n_3 \\ n_2 n_1 & n_2^2 & n_2 n_3 \\ n_3 n_1 & n_3 n_2 & n_3^2 \end{pmatrix}. \quad (1.7.26)$$

Die Projektion auf den Anteil eines Vektors senkrecht zu  $\vec{n}$  ist entsprechend

$$\hat{P}_\perp = \mathbb{1}_3 - \hat{P}_\parallel(\vec{n}). \quad (1.7.27)$$

Man rechnet leicht nach, daß aufgrund dieser Definitionen die Rechenregeln (*nachrechnen!*)

$$\hat{P}_\parallel(\vec{n})\hat{P}_\parallel(\vec{n}) = \hat{P}_\parallel(\vec{n}), \quad \hat{P}_\perp(\vec{n})\hat{P}_\perp(\vec{n}) = \hat{P}_\perp(\vec{n}), \quad \hat{P}_\perp(\vec{n})\hat{P}_\parallel(\vec{n}) = \hat{P}_\parallel(\vec{n})\hat{P}_\perp(\vec{n}) = \hat{0}, \quad (1.7.28)$$

wobei die **Nullmatrix**  $\hat{0}$  die Matrix mit lauter Nullen als Matrixelementen bezeichnet.

Schließlich ist noch

$$(\vec{n} \times \vec{v})^j = \sum_{k=1}^3 \epsilon_{jkl} n^k v^l, \quad (1.7.29)$$

d.h. wir benötigen die Matrix  $\hat{\epsilon}(\vec{n})$  mit

$$[\epsilon_{jkl}(\vec{n})] = \sum_{k=1}^3 \epsilon_{jkl} n^k = \begin{pmatrix} 0 & -n^3 & n^3 \\ n^3 & 0 & -n^1 \\ -n^2 & n^1 & 0 \end{pmatrix}. \quad (1.7.30)$$

Mit diesen Definitionen ergibt sich aus (1.7.24) für die entsprechende Gleichung in Komponenten

$$\vec{w} = \hat{D}_{\vec{n}}(\varphi) \quad \text{mit} \quad \hat{D}_{\vec{n}}(\varphi) = \hat{P}_\parallel(\vec{n}) + \cos \varphi \hat{P}_\perp(\vec{n}) + \sin \varphi \hat{\epsilon}(\vec{n}), \quad (1.7.31)$$

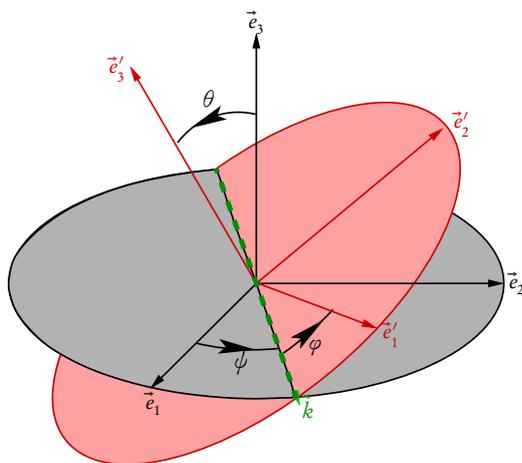
wobei wir die Multiplikation einer Matrix mit einer Zahl und die Addition zweier Matrizen komponentenweise verstehen:

$$\lambda \hat{A} = \lambda \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} \lambda A_{11} & \lambda A_{12} & \lambda A_{13} \\ \lambda A_{21} & \lambda A_{22} & \lambda A_{23} \\ \lambda A_{31} & \lambda A_{32} & \lambda A_{33} \end{pmatrix} \quad (1.7.32)$$

und

$$\hat{A} + \hat{B} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix} = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} & A_{13} + B_{13} \\ A_{21} + B_{21} & A_{22} + B_{22} & A_{23} + B_{23} \\ A_{31} + B_{31} & A_{32} + B_{32} & A_{33} + B_{33} \end{pmatrix}. \quad (1.7.33)$$

### 1.7.3 Euler-Winkel



Eine andere Parametrisierung von Drehungen, die allerdings erst für die Behandlung des Kreisels am Ende der Vorlesung relevant sein wird. Dort benötigt man sie für die Umrechnung zwischen Vektorkomponenten bzgl. einer rechtshändigen Orthonormalbasis ( $\vec{e}_j$ ), die im Bezugssystem eines raumfesten Beobachters ruht und einer rechtshändigen Orthonormalbasis ( $\vec{e}'_k$ ), die fest in dem betrachteten starren Körper verankert ist, wobei nur Drehungen um einen festgehaltenen Punkt, den wir als den Koordinatenursprung wählen, betrachtet werden. Zu irgendeinem festen Zeitpunkt ist offenbar die rechtshändige Orthonormalbasis ( $\vec{e}'_k$ ) gegenüber der rechtshändigen Orthonormalbasis ( $\vec{e}_j$ ) verdreht, und wir wollen die entsprechende Drehmatrix  $\hat{D}$  finden, so daß

$$\vec{e}'_k = \sum_{l=1}^3 D^l_k \vec{e}_l \quad (1.7.34)$$

ist. Dazu führen wir zunächst die Drehmatrizen für die Drehung um die drei Koordinatenachsen eines rechtshändigen kartesischen Koordinatensystems ein:

$$\begin{aligned}\hat{D}_1(\varphi) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}, \\ \hat{D}_2(\varphi) &= \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix}, \\ \hat{D}_3(\varphi) &= \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.\end{aligned}\tag{1.7.35}$$

Betrachten wir nun die geometrischen Verhältnisse der beiden rechtshändigen Orthonormalsysteme  $(\vec{e}_j)$  und  $(\vec{e}'_k)$  anhand der nebenstehenden Skizze: Die Schnittlinie der 12- mit der 1'2'-Ebene definiert die sogenannte Knotenlinie, die wir gestrichelt grün eingezeichnet haben. Der in ihre Richtung weisende Einheitsvektor  $\vec{k}$  ist nach der Rechte-Hand-Regel so orientiert, daß bei der Drehung des Orthonormalsystems  $(\vec{e}_j)$  um diese Achse die 3-Achse in die 3'-Achse gedreht wird, und zwar so, daß der Drehwinkel ins Intervall  $\vartheta \in [0, \pi]$  fällt. Das Orthonormalsystem  $(\vec{e}_j)$  wird nun nacheinander durch drei Drehungen in das Dreibein  $\vec{e}'_j$  verdreht, und zwar wie folgt: Zunächst erfolgt eine Drehung um die 3-Achse um den Drehwinkel  $\psi \in [0, 2\pi)$ , so daß die neue Einsachse  $\vec{e}''_1$  mit dem Knotenlinienvektor  $\vec{k}$  zusammenfällt:

$$\vec{e}''_j = \sum_{k=1}^3 \vec{e}_k [\hat{D}_3(\psi)]^k_j.\tag{1.7.36}$$

Sodann erfolgt eine Drehung um die Achse  $\vec{k} = \vec{e}''_1$  um den Winkel  $\vartheta$ , die dafür sorgt, daß die neue Dreiachse  $\vec{e}'''_3$  nunmehr mit der 3'-Achse zusammenfällt:

$$\vec{e}'''_l = \sum_{j=1}^3 \vec{e}''_j [\hat{D}_1(\vartheta)]^j_l = \sum_{j,k=1}^3 \vec{e}_k [\hat{D}_3(\psi)]^k_j [\hat{D}_1(\vartheta)]^j_l.\tag{1.7.37}$$

Schließlich wird noch um die Achse  $\vec{e}'''_3 = \vec{e}'_3$  um den Winkel  $\varphi$  gedreht, so daß schließlich auch die neuen Eins- und Zweiachsen jeweils auf  $\vec{e}'_1$  und  $\vec{e}'_2$  zu liegen kommen:

$$\begin{aligned}\vec{e}'_m &= \sum_{l=1}^3 \vec{e}'''_l [\hat{D}_3(\varphi)]^l_m \\ &= \sum_{j,k,l=1}^3 \vec{e}_k [\hat{D}_3(\psi)]^k_j [\hat{D}_1(\vartheta)]^j_l [\hat{D}_3(\varphi)]^l_m \\ &= \sum_{k=1}^3 \vec{e}_k [\hat{D}_3(\psi) \hat{D}_1(\vartheta) \hat{D}_3(\varphi)]^k_m = \sum_{k=1}^3 \vec{e}_k D^k_m.\end{aligned}\tag{1.7.38}$$

Für einen beliebigen Vektor  $\vec{x}$  folgt dann

$$\vec{x} = \sum_{m=1}^3 \vec{e}'_m x'^m = \sum_{k,m=1}^3 \vec{e}_k D^k_m x'^m \Rightarrow \bar{x} = \hat{D}(\psi, \vartheta, \varphi) \bar{x}',\tag{1.7.39}$$

wobei gemäß (1.7.38)

$$\hat{D}(\psi, \vartheta, \varphi) = \hat{D}_3(\psi) \hat{D}_1(\vartheta) \hat{D}_3(\varphi)\tag{1.7.40}$$

ist.

Es ist sehr wichtig zu bemerken, daß die Hintereinanderausführung von *Drehungen um verschiedene Achsen nicht kommutativ* ist, d.h. es ist

$$\hat{D}_3(\alpha)\hat{D}_1(\beta) \neq \hat{D}_1(\beta)\hat{D}_3(\alpha). \quad (1.7.41)$$

Wir schreiben schließlich die in **Euler-Winkeln**  $(\psi, \vartheta, \varphi)$  parametrisierte Drehmatrix (1.7.40) noch explizit auf. Man erhält diese Form einfach durch direkte Matrizenmultiplikation:

$$\hat{D} = \begin{pmatrix} \cos \psi \cos \varphi - \sin \psi \cos \vartheta \sin \varphi & -\cos \psi \sin \varphi - \sin \psi \cos \vartheta \cos \varphi & \sin \psi \sin \vartheta \\ \sin \psi \cos \varphi + \cos \psi \cos \vartheta \sin \varphi & -\sin \psi \sin \varphi + \cos \psi \cos \vartheta \cos \varphi & -\cos \psi \sin \vartheta \\ \sin \vartheta \sin \varphi & \sin \vartheta \cos \varphi & \cos \vartheta \end{pmatrix}. \quad (1.7.42)$$

## 1.8 Tensoren und ko- und kontravariante Komponenten

In diesem Abschnitt wollen wir noch einmal die Bedeutung der oben bzw. unten angeordneten Indizes betonen, da dies oft zu Schwierigkeiten führt. Insbesondere beim Vektorprodukt waren wir nicht sehr penibel mit der strikten Einhaltung der Regeln für die Schreibweise mit oberen und unteren Indizes. Dies liegt daran, daß wir meist mit kartesischen Basen gearbeitet haben, wo die Indexstellung unwichtig ist, was wir im folgenden auch genauer verstehen werden.

Dazu rekapitulieren wir nochmals das Transformationsverhalten von Vektoren bzgl. eines Basiswechsels und gehen dazu von zwei beliebigen Basen  $\vec{b}_j$  und  $\vec{b}'_k$  aus, die sich vermöge der Transformationsmatrix  $\hat{T}$  ineinander transformieren:

$$\vec{b}'_k = \sum_{j=1}^3 T^j_k \vec{b}_j, \quad \vec{b}_j = \sum_{k=1}^3 U^k_j \vec{b}'_k \quad \text{mit} \quad \hat{U} = \hat{T}^{-1}. \quad (1.8.1)$$

Man nennt im hier zu entwickelnden strikten **Ricci-Kalkül** alle Objekte mit einem unteren Index **kovariant**, d.h. sie transformieren sich definitionsgemäß wie die Basisvektoren gemäß (1.8.1). Im folgenden lassen wir die Summenzeichen über doppelt auftretende Indizes einfach weg (**Einstein-Summenkonvention**), d.h. treten doppelte Indizes auf, wird stets über dieses Indexpaar summiert. Im strikten Indexkalkül muß in einem solchen Paar immer zwingend ein Index oben und einer unten stehen.

Für einen Vektor  $\vec{v}$  führten wir die eindeutig bestimmten Komponenten  $v^j$  bzw.  $v'^k$  bzgl. der beiden Basis ein, d.h.

$$\vec{v} = v^j \vec{b}_j = v'^k \vec{b}'_k \quad (1.8.2)$$

ein. Setzen wir nun in den letzten mittleren Ausdruck die erste Gleichung in (1.8.1) ein, erhalten wir

$$\vec{v} = v^j T^k_j \vec{b}'_k \Rightarrow v'^k = T^k_j v^j \Rightarrow v^j = U^j_k v'^k. \quad (1.8.3)$$

Man bezeichnet Objekte mit einem oberen Index auch als **kontravariante Komponenten**, denn sie transformieren sich anders als Komponenten mit unteren Indizes (vgl. das Transformationsverhalten für die Basisvektoren in (1.8.1)).

Für das Skalarprodukt haben wir gesehen, daß wir nur die Komponenten

$$g_{jk} = \vec{b}_j \cdot \vec{b}_k \quad \text{bzw.} \quad g'_{lm} = \vec{b}'_l \cdot \vec{b}'_m \quad (1.8.4)$$

benötigen, denn es gilt für zwei beliebige Vektoren  $\vec{v}$  und  $\vec{w}$

$$\vec{v} \cdot \vec{w} = v^j w^k \vec{b}_j \cdot \vec{b}_k = g_{jk} v^j w^k = g'_{lm} v'^l w'^m. \quad (1.8.5)$$

Mit (1.8.3) folgt

$$g_{jk} v^j w^k = g'_{lm} T^l_j T^m_k v^j w^k. \quad (1.8.6)$$

Da dies für beliebige  $\vec{v}$  und  $\vec{w}$  gilt, muß folglich

$$g_{jk} = T^l_j T^m_k g'_{lm} \quad (1.8.7)$$

bzw.

$$g'_{lm} = U^j_l U^k_m g_{jk} \quad (1.8.8)$$

gelten. Die Komponenten des Skalarprodukts (bzw. einer allgemeinen Bilinearform) transformieren sich also bzgl. jedes der beiden Indizes wie die Basisvektoren.

Entsprechend führt man für Vektoren auch kovariante Komponenten ein, indem man definiert

$$v_j = g_{jk} v^k. \quad (1.8.9)$$

Dann ist zum einen

$$\vec{v} \cdot \vec{w} = g_{jl} v^j w^l = v_l w^l, \quad (1.8.10)$$

wobei wir  $g_{jl} = g_{lj}$  benutzt haben. Zum anderen transformieren sich aber die kovarianten Komponenten auch entsprechend der Transformationsvorschrift für solche, denn es gilt

$$v'_j = g'_{jk} v'^k \stackrel{(1.8.8)}{=} g_{lm} U^l_j U^m_k v'^k = U^l_j g_{lm} v^m \stackrel{(1.8.9)}{=} U^l_j v_l. \quad (1.8.11)$$

Dies entspricht aber genau der Transformationsvorschrift eines kovarianten Objekts, wie man es aufgrund der Indexstellung erwartet.

Entsprechend definieren wir die kontravarianten Komponenten  $g^{mn}$  des Skalarprodukts durch

$$g^{mn} g_{nk} = \delta^m_k, \quad (1.8.12)$$

d.h. betrachtet man  $(g_{jk})$  als Matrix  $\overleftrightarrow{g}$ , bilden die  $(g^{lm})$  gerade die inverse matrix, die wir mit  $\overleftarrow{g}$  bezeichnen.

Um zu zeigen, daß sich  $g^{lm}$  entsprechend der Indexstellung kontravariant transformiert, verwenden wir am bequemsten die Matrix-Vektorschreibweise: (1.8.8) schreibt sich damit in der Form (*warum?*)

$$\overleftarrow{g}' = \hat{U}^T \overleftrightarrow{g} \hat{U}. \quad (1.8.13)$$

Nun gilt mit den oben hergeleiteten Regeln für Matrixmultiplikation, -inversion und -transposition

$$\overleftrightarrow{g}' = (\overleftarrow{g}')^{-1} = \hat{U}^{-1} \overleftarrow{g}^{-1} \hat{U}^{T-1} = \hat{T} \overleftrightarrow{g} \hat{T}^T, \quad (1.8.14)$$

wobei wir  $\hat{U}^{-1} = \hat{T}$  benutzt haben. In der Indexschreibweise bedeutet (1.8.14) nun aber

$$g'^{mn} = T^m_j T^n_k g^{jk}, \quad (1.8.15)$$

d.h.  $g^{jk}$  transformiert sich bzgl. beider Komponenten kontravariant, wie es sein muß. Wir bezeichnen allgemein Symbole mit  $n$  unteren Indizes  $T_{k_1 \dots k_n}$  als kovariante Tensorcomponenten  $n$ -ter Stufe. Die Tensorcomponenten des Skalarprodukts (das in diesem Zusammenhang auch als die **Metrik** des Euklidischen Vektorraumes bezeichnet wird) bilden also kovariante Tensorcomponenten 2. Stufe.

Es fehlt nun nur noch eine manifest kovariante Beschreibung des Vektorprodukts. Motiviert durch unsere obige Betrachtung der Transformation der Volumenform auf Komponenten bzgl. beliebiger Koordinaten definieren wir dazu mit

$$g = \det \overleftrightarrow{g} \quad (1.8.16)$$

die kovarianten Komponenten des Levi-Civita-Tensors

$$\Delta_{abc} = \sqrt{g} \epsilon_{abc}, \quad (1.8.17)$$

wobei  $\epsilon_{abc}$  das gewöhnliche Levi-Civita-Symbol ist, das vollständig antisymmetrisch unter Vertauschung aller seiner Indizes ist und für das  $\epsilon_{123} = +1$  gilt. Wir zeigen, daß es sich kovariant transformiert. Es gilt

$$\Delta'_{abc} = \sqrt{g'} \epsilon_{abc} \quad (1.8.18)$$

Andererseits ist

$$\Delta_{abc} U^a_k U^b_l U^c_m = \sqrt{g} \epsilon_{abc} U^a_k U^b_l U^c_m = \sqrt{g} \det U \epsilon_{klm}. \quad (1.8.19)$$

Andererseits ist für  $\det \hat{U} > 0$

$$g' = \det \overleftrightarrow{g}' = \det(\hat{U}^T \overleftrightarrow{g} \hat{U}) = \det \overleftrightarrow{g} (\det \hat{U})^2 \Rightarrow \sqrt{g'} = \sqrt{g} \det \hat{U}. \quad (1.8.20)$$

Folglich ist gemäß (1.8.19)

$$\Delta_{abc} U^a_k U^b_l U^c_m = \sqrt{g'} \epsilon_{klm} = \Delta'_{klm}, \quad (1.8.21)$$

d.h.  $\Delta_{abc}$  transformiert sich entsprechend einem kovarianten Objekt mit drei unteren Indizes, wie behauptet. Das impliziert, daß für zwei Vektoren  $\vec{v}$  und  $\vec{w}$

$$x_j = \Delta_{jkl} v^k w^l \quad (1.8.22)$$

kovariante Komponenten sind. In der Tat ist

$$\begin{aligned} x'_a &= \Delta'_{abc} v'^b w'^c = \Delta_{klm} U^k_a U^l_b U^m_c T^b_o T^c_p v^o w^p \\ &= \Delta_{klm} U^k_a \delta^l_o \delta^m_p v^o w^p = U^k_a \Delta_{klm} v^l w^m = U^k_a x_k. \end{aligned} \quad (1.8.23)$$

Damit sind aber

$$x^j = g^{jk} x_k = g^{jk} \Delta_{klm} v^l w^m \quad (1.8.24)$$

kontravariante Komponenten eines Vektors, und wir haben entsprechend das Vektorprodukt mittels der Komponenten bzgl. beliebiger Basisvektoren ausgedrückt.

Wir können natürlich auch kontravariante Komponenten des Levi-Civita-Tensors definieren:

$$\Delta^{klm} = g^{ka} g^{lb} g^{mc} \Delta_{abc} = g^{ka} g^{lb} g^{mc} \sqrt{g} \epsilon_{abc} = \epsilon^{klm} \det \overleftrightarrow{g}^{\leftrightarrow{-1}} \sqrt{g} = \frac{1}{\sqrt{g}} \epsilon^{klm}, \quad (1.8.25)$$

wobei wir

$$\epsilon^{klm} = \epsilon_{klm} \quad (1.8.26)$$

definieren. Es ist zu betonen, daß mit dieser Definition  $\epsilon_{klm}$  und  $\epsilon^{klm}$  keine ko- bzw. kontravarianten Tensorkomponenten sind.

Wir bemerken noch, daß durch kovariante Tensorkomponenten  $n$ -ter Stufe eine von der Basis unabhängige multilineare Abbildung definiert wird, denn definieren wir für  $n$  beliebige Vektoren

$$T(\vec{v}, \vec{w}, \dots, \vec{z}) = T_{k_1 k_2 \dots k_n} v^{k_1} w^{k_2} \dots z^{k_n}, \quad (1.8.27)$$

so ist dieser Ausdruck offenbar von der Wahl der Basis unabhängig, weil sich die  $T_{k_1 k_2 \dots k_n}$  ko- und die  $v^{k_1}$  usw. kontravariant transformieren, denn es gilt demnach

$$\begin{aligned} T'_{k_1 k_2 \dots k_n} v'^{k_1} w'^{k_2} \dots z'^{k_n} &= U^{j_1}_{k_1} U^{j_2}_{k_2} \dots U^{j_n}_{k_n} T_{j_1 j_2 \dots j_n} T^{k_1}_{l_1} T^{k_2}_{l_2} \dots T^{k_n}_{l_n} v^{l_1} w^{l_2} \dots z^{l_n} \\ &= \delta^{j_1}_{l_1} \delta^{j_2}_{l_2} \dots \delta^{j_n}_{l_n} v^{l_1} w^{l_2} \dots z^{l_n} T_{j_1 j_2 \dots j_n} \\ &= T_{j_1 j_2 \dots j_n} v^{j_1} w^{j_2} \dots z^{j_n}. \end{aligned} \quad (1.8.28)$$

Die Definition (1.8.27) ist also tatsächlich unabhängig von der gewählten Basis, vorausgesetzt die  $T_{k_1 k_2 \dots k_n}$  transformieren sich für jeden Index kovariant.



## Kapitel 2

---

# Vektoranalysis

In diesem Kapitel kommen wir nun auf die Vektoranalysis zu sprechen. Hierbei verwenden wir die Darstellung der Vektoren in einer Ebene bzw. im Raum als Vektoren in  $E^2$  bzw.  $E^3$ , indem wir zunächst eine stets rechtshändig orientierte kartesische Basis  $(\vec{e}_1, \vec{e}_2)$  bzw.  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  verwenden und

$$\begin{aligned}\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 &\mapsto \bar{x} = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \quad \text{bzw.} \\ \vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3 &\mapsto \bar{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.\end{aligned}\tag{2.0.1}$$

### 2.1 Kurven in Ebene und Raum

Als erstes betrachten wir Kurven in der Ebene und im Raum, die wir durch beliebige Funktionen parametrisieren können. Diese Funktionen können wir dann ggf. differenzieren und integrieren, so daß die Methoden der Analysis für die Analyse von Kurven bzw. von Bahnkurven von Massenpunkten als Funktionen der Zeit (**Trajektorien**) zur Verfügung stehen. Wir gehen davon aus, daß die Begriffe der Ableitung und des Integrals von Funktionen von der Schule her bekannt sind.

#### 2.1.1 Ebene Kurven

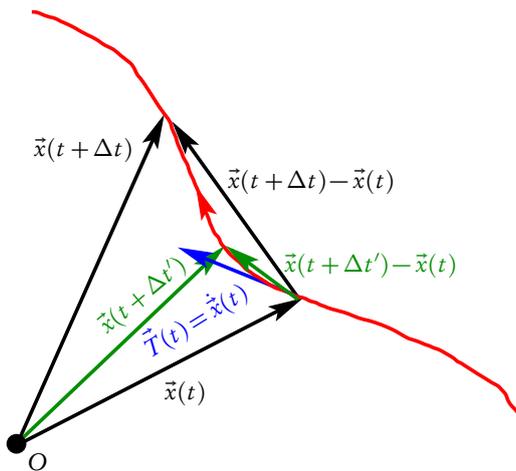
Wir betrachten als erstes Kurven in der Euklidischen Ebene  $\mathbb{R}^2$ . Nach Festlegung des Koordinatenursprungs können sie als Funktion der Ortsvektoren der Punkte auf der Kurve  $\bar{x} : \mathbb{R} \supseteq (t_0, t_1) \rightarrow \mathbb{R}^2$  dargestellt werden. Im folgenden nehmen wir an, diese Funktion sei im ganzen Intervall  $(t_0, t_1)$  stetig nach dem Parameter differenzierbar, d.h. daß beide Vektorkomponenten in allen Punkten  $t \in (t_0, t_1)$  Ableitungen besitzen und die Ableitungsfunktion dort stetig ist. Die Ableitung ist dabei durch den Grenzwert

$$\frac{d}{dt} \bar{x}(t) = \dot{\bar{x}} = \lim_{\Delta t \rightarrow 0} \frac{\bar{x}(t + \Delta t) - \bar{x}(t)}{\Delta t}\tag{2.1.1}$$

definiert. Es ist klar, daß wir diesen Vektor wieder in den Raum der geometrischen Vektoren abbilden können:

$$\frac{d}{dt} \vec{x}(t) = \dot{\vec{x}}(t) = \sum_{j=1}^3 \dot{x}_j(t) \vec{e}_j = \frac{d}{dt} \sum_{j=1}^3 x^j(t) \vec{e}_j.\tag{2.1.2}$$

Letzteres gilt, da die Basisvektoren  $(\vec{e}_j)$  fest vorgegebene konstante Vektoren sind und folglich nicht von  $t$  abhängen.



Jetzt wollen wir die lokalen Eigenschaften einer solchen Kurve durch geometrische vom Koordinatensystem unabhängige Größen charakterisieren. An einem Punkt  $t$  sei nun

$$\frac{d\vec{x}}{dt} \neq 0. \quad (2.1.3)$$

Wir nennen den dazugehörigen Punkt  $\vec{x}(t)$  auf der Kurve einen **regulären Punkt**. Offensichtlich besitzt der Vektor  $d\vec{x}/dt$  in jedem regulären Punkt der Kurve die geometrische Bedeutung eines **Tangentenvektors**. Dies zeigt die nebenstehende Skizze: Für endliche  $\Delta t$  ist der Zähler in (2.1.1) offenbar der Verbindungsvektor<sup>1</sup> zwischen den Punkten  $\vec{x}(t + \Delta t)$  und  $\vec{x}(t)$ . Lassen wir  $\Delta t$  kleiner werden, nähert sich die entsprechende Sekante immer mehr der Tangente an die Kurve im Punkt  $\vec{x}(t)$ , voraus-

gesetzt die Kurve ist differenzierbar. Wir definieren nun durch

$$\vec{T} = \left| \frac{d\vec{x}}{dt} \right|^{-1} \frac{d\vec{x}}{dt} \quad (2.1.4)$$

die **Tangenteneinheitsvektoren** an die Kurve, die an jedem regulären Punkt wohldefiniert sind. Leiten wir die Identität  $\vec{T}^2 = 1$  nach dem Parameter ab, erhalten wir

$$\vec{T} \cdot \frac{d\vec{T}}{dt} = 0. \quad (2.1.5)$$

Dies bedeutet, daß entweder  $d\vec{T}/dt$  ein auf dem Tangentenvektor senkrechter Vektor oder der Nullvektor ist. Wir betrachten nun den ersteren Fall weiter. Wir können dann in diesem Punkt durch

$$\vec{N} = \left| \frac{d\vec{T}}{dt} \right|^{-1} \frac{d\vec{T}}{dt} \quad (2.1.6)$$

einen zweiten zu  $\vec{T}$  senkrechten Einheitsvektor definieren. Das Paar  $\vec{T}$  und  $\vec{N}$  bezeichnen wir als das **die Kurve begleitende Zweibein**.

Bislang haben wir mit einem ganz allgemeinen Parameter  $t$  zur Beschreibung der Kurve gearbeitet. Dieser Parameter besitzt im allgemeinen keine besondere geometrische Bedeutung. Eine natürlichere Parametrisierung ist durch die **Bogenlänge** der Kurve, gemessen vom Anfangspunkt  $\vec{x}(t_0)$  gegeben. Der Zuwachs der Bogenlänge, der sich ergibt, wenn wir um ein infinitesimales Stückchen weiterrücken, ist offensichtlich durch

$$ds = \left| \frac{d\vec{x}}{dt} \right| dt \quad (2.1.7)$$

gegeben. Denken wir uns also die Kurve durch die Bogenlänge  $s$  parametrisiert, vereinfacht sich (2.1.4) zu

$$\vec{T} = \frac{d\vec{x}}{ds}. \quad (2.1.8)$$

Eine Gerade ist nun offenbar durch

$$g: \quad \vec{x}(s) = \vec{x}_0 + \vec{T}s \quad \text{mit} \quad \vec{T} = \text{const.} \quad \text{und} \quad |\vec{T}| = 1, \quad (2.1.9)$$

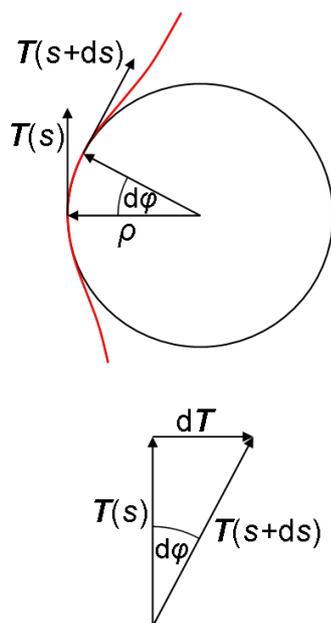
<sup>1</sup>repräsentiert durch die Komponenten bzgl. der kartesischen Basis

gegeben. Der Einheitstangentenvektor entlang einer Geraden ändert sich natürlich nicht. Folglich charakterisiert die Größe

$$\kappa = \left| \frac{d\vec{T}}{ds} \right| \quad (2.1.10)$$

die Abweichung der Kurve von einer Geraden, also die **Krümmung** der Kurve durch rein geometrische von der Wahl der Parametrisierung der Kurve unabhängige Größen.

Die Krümmung  $\kappa$  läßt sich nun noch geometrisch interpretieren. Wie wir oben gesehen haben, steht die infinitesimale Änderung des Tangentenvektors  $d\vec{T} = ds d\vec{T}/ds$  senkrecht auf  $\vec{T}$ . Für den Winkel  $d\varphi$  zwischen  $\vec{T} + d\vec{T}$  und  $\vec{T}$  gilt also (s. Abb. 2.1)



$$\sin(d\varphi) \simeq d\varphi = ds \left| \frac{d\vec{T}}{ds} \right| = ds \kappa. \quad (2.1.11)$$

Die Bogenlänge ist also  $ds = d\varphi/\kappa$ , d.h.  $\rho = 1/\kappa$  ist der Radius des sich an die Kurve im betrachteten Punkt anschmiegenden Kreises, den man in diesem Zusammenhang als **Krümmungskreis** bezeichnet. Die Größe  $\rho$  heißt **Krümmungsradius**. Der Ortsvektor des Mittelpunktes des Krümmungskreises an die Kurve in dem betrachteten Punkt liegt demnach bei

$$\vec{K} = \vec{x} + \rho \vec{N}. \quad (2.1.12)$$

Die Menge der Krümmungskreismittelpunkte bildet ihrerseits wieder eine Kurve, die sogenannte **Evolute** der Ausgangskurve.

Wir geben noch die expliziten Ausdrücke für diese Größen für eine beliebige Parametrisierung der Kurve, ausgedrückt durch die Komponenten bzgl. eines kartesischen Koordinatensystems an. Zunächst ist

$$\frac{ds}{dt} = \left| \frac{d\vec{x}}{dt} \right| = |\dot{\vec{x}}| = \sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}. \quad (2.1.13)$$

Weiter ist wegen (2.1.4) in dieser Schreibweise

$$\vec{T} = \frac{d\vec{x}}{ds} = \frac{\dot{\vec{x}}}{|\dot{\vec{x}}|} = \frac{1}{\sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}} \begin{pmatrix} \dot{x}^1 \\ \dot{x}^2 \end{pmatrix}. \quad (2.1.14)$$

Wir benötigen noch

$$\frac{d}{dt} |\dot{\vec{x}}| = \frac{d}{dt} \sqrt{\dot{\vec{x}}^2} = \frac{\ddot{\vec{x}} \cdot \dot{\vec{x}}}{|\dot{\vec{x}}|} = \frac{\dot{x}^1 \ddot{x}^1 + \dot{x}^2 \ddot{x}^2}{\sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}}, \quad (2.1.15)$$

und

$$\frac{d\vec{T}}{ds} = \frac{\dot{\vec{T}}}{|\dot{\vec{x}}|} = \frac{\ddot{\vec{x}} |\dot{\vec{x}}|^2 - \dot{\vec{x}} (\ddot{\vec{x}} \cdot \dot{\vec{x}})}{|\dot{\vec{x}}|^4} = \frac{\dot{x}^1 \ddot{x}^2 - \dot{x}^2 \ddot{x}^1}{[(\dot{x}^1)^2 + (\dot{x}^2)^2]^2} \begin{pmatrix} -\dot{x}^2 \\ \dot{x}^1 \end{pmatrix}, \quad (2.1.16)$$

woraus nach einiger Rechnung aus (2.1.10)

$$\kappa = \frac{|\dot{x}^1 \ddot{x}^2 - \dot{x}^2 \ddot{x}^1|}{\sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}^3} \quad (2.1.17)$$

folgt. Normierung von (2.1.16) liefert gemäß (2.1.6) schließlich für den Normalenvektor

$$\vec{N} = \frac{\text{sign}(\dot{x}^1 \ddot{x}^2 - \dot{x}^2 \ddot{x}^1)}{\sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}} \begin{pmatrix} -\dot{x}^2 \\ \dot{x}^1 \end{pmatrix}. \quad (2.1.18)$$

Für die Evolute erhalten wir wegen (2.1.12) und (2.1.18)

$$\vec{K} = \vec{x} + \frac{(\dot{x}^1)^2 + (\dot{x}^2)^2}{\dot{x}^1 \ddot{x}^2 - \dot{x}^2 \ddot{x}^1} \begin{pmatrix} -\dot{x}^2 \\ \dot{x}^1 \end{pmatrix}. \quad (2.1.19)$$

**Beispiel:** Für einen Kreis mit Radius  $r$  um den Ursprung des Koordinatensystems können wir die Parametrisierung

$$\vec{x} = r \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad t \in [0, 2\pi) \quad (2.1.20)$$

wählen. Hier ist eine Parametrisierung mit der Bogenlänge einfach, denn es ist offenbar

$$s(t) = \int_0^t dt' |\dot{\vec{x}}(t')| = r t. \quad (2.1.21)$$

Es ist also

$$\vec{x}(s) = r \begin{pmatrix} \cos(s/r) \\ \sin(s/r) \end{pmatrix}, \quad s \in [0, 2\pi r). \quad (2.1.22)$$

Der Tangenteneinheitsvektor an jedem Punkt ist also durch

$$\vec{T} = \frac{d\vec{x}}{ds} = \begin{pmatrix} -\sin(s/r) \\ \cos(s/r) \end{pmatrix} \quad (2.1.23)$$

gegeben und

$$\frac{d\vec{T}}{ds} = -\frac{1}{r} \begin{pmatrix} \cos(s/r) \\ \sin(s/r) \end{pmatrix}, \quad \vec{N} = -\begin{pmatrix} \cos(s/r) \\ \sin(s/r) \end{pmatrix} \quad (2.1.24)$$

Die Krümmung ist wegen (2.1.10)

$$\kappa = \frac{1}{\rho} = \left| \frac{d\vec{T}}{ds} \right| = \frac{1}{r}. \quad (2.1.25)$$

Das ist verständlich, denn der Schmiegekreis an den Kreis ist dieser Kreis selbst. Entsprechend ist die Evolute eines Kreises einfach ein einzelner Punkt, eben der Mittelpunkt des Kreises. In der Tat folgt aus (2.1.12) und (2.1.25) sofort  $\vec{K} = 0 = \text{const.}$

## 2.1.2 Raumkurven und Fresnetsche Formeln

Die Verallgemeinerung der Betrachtungen zu Kurven im Euklidischen  $\mathbb{R}^3$  ist nicht besonders schwierig. Auch hier definieren wir eine Kurve durch eine Parameterdarstellung  $\vec{x} : \mathbb{R} \supseteq (t_0, t_1) \rightarrow \mathbb{R}^3$  und ebenso wie für die ebenen Kurven ist ein intrinsischer Parameter durch die Bogenlänge der Kurve, gezählt vom Anfangspunkt  $\vec{x}(t_0)$  eine geometrische der Kurve immanenter (d.h. von der Wahl der Parametrisierung unabhängige) Größe. Betrachten wir also zunächst die Kurve wieder in dieser Parametrisierung. Dann ist der **Einheitstangentenvektor** durch

$$\vec{T} = \frac{d\vec{x}}{ds} = \frac{\dot{\vec{x}}}{|\dot{\vec{x}}|} \quad (2.1.26)$$

definiert, und der **Hauptnormaleneinheitsvektor** ist durch

$$\vec{N} = \frac{1}{\kappa} \frac{d\vec{T}}{ds}, \quad \kappa = \frac{1}{\rho} = \left| \frac{d\vec{T}}{ds} \right| \quad (2.1.27)$$

gegeben. Dabei ist  $\rho$  wieder der **Krümmungsradius** und  $\kappa = 1/\rho$  die **Krümmung der Kurve** an dem betrachteten Punkt. Wir ergänzen nun  $\vec{T}$  und  $\vec{N}$  durch die Definition des **Binormaleneinheitsvektors** vermöge

$$\vec{B} = \vec{T} \times \vec{N}. \quad (2.1.28)$$

Damit ist durch  $(\vec{T}, \vec{N}, \vec{B})$  eine allein durch die Geometrie der Kurve bestimmte (also von der Wahl der Parametrisierung unabhängige) rechtshändige Orthonormalbasis, das die Kurve **begleitende Dreibein**, definiert. Gegenüber den ebenen Kurven ist im Raum die Frage sinnvoll, ob die Kurve im Raum in einer Ebene bleibt oder nicht. Der Tangentenvektor  $\vec{T}$  und die Hauptnormale  $\vec{N}$  spannen die sogenannte **Schmiegeebene** auf. Ändert sich diese Ebene entlang der Kurve nicht, handelt es sich definitionsgemäß um eine **ebene Kurve im Raum**. Da der Binormalenvektor stets senkrecht auf der Schmiegeebene steht, ist dies genau dann der Fall, wenn  $d\vec{B}/ds = 0$  ist.

Die geometrische Bedeutung der Schmiegeebene wird klar, wenn man überlegt, wie sich die Kurve in der Umgebung eines Punktes  $\vec{x}_0 = \vec{x}(s_0)$  verhält. Dazu entwickeln wir die Parametrisierung der Kurve bis zu zweiter Ordnung um  $s_0$ :

$$\vec{x}(s_0 + \Delta s) = \vec{x}(s_0) + \Delta s \frac{d}{ds} \vec{x}(s_0) + \frac{\Delta s^2}{2} \frac{d^2}{ds^2} \vec{x}(s_0) + \mathcal{O}(\Delta s^3). \quad (2.1.29)$$

Nun ist gemäß (2.1.26) und (2.1.27)

$$\frac{d}{ds} \vec{x}(s_0) = \vec{T}(s_0), \quad \frac{d^2}{ds^2} \vec{x}(s_0) = \frac{d}{ds} \vec{T}(s_0) = \frac{1}{\rho} \vec{N}(t_0). \quad (2.1.30)$$

Dies in (2.1.29) eingesetzt liefert

$$\vec{x}(s_0 + \Delta s) = \vec{x}(s_0) + \Delta s \vec{T}(s_0) + \frac{\Delta s^2}{2} \frac{1}{\rho} \vec{N}(t_0) + \mathcal{O}(\Delta s^3). \quad (2.1.31)$$

In dieser Ordnung der Näherung bleibt also die Kurve in der durch  $\vec{T}(s_0)$  und  $\vec{N}(s_0)$  aufgespannten Ebene durch  $\vec{x}(s_0)$ , was die Bezeichnung Schmiegeebene rechtfertigt. Der Binormalenvektor ist dann der Normaleneinheitsvektor auf diese Ebene, wobei die Orientierung definitionsgemäß wie durch (2.1.28) angegeben gewählt wurde. Die Kurve bleibt demnach in der Tat in einer Ebene, wenn sich  $\vec{B}$  entlang der Kurve nicht ändert.

Wir suchen nun ein Maß für die Abweichung der Kurve von einer ebenen Kurve, also dafür, um wieviel sich die Kurve beim Fortschreiten um eine Bogenlängenänderung  $ds$  aus der Schmiegeebene herauswindet. Dafür bietet sich offensichtlich der Vektor  $d\vec{B}/ds$  an. Da das begleitende Dreibein eine Orthonormalbasis bildet, muß sich allerdings dieser Vektor als Linearkombination dieser Basisvektoren ausdrücken lassen. Wir wollen nun zeigen, daß er parallel zum Hauptnormalenvektor  $\vec{N}$  ist. Dazu differenzieren wir die Gleichung  $\vec{T} \cdot \vec{B} = 0 = \text{const}$  nach der Bogenlänge:

$$\frac{d\vec{T}}{ds} \cdot \vec{B} + \vec{T} \cdot \frac{d\vec{B}}{ds} \stackrel{(2.1.27)}{=} \kappa \vec{N} \cdot \vec{B} + \vec{T} \cdot \frac{d\vec{B}}{ds} \stackrel{(2.1.28)}{=} \vec{T} \cdot \frac{d\vec{B}}{ds} = 0. \quad (2.1.32)$$

Dies bedeutet, daß  $\vec{d}\vec{B}/ds$  senkrecht auf  $\vec{T}$  steht. Ebenso folgt aus  $\vec{B}^2 = 1 = \text{const}$ , daß dieser Vektor auch senkrecht auf  $\vec{B}$  steht. Damit muß aber

$$\frac{d\vec{B}}{ds} = -\chi \vec{N}, \quad \chi = -\vec{N} \cdot \frac{d\vec{B}}{ds} \quad (2.1.33)$$

sein. Die Größe  $\chi$  heißt **Windung oder Torsion**. Sie ist eine vorzeichenbehaftete Größe. Die geometrische Bedeutung wird klar, wenn wir (2.1.28) beachten. Aus der Formel für das doppelte Kreuzprodukt (1.4.32) folgt

$$-\vec{N} = \vec{T} \times \vec{B}, \quad (2.1.34)$$

so daß wir (2.1.33) auch in der Form

$$\frac{d\vec{B}}{ds} = \chi \vec{T} \times \vec{B} \quad (2.1.35)$$

schreiben können. Ist also  $\chi$  positiv, so ist die Änderung von  $\vec{B}$  in diesem Punkt bei infinitesimalem Fortschreiten entlang der Kurve um die Länge  $ds$  durch eine infinitesimale Drehung um den Winkel  $\chi ds$  um die durch  $\vec{T}$  definierte Drehachse im Sinne der Rechte-Hand-Regel gegeben. Die Größe  $\chi$  gibt also die lokale Ganghöhe der **Schraube** an, und zwar im Sinne einer **Rechtsschraube** falls  $\chi > 0$  und im Sinne einer **Linksschraube** falls  $\chi < 0$  ist.

Bevor wir ein charakteristisches Beispiel für diese Begriffsbildungen durchrechnen, wollen wir noch die **Frenet-Serret-Formeln** herleiten. Diese geben die Komponenten der Ableitungen des begleitenden Dreibeins nach der Bogenlänge der Kurve bzgl. der durch das begleitende Dreibein gegebenen Basis an. Aus (2.1.27) und (2.1.33) folgen bereits zwei der drei gesuchten Formeln

$$\frac{d\vec{T}}{ds} = \kappa \vec{N} \Rightarrow \vec{T} \cdot \frac{d\vec{T}}{ds} = 0, \quad \vec{N} \cdot \frac{d\vec{T}}{ds} = \kappa, \quad \vec{B} \cdot \frac{d\vec{T}}{ds} = 0, \quad (2.1.36)$$

$$\frac{d\vec{B}}{ds} = -\chi \vec{N} \Rightarrow \vec{T} \cdot \frac{d\vec{B}}{ds} = 0, \quad \vec{N} \cdot \frac{d\vec{B}}{ds} = -\chi, \quad \vec{B} \cdot \frac{d\vec{B}}{ds} = 0. \quad (2.1.37)$$

Daraus ergibt sich mit (2.1.34) aber auch sofort für die verbleibende Ableitung

$$\frac{d\vec{N}}{ds} = \frac{d}{ds}(\vec{B} \times \vec{T}) = \frac{d\vec{B}}{ds} \times \vec{T} + \vec{B} \times \frac{d\vec{T}}{ds} = -\chi \vec{N} \times \vec{T} + \kappa \vec{B} \times \vec{N} = \chi \vec{B} - \kappa \vec{T}, \quad (2.1.38)$$

wobei wir benutzt haben, daß  $(\vec{T}, \vec{N}, \vec{B})$  eine rechtshändige Orthonormalbasis bilden. Aus (2.1.38) folgt also

$$\vec{T} \cdot \frac{d\vec{N}}{ds} = -\kappa, \quad \vec{N} \cdot \frac{d\vec{N}}{ds} = 0, \quad \vec{B} \cdot \frac{d\vec{N}}{ds} = \chi. \quad (2.1.39)$$

Oft ist eine Parametrisierung der Kurve durch die Bogenlänge unbequem, so daß wir das begleitende Dreibein und die Krümmung und Torsion noch für eine beliebige Parametrisierung umschreiben wollen. Dazu benötigen wir nur

$$\frac{ds}{dt} = \left| \dot{\vec{x}} \right|, \quad (2.1.40)$$

wobei der Punkt über einem Symbol wieder die Ableitung nach dem beliebigen Parameter  $t$  bedeutet. Daraus folgt

$$\vec{T} = \frac{d\vec{x}}{ds} = \frac{\dot{\vec{x}}}{\left| \dot{\vec{x}} \right|}. \quad (2.1.41)$$

Weiter ist wegen  $|\dot{\vec{x}}| = \sqrt{\dot{x}^2}$

$$\frac{d}{dt} |\dot{\vec{x}}| = \frac{\ddot{\vec{x}} \cdot \dot{\vec{x}}}{|\dot{\vec{x}}|}. \quad (2.1.42)$$

Daraus folgt

$$\dot{\vec{T}} = \frac{\ddot{\vec{x}} \dot{x}^2 - \dot{\vec{x}} (\ddot{\vec{x}} \cdot \dot{\vec{x}})}{|\dot{\vec{x}}|^3} \stackrel{(1.4.32)}{=} \frac{\dot{\vec{x}} \times (\ddot{\vec{x}} \times \dot{\vec{x}})}{|\dot{\vec{x}}|^3}. \quad (2.1.43)$$

Zur Normierung dieses Vektors verwenden wir (1.4.1):

$$\left| \dot{\vec{T}} \right| = \frac{|\dot{\vec{x}}| |\ddot{\vec{x}} \times \dot{\vec{x}}|}{|\dot{\vec{x}}|^3} = \frac{|\ddot{\vec{x}} \times \dot{\vec{x}}|}{|\dot{\vec{x}}|^2}. \quad (2.1.44)$$

Der Hauptnormalenvektor ist also durch

$$\vec{N} = \frac{\dot{\vec{T}}}{|\dot{\vec{T}}|} = \frac{\dot{\vec{x}} \times (\ddot{\vec{x}} \times \dot{\vec{x}})}{|\dot{\vec{x}}| |\ddot{\vec{x}} \times \dot{\vec{x}}|} \quad (2.1.45)$$

gegeben und der Binormalenvektor durch

$$\vec{B} = \vec{T} \times \vec{N} = \frac{\dot{\vec{x}} \times \ddot{\vec{x}}}{|\dot{\vec{x}} \times \ddot{\vec{x}}|}. \quad (2.1.46)$$

Aus (2.1.36) ergibt sich für die Krümmung

$$\kappa = \vec{N} \cdot \frac{d\vec{T}}{ds} = \frac{\vec{N} \cdot \dot{\vec{T}}}{|\dot{\vec{x}}|} \stackrel{(2.1.43, 2.1.45)}{=} \frac{|\dot{\vec{x}} \times (\ddot{\vec{x}} \times \dot{\vec{x}})|^2}{|\dot{\vec{x}}|^5 |\ddot{\vec{x}} \times \dot{\vec{x}}|} = \frac{|\ddot{\vec{x}} \times \dot{\vec{x}}|}{|\dot{\vec{x}}|^3}. \quad (2.1.47)$$

Zur Berechnung der Torsion gehen wir von (2.1.37) aus:

$$\chi = -\vec{N} \cdot \frac{d\vec{B}}{ds} = -\frac{\vec{N} \cdot \dot{\vec{B}}}{|\dot{\vec{x}}|}. \quad (2.1.48)$$

Die Ableitung von  $\vec{B}$  ist

$$\dot{\vec{B}} = \frac{\dot{\vec{x}} \times \ddot{\vec{x}}}{|\dot{\vec{x}} \times \ddot{\vec{x}}|} - \frac{\vec{B}}{|\dot{\vec{x}} \times \ddot{\vec{x}}|} \frac{d}{dt} |\dot{\vec{x}} \times \ddot{\vec{x}}|. \quad (2.1.49)$$

Multiplikation mit  $\vec{N}$  liefert schließlich

$$\chi = \frac{\ddot{\vec{x}} \cdot (\dot{\vec{x}} \times \ddot{\vec{x}})}{|\dot{\vec{x}} \times \ddot{\vec{x}}|}. \quad (2.1.50)$$

Als **Beispiel** für diese geometrischen Betrachtungen, das besonders anschaulich ist, betrachten wir die Schraubenlinie

$$\vec{x}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ at \end{pmatrix}, \quad t \in [0, \infty), \quad r > 0, \quad a \in \mathbb{R}. \quad (2.1.51)$$

Wir haben

$$|\dot{\vec{x}}| = \sqrt{r^2 + a^2} \Rightarrow s = t \sqrt{r^2 + a^2}. \quad (2.1.52)$$

Wir können also in diesem Fall sehr einfach zur Parametrisierung mit der Bogenlänge übergehen, was die weiteren Rechnungen etwas erleichtert:

$$\vec{x}(s) = \begin{pmatrix} r \cos\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ r \sin\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ \frac{s}{\sqrt{r^2+a^2}} \end{pmatrix}. \quad (2.1.53)$$

Daraus ergibt sich mit (2.1.26-2.1.28) für das begleitende Dreibein

$$\vec{T} = \frac{1}{\sqrt{r^2+a^2}} \begin{pmatrix} -r \sin\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ r \cos\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ a \end{pmatrix}, \quad (2.1.54)$$

$$\vec{N} = \begin{pmatrix} -\cos\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ -\sin\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ 0 \end{pmatrix}, \quad (2.1.55)$$

$$\vec{B} = \frac{1}{\sqrt{r^2+a^2}} \begin{pmatrix} a \sin\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ -a \cos\left(\frac{s}{\sqrt{r^2+a^2}}\right) \\ r \end{pmatrix}, \quad (2.1.56)$$

Krümmung und Torsion der Kurve berechnen wir am bequemsten mit Hilfe der Gleichungen (2.1.27) bzw. (2.1.33):

$$\kappa = \frac{r}{r^2+a^2}, \quad \chi = \frac{a}{r^2+a^2}. \quad (2.1.57)$$

Die Kurve ist also für  $a > 0$  eine Rechts-,  $a < 0$  eine Linksschraube und für  $a = 0$  eben. In der Tat ergibt sich für  $a = 0$  ein Kreis in der  $12$ -Ebene mit Radius  $r$  um den Ursprung.

### 2.1.3 Anwendung auf die Bewegung eines Punktteilchens

Eine grundlegende Fragestellung der **klassischen Mechanik** ist die Beschreibung der Bewegung von **Punktteilchen** unter Einfluß von **Kräften**. Dabei versteht man unter einem Punktteilchen einen makroskopischen Körper *endlicher Ausdehnung*, dessen Ausdehnung allerdings für die betrachtete Bewegung irrelevant ist. Das Paradebeispiel für den Erfolg einer solchen **effektiven Beschreibung** von Körpern als Punktteilchen ist die **Himmelsmechanik**, die die eigentliche Motivation für die Entwicklung der klassischen Mechanik durch **Newton** war. Die Ausdehnung der Sonne und der Planeten in unserem Sonnensystem kann in der Tat für die Berechnung der Bewegung dieser Himmelskörper unter Einwirkung ihrer gegenseitigen **Gravitation** in sehr guter Näherung vernachlässigt werden.

Hier betrachten wir die einfachste Fragestellung dieser Art: Ein als **Massenpunkt** idealisierter Körper bewege sich unter Einfluß irgendwelcher vorgegebener Kräfte. Z.B. kann man bei der Planetenbewegung in einer ersten Näherung die Sonne als feststehenden Körper betrachten, da ihre Masse  $M$  sehr groß gegenüber der Masse  $m$  des Planeten ist. Wir wählen ein kartesisches Koordinatensystem mit dem Ursprung im Schwerpunkt der Sonne. Dann ist die Kraft auf den Planeten durch das **Newtonsche Gravitationsgesetz**

$$\vec{F}(\vec{x}) = -\gamma \frac{mM}{r^2} \frac{\vec{x}}{r} \quad (2.1.58)$$

gegeben. Dabei ist  $\vec{x}$  der **Ortsvektor** des Planeten, also der Vektor vom Schwerpunkt der Sonne zum Ort des Planeten,  $r = |\vec{x}|$  sein Betrag und  $\gamma$  die **Newtonsche Gravitationskonstante**. Die Gleichung (2.1.58) besagt, daß die Kraft stets vom Planeten in Richtung zur Sonne weist und proportional zur Masse des Planeten und der Sonne ist und mit dem **Quadrat des Abstands** abfällt. Aufgabe der **Dynamik** ist es nun, die **Position** des Planeten als Funktion der **Zeit** aus diesem Kraftgesetz zu bestimmen. Dazu verwenden wir die **Newtonsche Bewegungsgleichung**

$$m\ddot{\vec{x}} = \vec{F}(\vec{x}). \quad (2.1.59)$$

Dabei bedeutet  $\ddot{\vec{x}}$  die zweite Ableitung des Ortsvektors des Planeten nach der Zeit, also seine **Beschleunigung**. Es ist klar, daß die Lösung dieses Problems nicht einfach ist, denn die unbekannte Funktion  $\vec{x}(t)$  kommt mitsamt der zweiten Ableitung in der Gleichung vor, und wir suchen möglichst alle Funktionen, die diese Gleichung erfüllen, um zu sehen, welche Bahnen es bei dieser Bewegung gibt. Diese Aufgabenstellung ist typisch für die gesamte Physik. Es handelt sich um **Differentialgleichungen**, und ein wesentliches Ziel dieser Vorlesung ist es, Lösungsstrategien für diese Gleichungen zu entwickeln.

Hier betrachten wir die einfachere Aufgabe, die allgemeine **Kinematik** solcher Bewegungen von Punktteilchen zu verstehen. Wir nehmen also an, wir hätten eine Lösung  $\vec{x}(t)$  gefunden und untersuchen die geometrische und physikalische Bedeutung einer solchen Lösung. Zunächst einmal ist klar, daß  $\vec{x}(t)$  die **Bahnkurve** des Massenpunktes beschreibt, wobei die Zeit als Parameter für diese Bahnkurve dient. Es sind nun alle in diesem Abschnitt entwickelten geometrischen Analysemethoden anwendbar. Betrachten wir als erstes die **Momentangeschwindigkeit** des Teilchens. Sie ist durch die **Zeitableitung**

$$\vec{v}(t) = \dot{\vec{x}}(t) = \frac{d}{dt} \vec{x}(t). \quad (2.1.60)$$

Die geometrische Bedeutung wird dabei klar, wenn wir uns die Kurve als Funktion der Bogenlänge, wie oben beschrieben, parametrisiert denken. Wir führen also zunächst die neue Funktion  $\vec{x}(s)$  ein. Dabei ist die Bogenlänge durch das Differential

$$ds = dt |\dot{\vec{x}}(t)| \quad (2.1.61)$$

gegeben, und wir erhalten die Bogenlänge, von einem Anfangszeitpunkt  $t_0$  aus gerechnet durch **Integration**

$$s(t) = \int_{t_0}^t dt' |\dot{\vec{x}}(t')|. \quad (2.1.62)$$

In kartesischen Koordinaten ausgeschrieben lautet das Integral

$$s(t) = \int_{t_0}^t dt' \sqrt{\dot{x}_1^2(t') + \dot{x}_2^2(t') + \dot{x}_3^2(t')}. \quad (2.1.63)$$

Es ist klar, daß (außer für den trivialen Fall, daß das Teilchen über ein endliches Zeitintervall ruht) die Bogenlänge eine monoton wachsende Funktion der Zeit ist, denn der Integrand ist stets positiv.

Nun folgt aus dem Hauptsatz der Differential- und Integralrechnung, daß

$$\dot{s}(t) = \frac{d}{dt} s(t) = |\dot{\vec{x}}(t)| = |\vec{v}(t)| = v(t). \quad (2.1.64)$$

Andererseits ergibt die Kettenregel

$$\vec{v}(t) = \frac{d}{dt} \vec{x}[s(t)] = \dot{s}(t) \frac{d}{ds} \vec{x}[s(t)] = v(t) \vec{T}[s(t)]. \quad (2.1.65)$$

Dabei haben wir die Definition des Einheitstangentenvektors an die Kurve (2.1.26) verwendet. Dies macht den Geschwindigkeitsvektor intuitiv gut verständlich: Sein Betrag gibt die momentane pro Zeiteinheit zurückgelegte Strecke  $ds/dt$  an, und seine Richtung ist tangential zur Bahnkurve.

Kommen wir nun zur Beschleunigung. Sie ist die zeitliche Änderung der Geschwindigkeit pro Zeiteinheit, also

$$\vec{a}(t) = \dot{\vec{v}}(t) = \dot{v}(t)\vec{T}[s(t)] + v^2(t)\frac{d}{ds}\vec{T}[s(t)], \quad (2.1.66)$$

wobei wir von der Produkt- und Kettenregel Gebrauch gemacht haben. Wegen (2.1.27) können wir dies in der Form

$$\vec{a}(t) = \dot{v}(t)\vec{T}[s(t)] + \rho(t)v^2(t)\vec{N}[s(t)], \quad (2.1.67)$$

wobei  $\rho(t)$  der Krümmungsradius in dem momentanen Punkt der Bahn ist. Die Beschleunigung setzt sich demnach aus zwei Teilen zusammen: Der erste Term ist die **Tangentialbeschleunigung**, die sich aus der zeitlichen Änderung des Betrages der Geschwindigkeit ergibt und in tangentiale Richtung weist. Dabei kann freilich auch  $\dot{v}(t) < 0$  sein. In dem Fall wird das Teilchen „abgebremst“. Auch dies ist physikalisch gesehen eine Beschleunigung. Der zweite Term in (2.1.67) heißt **Zentripetalbeschleunigung**. Er tritt nur auf, wenn sich auch die *Richtung* der Geschwindigkeit ändert, also die Bahn in dem betreffenden Punkt von einer Geraden abweicht. Dabei ist  $\vec{N}[s(t)]$  ist der Hauptnormalenvektor an die Bahnkurve in dem betreffenden Punkt und stets  $\perp \vec{T}[s(t)]$ .

Betrachten wir als einfaches Beispiel die Bewegung auf einer Kreisbahn mit Radius  $R$  um den Ursprung in der 12-Ebene des kartesischen Koordinatensystems. Wir können sie wie folgt parametrisieren

$$\vec{r}(t) = R[\cos \varphi(t)\vec{e}_1 + \sin \varphi(t)\vec{e}_2]. \quad (2.1.68)$$

Dabei ist  $\varphi = \varphi(t)$  der Winkel des Ortsvektors zur 1-Achse des Koordinatensystems. Wir finden nun nacheinander die Geschwindigkeit und die Beschleunigung durch Ableiten nach der Zeit. Unter Verwendung der Kettenregel ergibt sich

$$\vec{v}(t) = \dot{\vec{r}}(t) = R\dot{\varphi}[-\sin \varphi(t)\vec{e}_1 + \cos \varphi(t)\vec{e}_2]. \quad (2.1.69)$$

Wir nehmen im folgenden der Einfachheit halber an, daß zu dem betrachteten Zeitpunkt  $\dot{\varphi}(t) > 0$  ist. Dann ist der Tangenteneinheitsvektor gerade durch die eckige Klammer gegeben, denn es ist

$$\vec{T}(t) = \frac{\vec{v}(t)}{|\vec{v}(t)|} = \frac{\vec{v}(t)}{R\dot{\varphi}} = -\sin \varphi(t)\vec{e}_1 + \cos \varphi(t)\vec{e}_2. \quad (2.1.70)$$

Die Beschleunigung ergibt sich wiederum durch Ableiten von (2.1.69) nach der Zeit

$$\vec{a}(t) = \dot{\vec{v}}(t) = R\ddot{\varphi}\vec{T}(t) - R\dot{\varphi}^2[\cos \varphi(t)\vec{e}_1 + \sin \varphi(t)\vec{e}_2]. \quad (2.1.71)$$

Andererseits ist

$$\frac{d\vec{T}}{ds} = \frac{d\vec{T}}{dt} \frac{1}{ds/dt} = -\dot{\varphi}[\cos \varphi(t)\vec{e}_1 + \sin \varphi(t)\vec{e}_2] \frac{1}{R\dot{\varphi}} = -\frac{1}{R}[\cos \varphi(t)\vec{e}_1 + \sin \varphi(t)\vec{e}_2] \quad (2.1.72)$$

Damit ist gemäß (2.1.27)

$$\vec{N}(t) = -\cos \varphi(t)\vec{e}_1 - \sin \varphi(t)\vec{e}_2 \quad \text{und} \quad \rho = \frac{1}{\kappa} = \left| \frac{d\vec{T}}{ds} \right|^{-1} = R \Rightarrow \rho = R. \quad (2.1.73)$$

Wie wir sehen, gilt also in der Tat (2.1.67).

Dieses einfache Beispiel liefert sogleich auch die anschauliche Bedeutung von (2.1.67): Lokal können wir die Bewegung des Massenpunktes entlang seiner Bahn als Bewegung auf dem Schmiegekreis in dem gerade durchlaufenen Punkt beschreiben: Die Zentripetalbeschleunigung ist auf den Mittelpunkt des Schmiegekreises hin gerichtet und zieht den Massenpunkt entsprechend der Krümmung der Kurve in die entsprechende Richtung. Der Betrag der Zentripetalbeschleunigung ist  $\rho\dot{\varphi}^2$ , wobei  $d\varphi = dt\dot{\varphi}$  der während des kleinen Zeitschritts  $dt$  durchlaufene Winkel ist (s. Skizze). Man nennt  $\dot{\varphi}$  auch die **momentane Winkelgeschwindigkeit** der Bewegung, entsprechend der Bedeutung in dem eben betrachteten Beispiel einer Kreisbewegung. Weiter zeigt der Vergleich von (2.1.67) mit (2.1.71), daß die Tangentialbeschleunigung in diesem Bild einer momentanen Bewegung auf dem Schmiegekreis durch die **momentane Winkelbeschleunigung**  $\ddot{\varphi}$  verursacht wird.

## 2.2 Skalare Felder, Gradient und Richtungsableitung

Im folgenden betrachten wir **Felder**. Sie sind von eminenter Bedeutung für die gesamte Physik. Wir beginnen in diesem Abschnitt mit **skalaren Feldern**. Sie ordnen jedem Punkt im Raum, gegeben durch den **Ortsvektor**  $\vec{x}$  bzgl. einem willkürlich vorgegebenen Punkt, dem Ursprung des Koordinatensystems, eine **reelle Größe** zu:  $\Phi : E^3 \rightarrow \mathbb{R}, \vec{x} \mapsto \Phi(\vec{x})$ .

Ein *Beispiel* ist die **Temperatur**: Wir können an jedem Ort im Raum die Temperatur messen und dadurch das **Temperaturfeld** definieren. Genauso ist die Dichte eines Mediums ein skalares Feld: An jedem Punkt des Raumes können wir die Masse  $\Delta m$  der Materie, in einem kleinen Volumenelement  $\Delta V$  um diesen Punkt bestimmen. Dann ist die mittlere Dichte in diesem Volumenelement  $\Delta m / \Delta V$ . Im Limes  $\Delta V \rightarrow 0$  erhalten wir die Dichte  $\rho(\vec{x})$ . In einem infinitesimal kleinen Volumenelement  $dV$  um den Punkt  $\vec{x}$  ist dann Materie von der Masse  $dm = dV \rho(\vec{x})$  enthalten. Wir werden im folgenden oft physikalische Beispiele aus der **Kontinuumsmechanik** wählen, um die mathematischen Betrachtungen zu veranschaulichen. Die eigentliche Kontinuumsmechanik (Hydrodynamik, Aerodynamik) ist allerdings (leider) nicht Gegenstand des Theoriezyklus<sup>2</sup>. Führen wir wieder eine beliebige rechtshändige Orthonormalbasis  $(\vec{e}_i)$  ein, können wir das Feld wieder als Funktion der drei Koordinaten bzw. als Funktion des entsprechenden Spaltenvektors  $\vec{x} \in \mathbb{R}^3$  ansehen. Dies impliziert das Transformationsverhalten des Feldes unter orthogonalen Transformationen. Bezeichnen wir nämlich mit  $\check{\Phi}$  das Skalarfeld als Funktion des Komponentenvektors  $\vec{x}$  bzgl. der Orthonormalbasis  $(\vec{e}_i)$  und mit  $\check{\Phi}'$  als Funktion des Komponentenvektors  $\vec{x}'$  bzgl. der Orthonormalbasis  $(\vec{e}'_i)$ <sup>2</sup>, so muß definitionsgemäß

$$\check{\Phi}(\vec{x}) = \check{\Phi}'(\vec{x}') = \Phi(\vec{x}) \quad (2.2.1)$$

gelten, denn es ist klar, daß z.B. die Dichte eines Gases nicht von der Wahl des Koordinatensystems abhängt sondern lediglich von dem Punkt im Raum, an dem wir diese Dichte messen. Betrachten wir die Basistransformation von einer zur anderen Orthonormalbasis gemäß (1.3.4), folgt wegen (1.3.5)

$$\vec{x}' = \hat{T}\vec{x}, \quad \check{\Phi}'(\vec{x}') = \check{\Phi}(\vec{x}) = \check{\Phi}(\hat{T}^{-1}\vec{x}'). \quad (2.2.2)$$

Wir bezeichnen alle Funktionen  $\mathbb{R}^3 \rightarrow \mathbb{R}$ , die sich bzgl. orthogonaler Transformationen ihrer Argumente so verhalten wie in (2.2.2) angegeben als **skalare Felder**. Sie beschreiben Größen als Funktionen vom Ort.

Wie für Funktionen einer reellen Veränderlichen können wir auch für Funktionen, die von mehr als einer Variablen abhängen, **Ableitungen** definieren, die die Änderungen der Funktionen bei kleinen Änderungen der Parameter beschreiben. Dazu definieren wir zunächst die **partiellen Ableitungen** der Funktion  $\check{\Phi}(\vec{x}) \equiv \check{\Phi}(x^1, x^2, x^3)$ . Dabei ist die partielle Ableitung nach der Variablen  $x^i$  durch den Limes des entsprechenden Differenzenquotienten bei Änderungen nur dieser einen Variablen  $x^i$  definiert, während die anderen beiden Variablen konstant gehalten werden. Für die partielle Ableitung nach  $x^1$  bedeutet dies

$$\frac{\partial}{\partial x^1} \check{\Phi}(x^1, x^2, x^3) = \partial_1 \check{\Phi}(x^1, x^2, x^3) = \lim_{\Delta x^1 \rightarrow 0} \frac{\check{\Phi}(x^1 + \Delta x^1, x^2, x^3) - \check{\Phi}(x^1, x^2, x^3)}{\Delta x^1} \quad (2.2.3)$$

und entsprechend analog für die partiellen Ableitungen  $\partial_2$  und  $\partial_3$  nach  $x^2$  bzw.  $x^3$ . Falls diese partiellen Ableitungen in einem gegebenen Punkt  $\vec{x} = (x^1, x^2, x^3)^T$  existiert, heißt die Funktion  $\check{\Phi}$  dort **partiell differenzierbar**. Ist sie in jedem Punkt eines bestimmten offenen Gebiets  $E^3 \subseteq \mathbb{R}^3$  partiell differenzierbar, heißt sie in diesem Gebiet **partiell differenzierbar**. Falls die partiellen Ableitungen in diesem Fall allesamt stetig sind, heißt die Funktion dort **stetig differenzierbar**. In der Physik gehen wir davon aus, daß die meisten dort relevanten Funktionen bzw. Felder **lokal stetig differenzierbar** sind, d.h. bis auf einige **singuläre Punkte**

<sup>2</sup>Gewöhnlich nehmen Physiker die subtile Unterscheidung in der Bezeichnung der Felder nicht vor und schreiben einfach  $\Phi(\vec{x})$ , d.h. sie verwenden dasselbe Symbol für die Funktion  $\Phi : E^3 \rightarrow \mathbb{R}$  und die Funktion  $\check{\Phi} : \mathbb{R}^3 \rightarrow \mathbb{R}$ . Es ist aber gerade zu Beginn der Beschäftigung mit Vektoranalysis wichtig, diese Unterscheidung in Erinnerung zu behalten. Freilich definieren sich bei vorgegebener Orthonormalbasis  $(\vec{e}_i)$ , auf die sich die Komponenten  $\vec{x}$  von  $\vec{x}$  beziehen, diese beiden Funktionen umkehrbar eindeutig wechselseitig.

existiert um jeden Punkt eines bestimmten Gebietes eine Kugel (bzw. **Umgebung**), in der die Funktion stetig differenzierbar ist. Manchmal bezeichnet man solche Funktionen auch als **lokal glatt**.

Es gelten für partielle Ableitungen freilich alle Rechenregeln wie für Ableitungen von Funktionen einer unabhängigen Variablen. Besonders wichtig ist, daß die partiellen Ableitungen **lineare Operatoren** sind, d.h. sind  $\tilde{\Phi}_1$  und  $\tilde{\Phi}_2$  lokal glatt, so gilt für beliebige  $\lambda_1, \lambda_2 \in \mathbb{R}$  in jedem regulären Punkt der Funktion

$$\partial_j(\lambda_1 \tilde{\Phi}_1 + \lambda_2 \tilde{\Phi}_2) = \lambda_1 \partial_j \tilde{\Phi}_1 + \lambda_2 \partial_j \tilde{\Phi}_2. \quad (2.2.4)$$

Weitere wichtige Ableitungsregeln sind die **Produkt- und Quotientenregel**

$$\begin{aligned} \partial_j(\tilde{\Phi}_1 \tilde{\Phi}_2) &= (\partial_j \tilde{\Phi}_1) \tilde{\Phi}_2 + \tilde{\Phi}_1 (\partial_j \tilde{\Phi}_2), \\ \partial_j \left( \frac{\tilde{\Phi}_1}{\tilde{\Phi}_2} \right) &= \frac{(\partial_j \tilde{\Phi}_1) \tilde{\Phi}_2 - \tilde{\Phi}_1 \partial_j \tilde{\Phi}_2}{\tilde{\Phi}_2^2}, \end{aligned} \quad (2.2.5)$$

wobei die Quotientenregeln natürlich nur an Stellen sinnvoll ist, wo  $\tilde{\Phi}_2 \neq 0$  ist.

Eine Erweiterung erfährt allerdings die Kettenregel. Dazu betrachten wir eine beliebige differenzierbare Kurve  $\bar{x}(t)$ , die durch ein Gebiet verläuft, wo das Skalarfeld  $\tilde{\Phi}(\bar{x})$  stetig differenzierbar ist. Dann können wir fragen, wie sich dieses Feld entlang dieser Bahnkurve ändert und die **totale Ableitung** bilden. Für diese gilt

$$\frac{d}{dt} \tilde{\Phi}[\bar{x}(t)] = \sum_{j=1}^3 \dot{x}_j \partial_j \tilde{\Phi}[\bar{x}(t)]. \quad (2.2.6)$$

Um diese Behauptung zu beweisen, betrachten wir den Differenzenquotienten

$$\frac{\Delta \tilde{\Phi}}{\Delta t} = \frac{\tilde{\Phi}[\bar{x}(t + \Delta t)] - \tilde{\Phi}[\bar{x}(t)]}{\Delta t}. \quad (2.2.7)$$

Setzen wir nun  $\Delta \bar{x} = \bar{x}(t + \Delta t) - \bar{x}(t)$ , ergibt sich

$$\frac{\Delta \tilde{\Phi}}{\Delta t} = \frac{\tilde{\Phi}(\bar{x} + \Delta x) - \tilde{\Phi}(\bar{x})}{\Delta t}. \quad (2.2.8)$$

Wir fügen addieren und subtrahieren nun im Zähler dieses Ausdrucks geeignete Zusatzterme, die Differenzen ergeben, wobei sich im Argument immer nur eine der Variablen  $x^j$  ändert:

$$\begin{aligned} \frac{\Delta \tilde{\Phi}}{\Delta t} &= \frac{\tilde{\Phi}(x^1 + \Delta x^1, x^2 + \Delta x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2 + \Delta x^2, x^3 + \Delta x^3)}{\Delta t} \\ &\quad + \frac{\tilde{\Phi}(x^1, x^2 + \Delta x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2, x^3 + \Delta x^3)}{\Delta t} \\ &\quad + \frac{\tilde{\Phi}(x^1, x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2, x^3)}{\Delta t}. \end{aligned} \quad (2.2.9)$$

Wir können nun den ersten Term wie folgt umformen:

$$\begin{aligned} &\frac{\tilde{\Phi}(x^1 + \Delta x^1, x^2 + \Delta x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2 + \Delta x^2, x^3 + \Delta x^3)}{\Delta t} \\ &= \frac{\tilde{\Phi}(x^1 + \Delta x^1, x^2 + \Delta x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2 + \Delta x^2, x^3 + \Delta x^3)}{\Delta x^1} \frac{\Delta x^1}{\Delta t}. \end{aligned} \quad (2.2.10)$$

Da  $\tilde{\Phi}$  voraussetzungsgemäß in einer Umgebung von  $\bar{x}$  stetig differenzierbar und auch die Kurve bei  $t$  differenzierbar ist, ergibt sich für diesen Ausdruck im Limes  $\Delta t \rightarrow 0$

$$\frac{\tilde{\Phi}(x^1 + \Delta x^1, x^2 + \Delta x^2, x^3 + \Delta x^3) - \tilde{\Phi}(x^1, x^2 + \Delta x^2, x^3 + \Delta x^3)}{\Delta t} \rightarrow \partial_1 \tilde{\Phi}[\bar{x}(t)] \dot{x}^1. \quad (2.2.11)$$

Genauso geht man für die beiden anderen Terme in (2.2.9) vor, woraus schließlich die Behauptung (2.2.6) folgt.

Nun wollen wir zeigen, daß diese totale Ableitung eine von der Wahl des kartesischen Koordinatensystems unabhängige Bedeutung besitzt. Dazu bemerken wir, daß (2.2.6) in der Form

$$\frac{d}{dt} \tilde{\Phi}[\bar{x}(t)] = \dot{\bar{x}}(t) \cdot \bar{\nabla} \tilde{\Phi}[\bar{x}(t)] \quad (2.2.12)$$

geschrieben werden kann, wenn wir

$$\bar{\nabla} \tilde{\Phi} = \begin{pmatrix} \partial_1 \tilde{\Phi} \\ \partial_2 \tilde{\Phi} \\ \partial_3 \tilde{\Phi} \end{pmatrix} \quad (2.2.13)$$

definieren. Dies legt nahe, daß  $\bar{\nabla} \tilde{\Phi}$  bzgl. *kartesischer Koordinaten* ein Vektor ist. Dies werden wir auch sogleich beweisen. Man nennt diesen Vektor den **Gradienten** des Skalarfeldes  $\Phi(\vec{x})$ :

$$\vec{\nabla} \Phi(\vec{x}) = \sum_{j=1}^3 \vec{e}_j \partial_j \tilde{\Phi}(\bar{x}). \quad (2.2.14)$$

Wir wollen nun zeigen, daß dieses **Vektorfeld** tatsächlich unabhängig von der Wahl des *kartesischen* Koordinatensystems ist. Betrachten wir dazu wieder die Basistransformation (1.3.4) für die Basisvektoren und die Komponenten (1.3.5). Verwenden wir nun (2.2.2), um die Summe auf der rechten Seite von (2.2.14) bzgl. der neuen Basis zu berechnen, folgt

$$\sum_{j=1}^3 \vec{e}'_j \partial'_j \tilde{\Phi}'(\bar{x}') = \sum_{j,k=1}^3 \vec{e}'_j \frac{\partial x^k}{\partial x'^j} \partial_k \tilde{\Phi}(\bar{x}). \quad (2.2.15)$$

Dabei haben wir die Kettenregel (2.2.6) für mehrere Veränderliche auf die partielle Ableitung nach  $x'^j$  verwendet. Gemäß (2.2.4) gilt (*warum?*)

$$\bar{x}' = \hat{T} \bar{x} \Rightarrow \bar{x} = \hat{U} \bar{x}' \Rightarrow \frac{\partial x^k}{\partial x'^j} = U^k_j. \quad (2.2.16)$$

Setzen wir dies in (1.3.15) ein, folgt

$$\sum_{j=1}^3 \vec{e}'_j \partial'_j \tilde{\Phi}'(\bar{x}') = \sum_{j,k=1}^3 \vec{e}'_j U^j_k \partial_k \tilde{\Phi}(\bar{x}). \quad (2.2.17)$$

Gemäß (1.3.4) folgt

$$\sum_{j=1}^3 \vec{e}'_j \partial'_j \tilde{\Phi}'(\bar{x}') = \sum_{k=1}^3 \vec{e}_k \partial_k \tilde{\Phi}(\bar{x}) \stackrel{(2.2.14)}{=} \vec{\nabla} \Phi(\vec{x}). \quad (2.2.18)$$

Dies zeigt, daß der Gradient tatsächlich unabhängig von der gewählten Orthonormalbasis ist. Dieser besitzt also eine von der kartesischen Basis unabhängige Bedeutung.

Betrachten wir die Parametrisierung der Kurve in (2.2.6) als Funktion der Bogenlänge, folgt

$$\frac{d}{ds}\Phi[\vec{x}(s)] = \frac{d\vec{x}}{ds} \cdot \vec{\nabla}\Phi[\vec{x}(s)] = \vec{T}(s) \cdot \vec{\nabla}\Phi[\vec{x}(s)]. \quad (2.2.19)$$

Dabei ist  $\vec{T}(s)$  der Tangenteneinheitsvektor an die Kurve. Die geometrische Bedeutung ist dabei klar: Bewegt man sich eine infinitesimale Strecke  $ds$  entlang der Kurve mit Tangenteneinheitsvektor  $\vec{T}(s)$  an der betreffenden Stelle, so ändert sich der Wert des Skalarfeldes  $ds\vec{T}(s) \cdot \vec{\nabla}\Phi[\vec{x}(s)]$ . Diese Änderung ist dabei freilich unabhängig von der Wahl des kartesischen Koordinatensystems, das wir dazu benutzen, um den Gradienten über die partielle Ableitungen der entsprechenden Funktion  $\check{\Phi}(\vec{x})$  zu berechnen. Man bezeichnet daher (2.2.19) als die **Richtungsableitung** des Skalarfeldes an dem betreffenden Punkt in der durch  $\vec{T}(s)$  gegebenen Richtung.

Aus der obigen Rechnung folgt noch das Verhalten der Darstellung des Gradienten durch die partiellen Ableitungen unter **orthogonalen Transformationen**. Wegen (1.3.17) ist nämlich

$$\partial'_j \check{\Phi}'(\vec{x}') = \sum_{k=1}^3 U_{kj} \partial_k \check{\Phi}(\vec{x}). \quad (2.2.20)$$

In Matrix-Vektor-Produktschreibweise bedeutet nun diese Gleichung

$$\vec{\nabla}' \check{\Phi}'(\vec{x}') = \hat{U}^T \vec{\nabla} \check{\Phi}(\vec{x}) = \hat{T} \vec{\nabla} \check{\Phi}(\vec{x}). \quad (2.2.21)$$

Das bedeutet, daß sich der Gradient bzgl. *orthogonaler Transformationen* tatsächlich wie ein Vektor verhält<sup>3</sup>.

### 2.2.1 Der Gradient in beliebigen Basen

Wir betrachten nun noch den Gradienten im allgemeineren Fall beliebiger Basen  $\vec{b}_k$  im strikten Ricci-Kalkül mit oberen und unteren Indizes wie in Abschnitt 1.8 eingeführt. Wir verwenden wieder die Einsteinsche Summationskonvention und betrachten wieder zwei Basen

$$\vec{b}_j = T^k_j \vec{b}'_k, \quad \vec{b}'_k = U^j_k \vec{b}_j \quad \text{mit} \quad \hat{U} = \hat{T}^{-1}. \quad (2.2.22)$$

Für ein Vektorfeld  $\Phi$  gilt dann

$$\check{\Phi}'(\vec{x}') = \check{\Phi}(\vec{x}) = \check{\Phi}(\hat{U}\vec{x}'), \quad (2.2.23)$$

wobei

$$\vec{x} = x^j \vec{e}_j = x'^k \vec{e}'_k, \quad \vec{x}' = \hat{T}\vec{x}, \quad \vec{x} = \hat{U}\vec{x}' \quad (2.2.24)$$

die Komponenten des Ortsvektors definieren.

Wir zeigen nun, daß die partiellen Ableitungen nach diesen Komponenten **kovariante Vektorfeldkomponenten** definieren. In der Tat folgt unter Verwendung der Kettenregel

$$\partial'_j \check{\Phi}' = \partial_k \check{\Phi} \frac{\partial x^k}{\partial x'^j} = U^k_j \partial_k \check{\Phi}, \quad (2.2.25)$$

wobei wir zur Abkürzung

$$\partial'_j = \frac{\partial}{\partial x'^j}, \quad \partial_k = \frac{\partial}{\partial x^k} \quad (2.2.26)$$

<sup>3</sup>Wir bemerken zum Schluß noch, daß sich der Gradient bzgl. allgemeiner Base nicht wie ein Vektor sondern wie ein sogenannte Kovektor, d.h. eine lineare Abbildung  $E^3 \rightarrow \mathbb{R}$  transformiert. Wir gehen darauf in dieser Vorlesung nicht genauer ein, da der allgemein kovariante Tensor-Index-Kalkül noch nicht benötigt wird. Der interessierte Leser sei auf [Hee05] verwiesen.

### 2.3 · Extremwertaufgaben

eingeführt haben. In der Tat zeigt (2.2.25), daß sich die  $\partial_k \tilde{\Phi}$  wie kovariante Vektorkomponenten, also wie die Basisvektoren transformieren. Führen wir nun gemäß (1.8.4) und (1.8.12) die ko- und kontravarianten Metrikkomponenten

$$g_{jk} = \vec{b}_j \cdot \vec{b}_k, \quad (g_{jk}) = \overleftrightarrow{g}, \quad (g^{jk}) = \overleftarrow{g} = \overleftrightarrow{g}^{-1} \quad (2.2.27)$$

ein, ergibt sich der Gradient als Vektorfeld

$$\vec{\nabla} \Phi = \vec{b}_j g^{jk} \partial_k \tilde{\Phi}. \quad (2.2.28)$$

Es ist klar, daß sich dabei die kontravarianten Komponenten eines Vektorfelds  $\vec{V}$  unter allgemeinen Basis-Transformationen gemäß

$$\vec{V}'(\vec{x}') = \hat{T} \vec{V}(\vec{x}) = \hat{T} \vec{V}(\hat{U} \vec{x}') \quad (2.2.29)$$

transformieren.

## 2.3 Extremwertaufgaben

Wie bei den Funktionen einer Veränderlichen (vgl. Anhang A.6.3) ist es auch für skalare Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  oft interessant, **lokale Extremstellen** zu suchen. Wir können dieses Problem auf das entsprechende Problem für Funktionen einer Veränderlichen zurückführen. Dazu definieren wir

$$g(t) = f(\vec{x}_0 + t\vec{n}). \quad (2.3.1)$$

Dabei sind  $\vec{x}_0$  und  $\vec{n}$  beliebige Vektoren in  $\mathbb{R}^n$ . Wir nehmen im folgenden an, die Funktion  $f$  sei zweimal stetig partiell differenzierbar. Aus den Betrachtungen in Anhang A.6.3 wird klar, daß  $f$  an der Stelle  $\vec{x}_0$  nur dann ein Extremum annehmen kann, wenn für alle  $\vec{n} \in \mathbb{R}^n$

$$\dot{g}(0) = \vec{n} \cdot \vec{\nabla} f(\vec{x}_0) = 0 \quad (2.3.2)$$

ist. Da wir für  $\vec{n}$  beliebige Vektoren einsetzen dürfen, folgt als *notwendige Bedingung für das Vorliegen eines Extremums* bei  $\vec{x}_0$ , daß der Gradient der skalaren Funktion  $f$  dort verschwinden muß

$$\vec{\nabla} f(\vec{x}_0) = \vec{0}. \quad (2.3.3)$$

Weiter wissen wir aus Anhang A.6.3, daß für das Vorliegen eines **lokalen Minimums** die Bedingung

$$\ddot{g}(0) = \sum_{i,j=1}^n n_i n_j \partial_i \partial_j f(\vec{x}_0) > 0 \quad (2.3.4)$$

*hinreichend* ist. Da  $f$  voraussetzungsgemäß zweimal stetig partiell differenzierbar ist, ist die **Hesse-Matrix**

$$H_{ij}(\vec{x}_0) = \partial_i \partial_j f(\vec{x}_0) \quad (2.3.5)$$

**symmetrisch**. Da (2.3.4) für alle  $\vec{n} \in \mathbb{R}^n$  gelten muß, bedeutet dies, daß ein lokales Minimum vorliegt, falls die durch die Hessematrix definierte **quadratische Form** für alle  $\vec{n} \neq \vec{0}$

$$Q(\vec{n}) = \sum_{i,j=1}^n H_{ij}(\vec{x}_0) n_i n_j > 0 \Leftrightarrow Q(\vec{n}) = \vec{n}^T \hat{H}(\vec{x}_0) \vec{n} > 0 \quad (2.3.6)$$

ist. Man nennt solche Bilinearformen **positiv definit**. Wir kommen auf Kriterien für positive Definitheit von Bilinearformen im nächsten Kapitel noch zurück.

Betrachten wir als einfachstes Beispiel die Funktion zweier Veränderlicher

$$f(\bar{x}) = 2(x^1)^2 + 3(x^2)^2. \quad (2.3.7)$$

Offenbar besitzt  $f$  in  $\bar{x}_0 = \bar{0}$  ein lokales Minimum. Dies können wir direkt dem eben hergeleiteten hinreichenden Kriterium entnehmen. Der Gradient ist

$$\bar{\nabla}f(\bar{x}) = \begin{pmatrix} \partial_1 f(\bar{x}) \\ \partial_2 f(\bar{x}) \end{pmatrix} = \begin{pmatrix} 4x^1 \\ 6x^2 \end{pmatrix}. \quad (2.3.8)$$

Der Gradient verschwindet für  $x^1 = x^{10} = x^2 = x^{20} = 0$ . Die Hesse-Matrix ist offenbar  $\hat{H}(\bar{x}_0) = \text{diag}(4, 6)$  und damit

$$Q(\bar{n}) = \bar{n}^T \hat{H}(\bar{x}_0) \bar{n} = 4n_1^2 + 6n_2^2 > 0 \quad (2.3.9)$$

für alle  $\bar{n} \neq 0$ . Die Hesse-Matrix ist also tatsächlich positiv definit, und damit besitzt  $f$  bei  $\bar{x}_0 = \bar{0}$  ein Minimum.

Nun ergibt sich bei Funktionen mehrerer Veränderlicher oft auch die Frage, wann sie unter bestimmten **Nebenbedingungen** extremal werden.

Ein praktisches Beispiel ist die Aufgabe, eine zylindrische Blechdose mit vorgegebenem Volumen  $V$  zu konstruieren, die eine möglichst kleine Oberfläche besitzt, d.h. den geringsten Materialaufwand erfordert. Ist  $R$  der Radius und  $h$  die Höhe der Blechdose, so sind Oberfläche und Volumen durch (*nachrechnen!*)

$$O(r, h) = 2\pi r h + 2\pi r^2, \quad V(r, h) = \pi r^2 h \quad (2.3.10)$$

gegeben. Wir suchen also diejenigen Werte für  $\bar{x} = (r, h)$ , daß  $O(r, h) = O(\bar{x})$  minimal wird unter der Bedingung, daß  $V(r, h) = V = \text{const}$  vorgegeben ist.

Wir lösen diese Aufgaben nun auf zwei Arten. In diesem Fall läßt sich nämlich die eine Variable (wir wählen hier  $h$ ) aus der Nebenbedingung als Funktion der anderen Variablen (hier also  $r$ ) darstellen:

$$h(r) = \frac{V}{\pi r^2}. \quad (2.3.11)$$

Demnach müssen wir also nur dasjenige  $r$  finden, für welches

$$\tilde{O}(r) = O[r, h(r)] = \frac{2V}{r} + 2\pi r^2 \quad (2.3.12)$$

minimal wird. Damit ist die Aufgabe auf eine Extremwertbestimmung für eine Funktion einer Veränderlichen  $\tilde{O}(r)$  (freilich nun ohne Nebenbedingungen) zurückgeführt. Zunächst muß die erste Ableitung verschwinden:

$$\tilde{O}'(r) = -\frac{2V}{r^2} + 4\pi r \stackrel{!}{=} 0. \quad (2.3.13)$$

Dies ist erfüllt für

$$r = r_0 = \left(\frac{V}{2\pi}\right)^{1/3} \Rightarrow h = h(r_0) = \frac{V}{\pi r_0} = 2r_0. \quad (2.3.14)$$

Nun prüfen wir noch nach, daß diese Lösung tatsächlich ein Minimum der Oberfläche unter der Nebenbedingung konstanten Volumens ist. Dazu berechnen wir

$$\tilde{O}''(r_0) = +\frac{4V}{r_0^3} + 4\pi = 12\pi > 0, \quad (2.3.15)$$

d.h.  $\tilde{O}$  wird tatsächlich für  $r = r_0$  minimal. Demnach besitzt diejenige Büchse vorgegebenen Volumens  $V$  die kleinste Oberfläche, für deren Höhe  $h = 2r_0 = (4V/\pi)^{1/3}$  ist.

In ähnlichen Extremwertaufgaben mit Nebenbedingungen kann es aber vorkommen, daß man nicht so einfach einen Teil der Variablen mit Hilfe der Nebenbedingungen eliminieren kann. In diesem Fall verwendet man die **Methode der Lagrange-Multiplikatoren**. Wir führen dies wieder anhand unseres Beispiels vor. Wir schreiben dazu die Nebenbedingung in der Form

$$N(r, b) = 0 \quad \text{mit} \quad N(r, b) = \pi r^2 b - V. \quad (2.3.16)$$

Dann betrachten wir die Extremwertaufgabe der Funktion

$$f(r, b, \lambda) = O(r, b) + \lambda N(r, b) \quad (2.3.17)$$

ohne Nebenbedingungen. Die notwendige Bedingung, daß diese Funktion extremal wird, verlangt, daß  $\bar{\nabla} f = (\partial_r f, \partial_b f, \partial_\lambda f)^T = \bar{0}$  wird. Aufgrund der Einführung des Lagrange-Parameters wird die Nebenbedingung also als eine der notwendigen Extremalbedingungen berücksichtigt. Wir erhalten aus diesen Bedingungen dieselbe Lösung für  $r$  und  $b$  wie oben in (2.3.14) angegeben. Auf hinreichende Bedingungen für das Vorliegen eines Extremums mit der Methode der Lagrange-Multiplikatoren gehen wir hier nicht ein. Die Frage, ob tatsächlich ein Minimum oder Maximum an der gefundenen Lösung der hinreichenden Bedingung vorliegt, ist also stets noch gesondert zu untersuchen.

## 2.4 Vektorfelder, Divergenz und Rotation

Es liegt nun nahe, neben skalaren ortsabhängigen Größen (Skalarfeldern) auch vektorielle ortsabhängige Größen (**Vektorfelder**) zu betrachten. Typische Beispiele sind das Geschwindigkeitsfeld  $\vec{v}(\vec{x})$  einer Flüssigkeit oder eines Gases: Es gibt die Geschwindigkeit eines sich gerade am Ort  $\vec{x}$  befindlichen kleinen Flüssigkeitselements an. Andere typische Beispiele sind das elektrische oder magnetische Feld<sup>4</sup>.

Mathematisch gesehen haben wir es mit einer Abbildung  $\vec{V} : E^3 \rightarrow V, \vec{x} \mapsto \vec{V}(\vec{x})$  zu tun. Um mit diesen Feldern bequem rechnen zu können, führen wir wieder eine beliebige kartesische rechtshändige Basis  $(\vec{e}_j)$  ein. Dann ist

$$\vec{V}(\vec{x}) = \sum_{j=1}^3 \vec{e}_j V^j(\vec{x}) \quad \text{mit} \quad \vec{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}. \quad (2.4.1)$$

Setzen wir wieder  $\bar{V} = (V^1, V^2, V^3)^T$ , impliziert also die Einführung der Orthonormalbasis  $(\vec{e}_j)$  eine umkehrbar eindeutige Abbildung des Vektorfeldes  $\vec{V} \in E^3$  auf ein Vektorfeld  $\bar{V} : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \vec{x} \mapsto \bar{V}(\vec{x})$ .

Das Transformationsverhalten unter orthogonalen Basistransformationen ist dann wieder gemäß (1.3.4) bzw. (1.3.5) gegeben:

$$\vec{V}(\vec{x}) = \bar{V}'(\vec{x}') = \hat{T} \bar{V}(\vec{x}) = \hat{T} \bar{V}(\hat{T}^{-1} \vec{x}'). \quad (2.4.2)$$

Im vorigen Abschnitt haben wir gesehen, daß bzgl. orthogonaler Transformationen der Differentialoperator<sup>5</sup>  $\bar{\nabla}$  Vektorcharakter besitzt.

Es liegt dann nahe, daß die **Divergenz eines Vektorfeldes**

$$\text{div } \vec{V}(\vec{x}) = \bar{\nabla} \cdot \vec{V}(\vec{x}) = \bar{\nabla} \cdot \bar{V}(\vec{x}) = \sum_{j=1}^3 \partial_j V_j(\vec{x}). \quad (2.4.3)$$

<sup>4</sup>I.a. werden solche Felder auch noch von der Zeit abhängen. Hier betrachten wir aber nur die Ortsabhängigkeit zu einem festen Zeitpunkt und lassen entsprechend das Zeitargument der Einfachheit halber weg.

<sup>5</sup>Aufgrund der Form des Symbols wird dieser Operator als **Nabla-Operator** bezeichnet. Der Name kommt von einem hebräischen harfenähnlichen Saiteninstrument her, das diese typische Form aufweist.

ein Skalarfeld ist. Wir können sofort zeigen, daß es sich dabei tatsächlich um ein Skalarfeld handelt, denn es gilt gemäß (2.2.21)

$$\vec{\nabla}' \cdot \vec{V}'(\vec{x}') = (\hat{T}^T \vec{\nabla}) \cdot \hat{T} \vec{V}(\vec{x}) = \vec{\nabla}^T \hat{T}^T \hat{T} \vec{V}(\vec{x}) = \vec{\nabla}^T \vec{V}(\vec{x}) = \vec{\nabla} \cdot \vec{V}(\vec{x}). \quad (2.4.4)$$

Die anschauliche Bedeutung dieses Differentialoperators werden wir weiter unten in Abschnitt 2.7 diskutieren.

Wir haben nun in (1.6.71) gesehen, daß das Vektorprodukt zweier Vektoren sich unter *orientierungserhaltenden orthogonalen Transformationen* (also *Drehungen*) wie ein Vektor verhält. Daraus folgt, daß die **Rotation eines Vektorfeldes**

$$\text{rot } \vec{V} = \vec{\nabla} \times \vec{V} = \sum_{j=1}^3 \vec{e}_j [\vec{\nabla} \times \vec{V}(\vec{x})]_j = \sum_{j,k,l=1}^3 \vec{e}_j \epsilon_{jkl} \partial_k V^l(\vec{x}) \quad (2.4.5)$$

ein Vektorfeld ist<sup>6</sup>. Auf die anschauliche Bedeutung der Rotation als **Wirbeldichte** einer Flüssigkeitsströmung werden wir weiter unten noch zurückkommen. In kartesischen Komponenten ausgeschrieben ergibt sich

$$\vec{\nabla} \times \vec{V} = \begin{pmatrix} \partial_2 V^3 - \partial_3 V^2 \\ \partial_3 V^1 - \partial_1 V^3 \\ \partial_1 V^2 - \partial_2 V^1 \end{pmatrix}. \quad (2.4.6)$$

### 2.4.1 Divergenz und Rotation in beliebigen Basen

Wir betrachten nun noch Divergenz und Rotation bzgl. einer beliebigen nicht notwendig kartesischen rechtshändigen Basis. Wir haben in Abschnitt 2.2.1 gesehen, daß die Partiiellen Ableitungen nach den Komponenten des Ortsvektors bzgl. beliebiger Basen sich wie kovariante Vektorkomponenten verhalten. Demnach sollte auch bzgl. beliebiger Koordinaten

$$\vec{\nabla} \cdot \vec{V} = \partial_j V^j \quad (2.4.7)$$

gelten. Um dies zu zeigen, verwenden wir eine beliebige kartesische Basis  $\vec{e}'_k$ . Mit

$$\vec{b}_j = T^k_j \vec{e}'_k, \quad \vec{e}'_k = U^j_k \vec{b}_j, \quad \hat{U} = \hat{T}^{-1} \quad (2.4.8)$$

folgt

$$\vec{\nabla} \cdot \vec{V} = \partial'_k V'^k = \frac{\partial q^j}{\partial x'^k} \partial_j (T^k_l V^l) = U^j_k T^k_l \partial_j V^l = \delta^j_l \partial_j V^l = \partial_j V^j. \quad (2.4.9)$$

Man beachte, daß es hier wichtig ist, daß die Matrizen der Basistransformation  $\hat{T}$  und  $\hat{U} = \hat{T}^{-1}$  konstant sind.

Für die Rotation müssen wir auf den Levi-Civita-Tensor (1.8.7) bzw. (1.8.25) zurückgreifen. In den kartesischen Koordinaten haben wir wegen  $g'_{jk} = g'^{jkl} = \delta_{jk} = \delta^{jk}$  und also  $g = \det \overleftrightarrow{g} = 1$  und weil  $V'_j = g'_{jk} V'^k = V^j$

$$\vec{\nabla} \times \vec{V} = \epsilon^{jkl} \vec{e}_j \partial'_k V'_l. \quad (2.4.10)$$

Da da für die kartesische Basis  $\Delta'^{jkl} = \epsilon^{jkl}$  ist und der  $\Delta^{jkl}$  kontravariante Tensorkomponenten sind, gilt folglich dieselbe Gleichung auch für die Komponenten bzgl. der allgemeinen Basis, d.h.

$$\vec{\nabla} \times \vec{V} = \Delta^{jkl} \vec{b}_j \partial_k V_l = \frac{1}{\sqrt{g}} \epsilon^{jkl} \vec{b}_j \partial_k V_l. \quad (2.4.11)$$

Wie bei der Divergenz können wir auch explizit mit Hilfe der Basistransformationen nachrechnen, daß in der Tat die rechten Seiten von (2.4.10) und (2.4.11) identisch sind. Wir müssen dabei nur voraussetzen, daß sowohl die kartesische als auch die allgemeine Basis rechtshändig orientiert sind, d.h. daß  $\det \hat{T} = 1 / \det \hat{U} > 0$  ist.

<sup>6</sup>In der englischsprachigen Literatur heißt die Rotation eines Vektorfeldes  $\text{curl}$ :  $\text{curl } \vec{V} = \vec{\nabla} \times \vec{V}$ .

## 2.5 Potentialfelder

Für die Physik ist eine sehr wichtige Fragestellung, in welchen Fällen wir ein vorgegebenes Vektorfeld  $\vec{V}(\vec{x})$  als **Gradient eines Skalarfeldes** schreiben können. In der Mechanik ist dies insbesondere für **Kraftfelder** interessant. In (2.1.58) haben wir z.B. das Kraftfeld für ein Partikelchen (Planet), das sich im Schwerfeld eines anderen im Ursprung des Koordinatensystems fixierten Punkt (Sonne) bewegt, angegeben. Die Newtonsche Gravitationskraft lautet demnach

$$\vec{F}_G(\vec{x}) = -\frac{\gamma m M}{r^2} \frac{\vec{x}}{r} \quad \text{mit} \quad r = |\vec{x}|. \quad (2.5.1)$$

Wir können nun fragen, ob diese Kraft der Gradient eines Skalarfeldes ist. Aus Gründen, die gegen Ende dieses Abschnitts klar werden, definiert man das Potential so, daß die Kraft durch den negativen Gradienten gegeben ist:

$$\vec{F}_G(\vec{x}) = -\vec{\nabla} \Phi_G(\vec{x}). \quad (2.5.2)$$

Dazu schreiben wir die erste Komponente dieses Ansatzes in kartesischen Koordinaten hin:

$$F_{G1} = F_G^1 = -\frac{\gamma m M}{[(x^1)^2 + (x^2)^2 + (x^3)^2]^{3/2}} x^1 \stackrel{!}{=} -\partial_1 \check{\Phi}_G(\vec{x}). \quad (2.5.3)$$

Wir haben also die *partielle Ableitung* bzgl.  $x^1$  als Funktion dieser Variablen gegeben. Die Komponenten  $x^2$  und  $x^3$  können wir dabei aufgrund der Definition der partiellen Ableitung als Konstanten ansehen. Wir müssen also den obigen Ausdruck nur bzgl.  $x^1$  *integrieren*, denn die Integration ist die Umkehrung der Ableitung. Kürzen wir das Vorzeichen auf beiden Seiten der Gleichung heraus, ergibt sich also aus (2.5.3)

$$\check{\Phi}_G(\vec{x}) = \gamma m M \int dx^1 \frac{x^1}{[(x^1)^2 + (x^2)^2 + (x^3)^2]^{3/2}}. \quad (2.5.4)$$

Dieses *unbestimmte* Integral lösen wir am bequemsten mit Hilfe der *Substitutionsmethode*. Wir führen dazu als Integrationsvariable  $u = (x^1)^2 + (x^2)^2 + (x^3)^2$  ein. Für das Differential folgt  $du = dx^1 \partial u / \partial x^1 = dx^1 2x^1$ , und folglich gilt

$$\check{\Phi}_G(\vec{x}) = \gamma m M \frac{1}{2} \int du u^{-3/2} = -\gamma m M u^{-1/2} + C_1(x^2, x^3) = -\frac{\gamma m M}{\sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}} + C_1(x^2, x^3). \quad (2.5.5)$$

Dabei haben wir berücksichtigt, daß ein unbestimmtes Integral nur bis auf eine von  $x^1$  unabhängige Funktion bestimmt ist. Diese Funktion kann aber offensichtlich noch von  $x^2$  und  $x^3$  abhängen. Deshalb haben wir die Bezeichnung  $C_1(x^2, x^3)$  gewählt.

Nun können wir die zweite Komponente von (2.5.2) betrachten. Es ergibt sich

$$F_{G2}(\vec{x}) = -\frac{\gamma m M}{[(x^1)^2 + (x^2)^2 + (x^3)^2]^{3/2}} x^2 = -\partial_2 \check{\Phi}_G(\vec{x}) \stackrel{(2.5.5)}{=} -\frac{\gamma m M}{[(x^1)^2 + (x^2)^2 + (x^3)^2]^{3/2}} x^2 - \partial_2 C_1(x^2, x^3). \quad (2.5.6)$$

Daraus folgt aber sofort

$$\partial_2 C_1(x^2, x^3) = 0 \Rightarrow C_1(x^2, x^3) = C_2(x^3). \quad (2.5.7)$$

Dabei haben wir verwendet, daß aus dem Verschwinden der partiellen Ableitung nach  $x^2$  folgt, daß die Funktion nicht von  $x^2$  abhängen kann. Es muß also  $C_1(x^2, x^3)$  eine Funktion allein von  $x^3$  sein, die wir mit  $C_2(x^3)$  bezeichnet haben. Wertet man schließlich noch die dritte Komponente von (2.5.2) aus, folgt auch noch, daß  $C_2(x^3) = C = \text{const}$  ist. Die Gravitationskraft besitzt also tatsächlich ein Potential, und zwar

$$\check{\Phi}_G(\vec{x}) = -\frac{\gamma m M}{\sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}} + C. \quad (2.5.8)$$

Die Konstante  $C$  hat keine weitere physikalische Bedeutung, denn bei der Bildung des Gradienten zur Berechnung der Kraft fällt sie stets weg. Es ist auch klar, daß das Potential ein Skalarfeld ist, denn wir können es unabhängig vom Koordinatensystem in der Form

$$\Phi_G(\vec{x}) = -\frac{\gamma m M}{|\vec{x}|} + C = -\frac{\gamma m M}{r} + C \quad (2.5.9)$$

schreiben.

Dieses Beispiel zeigt, daß es physikalisch wichtige Kraftfelder gibt, die tatsächlich ein Potential besitzen. Freilich ist die eben vorgeführte Methode i.a. zu kompliziert, um zu untersuchen, ob ein Kraftfeld ein **Potentialfeld** ist, d.h. ob es als (negativer) Gradient eines Skalarfeldes darstellbar ist. Im Rest dieses Abschnitts befassen wir uns mit der Frage nach Kriterien dafür, daß ein vorgegebenes Vektorfeld  $\vec{V}(\vec{x})$  ein solches **Potentialfeld** ist, ohne daß wir notwendig explizit das Potential ausrechnen müssen. Weiter wollen wir eine kompakte Formel für die Berechnung des Potentials finden, falls ein solches existiert.

Um eine **notwendige Bedingung** für die Existenz eines Potentials zu finden, müssen wir uns zunächst mit den **zweiten partiellen Ableitungen** eines vorgegebenen Skalarfeldes befassen. Es ist klar, daß diese wie bei Funktionen einer unabhängigen Veränderlichen einfach als Hintereinanderausführung zweier Ableitungen definieren lassen, also

$$\partial_j \partial_k \tilde{\Phi}(\vec{x}) = \frac{\partial}{\partial x^j} \left[ \frac{\partial}{\partial x^k} \tilde{\Phi}(\vec{x}) \right]. \quad (2.5.10)$$

Hier haben wir *zuerst* nach  $x^k$  und *dann* nach  $x^j$  differenziert. Die Frage ist nun, ob diese doppelte partielle Ableitung von dieser Reihenfolge abhängt. Die Antwort ist, daß partielle Ableitungen **kommutieren**, wenn die Funktion **zweimal stetig differenzierbar** ist.

Der Beweis erfordert ein wenig mehr Grundlagen aus der Analysis. Wir geben den Beweis hier, verschieben aber die Grundlagen in den Anhang. Der Einfachheit halber betrachten wir eine Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Wir nehmen an, sie sei zweimal partiell stetig differenzierbar. Wir wollen zeigen, daß dann

$$\partial_1 \partial_2 f(x^1, x^2) = \partial_2 \partial_1 f(x^1, x^2) \quad (2.5.11)$$

gilt. Betrachten wir nun für beliebige  $\Delta x^1$  und  $\Delta x^2$  die Funktion

$$F(x^1, x^2) = f(x^1 + \Delta x^1, x^2 + \Delta x^2) - f(x^1, x^2). \quad (2.5.12)$$

Weil diese Funktion voraussetzungsgemäß stetig differenzierbar ist, gibt es nach dem Mittelwertsatz der Differentialrechnung (vgl. Anhang A.6.3) eine Zahl  $\xi_1$  zwischen  $x^1$  und  $x^1 + \Delta x^1$ , so daß

$$F(x^1, x^2) = \Delta x^1 \partial_1 f(\xi_1, x^2) = [\partial_1 f(\xi_1, x^2 + \Delta x^2) - \partial_1 f(\xi_1, x^2)] \Delta x^1. \quad (2.5.13)$$

Da nach Voraussetzung auch die ersten Ableitungen stetig partiell differenzierbar sind, können wir nochmals den Mittelwertsatz der Differentialrechnung anwenden, d.h. es gibt eine Zahl  $\eta_1$  zwischen  $x^2$  und  $x^2 + \Delta x^2$ , so daß

$$F(x^1, x^2) = \partial_2 \partial_1 f(\xi_1, \eta_1) \Delta x^1 \Delta x^2. \quad (2.5.14)$$

Dieses Argument können wir aber auch genauso in umgekehrter Reihenfolge anwenden, d.h. den Mittelwertsatz zuerst auf die Variable  $x^2$  anwenden. Demnach existiert eine Zahl  $\eta_2$  zwischen  $x^2$  und  $x^2 + \Delta x^2$ , so daß

$$F(x^1, x^2) = \partial_2 F(x^1, \eta_2) \Delta x^2 = [\partial_2 f(x^1 + \Delta x^1, \eta_2) - \partial_2 f(x^1, \eta_2)] \Delta x^2. \quad (2.5.15)$$

Der Mittelwertsatz auf  $\partial_2 f$  angewandt liefert dann die Existenz einer Zahl  $\xi_2$  zwischen  $x^1$  und  $x^1 + \Delta x^1$ , so daß

$$F(x^1, x^2) = \partial_1 \partial_2 f(\xi_2, \eta_2) \Delta x^1 \Delta x^2. \quad (2.5.16)$$

Da dies für alle  $\Delta x^1$  und  $\Delta x^2$  gilt, folgt, daß

$$\partial_1 \partial_2 f(\xi_1, \eta_1) = \partial_2 \partial_1 f(\xi_2, \eta_2) \quad (2.5.17)$$

gilt. Lassen wir nun  $(\Delta x^1, \Delta x^2) \rightarrow (0, 0)$  streben, gilt  $(\xi_1, \eta_1) \rightarrow (x^1, x^2)$  und  $(\xi_2, \eta_2) \rightarrow (x^1, x^2)$ . Da voraussetzungsgemäß die zweiten partiellen Ableitungen stetig sind, folgt damit aus (2.5.17), daß tatsächlich (2.5.11).

Jetzt können wir ein **notwendiges Kriterium** für die Existenz eines Potentials für ein vorgegebenes Vektorfeld angeben. Sei also  $\vec{V}$  ein Vektorfeld, dessen Komponenten  $\bar{V}$  bzgl. einer kartesischen Basis stetig partiell differenzierbar sind. Angenommen, es sei ein Potentialfeld. Dann gilt

$$\bar{V}(\vec{x}) = -\bar{\nabla} \check{\Phi}(\vec{x}) \quad \text{bzw.} \quad V_j(\vec{x}) = -\partial_j \check{\Phi}(\vec{x}). \quad (2.5.18)$$

Nach Voraussetzung sind die Komponenten  $V_j$  stetig partiell differenzierbar, und damit  $\check{\Phi}$  zweimal stetig partiell differenzierbar. Nach dem eben bewiesenen Satz folgt

$$\partial_k \partial_j \check{\Phi} = \partial_j \partial_k \check{\Phi} \Rightarrow \partial_k V_j = \partial_j V_k. \quad (2.5.19)$$

Dies gilt für alle  $j, k \in \{1, 2, 3\}$ . Für  $k = j$  ist die Gleichung identisch erfüllt und trägt folglich keine Einschränkung an die Komponenten des Vektorfeldes für die Existenz eines Potentials bei. Für  $j \neq k$  impliziert aber (2.5.19)

$$\text{rot } \bar{V} = \bar{\nabla} \times \bar{V} = \begin{pmatrix} \partial_2 V^3 - \partial_3 V^2 \\ \partial_3 V^1 - \partial_1 V^3 \\ \partial_1 V^2 - \partial_2 V^1 \end{pmatrix} = \vec{0}. \quad (2.5.20)$$

Damit ein Vektorfeld ein Potential besitzt, muß also die Rotation in jedem regulären Punkt, also dort, wo es stetig partiell differenzierbar ist, verschwinden. Man rechnet leicht nach (*Übung!*), daß für das Gravitationskraftfeld (2.5.1) in der Tat außer im Ursprung des Koordinatensystems, wo das Feld eine Singularität besitzt,  $\text{rot } \vec{F}_G = \vec{0}$  gilt.

## 2.6 Wegintegrale und Potentialfelder

Als nächstes definieren wir **Wegintegrale** von Vektorfeldern. Es sei  $\vec{V}$  ein stetiges Vektorfeld und  $\mathcal{C} : \vec{x} : [t_1, t_2] \rightarrow E^3$  eine Kurve, die ganz im Inneren eines Bereiches verläuft, in dem  $\vec{V}$  definiert und stetig ist. Dann definieren wir das **Wegintegral entlang des Weges**  $\mathcal{C}$  durch

$$\int_{\mathcal{C}} d\vec{x} \cdot \vec{V}(\vec{x}) = \int_{t_1}^{t_2} dt \dot{\vec{x}}(t) \cdot \vec{V}[\vec{x}(t)]. \quad (2.6.1)$$

Es ist dabei klar, daß diese Definition nur von dem Weg selbst abhängt und nicht von der Parametrisierung, denn angenommen, es ist  $\vec{x}' : [\lambda_1, \lambda_2] \rightarrow E^3$  eine andere Parametrisierung desselben Weges, so gilt

$$\vec{x}'(\lambda) = \vec{x}[t(\lambda)] \rightarrow \frac{d}{d\lambda} \vec{x}'(\lambda) = \frac{d}{d\lambda} t(\lambda) \dot{\vec{x}}[t(\lambda)]. \quad (2.6.2)$$

Nach der Substitutionsformel für Integrale folgt daraus in der Tat, daß

$$\int_{\lambda_1}^{\lambda_2} d\lambda \frac{d}{d\lambda} \vec{x}'(\lambda) = \int_{t_1}^{t_2} dt \frac{d}{dt} t(\lambda) \frac{d}{d\lambda} t(\lambda) \dot{\vec{x}}(t) \cdot \vec{V}[\vec{x}(t)] = \int_{t_1}^{t_2} dt \dot{\vec{x}}(t) \cdot \vec{V}[\vec{x}(t)]. \quad (2.6.3)$$

Dabei haben wir im letzten Schritt den Satz von der Ableitung der Umkehrfunktion verwendet, demzufolge

$$\frac{d}{dt} t(\lambda) = \frac{1}{\frac{d}{d\lambda} t(\lambda)} \quad (2.6.4)$$

gilt. Außerdem ist offensichtlich das Wegintegral (2.6.1) ein **Skalar**, d.h. es kann mit Hilfe der Komponenten von  $\vec{V}$  bzgl. einer beliebigen kartesischen Basis berechnet werden:

$$\int_{\mathcal{C}} d\vec{x} \cdot \vec{V}(\vec{x}) = \int_{t_1}^{t_2} dt \dot{\vec{x}} \cdot \vec{V}[\vec{x}(t)] = \int_{t_1}^{t_2} dt \sum_{j=1}^3 \dot{x}_j V^j[\vec{x}(t)]. \quad (2.6.5)$$

Nehmen wir nun an  $\vec{V}$  sei ein Potentialfeld, d.h. es existiert ein Skalarfeld  $\Phi$ , so daß  $\vec{V} = -\vec{\nabla}\Phi$  ist. Berechnen wir nun das Wegintegral entlang eines beliebigen Weges, so folgt

$$\int_{\mathcal{C}} d\vec{x} \cdot \vec{V}(\vec{x}) = - \int_{\mathcal{C}} d\vec{x} \cdot \vec{\nabla}\Phi(\vec{x}) = - \int_{t_1}^{t_2} dt \dot{\vec{x}}(t) \cdot \vec{\nabla}\Phi[\vec{x}(t)]. \quad (2.6.6)$$

Nach der Kettenregel (2.2.12) folgt nun

$$\dot{\vec{x}}(t) \cdot \vec{\nabla}\Phi[\vec{x}(t)] = \dot{\vec{x}}(t) \cdot \check{\Phi}[\vec{x}(t)] = \frac{d}{dt} \check{\Phi}[\vec{x}(t)] = \frac{d}{dt} \Phi[\vec{x}(t)]. \quad (2.6.7)$$

Da  $\vec{V}$  voraussetzungsgemäß stetig ist, sind die partiellen Ableitungen von  $\check{\Phi}$  stetig und damit auch (2.6.7) als Funktion des Kurvenparameters  $t$ . Demnach können wir auf (2.6.6) den **Hauptsatz der Differential- und Integralrechnung** anwenden. Dies ergibt

$$\int_{\mathcal{C}} d\vec{x} \cdot \vec{V}(\vec{x}) = - \int_{t_1}^{t_2} dt \frac{d}{dt} \Phi[\vec{x}(t)] = - [\Phi[\vec{x}(t_2)] - \Phi[\vec{x}(t_1)]]. \quad (2.6.8)$$

Dies zeigt, daß für ein Potentialfeld das Wegintegral nur von **Anfangs- und Endpunkt** des Integrationsweges abhängt, nicht von dem Integrationsweg selbst. Dabei ist freilich darauf zu achten, daß dies nur für Wege gilt, die im Definitionsbereich des Vektorfeldes liegen, wo dieses stetig ist.

Aus (2.6.8) folgt auch, wie wir ggf. das Potential berechnen können. Dazu sei  $G \subseteq E^3$  ein beliebiges offenes Gebiet<sup>7</sup>. Existiert dann ein Punkt  $\vec{x}_0 \in G$ , so daß zu jedem  $\vec{x} \in G$  ein differenzierbarer Weg  $\mathcal{C}(\vec{x})$  mit Anfangspunkt  $\vec{x}_0$  und Endpunkt  $\vec{x}$  existiert, so ist gemäß (2.6.8)

$$\Phi(\vec{x}) = \check{\Phi}(\vec{x}) = - \int_{\mathcal{C}(\vec{x})} d\vec{x}' \cdot \vec{V}(\vec{x}') \quad (2.6.9)$$

ein Potential von  $\vec{V}$ .

Um dies zu zeigen, machen wir von der angenommenen **Wegunabhängigkeit** der Wegintegrale (2.6.9) Gebrauch. Dann ist nämlich die Funktion (2.6.9) unabhängig von der konkreten Wahl der Wege  $\mathcal{C}(\vec{x})$ . Da  $G$  offen ist, existiert um  $\vec{x} \in G$  eine Kugel  $K_\epsilon(\vec{x})$ , so daß  $K_\epsilon(\vec{x}) \subseteq G$ . Berechnen wir für das Potential in der Darstellung des Skalarfeldes  $\check{\Phi}$  bzgl. einer beliebigen Orthonormalbasis (2.6.9) die partielle Ableitung  $\partial_1 \check{\Phi}$ , indem wir die Definition als Limes anwenden. Dazu sei  $|\Delta x^1| < \epsilon$ . Dann liegt  $\vec{x} + \Delta x^1 \vec{e}_1 \in G$ , da konstruktionsgemäß die Kugel  $K_\epsilon(\vec{x})$  ganz in  $G$  liegt. Dann folgt aus (2.6.9)

$$\check{\Phi}(x^1 + \Delta x^1, x^2, x^3) - \check{\Phi}(x^1, x^2, x^3) = - \left[ \int_{\mathcal{C}(\vec{x} + \Delta x^1 \vec{e}_1)} d\vec{x}' \cdot \vec{V}(\vec{x}') - \int_{\mathcal{C}(\vec{x})} d\vec{x}' \cdot \vec{V}(\vec{x}') \right]. \quad (2.6.10)$$

Offenbar ist nun der Ausdruck in der eckigen Klammer das Wegintegral entlang eines in  $G$  gelegenen Weges, der von  $\vec{x}$  zu  $\vec{x} + \Delta x^1 \vec{e}_1$  verläuft. Da voraussetzungsgemäß das Wegintegral unabhängig von der konkreten

<sup>7</sup>Ein offenes Gebiet  $G \subseteq E^3$  ist dabei eine Teilmenge von  $E^3$ , für die zu jedem Punkt  $\vec{x} \in E^3$  eine Kugel  $K_\epsilon(\vec{x})$  mit Mittelpunkt bei  $\vec{x}$  und Radius  $\epsilon > 0$  existiert, so daß  $K_\epsilon(\vec{x}) \subseteq G$  gilt.

Form des Weges ist, können wir dieses Integral durch das Integral entlang der geraden Strecke (parallel zum Basisvektor  $\vec{e}_1$  ersetzen. Dieser Weg kann aber durch

$$s : \vec{x}'(t) = \vec{x} + t\Delta x^1 \vec{e}_1, \quad t \in [0, 1] \quad (2.6.11)$$

parametrisiert werden. Damit folgt

$$\begin{aligned} \tilde{\Phi}(x^1 + \Delta x^1, x^2, x^3) - \tilde{\Phi}(x^1, x^2, x^3) &= - \int_0^1 dt \Delta x^1 \vec{e}_1 \cdot \vec{V}(\vec{x} + t\Delta x^1 \vec{e}_1) \\ &= -\Delta x^1 \int_0^1 dt V^1(x^1 + t\Delta x^1, x^2, x^3). \end{aligned} \quad (2.6.12)$$

Dividieren wir diese Gleichung durch  $\Delta x^1$ , so folgt aus der Stetigkeit von  $V^1$  für den Limes  $\Delta x^1 \rightarrow 0$

$$\partial_1 \tilde{\Phi}(\vec{x}) = - \int_0^1 dt V^1(x^1, x^2, x^3) = -V^1(\vec{x}). \quad (2.6.13)$$

Genauso können wir auch für die beiden anderen partiellen Ableitungen zeigen, daß  $\partial_2 \tilde{\Phi}(\vec{x}) = -V^2(\vec{x})$  und  $\partial_3 \tilde{\Phi}(\vec{x}) = -V^3(\vec{x})$  gilt. Es ist also tatsächlich

$$\vec{V}(\vec{x}) = -\vec{\nabla} \tilde{\Phi}(\vec{x}) \Rightarrow \vec{V}(\vec{x}) = -\vec{\nabla} \Phi(\vec{x}). \quad (2.6.14)$$

Wir kommen auf die Frage, in welchen Fällen das Verschwinden der Rotation eines Vektorfeldes auch hinreichend für die Existenz eines Potentials für dieses Vektorfeldes in Abschnitt 2.8 noch zurück. Dazu benötigen wir aber Flächenintegrale und den Stokesschen Satz.

## 2.7 Flächenintegrale und der Stokessche Satz

In diesem Abschnitt führen wir **Flächenintegrale** über Vektorfeldern ein und besprechen den **Stokesschen Integralsatz**.

### 2.7.1 Orientierte Flächen im Raum

Analog zu den in Abschnitt 2.1 eingeführten Kurven im Raum führen wir nun **zweidimensionale Flächen** mit Hilfe ihrer **Parameterdarstellung** ein. Es sei dazu  $G \subseteq \mathbb{R}^2$  ein Bereich im  $\mathbb{R}^2$ . Dann beschreibt eine Abbildung  $\vec{x} : G \rightarrow E^3, (q_1, q_2) \mapsto \vec{x}(q_1, q_2)$  eine **Fläche**  $S$  im Raum. Wir gehen im folgenden davon aus, daß diese Abbildung stetig partiell differenzierbar ist. Man bezeichnet dann die Fläche als **glatt**.

Ein wichtiges *Beispiel* ist eine Kugelschale um den Ursprung eines rechtshändigen kartesischen Koordinatensystems, die wir durch **sphärische Koordinaten**  $\vartheta \in (0, \pi)$  und  $\varphi \in [0, 2\pi)$  parametrisieren. Wie wir der nebenstehenden Skizze entnehmen gilt für die Komponenten bzgl. kartesischer Koordinaten

$$\vec{x}(\vartheta, \varphi) = R \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}. \quad (2.7.1)$$

Wir bemerken, daß wir die durch  $\vartheta = 0$  bzw.  $\vartheta = \pi$  gegebenen Punkte auf der **Polarachse**, die wir hier wie konventionell üblich entlang der 3-Achse des kartesischen Koordinatensystems gewählt haben, ausnehmen müssen, weil dort offensichtlich der Winkel  $\varphi$  unbestimmt ist. Natürlich sind diese Punkte auf der Kugeloberfläche nicht irgendwie vor irgendeinem anderen Punkt ausgezeichnet. Es kommt also eine Art Singularität

allein aufgrund der Wahl der Parameter  $q_1 = \vartheta$  und  $q_2 = \varphi$ , die man auch als **generalisierte Koordinaten**<sup>8</sup> bezeichnet, zustande. Man spricht hier genauer von einer **Koordinatensingularität**, um deutlich zu machen, daß nicht notwendig eine irgendwie geartete tatsächlich Singularität bei der Fläche vorliegt. Solch eine tatsächlich Singularität kann z.B. ein „Knick“ oder eine „Spitze“ sein, wie z.B. bei der Kante bzw. der Ecke eines Würfels auftritt, wo die Parametrisierung nicht mehr differenzierbar ist oder die partiellen Ableitungen nicht stetig sind. Meistens muß man dann, wie beim Würfel, mehrere Parametrisierungen „aneinanderstückeln“. Man spricht dann von **stückweise glatten Flächen**.

Durch die Parametrisierung der Fläche durch generalisierte Koordinaten  $(q_1, q_2)$  sind durch die partiellen Ableitungen nun auch in jedem Punkt der zwei **Tangentenvektoren**

$$\vec{T}_j(q_1, q_2) = \frac{\partial}{\partial q_j} \vec{x}(q_1, q_2), \quad j \in \{1, 2\} \quad (2.7.2)$$

definiert. Die Parametrisierung der Fläche durch diese generalisierten Koordinaten heißt dann **regulär**, wenn diese beiden Tangentenvektoren **linear unabhängig** sind. Für unsere oben definierte Kugelschale gilt für die Komponenten dieser Tangentenvektoren (s. die obige Skizze)

$$\vec{T}_\vartheta(\varphi, \vartheta) = R \frac{\partial}{\partial \vartheta} \vec{x}(\vartheta, \varphi) = R \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix}, \quad \vec{T}_\varphi(\vartheta, \varphi) = R \frac{\partial}{\partial \varphi} \vec{x}(\vartheta, \varphi) = R \begin{pmatrix} -\sin \varphi \sin \vartheta \\ \cos \varphi \sin \vartheta \\ 0 \end{pmatrix}. \quad (2.7.3)$$

Diese Tangentenvektoren sind für die oben angegebenen Bereiche für  $\vartheta$  und  $\varphi$  linear unabhängig. Da für  $\vartheta = 0$  oder  $\vartheta = \pi$  der Tangentenvektor  $\vec{T}_\varphi = \vec{0}$  wird, ist dieser Punkt singulär. Dies reflektiert wieder die oben besprochene Koordinatensingularität der sphärischen Koordinaten.

In jedem regulären Punkt der Parametrisierung der Fläche können wir mit Hilfe des Kreuzproduktes den **Flächennormalenvektor** auf die Fläche an diesem Punkt definieren:

$$\vec{N}(q_1, q_2) = \vec{T}_1(q_1, q_2) \times \vec{T}_2(q_1, q_2). \quad (2.7.4)$$

Die Richtung des Normalenvektors hängt dabei von der Reihenfolge der generalisierten Koordinaten ab. Vertauschen wir  $q_1$  und  $q_2$ , kehrt sich wegen der Antisymmetrie des Vektorproduktes der Normalenvektor um. Durch die Wahl der Reihenfolge der beiden generalisierten Koordinaten prägen wir der Fläche also eine **Orientierung** auf und legen die Orientierung der Flächennormalenvektoren fest. Diese Wahl hängt von der Anwendung ab und ist mathematisch willkürlich. Wir kommen darauf im folgenden noch zurück.

Für die Kugelschale gilt

$$\vec{N} = \vec{T}_\vartheta \times \vec{T}_\varphi = R^2 \sin \vartheta \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix} \quad (2.7.5)$$

Durch die Wahl der Reihenfolge  $q_1 = \vartheta$  und  $q_2 = \varphi$  ist die Orientierung der Kugelfläche also so, daß die Normalenvektoren radial nach außen weisen<sup>9</sup>.

<sup>8</sup>Diese Art Koordinaten heißen „generalisiert“, weil sie keine Koordinaten eines Vektors bzgl. einer Basis bezeichnen. In dem hier betrachteten Beispiel haben die sphärischen Koordinaten immerhin noch eine konkrete geometrische Bedeutung als Winkel, die relativ zu einem kartesischen Koordinatensystem definiert sind. I.a. müssen generalisierte Koordinaten noch nicht einmal eine geometrische Bedeutung besitzen. Sie sind dann lediglich Parameter, die eindeutig die Punkte einer Fläche über die Funktion  $\vec{x}(q_1, q_2)$  „adressieren“. Im folgenden wird es auch darum gehen, Begriffe wie das Flächenintegral zu definieren, die konkrete geometrische Bedeutung besitzen und die Fläche unabhängig von der Wahl der generalisierten Koordinaten  $(q_1, q_2)$  charakterisieren.

<sup>9</sup>Dies entspricht einer Standardorientierung für geschlossene Flächen, die später im Zusammenhang mit den Volumenintegralen und dem Gaußschen Integralsatz noch wichtig werden wird: Betrachtet man ein Volumen  $V$  mit der dann notwendig geschlossenen Oberfläche  $S = \partial V$  als Rand, orientiert man diese Oberfläche so, daß die Normalenvektoren nach außen, d.h. von dem Volumen weg, zeigen.

### 2.7.2 Definition des Flächenintegrals

Wichtig für die Definition der Flächenintegrale ist die geometrische Bedeutung dieses Normalenvektors: Die beiden Tangentenvektoren (2.7.2) spannen die **Tangentialebene an die Fläche** an dem jeweiligen Punkt auf, und die „infinitesimalen Vektoren“  $dq_1 \vec{T}_1$  und  $dq_2 \vec{T}_2$  definieren ein infinitesimales Parallelogramm. Dann ist der **Flächenelementvektor**

$$d\vec{f} = dq_1 dq_2 \vec{T}_1 \times \vec{T}_2 = dq_1 dq_2 \vec{N} = d^2 q \vec{N} \quad (2.7.6)$$

ein auf der Tangentialebene senkrecht stehender Vektor, und seine Länge entspricht der Fläche des infinitesimalen Parallelogramms.

Das **Flächenintegral** eines Vektorfeldes  $\vec{V}$ , das in einem offenen Gebiet definiert und (stückweise) stetig ist, ist durch

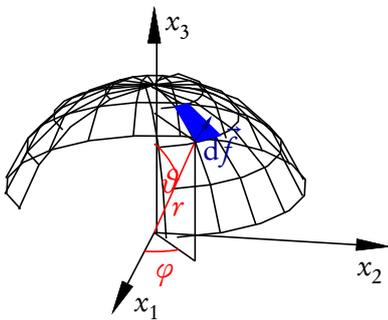
$$\int_S d\vec{f} \cdot \vec{V} = \int_G d^2 q \vec{N}(q_1, q_2) \cdot \vec{V}[\vec{x}(q_1, q_2)] \quad (2.7.7)$$

definiert. Dabei ist  $G \subseteq \mathbb{R}^2$  der Definitionsbereich der generalisierten Koordinaten, die die gesamte Fläche (ggf. bis auf einzelne Punkte) parametrisieren. Das Vorzeichen des Integrals hängt von der oben besprochenen Wahl der *Orientierung der Fläche* ab. Das Symbol  $S$  steht also stets für eine orientierte Fläche.

Natürlich können wir leicht auch den **Flächeninhalt** der Fläche ausrechnen:

$$A = \int_S |d\vec{f}| = \int_G d^2 q |\vec{N}(q_1, q_2)|. \quad (2.7.8)$$

Wir können z.B. leicht die Oberfläche einer Kugel ausrechnen. Dazu benötigen wir nur den oben berechneten Normalenvektor (2.7.5). Offenbar ist nämlich  $|\vec{N}| = R^2 \sin \vartheta$  (man beachte, daß wegen  $\vartheta \in (0, \pi)$  stets  $\sin \vartheta \geq 0$  ist). Damit folgt das aus der Elementargeometrie bekannte Resultat für die Kugeloberfläche:



$$\begin{aligned} A &= \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi R^2 \sin \vartheta = 2\pi R^2 \int_0^\pi d\vartheta \sin \vartheta \\ &= 2\pi R^2 [-\cos \vartheta]_0^\pi = 4\pi R^2. \end{aligned} \quad (2.7.9)$$

Ein wichtiges *Beispiel* für die Integration eines Vektorfeldes ist die Integration des Gravitationsfeldes einer Punktmasse über eine Kugel. Der Einfachheit halber lassen wir die Vorfaktoren weg und integrieren

$$\vec{V}(\vec{x}) = \frac{\vec{x}}{r^3}, \quad r = |\vec{x}|. \quad (2.7.10)$$

Für die Werte entlang der Kugeloberfläche erhalten wir

$$\vec{V}[\vec{x}(\vartheta, \varphi)] = \frac{1}{R^2} \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix} \quad (2.7.11)$$

und damit (*nachrechnen!*)

$$\int_S d\vec{f} \cdot \vec{V} = \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta \vec{N}(\vartheta, \varphi) \cdot \vec{V}[\vec{x}(\vartheta, \varphi)] = \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta = 4\pi. \quad (2.7.12)$$

### 2.7.3 Unabhängigkeit des Flächenintegrals von der Parametrisierung

In (2.7.6) haben wir das Flächenintegral auf die einfache Integration über die beiden generalisierten Koordinaten der Fläche zurückgeführt und die Vektoren durch ihre Komponenten bzgl. einer beliebigen rechtsorientierten kartesischen Basis ausgedrückt. Daß das Integral von der Wahl des rechtshändigen Orthonormalsystems unabhängig ist, rechnet man in analoger Weise wie oben bei den Wegintegralen nach und sei dem Leser zur *Übung* überlassen. Man muß nur beachten, daß die auftretenden Operationen wie Skalar- und Kreuzprodukte allesamt unabhängig von der Wahl der rechtshändigen Orthonormalbasis sind.

Nicht offensichtlich ist, daß das Flächenintegral auch unabhängig von der konkreten Wahl der Parametrisierung ist. Es sei also  $\vec{x}' : G' \rightarrow E^3, (q'_1, q'_2) \mapsto \vec{x}'(q'_1, q'_2)$  eine beliebige andere Parametrisierung derselben Fläche. Dies impliziert, daß wir die generalisierten Koordinaten  $(q_1, q_2)$  der ursprünglichen Parametrisierung als umkehrbar eindeutige Funktionen der neuen generalisierten Koordinaten  $(q'_1, q'_2)$  betrachten können.

Um zu zeigen, daß das Flächenintegral unabhängig von der Parametrisierung ist, untersuchen wir zuerst die Transformation der Flächennormalenvektoren zwischen den beiden Parametrisierungen und zeigen dann, daß die Definition des Flächenintegrals mit beiden Parametrisierungen tatsächlich dasselbe Resultat liefert. Für die Tangentenvektoren gilt

$$\vec{T}'_j = \frac{\partial \vec{x}'}{\partial q'_j} = \sum_{k=1}^2 \frac{\partial \vec{x}}{\partial q_k} \frac{\partial q_k}{\partial q'_j} = \sum_{k=1}^2 \vec{T}_k \frac{\partial q_k}{\partial q'_j}, \quad (2.7.13)$$

und daraus folgt für die Normalenvektoren

$$\vec{N}' = \vec{T}'_1 \times \vec{T}'_2 = \sum_{k,l=1}^2 \vec{T}_k \times \vec{T}_l \frac{\partial q_k}{\partial q'_1} \frac{\partial q_l}{\partial q'_2}. \quad (2.7.14)$$

In dieser Summe tragen nun nur die Terme bei, für die entweder  $k=1, l=2$  oder  $k=2, l=1$  ist. Schreiben wir also die Summe ausführlich hin, folgt

$$\vec{N}' = \vec{T}_1 \times \vec{T}_2 \frac{\partial q_1}{\partial q'_1} \frac{\partial q_2}{\partial q'_2} + \vec{T}_2 \times \vec{T}_1 \frac{\partial q_2}{\partial q'_1} \frac{\partial q_1}{\partial q'_2}. \quad (2.7.15)$$

Nun ist aber  $\vec{N} = \vec{T}_1 \times \vec{T}_2 = -\vec{T}_2 \times \vec{T}_1$ . Damit folgt

$$\vec{N}' = \left( \frac{\partial q_1}{\partial q'_1} \frac{\partial q_2}{\partial q'_2} - \frac{\partial q_2}{\partial q'_1} \frac{\partial q_1}{\partial q'_2} \right) \vec{N}. \quad (2.7.16)$$

Den Vorfaktor können wir nun wie folgt übersichtlicher schreiben. Dazu führen wir die **Jacobi-Matrix** der umkehrbar eindeutigen Transformation zwischen den beiden Sätzen von generalisierten Koordinaten  $(q_1, q_2) \leftrightarrow (q'_1, q'_2)$  ein

$$\frac{\partial(q_1, q_2)}{\partial(q'_1, q'_2)} = \begin{pmatrix} \partial q_1 / \partial q'_1 & \partial q_2 / \partial q'_1 \\ \partial q_1 / \partial q'_2 & \partial q_2 / \partial q'_2 \end{pmatrix} \quad (2.7.17)$$

Die Klammer in (2.7.16) ist offenbar die Determinante dieser Matrix, die **Jacobi-Determinante**:

$$\vec{N}' = \det \frac{\partial(q_1, q_2)}{\partial(q'_1, q'_2)} \vec{N}. \quad (2.7.18)$$

Es gilt also

$$\int_{G'} d^2 q' \vec{N}' \cdot \vec{V}[\vec{x}'(q'_1, q'_2)] = \int_{G'} d^2 q' \det \frac{\partial(q_1, q_2)}{\partial(q'_1, q'_2)} \vec{N} \cdot \vec{V}[\vec{x}'(q'_1, q'_2)]. \quad (2.7.19)$$

Da für jedes  $(q'_1, q'_2) \in G'$  stets  $\vec{x}'(q'_1, q'_2) = \vec{x}(q_1, q_2)$  für die entsprechenden generalisierten Koordinaten  $(q_1, q_2) \in G$  gilt, müssen wir zeigen, daß für beliebige Funktionen  $f : G \rightarrow \mathbb{R}$  gilt

$$\int_{G'} d^2 q' \det \frac{\partial(q_1, q_2)}{\partial(q'_1, q'_2)} f[q(q')] = \int_G d^2 q f(q) \quad (2.7.20)$$

gilt. Dies ist die Verallgemeinerung der Substitutionsformel einfacher Integrale auf Doppelintegrale. Diese Formel ist aber einfach zu verstehen. Dazu müssen wir nur bedenken, daß das Gebiet  $G' \subseteq \mathbb{R}^2$  umkehrbar eindeutig auf das Gebiet  $G \subseteq \mathbb{R}^2$  abgebildet wird, und zwar durch unsere Koordinatentransformation  $q : G' \rightarrow G, q' \mapsto q(q')$ . Durch die Koordinatenlinien  $q'_1 = \text{const}$  bzw.  $q'_2 = \text{const}$  wird das Gebiet  $G$  in infinitesimale Diagramme zerlegt, und diese besitzen die Flächen

$$dA = d^2 q' \left| \det \frac{\partial(q_1, q_2)}{\partial(q'_1, q'_2)} \right|. \quad (2.7.21)$$

Durch die Koordinatenlinien  $q_1$  und  $q_2$  wird das Gebiet  $G$  in infinitesimale Rechtecke zerlegt (wenn man  $(q_1, q_2)$  als kartesische Koordinaten in der Ebene auffaßt). Daraus folgt aber sofort, daß für stetige Funktionen in der Tat (2.7.20) gilt.

Vorausgesetzt, daß die Jacobi-Determinante der Transformation  $(q_1, q_2) \leftrightarrow (q'_1, q'_2)$  positiv ist, ist also durch die Unabhängigkeit des Flächenintegrals von der Wahl der Parametrisierung bewiesen, denn dann folgt aus (2.7.20)

$$\int_{G'} d^2 q' \vec{N}' \cdot \vec{V}[\vec{x}'(q'_1, q'_2)] = \int_G d^2 q \vec{N} \cdot \vec{V}[\vec{x}(q_1, q_2)] = \int_S d\vec{f} \cdot \vec{V}. \quad (2.7.22)$$

Wir müssen also für unsere Transformation voraussetzen, daß die Jacobi-Determinante positiv ist. Man bezeichnet solche Transformationen zwischen generalisierten Koordinaten einer Fläche als **orientierungserhaltende Transformationen**. Falls irgendeine gewählte Transformation eine negative Jacobi-Determinante ergibt, müssen wir nur die beiden neuen generalisierten umordnen.

#### 2.7.4 Koordinatenunabhängige Definition der Rotation

In Abschnitt 2.4 hatten wir die Rotation eines Vektorfeldes über seine kartesischen Komponenten definiert. Wir können nun die Rotation aber auch koordinatenunabhängig als Grenzwert eines Wegintegrals einführen. Dazu sei  $\vec{V} : G \rightarrow E^3$  ein auf einem (offenen) Gebiet  $G \subseteq E^3$  definiertes Vektorfeld mit partiell stetig differenzierbaren kartesischen Komponenten. Zu einem Punkt  $\vec{x}$  existiert dann eine Umgebung  $U$  (z.B. eine Kugel oder ein Quader), die vollständig in  $G$  liegt.

Wir betrachten nun eine Schar orientierter Flächen  $\Delta S \subseteq U$  mit dem dazu konsistent nach der Rechte-Hand-Regel orientierten Rand  $\partial \Delta S$  im Limes  $\Delta S \rightarrow 0$ , was bedeuten soll, daß die Fläche auf den Punkt  $\vec{x}$  zusammengezogen wird. Der Flächennormalenvektor sei im Limes  $\vec{N} \rightarrow \Delta A \vec{n}$ , wobei  $\Delta A$  der Flächeninhalt des infinitesimalen Flächenstücks sein soll und damit  $\vec{n}$  der **Flächennormaleneinheitsvektor**. Dann ist

$$\vec{n} \cdot \text{rot } \vec{V}(\vec{x}) = \lim_{\Delta S \rightarrow 0} \frac{1}{\Delta A} \int_{\partial \Delta S} d\vec{r} \cdot \vec{V}(\vec{r}). \quad (2.7.23)$$

Führen wir diese Prozedur nun mit drei Flächen mit Einheitsflächenvektoren  $\vec{n}_1, \vec{n}_2$  und  $\vec{n}_3$ , die ein rechtshändiges kartesischen Koordinatensystem (auch lokal im Punkt  $\vec{x}$ , was weiter unten im Zusammenhang mit **krummlinigen Orthogonalkoordinaten** noch wichtig wird) aufspannen, haben wir so die Rotation als Vektorfeld definiert.

Um zu zeigen, daß dies mit unserer obigen Definition in Abschnitt 2.4 übereinstimmt, wählen wir als Flächenelemente kleine Rechtecke parallel zu den Koordinatenachsen und berechnen das entsprechende Wegintegral

gemäß (2.7.23). Betrachten wir ein Quadrat parallel zur 12-Ebene, das wir im Gegenuhrzeigersinn durchlaufen. Eine bequeme Parametrisierung für die vier Seiten ist dann

$$\bar{r}_1(t) = \begin{pmatrix} x^1 + t \\ x^2 - \epsilon/2 \\ x^3 \end{pmatrix}, \quad \bar{r}_2(t) = \begin{pmatrix} x^1 + \epsilon/2 \\ x^2 + t \\ x^3 \end{pmatrix}, \quad \bar{r}_3(t) = \begin{pmatrix} x^1 - t \\ x^2 + \epsilon/2 \\ x^3 \end{pmatrix}, \quad \bar{r}_4(t) = \begin{pmatrix} x^1 - \epsilon/2 \\ x^2 - t \\ x^3 \end{pmatrix}, \quad (2.7.24)$$

wobei jeweils  $t \in (-\epsilon/2, \epsilon/2)$  durchläuft und  $\epsilon > 0$  so klein ist, daß das umschlossene Quadrat ganz in  $G$  liegt (s. nebenstehende Skizze). Dessen Einheitsnormalenvektor ist an jedem Punkt offenbar  $\vec{e}_3$ . Betrachten wir nun das Wegintegral entlang des ersten Wegstücks:

$$\begin{aligned} \int_{\mathcal{C}_1} d\vec{r} \cdot \vec{V}(\vec{r}) &= \int_{-\epsilon/2}^{\epsilon/2} dt \dot{\bar{r}}_1(t) \cdot \bar{V}[(x^1 + t, x^2 - \epsilon/2, x^3)] \\ &= \int_{-\epsilon/2}^{\epsilon/2} dt V^1(x^1 + t, x^2 - \epsilon/2, x^3) \\ &= V^1(\xi_1, x^2 - \epsilon/2, x^3)\epsilon. \end{aligned} \quad (2.7.25)$$

Dabei haben wir im letzten Schritt den Mittelwertsatz der Integralrechnung angewendet. Dabei ist  $\xi_1 \in (x^1 - \epsilon/2, x^1 + \epsilon/2)$ . Genauso folgt

$$\int_{\mathcal{C}_3} d\vec{r} \cdot \vec{V}(\vec{r}) = -V^1(\xi_1', x^2 + \epsilon/2, x^3)\epsilon. \quad (2.7.26)$$

Da voraussetzungsgemäß die Komponenten des Vektorfeldes  $\vec{V}$  stetig partiell differenzierbar sind, gibt es aufgrund des Zwischenwertsatzes der Differentialrechnung ein  $\xi_2 \in (x^2 - \epsilon/2, x^2 + \epsilon/2)$  mit

$$\begin{aligned} \int_{\mathcal{C}_1} d\vec{r} \cdot \vec{V}(\vec{r}) + \int_{\mathcal{C}_3} d\vec{r} \cdot \vec{V}(\vec{r}) &= -[V^1(\xi_1', x^2 + \epsilon/2, x^3) - V^1(\xi_1, x^2 - \epsilon/2, x^3)]\epsilon \\ &= -\epsilon^2 \partial_2 V^1(\xi_1'', \xi_2, x^3). \end{aligned} \quad (2.7.27)$$

Dabei ist  $\xi_1'' \in (x^1 - \epsilon/2, x^1 + \epsilon/2)$ . Lassen wir nun  $\Delta S \rightarrow 0$ , also  $\epsilon \rightarrow 0$  gehen folgt wegen  $\Delta A = \epsilon^2$

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left[ \int_{\mathcal{C}_1} d\vec{r} \cdot \vec{V}(\vec{r}) + \int_{\mathcal{C}_3} d\vec{r} \cdot \vec{V}(\vec{r}) \right] = -\partial_2 V^1(\bar{x}). \quad (2.7.28)$$

Analog zeigt man, daß

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left[ \int_{\mathcal{C}_2} d\vec{r} \cdot \vec{V}(\vec{r}) + \int_{\mathcal{C}_4} d\vec{r} \cdot \vec{V}(\vec{r}) \right] = +\partial_1 V^2(\bar{x}). \quad (2.7.29)$$

Gemäß der Definition (2.7.23) ist demnach

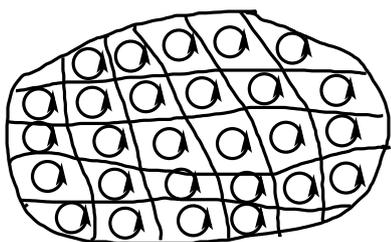
$$\vec{e}_3 \cdot \text{rot } \bar{V}(\bar{x}) = \partial_1 V^2(\bar{x}) - \partial_2 V^1(\bar{x}), \quad (2.7.30)$$

und das stimmt mit der Definition in (2.4.6) überein. Die übrigen Komponenten berechnen sich analog für Quadrate parallel zur 23- und 13-Ebene (*Übung!*).

## 2.7.5 Der Integralsatz von Stokes

Mit der Definition der Rotation über Wegintegrale im vorigen Abschnitt wird der **Integralsatz von Stokes** fast zu einer Selbstverständlichkeit. Für ein stetig partiell differenzierbares Vektorfeld gilt demnach für jede orientierte Fläche  $S$  mit gemäß der Rechte-Hand-Regel kompatibel mit der Flächenorientierung orientiertem Rand  $\partial S$

$$\int_S d\vec{f} \cdot \text{rot } \vec{V} = \int_{\partial S} d\vec{x} \cdot \vec{V}. \quad (2.7.31)$$



Um dies zu zeigen, muß man nur entsprechend der nebenstehenden Skizze die Fläche in viele kleine Flächenstücke unterteilen und auf jeder dieser Flächenstücke den Mittelwertsatz der Integralrechnung sowie die Definition der Rotation aus dem vorigen Abschnitt anwenden:

$$\int_{\Delta S_j} d\vec{f} \cdot \text{rot } \vec{V} = \Delta A_j \vec{n} \cdot \text{rot } \vec{V}(\vec{\xi}) = \int_{\partial \Delta S_j} d\vec{x} \cdot \vec{V}. \quad (2.7.32)$$

Bei der Summe über alle Flächenstücke heben sich die Wegintegrale über die inneren Linien weg, weil diese jeweils zweimal in entgegengesetztem Sinne durchlaufen werden, und es bleibt nur das Wegintegral über den Rand der Gesamtfläche übrig. Im Limes beliebig feiner Unterteilung der Flächenstücke ergibt sich aus (2.7.33) für die linke Seite wieder das Flächenintegral und die rechte Seite ist stets das Wegintegral über  $\partial S$ , womit der **Integralsatz von Stokes** bewiesen ist.

### 2.7.6 Der Greensche Satz in der Ebene

Wir können nun einen Integralsatz für **ebene Vektorfelder** als Spezialfall des Stokesschen Satzes herleiten. Wir wählen dazu die Ebene als 12-Ebene eines kartesischen Koordinatensystems im Raum. Dann ist für ein ebenes Vektorfeld

$$\vec{V}(\vec{x}) = \begin{pmatrix} V^1(x^1, x^2) \\ V^2(x^1, x^2) \\ 0 \end{pmatrix}. \quad (2.7.33)$$

Wir betrachten nun ein beliebiges offenes Gebiet  $S$  in der 12-Ebene mit im Gegenuhrzeigersinn orientierten Rand  $\partial S$ . Dieses Gebiet können wir aber genauso gut als Fläche im dreidimensionalen Raum ansehen. Die Flächennormaleneinheitsvektoren sind dann allesamt  $\vec{n} = \vec{e}_3$ .

Nun gilt gemäß (2.4.6)

$$\text{rot } \vec{V} = \begin{pmatrix} 0 \\ 0 \\ \partial_1 V^2 - \partial_2 V^1 \end{pmatrix}. \quad (2.7.34)$$

Dann spezialisiert sich der Stokessche Integralsatz auf

$$\int_S d\vec{f} \cdot \text{rot } \vec{V} = \int_S df(\partial_1 V^2 - \partial_2 V^1) = \int_{\partial S} d\vec{x} \cdot \vec{V}. \quad (2.7.35)$$

Das ist der **Integralsatz von Green in der Ebene**:

$$\int_S df(\partial_1 V^2 - \partial_2 V^1) = \int_{\partial S} d\vec{x} \cdot \vec{V}. \quad (2.7.36)$$

Dabei muß man nur die Konvention beachten, daß der Rand des ebenen Gebiets  $S$  so zu durchlaufen ist, daß in der Durchlaufrichtung betrachtet dieses Gebiet stets links liegt.

## 2.8 Das Poincaré-Lemma

Nun können wir die in Abschnitt 2.5 aufgeworfene Frage beantworten, unter welchen Umständen aus  $\text{rot } \vec{V} = 0$  folgt, daß  $\vec{V}$  ein Potentialfeld ist. In Abschnitt 2.6 haben wir gesehen, daß ein Potential existiert, wenn Wegintegrale in einem Gebiet unabhängig von der konkreten Form des Weges sind und nur von Anfangs- und Endpunkt der Wege abhängen. Sind nun  $\mathcal{C}_1$  und  $\mathcal{C}_2$ , die dieselben Punkte  $\vec{x}_1$  und  $\vec{x}_2$  zu Anfangs- und Endpunkt haben, so ist der Weg  $\mathcal{C} = \mathcal{C}_1 - \mathcal{C}_2$  geschlossen. Dabei bezeichnen wir mit  $-\mathcal{C}_2$  den in umgekehrter

Richtung durchlaufenen Weg  $\mathcal{C}_2$  und mit  $\mathcal{C}_1 - \mathcal{C}_2$  entsprechend den Weg, der zuerst entlang  $\mathcal{C}_1$  von  $\vec{x}_1$  zu  $\vec{x}_2$  und dann entlang  $-\mathcal{C}_2$  von  $\vec{x}_2$  zurück zu  $\vec{x}_1$  führt. Damit es also für ein stetig partiell differenzierbares Vektorfeld ein Potential gibt, ist es notwendig und hinreichend, daß für alle **geschlossenen Wege**

$$\int_{\mathcal{C}} d\vec{r} \cdot \vec{V} = 0 \quad (2.8.1)$$

gilt.

Ist nun also  $\text{rot } \vec{V} = 0$  in einem Gebiet, das so geartet ist, daß man zu jedem geschlossenen Weg  $\mathcal{C}$  eine Fläche  $S$  finden kann, so daß sein Rand  $\partial S = \mathcal{C}$  ist, folgt sofort aus dem Stokesschen Satz

$$0 = \int_S d\vec{f} \cdot \text{rot } \vec{V} = \int_{\partial S} d\vec{x} \cdot \vec{V} = \int_{\mathcal{C}} d\vec{x} \cdot \vec{V} = 0. \quad (2.8.2)$$

Neben der Bedingung, daß die Rotation verschwindet, muß also auch noch das Gebiet  $G$ , in dem  $\vec{V}$  stetig partiell differenzierbar sein muß die besagte Eigenschaft besitzen, daß jede geschlossene Kurve die Randkurve einer ganz in  $G$  gelegenen Fläche ist.

Diese Bedingung können wir auch so formulieren: Es muß möglich sein, eine jede geschlossene Kurve innerhalb von  $G$  stetig zu einem Punkt in  $G$  zu deformieren, denn bei dieser Deformation überstreicht die so definierte Schar von Kurven eine Fläche mit der ursprünglichen Kurve als Rand. Man nennt solche Gebiete **einfach zusammenhängend**.

I.a. wird ein Vektorfeld in einem **mehrfach zusammenhängenden Gebiet**, für das dort  $\text{rot } \vec{V} = 0$  gilt nur **lokal** ein Potential besitzen. Ist nämlich  $G$  offen, kann man um jeden Punkt eine ganz in  $G$  gelegene Kugelumgebung finden. Eine Kugel ist nun einfach zusammenhängend, und dort existiert dann auch ein Potential, und dieses ist dort gemäß Abschnitt 2.6 bis auf eine Konstante eindeutig bestimmt. I.a., wird es in mehrfach zusammenhängenden Gebieten nur lokale Potentiale geben, und diese sind i.a. nicht mehr eindeutig bestimmt. Ein schönes physikalisches Beispiel, das dies recht drastisch veranschaulicht, ist der **Potentialwirbel**. Das ist das Geschwindigkeitsfeld einer Strömung, das in kartesischen Koordinaten durch

$$\vec{V}(\vec{x}) = \frac{v_0}{(x^1)^2 + (x^2)^2} \begin{pmatrix} -x^2 \\ x^1 \\ 0 \end{pmatrix} \quad (2.8.3)$$

gegeben ist. Es ist offensichtlich in  $\vec{x} \in G = \mathbb{R}^3 \setminus \mathbb{R}(0, 0, 1)^T$  definiert (also überall außer entlang der 3-Achse) und stetig partiell differenzierbar. Außerdem gilt dort auch überall  $\text{rot } \vec{V} = 0$ , wie man mit Hilfe der Formel (2.4.6) sofort bestätigt (*Übung!*). Dieses Gebiet ist aber offensichtlich *nicht einfach zusammenhängend*, denn man kann jede Kurve, die die 3-Achse umschließt innerhalb von  $G$  nicht stetig zu einem Punkt zusammenziehen. Entsprechend schneidet jede Fläche mit einer solchen Kurve als Rand die 3-Achse und liegt folglich nicht ganz in  $G$ .

Betrachten wir nun als Kurve einen Kreis  $K_R$  um die 3-Achse in der 12-Ebene des kartesischen Koordinatensystems mit Radius  $R$ , den wir durch

$$\vec{x}(\varphi) = R \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \quad \varphi \in [-\pi, \pi] \quad (2.8.4)$$

parametrisieren. Es folgt

$$\frac{d}{d\varphi} \vec{x}(\varphi) = R \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \quad \vec{V}[\vec{x}(\varphi)] = \frac{v_0}{R} \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix} \quad (2.8.5)$$

und damit für das Wegintegral

$$\int_{K_R} d\vec{x} \cdot \vec{V}(\vec{x}) = \int_{-\pi}^{\pi} d\varphi \dot{\vec{x}}(\varphi) \cdot \vec{V}[\vec{x}(\varphi)] = v_0 R \int_{-\pi}^{\pi} d\varphi = 2\pi v_0 R. \quad (2.8.6)$$

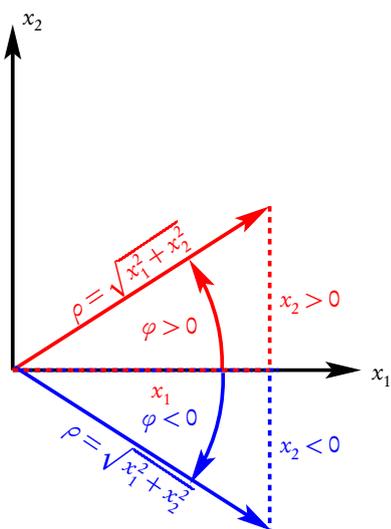
Obwohl also überall in  $G$  stets  $\text{rot } \vec{V} = 0$  gilt, verschwindet das Wegintegral entlang der geschlossenen Kreislinie nicht, und damit existiert kein in  $G$  eindeutig bestimmtes Potential.

Wir können andererseits aber in jedem **einfach zusammenhängenden Teilgebiet** ein solches Potential finden. Das größtmögliche solche Teilgebiet erhalten wir offenbar, wenn wir eine Halbebene mit der 3-Achse als Rand ausnehmen. Wir wählen willkürlich den Teil der 12-Ebene mit  $x^1 \leq 0$ , die wir mit  $H_{12}^{\leq}$  bezeichnen wollen. Es gibt nun in  $\tilde{G} = E^3 \setminus H_{12}^{x^1 \leq 0}$  keine geschlossenen stetigen Kurven, die die 3-Achse umlaufen, da diese unweigerlich unsere ausgeschlossene Halbebene schneiden müssten, und folglich ist dieses Gebiet einfach zusammenhängend, und entsprechend verschwinden die Wegintegrale von  $\vec{V}$  entlang geschlossener Wege in  $\tilde{G}$ . Beliebige Wege, die einen festgehaltenen Punkt  $\vec{x}_0 \in \tilde{G}$  mit einem beliebigen anderen Punkt  $\vec{x} \in \tilde{G}$  verbinden, ergeben dann denselben Wert für das Wegintegral, und dieses Integral definiert dann ein Potential für  $\vec{V}$  in dem auf  $\tilde{G}$  eingeschränkten Gebiet.

Um es zu berechnen können wir irgendeinen Punkt  $\vec{x}_0$  und irgendeinen Weg, der ihn innerhalb von  $\tilde{G}$  mit  $\vec{x}$  verbindet, verwenden, um das Potential zu bestimmen. Wir wählen als Anfangspunkt  $\vec{x}_0 = (1, 0, 0)^T$ . Um einen bequemen Verbindungsweg definieren zu können, definieren wir jetzt **Zylinderkoordinaten**. Offenbar können wir jeden Punkt in  $\tilde{G}$  umkehrbar eindeutig durch die Parameter  $\rho > 0$ ,  $\varphi \in (-\pi, \pi)$  und  $z \in \mathbb{R}$  festlegen, indem wir

$$\vec{x} = \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \\ z \end{pmatrix} \quad (2.8.7)$$

setzen. Man nennt  $\rho$ ,  $\varphi$  und  $z$  wieder generalisierte Koordinaten. Da die Fläche  $\rho = \text{const}$  ein Zylinder ist, heißen diese Koordinaten **Zylinderkoordinaten**. Wir kommen auf solche verallgemeinerten Koordinaten weiter unten noch ausführlicher zu sprechen.



Um nun  $\vec{x}_0$  mit einem Weg mit  $\vec{x}(\rho, \varphi, z)$  zu verbinden, der ein möglichst einfach zu berechnendes Wegintegral ergibt, beachten wir, daß gerade Linien entlang der 1-Achse und in  $z$ -Richtung keinen Beitrag zum Wegintegral liefern, denn beide Wegarten besitzen zu  $\vec{V}$  überall senkrechte Tangentenvektoren. Weiter wissen wir von unserer obigen Berechnung des Wegintegrals entlang einer Kreislinie in einer Ebene parallel zur 12-Ebene, daß Integrale entlang von solchen Kreislinienstücken trivial werden. Wir wählen also den folgenden Weg: Von  $\vec{x}_0 = (1, 0, 0)$  gehen wir zunächst entlang der 1-Achse zum Punkt  $(\rho, 0, 0)$ , dann entlang des Kreislinienstücks

$$\vec{x}(\varphi') = \begin{pmatrix} \rho \cos \varphi' \\ \rho \sin \varphi' \\ 0 \end{pmatrix}, \quad \varphi' \in [0, \varphi]. \quad (2.8.8)$$

Schließlich verbinden wir noch  $(\rho \cos \varphi, \rho \sin \varphi, 0)$  mit  $\vec{x} = (\rho \cos \varphi, \rho \sin \varphi, z)$  durch die entsprechende zur 3-Achse parallele Strecke. Wie oben gesagt, tragen das erste und das letzte gerade Teilstück dieses Weges  $\mathcal{C}$  nichts zum Wegintegral von  $\vec{V}$  bei, und wir finden schließlich

$$\Phi(\vec{x}) = - \int_{\mathcal{C}} d\vec{r} \cdot \vec{V}(\vec{r}) = - \int_0^{\varphi} d\varphi' \dot{\vec{x}}(\varphi') \cdot \vec{V}[\vec{x}(\varphi')] = -v_0 \varphi. \quad (2.8.9)$$

Den Winkel  $\varphi$  erhalten wir aufgrund der nebenstehenden Skizze aus

$$\varphi = \text{sign } x^2 \arccos\left(\frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}}\right). \quad (2.8.10)$$

Man muß dazu bemerken, daß dann  $\varphi \in (-\pi, \pi)$  zu liegen kommt, wenn man Punkte auf der negativen 1-Achse ausschließt. Das entspricht genau unserer Wahl des Gebiets  $\tilde{G}$ . Für Punkte auf der positiven 1-Achse ist eindeutig  $\varphi = 0$ , so daß die Unbestimmtheit der Signum-Funktion

$$\text{sign } x = \begin{cases} -1 & \text{für } x < 0, \\ +1 & \text{für } x > 0 \end{cases} \quad (2.8.11)$$

hierbei keine Rolle spielt.

Wir können nun leicht verifizieren, daß (2.8.9) mit (2.8.10) für  $\varphi$  tatsächlich ein Potential des Potentialwirbelfeldes in  $\tilde{G}$  ist. Es gilt nämlich nach der Kettenregel

$$\partial_1 \Phi = -\text{sign } x^2 \partial_1 \left( \frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}} \right) \arccos' \left( \frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}} \right). \quad (2.8.12)$$

Die Ableitung des arccos ist aber

$$\arccos' x = -\frac{1}{\sqrt{1-x^2}} \quad (2.8.13)$$

und damit

$$\begin{aligned} \partial_1 \Phi &= v_0 \text{sign } x^2 \frac{\sqrt{(x^1)^2 + (x^2)^2} - (x^1)^2 / \sqrt{(x^1)^2 + (x^2)^2}}{(x^1)^2 + (x^2)^2} \frac{1}{\sqrt{1 - (x^1)^2 / ((x^1)^2 + (x^2)^2)}} \\ &= v_0 \frac{x^2}{(x^1)^2 + (x^2)^2}, \end{aligned} \quad (2.8.14)$$

$$\partial_2 \Phi = -v_0 \text{sign } x^2 \frac{x^1 x^2}{((x^1)^2 + (x^2)^2)^{3/2}} \frac{1}{\sqrt{1 - (x^1)^2 / [(x^1)^2 + (x^2)^2]}} = -v_0 \frac{x^1}{(x^1)^2 + (x^2)^2},$$

$$\partial_3 \Phi = 0.$$

Daraus ergibt sich in der Tat

$$\vec{V} = -\vec{\nabla} \Phi = \frac{v_0}{(x^1)^2 + (x^2)^2} \begin{pmatrix} -x^2 \\ x^1 \\ 0 \end{pmatrix}. \quad (2.8.15)$$

Wählt man irgendeine andere Halbebene, findet man ein anderes Potential für das entsprechend geänderte einfach zusammenhängende Gebiet  $\tilde{G}'$ , das sich von dem soeben berechneten in Bereichen in  $\tilde{G} \cap \tilde{G}'$  um eine Konstante unterscheidet und entsprechend andere Wertebereiche für  $\varphi$  verwendet, so daß das Potential bei der entsprechenden Halbebene einen Sprung um  $\pm 2\pi v_0$  aufweist. Unser Potential ist entlang der negativen  $x^1$ -Achse  $-v_0 \pi$ , wenn man sich von positiven  $x^2$ -Werten her nähert und  $+v_0 \pi$ , wenn man sich von negativen  $x^2$ -Werten her nähert.

### 2.8.1 Der Energieerhaltungssatz

Wir wollen nun noch zeigen, daß für die Bewegung eines Teilchens in einem Kraftfeld, das ein Potential besitzt, der **Energieerhaltungssatz** gilt. Sei als  $\vec{F}(\vec{x})$  ein Kraftfeld, das ein Potential besitzt, d.h. für das ein Potential existiert:

$$\vec{F}(\vec{x}) = -\vec{\nabla} V(\vec{x}). \quad (2.8.16)$$

Die **Newtonsche Bewegungsgleichung** für die Bewegung eines Massenpunktes in diesem Kraftfeld lautet

$$m\ddot{\vec{x}} = \vec{F}(\vec{x}) = -\vec{\nabla}V(\vec{x}). \quad (2.8.17)$$

Dies ist eine **Differentialgleichung**, d.h. die unbekannte Funktion  $\vec{x}(t)$  kommt mitsamt ihren (zweiten) Ableitung in der Gleichung vor, und wir suchen Lösungen für diese Gleichung. Ein großer Teil der theoretischen Physik widmet sich daher der Lösung solcher Differentialgleichungen. Wie wir später noch ausführlicher besprechen werden, benötigt man außer der Differentialgleichung (2.8.17) noch **Anfangsbedingungen**, um die Gleichung eindeutig zu lösen. Da hier eine **Differentialgleichung 2. Ordnung** vorliegt, benötigen wir **Ort und Geschwindigkeit** zu einem Anfangszeitpunkt  $t_0$ , um die Bahnkurve durch Lösung von (2.8.17) zu bestimmen. Die **klassische Mechanik von Punktteilchen** beschäftigt sich letztlich also im wesentlichen mit der Lösung solcher **Anfangswertprobleme**. Wir denken uns also diese Anfangsbedingungen vorgegeben:

$$\vec{x}(t_0) = \vec{x}_0, \quad \dot{\vec{x}}(t_0) = \vec{v}(t_0) = \vec{v}_0. \quad (2.8.18)$$

I.a. ist es aber schwierig überhaupt Lösungen oder sogar alle möglichen Lösungen zu finden. Man kann aber für spezielle Fälle oft sehr allgemeine Eigenschaften über die Lösungen herleiten, ohne diese explizit zu kennen. Das sind u.a. die **Erhaltungssätze**. Wir wollen nun zeigen, daß ein solcher Erhaltungssatz für den Fall, daß die Kräfte nur vom Ort des Teilchens abhängen und ein Potential besitzen, d.h. (2.8.16) gilt, der **Energieerhaltungssatz** gilt. Dazu bemerken wir, daß wir das Potential gemäß (2.6.9) als Wegintegral entlang eines beliebigen Weges, der einen fest vorgegebenen Punkt  $\vec{x}_0$  mit dem Argument  $\vec{x}$  des Potentials verbindet. Die Form des Weges ist dabei gleichgültig, da für Potentialfelder das Wegintegral nur von Anfangs- und Endpunkt und nicht von der Form des Verbindungsweges abhängen. Wir können also insbesondere auch die Lösung der Bewegungsgleichung (2.8.18), die die Anfangsbedingungen (2.8.18) erfüllt, verwenden. Wir multiplizieren also die Bewegungsgleichung mit  $\dot{\vec{x}}$  und integrieren von  $t_0$  bis  $t$ . Für die linke Seite der Gleichung ergibt sich dabei

$$m \int_{t_0}^t dt' \dot{\vec{x}}(t') \cdot \ddot{\vec{x}}(t') = \frac{m}{2} \int_{t_0}^t dt' \frac{d}{dt'} \dot{\vec{x}}^2(t'), \quad (2.8.19)$$

denn gemäß der Produktregel der Differentiation gilt

$$\frac{d}{dt} \dot{\vec{x}}^2(t) = 2\dot{\vec{x}} \cdot \ddot{\vec{x}}. \quad (2.8.20)$$

Es gilt also

$$m \int_{t_0}^t dt' \dot{\vec{x}}(t') \cdot \ddot{\vec{x}}(t') = \frac{m}{2} [\dot{\vec{x}}^2(t) - \dot{\vec{x}}^2(t_0)] = \frac{m}{2} [\vec{v}^2(t) - \vec{v}_0^2]. \quad (2.8.21)$$

Für die rechte Seite von (2.2.17) ergibt diese Prozedur

$$-\int_{t_0}^t dt' \dot{\vec{x}}(t') \cdot \vec{\nabla}V[\vec{x}(t')] = -\int_{t_0}^t dt' \frac{d}{dt'} V[\vec{x}(t')] = -V[\vec{x}(t)] + V[\vec{x}(t_0)] = -V[\vec{x}(t)] + V(\vec{x}_0). \quad (2.8.22)$$

Gemäß der Bewegungsgleichung (2.2.17) sind aber die Resultate von (2.8.21) und (2.8.22) gleich, und es folgt

$$\frac{m}{2} [\vec{v}^2(t) - \vec{v}_0^2] = -V[\vec{x}(t)] + V(\vec{x}_0). \quad (2.8.23)$$

Sortieren nun die Gleichung so um, daß alle Größen zur Zeit  $t_0$  auf eine und alle Größen zur Zeit  $t$  auf die andere Seite kommen, folgt schließlich der **Energieerhaltungssatz**

$$\frac{m}{2} \vec{v}^2(t) + V[\vec{x}(t)] = \frac{m}{2} \vec{v}_0^2 + V(\vec{x}_0). \quad (2.8.24)$$

Die **Gesamtenergie**

$$E = \frac{m}{2} \vec{v}^2 + V(\vec{x}) = E_0 = \frac{m}{2} \vec{v}_0^2 + V(\vec{x}_0) \quad (2.8.25)$$

ist also für die Lösungen der Bewegungsgleichung **zeitlich konstant**. Die Herleitung hat gezeigt, daß die Existenz eines zeitunabhängigen Kraftpotentials dafür entscheidend ist. Man nennt solche Kräfte daher auch **konservativ**, weil die Energie erhalten ist.

## 2.9 Volumenintegrale, Divergenz und Gaußscher Integralsatz

In diesem Abschnitt definieren wir Volumenintegrale über Skalarfelder, geben die koordinatenunabhängige Definition der Divergenz an und beweisen den Gaußschen Integralsatz.

### 2.9.1 Definition des Volumenintegrals

Das **Volumenintegral** über ein Skalarfeld ist die Integration über einen dreidimensionalen Bereich  $V \subseteq E^3$ . Man kann dieses Gebiet wieder durch eine beliebige Parametrisierung mit drei generalisierten Koordinaten  $(q_1, q_2, q_3)$  beschreiben. Das Volumenelement ist dabei unabhängig von der Parametrisierung durch die entsprechende **Jacobi-Determinante** gegeben

$$d^3x = dq_1 dq_2 dq_3 \det \frac{\partial(x^1, x^2, x^3)}{\partial(q_1, q_2, q_3)} = d^3q (\vec{T}_1 \times \vec{T}_2) \cdot \vec{T}_3. \quad (2.9.1)$$

Dabei wählen wir die Reihenfolge der generalisierten Koordinaten wieder so, daß die Jacobideterminante positiv ist, d.h. die drei Koordinatenlinien in jedem regulären Punkt der Parametrisierung liefern drei linear unabhängige Tangentenvektoren

$$\vec{T}_j = \frac{\partial \vec{x}}{\partial q_j}, \quad (2.9.2)$$

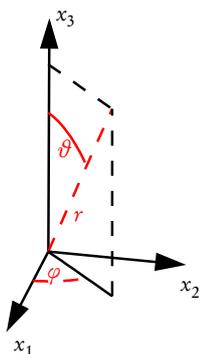
die relativ zum rechtshändigen kartesischen Koordinatensystem gleichorientiert sind also auch eine rechtehändige Basis sind. Analog wie bei den Flächenintegralen zeigt man (*Übung!*), daß dann das Volumenintegral

$$\int_V d^3x \Phi(\vec{x}) = \int_{\tilde{G}} d^3q \det \frac{\partial(x^1, x^2, x^3)}{\partial(q_1, q_2, q_3)} \tilde{\Phi}[\vec{x}(q)]. \quad (2.9.3)$$

Dabei ist  $\tilde{G} \subseteq \mathbb{R}^3$  ein Gebiet, das den Parameterbereich für die generalisierten Koordinaten  $q = (q_1, q_2, q_3)$  angibt.

Wichtige Beispiele für solche Bereiche sind Kugel und Zylinder. Die Standardparametrisierung für eine Kugel mit Radius  $R$  um den Ursprung sind die **Kugelkoordinaten**

$$\vec{x}(r, \vartheta, \varphi) = r \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}, \quad \vartheta \in (0, \pi), \quad r \in (0, R], \quad \varphi \in [0, 2\pi). \quad (2.9.4)$$



Wir bemerken gleich, daß hier die gesamte 3-Achse, die **Polarachse**, die den Werten  $\vartheta = 0$  bzw.  $\vartheta = \pi$  entspricht, ausgenommen wurde. Der Grund dafür ist, daß Kugelkoordinaten dort offenbar eine **Koordinatensingularität** aufweisen, wie wir schon oben bei der Parametrisierung der Kugelschale bemerkt haben.

Die Jacobi-Determinante bestimmen wir am einfachsten über die drei Tangentenvektoren der Koordinatenlinien

$$\begin{aligned}\bar{T}_r &= \partial_r \bar{x} = \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}, \\ \bar{T}_\vartheta &= \partial_\vartheta \bar{x} = r \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix}, \\ \bar{T}_\varphi &= \partial_\varphi \bar{x} = r \sin \vartheta \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}.\end{aligned}\tag{2.9.5}$$

Wir bemerken bereits hier, daß die drei hier auftretenden Spaltenvektoren in jedem Punkt eine **rechtshändige Orthonormalbasis**

$$\begin{aligned}\bar{e}_r &= \frac{1}{|\bar{T}_r|} \bar{T}_r = \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}, \\ \bar{e}_\vartheta &= \frac{1}{|\bar{T}_\vartheta|} \bar{T}_\vartheta = \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix}, \\ \bar{e}_\varphi &= \frac{1}{|\bar{T}_\varphi|} \bar{T}_\varphi = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}.\end{aligned}\tag{2.9.6}$$

bilden, d.h. es gilt (*nachrechnen!*)

$$\bar{e}_\varphi = \bar{e}_r \times \bar{e}_\vartheta, \quad \bar{e}_r \cdot \bar{e}_\vartheta = 0.\tag{2.9.7}$$

Dies macht die **Kugelkoordinaten** zu sog. **krummlinigen Orthogonalkoordinaten**, die wir in Abschnitt 2.10 genauer besprechen werden. Da

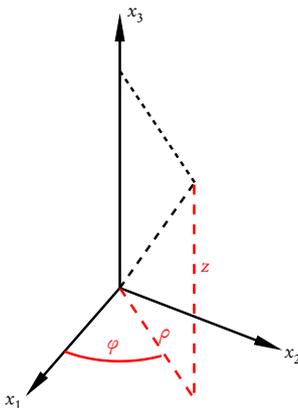
$$\det \frac{\partial(x^1, x^2, x^3)}{\partial(q_1, q_2, q_3)} = (\bar{T}_r \times \bar{T}_\vartheta) \cdot \bar{T}_\varphi = r^2 \sin \vartheta.\tag{2.9.8}$$

Wie wir sehen, verschwindet die Jacobi-Determinante für  $r = 0$  bzw.  $\vartheta = 0$  oder  $\vartheta = \pi$ , also entlang der 3-Achse. Dies zeigt, daß die Kugelkoordinaten dort tatsächlich eine Koordinatensingularität besitzen.

Ein weiteres Beispiel sind **Zylinderkoordinaten**  $(\rho, \varphi, z)$ . Für einen zylinderförmigen Bereich mit Radius  $R$  und Höhe  $h$  um den Ursprung des kartesischen Koordinatensystems ist die Standardparametrisierung

$$\bar{x}(\rho, \varphi, z) = \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \\ z \end{pmatrix}, \quad \rho \in (0, R], \quad \varphi \in [0, 2\pi), \quad z \in [-h/2, h/2].\tag{2.9.9}$$

Die Zylinderkoordinaten sind entlang der 3-Achse (die **Zylinderachse**) singulär. Auch sie sind krummlinige Orthogonalkoordinaten. Die Tangentenvektoren der



Koordinatenlinien sind

$$\begin{aligned}\vec{T}_\rho &= \vec{e}_\rho = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \\ \vec{T}_\varphi &= \rho \vec{e}_\varphi = \rho \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \\ \vec{T}_z &= \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\end{aligned}\tag{2.9.10}$$

und die Jacobi-Determinante ist

$$\det \frac{\partial(x^1, x^2, x^3)}{\partial(\rho, \varphi, z)} = \rho.\tag{2.9.11}$$

Als *Beispiele* für Volumenintegrale berechnen wir das Volumen einer Kugel und eines Zylinders mit Hilfe der beiden eben definierten Kugel- bzw. Zylinderkoordinaten. Für die Kugel gilt

$$\begin{aligned}V^{\text{Kugel}} &= \int_{K_R} d^3x = \int_0^R dr \int_0^\pi d\vartheta \int_0^{2\pi} \varphi r^2 \sin \vartheta \\ &= 2\pi \int_0^R dr \int_0^\pi d\vartheta r^2 \sin \vartheta \\ &= 2\pi \int_0^R dr r^2 [-\sin \vartheta]_0^\pi \\ &= 4\pi \int_0^R dr r^2 = \frac{4\pi}{3} R^3\end{aligned}\tag{2.9.12}$$

und für den Zylinder

$$V^{\text{Zylinder}} = \int_{Z(R,b)} d^3x = \int_0^R d\rho \int_0^{2\pi} d\varphi \int_{-b/2}^{b/2} dz \rho = b \int_0^R d\rho \int_0^{2\pi} d\varphi \rho = 2\pi b \int_0^R d\rho \rho = \pi R^2 b.\tag{2.9.13}$$

## 2.9.2 Die koordinatenunabhängige Definition der Divergenz

Ähnlich wie wir in Abschnitt 2.7.4 die Rotation durch den Limes eines Wegintegrals definiert haben, geben wir nun eine Definition der Divergenz über ein **Flächenintegral** an. Dazu definieren wir die Orientierung der Randfläche  $\partial B$  eines dreidimensionalen Bereiches  $B$  so, daß die Flächennormalenvektoren nach außen, also **von dem betrachteten Volumen weg** weisen. Für die Kugel ist die Randfläche in Kugelkoordinaten einfach durch  $r = R = \text{const}$  gegeben, und wir gelangen wieder zur Parametrisierung (2.7.1) der entsprechenden Kugelschale. Der Normalenvektor ist dabei stets  $\vec{e}_r$ , und weist entsprechend unserer Definition in die korrekte Richtung weg von der Kugel.

Beim Zylinder zerfällt die Randfläche in drei Teile, nämlich den Zylinder-Mantel  $\rho = R$  (generalisierte Koordinaten  $(\varphi, z)$ ; Flächennormalenvektor  $\vec{N} = R\vec{e}_\rho$  weist in die korrekte Richtung) sowie die Deckfläche ( $z = b/2$  mit generalisierten Koordinaten  $(\rho, \varphi)$ ,  $\vec{N} = \rho\vec{e}_3$ ) und die Bodenfläche, bei der wir die Orientierung umkehren müssen, damit der Flächennormalenvektor  $\vec{N} = -\rho\vec{e}_3$  ist, also aus dem Zylinder hinausweist.

Betrachten wir nun ein partiell stetig differenzierbares Vektorfeld, können wir die Divergenz durch

$$\text{div } \vec{V}(\vec{x}) = \lim_{\Delta V \rightarrow 0} \frac{1}{\text{vol}(\Delta V)} \int_{\partial \Delta V} d\vec{f} \cdot \vec{V}\tag{2.9.14}$$

definieren. Dabei ist  $\Delta V$  ein Volumenbereich mit dem Volumen  $\text{vol}(\Delta V)$ , der ganz im Definitionsbereich des Vektorfeldes liegt, und der Limes ist so zu verstehen, daß dieses Volumenelement auf den Punkt  $\vec{x}$  zusammengezogen wird.

Um zu zeigen, daß in kartesischen Koordinaten diese Definition mit der oben in (2.4.3) gegebenen Definition übereinstimmt, wählen wir  $\Delta V$  als Würfel der Kantenlänge  $\epsilon$  mit dem Mittelpunkt in  $\vec{x}$  und führen eine ähnliche Rechnung wie in Abschnitt 2.7.4 durch. Betrachten wir als Beispiel den Beitrag der beiden zur 12-Ebene parallelen Würfelflächen, die bei  $x^3 - \epsilon/2$  bzw.  $x^3 + \epsilon/2$  liegen. Mit dem Mittelwertsatz ergibt sich für diese beiden Flächenintegrale

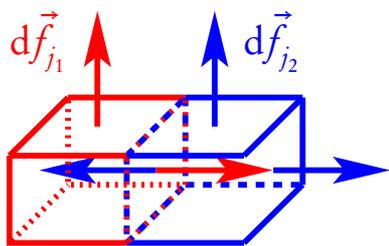
$$\epsilon^2 [V^3(\xi_1, \xi_2, x^3 + \epsilon/2) - V^3(\xi_1, \xi_2, x^3 - \epsilon/2)] = \epsilon^3 \partial_3 V^3(\xi_1'', \xi_2'', x^3). \quad (2.9.15)$$

Dies durch das Volumen  $\epsilon^3$  des Würfels dividiert und den Limes  $\epsilon \rightarrow 0$  gebildet liefert den Beitrag  $\partial_3 V^3(\vec{x})$ . Entsprechend erhält man für die beiden anderen Seitenflächen des Würfels die beiden anderen Beiträge, so daß (2.9.14) tatsächlich das gleiche Resultat wie (2.4.3) ergibt.

### 2.9.3 Der Gaußsche Integralsatz

Mit der obigen Definition der Divergenz ist der **Gaußsche Integralsatz** eine selbstverständliche Folgerung. Ist  $\vec{V}$  ein stetig partiell differenzierbares Vektorfeld,  $B$  ein Volumenbereich und  $\partial B$  sein entsprechend der Orientierungsvorschrift, daß die Flächennormalenvektoren aus dem Bereich  $B$  herauszeigen, so gilt

$$\int_B d^3x \operatorname{div} \vec{V}(\vec{x}) = \int_{\partial B} d\vec{f} \cdot \vec{V}(\vec{x}). \quad (2.9.16)$$



Um dies zu beweisen, müssen wir nur den Bereich  $B$  in viele kleine Teilvolumenelemente  $\Delta V^j$  zerlegen. Nach dem Mittelwertsatz der Integralrechnung ergibt sich dann analog wie beim Satz von Stokes

$$\int_{\Delta B_j} d^3x \operatorname{div} \vec{V} = \operatorname{vol}(\Delta B_j) \operatorname{div} \vec{V}(\vec{\xi}) \simeq \int_{\partial \Delta B_j} d\vec{f} \cdot \vec{V}(\vec{x}). \quad (2.9.17)$$

Dabei wird die letztere Beziehung im Limes  $\Delta B_j \rightarrow 0$ , also bei immer feinerer Unterteilung des Volumens  $B$  in Teilvolumina exakt. Summiert man nun die linke Seite dieser Gleichung auf, erhält man stets das Volumenintegral auf der linken Seite von (2.9.16), und auf der rechten Seite heben sich die Beiträge von den inneren Oberflächenteilen  $\partial \Delta B_j$  weg, da diese in der Summe stets zweimal mit unterschiedlicher Orientierung auftauchen (s. Skizze).

### 2.9.4 Die Greenschen Integralsätze im Raum

Die **Greenschen Integralsätze** im Raum sind Spezialfälle des Gaußschen Integralsatzes. Der **1. Greensche Integralsatz** ergibt sich, aus dem Gaußschen Integralsatz, indem wir als Vektorfeld

$$\vec{V}(\vec{x}) = \Phi_1(\vec{x}) \vec{\nabla} \Phi_2(\vec{x}) \quad (2.9.18)$$

setzen. Wir berechnen als erstes die Divergenz. Dazu verwenden wir am einfachsten die Darstellung in kartesischen Komponenten. Zunächst ist

$$V^j(\vec{x}) = \check{\Phi}_1(\vec{x}) \partial_j \check{\Phi}_2(\vec{x}). \quad (2.9.19)$$

Aus der Produktregel für Ableitungen folgt daraus

$$\operatorname{div} \vec{V}(\vec{x}) = \sum_{j=1}^3 \partial_j V^j(\vec{x}) = \sum_{j=1}^3 [\partial_j \check{\Phi}_1(\vec{x}) \partial_j \check{\Phi}_2(\vec{x}) - \check{\Phi}_1(\vec{x}) \partial_j^2 \check{\Phi}_2(\vec{x})]. \quad (2.9.20)$$

Dies können wir wieder in koordinatenunabhängiger Form als

$$\operatorname{div} \vec{V}(\vec{x}) = \vec{\nabla} \cdot \vec{V}(\vec{x}) = [\vec{\nabla} \Phi_1(\vec{x})] \cdot [\vec{\nabla} \Phi_2(\vec{x})] + \Phi_1(\vec{x}) \vec{\nabla}^2 \Phi_2(\vec{x}). \quad (2.9.21)$$

Der Differentialoperator  $\vec{\nabla}^2$  kommt so häufig in der Feldtheorie vor, daß man dafür ein eigenes Symbol  $\Delta$ , den **Laplace-Operator**, einführt. Wirkt es auf ein zweimal stetig partiell differenzierbares Skalarfeld, erhält man wieder ein Skalarfeld, und zwar

$$\Delta \Phi(\vec{x}) = \vec{\nabla} \cdot \vec{\nabla} \Phi(\vec{x}) = \operatorname{div} \operatorname{grad} \Phi(\vec{x}) = \vec{\nabla}^2 \Phi(\vec{x}). \quad (2.9.22)$$

In kartesischen Komponenten gilt ausgeschrieben

$$\Delta \Phi(\vec{x}) = \partial_1^2 \Phi(\vec{x}) + \partial_2^2 \Phi(\vec{x}) + \partial_3^2 \Phi(\vec{x}). \quad (2.9.23)$$

Setzen wir also dieses Vektorfeld in den Gaußschen Integralsatz ein, folgt der **1. Greensche Integralsatz**

$$\int_B d^3x \left\{ [\vec{\nabla} \Phi_1(\vec{x})] \cdot [\vec{\nabla} \Phi_2(\vec{x})] + \Phi_1(\vec{x}) \Delta \Phi_2(\vec{x}) \right\} = \int_{\partial B} d\vec{f} \cdot \Phi_1(\vec{x}) \vec{\nabla} \Phi_2(\vec{x}). \quad (2.9.24)$$

Der **2. Greensche Integralsatz** folgt daraus, indem wir in dieser Gleichung  $\Phi_1$  und  $\Phi_2$  vertauschen und das Resultat von der vorigen Gleichung abziehen. Dabei verschwindet im Volumenintegral der erste Term, weil dieser symmetrisch unter dieser Vertauschung ist:

$$\int_B d^3x [\Phi_1(\vec{x}) \Delta \Phi_2(\vec{x}) - \Phi_2(\vec{x}) \Delta \Phi_1(\vec{x})] = \int_{\partial B} d\vec{f} \cdot [\Phi_1(\vec{x}) \vec{\nabla} \Phi_2(\vec{x}) - \Phi_2(\vec{x}) \vec{\nabla} \Phi_1(\vec{x})]. \quad (2.9.25)$$

Dieser Satz wird oft in der **Potentialtheorie** gebraucht. Wir werden ihn in Abschnitt 2.12 verwenden, um die einfachste Aufgabe der Potentialtheorie zu lösen.

## 2.10 Krummlinige Orthogonalkoordinaten

In diesem Abschnitt betrachten wir **krummlinige Orthogonalkoordinaten**. Wir hatten weiter oben schon die **Kugel-** und **Zylinderkoordinaten** als die beiden wichtigsten Beispiele kennengelernt. Unser Ziel ist es, die Differentialoperatoren grad, div, rot und  $\Delta$  in allgemeinen krummlinigen Orthogonalkoordinaten auszudrücken.

### 2.10.1 Definition krummliniger Orthogonalkoordinaten

Wir betrachten irgendwelche generalisierten Koordinaten  $q = (q_1, q_2, q_3) \in G \subseteq \mathbb{R}^3$ , die durch eine Abbildung  $\vec{x} : G \rightarrow E^3$  einen bestimmten Raumbereich (ggf. auch den ganzen Raum) umkehrbar eindeutig parametrisieren. Hält man zwei dieser drei Koordinaten fest, erhält man die **Koordinatenlinien**, die in jedem Punkt von  $G$  drei Tangentenvektoren

$$\vec{T}_j(q) = \frac{\partial \vec{x}(q)}{\partial q_j} \quad (2.10.1)$$

definieren.

Diese generalisierten Koordinaten heißen **krummlinige Orthogonalkoordinaten**, falls in jedem Punkt diese Tangentenvektoren aufeinander senkrecht stehen, d.h. es gibt Funktionen  $g_j : G \rightarrow \mathbb{R}$  mit  $g_j(q) > 0$ , so daß

$$\vec{T}_j(q) \cdot \vec{T}_k(q) = g_j^2(q) \delta_{jk} \quad (2.10.2)$$

gilt. Wir definieren weiter die **Einheitsvektoren**

$$\vec{e}'_j(q) = \frac{1}{g_j(q)} \vec{T}_j(q). \quad (2.10.3)$$

Sie bilden in jedem Punkt offenbar **Orthonormalsysteme**. Im Gegensatz zu den kartesischen Koordinatensystemen<sup>10</sup> hängen sie aber i.a. vom Ort ab.

Im folgenden nehmen wir an, daß diese Einheitsvektoren in jedem Punkt ein **rechtshändiges Dreibein** sind, d.h. daß überall

$$\vec{e}'_1 \times \vec{e}'_2 = \vec{e}'_3, \quad \vec{e}'_2 \times \vec{e}'_3 = \vec{e}'_1, \quad \vec{e}'_3 \times \vec{e}'_1 = \vec{e}'_2 \quad (2.10.4)$$

gilt.

Für die Jacobi-Determinante der Transformation von kartesischen Koordinaten  $\bar{x}$  und den verallgemeinerten krummlinigen Orthogonalkoordinaten  $q$  folgt daraus sofort, daß

$$\det \frac{\partial(x^1, x^2, x^3)}{\partial(q_1, q_2, q_3)} = \vec{T}_1 \cdot (\vec{T}_2 \times \vec{T}_3) = g_1 g_2 g_3 \vec{e}'_1 \cdot (\vec{e}'_2 \times \vec{e}'_3) = g_1 g_2 g_3. \quad (2.10.5)$$

Im folgenden beschreiben wir skalare Felder und Vektorfelder bzgl. dieser krummlinigen Orthogonalkoordinaten:

$$\Phi(\vec{x}) = \check{\Phi}(q), \quad \vec{V}(\vec{x}) = \sum_{j=1}^3 \vec{e}_j(q) V^j(q). \quad (2.10.6)$$

Wir wollen nun die Differentialoperatoren grad, rot, div und  $\Delta$  durch diese Koordinaten ausdrücken. Dazu führen wir noch ein **rechtshändiges kartesisches Orthonormalsystem**  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  ein.

**Beispiele** für orthogonale krummlinige Koordinaten sind die schon oben betrachteten Kugel- und Zylinderkoordinaten, die wir im folgenden nochmals über die kartesischen Komponenten des Ortsvektors definieren

**Kugelkoordinaten**  $(r, \vartheta, \varphi)$ :

$$\bar{x} = r \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}. \quad (2.10.7)$$

Durch Ableiten nach den generalisierten Koordinaten ergeben sich zunächst die Tangentenvektoren an die Koordinatenlinien und die  $g_j$  zu

$$\begin{aligned} \bar{T}_r &= \frac{\partial \bar{x}}{\partial r} = \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}, & g_r &= |\bar{T}_r| = 1, \\ \bar{T}_\vartheta &= \frac{\partial \bar{x}}{\partial \vartheta} = r \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix}, & g_\vartheta &= |\bar{T}_\vartheta| = r, \\ \bar{T}_\varphi &= \frac{\partial \bar{x}}{\partial \varphi} = r \sin \vartheta \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, & g_\varphi &= |\bar{T}_\varphi| = r \sin \vartheta. \end{aligned} \quad (2.10.8)$$

Man prüft leicht nach, daß diese Vektoren tatsächlich orthogonal zueinander sind und in der angegebenen

<sup>10</sup>Es ist klar, daß die kartesischen Komponenten des Ortsvektors bzgl. einer beliebigen kartesischen Basis auch Spezialfälle krummliniger Orthogonalkoordinaten sind.

Reihenfolge ein rechtshändiges Basissystem bilden. Die normierten Basisvektoren sind

$$\begin{aligned}\vec{e}'_r &= \frac{\bar{T}_1}{g_1} = \begin{pmatrix} \cos \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix}, \\ \vec{e}'_\vartheta &= \frac{\bar{T}_2}{g_2} = \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix}, \\ \vec{e}'_\varphi &= \frac{\bar{T}_3}{g_3} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}.\end{aligned}\tag{2.10.9}$$

Die Jacobideterminante lautet gemäß (2.10.7) (*nachrechnen!*)

$$\det \left[ \frac{\partial(x, y, z)}{\partial(r, \vartheta, \varphi)} \right] = r^2 \sin \vartheta.\tag{2.10.10}$$

**Zylinderkoordinaten**  $(r, \varphi, z)$ :

$$\bar{x} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}.\tag{2.10.11}$$

Die Tangentenvektoren an die Koordinatenlinien und Normierungsfaktoren sind

$$\begin{aligned}\bar{T}_r &= \frac{\partial \bar{x}}{\partial r} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, & g_r &= |\bar{T}_r| = 1, \\ \bar{T}_\varphi &= \frac{\partial \bar{x}}{\partial \varphi} = r \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, & g_\varphi &= |\bar{T}_\varphi| = r, \\ \bar{T}_z &= \frac{\partial \bar{x}}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, & g_z &= |\bar{T}_z| = 1,\end{aligned}\tag{2.10.12}$$

und die normierten Basisvektoren lauten

$$\begin{aligned}\vec{e}'_r &= \frac{\bar{T}_1}{g_1} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \\ \vec{e}'_\varphi &= \frac{\bar{T}_2}{g_2} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \\ \vec{e}'_z &= \frac{\bar{T}_3}{g_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.\end{aligned}\tag{2.10.13}$$

Wir nun fahren zunächst mit allgemeinen Betrachtungen zu krummlinigen Orthonormalsystemen fort. Unser Ziel ist es, die Differentialoperatoren grad, div und rot mit Hilfe der generalisierten Koordinaten  $q_i$  auszudrücken, wobei alle Vektoren nach der dazugehörigen Orthonormalbasis  $\vec{e}'_i$  entwickelt werden.

Es sei im folgenden  $U(\vec{x})$  ein Skalar- und  $\vec{A}(\vec{x})$  ein Vektorfeld. Das Vektorfeld können wir dabei sowohl nach den kartesischen Basisvektoren  $\vec{e}_j$  als auch nach den zu den  $q_i$  gehörigen Orthonormalvektoren  $\vec{e}'_i$  entwickeln, d.h.

$$\vec{A} = \sum_{j=1}^3 A_j \vec{e}_j = \sum_{i=1}^3 A'_i \vec{e}'_i. \quad (2.10.14)$$

Wegen der Orthonormierung der  $\vec{e}_j$  und der  $\vec{e}'_i$  gilt

$$\begin{aligned} A_j &= \vec{e}_j \cdot \vec{A} = \sum_{i=1}^3 \vec{e}_j \cdot \vec{e}'_i A'_i, \\ A'_i &= \vec{e}'_i \cdot \vec{A} = \sum_{j=1}^3 \vec{e}'_i \cdot \vec{e}_j A_j. \end{aligned} \quad (2.10.15)$$

Bezeichnen wir die Skalarprodukte mit

$$C_{ji} = \vec{e}_j \cdot \vec{e}'_i, \quad (2.10.16)$$

entsteht die **orthogonale Transformationsmatrix**

$$\hat{C} = \begin{pmatrix} \frac{1}{g_1} \frac{\partial x}{\partial q_1} & \frac{1}{g_2} \frac{\partial x}{\partial q_2} & \frac{1}{g_3} \frac{\partial x}{\partial q_3} \\ \frac{1}{g_1} \frac{\partial y}{\partial q_1} & \frac{1}{g_2} \frac{\partial y}{\partial q_2} & \frac{1}{g_3} \frac{\partial y}{\partial q_3} \\ \frac{1}{g_1} \frac{\partial z}{\partial q_1} & \frac{1}{g_2} \frac{\partial z}{\partial q_2} & \frac{1}{g_3} \frac{\partial z}{\partial q_3} \end{pmatrix}. \quad (2.10.17)$$

Da wir voraussetzungsgemäß die generalisierten Koordinaten so anordnen, daß die entsprechenden Dreibeine an die Koordinatenlinien in allen Punkten rechtshändige Orthonormalbasen ergeben, gilt noch

$$\det \hat{C} = 1. \quad (2.10.18)$$

In der  $j$ -ten Spalte dieser Matrix steht also gerade der Spaltenvektor  $\vec{e}'_j$ . Mit dieser Matrix können wir (2.10.15) in der Form

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \hat{C} \begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix}, \quad \begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \hat{C}^T \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \quad (2.10.19)$$

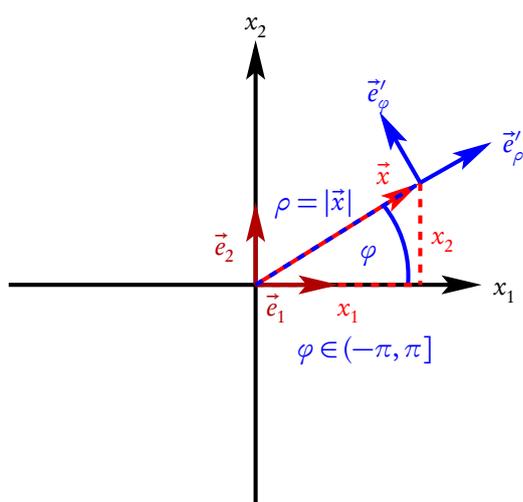
schreiben. Wendet man diese Gleichungen nacheinander an und bedenkt, daß sie für *beliebige* Vektoren  $\vec{A}$  gelten, folgt

$$\hat{C} \hat{C}^T = \hat{C}^T \hat{C} = \mathbb{1}_{3 \times 3} \Rightarrow \hat{C}^{-1} = \hat{C}^T, \quad (2.10.20)$$

d.h. die Transformationsmatrix ist eine **Orthogonalmatrix**. Da nach Voraussetzung  $\det C = +1$  ist, handelt es sich um eine **Drehung**. Das ist auch anschaulich klar: In jedem Raumpunkt geht das Dreibein  $\vec{e}'_i$  durch eine bestimmte Drehung aus dem kartesischen Dreibein  $\vec{e}_j$  hervor.

### 2.10.2 Polarkoordinaten in der Ebene

Natürlich kann man krummlinige Orthogonalkoordinaten auch in der Ebene einführen. Hier sollen die Einheitsvektoren  $\vec{e}'_j$  ( $j \in \{1, 2\}$ ) so orientiert sein, daß man durch Drehung von  $\vec{e}'_1$  **im Gegenuhrzeigersinn** um den Drehwinkel  $\pi/2$  auf  $\vec{e}'_2$  kommt. Alternativ kann man die Polarkoordinaten auch in der 12-Ebene eines rechtshändigen kartesischen Koordinatensystems im Raum auffassen, dann gilt für die entsprechenden Einheitsvektoren der positiv orientierte krummlinige Koordinaten in der Ebene stets  $\vec{e}'_1 \times \vec{e}'_2 = \vec{e}_3 = \vec{e}'_3$ . Man nennt solche krummlinigen Koordinaten im Raum auch „verallgemeinerte Zylinderkoordinaten“, da  $\vec{e}'_1$  und  $\vec{e}'_2$  für alle Komponenten  $x^3$  des Ortsvektors  $\vec{x}$  gleich sind und  $\vec{e}'_3$  ist ein konstanter (also ortsunabhängiger) Einheitsvektor.



Das wichtigste Beispiel für ebene krummlinige Orthogonalkoordinaten sind die **Polarkoordinaten**. Man verwendet zur Charakterisierung des Ortsvektors  $\vec{x}$  zunächst ein im obigen Sinne positiv orientiertes Orthonormalsystem  $(\vec{e}_1, \vec{e}_2)$  (s. nebenstehende Skizze) und definiert dann als krummlinige Koordinaten  $\rho = |\vec{x}| = \sqrt{(x^1)^2 + (x^2)^2}$  und den Winkel  $\varphi \in (-\pi, \pi]$  zwischen  $\vec{e}_1$  und  $\vec{x}$ , wobei man Winkel in der unteren Halbebene negativ zählt. Aus der Skizze lesen wir dann unmittelbar ab, daß

$$\vec{x} = \rho(\cos \varphi \vec{e}_1 + \sin \varphi \vec{e}_2) \quad (2.10.21)$$

ist. Die Tangentenvektoren an die Koordinatenlinien sind (*nachrechnen!*)

$$\begin{aligned} \vec{T}_\rho &= \frac{\partial \vec{x}}{\partial \rho} = \cos \varphi \vec{e}_1 + \sin \varphi \vec{e}_2, & g_\rho &= |\vec{T}_\rho| = 1, \\ \vec{T}_\varphi &= \frac{\partial \vec{x}}{\partial \varphi} = \rho(-\sin \varphi \vec{e}_1 + \cos \varphi \vec{e}_2), & g_\varphi &= |\vec{T}_\varphi| = \rho. \end{aligned} \quad (2.10.22)$$

Wir sehen, daß die Polarkoordinaten im Ursprung singular sind, weil dort  $T_\varphi = 0$  ist. Die entsprechenden orthogonalen Einheitsvektoren sind demnach

$$\begin{aligned} \vec{e}'_\rho &= \frac{1}{g_\rho} \vec{T}_\rho = \cos \varphi \vec{e}_1 + \sin \varphi \vec{e}_2, \\ \vec{e}'_\varphi &= \frac{1}{g_\varphi} \vec{T}_\varphi = -\sin \varphi \vec{e}_1 + \cos \varphi \vec{e}_2. \end{aligned} \quad (2.10.23)$$

Etwas schwieriger ist die umgekehrte Transformation von kartesischen Komponenten  $\bar{x} = (x^1, x^2)^T$  zu den Polarkoordinaten  $(\rho, \varphi)$ . Für  $\rho$  gilt, wie oben angegeben

$$\rho = |\vec{x}| = \sqrt{(x^1)^2 + (x^2)^2}. \quad (2.10.24)$$

Für den Winkel  $\varphi$ , der nur für  $\vec{x} \neq \vec{0}$  wohldefiniert ist, müssen wir simultan

$$\cos \varphi = \frac{x^1}{\rho}, \quad \sin \varphi = \frac{x^2}{\rho} \quad (2.10.25)$$

erfüllen. Anschaulich ist klar, daß die simultane Lösung dieser Gleichungen eindeutig ist, wenn man zusätzlich als Wertebereich für den Polarwinkel  $\varphi \in (-\pi, \pi]$  wählt. Die erste Gleichung ist offenbar nicht eindeutig zu lösen, denn aus ihr folgt, daß es zwei Lösungen

$$\varphi_\pm = \pm \arccos\left(\frac{x^1}{\rho}\right) \quad (2.10.26)$$

gibt. Dabei verstehen wir unter dem arccos den üblicherweise auch auf Taschenrechnern oder in den gängigen Programmiersprachen realisierten **Hauptwert**. Für alle  $\xi \in [-1, 1]$  ist  $\alpha = \arccos \xi \in [0, \pi]$ . Für diesen Wertebereich gilt nun aber  $\sin \alpha \geq 0$ . Wegen  $\sin^2 \alpha + \cos^2 \alpha = 1$  folgt daraus

$$\sin \alpha = +\sqrt{1 - \cos^2 \alpha} = \sqrt{1 - \xi^2}. \quad (2.10.27)$$

Setzen wir also (2.10.26) ein, erhalten wir für beide Vorzeichen

$$\sin(\varphi_{\pm}) = \pm \sin \left[ \arccos \left( \frac{x^1}{\rho} \right) \right] = \pm \sqrt{1 - \frac{(x^1)^2}{\rho^2}} = \pm \sqrt{\frac{\rho^2 - (x^1)^2}{\rho^2}} = \pm \sqrt{\frac{(x^2)^2}{\rho^2}} = \pm \frac{|x^2|}{\rho}. \quad (2.10.28)$$

Damit also auch die zweite Gleichung in (2.10.25) erfüllt ist, müssen wir für  $x^2 > 0$  in (2.10.26) das obere und für  $x^2 < 0$  das untere Vorzeichen wählen. Für  $x^2 = 0$  wählen wir definitionsgemäß stets das obere Vorzeichen, denn dann ist  $x^1/\rho = x^1/|x^1|$ , und man erhält dann für  $x^1 > 0$  stets  $\arccos(x^1/\rho) = \arccos 1 = 0$  und für  $x^1 < 0$  stets  $\arccos(x^1/\rho) = \arccos(-1) = \pi$ . Endgültig gilt also

$$\varphi = \begin{cases} \arccos(x^1/\rho) & \text{für } x^2 \geq 0, \\ -\arccos(x^1/\rho) & \text{für } x^2 < 0. \end{cases} \quad (2.10.29)$$

In der Literatur findet man oft auch Formeln, die  $\varphi$  mit Hilfe des arctan bestimmen. Das hat den Vorteil, daß man zur Bestimmung von  $\varphi$  nicht erst  $\rho$  berechnen muß. Allerdings benötigt man mehr Fallunterscheidungen bzgl. der Lage von  $\vec{x}$  in den unterschiedlichen Quadranten des kartesischen Koordinatensystems.

Dazu bemerken wir, daß für  $\xi \in \mathbb{R}$  definitionsgemäß  $\arctan \xi \in (-\pi/2, \pi/2)$  gilt. Aus der obigen Skizze lesen wir ab, daß für Ortsvektoren in der rechten offenen Halbebene, also  $x^1 > 0$  stets

$$\varphi = \arctan \left( \frac{x^2}{x^1} \right) \quad \text{für } x^1 > 0 \quad (2.10.30)$$

gilt. Falls  $x^1 < 0$  und  $x^2 > 0$ , ist  $\arctan(x^2/x^1) \in (-\pi/2, 0]$  und definitionsgemäß  $\varphi \in (\pi/2, \pi]$ . In diesem Fall müssen wir also

$$\varphi = \pi + \arctan \left( \frac{x^2}{x^1} \right) \quad \text{für } x^1 < 0, \quad x^2 \geq 0 \quad (2.10.31)$$

setzen. Für  $x^1 < 0$  und  $x^2 \leq 0$  ist  $\arctan(x^2/x^1) \in [0, \pi/2)$  und  $\varphi \in (-\pi, -\pi/2)$ . Demnach ist dann

$$\varphi = -\pi + \arctan \left( \frac{x^2}{x^1} \right) \quad \text{für } x^1 < 0, \quad x^2 \leq 0. \quad (2.10.32)$$

Bleibt schließlich der Fall  $x^1 = 0$ . Dann ist  $\varphi = \pi/2$  für  $x^2 > 0$  und  $\varphi = -\pi/2$  für  $x^2 < 0$  zu setzen.

### 2.10.3 Die Differentialoperatoren grad, div, rot und $\Delta$

Mit diesen Vorbereitungen können wir nun die Differentialoperatoren mittels der krummlinigen Orthonormalkoordinaten ausdrücken. Beginnen wir mit dem **Gradienten des Skalarfeldes**. Gemäß (2.10.19) gilt nach der Kettenregel unmittelbar

$$\begin{pmatrix} (\text{grad } U)_1 \\ (\text{grad } U)_2 \\ (\text{grad } U)_3 \end{pmatrix} = C^T \begin{pmatrix} \partial U / \partial x \\ \partial U / \partial y \\ \partial U / \partial z \end{pmatrix} = \begin{pmatrix} \frac{1}{g_1} \frac{\partial U}{\partial q_1} \\ \frac{1}{g_2} \frac{\partial U}{\partial q_2} \\ \frac{1}{g_3} \frac{\partial U}{\partial q_3} \end{pmatrix}, \quad (2.10.33)$$

denn wir hatten in Abschnitt 2.2 gezeigt, daß sich die Komponenten des Gradienten unter orthogonalen Transformationen wie ein Vektorfeld transformieren. Allerdings müssen wir hier beachten, daß jetzt die

Für die Berechnung der **Divergenz des Vektorfeldes** könnten wir ähnlich vorgehen. Es ist aber einfacher und anschaulicher, (2.9.14) auf den infinitesimalen Quader  $\Delta Q$ , der von den Tangentenvektoren der Koordinatenlinien, also

$$\frac{\partial \vec{x}}{\partial q_i} dq_i = \vec{T}_i dq_i = g_i \vec{e}_i' dq_i, \quad (2.10.34)$$

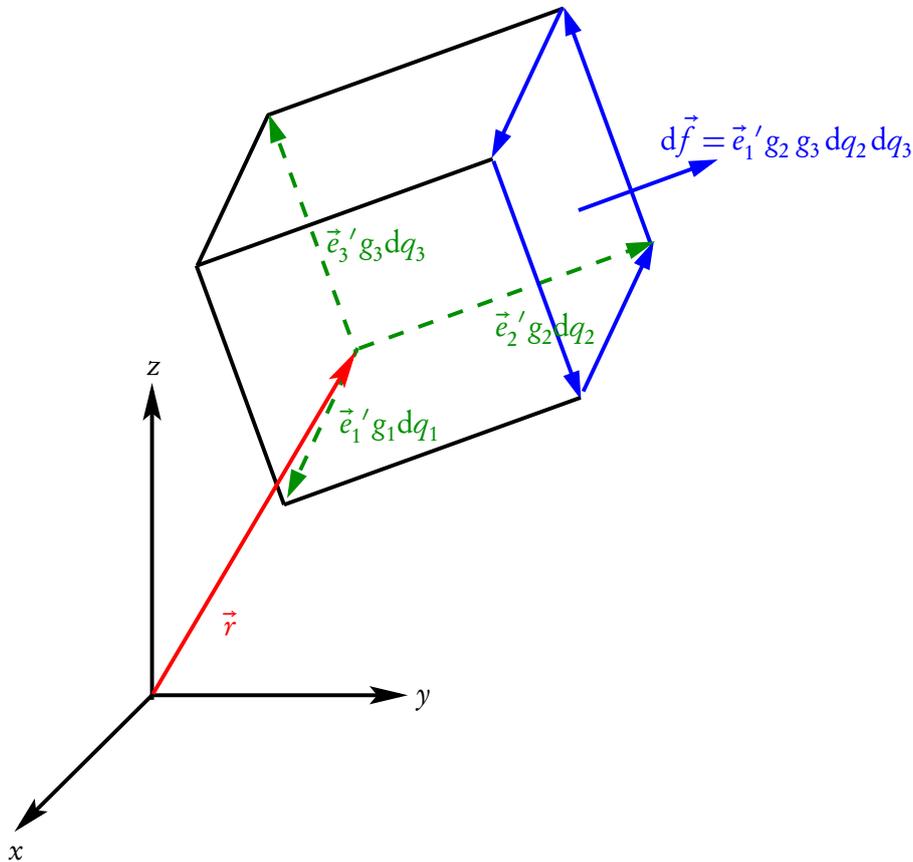
aufgespannt wird, anzuwenden. Für das Volumenintegral über diesen infinitesimalen Quader ergibt sich

$$\int_{\Delta Q} dV [\operatorname{div} \vec{A}(\vec{x})] = dV [\operatorname{div} \vec{A}(\vec{x})] = g_1 g_2 g_3 d^3 q \operatorname{div} \vec{A}. \quad (2.10.35)$$

Dabei haben wir das Volumenelement mit Hilfe von (2.9.1) und (2.10.5) ausgedrückt.

Das dazugehörige Flächenintegral über den Rand  $\partial \Delta Q$  unseres Quaders setzt sich aus den sechs infinitesimalen Seitenflächen des Quaders zusammen. Der Skizze unten entnehmen wir

$$\begin{aligned} \int_{\partial \Delta Q} d\vec{f} \cdot \vec{A}(\vec{x}) &= dq_2 dq_3 ([g_2 g_3 A_1]_{q_1+dq_1, q_2, q_3} - [g_2 g_3 A_1]_{q_1, q_2, q_3}) \\ &\quad + dq_3 dq_1 ([g_3 g_1 A_2]_{q_1, q_2+dq_2, q_3} - [g_3 g_1 A_2]_{q_1, q_2, q_3}) \\ &\quad + dq_1 dq_2 ([g_1 g_2 A_3]_{q_1, q_2, q_3+dq_3} - [g_1 g_2 A_3]_{q_1, q_2, q_3}) \\ &= d^3 q \left[ \frac{\partial (g_2 g_3 A_1)}{\partial q_1} + \frac{\partial (g_3 g_1 A_2)}{\partial q_2} + \frac{\partial (g_1 g_2 A_3)}{\partial q_3} \right]. \end{aligned} \quad (2.10.36)$$



Verwenden wir nun (2.10.35) und (2.10.36) in (2.9.14), erhalten wir schließlich

$$\begin{aligned} \operatorname{div} \vec{A} &= \frac{1}{g_1 g_2 g_3} \left[ \frac{\partial (g_2 g_3 A_1)}{\partial q_1} + \frac{\partial (g_3 g_1 A_2)}{\partial q_2} + \frac{\partial (g_1 g_2 A_3)}{\partial q_3} \right] \\ &= \frac{1}{g_1 g_2 g_3} \sum_{k=1}^3 \frac{\partial}{\partial q_k} \left( \frac{g_1 g_2 g_3}{g_k} A_k \right). \end{aligned} \quad (2.10.37)$$

Analog erhält man die Komponenten für die **Rotation des Vektorfeldes** aus (2.7.23) auf die drei infinitesimalen Rechteckflächen  $\Delta F_{ij}$ , die jeweils durch die Tangentenvektoren an die Koordinatenlinien  $\vec{T}_i dq_i = g_i \vec{e}'_i$

und  $\bar{T}_j dq_j = g_j \vec{e}_j'$  aufgespannt werden, anwendet, wobei nacheinander  $(i, j) \in \{(2, 3); (3, 1); (1, 2)\}$  gesetzt wird. Verwenden wir z.B. das erste Indexpaar, erhalten wir die erste Komponente der Rotation bzgl. der krummlinigen Koordinaten (s. wieder die Skizze oben). Das Flächenintegral ist

$$\int_{\Delta F_{23}} d\vec{f} \cdot \text{rot} \vec{A} = dq_2 dq_3 g_2 g_3 (\vec{e}_2' \times \vec{e}_3') \cdot \text{rot} \vec{A} = dq_2 dq_3 g_2 g_3 \vec{e}_1' \cdot \text{rot} \vec{A} = dq_2 dq_3 g_2 g_3 (\text{rot} \vec{A})'_1, \quad (2.10.38)$$

und das dazugehörige Linienintegral entlang des Randes  $\partial \Delta F_{23}$  lautet

$$\begin{aligned} \int_{\partial \Delta F_{23}} d\vec{x} \cdot \vec{A} &= dq_3 \left[ (g_3 A'_3)_{q_1, q_2 + dq_2, q_3} - (g_3 A'_3)_{q_1, q_2, q_3} \right] \\ &\quad - dq_2 \left[ (g_2 A'_2)_{q_1, q_2, q_3 + dq_3} - (g_2 A'_2)_{q_1, q_2, q_3} \right] \\ &= dq_2 dq_3 \left[ \frac{\partial (g_3 A'_3)}{\partial q_2} - \frac{\partial (g_2 A'_2)}{\partial q_3} \right]. \end{aligned} \quad (2.10.39)$$

Wegen des Stokesschen Integralsatzes sind (2.10.38) und (2.10.39) gleich, so daß sich schließlich

$$(\text{rot} \vec{A})'_1 = \frac{1}{g_2 g_3} \left[ \frac{\partial (g_3 A'_3)}{\partial q_2} - \frac{\partial (g_2 A'_2)}{\partial q_3} \right] \quad (2.10.40)$$

ergibt. Die beiden übrigen Komponenten finden wir auf analoge Weise durch Verwendung der anderen beiden infinitesimalen Rechteckflächen. Wir erhalten sie jedoch auch einfach durch **zyklische Vertauschung** der Indizes:

$$\begin{aligned} (\text{rot} \vec{A})'_2 &= \frac{1}{g_3 g_1} \left[ \frac{\partial (g_1 A'_1)}{\partial q_3} - \frac{\partial (g_3 A'_3)}{\partial q_1} \right], \\ (\text{rot} \vec{A})'_3 &= \frac{1}{g_1 g_2} \left[ \frac{\partial (g_2 A'_2)}{\partial q_1} - \frac{\partial (g_1 A'_1)}{\partial q_2} \right]. \end{aligned} \quad (2.10.41)$$

Es ist klar, daß für die Wahl kartesischer Koordinaten all diese Gleichungen die in Abschnitt 2.4 angegebene Form annehmen.

Wir stellen zur Übersicht die Differentialoperatoren in **Kugel- und Zylinderkoordinaten** zusammen.

**Kugelkoordinaten**  $(r, \vartheta, \varphi)$ :

$$\vec{\nabla} U = \text{grad} U = \vec{e}_r' \frac{\partial U}{\partial r} + \vec{e}_\vartheta' \frac{1}{r} \frac{\partial U}{\partial \vartheta} + \vec{e}_\varphi' \frac{1}{r \sin \vartheta} \frac{\partial U}{\partial \varphi}, \quad (2.10.42)$$

$$\vec{\nabla} \cdot \vec{A} = \text{div} \vec{A} = \frac{1}{r^2} \frac{\partial (r^2 A'_r)}{\partial r} + \frac{1}{r \sin \vartheta} \frac{\partial (\sin \vartheta A'_\vartheta)}{\partial \vartheta} + \frac{1}{r \sin \vartheta} \frac{\partial A'_\varphi}{\partial \varphi}, \quad (2.10.43)$$

$$\begin{aligned} \vec{\nabla} \times \vec{A} = \text{rot} \vec{A} &= \vec{e}_r' \frac{1}{r \sin \vartheta} \left[ \frac{\partial (\sin \vartheta A'_\varphi)}{\partial \vartheta} - \frac{\partial A'_\vartheta}{\partial \varphi} \right] + \vec{e}_\vartheta' \left[ \frac{1}{r \sin \vartheta} \frac{\partial A'_r}{\partial \varphi} - \frac{1}{r} \frac{\partial (r A'_\varphi)}{\partial r} \right] \\ &\quad + \vec{e}_\varphi' \frac{1}{r} \left[ \frac{\partial (r A'_\vartheta)}{\partial r} - \frac{\partial A'_r}{\partial \vartheta} \right]. \end{aligned} \quad (2.10.44)$$

Wenden wir schließlich (2.10.42) und (2.10.43) hintereinander an, erhalten wir für den **Laplaceoperator**

$$\begin{aligned} \Delta U = \vec{\nabla} \cdot \vec{\nabla} U = \text{div grad} U &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial U}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left( \sin \vartheta \frac{\partial U}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 U}{\partial \varphi^2} \\ &= \frac{1}{r} \frac{\partial^2}{\partial r^2} (r U) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left( \sin \vartheta \frac{\partial U}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 U}{\partial \varphi^2}. \end{aligned} \quad (2.10.45)$$

Zylinderkoordinaten  $(r, \varphi, z)$ :

$$\vec{\nabla} U = \text{grad } U = \vec{e}_r' \frac{\partial U}{\partial r} + \vec{e}_\varphi' \frac{1}{r} \frac{\partial U}{\partial \varphi} + \frac{\partial U}{\partial z} \vec{e}_z' \quad (2.10.46)$$

$$\vec{\nabla} \cdot \vec{A} = \text{div } \vec{A} = \frac{1}{r} \frac{\partial (rA_r')}{\partial r} + \frac{1}{r} \frac{\partial A_\varphi'}{\partial \varphi} + \frac{\partial A_z'}{\partial z}, \quad (2.10.47)$$

$$\vec{\nabla} \times \vec{A} = \text{rot } \vec{A} = \vec{e}_r' \left[ \frac{1}{r} \frac{\partial A_z'}{\partial \varphi} - \frac{\partial A_\varphi'}{\partial z} \right] + \vec{e}_\varphi' \left[ \frac{\partial A_r'}{\partial z} - \frac{\partial A_z'}{\partial r} \right] + \vec{e}_z' \frac{1}{r} \left[ \frac{\partial (rA_\varphi')}{\partial r} - \frac{\partial A_r'}{\partial \varphi} \right], \quad (2.10.48)$$

$$\Delta U = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial U}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 U}{\partial \varphi^2} + \frac{\partial^2 U}{\partial z^2}. \quad (2.10.49)$$

## 2.11 Solenoidalfelder und Vektorpotentiale

In den Abschnitten 2.6 und 2.8 haben wir die Frage untersucht, unter welchen Umständen ein vorgegebenes Vektorfeld  $\vec{V}$  der Gradient eines Skalarfeldes ist und wie ggf. dieses Skalarfeld, das **Potential** des Vektorfeldes, berechnet werden kann. Die Antwort war das **Lemma von Poincaré**, daß falls  $\vec{V}$  auf einem einfach zusammenhängenden Gebiet stetig partiell differenzierbar ist und dort  $\text{rot } \vec{V} = 0$  gilt, in diesem Gebiet ein bis auf eine Konstante eindeutig bestimmtes Potential  $\Phi$  existiert, so daß  $\vec{V} = -\text{grad } \Phi$  gilt. Das Potential ist dabei durch das Wegintegral entlang eines beliebigen Weges innerhalb dieses Gebiets von einem festen Punkt  $\vec{x}_0$  zum Punkt  $\vec{x}$  gegeben ist und das Skalarpotential bis auf eine Konstante eindeutig bestimmt ist.

Hier fragen wir nun, unter welchen Bedingungen ein stetig partiell differenzierbares Vektorfeld  $\vec{B}$  die Rotation eines anderen Vektorfeldes  $\vec{A}$  ist, d.h.

$$\vec{B}(\vec{x}) = \vec{\nabla} \times \vec{A}(\vec{x}). \quad (2.11.1)$$

Man nennt dann  $\vec{A}$  ein **Vektorpotential** für  $\vec{B}$ . Ein wichtiges Beispiel für Solenoidalfelder in der Physik sind **Magnetfelder**.

Zuerst leiten wir wieder eine notwendige Bedingung her. Dazu berechnen wir

$$\text{div } \vec{B} = \vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}). \quad (2.11.2)$$

Wäre  $\vec{\nabla}$  ein gewöhnlicher Vektor, würde die rechte Seite verschwinden. Wir zeigen nun, daß diese Annahme tatsächlich zutrifft, indem wir die Divergenz in kartesischen Koordinaten explizit ausrechnen. Zunächst gilt

$$\begin{aligned} B_1 = \partial_2 A_3 - \partial_3 A_2 &\Rightarrow \partial_1 B_1 = \partial_1 \partial_2 A_3 - \partial_1 \partial_3 A_2, \\ B_2 = \partial_3 A_1 - \partial_1 A_3 &\Rightarrow \partial_2 B_2 = \partial_2 \partial_3 A_1 - \partial_2 \partial_1 A_3, \\ B_3 = \partial_1 A_2 - \partial_2 A_1 &\Rightarrow \partial_3 B_3 = \partial_3 \partial_1 A_2 - \partial_3 \partial_2 A_1. \end{aligned} \quad (2.11.3)$$

Addieren wir nun die Gleichungen nach den Folgepfeilen, ergibt sich die Divergenz

$$\text{div } \vec{B} = \partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3 = \partial_1 \partial_2 A_3 - \partial_1 \partial_3 A_2 + \partial_2 \partial_3 A_1 - \partial_2 \partial_1 A_3 + \partial_3 \partial_1 A_2 - \partial_3 \partial_2 A_1. \quad (2.11.4)$$

Da voraussetzungsgemäß die partiellen Ableitungen der Komponenten von  $\vec{B}$  stetig sind, gilt dies notwendig auch für die zweiten Ableitungen der Komponenten von  $\vec{A}$ . Dann vertauschen aber die Ableitungsoperatoren, und daher folgt aus (2.11.4) tatsächlich, daß

$$\text{div } \vec{B} = 0 \quad (2.11.5)$$

ist. Für die Existenz eines Vektorpotentials ist also *notwendig* (2.11.5) erfüllt.

Wir wollen nun zeigen, daß (2.11.5) auch hinreichend ist für die **lokale Existenz eines Vektorpotentials**. Leider gibt es keine einfache Formel wie beim Skalarpotential, das wir wie oben gezeigt als Wegintegral darstellen können. Wir müssen also versuchen, die Gleichung (2.11.1) nach  $\vec{A}$  aufzulösen.

Dazu bemerken wir aber als erstes, daß offensichtlich das Vektorpotential nur **bis auf den Gradienten eines Skalarfeldes** bestimmt ist, denn wie wir oben gesehen haben, gilt für jedes zweimal stetig differenzierbare Skalarfeld  $\chi$

$$\vec{\nabla} \times \vec{\nabla} \chi = \text{rot grad } \chi = 0. \quad (2.11.6)$$

Das bedeutet, daß für jedes Vektorpotential  $\vec{A}$  von  $\vec{B}$  auch

$$\vec{A}' = \vec{A} + \vec{\nabla} \chi \quad (2.11.7)$$

ein Vektorpotential von  $\vec{B}$  ist. Wir können also zur Vereinfachung der Auflösung von (2.11.1) nach  $\vec{A}$  eine **Nebenbedingung** fordern. Wie wir sehen werden, erweist sich die Bedingung

$$A_3 = 0 \quad (2.11.8)$$

als sehr hilfreich. Um zu zeigen, daß man dies durch Wahl eines geeigneten Skalarfeldes  $\chi$  stets erreichen kann, nehmen wir an  $\vec{A}$  sei irgendein Vektorpotential von  $\vec{B}$  und suchen ein Skalarfeld  $\chi$ , so daß

$$A_3' = A_3 + \partial_3 \chi = 0 \quad (2.11.9)$$

ist. Diese Gleichung läßt sich aber sehr leicht lösen. Wir müssen sie nur bzgl.  $x^3$  integrieren:

$$\chi(x^1, x^2, x^3) = - \int_{x^{30}}^{x^3} dx'^3 A_3(x^1, x^2, x'^3). \quad (2.11.10)$$

Dabei gehen wir davon aus, daß es um den Punkt  $\vec{x}$  eine offene quaderförmige Umgebung parallel zu den Koordinatenachsen gibt, wo  $\vec{A}$  stetig differenzierbar ist. Dies ist sicher erfüllt, wenn  $\vec{B}$  in einem offenen Gebiet stetig ist. Dann liegt der Integrationsbereich von (2.11.10) ganz in dieser Umgebung, und das Integral ist daher eine Lösung von (2.11.9). Wir dürfen also annehmen, daß die Nebenbedingung (2.11.8) erfüllt ist.

In kartesischen Komponenten ausgeschrieben lautet dann die Gleichung (2.11.1)

$$\vec{B} = \vec{\nabla} \times \begin{pmatrix} A_1 \\ A_2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\partial_3 A_2 \\ \partial_3 A_1 \\ \partial_1 A_2 - \partial_2 A_1 \end{pmatrix}. \quad (2.11.11)$$

Aus der ersten Komponente dieser Gleichung folgt

$$A_2(x^1, x^2, x^3) = - \int_{x^{30}}^{x^3} dx'^3 B_1(x^1, x^2, x'^3) + A_2'(x^1, x^2) \quad (2.11.12)$$

und aus der zweiten

$$A_1(x^1, x^2, x^3) = \int_{x^{30}}^{x^3} dx'^3 B_1(x^1, x^2, x'^3) + A_1'(x^1, x^2). \quad (2.11.13)$$

Dabei haben wir beim Integrieren berücksichtigt, daß die Vorgabe der partiellen Ableitung nach  $x^3$  eine Funktion nur bis auf Funktionen, die von den beiden anderen unabhängigen Variablen  $x^1$  und  $x^2$  abhängen bestimmt ist. Diese können wir nun bestimmen, indem wir (2.11.12) und (2.11.13) in die dritte Komponente der Gleichung (2.11.11) einsetzen:

$$B_3(x^1, x^2, x^3) = - \int_{x^{30}}^{x^3} dx'^3 [\partial_1 B_1(x^1, x^2, x'^3) + \partial_2 B_2(x^1, x^2, x'^3)] + \partial_1 A_2'(x^1, x^2) - \partial_2 A_1'(x^1, x^2). \quad (2.11.14)$$

Wegen  $\operatorname{div} \vec{B} = \partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3 = 0$  folgt daraus

$$B_3(x^1, x^2, x^3) = \int_{x_0^3}^{x^3} dx'^3 \partial_3' B_3(x^1, x^2, x'^3) + \partial_1 A_2'(x^1, x^2) - \partial_2 A_1'(x^1, x^2). \quad (2.11.15)$$

Da voraussetzungsgemäß  $\vec{B}$  stetig partiell differenzierbar sein soll, können wir das Integral auswerten, woraus sich

$$\begin{aligned} B_3(x^1, x^2, x^3) &= B_3(x^1, x^2, x^3) - B_3(x^1, x^2, x_0^3) + \partial_1 A_2'(x^1, x^2) - \partial_2 A_1'(x^1, x^2) \\ &\Rightarrow \partial_1 A_2'(x^1, x^2) - \partial_2 A_1'(x^1, x^2) = B_3(x^1, x^2, x_0^3) \end{aligned} \quad (2.11.16)$$

ergibt. Dies zeigt, daß dank der Divergenzfreiheit des Vektorfeldes  $\vec{B}$  die Lösungen (2.11.14) tatsächlich konsistent mit der Existenz des Vektorpotentials  $\vec{A}$  mit  $A_3 = 0$  sind.

Die Nebenbedingung legt allerdings offensichtlich das Vektorpotential immer noch nicht eindeutig fest. Das ist unmittelbar einleuchtend, denn die Addition des Gradienten eines nur von  $x^1$  und  $x^2$  abhängigen Skalarfeldes ändert an den Eigenschaften des Vektorpotentials,  $\operatorname{rot} \vec{A} = \vec{B}$  und  $A_3 = 0$  zu erfüllen nichts. Wir können also in (2.11.13)  $A_1'(x^1, x^2) = 0$  setzen. Dann folgt aus (2.11.16)

$$A_1' = 0 \Rightarrow \partial_1 A_2'(x^1, x^2) = B_3(x^1, x^2, x_0^3). \quad (2.11.17)$$

Daraus folgt, daß durch die Wahl

$$A_2'(x^1, x^2) = \int_{x_0^2}^{x^2} dx'^2 B_3(x^1, x'^2, x_0^3) \quad (2.11.18)$$

das Vektorpotential vervollständigt wird. Sammeln wir also die Lösungsschritte (2.11.12), (2.11.17) und (2.11.18) erhalten wir als eine mögliche Lösung für das Vektorpotential

$$\begin{aligned} A_1(\vec{x}) &= \int_{x_0^3}^{x^3} dx'^3 B_2(x^1, x^2, x'^3), \\ A_2(\vec{x}) &= - \int_{x_0^2}^{x^2} dx'^2 B_1(x^1, x^2, x'^3) + \int_{x_0^2}^{x^2} dx'^2 B_3(x^1, x'^2, x_0^3), \\ A_3(\vec{x}) &= 0. \end{aligned} \quad (2.11.19)$$

## 2.12 Die Poisson-Gleichung und Green-Funktionen

In diesem Abschnitt beschäftigen wir uns mit einer typischen Fragestellung der Feldtheorie. Dazu betrachten wir das Newtonsche Gravitationsgesetz etwas genauer. Wie schon in Abschnitt 2.5 gesagt, ist die Gravitationskraft einer als punktförmig angenommenen im Ursprung des Koordinatensystems sitzenden Masse  $M$  auf eine bei  $\vec{x}$  andere Punktmasse  $m$ , die wir im folgenden **Probemasse** nennen, durch das Kraftfeld

$$\vec{F}(\vec{x}) = -\gamma m M \frac{\vec{x}}{r^3}, \quad r = |\vec{x}| \quad (2.12.1)$$

gegeben.

Dividieren wir diese Kraft durch die Masse der Probemasse, erhalten wir die **Gravitationsbeschleunigung**, die unabhängig von der Probemasse ist. Wir gelangen so zur Interpretation der Gravitation als **Feldwirkung**. Demnach erzeugt die Masse  $M$  ein Gravitationsfeld

$$\vec{g}(\vec{x}) = -\gamma M \frac{\vec{x}}{r^3}. \quad (2.12.2)$$

Man weist es einfach dadurch nach, daß man an die Stelle  $\vec{x}$  eine Probelmasse  $m$  setzt und die auf sie wirkende Gravitationskraft (2.12.1) mißt. In Abschnitt 2.5 haben wir auch gezeigt, daß dieses Feld ein Potentialfeld ist, das wegen (2.5.9) durch

$$\Phi(\vec{x}) = -\frac{\gamma M}{r} \quad (2.12.3)$$

gegeben ist. Wir nennen dieses Feld das **Gravitationspotential**. Dabei haben wir die willkürliche Konstante  $C = 0$  gesetzt. Dabei folgen wir der Konvention, daß das Gravitationspotential im Unendlichen verschwinden soll.

Wir fragen nun, wie die Gravitationskraft einer ausgedehnten Massenverteilung auf eine Punktladung zu berechnen ist. Dazu denken wir uns die Masse kontinuierlich gemäß der **Massendichteverteilung**  $\rho(\vec{x})$  über den Körper verteilt. Mit Newton gehen wir nun davon aus, daß jedes Massenelement am Punkt  $\vec{x}'$ , also  $dM = d^3x' \rho(\vec{x}')$  additiv zur Kraft auf eine Probeladung beiträgt. Dabei müssen wir nur beachten, daß der Ursprung dieser infinitesimalen **Quelle des Gravitationsfeldes** bei  $\vec{x}'$  und nicht im Koordinatenursprung setzt. Der entsprechende Beitrag zum Gravitationspotential ist demnach

$$d\Phi = -\frac{\gamma dM}{|\vec{x} - \vec{x}'|} \quad (2.12.4)$$

Insgesamt ist das Gravitationspotential der Massenverteilung demnach durch

$$\Phi(\vec{x}) = -\gamma \int_V d^3x' \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} \quad (2.12.5)$$

gegeben. Dabei nehmen wir an, daß  $V$  so gewählt ist, daß der gesamte Körper ganz im Inneren dieses Volumens liegt, also  $\rho(\vec{x}') = 0$  für  $\vec{x}' \in E^3 \setminus (V \cup \partial V)$  gilt.

Im Prinzip können wir nun das Gravitationspotential einer beliebigen Massenverteilung ausrechnen. In der Physik haben sich aber **lokale Feldgleichungen** als weitaus nützlicher erwiesen, d.h. wir suchen eine **partielle Differentialgleichung**, die das Gravitationspotential bzw. die Gravitationsbeschleunigung mit der Massenverteilung  $\rho(\vec{x})$  verknüpft.

Dazu betrachten wir noch einmal das Gravitationsfeld (2.12.2) einer einzelnen Punktmasse im Ursprung des Koordinatensystems. Eine solche Punktmasse ist in der Feldtheorie ein wenig problematisch, denn offenbar ist für solch eine Massenverteilung die Massendichte extrem singulär. Außer im Ursprung ist die Massendichte exakt 0, und da wir eine endliche Masse im Ursprung vereint haben, ist dort die Massendichte unendlich groß.

Wir können nun aber die im Ursprung konzentrierte Masse  $M$  durch ein **Flächenintegral** über das Gravitationsfeld (2.12.2) erhalten. Dazu sei  $V$  ein Volumen, bei dem der Koordinatenursprung durch eine kleine Kugel mit Radius  $\epsilon$  ausgespart ist. Sei  $V'$  nun also ein Volumen, das den Ursprung im Inneren enthält mit auf übliche Weise orientierter Randfläche  $\partial V'$  (Flächennormalenvektoren weisen aus dem Volumen heraus) und sei die Kugel mit Radius  $\epsilon$  um den Ursprung  $K_\epsilon(0) \subset V'$ . Wir wenden nun den Gaußschen Satz auf das Gravitationsfeld (2.12.2) auf das Volumen  $V = V' \setminus K_\epsilon(0)$  an. Dann besteht der auf die übliche Weise orientierte Rand von  $V$  zum einen aus dem Rand von  $V'$  und zum anderen aus der Kugeloberfläche  $S_\epsilon(0)$ , die so orientiert ist, daß die Normalenvektoren radial auf den Ursprung weisen, also ebenfalls weg vom Volumen  $V$ .

Durch direktes Nachrechnen in kartesischen Koordinaten weist man nun nach, daß

$$\operatorname{div} \vec{g}(\vec{x}) = \vec{\nabla} \cdot \vec{g}(\vec{x}) = 0 \quad (2.12.6)$$

gilt. Demnach folgt aus dem Gaußschen Satz

$$0 = \int_V d^3x \operatorname{div} \vec{g}(\vec{x}) = \int_{\partial V} d\vec{f} \cdot \vec{g}(\vec{x}) = \int_{\partial V'} d\vec{f} \cdot \vec{g}(\vec{x}) - \int_{S_\epsilon(0)} d\vec{f} \cdot \vec{g}(\vec{x}), \quad (2.12.7)$$

wobei wir in dem letzteren Flächenintegral die Orientierung wieder wie üblich mit den Normalenvektoren vom Kugelursprung weg gerichtet haben. Da für den Gaußschen Integralsatz allerdings die Orientierung über die Oberfläche der ausgeparten Kugel um den Ursprung umgekehrt zu richten ist, trägt das entsprechende Flächenintegral mit dem negativen Vorzeichen bei. Jedenfalls folgt aus (2.12.7)

$$\int_{\partial V'} d\vec{f} \cdot \vec{g}(\vec{x}) = \int_{S_\epsilon(0)} d\vec{f} \cdot \vec{g}. \quad (2.12.8)$$

Das rechtsstehende Integral läßt sich nun sehr leicht ausrechnen. Dazu verwenden wir unsere Standardparametrisierung der Kugelschale (2.7.1) und das entsprechende Resultat (2.7.12). Mit den entsprechenden Vorfaktoren erhalten wir also

$$\int_{\partial V'} d\vec{f} \cdot \vec{g}(\vec{x}) = -4\pi\gamma M. \quad (2.12.9)$$

Für irgendeine den Ursprung umschließende geschlossene Oberfläche  $\partial V'$ , erhalten wir also aus dem Gravitationspotential bis auf die Vorfaktoren  $-4\pi\gamma$  die in dem umschlossenen Volumen enthaltene felderzeugende Masse!

Nun liegt die Vermutung nahe, daß dies auch für die kontinuierliche Massenverteilung gilt. Dies führt auf die Idee, daß für eine kontinuierliche singularitätenfreie Massenverteilung das Gravitationspotential überall eine wohldefinierte Divergenz besitzt, und dann können wir den Gaußschen Integralsatz auf irgendein Volumen  $V$  anwenden, ohne Singularitäten durch kleine Kugeln ausschließen zu müssen. Diese Überlegung liefert

$$\int_V d^3x \operatorname{div} \vec{g}(\vec{x}) = \int_{\partial V} d\vec{f} \cdot \vec{g}(\vec{x}) \stackrel{!}{=} -4\pi\gamma \int_V d^3x \rho(\vec{x}) = -4\pi\gamma M_V, \quad (2.12.10)$$

wobei  $M_V$  die gesamte im Volumen  $V$  befindliche Masse ist. Da dies für jedes Volumen  $V$  gilt, können wir die Definition der Divergenz über ein Flächenintegral verwenden (s. Abschnitt 2.9.2). Ist nämlich unsere Hypothese (2.12.10) korrekt, so folgt

$$\operatorname{div} \vec{g}(\vec{x}) = \lim_{\Delta V \rightarrow 0} \frac{1}{\operatorname{vol}(\Delta V)} \int_{\partial \Delta V} d\vec{f}' \cdot \vec{g}(\vec{x}') = -4\pi\gamma \lim_{\Delta V \rightarrow 0} \frac{M_{\Delta V}}{\Delta V} = -4\pi\gamma \rho(\vec{x}). \quad (2.12.11)$$

Dies führt uns auf die Differentialgleichung

$$\operatorname{div} \vec{g}(\vec{x}) = -4\pi\gamma \rho(\vec{x}). \quad (2.12.12)$$

Nun erinnern wir uns, daß das Gravitationsfeld ein Potential besitzt. Es gilt demnach

$$\operatorname{div} \vec{g}(\vec{x}) = -\operatorname{div} \operatorname{grad} \Phi(\vec{x}) = -\Delta \Phi(\vec{x}) = -4\pi\gamma \rho(\vec{x}). \quad (2.12.13)$$

Dabei haben wir den in (2.9.22) eingeführten Laplace-Operator verwendet.

Um nun zu zeigen, daß die Annahme in (2.12.10) wirklich zutrifft, zeigen wir, daß (2.12.5) tatsächlich die **Poisson-Gleichung** (2.12.13) löst. Falls  $\vec{x} \notin V$  ist, können wir einfach den Laplaceoperator auf (2.12.5) anwenden, indem wir ihn in das Integral ziehen, denn für diesen Fall treten keine Singularitäten im Integranden auf. Nun ist aber

$$\Delta \frac{1}{|\vec{x} - \vec{x}'|} = \Delta' \frac{1}{|\vec{x} - \vec{x}'|} = 0, \quad \text{für } \vec{x} \neq \vec{x}', \quad (2.12.14)$$

wie man durch Bildung der partiellen Ableitungen direkt nachrechnet (*Übung!*). Daraus folgt, daß dann

$$\Delta \Phi(\vec{x}) = -\gamma \int_V d^3x' \rho(\vec{x}') \Delta \frac{1}{|\vec{x} - \vec{x}'|} = 0 \quad \text{für } \vec{x} \notin V \quad (2.12.15)$$

gilt. Da voraussetzungsgemäß für  $\vec{x} \notin V$  stets  $\rho(\vec{x}) = 0$  gelten soll, ist Die Poisson-Gleichung (2.12.13) erfüllt.

Nun betrachten wir den Fall, daß  $\vec{x} \in E^3$ . Dazu wenden wir den 2. Greenschen Satz (2.9.25) an, wobei wir  $\Phi_1 = \Phi$  und  $\Phi_2(\vec{x}') = 1/|\vec{x} - \vec{x}'|$  setzen. Als Integrationsvolumen wählen wir  $\tilde{V} = E^3 \setminus K_\epsilon(\vec{x})$ , wobei  $K_\epsilon(\vec{x})$  eine Kugel mit Radius  $\epsilon$  mit Mittelpunkt  $\vec{x}$  bezeichnet. Dann ist der Rand  $\partial\tilde{V}$  durch die Kugelschale  $S_\epsilon(\vec{x})$  gegeben. Entsprechend der Standardorientierung beim Gaußschen Integralsatz muß man für letztere die Normalvektoren in Richtung auf den Mittelpunkt  $\vec{x}$  weisend wählen (also vom Integrationsvolumen weg). Wir wollen aber die Standardorientierung verwenden und berücksichtigen dies durch das entsprechende zusätzliche Vorzeichen. Der Greensche Satz lautet für unseren Fall also

$$\int_{\tilde{V}} d^3\vec{x}' \left[ \Phi(\vec{x}') \Delta \frac{1}{|\vec{x} - \vec{x}'|} - \frac{1}{|\vec{x} - \vec{x}'|} \Delta \Phi(\vec{x}') \right] = \int_{\partial\tilde{V}} d\vec{f}' \cdot \left[ \Phi(\vec{x}') \vec{\nabla}' \frac{1}{|\vec{x} - \vec{x}'|} - \frac{1}{|\vec{x} - \vec{x}'|} \vec{\nabla}' \Phi(\vec{x}') \right]. \quad (2.12.16)$$

Im Volumenintegral auf der linken Seite fällt der erste Term wegen (2.12.14) weg und unter der Annahme, daß  $\Phi$  der Poisson-Gleichung  $\Delta\Phi = 4\pi\gamma\rho$  genügt, folgt

$$\int_{\tilde{V}} d^3\vec{x}' \left[ \Phi(\vec{x}') \Delta \frac{1}{|\vec{x} - \vec{x}'|} - \frac{1}{|\vec{x} - \vec{x}'|} \Delta \Phi(\vec{x}') \right] = -4\pi\gamma \int_{\tilde{V}} d^3\vec{x}' \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|}. \quad (2.12.17)$$

Die Oberflächenintegrale werten wir nun aus, indem wir  $S_\epsilon(\vec{x})$  mit Kugelkoordinaten parametrisieren:

$$\vec{x}' = \vec{x} + \epsilon \begin{pmatrix} \cos\varphi \sin\vartheta \\ \sin\varphi \sin\vartheta \\ \cos\vartheta \end{pmatrix}. \quad (2.12.18)$$

Der Flächenelementvektor berechnet sich gemäß (2.7.6) zu (Übung!)

$$d\vec{f}' = d\vartheta d\varphi \epsilon^2 \sin\vartheta \begin{pmatrix} \cos\varphi \sin\vartheta \\ \sin\varphi \sin\vartheta \\ \cos\vartheta \end{pmatrix}, \quad (2.12.19)$$

mit der Standardorientierung vom Kugelzentrum weg, was wir wie bereits oben erwähnt im Flächenintegral auf der rechten Seite von (2.12.16) durch ein zusätzliches Vorzeichen berücksichtigen müssen. Für das Flächenintegral benötigen wir noch

$$\vec{\nabla}' \frac{1}{|\vec{x} - \vec{x}'|} = -\frac{\vec{x}' - \vec{x}}{|\vec{x} - \vec{x}'|^3} \Rightarrow d\vec{f}' \cdot \vec{\nabla}' \frac{1}{|\vec{x} - \vec{x}'|} = -d\vartheta d\varphi \sin\vartheta \quad (2.12.20)$$

und

$$d\vec{f}' \cdot \vec{\nabla}' \Phi(\vec{x}') = \epsilon^2 \partial_\epsilon \check{\Phi}(\vec{x}). \quad (2.12.21)$$

Dabei haben wir verwendet, daß in Kugelkoordinaten  $(\epsilon, \vartheta, \varphi)$

$$\vec{e}_r = \vec{T}_r = \begin{pmatrix} \cos\varphi \sin\vartheta \\ \sin\varphi \sin\vartheta \\ \cos\vartheta \end{pmatrix} \quad (2.12.22)$$

und somit wegen der Kettenregel

$$\vec{e}_r \cdot \vec{\nabla}' \Phi(\vec{x}') = \frac{\partial \vec{x}}{\partial \epsilon} \cdot \vec{\nabla}' \Phi(\vec{x}') = \partial_\epsilon \check{\Phi}(\vec{x}'). \quad (2.12.23)$$

Setzen wir also (2.12.19-2.12.23) in der rechten Seite von (2.12.16) ein, so folgt

$$\int_{\partial\tilde{V}} d\vec{f}' \cdot \left[ \Phi(\vec{x}') \vec{\nabla}' \frac{1}{|\vec{x} - \vec{x}'|} - \frac{1}{|\vec{x} - \vec{x}'|} \vec{\nabla}' \Phi(\vec{x}') \right] = -\int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin\vartheta \left[ -\check{\Phi}(\vec{x}') - \epsilon \partial_\epsilon \check{\Phi}(\vec{x}') \right]. \quad (2.12.24)$$

Da  $\tilde{\Phi}$  zweimal partiell stetig differenzierbar ist, verschwindet das zweite Integral für  $\epsilon \rightarrow 0$  und nach dem Zwischenwertsatz ergibt das erste

$$\int_{\partial\tilde{V}} d\vec{f}' \cdot \left[ \Phi(\vec{x}') \vec{\nabla}' \frac{1}{|\vec{x} - \vec{x}'|} - \frac{1}{|\vec{x} - \vec{x}'|} \vec{\nabla}' \Phi(\vec{x}') \right] \stackrel{\epsilon \rightarrow 0}{=} 4\pi\Phi(\vec{x}). \quad (2.12.25)$$

Gleichsetzen mit der rechten Seite von (2.12.15) liefert schließlich das gewünschte Resultat

$$\Phi(\vec{x}) = -\gamma \lim_{\epsilon \rightarrow 0} \int_{\tilde{V}} d^3x' \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} = -\gamma \int_V d^3x' \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|}, \quad (2.12.26)$$

wobei wir verwendet haben, daß  $\rho(\vec{x}')$  nur im Inneren von  $V$  von 0 verschieden ist.

Der obige Beweis zeigt aber, daß dieses Resultat auch dann noch gilt, wenn  $\rho(\vec{x}')$  im Unendlichen schnell genug abfällt, so daß die oben betrachteten Integrale über  $\tilde{V}$  allesamt existieren.

In gewisser Weise haben wir eine Umkehroperation zum Laplaceoperator gefunden, denn wie wir gezeigt haben, folgt aus

$$\Delta\Phi_1(\vec{x}) = \Phi_2(\vec{x}), \quad (2.12.27)$$

daß

$$\Phi_1(\vec{x}) = \int_{E^3} d^3x' G(\vec{x}, \vec{x}') \Phi_2(\vec{x}') \quad \text{mit} \quad G(\vec{x}, \vec{x}') = -\frac{1}{4\pi|\vec{x} - \vec{x}'|} \quad (2.12.28)$$

gilt, sofern nur  $\Phi_2(\vec{x})$  hinreichend schnell verschwindet, so daß das Volumenintegral existiert. Dabei erfüllt die Lösung  $\Phi_1$  die **Randbedingung**, daß sie im Unendlichen verschwindet. Unter dieser Bedingung ist (2.12.28) eine eindeutige Lösung der Poisson-Gleichung (2.12.27). Man nennt in diesem Zusammenhang  $G$  eine **Green-Funktion des Laplace-Operators**<sup>11</sup>.

## 2.13 Der Helmholtzsche Zerlegungssatz der Vektoranalysis

In der Physik muß man oft ein Vektorfeld aus der Vorgabe seiner Divergenz („Quellen“) und seiner Rotation („Wirbeln“) bestimmen.

Gegeben sei ein skalares Feld  $\rho(\vec{x})$  und ein Vektorfeld  $\vec{w}(\vec{x})$ , und wir fragen nach der Existenz eines Vektorfeldes  $\vec{V}$ , so daß

$$\text{div } \vec{V} = \rho, \quad \text{rot } \vec{V} = \vec{w} \quad (2.13.1)$$

gilt. Es ist klar, daß für die (zumindest lokale) Existenz eines solchen Vektorfeldes  $\vec{V}$  die Konsistenzbedingung

$$\text{div } \vec{w} = 0 \quad (2.13.2)$$

erfüllt sein muß, d.h.  $\vec{w}$  ist ein Solenoidalfeld, dessen Vektorpotential  $\vec{V}$  sein soll.

Die Betrachtungen in den vorigen Abschnitten legen die Zerlegung des Vektorfeldes in zwei Anteile nahe:

$$\vec{V} = \vec{V}_1 + \vec{V}_2 \quad \text{mit} \quad \text{rot } \vec{V}_1 = 0 \quad \text{und} \quad \text{div } \vec{V}_2 = 0, \quad (2.13.3)$$

d.h. wir spalten  $\vec{V}$  in ein Potentialfeld  $\vec{V}_1$  und ein Solenoidalfeld  $\vec{V}_2$  auf, so daß

$$\text{div } \vec{V} = \text{div } \vec{V}_1 = \rho, \quad \text{rot } \vec{V}_1 = 0, \quad (2.13.4)$$

$$\text{rot } \vec{V} = \text{rot } \vec{V}_2 = \vec{w}, \quad \text{div } \vec{V}_2 = 0 \quad (2.13.5)$$

<sup>11</sup>Green-Funktionen von linearen Differentialoperatoren werden Ihnen in allen Theorie-Vorlesungen während des gesamten Physikstudiums begegnen, insbesondere in der Elektrodynamik, Quantenmechanik und schließlich der Quantenfeldtheorie.

gilt (vgl. (2.13.1)). Im folgenden zeigen wir, daß eine solche Zerlegung existiert, vorausgesetzt es sind bestimmte Glattheitsbedingungen und ein hinreichend schnelles Abfallen der Quellen  $\rho$  und  $\vec{w}$  im Unendlichen erfüllt. Wie wir gleich sehen werden, ist dafür hinreichend, daß diese Funktionen stetig differenzierbar und die Volumenintegrale

$$\int_{E^3} d^3x \frac{|\rho(\vec{x})|}{\sqrt{\epsilon^2 + \vec{x}^2}}, \quad \int_{E^3} d^3x \frac{|\vec{w}(\vec{x})|}{\sqrt{|\epsilon^2 + \vec{x}^2|}} \quad (2.13.6)$$

für  $\epsilon > 0$  existieren. Das bedeutet, daß  $|\rho|$  und  $|\vec{w}|$  im Unendlichen schneller als  $1/|\vec{x}|^2$  verschwinden müssen (*warum?*). Fordert man zusätzlich noch, daß  $\vec{V} \rightarrow 0$  für  $|\vec{x}| \rightarrow \infty$ , ist  $\vec{V}$  eindeutig durch die Vorgabe der Quellen  $\vec{w}$  und  $\rho$  bestimmt.

### 2.13.1 Bestimmung des Potentialfeldanteils

Wir beginnen mit der Aufgabe,  $\vec{V}_1$  zu bestimmen. Wegen  $\text{rot } \vec{V}_1 = 0$  existiert nach dem Poincaréschen Lemma (zumindest lokal) ein skalares Feld  $\Phi$ , so daß

$$\vec{V}_1 = -\text{grad } \Phi \quad (2.13.7)$$

gilt. Setzen wir dies in (2.13.4) ein, ergibt sich

$$\text{div } \vec{V}_1 = -\text{div grad } \Phi = -\Delta \Phi = \rho. \quad (2.13.8)$$

Dies bezeichnet man als die inhomogene Potentialgleichung oder auch als **Poisson-Gleichung**. Diese haben wir im vorigen Abschnitt gelöst. Verwenden wir also (2.12.27) und (2.12.28), ergibt sich bereits die Lösung

$$\Phi(\vec{x}) = \int_{E^3} d^3x' \frac{\rho(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}. \quad (2.13.9)$$

Der Potentialfeldanteil selbst berechnet sich durch Gradientenbildung gemäß unseres Ansatzes (2.13.7):

$$\vec{V}_1(\vec{x}) = -\text{grad } \Phi(\vec{x}) = \int_{E^3} d^3x' \frac{\rho(\vec{x}')}{4\pi} \frac{\vec{x} - \vec{x}'}{|\vec{x} - \vec{x}'|^3}. \quad (2.13.10)$$

### 2.13.2 Bestimmung des Solenoidalfeldanteils

Wir müssen nun noch den Solenoidalfeldanteil des Ausgangsvektorfeldes bestimmen, d.h. wir haben (2.13.5) zu lösen. Da  $\text{div } \vec{V}_2 = 0$  ist, existiert aufgrund unserer Betrachtungen in Abschnitt 2.11 ein **Vektorpotential**  $\vec{A}$ , so daß

$$\vec{V}_2 = \text{rot } \vec{A}. \quad (2.13.11)$$

Wie wir ebenfalls in Abschnitt 2.11 erläutert haben, ist  $\vec{A}$  nur bis auf ein Potentialvektorfeld bestimmt, so daß wir eine Nebenbedingung fordern können. Diesmal wird sich die folgende als **Coulomb-Bedingung**<sup>12</sup>

$$\text{div } \vec{A} = 0 \quad (2.13.12)$$

als besonders bequem erweisen. Setzen wir nämlich (2.13.11) in die erste Gleichung von (2.13.5) ein, finden wir für diese Nebenbedingung (*nachrechnen!*)

$$\text{rot rot } \vec{A} = \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - (\vec{\nabla} \cdot \vec{\nabla})\vec{A} = \text{grad div } \vec{A} - \Delta \vec{A} = -\Delta \vec{A} \stackrel{!}{=} \vec{w}. \quad (2.13.13)$$

<sup>12</sup>Diese Bezeichnung kommt aus der Elektrodynamik, wo statische Magnetfelder ein Vektorpotential besitzen.

Es sei auch an dieser Stelle betont, daß diese Gleichung **nur auf kartesische Komponenten** des Vektorfeldes  $\vec{A}$  angewandt werden darf, da nur für kartesische Orthonormalsysteme die Einheitsvektoren ortsunabhängig sind.

Wir werden also wieder auf **Poisson-Gleichungen** für die **kartesischen** Komponenten von  $\vec{A}$  geführt, so daß wir wieder die Green-Funktion des Laplace-Operators anwenden können:

$$\vec{A}(\vec{x}) = \int_{E^3} d^3x' \frac{\vec{w}(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}. \quad (2.13.14)$$

Allerdings müssen wir uns nun zuerst vergewissern, ob diese Lösung auch tatsächlich die Coulomb-Bedingung (2.13.12) erfüllt, die wir ja verlangt haben, um für  $\vec{A}$  eine Poisson-Gleichung zu erhalten. In der Tat gilt

$$\operatorname{div} \vec{A}(\vec{x}) = \int_{E^3} d^3x' \frac{1}{4\pi} \vec{w}(\vec{x}') \vec{\nabla}_x \frac{1}{|\vec{x} - \vec{x}'|} = \int_{E^3} d^3x' \frac{1}{4\pi} \vec{w}(\vec{x}') \left( -\vec{\nabla}_{x'} \frac{1}{|\vec{x} - \vec{x}'|} \right). \quad (2.13.15)$$

Wegen der Konsistenzbedingung (2.13.2) können wir dies mit Hilfe des Gaußschen Satzes in ein Oberflächenintegral umwandeln. Da der Rand des gesamten Raumes  $E^3$  aber im „Unendlichen“ liegt, können wir unter der Annahme, daß  $\vec{w}$  im Unendlichen hinreichend schnell verschwindet, schließen, daß dieses Oberflächenintegral tatsächlich verschwindet. Nach der Produktregel gilt

$$\operatorname{div}_{x'} \left[ \frac{\vec{w}(\vec{x}')}{|\vec{x} - \vec{x}'|} \right] = [\operatorname{div}_{x'} \vec{w}(\vec{x}')] \frac{1}{|\vec{x} - \vec{x}'|} + \vec{w}(\vec{x}') \cdot \operatorname{grad}_{x'} \frac{1}{|\vec{x} - \vec{x}'|}. \quad (2.13.16)$$

Wir haben also nach dem Gaußschen Satz

$$\operatorname{div} \vec{A}(\vec{x}) = -\frac{1}{4\pi} \int_{E^3} d^3x' \operatorname{div}_{x'} \left[ \frac{\vec{w}(\vec{x}')}{|\vec{x} - \vec{x}'|} \right] = \lim_{R \rightarrow \infty} \frac{1}{4\pi} \int_{\partial K_R} d\vec{f}(\vec{x}') \frac{\vec{w}(\vec{x}')}{|\vec{x} - \vec{x}'|} = 0, \quad (2.13.17)$$

falls  $\vec{w}(\vec{x}')$  im Unendlichen schneller als  $\mathcal{O}(|\vec{x}'|^{-1})$  verschwindet. Also ist die Coulomb-Eichbedingung erfüllt und somit (2.13.14) tatsächlich eine Lösung für das Vektorpotential. Für den Solenoidalfeldanteil finden wir schließlich

$$\begin{aligned} \vec{V}_2(\vec{x}) &= \operatorname{rot} \vec{A}(\vec{x}) = \int_{E^3} d^3x' \vec{\nabla}_x \times \frac{\vec{w}(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} \\ &= - \int_{E^3} d^3x' \vec{w}(\vec{x}') \times \frac{\vec{x} - \vec{x}'}{4\pi|\vec{x} - \vec{x}'|^3} = \int_{E^3} d^3x' \vec{w}(\vec{x}') \times \frac{\vec{x} - \vec{x}'}{4\pi|\vec{x} - \vec{x}'|^3}. \end{aligned} \quad (2.13.18)$$

Damit haben wir den **Helmholtzischen Zerlegungssatz** für Vektorfelder bewiesen. Wir fassen ihn noch einmal übersichtlich zusammen:

Seien  $\rho(\vec{x})$  und  $\vec{w}(\vec{x})$  im  $E^3$  definierte Felder, die im Unendlichen schneller als mit der Ordnung  $\mathcal{O}(|\vec{x}|^{-1})$  abfallen und erfülle  $\vec{w}$  die Bedingung  $\operatorname{div} \vec{w} = 0$ . Dann läßt sich ein gegebenes Vektorfeld  $\vec{V}$  eindeutig in einen **Potentialfeldanteil** und einen **Solenoidalfeldanteil** zerlegen, so daß

$$\vec{V} = \vec{V}_1 + \vec{V}_2, \quad \operatorname{div} \vec{V} = -\rho, \quad \operatorname{rot} \vec{V} = \vec{w}, \quad \text{wobei} \quad (2.13.19)$$

$$\vec{V}_1 = -\operatorname{grad} \Phi \quad \text{mit} \quad \Phi(\vec{x}) = \int_{E^3} d^3x' \frac{\rho(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}, \quad (2.13.20)$$

$$\vec{V}_2 = \operatorname{rot} \vec{A} \quad \text{mit} \quad \vec{A}(\vec{x}) = \int_{E^3} d^3x' \frac{\vec{w}(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}. \quad (2.13.21)$$

Die Felder selbst berechnen sich aus der **Quelle**  $\rho$  und der **Wirbelstärke**  $\vec{w}$  des Vektorfeldes zu

$$\vec{V}_1(\vec{x}) = \int_{E^3} d^3x' \frac{\rho(\vec{x}')}{4\pi} \frac{\vec{x} - \vec{x}'}{|\vec{x} - \vec{x}'|^3}, \quad \vec{V}_2(\vec{x}) = \int_{E^3} d^3x' \vec{w}(\vec{x}') \times \frac{\vec{x} - \vec{x}'}{4\pi|\vec{x} - \vec{x}'|^3}. \quad (2.13.22)$$

## 2.14 Allgemein kovariante Vektoranalysis

Zum Abschluß unserer Darstellung der Vektoranalysis betrachten wir noch die Differentialoperatoren in allgemeinen krummlinigen Koordinaten (oder **generalisierte Koordinaten**), die wir mit  $q = (q^1, q^2, q^3)$  bezeichnen wollen. Wir bemerken, daß dies i.a. *keine* kontravarianten Vektorkomponenten sind (obwohl freilich solche Vektorkomponenten ebenfalls diese allgemeinste Darstellung der Vektoranalysis erfüllen). Der Ortsvektor ist dann eine Funktion dieser verallgemeinerten Koordinaten  $\vec{x} = \vec{x}(q)$ . Diese Koordinaten beschreiben nun i.a. Teilgebiete des gesamten Euklidischen Raumes. Die generalisierten Koordinaten sollen nun so geartet sein, daß die Vektoren

$$\vec{g}_k(q) = \partial_k \vec{x}(q) = \frac{\partial}{\partial q^k} \vec{x}(q) \quad (2.14.1)$$

linear unabhängig sind und folglich eine Basis bilden. Man bezeichnet diese Vektoren als die zu den generalisierten Koordinaten  $q$  gehörigen **holonomen Basisvektoren**. Sie sind offenbar die Tangentialvektoren der **Koordinatenlinien** in dem Punkt  $\vec{x}(q)$ , d.h. derjenigen drei Kurven, die sich durch Festhalten zweier der drei generalisierten Koordinaten ergeben.

Nun können wir den allgemeinen Ricci-Kalkül anwenden, müssen jedoch beachten, daß die Basisvektoren jetzt ortsabhängig sind. Insbesondere sind auch die Metriktensorkomponenten

$$g_{jk} = \vec{g}_j \cdot \vec{g}_k \quad (2.14.2)$$

von  $q$  und damit vom Ort abhängig.

Um die Differentialoperatoren grad, div, rot bestimmen zu können, führen wir am einfachsten auch kartesische Basisvektoren  $\vec{e}'_k$  ein. Die Komponenten des Ortsvektors bezeichnen wir mit  $x'^k$ . Die Matrizen des Basiswechsels zwischen den holonomen Basisvektoren und der kartesischen Basis sind nun ebenfalls ortsabhängig:

$$\vec{g}_j = T^k{}_j \vec{e}'_k, \quad \vec{e}'_k = U^j{}_k \vec{g}_j, \quad (2.14.3)$$

denn die  $\vec{e}'_k$  sind unabhängig, die  $\vec{g}_j$  aber abhängig vom Ort. Wir müssen also bei der Bildung von Ableitungen diese Ortsabhängigkeit beachten. Als erstes bemerken wir, daß wegen  $\vec{e}'_k = \text{const}$

$$\vec{g}_j = \partial_j \vec{x} = \frac{\partial x'^k}{\partial q^j} \vec{e}_j \quad (2.14.4)$$

und damit

$$T^k{}_j = \frac{\partial x'^k}{\partial q^j} = \partial_j x'^k \quad (2.14.5)$$

ist. Aus der Kettenregel folgt nun

$$\frac{\partial q^j}{\partial q^k} = \delta^j_k = \frac{\partial q^j}{\partial x'^l} \frac{\partial x'^l}{\partial q^k} = (\partial'_l q^j)(\partial_j x'^l) = (\partial'_l q^j) T^l{}_j \quad (2.14.6)$$

und damit wegen der Eindeutigkeit der Umkehrmatrix

$$U^j{}_k = \partial'_k q^j = \frac{\partial q^j}{\partial x'^k}. \quad (2.14.7)$$

Mit (2.14.2) folgt

$$g_{jk} = \vec{g}_j \cdot \vec{g}_k = T^a{}_j T^b{}_k \vec{e}'_a \cdot \vec{e}'_b = \delta_{ab} T^a{}_j T^b{}_k \quad (2.14.8)$$

bzw. in Matrizen=Schreibweise

$$(g_{jk}) = \overleftrightarrow{g} = \hat{T}^T \hat{T}. \quad (2.14.9)$$

Für die kontravarianten Komponenten der Metrik folgt daraus

$$\overleftrightarrow{g} = \overleftrightarrow{g}^{-1} = \hat{T}^{-1} \hat{T}^{-1T} = \hat{U} \hat{U}^T, \quad (2.14.10)$$

d.h.

$$g^{jk} = U^j{}_a U^k{}_b \delta^{ab}. \quad (2.14.11)$$

Im folgenden ist es nützlich, neben den **kovarianten Basisvektoren**  $\vec{g}_j$  durch

$$\vec{g}^k = g^{kj} \vec{g}_j \quad (2.14.12)$$

auch **kontravariante Basisvektoren** einzuführen. Es gilt

$$\vec{g}^j \cdot \vec{g}_k = g^{jl} \vec{g}_l \cdot \vec{g}_k = g^{jl} g_{lk} = \delta^l_k. \quad (2.14.13)$$

Damit können wir einen beliebigen Vektor  $\vec{v}$  durch ko- und kontravariante Komponenten ausdrücken, d.h. es folgt

$$\vec{v} = v^j \vec{g}_j \Rightarrow \vec{g}^k \cdot \vec{v} = v^j \vec{g}^k \cdot \vec{g}_j = v^j \delta^k_j = v^k \quad (2.14.14)$$

bzw.

$$\vec{v} = v_j \vec{g}^j \Rightarrow \vec{g}_k \cdot \vec{v} = \vec{g}_k \cdot \vec{g}^j v_j = \delta^j_k v_j = v_k. \quad (2.14.15)$$

Betrachten wir nun den **Gradienten eines Skalarfeldes**. Dazu verwenden wir zunächst die Darstellung in kartesischen Koordinaten und verwenden die eben entwickelten Beziehungen, um auf die krummlinigen Koordinaten umzurechnen. Es ist klar, daß für die kartesischen Koordinaten wegen  $g'_{jk} = \delta_{jk}$  und  $g'^{jk} = \delta^{jk}$  stets  $\vec{e}'^k = \vec{e}'_k$  gilt, d.h. wir können

$$\vec{\nabla} \Phi = \vec{e}'^k \partial'_k \Phi \quad (2.14.16)$$

schreiben. Nun gilt

$$\vec{e}'^k = \delta^{kl} \vec{e}'_l = \delta^{kl} U^a{}_l \vec{g}_a \quad (2.14.17)$$

und

$$\partial'_k \Phi = \partial'_k q^j \partial_j \Phi \stackrel{(2.14.7)}{=} U^j{}_k \partial_j \Phi. \quad (2.14.18)$$

Setzen wir also (2.14.17) und (2.14.16) in (2.14.16) ein, erhalten wir

$$\vec{\nabla} \Phi = \delta^{kl} U^a{}_l \vec{g}_a U^j{}_k \partial_j \Phi \stackrel{(2.14.11)}{=} g^{aj} \vec{g}_a \partial_j \Phi = \vec{g}^j \partial_j \Phi. \quad (2.14.19)$$

Dies bedeutet, daß die partiellen Ableitungen eines Skalarfeldes nach den generalisierten Koordinaten **kovariante Vektorkomponenten** bzgl. der Basis  $\vec{g}_k$  bilden.

Nach dieser Vorarbeit können wir nun auch die **allgemeine Kovarianz** dieses Formalismus betrachten, d.h. das Verhalten der diversen Größen unter beliebigem Wechsel von einem System generalisierter Koordinaten  $q^j$  zu einem anderen  $q'^k$  betrachten. Dabei gehen wir davon aus, daß diese Koordinaten zumindest in einer bestimmten (offenen) Region des Vektorraums den Ortsvektor umkehrbar eindeutig bestimmen. Dann kann man auch in jedem Punkt eine umkehrbar eindeutige Abbildung zwischen den  $q^j$  und  $q'^k$  angeben. Nun gilt

$$\vec{g}'_k = \partial'_k \vec{x} = \frac{\partial}{\partial q'^k} \vec{x} = \partial'_k q^j \partial_j \vec{x} = \partial'_k q^j \vec{g}_j \Rightarrow U^j{}_k = \partial'_k q^j. \quad (2.14.20)$$

Es ist daraus unmittelbar klar, daß die Metrikkomponenten  $g_{jk} = g_{kj}$  sich auch bzgl. solcher allgemeiner Koordinatentransformationen kovariant transformieren

$$g'_{jk} = \vec{g}'_j \cdot \vec{g}'_k = U^a_j U^b_k \vec{g}_a \cdot \vec{g}_b = U^a_j U^b_k g_{ab}. \quad (2.14.21)$$

Wie oben können wir nun auch die kontravarianten Basisvektoren

$$\vec{g}^j = g^{jk} \vec{g}_k \quad (2.14.22)$$

eingeführen, wobei  $g^{jk}$  wieder die Komponenten der zu  $g_{lm}$  inversen Matrix bezeichnen, d.h. es gilt

$$\vec{g}'^j = g'^{ja} \vec{g}'_a = T^j_b T^a_c g^{bc} U^d_a \vec{g}_d = T^j_b \delta_c^d g^{bc} \vec{g}_d = T^j_b g^{bc} \vec{g}_c = T^j_b \vec{g}^b, \quad (2.14.23)$$

wobei wir wieder  $\hat{T} = \hat{U}^{-1}$  definiert haben, d.h.

$$T^j_k = \partial_k q'^j. \quad (2.14.24)$$

Damit ist sofort klar, daß der Gradient eines Skalarfeldes

$$\vec{\nabla}\Phi = \vec{g}^j \partial_j \Phi \quad (2.14.25)$$

ein Vektorfeld ist.

Ebenso definiert die partielle Ableitung eines Vektorfeldes nach  $q^j$  ein Objekt, das sich bzgl. des entstehenden freien Index'  $j$  allgemein kovariant transformiert, denn es gilt

$$\partial'_j \vec{V} = \partial'_j q^k \partial_k \vec{V} = U^k_j \partial_k \vec{V}. \quad (2.14.26)$$

Um die Komponenten von  $\partial_j \vec{V}$  bzgl. der Basis  $\vec{g}_k$  zu finden, müssen wir nun beachten, daß die  $\vec{g}_j$  von den  $q^k$  abhängen:

$$\partial_j \vec{V} = (\partial_j V^k) \vec{g}_k + V^k (\partial_j \vec{g}_k) \quad (2.14.27)$$

Es ist klar, daß man den Vektor  $\partial_j \vec{g}_k$  wie jeden Vektor nach den Basisvektoren  $\vec{g}_j$  entwickeln kann. Wir definieren damit die sogenannten **Christoffel-Symbole**  $\Gamma^l_{jk}$

$$\partial_j \vec{g}_k = \partial_j \partial_k \vec{x} = \Gamma^l_{jk} \vec{g}_l \Rightarrow \Gamma^l_{jk} = \vec{g}^l \cdot (\partial_j \vec{g}_k). \quad (2.14.28)$$

Wir halten fest, daß wegen der Vertauschbarkeit partieller Ableitungen für zweimal stetig partiell differenzierbare Funktionen das Christoffelsymbol in den beiden unteren Indizes symmetrisch ist

$$\Gamma^l_{jk} = \Gamma^l_{kj}. \quad (2.14.29)$$

Wir bemerken jedoch sogleich, daß trotz der Notation mit hoch- und tiefgestellten Indizes das Christoffel-symbol *keine* gemischten Tensorkomponenten definiert, denn es ist

$$\begin{aligned} \Gamma^a_{bc} &= \vec{g}^a \cdot (\partial_b \vec{g}'_c) = T^a_j \vec{g}^j \cdot U^k_b \partial_k (U^l_c \vec{g}_l) \\ &= T^a_j \vec{g}^j U^k_b \cdot (U^l_c \partial_k \vec{g}_l + \vec{g}_l \partial_k U^l_c) \\ &= T^a_j U^k_b U^l_c \Gamma^j_{kl} + T^a_j U^k_b \partial_k U^j_c. \end{aligned} \quad (2.14.30)$$

Den letzten Term können wir noch vereinfachen vermöge

$$U^k_b \partial_k U^j_c = U^k_b T^d_k \partial'_d U^j_c = \delta^d_b \partial'_d U^j_c = \partial'_b U^j_c = \partial'_b \partial'_c q^j \quad (2.14.31)$$

und damit

$$\Gamma'^a{}_{bc} = T^a{}_j U^k{}_b U^l{}_c \Gamma^j{}_{kl} + T^a{}_j \partial'_b \partial'_c q^j. \quad (2.14.32)$$

Setzen wir nun (2.14.28) in (2.14.27) ein, folgt

$$\partial'_j \vec{V} = \vec{g}_l (\partial'_j V^l + \Gamma^l{}_{jk} V^k) =: \vec{g}_l \nabla'_j V^l. \quad (2.14.33)$$

Man bezeichnet den Operator  $\nabla'_j$  als **kovariante Ableitung**, denn mit (2.14.20) folgt

$$\partial'_j \vec{V} = \vec{g}'_k \nabla'_j V'^k = U^b{}_k \vec{g}_b \nabla'_j V'^k = U^a{}_j \partial_a \vec{V} = U^a{}_j \vec{g}_b \nabla_a V^b \Rightarrow U^a{}_j \nabla_a V^b = U^b{}_k \nabla'_j V'^k. \quad (2.14.34)$$

Multiplizieren mit  $T^l{}_b$  liefert

$$\nabla'_j V'^l = U^a{}_j T^l{}_b \nabla_a V^b, \quad (2.14.35)$$

d.h. genau das Transformationsverhalten für gemischt ko- und kontravariante Tensorkomponenten.

Wir können freilich auch kovariante Ableitungen von kovarianten Vektorkomponenten definieren. Dazu schreiben wir

$$\vec{V} = \vec{g}^j V_j \Rightarrow \partial_k \vec{V} = \vec{g}^j \partial_k V_j + V_j \partial_k \vec{g}^j. \quad (2.14.36)$$

Die Ableitung des kontravarianten Basisvektors läßt sich nun auf die Christoffel-Symbole zurückführen, denn es gilt

$$\vec{g}^j \cdot \vec{g}_l = \delta^j_l \Rightarrow \partial_k (\vec{g}^j \cdot \vec{g}_l) = 0 = (\partial_k \vec{g}^j) \cdot \vec{g}_l + \vec{g}^j \cdot \partial_k \vec{g}_l \stackrel{(2.14.38)}{=} (\partial_k \vec{g}^j) \cdot \vec{g}_l + \Gamma^j{}_{kl} \Rightarrow \partial_k \vec{g}^j = -\Gamma^j{}_{kl} \vec{g}^l. \quad (2.14.37)$$

Dies in (2.14.36) eingesetzt ergibt

$$\partial_k \vec{V} = \vec{g}^l (\partial_k V_l - \Gamma^j{}_{kl} V_j) =: \vec{g}^l \nabla_k V_l. \quad (2.14.38)$$

Wir wollen nun zeigen, daß wir die Christoffelsymbole auch allein mit Hilfe der Komponenten des Metrikensors berechnen können. In der Tat folgt mit (2.14.28)

$$\partial_j g_{kl} = \partial_j (\vec{g}_k \cdot \vec{g}_l) = (\partial_j \vec{g}_k) \cdot \vec{g}_l + \vec{g}_k \cdot \partial_j \vec{g}_l = \Gamma^i{}_{jk} g_{il} + \Gamma^i{}_{jl} g_{ik} = \Gamma_{ljk} + \Gamma_{kjl}, \quad (2.14.39)$$

wobei wir zur Abkürzung auf die Christoffelsymbole die übliche Regel zum „Indexziehen“  $\Gamma_{ljk} = g_{li} \Gamma^i{}_{jk}$  anwenden, obwohl es sich *nicht* um Tensorkomponenten handelt. Jedenfalls können wir nun (2.14.39) mit zyklisch vertauschten Indizes hinschreiben:

$$\partial_k g_{lj} = \Gamma_{jkl} + \Gamma_{ljk}, \quad \partial_l g_{jk} = \Gamma_{kjl} + \Gamma_{jkl}, \quad (2.14.40)$$

wobei wir ausgenutzt haben, daß gemäß (2.14.29) die Christoffelsymbole bzgl. Vertauschen der hinteren beiden Indizes symmetrisch sind. Mit (2.14.39) folgt daraus

$$\Gamma_{ljk} = \frac{1}{2} (\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}) \Rightarrow \Gamma^i{}_{jk} = \frac{1}{2} g^{il} (\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}). \quad (2.14.41)$$

Nun können wir auch die Divergenz eines Vektorfeldes bestimmen. In kartesischen Koordinaten ist  $\partial_j V^j$  ein Skalar und definiert die Divergenz des Vektorfeldes  $\vec{V}$ . Da in kartesischen Koordinaten die Christoffel-Symbole allesamt verschwinden, können wir folglich einfach definieren

$$\vec{\nabla} \cdot \vec{V} = \text{div } \vec{V} = \nabla_j V^j. \quad (2.14.42)$$

Da diese Gleichung nämlich in kartesischen Koordinaten gilt, gilt sie auch in allgemeinen Krümmigen Koordinaten, da  $\nabla_j V^k$  gemischt ko- und kontravariante Tensorkomponenten bilden. Entsprechend ist die

Kontraktion, also Gleichsetzen der Indizes, wobei wir die Einsteinsche Summationskonvention anwenden, also die Summe über  $j$  ausgeführt wird, ein Skalarfeld bzgl. allgemeiner Koordinatentransformationen. Der Ausdruck (2.14.42) läßt sich aber noch vereinfachen. Es ist nämlich

$$\nabla_j V^j = \partial_j V^j + \Gamma^j_{jk} V^k. \quad (2.14.43)$$

Mit (2.14.41) folgt

$$\Gamma^j_{jk} = \frac{1}{2} g^{jl} (\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}) = \frac{1}{2} g^{jl} \partial_k g_{jl}. \quad (2.14.44)$$

Nun ist  $g^{jk}$  die Umkehrmatrix von  $g_{jk}$ . Aus (1.6.48) und (1.6.49) folgt dann

$$\partial_j g = (\partial_j g_{kl}) \frac{\partial g}{\partial g_{kl}} = (\partial_j g_{kl}) g^{kl} g \quad (2.14.45)$$

Also ist gemäß (2.14.44)

$$\Gamma^j_{jk} = \frac{1}{2g} \partial_k g = \frac{1}{\sqrt{g}} \partial_k \sqrt{g}. \quad (2.14.46)$$

Daraus folgt für (2.14.43)

$$\vec{\nabla} \cdot \vec{V} = \nabla_k V^k = \frac{1}{\sqrt{g}} (\sqrt{g} \partial_k V^k + V^k \partial_k \sqrt{g}) = \frac{1}{\sqrt{g}} \partial_k (\sqrt{g} V^k). \quad (2.14.47)$$

Schließlich können wir auch den Levi-Civita-Tensor aus (1.8.17) übernehmen, denn mit demselben Beweis sind auch

$$\Delta_{jkl} = \sqrt{g} \epsilon_{jkl}, \quad \Delta^{jkl} = \frac{1}{\sqrt{g}} \epsilon^{jkl} \quad (2.14.48)$$

allgemein ko- bzw. kontravariante Tensorkomponenten<sup>13</sup>.

Für die Rotation eines Vektorfeldes finden wir dann

$$(\text{rot } \vec{V})^j = \Delta^{jkl} \nabla_k V_l = \frac{1}{2} \Delta^{jkl} (\nabla_k V_l - \nabla_l V_k). \quad (2.14.49)$$

Nun gilt

$$\nabla_k V_l - \nabla_l V_k = \partial_k V_l - \Gamma^j_{kl} V_j - \partial_l V_k - \Gamma^j_{lk} V_j = \partial_k V_l - \partial_l V_k. \quad (2.14.50)$$

Für die antisymmetrisierte kovariante Ableitung kovarianter Vektorkomponenten darf man also auch gewöhnliche partielle Ableitungen schreiben:

$$(\text{rot } \vec{V})^j = \frac{1}{2} \Delta^{jkl} (\partial_k V_l - \partial_l V_k) = \Delta^{jkl} \partial_k V_l. \quad (2.14.51)$$

Es ist klar, daß wir in diesem Formalismus auch die verschiedenen Integrale der Vektoranalysis in allgemein kovariant formulieren können. Beginnen wir mit dem **Wegintegral eines Vektorfeldes**. Offenbar können wir einen Weg durch eine differenzierbare Funktion  $\mathcal{C} : t \mapsto q^k(t)$  mit  $t \in (a, b)$  der generalisierten Koordinaten beschreiben, und für ein Vektorfeld gilt

$$\int_{\mathcal{C}} d\vec{x} \cdot \vec{V} = \int_a^b dt \frac{dq^k}{dt} \partial_k \vec{x} \cdot \vec{V} = \int_a^b dt \frac{dq^k}{dt} \vec{g}_k \cdot \vec{V} = \int_a^b dt \frac{dq^k}{dt} V_k = \int_a^b dt \frac{dq^k}{dt} g_{kl} V^l. \quad (2.14.52)$$

<sup>13</sup>Dies gilt allerdings nur für orientierungserhaltende allgemeine Koordinatentransformationen, also für solche Transformationen, für die  $\det \hat{T} > 0$  ist. Das läßt sich allerdings offenbar stets durch geeignete Anordnung der neuen Koordinaten erreichen, so daß dies keine wesentliche Einschränkung darstellt.

Das **Volumenintegral über ein Skalarfeld** ergibt sich ebenso einfach aus der Verallgemeinerung des Vorgehens bei seiner Definition mit Hilfe kartesischer Koordinaten in Abschnitt 2.9. Wir müssen nur (2.9.1) mit Hilfe des Levi-Civita-Tensor allgemein kovariant schreiben. Ist das Volumen  $V$  durch irgendein Gebiet  $G \subseteq \mathbb{R}^3$  der generalisierten Koordinaten  $q^k$  definiert ist offenbar

$$\int_V d^3 \vec{x} \Phi = \int_G d^3 q \sqrt{g} \Phi \quad (2.14.53)$$

gegeben, denn es ist mit den kartesischen Koordinaten und der entsprechenden Transformationsmatrix zu den beliebigen generalisierten Koordinaten  $q^k$

$$d^3 \vec{x} = d^3 q \epsilon_{ijk} \partial_1 x'^i \partial_2 x'^j \partial_3 x'^k = d^3 q \det \hat{T} = d^3 q \sqrt{g}. \quad (2.14.54)$$

Ebenso ergibt sich aus der Parametrisierung einer Fläche im Raum mit Hilfe der generalisierten Koordinaten  $\mathcal{F} : (u_1, u_2) \mapsto q^k(u_1, u_2)$ , wobei  $(u_1, u_2) \in F \subseteq \mathbb{R}^2$  mit Hilfe des Levi-Civita-Tensors für das **Flächenintegral über ein Vektorfeld**

$$\int_{\mathcal{F}} d^2 \vec{F} \cdot \vec{V} = \int_F d^2 u \Delta_{jkl} \frac{\partial q^j}{\partial u^1} \frac{\partial q^k}{\partial u^2} V^l = \int_F d^2 u \sqrt{g} \epsilon_{jkl} \frac{\partial q^j}{\partial u^1} \frac{\partial q^k}{\partial u^2} V^l. \quad (2.14.55)$$

Es ist nun klar, daß die **Integralsätze** auch im Sinne dieses kovarianten Kalküls gelten. Betrachten wir zuerst den **Gaußschen Integralsatz**. Sei dazu  $\mathcal{V} \subseteq E^3$  ein Bereich („Volumen“) im Euklidischen Raum und  $V \subseteq \mathbb{R}^3$  der entsprechende Bereich für die generalisierten Koordinaten  $q^k$ . Dann entspricht der Rand  $\partial \mathcal{V}$  dem Rand  $\partial V$ . Dieser Rand  $\partial V$  sei wieder durch zwei generalisierte Koordinaten  $(u^1, u^2) \in F \subseteq \mathbb{R}^2$  parametrisiert. Wir behaupten nun, daß nun für allgemein kovariante Vektorkomponenten

$$\int_{\mathcal{V}} d^3 \vec{x} \vec{\nabla} \cdot \vec{V} = \int_{\partial \mathcal{V}} d^2 \vec{F} \cdot \vec{V} \quad (2.14.56)$$

gilt. Zum Beweis müssen wir nur (2.14.47), (2.14.53) und (2.14.55) verwenden:

$$\int_{\mathcal{V}} d^3 \vec{x} \vec{\nabla} \cdot \vec{V} = \int_V d^3 q \sqrt{g} \frac{1}{\sqrt{g}} \partial_k (\sqrt{g} V^k) = \int_V d^3 q \partial_k (\sqrt{g} V^k). \quad (2.14.57)$$

Nun können auf den letzten Ausdruck den Gaußschen Integralsatz wie für kartesische Komponenten hergeleitet anwenden, denn es handelt sich ja um den Ausdruck der Divergenz mit einfachen partiellen Ableitungen wie in diesem Spezialfall, d.h. es gilt

$$\int_{\mathcal{V}} d^3 \vec{x} \vec{\nabla} \cdot \vec{V} = \int_F d^2 u \epsilon_{ijk} \frac{\partial q^i}{\partial u^1} \frac{\partial q^j}{\partial u^2} (\sqrt{g} V^k) = \int_F d^2 u \Delta_{ijk} \frac{\partial q^i}{\partial u^1} \frac{\partial q^j}{\partial u^2} V^k = \int_{\partial \mathcal{V}} d^2 \vec{F} \cdot \vec{V}, \quad (2.14.58)$$

und das ist (2.14.56).

Ebenso folgt der **Stokessche Integralsatz**. Sei dazu  $\mathcal{F}$  wieder eine Fläche mit der Parametrisierung  $q^k(u^1, u^2)$ ,  $(u^1, u^2) \in F \subseteq \mathbb{R}^2$  und  $\partial \mathcal{F}$  ihr Rand parametrisiert durch  $q^k(t)$  mit  $t \in (a, b)$ . Dann folgt mit (2.14.55) und (2.14.51)

$$\int_{\mathcal{F}} d^2 \vec{F} \cdot (\vec{\nabla} \times \vec{V}) = \int_F d^2 u \Delta_{ijk} \frac{\partial q^i}{\partial u^1} \frac{\partial q^j}{\partial u^2} \Delta^{klm} \partial_l V_m. \quad (2.14.59)$$

Mit (2.14.48) folgt aber

$$\int_{\mathcal{F}} d^2 \vec{F} \cdot (\vec{\nabla} \times \vec{V}) = \int_F d^2 u \epsilon_{ijk} \frac{\partial q^i}{\partial u^1} \frac{\partial q^j}{\partial u^2} \epsilon^{klm} \partial_l V_m, \quad (2.14.60)$$

und hierauf läßt sich nun wieder der Stokessche Integralsatz für kartesische Vektorkomponenten anwenden:

$$\int_{\mathcal{F}} d^2 \vec{F} \cdot (\vec{\nabla} \times \vec{V}) = \int_{\partial F} dq^k V_k = \int_a^b dt \frac{dq^k}{dt} V_k = \int_{\partial \mathcal{F}} d\vec{x} \cdot \vec{V}, \quad (2.14.61)$$

und das ist der Stokessche Integralsatz.

### 2.14.1 Konsistenz mit dem Kalkül in krummlinigen Orthonormalkoordinaten

Wir wollen nun noch zeigen, daß der obige allgemein-kovariante Kalkül in holonomen Koordinaten auch geeignet ist, unsere früheren Resultate für den Fall krummliniger Orthonormalkoordinaten in Abschnitt 2.10 herzuleiten. Wir geben dazu die Einsteinsche Summationskonvention in diesem Abschnitt auf und schreiben wieder die Summenzeichen explizit hin. Für krummlinige Orthogonalkoordinaten  $q^k$  gilt natürlich zunächst der allgemeine Kalkül. Sie sind nur dadurch speziell, daß die holonomen Basisvektoren  $\vec{g}_k$  orthogonal zueinander sind, d.h. es gilt

$$(g_{jk}) = (\vec{g}_j \cdot \vec{g}_k) = \text{diag}(g_1^2, g_2^2, g_3^2) \Rightarrow g^{jk} = \text{diag}(g_1^{-2}, g_2^{-2}, g_3^{-2}). \quad (2.14.62)$$

Wir führen dann die normierten Basisvektoren

$$\vec{e}'_k = \frac{1}{g_k} \vec{g}_k \quad (2.14.63)$$

ein. Die  $\vec{e}'_k$  bilden dann ein kartesisches Koordinatensystem, das allerdings von den  $q^k$  abhängt. Für die entsprechenden Metrikkomponenten gilt aber offenbar

$$g'_{jk} = \delta_{jk} \Rightarrow g'^{jk} = \delta^{jk} = g_j^2 g^{jk}. \quad (2.14.64)$$

Dabei ist in dieser Gleichung über  $j$  natürlich *nicht* zu summieren. Deshalb müssen wir hier und im folgenden, wie oben bemerkt, die Einsteinsche Summationskonvention aufgeben. Wir nehmen weiter wie in Abschnitt 2.10 an, daß die Reihenfolge der  $q^k$  so gewählt ist, daß die  $\vec{e}'_k$  überall ein rechtshändiges Dreibein bilden, d.h.  $\vec{e}_3 = \vec{e}_1 \times \vec{e}_2$ . Für die kontravariante holonome Basis folgt daraus

$$\vec{g}^k = \sum_j g^{jk} \vec{g}_k = \frac{1}{g_k^2} \vec{g}_k = \frac{1}{g_k} \vec{e}'_k = \frac{1}{g_k} \vec{e}'^k. \quad (2.14.65)$$

Für die ko- und kontravarianten Vektorkomponenten bzgl. der holonomen Basis folgt nun

$$\vec{V} = \sum_{k=1}^3 V^k \vec{g}_k = \sum_{k=1}^3 V^k g_k \vec{e}'_k = \sum_{k=1}^3 V_k \vec{g}^k = \sum_{k=1}^3 \frac{1}{g_k} V_k \vec{e}'^k \Rightarrow V'^k = V'_k = V^k g_k = \frac{1}{g_k} V_k. \quad (2.14.66)$$

Weiter ist

$$g = \det \vec{g} = g_1^2 g_2^2 g_3^2 \Rightarrow \sqrt{g} = g_1 g_2 g_3. \quad (2.14.67)$$

Jetzt brauchen wir nur die oben hergeleiteten Formeln für die Komponenten bzgl. der holonomen Basis und die dazugehörigen Basisvektoren mit Hilfe von (2.14.63-2.14.67) auf die entsprechenden Komponenten bzgl. der orthonormierten Basis umzuschreiben. Für den Gradienten benötigen wir nur (2.14.25) und (2.14.65):

$$\vec{\nabla} \Phi = \sum_{j=1}^3 \vec{g}^j \partial_j \Phi = \sum_{j=1}^3 \frac{1}{g_j} \vec{e}'^j \partial_j \Phi \Rightarrow (\vec{\nabla} \Phi)_j = (\vec{\nabla} \phi)^j = \frac{1}{g_j} \partial_j \Phi, \quad (2.14.68)$$

was mit (2.10.33) übereinstimmt.

Für die **Divergenz** eines Vektorfeldes gehen wir von (2.14.47) aus und erhalten wieder mit (2.14.66) und (2.14.67)

$$\vec{\nabla} \cdot \vec{V} = \frac{1}{\sqrt{g}} \sum_{j=1}^3 \partial_{j=1}^3 (g V^j) = \frac{1}{g_1 g_2 g_3} \sum_{j=1}^3 \partial_j \left( \frac{g_1 g_2 g_3}{g_j} V'^j \right) = \frac{1}{g_1 g_2 g_3} \sum_{j=1}^3 \partial_j \left( \frac{g_1 g_2 g_3}{g_j} V'_j \right), \quad (2.14.69)$$

was (2.10.37) bestätigt.

Für die **Rotation** finden wir schließlich mit (2.14.51), (2.14.48) und (2.14.63-2.14.67)

$$\vec{\nabla} \times \vec{V} = \sum_{i,j,k=1}^3 \Delta^{ijk} \vec{g}_i \partial_j V_k = \frac{1}{g_1 g_2 g_3} \sum_{i,j,k=1}^3 \epsilon^{ijk} g_i \vec{e}'_i \partial_j (g_k V'_k). \quad (2.14.70)$$

Es ist also

$$(\vec{\nabla} \times \vec{V})'_i = (\vec{\nabla} \times \vec{V})'^i = \frac{1}{g_1 g_2 g_3} \sum_{j,k} \epsilon^{ijk} g_i \partial_j (g_k V'_k). \quad (2.14.71)$$

Schreibt man die Komponenten einzeln aus, erkennt man sofort (*Übung*), daß dies in der Tat mit (2.10.40-2.10.41) übereinstimmt.

## Kapitel 3

# Tensoren und Tensorfelder

In diesem Kapitel erweitern wir unsere Betrachtungen über Vektorräume und „lineare Strukturen“ noch um den Begriff der **Multilinearformen** oder **Tensoren** bzw. Tensorfelder. Der Daß wir diese eigentlich eher zur linearen Algebra gehörigen Teilgebiete der linearen Algebra erst jetzt besprechen und nicht bereits in Kapitel 1 ist darin begründet, daß wir die für die Physik besonders wichtige **Hauptachsentransformation** von symmetrischen Tensoren 2. Stufe einfacher mit Hilfe analytischer Methoden als mit rein algebraischen Hilfsmitteln behandeln können.

In der Physik spielen Tensorgrößen eine große Rolle. In der Vorlesung Mechanik 1 begegnet uns der **Trägheitstensor** eines starren Körpers, und für die entsprechenden Anwendungen benötigen wir bereits die Hauptachsentransformation. In der Kontinuumsmechanik kommen Tensoren zweiter Stufe ebenfalls häufig vor, meist auch als **Tensorfelder**. So gibt es z.B. den **Dehnungstensor**, der die Verzerrung eines elastischen Körpers bei Anlegen äußerer Kräfte an seiner Oberfläche, die wiederum durch den **Spannungstensor** entlang dieser Oberfläche beschrieben werden, angibt. Der Name Tensor für solche mathematischen Konstrukte geht aus der Bedeutung des Dehnungstensors hervor (lat. tenere=strecken, verzerren).

### 3.1 Multilinearformen

Wir betrachten einen beliebigen euklidischen Vektorraum  $V$  der endlichen Dimension  $n$ . Dann heißt eine Abbildung  $A : V^k \rightarrow \mathbb{R}$  eine **Multilinearform** oder **Tensor**  $k$ -ter Stufe, wenn sie in jedem Argument linear ist, d.h. es gilt für alle Vektoren  $\vec{x}_1, \dots, \vec{x}_k, \vec{y}_1, \vec{y}_2$  und reelle Zahlen  $\lambda_1, \lambda_2$

$$A(\vec{x}_1, \dots, \vec{x}_j, \lambda_1 \vec{y}_1 + \lambda_2 \vec{y}_2, \vec{x}_{j+1}, \dots, \vec{x}_k) = \lambda_1 A(\vec{x}_1, \dots, \vec{x}_{j-1}, \vec{y}_1, \vec{x}_{j+1}, \dots, \vec{x}_k) + \lambda_2 A(\vec{x}_1, \dots, \vec{x}_{j-1}, \vec{y}_2, \vec{x}_{j+1}, \dots, \vec{x}_k). \quad (3.1.1)$$

Sei nun  $(\vec{e}_j)$  mit  $j \in \{1, \dots, n\}$  eine beliebige kartesische Basis. Dann folgt aus der Multilinearität der Abbildung  $A$ , daß diese bereits durch die  $n^k$  reellen Zahlen

$$A_{j_1 j_2 \dots j_k} = A(\vec{e}_{j_1}, \vec{e}_{j_2}, \dots, \vec{e}_{j_k}), \quad j_i \in \{1, \dots, n\} \quad (3.1.2)$$

vollständig und eindeutig charakterisiert ist. Man nennt  $T_{j_1 \dots j_k}$  **kartesische Tensorkomponenten** des Tensors  $T$ .

Da  $(\vec{e}_j)$  nämlich eine kartesische Basis ist, gilt

$$A(\vec{a}, \vec{b}, \dots, \vec{z}) = \sum_{j_1, \dots, j_k=1}^n A_{j_1, j_2, \dots, j_k} a_{j_1} b_{j_2} \dots z_{j_k}, \quad (3.1.3)$$

wobei  $a_j = \vec{e}_j \cdot \vec{a}$  die  $j$ -te Komponente des Vektors  $\vec{a}$  bzgl. der kartesischen Basis  $(\vec{e}_j)$  bezeichnet.

Ein Spezialfall sind Tensoren 1. Stufe, die man kurz auch **Linearformen** nennt. Sei also  $L$  eine Linearform. Sie ist eindeutig durch ihre Tensorkomponenten  $L_j = L(\vec{e}_j)$  bzgl. der Basis  $(\vec{e}_j)$  bestimmt, und es gilt

$$L(\vec{x}) = \sum_{j=1}^n L_j x_j. \quad (3.1.4)$$

Definieren wir nun den Vektor

$$\vec{L} = \sum_{j=1}^n L_j \vec{e}_j, \quad (3.1.5)$$

so folgt daraus sofort

$$L(\vec{x}) = \vec{L} \cdot \vec{x}. \quad (3.1.6)$$

Gibt man umgekehrt  $\vec{L} \in V$  beliebig vor, definiert (3.1.6) eine Linearform  $L$ . Es gibt also eine umkehrbar eindeutige Abbildung zwischen Linearformen auf einem euklidischen Vektorraum und den Vektoren dieses Vektorraums.

Wir betrachten nun das Transformationsverhalten der Tensorkomponenten bzgl. orthogonaler Transformationen, die den Wechsel von einer kartesischen Basis  $(\vec{e}_j)$  zu einer anderen kartesischen Basis  $(\vec{e}'_j)$  beschreiben.

Es sei also

$$\vec{e}_j = \sum_{i=1}^n T_{ij} \vec{e}'_i, \quad \vec{e}'_i = \sum_{j=1}^n U_{ji} \vec{e}_j. \quad (3.1.7)$$

Aus Abschnitt 1.3 wissen wir, daß für die entsprechenden Matrizen die Orthogonalitätsbeziehung

$$\hat{U} = \hat{T}^{-1} = \hat{T}^T \Leftrightarrow T_{ij} = \vec{e}'_i \cdot \vec{e}_j = U_{ji} \quad (3.1.8)$$

gilt.

Damit ist es aber sehr leicht, die Tensorkomponenten bzgl. der einen Basis in durch die Tensorkomponenten der jeweils anderen Basis umzurechnen, denn es gilt wegen der Multilinearität der Abbildung

$$A'_{i_1 \dots i_k} = A(\vec{e}'_{i_1}, \dots, \vec{e}'_{i_k}) = \sum_{j_1, \dots, j_k=1}^n U_{j_1 i_1} \dots U_{j_k i_k} A_{j_1 \dots j_k} = \sum_{j_1, \dots, j_k=1}^n T_{i_1 j_1} \dots T_{i_k j_k} A_{j_1 \dots j_k}. \quad (3.1.9)$$

Die kartesischen Tensorkomponenten transformieren sich also gemäß (1.3.5) bzgl. jedes Index' analog zu den kartesischen Komponenten eines Vektors  $\vec{x}$ .

Offensichtlich kann man aus zwei Tensoren  $k$ -ter  $A$  und  $B$  durch Linearkombination einen neuen Tensor  $C = \lambda A + \mu B$  bilden, indem man für alle  $\vec{x}_1, \dots, \vec{x}_k \in V$

$$C(\vec{x}_1, \dots, \vec{x}_k) = \lambda A(\vec{x}_1, \dots, \vec{x}_k) + \mu B(\vec{x}_1, \dots, \vec{x}_k) \quad (3.1.10)$$

definiert. Offensichtlich bilden mit dieser Definition die Tensoren  $k$ -ter Stufe einen  $n^k$ -dimensionalen Vektorraum, den wir mit  $L_k(V)$  bezeichnen.

Weiter können wir aber auch ein äußeres **Tensorprodukt** definieren. Sei  $A \in L_{k_1}(V)$  und  $B \in L_{k_2}$ , so definiert man  $C = A \otimes B \in L_{k_1+k_2}(V)$  dadurch, daß man für alle  $\vec{x}_1, \dots, \vec{x}_{k_1}, \vec{y}_1, \dots, \vec{y}_{k_2} \in V$

$$C(\vec{x}_1, \dots, \vec{x}_{k_1}, \vec{y}_1, \dots, \vec{y}_{k_2}) = A(\vec{x}_1, \dots, \vec{x}_{k_1}) B(\vec{y}_1, \dots, \vec{y}_{k_2}) \quad (3.1.11)$$

setzt.

Insbesondere kann man mit der Identifikation der Linearformen mit Vektoren vermöge (3.1.4) die  $k$ -fachen Produkte der kartesischen Basisvektoren  $(\vec{e}_j)$

$$E_{j_1, \dots, j_k} = \vec{e}_{j_1} \otimes \dots \otimes \vec{e}_{j_k} \quad (3.1.12)$$

bilden. Aus der Definition (3.1.2) der Tensorkomponenten bzgl. dieser Basis ist klar (*nachprüfen!*), daß

$$A = \sum_{j_1, \dots, j_k=1}^n A_{j_1 \dots j_k} E_{j_1, \dots, j_k} \quad (3.1.13)$$

gilt. Daraus folgt, daß für jedes Orthonormalsystem  $(\vec{e}_j)$  die Tensoren (3.1.12) eine **Basis** des Tensorraumes  $L_k(V)$  bilden. Außerdem folgt aus dem Transformationsverhalten der Basisvektoren unter orthogonalen Transformationen (3.1.7), daß stets

$$A = \sum_{j_1, \dots, j_k=1}^n A_{j_1 \dots j_k} E_{j_1, \dots, j_k} = \sum_{j_1, \dots, j_k=1}^n A'_{j_1 \dots j_k} E'_{j_1, \dots, j_k} \quad (3.1.14)$$

gilt.

Schließlich können wir für Tensoren mindestens 2. Stufe noch die **Kontraktion** oder **Spurbildung** definieren. Sei also  $A \in L_k(V)$  mit  $k \geq 2$ . Dann ist

$$B = \sum_{i, j_1, \dots, j_{k-2}} A_{i i j_1, \dots, j_{k-2}} E_{j_1, \dots, j_{k-2}} \in L_{k-2}(V) \quad (3.1.15)$$

ein Tensor  $(k-2)$ -ter Stufe. Das Transformationsverhalten der Basisvektoren und Tensorkomponenten unter orthogonalen Transformationen zeigt sofort, daß diese Definition unabhängig von der Wahl der Basis ist. Man nennt die Operation (3.1.16) auch **Kontraktion bzgl. der ersten beiden Tensorindizes**. Es ist klar, daß man analog auch ein beliebiges anderes Paar von Indizes oder mehrere Indizes zugleich kontrahieren kann. Man erhält dann immer Tensoren der entsprechenden niedrigeren Stufe.

Die Kontraktion eines Tensors zweiter Stufe ergibt definitionsgemäß eine reelle Zahl, die man als **Spur** (engl. trace) des Tensors bezeichnet, d.h. für  $T \in L_2(V)$

$$\text{Tr} A = \sum_{i=1}^k A_{ii}. \quad (3.1.16)$$

## 3.2 Symmetrische Bilinearformen und Hauptachsentransformationen

Die Tensoren 2. Stufe bezeichnet man als **Bilinearform**. Besonders wichtig sind die **symmetrischen Bilinearformen**. Dabei heißt eine Bilinearform  $B \in L_k(V)$  symmetrisch, wenn für alle  $\vec{x}_1, \vec{x}_2 \in V$

$$B(\vec{x}_1, \vec{x}_2) = B(\vec{x}_2, \vec{x}_1) \quad (3.2.1)$$

gilt. Insbesondere bilden die Komponenten dieser Bilinearform wegen

$$B_{ij} = B(\vec{e}_i, \vec{e}_j) = B(\vec{e}_j, \vec{e}_i) = B_{ji}. \quad (3.2.2)$$

eine **symmetrische Matrix**, die wir mit  $\overleftrightarrow{B}$  bezeichnen wollen. Dann gilt mit den entsprechenden Spaltenvektoren  $\vec{x} = (x_j)$  und  $\vec{y} = (y_j)$  aus den Vektorkomponenten von  $\vec{x}$  und  $\vec{y}$  bzgl. der kartesischen Basis  $(\vec{e}_j)$

$$B(\vec{x}, \vec{y}) = \vec{x}^T \overleftrightarrow{B} \vec{y}. \quad (3.2.3)$$

Aus dem Transformationsverhalten von Tensor- und Vektorkomponenten gemäß (3.1.9)

$$\overleftrightarrow{B}' = \hat{T} \overleftrightarrow{B} \hat{T}^T, \quad \vec{x}' = \hat{T} \vec{x}, \quad \vec{y}' = \hat{T} \vec{y} \quad (3.2.4)$$

sowie der Orthogonalität der Transformationsmatrix ( $\hat{T}^T = \hat{T}^{-1}$ ) die Invarianz von (3.2.3) unter orthogonalen Basistransformationen

$$\vec{x}'^T \overset{\leftrightarrow}{B} \vec{y}' = (\hat{T}\vec{x})^T \hat{T} \overset{\leftrightarrow}{B} \hat{T}^T \hat{T} \vec{y} = \vec{x}^T \hat{T}^T \hat{T} \overset{\leftrightarrow}{B} \vec{y} = \vec{x}^T \overset{\leftrightarrow}{B} \vec{y}. \quad (3.2.5)$$

Dies ergibt sich auch sofort daraus, daß links in (3.2.3) eine von der Basis unabhängige Größe steht.

Im folgenden wollen wir die Frage klären, wie man die Matrix  $\overset{\leftrightarrow}{B}$  durch Wahl der Orthonormalbasis in eine möglichst einfache Form bringen kann. Dazu betrachten wir das **Eigenwertproblem** zu dieser Matrix  $\overset{\leftrightarrow}{B}$ . Wir bezeichnen einen Vektor  $\vec{e} \neq 0$  als **Eigenvektor** der Matrix  $\overset{\leftrightarrow}{B}$ , wenn es ein  $\lambda \in \mathbb{R}$  gibt, so daß

$$\overset{\leftrightarrow}{B} \vec{e} = \lambda \vec{e} \quad (3.2.6)$$

und bezeichnen dann  $\lambda$  als **Eigenwert** der Matrix  $\overset{\leftrightarrow}{B}$ .

Da weiter voraussetzungsgemäß  $\overset{\leftrightarrow}{B}^T = \overset{\leftrightarrow}{B}$  ist, folgt durch Transponieren der Gleichung (3.2.6)

$$\vec{e}^T \overset{\leftrightarrow}{B}^T = \vec{e}^T \overset{\leftrightarrow}{B} = \lambda \vec{e}^T. \quad (3.2.7)$$

Seien nun  $\vec{e}_1$  und  $\vec{e}_2$  zwei Eigenvektoren zu den Eigenwerten  $\lambda_1$  und  $\lambda_2$ . Dann folgt aus (3.2.6)

$$B(\vec{e}_1, \vec{e}_2) = \vec{e}_1^T \overset{\leftrightarrow}{B} \vec{e}_2 = \lambda_2 \vec{e}_1^T \vec{e}_2 = \lambda_2 \vec{e}_1 \cdot \vec{e}_2 \quad (3.2.8)$$

und mit (3.2.7)

$$B(\vec{e}_1, \vec{e}_2) = \vec{e}_1^T \overset{\leftrightarrow}{B} \vec{e}_2 = \lambda_1 \vec{e}_1^T \vec{e}_2 = \lambda_1 \vec{e}_1 \cdot \vec{e}_2 \quad (3.2.9)$$

Subtrahieren wir voneinander, folgt

$$(\lambda_1 - \lambda_2) \vec{e}_1 \cdot \vec{e}_2 = 0. \quad (3.2.10)$$

Falls  $\lambda_1 \neq \lambda_2$ , folgt daraus  $\vec{e}_1 \cdot \vec{e}_2 = 0$ , d.h. zwei Eigenvektoren einer symmetrischen Matrix zu verschiedenen Eigenwerten stehen stets orthogonal zueinander.

Weiter stellen wir fest, daß die Eigenvektoren zu einem Eigenwert  $\lambda$  einen Untervektorraum aufspannen. Seien nämlich  $\vec{e}_1$  und  $\vec{e}_2$  Eigenvektoren zum gleichen Eigenwert  $\lambda$ , so gilt für beliebige Zahlen  $\mu_1, \mu_2 \in \mathbb{R}$

$$\overset{\leftrightarrow}{B}(\mu_1 \vec{e}_1 + \mu_2 \vec{e}_2) = \mu_1 \overset{\leftrightarrow}{B} \vec{e}_1 + \mu_2 \overset{\leftrightarrow}{B} \vec{e}_2 = \mu_1 \lambda \vec{e}_1 + \mu_2 \lambda \vec{e}_2 = \lambda(\mu_1 \vec{e}_1 + \mu_2 \vec{e}_2), \quad (3.2.11)$$

d.h. auch jede Linearkombination von Eigenvektoren zum Eigenwert  $\lambda$  sind wieder Eigenvektoren zu diesem Eigenwert. Wir bezeichnen den Untervektorraum der Eigenvektoren von  $\overset{\leftrightarrow}{B}$  zum Eigenwert  $\lambda$  als **Eigenraum zum Eigenwert  $\lambda$**  und bezeichnen ihn mit  $\text{Eig}(\overset{\leftrightarrow}{B}, \lambda)$ .

Nun können wir im Unterraum  $\text{Eig}(\overset{\leftrightarrow}{B}, \lambda)$  eine Orthonormalbasis wählen und zu einer Orthonormalbasis für den ganzen Vektorraum ergänzen. Da die Vektoren in Eigenräumen zu verschiedenen Eigenwerten gemäß (3.2.10) aufeinander senkrecht stehen, können wir zuerst zu jedem Eigenraum eine Orthonormalbasis bilden und dann diese Vektoren zusammenfassen. Diese sind alle Einheitsvektoren und paarweise orthogonal zueinander. Wir können dann sicher auch diese Eigenvektoren von  $\overset{\leftrightarrow}{B}$  zu einer Orthonormalbasis des ganzen  $\mathbb{R}^n$  vervollständigen.

Es erhebt sich nun die Frage, ob es eine Orthonormalbasis gibt, die *nur* aus Eigenvektoren von  $B$  besteht. Dies ist besonders bequem, weil dann die Matrix  $\overset{\leftrightarrow}{B}$  bzgl. dieser Basis eine **Diagonalmatrix** ist, denn dann gilt

$$B_{ij} = B(\vec{e}_i, \vec{e}_j) = \lambda_j \vec{e}_i \cdot \vec{e}_j = \lambda_j \delta_{ij}, \quad (3.2.12)$$

d.h.  $B_{ij} = 0$  für  $i \neq j$ , und  $B_{jj} = \lambda_j$ . Die Matrix  $\overleftrightarrow{B}$  bzgl. einer Basis von Eigenvektoren ist also in der Tat eine Diagonalmatrix, und auf der Diagonalen stehen die Eigenwerte dieser Matrix.

Wir zeigen nun, daß es für *symmetrische Matrizen* stets eine Orthonormalbasis aus Eigenvektoren gibt, d.h. man kann stets eine orthogonale Transformation der ursprünglichen Orthonormalbasis  $(\vec{e}_j)$  zu einer Orthonormalbasis  $(\vec{e}'_j)$  finden, so daß  $\overleftrightarrow{B} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  ist. Man bezeichnet eine solche orthogonale Transformation als **Hauptachsentransformation**. Diese Bezeichnung wird im folgenden noch klar.

Den Beweis führen wir durch vollständige Induktion bzgl. der Dimension  $n$ . Für  $n = 1$  ist nichts zu beweisen, denn dann besteht  $\overleftrightarrow{B}$  einfach aus einer Matrix mit einer Zahl, und folglich ist ein Einheitsvektor in dem eindimensionalen Vektorraum stets Eigenvektor von  $\overleftrightarrow{B}$ . Nehmen wir nun an, die Behauptung sei für einen Vektorraum der Dimension  $n$  korrekt. Dann müssen wir zeigen, daß dies auch für einen Vektorraum der Dimension  $n + 1$  zutrifft. Sei also  $B$  eine Bilinearform auf einem  $(n + 1)$ -dimensionalen Vektorraum. Dann zeigen wir, daß es wenigstens einen Eigenvektor von  $\overleftrightarrow{B}$  gibt. Dazu betrachten wir die **quadratische Form**

$$Q(\vec{x}) = B(\vec{x}, \vec{x}) = \vec{x}^T \overleftrightarrow{B} \vec{x}. \quad (3.2.13)$$

Die Einheitskugelschale, die durch die Nebenbedingung  $N(\vec{x}) = 1 - \vec{x} \cdot \vec{x} = 1 - \vec{x}^T \vec{x} = 0$  definiert ist, ist offenbar eine abgeschlossene und beschränkte Menge von Vektoren, d.h. jede konvergente Folge von Einheitsvektoren konvergiert gegen einen Einheitsvektor (*warum?*). Weiter ist die quadratische Form stetig und beliebig oft stetig partiell differenzierbar (wobei Ableitung einer Ordnung  $\geq 3$  offenbar allesamt verschwinden) und damit auch stetig. Nach dem Satz vom Maximum, der für stetige Funktionen mehrerer Veränderlicher genauso gilt wie für solche einer Veränderlichen (und auch in analoger Weise wie in Anhang A.5 bewiesen werden kann), besitzt  $Q(\vec{x})$  unter der Nebenbedingung  $N(\vec{x}) = 0$  ein Maximum.

Wie in Abschnitt 2.3 gezeigt, muß dann *notwendig* die Funktion

$$g(\vec{x}, \lambda) = Q(\vec{x}) + \lambda(1 - \vec{x}^T \vec{x}) \quad (3.2.14)$$

einen verschwindenden Gradienten besitzen, d.h. es muß ein  $\vec{x} = \vec{e}$  geben, so daß (*nachrechnen!*)

$$\frac{\partial}{\partial x_i} g(\vec{e}, \lambda) = 2 \left( \sum_{j=1}^n B_{ij} e_j - \lambda e_i \right) = 0, \quad \partial_\lambda g(\vec{e}, \lambda) = 1 - \vec{e}^T \vec{e} = 0. \quad (3.2.15)$$

Der erste Satz von Gleichungen für  $i \in \{1, 2, \dots, n + 1\}$  ergibt die Eigenwertgleichung (3.2.6), d.h. es existiert ein Eigenvektor  $\vec{e}$ .

Wir können dann eine Orthonormalbasis  $(\vec{e}_i)_{i \in \{1, 2, \dots, n+1\}}$  mit  $\vec{e}_{n+1} = \vec{e}$  konstruieren. Für die entsprechende darstellende Matrix unserer Bilinearform bzgl. dieser Basis gilt dann aber

$$B_{n+1,j} = B_{j,n+1} = \lambda \delta_{j,n+1}. \quad (3.2.16)$$

Die Matrix hat bzgl. dieser Basis also die Gestalt

$$\overleftrightarrow{B} = \begin{pmatrix} \tilde{B} & \vec{0}_n \\ \vec{0}_n^T & \lambda \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} & 0 \\ B_{21} & B_{22} & \dots & B_{2n} & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ B_{n1} & B_{n2} & \dots & B_{nn} & 0 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix}. \quad (3.2.17)$$

Damit ist das Problem, diese Matrix weiter zu diagonalisieren offenbar auf den  $n$ -dimensionalen Unterraum, der von den  $(\vec{e}_j)_{j \in \{1, 2, \dots, n\}}$  aufgespannt wird, reduziert. Nach Induktionsvoraussetzung kann dies durch Wahl einer Basis aus Eigenvektoren dieses Unterraumes erreicht werden, und damit ist die Behauptung bewiesen:

Für eine symmetrische Matrix gibt es stets eine Orthonormalbasis aus Eigenvektoren.

Nun wenden wir uns der Frage zu, wie man das Problem lösen kann, diese Basis zu finden. Sei nun also wieder  $\overleftrightarrow{B}$  die darstellende Matrix einer Bilinearform im  $n$ -dimensionalen Vektorraum. Nun schreiben wir die Eigenwertgleichung (3.2.6) in der Form

$$(\overleftrightarrow{B} - \lambda \mathbb{1}_n) \vec{e} = \vec{0}. \quad (3.2.18)$$

Damit dieses lineare Gleichungssystem eine Lösung  $\vec{e} \neq \vec{0}$  besitzt, muß die Determinante der in Klammern stehenden Matrix verschwinden, d.h.

$$P_B(\lambda) = \det(\overleftrightarrow{B} - \lambda \mathbb{1}_n) = 0, \quad (3.2.19)$$

weil anderenfalls nach der Cramerschen Regel (1.6.55) das Gleichungssystem (3.2.18) nur die triviale Lösung  $\vec{e} = \vec{0}$  besitzen würde.

Nun ist die durch (3.2.19) gegebene Funktion  $P_B$  ein Polynom  $n$ -ten Grades,

$$P_B(\lambda) = (-1)^n \lambda^n + \dots, \quad (3.2.20)$$

das **charakteristische Polynom** der Matrix  $\overleftrightarrow{B}$ .

Die Eigenwerte dieser Matrix sind also die Nullstellen des charakteristischen Polynoms. Aus dem obigen Beweis zur Hauptachsentransformation wissen wir, daß es stets genügend reelle Lösungen zu dieser Gleichung gibt, um eine Orthonormalbasis von Eigenvektoren zu bestimmen.

Als einfaches *Beispiel* betrachten wir die symmetrische Matrix in  $\mathbb{R}^{2 \times 2}$

$$\overleftrightarrow{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (3.2.21)$$

Um die Hauptachsentransformation zu bewerkstelligen, also eine Orthonormalbasis aus Eigenvektoren zu finden, berechnen wir zuerst das charakteristische Polynom

$$P_B(\lambda) = \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda-2). \quad (3.2.22)$$

Die Eigenwerte der Matrix ergeben sich als Nullstellen dieses Polynoms. Man liest unmittelbar ab, daß dies  $\lambda_1 = 0$  und  $\lambda_2 = 2$  sind.

Die Eigenvektoren findet man dann durch Lösung der entsprechenden Eigenwertgleichungen zu diesen Eigenwerten, also

$$\overleftrightarrow{B} \vec{e}'_1 = 0. \quad (3.2.23)$$

Es ist sofort klar, daß dies für

$$\vec{e}'_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (3.2.24)$$

erfüllt ist, wobei wir den Vektor bereits auf 1 normiert haben. Für den zu  $\lambda_2 = 2$  gehörigen Eigenvektor folgt

$$(\overleftrightarrow{B} - 2\mathbb{1}_2) \vec{e}'_2 = 2\vec{e}'_2. \quad (3.2.25)$$

Dies ist offensichtlich für

$$\vec{e}'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (3.2.26)$$

der Fall. Offenbar ist  $(\vec{e}'_1, \vec{e}'_2)$  in der Tat eine Orthonormalbasis des  $\mathbb{R}^2$  (*nachprüfen!*).

Die Transformationsmatrix  $\hat{T}$  ergibt sich gemäß (3.1.7) durch

$$T_{ij} = \underline{e}'_i \cdot \underline{e}_j, \quad (3.2.27)$$

d.h.

$$\hat{T} = \frac{1}{\sqrt{2}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = (\underline{e}'_1, \underline{e}'_2)^T. \quad (3.2.28)$$

In der Tat erhält man durch einfaches Nachrechnen, daß die darstellende Matrix der Bilinearform  $B$  bzgl. der neuen Basis diagonal ist. Gemäß (3.2.4) ist nämlich

$$\overset{\leftrightarrow}{B} = \hat{T} \overset{\leftrightarrow}{B} \hat{T}^T = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}. \quad (3.2.29)$$

### 3.3 Kegelschnitte im $\mathbb{R}^2$

Kommen wir noch auf den Namen „Hauptachsentransformation“ für die Diagonalisierung der darstellenden Matrix  $\overset{\leftrightarrow}{B}$  einer symmetrischen Bilinearform  $B$ , die sich aus einer geometrischen Betrachtung ergibt.

Betrachten wir zunächst  $\mathbb{R}^2$  und eine beliebige symmetrische Bilinearform  $B \in L_2(\mathbb{R}^2)$  und  $Q(\vec{x}) = B(\vec{x}, \vec{x})$  die dazugehörige quadratische Form. Sei weiter  $\vec{a} \in \mathbb{R}^2$  beliebig vorgeben. Wir betrachten nun Kurven, die implizit durch die Bedingung

$$Q(\vec{x}) + \vec{a} \cdot \vec{x} = 1 \quad (3.3.1)$$

definiert sind.

Wir wollen die Eigenschaften der so definierten Kurven untersuchen. Dies können wir zunächst dadurch vereinfachen, daß wir als Basis des kartesischen Koordinatensystems eine Orthonormalbasis aus Eigenvektoren wählen. Wir setzen also im folgenden voraus, daß  $\overset{\leftrightarrow}{B} = \text{diag}(\lambda_1, \lambda_2)$  gilt. Weiter können wir noch den Koordinatenursprung beliebig verschieben, also die Vektoren  $\vec{x} = \vec{\xi} + \vec{c}$  mit  $\vec{c} = \text{const}$  wählen. In diesem Koordinatensystem ergibt sich für (3.3.1)

$$Q(\vec{x}) + \vec{a} \cdot \vec{x} = \lambda_1(\xi_1 + c_1)^2 + \lambda_2(\xi_2 + c_2)^2 + a_1(\xi_1 + c_1) + a_2(\xi_2 + c_2) = 1. \quad (3.3.2)$$

Wir müssen nun Fallunterscheidungen bzgl. der Eigenwerte  $\lambda_1, \lambda_2$  vornehmen.

$\lambda_1, \lambda_2 \neq 0$ . Dann können wir durch geeignete Wahl von  $\vec{c}$  die in  $\vec{\xi}$  linearen Terme eliminieren. Dazu multiplizieren wir den quadratischen Teil aus und sortieren nach Potenzen der Komponenten von  $\vec{\xi}$ :

$$\begin{aligned} Q(\vec{x}) + \vec{a} \cdot \vec{x} &= \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 \\ &+ (2\lambda_1 c_1 + a_1)\xi_1 + (2\lambda_2 c_2 + a_2)\xi_2 \\ &+ \lambda_1 c_1^2 + \lambda_2 c_2^2 + a_1 c_1 + a_2 c_2. \end{aligned} \quad (3.3.3)$$

Dann können wir offenbar durch die Wahl

$$c_1 = -\frac{a_1}{2\lambda_1}, \quad c_2 = -\frac{a_2}{2\lambda_2} \quad (3.3.4)$$

den in  $\vec{\xi}$  linearen Anteil zum Verschwinden bringen. Dann wird

$$Q(\vec{x}) + \vec{a} \cdot \vec{x} = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 - \frac{a_1^2}{4\lambda_1} - \frac{a_2^2}{4\lambda_2}. \quad (3.3.5)$$

Die Bedingung (3.3.2) lautet dann nach einfachen Umformungen

$$\lambda'_1 \xi_1^2 + \lambda'_2 \xi_2^2 = 1, \quad (3.3.6)$$

wobei

$$\lambda'_j = \lambda_j \left( 1 + \frac{a_1^2}{4\lambda_1} + \frac{a_2^2}{4\lambda_2} \right)^{-1}, \quad j \in \{1, 2\}. \quad (3.3.7)$$

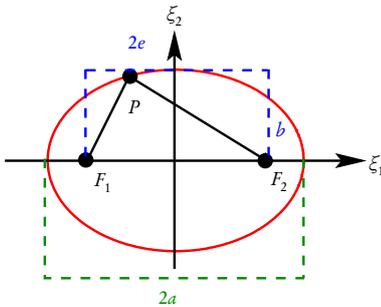
Dabei nehmen wir an, daß die Klammer nicht verschwindet. Diesen Fall betrachten wir unten gesondert.

Als weitere Fallunterscheidung betrachten weiter  $\lambda_1, \lambda_2 > 0$ . Dann haben wir es mit einer geschlossenen Kurve zu tun, die wir durch

$$\xi_1 = \frac{1}{\sqrt{\lambda'_1}} \cos \varphi, \quad \xi_2 = \frac{1}{\sqrt{\lambda'_2}} \sin \varphi, \quad \varphi \in [0, 2\pi) \quad (3.3.8)$$

parametrisieren können. Es handelt sich um eine **Ellipse** mit den Halbachsen  $a = 1/\sqrt{\lambda_1}$  und  $b = 1/\sqrt{\lambda_2}$ . Mit diesen Bezeichnungen wird die Parameterform (3.3.6) zu

$$\left( \frac{\xi_1}{a} \right)^2 + \left( \frac{\xi_2}{b} \right)^2 = 1. \quad (3.3.9)$$



Daß es sich tatsächlich um eine Ellipse handelt, sieht man wie folgt ein: Nehmen wir an, daß  $a \geq b$  sei (andernfalls können wir dies durch Vertauschen der Basisvektoren erreichen). Sei dann der **Brennpunktstrecke** durch  $e = \sqrt{a^2 - b^2}$  definiert. Die Brennpunkte sind dann durch  $\xi_1 = \pm e, \xi_2 = 0$  definiert. Um zu zeigen, daß die durch (3.3.9) beschriebene Kurve tatsächlich eine Ellipse ist, müssen wir zeigen, daß die Summe der Abstände eines Punktes auf der Kurve von den beiden Brennpunkten konstant ist (s. Skizze). In der Tat gilt für Vektoren  $\bar{\xi}$ , die (3.3.9) erfüllen (*nachrechnen!*)

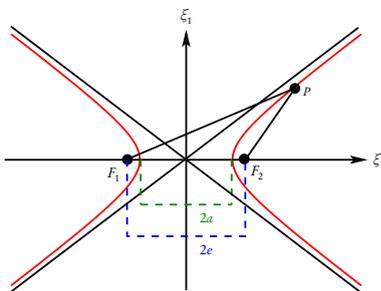
$$|F_1 P| + |F_2 P| = \sqrt{(\xi_1 + e)^2 + \xi_2^2} + \sqrt{(\xi_1 - e)^2 + \xi_2^2} = 2a = \text{const.} \quad (3.3.10)$$

$\lambda_1 > 0$  und  $\lambda_2 < 0$ . Dann setzen wir  $a = 1/\sqrt{\lambda_1}$  und  $b = 1/\sqrt{-\lambda_2}$ . Dann nimmt (3.3.6) die Form

$$\left( \frac{\xi_1}{a} \right)^2 - \left( \frac{\xi_2}{b} \right)^2 = 1 \quad (3.3.11)$$

an. Diese Kurve können wir offensichtlich durch

$$\xi_1 = \pm a \cosh \eta, \quad \xi_2 = b \sinh \eta, \quad \eta \in \mathbb{R} \quad (3.3.12)$$



parametrisieren. Die Kurve zerfällt offenbar in zwei Äste gemäß der unterschiedlichen Wahl der Vorzeichen von  $\xi_1$ . Wir zeigen nun, daß es sich um eine **Hyperbel** handelt. Sei dazu  $e = \sqrt{a^2 + b^2}$ . Dann sind die Brennpunkte durch  $\xi_1 = \pm e, \xi_2 = 0$  gegeben. In der Tat gilt (*nachrechnen!*)

$$||F_1 P| - |F_2 P|| = \left| \sqrt{(\xi_1 + e)^2 + \xi_2^2} - \sqrt{(\xi_1 - e)^2 + \xi_2^2} \right| = 2a. \quad (3.3.13)$$

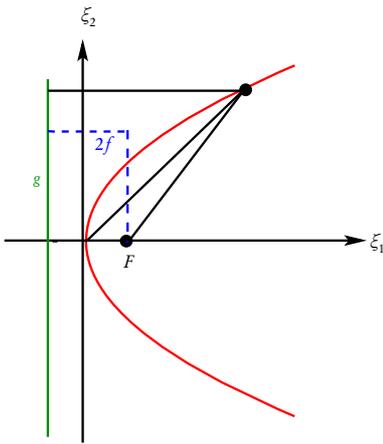
Schließlich betrachten wir den oben ausgeschlossenen Fall, daß die Klammer in (3.3.7) verschwindet. Dann nimmt die Gleichung (3.3.2) die Form

$$\left( \frac{\xi_1}{a} \right)^2 - \left( \frac{\xi_2}{b} \right)^2 = 0 \quad (3.3.14)$$

an. Damit entartet in diesem Fall die Kurve zu den beiden Geraden

$$\xi_2 = \pm \frac{a}{b} \xi_1. \quad (3.3.15)$$

### 3.3 · Kegelschnitte im $\mathbb{R}^2$



Schließlich fehlt noch der Fall  $\lambda_1 = 0, \lambda_2 \neq 0$ . Wenn ein Eigenwert von  $\vec{B}$  verschwindet, können wir stets erreichen, daß dies  $\lambda_1$  ist, indem wir ggf. die Basisvektoren vertauschen. Dann nimmt (3.3.3) die Form

$$Q(\vec{x}) + \vec{a} \cdot \vec{x} = \lambda_2 \xi_2^2 + a_1 \xi_1 + (2\lambda_2 c_2 + a_2) \xi_2 + \lambda_2 c_2^2 + a_2 c_2 \quad (3.3.16)$$

an. Setzen wir nun  $c_2 = -a_2/(2\lambda_2)$ , vereinfacht sich dies zu

$$Q(\vec{x}) + \vec{a} \cdot \vec{x} = \lambda_2 \xi_2^2 + a_1 \xi_1 - \frac{a_2^2}{4\lambda_2} + a_1 c_1. \quad (3.3.17)$$

Die Kurve (3.3.1) können wir dann eindeutig mit  $\xi_2 \in \mathbb{R}$  parametrisieren, und es ergibt sich für  $a_1 \neq 0$  und die Wahl  $c_1 = 1/a_1 + a_2^2/(4a_1\lambda_2)$

$$\xi_1 = \frac{\lambda_2}{a_1} \xi_2^2. \quad (3.3.18)$$

Wir können weiter annehmen, daß  $\lambda_2/a_1 > 0$  ist. Andernfalls müssen wir in der Eigenbasis von  $\vec{B}$  nur  $\vec{e}_1$  durch  $-\vec{e}_1$  ersetzen. Setzen wir nun  $a_1/\lambda_2 = 4f$ , erhalten wir

$$\xi_1 = \frac{\xi_2^2}{4f}. \quad (3.3.19)$$

Es handelt sich um eine **Parabel**. Dazu definieren wir den Brennpunkt als  $F : \xi_1 = f, \xi_2 = 0$  und die Leitgerade  $g : \xi_1 = -f$  und zeigen, daß für jeden Punkt auf der Kurve (3.3.19) der Abstand von der Leitgeraden gleich dem Abstand vom Brennpunkt ist (s. Skizze). In der Tat gilt für den Abstand des Punktes  $P : (\xi_1, \xi_2)$  auf der Kurve vom Brennpunkt

$$|FP| = \sqrt{(\xi_1 - f)^2 + \xi_2^2} = \sqrt{(\xi_1 - f)^2 + 4f\xi_1} = \xi_1 + f, \quad (3.3.20)$$

und dies ist auch der Abstand des Punktes auf der Kurve von der Leitgeraden, und somit ist die Kurve in der Tat eine Parabel. Wir sehen also, daß die Eigenvektoren der Matrix  $\vec{B}$  für den Fall der Ellipse (also beide Eigenwerte positiv) die Richtung der Hauptachsen charakterisieren. Dies erklärt die Bezeichnung „Hauptachsentransformation“ für die Diagonalisierung der darstellenden Matrizen von Bilinearformen (also Tensoren 2. Stufe).

Eine alternative Beschreibung der Kegelschnitte erhält man, indem man den Ursprung des Koordinatensystems in einen der Brennpunkte von Ellipse oder Hyperbel bzw. den Brennpunkt im Falle einer Parabel legt. Diese Darstellung erhält man z.B. bei der üblichen Behandlung des Keplerproblems, denn in der Näherung, daß man die Sonne als festes Zentrum betrachtet, legt man das Koordinatensystem gewöhnlich in den Mittelpunkt der Sonne. Der betrachtete Himmelskörper bewegt sich dann auf einem Kegelschnitt, bei dem die Sonne in einem Brennpunkt liegt. Dann verwenden wir Polarkoordinaten, wobei  $\varphi = 0$  dem **sonnennächsten Punkt (Perihelion)** der Bahn entspricht. Wir wollen die entsprechenden Darstellungen der Kegelschnitte als Funktion  $r = r(\varphi)$  aus den oben besprochenen geometrischen Eigenschaften der Kegelschnitte herleiten.

Beginnen wir mit der **Ellipse**. Gemäß Abbildung 3.1 gilt

$$\vec{F_1P} = 2e\vec{e}_1 + \vec{x}, \quad \vec{F_2P} = \vec{r}. \quad (3.3.21)$$

Um die Polardarstellung der Ellipse zu finden, verwenden wir (3.3.13). Zunächst gilt wegen (3.3.21)

$$|F_1P|^2 = (2e\vec{e}_1 + \vec{x})^2 = 4e^2 + 4er \cos \varphi + r^2, \quad (3.3.22)$$

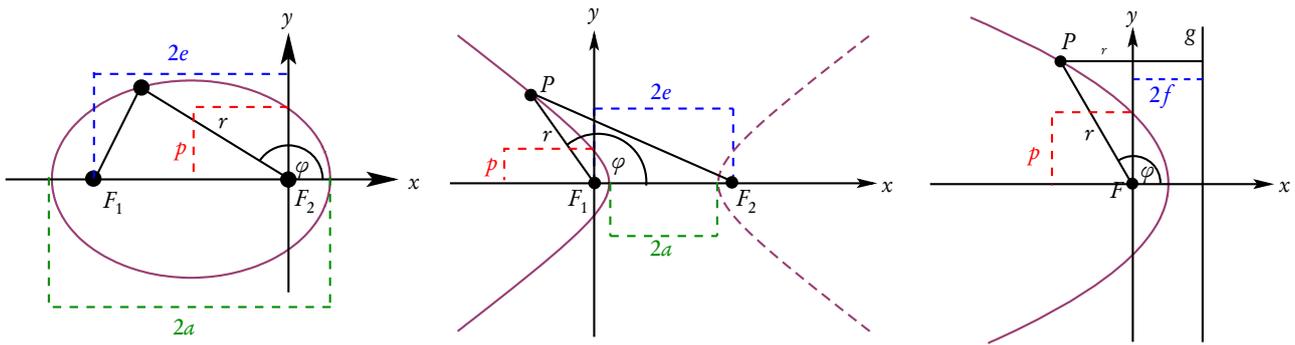


Abbildung 3.1: Zur Herleitung der Kegelschnittgleichung in Polarkoordinaten mit einem Brennpunkt im Koordinatenursprung.

und damit lautet (3.3.13)

$$|F_1P| + |F_2P| = \sqrt{4e^2 + 4er \cos \varphi + r^2} + r = 2a. \quad (3.3.23)$$

Dies lösen wir nach  $r$  auf. Wegen  $a^2 - e^2 = b^2$  ergibt sich

$$(2a - r)^2 = 4e^2 + 4er \cos \varphi + r^2 \Rightarrow b^2 = r(a + e \cos \varphi). \quad (3.3.24)$$

Dies ergibt

$$r(\varphi) = \frac{p}{1 + \epsilon \cos \varphi} \quad \text{mit} \quad \epsilon = \frac{e}{a}, \quad p = \frac{b^2}{a}. \quad (3.3.25)$$

Dabei heißt  $0 \leq \epsilon < 1$  die **exzentrische Anomalie** der Ellipse, denn dieser Parameter gibt die Abweichung vom Fall einer Kreisbahn  $\epsilon = 0$  an, und  $p$  heißt Halbparameter der Ellipse. Dessen geometrische Bedeutung ist in Abb. 3.1 angegeben. In der Tat ergibt sich für  $\varphi = \pi/2$  aus (3.3.25)  $r(\pi/2) = p$ .

Ebenso können wir bei der **Hyperbel** vorgehen. Wir erhalten wieder dieselbe Gleichung (3.3.25) (*nachrechnen!*). Nur ist jetzt  $\epsilon = e/a > 1$ . Da stets  $r > 0$  sein muß, ist der Winkelbereich auf  $|\varphi| \leq \varphi_{\max}$  bestimmt, wobei

$$\epsilon \cos \varphi_{\max} = -1 \quad \varphi_{\max} = -\arccos\left(\frac{1}{\epsilon}\right) \quad (3.3.26)$$

gegeben ist. Freilich ergibt sich nur der linke Ast der Hyperbel. Der rechte wird durch die entsprechende Polargleichung (*nachrechnen!*)

$$r(\varphi) = \frac{p}{1 - \epsilon \cos \varphi} \quad (3.3.27)$$

beschrieben.

Bei der **Parabel** lesen wir aus Abb. 3.1 aufgrund der geometrischen Eigenschaft, daß ein Punkt auf der Parabel vom Brennpunkt gleich weit entfernt ist wie von der Leitgeraden  $g$

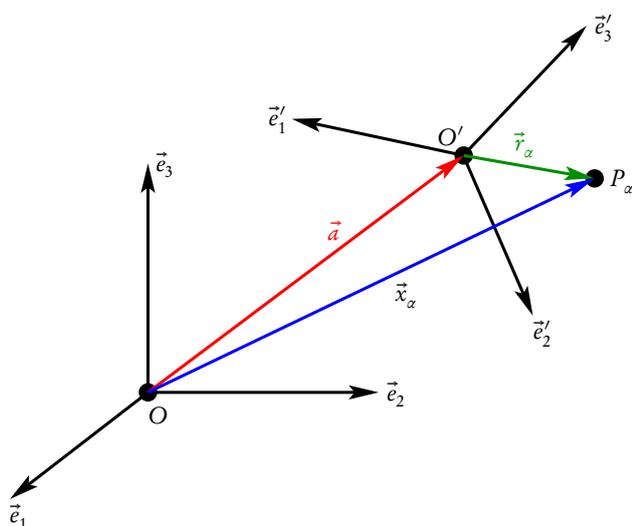
$$2f + r \cos(\pi - \varphi) = r \Rightarrow r(\varphi) = \frac{2f}{1 + \cos \varphi}, \quad (3.3.28)$$

d.h. es gilt wieder die Gleichung (3.3.25), wobei  $\epsilon = 1$  und  $p = 2f$  ist. Daß für die Parabel der Halbparameter gerade  $p = 2f$  ist, folgt auch sofort aus der geometrischen Bedeutung der Parabel (vgl. Abb. 3.1).

### 3.4 Elementare Kreiseltheorie

In diesem Abschnitt zeigen wir anhand der Beispiels der Kinematik des starren Körpers eine Anwendung der bisher behandelten Vektor- und Tensoranalysis.

Wir modellieren den **starrten Körper** als ein System von Partikeln, die durch sehr starke Zentralkräfte zusammengehalten werden, so daß von Deformationen durch außen angreifende Kräfte abgesehen werden kann, d.h. aufgrund der Bindungskräfte ist die relative Lage der Massenpunkte fixiert.



Zunächst ist die Lage von  $N$  Massenpunkten durch  $3N$  kartesische Komponenten der entsprechenden Ortsvektoren relativ zu einem in einem Inertialsystem fixierten Punkt gegeben. Wir bezeichnen dieses Inertialsystem im folgenden als **raumfestes Bezugssystem** und definieren einen Ursprung  $O$  und ein rechtshändiges kartesisches Basissystem  $(\vec{e}_l)$ .

Freilich ist die Angabe aller  $3N$  kartesischen Komponenten für jeden Massenpunkt innerhalb des starren Körpers unnötig, da ja die relative Lage der Massenpunkte festliegt. Um die Lage des starren Körpers vollständig zu charakterisieren, benötigen wir nur die Lage eines bzgl. des Körpers fixierten Punktes  $O'$  und die Orientierung eines starr mit dem Körper verbundenen rechtshändigen kartesischen Basissystems  $(\vec{e}'_k)$ . Dies bezeichnen wir im folgenden als

**körperfestes Bezugssystem**. Wir bezeichnen den Ortsvektor des körperfesten Punktes  $O'$  mit  $\vec{a}$ . Wie wir in Abschnitt 1.3 gesehen haben, sind die beiden rechtshändigen Orthonormalbasen des raum- und körperfesten Bezugssystems durch eine orthogonale Matrix  $\hat{D}$  mit  $\hat{D}^{-1} = \hat{D}^T$  und  $\det \hat{D} = 1$  miteinander verknüpft:

$$\vec{e}'_k = \sum_{j=1}^3 D_{lk} \vec{e}_l \quad \text{und} \quad \vec{e}_l = \sum_{k=1}^3 D_{lk} \vec{e}'_k. \quad (3.4.1)$$

Wie wir unten sehen werden, ist es zweckmäßig, diese Drehung durch die drei **Euler-Winkel**  $(\psi, \theta, \phi)$  wie in Abschnitt 1.7.3 beschrieben zu parametrisieren.

Folglich besitzt der starre Körper genau 6 voneinander unabhängige Freiheitsgrade, nämlich die drei Komponenten  $\vec{a}$  des Ortsvektors des körperfesten Bezugspunktes  $O'$  und die drei Euler-Winkel der Drehmatrix  $\hat{D}$ . Die Ortsvektoren der einzelnen Massenpunkte bzgl. des Raumfesten Bezugspunkt  $O$  bezeichnen wir mit  $\vec{x}_\alpha$  ( $\alpha \in \{1, \dots, N\}$ ) und relativ zum körperfesten Bezugspunkt  $O'$  mit  $\vec{r}_\alpha$ . Dann gilt

$$\vec{x}_\alpha = \vec{a} + \vec{r}_\alpha. \quad (3.4.2)$$

Um die Dynamik des starren Körpers zu beschreiben, benötigen wir nun oft die Umrechnung von raumfesten Koordinaten  $\vec{V}$  und körperfesten Koordinaten  $\vec{V}'$  eines beliebigen Vektors

$$\vec{V} = \sum_{k=1}^3 V'_k \vec{e}'_k = \sum_{l=1}^3 V_l \vec{e}_l. \quad (3.4.3)$$

Bzgl. der Zeitabhängigkeit ist zu beachten, daß die  $\vec{e}'_k$  zeitabhängig und die  $\vec{e}_l$  zeitunabhängig sind. Mit (3.4.1) folgt

$$\vec{V} = \hat{D} \vec{V}' \quad \text{und} \quad \vec{V}' = \hat{D}^T \vec{V}, \quad (3.4.4)$$

und es gilt für die Zeitableitung der raumfesten Vektorkomponenten

$$\dot{\bar{V}} = \dot{\hat{D}}\bar{V}' + \hat{D}\dot{\bar{V}}' = \dot{\hat{D}}\hat{D}^T\bar{V} + \hat{D}\dot{\bar{V}}'. \quad (3.4.5)$$

Nun bemerken wir, daß aus der Orthogonalitätsbeziehung

$$\hat{D}\hat{D}^T = \mathbb{1}_3 \quad (3.4.6)$$

durch Ableitung nach der Zeit

$$\dot{\hat{D}}\hat{D}^T + \hat{D}\dot{\hat{D}}^T = 0 \Rightarrow \dot{\hat{D}}\hat{D}^T = -\hat{D}\dot{\hat{D}}^T = -(\dot{\hat{D}}\hat{D}^T)^T \quad (3.4.7)$$

folgt. Das bedeutet, daß

$$\hat{\Omega} = \dot{\hat{D}}\hat{D}^T = -\hat{\Omega}^T \quad (3.4.8)$$

ist, d.h. daß  $\hat{\Omega}$  eine **antisymmetrische Matrix** ist. Wir können sie folglich mit Hilfe des Levi-Civita-Symbols durch Vektorkomponenten  $\bar{\omega}$  parametrisieren gemäß

$$\Omega_{jk} = -\sum_{l=1}^3 \epsilon_{jkl} \omega_l = +\sum_{l=1}^3 \omega_{jlk} \omega_l, \quad (3.4.9)$$

wobei die Wahl des Vorzeichens Konvention ist. In Matrixform geschrieben, lautet diese Gleichung

$$\hat{\Omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \quad (3.4.10)$$

Dann gilt für den ersten Term in (3.4.5)

$$\left(\dot{\hat{D}}\hat{D}^T\bar{V}\right)_j = \left(\hat{\Omega}\bar{V}\right)_j = \sum_{k=1}^3 \Omega_{jk} V_k = \sum_{k,l=1}^3 \epsilon_{jlk} \omega_l V_k = (\bar{\omega} \times \bar{V})_j. \quad (3.4.11)$$

Da sich die Komponenten eines Kreuzprodukts wie Vektorkomponenten transformieren, gilt

$$\bar{\omega} \times \bar{V} = \hat{D}(\bar{\omega}' \times \bar{V}'). \quad (3.4.12)$$

Wir können also

$$\dot{\bar{V}} = \hat{D}(\dot{\bar{V}}' + \bar{\omega}' \times \bar{V}') \quad (3.4.13)$$

schreiben. Es gilt demnach

$$\dot{\bar{V}} = \sum_{j=1}^3 \dot{V}_j \vec{e}_j = \sum_{j,k=1}^3 \vec{e}_j D_{jk} (\dot{\bar{V}}' + \bar{\omega}' \times \bar{V}')_k = \sum_{k=1}^3 \vec{e}'_k (\dot{\bar{V}}' + \bar{\omega}' \times \bar{V}')_k. \quad (3.4.14)$$

Die Komponenten von  $\dot{\bar{V}}$  bzgl. des körperfesten Bezugssystems weist also einen Zusatzterm  $\bar{\omega}' \times \bar{V}'$  aufgrund der Zeitabhängigkeit der körperfesten Basis ( $\vec{e}'_j$ ) auf, die von der zeitabhängigen Rotation dieser Basis gegen die zeitabhängige raumartige Basis herrührt. Entsprechend ist  $\bar{\omega}$  die **momentane Winkelgeschwindigkeit** dieser Drehung, d.h. die Richtung dieses Vektors gibt die Richtung der Drehachse im Sinne der Rechten-Hand-Regel an und sein Betrag die momentane Winkelgeschwindigkeit der Drehung.

Definieren wir nun die Geschwindigkeit der Relativvektoren  $\vec{r}_\alpha$  als  $\vec{v}_\alpha^{(\text{rel})} = \dot{\vec{r}}_\alpha$  gilt wegen  $\vec{r}'_\alpha = \text{const}$ , d.h.  $\dot{\vec{r}}'_\alpha = 0$  aufgrund von (3.4.15)

$$\vec{v}_\alpha^{(\text{rel})'} = \vec{\omega}' \times \vec{r}'_\alpha. \quad (3.4.15)$$

Nun wollen wir die gesamte **kinetische Energie** eines starren Körpers berechnen. Zunächst gilt definitionsgemäß

$$T = \sum_{\alpha=1}^N \frac{m_\alpha}{2} \dot{\vec{x}}_\alpha^2 = \sum_{\alpha=1}^N \frac{m_\alpha}{2} (\dot{\vec{a}} + \dot{\vec{r}}_\alpha)^2 = \sum_{\alpha=1}^N \frac{m_\alpha}{2} (\dot{\vec{a}}^2 + 2\dot{\vec{a}} \cdot \dot{\vec{r}}_\alpha + \dot{\vec{r}}_\alpha^2) \quad (3.4.16)$$

Der erste und der letzte Term lassen sich leicht interpretieren: Der erste Term ist die **kinetische Translationsenergie** aufgrund der momentanen Translationsgeschwindigkeit  $\dot{\vec{a}}$  des Körpers, also die Änderung der Lage des körperfesten Bezugspunktes  $O'$  gegenüber der Lage des raumfesten Bezugspunktes  $O$ . Der dritte Term ist die kinetische **Rotationsenergie**, die aus der Rotation der Massenpunkte, aus denen wir uns den Körper zusammengesetzt denken, um den körperfesten Bezugspunkt  $O'$  resultiert. Diese Eigendrehung des Körpers um einen körperfesten Bezugspunkt bezeichnen wir auch als **Spin**.

Der mittlere Term ist schwieriger zu interpretieren. Es handelt sich um eine Art „Spin-Bahn-Kopplung“. Wir wollen nun das körperfeste Bezugssystem so einfach wie möglich wählen und betrachten dazu zunächst den mittleren Term. Hier kommt der Vektor

$$\vec{r}_S = \frac{1}{M} \sum_{\alpha=1}^N m_\alpha \vec{r}_\alpha \quad \text{mit} \quad M = \sum_{\alpha=1}^N m_\alpha, \quad (3.4.17)$$

also der Ortsvektor vom körperfesten Bezugspunkt  $O'$  zum **Schwerpunkt** des Körpers bzw. die **Schwerpunktsgeschwindigkeit**

$$\vec{v}_S^{(\text{rel})} = \dot{\vec{r}}_S \quad (3.4.18)$$

vor. Wir können nun bequemerweise den körperfesten Bezugspunkt  $O'$  in den Schwerpunkt des Körpers legen. Dann wird

$$\vec{r}_S = 0, \quad \vec{v}_S^{(\text{rel})} = 0, \quad \vec{a} = \dot{\vec{x}}_S = \frac{1}{M} \sum_{\alpha=1}^N m_\alpha \ddot{\vec{x}}_\alpha. \quad (3.4.19)$$

Im folgenden nehmen wir an, daß wir  $O'$  entsprechend in den Schwerpunkt des Körpers gelegt haben, und die kinetische Energie (3.4.17) vereinfacht sich zu

$$T = T_{\text{trans}} + T_{\text{spin}} = \frac{M}{2} \dot{\vec{x}}_S^2 + \sum_{\alpha=1}^N \frac{m_\alpha}{2} \dot{\vec{r}}_\alpha^2. \quad (3.4.20)$$

Es deutet sich hier bereits an, daß sich die Dynamik der Translationsbewegung von der Dynamik der Rotationsbewegung zumindest für den Fall entkoppelt, d.h. die Translationsbewegung hängt nur von den raumfesten Komponenten des Schwerpunktsvektors  $\vec{x}_S$  und die der Spinbewegung nur von den Komponenten von  $\vec{r}_\alpha$ , d.h. nur von den Eulerwinkeln, die die Drehung zwischen raum- und körperfestem Basissystem parametrisieren.

Der Spinanteil der kinetischen Energie läßt sich demnach offenbar am einfachsten mit den körperfesten Komponenten beschreiben. Wegen (3.4.15) gilt

$$T_{\text{rot}} = \sum_{\alpha=1}^N \frac{m_\alpha}{2} (\vec{\omega}' \times \vec{r}'_\alpha)^2 = \frac{1}{2} \sum_{j,k=1}^3 \omega'_j \omega'_k \sum_{\alpha=1}^N \frac{m_\alpha}{2} (\vec{r}'_\alpha^2 \delta_{jk} - r'_{\alpha,j} r'_{\alpha,k}). \quad (3.4.21)$$

Dabei ergibt sich die letztere Form wie folgt. Zunächst dürfen wir „Punkt und Kreuz in einem Spatprodukt vertauschen“, d.h. es gilt

$$(\vec{\omega}' \times \vec{r}'_\alpha)^2 = (\vec{\omega}' \times \vec{r}'_\alpha) \cdot (\vec{\omega}' \times \vec{r}'_\alpha) = \vec{\omega}' \cdot [\vec{r}'_\alpha \times (\vec{\omega}' \times \vec{r}'_\alpha)]. \quad (3.4.22)$$

Mit Hilfe der „BAC-CAB-Formel“ folgt weiter

$$(\vec{\omega}' \times \vec{r}'_\alpha)^2 = \vec{\omega}' \cdot [\vec{\omega}' \vec{r}'_\alpha{}^2 - \vec{r}'_\alpha (\vec{\omega}' \cdot \vec{r}'_\alpha)] = \vec{\omega}'^2 \vec{r}'_\alpha{}^2 - (\vec{\omega}' \cdot \vec{r}'_\alpha). \quad (3.4.23)$$

Schreibt man dies in die Indexschreibweise um, erhält man folglich in der Tat (3.4.21).

Definieren wir nun die raumfesten Komponenten des **Trägheitstensors** durch

$$\Theta'_{jk} = \sum_{\alpha=1}^N m_\alpha (\vec{r}'_\alpha{}^2 \delta_{jk} - r'_{\alpha,j} r'_{\alpha,k}), \quad (3.4.24)$$

folgt kurz

$$T_{\text{spin}} = \frac{1}{2} \vec{\omega}'^T \overset{\leftarrow}{\Theta} \vec{\omega}'. \quad (3.4.25)$$

Da die  $\vec{r}'_\alpha$  zeitlich konstant sind, sind auch die raumfesten Komponenten des Trägheitstensors räumlich konstant. Außerdem folgt aus der Definition von  $T_{\text{spin}}$ , daß es sich um einen positiv definiten symmetrischen Tensor zweiter Stufe handelt, denn es ist offenbar stets  $T_{\text{spin}} \geq 0$ .

Nun können wir wegen des Theorems von der Hauptachsentransformation die Orientierung der körperfesten Basis  $(\vec{e}'_k)$  stets so wählen, daß  $\overset{\leftarrow}{\Theta} = \text{diag}(A, B, C)$  mit den drei **Hauptträgheitsmomenten**  $A, B, C > 0$  gilt. Die entsprechende Orthonormalbasis besteht aus den **Hauptträgheitsachsen** des Körpers. Diese Richtungen sowie die Werte der Hauptträgheitsmomente ist allein durch die relative Lage der Massenpunkte, die den Körper bilden sowie deren Massen bestimmt. Im folgenden nehmen wir an, daß wir die körperfeste Orthonormalbasis bereits durch die Hauptträgheitsachsen definiert haben. Damit haben wir die einfachst mögliche Wahl des körperfesten Bezugspunktes zu  $O' = S'$  und die Wahl der körperfesten Basis gefunden.

Wir wollen nun die Bewegung des starren Körpers für den einfachsten Fall, daß keine äußeren Kräfte auf ihn einwirken, beschreiben. Dazu betrachten wir die Newtonschen Bewegungsgleichungen für die **Bindungskräfte**, die den Körper zusammenhalten:

- (i) Die gesamte Kraft  $\vec{F}_\alpha$ , die aufgrund der Wechselwirkung des Massenpunktes am Ort  $P_\alpha$  mit den übrigen Massenpunkten  $P_\beta$  ( $\beta \neq \alpha$ ) auf  $P_\alpha$  ausgeübt werden, ist durch die Summe von **Paarwechselwirkungen** gegeben:

$$\vec{F}_\alpha = \sum_{\beta \neq \alpha} \vec{F}_{\alpha,\beta}. \quad (3.4.26)$$

Aufgrund des **3. Newtonschen Postulats** („**actio=reactio**“) gilt

$$\vec{F}_{\alpha,\beta} = -\vec{F}_{\beta,\alpha}. \quad (3.4.27)$$

- (ii) Die Paarwechselwirkungskräfte sind Zentralkräfte, wirken also stets in Richtung der Verbindungslinie zwischen den Massenpunkten (in unserem Fall stark anziehend, denn der Körper soll ja starr sein, d.h. alle betrachteten äußeren Kräfte sollen klein gegen die Wechselwirkungskräfte sein, so daß der Körper durch sie als nicht deformiert betrachtet werden kann):

$$\vec{F}_{\alpha,\beta} \propto \vec{x}_\alpha - \vec{x}_\beta = \vec{r}_\alpha - \vec{r}_\beta. \quad (3.4.28)$$

Nach dem **2. Newtonschen Postulat** gilt für den  $\alpha$ -ten Massenpunkt die Bewegungsgleichung

$$m_\alpha \ddot{\vec{x}}_\alpha = m_\alpha \dot{\vec{v}}_\alpha = m_\alpha \dot{\vec{v}}_S + m_\alpha \dot{\vec{v}}_\alpha^{(\text{rel})} = \sum_{\beta \neq \alpha} \vec{F}_{\alpha,\beta}. \quad (3.4.29)$$

Wir müssen nun aber nicht all diese Bewegungsgleichungen lösen. Aufgrund unserer Annahme, daß es sich um einen starren Körper handelt brauchen wir aufgrund unserer obigen Überlegungen ja nur Bewegungsgleichungen für den Schwerpunktsvektor und die drei Eulerwinkel der Drehung, die die Orientierung der körperfesten Hauptträgheitsachsen des Körpers bzgl. der raumfesten Basis parametrisiert. Wegen (3.4.17-3.4.19) empfiehlt es sich offenbar (3.4.29) auch noch über alle Massenpunkte  $P_\alpha$  zu summieren. Wegen (3.4.27) verschwindet die entsprechende Summe über alle Kräfte. Demnach gilt also

$$\sum_{\alpha=1}^N m_\alpha \ddot{\vec{x}}_\alpha = M \ddot{\vec{x}}_S = M \dot{\vec{v}}_S = \vec{0}. \quad (3.4.30)$$

Die Schwerpunktsbewegung ist also geradlinig gleichförmig und unabhängig von der Spinbewegung des Körpers um seinen Schwerpunkt. Durch einmalige Integration von (3.4.30) nach der Zeit folgt der **Impulserhaltungssatz**

$$\vec{P} = \sum_{\alpha=1}^N m_\alpha \dot{\vec{x}}_\alpha = \sum_{\alpha=1}^N m_\alpha \vec{v}_\alpha = M \vec{v}_S = \text{const.} \quad (3.4.31)$$

Die Schwerpunktsbewegung ist also durch

$$\vec{x}_S(t) = \frac{1}{M} \vec{P} t + \vec{x}_{S,0} \quad (3.4.32)$$

vollständig beschrieben. Nun können wir den Ursprung des raumfesten Bezugssystems offenbar so wählen, daß

$$\vec{x}_{S,0} = 0 \quad (3.4.33)$$

ist, d.h. der raumfeste Bezugspunkt  $O$  wird in den Schwerpunkt des starren Körpers zur Zeit  $t = 0$  gelegt.

Um nun auch eine Bewegungsgleichung für die Spinbewegung, also die Eigenrotation des Körpers um seinen Schwerpunkt, zu finden, betrachten wir den **Gesamtdrehimpuls** des starren Körpers, also die Größe

$$\begin{aligned} \vec{J} &= \sum_{\alpha=1}^N m_\alpha \vec{x}_\alpha \times \vec{v}_\alpha = \sum_{\alpha=1}^N m_\alpha (\vec{x}_S + \vec{r}_\alpha) \times (\vec{v}_S + \vec{v}_\alpha^{(\text{rel})}) \\ &= M \vec{x}_S \times \vec{v}_S + \vec{x}_S \times \sum_{\alpha} m_\alpha \vec{v}_S^{(\text{rel})} + \left( \sum_{\alpha=1}^N m_\alpha \vec{r}_\alpha \right) \times \vec{v}_S + \sum_{\alpha=1}^N m_\alpha \vec{r}_\alpha \times \vec{v}_\alpha^{(\text{rel})}. \end{aligned} \quad (3.4.34)$$

In dem Ausdruck verschwinden nun wegen (3.4.32) und (3.4.33) der erste und wegen (3.4.17-3.4.19) auch der zweite und der dritte Term. In diesem Bezugssystem ist also der Gesamtdrehimpuls des Körpers identisch mit dem Spin des Körpers:

$$\vec{J} = \sum_{\alpha=1}^N m_\alpha \vec{r}_\alpha \times \vec{v}_\alpha^{(\text{rel})}. \quad (3.4.35)$$

Um nun eine Bewegungsgleichung für diese Drehbewegung, also für die Euler-Winkel zu gewinnen, bilden wir von dieser Beziehung die Zeitableitung. Wegen  $\dot{\vec{r}}_\alpha = \vec{v}_\alpha^{(\text{rel})}$  folgt

$$\dot{\vec{J}} = \sum_{\alpha=1}^N m_\alpha \vec{r}_\alpha \times \dot{\vec{v}}_\alpha^{(\text{rel})} = \sum_{\alpha=1}^n \vec{r}_\alpha \times \vec{F}_\alpha = \vec{\tau}. \quad (3.4.36)$$

Die Vektorgröße  $\vec{\tau}$  ist das **Gesamtdrehmoment bzgl. des Schwerpunktes**. Im letzten Schritt haben wir (3.4.30) verwendet, also  $\dot{\vec{v}}_\alpha^{(\text{rel})} = \dot{\vec{v}}_\alpha + \dot{\vec{v}}_S = \dot{\vec{v}}_\alpha$ . Aus (3.4.26 und (3.4.27) folgt nun aber

$$\dot{\vec{J}} = \sum_{\alpha=1}^N \sum_{\beta \neq \alpha} \vec{r}_\alpha \times \vec{F}_\beta = \sum_{\alpha, \beta; \alpha \neq \beta} \vec{r}_\alpha \times \vec{F}_{\alpha, \beta}. \quad (3.4.37)$$

Wir können nun in der letzteren Summe die Summationsindizes  $\alpha$  und  $\beta$  vertauschen. Dann folgt mit (3.4.28) schließlich

$$\dot{\vec{J}} = \sum_{\alpha, \beta; \alpha \neq \beta} \vec{r}_\beta \times \vec{F}_{\beta, \alpha} = - \sum_{\alpha, \beta; \alpha \neq \beta} \vec{r}_\alpha \times \vec{F}_{\alpha, \beta}. \quad (3.4.38)$$

Addieren wir nun (3.4.38) zu (3.4.37) und dividieren das Resultat zu (3.4.37), folgt schließlich

$$\dot{\vec{J}} = \vec{0}. \quad (3.4.39)$$

Dies ist der **Drehimpulserhaltungssatz**: Für ein geschlossenes System von Massenpunkten, die einzig durch zentrale Paarwechselwirkungskräfte miteinander interagieren, ist der Drehimpuls zeitlich konstant<sup>1</sup>.

Die Bewegungsgleichungen für die Spinbewegung des Körpers um seinen Schwerpunkt leiten wir nun in zwei Schritten her: Zunächst drücken wir  $\vec{J}$  und die Erhaltungsgleichung (3.4.39) in körperfesten Koordinaten aus. Wie wir gleich sehen werden, ergeben sich daraus Bewegungsgleichungen für die körperfesten Komponenten  $\vec{\omega}'$  der Winkelgeschwindigkeit, die **Eulerschen Kreiselgleichungen**, die man für den Spezialfall des **symmetrischen Kreisels**<sup>2</sup> sogar geschlossen mit elementaren Funktionen lösen kann. Im zweiten Schritt leiten wir dann die Beziehung zwischen den körperfesten Komponenten  $\vec{\omega}'$  und den Zeitableitungen der Eulerwinkel her, die wir im Falle des symmetrischen Kreisels durch einfaches Integrieren ebenfalls geschlossen als Funktion der Zeit angeben können. Da auch die raumfeste Drei-Achse ausgezeichnet ist, legen wir dazu die Orientierung des raumfesten Koordinatensystems so fest, daß für den erhaltenen Drehimpuls  $\vec{J} = (0, 0, J)^T$  mit  $J > 0$  gilt.

Wir führen dieses Programm nun Schritt für Schritt aus. Setzen wir also (3.4.16) in (3.4.35) ein, erhalten wir für die körperfesten Komponenten des Drehimpulses

$$\vec{J}' = \sum_{\alpha=1}^N m_\alpha \vec{r}'_\alpha \times (\vec{\omega}' \times \vec{r}'_\alpha) = \sum_{\alpha=1}^N m_\alpha \left[ (\vec{r}'_\alpha)^2 \vec{\omega}' - \vec{r}'_\alpha (\vec{r}'_\alpha \cdot \vec{\omega}') \right]. \quad (3.4.40)$$

Im letzten Schritt haben wir wieder die „BAC-CAP-Formel“ für das doppelte Kreuzprodukt verwendet. Betrachten wir die Definition der körperfesten Komponenten des Trägheitstensors (3.4.24), zeigt sich, daß wir diese Beziehung in der kompakten Form

$$\vec{J}' = \overset{\leftarrow}{\Theta} \vec{\omega}' \quad (3.4.41)$$

schreiben können. Mit der allgemeinen Formel (3.4.14) für die körperfesten Komponenten der Zeitableitung eines Vektors lautet die Drehimpulserhaltungsgleichung (3.4.39)

$$\dot{\vec{J}}' + \vec{\omega}' \times \vec{J}' = 0. \quad (3.4.42)$$

Setzt man hierin (3.4.41) ein und weil  $\overset{\leftarrow}{\Theta} = \text{diag}(A, B, C) = \text{const}$  ist lautet diese Gleichung in Komponenten ausgeschrieben (*nachrechnen!*)

$$\begin{pmatrix} A\dot{\omega}'_1 \\ B\dot{\omega}'_2 \\ C\dot{\omega}'_3 \end{pmatrix} = \begin{pmatrix} (B-C)\omega'_2\omega'_3 \\ (C-A)\omega'_3\omega'_1 \\ (A-B)\omega'_1\omega'_2 \end{pmatrix}. \quad (3.4.43)$$

<sup>1</sup>Wie die obigen Herleitungen des Impuls- und Drehimpulserhaltungssatzes zeigen, gelten diese allgemein, d.h. auch wenn der Körper nicht starr ist, also für beliebige Relativbewegungen der Körper untereinander, solange die Wechselwirkungskräfte zwischen den Massenpunkten zentrale Paarwechselwirkungen sind.

<sup>2</sup>Ein symmetrischer Kreisler liegt definitionsgemäß dann vor, wenn von den drei Hauptträgheitsmomenten  $A, B$  und  $C$  mindestens zwei miteinander übereinstimmen. Da bei der Einführung der Eulerwinkel die körperfeste Drei-Achse eine ausgezeichnete Rolle einnimmt, wählen wir die entsprechende Symmetrieachse des Körpers, die sog. **Figurenachse**, in Richtung von  $\vec{e}'_3$ . Dann gilt  $A = B$ . Falls sogar  $A = B = C$  gilt, liegt ein sog. **Kugelkreisler** vor. Wir werden aber nur  $A = B$  voraussetzen. Wir bemerken noch, daß sich diese Symmetrien auf den Trägheitstensor bezieht. Der tatsächliche Kreislerkörper muß je nach der Massenverteilung nicht entsprechend symmetrisch sein. Wenn der Körper allerdings homogen ist, d.h. eine räumlich konstante Dichte aufweist, entsprechen die geometrischen Symmetrieachsen des Körpers den Hauptträgheitsachsen.

Dies sind die **Eulerschen Kreiselgleichungen**. Dieses nichtlineare Differentialgleichungssystem läßt sich i.a. nicht analytisch lösen.

Wir betrachten daher den oben erläuterten Spezialfall des **symmetrischen Kreisels**, für den  $A = B$  gilt. Dann vereinfachen sich die Eulerschen Kreiselgleichungen zu

$$\begin{pmatrix} A\dot{\omega}'_1 \\ A\dot{\omega}'_2 \\ C\dot{\omega}'_3 \end{pmatrix} = \begin{pmatrix} (A-C)\omega'_2\omega'_3 \\ (C-A)\omega'_3\omega'_1 \\ 0 \end{pmatrix}. \quad (3.4.44)$$

Aufgrund der dritten Komponente folgt, daß  $\omega'_3 = \text{const.}$  Die Komponente der Winkelgeschwindigkeit um die körperfeste 3'-Achse, die **Figurenachse** des Kreisels, ist also konstant. Wir können die Richtung der Dreiachse stets so wählen, daß  $\omega'_3 > 0$  ist.

Wegen  $\omega'_3 = \text{const.}$  stellen die ersten beiden Komponenten des Differentialgleichungssystems (3.4.44) effektiv ein *lineares* System dar und lassen sich folglich geschlossen lösen. Wir setzen zur Abkürzung  $\alpha = (C-A)\omega_3/A$ . Dann lauten die ersten beiden Gleichungen in (3.4.44)

$$\dot{\omega}'_1 = -\alpha\omega'_2, \quad \dot{\omega}'_2 = \alpha\omega'_1. \quad (3.4.45)$$

Leiten wir nun die erste Gleichung nach der Zeit ab und verwenden im Resultat die zweite Gleichung, ergibt sich

$$\ddot{\omega}'_1 = -\alpha^2\omega'_1. \quad (3.4.46)$$

Dies ist die Bewegungsgleichung für einen harmonischen Oszillator. Die Lösungen solcher linearer Differentialgleichungen mit konstanten Koeffizienten werden in Kapitel 5.2 ausführlich besprochen. Demnach lautet die allgemeine Lösung

$$\omega'_1(t) = \omega'_\perp \cos(\alpha t + \delta), \quad \omega'_\perp, \delta = \text{const.} \quad (3.4.47)$$

Mit der ersten Gleichung in (3.4.45) erhalten wir

$$\omega'_2(t) = -\frac{1}{\alpha}\dot{\omega}'_1(t) = \omega'_\perp \sin(\alpha t + \delta). \quad (3.4.48)$$

Insgesamt ist also

$$\vec{\omega}' = \begin{pmatrix} \omega'_\perp \cos(\alpha t + \delta) \\ \omega'_\perp \sin(\alpha t + \delta) \\ \omega'_3 \end{pmatrix} \quad (3.4.49)$$

und

$$\vec{J}' = \overset{\leftrightarrow}{\Theta}' \vec{\omega}' = \begin{pmatrix} A\omega'_\perp \cos(\alpha t + \delta) \\ A\omega'_\perp \sin(\alpha t + \delta) \\ C\omega'_3 \end{pmatrix}. \quad (3.4.50)$$

Aus (3.4.49) und (3.4.50) folgt, daß die Figurenachse, Winkelgeschwindigkeit und der Drehimpuls stets in einer Ebene liegen, denn durch direktes Nachrechnen findet man

$$\vec{f}' \cdot (\vec{\omega}' \times \vec{J}') = \vec{f}' \cdot (\vec{\omega}' \times \vec{J}') = 0. \quad (3.4.51)$$

Außerdem präzedieren, vom körperfesten Bezugssystem aus betrachtet, die Winkelgeschwindigkeit und der Drehimpuls mit derselben Kreisfrequenz gerade Kreiskegel um die Figurenachse, und zwar für  $\alpha > 0$  gegen und für  $\alpha < 0$  im Uhrzeigersinn. Der zu  $\vec{\omega}'$  gehörige Kegel heißt **Polkegel**. Der entsprechende zeitlich konstante Winkel zur Figurenachse  $\vec{f}' = \vec{e}'_3$  berechnet sich zu

$$\cos \angle(\vec{\omega}', \vec{e}'_3) = \frac{\omega'_3}{\sqrt{\omega'^2_\perp + \omega'^2_3}}. \quad (3.4.52)$$

Der Betrag der Winkelgeschwindigkeit  $\omega = |\vec{\omega}| = |\vec{\omega}'| = \sqrt{\omega_3'^2 + \omega_\perp'^2} = \text{const.}$  Ebenso sind die Winkel zwischen Winkelgeschwindigkeit und Drehimpuls und zwischen Figurenachse und Drehimpuls zeitlich konstant:

$$\begin{aligned}\cos \angle(\vec{\omega}, \vec{J}) &= \frac{\vec{\omega} \cdot \vec{J}}{\omega |\vec{J}|} = \frac{A\omega_\perp'^2 + C\omega_3'^2}{\omega \sqrt{A^2\omega_\perp'^2 + C^2\omega_3'^2}}, \\ \cos \angle(\vec{f}, \vec{J}) &= \frac{J_3'}{|\vec{J}'|} = \frac{C\omega_3'}{\sqrt{A^2\omega_\perp'^2 + C^2\omega_3'^2}} = \frac{\omega_3'}{\sqrt{(A/C)^2\omega_\perp'^2 + \omega_3'^2}}.\end{aligned}\quad (3.4.53)$$

Da der Cosinus im hier relevanten Winkelbereich zwischen 0 und  $\pi/2^3$  monoton fallend ist, gilt

$$\begin{aligned}\text{oblater Kreisel: } A < C &\Rightarrow \angle(\vec{\omega}, \vec{f}) < \angle(\vec{f}, \vec{J}), \\ \text{prolater Kreisel: } A > C &\Rightarrow \angle(\vec{\omega}, \vec{f}) > \angle(\vec{f}, \vec{J}).\end{aligned}\quad (3.4.54)$$

Um nun die Bewegung weiter zu untersuchen und vor allem ein Bild im raumfesten Bezugssystem zu gewinnen, das ja der Beobachtungssituation bei den üblichen Kreiselexperimenten entspricht, müssen wir im zweiten Schritt unserer Rechnung die Euler-Winkel einführen.

Dazu gehen wir von der expliziten Form (1.7.42) aus. Wir berechnen zunächst mit (3.4.8) die Komponenten der Winkelgeschwindigkeit bzgl. der raumfesten Koordinaten. Nach einer recht umfangreichen Rechnung (*Übung!*), die die Bildung der Zeitableitungen von (1.7.42) und die Matrixmultiplikation in (3.4.8) verlangt, ergibt sich

$$\hat{\Omega} = \begin{pmatrix} 0 & -\dot{\psi} - \dot{\varphi} \cos \vartheta & \dot{\vartheta} \sin \psi - \dot{\varphi} \sin \vartheta \cos \psi \\ \dot{\psi} + \dot{\varphi} \cos \vartheta & 0 & -\dot{\vartheta} \cos \psi - \dot{\varphi} \sin \vartheta \sin \psi \\ -\dot{\vartheta} \sin \psi + \dot{\varphi} \sin \vartheta \cos \psi & \dot{\vartheta} \cos \psi + \dot{\varphi} \sin \vartheta \sin \psi & 0 \end{pmatrix}, \quad (3.4.55)$$

und der Vergleich mit (3.4.10) liefert schließlich die raumfesten Komponenten der Winkelgeschwindigkeit

$$\vec{\omega} = \begin{pmatrix} \dot{\vartheta} \cos \psi + \dot{\varphi} \sin \vartheta \sin \psi \\ \dot{\vartheta} \sin \psi - \dot{\varphi} \sin \vartheta \cos \psi \\ \dot{\psi} + \dot{\varphi} \cos \vartheta \end{pmatrix}. \quad (3.4.56)$$

Die körperfesten Komponenten erhalten wir vermöge der allgemeinen Transformationsformel (3.4.4) zu

$$\vec{\omega}' = \hat{D}^t \vec{\omega} = \begin{pmatrix} \dot{\vartheta} \cos \varphi + \dot{\psi} \sin \vartheta \sin \varphi \\ -\dot{\vartheta} \sin \varphi + \dot{\psi} \sin \vartheta \cos \varphi \\ \dot{\varphi} + \dot{\psi} \cos \vartheta \end{pmatrix}. \quad (3.4.57)$$

Weiter gilt für die raumfesten Komponenten des Drehimpulses wegen (3.4.39)  $\dot{\vec{J}} = 0$ , und wir wählen die Orientierung der raumfesten Orthonormalbasis so, daß  $\vec{J} = (0, 0, J)^T$ . Dann gilt wegen (3.4.4) und (3.4.41)

$$\vec{J}' = \hat{D}^T \vec{J} = J \begin{pmatrix} \sin \varphi \sin \vartheta \\ \cos \varphi \sin \vartheta \\ \cos \vartheta \end{pmatrix} = \begin{pmatrix} A\omega_1' \\ A\omega_2' \\ C\omega_3' \end{pmatrix} = \text{const.} \quad (3.4.58)$$

<sup>3</sup>Wir erinnern uns, daß wir oben die Orientierung der körperfesten 3'-Achse so gewählt hatten, daß  $\omega_3' > 0$  ist und folglich die drei Cosinus in (3.4.52) und (3.4.53) positiv sind.

Oben haben wir gesehen, daß aufgrund der Eulerschen Kreiselgleichungen für den symmetrischen Kreisel  $\omega'_3 = \text{const}$  ist. Da auch  $J = \text{const}$  ist, muß wegen der dritten Komponente der Gleichung (3.4.58) auch  $\vartheta = \text{const}$  sein. Damit vereinfacht sich (3.4.57) zu

$$\bar{\omega}' = \begin{pmatrix} \dot{\psi} \sin \vartheta \sin \varphi \\ \dot{\psi} \sin \vartheta \cos \varphi \\ \dot{\varphi} + \dot{\psi} \cos \vartheta \end{pmatrix} = \begin{pmatrix} \omega'_\perp \cos(\alpha t + \delta) \\ \omega'_\perp \sin(\alpha t + \delta) \\ \omega'_3 \end{pmatrix}. \quad (3.4.59)$$

Dabei haben wir im letzten Schritt die oben gefundenen Lösungen (3.4.47) und (3.4.48) eingesetzt. Aus den beiden ersten Komponenten von (3.4.60) folgt, daß

$$\omega'^2_\perp = \dot{\psi}^2 \sin^2 \vartheta = \text{const} \quad (3.4.60)$$

ist. Wegen

$$\vartheta(t) = \vartheta_0 = \text{const} \quad (3.4.61)$$

ist demzufolge  $\dot{\psi} = \Omega_1 = \text{const}$  und damit

$$\psi(t) = \Omega_1 t + \psi_0. \quad (3.4.62)$$

. Aus der dritten Komponente von (3.4.59) folgt damit weiter

$$\varphi(t) = (\omega'_3 - \Omega_1 \cos \vartheta)t + \varphi_0 = \Omega_2 t + \varphi_0. \quad (3.4.63)$$

Durch Einsetzen all dieser Beziehungen in (3.4.59) und Vergleich der Komponenten folgt

$$\Omega_1 \sin \vartheta = \omega'_\perp, \quad \varphi(t) = \Omega_2 t + \varphi_0 = \frac{\pi}{2} - \alpha t - \delta \Rightarrow \Omega_2 = -\alpha, \quad \varphi_0 = \frac{\pi}{2} - \delta. \quad (3.4.64)$$

Damit erhalten wir schließlich für die raumfesten Komponenten der Winkelgeschwindigkeit

$$\bar{\omega} = \begin{pmatrix} \Omega_2 \sin \vartheta_0 \sin(\Omega_1 t + \psi_0) \\ -\Omega_2 \sin \vartheta_0 \cos(\Omega_1 t + \psi_0) \\ \Omega_1 + \Omega_2 \sin \vartheta_0 \end{pmatrix}. \quad (3.4.65)$$

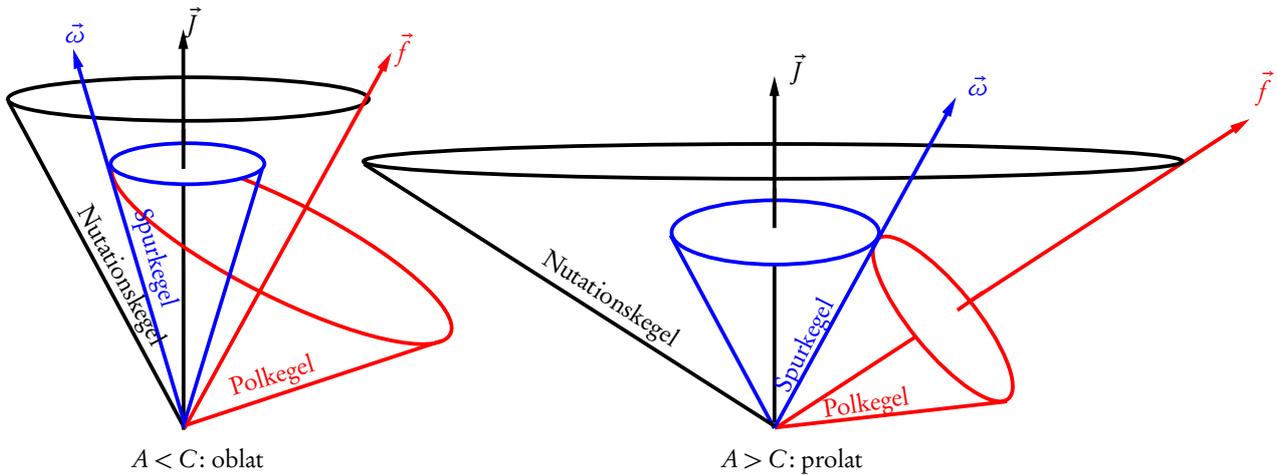
Vom raumfesten Bezugssystem aus betrachtet bewegt sich also die Winkelgeschwindigkeit auf einem Kreiskegel, dem sog. **Spurkegel**, um die raumfeste 3-Achse, also um die Richtung des konstanten Gesamtdrehimpulses, und zwar mit der Kreisfrequenz  $\Omega_1$ .

Schließlich gilt für den Richtungsvektor der Figurenachse  $\vec{f} = \vec{e}'_3$  bzgl. des raumfesten Systems

$$\vec{f} = \hat{D}^T \vec{f}' = \hat{D}^T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin \vartheta_0 \sin(\Omega_1 t + \psi_0) \\ -\sin \vartheta_0 \cos(\Omega_1 t + \psi_0) \\ \cos \vartheta_0 \end{pmatrix}. \quad (3.4.66)$$

Wie auch aus der geometrischen Bedeutung der Euler-Winkel hervorgeht, ist  $\vartheta_0$  der Winkel zwischen Figuren- und Drehimpulsachse. Die Figurenachse präzediert wie die Winkelgeschwindigkeit vom raumfesten System aus betrachtet mit derselben Kreisfrequenz  $\Omega_1$  um die raumfeste Drehimpulsachse. Der entsprechende Kegel heißt **Nutationskegel**.

Unter Berücksichtigung der bereits oben in (3.4.54) bestimmten Winkelverhältnisse ergibt sich also das folgende Bild der Bewegung des kräftefreien symmetrischen Kreisels:



Vom raumfesten Bezugssystem aus betrachtet präzedieren die Figurenachse (Nutationskegel) und die Winkelgeschwindigkeit (Spurkegel) mit derselben Kreisfrequenz  $\Omega_1$  um den konstanten Drehimpuls. Diese drei Vektoren liegen dabei stets in einer Ebene, die folglich ebenfalls mit der Kreisfrequenz  $\Omega_1$  um die Drehimpulsachse rotiert. Vom raumfesten Bezugssystem aus betrachtet präzediert die Winkelgeschwindigkeit mit der Kreisfrequenz  $\alpha = -\Omega_2$  um die Figurenachse (Polkegel). Vom raumfesten System aus betrachtet rollt also der Polkegel um den Spurkegel, wobei letzterer für einen oblaten (prolaten) Kreisel innerhalb (außerhalb) des Polkegels liegt, wie in der obigen Abbildung gezeigt.

### 3.5 Allgemein kovariante Tensorschreibweise

Wir können nun ohne großen Aufwand in Fortführung von Abschnitt 2.14 auch die Tensoren bzw. Tensorfelder allgemeiner Stufe bzgl. beliebiger verallgemeinerter („krummliniger“) Koordinaten schreiben. Wir verwenden in diesem Abschnitt wieder ausschließlich die *holonomen ko- und kontravarianten Basisvektoren* (2.14.1) und (2.14.22) sowie die *Einsteinsche Summationskonvention*. Einen Tensor  $n$ -ter Stufe können wir dann sowohl durch seine ko- bzw. kontravarianten Komponenten charakterisieren. Mit der Tensorproduktschreibweise gilt z.B. für einen Tensor dritter Stufe

$$T = T_{ijk} \vec{g}^i \otimes \vec{g}^j \otimes \vec{g}^k = T^{ijk} \vec{g}_i \otimes \vec{g}_j \otimes \vec{g}_k = T_j^{i \ k} \vec{g}_i \otimes \vec{g}^j \otimes \vec{g}_k \quad \text{etc.} \quad (3.5.1)$$

Wir können durch Ableitung von Tensorfeldern mit Hilfe von (2.14.28) und (2.14.37) ein neues Tensorfeld einer um eins höheren Stufe definieren, denn offenbar ist für ein Tensorfeld  $T$

$$A = \vec{\nabla} \otimes T := \vec{g}^l \otimes \partial_l T = \vec{g}^l \otimes \partial_l (T_{ijk} \vec{g}^i \otimes \vec{g}^j) \quad (3.5.2)$$

wieder ein Tensor. Dies macht man sich anhand des Transformationsverhaltens unter allgemeinen Koordinatentransformationen genauso klar, wie mit der Rechnung (2.14.26) für den Fall eines Vektors (der im allgemeinen Tensorkalkül auch ein Tensor 1. Stufe ist).

Man kann nun (3.5.2) mit Hilfe von (2.14.37) durch die **kovariante Ableitung** der Tensorkomponenten ausdrücken. Offenbar muß man dazu auf die Tensorprodukte der Basisvektoren einfach die Produktregel anwenden. Man erhält für das Beispiel eines Tensors 3. Stufe (*nachrechnen!*)

$$A_{lijk} = \nabla_l T_{ijk} = \partial_l T_{ijk} - \Gamma^a_{li} T_{ajk} - \Gamma^a_{lj} T_{iak} - \Gamma^a_{lk} T_{ija}. \quad (3.5.3)$$

Die Metrik ist nun auch ein spezieller Tensor 2. Stufe:

$$\vec{g} = g_{ij} \vec{g}^i \otimes \vec{g}^j = g^{ij} \vec{g}_i \otimes \vec{g}_j. \quad (3.5.4)$$

Nun sind bzgl. kartesischen Koordinaten  $q'^k = x^k$  die Metrik-Komponenten durch  $g'_{ij} = \delta_{ij} = \text{const}$  gegeben. Gemäß (2.14.41) verschwinden bzgl. der kartesischen Koordinaten die Christoffelsymbole, und es gilt  $\nabla'_k g'_{ij} = \partial'_k g'_{ij} = 0$ . Demnach folgt für allgemeine Tensorkomponenten bzgl. einer holonomen Basis

$$\nabla_k g_{ij} = 0. \quad (3.5.5)$$

Das rechnen wir auch mit Hilfe der Christoffel-Symbole unter Verwendung von (2.14.41) direkt nach:

$$\begin{aligned} \nabla_k g_{ij} &= \partial_k g_{ij} - \Gamma^a_{ki} g_{aj} - \Gamma^a_{kj} g_{ia} \\ &= \partial_k g_{ij} - \Gamma_{jki} - \Gamma_{ikj} \\ &= \partial_k g_{ij} - \frac{1}{2}(\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ik}) - \frac{1}{2}(\partial_k g_{ij} + \partial_j g_{ik} - \partial_i g_{jk}) \\ &= 0. \end{aligned} \quad (3.5.6)$$

Weiter ist klar, daß wir Tensoren **verjüngen** können, denn offenbar führt das Gleichsetzen eines hoch- und eines tiefgestellten Index' und Summation über diesen Index in den Komponenten eines Tensors  $n$ -ter Stufe zu Komponenten eines Tensors  $(n-2)$ -ter Stufe, z.B. sind  $V_j = T^i_{ij} = g^{ik} T_{ikj}$  kovariante Vektorkomponenten eines Vektors.

Ebenso kann man aus einem Tensor  $n$ -ter Stufe und einem Tensor  $m$ -ter Stufe vermöge  $(n+m)$ -ter Stufe definieren: (**Tensorprodukt**):

$$S \otimes T = S_{ijk\dots} T_{lmn\dots} \vec{g}^i \otimes \vec{g}^j \otimes \vec{g}^k \otimes \dots \otimes \vec{g}^l \otimes \vec{g}^m \otimes \vec{g}^n \otimes \dots \quad (3.5.7)$$

Aus dieser Sichtweise kann man die Rotation eines Vektorfeldes zunächst als antisymmetrisierte kovariante Ableitung, die zu einem Tensorfeld 2. Stufe führt, definieren. Dabei ist wegen (2.14.50) die Verwendung kovarianter Komponenten besonders bequem

$$T_{jk} = \nabla_j V_k - \nabla_k V_j = \partial_j V_k - \partial_k V_j. \quad (3.5.8)$$

Die Rotation ergibt sich dann durch das Tensorprodukt mit dem Levi-Civita-Tensor und geeigneter Verjüngung gemäß

$$(\text{rot } \vec{V})^i = \frac{1}{2} \Delta^{ijk} T_{jk} = \Delta^{ijk} \partial_j V_k. \quad (3.5.9)$$

Zum Schluß bemerken wir noch, daß all diese Betrachtungen sich auf endlichdimensionale Vektorräume beliebiger Dimension ausdehnen lassen. Dort spielen dann die vollständig antisymmetrischen Tensoren eine besondere Rolle. In einem  $d$ -dimensionalen Raum gibt es offensichtlich vollständig antisymmetrische Tensoren bis zu  $d$ -ter Stufe, und aus einem vollständig antisymmetrischen Tensor  $n < d$ -ter Stufe  $T$  kann man mit Hilfe des Levi-Civita-Tensors einen vollständig antisymmetrischen Tensor  $(d-n)$ -ter Stufe  ${}^\dagger T$  über seine Komponenten definieren, das sog. **Hodge-Dual** des Tensors  $T$ :

$$({}^\dagger T)^{i_1 \dots i_{d-n+1} \dots i_d} = \frac{1}{n!} \Delta^{i_1 i_2 \dots i_d} T_{i_1 \dots i_n}. \quad (3.5.10)$$

Man kann zeigen, daß zweimaliges Hodge-Dualisieren zum ursprünglichen Tensor zurückführt, d.h. es gilt

$${}^{\dagger\dagger} T = (-1)^{n(d-n)} T. \quad (3.5.11)$$

Zum Beweis verwenden wir, daß für die Einheitsmatrix  $\det \mathbb{1}_d = 1$  ist. Durch Vertauschen zweier Spalten oder Zeilen wechselt die Determinante jeweils das Vorzeichen. Wir können also schreiben

$$\epsilon^{i_1 \dots i_d} = \det \begin{pmatrix} \delta_1^{i_1} & \delta_2^{i_1} & \dots & \delta_d^{i_1} \\ \vdots & \vdots & \dots & \vdots \\ \delta_1^{i_d} & \delta_2^{i_d} & \dots & \delta_d^{i_d} \end{pmatrix}. \quad (3.5.12)$$

Dann folgt aus der Produktregel für Determinanten

$$\epsilon_{j_1 \dots j_d} \epsilon^{i_1 \dots i_d} = \det \left[ \begin{pmatrix} \delta_1^{i_1} & \delta_2^{i_1} & \dots & \delta_d^{i_1} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_1^{i_d} & \delta_2^{i_d} & \dots & \delta_d^{i_d} \end{pmatrix} \begin{pmatrix} \delta_{j_1}^1 & \delta_{j_2}^1 & \dots & \delta_{j_d}^1 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{j_1}^d & \delta_{j_2}^d & \dots & \delta_{j_d}^d \end{pmatrix} \right] = \det \begin{pmatrix} \delta_{j_1}^{i_1} & \delta_{j_1}^{i_2} & \dots & \delta_{j_1}^{i_d} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_1} & \delta_{j_d}^{i_2} & \dots & \delta_{j_d}^{i_d} \end{pmatrix}. \quad (3.5.13)$$

Entwickeln wir die Determinante nach der ersten Spalte, ergibt sich nach dem Determinantenentwicklungssatz

$$\epsilon_{j_1 \dots j_d} \epsilon^{i_1 \dots i_d} = \delta_{j_1}^{i_1} \det \begin{pmatrix} \delta_{j_2}^{i_2} & \dots & \delta_{j_2}^{i_d} \\ \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_2} & \dots & \delta_{j_d}^{i_d} \end{pmatrix} - \delta_{j_2}^{i_1} \det \begin{pmatrix} \delta_{j_1}^{i_2} & \dots & \delta_{j_1}^{i_d} \\ \delta_{j_3}^{i_2} & \dots & \delta_{j_d}^{i_d} \\ \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_2} & \dots & \delta_{j_d}^{i_d} \end{pmatrix} + \dots \quad (3.5.14)$$

Setzt man nun  $i_1 = j_1$  und summiert entsprechend der Einsteinschen Summenkonvention über den nunmehr doppelt auftretenden Index, erhält man für den ersten Term  $\delta_{j_1}^{j_1} = d$ . Die übrigen Terme ergeben dann  $(d-1)$ -mal das Negative der Determinante im ersten Term, d.h. wir haben insgesamt

$$\epsilon_{j_1 \dots j_d} \epsilon^{i_1 \dots i_d} = \det \begin{pmatrix} \delta_{j_2}^{i_2} & \dots & \delta_{j_2}^{i_d} \\ \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_2} & \dots & \delta_{j_d}^{i_d} \end{pmatrix}. \quad (3.5.15)$$

So können wir sukzessive verfahren, um die Verjüngung bzgl. der ersten  $n$  Indizes auszuführen. Wir erhalten dann

$$\epsilon_{i_1 \dots i_n j_{n+1} \dots j_d} \epsilon^{i_1 \dots i_d} = n! \det \begin{pmatrix} \delta_{j_{n+1}}^{i_{n+1}} & \dots & \delta_{j_{n+1}}^{i_d} \\ \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_{n+1}} & \dots & \delta_{j_d}^{i_d} \end{pmatrix}. \quad (3.5.16)$$

Mit (3.5.10) erhalten wir damit

$$\begin{aligned} (\dagger\dagger T)_{j_1 \dots j_n} &= \frac{1}{(d-n)!} \Delta_{j_{n+1} \dots j_d j_1 \dots j_n} (\dagger T)^{j_{n+1} \dots j_d} \\ &= \frac{1}{n!(d-n)!} \Delta_{j_{n+1} \dots j_d j_1 \dots j_n} \Delta^{i_1 \dots i_n j_{n+1} \dots j_d} T_{i_1 \dots i_n} \\ &= \frac{(-1)^{n(d-n)}}{n!(n-d)!} \epsilon_{j_{n+1} \dots j_d j_1 \dots j_n} \epsilon^{j_{n+1} \dots j_d i_1 \dots i_n} T_{i_1 \dots i_n} \\ &= \frac{(-1)^{n(d-n)}}{n!} \det \begin{pmatrix} \delta_{j_1}^{i_1} & \dots & \delta_{j_1}^{i_d} \\ \vdots & \ddots & \vdots \\ \delta_{j_d}^{i_1} & \dots & \delta_{j_d}^{i_d} \end{pmatrix} T_{i_1 \dots i_n} \\ &= (-1)^{n(d-n)} T_{j_1 \dots j_n} \Rightarrow \dagger\dagger T = (-1)^{n(d-n)} T, \end{aligned} \quad (3.5.17)$$

und das wollten wir zeigen.

Wir bemerken noch, daß die kovarianten Ableitungen vollständig antisymmetrischer Tensoren besondere Eigenschaften besitzen. Das hängt damit zusammen, daß die Christoffel-Symbole in den unteren Indizes symmetrisch sind. Mit (3.5.3) folgt

$$\nabla_{j_{n+1}} T_{j_1 \dots j_n} = \partial_{j_{n+1}} T_{j_1 \dots j_n} - \Gamma_{j_{n+1} j_1}^{i_1} T_{i_1 j_2 \dots j_n} - \Gamma_{j_{n+1} j_2}^{i_2} T_{j_1 i_2 j_3 \dots j_n} - \dots \quad (3.5.18)$$

Kontrahieren wir dies mit dem Levi-Civita-Tensor, erhalten wir Tensorkomponenten, die nur kovariante Ableitungen enthalten, denn die Kontraktion über die beiden Indizes, die in den Christoffel-Symbolen in (3.5.18) verschwindet, weil die Christoffel-Symbole bzgl. Vertauschens der beiden unteren Indizes symmetrisch, aber die Komponenten des Levi-Civita-Tensors jeweils

$$S^{j_{n+2}\dots j_d} = \frac{1}{n!} \Delta^{j_1\dots j_d} \nabla_{j_{n+1}} T_{j_1\dots j_n} = \frac{1}{n!} \Delta^{j_1\dots j_d} \partial_{j_{n+1}} T_{j_1\dots j_n}. \quad (3.5.19)$$

Hodge-Dualisieren liefert einen vollständig antisymmetrischen Tensor  $(n+1)$ -ter Stufe mit den Komponenten

$$\begin{aligned} (dT)_{j_1\dots j_{n+1}} &= \frac{(-1)^{(n+1)(d-n-1)}}{(d-n-1)!} \Delta_{j_1\dots j_d} S^{j_{n+2}\dots j_d} = \frac{(-1)^{(n+1)(d-n-1)}}{(d-n-1)!} \epsilon_{j_1\dots j_d} \epsilon^{i_1\dots i_n i_{n+1} j_{n+2}\dots j_d} \partial_{i_{n+1}} T_{i_1\dots i_n} \\ &= \frac{1}{n!} \det \begin{pmatrix} \delta_{j_1}^{i_1} & \dots & \delta_{j_1}^{i_{n+1}} \\ \vdots & \ddots & \vdots \\ \delta_{j_{n+1}}^{i_1} & \dots & \delta_{j_{n+1}}^{i_{n+1}} \end{pmatrix} \partial_{i_{n+1}} T_{i_1\dots i_n} \\ &= \partial_{j_1} T_{j_2\dots j_{n+1}} - \partial_{j_2} T_{j_1 j_3\dots j_{n+1}} + \partial_{j_3} T_{j_1 j_2 j_4\dots j_{n+1}} - \dots + (-1)^{n+1} \partial_{j_{n+1}} T_{j_1\dots j_n}. \end{aligned} \quad (3.5.20)$$

Man nennt dieses Konstrukt eine **alternierende Differentialform**.

Wir können nun weiter invariante Integrale über beliebige **vollständig antisymmetrische Tensorfelder** definieren. Dazu definiert man für  $n \leq d$  zunächst  $n$ -dimensionale **Untermannigfaltigkeiten oder Hyperflächen** des  $d$ -dimensionalen Euklidischen Raums durch eine Parametrisierung  $q^j = q^j(u^1, \dots, u^n)$ , wobei  $(u^k) \in G \subseteq \mathbb{R}^n$ . Dann wird das Integral eines Tensorfeldes über die betreffende Hyperfläche durch

$$\int_{\mathcal{F}} d^n F \cdot T = \int_{\mathcal{F}} d^n F^{j_1\dots j_n} T_{j_1\dots j_n} = \frac{1}{n!} \int_G d^n u \det \left( \frac{\partial(q^{j_1}, \dots, q^{j_n})}{\partial(u^1, \dots, u^n)} \right) T_{j_1\dots j_n}. \quad (3.5.21)$$

Wir können nun den **Satz von Stokes** auf beliebige Differentialformen erweitern. Es sei dazu  $\mathcal{F}$  eine  $n$ -dimensionale Hyperfläche mit dem Rand  $\partial \mathcal{F}$  und  $T_{j_1\dots j_{n-1}}$  ein vollständig antisymmetrisches Tensorfeld  $(n-1)$ -ster Stufe. Dann gilt bei geeigneter Orientierung der Koordinaten der **Satz von Stokes**

$$\int_{\mathcal{F}} d^n F \cdot (dT) = \int_{\partial \mathcal{F}} d^{n-1} F \cdot T. \quad (3.5.22)$$

Wir wollen diesen Satz nur für den einfachsten Fall beweisen, daß  $\mathcal{F}$  auf einem Hyperquader  $Q = [a^1, b^1] \times [a^2, b^2] \times \dots \times [a^n, b^n]$  im Raum der Parameter  $(u^k)$  definierbar ist. Man kann dann beliebige Gebiete durch immer feinmaschigere Netze solcher Hyperquader beliebig genau approximieren und den Satz auf allgemeinere Fälle verallgemeinern. Der Rand  $\partial \mathcal{F}$  ist dann durch die  $2n$  Hyperquader  $u^k = a^k = \text{const}$  und  $u^k = b^k = \text{const}$  gegeben. Nun ist gemäß (3.5.21) und (3.5.20)

$$\begin{aligned} \int_{\mathcal{F}} d^n F \cdot (dT) &= \int_Q d^n u \frac{1}{(n-1)!} \det \left[ \begin{pmatrix} \frac{\partial q^{j_1}}{\partial u^1} & \dots & \frac{\partial q^{j_1}}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial q^{j_n}}{\partial u^1} & \dots & \frac{\partial q^{j_n}}{\partial u^n} \end{pmatrix} \begin{pmatrix} \delta_{j_1}^{i_1} & \dots & \delta_{j_1}^{i_n} \\ \vdots & \ddots & \vdots \\ \delta_{j_n}^{i_1} & \dots & \delta_{j_n}^{i_n} \end{pmatrix} \right] \partial_{i_n} T_{i_1\dots i_{n-1}} \\ &= \int_Q d^n u \frac{1}{(n-1)!} \det \begin{pmatrix} \frac{\partial q^{j_1}}{\partial u^1} & \dots & \frac{\partial q^{j_1}}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial q^{j_n}}{\partial u^1} & \dots & \frac{\partial q^{j_n}}{\partial u^n} \end{pmatrix} \partial_{j_n} T_{j_1\dots j_{n-1}}. \end{aligned} \quad (3.5.23)$$

Entwickeln der Determinante nach der letzten Zeile ergibt dann  $n$  Integrale. Das erste lautet

$$\begin{aligned}
 I_1 &= \frac{1}{(n-1)!} \int_Q d^n u \det \begin{pmatrix} \frac{\partial q^{j_1}}{\partial u^2} & \cdots & \frac{\partial q^{j_1}}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial q^{j_{n-1}}}{\partial u^2} & \cdots & \frac{\partial q^{j_{n-1}}}{\partial u^n} \end{pmatrix} \frac{\partial q^{j_n}}{\partial u^1} \partial_{j_n} T_{j_2 \cdots j_{n-1}} \\
 &= \frac{1}{(n-1)!} \int_Q d^n u \det \begin{pmatrix} \frac{\partial q^{j_1}}{\partial u^2} & \cdots & \frac{\partial q^{j_1}}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial q^{j_{n-1}}}{\partial u^2} & \cdots & \frac{\partial q^{j_{n-1}}}{\partial u^n} \end{pmatrix} \frac{\partial}{\partial u^1} T_{j_2 \cdots j_{n-1}}[q(u)].
 \end{aligned} \tag{3.5.24}$$

Wir können nun die Integration über  $u^1$  ausführen und erhalten

$$\begin{aligned}
 I_1 &= \frac{1}{(n-1)!} \int_{Q_1} dq^2 \cdots dq^n \det \begin{pmatrix} \frac{\partial q^{j_1}}{\partial u^2} & \cdots & \frac{\partial q^{j_1}}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial q^{j_{n-1}}}{\partial u^2} & \cdots & \frac{\partial q^{j_{n-1}}}{\partial u^n} \end{pmatrix} \\
 &\quad \times \left\{ T_{j_2 \cdots j_{n-1}}[q(q^1 = b^1, q^2, \dots, q^n)] - T_{j_2 \cdots j_{n-1}}[q(q^1 = a^1, q^2, \dots, q^n)] \right\}
 \end{aligned} \tag{3.5.25}$$

Dabei ist der  $(n-1)$ -dimensionale Quader  $Q_1 = (a^2, b^2) \times [a^3, b^3] \times \cdots \times [a^n, b^n]$ . Ebenso kann man mit den übrigen Integralen  $I_2, \dots, I_n$  verfahren und erhält schließlich

$$\sum_{j=1}^n I_j = \int_{\partial \mathcal{F}} d^{n-1} F \cdot T, \tag{3.5.26}$$

und das war zu zeigen.

## Kapitel 4

---

# Komplexe Zahlen

Die Erweiterung von den reellen zu den **komplexen Zahlen** ist durch die Forderung nach der Lösbarkeit von **Polynomgleichungen** motiviert. Während sich reelle lineare Gleichungen der Form  $ax + b = 0$  für  $a \neq 0$  noch im Rahmen der reellen Zahlen eindeutig lösen lassen, denn offenbar wird die obige Gleichung dann durch  $x = -b/a$  (eindeutig) gelöst, ist dies schon für quadratische Gleichungen nicht mehr der Fall. Dies wird durch die Einführung der **imaginären Einheit** behoben, wie wir gleich im nächsten Abschnitt sehen werden. Zugleich bilden die komplexen in Analogie zu den reellen Zahlen algebraisch gesehen einen **Zahlenkörper**, und man kann Konvergenzfragen ebenfalls vollkommen analog behandeln wie die von reellen Zahlen, und die komplexen Zahlen sind bzgl. Grenzwertbildung ebenso abgeschlossen wie die reellen Zahlen. In diesem Skript werden wir nur die wichtigsten algebraischen Eigenschaften der komplexen Zahlen begründen und die wichtigsten **elementaren Funktionen** über ihre Potenzreihen und die Bildung von Umkehrfunktionen definieren. Die eigentliche **Funktionentheorie** werden wir hier nicht behandeln. Der interessierte Leser sei dazu auf die Literatur verwiesen, z.B. auf das Skript [CH10].

### 4.1 Definition der komplexen Zahlen

Bei der Lösung quadratischer Gleichungen der Form

$$x^2 + px + q = 0 \tag{4.1.1}$$

stoßen wir auf das Problem, daß für  $x \in \mathbb{R}$  stets  $x^2 \geq 0$  gilt, d.h. im Rahmen der reellen Zahlen können wir keine Quadratwurzeln aus negativen Zahlen ziehen. Die Lösungsstrategie für die Gleichung (4.1.1) besteht darin, eine **quadratische Ergänzung** auszuführen. Offenbar gilt nämlich

$$x^2 + px + q = \left(x + \frac{p}{2}\right)^2 - \frac{p^2}{4} + q. \tag{4.1.2}$$

Die Gleichung (4.1.1) ist also äquivalent zu der Gleichung

$$\left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q. \tag{4.1.3}$$

Wollen wir diese Gleichung nach  $x$  auflösen, müssen wir die Wurzel aus der rechten Seite ziehen können. Im Bereich der reellen Zahlen ist das offensichtlich nur möglich, wenn  $p^2/4 - q \geq 0$  ist. Dann besitzt die Gleichung entweder eine (falls  $p^2/4 - q = 0$  ist) oder (für  $p^2/4 - q > 0$ ) zwei Lösungen. Dies schreiben wir dann kurz als

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \tag{4.1.4}$$

Wir versuchen nun die reellen Zahlen einfach dadurch zu erweitern, daß wir eine neue zunächst rein symbolisch zu verstehende „Zahl“  $i$ , die **imaginäre Einheit**, einführen, für die

$$i^2 = -1 \quad (4.1.5)$$

gelten soll. Dann hätte für  $a > 0$  die Gleichung  $x^2 = -a$  die beiden Lösungen  $x = \pm i\sqrt{a}$ , wobei wir voraussetzen, daß die **komplexen Zahlen**, die allgemein von der Form

$$z = x + iy, \quad x, y \in \mathbb{R} \quad (4.1.6)$$

sein sollen, die gewöhnlichen Rechenregeln wie für reelle Zahlen gelten, also die sogenannten Axiome eines **Zahlenkörpers** erfüllen. Dabei soll eine komplexe Zahl definitionsgemäß durch ihren **Real- und Imaginärteil**  $x$  bzw.  $y$  eindeutig bestimmt sein. Wir schreiben

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y. \quad (4.1.7)$$

Nehmen wir dies an, so folgt für die **Addition** zweier komplexer Zahlen

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2), \quad (4.1.8)$$

wobei wir mehrfach das Assoziativ-, Kommutativ- und Distributivgesetz verwendet haben und wir  $i$  wie eine gewöhnliche Variable behandelt haben. Die Menge aller komplexen Zahlen nennen wir  $\mathbb{C}$ .

Die Regel für die Multiplikation folgt ebenso durch formales Ausmultiplizieren:

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2). \quad (4.1.9)$$

Dabei haben wir im zweiten Term des Realteils die definierende Eigenschaft (4.1.5) der imaginären Einheit benutzt.

Wir berechnen gleich noch die Potenzen von  $i$ :

$$i^2 := -1, \quad i^3 = (i^2)i = -i, \quad i^4 = (i^2)(i^2) = 1 \cdot 1 = 1, \dots \quad (4.1.10)$$

Als weitere Operation an einer einzelnen komplexen Zahl ist noch die **komplexe Konjugation** nützlich. Sie ist so definiert, daß die konjugiert komplexe Zahl  $z^*$  von  $z$  denselben Real- und den entgegengesetzt gleichen Imaginärteil wie  $z$  haben soll, d.h. durch

$$z^* = x - iy. \quad (4.1.11)$$

Offensichtlich ist  $(z^*)^* = z$  für alle  $z \in \mathbb{C}$ . Weiter ist  $\mathbb{R} \subset \mathbb{C}$ , denn die komplexen Zahlen mit verschwindendem Imaginärteil sind umkehrbar eindeutig auf  $\mathbb{R}$  abbildbar. Offenbar ist  $z \in \mathbb{R}$  genau dann, wenn  $\operatorname{Im} z = 0$ , was zugleich  $z^* = z$  impliziert. Wir haben weiter

$$\operatorname{Re} z = \frac{z + z^*}{2}, \quad \operatorname{Im} z = \frac{z - z^*}{2i}. \quad (4.1.12)$$

Weiter rechnet man leicht nach (*Übung!*), daß

$$(z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 z_2)^* = z_1^* z_2^* \quad (4.1.13)$$

ist. Das Produkt einer komplexen Zahl mit ihrem konjugiert Komplexen ist

$$z z^* = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 - i^2 y^2 = x^2 + y^2 \geq 0. \quad (4.1.14)$$

Den Betrag der komplexen Zahl definieren wir als

$$|z| = \sqrt{x^2 + y^2} = \sqrt{z z^*}. \quad (4.1.15)$$

Schließlich können wir auch die Division im Bereich der komplexen Zahlen betrachten. Sei dazu  $z_2 \neq 0$  und  $z_1 \in \mathbb{C}$  beliebig. Dann gilt

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{x_2^2 + y_2^2} = \left( \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} \right) + i \left( \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2} \right). \quad (4.1.16)$$

Da  $z_2 \neq 0$  ist offenbar auch  $x_2^2 + y_2^2 \neq 0$  und also die Division für  $z_2 \neq 0$  durch die soeben berechneten Real- und Imaginärteile wohldefiniert.

Die reellen Zahlen können wir geometrisch durch eine Zahlengerade veranschaulichen. Entsprechend kann man die komplexen Zahlen geometrisch interpretieren, wenn man das Zahlenpaar  $(x, y) = (\operatorname{Re} z, \operatorname{Im} z)$  als Komponenten bzgl. eines kartesischen Koordinatensystems in der Euklidischen Ebene interpretiert. Dies ist die **Gaußsche Zahlenebene**. Es ist klar, daß  $|z|$  geometrisch die Länge des entsprechenden  $z$  repräsentierenden Ortsvektors in der Gaußschen Zahlenebene ist (s. Abb. 4.1).

Wir können nun diesen Vektor durch Polarkoordinaten  $(r, \varphi)$  darstellen. Offenbar ist  $r = |z|$ , und es gilt definitionsgemäß

$$z = x + iy = r(\cos \varphi + i \sin \varphi) = |z|(\cos \varphi + i \sin \varphi). \quad (4.1.17)$$

Definieren wir den Bereich für den Polarwinkel als  $(-\pi, \pi]$ , so errechnet sich dieser Winkel gemäß (s. Abschnitt 2.10.2)

$$\varphi = \arg z = \begin{cases} \operatorname{sign} y \arccos\left(\frac{x}{\sqrt{x^2 + y^2}}\right) & \text{falls } y \neq 0, \\ 0 & \text{falls } y = 0, x > 0, \\ \pi & \text{falls } y = 0, x < 0. \end{cases} \quad (4.1.18)$$

Man nennt  $\varphi$  auch das **Argument der komplexen Zahl** ( $\arg z$ ).

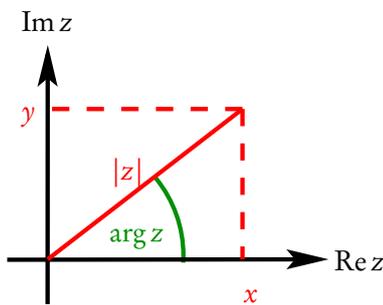


Abbildung 4.1: Zur Gaußschen Zahlenebene und Polarform einer komplexen Zahl.

## 4.2 Potenzreihen

Die Konvergenz von Folgen und Reihen kann nun wörtlich wie für die entsprechenden Begriffe im Reellen definiert werden. Natürlich fallen alle Begriffe weg, die die Anordnungsrelationen von reellen Zahlen verwenden wie Monotoniekriterien usw. Andererseits sind natürlich alle Sätze über absolut konvergente Reihen anwendbar, und wegen der Dreiecksungleichung sind komplexe Folgen und Reihen genau dann konvergent, wenn ihr Real- und Imaginärteil konvergent sind.

Nun definieren wir noch einige elementare Funktionen, die wir schon aus der reellen Analysis kennen, auch für komplexe Zahlen. Dies geschieht am bequemsten über **Potenzreihen**. Die Potenzreihen weisen nun im Komplexen dieselben Konvergenzeigenschaften wie im Reellen auf, d.h. sie sind in jedem abgeschlossenen Gebiet in der komplexen Zahlenebene absolut konvergent, die ganz im Inneren des Konvergenzbereichs liegen, und das für die Potenzreihe

$$f(z) = \sum_{j=0}^{\infty} a_j (z - a)^j \quad (4.2.1)$$

ein Kreis um  $a$  mit dem **Konvergenzradius**

$$r = \lim_{j \rightarrow \infty} \left| \frac{a_j}{a_{j+1}} \right|, \quad (4.2.2)$$

falls der Limes existiert. Vgl. dazu die hier vollständig ins Komplexe Übertragbare Herleitung und Diskussion der Gl. (A.7.87).

Wir beginnen mit der Exponentialfunktion und übernehmen die entsprechende Potenzreihe einfach von der entsprechenden reellen Funktion als Definition für die Exponentialfunktion (A.7.66) im Komplexen

$$\exp z = \sum_{j=0}^{\infty} \frac{z^j}{j!}. \quad (4.2.3)$$

Auch sie konvergiert für alle  $z \in \mathbb{C}$ .

Weiter benötigen wir noch die trigonometrischen Funktionen. Auch ihre Potenzreihen übernehmen wir aus dem Reellen, d.h. mit (A.7.72) bzw. (A.7.74) folgt (*nachrechnen!*)

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j}}{(2j)!} \quad (4.2.4)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j+1}}{(2j+1)!}. \quad (4.2.5)$$

Berechnen wir nun

$$\begin{aligned} \exp(iz) &= 1 + iz + \frac{(iz)^2}{2!} + \frac{(iz)^3}{3!} + \dots \\ &= \left(1 - \frac{z^2}{z} + \dots\right) + i \left(z - \frac{z^3}{3!} + \dots\right) \\ &= \left(\sum_{j=0}^{\infty} (-1)^j \frac{z^{2j}}{(2j)!}\right) + i \left(\sum_{j=0}^{\infty} (-1)^j \frac{z^{2j+1}}{(2j+1)!}\right). \end{aligned} \quad (4.2.6)$$

Dabei haben wir die Reihe so umgeordnet, daß wir in einem Term den Faktor  $i$  ausklammern konnten. Das ist bei Potenzreihen erlaubt, da sie in jedem kompakten Bereich der komplexen Ebene absolut konvergiert. Vergleichen wir nun die Reihen in den Klammern der Gleichung (4.2.6) mit (4.2.4) und (4.2.5), erhält man die **Eulersche Formel**

$$\exp(iz) = \cos z + i \sin z. \quad (4.2.7)$$

Für die Polardarstellung der komplexen Zahl (4.1.17) folgt damit

$$z = |z| \exp(i\varphi). \quad (4.2.8)$$

Da für die Exponentialfunktion auch im Komplexen die Formel

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2) \quad (4.2.9)$$

gilt, wie man mit Hilfe der Reihe (4.2.3) beweisen kann (vgl. (A.7.69)), erleichtert dies die Rechnung mit trigonometrischen Funktionen erheblich. Z.B. folgt genau wie (4.2.7) auch die Gleichung

$$\exp(-iz) = \cos z - i \sin z. \quad (4.2.10)$$

Wir haben damit

$$\cos z = \frac{1}{2} [\exp(iz) + \exp(-iz)], \quad \sin z = \frac{1}{2i} [\exp(iz) - \exp(-iz)]. \quad (4.2.11)$$

Dies erinnert an die Definition der Hyperbelfunktionen

$$\cosh z = \frac{1}{2} [\exp(z) + \exp(-z)], \quad \sinh z = \frac{1}{2} [\exp(z) - \exp(-z)]. \quad (4.2.12)$$

Vergleicht man (4.2.11) mit diesen Definitionen folgt sofort, daß

$$\cosh(iz) = \cos z, \quad \sinh(iz) = i \sin z \quad (4.2.13)$$

gilt. Die trigonometrischen und Hyperbelfunktionen sind im Komplexen also bis auf Konstanten im wesentlichen die gleichen Funktionen, und beide sind durch die Exponentialfunktion definiert.

Genauso folgt aus (4.2.11)

$$\cos(iz) = \cosh z, \quad \sin(iz) = i \sinh z. \quad (4.2.14)$$

Als Anwendungsbeispiel leiten wir noch die Additionstheoreme für die trigonometrischen Funktionen für reelle Argumente aus (4.2.9) ab. Es gilt nämlich einerseits wegen der Eulerschen Formel (4.2.7) und (4.2.8)

$$\exp[i(\varphi_1 + \varphi_2)] = \cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2) \quad (4.2.15)$$

und andererseits

$$\begin{aligned} \exp[i(\varphi_1 + \varphi_2)] &= \exp(i\varphi_1) \exp(i\varphi_2) = (\cos \varphi_1 + i \sin \varphi_1)(\cos \varphi_2 + i \sin \varphi_2) \\ &= (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i(\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2). \end{aligned} \quad (4.2.16)$$

Vergleicht man nun Real- und Imaginärteil von (4.2.15) und (4.2.16), folgen die bekannten Additionstheoreme

$$\begin{aligned} \cos(\varphi_1 + \varphi_2) &= \cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2, \\ \sin(\varphi_1 + \varphi_2) &= \sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2. \end{aligned} \quad (4.2.17)$$

Es ist leicht zu zeigen, daß diese Additionstheoreme auch allgemein für beliebige komplexe Argumente gelten (*Übung*).



## Kapitel 5

# Gewöhnliche Differentialgleichungen

Die Entwicklung von Lösungsverfahren für **Differentialgleichungen** stellt eines der wichtigsten mathematischen Hilfsmittel für die Physik dar, denn die Naturgesetze werden durch eben solche Gleichungen beschrieben. Dabei stellen uns **gewöhnliche Differentialgleichungen** vor die Aufgabe, unbekannte Funktionen einer Veränderlichen (in der klassischen Mechanik ist das die Zeit  $t$ ) aus Gleichungen zu bestimmen, die diese Funktion und Ableitungen dieser Funktion enthalten. Die höchste Ordnung der Ableitung bestimmt die **Ordnung der Differentialgleichung**. Die allgemeine Differentialgleichung  $n$ -ter Ordnung  $n \in \mathbb{N}$  nimmt dann die Form

$$F(t, f, \dot{f}, \dots, f^{(n)}) = 0 \quad (5.0.1)$$

an, wobei  $f^{(j)}$  die  $j$ -te Ableitung der Funktion  $f$  nach der Zeit bedeutet. Je nach Anwendung kann es erforderlich sein, alle Lösungen oder nur bestimmte, die durch die Anwendung vorgeschriebene Nebenbedingungen erfüllen.

**Systeme von Differentialgleichungen**  $n$ -ter Ordnung sind entsprechend gekoppelte Differentialgleichungen für mehrere Funktionen  $f_1, \dots, f_k$ , wobei die die höchste Ableitung, die in diesen Gleichungen vorkommt, die  $n$ -te ist.

In der klassischen Mechanik haben wir es i.a. mit Differentialgleichungen oder Systemen von Differentialgleichungen 2. Ordnung zu tun, denn die **Newtonsche Bewegungsgleichung** für einen Massenpunkt, auf den irgendwelche vorgegebenen Kräfte wirken, lautet

$$m\ddot{\vec{x}} = \vec{F}(t, \vec{x}, \dot{\vec{x}}). \quad (5.0.2)$$

Ein Beispiel ist die Bewegung eines Planeten um die sehr viel schwerere Sonne, deren Bewegung wir näherungsweise vernachlässigen dürfen. Setzen wir die Sonne in den Ursprung, ergibt sich mit dem Newtonschen Gravitationsgesetz die Gleichung

$$m\ddot{\vec{x}} = -\gamma m M \frac{\vec{x}}{|\vec{x}|^3}. \quad (5.0.3)$$

Hier hängt die Kraft nicht von der Geschwindigkeit ab. Dabei ist  $m$  die Masse des Planeten,  $M$  die der Sonne und  $\gamma$  die Newtonsche Gravitationskonstante.

Ein Beispiel für einen solchen Fall ist die Bewegung eines geladenen Teilchens in einem elektromagnetischen Feld (in Gaußschen Einheiten für das elektromagnetische Feld):

$$m\ddot{\vec{x}} = q \left[ \vec{E}(t, \vec{x}) + \frac{\dot{\vec{x}}}{c} \times \vec{B}(t, \vec{x}) \right]. \quad (5.0.4)$$

Dabei ist  $m$  die Masse des Teilchens,  $q$  seine Ladung und  $\vec{E}$  und  $\vec{B}$  das durch irgendwelche anderen Ladungen und Ströme erzeugte elektrische bzw. magnetische Feld.

In der Physik suchen wir Lösungen solcher Bewegungsgleichungen unter Vorgabe bestimmter **Anfangsbedingungen**, d.h. man gibt zur Zeit  $t = 0$  die Werte für Ort  $\vec{x}(0) = \vec{x}_0$  und Geschwindigkeit  $\dot{\vec{x}}(0) = \vec{v}_0$  des Teilchens vor und sucht Lösungen für die Differentialgleichung, die zusätzlich diese Anfangsbedingungen

$$\vec{x}(0) = \vec{x}_0, \quad \dot{\vec{x}}(0) = \vec{v}_0 \quad (5.0.5)$$

erfüllen.

Aus der physikalischen Fragestellung heraus erwarten wir, daß diese Differentialgleichungen stets eine Lösung besitzen (**Existenz**) und bei Vorgabe der Anfangsbedingungen (5.0.5) eindeutig (**Eindeutigkeit**) sind. Der Beweis entsprechender **Existenz- und Eindeutigkeitsätze** für Differentialgleichungen bei genauer Bestimmung der Eigenschaften der involvierten Funktionen (wie  $F$  in (5.0.1) oder die Kräfte in der Newtonschen Bewegungsgleichung) sind Klassiker der Analysis und finden sich in vielen mathematischen Lehrbüchern. In diesem Skript beschränken wir uns auf die Lösungsmethoden für die einfachsten Typen von Differentialgleichungen. Von der umfangreichen Spezialliteratur sei hier nur auf [Col90, Bro03] verwiesen.

## 5.1 Differentialgleichungen 1. Ordnung

Als einfachsten Typ von Differentialgleichungen betrachten wir Differentialgleichungen 1. Ordnung. Für sie gibt es mannigfaltige Lösungsverfahren, von denen wir hier nur einige der wichtigsten besprechen. Die allgemeinste Form des **Anfangswertproblems** ist

$$F(t, x, \dot{x}) = 0, \quad x(t_0) = x_0. \quad (5.1.1)$$

Natürlich läßt sich in dieser allgemeinsten Form wenig über die Lösungen der Differentialgleichung aussagen. Im folgenden betrachten wir die einfachsten Fälle, in denen sich die Lösung des Anfangswertproblems zumindest in impliziter Form auf **Integrationen** zurückführen läßt.

### 5.1.1 Separierbare Differentialgleichungen

Man spricht von **separierbaren Differentialgleichungen 1. Ordnung**, wenn sie sich auf die Form

$$p(t) + \dot{x}q(x) = 0 \quad (5.1.2)$$

zurückführen lassen. Dann genügt bereits eine beliebige Stammfunktion der Funktionen  $p$  und  $q$  (wobei im Einzelfall freilich stets auf deren Definitionsbereich und die damit verbundene Existenz der Integrale zu achten ist). Definieren wir nämlich die Stammfunktionen zu

$$P(t) = \int_{t_0}^t dt' p(t'), \quad Q(x) = \int_{x_0}^x dx' q(x), \quad (5.1.3)$$

können wir wegen des Hauptsatzes der Differential- und Integralrechnung für (5.1.2)

$$\dot{P}(t) + \dot{x} \frac{d}{dx} Q(x) = 0 \quad (5.1.4)$$

schreiben. Mit der Kettenregel wird dies zu

$$\frac{d}{dt} Q[x(t)] = -\dot{P}(t), \quad (5.1.5)$$

integrieren wir beide Gleichungen von  $t = t_0$  bis  $t$  erhalten wir unter Berücksichtigung der Anfangsbedingung bereits die Lösung in impliziter Form

$$Q(x) = -P(t). \quad (5.1.6)$$

Diese Lösung müssen wir nun nur noch nach  $x$  auflösen, um die Lösung der Differentialgleichung zu erhalten. Da wir die Anfangsbedingung bereits eingearbeitet haben, ist auch diese dann automatisch erfüllt.

Als *Beispiel* betrachten wir die Differentialgleichung des **radioaktiven Zerfalls**. Dazu gehen wir davon aus, daß die entsprechende Zerfallsrate (also die Anzahl der Atomkerne, die pro Zeiteinheit zerfällt) unabhängig von den äußeren Umständen (Temperatur, Druck usw.) und proportional zur Zahl  $N(t)$  der zur Zeit  $t$  vorhandenen Kerne ist. Zur Zeit  $t_0 = 0$  seien  $N_0$ . In eine Gleichung gebracht, haben wir es also mit dem Anfangswertproblem einer Differentialgleichung 1. Ordnung zu tun,

$$\dot{N}(t) = -\lambda N(t), \quad \lambda = \text{const.} \quad N(0) = N_0. \quad (5.1.7)$$

Diese Gleichung ist tatsächlich vom separablen Typ, denn wir können sie in der Form

$$\lambda + \dot{N} \frac{1}{N} \quad (5.1.8)$$

schreiben. Bis auf die Bezeichnung der Variablen  $N$  statt  $x$  ist das in der Tat in der Form (5.1.2). Nach der allgemeinen Lösungsformel (5.1.6) benötigen wir die Stammfunktionen von  $p(t) = \lambda$  und  $q(N) = 1/N$ :

$$P(t) = \int_0^t dt' \lambda = \lambda t, \quad N(t) = \int_0^N dN' \frac{1}{N'} = \ln \left( \frac{N}{N_0} \right). \quad (5.1.9)$$

Die Lösung lautet also gemäß (5.1.6)

$$\ln \left( \frac{N(t)}{N_0} \right) = -\lambda t \Rightarrow N(t) = N_0 \exp(-\lambda t). \quad (5.1.10)$$

Wir erhalten also das bekannte **Exponentialgesetz vom radioaktiven Zerfall**. Man bezeichnet als **Lebensdauer**  $\tau$  des betreffenden Atomkerns die Zeit nach der die Anzahl der Kerne auf  $N(\tau)/N_0 = e$  abgefallen ist. Aus unserer Lösung (5.1.10) folgt sofort, daß  $\tau = 1/\lambda$  ist. Die **Halbwertszeit**  $\tau_{1/2}$  gibt hingegen die Zeit an, nachdem die anfänglich vorhandene Menge an radioaktivem Material zur Hälfte zerfallen ist, d.h.

$$\ln \left( \frac{1}{2} \right) = -\ln 2 = -\lambda \tau_{1/2} \Rightarrow \tau_{1/2} = \frac{\ln 2}{\lambda} = \tau \ln 2 \approx 0,693 \tau. \quad (5.1.11)$$

### 5.1.2 Homogene Differentialgleichungen 1. Ordnung

Angenommen, die Differentialgleichung besitzt die spezielle Form

$$\dot{x} = f(t, x), \quad (5.1.12)$$

wobei die Funktion  $f$  die **Homogenitätseigenschaft**

$$f(\lambda t, \lambda x) = f(t, x) \quad (5.1.13)$$

besitzt. Dann können wir durch die Substitution

$$x(t) = t \alpha(t) \quad (5.1.14)$$

die Differentialgleichung in eine separable Gleichung transformieren. In der Tat ist dann wegen der Homogenitätseigenschaft (5.1.13)

$$\dot{x} = \alpha + t \dot{\alpha} = f(t, t \alpha) = f(1, \alpha) = \tilde{f}(\alpha). \quad (5.1.15)$$

Etwas umgeformt erhalten wir die folgende Gleichung für  $\alpha$ :

$$-t + \frac{\dot{\alpha}}{\tilde{f}(\alpha) - \alpha} = 0, \quad (5.1.16)$$

die tatsächlich von der separablen Form (5.1.2) ist.

### 5.1.3 Exakte Differentialgleichung 1. Ordnung

Angenommen die Differentialgleichung sei von der Form

$$\alpha(t, x)\dot{x} + \beta(t, x) = 0, \quad (5.1.17)$$

und es existiere eine Funktion  $g(t, x)$ , so daß

$$\alpha(t, x) = -\frac{\partial}{\partial t}g(t, x), \quad \beta(t, x) = -\frac{\partial}{\partial x}g(t, x). \quad (5.1.18)$$

Betrachten wir also  $\vec{A} = (\alpha, \beta)$  als zweidimensionales Vektorfeld mit  $\vec{\xi} = (t, x)$  als ebene unabhängige Koordinaten, verlangen wir, daß  $g$  ein Potential von  $\vec{A}$  ist. Nach dem **Lemma von Poincaré** (vgl. Abschnitt 2.8) ist das in einfach zusammenhängenden Gebieten der Fall, wenn  $\partial_x \alpha = \partial_t \beta$  gilt, und das Potential ist dann durch das entsprechende Wegintegral gegeben (vgl. Abschnitt 2.6). Jedenfalls können wir dann (5.1.17) in der Form

$$\frac{d\vec{\xi}}{dt} \cdot \vec{\nabla} g = \frac{d}{dt}g[t, x(t)] = 0 \quad (5.1.19)$$

schreiben. Die allgemeine Lösung ist demnach in impliziter Form sofort durch

$$g[t, x(t)] = C = \text{const.} \quad (5.1.20)$$

gegeben. Die Integrationskonstante  $C$  bestimmt sich aus der Anfangsbedingung  $x(t_0) = x_0$ .

### 5.1.4 Homogene lineare Differentialgleichung 1. Ordnung

Eine Differentialgleichung 1. Ordnung heißt **linear**, wenn sie von der Form

$$\dot{x} + p(t)x = q(t) \quad (5.1.21)$$

ist und **homogen**, wenn  $q(t) = 0$  ist. Wir beschäftigen uns zuerst mit dem homogenen Fall

$$\dot{x} + p(t)x = 0. \quad (5.1.22)$$

Dividiert man diese Gleichung durch  $x$ , erkennt man, daß sie vom separablen Typ ist. Jedenfalls können wir sie auf die Form

$$\frac{d}{dt} \ln\left(\frac{x}{x_0}\right) = -p(t) \quad (5.1.23)$$

bringen. Berücksichtigen wir die Anfangsbedingung  $x(t_0) = x_0$  und integrieren diese Gleichung bzgl.  $t$ , erhalten wir nach einer einfachen Umformung die Lösung in der Form

$$x(t) = x_0 \exp\left[-\int_{t_0}^t dt' p(t')\right]. \quad (5.1.24)$$

### 5.1.5 Inhomogene lineare Differentialgleichung 1. Ordnung

Kommen wir nun auf (5.1.21) mit  $q \neq 0$  zurück. Zuerst bemerken wir, daß wegen der Linearität der Differentialgleichung die Differenz zweier Lösungen  $x_1$  und  $x_2$  die homogene Gleichung (5.1.23) löst:

$$\dot{x}_1(t) + p(t)x_1(t) = q(t), \quad \dot{x}_2(t) + p(t)x_2(t) = q(t) \Rightarrow \frac{d}{dt}(x_1 - x_2) + p(t)(x_1 - x_2) = 0. \quad (5.1.25)$$

Die allgemeine Lösung der inhomogenen Gleichung ist also durch die Summe aus der allgemeinen Lösung der homogenen Gleichung und einer beliebigen Lösung der inhomogenen Gleichung gegeben. Sei also  $\tilde{x} \neq 0$  eine Lösung der homogenen Gleichung. Dann können wir eine spezielle Lösung der inhomogenen Gleichung finden, indem wir den **Ansatz der Variation der Konstanten**

$$x(t) = y(t)\tilde{x}(t) \quad (5.1.26)$$

vornehmen. Wegen  $\dot{\tilde{x}} = -p$  folgt dann nämlich aus der Produktregel

$$\dot{x} + p = y\dot{\tilde{x}} + y\dot{\tilde{x}} + py\tilde{x} = y\dot{\tilde{x}} \stackrel{!}{=} q. \quad (5.1.27)$$

Damit finden wir die Lösung für  $y$  durch eine einfache Integration. Da wir nur irgendeine Lösung der inhomogenen Gleichung benötigen, können wir zusätzlich  $y(t_0) = 0$  fordern. Dann folgt

$$y(t) = \int_{t_0}^t dt' \frac{q(t')}{\tilde{x}(t')}. \quad (5.1.28)$$

Dann wird nach der Überlegung oben das Anfangswertproblem der inhomogenen Gleichung mit  $x(t_0) = x_0$  durch

$$x(t) = \tilde{x}(t) + \tilde{x}(t) \int_{t_0}^t dt' \frac{q(t')}{\tilde{x}(t')}, \quad (5.1.29)$$

wobei gemäß (5.1.24)

$$\tilde{x}(t) = x_0 \exp \left[ - \int_{t_0}^t dt' p(t') \right]. \quad (5.1.30)$$

die Lösung der homogenen Gleichung (5.1.22) ist, die die Anfangsbedingung  $\tilde{x}(t_0) = x_0$  erfüllt.

## 5.2 Lineare Differentialgleichungen 2. Ordnung

Lineare Differentialgleichungen besitzen eine besonders einfache Lösungsstruktur. In der Physik kann man auch oft kompliziertere Probleme durch lineare Differentialgleichungen nähern (für ein Beispiel s.u. den Abschnitt zum harmonischen Oszillator als Näherung für das mathematische Pendel). Hier betrachten wir kurz die ganz allgemeine Struktur der allgemeinen Lösungen für lineare DGLn 2. Ordnung.

Die allgemeine lineare Differentialgleichung (DGL) 2. Ordnung lautet

$$\ddot{x}(t) + A(t)\dot{x}(t) + B(t)x(t) = C(t). \quad (5.2.1)$$

Dabei sind  $A$ ,  $B$  und  $C$  vorgegebene Funktionen der unabhängigen Variablen  $t$  und  $x(t)$  die gesuchte Funktion. Hier besprechen wir nur die wichtigsten Grundlagen über die Struktur der Lösungen solcher linearer Differentialgleichungen. Konkrete Beispiele liefern die in den nächsten Abschnitten behandelten harmonischen Oszillatoren verschiedener Art.

Man nennt die obige Differentialgleichung (5.2.1) **homogen**, wenn  $C(t) = 0$  und entsprechend **inhomogen**, wenn  $C(t) \neq 0$ . Wir betrachten zuerst die Lösungsstruktur der **homogenen Gleichung**

$$\ddot{x}(t) + A(t)\dot{x}(t) + B(t)x(t) = 0. \quad (5.2.2)$$

Da die Ableitungsoperation linear ist, d.h. für irgendwelche zwei Funktionen  $x_1(t)$  und  $x_2(t)$ , die mindestens zweimal differenzierbar sind,

$$\frac{d}{dt}[C_1x_1(t) + C_2x_2(t)] = C_1\dot{x}_1 + C_2\dot{x}_2, \quad \frac{d^2}{dt^2}[C_1x_1(t) + C_2x_2(t)] = C_1\ddot{x}_1 + C_2\ddot{x}_2 \quad (5.2.3)$$

gilt, ist für zwei Lösungen  $x_1$  und  $x_2$  von (5.2.2) auch die **Linearkombination**

$$x(t) = C_1 x_1(t) + C_2 x_2(t) \quad (5.2.4)$$

mit  $C_1, C_2 = \text{const}$  eine weitere Lösung. Es ist weiter klar, daß wir zur eindeutigen Festlegung der Lösung **Anfangsbedingungen** fordern müssen, d.h. wir verlangen von der Lösung  $x$  der DGL zusätzlich, daß sie und ihre erste Ableitung bei  $t = 0$  bestimmte Werte annimmt:

$$x(0) \stackrel{!}{=} x_0, \quad \dot{x}(0) \stackrel{!}{=} v_0. \quad (5.2.5)$$

Nehmen wir an, wir hätten zwei Lösungen  $x_1$  und  $x_2$  gefunden, können wir versuchen, die Konstanten  $C_1$  und  $C_2$  in der Linearkombination (5.2.4) so zu bestimmen, daß diese Anfangsbedingungen (5.2.5) gelten, d.h. wir müssen das lineare Gleichungssystem

$$\begin{aligned} C_1 x_1(0) + C_2 x_2(0) &= x_0 \\ C_1 \dot{x}_1(0) + C_2 \dot{x}_2(0) &= v_0 \end{aligned} \quad (5.2.6)$$

nach  $C_1$  und  $C_2$  auflösen. Multiplizieren wir die erste Gleichung mit  $\dot{x}_2(0)$  und die zweite mit  $x_2(0)$  und subtrahieren die beiden entstehenden Gleichungen, finden wir

$$C_1 [x_1(0)\dot{x}_2(0) - x_2(0)\dot{x}_1(0)] = \dot{x}_2(0)x_0 - x_2(0)v_0. \quad (5.2.7)$$

Wenn die eckige Klammer nicht verschwindet, können wir nach  $C_1$  auflösen:

$$C_1 = \frac{\dot{x}_2(0)x_0 - x_2(0)v_0}{x_1(0)\dot{x}_2(0) - x_2(0)\dot{x}_1(0)}. \quad (5.2.8)$$

Unter derselben Voraussetzung können wir auf ähnliche Weise auch  $C_2$  berechnen:

$$C_2 = \frac{v_0 x_1(0) - x_0 \dot{x}_1(0)}{x_1(0)\dot{x}_2(0) - x_2(0)\dot{x}_1(0)}. \quad (5.2.9)$$

Wir untersuchen nun noch, wann die Bedingung, daß der Nenner in (5.2.8) und (5.2.9) für die Lösungen  $x_1$  und  $x_2$  der DGL nicht verschwindet, erfüllt ist. Es handelt sich um die **Determinante der Koeffizientenmatrix** des linearen Gleichungssystems (5.2.6). Um diesen Ausdruck näher zu untersuchen, definieren wir für die beiden Lösungen  $x_1$  und  $x_2$  die **Wronski-Determinante** genannte Größe

$$W(t) = \det \begin{pmatrix} x_1(t) & x_2(t) \\ \dot{x}_1(t) & \dot{x}_2(t) \end{pmatrix} = x_1(t)\dot{x}_2(t) - x_2(t)\dot{x}_1(t). \quad (5.2.10)$$

Berechnen wir die Zeitableitung, finden wir mit Hilfe der Produktregel nach einiger Rechnung

$$\dot{W}(t) = x_1(t)\ddot{x}_2(t) - x_2(t)\ddot{x}_1(t). \quad (5.2.11)$$

Jetzt verwenden wir, daß  $x_1$  und  $x_2$  Lösungen der DGL (5.2.2) sind und setzen

$$\ddot{x}_j(t) = -A(t)\dot{x}_j(t) - B(t)x_j(t), \quad j \in \{1, 2\} \quad (5.2.12)$$

in (5.2.11) ein. Das ergibt

$$\dot{W}(t) = -A(t)[x_1(t)\dot{x}_2(t) - x_2(t)\dot{x}_1(t)] = -A(t)W(t). \quad (5.2.13)$$

Dies ist eine lineare Differentialgleichung 1. Ordnung, die sich durch Trennen der Variablen lösen läßt. Teilen wir also (5.2.13) durch  $W(t)$ , erhalten wir

$$\frac{\dot{W}(t)}{W(t)} = -A(t). \quad (5.2.14)$$

Integration dieser Gleichung bzgl.  $t$  von  $t = 0$  bis  $t$ , liefert

$$\ln\left(\frac{W(t)}{W(0)}\right) = -\int_0^t dt' A(t'). \quad (5.2.15)$$

Lösen wir dies nach  $W(t)$  auf, finden wir schließlich

$$W(t) = W(0) \exp\left[-\int_0^t dt' A(t')\right]. \quad (5.2.16)$$

Das bedeutet aber, daß entweder  $W(t) = 0 = \text{const}$  ist (nämlich wenn  $W(0) = 0$ ) oder  $W(t) \neq 0$  für alle  $t \in \mathbb{R}$  gilt.

Untersuchen wir deshalb weiter, was es für die Lösungen  $x_1$  und  $x_2$  der DGL bedeutet, wenn  $W(t) = 0$  für alle  $t$  gilt. Aus der Definition der Wronski-Determinante (5.2.10) folgt dann

$$x_1(t)\dot{x}_2(t) - x_2(t)\dot{x}_1(t) = 0 \Rightarrow \frac{\dot{x}_1(t)}{x_1(t)} = \frac{\dot{x}_2(t)}{x_2(t)}. \quad (5.2.17)$$

Auch diese Gleichung können wir wieder bzgl.  $t$  von 0 bis  $t$  integrieren, und das ergibt

$$\ln\left(\frac{x_1(t)}{x_1(0)}\right) = \ln\left(\frac{x_2(t)}{x_2(0)}\right) \Rightarrow x_2(t) = \frac{x_2(0)}{x_1(0)} x_1(t). \quad (5.2.18)$$

Das bedeutet aber, daß  $W(0) = 0$  genau dann, wenn  $x_2(t) = C x_1(t)$  mit  $C = x_2(0)/x_1(0) = \text{const}$  ist.

Die Anfangsbedingungen (5.2.5) sind also durch die Linearkombination (5.2.4) genau dann immer erfüllbar, wenn die Lösungen  $x_1$  und  $x_2$  **linear unabhängig** sind und daher  $W(0) \neq 0$  ist. Um also die **allgemeine Lösung der homogenen DGL** zu finden, müssen wir nur irgendwelche zwei linear unabhängigen Lösungen finden. Die allgemeine Lösung ist dann durch die allgemeine Linearkombination (5.2.4) gegeben.

Kommen wir nun auf die inhomogene Gleichung (5.2.1) zurück. Nehmen wir wieder an, daß  $x_1^{(\text{inh})}(t)$  und  $x_2^{(\text{inh})}(t)$  Lösungen dieser inhomogenen Gleichung ist. Wegen der Linearität der Ableitungsoperation und der Linearität der linken Seite der inhomogenen Gleichung erfüllt dann offenbar  $x_1^{(\text{inh})}(t) - x_2^{(\text{inh})}(t)$  die *homogene DGL*. Das bedeutet aber, daß bei Kenntnis von zwei linear unabhängigen Lösungen  $x_1^{(\text{hom})}(t)$  und  $x_2^{(\text{hom})}(t)$  der *homogenen DGL* diese Differenz durch eine Linearkombination dieser Lösungen gegeben sein muß. Es ist also die *allgemeine* Lösung der inhomogenen Gleichung durch

$$x(t) = C_1 x_1^{(\text{hom})}(t) + C_2 x_2^{(\text{hom})}(t) + x_1^{(\text{inh})}(t) \quad (5.2.19)$$

gegeben. Haben wir also die allgemeine Lösung der homogenen Gleichung gefunden, genügt es, nur eine einzige spezielle Lösung der inhomogenen DGL zu kennen, um alle Lösungen in der Form (5.2.19) angeben zu können. Es ist klar, daß auch hier die Anfangsbedingungen (5.2.5) durch die entsprechende Berechnung der Integrationskonstanten  $C_1$  und  $C_2$  stets erfüllbar sind, wenn nur  $x_1^{(\text{hom})}(t)$  und  $x_2^{(\text{hom})}(t)$  linear unabhängig sind und also die Wronski-Determinante  $W(0) \neq 0$  ist.

### 5.3 Der ungedämpfte harmonische Oszillator

Bei vielen typischen Bewegungsgleichungen der klassischen Mechanik tritt der Fall auf, daß ein Massepunkt sich in einem **Kräftepotential** bewegt. Wir betrachten eindimensionale Bewegungen entlang der  $x$ -Achse eines kartesischen Koordinatensystems. Dann ist die Kraft durch die Ableitung des Potentials gegeben:

$$F(x) = -V'(x). \quad (5.3.1)$$

Die Newtonsche Bewegungsgleichung für solch einen Massenpunkt lautet demnach

$$m\ddot{x} = -V'(x). \quad (5.3.2)$$

Um die Bahn der Bewegung als Funktion der Zeit zu erhalten, müssen wir also eine **Differentialgleichung** zweiter Ordnung lösen, d.h. wir suchen die Ortskoordinate  $x$  als Funktion von  $t$ . Dabei ergibt sich eine ganze Schar von Lösungen. Um die Bewegung des Massenpunktes eindeutig festzulegen, müssen wir noch **Anfangsbedingungen** fordern, d.h. wir müssen zu einem vorgegebenen Zeitpunkt, den wir bequemlichkeitshalber bei  $t = 0$  wählen, **Ort und Geschwindigkeit** des Massenpunktes vorgeben. Wir verlangen also von der Lösung der Bewegungsgleichung (5.3.2), daß die Anfangsbedingungen

$$x(0) = x_0, \quad \dot{x}(0) = v_0 \quad (5.3.3)$$

erfüllt sind. Wir wissen bereits, daß für die Bewegungsgleichung (5.3.2) der **Satz von der Energieerhaltung** gilt, denn multiplizieren wir (5.3.2) mit  $\dot{x}$  und bringen alle Ausdrücke auf die linke Seite der Gleichung, erhalten wir

$$m\dot{x}\ddot{x} + \dot{x}V'(x) = 0. \quad (5.3.4)$$

Es ist aber leicht zu sehen, daß dies eine totale Zeitableitung ist, denn es gilt

$$\frac{d}{dt}(\dot{x}^2) = 2\dot{x}\ddot{x}, \quad \frac{d}{dt}V(x) = \dot{x}V'(x). \quad (5.3.5)$$

Wir können also (5.3.4) in der Form

$$\frac{d}{dt} \left[ \frac{m}{2} \dot{x}^2 + V(x) \right] = 0 \quad (5.3.6)$$

schreiben. Das bedeutet aber, daß der Ausdruck in den eckigen Klammern, die **Gesamtenergie** des Massenpunktes, für alle Lösungen der Bewegungsgleichung (5.3.2) zeitlich konstant ist:

$$E = \frac{m}{2} \dot{x}^2 + V(x) = \frac{m}{2} v_0^2 + V(x_0) = \text{const.} \quad (5.3.7)$$

Dabei haben wir die Anfangsbedingung (5.3.4) eingesetzt, um den Wert der Gesamtenergie zu bestimmen.

Nun ist die **kinetische Energie**

$$E_{\text{kin}} = \frac{m}{2} \dot{x}^2 \geq 0. \quad (5.3.8)$$

In Abb. 5.1 haben wir ein beliebiges Potential als Funktion der Ortskoordinate aufgezeichnet und den konstanten Wert der Gesamtenergie als horizontale Linie eingetragen. Da  $E_{\text{kin}} \geq 0$  muß für alle Zeiten

$$E \geq V(x) \quad (5.3.9)$$

gelten. D.h. für eine durch die Anfangsbedingungen gegebene Gesamtenergie  $E$  kann sich das Teilchen nur dort aufhalten, wo das Potential unterhalb der Linie für die Gesamtenergie verläuft.

Nun kommt es oft vor, daß das Potential bei einer Stelle  $x_0$  ein **lokales Minimum** aufweist. In diesem Minimum ist  $V'(x_0) = 0$ . Ist dann die Energie  $E$  so gewählt, daß die Linie  $E = \text{const}$  das Potential an zwei Stellen

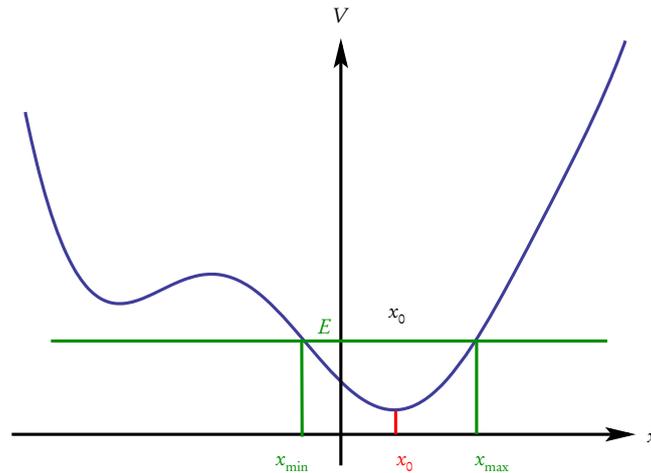


Abbildung 5.1: Bewegung in einer Potentialmulde.

$x_{\min}$  und  $x_{\max}$  schneidet, muß das Teilchen in dem Bereich  $[x_{\min}, x_{\max}]$  bleiben, denn aufgrund der Differentialgleichung muß die Ortskoordinate als Funktion der Zeit mindestens zweimal differenzierbar sein und ist daher stetig. Der Massenpunkt kann also nicht über eine Potentialbarriere einfach in einen anderen Bereich springen, wo wieder  $E > V(x)$  gilt. Das Teilchen ist also in dem besagten Intervall gefangen. Ist dieser Bereich nicht zu groß, reicht es weiter aus, das Potential um  $x_0$  in eine **Potenzreihe** zu entwickeln und nur die Terme bis zur zweiten Ordnung mitzunehmen. Angenommen, das Potential ist mindestens dreimal stetig differenzierbar, können wir schreiben (**Taylor-Entwicklung**)

$$V(x) = V(x_0) + \frac{1}{2} V''(x_0)(x - x_0)^2 + \mathcal{O}[(x - x_0)^3]. \quad (5.3.10)$$

Dabei bedeutet das **Landau-Symbol**  $\mathcal{O}[(x - x_0)^3]$ , daß der nächste Term in der Potenzreihenentwicklung von der Größenordnung  $(x - x_0)^3$  ist. Es ist klar, daß der Term linear zu  $(x - x_0)$  verschwindet, weil voraussetzungsgemäß  $V$  an der Stelle  $x_0$  ein Minimum besitzt. Außerdem nehmen wir an, daß  $V''(x_0) > 0$  ist.

Für die Kraft folgt dann

$$F(x) = -V'(x) = -V''(x_0)(x - x_0) + \mathcal{O}[(x - x_0)^2]. \quad (5.3.11)$$

Für nicht zu große Abweichungen der Lage des Massenpunktes von  $x_0$  können wir also näherungsweise die vereinfachte Bewegungsgleichung

$$m\ddot{x} = -D(x - x_0) \quad \text{mit} \quad D = V''(x_0) > 0. \quad (5.3.12)$$

betrachten.

Wählen wir das Koordinatensystem noch so, daß  $x_0 = 0$  ist, erhalten wir die relativ einfache Bewegungsgleichung eines **harmonischen Oszillators**:

$$m\ddot{x} = -Dx. \quad (5.3.13)$$

Wir können diese Situation durch ein reales System in sehr guter Näherung realisieren, indem wir einen Massenpunkt an eine Feder hängen. Für nicht zu große Auslenkungen der Feder aus ihrer Gleichgewichtslage ist die von der Feder ausgeübte Kraft proportional zur Auslenkung ( $|F_{\text{Feder}}| = D\Delta x$ , wo  $\Delta x$  die Dehnung der Feder aus ihrer Ruhelage ist). Der Gleichgewichtspunkt  $x_0 = 0$  ist dann dadurch gegeben, daß dort die Federkraft die Schwerkraft  $mg$  gerade kompensiert. Die Feder wirkt immer der Auslenkung entgegen, und die gesamte Kraft auf den Massenpunkt ist dann durch  $F_x = -Dx$  gegeben. Dabei rührt das Vorzeichen in dieser Gleichung daher, daß die Feder immer der Auslenkung aus der Gleichgewichtslage entgegenwirkt.

Zur Lösung dieser Gleichung beachten wir, daß es sich um eine lineare homogene Differentialgleichung (DGL) zweiter Ordnung handelt. Wir können uns also auf die allgemeinen Betrachtungen im vorigen Abschnitt stützen. Aus den dortigen Überlegungen folgt, daß die allgemeine Lösung von der Form

$$x(t) = C_1 x_1(t) + C_2 x_2(t) \quad (5.3.14)$$

ist, wobei  $x_1$  und  $x_2$  irgendwelche zwei linear unabhängige Lösungen der Gleichung sind, d.h. es muß  $x_1/x_2 \neq \text{const}$  sein, und beide Funktionen müssen die DGL lösen. Ein Blick auf (5.3.13) zeigt, daß ein Ansatz mit trigonometrischen Funktionen

$$x_1(t) = C_1 \sin(\omega_0 t), \quad x_2(t) = C_2 \cos(\omega_0 t) \quad (5.3.15)$$

erfolgsversprechend ist, denn es gilt

$$\dot{x}_1 = C_1 \omega_0 \cos(\omega_0 t), \quad \ddot{x}_1 = -C_1 \omega_0^2 \sin(\omega_0 t) \quad (5.3.16)$$

und

$$\dot{x}_2 = -C_2 \omega_0 \sin(\omega_0 t), \quad \ddot{x}_2 = -C_2 \omega_0^2 \cos(\omega_0 t). \quad (5.3.17)$$

Setzt man diese Ansätze in (5.3.13) ein, erkennt man sofort, daß beides Lösungen der Differentialgleichung sind, und zwar für

$$\omega_0 = \sqrt{\frac{D}{m}}. \quad (5.3.18)$$

Die allgemeine Lösung der DGL (5.3.13) lautet also

$$x(t) = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t). \quad (5.3.19)$$

Um diese Lösung etwas einfacher analysieren zu können, bringen wir sie noch in eine etwas einfachere Form, und zwar versuchen wir Konstanten  $\hat{x} \geq 0$  und  $\varphi_0$  so zu bestimmen, daß

$$x(t) = \hat{x} \cos(\omega_0 t - \varphi_0) \quad (5.3.20)$$

gilt. Ausnutzen des Additionstheorems für den Cosinus liefert

$$x(t) = \hat{x} [\cos \varphi_0 \cos(\omega_0 t) + \sin \varphi_0 \sin(\omega_0 t)]. \quad (5.3.21)$$

Vergleicht man dies mit (5.3.19) folgt, daß dann

$$C_1 = \hat{x} \cos \varphi_0, \quad C_2 = \hat{x} \sin \varphi_0 \quad (5.3.22)$$

gelten muß. Es ist klar, daß man dies als Gleichung für die Komponenten eines Vektors  $(C_1, C_2)$  in der Ebene, ausgedrückt durch seine Polarkoordinaten  $(\hat{x}, \varphi_0)$  ansehen kann (s. nebenstehende Abbildung). Quadriert man jedenfalls diese beiden Gleichungen, erhält man

$$\hat{x}^2 (\cos^2 \varphi_0 + \sin^2 \varphi_0) = \hat{x}^2 = C_1^2 + C_2^2 \Rightarrow \hat{x} = \sqrt{C_1^2 + C_2^2}. \quad (5.3.23)$$

Aus dem Bild liest man weiter ab, daß

$$\varphi_0 = \text{sign } C_2 \arccos\left(\frac{C_1}{\hat{x}}\right) \in (-\pi, \pi) \quad (5.3.24)$$

gegeben ist. Das einzige Problem mit dieser Formel ist, daß für  $C_2 = 0$  und  $C_1 \neq 0$  ein unbestimmtes Ergebnis herauskommt. Man hat dann aber  $\cos \varphi_0 = C_1/|C_1| = \pm 1$ . Für  $C_1 > 0$  erhält man dann immer noch eindeutig  $\varphi_0 = 0$ . Für  $C_1 < 0$  wären aber zwei Lösungen  $\varphi_0 = \pm \pi$  korrekt. Man kann in diesem Fall einfach eine von beiden Möglichkeiten wählen, z.B.  $\varphi_0 = +\pi$ . Diese Gleichung zur Berechnung des Polarwinkels liefert im Gegensatz zu der in der Literatur oft zu findenden Formel " $\varphi_0 = \arctan(C_2/C_1)$ " stets den korrekten Winkel, ohne daß man sich genauere Gedanken machen muß, in welchem Quadranten der gerade betrachtete Punkt liegt.

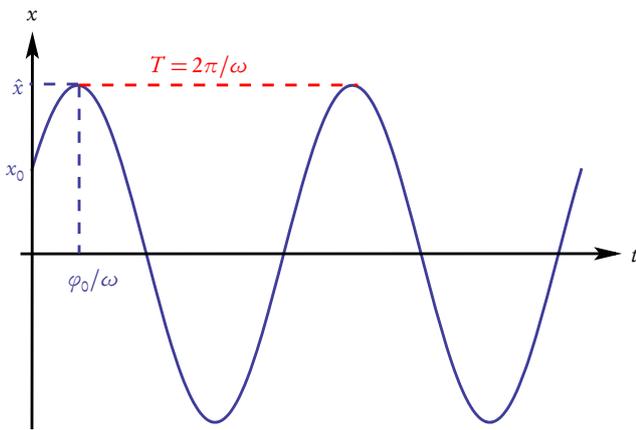


Abbildung 5.2: Lösung zum harmonischen Oszillator mit den Kenngrößen  $\hat{x}$  (Amplitude),  $\varphi_0$  (Anfangsphase) und  $T$  (Periodendauer).

Die Konstanten  $C_1$  und  $C_2$  in (5.3.19) lassen sich aus den Anfangsbedingungen (5.3.3) bestimmen. Wir verlangen also

$$\begin{aligned} x(0) &= C_1 = x_0, \\ \dot{x}(0) &= C_2 \omega_0 = v_0 \Rightarrow C_2 = \frac{v_0}{\omega_0}. \end{aligned} \quad (5.3.25)$$

Die Lösung für das Anfangswertproblem lautet also

$$x(t) = x_0 \cos(\omega_0 t) + \frac{v_0}{\omega_0} \sin(\omega_0 t). \quad (5.3.26)$$

Für die Lösungsform (5.3.20) ergibt sich aus (5.3.23) und (5.3.24)

$$\begin{aligned} \hat{x} &= \sqrt{x_0^2 + \frac{v_0^2}{\omega^2}}, \\ \varphi_0 &= \text{sign } v_0 \arccos\left(\frac{x_0}{\hat{x}}\right). \end{aligned} \quad (5.3.27)$$

Es ergibt sich also insgesamt eine um  $\Delta t = \varphi_0/\omega_0$  entlang der  $t$ -Achse verschobene cos-Funktion mit der **Periodendauer**  $T$ , wobei

$$\omega_0 T = 2\pi \quad T = \frac{2\pi}{\omega_0}. \quad (5.3.28)$$

Die **Frequenz**, also die Anzahl der Schwingungen pro Zeiteinheit ist durch

$$f = \frac{1}{T} = \frac{\omega_0}{2\pi} \quad (5.3.29)$$

gegeben. Der Massepunkt schwingt zwischen den Werten  $\pm \hat{x}$  hin und her. Diese Maximalabweichung von der Ruhelage  $\hat{x}$  heißt **Amplitude** der Schwingung (vgl. Abb. 5.2). Wir bemerken, daß die Periodendauer der Schwingung unabhängig von der Amplitude ist. Man bezeichnet solche Schwingungen als **harmonische Schwingungen**. Schwingungen sind nur dann strikt harmonisch, wenn die Kraft exakt proportional zur Auslenkung von der Ruhelage ist. Für allgemeinere Kraftgesetze liegt dieser Fall nur näherungsweise für **kleine Amplituden** vor.

## 5.4 Der gedämpfte harmonische Oszillator

Im allgemeinen wird die Bewegung eines Massepunktes auch irgendwelchen Reibungsprozessen unterliegen. Um zu sehen, welche Auswirkungen die Reibung besitzt, untersuchen wir den besonders einfachen Fall der **Stokesschen Reibung**, wo die Reibungskraft proportional zur Geschwindigkeit  $v = \dot{x}$  ist. Die Bewegungsgleichung lautet dann

$$m\ddot{x} = -Dx - \beta\dot{x}. \quad (5.4.1)$$

In die Normalform gebracht ergibt sich wieder eine lineare homogene Differentialgleichung:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0 \quad \text{mit} \quad \gamma = \frac{\beta}{2m}, \quad \omega_0^2 = \frac{D}{m}. \quad (5.4.2)$$

Wir werden gleich sehen, daß die willkürlich erscheinende Einführung des Faktors 2 im Reibungsterm einige Formeln ein wenig übersichtlicher macht.

Um diese Gleichung zu lösen, bemerken wir, daß sich die Exponentialfunktion beim Differenzieren „reproduziert“. Daher erscheint der Ansatz für die Lösung der Gleichung (5.4.2)

$$x(t) = A \exp(\lambda t), \quad \lambda = \text{const} \quad (5.4.3)$$

erfolgsversprechend. In der Tat ist

$$\dot{x}(t) = A\lambda \exp(\lambda t), \quad \ddot{x}(t) = A\lambda^2 \exp(\lambda t). \quad (5.4.4)$$

Setzt man also den Ansatz (5.4.3) in (5.4.2) ein, ergibt sich

$$A \exp(\lambda t)(\lambda^2 + 2\gamma\lambda + \omega_0^2) = 0. \quad (5.4.5)$$

Demnach erhalten wir also (nichttriviale) Lösungen, d.h. für  $A \neq 0$ , für die Lösungen der quadratischen Gleichung

$$\lambda^2 + 2\gamma\lambda + \omega_0^2 = 0 \Rightarrow (\lambda + \gamma)^2 + \omega_0^2 - \gamma^2 = 0. \quad (5.4.6)$$

Wir müssen nun mehrere Fälle unterscheiden, je nachdem, ob die Lösungen reell oder komplex sind und ob das quadratische Polynom zwei einfache oder eine doppelte Nullstelle besitzt:

1.  $\omega_0 > \gamma$ : Zwei zueinander konjugiert komplexe Nullstellen,
2.  $\omega_0 < \gamma$ : Zwei einfache reelle Nullstellen,
3.  $\omega_0 = \gamma$ : Eine doppelte Nullstelle.

Wir behandeln diese Fälle nun nacheinander ausführlich. Wir benötigen hierzu die kurze Einführung in die komplexen Zahlen in Abschnitt 4.

### 5.4.1 Schwingfall ( $\omega_0 > \gamma$ )

In diesem Fall besitzt die Gleichung (5.4.6) die beiden zueinander konjugiert komplexen Lösungen

$$\lambda_{1,2} = -\gamma \pm i\sqrt{\omega_0^2 - \gamma^2} = -\gamma \pm i\omega \quad \text{mit} \quad \omega = \sqrt{\omega_0^2 - \gamma^2} > 0. \quad (5.4.7)$$

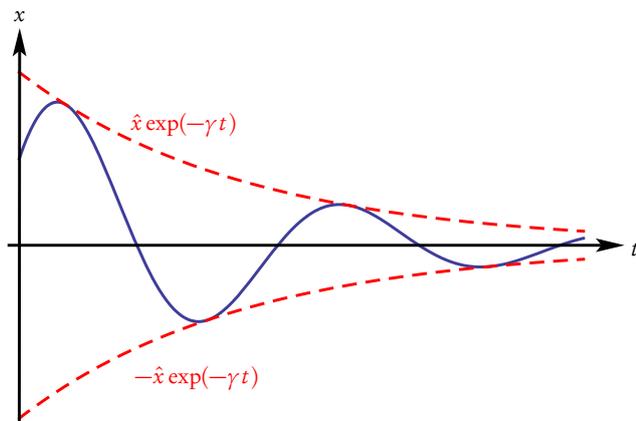


Abbildung 5.3: Lösung zum gedämpften harmonischen Oszillator. Für  $\omega_0 > \gamma$  schwingt der Massenpunkt wieder sinusförmig auf und ab, aber die Amplitude ist exponentiell gedämpft. Die Dämpfungsrate ist  $\gamma$ .

Wir haben damit offenbar zwei linear unabhängige Lösungen über unseren Ansatz (5.4.3) gefunden, und die allgemeine Lösung ergibt sich als deren Linearkombination

$$x(t) = \exp(-\gamma t)[C_1 \exp(-i\omega t) + C_2 \exp(i\omega t)]. \quad (5.4.8)$$

Da selbstverständlich nur reelle Lösungen physikalisch sinnvoll sind, muß offenbar

$$C_1 = C_{1,R} + iC_{1,I}, \quad C_2 = C_1^* = C_{1,R} - iC_{1,I} \quad (5.4.9)$$

mit  $C_{1,R}, C_{1,I} \in \mathbb{R}$  gelten. Wir können also (5.4.8) in der Form

$$\begin{aligned} x(t) &= \exp(-\gamma t) \\ &\times \{C_1 \exp(-i\omega t) + [C_1 \exp(-i\omega t)]^*\} \quad (5.4.10) \\ &= \exp(-\gamma t) 2 \operatorname{Re}[C_1 \exp(-i\omega t)] \end{aligned}$$

schreiben.

Mit der Eulerschen Formel folgt

$$\operatorname{Re}[C_1 \exp(-i\omega t)] = \operatorname{Re}[(C_{1,R} + iC_{1,I})(\cos(\omega t) - i\sin(\omega t))] = C_{1,R} \cos(\omega t) + C_{1,I} \sin(\omega t). \quad (5.4.11)$$

Benennen wir die *reellen* Konstanten zu

$$\tilde{C}_1 = 2C_{1,R}, \quad \tilde{C}_2 = 2C_{1,I} \quad (5.4.12)$$

um, erhalten wir als allgemeine *reelle* Lösung

$$x(t) = \exp(-\gamma t) [\tilde{C}_1 \cos(\omega t) + \tilde{C}_2 \sin(\omega t)]. \quad (5.4.13)$$

Die Anfangsbedingungen (5.3.3) gestatten wieder die eindeutige Bestimmung der Integrationskonstanten:

$$x(0) = \tilde{C}_1 \stackrel{!}{=} x_0, \quad \dot{x}(0) = -\gamma \tilde{C}_1 + \tilde{C}_2 \omega \stackrel{!}{=} v_0. \quad (5.4.14)$$

Die Lösung dieses linearen Gleichungssystems ist offenbar

$$\tilde{C}_1 = x_0, \quad \tilde{C}_2 = \frac{v_0 + \gamma x_0}{\omega}, \quad (5.4.15)$$

und die Lösung des Anfangswertproblems lautet also

$$x(t) = \exp(-\gamma t) \left[ x_0 \cos(\omega t) + \frac{v_0 + \gamma x_0}{\omega} \sin(\omega t) \right]. \quad (5.4.16)$$

Analog zu unserem Vorgehen oben beim ungedämpften Oszillator können wir die eckige Klammer auch in Form einer einzelnen Cosinus-Funktion gemäß

$$x(t) = \exp(-\gamma t) \hat{x} \cos(\omega t - \varphi_0) \quad (5.4.17)$$

schreiben.

Das Auffinden der Amplitude  $\hat{x}$  und der Phasenverschiebung  $\varphi_0$  wird aber durch die komplexe Rechnung erheblich abgekürzt. Dazu gehen wir von (5.4.10) aus und verwenden die „Polardarstellung“ der komplexen Zahl  $C_1$ :

$$C_1 = |C_1| \exp(i\varphi_0) = 2\hat{x} \exp(i\varphi_0), \quad \hat{x} = |C_1|/2. \quad (5.4.18)$$

Einsetzen in (5.4.10) liefert dann (5.4.17). Mit (5.4.12) und (5.4.15) folgt wegen (5.4.18) sofort gemäß (4.1.17) und (4.1.18)

$$\hat{x} = \sqrt{\tilde{C}_1^2 + \tilde{C}_2^2} = \frac{\sqrt{x_0^2 \omega^2 + (v_0 + \gamma x_0)^2}}{\omega}, \quad \varphi_0 = \operatorname{sign}(v_0 + \gamma x_0) \arccos\left(\frac{x_0}{\hat{x}}\right) \quad (5.4.19)$$

ist. Wir haben also insgesamt eine periodische Schwingung mit der durch die Dämpfung verringerten Kreisfrequenz  $\omega = \sqrt{\omega_0^2 - \gamma^2}$ , deren Amplitude exponentiell abfällt. In der **Dämpfungszeit**  $t_d = 1/\gamma$  verringert sich die Amplitude um einen Faktor  $1/e = \exp(-1) \approx 1/2.718$ . Der Massepunkt bewegt sich stets innerhalb der Einhüllenden  $\pm \hat{x} \exp(-\gamma t)$  (vgl. Abb 5.3).

### 5.4.2 Kriechfall ( $\omega_0 < \gamma$ )

In diesem Fall besitzt die quadratische Gleichung (5.4.6) zwei *verschiedene* reelle Lösungen

$$\lambda_{1,2} = -\gamma_{1,2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}, \quad (5.4.20)$$

und wir haben direkt die allgemeine Lösung der Bewegungsgleichung in der reellen Form

$$x(t) = C_1 \exp(-\gamma_1 t) + C_2 \exp(-\gamma_2 t). \quad (5.4.21)$$

Da offenbar  $\gamma_2 > \gamma_1 > 0$  ist, ist  $x(t) \rightarrow 0$  für  $t \rightarrow \infty$ . Die Dämpfung ist hierbei so stark, daß es zu keinerlei Schwingungen kommt. Der Massepunkt läuft gegen den Gleichgewichtspunkt bei  $x = 0$ . Für große  $t$  dominiert der Term mit der kleineren Dämpfungskonstante  $\gamma_1$ .

Die Anfangsbedingungen (5.3.3) liefern für die Integrationskonstanten  $C_1$  und  $C_2$  die Gleichungen

$$x(0) = C_1 + C_2 \stackrel{!}{=} x_0, \quad \dot{x}(0) = -C_1 \gamma_1 - C_2 \gamma_2 \stackrel{!}{=} v_0. \quad (5.4.22)$$

Die Lösung des linearen Gleichungssystems ergibt

$$C_1 = -\frac{v_0 + \gamma_2 x_0}{\gamma_1 - \gamma_2} = +\frac{v_0 + \gamma_2 x_0}{2\sqrt{\gamma^2 - \omega_0^2}}, \quad C_2 = \frac{v_0 + \gamma_1 x_0}{\gamma_1 - \gamma_2} = -\frac{v_0 + \gamma_1 x_0}{2\sqrt{\gamma^2 - \omega_0^2}}. \quad (5.4.23)$$

Die Lösung für das Anfangsproblem lautet also

$$x(t) = \frac{1}{2\sqrt{\gamma^2 - \omega_0^2}} [(v_0 + \gamma_2 x_0) \exp(-\gamma_1 t) - (v_0 + \gamma_1 x_0) \exp(-\gamma_2 t)]. \quad (5.4.24)$$

Setzt man hierin für  $\gamma_1$  und  $\gamma_2$  (5.4.20) ein, erhält man das Resultat in der Form

$$x(t) = \exp(-\gamma t) \left[ x_0 \cosh(\sqrt{\gamma^2 - \omega_0^2} t) + \frac{v_0 + \gamma x_0}{\sqrt{\gamma^2 - \omega_0^2}} \sinh(\sqrt{\gamma^2 - \omega_0^2} t) \right]. \quad (5.4.25)$$

Dabei haben wir die Definition der Hyperbelfunktionen

$$\cosh z = \frac{\exp z + \exp(-z)}{2}, \quad \sinh z = \frac{\exp z - \exp(-z)}{2} \quad (5.4.26)$$

benutzt, die übrigens nicht nur im Reellen sondern auch für alle  $z \in \mathbb{C}$  gilt. Daraus folgt (wieder für  $z \in \mathbb{C}$ )

$$\cosh(iz) = \frac{\exp(iz) + \exp(-iz)}{2} = \cos z, \quad \sinh(iz) = \frac{\exp(iz) - \exp(-iz)}{2} = i \sin z. \quad (5.4.27)$$

Wir können also die Lösung für den Schwingfall (5.4.8) gewinnen, indem wir einfach in (5.4.25) überall

$$\sqrt{\gamma^2 - \omega_0^2} = i\sqrt{\omega_0^2 - \gamma^2} = i\omega \quad (5.4.28)$$

setzen. Denn dann liefert die Anwendung der Formeln (5.4.27) in (5.4.25) in der Tat sofort (5.4.8).

### 5.4.3 Aperiodischer Grenzfall ( $\omega_0 = \gamma$ )

Hier hat die quadratische Gleichung (5.4.6) nur die eine (reelle) Lösung

$$\lambda = -\gamma = -\omega_0. \quad (5.4.29)$$

Wir erhalten also auch mit dem Exponentialansatz (5.4.3) nur eine Lösung. Um die vollständige Lösung der Bewegungsgleichung zu finden, benötigen wir allerdings noch eine zweite linear unabhängige Lösung. Statt direkt eine solche Lösung zu konstruieren, ist es einfacher, für festgehaltene Anfangsbedingungen den Limes  $\omega_0 \rightarrow \gamma$  von (5.4.25) zu berechnen. Dabei macht nur der zweite Term in der eckigen Klammer Probleme, weil dessen Nenner im Limes dort verschwindet. Hier führt die Potenzreihenentwicklung der sinh-Funktion zum Ziel. Es gilt

$$\sinh z = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \sum_{j=0}^{\infty} \frac{z^{2j+1}}{(2j+1)!}. \quad (5.4.30)$$

Auf den fraglichen Term in (5.4.25) angewandt, ergibt diese Reihenentwicklung

$$\frac{\sinh(\sqrt{\gamma^2 - \omega_0^2} t)}{\sqrt{\gamma^2 - \omega_0^2}} = \frac{\sqrt{\gamma^2 - \omega_0^2} t + \mathcal{O}(\sqrt{\gamma^2 - \omega_0^2}^3)}{\sqrt{\gamma^2 - \omega_0^2}} = t + \mathcal{O}\left[\left(\sqrt{\gamma^2 - \omega_0^2}\right)^2\right]. \quad (5.4.31)$$

Die übrigen Terme in (5.4.25) sind unproblematisch, und wir können für diese einfach  $\omega_0 = \gamma$  setzen, weil alle vorkommenden Funktionen stetig sind. Wenden wir also (5.4.31) in (5.4.25) an und führen den Grenzwert  $\omega_0 \rightarrow \gamma$  aus, erhalten wir schließlich

$$x(t) = [x_0 + (v_0 + \gamma x_0)t] \exp(-\gamma t). \quad (5.4.32)$$

Man prüft leicht nach, daß diese Funktion tatsächlich die Bewegungsgleichung (5.4.2) für  $\omega_0 = \gamma$  löst und die Anfangsbedingungen (5.3.3) erfüllt.

Auch hier fällt die Auslenkung des Massenpunktes unabhängig von der Anfangsgeschwindigkeit stets zum Gleichgewichtspunkt  $x = 0$  ab, denn sowohl die Funktion  $\exp(-\gamma t)$  als auch  $t \exp(-\gamma t)$  streben gegen 0 für  $t \rightarrow \infty$ . Es zeigt sich, daß insgesamt das Abklingen der Anfangsauslenkung und -geschwindigkeit im aperiodischen Grenzfall am schnellsten vonstatten geht. Dies hat praktische Bedeutung für die Konstruktion von Meßinstrumenten wie (analogen) Galvanometern zur Spannungs- und Strommessung in der Elektrotechnik.

### 5.4.4 Direkte Lösung im aperiodischen Grenzfall

Alternativ zu der im vorigen Abschnitt vorgestellten Methode, den aperiodischen Grenzfall des gedämpften Oszillators als Grenzwert für den überdämpften Fall zu behandeln, können wir auch die Differentialgleichung direkt lösen. Setzen wir also in (5.4.2)  $\omega_0 = \gamma$ , erhalten wir

$$\ddot{x} + 2\gamma\dot{x} + \gamma^2 x = 0. \quad (5.4.33)$$

Der Exponentialansatz (5.4.3) liefert nur die eine Lösung für  $\lambda = -\gamma$ . Um die vollständige Lösung, also eine zweite linear unabhängige Lösung zu finden, machen wir stattdessen den Ansatz („**Variation der Konstanten**“)

$$x(t) = f(t) \exp(-\gamma t). \quad (5.4.34)$$

Die Ableitungen sind wegen der Produktregel

$$\dot{x} = (\dot{f} - \gamma f) \exp(-\gamma t), \quad \ddot{x} = (\ddot{f} - 2\gamma\dot{f} + \gamma^2 f) \exp(-\gamma t). \quad (5.4.35)$$

Setzen wir also unseren Ansatz in (5.4.33) ein, erhalten wir

$$\exp(-\gamma t)[\ddot{f} - 2\gamma\dot{f} + \gamma^2 f + 2\gamma(\dot{f} - \gamma f) + \gamma^2 f] = \exp(-\gamma t)\ddot{f} \stackrel{!}{=} 0. \quad (5.4.36)$$

Das bedeutet, daß  $f$  die Differentialgleichung

$$\ddot{f} = 0 \quad (5.4.37)$$

erfüllen muß. Ihre allgemeine Lösung finden wir durch zweimaliges Integrieren nach der Zeit. Es ergibt sich

$$f(t) = C_1 + C_2 t, \quad (5.4.38)$$

und demnach lautet die allgemeine Lösung für die Differentialgleichung (5.4.33)

$$x(t) = (C_1 + C_2 t)\exp(-\gamma t). \quad (5.4.39)$$

Anpassung der Integrationskonstanten  $C_1$  und  $C_2$  an die Anfangsbedingungen (5.3.3) liefert wieder die bereits oben gefundene Lösung (5.4.32).

## 5.5 Der getriebene gedämpfte Oszillator

Wir schließen unsere Betrachtung der harmonischen Oszillatoren mit der Behandlung der Bewegungsgleichung für den Fall, daß zusätzlich zur Reibungs- und harmonischen Kraft noch eine zeitabhängige äußere **treibende Kraft** an dem Massenpunkt angreift. Dabei beschränken wir uns auf den Fall einer **harmonischen Zeitabhängigkeit der äußeren Kraft**. Die Bewegungsgleichung lautet dann

$$m\ddot{x} = -m\omega_0^2 x - 2m\gamma\dot{x} + mA \cos(\Omega t). \quad (5.5.1)$$

Dies in die Normalform für lineare Differentialgleichungen 2. Ordnung gebracht liefert die **inhomogene Gleichung**

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = A \cos(\Omega t). \quad (5.5.2)$$

Wir bemerken als erstes, daß wegen der Linearität des Differentialoperators auf der linken Seite die Differenz zweier Lösungen dieser inhomogenen Gleichung wieder die homogene Gleichung löst. Die allgemeine Lösung der inhomogenen Gleichung ist also durch die Summe aus der allgemeinen Lösung der homogenen Gleichung und einer beliebigen speziellen Lösung der inhomogenen Gleichung gegeben:

$$x(t) = C_1 x_1^{(\text{hom})}(t) + C_2 x_2^{(\text{hom})}(t) + x^{(\text{inh})}(t). \quad (5.5.3)$$

Dabei sind  $x_1$  und  $x_2$  beliebige zueinander linear unabhängige Lösungen der homogenen Differentialgleichung, die wir im vorigen Abschnitt für die drei Fälle (Schwingfall, Kriechfall, aperiodischer Grenzfall) gefunden haben:

$$\begin{cases} x_1(t) = \exp(-\gamma t) \cos(\omega t), & x_2(t) = \exp(-\gamma t) \sin(\omega t) & \text{für } \omega_0 > \gamma, \\ x_1(t) = \exp(-\gamma_1 t), & x_2(t) = \exp(-\gamma_2 t) & \text{für } \omega_0 < \gamma, \\ x_1(t) = \exp(-\gamma t), & x_2(t) = t \exp(-\gamma t) & \text{für } \omega_0 = \gamma. \end{cases} \quad (5.5.4)$$

Dabei ist im Schwingfall  $\omega = \sqrt{\omega_0^2 - \gamma^2}$  und im Kriechfall  $\gamma_{12} = \gamma \pm \sqrt{\gamma^2 - \omega_0^2}$ .

Diese Lösungen werden allesamt für  $t \gg 1/\gamma$  (bzw. im Kriechfall für  $t \gg 1/\gamma_2$ ) exponentiell weggedämpft werden. Für große Zeiten wird die Lösung also durch die spezielle Lösung der inhomogenen Gleichung dominiert. Man spricht vom **eingeschwungenen Zustand**, und wir interessieren uns im folgenden für diesen Zustand.

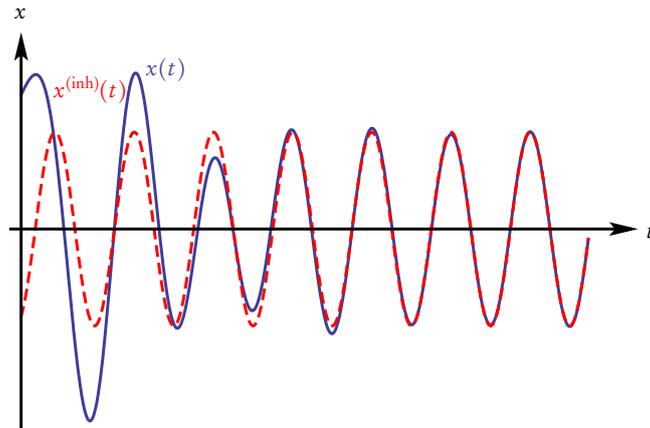


Abbildung 5.4: Lösung zum getriebenen gedämpften harmonischen Oszillator im Schwingfall  $\omega_0 > \gamma$ . Für  $t \gg 1/\gamma$  werden die Eigenschwingungen, also der Anteil der Lösung der homogenen Gleichung in (5.5.3) merklich gedämpft, und die Bewegung geht in den durch die spezielle Lösung der inhomogenen Schwingung eingeschwingenen Zustand über.

### 5.5.1 Spezielle Lösung der inhomogenen Gleichung

Das Auffinden einer speziellen Lösung für die inhomogene Gleichung wird nun dadurch wesentlich erleichtert, daß sowohl die unabhängige Variable, die Zeit  $t$ , als auch die Koeffizienten auf der linken Seite in (5.5.2) reelle Zahlen sind. Wir können also die linke Seite der Gleichung (5.5.2) als Realteil derselben Gleichung einer komplexen Funktion  $z(t) = x(t) + iy(t)$  ansehen:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = \operatorname{Re}(\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z). \quad (5.5.5)$$

Die rechte Seite der Gleichung können wir aber auch als Realteil ausdrücken, denn wegen der Eulerschen Formel gilt

$$A \cos(\Omega t) = \operatorname{Re}[A \exp(-i\Omega t)]. \quad (5.5.6)$$

Wir können also die etwas einfacher zu lösende komplexe Gleichung

$$\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z = A \exp(-i\Omega t) \quad (5.5.7)$$

betrachten. Die Wahl des negativen Vorzeichens im Exponenten auf der rechten Seite ist dabei willkürlich und entspricht der in den meisten Lehrbüchern verwendeten Konvention in der theoretischen Physik.

Jedenfalls legt die Form der Gleichung (5.5.7) den Ansatz

$$z(t) = AB \exp(-i\Omega t) \quad (5.5.8)$$

nahe. Dabei ist  $B$  eine zu bestimmende komplexwertige Konstante. Setzen wir also (5.5.8) in (5.5.7) ein, finden wir

$$AB \exp(-i\Omega t)(-\Omega^2 - 2i\gamma\Omega + \omega_0^2) = A \exp(-i\Omega t). \quad (5.5.9)$$

Auflösen nach  $B$  liefert

$$B = \frac{1}{\omega_0^2 - \Omega^2 - 2i\gamma\Omega}. \quad (5.5.10)$$

Um nun einfacher  $x(t) = \operatorname{Re} z(t)$  bestimmen zu können, machen wir noch den Zähler reell, indem wir den Bruch mit dem konjugiert Komplexen des Nenners erweitern:

$$B = \frac{\omega_0^2 - \Omega^2 + 2i\gamma\Omega}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}. \quad (5.5.11)$$

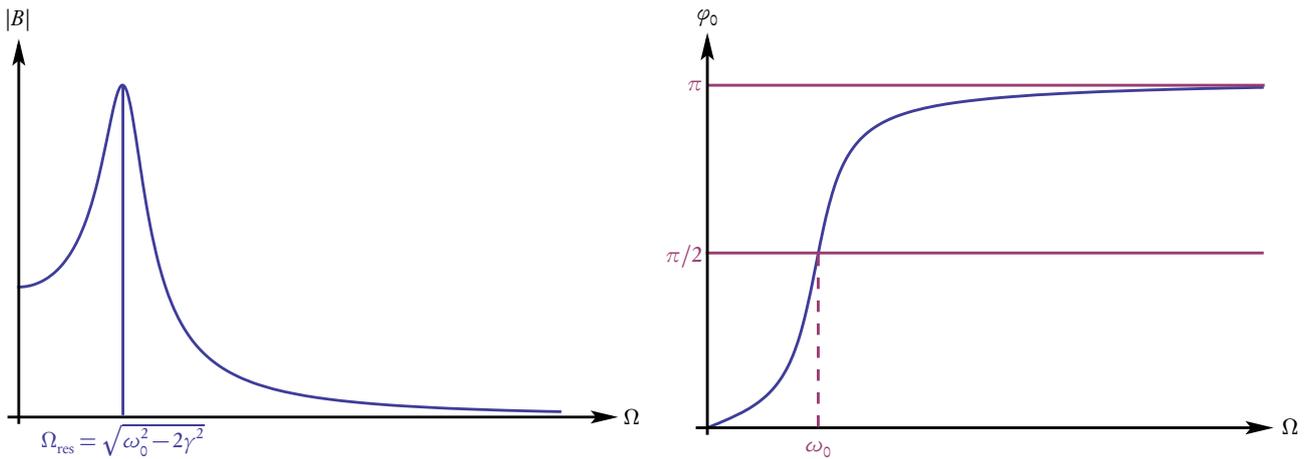


Abbildung 5.5: Amplitudenfaktor (links)  $|B|$  und Phasenverschiebung  $\varphi_0$  für den eingeschwungenen Zustand (5.5.13).

Wir können nun  $B$  in die Polardarstellung bringen:

$$B = |B| \exp(i\varphi_0), \quad |B| = \frac{1}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}}, \quad \varphi_0 = + \arccos\left(\frac{\omega_0^2 - \Omega^2}{|B|}\right). \quad (5.5.12)$$

Setzen wir dies in unseren Ansatz ein, ergibt sich für die spezielle Lösung der inhomogenen Gleichung schließlich

$$x^{(\text{inh})}(t) = \text{Re } z(t) = A|B| \text{Re} \{ \exp[-i(\Omega t - \varphi_0)] \} = A|B| \cos(\Omega t - \varphi_0). \quad (5.5.13)$$

Im eingeschwungenen Zustand schwingt also der Massenpunkt mit derselben Frequenz wie die äußere harmonische Kraft, und zwischen der Kraft und der Schwingung des Massenpunktes besteht eine *stets positive* Phasenverschiebung  $\varphi_0$  (vgl. Abb 5.5, rechts). Als Funktion von  $\Omega$  ist  $\varphi_0$  monoton wachsend. Für  $\Omega = \omega_0$  wird  $\varphi_0 = \pi/2$ , und für  $\Omega \rightarrow \infty$  strebt  $\varphi_0 \rightarrow \pi$ .

Bei vorgegebener Amplitude der äußeren Kraft  $mA$  ist die Amplitude des eingeschwungenen Zustands durch  $|B|$  gemäß (5.5.12) gegeben. In Abb. 5.5 (links) haben wir diesen Proportionalitätsfaktor als Funktion der Kreisfrequenz der antreibenden Kraft  $\Omega$  geplottet. Er weist in dem gewählten Fall ein ausgeprägtes Maximum bei einer **Resonanzfrequenz**  $\Omega_{\text{res}}$  auf, die wir im nächsten Abschnitt ausrechnen wollen.

## 5.5.2 Amplitudenresonanzfrequenz

Die Lösung (5.5.13) zeigt, daß die Amplitude der Schwingung im eingeschwungenen Zustand zur Amplitude der äußeren Kraft proportional ist und der Proportionalitätsfaktor  $|B|$  gemäß (5.5.12) durch die Parameter des gedämpften freien Oszillators, also  $\omega_0$  und  $\gamma$  und die Kreisfrequenz  $\Omega$  der äußeren Kraft allein bestimmt ist (s. Abb 5.5, links).

Wir untersuchen nun, für welches  $\Omega$  dieser Proportionalitätsfaktor maximal wird, d.h. für welche Frequenz der äußeren Kraft die Amplitude der Schwingung bei festgehaltener Amplitude der äußeren Kraft am größten wird.

Dazu müssen wir das Minimum des Ausdrucks unter der Wurzel in (5.5.12) suchen, d.h. wir müssen die Funktion

$$f(\Omega) = (\Omega^2 - \omega_0^2)^2 + 4\gamma^2\Omega^2 \quad (5.5.14)$$

untersuchen. Dazu bilden wir die Ableitung

$$\frac{d}{d\Omega} f(\Omega) = f'(\Omega) = 4\Omega(\Omega^2 - \omega_0^2 + 2\gamma^2). \quad (5.5.15)$$

Mögliche Minima ergeben sich als die Nullstellen dieser Ableitung. Offenbar ist die Lösung entweder  $\Omega = 0$ , d.h. es wirkt eine zeitlich konstante äußere Kraft. Es liegt dann offenbar ein Minimum vor, wenn  $\omega_0^2 < 2\gamma^2$  ist, denn es gilt

$$f''(\Omega) = 4(2\gamma^2 - \omega_0^2 + 3\Omega^2). \quad (5.5.16)$$

Es ist also  $f''(0) = 4(2\gamma^2 - \omega_0^2)$ , und dies ist positiv für  $\omega_0^2 < 2\gamma^2$ .

Falls  $\omega_0^2 > 2\gamma^2$ , nimmt (5.5.15) eine Nullstelle bei der **Resonanzfrequenz**

$$\Omega_{\text{res}} = \sqrt{\omega_0^2 - 2\gamma^2} \quad (5.5.17)$$

an. Dort gilt  $f''(\Omega_{\text{res}}) = 8(\omega_0^2 - 2\gamma^2) > 0$ , und es liegt also auch in diesem Fall ein Minimum für  $f$  und also ein Maximum der Amplitude vor.

Man nennt daher  $\Omega_{\text{res}}$  die **Amplitudenresonanzfrequenz**, weil dies die Frequenz ist, bei der die Amplitude der eingeschwungenen Bewegung maximal wird. Interessanterweise ist sie weder durch die Eigenfrequenz des ungedämpften Oszillators  $\omega_0$  noch durch die Eigenfrequenz der gedämpften Schwingung  $\omega = \sqrt{\omega_0^2 - \gamma^2}$  im Schwingfall  $\omega_0 > \gamma$  gegeben sondern durch die kleinere Amplitudenresonanzfrequenz (5.5.17).

### 5.5.3 Energieresonanz

Eine andere Frage ist es, bei welcher Frequenz der äußeren Kraft, diese die größte mittlere Leistung aufbringen muß, um den Massenpunkt in dem erzwungenen eingeschwungenen Schwingungszustand zu halten. Dazu berechnen wir die über eine Periode  $T = 2\pi/\Omega$  gemittelte Leistung der äußeren Kraft

$$\bar{P} = \frac{1}{T} \int_0^T dt m A \cos(\Omega t) \dot{x}^{(\text{inh})}(t). \quad (5.5.18)$$

Nun ist gemäß (5.5.13)

$$P(t) = m A \cos(\Omega t) \dot{x}^{(\text{inh})}(t) = -m A^2 |B| \Omega \cos(\Omega t) \sin(\Omega t - \varphi_0). \quad (5.5.19)$$

Mit dem Additionstheorem für den Sinus ist

$$P(t) = -m A^2 |B| \Omega \cos(\Omega t) [\sin(\Omega t) \cos \varphi_0 - \cos(\Omega t) \sin \varphi_0]. \quad (5.5.20)$$

Zur einfacheren Integration bemerken wir, daß

$$\cos(\Omega t) \sin(\Omega t) = \frac{1}{2} \sin(2\Omega t), \quad \cos^2(\omega t) = \frac{1}{2} [1 + \cos(2\Omega t)] \quad (5.5.21)$$

gilt, was man sofort durch Anwendung der Additionstheoreme nachweist. Da weiter

$$\int_0^T dt \sin(2\Omega t) = -\frac{\cos(2\Omega t)}{2\Omega} \Big|_{t=0}^{t=T} = 0, \quad \int_0^T dt \cos(2\Omega t) = \frac{\sin(2\Omega t)}{2\Omega} \Big|_{t=0}^{t=T} = 0 \quad (5.5.22)$$

gilt erhalten wir also durch Einsetzen von (5.5.20) in (5.5.18) für die mittlere Leistung

$$\bar{P} = \frac{1}{2} m A^2 |B| \Omega \sin \varphi_0. \quad (5.5.23)$$

Aus (5.5.12) folgt, daß  $\varphi_0 \in [0, \pi]$  und also  $\sin \varphi_0 > 0$

$$\sin \varphi_0 = \sqrt{1 - \cos^2 \varphi_0} = \frac{2\gamma\Omega}{\sqrt{(\Omega^2 - \omega_0^2) + 4\gamma^2\Omega^2}} = 2\gamma\Omega |B|. \quad (5.5.24)$$

Damit wird (5.5.23)

$$\bar{P} = mA^2|B|^2\gamma\Omega^2. \quad (5.5.25)$$

Wir fragen nun, bei welcher Kreisfrequenz  $\Omega$  der äußeren Kraft bei vorgegebenen Parametern des Oszillators und der Amplitude  $A$  der äußeren Kraft maximal wird. Man spricht in diesem Fall von **Energieresonanz**, denn dort ist die mittlere Leistungsaufnahme des Massenpunktes aus der äußeren Kraft maximal. Wir haben also diesmal das Maximum der Funktion

$$g(\Omega) = \Omega^2|B|^2 = \frac{\Omega^2}{(\Omega^2 - \omega_0^2)^2 + 4\gamma^2\Omega^2} \quad (5.5.26)$$

zu suchen. Nach einiger Rechnung findet man für die Ableitung

$$g'(\Omega) = \frac{2\Omega(\Omega^4 - \omega_0^4)}{[(\Omega^2 - \omega_0^2)^2 + 4\gamma^2\Omega^2]^2}. \quad (5.5.27)$$

Offenbar liegt bei  $\Omega = 0$  ein Minimum vor, denn dort ist  $g'$  lokal monoton fallend. Ein Maximum erhalten wir bei  $\Omega = \omega_0$ , denn dort ist  $g'$  offenbar lokal monoton wachsend. Die größte mittlere Leistung wird also vom Oszillator aufgenommen, wenn  $\Omega = \omega_0$  ist, also die Schwingungsfrequenz der äußeren Kraft der Eigenfrequenz des *ungedämpften* Oszillators entspricht. Wie (5.5.12) zeigt, ist dort gerade die Phasenverschiebung der eingeschwungenen Bewegung gegenüber der Phase der antreibenden Kraft  $\varphi_0 = \pi/2$ .

Wir bemerken noch, daß für verschwindende Dämpfung  $\gamma = 0$  bei  $\Omega = \omega_0$  der Faktor  $|B|$  unendlich wird. Das ist die sogenannte **Resonanzkatastrophe**.

#### 5.5.4 Lösung des Anfangswertproblems

Wir kommen schließlich auf die Lösung des Anfangswertproblems für den getriebenen harmonischen Oszillators zurück, die zur vollständigen Beschreibung der Bewegung bei beliebig vorgegebenen Anfangsbedingungen (5.3.3) dient und nicht nur den eingeschwungenen Zustand liefert. Man spricht auch vom **Einschwingvorgang**.

Wir müssen nur für die allgemeine Lösung (5.5.3) die Integrationskonstanten  $C_1$  und  $C_2$  aus den Anfangsbedingungen (5.4.1) durch Lösung des entsprechenden linearen Gleichungssystems zu bestimmen. Wir geben das Ergebnis für die oben diskutierten drei Fälle an

**Schwingfall** ( $\omega_0 > \gamma$ ):

$$\begin{aligned} x(t) = & \left[ x_0 - \frac{A(\omega_0^2 - \Omega^2)}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \cos(\omega t) \right] \exp(-\gamma t) \\ & + \frac{1}{\omega} \left[ v_0 + \gamma x_0 - \frac{A\gamma(\Omega^2 + \omega_0^2)}{(\omega_0^2 - \Omega^2)^2 + 4\gamma\Omega^2} \right] \sin(\omega t) \exp(-\gamma t) \\ & + A \frac{\omega_0^2 - \Omega^2}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \cos(\Omega t) + \frac{2A\gamma\Omega}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \sin(\Omega t). \end{aligned} \quad (5.5.28)$$

**Kriechfall** ( $\omega_0 < \gamma$ ): Setzen wir zur Abkürzung  $\alpha = \sqrt{\gamma^2 - \omega_0^2}$ , erhalten wir für diesen Fall

$$\begin{aligned} x(t) = & \left[ x_0 - \frac{A(\omega_0^2 - \Omega^2)}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \cosh(\alpha t) \right] \exp(-\gamma t) \\ & + \frac{1}{\alpha} \left[ v_0 + \gamma x_0 - \frac{A\gamma(\Omega^2 + \omega_0^2)}{(\omega_0^2 - \Omega^2)^2 + 4\gamma\Omega^2} \right] \sinh(\alpha t) \exp(-\gamma t) \\ & + A \frac{\omega_0^2 - \Omega^2}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \cos(\Omega t) + \frac{2A\gamma\Omega}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2} \sin(\Omega t). \end{aligned} \quad (5.5.29)$$

Zu dieser Lösung können wir auch wieder gelangen, indem wir in (5.5.28)  $\omega = i\alpha$  setzen und die in Formeln (4.2.14)

$$\cos(iz) = \cosh z, \quad \sin(iz) = i \sinh z \quad (5.5.30)$$

verwenden.

**Aperiodischer Grenzfall ( $\omega_0 = \gamma$ ):**

$$x(t) = \exp(-\gamma t) \left[ x_0 - \frac{A(\gamma^2 - \Omega^2)}{(\gamma^2 + \Omega^2)^2} + \left( v_0 + \gamma x_0 - \frac{A\gamma}{\gamma^2 + \Omega^2} \right) t \right] + A \frac{(\gamma^2 - \Omega^2) \cos(\Omega t) + 2\gamma\Omega \sin(\Omega t)}{(\Omega^2 + \gamma^2)^2}. \quad (5.5.31)$$

### 5.5.5 Resonant angetriebener ungedämpfter Oszillator

Wie wir oben gesehen haben, müssen wir den Fall des angetriebenen ungedämpften harmonischen Oszillators gesondert behandeln, wenn die Frequenz  $\Omega$  der antreibenden Kraft der Eigenfrequenz  $\omega_0$  des Oszillators entspricht, da dann wegen  $B$  wegen  $\gamma = 0$  für  $\Omega = \omega_0$  aufgrund von (5.5.11) nicht definiert ist, also der Ansatz (5.5.8) nicht zum Ziel führt.

Wir suchen also eine Beliebige Lösung der inhomogenen Gleichung

$$\ddot{x} + \omega_0^2 x = a \cos(\omega t) = a \operatorname{Re} \exp(-i\omega_0 t). \quad (5.5.32)$$

Dazu suchen wir zunächst die Lösung der entsprechenden komplexifizierten Gleichung

$$\ddot{z} + \omega_0^2 z = a \cos(\omega t) = a \exp(-i\omega_0 t). \quad (5.5.33)$$

Da (5.5.8) nicht zum Ziel führt, machen wir den etwas abgewandelten „Ansatz vom Typ der rechten Seite“

$$z(t) = a f(t) \exp(-i\omega_0 t) \quad (5.5.34)$$

mit dem Ziel,  $f$  zu bestimmen. Dazu setzen wir den Ansatz in (5.5.33) ein. Nach einer einfachen Umformung führt dies auf die lineare inhomogene Differentialgleichung mit konstanten Koeffizienten

$$\ddot{f} - 2i\omega_0 \dot{f} = 1. \quad (5.5.35)$$

Wir benötigen nur eine Lösung dieser Gleichung. Offenbar führt hier der Ansatz

$$f(t) = C t \quad (5.5.36)$$

zum Ziel. Setzt man dies nämlich in (5.5.35) ein, erhält man

$$-2i\omega_0 C = 1 \Rightarrow C = \frac{i}{2\omega_0}. \quad (5.5.37)$$

Wir erhalten also eine Lösung der Gl. (5.5.33) durch Einsetzen von (5.5.36) und (5.5.37) in den Ansatz (5.5.34):

$$z(t) = \frac{ia}{2\omega_0} t \exp(-i\omega_0 t) \quad (5.5.38)$$

und damit eine Lösung von (5.5.32)

$$x(t) = \operatorname{Re}[z(t)] = \operatorname{Re} \left[ \frac{ia}{2\omega_0} t \exp(-i\omega_0 t) \right] = \frac{a}{2\omega_0} t \sin(\omega_0 t). \quad (5.5.39)$$

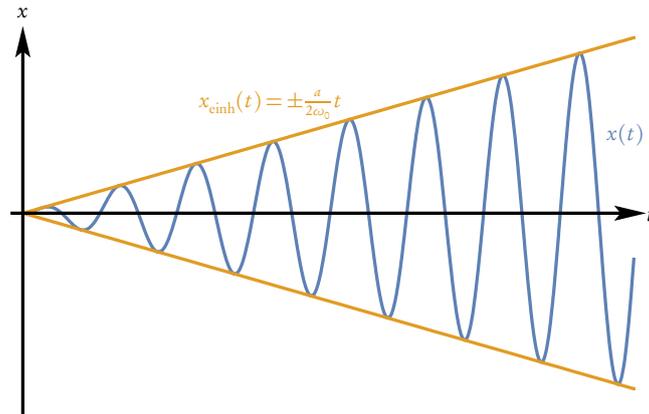


Abbildung 5.6: Lösung zum resonant angetriebenen und gedämpften harmonischen Oszillator mit den Anfangsbedingungen  $x(0) = 0$ ,  $\dot{x}(0) = 0$ . Dann ergibt sich die Lösung zu (5.5.39). Die Einhüllende ist durch die Geraden  $x_{\text{einh}}(t) = \pm \frac{a}{2\omega_0} t$  gegeben.

Die allgemeine Lösung von (5.5.32) ergibt sich also aus der Überlagerung der allgemeinen Lösung der entsprechenden homogenen Gleichung, deren Lösung wir oben in (5.3.19) gefunden haben:

$$x(t) = \frac{a}{2\omega_0} t \sin(\omega_0 t) + C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t). \quad (5.5.40)$$

Die Integrationskonstanten  $C_1$  und  $C_2$  bestimmen sich wieder aus den Anfangsbedingungen. Unabhängig von diesen wächst die Amplitude beständig an. Man spricht daher auch von der **Resonanzkatastrophe** (vgl. Abb. 5.6). Natürlich kommt so etwas in der Natur nicht wirklich vor, da i.a. die Kräfte für große Auslenkungen nicht mehr harmonisch sind und daher die Beschreibung als harmonischer Oszillator für große Zeiten ungültig wird.

### 5.5.6 Allgemeine äußere Kräfte und die $\delta$ -Distribution

Zum Abschluß des Kapitels über den harmonischen Oszillator wollen wir noch die Frage beantworten, wie man den Fall beliebig vorgegebener äußerer Kräfte behandeln kann. Wir benötigen dazu nur eine spezielle Lösung der inhomogenen Gleichung

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = a(t), \quad (5.5.41)$$

wobei  $a$  eine vorgegebene Funktion ist (die äußere Kraft ist  $F(t) = ma(t)$ ). Die Idee zu einer Lösung ist, daß sich die Auslenkung  $x(t)$  zur Zeit  $t$  als eine Art „kontinuierlicher Linearkombination“ der rechten Seite der Gleichung ergibt. Das physikalische Bild hinter dieser Idee ist, daß zu jedem Zeitpunkt die äußere Kraft den harmonischen Oszillator neu anstößt. Diese Überlegung führt uns auf den Ansatz

$$x(t) = \int_{-\infty}^{\infty} dt' G(t, t') a(t'). \quad (5.5.42)$$

Das **Kausalitätsprinzip** verlangt nun, daß die Auslenkung zur Zeit  $t$  nur von der äußeren Kraft zu früheren Zeitpunkten  $t' < t$  abhängen darf. MaW. die äußere Kraft zu einer späteren Zeit kann nicht in die Vergangenheit zurückwirken. Dies verlangt, daß

$$G(t, t') = 0 \quad \text{für} \quad t' > t. \quad (5.5.43)$$

Die untere Integrationsgrenze, also den Anfangszeitpunkt der Bewegung, haben wir bei  $t_0 \rightarrow -\infty$  gewählt, damit wir genau die Lösung erhalten, für die die von spezifischen Anfangsbedingungen abhängigen Einschwingvorgänge zu jedem endlichen Zeitpunkt bereits abgeklungen sind.

Wir setzen nun diesen Ansatz in (5.5.41) ein. Das führt auf die Gleichung

$$\int_{-\infty}^{\infty} dt' [\partial_t^2 G(t, t') + 2\gamma \partial_t G(t, t') + \omega_0^2 G(t, t')] a(t') = a(t). \quad (5.5.44)$$

Dies verlangt nun, daß der Ausdruck in den eckigen Klammern eine „Funktion“  $\delta(t, t')$  ist, die für beliebige stetige Funktionen  $a$  stets<sup>1</sup>

$$\int_{-\infty}^{\infty} dt' \delta(t, t') a(t') = a(t) \quad (5.5.45)$$

gilt. Man kann zeigen, daß es eine solche Funktion im strengen Sinne nicht gibt. Wir können aber Näherungen für eine solche Funktion definieren. Die einfachste Näherung ist,

$$\delta_\epsilon(t, t') = \begin{cases} \frac{1}{\epsilon} & \text{falls } t' \in (t - \epsilon/2, t + \epsilon/2), \\ 0 & \text{sonst,} \end{cases} \quad (5.5.46)$$

wobei  $\epsilon > 0$  beliebig ist. Mit dem Mittelwertsatz der Integralrechnung folgt dann

$$\int_{-\infty}^{\infty} dt' \delta_\epsilon(t, t') a(t') = \frac{1}{\epsilon} \int_{t-\epsilon/2}^{t+\epsilon/2} dt' a(t') = \frac{1}{\epsilon} a(\tau) \epsilon = a(\tau), \quad \tau \in (t - \epsilon/2, t + \epsilon/2). \quad (5.5.47)$$

Lassen wir nun  $\epsilon \rightarrow 0^+$  gehen, erhalten wir wegen der Stetigkeit von  $a$  im Limes für das Integral tatsächlich  $a(t)$ ; die Funktionen  $\delta_\epsilon$  streben allerdings nicht gegen eine gewöhnliche Funktion. Anschaulich ergibt sich eine Funktion, die überall verschwindet außer an der Stelle  $t' = t$ , wo sie unendlich wird. Der Grenzwert  $\epsilon \rightarrow 0^+$  ergibt nur im Sinne des Integrals (5.5.45) einen Sinn, d.h. man führt *zuerst* die Integration in (5.5.45) mit der regularisierten Funktion  $\delta_\epsilon$  aus und läßt *dann*  $\epsilon \rightarrow 0^+$  streben. Dies nennt man einen **Grenzwert im schwachen Sinne** und den Grenzwert  $\delta(t, t')$  eine **verallgemeinerte Funktion** oder **Distribution**<sup>2</sup>.

Offensichtlich können wir die Schreibweise für die  $\delta$ -Distribution noch etwas vereinfachen. Dazu definieren wir zunächst

$$\delta_\epsilon(t) = \begin{cases} \frac{1}{\epsilon} & \text{für } t \in (-\epsilon/2, \epsilon/2) \\ 0 & \text{sonst.} \end{cases} \quad (5.5.48)$$

Dann gilt offensichtlich (*nachprüfen!*)

$$\delta_\epsilon(t, t') = \delta_\epsilon(t - t'), \quad (5.5.49)$$

und entsprechend schreiben wir auch

$$\delta(t, t') = \delta(t - t'). \quad (5.5.50)$$

Dann verlangt offenbar (5.5.44), daß

$$\partial_t^2 G(t, t') + 2\gamma \partial_t G(t, t') + \omega_0^2 G(t, t') = \delta(t - t'). \quad (5.5.51)$$

Daraus folgt, daß man auch mit dem Ansatz  $G(t, t') = G(t - t')$  auskommen sollte und daß dann

$$\ddot{G}(t) + 2\gamma \dot{G}(t) + \omega_0^2 G(t) = \delta(t) \quad (5.5.52)$$

gelten sollte.

Wir suchen nun also eine Funktion, die die folgenden Eigenschaften besitzt:

<sup>1</sup>Im mathematisch strikten Sinne muß man an die möglichen Funktionen  $a$  noch weitere Forderungen stellen, insbesondere, daß sie im Unendlichen hinreichend schnell verschwinden, damit die uneigentlichen Integrale in unserer Betrachtung existieren. Wir gehen darauf hier nicht genauer ein.

<sup>2</sup>Diese Idee wurde u.a. von P. A. M. Dirac im Zusammenhang mit seinen Pionierarbeiten zur Quantenmechanik in die theoretische Physik eingeführt. Die Mathematiker kritisierten anfangs diese Idee heftig, weil sie mathematisch nicht streng begründet war. Die Nützlichkeit der Idee in der praktischen Anwendung führte schließlich die Mathematiker dazu, diese Idee zu formalisieren. Daraus ist ein ganzer Zweig der modernen Mathematik entstanden, die **Funktionalanalysis**.

- $G(t) = 0$  für  $t < 0$  (Kausalitäts- oder **Retardierungsbedingung**)
- $G$  erfüllt für  $t > 0$  die homogene Differentialgleichung

Betrachten wir der Einfachheit halber den Schwingfall. Dann lautet die allgemeine Lösung für  $t > 0$  gemäß (5.4.13)

$$G(t) = \begin{cases} \exp(-\gamma t)[A \cos(\omega t) + B \sin \omega t] & \text{für } t > 0, \\ 0 & \text{für } t \leq 0. \end{cases} \quad (5.5.53)$$

Wir müssen nun die beiden Integrationskonstanten  $A$  und  $B$  bestimmen, so daß die Singularität der  $\delta$ -Funktion bei  $t = 0$  in (5.5.52) auftritt. Wir benötigen offenbar zwei Bedingungen, um beide Konstanten eindeutig bestimmen zu können. Dazu überlegen wir uns, daß Ableitungen von Funktionen i.a. mehr Singularitäten besitzen als die abgeleitete Funktion. Eine stetige Funktion, die an der Stelle  $t = 0$  einen Knick besitzt, ist dort sicher nicht differenzierbar, und wenn sie überall sonst differenzierbar ist, macht die Ableitungsfunktion bei  $t = 0$  einen Sprung, weil die Steigung der Tangenten an der Knickstelle sich abrupt ändert. Wir erwarten also, daß die Singularität in (5.5.52) bei  $t = 0$  in der höchsten vorkommenden Ableitung, also  $\dot{G}$  auftreten sollte, während  $G$  dort lediglich einen endlichen Sprung und  $G$  sogar stetig sein sollte (aber einen Knick bei  $t = 0$  aufweisen sollte).

Als erste Bedingung zur Bestimmung der Integrationskonstanten in (5.5.53) verlangen wir also, daß  $G$  stetig bei  $t = 0$  sein soll. Das führt auf

$$A = 0 \Rightarrow G(t) = B \exp(-\gamma t) \sin(\omega t) \quad \text{für } t > 0. \quad (5.5.54)$$

Schließlich benötigen wir noch eine zweite Bedingung. Dazu integrieren wir (5.5.52) über ein kleines Intervall  $(-\epsilon, \epsilon)$ . Für  $t < 0$  ist  $G(t) = \dot{G}(t) = \ddot{G}(t) = 0$ . Daraus ergibt sich beim Integrieren

$$\dot{G}(\epsilon) + 2\gamma G(\epsilon) + \omega_0^2 \int_{-\epsilon}^{\epsilon} dt G(t) = 1. \quad (5.5.55)$$

Lassen wir nun  $\epsilon \rightarrow 0^+$  streben, verschwindet im Limes sowohl  $G(\epsilon)$  als auch das Integral wegen der angenommenen Stetigkeit von  $G$  bei  $t = 0$ . Es ergibt sich zur Bestimmung der verbliebenen Integrationskonstante also die Bedingung

$$\dot{G}(0^+) = 1. \quad (5.5.56)$$

Nun folgt aus (5.5.55) für  $\epsilon \rightarrow 0^+$  für  $t > 0$

$$\dot{G}(t) = B \exp(-\gamma t)[- \gamma \sin \omega t + \omega \cos \omega t] \rightarrow B \omega \quad \text{für } t \rightarrow 0^+. \quad (5.5.57)$$

Aus (5.5.56) folgt damit  $B = 1/\omega$ , und

$$G(t) = \begin{cases} \frac{\exp(-\gamma t)}{\omega} \sin(\omega t) & \text{for } t > 0, \\ 0 & \text{for } t \leq 0. \end{cases} \quad (5.5.58)$$

Diese Funktion nennt man die **retardierte Greensche Funktion** für den linearen Differentialoperator

$$\hat{L} = \frac{d^2}{dt^2} + 2\gamma \frac{d}{dt} + \omega_0^2. \quad (5.5.59)$$

Die ursprüngliche Differentialgleichung (5.5.41) können wir dann in der Form

$$\hat{L}x(t) = a(t) \quad (5.5.60)$$

und die allgemeine Lösung gemäß (5.5.42)

$$x(t) = \exp(-\gamma t)[C_1 \cos(\omega t) + C_2 \sin(\omega t)] + \int_{-\infty}^{\infty} dt' G(t-t')a(t'). \quad (5.5.61)$$

Dabei ist der erste Term die allgemeine Lösung der homogenen Differentialgleichung  $\hat{L}x(t) = 0$  und das Integral die spezielle Lösung der inhomogenen Gleichung.

Wir können nun zeigen, daß dies tatsächlich eine Lösung ist. Dazu bemerken wir, daß das Integral effektiv nur bis  $t$  läuft, da für  $t' > t$  die Green-Funktion konstruktionsgemäß wegen der Retardierungsbedingung verschwindet. Wir haben also für die Lösung der inhomogenen Gleichung

$$\begin{aligned} x_{\text{inh}}(t) &= \int_{-\infty}^t dt' G(t-t')a(t') \\ \dot{x}_{\text{inh}}(t) &= G(0^+)a(t) + \int_{-\infty}^t dt' \partial_t G(t-t')a(t') = \int_{-\infty}^t dt' \partial_t G(t-t')a(t'), \\ \ddot{x}_{\text{inh}}(t) &= \partial_t G(0^+)a(t) + \int_{-\infty}^t dt' \partial_t^2 G(t-t')a(t'). \end{aligned} \quad (5.5.62)$$

Konstruktionsgemäß ist nun  $\partial_t G(0^+) = 1$ , und damit finden wir schließlich

$$\hat{L}x(t) = a(t) + \int_{-\infty}^t dt' [\hat{L}G(t-t')]a(t') = a(t), \quad (5.5.63)$$

denn für  $t' < t$  erfüllt  $G(t-t')$  als Funktion von  $t$  die homogene Differentialgleichung  $\hat{L}G(t-t') = 0$ . Dies zeigt, daß unsere mathematisch recht unscharf begründeten Manipulationen mit der  $\delta$ -Funktion tatsächlich zum Ziel führen, die inhomogene Differentialgleichung für beliebige rechte Seiten zu lösen.

Dieses Konzept der Green-Funktionen spielt eine weit größere Rolle bei der Lösung partieller Differentialgleichungen wie sie in der Feldtheorie (z.B. der Elektrodynamik im 3. Semester) auftreten als in der klassischen Mechanik.



## Anhang A

---

# Zusammenfassung der Analysis für reelle Funktionen einer Variablen

In diesem Kapitel stellen wir einige Grundlagen zusammen, die wir im folgenden voraussetzen wollen. Dies umfaßt den Umgang mit Mengen und reellen Zahlen sowie die Analysis für Funktionen einer reellen Veränderlichen. Im Bedarfsfall sollte dieser Stoff der Schulmathematik selbständig nachgearbeitet werden.

### A.1 Mengen und reelle Zahlen

Die moderne Mathematik versteht sich als die Lehre über abstrakte Strukturen und basiert (nahezu) vollständig auf der Mengenlehre, die wir hier in einer sehr naiven Auffassung verwenden. Demnach ist eine **Menge** einfach eine Zusammenfassung irgendwelcher (realer oder abstrakter) Gegenstände. Eine Menge  $M$  ist demnach dadurch definiert, daß man von beliebigen Gegenstand  $x$  sagen kann, ob er zur Menge gehört ( $x \in M$ : „ $x$  ist in  $M$  enthalten“ oder „ $x$  ist Element der Menge  $M$ “) oder nicht ( $x \notin M$ ).

Eine Menge  $M$  kann zum einen durch einfache Aufzählung der in ihr enthaltenen Elemente definiert werden. Z.B. definiert man als die Menge der natürlichen Zahlen  $\mathbb{N} = \{1, 2, 3, \dots\}$  oder  $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ . Sie kann aber auch durch die Eigenschaft ihrer Elemente definiert sein. Z.B. kann man die Menge aller natürlichen Zahlen, die kleiner als 7 sind zum einen einfach durch Aufzählung  $M = \{1, 2, 3, 4, 5, 6\}$  oder durch die betreffende Eigenschaft der Elemente  $M = \{n \in \mathbb{N} | n < 7\} = \{n \in \mathbb{N} | n \leq 6\}$ , was „alle natürlichen Zahlen  $n$  mit der Eigenschaft  $n < 7$  bzw.  $n \leq 6$ “ bedeutet, definiert werden.

Einige nützliche Notationen sind noch die **Teilmenge** bzw. die **Obermenge**. Man sagt eine Menge  $M'$  ist Teilmenge der Menge  $M$ ,  $M' \subseteq M$ , wenn aus  $x \in M'$  stets folgt, daß auch  $x \in M$  ist. Man sagt in diesem Fall auch, daß  $M$  Obermenge von  $M'$  ist,  $M \supseteq M'$ .

Auch das Rechnen mit reellen Zahlen wollen wir als bekannt voraussetzen. Wir deuten nur kurz einige Grundlagen an. Die reellen Zahlen werden im wesentlichen durch die Rechenregeln für Addition und Multiplikation und deren jeweilige Umkehrungen, also Subtraktion und Division definiert. Diese algebraischen Regeln definieren, was die Mathematiker als **Zahlenkörper** bezeichnen.

Zunächst bildet die Menge der **reellen Zahlen**  $\mathbb{R}$  zusammen mit der Addition als **Abbildung** zweier reeller Zahlen auf eine reelle Zahl eine **Abelsche Gruppe**, d.h. es gelten die folgenden „Rechenregeln“.

1. Für alle  $a, b, c \in \mathbb{R}$  gilt für die Addition stets das **Assoziativgesetz**

$$(a + b) + c = a + (b + c). \quad (\text{A.1.1})$$

2. Es existiert genau eine Zahl  $0 \in \mathbb{R}$ , so daß für alle  $a \in \mathbb{R}$

$$a + 0 = a \quad (\text{A.1.2})$$

gilt. Diese Zahl 0 (Null) ist das **neutrale Element der Addition**.

3. Zu jeder Zahl  $a \in \mathbb{R}$  existiert eine Zahl  $(-a) \in \mathbb{R}$ , so daß

$$a + (-a) = 0. \quad (\text{A.1.3})$$

Man nennt  $(-a)$  das inverse Element zu  $a$  bzgl. der Addition.

4. Für alle  $a, b \in \mathbb{R}$  gilt für die Addition das **Kommutativgesetz**

$$a + b = b + a. \quad (\text{A.1.4})$$

Man schreibt abkürzend auch  $a + (-b) = a - b$ , d.h. die Subtraktion ist auf die Addition und Bildung des inversen Elementz bzgl. der Addition zurückgeführt.

Es gibt noch eine weitere elementare Verknüpfung, die **Multiplikation** (Abbildung zweier reeller Zahlen auf eine reelle Zahl). Auf der Menge  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  (die reellen Zahlen ohne die Null) bildet auch diese Verknüpfung eine Abelsche Gruppe.

1. Für alle  $a, b, c \in \mathbb{R}$  gilt für die Multiplikation stets das **Assoziativgesetz**

$$(ab)c = a(bc). \quad (\text{A.1.5})$$

2. Es existiert genau eine Zahl  $1 \in \mathbb{R}$ , so daß für alle  $a \in \mathbb{R}$

$$a1 = a \quad (\text{A.1.6})$$

gilt. Diese Zahl 1 (Eins) ist das **neutrale Element der Multiplikation**.

3. Zu jeder Zahl  $a \in \mathbb{R}^*$  existiert genau eine Zahl  $a^{-1} \in \mathbb{R}$ , so daß

$$a(a^{-1}) = 1 \quad (\text{A.1.7})$$

ist. Man nennt  $a^{-1}$  das inverse Element zu  $a$  bzgl. der Multiplikation. Die Null besitzt kein inverses Element bzgl. der Multiplikation („durch Null kann man nicht teilen!“)

4. Für alle  $a, b \in \mathbb{R}$  gilt für die Multiplikation das **Kommutativgesetz**

$$ab = ba. \quad (\text{A.1.8})$$

Die Multiplikation mit dem Inversen bezeichnet man auch als Division und definiert für alle  $a \in \mathbb{R}$  und alle  $b \in \mathbb{R}^*$

$$a(b^{-1}) = \frac{a}{b} = a/b. \quad (\text{A.1.9})$$

Weiter gilt noch für die Verknüpfung der beiden Rechenoperationen das **Distributivgesetz**, d.h. für alle  $a, b, c \in \mathbb{R}$  gilt

$$a(b + c) = ab + ac. \quad (\text{A.1.10})$$

Man kann mit diesen Axiomen alle weiteren algebraischen Rechenregeln herleiten.

Die reellen Zahlen sind dadurch aber noch nicht vollständig charakterisiert. Als weiteres Element gibt es eine **Anordnungsrelation**, d.h. man kann von zwei reellen Zahlen sagen, ob die eine kleiner oder größer als die andere ist. Wir definieren für zwei Zahlen  $a, b \in \mathbb{R}$  also eine Relation  $a \leq b$  ( $a$  ist kleiner oder gleich  $b$ ), die folgende Regeln erfüllt

- Für alle  $a, b, c \in \mathbb{R}$  gilt

$$a \leq b \Rightarrow a + c \leq b + c. \quad (\text{A.1.11})$$

- Für alle  $a, b \in \mathbb{R}$  und alle  $c > 0$  gilt

$$a \leq b \Rightarrow ac \geq bc. \quad (\text{A.1.12})$$

- Seien  $a, b \in \mathbb{R}$  mit  $a > 0$  und  $b > 0$ . Dann existiert stets eine natürliche Zahl  $n \in \mathbb{N} \subset \mathbb{R}$ , so daß

$$b \leq na. \quad (\text{A.1.13})$$

Die bis jetzt gegebenen Gesetze definieren einen **archimedisch angeordneten Zahlkörper**. Allerdings erfüllt auch die Menge der rationalen Zahlen  $\mathbb{Q}$ , also die Teilmenge der reellen Zahlen, die sich als „Bruch“ zweier ganzer Zahlen schreiben lassen, diese Axiome. Um  $\mathbb{R}$  eindeutig zu charakterisieren, benötigen wir noch die **Vollständigkeit** bzgl. der Grenzwertbildung von Folgen. Darauf gehen wir im folgenden Abschnitt genauer ein.

Für das folgende benötigen wir noch den Begriff des **Betrags** einer reellen Zahl. Dieser ist definiert als

$$|a| = \begin{cases} a & \text{falls } a \geq 0, \\ (-a) & \text{falls } a < 0. \end{cases} \quad (\text{A.1.14})$$

Offensichtlich ist  $|a| \geq 0$  für alle  $a \in \mathbb{R}$ . Es gelten die Rechenregeln

$$|ab| = |a||b|, \quad |a + b| \leq |a| + |b|. \quad (\text{A.1.15})$$

Die letztgenannte Ungleichung heißt **Dreiecksungleichung**.

## A.2 Folgen und Grenzwerte

Eine Abbildung  $\mathbb{N} \rightarrow \mathbb{R}$ ,  $n \mapsto a_n$  bezeichnet man als reelle Zahlenfolge, d.h. jeder natürlichen Zahl  $n$  wird eindeutig eine reelle Zahl  $a_n$  zugeordnet. Wir bezeichnen eine solche Zahlenfolge abkürzend auch als  $(a_n)$  oder genauer  $(a_n)_{n \in \mathbb{N}}$ .

Eine Folge reeller Zahlen  $(a_n)$  konvergiert gegen eine reelle Zahl  $a$  genau dann, wenn es zu jedem reellen  $\epsilon > 0$  eine natürliche Zahl  $N$  gibt, so daß  $|a_n - a| < \epsilon$  für alle  $n > N$  gilt. Anschaulich bedeutet das, daß man die Abweichung der Folgenglieder mit hinreichend großem Index  $n$  von der Zahl  $a$  beliebig klein machen kann. Man schreibt dann auch

$$\lim_{n \rightarrow \infty} a_n = a. \quad (\text{A.2.1})$$

Eine Folge, für die eine solche Zahl  $a$  existiert, heißt **konvergent** und  $a$  der Grenzwert der Folge.

Wir können nun die konvergenten Folgen auch noch charakterisieren, wenn wir den Grenzwert nicht kennen. Zunächst nehmen wir an, die Folge  $(a_n)$  sei konvergent mit Grenzwert  $a$ . Sei weiter  $\epsilon > 0$  eine beliebige positive reelle Zahl. Dann können wir definitionsgemäß eine Zahl  $N$  finden, so daß  $|a_n - a| < \epsilon/2$  für alle  $n > N$  ist. Seien dann  $n_1, n_2 > N$  natürliche Zahlen. Dann folgt

$$|a_{n_1} - a_{n_2}| = |(a_{n_1} - a) + (a - a_{n_2})| \leq |a_{n_1} - a| + |a - a_{n_2}| < \epsilon/2 + \epsilon/2 = \epsilon. \quad (\text{A.2.2})$$

Ist also  $a_n$  konvergent, kann man zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  finden, so daß

$$|a_{n_1} - a_{n_2}| < \epsilon \quad \text{für alle } n_1, n_2 > N \quad (\text{A.2.3})$$

gilt. Folgen mit dieser Eigenschaft heißen **Cauchy-Folgen**. Anschaulich bedeutet das, daß die Folgenglieder  $(a_n)$  für hinreichend große  $N$  beliebig nahe beieinander liegen, wenn nur ihre Indizes  $> N$  sind. Anschaulich

liegt es daher nahe, zu denken, daß alle Cauchy-Folgen konvergent sind. Dies ist aber **genau für die reellen Zahlen**  $\mathbb{R}$  der Fall. Man sagt auch die reellen Zahlen sei **abgeschlossen bzgl. der Grenzwertbildung**. Man kann zeigen, daß  $\mathbb{R}$  eindeutig als **abgeschlossener Archimedisch angeordneter Zahlenkörper** charakterisiert ist.

Man beachte, daß die rationalen Zahlen  $\mathbb{Q} \subset \mathbb{R}$ , der alle algebraischen Eigenschaften wie die reellen Zahlen besitzt, *nicht abgeschlossen bzgl. der Grenzwertbildung* sind. Z.B. gibt es rationale Zahlenfolgen, die gegen  $\sqrt{2}$  konvergieren, und wir wissen, daß  $\sqrt{2}$  keine rationale Zahl ist.

Wir bemerken noch, daß die Grenzwertbildung mit der Addition und Multiplikation „verträglich“ sind, d.h. für konvergente Zahlenfolgen  $(a_n)$  und  $(b_n)$  mit den Grenzwerten  $a$  bzw.  $b$  gilt

$$\lim_{n \rightarrow \infty} a_n + b_n = a + b, \quad \lim_{n \rightarrow \infty} a_n b_n = ab. \quad (\text{A.2.4})$$

Ist auch noch  $b_n \neq 0$  für alle  $n \in \mathbb{N}$  und auch der Grenzwert  $b \neq 0$ , gilt auch

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{a}{b}. \quad (\text{A.2.5})$$

Wir wollen als Beispiel für die Anwendung dieser Begriffe einen wichtigen Grenzwertsatz beweisen. Dazu definieren wir zunächst, daß eine Folge **monoton wachsend** ist, wenn für alle  $n_1 < n_2$  stets  $a_{n_2} \geq a_{n_1}$  ist. Eine Folge heißt **nach oben beschränkt**, wenn es eine Zahl  $M \in \mathbb{R}$  gibt, so daß  $a_n \leq M$  für alle  $n \in \mathbb{N}$  gilt.

Es gilt der folgende Grenzwertsatz: Eine nach oben beschränkte monoton wachsende Folge ist konvergent. Zum *Beweis* bemerken wir, daß wegen der Monotonie der Folge alle Folgenglieder im Intervall  $I_0 = [a_1, K]$  liegen. Dabei bedeuten die eckigen Klammern, daß die Endpunkte in dem Intervall enthalten sein sollen (**abgeschlossenes Intervall**). Jetzt zerlegen wir das Intervall in zwei Hälften  $[a_1, (a_1 + K)/2]$  und  $[(a_1 + K)/2, K]$ . Falls in der rechten Hälfte noch Folgenglieder liegen, bezeichnen wir diese mit  $I_1$ . Andernfalls ist  $I_1$  die linke Hälfte. Da die Folge monoton wachsend ist und wenigstens ein Folgenglied  $a_{n_1}$  in  $I_1$  liegt, liegen auch alle Folgenglieder  $a_n$  mit  $n > n_1$  in diesem Intervall  $I_1$ . Nach Konstruktion gibt es aber keine Folgenglieder, die größer sind als die rechte Grenze dieses Intervalls. So können wir nun beliebig oft fortfahren und so immer kleinere Intervalle  $I_k$  bilden, so daß es stets  $n_k \in \mathbb{N}$  gibt, so daß  $a_n \in I_k$  für alle  $n > n_k$  gilt, aber keine Folgenglieder größer als die jeweils rechte Grenze der Intervalle  $I_k$  sind. Das Intervall  $I_k$  hat offenbar die Länge  $L_k = (K - a_1)/2^k$ . Da für alle  $n > n_k$  die Folgenglieder  $a_n \in I_k$  sind, gilt also für alle  $n, n' > n_k$ , daß  $|a_n - a_{n'}| < L_k$  ist. Sei nun  $\epsilon > 0$  beliebig vorgegeben. Dann können wir offenbar ein  $k \in \mathbb{N}$  finden, so daß  $L_k < \epsilon$  ist, denn aus

$$L_k = \frac{K - a_1}{2^k} < \epsilon \Leftrightarrow 2^k > \frac{K - a_1}{\epsilon}, \quad (\text{A.2.6})$$

und wir können zweifelsohne ein  $k$  finden, das diese Bedingung erfüllt. Folglich gibt es zu jedem  $\epsilon > 0$  ein  $k \in \mathbb{N}$ , so daß  $|a_n - a_{n'}| < \epsilon$  für alle  $n, n' > n_k$ . Damit ist aber die Folge nach dem Cauchyschen Konvergenzkriterium konvergent, und das war zu zeigen.

Wichtig ist noch der **Satz von Bolzano-Weierstraß**. Sei  $(a_n)$  eine beschränkte Folge, d.h. es gibt Zahlen  $m < M \in \mathbb{R}$ , so daß  $m \leq a_n \leq M$  für alle  $n \in \mathbb{N}$ . Sei weiter  $(n_k)$  eine Folge mit  $n_k \in \mathbb{N}$  und  $n_k \rightarrow \infty$  für  $k \rightarrow \infty$ . Dann heißt die Folge  $(a'_{n_k}) = a_{n_k}$  **Teilfolge** von  $(a_n)$ . Der Satz von Bolzano-Weierstraß besagt nun, daß jede beschränkte Folge stets wenigstens eine konvergente Teilfolge enthält. Zum *Beweis* betrachten wir das Intervall  $I_0 = [m, M]$ , indem voraussetzungsgemäß alle (unendlich vielen) Folgenglieder enthalten sind. Jetzt betrachten wir den Mittelpunkt  $x_1 = (m + M)/2$ . Liegen dann im linken Teilintervall  $[m, x_1]$  wieder unendlich viele Folgenglieder nennen wir dieses Teilintervall  $I_1$ . Andernfalls müssen im rechten Teilintervall unendlich viele Folgenglieder liegen, und wir nennen dieses Teilintervall  $I_1$ . So können wir beliebig oft verfahren. Wir haben dann eine Folge von Intervallen  $(I_k)$ , deren Länge  $l_k = (M - m)/2^k$  ist. In jedem Intervall  $I_k$  liegen unendlich viele Folgenglieder, und wir können daher ein  $n_k \in \mathbb{N}$  finden, so daß  $a_{n_k} \in I_k$  liegt und daß für  $k > k'$  stets  $n_k > n_{k'}$  ist. Dann strebt sicher  $n_k \rightarrow \infty$ , wenn  $k \rightarrow \infty$  ist. Es ist also  $(a'_{n_k})_k = (a_{n_k})_k$  eine

Teilfolge. Wir zeigen nun, daß diese konvergent ist, indem wir nachweisen, daß sie eine Cauchy-Folge ist. Sei dazu  $\epsilon > 0$  beliebig vorgegeben. Dann sei  $k_0$  so groß, daß  $l_k < \epsilon$  für alle  $k > k_0$  ist. Das ist sicher immer möglich, da  $l_k = (M - m)/2^k \rightarrow 0$  für  $k \rightarrow \infty$  ist. Nun ist für zwei beliebige natürliche Zahlen  $k_1, k_2 > k_0$  die Folgenglieder  $a'_{k_1}, a'_{k_2} \in I_{k_0}$ , denn für  $k > k_0$  ist aufgrund unserer Konstruktion der Intervalle  $I_k \subset I_{k_0}$ . Dann ist aber  $|a'_{k_1} - a'_{k_2}| < \epsilon$ . Für jedes  $\epsilon > 0$  existiert also ein  $k_0 \in \mathbb{N}$ , so daß  $|a'_{k_1} - a'_{k_2}| < \epsilon$  für alle  $k_1, k_2 > k_0$ , und folglich ist  $(a'_k)_k$  konvergent. Da dies eine Teilfolge von  $(a_n)_n$  ist, ist damit der Satz von Bolzano-Weierstraß bewiesen.

Eine weitere wichtige Art von Folgen sind die **Reihen**. Sei dazu  $(a_n)$  eine beliebige Folge reeller Zahlen. Dann bezeichnet man als **Teilsammenfolge** die durch

$$s_n = a_1 + a_2 + \dots + a_n = \sum_{j=1}^n a_j \quad (\text{A.2.7})$$

definierte Folge  $(s_n)$ . Falls diese Teilsammenfolge zu einem Grenzwert  $s$  konvergiert, spricht man von einer **konvergenten unendlichen Reihe** und schreibt

$$s = \lim_{n \rightarrow \infty} s_n = \sum_{j=1}^{\infty} a_j. \quad (\text{A.2.8})$$

Eine Reihe heißt **absolut konvergent**, wenn

$$\sum_{j=1}^{\infty} |a_n| \quad (\text{A.2.9})$$

existiert. Es ist klar, daß eine absolut konvergente unendliche Reihe immer konvergent ist. Um das einzusehen, wenden wir das Cauchysche Konvergenzkriterium auf die Teilsammenfolge an. Demnach ist die Reihe (A.2.8) konvergent, wenn es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so daß

$$|s_{n_1} - s_{n_2}| < \epsilon \quad \text{für alle } n_1, n_2 > N. \quad (\text{A.2.10})$$

Wir können nun ohne Beschränkung der Allgemeinheit annehmen, daß  $n_2 > n_1 > N$  ist. Dann bedeutet (A.2.10)

$$\left| \sum_{j=n_1+1}^{n_2} a_j \right| < \epsilon. \quad (\text{A.2.11})$$

Da nun voraussetzungsgemäß die Reihe absolut konvergent ist, können wir sogar ein  $N \in \mathbb{N}$  finden, so daß die die Teilsammenfolge von (A.2.9) das Cauchy-Kriterium erfüllt, d.h.

$$\sum_{j=n_1+1}^{n_2} |a_j| < \epsilon \quad \text{für alle } n_2 > n_1 > N. \quad (\text{A.2.12})$$

Wegen der Dreiecksungleichung ist dann aber

$$\left| \sum_{j=n_1+1}^{n_2} a_j \right| \leq \sum_{j=n_1+1}^{n_2} |a_j| < \epsilon \quad \text{für alle } n_2 > n_1 > N. \quad (\text{A.2.13})$$

Das bedeutet aber, daß die Reihe (A.2.8) konvergent ist, wenn sie absolut konvergent ist. *Die Umkehrung gilt allerdings nicht!*

Für absolut konvergente Reihen gilt der **Umordnungssatz**: Sei  $P : \mathbb{N} \rightarrow \mathbb{N}$  eine beliebige umkehrbar eindeutige Abbildung und sei die Reihe  $\sum a_n$  absolut konvergent mit dem Grenzwert  $\sum_{n=1}^{\infty} a_n = a$ . Dann gilt auch

$$\sum_{k=1}^{\infty} a_{P(k)} = a. \quad (\text{A.2.14})$$

**Beweis:** Da die Reihe voraussetzungsgemäß absolut konvergent ist, gibt es zu jedem  $\epsilon > 0$  ein  $n_0 \in \mathbb{N}$  mit

$$\sum_{k=n_0}^{\infty} |a_k| < \frac{\epsilon}{2}. \quad (\text{A.2.15})$$

Dann ist

$$\left| a - \sum_{k=1}^{n_0-1} a_k \right| \leq \sum_{k=n_0}^{\infty} |a_k| < \frac{\epsilon}{2}. \quad (\text{A.2.16})$$

Da  $P : \mathbb{N} \rightarrow \mathbb{N}$  umkehrbar eindeutig ist, gibt es ein  $N \in \mathbb{N}$ , so daß  $\{P(1), P(2), \dots, P(N)\} \supset \{1, 2, \dots, n_0 - 1\}$ . Dann gilt für alle  $m \geq N$

$$\left| \sum_{k=1}^m a_{P(k)} - a \right| \leq \left| \sum_{k=1}^m a_{P(k)} - \sum_{k=1}^{n_0-1} a_k \right| + \left| \sum_{k=1}^{n_0-1} a_k - a \right| \leq \sum_{k=n_0}^{\infty} |a_k| + \frac{\epsilon}{2} < \epsilon, \quad (\text{A.2.17})$$

wobei wir (A.2.15) und (A.2.16) verwendet haben. Damit ist aber klar, daß tatsächlich (A.2.14) gilt, und das war zu zeigen.

Im folgenden beweisen wir einige **Konvergenzkriterien** für Reihen, also notwendige Bedingungen für die Konvergenz von Reihen.

**Leibnizsches Konvergenzkriterium für alternierende Reihen:** Es sei  $(a_k)$  eine monoton fallende Nullfolge mit  $a_k \geq 0$ . Dann ist die „alternierende Reihe“  $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$  konvergent.

**Beweis:** Wir betrachten die Partialsummen

$$S_k = \sum_{n=1}^k (-1)^{n-1} a_n. \quad (\text{A.2.18})$$

Dann gilt für die Partialsummenfolge mit geraden Indizes

$$S_{2k+2} - S_{2k} = (-1)^{2k} a_{2k+1} + (-1)^{2k+1} a_{2k+2} = a_{2k+1} - a_{2k+2} \geq 0, \quad (\text{A.2.19})$$

da die Folge  $(a_j)_{j \in \mathbb{N}}$  voraussetzungsgemäß monoton fallend ist. Folglich ist die Folge  $(S_{2k})_{k \in \mathbb{N}}$  monoton wachsend, d.h. für alle  $k$  ist

$$S_2 \leq S_4 \leq \dots \leq S_{2k}. \quad (\text{A.2.20})$$

Für die ungeraden Partialsummenfolglglieder ist hingegen

$$S_{2k+3} - S_{2k+1} = (-1)^{2k+1} a_{2k+2} + (-1)^{2k+2} a_{2k+3} = -a_{2k+2} + a_{2k+3} \leq 0 \quad (\text{A.2.21})$$

und damit diese Teilsummenfolge monoton fallend, d.h. für alle  $k \in \mathbb{N}$ .

$$S_1 \geq S_3 \geq S_5 \geq \dots \geq S_{2k+1}. \quad (\text{A.2.22})$$

Andererseits ist

$$S_{2k+1} - S_{2k} = a_{2k+1} \geq 0 \quad (\text{A.2.23})$$

und folglich für alle  $k$

$$S_{2k} \leq S_{2k+1} \leq S_1. \quad (\text{A.2.24})$$

Da  $(S_{2k})$  somit gemäß (A.2.19) monoton wachsend und wegen (A.2.24) nach oben beschränkt ist, ist diese Folge konvergent. Es sei

$$\lim_{k \rightarrow \infty} S_{2k} = S_1. \quad (\text{A.2.25})$$

Aus (A.2.23) folgt andererseits für alle  $k \in \mathbb{N}$

$$S_{2k+1} \geq S_{2k} \geq S_2 \quad (\text{A.2.26})$$

Damit ist also  $(S_{2k+1})$  eine gemäß (A.2.22) monoton fallende und wegen (A.2.26) nach unten beschränkte Folge damit ebenfalls konvergent. Es sei

$$\lim_{k \rightarrow \infty} S_{2k+1} = S_2. \quad (\text{A.2.27})$$

Offenbar gilt  $S_1 = S_2$ , denn es ist

$$S_{2k+1} - S_{2k} = a_{2k+1} \rightarrow 0 \quad \text{für } k \rightarrow \infty, \quad (\text{A.2.28})$$

da die Folge  $(a_k)$  voraussetzungsgemäß eine Nullfolge ist. Demnach konvergieren  $(S_{2k})$  und  $S_{2k+1}$  gegen denselben Grenzwert  $S = S_1 = S_2$ . Das bedeutet, daß es zu  $\epsilon > 0$  Zahlen  $N_1 \in \mathbb{N}$  und  $N_2 \in \mathbb{N}$  gibt, so daß  $|S_{2k} - S| < \epsilon$  für alle  $2k \geq N_1$  und  $|S_{2k+1} - S| < \epsilon$  für alle  $2k + 1 \geq N_2$ . Setzen wir dann  $N = \max(N_1, N_2)$ , so ist  $|S_k - S| < \epsilon$  für alle  $k > N$ , und folglich ist  $(S_k)$  konvergent, und das war zu zeigen.

Damit können wir auch ein *Beispiel* für eine konvergente aber nicht absolut konvergente Reihe angeben. Betrachten wir die Reihe  $\sum_{k=1}^{\infty} (-1)^{k+1}/k = 1 - 1/2 + 1/3 + \dots$ , so ist diese nach dem Leibnizschen Konvergenzkriterium konvergent. Dies ist jedoch nicht der Fall für die Reihe  $\sum_{k=1}^{\infty} 1/k = 1 + 1/2 + 1/3 + \dots$ . Betrachten wir nämlich die Partialsummen dieser letztgenannten Reihe, folgt

$$S_{2k} = 1 + 1/2 + (1/3 + 1/4) + (1/5 + 1/6 + 1/7 + 1/8) + \dots + [1/(2^{k-1} + 1) + \dots + 1/2^k]. \quad (\text{A.2.29})$$

Jede Klammer ist nun aber  $\geq K/2^j$  mit  $K = 2^j - 2^{j-1} = 2^{j-1}(2 - 1) = 2^{j-1}$ , also  $\geq 1/2$  und damit

$$S_{2k} \geq 1 + \frac{k}{2} \rightarrow \infty \quad \text{für } k \rightarrow \infty, \quad (\text{A.2.30})$$

und folglich divergiert die Reihe  $\sum_{k=1}^{\infty} 1/k \rightarrow \infty$ .

Wir zeigen nun auch, daß wir die entsprechende alternierende Reihe so umordnen können, daß sie divergiert. Dazu schreiben wir ihre Glieder in der folgenden Reihenfolge

$$\begin{aligned} & 1 - 1/2 + 1/3 - 1/4 \\ & \quad + (1/5 + 1/7) - 1/6 \\ & \quad + (1/9 + 1/11 + 1/13 + 1/15) - 1/8 \\ & \quad + \dots \\ & \quad + [1/(2^j + 1) + 1/(2^j + 3) + \dots + 1/(2^{j+1} - 1)] - 1/(2n + 2) \\ & \quad + \dots \end{aligned} \quad (\text{A.2.31})$$

Für die Klammern gilt

$$1/(2^j + 1) + 1/(2^j + 3) + \dots + 1/(2^{j+1} - 1) \geq (2^{n+1} - 2^n - 1)/2^{n+1} = (2^n - 1)/2^{n+1} \geq 2^{n-1}/2^{n+1} = 1/4. \quad (\text{A.2.32})$$

Damit ist aber jede der Zeilen in (A.2.31) größer als 0, und folglich die Umordnung gegen  $\infty$  divergent.

Ein wichtiges **Kriterium für die absolute Konvergenz** einer Reihe ist das **Vergleichskriterium**. Dazu sei  $a_j$  eine beliebige Folge und  $b_j$  eine Folge mit  $b_j \geq 0$ , so daß  $|a_j| < b_j$  für alle  $j \in \mathbb{N}$  ist. Ist dann die aus den  $b_j$  gebildete Reihe konvergent, so ist die aus den  $a_j$  gebildete Reihe absolut konvergent.

**Beweis:** Die Teilsummenfolge der aus den  $b_j$  gebildeten Reihe

$$s_n^{(b)} = \sum_{j=1}^n b_j \quad (\text{A.2.33})$$

ist monoton wachsend und voraussetzungsgemäß konvergent. Sei  $s^{(b)}$  der Grenzwert. Offenbar gilt  $s_n^{(b)} \leq s^{(b)}$  für alle  $n \in \mathbb{N}$ . Da aber  $|a_j| < b_j$  für alle  $j \in \mathbb{N}$  ist, gilt für die Teilsummenfolge

$$s_n^{(|a|)} = \sum_{j=1}^n |a_j| \leq \sum_{j=1}^n b_j \leq s^{(b)}. \quad (\text{A.2.34})$$

Folglich ist die monoton wachsende Folge  $s_n^{(|a|)}$  nach oben beschränkt und damit nach dem oben bewiesenen Konvergenzkriterium für monoton wachsende Folgen ebenfalls konvergent. Damit ist die aus  $a_j$  gebildete Reihe absolut konvergent, was zu beweisen war.

Ein wichtiges Beispiel ist die **geometrische Reihe**. Sei dazu  $q \neq 0$ . Die **geometrische Folge** ist dann durch  $a_j = q^j$  für  $j \in \mathbb{N}$  definiert, und die geometrische Reihe wird aus dieser Folge gebildet. Hier liegt der seltene Fall vor, daß wir die Teilsummenfolge explizit ausrechnen können, denn es gilt

$$s_n = \sum_{j=1}^n q^j = q + q^2 + \dots + q^n. \quad (\text{A.2.35})$$

Multiplizieren wir diese Gleichung mit  $q$ , folgt

$$qs_n = \sum_{j=1}^n q^{j+1} = q^2 + q^3 + \dots + q^{n+1}. \quad (\text{A.2.36})$$

Subtrahieren wir beide Gleichungen, ergibt sich

$$s_n - qs_n = (1 - q)s_n = q - q^{n+1}. \quad (\text{A.2.37})$$

Falls nun  $q \neq 1$  ist, folgt daraus

$$s_n = \frac{q - q^{n+1}}{1 - q}. \quad (\text{A.2.38})$$

Für  $q = 1$  ist offensichtlich  $s_n = n$ . Offensichtlich ist diese Teilsummenfolge genau dann konvergent, wenn  $|q| < 1$  ist, denn andernfalls wächst  $q^{n+1}$  für  $n \rightarrow \infty$  über alle Grenzen. Für  $|q| < 1$  ist  $\lim_{n \rightarrow \infty} q^n = 0$  (*Beweis als Übungsaufgabe*) und damit

$$\sum_{j=1}^{\infty} q^j = \frac{q}{1 - q} \quad \text{falls } |q| < 1. \quad (\text{A.2.39})$$

Die geometrische Reihe ist nützlich, um zusammen mit dem oben hergeleiteten Vergleichskriterium für absolute Konvergenz von Reihen einfache Kriterien für die absolute Konvergenz von Reihen zu liefern, das **Quotienten-** und das **Wurzelkriterium**. Gemäß dem Vergleichskriterium ist nämlich die aus der Folge  $(a_j)$  gebildete Reihe absolut konvergent, wenn es reelle Zahlen  $A > 0$  und  $0 < q < 1$  und eine Zahl  $N \in \mathbb{N}$  gibt, so daß

$$|a_j| < Aq^j \quad \text{für alle } j > N \quad (\text{A.2.40})$$

gilt.

Angenommen, für die  $a_j$  gilt

$$\left| \frac{a_{j+1}}{a_j} \right| < q < 1 \quad \text{für alle } j > N. \quad (\text{A.2.41})$$

Dann folgt

$$|a_{N+2}| < |a_{N+1}|q, \quad |a_{N+3}| < |a_{N+2}|q < |a_{N+1}|q^2, \quad \dots, \quad |a_{N+j}| < |a_{N+1}|q^{j-1} \quad \text{für alle } j \in \mathbb{N}. \quad (\text{A.2.42})$$

Setzt man also  $A = |a_{N+1}|q^{-N-1}$ , so folgt daraus, daß

$$|a_k| < |a_{N+1}|q^{k-N-1} = Aq^k \quad \text{für alle } k > N. \quad (\text{A.2.43})$$

Demnach ist gemäß dem Vergleichskriterium (A.2.40) die aus der Folge  $(a_j)$  gebildete Reihe absolut konvergent.

Offensichtlich ist (A.2.41) insbesondere dann erfüllt, wenn der Grenzwert

$$0 \leq \lim_{j \rightarrow \infty} \left| \frac{a_{j+1}}{a_j} \right| = q' < 1 \quad (\text{A.2.44})$$

ist. Denn dann gibt es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so daß

$$\left| \left| \frac{a_{j+1}}{a_j} \right| - q' \right| < \epsilon \quad \text{für alle } j > N. \quad (\text{A.2.45})$$

Das bedeutet aber, daß für alle  $j > N$

$$\left| \frac{a_{j+1}}{a_j} \right| < q' + \epsilon \quad (\text{A.2.46})$$

gilt. Wählen wir nun  $\epsilon$  so klein, daß  $q = q' + \epsilon < 1$ , also  $0 < \epsilon < 1 - q'$  ist, ist folglich (A.2.41) mit diesem  $0 < q < 1$  erfüllt, und die Reihe ist konvergent.

Dies ist das **Quotientenkriterium**: Falls (A.2.44) erfüllt ist, ist die aus  $(a_j)$  gebildete Reihe absolut konvergent.

Ebenso ergibt sich das **Wurzelkriterium**. Offensichtlich ist nämlich (A.2.40) genau dann erfüllt, wenn für alle  $j > N$

$$|a_j|^{1/j} < A^{1/j}q \quad (\text{A.2.47})$$

gilt. Ist dabei  $q < 1$ , ist die aus der Folge  $(a_j)$  gebildete Reihe absolut konvergent. Dies ist insbesondere der Fall, wenn

$$\lim_{j \rightarrow \infty} |a_j|^{1/j} = q' < 1 \quad (\text{A.2.48})$$

gilt.

Besonders wichtig sind nun die **Potenzreihen**, die es gestatten außer den rein algebraisch definierbaren **Polynomen** auch allgemeinere Funktionen zu definieren. Dazu betrachtet man Reihen der Form

$$f(x) = \sum_{j=0}^{\infty} a_j x^j. \quad (\text{A.2.49})$$

Die Teilsummen sind Polynome in  $x$ . Es stellt sich freilich sofort die Frage, ob die Reihe für irgendein  $x \neq 0$  konvergiert und ggf. für welche  $x \in \mathbb{R}$  dies der Fall ist.

Das Quotientenkriterium liefert hier eine Antwort: Die Potenzreihe ist sicher für diejenigen  $x$  absolut konvergent, für die

$$\lim_{j \rightarrow \infty} \left| \frac{a_{j+1} x^{j+1}}{a_j x^j} \right| < 1 \quad (\text{A.2.50})$$

ist (vorausgesetzt der Grenzwert existiert). Das kann man aber sofort auch umschreiben:

$$|x| \lim_{j \rightarrow \infty} \left| \frac{a_{j+1}}{a_j} \right| < 1 \Rightarrow |x| < \lim_{j \rightarrow \infty} \left| \frac{a_j}{a_{j+1}} \right| = R. \quad (\text{A.2.51})$$

Falls der rechts stehende Grenzwert  $R$  existiert, ist demnach die Potenzreihe (A.2.49) für alle  $|x| < R$  absolut konvergent. Falls  $R \rightarrow \infty$ , ist die Potenzreihe sogar für alle  $x \in \mathbb{R}$  absolut konvergent.

### A.3 Satz vom Supremum und Infimum

In diesem Abschnitt beleuchten wir nochmals die Vollständigkeitseigenschaften der reellen Zahlen von einem etwas anderen Blickwinkel. Wir betrachten dazu eine beliebige Teilmenge der reellen Zahlen:  $A \subseteq \mathbb{R}$ . Diese Menge heißt von oben (unten) beschränkt, wenn es eine Zahl  $M$  ( $m$ ) gibt, so daß für alle  $x \in A$  stets  $x < M$  ( $x > m$ ) gilt. Dabei müssen die **obere Schranke**  $M$  (**untere Schranke**  $m$ ) selbst nicht notwendig zu  $A$  gehören. Man nennt nun die kleinste obere (größte untere) Schranke von  $A$  das **Supremum** (**Infimum**) der Menge  $A$  (in Formeln  $\sup A$  bzw.  $\inf A$ ).

Um die Darstellung klarer zu machen, betrachten wir im folgenden nach eine nach oben beschränkte Menge  $A$  und zeigen, daß diese stets ein Supremum in  $\mathbb{R}$  besitzt. Analog besitzt eine nach unten beschränkte Menge  $A$  ein Infimum. Der Beweis der letzteren Behauptung verläuft ebenfalls vollständig analog zum Satz vom Supremum.

Sei also  $A \subseteq \mathbb{R}$  nicht leer und besitze eine obere Schranke  $M$ . Sei nun  $x_1 \in A$  beliebig gewählt. Ist dann  $x_1$  eine obere Schranke von  $A$ , ist diese Zahl offenbar bereits das Supremum dieser Menge, denn es kann dann keine kleinere obere Schranke als  $x_1$  geben. Ist  $x_1$  keine obere Schranke von  $A$ , betrachten wir das Intervall  $I_1 = [x_1, M]$ . Da voraussetzungsgemäß für alle  $x \in A$  immer  $x \leq M$  gilt und  $x_1$  keine obere Schranke ist, existiert in  $I_1$  noch wenigstens eine Zahl  $x \in M$  mit  $x_1 < x \leq M$ . Betrachten wir nun den Mittelpunkt des Intervalls  $I_1$ , also  $x_2 = (x_1 + M)/2$ . Falls  $x_2$  obere Schranke von  $A$  ist, setzen wir  $I_2 = [x_1, x_2]$ , andernfalls  $I_2 = [x_2, M]$ . Dann ist in jedem Fall  $A \cap I_2 \neq \emptyset$  und die obere Grenze des Intervalls  $I_2$  ist eine obere Schranke von  $A$ . Dieses Verfahren können wir nun beliebig oft wiederholen, und es entsteht eine Folge von Intervallen  $I_n = [a_n, b_n]$  mit  $0 \leq b_n - a_n = (M - x_1)/2^{n-1}$ . Für  $n \rightarrow \infty$  wird das Intervall beliebig klein, und  $a_n$  und  $b_n$  konvergieren demnach zu einem Wert  $M'$ . Offenbar ist  $M'$  eine obere Schranke von  $A$ , denn es gilt für alle  $x \in A$  stets  $x \leq b_n$  für alle  $n \in \mathbb{N}$ . Demnach ist aber auch  $x \leq \lim_{n \rightarrow \infty} b_n = M'$  und also  $M'$  eine obere Schranke. Wir müssen nun noch zeigen, daß  $M'$  die kleinste obere Schranke, also  $\sup A$  ist. Betrachten wir dazu eine beliebige Zahl  $M'' < M$  und nehmen an, sie sei obere Schranke von  $A$ . Es ist klar, daß dann für alle  $n \in \mathbb{N}$  stets  $M'' \in I_n$  gelten muß. Da nun aber die linken Grenzen  $a_n$  der Intervalle  $I_n$  monoton wachsen und gegen  $M$  streben, muß es ein  $n_0 \in \mathbb{N}$  geben, so daß  $a_n > M''$  für alle  $n \geq n_0$  gilt, und damit kann  $M''$  nicht in  $I_n$  liegen. Dann ist aber für diese  $n$  immer  $M'' \notin I_n$  im Widerspruch zur Annahme, daß  $M''$  obere Schranke von  $A$  ist. Folglich muß  $M' = \sup A$ , also kleinste obere Schranke von  $A$  sein.

Wir betonen nochmals, daß wir hier wesentlich die **Vollständigkeit** der reellen Zahlen unter Grenzwertbildungen verwendet haben, denn sie garantiert, daß die eben konstruierte Zahl  $M' \in \mathbb{R}$  gilt.

### A.4 Lineare und quadratische Gleichungen

Zu den klassischen Aufgaben der Algebra gehört die Auflösung von Gleichungen. Die einfachste Art von Gleichungen sind die **linearen Gleichungen**. Seien dazu  $a, b \in \mathbb{R}$  irgendwelche Zahlen. Dann fragen wir, ob es eine Zahl  $x \in \mathbb{R}$  gibt, die die Gleichung

$$ax + b = 0 \tag{A.4.1}$$

erfüllt. Zunächst können wir die Gleichung vereinfachen, indem wir auf beiden Seiten  $b$  subtrahieren:

$$(ax + b) - b = -b \Rightarrow ax + (b - b) = ax = -b. \tag{A.4.2}$$

Diese Gleichung läßt sich nun offenbar genau dann eindeutig nach  $x$  auflösen, wenn es ein Inverses von  $a$  bzgl. der Multiplikation gibt, d.h. falls  $a \neq 0$  ist. Dann folgt

$$x = -\frac{b}{a}. \tag{A.4.3}$$

Falls nun  $a = 0$  ist, so ist  $ax = 0x = 0$  für alle  $x \in \mathbb{R}$ . In diesem Falle wird die Gleichung durch alle Zahlen  $x \in \mathbb{R}$  gelöst, falls  $b = 0$  ist. Falls  $b \neq 0$  ist, besitzt die Gleichung keine Lösung.

Kommen wir nun auf die quadratischen Gleichungen zu sprechen. Dies sind Gleichungen der Form

$$x^2 + qx + p = 0. \quad (\text{A.4.4})$$

Um sie zu lösen, erinnern wir an die **binomische Formel**

$$(a + b)^2 = a^2 + 2ab + b^2, \quad (\text{A.4.5})$$

die man sehr leicht durch Anwendung des Distributivgesetzes nachrechnet. Setzen wir hierin  $a = x$  und  $b = q/2$ , finden wir

$$\left(x + \frac{q}{2}\right)^2 = x^2 + qx + \frac{q^2}{4}. \quad (\text{A.4.6})$$

Verwenden wir dies in (A.4.4), ergibt sich die Gleichung

$$\left(x + \frac{q}{2}\right)^2 + p - \frac{q^2}{4} = 0. \quad (\text{A.4.7})$$

Dies ergibt

$$\left(x + \frac{q}{2}\right)^2 = \frac{q^2}{4} - p. \quad (\text{A.4.8})$$

Wir benötigen nun lediglich eine **Umkehrfunktion** des Quadrierens, die **Quadratwurzel**. Ohne Beweis bemerken wir, daß für  $a \geq 0$  die **Quadratwurzelfunktion** wohldefiniert ist, d.h. zu jedem reellen  $a \geq 0$  existiert genau eine reelle Zahl  $b \geq 0$  mit  $b = \sqrt{a}$ , die dadurch definiert ist, daß  $b^2 = a$  ist. Freilich gilt dann auch  $(-b)^2 = b^2 = a$ . Für  $a \neq 0$  gibt es also zwei Lösungen der Gleichung  $b^2 = a$ , nämlich  $b = +\sqrt{a} \geq 0$  und  $b = -\sqrt{a} \leq 0$ .

Dies wenden wir nun auf (A.4.8) an. Falls also  $q^2/4 - p > 0$  ist, existieren stets zwei Lösungen

$$x_1 + \frac{q}{2} = \sqrt{\frac{q^2}{4} - p} \Rightarrow x_1 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} - p} \quad (\text{A.4.9})$$

und

$$x_2 + \frac{q}{2} = -\sqrt{\frac{q^2}{4} - p} \Rightarrow x_2 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} - p}. \quad (\text{A.4.10})$$

Für  $q^2/4 = p$  gibt es nur eine Lösung  $x_1 = x_2 = -q/2$ . Wir bemerken weiter, daß stets

$$x^2 + qx + p = (x - x_1)(x - x_2) \quad (\text{A.4.11})$$

gilt, falls die Gleichung (A.4.4) lösbar ist. Dies zeigen wir durch Ausmultiplizieren:

$$(x - x_1)(x - x_2) = x^2 - (x_1 + x_2)x + x_1x_2. \quad (\text{A.4.12})$$

Setzt man die Lösungen (A.4.9) und (A.4.10) ein, finden wir durch einfaches Ausrechnen in der Tat

$$-(x_1 + x_2) = q, \quad x_1x_2 = p. \quad (\text{A.4.13})$$

Wie wir gesehen haben, gibt es im Reellen keine Lösung der quadratischen Gleichung (A.4.4), falls  $q^2/4 - p < 0$  ist, denn für alle  $a \in \mathbb{R}$  ist  $a^2 \geq 0$  und  $a^2 = 0 \Leftrightarrow a = 0$ . Wir werden in einem späteren Kapitel dieser Vorlesung sehen, wie man die reellen Zahlen erweitern kann, so daß auch in diesem Fall die quadratische Gleichung lösbar ist. Dies führt dann auf den **Körper der komplexen Zahlen**.

## A.5 Funktionen und Stetigkeit

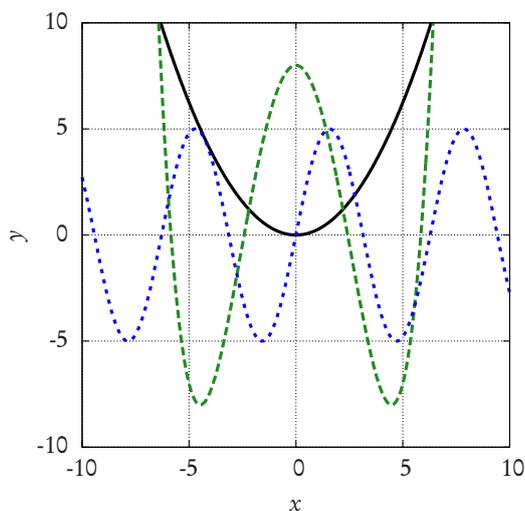


Abbildung A.1: Beispiele für Funktionen:  $y = x^2/4$  (schwarz durchgezogen),  $y = (0,2x^2 - 1)^2 - 8$  (grün gestrichelt) und  $y = 5 \sin x$  (blau gepunktet).

Eine **Funktion einer reellen Veränderlichen** ist eine eindeutige Abbildung von einer Teilmenge der reellen Zahlen ( $D \subseteq \mathbb{R}$ ) in die reellen Zahlen. Dies schreiben wir in der Form  $f : D \rightarrow \mathbb{R}$ . Jeder Zahl  $x \in D$  wird eine Zahl  $y \in \mathbb{R}$  zugeordnet, nämlich  $y = f(x)$ . Veranschaulichen können wir uns Funktionen, indem wir  $(x, y)$  als Koordinaten eines kartesischen Koordinatensystems auffassen. Die Kurve, die aus den entsprechenden Punkten  $(x, f(x))$  besteht, heißt **Graph der Funktion  $f$** .

In den Beispielen im obigen Bild sind alle Funktionen Funktionen auf ganz  $\mathbb{R}$  definiert. Außerdem weist ihr Graph keinerlei Unterbrechungen oder ähnliche Anomalien auf. Mathematisch ist eine Voraussetzung dafür, daß eine Funktion stetig ist. Anschaulich bedeutet die **Stetigkeit einer Funktion** an der Stelle  $x_0$ , daß sich der Funktionswert um  $x_0$  nur wenig ändert, wenn man sie in einer kleinen Umgebung um  $x_0$  betrachtet. Formal läßt sich dies so definieren:

Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt stetig an der Stelle  $x_0 \in D$ , wenn es zu jedem  $\epsilon > 0$  ein  $\delta > 0$  gibt, so daß für alle  $x \in D$  mit  $|x - x_0| < \delta$  stets  $|f(x) - f(x_0)| < \epsilon$  ist.

Anders formuliert heißt das, daß die Funktionswerte für hinreichend kleine Umgebungen um  $x_0$  beliebig wenig von dem

Funktionswert  $f(x_0)$  bei  $x_0$  abweichen.

Es ist nun leicht zu zeigen, daß für zwei in  $x_0$  stetige Funktionen  $f$  und  $g$  auch die Funktion  $f + g$  stetig ist. Das sieht man sofort wie folgt ein: Voraussetzungsgemäß gibt es zu jedem  $\epsilon > 0$  ein  $\delta_1 > 0$ , so daß  $|f(x) - f(x_0)| < \epsilon/2$  für alle  $x \in D$  mit  $|x - x_0| < \delta_1$ . Ebenso gibt es auch ein  $\delta_2 > 0$ , so daß  $|g(x) - g(x_0)| < \epsilon/2$ . Setzen wir nun  $\delta = \min(\delta_1, \delta_2)$ , dann gilt wegen der Dreiecksungleichung für alle  $x < \delta$

$$\begin{aligned} |[f(x) + g(x)] - [f(x_0) + g(x_0)]| &= |[f(x) - f(x_0)] + [g(x) - g(x_0)]| \\ &\leq |f(x) - f(x_0)| + |g(x) - g(x_0)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned} \tag{A.5.1}$$

d.h. zu jedem  $\epsilon > 0$  können wir ein  $\delta > 0$  finden, so daß für alle  $x \in D$  mit  $|x - x_0| < \delta$  stets

$$|[f(x) + g(x)] - [f(x_0) + g(x_0)]| < \epsilon \tag{A.5.2}$$

gilt. Das bedeutet aber gemäß der Definition, daß die Funktion  $f + g$  stetig ist.

Ebenso beweist man, daß mit zwei Funktionen  $f$  und  $g$ , die stetig in  $x_0$  sind, auch das Produkt  $f g$  und (vorausgesetzt  $g(x_0) \neq 0$ ) der Quotient  $f/g$  stetig in  $x_0$  sind (*Beweis als Übung!*).

Offensichtlich sind eine konstante Funktion  $f(x) = A = \text{const}$  sowie die Funktion  $f(x) = x$  in allen  $x \in \mathbb{R}$  definiert und stetig. Aus den oben angegebenen Sätzen folgt, dann, daß auch jedes Polynom

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n = \sum_{k=0}^n a_k x^k \tag{A.5.3}$$

stetig ist.

Wichtig ist der **Zwischenwertsatz**. Sei dazu  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Weiter sei  $f(a) < f(b)$ . Dann gibt es zu jedem  $u \in [f(a), f(b)]$  ein  $\xi \in [a, b]$ , so daß  $f(\xi) = u$  ist (entsprechendes gilt auch falls  $f(a) > f(b)$ )

ist). Zum Beweis konstruieren wir  $\xi$  mittels Intervallschachtelung. Im ersten Schritt betrachten wir also den Mittelpunkt des Ausgangsintervalls  $\xi_1 = (a + b)/2$ . Falls dann  $f(\xi_1) = u$ , haben wir  $\xi$  schon gefunden. Andernfalls, setzen wir  $a_1 = \xi_1$ ,  $b_1 = b$  falls  $f(\xi_1) < u$  und  $a_1 = a$  und  $b_1 = \xi_1$  falls  $f(\xi_1) > u$ . Wir haben dann wieder ein Intervall  $[a_1, b_1]$ , für das  $f(a_1) \leq u \leq f(b_1)$  ist. Mit diesem Intervall verfahren wir nun genauso. Dadurch entsteht eine Folge von Intervallen  $[a_n, b_n]$ , für die stets  $f(a_n) \leq u \leq f(b_n)$  ist. Die Folge  $a_n$  ist monoton wachsend und die Folge  $b_n$  monoton fallend. Die Länge des Intervalls ist  $|b_n - a_n| = (b - a)/2^n$ , und folglich konvergieren beide Folgen gegen denselben Grenzwert  $\xi$ . Da weiter  $f$  voraussetzungsgemäß im ganzen Intervall  $[a, b]$  stetig ist, muß folglich  $f(\xi) = u$  sein, und das war zu zeigen.

Wichtig ist noch der **Satz vom Maximum und Minimum**: Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Dann nimmt  $f$  an wenigstens einer Stelle  $\xi \in [a, b]$  ein Maximum (Minimum) an. Dabei ist das Maximum durch  $\sup_{x \in [a, b]} f(x)$  (Minimum durch  $\inf_{x \in [a, b]} f(x)$ ) definiert.

Wir beweisen den Satz für das Maximum. Der Beweis für das Minimum folgt analog. Da die Funktion im gesamten (abgeschlossenen!) Intervall definiert ist, ist ihr Bildbereich  $f([a, b])$  eine beschränkte Menge und besitzt demnach gemäß Abschnitt A.3 ein Supremum, d.h. es gilt  $f(x) \leq M$  für alle  $x \in [a, b]$  mit  $M = \sup_{x \in [a, b]} f(x)$ . Wie wir beim Beweis des Satzes vom Infimum und Supremum in Abschnitt A.3 gesehen haben, existiert dann eine Zahlenfolge  $x_n \in [a, b]$ , so daß  $\lim_{n \rightarrow \infty} f(x_n) = M$ .

## A.6 Differentialrechnung für Funktionen einer reellen Veränderlichen

Die Differentialrechnung beschäftigt sich mit **lokalen Eigenschaften** von Funktionen, widmet sich also der Untersuchung des Änderungsverhaltens von Funktionen in kleinen Umgebungen eines beliebigen Punktes im Definitionsbereich der Funktion.

### A.6.1 Definition der Ableitung einer Funktion

Es sei  $f : D \rightarrow \mathbb{R}$  mit einem offenen Definitionsbereich  $D \subseteq \mathbb{R}$ , d.h. wir nehmen an, daß mit  $x_0 \in D$  auch eine bestimmte Umgebung um diesen Punkt in  $D$  enthalten sei, d.h. daß es ein  $\delta > 0$  gibt, so daß für alle  $x$  mit  $|x - x_0| < \delta$  auch  $x \in D$  gilt.

In diesem Fall besitzt die Funktion  $f$  den **Grenzwert**  $A \in \mathbb{R}$  für  $x \rightarrow x_0$ , wenn es zu jedem  $\epsilon > 0$  ein  $\delta > 0$  gibt, so daß

$$|f(x) - A| < \epsilon \quad \text{für alle } x \text{ mit } |x - x_0| < \delta. \quad (\text{A.6.1})$$

In dem Fall schreiben wir

$$\lim_{x \rightarrow x_0} f(x) = A. \quad (\text{A.6.2})$$

Wir bemerken, daß unter den obigen Voraussetzungen eine Funktion genau dann stetig in  $x_0$  ist, wenn

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) \quad (\text{A.6.3})$$

gilt.

Die **Ableitung** der Funktion  $f$  an der Stelle  $x_0$  ist dann definiert als

$$f'(x_0) = \frac{d}{dx} f(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}, \quad (\text{A.6.4})$$

vorausgesetzt, dieser Grenzwert existiert. In dem Fall heißt die Funktion  $f$  **differenzierbar** an der Stelle  $x_0$ . Sie heißt differenzierbar in einem Bereich  $D' \subseteq D$  falls  $f'(x)$  in jedem Punkt  $x \in D'$  differenzierbar ist.

Wichtige Beispiele sind

- $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = c = \text{const.}$  Dann gilt für jedes  $x_0 \in \mathbb{R}$

$$\frac{f(x) - f(x_0)}{x - x_0} = 0 \Rightarrow f'(x_0) = 0. \quad (\text{A.6.5})$$

Die konstante Funktion ist also in ganz  $\mathbb{R}$  differenzierbar und ihre Ableitung verschwindet.

- $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x$ . Dann folgt

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = \lim_{x \rightarrow x_0} 1 = 1. \quad (\text{A.6.6})$$

- $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x^2$

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \rightarrow x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} = \lim_{x \rightarrow x_0} (x + x_0) = 2x_0. \quad (\text{A.6.7})$$

Die Ableitung besitzt auch eine **geometrische Bedeutung**. Betrachten wir dazu den Graphen einer Funktion  $y = f(x)$ . Dann bedeutet  $[f(x) - f(x_0)] / (x - x_0)$  die Steigung der Sekante durch die Punkte  $[x, f(x)]$  und  $[x_0, f(x_0)]$ . Die Ableitung ergibt, falls sie existiert, entsprechend die Steigung der **Tangente** an die durch den Funktionsgraphen gegebenen Kurve.

## A.6.2 Formeln zur Ableitung

Im folgenden leiten wir einige allgemeine Formeln zur Ableitung von Funktionen her. Im folgenden nehmen wir an, daß alle beteiligten Funktionen im Punkt  $x_0$  ihres Definitionsbereiches differenzierbar sind. Dann gilt

$$\frac{d}{dx}(f + g)(x_0) = f'(x_0) + g'(x_0). \quad (\text{A.6.8})$$

Zum Beweis müssen wir nur den entsprechenden Limes betrachten

$$\begin{aligned} \frac{d}{dx}(f + g)(x_0) &= \lim_{x \rightarrow x_0} \frac{f(x) + g(x) - [f(x_0) + g(x_0)]}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \left[ \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \right] \\ &= f'(x_0) + g'(x_0). \end{aligned} \quad (\text{A.6.9})$$

Mit zwei Funktionen besitzt also auch deren Summe an der Stelle  $x_0$  eine Ableitung, und diese berechnet sich als die Summe der Ableitungen. Ebenso einfach zeigt man, daß für eine beliebige Konstante  $a \in \mathbb{R}$

$$\frac{d}{dx}(af)(x_0) = af'(x_0) \quad (\text{A.6.10})$$

gilt. Zusammen mit (A.6.8) bedeutet dies, daß die Ableitung eine **lineare Operation** ist, d.h. für beliebigen Funktionen  $f, g$ , die bei  $x_0$  differenzierbar sind und  $a, b \in \mathbb{R}$  gilt

$$\frac{d}{dx}[af(x_0) + bf(x_0)] = a \frac{d}{dx}f(x_0) + b \frac{d}{dx}f(x_0) = af'(x_0) + bf'(x_0). \quad (\text{A.6.11})$$

Weiter gilt die **Produktregel**

$$\frac{d}{dx}(fg)(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0). \quad (\text{A.6.12})$$

Zum Beweis verwenden wir wieder die Definition der Ableitung:

$$\begin{aligned} & \frac{f(x_0 + \Delta x)g(x_0 + \Delta x) - f(x_0)g(x_0)}{\Delta x} \\ &= \frac{[f(x_0 + \Delta x) - f(x_0)]g(x_0 + \Delta x) + f(x_0)[g(x_0 + \Delta x) - g(x_0)]}{\Delta x}. \end{aligned} \quad (\text{A.6.13})$$

Nun ist eine Funktion, die in  $x_0$  differenzierbar ist, dort auch stetig, und damit ergibt sich aus (A.6.13) im Limes  $\Delta x \rightarrow 0$  die Produktregel (A.6.12).

Um zu zeigen, daß eine in  $x_0$  differenzierbare Funktion  $f$  auch stetig ist, gehen wir auf die Definition des Limes zurück. Da  $f$  in  $x_0$  differenzierbar ist, gibt es demnach zu jedem  $\epsilon_0$  ein  $\delta > 0$ , so daß

$$\left| \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - f'(x_0) \right| < \epsilon \quad \text{falls} \quad |\Delta x| < \delta. \quad (\text{A.6.14})$$

Wir können nun diese Ungleichung mit  $|\Delta x|$  multiplizieren:

$$|f(x_0 + \Delta x) - f(x_0) - \Delta x f'(x_0)| < \epsilon |\Delta x|. \quad (\text{A.6.15})$$

Für  $\Delta x \rightarrow 0$  strebt die rechte Seite der Ungleichung sowie  $\Delta x f'(x_0)$  unter dem Betrag auf der linken Seite gegen 0. Damit ist aber

$$\lim_{\Delta x \rightarrow 0} |f(x_0 + \Delta x) - f(x_0)| = 0 \Leftrightarrow \lim_{\Delta x \rightarrow x_0} f(x_0 + \Delta x) = f(x_0), \quad (\text{A.6.16})$$

d.h.  $f$  ist an der Stelle  $x_0$  stetig, und das war zu zeigen.

Weiter können wir nun auch die **Kettenregel** beweisen. Seien  $f$  und  $g$  Funktionen, wobei  $g$  bei  $x_0$  und  $f$  bei  $g(x_0)$  differenzierbar seien. dann gilt

$$\frac{d}{dx} f[g(x_0)] = f'[g(x_0)]g'(x_0). \quad (\text{A.6.17})$$

Zum Beweis verwenden wir wieder die Definition der Ableitung

$$\frac{f[g(x_0 + \Delta x)] - f[g(x_0)]}{\Delta x} = \frac{f[g(x_0 + \Delta x)] - f[g(x_0)]}{g(x_0 + \Delta x) - g(x_0)} \frac{g(x_0 + \Delta x) - g(x_0)}{\Delta x}. \quad (\text{A.6.18})$$

Da nach der gerade durchgeführten Überlegung  $g$  an der Stelle  $x_0$  stetig ist, folgt daraus für  $\Delta x \rightarrow 0$  die Kettenregel (A.6.17).

Mit der Produktregel können wir nun leicht mittels **vollständiger Induktion** die Formel

$$\frac{d}{dx} x^n = nx^{n-1}, \quad n \in \mathbb{N}_0 \quad (\text{A.6.19})$$

beweisen, wobei  $x \in \mathbb{R}$  beliebig sein darf, d.h. Potenzfunktionen (und damit auch Polynome) sind überall differenzierbar. Zum Beweis bemerken wir, daß die Formel sicher für  $n = 0$  gilt. Denn dann ist  $f(x) = x^0 = 1 = \text{const}$ , und die Ableitung verschwindet gemäß (A.6.5). Nehmen wir nun an, die Formel gilt für  $n = k$ . Dann folgt mit der Produktregel

$$\frac{d}{dx} x^{k+1} = \frac{d}{dx} (x x^k) = \frac{dx}{dx} x^k + x \frac{d}{dx} x^k = x^k + x k x^{k-1} = (k+1)x^k, \quad (\text{A.6.20})$$

und das ist genau (A.6.19) für  $n = k + 1$ , und damit ist die Formel nach dem Beweisprinzip der vollständigen Induktion bewiesen.

Weiter ist die Funktion  $f(x) = 1/x$  für  $x \neq 0$  ebenfalls differenzierbar, denn es gilt

$$\frac{1}{\Delta x} \left( \frac{1}{x + \Delta x} - \frac{1}{x} \right) = \frac{1}{\Delta x} \left( \frac{x - (x + \Delta x)}{x(x + \Delta x)} \right) = -\frac{1}{x(x + \Delta x)}. \quad (\text{A.6.21})$$

Mit  $\Delta x \rightarrow 0$  erhalten wir damit für alle  $x \neq 0$

$$\frac{d}{dx} \left( \frac{1}{x} \right) = -\frac{1}{x^2}. \quad (\text{A.6.22})$$

Verwenden wir weiter die Produkt- und die Kettenregel auf die Funktion  $f/g = f \cdot 1/g$ , wobei  $f$  und  $g$  in  $x_0$  differenzierbar und  $g(x_0) \neq 0$  sei, erhalten wir die **Quotientenregel**

$$\frac{d}{dx} \left( \frac{f(x_0)}{g(x_0)} \right) = \frac{d}{dx} \left( f(x_0) \frac{1}{g(x_0)} \right) = \frac{f'(x_0)}{g(x_0)} + f(x_0) \frac{d}{dx} \frac{1}{g(x_0)} = \frac{f'(x_0)}{g(x_0)} - f(x_0) \frac{g'(x_0)}{g^2(x_0)}. \quad (\text{A.6.23})$$

Nochmals zusammengefaßt ergibt sich also

$$\frac{d}{dx} \left( \frac{f(x_0)}{g(x_0)} \right) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}. \quad (\text{A.6.24})$$

Damit können wir nun auch für  $n \in \mathbb{N}$  die Funktionen  $1/x^n$  für alle  $x \neq 0$  ableiten:

$$\frac{d}{dx} \frac{1}{x^n} = \frac{0 - nx^{n-1}}{x^{2n}} = -nx^{-n-1}. \quad (\text{A.6.25})$$

Es gilt also die Ableitungsregel

$$\frac{d}{dx} x^k = kx^{k-1} \quad \text{für alle } k \in \mathbb{Z}. \quad (\text{A.6.26})$$

Schließlich betrachten wir noch eine Funktion  $f : I \rightarrow \mathbb{R}$ , die in einem Intervall  $I$  strikt monoton ist, d.h. es gilt für alle  $x_1 < x_2 \in I$  stets  $f(x_1) < f(x_2)$  (**strikt monoton wachsende**) oder  $f(x_1) > f(x_2)$  (**strikt monoton fallende**) Funktion. Es ist dann klar, daß in diesem Intervall die Zuordnung des Funktionswerts  $y = f(x)$  zum Bildpunkt  $x$  umkehrbar eindeutig ist, d.h. zu jedem Wert  $y$  im Wertebereich der Funktion existiert genau ein  $x \in I$  mit  $f(x) = y$ , d.h. diese Gleichung ist eindeutig nach  $x$  auflösbar. Dies ermöglicht die Definition der **Umkehrfunktion** der Funktion  $f$ :  $f^{-1} : f(I) \rightarrow I$ . Dabei ist  $f(I) = \{y \mid \text{es gibt ein } x \in I \text{ mit } f(x) = y\}$  die **Wertemenge** der Funktion  $f$ . Wir nehmen nun an, die Funktion  $f$  sei differenzierbar im Intervall  $I$ . Aufgrund der Definition der Umkehrfunktion gilt für alle  $x \in I$

$$f^{-1}[f(x)] = x. \quad (\text{A.6.27})$$

Da die rechte Seite differenzierbar und auch  $f$  differenzierbar ist, ist demnach auch  $f^{-1}$  differenzierbar und nach der Kettenregel gilt

$$(f^{-1})'[f(x)]f'(x) = 1 \Rightarrow (f^{-1})'[f(x)] = \frac{1}{f'(x)}. \quad (\text{A.6.28})$$

Mit Hilfe dieser Formel können wir die Ableitung der Umkehrfunktionen monotoner Funktionen finden, wenn wir die Ableitung der Funktion selbst kennen.

Als Beispiel betrachten wir die Exponentialfunktion  $\exp : \mathbb{R} \rightarrow \mathbb{R}$ . Sie ist auf ganz  $\mathbb{R}$  strikt monoton wachsend und positiv besitzt daher eine Umkehrfunktion, den **natürlichen Logarithmus**  $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ . Die Exponentialfunktion kann man sich dadurch definiert denken, daß  $\exp' x = \exp x$  und  $\exp 0 = 1$  gelten soll. Wir werden

unten sehen, daß dies die Berechnung der Exponentialfunktion in Form einer Potenzreihe erlaubt. Aus dieser Ableitungsregel folgt nun

$$\ln(\exp x) = x \Rightarrow \ln'(\exp x) = \frac{1}{\exp' x} = \frac{1}{\exp x}. \quad (\text{A.6.29})$$

Setzt man nun  $y = \exp x$ , folgt

$$\ln' y = \frac{1}{y}, \quad (\text{A.6.30})$$

womit die Ableitungsregel für den natürlichen Logarithmus bestimmt ist.

Wir geben der Vollständigkeit halber noch die Ableitung der **trigonometrischen Funktionen** und ihrer jeweiligen Umkehrfunktionen an<sup>1</sup>. Wir beginnen mit dem Sinus. Aus dem Additionstheorem für den Sinus folgt

$$\frac{\sin(x+h) - \sin(x)}{h} = \frac{\sin x(\cos h - 1) + \cos x \sin h}{h}. \quad (\text{A.6.31})$$

Aus dem Doppelwinkeltheorem für den Cosinus erhalten wir

$$\cos h = \cos\left(2\frac{h}{2}\right) = \cos^2\left(\frac{h}{2}\right) - \sin^2\left(\frac{h}{2}\right) = 1 - 2\sin^2\left(\frac{h}{2}\right) \quad (\text{A.6.32})$$

und damit

$$\frac{\sin(x+h) - \sin(x)}{h} = \frac{\cos x \sin h - 2\sin^2(h/2)}{h}. \quad (\text{A.6.33})$$

Wir behaupten nun, daß

$$\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1. \quad (\text{A.6.34})$$

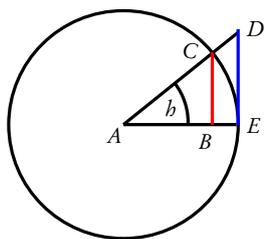
Dann gilt noch

$$\lim_{h \rightarrow 0} \frac{\sin^2(h/2)}{h} = \lim_{h \rightarrow 0} 2\sin(h/2) \frac{\sin(h/2)}{h/2} = 0 \cdot 1 = 0. \quad (\text{A.6.35})$$

Damit folgt aus (A.6.33)

$$\sin' x = \cos x. \quad (\text{A.6.36})$$

Nun müssen wir noch (A.6.34) beweisen. Dazu betrachten wir die nebenstehende Abbildung. Offenbar ist die Fläche des Dreiecks rechtwinkligen  $ABC$  kleiner als die des Kreissegments  $AEC$  und dessen Fläche wiederum kleiner als die des rechtwinkligen Dreiecks  $AED$ . Der Radius des Kreises sei 1. Dann gilt demnach wegen der Definition von Sinus und Cosinus für rechtwinklige Dreiecke



$$\frac{1}{2} \cos h \sin h \leq \frac{1}{2} h \leq \frac{1}{2} \tan h. \quad (\text{A.6.37})$$

Für  $0 < h < \pi/2$  ist  $\sin h > 0$ , und Multiplikation der Ungleichung mit  $2/\sin h$  ergibt

$$\cos h \leq \frac{h}{\sin h} \leq \frac{1}{\cos h} \Rightarrow \cos h \leq \frac{\sin h}{h} \leq \frac{1}{\cos h}. \quad (\text{A.6.38})$$

Nun ist  $\cos h = 1$  und damit folgt im Limes  $h \rightarrow 0^+$  (A.6.34). Da weiter  $\cos(-h) = \cos h$  und  $\sin(-h) = -\sin h$  ist, gilt (A.6.38) auch für  $-\pi/2 < h < 0$ , und damit der Limes (A.6.34) auch für  $h \rightarrow 0^-$ . Damit ist aber gezeigt, daß  $\sin x$  überall differenzierbar ist und (A.6.36) gilt.

<sup>1</sup>Eine strikte Begründung im Rahmen der Analysis erfolgt gewöhnlich über die Definition der trigonometrischen Funktionen  $\sin$  und  $\cos$  über ihre Potenzreihen, aus denen man alle Eigenschaften, die wir aus der Geometrie gewohnt sind, streng herleiten kann. Wir verzichten hier auf diese Begründung. Der interessierte Leser sei auf [For13] verwiesen.

Für den Cosinus erhalten wir

$$\begin{aligned}\cos' x &= \lim_{h \rightarrow 0} \frac{\cos(x+h) - \cos(x)}{h} = \lim_{h \rightarrow 0} \frac{\cos x(\cos h - 1) - \sin x \sin h}{h} \\ &= \lim_{h \rightarrow 0} \frac{2 \cos x \sin^2(h/2) - \sin x \sin h}{h} = -\sin x,\end{aligned}\tag{A.6.39}$$

wobei wir wieder (A.6.32 - A.6.35) verwendet haben. Der Cosinus ist also überall differenzierbar, und es gilt

$$\cos' x = -\sin x.\tag{A.6.40}$$

Mit (A.6.36) und (A.6.40) können wir mit Hilfe der Quotientenregel auch den Tangens ableiten. Für  $x \neq (2n+1)\pi/2, n \in \mathbb{Z}$  folgt

$$\frac{d}{dx} \tan x = \frac{d}{dx} \frac{\sin x}{\cos x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x.\tag{A.6.41}$$

Die trigonometrischen Funktionen sind periodisch und haben daher keine eindeutig definierten **Umkehrfunktionen**. Gewöhnlich definiert man die Umkehrfunktion auf einem eingeschränkten Intervall für den Winkel. Beginnen wir mit dem Cosinus: Dieser ist im Intervall  $[0, \pi]$  streng monoton fallend. Entsprechend definieren wir als Umkehrfunktion  $\arccos: [-1, 1] \rightarrow [0, \pi]$  durch  $\cos(\arccos x) = x$ . Wegen der strikten Monotonie des  $\cos$  im Intervall  $[0, \pi]$  ist durch die Einschränkung  $\arccos x \in [0, \pi]$  der Wert der  $\arccos$ -Funktion eindeutig bestimmt. Mit dem Satz von der Ableitung der Umkehrfunktion folgt aus  $y = \arccos x$ :

$$\frac{d}{dx} \arccos x = \frac{1}{\cos' y} = -\frac{1}{\sin y}.\tag{A.6.42}$$

Nun ist für  $y \in [0, \pi]$  der Sinus positiv, so daß dort  $\sin y = +\sqrt{1 - \cos^2 y}$  ist. Für  $x \in (-1, 1)$ , also  $y \in (0, \pi)$  ist  $\sin y \neq 0$ , und folglich gilt demnach

$$\frac{d}{dx} \arccos x = -\frac{1}{\sqrt{1 - \cos^2 y}} = -\frac{1}{\sqrt{1 - x^2}}.\tag{A.6.43}$$

Demnach haben wir die Ableitungsregel

$$\arccos' x = -\frac{1}{\sqrt{1 - x^2}}, \quad x \in (-1, 1).\tag{A.6.44}$$

Die Ableitung divergiert offensichtlich für  $x \rightarrow \pm 1$ , d.h. der Funktionsgraph des  $\arccos$  besitzt für  $x \rightarrow \pm 1$  senkrechte Tangenten. Die Umkehrfunktion des Sinus definieren wir im Winkelbereich  $x \in [-\pi/2, \pi/2]$ , wo der Sinus strikt monoton wachsend ist. Dann ist  $\arcsin: [-1, 1] \rightarrow [-\pi/2, \pi/2]$  durch  $\sin(\arcsin x) = x$  eindeutig definiert. Analog wie beim Cosinus zeigt man durch Ableiten der Umkehrfunktion (*Übung!*)

$$\arcsin' x = +\frac{1}{\sqrt{1 - x^2}}, \quad x \in (-\pi/2, \pi/2).\tag{A.6.45}$$

Der Tangens ist im offenen Intervall  $(-\pi/2, \pi/2)$  strikt monoton wachsend und nimmt beliebige reelle Werte an, denn  $\lim_{x \rightarrow \pm\pi/2} \tan x = \pm\infty$ . Entsprechend definieren wir die Umkehrfunktion  $\arctan: \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  eindeutig durch  $\tan(\arctan x) = x$ . Über die Ableitung der Umkehrfunktion finden wir mit (A.6.41) (*Übung*)

$$\arctan' x = \frac{1}{1 + x^2}.\tag{A.6.46}$$

Nützlich sind oft auch noch die sog. **Hyperbelfunktionen**. Sie werden mit Hilfe der Exponentialfunktion durch

$$\sinh x = \frac{\exp x - \exp(-x)}{2}, \quad \cosh x = \frac{\exp x + \exp(-x)}{2}, \quad \tanh x = \frac{\sinh x}{\cosh x} = \frac{\exp x - \exp(-x)}{\exp x + \exp(-x)} \quad (\text{A.6.47})$$

für  $x \in \mathbb{R}$  definiert. Sie heißen sinus hyperbolicus, cosinus hyperbolicus bzw. tangens hyperbolicus. Durch direktes Nachrechnen zeigt man, daß für alle  $x \in \mathbb{R}$

$$\cosh^2 x - \sinh^2 x = 1 \quad (\text{A.6.48})$$

gilt. Die Kurve  $\vec{r}(t) = [\cosh t, \sinh t]$ ,  $t \in \mathbb{R}$  beschreibt den Ast einer **Hyperbel**, woher der Name für diese Funktionen stammt. Dies ist in analogie zu den trigonometrischen Funktionen, für die ja  $\vec{r}(t) = [\cos t, \sin t]$  für  $t \in [0, 2\pi)$  einen Einheitskreis beschreibt.

Aus der Ableitung der Exponentialfunktion  $\exp' x = \exp x$  folgt sofort (*nachrechnen!*)

$$\sinh' x = \cosh x, \quad \cosh' x = \sinh x. \quad (\text{A.6.49})$$

Für die Ableitung des tanh verwendet man wieder die Quotientenregel, was zu

$$\tanh' x = 1 - \tanh^2 x = \frac{1}{\cosh^2 x} \quad (\text{A.6.50})$$

führt.

Der sinh ist wegen  $\sinh' x = \cosh x > 0$  streng monoton wachsend und daher überall eindeutig umkehrbar. Die entsprechende Umkehrfunktion  $\operatorname{arsinh} : \mathbb{R} \rightarrow \mathbb{R}$  (area sinus hyperbolicus) ist demnach eindeutig durch  $\sinh(\operatorname{arsinh} x) = x$  definiert. Über die Ableitung der Umkehrfunktion finden wir (*nachrechnen!*)

$$\operatorname{arsinh}' x = \frac{1}{x^2 + 1}. \quad (\text{A.6.51})$$

Der cosh ist eine gerade Funktion und für  $x > 0$  strikt monoton wachsend (für  $x < 0$  strikt monoton fallend). Es gilt stets  $\cosh x \geq 1$ . Daher wird die Umkehrfunktion, der area cosinus hyperbolicus, eindeutig durch  $\operatorname{arcosh} : \mathbb{R}_{\geq 1} \rightarrow \mathbb{R}_{\geq 0}$  via  $\cosh(\operatorname{arcosh} x) = x$  definiert. Die Ableitung der Umkehrfunktion liefert wieder (*nachrechnen!*)

$$\operatorname{arcosh}' x = \frac{1}{x^2 - 1}. \quad (\text{A.6.52})$$

Der tanh ist wegen  $\tanh' x = 1/\cosh^2 x > 0$  überall strikt monoton wachsend und folglich eindeutig umkehrbar. Der Wertebereich des  $\operatorname{artanh}$  ist offensichtlich  $(-1, 1)$ , denn offensichtlich gilt  $\lim_{x \rightarrow \pm\infty} \tanh x = \pm 1$ . Damit ist durch  $\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}$  via  $\tanh(\operatorname{artanh} x) = x$  der area tangens hyperbolicus eindeutig definiert, und mit (A.6.50) erhält man über die Ableitung der Umkehrfunktion (*nachrechnen!*)

$$\operatorname{artanh}' x = \frac{1}{1 - x^2}. \quad (\text{A.6.53})$$

### A.6.3 Kurvendiskussionen, Mittelwertsatz der Differentialrechnung

Wir besprechen als wichtige Anwendung der Differentialrechnung die Untersuchung der **lokalen Eigenschaften** von Funktionen einer reellen Veränderlichen.

Eine Funktion  $f : (a, b) \rightarrow \mathbb{R}$  besitzt in  $\xi \in (a, b)$  ein **lokales Maximum (lokales Minimum)**, wenn es ein  $\epsilon > 0$  gibt, so daß für alle  $x \in (a, b)$  mit  $|x - \xi| < \epsilon$  stets  $f(x) \leq f(\xi)$  ( $f(x) \geq f(\xi)$ ) gilt.

<sup>2</sup>Dies zeigt man am einfachsten über die Taylorreihe von cosh, vgl. Abschnitt A.7.6

Besitzt  $f$  in  $\xi$  ein lokales Maximum oder Minimum sagt, man daß  $f$  in  $\xi$  ein **lokales Extremum** aufweist.

Es gilt der **Satz**: Sei  $f : (a, b) \rightarrow \mathbb{R}$  stetig und wenigstens einmal differenzierbar. Besitzt dann  $f$  in  $\xi \in (a, b)$  ein lokales Extremum, so gilt  $f'(\xi) = 0$ .

Zum Beweis nehmen wir an  $f$  besitze ein lokales Maximum bei  $x = \xi$ . Da  $f$  voraussetzungsgemäß differenzierbar in  $x = \xi$  ist, gilt

$$f'(\xi) = \lim_{h \rightarrow 0^+} \frac{f(\xi + h) - f(\xi)}{h} \leq 0, \quad (\text{A.6.54})$$

denn voraussetzungsgemäß nimmt bei Einschränkung auf eine hinreichend kleine Umgebung  $f$  an der Stelle  $\xi$  den größten Wert an. Ebenso argumentiert man aber, daß

$$f'(\xi) = \lim_{h \rightarrow 0^-} \frac{f(\xi + h) - f(\xi)}{h} \geq 0, \quad (\text{A.6.55})$$

Nun können aber (A.6.54) und (A.6.55) nur beide gelten, wenn  $f'(\xi) = 0$ . Analog zeigt man die Behauptung, falls die Funktion ein Minimum in  $(a, b)$  besitzt.

**Bemerkung:** Die Bedingung  $f'(\xi) = 0$  ist nur *notwendig aber nicht hinreichend* für das Vorliegen eines Extremums. Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$  ist z.B. strikt monoton wachsend, besitzt also sicher kein lokales Extremum. Andererseits gilt  $f'(x) = 3x^2$  und folglich  $f'(0) = 0$ . Die Ableitung von  $f$  verschwindet bei  $x = 0$ , besitzt dort aber weder ein Minimum noch ein Maximum.

Als nächstes wollen wir den **Satz von Rolle** beweisen. Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig auf  $(a, b)$  differenzierbar. Gilt dann  $f(a) = f(b)$ , existiert mindestens eine Stelle  $\xi \in (a, b)$  mit  $f'(\xi) = 0$ .

**Beweis:** Ist  $f(x) = \text{const}$ , ist die Behauptung trivial, denn dann ist  $f'(x) = 0$  für alle  $x \in (a, b)$ . Falls dies nicht der Fall ist, gibt es wenigstens eine Stelle  $x_0$  mit  $f(x_0) > f(a) = f(b)$  oder  $f(x_0) < f(a) = f(b)$ . Da die Funktion  $f$  in dem abgeschlossenen Intervall  $[a, b]$  voraussetzungsgemäß stetig ist, nimmt sie nach dem Satz vom Maximum und Minimum (vgl. Abschnitt A.5) irgendwo in diesem Intervall das Infimum und das Supremum ihres Bildbereiches an. Im ersten Fall muß offenbar das Supremum und im zweiten Fall das Infimum an wenigstens einer Stelle  $\xi$  im offenen Intervall  $(a, b)$  angenommen werden, d.h. es liegt allemal ein Extremum in  $(a, b)$  vor, und nach dem eben bewiesenen Satz ist  $f'(\xi) = 0$ .

Schließlich gilt der **Mittelwertsatz der Differentialrechnung**: Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und auf  $(a, b)$  differenzierbar. Dann gibt es eine Stelle  $\xi \in (a, b)$  mit

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \quad (\text{A.6.56})$$

Zum Beweis wenden wir den Satz von Rolle auf die Funktion

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \quad (\text{A.6.57})$$

an. In der tat erfüllt diese Funktion die Voraussetzungen des Satzes von Rolle. Da  $f$  auf  $[a, b]$  stetig und auf  $(a, b)$  differenzierbar ist, gilt dies sicher auch für  $g$ . Außer dem ist  $g(a) = g(b) = f(a)$ . Folglich gibt es eine Stelle  $\xi \in (a, b)$  mit  $g'(\xi) = 0$ . Nun ist aber

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \Rightarrow 0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}, \quad (\text{A.6.58})$$

und das war zu zeigen.

**Satz von der Monotonie:** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und auf  $(a, b)$  differenzierbar. Dann ist  $f$  genau dann strikt monoton wachsend (monoton fallend) falls  $f'(x) > 0$  ( $f'(x) < 0$ ).

**Beweis:** Sei  $f$  strikt monoton wachsend und  $x \in (a, b)$ . Dann gibt es ein  $\epsilon > 0$ , so daß  $x + b \in [a, b]$  für alle  $b \in (-\epsilon, \epsilon)$ . Für diese  $b$  gilt dann wegen des strikt monotonen Wachsens von  $f$

$$\frac{f(x+b) - f(x)}{b} > 0. \quad (\text{A.6.59})$$

Da  $f$  in  $(a, b)$  differenzierbar ist, erhält man durch Grenzwertbildung  $b \rightarrow 0$ , daß  $f'(x) > 0$ .

Sei nun umgekehrt  $f'(x) > 0$ . Wir nehmen nun an, daß  $f$  nicht strikt monoton wachsend ist. Dann gibt es zwei Stellen  $x_1 < x_2 \in (a, b)$  mit  $f(x_1) \geq f(x_2)$ . Nach dem oben bewiesenen Mittelwertsatz der Differentialrechnung gibt es dann eine Stelle  $\xi \in (a, b)$  mit

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq 0, \quad (\text{A.6.60})$$

und das steht im Widerspruch zur Voraussetzung, daß  $f'(x) > 0$  im ganzen offenen Intervall  $(a, b)$ . Damit muß also  $f$  strikt monoton wachsend sein.

Der Beweis für strikt monoton fallende Funktionen erfolgt analog.

Schließlich geben wir noch ein *hinreichendes Kriterium* für ein lokales Extremum an. Sei  $f : [a, b]$  stetig und auf  $(a, b)$  zweimal differenzierbar. Ist dann  $f'(\xi) = 0$  für ein  $\xi \in (a, b)$  und  $f''(\xi) < 0$  ( $f''(\xi) > 0$ ), so besitzt  $f$  bei  $x = \xi$  ein lokales Maximum (Minimum).

Wir beweisen die Behauptung für den Fall  $f''(\xi) < 0$ . Es gilt

$$f''(\xi) = \lim_{b \rightarrow 0} \frac{f'(\xi + b) - f'(\xi)}{b} = \lim_{b \rightarrow 0} \frac{f'(\xi + b)}{b} < 0. \quad (\text{A.6.61})$$

Dies bedeutet, daß es ein  $\epsilon > 0$  gibt, so daß  $f'(\xi + b)/b < 0$  für  $b \in (-\epsilon, \epsilon) \setminus \{0\}$ . Damit ist aber  $f'(x) < 0$  für  $x \in (\xi, \xi + \epsilon]$  und  $f'(x) > 0$  für  $x \in [\xi - \epsilon, \xi)$ . Damit ist  $f$  zumindest in einer kleinen Umgebung von  $\xi$  für  $x < \xi$  strikt monoton wachsend und für  $x > \xi$  strikt monoton fallend. Damit muß  $f$  bei  $\xi$  ein lokales Maximum besitzen.

Der Beweis, daß für  $f''(\xi) > 0$  und  $f'(\xi) = 0$  ein Minimum vorliegt, folgt analog.

Um zu zeigen, daß das obige Kriterium *hinreichend aber nicht notwendig* ist, betrachten wir die Funktion  $f(x) = x^4$ . Sie besitzt offenbar bei  $x = 0$  ein lokales (sogar ein absolutes) Minimum. In der Tat ist  $f'(x) = 4x^3$  und damit  $f'(0) = 0$ , wie es der Satz von lokalen Extrema differenzierbarer Funktionen sagt. Allerdings gilt weiter  $f''(x) = 12x^2$  und folglich  $f''(0) = 0$ , d.h. obwohl ein Minimum vorliegt ist  $f''(0)$  *nicht*  $> 0$ .

## A.7 Integralrechnung

Die Integralrechnung ist, wie wir sehen werden, in gewissem Sinne die Umkehrung der Differentialrechnung. Sie beschäftigt sich generell mit dem geometrischen Inhaltsbegriff (Bogenlänge einer Kurve, Flächeninhalte von Flächen, Volumen von Körpern) und ist daher eng mit der Maßtheorie verknüpft. In dieser Vorlesung behandeln wir nur das **Riemann-Integral**. Für an der modernen Lebesgue-Integrationstheorie Interessierte sei auf [Wei80] verwiesen, wo man eine Darstellung findet, die nicht die allgemeine Maßtheorie voraussetzt.

### A.7.1 Definition des Riemann-Integrals

Das Ziel der Integralrechnung für reellwertige Funktionen einer reellen Veränderlichen ist, anschaulich formuliert, die Berechnung des Flächeninhaltes der Fläche unter der Kurve der Funktion  $f : [a, b] \rightarrow \mathbb{R}$ , wobei  $[a, b] \subset \mathbb{R}$  ein endliches Intervall ist. Zunächst setzen wir nur voraus, daß die Funktion  $f$  beschränkt sei, d.h. es gibt Zahlen  $m$  und  $M$ , so daß  $m \leq f(x) \leq M$  für alle  $x \in [a, b]$ . Die grundlegende Idee, den Flächeninhalt zu bestimmen ist, das Intervall  $[a, b]$  in  $n$  Teilintervalle  $[x_{j-1}, x_j]$  mit  $j \in \{0, 1, \dots, n\}$  zu zerlegen,

wobei  $a = x_0 < x_1 < \dots < x_n = b$  ist und die Funktion durch eine entsprechende **Treppenfunktion** zu approximieren. Dabei ist eine Treppenfunktion eine Funktion, die auf jedem Teilintervall  $[x_{j-1}, x_j]$  konstant ist.

Nun kann man für jedes Teilintervall das Infimum bzw. das Supremum der Funktionswerte in diesem Intervall verwenden. Nach dem Satz vom Infimum und Supremum existieren nämlich beide Größen, weil ja voraussetzungsgemäß die Funktion  $f$  beschränkt ist. Dann bezeichnen wir als **Unter- und Obersumme** der Funktion  $f$  bzgl. der Zerlegung  $[x_{j-1}, x_j]$  des Intervalls  $[a, b]$  die Größen

$$S_u = \sum_{j=1}^n (x_j - x_{j-1}) \inf\{f([x_{j-1}, x_j])\}, \quad S_o = \sum_{j=1}^n (x_j - x_{j-1}) \sup\{f([x_{j-1}, x_j])\}. \quad (\text{A.7.1})$$

Es ist anschaulich klar, daß der Flächeninhalt der Fläche unter der Kurve für alle Zerlegungen des Intervalls  $[a, b]$  stets zwischen diesen beiden Werten liegen muß und daß wir den Flächeninhalt erhalten, indem wir die Unterteilung immer feiner machen, d.h. wir lassen die Anzahl der Intervalle  $n \rightarrow \infty$  gehen, wobei zugleich  $\max_j \in \{1, \dots, n\} (x_j - x_{j-1}) \rightarrow 0$  gehen soll. Der Flächeninhalt existiert dann sicher, wenn  $S_u$  und  $S_o$  gegen denselben Wert streben. Wir sagen dann, die Funktion  $f$  sei **Riemannintegrierbar** über das Intervall  $[a, b]$  und wir definieren das Integral als den entsprechenden Grenzwert

$$\int_a^b dx f(x) = \lim_{n \rightarrow \infty} S_u = \lim_{n \rightarrow \infty} S_o. \quad (\text{A.7.2})$$

Da bei einer Verfeinerung der Zerlegung von  $[a, b]$  die Untersummen stets größer und die Obersummen stets kleiner werden (*warum?*), kann man das Integral auch als  $\sup S_u = \inf S_o$  definieren, wobei Supremum bzw. Infimum über alle möglichen Zerlegungen zu nehmen ist.

Wir bemerken sogleich, daß wenn  $f, g : [a, b] \rightarrow \mathbb{R}$  Riemann-integrierbar sind, auch die Linearkombination  $\lambda_1 f + \lambda_2 g$  für beliebige  $\lambda_1, \lambda_2 \in \mathbb{R}$  Riemann-integrierbar sind und daß dann

$$\int_a^b dx [\lambda_1 f(x) + \lambda_2 g(x)] = \lambda_1 \int_a^b dx f(x) + \lambda_2 \int_a^b dx g(x) \quad (\text{A.7.3})$$

gilt. Der einfache Beweis sei dem Leser zur *Übung* überlassen.

Nehmen wir weiter an  $f : [a, c]$  sei über die Intervalle  $[a, b]$  und  $[b, c]$  Riemann-integrierbar. Man zeigt dann leicht (*Übung*), daß  $f$  dann auch über  $[a, c]$  Riemann-integrierbar ist und daß dann

$$\int_a^c dx f(x) = \int_a^b dx f(x) + \int_b^c dx f(x). \quad (\text{A.7.4})$$

**Definition:** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beliebige Funktion. Dann definieren wir

$$f_+(x) = \begin{cases} f(x) & \text{falls } f(x) > 0, \\ 0 & \text{falls } f(x) \leq 0, \end{cases} \quad (\text{A.7.5})$$

$$f_-(x) = \begin{cases} -f(x) & \text{falls } f(x) < 0, \\ 0 & \text{falls } f(x) \geq 0. \end{cases}$$

Offenbar gilt dann  $f = f_+ - f_-$  und  $|f| = f_+ + f_-$ .

**Satz:** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Riemann-integrierbare Funktion. Dann sind auch  $f_+, f_-, |f|$  und  $f^2$  Riemann-integrierbar.

Beweis: Nach Voraussetzung ist  $\sup S_u = \inf S_o$ , wobei Supremum und Infimum der Ober- bzw. Untersumme stets über alle Zerlegungen des Intervalls  $[a, b]$  zu nehmen sind. Das bedeutet aber, daß es zu jedem  $\epsilon > 0$  Treppenfunktionen  $\varphi$  und  $\psi$  gibt, so daß  $\varphi \leq f \leq \psi$  ist, so daß

$$0 \leq \int_a^b dx[\psi(x) - \varphi(x)] < \epsilon \quad (\text{A.7.6})$$

ist. Wegen  $\varphi \leq f \leq \psi$  folgt aus (A.7.5) sofort, daß  $\varphi_+ \leq f_+ \leq \psi_+$  und  $\psi_- \leq f_- \leq \varphi_-$ . Nun sind  $\varphi_{\pm}$  und  $\psi_{\pm}$  wieder Treppenfunktionen, und es gilt

$$\begin{aligned} 0 \leq \int_a^b dx[\psi_+(x) - \varphi_+(x)] &\leq \int_a^b dx[\psi(x) - \varphi(x)] < \epsilon, \\ 0 \leq \int_a^b dx[\varphi_-(x) - \psi_-(x)] &\leq \int_a^b dx[\psi(x) - \varphi(x)] < \epsilon. \end{aligned} \quad (\text{A.7.7})$$

Zu jedem  $\epsilon > 0$  gibt es also Zerlegungen des Intervalls, so daß Ober- und Untersumme von  $f_{\pm}$  um weniger als  $\epsilon$  voneinander abweichen. Folglich stimmen das Supremum der Unter- und das Infimum der Obersummen überein, und die Funktionen  $f_{\pm}$  sind folglich Riemann-integrierbar. Da mit ist auch  $|f| = f_+ + f_-$  Riemann-integrierbar.

Weiter ist  $f^2 = |f|^2$ . Da  $|f|$  Riemann-integrierbar ist, gibt es zu jedem  $\epsilon > 0$  Treppenfunktionen  $\tilde{\psi}, \tilde{\varphi} : [a, b] \rightarrow \mathbb{R}$  mit  $\tilde{\varphi} \leq |f| \leq \tilde{\psi}$ , so daß

$$\int_a^b dx[\psi(x) - \varphi(x)] < \frac{\epsilon}{2M} \quad \text{mit} \quad M = \sup_{x \in [a, b]} |f(x)| - \inf_{x \in [a, b]} |f(x)|. \quad (\text{A.7.8})$$

Dabei gehen wir davon aus, daß  $f$  nicht auf  $[a, b]$  konstant ist. In dem Fall wäre auch  $f^2$  konstant und somit auch  $f^2$  integrierbar, und die Behauptung des Satzes also erfüllt. Da nun die Funktion  $g(x) = x^2$  für  $x \geq 0$  monoton wachsend ist, gilt auch

$$\tilde{\varphi}^2 \leq |f|^2 \leq \tilde{\psi}^2, \quad (\text{A.7.9})$$

und  $\tilde{\varphi}$  und  $\tilde{\psi}$  sind Treppenfunktionen. Nach dem Mittelwertsatz der Differentialrechnung, angewandt auf  $g(x) = x^2$ , gilt für beliebige Zahlen  $\tilde{\psi}^2 - \tilde{\varphi}^2 = g'(\xi)(\psi - \varphi) = 2\xi(\psi - \varphi) \leq 2M(\psi - \varphi)$  mit  $\varphi \leq \xi \leq \psi$ . Damit ist also

$$0 \leq \int_a^b dx[\tilde{\psi}^2(x) - \tilde{\varphi}^2(x)] \leq 2M \int_a^b dx[\tilde{\psi}(x) - \tilde{\varphi}(x)] \leq \epsilon. \quad (\text{A.7.10})$$

Folglich gibt es zu jedem  $\epsilon$  Treppenfunktionen  $\tilde{\varphi}^2$  und  $\tilde{\psi}^2$  mit  $\tilde{\varphi}^2 \leq f^2 \leq \tilde{\psi}^2$ , die (A.7.10) erfüllen.

**Satz:** Seien weiter  $f, g : [a, b] \rightarrow \mathbb{R}$  und Riemann-integrierbare Funktionen. Dann ist auch  $f g$  Riemann-integrierbar.

**Beweis:** Nach den obigen Sätzen sind  $f + g$  und  $f - g$  sowie  $(f + g)^2$  und  $(f - g)^2$  Riemann-integrierbar. Damit ist aber auch

$$f g = \frac{1}{4}[(f + g)^2 - (f - g)^2] \quad (\text{A.7.11})$$

Riemann-integrierbar.

**Satz Stetige Funktionen**  $f : [a, b] \rightarrow \mathbb{R}$  sind Riemann-integrierbar.

**Beweis:** Dazu zeigen wir zuerst, daß in diesem Fall  $f$  sogar **gleichmäßig** stetig ist, d.h. zu jedem  $\epsilon > 0$  existiert ein  $\delta > 0$ , so daß für alle  $x, x' \in [a, b]$  mit  $|x - x'| < \delta$  stets  $|f(x) - f(x')| < \epsilon$  ist. Zum Beweis nehmen wir an, daß  $f$  zwar stetig aber nicht gleichmäßig stetig ist. Dann gibt es zu einem  $\epsilon > 0$  zu jedem  $n \in \mathbb{N}$  Zahlen  $x_n, x'_n \in [a, b]$  mit  $|x_n - x'_n| < 1/n$ , so daß  $|f(x_n) - f(x'_n)| \geq \epsilon$ . Die Zahlenfolgen  $(x_n)_n$  und  $(x'_n)_n$  sind beschränkt

und besitzen daher konvergente Teilfolgen  $(x_{n_k})_k$  bzw.  $(x'_{n_k})_k$ . Wegen der Voraussetzung, daß  $|x_n - x'_n| < 1/n$  gilt  $\xi := \lim_{k \rightarrow \infty} x_{n_k} = \lim_{k \rightarrow \infty} x'_{n_k}$ . Folglich existiert auch  $\lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{k \rightarrow \infty} f(x'_{n_k}) = f(\xi)$ , weil  $f$  voraussetzungsgemäß stetig auf dem Intervall  $[a, b]$  ist. Andererseits soll aber gemäß unserer obigen Annahme  $|f(x_{n_k}) - f(x'_{n_k})| \geq \epsilon$  für alle  $k \in \mathbb{N}$  gelten. Dann wären aber die Folgen  $(f(x_{n_k}))_k$  und  $(f(x'_{n_k}))_k$  wiederum nicht konvergent im Widerspruch zu unserer gerade durchgeführten Argumentation. Folglich muß also  $f$  auf  $[a, b]$  gleichmäßig stetig sein.

Nun können wir leicht zeigen, daß eine auf  $[a, b]$  stetige Funktion Riemann-integrierbar ist. Sei dazu  $\epsilon > 0$  beliebig gewählt. Dann gibt es ein  $\delta > 0$ , so daß  $|f(x) - f(x')| < \epsilon/(b-a)$  ist, wenn nur  $|x - x'| < \delta$ . Es sei nun  $I_j = (x_{j-1}, x_j)$  eine so feine Zerlegung des Intervalls  $[a, b]$ , so daß  $\max_j (x_j - x_{j-1}) < \delta$  ist. Weiter gibt es nach dem Satz vom Minimum und Maximum (s. Abschnitt A.5)  $\xi_j \in I_j$  und  $\xi'_j \in I_j$ , so daß  $\inf_{x \in I_j} f(x) = f(\xi_j)$  und  $\sup_{x \in I_j} f(x) = f(\xi'_j)$ . Dann ist sicher  $|\xi_j - \xi'_j| < \delta$  und folglich

$$0 \leq S_o - S_u = \sum_{j=1}^n (x_j - x_{j-1}) [f(\xi'_j) - f(\xi_j)] < \frac{\epsilon}{b-a} \sum_{j=1}^n (x_j - x_{j-1}) = \epsilon. \quad (\text{A.7.12})$$

Folglich kann man durch beliebige Verfeinerung der Zerlegung Ober- und Untersumme beliebig aneinander annähern. Folglich existieren das Infimum der Ober- und das Supremum der Untersummen, und beide sind gleich. Folglich ist  $f$  über das Intervall  $[a, b]$  Riemann-integrierbar.

### A.7.2 Der Mittelwertsatz der Integralrechnung

**Mittelwertsatz der Integralrechnung:** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und  $\varphi : [a, b] \rightarrow \mathbb{R}$  mit  $\varphi \geq 0$  integrierbar. Dann gibt es eine Zahl  $\xi \in [a, b]$ , so daß

$$\int_a^b dx f(x) \varphi(x) = f(\xi) \int_a^b dx \varphi(x). \quad (\text{A.7.13})$$

**Beweis:** Nach den obigen Sätzen zur Riemann-Integrierbarkeit ist mit den obigen Voraussetzungen  $f\varphi$  Riemann-integrierbar. Da  $f$  auf dem abgeschlossenen Intervall stetig ist, existieren  $M = \sup_{x \in [a, b]} f(x)$  und  $m = \inf_{x \in [a, b]} f(x)$ . Da weiter  $\varphi \geq 0$  ist, folgt  $m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x)$ , und damit

$$m \int_a^b dx \varphi(x) \leq \int_a^b f(x) \varphi(x) \leq M \int_a^b dx \varphi(x). \quad (\text{A.7.14})$$

Da auch  $\int_a^b dx \varphi(x) \geq 0$ , gibt es demnach eine Zahl  $\mu$  mit  $m \leq \mu \leq M$ , so daß

$$\int_a^b dx f(x) \varphi(x) = \mu \int_a^b dx \varphi(x). \quad (\text{A.7.15})$$

Da  $f : [a, b] \rightarrow \mathbb{R}$  stetig ist, nimmt  $f$  die Werte  $m$  und  $M$  tatsächlich für irgendwelche Werte  $\xi_1, \xi_2 \in [a, b]$  tatsächlich an:  $f(\xi_1) = m$ ,  $f(\xi_2) = M$ . Wegen des Zwischenwertsatzes für stetige Funktionen (vgl. Abschnitt A.5) gibt es eine Zahl  $\xi$  zwischen  $\xi_1$  und  $\xi_2$  mit  $f(\xi) = \mu$ . Wegen (A.7.15) ist damit der Zwischenwertsatz der Integralrechnung bewiesen.

### A.7.3 Der Hauptsatz der Analysis

Nach dem vorigen Abschnitt können wir für eine auf  $[a, b]$  stetige Funktion  $f$  die **Integralfunktion**

$$F(x) = \int_a^x dx' f(x') \quad (\text{A.7.16})$$

definieren. Wir wollen nun den **Hauptsatz der Analysis** beweisen, demzufolge dann  $F$  eine auf  $[a, b]$  differenzierbare Funktion ist und

$$F'(x) = f(x) \quad (\text{A.7.17})$$

gilt. Offenbar ist für  $\Delta x > 0$

$$F(x + \Delta x) - F(x) = \int_x^{x+\Delta x} dx' f(x'). \quad (\text{A.7.18})$$

Sei nun  $m = \inf_{x' \in [x, x+\Delta x]} f(x')$  und  $M = \sup_{x' \in [x, x+\Delta x]} f(x')$ . Beide Werte existieren, und es gibt nach dem Satz von Bolzano-Weierstraß Zahlen  $\xi_{\Delta x}, \xi'_{\Delta x} \in [x, x + \Delta x]$ , so daß  $m = f(\xi_{\Delta x})$  und  $M = f(\xi'_{\Delta x})$ . Demnach ist

$$\int_x^{x+\Delta x} dx f(\xi_{\Delta x}) \leq F(x + \Delta x) - F(x) \leq \int_x^{x+\Delta x} dx f(\xi'_{\Delta x}) \quad (\text{A.7.19})$$

oder nach Ausführung der Integrale

$$\Delta x f(\xi_{\Delta x}) \leq F(x + \Delta x) - F(x) \leq \Delta x f(\xi'_{\Delta x}). \quad (\text{A.7.20})$$

Dividieren wir durch  $\Delta x$ , erhalten wir

$$f(\xi_{\Delta x}) \leq \frac{F(x + \Delta x) - F(x)}{\Delta x} \leq f(\xi'_{\Delta x}). \quad (\text{A.7.21})$$

Für  $\Delta x \rightarrow 0$  gehen nun  $\xi_{\Delta x}, \xi'_{\Delta x} \rightarrow x$  und wegen der Stetigkeit von  $f$  entsprechen  $f(\xi_{\Delta x}), f(\xi'_{\Delta x}) \rightarrow f(x)$ , womit (A.7.17) bewiesen ist. Analog gehen wir für  $\Delta x < 0$  vor. Wir definieren dabei ganz allgemein für Integrale

$$\int_a^b dx f(x) = - \int_b^a dx f(x) \quad \text{falls } a < b. \quad (\text{A.7.22})$$

Für  $\Delta x < 0$  müssen wir dann nur die obige Betrachtung für das Integral

$$\int_x^{x+\Delta x} dx f(x) = - \int_{x+\Delta x}^x dx f(x) \quad (\text{A.7.23})$$

anstellen. Dann ergibt sich auch für den Grenzwert  $\Delta x \rightarrow 0$  für  $\Delta x < 0$ , daß  $F'(x) = f(x)$  ist. Damit ist  $F$  differenzierbar und besitzt die Ableitung  $F' = f$  für alle  $x \in [a, b]$ .

Wir bemerken noch, daß demnach für jede **Stammfunktion**  $\tilde{F} : [a, b] \rightarrow \mathbb{R}$  von  $f$ , gilt

$$\int_a^b dx f(x) = \tilde{F}(b) - \tilde{F}(a). \quad (\text{A.7.24})$$

Dabei heißt  $\tilde{F}$  Stammfunktion zu  $f$ , falls

$$\tilde{F}'(x) = f(x) \quad (\text{A.7.25})$$

gilt. Zum Beweis zeigen wir, daß  $\tilde{F}'(x) = F'(x) + \tilde{F}'(a)$  ist. Es ist nämlich  $\tilde{F}'(x) = F'(x) = f(x)$  und also  $\tilde{F}'(x) - F'(x) = 0$ . Nun muß eine auf  $[a, b]$  differenzierbare Funktion, deren Ableitung überall im Intervall  $[a, b]$  verschwindet, konstant sein. Da  $\tilde{F}'(a) = 0$  ist, gilt also die Behauptung.

Man bezeichnet eine beliebige Stammfunktion von  $f$  auch als **unbestimmtes Integral** und schreibt kurz

$$\tilde{F}(x) = \int dx f(x) + C, \quad C = \text{const.} \quad (\text{A.7.26})$$

### A.7.4 Integrationstechniken

Mit dem Hauptsatz der Differential- und Integralrechnung können wir nun sehr viele Integrale berechnen, indem wir eine beliebige Stammfunktion zu der zu integrierenden Funktion suchen. Dafür gibt es einige Rechenstechniken, die wir hier kurz zusammenfassen. Zunächst kann man sich aus jeder Ableitungsregel eine Integrationsregel ableiten. Z.B. folgt aus der Ableitung von Potenzfunktionen  $(x^n)' = nx^{n-1}$  die Regel

$$\int dx x^n = \frac{x^{n+1}}{n+1} + C. \quad (\text{A.7.27})$$

Dies gilt für alle  $n \in \mathbb{Z} \neq \{-1\}$ . Die Ausnahme ist  $n = -1$ . Aber auch hier kennen wir eine Stammfunktion:

$$\int dx \frac{1}{x} = \ln x + C. \quad (\text{A.7.28})$$

Für  $n < 0$  müssen wir bei der Anwendung des Hauptsatzes aufpassen: Die entsprechenden Funktionen  $f(x) = x^n$  besitzen nämlich dann bei  $x = 0$  eine Singularität und sind insbesondere unstetig. Sie können daher *nur* über Intervalle  $[a, b]$  integriert werden für die  $0 \notin [a, b]$  ist, und nur dann liefert der Hauptsatz der Analysis den korrekten Wert für das Integral!

Weitere Grundintegrale sind

$$\begin{aligned} \int dx \sin x &= -\cos x + C, & \int dx \cos x &= \sin x + C, \\ \int dx \frac{1}{\cos^2 x} &= \int dx (1 + \tan^2 x) = \tan x + C, & \int dx \frac{1}{\sin^2 x} &= \int dx (1 + \cot^2 x) = -\cot x + C, \\ \int dx \exp x &= \exp x + C, & \int dx \sinh x &= \cosh x + C, & \int dx \cosh x &= \sinh x + C, \\ \int dx \frac{1}{\cosh^2 x} &= \int dx (1 - \tanh^2 x) = \tanh x + C, \\ \int dx \frac{1}{\sinh^2 x} &= \int dx (\coth^2 x - 1) = -\coth x + C. \\ \int dx \frac{1}{\sqrt{1-x^2}} &= \arcsin x + C = -\arccos x + C', & \int dx \frac{1}{1+x^2} &= \arctan x + C = -\operatorname{arccot} x + C', \\ \int dx \frac{1}{1-x^2} &= \operatorname{artanh} x + C \quad \text{falls } x < 1, & \int dx \frac{1}{1-x^2} &= \operatorname{arcoth} x + C \quad \text{falls } x > 1, \\ \int dx \frac{1}{\sqrt{1+x^2}} &= \operatorname{arsinh} x + C, & \int dx \frac{1}{\sqrt{x^2-1}} &= \operatorname{arcosh} x + C \quad \text{falls } x > 1. \end{aligned} \quad (\text{A.7.29})$$

Aus allgemeinen Regeln der Differentialrechnung erhalten wir entsprechend Rechenregeln für unbestimmte Integrale. Aus der Produktregel der Integralrechnung folgt die **Regel von der partiellen Integration**. Aus  $(uv)' = u'v + v'u$  folgt

$$\int dx u'(x)v(x) = u(x)v(x) - \int dx u(x)v'(x) + C. \quad (\text{A.7.30})$$

Als Anwendungsbeispiel betrachten wir das Integral von  $x \exp x$ . Setzen wir  $u'(x) = \exp x$  und  $v(x) = x$ , gilt  $u(x) = \exp x$  und  $v'(x) = 1$ . Damit ist

$$\int dx x \exp x = x \exp x - \int dx \exp x = (x-1)\exp x + C. \quad (\text{A.7.31})$$

In der Tat findet man durch Ableiten, daß tatsächlich  $[(x-1)\exp x]' = \exp x + (x-1)\exp x = x \exp x$  gilt. Ein weiteres Beispiel ist  $\int dx \ln x$ . Hier setzen wir  $u'(x) = 1$  und  $v(x) = \ln x$ . Dann ist  $u(x) = x$  und  $v'(x) = 1/x$ . Damit liefert (A.7.30)

$$\int dx \ln x = x \ln x - \int dx 1 = x(\ln x - 1) + C. \quad (\text{A.7.32})$$

In der Tat folgt aus der Produktregel  $[x(\ln x - 1)]' = (\ln x - 1) + x/x = \ln x$ .

Die Kettenregel der Differentialrechnung liefert die **Substitutionsregel für Integrale**. Sei dazu  $F$  eine Stammfunktion von  $f$ .

$$\frac{d}{du} F[x(u)] = \frac{d}{dx} F[x(u)] \frac{dx(u)}{du} x = f[x(u)] \frac{dx(u)}{du} \Rightarrow \int du \frac{dx(u)}{du} f[x(u)] = F[u(x)] + C. \quad (\text{A.7.33})$$

Diese Regel läßt sich leichter merken, wenn man formal  $du = dx u'(x)$  schreibt. Dann wird (A.7.34)

$$\int du \frac{dx}{du} f[x(u)] = \left[ \int dx f(x) \right]_{x=x(u)}. \quad (\text{A.7.34})$$

Als Beispiel betrachten wir die Aufgabe, die Fläche eines Halbkreises um den Ursprung mit Radius  $r$  zu berechnen. Für den Halbkreis gilt  $x^2 + y^2 = r^2$  bzw.  $y(x) = \sqrt{r^2 - x^2}$  mit  $x \in [-r, r]$ . Damit ist die Fläche

$$A = \int_{-r}^r dx \sqrt{r^2 - x^2}. \quad (\text{A.7.35})$$

Um das Integral zu berechnen, substituieren wir nun  $x = r \cos u$ . Dann ist  $dx = -du r \sin u$ . Für  $x = -r$  ist offenbar  $u = \pi$  und  $x = r$  für  $u = 0$ . Damit folgt

$$A = \int_{\pi}^0 du (-r \sin u) \sqrt{r^2 - r^2 \cos^2 u} = r^2 \int_0^{\pi} du \sin^2 u. \quad (\text{A.7.36})$$

Um schließlich dieses Integral zu berechnen, bemerken wir, daß

$$\cos(2u) = \cos^2 u - \sin^2 u = 1 - 2 \sin^2 u \Rightarrow \sin^2 u = \frac{1}{2} [1 - \cos(2u)]. \quad (\text{A.7.37})$$

Dies in (A.7.36) eingesetzt liefert

$$A = \frac{r^2}{2} \int_{-\pi}^{\pi} [1 - \cos(2u)] = \frac{r^2}{2} \left[ u - \frac{1}{2} \sin(2u) \right]_{u=0}^{u=\pi} = \frac{\pi r^2}{2}. \quad (\text{A.7.38})$$

### A.7.5 Funktionenfolgen und -reihen

In diesem Abschnitt untersuchen wir Funktionen, die sich als Grenzwerte von Folgen oder Reihen ergeben, deren Glieder selbst Funktionen sind. Insbesondere interessiert uns die Frage, ob Folgen oder Reihen stetiger Funktionen selbst gegen stetige Funktionen konvergieren und ob man ggf. Operationen wie die Differentiation und Integration mit der Grenzwertbildung vertauschen darf. Dabei ist es wichtig zu beachten, in welchem Sinne die Funktionenfolgen und -reihen konvergent sein sollen. Dazu definieren wir zwei Begriffe:

Es sei  $(f_n(x))_n$  eine Folge von Funktionen  $f_n : D \rightarrow \mathbb{R}$  mit einem beliebigen Definitionsbereich  $D \subseteq \mathbb{R}$ . Wir sagen, die Funktionenfolge konvergiere **punktweise** im Definitionsbereich  $D$  gegen eine Funktion  $f : D \rightarrow \mathbb{R}$ , wenn sie für jedes  $x \in D$  konvergiert. D.h. zu jedem  $x$  existiert zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so daß  $|f_n(x) - f(x)| < \epsilon$  für alle  $n > N$ .

Davon zu unterscheiden ist die stärkere Forderung nach **gleichmäßiger Konvergenz** gegen die Funktion  $f$ . Dazu verlangen wir, daß zu jedem  $\epsilon$  ein  $N \in \mathbb{N}$  existiert, so daß für alle  $x \in D$  stets  $|f_n(x) - f(x)| < \epsilon$  für alle  $n > N$  gilt.

Im Unterschied zur punktweisen Konvergenz, muß also zu jedem  $\epsilon$  das  $N$  unabhängig von  $x \in D$  gewählt werden können, d.h. die Forderung nach gleichmäßiger Konvergenz ist stärker als die nach punktwieser.

Man kann die gleichmäßige Konvergenz auch noch anders beschreiben. Dies macht die Arbeit mit diesem Begriff erheblich einfacher. Dazu führen wir einen Abstandsbegriff, eine **Norm**, für Funktionen ein, und zwar die **Supremumsnorm**. Sei also  $f : D \rightarrow \mathbb{R}$  eine Funktion mit beliebigem Definitionsbereich  $D \subseteq \mathbb{R}$ . Dann ist die **Supremumsnorm** durch

$$\|f\| = \sup_{x \in D} |f(x)|. \quad (\text{A.7.39})$$

Aus der Definition geht unmittelbar hervor, daß die folgenden Regeln für eine Norm gelten (*Beweis als Übung*):

$$\|f\| \geq 0, \quad \|f\| = 0 \Leftrightarrow f(x) = 0 \quad \text{für alle } x \in D, \quad (\text{A.7.40})$$

$$\|\lambda f\| = |\lambda| \|f\|, \quad (\text{A.7.41})$$

$$\|f + g\| \leq \|f\| + \|g\|. \quad (\text{A.7.42})$$

Eine Funktionenfolge ist dann genau dann gleichmäßig konvergent, wenn sie im Sinne dieser Norm konvergiert. Nehmen wir an die Funktionenfolge  $(f_n)_n$  konvergiert gleichmäßig gegen  $f$ . Dann gibt es voraussetzungsgemäß zu jedem  $\epsilon \in \mathbb{R}$  ein  $N \in \mathbb{N}$ , so daß für alle  $x$  stets  $|f_n(x) - f(x)| < \epsilon/2$  gilt, wenn nur  $n > N$  ist. Dann ist aber auch  $\|f_n - f\| = \sup_{x \in D} |f_n(x) - f(x)| \leq \epsilon/2 < \epsilon$ . Kurz: Falls  $(f_n)_n$  gleichmäßig konvergiert, gibt es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so daß  $\|f_n - f\| < \epsilon$  für alle  $n > N$  gilt. Das bedeutet aber, daß die Funktionenfolge im Sinne der Supremumsnorm gegen  $f$  konvergiert.

Ist umgekehrt im Sinne die Funktionenfolge  $(f_n)$  im Sinne der Supremumsnorm konvergent gegen  $f$ , so gibt es zu jedem  $\epsilon > 0$  ein  $n \in \mathbb{N}$ , so daß  $\|f_n - f\| = \sup_{x \in D} |f_n(x) - f(x)| < \epsilon$  für alle  $n > N$ . Dann gilt aber für alle  $x \in D$  offenbar  $|f_n(x) - f(x)| < \epsilon$ , d.h. die Funktionenfolge ist gleichmäßig konvergent.

**Satz von der Stetigkeit stetiger Funktionenfolgen:** Seien alle Glieder der  $f_n : D \rightarrow \mathbb{R}$  eine gleichmäßig gegen  $f$  konvergenten Funktionenfolge stetig. Dann ist auch  $f$  stetig.

**Beweis:** Wegen der gleichmäßigen Konvergenz der Funktionenfolge  $(f_n)_n$  gegen  $f$ , gibt es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so daß für alle  $x \in D$  für alle  $n > N$  stets  $|f_n(x) - f(x)| < \epsilon/3$  gilt. Sei nun  $n > N$  festgehalten. Da weiter die  $f_n$  stetig in  $x \in D$  sind, gibt es zu jedem  $\epsilon > 0$  ein  $\delta > 0$ , so daß für alle  $x' \in D$  mit  $|x - x'| < \delta$  stets  $|f_n(x) - f_n(x')| < \epsilon/3$  ist. Damit folgt nun aber, daß

$$|f(x) - f(x')| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x')| + |f_n(x') - f(x')|. \quad (\text{A.7.43})$$

Wegen der Wahl von  $n > N$  sind die beiden letzten Terme beide  $< \epsilon/3$ , während der mittlere für  $|x - x'| < \delta$  ebenfalls  $< \epsilon/3$  ist. Wir haben also zu einem beliebigen  $\epsilon > 0$  ein  $\delta > 0$  gefunden, so daß für alle  $x' \in D$  mit  $|x - x'| < \delta$  stets  $|f(x) - f(x')| < \epsilon$  gilt. Damit ist  $f$  definitionsgemäß stetig, und das war zu zeigen.

Oft ist es wichtig zu wissen, ob man bestimmte Operationen an Funktionenfolgen mit der Limesbildung vertauschen darf. Hier interessieren uns Integration und Differentiation.

**Integration von Funktionenfolgen:** Seien  $f_n : [a, b] \rightarrow \mathbb{R}$  stetige Funktionen und die Funktionenfolge  $(f_n)_n$  auf  $[a, b]$  gleichmäßig konvergent gegen eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$ . Dann gilt

$$\int_a^b dx f(x) = \int_a^b dx \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \int_a^b dx f_n(x). \quad (\text{A.7.44})$$

Kurz gesagt darf man bei gleichmäßiger Konvergenz von Funktionenfolgen die Grenzwertbildung und die Integration vertauschen.

**Beweis:** Nach dem vorigen Satz ist  $f$  stetig auf  $[a, b]$  und folglich Riemann-integrierbar. Wegen der gleichmäßigen Konvergenz der Funktionenfolge  $(f_n)_n$  gegen  $f$  gilt

$$\left| \int_a^b dx f(x) - \int_a^b dx f_n(x) \right| \leq \int_a^b dx |f(x) - f_n(x)| \leq (b-a) \|f - f_n\| \xrightarrow{n \rightarrow \infty} 0, \quad (\text{A.7.45})$$

d.h. der Limes der Integrale existiert, und es gilt (A.7.44) und das war zu zeigen.

**Differentiation von Funktionenfolgen:** Seien  $f_n : [a, b] \rightarrow \mathbb{R}$  stetig differenzierbare Funktionen und die Funktionenfolge  $(f_n)_n$  punktweise gegen eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  konvergent. Die Folge  $(f'_n)_n$  der Ableitungen konvergiere gleichmäßig. Dann ist  $f$  differenzierbar, und es gilt

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x). \quad (\text{A.7.46})$$

MaW. darf man unter den angegebenen Voraussetzung Grenzwertbildung und Differentiation vertauschen.

**Beweis:** Es sei  $\tilde{f}(x) = \lim_{n \rightarrow \infty} f'_n(x)$  für  $x \in [a, b]$ . Da die  $f'_n$  allesamt stetig sind und die Folge der Ableitungen voraussetzungsgemäß gleichmäßig konvergiert, ist  $\tilde{f}$  nach dem oben bewiesenen Satz stetig und besitzt eine Stammfunktion, denn nach dem Satz über die Vertauschbarkeit der Integration mit der Grenzwertbildung für gleichmäßig konvergente Folgen stetiger Funktionen gilt

$$\int_a^x dx' \tilde{f}(x') = \lim_{n \rightarrow \infty} \int_a^x dx' f'_n(x) = \lim_{n \rightarrow \infty} [f_n(x) - f_n(a)]. \quad (\text{A.7.47})$$

Wegen der vorausgesetzten punktweisen Konvergenz von  $(f_n)_n$  gegen  $f$  gilt also

$$\int_a^x dx' \tilde{f}(x') = f(x) - f(a). \quad (\text{A.7.48})$$

Ableiten dieser Gleichung nach  $x$  liefert nach dem Hauptsatz der Differential- und Integralrechnung schließlich  $f'(x) = \tilde{f}(x)$ , und das war zu zeigen.

Schließlich definieren wir eine **Funktionenreihe**  $\sum_{k=1}^{\infty} f_k(x)$  als gleichmäßig konvergent gegen  $f$ , wenn ihre **Partialsammenfolge**  $(S_n(x))_n$ ,

$$S_n(x) = \sum_{k=1}^n f_k(x), \quad (\text{A.7.49})$$

gleichmäßig (also im Sinne der Supremumsnorm) gegen  $f$  konvergiert.

**Weierstraßsches Konvergenzkriterium:** Seien  $f_n : D \rightarrow \mathbb{R}$  Funktionen und

$$\sum_{k=1}^{\infty} \|f_k\| \quad \text{konvergent}, \quad (\text{A.7.50})$$

so ist die Funktionenreihe  $\sum f_n$  absolut und gleichmäßig konvergent.

**Beweis:** Zunächst zeigen wir, daß die Reihe punktweise absolut konvergiert. Sei also  $x \in D$ . Dann gilt  $|f_k(x)| \leq \|f_k\|$ . Da voraussetzungsgemäß (A.7.50) gilt, ist also  $\sum_k |f_k(x)|$  konvergent. Nun definieren wir im Sinne dieser *punktweisen* Konvergenz

$$f(x) = \sum_{k=1}^{\infty} f_k(x) \quad (\text{A.7.51})$$

und zeigen, daß die Funktionenreihe sogar gleichmäßig gegen  $F$  konvergiert. Seien dazu die Partialsammen durch (A.7.49) definiert. Dann gilt

$$|S_n(x) - f(x)| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\|. \quad (\text{A.7.52})$$

Wegen (A.7.50) können wir nun zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  finden, so daß die letztere Summe  $< \epsilon$  für alle  $n > N$  ist. Damit gilt für alle  $x \in D$

$$|S_n(x) - f(x)| < \epsilon \quad \text{für } n > N, \quad (\text{A.7.53})$$

und folglich ist die Funktionenreihe tatsächlich nicht nur punktweise sondern sogar gleichmäßig gegen  $f$  konvergent, und das war zu zeigen.

### A.7.6 Taylor-Entwicklung und Potenzreihen

Sei also  $f : D \rightarrow \mathbb{R}$  mit offenem Definitionsbereich  $D \subseteq \mathbb{R}$  eine Funktion, die in einem abgeschlossenen Intervall  $I \subset D$ , mindestens  $(n + 1)$ -mal ( $n \in \mathbb{N}_0$ ) stetig differenzierbar ist und sei  $a \in D$ . Wir suchen dann eine Näherung von  $f$  durch ein **Polynom**  $n$ -ten Grades für Argumente „in der Nähe“ von  $a$ . Sei  $x \in I$ . Dann gibt es Koeffizienten  $c_k \in \mathbb{R}$ , so daß

$$f(x) = \sum_{k=0}^n c_k (x-a)^k + R_n(x, a) \quad (\text{A.7.54})$$

ist, wobei das Restglied  $R_n(x, a) = \mathcal{O}[(x-a)^{n+1}]$  ist, d.h.

$$\lim_{n \rightarrow \infty} \frac{R_n(x, a)}{(x-a)^{n+1}} = \text{const.} \quad (\text{A.7.55})$$

Um die Koeffizienten  $c_k$  in (A.7.54) zu bestimmen, leiten wir diese Gleichung  $k$ -mal ( $k \in \{0, 1, \dots, n\}$ ) ab und setzen danach  $x = a$ :

$$\left. \frac{d^k}{dx^k} f(x) \right|_{x=a} =: f^{(k)}(a) = k! c_k, \quad (\text{A.7.56})$$

denn wegen (A.7.55) verschwindet die  $k$ -te Ableitung des Restglieds für  $x = a$ . Damit wird

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_n(x, a). \quad (\text{A.7.57})$$

Dies ist die **Taylor'sche Formel**.

Zum Beweis bemerken wir, daß wegen der Stetigkeit von  $f'(x)$

$$f(x) = f(a) + \int_a^x dx' f'(x') \quad (\text{A.7.58})$$

gilt. Ist  $n \geq 2$ , können wir eine partielle Integration in (A.7.58) vornehmen. Dazu setzen wir

$$u'(x') = 1, \quad v(x') = f'(x'), \quad u(x') = x' - x,$$

so daß

$$f(x) = f(a) + (x-a)f'(a) - \int_a^x dx' (x'-x)f''(x') \quad (\text{A.7.59})$$

resultiert. Dies wiederholen wir noch weitere  $n - 1$ -mal. Dann entsteht

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k + \underbrace{(-1)^n \int_a^x \frac{(x'-x)^n}{n!} f^{(n+1)}(x')}_{R_n(x, a)}. \quad (\text{A.7.60})$$

Da voraussetzungsgemäß  $f^{(n+1)}$  im Intervall  $(a, x)$  bzw.  $(x, a)$  stetig ist und die Funktion  $x' \mapsto (x' - x)^n$  dort entweder stets negativ oder positiv ist, gibt es nach dem Mittelwertsatz der Integralrechnung wenigstens ein  $\xi \in (a, x)$  bzw.  $\xi \in (x, a)$ , so daß

$$R_n(x, a) = (-1)^n \frac{f^{(n+1)}(\xi)}{n!} \int_a^x dx' (x' - x)^n = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - a)^{n+1}. \quad (\text{A.7.61})$$

Dies ist die **Restgliedformel von Lagrange**. Man kann freilich den Mittelwertsatz der Integralrechnung auch direkt auf die Funktion  $x' \mapsto (x' - x)^n f^{(n+1)}(x')$  anwenden. Dann folgt die **Restgliedformel von Cauchy**:

$$R_n(x, a) = (-1)^n \frac{f^{(n+1)}(\xi')}{n!} (x - \xi')^n \int_a^x dx' = \frac{f^{(n+1)}(\xi')}{n!} (x - \xi')^n (x - a). \quad (\text{A.7.62})$$

Dabei ist  $\xi' \in (a, x)$  bzw.  $\xi' \in (x, a)$ . Aus der Lagrangeschen Restgliedformel (A.7.61) folgt in der Tat (A.7.55), womit die Taylorsche Formel bewiesen ist.

Falls nun  $f$  im Intervall  $(a, x)$  bzw.  $(x, a)$  sogar beliebig oft stetig differenzierbar ist und ein  $S > 0$  existiert, so daß  $|f^{(n+1)}(\xi)| < S$  für alle  $\xi \in (a, x)$  bzw.  $\xi \in (x, a)$  ist, folgt aus (A.7.61)<sup>3</sup>

$$|R_n(x, a)| < S \frac{|x - a|^{n+1}}{(n+1)!} \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (\text{A.7.63})$$

In diesem Fall besitzt  $f$  die **Taylor-Entwicklung**

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k. \quad (\text{A.7.64})$$

Als **Beispiel** betrachten wir die Taylor-Entwicklung von  $x \mapsto \exp(x)$  um  $a = 0$ . Da

$$f^{(k+1)}(x) = \exp(x), \quad (\text{A.7.65})$$

diese Funktion die Bedingungen für die Taylor-Entwicklung für jedes  $x \in \mathbb{R}$  erfüllt, gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (\text{A.7.66})$$

Setzen wir in (A.7.64)  $x = a + y$ , können wir symbolisch

$$f(a + y) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} y^k = \exp\left(y \frac{d}{da}\right) f(a) \quad (\text{A.7.67})$$

schreiben. Dabei ist die Exponentialfunktion des **Differentialoperators** durch die formale Reihe (A.7.66)

$$\exp\left(y \frac{d}{da}\right) = \sum_{k=0}^{\infty} \frac{y^k}{k!} \frac{d^k}{da^k} \quad (\text{A.7.68})$$

definiert. Wendet man diesen Operator auf  $f$  an, ergibt sich in der Tat die Taylor-Reihe von  $f$  in der in (A.7.67) angegebenen Form.

Für die Exponentialfunktion selbst gilt  $f^{(k)}(a) = \exp a$  für alle  $k \in \mathbb{N}$ . Daraus folgt mit (A.7.67)

$$\exp(a + y) = \exp a \sum_{k=0}^{\infty} \frac{y^k}{k!} = \exp a \exp y. \quad (\text{A.7.69})$$

<sup>3</sup>Um zu zeigen, daß für beliebige  $y > 0$   $\lim_{n \rightarrow \infty} y^n/n! = 0$  gilt, bemerken wir, daß für  $n > 2y$  und  $N \in \mathbb{N}$  die Abschätzung  $0 \leq y^{n+N}/(n+N)! \leq y^n/n!(1/2)^N$  gilt. Der letzte Ausdruck strebt aber für  $N \rightarrow \infty$  gegen 0.

Wichtig sind noch die Taylor-Entwicklungen der trigonometrischen Funktionen. Betrachten wir zuerst den Sinus:

$$f(x) = \sin x, \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad \dots \quad (\text{A.7.70})$$

Es gilt also

$$f^{(2k)}(0) = 0, \quad f^{(2k+1)}(x) = (-1)^k, \quad k \in \{0, 1, 2, \dots\}. \quad (\text{A.7.71})$$

Die Taylor-Reihe für den Sinus ist also

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}. \quad (\text{A.7.72})$$

Ebenso finden wir für den Cosinus

$$f(x) = \cos x, \quad f'(x) = -\sin x, \quad f''(x) = -\cos x, \quad \dots, \quad (\text{A.7.73})$$

und damit die Taylorreihe

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}. \quad (\text{A.7.74})$$

Es ist klar, daß wegen  $-1 \leq \sin x \leq 1$  und  $-1 \leq \cos x \leq 1$  aufgrund der Restgliedformel die Taylorreihen (A.7.72) und (A.7.74) sicher überall in  $\mathbb{R}$  und tatsächlich gegen  $\sin x$  bzw.  $\cos x$  konvergieren.

Die Additionstheoreme der trigonometrischen Funktionen können ebenfalls leicht mit Hilfe von (A.7.67) bewiesen werden. Für den  $\cos$  bemerken wir, daß

$$f^{(k)}(a) = \begin{cases} (-1)^{j+1} \sin a & \text{falls } k = 2j + 1, \\ (-1)^j \cos a & \text{falls } k = 2j, \end{cases} \quad j \in \{0, 1, 2, \dots\}. \quad (\text{A.7.75})$$

Daraus folgt mit (A.7.67)

$$\cos(a+y) = \cos a \sum_{j=0}^{\infty} (-1)^j \frac{y^{2j}}{(2j)!} + \sin a \sum_{j=0}^{\infty} (-1)^{j+1} \frac{y^{2j+1}}{(2j+1)!} = \cos a \cos y - \sin a \sin y, \quad (\text{A.7.76})$$

wobei wir (A.7.72) und (A.7.74) benutzt haben. Außerdem haben wir stillschweigend die unendliche Reihe umgeordnet, indem wir gerade und ungerade Potenzen von  $y$  zusammengefaßt haben. Dies ist erlaubt, weil Potenzreihen im Inneren ihres Definitionsbereichs absolut konvergent sind. Auf einen formalen Beweis verzichten wir hier.

Wir bemerken, daß wir offenbar die Taylor-Reihen von beliebig oft stetig differenzierbaren Funktionen gliedweise differenzieren und integrieren dürfen, d.h. die Summation der unendlichen Reihe vertauscht in diesem Fall mit der Differentiation und Integration.

Betrachten wir als ein weiteres Beispiel die geometrische Reihe

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad \text{für } -1 < x < 1. \quad (\text{A.7.77})$$

Dies läßt sich unmittelbar aus (A.2.39) herleiten, denn demnach gilt

$$\sum_{k=0}^{\infty} x^k = 1 + \sum_{k=1}^{\infty} x^k = 1 + \frac{x}{1-x} = \frac{1}{1-x}. \quad (\text{A.7.78})$$

Natürlich können wir (A.7.77) auch durch Taylorentwicklung der Funktion  $f : (-1, 1)$ ,  $f(x) = 1/(1-x)$  herleiten (*Übung!*). Das Argument der Konvergenz aufgrund der Taylorschen Restgliedformel gilt hier offensichtlich nur für abgeschlossene Intervalle  $[a, b]$  mit  $-1 < a < b < 1$ , denn die Funktion und ihre Ableitungen sind offenbar bei  $x = 1$  singulär und unbeschränkt für  $x \rightarrow 1$ .

Nun gilt (*nachrechnen!*)

$$\int_0^x dx' \frac{1}{1-x'} = -\ln(1-x) \quad \text{für } x < 1. \quad (\text{A.7.79})$$

Da die Potenzreihe (A.7.77) nur für  $|x| < 1$  existiert, dürfen wir diese Integrationsformel nur für solche  $|x|$  anwenden. Dann erhalten wir

$$\ln(1-x) = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = -\sum_{k=1}^{\infty} \frac{x^k}{k} \quad \text{für } |x| < 1. \quad (\text{A.7.80})$$

Dabei haben wir im letzten Schritt einfach  $k+1$  durch  $k$  ersetzt und die Summationsgrenzen entsprechend angepaßt. Ersetzen wir in (A.7.80)  $x$  durch  $-x$  erhalten wir

$$\ln(1+x) = \ln[1-(-x)] = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \quad \text{für } |x| < 1. \quad (\text{A.7.81})$$

Dies können wir auch wieder durch Taylorentwicklung der Funktion  $\ln(1+x)$  beweisen (*Übung*).

Wir bemerken nun umgekehrt, daß wir auch Funktionen durch Potenzreihen definieren können. Sei zunächst durch eine formale Potenzreihe

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \quad (\text{A.7.82})$$

eine Funktion  $f$  definiert. Wir studieren nun das Konvergenzverhalten solcher Potenzreihen genauer. Wir interessieren uns für die absolute Konvergenz.

**Satz:** Sei die Reihe (A.7.82) für ein  $x = r > 0$  konvergent. Dann ist sie auf  $[-r, r]$  gleichmäßig und absolut konvergent.

**Beweis:** Wir definieren  $f_k : [-r, r] \rightarrow \mathbb{R}$ ,  $f_k(x) = c_k x^k$ . Nun gilt

$$\|f_k\| = \sup_{x \in [-r, r]} |c_k x^k| \leq |c_k| r^k. \quad (\text{A.7.83})$$

Nach dem Majorantenkriterium ist demnach die Reihe  $\sum_k \|f_k\|$  konvergent und nach dem Weierstraßschen Konvergenzkriterium folglich (A.7.82) auf  $[-r, r]$  absolut und gleichmäßig konvergent, und das war zu zeigen.

Definieren wir nun die Menge

$$K = \{x \in \mathbb{R} \mid (\text{A.7.82}) \text{ ist absolut konvergent}\}. \quad (\text{A.7.84})$$

Ist dann  $K$  nach oben beschränkt, existiert nach dem Satz vom Supremum

$$R = \sup K. \quad (\text{A.7.85})$$

Nach dem gerade bewiesenen Satz ist dann (A.7.82) im *offenen* Intervall  $(-R, R)$  absolut konvergent und in jedem ganz in diesem Intervall gelegenen *abgeschlossenen* Intervall absolut und gleichmäßig konvergent. Man nennt  $R$  daher den **Konvergenzradius** der Potenzreihe.

Falls  $K$  nicht nach oben beschränkt ist, ist die Reihe für alle  $x \in \mathbb{R}$  absolut konvergent, und in jedem endlichen Intervall gleichmäßig und absolut konvergent.

Zur praktischen Berechnung von  $R$  kann man sehr oft das Quotientenkriterium (A.2.44) heranziehen. Demnach ist die Reihe konvergent falls

$$\lim_{k \rightarrow \infty} \left| \frac{c_{k+1} x^{k+1}}{c_k x^k} \right| = |x| \lim_{k \rightarrow \infty} \frac{c_{k+1}}{c_k} < 1 \quad (\text{A.7.86})$$

ist. Ist der Limes auf der rechten Seite 0, bedeutet dies, daß die Potenzreihe für alle  $x \in \mathbb{R}$  konvergiert. Andernfalls ist die Reihe für

$$|x| < R \quad \text{mit} \quad R = \lim_{k \rightarrow \infty} \left| \frac{c_k}{c_{k+1}} \right| \quad (\text{A.7.87})$$

konvergent. Falls der Grenzwert nicht existiert, muß die Konvergenz der Potenzreihe (A.7.81) gesondert untersucht werden. Da dies in der Praxis selten vorkommt, gehen wir nicht genauer darauf ein<sup>4</sup>.

**Satz:** Sei der Konvergenzradius der Potenzreihe (A.7.82)  $R > 0$ . Dann ist auch der Konvergenzradius der Potenzreihen

$$\tilde{f}(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}, \quad (\text{A.7.88})$$

$$\tilde{F}(x) = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1} \quad (\text{A.7.89})$$

wieder  $R$ . Außerdem ist  $f$  differenzierbar und integrierbar, und es gilt

$$f'(x) = \tilde{f}(x), \quad (\text{A.7.90})$$

$$F(x) = \int_0^x dx' f(x') = \tilde{F}(x). \quad (\text{A.7.91})$$

Kurz gesagt dürfen Potenzreihen im Inneren ihres Konvergenzbereichs gliedweise differenziert und integriert werden.

**Beweis:** Sei  $r < R$ . Nach den obigen Betrachtungen zur gleichmäßigen und absoluten Konvergenz von Potenzreihen ist (A.7.82) auf dem abgeschlossenen Intervall  $[-r, r]$  gleichmäßig konvergent. Da die Reihenglieder  $f_k(x) = c_k x^k$  allesamt stetig sind, darf nach dem entsprechenden Satz über gleichmäßig konvergente Reihen Integration und Summation vertauscht werden. Damit ist bereits die Behauptung bzgl. der Integration bewiesen.

Hinsichtlich der Ableitung müssen wir nur zeigen, daß der Konvergenzradius der Reihe (A.7.88) der gleiche wie der ursprünglichen Reihe (A.7.82) ist, denn dann ist (A.7.88) auf jedem kompakten Intervall innerhalb des Konvergenzbereichs gleichmäßig konvergent, und nach dem entsprechenden Satz zur Ableitung gleichmäßig konvergenter Funktionenfolgen gilt (A.7.90). Seien nun  $r$  und  $r'$  Zahlen, die  $0 < r < r' < R$ . Dann ist  $\sum_k c_k r'^k$  absolut konvergent. Dann existiert  $\max_k |c_k r'^k| = M > 0$ . Sei nun  $q = r/r' < 1$ . Für alle  $x \in [-r, r]$  gilt dann

$$|k c_k x^{k-1}| \leq k |c_k| r^{k-1} \leq k |c_k| r'^{-1} k q^{k-1} \leq M k q^{k-1}. \quad (\text{A.7.92})$$

Die Reihe  $\sum_k k q^{k-1}$  ist nach dem Quotientenkriterium konvergent, denn mit  $a_k k q^{k-1}$  gilt

$$\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = \lim_{k \rightarrow \infty} \frac{k+1}{k} q = q < 1. \quad (\text{A.7.93})$$

<sup>4</sup>Haben wir es mit Potenzreihen zu tun, bei denen nur gerade oder ungerade Potenzen vorkommen, wie bei den Potenzreihen für  $\sin$  und  $\cos$  (A.7.72) und (A.7.74), können wir das Quotientenkriterium auf  $c_j/c_{j+2}$  anwenden. Dann ist offenbar  $\lim_{j \rightarrow \infty} |c_j/c_{j+1}| = R^2$ , falls der Limes existiert und demnach die Reihe für  $x^2 < R^2$ , d.h. wieder für  $|x| < R$  konvergent.

Mit  $g_k(x) = kc_k x^{k-1}$  ist also

$$\|g_k\| \leq M k q^{k-1} \quad (\text{A.7.94})$$

und folglich (A.7.88) tatsächlich gleichmäßig auf  $[r, r]$  und absolut konvergent, und die Behauptung damit bewiesen.

## A.8 Die strikte Definition der trigonometrischen Funktionen

Bisher haben wir die trigonometrischen Funktionen Sinus, Cosinus und Tangens nur geometrisch definiert, und bei der Herleitung von Rechenregeln wie Additionstheoremen und Ableitungen geometrische Betrachtungen verwendet, die nicht streng bewiesen wurden. Wir geben daher eine Definition der Funktionen  $\cos$  und  $\sin$  durch ihre Potenzreihen und leiten alle Eigenschaften dieser Funktionen mit Hilfe der oben bewiesenen Sätze für Potenzreihen her.

Wir definieren also die Funktionen durch die Potenzreihen (A.7.72) und (A.7.74), die sich oben unter Verwendung der heuristisch begründeten Ableitungsregeln ergeben haben:

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}, \quad (\text{A.8.1})$$

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}. \quad (\text{A.8.2})$$

Den Bereich absoluter Konvergenz können wir mit dem Quotientenkriterium ermitteln. Sei für die Cosinusreihe  $a_k = (-1)^k x^{2k} / (2k)!$  dann gilt

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k \rightarrow \infty} \frac{|x|^{2k+2} (2k)!}{|x|^{2k} (2k+2)!} = x^2 \lim_{k \rightarrow \infty} \frac{1}{(2k+1)(2k+2)} = 0. \quad (\text{A.8.3})$$

Das bedeutet, daß aufgrund der im vorigen Abschnitt bewiesenen Sätze die Potenzreihe für alle  $x \in \mathbb{R}$  absolut und somit in jedem beschränkten Intervall gleichmäßig konvergiert und folglich gliedweise differenziert und integriert werden darf. Auch die Sinusreihe ist überall konvergent, wie man auf analoge Weise wie eben bei der Cosinusreihe nachrechnet.

Daraus ergeben sich sogleich die Ableitungsregeln, die wir vorher auf naive Weise berechnet haben:

$$\cos' x = -\sin x, \quad \sin' x = \cos x. \quad (\text{A.8.4})$$

Als nächstes wollen wir einige Eigenschaften dieser Funktionen beweisen. Als erstes definieren wir  $\pi/2$  als Nullstelle des Cosinus im Intervall  $[0, 2]$ . Dazu zeigen wir, daß der Cosinus dort genau eine Nullstelle besitzt. Aus der Potenzreihe folgt sofort, daß  $\cos 0 = 1$  ist. Wir zeigen nun, daß  $\cos 2 < 0$  ist. Es gilt

$$a_k = \frac{2^{2k}}{(2k)!} - \frac{2^{2k+2}}{(2k+2)!} = \frac{2^{2k}}{(2k)!} \left( 1 - \frac{4}{(2k+1)(2k+2)} \right). \quad (\text{A.8.5})$$

Die Klammer ist monoton wachsend als Funktion von  $k$ , und für  $k = 1$  wird sie  $2/3$ , d.h. es gilt  $a_k > 0$  für  $k \geq 1$ . Nun gilt

$$\cos 2 = 1 - a_1 - a_3 - \dots = 1/3 - \sum_{k=0}^{\infty} a_{2k+1} < 0. \quad (\text{A.8.6})$$

Da der Cosinus aufgrund der Definition als Potenzreihe stetig ist, gibt es aufgrund des Zwischenwertsatzes wegen  $\cos 0 = 1 > 0$  und  $\cos 2 < 0$  wenigstens ein  $\xi \in (0, 2)$  mit  $\cos \xi = 0$ .

Um zu zeigen, daß es nur genau eine Nullstelle in diesem Intervall gibt, zeigen wir, daß  $\cos x$  dort streng monoton fallend ist. Dazu müssen wir nur zeigen, daß  $\cos' x = -\sin x$  überall in  $(0, 2)$  negativ ist. In der Tat ist für alle  $x \in (0, 2)$  und alle  $k \in \mathbb{N}_0$

$$b_k = \frac{x^{2k+1}}{(2k+1)!} - x^{2k+3}(2k+3)! = \frac{x^{2k+1}}{(2k+1)!} \left( 1 - \frac{x^2}{(2k+2)(2k+3)} \right) > 0. \quad (\text{A.8.7})$$

Damit ist für  $x \in (0, 2)$

$$\sin x = \sum_{k=0}^{\infty} b_{2k} > 0 \Rightarrow \cos' x = -\sin x < 0 \quad (\text{A.8.8})$$

und damit  $\cos x$  in  $(0, 2)$  strikt monoton fallend. Damit gibt es also in diesem Intervall genau eine Nullstelle, die wir mit  $\pi/2$  bezeichnen.

Die Additionstheoreme lassen sich nun über den Taylorsche Lehrrsatz herleiten. Z.B. ist

$$\cos(x_1 + x_2) = \sum_{k=0}^{\infty} \frac{x_2^k}{k!} \frac{d^k}{dx_1^k} \cos x_1. \quad (\text{A.8.9})$$

Nun ist

$$\frac{d^{2j+1}}{dx_1^{2j+1}} \cos x_1 = (-1)^{j+1} \sin x_1 \quad \text{und} \quad \frac{d^{2j}}{dx_1^{2j}} \cos x_1 = (-1)^j \cos x_1. \quad (\text{A.8.10})$$

Die Potenzreihe (A.8.9) ist für alle  $x_2$  absolut konvergent, wie man wie oben bei der Cosinus-Reihe nachrechnet. Folglich kann man diese Reihe nach dem Umordnungssatz beliebig umordnen, ohne daß sich an ihrem Wert etwas ändert. Wir können also in (A.8.9) alle Glieder mit geraden und mit ungeraden  $k$  zusammenfassen. Mit (A.8.10) ergibt dies

$$\cos(x_1 + x_2) = \cos x_1 \sum_{j=0}^{\infty} \frac{(-1)^j x_2^{2j}}{(2j)!} - \sin x_1 \sum_{j=0}^{\infty} \frac{(-1)^j x_2^{2j+1}}{(2j+1)!} = \cos x_1 \cos x_2 - \sin x_1 \sin x_2. \quad (\text{A.8.11})$$

Dabei haben wir uns im letzten Schritt der Definitionen (A.8.1) und (A.8.2) des Cosinus und Sinus bedient. Aus den Ableitungsregeln folgt daraus sofort

$$\sin(x_1 + x_2) = -\frac{d}{dx_1} \cos(x_1 + x_2) = \sin x_1 \cos x_2 + \cos x_1 \sin x_2. \quad (\text{A.8.12})$$

Als nächstes beweisen wir die Formel

$$\cos^2 x + \sin^2 x = 1. \quad (\text{A.8.13})$$

Dazu bilden wir die Ableitung mit Hilfe der Produktregel:

$$\frac{d}{dx} (\cos^2 x + \sin^2 x) = -2 \sin x \cos x + 2 \sin x \cos x = 0. \quad (\text{A.8.14})$$

Damit ist klar, daß  $\cos^2 x + \sin^2 x = \text{const}$  ist, und für  $x = 0$  folgt, daß die Konstante 1 sein muß und damit (A.8.13).

Wegen  $\cos(\pi/2) = 0$  ist also  $\sin^2(\pi/2) = 1$ , und da für  $x \in (0, 2)$  stets  $\sin x > 0$  ist, muß demnach  $\sin \pi/2 = 1$  gelten. Daraus folgt dann mit den Additionstheoremen (A.8.11)

$$\cos(x + \pi/2) = \cos x \cos(\pi/2) - \sin x \sin(\pi/2) = -\sin x, \quad (\text{A.8.15})$$

$$\sin(x + \pi/2) = \sin x \cos(\pi/2) + \cos x \sin(\pi/2) = \cos x. \quad (\text{A.8.16})$$

### A.8 · Die strikte Definition der trigonometrischen Funktionen

Da  $\sin x \geq 0$  für  $x \in [0, \pi/2]$  ist  $\cos x \leq 0$  für  $x \in [\pi/2, \pi]$ , und wegen  $\cos x \geq 0$  für  $x \in [0, \pi/2]$  ist  $\sin x \geq 0$  für  $x \in [\pi/2, \pi]$ . Wegen der Ableitungsregeln für Sinus und Cosinus folgt daraus weiter, daß in  $[\pi/2, \pi]$  der Sinus und Cosinus beide monoton fallend sind. Setzt man in (A.8.15) und (A.8.16) jeweils  $x = \pi/2$  folgt

$$\cos \pi = -\sin(\pi/2) = -1, \quad \sin \pi = \cos(\pi/2) = 0. \quad (\text{A.8.17})$$

Damit haben wir wieder mit den Additionstheoremen (A.8.11) und (A.8.12)

$$\cos(x + \pi) = -\cos x, \quad \sin(x + \pi) = -\sin x. \quad (\text{A.8.18})$$

Folglich ist in  $[\pi, 3\pi/2]$  stets  $\cos x \leq 0$  und  $\sin x \leq 0$  und der Cosinus monoton wachsend und der Sinus monoton fallend. Verwendet man diese Information in (A.8.15) und (A.8.16) folgt, daß im Intervall  $[3\pi/2, 2\pi]$  stets  $\cos x \geq 0$  und  $\sin x \leq 0$  und Cosinus und Sinus monoton wachsend sind.

Für  $x = \pi/2$  bzw.  $x = \pi$  folgt aus (A.8.18) mit (A.8.17)

$$\cos(3\pi/2) = 0, \quad \sin(3\pi/2) = -1, \quad \cos(2\pi) = 1, \quad \sin(2\pi) = 0. \quad (\text{A.8.19})$$

Daraus folgt, daß Cosinus und Sinus **periodische Funktionen** mit der Periode  $2\pi$  sind, d.h.

$$\cos(x + 2\pi) = \cos x, \quad \sin(x + 2\pi) = \sin x. \quad (\text{A.8.20})$$

Außerdem ist aufgrund der eben hergeleiteten Monotonieeigenschaften dieser Funktionen  $2\pi$  die kleinste Periode beider Funktionen. Damit haben wir die wesentlichsten Eigenschaften von Cosinus und Sinus streng hergeleitet.

Daß diese Funktionen die bisher zu ihrer Definition verwendete geometrische Bedeutung besitzen, wird in Kapitel 2 bei der Behandlung der Parameterdarstellung von Kurven in der Ebene und im Raum gezeigt.



# Literaturverzeichnis

- [BK88] D. Bourne, P. Kendall, *Vektoranalysis*, 2. Aufl., B. G. Teubner, Stuttgart (1988).
- [Bro03] R. Bronson, *Differential Equations*, Schaum's Easy Outlines, McGraw-Hill, New York, Chicago, San Francisco (2003).
- [CH10] W. Cassing, H. van Hees, *Mathematische Methoden für Physiker*, Universität Gießen (2010).  
<http://fias.uni-frankfurt.de/~hees/publ/maphy.pdf>
- [Col90] L. Collatz, *Differentialgleichungen*, Teubner (1990).
- [For13] O. Forster, *Analysis 1*, 11. Aufl., Springer Spektrum, Wiesbaden (2013).
- [Gre03] W. Greiner, *Klassische Mechanik 1*, 7. Aufl., Verlag Harri Deutsch, Frankfurt/Main (2003).
- [Gro05] S. Großmann, *Mathematischer Einführungskurs für die Physik*, Vieweg+Teubner Verlag (2005).  
<http://dx.doi.org/10.1007/978-3-8348-8347-6>
- [Hee05] H. van Hees, *Klassische Vektoranalysis*, FAQ der Newsgroup de.sci.physik (2005).  
<http://theory.gsi.de/~vanhees/faq/vekanal/vekanal.html>
- [Joo89] G. Joos, *Lehrbuch der theoretischen Physik*, 15. Aufl., Aulaverlag, Wiesbaden (1989).
- [Nol13] W. Nolting, *Grundkurs Theoretische Physik 1: Klassische Mechanik*, 10. Aufl., Springer Spektrum, Berlin, Heidelberg (2013).  
<http://dx.doi.org/10.1007/978-3-642-29937-7>
- [Ris09] D. Rischke, *Theoretische Physik 1: Klassische Mechanik*, Johann-Wolfgang-Goethe-Universität, Frankfurt/Main (2009).  
<http://www.th.physik.uni-frankfurt.de/~drischke/Skript.pdf>
- [Sau73] F. Sauter, *Becker/Sauter Theorie der Elektrizität 1*, 21. Aufl., B. G. Teubner, Stuttgart (1973).  
<http://dx.doi.org/10.1007/978-3-322-96789-3>
- [Sch07] F. Scheck, *Theoretische Physik 1: Mechanik*, 8. Aufl., Springer-Verlag, Berlin, Heidelberg, New York (2007).
- [SH99] M. R. Spiegel, C. Hipp, *Einführung in die höhere Mathematik*, Schaum's Outline - Überblicke/Aufgaben, McGraw-Hill Book Company (1999).
- [Som92] A. Sommerfeld, *Vorlesungen über Theoretische Physik II, Mechanik der deformierbaren Medien*, Verlag Harri Deutsch, Frankfurt/M. (1992).
- [Wei80] J. Weidmann, *Linear Operators in Hilbert Space*, Springer Verlag, New York, Berlin, Heidelberg (1980).