

Teil II

Mathematische Grundlagen der Theoretischen Physik

Kapitel 9

Die Integralsätze von Gauß und Stokes

9.1 Flächenintegrale und Vektorfluss

In Kapitel 8.4 haben wir gelernt, wie wir Skalarfelder in beliebigen Dimensionen integrieren können. Analog zum Wegintegral (siehe Kapitel 8.1), bei dem Vektorfelder längs eines Weges integriert werden, können wir auch Vektorfelder über (gerichtete) Flächen integrieren. Dies wird z.B. benötigt, um den Fluss durch eine Fläche zu bestimmen. Dies führt auf das Konzept des *Flussintegrals*, das man als zweidimensionale Verallgemeinerung des Wegintegrals auffassen kann. Die Bezeichnung “Fluss” ist durch die Hydrodynamik motiviert. Man kann sich tatsächlich eine durch die Oberfläche strömende Flüssigkeit vorstellen.

Wir betrachten eine Fläche im dreidimensionalen Raum¹, die beliebig geformt sein kann. Ihr Flächeninhalt (die Oberfläche) sei F . Wir stellen uns vor, dass die Fläche aus infinitesimalen Flächenelementen df besteht (siehe Abb. 9.1.1), so dass

$$F = \int df.$$

Definition 9.1.1 (Normale, Flächenvektor).

Auf jedem infinitesimalen Flächenelement df errichten wir einen senkrecht stehenden Einheitsvektor \underline{n} . \underline{n} heißt auch **Normale**.

$\underline{df} = \underline{n}df$ heißt (infinitesimaler) **Flächenvektor**.

Bemerkung: Bisher sind für die Normale noch zwei Richtungen möglich. Für geschlossene Flächen (z.B. Kugeloberflächen) zeigt die Normale jedoch per Konvention immer nach außen.

Definition 9.1.2 (Vektorfluss).

Seien nun ein Vektorfeld $\underline{A}(\underline{r})$ und eine Fläche F gegeben. Dann bezeichnet man

$$\int_F \underline{A} \cdot \underline{df} = \int_F \underline{A} \cdot \underline{n}df$$

¹Die meisten Überlegungen lassen sich auf Hyperflächen beliebiger Dimension übertragen.

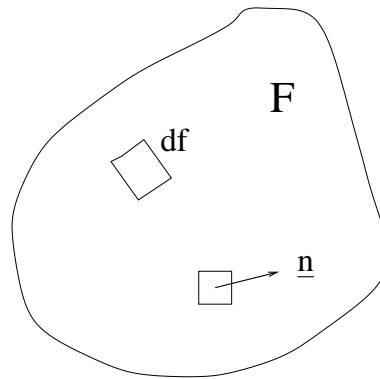


Abbildung 9.1.1: Eine Fläche F bestehend aus (infinitesimalen) Flächenelementen df mit Normalenvektoren \underline{n} .

als den **Fluss von \underline{A} durch F** . Dabei wird das Integral über die Oberfläche genommen.

Der Vektorfluss ist eine skalare Größe, die von \underline{A} und der Art und Lage der Fläche bestimmt wird (siehe Abb. 9.1.2).

Die Darstellung auf der rechten Seite zeigt, dass man letztlich ein gewöhnliches zweidimensionales Integral über ein Skalarfeld zu berechnen hat. Man beachte aber dabei, dass der Normalenvektor \underline{n} i.a. keine Konstante ist, sondern sich von Position zu Position auf der Fläche ändert!

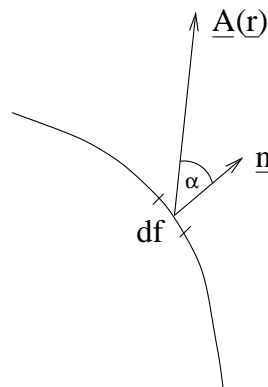


Abbildung 9.1.2: Fluss des Vektorfeldes \underline{A} durch die Fläche F .

Sei α der Winkel zwischen \underline{A} und df (bzw. \underline{n}) (siehe Abb. 9.1.2). Dann ist $\underline{A} \cdot df = |\underline{A}| \cos \alpha df$. Ist \underline{A} an der betrachteten Stelle senkrecht zu df , d.h. \underline{A} liegt in der Fläche, so ist $\underline{A} \cdot df = 0$. Zum Vektorfluss trägt also nur der durch die (gekrümmte) Fläche jeweils senkrecht hindurchtretende Anteil von \underline{A} bei.

Beispiel 9.1.1.

Als Beispiel betrachten wir eine Kugelschale K_R vom Radius R um den Ursprung. Der Normalenvektor zeigt dann immer radial nach außen, d.h. wir haben $\underline{n} = \hat{r} = \underline{r}/r$. Das gerichtete Flächenelement ist daher (in Kugelkoordinaten) gegeben durch

$$\underline{df} = R^2 \sin \theta d\theta d\varphi \hat{r} = R^2 d\Omega \hat{r},$$

wobei wir das **Raumwinkelement**

$$d\Omega = \sin \theta d\theta d\phi$$

eingeführt haben. Die Form des Flächenelements macht man sich leicht aus der allgemeinen Form des Volumenelements in Kugelkoordinaten klar, da das Flächenelement keine "Ausdehnung" in radialer Richtung hat.

Wir wollen nun ein Radialfeld der Form

$$\underline{A}(\underline{r}) = A(r)\hat{r}$$

über die Kugelschale integrieren. Felder solcher Form nennt man **Radialfelder**. Sie zeigen immer in radiale Richtung und ihr Betrag hängt nur vom Abstand r ab, nicht von der Richtung. Dieser Fall tritt in der Physik sehr häufig auf, z.B. bei Gravitations- oder elektrischen Feldern. Für das Flussintegral erhalten wir nun:

$$\begin{aligned} \int_{K_R} \underline{A}(\underline{r}) \cdot \underline{df} &= \int_0^{2\pi} d\varphi \int_0^\pi d\theta R^2 \sin \theta \underline{A}(\underline{r}) \cdot \hat{r} \\ &= A(R)R^2 \int_0^{2\pi} d\varphi \int_0^\pi d\theta R^2 \sin \theta \\ &= 4\pi R^2 A(R). \end{aligned}$$

Im Integral der ersten Zeilen haben wir die Parametrisierung der Fläche in Kugelkoordinaten benutzt, in der zweiten Zeile die spezielle Form des Vektorfeldes, insbesondere die Tatsache, dass sein Betrag auf der Kugelschale konstant ist. Letzteres macht auch das Endergebnis klar, denn offensichtlich ist der Wert des Flussintegrals gegeben durch das Produkt der Kugelfläche und dem (konstanten) Wert des Feldes auf dieser.

9.2 Der Greensche Satz in der Ebene

Wir betrachten eine Fläche F , deren Rand durch die Randkurve $\mathcal{C} = \partial F$ gegeben ist. Wir nehmen o.B.d.A. an, dass es für jeden Wert von x genau eine obere und eine untere Grenze gibt². Dies definiert die untere bzw. obere Grenzkurve $f_1(x)$ bzw. $f_2(x)$ (siehe Abb. 9.2.1) und für alle $x \in [a, b]$ gilt $f_1(x) \leq y \leq f_2(x)$, falls der Punkt (x, y) in der Fläche F liegt.

²Ansonsten muß man die kompliziertere Fläche aus solchen einfachen zusammensetzen.

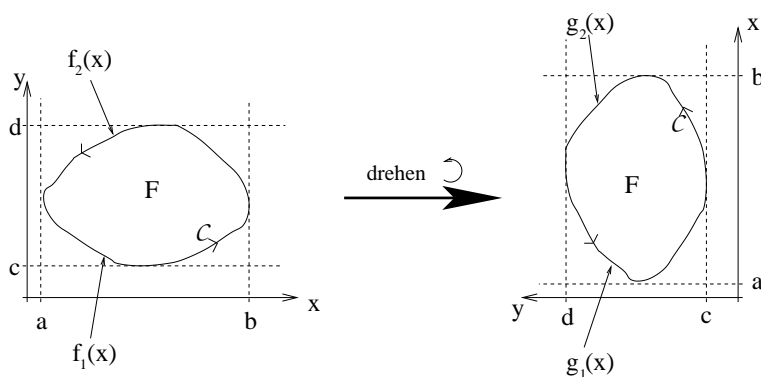


Abbildung 9.2.1: Zur Herleitung des Greenschen Satzes.

Es sei nun $V(x, y)$ eine Funktion, die auf F stetig partiell differenzierbar sei. Dann gilt:

$$\begin{aligned}
 \int_F dx dy \frac{\partial V(x, y)}{\partial y} &= \int_a^b dx \left[\int_{f_1(x)}^{f_2(x)} dy \frac{\partial V(x, y)}{\partial y} \right] \\
 &= \int_a^b dx [V(x, f_2(x)) - V(x, f_1(x))] \\
 &= - \int_b^a dx V(x, f_2(x)) - \int_a^b dx V(x, f_1(x)) \\
 &= - \oint_C dx V(x, y),
 \end{aligned}$$

d.h. das Flächenintegral kann in ein Wegintegral umgeschrieben werden!

Wir drehen nun die Fläche, wie in Fig. 9.2.1 dargestellt und gehen in der gedrehten Version analog vor:

$$\int_F dx dy \frac{\partial V(x, y)}{\partial x} = \oint_C dy V(x, y).$$

Wir fassen nun diese beiden Teilergebnisse zusammen:

Satz 9.2.1 (Greenscher Satz).

Der **Satz von Green** für Funktionen $V_1(x, y)$, $V_2(x, y)$ die auf der zweidimensionalen Fläche F stetig partiell differenzierbar sind, lautet

$$\boxed{\int_F dx dy \left[\frac{\partial V_2(x, y)}{\partial x} - \frac{\partial V_1(x, y)}{\partial y} \right] = \oint_C [V_1(x, y) dx + V_2(x, y) dy]}$$

wobei $C = \partial F$ die Randkurve von F ist.

Wir wollen zwei Anwendungen dieses Satzes diskutieren.

1. Sei

$$V_1(x, y) = \frac{\partial \phi}{\partial x}, \quad V_2(x, y) = \frac{\partial \phi}{\partial y}.$$

Dann verschwindet das Flächenintegral im Greenschen Satz, da

$$\frac{\partial V_2}{\partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial V_1}{\partial y}$$

ist. Außerdem gilt mit $\underline{r} = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$V_1 dx + V_2 dy = d\phi = \text{grad } \phi \cdot d\underline{r}.$$

Somit lautet die Aussage des Greenschen Satzes in diesem Fall

$$\oint_C d\phi = 0.$$

Dies wissen wir natürlich schon, da $\underline{V} := \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ per Definition ein konservatives Feld mit Potential $-\phi$ ist.

2. Für $\underline{V} := \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ gilt

$$\oint_C \underline{V} \cdot d\underline{r} = \int_F \text{rot } \underline{V} \cdot d\underline{f},$$

wobei $d\underline{f} = \underline{n} dx dy$ senkrecht zur $x - y$ -Ebene steht und auf der rechten Seite \underline{V} als dreidimensionaler Vektor mit 3. Komponente $V_3 = 0$ zu interpretieren ist.

3. Wir können mit Hilfe des Greenschen Satzes Flächen berechnen. Dazu betrachten wir die Funktionen $V_1(x, y) = -y$ und $V_2(x, y) = x$. Dann gilt:

$$\oint_C [-y dx + x dy] = \int \int_F dx dy \left[\frac{\partial x}{\partial x} - \left(-\frac{\partial y}{\partial y} \right) \right] = 2 \int \int_F dx dy = 2F.$$

Als Beispiel bestimmen wir die Fläche einer Ellipse, die durch

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{9.2.1}$$

definiert ist³. Eine Parameterdarstellung ist durch $\underline{r}(t) = (a \cos t, b \sin t)$ gegeben ($t \in [0, 2\pi]$). Dann gilt:

$$F_{\text{Ellipse}} = \frac{1}{2} \oint_{\text{Ellipse}} [-y dx + x dy] = \int_0^{2\pi} [-b \sin t (-a \sin t) + a \cos t b \cos t] dt = ab\pi.$$

³D.h. alle (x, y) , die diese Gleichung erfüllen, liegen auf der Ellipse.

9.3 Integraldarstellung der Rotation

Die Rotation eines Vektorfeldes $\underline{A}(\underline{r})$ läßt sich als Wegintegral ausdrücken. Diese Darstellung wird in den folgenden Abschnitten sehr nützlich sein.

Wir betrachten einen geschlossenen Weg \mathcal{C} um den Punkt \underline{r} , an dem wir die Rotation bestimmen wollen. Dieser Weg schließt die Fläche ΔF ein. Ist diese hinreichend klein, so können wir sie als eben ansehen und durch den Normalenvektor \underline{n} charakterisieren, der senkrecht auf ihr steht. Die Komponente der Rotation parallel zu \underline{n} läßt sich dann folgendermaßen bestimmen:

$$\underline{n} \cdot \text{rot } \underline{A}(\underline{r}) = \lim_{\Delta F \rightarrow 0} \frac{1}{\Delta F} \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r}.$$

Um die Rotation vollständig zu bestimmen, kann man z.B. Wege nehmen, die in der $x-y$ -, $x-z$ - und $y-z$ -Ebene liegen, wobei dann \underline{n} der Einheitsvektor in z -, y - und x -Richtung ist.

Zum Beweis dieser Integraldarstellung betrachten wir o.B.d.A. ein infinitesimales rechteckiges Flächenelement (Abb. 9.3.1) in der $y-z$ -Ebene. Das Wegintegral längs des Weges $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ läßt sich dann folgendermaßen berechnen:

$$\begin{aligned} \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} &= A_y(x, y, z - \frac{1}{2}dz)dy + A_z(x, y + \frac{1}{2}dy, z)dz \\ &\quad - A_y(x, y, z + \frac{1}{2}dz)dy - A_z(x, y - \frac{1}{2}dy, z)dz \\ &= \frac{\partial A_z(\underline{r})}{\partial y} dydz - \frac{\partial A_y(\underline{r})}{\partial z} dydz \\ &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \Delta F = (\text{rot } \underline{A})_x \Delta F. \end{aligned}$$

Dabei haben wir angenommen, dass die Kantenlängen dy und dz infinitesimal sind und deshalb der Wert des Vektorfeldes auf einer Kante konstant ist. Hierbei haben wir den Funktionswert in der Kantenmitte gewählt. Für die Kante $1 \rightarrow 2$ sind die x - und z -Koordinaten konstant (nämlich x und $z - \frac{1}{2}dz$) und die y -Koordinate variiert von $y - \frac{1}{2}dy$ nach $y + \frac{1}{2}dy$. Die Kantenmitte ist also bei $(x, y, z - \frac{1}{2}dz)$ und deshalb ist $A_y(x, y, z - \frac{1}{2}dz)$ zu wählen⁴. Beim Übergang zur dritten Zeile sind wir analog vorgegangen wie beim Beweis des Satzes von Gauß. Hier wurden die auftretenden Beiträge Taylor-entwickelt, wobei sich die konstanten Terme wegheben.

Somit haben wir für $\underline{n} = \hat{x}$ gezeigt

$$\lim_{\Delta F \rightarrow 0} \frac{1}{\Delta F} \oint_{\Delta \mathcal{C}} \underline{A} \cdot d\underline{r} = (\text{rot } \underline{A})_x = \underline{n} \cdot \text{rot } \underline{A}.$$

Mit einem analogen Argument läßt sich diese Aussage auch für die y - und z -Komponenten beweisen, womit (***) gezeigt wäre, und damit der Stokessche Integralsatz.

⁴Es trägt nur die y -Komponente des Feldes bei, da sich auf dem Weg nur die y -Koordinate ändert und deshalb $\underline{A} \cdot d\underline{r} = A_y dy$ ist.

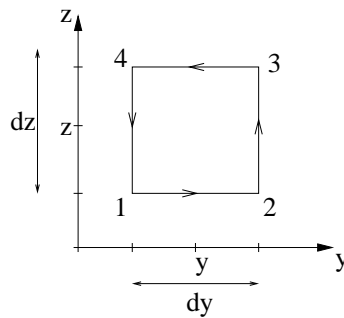


Abbildung 9.3.1: Zum Beweis der Integraldarstellung der Rotation.

9.4 Integralsatz von Stokes

Es sei $\underline{A}(r)$ ein beliebiges Vektorfeld. Aus Kapitel 8.2 wissen wir, dass gilt (falls der Definitionsbereich von \underline{A} einfach-zusammenhängend ist):

$$\underline{A} = -\text{grad } \varphi \iff \text{rot } \underline{A} = 0 \iff \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = 0,$$

wobei die rechte Seite für alle geschlossenen Wege \mathcal{C} gelten soll.

Was passiert für $\text{rot } \underline{A} \neq 0$? Man erwartet, dass dann i.a. auch $\oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} \neq 0$ sein wird. Dies bringt der Integralsatz von Stokes (in quantitativer Form) zum Ausdruck.

Satz 9.4.1 (Stokes'scher Integralsatz).

Sei F eine beliebige (gekrümmte) Fläche im \mathbb{R}^3 mit Rand \mathcal{C} , d.h. der Rand ist eine geschlossene Kurve (siehe Abb. 9.4.1a). Dann gilt:

$$\boxed{\oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = \int_F \text{rot } \underline{A} \cdot d\underline{f}.}$$

Dabei ist das infinitesimale Flächenelement $d\underline{f} = \underline{n}df$ gemäß einer rechten-Hand-Regel (siehe Abb. 9.4.1b) orientiert.

Der Satz von Stokes besagt also, dass das Wegintegral von \underline{A} längs des geschlossenen Weges \mathcal{C} gleich dem Fluss von $\text{rot } \underline{A}$ durch die von \mathcal{C} berandete Fläche F ist.

Beim Beweis gehen wir ähnlich vor wie beim Gaußschen Satz. Zunächst unterteilen wir die Fläche F in zwei Teile F_1 und F_2 . Diese Teile werden von den geschlossenen Kurven \mathcal{C}_1 und \mathcal{C}_2 berandert (Abb. 9.4.2a). Dann gilt:

$$\oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = \oint_{\mathcal{C}_1} \underline{A} \cdot d\underline{r} + \oint_{\mathcal{C}_2} \underline{A} \cdot d\underline{r}.$$

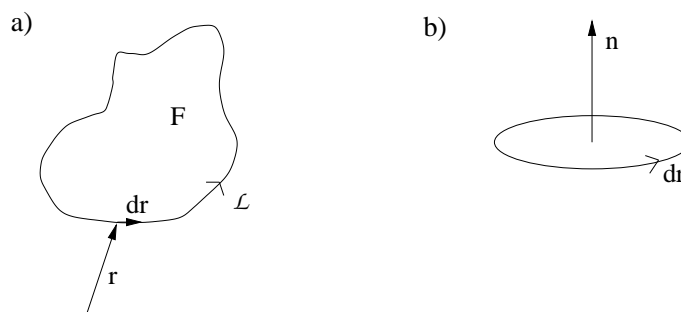


Abbildung 9.4.1: Zum Satz von Stokes. a) Fläche F mit Rand C . b) Orientierung der Flächennormalen (rechte-Hand-Regel).

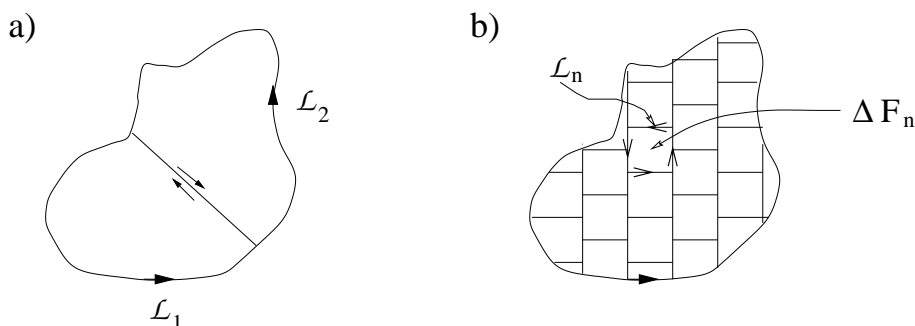


Abbildung 9.4.2: Zum Beweis des Stokesschen Integralsatzes.

Dabei haben wir schon ausgenutzt, dass die Schnittfläche von C_1 und C_2 in unterschiedlichen Richtungen durchlaufen wird und sich die entsprechenden Beiträge deshalb wegheben.

Die obigen Teilungsprozedur setzen wir nun immer weiter fort (Abb. 9.4.2b). Dann erhalten wir:

$$\begin{aligned} \oint_C \underline{A} \cdot d\underline{r} &= \sum_n \oint_{C_n} \underline{A} \cdot d\underline{r} = \sum_n \left[\frac{1}{\Delta F_n} \oint_{C_n} \underline{A} \cdot d\underline{r} \right] \Delta F_n \\ &\rightarrow \int_F \text{rot } \underline{A} \cdot \underline{n} df \end{aligned}$$

wobei wir beim Übergang zur zweiten Zeile die Summe wieder als Zwischensumme eines Integrals interpretiert haben.

Der Integralsatz von Stokes hat zahlreiche wichtige Anwendungen in der Physik, insbesondere in der Elektrodynamik im Zusammenhang mit den Maxwell'schen Gleichungen.

9.5 Integraldarstellung der Divergenz

Eine andere Interpretation der Divergenz erhält man durch Betrachtung des Vektorflusses durch die Oberfläche ΔF eines Volumenelements ΔV (siehe Abb. 9.5.1):

$$\operatorname{div} \underline{A} = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \int_{\Delta F} \underline{A} \cdot d\underline{f}.$$

Die Divergenz des Vektorfeldes kann also als Grenzfall des Flusses durch die Oberfläche eines kleinen Volumenelements aufgefasst werden. Dies quantifiziert unsere anschauliche Interpretation der Divergenz als ‘‘Quellstärke’’ des Feldes. Befindet sich eine Quelle in dem betrachteten Volumen, so strömt mehr heraus als herein. Dann ist $\operatorname{div} \underline{A} > 0$. Entsprechend ist bei einer Senke $\operatorname{div} \underline{A} < 0$.

Abbildung 9.5.1: Fluß durch ΔV

Zum Beweis der Integraldarstellung betrachten wir den Fluss durch die Oberfläche eines Würfels mit den Kantenlängen dx , dy und dz um $\underline{r} = (x, y, z)$ (siehe Abb. 9.5.2). Auf der Vorderseite F_1 zeige der Normalenvektor \underline{n} in x -Richtung, auf der Rückseite F_2 in $-x$ -Richtung. Die Flüsse durch F_1 und F_2 sind dann

$$\text{Fluss durch } F_1 = A_x\left(x + \frac{1}{2}dx, y, z\right)dydz,$$

$$\text{Fluss durch } F_2 = A_x\left(x - \frac{1}{2}dx, y, z\right)dydz$$

wobei $A_x(\underline{r}) = A_x(x, y, z)$ die x -Komponente des Vektorfeldes am Ort \underline{r} , dem Zentrum des Würfels, bezeichnet. Wir haben hier angenommen, dass das Vektorpotential auf den Flächen als konstant angesehen werden kann, wobei der Wert durch den Wert in der Mitte gegeben ist.

Somit erhält man für den Fluss durch die beiden Flächen F_1 und F_2 :

$$\begin{aligned} \int_{F_1+F_2} \underline{A} \cdot d\underline{f} &= \left[A_x\left(x + \frac{1}{2}dx, y, z\right) - A_x\left(x - \frac{1}{2}dx, y, z\right) \right] dydz \\ &= \left[\left(A_x(x, y, z) + \frac{\partial A_x}{\partial x} \frac{1}{2}dx \right) - \left(A_x(x, y, z) - \frac{\partial A_x}{\partial x} \frac{1}{2}dx \right) \right] dydz \\ &= \frac{\partial A_x(\underline{r})}{\partial x} dx dy dz. \end{aligned}$$

Dabei haben wir beim Übergang zur zweiten Zeile ausgenutzt, dass der Würfel infinitesimal ist und wir deshalb die Feldwerte durch eine Taylor-Entwicklung erster Ordnung approximieren können.

Analog erhalten wir dann für den Fluss in y - und z -Richtung:

$$\begin{aligned} y\text{-Richtung:} & \int_{F_3+F_4} \underline{A} \cdot d\underline{f} = \frac{\partial A_y(\underline{r})}{\partial y} dx dy dz, \\ z\text{-Richtung:} & \int_{F_5+F_6} \underline{A} \cdot d\underline{f} = \frac{\partial A_z(\underline{r})}{\partial z} dx dy dz. \end{aligned}$$

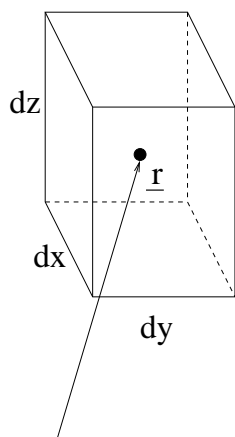


Abbildung 9.5.2: Zum Gaußschen Satz: Fluss durch einen infinitesimalen Würfel mit den Kantenlängen dx , dy und dz . Das Zentrum des Würfels befindet sich bei $\underline{r} = (x, y, z)$.

Fasst man nun diese drei Teilergebnisse zusammen, so folgt die angegebene Integraldarstellung der Divergenz.

9.6 Gaußscher Satz

Ziel des Gaußschen Integralsatzes ist es, eine andere Darstellung für ein Flussintegral der Form $\oint \underline{A} \cdot d\underline{f}$ anzugeben.

Satz 9.6.1 (Gaußscher Integralsatz).

Sei $\underline{A}(\underline{r})$ ein Vektorfeld und F eine beliebige geschlossene Fläche, die ein Volumen V umschließt. Dann gilt:

$$\oint_F \underline{A} \cdot d\underline{f} = \int_V \operatorname{div} \underline{A} dV.$$

Der **Gaußsche Integralsatz** erlaubt also die Umwandlung eines Flussintegrals in ein Volumenintegral über die Divergenz des Vektorfeldes über das von der Fläche eingeschlossene Volumen.

Zum Beweis teilen wir zunächst das Volumen V in zwei (beliebige) Teilvolumina V_1 und V_2 (siehe Abb. 9.6.1). Die neue Oberfläche der beiden Teilvolumina bezeichnen wir mit F_1 bzw. F_2 . Man beachte, dass eine gemeinsame Trennfläche zwischen den beiden Teilvolumina entsteht. Da die Normalen jeweils aus dem Volumen herauszeigen, ist auf der gemeinsamen Trennfläche $\underline{n}_1 = -\underline{n}_2$ (siehe Abb. 9.6.1) und somit

$$\oint_F \underline{A} \cdot d\underline{f} = \oint_{F_1} \underline{A} \cdot d\underline{f} + \oint_{F_2} \underline{A} \cdot d\underline{f},$$

da sich die Beiträge auf der Trennfläche wegheben.

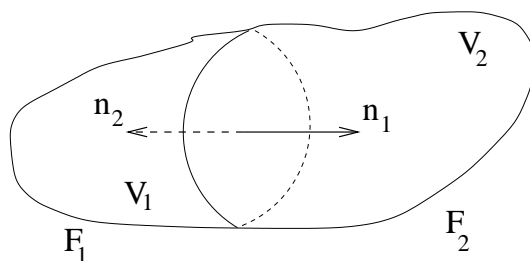


Abbildung 9.6.1: Zum Gaußschen Satz: Unterteilung des Volumens V in zwei Teilvolumina V_1 und V_2 . $\underline{n}_1, \underline{n}_2$ sind die Normalenvektoren auf der Trennfläche.

Wir setzen nun diese Teilungsprozedur immer weiter fort und erhalten so eine Einteilung in viele kleine Volumina ΔV_n mit Oberfläche F_n :

$$\begin{aligned} \oint_F \underline{A} \cdot d\underline{f} &= \sum_{n=1}^N \oint_{F_n} \underline{A} \cdot d\underline{f} = \sum_{n=1}^N \left[\frac{1}{\Delta V_n} \oint_{F_n} \underline{A} \cdot d\underline{f} \right] \Delta V_n \\ &\xrightarrow{\Delta V \rightarrow 0} \int_V \left(\lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \oint_{F(\Delta V)} \underline{A} \cdot d\underline{f} \right) dV = \int_V \operatorname{div} \underline{A} dV \end{aligned}$$

wobei $F(\Delta V)$ die Oberfläche des Würfels ΔV um den Punkt \underline{r} ist. Im letzten Schritt haben wir die Integraldarstellung der Divergenz aus Abschnitt 9.5 verwendet.

Eine wichtige Anwendung des Gaußschen Integralsatzes ist die Herleitung der sogenannten **Kontinuitätsgleichung**

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \underline{j} = 0.$$

Gleichungen dieses Typs treten in der Physik immer wieder auf, z.B. in der Hydrodynamik, denn sie der mathematische Ausdruck für ein lokales (d.h. von \underline{r} abhängiges) Erhaltungsgesetz. $\rho(\underline{r}, t)$ bezeichnet dabei immer eine Dichte (z.B. eine Massen- oder Ladungsdichte) und $\underline{j}(\underline{r}, t)$ die zugehörige Stromdichte (z.B. den Massen- oder elektrischen Strom). In den Übungen (Aufg. 19) werden wir die Kontinuitätsgleichung aus sehr allgemeinen Überlegungen ableiten.

Kapitel 10

Fourierreihen und Integraltransformationen

10.1 Funktionenräume

Wir haben schon in den Übungen (Aufgabe 4 und 12) gesehen, dass man auch Funktionen auch als Vektoren interpretieren kann, nicht nur Elemente des \mathbb{R}^n . Dies ist oft sehr nützlich, z.B. in der Quantenmechanik. Dabei sind Addition und Multiplikation mit einem Skalar wie üblich definiert:

$$h = f + \alpha g, \quad \text{d.h. } h(x) = f(x) + \alpha g(x) \text{ für alle } x.$$

Ein Beispiel ist die Menge

$$C^m(X) := \{m\text{-fach stetig differenzierbaren Funktionen auf } X\}.$$

Die bekannten Definitionen und Konzepte wie ‘Basis’ etc. lassen sich übertragen. Oft sind Funktionenräume unendlich-dimensional. Außerdem gibt es weitere nützliche Strukturen, die wir im Folgenden vorstellen wollen.

10.1.1 Metrische und normierte Räume

Zunächst wollen wir den bekannten Abstands begriff verallgemeinern.

Definition 10.1.1 (metrischer Raum, Metrik).

Ein **metrischer Raum** ist eine Menge M , auf der Abstände mittels einer Metrik definiert sind. Dabei heißt eine Abbildung $d : M \times M \rightarrow \mathbb{R}$ **Metrik**, wenn für alle $x, y, z \in M$ gilt:

1. $d(x, y) \geq 0$ (Nichtnegativität)
2. $d(x, y) = d(y, x)$ (Symmetrie)
3. $d(x, y) = 0$ genau dann, wenn $x = y$ (Eindeutigkeit)
4. $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Neben der Metrik gibt es noch eine weitere wichtige Struktur.

Definition 10.1.2 (normierter Raum, Norm).

Eine Abbildung $\|\dots\| : V \rightarrow \mathbb{R}$ auf einem Vektorraum V heißt **Norm**, wenn für alle $x, y, z \in V$ gilt:

1. $\|x\| \geq 0$ (Nichtnegativität)
2. $\|x - y\| = 0$ genau dann, wenn $x = \underline{0}$ (Eindeutigkeit)
3. $\|\lambda x\| = |\lambda| \|x\|$ mit $\lambda \in \mathbb{R}$ (Skalierung)
4. $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

In diesem Fall bezeichnet man V auch als **normierten Raum**.

Bemerkung. Hat man eine Norm $\|\dots\|$, so definiert

$$d(x, y) := \|x - y\|$$

eine Metrik.

Beispiel 10.1.1. Ein wichtiges Beispiel, das z.B. in der Quantenmechanik eine zentrale Rolle spielt, ist der Raum der quadratisch-integriblen Funktionen auf $X \subset \mathbb{R}$,

$$L^2(\mathbb{R}) := \left\{ f : X \rightarrow \mathbb{C} \mid \int_X |f(x)|^2 dx < \infty \right\}$$

mit der Norm

$$\|f\|_2 := \sqrt{\int_X |f(x)|^2 dx}.$$

10.1.2 Skalarprodukt

Definition 10.1.3 (Skalarprodukt).

Es sei X ein Vektorraum über \mathbb{C} . Ein **Skalarprodukt** (\dots, \dots) ist eine Abbildung $X \times X \rightarrow \mathbb{C}$ mit folgenden Eigenschaften (für alle $x, y, z \in X$ und $\lambda \in \mathbb{C}$):

1. $(x, y) = \overline{(y, x)}$,
2. $(\lambda x, y) = \bar{\lambda}(x, y)$,
 $(x, \lambda y) = \lambda(x, y)$,
3. $(x + y, z) = (x, z) + (y, z)$,
4. $(x, x) \geq 0$,
5. $(x, x) = 0 \Rightarrow x = 0$.

Dabei bezeichnet \bar{a} das komplex-Konjugierte von a .

Eine unmittelbare Folgerung aus Eigenschaft 1. ist, dass $(x, x) \in \mathbb{R}$ ist.

Eigenschaft 2. bedeutet, dass das Skalarprodukt antilinear im ersten Argument ist und linear im zweiten. Dies ist eine Konvention, die in der Physik häufig benutzt wird.

Aus Eigenschaft 3. folgt wegen 1., dass auch $(x, y + z) = (x, y) + (x, z)$ gilt.

Weitere Folgerungen aus den Eigenschaften sind die

- **Cauchy-Schwarz-Ungleichung**

$$|(x, y)|^2 \leq (x, x)(y, y),$$

- **Minkowski-Ungleichung**

$$\sqrt{(x + y, x + y)} \leq \sqrt{(x, x)} + \sqrt{(y, y)}.$$

Definition 10.1.4 (Hilbertraum).

Ein Vektorraum mit Skalarprodukt heisst auch **Prähilbertraum**. Ist der Raum vollständig, d.h. jede Cauchy-Folge konvergiert gegen ein Element des Raumes, so heisst er **Hilbertraum**.

Bemerkung.

1. Durch $\|f\| := \sqrt{(f, f)}$ ist eine Norm definiert.
2. Die Definitionen gelten analog für reelle Räume, wobei man überall die komplexe Konjugation weglässt. Für die Quantenmechanik sind aber komplexe Räume wichtig.

Beispiel 10.1.2.

Wir können in $L^2(X)$ durch

$$(f, g) := \int_X \bar{f}(x)g(x)dx$$

ein Skalarprodukt definieren.

Eine Verallgemeinerung hiervon ist

$$(f, g) := \int_X w(x)\bar{f}(x)g(x)dx$$

mit einer nicht-negativen Gewichtsfunktion $w(x)$ (z.B. $w(x) = e^{-x^2}$).

10.2 Fourierreihen

Im diesem Abschnitt wollen wir untersuchen, inwieweit sich periodische Funktionen nach ‘Teilfrequenzen’ zerlegen lassen. Dieses Verfahren ist auch unter dem Namen **Fourier-Analyse** bekannt.

Zunächst wollen wir noch einmal an die Definition von periodischen Funktionen erinnern (siehe Kap. 4.2.4):

Definition 10.2.1 (Periodische Funktionen).

Die $f(t)$ ist heißt **periodisch mit der Periode** T ($T > 0$), wenn für alle t gilt:

$$f(t + T) = f(t).$$

Der kleinste Wert von $T > 0$, der dies erfüllt, heißt **kleinste Periode** oder einfach nur **die Periode** von $f(t)$.

Prototypen periodischer Funktionen sind die komplexe Exponentialfunktion $e^{i\omega t}$ und die trigonometrischen Funktionen $\sin \omega t$ und $\cos \omega t$, die jeweils die Periode $T = \frac{2\pi}{\omega}$ haben.

Im folgenden werden wir zur Vereinfachung $x := \omega t$ als Variable wählen, d.h. wir betrachten nur die normierte Periode $T = 2\pi$ der Funktionen $\sin x$, $\cos x$. Am Ende des Abschnittes werden wir dann der Vollständigkeit halber die allgemeinen Formeln für beliebige Perioden T angeben.

Definition 10.2.2 (Fourierreihen).

Eine Reihe der Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

bezeichnet man als **Fourierreihe**. Man beachte, dass diese Reihe 2π -periodisch ist.

Im folgenden wollen wir die Frage untersuchen, wann sich eine periodische Funktion als Fourierreihe darstellen läßt.

Satz 10.2.1 (Fourierreihendarstellung periodischer Funktionen).

Die Funktion $f(x)$ sei im Intervall $[-\pi, \pi]$ definiert. Weiterhin sollen die sog. **Dirichlet-Bedingungen** erfüllt sein:

- a) $f(x)$ ist 2π -periodisch, d.h. $f(x + 2\pi) = f(x)$
- b) $f(x)$ und $f'(x)$ stückweise stetig in $[-\pi, \pi]$

Durch die Bedingung a) ist $f(x)$ für alle reellen Zahlen x eindeutig definiert. Bedingung b) bedeutet, dass $f(x)$ und $f'(x)$ im Intervall $[-\pi, \pi]$ bis auf endlich viele Punkte stetig sind.

Sind die Dirichlet-Bedingungen erfüllt, dann gilt: $f(x)$ ist darstellbar als **Fourierreihe**, d.h.

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

mit den Koeffizienten (**Fourier-Koeffizienten**)

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx & (n = 0, 1, \dots), \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx & (n = 1, 2, \dots). \end{aligned}$$

Das bedeutet präziser: Die Fourierreihe konvergiert gegen den Wert

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx] \rightarrow \begin{cases} f(x), & \text{wenn } f \text{ stetig bei } x \text{ ist} \\ \frac{f(x+0)+f(x-0)}{2}, & \text{wenn } f \text{ unstetig bei } x \text{ ist.} \end{cases}$$

Dabei bezeichnen $f(x+0)$ und $f(x-0)$ den rechts- bzw. linksseitigen Grenzwert (siehe Abb. 10.2.1).

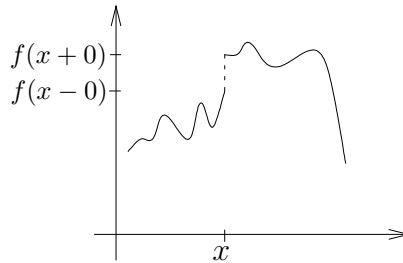


Abbildung 10.2.1: Unstetige Funktion

Bemerkungen:

1. Die Dirichlet-Bedingungen sind hinreichend, aber nicht notwendig. Für die physikalische Praxis reicht dies i.a. aus.
2. Als Periodenintervall kann jedes Intervall $[x_0, x_0 + 2\pi]$ gewählt werden, z. B. auch $[0, 2\pi]$.
3. Das Intervall $[-\pi, \pi]$ ist aber zweckmäßig, da symmetrisch um 0. Denn man sieht sofort:

$$\begin{aligned} a_n &= 0, & \text{wenn } f(x) &= -f(-x) & \text{(ungerade Funktion)} \\ b_n &= 0, & \text{wenn } f(x) &= f(-x) & \text{(gerade Funktion).} \end{aligned}$$

Das Interessante an obigem Satz ist, dass sich *jede* periodische Funktion (mit Dirichlet-Bedingungen) durch eine Reihe mit $\cos nx$ und $\sin nx$ darstellen lässt (Beweis: \Leftarrow Mathematik).

Eine andere Interpretation ist die über den Funktionenraum

$$\left\{ \frac{1}{\sqrt{2}}, \sin nx, \cos nx \mid n = 1, 2, \dots \right\}.$$

Dieser bildet ein Orthonormalsystem für periodische Funktionen in $L^2([-\pi, \pi])$ mit dem Skalarprodukt

$$(f, g) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx$$

Wir nehmen nun an, eine periodische Funktion lässt sich als Fourierreihe mit noch unbekanntem Koeffizienten a_n, b_n darstellen. Diese können wir dann folgendermassen berechnen:

$$\begin{aligned} (f, \cos mx) &= \left(\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx], \cos mx \right) \\ &= \left(\frac{a_0}{2}, \cos mx \right) + \sum_{n=1}^{\infty} a_n (\cos nx, \cos mx) + \sum_{n=1}^{\infty} b_n (\sin nx, \cos mx) \\ &= a_m. \end{aligned}$$

Im letzten Schritt haben wir ausgenutzt, dass die Funktionen orthonormiert sind, d.h. insbesondere ist $(\sin nx, \cos mx) = 0$ und $(\cos nx, \cos mx) = \delta_{n,m}$. Wenn man nun das Skalarprodukt $(f, \cos mx)$ explizit als Integral ausschreibt, erhält man die oben angegebene Darstellung der Koeffizienten. Analog bestimmt man die anderen Koeffizienten zu $b_m = (f, \sin mx)$.

Den Beweis, dass sich die periodische Funktion f tatsächlich in der angegebenen Reihenform schreiben lässt, überlassen wir der Mathematik. Hier haben wir gezeigt, wie sich die Koeffizienten bestimmen lassen, wenn dies der Fall ist.

Die Tatsache, dass sich beliebige periodische Funktionen durch die Elemente des oben angegebenen Orthonormalsystems darstellen lassen, gibt Anlass zu folgender Definition:

Definition 10.2.3 (vollständiges Funktionensystem).

Da sich beliebige 2π -periodische Funktionen als Reihe darstellen lassen, bezeichnet man die Menge $\{\cos nx, \sin nx \mid n = 0, 1, \dots\}$ als **vollständig** (zur Darstellung periodischer Funktionen).

Man kann auch die Taylorentwicklung analytischer Funktionen in diesem Sinne verstehen. Statt nach den (vollständigen) periodischen Funktionen $\cos nx$ und $\sin nx$ zu entwickeln, entwickelt man hier nach den Funktion x^n .

Vielfach ist eine *komplexe Darstellung* der Fourierreihe zweckmäßig und einfacher. Mit den bekannten Beziehungen

$$\begin{aligned} e^{\pm inx} &= \cos nx \pm i \sin nx, \\ \cos nx &= \frac{1}{2}(e^{inx} + e^{-inx}), \quad \sin nx = \frac{1}{2i}(e^{inx} - e^{-inx}), \end{aligned}$$

sieht man sofort ein, dass auch $\{e^{inx} \mid n = -\infty, \dots, \infty\}$ vollständig ist. Es gilt:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

mit

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

Die reelle und die komplexe Darstellung lassen sich leicht ineinander umrechnen:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + \sum_{n=1}^{\infty} [c_n e^{inx} + c_{-n} e^{-inx}] \\ &= c_0 + \sum_{n=1}^{\infty} [c_n (\cos nx + i \sin nx) + c_{-n} (\cos nx - i \sin nx)] \\ &= c_0 + \sum_{n=1}^{\infty} [(c_n + c_{-n}) \cos nx + i(c_n - c_{-n}) \sin nx] \end{aligned}$$

woraus folgt

$$c_0 = \frac{1}{2} a_0, \quad a_n = c_n + c_{-n}, \quad b_n = i(c_n - c_{-n}).$$

Die Funktionen e^{inx} bilden ein Orthonormalsystem bzgl. des Skalarproduktes

$$(f, g) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{f}(x) g(x) dx,$$

denn es gilt:

$$\begin{aligned} n \neq m : \quad (e^{imx}, e^{inx}) &= \frac{1}{2\pi} \frac{e^{i(n-m)x}}{i(n-m)} \Big|_{-\pi}^{\pi} = \frac{e^{i(n-m)\pi} - e^{-i(n-m)\pi}}{i(n-m)} = 0 \\ n = m : \quad (e^{inx}, e^{inx}) &= \frac{1}{2} \int_{-\pi}^{\pi} dx = 1. \end{aligned}$$

Dieses Ergebnis läßt sich mit Hilfe des Kronecker-Symbols kompakt zusammenfassen:

$$(e^{imx}, e^{inx}) = \delta_{n,m}.$$

Wir wollen nun noch einmal die Fourierdarstellung beweisen, diesmal in der komplexen Variante. Wieder nehmen wir an, dass sich die Funktion $f(x)$ als eine Fourierreihe schreiben läßt. Wir zeigen jetzt, dass die Entwicklungskoeffizienten gerade die in dem obigen Satz behauptete Form haben. Dabei verwenden wir wieder die Orthonormalität der Funktionen e^{inx} :

$$\frac{1}{2} \int_{-\pi}^{\pi} dx f(x) e^{-imx} = (e^{imx}, f) = \left(e^{imx}, \sum_{n=-\infty}^{\infty} c_n e^{inx} \right) = \sum_{n=-\infty}^{\infty} c_n (e^{imx}, e^{inx}) = c_m.$$

Beispiel 10.2.1. Wir betrachten die Funktion $f(x) = x$ in $[0, 2\pi]$, die wir uns 2π -periodisch auf ganz \mathbb{R} fortgesetzt denken (siehe Abb. 10.2.2)¹. Der Graph von $f(x)$ ähnelt dann einem Sägezahn.

$$n \neq 0: \quad c_n = \frac{1}{2\pi} \int_0^{2\pi} x e^{-inx} dx = \frac{x e^{-inx}}{-2\pi i n} \Big|_0^{2\pi} + \frac{1}{2\pi i n} \underbrace{\int_0^{2\pi} e^{-inx} dx}_{=0} = \frac{1}{-in} = \frac{i}{n}$$

$$n = 0: \quad c_0 = \frac{1}{2\pi} \int_0^{2\pi} x dx = \frac{x^2}{4\pi} \Big|_0^{2\pi} = \pi$$

Also haben wir folgende Darstellung von f als Fourierreihe:

$$x \Big|_{2\pi\text{-periodisch}} = \pi + \sum_{n \neq 0} \frac{i}{n} e^{inx} = \pi - \sum_{n=1}^{\infty} \frac{2}{n} \sin(nx),$$

wobei wir bei der letzten Umformung die Terme zu n und $-n$ zusammengefaßt haben. Man beachte, dass die Reihe an den Unstetigkeitspunkten $x = 0 \pmod{2\pi}$ (d. h. x ist ein Vielfaches von 2π) tatsächlich den Wert $\frac{f(x+0)+f(x-0)}{2} = \frac{0+2\pi}{2} = \pi$ interpoliert!

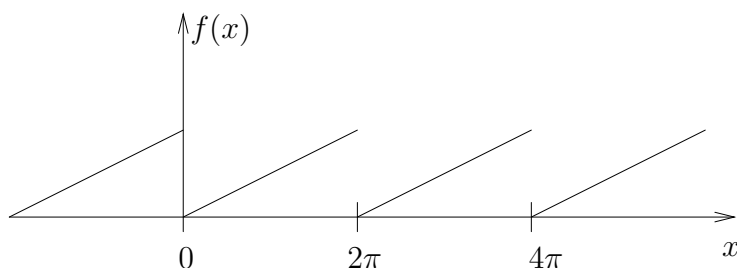


Abbildung 10.2.2: Sägezahn

Bemerkung:

Verwendet man statt der vollen Fourierreihe nur endlich viele Funktionen e^{inx} , so liefert dies eine Approximation der periodischen Funktion. Diese wird umso schlechter, je schneller sich die Funktion bei x ändert. Insbesondere an Unstetigkeiten tritt das sogenannte Gibbs-Phänomen auf.

Zum Abschluss wollen wir noch die Fourierdarstellung einer beliebigen T -periodischen Funktion ($T > 0$) angeben.

Satz 10.2.2 (Fourierreihendarstellung T -periodischer Funktionen).

Die Funktion $f(x)$ sei T -periodisch und erfülle die Dirichlet-Bedingungen im Intervall $[-\frac{T}{2}, \frac{T}{2}]$. Dann hat $f(x)$ folgende Fourierdarstellung:

¹Natürlich ist f auch periodisch mit Periodizitätsintervall $[-\pi, \pi]$. Die folgenden Rechnungen sind aber einfacher in $[0, 2\pi]$.

$$\begin{aligned}
 f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \frac{2\pi nx}{T} + b_n \sin \frac{2\pi nx}{T} \right] \\
 &= \sum_{n=-\infty}^{\infty} c_n e^{i \frac{2\pi nx}{T}}
 \end{aligned}$$

mit den Fourier-Koeffizienten

$$\begin{aligned}
 a_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(x) \cos \left(\frac{2\pi nx}{T} \right) dx & (n = 0, 1, \dots), \\
 b_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(x) \sin \left(\frac{2\pi nx}{T} \right) dx & (n = 1, 2, \dots), \\
 c_n &= \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-i \frac{2\pi nx}{T}} dx & (n = -\infty, \dots, \infty).
 \end{aligned}$$

Wie man sieht, geht dies für den Spezialfall $T = 2\pi$ in die bekannten Formeln über.

10.3 Fourier-Transformation

Im vorigen Abschnitt haben wir gesehen, dass sich periodische Funktionen als Fourierreihen darstellen lassen. Was passiert aber für Funktionen, die gar nicht periodisch sind? Dieser Fall entspricht gerade dem Grenzübergang $T \rightarrow \infty$.

Wir betrachten zunächst die Funktion $\cos n\omega t = \cos\left(\frac{2\pi n}{T}t\right)$, die periodisch mit Periode T ist. Wenn $T \rightarrow \infty$ liegen die Punkte $\omega_n = \frac{2\pi n}{T}$ immer dichter. Der Limes $T \rightarrow \infty$ entspricht dann dem Übergang von einer Fourier-Summe zu einem Fourier-Integral. Dies wollen wir im folgenden genauer betrachten.

Satz 10.3.1 (Fourier-Transformation).

Die Funktion $f(x)$ erfülle folgende Voraussetzungen:

- 1) $f(x)$ genügt den Dirichlet-Bedingungen im Intervall $] -\infty, \infty[$,
- 2) $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$.

Unter diesen Voraussetzungen gilt:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega$$

mit der sog. **Fourier-Transformierten**

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx.$$

Man beachte, dass die oben genannten Voraussetzungen hinreichend, aber nicht notwendig sind. In der Praxis sind sie aber meist erfüllt.

Bemerkung. F ist Fourier-Transformierte von f und f ist Fouriertransformierte von F !

Bemerkung. Wir haben hier eine symmetrische Aufteilung der Koeffizienten (d. h. $\frac{1}{\sqrt{2\pi}}$) vor dem Integral gewählt. Häufig werden andere Konventionen verwendet, d. h.

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} d\omega \dots \quad \text{und} \quad F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \dots \\ \text{oder} \quad f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \dots \quad \text{und} \quad F(\omega) = \int_{-\infty}^{\infty} dx \dots \end{aligned}$$

Die entsprechenden Fourier-Transformierten unterscheiden sich um einen Faktor ($\frac{1}{\sqrt{2\pi}}$ bzw. $\sqrt{2\pi}$). Dies muß man z. B. beim Nachschlagen in Tabellen immer beachten.

Ebenso, ob die Fourier-Transformierte als Integral über $e^{-i\omega x}$ oder $e^{+i\omega x}$ definiert ist. In letzterem Fall ändert sich natürlich auch das entsprechende Vorzeichen in der Fourier-Darstellung von $f(x)$!

Im folgenden zeigen wir, wie sich die Fourier-Transformation aus dem Grenzfall $T \rightarrow \infty$ der Fourier-Reihe

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n x} \quad \text{mit} \quad \omega_n = \frac{2\pi n}{T}, \\ c_n &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) e^{-i\omega_n x} dx \end{aligned}$$

ergibt, wobei $f(x) = f(x + T)$.

Wir machen nun den Grenzübergang $T \rightarrow \infty$ und $n \rightarrow \infty$ so, dass $\omega_n = \omega$ konstant gehalten wird. Dann ist auch $\Delta\omega_n = \omega_{n+1} - \omega_n = \frac{2\pi}{T} =: \Delta\omega$ konstant und beim Grenzübergang geht $\Delta\omega \rightarrow d\omega$. Für die Fourierreihe folgt:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n \frac{T}{2\pi} e^{i\omega_n x} \Delta\omega = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} F(\omega_n) e^{i\omega_n x} \Delta\omega \\ &\rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega \end{aligned}$$

wobei wir definiert haben

$$F(\omega_n) := \frac{T}{\sqrt{2\pi}} c_n = \frac{1}{\sqrt{2\pi}} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) e^{-i\omega_n x} dx \xrightarrow{t, n \rightarrow \infty} F(\omega).$$

Damit haben wir gezeigt, dass sich die Fourier-Transformierte als Grenzfall der Fourierreihe einer Funktion mit unendlicher Periode ergibt.

Bemerkung. In der Physik interessieren uns häufig nur reelle Funktionen $f(x)$.

Dann gilt: $F(\omega)^* = F(-\omega)$

Beweis:

$$\begin{aligned} F(\omega)^* &= \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx \right)^* = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{f^*(x)}_{f(x)} \underbrace{(e^{-i\omega x})^*}_{e^{i\omega x}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx = F(-\omega) \end{aligned}$$

Weitere solcher nützlicher Identitäten werden wir in den Übungen kennenlernen (Aufgabe 23). Als Beispiel wollen wir die Fouriertransformierte der Blockfunktion (bzw. der charakteristischen Funktion des Intervalls $[-a, a]$) (siehe Abb. 10.3.1)

$$f(x) = \begin{cases} 1 & |x| < a \\ 0 & |x| > a \end{cases}$$

berechnen:

$$\begin{aligned} F(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx = \frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{i\omega x} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_0^a \cos(\omega x) dx \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin(\omega x)}{\omega} \Big|_0^a = \sqrt{\frac{2}{\pi}} \frac{\sin(\omega a)}{\omega}. \end{aligned}$$

Beim Übergang zur zweiten Zeile haben wir ausgenutzt, dass das Integral über den Imaginärteil $\sin(\omega x)$ verschwindet, da der Sinus ungerade ist. Abb. 10.3.2 zeigt die Form der Fouriertransformierten. Das Maximum bei $x = 0$ bezeichnet man manchmal auch mit dem englischen Ausdruck **Peak** und die kleineren Maxima im oszillierenden Teil als **Wiggle**.

Speziell für $x = 0$ haben wir:

$$\begin{aligned} 1 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) d\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin(\omega a)}{\omega} d\omega = \frac{2}{\pi} \int_0^{\infty} \frac{\sin(\omega a)}{\omega} d\omega \\ &\stackrel{u=\omega a}{=} \frac{2}{\pi} \int_0^{\infty} \frac{\sin u}{u} du \end{aligned}$$

und somit

$$\int_0^{\infty} \frac{\sin u}{u} du = \frac{\pi}{2}.$$

Ähnlich wie die Fourierreihe bei der Berechnung komplizierter Reihen hilfreich sein kann, hilft die Fouriertransformation manchmal bei der Berechnung von Integralen.

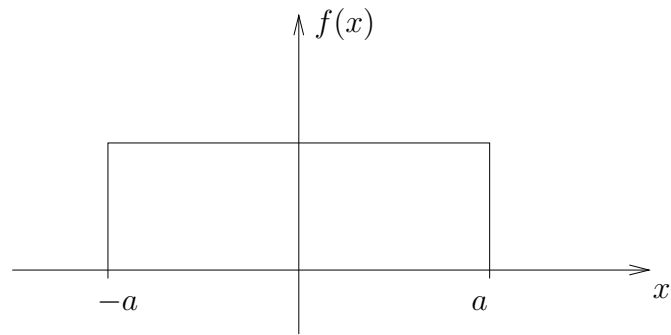


Abbildung 10.3.1: Blockfunktion

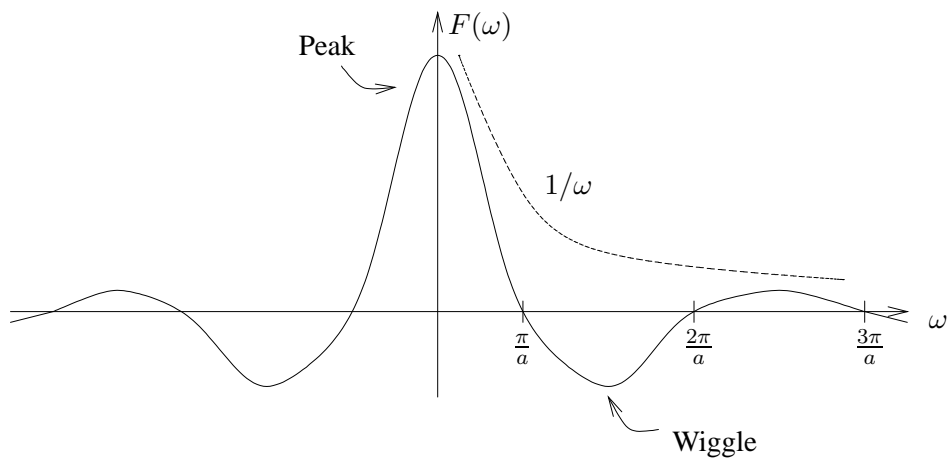


Abbildung 10.3.2: Fouriertransformierte der Blockfunktion

10.3.1 Delta-Funktion

Wir haben bereits häufig das Kronecker-Symbol

$$\delta_{jl} = \begin{cases} 1 & \text{falls } j = l \\ 0 & \text{falls } j \neq l \end{cases}$$

gewinnbringend verwendet. Eine ähnliche Größe kann man im kontinuierlichen Fall definieren.

Definition 10.3.1 (Delta-Funktion).

Wir definieren die sog. **Delta-Funktion** durch folgende Eigenschaften:

$$\delta(x - x_0) = 0 \quad \text{für alle } x \neq x_0,$$

und

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0)$$

wobei die zweite Eigenschaft für alle Funktionen $f(x)$ gelten soll, die stetig bei x_0 sind.

Bem.: Für analytische Funktionen $f(x)$ kann die zweite Bedingung abgeschwächt werden zu $\int_{-\infty}^{\infty} \delta(x - x_0)dx = 1$.

Dies ist eine heuristische Definition!! Eigentlich ist δ keine Funktion, sondern eine sogenannte **Distribution** (\rightarrow Mathematik): Der Funktion f wird ihr Funktionswert $f(x_0)$ an der Stelle x_0 zugeordnet. Als Funktion ist $\delta(x - x_0)$ hochgradig singular, da sie nur an einem Punkt x_0 von Null verschieden ist, aber trotzdem ein endliches Integral besitzt. $\delta(x - x_0)$ kann deshalb in x_0 keinen endlichen Wert annehmen, sondern wird dort unendlich. In Abb. 10.3.3 ist angedeutet, wie man sich $\delta(x - x_0)$ vorzustellen hat. Man kann die Delta-Funktion als das kontinuierliche Analogon des Kronecker-Deltas ansehen, insbesondere wenn man beachtet das $\delta_{a,b} = \delta_{a-b,0}$.

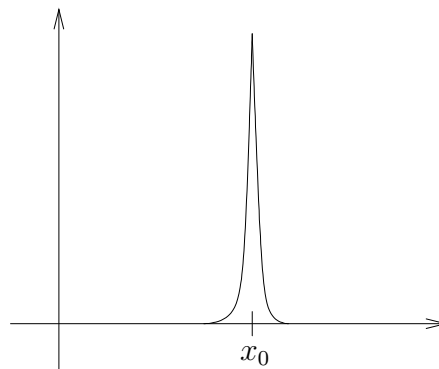


Abbildung 10.3.3: „Graph“ von $\delta(x - x_0)$

Die Delta-Funktion spielt eine wichtige Rolle in der Physik, denn sie hat viele Anwendungen. Ein Beispiel hierfür liefern die Punktmassen, von denen wir sehr oft gesprochen haben.

Für die Massendichte $\rho(x)$ einer Punktmasse m am Ort x_0 gilt:

$$\begin{aligned} \rho(x) &= 0 \quad \text{für } x \neq x_0, \\ \text{aber } m &= \int_{-\infty}^{\infty} \rho(x) dx. \end{aligned}$$

Somit sieht man, dass die Massendichte einer Punktmasse durch

$$\rho = m\delta(x - x_0)$$

gegeben ist. Mit Hilfe der Deltafunktion kann man also im Prinzip diskrete und kontinuierliche Systeme in einheitlicher Weise beschreiben.

Das Fourier-Integral führt auf eine „Darstellung“ der δ -Funktion. Setzt man nämlich die Fouriertransformierte explizit in die Fourierdarstellung ein, so folgt:

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{i\omega x} F(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{i\omega x} \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} e^{-i\omega y} f(y) \\ &= \int_{-\infty}^{\infty} f(y) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-y)} d\omega \right] dy \stackrel{!}{=} f(x). \end{aligned}$$

Hieraus liest man das wichtige Ergebnis

$$\boxed{\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-y)} d\omega = \delta(x-y)}$$

ab. Dies kann man als Verallgemeinerung von $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x} dx = \delta_{n,m}$ ansehen.

Zum Abschluss wollen wir noch eine weitere Darstellung von $\delta(x)$ angeben. Die Funktionenfolge $(\delta_\alpha)_\alpha$ mit

$$\delta_\alpha(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega x} e^{-\alpha|\omega|} = \frac{1}{2\pi} \frac{2\alpha}{\alpha^2 + x^2}$$

konvergiert nämlich für $\alpha \rightarrow 0$ gegen die Deltafunktion $\delta(x)$. Dies kann man zeigen, indem man verifiziert, dass für $\alpha \rightarrow 0$ die beiden definierenden Eigenschaften der Deltafunktion erfüllt sind. Dies wird in den Übungen genauer betrachtet, zusammen mit anderen Darstellungen und den wichtigsten Rechenregeln (siehe Aufgaben 26 und 27).

10.3.2 Faltungsintegral

Ein Integral der Form

$$h(x) := \int f(x-y)g(y)dy$$

bezeichnet man als **Faltungsintegral** bzw. als **Faltung** der Funktionen f und g . Es gilt folgender Zusammenhang des Faltungsintegrals mit den Fouriertransformierten F und G von f bzw. g :

$$\int_{-\infty}^{\infty} f(x-y)g(y)dy = \int_{-\infty}^{\infty} F(\omega)G(\omega)e^{i\omega x}d\omega.$$

Dies kann man auch so formulieren: Die Fourier-Transformierte des Faltungsintegrals ist das Produkt der Transformierten der gefalteten Funktionen:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} h(x)e^{-i\omega x}dx = F(\omega)G(\omega).$$

Der Beweis der ersten Aussage folgt durch Einsetzen der jeweiligen Fouriertransformierten in das Faltungsintegral und anschließender Vertauschung der Integrationsreihenfolge:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x-y)g(y)dy &= \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\omega d\omega'}{2\pi} F(\omega)e^{i\omega(x-y)}G(\omega')e^{i\omega'y} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\omega d\omega'}{2\pi} F(\omega)G(\omega') \underbrace{\int_{-\infty}^{\infty} dy e^{iy(\omega'-\omega)}}_{2\pi\delta(\omega'-\omega)} e^{i\omega x} \\ &= \int_{-\infty}^{\infty} d\omega F(\omega)e^{i\omega x} \int_{-\infty}^{\infty} d\omega' G(\omega')\delta(\omega'-\omega) \\ &= \int_{-\infty}^{\infty} d\omega F(\omega)G(\omega)e^{i\omega x}. \end{aligned}$$

Die zweite Aussage erhält man, indem man die erste Fourier-transformiert:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} h(x)e^{-i\omega x}dx &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} F(\omega')G(\omega')e^{i\omega'x}d\omega' \right] e^{-i\omega x}dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' F(\omega')G(\omega') \underbrace{\int_{-\infty}^{\infty} dx e^{i(\omega'-\omega)x}}_{2\pi\delta(\omega'-\omega)} \\ &= \int_{-\infty}^{\infty} d\omega' F(\omega')G(\omega')\delta(\omega'-\omega) \\ &= F(\omega)G(\omega). \end{aligned}$$

10.3.3 Anwendung: DGL

Als physikalische Anwendung der Fourier-Transformation betrachten wir die DGL für eine erzwungene Schwingung:

$$\ddot{x} + a\dot{x} + bx = f(t).$$

Dabei ist $x(t)$ die Auslenkung aus der Ruhelage zur Zeit t und $f(t)$ die (zeitabhängige) äußere Kraft. Die Terme auf der linken Seite repräsentieren die Beschleunigung, die Dämpfung durch eine geschwindigkeitsabhängige Reibungskraft und die Rückstellkraft (Hooke'sches Gesetz).

Wir wollen nun diese DGL mit Hilfe der Fourier-Transformation ganz allgemein lösen. Dazu stellen wir die Auslenkung $x(t)$ durch ihre Fourier-Transformierte $\tilde{x}(\omega)$ dar:

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} \tilde{x}(\omega) d\omega.$$

Nun können wir die Fourier-Transformierten der Zeitableitungen bestimmen:

$$\dot{x}(t) = \frac{1}{\sqrt{2\pi}} \frac{d}{dt} \int_{-\infty}^{\infty} e^{i\omega t} \tilde{x}(\omega) d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\partial}{\partial t} (e^{i\omega t} \tilde{x}(\omega)) d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} i\omega e^{i\omega t} \tilde{x}(\omega) d\omega.$$

Hieraus lesen wir ab:

$$\text{FT}(\dot{x}) = i\omega \text{FT}(x),$$

wobei $\text{FT}(\dot{x})$ die Fourier-Transformierte von \dot{x} bezeichnet. Analog folgt

$$\text{FT}(\ddot{x}) = -\omega^2 \text{FT}(x).$$

Wir sehen hier die wesentliche Vereinfachung, die wir durch Betrachtung der Fourier-Transformierten erzielt haben: Nach Fouriertransformation werden aus Ableitungen einfache Multiplikationen (mit $i\omega$)!!!

Wir wenden nun die Fouriertransformation auf die gesamte DGL an, d.h. wir multiplizieren beide Seiten der Gleichung mit $e^{-i\omega t}$ und integrieren dann. Unter Ausnutzung der Linearität sowohl der DGL als auch der Fourier-Transformation erhalten wir dann:

$$\begin{aligned} F(\omega) := \text{FT}(f(t)) &= \text{FT}(\ddot{x} + a\dot{x} + bx) = \text{FT}(\ddot{x}) + a \text{FT}(\dot{x}) + b \text{FT}(x) \\ &= -\omega^2 \text{FT}(x) + ai\omega \text{FT}(x) + b \text{FT}(x) = (b + ai\omega - \omega^2) \text{FT}(x) \\ &= (b + ai\omega - \omega^2) \tilde{x}(\omega). \end{aligned}$$

Somit können wir die Fourier-Transformierte $\tilde{x}(\omega)$ von $x(t)$ explizit bestimmen:

$$\tilde{x}(\omega) = \text{FT}(x) = \frac{F(\omega)}{b + ai\omega - \omega^2}.$$

Durch Rücktransformation erhalten wir dann die gesuchte Lösung

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \tilde{x}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \frac{F(\omega) e^{i\omega t}}{b + ai\omega - \omega^2}.$$

Zur expliziten Berechnung muss natürlich die antreibende Kraft $f(t)$ (und damit ihre Fourier-Transformierte $F(\omega)$) angegeben werden. Später werden wir sehen, wie wir solche Integrale mit Hilfe der Funktionentheorie berechnen können.

10.4 Laplace-Transformation

Wir betrachten kurz eine weitere Integraltransformation, die häufig in physikalischen Problemen gewinnbringend verwendet wird.

Es sei $f(t)$ auf $[0, \infty[$ definiert und integrierbar. Wir nehmen im folgenden o.B.d.A. an, dass $f(t) = 0$ für $t < 0$. Dann ist die **Laplace-Transformierte** von f definiert durch

$$F(p) := L(f) := \int_0^{\infty} dt f(t) e^{-pt},$$

wobei $\operatorname{Re}(p) > 0$ sein soll.

Beispielsweise gilt dann

$$L(1) = \int_0^{\infty} dt e^{-pt} = \frac{1}{p}$$

und

$$L(e^{iat}) = \frac{1}{p - ia}.$$

Ähnlich wie bei der Fourier-Transformation lassen sich wieder die Transformaten der Ableitungen von f leicht bestimmen:

$$\begin{aligned} L\left(\frac{df}{dt}\right) &= \int_0^{\infty} dt e^{-pt} \frac{df}{dt} = e^{-pt} f(t) \Big|_0^{\infty} - \int_0^{\infty} dt f(t) (-p) e^{-pt} \\ &= -f(0) + pL(f). \end{aligned}$$

Dabei haben wir im zweiten Schritt partiell integriert und angenommen, dass $\lim_{t \rightarrow \infty} e^{-pt} f(t) = 0$ ist.

Dieses Resultat kann natürlich auf höhere Ableitungen verallgemeinert und zur Lösung von DGL eingesetzt werden.

Die Umkehrung der Laplace-Transformation ist nicht ganz so einfach wie im Fall der Fourier-Transformation. Wir geben hier nur das Ergebnis an, das sog. **Bromwich-Integral**:

$$f(t) = L^{-1}(F) := \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} dz F(z) e^{zt},$$

wobei $t > 0$, $c > k$ wenn für $F = L(f)$ gilt, dass $\operatorname{Re}(p) > k$ ist. Wie man solche komplexen Wegintegral berechnet, lernen wir später im Kapitel über Funktionentheorie (siehe Kap. 16).

In der Praxis bestimmt man die Umkehrung meist mittels Tabellen aus Linearkombinationen bekannter Umkehrungen.

Kapitel 11

Koordinatensysteme

11.1 Gebräuchliche Koordinatensysteme

Bisher haben wir im \mathbb{R}^3 die kartesischen Koordinaten x, y, z (bzw. x_1, x_2, x_3), die Zylinderkoordinaten ρ, φ, z und die Kugelkoordinaten r, φ, ϑ kennengelernt. Daneben gibt es noch viele weitere Koordinaten, die in einzelnen Fällen hilfreich sein können!

Streng genommen haben wir aber meist nicht wirklich Zylinder- oder Kugelkoordinaten verwendet, d.h. Vektoren in der *Basis* dieser Koordinatensysteme dargestellt, sondern nur die *Komponenten* der kartesischen Koordinaten mit Hilfe der Variablen ρ, φ, z bzw. r, φ, ϑ *parametrisiert*.

Wir wollen in diesem Kapitel ganz allgemein herleiten, wie sich Größen in unterschiedlichen Koordinatensystemen beschreiben lassen. Wir konzentrieren uns auf den Fall des \mathbb{R}^3 , aber alle Überlegungen lassen sich in offensichtlicher Weise auf beliebige Dimensionen verallgemeinern. Bei unseren Betrachtungen werden uns die kartesischen Koordinaten als Referenzsystem dienen. Zur Vereinfachung bezeichnen wir sie im folgenden mit x_1, x_2, x_3 statt x, y, z , d.h. der Ortsvektor

ist durch $\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ gegeben.

Es seien nun u_1, u_2, u_3 beliebige andere Koordinaten. Wir stellen die kartesischen Koordinaten als Funktion dieser neuen Koordinaten dar:

$$\begin{aligned}x_1 &= x_1(u_1, u_2, u_3), \\x_2 &= x_2(u_1, u_2, u_3), \\x_3 &= x_3(u_1, u_2, u_3),\end{aligned}$$

wobei wir den Funktionsnamen gleich der entsprechenden kartesischen Komponente gewählt haben! Dieses Gleichungssystem ist nach den u_j auflösbar, falls für die Jacobi- (bzw. Funktional-) Determinante (vgl. Kap. 8.4) gilt:

$$\frac{\partial(x_1, x_2, x_3)}{\partial(u_1, u_2, u_3)} := \det \left(\frac{\partial x_j}{\partial u_i} \right)_{ji} \neq 0.$$

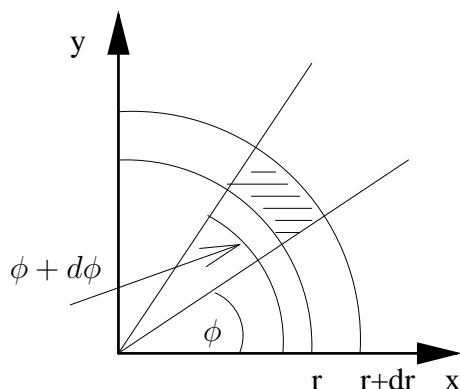


Abbildung 11.1.1: Koordinatenlinien in ebenen Polarkoordinaten. Die r -Linien sind radiale Gerade mit Winkel φ_0 und die φ -Linien Kreise vom Radius r_0 .

Wir betrachten nun die **Koordinatenlinien** (vgl. Kap. 8.4), die man erhält wenn man alle Koordinaten bis auf eine fixiert und diese freie Koordinate alle erlaubten Werte durchlaufen lässt. Die Koordinatenlinien durch den festen Punkt $\underline{r}^{(0)} = \underline{r}(u_1^{(0)}, u_2^{(0)}, u_3^{(0)})$ nennen wir

$$\underline{r}_1(u_1) = \underline{r}(u_1, u_2^{(0)}, u_3^{(0)}), \quad \underline{r}_2(u_1) = \underline{r}(u_1^{(0)}, u_2, u_3^{(0)}), \quad \underline{r}_3(u_1) = \underline{r}(u_1^{(0)}, u_2^{(0)}, u_3).$$

Schneiden sich die Koordinatenlinien paarweise in einem rechten Winkel, so spricht man von **rechtwinkligen** oder **orthogonalen Koordinaten**.

In kartesischen Koordinaten sind die Koordinatenlinien Geraden parallel zur korrespondierenden Achse, z.B. ist $\underline{r}_1(x_1) = \underline{r}(x_1, x_2^{(0)}, x_3^{(0)})$ eine Parallele zur x -Achse.

Abb. 11.1.1 zeigt die Koordinatenlinien $\underline{r}_1(r) = \underline{r}(r, \varphi_0)$ (r -Linien) und $\underline{r}_2(r) = \underline{r}(r_0, \varphi)$ (φ -Linien) der ebenen Polarkoordinaten.

Die Basisvektoren der kartesischen Koordinaten sind die Einheitsvektoren $\underline{e}_1, \underline{e}_2, \underline{e}_3$ in Richtung der Koordinatenachsen. Sie stimmen mit den Tangentenvektoren an die Koordinatenlinien überein! Dies wollen wir auf beliebige Koordinaten übertragen und definieren daher als Basisvektoren \underline{e}_{u_j} der Koordinaten u_1, u_2, u_3

$$\underline{e}_{u_j} := \frac{1}{h_{u_j}} \frac{\partial \underline{r}}{\partial u_j} \Big|_{\underline{r}=\underline{r}_0} \quad \text{mit} \quad h_{u_j} = \left| \frac{\partial \underline{r}}{\partial u_j} \right|.$$

Damit ist \underline{e}_{u_j} also ein normierter Tangentenvektor \underline{T}_j an die u_j -Linie durch \underline{r}_0 . Man beachte, dass daher die \underline{e}_{u_j} noch vom Punkt \underline{r}_0 abhängen können! Die Normierungsfaktoren h_{u_j} werden wir später explizit benötigen!

Für orthogonale Koordinaten, auf die wir uns hier beschränken wollen, gilt

$$\underline{e}_{u_j} \cdot \underline{e}_{u_l} = \delta_{jl}$$

Daher bilden die $\{\underline{e}_{u_j}\}$ eine Orthonormalbasis des \mathbb{R}^3 . I.a. benutzt man die Konvention, dass dieses System rechtshändig ist, d.h. es gilt

$$\underline{e}_{u_1} = \underline{e}_{u_2} \times \underline{e}_{u_3} \quad + \text{zyklische Indexvertauschungen.}$$

Under construction !

Abbildung 11.1.2: Die Richtung der Basisvektoren \underline{e}_{u_j} ändert sich i.a. mit der Position.

An dieser Stelle sei nochmals betont, dass sich i.a. die Richtung der Basisvektoren \underline{e}_{u_j} , und damit die Orientierung der Basis, als Funktion von \underline{r}_0 ändert (siehe Abb. 11.1.2). Man mache sich das am Beispiel der ebenen Polarkoordinaten klar (siehe Abb. 11.1.1). Der Basisvektor \underline{e}_r zeigt als Tangentenvektor an die r -Linie immer in radiale Richtung, die aber mit φ_0 variiert. Die kartesischen Koordinaten bilden hier eine Ausnahme, da die Koordinatenlinien durch verschiedenen Punkte parallel sind!

Wir wollen diese allgemeinen Betrachtungen auf die bereits bekannten Fälle der Zylinder- und Kugelkoordinaten spezialisieren. Die Ergebnisse werden wir dann verwenden, um z.B. die Form der Differentialoperatoren in diesen Systemen herzuleiten.

11.1.1 Zylinderkoordinaten

Wir hatten die Zylinderkoordinaten in Kapitel 1.4.3 eingeführt. Präziser gesagt, haben wir dort nur die 'Zylindervariablen' ρ, φ, x_3 definiert und mit ihnen die kartesischen Koordinaten parametrisiert. In unserer jetzigen Notation (mit $u_1 = \rho, u_2 = \varphi$ und $u_3 = x_3$) lauten die dort abgeleiteten Beziehungen:

$$\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \\ x_3 \end{pmatrix} \quad \text{mit} \quad \begin{array}{l} \rho = \sqrt{x_1^2 + x_2^2} \in [0, \infty[\\ \varphi = \arctan \frac{x_2}{x_1} \in [0, 2\pi[\\ x_3 \in \mathbb{R} \end{array} .$$

Die Koordinatenlinien lassen sich analog zum oben betrachteten Fall der ebenen Polarkoordinaten leicht bestimmen. Die ρ -Linien

$$\underline{r}_1(\rho) = \begin{pmatrix} \rho \cos \varphi_0 \\ \rho \sin \varphi_0 \\ x_3^{(0)} \end{pmatrix}$$

sind radiale Halbgeraden in Höhe $x_3^{(0)}$ im Winkel φ_0 . Die φ -Linien sind Kreise um den Ursprung und die x_3 -Linien sind Parallelen zur z -Achse.

Die Basisvektoren der Zylinderkoordinaten sind daher gegeben durch

$$\begin{aligned}
 \underline{e}_\rho &= \frac{1}{h_\rho} \frac{\partial \underline{r}}{\partial \rho} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, & h_\rho &= 1, \\
 \underline{e}_\varphi &= \frac{1}{h_\varphi} \frac{\partial \underline{r}}{\partial \varphi} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, & h_\varphi &= \rho, \\
 \underline{e}_{x_3} &= \frac{1}{h_{x_3}} \frac{\partial \underline{r}}{\partial x_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \underline{e}_3, & h_{x_3} &= 1.
 \end{aligned}$$

Man beachte, dass wir hier die Basisvektoren durch ihre Darstellung in kartesischen Koordinaten (als 3-komponentige Vektoren) angegeben haben.

Da

$$\underline{e}_\rho \times \underline{e}_\varphi = \underline{e}_{x_3},$$

wie man leicht nachrechnet, bilden $(\underline{e}_\rho, \underline{e}_\varphi, \underline{e}_{x_3})$ ein Rechtssystem.

11.1.2 Kugelkoordinaten

Die Kugelkoordinaten ρ, ϑ, φ hatten wir ebenfalls in Kapitel 1.4.4 kennengelernt¹. Auch sie hatten wir bisher im Wesentlichen zur Parametrisierung der kartesischen Komponenten benutzt:

$$\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix} \quad \text{mit} \quad \begin{aligned} \rho &= \sqrt{x_1^2 + x_2^2 + x_3^2} \in [0, \infty[\\ \vartheta &= \arctan \frac{\sqrt{x_1^2 + x_2^2 + x_3^2}}{x_3} \in [0, \pi] \\ \varphi &= \arctan \frac{x_2}{x_1} \in [0, 2\pi[\end{aligned}$$

Wir können nun die Basisvektoren der Kugelkoordinaten im kartesischen System angeben:

$$\begin{aligned}
 \underline{e}_r &= \frac{1}{h_r} \frac{\partial \underline{r}}{\partial r} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}, & h_r &= 1, \\
 \underline{e}_\vartheta &= \frac{1}{h_\vartheta} \frac{\partial \underline{r}}{\partial \vartheta} = \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ -\sin \vartheta \end{pmatrix}, & h_\vartheta &= r, \\
 \underline{e}_\varphi &= \frac{1}{h_\varphi} \frac{\partial \underline{r}}{\partial \varphi} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, & h_\varphi &= r \sin \vartheta.
 \end{aligned}$$

Da

$$\underline{e}_r \times \underline{e}_\vartheta = \underline{e}_\varphi,$$

bilden $(\underline{e}_r, \underline{e}_\vartheta, \underline{e}_\varphi)$ ein Rechtssystem.

¹Siehe auch Aufgabe 3 der Übungen.

11.2 Bestimmung von Vektorkomponenten

Definition 11.2.1 ((Vektor-)Komponenten).

Da die Vektoren $\{\underline{e}_{u_j}\}$ eine Basis bilden, kann man beliebige Vektoren \underline{F} in ihr darstellen:

$$\underline{F} = \sum_j F_{u_j} \underline{e}_{u_j}.$$

Die Zahlen F_{u_j} nennt man die **Komponenten** von \underline{F} bzgl. der Basis $\{\underline{e}_{u_j}\}$.

Für orthogonale Koordinaten ist $\{\underline{e}_{u_j}\}$ sogar eine Orthogonalbasis. Dann können wir die Komponente F_{u_l} einfach als Projektion von \underline{F} auf die Richtung \underline{e}_{u_l} (die Richtung der u_l -Koordinatenlinie) berechnen:

$$\underline{F} \cdot \underline{e}_{u_l} = \sum_j F_{u_j} \underline{e}_{u_j} \cdot \underline{e}_{u_l} = \sum_j F_{u_j} \delta_{jl} = F_{u_l}.$$

Die kartesischen Komponenten F_j , d.h. die Komponenten bzgl. der Basis $\{e_x, e_y, e_z\}$, sind offensichtlich ein Spezialfall der obigen Definition. Meist sind Vektoren in kartesischen Koordinaten gegeben:

$$\underline{F} = \sum_j F_j \underline{e}_j,$$

wobei wir vereinfachend \underline{e}_j für \underline{e}_{x_j} schreiben. Die kartesischen Komponenten fasst man dann

auch als Spaltenvektor $\underline{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$ zusammen.

Beispiel 11.2.1.

1. Die Komponenten des Ortsvektors \underline{r} in Zylinderkoordinaten sind gegeben durch:

$$\begin{aligned} r_\rho &= \underline{r} \cdot \underline{e}_\rho = \cos \varphi (\rho \cos \varphi) + \sin \varphi (\rho \sin \varphi) = \rho (\cos^2 \varphi + \sin^2 \varphi) = \rho, \\ r_\varphi &= \underline{r} \cdot \underline{e}_\varphi = (-\sin \varphi) \rho \cos \varphi + \cos \varphi (\rho \sin \varphi) = 0, \\ r_{x_3} &= \underline{r} \cdot \underline{e}_{x_3} = x_3. \end{aligned}$$

Dabei haben wir die explizite Form der Basisvektoren $\underline{e}_\rho, \underline{e}_\varphi, \underline{e}_{x_3}$ ausgenutzt, die wir in Kapitel 11.1.1 angegeben hatten. Somit erhalten wir also die Darstellung von \underline{r} in Zylinderkoordinaten als

$$\underline{r} = \rho \underline{e}_\rho + x_3 \underline{e}_{x_3}.$$

2. Analog können wir nun mit Hilfe der Ergebnisse aus Kapitel 11.1.2 den Ortsvektor in Kugelkoordinaten darstellen. Man sieht schnell, dass $\underline{r} \cdot \underline{e}_\varphi = 0 = \underline{r} \cdot \underline{e}_\vartheta$ ist und so ergibt sich

$$\underline{r} = r \underline{e}_r.$$

Die Darstellung in Zylinder- und insbesondere in Kugelkoordinaten ist also viel einfacher, als die Darstellung $\underline{r} = x_1 \underline{e}_{x_1} + x_2 \underline{e}_{x_2} + x_3 \underline{e}_{x_3}$ in kartesischen Koordinaten!

3. Abschließend wollen wir noch den Vektor $\underline{F} = x_3 \underline{e}_{x_1} = \begin{pmatrix} x_3 \\ 0 \\ 0 \end{pmatrix}$ in Kugelkoordinaten

darstellen.

Mit $x_3 = r \cos \vartheta$ folgt:

$$F_r = \underline{F} \cdot \underline{e}_r = \begin{pmatrix} r \cos \vartheta \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = r \sin \vartheta \cos \vartheta \cos \varphi,$$

$$F_\vartheta = \underline{r} \cdot \underline{e}_\vartheta = r \cos^2 \vartheta \cos \varphi,$$

$$F_\varphi = \underline{r} \cdot \underline{e}_\varphi = -r \cos \vartheta \sin \varphi.$$

Die Darstellung von \underline{F} in Kugelkoordinaten ist also wesentlich komplizierter als die in kartesischen Koordinaten.

Als Nächstes wollen wir uns mit dem Differenzieren von Vektoren beschäftigen. Da die Koordinatenlinien i.a. keine Geraden sind (außer im kartesischen Fall), ändern die Basisvektoren ihre Richtung und müssen daher auch differenziert werden, nicht nur die Komponenten, wie im kartesischen Fall. Bei Ableitung nach den Koordinaten u_l gilt daher:

$$\frac{\partial \underline{F}}{\partial u_j} = \sum_l \left(\frac{\partial F_{u_l}}{\partial u_j} \underline{e}_{u_l} + F_{u_l} \frac{\partial \underline{e}_{u_l}}{\partial u_j} \right).$$

Bei der Ableitung nach anderen Parametern (z.B. der Zeit) geht man unter Berücksichtigung der Kettenregel analog vor.

Beispiel 11.2.2. Wir wollen die Geschwindigkeit $\underline{v}(t)$ aus $\underline{r}(t)$ bestimmen.

In kartesischen Koordinaten sind die Basisvektoren \underline{e}_j konstant und wir haben die bekannte Formel

$$\underline{v}(t) = \dot{\underline{r}}(t) = \sum_j \dot{x}_j(t) \underline{e}_j,$$

d.h. wir erhalten die Komponenten der Geschwindigkeit durch Zeitableitung der Komponenten von \underline{r} .

In Zylinderkoordinaten ist das nicht so! Wir verwenden die in Beispiel 11.2.1 abgeleitete Darstellung des Ortsvektors in Zylinderkoordinaten und erhalten zunächst

$$\begin{aligned} \underline{v}(t) &= \dot{\underline{r}}(t) = \frac{d}{dt} (\rho \underline{e}_\rho + x_3 \underline{e}_{x_3}) \\ &= \dot{\rho} \underline{e}_\rho + \rho \dot{\underline{e}}_\rho + \dot{x}_3 \underline{e}_{x_3} + x_3 \dot{\underline{e}}_{x_3}. \end{aligned}$$

Wir benötigen nun die Zeitableitungen der Basisvektoren. Diese können wir uns aus den Ergebnissen in Kapitel 11.1.1 beschaffen, denn es gilt z.B.

$$\underline{e}_\rho = \frac{d}{dt} \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix} = \begin{pmatrix} -\dot{\varphi} \sin \varphi \\ \dot{\varphi} \cos \varphi \\ 0 \end{pmatrix} = \dot{\varphi} \underline{e}_\varphi.$$

Offensichtlich gilt $\dot{\underline{e}}_{x_3} = 0$ und somit erhalten wir für die Geschwindigkeit

$$\underline{v}(t) = \dot{\rho}\underline{e}_\rho + \rho\dot{\varphi}\underline{e}_\varphi + \dot{x}_3\underline{e}_{x_3}.$$

Dieses Ergebnis kann man alternativ auch folgendermaßen herleiten: Zunächst drückt man \underline{r} im kartesischen System mit Hilfe der Zylindervariablen aus:

$$\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \\ x_3 \end{pmatrix}.$$

Hieraus folgt

$$\dot{\underline{r}} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} \dot{\rho} \cos \varphi - \rho \dot{\varphi} \sin \varphi \\ \dot{\rho} \sin \varphi + \rho \dot{\varphi} \cos \varphi \\ \dot{x}_3 \end{pmatrix}.$$

Die Projektionen $\dot{\underline{r}} \cdot \underline{e}_\rho$ etc. ins Basissystem der Zylinderkoordinaten liefern dann wieder $\underline{v}(t) = \dot{\rho}\underline{e}_\rho + \rho\dot{\varphi}\underline{e}_\varphi + \dot{x}_3\underline{e}_{x_3}$.

Bei der obigen Rechnung haben wir die Ableitungen der Basisvektoren benötigt. Allgemein ist folgendes Ergebnis nützlich:

Es sei $\underline{e}(t)$ ein beliebiger Einheitsvektor, dessen Richtung von der Zeit abhängt. Dann gilt wegen $1 = \underline{e}(t) \cdot \underline{e}(t) = e^2(t)$ (wobei $e(t) = |\underline{e}(t)|$ den Betrag des Vektors bezeichnet)

$$0 = \frac{d}{dt}(\underline{e}(t) \cdot \underline{e}(t)) = 2\underline{e}(t) \cdot \dot{\underline{e}}(t)$$

und somit ist

$$\dot{\underline{e}}(t) \perp \underline{e}(t).$$

Die Ableitung eines Einheitsvektors steht immer senkrecht auf diesem!

In obigem Beispiel wissen wir daher, dass sich z.B. $\dot{\underline{e}}_\rho$ darstellen lassen muß als $\dot{\underline{e}}_\rho = a\underline{e}_\varphi + b\underline{e}_{x_3}$ mit Koeffizienten a, b . Diese lassen sich manchmal durch andere Überlegungen bestimmen.

Bemerkung.

- Es besteht ein Unterschied zwischen der bloßen Verwendung krummliniger Koordinaten (genauer: Variablen) und der Darstellung eines Vektors in einem solchen Basissystem!
- Die Spezifikation eines Vektors durch seine Komponenten macht nur bei Angabe des Basissystems Sinn! Wie wir an den Beispielen gesehen haben, sind die Komponenten eines Vektors in kartesischen Koordinaten als die in Kugelkoordinaten.

Beispiel 11.2.3. Wir wollen diese wichtigen Bemerkungen noch einmal an Hand eines Beispiels diskutieren. Dazu betrachten wir den Vektor $\underline{a} = 5x_1\underline{e}_1 + 5x_2\underline{e}_2 + 5x_3\underline{e}_3$ in der kartesischen Basis $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$. Dort hat der die Komponenten $(5x_1, 5x_2, 5x_3)$. Wir können nun diese Komponenten (!) auch in Kugelvariablen ausdrücken als $(5r \sin \vartheta \cos \varphi, 5r \sin \vartheta \sin \varphi, 5r \cos \vartheta)$. Man beachte:

Dies sind immer noch die *kartesischen* Komponenten, nun aber parametrisiert durch die Variablen der Kugelkoordinaten!

In der Basis $\{\underline{e}_r, \underline{e}_\vartheta, \underline{e}_\varphi\}$ der Kugelkoordinaten (und in den entsprechenden Variablen ausgedrückt) haben wir aber $\underline{a} = 5r\underline{e}_r$ mit den Komponenten $(5r, 0, 0)$. Diese können wir auch noch mit den kartesischen Variablen ausdrücken als $(5\sqrt{x_1^2 + x_2^2 + x_3^2}, 0, 0)$.

Bemerkung. Wenn nicht anders angegeben, sind mit ‘Komponenten’ i.a. die kartesischen Komponenten gemeint, z.B. in der komponentenweisen Definition eines Vektors.

11.3 Differentialoperatoren in krummlinigen Koordinaten

Als nächstes wollen wir die Frage klären, wie die Differentialoperatoren Gradient, Rotation, Divergenz und der Laplace-Operator in anderen orthogonalen Basissystemen aussehen.

11.3.1 Gradient

Es sei $\phi(\underline{r})$ ein Skalarfeld. Die Komponente des Gradienten $\text{grad } \phi$ in \underline{e}_{u_j} -Richtung ist gegeben durch

$$\begin{aligned} (\text{grad } \phi)_{u_j} &= (\text{grad } \phi) \cdot \underline{e}_{u_j} = (\text{grad } \phi) \cdot \left(\frac{1}{h_{u_j}} \frac{\partial \underline{r}}{\partial u_j} \right) \\ &= \frac{1}{h_{u_j}} \sum_l \frac{\partial \phi}{\partial x_l} \frac{\partial x_l}{\partial u_j} = \frac{1}{h_{u_j}} \frac{\partial \phi}{\partial u_j}, \end{aligned}$$

und somit gilt im neuen Basissystem

$$\boxed{\text{grad} = \sum_j \underline{e}_{u_j} \frac{1}{h_{u_j}} \frac{\partial}{\partial u_j}}.$$

Speziell für die uns bekannten Koordinaten erhalten wir hieraus explizit

kartesisch:	$\text{grad } \phi(x_1, x_2, x_3) = \sum_j \underline{e}_j \frac{\partial \phi}{\partial x_j},$
Zylinder:	$\text{grad } \phi(\rho, \varphi, x_3) = \left[\underline{e}_\rho \frac{\partial}{\partial \rho} + \frac{1}{\rho} \underline{e}_\varphi \frac{\partial}{\partial \varphi} + \underline{e}_{x_3} \frac{\partial}{\partial x_3} \right] \phi(\rho, \varphi, x_3),$
Kugel:	$\text{grad } \phi(r, \vartheta, \varphi) = \left[\underline{e}_r \frac{\partial}{\partial r} + \frac{1}{r} \underline{e}_\vartheta \frac{\partial}{\partial \vartheta} + \frac{1}{r \sin \vartheta} \underline{e}_\varphi \frac{\partial}{\partial \varphi} \right] \phi(r, \vartheta, \varphi).$

11.3.2 Divergenz

Aus dem allgemeinen Ergebnis für den Gradienten angewandt auf die Koordinaten selbst folgt zunächst eine Darstellung der Basisvektoren durch den Gradienten:

$$h_{u_j} \text{grad } u_j = h_{u_j} \sum_l \underline{e}_{u_l} \frac{1}{h_{u_l}} \frac{\partial u_j}{\partial u_l} = h_{u_j} \sum_l \underline{e}_{u_l} \frac{1}{h_{u_l}} \delta_{jl} = \underline{e}_{u_j}.$$

Wir betrachten nun ein Skalarfeld $\underline{A} = \sum_j A_{u_j}(u_1, u_2, u_3) \underline{e}_{u_j}$, dessen Komponenten A_{u_j} in den krummlinigen Koordinaten bekannt sind.

Zur Vereinfachung untersuchen wir zunächst nur die Wirkung der Divergenz auf den ersten Term $A_{u_1} \underline{e}_{u_1}$:

$$\begin{aligned} \text{div}(A_{u_1} \underline{e}_{u_1}) &= \text{div}((\underline{e}_{u_2} \times \underline{e}_{u_3}) A_{u_1}) \\ &= \text{div}(h_{u_2} h_{u_3} A_{u_1} (\text{grad } u_2 \times \text{grad } u_3)) \\ &= \text{grad}(h_{u_2} h_{u_3} A_{u_1}) \cdot (\text{grad } u_2 \times \text{grad } u_3) + h_{u_2} h_{u_3} A_{u_1} \text{div}(\text{grad } u_2 \times \text{grad } u_3) \\ &= \frac{1}{h_{u_2} h_{u_3}} \underline{e}_{u_1} \cdot \text{grad}(h_{u_2} h_{u_3} A_{u_1}) \\ &= \frac{1}{h_{u_2} h_{u_3}} \frac{\partial(h_{u_2} h_{u_3} A_{u_1})}{\partial u_1}. \end{aligned}$$

Dabei haben wir im ersten Schritt ausgenutzt, dass die \underline{e}_{u_j} ein rechtshändiges Orthonormalsystem bilden, und im zweiten Schritt die oben angegebene Darstellung der Basisvektoren durch den Gradienten. Im dritten Schritt wurde die allgemeine Produktregel

$$\text{div}(\phi \underline{A}) = \underline{A} \cdot \text{grad } \phi + \phi \text{div } \underline{A},$$

die für beliebige Skalarfelder ϕ und Vektorfelder \underline{A} gilt, benutzt. Eine ähnliche Identität für die Divergenz eines Kreuzproduktes wurde beim Übergang zur 4. Zeile ausgenutzt:

$$\text{div}(\underline{A} \times \underline{B}) = \underline{B} \cdot \text{rot } \underline{A} - \underline{A} \cdot \text{rot } \underline{B}.$$

Speziell für $\underline{A} = \text{grad } u_2$ und $\underline{B} = \text{grad } u_3$ folgt wegen $\text{rot}(\text{grad } \phi) = 0$ (siehe Aufgabe 15):

$$\text{div}(\text{grad } u_2 \times \text{grad } u_3) = 0.$$

Die Rechnung für die anderen Terme geht analog. Als Endergebnis für die Divergenz in einem beliebigen orthogonalen Koordinatensystem erhalten wir daher

$$\boxed{\text{div } \underline{A}(u_1, u_2, u_3) = \frac{1}{h_{u_1} h_{u_2} h_{u_3}} \left[\frac{\partial}{\partial u_1} (h_{u_2} h_{u_3} A_{u_1}) + \frac{\partial}{\partial u_2} (h_{u_1} h_{u_3} A_{u_2}) + \frac{\partial}{\partial u_3} (h_{u_1} h_{u_2} A_{u_3}) \right].}$$

Speziell für die uns bekannten Koordinaten erhalten wir hieraus explizit

kartesisch:	$\operatorname{div} \underline{A} = \sum_j \frac{\partial A_j}{\partial x_j},$
Zylinder:	$\operatorname{div} \underline{A}(\rho, \varphi, x_3) = \frac{1}{\rho} \frac{\partial(\rho A_\rho)}{\partial \rho} + \frac{1}{\rho} \frac{\partial A_\varphi}{\partial \varphi} + \frac{\partial A_{x_3}}{\partial x_3},$
Kugel:	$\operatorname{div} \underline{A}(r, \vartheta, \varphi) = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin \vartheta} \frac{\partial(\sin \vartheta A_\vartheta)}{\partial \vartheta} + \frac{1}{r \sin \vartheta} \frac{\partial A_\varphi}{\partial \varphi}.$

Dabei haben wir das Vektorfeld jeweils in der entsprechenden Basis dargestellt, z.B. für Zylinderkoordinaten als $\underline{A}(\rho, \varphi, x_3) = A_\rho \underline{e}_\rho + A_\varphi \underline{e}_\varphi + A_{x_3} \underline{e}_{x_3}$.

Beispiel 11.3.1. Als Beispiel für die Nützlichkeit dieser Ergebnisse wollen wir die Divergenz des Vektorfeldes $\underline{F} = \frac{1}{r^2} \underline{e}_r$ berechnen, das also in der Basis der Kugelkoordinaten gegeben ist. Es beschreibt z.B. die Gravitationskraft zwischen zwei Punktmassen.

Die Komponenten von \underline{F} in Kugelkoordinaten sind $F_r = \frac{1}{r^2}$ und $F_\vartheta = F_\varphi = 0$. Somit gilt

$$\operatorname{div} \underline{F} = \frac{1}{r^2} \frac{\partial(r^2 F_r)}{\partial r} = 0.$$

11.3.3 Rotation

Wir betrachten zunächst nur den Anteil $A_{u_1} \underline{e}_{u_1}$ des Vektorfeldes $\underline{A} = \sum_j A_{u_j} \underline{e}_{u_j}$ und gehen ähnlich vor wie bei der Berechnung der Divergenz:

$$\begin{aligned} \operatorname{rot}(A_{u_1} \underline{e}_{u_1}) &= \operatorname{rot}(h_{u_1} A_{u_1} \operatorname{grad} u_1) \\ &= h_{u_1} A_{u_1} \operatorname{rot}(\operatorname{grad} u_1) + (\operatorname{grad}(h_{u_1} A_{u_1})) \times \operatorname{grad} u_1 \\ &= -\operatorname{grad} u_1 \times (\operatorname{grad}(h_{u_1} A_{u_1})) \\ &= -\frac{1}{h_{u_1}} \underline{e}_{u_1} \times \left(\sum_l \frac{1}{h_{u_l}} \frac{\partial(h_l A_{u_1})}{\partial u_l} \underline{e}_{u_l} \right) \\ &= \frac{1}{h_{u_1} h_{u_3}} \frac{\partial(h_1 A_{u_1})}{\partial u_3} \underline{e}_{u_2} - \frac{1}{h_{u_1} h_{u_2}} \frac{\partial(h_1 A_{u_1})}{\partial u_2} \underline{e}_{u_3}. \end{aligned}$$

Dabei haben wir die Identität

$$\operatorname{rot}(\phi \underline{A}) = \phi \operatorname{rot} \underline{A} + \operatorname{grad} \phi \times \underline{A}$$

benutzt, die für beliebige Skalarfelder ϕ und Vektorfelder \underline{A} gilt.

Die Rechnung für die anderen Beiträge geht analog. Insgesamt kann man die Rotation kompakt in folgender Form als Determinante schreiben:

$\operatorname{rot} A(u_1, u_2, u_3) = \begin{vmatrix} h_{u_1} \underline{e}_{u_1} & h_{u_2} \underline{e}_{u_2} & h_{u_3} \underline{e}_{u_3} \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_{u_1} A_{u_1} & h_{u_2} A_{u_2} & h_{u_3} A_{u_3} \end{vmatrix}.$
--

Diese Regel ist analog zur bekannten Merkregel für das Kreuzprodukt zu interpretieren: Man entwickelt die Determinante formal nach der ersten Zeile, die ja Vektoren als Elemente enthält, und erhält so die Komponenten der Rotation. Bei der Entwicklung ist zu beachten, dass man die Reihenfolge der Faktoren nicht vertauschen darf, da die zweite Zeile Differentialoperatoren enthält, die auf die entsprechenden Elemente der dritten Zeile wirken.

Alternativ können wir das allgemeine Ergebnis auch mit Hilfe des Levi-Cevitá-Symbols ϵ_{ijk} (vgl. Auf. 7) schreiben als

$$(\operatorname{rot} A(u_1, u_2, u_3))_{u_i} = \sum_{j,k} \epsilon_{ijk} h_{u_i} \frac{\partial(h_{u_k} A_{u_k})}{\partial u_j}.$$

Dabei ist dann $\operatorname{rot} A = \sum_i (\operatorname{rot} A)_{u_i} \underline{e}_{u_i}$.

Wir spezialisieren dieses allgemeine Resultat wieder auf die wichtigsten Fälle:

<p>kartesisch: $(\operatorname{rot} \underline{A}(x_1, x_2, x_3))_i = \sum_{j,k} \epsilon_{ijk} \frac{\partial A_k}{\partial x_j},$</p> <p>Zylinder: $\operatorname{rot} \underline{A}(\rho, \varphi, x_3) = \left(\frac{1}{\rho} \frac{\partial A_{x_3}}{\partial \varphi} - \frac{\partial A_\varphi}{\partial x_3} \right) \underline{e}_\rho + \left(\frac{\partial A_\rho}{\partial x_3} - \frac{\partial A_{x_3}}{\partial \rho} \right) \underline{e}_\varphi$ $+ \frac{1}{\rho} \left(\frac{\partial(\rho A_\varphi)}{\partial \rho} - \frac{\partial A_\rho}{\partial \varphi} \right) \underline{e}_{x_3},$</p> <p>Kugel: $\operatorname{rot} \underline{A}(r, \vartheta, \varphi) = \frac{1}{r \sin \vartheta} \left(\frac{\partial(\sin \vartheta A_\varphi)}{\partial \vartheta} - \frac{\partial A_\vartheta}{\partial \varphi} \right) \underline{e}_r + \left(\frac{1}{r \sin \vartheta} \frac{\partial A_r}{\partial \varphi} - \frac{1}{r} \frac{\partial(r A_\varphi)}{\partial r} \right) \underline{e}_\vartheta$ $+ \frac{1}{r} \left(\frac{\partial(r A_\vartheta)}{\partial r} - \frac{1}{r} \frac{\partial A_r}{\partial \vartheta} \right) \underline{e}_\varphi.$</p>

Dabei bezeichnet wieder ϵ_{ijk} das Levi-Cevitá-Symbol.

11.3.4 Laplace-Operator

Abschließend wollen wir noch die Ergebnisse für den Laplace-Operator ohne Rechnung angeben. Im Prinzip kann man sie aus den bereits abgeleiteten Darstellungen des Gradienten und der Divergenz über $\Delta \phi = \operatorname{div}(\operatorname{grad} \phi)$ gewinnen.

Das allgemeine Resultat ist

$$\begin{aligned} \Delta \phi(u_1, u_2, u_3) &= \operatorname{div}(\operatorname{grad} \phi(u_1, u_2, u_3)) \\ &= \frac{1}{h_{u_1} h_{u_2} h_{u_3}} \left[\frac{1}{\partial u_1} \left(\frac{h_{u_2} h_{u_3}}{h_{u_1}} \frac{\partial \phi}{\partial u_1} \right) + \frac{1}{\partial u_2} \left(\frac{h_{u_1} h_{u_3}}{h_{u_2}} \frac{\partial \phi}{\partial u_2} \right) + \frac{1}{\partial u_3} \left(\frac{h_{u_1} h_{u_2}}{h_{u_3}} \frac{\partial \phi}{\partial u_3} \right) \right]. \end{aligned}$$

Under construction !

Abbildung 11.4.1: Parametrisierung einer Raumkurve durch die Bogenlänge.

Speziell für die drei wichtigsten Koordinatensysteme bedeutet dies:

$$\text{kartesisch: } \Delta\phi(x_1, x_2, x_3) = \sum_j \frac{\partial^2 \phi}{\partial x_j^2},$$

$$\text{Zylinder: } \Delta\phi(\rho, \varphi, x_3) = \left[\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial x_3^2} \right] \phi(\rho, \varphi, x_3),$$

$$\text{Kugel: } \Delta\phi(r, \vartheta, \varphi) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial \phi}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 \phi}{\partial \varphi^2}.$$

11.4 Bogen-, Flächen-, Volumenelemente

11.4.1 Raumkurven

Wir betrachten eine Raumkurve, die durch die Parametrisierung $\underline{r}(t)$ beschrieben wird. Beispiele hatten wir schon im Abschnitt über Wegintegrale kennengelernt.

Eine natürliche Parametrisierung ist durch die Bogenlänge gegeben, d.h. dem zurückgelegten Weg entlang der Kurve. Dieser wird relativ zu einem festen Bezugspunkt $\underline{r}_0 = \underline{r}(t=0)$ gemessen (Abb. 11.4.1).

Die infinitesimale Bogenlänge ds können wir einfach in kartesischen Koordinaten ausdrücken:

$$(ds)^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2,$$

woraus folgt

$$s(t) = \int_0^t ds(t') = \int_0^t dt' \sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2 + \left(\frac{dx_3}{dt}\right)^2}.$$

Nach Definition der Geschwindigkeit gilt außerdem für ihren Betrag

$$v(t) = \frac{ds}{dt}.$$

Der (normierte) **Tangentenvektor** an die Kurve ist

$$\underline{T}(t) = \frac{\underline{v}(t)}{v(t)}$$

gegeben, was sich mit Hilfe der Bogenlänge in der Form

$$\underline{T} = \frac{d\underline{r}}{ds}$$

schreiben läßt.

Neben der Tangenten gibt es noch zwei weitere Vektoren, die die Raumkurve lokal charakterisieren. Der erste ist der **Hauptnormalenvektor** (oder **Krümmungsvektor**)

$$\underline{H} = \frac{1}{\kappa} \frac{d\underline{T}}{ds}$$

wobei κ durch die Forderung $|\underline{H}| = 1$ bestimmt ist. κ nennt man auch **Krümmung**. Man kann sie auch über

$$\kappa = \frac{|\dot{\underline{r}} \times \ddot{\underline{r}}|}{|\dot{\underline{r}}|}$$

bestimmen.

Aus der Normiertheit von \underline{T} folgt

$$0 = \frac{d}{ds} \underline{T}^2 = 2\underline{T} \cdot \frac{d\underline{T}}{ds}$$

d.h. der Hauptnormalenvektor steht senkrecht auf der Tangenten.

Um ein lokales Orthogonalsystem zu erhalten, benötigen wir noch einen Vektor, der senkrecht auf \underline{T} und \underline{H} steht. Diesen nennt man auch **Binormalenvektor**:

$$\underline{B} := \underline{T} \times \underline{H}.$$

Wir können die Bogenlänge auch in beliebigen Koordinaten u_j ausdrücken:

$$ds = \sqrt{(h_{u_1} du_1)^2 + (h_{u_2} du_2)^2 + (h_{u_3} du_3)^2}.$$

Speziell in Zylinderkoordinaten haben wir z.B.

$$(ds)^2 = (d\rho)^2 + \rho^2(d\varphi)^2 + (dz)^2.$$

11.4.2 Flächen- und Volumenelemente

Wir wollen nun noch die allgemeine Form von Flächen- und Volumenelementen in krummlinigen (orthogonalen) Koordinaten ableiten.

Zunächst betrachten wir eine Koordinatenfläche, die durch die Parametrisierung $\underline{r}(u_1, u_2, u_3^{(0)})$ gegeben sei. Dann ist das Flächenelement gegeben durch

$$dA(u_1, u_2) = |(h_{u_1} \underline{e}_{u_1}) \times (h_{u_2} \underline{e}_{u_2})| du_1 du_2.$$

Analoge Ergebnisse erhält man, wenn man andere Koordinaten als u_3 fixiert.

Ein Beispiel ist der Zylindermantel, der sich als Koordinatenfläche ergibt, wenn man $\rho = \rho_0$ fixiert. Dann gilt

$$dA(\varphi, x_3) = \rho_0 d\varphi dx_3.$$

Das Volumenelement ist allgemein durch

$$dV = |h_{u_1} h_{u_2} h_{u_3}| du_1 du_2 du_3.$$

gegeben. Hieraus erhält man dann für die Kugel- und Zylinderkoordinaten die in Kapitel 8.4 mit Hilfe des Transformationssatzes abgeleiteten Ergebnisse.

Kapitel 12

Operatoren und Eigenwerte

12.1 Eigenwerte von Matrizen

Definition 12.1.1 (Eigenwert, Eigenvektor).
Es sei $\underline{\underline{A}}$ eine quadratische Matrix. Gilt dann

$$\underline{\underline{A}} \cdot \underline{v} = \lambda \underline{v}$$

so heißt λ **Eigenwert** von $\underline{\underline{A}}$ und \underline{v} **Eigenvektor** (zum Eigenwert λ)¹. Dabei ist λ eine reelle oder komplexe Zahl.

Bem.: Die Eigenvektoren zum Eigenwert λ sind nicht eindeutig festgelegt, da mit \underline{v} auch jedes Vielfache $\alpha \underline{v}$ (mit $\alpha \neq 0$) Eigenvektor zum gleichen Eigenwert ist.

Die Menge aller Eigenwerte einer Matrix bezeichnet man als **Spektrum** der Matrix, die Menge aller Eigenwerte und Eigenvektoren als **Eigensystem**.

Beispiel 12.1.1. 1. Bei Drehungen gehen Punkte auf der Drehachse in sich über: $\underline{\underline{A}}\underline{v} = \underline{v}$.

2. Wir betrachten die Matrix $\underline{\underline{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Dann gilt

$$\underline{\underline{A}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \underline{\underline{A}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die Matrix $\underline{\underline{A}}$ hat also die Eigenwerte 1 und 2 mit den zugehörigen Eigenvektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
bzw. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Wir wollen uns nun mit der Frage beschäftigen, wie man die Eigenwerte und -vektoren konkret berechnen kann. Dies bezeichnet man als **Eigenwertproblem**.

¹Man beachte, dass man im Englischen von *eigenvalue* und *eigenvector* spricht!

Die Eigenwerte und -vektoren sind als Lösungen des linearen Gleichungssystems

$$(\underline{A} - \lambda \mathbb{1}) \underline{v} = 0$$

definiert. Man beachte, dass man hierbei auch das λ a priori nicht kennt! Das Gleichungssystem hat nicht-triviale Lösungen, falls

$$\det(\underline{A} - \lambda \mathbb{1}) = 0$$

ist. Ansonsten existiert nur die triviale Lösung $\underline{v} = \underline{0}$. Diese Gleichung bezeichnet man auch als **Säkulargleichung**. Ist \underline{A} eine $n \times n$ -Matrix, so liefert die Determinante ein Polynom n -ten Grades in λ , das sog. **charakteristische Polynom** $p(\lambda)$.

Nach dem Fundamentalsatz der Algebra hat ein Polynom n -ten Grades n Nullstellen, die aber komplex sein können. Wir bezeichnen die Nullstellen von $p(\lambda)$ mit $\lambda_1, \lambda_2, \dots, \lambda_n$. Sind einige der λ_j gleich, so spricht man von **Nullstellen höherer Ordnung** bzw. **Entartung**. Man charakterisiert diese durch die **algebraische Multiplizität** oder **algebraische Vielfachheit**. Ist z.B. $\lambda_1 = \lambda_2$, so hat der Eigenwert λ_1 die algebraische Multiplizität 2.

Sind die Eigenwerte λ_j bekannt, so kann man die zugehörigen Eigenvektoren \underline{v}_j als Lösungen des linearen Gleichungssystems

$$\underline{A} \cdot \underline{v}_j = \lambda_j \underline{v}_j$$

explizit bestimmen.

Beispiel 12.1.2. Als Beispiel betrachten wir die Matrix

$$\underline{A} = \begin{pmatrix} 10 & -3 \\ -3 & 2 \end{pmatrix}.$$

Zunächst bestimmen wir mit Hilfe der Säkulargleichung das charakteristische Polynom:

$$p(\lambda) := \det(\underline{A} - \lambda \mathbb{1}) = \det \begin{pmatrix} 10 - \lambda & -3 \\ -3 & 2 - \lambda \end{pmatrix} = (10 - \lambda)(2 - \lambda) - (-3)(-3) = \lambda^2 - 12\lambda + 11.$$

$p(\lambda)$ hat die Nullstellen $\lambda_1 = 1$ und $\lambda_2 = 11$, die also jeweils die algebraische Vielfachheit 1 haben und die Eigenwerte von \underline{A} sind.

Einen Eigenvektor zu λ_1 finden wir durch Lösung des linearen Systems

$$(\underline{A} - \lambda_1 \mathbb{1}) \underline{v}_1 = 0.$$

Explizit lautet dieses System

$$\begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 0.$$

Die Lösung ist $x_1 = t$ und $y_1 = 3t$ mit beliebigem $t \in \mathbb{R}$, $t \neq 0$. Somit haben wir die Eigenvektoren zum Eigenwert $\lambda_1 = 1$ bestimmt:

$$\underline{v}_1 = \begin{pmatrix} t \\ 3t \end{pmatrix}.$$

Bei der Bestimmung des Eigenvektors zum Eigenwert $\lambda_2 = 11$ gehen wir analog vor. Hier haben wir das System

$$(\underline{A} - \lambda_2 \mathbb{1}) \underline{v}_2 = \begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = 0$$

zu lösen. Wir erhalten $x_2 = t$ und $y_2 = -\frac{t}{3}$.

In der Regel gibt man die normierten Einheitsvektoren an, also hier

$$\underline{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \underline{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Man beachte, dass offensichtlich v_1 und v_2 senkrecht aufeinander stehen. Wir werden später Kriterien dafür angeben, wann dies der Fall ist.

Bemerkung. Allgemein sind die Eigenvektoren zu verschiedenen Eigenwerte linear unabhängig!

Bemerkung. Es gibt nicht immer gleich viele Eigenvektoren wie Eigenwerte. Zum Beispiel hat die Matrix

$$\underline{A} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$$

den entarteten Eigenwert $\lambda = -1$ (mit algebraischer Multiplizität 2), aber nur einen Eigenvektor $\underline{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Die Zahl der Eigenvektoren zu einem bestimmten Eigenwert nennt man seine **geometrische Multiplizität** oder **geometrische Vielfachheit**. Sie ist kleiner oder gleicher der algebraischen Multiplizität.

Wenn wir die Eigenwerte und Eigenvektoren einer Matrix \underline{A} kennen, können wir die Matrix unter bestimmten Umständen **diagonalisieren**, d.h. durch eine Ähnlichkeitstransformation² in Diagonalgestalt bringen. Dies entspricht einem Basiswechsel.

Wir betrachten hierzu eine $n \times n$ -Matrix \underline{A} , deren Eigenwerte λ_j und zugehörigen Eigenvektoren \underline{v}_j wir kennen:

$$\underline{A} \cdot \underline{v}_j = \lambda_j \underline{v}_j.$$

Wir nehmen nun an, dass alle Eigenwerte reell und die Eigenvektoren \underline{v}_j paarweise orthogonal sind. Außerdem sollen sie normiert sein, d.h. $|\underline{v}_j| = 1$.

Wir definieren nun die Matrix

$$\underline{U} := (\underline{v}_1, \dots, \underline{v}_n),$$

deren Spalten aus den Eigenvektoren gebildet werden, d.h. es gilt $U_{jl} = (\underline{v}_l)_j$.

Für diese Matrix gilt

$$\underline{U}^\dagger \cdot \underline{U} = \underline{U} \cdot \underline{U}^\dagger = \begin{pmatrix} \overline{\underline{v}_1} \cdot \underline{v}_1 & \overline{\underline{v}_1} \cdot \underline{v}_2 & \dots \\ \overline{\underline{v}_2} \cdot \underline{v}_1 & \overline{\underline{v}_2} \cdot \underline{v}_2 & \dots \\ \vdots & \vdots & \ddots \\ & & & \overline{\underline{v}_n} \cdot \underline{v}_n \end{pmatrix} = \mathbb{1},$$

²Das sind Transformationen der Form $\underline{M}^{-1} \cdot \underline{A} \cdot \underline{M}$.

wobei

$$\underline{\underline{U}}^\dagger := (\overline{\underline{\underline{U}}})^T,$$

die adjungierte Matrix (oder auch hermitesch konjugierte Matrix) ist³. Sie entsteht durch Bildung der Transposition der Matrix, die aus den komplex-konjugierten Elementen $\overline{U_{jl}}$ der Matrix $\underline{\underline{U}}$ besteht. Ist die Matrix reell, so ist die Adjungierte gleich der Transponierten.

Daher gilt für die oben definierte Matrix $\underline{\underline{U}}$, dass

$$\underline{\underline{U}}^{-1} = \underline{\underline{U}}^\dagger.$$

Eine solche Matrix nennt man unitär, siehe Kapitel 14.5.2.

Außerdem sieht man schnell ein, dass per Konstruktion von $\underline{\underline{U}}$ aus den Eigenvektoren von $\underline{\underline{A}}$ gilt:

$$\underline{\underline{A}} \cdot \underline{\underline{U}} = (\lambda_1 v_1, \dots, \lambda_n v_n) = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}}$$

mit der Diagonalmatrix

$$\underline{\underline{\Lambda}} := \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Somit gilt wegen der Unitarität von $\underline{\underline{U}}$

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} = \underline{\underline{\Lambda}} \quad \text{bzw.} \quad \underline{\underline{A}} = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{U}}^\dagger.$$

Man sagt auch: $\underline{\underline{U}}$ diagonalisiert $\underline{\underline{A}}$.

Komponentenweise lautet die letzte Gleichung

$$(\underline{\underline{A}})_{jm} = \sum_k \lambda_k (v_k)_j (\overline{v_k})_m.$$

Die Matrix $\underline{\underline{A}}$ läßt sich also mit Hilfe ihrer Eigenwerte und Eigenvektoren darstellen. Dies bezeichnet man als **Spektralzerlegung** oder **Spektraldarstellung**.

Speziell für $\underline{\underline{A}} = \mathbb{1}$ gilt dann

$$\sum_k (v_k)_j (\overline{v_k})_m = \delta_{jm}.$$

Dies bezeichnet man auch als **Zerlegung der Einheit** oder **Vollständigkeitsrelation**.

Bemerkung. Beachte, dass sich nicht alle Matrizen auf diese Art diagonalisieren lassen! Für eine spezielle Klasse von Matrizen läßt sich dies aber allgemein beweisen.

Definition 12.1.2 (Normale Matrizen).

Gilt

$$[\underline{\underline{A}}, \underline{\underline{A}}^\dagger] := \underline{\underline{A}} \cdot \underline{\underline{A}}^\dagger - \underline{\underline{A}}^\dagger \cdot \underline{\underline{A}} = 0,$$

so heißt die Matrix $\underline{\underline{A}}$ **normal**.

Dabei ist $[A, B] = AB - BA$ der sogenannte **Kommutator**.

Insbesondere sind also hermitesche ($\underline{\underline{A}}^\dagger = \underline{\underline{A}}$) und symmetrische ($\underline{\underline{A}}$ reell mit $\underline{\underline{A}}^T = \underline{\underline{A}}$) Matrizen normal.

³Siehe später in Kapitel 14.5.2.

Es gilt: Normale Matrizen sind unitär diagonalisierbar (d.h. $\underline{U}^\dagger \cdot \underline{A} \cdot \underline{U} = \underline{\Lambda}$) und haben eine Spektraldarstellung.

Bemerkung. Die Diagonalisierung entspricht dem Übergang zu einem neuen Koordinatensystem, d.h. einem Basiswechsel. Es gilt

$$\underline{U}^\dagger \cdot \underline{A} \cdot \underline{U} (\underline{U}^\dagger \underline{v}) = \lambda (\underline{U}^\dagger \underline{v}),$$

also mit $\underline{U}^\dagger \cdot \underline{A} \cdot \underline{U} = \underline{\Lambda}$ und der Abkürzung $\tilde{v} := \underline{U}^\dagger \underline{v}$

$$\underline{\Lambda} \cdot \tilde{v} = \lambda \tilde{v}.$$

In der Basis der Eigenvektoren \tilde{v} hat \underline{A} also Diagonalform!

Da sich die Spur und die Determinante bei Basiswechseln nicht ändern, gilt

$$\begin{aligned} \text{Spur } \underline{A} &= \sum_{j=1}^n \lambda_j = \sum_j \lambda_j p_j, \\ \det \underline{A} &= \prod_{j=1}^n \lambda_j = \prod_j \lambda_j^{p_j}. \end{aligned}$$

Dabei ist p_j die algebraische Vielfachheit des Eigenwerts⁴ λ_j . Die **Spur** einer $n \times n$ -Matrix \underline{A} ist definiert durch

$$\text{Spur } \underline{A} := \sum_{j=1}^n A_{jj},$$

d.h. als Summe der Diagonalelemente.

Bemerkung. Die Eigenwerte der Matrix \underline{A} und ihrer Transponierten \underline{A}^T stimmen überein. Dies folgt aus der allgemeingültigen Identität $\det \underline{A}^T = \det \underline{A}$, die die Gleichheit der charakteristischen Polynome impliziert. I.a. stimmen aber die zugehörigen Eigenvektoren nicht überein! Dies führt zu folgender Definition:

Definition 12.1.3 (Rechte und linke Eigenvektoren).

Die bisher betrachteten Eigenvektoren \underline{v} sind **rechte Eigenvektoren** für die gilt

$$\underline{A} \cdot \underline{v} = \lambda \underline{v}.$$

In Analogie hierzu definiert man **linke Eigenvektoren** \tilde{v} als die Eigenvektoren von \underline{A}^T :

$$\underline{A}^T \tilde{v} = \lambda \tilde{v}.$$

Die Bezeichnung erklärt sich aus folgender Identität:

$$\tilde{v}^T \underline{A} = (\underline{A}^T \cdot \tilde{v})^T = (\lambda \tilde{v})^T = \lambda \tilde{v}^T,$$

⁴In der entsprechenden Darstellung ist nur noch über die *unterschiedlichen* Eigenwerte zu summieren bzw. zu multiplizieren!

d.h. ein linker Eigenvektor wird bei Multiplikation von links an die Matrix bis auf einen Faktor reproduziert.

Für normale Matrizen, also insbesondere für hermitesche und symmetrische Matrizen, stimmen rechter und linker Eigenvektor überein.

12.1.1 Theoreme zum Eigenwertproblem

Wir geben im folgenden eine Sammlung wichtiger und nützlicher Aussagen zum Eigenwertproblem an, die wir aber nicht beweisen wollen.

Satz 12.1.1 (Eigenwertprobleme).

1. Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.
2. **Hermitesche Matrizen**, d.h. Matrizen für die $\underline{A}^\dagger = \underline{A}$ gilt, haben *reelle* Eigenwerte und die (normierten) Eigenvektoren bilden eine *Orthonormalbasis*. Sie sind mit Hilfe der unitären Matrix \underline{U} der Eigenvektoren diagonalisierbar und haben eine Spektraldarstellung.
3. Eine analoge Aussage gilt für **symmetrische Matrizen**, d.h. alle Elemente A_{jl} sind reell und es gilt $\underline{A}^T = \underline{A}$. In diesem Fall kann \underline{A} mit Hilfe einer **orthogonalen Matrix** \underline{U} diagonalisiert werden, d.h. $\underline{U}^{-1} = \underline{U}^T$.
4. Kommutierende, hermitesche Matrizen $\underline{A}, \underline{B}$ (mit $[\underline{A}, \underline{B}] = 0$) haben ein gemeinsames Eigenvektorensystem $\underline{v}_1, \dots, \underline{v}_n$, d.h. es gilt

$$\underline{A} \cdot \underline{v}_j = \lambda_j \underline{v}_j \quad \text{und} \quad \underline{B} \cdot \underline{v}_j = \tilde{\lambda}_j \underline{v}_j.$$

Man beachte, dass die Eigenwerte λ_j und $\tilde{\lambda}_j$ aber i.a. verschieden sind!

5. Die (normierten) Eigenvektoren einer normalen Matrix bilden eine Orthonormalbasis. Sie sind wie hermitesche Matrizen unitär diagonalisierbar und haben eine Spektraldarstellung.
6. Alle normalen Matrizen sind diagonalisierbar, aber nicht alle diagonalisierbaren Matrizen sind normal.
7. Nicht alle diagonalisierbaren Matrizen haben eine Spektraldarstellung.

12.2 Operatoren

Wir betrachten zwei Vektorräume V und W . Allgemein versteht man unter einem **Operator** \hat{A} eine Abbildung⁵ $\hat{A} : V \rightarrow W$. Im folgenden werden wir, sofern nicht anders angegeben oder aus

⁵Operatoren kennzeichnen wir oft durch ein $\hat{}$, z.B. \hat{A} .

dem Zusammenhang offensichtlich, vor allem den Fall $V = W$ behandeln. Außerdem wollen wir vor allem **lineare Operatoren** betrachten, die durch

$$\hat{A}(x + \alpha y) = \hat{A}x + \alpha \hat{A}y \quad \text{für alle } x, y \in V \text{ und } \alpha \in \mathbb{R} \text{ oder } \mathbb{C}$$

charakterisiert sind. Hierbei haben wir schon eine häufig verwendete Konvention benutzt, nämlich die, dass man bei Operatoren oft $\hat{A}x$ statt $\hat{A}(x)$ schreibt. Dies kennen natürlich schon von Differentialoperatoren wie $\frac{d}{dx}$ oder grad .

Im folgenden werden wir die Notation für den Fall eines komplexen Vektorraumes benutzen. Wenn nicht anders angegeben, gelten analoge Ergebnisse für den reellen Fall, man muß lediglich die komplexe Konjugation weglassen!

Bisher haben wir den Fall $V = \mathbb{R}^n$ betrachtet, wo wir \hat{A} durch eine Matrix \underline{A} darstellen können. Die für Matrizen angestellten Überlegungen lassen sich dann übertragen, wenn wir ein Skalarprodukt (x, y) (mit $x, y \in V$) zur Verfügung haben (siehe Kapitel 10.1.2) und folgende Identifikationen benutzen:

$$\underline{x} \cdot \underline{y} \leftrightarrow (x, y), \quad \underline{x} \cdot \underline{A} \cdot \underline{y} \leftrightarrow (x, \hat{A}y), \quad \text{etc.}$$

Dies werden wir gleich weiter ausarbeiten. Zunächst wollen wir Beispiele für lineare Operatoren kennenlernen.

Beispiel 12.2.1.

1. Das Skalarprodukt mit einem festen $x_0 \in V$ definiert einen linearen Operator: $\hat{A}x := (x_0, x)$. In diesem Fall ist $W = \mathbb{R} \neq V$.
2. Die Ableitung $\hat{A}f = f'$ ist ein linearer Operator.

Definition 12.2.1 (adjungierte und hermitesche Operatoren).

Ist $\hat{A} : V \rightarrow V$ ein linearer Operator, so bezeichnet man den Operator $\hat{A}^\dagger : V \rightarrow V$ charakterisiert durch

$$(x, \hat{A}y) = (\hat{A}^\dagger x, y) \quad \text{für alle } x, y \in V$$

als den **adjungierten Operator** oder auch **hermitesch konjugierten Operator**.

Ist $\hat{A}^\dagger = \hat{A}$, so nennt man \hat{A} auch **hermitesch** oder **selbstadjungiert**⁶.

Wir können nun Darstellungen des Operators \hat{A} angeben. Dazu sei $\{\varphi_k\}$ eine Orthonormalbasis von V . Beliebige Elemente $x \in V$ lassen sich daher als Linearkombinationen $x = \sum_k c_k \varphi_k$ darstellen. Wir können hiermit z.B. die Identität $\hat{A}x = y$ in eine Matrixgleichung übersetzen:

$$\sum_k \tilde{c}_k \varphi_k = y = \hat{A}x = \hat{A} \sum_k c_k \varphi_k = \sum_k c_k \hat{A} \varphi_k.$$

⁶In der Mathematik gibt es einen winzigen Unterschied zwischen *hermitesch* und *selbstadjungiert*, der in der Physik aber in der Regel unwichtig ist.

Nun bilden wir das Skalarprodukt dieser Identität mit φ_j :

$$\tilde{c}_j = \sum_k \tilde{c}_k \delta_{jk} = \sum_k \tilde{c}_k (\varphi_j, \varphi_k) = \sum_k c_k (\varphi_j, \hat{A}\varphi_k) =: \sum_k c_k A_{jk}$$

wobei wir die Abkürzung $A_{jk} := (\varphi_j, \hat{A}\varphi_k)$ eingeführt haben. Somit läßt sich die Identität nach Wahl einer Basis auch in Matrixform darstellen als

$$\sum_k A_{jk} c_k = \tilde{c}_j \quad \text{bzw.} \quad \underline{A} \cdot \underline{c} = \underline{\tilde{c}},$$

wobei die Elemente A_{jk} zu einer Matrix zusammengefasst und die c_k, \tilde{c}_j als Komponenten eines (Spalten-)Vektors aufgefasst haben. Man beachte, dass Matrix und Vektoren im Prinzip unendlich sein können, wenn es sich bei V nicht um einen endlichdimensionalen Vektorraum handelt. Weiterhin gilt:

$$(x, \hat{A}y) = \left(\sum_j c_j \varphi_j, \hat{A} \sum_l c'_l \varphi_l \right) = \sum_{j,l} \bar{c}_j c'_l (\varphi_j, \hat{A}\varphi_l) = \sum_{j,l} \bar{c}_j A_{jl} c'_l.$$

Für $x = y$ nennt man dies auch **Erwartungswert** von \hat{A} .

Es sei nun $\{\psi_k\}$ eine andere Orthonormalbasis von V . Der Basis- oder Darstellungswechsel $\{\varphi_k\} \rightarrow \{\psi_k\}$ wird dann durch einen unitären Operator \hat{U} vermittelt:

$$\psi_j = \hat{U}\varphi_j \quad \text{mit} \quad \hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}$$

mit der Matrixdarstellung

$$U_{jl} = (\varphi_j, \hat{U}\varphi_l) = (\varphi_j, \psi_l).$$

Es gilt wegen $\hat{U}^\dagger \hat{U} = \mathbb{1}$ für alle $x, y \in V$:

$$(x, y) = (x, \hat{U}^\dagger \hat{U}y) = (\hat{U}x, \hat{U}y).$$

Skalarprodukte bleiben daher bei einem solchen Basiswechsel unverändert (*invariant*), weshalb man ihn als Drehung im Vektorraum interpretieren kann.

In Komponenten gilt:

$$\begin{aligned} (\psi_k, x) &= \sum_l (\psi_k, \varphi_l) = \sum_l (U^\dagger)_{kl} c_l = \sum_l (U^\dagger)_{kl} (\varphi_l, x), \\ (\psi_j, \hat{A}\psi_l) &= (U^\dagger \hat{A} U)_{jl} = \sum_{k,m} (U^\dagger)_{jk} (\varphi_k, \hat{A}\varphi_m) U_{ml}. \end{aligned}$$

12.3 Eigenwertproblem für Operatoren

Analog zum Fall von Matrizen können wir auch bei Operatoren Eigenwerte und Eigenvektoren suchen. Das entsprechende Eigenwertproblem lautet dann

$$\hat{A}f = \lambda f.$$

Im folgenden werden wir uns vorstellen, dass der Vektorraum aus Funktionen besteht. Dies haben wir hier schon durch die Notation 'f' statt 'x' für den Eigenvektor angedeutet. Natürlich gelten aber alle folgenden Aussagen für beliebige zugrundeliegende Vektorräume.

Beispiel 12.3.1. Ein typisches Beispiel ist der Differentialoperator

$$\hat{A} = \frac{d^2}{dx^2}.$$

Dieser hat die Eigenvektoren $f_k := \sin(kx)$ mit den zugehörigen Eigenwerten $\lambda_k = -k^2$.

Im vorigen Kapitel haben wir gesehen, dass wir ein Operatorproblem in ein Matrixproblem übersetzen können. Daher ist es nicht überraschend, dass sich die Aussagen für das Eigenwertproblem von Matrizen auf Operatoren übertragen lassen. Insbesondere gilt:

1. Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.
2. Normale Operatoren ($[\hat{A}, \hat{A}^\dagger] = 0$, d.h. $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$) besitzen eine Orthonormalbasis $\{\varphi_j\}$ aus Eigenvektoren: $\hat{A}\varphi_j = \lambda_j\varphi_j$.
3. Selbstdjungierte Operatoren haben reelle Eigenwerte:

$$\lambda_j = (\varphi_j, \lambda_j\varphi_j) = (\varphi_j, \hat{A}\varphi_j) = (\hat{A}\varphi_j, \varphi_j) = (\lambda_j\varphi_j, \varphi_j) = \overline{\lambda_j} (\varphi_j, \varphi_j) = \overline{\lambda_j}.$$

Dabei haben wir, neben der Tatsache dass $\{\varphi_j\}$ eine Orthonormalbasis aus Eigenvektoren ist, im dritten Schritt die Selbstdjungiertheit von \hat{A} ausgenutzt.

4. Unitäre Operatoren beschreiben Darstellungswechsel. Im Eigenvektorsystem wird der Operator diagonalisiert:

$$(\varphi_j, \hat{A}\varphi_l) = \lambda_l(\varphi_j, \varphi_l) = \lambda_l\delta_{jl}.$$

Der Operator transformiert sich daher wie

$$\underbrace{(\hat{U}^\dagger \hat{A} \hat{U})}_{=\hat{\Lambda}} (\hat{U}^\dagger f) = \lambda (\hat{U}^\dagger f)$$

wobei $\hat{\Lambda}$ diagonal ist.

5. Hermitesche, miteinander kommutierende Operatoren haben ein gemeinsames Eigenvektorsystem.

Bem.: Dies wird in der Quantenmechanik wichtig, wo kommutierende Operatoren Größen entsprechen, die man gleichzeitig "scharf" messen kann.

Es sei \hat{A} ein hermitescher Operator und $f = \sum_j c_j \varphi_j$ ein beliebiger Vektor, wobei $\{\varphi\}$ eine Orthonormalbasis aus Eigenvektoren von \hat{A} ist. Dann ist der **Erwartungswert** von \hat{A} (bzgl. f) gegeben durch

$$\begin{aligned} (f, \hat{A}f) &= \sum_{j,l} \bar{c}_j c_l (\varphi_j, \hat{A}\varphi_l) = \sum_{j,l} \bar{c}_j c_l \lambda_l (\varphi_j, \varphi_l) = \sum_{j,l} \bar{c}_j c_l \lambda_l \delta_{jl} \\ &= \sum_j \lambda_j |c_j|^2. \end{aligned}$$

Es tragen also i.a. alle Eigenvektoren und Eigenwerte zum Erwartungswert bei!

12.4 Operatoren in der Quantenmechanik

Wir haben in Kapitel 10.1.2 schon den Begriff des Hilbertraumes kennengelernt. Hierbei handelt es sich um einen Vektorraum, auf dem ein Skalarprodukt definiert ist und der darüberhinaus vollständig ist.

In der mathematischen Beschreibung der Quantenmechanik wird ein *quantenmechanischer Zustand* mit einem Vektor eines geeigneten Hilbertraumes identifiziert. *Meßgrößen (Observablen)* entsprechen dann selbstadjungierten linearen Operatoren in diesem Hilbertraum. Die möglichen *Meßwerte* sind dann die Eigenwerte des entsprechenden Operators.

Definition 12.4.1 (Dirac-Notation).

Dirac hat eine sehr elegante und zweckmässige Notation, die sog. **Dirac-Notation**, erfunden, mit der sich besonders elegant rechnen läßt.

Vektoren im Hilbertraum schreibt man dabei als $|a\rangle$, den zu ihm hermitesch adjungierten Vektor, den man auch als **dualen Vektor** bezeichnet als $\langle a| := (|a\rangle)^\dagger$. Im \mathbb{R}^n entspricht $|a\rangle$ einem Spaltenvektor und $\langle a|$ dem zugehörigen Zeilenvektor⁷.

Man bezeichnet $\langle a|$ auch als **Bra(-Vektor)** und $|a\rangle$ als **Ket(-Vektor)**. Zusammengenommen ergibt dies (fast) das englische Wort *bracket* für Klammer.

In Dirac-Notation lautet das Eigenwertproblem für eine Operator \hat{A}

$$\hat{A}|a\rangle = a|a\rangle.$$

Hierbei haben wir schon die Standardkonvention benutzt, dass man den Eigenwert mit dem gleichen Buchstaben wie den Vektor bezeichnet. Auf Grund der Notation ist hier eine Verwechslung fast ausgeschlossen.

Das Skalarprodukt zweier Vektoren ist durch

$$\langle a|b\rangle$$

⁷Dies sollte man im folgenden immer im Hinterkopf behalten!

gegeben. Dies ist analog zum Standardskalarprodukt im \mathbb{R}^n . In Vektorschreibweise hatten wir dieses immer als $\underline{a} \cdot \underline{b}$ geschrieben. Es gilt aber

$$\underline{a} \cdot \underline{b} = \sum_{j=1}^n a_j b_j = \underline{a}^T \underline{b}$$

wobei wir die rechte Seite als Matrixmultiplikation der $1 \times n$ -Matrix \underline{a}^T mit der $n \times 1$ -Matrix \underline{b} zu interpretieren haben.

Mit dem Skalarprodukt können wir auch die Norm angeben:

$$||\underline{a}||^2 := \langle a|a \rangle.$$

Der Erwartungswert im Zustand $|a\rangle$ ist durch $\langle a|\hat{A}|a\rangle$ gegeben, falls $|a\rangle$ normiert ist.

Der Bra-Vektor $\langle a|$ entspricht der *abstrakten* Darstellung des quantenmechanischen Zustandes. Wählt man eine geeignete Basis, so kommt man z.B. zu den sog. *Wellenfunktionen*.

Es sei nun wieder \hat{A} ein selbstadjungierter Operator mit Eigensystem $\hat{A}|a_n\rangle = a_n|a_n\rangle$. Ein beliebiger Vektor (Zustand) läßt sich dann schreiben als

$$|f\rangle = \sum_n f_n |a_n\rangle \quad \text{mit} \quad f_n = \langle a_n|f\rangle.$$

Die Darstellung der Entwicklungskoeffizienten f_n folgt wie früher, da $\{|a_n\rangle\}$ eine Orthonormalbasis ist.

Wir können jetzt die Spektraldarstellung von \hat{A} in sehr eleganter Weise angeben:

$$\hat{A} = \sum_n a_n |a_n\rangle \langle a_n|.$$

Dies sieht man folgendermaßen. Für eine beliebiges $|f\rangle = \sum_n f_n |a_n\rangle$ gilt:

$$\begin{aligned} \left(\sum_n a_n |a_n\rangle \langle a_n| \right) |f\rangle &= \sum_n a_n |a_n\rangle \underbrace{\langle a_n|f\rangle}_{=f_n} = \sum_n f_n a_n |a_n\rangle = \sum_n f_n \hat{A} |a_n\rangle \\ &= \hat{A} |f\rangle. \end{aligned}$$

Diese Rechnung wesentlichen Prinzipien des Arbeitens mit der Dirac-Notation. Zunächst haben wir alle Bra's und Ket's gewissermaßen ausmultipliziert, unter Beachtung der Reihenfolge der Faktoren. Danach haben wir nach auftretenden Größen der Form $\langle b|a\rangle$ gesucht und diese als Skalarprodukte interpretiert (und berechnet).

Speziell für den Einheitsoperator $\mathbb{1}$ erhalten wir wieder die **Zerlegung der Einheit** oder **Vollständigkeitsrelation**:

$$\mathbb{1} = \sum_n |a_n\rangle \langle a_n|.$$

Kapitel 13

Differentialgleichungen II

In diesem Kapitel wollen wir uns weiter mit Differentialgleichungen beschäftigen. In Kapitel 6 hatten wir die Grundlagen der Theorie der gewöhnlichen DGL kennengelernt. Zunächst wollen wir diese Thema etwas vertiefen und uns mit Systemen von DGL und einigen speziellen DGL der Physik genauer beschäftigen. Zum Abschluß werden wir dann partielle Differentialgleichungen betrachten.

13.1 Systeme von Differentialgleichungen

In der Physik hat man es sehr oft mit Systemen von DGL zu tun, d.h. mehreren DGL, die untereinander gekoppelt sind. So etwas passiert z.B. immer dann, wenn man Systeme aus mehreren Teilchen betrachtet, die miteinander wechselwirken (d.h. Kräfte aufeinander ausüben). Über diese Wechselwirkungskräfte sind die Newtonschen Bewegungsgleichungen der Teilchen dann miteinander verkoppelt.

Es gibt aber auch andere Gründe, warum man an Systemen von DGL interessiert sein kann. Als weitere Motivation betrachten wir die folgende DGL 3. Ordnung:

$$y''' - 3y'' - 9y' - 5y = 0.$$

Wir definieren nun neue Funktionen durch

$$y_1(x) := y(x), \quad y_2(x) := y'(x), \quad y_3(x) := y''(x).$$

Für dies gilt dann offensichtlich folgendes System von DGL:

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= y_3, \\ y_3' &= 3y_3 + 9y_2 + 5y_1, \end{aligned}$$

wobei wir die letzte Gleichung aus der ursprünglichen DGL 3. Ordnung erhalten, wenn wir sie nach $y''' = y_3'$ auflösen und die auftretenden Ableitungen durch die neu definierten Funktionen ersetzen.

In Matrixform können wir dieses System kompakt schreiben als

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & 9 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Dies ist jetzt ein DGL-System, allerdings ist es von 1. Ordnung, höhere Ableitungen treten nicht mehr auf.

Allgemein kann man mit diesem Trick eine DGL n-ter Ordnung in n gekoppelte DGL 1. Ordnung verwandeln.

Ein allgemeines Differentialgleichungssystem (1. Ordnung) hat die Form

$$\underline{y}' = \underline{f}(\underline{y}, x) \quad \text{mit} \quad \underline{y}(x) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Hinzu kommt u.U. noch eine Anfangsbedingung $\underline{y}(x_0) = \underline{y}_0$.

Ist $\underline{f}(\underline{y}, x) = \underline{f}(\underline{y})$, so spricht man von einem **autonomen System**.

Im Fall $n = 1$ reduziert sich dies auf $y' = f(y, x)$, die allgemeine DGL 1. Ordnung wie wir sie bereits in Kapitel 6 kennengelernt haben.

13.1.1 Lineare Differentialgleichungssysteme

Zunächst betrachten wir den Fall eines autonomen Systems, bei dem die Funktion $\underline{f}(\underline{y})$ linear ist. Dann können wir das System auch als Matrixgleichung der Form

$$\underline{y}' = \underline{A} \cdot \underline{y} + \underline{B}$$

mit $n \times n$ -Matrizen \underline{A} und \underline{B} schreiben, deren Elemente Konstanten sind.

Zunächst betrachten wir wieder den homogenen Fall

$$\underline{y}' = \underline{A} \cdot \underline{y}$$

und machen analog zum Fall $n = 1$ den Ansatz

$$\underline{y}_j(x) = e^{\lambda_j x} \underline{u}_j.$$

Dann ist $\underline{y}_j' = \lambda_j \underline{y}_j$ und somit

$$\lambda_j \underline{y}_j = \underline{A} \cdot \underline{y}_j.$$

Dies ist eine Eigenwertgleichung! Somit liefert unser Ansatz eine Lösung des Systems, falls λ_j ein Eigenwert von \underline{A} ist und \underline{u}_j der zugehörige Eigenvektor.

Gibt es n unterschiedliche Eigenwerte und linear unabhängige Eigenvektoren, so lautet die allgemeine Lösung

$$\underline{y}(x) = \sum_{j=1}^n c_j \underline{y}_j(x)$$

mit Konstanten c_j , die durch die Anfangsbedingung $\underline{y}(x_0) = \underline{y}_0$ bestimmt sind.

Beispiel 13.1.1. Als ein Beispiel betrachten wir die Schwingungsgleichung

$$\ddot{y} + \omega^2 y = 0.$$

Mit $y_1 := y$ und $y_2 = \dot{y}$ erhält man das System

$$\dot{\underline{y}} = \underline{A} \cdot \underline{y} \quad \text{mit} \quad \underline{A} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}.$$

Die Eigenwerte von \underline{A} sind

$$\lambda_{1,2} = \pm i\omega \quad \text{mit Eigenvektor} \quad \underline{u}_{1,2} = \begin{pmatrix} 1 \\ \pm i\omega \end{pmatrix}.$$

Damit lautet die allgemeine Lösung

$$\underline{y}(t) = c_1 \begin{pmatrix} 1 \\ i\omega \end{pmatrix} e^{i\omega t} + c_2 \begin{pmatrix} 1 \\ -i\omega \end{pmatrix} e^{-i\omega t}.$$

Die erste Komponente entspricht der Lösung, die wir früher schon bestimmt hatten. Die zweite ist deren Ableitung, entspricht also der Geschwindigkeit, wenn man $y_1(t)$ als Auslenkung interpretiert.

Bemerkung. Wenn zwei Eigenwerte $\lambda_j = \lambda_l$ übereinstimmen, also entartet sind, und nur ein Eigenvektor \underline{u}_j existiert, dann macht man den Ansatz

$$\underline{y}_l(x) = x \underline{u}_l(x) + e^{\lambda_j x} \underline{v} = e^{\lambda_j x} (x \underline{u}_l + \underline{v}).$$

Eine kurze Rechnung liefert dann folgende Bedingung an den noch unbekanntem Vektor \underline{v} :

$$(\underline{A} - \lambda_j \mathbb{1}) \cdot \underline{v} = \underline{u}_j.$$

Diese Gleichung hat eine nicht-triviale Lösung für \underline{v} (und damit \underline{y}_j), da ja $\det(\underline{A} - \lambda_j \mathbb{1}) = 0$ ist, da λ_j Eigenwert von \underline{A} ist.

Bei höheren Entartungen macht man dann den Ansatz

$$\underline{y}_k(x) = e^{\lambda_j x} (x^2 \underline{u}_j + x \underline{v} + \underline{w})$$

etc.

Um die Lösung des inhomogenen Falls ($\underline{B} \neq 0$) zu erhalten, geht man wie früher im Fall $n = 1$ vor. Man bestimmt eine *spezielle* Lösung des inhomogenen Systems, z.B. durch Raten, Variation der Konstanten etc. Die allgemeine Lösung erhält man dann als Summe der speziellen und der allgemeinen Lösung des homogenen Problems (wie oben angegeben).

13.1.2 Dynamische Systeme

Wir betrachten nun ein allgemeines DGLsystem $\dot{\underline{y}} = \underline{f}(\underline{y}, t)$ und wollen die unabhängige Variable t als ‘Zeit’ interpretieren. Dabei sei die Funktion $\underline{f}(\underline{y}, t)$ nicht notwendig linear. Ein solches Problem bezeichnet man auch als **dynamisches System**. Die Lösungen im nichtlinearen Fall haben häufig interessante Eigenschaften (Fraktale etc.)!

Wir wollen uns hier nur mit dem autonomen Fall

$$\dot{\underline{y}} = \underline{f}(\underline{y})$$

beschäftigen. Wenn wir \underline{y} als den Ortsvektor eines Teilchens interpretieren, dann liefert die DGL die zu einer Position gehörige Geschwindigkeit. Dies macht die Bezeichnung “dynamisches System” plausibel.

Was passiert nun, falls an einem Punkt \underline{y}_0 gilt: $\underline{f}(\underline{y}_0) = 0$? Offensichtlich ist dann $\dot{\underline{y}} = 0$ und daher ändert sich \underline{y} nicht mehr, wenn das Teilchen einmal den Punkt \underline{y}_0 erreicht hat! Solche Punkte bezeichnet man daher auch als **Fixpunkte**. I.a. werden sie erst nach langer Zeit, d.h. im Grenzfall $t \rightarrow \infty$ erreicht.

Je nach Verhalten in der Nähe des Fixpunktes unterscheidet man

- **anziehende** oder **attraktive Fixpunkte**,
- **abstoßende** oder **repulsive Fixpunkte**.

Man spricht in diesem Zusammenhang auch von der **Stabilität** des Fixpunktes und nennt attraktive Fixpunkte **stabil** und repulsive Fixpunkte **instabil**. Über die Stabilität eines Fixpunktes erhalten wir Aufschluß durch eine Taylorentwicklung in seiner Nähe. Da $\underline{f}(\underline{y}_0) = 0$ ist, folgt:

$$\underline{f}(\underline{y}) \approx \underline{A} \cdot (\underline{y} - \underline{y}_0) \quad \text{mit} \quad A_{jl} = \left. \frac{\partial f_j}{\partial y_l} \right|_{\underline{y}=\underline{y}_0}.$$

Mit der Abkürzung $\underline{z} := \underline{y} - \underline{y}_0$ erhält man dann die Gleichung

$$\dot{\underline{z}} = \underline{A} \cdot \underline{z},$$

die das Verhalten des dynamischen Systems in der Nähe des Fixpunktes beschreibt. \underline{A} heißt **Stabilitätsmatrix**, ihre Eigenwerte charakterisieren das Verhalten nahe dem Fixpunkt.

Beispiel 13.1.2. Als Beispiel betrachten wir im Fall $n = 1$ die **logistische Gleichung**

$$\dot{y} = \alpha y(1 - y)$$

mit $\alpha > 0$, die z.B. in der Populationsdynamik auftritt. Sie enthält zwei konkurrierende Terme y und $-y^2$, die zum Wachstum bzw. Schrumpfen (Zerfall) von y führen. Ist y klein, so kann man y^2 gegenüber y vernachlässigen. Dann dominiert das Wachstum. Ist y aber groß, so dominiert der quadratische Term, d.h. der Zerfall.

Die Funktion $f(y) = \alpha y(1 - y)$ hat zwei Fixpunkte, die man direkt ablesen kann. Es sind $y_0 = 0$ und $y_1 = 1$. Führt man die oben beschriebene Entwicklung durch, so erhält man für y_0 :

$$A_0 = f'(y = 0) = \alpha(1 - 2y)|_{y=0} = \alpha$$

und somit die linearisierte Gleichung für $z = y - y_0 = y$

$$\dot{z} = A_0 z = \alpha z.$$

Deren Lösung ist

$$z_{\text{lin}}^{(0)}(t) = C e^{\alpha t}.$$

Analog erhält man für y_1 zunächst

$$A_1 = f'(y = 1) = \alpha(1 - 2y)|_{y=1} = -\alpha$$

und hieraus die linearisierte Gleichung für $z = y - y_1 = y - 1$

$$\dot{z} = A_1 z = -\alpha z,$$

deren Lösung durch

$$z_{\text{lin}}^{(1)}(t) = C e^{-\alpha t}.$$

gegeben ist.

Da $\alpha > 0$, wachsen Abweichungen vom Fixpunkt $y_0 = 0$ schnell (exponentiell!) an, während Abweichungen vom Fixpunkt $y_1 = 1$ schnell kleiner werden. Daher ist y_0 abstoßend, also instabil, und y_1 anziehend und somit stabil.

13.2 DGL als Eigenwertproblem

Differentialgleichungen haben manchmal die Form von Eigenwertproblemen. Deren typische Form ist

$$\hat{D}y(x) = \lambda y(x),$$

wobei \hat{D} ein Operator ist, der Ableitungen enthält. In der Regel fordert man noch zusätzliche Randbedingungen, so dass die DGL nicht unbedingt für alle Werte von λ eine Lösung hat. Existiert eine Lösung $y_\lambda(x)$, so heißt das entsprechende λ Eigenwert und $y_\lambda(x)$ Eigenfunktion oder **Eigenlösung** von \hat{D} .

Ein wichtiges Beispiel ist die Schwingungsgleichung, die sich für $\hat{D} = \frac{d^2}{dx^2}$ ergibt. Wir werden darauf am Ende dieses Abschnitts zurückkommen. Vorher wollen wir uns mit einem allgemeineren Problem beschäftigen.

Definition 13.2.1 (Sturm-Liouville-Problem).

Wir betrachten die Funktionen $q(x)$, $p(x)$ und $r(x)$, die jeweils im Intervall $[a, b]$ definiert sind. Dabei sei $q(x)$ stetig, $p(x) > 0$ und zweimal stetig differenzierbar und $r(x) > 0$ und stetig in $]a, b[$.

Das Eigenwertproblem

$$\hat{S}y(x) + \lambda r(x)y(x) = 0$$

bezeichnet man als **Sturm-Liouville-Problem**, wobei \hat{S} mit

$$\hat{S}y(x) := \frac{d}{dx} (p(x)y'(x)) + q(x)y(x)$$

als **Sturm-Liouville-Differentialoperator** bezeichnet wird.

Zusätzlich fordert man die sog. **Sturm-Liouville-Randbedingungen**. Hierbei soll für zwei beliebige, unterschiedliche Lösungen $u(x)$ und $v(x)$ des Sturm-Liouville-Problems gelten:

$$p(x) [u(x)v'(x) - u'(x)v(x)] \Big|_a^b = 0.$$

Dabei ist $\dots \Big|_a^b$ wie bei der Integration zu interpretieren als die Differenz der Funktionswerte an der Stelle b und a . Der Grund für diese recht ungewöhnliche Wahl der Randbedingungen wird gleich etwas klarer werden.

Man kann zeigen, dass das Sturm-Liouville-Problem unter den oben angegebenen Voraussetzungen (inklusive der Randbedingungen) reelle Eigenwerte λ_n hat mit zugehörigen Eigenlösungen $y_n(x)$. Diese erfüllen folgende Orthogonalitätsbedingung

$$(y_n, y_m) := \int_a^b dx r(x) y_n(x) y_m(x) = 0 \quad \text{für } n \neq m.$$

Dies kann man relativ einfach beweisen. Zunächst gilt, da $y_n(x)$ eine Lösung der Sturm-Liouville-DGL ist,

$$(p(x)y_n'(x))' = -[q(x) + \lambda_n r(x)] y_n(x).$$

Nun multiplizieren wir diese Gleichung mit $y_m(x)$ und integrieren dann über das Intervall $[a, b]$. Eine analoge Beziehung erhalten wir, wenn wir in obiger Vorgehensweise n und m vertauschen. Bildet man die Differenz dieser beiden Gleichungen, so folgt

$$p(x) [y_m(x)y_n'(x) - y_n(x)y_m'(x)] \Big|_a^b = (\lambda_m - \lambda_n) \int_a^b dx r(x) y_n(x) y_m(x).$$

Die linke Seite verschwindet auf Grund der Sturm-Liouville-Randbedingungen. Da $\lambda_m \neq \lambda_n$ ist, muß das Integral verschwinden. An dieser Stelle erkennen wir einen Grund für die ungewöhnliche Wahl der Randbedingungen.

Beispiel 13.2.1. Als ein wichtiges Beispiel wollen wir auf die Schwingungsgleichung zurückkommen. Hier ist offensichtlich

$$p(x) = 1, \quad q(x) = 0, \quad r(x) = 1.$$

Als Intervall wählen wir $[-\pi, \pi]$, d.h. $b = -a = \pi$. Mit der Wahl $y(\pm\pi) = 0$ sind offensichtlich die Sturm-Liouville-Randbedingungen erfüllt.

Die Lösungen dieses Sturm-Liouville-Problems sind durch

$$\lambda_n = n^2 \quad \text{und} \quad y_n(x) = A \sin(nx)$$

gegeben. Dies sieht man folgendermaßen: Ohne Berücksichtigung der Randbedingungen lautet die allgemeine Lösung für $\lambda \geq 0$

$$y(x) = A \sin(\sqrt{\lambda}x) + B \cos(\sqrt{\lambda}x).$$

Die Randbedingungen lassen sich nur erfüllen, falls $B = 0$ ist und $\sqrt{\lambda} = n \in \mathbb{N}$. Daher haben die Eigenwerte in diesem Fall die Form $\lambda_n = n^2$ mit $n \in \mathbb{N}$.

Die Eigenlösungen erfüllen die Orthogonalitätsbedingungen

$$(y_n, y_m) = \int_{-\pi}^{\pi} \sin(nx) \sin(mx) dx = 0 \quad \text{für } n \neq m,$$

was wir früher schon explizit gezeigt hatten.

13.3 Spezielle DGL

13.3.1 Legendre'sche DGL

Die Legendre'sche DGL lautet

$$(1 - x^2)y''(x) - 2xy'(x) + l(l + 1)y(x) = 0,$$

wobei $x \in [-1, 1]$ und $l \in \mathbb{Z}$ sein soll¹. Die Lösungen dieser Gleichung bezeichnet man als **Legendre-Polynome** $P_l(x)$ zu den Eigenwerten $l(l + 1)$.

Dies kann man explizit nachprüfen mit dem Potenzreihenansatz $y(x) = \sum_{n=0}^{\infty} a_n x^n$. Explizit lauten die ersten Legendre-Polynome

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, \quad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x.$$

Für $l < 0$, $l \in \mathbb{Z}$ gilt:

$$P_{-l-1} = P_l.$$

Die Legendre-Polynome genügen der Orthogonalitätsbeziehung

$$\int_{-1}^1 dx P_l(x) P_n(x) = \frac{1}{2l + 1} \delta_{ln}.$$

Man kann die Legendre-Polynome durch Ableitung aus einer sog. **erzeugenden Funktion** erhalten. Diese ist definiert durch

$$\phi(x, h) := \sum_{l=0}^{\infty} P_l(x) h^l.$$

¹Es gibt auch Lösung für $l \in \mathbb{R}$. Diese haben die Form von Potenzreihen und werden hier nicht betrachtet.

Damit gilt dann

$$P_n(x) = \frac{1}{n!} \frac{\partial^n}{\partial h^n} \phi(x, h) \Big|_{h=0}.$$

Man kann zeigen, dass $\phi(x, h)$ explizit gegeben ist durch

$$\phi(x, h) = \frac{1}{\sqrt{1 - 2xh + h^2}} \quad (|h| < 1).$$

Hieraus kann man dann durch Berechnung der n -ten Ableitung an der Stelle $h = 0$ das n -te Legendre-Polynom bestimmen.

Eine andere Möglichkeit zur Berechnung von $P_n(x)$ liefert die Rekursionsgleichung

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

Hiermit kann man aus $P_0(x) = 1$ und $P_1(x) = x$ rekursiv alle höheren Legendre-Polynome bestimmen.

Neben der Legendre-DGL tritt auch manchmal die **verallgemeinerte Legendresche DGL**

$$\frac{d}{dx} \left((1-x^2)y'(x) \right) - \left(-\frac{m^2}{1-x^2} + l(l+1) \right) y(x) = 0,$$

auf, wobei $m \in \mathbb{Z}$ mit $m^2 \leq l^2$ ist. Für $m = 0$ reduziert sich dies zur Legendreschen DGL.

Die Lösungen der verallgemeinerten Legendreschen DGL sind durch die **zugeordneten Legendre-Polynome**

$$P_l^m(x) := (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \quad (0 \leq m \leq l)$$

gegeben. Für $-l < m < l$ gilt dabei

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x).$$

13.3.2 Kugelflächenfunktionen

Eine sehr wichtige Gleichung der Physik ist die **Laplace-Gleichung**

$$\Delta f = 0,$$

die z.B. im Zusammenhang mit Wellenphänomenen auftaucht. Streng genommen handelt es sich um eine partielle DGL (siehe Kapitel 13.4), die wir aber im folgenden auf gewöhnliche DGLen zurückführen werden.

Wir wollen speziell die Lösung der Laplace-Gleichungen in Kugelkoordinaten $f(r, \vartheta, \varphi)$ untersuchen. Dazu machen wir den (Separations-)Ansatz (siehe später in Kapitel 13.5.4)

$$f(r, \vartheta, \varphi) = R(r)S(\vartheta)T(\varphi),$$

d.h. wir suchen Lösungen, die sich als Produkt von Funktionen nur einer Variablen schreiben lassen!

Zunächst setzen wir diesen Ansatz in die Laplace-Gleichung ein, wobei wir den Laplace-Operator Δ zweckmäßigerweise in Kugelkoordinaten darstellen (siehe Kapitel 11.3.4):

$$\begin{aligned} 0 &= \Delta f(r, \vartheta, \varphi) \\ &= \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} \right] f(r, \vartheta, \varphi) \\ &= \frac{RST}{r^2} \left[\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{\sin^2 \vartheta} \left(\frac{\sin \vartheta}{S} \frac{d}{d\vartheta} \left(\sin \vartheta \frac{dS}{d\vartheta} \right) + \frac{1}{T} \frac{d^2 T}{d\varphi^2} \right) \right]. \end{aligned}$$

Wir betrachten nun die beiden Terme in der eckigen Klammer. Der erste hängt nur von r ab, der zweite nur von ϑ und φ . Daher kann die Gleichung für beliebige Werte von r , ϑ und φ nur dann erfüllt sein, wenn beide konstant sind, genauer

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = \alpha \quad \text{und} \quad \frac{1}{\sin^2 \vartheta} \left(\frac{\sin \vartheta}{S} \frac{d}{d\vartheta} \left(\sin \vartheta \frac{dS}{d\vartheta} \right) + \frac{1}{T} \frac{d^2 T}{d\varphi^2} \right) = -\alpha$$

mit einer noch unbekanntenen Konstanten α .

Die Gleichung für r wollen wir hier nicht weiter betrachten. Für die zweite Gleichung gilt nach der elementaren Umformung in

$$\left(\frac{\sin \vartheta}{S} \frac{d}{d\vartheta} \left(\sin \vartheta \frac{dS}{d\vartheta} \right) + \alpha \sin^2 \vartheta \right) + \frac{1}{T} \frac{d^2 T}{d\varphi^2} = 0$$

ein ähnliches Argument: Der erste Teil hängt nur von ϑ ab, der zweite nur von φ . Daher gilt

$$\frac{\sin \vartheta}{S} \frac{d}{d\vartheta} \left(\sin \vartheta \frac{dS}{d\vartheta} \right) + \alpha \sin^2 \vartheta = \beta \quad \text{und} \quad \frac{1}{T} \frac{d^2 T}{d\varphi^2} = -\beta,$$

mit einer Konstanten β .

Die zweite Gleichung kann direkt gelöst werden:

$$T(\varphi) = t_1 e^{i\sqrt{\beta}\varphi} + t_2 e^{-i\sqrt{\beta}\varphi}.$$

Die erste Gleichung formen wir zunächst mit Hilfe der Substitution $x := \cos \vartheta$ um in

$$\frac{d}{dx} \left((1-x^2) \frac{dS}{dx} \right) + \left(\alpha - \frac{\beta}{1-x^2} \right) S(x) = 0.$$

Dies ist eine verallgemeinerte Legendre-DGL. Damit eine Lösung existiert, müssen α und β die Form

$$\alpha = l(l+1) \quad \text{und} \quad \beta = m^2$$

haben mit ganzzahligen l und m . Es gilt dann

$$S(\vartheta) = P_l^m(\cos \vartheta).$$

Somit haben wir den Winkelanteil der Lösung der Laplace-Gleichung in Kugelkoordinaten bestimmt. Er wird erzeugt durch die sog. Kugelflächenfunktionen.

Definition 13.3.1 (Kugelflächenfunktionen).

Die **Kugelflächenfunktionen** sind definiert durch

$$Y_{lm}(\vartheta, \varphi) := \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos \vartheta) e^{im\varphi}$$

Man beachte hierbei, dass es in der Literatur durchaus unterschiedliche Konventionen für das Vorzeichen und die Normierung gibt!

Explizit lauten die ersten Kugelflächenfunktionen

$$Y_{00} = \sqrt{\frac{1}{4\pi}}, \quad Y_{11} = -\sqrt{\frac{3}{8\pi}} \sin \vartheta e^{i\varphi}, \quad Y_{10} = \sqrt{\frac{3}{4\pi}} \cos \vartheta.$$

Außerdem gilt

$$Y_{l,-m} = (-1)^m \overline{Y_{lm}}.$$

Die Kugelflächenfunktionen erfüllen die Orthogonalitätsbeziehungen

$$\int d\Omega \overline{Y_{l'm'}}(\vartheta, \varphi) Y_{lm}(\vartheta, \varphi) = \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta \sin \vartheta \overline{Y_{l'm'}}(\vartheta, \varphi) Y_{lm}(\vartheta, \varphi) = \delta_{ll'} \delta_{mm'}.$$

13.3.3 Bessel'sche DGL

Macht man einen analogen Separationsansatz für die Laplace-Gleichung in Zylinderkoordinaten, so kommt man für den ρ -abhängigen Teil der Lösung zur **Bessel'sche DGL**

$$x \frac{d}{dx} \left(x \frac{dy}{dx} \right) + (x^2 - p^2) y(x) = 0.$$

Eine Klasse von Lösungen sind die **Besselfunktionen 1. Art der Ordnung p**:

$$J_p(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)} \left(\frac{x}{2}\right)^{2n+p}.$$

Dabei ist $\Gamma(x)$ die **Gamma-Funktion**. Für positive, ganzzahlige $x = n + 1$ gilt $\Gamma(n + 1) = n!$.

Weitere Lösungen sind durch die **Besselfunktionen 2. Art**

$$Y_p(x) = \lim_{\nu \rightarrow p} \frac{\cos(\pi\nu) J_\nu(x) - J_{-\nu}(x)}{\sin(\pi\nu)}$$

gegeben.

13.3.4 Hermite'sche DGL

Die **Hermite'sche DGL**

$$\frac{d}{dx} \left(e^{-x^2} \frac{dy}{dx} \right) + 2ne^{-x^2} y(x) = 0$$

wird durch die **Hermite-Polynome** $H_n(x)$ gelöst. Diese erhält man aus der erzeugenden Funktion

$$\phi(x, h) = \sum_{n=0}^{\infty} H_n(x) \frac{h^n}{n!} = e^{-h^2+2hx}.$$

Sie genügen der Orthogonalitätsrelation

$$\int_{-\infty}^{\infty} dx e^{-x^2} H_n(x) H_m(x) = \sqrt{\pi} 2^n n! \delta_{nm}.$$

13.3.5 Laguerre'sche DGL

Die **Laguerre'sche DGL** lautet

$$\frac{d}{dx} \left(x e^{-x} \frac{dy}{dx} \right) + n e^{-x} y(x) = 0.$$

Ihre Lösungen sind die **Laguerre-Polynome** L_n , die durch die erzeugende Funktion

$$\phi(x, h) = \sum_{n=0}^{\infty} L_n(x) h^n = \frac{1}{1-h} e^{-\frac{xh}{1-h}}$$

bestimmt sind.

13.4 Partielle DGL

Partielle Differentialgleichungen (pDGL) enthalten (partielle) Ableitungen nach mehreren Variablen, z.B. $\frac{\partial u(x,t)}{\partial x}$ und $\frac{\partial u(x,t)}{\partial t}$. Sie sind in der Regel deutlich schwieriger als gewöhnliche DGL und selbst für Gleichungen 1. Ordnung (bei denen nur erste Ableitungen auftauchen) gibt es keine so allgemeine Theorie wie für gewöhnliche DGL.

In der Physik am wichtigsten sind partielle DGL 2. Ordnung, bei der höchstens 2. Ableitungen auftreten. Ihre allgemeine Form ist im Falle von zwei unabhängigen Variablen x und y :

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + d \frac{\partial u}{\partial x} + e \frac{\partial u}{\partial y} + fu = R(x, y).$$

Dabei können alle auftretenden Koeffizienten wie die gesuchte Funktion $u(x, y)$ ebenfalls Funktionen von x und y sein, also $a = a(x, y)$ etc.

Man unterscheidet drei Typen : **elliptische**, **parabolische** und **hyperbolische** pDGL 2. Ordnung. Diese liegen vor, wenn folgende Bedingungen erfüllt sind:

$$\begin{array}{ll} \text{hyperbolisch:} & b^2 > 4ac, \\ \text{parabolisch:} & b^2 = 4ac, \\ \text{elliptisch:} & b^2 < 4ac. \end{array}$$

Da a , b und c Funktionen sind, kann der Typ in verschiedenen Bereichen des Definitionsbereiches unterschiedlich sein!

Häufig hat man es mit einem **Randwertproblem** zu tun. Hierbei muß die gesuchte Lösung u noch am Rand ∂A des Definitionsbereiches A gewisse Bedingungen erfüllen. Man unterscheidet dabei folgende Typen:

- **Dirichlet-Randbedingungen:** Hier sind die Funktionswerte am Rand vorgegeben, also $u(\partial A)$.
- **Neumann-Randbedingungen:** Hier ist die Ableitung der Funktion am Rand in Richtung der Normalen vorgegeben, d.h. $\underline{n} \cdot \text{grad } u|_{\partial A}$.
- **Cauchy-Randbedingungen:** Hierbei handelt es sich um eine (gewichtete) Kombination von Dirichlet- und Neumann-Randbedingungen.

13.4.1 Wichtige pDGL der Physik

Im folgenden wollen wir die wichtigsten pDGL der Physik aufzählen und klassifizieren.

- Die **Laplace-Gleichung**

$$\Delta u = 0$$

haben wir bereits kennengelernt. In zwei Dimensionen lautet sie expliziter

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Die zugehörige inhomogene Gleichung

$$\Delta u = \rho(x, y)$$

heißt **Poisson-Gleichung**. Sie beschreibt z.B. das elektrische Potential einer Ladungsverteilung $\rho(x, y)$.

Die Poisson-Gleichung ist vom elliptischen Typ.

- Die **Diffusions-Gleichung** oder auch **Wärmeleitungsgleichung** lautet

$$\Delta u(x, y, t) = \frac{1}{\kappa} \frac{\partial u}{\partial t}.$$

Dabei ist $\kappa > 0$ eine Konstante.

Es handelt sich um eine parabolische DGL. Für $t \rightarrow \infty$ erhält man wegen

$$\lim_{t \rightarrow \infty} \frac{\partial u}{\partial t} = 0$$

eine statische (d.h. zeitunabhängige) Lösung. Diese erfüllt dann die Laplace-Gleichung!

- Die **Wellengleichung**

$$\Delta u(x, y, t) - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0$$

beschreibt die Ausbreitung der Ausbreitung von Wellen mit der Fortpflanzungsgeschwindigkeit v . Sie ist vom hyperbolischen Typ.

- Die **Schrödinger-Gleichung**

$$-\frac{\hbar^2}{2m} \Delta u(\underline{r}, t) + V(\underline{r})u(\underline{r}, t) = i\hbar \frac{\partial u}{\partial t}$$

ist die grundlegende Gleichung der Quantenmechanik. Sie beschreibt die quantenmechanische Wellenfunktion $u(\underline{r}, t) = u(x, y, z, t)$ eines (nichtrelativistischen) Teilchens der Masse m , das sich im Potential $V(\underline{r})$ bewegt. $\hbar = h/2\pi$ ist das Wirkungsquantum.

In den obigen Beispielen haben wir in der Regel den zweidimensionalen Fall (zwei Raumdimensionen x und y) angegeben. Die Verallgemeinerung auf den n -dimensionalen Fall ist in allen Beispielen offensichtlich.

13.5 Lösungsverfahren für pDGL; Green'sche Funktionen

Wie schon erwähnt, gibt es für pDGL weniger allgemeine Aussagen oder Lösungsverfahren als für gewöhnliche DGL. In der Praxis werden daher häufig numerische Methoden wie Relaxationsverfahren oder Fast Fourier Transformation (FFT) eingesetzt. Hier wollen wir aber einige Beispiele für analytische Methoden vorstellen.

13.5.1 Integraldarstellung

Wir betrachten die eindimensionale Diffusionsgleichung²

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}$$

mit der Rand- bzw. Anfangsbedingung $u(x, 0) = f(x)$. Sie beschreibt die zeitliche Entwicklung der zur Zeit $t = 0$ vorgegebenen Wärmeverteilung $u(x, t)$.

In den Übungen (Aufgabe 32) werden wir mit Hilfe der Fourier-Transformation zeigen, dass sich Lösung in folgender Form schreiben läßt:

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} dy e^{-\frac{(x-y)^2}{4t}} f(y).$$

Den Faktor $e^{-\frac{(x-y)^2}{4t}}$ bezeichnet man in diesem Zusammenhang auch als **Hitzekern** (“heat kernel”).

13.5.2 Integraltransformation

Wir betrachten die Poisson-Gleichung in drei Raumdimensionen:

$$\Delta u(\underline{r}) = -4\pi q \delta^{(3)}(\underline{r}),$$

wobei

$$\delta^{(3)}(\underline{r}) := \delta(x)\delta(y)\delta(z).$$

Physikalisch beschreibt dann $u(\underline{r})$ das elektrische Potential einer Punktladung q am Ursprung $\underline{r} = 0$. Wir werden später sehen, wie man aus der Lösung dieses speziellen Problems zur Lösung für eine beliebige Inhomogenität $\rho(\underline{r})$ kommt.

Zur Lösung dieser Gleichung wenden wir eine Fouriertransformation auf alle drei Variablen x , y , z an, z.B.

$$\tilde{u}(p_x, p_y, p_z) := \text{FT}(u(\underline{r})) = \frac{1}{\sqrt{2\pi}^3} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz e^{-ip_x x} e^{-ip_y y} e^{-ip_z z} u(x, y, z).$$

Man beachte, dass es sich um drei getrennte Transformationen handelt. Die Variablen im Fourierraum haben wir dabei mit p_x , p_y und p_z bezeichnet. Wenn wir sie als Komponenten eines Vektors \underline{p} interpretieren, können wir dies auch kompakter schreiben als

$$\tilde{u}(\underline{p}) = \frac{1}{\sqrt{2\pi}^3} \int d^3 r e^{-i\underline{p} \cdot \underline{r}} u(\underline{r}).$$

²Wobei wir o.B.d.A. $\kappa = 1$ setzen.

Wendet man nun die Fouriertransformation auf die obige Poisson-Gleichung an, so erhält man nach kurzer Rechnung

$$-p^2 \tilde{u}(\underline{p}) = -4\pi q \frac{1}{\sqrt{2\pi}^3},$$

und somit

$$\tilde{u}(\underline{p}) = \sqrt{\frac{2}{\pi}} \frac{q}{p^2}.$$

Der Faktor $-p^2 = -p_x^2 - p_y^2 - p_z^2$ entsteht dabei durch Anwendung des Laplace-Operators $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ auf die Fouriertransformierte.

Durch Rücktransformation können wir nun die Lösung explizit bestimmen:

$$\begin{aligned} u(\underline{r}) &= \frac{1}{\sqrt{2\pi}^3} \int d^3p \tilde{u}(\underline{p}) e^{i\underline{p}\cdot\underline{r}} \\ &= \frac{q}{2\pi^2} \int_0^\infty dp \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta p^2 \sin \vartheta \frac{1}{p^2} e^{ipr \cos \vartheta} \\ &= \frac{q}{r}. \end{aligned}$$

Hier haben wir die Details der Rechnung ausgelassen. Beim Übergang zur zweiten Zeile wurden für die Berechnung des Integrals Kugelkoordinaten ($p = |\underline{p}|$, ϑ , φ) im p -Raum eingeführt. Dabei wurde die p_z -Achse parallel zur Richtung von \underline{r} gewählt. Bei der Berechnung des Integrals wurde außerdem $\int_0^\infty \frac{\sin y}{y} dy = \frac{\pi}{2}$ verwendet.

Aus physikalischer Sicht ist das Ergebnis natürlich nicht überraschend, denn es handelt sich um das bekannte Coulomb-Potential einer Punktladung q im Ursprung!

13.5.3 Green'sche Funktion

Wir wollen jetzt ein wichtiges Verfahren diskutieren, das zur Lösung allgemeiner inhomogener DGL mit gegebenen Randbedingungen, insbesondere aber auch von pDGL.

Wir wollen hier speziell den eindimensionalen Fall betrachten, und zwar für ein Sturm-Liouville-Problem vom Typ

$$\hat{D}y(x) = f(x) \quad \text{mit } y(a) = y(b) = 0$$

und dem Sturm-Liouville-Operator

$$\hat{D}y = (p(x)y'(x))' + a(x)y.$$

Die Wahl der Randbedingungen ist keine einschneidende Einschränkung. Für den Fall $y(a), y(b) \neq 0$ löst man zuerst die homogene Gleichung $\hat{D}y_h = 0$ mit $y_h(a) = y_a$, $y_h(b) = y_b$. Ist $y(x)$ die Lösung des oben angegebenen Problems mit $y(a) = y(b) = 0$, so ist

$$\hat{y}(x) = y(x) + y_h(x)$$

die Lösung von $\hat{D}y(x) = f(x)$ mit $\hat{y}(a) = y_a$ und $\hat{y}(b) = y_b$.

Wir kommen nun zur Lösung von $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$.

Definition 13.5.1 (Green'sche Funktion).

Die **Green'sche Funktion** (oder *Einflussfunktion*) der DGL $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$ ist die Lösung $G(x, z)$ der Gleichung

$$\begin{aligned} \hat{D}G(x, z) &= \delta(x - z) \\ G(a, z) = G(b, z) &= 0 \quad (a < x, z < b). \end{aligned}$$

$G(x, z)$ stellt eine Art *Elementarlösung* dar, aus der sich die Lösung von $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$ bestimmen läßt. Dies sieht man folgendermaßen: Zunächst gilt, nach Definition der Green'schen Funktion,

$$(pG')' + aG = \delta(x - z).$$

Diese Gleichung multiplizieren wir mit $y(x)$ und subtrahieren hiervon die betrachtete Sturm-Liouville-Gleichung,

$$G(py')' + ayG = G(x, z)f(x),$$

die wir noch mit $G(x, z)$ multipliziert haben. Dies ergibt zunächst

$$[p(yG' - y'G)]' = y\delta(x - z) - Gf$$

und nach Integration über x über das Intervall $[a, b]$

$$p(yG' - y'G)\Big|_a^b = \int_a^b dx y(x)\delta(x - z) - \int_a^b dx G(x, z)f(x) = y(z) - \int_a^b dx G(x, z)f(x).$$

Die linke Seite verschwindet auf Grund der Randbedingungen. Somit erhalten wir folgende Darstellung der Lösung des Sturm-Liouville-Problems $\hat{D}y(x) = f(x)$, $y(a) = y(b) = 0$ durch die Green'sche Funktion:

$$y(x) = \int_a^b dx G(x, z)f(x)$$

für $z \in]a, b[$. Dies erlaubt uns, bei Kenntnis der Green'schen Funktion, die DLG für beliebige Inhomogenitäten $f(x)$ explizit zu lösen!

Beispiel 13.5.1. Als Beispiel betrachten wir die Poisson-Gleichung³

$$\Delta u = -4\pi\rho(x, y).$$

Die Green'sche Funktion hatten wir im Prinzip schon in Kapitel 13.5.2 bestimmt, wenn man dort $q = 1$ setzt:

$$G(\underline{r}, \underline{r}') = \frac{1}{|\underline{r} - \underline{r}'|}.$$

Somit lautet die allgemeine Lösung der Poisson-Gleichung

$$u(\underline{r}) = \int_{\mathbb{R}^3} d^3r' \frac{\rho(\underline{r}')}{|\underline{r} - \underline{r}'|}.$$

³Wir haben hier die Inhomogenität etwas anders normiert, um die Poisson-Gleichung in eine 'physikalische' Form zu bringen.

13.5.4 Separation der Variablen

Ein Verfahren, mit dem sich in vielen Fällen Lösungen von pDGL bestimmen lassen, beruht auf einem **Separationsansatz**, z.B.

$$u(x, y, t) = X(x)Y(y)T(t).$$

Wir haben dies bereits bei der Herleitung der Kugelflächenfunktionen in Kapitel 13.3.2 angewendet. Das dort beschriebene Vorgehen ist typisch für dieses Verfahren. Wir wollen uns daher auf ein paar allgemeine Bemerkungen beschränken.

- Ein Separationssatz bietet sich vor allem dann an, wenn der Rand durch Koordinatenlinien oder -flächen darstellbar ist. Daher ist beim Separationsansatz die Wahl geeigneter Koordinaten, in denen man die Separation durchführt, wichtig.
- Durch den Ansatz wird die pDGL in mehrere gewöhnliche DGL zerlegt, die oft die Form von Eigenwertproblemen haben. Diese DGL sind durch gemeinsame Konstanten miteinander verbunden.
- Die allgemeine Lösung erhält man als Linearkombination von Produkten aller Teillösungen.
- Randbedingungen schränken die Lösungen ein, z.B. legen sie die erlaubten Eigenwerte fest oder bestimmen die Koeffizienten der Linearkombinationen.
- Manchmal gibt es nur für bestimmte Parameterwerte der DGL zulässige Lösungen.

Kapitel 14

Symmetrien und Gruppen

14.1 Symmetrien

Symmetrien oder Invarianzen spielen in der Physik eine extrem wichtige Rolle, da sie z.B. eine Vereinfachung der Beschreibung erlauben. Dabei gibt es unterschiedliche Klassen von Symmetrien, die eine Rolle spielen.

1. **Symmetrien von Objekten:** Z.B. ist ein Kreis invariant unter beliebigen Drehungen um seinen Mittelpunkt. Ein Quadrat ist dagegen nur invariant unter Drehungen um den Mittelpunkt, wenn der Drehwinkel ein Vielfaches von $\frac{\pi}{2}$ beträgt. Ein gleichseitiges Dreieck schließlich ist invariant unter Drehungen um den Schwerpunkt um Vielfache von $\frac{2\pi}{3}$.
2. **Symmetrien der Naturgesetze:** Wir nehmen an, dass die Naturgesetze zeitlich unveränderlich sind, d.h. ein (ideales) Experiment liefert heute und nächste Woche das gleiche Ergebnis. Die Naturgesetze sind daher invariant unter zeitlichen Verschiebungen. In analoger Weise sind sie auch unter räumlichen Verschiebungen invariant, d.h. die Naturgesetze in Köln und New York sind die gleichen. Offensichtlich besteht auch Invarianz unter beliebigen Kombinationen von zeitlichen und räumlichen Verschiebungen. Eine wichtige Konsequenz solcher Invarianzen sind sog. *Erhaltungssätze*, z.B. impliziert die Invarianz unter zeitlichen Verschiebungen den Energieerhaltungssatz.

Allgemein versteht man unter einer **Symmetrie** eine Operation, die ein Objekt in eine von ihm nicht unterscheidbare Form überführt. Diese abstrakte Definition wird von den oben angegebenen Beispiele gut illustriert.

Grundsätzlich ist es eine Bestimmung aller dieser Operation (Rotationen, Verschiebungen, Spiegelungen etc.) bei jedem Problem sehr wichtig.

Die Gruppentheorie erlaubt nun die Untersuchung der Gesamtheit solcher Operationen. Die Elemente der (mathematischen) Gruppe entsprechen dabei den Symmetrioperationen.

14.2 Gruppen

Definition 14.2.1 (Gruppe).

Eine **Gruppe** ist eine (nichtleere) Menge G , für deren Elemente eine Verknüpfung \cdot definiert ist, die üblicherweise als (Gruppen-)Multiplikation¹ bezeichnet wird.

Die Gruppenmultiplikation muß folgende Eigenschaften haben (Man beachte, dass statt $a \cdot b$ oft einfach ab geschrieben wird!):

1. Abgeschlossenheit: Für alle $a, b \in G$ ist auch $ab \in G$.
2. Assoziativität: Sind $a, b, c \in G$, so gilt: $(ab)c = a(bc)$.
3. Einheitselement (neutrales Element): Es gibt ein $e \in G$, so dass für alle $a \in G$ gilt: $ae = a$ und $ea = a$.
4. Inverses Element: Zu jedem $a \in G$ existiert ein $b \in G$ mit $ab = e$ und $ba = e$. Man bezeichnet dieses Element auch als das zu a Inverse und schreibt $b = a^{-1}$.
Bem.: Wie man leicht überprüft, gilt $(ab)^{-1} = b^{-1}a^{-1}$.

Die Anzahl der Elemente einer Gruppe heißt **Ordnung** der Gruppe. **Diskrete Gruppen** haben endlich oder abzählbar unendlich viele Elemente. **Kontinuierliche Gruppen** haben dagegen überabzählbar viele Elemente.

Gilt zusätzlich

5. Kommutativität: Für alle $a, b \in G$ ist $ab = ba$.

so heißt die Gruppe **kommutativ** oder **abelsch**.

Beispiel 14.2.1. Wir betrachten einige Beispiele für Gruppen.

1. Die ganzen Zahlen \mathbb{Z} bilden bzgl. der Addition eine diskrete, abelsche Gruppe.
2. $\{-1, 1\}$ bzgl. der (üblichen) Multiplikation.
3. $\{e^{i\alpha} \mid \alpha \in \mathbb{R}\}$ mit der Multiplikation.
4. Alle Vektorräume sind bzgl. der Vektoraddition abelsche Gruppen.
5. Die natürlichen Zahlen \mathbb{N} mit der Addition sind bilden keine Gruppe, da die Inversen nicht enthalten sind (z.B. -2 , das Inverse von 2 bzgl. der Addition).

Endliche Gruppen kann man auch durch eine **Gruppenmultiplikationstabelle** oder **Gruppentafel** definieren. Für Beispiel 2) lautet diese z.B.

\cdot	1	-1
1	1	-1
-1	-1	1

¹Selbst dann, wenn es sich eigentlich um eine Addition handelt!

Gruppen mit einer analogen Gruppentafel bezeichnet man mit Z_2 .
Die Gruppe Z_4 ist durch die Gruppentafel

·	1	a	b	c
1	1	a	b	c
a	a	b	c	1
b	b	c	1	a
c	c	1	a	b

definiert.

Anhand solcher Gruppentafeln lassen sich die Gruppeneigenschaften leicht überprüfen.

Definition 14.2.2 (Untergruppe).

Ist eine Teilmenge $H \subset G$ selbst eine Gruppe, so nennt man sie auch **Untergruppe von G** .

Bem.: Man muß nicht alle Gruppeneigenschaften nachweisen. Es genügt, neben der Abgeschlossenheit von H , zu zeigen, dass das neutrale Element und die Inversen *in* H enthalten sind. Ihre Existenz ist bereits klar, da G eine Gruppe ist.

Definition 14.2.3 (direktes Produkt).

Aus zwei Gruppen G_1 und G_2 kann man eine neue Gruppe $G := G_1 \times G_2$ konstruieren, das **direkte Produkt** von G_1 und G_2 , mit den Elementen $g := (a, b)$, wobei $a \in G_1$ und $b \in G_2$, und der Multiplikation

$$g_1 g_2 = (a_1, b_1)(a_2, b_2) := (a_1 a_2, b_1 b_2).$$

Auf Grund dieser Definition “erbt” G die Gruppeneigenschaften seiner Bausteine G_1 und G_2 .

14.3 Konjugations- und Nebenklassen

Im den folgenden Abschnitten wollen wir einige Möglichkeiten vorstellen, mit denen sich die Elemente von Gruppen “klassifizieren” lassen.

14.3.1 Äquivalenzrelationen

Definition 14.3.1 (Äquivalenzrelation).

Eine Verknüpfung ‘ \sim ’ zwischen den Elementen einer Menge heißt **Äquivalenzrelation**, falls sie folgende Eigenschaften hat:

1. reflexiv: $a \sim a$
2. symmetrisch: $a \sim b \iff b \sim a$
3. transitiv: Gilt $a \sim b$ und $b \sim c$, dann auch $a \sim c$.

Eine Menge von Objekten, die zueinander äquivalent sind, bezeichnet man als **Äquivalenzklasse** bzw. nur als **Klasse**. Solche Klassen sind nützlich, um Mengen weiter zu unterteilen (“Klassifikation”).

Ein Beispiel ist der Verknüpfung “. . . hört die gleiche Vorlesung wie. . .”. Die Reflexivität ist offensichtlich, da natürlich jeder die gleiche Vorlesung wie er selbst hört. Auch die Symmetrie ist klar, ebenso die Transitivität: Wenn Albert die gleiche Vorlesung wie Berta hört und Berta die gleiche Vorlesung wie Claudia, dann hören auch Albert und Claudia die gleiche Vorlesung!

14.3.2 Konjugationsklassen

Definition 14.3.2 (Konjugationsklassen).

Es sei G eine Gruppe und $a, b \in G$. Existiert dann ein $g \in G$ mit

$$a = gbg^{-1},$$

so nennt man a und b **konjugiert zueinander** und schreibt $a \sim b$. Offenbar definiert dies eine Äquivalenzrelation.

Die zueinander konjugierten Elemente einer Gruppe bilden eine **Konjugationsklasse**, von denen es natürlich mehrere geben kann. Unterschiedliche Klassen haben keine gemeinsamen Elemente!

Ein Spezialfall sind die abelschen Gruppen. Hier bildet jedes Element seine eigene Konjugationsklasse, da

$$a = gbg^{-1} = gg^{-1}b = b.$$

14.3.3 Nebenklassen

Mit Hilfe einer Untergruppe $H \subset G$ kann man G in durchschnittsfreie Teilmengen aufteilen, die “Nebenklassen”.

Definition 14.3.3 (Nebenklasse).

Sei $a \in G$. Dann heißt

$$Ha = \{ha \mid h \in H\}$$

rechte Nebenklasse von a . Analog definiert man die **linke Nebenklasse** aH .

G ist dann die Vereinigungsmenge aller unterschiedlichen Nebenklassen:

$$G = He \cup Ha_1 \cup Ha_2 \cup \dots$$

Bemerkung.

1. Offensichtlich ist $He = H$.
2. Nebenklassen sind i.a. nur einfache Mengen, keine Gruppen!!
3. Alle Nebenklassen enthalten die gleiche Anzahl n_H von Elementen, wobei n_H die Anzahl der Elemente von H ist.

4. Satz von Lagrange: Die Anzahl der Nebenklassen ist durch n_G/n_H gegeben, wobei n_G die Anzahl der Elemente von G ist.

Folgerung: Ist n_G eine Primzahl, so hat G keine nichttriviale Untergruppe.

14.4 Spezielle Untergruppen

Wir stellen kurz einige wichtige Typen von Untergruppen vor:

Definition 14.4.1 (Spezielle Untergruppen).

- a) $gHg^{-1} = \{ghg^{-1} | h \in H\}$ heißt auch **konjugierte Untergruppe** zu H bzgl. g .
- b) Ist $gHg^{-1} = H$, d.h. $gH = Hg$, so heißt H **Normalteiler** oder **invariante Untergruppe**.
Die Ordnung des Normalteilers muß echter Teiler der Ordnung n_G der Gruppe sein. Dies ist nützlich beim "Erraten" von Normalteilern!
Gruppen ohne Normalteil (außer den trivialen G und $\{e\}$) nennt man **einfach**, solche ohne *abelschen* Normalteiler **halbeinfach**.
- c) Als **Zentrum** bezeichnet man die Untergruppe von G , deren Elemente mit allen Elemente aus G vertauschen. Das Zentrum ist auch ein Normalteiler.
Abelsche Gruppen sind mit ihrem Zentrum identisch.
- d) Die Menge aller Nebenklassen eines Normalteilers N ist selbst eine Gruppe, die man als **Quotienten- oder Faktorgruppe** G/N bezeichnet:

$$G/N := \{Ne, Na_1, Na_2, \dots\},$$

wobei für die Multiplikation zweier Elemente $x \in Na$ und $y \in Nb$ (d.h. $x = n_1a$ und $y = n_2b$ mit $n_1, n_2 \in N$) gilt:

$$xy := (n_1a)(n_2b) = n_1n_2(ab) = n_1n_2n_3c \in Nc.$$

14.5 Wichtige Gruppen

14.5.1 Permutationsgruppe

Definition 14.5.1 (Permutationsgruppe S_n).

Die Menge aller möglichen Permutationen von n Elementen ist eine Gruppe der Ordnung $n!$, die sog. **Permutationsgruppe** S_n . Dabei ist die Gruppenmultiplikation als Hintereinanderausführung zweier Permutationen definiert.

Wir hatten schon in Kapitel 2.2 einige Eigenschaften von Permutationen kennengelernt. Ausgehend von der Identität $(1\ 2\ 3\ 4)$ (im Falle $n = 4$) können wir z.B. die Permutation $1 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 4, 4 \rightarrow 2$ beschreiben durch die Aufzählung der Bildmenge $(1\ 3\ 4\ 2)$ oder ausführlicher durch $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$.

Die Multiplikation, also Hintereinanderausführung zweier Permutationen sieht dann z.B. folgendermaßen aus ($n = 3$):

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Dabei wird der rechte ‘Faktor’ zuerst angewendet und auf dessen Ergebnis dann der linke Faktor. Z.B. wird die ‘1’ zunächst auf die ‘2’ abgebildet und diese dann auf die ‘3’.

Vertauschen wir die Reihenfolge der Faktoren, so erhalten wir

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Die Permutationsgruppe S_n ist also nicht kommutativ!

Die Menge (siehe auch Kap. 2.2)

$$A_n := \{P \in S_n \mid P \text{ ist gerade Permutation}\}$$

ist eine Untergruppe von S_n , die sog. **alternierende Gruppe**. Sie ist sogar Normalteiler von S_n .

Die Menge

$$A_n := \{P \in S_n \mid P \text{ ist zyklische Permutation}\}$$

ist ebenfalls eine Untergruppe von S_n , die man als **zyklische Gruppe** bezeichnet. Genauer ist A_n eine abelsche Untergruppe mit n Elementen.

Z.B. im Fall $n = 4$ sind ist die zyklische Gruppe gegeben durch

$$A_n := \{(1234), (4123), (3412), (2341)\}.$$

Die Menge $\{0, 1, 2, \dots, n-1\}$ mit der Addition modulo n (d.h. man betrachtet nur den Rest bei Division durch n , z.B. ist dann $(n-1) + 3 = n+2 \equiv 2 \pmod{n}$) ist eine zyklische Gruppe, die man Z_n nennt. Die Multiplikationstabellen im Fall $n = 2$ und $n = 4$ hatten wir schon in Kapitel 14.2 angegeben.

14.5.2 Matrixgruppen

Definition 14.5.2 (Matrixgruppen).

Eine Reihe wichtiger Gruppen sind Teilmengen der Menge $\mathcal{M}(n, n)$ aller $n \times n$ -Matrizen. In Kapitel 14.6 werden wir sehen, dass sie auch Repräsentanten von abstrakten Gruppen mit der gleichen Struktur sind.

- Die **orthogonale Gruppe**

$$O(n) := \{ \underline{M} \in \mathcal{M}(n, n) \mid \underline{M}^T \cdot \underline{M} = \mathbb{1} = \underline{M} \cdot \underline{M}^T \}$$

aller orthogonalen $n \times n$ -Matrizen. Dabei heißt eine Matrix \underline{M} **orthogonal**, falls gilt

$$\underline{M}^{-1} = \underline{M}^T,$$

d.h. die transponierte Matrix ist gleich der Inversen.

Orthogonale Matrizen lassen Skalarprodukte invariant:

$$(\underline{M} \underline{a}) \cdot (\underline{M} \underline{b}) = (\underline{M} \underline{a})^T \cdot (\underline{M} \underline{b}) = \underline{a}^T \underline{M}^T \underline{M} \underline{b} = \underline{a}^T \underline{b} = \underline{a} \cdot \underline{b}.$$

Dabei haben wir ausgenutzt, dass sich das Skalarprodukt $\underline{a} \cdot \underline{b}$ zweier Vektoren auch als Matrixmultiplikation $\underline{a}^T \underline{b}$ interpretieren läßt und das allgemein gilt $(\underline{A} \underline{B})^T = \underline{B}^T \underline{A}^T$. Anschaulich umfasst daher die Gruppe $O(n)$ alle Drehungen und Drehspiegelungen im \mathbb{R}^n . Es gilt

$$1 = \det(\mathbb{1}) = \det(\underline{M}^T \cdot \underline{M}) = \det(\underline{M}^T) \det(\underline{M}) = (\det \underline{M})^2.$$

Somit haben orthogonale Matrizen die Determinante $\det \underline{M} = \pm 1$. Matrizen mit $\det \underline{M} = 1$ beschreiben reine Drehungen, die mit $\det \underline{M} = -1$ Drehspiegelungen.

- Die **spezielle orthogonale Gruppe**

$$SO(n) := \{ \underline{M} \in O(n) \mid \det(\underline{M}) = 1 \}$$

beschreibt reine Drehungen im \mathbb{R}^n .

- Die **unitäre Gruppe**

$$U(n) := \{ \underline{U} \in \mathcal{M}(n, n) \mid \underline{U}^\dagger \cdot \underline{U} = \mathbb{1} = \underline{U} \cdot \underline{U}^\dagger \}$$

aller unitären $n \times n$ -Matrizen. Dabei heißt eine Matrix \underline{U} **unitär**, falls gilt

$$\underline{U}^{-1} = \underline{U}^\dagger,$$

wobei die **adjungierte Matrix** (oder auch **hermitesch konjugierte Matrix**) definiert ist durch

$$\underline{U}^\dagger := (\bar{\underline{U}})^T,$$

d.h. man bildet die Transponierte der Matrix, die aus den komplex-konjugierten Elementen der Matrix \underline{U} besteht. Ist die Matrix reell, so ist die Adjungierte gleich der Transponierten.

Wie im Fall der orthogonalen Gruppe folgt wieder, dass $\det(\underline{U}) = \pm 1$ ist.

- Die **spezielle unitäre Gruppe**

$$SU(n) := \{ \underline{U} \in U(n) \mid \det(\underline{U}) = 1 \}.$$

- Die **allgemeine lineare Gruppe** (engl. “general linear group”)

$$GL(n, \mathbb{R}) := \{ \underline{M} \in \mathcal{M}(n, n) \mid \det(\underline{M}) \neq 0 \}$$

umfaßt die allgemeinen, nicht singulären (d.h. invertierbaren), linearen Transformationen $\underline{x} \mapsto \underline{x}' = \underline{M} \underline{x}$ im \mathbb{R}^n .

14.6 Darstellungen

Gruppen können auf verschiedene Weise realisiert werden. Erfolgt die Realisierung durch Matrizen, so spricht man von einer **Darstellung** der Gruppe.

Beispiel 14.6.1. Wir hatten schon am Anfang dieses Kapitels argumentiert, dass die Drehungen im \mathbb{R}^n eine Gruppe bilden. Wählt man nun eine Basis, so entsprechen den Vektoren als n -komponentige Spaltenvektoren und deren Drehung der Multiplikation mit einer orthogonalen $n \times n$ -Matrix (mit Determinante 1). In diesem Sinne ist die Matrixgruppe $SO(n)$ als Darstellung einer abstrakten Gruppe zu interpretieren, die den Drehungen im \mathbb{R}^n entspricht.

Definition 14.6.1 (Darstellung).

Eine **Darstellung** D der Gruppe G ist eine Abbildung

$$D : G \rightarrow GL(n, \mathbb{C}) \quad \text{mit} \quad \underline{D}(a)\underline{D}(b) = \underline{D}(ab).$$

Eine Abbildung mit dieser Eigenschaft bezeichnet man auch als **Gruppenhomomorphismus**. Offensichtlich folgt hieraus bereits, dass $\underline{D}(e) = \mathbb{1}$ sein muß.

Darstellungen sind nicht eindeutig, man kann immer äquivalente Darstellungen D' durch eine **Ähnlichkeitstransformation**

$$\underline{D}'(a) := \underline{V} \underline{D}(a) \underline{V}^{-1}$$

mit einer nicht-singulären Matrix \underline{V} erzeugen. Das diese die gleichen Gruppeneigenschaften hat, sieht man folgendermaßen: Aus $\underline{D}(a)\underline{D}(b) = \underline{D}(ab)$ folgt

$$\underline{D}'(ab) = \underline{V} \underline{D}(ab) \underline{V}^{-1} = \underline{V} \underline{D}(a)\underline{D}(b) \underline{V}^{-1} = \underline{V} \underline{D}(a) \underline{V}^{-1} \underline{V} \underline{D}(b) \underline{V}^{-1} = \underline{D}'(a)\underline{D}'(b).$$

Wir betrachten noch einige Spezialfälle:

- Bei einer **unitären Darstellung** sind die Matrizen $\underline{D}(a)$ unitär. Dann gilt auch

$$\underline{D}(a^{-1}) = (\underline{D}(a))^\dagger = (\underline{D}(a))^{-1}.$$

- Bei einer **treuen Darstellung** entspricht jedem Gruppenelement genau eine Matrix und umgekehrt.
- Eine **reguläre Darstellung** einer endlichen Gruppe der Ordnung n ist durch eine $n \times n$ -Matrixdarstellung gegeben. Reguläre Darstellungen sind auch treu.
- Eine **reduzible Darstellung** hat z.B. die Form

$$\underline{\underline{D}}(a) = \begin{pmatrix} \underline{\underline{A}}(a) & 0 \\ 0 & \underline{\underline{B}}(a) \end{pmatrix},$$

d.h. sie zerfällt in unabhängige Teile, die jeweils für sich eine Gruppe darstellen. Ist das nicht der Fall, so spricht man von einer **irreduziblen Darstellung**.

14.7 Kontinuierliche Gruppen

Definition 14.7.1 (kontinuierlichen Gruppen).

Bei einer **kontinuierlichen Gruppe** G können Gruppenelemente $g(t) \in G$ durch stetige Änderungen eines Parameters t ineinander übergeführt werden. Dabei kann t auch ein Vektor sein, der für mehrere reelle Parameter steht.

Auf Grund der Gruppeneigenschaft gilt

$$g(t_1)g(t_2) = g(t_3) \quad \text{mit } t_3 = h(t_1, t_2).$$

Beispiel 14.7.1. Ein Beispiel ist die unitäre Gruppe $U(1)$, deren Elemente die Form $g(\varphi) = e^{i\varphi}$ haben: $U(1) = \{e^{i\varphi} | \varphi \in \mathbb{R}\}$. In diesem Fall ist wegen

$$g(\varphi_1)g(\varphi_2) = e^{i\varphi_1}e^{i\varphi_2} = e^{i(\varphi_1+\varphi_2)} = g(\varphi_1 + \varphi_2)$$

h durch

$$h(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2$$

gegeben.

Definition 14.7.2 (Lie-Gruppe).

Bei einer **Lie-Gruppe** ist die Funktion $h(t_1, t_2)$ analytisch in ihren Argumenten.

Beispiel 14.7.2. Die Elemente $g(\varphi)$ der speziellen orthogonalen Gruppe $SO(2)$ beschreiben Drehungen um die z -Achse um den Winkel φ und können allgemein in der Form (vgl. Aufg. 34)

$$g(\varphi) = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$$

mit $\varphi \in [0, 2\pi[$ geschrieben werden.

Von Drehungen um die gleiche Achse erwartet man, dass $g(\varphi_1)g(\varphi_2) = g(\varphi_1 + \varphi_2)$ gilt. Dies kann man explizit an Hand der angegebenen Darstellung nachprüfen². Daher ist wieder $h(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2$.

Offensichtlich kommutieren die Drehungen miteinander. Dies gilt aber nicht mehr, wenn man Drehungen um *unterschiedliche* Achsen betrachtet!

Beispiel 14.7.3. Ein weiteres Beispiel sind die Translationen

$$T(a) : x \mapsto x' = x + a$$

für die offensichtlich auch

$$T(a)T(b) = T(a + b)$$

gilt. Es handelt sich also um eine abelsche Gruppe.

14.8 Generatoren und Lie-Algebra

Um zu betonen, dass der Parameter t auch ein Vektor sein kann, schreiben wir im folgenden $g(\underline{a})$ statt $g(t)$, mit einem n -komponentigen Vektor \underline{a} . Außerdem nehmen wir o.B.d.A. an, dass $g(\underline{a} = \underline{0}) = e$, dem neutralen Element der Gruppe G , ist.

Definition 14.8.1 (Generatoren einer Lie-Gruppe).

Dann kann man $g(\underline{a})$ nach Taylor entwickeln (bis 1. Ordnung):

$$g(\underline{a}) = g(\underline{0}) + \sum_{k=1}^n a_k \left. \frac{\partial g(\underline{a})}{\partial a_k} \right|_{\underline{a}=\underline{0}} + \mathcal{O}(a^2).$$

Dann definiert man die **Generatoren** X_k der Lie-Gruppe durch

$$iX_k := \left. \frac{\partial g(\underline{a})}{\partial a_k} \right|_{\underline{a}=\underline{0}},$$

wobei der Faktor i einer verbreiteten Konvention entspricht.

Ist g Darstellungsmatrix einer Gruppe, dann sind auch die Generatoren Matrizen. Hierzu werden wir gleich ein Beispiel kennenlernen.

Speziell für unitäre oder orthogonale Gruppen gilt:

$$\begin{aligned} e &= g(\underline{a})g(\underline{a})^{-1} = g(\underline{a})g(\underline{a})^\dagger \\ &= \left(e + i \sum_k a_k X_k + \mathcal{O}(a^2) \right) \left(e - i \sum_k a_k X_k^\dagger + \mathcal{O}(a^2) \right) \\ &= e + i \sum_k a_k (X_k - X_k^\dagger) + \mathcal{O}(a^2). \end{aligned}$$

²Bei der Rechnung gehen die Additionstheoreme für Winkelfunktionen ein!

In der Ordnung $\mathcal{O}(a)$ gilt daher

$$X_k = X_k^\dagger,$$

d.h. die Generatoren sind hermitesch.

Weiter gilt für unitäre oder orthogonale Gruppen

$$g(\underline{a}) = \exp\left(i \sum_{k=1}^n a_k X_k\right)$$

wobei die rechte Seite durch die Reihendarstellung der Exponentialfunktion definiert ist. Aus der allgemeinen Identität für Matrizen (bzw. Operatoren)³

$$\det(\exp(\underline{A})) = \exp(\text{Spur } \underline{A}),$$

wobei die **Spur** einer $n \times n$ -Matrix \underline{A} definiert ist durch

$$\text{Spur } \underline{A} := \sum_{j=1}^n A_{jj},$$

folgt speziell für die Wahl⁴ $\underline{A} = \ln \underline{Q}$

$$1 = \det g = \exp(\text{Spur } \ln g) = \exp\left[\text{Spur}\left(i \sum_k a_k X_k\right)\right].$$

Die erste Identität gilt auf Grund der Unitarität von g . Somit erhalten wir

$$\text{Spur } X_k = 0.$$

Wir fassen die obigen Ergebnisse zusammen:

Satz 14.8.1. Die Generatoren von unitären und orthogonalen Gruppen sind hermitesch und spurlos.

Wir wollen diese abstrakten Definitionen an einigen konkreten Beispielen diskutieren.

Beispiel 14.8.1.

1. Für die Gruppe $U(1)$ mit der Darstellung $g(\varphi) = e^{i\varphi}$ ist $n = 1$ und es gibt daher nur einen Generator. Wegen

$$\left. \frac{\partial g(\varphi)}{\partial \varphi} \right|_{\varphi=0} = \left. i e^{i\varphi} \right|_{\varphi=0} = i$$

ist dieser durch

$$X = 1$$

gegeben.

³Die man z.B. durch Diagonalisierung beweisen kann.

⁴Wobei die rechte Seite wieder über die Reihenentwicklung definiert ist.

2. Für die spezielle orthogonale Gruppe $SO(2)$ ist

$$g(\varphi) = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$$

und somit wieder $n = 1$ und

$$iX = \left. \frac{\partial g(\varphi)}{\partial \varphi} \right|_{\varphi=0} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \Big|_{\varphi=0} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

also

$$X = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Hiermit können wir jetzt auch das oben angegebene allgemeine Resultat

$$g(\varphi) = \exp(i\varphi X)$$

überprüfen:

$$\begin{aligned} \exp(i\varphi X) &= \sum_{k=0}^{\infty} \frac{\varphi^k}{k!} (iX)^k = \mathbb{1} \sum_{k=0}^{\infty} \frac{(-1)^k \varphi^{2k}}{(2k)!} + iX \sum_{k=0}^{\infty} \frac{(-1)^k \varphi^{2k+1}}{(2k+1)!} \\ &= \mathbb{1} \cos \varphi + iX \sin \varphi = \begin{pmatrix} \cos \varphi & 0 \\ 0 & \cos \varphi \end{pmatrix} + \begin{pmatrix} 0 & \sin \varphi \\ -\sin \varphi & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = g(\varphi), \end{aligned}$$

wobei wir $(iX)^2 = \mathbb{1}$ ausgenutzt haben, was man leicht verifiziert.

3. Für die Gruppe der Translationen $T(a) : x \mapsto x' = x + a$ gilt:

$$X = -i \frac{d}{dx}.$$

4. In den Übungen (Aufgabe 35) untersuchen wir den Fall der $SU(2)$, die drei Generatoren besitzt.

Definition 14.8.2 (Lie-Algebra).

Für die Generatoren einer Lie-Gruppe gilt

$$[X_j, X_k] = i \sum_l c_{jk}^l X_l,$$

wobei der **Kommutator** definiert ist als

$$[A, B] := AB - BA.$$

Die Zahlen c_{jk}^l heißen **Strukturkonstanten**.

Die Generatoren bilden also einen Vektorraum bzw. sogar eine Algebra, die man **Lie-Algebra** der Gruppe bezeichnet.

Kapitel 15

Wahrscheinlichkeitstheorie und Statistik

Wahrscheinlichkeitstheorie und die damit verwandte Statistik spielen in vielen Bereichen der Physik eine wichtige Rolle, vor allem bei der Analyse von Meßdaten und in der statistischen Physik.

15.1 Grundbegriffe der Wahrscheinlichkeitstheorie

Unter ‘Zufall’ versteht man im Alltag Ereignisse, die man nicht beeinflussen oder vorhersehen kann.

Definition 15.1.1 (Ereignis, Experiment, Zufallsexperiment).

Ein **Ereignis** kann eintreffen (wahr sein) oder auch nicht (falsch sein). Ein **Experiment** (oder **Versuch**) stellt fest, ob ein Ereignis eintritt oder nicht. Ein **Zufallsexperiment** ist ein Experiment mit ungewissem Ausgang.

Ein Beispiel für ein Zufallsexperiment ist die Ziehung der Lottozahlen. Dagegen würde man das Fallenlassen eines Balles aus einer Höhe von 2 m nicht als Zufallsexperiment bezeichnen, da der Ausgang nicht ungewiß ist.

Wir geben nun eine intuitive Definition von ‘Wahrscheinlichkeit’.

Definition 15.1.2 (Wahrscheinlichkeit).

Ein Experiment resultiert mit **Wahrscheinlichkeit** $P(E)$ im Ereignis E , wenn wir erwarten (!), dass bei N -maliger Wiederholung im Limes $N \rightarrow \infty$ das Ereignis $N(E) = NP(E)$ mal eintreffen wird.

Bem.: Da man nie unendlich viele Versuche durchführen kann, muß man die Wahrscheinlichkeit theoretisch begründen oder postulieren. Man spricht daher auch von **a-priori Wahrscheinlichkeit**.

Ein Beispiel ist ein perfekter Würfel, bei dem alle Seiten etc. exakt gleich sind. Hier erwarten wir, dass alle Zahlen gleich oft als Ergebnis auftauchen. Daher ist z.B. die a-priori Wahrscheinlichkeit für eine Sechs durch $1/6$ gegeben.

Definition 15.1.3 (relative Häufigkeit).

Ein mit der Wahrscheinlichkeit eng verwandter Begriff ist die **relative Häufigkeit**. Tritt das Ereignis E bei N -maliger Wiederholung des Experiments $N(E)$ mal auf, so ist seine relative Häufigkeit

$$h_N(E) := \frac{N(E)}{N}.$$

Ist ein Versuch unter gleichen Bedingungen beliebig oft wiederholbar, so gilt natürlich

$$\lim_{N \rightarrow \infty} h_N(E) = P(E).$$

Bem.: Die Aussage “Diese Glühbirne hat mit der Wahrscheinlichkeit 0.99 eine Lebensdauer von mehr als 100 Stunden” macht u.U. Sinn, obwohl man sie nicht durch Messung der relativen Häufigkeit bestätigen kann, da man die Lebensdauer einer Birne nur einmal messen kann!

Definition 15.1.4 (Elementarereignis, Stichprobenraum).

Als **Elementarereignisse** bezeichnet man die Ereignisse, die nicht gleichzeitig zutreffen können. Als **Stichprobenraum** oder **Grundgesamtheit** oder **Wahrscheinlichkeitsraum** bezeichnet man die Menge S aller Elementarereignisse. Dann entspricht ein Ereignis E eines Versuchs mit Wahrscheinlichkeitsraum S einer Teilmenge $E = \{A, B, \dots\} \subset S$. Das Ereignis E tritt genau dann ein, wenn der Versuch mit einem $X \in E$ ausgeht.

Beispiel 15.1.1. Wir wollen uns diese Definitionen am Beispiel eines Würfels klar machen. Hier ist der Stichprobenraum offensichtlich durch

$$S = \{1, 2, 3, 4, 5, 6\}$$

gegeben. Dann ist “Augenzahl = 6” ein Elementarereignis, aber “Augenzahl = 5 oder 6” nicht. Andere Ereignisse, die keine Elementarereignisse sind, sind z.B.

$$\begin{aligned} E_1 &:= \{1, 2, 3\} : && \text{Augenzahl} \leq 3, \\ E_2 &:= \{2, 4, 6\} : && \text{gerade Augenzahl.} \end{aligned}$$

Die Definition 15.1.4 legt nahe, dass man Probleme der Wahrscheinlichkeitstheorie mit Hilfe der Mengenlehre darstellen kann.

Definition 15.1.5 (Wahrscheinlichkeit und Mengenlehre).

Wir betrachten einen Wahrscheinlichkeitsraum S und Ereignisse $E_1, E_2 \subset S$. Da es sich um Teilmengen handelt, kann man folgende mengentheoretische Operationen anwenden

- Vereinigungsmenge $E = E_1 \cup E_2$: entweder E_1 oder E_2 (oder beide) treten ein (“Ereignis E_1 oder E_2 ”);
- Schnittmenge $E = E_1 \cap E_2$: sowohl E_1 als auch E_2 treten ein (“Ereignis E_1 und E_2 ”);
- Komplementärmenge $\bar{E} := S \setminus E$: alle Elemente von S , die nicht in E enthalten sind (Gegenereignis);

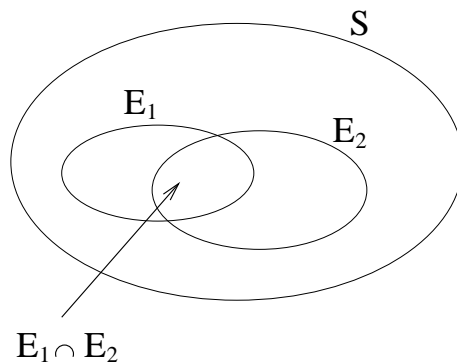


Abbildung 15.1.1: Grafische Darstellung von Mengen mit Hilfe eines Venn-Diagramms am Beispiel der Schnittmengenbildung.

- Differenzmenge $E = E_1 \setminus E_2$: Elemente in E_1 aber nicht in E_2 ;
- sicheres Ereignis: Da $S \subset S$ kann man den Stichprobenraum selbst als Ereignis interpretieren. Per definition ist dann $P(S) = 1$. Man nennt S daher auch das **sichere Ereignis**.
- unmögliches Ereignis: Da $\emptyset \subset S$ entspricht die leere Menge ebenfalls einem Ereignis, dem **unmöglichen Ereignis** mit $P(\emptyset) = 0$.

Wie in der Mengenlehre lassen sich auch in der Wahrscheinlichkeitstheorie grafische Darstellungen von Menge mit Hilfe von **Venn-Diagrammen** (siehe Abb. 15.1.1) nutzbringend anwenden

Ein weiteres nützliches Ergebnis der Mengenlehre sind die **Regeln von de Morgan**,

$$\overline{E_1 \cup E_2} = \bar{E}_1 \cap \bar{E}_2, \quad \overline{E_1 \cap E_2} = \bar{E}_1 \cup \bar{E}_2,$$

zur Bildung von Komplementärmenge.

15.1.1 Kolmogorow-Axiome

Kolmogorow hat ein Axiomensystem für die Wahrscheinlichkeitstheorie vorgeschlagen. Aufbauend auf den oben angegebenen Definitionen lauten die wichtigsten Axiome:

1. Jedes Ereignis E hat eine Wahrscheinlichkeit $P(E) \geq 0$.
2. Das sichere Ereignis S in einem Wahrscheinlichkeitsraum S hat die Wahrscheinlichkeit $P(S) = 1$.
3. Für disjunkte (unvereinbare) Ereignisse (d.h. $E_1 \cap E_2 = \emptyset$) gilt

$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

Aus diesen Axiomen folgen bereits zahlreiche Eigenschaften:

- Die Wahrscheinlichkeit des unmöglichen Ereignisses ist $P(\emptyset) = 0$, denn:

$$1 = P(S) = P(S \cup \emptyset) = P(S) + P(\emptyset) = 1 + P(\emptyset).$$

- Da $E \cup \bar{E} = S$ und $E \cap \bar{E} = \emptyset$, gilt

$$1 = P(E \cup \bar{E}) = P(E) + P(\bar{E})$$

und somit

$$P(\bar{E}) = 1 - P(E).$$

- Für paarweise unabhängige Ereignisse gilt

$$P(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_{j=1}^n P(E_j).$$

- Sind die Ereignisse nicht unabhängig, so gilt

$$P(E_1 \cup E_2) \leq P(E_1) + P(E_2).$$

- Ist $E_1 \subset E_2$, so ist $P(E_1) \leq P(E_2)$.

- Ist $E_1 \cap E_2 \neq \emptyset$, so gilt:

$$P(E_1 \cup E_2) \leq P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

15.1.2 Bedingte Wahrscheinlichkeiten

Definition 15.1.6 (bedingte Wahrscheinlichkeit).

Die **bedingte Wahrscheinlichkeit** oder **konditionale Wahrscheinlichkeit** $P(E_1|E_2)$ (gesprochen: “ $P(E_1$ wenn E_2)”) ist definiert als die Wahrscheinlichkeit, mit der E_1 eintritt, falls E_2 eingetreten ist (mit $E_1, E_2 \subset S$).

Später werden wir noch eine mathematische Definition von $P(E_1|E_2)$ geben.

Ein einfaches Beispiel ist die Wahrscheinlichkeit, dass es regnet, wenn man keinen Regenschirm mitgenommen hat oder die Wahrscheinlichkeit, eine Sechs zu würfeln, wenn die Augenzahl gerade ist.

Wir nennen einige offensichtliche Eigenschaften der bedingten Wahrscheinlichkeit.

- $P(E|S) = P(E)$, da S das sichere Ereignis ist.
- Sind E_1 und E_2 Ereignisse, die sich ausschließen, so ist $P(E_1|E_2) = 0$.

- Ist $E_2 \subset E_1$, so ist $P(E_1|E_2) = 1$.

Ein Beispiel hierfür ist die Wahrscheinlichkeit eine gerade Zahl zu würfeln, wenn man eine Sechs gewürfelt hat.

Diese intuitiven Eigenschaften führen zu folgender mathematischer Definition der bedingten Wahrscheinlichkeit :

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)},$$

falls $P(E_2) \neq 0$. Diese Definition erfüllt die obigen Eigenschaften, wie man leicht nachprüft.

Beispiel 15.1.2. Als Beispiel berechnen wir die Wahrscheinlichkeit, die Augenzahl Zwei zu würfeln, falls die gewürfelte Augenzahl gerade ist:

$$P(\text{“Augenzahl 2”}|\text{“Augenzahl gerade”}) = \frac{\text{“Augenzahl 2 und Augenzahl gerade”}}{\text{“Augenzahl gerade”}} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Dieses Ergebnis erwartet man natürlich anschaulich. Wenn die Augenzahl gerade ist, so kommen nur 2, 4 oder 6 in Frage. Da alle mit gleicher Wahrscheinlichkeit auftreten, beträgt diese für jede der drei Ereignisse $1/3$.

Aus der Definition der bedingten Wahrscheinlichkeit erhält man eine andere Darstellung für die Wahrscheinlichkeit des gleichzeitigen Eintretens der Ereignisse E_1 und E_2 :

$$P(E_1 \cap E_2) = P(E_1|E_2)P(E_2) = P(E_2|E_1)P(E_1).$$

Satz 15.1.1 (Satz von Bayes).

Wir zerlegen den Wahrscheinlichkeitsraum S in sich einander ausschließende¹ Ereignisse E_j : $S = \dot{\cup}_j E_j$. Dann gilt für ein beliebiges Ereignis A :

$$\begin{aligned} P(A) &= P(A \cap S) = P(A \cap (\dot{\cup}_j E_j)) = P(\cup_j (A \cap E_j)) = \sum_j P(A \cap E_j) \\ &= \sum_j P(E_j)P(A|E_j). \end{aligned}$$

Hieraus erhalten wir den **Satz von Bayes**:

$$P(E_l|A) = \frac{P(A|E_l)P(E_l)}{\sum_j P(A|E_j)P(E_j)}.$$

¹Dabei soll der Punkt über dem Vereinigungszeichen \cup andeuten, dass $E_j \cap E_l = \emptyset$ für alle j, l .

Definition 15.1.7 (statistische Unabhängigkeit).

Zwei Ereignisse E_1 und E_2 heißen **statistisch unabhängig**, wenn das Eintreten von E_1 die Wahrscheinlichkeit des Eintreffens von E_2 nicht beeinflusst, oder umgekehrt. Somit gilt:

$$P(E_2|E_1) \stackrel{!}{=} P(E_2) \quad \text{oder} \quad P(E_1|E_2) \stackrel{!}{=} P(E_1).$$

Hieraus folgt:

$$\boxed{P(E_1 \cap E_2) = P(E_1)P(E_2) \quad \text{für unabhängige Ereignisse.}}$$

Bemerkung. Hat man es mit mehreren paarweise statistisch unabhängigen Ereignissen E_j zu tun, so gilt

$$P(\cup_j E_j) = \prod_j P(E_j).$$

Sind E_1 und E_2 statistisch unabhängig, so gilt außerdem:

$$\begin{aligned} P(E_1 \cup E_2) &= P(E_1) + P(E_2) - P(E_1 \cap E_2) \\ &= P(E_1) + P(E_2) - P(E_1)P(E_2). \end{aligned}$$

15.1.3 Gleichverteilte Elementarwahrscheinlichkeiten und Kombinatorik

Wir betrachten den in der Praxis wichtigen Fall gleichwahrscheinlicher (gleichverteilter) Elementarereignisse genauer. Das Standardbeispiel hierfür ist der perfekte Würfel, bei dem alle Zahlen jeweils mit der Wahrscheinlichkeit $1/6$ auftreten.

Wir betrachten konkret einen Wahrscheinlichkeitsraum $S = \{A_1, A_2, \dots, A_n\}$, bei dem die Elementarereignisse $A \in S$ mit gleicher Wahrscheinlichkeit

$$P(A) = \frac{1}{|S|} = \frac{1}{n} =: p$$

aufzutreten. Dabei bezeichnet $|S|$ die Anzahl der Elemente der Menge S . Für ein beliebiges Ereignis $E \subset S$ gilt dann:

$$\boxed{P(E) = \sum_{A \in E} p = |E|p = \frac{|E|}{|S|} = \frac{\text{Anzahl günstiger Ereignisse}}{\text{Anzahl möglicher Ereignisse}},}$$

wobei mit der ‘Anzahl günstiger Ereignisse’ die Ereignisse in E gemeint sind.

Beispiel 15.1.3. Als Beispiel betrachten wir wieder einen Würfel:

$$P(\text{“ungerade Augenzahl”}) = P(\{1, 3, 5\}) = \frac{|\{1, 3, 5\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{3}{6} = \frac{1}{2}.$$

Analoge Betrachtungen kann man z.B. auch für Lotto-Ziehungen anstellen.

Im Fall gleichverteilter Elementarwahrscheinlichkeiten lässt sich die Berechnung von Wahrscheinlichkeiten i.a. auf kombinatorische Probleme (“Abzählregeln”) zurückführen. Wir stellen hier die wichtigsten zusammen.

- **Variationen:** Die Anzahl der Möglichkeiten, aus n unterschiedlichen Zeichen Worte mit k Zeichen zu bilden, wobei Wiederholungen erlaubt sind:

$$V_k = n^k .$$

So gibt es z.B. $10^2 = 100$ zweistellige Dezimalzahlen (mit ‘00’).

Ohne Wiederholungen gibt es

$$\tilde{V}_k = \frac{n!}{(n-k)!}$$

verschiedene Möglichkeiten.

- **Permutationen:** Alle Anordnungen von n verschiedenen Elementen. Ihre Anzahl beträgt

$$P_n = n! .$$

- **Permutationen von Gruppen:** Alle unterschiedlichen Anordnungen von insgesamt n Elementen, die in k Untergruppen von jeweils n_j gleichen Elementen unterteilt sind. Ihre Anzahl beträgt

$$P_{n,k} = \frac{n!}{\prod_{j=1}^k n_j!} .$$

So gibt es $\frac{5!}{2!2!1!} = 30$ unterschiedliche Möglichkeiten, die fünf Buchstaben $\{a, a, b, b, c\}$ anzuordnen.

- **Kombinationen:** Auswahlmöglichkeiten von k Elementen aus einer Menge von n Elementen (ohne Wiederholungen!). Ihre Anzahl beträgt

$$K_k = \binom{n}{k} .$$

Beim Lotto (“6 aus 49”) gibt es demnach $\binom{49}{6}$ verschiedene Möglichkeiten.

15.2 Zufallsvariablen und Wahrscheinlichkeitsverteilungen

Zur Motivation betrachten wir zunächst einige Beispiele.

- **Glückspiel:** Eine Münze wird geworfen. Bei “Kopf” erhalten wir einen Euro, bei “Zahl” müssen wir einen Euro bezahlen. Wie sieht die Gewinn-Verlust-Verteilung nach N Würfeln aus?

- **Random Walk** (Zufallsweg): Ein Teilchen bekommt pro Zeiteinheit einen zufälligen Stoß nach links oder rechts und bewegt sich daher zufällig jeweils mit Wahrscheinlichkeit $1/2$ um eine Längeneinheit nach links oder rechts (Brownsche Bewegung, Diffusion). Wie sieht die Ortsverteilung nach N Schritten aus?

Definition 15.2.1 (Zufallsvariable).

Eine **Zufallsvariable** X ist eine Abbildung von Ereignissen eines Ereignisraums S auf die reellen Zahlen:

$$X : S \rightarrow \mathbb{R}, \quad E \mapsto X(E)$$

mit der Wertemenge $X(E) = \{x_1, \dots, x_n\} =: W$.

In unseren Beispielen ist jeweils $S = \{(E_1, \dots, E_N) \mid E_j = \pm 1\}$ und $X(E_1, \dots, E_N) = \sum_{j=1}^N E_j$. Beim Münzwurf haben wir die Identifikation “Kopf” = $+1$ und “Zahl” = -1 gezählt, während beim Random Walk $+1$ für einen Schritt nach rechts und -1 für einen Schritt nach links steht. Die Größe $\sum_{j=1}^N E_j$ entspricht dann dem Gewinn/Verlust nach N Spielen bzw. der Teilchenposition nach N Schritten.

Wir sehen an diesen Beispielen, dass X auch wieder eine Zufallsgröße ist! Wir können nun nach den Eigenschaften dieser Zufallsgröße fragen.

Definition 15.2.2 (Wahrscheinlichkeitsverteilung).

Es sei $f(x)$ die Wahrscheinlichkeit, mit der die Zufallsgröße X den Wert $x \in W$ annimmt. Die Funktion $f : W \rightarrow [0, \infty[$ heißt dann die **Wahrscheinlichkeitsverteilung** von X (oder kurz: Verteilung).

Die Wahrscheinlichkeitsverteilung kann anhand der Wahrscheinlichkeiten für die Ereignisse $E \subset S$ bestimmt werden: Offenbar nimmt X den Wert x immer genau dann an, wenn der Versuch mit einem Ergebnis aus dem Urbild $E_X = X^{-1}(\{x\})$ von x unter X endet. E_X ist Teilmenge von S und demnach ein Ereignis, dessen Wahrscheinlichkeit genau $f(x)$ ist. In Formeln ausgedrückt bedeutet dies:

$$f(x) = P(E_X) = P(X^{-1}(\{x\})) .$$

Wir illustrieren diese abstrakten Überlegungen an unserem konkreten Beispiel vom Anfang des Kapitels.

Beispiel 15.2.1. Im Fall $N = 2$ ist der Ereignisraum $S = \{--, -+, +-, ++\}$, wobei $+$ und $-$ Kurzschreibweisen für $+1$ und -1 sind (die wiederum für “Kopf” oder Zahl bzw. “Schritt nach rechts” oder “Schritt nach links stehen!). Alle vier Ereignisse sind gleichwahrscheinlich und treten mit der Wahrscheinlichkeit $1/4$ auf.

Die Zufallsgröße X (Gewinn/Verlust bzw. Position) ist gegeben durch

$$X(--)= -2 =: x_1, \quad X(-+)= X(+)= 0 =: x_2, \quad X(++)= +2 =: x_3 .$$

Die Wahrscheinlichkeitsverteilung

$$\begin{aligned} f(x_1) &= P(\{--\}) = \frac{1}{4}, \\ f(x_2) &= P(\{-+, +-\}) = \frac{1}{2}, \\ f(x_3) &= P(\{++\}) = \frac{1}{4}, \end{aligned}$$

ist dann nicht gleichverteilt, im Unterschied zu den Elementarwahrscheinlichkeiten!

Bemerkung. Sind die Elementarwahrscheinlichkeiten $P(E)$ mit $E \in S$ normiert, d.h.

$$\sum_{E \in S} P(E) = 1,$$

so gilt dies auch automatisch für die Wahrscheinlichkeitsverteilung, denn

$$\sum_{x \in W} f(x) = \sum_{x \in W} P(X^{-1}(\{x\})) = P\left(\bigcup_{x \in W} X^{-1}(\{x\})\right) = P(S) = 1.$$

Dabei haben wir im zweiten Schritt die Tatsache ausgenutzt, dass die Urbilder aller $x \in W$ disjunkt sind und im dritten Schritt, dass sie ganz S überdecken.

Wir fassen noch einmal das Wesentliche zusammen:

Eine diskrete Zufallsgröße X mit Werten $W \subset \mathbb{R}$ wird durch eine normierte ($\sum_x f(x) = 1$), nicht-negative Wahrscheinlichkeitsverteilung $f : W \rightarrow \mathbb{R}$ beschrieben. Dabei ist $f(x)$ die Wahrscheinlichkeit, dass X bei einem Versuch den Wert x annimmt.

Definition 15.2.3 (Verteilungsfunktion).

Die **Verteilungsfunktion** $F_X(t)$ ist definiert als die Wahrscheinlichkeit, mit der die Zufallsvariable X einen Wert $x \leq t$ annimmt:

$$F_X(t) := P(X \leq t).$$

Per Definition ist $F_X(t)$ also monoton wachsend!

Hiermit kann man Wahrscheinlichkeiten für Teilintervalle bestimmen:

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F_X(x_2) - F_X(x_1).$$

Haben wir es nur mit einer Zufallsvariablen zu tun, so dass keine Verwechslungsgefahr besteht, so schreiben wir F statt F_X .

Wir haben die Funktion $F(x)$ so definiert, dass sie rechtsseitig stetig ist: $\lim_{\epsilon \rightarrow 0} F(x+\epsilon) = F(x)$. Für eine diskrete Zufallsvariable, die die Werte x_j mit Wahrscheinlichkeit P_j annimmt, ist F eine Stufenfunktion:

$$F(x) = \sum_{x_j \leq x} P_j \quad \text{und} \quad \sum_j P_j = 1.$$

Stetige Zufallsvariablen nehmen beliebige, meist kontinuierliche reelle Werte an. Für die Wahrscheinlichkeitsdichte (Verteilungsdichte) $f(x) > 0$ gilt:

$$F(x) = \int_{-\infty}^x f(t)dt \quad \text{mit} \quad F(\infty) = \int_{-\infty}^{\infty} f(t)dt = 1.$$

Ist $F(x)$ differenzierbar, so können wir diese Beziehung umkehren:

$$f(x) = \frac{dF}{dx}.$$

15.2.1 Erwartungswerte und Momente

Oft genügt es, statt der gesamten Wahrscheinlichkeitsverteilung nur gewisse typische Parameter zu kennen, die diese charakterisieren.

Definition 15.2.4 (Mittelwert, Erwartungswert, Momente).

Der **Mittelwert** der Verteilung $f(x)$ ist definiert durch

$$\langle X \rangle := \sum_j P_j x_j = \sum_{x_j \in W} f(x_j) x_j,$$

wobei P_j die Wahrscheinlichkeit ist, mit der die Zufallsvariable X den Wert x_j annimmt. Statt $\langle x \rangle$ schreibt man oft auch \bar{x} .

Für eine kontinuierliche Verteilung ist

$$\langle X \rangle := \int_{-\infty}^{\infty} x f(x) dx.$$

Die Verallgemeinerung hiervon ist der **Erwartungswert** einer Funktion $g(X)$ der Zufallsvariablen X :

$$\langle g(X) \rangle := \sum_j P_j g(x_j) \quad \text{bzw.} \quad \langle g(X) \rangle := \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Ein einfaches Beispiel ist $\langle 1 \rangle = \int_{-\infty}^{\infty} f(x) = 1$ auf Grund der Normierung. Der Erwartungswert der Funktion $g(X) = X$ ist gerade der Mittelwert der Verteilung.

Als **n -tes Moment** der Verteilung bezeichnet man den Erwartungswert μ_n von X^n :

$$\mu_n := \langle X^n \rangle.$$

Das erste Moment $\mu_1 =: \mu$ entspricht dem Mittelwert.

Zur Charakterisierung der Verteilung ist die **Varianz** sehr nützlich, die man auch als **Streuung** oder **mittlere quadratische Abweichung** bezeichnet:

$$\text{Var}(X) := \sigma^2 := \langle (X - \langle X \rangle)^2 \rangle = \int_{-\infty}^{\infty} f(x)(x - \langle x \rangle)^2 dx.$$

Die letzte Identität gilt für kontinuierliche Verteilungen und ist für diskrete Verteilungen durch $\sum_j f(x_j)(x_j - \langle X \rangle)^2$ zu ersetzen. $\sigma := \sqrt{\sigma^2}$ heißt **Standardabweichung**. Die Varianz läßt sich auch anders berechnen:

$$\begin{aligned} \langle (X - \langle X \rangle)^2 \rangle &= \langle (X - \mu)^2 \rangle = \langle (X^2 - 2\mu X + \mu^2) \rangle = \langle X^2 \rangle - 2\mu \langle X \rangle + \mu^2 = \langle X^2 \rangle - \mu^2 \\ &= \langle X^2 \rangle - \langle X \rangle^2. \end{aligned}$$

Definition 15.2.5 (Charakteristische Funktion).

Die **charakteristische Funktion** ϕ_X einer Verteilungsdichte $f_X(x)$ ist definiert durch

$$\phi_X(t) := \int_{-\infty}^{\infty} e^{itx} f_X(x) dx = \langle e^{itX} \rangle,$$

also als Erwartungswert von e^{itX} . Man beachte, dass dies gerade der Fouriertransformierten von f_X entspricht².

Die charakteristische Funktion (und damit die Verteilungsfunktion als ihre Fourier(rück)transformierte) sind durch die Momente eindeutig bestimmt, denn

$$\phi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \langle X^n \rangle = 1 + it \langle X \rangle - \frac{1}{2} t^2 \langle X^2 \rangle + \dots$$

Dies gilt natürlich nur, falls die Momente auch existieren (endlich sind).

15.3 Spezielle Wahrscheinlichkeitsverteilungen

15.3.1 Gleichverteilung

Die **Gleichverteilung**³ ist die einfachste Verteilung einer stetigen Zufallsvariablen X : Jeder Wert in einem Intervall $[a, b] \subset \mathbb{R}$ sei gleichwahrscheinlich, d.h. die Wahrscheinlichkeitsverteilung der Gleichverteilung lautet

$$f(x) = \frac{1}{b-a} \quad \text{falls } x \in [a, b].$$

Die zugehörige Verteilungsfunktion, also die Wahrscheinlichkeit, dass die Zufallsvariable einen Wert $\leq x$ annimmt, ist

$$F(x) = \begin{cases} 0 & \text{für } x < a, \\ \frac{x-a}{b-a} & \text{für } a \leq x < b, \\ 1 & \text{für } b \leq x. \end{cases}$$

²Man beachte, dass eine andere Konvention für den Normierungsfaktor als in Kapitel 10.3 benutzt wird.

³englisch: *uniform distribution*

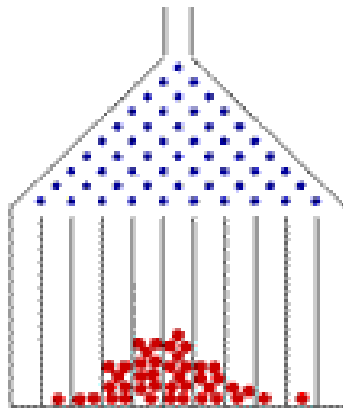


Abbildung 15.3.1: Das Galtonsche Nagelbrett: Kugeln werden oben eingefüllt und auf dem Weg nach unten an Nägeln zufällig nach rechts oder links abgelenkt.

Mittelwert und Varianz der Gleichverteilung sind durch

$$\langle X \rangle = \frac{1}{2}(a + b), \quad \sigma^2 = \frac{1}{12}(b - a)^2$$

gegeben. Die Momente μ_k lassen sich ebenfalls relativ einfach bestimmen.

15.3.2 Binomialverteilung

Die **Binomialverteilung**, manchmal auch als **Bernoulli-Verteilung** bezeichnet, lässt sich relativ einfach durch das sog. **Galtonsche Nagelbrett** realisieren (siehe Abb. 15.3.1). Dabei werden fallende Kugeln durch regelmässig angeordnete Nägel zufällig nach rechts oder links abgelenkt. Die Verteilung der Kugel in den Auffangbehältern wird dann durch die Binomialverteilung beschrieben.

Zur Verallgemeinerung betrachten wir ein Experiment, bei dem das Ereignis entweder ‘0’ (oder “wahr”) oder ‘1’ (oder “falsch”) ist⁴. Die beiden Ereignisse sollen dabei mit den Wahrscheinlichkeiten $P(1) =: p$ und $P(0) = 1 - p =: q$ auftreten. Wiederholt man das Experiment N mal, so gibt es 2^N mögliche Ereignisse, z.B. $(0, 0, 1, 0, \dots)$ etc. Die Wahrscheinlichkeit, bei N Versuchen k -mal den Wert 1 zu bekommen, ist durch die **Binomialverteilung** (Abb. 15.3.2)

$$P(X = k) = \binom{N}{k} p^k (1 - p)^{N-k} =: f_N(k)$$

gegeben, denn $p^k (1 - p)^{N-k}$ ist die Wahrscheinlichkeit für k Einsen und $N - k$ Nullen und $\binom{N}{k}$ die Anzahl der möglichen Reihenfolgen (siehe Kapitel 15.1.3).

⁴Die folgenden Überlegungen gelten für beliebige binäre Entscheidungen!

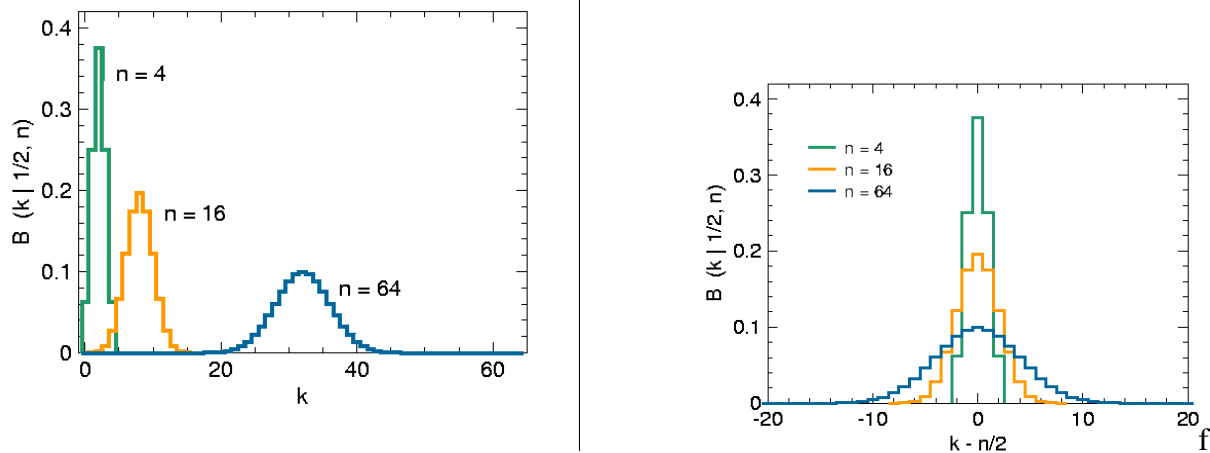


Abbildung 15.3.2: Die Binomialverteilung für den Fall $p = \frac{1}{2}$ und $N = 4, 16, 64$ (links). Die rechte Abbildung zeigt die selben Kurven so verschoben, dass der Mittelwert bei 0 liegt.

Die Verteilung ist bereits normiert, denn nach dem binomischen Satz gilt

$$\sum_{k=0}^N f_N(k) = \sum_{k=0}^N \binom{N}{k} p^k (1-p)^{N-k} = (p + (1-p))^N = 1.$$

Mittelwert und Varianz sind durch

$$\mu = \langle X \rangle = pN, \quad \mu_2 = \langle X^2 \rangle = pN + p^2 N(N-1), \quad \sigma^2 = Np(1-p)$$

gegeben.

Dieses Ergebnis illustriert das **Gesetz der großen Zahlen**, denn

$$\frac{\sigma}{\mu} \sim \frac{1}{\sqrt{N}} \rightarrow 0,$$

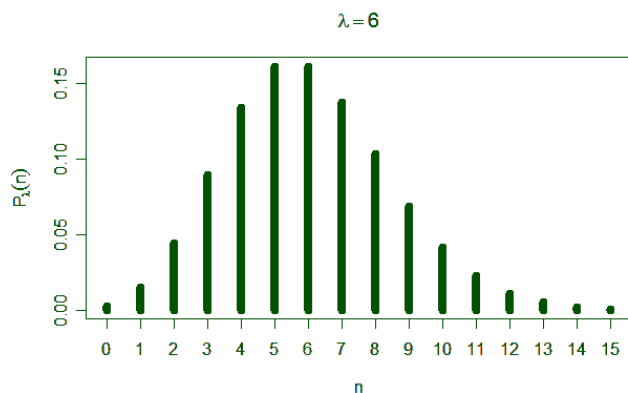
d.h. je mehr Versuche man macht, umso genauer wird das Ergebnis, denn die relative Schwankung $\frac{\sigma}{\mu}$ strebt gegen Null.

Man spricht auch von der **\sqrt{n} -Regel**: Die Anzahl n eingetretener unabhängiger Ereignisse schwankt mit $\sigma = \sqrt{n}$.

Dies wollen wir an einem Beispiel illustrieren.

Beispiel 15.3.1. Wir zählen die Regentropfen, die in fester Zeit auf eine Fläche A fallen. Ihre Anzahl n ist bei verschiedenen Versuchen unterschiedlich und schwankt um den Mittelwert $\langle n \rangle$. Es stellt sich die Frage, wie groß diese Schwankung σ ist!

Wir betrachten daher $N \gg n$ Tropfen, die mit Wahrscheinlichkeit $p \ll 1$ die Fläche treffen. Die Tropfenzahl n ist dann die binomialverteilte Zufallsgröße $n = \sum_{j=1}^N x_j$ mit $x_j = 1$ mit Wahrscheinlichkeit p und $x_j = 0$ mit Wahrscheinlichkeit $1-p$. Der Mittelwert ist dann $\langle n \rangle = pN$ und $\sigma = \sqrt{Np(1-p)} \approx \sqrt{pN} = \sqrt{\langle n \rangle}$ unabhängig von N und p !!

Abbildung 15.3.3: Die Poissonverteilung mit $\mu = 6$.

15.3.3 Poisson-Verteilung

Läßt man in der Binomialverteilung die Zahl der Versuche $N \rightarrow \infty$ gegen Unendlich gehen, hält aber gleichzeitig den Mittelwert μ konstant (d.h. die Einzelwahrscheinlichkeit $p = \frac{\mu}{N}$ strebt gegen Null), dann gilt mit $p = \frac{\mu}{N}$:

$$\begin{aligned} \lim_{N \rightarrow \infty} \binom{N}{k} \left(\frac{\mu}{N}\right)^k \left(1 - \frac{\mu}{N}\right)^{N-k} &= \lim_{N \rightarrow \infty} \binom{N}{k} \frac{\mu^k}{(N - \mu)^k} \left(1 - \frac{\mu}{N}\right)^N \\ &= \lim_{N \rightarrow \infty} \frac{N!}{(N - k)!(N - \mu)^k} \frac{\mu^k}{k!} \underbrace{\left(1 - \frac{\mu}{N}\right)^N}_{\rightarrow e^{-\mu}} \\ &= \frac{\mu^k}{k!} e^{-\mu}. \end{aligned}$$

Dabei haben wir z.B. benutzt, dass

$$\frac{N!}{(N - k)!} = N(N - 1) \cdots (N - k + 1) \approx N^k,$$

falls $N \gg 1$.

Somit erhalten wir in diesem Limes aus der Binomialverteilung die **Poisson-Verteilung**

$$P(X = k) = \frac{\mu^k}{k!} e^{-\mu},$$

wobei nun μ ein freier Parameter ist und $\sigma^2 = \mu$.

Beispiel 15.3.2. Wir betrachten $N \gg 1$ radioaktive Atomkerne. Ein Kern zerfällt mit der Wahrscheinlichkeit $\epsilon \ll 1$ im Zeitintervall Δ . Wenn im Mittel 6 Zerfälle in der Zeit Δ beobachtet werden, mit welcher Wahrscheinlichkeit misst man dann 10 Zerfälle in der Zeit Δ ?

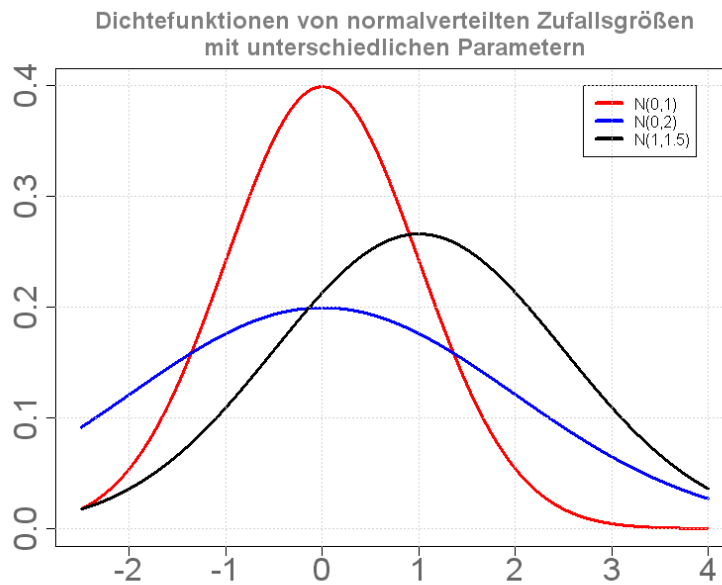


Abbildung 15.3.4: Gauß-Verteilungen mit unterschiedlichen Mittelwerten ($\mu = 0, 1$) und Standardabweichungen ($\sigma = 1, 1.5, 2$).

Im Prinzip wird der Prozeß durch eine Binomialverteilung mit $k = 10$ und $p = \epsilon$ beschrieben. Da $N \gg 1$ ist, können wir diese durch eine Poisson-Verteilung mit $\mu = N\epsilon = 6$ approximieren:

$$P(k = 10) = \frac{6^{10}}{10!} e^{-6} \approx 0.0007 \quad (\approx 1/1400),$$

d.h. im Mittel werden bei einem von 1400 Beobachtungen (der Dauer Δ) 10 Zerfälle gemessen. Man beachte, dass für die obigen Überlegungen weder die Kenntnis von N noch die von ϵ notwendig war!

15.3.4 Normalverteilung (Gauß-Verteilung)

Die **Normalverteilung** oder **Gauß-Verteilung** ist neben der Gleichverteilung die wichtigste Verteilung. Einen Grund hierfür werden wir im nächsten Abschnitt 15.4 kennenlernen. Sie ist gegeben durch (Abb. 15.3.4)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

bzw.

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x dx' \exp\left(-\frac{(x'-\mu)^2}{2\sigma^2}\right) =: \frac{1}{2} + \operatorname{erf}\left(\frac{x-\mu}{\sigma}\right),$$

wobei $\operatorname{erf}(x)$ die **Fehlerfunktion (Error function)** ist.

Das Maximum der Verteilung liegt bei $x = \mu$ und σ bestimmt die Breite der Verteilung. Dabei liegt ein Wert immer mit der Wahrscheinlichkeit 68.3% im Intervall $[\mu - \sigma, \mu + \sigma]$.

Wir können auch die Momente berechnen. Zunächst ist offensichtlich

$$\langle (X - \mu)^{2k+1} \rangle = 0 \quad (k = 0, 1, 2, \dots).$$

Dies sieht man am Einfachsten im Fall $\mu = 0$ (den man nach einer Substitution $x \rightarrow x - \mu$ immer erreichen kann): Hier hat man eine ungerade Funktion über ein symmetrisches Intervall zu integrieren. Für $k = 0$ folgt hieraus

$$\langle X \rangle = \mu,$$

d.h. der Mittelwert stimmt mit dem Maximum überein.

Um die anderen Momente zu bestimmen, definieren wir zunächst die neue Zufallsvariable $Y := \frac{X - \mu}{\sigma}$. Diese ist auch normalverteilt mit Mittelwert $\mu_Y = 0$ und Varianz $\sigma_Y^2 = 1$. Für deren Momente gilt dann ($k = 0, 1, 2, \dots$):

$$\begin{aligned} \langle Y^{2k} \rangle &= \sqrt{12}\pi \int_{-\infty}^{\infty} dy y^{2k} e^{-y^2/2} = \sqrt{12}\pi (-2)^k \left(\frac{\partial^k}{\partial a^k} \int_{-\infty}^{\infty} dy y^{2k} e^{-ay^2/2} \right) \Big|_{a=1} \\ &= (-2)^k \left(\frac{\partial^k}{\partial a^k} \frac{1}{\sqrt{a}} \right) \Big|_{a=1} = (2k - 1)!! := (2k - 1)(2k - 3) \cdots 3 \cdot 1. \end{aligned}$$

Dabei haben wir den Feynman-Trick der Ableitung nach einem Parameter (siehe Kap. 5.3.3) verwendet.

Somit erhalten wir für die geraden Momente der Normalverteilung:

$$\langle (X - \mu)^{2k} \rangle = (2k - 1)!! \sigma^{2k}.$$

Die entsprechenden ungeraden Momente verschwinden. Für $k = 1$ erhalten wir

$$\langle (X - \mu)^2 \rangle = \sigma^2,$$

d.h. der in der Normalverteilung auftretende Parameter σ entspricht tatsächlich der Standardabweichung.

15.4 Zentraler Grenzwertsatz

Wir haben schon betont, dass die Normal- oder Gaußverteilung in der Praxis sehr wichtig ist. Der Grund hierfür ist der folgende sog. **Zentrale Grenzwertsatz**.

Satz 15.4.1 (Zentraler Grenzwertsatz).

Es seien $X_1, X_2, X_3, \dots, X_N$ unabhängige Zufallsvariablen, deren dritten Momente beschränkt

sind, d.h. es gibt ein $C \in \mathbb{R}$ mit $\langle X_j^3 \rangle < C$ für alle j . Die Mittelwerte bzw. Varianzen der Verteilungen seien μ_j bzw. σ_j^2 . Dann ist die Zufallsvariable

$$X := \frac{1}{N} \sum_{j=1}^N X_j$$

im Limes $N \rightarrow \infty$ gaußverteilt mit Mittelwert

$$\mu = \frac{1}{N} \sum_{j=1}^N \mu_j \quad \left(\text{d.h. } \langle X \rangle = \frac{1}{N} \sum_{j=1}^N \langle X_j \rangle \right)$$

und Varianz

$$\sigma^2 = \frac{1}{N} \left(\frac{1}{N} \sum_{j=1}^N \sigma_j^2 \right).$$

Hierzu einige Bemerkungen:

1. In Worten bedeutet dies: Eine additiv aus vielen zufälligen Beiträgen zusammengesetzte Zufallsgröße ist in guter Näherung gaußverteilt.
2. Dabei sollen alle Beiträge substantiell beitragen, nicht nur einige wenige! Dies wird durch die Voraussetzung $\langle X_j^3 \rangle < C$ garantiert, aus der auch die Beschränktheit der 1. Momente folgt.
3. Diese Bedingung garantiert auch die Konvergenz, kann aber sogar noch abgeschwächt werden.
4. Zwei Beispiele für mögliche Anwendungen sind:
 - (i) Die Position eines Random Walkers bzw. binomialverteilter Größen im Allgemeinen (die Binomialverteilung nähert sich für große N und große $Np(1-p)$ einer Normalverteilung an.
 - (ii) Der Mittelwert einer Zufallsgröße aus $N \gg 1$ Messungen.
5. Auf Grund der Linearität des Erwartungswertes ist die Aussage über die Mittelwerte klar. Aber: Die Varianz konvergiert nicht gegen das Mittel $\frac{1}{N} \sum_{j=1}^N \sigma_j^2$ der Varianzen, sondern ist um den Faktor $\frac{1}{N}$ kleiner!! Für gleichverteilte X_j ist die Standardabweichung von X gegenüber der Standardabweichung σ_0 der X_j um einen Faktor \sqrt{N} kleiner. Dies ist eine Variante der \sqrt{N} -Regel.
6. In der Praxis ist die Aussage oft schon für relativ kleine Werte von N ($N \approx 10$) sehr gut erfüllt!
7. In Experimenten setzt sich der Fehler einer Meßgröße aus vielen Beiträgen zusammen, z.B. kaum kontrollierbaren Temperatur- und Druckschwankungen, Erschütterungen oder Ungenauigkeiten im Versuchsaufbau. Den unkontrollierbaren Anteil, den man nicht durch

geeignetes Eichen bestimmen kann, nennt man auch **statistischen Fehler**, im Gegensatz zum **systematischen Fehler**. Nach dem Zentralen Grenzwertsatz erwartet man, dass statistische Fehler gaußverteilt sind.

Zum Beweis des Zentralen Grenzwertsatzes betrachten wir die charakteristische Funktion ϕ der Zufallsgröße

$$\tilde{X} := X - \mu = \frac{1}{N} \sum_{j=1}^N \tilde{X}_j \quad \text{mit } \tilde{X}_j := X_j - \mu_j.$$

Wie die X_j sind auch die Zufallsvariablen \tilde{X}_j statistisch unabhängig und

$$\langle \tilde{X}_j \rangle = 0 \quad \text{und} \quad \langle \tilde{X} \rangle = 0 \quad (j = 1, \dots, N)$$

Die charakteristische Funktion ist dann gegeben durch

$$\begin{aligned} \phi(t) &= \langle e^{it\tilde{X}} \rangle = \int d^N x \tilde{f}(x_1, \dots, x_N) e^{it \frac{1}{N} \sum_{j=1}^N x_j} \\ &= \int d^N x \tilde{f}(x_1) \cdots \tilde{f}(x_N) e^{i \frac{t}{N} x_1} \cdots e^{i \frac{t}{N} x_N} = \prod_{j=1}^N \int dx_j \tilde{f}(x_j) e^{i \frac{t}{N} x_j} \\ &= \prod_{j=1}^N \langle e^{it\tilde{X}_j} \rangle = \prod_{j=1}^N \phi_j \left(\frac{t}{N} \right). \end{aligned}$$

Dabei ist $\tilde{f}(x_1, \dots, x_N)$ die Wahrscheinlichkeitsverteilung von \tilde{X} , die wir beim Übergang zur zweiten Zeile auf Grund der statistischen Unabhängigkeit der \tilde{X}_j als Produkt deren Wahrscheinlichkeitsverteilungen $\tilde{f}(x_j)$ geschrieben haben.

Wir sehen also, dass sich die charakteristische Funktion $\phi(t)$ von \tilde{X} als Produkt der charakteristischen Funktionen ϕ_j der \tilde{X}_j ergibt, allerdings mit reskalierten Argumenten.

Als nächsten Schritt untersuchen wir das Verhalten von $\phi_j \left(\frac{t}{N} \right)$ für große N . Dazu überlegen wir zunächst, dass nach Taylor gilt:

$$e^{i \frac{t}{N} \tilde{X}_j} = 1 + i \frac{t \tilde{X}_j}{N} - \frac{1}{2} \frac{(t \tilde{X}_j)^2}{N^2} - \frac{i e^{i \frac{t}{N} \tilde{X}_j \theta}}{6} \frac{(t \tilde{X}_j)^3}{N^3},$$

wobei wir das Restglied der Taylor-Entwicklung auf Grund der Voraussetzung über die Beschränktheit der 3. Momente abschätzen können durch

$$\left\langle \left| \frac{i e^{i \frac{t}{N} \tilde{X}_j \theta}}{6} \frac{(t \tilde{X}_j)^3}{N^3} \right| \right\rangle = \frac{t^3}{6N^3} \left\langle |\tilde{X}_j|^3 \right\rangle < \tilde{C}.$$

Somit gilt für die charakteristische Funktion von \tilde{X}_j :

$$\begin{aligned}\phi_j\left(\frac{t}{N}\right) &\approx \left\langle 1 + i\frac{t\tilde{X}_j}{N} - \frac{1}{2}\frac{(t\tilde{X}_j)^2}{N^2} \right\rangle \\ &= \underbrace{\langle 1 \rangle}_{=1} + \frac{it}{N} \underbrace{\langle \tilde{X}_j \rangle}_{=0} - \frac{t^2}{2N^2} \underbrace{\langle \tilde{X}_j^2 \rangle}_{=\sigma_j^2} \\ &= 1 - \frac{t^2\sigma_j^2}{2N^2},\end{aligned}$$

wobei wir $\langle \tilde{X}_j^2 \rangle = \langle (X_j - \mu_j)^2 \rangle = \sigma_j^2$ benutzt haben. Hiermit können wir nun die charakteristische Funktion $\phi(t)$ von \tilde{X} für große N abschätzen:

$$\begin{aligned}\phi(t) &= \prod_{j=1}^N \phi_j\left(\frac{t}{N}\right) = \prod_{j=1}^N \left(1 - \frac{t^2\sigma_j^2}{2N^2}\right) + \mathcal{O}\left(\frac{1}{N^3}\right) \\ &= 1 - \sum_{j=1}^N \frac{t^2\sigma_j^2}{2N^2} + \mathcal{O}\left(\frac{1}{N^3}\right) = 1 - \frac{t^2\bar{\sigma}^2}{2N} + \mathcal{O}\left(\frac{1}{N^3}\right) \\ &= e^{-\frac{\bar{\sigma}^2 t^2}{2N}} + \mathcal{O}\left(\frac{1}{N^3}\right),\end{aligned}$$

wobei wir die Abkürzung $\bar{\sigma}^2 := \frac{1}{N} \sum_{j=1}^N \sigma_j^2$ eingeführt haben.

Durch Fourier-Transformation erhalten wir aus $\phi(t)$ die zugehörige Wahrscheinlichkeitsverteilung $\tilde{f}(x)$:

$$\tilde{f}(x) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2/N}} e^{-\frac{x^2}{2\bar{\sigma}^2/N}}.$$

Dies ist eine Gaußverteilung mit Standardabweichung $\sigma = \frac{\bar{\sigma}}{\sqrt{N}}$. Die Verteilung von $X = \tilde{X} + \mu$ ist dann durch

$$f(x) = \tilde{f}(x - \mu)$$

gegeben und somit ebenfalls eine Gaußverteilung, wie im Zentralen Grenzwertsatz behauptet.

Kapitel 16

Funktionentheorie

Die **Funktionentheorie** beschäftigt sich mit der Analysis von Funktionen $f(z)$ einer komplexen Variablen z . Die englische Bezeichnung *complex variable theory* ist etwas präziser.

Im Prinzip handelt es sich um einen Spezialfall der Analysis im \mathbb{R}^2 , denn über

$$f(z) = u(x, y) + iv(x, y) \quad \text{mit } z = x + iy, \quad u, v \text{ reellwertig}$$

läßt sich das Problem auf die Untersuchung zweier reellwertiger Funktionen u, v der reellen Variablen x, y zurückführen.

Allerdings haben die komplexen Zahlen \mathbb{C} mehr Struktur als der \mathbb{R}^2 . Insbesondere können wir durch komplexe Zahlen dividieren, während dies für Elemente des \mathbb{R}^2 nicht geht. Dies hat weitreichende Konsequenzen.

16.1 Komplexe Differenzierbarkeit

Zunächst wollen wir feststellen, dass sich viele Begriffe und Konzepte der reellen Analysis in natürlicher Weise auf die komplexen Zahlen übertragen lassen, z.B. die Stetigkeit.

Bei der Differenzierbarkeit ergeben sich aber auf Grund der oben erwähnten zusätzlichen Struktur von \mathbb{C} neue Aspekte!

Definition 16.1.1 (komplexe Differenzierbarkeit).

Eine Funktion f heißt **komplex differenzierbar** im Punkt $z \in \mathbb{C}$, falls

$$\lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} =: \frac{df}{dz}(z) =: f'(z)$$

existiert.

Ist f in einem Gebiet¹ $G \subset \mathbb{C}$ komplex differenzierbar (und eindeutig), so heißt f **analytisch** oder auch **holomorph** oder **regulär** in G .

¹Ein Gebiet G ist eine offene zusammenhängende Menge.

Beispiel 16.1.1. Wir betrachten zwei Beispiele.

1. Die Funktion $f(z) = z^n$ ist holomorph, denn

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)^n - z^n}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{z^n + n z^{n-1} \Delta z + \mathcal{O}((\Delta z)^2) - z^n}{\Delta z} = n z^{n-1}.$$

Dabei haben wir, wie im reellen Fall, den binomischen Satz verwendet.

2. Wir betrachten die Funktion

$$f(z) = |z|^2 = z\bar{z} = x^2 + y^2.$$

Diese ist natürlich reell differenzierbar, wie man an der letzten Darstellung sieht. Ist sie aber auch komplex differenzierbar? Dazu betrachten wir für $\Delta z = \epsilon e^{i\varphi}$ (mit $\epsilon \in \mathbb{R}$) den Grenzwert

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{(z + \Delta z)^n - z^n}{\Delta z} &= \lim_{\epsilon \rightarrow 0} \frac{(z + \epsilon e^{i\varphi})(\bar{z} + \epsilon e^{-i\varphi}) - z\bar{z}}{\epsilon e^{i\varphi}} \\ &= \lim_{\epsilon \rightarrow 0} \frac{z\bar{z} + z\epsilon e^{-i\varphi} + \bar{z}\epsilon e^{i\varphi} + \epsilon^2 - z\bar{z}}{\epsilon e^{i\varphi}} \\ &= z e^{-2i\varphi} + \bar{z}. \end{aligned}$$

Dieser hängt also von der Richtung φ ab! Daher existiert der Grenzwert $\lim_{\Delta z \rightarrow 0} \dots$ nicht (für alle z). Daher ist $f(z) = |z|^2$ nicht komplex differenzierbar und zwar für $z \in \mathbb{C}$, nicht nur für $z = 0$.

16.1.1 Cauchy-Riemann-Bedingungen

Wir wollen nun einen Zusammenhang zwischen der reellen und der komplexen Differenzierbarkeit herstellen. Dazu schreiben wir unsere komplexe Funktion wieder in der Form

$$f(z) = u(x, y) + iv(x, y)$$

mit $z = x + iy$ und reellwertigen Funktionen u und v .

Wir berechnen nun die partiellen Ableitungen nach x und y :

$$\begin{aligned} \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} &= \frac{\partial f}{\partial x} = \frac{df}{dz} \frac{\partial z}{\partial x} = \frac{df}{dz} \frac{\partial(x + iy)}{\partial x} = \frac{df}{dz}, \\ \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} &= \frac{\partial f}{\partial y} = \frac{df}{dz} \frac{\partial z}{\partial y} = \frac{df}{dz} \frac{\partial(x + iy)}{\partial y} = i \frac{df}{dz}, \end{aligned}$$

wobei wir jeweils im dritten Schritt die Kettenregel verwendet haben.

Aus dem Vergleich dieser beiden Gleichungen folgt nun

$$i \left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) = \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y}$$

und somit, da u und v reell sind, die sog. **Cauchy-Riemann-Bedingungen** oder auch **Cauchy-Riemann-DGL**

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.}$$

Man sieht sofort, dass u und v die Laplace-Gleichung erfüllen müssen, d.h.

$$\Delta u = 0, \quad \Delta v = 0.$$

Als Faustregel kann man festhalten, dass analytische Funktionen kein \bar{z} enthalten dürfen, denn es gilt mit $x = \frac{1}{2}(z + \bar{z})$ und $y = \frac{1}{2i}(z - \bar{z})$:

$$\begin{aligned} \frac{\partial f}{\partial \bar{z}} &= \frac{\partial u}{\partial x} \frac{\partial x}{\partial \bar{z}} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial \bar{z}} + i \left(\frac{\partial v}{\partial x} \frac{\partial x}{\partial \bar{z}} + \frac{\partial v}{\partial y} \frac{\partial y}{\partial \bar{z}} \right) \\ &= \frac{1}{2} \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) + \frac{i}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) = 0, \end{aligned}$$

wobei im letzten Schritt die Cauchy-Riemann-Bedingungen verwendet wurden.

16.1.2 Potenzreihen

Eine in einem Gebiet G holomorphe Funktion f ist dort sogar beliebig oft differenzierbar und kann durch eine Potenzreihe

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$$

dargestellt werden, die man auch als **Taylor-MacLaurin-Formel** bezeichnet. Die Reihe konvergiert in einem Kreis vom Radius R um z_0 .

Ist f in G analytisch, so existieren alle Ableitungen $f^{(n)}(z_0)$ und daher kann f um jeden Punkt in eine Potenzreihe entwickelt werden. Deren Konvergenz ist aber nur im Inneren von G garantiert, am Rand kann es Probleme geben.

Beispiel 16.1.2. Als Beispiel betrachten wir die Funktion

$$f(z) = \frac{1}{1+z^2} = 1 - z^2 + z^4 - z^6 + \dots,$$

wobei sich die Potenzreihendarstellung aus der bekannten geometrischen Reihe ergibt. Die Reihe konvergiert daher aber nur für $|z| < 1$, obwohl z.B. $f(z = \pm 1) = \frac{1}{2}$ wohldefiniert ist. Der Grund sind die Singularitäten von f bei $z = \pm i$, weshalb der Konvergenzradius nicht größer als 1 sein kann.

16.2 Komplexe Wegintegrale und Integralsatz von Cauchy

16.2.1 Komplexe Wegintegrale

Komplexe Wegintegrale sind analog zu den Wegintegralen im \mathbb{R}^2 definiert (siehe Kapitel 8.1). Haben wir eine Parametrisierung $z(t): [a, b] \rightarrow \mathbb{C}$ des Weges γ , längs dem wir integrieren wollen, so können wir das komplexe Wegintegral folgendermaßen berechnen:

$$\int_{\gamma} f(z) dz = \int_a^b f(z(t)) z'(t) dt.$$

Die bekannten Rechenregeln übertragen sich, es ist z.B.

$$\int_{\gamma_1 + \gamma_2} f(z) dz = \int_{\gamma_1} f(z) dz + \int_{\gamma_2} f(z) dz$$

etc.

Die wichtigsten Parametrisierungen sind wieder Geraden und Kreisbögen:

1. Gerade von z_0 nach z_1 : $z(t) = z_0 + t(z_1 - z_0)$ mit $t \in [0, 1]$.
2. Kreisbogen vom Radius r um z_0 : $z(t) = z_0 + r e^{it}$ mit $t \in [a, b]$.
Für einen Vollkreis ist z.B. $a = 0$ und $b = 2\pi$.

Über die Zerlegung $f(z) = u(x, y) + iv(x, y)$ können wir einen Zusammenhang mit reellen Integralen herstellen. Wegen

$$dz = \left(\frac{dx}{dt} + i \frac{dy}{dt} \right) dt =: (x' + iy') dt$$

gilt:

$$\begin{aligned} \int_{\gamma} f(z) dz &= \int_a^b (u + iv)(x' + iy') dt = \int_a^b (ux' - vy')(x' + iy') dt + i \int_a^b (uy' + vx') dt \\ &= \int_{\gamma} (u dx - v dy) + i \int_{\gamma} (u dy + v dx). \end{aligned}$$

Damit läßt sich das komplexe Wegintegral durch zwei reelle Wegintegrale ausdrücken.

16.2.2 Integralsatz von Cauchy

Satz 16.2.1 (Integralsatz von Cauchy).

Es sei γ ein geschlossener Weg (ohne Kreuzungen) und f sei analytisch im von γ eingeschlossenen einfach-zusammenhängenden Gebiet G . Dann gilt der **Integralsatz von Cauchy**:

$$\oint_{\gamma} f(z) dz = 0$$

Bemerkungen:

1. Dies ist das zentrale Resultat der Funktionentheorie, aus dem viele weitere Aussagen folgen!

2. Die Aussage gilt in einem beliebigen Gebiet G für alle null-homologen Wege γ . Dies sind Wege, die für alle $z_0 \notin G$ die Umlaufzahl 0 haben.

Beweis: Auf Grund der reellen Darstellung, die wir am Ende von Kapitel 16.2.1 abgeleitet haben, gilt:

$$\int_{\gamma} f(z)dz = \int_{\gamma} (udx - vdy) + i \int_{\gamma} (udy + vdx).$$

Mit Hilfe des Green'schen Satz mit $p = u$ und $q = -v$ (siehe Kapitel 9.2) gilt dann:

$$\int_{\gamma} (udx - vdy) = \int_F dx dy \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) = 0,$$

wobei der Integrand des Flächenintegrals auf Grund der Cauchy-Riemann-Bedingungen verschwindet. Analog folgt

$$\int_{\gamma} (udy + vdx) = 0,$$

womit der Integralsatz bewiesen ist.

Die vielleicht wichtigste Anwendung des Integralsatzes von Cauchy besteht in der Deformation von Integrationswegen. Er erlaubt nämlich, die Integration längs eines Weges γ_1 durch die Integration längs eines anderen Weges γ_2 zu ersetzen, wenn beide im Gebiet G verlaufen (siehe Abb. 16.2.1). Dies ist möglich, da das Integral längs des kombinierten (geschlossenen!) Weges $\gamma_1 - \gamma_2 = \gamma_1 + (-\gamma_2)$ verschwindet:

$$\int_{\gamma_1} f(z)dz = \oint_{\gamma_1 - \gamma_2} f(z)dz + \int_{\gamma_2} f(z)dz = \int_{\gamma_2} f(z)dz.$$

Eine solche Umformung ist natürlich dann besonders nützlich, wenn sich das neue Integral leichter berechnen läßt. In der Praxis bedeutet dies, dass man in der Funktionentheorie wirklich mit den beiden oben angegebenen Standardwegen (Gerade, Kreisbogen) auskommt, da sich in sehr vielen Fällen die Berechnung von komplexen Wegintegralen durch Deformation auf einen dieser Fälle (oder eine Kombination von ihnen) zurückführen läßt!

Beispiel 16.2.1. Als konkretes Beispiel betrachten wir das wichtige Integral

$$I_n(a, b) := \oint_{\gamma} \frac{b}{(z-a)^n} dz.$$

Dabei ist γ ein beliebiger geschlossener Weg, der den Punkt a umschließt. Da der Integrand dort eine Polstelle (n -ter Ordnung) hat, handelt es sich eigentlich schon um eine fortgeschrittenere Anwendung!

Als Definitionsbereich G von $f(z)$ wählen wir hier die komplexe Ebene mit einem Schnitt wie in Abbildung 16.2.2 dargestellt. Damit ist G ein einfach-zusammenhängendes Gebiet. Man beachte,

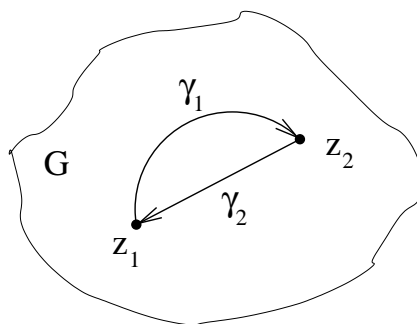


Abbildung 16.2.1: Deformation von Integrationswegen mit Hilfe des Integralsatzes von Cauchy.

dass die Funktion f entlang des Schnittes stetig ist (außer bei $z = a$). Daher spielt es keine Rolle, wenn man aus Wegen, die den Schnitt kreuzen würden, einen Punkt herausnimmt!

Wir wollen den Weg in einen Kreisweg $\tilde{\gamma}$ deformieren. Dies machen wir, wie in Abbildung 16.2.2 dargestellt. Das Integral längs des ganzen Weges $\gamma + \gamma_1 + \gamma_2 - \tilde{\gamma}$ verschwindet auf Grund des Integralsatzes von Cauchy. Die Beiträge der beiden Geradenstücke γ_1 und γ_2 heben sich weg, da der Integrand dort stetig ist und die Wege in entgegengesetzte Richtung durchlaufen werden. Damit gilt:

$$\oint_{\gamma} \frac{b}{(z-a)^n} dz = \oint_{\tilde{\gamma}} \frac{b}{(z-a)^n} dz.$$

Wir parametrisieren den Kreisweg $\tilde{\gamma}$ durch

$$z(t) = a + re^{it} \quad \text{mit } t \in [0, 2\pi],$$

weshalb

$$dz = ire^{it} dt \quad \text{und} \quad (z-a)^n = r^n e^{int}$$

ist. Somit erhalten wir

$$\begin{aligned} I_n(a, b) &= \int_0^{2\pi} \frac{ibre^{it}}{r^n e^{int}} dt = \frac{ib}{r^{n-1}} \int_0^{2\pi} e^{i(n-1)t} dt = \begin{cases} 0 & n \neq 1 \\ 2\pi ib & n = 1 \end{cases} \\ &= 2\pi ib \delta_{n,1}. \end{aligned}$$

Dieses Ergebnis ist sehr wichtig und wird später noch einige Male verwendet.

16.3 Integralformel von Cauchy

Wir betrachten wieder eine Funktion f , die im (einfach-zusammenhängenden) Gebiet G analytisch ist. Außerdem sei $z_0 \in G$ ein beliebiger Punkt in G . Dann können wir f darstellen als

$$f(z) = f(z_0) + (z - z_0)g(z)$$

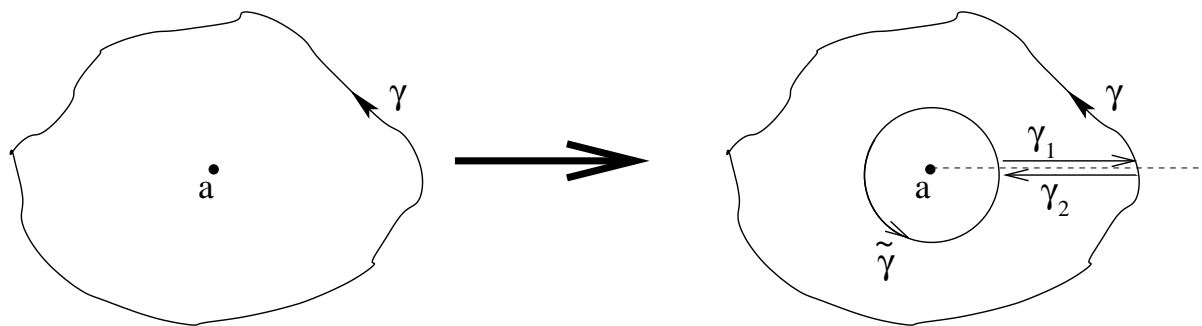


Abbildung 16.2.2: Deformation des Integrationsweges zur Berechnung des Integrals $I_n(a, b)$ in einen Kreisweg. Der Definitionsbereich G von f ist die komplexe Ebene ohne die gestrichelte Gerade.

mit einer analytischen Funktion $g(z)$. Hieraus folgt

$$\frac{f(z)}{z - z_0} = \frac{f(z_0)}{z - z_0} + g(z).$$

Nun integrieren wir diese Identität über den Rand $\gamma := \partial G$ des Gebietes G :

$$\oint_{\gamma} \frac{f(z)}{z - z_0} dz = \oint_{\gamma} \frac{f(z_0)}{z - z_0} dz + \oint_{\gamma} g(z) dz = f(z_0) \oint_{\gamma} \frac{dz}{z - z_0} dz = 2\pi i f(z_0).$$

Dabei haben wir den Integralsatz von Cauchy benutzt, aus dem das Verschwinden des Integrals über die analytische Funktion g folgt. Im letzten Schritt wurde das Ergebnis aus Beispiel 16.2.1 verwendet. Somit haben wir folgendes Ergebnis abgeleitet:

Satz 16.3.1 (Integralformel von Cauchy).

Ist $\gamma := \partial G$ der Rand eines (einfach-zusammenhängenden) Gebietes G , so gilt für analytische Funktionen $f: G \rightarrow \mathbb{C}$ und beliebige Punkte $z_0 \in G$ aus dem Inneren von G die **Integralformel von Cauchy**

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

Dies bedeutet, dass für eine analytische Funktion alle Funktionswerte im Inneren des Analytizitätsbereiches durch die Funktionswerte auf dem Rand bestimmt sind!

Beispiel 16.3.1. Die Funktion $f(z)$ sei analytisch für $|z| \leq 1$ und hat am Kreisrand $c(\varphi)e^{i\varphi}$ die Werte $f(z) = f(e^{i\varphi}) = e^{2i\varphi} - 1$. Dann gilt für z im Inneren des Kreises, d.h. $|z| < 1$, nach der

Integralformel von Cauchy:

$$\begin{aligned}
 2\pi i f(z) &= \oint_c \frac{f(z')}{z' - z} dz' = \int_0^{2\pi} \frac{f(e^{i\varphi})}{e^{i\varphi} - z} i e^{i\varphi} d\varphi \\
 &= i \int_0^{2\pi} \frac{e^{2i\varphi} - 1}{1 - z e^{-i\varphi}} d\varphi i \int_0^{2\pi} e^{2i\varphi} - 1 (1 + z e^{-i\varphi} + z^2 e^{-2i\varphi} + \dots) d\varphi \\
 &= i \int_0^{2\pi} d\varphi (-1 + z^2 + z e^{i\varphi} + \dots) \\
 &= 2\pi i (z^2 - 1).
 \end{aligned}$$

Dabei haben wir im 3. Schritt den Nenner mit Hilfe der geometrischen Reihe entwickelt. Dies geht, da $|z e^{-i\varphi}| < 1$ ist.

Insgesamt sehen wir also, dass die gesuchte Funktion (für $|z| \leq 1$) gegeben ist durch

$$f(z) = z^2 - 1.$$

Durch Ableitung der Integralformel nach z_0 erhält man

$$\frac{d^n f}{dz_0^n}(z_0) = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz,$$

d.h. man kann Ableitung durch ein Integral ausdrücken!

Das Ergebnis aus Beispiel 16.2.1 ist ein Spezialfall dieser verallgemeinerten Integralformel von Cauchy für $f(z) = \frac{b}{(z-a)^n}$.

16.3.1 Laurentreihen

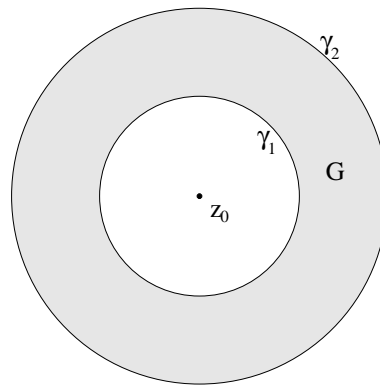
Im Analytizitätsgebiet G kann eine Funktion $f(z)$ durch eine Potenzreihe dargestellt werden, deren Entwicklungspunkt z_0 in G liegt. Manchmal hat man aber die Situation vorliegen, dass f nur analytisch in einem Kreisring mit den Rändern γ_1 und γ_2 um z_0 ist (siehe Abb. 16.3.1).

Dann läßt sich zeigen, dass man $f(z)$ in eine sog. **Laurentreihe** entwickeln kann. Diese hat die Form

$$\begin{aligned}
 f(z) &= \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} \\
 \text{mit } a_n &= \frac{1}{2\pi i} \oint_{\gamma_2} \frac{f(z)}{(z - z_0)^{n+1}} dz, \quad b_n = \frac{1}{2\pi i} \oint_{\gamma_1} f(z) (z - z_0)^n dz.
 \end{aligned}$$

Die erste Summe bezeichnet man auch als **Nebenteil**, die zweite als **Hauptteil**.

Die Laurentreihe konvergiert zwischen γ_1 und γ_2 . Die maximale Größe der Kreise wird dabei durch die Singularitäten von $f(z)$ beschränkt.

Abbildung 16.3.1: Die Funktion f ist analytisch im Kreisring G um z_0 .

Beispiel 16.3.2. Als Beispiel wollen wir Laurentreihen der Funktion $f(z) = \frac{1}{z-2}$ um $z_0 = 0$ untersuchen.

Zunächst betrachten wir das Gebiet $G_1 = \{z \mid |z| < 2\}$, d.h. einen Kreisring mit äußerem Radius 2 und innerem Radius 0. Unter Verwendung der geometrischen Reihe erhalten wir dort

$$\frac{1}{z-2} = -\frac{1}{2} \frac{1}{1-z/2} = -\frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n.$$

Hier stimmt also die Laurentreihe mit der Taylorreihe überein, der Hauptteil verschwindet.

Wir betrachten nun das Gebiet $G_2 = \{z \mid |z| > 2\}$, d.h. einen Kreisring mit innerem Radius 2 und unendlichem äußerem Radius. Dann können wir wieder die geometrische Reihe benutzen, um eine Reihendarstellung von f in G_2 zu erhalten:

$$\frac{1}{z-2} = \frac{1}{z} \frac{1}{1-2/z} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{2}{z}\right)^n = \frac{1}{z} \sum_{n=1}^{\infty} \frac{2^{n-1}}{z^n}.$$

Hierbei war entscheidend, dass $|\frac{2}{z}| < 1$ ist. In G_2 verschwindet also der Nebenteil.

Wir haben hier die geometrische Reihe als Abkürzung verwendet. Die gleichen Ergebnisse hätte man natürlich auch mit der Integraldarstellung der Koeffizienten a_n und b_n bekommen, allerdings wäre die Rechnung aufwändiger gewesen.

16.4 Residuensatz

Definition 16.4.1 (meromorphe Funktion).

Eine Funktion $f: G \rightarrow \mathbb{C}$ heißt **meromorph**, falls sie in G analytisch bis auf endlich viele isolierte Singularitäten (Polstellen) ist.

Satz 16.4.1 (Residuensatz).

Es sei f in einem Gebiet G meromorph und am Rand $\partial G =: \gamma$ analytisch. Dann gilt der **Residuensatz**

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_n \operatorname{Res}(z_n),$$

wobei z_n die Polstellen von f sind und

$$\operatorname{Res}(z_n) = \frac{1}{2\pi i} \oint_{\gamma_n} f(z) dz$$

das **Residuum** von f in z_n . Dabei ist γ_n ein Weg, der außer z_n keine weiteren Singularitäten einschließt. Das Residuum entspricht also gerade dem Koeffizienten b_1 in der Laurent-Reihe von f . Somit verschwindet das Residuum in allen Punkten, in denen f analytisch ist.

Beweis: Zunächst deformieren wir den Weg γ , wie in Abb. 16.4.1 gezeigt. Da sich die Beiträge von den Geradenstücken jeweils wegheben, zerfällt der deformierte Weg in einzelne Wege γ_n , die jeweils nur eine Singularität z_n umschließen. Somit gilt:

$$\oint_{\gamma} f(z) dz = \sum_n \oint_{\gamma_n} f(z) dz.$$

Nun betrachten wir die Laurent-Reihe um z_n :

$$f(z) = \sum_{m=0}^{\infty} a_m (z - z_n)^m + \sum_{m=1}^{\infty} \frac{b_m}{(z - z_n)^m}$$

Da $(z - z_n)^m$ analytisch ist, verschwinden die Beiträge vom Nebenteil:

$$\oint_{\gamma_n} a_m (z - z_n)^m dz = 0.$$

Der Hauptteil liefert einen Beitrag, siehe Beispiel 16.2.1:

$$\oint_{\gamma_n} \frac{b_m}{(z - z_n)^m} dz = 2\pi i b_m \delta_{m,1}.$$

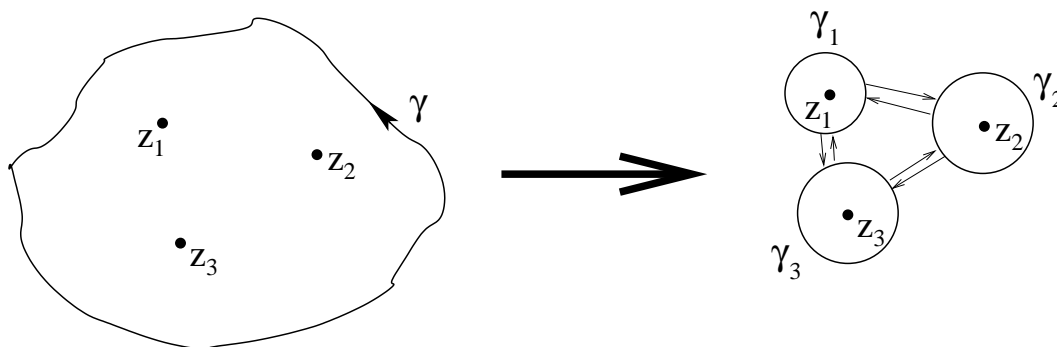


Abbildung 16.4.1: Zum Beweis des Residuensatzes deformieren wir den Weg γ in vielen einzelne Wege γ_n , die jeweils nur eine Singularität z_n umschließen.

Somit folgt

$$\oint_{\gamma_n} f(z) dz = 2\pi i b_1 = 2\pi i \operatorname{Res}(z_n).$$

In der Praxis ist die Berechnung eines Residuums über ein komplexes Wegintegral relativ unhandlich. In vielen Fällen kann es aber einfacher bestimmen.

- Falls f in z_0 einen Pol 1. Ordnung hat, so gilt

$$\operatorname{Res}(z_0) = \lim_{z \rightarrow z_0} (z - z_0) f(z).$$

Als Beispiel betrachten wir die Funktion $f(z) = \frac{e^z}{z}$, die bei $z_0 = 0$ einen einfachen Pol besitzt. Dann gilt

$$\operatorname{Res}(z = 0) = \lim_{z \rightarrow 0} z f(z) = \lim_{z \rightarrow 0} z \frac{e^z}{z} = e^0 = 1.$$

- Falls $f(z)$ von der Form

$$f(z) = \frac{g(z)}{(z - z_0)^n}$$

ist, mit $g(z)$ analytisch bei z_0 , so gilt

$$\operatorname{Res}(z = 0) = \frac{1}{(n-1)!} g^{(n-1)}(z_0) \quad \text{mit } g(z) = (z - z_0)^n f(z).$$

Dies sieht man folgendermaßen: Da g analytisch ist, können wir es durch eine Potenzreihe

$$g(z) = \sum_{m=0}^{\infty} \frac{1}{m!} g^{(m)}(z_0) (z - z_0)^m$$

darstellen. Somit ist $f(z)$ von der Form

$$f(z) = \text{analytische Terme } \mathcal{O}(z - z_0) + \frac{1}{(n-1)!} \frac{1}{z - z_0} g^{(n-1)}(z_0) + \text{Pole höherer Ordnung}.$$

Bilden wir hierüber das Wegintegral längs eines geschlossenen Weges, so fallen die analytischen Terme wegen des Integralsatzes von Cauchy weg und die Pole höherer Ordnung wegen Beispiel 16.2.1.

Als wichtige Anwendung des Residuensatzes wollen zeigen, wie wir mit ihm uneigentliche Integrale der Form $\int_{-\infty}^{\infty} f(x) dx$ berechnen können. Dazu betrachten wir zunächst das komplexe Wegintegral längs des in Abb. 16.4.2 dargestellten geschlossenen Weges C_R . Dieser besteht aus dem reellen Intervall $A_R := [-R, R]$ und einem Halbkreisbogen vom Radius R in der oberen Halbebene. Dann gilt:

$$\oint_{C_R} f(z) dz = \int_{A_R} f(z) dz + \int_{B_R} f(z) dz = \int_{-R}^R f(z) dz + \int_0^{\pi} d\varphi f(Re^{i\varphi}) iRe^{i\varphi}.$$

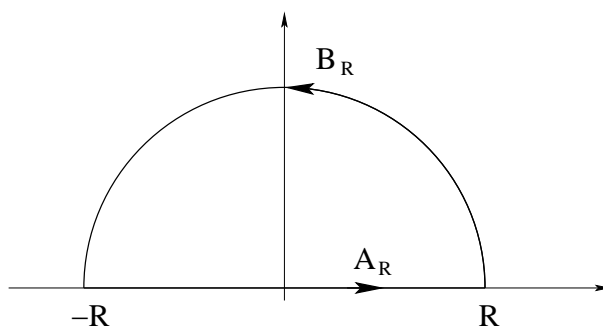


Abbildung 16.4.2: Integrationsweg $C_R = A_R + B_R$ zur Berechnung von reellen uneigentlichen Integralen.

Dabei haben wir die explizite Parametrisierung des Weges B_R verwendet. Falls nun $f(z)$ schnell genug abfällt, d.h. präziser

$$\lim_{R \rightarrow \infty} Rf(Re^{i\varphi}) = 0$$

für alle $\varphi \in [0, \pi]$, so folgt

$$\int_{-\infty}^{\infty} f(x)dx = \lim_{R \rightarrow \infty} \oint_{C_R} f(z)dz.$$

Das Integral auf der rechten Seite lässt sich oft einfach berechnen, z.B. mit Hilfe des Residuensatzes. Dann gilt

$$\int_{-\infty}^{\infty} f(x)dx = 2\pi i \sum_{n, \operatorname{Re}(z_n) > 0} \operatorname{Res}(z_n),$$

wobei nur über die Pole z_n in der oberen Halbebene (d.h. $\operatorname{Re}(z_n) > 0$) zu summieren ist!

Eine analoge Überlegung gilt natürlich, wenn man den Weg in der unteren Halbebene durch einen Halbkreis schließt. Dann tragen natürlich nur die dort liegenden Residuen bei.

Beispiel 16.4.1. Als konkretes Beispiel betrachten wir die Funktion $f(x) = \frac{1}{1+x^2}$, für die gilt

$$\lim_{R \rightarrow \infty} Rf(Re^{i\varphi}) = \lim_{R \rightarrow \infty} \frac{R}{1 + R^2 e^{2i\varphi}} = 0.$$

Daher sind die Voraussetzungen der obigen Betrachtung erfüllt. Wegen

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

hat $f(z)$ zwei Pole $z_{\pm} = \pm i$, von denen aber nur $z_+ = i$ im Inneren des Weges C_R liegt. Das Residuum an diesem Punkt ist durch $\operatorname{Res}(z_+) = \frac{1}{2i}$ gegeben und somit

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = 2\pi i \operatorname{Res}(z_+) = \pi.$$

Das gleiche Ergebnis erhält man natürlich auch durch direkte Integration, denn

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \arctan x \Big|_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

Kapitel 17

Tensorrechnung

Neben Skalaren und Vektoren treten in vielen Bereichen der Physik sog. *Tensoren* auf. In der Experimentalphysik-Vorlesung haben Sie schon den *Trägheitstensor* kennengelernt. Eine besondere Rolle spielen Tensoren aber in allgemeinen Relativitätstheorie.

17.1 Definition

Bevor wir eine genaue Definition geben, wollen wir ein Beispiel betrachten.

Beispiel 17.1.1. Wir betrachten die lineare Vektorfunktion $\underline{b} = T\underline{a}$, bei der jedem Vektor \underline{a} in linearer Weise ein Vektor \underline{b} zugeordnet wird. Für den Fall, dass es sich jeweils um Elemente des \mathbb{R}^3 handelt, schreibt man auch komponentenweise

$$\begin{aligned}b_1 &= t_{11}a_1 + t_{12}a_2 + t_{13}a_3, \\b_2 &= t_{21}a_1 + t_{22}a_2 + t_{23}a_3, \\b_3 &= t_{31}a_1 + t_{32}a_2 + t_{33}a_3,\end{aligned}$$

oder kürzer

$$b_i = t_{ij}a_j := \sum_{j=1}^3 t_{ij}a_j.$$

Dabei wurde die **Einstein'sche Summenkonvention** verwendet. Diese besagt, dass über doppelt vorkommende Indizes zu summieren ist. Die Summationsgrenze ergeben sich dabei aus dem Zusammenhang. Diese Konvention ist sehr praktisch, wenn man es, wie in der allgemeinen Relativitätstheorie, mit vielen Matrizen, Tensoren etc. zu tun hat. Man muß allerdings sehr aufpassen, z.B. kann man in einem Ausdruck wie $t_{ij}a_j / (s_{ij}a_j)$ nicht einfach a_j kürzen! In diesem Kapitel werden wir die Summenkonvention durchgehend verwenden!

In dem Beispiel haben wir die lineare Abbildung T durch eine Matrixmultiplikation dargestellt. Unser Tensor ist daher nichts anderes als eine Matrix¹.

¹Wegen der späteren Verallgemeinerung des Tensorbegriffes schreiben wir hier T statt \underline{T} .

In der Physik charakterisiert man allgemein Tensoren durch ihr Verhalten unter orthogonalen Transformationen, z.B. Drehungen. Wir wenden daher auf \underline{a} und \underline{b} eine orthogonale Transformation $\underline{R} = (r_{ij})$, d.h. $\underline{R}^{-1} = \underline{R}^t$ bzw.

$$\underline{R} \cdot \underline{R}^t = \mathbb{1}, \quad \text{bzw.} \quad r_{ij}r_{il} = \delta_{jl},$$

(wobei wir die Summenkonvention verwendet haben) also eine Drehung oder Drehspiegelung, an:

$$a'_i := r_{lj}a_j, \quad b'_k := r_{ki}b_i,$$

wobei die transformierten (“gedrehten”) Vektoren durch einen Strich gekennzeichnet sind, bzw. die Umkehrung

$$a_j = r_{lj}a'_l, \quad b_i = r_{ki}b'_k.$$

Diese Umkehrung folgt aus der Orthogonalität der Matrix \underline{R} .

Wir fordern nun, dass im gedrehten System der gleiche formale Zusammenhang zwischen den Vektoren wie im ursprünglichen System bestehen soll, d.h.

$$b'_k = t'_{kl}a'_l.$$

Es stellt sich dann die Frage, wie der transformierte Tensor $T' = (t'_{kl})$ mit dem ursprünglichen $T = (t_{kl})$ zusammenhängt. Dies zeigt folgende kurze Rechnung:

$$t'_{kl}a'_l = b'_k = r_{ki}b_i = r_{ki}(t_{ij}a_j) = r_{ki}t_{ij}(r_{lj}a'_l) = (r_{ki}r_{lj}t_{ij})a'_l.$$

Somit lesen wir ab:

$$t'_{kl} = r_{ki}r_{lj}t_{ij}.$$

Dies definiert das Transformationsverhalten eines **Tensors 2. Stufe**, d.h. alle Größen, die sich unter Drehungen genau so transformieren, sind Tensoren.

Man beachte, dass dieses Transformationsverhalten analog zu dem des Produktes $b_i a_j$ zweier Vektoren² ist:

$$b'_k a'_l = r_{ki}r_{lj}b_i a_j.$$

Definition 17.1.1 (Tensoren).

Ein **Tensor 2. Stufe** transformiert sich unter orthogonalen Transformationen $\underline{R} = (r_{ij})$ wie das Produkt zweier Vektorkomponenten:

$$t'_{kl} = r_{ki}r_{lj}t_{ij}.$$

Allgemein ist ein **Tensor n-ter Stufe** (vom Rang n oder der Ordnung n) eine Menge von Termen mit n Indizes und analogem Transformationsverhalten

²Genauer müsste man sagen, zum Transformationsverhalten des Produktes zweier Vektorkomponenten!

$$t'_{klmn\dots} = r_{ka}r_{lb}r_{mc}r_{nd}\cdots t_{abcd\dots},$$

d.h. mit einem Faktor r_{ij} für jeden Index. Speziell haben wir

$$\begin{aligned} n = 0 & : \text{ Skalare} && (1 \text{ Komponente im } \mathbb{R}^3), \\ n = 1 & : \text{ Vektoren} && (3 \text{ Komponenten im } \mathbb{R}^3), \\ n = 2 & : \text{ Matrizen} && (3^2 = 9 \text{ Komponenten im } \mathbb{R}^3), \end{aligned}$$

d.h. die Fälle $n = 1, 2, 3$ reduzieren sich auf bereits bekannte Objekte.

Etwas formaler sind Tensoren multilineare Abbildungen des Raumes $V_1 \otimes V_2 \otimes \cdots \otimes V_n$, dem **direkten Produkt** oder **Tensorprodukt** der Vektorräume V_1, \dots, V_n , in die reellen oder komplexen Zahlen. Der Tensor "erbt" dann das Transformationsverhalten der einzelnen Räume.

Man unterscheidet häufig zwischen **symmetrischen** und **antisymmetrischen Tensoren**

$$t_{ik} = \pm t_{ki},$$

wobei das obere (untere) Vorzeichen für den symmetrischen (antisymmetrischen) Fall steht. Bei Tensoren höherer Stufe muß man zusätzlich angeben, bzgl. welcher Indizes eine (Anti-)Symmetrie besteht. So bedeutet

$$t_{ijkl\dots} = \pm t_{jikl\dots}$$

(Anti-)Symmetrie in den ersten beiden Indizes.

Man kann Tensoren 2. Stufe immer in einen symmetrischen und einen antisymmetrischen Anteil aufspalten:

$$t_{ij} = \underbrace{\frac{1}{2}(t_{ij} + t_{ji})}_{\text{symmetrisch}} + \underbrace{\frac{1}{2}(t_{ij} - t_{ji})}_{\text{antisymmetrisch}}.$$

Für Tensoren höherer Stufe geht das für zwei beliebige Indizes.

Beispiel 17.1.2. Wir geben einige wichtige Beispiele für Tensoren.

1. Das Kronecker-Delta δ_{ij} ist ein symmetrischer Tensor 2. Stufe mit

$$\delta'_{ij} = r_{is}r_{jt}\delta_{st} = r_{is}r_{js} = \delta_{ij},$$

d.h. δ_{ij} ist invariant unter orthogonalen Transformationen.

2. Im \mathbb{R}^2 haben wir mit $\sigma_{12} = -\sigma_{21} = 1$ und $\sigma_{11} = \sigma_{22} = 0$ den sog. antisymmetrischen Tensor 2. Stufe

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

3. Aus den Übungen (vgl. Aufgabe 7) kennen wir bereits das **Levi-Cevita-Symbol**

$$\epsilon_{ijk} := \begin{cases} 1, & (ijk) \text{ gerade Permutation von } (123), \\ -1, & (ijk) \text{ ungerade Permutation von } (123), \\ 0, & \text{sonst (d.h. mind. zwei gleiche Indizes)}. \end{cases}$$

Es ist also $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ und $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$, alle 21 anderen Komponenten verschwinden.

Das Levi-Cevita-Symbol bezeichnet man auch als ϵ -**Tensor** oder den **vollständig antisymmetrischen Tensor 3. Stufe**.

4. Als physikalischen Beispiel wollen wir den **Trägheitstensor** $I = (I_{ij})$ nennen, der den Zusammenhang zwischen Drehimpuls \underline{L} und Winkelgeschwindigkeit $\underline{\omega}$ für einen starren Körper beschreibt:

$$L_i = I_{ij}\omega_j = \int_V dV \rho(x_1, x_2, x_3) (r^2\delta_{ij} - x_i x_j) \omega_j,$$

wobei $\rho(\underline{r})$ die Massendichte des Körpers ist und $r^2 = x_l x_l = x_1^2 + x_2^2 + x_3^2$. Explizit hat er die Form

$$I = \int_V dV \rho(x_1, x_2, x_3) \begin{pmatrix} x_2^2 + x_3^2 & -x_1 x_2 & -x_1 x_3 \\ -x_2 x_1 & x_1^2 + x_3^2 & -x_2 x_3 \\ -x_3 x_1 & -x_3 x_2 & x_1^2 + x_2^2 \end{pmatrix},$$

d.h. der Trägheitstensor ist ein symmetrischer Tensor 2. Stufe. Das Trägheitsmoment um eine Achse (durch den Ursprung) in Richtung \underline{n} ist dann durch $I_{\underline{n}} = n_i I_{ij} n_j$ gegeben.

17.2 Rechenregeln für Tensoren

Wir stellen hier kurz die wichtigsten Rechenregeln für Tensoren zusammen.

Die Addition (Subtraktion) zweier Vektoren (gleicher Stufe!) ist analog zur Addition von Matrizen komponentenweise definiert:

$$c_{ik} = a_{ik} \pm b_{ik}.$$

Entsprechendes gilt für Tensoren beliebiger Stufe.

Im Gegensatz zur Addition unterscheidet sich die Multiplikation von Tensoren von der von Vektoren oder Matrizen.

Das **direkte Produkt** oder **Tensorprodukt**

$$r_{jklm} = a_{jk} \otimes b_{lm} := a_{jk} b_{lm}$$

führt nämlich auf Tensoren höherer Stufe³. Es ist nicht nur für Tensoren gleicher Stufe definiert, die Definition läßt sich analog auf beliebige Produkte verallgemeinern. Allgemein gilt:

$$n\text{-Stufe} \otimes m\text{-Stufe} = (n + m)\text{-Stufe} .$$

Eine weitere neue Operation ist die sogenannte **Verjüngung** oder **Kontraktion**. Hierbei werden Indizes aussummiert, z.B. im Falle eines Tensors 4. Stufe, bei dem die ersten beiden Indizes verjüngt werden:

$$s_{lm} = r_{jjlm} \quad \left(= \sum_j r_{jjlm} \right) .$$

Die Verjüngung, die über zwei beliebige Indizes erfolgen kann, führt also auf einen Tensor der Stufe $n - 2$.

Die Matrixmultiplikation können wir nun als Tensorprodukt mit anschließender Verjüngung darstellen, z.B. für das Produkt einer Matrix und eines Vektors:

$$t_{ij} \otimes a_k = t_{ij}a_k =: s_{ijk} \quad \longrightarrow \quad b_i = t_{ij}a_j = s_{ijj} .$$

Dies bezeichnet man auch als **Überschiebung** zweier Tensoren.

Analog ist das Skalarprodukt die Überschiebung zweier Tensoren 1. Stufe:

$$a_i \otimes b_j = a_i b_j \quad \longrightarrow \quad a_i b_i = \underline{a} \cdot \underline{b} .$$

Im Fall eines Tensors 2. Stufe entspricht die Verjüngung der Spurbildung:

$$\text{Spur}(t_{jk}) = t_{jj} .$$

17.3 Differentialoperationen und Tensoren

Der Gradient eines Skalars (also eines Tensors 0. Stufe) ist ein Tensor 1. Stufe, denn:

$$\frac{\partial \phi}{\partial x'_i} = \frac{\partial \phi}{\partial x_j} \frac{\partial x_j}{\partial x'_i} = \frac{\partial \phi}{\partial x_j} \frac{\partial (r_{lj} x'_l)}{\partial x'_i} = r_{lj} \frac{\partial \phi}{\partial x_j} \underbrace{\frac{\partial x'_l}{\partial x'_i}}_{=\delta_{li}} = r_{ij} \frac{\partial \phi}{\partial x_j} .$$

Analog sieht man, dass

$$\left(\frac{\partial b_i}{\partial x_j} \right) = \begin{pmatrix} \frac{\partial b_1}{\partial x_1} & \frac{\partial b_2}{\partial x_1} & \frac{\partial b_3}{\partial x_1} \\ \frac{\partial b_1}{\partial x_2} & \frac{\partial b_2}{\partial x_2} & \frac{\partial b_3}{\partial x_2} \\ \frac{\partial b_1}{\partial x_3} & \frac{\partial b_2}{\partial x_3} & \frac{\partial b_3}{\partial x_3} \end{pmatrix}$$

ein Tensor 2. Stufe ist, d.h.

$$\frac{\partial b'_i}{\partial x'_j} = r_{jk} r_{il} \frac{\partial b_k}{\partial x_l} .$$

³Mit Ausnahme des Falls Tensoren 0. Stufe.

Analoges gilt für die Ableitungen von Tensoren höherer Stufe.

Abschließend wollen wir noch den Zusammenhang mit Divergenz und Rotation beleuchten. Mit der Abkürzung

$$\partial_i := \frac{\partial}{\partial x_i}$$

läßt sich die Divergenz als Verjüngung der Ableitung $\partial_i b_j$ (eines Tensors 2. Stufe!) interpretieren:

$$\partial_i b_j \longrightarrow \partial_i b_i = \operatorname{div} \underline{b}.$$

Die Rotation läßt sich mit dem Levi-Cevita-Symbol als zweifache Verjüngung eines Tensors 5. Stufe auffassen:

$$\epsilon_{ijk} \partial_m b_n \longrightarrow \epsilon_{ijk} \partial_j b_k = \operatorname{rot} \underline{b}.$$