

Der mathematische Werkzeugkasten

John Smith, 2014–2015
Lizenz: Creative Commons CC0

Inhaltsverzeichnis

1	Lineare Algebra	1
1.1	Komplexe Zahlen	1
1.2	Das Skalarprodukt	2
1.3	Das äußere Produkt	3
1.4	Das geometrische Produkt	3
1.5	Das Tensorprodukt	5
1.6	Allgemeine Basen	7
1.7	Basiswechsel	7
1.8	Eigenwerte	8
1.9	Multilinearformen	9
2	Analysis	10
2.1	Vektorwertige Funktionen	10
2.2	Funktionen mit mehreren Argumenten	10
2.3	Die Kettenregel	11
2.4	Kurvenintegrale	11
2.5	Kurvenintegrale zweiter Art	12
2.6	Variationsrechnung	13
2.7	Extrema mit Nebenbedingungen	15
2.8	Die Gâteaux-Ableitung	16
3	Differentialgeometrie	16
3.1	Integralsätze	16
3.2	Die Richtungsableitung	18
3.3	Koordinatensysteme	18
3.4	Der Gradient	19
3.5	Volumina	20
3.6	Vektorfelder	21
3.7	Die Volumenform	22
3.8	Der Pullback	24
3.9	Der Pushforward	25
4	Funktionalanalysis	26
4.1	Bra-Ket-Notation	26
4.2	Funktionen	27
4.3	Schwingungen	28
4.4	Basen	28
4.5	Integraltransformationen	30
4.6	Das Stieltjes-Integral	30
4.7	Fraktionale Iterationen	30
5	Stochastik	32
5.1	Ereignisse	32
5.2	Zufallsgrößen	33
5.3	Stochastische Prozesse	33
6	Funktionentheorie	34
6.1	Holomorphie	34
6.2	Integralsatz von Cauchy	34

1 Lineare Algebra

1.1 Komplexe Zahlen

Die Winkelfunktionen stehen zur Exponentialfunktion in Beziehung. Es ist

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix}),$$

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}).$$

Mit diesen Formeln lassen sich die Ableitungen der Winkelfunktionen gewinnen. Z.B. ist

$$\begin{aligned}\sin' x &= \frac{1}{2i}(e^{ix} - e^{-ix})' = \frac{1}{2i}(ie^{ix} + ie^{-ix}) \\ &= \frac{1}{2}(e^{ix} + e^{-ix}) = \cos x.\end{aligned}$$

Auch die Additionstheoreme lassen sich über die Exponentialfunktion herleiten. Die Sinusschwingung

$$u = \hat{u} \sin(\omega t + \varphi_0)$$

lässt sich alternativ schreiben als

$$\underline{u} = \hat{u} e^{i\omega t + i\varphi_0}.$$

Die ursprüngliche Schwingung ist dann gegeben durch den Imaginärteil

$$u = \text{Im}(\underline{u}).$$

Die Schwingung lässt sich damit als Drehzeiger interpretieren, der sich mit der Kreisfrequenz ω entgegen dem Uhrzeigersinn dreht. Die ursprüngliche Sinus-Funktion ist einfach die Projektion auf die y -Achse.

Dieser Formalismus ist in dem Sinn vorteilhaft, als dass sich hiermit zwei Schwingungen gleicher Kreisfrequenz addieren lassen, indem man die Drehzeiger wie Vektoren addiert. Es entsteht wieder eine Sinus-Schwingung mit der gleichen Kreisfrequenz.

Man betrachtet die Zeiger (welche komplexe Zahlen sind) einfach zum Zeitpunkt $t = 0$. Selbst diese Zeiger kodieren sämtliche Information über die Schwingung, nämlich die Anfangsphase und die Amplitude. Man muss also lediglich die Summe

$$\underline{u}_1(0) + \underline{u}_2(0) = \hat{u}_1 e^{i\varphi_1} + \hat{u}_2 e^{i\varphi_2}$$

berechnen. Dazu ist es notwendig die komplexen Zahlen von der Polarform in die algebraische Form umzurechnen. Dann kann man sie einfach addieren. Die Summe muss anschließend wieder in die Polarform umgerechnet werden. In dieser Form können Amplitude und Anfangsphase abgelesen werden und man ist fertig.

Was haben wir jetzt gelernt? Wir haben gelernt, dass sich Schwingungen als Vektoren in Polarkoordinaten interpretieren lassen. Will man diese addieren, so geht man einen Weg, den man in der Mathematik

häufiger finden kann. Man transformiert erst von Polarkoordinaten in euklidische Koordinaten, führt die Addition aus und transformiert anschließend zurück in Polarkoordinaten.

Um dies möglichst zügig auszuführen, sollte man die Transformationsformeln immer parat haben. Die Transformationsformeln von Polarkoordinaten in euklidische Koordinaten sind

$$\begin{aligned}x &= r \cos \varphi, \\y &= r \sin \varphi.\end{aligned}$$

Die Transformationsformeln von euklidischen Koordinaten in Polarkoordinaten sind

$$\begin{aligned}r &= \sqrt{x^2 + y^2}, \\ \varphi &= \operatorname{sgn}(y) \arccos(x/r).\end{aligned}$$

Die letzte Gleichung ist nur für $\varphi = \pi$ ungültig.

Zwei Dinge spielen beim Rechnen mit komplexen Zahlen eine wesentliche Rolle. Das ist zum einen die Umwandlung von kartesischen Koordinaten in Polarkoordinaten und umgekehrt. Zum anderen ist es die Verwendung der eulerschen Formel.

Man kann die Berechnung von Funktionswerten der komplexen Exponentialfunktion auf die Berechnung von Funktionswerten der reellen Exponentialfunktion zurückführen. Sei $a := \operatorname{Re}(z)$ und $b := \operatorname{Im}(z)$. Es ist

$$e^z = e^{(a+bi)} = e^a e^{bi} = e^a (\cos b + i \sin b).$$

Damit hat man also $\operatorname{Re}(e^z) = e^a \cos b$ und $\operatorname{Im}(e^z) = e^a \sin b$ gewonnen. Realteil und Imaginärteil können separat ausgerechnet werden, ohne komplexe Zahlen benutzen zu müssen. Die eulersche Formel hat uns diese Zurückführung ermöglicht.

Die Berechnung des komplexen Logarithmus lässt sich ebenfalls auf die Berechnung des reellen Logarithmus zurückführen. Dazu stellt man z in Polarform dar. Sei also $r = |z|$ und $\varphi = \arg(z)$. Es ist

$$\begin{aligned}w &= \ln(z) = \ln(re^{i\varphi}) = \ln(r) + \ln(e^{i\varphi}) \\ &= \ln(r) + i\varphi.\end{aligned}$$

Ersetzt man φ gegen $\varphi + 2\pi$, so erhält man ebenfalls z , jedoch ein anderes w . Das Ergebnis der Logarithmusfunktion ist also nicht eindeutig. Sie hat unendlich viele sogenannte Zweige. Man verlangt nun $-\pi < \varphi$ und $\varphi \leq \pi$. Dann definiert man den Hauptzweig als

$$\ln(z) := \ln(r) + i\varphi.$$

Die mengenwertige Logarithmusfunktion mit allen Zweigen ist dann

$$\operatorname{Ln}(z) := \ln(z) + 2\pi ki. \quad (k \in \mathbf{Z})$$

Man muss beachten, dass die Regeln $\ln(xy) = \ln(x) + \ln(y)$ und $\ln(x^y) = y \ln(x)$ im Allgemeinen nicht mehr gültig sind. Die hier dargestellte Rechnung für den Logarithmus funktioniert aber. Man kann sich durch die Probe überzeugen. Es ist

$$e^w = e^{\ln(r) + i\varphi + 2\pi ki} = e^{\ln(r)} e^{i\varphi} e^{2\pi ki} = r e^{i\varphi} = z.$$

1.2 Das Skalarprodukt

Die Menge der komplexen Zahlen mit der Addition zusammen ist isomorph zum Vektorraum $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$. Das heißt *alles*, was man mit Vektoren machen kann, ist auch mit komplexen Zahlen möglich. Zu jedem Punkt in der Menge $\mathbf{R} \times \mathbf{R}$ gehört genau ein Vektor, der vom Koordinatenursprung auf diesen Punkt zeigt. Zu jedem Punkt gehört außerdem genau eine komplexe Zahl.

Sei V ein zweidimensionaler reeller Vektorraum mit einem Skalarprodukt $V \times V \rightarrow \mathbf{R}$, welches für $a, b \in V$ mit $\langle a, b \rangle$ notiert wird. Wie dieses spezifiziert ist, ist egal, solange es die Skalarproduktaxiome erfüllt.

Für jeden Vektorraum lässt sich eine Basis finden. (Das schließt wegen *vacuous truth* auch den Nullvektorraum mit ein, was aber eine Nebensächlichkeit darstellt). Über das Gram-Schmidt-Verfahren lässt sich daher jedem Skalarproduktraum eine Orthonormalbasis geben.

Sei $B := (e_1, e_2)$ eine Orthonormalbasis des reellen Vektorraumes V . Jeder Vektor $a \in V$ kann als

$$a = a_1 e_1 + a_2 e_2 \tag{1}$$

dargestellt werden und es gilt $a_k = \langle e_k, a \rangle$.

Ist $a := a_1 e_1 + a_2 e_2$ und $b := b_1 e_1 + b_2 e_2$, so gilt

$$\langle a, b \rangle = a_1 b_1 + a_2 b_2. \tag{2}$$

Ein Vektor $a \in V$ lässt sich auch als komplexe Zahl

$$a = a_1 + a_2 i \tag{3}$$

interpretieren, wobei man $e_1 := 1$ und $e_2 := i$ setzt. Für $a, b \in \mathbf{C}$ muss hierbei

$$\langle a, b \rangle := \operatorname{Re}(\bar{a}b). \tag{4}$$

sein, um Formel (2) zu erhalten.

Weiterhin gilt die Formel

$$\langle a, b \rangle = |a||b| \cos \varphi \tag{5}$$

wobei φ der Winkel zwischen den beiden Vektoren ist. Die Betrag eines Vektors ist

$$|a| = \sqrt{\langle a, a \rangle} = \sqrt{a_1^2 + a_2^2}. \tag{6}$$

Der Betrag eines Vektors ist die Länge dieses Vektors. Die Formel wird durch den Satz des Pythagoras motiviert. Der Betrag einer komplexen Zahl stimmt mit dem Betrag des Vektors überein, wenn man die komplexe Zahl als Vektor interpretiert. Es ist der Abstand vom Koordinatenursprung zu dem Punkt, welcher durch die komplexe Zahl gegeben ist.

Zwei Vektoren a und b lassen sich auch als Punkte in einem Raum interpretieren. Der Abstand ist dann gegeben durch

$$d(a, b) = |a - b|. \tag{7}$$

Diese erfüllen die Metrik-Axiome. Somit ist $d(a, b)$ eine vernünftige Metrik und (V, d) wird zu einem metrischen Raum.

Den hier vorgestellten Formalismus kann man nun sofort auf den n -dimensionalen euklidischen Vektorraum \mathbf{R}^n verallgemeinern. Jeder Vektor lässt sich als Linearkombination

$$a = a_1 e_1 + \dots + a_n e_n$$

darstellen. Das Skalarprodukt ist

$$\langle a, b \rangle = a_1 b_1 + \dots + a_n b_n.$$

Die Formeln

$$\langle a, b \rangle = |a||b| \cos \varphi, \quad (8)$$

$$|a| = \sqrt{\langle a, a \rangle}, \quad (9)$$

$$d(a, b) = |a - b| \quad (10)$$

bleiben weiterhin gültig.

Für das Skalarprodukt gibt es wichtige Rechenregeln. Es ist kommutativ und bilinear, jedoch nicht assoziativ. Was bedeutet bilinear? Nun ja, es bedeutet, dass eine Funktion linear in jedem ihrer zwei Argumente ist. Allgemeiner ist multilinear auf diese Art definiert. Der Ableitungsoperator ist linear. Das bestimmte Integral ist linear. Das Summenzeichen ist linear. Proportionale Funktionen sind linear. Lineare Funktionen sind in diesem Sinn im Allgemeinen aber nicht linear.

Lineare Differentialgleichungen heißen linear, weil sie sich in der Form

$$Df(x) = g(x)$$

schreiben lassen. Dabei ist g eine gegebene Funktion und f die gesuchte Funktion. Der Operator D ist ein linearer Differentialoperator. Z.B. lässt sich

$$f''(x) + 2f'(x) + f(x) = 4$$

auch als

$$Df(x) = \left(\frac{d^2}{dx^2} + 2 \frac{d}{dx} + 1 \right) f(x) = 4$$

schreiben.

1.3 Das äußere Produkt

Ein weiteres wichtiges Produkt ist das äußere Produkt. Es lässt sich mit der Formel

$$\underline{a} \wedge \underline{b} = (a_1 b_2 - a_2 b_1) \underline{e}_1 \wedge \underline{e}_2$$

berechnen. Das äußere Produkt ist antikommutativ, es ist

$$\underline{a} \wedge \underline{b} = -\underline{b} \wedge \underline{a}.$$

Damit ergibt sich $\underline{a} \wedge \underline{a} = 0$. Allgemeiner ist $\underline{a} \wedge \underline{b} = 0$, wenn \underline{a} und \underline{b} kollinear sind. Zwei Vektoren sind kollinear, wenn sie in die selbe Richtung oder die entgegengesetzte Richtung zeigen.

Das äußere Produkt ist außerdem bilinear und assoziativ. Diese Rechenregeln genügen, um mit dem äußeren Produkt zu rechnen. Ein Beispiel.

$$\begin{aligned} \underline{a} \wedge \underline{b} &= (a_1 e_1 + a_2 e_2) \wedge (b_1 e_1 + b_2 e_2) \\ &= a_1 e_1 \wedge (b_1 e_1 + b_2 e_2) + a_2 e_2 \wedge (b_1 e_1 + b_2 e_2) \\ &= a_1 e_1 \wedge b_1 e_1 + a_1 e_1 \wedge b_2 e_2 \\ &\quad + a_2 e_2 \wedge b_1 e_1 + a_2 e_2 \wedge b_2 e_2 \\ &= 0 + a_1 e_1 \wedge b_2 e_2 + a_2 e_2 \wedge b_1 e_1 + 0 \\ &= a_1 b_2 e_1 \wedge e_2 + a_2 b_1 e_2 \wedge e_1 \\ &= a_1 b_2 e_1 \wedge e_2 - a_2 b_1 e_1 \wedge e_2 \\ &= (a_1 b_2 - a_2 b_1) e_1 \wedge e_2 \end{aligned}$$

Für das äußere Produkt gilt die Formel

$$\underline{a} \wedge \underline{b} = |a||b|(\sin \varphi) \underline{e}_1 \wedge \underline{e}_2.$$

Das äußere Produkt ist der Flächeninhalt des Parallelogramms, welches die beiden Vektoren aufspannen.

1.4 Das geometrische Produkt

Das geometrische Produkt ist definiert als

$$\underline{a} \underline{b} = \langle \underline{a}, \underline{b} \rangle + \underline{a} \wedge \underline{b}.$$

Das geometrische Produkt ist bilinear und assoziativ. Es ist jedoch weder kommutativ noch antikommutativ. Für rechtwinklige Vektoren reduziert sich das geometrische Produkt auf das äußere Produkt. Für kollineare Vektoren reduziert sich das geometrische Produkt auf das Skalarprodukt. Wenn man eine Orthonormalbasis $B = (e_1, e_2)$ wählt, gilt daher

$$\begin{aligned} e_1 e_2 &= -e_2 e_1, \\ e_1 e_1 &= 1, \\ e_2 e_2 &= 1. \end{aligned}$$

Allgemeiner rechnet man bei Orthonormalbasen nach dem folgenden Prinzip. Führt man z.B. in

$$e_1 e_3 e_4 e_1 e_2 e_2$$

eine Transposition von zwei *ungleichen* Basisvektoren aus, so ändert sich das Vorzeichen. Außerdem darf man wegen des Assoziativgesetzes die Klammern beliebig setzen. Zwei nebeneinander liegende gleiche Basisvektoren heben sich zu eins auf. Z.B. ist

$$\begin{aligned} e_1 e_3 e_4 e_1 e_2 e_2 &= e_1 e_3 e_4 e_1 \\ &= -e_1 e_1 e_4 e_3 = -e_4 e_3 = e_3 e_4. \end{aligned}$$

Diese Vereinfachungen muss man ständig machen, wenn man mit dem geometrischen Produkt rechnet. Multivektoren der Form

$$a + b e_1 e_2$$

sind isomorph zu den komplexen Zahlen. Es ist $e_1 e_2 = e_1 \wedge e_2 = i$. Z. B. ist

$$\begin{aligned} i^2 &= (e_1 e_2)^2 = (e_1 e_2)(e_1 e_2) = e_1 e_2 e_1 e_2 \\ &= -e_1 e_1 e_2 e_2 = -1. \end{aligned}$$

Mit der eulerschen Formel kann man schreiben

$$\begin{aligned} \underline{a} \underline{b} &= \langle \underline{a}, \underline{b} \rangle + \underline{a} \wedge \underline{b} = |\underline{a}| |\underline{b}| \cos \varphi + e_1 e_2 |\underline{a}| |\underline{b}| \sin \varphi \\ &= |\underline{a}| |\underline{b}| \cos \varphi + i |\underline{a}| |\underline{b}| \sin \varphi = |\underline{a}| |\underline{b}| e^{i\varphi}. \end{aligned}$$

Mit dem äußeren Produkt können wir einen Formalismus aufbauen, der alle assoziativen hyperkomplexen Zahlensysteme zu einem einzigen System zusammenfasst.

Was das soll? Nun ja, man kann jetzt austesten, was passiert, wenn man einen gewöhnlichen Vektor mit der imaginären Einheit multipliziert. Es ist

$$\begin{aligned} i(a_1 e_1 + a_2 e_2) &= e_1 e_2 (a_1 e_1 + a_2 e_2) \\ &= a_1 e_1 e_2 e_1 + a_2 e_1 e_2 e_2 = -a_1 e_1 e_1 e_2 + a_2 e_1 \\ &= -a_1 e_2 + a_2 e_1 = a_2 e_1 - a_1 e_2. \end{aligned}$$

bzw.

$$i \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}.$$

Der Vektor wird um 90 Grad im Uhrzeigersinn gedreht. Das geometrische Produkt ist nicht kommutativ. Wenn man von der anderen Seite mit i multipliziert, so dreht man gegen den Uhrzeigersinn. Bei komplexen Zahlen ist es egal, weil die Multiplikation ja kommutativ ist. Es ist

$$\begin{aligned} i(a_1 + a_2 i) &= e_1 e_2 (a_1 + a_2 e_1 e_2) \\ &= a_1 e_1 e_2 - a_2 = a_1 i - a_2 = -a_2 + a_1 i \end{aligned}$$

Die komplexe Zahl wird um 90 Grad gegen den Uhrzeigersinn gedreht. Multipliziert man von der anderen Seite mit i , so erhält man das selbe Ergebnis. Die komplexen Zahlen sind ein Spezialfall, für den das geometrische Produkt kommutativ ist, obwohl es im Allgemeinen nicht kommutativ ist.

Der Isomorphismus, der einen Vektor in eine komplexe Zahl überführt ist das geometrische Produkt mit e_1 . Denn es ist

$$\begin{aligned} e_1 \underline{a} &= e_1 (a_1 e_1 + a_2 e_2) = a_1 e_1 e_1 + a_2 e_1 e_2 \\ &= a_1 + a_2 e_1 e_2 = a_1 + a_2 i. \end{aligned}$$

Was macht die komplexen Zahlen eigentlich aus? Bezüglich der Addition sind es nur ganz gewöhnliche Vektoren. Völlig uninteressant. Das besondere ist die seltsame Multiplikation. Mit der Multiplikation von zwei können wir die Länge einer Zahl (interpretiert als Vektor vom Ursprung auf einen Punkt im Zahlenstrahl) verdoppeln. Durch die Multiplikation mit einer komplexen Zahl können wir eine Zahl nicht nur verlängern, wir können sie auch Drehen.

Wenn man einen Vektorraum mit einer bilinearen Multiplikation ausstattet, entsteht etwas, das man auch als Algebra bezeichnet. Der euklidische Vektorraum zusammen mit dem geometrischen Produkt wird dadurch zu einer assoziativen Algebra, die Clifford-Algebra genannt wird. Mit dem geometrischen Produkt können wir Vektoren drehen.

Das Vektorprodukt lässt sich mit

$$\underline{a} \times \underline{b} = (\underline{a} \wedge \underline{b}) I = (\underline{a} \wedge \underline{b}) e_1 e_2 e_3$$

berechnen. Diese Formel gilt nur, wenn man keine Rauminversion zulässt. Das Ergebnis des Vektorproduktes ist eigentlich ein sogenannter Pseudovektor. Dieser entsteht, weil man die Bilinearität des Vektorproduktes fordert.

Ein Vektor ändert bei Rauminversion im \mathbf{R}^3 sein Vorzeichen, ein Pseudovektor tut das nicht. Rauminversion bedeutet, dass jeder Basisvektor e_k gegen $-e_k$ ausgetauscht wird. Das rechtshändige Koordinatensystem wird dadurch zu einem linkshändigen. Bei einem Vektor ist nun

$$a_1(-e_1) + a_2(-e_2) + a_3(-e_3) = -\underline{a}.$$

Für I bekommt man

$$(-e_1)(-e_2)(-e_3) = -e_1 e_2 e_3 = -I.$$

Das Vektorprodukt ist bilinear. Man erhält

$$(-\underline{a}) \times (-\underline{b}) = (-1)(-1)\underline{a} \times \underline{b} = \underline{a} \times \underline{b}.$$

Eigentlich müsste man aber

$$(-\underline{a}) \wedge (-\underline{b})(-I) = -(\underline{a} \wedge \underline{b}) I$$

erhalten. Das äußere Produkt ist auch bilinear. Allerdings besteht hier kein Problem, weil ja die Basisbivektoren $e_i \wedge e_j = (-e_i) \wedge (-e_j)$ im Ergebnis des Produktes enthalten sind.

Abschließend sollen noch einige nützliche Definitionen formuliert werden. Der Projektionsoperator $[A]_k$ projiziert den Anteil vom Grad k heraus. Vom Grad null sind Skalare, vom Grad eins sind Vektoren und Kovektoren, vom Grad zwei sind Bivektoren usw. Bei negativem k kann man $[A]_k = 0$ setzen. Es ist z.B.

$$\begin{aligned} [a + be_1 + ce_1 e_2]_0 &= a, \\ [a + be_1 + ce_1 e_2]_1 &= be_1, \\ [a + be_1 + ce_1 e_2]_2 &= ce_1 e_2. \end{aligned}$$

Man definiert nun

$$\begin{aligned} A \wedge B &:= \sum_{i,j} [[A]_i [B]_j]_{i+j}, \\ A \lrcorner B &:= \sum_{i,j} [[A]_i [B]_j]_{j-i}, \\ \langle A, B \rangle &:= \sum_{i,j} [[A]_i [B]_j]_0, \\ A \lrcorner B &:= \sum_{i,j} [[A]_i [B]_j]_{i-j}. \end{aligned}$$

Man projiziert Anteile heraus, bildet von diesen das geometrische Produkt und projiziert dann wieder der Formel entsprechend den Anteil heraus. Darüber wird aufsummiert.

Bei $A \wedge B$ handelt es sich um die Verallgemeinerung des äußeren Produktes. Das Produkt $A \lrcorner B$ wird Linkskontraktion genannt und ist eine Verallgemeinerung des Skalarproduktes. Das Produkt $A \llcorner B$ wird Rechtskontraktion genannt und ist auch eine Verallgemeinerung des Skalarproduktes. Das Produkt $\langle A, B \rangle$ ist eine Verallgemeinerung des Skalarproduktes, die auch tatsächlich immer einen Skalar liefert.

Es gibt auch die Schreibweise $i_X \omega := X \lrcorner \omega$. Mit der Rechtskontraktion würde das keinen Sinn ergeben, weil X ein Vektor und ω z.B. vom Grad zwei sein soll. Mit der Rechtskontraktion würde man dann einen negativen Index zur Projektion erhalten.

In der Physik sind die Pauli-Matrizen von Bedeutung. Es ist

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

Für die Multiplikation von Pauli-Matrizen muss man sich eine Multiplikationstabelle merken. Wenn man das geometrische Produkt zur Multiplikation verwendet, dann entsprechen die Pauli-Matrizen σ_k jedoch einfach den Basisvektoren e_k . Die Einheitsmatrix entspricht der Zahl eins und i entspricht $e_1 e_2 e_3$. Der Vorteil ist, dass man sich keine Multiplikationstabelle merken muss.

Verwendet man die Basis des Minkowskiraums mit dem metrischen Tensor $g = \text{diag}(1, -1, -1, -1)$, so lassen sich die Dirac-Matrizen γ^k ebenfalls gegen die Basisvektoren e^k ersetzen. Die Basisvektoren bilden jedoch keine Orthonormalbasis, man rechnet stattdessen $(e_k)^2 = g^{kk}$ mit $g^{kk} = g_{kk}$. Das Produkt von Dirac-Matrizen entspricht dem geometrischen Produkt von Basisvektoren. Da die Basisvektoren eine Orthogonalbasis bilden, ist das geometrische Produkt der Basisvektoren antikommutativ.

1.5 Das Tensorprodukt

Das Tensorprodukt ist bilinear und assoziativ. Jedoch ist es weder kommutativ noch antikommutativ. Mit diesen Rechenregeln ist es möglich das Tensorprodukt von zwei Vektoren auszurechnen. Es ergibt sich

$$\begin{aligned} \underline{a} \otimes \underline{b} &= (a_1 e_1 + a_2 e_2) \otimes (b_1 e_1 + b_2 e_2) \\ &= a_1 e_1 \otimes (b_1 e_1 + b_2 e_2) + a_2 e_2 \otimes (b_1 e_1 + b_2 e_2) \\ &= a_1 e_1 \otimes b_1 e_1 + a_1 e_1 \otimes b_2 e_2 \\ &\quad + a_2 e_2 \otimes b_1 e_1 + a_2 e_2 \otimes b_2 e_2 \\ &= a_1 b_1 e_1 \otimes e_1 + a_1 b_2 e_1 \otimes e_2 \\ &\quad + a_2 b_1 e_2 \otimes e_1 + a_2 b_2 e_2 \otimes e_2. \end{aligned}$$

Das Tensorprodukt macht aus zwei Vektoren eine Matrix, die als Tensor bezeichnet wird. Es ist

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix}.$$

Dabei ist folgendes wichtig. Der Tensor selbst ist wie ein Vektor unabhängig von der gewählten Basis. Wechselt man die Basis gegen eine andere aus, so ändern sich die Einträge der Matrix. Der Tensor selbst bleibt aber der gleiche.

Es verhält sich wie bei physikalischen Größen. In einer anderen Einheit ergibt sich ein anderer Zahlenwert. Die Größe, welche aus Zahlenwert und Einheit besteht, bleibt aber gleich.

Sei $B = (e_1, e_2)$ eine Basis und sei $B' = (e'_1, e'_2)$ eine zweite Basis. Ein und der selbe Vektor kann in zwei unterschiedlichen Basen ausgedrückt werden. Seien a_k die Koordinaten bezüglich der Basis B und a'_k die Koordinaten bezüglich der Basis B' . Es ist

$$\underline{a} = a_1 e_1 + a_2 e_2 = a'_1 e'_1 + a'_2 e'_2.$$

Dabei stellt sich natürlich die Frage, wie man von den alten Koordinaten auf die neuen kommt. Dieses Problem wird unter dem Stichwort Basiswechsel gelöst. Bei Tensoren kann man auch einen Basiswechsel machen.

Nun ist es auch möglich $\underline{a} \otimes \underline{b} \otimes \underline{c}$ zu bilden. Wie soll man sich das Ergebnis vorstellen? Nun ja, es ist eine dreidimensionale Tabelle. Man kann sich ein Regal mit Zeilen und Spalten vorstellen, bei dem man in jedem Fach noch schrittweise in die Tiefe gehen kann.

Tensoren bilden mit der Addition wieder einen Vektorraum. Alles was man mit Vektoren tun kann ist also auch mit Tensoren möglich.

Anstelle von

$$T = \sum_{i,j} T_{ij} e_i \otimes e_j$$

schreibt man einen Tensor auch kürzer als T_{ij} . Dabei hat man sich natürlich stillschweigend auf die Basis $B = (e_1, e_2)$ festgelegt. Wir wollen diese Kurzschreibweise im Folgenden auch verwenden. Das Tensorprodukt wird nun kürzer

$$T_{ij} = a_i b_j$$

geschrieben. Wir können ein Tensorprodukt aus beliebig vielen Vektoren bilden. Z.B.

$$T_{ijkl} = a_i b_j c_k d_l.$$

Auch aus Tensoren lassen sich Tensorprodukte bilden. Z.B.

$$T_{ijkl} = A_{ij} B_{kl}.$$

Man kann einen Tensor nun kontrahieren. Bei der Kontraktion setzt man zwei Indizes gleich und summiert anschließend über diesen Index auf. Die Kontraktion von T_{ij} ist z.B.

$$\sum_k T_{kk} = T_{11} + T_{22}.$$

Das Skalarprodukt ist ein Spezialfall von Tensorprodukt gefolgt von Kontraktion. Das Skalarprodukt von a_i und b_j ist

$$\sum_k a_k b_k.$$

Wenn wir keine Orthonormalbasis verwenden, dann lässt sich das Skalarprodukt so nicht bilden. Man muss dann zuerst einen Basiswechsel in eine Orthonormalbasis machen oder man verwendet den Formalismus mit dem metrischen Tensor.

Der Formalismus mit dem metrischen Tensor lässt sich allgemein bei Kontraktionen verwenden. Wenn man keine Orthonormalbasis hat, muss man bei der Kontraktion bloß darauf achten, dass ein Index oben und der andere unten steht. Ein beliebiger Index lässt sich mit dem metrischen Tensor senken und mit dem inversen metrischen Tensor heben.

Die Multiplikation von Matrizen ist ein Spezialfall von Tensorprodukt gefolgt von Kontraktion. Seien a_{ik} und b_{lj} zwei Matrizen. Das Produkt ist

$$c_{ij} = \sum_k a_{ik} b_{kj}.$$

Ein antisymmetrischer Tensor ist ein Tensor mit der Eigenschaft $T_{ij} = -T_{ji}$. Die Hauptdiagonale von einem antisymmetrischen Tensor besteht zwingend aus Nullen. Von zwei Vektoren $\underline{a}, \underline{b}$ kann man einen antisymmetrischen Tensor auf folgende Art bilden. Man rechnet

$$\begin{aligned} \underline{a} \otimes \underline{b} - \underline{b} \otimes \underline{a} &= \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix} - \begin{bmatrix} a_1 b_1 & a_2 b_1 \\ a_1 b_2 & a_2 b_2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & (a_1 b_2 - a_2 b_1) \\ (a_2 b_1 - a_1 b_2) & 0 \end{bmatrix} \\ &= (a_1 b_2 - a_2 b_1) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ &= (a_1 b_2 - a_2 b_1)(e_1 \otimes e_2 - e_2 \otimes e_1). \end{aligned}$$

Mit der Definition

$$e_i \wedge e_j := e_i \otimes e_j - e_j \otimes e_i$$

erhält man

$$\underline{a} \wedge \underline{b} = \underline{a} \otimes \underline{b} - \underline{b} \otimes \underline{a}.$$

Die Formel ist auch für Vektoren mit mehr als zwei Komponenten gültig. Das bestätigt man durch die Rechnung

$$\begin{aligned} \underline{a} \otimes \underline{b} - \underline{b} \otimes \underline{a} &= \sum_{i,j} a_i b_j e_i \otimes e_j - \sum_{i,j} b_i a_j e_i \otimes e_j \\ &= \sum_{i,j} (a_i b_j - a_j b_i) e_i \otimes e_j \\ &= \sum_{i < j} (a_i b_j - a_j b_i) (e_i \otimes e_j - e_j \otimes e_i) \\ &= \sum_{i < j} (a_i b_j - a_j b_i) e_i \wedge e_j = \underline{a} \wedge \underline{b}. \end{aligned}$$

Ab jetzt sollen Vektoren nicht mehr unterstrichen werden. Außerdem sollen die Vektoren jetzt mit einem Index nummeriert werden. Das ist leider etwas verwirrend, da a_k bisher Komponenten waren und ab jetzt Vektoren sind. Man definiert

$$\text{Alt}_n(a_1 \otimes \dots \otimes a_n) := \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)} \otimes \dots \otimes a_{\sigma(n)}.$$

Für Vektoren a_k gilt allgemein die Formel

$$a_1 \wedge \dots \wedge a_n = n! \text{Alt}_n(a_1 \otimes \dots \otimes a_n).$$

Wenn man das Epsilon-Symbol verwendet, dann kann man auch alternativ schreiben

$$\text{Alt}_n(a_1 \otimes \dots \otimes a_n) = \frac{1}{n!} \sum_{i_1, \dots, i_n} \varepsilon_{i_1 \dots i_n} a_{i_1} \otimes \dots \otimes a_{i_n}.$$

An dieser Stelle soll noch festhalten werden, dass für einen antisymmetrischen Tensor gilt

$$A = \sum_{i,j} a_{ij} e_i \otimes e_j = \frac{1}{2} \sum_{i,j} a_{ij} e_i \wedge e_j = \sum_{i < j} a_{ij} e_i \wedge e_j.$$

Allgemeiner ist

$$\begin{aligned} A &= \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} e_{i_1} \otimes \dots \otimes e_{i_n} \\ &= \frac{1}{n!} \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} e_{i_1} \wedge \dots \wedge e_{i_n} \\ &= \sum_{i_1 < \dots < i_n} a_{i_1, \dots, i_n} e_{i_1} \wedge \dots \wedge e_{i_n}. \end{aligned}$$

Für einen antisymmetrischen Tensor mit den Komponenten $a[i_1, \dots, i_n] := a_{i_1 \dots i_n}$ gilt

$$a[i_1, \dots, i_n] = \text{sgn}(\sigma) a[\sigma(i_1), \dots, \sigma(i_n)].$$

Ein Tensor lässt sich nun auf einen Vektor anwenden, indem immer über den letzten Index kontrahiert wird. Für $A := \sum_{ij} a_{ij} e_i \otimes e_j$ und $v := \sum_k v_k e_k$ hat man nun

$$A(v) = \langle A, v \rangle = \sum_i a_{ik} v_k e_i.$$

Somit gilt $\langle e_i \otimes e_j, e_k \rangle = e_i \delta_{jk}$. Weiterhin gilt

$$\begin{aligned} T(v, w) &= \langle \langle T, w \rangle, v \rangle, \\ T(u, v, w) &= \langle \langle \langle T, w \rangle, v \rangle, u \rangle \end{aligned}$$

usw. Allgemein gilt dann die Formel

$$\langle T \otimes e_i, e_j \rangle = T \delta_{ij}.$$

Sinnvollerweise definiert man auch

$$\langle e_i, e_j \otimes T \rangle := T \delta_{ij}.$$

Für einen Tensor A zweiter Stufe ergibt sich dann

$$\langle A, v \rangle = \langle v, A^T \rangle.$$

Einen Tensor zweiter Stufe transponiert man wie eine Matrix, d.h. man definiert

$$\left(\sum_{ij} a_{ij} e_i \otimes e_j\right)^T := \sum_{ij} a_{ji} e_i \otimes e_j.$$

Allgemein verwendet man eine Permutation σ , welche die Indizes permutiert. Man notiert am besten T^σ . Speziell ist dann $A^T = A^{(21)}$.

1.6 Allgemeine Basen

Bei Orthonormalbasen stehen die Basisvektoren jeweils senkrecht aufeinander und haben den Betrag eins. Die Basisvektoren einer Orthonormalbasis $B = (e_1, e_2)$ wollen wir mit e_k bezeichnen.

Wir lassen diese Forderung nun fallen und bezeichnen mit f_k die Basisvektoren einer allgemeinen Basis $F = (f_1, f_2)$. Es ist möglich, den Vektor f_k als Linearkombination aus den e_k auszudrücken. Es ist

$$\begin{aligned} f_1 &= f_{11}e_1 + f_{12}e_2, \\ f_2 &= f_{21}e_1 + f_{22}e_2. \end{aligned}$$

Das können wir kompakt zusammenfassen zu

$$f_k = f_{k1}e_1 + f_{k2}e_2$$

bzw. noch kompakter zu

$$f_k = \sum_i f_{ki} e_i.$$

Wir wollen jetzt ein Skalarprodukt berechnen. Es ergibt sich

$$\begin{aligned} \langle \underline{a}, \underline{b} \rangle &= \langle a_1 f_1 + a_2 f_2, b_1 f_1 + b_2 f_2 \rangle \\ &= \langle a_1 f_1, b_1 f_1 \rangle + \langle a_1 f_1, b_2 f_2 \rangle \\ &\quad + \langle a_2 f_2, b_1 f_1 \rangle + \langle a_2 f_2, b_2 f_2 \rangle \\ &= a_1 b_1 \langle f_1, f_1 \rangle + a_1 b_2 \langle f_1, f_2 \rangle \\ &\quad + a_2 b_1 \langle f_2, f_1 \rangle + a_2 b_2 \langle f_2, f_2 \rangle \\ &= a_1 b_1 g_{11} + a_1 b_2 g_{12} + a_2 b_1 g_{21} + a_2 b_2 g_{22} \end{aligned}$$

Die Skalarprodukte $g_{ij} = \langle f_i, f_j \rangle$ kann man zum metrischen Tensor zusammenfassen. Der metrische Tensor ist also die Matrix

$$g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \langle f_1, f_1 \rangle & \langle f_1, f_2 \rangle \\ \langle f_2, f_1 \rangle & \langle f_2, f_2 \rangle \end{bmatrix}.$$

Mit $\underline{a}' = a_1 e_1 + a_2 e_2$ wollen wir den Koordinatenvektor zum Vektor $\underline{a} = a_1 f_1 + a_2 f_2$ bezeichnen. Wir können das Skalarprodukt nicht mehr mit

$$\langle \underline{a}', \underline{b}' \rangle = \sum_i a_i b_i$$

berechnen, sondern müssen es mit

$$\langle \underline{a}, \underline{b} \rangle = \langle g \underline{a}', \underline{b}' \rangle = \sum_{i,j} g_{ij} a_i b_j$$

berechnen. Ab jetzt wollen wir die Indizes aller Koordinaten nach oben stellen und schreiben

$$\begin{aligned} \underline{a} &= a^1 f_1 + a^2 f_2, \\ \underline{a}' &= a^1 e_1 + a^2 e_2. \end{aligned}$$

Zu den Koordinaten a^k des Vektors \underline{a} bezüglich der Basis F gibt es duale Koordinaten a_k . Der Duale Koordinatenvektor von \underline{a}' ist $g \underline{a}'$. Das bedeutet der Koordinatenvektor wird mit dem metrischen Tensor multipliziert. Es ist also

$$a_k = \sum_{i,j} g_{ki} a^i.$$

Außerdem gibt es eine zu F duale Basis $F^* = (f^1, f^2)$. Es ist

$$\begin{aligned} f_k &= \sum_{i,j} g_{ki} f^i, \\ f^k &= \sum_{i,j} g^{ki} f_i. \end{aligned}$$

Dabei ist $g^{-1} = (g^{ij})$ die inverse Matrix von $g = (g_{ij})$. Hierfür gibt es die Formel

$$g^{-1} = \begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix} = \frac{1}{g_{11}g_{22} - g_{12}g_{21}} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix}.$$

Mit diesem Formalismus lässt sich das Skalarprodukt auch durch

$$\langle \underline{a}, \underline{b} \rangle = a^1 b_1 + a^2 b_2$$

berechnen. Allgemeiner ist

$$\langle \underline{a}, \underline{b} \rangle = \sum_i a^i b_i = \sum_{i,j} g_{ij} a^i b^j.$$

Ein Vektor lässt sich sowohl als Linearkombinationen mit der Basis als auch als Linearkombinationen mit der Dualbasis ausdrücken. Es ist

$$\underline{a} = a^1 f_1 + a^2 f_2 = a_1 f^1 + a_2 f^2.$$

1.7 Basiswechsel

Sei $B = (b_1, b_2)$ eine Basis und $B' = (b'_1, b'_2)$ eine zweite Basis. Seien x_k die Koordinaten des Vektors \underline{v} bezüglich der Basis B und x'_k die Koordinaten bezüglich der Basis B' . Mit \underline{x} wollen wir den Koordinatenvektor der Koordinaten x_k bezeichnen. Das sind die Koordinaten x_k als Tupel geschrieben, jedoch nicht der eigentliche Vektor \underline{v} . Es stellt sich die Frage, wie man von den Koordinaten x_k zu den Koordinaten x'_k gelangt.

Die Basisvektoren können als Linearkombination aus den gestrichenen Basisvektoren ausgedrückt werden. Es ist

$$b_k = \sum_i a_{ik} b'_i.$$

Man kann das auch als Matrizenmultiplikation schreiben. Sei $A = (a_{ij})$. Da man bei der Matrizenmultiplikation über den zweiten Index summiert, müssen wir die Matrix transponieren. Kompakter schreibt man nun

$$B^T = A^T B'^T.$$

Mit B^T ist der Spaltenvektor zu B gemeint, da es sich bei B um einen Zeilenvektor handelt. Daraus erhält man mit der allgemeinen Regel $(AB)^T = B^T A^T$ auch

$$B = B' A.$$

Den Vektor können wir sowohl in der alten als auch in der neuen Basis ausdrücken. Es ist

$$\underline{v} = \sum_k x_k b_k = \sum_k x'_k b'_k.$$

Man ersetzt die b_k nun und erhält

$$\begin{aligned} \underline{v} &= \sum_k x_k \sum_i a_{ik} b'_i = \sum_i \sum_k a_{ik} x_k b'_i \\ &= \sum_k \sum_i a_{ki} x_i b'_k. \end{aligned}$$

Durch einen Koeffizientenvergleich stellt man fest

$$x'_k = \sum_i a_{ki} x_i.$$

Das kann man kompakt als

$$\underline{x}' = A \underline{x}.$$

formulieren. Die gesamte Herleitung lässt sich genauso kompakt schreiben. Unter Verwendung von $B = B' A$ ergibt sich

$$\underline{v} = B' \underline{x}' = B \underline{x} = B' A \underline{x}.$$

Durch einen Koeffizientenvergleich oder durch Multiplikation mit B'^{-1} erhält man wieder

$$\underline{x}' = A \underline{x}.$$

Wenn wir also A kennen, so ist unser Problem schon gelöst. Kennen wir A^{-1} , so müssen wir daraus noch A bestimmen.

Es kann auch der Fall bestehen, dass man die Koordinaten der Basisvektoren kennt. Man kann dann B auch zu einer Matrix expandieren. Es ist

$$B = (b_1, b_2) = \left(\begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}, \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} \right) = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Die Gleichung $B' A = B$ lässt sich nun als lineares Gleichungssystem interpretieren. Dabei sind B' und B gegeben und A ist gesucht. Man hat also

$$\left[\begin{array}{cc|cc} b'_{11} & b'_{12} & b_{11} & b_{12} \\ b'_{21} & b'_{22} & b_{21} & b_{22} \end{array} \right].$$

Mit dem Gauß-Jordan-Verfahren erhält man

$$\left[\begin{array}{cc|cc} 1 & 0 & a_{11} & a_{12} \\ 0 & 1 & a_{21} & a_{22} \end{array} \right]$$

und kann die Einträge von A direkt ablesen. Eigentlich handelt es sich um zwei Gleichungssysteme. Das erste davon ist

$$\begin{bmatrix} b'_{11} & b'_{12} \\ b'_{21} & b'_{22} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}.$$

Alternativ kann man B'^{-1} bestimmen und dann

$$A = B'^{-1} B$$

berechnen.

Ein besonders einfacher Fall ist der, dass der Basisvektor b'_k zum Basisvektor b_k um einen Winkel verdreht ist. Für jedes k soll dabei der gleiche Drehwinkel φ bestehen. Die Matrix A ist dann orthogonal. Man hat in diesem Fall $A^{-1} = A^T$.

Die Transformationsmatrix A von den alten in die neuen Koordinaten nennt man auch $T_{B'}^B$. Man schreibt dann $\underline{x}' = T_{B'}^B \underline{x}$. Dass B' im Index unten steht und B im Index oben, ist sinnvoll. Es ist ja $T_{B'}^B = B'^{-1} B$. Die Matrix B' steht sozusagen im Nenner.

1.8 Eigenwerte

Sei A eine quadratische Matrix. Normalerweise dreht und skaliert die Matrix A einen Vektor \underline{v} . Man kann nun die Frage nach Vektoren \underline{v} stellen, für die sich die Anwendung der Matrix wie eine Skalierung verhält, ohne den Vektor zu drehen. Gesucht sind also Vektoren \underline{v} , für die

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

ist. Dieses Problem wird Eigenwertproblem genannt. Der Vektor \underline{v} wird Eigenvektor der Matrix A genannt und λ ist der dazu gehörige Eigenwert.

Beide Seiten der Gleichung lassen sich mit einer Zahl multiplizieren. Man hat dann $A(r\underline{v}) = \lambda(r\underline{v})$. Nun ist $\underline{v}' = r\underline{v}$ wieder ein Vektor. Hat man einen Eigenvektor gefunden, so lassen sich also beliebig viele produzieren, die alle kollinear sind.

Die Umformung der Gleichung ergibt

$$(A - \lambda E) \underline{v} = 0.$$

Das lässt sich als homogenes LGS interpretieren. Ein homogenes LGS $A \underline{v} = 0$ hat nur dann nichttriviale Lösungen, wenn $\det A = 0$ ist. Es muss also

$$\det(A - \lambda E) = 0$$

sein. Sei A nun eine 2×2 -Matrix. Es ist dann

$$\det \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = 0.$$

Damit erhält man das charakteristische Polynom

$$\lambda^2 - \text{Spur}(A)\lambda + \det(A) = 0.$$

In diesem Fall ist es eine quadratische Gleichung. Mit den beiden Lösungen λ_1, λ_2 löst man nun die Gleichungssysteme

$$\begin{aligned}(A - \lambda_1 E)\underline{v} &= 0, \\ (A - \lambda_2 E)\underline{v} &= 0.\end{aligned}$$

Da diese unterbestimmt sind, sucht man sich Lösungen $\underline{v}_1, \underline{v}_2$ aus, wo z.B. $v_{11} = 1$ ist. Damit kann man leichter rechnen.

Jetzt kann man eine Eigenzerlegung $A = TDT^{-1}$ vornehmen. Dabei ist $D = \text{diag}(\lambda_1, \lambda_2)$. Die Matrix T ist

$$T = [\underline{v}_1, \underline{v}_2] = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}.$$

Man stellt nun fest, dass

$$A^2 = (TDT^{-1})^2 = TDT^{-1}TDT^{-1} = TD^2T^{-1}.$$

Im Allgemeinen hat man für eine natürliche Potenz

$$A^n = TD^nT^{-1}.$$

Dieses Prinzip funktioniert erstaunlicherweise sogar für reelle Potenzen. Die Formel ist von praktischer Bedeutung, da sie es ermöglicht, große Potenzen einer Matrix ohne großen Aufwand zu berechnen. Es ist nämlich $D^n = \text{diag}(\lambda_1^n, \lambda_2^n)$.

Man kann sogar allgemein eine Funktion f auf die Matrix A anwenden. Sei f eine Funktion, die sich als Potenzreihe darstellen lässt. Man hat dann

$$\begin{aligned}f(A) &= \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} a_k TD^kT^{-1} \\ &= T \left(\sum_{k=0}^{\infty} a_k D^k \right) T^{-1} = Tf(D)T^{-1}.\end{aligned}$$

Außerdem ist $f(D) = \text{diag}(f(\lambda_1), f(\lambda_2))$.

Auch mit dem Satz von Cayley-Hamilton lässt sich die Anwendung einer Funktion auf eine Matrix berechnen. Der Satz von Cayley-Hamilton sagt aus, dass eine Matrix eine Nullstelle ihres eigenen charakteristischen Polynoms ist. Es ist also

$$A^2 - \text{Spur}(A)A + \det(A)E = 0.$$

Mit dieser Formel hat man

$$A^2 = \text{Spur}(A)A - \det(A)E = pA + qE.$$

Damit ergibt sich

$$\begin{aligned}A^3 &= A^2A = (pA + qE)A = pA^2 + qA \\ &= p(pA + qE) + qA = p^2A + qA + pqE \\ &= (p^2 + q)A + pqE = p'A + q'E.\end{aligned}$$

Das setzt sich für höhere Potenzen so fort. Es gibt zwei Zahlen p, q , sodass

$$A^n = pA + qE$$

ist. Bei Partialsummen von Potenzreihen kann man diese Methode auch anwenden. Die Methode funktioniert sogar für Potenzreihen. Sei f eine Funktion, die man als Potenzreihe darstellen kann. Es gibt zwei Zahlen p, q , sodass

$$f(A) = pA + qE$$

ist. Die Frage ist jetzt natürlich, wie man auf diese Zahlen p, q kommt. Weil die gleiche Argumentation auch für die Eigenwerte funktioniert, erhält man

$$\begin{aligned}f(\lambda_1) &= p\lambda_1 + q, \\ f(\lambda_2) &= p\lambda_2 + q.\end{aligned}$$

Das ist ein einfaches lineares Gleichungssystem in den Unbekannten p, q . Man erhält die Lösungen

$$\begin{aligned}p &= \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1}, \\ q &= f(\lambda_1) - p\lambda_1.\end{aligned}$$

Wenn die beiden Eigenwerte zusammenfallen, kann man p mit der angegebenen Formel nicht berechnen. Man sieht aber, dass der Ausdruck beim zusammenfallen zur Ableitung wird. Im Fall $\lambda_1 = \lambda_2$ ist daher

$$p = f'(\lambda).$$

1.9 Multilinearformen

Die Dachprodukte vom Grad k bilden k -Blades. Durch Linearkombination von k -Blades lassen sich alle k -Vektoren bilden. Solche k -Vektoren kann man nun auch als alternierende Multilinearformen auffassen.

Multilinearformen sind einfach lineare Abbildungen, wenn man alle Argumente bis auf eines konstant hält. Alternierend bedeutet, dass sich das Vorzeichen umdreht, wenn zwei Argumente vertauscht werden.

Eine Multilinearform ist nun aber eine Funktion, die Vektoren als Argumente und eine Zahl als Wert hat. Wie kann man k -Vektoren auf Vektoren anwenden?

Dazu betrachtet man zunächst 1-Formen. Das Skalarprodukt schreibt man $\langle u, v \rangle$. Wenn aber keine Orthonormalbasis vorliegt, so muss man den Formalismus mit Kovektoren verwenden. Es lässt sich dann die Formel

$$\langle e^i, e_j \rangle = \delta_{ij}$$

verwenden. Man berechnet dann

$$\omega(v) := \langle \omega, v \rangle = \omega_1 v^1 + \omega_2 v^2$$

wobei ω_k die Komponenten von ω bezüglich der Dualbasis und v^k die Komponenten von v bezüglich der normalen Basis sind. Eben dieses ω ist die Linearform.

Ziel ist es nun, diesen Formalismus auf alternierenden Multilinearformen zu übertragen. Das ist natürlich völlig analog zur Kontraktion von Tensoren.

Allgemein wird die Anwendung einer alternierenden Multilinearform auf einen Vektor mit dem inneren Produkt formuliert. Es gilt die Rechenregel

$$\omega(v_1, \dots, v_n) = (v_1 \lrcorner \omega)(v_2, \dots, v_n).$$

Für eine 1-Form ist einfach $v \lrcorner \omega = \langle \omega, v \rangle$.

Sei a der Grad von A . Mit der Formel

$$v \lrcorner (A \wedge B) = (v \lrcorner A) \wedge B + (-1)^a A \wedge (v \lrcorner B)$$

lässt sich jede Berechnung auf 1-Formen zurückführen. Für eine 2-Form gilt einfach

$$\begin{aligned} \omega(u, v) &= (\omega_{12} e^1 \wedge e^2)(u, v) \\ &= \omega_{12} u^1 v^2 - \omega_{12} u^2 v^1. \end{aligned}$$

Es gibt auch eine allgemeine Formel mit Permutationen und Signum.

Allgemeiner kann man rein kovariante Tensoren als Multilinearformen interpretieren. Man definiert dazu die Anwendung auf Vektoren über die Tensorkontraktion. D.h. man definiert

$$\begin{aligned} T(a, b) &:= \sum_{i,j} T_{ij} a^i b^j, \\ T(a, b, c) &:= \sum_{i,j,k} T_{ijk} a^i b^j c^k \end{aligned}$$

usw. Die alternierenden Multilinearformen entsprechen dann genau den antisymmetrischen Tensoren. Wenn also $\omega_{ij} = -\omega_{ji}$ und $\omega_{kk} = 0$ ist, so ergibt sich z.B.

$$\begin{aligned} \omega_{12} e^1 \wedge e^2 &= \omega_{12} e^1 \otimes e^2 - \omega_{12} e^1 \otimes e^2 \\ &= \sum_{i,j} \omega_{ij} e^i \otimes e^j. \end{aligned}$$

Damit erhält man

$$\omega(a, b) = \sum_{i,j} \omega_{ij} a^i b^j = \omega_{12} a^1 b^2 - \omega_{21} a^2 b^1.$$

Kontrahiert man nicht über alle Indizes so lassen sich Tensoren allgemeiner als multilineare Abbildungen interpretieren. Wichtig ist nur, dass bei der Kontraktion immer ein Index oben und der zweite unten steht.

2 Analysis

2.1 Vektorwertige Funktionen

Das Ableiten einer vektorwertigen Funktion ist einfach, da man die Linearität des Differentialoperators ausnutzen kann. Sind e_1 und e_2 die Basisvektoren, so lässt sich die vektorwertige Funktion schreiben als

$$\underline{v}(x) = e_1 v_1(x) + e_2 v_2(x).$$

Es kann sich z.B. um eine Parameterkurve handeln. Die e_k sind wie konstante Faktoren zu behandeln. Somit ergibt sich

$$\begin{aligned} D\underline{v}(x) &= D(e_1 v_1) + D(e_2 v_2) \\ &= e_1 Dv_1 + e_2 Dv_2. \end{aligned}$$

Ein Vektorfeld wird daher abgeleitet, indem es komponentenweise abgeleitet wird. Falls die Basisvektoren e_k auch von x abhängig sind, muss man die Produktregel zu benutzen. Es ist z.B.

$$D(e_1 v_1) = (De_1)v_1 + e_1 Dv_1.$$

Da Matrizen ebenfalls Vektoren sind, können auch sie komponentenweise abgeleitet werden. Das selbe gilt für Tensoren. Die Produktregel gilt außerdem für Skalarprodukte. Es ist

$$\begin{aligned} D\langle \underline{a}, \underline{b} \rangle &= D\left(\sum_k a_k b_k\right) = \sum_k D(a_k b_k) \\ &= \sum_k [(Da_k)b_k + a_k Db_k] \\ &= \sum_k (Da_k)b_k + \sum_k a_k Db_k \\ &= \langle Da, b \rangle + \langle a, Db \rangle. \end{aligned}$$

Weiterhin gilt die Produktregel für Tensorprodukte und damit auch für äußere Produkte.

Da das bestimmte Integral auch linear ist, können vektorwertige Funktionen auch komponentenweise integriert werden, solange die Basisvektoren konstant sind. Allgemein gilt die Argumentation für beliebige lineare Operatoren.

2.2 Funktionen mit mehreren Argumenten

Man kann anstelle einer Funktion $f(x)$ mit einem Argument auch eine Funktion mit zwei Argumenten $f(x_1, x_2)$ definieren. Wie leitet man eine solche Funktion ab? Zunächst kann man die partiellen Ableitungen von f berechnen. Dabei tut man so, als wären alle Variablen, bis auf eine, Konstanten. Dann leitet man die Funktion ganz normal nach dieser Variable ab. Zum Beispiel ist

$$\begin{aligned} \frac{\partial}{\partial x_1} f(x_1, x_2) &= \frac{\partial}{\partial x_1} (x_1^2 x_2 + x_1 x_2^2) \\ &= \frac{d}{dx} (x^2 a + x a^2) = 2xa + a^2 \\ &= 2x_1 x_2 + x_2^2. \end{aligned}$$

Kurze Schreibweisen für partielle Ableitungen sind

$$D_1 f = \partial_1 f = \frac{\partial}{\partial x_1} f.$$

Die partiellen Ableitungen lassen sich zu einem Vektor zusammenfassen. Dieser Vektor wird dann als Nabla-Operator bezeichnet. Es ist

$$\nabla = D_1 e_1 + D_2 e_2$$

bzw. allgemeiner

$$\nabla = D_1 e_1 + \dots + D_n e_n.$$

Die Argumente einer Funktion lassen sich auch zu einem Vektor zusammenfassen. Man schreibt dann kurz

$$f(\underline{x}) = f(x_1, \dots, x_n).$$

Eine Funktion kann einer Zahl auch einen Vektor zuordnen. Ein Beispiel ist die Parameterkurve

$$\underline{f}(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}.$$

Es ergibt sich der Einheitskreis. Eine solche Funktion wird abgeleitet, indem sie nach jeder Komponente abgeleitet wird. Wertet man die Ableitung an der Stelle t aus, so erhält man einen Tangentialvektor.

Nun kann man doch auch eine Funktion definieren, welche sowohl mehrere Argumente hat, als auch vektorwertig ist. Man schreibt dann $\underline{f}(\underline{x})$ oder auch einfach wieder $f(\underline{x})$ bzw. $f(\underline{x})$.

Wie leitet man eine solche Funktion ab? Man kann bezüglich einer Komponente nach einer Variable partiell ableiten. Alle Ableitungen fasst man zur Jacobi-Matrix zusammen. Die Jacobi-Matrix verallgemeinert also den Nabla-Operator. Die Jacobi-Matrix ist definiert durch

$$J = Df = (D_j f_i).$$

Dabei sind $f_i(\underline{x})$ die Komponenten der vektorwertigen Funktion f .

Die Tangente der reellen Funktion f an der Stelle x_0 ist durch die Gleichung

$$T(x) = f'(x_0)(x - x_0) + f(x_0)$$

gegeben. Allgemeiner berechnet man den Tangentialraum der Funktion f an der Stelle \underline{x}_0 mit

$$T(\underline{x}) = Df(\underline{x}_0)(\underline{x} - \underline{x}_0) + f(\underline{x}_0).$$

Der Tangentialraum ist eigentlich ein Vektorraum der seinen Ursprung am Berührungspunkt hat. Daher sollte man hier vom *affinen* Tangentialraum sprechen.

Mit $Df(\underline{x}_0)$ ist die Jacobi-Matrix von f an der Stelle x_0 gemeint. Für eine reellwertige Funktion ergibt sich der Spezialfall

$$T(\underline{x}) = \langle \nabla f(\underline{x}_0), \underline{x} - \underline{x}_0 \rangle + f(\underline{x}_0).$$

Mit den spitzen Klammern ist das Standardskalarprodukt gemeint. Wenn man die kanonische Basis verwendet, dann ist das Skalarprodukt ein Spezialfall der Matrizenmultiplikation.

Man kann eine Funktion $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ als Koordinatentransformation interpretieren. Dafür ist es wichtig, dass die Jacobi-Matrix von f an jeder Stelle invertierbar ist, dass also die Determinante der Matrix nirgendwo verschwindet. Invertierbare Matrizen bezeichnet man auch als regulär bzw. nicht singulär.

Die partiellen Ableitungen $D_k f(\underline{x})$ sind die Tangentialvektoren des Koordinatennetzes an der Stelle \underline{x} .

Ein Beispiel für eine solche Koordinatentransformation ist die Transformation von Polarkoordinaten in rechtwinklige Koordinaten

$$f(r, \varphi) = \begin{bmatrix} f_1(r, \varphi) \\ f_2(r, \varphi) \end{bmatrix} = \begin{bmatrix} r \cos \varphi \\ r \sin \varphi \end{bmatrix}.$$

Bei Skalarfeldern handelt es sich um Funktionen in mehreren Variablen. Vektorfelder sind vektorwertige Funktionen in mehreren Variablen. Parameterkurven sind vektorwertige Funktionen in einer Variable. Parameterflächen sind vektorwertige Funktionen in zwei Variablen.

2.3 Die Kettenregel

Sei $f(\underline{x})$ eine reellwertige Funktion in mehreren Variablen und sei $\underline{u}(t)$ eine dazu passende vektorwertige Funktion in einer Variablen. Die Kettenregel lautet

$$[f(\underline{u})]'(t) = \sum_k D_k f(\underline{u}) u'_k(t)$$

bzw. alternativ

$$\frac{d}{dt}[f(\underline{u})] = \sum_k \frac{\partial f}{\partial u_k} \frac{du_k}{dt}.$$

Das lässt sich mit Nabla-Operator und Skalarprodukt auch kurz schreiben als

$$[f(\underline{u})]'(t) = \langle \nabla f(\underline{u}), \underline{u}'(t) \rangle.$$

Ist \underline{u} auch eine Funktion von mehreren Variablen, so lässt sich \underline{u} als Funktion einer einzigen Variablen auffassen wenn man die entsprechende partielle Ableitung bildet. Damit ergibt sich

$$\frac{\partial f}{\partial x_i} = \sum_k \frac{\partial f}{\partial u_k} \frac{\partial u_k}{\partial x_i} = \sum_k D_k f(\underline{u}) D_i u_k$$

bzw. kurz

$$D_i[f(\underline{u})] = \langle \nabla f(\underline{u}), D_i \underline{u} \rangle.$$

Am allgemeinsten kann man die Kettenregel mit Jacobi-Matrizen formulieren. Es ist dann

$$D[f(\underline{u})] = Df(\underline{u}) Du.$$

Das ist völlig analog zur einfachen Kettenregel

$$[f(u)]' = f'(u) u'.$$

2.4 Kurvenintegrale

Man stelle sich eine Kurve in der Ebene vor. Diese Kurve kann durch eine vektorwertige Funktion $\underline{x}(t)$ angegeben werden. Das wird Parameterdarstellung der Kurve genannt, wobei es sich bei t um den Parameter handelt.

Als Bogenlänge wird der Weg bezeichnet, den man zurücklegt, wenn man sich auf der Kurve bewegt. Man kann sich dabei einen Fußweg, eine Straße, oder einen Fluss vorstellen. Wie kann man die Bogenlänge einer solchen Kurve berechnen? Unter einer stark vergrößernden Lupe wird die Kurve aussehen wie eine Gerade. Das ist nicht immer der Fall, aber wir wollen es voraussetzen. Bei einer Gerade kann man nun den Pythagoras benutzen. Für die kleinen Geradenstücke erhält man also

$$ds = \sqrt{dx_1^2 + dx_2^2}.$$

Man dividiert auf beiden Seiten durch dt und erhält

$$s'(t) = \sqrt{x_1'(t)^2 + x_2'(t)^2} = |\underline{x}'(t)|.$$

Integration auf beiden Seiten bringt

$$L = s_2 - s_1 = \int_{t_1}^{t_2} |\underline{x}'(t)| dt.$$

Als Beispiel soll die Bogenlänge eines Kreises berechnet werden. Als Parameterkurve wird

$$\underline{x}(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}$$

mit $0 \leq t \leq 2\pi$ gewählt. Es ist

$$\underline{x}'(t) = \begin{bmatrix} -\sin t \\ \cos t \end{bmatrix}$$

und $|\underline{x}'(t)| = \sqrt{\sin^2 t + \cos^2 t} = 1$. Damit ergibt sich

$$L = \int_0^{2\pi} 1 dt = 2\pi.$$

Man stelle sich nun eine Gerade vor, auf der sich gleichmäßig verteilt Staub befindet. Ein Sternenstaubsammler bewegt sich nun auf der Gerade entlang. Die Masse an eingesammeltem Staub soll berechnet werden. Man multipliziert dazu die Länge L mit der konstanten Staubbichte. Wenn die Staubbichte $f(x)$ aber nicht konstant ist, so berechnet man die eingesammelte Masse mit

$$m = \int_0^s f(x) dx.$$

Was ist nun, wenn der Staub nicht auf einer Geraden liegt, sondern auf einer Kurve? Jeder Punkt auf der Kurve ist durch den Weg $s(t)$ bestimmt, welchen man zurücklegen muss, um zu diesem Punkt zu gelangen. Zu jedem Punkt gehört außerdem eine Staubbichte. Die eingesammelte Masse berechnet man also mit

$$m = \int_0^{s(t)} f(s) ds = \int_0^t f(s(t)) s'(t) dt.$$

Der letzte Term ergibt sich durch die Substitutionsregel.

Mit der Feststellung $f(s) = f(\underline{x})$ und der Formel $s'(t) = |\underline{x}'(t)|$ ergibt sich schließlich

$$m = \int_0^t f(\underline{x}(t)) |\underline{x}'(t)| dt.$$

Ein solches Integral über eine Kurve bezeichnet man, wie der Name schon sagt, als Kurvenintegral. Die innere Geometrie der Kurve stimmt ja mit der inneren Geometrie einer Gerade überein. Aus diesem Grund stimmt das Kurvenintegral aus Sicht der Bogenlänge mit dem gewöhnlichen Integral überein.

Mit dem Spezialfall $f(s) = 1$ erhält man die Bogenlänge.

2.5 Kurvenintegrale zweiter Art

Ein Kurvenintegral lässt sich so interpretieren, dass die Kurve dabei in ein Skalarfeld eingebettet ist. Jedem Punkt der Kurve kann damit ein Skalar zugeordnet werden. Was ist nun, wenn die Kurve in ein Vektorfeld eingebettet ist? Jedem Punkt der Kurve wird ein Vektor zugeordnet. Wie berechnet man dann das Kurvenintegral?

Nun man kann doch eine Analogiebetrachtung machen und einfach das Produkt gegen ein Skalarprodukt austauschen. Dafür ist es auch notwendig, die Betragsstriche zu entfernen, da man sonst eine Skalarmultiplikation hätte. Sei \underline{v} ein Vektorfeld. Man definiert

$$\int_C \langle \underline{v}(\underline{x}), d\underline{x} \rangle := \int_a^b \langle \underline{v}(\underline{x}(t)), \underline{x}'(t) \rangle dt.$$

Sei $U(\underline{x})$ nun ein beliebiges Skalarfeld. Wenn $\underline{v} = \nabla U$ ist, so heißt \underline{v} Potentialfeld und U ist das dazu gehörige Potential. Man kann nun rechnen

$$\langle \underline{v}, d\underline{x} \rangle = \langle \nabla U, d\underline{x} \rangle = \nabla^b U = dU.$$

Die Kurve C habe den Anfangspunkt $A = \underline{x}(a)$ und den Endpunkt $B = \underline{x}(b)$. Nach dem Satz von Stokes ist

$$\int_C dU = \int_{R(C)} U = U(B) - U(A).$$

Für ein Potentialfeld ist das Kurvenintegral also nur von Randpunkten der Kurve abhängig, jedoch nicht von der Kurve selbst. Das Kurvenintegral ist für ein Potentialfeld *wegunabhängig*.

Damit folgt sofort das Verschwinden des Kurvenintegrals für eine geschlossene Kurve, es ist

$$\oint_C \langle \nabla U, d\underline{x} \rangle = 0.$$

In der Physik wird man noch eine Proportionalitätskonstante einschieben. Das Potentialfeld ist dann $\underline{v} = c\nabla U$. Die Formel lautet dementsprechend

$$\int_C \langle \underline{v}, d\underline{x} \rangle = cU(B) - cU(A).$$

Eine Potentialfeld ist rotationsfrei, denn es ist

$$\nabla \wedge \underline{v} = \nabla \wedge \nabla U = (\nabla \wedge \nabla)U = 0.$$

Verlangt man nun zusätzlich die Gültigkeit der Laplace-Gleichung $\Delta U = 0$, so ergibt sich mit $\Delta U = \langle \nabla, \nabla U \rangle$ dann

$$\nabla \underline{v} = \langle \nabla, \underline{v} \rangle + \nabla \wedge \underline{v} = 0.$$

Wenn \underline{v} ein Vektorfeld in der Ebene ist, so stellt diese Bedingung gerade die Cauchy-Riemann-PDGn dar, wie im Abschnitt Holomorphie erläutert wird. Ein elektrostatisches Feld in einer ladungsfreien Ebene kann also durch eine holomorphe Funktion beschrieben werden. Charakteristisch sind dabei die Feldlinien, die überall rechtwinklig auf den Äquipotential-Isolinien stehen.

Im Raum kann man komplexe Funktionen leider nicht mehr benutzen. Da das Vektorfeld \underline{v} aber die Gleichung $\nabla \underline{v} = 0$ erfüllt, muss es ähnliche Eigenschaften haben. Unter anderem stehen die Feldlinien rechtwinklig auf den Äquipotentialflächen.

2.6 Variationsrechnung

Die kürzeste Verbindung zweier Punkte in der Ebene ist eine Gerade. Wie sieht es aber mit der kürzesten Verbindung zweier Punkte auf einer gekrümmten Oberfläche aus? Zur Berechnung einer solchen Verbindungskurve benötigt man die Variationsrechnung.

In der Ebene hat man eine Funktion $L(x, y, y') = \sqrt{1 + y'^2}$, welche die Ableitung der Bogenlänge $S(f)$ ist. Dabei soll $y = f(x)$ sein. Die Funktion L wird allgemein als Lagrange-Funktion bezeichnet. Die Bogenlänge ist gegeben durch

$$S(f) = \int_a^b L(x, y, y') dx.$$

Wir fragen nun nach einer Funktion f , für welche $S(f)$ am kleinsten ist.

Wir tun nun so, als ob wir diese Funktion f schon kennen. Sei g eine Funktion, die auf $[a, b]$ zweimal stetig differenzierbar ist. Weiter soll $g(a) = 0$ und $g(b) = 0$ sein. Sei p eine Zahl, die auch negativ sein darf. Man kann nun die Funktionen $h(x) = f(x) + pg(x)$ bilden, welche die Funktion f etwas variieren.

Was das soll? Wir können die Bedingung $S(h) = \min$ nun als ein gewöhnliches Extremwertproblem $S(p) = \min$ ansehen. Und für ein solches muss beim Minimum ja die notwendige Bedingung $S'(p) = 0$ gültig sein. Das Minimum erhält man für $p = 0$. Damit muss $S'(0) = 0$ sein.

Wir benutzen nun die Rechenregel

$$\frac{d}{dp} \int_a^b f(x, p) dx = \int_a^b \frac{\partial}{\partial p} f(x, p) dx.$$

Diese Rechenregel ist ein Spezialfall der Rechenregel für Parameterintegrale. Bei p handelt es sich um den

Parameter. In diesem Fall ist damit

$$S'(p) = \int_a^b \frac{\partial L}{\partial p} dx.$$

Man benutzt nun die Kettenregel und erhält

$$S'(p) = \int_a^b \frac{\partial L}{\partial y} g + \frac{\partial L}{\partial y'} g' dx.$$

Mit der Produktregel hat man nun

$$\begin{aligned} \int_a^b \frac{\partial L}{\partial y'} g' dx &= \left[g \frac{\partial L}{\partial y'} \right]_a^b - \int_a^b g \frac{d}{dx} \frac{\partial L}{\partial y'} dx \\ &= - \int_a^b g \frac{d}{dx} \frac{\partial L}{\partial y'} dx. \end{aligned}$$

Insgesamt ist also

$$S'(p) = \int_a^b \left[\frac{\partial L}{\partial y} - \frac{d}{dx} \frac{\partial L}{\partial y'} \right] g dx$$

Die Argumentation ist nun folgende. Da die Gleichung $S'(0) = 0$ für jede Funktion g gelten soll, muss notwendigerweise der andere Faktor gleich null sein. Dass diese Argumentation korrekt ist, ist Inhalt des Fundamentallemmas der Variationsrechnung. Damit erhält man die

Euler-Lagrange-Gleichung

$$\frac{\partial L}{\partial y} - \frac{d}{dx} \frac{\partial L}{\partial y'} = 0.$$

Diese Gleichung stellt eine notwendige Bedingung für $S(f) = \min$ oder $S(f) = \max$ dar.

Man kann die Gleichung noch verallgemeinern. Sei $\underline{v}(t)$ eine vektorwertige Funktion und sei $L(t, v_1, \dots, v_n, v'_1, \dots, v'_n)$ die Lagrange-Funktion. Die Gleichungen

$$\frac{\partial L}{\partial v_k} - \frac{d}{dt} \frac{\partial L}{\partial v'_k} = 0$$

stellen eine notwendige Bedingung für $S(\underline{v}) = \min$ oder $S(\underline{v}) = \max$ dar. Man kann die Euler-Lagrange-Gleichung also komponentenweise formulieren. Trotzdem kann dabei aber ein System von gekoppelten Differentialgleichungen entstehen. In einem solchen Fall ist es nicht möglich, eine der Gleichungen ohne die anderen zu lösen.

Man kann z.B. die kürzeste Verbindung in der Ebene suchen. Man erhält

$$\begin{aligned} \frac{\partial L}{\partial y} &= 0, \quad \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}}, \\ \frac{d}{dx} \frac{\partial L}{\partial y'} &= y'' \frac{\sqrt{1 + y'^2} - y'^2 (\sqrt{1 + y'^2})^{-1}}{1 + y'^2} \\ &= y'' \frac{1}{(1 + y'^2)^{3/2}}. \end{aligned}$$

Damit ergibt sich als notwendige Bedingung die Differentialgleichung

$$y'' = 0.$$

Man hätte anstelle des Bogenlängenfunktionals auch das Energiefunktional verwenden können. Man kommt so mit weniger Rechenaufwand auf die gleiche Differentialgleichung. Die Lösungen dieser Gleichung sind die Funktionen

$$y = mx + n.$$

Ein Beispiel aus der Physik ist der freie Fall. Es ist $L = T - V$, wobei mit T die kinetische und mit V die potentielle Energie gemeint ist. Weiterhin ist $v = s'(t)$. Man hat also

$$L(t, s, v) = \frac{1}{2}mv^2 - mg(-s).$$

Man rechnet nun

$$\frac{\partial L}{\partial s} = mg, \quad \frac{\partial L}{\partial v} = mv, \quad \frac{d}{dt} \frac{\partial L}{\partial v} = mv' = ms''.$$

Als notwendige Bedingung ergibt sich die Differentialgleichung

$$s''(t) = g.$$

Die Lösung dieser Gleichung ist

$$s = \frac{1}{2}gt^2 + v_0t + s_0.$$

Auch für Felder gibt es eine Euler-Lagrange-Gleichung. Dazu ist es notwendig, die Lagrangedichte l von einem Feld $\varphi(t, x_1, x_2, x_3)$ einzuführen. Es ist

$$L = \int_V l dx_1 dx_2 dx_3.$$

Setzt man noch $x_0 = ct$, so ergibt sich

$$S(\varphi) = \int_H l d^4x.$$

Die Lagrangedichte ist von mehreren Argumenten abhängig, man hat

$$l(t, \varphi, D_0\varphi, D_1\varphi, D_2\varphi, D_3\varphi).$$

Mit $D_k\varphi$ ist die partielle Ableitung von φ nach x_k gemeint.

Die Euler-Lagrange-Gleichung ist

$$\frac{\partial l}{\partial \varphi} - \sum_{k=0}^3 \frac{\partial}{\partial x_k} \frac{\partial l}{\partial (D_k\varphi)} = 0.$$

Als Beispiel soll das elektrische Potentialfeld dienen. Bei der Energiedichte $w = (1/2)\varepsilon_0 E^2$ handelt es sich um die Lagrangedichte. Es ist also $l = w$. Als Berechnungsformel gibt es auch

$$l = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu}.$$

Die magnetische Feldkonstante μ_0 hat natürlich nichts mit dem Index μ zu tun. Man könnte auch $F_{\alpha\beta}$ oder

F_{ij} schreiben. Es ist üblich, die Indizes von null bis drei gehen zu lassen und nicht von eins bis vier.

Der Feldstärke-Tensor ist

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & 0 & 0 \\ E_2/c & 0 & 0 & 0 \\ E_3/c & 0 & 0 & 0 \end{bmatrix}.$$

Mit dem metrischen Tensor $g = \text{diag}(1, -1, -1, -1)$ senkt man die Indizes und erhält

$$F_{\mu\nu} = \begin{bmatrix} 0 & E_1/c & E_2/c & E_3/c \\ -E_1/c & 0 & 0 & 0 \\ -E_2/c & 0 & 0 & 0 \\ -E_3/c & 0 & 0 & 0 \end{bmatrix}.$$

Es ergibt sich

$$F^{\mu\nu} F_{\mu\nu} = \frac{2}{c^2} E^2 = \frac{2}{c^2} (E_1^2 + E_2^2 + E_3^2).$$

Nutzt man nun noch $\mu_0 \varepsilon_0 c^2 = 1$ aus, so gelangt man zu

$$l = \frac{1}{2} \varepsilon_0 E^2 = \frac{1}{2} \varepsilon_0 |\nabla \varphi|^2.$$

Man rechnet nun

$$\begin{aligned} \frac{\partial l}{\partial (D_1\varphi)} &= \frac{\varepsilon_0}{2} \frac{\partial}{\partial (D_1\varphi)} \sum_{k=1}^3 (D_k\varphi)^2 = \varepsilon_0 D_1\varphi, \\ \frac{\partial}{\partial x_1} \frac{\partial l}{\partial (D_1\varphi)} &= \varepsilon_0 D_1^2\varphi, \quad \frac{\partial l}{\partial \varphi} = 0. \end{aligned}$$

Die notwendige Bedingung ist also

$$\varepsilon_0 (D_1^2\varphi + D_2^2\varphi + D_3^2\varphi) = 0$$

oder kurz $\varepsilon_0 \Delta\varphi = 0$. Es ist also

$$\Delta\varphi = 0.$$

Diese partielle Differentialgleichung wird Laplace-Gleichung genannt. Es ist die Gleichung, welche elektrostatische Felder in einem ladungsfreien Raum beschreibt, wenn das Potentialfeld am Rand dieses Raumes bekannt ist. Auf jeden Fall wird das Potentialfeld die Laplace-Gleichung erfüllen, auch wenn es nicht auf dem Rand bekannt ist.

Die Lagrangedichte für das elektromagnetische Feld im Vakuum ist

$$l = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} - J_\mu A^\mu.$$

Hierbei ist

$$F_{\mu\nu} = D_\mu A_\nu - D_\nu A_\mu.$$

Man kann auf beiden Seiten zweimal mit dem inversen metrischen Tensor multiplizieren, womit die Indizes gehoben werden. Man erhält

$$F^{\mu\nu} = D^\mu A^\nu - D^\nu A^\mu.$$

An dieser Stelle sollen noch einige Kurznotationen eingeführt werden. Wenn man den Gradient nicht nach den x_k , sondern nach den a_k bilden will, dann beschreibt man das am besten durch

$$\nabla[a] = \frac{\partial}{\partial a} = \sum_k e_k \frac{\partial}{\partial a_k}.$$

Analog kann man auch $D[a_k] = \frac{\partial}{\partial a_k}$ definieren. Damit schreibt man kurz

$$\nabla[\nabla\varphi] = \frac{\partial}{\partial \nabla\varphi} = \sum_k e_k \frac{\partial}{\partial D_k\varphi}.$$

Somit kann man schreiben

$$\langle \nabla, \nabla[\nabla\varphi]l \rangle = \sum_k \frac{\partial}{\partial x_k} \frac{\partial l}{\partial D_k\varphi}.$$

Die Euler-Lagrange-Gleichung lautet nun kurz

$$D[\varphi]l - \langle \nabla, \nabla[\nabla\varphi]l \rangle = 0.$$

Die gewöhnliche Euler-Lagrange-Gleichung ist ein Spezialfall dieser Gleichung. Man erhält sie, wenn man anstelle der Lagrangedichte die gewöhnliche Lagrange-funktion einsetzt. Bei einem Vektorfeld kann die allgemeine Euler-Lagrange-Gleichung wieder komponentenweise formuliert werden.

2.7 Extrema mit Nebenbedingungen

Man denke sich einen Hügel. Der Funktionswert $y = f(x_1, x_2)$ kann als Höhe des Hügels an der Position (x_1, x_2) interpretiert werden. Nun wird durch die Gleichung $g(x_1, x_2) = 0$ ein Weg auf der (x_1, x_2) -Ebene festgelegt, den man entlang wandert. Man hätte auch $g_1(x_1, x_2) = c$ festlegen können. Aber dann kann man ja $g := g_1 - c$ definieren und hat wieder $g(x_1, x_2) = 0$. Dieser Weg wird nun auf die Fläche $f(x_1, x_2)$ hinauf projiziert. Man fragt sich nun, wo das Maximum des Weges auf der Fläche $f(x_1, x_2)$ ist.

Die Höhe beim Maximum sei h und die Position des Maximumspunktes sei $\underline{a} = (a_1, a_2)$. Durch \underline{a} geht eine Isolinie, die implizit durch $f(x_1, x_2) = h$ beschrieben wird. Weiterhin geht durch \underline{a} die Isolinie, die implizit durch $g(x_1, x_2) = 0$ beschrieben wird. Man kann ja auch $g(x_1, x_2)$ als eine Fläche interpretieren. Die beiden Isolinien berühren sich, d.h. sie haben eine gemeinsame Tangente.

Da der Gradient immer rechtwinklig auf der Isolinie steht, sind die beiden Gradienten $\nabla f(\underline{a})$ und $\nabla g(\underline{a})$ kollinear zueinander. Man sucht also eine Stelle \underline{x} wo die Bedingung

$$\nabla f(\underline{x}) = -\lambda \nabla g(\underline{x})$$

erfüllt ist. Das Minuszeichen ist nur Konvention und hat keine weitere Bedeutung. Die Gradienten können gleichsinnig parallel oder entgegengesetzt parallel sein. Durch Umformen erhält man die Bedingung

$$\nabla(f + \lambda g) = 0.$$

Diese Bedingung kann man in die beiden Gleichungen

$$\begin{aligned} D_1(f + \lambda g) &= 0, \\ D_2(f + \lambda g) &= 0. \end{aligned}$$

aufspalten. Zusammen mit der Nebenbedingung $g(x_1, x_2) = 0$ bilden die Gleichungen ein nichtlineares Gleichungssystem, welches die notwendige Bedingung für einen Extrempunkt angibt.

Der Skalierungsfaktor λ heißt Lagrange-Multiplikator. Er gehört auch zur Lösung des Gleichungssystems, ist aber im Ergebnis nicht weiter wichtig.

Die Nebenbedingung kann man auch durch $D_\lambda(f + \lambda g) = 0$ beschreiben. Man definiert nun die Funktion

$$L(x_1, x_2, \lambda) := f(x_1, x_2) + \lambda g(x_1, x_2).$$

Man kann das Gleichungssystem jetzt mit $\nabla L = 0$ abkürzen, wobei der Gradient über alle drei Variablen zu bilden ist. Die Funktion L wird Lagrangefunktion genannt.

Die drei Bedingungen, die erfüllt sein müssen, sind

$$D_1 L = D_2 L = D_\lambda L = 0.$$

Dabei ist noch etwas zu beachten. Wenn $\nabla g = 0$ und $\nabla f \neq 0$ ist, können die Bedingungen nicht erfüllt sein. Der Nullvektor ist ja zu jedem Vektor rechtwinklig, kann also nicht zu ∇f parallel sein. Trotzdem kann an einer solchen Stelle ein Extrempunkt vorliegen.

Das Verfahren funktioniert auch, wenn man eine Funktion von m Variablen hat. Dann ist $\underline{x} = (x_1, \dots, x_m)$. Man kann immer noch $\nabla L = 0$ schreiben. Man erhält jedoch ein System von $m + 1$ Gleichungen. Die Bedingungen sind dann

$$D_1 L = D_2 L = \dots = D_n L = D_\lambda L = 0.$$

Wenn man n Nebenbedingungen $g_k(\underline{x}) = 0$ hat, dann kann man immer noch $\nabla L = 0$ schreiben. Man hat dann jedoch ein System von $m + n$ Gleichungen.

Die Lagrangefunktion ist

$$\begin{aligned} L(x_1, \dots, x_m, \lambda_1, \dots, \lambda_n) \\ = f(x_1, \dots, x_m) + \sum_{k=1}^n \lambda_k g_k(x_1, \dots, x_m). \end{aligned}$$

Man bekommt die Bedingungen

$$\begin{aligned} \frac{\partial L}{\partial x_k} &= 0 \quad (1 \leq k \leq m), \\ \frac{\partial L}{\partial \lambda_k} &= 0 \quad (1 \leq k \leq n). \end{aligned}$$

An dieser Stelle soll noch eine alternative Formulierungsweise erwähnt werden. Da ∇f und ∇g kollinear sind, gilt ja $df \wedge dg = 0$. Das äußere Produkt muss man ausmultiplizieren. Man erhält dann einen Ausdruck der Form

$$\sum_{i < j} a_{ij} dx_i \wedge dx_j = 0.$$

Damit ergibt sich jeweils $a_{ij} = 0$. Ein oberflächlicher Vorteil ist dabei, dass kein lagrangescher Multiplikator vorkommt. Ein tiefer greifender Vorteil ist, dass dieser Rechenweg manchmal eleganter ist.

2.8 Die Gâteaux-Ableitung

Variationsaufgaben lassen sich auch mit der sogenannten Variationsableitung formulieren. Bei dieser Ableitung wird ein Funktional nicht nach einer reellen Variablen, sondern nach einer Funktion abgeleitet.

Die Gâteaux-Ableitung ist nun ein Konzept, das die Richtungsableitung verallgemeinert. Wichtig ist dabei, dass Variationsableitungen auch Gâteaux-Ableitungen sind.

Die Richtungsableitung ist definiert durch

$$\begin{aligned} D_v f &:= \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h} \\ &= \left[\frac{d}{dh} f(x + hv) \right]_{h=0}. \end{aligned}$$

Wenn f aber ein Funktional ist, so sind doch $x(t)$ und $v(t)$ Funktionen. In diesem Fall spricht man von der Gâteaux-Ableitung. Diese sollte linear und stetig bezüglich v sein und man schreibt dann auch $\delta f v$ anstelle von $D_v f$. Die Notation suggeriert somit

$$\delta f(v_1 + v_2) = \delta f v_1 + \delta f v_2$$

und $\delta f(rv) = r\delta f v$. Im Fall eines endlichdimensionalen Vektorraumes ist $\delta f v = \langle \nabla f, v \rangle$. Man fragt sich jetzt natürlich sofort inwiefern sich das auf Hilberträume übertragen lässt.

Die notwendige Bedingung für ein Maximum ist nun, dass die Richtungsableitung in jede beliebige Richtung verschwindet. D.h. wenn x_0 eine Maximumsstelle sein soll, so muss notwendig $(D_v f)(x_0) = 0$ für alle v sein.

Analog lässt sich $\delta f v = 0$ für alle Funktionen $v(t)$ verlangen. Damit ergibt sich der gleiche Ansatz wie bei der Herleitung der Euler-Lagrange-Gleichung.

3 Differentialgeometrie

3.1 Integralsätze

Der gaußsche Integralsatz in der Ebene (Satz von Green), der gaußsche Integralsatz im Raum und der Integralsatz von Stokes. Diese Sätze lassen sich zu einem einzigen vereinigen. Der Dafür notwendige Formalismus soll hier erläutert werden.

Ab jetzt wollen wir Vektoren nicht mehr unterstreichen. Das ist einerseits ungünstig, weil wir Skalarfelder nicht mehr so gut von Vektorfeldern unterscheiden können. Andererseits spart es Schreibaufwand.

Zur Vereinfachung wird zunächst eine Orthonormalbasis gewählt. Wir verwenden nun die Operatoren

$$\begin{aligned} v^\flat &= (v_1 e_1 + v_2 e_2)^\flat = v_1 dx_1 + v_2 dx_2, \\ \omega^\sharp &= (\omega_1 dx_1 + \omega_2 dx_2)^\sharp = \omega_1 e_1 + \omega_2 e_2. \end{aligned}$$

Außerdem wird dx_1 mit e_1 und dx_2 mit e_2 identifiziert. Wenn wir keine Orthonormalbasis haben, dann sind die Differentiale jedoch die dualen Basisvektoren. Die Umrechnung geht dann über den metrischen Tensor.

Der gaußsche Integralsatz in der Ebene lautet

$$\iint_B \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} dx_1 dx_2 = \oint_{\partial(B)} (v_1 dx_1 + v_2 dx_2).$$

Das lässt sich kurz schreiben. Es ist dann

$$\iint_B \nabla^\flat \wedge v^\flat = \oint_{\partial(B)} v^\flat.$$

Sei $\omega = v^\flat$. Der Integralsatz lautet mit dieser Abkürzung

$$\iint_B \nabla^\flat \wedge \omega = \oint_{\partial(B)} \omega.$$

Sei $x = (x_1, \dots, x_n)$. Die Cartan-Ableitung von einem Skalarfeld $f(x)$ ist

$$df = \nabla^\flat f = \sum_k D_k f dx_k = \sum_k \frac{\partial f}{\partial x_k} dx_k.$$

Außerdem gilt die Formel

$$d(f dx_1 \wedge \dots \wedge dx_n) = df \wedge dx_1 \wedge \dots \wedge dx_n.$$

Diese beiden Rechenregeln genügen, um die Cartan-Ableitung zu berechnen. Wir rechnen nun

$$\begin{aligned} d\omega &= d(v^\flat) = d(v_1 dx_1 + v_2 dx_2) \\ &= dv_1 \wedge dx_1 + dv_2 \wedge dx_2 \\ &= (D_1 v_1 dx_1 + D_2 v_1 dx_2) \wedge dx_1 \\ &\quad + (D_1 v_2 dx_1 + D_2 v_2 dx_2) \wedge dx_2 \\ &= D_2 v_1 dx_2 \wedge dx_1 + D_1 v_2 dx_1 \wedge dx_2 \\ &= -D_2 v_1 dx_1 \wedge dx_2 + D_1 v_2 dx_1 \wedge dx_2 \\ &= (D_1 v_2 - D_2 v_1) dx_1 \wedge dx_2. \end{aligned}$$

Wenn ω ein 1-Kovektor (eine 1-Form) ist, dann ist also

$$d\omega = \nabla^b \wedge \omega.$$

Der Integralsatz von Gauß in der Ebene lässt sich nun kurz fassen zu

$$\iint_B d\omega = \oint_{R(B)} \omega.$$

Sei v ein Vektorfeld. Der Integralsatz von Gauß im Raum lautet

$$\iiint_V \operatorname{div} v \, dV = \iint_{R(V)} \langle v, dA \rangle$$

Dabei soll $\langle v, dA \rangle$ das Standardskalarprodukt sein. Ausgeschrieben lautet der Satz

$$\begin{aligned} & \iiint_V D_1 v_1 + D_2 v_2 + D_3 v_3 \, dx_1 dx_2 dx_3 \\ &= \iint_{R(V)} (v_1 dx_2 dx_3 + v_2 dx_3 dx_1 + v_3 dx_1 dx_2). \end{aligned}$$

Nun wird $dx_1 dx_2$ gegen $dx_1 \wedge dx_2$ ersetzt usw. und $dx_1 dx_2 dx_3$ gegen $dx_1 \wedge dx_2 \wedge dx_3$ ersetzt. Nun ist z.B.

$$\begin{aligned} d(v_1 dx_2 \wedge dx_3) &= dv_1 \wedge dx_2 \wedge dx_3 \\ &= (D_1 v_1 dx_1 + D_2 v_2 dx_2 + D_3 v_3 dx_3) \wedge dx_2 \wedge dx_3 \\ &= D_1 v_1 dx_1 \wedge dx_2 \wedge dx_3 + 0 + 0 \\ &= D_1 v_1 dx_1 \wedge dx_2 \wedge dx_3. \end{aligned}$$

Dabei wurde $dx_k \wedge dx_k = 0$ verwendet. Mit

$$\omega = v_1 dx_2 \wedge dx_3 + v_2 dx_3 \wedge dx_1 + v_3 dx_1 \wedge dx_2$$

erhält man den Integralsatz in der Form

$$\iiint_V d\omega = \iint_{R(V)} \omega.$$

Der Integralsatz von Stokes lautet

$$\iint_B \langle \operatorname{rot} v, dA \rangle = \oint_{R(B)} \langle v, dr \rangle.$$

Es ist $dA = (dx_2 dx_3, dx_3 dx_1, dx_1 dx_2)$. Wir können das umschreiben. Es ist

$$\begin{aligned} \langle \operatorname{rot} v, dA \rangle &= (D_2 v_3 - D_3 v_2) dx_2 \wedge dx_3 \\ &+ (D_3 v_1 - D_1 v_3) dx_3 \wedge dx_1 \\ &+ (D_1 v_2 - D_2 v_1) dx_1 \wedge dx_2 \\ &= (D_1 dx_1 + D_2 dx_2 + D_3 dx_3) \\ &\wedge (v_1 dx_1 + v_2 dx_2 + v_3 dx_3) \\ &= \nabla^b \wedge v^b \end{aligned}$$

Auch hier kann man nachrechnen, dass gilt

$$d(v^b) = \nabla^b \wedge v^b.$$

Außerdem ist

$$\langle v, dr \rangle = v^b = v_1 dx_1 + v_2 dx_2 + v_3 dx_3.$$

Definiert man noch $\omega = v^b$, so ergibt sich wieder

$$\iint_B d\omega = \oint_{R(B)} \omega.$$

Der Hauptsatz der Analysis lautet

$$\int_a^b f'(x) \, dx = f(b) - f(a).$$

Da es nur eine Variable gibt, stimmen gewöhnliche Ableitung und partielle Ableitung überein. Es ist also $D_x f(x) = f'(x)$. Damit ist

$$df = D_x f(x) \, dx = f'(x) \, dx.$$

Außerdem ist wollen wir die Rechnung

$$\int_{R[a,b]} f = \int_{[b]-[a]} f = \int_{[b]} f + \int_{-[a]} f = f(b) - f(a)$$

als gültig annehmen. Das ist wieder so ein Formalismus, der *Integration über Ketten* genannt wird. Der Hauptsatz lässt sich jetzt umschreiben zu

$$\int_{[a,b]} df = \int_{R[a,b]} f.$$

Der Hauptsatz und alle Integralsätze lassen sich also mit

$$\int_B d\omega = \int_{R(B)} \omega$$

zusammenfassen.

An dieser Stelle sollen noch zwei wichtige Eigenschaften der Cartan-Ableitung hergeleitet werden. Seien A, B Differentialformen und sei A vom Grad r . Für die Cartan-Ableitung gilt

$$d(A \wedge B) = dA \wedge B + (-1)^r A \wedge dB.$$

Diese seltsame Regel sieht aus wie die Produktregel. Es handelt sich tatsächlich um die Produktregel. Um das einzusehen muss man zunächst $d\omega = \nabla^b \wedge \omega$ schreiben. Sei P vom Grad p und Q vom Grad q . Für das äußere Produkt gilt die Regel

$$P \wedge Q = (-1)^{pq} Q \wedge P.$$

Man geht nun in Analogie zu $D(fg) = (Df)g + fDg$ für die gewöhnliche Ableitung vor. Dazu ist es notwendig einmal $\nabla^b \wedge A$ zu verdrehen. Man macht die Rechnung

$$\begin{aligned} d(A \wedge B) &= \nabla^b \wedge (A \wedge B) \\ &= (\nabla^b \wedge A) \wedge B + (-1)^r A \wedge \nabla^b \wedge B \\ &= dA \wedge B + (-1)^r A \wedge dB. \end{aligned}$$

Für alle ω ist $dd\omega = 0$. Sei v vom Grad eins. Mit $v \wedge v = 0$ macht man die Rechnung

$$dd\omega = \nabla^b \wedge \nabla^b \omega = 0 \wedge \omega = 0.$$

Sei r ein Skalar. Man verwende nun die Rechenregeln $a \wedge r = ra$ und $r(a \wedge b) = ra \wedge b = a \wedge rb$. Für eine null-Form ergibt sich

$$\nabla^b \wedge f = \nabla^b f = df.$$

Für eine eins-Form $\omega = f dx_1 \wedge dx_2$ erhält man

$$\begin{aligned} \nabla^b \wedge \omega &= \nabla^b \wedge f dx_1 \wedge dx_2 = \nabla^b f \wedge dx_1 \wedge dx_2 \\ &= df \wedge dx_1 \wedge dx_2 = d\omega. \end{aligned}$$

Die Formel

$$dA = \nabla^b \wedge A$$

gilt also im allgemeinen, wo A eine null-Form, eins-Form, zwei-Form usw. sein kann.

3.2 Die Richtungsableitung

Von dem Folgenden sollte man sich unbedingt eine Skizze auf einem Blatt Papier machen. Die partiellen Ableitungen lassen sich leicht erklären. Man stellt sich ein Skalarfeld $f(x_1, x_2)$ als eine Oberfläche vor. Jeder Stelle $\underline{x} = (x_1, x_2)$ wird dabei eine Höhe zugeordnet. Nun macht man parallel zu einer der Koordinatenachsen einen Schnitt, indem man eine der Variablen konstant hält. Halten wir $x_2 = b$ konstant. Es ergibt sich eine Schnittkurve. Diese ist eine reelle Funktion $g(x) = f(x, b)$ mit der nicht festgehaltenen Koordinate als Variable. Von der Funktion g kann dann die Ableitung an der Stelle a bestimmt werden. Das ist dann der Anstieg von f an der Stelle $\underline{x} = (a, b)$ in Richtung der Koordinatenachse $(x_1, 0)$. Dieser Anstieg ist die partielle Ableitung

$$D_1 f(a, b) = \frac{\partial f}{\partial x_1}(a, b).$$

Nun liegt es nicht fern sich zu fragen, wie der Anstieg in eine beliebige Richtung bestimmt wird. Das wäre dann eine wesentliche Verallgemeinerung der partiellen Ableitung. Um die Ableitung in eine beliebige Richtung zu bilden dient, wie der Name schon sagt, die Richtungsableitung. Sei \underline{v} ein Einheitsvektor. Die Richtungsableitung ist definiert durch

$$D_{\underline{v}} f(\underline{x}) := \lim_{h \rightarrow 0} \frac{f(\underline{x} + h\underline{v}) - f(\underline{x})}{h}.$$

Es gibt eine Formel für die praktische Berechnung, nämlich

$$D_{\underline{v}} f(\underline{x}) = \langle \nabla f(\underline{x}), \underline{v} \rangle.$$

Das ist eine überraschend einfache Formel. Der wesentliche Berechnungsaufwand liegt in der Bestimmung des Gradienten an der Stelle \underline{x} .

Die partiellen Ableitungen sind tatsächlich Spezialfälle der Richtungsableitung. Sei z.B. $\underline{v} = e_1$. Man erhält dann

$$\begin{aligned} D_{\underline{v}} f &= \langle e_1 D_1 f + e_2 D_2 f, e_1 \rangle \\ &= \langle e_1, e_1 \rangle D_1 f + \langle e_2, e_1 \rangle D_2 f \\ &= D_1 f + 0 = D_1 f. \end{aligned}$$

3.3 Koordinatensysteme

Polarkoordinaten können durch eine Transformation von Polarkoordinaten in kartesische Koordinaten dargestellt werden. Es ist

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = f(r, \varphi) = \begin{bmatrix} r \cos \varphi \\ r \sin \varphi \end{bmatrix}.$$

Seien $x = (x_1, x_2)$ die kartesischen Koordinaten und $u = (r, \varphi)$ die Polarkoordinaten. Allgemein ist eine Transformation von gekrümmten in kartesische Koordinaten gegeben durch $x = f(u)$.

Damit überall zurücktransformiert werden kann, muss die Jacobi-Matrix $J = Df(u)$ an jeder Stelle regulär sein. Es muss also überall $\det J \neq 0$ sein. Das ist bei den Polarkoordinaten nur für $x \neq 0$ der Fall, denn dort lässt sich ja der Winkel nicht bestimmen. Bei Polarkoordinaten ist $\det J = r$.

Die partiellen Ableitungen $D_k f(u)$ sind die Tangentialvektoren des Koordinatennetzes. Die Tangentialvektoren spannen einen Vektorraum, den sogenannten Tangentialraum auf. Die Tangentialvektoren sind die Basisvektoren des Tangentialraums. Wenn die Jacobi-Matrix singulär wird, so gelingt es den Tangentialvektoren nicht, den Tangentialraum aufzuspannen.

Man kann jedem Punkt der Bildmenge $\text{Bild}(f)$ also einen Tangentialraum anheften. Das Problem dabei ist, dass die Basis im Allgemeinen keine Orthonormalbasis sein wird. Uns würde es schon reichen, wenn es eine Orthogonalbasis wäre, denn dann könnten wir sie einfach normieren, indem wir jeden Basisvektor durch seinen Betrag teilen. Aber nicht mal das ist der Fall. Es gibt jetzt zwei Möglichkeiten. Entweder man verwendet ein Orthogonalisierungsverfahren oder man verwendet den Formalismus mit dem metrischen Tensor. Die zweite Möglichkeit scheint eleganter zu sein.

Wir werden mit

$$e_k = D_k x = \partial_k x = \frac{\partial x}{\partial u^k}$$

die Basisvektoren bezeichnen. Mit

$$e^k = dx^k = D^k x = \partial^k x = \frac{\partial x}{\partial u_k}$$

werden die dualen Basisvektoren bezeichnet. Die Basisvektoren lassen sich, wie schon gesagt mit

$$e_k = \frac{\partial x}{\partial u^k}.$$

berechnen. Die Jacobi-Matrix lautet

$$J_{ij} = \frac{\partial x_i}{\partial u^j}.$$

Die Jacobi-Matrix enthält also die gesamte Information über die Basisvektoren und damit auch die gesamte Information über den Tangentialraum. Die Basisvektoren sind die Spaltenvektoren der Jacobi-Matrix.

Die Einträge des metrischen Tensors sind ja die Skalarprodukte der Basisvektoren. Es ist

$$g_{ij} = \langle D_i x, D_j x \rangle = \left\langle \frac{\partial x}{\partial u^i}, \frac{\partial x}{\partial u^j} \right\rangle.$$

Der metrische Tensor lässt sich auch als Gramsche Matrix von J berechnen. Es ist

$$g = \text{Gram}(J) = J^T J.$$

Die Umkehrfunktion der Transformation $x(u)$ ist $u(x)$. Mit dieser bestimmt man

$$(J^{-1})_{ij} = \frac{\partial u^i}{\partial x_j}.$$

Das Kronecker-Delta ist definiert durch

$$\delta_{ij} = [i = j] = \begin{cases} 1 & \text{wenn } i = j, \\ 0 & \text{sonst.} \end{cases}$$

Die Kronecker-Deltas sind die Einträge der Einheitsmatrix. Es ist noch zu überprüfen, ob J^{-1} wirklich die inverse Matrix von J ist. Das schafft die Rechnung

$$\begin{aligned} (J^{-1}J)_{ij} &= \sum_k (J^{-1})_{ik} J_{kj} \\ &= \sum_k \frac{\partial u^i}{\partial x_k} \frac{\partial x_k}{\partial u^j} = \frac{\partial u^i}{\partial u^j} = \delta_{ij}. \end{aligned}$$

Dabei haben wir uns die Kettenregel zu nutze gemacht. Diese Rechnung lässt sich auch kompakt formulieren. Sei u die Variable und $f(u)$ die Funktion mit $x = f(u)$. Es ist

$$J^{-1}J = Df^{-1}(f)Df = D[f^{-1}(f)] = Du = E.$$

Es ist also tatsächlich

$$J^{-1}J = E.$$

Die inverse Matrix des metrischen Tensors lässt sich mit der inversen Jacobi-Matrix ausdrücken, denn es ist

$$g^{-1} = (J^T J)^{-1} = J^{-1} (J^T)^{-1} = J^{-1} (J^{-1})^T.$$

Das heißt

$$g^{-1} = \text{Gram}((J^{-1})^T).$$

Mit den dualen Basisvektoren geht man analog vor und definiert

$$g^{ij} = \langle e^i, e^j \rangle.$$

Es ist nun so, dass ein Basisvektor zu jedem dualen Basisvektor außer dem mit dem gleichen Index rechtwinklig ist. Haben Basisvektor und dualer Basisvektor

den gleichen Index, so ist ihr Skalarprodukt eins. Man fasst diese beiden Feststellungen zusammen zu

$$\langle e^i, e^j \rangle = \delta_{ij}.$$

Weiterhin können wir Indizes mit dem metrischen Tensor ja heben und senken. Daher ergibt sich

$$\begin{aligned} \langle e^i, e^j \rangle &= \left\langle \sum_k g^{ik} e_k, e_j \right\rangle \\ &= \sum_k g^{ik} \langle e_k, e_j \rangle = \sum_k g^{ik} g_{kj}. \end{aligned}$$

In Matrixschreibweise ist das

$$E = (\delta_{ij}) = g' g.$$

Bei $g' = (g^{ij})$ muss es sich also um g^{-1} handeln.

Für den metrischen Tensor haben wir die Formel

$$g_{ij} = \sum_k \frac{\partial x_k}{\partial u^i} \frac{\partial x_k}{\partial u^j}.$$

Wir wollen so eine Formel auch für g^{ij} haben.

Mit $g^{ij} = J^{-1}(J^{-1})^T$ bekommt man

$$g^{ij} = \sum_k \frac{\partial u^i}{\partial x_k} \frac{\partial u^j}{\partial x_k}.$$

3.4 Der Gradient

Sei $a(u)$ ein Skalarfeld. In gekrümmten Koordinaten lässt sich der Gradient nicht mehr wie in kartesischen Koordinaten berechnen. Wir finden eine allgemeine Formel für den Gradient, indem wir ihn in kartesischen Koordinaten betrachten. Sei (b_k) eine Orthonormalbasis. Es ist

$$\begin{aligned} \text{grad } a &= \sum_k \frac{\partial a}{\partial x_k} b_k \\ &= \sum_k \frac{\partial a}{\partial x_k} \frac{\partial u}{\partial x_k} = \sum_{k,i} \frac{\partial a}{\partial u^i} \frac{\partial u^i}{\partial x_k} \frac{\partial u}{\partial x_k} \\ &= \sum_{k,i,j} \frac{\partial a}{\partial u^i} \frac{\partial u^i}{\partial x_k} \frac{\partial u^j}{\partial x_k} e_j = \sum_{i,j} \frac{\partial a}{\partial u^i} g^{ij} e_j. \end{aligned}$$

Der Gradient soll als Beispiel in Zylinderkoordinaten $u = (r, \varphi, z)$ berechnet werden. In Zylinderkoordinaten berechnet man $g = \text{diag}(1, r^2, 1)$. Damit ergibt sich $g^{-1} = \text{diag}(1, r^{-2}, 1)$. Der Gradient ist also

$$\text{grad } a = \frac{\partial a}{\partial r} e_r + \frac{1}{r^2} \frac{\partial a}{\partial \varphi} e_\varphi + \frac{\partial a}{\partial z} e_z.$$

Die Basisvektoren sind $e_r = (\cos \varphi, \sin \varphi, 0)$ und $e_\varphi = (-r \sin \varphi, r \cos \varphi, 0)$ sowie $e_z = (0, 0, 1)$. Damit ergeben sich die Beträge $|e_r| = 1$ und $|e_\varphi| = r$ sowie $|e_z| = 1$. Mit den normierten Basisvektoren

$$\hat{e}_k = \frac{e_k}{|e_k|}$$

lautet der Gradient nun

$$\text{grad } a = \frac{\partial a}{\partial r} \hat{e}_r + \frac{1}{r} \frac{\partial a}{\partial \varphi} \hat{e}_\varphi + \frac{\partial a}{\partial z} \hat{e}_z.$$

Eine elegante Formulierung des Gradienten soll noch vorgestellt werden. Dazu erinnern wir uns an die musikalischen Operatoren. Allgemein ist

$$v^b = \left(\sum_k v^k e_k \right)^b = \sum_{k,i} g_{ki} v^k dx^i,$$

$$\omega^\sharp = \left(\sum_k \omega_k dx^k \right)^\sharp = \sum_{k,i} g^{ki} \omega_k e_i.$$

Die Cartan-Ableitung eines Skalarfeldes a berechnet man in kartesischen Koordinaten mit $da = \nabla^b a$. Man kann jetzt auf beiden Seiten der Gleichung heben und erhält

$$\text{grad } a = (da)^\sharp.$$

Diese schöne Formel ist auch in gekrümmten Koordinaten gültig. Es ergibt sich nämlich

$$(da)^\sharp = \left(\sum_i \frac{\partial a}{\partial u^i} dx^i \right)^\sharp = \sum_{i,j} \frac{\partial a}{\partial u^i} g^{ij} e_j.$$

3.5 Volumina

In kartesischen Koordinaten berechnet man Abstände mit dem Satz des Pythagoras. In gekrümmten Koordinaten geht das nicht mehr. Man denke sich eine Gerade in kartesischen Koordinaten. Nach der Transformation in gekrümmte Koordinaten sieht die Gerade auch gekrümmt aus. Wie berechnen wir nun davon die Länge? Nun ja, unter einer starken Lupe sieht die Kurve ja noch gerade aus. Es ist nicht mehr

$$s^2 = \Delta x^2 + \Delta y^2.$$

Jedoch haben wir noch

$$ds^2 = dx^2 + dy^2.$$

Die kleinen Teile muss man nur alle aufsummieren. Es ergibt sich

$$s = \int \sqrt{dx^2 + dy^2} = \int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt.$$

Die allgemeine Formel lautet

$$s = \int_{t_1}^{t_2} |\underline{x}'(t)| dt.$$

Das Quadrat des Wegelements ist

$$ds^2 = \sum_k dx_k^2.$$

Bisher haben wir uns immer noch in kartesischen Koordinaten bewegt. Die Transformation ist $x(u)$. Zur

Abkürzung wird die Schreibweise $\dot{x} = x'(t)$ benutzt. Mit der Kettenregel ergibt sich

$$\begin{aligned} \dot{x}_1^2 + \dot{x}_2^2 &= (D_1 x_1 \dot{u}_1 + D_2 x_1 \dot{u}_2)^2 + (D_1 x_2 \dot{u}_1 + D_2 x_2 \dot{u}_2)^2 \\ &= (D_1 x_1)^2 \dot{u}_1^2 + 2D_1 x_1 D_2 x_1 \dot{u}_1 \dot{u}_2 + (D_2 x_1)^2 \dot{u}_2^2 \\ &\quad + (D_1 x_2)^2 \dot{u}_1^2 + 2D_1 x_2 D_2 x_2 \dot{u}_1 \dot{u}_2 + (D_2 x_2)^2 \dot{u}_2^2 \\ &= [(D_1 x_1)^2 (D_1 x_2)^2] \dot{u}_1^2 \\ &\quad + 2[D_1 x_1 D_2 x_1 + D_1 x_2 D_2 x_2] \dot{u}_1 \dot{u}_2 \\ &\quad + [(D_2 x_1)^2 + (D_2 x_2)^2] \dot{u}_2^2 \\ &= g_{11} \dot{u}_1^2 + 2g_{12} \dot{u}_1 \dot{u}_2 + g_{22} \dot{u}_2^2. \end{aligned}$$

Damit ist das Wegelement

$$ds^2 = \sum_{i,j} g_{ij} du^i du^j.$$

Mit dem metrischen Tensor senkt man einen der Indizes und erhält

$$ds^2 = \sum_k du^k du_k.$$

Die Formel für die Bogenlänge ist

$$s = \int_{t_1}^{t_2} \sqrt{\langle gu', u' \rangle} dt.$$

Als Beispiel soll der Umfang eines Kreises in Polarkoordinaten berechnet werden. Der metrische Tensor ist $g = (1, r^2)$. Man rechnet

$$\begin{aligned} u(t) &= r e_1 + t e_2, & u'(t) &= e_2, \\ gu'(t) &= r^2 e_2, & \langle gu', u' \rangle &= r^2. \end{aligned}$$

Es ergibt sich

$$s = \int_0^{2\pi} r dt = 2\pi r.$$

Sei $h(u)$ eine Koordinatentransformation mit Definitionsbereich $M \subseteq \mathbf{R}^n$ und Zielmenge \mathbf{R}^n . Sei $V = h(M)$. Wenn $h(u)$ bestimmte Voraussetzungen erfüllt, dann ist nach dem Transformationssatz

$$V = \int_V dV = \int_M |\det Dh(u)| du = \int_M |\det J| du.$$

Der Transformationssatz kann eine Erleichterung bei der Berechnung des Volumens bringen. Berechnen wir z.B. den Flächeninhalt eines Kreises ohne Transformationssatz. Ein viertel Kreis hat den Flächeninhalt

$$\begin{aligned} \frac{A}{4} &= \int_B dx dy = \int_B dy dx = \int_0^r \int_0^{y(x)} dy dx \\ &= \int_0^r y(x) dx = \int_0^r \sqrt{r^2 - x^2} dx \\ &= \frac{1}{2} \left[r^2 \arcsin\left(\frac{x}{r}\right) + x \sqrt{r^2 - x^2} \right]_0^r \\ &= \frac{1}{2} r^2 \frac{\pi}{2} = \frac{1}{4} \pi r^2. \end{aligned}$$

Jetzt die Berechnung des Flächeninhalts mit Transformationssatz. Bei Polarkoordinaten hat man $u_1 = r$ und $u_2 = \varphi$ sowie $\det h(u) = r$.

$$\begin{aligned} A &= \int_A dx dy = \int_0^{2\pi} \int_0^r r dr d\varphi \\ &= \int_0^{2\pi} \frac{1}{2} r^2 d\varphi = \frac{1}{2} r^2 2\pi = \pi r^2. \end{aligned}$$

Es ist notwendig, zuvor die Determinante der Jacobi-Matrix zu bestimmen. Der zentrale Punkt ist hier aber, dass man auf die Stammfunktion von $\sqrt{r^2 - x^2}$ verzichten kann.

Was ist, wenn man eine Parameterfläche in einem Raum hat und man den Flächeninhalt mit einem Doppelintegral bestimmen möchte? Bei den euklidischen Koordinaten wissen wir gar nicht, wie wir das integrieren sollen. Es sind ja drei Differentiale, obwohl man für ein Doppelintegral zwei benötigt.

Besser ist es, über die Parameter der Fläche zu integrieren. Da hat man auch zwei Differentiale. Beim Transformationssatz ergibt sich die Beschränkung, dass die Jacobimatrix quadratisch sein muss. Wir machen nun die Rechnung

$$\begin{aligned} |\det J| &= \sqrt{\det J \det J} = \sqrt{\det J^T \det J} \\ &= \sqrt{\det(J^T J)} = \sqrt{\det g}. \end{aligned}$$

Man bekommt nun die allgemeine Formel

$$V = \int_M \sqrt{\det g} du.$$

Man kann außerdem die Formel

$$|v_1 \wedge \dots \wedge v_n|^2 = \det(\langle v_i, v_j \rangle)$$

benutzen. Mit $x = h(u)$ erhält man

$$\sqrt{\det g} = |D_1 x \wedge \dots \wedge D_n x|.$$

Wenn $h(u)$ nur von der Variable $u = t$ abhängig ist, dann ergibt sich der Spezialfall $|D_1 x| = |x'(t)|$. Damit erhält man die schon bekannte Formel für die Bogenlänge.

Der Transformationssatz ist eine Verallgemeinerung der Substitutionsregel. Die Substitutionsregel erhält man wiederum aus der Kettenregel. Sei $h(x)$ eine reelle Funktion. Das Differential ist $dh = h'(x)dx$. Man erhält

$$\int_{h(a)}^{h(b)} f(h) dh = \int_a^b f(x) h'(x) dx.$$

Wenn wir die Rechnung auf höhere Dimensionen übertragen wollen, dann müssen wir zunächst schreiben

$$\int_B dx_1 dx_2 = \int_B dx_1 \wedge dx_2.$$

Wenn $x_k(u)$ eine Funktion der u_k ist, dann erhält man

$$dx_k = D_1 x_k du_1 + D_2 x_k du_2.$$

Einsetzen bringt

$$\begin{aligned} dx_1 \wedge dx_2 &= (D_1 x_1 du_1 + D_2 x_1 du_2) \\ &\wedge (D_1 x_2 du_1 + D_2 x_2 du_2) \\ &= (D_1 x_1 D_2 x_2 - D_2 x_1 D_1 x_2) du_1 \wedge du_2 \\ &= \det(J) du_1 \wedge du_2. \end{aligned}$$

Für eine quadratische Jacobi-Matrix bekommt man allgemein

$$dx_1 \wedge \dots \wedge dx_n = \det(J) du_1 \wedge \dots \wedge du_n.$$

3.6 Vektorfelder

Das Ableiten von Skalarfeldern ist unproblematisch. Beim Ableiten von Vektorfeldern muss man jedoch beachten, dass die Basisvektoren nun vom Ort abhängig sind. Daher verwendet man die Produktregel der Differentialrechnung. Man erhält

$$D_i v = D_i \sum_j v^j e_j = \sum_j (e_j D_i v^j + v^j D_i e_j).$$

Man kommt jetzt auf die Idee, die Ableitungen $D_i e_j$ als Linearkombination der Basisvektoren e_k darzustellen. Das sollte möglich sein, weil die Ableitung eines Vektors wieder ein Vektor sein wird. Die gekrümmte Fläche $x(u)$ wird mit der Orthonormalbasis (b_k) dargestellt. Man schreibt

$$\begin{aligned} e_j &= D_j x = D_j \sum_l b_l x_l \\ &= \sum_l b_l D_j x_l = \sum_l \frac{\partial x}{\partial x_l} D_j x_l. \end{aligned}$$

Mit der Kettenregel erhält man

$$D_i e_j = \sum_{k,l} \frac{\partial x}{\partial u^k} \frac{\partial u^k}{\partial x_l} D_i D_j x_l = \sum_{k,l} e_k \frac{\partial u^k}{\partial x_l} D_i D_j x_l.$$

Man definiert nun die Abkürzung

$$\Gamma_{ij}^k := \sum_l \frac{\partial u^k}{\partial x_l} D_i D_j x_l.$$

Die Γ_{ij}^k werden als Christoffelsymbole bezeichnet. Damit ist die Linearkombination

$$D_i e_j = \sum_k \Gamma_{ij}^k e_k.$$

Die Ableitung des Vektorfeldes ist hiermit

$$\begin{aligned} D_i v &= \sum_j (e_j D_i v^j + v^j \sum_k \Gamma_{ij}^k e_k) \\ &= \sum_j e_j D_i v^j + \sum_{k,j} v^j \Gamma_{ij}^k e_k. \end{aligned}$$

Da der Index j in der ersten Summe durch die Summe gebunden ist, kann man ihn durch k austauschen. Man

erhält

$$\begin{aligned} D_i v &= \sum_k e_k D_i v^k + \sum_{k,j} v^j \Gamma_{ij}^k e_k \\ &= \sum_k e_k (D_i v^k + \sum_j v^j \Gamma_{ij}^k). \end{aligned}$$

Man definiert nun die kovariante Ableitung

$$\nabla_i v^k := D_i v^k + \sum_j v^j \Gamma_{ij}^k.$$

Damit ist

$$D_i v = \sum_k e_k \nabla_i v^k.$$

Die Schreibweise $\nabla_i v^k$ ist etwas unglücklich, da die Information des gesamten Vektorfeldes notwendig ist, um die Ableitung bilden zu können. Die Komponente v^k allein reicht dafür nicht aus. Man muss $\nabla_i v^k = (\nabla_i v)^k$ lesen. Wenn man $\nabla_i v^k = \nabla_i(v^k)$ liest, dann ergibt das keinen Sinn.

Die Richtungsableitung soll nun auf Vektorfelder verallgemeinert werden. Sei a ein Vektor. Für ein Skalarfeld f ist die Richtungsableitung

$$\begin{aligned} D_a f &= \langle a, \text{grad } f \rangle = \left\langle \sum_k a^k e_k, \sum_k e^k D_k f \right\rangle \\ &= \sum_k a^k D_k f. \end{aligned}$$

Obwohl der Gradient die $e^k \equiv dx^k$ enthält, ist es bei der Richtungsableitung wieder unerheblich, ob man kartesische Koordinaten hat oder nicht.

Die Richtungsableitung von einem Vektorfeld in kartesischen Koordinaten kann man bilden, indem man die Richtungsableitung komponentenweise bildet. D.h. man kann die Komponenten eines Vektorfeldes als Skalarfelder betrachten. Bei gekrümmten Koordinaten macht meine eine Rechnung, die zur Rechnung oben analog ist. Man erhält zunächst

$$D_a e_j = \sum_{k,i} a^i \Gamma_{ij}^k e_k.$$

Man definiert

$$\nabla_a v^k := D_a(v^k) + \sum_{i,j} a^i v^j \Gamma_{ij}^k.$$

Damit hat man

$$D_a v = \sum_k e_k \nabla_a v^k.$$

Man kann nachrechnen, dass $\nabla_{e_i} v^k = \nabla_i v^k$ ist. Um das zu überprüfen macht man die Rechnung $(e^i)^l = \langle e^i, e_l \rangle = \delta_{il}$. Damit ist

$$\nabla_{e_i} v^k = D_{e_i}(v^k) + \sum_{l,j} (e^i)^l v^j \Gamma_{il}^k = D_i v^k + \sum_j v^j \Gamma_{ij}^k.$$

Wir wollen nun herausfinden, wie man Kovektoren ableitet. Mit der Formel $D_i e_j = \sum_k \Gamma_{ij}^k e_k$ lassen sich umgekehrt die Christoffelsymbole als Projektion darstellen. Es ist $\Gamma_{ij}^k = \langle e^k, D_i e_j \rangle$. Mit der Produktregel ist weiterhin

$$0 = D_i \langle e^k, e_j \rangle = \langle D_i e^k, e_j \rangle + \langle e^k, D_i e_j \rangle.$$

Damit erhält man

$$D_i e^k = - \sum_j \Gamma_{ij}^k e^j$$

Mit dieser Formel findet man

$$\nabla_i v_k = D_i v_k - \sum_j v_j \Gamma_{ik}^j.$$

3.7 Die Volumenform

Wir wollen den Hodge-Stern-Operator definieren. Da dieser Operator etwas kompliziert ist, wollen wir das zunächst auf Orthonormalbasen beschränken. Der Hodge-Operator ist linear, es ist $*(v+w) = *v + *w$ und $*(rv) = r(*v)$.

Sei n die Dimension des Raums. Man definiert

$$*(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(p)}) := \text{sgn}(\sigma) e_{\sigma(p+1)} \wedge \dots \wedge e_{\sigma(n)}.$$

Es ergibt sich der Spezialfall

$$*(e_1 \wedge \dots \wedge e_p) = e_{p+1} \wedge \dots \wedge e_n.$$

Ein Beispiel. Für $n = 4$ soll $*(e_1 \wedge e_4)$ berechnet werden. Man ergänzt die fehlenden Vektoren auf der anderen Seite und erhält

$$*(e_1 \wedge e_4) = s e_2 \wedge e_3.$$

Das Vorzeichen s bestimmt man, indem man die Permutation (1423) per Transposition sortiert und bei jeder Transposition das Vorzeichen wechselt. Man rechnet

$$s = \text{sgn}(1423) = -\text{sgn}(1324) = \text{sgn}(1234) = 1.$$

Man definiert

$$I_n := e_1 e_2 \dots e_n = e_1 \wedge \dots \wedge e_n.$$

Damit definiert man noch $*1 := I_n$ und $*I_n := 1$. Wenn sich n aus dem Kontext ergibt, dann schreibt man auch $I := I_n$.

Außerdem legt man $I_0 := 1$ fest.

Sei A vom Grad r und B vom Grad s . Dann gilt

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

Man kann I_n in drei Teile zerlegen. Es ist

$$I_n = I_{k-1} e_k J_{n-k} = I_{k-1} \wedge e_k \wedge J_{n-k}.$$

Damit erhält man

$$\begin{aligned} I_n e_k &= (-1)^{n-k} I_{k-1} J_{n-k} e_k e_k \\ &= (-1)^{n-k} I_{k-1} J_{n-k}. \end{aligned}$$

Anders herum ist

$$e_k I_n = (-1)^{k-1} e_k e_k I_{k-1} J_{n-k} = (-1)^{k-1} I_{k-1} J_{n-k}.$$

Weiterhin ist

$$*e_k = \text{sgn}(k1 \dots \hat{k} \dots n) I_{k-1} J_{n-k}.$$

Das Zirkumflex bedeutet, dass k dort ausgelassen wird. Man rechnet nun

$$\text{sgn}(1 \dots n) = (-1)^{k-1} \text{sgn}(k1 \dots \hat{k} \dots n).$$

Ein Vergleich bringt

$$*e_k = e_k I_n = (-1)^{n-1} I_n e_k.$$

Für einen Vektor v erhält man also

$$*v = v I_n = (-1)^{n-1} I_n v.$$

Sei a_k ein Vektor aus (e_k) und seien die a_k paarweise verschieden. Man definiert noch die Reversion

$$R(a_1 \dots a_n) := (a_n \dots a_1).$$

Man stellt fest, dass $R(I_n) = (-1)^{(n/2)(n-1)} I_n$ ist. Es ergibt sich weiterhin

$$R(I_p) I_n = R(I_p) I_p I_{n-p} = I_{n-p}.$$

Außerdem ist ja $*I_p = I_{n-p}$. Ein Vergleich bringt $*I_p = R(I_p) I_n$. Wenn nun A eine Permutation von I_p ist, dann hat man $A = s I_p$, wobei s ein Vorzeichen ist. Man erhält also

$$\begin{aligned} *A &= *(s I_p) = s *I_p = s R(I_p) I_n \\ &= R(s I_p) I_n = R(A) I_n. \end{aligned}$$

Das legt den Verdacht nahe, dass die Formel $*A = R(A) I_n$ allgemein gültig ist. Mit der Formel $R(AB) = R(B) R(A)$ kann man auch rechnen

$$*(AB) = R(AB) I_n = R(B) R(A) I_n = R(B) *A.$$

Wir benutzen nun die Formel

$$e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(n)} = \text{sgn}(\sigma) e_1 \wedge \dots \wedge e_n.$$

Damit rechnet man

$$\begin{aligned} &(e_{\sigma(p)} \wedge \dots \wedge e_{\sigma(1)}) I_n \\ &= (e_{\sigma(p)} \wedge \dots \wedge e_{\sigma(1)}) (e_1 \wedge \dots \wedge e_n) \\ &= \text{sgn}(\sigma) (e_{\sigma(p)} \wedge \dots \wedge e_{\sigma(1)}) (e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(n)}) \\ &= \text{sgn}(\sigma) e_{\sigma(p+1)} \wedge \dots \wedge e_{\sigma(n)}. \end{aligned}$$

Damit ist die Formel $*A = R(A) I_n$ bestätigt. Wenn A vom Grad r ist, dann ist $R(A) = (-1)^{(r/2)(r-1)} A$.

Außerdem ist

$$I_n^2 = (-1)^{(n/2)(n-1)} R(I_n) I_n = (-1)^{(n/2)(n-1)}.$$

Weiterhin ist

$$**A = *(R(A) I_n) = R(I_n) A I_n = s R(I_n) I_n A = s A.$$

Für das Vorzeichen gilt

$$\begin{aligned} A I_n &= s I_n A = s R(R(A) R(I_n)) \\ &= s (-1)^{r(r-1)/2} (-1)^{(n(n-1)/2)} R(A I_n). \end{aligned}$$

Der Grad von $A I_n$ ist $n - r$. Multipliziert man auf beiden Seiten mit $A I_n$, so ergibt sich

$$\begin{aligned} (A I_n)^2 &= (-1)^{(n-r)(n-r-1)/2} \\ &= s (-1)^{r(r-1)/2} (-1)^{n(n-1)/2}. \end{aligned}$$

Man formt nach s um. Den Exponenten vereinfacht man zu

$$n^2 + r^2 - n - nr = n(n-1) + r(r-n).$$

Damit erhält man

$$**A = (-1)^{r(n-r)} A.$$

Wie berechnet man den Hodge-Operator nun, wenn man keine Orthonormalbasis hat?

Ab jetzt sollen die e_k nicht mehr unbedingt ein Orthonormalsystem bilden. Die Formel $e_1 e_2 = e_1 \wedge e_2 = -e_2 e_1$ können wir leider nicht mehr benutzen. Diesem Problem soll im Folgenden Abhilfe geschaffen werden.

Das geometrische Produkt von Basisvektoren ist

$$e_1 e_2 = \langle e_1, e_2 \rangle + e_1 \wedge e_2.$$

Da $e_1 \wedge e_2$ senkrecht zu den beiden Basisvektoren steht, ist

$$|e_1|^2 |e_2|^2 = \langle e_1, e_2 \rangle^2 + |e_1 \wedge e_2|^2.$$

Mit dem metrischen Tensor erhält man

$$g_{11} g_{22} = g_{12} g_{12} + |e_1 \wedge e_2|^2$$

bzw. $|e_1 \wedge e_2|^2 = \det g$. Der Basisbivektor hat also einen Flächeninhalt der von eins abweichen kann. Man bildet deswegen den normierten Bivektor

$$I = \frac{1}{\sqrt{\det g}} e_1 \wedge e_2.$$

Es gibt auch noch $I' = \sqrt{\det g} e^1 \wedge e^2$. Welchen man benutzt ist unerheblich, denn es ist $I' = I$. Man kann den Strich also weglassen.

Für höherdimensionale Räume hat man

$$I = \frac{1}{\sqrt{\det g}} e_1 \wedge \dots \wedge e_n.$$

Alternativ ist

$$I' = \sqrt{\det g} e^1 \wedge \dots \wedge e^n.$$

Mit der Formel

$$a_1 \wedge \dots \wedge a_n = \det(a_1, \dots, a_n) e_1 \wedge \dots \wedge e_n$$

macht man die Rechnung

$$\bigwedge_k e^k = \bigwedge_k \sum_{i_k} g^{i_k k} e_{i_k} \\ = \det(g^{i_1 1}, \dots, g^{i_n n}) \bigwedge_k e_k = \det(g) \bigwedge_k e_k.$$

Damit bekommt man $I' = I$.

Sei v ein Vektorfeld. Das geometrische Produkt vI soll berechnet werden. Es ist nun problematisch, dass die Basisvektoren nicht rechtwinklig aufeinander stehen. Jedoch steht e_1 rechtwinklig zu e^2 und e_2 zu e^1 . Das wollen wir ausnutzen. Man hat $e^i e_j = -e_j e^i$ für $i \neq j$.

Der Vektor e_2 lässt sich als Linearkombination von e_1 und e^2 darstellen. Es ist ja $e^2 = g^{21}e_1 + g^{22}e_2$. Umformen bringt

$$e_2 = \frac{1}{g^{22}}(e^2 - g^{21}e_1).$$

Damit erhält man

$$e_1 \wedge e_2 = \frac{1}{g^{22}}(e_1 \wedge e^2 - g^{21}e_1 \wedge e_1) = \frac{1}{g^{22}}e_1 \wedge e^2.$$

Man rechnet nun

$$e_1(e_1 \wedge e_2) = \frac{1}{g^{22}}e_1(e_1 \wedge e^2) = \frac{1}{g^{22}}e_1 e_1 e^2 \\ = \frac{1}{g^{22}}\langle e_1, e_1 \rangle e^2 = \frac{g_{11}}{g^{22}}e^2.$$

Man benutzt nun die Formeln

$$\begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix} = \frac{1}{\det(g)} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix}, \\ \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \det(g) \begin{bmatrix} g^{22} & -g^{12} \\ -g^{21} & g^{11} \end{bmatrix}.$$

Damit erhält man

$$e_1(e_1 \wedge e_2) = \det(g) e^2.$$

Über analoge Rechnung erhält man die anderen Formeln

$$e_2(e_1 \wedge e_2) = -\det(g) e^1, \\ e^1(e^1 \wedge e^2) = \det(g)^{-1} e_2, \\ e^2(e^1 \wedge e^2) = -\det(g)^{-1} e_1.$$

Damit ergibt sich

$$*v = vI = \sqrt{\det g} (v^1 e^2 - v^2 e^1) \\ = \sqrt{\det g} (v^1 dx^2 - v^2 dx^1).$$

Mann kann noch die Indizes senken und erhält

$$*v = \sqrt{\det g} \sum_{i,j} \varepsilon_{ij} v^i dx^j \\ = \sqrt{\det g} \sum_{i,j} \sum_k \varepsilon_{ij} g^{ik} v_k dx^j.$$

Weiterhin ist ja noch $*1 = 1I = I$.

Man verwendet nun

$$(e_1 \wedge e_2)(e_1 \wedge e_2) = -\det(e_1, e_2) = -\det(g).$$

Damit rechnet man

$$*I = R(I)I = \frac{1}{\det g} (e_2 \wedge e_1)(e_2 \wedge e_1) \\ = -\frac{1}{\det g} (e_1 \wedge e_2)(e_2 \wedge e_1) \\ = -\frac{1}{\det g} (-1) \det(g) = 1.$$

Die Formel für Orthogonalbasen lässt sich sehr leicht herleiten. Sei $b_i = \frac{e_i}{|e_i|}$. Die Rechnung ist

$$*(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(k)}) \\ = |e_{\sigma(1)}| \dots |e_{\sigma(k)}| *(b_{\sigma(1)} \wedge \dots \wedge b_{\sigma(k)}) \\ = |e_{\sigma(1)}| \dots |e_{\sigma(k)}| \operatorname{sgn}(\sigma) (b_{\sigma(k+1)} \wedge \dots \wedge b_{\sigma(n)}) \\ = \frac{|e_{\sigma(1)}| \dots |e_{\sigma(k)}|}{|e_{\sigma(k+1)}| \dots |e_{\sigma(n)}|} \operatorname{sgn}(\sigma) \\ (e_{\sigma(k+1)} \wedge \dots \wedge e_{\sigma(n)}).$$

Die Reihenfolge der Basisvektoren ist bei der Determinante des metrischen Tensors nicht wichtig. Man hat

$$|e_{\sigma(k+1)}| \dots |e_{\sigma(n)}| = \frac{\sqrt{\det g}}{|e_{\sigma(1)}| \dots |e_{\sigma(k)}|}.$$

Damit ergibt sich

$$*(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(k)}) \\ = g_{\sigma(1)\sigma(1)} \dots g_{\sigma(k)\sigma(k)} \frac{1}{\sqrt{\det g}} \operatorname{sgn}(\sigma) \\ (e_{\sigma(k+1)} \wedge \dots \wedge e_{\sigma(n)}).$$

Analog dazu ist

$$*(e^{\sigma(1)} \wedge \dots \wedge e^{\sigma(k)}) \\ = g^{\sigma(1)\sigma(1)} \dots g^{\sigma(k)\sigma(k)} \sqrt{\det g} \operatorname{sgn}(\sigma) \\ (e^{\sigma(k+1)} \wedge \dots \wedge e^{\sigma(n)}).$$

3.8 Der Pullback

Bei der Substitutionsregel der Integralrechnung wird eine Substitution vorgenommen. Dabei wird die Variable u gegen die Funktion $u(x)$ ausgetauscht. Die Funktion $f(u)$ wird daher zur Verkettung

$$(f \circ u)(x) = f(u(x)).$$

Beim Differential muss die Substitution natürlich auch ausgeführt werden. Es ergibt sich

$$du(x) = u'(x) dx.$$

Verwendet man die Iverson-Klammer bzw. charakteristische Funktion, so sieht man, dass die Substitution

auch für die Variable zwischen den Grenzen vorgenommen werden muss. Es ist

$$\begin{aligned}
& \int_{u(a)}^{u(b)} f(u) du \\
&= \int_{-\infty}^{\infty} [u(a) \leq u][u \leq u(b)] f(u) du \\
&= \int_{-\infty}^{\infty} [u(a) \leq u(x)][u(x) \leq u(b)] f(u(x)) u'(x) dx \\
&= \int_{-\infty}^{\infty} [a \leq x][x \leq b] f(u(x)) u'(x) dx \\
&= \int_a^b f(u(x)) u'(x) dx.
\end{aligned}$$

Wenn $u(x)$ nicht monoton steigend ist, dann kann man die Ungleichungen nicht so einfach kürzen. Die Substitutionsregel funktioniert in einem solchen Fall aber trotzdem. Der kanonische Beweis der Substitutionsregel ist äußerst elegant, liefert jedoch keine so anschauliche Begründung für die Modifikation der Grenzen.

Die Substitution wird auch als Pullback bezeichnet, wenn der Rechenformalismus auf Differentialformen übertragen wird. Für eine Funktion ist der Pullback gegeben durch

$$u^* f = f \circ u.$$

Der Pullback ist linear. Es ist

$$u^*(\omega_1 + \omega_2) = u^*\omega_1 + u^*\omega_2,$$

und wenn r eine konstante Zahl ist, so ist

$$u^*(r\omega) = ru^*(\omega).$$

Der Pullback verteilt sich außerdem auf Faktoren des äußeren Produktes. Man sagt, er ist multiplikativ. Es ist

$$u^*(\omega_1 \wedge \omega_2) = u^*\omega_1 \wedge u^*\omega_2.$$

Wenn f eine 1-Form ist, so ist ja $f \wedge \omega = f\omega$. Somit ergibt sich auch die Rechenregel

$$u^*(f\omega) = (u^*f)(u^*\omega) = (f \circ u)u^*\omega.$$

Eine weitere wichtige Rechenregel ist die Verträglichkeit mit der Cartan-Ableitung. Es ist

$$u^*(d\omega) = d(u^*\omega).$$

Schließlich gilt noch die Regel

$$(u_1 \circ u_2)^* = u_2^* \circ u_1^*.$$

Setzt man $\omega := f(u)du$, so lässt sich die Substitutionsregel nun umformulieren zu

$$\int_{u[a,b]} \omega = \int_{[a,b]} u^*\omega.$$

Wenn u ein Diffeomorphismus ist, so gilt diese Formel allgemein und enthält den Transformationssatz. Die allgemeine Formel lautet

$$\int_{u(B)} \omega = \int_B u^*\omega.$$

Voraussetzung ist, dass die Orientierung erhalten bleibt, damit man beim Transformationssatz die Betragstriche entfernen kann. Die Formel gilt auch, wenn u^{-1} eine diffeomorphe Karte ist. Da der \mathbf{R}^n ein Koordinatensystem bezüglich einer Orthonormalbasis ist, gilt dort

$$\int_B f(\underline{x}) dx_1 \wedge \dots \wedge dx_n = \int_B f(\underline{x}) dx_1 \dots dx_n.$$

Integrale auf Mannigfaltigkeiten können damit auf gewöhnliche mehrdimensionale Integrale zurückgeführt werden.

3.9 Der Pushforward

Bei $v(f)$ handelt es sich um eine in der Differentialgeometrie übliche Schreibweise für die Richtungsableitung $D_v f$. Man interpretiert das als Anwendung des Vektors v auf die Funktion f . Es ist

$$v(f) = df(v) = \langle df, v \rangle = \sum_k v^k D_k f.$$

Formal kann man aber auch gleich

$$v = \sum_k v^k \frac{\partial}{\partial x^k} = \sum_k v^k D_k$$

schreiben. Dann ist

$$v(f) = v f = \left(\sum_k v^k D_k \right) f = \sum_k v^k D_k f.$$

Seien M und N zwei Mannigfaltigkeiten und sei f eine Funktion auf N . Sei weiterhin F eine Funktion von M nach N . Sei v ein Vektor aus dem Tangentialraum von M am Punkt p . Der *Pushforward* ist definiert durch

$$(F_* v)(f) := v(f \circ F).$$

Mit der Kettenregel rechnet man nun

$$\begin{aligned}
v(f \circ F) &= \sum_k v^k D_k (f \circ F) \\
&= \sum_{k,i} v^k D_i f(F_i) D_k F_i = \sum_{k,i} v^k \frac{\partial f}{\partial F_i} \frac{\partial F_i}{\partial x^k}.
\end{aligned}$$

Man kann den Pushforward als Verallgemeinerung der Richtungsableitung ansehen. Wenn man $F = \text{id}$ setzt, dann erhält man ja wieder die Richtungsableitung.

4 Funktionalanalysis

4.1 Bra-Ket-Notation

Ein Vektor lässt sich als Linearkombination von Basisvektoren darstellen. Es ist

$$\underline{a} = a_1 e_1 + a_2 e_2.$$

Eine alternative Schreibweise dafür ist

$$|a\rangle = a_1 |e_1\rangle + a_2 |e_2\rangle.$$

Bezeichnet man mit $|1\rangle$ und $|2\rangle$ die beiden Basisvektoren, so schreibt man auch

$$|a\rangle = a_1 |1\rangle + a_2 |2\rangle.$$

Die Zahlen a_1 und a_2 sollen nun komplexe Zahlen sein. Damit wird der Vektorraum zu einem komplexen Vektorraum. Wir können auch für einen komplexen Vektorraum ein Skalarprodukt definieren. Das Skalarprodukt darf nie negativ werden, weil die Wurzel aus dem Skalarprodukt immer ein anschauliches Ergebnis sein soll. Würde man das Skalarprodukt als

$$\langle a, b \rangle = \sum_{k=1}^n a_k b_k$$

definieren, so wäre $\langle a, b \rangle \geq 0$ nicht immer gewährleistet. Wir schaffen dem Abhilfe, indem die a_k konjugiert werden. Das heißt, der Imaginärteil der a_k wird negiert.

Definition. Das Standardskalarprodukt von Vektoren aus einem komplexen Vektorraum ist

$$\langle a, b \rangle = \sum_{k=1}^n \bar{a}_k b_k.$$

Das Skalarprodukt ist nach dieser Definition nicht bilinear, sondern sesquilinear. Es ist auch nicht mehr symmetrisch, sondern hermitisch. Es ist

$$\begin{aligned} \langle ra, b \rangle &= \bar{r} \langle a, b \rangle, \\ \langle a, rb \rangle &= r \langle a, b \rangle, \\ \langle b, a \rangle &= \overline{\langle a, b \rangle}. \end{aligned}$$

Die Konjugation kann anstelle von \bar{z} alternativ auch durch z^* ausgedrückt werden.

Man definiert auch

$$\langle a| = [\bar{a}_1, \bar{a}_2], \quad |a\rangle = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Das Skalarprodukt $\langle a|b \rangle$ lässt sich nun durch eine Matrizenmultiplikation bilden. Man definiert außerdem die Adjunktion H als

$$[\bar{a}_1, \bar{a}_2]^H = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}^H = [\bar{a}_1, \bar{a}_2].$$

In Bra-Ket-Notation ist also $\langle a|^H = |a\rangle$ und $|a\rangle^H = \langle a|$. Die Adjunktion soll auch für eine Matrix definiert werden. Sei

$$A^H = (a_{ij})^H = (\bar{a}_{ji}).$$

Sei außerdem $\bar{A} = (\bar{a}_{ij})$. Man schreibt dann auch $A^H = \bar{A}^T$. Die Transposition einer Matrix ist der Spezialfall der Adjunktion. Die Adjunktion geht nämlich in die Transposition über, wenn nur reelle Zahlen verwendet werden. Für eine reelle Matrix ist nämlich $\bar{A} = A$.

Die Adjunktion hat fast die gleichen Rechenregeln, wie die Transposition. Die Rechenregeln sind

$$\begin{aligned} (A + B)^H &= A^H + B^H \\ (AB)^H &= B^H A^H \\ (rA)^H &= \bar{r} A^H \\ (A^H)^{-1} &= (A^{-1})^H \\ (A^H)^H &= A \\ \det(A^H) &= \overline{\det(A)} \end{aligned}$$

Durch den Formalismus mit Zeilenvektoren und Spaltenvektoren kann man das Skalarprodukt durch eine Matrizenmultiplikation von einem Zeilenvektor mit einem Spaltenvektor ausdrücken.

Sei A eine quadratische Matrix. Die Matrizenmultiplikation ist ja assoziativ. Es ist also

$$(\langle v|A)|w\rangle = \langle v|(A|w\rangle).$$

Man kann das auch explizit nachrechnen. Wegen dem Assoziativgesetz schreibt man anstelle von $a(bc)$ auch einfach abc . Das wollen wir bei der Multiplikation von Matrizen auch machen. Wir schreiben also einfach

$$\langle v|A|w\rangle.$$

Für Matrizen ist ja

$$(ABC)^H = C^H B^H A^H.$$

Es ergibt sich daher

$$(\langle v|A|w\rangle)^H = |w\rangle^H A^H \langle v|^H = \langle w|A^H|v\rangle.$$

Außerdem ist

$$(\langle v|w\rangle)^H = |w\rangle^H \langle v|^H = \langle w|v\rangle.$$

Die Komponenten eines Vektors bekommt man durch Skalarprodukte mit den Basisvektoren der Orthonormalbasis. Es ist

$$a_k = \langle e_k|a\rangle.$$

Damit ist

$$|a\rangle = \sum_{k=1}^n a_k |e_k\rangle = \sum_{k=1}^n \langle e_k|a\rangle |e_k\rangle.$$

Da die Multiplikation mit einem Skalar kommutativ ist, können die Faktoren in $a_k|e_k\rangle$ zu $|e_k\rangle a_k$ vertauscht werden. Man klammert $|a\rangle$ dann aus der Summe aus. Übrig bleibt der Operator

$$I = \sum_{k=1}^n |e_k\rangle\langle e_k|.$$

Der Operator I ist der Identitätsoperator. Es ist $I|a\rangle = |a\rangle$.

4.2 Funktionen

Ein Vektor kann als Punkt im \mathbf{R}^n angesehen werden. Zu jedem Punkt gehört genau ein Vektor, welcher vom Koordinatenursprung auf diesen Punkt zeigt. Dieser Punkt ist ein Tupel (a_1, \dots, a_n) von Koordinaten. Dann kann doch auch jede endliche Folge (a_k) als Vektor angesehen werden, denn die Folge ist ja auch ein Tupel.

Ein Tupel ist eine Familie mit der Indexmenge $I = \{k \in \mathbf{N} \mid k \leq n\}$. Wir können doch auch die Indexmenge $I = \mathbf{N}$ wählen, oder auch $I = \mathbf{R}$. Die Familien (a_k) mit $I = \mathbf{R}$, das sind doch die reellen Funktionen. Wir setzen $x = k$ und $f(x) = a_k$.

Sind Funktionen so etwas wie Vektoren?

Die Notation $f = x \mapsto 2x$ soll bedeuten, dass f eine Funktion mit $f(x) = 2x$ ist. Eigentlich müsste man zu jeder Funktion auch den Definitionsbereich und die Zielmenge angeben, aber wenn wir verlangen, dass die Funktionen alle aus dem selben Funktionenraum entstammen, so ist dies nicht notwendig.

Seien f, g reelle Funktionen. Ein Skalarprodukt ist z.B.

$$\langle f, g \rangle = \int_a^b f(x)g(x) \, dx.$$

Seien f, g Funktionen mit komplexen Werten. Ein Skalarprodukt ist z.B.

$$\langle f, g \rangle = \int_a^b \overline{f(x)}g(x) \, dx.$$

Analog zu Vektoren wird mit dem Skalarprodukt die Norm

$$\|f\| = \sqrt{\langle f, f \rangle}$$

definiert. Analog zu Vektoren werden Funktionen f, g mit $\langle f, g \rangle = 0$ als orthogonal zueinander bezeichnet. Analog zu Vektoren wird eine Funktion f normiert genannt, wenn $\|f\| = 1$ ist.

In einem Hilbertraum gibt es Orthonormalbasen. Analog zu Vektoren kann jede Funktion aus dem Raum in eine uneigentliche Linearkombination

$$f = \sum_{k \in I} f_k e_k = \sum_{k \in I} \langle e_k, f \rangle e_k.$$

zerlegt werden. Die Funktionen $e_k(x)$ sind die Basisvektoren, die f_k sind reelle oder komplexe Zahlen.

Beispiel: Sei

$$B = \{x \mapsto 1, x \mapsto \sqrt{2} \cos(k\omega x), \\ x \mapsto \sqrt{2} \sin(k\omega x) \mid k \in \mathbf{N}\}.$$

Die Menge B ist eine Orthonormalbasis, wenn sie als unendliches Tupel aufgeschrieben wird. Das Skalarprodukt ist

$$\langle f, g \rangle = \frac{1}{T} \int_0^T f(x)g(x) \, dx$$

mit $T = 2\pi/\omega$. Wir normieren jetzt z.B. auf $\omega = 1$ und rechnen nach

$$\begin{aligned} \langle x \mapsto 1, x \mapsto \sqrt{2} \cos(x) \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \sqrt{2} \cos x \, dx \\ &= \frac{1}{2\pi} \sqrt{2} [\sin x]_0^{2\pi} = \frac{1}{2\pi} \sqrt{2} [\sin(2\pi) - \sin 0] = 0. \end{aligned}$$

Die beiden Funktionen sind tatsächlich orthogonal. Außerdem ist z.B.

$$\|x \mapsto 1\|^2 = \frac{1}{2\pi} \int_0^{2\pi} dx = \frac{1}{2\pi} [x]_0^{2\pi} = 1.$$

Eine Funktion aus dem Funktionenraum wird mit B zerlegt in die uneigentliche Linearkombination

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k}{\sqrt{2}} \sqrt{2} \cos(k\omega x) + \frac{b_k}{\sqrt{2}} \sqrt{2} \sin(k\omega x) \right].$$

Die Koeffizienten sind

$$\begin{aligned} \frac{a_0}{2} &= \langle x \mapsto 1, f \rangle, \\ \frac{a_k}{\sqrt{2}} &= \langle x \mapsto \sqrt{2} \cos(k\omega x), f \rangle, \\ \frac{b_k}{\sqrt{2}} &= \langle x \mapsto \sqrt{2} \sin(k\omega x), f \rangle. \end{aligned}$$

Sei der Vektor v dargestellt als Linearkombination von Basisvektoren

$$v = v_1 e_1 + v_2 e_2.$$

und sei $B = (e_1, e_2)$ eine Orthonormalbasis. Für Vektoren gilt der Satz des Pythagoras in der Form

$$|v|^2 = v_1^2 + v_2^2.$$

Allgemeiner gilt

$$|v|^2 = \sum_{k=1}^n v_k^2.$$

Analog gibt es für Funktionen aus einem Hilbertraum einen Satz des Pythagoras, der als Parsevalsche Gleichung bezeichnet wird. Es ist

$$\|f\|^2 = \sum_{k \in I} f_k^2$$

mit $f_k = \langle e_k, f \rangle$.

Bei der Multiplikation eines Vektors mit einer Matrix entsteht ein neuer Vektor. Für Funktionen gibt es ein analoges Konzept. Wird ein linearer Operator auf eine Funktion angewendet, so entsteht dabei eine neue Funktion. Zu den linearen Operatoren gehören die Integraltransformationen.

Ein Beispiel ist die Laplace-Transformation. Die Laplace-Transformation ist

$$L(f)(s) = \int_0^\infty f(x)e^{-sx} dx.$$

Die Laplace-Transformation von $f(x) = 1$ ist z.B.

$$\begin{aligned} L(x \mapsto 1)(s) &= \int_0^\infty e^{-sx} dx = -\frac{1}{s} \int_0^\infty e^x dx \\ &= -\frac{1}{s} \lim_{a \rightarrow \infty} (e^{-a} - 1) = \frac{1}{s}. \end{aligned}$$

Aus $x \mapsto 1$ macht die Transformation also $s \mapsto 1/s$.

Drehmatrizen drehen Vektoren, ändern ihren Betrag jedoch nicht. Das Analogon zu Drehmatrizen sind orthogonale bzw. unitäre Operatoren. Unitäre Operatoren erhalten die Norm. Ein Beispiel dafür ist die Fourier-Transformation F . Es ist $\|F(f)\| = \|f\|$.

Für unitäre Operatoren ist $T^{-1} = T^H$, die inverse Transformation ist die adjungierte Transformation. Z.B. ist die Fouriertransformation

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

Sei $z = re^{i\varphi}$ eine komplexe Zahl in Polarform. Die konjugierte Zahl ist $\bar{z} = re^{-i\varphi}$. Auf diese Art können wir auch die Transformation adjungieren. Man erhält

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

4.3 Schwingungen

Eine harmonische Schwingung hat die Gleichung

$$u = \hat{u} \sin(\omega t + \varphi_0).$$

Für den Effektivwert ergibt sich $U_{\text{eff}} = \|u\|$. Mit dieser Interpretation des Effektivwertes hat man die Möglichkeit geometrische Mittel zu benutzen. Z.B. kann man nun die Dreiecksungleichung

$$\|u_1 + u_2\| \leq \|u_1\| + \|u_2\|.$$

verwenden. Die Summe zweier Schwingungen wird immer einen kleineren Effektivwert haben, als die Summe der Effektivwerte der Schwingungen. Dabei scheint es egal zu sein, ob diese Schwingungen harmonisch sind oder nicht.

Weiterhin erhält man mit der ersten binomischen Formel

$$\|u_1 + u_2\|^2 = \|u_1\|^2 + 2\langle u_1, u_2 \rangle + \|u_2\|^2.$$

Wir definieren die Mischspannung $u_m := u + \bar{u}$, wobei \bar{u} der konstante Gleichanteil sein soll. Man sieht leicht ein, dass $\langle u, \bar{u} \rangle = 0$ ist. Außerdem ist $\|\bar{u}\| = \bar{u}$. Man erhält also

$$\|u_m\|^2 = \|u\|^2 + \bar{u}^2.$$

Wenn u_1, u_2 zwei harmonische Schwingungen mit gleicher Kreisfrequenz sind, dann ist

$$\langle u_1, u_2 \rangle = \|u_1\| \|u_2\| \cos \Delta\varphi.$$

Eine beliebige Schwingung kann nun als Fourierreihe von Grundschnungen dargestellt werden. Dann kann man den Effektivwert auch mit der parsevalschen Gleichung berechnen.

Z.B. ist die Fourierreihe der Rechteckschwingung

$$u = \frac{4\hat{u}}{\pi} \sum_{k=1}^\infty \frac{\sin((2k-1)\omega t)}{2k-1}.$$

Mit der parsevalschen Gleichung ist

$$\begin{aligned} \|u\|^2 &= \frac{16\hat{u}^2}{\pi^2} \sum_{k=1}^\infty \left\| \frac{\sin((2k-1)\omega t)}{2k-1} \right\|^2 \\ &= \frac{16\hat{u}^2}{\pi^2} \sum_{k=1}^\infty \frac{1}{2(2k-1)^2} = \frac{8\hat{u}^2}{\pi^2} \sum_{k=1}^\infty \frac{1}{(2k-1)^2} \\ &= \frac{8\hat{u}^2}{\pi^2} \frac{\pi^2}{8} = \hat{u}^2. \end{aligned}$$

Man erhält also $U_{\text{eff}} = \hat{u}$.

Es wird nun klarer, warum bei einer harmonischen Schwingung

$$U_{\text{eff}} = \frac{1}{\sqrt{2}} \hat{u}$$

ist. Es ist so, weil nicht die Schwingungen $\sin(\omega t + \varphi_0)$ die normierten Basisvektoren darstellen, sondern die Schwingungen $\sqrt{2} \sin(\omega t + \varphi_0)$. Das sieht natürlich etwas komisch aus. Verwendet man aber die komplexe Darstellung einer Schwingung, so hat man $\exp(j\omega t + j\varphi_0)$ als normierte Basisvektoren. Mit diesen ist

$$\underline{u} = \|\underline{u}\| e^{j\omega t + j\varphi_0}.$$

Man beachte aber, dass dies kein Amplitudenzeiger sondern ein Effektivwertzeiger ist.

4.4 Basen

Bisher waren die Basen immer abzählbar. Was ist, wenn die Basis nicht abzählbar ist? Die uneigentliche Linearkombination ist dann keine Summe, sondern ein Integral.

Für die Formulierung benötigt man ein Hilfsmittel, das sich Delta-Distribution nennt. Die Delta-Distribution $\delta(x)$ kann man sich als eine Funktion vorstellen, die für $x \neq 0$ den Wert null und bei $x = 0$ den Wert unendlich hat. Der Flächeninhalt der Fläche unter der

Kurve (d.h. zwischen Graph und x -Achse) soll aber trotzdem gleich eins sein.

Für die Delta-Distribution gilt

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

Man definiert

$$\delta_a(x) := \delta(x-a).$$

Das Skalarprodukt definiert man mit

$$\langle f, g \rangle = \int_{-\infty}^{\infty} \overline{f(x)} g(x) dx.$$

Man kann nun überprüfen, dass $\langle \delta_a, \delta_b \rangle = \delta(a-b)$ ist. Das ist analog zu $\langle e_i, e_j \rangle = \delta_{ij}$. Die Menge der δ_a kann man als Orthonormalbasis interpretieren. Man schreibt nun kürzer $f(a) = \langle \delta_a, f \rangle$.

Damit kann man eine uneigentliche Linearkombination formulieren. Die Funktion f zerlegt man zu

$$f = \int_{-\infty}^{\infty} f(a) \delta_a da = \int_{-\infty}^{\infty} \langle \delta_a, f \rangle \delta_a da.$$

Wir können nun $f = |f\rangle$ aus dem Integral ausklammern und setzen

$\delta_a = |\delta_a\rangle$. Übrig bleibt der Identitätsoperator

$$I = \int_{-\infty}^{\infty} |\delta_a\rangle \langle \delta_a| da.$$

Dieser Operator verändert eine Funktion nicht. Es ist $I f = f$.

Um Schreibaufwand zu sparen, wollen wir das Intervall $(-\infty, \infty)$ mit Ω bezeichnen. Das Skalarprodukt kann man auch auf folgende Art ausrechnen:

$$\begin{aligned} \langle f, g \rangle &= \langle f | \int_{\Omega} g(a) |\delta_a\rangle da = \int_{\Omega} g(a) \langle f, \delta_a \rangle da \\ &= \int_{\Omega} g(a) \overline{\langle \delta_a, f \rangle} da = \int_{\Omega} g(a) \overline{f(a)} da. \end{aligned}$$

Fast die gleiche Rechnung macht man, wenn man bei $\langle f, g \rangle = \langle f | I | g \rangle$ für den Identitätsoperator einsetzt.

Da die inverse Fourier-Transformation unitär ist, kann man sie auf die Basisvektoren δ_a anwenden, um neue Basisvektoren zu erhalten. Es ergibt sich

$$e_a(x) = F^{-1}(\delta_a)(x) = \frac{1}{\sqrt{2\pi}} e^{iax}.$$

Tatsächlich ist nun

$$\int_{-\infty}^{\infty} e^{-iax} e^{ibx} dx = 2\pi \delta(a-b).$$

Man hat also $\langle e_a, e_b \rangle = \delta(a-b)$, d.h. die Funktionen e_a sind orthonormal.

Bei diagonalisierbaren Matrizen ist es möglich eine Eigenzerlegung der Matrix vorzunehmen. Im Folgenden wird erläutert, wie das beim Ableitungsoperator gemacht wird.

Der Differentialoperator $D(f) = x \mapsto f'(x)$ hat die Eigenwerte λ . Zum Eigenwert λ gehört der Eigenvektor $e^{\lambda x}$. Der Operator D muss nun zerlegt werden zu

$$D = T \circ S \circ T^{-1}.$$

Wenn man einen Basisvektor einer Orthonormalbasis mit einer Diagonalmatrix multipliziert, so wird der Basisvektor mit der entsprechenden Komponente der Matrix multipliziert. Es ergibt sich

$$\text{diag}(a_{11}, \dots, a_{nn}) e_k = a_{kk} e_k.$$

Dieses Prinzip übertragen wir analog auf den diagonalen Operator S . Man hat dann $S(\delta_r) = \lambda_r \delta_r$. Wir ordnen die Eigenwerte nun in der Reihenfolge an, bei der $\lambda_r = -r$ ist. Wenn man S nun auf eine Funktion anwendet, so erhält man

$$S(f) = x \mapsto (-x)f(x).$$

Für die Transformation T muss dann $T\delta_r = x \mapsto e^{-rx}$ sein. Man rechnet nun

$$\begin{aligned} Tf &= T \int_{-\infty}^{\infty} f(a) \delta_a da = \int_{-\infty}^{\infty} f(a) T\delta_a da \\ &= \int_{-\infty}^{\infty} f(a) (x \mapsto e^{-ax}) da \\ &= \int_{-\infty}^{\infty} f(a) \int_{-\infty}^{\infty} e^{-ab} \delta_b db da \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a) e^{-ab} da \delta_b db \\ &= x \mapsto \int_{-\infty}^{\infty} f(a) e^{-ax} da. \end{aligned}$$

Den Definitionsbereich von f schränkt man nun zu $x \geq 0$ ein. Für $x < 0$ kann man ja einfach $f(x) = 0$ definieren. Unter dieser Voraussetzung erhält man $T = L$ wobei mit L die Laplace-Transformation gemeint ist.

Z.B. soll $F(s) = 1/s$ abgeleitet werden. Man erhält

$$\begin{aligned} D \left\{ \frac{1}{s} \right\} &= L S L^{-1} \left\{ \frac{1}{s} \right\} = L S \{1\} \\ &= L \{-x\} = -L\{x\} = -\frac{1}{s^2}. \end{aligned}$$

Wenn man die Eigenwerte stattdessen in der Reihenfolge anordnet, dass $\lambda_r = r$ ist, ergibt sich $T\delta_r = x \mapsto e^{rx}$. In diesem Fall ist $T = L^{-1}$. Sei $R(f) = x \mapsto xf(x)$. Die Zerlegung lautet dann $D = L^{-1} R L$. Nun ist ja

$$D^2 = D D = L^{-1} R L L^{-1} R L = L^{-1} R^2 L.$$

Analog zu Matrizen müsste das auch für beliebige Potenzen möglich sein. Man rechnet nun

$$\begin{aligned} D^{-a} \{f(x)\} &= L^{-1} R^{-a} L \{f(x)\} \\ &= L^{-1} \{s^{-a} L \{f(x)\}\}. \end{aligned}$$

Nach den Rechenregeln der Laplace-Transformation ist aber

$$s^{-a} = \frac{1}{\Gamma(a)} L\{x^{a-1}\}.$$

Mit der Rechenregel $L(f)L(g) = L(f * g)$ ergibt sich

$$D^{-a}\{f(x)\} = \frac{1}{\Gamma(a)} L^{-1}\{L\{\int_0^x (x-t)^{a-1} f(t) dt\}\}.$$

Somit ist

$$D^{-a}\{f(x)\} = \frac{1}{\Gamma(a)} \int_0^x (x-t)^{a-1} f(t) dt.$$

Dieser Operator heißt Riemann-Liouville-Integral und stellt einen fraktionalen Integraloperator dar.

4.5 Integraltransformationen

Sei T eine Integraltransformation. Mit $T(f)$ oder auch Tf ist die Applikation von T auf die Funktion f gemeint. Eine Integraltransformation hat die Form

$$Tf = x \mapsto \int_{\Omega} K(a, x) f(a) da.$$

Die Funktion $K(a, x)$ wird Integralkern genannt. Die Transformation kann man als uneigentliche Linearkombination darstellen. Man erhält dann

$$Tf = \int_D \int_{\Omega} K(a, b) |\delta_b\rangle f(a) dadb.$$

Ersetzt man noch $f(a)$ gegen $\langle \delta_a, f \rangle$, so erhält man

$$Tf = \int_D \int_{\Omega} K(a, b) |\delta_b\rangle \langle \delta_a, f \rangle dadb.$$

Man klammert jetzt $f = |f\rangle$ aus. Der Integraloperator bleibt übrig. Es ergibt sich

$$T = \int_D \int_{\Omega} K(a, b) |\delta_b\rangle \langle \delta_a| dadb.$$

Man kann die Transformation auf die Delta-Distribution anwenden. Es ist

$$T\delta_r = x \mapsto K(r, x).$$

4.6 Das Stieltjes-Integral

Sei g eine Funktion mit stetiger Ableitung. Dann ist

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx.$$

Die Funktion g wird Integrator genannt. Beim Stieltjes-Integral sind aber auch Integratoren möglich, die z.B. Treppenfunktionen sind. Sei g eine Treppenfunktion, die im Intervall (a, b) an den Stellen x_k Sprünge der Höhe c_k macht. Wenn c_k negativ ist, dann macht g

an der Stelle x_k einen Sprung nach unten. Das Stieltjes-Integral ist

$$\int_a^b f(x) dg(x) = \sum_{k=1}^n f(x_k) c_k.$$

Die verallgemeinerte Regel zur partiellen Integration ist

$$\int_a^b f(x) dg(x) = [f(x)g(x)]_a^b - \int_a^b g(x) df(x).$$

Das Stieltjes-Integral ist linear im Integranden und im Integrator. Es ist

$$\begin{aligned} d(g(x) + h(x)) &= dg(x) + dh(x), \\ d(cg(x)) &= c dg(x). \end{aligned}$$

Damit kann man Treppenfunktionen mit stetig differenzierbaren Funktionen überlagern, wobei verschiedene Regelfunktionen entstehen. Wenn der Integrator Knickstellen hat, die Funktion f aber stetig differenzierbar ist, so kann man das Stieltjes-Integral über partielle Integration auf ein Riemann-Integral zurückführen.

Sei Ω ein Intervall, das die Stelle null enthält, wobei null nicht auf dem Rand von Ω liegen soll. Mit der Heaviside-Funktion $H(x)$ ist

$$\int_{\Omega} f(x) dH(x) = f(0).$$

Man kann $\delta(x) dx$ also als $dH(x)$ interpretieren. Wenn die Heaviside-Funktion eine Ableitung hätte, dann wäre diese die Delta-Distribution. Es ist nun ja

$$H(x) = \frac{1}{2}(\operatorname{sgn} x + 1).$$

Damit ergibt sich $dH(x) = (1/2)d\operatorname{sgn}(x)$. Außerdem beachtet man die Rechenregel

$$\operatorname{sgn}(ax) = \operatorname{sgn}(a) \operatorname{sgn}(x).$$

Man erhält

$$dH(ax) = \operatorname{sgn}(a) dH(x).$$

Es gilt auch die Rechenregel

$$|a| = \operatorname{sgn}(a)a = \frac{a}{\operatorname{sgn}(a)}.$$

Verwendet man nun noch $d(ax) = a dx$, so erhält man

$$\delta(ax) = \frac{1}{|a|} \delta(x).$$

4.7 Fraktionale Iterationen

Sei f eine Funktion, und man setze $D(f) = \operatorname{Bild}(f)$. Es lassen sich nun die Iterationen rekursiv definieren. Sei

$$\begin{aligned} f^0(x) &:= \operatorname{id}(x) = x, \\ f^n(x) &:= f^{n-1}(f(x)). \end{aligned}$$

Die Iterationen sind Potenzen bezüglich des Verkettungsoperators. Es ist $f^2 = f \circ f$ usw. Man kann sich nun die Frage stellen, ob sich so etwas wie $f^{1/2}$ bestimmen lässt. Allgemeiner f^c , wobei c eine komplexe Zahl ist.

Sei $C_f(g) := g \circ f$. Man rechnet nun

$$C_f(rg) = (rg) \circ f = x \mapsto rg(f(x)) = r(g \circ f)$$

und

$$\begin{aligned} C_f(g_1 + g_2) &= (g_1 + g_2) \circ f \\ &= x \mapsto g_1(f(x)) + g_2(f(x)) = g_1 \circ f + g_2 \circ f \\ &= C_f(g_1) + C_f(g_2). \end{aligned}$$

Bei C_f handelt es sich um einen linearen Operator, und $C_f^n(g)$ ist $g \circ f^n$, da Verkettungen assoziativ sind. Der Operator C_f wird *Kompositionoperator* genannt.

Sei nun Ψ ein Eigenvektor von C_f . Das zugehörige Eigenwertproblem $C_f\Psi = \lambda\Psi$ wird *Schrödersche Gleichung* genannt und stellt eine Funktionalgleichung dar. Hat man Ψ gefunden, so ergibt sich damit auch

$$f^r(x) = \Psi^{-1}(\lambda^r \Psi(x)).$$

Sei nun u ein Fixpunkt von f und sei $0 < |f'(u)| < 1$. Unter bestimmten Umständen ist dann

$$\Psi(x) = \lim_{n \rightarrow \infty} \frac{f^n(x) - u}{f'(u)^n}$$

und $\lambda = f'(u)$. Die Eigenfunktion $\Psi(x)$ wird dann *Funktion von Koenigs* genannt.

Jede gewöhnliche Potenz kann als Iteration formuliert werden. Sei z.B. $f(x) = ax$, dann ist $f^2(x) = a(a(x)) = a^2x$, da die Multiplikation assoziativ ist. Zu jeder Funktion f gehört nun genau eine Zahl a , womit sich eine Bijektion ergibt. Die linearen Funktionen und die reellen Zahlen sind also isomorph bezüglich der „Potenzierung“.

Für eine reelle Zahl r wäre dann aber

$$a^r x = f^r(x).$$

Diese Vorgehensweise hängt offensichtlich nur davon ab, dass die Multiplikation assoziativ ist. Man kann auch $f(x) = Dx$ setzen, wobei D der Ableitungsoperator und $x(t)$ eine Funktion ist. Somit können fraktionale Ableitungen als Spezialfall von fraktionalen Iterationen betrachtet werden.

Wenn u ein Fixpunkt von f ist, dann ist $f(u) = u$ und somit auch $f^n(u) = u$. Damit erscheint es vernünftig, auch $f^r(u) = u$ für eine reelle Zahl r zu verlangen.

In der Umgebung von u sollte sich f nun durch eine affine Funktion $g(x) = b + ax$ approximieren lassen. Man hat nun

$$\begin{aligned} g^2(x) &= b + ab + a^2x, \\ g^3(x) &= b + ab + a^2b + a^3x, \\ g^4(x) &= b + ab + a^2b + a^3b + a^4x \end{aligned}$$

und daher

$$g^n(x) = \left[\sum_{k=0}^{n-1} a^k \right] b + a^n x = \frac{a^n - 1}{a - 1} b + a^n x.$$

Jetzt lässt sich wieder $g^r(x)$ für eine reelle Zahl r berechnen. Für $r = -1$ ergibt sich die Umkehrfunktion von g , womit die hier dargestellte Vorgehensweise eine gewisse Sinnhaftigkeit erhält. Setzt man nun $g(x) = u + a(x - u)$ bzw. $b = u - au$, so ergibt sich die äquivalente, aber wesentlich einfachere Formel

$$g^n(x) = u + a^n(x - u).$$

Allgemeiner kann man f durch eine Taylorreihe approximieren. Sei $a := f'(u)$, $b := f''(u)$. Man erhält weiter

$$g^n(x) = u + a^n(x - u) + \frac{a^n - 1}{a - 1} a^{n-1} \frac{b}{2} (x - u)^2 + \dots$$

Die nächsten Summanden werden wesentlich komplizierter sein. Ohne System wird man händisch noch ein bis zwei Summanden formulieren können. Eine numerische Berechnung beliebiger Genauigkeit ist damit nicht möglich.

Als nächstes soll eine weitere Methode vorgestellt werden. Die Funktion f soll sich nun als Potenzreihe mit $f(x) = \sum_{k=0}^{\infty} a_k x^k$ darstellen lassen. Die Potenzreihe ist aber als Skalarprodukt $\langle A_1, v(x) \rangle$ mit $A_1 = (a_0, a_1, \dots)$ und $v(x) = (x^0, x^1, \dots)$ interpretierbar. Die Idee ist nun, auch die Potenzen $f(x)^i$ damit auszudrücken. Dann muss $A_0 = (1, 0, 0, \dots)$ sein. Sagen wir, die A_i sollen Spaltenvektoren sein, so lassen sie sich zur Matrix A zusammenfassen. Diese Matrix ist also von folgender Gestalt.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ a_0 & a_1 & a_2 & a_3 & \dots \\ a_0^2 & 2a_0a_1 & a_1^2 + 2a_0a_2 & 2a_1a_2 + 2a_0a_3 & \dots \\ a_0^3 & 3a_0^2a_1 & 3a_0a_1^2 + 3a_0^2a_2 & \dots & \dots \\ a_0^4 & 4a_0^3a_1 & 6a_0^2a_1^2 + 4a_0^3a_2 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Allgemein ist

$$A_{jk} = \frac{1}{k!} [D^k f(x)^j]_{x=0}.$$

Umgekehrt hat man mit dieser Matrix

$$f(x)^j = \sum_{k=0}^{\infty} A_{kj} x^k.$$

Mit dieser Matrix ergibt sich nun

$$Av(x) = A[1, x, x^2, \dots] = [1, f(x), f(x)^2, \dots].$$

Also kurz $Av(x) = v(f(x))$. Damit ist aber

$$A^n v(x) = v(f^n(x)).$$

Diese Gleichung stellt eine Beziehung zwischen der Matrixpotenz und der Iteration von f dar. Die Matrix $A(f)$ wird *Carleman-Matrix* der Funktion f genannt. Anders ausgedrückt erhält man die Formeln $A(f \circ g) = A(f)A(g)$ und $A(f^n) = A(f)^n$. Wegen $f(x) = [A(f)]_1 v(x)$ erhält man

$$f^n(x) = [A(f)^n]_1 v(x).$$

Um eine Approximation zu erhalten, beschränkt man sich auf eine quadratische Teilmatrix von A . Möchte man für n eine reelle Zahl einsetzen, so muss man nur die üblichen Methoden zur Berechnung von Matrixpotenzen mit reellem Exponenten benutzen.

Unter Verwendung der binomischen Reihe und der binomischen Formel ergibt sich

$$\begin{aligned} x^m &= (x - 1 + 1)^m = \sum_{n=0}^{\infty} \binom{m}{n} (x - 1)^n \\ &= \sum_{m=0}^{\infty} \binom{m}{n} \sum_{n=0}^k \binom{n}{k} (-1)^{n-k} x^k. \end{aligned}$$

Und für Iterationen erhält man analog

$$f^m(x) = \sum_{m=0}^{\infty} \binom{m}{n} \sum_{n=0}^k \binom{n}{k} (-1)^{n-k} f^k(x).$$

Um der Frage auf den Grund zu gehen, unter welchen Voraussetzungen die Formeln konvergente Lösungen liefern und wann die Ergebnisse übereinstimmen, benötigt man Kenntnisse in Funktionentheorie und Funktionalanalysis.

5 Stochastik

5.1 Ereignisse

Wenn man einen Würfel wirft, so erhält man eine der Zahlen aus

$$\{1, 2, 3, 4, 5, 6\}$$

als Ergebnis. Man kann jetzt jeder Zahl die Wahrscheinlichkeit $1/6$ zuordnen. Diese Vorgehensweise ist aber von einer gewissen Mangelhaftigkeit betroffen, weil man so nicht alles elegant formulieren kann, was denkbar ist. Es ist viel nützlicher, zuerst sogenannte Ereignisse aus den Ergebnissen zu bilden. Zu den Ergebnissen gehören die elementaren Ereignisse

$$\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}.$$

Jedes elementare Ereignis A hat nun die Wahrscheinlichkeit $P(A) = 1/6$. Wenn man sich jetzt nur dafür interessiert, ob eine Eins oder eine Zwei gewürfelt wurde, so kann man die Vereinigungsmenge dieser Ereignisse bilden. Man erhält das Ereignis

$$A = \{1\} \cup \{2\} = \{1, 2\}.$$

Man sagt, ein Ereignis A ist eingetreten, wenn eines der Elemente aus A das Ergebnis des Würfelwurfes ist. Für zwei disjunkte Ereignisse gilt nun die Formel

$$P(A \cup B) = P(A) + P(B).$$

Damit erhält man in diesem Fall

$$\begin{aligned} P(\{1, 2\}) &= P(\{1\} \cup \{2\}) = P(\{1\}) + P(\{2\}) \\ &= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

Ein Zufallsexperiment, bei dem jedem elementaren Ereignis die gleiche Wahrscheinlichkeit zugeordnet wird, wird allgemein als Laplace-Experiment bezeichnet. Das klassische Würfeln und der Münzwurf sind Laplace-Experimente. Die Menge aller Ergebnisse soll mit dem Buchstaben Ω bezeichnet werden. Für Laplace-Experimente gilt allgemein die Formel

$$P(A) = \frac{|A|}{|\Omega|}.$$

Dabei ist $|A|$ die Anzahl der Elemente von A und $|\Omega|$ die Anzahl der Elemente von Ω . In Wirklichkeit gibt es eigentlich keine Laplace-Experimente. In der Realität weichen Wahrscheinlichkeitsverteilungen immer geringfügig von der Gleichverteilung ab.

Wenn man zwei Würfel auf einmal wirft, so kann man die Ergebnismenge aus den Tupeln (e_1, e_2) bilden, wobei e_1 das Ergebnis des ersten Würfels und e_2 das Ergebnis des zweiten Würfels ist. Die Anzahl der Tupel ist

$$|\Omega \times \Omega| = |\Omega| \times |\Omega| = 6 \times 6 = 36.$$

Jedes elementare Ereignis hat also die Wahrscheinlichkeit $1/36$.

Wenn sicher ist, dass das Ereignis B eintritt, so schreibt man $P(A|B)$ anstelle von $P(A)$. Man fragt also nach der Wahrscheinlichkeit für das Eintreten von A , wenn man das Eintreten von B voraussetzen kann. Das wird als bedingte Wahrscheinlichkeit bezeichnet. Es gilt die Formel

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Zwei Ereignisse heißen stochastisch unabhängig, wenn

$$P(A \cap B) = P(A)P(B)$$

gilt. Das ist der Fall, wenn das Eintreten A nichts mit dem Eintreten von B zu tun hat.

Wenn man das gleiche Zufallsexperiment mehrmals hintereinander durchführt oder verschiedene Zufallsexperimente nacheinander durchführt, dann spricht man von einem mehrstufigen Zufallsexperiment. Für mehrstufige Zufallsexperimente kann man Baumdiagramme als Hilfsmittel benutzen. Es gelten Pfadregeln. Man kann z.B. nach der Wahrscheinlichkeit fragen, mit einer Münze erst Zahl zu werfen und dann mit einem Würfel eine Sechs zu würfeln. Die Wahrscheinlichkeit dafür ist $1/12$. Ein mehrstufiges Zufallsexperiment ist also ein zusammengesetztes Zufallsexperiment, welches aus mehreren kleinen Zufallsexperimenten besteht, die nacheinander durchgeführt werden. Auch beim gleichzeitigen Werfen von mehreren Würfeln kann man sich vorstellen, dass diese Würfel in einer festen Reihenfolge nacheinander geworfen werden.

Ein Zufallsexperiment mit der Ergebnismenge $\{0, 1\}$ wird als Bernoulli-Versuch bezeichnet. Bei Bernoulli-Versuchen schreibt man kurz $p := P(\{1\})$. Damit ergibt sich auch $P(\{0\}) = 1 - p$. Wenn man einen Bernoulli-Versuch mehrmals hintereinander durchführt, so spricht man von einer Bernoulli-Kette. Eine Bernoulli-Kette kann als mehrstufiges Zufallsexperiment oder auch als ein sogenannter stochastischer Prozess interpretiert werden.

Da die Ereignisse der Ergebnismenge gewisse Axiome erfüllen, bezeichnet man die Ergebnismenge auch als Ereignisraum. Damit sind nicht die Axiome von Kolmogorow gemeint, die gelten für Wahrscheinlichkeiten und nicht für Ereignisse. Ereignisräume sollen mit dem Buchstaben Σ bezeichnet werden.

Ein Ereignisraum ist eine Teilmenge der Potenzmenge von Ω . Daher ist $2^{|\Omega|}$ die maximale Anzahl der Elemente in Σ . Wenn man jedem Ergebnis ein dazugehöriges elementares Ereignis gibt, dann wird diese Anzahl auch erreicht.

5.2 Zufallsgrößen

Es hat sich als fruchtbar erwiesen sogenannte Zufallsgrößen zu betrachten. Das sind Funktionen von der

Ergebnismenge in die reellen Zahlen. Eine diskrete Teilmenge der reellen Zahlen ist natürlich auch eine Ergebnismenge, aber man kommt auf sinnvolle Ideen, wenn man stattdessen eine Funktion

$$X: \Omega \rightarrow \mathbf{R}$$

betrachtet. Anstelle von $y = f(x)$ schreibt man hier $x = X(\omega)$. Wenn man schon eine diskrete Teilmenge der reellen Zahlen vorliegen hat und falls man den Begriff der Zufallsgröße nicht braucht, so kann man die identische Funktion $X = \text{id}$ als Zufallsgröße wählen.

Jetzt stellt sich natürlich die Frage, wie man die Wahrscheinlichkeit für das Eintreten von x angibt. Wahrscheinlichkeiten sind ja eigentlich nur für Ereignisse definiert, und genau so muss man auch Vorgehen. Man wählt eine Bildmenge aus, in diesem Fall $\{x\}$, und bestimmt die Urbildmenge. Die Urbildmenge ist nun aber ein Ereignis, und von diesem kann man die Wahrscheinlichkeit bestimmen. Man definiert also

$$P(\{x\}) := P(X^{-1}(\{x\})).$$

Allgemeiner definiert man

$$P(A \subseteq \text{Bild}(X)) := P(X^{-1}(A)).$$

Nun kann man auch die sogenannte Verteilungsfunktion definieren. Man definiert

$$F(x) := P(\{a \in \text{Bild}(X) | a \leq x\}).$$

Oft wird auch die Kurzschreibweise $F(x) = P(X \leq x)$ verwendet. Die Wahrscheinlichkeit einer Teilmenge der Bildmenge ist immer noch ein sogenanntes Maß. Damit gilt für disjunkte Teilmengen A, B der Bildmenge von X auch jetzt noch die Formel

$$P(A \cup B) = P(A) + P(B).$$

Damit ergibt sich

$$F(x) = P(\{a | a \leq x\}) = P\left(\bigcup_{a \leq x} \{a\}\right) = \sum_{a \leq x} P(\{a\}).$$

5.3 Stochastische Prozesse

Betrachten wir eine Menge von Zuständen s_k . Für jeden Zustand gibt es nun Pfeile $a_{ki} := (s_k, s_i)$ auf andere Zustände. Ein Pfeil der auf den aktuellen Zustand zurückführt, ist auch zugelassen. Jeder Pfeil wird nun mit einer Wahrscheinlichkeit p_{ki} behaftet, so dass die Summe aller Wahrscheinlichkeiten bezüglich des Zustandes s_k eins ist. Man beginnt nun bei einem Anfangszustand. Durch den Zufall wird ein Pfeil ausgewählt, so dass man bei einem neuen Zustand landet. Dann wird durch Zufall wieder ein Pfeil ausgewählt und man landet bei einem neuen Zustand usw. Ein solches Zufallsexperiment bezeichnet man als *Markov-Kette*. Auf ähnliche Art lassen sich auch kompliziertere Systeme, wie endliche Automaten oder zelluläre Automaten, mit Zufall behaften.

Eine solche Markov-Kette wird als Gedächtnislos bezeichnet, da die Wahrscheinlichkeiten für den Übergang in einen anderen Zustand nur vom aktuellen Zustand abhängig sind.

Ein Beispiel für eine Markov-Kette ist die Bernoulli-Kette. Der Zustandsraum ist hier $Z = \mathbb{N}_0$. Die Zustände $s_k = k$ sind die Anzahl der Erfolge. Jeder Zustand k hat zwei Pfeile, einen auf sich selbst, und den anderen auf $k + 1$. Die Wahrscheinlichkeit für den Übergang zu $k + 1$ ist für alle Zustände gleich und soll mit p bezeichnet werden. Die Wahrscheinlichkeit, im aktuellen Zustand zu verbleiben, ist damit für alle Zustände $p - 1$.

Eine einfache Erweiterung wäre nun z.B. die ganzen Zahlen als Zustände und Hinzufügung eines Abstieges von k zu $k - 1$. Man erhält die sogenannte Irrfahrt auf den ganzen Zahlen.

6 Funktionentheorie

6.1 Holomorphie

Ein Kriterium für die komplexe Differenzierbarkeit sind die Cauchy-Riemannschen Differentialgleichungen. Sei $x = \operatorname{Re}(z)$ und $y = \operatorname{Im}(z)$. Die komplexe Funktion $f(z)$ lässt sich in den Realteil u und den Imaginärteil v auftrennen. Es ist dann

$$f(z) = f(x + iy) = u(x, y) + iv(x, y).$$

Somit lässt sich f als Vektorfeld auffassen, welches jedem Tupel (x, y) ein Tupel (u, v) zuordnet. Die Funktion f ist holomorph, wenn sie als Vektorfeld total differenzierbar ist und wenn u, v die Cauchy-Riemannschen Differentialgleichungen erfüllen. Die Differentialgleichungen sind $D_x u = D_y v$ und $D_y u = -D_x v$.

Sei nun $x_1 = x$ und $x_2 = y$. Sei außerdem $u_1 = u$ und $u_2 = -v$. Die Differentialgleichungen lassen sich schreiben als

$$\begin{aligned} D_1 u_1 + D_2 u_2 &= 0, \\ D_1 u_2 - D_2 u_1 &= 0. \end{aligned}$$

Damit erhält man die Bedingungen $\langle \nabla, u \rangle = 0$ und $\nabla \wedge u = 0$. Bei u handelt es sich also um ein quellenfreies und rotationsfreies Vektorfeld. Mit

$$\nabla u = \langle \nabla, u \rangle + \nabla \wedge u$$

lassen sich die beiden Bedingungen zur Gleichung $\nabla u = 0$ zusammenfassen.

6.2 Integralsatz von Cauchy

Man stelle sich eine komplexwertige Funktion $f(t)$ vor. Wie berechnet man das bestimmte Integral einer solchen Funktion? Nun ja, eine komplexwertige Funktion lässt sich doch als vektorwertige Funktion interpretieren. Die Funktion $f(t)$ wird in Realteil und Imaginärteil aufgetrennt, sie wird also in der Form $f(t) = u(t) + iv(t)$ dargestellt. Da das Integral ein linearer Operator ist, lässt es sich vor die Bildung der Linearkombination ziehen. D.h. man rechnet

$$\begin{aligned} \int_{t_1}^{t_2} f(t) dt &= \int_{t_1}^{t_2} u(t) + iv(t) dt \\ &= \int_{t_1}^{t_2} u(t) dt + i \int_{t_1}^{t_2} v(t) dt. \end{aligned}$$

Jetzt stellt sich natürlich sofort die Frage, ob es auch möglich ist, für eine komplexe Funktion ein bestimmtes Integral zu berechnen. Das ist etwas komplizierter, man muss ein sogenanntes komplexes Kurvenintegral definieren. Dafür bekommt man jedoch einfache analoge Rechenregeln, ein sehr nützliches Werkzeug und höchst interessante Resultate. Die analogen Rechenregeln bekommt man deshalb, weil das Wegintegral unter einfachen Voraussetzungen wegunabhängig ist.

Sei nun γ ein Weg. Dieser soll jedoch nicht durch eine vektorwertige Funktion parametrisiert werden, sondern durch eine komplexwertige. Das Integral wird analog zum Kurvenintegral zweiter Art definiert durch

$$\int_{\gamma} f(z) dz := \int_{t_1}^{t_2} f(\gamma(t)) \gamma'(t) dt.$$

Der Unterschied ist, dass die Multiplikation beim Kurvenintegral zweiter Art ein Skalarprodukt ist, es sich hier aber um die Multiplikation komplexer Zahlen handelt. Man kann auch immer $t_1 = 0$ und $t_2 = 1$ wählen. Der Weg ist dann so parametrisiert, dass $\gamma(0)$ der Anfangspunkt ist und $\gamma(1)$ der Endpunkt. Das lässt sich natürlich durch triviale Uparametrisierung immer erreichen.

Man rechnet nun

$$dz = d(x + iy) = dx + i dy.$$

Außerdem rechnet man

$$\begin{aligned} f(z) dz &= (u + iv)(dx + i dy) \\ &= u dx - v dy + i(udx + vdy) \\ &= u_1 dx_1 + u_2 dx_2 + i(u_1 dx_2 - u_2 dx_1) \\ &= (u_1, u_2)^b + i(-u_2, u_1)^b. \end{aligned}$$

Man wendet jetzt den gaußschen Integralsatz in der Ebene rückwärts an und erhält

$$\begin{aligned} \int_{R(B)} f(z) dz \\ = \int_B \nabla^b \wedge (u_1, u_2)^b + i \int_B \nabla^b \wedge (-u_2, u_1)^b \end{aligned}$$

Umformen und ausnutzen der Quellenfreiheit und Rotationsfreiheit bringt

$$\begin{aligned} \nabla^b \wedge (u_1, u_2)^b &= (D_1 u_2 - D_2 u_1) dx_1 \wedge dx_2 = 0, \\ \nabla^b \wedge (-u_2, u_1)^b &= (D_1 u_1 + D_2 u_2) dx_1 \wedge dx_2 = 0. \end{aligned}$$

Damit erhält man schließlich

$$\oint_{\gamma} f(z) dz = 0.$$

Dieses Resultat wird als Integralsatz von Cauchy bezeichnet. Es ist wichtig, dass die Funktion f auf einem einfach zusammenhängenden Gebiet definiert ist, und dass die Kurve γ innerhalb dieses Gebietes liegt. Wenn das nicht der Fall ist, dann kann man den gaußschen Integralsatz nicht verwenden. In der Tat lassen sich dann auch Gegenbeispiele finden.

Man denke sich jetzt einen *nicht* geschlossenen Weg γ mit Anfangspunkt a und Endpunkt b . Außerdem denke man sich einen konstanten Weg β , der den Weg γ abschließt, d.h. die beiden Punkte verbindet. Den Weg γ kann man nun variieren. Nach dem Integralsatz wird aber immer

$$\int_{\beta} f(z) dz + \int_{\gamma} f(z) dz = K + \int_{\gamma} f(z) dz = 0$$

sein. Daher ist

$$\int_{\gamma} f(z) dz = \text{const.}$$

Das Wegintegral hängt nur von Anfangspunkt a und Endpunkt b ab, nicht jedoch von der Wahl des Weges. Es ist wegunabhängig. Daher kann man auch schreiben

$$\int_a^b f(z) dz := \int_{\gamma} f(z) dz,$$

wobei man sich einen Weg von a nach b aussuchen darf.

Die Wegunabhängigkeit von komplexen Wegintegralen für holomorphe Funktionen scheint analog zur Wegunabhängigkeit von Wegintegralen zweiter Art für Potentialfelder zu sein. Jedoch gibt es zwei wichtige Unterschiede. Zum einen sind komplexe Wegintegrale vektorwertig, wenn man die komplexen Zahlen als Vektoren interpretiert. Zum anderen benötigt man als Voraussetzung zusätzlich zur Rotationsfreiheit auch noch die Quellenfreiheit.

Es gibt also drei Arten von Wegintegralen. Eines mit skalarer Multiplikation, eines mit Skalarprodukt und eines mit komplexer Multiplikation.

Der Hauptsatz gilt auch im Komplexen, was äußerst nützlich ist. Sei F eine Stammfunktion von f , d.h. die komplexe Ableitung von F muss f sein. Dann gilt

$$\int_a^b f(z) dz = F(b) - F(a).$$

Außerdem ist eine Stammfunktion F gegeben durch

$$F(z) = \int_a^z f(z) dz.$$