

Mathematische Methoden 1

PHB.01106UB

Vorlesungsskriptum WS 2021/22

Assoz.-Prof. Dr. Peter Puschnig

Institut für Physik, Fachbereich Theoretische Physik

Karl-Franzens-Universität Graz

Universitätsplatz 5, A-8010 Graz

peter.puschnig@uni-graz.at

<https://homepage.uni-graz.at/de/peter.puschnig/>

Graz, 14. 10. 2021

Über dieses Skriptum

Dieses Vorlesungsskriptum ist aus handschriftlichen Unterlagen von Prof. Leopold Mathelitsch aus dem WS 2010/2011 entstanden, und lehnt sich in seiner Vorgehensweise weitgehend an das Buch "Mathematische Methoden in der Physik" von Christian B. Lang und Norbert Pucker an (siehe auch <https://homepage.uni-graz.at/de/christian.lang/lehre-teaching/my-books/>). An manchen Stellen wurde auch Anleihe an den Vorlesungsunterlagen von Prof. Ganster genommen, der diese Vorlesung im WS 2011/12 abgehalten hat: http://www.math.tugraz.at/~ganster/differenzialrechnung_ws_2011.html

Warum ein/e Physiker/in Mathematik benötigt

Wenn Sie diese Vorlesung besuchen, dann sind Sie höchstwahrscheinlich an *Physik* interessiert und stellen sich vielleicht die Frage, warum Sie in den ersten Semestern so viele *Mathematik*-Vorlesungen besuchen müssen. Fakt ist nunmal, dass sich Naturgesetze einfach am besten in der Sprache der Mathematik formulieren lassen, und wer physikalische Zusammenhänge verstehen will, muss auch die Sprache der Mathematik beherrschen. Wir bilden also Naturgesetze auf mathematische Gleichungen ab, lösen diese Gleichungen dann in der Welt der Mathematik, und übersetzen am Ende die gewonnen Ergebnisse wieder in die Welt der Physik zurück. Aus dieser Vorgehensweise ergibt sich auch ein wichtiger Unterschied im Zugang zum Fach Mathematik, wie er von MathematikerInnen und PhysikerInnen praktiziert wird. Während sich ein/e Mathematiker/in sich damit beschäftigt, mathematische Sätze zu formulieren und zu beweisen, die *Existenz* von Lösungen einer bestimmten Gleichung nachzuweisen, benutzt ein/e Physiker/in die Mathematik sehr oft einfach als Werkzeug, um die Lösung eines physikalischen Problems zu erhalten. Die Existenz einer Lösung allein reicht dem/der Physiker/in da nicht aus, er/sie will *die* Lösung der betreffenden Gleichung erhalten, die – wenn die mathematischen Symbole wieder als physikalische Größen interpretiert werden – auch in der Welt der Physik Sinn ergibt.

Dass Mathematik mehr ist als ein einfaches Werkzeug der Physik darstellt, ist unbestritten. Manchmal passieren in dem Wechselspiel zwischen Mathematik und Physik auch ganz erstaunliche Dinge. Beispielsweise bei der Entdeckung des Positrons, des Antiteilchens des Elektrons. Hier hatte Paul Adrien Dirac 1928 eine relativistische Formulierung der quantenmechanischen Wellenfunktion des Elektrons gefunden, die Dirac-Gleichung, die als Lösung auch Teilchen mit positiver Ladung zuließ. Dies führte zur *Vorhersage* von Positronen, die dann im Jahre 1932 tatsächlich von Carl Anderson experimentell bestätigt wurde. In einem anderen historischen Beispiel des 20. Jahrhunderts bediente sich Albert Einstein 1915 bei der Aufstellung seiner Allgemeinen Relativitätstheorie maßgeblich bei den *Vorarbeiten* des Mathematikers Bernhard Riemann, der sich Jahre davor schon mit der Geometrie von Nicht-Euklidischen Räumen beschäftigte. In vielen Fällen liefern aber auch Problemstellungen, die in der Physik auftauchen, wichtige *Impulse zur Weiterentwicklung* der Mathematik. Die Entwicklung der Differenzialrechnung durch Newton und Leibniz

wurde nicht zuletzt dadurch angetrieben, weil sich die Gesetze der klassischen Mechanik in Form von Differenzialgleichungen anschreiben lassen. Auch bei dem aktuellen Versuch Quantentheorie und Gravitation in einer gemeinsamen Theorie zu vereinheitlichen bringen physikalische Fragestellungen an der Schnittstelle zwischen Mathematik und Physik, also in der Disziplin *Mathematische Physik*, neue Konzepte in die Mathematik ein.

Aber zurück auf den Boden der Realität eines Physikstudiums. Da sich alle bekannten physikalischen Theorien in der Sprache der Mathematik elegant ausdrücken lassen, könnte man meinen, dass es wohl das beste wäre, die ersten paar Semester *ausschließlich* dem Erlernen mathematischer Konzepte zu widmen. Ausgestattet mit dem nötigen mathematischen Rüstzeug, könnte man dann in weiterer Folge physikalische Theorien wie die klassische Mechanik und Elektrodynamik, die Quantenmechanik, die Relativitätstheorie, Quantenfeldtheorien usw. behandeln. Der Nachteil einer solchen Vorgehensweise wäre natürlich, dass die Gefahr besteht, dass die Motivation, sich mathematische Fertigkeiten anzueignen, enden wollend ist, solange der/die angehende Physiker/in nicht erkennt, wozu denn die mathematischen Werkzeuge in der Physik zu gebrauchen sind. In allen gängigen Physikcurricula an Universitäten wird daher ein didaktischer Mittelweg beschritten. Die wichtigsten mathematischen Methoden werden am Anfang des Studiums vermittelt und parallel dazu werden physikalische Konzepte transportiert, zunächst aus dem Blickwinkel der experimentellen Physik und anschließend in einem Theoriezyklus aus dem Blickwinkel der theoretischen Physik mit einer stärkeren mathematischen Gewichtung.

Diese Vorlesung "*Mathematische Methoden 1*" bildet die Grundlage für die weitere Mathematikausbildung. *Folgen und Reihen* (Kapitel 1) werden Ihnen im Laufe Ihres Physikstudiums immer wieder begegnen, etwa bei der Lösung von Differenzialgleichungen. Auch *komplexe Zahlen* (Kapitel 2) spielen eine wesentliche Rolle bei vielen physikalischen Problemen oder helfen diese zu vereinfachen.¹ Die Schrödinger-Gleichung etwa ist eine wesentlich komplexe partielle Differenzialgleichung, wie es MathematikerInnen ausdrücken würden, und kann ohne Kenntnis komplexer Zahlen nicht angeschrieben werden. Im Kapitel 3 wird eine Einführung in die *Lineare Algebra* gegeben. Insbesondere widmen wir uns hier der Lösung von linearen Gleichungssystemen, und dem Umgang mit Vektoren und Matrizen. Im Kapitel 4 (*Differenzialrechnung*) werden wir die Differenzialrechnung, die Sie (hoffentlich) aus der Schule kennen, auf Funktionen mehrerer Veränderlicher erweitern, was die Grundlage bildet für die Behandlung der klassischen Mechanik sowie von klassischen Feldtheorien, also etwa Strömungsmechanik oder die Elektrodynamik. Als ausgebildeter theoretischer Physiker, werde ich versuchen, den Schwerpunkt auf die *Anwendung der erlernten mathematischen Konzepte* zu legen, während ich auf die strenge Beweisführung von mathematischen Sätzen in meiner Vorlesung großteils verzichten werde.

¹Auf den ersten Blick scheinen sich der Terminus "komplexe Zahl" und der Ausdruck "vereinfachen" zu widersprechen, aber es ist in der Tat so, dass sich viele Zusammenhänge mithilfe komplexer Zahlen einfacher ausdrücken lassen.

Inhaltsverzeichnis

1	Folgen und Reihen	1
1.1	Überblick	1
1.2	Metrische Räume	7
1.2.1	Intervalle	7
1.2.2	Norm	8
1.2.3	Metrische Räume	8
1.3	Folgen	10
1.4	Reihen	14
1.4.1	Die vollständige Induktion	14
1.4.2	Konvergenz von Reihen	16
1.4.3	Elementare Reihen	17
1.4.4	Konvergenztests	21
1.5	Potenzreihen	27
1.5.1	Konvergenzradius	27
1.5.2	Taylor-Reihe	29
1.5.3	Eigenschaften von Potenzreihen	31
1.5.4	Beispiele für Taylor-Reihen	32
1.5.5	Anwendungsbeispiele	33
1.5.6	Regel von De L'Hospital	36
2	Komplexe Zahlen	39
2.1	Die komplexe Ebene	39
2.2	Komplexe Reihen	43
2.3	Funktionen komplexer Variablen	44
2.3.1	Exponentialfunktion und trigonometrische Funktionen	44
2.3.2	Wurzeln	49
2.3.3	Logarithmus	50
2.3.4	Überblick über einige elementare komplexwertige Funktionen	51
2.3.5	Umkehrfunktionen der trigonometrischen Funktionen	53

3	Lineare Algebra	55
3.1	Lineare Gleichungssysteme	55
3.1.1	Determinanten	57
3.1.2	Lösung eines linearen Gleichungssystems	61
3.2	Vektoren und ihre Algebra	64
3.2.1	Schreibweisen und Definitionen	64
3.2.2	Vektoralgebra	65
3.2.3	Das Skalarprodukt	66
3.2.4	Darstellung von Vektoren	67
3.2.5	Das Vektorprodukt	69
3.2.6	Analytische Geometrie	70
3.3	Matrizen	72
3.3.1	Lineare Algebra der Matrizen	72
3.3.2	Die inverse Matrix	74
3.3.3	Lineare Abhängigkeit	77
3.3.4	Rang einer Matrix	78
3.4	Das Eigenwertproblem	81
4	Differenzialrechnung	87
4.1	Die lineare Näherung	87
4.1.1	Differenzenquotient und Differenzialquotient	87
4.1.2	Totales Differenzial	91
4.1.3	Mittelwertsatz der Differenzialrechnung	93
4.1.4	Ableitung der Umkehrfunktion	94
4.2	Funktionen mehrerer Variablen	95
4.2.1	Motivation	95
4.2.2	Partielle Ableitung	96
4.2.3	Variablentransformation	99
4.2.4	Taylor-Reihe für mehrdimensionale Funktionen	101
4.2.5	Totales Differenzial	103
4.2.6	Richtungsableitung und Gradient	105
4.2.7	Jacobi-Matrix	110
4.3	Methoden der Differenziation	112
4.3.1	Produktregel, Quotientenregel, und Kettenregel	113
4.3.2	Implizite Differenziation	118
4.4	Extremwertwertberechnungen	120
4.4.1	Funktionen einer Veränderlichen	121

4.4.2	Funktionen mehrerer Veränderlicher	123
4.5	Extremwertwertberechnungen mit Nebenbedingungen	127
4.5.1	Elimination	128
4.5.2	Lagrangesche Multiplikatoren	129

CDF Demonstrationen

1	Euler-Zahl	12
2	Rekursive Folge für π	13
3	Elementare Reihen	19
4	Geometrische Reihe	20
5	Integraltest	26
6	Taylorreihen	33
7	Komplexe Addition	40
8	Komplexe Multiplikation	42
9	Euler-Gleichung	45
10	Komplexe Reihen	48
11	Komplexe Wurzeln	50
12	Komplexe Funktionen	52
13	Mandelbrotfolge	53
14	Regel von Sarrus	58
15	Entwicklung von Determinanten	60
16	Lösung eines 3x3 Gleichungssystems	62
17	Illustration der Matrixmultiplikation	73
18	Rotation in der Ebene mit Matrizen	76
19	Eigenschaften einer 3x3 Matrix	79
20	Eigenvektoren in 2D	82
21	Differenzialquotient	88
22	1. und 2. Ableitung	90
23	Partielle Ableitung	97
24	Totales Differenzial	104
25	Richtungsableitung	109
26	Gradient	109

27 Gradient – Teil 2	110
28 Extrema in 2D	124
29 Lagrange Multiplikator	132

Kapitel 1

Folgen und Reihen

1.1 Überblick

In diesem Einleitungsteil werden wir Folgen und Reihen kurz definieren, damit wir wissen, womit wir uns beschäftigen werden. Anschließend werden wir anhand einiger Beispiele versuchen zu zeigen, wofür Folgen und Reihen nützlich sind.

Eine Folge ist definiert als geordnete Menge von Elementen. Was das heißt, können Sie an folgenden zwei einfachen Beispielen sehen:

$$\begin{array}{ccccc} & & b & & \\ a & e & & d & \\ & c & f & & \end{array} \qquad c, e, a, b, f, d$$

Intuitiv ist die linke Menge nicht geordnet, während die rechte Menge einen höheren Ordnungsgrad aufweist. Natürlich können wir auch für die rechte Menge einen noch höheren Ordnungsgrad herstellen, zum Beispiel indem wir sie alphabetisch ordnen. Das wesentliche an der rechten Menge ist allerdings, dass die *Reihenfolge* eindeutig gegeben ist. Das heißt, die Elemente haben eindeutig einen Vorgänger bzw. Nachfolger, und daher können wir die Elemente durchnummerieren. Dieses Durchnummerieren ist das Essentielle einer **Folge** und kann mathematisch folgendermaßen definiert werden.

Eine Folge ist eine Abbildung (oder Zuordnung) einer Teilmenge N der natürlichen Zahlen \mathbb{N} auf eine Menge M :

$$N \longrightarrow M, \text{ im obigen Beispiel: } \begin{array}{cccccc} c & e & a & b & f & d \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 1 & 2 & 3 & 4 & 5 & 6 \end{array} \qquad (1.1)$$

Ist die Menge N gleich den gesamten natürlichen Zahlen, so handelt es sich um eine unendliche Folge, wenn $N \subset \mathbb{N}$ dann nennt man diese Abbildung eine endliche Folge. Man schreibt für eine Folge $\{a_n\}$

oder genauer $\{a_n\}_{n \in \mathbb{N}}$, wobei a_n das n -te Glied der Folge ist. In unserem obigen Beispiel lauten die Folgenglieder also $a_1 = c$, $a_2 = e$, $a_3 = a$, $a_4 = b$, $a_5 = f$, und $a_6 = d$.

Es gibt nun mehrere Möglichkeiten, eine Folge anzugeben. Als erstes kann man die Elemente *explizit* anführen

$$a_n = 1, \frac{1}{2}, \frac{1}{3}, \dots$$

Eine explizite Angabe der Folgenglieder ist nicht nur bei endlichen Folgen möglich wie im obigen Beispiel, sondern auch bei unendlichen Folgen:

$$a_n = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

wobei die Punkte "...“ die Unendlichkeit der Folge andeuten. Eine weitere Möglichkeit Folgen zu definieren, besteht in der Angabe eines *allgemeinen Folgenglieds* a_n der Folge, also

$$a_n = \frac{1}{n}.$$

Schließlich besteht auch die Möglichkeit, eine Folge *rekursiv* anzugeben, d.h. das n -te Element der Folge wird aus dem vorigen (oder aus mehreren vorigen) Folgengliedern berechnet, also

$$a_{n+1} = a_n - \frac{1}{n(n+1)}.$$

Hierbei müssen wir natürlich auch das Startglied definieren, um die Folge eindeutig festzulegen, also etwa $a_1 = 1$. Wie wir leicht überprüfen können, führt auch dieses Beispiel einer rekursiv definierten Folge auf die gleiche Folge $a_n = \frac{1}{n}$. Andere Beispiele für Folgen sind etwa:

$$\begin{aligned} a_n &= 7 \\ b_n &= (-1)^n \\ c_n &= 1, \frac{1}{2}, 2, \frac{1}{3}, 1, \frac{1}{4}, \dots \\ a_n &= \left(1 + \frac{1}{n}\right)^n \\ a_{n+1} &= 2.5a_n(1 - a_n), a_0 = 0.5 \end{aligned}$$

Zusammenfassend können wir also nochmals festhalten, dass eine *Folge eine geordnete, linearisierte Menge von Elementen ist, deren Ordnung durch Durchnummerieren zustande kommt*.

Was verstehen wir nun unter einer **Reihe**? Wir können aus jeder Folge eine weitere, spezielle Folge

bilden, nämlich durch teilweises Aufsummieren der Glieder der ursprünglichen Folge a_n , also

$$\begin{aligned} a_1 &\longrightarrow s_1 = a_1 \\ a_2 &\longrightarrow s_2 = a_1 + a_2 \\ a_3 &\longrightarrow s_3 = a_1 + a_2 + a_3 \\ &\vdots \\ a_n &\longrightarrow s_n = \sum_{i=1}^n a_i \end{aligned}$$

Die Summe

$$S = a_1 + a_2 + a_3 + \cdots = \sum_{i=1}^{\infty} a_i$$

nennt man *unendliche Reihe*, und

$$S_n = a_1 + a_2 + a_3 + \cdots + a_n = \sum_{i=1}^n a_i$$

nennt man *Teil- oder Partialsummen*. Eine unendliche Reihe können wir natürlich nicht direkt aufsummieren, weil das ja auch unendlich lange dauern würde. Wir können allerdings über die Partialsummen die Summe der unendlichen Reihe bestimmen. Existiert nämlich der *Grenzwert (limes)* der Partialsummen

$$\lim_{n \rightarrow \infty} S_n = S,$$

so nennt man S die Summe der unendlichen Reihe, und die Reihe heißt dann *konvergent*. Existiert kein Grenzwert, das heißt der Limes von S_n strebt gegen keinen endlichen Wert, so nennt man die Reihe *divergent*.

Bilden wir beispielsweise aus der Folge $a_n = \frac{1}{n}$, die wir oben kennengelernt haben, eine Reihe, dann erhalten wir die sogenannte *harmonische Reihe*. Wie wir gleich zeigen werden, ist die harmonische Reihe divergent, was beim ersten Hinsehen vielleicht etwas unerwartet scheint. Im Kapitel 1.4 werden wir genaue Algorithmen kennenlernen, mit deren Hilfe die Konvergenz von Reihen überprüft werden kann. An dieser Stelle folgt zunächst ein kurzer Beweis für die Divergenz der harmonischen Reihe:

$$S = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{=\frac{7}{12} > \frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7}}_{=\frac{107}{210} > \frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12}}_{=\frac{2021}{3960} > \frac{1}{2}} + \cdots$$

Aus dieser Darstellung ersehen wir, dass es uns immer gelingen wird so viele Glieder zusammenzufas-

sen, so dass deren Teilsumme größer als $\frac{1}{2}$ sein wird. Damit gilt für die Summe

$$S > k \frac{1}{2},$$

wobei $k \rightarrow \infty$ und damit auch $S \rightarrow \infty$ folgt. Gegen diese "Beweisführung" könnte man argumentieren, dass es ab einem gewissen n nicht mehr möglich ist, Reihenglieder so zusammenzufassen, dass deren Summe größer als $\frac{1}{2}$ ist. Dass dies aber tatsächlich für beliebig große n immer möglich ist, sieht man bei folgender Aufteilung der Summe:

$$S = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{> 2\frac{1}{4} = \frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{> 4\frac{1}{8} = \frac{1}{2}} + \underbrace{\frac{1}{9} + \dots + \frac{1}{16}}_{> 8\frac{1}{16} = \frac{1}{2}} + \underbrace{\frac{1}{17} + \dots + \frac{1}{32}}_{> 16\frac{1}{32} = \frac{1}{2}} + \dots$$

wodurch wir die Divergenz der harmonischen Reihe rigoros gezeigt haben.

Ein Beispiel für eine konvergente Reihe ist etwa

$$S = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$$

Wie wir leicht überprüfen können, lauten hier die Partialsummen

$$\begin{aligned} S_1 &= \frac{1}{2} = 1 - \left(\frac{1}{2}\right)^1 \\ S_2 &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} = 1 - \left(\frac{1}{2}\right)^2 \\ S_3 &= \frac{3}{4} + \frac{1}{8} = \frac{7}{8} = 1 - \left(\frac{1}{2}\right)^3 \\ &\vdots \\ S_n &\longrightarrow 1 - \left(\frac{1}{2}\right)^n \end{aligned}$$

Damit können wir den Grenzwert der Partialsummen s_n , das heißt die Summe der Reihe, einfach bestimmen

$$S = \lim_{n \rightarrow \infty} \left[1 - \left(\frac{1}{2}\right)^n \right] = 1.$$

Andere Beispiele für unendliche Reihen, begegnen uns bei bestimmten Dezimalzahlen:

$$\frac{1}{3} = \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \dots = \sum_{i=1}^{\infty} \frac{3}{10^i}$$

Bevor wir in den nächsten Kapiteln auf Folgen (1.3) und Reihen (1.4) genauer eingehen, wollen wir uns zuvor einige konkrete Beispiele anschauen, anhand derer man den praktischen Nutzen von Folgen und Reihen bereits erahnen kann. Jeder/e hat schon einmal mit einem Taschenrechner gearbeitet, aber haben Sie sich schon einmal die Frage gestellt, wie der Rechner etwa die Quadratwurzel einer Zahl $\sqrt{2}$ oder den Wert einer Winkelfunktion berechnet? Betrachten wir die rekursiv definierte Folge

$$a_0 = 1, \quad a_{n+1} = \frac{1}{2} \left(a_n + \frac{x}{a_n} \right)$$

Man kann zeigen, dass diese Folge als Grenzwert die Zahl \sqrt{x} besitzt. Wir schreiben für $x = 2$ einige Folgenglieder an

n	a_n	$ a_n - \sqrt{2} $
0	$1 = 1.00000000$	0.41421356
1	$\frac{1}{2} \left(1 + \frac{2}{1} \right) = \frac{3}{2} = 1.50000000$	0.08578644
2	$\frac{1}{2} \left(\frac{3}{2} + \frac{2}{\frac{3}{2}} \right) = \frac{17}{12} = 1.41666667$	0.00245310
3	$\frac{1}{2} \left(\frac{17}{12} + \frac{2}{\frac{17}{12}} \right) = \frac{577}{408} = 1.41421569$	0.00000212
4	$\frac{1}{2} \left(\frac{577}{408} + \frac{2}{\frac{577}{408}} \right) = \frac{665857}{470832} = 1.41421356$	0.00000000

Wie wir sehen konvergiert diese Folge tatsächlich schon nach einigen Folgengliedern sehr gut gegen den numerischen Wert $\sqrt{2}$. Dieses Verfahren, die Quadratwurzel einer Zahl zu berechnen wird auch *Babylonisches Wurzelziehen* genannt, da es bereits den Bewohnern des antiken Babylon bekannt gewesen sein dürfte.

Als ein Beispiel für die Anwendung von Reihen nehmen wir an, wir wollen einen numerischen Wert für $\sin(3\pi/8) = 0.923879532511$. Folgende Optionen fallen uns ein

- Sinus = Gegenkathete / Hypotenuse: $\sin \alpha = \frac{a}{c}$
- Sehnenabschnitt im Einheitskreis
- Mittels Sinussatz: $\sin \alpha : \sin \beta : \sin \gamma = a : b : c$
- Aus Tabellenwerk
- Mithilfe der Reihendarstellung der Sinusfunktion

Uns interessiert hier vor allem die letzte Variante, nicht nur weil wir uns in dieser Vorlesung für Reihen interessieren, sondern auch weil die Reihendarstellung besonders dazu geeignet ist, von Computern

abgearbeitet zu werden. Wir werden nämlich sehen, dass die Reihendarstellung der Sinusfunktion durch folgenden Ausdruck gegeben ist

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots = \sum_{i=0}^{\infty} (-1)^i \frac{x^{2i+1}}{(2i+1)!}$$

Wenn wir hier für $x = 3\pi/8$ setzen, erhalten wir

$$\sin 3\pi/8 = 3\pi/8 - \frac{(3\pi/8)^3}{3!} + \frac{(3\pi/8)^5}{5!} - \dots = \sum_{i=0}^{\infty} (-1)^i \frac{(3\pi/8)^{2i+1}}{(2i+1)!}$$

Das heißt, wir können folgende Teilsummen berechnen

$$\begin{aligned} s_1 &= 1.1780972450 = \frac{3\pi}{8} \\ s_3 &= 0.9055811415 = 3\pi/8 - \frac{(3\pi/8)^3}{3!} \\ s_5 &= 0.9244925752 = 3\pi/8 - \frac{(3\pi/8)^3}{3!} + \frac{(3\pi/8)^5}{5!} \\ s_7 &= 0.9238676365 = 3\pi/8 - \frac{(3\pi/8)^3}{3!} + \frac{(3\pi/8)^5}{5!} - \frac{(3\pi/8)^7}{7!} \end{aligned}$$

und uns so sukzessive dem exakten Ergebnis von $0.923879532511 \dots$ annähern.

Eine andere Anwendung von Reihen taucht bei der Berechnung von Integralen auf. Ein bestimmtes Integral

$$\int_a^b f(x) dx$$

kann bekannterweise ja als Fläche unter der Funktion $f(x)$ interpretiert werden. Wenn keine analytische Form der Stammfunktion $F(x)$ bekannt ist, so können wir das bestimmte Integral näherungsweise als Summe von schmalen Rechtecken der Breite Δx_i und der Höhe $f(x_i)$ berechnen, also

$$\int_a^b f(x) dx \simeq \sum_i^n f(x_i) \Delta x_i,$$

wobei das Ergebnis umso genauer wird, je kleiner die Intervalle Δx_i gewählt werden. Aber auch unbestimmte Integrale können mithilfe von Reihen näherungsweise berechnet werden. Nehmen wir etwa das unbestimmte Integral.

$$\int \sin(x^4) dx = ?$$

Da wir für den Sinus bereits eine Reihendarstellung kennen, können wir dort einfach x durch x^4

ersetzen und erhalten

$$\int \sin(x^4) dx = \int \left(x^4 - \frac{x^{12}}{3!} + \frac{x^{20}}{5!} - \dots \right) dx$$

Weil wir wissen, wie Potenzen in x zu integrieren sind, erhalten wir schließlich eine Reihendarstellung für unser gesuchtes Integral

$$\int \sin(x^4) dx = \frac{x^5}{5} - \frac{x^{13}}{3!13} + \frac{x^{21}}{5!21} - \dots$$

Es gäbe noch zahlreiche weitere Beispiele, wie viele Probleme in der Mathematik – und vor allem auch in der Physik – mithilfe von Folgen und Reihen gelöst werden können.

1.2 Metrische Räume

Bevor wir uns in den folgenden zwei Kapiteln mit der Definition und der Berechnung von Grenzwerten von Folgen und Reihen auseinandersetzen werden, müssen wir an dieser Stelle einige grundlegende Begriffe der Metrik widmen, das heißt, wir werden uns die Frage stellen, wie in der Mathematik der *Abstand* zweier Zahlen definiert ist.

1.2.1 Intervalle

Bei den Zahlenmengen der natürlichen, ganzen, rationalen und reellen Zahlen, also \mathbb{N} , \mathbb{Z} , \mathbb{Q} , und \mathbb{R} gibt es jeweils eine einfache Ordnungsrelation $a < b$ bzw. $a \leq b$, mit deren Hilfe wir besondere Teilmengen der reellen Zahlen, nämlich *Intervalle*, definieren können:

$$\begin{aligned} (a, b) &= \{x \in \mathbb{R} : a < x < b\} \quad (\text{offenes Intervall}) \\ [a, b] &= \{x \in \mathbb{R} : a \leq x \leq b\} \quad (\text{abgeschlossenes Intervall}) \\ [a, b) &= \{x \in \mathbb{R} : a \leq x < b\} \quad (\text{halboffenes Intervall}) \\ (a, b] &= \{x \in \mathbb{R} : a < x \leq b\} \quad (\text{halboffenes Intervall}) \\ (a, \infty) &= \{x \in \mathbb{R} : a < x\} \\ [a, \infty) &= \{x \in \mathbb{R} : a \leq x\} \\ (-\infty, a) &= \{x \in \mathbb{R} : x < a\} \\ (-\infty, a] &= \{x \in \mathbb{R} : x \leq a\} \end{aligned}$$

1.2.2 Norm

Eine Norm (von lateinisch *norma* "Richtschnur") ist eine Abbildung, die einem mathematischen Objekt, beispielsweise einer Zahl, einem Vektor, einer Matrix, einer Folge oder einer Funktion, eine Zahl zuordnet, die auf gewisse Weise die Größe oder "Länge" des Objekts beschreiben soll. In der Menge der reellen Zahlen \mathbb{R} erfüllt der **Absolutbetrag** diese Aufgabe:

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases}$$

Geometrisch betrachtet misst der Absolutbetrag den Abstand der Zahl x vom Ursprung. Folgende einfach zu beweisende Rechenregeln sind erfüllt:

$$\begin{aligned} \forall x \in \mathbb{R} : |x| &\geq 0, |x| = \sqrt{x^2} \\ |xy| &= |x||y|, \left| \frac{x}{y} \right| = \frac{|x|}{|y|} \\ |x + y| &\leq |x| + |y| \dots \text{Dreiecksungleichung} \\ |x| < a &\Leftrightarrow -a < x < a, \text{ wobei } a > 0 \\ |x| > a &\Leftrightarrow (x > a) \vee (x < -a), \text{ wobei } a > 0 \end{aligned}$$

Beispiele. (1) Für welche x gilt $|x - 2| \leq 4$? (2) Bestimme alle $x \in \mathbb{R}$, welche der Ungleichung $|x - 1| + |x + 2| \leq 4$ genügen.

1.2.3 Metrische Räume

Unter einem metrischen Raum versteht man eine Menge, auf der eine Metrik, das heißt eine Abstandsfunktion, definiert ist, die je zwei Elementen des Raums einen nicht negativen reellen Wert zuordnet, der als Abstand der beiden Elemente voneinander aufgefasst werden kann. Mittels der Norm kann in \mathbb{R} (und allen Teilmengen von \mathbb{R}) ein Abstandsbegriff (**Metrik**) eingeführt werden.

Definition. Wir definieren den Abstand zweier reeller Zahlen x und y als $d(x, y) = |x - y|$. Diese Metrik erfüllt offenbar folgende drei Bedingungen, die für die axiomatische Definition einer Metrik erforderlich sind:

$$(M1) \quad d(x, y) \geq 0, d(x, y) = 0 \Leftrightarrow x = y$$

$$(M2) \quad d(x, y) = d(y, x)$$

$$(M3) \quad d(x, z) \leq d(x, y) + d(y, z) \dots \text{Dreiecksungleichung}$$

Beweis. (M1) und (M2) folgen direkt aus der Definition des Absolutbetrags. Für (M3):

$$d(x, z) = |x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = d(x, y) + d(y, z).$$

Beispiel. Wir beweisen für die sogenannte *triviale* Metrik definiert durch $d(x, y) = 1$ für $x \neq y$ und $d(x, y) = 0$ für $x = y$ die Axiome (M1)–(M3).

Im weiteren Verlauf dieser Vorlesung werden wir auch komplexe Zahlen behandeln und auch mehrdimensionale Funktionen kennenlernen. In diesem Zusammenhang erwähnen wir bereits an dieser Stelle, dass sich auch in der Menge der komplexen Zahlen \mathbb{C} genauso wie in Vektorräumen etwa in der Ebene \mathbb{R}^2 eine Metrik definieren lässt. Im zweidimensionalen Raum (Ebene) etwa können wir eine Metrik über den *Euklidischen Abstand* zweier Punkte $s = (x_s, y_s)$ und $t = (x_t, y_t)$ definieren: $d(s, t) = \sqrt{(x_s - x_t)^2 + (y_s - y_t)^2}$. Eine andere Metrik ergibt sich auch durch folgende Wahl der Abstandsfunktion (*Manhattan Metrik*): $d(s, t) = |x_s - x_t| + |y_s - y_t|$. Wie sich leicht zeigen lässt, erfüllen beide Metriken tatsächlich die Axiome (M1)–(M3).

Wie schon eingangs erwähnt versteht man unter einem metrischen Raum eine Menge, auf der eine Metrik definiert ist, also:

Definition. Sei X eine Menge und $d : X \times X \rightarrow \mathbb{R}$ eine Abbildung, welche die Eigenschaften (M1)–(M3) erfüllt, dann heißt d eine Metrik auf X und das Paar (X, d) heißt **metrischer Raum**.

Bemerkung. Mit der Metrik $d(x, y) = |x - y|$ ist \mathbb{R} also ein metrischer Raum.

Um die Konvergenz von Folgen und Reihen mathematisch sauber definieren zu können, werden wir uns noch folgender Definitionen bedienen.

Ist (X, d) ein metrischer Raum, $x_0 \in X$ und ε eine reelle, positive Zahl $\varepsilon > 0$, dann heißt die Menge $K(x_0, \varepsilon) = \{x \in X : d(x_0, x) < \varepsilon\}$ die **offene ε -Kugel** um x_0 mit Radius ε . Analog dazu bezeichnen wir die Menge $B(x_0, \varepsilon) = \{x \in X : d(x_0, x) \leq \varepsilon\}$ als die **abgeschlossene ε -Kugel** um x_0 mit Radius ε . Bemerkung: Im Falle (\mathbb{R}, d) entspricht diese Definition dem offenen Intervall $(x_0 - \varepsilon, x_0 + \varepsilon)$ bzw. dem abgeschlossenen Intervall $[x_0 - \varepsilon, x_0 + \varepsilon]$.

Sei (X, d) ein metrischer Raum, dann heißt die Teilmenge $G \subseteq X$ eine **offene Menge**, wenn zu jedem $x \in G$ ein zugehöriges $\varepsilon_x > 0$ existiert mit der Eigenschaft $x \in K(x, \varepsilon_x) \subseteq G$. Per definitionem ist die leere Menge $\{\}$ eine offene Menge. Mit dieser Definition ist in \mathbb{R} eine Teilmenge $G \subseteq \mathbb{R}$ also genau dann offen, wenn mit jedem $x \in G$ zugleich ein geeignetes offenes Intervall um x ganz in G liegt.

Sei (X, d) ein metrischer Raum, dann heißt die Teilmenge $A \subseteq X$ eine **abgeschlossene Menge**, wenn $X \setminus A$ eine offene Menge ist. Das heißt, zu jedem $x \notin A$ gibt es ein zugehöriges $\varepsilon_x > 0$, sodass $K(x, \varepsilon) \cap A = \{\}$.

Für $X = \mathbb{R}$ gilt, dass \mathbb{R} und jedes offene Intervall (a, b) eine offene Menge darstellt. Jedes abgeschlossene Intervall $[a, b]$ hingegen bildet eine abgeschlossene Menge. Ein halboffenes Intervall $[a, b)$ ist weder offen noch abgeschlossen.

Eine sehr wichtige Klasse von Teilmengen in der Analysis wird von **kompakten Teilmengen** gebildet. Eine Menge $C \subseteq \mathbb{R}$ heißt kompakt, wenn C beschränkt und abgeschlossen ist. Aus dieser Definition folgt, dass jedes abgeschlossene Intervall $[a, b]$ kompakt ist. \mathbb{R} wiederum ist nicht kompakt, weil nicht beschränkt, und ein offenes Intervall (a, b) ist ebenso nicht kompakt, weil nicht abgeschlossen.

1.3 Folgen

In dem Einführungskapitel 1.1 haben wir ja bereits Folgen $\{a_n\}$ als eine geordnete, durchnummerierte Menge kennengelernt, oder genauer eine Zuordnung der natürlichen Zahlen \mathbb{N} zur Menge $\{a_n\}$. Bevor wir in nächsten Unterkapitel den Grenzwert einer Folge definieren, sind folgende Definitionen hilfreich:

Bei einer **monoton steigenden** Folge ist jedes Folgenglied größer als sein Vorgänger, d.h. $\forall n \in \mathbb{N} : a_{n+1} \geq a_n$.

Bei einer **monoton fallenden** Folge ist jedes Folgenglied kleiner als sein Vorgänger, d.h. $\forall n \in \mathbb{N} : a_{n+1} \leq a_n$.

Bei einer **alternierenden** Folge wechselt das Vorzeichen von einem Folgenglied zum nächsten, d.h. $\forall n \in \mathbb{N} : a_{n+1}a_n < 0$.

Definition. Eine Element A nennt man **Häufungspunkt** einer Menge M , wenn in einer auch noch so kleinen Umgebung von A (offene ε -Kugel) unendlich viele Elemente von M liegen.

Es ist zu beachten, dass ein Häufungspunkt A der Menge M nicht notwendigerweise selbst zur Menge M gehören muss! Die Folge $\{a_n\} = \frac{1}{n}$ besitzt etwa den Häufungspunkt 0, der nicht Teil der Menge ist. Die alternierende Folge $\{a_n\} = (-1)^n$ besitzt zwei Häufungspunkte, nämlich -1 und $+1$. In diesem Falls sind beide Häufungspunkte auch Teil der Folge.

Definition. Besitzt eine Folge $\{a_n\}$ **genau einen Häufungspunkt**, dann nennt man diesen Häufungspunkt den **Grenzwert** (lat. *limes*) dieser Folge und bezeichnet die Folge als **konvergent**. Mathematisch ausgedrückt, schreiben wir:

$$a = \lim_{n \rightarrow \infty} a_n \Leftrightarrow \forall \varepsilon > 0 \exists N_\varepsilon : \forall n > N_\varepsilon : |a - a_n| < \varepsilon$$

In Worten: a nennt man Grenzwert der Folge a_n genau dann, wenn für alle positiven, beliebig kleinen reellen Zahlen ε eine natürliche Zahl N_ε existiert, sodass für alle natürlichen Zahlen n größer als N_ε der Abstand zwischen dem Grenzwert a und dem Folgenglied a_n beliebig klein wird.

Beispiel. Nach dieser Definition ist die Folge $\{a_n\} = \frac{1}{n}$ konvergent und hat den Grenzwert 0. Die Folge $\{a_n\} = (-1)^n$ hingegen besitzt keinen Grenzwert. Beweis erfolgt durch Einsetzen in die obige Definition.

Ist der Grenzwert einer Folge 0, so nennt man diese Folge eine **Nullfolge**. Kennt man den Grenzwert einer Folge $\lim_{n \rightarrow \infty} b_n = b$, so kann man aus dieser Folge durch Subtraktion des Grenzwerts immer eine Nullfolge machen, d.h. $\{c_n\} = \{b_n - b\}$, und $\lim_{n \rightarrow \infty} c_n = 0$.

Ganz allgemein lässt sich zeigen, dass man mit Grenzwerten, sofern sie existieren, ganz ähnlich rechnen kann wie mit gewöhnlichen reellen Zahlen. Das heißt, falls die Folgen $\{a_n\}$ und $\{b_n\}$ konvergent sind, also die Grenzwerte $\lim_{n \rightarrow \infty} a_n = A$ und $\lim_{n \rightarrow \infty} b_n = B$ existieren, so gelten folgende Rechenregeln:

$$\lim_{n \rightarrow \infty} (\alpha a_n \pm \beta b_n) = \alpha \lim_{n \rightarrow \infty} a_n \pm \beta \lim_{n \rightarrow \infty} b_n \quad (1.2)$$

$$\lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n \quad (1.3)$$

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n} \quad (\text{falls } \lim_{n \rightarrow \infty} b_n \neq 0) \quad (1.4)$$

Um die Bestimmung von Grenzwerten zu erleichtern, geben wir hier einige bekannte Grenzwerte an:

$$\lim_{n \rightarrow \infty} n^\alpha = \begin{cases} \infty & \text{wenn } \alpha > 0 \\ 1 & \text{wenn } \alpha = 0 \\ 0 & \text{wenn } \alpha < 0 \end{cases} \quad (1.5)$$

$$\lim_{n \rightarrow \infty} q^n = \begin{cases} \infty & \text{wenn } q > 1 \\ 1 & \text{wenn } q = 1 \\ 0 & \text{wenn } -1 < q < 1 \\ \text{existiert nicht} & \text{wenn } q \leq -1 \end{cases} \quad (1.6)$$

$$\lim_{n \rightarrow \infty} c^{\frac{1}{n}} = 1 \quad \text{wenn } c > 0 \quad (1.7)$$

$$\lim_{n \rightarrow \infty} n^{\frac{1}{n}} = 1 \quad (1.8)$$

Mit den oben erwähnten Rechenregeln für Grenzwerte und der Liste elementarer Grenzwerte sind wir nun in der Lage die Grenzwerte von vielen Folgen zu berechnen. Wir werden zu einem späteren Zeitpunkt noch eine sehr effiziente Methode kennenlernen, um Grenzwerte von Folgen zu berechnen, nämlich die *Regel von de L'Hospital*.

Beispiel. Sind die unten angeführten Folgen konvergent? Wenn ja, wie lautet ihr Grenzwert?

$$(a) \lim_{n \rightarrow \infty} \frac{3n - 1}{2(n + 1)}, (b) \lim_{n \rightarrow \infty} \frac{2n^3 - n}{5n^3 - 2n}, (c) \lim_{n \rightarrow \infty} \frac{n!}{n^n}$$

Auch die Eulersche Zahl e bzw. die Exponentialfunktion e^x lässt sich als Grenzwert einer Folge definieren

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = 2.718281828459045 \dots \equiv e \tag{1.9}$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x \text{ für } x \in \mathbb{R}, \tag{1.10}$$

was auch in der nachfolgenden CDF-Anwendung visualisiert wird.¹

CDF 1. Euler-Zahl <https://moodle.uni-graz.at/course/view.php?id=95124>

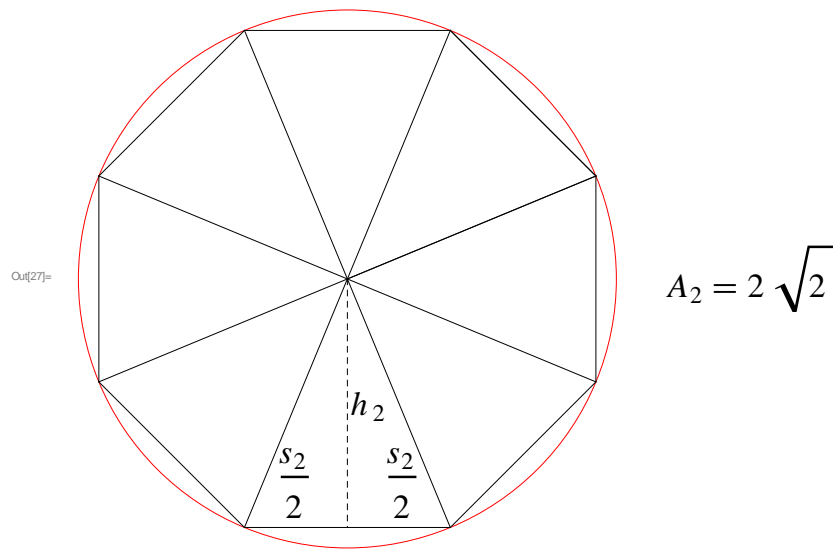
Neben dieser Folge, $a_n = \left(1 + \frac{1}{n}\right)^n$, lassen sich noch weitere Folgen angeben, deren Grenzwert ebenfalls die Eulersche Zahl e ist, etwa $b_n = \frac{n}{\sqrt[n]{n!}}$ oder $c_n = \sum_{k=1}^n \frac{1}{k!}$ (c_n ist eigentlich eine Reihe, mehr dazu später). Das soll diese CDF-Demonstration veranschaulichen. Es zeigt sich insbesondere auch, dass die Geschwindigkeit, mit der eine Folge gegen ihren Grenzwert konvergiert, von Folge zu Folge sehr verschieden sein kann.

n	$a_n = \left(1 + \frac{1}{n}\right)^n$	$e - a_n$	$b_n = \frac{n}{\sqrt[n]{n!}}$	$e - b_n$	$c_n = \sum_{k=1}^n \frac{1}{k!}$	$e - c_n$
1	2	0.7182818285	1	1.718281828	2	0.7182818285
2	$\frac{9}{4}$	0.4682818285	$\sqrt{2}$	1.304068266	$\frac{5}{2}$	0.2182818285
3	$\frac{64}{27}$	0.3479114581	$\frac{3^{2/3}}{2^{1/3}}$	1.067318204	$\frac{8}{3}$	0.05161516179
4	$\frac{625}{256}$	0.2768755785	$2 \left(\frac{2}{3}\right)^{1/4}$	0.9110778212	$\frac{65}{24}$	0.009948495126
5	$\frac{7776}{3125}$	0.2299618285	$\frac{5^{4/5}}{2^{3/5} 3^{1/5}}$	0.7990220803	$\frac{163}{60}$	0.001615161792
6	$\frac{117\ 649}{46\ 656}$	0.1966554567	$\frac{2^{1/3} 3^{2/3}}{5^{1/6}}$	0.7141366989	$\frac{1957}{720}$	0.0002262729035
7	$\frac{2\ 097\ 152}{823\ 543}$	0.1717821314	$\frac{7^{6/7}}{2^{4/7} 3^{2/7} 5^{1/7}}$	0.6472851996	$\frac{685}{252}$	0.00002786020508
8	$\frac{43\ 046\ 721}{16\ 777\ 216}$	0.1524973145	$\frac{4 \left(\frac{2}{3}\right)^{1/8}}{3^{1/4}}$	0.5930812690	$\frac{109\ 601}{40\ 320}$	$3.058617775 \times 10^{-6}$

¹Das Computable Document Format (CDF) ist ein von Wolfram Research entwickeltes Dokumentenformat für interaktive Dokumente. Zum Erstellen von CDF-Dokumenten benötigt man Mathematica. CDF-Dokumente können allerdings nicht nur mit Mathematica gelesen werden, sondern Wolfram Research bietet für das Lesen von CDF-Dokumenten den **CDF-Player** kostenlos zum Download an.

CDF 2. Rekursive Folge für π <https://moodle.uni-graz.at/course/view.php?id=95124>

Schreibt man in einen Kreis mit dem Radius 1 ein Quadrat ein, so hat dieses die Fläche $A_1 = 2$. Als zweites Glied der Folge betrachtet man ein regelmäßiges 8-Eck, dessen Fläche sich zu $A_2 = 2\sqrt{2}$ ergibt (siehe Abbildung). Die weiteren Glieder der Folge sind die Flächeninhalte eines eingeschriebenen 16-Ecks ($n = 3$), 32-Ecks ($n = 4$) usw., deren Grenzwert die Zahl π ist. Die Glieder dieser rekursiven Folge werden in dieser CDF-Anwendung berechnet. Hinweis: $A_1 = 4 \cdot \frac{1}{2}s_1h_1 = 2$, wobei $s_1 = \sqrt{2}$ und $h_1 = \sqrt{1 - (\frac{s_1}{2})^2}$ die Grundfläche bzw. Höhe eines der 4 Dreiecke ist, in das das Quadrat unterteilt wird. Der jeweils nächste Flächeninhalt ist dann $A_{k+1} = 2^{k+2} \cdot \frac{1}{2}s_{k+1}h_{k+1}$, wobei $s_{k+1} = \sqrt{(1 - h_k)^2 + (\frac{s_k}{2})^2}$ und $h_{k+1} = \sqrt{1 - (\frac{s_{k+1}}{2})^2}$.



n	A_n	A_n	$\pi - A_n$
1	2	2.000000000000000	1.14159265359
2	$2\sqrt{2}$	2.82842712474619	0.313165528844
3	$4\sqrt{2 - \sqrt{2}}$	3.06146745892072	0.0801251946691
4	$8\sqrt{2 - \sqrt{2 + \sqrt{2}}}$	3.12144515225805	0.0201475013317
5	$16\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}}}$	3.13654849054594	0.00504416304385
6	$32\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2}}}}}$	3.14033115695475	0.00126149663504
7	$64\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2}}}}}}$	3.14127725093277	0.000315402657020
8	$128\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2}}}}}}}$	3.14151380114430	0.0000788524454922

1.4 Reihen

1.4.1 Die vollständige Induktion

Bevor wir uns Reihen und deren Konvergenzeigenschaften widmen, behandeln wir hier ein Beweisverfahren (die vollständige Induktion), das für Aussagen $A(n)$, die von einer natürlichen Zahl n abhängen, anwendbar ist, z.B.

$$A(n) \quad : \quad 1 + 2 + 3 + \dots = \frac{n(n+1)}{2} \quad \text{für } n \geq 1$$

Beweise nach dem Prinzip der vollständigen Induktion beinhalten immer die folgenden vier Beweisschritte:

- (I) **Induktionsanfang:** Die Aussage $A(n)$ ist wahr für einen Anfangswert $n = n_0$ (sehr oft $n_0 = 1$).
- (II) **Induktionsvoraussetzung:** $A(n)$ sei wahr für $n \geq n_0$.
- (III) **Induktionsbehauptung:** $A(n+1)$ ist wahr.
- (IV) **Induktionsbeweis:** Gelingt es allgemein zu beweisen, dass aus der Gültigkeit der Induktionsvoraussetzung (II) die Induktionsbehauptung (III) folgt, dann ist $A(n)$ für all $n \geq n_0$ erfüllt.

Beispiel. Wir beweisen mit Hilfe von vollständiger Induktion, dass gilt:

$$A(n) = 1 + 2 + 3 + \dots = \sum_{k=1}^n k = \frac{n(n+1)}{2}.$$

Die vier Schritte des Beweisverfahrens lauten für diesen Fall:

- (I) **Induktionsanfang:** linke Seite = rechte Seite für $A(1)$

$$n = 1 : \sum_{k=1}^1 k = 1 = \frac{1 \cdot 2}{2}$$

- (II) **Induktionsvoraussetzung** = die Aussage, die zu beweisen ist, also

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

(III) **Induktionsbehauptung:** Wir ersetzen in der Induktionsvoraussetzung $n \rightarrow n + 1$, also

$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}$$

(IV) **Induktionsbeweis:** Umformen der Induktionsbehauptung (III) unter Verwendung der Induktionsvoraussetzung (II)

$$\sum_{k=1}^{n+1} k = \left(\sum_{k=1}^n k \right) + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}.$$

Somit ist der Beweis der Aussage für alle $n \geq 1$ gelungen.

Beispiel. Als weiteres Beispiel für vollständige Induktion wollen wir die sogenannte Bernoulli-Ungleichung beweisen

$$(1+x)^n > 1+nx \quad \text{für} \quad x > -1, \quad x \neq 0 \quad \text{und} \quad n \geq 2.$$

Die vier Schritte des Beweisverfahrens lauten für diesen Fall:

(I) **Induktionsanfang:** linke Seite = rechte Seite für $n = 2$

$$(1+x)^2 = 1+2x+x^2 > 1+2x \quad \text{wegen} \quad x \neq 0.$$

(II) **Induktionsvoraussetzung** = die Aussage, die zu beweisen ist, also

$$(1+x)^n > 1+nx$$

(III) **Induktionsbehauptung:** Wir ersetzen in der Induktionsvoraussetzung $n \rightarrow n + 1$, also

$$(1+x)^{n+1} > 1+(n+1)x$$

(IV) **Induktionsbeweis:** Umformen der Induktionsbehauptung (III) unter Verwendung der Induktionsvoraussetzung (II)

$$(1+x)^{n+1} = (1+x)^n(1+x) > (1+nx)(1+x) = 1+nx+x+nx^2 > 1+(n+1)x,$$

weil $1+x > 0$ und $nx^2 > 0$. Somit ist der Beweis der Aussage für alle $n \geq 2$ gelungen.

1.4.2 Konvergenz von Reihen

Im Einführungskapitel 1.1 hatten wir ja schon unendliche Reihen kennengelernt, die wir durch Aufsummieren der Glieder einer Folge $\{a_k\}$ gewonnen haben. Wir definieren: Sei $\{a_k\}$ eine Folge, dann bilden wir die zugehörige Folge $\{s_n\}$ mit $s_n = \sum_{k=1}^n a_k$, die wir als die **n -te Partialsumme** bezeichnen. Eine unendliche Reihe

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \dots$$

ist konvergent, wenn die Partialsummen s_n einen Grenzwert S besitzen. S nennt man dann die Summe der Reihe und schreibt

$$S = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k = \sum_{k=1}^{\infty} a_k$$

Es ist von großer Bedeutung zu wissen, ob eine unendliche Reihe konvergiert oder nicht konvergiert (=divergiert). Betrachten wir das Beispiel

$$\sum_{k=0}^{\infty} (-1)^k = 1 - 1 + 1 - 1 + \dots = \underbrace{1 - 1}_{=0} + \underbrace{1 - 1}_{=0} + \dots = 0$$

Wir können natürlich Summanden auch anders zusammenfassen,

$$\sum_{k=0}^{\infty} (-1)^k = 1 - 1 + 1 - 1 + \dots = 1 + \underbrace{(-1 + 1)}_{=0} + \underbrace{(-1 + 1)}_{=0} + \dots = 1$$

Der Grund für diesen offensichtlichen Widerspruch ist die Tatsache, dass diese Reihe nicht konvergiert. Die Folge der Partialsummen hat zwei Häufungspunkte nämlich 0 und +1.

In einem anderen Beispiel betrachten wir die Reihe, die offensichtlich divergiert, von der wir aber naiv annehmen wollen, dass sie eine endliche Summe S besitzt

$$S = 1 + 2 + 4 + 8 + \dots$$

Multiplizieren wir beide Seiten der obigen Gleichung mit 2, so erhalten wir

$$2S = 2 + 4 + 8 + 16 + \dots$$

Die Subtraktion dieser beiden Gleichungen führt zu dem überraschenden Ergebnis $S = -1$! Dieser offensichtliche Widerspruch ergibt sich aus der Tatsache, dass wir angenommen haben, dass die Reihe konvergiert. Und nur in diesem Fall dürfen wir mit Summen von unendlichen Reihen wie mit gewöhnlichen Zahlen rechnen. Falls die Reihe divergiert, kann das zu unsinnigen Ergebnissen führen.

Bevor wir in den nächsten beiden Kapiteln einige elementare Reihen auf deren Konvergenz untersuchen werden, und einige Kriterien kennenlernen, mit deren Hilfe die Konvergenz (oder Divergenz) einer Reihe überprüft werden kann, wollen wir hier noch einige Typen von Reihen definieren.

Absolut konvergent. Wir bezeichnen eine Reihe als absolut konvergent, wenn auch die Reihe gebildet aus den Absolutbeträgen konvergiert, also

$$\sum_{k=0}^{\infty} |a_k| = |a_1| + |a_2| + |a_3| + \dots < \infty$$

Unbedingt konvergent. Wir nennen eine Reihe unbedingt konvergent, wenn sie unabhängig von der Ordnung der Summanden gegen dieselbe Summe konvergiert. Umgekehrt heißt eine Reihe bedingt konvergent, wenn sie nur für bestimmte Anordnungen der Reihenglieder konvergiert.

Alternierend. Eine Reihe heißt alternierend, wenn bei zwei aufeinanderfolgende Glieder der Reihe das Vorzeichen wechselt.

1.4.3 Elementare Reihen

Arithmetische Reihe.

$$\sum_{k=1}^{\infty} k = 1 + 2 + 3 + 4 + \dots$$

Die oben angeführte Reihe ist eigentlich nur der Spezialfall einer allgemeinen arithmetischen Reihe $\sum_{k=0}^{\infty} (a_0 + kd)$ mit $a_0 = 1$ und $d = 1$. Für unseren Spezialfall kann man durch *vollständige Induktion* leicht zeigen, dass die Partialsumme den Wert

$$s_n = \sum_{k=1}^n k = \frac{n(n+1)}{2}$$

hat. Wie unschwer zu erkennen, divergiert also die arithmetische Reihe. Wir erahnen, dass ein *notwendiges* Kriterium für die Konvergenz von Reihen ist, dass die Glieder der Reihe eine Nullfolge bilden. Das stellt aber keineswegs ein *hinreichendes* Kriterium dar, wie wir gleich am nächsten Beispiel sehen werden.

Harmonische Reihe.

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

Diese Reihe haben wir bereits im Einführungskapitel 1.1 kennengelernt. Hier konnten wir durch geeignetes Zusammenfassen von Summanden in der Art:

$$S = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{> 2 \cdot \frac{1}{4} = \frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{> 4 \cdot \frac{1}{8} = \frac{1}{2}} + \underbrace{\frac{1}{9} + \dots + \frac{1}{16}}_{> 8 \cdot \frac{1}{16} = \frac{1}{2}} + \underbrace{\frac{1}{17} + \dots + \frac{1}{32}}_{> 16 \cdot \frac{1}{32} = \frac{1}{2}} + \dots > k \frac{1}{2}$$

zeigen, dass die Partialsummen über jede Grenze wachsen, d.h. dass auch die harmonische Reihe divergiert. Wir betonen nochmals, dass der Umstand, dass die Folge $\{\frac{1}{k}\}$ eine Nullfolge ist nicht notwendig die Konvergenz der dazugehörigen Reihe bedingt.

Beachte, dass folgende *notwendige* bzw. *hinreichende* Kriterien angegeben werden können:

$$\sum_k a_k \dots \text{konvergent} \Rightarrow \lim_{k \rightarrow \infty} a_k = 0$$

$$\sum_k a_k \dots \text{divergent} \Leftarrow \lim_{k \rightarrow \infty} a_k \neq 0$$

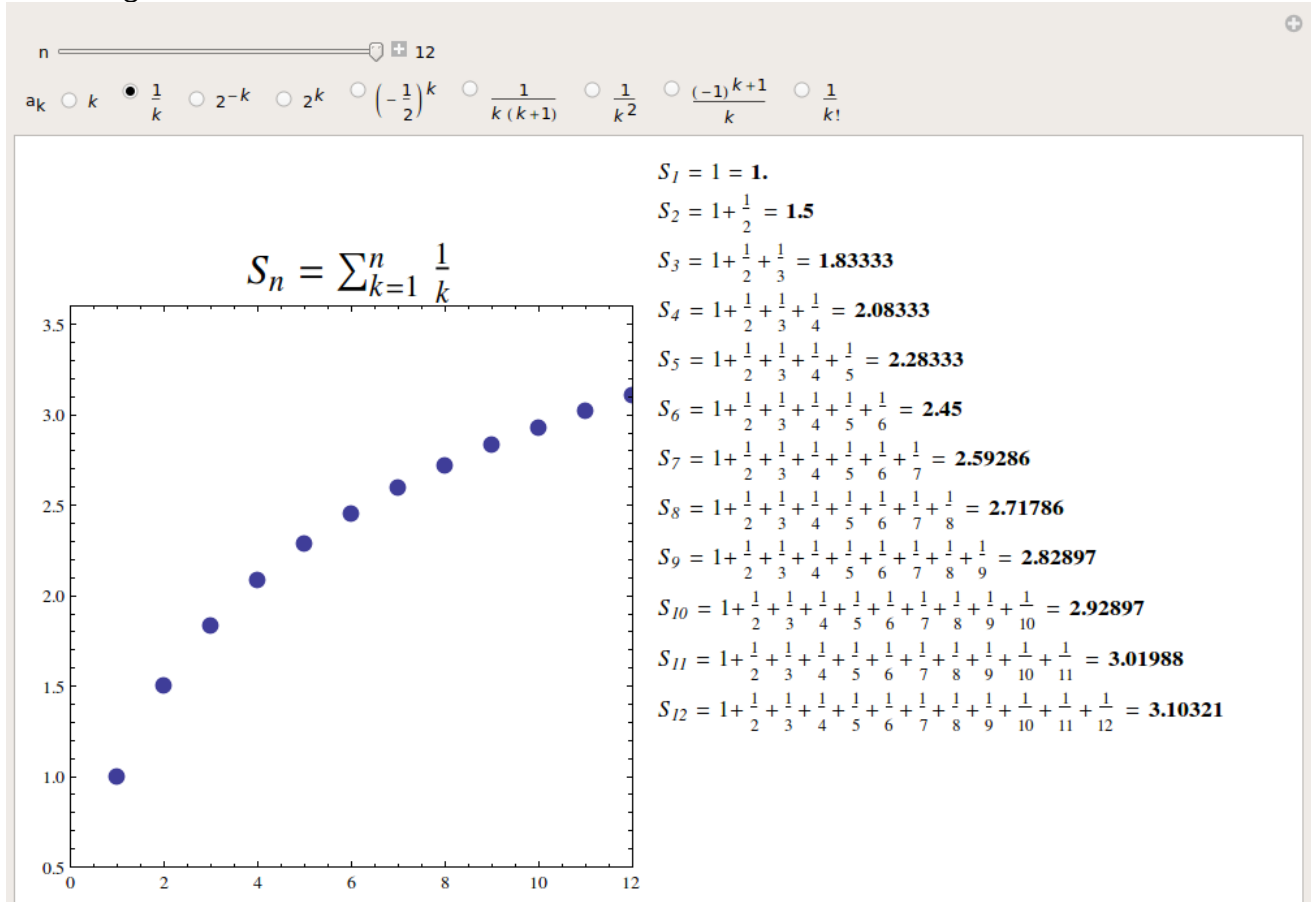
aber, dass die Umkehrung **nicht** gilt!

$$\sum_k a_k \dots \text{divergent} \not\Rightarrow \lim_{k \rightarrow \infty} a_k > 0$$

$$\sum_k a_k \dots \text{konvergent} \Leftarrow \lim_{k \rightarrow \infty} a_k = 0.$$

CDF 3. Elementare Reihen Beispiele_fuer_Reihen.cdf

In dieser CDF-Anwendung können für verschiedene elementare Reihen Partialsummen s_n berechnet und grafisch dargestellt werden.



Geometrische Reihe.

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + q^4 + \dots$$

Wir betrachten zunächst den Fall $q = 1$. In diesem Fall ist die Partialsumme $s_n = 1 + n$ und die Reihe divergiert. Für $q \neq 0$ können wir ebenfalls einen geschlossenen Ausdruck für die Partialsumme erhalten, indem wir die zwei Ausdrücke $s_n = 1 + q + q^2 + \dots + q^n$ und $qs_n = q + q^2 + \dots + q^{n+1}$ kombinieren, was zu folgendem Ergebnis führt:

$$s_n = \frac{1 - q^{n+1}}{1 - q}$$

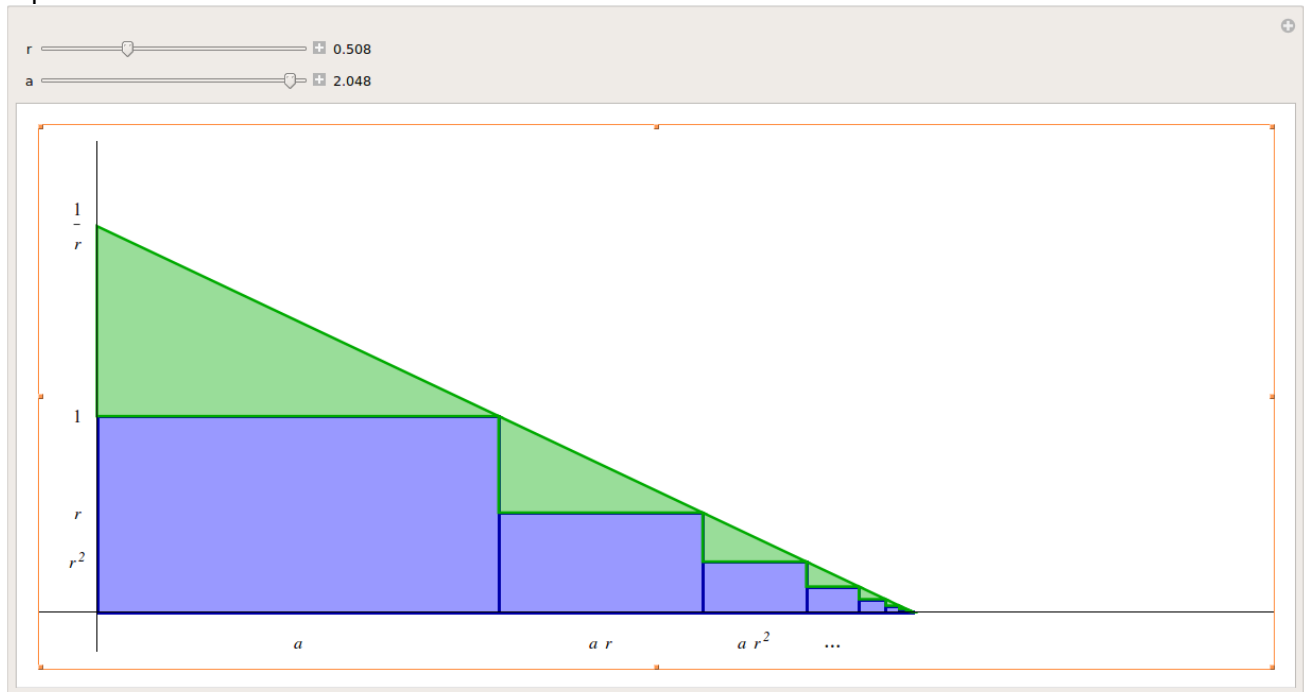
Aufgrund der Eigenschaften der geometrischen Folge zusammengefasst in Gleichung (1.6) sehen wir, dass die geometrische Reihe für $|q| < 1$ konvergiert, während sie für $|q| \geq 1$ divergiert. Im Falle der

Konvergenz lautet die Summe der geometrischen Reihe:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \text{ für } |q| < 1 \quad (1.11)$$

CDF 4. Geometrische Reihe [SumOfAGeometricSeries.cdf](#)

In dieser CDF-Anwendung wird die Partialsumme der geometrischen Reihe mithilfe einer geometrischen Interpretation ermittelt.



All the green triangles are similar, and the big triangle is similar to the green triangles.

The ratio of the big triangle's sides is equal to the ratio of the sides for those triangles, so that $\frac{a+ar+ar^2+\dots+ar^n+\dots}{\frac{1}{r}} = \frac{ar}{1-r}$, and then $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$.

Teleskop-Reihe.

$$\sum_{k=2}^{\infty} \frac{1}{(k-1)k} = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots$$

Auch diese Reihe ist konvergent, was man sehen kann, indem man die Reihenglieder in der Art $\frac{1}{(k-1)k} = \frac{1}{k-1} - \frac{1}{k}$ aufspaltet. Dann erkennt man, dass sich die Partialsumme schreiben lässt als:

$$s_n = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{n-1} - \frac{1}{n} = 1 - \frac{1}{n}$$

Das heißt, je zwei Nachbarglieder (außer dem ersten und dem letzten) heben sich gegenseitig auf. Diesen Vorgang nennt man Teleskopieren einer Summe, abgeleitet vom Ineinanderschieben zweier

oder mehrerer zylindrischer Rohre. Aus der Partialsumme ist unschwer zu erkennen, dass diese Reihe konvergiert und die Summe $S = 1$ aufweist.

1.4.4 Konvergenztests

Vergleichskriterium. Ausgangspunkt für viele Konvergenzbeweise ist das sogenannte Vergleichskriterium. Ist bekannt, dass die Reihe $\sum_{n=1}^{\infty} b_n$, die aus positiven Gliedern $b_n > 0$ besteht, konvergiert, so konvergiert auch eine Reihe $\sum_{n=1}^{\infty} a_n$ dann, wenn für fast alle Glieder die Ungleichung $|a_n| \leq b_n$ erfüllt ist. Die Reihe b_n nennt man dann eine *konvergente Majorante* von a_n . Umgekehrt lässt sich zeigen, dass eine Reihe $\sum_{n=1}^{\infty} a_n$ divergiert, wenn für fast alle Glieder die Ungleichung $|a_n| \geq b_n$ erfüllt ist, und bekannt ist, dass $\sum_{n=1}^{\infty} b_n$ divergiert. Die Reihe b_n nennt man dann eine *divergente Minorante* von a_n .

Als Beispiel betrachten wir die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$$

Diese Reihe hat große Ähnlichkeit mit der konvergenten Teleskopreihe, deren Reihenglieder ja $b_k = \frac{1}{(k-1)k}$ für $k \geq 2$ sind. Setzen wir hier noch $b_k = 1$, dann gilt $a_k = \frac{1}{k^2} \leq \frac{1}{(k-1)k} = b_k \forall k$, was die Konvergenz der obigen Reihe beweist.

In einem anderen Beispiel untersuchen wir die Reihe

$$\sum_{k=1}^{\infty} \frac{1 + \sqrt{k}}{k} = 2 + \frac{1 + \sqrt{2}}{2} + \frac{1 + \sqrt{3}}{3} + \dots$$

Hier benutzen wir die harmonische Reihe als divergente Minorante in der Art: $|a_k| = \frac{1 + \sqrt{k}}{k} \geq \frac{1}{k} = b_k$, um die Divergenz der Reihe zu beweisen.

Wurzelkriterium. Bei diesem Kriterium wird die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$ untersucht, indem der Grenzwert der Folge

$$\rho = \lim_{k \rightarrow \infty} \rho_k = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \begin{cases} \rho < 1 & \text{die Reihe ist absolut konvergent} \\ \rho = 1 & \text{keine Aussage ist möglich} \\ \rho > 1 & \text{die Reihe divergiert} \end{cases}$$

Der Beweis läuft über die Konvergenz der geometrischen Reihe, die als konvergente Majorante fungiert (siehe z.B. Skriptum von Prof. Ganster [pdf](#)).

Als Beispiel betrachten wir die Reihe

$$\sum_{k=1}^{\infty} \frac{k^4}{e^k}$$

Die Berechnung des Grenzwerts ρ zeigt, dass die Reihe konvergent ist:

$$\rho = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \lim_{k \rightarrow \infty} \frac{1}{e} \left(\sqrt[k]{k} \right)^4 = \frac{1}{e} \left(\lim_{k \rightarrow \infty} \sqrt[k]{k} \right)^4 = \frac{1}{e} < 1$$

Quotientenkriterium. Ähnlich wie beim Wurzelkriterium, wird auch bei diesem Kriterium die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$ untersucht, indem der Grenzwert einer Folge ausgewertet wird:

$$\rho = \lim_{k \rightarrow \infty} \rho_k = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = \begin{cases} \rho < 1 & \text{die Reihe ist absolut konvergent} \\ \rho = 1 & \text{keine Aussage ist möglich} \\ \rho > 1 & \text{die Reihe divergiert} \end{cases}$$

Der Beweis des Quotientenkriteriums benutzt die Konvergenz der geometrischen Reihe, die in der Beweisführung wieder als konvergente Majorante eingesetzt wird (siehe z.B. Skriptum von Prof. Ganster [pdf](#)).

Als Beispiel betrachten wir die Reihe

$$\sum_{k=1}^{\infty} \frac{k!}{k^k}$$

Die Berechnung des Grenzwerts von ρ_k zeigt, dass die Reihe konvergent ist:

$$\rho_k = \left| \frac{a_{k+1}}{a_k} \right| = \frac{(k+1)!}{(k+1)^{k+1}} \frac{k^k}{k!} = \frac{k^k}{(k+1)^k} = \left(\frac{k}{k+1} \right)^k = \frac{1}{\left(1 + \frac{1}{k}\right)^k} \rightarrow \frac{1}{e} < 1$$

Verdichtungssatz von Cauchy. Sei $\sum_{k=1}^{\infty} a_k$ eine Reihe mit positiven $a_k > 0$ und monoton fallenden $a_{k+1} < a_k$ Reihengliedern für alle k . Dann besagt der Satz, dass $\sum_{k=1}^{\infty} a_k$ genau dann konvergiert, wenn die Reihe $\sum_{k=1}^{\infty} 2^k a_{2^k}$ konvergiert. Für den Beweis dieses Satzes sei auf weiterführende Literatur verwiesen, z.B. Konrad Königsberger: Analysis 1.

Als prototypisches Beispiel wenden wir dieses Konvergenzkriterium auf die verallgemeinerte harmo-

nische Reihe mit einem festen $\alpha > 0$ an.

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} = 1 + \frac{1}{2^{\alpha}} + \frac{1}{3^{\alpha}} + \frac{1}{4^{\alpha}} + \dots$$

Wir wissen ja bereits, dass die harmonische Reihe ($\alpha = 1$) divergiert. Wir haben auch schon festgestellt, dass die Reihe mit $\alpha = 2$ konvergiert. Mithilfe des Verdichtungssatzes von Cauchy können wir nun das Konvergenzverhalten für beliebige $\alpha > 0$ beurteilen:

$$\sum_{k=1}^{\infty} 2^k a_{2^k} = \sum_{k=1}^{\infty} 2^k \frac{1}{(2^k)^{\alpha}} = \sum_{k=1}^{\infty} (2^{1-\alpha})^k = \sum_{k=1}^{\infty} q^k$$

Wir sehen also, dass die Reihe genau dann konvergiert, wenn die entsprechende geometrische Reihe mit $q = 2^{1-\alpha}$ konvergiert, was für $q < 1$, das heißt für $\alpha > 1$, der Fall ist.

Leibniz-Kriterium. Dieses Kriterium bezieht sich nur auf **alternierende** Reihen. Und zwar ist eine alternierende Reihe genau dann konvergent, wenn (i) die Absolutwerte der Reihenglieder monoton fallend sind, und (ii) die Glieder eine Nullfolge bilden.

Wir wollen den Beweis des Leibniz-Kriteriums skizzieren. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass das erste Glied der Reihe positiv ist. Zunächst betrachten wir die Folgen der Partialsummen mit ungeraden Indizes

$$\begin{aligned} s_1 &= a_1 \\ s_3 &= s_1 - \underbrace{(a_2 - a_3)}_{\geq 0} \leq s_1 \\ s_5 &= s_3 - \underbrace{(a_4 - a_5)}_{\geq 0} \leq s_3 \\ &\vdots \\ s_{2n+1} &= s_{2n-1} - \underbrace{(a_{2n} - a_{2n+1})}_{\geq 0} \leq s_{2n-1} \end{aligned}$$

Die ungeraden Partialsummen bilden also eine *monoton fallende* Folge, die nach unten begrenzt ist durch $s_2 = a_1 - a_2$. Umgekehrt bildet die Teilfolge der Partialsummen mit geraden Indizes eine

monoton steigende Folge, die nach oben durch $s_1 = a_1$ begrenzt ist.

$$\begin{aligned}
 s_2 &= a_1 - a_2 \\
 s_4 &= s_2 + \underbrace{(a_3 - a_4)}_{\geq 0} \geq s_2 \\
 s_6 &= s_4 + \underbrace{(a_5 - a_6)}_{\geq 0} \geq s_4 \\
 &\vdots \\
 s_{2n+2} &= s_{2n} + \underbrace{(a_{2n+1} - a_{2n+2})}_{\geq 0} \geq s_{2n}
 \end{aligned}$$

Nun wissen wir, dass eine monoton steigende (bzw. fallende) und nach oben (bzw. unten) beschränkte Folge einen Grenzwert besitzt, also

$$\lim_{n \rightarrow \infty} s_{2n+1} = S^* \quad \text{und} \quad \lim_{n \rightarrow \infty} s_{2n} = S^{**}$$

Bilden wir nun die Differenz der beiden Grenzwerte

$$S^* - S^{**} = \lim_{n \rightarrow \infty} (s_{2n+1} - s_{2n}) = \lim_{n \rightarrow \infty} a_{2n+1} = 0,$$

so stellen wir fest, dass die beiden Grenzwerte gleich sind $S \equiv S^* = S^{**}$ und die Summe der Reihe bilden, da nach unserer Voraussetzung die Glieder der Reihe eine Nullfolge bilden.

Als Beispiel für die Anwendung des Leibniz-Kriteriums betrachten wir die alternierende Reihe

$$\sum_{k=1}^{\infty} (-1)^k \frac{\sqrt{3k-1}}{\sqrt{k(k+1)}}$$

Damit das Leibniz-Kriterium erfüllt ist, müssen die Beträge der Reihenglieder (i) monoton fallend sein, und (ii) eine Nullfolge bilden. Zunächst zeigen wir (i)

$$\begin{aligned}
 \frac{\sqrt{3k-1}}{\sqrt{k(k+1)}} &> \frac{\sqrt{3(k+1)-1}}{\sqrt{(k+1)(k+2)}} = \frac{\sqrt{3k+2}}{\sqrt{(k+1)(k+2)}} \\
 \frac{3k-1}{k} &> \frac{3k+2}{k+2} \\
 3k &> 2 \quad \dots \text{gilt } \forall k
 \end{aligned}$$

Als zweites müssen wir noch überprüfen, ob die Reihenglieder eine Nullfolge bilden:

$$a_k = \frac{\sqrt{3k-1}}{\sqrt{k(k+1)}} < \frac{\sqrt{3k}}{\sqrt{k(k+1)}} = \frac{\sqrt{3}}{\sqrt{k+1}} \rightarrow 0$$

Somit haben wir gezeigt, dass die Reihe konvergiert.

Integraltest. Beim Integraltest wird die Konvergenz einer Reihe $\sum_{k=1}^{\infty} a_k$ untersucht, indem das korrespondierende Integral ausgewertet wird:

$$\lim_{n \rightarrow \infty} \int_{k_0}^n a(k) dk = \int_{k_0}^{\infty} a(k) dk$$

Voraussetzung für die Anwendung des Integraltests ist, dass die Reihenglieder eine monoton fallende Folge bilden. Hierbei wird ein Vergleich der Fläche unter der Funktion $a(k)$ mit der Summe der Rechtecksflächen der Höhe a_k und der Breite $\Delta k = 1$, also $\sum_{k=1}^{\infty} a_k \Delta k$, angestellt. Daraus kann man ableiten, dass die Existenz des obigen Integrals ein Kriterium für die Konvergenz der Reihe ist: Konvergiert das Integral, so konvergiert auch die Reihe. Divergiert das Integral, so divergiert auch die Reihe.

Es ist zu beachten, dass die untere Grenze des Integrals, k_0 , für die Konvergenz (Divergenz) der Reihe unerheblich ist, da ja eine beliebige, endliche Zahl von Reihengliedern weggelassen werden darf, ohne dass sich das Konvergenzverhalten der Reihe ändert. Wichtig ist nur das Verhalten des bestimmten Integrals an der oberen Grenze $n \rightarrow \infty$.

Als Beispiel für die Anwendung des Integraltests betrachten wir wiederum die verallgemeinerte harmonische Reihe mit einem festen $\alpha > 0$

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} = 1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \frac{1}{4^\alpha} + \dots$$

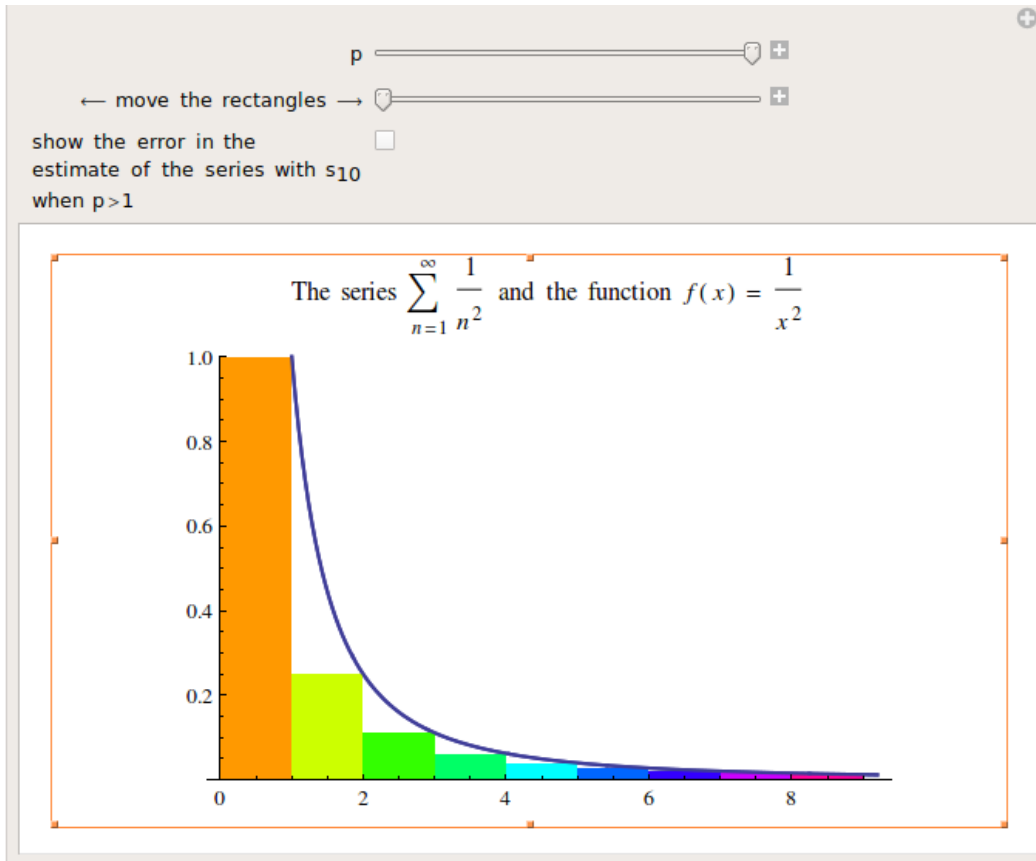
Die Reihenglieder sind monoton fallend, und wir können das entsprechende Integral auswerten

$$\lim_{n \rightarrow \infty} \int_{k_0}^n k^{-\alpha} dk = \lim_{n \rightarrow \infty} \begin{cases} k^{1-\alpha} \Big|_{k_0}^n \rightarrow \infty & \text{für } \alpha < 1 \\ \ln k \Big|_{k_0}^n \rightarrow \infty & \text{für } \alpha = 1 \\ k^{1-\alpha} \Big|_{k_0}^n \rightarrow k_0^{1-\alpha} < \infty & \text{für } \alpha > 1 \end{cases}$$

Wir erhalten das bereits bekannte Ergebnis, dass die Reihe für $0 < \alpha \leq 1$ divergiert, und für $\alpha > 1$ konvergiert.

CDF 5. Integraltest <https://moodle.uni-graz.at/course/view.php?id=95124>

In dieser CDF-Anwendung wird das Integralkriterium für Reihen der Form $\sum_{n=1}^{\infty} \frac{1}{k^p}$ graphisch veranschaulicht.



The area of the colored rectangles represents the sum of the series $\sum_{n=1}^{\infty} a_n$. With the rectangles in the starting (left) position, you can see that $\sum_{n=2}^{\infty} a_n \leq \int_1^{\infty} f(x) dx$. Slide the rectangles to the right to see that $\int_1^{\infty} f(x) dx \leq \sum_{n=1}^{\infty} a_n$. Thus the series and the integral converge or diverge together. For a convergent p -series (or any convergent series satisfying the criteria of the integral test), the inequality $s_N + \int_{N+1}^{\infty} f(x) dx \leq \sum_{n=1}^{\infty} a_n \leq s_N + \int_N^{\infty} f(x) dx$ holds for all N , where s_N is the N^{th} partial sum. An option in this Demonstration lets you see this fact using $N = 10$. The gray rectangles represent s_{10} , and the colored rectangles represent the "tail", $T_{10} = \sum_{n=11}^{\infty} a_n$.

1.5 Potenzreihen

In diesem Kapitel wollen wir nun Reihen betrachten, bei denen die Reihenglieder a_k von einer unabhängigen Variable x abhängen, also $a_k = a_k(x)$. Solche Objekte nennen wir dann *Funktionenreihen*. Innerhalb dieser Funktionenreihen bilden Potenzreihen eine spezielle und besonders wichtige Untergruppe. Sie sind dadurch gekennzeichnet, dass die unabhängige Variable x in aufsteigenden Potenzen vorkommt, und wir schreiben Potenzreihen in der allgemeinen Form

$$\sum_{k=0}^{\infty} a_k(x - x_0)^k = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \dots$$

Hierbei bezeichnen wir x_0 als den Entwicklungspunkt, und a_k als den k -ten Entwicklungskoeffizienten. Sehr häufig werden wir auch dem speziellen Fall begegnen, dass der Entwicklungspunkt gleich dem Ursprung ist, also $x_0 = 0$. In diesem Fall vereinfacht sich die Potenzreihendarstellung:

$$\sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

1.5.1 Konvergenzradius

Wählen wir einen bestimmten, festen Wert für die unabhängige Variable x , dann haben es wir wiederum mit einer bereits hinlänglich bekannten unendlichen Reihe zu tun, deren Konvergenz wir mit den in Kapitel 1.4.4 besprochenen Kriterien überprüfen können. Im allgemeinen wird das Konvergenzverhalten natürlich von der Wahl von x abhängen. Wir werden sehen, dass Potenzreihen in einem symmetrischen Intervall um den Entwicklungspunkt x_0 konvergieren, und wir bezeichnen dieses Intervall als den *Konvergenzradius* der Potenzreihe. Diesen Sachverhalt wollen wir anhand von einigen Beispielen verdeutlichen.

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k x^k = 1 + \frac{3}{4}x + \frac{9}{16}x^2 + \dots$$

Für welche Werte von x konvergiert diese Reihe? Wir können hier das *Wurzelkriterium* anwenden

$$\rho = \lim_{k \rightarrow \infty} \rho_k = \lim_{k \rightarrow \infty} \sqrt[k]{\left|\left(\frac{3}{4}\right)^k x^k\right|} = \lim_{k \rightarrow \infty} \left|\frac{3}{4}x\right| = \left|\frac{3}{4}x\right|$$

und finden, dass die Reihe für $\rho < 1$, also für $|x| < \frac{4}{3}$ konvergiert, und für $\rho > 1$, also für $|x| > \frac{4}{3}$ divergiert. Da das Wurzelkriterium für $\rho = 1$ keine Aussage zulässt, müssen wir die Ränder des Konvergenzintervalls, also $x = +\frac{4}{3}$ und $x = -\frac{4}{3}$ noch gesondert untersuchen. Für $x = +\frac{4}{3}$ erhalten wir

die Reihe $1 + 1 + 1 + \dots$, die divergiert, und für $x = -\frac{4}{3}$ die alternierende Reihe $1 - 1 + 1 - 1 + \dots$, die ebenfalls divergiert. Zusammenfassend finden wir somit, dass die Reihe in dem offenen Intervall $(-\frac{4}{3}, +\frac{4}{3})$ konvergiert.

Im nächsten Beispiel bestimmen wir den Konvergenzradius der folgenden Reihe:

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Hier wenden wir am besten das *Quotientenkriterium* an und finden

$$\rho = \lim_{k \rightarrow \infty} \rho_k = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k \rightarrow \infty} \left| \frac{x^{k+1} k!}{(k+1)! x^k} \right| = \lim_{k \rightarrow \infty} \left| \frac{x}{k+1} \right| = |x| \lim_{k \rightarrow \infty} \frac{1}{k+1} = 0.$$

Hier gilt, dass $\rho = 0 < 1 \forall x \in \mathbb{R}$, das heißt, die Reihe konvergiert für alle reelle Zahlen x und der Konvergenzradius ist ∞ .

Als drittes Beispiel untersuchen wir noch folgende Potenzreihe mit dem Entwicklungspunkt $x_0 = 3$

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{(x-3)^k}{2k} = \frac{x-3}{2} - \frac{(x-3)^2}{4} + \frac{(x-3)^3}{6} - \dots$$

Wir wenden wieder das *Wurzelkriterium* an und finden

$$\rho = \lim_{k \rightarrow \infty} \rho_k = \lim_{k \rightarrow \infty} \sqrt[k]{\left| (-1)^{k+1} \frac{(x-3)^k}{2k} \right|} = |x-3| \lim_{k \rightarrow \infty} \sqrt[k]{\frac{1}{2k}} = \frac{|x-3|}{\lim_{k \rightarrow \infty} \sqrt[k]{2} \cdot \lim_{k \rightarrow \infty} \sqrt[k]{k}} = |x-3|.$$

Die Reihe konvergiert für $\rho < 1$, also für $|x-3| < 1$, das heißt in dem Intervall $2 < x < 4$. Was passiert an den Rändern des Intervalls? Für $x = 2$ erhalten wir die Reihe $\frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{k}$, also die harmonische Reihe, die bekanntermaßen divergiert. Am oberen Ende des Konvergenzintervalls, $x = 4$, erhalten wir die alternierende harmonische Reihe $\frac{1}{2} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k}$, die nach dem Leibnizkriterium konvergent ist. Zusammenfassend gilt also für den Konvergenzbereich der Reihe: $2 < x \leq 4$.

Wir halten fest: Potenzreihen sind spezielle Funktionenreihen, in denen die unabhängige Variable x in Potenzen vorkommt. Die Menge der x , für die die Reihe konvergiert, nennen wir den Konvergenzbereich. Wir können den Konvergenzbereich bestimmen, indem wir die bereits bekannten Kriterien zur Bestimmung der Konvergenz von Reihen anwenden. Bei Potenzreihen ist dieser Konvergenzbereich symmetrisch um den Entwicklungspunkt der Reihe, und heißt deshalb auch Konvergenzradius. Innerhalb des Konvergenzbereichs, konvergiert die Potenzreihe gegen die Summenfunktion $f(x)$.

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \dots$$

Abhängig von der betrachteten Potenzreihe kann der Konvergenzradius 0, endlich oder unendlich sein.

1.5.2 Taylor-Reihe

In der Praxis stellt sich weniger oft die Frage, welche Summe eine Potenzreihe besitzt, als vielmehr die, wie für eine gegebene Funktion $f(x)$ die dazugehörige Potenzreihe bestimmt werden kann? Eine solche *Reihenentwicklung* ist in den meisten Fällen möglich, auf Gegenbeispiele wird zu einem späteren Zeitpunkt eingegangen. Bevor wir die allgemeine *Formel von Taylor* kennenlernen, mit der wir eine Funktion $f(x)$ in eine Potenzreihe (=Taylor-Reihe) entwickeln können, wollen wir anhand der Funktion $f(x) = \cos x$ die grundsätzliche Vorgehensweise erläutern. Das heißt, wir stehen vor der Aufgabe, die zunächst unbekanntem Entwicklungskoeffizienten a_0, a_1, a_2, \dots zu ermitteln

$$\cos x = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

Wir haben uns also für den Entwicklungspunkt $x_0 = 0$ entschieden. Setzen wir in der obigen Gleichung $x = x_0 = 0$, so erhalten wir $1 = a_0$. Um die weiteren Koeffizienten zu bestimmen, differenzieren wir beide Seiten der Gleichung schrittweise, und erhalten

$$\begin{array}{ll} \frac{d}{dx} : & -\sin x = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 \dots & x = 0 : & 0 = a_1 = 1!a_1 \\ \frac{d^2}{dx^2} : & -\cos x = 2a_2 + 6a_3x + 12a_4x^2 \dots & x = 0 : & -1 = 2a_2 = 2!a_2 \\ \frac{d^3}{dx^3} : & \sin x = 6a_3 + 24a_4x + \dots & x = 0 : & 0 = 6a_3 = 3!a_3 \\ \frac{d^4}{dx^4} : & \cos x = 24a_4 + \dots & x = 0 : & 1 = 24a_4 = 4!a_4 \end{array}$$

Hieraus können wir die allgemeine Form des k -ten Koeffizienten bereits ersehen, und erhalten für die Potenzreihenentwicklung der Funktion $\cos x$ um den Entwicklungspunkt $x_0 = 0$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

Wie sieht es um die Konvergenz der Reihe aus? Wir wenden das Quotientenkriterium an und erhalten

$$\rho = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k \rightarrow \infty} \left| \frac{x^{2k+2}}{(2k+2)!} \frac{(2k)!}{x^{2k}} \right| = \lim_{k \rightarrow \infty} \left| \frac{x^2}{(2k+1)(2k+2)} \right| = x^2 \lim_{k \rightarrow \infty} \frac{1}{(2k+1)(2k+2)} = 0.$$

Das heißt, die gefundene Taylor-Reihe für die Funktion $\cos x$ konvergiert auf ganz \mathbb{R} .

Wir wollen nun dasselbe Verfahren zur Ermittlung der Potenzreihe auf eine allgemeine Funktion $f(x)$

anwenden, wobei wir als Entwicklungspunkt $x = x_0$ wählen.

$$\begin{aligned}
 f(x) &= a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_k(x - x_0)^k + \dots \\
 f'(x) &= a_1 + 2a_2(x - x_0) + \dots + ka_k(x - x_0)^{k-1} + \dots \\
 f''(x) &= 2a_2 + \dots + k(k-1)a_k(x - x_0)^{k-2} + \dots \\
 &\vdots \\
 f^{(k)}(x) &= k! a_k + \dots
 \end{aligned}$$

Setzen wir in den obigen Gleichungen jeweils $x = x_0$, so erhalten wir

$$\begin{aligned}
 f(x_0) &= a_0 \\
 f'(x_0) &= a_1 \\
 f''(x_0) &= 2a_2 \\
 &\vdots \\
 f^{(k)}(x_0) &= k! a_k
 \end{aligned}$$

wobei wir mit $f^{(k)}(x_0)$ die k -te Ableitung der Funktion $f(x)$ an der Stelle x_0 bezeichnet haben:

$$f^{(k)}(x_0) \equiv \left. \frac{d^k f(x)}{dx^k} \right|_{x=x_0}.$$

Das heißt, wir erhalten für die *Taylor-Reihe* – oder auch Formel von Taylor – folgenden Ausdruck

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \frac{f'''(x_0)}{3!} (x - x_0)^3 + \dots \quad (1.12)$$

Für den Spezialfall, dass der Entwicklungspunkt gleich $x_0 = 0$ ist, erhalten wir die etwas einfachere Form der Taylor-Formel, die man auch als *MacLaurin-Reihe* bezeichnet:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!} x^2 + \frac{f'''(0)}{3!} x^3 + \dots \quad (1.13)$$

Wir betonen an dieser Stelle, dass die Taylor-Formel zunächst formal für jede Funktion $f(x)$, die ausreichend oft differenzierbar ist, definiert werden kann. Die Konvergenz der Reihe, das heißt das **Konvergenzgebiet, muss aber in jedem Fall überprüft werden**. Insbesondere gibt es Funktionen, deren formale Potenzreihe als Konvergenzgebiet nur den Entwicklungspunkt hat! Ein Beispiel für eine solche Funktion ist etwa $e^{-\frac{1}{x^2}}$. Die Funktion und alle ihre Ableitungen haben bei $x = 0$ den Wert 0, und die Potenzreihe lautet daher $e^{-\frac{1}{x^2}} = 0 + 0 + 0 + \dots$, wodurch der Konvergenzradius dieser Potenzreihe 0 ist. Nach diesem eher abschreckenden Gegenbeispiel wollen wir aber in den nächsten

Kapitel, einige sehr nützliche Eigenschaften von Potenzreihen aufzählen sowie die Taylor-Reihen und deren Konvergenzgebiete einiger wichtiger elementarer Funktionen angeben.

1.5.3 Eigenschaften von Potenzreihen

In ihrem Konvergenzgebiet verhalten sich Potenzreihen wie "gewöhnliche" Funktionen, was sich in folgenden Eigenschaften widerspiegelt.

1. **Eindeutigkeit.** Für eine Funktion $f(x)$ und den Entwicklungspunkt x_0 gibt es genau eine Potenzreihendarstellung.
2. **Gliedweise Addition, Subtraktion, Multiplikation, Division.** Potenzreihen können gliedweise addiert, subtrahiert oder multipliziert werden. Bei der Division darf die Nennerreihe keine Nullstellen am Entwicklungspunkt x_0 haben.
3. **Gliedweise Differenzier- und Integrierbarkeit.** Die gliedweise Ableitung (bzw. Integration) einer Potenzreihen konvergiert gegen die Ableitung (bzw. Integral) der ursprünglichen Reihe

Restglied. Wenn wir uns die Frage stellen, wie groß ist die Abweichung der n -ten Partialsumme $S_n(x)$ von der Summenfunktion $f(x)$ ist, so können wir das mithilfe des Restglieds $R_n(x)$ abschätzen.

$$R_n(x) = f(x) - S_n(x) = f(x) - \sum_{k=0}^n a_k (x - x_0)^k$$

Zunächst ist klar, dass die Reihe konvergiert, wenn das Restglied für $n \rightarrow \infty$ verschwindet. Für *alternierende* Reihen ist das Restglied $R_n(x)$ immer kleiner als das $(n + 1)$ -te Reihenglied, also

$$|R_n(x)| \leq |a_{n+1}(x - x_0)^{n+1}|$$

Für eine *allgemeine* Reihe lässt sich ebenfalls das Restglied abschätzen (Lagrangesches Restglied), und man kann zeigen, dass gilt:

$$|R_n(x)| \leq \left| \frac{(x - x_0)^{n+1}}{(n + 1)!} \right| \max |f^{(n+1)}(t)|$$

Hier muss der Maximalwert der $(n + 1)$ -ten Ableitung in dem Intervall $x \leq t \leq x_0$ bzw. $x_0 \leq t \leq x$ eingesetzt werden.

1.5.4 Beispiele für Taylor-Reihen

Bei den folgenden Beispielen handelt es sich jeweils um Taylor-Reihen um den Entwicklungspunkt $x_0 = 0$, also um MacLaurin-Reihen, die wir auch als Übung durch Anwendung von Gleichung 1.13 berechnen können.

$$\begin{aligned}
 e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots && \text{für } x \in \mathbb{R} \\
 \sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots && \text{für } x \in \mathbb{R} \\
 \cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots && \text{für } x \in \mathbb{R} \\
 \frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n &= 1 + x + x^2 + x^3 + \dots && \text{für } |x| < 1 \\
 \ln(1+x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n} &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots && \text{für } -1 < x \leq 1 \\
 (1+x)^p &= \sum_{n=0}^{\infty} \binom{p}{n} x^n &= 1 + px + \frac{p(p-1)}{2!} x^2 + \dots && \text{für } |x| < 1
 \end{aligned}$$

Das letzte Beispiel in der obigen Tabelle ist die sogenannte *Binomialreihe* und bedarf einer näheren Erläuterung. Ist der Parameter p eine positive ganze Zahl, so bricht die Reihe nach einer endlichen Anzahl an Gliedern – genauer nach p Gliedern – ab. Das sieht man beispielsweise, wenn wir für $p = 3$ setzen:

$$(1+x)^3 = \sum_{n=0}^{\infty} \binom{3}{n} x^n = 1 + 3x + \frac{3 \cdot 2}{2!} x^2 + \frac{3 \cdot 2 \cdot 1}{3!} x^3 + \underbrace{\frac{3 \cdot 2 \cdot 1 \cdot 0}{4!} x^4 + \dots}_{=0}$$

Ist p negativ oder nichtganzzahlig, so handelt es sich um eine unendliche Reihe, die dementsprechend nicht abbricht. Setzen wir etwa $p = -1$, so erhalten wir das schon bekannte Ergebnis

$$\begin{aligned}
 (1+x)^{-1} &= \sum_{n=0}^{\infty} \binom{-1}{n} x^n = 1 - x + \frac{(-1)(-2)}{2!} x^2 + \frac{(-1)(-2)(-3)}{3!} x^3 + \dots \\
 &= 1 - x + x^2 - x^3 + \dots
 \end{aligned}$$

Wenn wir wiederum $p = \frac{1}{2}$ setzen, erhalten wir folgende unendliche Reihe

$$\begin{aligned}
 (1+x)^{\frac{1}{2}} &= \sum_{n=0}^{\infty} \binom{1/2}{n} x^n = 1 + \frac{1}{2}x + \frac{(1/2)(-1/2)}{2!} x^2 + \frac{(1/2)(-1/2)(-3/2)}{3!} x^3 + \dots \\
 &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots
 \end{aligned}$$

Und zum Abschluss betrachten wir noch den sehr ähnlichen Fall $p = -\frac{1}{2}$:

$$\frac{1}{\sqrt{1+x}} = \sum_{n=0}^{\infty} \binom{-1/2}{n} x^n = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \dots$$

CDF 6. Taylorreihen <https://moodle.uni-graz.at/course/view.php?id=95124>
 In dieser CDF-Anwendung werden Taylorreihen für verschiedene elementare Funktionen bis zu einer vorgegebenen Ordnung und um einen wählbaren Entwicklungspunkt berechnet und graphisch dargestellt.

expansion point xpt +
 terms (highest degree) terms +
 x axis xmax +

f(x) =
 graph all polynomials showall
 show opts polynomial error neither

$f(x) = \sin(x)$

$T_7(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040}$

1.5.5 Anwendungsbeispiele

An dieser Stelle wollen wir exemplarisch einige Anwendungen von Potenzreihen bzw. Taylor-Reihen Entwicklungen diskutieren.

Numerik. Wie schon im Einführungskapitel 1.1 angedeutet, erlaubt etwa die Reihendarstellung einer Winkelfunktion die numerische Auswertung auf einem Taschenrechner oder Computer, der ja nur die Rechenoperationen Addition, Subtraktion, Multiplikation und Division durchführen kann. In

ähnlicher Weise können wir auch – ohne Zuhilfenahme eines Taschenrechners – Ausdrücke der Art $\sqrt[3]{1005}$ berechnen. Dazu schreiben wir

$$\sqrt[3]{1005} = 10(1 + 0.005)^{\frac{1}{3}} = 10(1 + x)^{\frac{1}{3}} = 10\left(1 + \frac{x}{3} - \frac{x^2}{9} + \dots\right) = 10(1 + 0.001\dot{6} - 0.000002\dot{7} + \dots)$$

Wie genau ist das Ergebnis, das wir so erhalten haben? Es handelt sich um eine alternierende Reihe, also ist das Restglied $R_n(x)$ einfach abzuschätzen. Für den Fall, dass wir Terme bis zur Ordnung $\mathcal{O}(x^2)$ berücksichtigen liefert die Fehlerabschätzung:

$$|R_2(x = 0.005)| \leq |a_3 x^3| = \frac{5}{81} 0.005^3 \simeq 2 \times 10^{-8}$$

In ähnlicher Weise können wir nun auch Logarithmen berechnen. Dazu halten wir zunächst fest, dass wir mit Reihendarstellung von $\ln(1 + x)$ nun auch endlich wissen, was die Summe der alternierenden harmonischen Reihe ist! Bisher wussten wir ja lediglich (Leibnizkriterium), dass sie konvergiert, aber nicht gegen welche Summe. Setzen wir nämlich in $\ln(1 + x)$ den Funktionswert $x = 1$, so erhalten wir die alternierende harmonische Reihe!

$$\ln 2 = \ln(1 + 1) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

Diese Reihe konvergiert sehr langsam! Das ist auch dadurch ersichtlich, weil sich der Punkt $x = 1$ genau am Rand des Konvergenzgebiets von $\ln(1 + x)$ befindet und die Konvergenz von Potenzreihen in Richtung Rand des Konvergenzgebiets generell langsamer wird. Hier etwa beträgt nach 10 Gliedern der alternierenden Reihe der Fehler $R_{10} \leq \frac{1}{11}$, auch nach 1000 Termen ist die Summe nur bis zur 3. Dezimalstelle genau, $R_{1000} \leq \frac{1}{1001}$. Haben wir allerdings einen genauen Wert für $\ln 2 = 0.69314718\dots$ gefunden, so können wir mit Hilfe der Reihendarstellung des Logarithmus, auch numerische Werte für beliebige Argumente erhalten, indem wir das Argument in Potenzen von 2 faktorisieren

$$\begin{aligned} \ln 47 &= \ln(32 \times 1.46875) = \ln 2^5 + \ln 1.46875 \\ &= 5 \ln 2 + \ln(1 + 0.46875) \approx 3.465736 + 0.46875 - \frac{1}{2} 0.46875^2 + \dots \end{aligned} \quad (1.14)$$

Kleine "Störungen". Sehr oft kennt man in der Physik zwar formal den exakten Ausdruck für ein bestimmtes physikalisches Problem, man kann aber die entsprechenden Gleichungen nicht lösen. Abhilfe kann hier oft die Tatsache schaffen, dass man den gefunden Ausdruck in eine Taylorreihe einer "kleinen" Größe – oder *Störung* – entwickelt. Um ein konkretes Beispiel zu geben, betrachten wir den Ausdruck für die *relativistische Energie* eines freien Teilchens mit der Ruhemasse m und der

Geschwindigkeit v

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} = mc^2 \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = mc^2 (1 - \varepsilon)^{-\frac{1}{2}}$$

Hierbei ist c die Lichtgeschwindigkeit, und wir haben die Größe $\varepsilon = \frac{v^2}{c^2}$ eingeführt, die für Geschwindigkeiten sehr viel kleiner als die Lichtgeschwindigkeit eine kleine Größe darstellt. Mit den Ergebnissen aus den vorangegangenen Kapiteln können wir nun den Klammerausdruck in Potenzen von ε entwickeln und erhalten

$$E \approx mc^2 \left(1 + \frac{\varepsilon}{2} + \frac{3\varepsilon^2}{8} + \dots\right) = mc^2 + \frac{1}{2}mv^2 + \frac{3mv^4}{8c^2} + \dots$$

Die Bedeutung des ersten Terms ist die Ruheenergie des Teilchens, der zweite Term ist die kinetische Energie des Teilchens, so wie sie aus der Newtonschen Mechanik bekannt ist, und die weiteren Terme sind relativistische Korrekturen für die kinetische Energie, die von der Massenzunahme bei hohen Geschwindigkeiten herrühren.

Lösung von Integralen. Auch diese Anwendung von Taylorreihen haben wir bereits im ersten Kapitel dieses Skriptums vorweggenommen. Sehr oft können wir keine analytische Form der Stammfunktion $F(x)$ einer Funktion $f(x)$ angeben. Kennen wir allerdings die Taylorreihe der Funktion $f(x)$, so können wir – innerhalb des Konvergenzgebiets – gliedweise integrieren, und so eine Reihendarstellung der Integralfunktion gewinnen. Betrachten wir etwa die Gaußsche Glockenkurve $f(x) = e^{-x^2}$. Um ihre Reihendarstellung zu finden müssen wir lediglich in der bekannten Taylorreihe von e^x das x durch $-x^2$ ersetzen und erhalten

$$f(x) = e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots$$

Die gliedweise Integration ergibt

$$F(x) = x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \dots$$

Damit können wir etwa folgendes bestimmte Integral auswerten

$$\int_0^1 f(x) dx = F(1) - F(0) = 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \dots \approx 0.742857$$

Das exakte Ergebnis würde hier den Wert 0.746824 liefern.

1.5.6 Regel von De L'Hospital

Die Potenzreihendarstellung von Funktionen erlaubt es auch Ausdrücke wie

$$f(x) = \frac{\sin x}{x}$$

zu behandeln, die an einzelnen Stellen, in diesem Beispiel bei $x = 0$, nicht definiert sind. Sowohl die Zähler- als auch die Nennerfunktion verschwindet im $\lim_{x \rightarrow 0}$, und wir erhalten den *unbestimmten Ausdruck* " $\frac{0}{0}$ ". Setzen wir die Reihendarstellung der Sinusfunktion ein, so können wir den Grenzwert ganz einfach bestimmen

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots}{x} = \lim_{x \rightarrow 0} \left(1 - \frac{x^2}{3!} + \frac{x^4}{5!} + \dots \right) = 1$$

Anhand dieses Beispiels können wir eine allgemeine Regel, die *Regel von De L'Hospital* ableiten, die es uns erlaubt die Grenzwerte von solchen unbestimmten Formen zu berechnen. Betrachten wir dazu den Quotient zweier Funktionen $f(x)$ und $g(x)$, der an einer Stelle a die unbestimmte Form " $\frac{0}{0}$ " aufweist. Das heißt, dass $\lim_{x \rightarrow a} f(x) = 0$ und $\lim_{x \rightarrow a} g(x) = 0$ gilt. Entwickeln wir nun $f(x)$ und $g(x)$ jeweils in eine Taylorreihe, so sehen wir

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f(a) + f'(a)(x-a) + \dots}{g(a) + g'(a)(x-a) + \dots} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)},$$

dass wir den Grenzwert dadurch bestimmen können, indem wir **Zähler und Nenner getrennt voneinander differenzieren**. Sollte der erhaltene Ausdruck an der Stelle a wieder unbestimmt sein, so wiederholen wir das Spiel und differenzieren Zähler und Nennerfunktion erneut. In unserem obigen Beispiel ergibt die Anwendung der Regel von de L'Hospital also

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1$$

Auf analoge Weise können wir auch folgende unbestimmte Form auswerten

$$\lim_{x \rightarrow 1} \frac{\ln x}{1-x} = \lim_{x \rightarrow 1} \frac{\frac{1}{x}}{-1} = - \lim_{x \rightarrow 1} \frac{1}{x} = -1$$

Die Regel von de L'Hospital lässt sich auch auf andere unbestimmte Formen übertragen. Man kann zeigen, dass Ausdrücke der Form " $\frac{\infty}{\infty}$ " in gleicher Weise behandelt werden können wie " $\frac{0}{0}$ ", also etwa

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x^c} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x}}{cx^{c-1}} = \frac{1}{c} \lim_{x \rightarrow \infty} \frac{1}{x^c} = 0 \text{ für } c > 0$$

Aus diesem Beispiel sehen wir, dass jede Potenz x^c für $c > 0$ im Limes $x \rightarrow \infty$ stärker ansteigt als der Logarithmus.

Auch andere unbestimmte Formen wie " $0 \cdot \infty$ ", " $\infty - \infty$ ", " 0^∞ ", " 1^∞ ", und " ∞^0 " können mithilfe der Regel von de L'Hospital behandelt werden, sie müssen allerdings vor ihrer Anwendung noch geeignet umgeformt werden:

$$\begin{aligned} 0 \cdot \infty : f \cdot g &\longrightarrow \frac{f}{g^{-1}} \\ \infty - \infty : f - g &\longrightarrow \frac{g^{-1} - f^{-1}}{(fg)^{-1}} \\ 0^\infty, 1^\infty, \infty^0 : f^g &\longrightarrow \exp(g \ln f) \end{aligned}$$

Wir demonstrieren diese "Rezepte" zur Anwendung der Regel von de L'Hospital anhand je eines Beispiels.

$$0 \cdot \infty : \lim_{x \rightarrow 0} \sqrt{x} \ln x = \lim_{x \rightarrow 0} \frac{\ln x}{x^{-\frac{1}{2}}} = \lim_{x \rightarrow 0} \frac{\frac{1}{x}}{-\frac{1}{2}x^{-\frac{3}{2}}} = -2 \lim_{x \rightarrow 0} x^{\frac{1}{2}} = 0.$$

$$\infty - \infty : \lim_{x \rightarrow 1} \left(\frac{x}{x-1} - \frac{1}{\ln x} \right) = \lim_{x \rightarrow 1} \frac{\ln x - \frac{x-1}{x}}{\frac{1}{x-1} - \frac{1}{\ln x}} = \lim_{x \rightarrow 1} \frac{x \ln x - x + 1}{(x-1) \ln x} = \lim_{x \rightarrow 1} \frac{\ln x}{\ln x + 1 - \frac{1}{x}} = \lim_{x \rightarrow 1} \frac{\frac{1}{x}}{\frac{1}{x} + \frac{1}{x^2}} = \frac{1}{2}.$$

$$\infty^0 : \lim_{x \rightarrow \infty} x^{\frac{1}{x}} = \lim_{x \rightarrow \infty} e^{\frac{1}{x} \ln x} = \exp \lim_{x \rightarrow \infty} \frac{1}{x} \ln x = e^0 = 1.$$

Kapitel 2

Komplexe Zahlen

Die komplexen Zahlen erweitern den Zahlenbereich der reellen Zahlen derart, dass quadratische Gleichungen in der Art $x^2 + 1 = 0$ lösbar werden. Dies gelingt durch Einführung einer neuen Zahl i mit der Eigenschaft $i^2 = -1$. Diese Zahl i wird als **imaginäre Einheit** bezeichnet. Komplexe Zahlen werden meist in der Form $z = x + iy$ dargestellt, wobei x und y reelle Zahlen. Auf die so dargestellten komplexen Zahlen lassen sich die üblichen Rechenregeln für reelle Zahlen anwenden, wobei i^2 stets durch -1 ersetzt werden kann und umgekehrt. Für die Menge der komplexen Zahlen wird das Symbol \mathbb{C} verwendet.

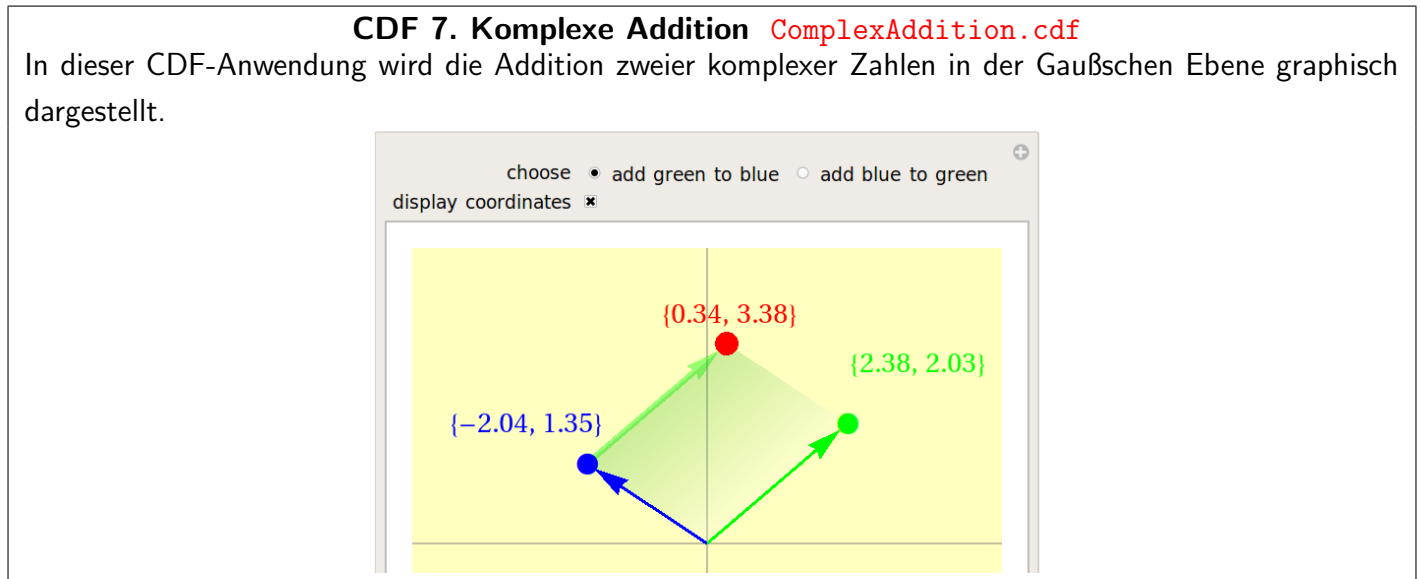
Die komplexen Zahlen bilden einen Körper mit einer Reihe nützlicher Eigenschaften, die sich auch in vielen Bereichen der Physik als äußerst nützlich erwiesen haben. Einer der Gründe für diese positiven Eigenschaften ist dass jede algebraische Gleichung vom Grad größer Null über den komplexen Zahlen eine Lösung besitzt (*Fundamentalsatzes der Algebra*). Ein weiterer Grund ist ein Zusammenhang zwischen *trigonometrischen Funktionen und der Exponentialfunktion*, der über die komplexen Zahlen hergestellt werden kann (siehe Kapitel 2.3.1). Ferner ist jede auf einer offenen Menge einmal komplex differenzierbare Funktion dort auch *beliebig oft differenzierbar* (anders als in der Mathematik der reellen Zahlen). Die Eigenschaften von Funktionen mit komplexen Argumenten sind Gegenstand der Funktionentheorie.

2.1 Die komplexe Ebene

Um eine komplexe Zahl $z = x + iy$ anzuschreiben, benötigen wir also zwei reelle Zahlen x und y , wobei wir x als den Realteil und y als den Imaginärteil von z bezeichnen, und wir schreiben

$$x = \operatorname{Re} z \text{ und } y = \operatorname{Im} z$$

Zwei komplexe Zahlen $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$ sind genau gleich, wenn ihre Realteile und ihre Imaginärteile gleich sind, also $x_1 = x_2$ und $y_1 = y_2$.



Für die graphische Darstellung einer komplexen Zahl $z = x + iy$ tragen wir analog zur analytischen Geometrie den Punkt mit den Koordinaten (x, y) in eine Ebene ein. Die x -Achse nennen wir die *reelle Achse*, während wir die y -Achse als die *imaginäre Achse* bezeichnen. Diese Darstellungsweise heißt "komplexe Ebene" oder "Gaußsche Ebene" oder auch "Arganddiagramm". Gleich wie in der analytischen Geometrie können wir für die Darstellung der komplexen Zahl $z = x + iy$ nicht nur *kartesische Koordinaten* verwenden, sondern wir können auch ebenso die *Polarkoordinaten* r und φ benutzen, um die Position von z in der komplexen Ebene festzulegen. Der Zusammenhang zwischen der kartesischen und der Polardarstellung ist dabei durch folgende Gleichungen festgelegt:

$$\begin{aligned} x &= r \cos \varphi \\ y &= r \sin \varphi \\ r &= \sqrt{x^2 + y^2} \\ \varphi &= \arctan \frac{y}{x} \end{aligned}$$

Dabei ist zu beachten, dass der Radius r , den wir auch als **Betrag der komplexen Zahl** z bezeichnen, eine positive reelle Zahl $r \geq 0$ ist, und der Winkel φ , den wir das **Argument von** z nennen, auf das Intervall $[0, 2\pi)$ beschränkt ist. Mit den oben genannten Relationen können wir nun eine komplexe Zahl z auch in folgender Form, in der Polardarstellung, angeben:

$$z = x + iy = r(\cos \varphi + i \sin \varphi)$$

Den **Betrag einer komplexen Zahl** $r = |z|$ interpretieren wir also als den euklidischen Abstand des Punktes (x, y) vom Ursprung. Mit dieser Festlegung erfüllt der Betrag einer komplexen Zahl $|z|$ auch die Eigenschaften einer Norm (vgl. Kapitel 1.2.2). Der Abstand zweier komplexer Zahlen z und w wird daher durch $d(z, w) = |z - w|$ berechnet. Die Menge der komplexen Zahlen \mathbb{C} zusammen mit der Norm $|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}$ bildet einen metrischen Raum.

Das **Argument einer komplexen Zahl** $\varphi = \arg z$ bestimmen wir als den Polarwinkel ausgehend von der positiven reellen Achse gezählt im Uhrzeigersinn. Bei der Berechnung des Polarwinkels φ über den $\arctan \frac{y}{x}$ müssen wir allerdings beachten, dass der resultierende Winkel im richtigen Quadranten liegt.

Ein Beispiel, um diesen Sachverhalt zu verdeutlichen:

$$z = -2 - 2i$$

Der Betrag ist also $r = \sqrt{(-2)^2 + (-2)^2} = \sqrt{8}$ und für das Argument erhalten wir mithilfe eines Taschenrechners $\varphi = \arctan \frac{-2}{-2} = \frac{\pi}{4}$ radians. Das ist aber nicht die Lösung, die wir suchen, da unsere Zahl nicht im ersten, sondern im dritten Quadranten liegt! Den tatsächlichen Wert des Arguments erhalten wir hier, indem wir zu φ noch den Wert π addieren, also $\varphi = \frac{5\pi}{4}$. Genauer gesagt ist es sogar noch so, dass wir jeden ganzzahlige Vielfache von 2π noch addieren können ohne die resultierende komplexe Zahl zu ändern. Das heißt, das Argument einer komplexen Zahl ist nicht eindeutig, sondern nur bis auf $2n\pi$ mit $n \in \mathbb{Z}$ festgelegt. In unserem Beispiel also $\varphi = \frac{5\pi}{4} + 2n\pi$.

Bevor wir im folgenden einige Beispiele zum Umgang mit komplexen Zahlen üben werden, führen wir noch den Begriff der zu z **komplex konjugierten** Zahl \bar{z} ein. Es gilt

$$\begin{aligned} z &= x + iy \\ \bar{z} &= x - iy \end{aligned}$$

Das heißt die komplex konjugierte Zahl \bar{z} , in der Physik schreibt man oft auch z^* , unterscheidet sich von z nur im Vorzeichen des Imaginärteils. Mithilfe von \bar{z} können wir nun auch den Betrag der komplexen Zahl in folgender Weise angeben:

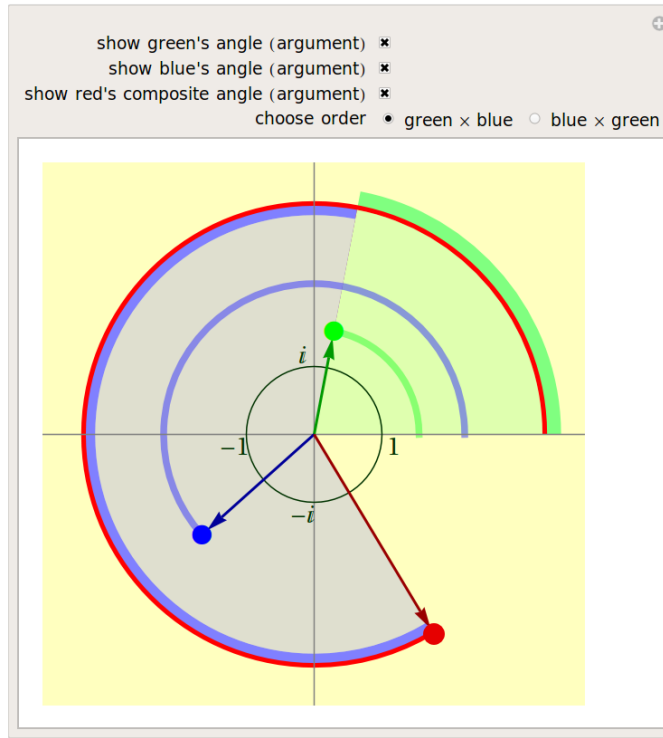
$$|z| = \sqrt{z\bar{z}}$$

Rechenbeispiele. Wir berechnen den Betrag $|z|$ sowie jeweils den Real- und Imaginärteil der folgenden komplexen Ausdrücke:

$$(a) z = \frac{3 + 2i}{1 - i} \quad (b) z = (1 + i)^2 \quad (c) z = \frac{i + 1}{i - 1}$$

CDF 8. Komplexe Multiplikation `ComplexMultiplication.cdf`

In dieser CDF-Anwendung wird die Multiplikation zweier komplexer Zahlen in der Gaußschen Ebene graphisch dargestellt.



Da komplexe Zahlen immer Paaren reeller Zahlen entsprechen, sind Gleichungen, in denen komplexe Zahlen vorkommen, auch immer Paare von reellen Gleichungen, je eine für den Real- und eine für den Imaginärteil. Lösungsmengen von Gleichungen mit komplexen Zahlen beschreiben daher Objekte in der Gaußschen Ebene. Betrachten wir etwa die Gleichung

$$z^2 = -i \iff (x + iy)^2 = -2i,$$

die in folgende zwei Gleichungen für den Real und Imaginärteil zerfällt:

$$x^2 - y^2 = 0 \quad \text{und} \quad xy = -1,$$

für die wir zwei Lösungen angeben können $x_1 = 1, y_1 = -1$ und $x_2 = -1, y_2 = 1$. das heißt die beiden komplexen Zahlen $z_1 = 1 - i$ und $z_2 = -1 + i$ lösen die obige Gleichung.

Als weitere Beispiele betrachten wir die folgenden Gleichungen bzw. Ungleichungen und berechnen die Lösungsmenge und deren Interpretation in der komplexen Zahlenebene:

$$(a) |z| \leq 2 \quad (b) |z + 2i| = 1 \quad (c) 0 \leq \arg(z) \leq \frac{\pi}{4} \quad (d) |z + i| + |z - i| = 6$$

2.2 Komplexe Reihen

Da die komplexen Zahlenmenge \mathbb{C} zusammen mit der Metrik $d(z, w) = |z - w|$ einen metrischen Raum bildet, können wir die Konvergenz von komplexen Folgen und Reihen analog zur Konvergenz von reellen Folgen und Reihen definieren. Wir müssen lediglich den reellen Absolutbetrag $|x|$ durch den komplexen Absolutbetrag $|z|$ ersetzen.

Eine Folge $\{z_n\}$ komplexer Zahlen heißt konvergent gegen $z \in \mathbb{C}$, wenn $|z_n - z| < \varepsilon$ für fast alle $n \in \mathbb{N}$ gilt. Geometrisch bedeutet dies, dass fast alle Folgenglieder in einer Kreisscheibe mit dem Radius ε um den Mittelpunkt z liegen.

Auch **komplexe Reihen**, also unendliche Reihen, deren Reihenglieder komplexe Zahlen darstellen, können mit denselben Methoden wie bei reellen Reihen untersucht werden, wenn man jeweils die Definition des Absolutbetrags einer komplexen Zahl verwendet. Das heißt Kriterien wie Wurzel- oder Quotientenkriterium (vgl. Kapitel 1.4.4) können auch für komplexe Reihen angewendet werden, um deren absolute Konvergenz bzw. Divergenz zu zeigen. Zudem können wir jede Folge oder jede Reihe auch in Real- und Imaginärteil zerlegen, also etwa $z_n = x_n + iy_n$, wobei x_n und y_n zwei reelle Folgen darstellen. Es gilt folgender

Satz. Sei $\{z_n\}$ eine komplexe Folge mit den Folgengliedern $z_n = x_n + iy_n$ und dem Grenzwert $z = x + iy$, dann gilt: Die Folge $\{z_n\}$ konvergiert genau dann gegen $z = x + iy$, wenn auch die beiden reellen Folgen $\{x_n\}$ und $\{y_n\}$, gebildet aus den Real- und Imaginärteilen von z_n gegen x bzw. y konvergieren:

$$\left(\lim_{n \rightarrow \infty} z_n = z\right) \Leftrightarrow \left(\lim_{n \rightarrow \infty} x_n = x\right) \wedge \left(\lim_{n \rightarrow \infty} y_n = y\right)$$

Beispiel. Wir betrachten die komplexe geometrische Reihe

$$z = \sum_{k=0}^{\infty} \left(\frac{1-i}{2}\right)^k = 1 + \frac{1-i}{2} + \left(\frac{1-i}{2}\right)^2 + \dots$$

und untersuchen ihre Konvergenz mithilfe des Quotientenkriteriums:

$$\rho = \lim_{n \rightarrow \infty} \rho_n = \lim_{n \rightarrow \infty} \left| \frac{(1-i)^{n+1}}{2^{n+1}} \frac{2^n}{(1-i)^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{1-i}{2} \right| = \left| \frac{1-i}{2} \right| = \frac{|1-i|}{|2|} = \frac{\sqrt{2}}{2} < 1.$$

Die Reihe ist also **absolut konvergent**. Mit der bekannten Reihenformel der geometrischen Reihe (siehe Gleichung 1.11, erhalten wir somit mit dem komplexen $q = \frac{1-i}{2}$

$$z = \frac{1}{1 - \frac{1-i}{2}} = \frac{2}{2 - 1 + i} = \frac{2}{1 + i} = \frac{2}{1 + i} \frac{1 - i}{1 - i} = 1 - i.$$

Komplexe Potenzreihen. Um die Definition von Potenzreihen, wie wir sie in Kapitel 1.5 kennengelernt haben, auf komplexe Potenzreihen zu erweitern, müssen wir einfach Potenzen von z ansetzen, also

$$\sum_{k=0}^{\infty} a_k z^k = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \dots$$

beziehungsweise für einen allgemeinen Entwicklungspunkt z_0 gilt dementsprechend

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k = a_0 + a_1 (z - z_0) + a_2 (z - z_0)^2 + a_3 (z - z_0)^3 + \dots$$

Wie wir durch die Bezeichnung *Konvergenzradius* bereits vorweggenommen haben, konvergieren Potenzreihen in einem kreisförmigen Gebiet der komplexen Ebene um den Entwicklungspunkt z_0 mit dem Radius R . Die Aussagen, die wir in Kapitel 1.5 über die Konvergenz von Potenzreihen gemacht haben, sind weiterhin gültig, wenn wir entsprechend den Begriff *Konvergenzintervall* durch *Konvergenzradius* ersetzen.

2.3 Funktionen komplexer Variablen

2.3.1 Exponentialfunktion und trigonometrische Funktionen

Wie können wir die komplexe Exponentialfunktion definieren?

$$f(z) = e^z$$

Die Antwort ist, dass wir die Taylor-Reihenentwicklung der Exponentialfunktion auf komplexe Potenzen erweitern, also

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \text{ für } z \in \mathbb{C}$$

Wir setzen zunächst in diese Reihendarstellung ein *rein imaginäres* Argument $z = i\varphi$ ein, also

$$e^{i\varphi} = 1 + i\varphi + \frac{(i\varphi)^2}{2!} + \frac{(i\varphi)^3}{3!} + \frac{(i\varphi)^4}{4!} + \dots$$

Wir erkennen, dass gerade Potenzen von $(i\varphi)$ eine reelle Zahl ergeben, während ungerade Potenzen von $(i\varphi)$ eine imaginäre Zahl liefern. Dem zufolge ordnen wir die Glieder der Potenzreihe so um, dass wir eine *reelle* und eine *imaginäre* Reihe erhalten:

$$e^{i\varphi} = \left(1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} + \dots\right) + \left(\frac{\varphi}{1!} - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} + \dots\right),$$

die wir als Reihendarstellung von $\cos \varphi$ und $\sin \varphi$ identifizieren! (vgl. Kapitel 1.5.4)

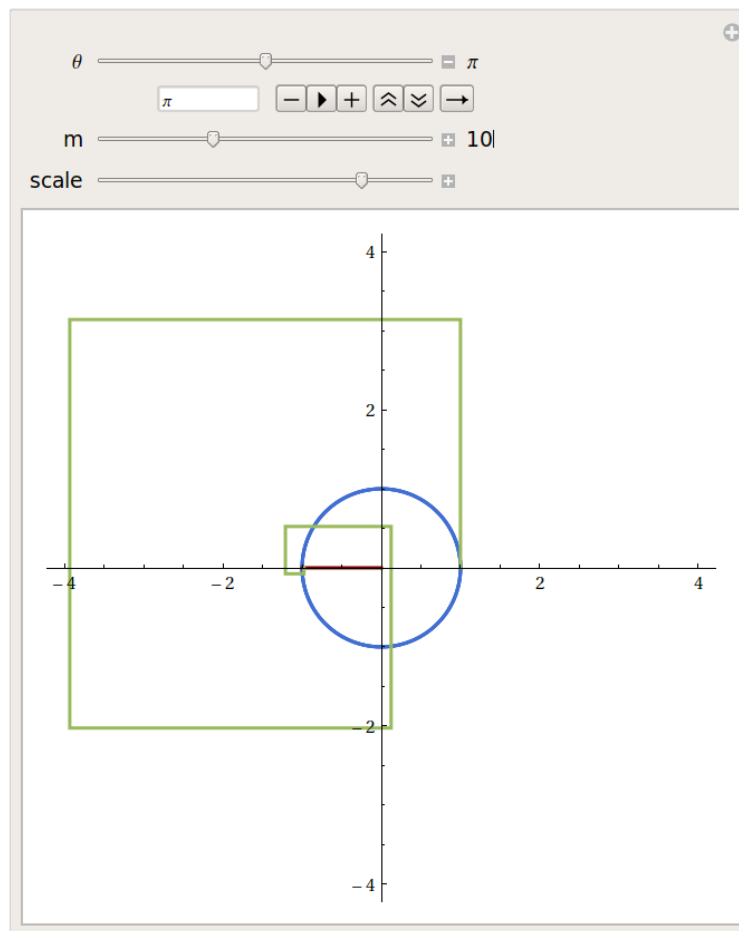
$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad (2.1)$$

Diese Beziehung, die einen Zusammenhang zwischen der Exponentialfunktion und den trigonometrischen Funktionen \cos und \sin herstellt, wird auch als **Eulersche Formel** bezeichnet. Setzen wir speziell $\varphi = \pi$, so erhalten wir eine berühmtesten – und zugleich ästhetischsten – Gleichungen der Mathematik, die **Eulersche Identität**¹

$$e^{i\pi} + 1 = 0 \quad (2.2)$$

CDF 9. Euler-Gleichung `EulersIdentity.cdf`

Diese CDF-Anwendung erlaubt die Darstellung Euler-Identität $e^{i\pi} = -1$ mit Hilfe der Reihendarstellung in der komplexen Ebene.



¹Die Eulersche Identität stellt einen Zusammenhang zwischen fünf der bedeutendsten mathematischen Konstanten her: Der Eulerschen Zahl e , der Kreiszahl π , der imaginären Einheit i , sowie den neutralen Elementen der Multiplikation 1 und der Addition 0. Darüber hinaus kommen auch die drei Rechenarten Addition, Multiplikation und Potenzrechnung jeweils genau einmal vor.

Mithilfe der Eulerschen Formel können wir nun eine komplexe Zahl auch in einer neuen Schreibweise angeben:

$$z = x + iy = r(\cos \varphi + i \sin \varphi) = re^{i\varphi}$$

Diese Schreibweise $z = re^{i\varphi}$ erweist sich als besonders nützlich, um eine Vielzahl von Rechenoperation zu vereinfachen. Dazu einige Beispiele.

Potenzen komplexer Zahlen lassen sich sehr effizient auswerten

$$(1 - i)^8 = \left(\sqrt{2}e^{i\frac{7\pi}{4}}\right)^8 = 2^{\frac{1}{2} \cdot 8} e^{i\frac{7\pi}{4} \cdot 8} = 16e^{i14\pi} = 16(\cos 14\pi + i \sin 14\pi) = 16.$$

Ganz allgemein gilt also die **Formel von De Moivre**

$$z^n = (re^{i\varphi})^n = r^n e^{in\varphi},$$

beziehungsweise wir können Produkte oder Quotienten zweier komplexer Zahlen ebenfalls einfach auswerten:

$$z_1 z_2 = r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)} \quad \text{und} \quad \frac{z_1}{z_2} = \frac{r_1 e^{i\varphi_1}}{r_2 e^{i\varphi_2}} = \frac{r_1}{r_2} e^{i(\varphi_1 - \varphi_2)}$$

Beispielsweise

$$\frac{(1 - i)^2}{1 + i} = \frac{(\sqrt{2}e^{-i\frac{\pi}{4}})^2}{\sqrt{2}e^{i\frac{\pi}{4}}} = \frac{2e^{-i\frac{\pi}{2}}}{\sqrt{2}e^{i\frac{\pi}{4}}} = \sqrt{2}e^{-i\frac{3\pi}{4}} = -1 - i$$

Winkeladditionstheoreme lassen sich auch mit Hilfe der Eulerschen Formel relativ einfach ableiten

$$\begin{aligned} e^{i(\varphi_1 + \varphi_2)} &= e^{i\varphi_1} e^{i\varphi_2} \\ \cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2) &= (\cos \varphi_1 + i \sin \varphi_1)(\cos \varphi_2 + i \sin \varphi_2) \\ \Rightarrow \cos(\varphi_1 + \varphi_2) &= \cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 \\ \Rightarrow \sin(\varphi_1 + \varphi_2) &= \sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2 \end{aligned}$$

(2.3)

Ganz allgemein können wir mithilfe der Beziehung

$$e^{z_1 + z_2} = e^{z_1} e^{z_2}$$

die wir im übrigen durch Einsetzen in die entsprechenden Taylor-Reihe beweisen können, viele Ausdrücke berechnen. Zunächst halten wir fest, dass die Exponentialfunktion mit komplexem Argument

$z = x + iy$ mithilfe der obigen Beziehung geschrieben werden kann als ein Produkt einer Exponentialfunktion mit reellem Argument und Winkelfunktionen mit dem Imaginärteil y als Argument:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y).$$

Wir können nun auch die Winkelfunktionen für komplexe Argumente angeben, indem wir die Ausdrücke für e^{iz} und e^{-iz} geeignet kombinieren.

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

$$\tan z = \frac{1}{i} \frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}}, \quad \cot z = i \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}}$$

Wir können damit etwa ausrechnen, was der $\sin i$ ergibt:

$$\sin i = \frac{1}{2i}(e^{-1} - e^{+1}) = -\frac{i}{2} \left(\frac{1}{e} - e \right),$$

Als Übung, können wir mit den obigen Definition auch folgende Identitäten herleiten:

$$\sin iy = i \sinh y \quad \text{und} \quad \cos iy = \cosh y$$

Wir können nun auch beliebige Ausdrücke der Form $z_1^{z_2}$ auswerten und erhalten das vielleicht überraschende Ergebnis, dass i^i eine reelle Zahl ist:

$$i^i = \left(e^{i\frac{\pi}{2}} \right)^i = e^{-\frac{\pi}{2}} = 0.20788 \dots$$

Hierbei gilt es allerdings zu beachten, dass das nur eine von unendlich vielen möglichen Werten von i^i ist! Warum das so ist, sehen wir, wenn wir berücksichtigen, dass das Argument einer komplexen Zahl nur bis auf ein ganzzahliges Vielfaches von 2π festgelegt ist. Die obige Rechnung lautet dann:

$$i^i = \left(e^{i\left(\frac{\pi}{2} + 2n\pi\right)} \right)^i = e^{-\left(\frac{\pi}{2} + 2n\pi\right)}$$

Diese Vieldeutigkeit ergibt sich aus der Berechnung einer Potenz mit nicht ganzzahligem Exponenten, also einer *Wurzel*, welche im folgenden Kapitel 2.3.2 näher besprochen wird.

Anwendung: Wellendarstellung. Zum Abschluss dieses Kapitels erwähnen wir noch eine wichtige Anwendung der komplexen Exponentialfunktion zur Darstellung von Wellenphänomenen in der Physik. Eine ebene Welle, die sich in die positive x -Richtung ausbreitet, wird durch Angabe der

Amplitude a_0 , der Kreisfrequenz ω , der Wellenzahl k und des Phasenwinkel φ charakterisiert:

$$a(x, t) = a_0 \cos(kx - \omega t + \varphi)$$

Die Phasengeschwindigkeit c stellt dabei einen Zusammenhang zwischen Kreisfrequenz und Wellenzahl her (*Dispersionsbeziehung*)

$$\omega = ck.$$

Mithilfe der Exponentialdarstellung können wir die Welle also auch anschreiben als

$$a(x, t) = \operatorname{Re} [a_0 e^{i(kx - \omega t + \varphi)}] = \operatorname{Re} [A e^{i(kx - \omega t)}],$$

wobei wir die komplexe Amplitude $A = a_0 e^{i\varphi}$ eingeführt haben. Wollen wir nun etwa die Superposition von zwei Wellen mit gleicher Frequenz berechnen, so müssen wir lediglich die komplexen Amplituden addieren

$$a_1(x, t) + a_2(x, t) = A_1 e^{i(kx - \omega t)} + A_2 e^{i(kx - \omega t)} = (A_1 + A_2) e^{i(kx - \omega t)},$$

und am Ende der Rechnung den Realteil des Ausdrucks berechnen, um auf die physikalisch relevante Größe zurückzukommen. Ein weiterer Vorteil dieser Exponentialnotation liegt in der Tatsache, dass ohne Einführung eines neuen Parameters auch abklingende (gedämpfte) Schwingungen oder Wellen beschrieben werden können. Dies kann erreicht werden, indem die Wellenzahl k auch einen positiven, nicht verschwindenden Imaginärteil erhält, der die Rolle des Dämpfungskoeffizienten übernimmt, $\gamma = \operatorname{Im}(k)$.

$$a(x, t) = A e^{i(kx - \omega t)} = A e^{i[(\operatorname{Re}(k) + i\operatorname{Im}(k))x - \omega t]} = A e^{-\gamma x} e^{i[\operatorname{Re}(k)x - \omega t]}$$

CDF 10. Komplexe Reihen [InfiniteSeriesExplorer.cdf](#)

Diese CDF-Anwendung erlaubt die Darstellung von verschiedenen unendlichen komplexen Reihen und die Angabe der dazugehörigen komplexen Funktion.

The screenshot shows the InfiniteSeriesExplorer interface with the following settings:

- range of n: 0, 1, 2, ...
- variable: x
- coefficient: $1/n!$

The displayed series is:

$$\sum_{n=0, 1, 2, \dots} \frac{(ix)^n}{n!} =$$

$$1 + ix - \frac{x^2}{2} - \frac{ix^3}{6} + \frac{x^4}{24} + \dots =$$

$$e^{ix}$$

2.3.2 Wurzeln

Mithilfe der Exponentialdarstellung können wir auch die Mehrdeutigkeit von Wurzeln von komplexen Zahlen besser verstehen. Wir betrachten also die Lösung der folgenden komplexen Gleichung

$$z = w^n,$$

wobei z und w komplexe Zahlen, und n eine natürlich Zahl darstellt. Lösen wir diese Gleichung nach w auf, so erhalten wir die n -te Wurzel aus z .

$$w = z^{\frac{1}{n}} = \sqrt[n]{z},$$

Betrachten wir ein Beispiel: was ist die 4-te Wurzel aus $z = 1$? Offensichtlich könne wir 4 Wurzeln angeben, nämlich $w_1 = 1$, $w_2 = -1$, $w_3 = i$, und $w_4 = -i$! Drücken wir "1" als $1 = e^{i(2k\pi)}$ aus, wobei wir die Vieldeutigkeit des komplexen Arguments berücksichtigt haben mit $k \in \mathbb{Z}$, so erhalten wir

$$\sqrt[4]{1} = [e^{i(2k\pi)}]^{\frac{1}{4}} = e^{i\frac{k\pi}{2}} = \begin{cases} e^{i\frac{0\pi}{2}} = 1 \\ e^{i\frac{1\pi}{2}} = i \\ e^{i\frac{2\pi}{2}} = -1 \\ e^{i\frac{3\pi}{2}} = -i \\ \vdots \end{cases}$$

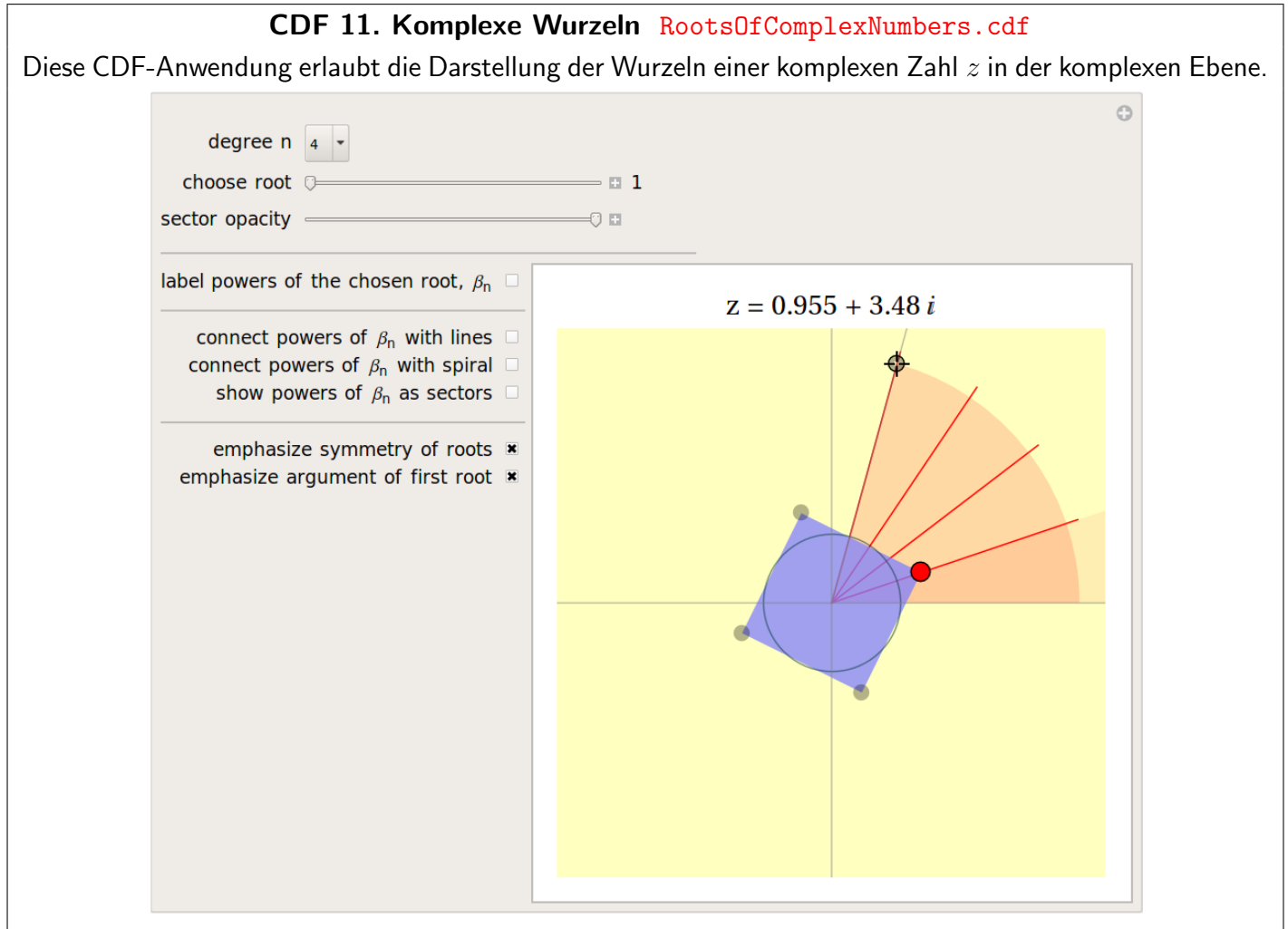
Wir erhalten also die schon oben erwähnten vier Wurzeln $1, i, -1, -i$. Ganz allgemein kann man sich leicht überlegen, dass es genau n Einheitswurzeln $\sqrt[n]{1}$ gibt, die am Rand eines Kreises in der komplexen Ebene mit dem Radius 1 liegen. Wollen wir die Wurzel aus einer beliebigen komplexen Zahl z ziehen, so schreiben wir z zunächst in Exponentialdarstellung an und gehen analog zum obigen Beispiel vor:

$$\sqrt[n]{z} = [re^{i(\varphi+2k\pi)}]^{\frac{1}{n}} = \sqrt[n]{r}e^{i(\frac{\varphi}{n}+\frac{2k\pi}{n})} \text{ für } k = 0, 1, \dots, (n-1)$$

Betrachten wir zwei Beispiele und berechnen die 3. Wurzel aus $z = 1 + i$ und die 4. Wurzel aus $z = i$

$$\sqrt[3]{1+i} = [\sqrt{2}e^{i(\frac{\pi}{4}+2k\pi)}]^{\frac{1}{3}} = \sqrt[3]{2}e^{i(\frac{\pi}{12}+\frac{2k\pi}{3})} = \begin{cases} \sqrt[3]{2}e^{i\frac{\pi}{12}} = 1.08422 + 0.290515i \\ \sqrt[3]{2}e^{i\frac{3\pi}{4}} = -0.793701 + 0.793701i \\ \sqrt[3]{2}e^{i\frac{17\pi}{12}} = -0.290515 - 1.08422i \end{cases}$$

$$\sqrt[4]{i} = [e^{i(\frac{\pi}{2}+2k\pi)}]^{\frac{1}{4}} = e^{i(\frac{\pi}{8}+\frac{k\pi}{2})} = \begin{cases} e^{i\frac{\pi}{8}} = 0.92388 + 0.382683i \\ e^{i\frac{5\pi}{8}} = -0.382683 + 0.92388i \\ e^{i\frac{9\pi}{8}} = -0.92388 - 0.382683i \\ e^{i\frac{13\pi}{8}} = 0.382683 - 0.92388i \end{cases}$$



2.3.3 Logarithmus

Genauso wie die Wurzel die Umkehrfunktion der Potenzen darstellt, so ist der Logarithmus die Umkehrfunktion zur Exponentialfunktion

$$z = e^w \iff w = \ln z.$$

Verallgemeinert man den Logarithmus auch auf komplexe Zahlen, so ergeben sich wieder Vieldeutigkeiten, die von der Vieldeutigkeit der komplexen Phase von z herrühren

$$w = \ln z = \ln [re^{i(\varphi+2k\pi)}] = \ln r + i\varphi + 2k\pi i, \quad k = 0, \pm 1, \pm 2, \dots$$

Wir sehen also, dass der Logarithmus einer komplexen Zahl z unendlich viele Lösungen hat, die sich im Imaginärteil um ganzzahlige Vielfache von 2π unterscheiden. Deshalb beschränkt man sich bei

der Angabe der Logarithmus in der Praxis sehr oft auf den *Hauptwert des Logarithmus*, der wie folgt definiert ist

$$\ln z = \ln r + i\varphi.$$

Damit können wir beispielsweise auch den Logarithmus einer negative Zahl berechnen

$$\ln(-e) = \ln [e \cdot e^{i\pi}] = 1 + i\pi$$

2.3.4 Überblick über einige elementare komplexwertige Funktionen

Wurzelfunktion

$$\begin{aligned}\sqrt{z} &= (re^{i\varphi})^{\frac{1}{2}} = \sqrt{r}e^{i(\frac{\varphi}{2}+k\pi)} \\ \operatorname{Re}\sqrt{z} &= \sqrt{r} \cos \frac{\varphi}{2} \\ \operatorname{Im}\sqrt{z} &= \sqrt{r} \sin \frac{\varphi}{2} \\ |\sqrt{z}| &= \sqrt{r}\end{aligned}$$

Exponentialfunktion

$$\begin{aligned}e^z &= e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y) \\ \operatorname{Re} e^z &= e^x \cos y \\ \operatorname{Im} e^z &= e^x \sin y \\ |e^z| &= e^x\end{aligned}$$

Cosinusfunktion

$$\begin{aligned}\cos z &= \frac{1}{2} [e^{iz} + e^{-iz}] = \frac{1}{2} [e^{i(x+iy)} + e^{-i(x+iy)}] = \cos x \cosh y - i \sin x \sinh y \\ \operatorname{Re} \cos z &= \cos x \cosh y \\ \operatorname{Im} \cos z &= -\sin x \sinh y \\ |\cos z| &= \sqrt{\cos^2 x \cosh^2 y + \sin^2 x \sinh^2 y}\end{aligned}$$

Sinusfunktion

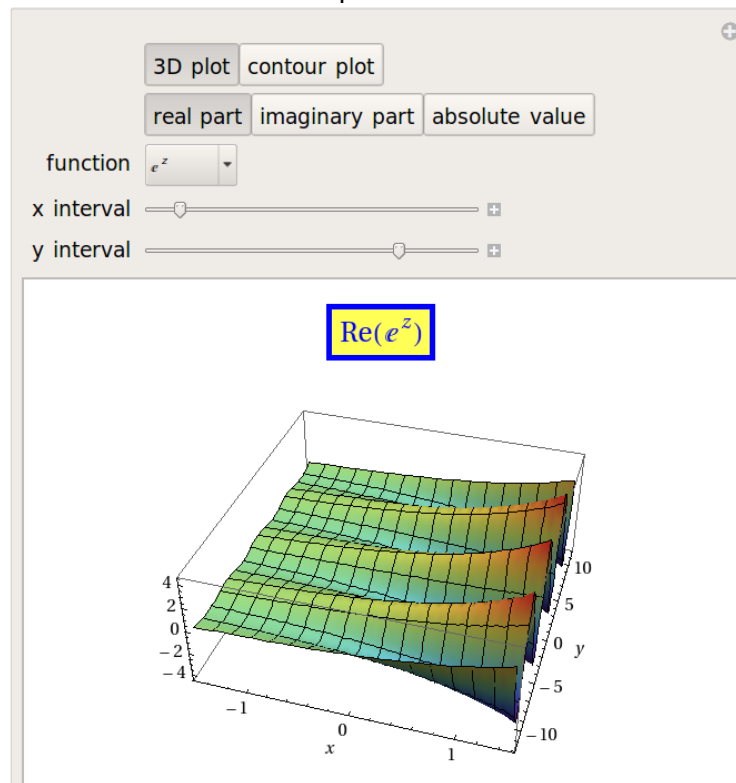
$$\begin{aligned}\sin z &= \frac{1}{2i} [e^{iz} - e^{-iz}] = \frac{1}{2i} [e^{i(x+iy)} - e^{-i(x+iy)}] = \sin x \cosh y + i \cos x \sinh y \\ \operatorname{Re} \cos z &= \sin x \cosh y \\ \operatorname{Im} \cos z &= \cos x \sinh y \\ |\cos z| &= \sqrt{\sin^2 x \cosh^2 y + \cos^2 x \sinh^2 y}\end{aligned}$$

Natürlicher Logarithmus

$$\begin{aligned}\ln z &= \ln [re^{i(\varphi+2k\pi)}] = \ln r + i\varphi + 2ik\pi \\ \operatorname{Re} \sqrt{z} &= \ln r \\ \operatorname{Im} \sqrt{z} &= \varphi + 2k\pi \\ |\sqrt{z}| &= \sqrt{(\ln r)^2 + \varphi^2}\end{aligned}$$

CDF 12. Komplexe Funktionen [ComplexFunctions.cdf](#)

Diese CDF-Anwendung erlaubt die Darstellung von Real-, Imaginärteil bzw. des Absolutbetrages von verschiedenen elementaren Funktionen über der komplexen Ebene.



2.3.5 Umkehrfunktionen der trigonometrischen Funktionen

Suchen wir etwa nach der Umkehrfunktion der Sinusfunktion

$$z = \sin w \iff w = \arcsin z,$$

so finden wir unter Zuhilfenahme der Exponentialdarstellung der Sinusfunktion und Auflösung der quadratischen Gleichung in (e^{iw})

$$\begin{aligned} z &= \frac{1}{2i}(e^{iw} - e^{-iw}) \\ &\vdots \\ \arcsin z = w &= \frac{1}{i} \ln \left[iz \pm \sqrt{1 - z^2} \right] + 2k\pi, k \in \mathbb{Z} \end{aligned}$$

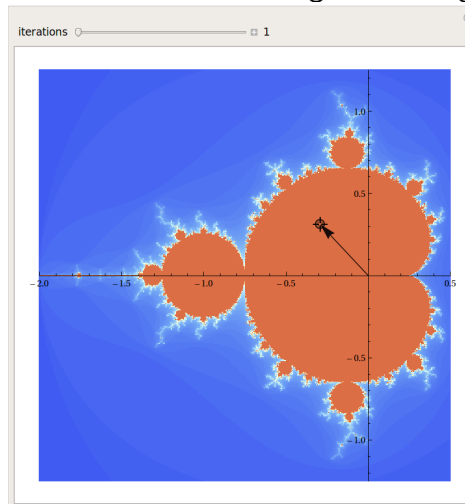
Eine analoge Rechnung für $w = \arccos z$ führt auf

$$\arccos z = \frac{1}{i} \ln \left[z \pm \sqrt{z^2 - 1} \right] + 2k\pi, k \in \mathbb{Z}$$

Mit diesen Darstellungen der Umkehrfunktionen können wir den Wert der Funktionen Arcussinus und Arcuscosinus für beliebige komplexe Zahlen angeben.

CDF 13. Mandelbrotfolge <https://moodle.uni-graz.at/course/view.php?id=95124>

Diese CDF-Anwendung stellt die sogenannte Mandelbrot-Menge \mathbb{M} , eine Teilmenge der komplexen Zahlen \mathbb{C} dar. \mathbb{M} ist die Menge aller komplexen Zahlen c , für welche die rekursiv definierte Folge $z_{n+1} = z_n^2 + c$ mit dem Anfangsglied $z_0 = 0$ beschränkt bleibt. Die grafische Darstellung dieser Menge erfolgt in der komplexen Ebene. Die Punkte der Menge \mathbb{M} werden hier braun dargestellt und der Rest farbig, wobei die Farbe eines Punktes außerhalb von \mathbb{M} den Grad der Divergenz der zugehörigen Folge widerspiegelt.



Kapitel 3

Lineare Algebra

Die lineare Algebra (auch Vektoralgebra) ist ein wichtiges Teilgebiet der Mathematik, das sich mit Vektorräumen und linearen Abbildungen zwischen diesen beschäftigt. Dies schließt insbesondere auch die Betrachtung von linearen Gleichungssystemen und Matrizen mit ein. Will man Lineare Algebra "mathematisch exakt" betreiben, so müssen zunächst Begriffe wie Vektorräume, lineare Abbildung etc. eingeführt werden (siehe etwa Klaus Jänich, Lineare Algebra, Springer-Lehrbuch). Wir wollen uns jedoch hier der linearen Algebra auf einem sehr pragmatischen Weg nähern, nämlich über die Lösung von linearen Gleichungssystemen. Auch die weiteren Abschnitte des Kapitels 3 entsprechen weitestgehend der Darstellung in Kapitel 3 des Lehrbuchs Lang und Pucker, Mathematische Methoden in der Physik.

3.1 Lineare Gleichungssysteme

Ein lineares Gleichungssystem (kurz LGS) ist in der linearen Algebra eine Menge linearer Gleichungen mit einer oder mehreren Unbekannten, die alle gleichzeitig erfüllt sein sollen. Ein entsprechendes System für zwei Unbekannte x und y sieht beispielsweise wie folgt aus:

$$\begin{aligned}x + 3y &= 6 \\2x - y &= 5.\end{aligned}$$

Wir wissen bereits aus der Schule, wie wir ein solches System aus zwei Gleichungen lösen. Wir können etwa die erste Gleichung auf $x = 6 - 3y$ umformen, und dann in die zweite Gleichung einsetzen, was auf $y = 1$ führt. Einsetzen dieses Werts für y ergibt dann $x = 3$. Geometrisch können wir die beiden Gleichungen dieses LGS jeweils als Geraden in \mathbb{R}^2 interpretieren. Der Schnittpunkt dieser beiden Geraden ist die Lösung des LGS, hier der Punkt $(x, y) = (3, 1)$. Dieser geometrischen Interpretation

folgend ist es klar, dass es für ein LGS in \mathbb{R}^2 grundsätzlich drei Lösungsarten gibt: (1) Die beiden Geraden sind nicht parallel, wonach es genau eine Lösung (Schnittpunkt) gibt, (2) die beiden Geraden sind parallel und gehen nicht durch denselben Punkt, dann besitzt das LGS keine Lösung, und (3) die beiden Geraden sind deckungsgleich, dann sind alle Punkte der Gerade Lösung des LGS.

Allgemein können wir ein Gleichungssystem in den zwei Unbekannten $x_1 \equiv x$ und $x_2 \equiv y$ mit allgemeinen Koeffizienten wie folgt anschreiben

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

wobei $a_{11}, a_{12}, a_{21}, a_{22}$ bzw. b_1, b_2 beliebige reelle Zahlen darstellen sollen. Wir lösen dieses LGS wieder, indem wir aus der ersten Gleichung beispielsweise x_2 ausrechnen

$$x_2 = \frac{1}{a_{12}} (b_1 - a_{11}x_1),$$

und dieses Ergebnis dann in die zweite Gleichung einsetzen. Dies führt dann auf

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}},$$

wodurch wir durch Rückeinsetzen auch x_2 erhalten

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}.$$

Für die hier auftretende Kombination der Koeffizienten führt man den Begriff der *Determinante* ein

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \equiv a_{11}a_{22} - a_{21}a_{12}.$$

Somit können wir uns Ergebnis für x_1 und x_2 auch folgendermaßen anschreiben

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \quad (3.1)$$

Dieses Ergebnis ist ein Spezialfall der *Cramerschen Regel* für den \mathbb{R}^2 . Zuvor beschäftigen wir uns jedoch mit den Eigenschaften von Determinaten und erweitern deren Definition von 2×2 Zahlen auf $n \times n$ Zahlen.

3.1.1 Determinanten

In der linearen Algebra ist die Determinante eine spezielle Funktion, die einer quadratischen $n \times n$ Matrix (eine Anordnung von Zahlen in Zeilen und Spalten) einen Skalar (eine Zahl) zuordnet. Mit Hilfe von Determinanten kann man beispielsweise feststellen, ob ein lineares Gleichungssystem eindeutig lösbar ist, und kann die Lösung mit Hilfe der *Cramerschen Regel* explizit angeben. Wie wir anhand von Gleichung 3.1 sehen, ist das Gleichungssystem genau dann eindeutig lösbar, wenn die Determinante der Koeffizientenmatrix ungleich null ist. Verschwindet die Determinante, hat das LGS entweder keine Lösung (parallele Geraden) oder unendlich viele Lösungen (die beiden Gleichungen beschreiben ein und dieselbe Gerade).

Wie wir die Determinante einer 2×2 Matrix berechnen (man sagt: 2-Determinante) wissen wir bereits. Folgende Schreibweisen sind dafür gebräuchlich:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \equiv a_{11}a_{22} - a_{21}a_{12}. \quad (3.2)$$

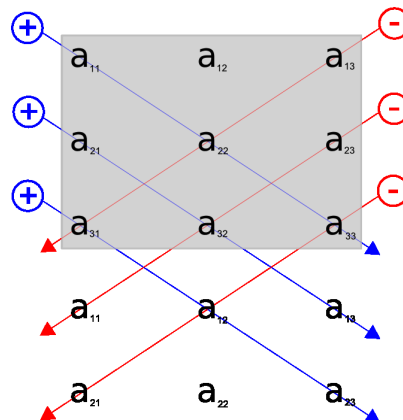
Die 1-Determinante ist ganz einfach das Element selbst, also

$$\det(a) = a, \quad (3.3)$$

während die 3-Determinante nach der sogenannten Regel von Sarrus berechnet werden kann (benannt nach dem französischen Mathematiker Pierre Sarrus)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \equiv +a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21} \quad (3.4)$$

Mnemotechnisch ist folgendes Schema hilfreich (Bildquelle: Pajs, cs.wikipedia)



Dazu hängt man die ersten beiden Zeilen unten an die Matrix an, und bildet dann nach dem Muster in der oben stehenden Abbildung die durch die Striche angedeuteten Produkte, wobei die 3 Produkte von links oben nach rechts unten ein positives Vorzeichen, und die drei Produkte von links unten nach rechts oben ein negatives Vorzeichen erhalten.

CDF 14. Regel von Sarrus [33DeterminantsUsingDiagonals.cdf](#)
 Interaktive Demonstration der Regel von Sarrus.

new numbers

$$\begin{vmatrix} -7 & 3 & 1 \\ -4 & 2 & -1 \\ -2 & -1 & 1 \end{vmatrix} = -14 + 6 + 4 - (-4 + -7 + -12) = 19$$

Für $n > 3$ ist die n -Determinante nach der *Leibniz'schen Formel* wie folgt definiert:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdots & \cdot \\ a_{31} & \cdot & \cdot & \cdots & \cdot \\ \vdots & \cdot & \cdot & \cdots & \vdots \\ a_{n1} & \cdot & \cdot & \cdots & a_{nn} \end{vmatrix} = \sum_{P[\alpha\beta\gamma\cdots\omega]} \varepsilon_{\alpha\beta\gamma\cdots\omega} a_{1\alpha} a_{2\beta} a_{3\gamma} \cdots a_{n\gamma}. \quad (3.5)$$

Die Summe läuft dabei über alle $n!$ Permutationen der Indizes $\{\alpha, \beta, \gamma, \cdots, \omega\} = \{1, 2, 3, \cdots, n\}$, und der Faktor ε ist entweder $+1$ oder -1 je nachdem, ob es sich um eine gerade oder ungerade Permutation von $\{1, 2, 3, \cdots, n\}$ handelt.

Beispiel. Wir erläutern die Bedeutung von Gleichung 3.5 anhand von einigen Beispielen. Für $n = 2$ gibt es $2! = 2$ Permutationen von $1, 2$, nämlich $1, 2$ und $2, 1$. Erstere ist unverändert, also null-mal vertauscht und erhält ein $\varepsilon_{12} = +1$, und zweitere geht durch einmaliges Vertauschen hervor und verhält den Faktor $\varepsilon_{21} = -1$. Damit erhalten wir nach 3.5

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \varepsilon_{12} a_{11} a_{22} + \varepsilon_{21} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21},$$

was natürlich mit unsere ursprünglichen Definition 3.2 übereinstimmt.

Für $n = 3$ gibt es bereits $3! = 6$ Permutationen von $\{1, 2, 3\}$, nämlich: $\{1, 2, 3\}$, $\{2, 1, 3\}$, $\{1, 3, 2\}$, $\{3, 2, 1\}$, $\{2, 3, 1\}$, und $\{3, 1, 2\}$, mit den folgenden ε -Faktoren: $\varepsilon_{123} = +1$, $\varepsilon_{213} = -1$, $\varepsilon_{132} = -1$, $\varepsilon_{321} = -1$, $\varepsilon_{231} = +1$, und $\varepsilon_{312} = +1$. Die Anwendung der Leibniz'schen Formel 3.5 liefert somit

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \varepsilon_{123} a_{11} a_{22} a_{33} + \varepsilon_{213} a_{12} a_{21} a_{33} + \varepsilon_{132} a_{11} a_{23} a_{32} \\ &+ \varepsilon_{321} a_{13} a_{22} a_{31} + \varepsilon_{231} a_{12} a_{23} a_{31} + \varepsilon_{312} a_{13} a_{21} a_{32} \\ &= +a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}. \end{aligned}$$

Auch dieses Ergebnis stimmt natürlich mit der Formel von Sarrus 3.4 überein.

Für $n = 4$ gibt es bereits $4! = 24$ Permutationen von $\{1, 2, 3, 4\}$, für $n = 5$ entsprechend $5! = 120$ Permutationen von $\{1, 2, 3, 4, 5\}$, und beispielsweise für die Berechnung einer 10-Determinante bedeutet die Leibniz'sche Formel bereits die Summe über $10! = 3628800$ Produkte!

Eine andere Möglichkeit, Determinanten zu berechnen ist die sogenannte *Laplace-Entwicklung*. Dies ist eine rekursive Vorschrift, in der die Berechnung der n -Determinante auf die Berechnung von n Unterdeterminanten der Größe $(n - 1) \times (n - 1)$ zurückgeführt wird. Hierzu wird die sogenannte *Minor* M_{ij} der $n \times n$ Matrix a_{ij} eingeführt als Determinante der $(n - 1) \times (n - 1)$ -Matrix, die durch Wegstreichen der i -ten Zeile und j -ten Spalte von a_{ij} hervorgeht. Die Laplace-Entwicklung kann entweder als Entwicklung nach einer beliebigen Spalte j aufgefasst werden,

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij} = (-1)^{1+j} a_{1j} M_{1j} + (-1)^{2+j} a_{2j} M_{2j} + \cdots + (-1)^{n+j} a_{nj} M_{nj}, \quad (3.6)$$

oder als Entwicklung nach einer beliebigen Zeile i

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_{ij} = (-1)^{i+1} a_{i1} M_{i1} + (-1)^{i+2} a_{i2} M_{i2} + \cdots + (-1)^{i+n} a_{in} M_{in}. \quad (3.7)$$

Beispiel. Wir erläutern die Laplace-Entwicklung anhand einer 3-Determinante und entscheiden uns für eine Entwicklung nach der 1. Spalte der Matrix. Nach Gleichung 3.6 ergibt sich somit mit $j = 1$

$$\begin{aligned}
 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= (-1)^{1+1}a_{11}M_{11} + (-1)^{2+1}a_{21}M_{21} + (-1)^{3+1}a_{31}M_{31} \\
 &= +a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\
 &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}.
 \end{aligned}$$

Dieses Ergebnis stimmt natürlich wieder mit der der Gleichung 3.4 (Regel von Sarrus) überein.

CDF 15. Entwicklung von Determinanten [33DeterminantsByExpansion.cdf](#)
 Interaktive Demonstration zur Entwicklung von Determinanten.

expand along:
 row
 column

Expand along row 1:

$$\begin{vmatrix} -1 & 2 & 3 \\ 4 & -5 & 6 \\ 7 & 8 & -9 \end{vmatrix} =$$

$$-1 \begin{vmatrix} -5 & 6 \\ 8 & -9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 7 & -9 \end{vmatrix} + 3 \begin{vmatrix} 4 & -5 \\ 7 & 8 \end{vmatrix} =$$

$$-1(-3) - 2(-78) + 3(67) =$$

$$3 + 156 + 201 =$$

$$360$$

Mathematisch gesehen eigentlich wichtiger als diese Berechnungsvorschriften für Determinanten (Leibniz-Formel 3.5 oder Laplace-Entwicklungen 3.6 oder 3.7) sind eine Reihe von Eigenschaften von Determinanten, von denen wir hier die wichtigsten zusammenfassen:

Eigenschaften von Determinanten. Die n -Determinante $D = \det(a_{ij})$ einer $n \times n$ Matrix hat folgende Eigenschaften:

- (D1) Die Determinante verschwindet ($D = 0$), wenn zumindest eine Zeile oder eine Spalte verschwindet.
- (D2) Die Determinante verschwindet ($D = 0$), wenn zwei Zeilen oder zwei Spalten gleich oder zueinander proportional sind.
- (D3) Wenn man zwei Zeilen oder zwei Spalten miteinander vertauscht, so wechselt die Determinante ihr Vorzeichen ($D \rightarrow -D$).
- (D4) Eine Determinante bleibt unverändert, wenn alle Zeilen mit den Spalten vertauscht werden: $\det(a_{ij}) = \det(a_{ji})$.
- (D5) Eine Determinante bleibt unverändert, wenn zu einer Zeile (oder Spalte) das Vielfache einer *anderen* Zeile (bzw. Spalte) addiert wird.
- (D6) Die Determinante ist eine n -lineare Funktion. Das bedeutet folgendes: Drückt man die j -te Spalte einer Matrix a_{ij} aus durch die Linearkombination von zwei Spaltenvektoren v_i und w_i also , $a_{ij} = \alpha v_i + \beta w_i$, so ist die Determinante von a_{ij} gleich der Linearkombination der zwei Determinanten, bei denen die j -te Spalte durch v_i bzw. w_i ersetzt wurde:

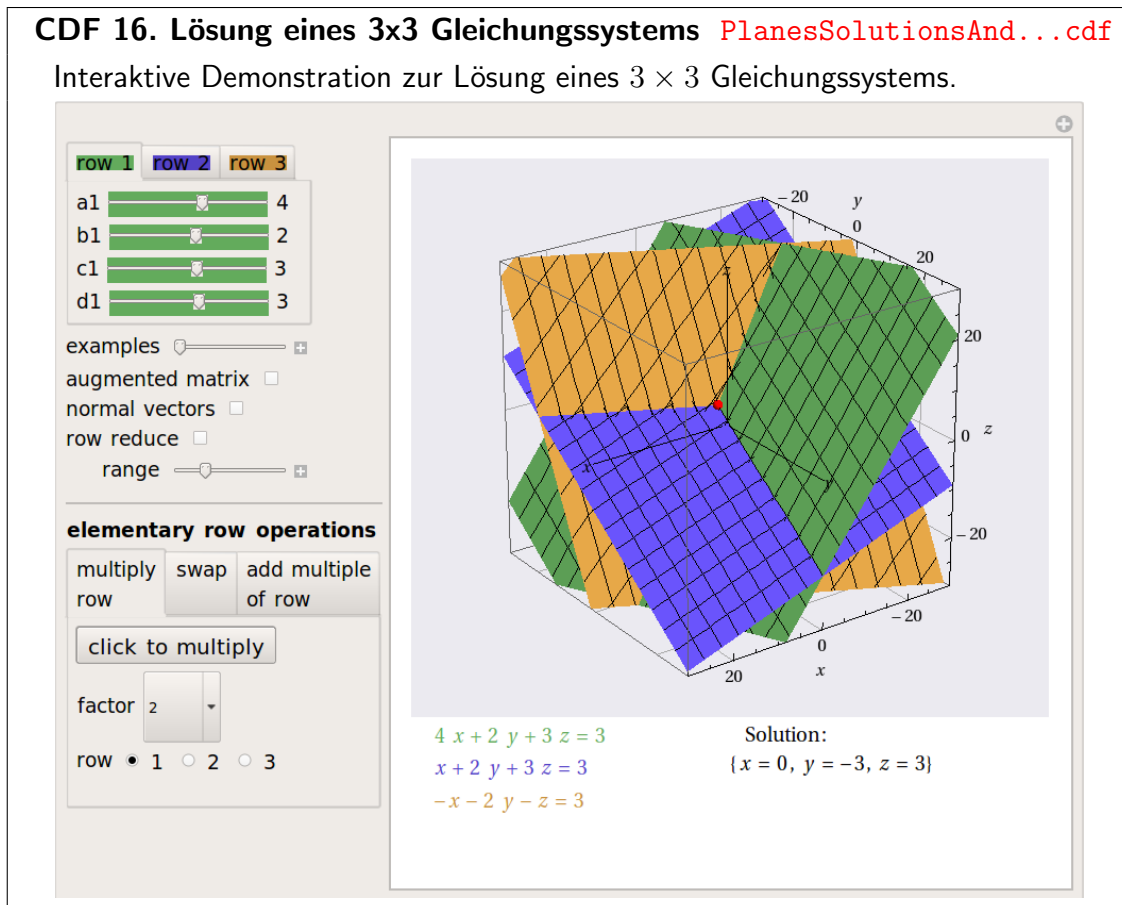
$$\begin{vmatrix} a_{11} & \cdots & (\alpha v_1 + \beta w_1) & \cdots & a_{1n} \\ a_{21} & \cdots & (\alpha v_2 + \beta w_2) & \cdots & a_{2n} \\ a_{31} & \cdots & (\alpha v_3 + \beta w_3) & \cdots & a_{3n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & (\alpha v_n + \beta w_n) & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & v_1 & \cdots & a_{1n} \\ a_{21} & \cdots & v_2 & \cdots & a_{2n} \\ a_{31} & \cdots & v_3 & \cdots & a_{3n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & v_n & \cdots & a_{nn} \end{vmatrix} + \beta \begin{vmatrix} a_{11} & \cdots & w_1 & \cdots & a_{1n} \\ a_{21} & \cdots & w_2 & \cdots & a_{2n} \\ a_{31} & \cdots & w_3 & \cdots & a_{3n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & w_n & \cdots & a_{nn} \end{vmatrix} \quad (3.8)$$

3.1.2 Lösung eines linearen Gleichungssystems

Mit Hilfe von Determinanten können wir die Lösung(en) eines linearen Gleichungssystems auf systematische Art und Weise bestimmen. Wir demonstrieren dieses Verfahren anhand eines Systems von 3 Gleichungen für drei Unbekannte.

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{23}z &= b_2 \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned} \quad (3.9)$$

Geometrisch betrachtet beschreibt jede der drei Gleichungen eine Ebene im \mathbb{R}^3 . Im allgemeinen schneiden sich diese 3 Ebenen in einem Punkt (x_0, y_0, z_0) , den wir dann als die Lösung des linearen Gleichungssystems bezeichnen. Es ist aber auch möglich, dass die Lösung eine Gerade bzw. Ebene im Raum beschreibt, oder auch dass keine Lösung existiert. Eine geometrische Veranschaulichung dieser Möglichkeiten wird in der CDF-Demonstration 16 gegeben. Mathematisch können wir die Lösungen bzw. die Arten der Lösung mit Hilfe von Determinanten charakterisieren.



Dazu benötigen wir die Determinante D der Koeffizienten a_{ij} , sowie drei weitere Determinanten D_1 , D_2 , und D_3 , die man erhält wenn jeweils eine Spalte von a_{ij} durch den Spaltenvektor b_i der rechten Seite des LGS ausgetauscht wird:

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}. \quad (3.10)$$

Die *Cramersche Regel* besagt, dass es für den Fall $D \neq 0$ genau eine Lösung gibt, die wie folgt

berechnet werden kann

$$D \neq 0 \quad \Rightarrow \quad x_0 = \frac{D_1}{D}, \quad y_0 = \frac{D_2}{D}, \quad z_0 = \frac{D_3}{D}. \quad (3.11)$$

Beweis. Wir zeigen die Cramersche Regel anhand des 3×3 LGS definiert in Gleichung 3.9. Bringen wir den Vektor b_i auf die linke Seite, so können wir 3.9 auch folgendermaßen anschreiben

$$x \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + \begin{pmatrix} ya_{12} - b_1 \\ ya_{22} - b_2 \\ ya_{32} - b_3 \end{pmatrix} + z \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hierbei haben wir uns entschieden, den Spaltenvektor b_i beim 2. Summanden abzuziehen.¹ Diese Gleichung bedeutet jedoch nichts anderes, als dass die drei Spaltenvektoren

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \begin{pmatrix} ya_{12} - b_1 \\ ya_{22} - b_2 \\ ya_{32} - b_3 \end{pmatrix}, \quad \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

nicht unabhängig voneinander sind, sondern linear abhängig (proportional) sind. Mit der Eigenschaft (D2) von Determinanten muss folgende Determinante somit den Wert 0 haben

$$\begin{vmatrix} a_{11} & ya_{12} - b_1 & a_{13} \\ a_{21} & ya_{22} - b_2 & a_{23} \\ a_{31} & ya_{32} - b_3 & a_{33} \end{vmatrix} = 0.$$

Wegen der n -Linearität von Determinanten – Eigenschaft (D6) – gilt weiters

$$y \underbrace{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}_{=D} - \underbrace{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}_{=D_2} = 0,$$

womit wir die Cramersche Regel für $y = D_2/D$ gezeigt haben. Analog zeigt man $x = D_1/D$ bzw. $z = D_3/D$, indem wir den Spaltenvektor b_i in die erste bzw. dritte Spalte bringen.

”Die Cramersche Regel ist eine besonders unpraktische Methode zur Berechnung der Lösung linearer Gleichungssysteme. Sie ist aber trotzdem von großem mathematischen Interesse, weil sie zeigt, wie sich die Lösung verändert, wenn man die Daten des LGS, also die a_{ij} oder die b_i , ändert.” (Zitat: Jänich, Lineare Algebra, Springer-Lehrbuch).

¹Genauso gut hätten wir b_i auch beim ersten oder dritten Summanden abziehen können.

Was passiert, wenn die Determinante verschwindet, d.h. $D = 0$ gilt? In diesem Fall gibt es zwei Möglichkeiten: (i) Zumindest eine der Determinanten $D_i \neq 0$. Dann sind die Gleichungen widersprüchlich, und es gibt keine Lösung. (ii) *Alle* $D_i = 0$. Dann sind die Gleichungen linear abhängig, und die Lösungsmenge ist eine leere Menge oder eine unendliche Punktmenge (zum Beispiel eine Gerade oder eine Ebene). Genaueres folgt dann im Abschnitt 3.3.4 über den Rang von Matrizen. Zuvor verlassen wir aber das Thema Determinanten und Matrizen und wenden uns Vektoren zu, die man auch als einspaltige Matrizen auffassen kann.

3.2 Vektoren und ihre Algebra

Im allgemeinen Sinn versteht man unter einem Vektor ein Element eines Vektorraums, das heißt ein Objekt, das zu anderen Vektoren addiert und mit Zahlen, die als Skalare bezeichnet werden, multipliziert werden kann. Im engeren Sinne versteht man unter einem Vektor in der analytischen Geometrie ein mathematisches Objekt, das eine Parallelverschiebung in der Ebene oder im Raum beschreibt. In kartesischen Koordinaten werden Vektoren durch Zahlenpaare (in der Ebene, \mathbb{R}^2) bzw. -tripel (im Raum, \mathbb{R}^3) dargestellt, die oft untereinander (als "Spaltenvektoren") geschrieben werden. Davon ausgehend bezeichnen wir ein n -Tupel reeller Zahlen also ein Element des \mathbb{R}^n ebenfalls als Vektor. In der klassischen Physik bezeichnen wir als *Vektoren* eine physikalische Größe, die durch einen Betrag und eine Richtung gekennzeichnet ist (z.B.: Geschwindigkeit), während *Skalare* physikalisch Größen bezeichnen, die keine Richtung aufweisen (z.B.: Temperatur).

3.2.1 Schreibweisen und Definitionen

Namen und Darstellung von Vektoren. Als *symbolischen Namen* für Vektoren verwenden wir in diesem Skriptum fettgedruckte, kursive Buchstaben, z.B. $\mathbf{A}, \mathbf{B}, \mathbf{C}$ oder $\mathbf{a}, \mathbf{b}, \mathbf{c}$ usw. Wenn wir konkret die Koordinaten eines Vektors angeben, also die *Darstellung eines Vektors* in einem bestimmten Koordinatensystem, dann verwenden wird denselben Buchstaben, allerdings mit einem *Index*, der die Werte 1, 2, oder 3 annehmen kann:

$$\{\mathbf{A}\}_k = A_k, \quad k = 1, 2, 3$$

Die geschwungene Klammer mit dem Index auf der linken Seite soll ausdrücken, dass wir von dem symbolischen Ausdruck \mathbf{A} in die Koordinatendarstellung übergehen.

Betrag und Norm von Vektoren. Die *Länge* oder den *Betrag eines Vektors* \mathbf{A} bezeichnen wir mit $|\mathbf{A}|$ oder einfach A und es gilt:

$$|\mathbf{A}| = A = \sqrt{A_1^2 + A_2^2 + A_3^2} \quad (3.12)$$

Die Länge (Betrag) eines Vektors ein Skalar ist, das heißt, hängt nicht von der Wahl des Koordinatensystems ab.

Einstein'sche Summenkonvention. Das häufige Auftreten von bestimmten Summationen in der Vektorrechnung hat zur Aufstellung des so genannten Summationsübereinkommens geführt. Diese Konvention macht das Schreiben von Summenzeichen überflüssig. Es besteht in der Festsetzung, dass *über jeden Index, der in einem Produkt zweimal vorkommt, von 1 bis 3 zu summieren ist*. Als Beispiel geben wir nochmals den Betrag eines Vektors nach dieser Konvention an:

$$|\mathbf{A}|^2 = A_i A_i = \sum_{i=1}^3 A_i A_i = A_1 A_1 + A_2 A_2 + A_3 A_3.$$

Es ist zu beachten, dass die Wahl des Summationsindex beliebig ist, also folgende Ausdrücke identisch sind:

$$|\mathbf{A}| = \sqrt{A_i A_i} = \sqrt{A_j A_j} = \sqrt{A_k A_k}.$$

Nicht zulässig wäre obigen Ausdruck in folgender Weise zu schreiben:

$$|\mathbf{A}| = \sqrt{A_i^2}, \quad (\text{falsch!}),$$

weil ja der Summationsindex i nur einmal auftaucht, und somit nicht klar ist, worüber zu summieren ist.

3.2.2 Vektoralgebra

Hier fassen wir einfache Rechenregeln für das Addieren und Subtrahieren von Vektoren sowie die Multiplikation von Vektoren mit Skalaren zusammen. Es ist zu beachten, dass für die konkrete Rechnung von der symbolischen Schreibweise, z.B. $\mathbf{A} + \mathbf{B}$, immer auf die Darstellung des Vektors in einem Koordinatensystem übergegangen wird, also:

$$\{\mathbf{A} + \mathbf{B}\}_k = A_k + B_k \quad (3.13)$$

$$\{\mathbf{A} - \mathbf{B}\}_k = A_k - B_k. \quad (3.14)$$

Weiters lässt sich leicht zeigen, dass die Vektoraddition kommutativ und assoziativ ist:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \quad (3.15)$$

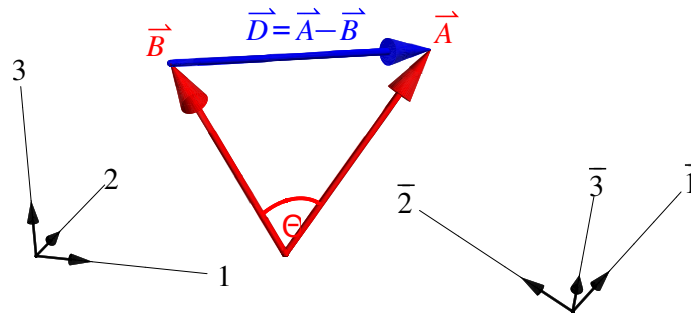
$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}). \quad (3.16)$$

Unter dem Produkt eines Vektors \mathbf{A} mit einem Skalar λ , symbolisch schreiben wir $\lambda\mathbf{A}$, verstehen wir jenen Vektor, dessen Koordinaten sich durch die Multiplikation der entsprechenden Koordinaten von \mathbf{A} mit λ ergeben:

$$\{\lambda\mathbf{A}\}_k = \lambda A_k. \quad (3.17)$$

3.2.3 Das Skalarprodukt

Wir suchen zunächst den Winkel θ zwischen zwei Vektoren \mathbf{A} und \mathbf{B} und verwenden dazu das aus den Vektoren \mathbf{A} , \mathbf{B} , und $\mathbf{D} = \mathbf{A} - \mathbf{B}$ gebildete Dreieck (siehe Abbildung).



Aus dem Cosinussatz folgt:

$$D^2 = A^2 + B^2 - 2AB \cos \theta.$$

Andererseits folgt aus der Bildung des Betragsquadrats von $\mathbf{D} = \mathbf{A} - \mathbf{B}$ in einem Koordinatensystem 1,2,3:

$$\begin{aligned} D^2 &= D_i D_i = (A_i - B_i)(A_i - B_i) \\ &= A_i A_i + B_i B_i - 2A_i B_i \\ &= A^2 + B^2 - 2A_i B_i. \end{aligned}$$

Der Vergleich der obigen beiden Gleichungen zeigt, dass gilt

$$A_i B_i = AB \cos \theta, \quad (\text{Summenkonvention!})$$

In symbolischer Schreibweise führen wir für das sogenannte *skalare Produkt* folgende Schreibweise ein:

$$\mathbf{AB} = AB \cos \theta. \quad (3.18)$$

Das skalare Produkt zweier Vektoren \mathbf{AB} ist, wie der Name schon nahelegt, ein Skalar, d.h. invariant in Bezug auf Änderung des Koordinatensystems. Diese Tatsache sieht man daran, dass die rechte Seite von Gl. 3.18 das Produkt der Längen A und B der Vektoren und des Cosinus des eingeschlossenen Winkel θ ist. Alle diese Größen sind unabhängig von der Wahl des Koordinatensystems. Das skalare Produkt ist kommutativ, distributiv, aber *nicht* assoziativ:

$$\mathbf{AB} = \mathbf{BA} \quad (3.19)$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \quad (3.20)$$

$$\mathbf{A}(\mathbf{BC}) \neq (\mathbf{AB})\mathbf{C}. \quad (3.21)$$

Dass 3.21 nicht gilt, ist unmittelbar einsichtig, weil die linke Seite ja einen Vektor parallel zu \mathbf{A} darstellt, während die rechte Seite ein Vektor parallel zu \mathbf{C} ist. Für beliebige Vektoren \mathbf{A} und \mathbf{C} kann somit 3.21 nicht gelten.

Aus 3.18 folgt, dass das skalare Produkt verschwindet, wenn entweder \mathbf{A} oder \mathbf{B} der Nullvektor ist (Vektor mit Länge Null), oder der $\cos \theta = 0$, das heißt der Winkel $\theta = \frac{\pi}{2}$, gleichbedeutend mit der Aussage, dass \mathbf{A} senkrecht auf \mathbf{B} steht.

3.2.4 Darstellung von Vektoren

Mithilfe des Skalarprodukts können wir die *Koordinate* und die *Komponente* eines Vektors \mathbf{A} in Richtung eines Einheitsvektors \mathbf{e} berechnen, wobei $|\mathbf{e}| = 1$.

Unter der *Koordinate* des Vektors \mathbf{A} in Richtung eines Einheitsvektors \mathbf{e} verstehen wir das skalare Produkt (die Zahl)

$$\mathbf{Ae} = A \cos \theta = A_i e_i \quad (3.22)$$

Unter der *Komponente* des Vektors \mathbf{A} in Richtung eines Einheitsvektors \mathbf{e} verstehen wir den Vektor

$$\mathbf{A}' = (\mathbf{Ae})\mathbf{e} \quad (3.23)$$

$$\{\mathbf{A}'\}_k = A_i e_i e_k. \quad (3.24)$$

Jeder Vektor \mathbf{A} lässt sich in eindeutiger Weise als Summe von zwei Vektoren \mathbf{A}' und \mathbf{A}'' darstellen,

von denen der erste parallel und der zweite senkrecht zu einem gegebenen Einheitsvektor \mathbf{e} ist:

$$\mathbf{A} = \mathbf{A}' + \mathbf{A}'' \quad (3.25)$$

$$\mathbf{A}' = (\mathbf{A}\mathbf{e})\mathbf{e} \quad (3.26)$$

$$\mathbf{A}'' = \mathbf{A} - (\mathbf{A}\mathbf{e})\mathbf{e}. \quad (3.27)$$

Definition. Wir bezeichnen eine *orthogonale Basis* als ein System von drei Einheitsvektoren \mathbf{e}^1 , \mathbf{e}^2 , und \mathbf{e}^3 , die paarweise aufeinander senkrecht stehen

$$\mathbf{e}^i \mathbf{e}^k = \delta_{ik}. \quad (3.28)$$

Hierbei haben wir das so genannte *Kronecker'sche Delta* verwendet, das 1 ist wenn die Indizes $i = k$ und 0 ist, wenn $i \neq k$. Zu beachten ist auch, dass die hochgestellten Indizes zum *Namen* des Vektors gehören. Schreibt man die Gleichung 3.28 ausführlicher, dann sind das die folgenden 9 Gleichungen:

$$\begin{array}{lll} \mathbf{e}^1 \mathbf{e}^1 = 1 & \mathbf{e}^1 \mathbf{e}^2 = 0 & \mathbf{e}^1 \mathbf{e}^3 = 0 \\ \mathbf{e}^2 \mathbf{e}^1 = 0 & \mathbf{e}^2 \mathbf{e}^2 = 1 & \mathbf{e}^2 \mathbf{e}^3 = 0 \\ \mathbf{e}^3 \mathbf{e}^1 = 0 & \mathbf{e}^3 \mathbf{e}^2 = 0 & \mathbf{e}^3 \mathbf{e}^3 = 1 \end{array}$$

Eine *rechtsorientierte orthogonale Basis* liegt vor, wenn \mathbf{e}^1 , \mathbf{e}^2 , und \mathbf{e}^3 eine orthogonale Basis bilden, und wenn eine Drehung von \mathbf{e}^1 in Richtung \mathbf{e}^2 verbunden mit einem Fortschreiten in Richtung der Orientierung von \mathbf{e}^3 die Bewegung einer Rechtsschraube ergibt.

Stimmen die Richtungen und Orientierungen der Einheitsvektoren einer orthogonalen Basis mit den Richtungen und Koordinatenachsen 1,2,3 eines Koordinatensystems überein, dann gilt:

$$\mathbf{e}_i^k = \delta_{ik}. \quad (3.29)$$

Bei dieser Gleichung in Koordinatenschreibweise ist zu beachten, dass der hochgestellte Index k den Namen des Einheitsvektors bezeichnet, während der tiefgestellte Index i die i -te Koordinate anzeigt.

Satz. Die skalaren Produkte eines Vektors \mathbf{A} mit den Einheitsvektoren \mathbf{e}^k in Richtung der Koordinatenachsen liefern die Koordinaten A_k des Vektors \mathbf{A} bezüglich des gegebenen Koordinatensystems:

$$\mathbf{A}\mathbf{e}^k = A_i \mathbf{e}_i^k = A_i \delta_{ik} = A_k. \quad (3.30)$$

Somit lässt sich jeder Vektor als Summe seiner drei Komponenten in Richtung der Achsen eines

Koordinatensystems darstellen:

$$\mathbf{A} = \mathbf{A}^1 + \mathbf{A}^2 + \mathbf{A}^3 = A_1\mathbf{e}^1 + A_2\mathbf{e}^2 + A_3\mathbf{e}^3 \quad (3.31)$$

$$\mathbf{A}^1 = (\mathbf{A}\mathbf{e}^1)\mathbf{e}^1 = A_1\mathbf{e}^1 \quad (3.32)$$

$$\mathbf{A}^2 = (\mathbf{A}\mathbf{e}^2)\mathbf{e}^2 = A_2\mathbf{e}^2 \quad (3.33)$$

$$\mathbf{A}^3 = (\mathbf{A}\mathbf{e}^3)\mathbf{e}^3 = A_3\mathbf{e}^3 \quad (3.34)$$

3.2.5 Das Vektorprodukt

Vermutlich kennen Sie das Vektorprodukt (oder Kreuzprodukt) zweier Vektoren nach folgender Definition:

$$\mathbf{A} \times \mathbf{B} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \times \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = \begin{pmatrix} A_2B_3 - A_3B_2 \\ A_3B_1 - A_1B_3 \\ A_1B_2 - A_2B_1 \end{pmatrix}$$

Wir geben hier einige wichtige Eigenschaften des Vektorprodukts an.

Satz. Bei Vertauschen der Reihenfolge der Vektoren ändert das vektorielle Produkt sein Vorzeichen (d.h. das vektorielle Produkt ist nicht kommutativ)

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A} \quad (3.35)$$

Bsp. Es ist die Fläche des durch die Punkte $A_i = (1, 2, 1)$, $B_i = (2, 7, 5)$, und $C_i = (5, 4, 2)$ bestimmten Dreiecks zu berechnen.

Satz. Das vektorielle Produkt $\mathbf{A} \times \mathbf{B}$ der Vektoren \mathbf{A} und \mathbf{B} steht senkrecht auf beide Vektoren \mathbf{A} und \mathbf{B} , d.h. es gilt:

$$(\mathbf{A} \times \mathbf{B})\mathbf{A} = 0 \quad \text{und} \quad (\mathbf{A} \times \mathbf{B})\mathbf{B} = 0 \quad (3.36)$$

Bsp. Wir untersuchen, ob die folgenden drei Vektoren in einer Ebene liegen: $A_i = (4, 5, 2)$, $B_i = (2, -2, 1)$, $C_i = (-1, 0, 2)$

Satz. Der Betrag des vektoriellen Produkts $\mathbf{A} \times \mathbf{B}$ der Vektoren \mathbf{A} und \mathbf{B} ist gegeben durch:

$$|\mathbf{A} \times \mathbf{B}| = \sqrt{A^2 B^2 - (\mathbf{A}\mathbf{B})^2} \quad (3.37)$$

Der Betrag des vektoriellen Produkts $\mathbf{A} \times \mathbf{B}$ der Vektoren \mathbf{A} und \mathbf{B} ist gleich der Fläche des von \mathbf{A} und \mathbf{B} gebildeten Parallelogramms

$$|\mathbf{A} \times \mathbf{B}| = AB \sin \theta \quad (3.38)$$

Zum Beweis der Eigenschaften 3.35–3.38 des vektoriellen Produkts beweisen siehe z.B. mein Skriptum [Vektoranalysis](#)

Das gemischte Produkt (oder auch Spatprodukt). Das gemischte Produkt der Vektoren $\mathbf{A}, \mathbf{B}, \mathbf{C}$ wird symbolisch mit $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ bezeichnet und ist definiert durch

$$(\mathbf{A}, \mathbf{B}, \mathbf{C}) = (\mathbf{A} \times \mathbf{B})\mathbf{C} = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} \quad (3.39)$$

Das Spatprodukt ist ein Skalar, der das Volumen des durch die Vektoren $\mathbf{A}, \mathbf{B}, \mathbf{C}$ aufgespannten Parallelepipeds angibt.

Bsp. Berechne das Volumen des von den Vektoren $A_i = (0, 3, 0)$, $B_i = (-2, 1, 0)$, und $C_i = (0, 2, 4)$ aufgespannten Parallelepipeds.

3.2.6 Analytische Geometrie

Mit Hilfe von Vektoren kann die analytische Beschreibung von geometrischen Sachverhalten sehr einfach ausgedrückt werden. Wir beschränken uns in diesem Abschnitt auf die Darstellung von Geraden in der Ebene (\mathbb{R}^2) und im Raum (\mathbb{R}^3) bzw. die Darstellung von Ebenen im \mathbb{R}^3 .

Der Ortsvektor. Der Vektor, der den Ursprung mit dem Punkt (x, y, z) verbindet, nennen wir den Ortsvektor (er ist somit ein *gebundener* Vektor) und ist gegeben durch

$$\mathbf{r} = x \mathbf{e}^1 + y \mathbf{e}^2 + z \mathbf{e}^3 \quad (\mathbb{R}^3) \quad (3.40)$$

$$\mathbf{r} = x \mathbf{e}^1 + y \mathbf{e}^2 \quad (\mathbb{R}^2) \quad (3.41)$$

Darstellung von Geraden. Die *Parameterdarstellung einer Geraden* in \mathbb{R}^2 oder \mathbb{R}^3 erhalten wir,² indem wir den Ortsvektor als Funktion *eines* Parameters $t \in \mathbb{R}$ auffassen

$$\mathbf{r}(t) = \mathbf{r}_0 + t \mathbf{a}. \quad (3.42)$$

Hierbei ist \mathbf{r}_0 ein beliebiger Punkt, der auf der Geraden liegt, und \mathbf{a} ist ein Vektor, der die Richtung der Gerade angibt.

In \mathbb{R}^2 , können wir eine Gerade auch in folgender *impliziter* Form angeben

$$(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n} = 0. \quad (3.43)$$

Hierbei ist \mathbf{n} der Normalenvektor, also ein Vektor, der senkrecht auf die Gerade steht: $\mathbf{a} \perp \mathbf{n} \Leftrightarrow \mathbf{a} \cdot \mathbf{n} = 0$.

Schließlich gibt es im \mathbb{R}^3 noch eine weitere Möglichkeit, eine Gerade anzugeben, nämlich *implizit* als die Menge aller Punkte, die durch \mathbf{r}_0 gehen und parallel zu \mathbf{a} sind. Dementsprechend muss das Kreuzprodukt von $\mathbf{r} - \mathbf{r}_0$ mit \mathbf{a} verschwinden

$$(\mathbf{r} - \mathbf{r}_0) \times \mathbf{a} = 0. \quad (3.44)$$

Diese Vektorgleichung beschreibt eigentlich ein System von 3 linearen Gleichungen, von denen 2 linear unabhängig sind, und deren Schnittgerade die dargestellte Gerade ist.

Darstellung von Ebenen. Die *Parameterdarstellung einer Ebene* im \mathbb{R}^3 erhalten wir, indem wir den Ortsvektor als Funktion von *zwei* Parametern $s \in \mathbb{R}$ und $t \in \mathbb{R}$ auffassen

$$\mathbf{r}(s, t) = \mathbf{r}_0 + s \mathbf{u} + t \mathbf{v}. \quad (3.45)$$

Hierbei ist \mathbf{r}_0 ein beliebiger Punkt, der sich in der Ebene befindet, und die beiden Vektoren \mathbf{u} und \mathbf{v} , die nicht parallel zueinander sein dürfen, sind zwei Richtungsvektoren, die in der Ebene liegen.

Die *implizite Form einer Ebenengleichung* in \mathbb{R}^3 erhalten wir als Menge aller Punkte, deren Skalarprodukt mit dem Normalenvektor \mathbf{n} verschwindet

$$(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n} = 0. \quad (3.46)$$

Hierbei ist der Normalenvektor \mathbf{n} ein Vektor, der senkrecht auf die Ebene steht, und z.B. über ein Vektorprodukt zweier unabhängiger Richtungsvektoren aufgestellt werden kann: $\mathbf{n} = \mathbf{u} \times \mathbf{v}$.

²Diese Darstellung ist sogar in beliebig vielen Dimensionen also im \mathbb{R}^n möglich.

3.3 Matrizen

In der Mathematik versteht man unter einer Matrix (Plural Matrizen) eine rechteckige Anordnung (Tabelle) von Zahlen. Mit diesen Objekten lässt sich dann in bestimmter Weise rechnen, indem man Matrizen addiert oder miteinander multipliziert. Matrizen können beliebige Dimensionalität besitzen. Matrizen sind ein Schlüsselkonzept der linearen Algebra und tauchen in fast allen Gebieten der Mathematik auf. Sie werden insbesondere dazu benutzt, lineare Abbildungen darzustellen und lineare Gleichungssysteme zu beschreiben und zu lösen.

Eine Matrix, die wir in fettgedruckten, nicht kursiven Großbuchstaben notieren, hat n Zeilen und m Spalten schreiben wir in folgender Form:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdot & \cdots & \cdot \\ a_{31} & \cdot & \cdot & \cdots & \cdot \\ \vdots & \cdot & \cdot & \cdots & \vdots \\ a_{n1} & \cdot & \cdot & \cdots & a_{nm} \end{pmatrix} = (a_{ij}) \quad (3.47)$$

3.3.1 Lineare Algebra der Matrizen

Für Matrizen definieren wir die folgenden Rechenregeln (Lineare Algebra der Matrizen)

1. Zwei Matrizen sind gleich, wenn sie in allen Elementen übereinstimmen:

$$\mathbf{A} = \mathbf{B} \Leftrightarrow \forall i, j : a_{ij} = b_{ij}.$$

2. Matrizen werden elementweise addiert. Diese Addition ist kommutativ und assoziativ:

$$\begin{aligned} \mathbf{A} + \mathbf{B} = \mathbf{C} &\Leftrightarrow \forall i, j : a_{ij} + b_{ij} = c_{ij} \\ \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}, \quad \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} \end{aligned}$$

3. Bei der Multiplikation einer Matrix mit einer Zahl α wird jedes Element der Matrix mit dieser Zahl multipliziert. Diese Operation ist kommutativ:

$$\mathbf{B} = \alpha \mathbf{A} \Leftrightarrow \forall i, j : b_{ij} = \alpha a_{ij}, \quad \alpha \mathbf{A} = \mathbf{A} \alpha.$$

4. Das *Produkt zweier Matrizen* ist nur dann möglich, wenn die linksstehende Matrix gleich viele Spalten hat wie die rechtsstehende Matrix Zeilen. Das Ergebnis ist eine Matrix mit so vielen Zeilen wie die linke und so vielen Spalten wie die rechte Matrix:

$$\underbrace{\mathbf{A}}_{n \times k} \cdot \underbrace{\mathbf{B}}_{k \times m} = \underbrace{\mathbf{C}}_{n \times m}, \quad \mathbf{A} \cdot \mathbf{B} = \mathbf{C} \Leftrightarrow c_{ij} = \sum_k a_{ik} b_{kj}.$$

Für die Multiplikation und Addition gilt auch ein distributives Gesetz: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{C}$.

Die Matrixmultiplikation ist im allgemeinen aber nicht kommutativ: $\mathbf{AB} \neq \mathbf{BA}$!

5. Es gibt ein neutrales Element in Bezug auf die Matrixaddition, $\mathbf{0}$, sowie ein neutrales Element in Bezug auf die Matrixmultiplikation, $\mathbf{1}$:

$$\mathbf{A} + \mathbf{0} = \mathbf{A}, \quad \mathbf{A}\mathbf{0} = \mathbf{0}\mathbf{A} = \mathbf{0}, \quad \mathbf{A}\mathbf{1} = \mathbf{1}\mathbf{A} = \mathbf{A}.$$

Das Nullelement $\mathbf{N} = \mathbf{0}$ sowie das Einselement $\mathbf{E} = \mathbf{1}$ ist dabei wie folgt definiert:

$$\mathbf{N} = \mathbf{0} \Leftrightarrow \forall i, j : n_{ij} = 0$$

$$\mathbf{E} = \mathbf{1} \Leftrightarrow \forall i, j : e_{ij} = \delta_{ij}, \quad \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

CDF 17. Illustration der Matrixmultiplikation [MatrixMultiplication.cdf](#)
 Interaktive Demonstration zur Matrixmultiplikation.

A's rows: 2
 A's columns: 3
 B's rows: 3
 B's columns: 3

new matrices
 row: [slider]
 column: [slider]

$$A = \begin{pmatrix} 2 & 3 & -5 \\ -3 & 2 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 1 & 0 \\ -2 & 1 & -5 \\ -2 & 4 & -3 \end{pmatrix}$$

$$AB = \begin{pmatrix} 12 & -15 & 0 \\ -12 & -9 & -4 \end{pmatrix}$$

$$\begin{aligned} & (2)(1) + (3)(1) + (-5)(4) \\ & = \\ & 2 + 3 - 20 \\ & = \\ & -15 \end{aligned}$$

LGS und Matrizen. Eine naheliegende Anwendung von Matrizen und der oben definierten Algebra ist, dass wir lineare Gleichungssysteme wie unser anfängliches Beispiel

$$\begin{aligned} x_1 + 3x_2 &= 6 \\ 2x_1 - x_2 &= 5. \end{aligned}$$

mit Hilfe von Matrizen sehr kompakt anschreiben können:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \quad \mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}.$$

Hier bei ist \mathbf{A} eine 2×2 Matrix während \mathbf{x} und \mathbf{b} jeweils 2×1 Matrizen sind, also einspaltige Matrizen, die wir auch als Spaltenvektoren bezeichnen wollen. Um dies zu verdeutlichen, verwenden wir für diese Spaltenvektoren fettgedruckte, kursive, und kleine Buchstaben.

3.3.2 Die inverse Matrix

Wir betrachten das LGS in der Form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b},$$

für quadratische Matrizen \mathbf{A} (d.h. gleich viele Gleichungen wie Unbekannte). Stünde an Stelle der Matrix \mathbf{A} einfach eine Zahl a , so könnten wir die Lösung einfach angeben,

$$a \mathbf{x} = \mathbf{b} \quad \Rightarrow \quad \mathbf{x} = a^{-1} \mathbf{b}.$$

Wir stellen uns nun die Frage, ob es auch für Matrizen ein solches inverses Element in Bezug auf die Multiplikation gibt, oder anders ausgedrückt, ob es eine zu \mathbf{A} *inverse Matrix* " \mathbf{A}^{-1} " mit folgender Eigenschaft gibt

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{1}. \quad (3.48)$$

Man kann zeigen, dass die inverse Matrix \mathbf{A}^{-1} existiert, sofern die Determinante von \mathbf{A} nicht verschwindet. Die Elemente $(\mathbf{A}^{-1})_{ij}$ der inversen Matrix sind durch folgende Gleichung gegeben:

$$(\mathbf{A}^{-1})_{ij} = \frac{(-1)^{i+j} M_{ji}}{\det \mathbf{A}} = \frac{C_{ji}}{\det \mathbf{A}}, \quad (3.49)$$

wobei wir neben dem bereits bekannten Minor M_{ji} von \mathbf{A} (siehe Gl. 3.6) auch die so genannte Kofaktorenmatrix $C_{ij} = (-1)^{i+j} M_{ij}$ eingeführt haben. Beachte, dass für die Berechnung der i -ten Zeile und j -ten Spalte der inversen Matrix, die i -te *Spalte* und j -te *Zeile* des Minors bzw. der Kofaktorenmatrix benötigt wird!

Beispiel. Wir betrachten die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix},$$

und berechnen zunächst ihre Kofaktorenmatrix:

$$C_{11} = (-1)^{1+1} \cdot (-1) = -1, \quad C_{12} = (-1)^{1+2} \cdot 3 = -3, \quad C_{21} = (-1)^{2+1} \cdot 2 = -2, \quad C_{22} = (-1)^{2+2} \cdot 1 = 1.$$

Mit der Determinante $\det \mathbf{A} = -1 - 6 = -7$ erhalten wir nach Gleichung 3.49,

$$\mathbf{A}^{-1} = -\frac{1}{7} \begin{pmatrix} -1 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix}.$$

Wir überprüfen leicht, dass gilt

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Lösung durch Matrixinversion. Wir lösen das nun das LGS $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, indem wir beide Seiten der Gleichung von links mit der inversen Matrix \mathbf{A}^{-1} und erhalten

$$\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} \quad \Rightarrow \quad \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}.$$

Wir müssen also lediglich die inverse Matrix \mathbf{A}^{-1} mit dem Spaltenvektor \mathbf{b} (rechte Seite des LGS) multiplizieren. Für unser altbekanntes Beispiel erhalten wir nach dieser Vorschrift

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 5 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 21 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Wir bemerken nochmals, dass dieses Lösungsverfahren nur für Punktlösungen zum Ziel führt. Matrizen, die linear voneinander abhängige Gleichungen beschreiben, können nicht invertiert werden, da ihre Determinante verschwindet. Was wir unter dem Begriff *lineare Abhängigkeit* genauer verstehen, untersuchen wir im nächste Kapitel 3.3.3, was uns in weitere Folge auf den *Rang einer Matrix* führt (Kapitel 3.3.4), der den Grad der linearen Abhängigkeit ausdrückt.

Zuvor lernen wir jedoch noch einige weitere Begriffe kennen, die in Zusammenhang mit Matrizen wichtig sind.

Spur einer Matrix. Unter der Spur einer Matrix (englisch: trace) versteht man die Summe ihrer Diagonalterme, und wir schreiben:

$$\text{Sp } \mathbf{A} \equiv \text{tr } \mathbf{A} = \sum_i a_{ii}.$$

Transponierte Matrix. Die transponierte Matrix entspricht der um die Diagonale gespiegelten Matrix. Wir kennzeichnen die transponierte Matrix durch ein hochgestelltes "T". Nach dieser Definition gilt also:

$$(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji}.$$

Symmetrische Matrix. Bei einer symmetrischen Matrix ist die Matrix gleich der transponierten Matrix

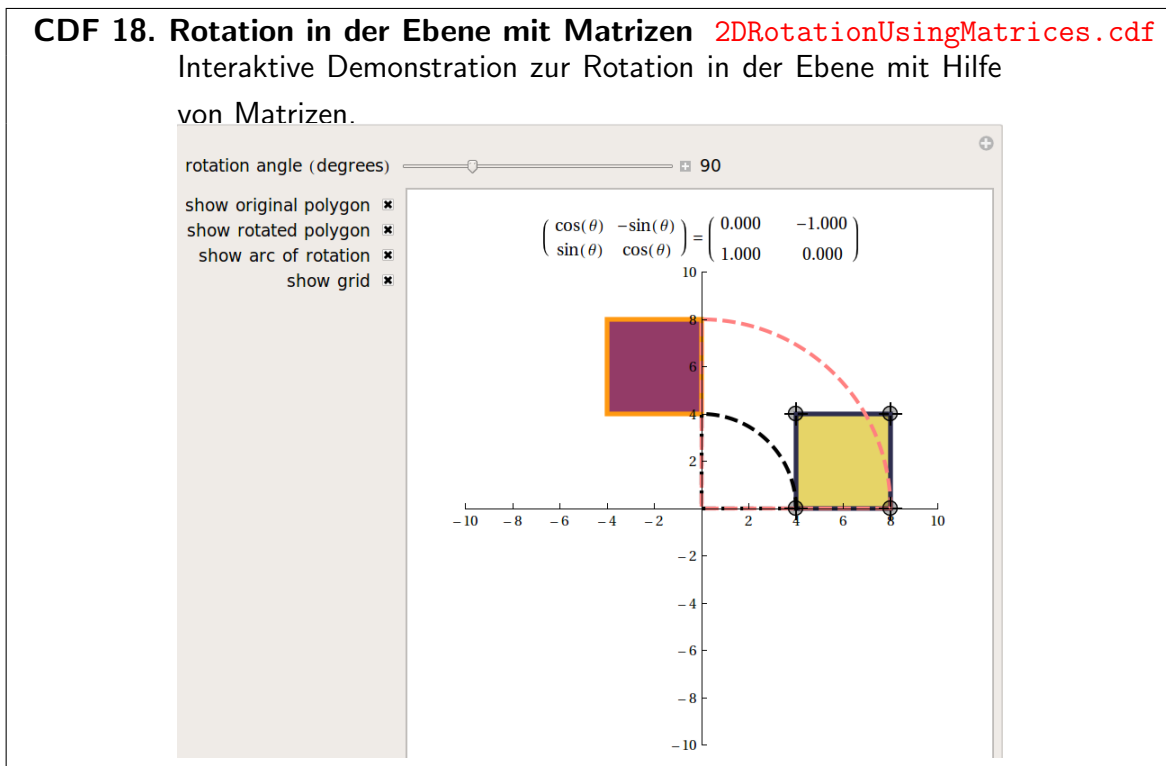
$$\mathbf{A}^T = \mathbf{A} \quad \Leftrightarrow \quad A_{ij} = A_{ji}.$$

Orthogonale Matrix. Wir bezeichnen eine Matrix \mathbf{A} als orthogonal, wenn ihre transponierte Matrix gleich ihrer inversen Matrix ist, also $\mathbf{A}^T = \mathbf{A}^{-1}$. Für orthogonale Matrizen gilt somit

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^T = \mathbf{1}.$$

Ein Beispiel für orthogonale Matrizen sind Drehmatrizen $\mathbf{R}(\varphi)$, die etwa für Drehungen um einen Winkel φ in der Ebene folgende Form haben

$$\mathbf{R}(\alpha) = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}.$$



Diagonalmatrix. Eine quadratische Matrix \mathbf{A} nennen wir diagonal, wenn alle Nichtdiagonalelemente verschwinden, also $(\mathbf{A})_{ij} = 0$ für $i \neq j$.

Singuläre Matrix. Eine Matrix \mathbf{A} , deren Determinante verschwindet, $\det \mathbf{A} = 0$, nennen wir singular.

3.3.3 Lineare Abhängigkeit

Eine Familie von Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ nennen wir *linear unabhängig*, wenn sich der Nullvektor nur durch eine Linearkombination der Vektoren erzeugen lässt, in der alle Koeffizienten α_i der Kombination auf den Wert Null gesetzt werden. (oder: Die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sind genau dann linear unabhängig, wenn sich keiner von ihnen als Linearkombination der anderen darstellen lässt.)

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0}. \quad (3.50)$$

Anders ausgedrückt gilt, die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sind genau dann *linear abhängig*, wenn es Koeffizienten α_i gibt, von denen mindestens einer ungleich 0 ist, so dass Gleichung 3.50 erfüllt ist. Es gilt weiters: In einem d -dimensionalen Vektorraum ist eine Familie aus mehr als d Vektoren immer linear abhängig. Beispielsweise sind 3 Vektoren in der Ebene ($d = 2$) sicher linear abhängig, d.h. einer der 3 Vektoren kann als Linearkombination der beiden anderen dargestellt werden.

Beispiel. Wir zeigen, dass die Vektoren

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

aus \mathbb{R}^2 linear unabhängig sind. Beweis: Für $\alpha, \beta \in \mathbb{R}$ gelte

$$\alpha \mathbf{u} + \beta \mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} \alpha - 3\beta \\ \alpha + 2\beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also

$$\alpha - 3\beta = 0 \quad \wedge \quad \alpha + 2\beta = 0.$$

Dieses Gleichungssystem ist nur für die Lösung $\alpha = 0, \beta = 0$ (die sogenannte triviale Lösung) erfüllt; d. h. \mathbf{u} und \mathbf{v} sind linear unabhängig.

Beispiel. Auf analoge Weise zeigen wir, dass die Vektoren

$$\mathbf{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} -6 \\ 4 \end{pmatrix}$$

aus \mathbb{R}^2 linear *abhängig* sind. Beweis: Für $\alpha, \beta \in \mathbb{R}$ gelte

$$\alpha \mathbf{u} + \beta \mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad \alpha \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \beta \begin{pmatrix} -6 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} 3\alpha - 6\beta \\ -2\alpha + 4\beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also

$$3\alpha - 6\beta = 0 \wedge -2\alpha + 4\beta = 0.$$

Dieses Gleichungssystem hat für jedes $\alpha, \beta \in \mathbb{R}$, das die Gleichung $\alpha = 2\beta$ erfüllt, eine nichttriviale Lösung. Damit sind \mathbf{u} und \mathbf{v} linear abhängig.

Lineare Hülle. Die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ seien Elemente eines N -dimensionalen Vektorraums (z.B. $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{R}^N$), dann nennen wir die Menge aller Linearkombinationen

$$\{\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n \mid \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}\},$$

die *lineare Hülle* der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Die *Dimension* der linearen Hülle ist definiert als die Anzahl d der linear unabhängigen Vektoren aus der Familie $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. In unseren obigen Beispielen ist die lineare Hülle der Vektoren $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ mit $\mathbf{u} = (1, 1)$ und $\mathbf{v} = (-3, 2)$ der gesamte \mathbb{R}^2 , somit die Dimension der linearen Hülle $d = 2$ gleich der Dimension des Vektorraums. In dem zweiten Beispiel waren die Vektoren $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ mit $\mathbf{u} = (3, -2)$ und $\mathbf{v} = (-6, 4)$ ja linear abhängig. Daher ist ihre lineare Hülle eine Gerade durch den Ursprung in Richtung $\mathbf{u} = (3, -2)$, und die Dimension dieser linearen Hülle ist $d = 1 < 2 = N$.

3.3.4 Rang einer Matrix

Wir betrachten nun die Zeilen einer Matrix als Vektoren und fragen uns, ob diese Zeilenvektoren einer Matrix linear unabhängig sind oder nicht? Für eine Matrix \mathbf{A} definiert man den Zeilenraum $ZR(\mathbf{A})$ als die lineare Hülle der Zeilenvektoren aus \mathbf{A} . Die Dimension des Zeilenraums bezeichnet man als Zeilenrang. Analog definiert man den Spaltenraum $SR(\mathbf{A})$ und den Spaltenrang durch die Spaltenvektoren. Man kann zeigen, dass für Matrizen mit Einträgen aus einem Körper (z.B. reelle Zahlen) der Zeilen- und Spaltenrang jeder Matrix gleich ist, und spricht deshalb vom (wohldefinierten) Rang der Matrix, $\text{Rang}(\mathbf{A})$.³

Beispiel. Wir bestimmen den Rang der folgenden Matrizen

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 4 \\ 0 & 10 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 6 & 4 \\ 0 & 3 & 2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 4 & -1 \end{pmatrix}.$$

³Der Rang einer Matrix ist auch gleich der Ordnung der größten nichtverschwindenden Unterdeterminante der Matrix.

Die drei Zeilenvektoren der Matrix \mathbf{A} , $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (0, 5, 3)$, und $\mathbf{v}_3 = (0, 10, 2)$, sind linear unabhängig, daher ist die Dimension ihrer linearen Hülle $d = 3$, und somit gilt $\text{Rang}(\mathbf{A}) = 3$. Für die Matrix \mathbf{B} sind zwei der drei Zeilenvektoren, $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (0, 6, 4)$, und $\mathbf{v}_3 = (0, 3, 2)$, linear abhängig ($\mathbf{v}_2 = 2\mathbf{v}_3$). Daher ist die Dimension d der linearen Hülle von $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ gleich 2 und es gilt $\text{Rang}(\mathbf{B}) = 2$. Schließlich betrachten wir die Matrix \mathbf{C} , die 3 Zeilen und 2 Spalten aufweist. Ihre drei Zeilenvektoren $\mathbf{v}_1 = (2, 3)$, $\mathbf{v}_2 = (0, 1)$, und $\mathbf{v}_3 = (4 - 1)$ sind daher Elemente eines zweidimensionalen Vektorraums (\mathbb{R}^2) und die Dimension ihrer linearen Hülle beträgt damit maximal $d \leq N = 2$. Tatsächlich sind in diesem Beispiel zwei der drei Vektoren linear unabhängig, die Dimension der linearen Hülle ist somit 2 und es gilt $\text{Rang}(\mathbf{C}) = 2$. Hinweis: Wir hätten den Rang der Matrizen \mathbf{A} , \mathbf{B} , und \mathbf{C} ebenso bestimmen können, indem wir die *Spaltenvektoren* der Matrizen untersucht hätten. Auch auf diesem Weg gelangen wir zu den gleichen Ergebnissen: $\text{Rang}(\mathbf{A}) = 3$, $\text{Rang}(\mathbf{B}) = 2$, und $\text{Rang}(\mathbf{C}) = 2$.

Quadratische $n \times n$ Matrizen \mathbf{A} nennen wir *regulär*, wenn ihr Rang gleich n ist (gleichbedeutend mit $\det \mathbf{A} \neq 0$), und *singulär*, wenn ihr Rang kleiner als n ist (gleichbedeutend mit $\det \mathbf{A} = 0$).

CDF 19. Eigenschaften einer 3x3 Matrix `3x3MatrixExplorer.cdf`
 Interaktive Demonstration zu Eigenschaften einer 3×3 -Matrix.

<u>Matrix</u>	<u>Transpose</u>	<u>Inverse</u>
$\begin{pmatrix} -1 & 3 & 2 \\ 1 & 1 & 3 \\ 3 & 2 & 2 \end{pmatrix}$	$\begin{pmatrix} -1 & 1 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 2 \end{pmatrix}$	$\begin{pmatrix} -0.17 & -0.09 & 0.30 \\ 0.30 & -0.35 & 0.22 \\ -0.04 & 0.48 & -0.17 \end{pmatrix}$
<u>Trace Determinant Rank</u>		
2 23 3		
<u>Eigenvalues</u>		<u>Eigenvectors</u>
$\begin{pmatrix} 5.59 \\ -1.80 + 0.94 i \\ -1.80 - 0.94 i \end{pmatrix}$		$\begin{pmatrix} 0.46 & 0.55 & 0.69 \\ 0.64 & 0.17 - 0.43 i & -0.50 + 0.35 i \\ 0.64 & 0.17 + 0.43 i & -0.50 - 0.35 i \end{pmatrix}$

integer entries in matrix

Unser Ausgangspunkt in diesem Kapitel war ja die Lösung von linearen Gleichungssystemen

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}.$$

Mithilfe des hier eingeführten Begriffs des Ranges einer Matrix können wir nun systematisch angeben, in welchem Fall ein Gleichungssystem mit m Gleichungen für n Unbekannte, $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ lösbar ist, und

wenn ja, welcher Art die zu erwartende Lösung ist bzw. welcher Art die zu erwartenden Lösungen sind. Dazu betrachten wir eine Matrix \mathbf{M} , die wir aus der $m \times n$ Matrix \mathbf{A} und dem Vektor \mathbf{b} in folgender Weise bilden:

$$\mathbf{M} = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} \quad (3.51)$$

Mit Hilfe des Rangs, $\text{Rang}(\mathbf{M})$, dieser $m \times (n+1)$ -Matrix können wir nun folgende Aussagen machen:

- $\text{Rang}(\mathbf{M}) = \text{Rang}(\mathbf{A}) = n$: Es gibt genau eine Punktlösung.
- $\text{Rang}(\mathbf{M}) > \text{Rang}(\mathbf{A})$: Die Gleichungen sind widersprüchlich, es gibt keine Lösung.
- $\text{Rang}(\mathbf{M}) = \text{Rang}(\mathbf{A}) = r < n$: Es gibt eine Schar von Lösungen, die durch $(n-r)$ Variablen parametrisierbar ist. Oder anders ausgedrückt: Die Lösungen spannen einen $(n-r)$ -dimensionalen Vektorraum auf.

Beispiel 1: Wir haben zwei Gleichungen mit zwei Unbekannten

$$\begin{aligned} x - 2y &= 1 \\ -2x + 4y &= 2 \end{aligned}$$

Es gilt:

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & 2 \end{pmatrix},$$

und wegen $\text{Rang}(\mathbf{M}) = 2 > 1 = \text{Rang}(\mathbf{A})$ hat dieses Gleichungssystem keine Lösung. Geometrisch betrachtet handelt es sich um 2 parallele Geraden, die nicht durch denselben Punkt verlaufen.

Beispiel 2: Wir betrachten wieder zwei Gleichungen mit zwei Unbekannten mit leicht verändertem Vektor \mathbf{b}

$$\begin{aligned} x - 2y &= -1 \\ -2x + 4y &= 2 \end{aligned}$$

Nun gilt

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 4 & 2 \end{pmatrix},$$

und wegen $\text{Rang}(\mathbf{M}) = \text{Rang}(\mathbf{A}) = 1$ hat dieses Gleichungssystem eine $(n-r) = 2-1 = 1$ parametrische Lösungsschar. Wir berechnen beispielsweise aus der ersten Gleichung $x = t$, $y(t) = \frac{1}{2} + \frac{t}{2}$. Geometrisch betrachtet handelt es sich um 2 parallele und deckungsgleiche Geraden.

Beispiel 3: Wir betrachten zum Abschluss noch ein System von drei Gleichungen für drei Unbekannte

$$\begin{aligned}x + y + z &= 1 \\x - 2y - 3z &= 3 \\2x - y - 2z &= 4\end{aligned}$$

Hier ist

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & -3 \\ 2 & -1 & -2 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -2 & -3 & 3 \\ 2 & -1 & -2 & 4 \end{pmatrix}.$$

Wegen $\text{Rang}(\mathbf{M}) = \text{Rang}(\mathbf{A}) = 2$ hat dieses Gleichungssystem eine $(n-r) = 3-2 = 1$ parametrische Lösungsschar. Wir berechnen beispielsweise aus den ersten beiden Gleichung $x(t) = t, y(t) = 6 - 4t, z(t) = -5 + 3t$. Geometrisch betrachtet handelt es um drei Ebenen, die eine gemeinsame Schnittgerade aufweisen.

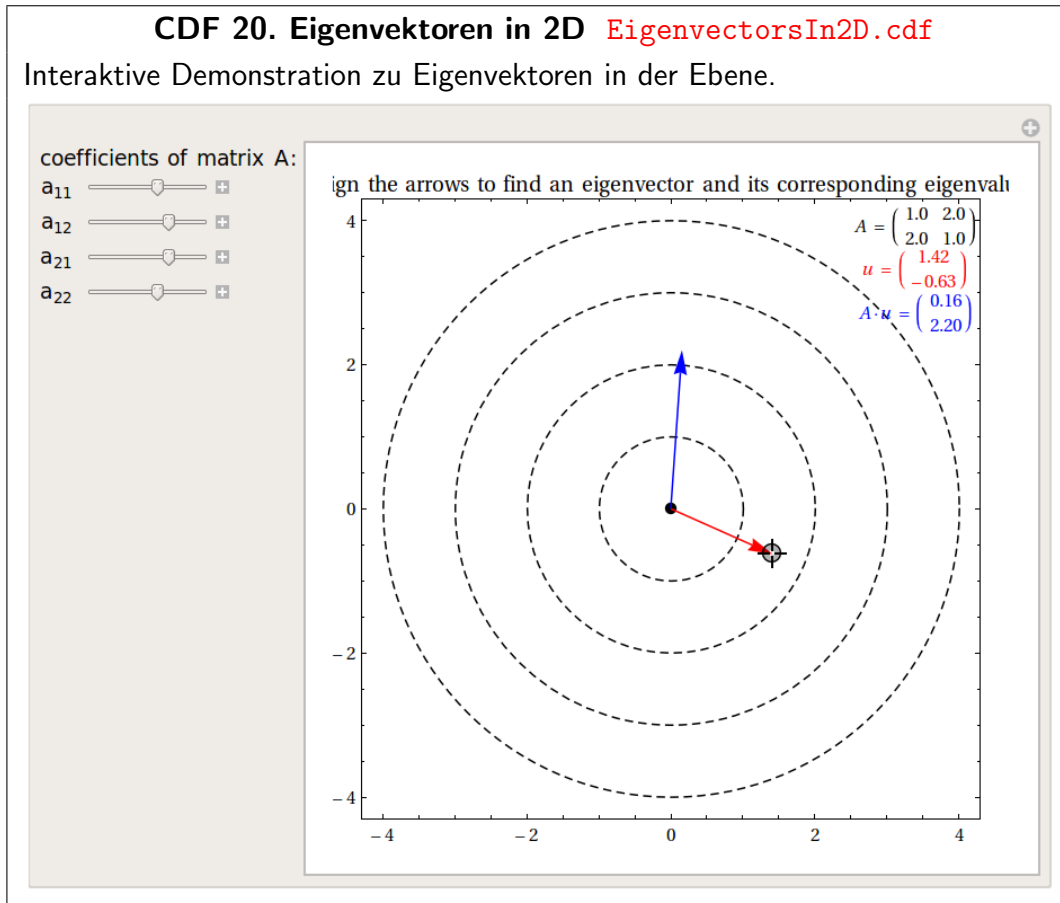
3.4 Das Eigenwertproblem

Zum Abschluss des Themas "Lineare Algebra" beschäftigen wir uns mit dem sogenannten Eigenwertproblem. Ein *Eigenvektor* einer Matrix \mathbf{A} ist ein vom Nullvektor verschiedener Vektor \mathbf{v} , dessen Richtung durch die Multiplikation mit der Matrix nicht verändert wird.

$$\mathbf{A} \cdot \mathbf{v} = \lambda \mathbf{v}. \tag{3.52}$$

Man bezeichnet den Proportionalitätsfaktor λ als Eigenwert der Abbildung. Eigenwertgleichungen wie in 3.52 definiert, spielen in vielen Teilgebieten der Physik eine wichtige Rolle. In der Mechanik etwa beschreiben die Eigenwerte der dynamischen Matrix die möglichen Frequenzen eines schwingungsfähigen Systems, oder die Eigenwerte und Eigenvektoren beschreiben die Hauptträgheitsmomente und -achsen eines starren Körpers. Eigenwerte spielen auch in der Quantenmechanik eine besondere Rolle. Physikalische Größen wie z.B. der Drehimpuls werden hier durch Operatoren repräsentiert. Messbar sind nur die Eigenwerte der Operatoren. Hat z. B. der Hamiltonoperator, der die Energie eines quantenmechanischen Systems repräsentiert, ein diskretes Spektrum, so kann die Energie nur diskrete Werte

annehmen, was z. B. für die Energieniveaus in einem Atom typisch ist. So stellen bei den Lösungen der bekannten Schrödingergleichung (im Jahr 1926 durch den Physiker Erwin Schrödinger aufgestellt) die Eigenwerte die erlaubten Energiewerte der Elektronen und die Eigenfunktionen die zugehörigen Wellenfunktionen der Elektronen dar.



Charakteristische Gleichung. Bringen wir in der Eigenwertgleichung 3.52 den Term $\lambda \mathbf{v}$ auf die linke Seite, so erhalten wir ein *homogenes*, lineares Gleichungssystem

$$(\mathbf{A} - \lambda \mathbf{1}) \cdot \mathbf{v} = \mathbf{0}. \quad (3.53)$$

Aus den vorangegangenen Kapiteln wissen wir, dass nichttriviale Lösungen ($\mathbf{v} \neq \mathbf{0}$) nur dann existieren, wenn die Determinante der Koeffizientenmatrix verschwindet.

$$\det(\mathbf{A} - \lambda \mathbf{1}) = 0. \quad (3.54)$$

Für eine $n \times n$ Matrix \mathbf{A} führt diese Bedingung auf ein Polynom n -ten Grades in λ , das man als das *charakteristische Polynom* des Eigenwertproblems bezeichnet. Da ein Polynom n -ten Grades im allge-

meinen n Nullstellen besitzt, gibt es im allgemeinen auch n Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n$ und dazugehörige n Eigenvektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

Beispiel. Wir wollen die Eigenwerte und Eigenvektoren der folgenden 2×2 Matrix bestimmen

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$$

Die charakteristische Gleichung und deren Lösungen lauten

$$\det \begin{pmatrix} 1 - \lambda & 2 \\ -1 & 3 - \lambda \end{pmatrix} = 0 \quad \Rightarrow \quad \lambda^2 - 4\lambda + 5 = 0 \quad \Rightarrow \quad \lambda_{1,2} = 2 \pm i.$$

Die dazugehörigen Eigenvektoren erhalten wir, indem wir die Eigenvektoren $\lambda_1 = 2 + i$ bzw. $\lambda_2 = 2 - i$ in die Eigenwertgleichung einsetzen, und wir erhalten

$$\mathbf{v}_1 = \begin{pmatrix} 1 - i \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 + i \\ 1 \end{pmatrix}$$

Es ist zu beachten, dass die Eigenwertgleichung nur die Richtung der Eigenvektoren festlegt, nicht aber deren Länge, oder anders ausgedrückt, es gibt eine 1-parametrische Schar von Lösungen. Wir sehen auch, dass die Eigenwerte und Eigenvektoren im allgemein komplexe Größen sind.

In einem zweiten Beispiel berechnen wir die Eigenwerte und Eigenvektoren der folgenden 2×2 Matrix

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Die charakteristische Gleichung und deren Lösungen lauten

$$\det \begin{pmatrix} 3 - \lambda & -1 \\ -1 & 3 - \lambda \end{pmatrix} = 0 \quad \Rightarrow \quad (3 - \lambda)^2 - 1 = 0 \quad \Rightarrow \quad \lambda_{1,2} = 3 \pm 1.$$

Die dazugehörigen Eigenvektoren erhalten wir, indem wir die Eigenvektoren $\lambda_1 = 2$ bzw. $\lambda_2 = 4$ in die Eigenwertgleichung einsetzen, und wir erhalten

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

In diesem Beispiel erhalten wir *reelle* Eigenwerte und Eigenvektoren. Zusätzlich sehen wir, dass die beiden *Eigenvektoren normal aufeinander* stehen, d.h. $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Das ist kein Zufall, sondern liegt

daran, dass es sich bei der Matrix \mathbf{A} in diesem Beispiel um eine *symmetrische* Matrix handelt, d.h. $A_{ij} = A_{ji}$.

Die Eigenwerte einer symmetrischen Matrix sind stets reell. Wir beweisen diesen Sachverhalt, indem wir zunächst annehmen, dass der Eigenwert λ einen Realteil α und Imaginärteil β besitzt, und auch der Eigenvektor \mathbf{v} in Real und Imaginärteil aufgespalten werden kann:

$$\begin{aligned}\lambda &= \alpha + i \cdot \beta & (\alpha, \beta \in \mathbb{R}) \\ \mathbf{v} &= \vec{p} + i \cdot \vec{q} & (\vec{p}, \vec{q} \in \mathbb{R}^3)\end{aligned}$$

Trennen wir die Eigenwertgleichung

$$\mathbf{A} \cdot \mathbf{v} = \lambda \mathbf{v} \quad \Leftrightarrow \quad \mathbf{A} \cdot (\mathbf{p} + i\mathbf{q}) = (\alpha + i\beta)(\mathbf{p} + i\mathbf{q})$$

in Real- und Imaginärteil auf, so erhalten wir die zwei Gleichungen

$$\begin{aligned}\mathbf{A} \cdot \mathbf{p} &= \alpha \mathbf{p} - \beta \mathbf{q} \\ \mathbf{A} \cdot \mathbf{q} &= \alpha \mathbf{q} + \beta \mathbf{p}\end{aligned}$$

Multiplizieren wir die erste Gleichung von links mit einem Zeilenvektor \mathbf{q} bzw. die zweite Gleichung von links mit dem Zeilenvektor \mathbf{p} , so erhalten wir

$$\begin{aligned}\mathbf{q} \cdot \mathbf{A} \cdot \mathbf{p} &= \alpha \mathbf{q} \cdot \mathbf{p} - \beta \mathbf{q} \cdot \mathbf{q} \\ \mathbf{p} \cdot \mathbf{A} \cdot \mathbf{q} &= \alpha \mathbf{p} \cdot \mathbf{q} + \beta \mathbf{p} \cdot \mathbf{p}\end{aligned}$$

Weil für einen symmetrischen Tensor $A_{nm} = A_{mn}$ gilt, ergibt die Subtraktion der ersten von der zweiten Gleichung

$$0 = 0 + 2\beta(\mathbf{p} \cdot \mathbf{p} + \mathbf{q} \cdot \mathbf{q}) \quad \Rightarrow \quad \beta \cdot (p^2 + q^2) = 0.$$

Da der Betrag des Vektors $\vec{v} \neq 0$ ist, gilt sicher auch $p^2 + q^2 > 0$, und damit folglich, dass $\beta = 0$ sein muss. Damit ist gezeigt, dass der Imaginärteil des Eigenwerts verschwindet, und der Eigenwert somit rein reell ist.

Die zu verschiedenen Eigenwerten gehörigen Eigenvektoren einer symmetrischen Matrix sind stets orthogonal. Zum Beweis dieses Satzes nehmen wir an, λ_1 und λ_2 seien zwei verschiedene Eigenwerte ($\lambda_1 \neq \lambda_2$) und \mathbf{v}_1 und \mathbf{v}_2 seien die zugehörigen Eigenvektoren der symmetrischen Matrix

A. Dann gilt

$$\mathbf{A} \cdot \mathbf{v}_1 = \lambda_1 \mathbf{v}_1$$

$$\mathbf{A} \cdot \mathbf{v}_2 = \lambda_2 \mathbf{v}_2$$

Multiplizieren wir die erste dieser Eigenwertgleichungen von links mit dem Zeilenvektor \mathbf{v}_2 und die zweite Gleichung mit \mathbf{v}_1 , so erhalten wir

$$\mathbf{v}_2 \cdot \mathbf{A} \cdot \mathbf{v}_1 = \lambda_1 \mathbf{v}_2 \cdot \mathbf{v}_1$$

$$\mathbf{v}_1 \cdot \mathbf{A} \cdot \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2$$

Unter Berücksichtigung der Symmetrie von $A_{nm} = A_{mn}$ ergibt die Subtraktion dieser Gleichungen

$$0 = (\lambda_1 - \lambda_2) \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Da wir vorausgesetzt haben, dass $\lambda_1 \neq \lambda_2$ ist, muss daher $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ sein, was gleichbedeutend mit der zu beweisenden Aussage ist, dass \mathbf{v}_1 orthogonal auf \mathbf{v}_2 steht.

Kapitel 4

Differenzialrechnung

Die Differenzialrechnung geht zurück auf Isaac Newton (1643 – 1727) und auf Gottfried Wilhelm Leibniz (1646 – 1716), siehe zum Beispiel <http://de.wikipedia.org/wiki/Differentialrechnung> für einen kurzen historischen Überblick. Während Newton's Motivation für die Entwicklung der Differenzialrechnung aus der Berechnung der *Momentangeschwindigkeit* von Körpern herrührte, suchte Leibniz nach einer Möglichkeit *Tangenten* an beliebige Kurven legen zu können.

4.1 Die lineare Näherung

4.1.1 Differenzenquotient und Differenzialquotient

Wir wollen uns der Thematik zunächst nach dem Leibniz'schen Zugang nähern. Wenn wir für eine Funktion $f(x)$ an einer Stelle x_0 eine Tangente berechnen wollen, so können wir diese näherungsweise beschreiben als eine *Sekante* durch die Punkte $\{x_0, f(x_0)\}$ und $\{x_0 + \Delta x, f(x_0 + \Delta x)\}$. Die Steigung dieser Sekante ist dann gegeben als Quotient der Differenzen auf der Ordinatenachse, Δy , und der Differenzen auf der Abszissenachse, Δx .

$$\tan \alpha = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

Dieser Ausdruck wird als *Differenzenquotient* bezeichnet. Macht man nun Δx immer kleiner, so wird aus der Sekante im Grenzübergang $\Delta x \rightarrow 0$ eine *Tangente* mit der Steigung

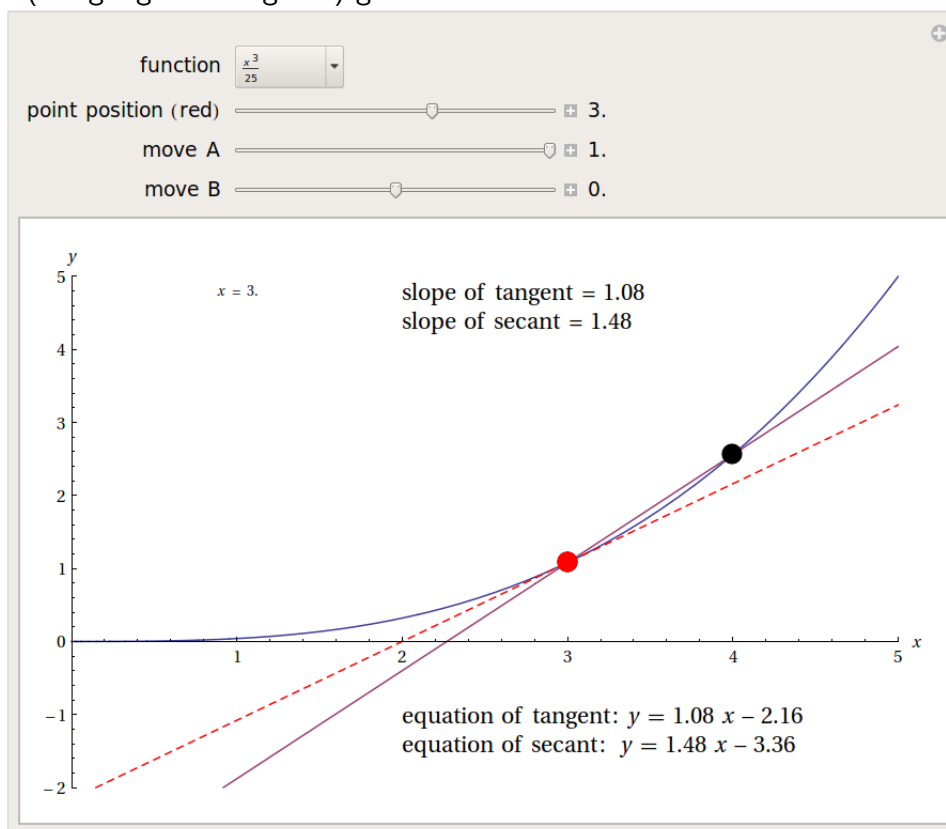
$$\tan \alpha = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

Wir bezeichnen diesen Ausdruck nun als *Differenzialquotient* oder auch als *1. Ableitung* der Funktion $f(x)$ and der Stelle x_0 , und schreiben

$$f'(x_0) = \left. \frac{df}{dx} \right|_{x=x_0} = \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta f}{\Delta x} \right)_{x=x_0}$$

CDF 21. Differenzialquotient Differenzial_vs_Differenzen.cdf

Mit dieser CDF-Anwendung können Sie den Differenzenquotienten (Steigung der Sekante) bzw. den Differenzialquotienten (Steigung der Tangente) grafisch darstellen.



Beispiele. Betrachten wir die Funktion $f(x) = x^2$, und setzen in die Definition des Differenzialquotienten ein, so erhalten wir

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} 2x + \lim_{\Delta x \rightarrow 0} \Delta x = 2x$$

Mit einer ähnlichen Rechnung erhalten wir für allgemeine Potenzen $f(x) = x^n$

$$f(x) = x^n \Rightarrow f'(x) = nx^{n-1}.$$

In ähnlicher Weise können wir auch die Ableitung der Exponentialfunktion $f(x) = e^x$ über die Definition des Differentialquotienten bestimmen:

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{e^{x+\Delta x} - e^x}{\Delta x} = e^x \underbrace{\lim_{\Delta x \rightarrow 0} \frac{e^{\Delta x} - 1}{\Delta x}}_{=1} = e^x$$

Was wir in obiger Gleichung zeigen müssen ist, dass der Grenzwert tatsächlich den Wert 1 hat. Mit unserem Wissen der Taylorreihendarstellung der Exponentialfunktion oder der Regel von De L'Hospital wäre das natürlich ein Leichtes. Beide Methoden setzen aber die Kenntnis der Ableitung der Exponentialfunktion voraus, was wir vermeiden wollen. Wir können uns aber folgender Abschätzung der Exponentialfunktion bedienen, die auf jedem Fall in dem Intervall $x \in [-1, 1]$ erfüllt ist.

$$1 + x \leq e^x \leq \frac{1}{1 - x}$$

Der linke Teil der Abschätzung folgt aus der Folgendarstellung der Exponentialfunktion und der Bernoulli-schen Ungleichung $(1 + x)^n \geq 1 + nx$. Der rechte Teil der Ungleichung folgt ebenso aus der Bernoulli'schen Ungleichung kombiniert mit der Tatsache, dass für ein $u < 1$ gilt:

$$1 + u \leq \frac{1}{1 - u}$$

Formen wir die obige Abschätzung der Exponentialfunktion nach unten und oben geeignet um, so können wir den für die 1. Ableitung gebrauchten Grenzwert ermitteln:

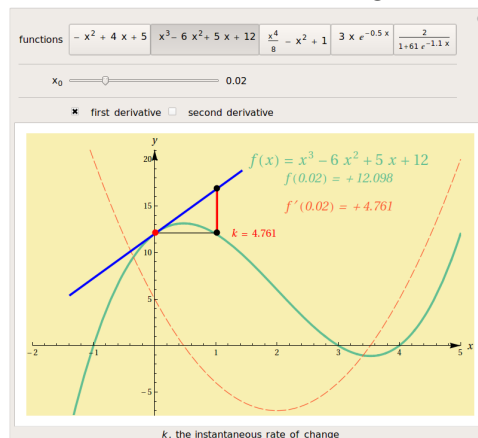
$$\begin{aligned} 1 + u &\leq e^u &\leq \frac{1}{1-u} &| - 1 \\ u &\leq e^u - 1 &\leq \frac{u}{1-u} &| : u \\ 1 &\leq \frac{e^u - 1}{u} &\leq \frac{1}{1-u} &| \lim_{u \rightarrow 0} \\ 1 &\leq \lim_{u \rightarrow 0} \frac{e^u - 1}{u} &\leq 1 & \end{aligned}$$

Somit haben wir gezeigt, dass die 1. Ableitung der Exponentialfunktion wiederum die Exponentialfunktion ergibt. Auf ähnliche Art und Weise können die Ableitungsregeln für weitere elementare Funktionen aus der Definition des Differenzialquotienten gewonnen werden:

$f(x)$	$f'(x)$	Einschränkung
a	0	
x	1	
x^b	bx^{b-1}	
$\frac{1}{a+bx}$	$-\frac{b}{(a+bx)^2}$	
e^x	e^x	
e^{ax}	ae^{ax}	
$a^x = e^{x \ln a}$	$a^x \ln a$	$a > 0$
$\ln x$	$\frac{1}{x}$	$x > 0$
$x(\ln x - 1)$	$\ln x$	$x > 0$
$\sin x$	$\cos x$	
$\cos x$	$-\sin x$	
$\tan x$	$\frac{1}{(\cos x)^2} = 1 + (\tan x)^2$	
$\cot x$	$-\frac{1}{(\sin x)^2} = -1 - (\cot x)^2$	
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$-1 < x < +1$
$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$	$-1 < x < +1$
$\arctan x$	$\frac{1}{1+x^2}$	
$\text{arccot } x$	$-\frac{1}{1+x^2}$	
$\sinh x$	$\cosh x$	
$\cosh x$	$\sinh x$	
$\text{arsinh } x$	$\frac{1}{\sqrt{1+x^2}}$	
$\text{arcosh } x$	$\frac{1}{\sqrt{x^2-1}}$	$x > 1$
$\text{artanh } x$	$\frac{1}{1-x^2}$	$ x < 1$
$\text{arcoth } x$	$\frac{1}{1-x^2}$	$ x > 1$

CDF 22. 1. und 2. Ableitung [ErsteUndZweiteAbleitung.cdf](#)

Mit dieser CDF-Anwendung können Sie die 1. und 2. Ableitung einer Funktion grafisch darstellen.



4.1.2 Totales Differenzial

Da die Ableitung einer Funktion ein zentrales Thema dieser Vorlesung ist, betrachten wir die Thematik noch von einem anderen Blickwinkel. Und zwar können wir die Ableitungsfunktion einer Funktion $f(x)$ auch so definieren, dass wir mithilfe der Ableitung $f'(x)$ die beste *lineare Näherung* von $f(x)$ in einer Umgebung um die Stelle x angeben können. Das heißt wir können die Änderung der Funktion durch die Ableitung ausdrücken:

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \mathcal{O}((\Delta x)^2). \quad (4.1)$$

Hierbei fordern wir, dass der Fehler der linearen Näherung zumindest quadratisch mit der Größe Δx abnimmt. Beispielsweise heißt das, verkleinern wir den Abstand von x um einen Faktor 5, so nimmt der Fehler $\Delta f(x) = f(x + \Delta x) - f(x)$ zumindest um einen Faktor 25 kleiner werden. Existiert eine solche Funktion $f'(x)$, dann nennen wir sie die 1. Ableitung der Funktion $f(x)$, und die Funktion $f(x)$ ist an der Stelle x differenzierbar. Wir nennen nun diese lineare Näherung der Änderung der Funktion das **Totale Differenzial** das wir wie folgt anschreiben

$$df(x) = f'(x)\Delta x = f'(x)dx$$

Im letzten Schritt haben wir die endliche Differenz Δx durch die infinitesimale Differenz dx ersetzt.

Betrachten wir als Beispiel wieder die Funktion $f(x) = x^2$, dann lautet das totale Differenzial

$$df = 2x dx$$

An jeder beliebigen Stelle x gilt dann, dass die lineare Näherung für die Änderung der Funktion df die tatsächliche Änderung bis zur Ordnung $\mathcal{O}((\Delta x)^2)$ beschreibt.

$$\Delta f(x) = f(x + dx) - f(x) = (x + dx)^2 - x^2 = 2x dx + (dx)^2 = df + \mathcal{O}((dx)^2)$$

Betrachten wir hingegen die Betragsfunktion

$$f(x) = |x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases}$$

dann können wir an der Stelle $x = 0$ keine Zahl $f'(0)$ angeben, so dass sie die Änderung der Funktion in einer Umgebung $(-dx, dx)$ um den Punkt $x = 0$ in linearer Näherung beschreibt. Wir finden also, dass die Ableitung der Betragsfunktion an der Stelle $x = 0$ nicht existiert, oder anders ausgedrückt,

dass die Funktion an der Stelle $x = 0$ nicht differenzierbar ist.

$$f'(x) = \begin{cases} -1 & \text{falls } x < 0 \\ \text{nicht definiert} & \text{falls } x = 0 \\ +1 & \text{falls } x > 0 \end{cases}$$

Die Differenzierbarkeit einer Funktion können wir auch mithilfe des Differenzialquotienten definieren. Und zwar betrachten wir beim Grenzübergang getrennt den *rechtsseitigen* und den *linksseitigen* Grenzwert

$$f'(x^+) = \lim_{\Delta x \rightarrow 0^+} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

$$f'(x^-) = \lim_{\Delta x \rightarrow 0^-} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

Beim rechtsseitigen Grenzwert, $f'(x^+)$, ist $\Delta x > 0$ und wir nähern uns von der rechten Seite. Existiert dieser Grenzwert, so nennen wir die Funktion *rechtsseitig differenzierbar*. Analog bezeichnen wir eine Funktion *linksseitig differenzierbar*, wenn der Grenzwert $f'(x^-)$ für $\Delta x < 0$ existiert. Existieren beide Grenzwerte und gilt $f'(x^+) = f'(x^-)$, dann nennen wir die Funktion *differenzierbar*. Funktionen, deren Ableitungsfunktion stetig ist, nennen wir *stetig differenzierbar*. Dazu zwei Beispiele:

- Die Betragsfunktion $f(x) = |x|$ ist stetig. Ihre linksseitige Ableitung an der Stelle $x = 0$ ist -1 , die rechtsseitige Ableitung ist $+1$. Damit ist die Funktion an der Stelle $x = 0$ nicht differenzierbar.
- Die erste Ableitung der Funktion

$$f(x) = \begin{cases} 0 & \text{falls } x < 0 \\ x^2 & \text{falls } x \geq 0 \end{cases}$$

ist auf ganz \mathbb{R} stetig, damit ist $f(x)$ stetig differenzierbar. Die 2. Ableitung der Funktion ist allerdings an der Stelle $x = 0$ nicht definiert. Salopp gesprochen, weist die 1. Ableitung an der Stelle $x = 0$ einen Knick auf (genauso wie die Betragsfunktion).

Schreibweise. Wir fassen zusammen. Für die 1. Ableitung einer Funktion haben wir mehrere gleichbedeutende Schreibweisen kennengelernt

$$f'(x) \equiv \frac{d}{dx} f(x) \equiv \frac{df(x)}{dx} \equiv f^{(1)}(x)$$

Für die 2. bzw. n -te Ableitung schreiben wir

$$f''(x) \equiv \frac{d}{dx} \frac{d}{dx} f(x) \equiv \frac{d^2}{dx^2} f(x) \equiv \frac{d^2 f(x)}{dx^2} \equiv f^{(2)}(x)$$

$$f^{(n)}(x) \equiv \frac{d}{dx} \dots \frac{d}{dx} f(x) \equiv \frac{d^n}{dx^n} f(x) \equiv \frac{d^n f(x)}{dx^n}$$

4.1.3 Mittelwertsatz der Differenzialrechnung

Der Mittelwertsatz der Differenzialrechnung sagt aus, dass für eine Funktion $f(x)$, die auf dem Intervall $[a, b]$ stetig ist, und die auf dem offenen Intervall (a, b) auch differenzierbar ist gilt:

$$\exists \xi \in [a, b] : f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

In Worten bedeutet das, dass es zwischen den Intervallgrenzen a und b eine Zahl ξ geben muss, sodass die Steigung der Funktion an der Stelle ξ gleich der Steigung der Verbindungsgerade zwischen den Punkten $\{a, f(a)\}$ und $\{b, f(b)\}$ ist. Anhand einer Skizze können wir diesen Sachverhalt unmittelbar einsehen, solange wir fordern, dass die Funktion zwischen a und b keine Sprünge aufweist (Stetigkeit) und auch keine "Knicke" hat (Differenzierbarkeit).

Anwendungen. Mithilfe des Mittelwertsatzes können wir zeigen, dass eine Funktion im Intervall (a, b) *streng monoton wachsend* ist, wenn $\forall x \in (a, b)$ gilt: $f'(x) > 0$. Wählen wir $a \leq x_1 < x_2 \leq b$ und wenden den Mittelwertsatz auf x_1 und x_2 an, so finden wir

$$f(x_2) - f(x_1) = \underbrace{f'(\xi)}_{>0} \underbrace{(x_2 - x_1)}_{>0} > 0,$$

womit $f(x_2) > f(x_1)$, und damit die strenge Monotonie gezeigt ist.

Als zweite Anwendung des Mittelwertsatzes wollen wir die bereits bekannte Ungleichung

$$1 + x < e^x < \frac{1}{1 - x}, \quad \forall x \in (0, 1)$$

zeigen. Wir wenden den Mittelwertsatz auf ein ξ im Intervall $[0, x]$ an

$$e^\xi = \frac{e^x - e^0}{x - 0}, \quad \text{mit } 0 < \xi < x$$

Da die Exponentialfunktion streng monoton wachsend ist, gilt folgende Ungleichung

$$1 = e^0 < e^x = \frac{e^x - 1}{x} < e^x$$

Der linke Teil der Ungleichung führt auf $1 + x < e^x$ während der rechte Teil der Ungleichung auf $e^x < \frac{1}{1-x}$ umgeformt werden kann. Damit ist die obige Ungleichung bewiesen.

4.1.4 Ableitung der Umkehrfunktion

Die Umkehrfunktion $f^{-1}(x)$ einer Funktion $f(x)$ ist so definiert, dass die Anwendung der Funktion f^{-1} auf das Ergebnis der Funktion f wieder auf den ursprünglichen Funktionswert x zurückführt $f^{-1}(f(x)) = x$, oder anders ausgedrückt:

$$x \xrightarrow{f} y = f(x) \xrightarrow{f^{-1}} x$$

Die Umkehrfunktion bildet also y auf x ab. Damit können wir uns auch den Graphen der Funktion f^{-1} auch als Spiegelung an der Geraden $y = x$ vorstellen. Somit ist die Steigung der Umkehrfunktion gleich der reziproken Steigung der Funktion f

$$\tan \alpha = f'(x) = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{(f^{-1})'(y)}$$

Als Beispiel untersuchen wir die Funktion $f(x) = y = x^2$ deren Ableitung $f'(x) = \frac{dy}{dx} = 2x$ ist. Für die Umkehrfunktion finden wir $f^{-1}(y) = x = \sqrt{y}$, und somit erhalten wir für die 1. Ableitung $(f^{-1})'(y) = \frac{dx}{dy} = \frac{1}{2\sqrt{y}}$. Somit überprüfen wir

$$2x = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{\frac{1}{2\sqrt{y}}} = 2x.$$

Wir können den Zusammenhang zwischen den Ableitungen von Funktionen und ihrer Umkehrfunktionen also dazu benutzen, um eine der beiden Ableitungen durch die jeweils andere auszudrücken. Wollen wir etwa die Ableitung der Umkehrfunktion des Tangens berechnen, so finden wir

$$y = \tan x = \frac{\sin x}{\cos x}, \quad \frac{dy}{dx} = 1 + (\tan x)^2,$$

$$x = \arctan y, \quad \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{1 + (\tan x)^2} = \frac{1}{1 + y^2} \quad (4.2)$$

4.2 Funktionen mehrerer Variablen

4.2.1 Motivation

Bis jetzt hatten wir es mit Funktionen einer veränderlichen Variable zu tun. Beispiele für solche Funktionszusammenhänge in der Physik sind etwa der Ort s eines Massenpunktes zum Zeitpunkt t , also $s(t)$. Bekanntlich ist dann die Momentangeschwindigkeit die 1. Ableitung, $v(t) = \frac{ds}{dt}$, und die zweite Ableitung die Beschleunigung $a(t) = \frac{d^2s}{dt^2}$. Der zeitliche Verlauf der Bewegung wird dann über die Newton'sche Bewegungsgleichung beschrieben

$$m \frac{d^2s}{dt^2} = F(t).$$

Bei sehr vielen Fragestellungen in der Physik hängt aber die betrachtete physikalische Größe von mehr als einer unabhängigen Variable ab. Wollen wir etwa die Ausbreitung einer ebenen Welle beschreiben, so hängt die Amplitude u von dem Ort x und von der Zeit t ab, also $u(x, t)$. Die bekannte Form der Lösung $u(x, t) = u_0 e^{i(kx - \omega t)}$ erhalten wir als Lösung der Wellengleichung

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0.$$

In dieser sogenannten *partiellen Differenzialgleichung* wird ein Zusammenhang hergestellt zwischen der 2. Ableitung der Funktion $u(x, t)$ in Bezug auf die Zeit t und den Ort x , wobei c die Ausbreitungsgeschwindigkeit der Welle bezeichnet. Die Verwendung des Symbols " ∂ " anstelle des gewöhnlichen Zeichens " d " soll verdeutlichen, dass es sich hierbei um eine *partielle Ableitung* handelt. Im Laufe des Physikstudiums werden Ihnen noch zahlreiche weitere Fälle begegnen, die partielle Ableitungen von Funktionen mehrerer Veränderlicher erfordern. Beispielsweise folgt die Temperaturverteilung $T(x, y, z, t)$ als Funktion der Ortsposition (x, y, z) und der Zeit t der Wärmeleitungsgleichung, die wiederum eine partielle Differenzialgleichung darstellt

$$\frac{1}{\lambda} \frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2},$$

wobei λ den Wärmeleitkoeffizienten darstellt. Auch in anderen Teilen der Thermodynamik, in der Elektrodynamik und in der Quantenmechanik kommen wir ohne partielle Differenzialgleichungen nicht aus. Sie stellen jeweils einen Zusammenhang zwischen verschiedenen partiellen Ableitungen der gesuchten Funktion her, also Ableitungen in Bezug auf unterschiedliche unabhängige Variable.

$$\text{Elektrostatik: } \frac{\partial E_x(x, y, z)}{\partial x} + \frac{\partial E_y(x, y, z)}{\partial y} + \frac{\partial E_z(x, y, z)}{\partial z} = \frac{\rho(x, y, z)}{\varepsilon_0}$$

$$\text{Schrödinger-Gleichung: } i\hbar \frac{\partial \Psi(x, y, z, t)}{\partial t} = -\frac{\hbar^2}{2m} \left[\frac{\partial^2 \Psi(x, y, z, t)}{\partial x^2} + \frac{\partial^2 \Psi(x, y, z, t)}{\partial y^2} + \frac{\partial^2 \Psi(x, y, z, t)}{\partial z^2} \right]$$

4.2.2 Partielle Ableitung

Im Rahmen dieser Vorlesung werden Sie nicht lernen, wie man die Lösung zu den obigen Differenzialgleichungen findet, sie werden aber erfahren, worum es sich bei einer partiellen Ableitung handelt, und wie man sie bildet. Betrachten wir also zunächst eine Funktion f , die jedem Punkt in der Ebene $(x, y) \in \mathbb{R}^2$ eine reelle Zahl z zuordnet.

$$f : \mathbb{R}^2 \mapsto \mathbb{R} : (x, y) \mapsto z \equiv f(x, y)$$

Den Graphen einer solchen Funktion von zwei unabhängigen Variablen x und y können wir uns als Fläche über der (x, y) -Ebene vorstellen.

Beispiel. Wir veranschaulichen uns zunächst anhand eines Beispiels die Bedeutung der partiellen Ableitung. Dazu betrachten wir die Funktion

$$f(x, y) = \sin(x - y) + \cos x.$$

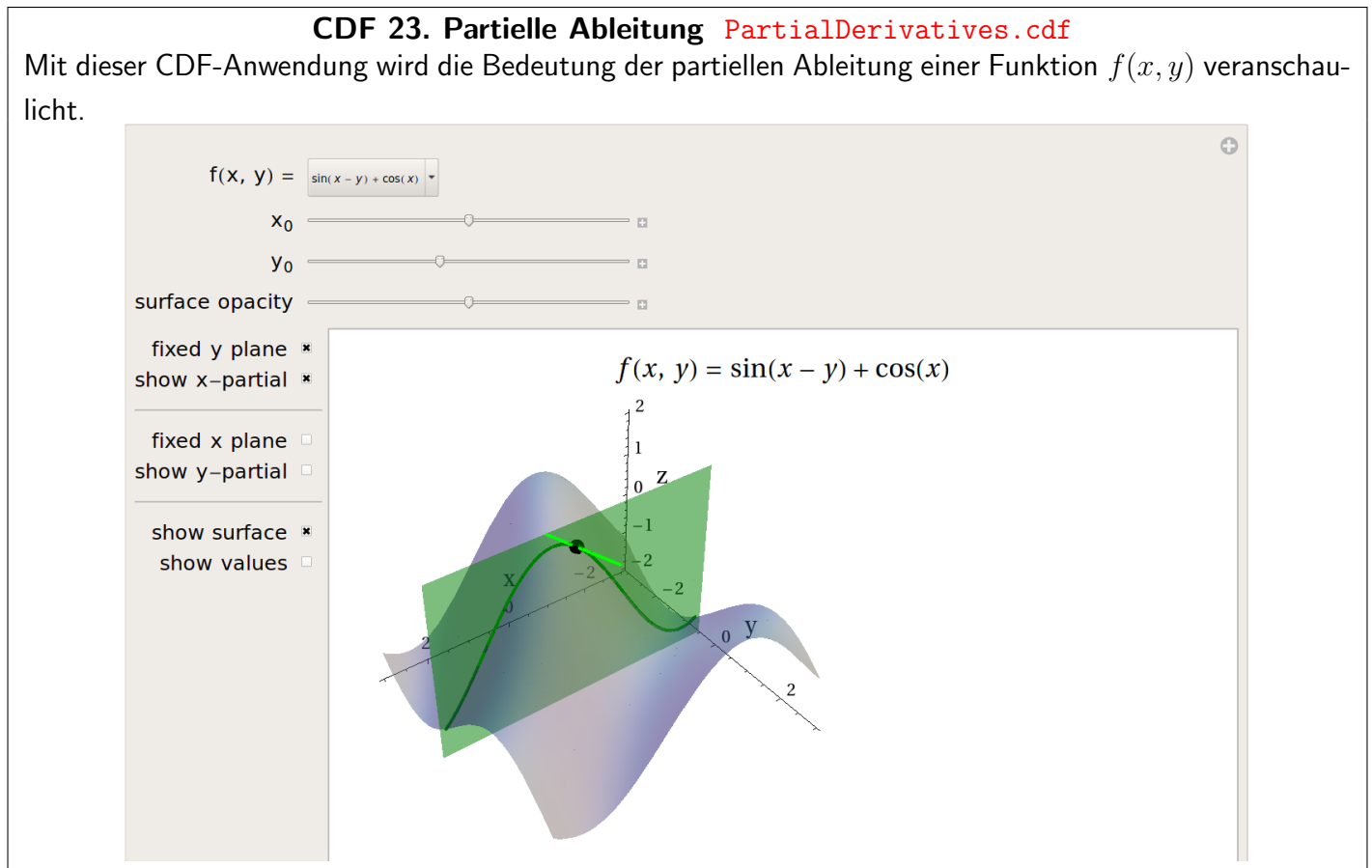
Halten wir die y -Koordinate fest und setzen sie auf den Wert y_0 , so haben wir es nur mehr mit einer "gewöhnlichen" Funktion von einer Variable x zu tun. Und genau das verstehen wir unter *partieller Ableitung*: Wir bilden die Ableitung der Funktion nach einer Variablen, wobei wir alle anderen Variablen – in unserem Beispiel nur y – festhalten

$$\frac{\partial}{\partial x} f(x, y_0) \equiv \frac{\partial f(x, y_0)}{\partial x} \equiv \partial_x f(x, y_0) \equiv f_x(x, y_0) = \cos(x - y_0) - \sin x.$$

Wir haben hier vier verschiedene Schreibweisen eingeführt, die allesamt denselben Sachverhalt ausdrücken. Die ersten drei (von links gesehen) werden vorwiegend in der Physik, die vierte, $f_x(x, y_0)$, ist vor allem in der Mathematik sehr gebräuchlich. Anschaulich können wir uns die partielle Ableitung nach x vorstellen, indem wir die Fläche $z = f(x, y)$ mit der Ebene $y = y_0$ schneiden und dann die entstandene Schnittkurve entlang der x -Achse ableiten. Analog finden wir die partielle Ableitung nach y , indem wir die Fläche $z = f(x, y)$ mit der Ebene $x = x_0$ schneiden, und dann die erhaltene Kurve entlang der y -Achse differenzieren. In der Rechnung halten wir also $x = x_0$ fest und bekommen:

$$\frac{\partial}{\partial y} f(x_0, y) \equiv \frac{\partial f(x_0, y)}{\partial y} \equiv \partial_y f(x_0, y) \equiv f_y(x_0, y) = -\cos(x_0 - y).$$

Zur Übung haben wir wieder die vier möglichen Schreibweisen angegeben. In Zukunft beschränken wir uns natürlich auf eine davon.



Beispiel. Wir berechnen alle möglichen 1. und 2. Ableitungen der Funktion $f(x, y) = x^2y^3 + e^{-(x^2+y^2)}$

$$\begin{aligned}
 f_x &= 2xy^3 - 2xe^{-(x^2+y^2)} \\
 f_y &= 3x^2y^2 - 2ye^{-(x^2+y^2)} \\
 f_{xx} &= 2y^3 + (4x^2 - 2)e^{-(x^2+y^2)} \\
 f_{xy} &= 6xy^2 + 4xye^{-(x^2+y^2)} \\
 f_{yx} &= 6xy^2 + 4xye^{-(x^2+y^2)} \\
 f_{yy} &= 6x^2y + (4y^2 - 2)e^{-(x^2+y^2)}
 \end{aligned}$$

Einer kompakteren Schreibweise willen haben wir hier darauf verzichtet, die jeweils festgehaltene Variable auf y_0 bzw. x_0 umzubenennen. Weiters bemerken wir, dass die beiden gemischten 2. Ableitungen f_{xy} und f_{yx} das gleiche Ergebnis liefern. Ganz allgemein gilt:

Satz von Schwarz. Sind die gemischten partiellen Ableitungen f_{xy} und f_{yx} einer Funktion f an der Stelle (x, y) stetig, dann kann man die Reihenfolge der Ableitungen vertauschen:

$$f_{xy}(x, y) = f_{yx}(x, y) \iff \frac{\partial^2 f(x, y)}{\partial x \partial y} = \frac{\partial^2 f(x, y)}{\partial y \partial x}.$$

Bei den bisherigen Beispielen hatten wir es mit Funktionen $f(x, y)$, also Abbildungen der Art $f : \mathbb{R}^2 \mapsto \mathbb{R}$ zu tun. Das Konzept der partiellen Ableitung können wir aber ganz leicht auch auf Funktionen von n unabhängigen Variablen erweitern.

$$f : \mathbb{R}^n \mapsto \mathbb{R} : (x_1, x_2, \dots, x_i, \dots, x_n) \mapsto f(x_1, x_2, \dots, x_i, \dots, x_n)$$

Die partielle Ableitung einer solchen Funktion nach der Variable x_i erhalten wir – analog zu unserer Vorgehensweise bei $f(x, y)$, indem wir alle anderen unabhängigen Variablen bei der Ableitung festhalten.

$$\left(\frac{\partial f(x_1, x_2, \dots, x_i, \dots, x_n)}{\partial x_i} \right)_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n}$$

Hierbei haben wir die übliche Schreibweise eingeführt, in der die festgehaltenen Variablen explizit angegeben werden.

Beispiel: Wir betrachten die Funktion

$$f(x, y, z, t) = e^{i(2x+3y+z-4t)}$$

und bilden einige partielle Ableitungen:

$$\begin{aligned} \left(\frac{\partial f(x, y, z, t)}{\partial x} \right)_{y, z, t} &= 2ie^{i(2x+3y+z-4t)} \\ \left(\frac{\partial f(x, y, z, t)}{\partial y} \right)_{x, z, t} &= 3ie^{i(2x+3y+z-4t)} \\ \left(\frac{\partial f(x, y, z, t)}{\partial t} \right)_{x, y, z} &= -4ie^{i(2x+3y+z-4t)} \end{aligned}$$

In der Praxis werden wir zwar oft auch eine verkürzte Schreibweise benutzen

$$\begin{aligned} f_x(x, y, z, t) &= 2ie^{i(2x+3y+z-4t)} \\ f_y(x, y, z, t) &= 3ie^{i(2x+3y+z-4t)} \\ f_t(x, y, z, t) &= -4ie^{i(2x+3y+z-4t)}, \end{aligned}$$

wir sollten dabei aber nicht vergessen, dass wir bei partiellen Ableitungen immer berücksichtigen müssen, welche Variablen festgehalten werden. Das ist besonders wichtig, wenn wir Funktionen von einem Satz von unabhängigen Variablen auf einen anderen Satz von Variablen umschreiben, also eine Variablentransformation durchführen.

4.2.3 Variablentransformation

Sehr oft lassen sich physikalische Zusammenhänge vereinfacht darstellen, wenn man die Symmetrie des Problems berücksichtigt, und spezielle Koordinaten verwendet. Wir betrachten dazu zwei sehr häufig auftretende Beispiele, nämlich Polar- oder Zylinderkoordinaten und Kugelkoordinaten.

Polarkoordinaten. Wie bereits im Kapitel über Komplexe Zahlen erwähnt können wir einen Punkt (x, y) in der Ebene entweder durch Angabe der kartesischen Koordinaten x und y festlegen, oder den Abstand vom Ursprung r und den Polarwinkel φ angeben.

$$\begin{aligned} x &= r \cos \varphi & r &= \sqrt{x^2 + y^2} \\ y &= r \sin \varphi & \varphi &= \arctan \frac{y}{x} \end{aligned}$$

Betrachten wir jetzt eine Funktion

$$z = \frac{x^2}{\sqrt{x^2 + y^2}} = \frac{x^2}{r} = r \cos^2 \varphi,$$

so können wir die partiellen Ableitungen nach x , y , r , und φ berechnen. Dabei ist zu beachten, dass das Ergebnis davon abhängt, welche Variable wir festhalten.

$$\begin{aligned} \left(\frac{\partial z}{\partial x}\right)_y &= \frac{2x\sqrt{x^2 + y^2} - x^2 \frac{2x}{2\sqrt{x^2 + y^2}}}{x^2 + y^2} = \frac{2x}{r} - \frac{x^3}{r^3} \\ \left(\frac{\partial z}{\partial x}\right)_r &= \frac{2x}{r} \\ \left(\frac{\partial z}{\partial r}\right)_x &= -\frac{x^2}{r^2} = -\cos^2 \varphi \\ \left(\frac{\partial z}{\partial r}\right)_\varphi &= \cos^2 \varphi \end{aligned}$$

Wir sehen also, dass partielle Ableitungen davon abhängen, welche zwei Variablen wir gewählt haben, um z auszudrücken. Mathematisch gesehen handelt es sich ja bei den Funktionen $z(x, y)$ und $z(x, r)$ bzw. bei $z(x, r)$ und $z(r, \varphi)$ um jeweils andere Funktionszusammenhänge, wodurch es auch nicht

überraschen sollte, dass sich ihre partiellen Ableitungen unterscheiden. Streng genommen müssten wir also auch jeweils andere Symbole für diese Funktionen einführen, also $z_1(x, y)$, $z_2(x, r)$, und $z_3(r, \varphi)$. Wenn z aber zum Beispiel die ein und dieselbe physikalische Größe ausdrücken soll, dann wird das in der physikalischen Anwendung nicht gemacht, und man muss bei partiellen Ableitungen nicht nur angeben, nach welcher Variable abgeleitet wird, sondern auch welche Variable(n) festgehalten werden.

Kugelkoordinaten. In Kugelkoordinaten wird ein Punkt (x, y, z) im dreidimensionalen Raum durch seinen Abstand vom Ursprung r und durch zwei Winkel angegeben, den Polarwinkel θ und den Azimutwinkel φ . Wenn der Abstand vom Ursprung konstant ist (auf einer Sphäre = Kugeloberfläche), benötigt man nur die zwei Winkel, um einen Punkt eindeutig zu bezeichnen, und spricht dann von sphärischen Koordinaten.

$$\begin{aligned} x &= r \sin \theta \cdot \cos \varphi & r &= \sqrt{x^2 + y^2 + z^2} \\ y &= r \sin \theta \cdot \sin \varphi & \theta &= \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ z &= r \cos \theta & \varphi &= \arctan \frac{y}{x} \end{aligned}$$

Wir betrachten folgende Funktion (das elektrostatische Potential eines Dipols), zum einen gegeben als Funktion der kartesischen Koordinaten x , y , und z , und zum anderen in Kugelkoordinaten,

$$V(x, y, z) = \frac{z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \quad V(r, \theta, \varphi) = \frac{\cos \theta}{r^2}$$

und berechnen die partiellen Ableitungen nach x , y , und z bzw. r , θ , und φ .

$$\begin{aligned} \left(\frac{\partial V}{\partial x}\right)_{y,z} &= -\frac{3xz}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \\ \left(\frac{\partial V}{\partial y}\right)_{x,z} &= -\frac{3yz}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \\ \left(\frac{\partial V}{\partial z}\right)_{x,y} &= \frac{(x^2 + y^2 + z^2)^{\frac{3}{2}} - 3z^2(x^2 + y^2 + z^2)^{\frac{1}{2}}}{(x^2 + y^2 + z^2)^3} = \frac{x^2 + y^2 - 2z^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \\ \left(\frac{\partial V}{\partial r}\right)_{\theta,\varphi} &= -2\frac{\cos \theta}{r^3} \\ \left(\frac{\partial V}{\partial \theta}\right)_{r,\varphi} &= -\frac{\sin \theta}{r^2} \\ \left(\frac{\partial V}{\partial \varphi}\right)_{r,\theta} &= 0 \end{aligned}$$

4.2.4 Taylor-Reihe für mehrdimensionale Funktionen

Wir können auch Funktionen mehrerer Veränderlicher in eine Potenzreihe entwickeln. Bevor wir die allgemeine Taylor-Formel für n -dimensionale Funktionen kennenlernen, demonstrieren wir die Vorgehensweise anhand eines Beispiels, und betrachten die Funktion $f(x, y) = e^x \sin y$ über \mathbb{R}^2 . Wir suchen nach einer Potenzreihendarstellung mit dem Entwicklungspunkt $\{0, 0\}$ in folgender Form

$$f(x, y) = e^x \sin y \stackrel{!}{=} a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \dots$$

Analog zur Herleitung der Taylor-Formel für eine Funktion mit nur einer unabhängigen Variable, bilden wir nacheinander verschiedene partielle Ableitungen der obigen Gleichung und setzen anschließend auf beiden Seiten der Gleichung den Entwicklungspunkt $x = 0$ und $y = 0$ ein. Daraus erhalten wir

$$0 = a_{00}, \quad 0 = a_{01}, \quad 1 = a_{01}, \quad 0 = 2a_{20}, \quad 1 = a_{11}, \quad 0 = 2a_{02},$$

wodurch wir für die Taylor-Reihe bis zur quadratischen Ordnung folgenden Ausdruck bekommen¹

$$e^x \sin y = y + xy + \dots$$

Verallgemeinern wir die obige Vorgehensweise auf einer beliebige Funktion $f(x, y)$ und suchen nach der Potenzreihe um den Entwicklungspunkt $\{x_0, y_0\}$, so erhalten wir die Formel von Taylor für Funktionen von 2 Veränderlichen

$$f(x, y) = \sum_{k=0}^{\infty} \frac{1}{k!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^k f(x_0, y_0) \quad (4.3)$$

Die Bedeutung des *Differenzialoperators* in eckigen Klammern sei anhand eines Beispiels erläutert. Das Reihenglied für $k = 2$ lautet etwa

$$\left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^2 = (x - x_0)^2 \frac{\partial^2}{\partial x^2} + 2(x - x_0)(y - y_0) \frac{\partial^2}{\partial x \partial y} + (y - y_0)^2 \frac{\partial^2}{\partial y^2}$$

Das heißt wir lassen die Differenzialoperatoren $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial x \partial y}$, und $\frac{\partial^2}{\partial y^2}$ auf die Funktion $f(x, y)$ anwenden, oder anders ausgedrückt wir müssen die partiellen Ableitungen f_{xx} , f_{xy} , und f_{yy} berechnen. Erst **nach** der Ableitung setzen wir den Entwicklungspunkt ein. Wir können leicht überprüfen, dass die Form der Gleichung 4.3 identisch ist mit der zuvor abgeleiteten Taylor-Entwicklung für die Funktion $e^x \sin y$. Die Form von 4.3 kann auch problemlos auf Funktionen mit mehreren Veränderlichen verallgemeinert

¹Um dieses Ergebnis zu erhalten, hätten wir auch das Produkt der Taylor-Entwicklungen der Funktionen e^x und $\sin y$ bilden können.

werden

$$f(x, y, z, \dots) = \sum_{k=0}^{\infty} \frac{1}{k!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} + (z - z_0) \frac{\partial}{\partial z} + \dots \right]^k f(x_0, y_0, z_0, \dots) \quad (4.4)$$

Beispiel. Wir wollen die Taylor-Reihe der Funktion $f(x, y, z) = \sin(x-y) \ln(1+z)$ bis zur 2. Ordnung um den Punkt $(x_0, y_0, z_0) = (0, 0, 0)$ auf zwei Arten berechnen. In der ersten Variante verwenden wir die allgemeine Taylor-Form aus Gleichung 4.4. In einer zweiten Methode verwenden wir die bekannten Taylor-Reihen der Funktionen $\sin(x)$ und $\ln(1+z)$, ersetzen in der Sinusreihe x durch $x-y$ und multiplizieren mit der Reihe für den Logarithmus. Aber zunächst zur Variante 1, bei der wir alle partiellen Ableitungen der Funktion $f(x, y, z)$ bis zur 2. Ordnung benötigen

$$\begin{aligned} f &= \sin(x-y) \ln(1+z), & f(0,0,0) &= 0 \\ f_x &= \cos(x-y) \ln(1+z), & f_x(0,0,0) &= 0 \\ f_y &= -\cos(x-y) \ln(1+z), & f_y(0,0,0) &= 0 \\ f_z &= \frac{\sin(x-y)}{1+z}, & f_z(0,0,0) &= 0 \end{aligned}$$

$$\begin{aligned} f_{xx} &= -\sin(x-y) \ln(1+z), & f_{xx}(0,0,0) &= 0 \\ f_{xy} &= \sin(x-y) \ln(1+z), & f_{xy}(0,0,0) &= 0 \\ f_{xz} &= \frac{\cos(x-y)}{1+z}, & f_{xz}(0,0,0) &= 1 \\ f_{yy} &= -\sin(x-y) \ln(1+z), & f_{yy}(0,0,0) &= 0 \\ f_{yz} &= -\frac{\cos(x-y)}{1+z}, & f_{yz}(0,0,0) &= -1 \\ f_{zz} &= -\frac{\sin(x-y)}{(1+z)^2}, & f_{zz}(0,0,0) &= 0 \end{aligned}$$

Gleichung 4.4 lautet explizit

$$\begin{aligned} f(x, y, z) &= f + (f_x x + f_y y + f_z z) + \frac{1}{2} (f_{xx} x^2 + f_{yy} y^2 + f_{zz} z^2 + 2f_{xy} xy + 2f_{xz} xz + 2f_{yz} yz) + \dots \\ f(x, y, z) &= xz - yz + \dots \end{aligned}$$

Wir überprüfen noch dieses Ergebnis, indem wir die bekannten Taylor-Reihen für $\sin(x-y)$ und $\ln(1+z)$ verwenden:

$$\begin{aligned} \sin(x-y) &= (x-y) - \frac{(x-y)^3}{3!} + \dots \\ \ln(1+z) &= z - \frac{z^2}{2} + \dots \\ \sin(x-y) \ln(1+z) &= xz - yz + \dots \end{aligned} \quad (4.5)$$

Wir sehen also, dass wir mit dieser Variante viel schneller zum Ziel gekommen sind.

4.2.5 Totales Differenzial

Wir können die Definition des totalen Differenzials, wie wir sie für eine Funktion einer Veränderlichen kennengelernt haben, auch auf 2 und mehr Dimensionen erweitern. Im Fall einer Funktion $f(x, y)$ von zwei Veränderlichen, x und y , beschreibt das totale Differenzial

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

eine Tangentialebene an die Fläche $z = f(x, y)$ im Punkt $\{x, y, f(x, y)\}$. Das totale Differenzial beschreibt also die beste Näherung der Funktion $f(x, y)$ bis zur linearen Ordnung in dx und dy :

$$\Delta f(x, y) = df + \mathcal{O}((dx)^2, dx dy, (dy)^2)$$

Die Gleichung der Tangentialebene an den Punkt $\{x_0, y_0, f(x_0, y_0)\}$ können wir dann wie folgt angeben

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

Beispiel. Wir bestimmen das totale Differenzial der Funktion $f(x, y) = 1 - x^2 - y^2$ sowie die Tangentialebene an den Punkt $x_0 = \frac{1}{2}$ und $y_0 = \frac{1}{2}$. Dazu benötigen wir die partiellen Ableitungen

$$f_x(x, y) = -2x \quad f_y(x, y) = -2y$$

und erhalten für das totale Differenzial

$$df = -2x dx - 2y dy.$$

Für die Gleichung der Tangentialebene finden wir

$$z = \frac{1}{2} - \left(x - \frac{1}{2}\right) - \left(y - \frac{1}{2}\right)$$

Verallgemeinerung auf \mathbb{R}^n . Für eine Funktion f , die von n unabhängigen Variablen x_1, x_2, \dots, x_n abhängt, $f(x_1, x_2, \dots, x_n)$, lautet das **Totale Differenzial**

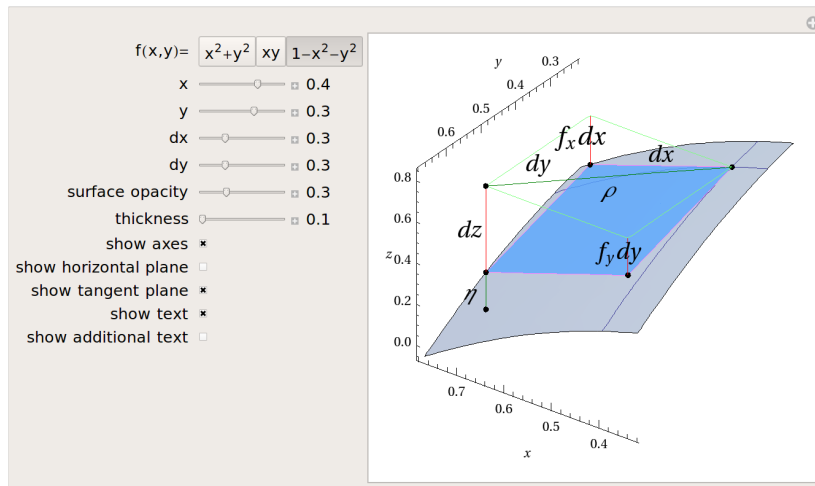
$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n \equiv \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i. \quad (4.6)$$

Die Gleichung der "Tangentialebene" an die Hyperfläche $z = f(x_1, x_2, \dots, x_n)$ in dem Punkt $\{a_1, a_2, \dots, a_n\}$ lautet dann

$$z = f(a_1, a_2, \dots, a_n) + \sum_{i=1}^n \frac{\partial f(a_1, a_2, \dots, a_n)}{\partial x_i} (x_i - a_i).$$

CDF 24. Totales Differenzial `TotalDifferential.cdf`

Die geometrische Interpretation des totalen Differenzials als Tangentialebene wird hier verdeutlicht.



Beispiel: Mit Hilfe des totalen Differenzials einer Funktion können wir die lineare Änderung der Funktion in der Umgebung eines vorgegebenen Punktes sehr schnell abschätzen. Nehmen wir als Beispiel die Formel für das Volumen eines Zylinders $V = r^2 \pi h$, wobei r den Radius und h die Höhe bezeichnet. Wie ändert sich das Volumen, wenn wir den Radius bzw. die Höhe leicht ändern? Diese Frage können wir mithilfe des totalen Differenzials beantworten:

$$dV = \frac{\partial V}{\partial r} r + \frac{\partial V}{\partial h} h = 2r\pi h dr + r^2 \pi dh$$

Nehmen wir an, der Ausgangsradius sei $r = 5$ cm, die Höhe $h = 10$ cm, was ein Volumen von $V = 785.398$ cm³ ergibt. Wie ändert sich das Volumen, wenn wir den Radius um $\Delta r = 0.1$ cm und die Höhe um $\Delta h = 0.1$ cm ändern? Natürlich können wir das neue exakte Volumen berechnen und erhalten $V + \Delta V = (r + \Delta r)^2 \pi (h + \Delta h) = 825.299$ cm³. Wir können die Änderung $\Delta V = 39.90$ cm³ auch abschätzen:

$$\Delta V = dV + \mathcal{O}(dr^2, drdh, dh^2) = 2r\pi h dr + r^2 \pi dh = 31.41 + 7.85 = 39.27 \text{ cm}^3$$

Wir sehen also, dass die lineare Näherung für sehr kleine Abweichungen sehr gut funktioniert und lernen auch, dass die Änderung des Radius einen viel größeren Einfluss auf die Volumenänderung hat als eine Änderung der Höhe.

Fehlerfortpflanzung: Das totale Differenzial einer Funktion ist auch Ausgangspunkt für die Betrachtung der Fehlerfortpflanzung bei Messungenauigkeiten. Als ein ganz einfaches Beispiel nehmen wir die Bestimmung der Erdbeschleunigung g durch einen Fallversuch, bei dem wir die Fallhöhe h und die Fallzeit t messen. Bekanntermaßen gilt ja der Zusammenhang $g = \frac{2s}{t^2}$. Das totale Differenzial lautet also

$$dg = \frac{\partial g}{\partial s} ds + \frac{\partial g}{\partial t} dt = \frac{2}{t^2} ds - \frac{4s}{t^3} dt$$

Ersetzen wir die Differenziale durch endliche Größen, die Messungenauigkeiten, $ds \rightarrow \Delta s$, $dt \rightarrow \Delta t$, und $dg \rightarrow \Delta g$, dann erhalten wir für den *maximalen absoluten Fehler* Δg bzw. den *maximalen relativen Fehler* $\frac{\Delta g}{g}$

$$\begin{aligned} \Delta g &= \left| \frac{2}{t^2} \right| \Delta s + \left| \frac{4s}{t^3} \right| \Delta t, \\ \frac{\Delta g}{g} &= \frac{\Delta s}{s} + 2 \frac{\Delta t}{t}. \end{aligned}$$

4.2.6 Richtungsableitung und Gradient

Unter Richtungsableitung einer von mehreren Variablen abhängigen Funktion verstehen wir die Änderungsrate dieser Funktion in einer durch einen Einheitsvektor \vec{v} vorgegebenen Richtung. Wir betrachten eine Funktion

$$f : \mathbb{R}^n \mapsto \mathbb{R} : (x_1, x_2, \dots, x_n) \mapsto z \equiv f(x_1, x_2, \dots, x_n)$$

Die **Richtungsableitung** an einem Punkt $\vec{x} = (x_1, x_2, \dots, x_n)$, die wir mit $D_{\vec{v}}f(\vec{x})$ bezeichnen, erhalten wir dann aus folgendem Grenzwert

$$D_{\vec{v}}f(\vec{x}) = \lim_{h \rightarrow 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h} \quad (4.7)$$

Existiert der Grenzwert, so nennen wir die Funktion f am Punkt \vec{x} in die Richtung \vec{v} mit $|\vec{v}| = 1$ beidseitig differenzierbar. Wir bemerken, dass partielle Ableitungen spezielle Richtungsableitungen darstellen, nämlich Ableitungen in Richtung einer der kartesischen Koordinaten x_1, x_2, \dots oder x_n . Beispielsweise ist die partielle Ableitung der nach der i -ten kartesischen Koordinate gleich

$$D_{(0, \dots, 1, \dots, 0)}f(\vec{x}) = \lim_{h \rightarrow 0} \frac{f(\vec{x} + h(0, \dots, 1, \dots, 0)) - f(\vec{x})}{h} = \frac{\partial f}{\partial x_i}(\vec{x})$$

Wie können wir aber Richtungsableitungen in eine beliebige Richtung \vec{v} berechnen? Mithilfe der Definition des Totalen Differenzials in Gleichung 4.6 können wir den Zähler in Gleichung 4.7 über das

totale Differenzial ausdrücken und finden

$$D_{\vec{v}}f(\vec{x}) = \lim_{h \rightarrow 0} \frac{\sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i h}{h} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i = (\vec{\nabla} f) \cdot \vec{v}. \quad (4.8)$$

Das heißt, wir können die Richtungsableitung als Skalarprodukt zweier Vektoren verstehen, nämlich zwischen dem Vektor $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$, den wir als *Gradienten* der Funktion $f(x_1, x_2, \dots, x_n)$ bezeichnen, und dem Einheitsvektor (v_1, v_2, \dots, v_n) in die betrachtete Richtung.

Gradient. Der Gradient einer partiell differenzierbaren Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}$ im Punkt $\vec{p} = (p_1, p_2, \dots, p_n)$ ist der Vektor der partiellen Ableitungen in diesem Punkt:

$$(\vec{\nabla} f)(\vec{p}) \equiv (\text{grad } f)(\vec{p}) = \begin{pmatrix} \partial_{x_1} f(\vec{p}) \\ \partial_{x_2} f(\vec{p}) \\ \vdots \\ \partial_{x_n} f(\vec{p}) \end{pmatrix}$$

Hier haben wir das Symbol "∇" eingeführt, das man als den "Nabla"-Operator bezeichnet, also

$$\vec{\nabla} = \begin{pmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \vdots \\ \partial_{x_n} \end{pmatrix} \equiv \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

Mit Hilfe des Nabla-Operators können wir Gleichung 4.6 für das totale Differenzial der Funktion f auch als Skalarprodukt des Gradientenvektors $\vec{\nabla} f$ mit dem Vektor der Differenziale $d\vec{x} = (x_1, x_2, \dots, x_n)$ schreiben

$$df = (\vec{\nabla} f) \cdot d\vec{x} \quad (4.9)$$

Totale Differenzierbarkeit von Funktionen auf \mathbb{R}^n . Wir nennen eine Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}$ im Punkt $\vec{x} = (x_1, x_2, \dots, x_n)$ (total) differenzierbar, wenn das totale Differenzial die Änderungen der Funktion f im Punkt \vec{x} in linearer Näherung beschreibt. Oder anders ausgedrückt, der Fehler ist von höherer als erster Ordnung in $\Delta\vec{x} = (\Delta x_1, \Delta x_2, \dots, \Delta x_n)$

$$f(\vec{x} + \Delta\vec{x}) = f(\vec{x}) + (\vec{\nabla} f) \cdot \Delta\vec{x} + \mathcal{O}(|\Delta\vec{x}|^2) \quad (4.10)$$

Diese Definition der Differenzierbarkeit ist analog zur Definition, wie wir sie bereits für eindimensionale

Funktionen kennengelernt haben 4.1, und die hier für Vergleichszwecke noch einmal angegeben wird:

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \mathcal{O}((\Delta x)^2).$$

Ein wichtiger Hinweis: Die Existenz der partiellen Ableitungen allein sagt noch nichts über die totale Differenzierbarkeit einer Funktion aus! Das heißt, es gibt Funktionen, von denen zwar die partiellen Ableitungen existieren, die aber im Sinne der Gleichung 4.10 nicht differenzierbar sind. Im folgenden wollen wir uns aber auf solche Fälle beschränken, für die die totale Differenzierbarkeit gewährleistet ist, und geben folgendes hinreichendes Kriterium für die Differenzierbarkeit an:

Sind die *partiellen Ableitungen* einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in einer Umgebung eines Punktes \vec{p} *stetig*, so ist f im Punkt \vec{p} *differenzierbar*. Es gelten folgende Implikationen

$$\begin{aligned} \text{stetige partielle Differenzierbarkeit} &\Rightarrow \text{totale Differenzierbarkeit} \\ \text{totale Differenzierbarkeit} &\Rightarrow \text{Differenzierbarkeit in jede Richtung} \\ \text{Differenzierbarkeit in jede Richtung} &\Rightarrow \text{partielle Differenzierbarkeit,} \end{aligned}$$

jedoch keine der Umkehrungen!

Beispiel: Wir geben ein Beispiel für eine Funktion an, die zwar partiell differenzierbar ist, aber nicht stetig differenzierbar, und von der nicht alle Richtungsableitungen existieren.

$$f(x, y) = \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

Von Interesse ist hier nur die Differenzierbarkeit und Stetigkeit am Ursprung $(0, 0)$. Überall sonst ist die Funktionen stetig differenzierbar. Die partielle Ableitung von $f(x, y)$ entlang der x -Achse finden wir, indem wir $y = 0$ setzen, also $f(x, 0) = 0$. Damit ist auch die beidseitige partielle Ableitung in x -Richtung identisch Null, also $f_x(x, 0) = 0$. Analoges gilt für die Ableitung in y -Richtung, also $f_y(0, y) = 0$. Die Funktion $f(x, y)$ ist jedoch an $(0, 0)$ nicht stetig, und damit nicht total differenzierbar, und auch die Richtungsableitungen in andere Richtungen als in die Koordinatenrichtungen existieren nicht. Setzen wir etwa $t \equiv x = y$, so finden wir $f(t, t) = 1$, andererseits finden wir für $f(t, -t) = -1$. Damit ist diese Funktion ein Beispiel dafür, dass die Umkehrung der oben angeführten Implikationen nicht gilt.

Für Funktionen, die jedoch stetig partiell differenzierbar sind – und mit solchen wollen wir uns im weiteren auseinandersetzen – können wir folgende wichtige **Eigenschaften des Gradienten der Funktion** feststellen.

1. Für stetig partiell differenzierbare Funktionen f kann die Richtungsableitung $D_{\vec{v}}f(\vec{x})$ in eine

Richtung \vec{v} durch den Gradienten ausgedrückt werden

$$D_{\vec{v}}f(\vec{x}) = (\vec{\nabla}f)\vec{v}$$

2. Der Betrag des Gradientenvektors $|\vec{\nabla}f|$ gibt die maximale Änderung der Funktion f an.

3. Gilt $|\vec{\nabla}f| \neq 0$, dann gibt der Einheitsvektor

$$\vec{e} = \frac{\vec{\nabla}f}{|\vec{\nabla}f|}$$

die Richtung des steilsten Anstiegs der Funktion f an. Anders ausgedrückt: Die Richtungsableitung wird maximal, wenn man in Richtung des Gradienten ableitet.

Beispiel. Wir betrachten eine Funktion $f : \mathbb{R}^2 \mapsto \mathbb{R}$ gegeben durch $f(x, y) = \sin(x - y) + \cos x$, und suchen nach der Richtungsableitung am Punkt $\vec{p} = (\frac{\pi}{2}, \frac{\pi}{3})$ in die Richtung $\vec{v} = \frac{1}{\sqrt{2}}(1, 1)$. Wir berechnen zunächst den Gradienten

$$\vec{\nabla}f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos(x - y) - \sin x \\ -\cos(x - y) \end{pmatrix} \Big|_{x=\frac{\pi}{2}, y=\frac{\pi}{3}} = \begin{pmatrix} \frac{\sqrt{3}}{2} - 1 \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$$

Multiplizieren wir den Gradienten am Punkt $(\frac{\pi}{2}, \frac{\pi}{3})$ skalar mit dem Richtungsvektor $\frac{1}{\sqrt{2}}(1, 1)$ so erhalten wir für die Richtungsableitung

$$D_{\vec{v}}f(\vec{x}) = \begin{pmatrix} \frac{\sqrt{3}}{2} - 1 \\ -\frac{\sqrt{3}}{2} \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\frac{1}{\sqrt{2}} = -0.707107$$

Wir berechnen noch den Betrag des Gradienten am Punkt $(\frac{\pi}{2}, \frac{\pi}{3})$

$$|\vec{\nabla}f|_{x=\frac{\pi}{2}, y=\frac{\pi}{3}} = \left| \begin{pmatrix} \frac{\sqrt{3}}{2} - 1 \\ -\frac{\sqrt{3}}{2} \end{pmatrix} \right| = \sqrt{\frac{5}{2} - \sqrt{3}} = 0.876327$$

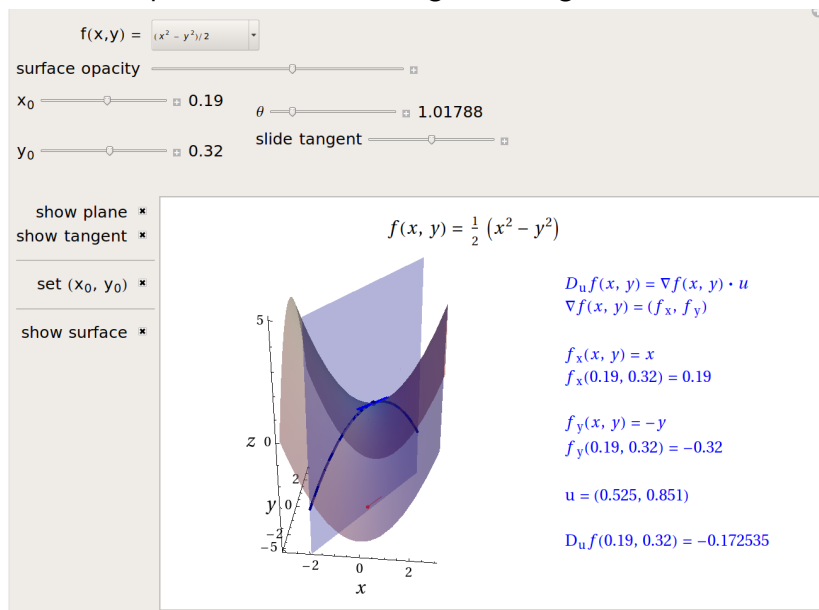
sowie die Richtung des steilsten Anstiegs an diesem Punkt

$$\vec{e} = \frac{\vec{\nabla}f}{|\vec{\nabla}f|} = \frac{1}{\sqrt{\frac{5}{2} - \sqrt{3}}} \begin{pmatrix} \frac{\sqrt{3}}{2} - 1 \\ -\frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} -0.152882 \\ -0.988244 \end{pmatrix}$$

Diese Ergebnisse sowie Richtungsableitungen und Gradienten von weiteren Funktionen können sie auch in den folgenden interaktiven Mathematica-Demonstrationen visualisieren:

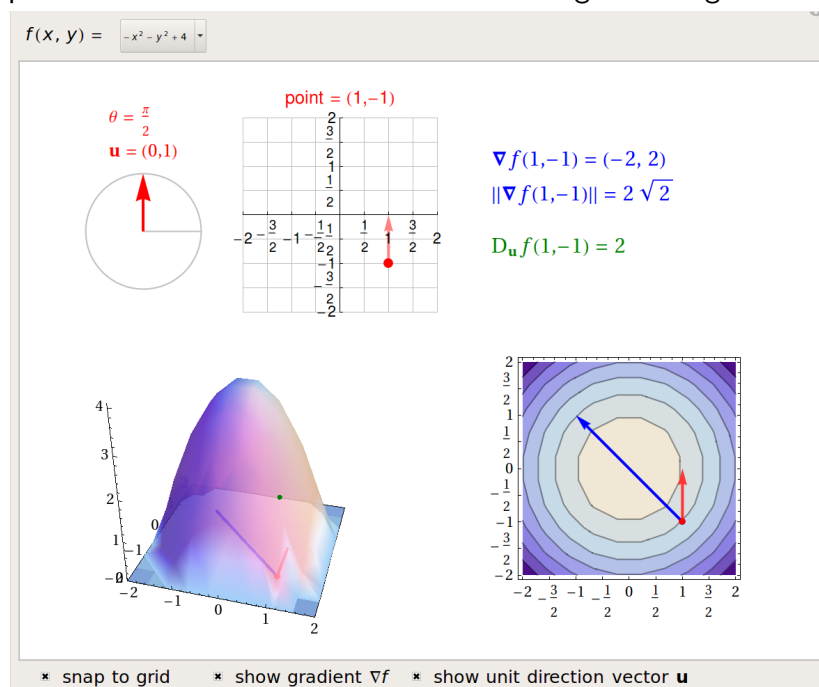
CDF 25. Richtungsableitung <https://moodle.uni-graz.at/course/view.php?id=95124>

Hier wird die geometrische Interpretation der Richtungsableitung verdeutlicht.



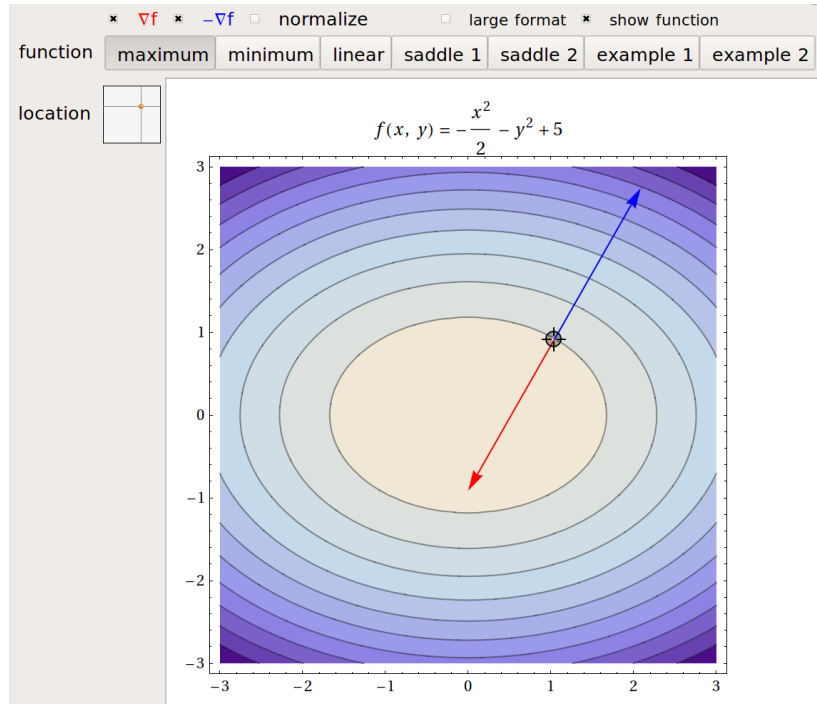
CDF 26. Gradient <https://moodle.uni-graz.at/course/view.php?id=95124>

Eine weitere CDF-Applikation zum Thema Gradient und Richtungsableitung:



CDF 27. Gradient – Teil 2 <https://moodle.uni-graz.at/course/view.php?id=95124>

... und noch eine CDF-Applikation zum Thema Gradient.



4.2.7 Jacobi-Matrix

Bisher hatten wir es mit Funktionen der Art $f : \mathbb{R}^n \mapsto \mathbb{R}$ zu tun. Solche Abbildungen bezeichnet man in der Physik als *Skalarfelder*. Speziell für $n = 3$, also $f : \mathbb{R}^3 \mapsto \mathbb{R}$ können wir uns solche Funktionen als Abbildungen vorstellen, die jedem Punkt (x, y, z) im dreidimensionalen Raum *eine* Zahl, also einen *Skalar* zuordnen. Beispiele für skalare Felder wären etwa die räumliche Temperatur- oder Druckverteilung, das Potenzial der Schwerkraft über der Erde, die elektrostatische Potenzial eines geladenen Körpers, oder die Aufenthaltswahrscheinlichkeitsdichte eines quantenmechanischen Teilchens.² In der Physik gibt es aber auch eine Reihe von physikalischen Größen, die durch Vektoren beschrieben werden. Bei der Strömung einer Flüssigkeit etwa, wird jedem Punkt (x, y, z) ein Geschwindigkeitsvektor (v_x, v_y, v_z) zugewiesen. Es handelt sich also um ein *Vektorfeld*, das man mathematisch als Abbildung $f : \mathbb{R}^3 \mapsto \mathbb{R}^3$ beschreiben kann. Weitere Beispiele von Vektorfeldern wären das Gravitations-Kraftfeld, oder das elektrische Feld einer Ladungsverteilung.

Um nun das totale Differenzial, also die lineare Näherung, einer allgemeinen Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ ausdrücken zu können, definieren wir die *Jacobi-Matrix* (benannt nach Carl Gustav Jacob Jacobi;

²Die stationäre Wellenfunktion ist eine komplexwertige Funktion der Art $\Psi : \mathbb{R}^3 \mapsto \mathbb{C}$.

auch Funktionalmatrix oder Ableitungsmatrix genannt) in der folgenden Form:

$$J_f(\vec{x}_0) = \left(\begin{array}{cccc} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{array} \right) \Bigg|_{\vec{x}_0} \equiv \frac{\partial(f_1 \cdots f_m)}{\partial(x_1 \cdots x_n)} \Bigg|_{\vec{x}_0}$$

Die Jacobi-Matrix ist also nichts anderes als eine $m \times n$ Matrix sämtlicher erster partieller Ableitungen. Wir identifizieren die i -te Zeile der Jacobi-Matrix mit dem Gradientenvektor der i -ten Vektorkomponente f_i . Ganz ähnlich wie wir den Gradienten einer Funktion dazu benutzt haben, das totale Differenzial, also ihre lineare Näherung, anzugeben (siehe Gleichung 4.9), so können wir mithilfe der Jacobi-Matrix das totale Differenzial einer Vektorfunktion $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ angeben

$$d\vec{f} = J_f(x) \cdot d\vec{x}, \quad (4.11)$$

wobei wir die Multiplikation als Matrix-Vektor Produkt verstehen.

Beispiel: Wir betrachten die Funktion $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$

$$\vec{f}(x, y, z) \equiv \begin{pmatrix} f_1(x, y, z) \\ f_2(x, y, z) \end{pmatrix} = \begin{pmatrix} x^2 + y^2 + z \sin x \\ z^2 + z \sin y \end{pmatrix},$$

und wollen die Jacobi-Matrix aufstellen, sowie das totale Differenzial der Funktion mithilfe der Gleichung 4.11 angeben. Bilden wir alle partiellen Ableitungen, so finden wir für die Jacobi-Matrix

$$J_f(x, y, z) = \begin{pmatrix} 2x + z \cos z & 2y & \sin x \\ 0 & z \cos y & 2z + \sin y \end{pmatrix}$$

und können das totale Differenzial nach Gleichung 4.11 anschreiben als

$$\begin{pmatrix} df_1 \\ df_2 \end{pmatrix} = \begin{pmatrix} 2x + z \cos z & 2y & \sin x \\ 0 & z \cos y & 2z + \sin y \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix},$$

und finden schließlich

$$\begin{aligned} df_1 &= (2x + z \cos z)dx + 2ydy + \sin x dz \\ df_2 &= z \cos y dy + (2z + \sin y)dz \end{aligned}$$

Jacobi-Matrix für Polarkoordinaten. Wir betrachten die Transformation von kartesischen Koordinaten auf Polarkoordinaten also eine Abbildung $\mathbb{R}^2 \mapsto \mathbb{R}^2$

$$\begin{pmatrix} x(r, \varphi) \\ y(r, \varphi) \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix},$$

und wollen wiederum die Jacobi-Matrix aufstellen,

$$J(r, \varphi) = \begin{pmatrix} \partial_r x & \partial_\varphi x \\ \partial_r y & \partial_\varphi y \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

Jacobi-Matrix für Kugelkoordinaten. Zum Abschluss dieses Kapitels berechnen wir noch die Jacobi-Matrix für Transformation von kartesischen Koordinaten auf Kugelkoordinaten also für folgende Abbildung von $\mathbb{R}^3 \mapsto \mathbb{R}^3$

$$\begin{pmatrix} x(r, \theta, \varphi) \\ y(r, \theta, \varphi) \\ z(r, \theta, \varphi) \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix}.$$

Die Jacobi-Matrix für diese Transformation hat dann folgende Gestalt:

$$J(r, \theta, \varphi) = \begin{pmatrix} \partial_r x & \partial_\theta x & \partial_\varphi x \\ \partial_r y & \partial_\theta y & \partial_\varphi y \\ \partial_r z & \partial_\theta z & \partial_\varphi z \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix} \quad (4.12)$$

4.3 Methoden der Differenziation

Für die praktische Berechnung von (partiellen) Ableitungen von Funktionen gehen wir üblicherweise nicht von der Definition des Differenzialquotienten aus. Hingegen verwenden wir unsere Kenntnis über die Ableitung einiger elementarer Funktionen und einfach anwendbare Ableitungsregeln, die es uns erlauben, auch die Ableitungen von komplizierteren zusammengesetzten Funktionen zu berechnen. Einige dieser Ableitungsregeln sind uns bereits wohl bekannt und wir haben sie im bisherigen Verlauf der Vorlesung vielfach eingesetzt. Dazu gehören die *Produktregel*, die *Quotientenregel*, sowie die *Kettenregel*. Wir werden diese Regeln an dieser Stelle nochmals wiederholen bzw. auch im Lichte von partiellen Ableitungen betrachten. Zudem lernen wir auch eine Methode kennen, mit deren Hilfe wir die Ableitungen von Funktionen berechnen können, für die wir den Funktionszusammenhang nicht explizit angeben können, sondern die nur *implizit* geben sind.

4.3.1 Produktregel, Quotientenregel, und Kettenregel

Produktregel. Die Produktregel für Funktionen $f : \mathbb{R} \mapsto \mathbb{R}$, die sich als Produkt zweier Funktionen $u(x)$ und $v(x)$ darstellen lässt, lautet

$$f(x) = u(x)v(x), \quad f'(x) = u'(x)v(x) + u(x)v'(x).$$

Quotientenregel. Die Quotientenregel für Funktionen $f : \mathbb{R} \mapsto \mathbb{R}$, die sich als Quotient zweier Funktionen $u(x)$ und $v(x)$ darstellen lässt, lautet

$$f(x) = \frac{u(x)}{v(x)}, \quad f'(x) = \frac{u'(x)v(x) - u(x)v'(x)}{v^2(x)}.$$

Kettenregel. Die Kettenregel trifft Aussagen über die Ableitung einer Funktion, die sich als *Verkettung von zwei differenzierbaren Funktionen* darstellen lässt. Die Kettenregel ergibt, dass eine solche Funktion selbst wieder differenzierbar ist und man ihre Ableitung erhält, indem man die beiden miteinander verketteten Funktionen separat ableitet und – ausgewertet an den richtigen Stellen – miteinander multipliziert. Wir betrachten also die verkettete Funktion $f(x) = g(h(x))$

$$\begin{array}{ccccc} \mathbb{R} & \mapsto & \mathbb{R} & \mapsto & \mathbb{R} \\ x & \xrightarrow{h} & h(x) \equiv y & \xrightarrow{g} & f(y) \end{array}$$

Die Kettenregel besagt, dass wir die Ableitung $f'(x)$ erhalten, indem wir folgenden Ausdruck berechnen:

$$f'(x) = g'(y)h'(x) \equiv g'(h(x))h'(x).$$

Beispiel: Wir wollen die Funktion $f(x) = (\sin x)^4$ ableiten. Wir können f in die zwei Teilfunktionen $h(x) = \sin x$ und $g(y) = y^4$ zerlegen. Damit erhalten wir

$$f'(x) = [(\sin x)^4]' = 4y^3 \cos x = 4(\sin x)^3 \cos x.$$

Verallgemeinerung der Kettenregel – Teil 1. Soweit, so bekannt. Wir wollen jetzt die Kettenregel auch auf Funktionen mehrerer Veränderlicher anwenden. Als ersten Schritt in diese Richtung betrachten wir verkettete Funktionen der Art $\mathbb{R} \mapsto \mathbb{R}^n \mapsto \mathbb{R}$

$$\left. \begin{array}{l} x \xrightarrow{h_1} h_1(x) \\ x \xrightarrow{h_2} h_2(x) \\ \vdots \\ x \xrightarrow{h_n} h_n(x) \end{array} \right\} \xrightarrow{f} f(h_1, h_2, \dots, h_n)$$

Um ein Beispiel zu geben. Die Funktion $y = \sqrt{x-2}e^{-x^2} \sin x$ können wir mit $h_1(x) = \sqrt{x-2}$, $h_2(x) = e^{-x^2}$, und $h_3(x) = \sin x$ in die Form $y = f(h_1, h_2, h_3) = h_1 h_2 h_3$ bringen. Um die Ableitung $\frac{dy}{dx}$ zu bilden, berechnen wir zunächst das totale Differenzial der Funktion $f(h_1, h_2, \dots, h_n)$

$$df = \frac{\partial f}{\partial h_1} dh_1 + \frac{\partial f}{\partial h_2} dh_2 + \dots + \frac{\partial f}{\partial h_n} dh_n,$$

und ersetzen die Differenziale dh_1, dh_2, \dots, dh_n durch die jeweiligen totalen Differenziale der eindimensionalen Funktionen $h_1(x), h_2(x), \dots, h_n(x)$, also $dh_1 = \frac{dh_1}{dx} dx$ usw. Damit erhalten wir

$$\begin{aligned} df &= \frac{\partial f}{\partial h_1} \frac{dh_1}{dx} dx + \frac{\partial f}{\partial h_2} \frac{dh_2}{dx} dx + \dots + \frac{\partial f}{\partial h_n} \frac{dh_n}{dx} dx \\ &= \left(\frac{\partial f}{\partial h_1} \frac{dh_1}{dx} + \frac{\partial f}{\partial h_2} \frac{dh_2}{dx} + \dots + \frac{\partial f}{\partial h_n} \frac{dh_n}{dx} \right) dx \end{aligned}$$

und identifizieren den Ausdruck in Klammern somit als die Ableitung der Funktion nach x

$$\frac{df}{dx} = \frac{\partial f}{\partial h_1} \frac{dh_1}{dx} + \frac{\partial f}{\partial h_2} \frac{dh_2}{dx} + \dots + \frac{\partial f}{\partial h_n} \frac{dh_n}{dx} \quad (4.13)$$

Wenden wir diese Gleichung auf das obige Beispiel an, so finden wir

$$\begin{aligned} \frac{df}{dx} &= h_2 h_3 \frac{dh_1}{dx} + h_1 h_3 \frac{dh_2}{dx} + h_1 h_2 \frac{dh_3}{dx} \\ &= e^{-x^2} \sin x \frac{1}{2\sqrt{x-2}} + \sqrt{x-2} \sin x \left(-2xe^{-x^2} \right) + \sqrt{x-2} e^{-x^2} \cos x \end{aligned}$$

Verallgemeinerung der Kettenregel – Teil 2. In einem zweiten Schritt wollen wir jetzt die Kettenregel auch auf Funktionen mehrerer Veränderlicher anwenden, die wir wieder analog zu oben in der folgenden Form zerlegen: $\mathbb{R}^m \mapsto \mathbb{R}^n \mapsto \mathbb{R}$

$$\left. \begin{array}{l} (x_1, x_2, \dots, x_m) \xrightarrow{h_1} h_1(x_1, x_2, \dots, x_m) \\ (x_1, x_2, \dots, x_m) \xrightarrow{h_2} h_2(x_1, x_2, \dots, x_m) \\ \vdots \\ (x_1, x_2, \dots, x_m) \xrightarrow{h_n} h_n(x_1, x_2, \dots, x_m) \end{array} \right\} \xrightarrow{f} f(h_1, h_2, \dots, h_n)$$

Ein Beispiel für eine solche Funktion $f : \mathbb{R}^m \mapsto \mathbb{R}$ wäre etwa

$$f(x_1, x_2, x_3) = \frac{e^{-x_1^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \sin(x_2 - x_3)$$

mit den Funktionen

$$h_1(x_1, x_2, x_3) = e^{-x_1^2}, \quad h_2(x_1, x_2, x_3) = \sqrt{x_1^2 + x_2^2 + x_3^2}, \quad h_3(x_1, x_2, x_3) = \sin(x_2 - x_3),$$

$$f(h_1, h_2, h_3) = \frac{h_1}{h_2} h_3.$$

Wir bilden wieder das totale Differenzial der Funktion f in Bezug auf die Größen h_1, h_2, \dots, h_n

$$df = \frac{\partial f}{\partial h_1} dh_1 + \frac{\partial f}{\partial h_2} dh_2 + \dots + \frac{\partial f}{\partial h_n} dh_n,$$

und ersetzen die Differenziale dh_1, dh_2, \dots, dh_n durch die jeweiligen totalen Differenziale der eindimensionalen Funktionen $h_1(x), h_2(x), \dots, h_n(x)$, also etwa

$$dh_1 = \frac{\partial h_1}{\partial x_1} dx_1 + \frac{\partial h_1}{\partial x_2} dx_2 + \dots + \frac{\partial h_1}{\partial x_m} dx_m \equiv \left(\vec{\nabla}_{\vec{x}} h_1 \right) d\vec{x}$$

Damit erhalten wir für das totale Differenzial

$$\begin{aligned} df &= \frac{\partial f}{\partial h_1} \left(\vec{\nabla}_{\vec{x}} h_1 \right) d\vec{x} + \frac{\partial f}{\partial h_2} \left(\vec{\nabla}_{\vec{x}} h_2 \right) d\vec{x} + \dots + \frac{\partial f}{\partial h_n} \left(\vec{\nabla}_{\vec{x}} h_n \right) d\vec{x} \\ &= \left[\frac{\partial f}{\partial h_1} \frac{\partial h_1}{\partial x_1} + \frac{\partial f}{\partial h_2} \frac{\partial h_2}{\partial x_1} + \dots + \frac{\partial f}{\partial h_n} \frac{\partial h_n}{\partial x_1} \right] dx_1 \\ &+ \dots \\ &+ \left[\frac{\partial f}{\partial h_1} \frac{\partial h_1}{\partial x_m} + \frac{\partial f}{\partial h_2} \frac{\partial h_2}{\partial x_m} + \dots + \frac{\partial f}{\partial h_n} \frac{\partial h_n}{\partial x_m} \right] dx_m. \end{aligned}$$

Die Terme in eckigen Klammern identifizieren wir als die gesuchten partiellen Ableitungen der Funktion f nach den Variablen x_1, x_2, \dots, x_m . Die partielle Ableitung nach x_j ist also

$$\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial h_1} \frac{\partial h_1}{\partial x_j} + \frac{\partial f}{\partial h_2} \frac{\partial h_2}{\partial x_j} + \dots + \frac{\partial f}{\partial h_n} \frac{\partial h_n}{\partial x_j} \quad (4.14)$$

Wir wenden die soeben gefunden Kettenregel zur Bildung der partiellen Ableitung auf die oben erwähnte Beispielfunktion an, und wollen alle drei partiellen Ableitungen $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$, und $\frac{\partial f}{\partial x_3}$ berechnen. Wir finden also zunächst mit $f(h_1, h_2, h_3) = \frac{h_1}{h_2} h_3$

$$\frac{\partial f}{\partial h_1} = \frac{h_3}{h_2}, \quad \frac{\partial f}{\partial h_2} = -\frac{h_1}{h_2^2} h_3, \quad \frac{\partial f}{\partial h_3} = \frac{h_1}{h_2}.$$

Jetzt bilden wir noch die partiellen Ableitungen der Funktionen h_1 , h_2 , und h_3 nach den Variablen x_1 , x_2 , und x_3 , also die insgesamt 9 partielle Ableitungen:

$$\begin{aligned} \frac{\partial h_1}{\partial x_1} &= -2x_1 e^{-x_1^2} & \frac{\partial h_2}{\partial x_1} &= \frac{x_1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} & \frac{\partial h_3}{\partial x_1} &= 0 \\ \frac{\partial h_1}{\partial x_2} &= 0 & \frac{\partial h_2}{\partial x_2} &= \frac{x_2}{\sqrt{x_1^2 + x_2^2 + x_3^2}} & \frac{\partial h_3}{\partial x_2} &= \cos(x_2 - x_3) \\ \frac{\partial h_1}{\partial x_3} &= 0 & \frac{\partial h_2}{\partial x_3} &= \frac{x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}} & \frac{\partial h_3}{\partial x_3} &= -\cos(x_2 - x_3) \end{aligned}$$

Damit erhalten wir schließlich für die partiellen Ableitungen der Funktion f

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= -2x_1 \frac{h_1 h_3}{h_2} - \frac{h_1}{h_2^2} h_3 \frac{x_1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\ \frac{\partial f}{\partial x_2} &= -\frac{h_1}{h_2^2} h_3 \frac{x_2}{\sqrt{x_1^2 + x_2^2 + x_3^2}} + \frac{h_1}{h_2} \cos(x_2 - x_3) \\ \frac{\partial f}{\partial x_3} &= -\frac{h_1}{h_2^2} h_3 \frac{x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}} - \frac{h_1}{h_2} \cos(x_2 - x_3) \end{aligned}$$

Bei diesem Beispiel hätten wir natürlich genauso gut direkt die partielle Ableitung der Funktion $f(x_1, x_2, x_3)$ bilden können, was auf das gleiche Resultat geführt hätte, wie wir leicht als Übung zeigen können.

Beispiel: Als abschließendes Beispiel zur Anwendung der Kettenregel, und um einen Zusammenhang zur Jacobi-Matrix herzustellen, betrachten wir noch folgende Funktion

$$f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}},$$

von der wir die partiellen Ableitungen nach den Kugelkoordinaten (r, θ, φ) berechnen wollen. Dazu schreiben wir das totale Differenzial der Funktion f zunächst in der Form

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$

Für die totalen Differenziale der kartesischen Koordinaten als Funktion der Kugelkoordinaten finden wir (siehe Kapitel 4.2.7)

$$\begin{aligned} dx &= \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta + \frac{\partial x}{\partial \varphi} d\varphi = \sin \theta \cos \varphi dr + r \cos \theta \cos \varphi d\theta - r \sin \theta \sin \varphi d\varphi \\ dy &= \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta + \frac{\partial y}{\partial \varphi} d\varphi = \sin \theta \sin \varphi dr + r \cos \theta \sin \varphi d\theta + r \sin \theta \cos \varphi d\varphi \\ dz &= \frac{\partial z}{\partial r} dr + \frac{\partial z}{\partial \theta} d\theta + \frac{\partial z}{\partial \varphi} d\varphi = \cos \theta dr - r \sin \theta d\theta \end{aligned}$$

Das heißt mithilfe der Jacobi-Matrix $J(r, \theta, \varphi)$ aus Gleichung 4.12 hätten wir obige Gleichungen auch vereinfacht schreiben können, als:

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = J(r, \theta, \varphi) \cdot \begin{pmatrix} dr \\ d\theta \\ d\varphi \end{pmatrix}$$

Setzen wir die Differenziale dx , dy , und dz in das totale Differenzial df ein, und ordnen die Terme nach den neuen Differenzialen dr , $d\theta$, und $d\varphi$, so erhalten wir

$$\begin{aligned} df &= \left(\frac{\partial f}{\partial x} \sin \theta \cos \varphi + \frac{\partial f}{\partial y} \sin \theta \sin \varphi + \frac{\partial f}{\partial z} \cos \theta \right) dr \\ &+ \left(\frac{\partial f}{\partial x} r \cos \theta \cos \varphi + \frac{\partial f}{\partial y} r \cos \theta \sin \varphi - \frac{\partial f}{\partial z} r \sin \theta \right) d\theta \\ &+ \left(-\frac{\partial f}{\partial x} r \sin \theta \sin \varphi + \frac{\partial f}{\partial y} r \sin \theta \cos \varphi \right) d\varphi. \end{aligned}$$

Hierbei identifizieren wir die Terme in runden Klammern als die partiellen Ableitungen der Funktion f nach den Kugelkoordinaten r , θ , und φ . Mit

$$\frac{\partial f}{\partial x} = -\frac{x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \quad \frac{\partial f}{\partial y} = -\frac{y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \quad \frac{\partial f}{\partial z} = -\frac{z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}$$

finden wir für die partiellen Ableitungen nach r , θ , und φ :

$$\begin{aligned}\frac{\partial f}{\partial r} &= -\frac{r}{r^3} (\sin^2 \theta \cos^2 \varphi + \sin^2 \theta \sin^2 \varphi + \cos^2 \theta) = -\frac{1}{r^2} \\ \frac{\partial f}{\partial \theta} &= -\frac{r^2}{r^3} (\sin \theta \cos \theta \cos^2 \varphi + \sin \theta \cos \theta \sin^2 \varphi - \sin \theta \cos \theta) = 0 \\ \frac{\partial f}{\partial \varphi} &= -\frac{r^2}{r^3} (-\sin^2 \theta \sin \varphi \cos \varphi + \sin^2 \theta \sin \varphi \cos \varphi) = 0\end{aligned}$$

4.3.2 Implizite Differenziation

Bisher hatten wir es immer mit *expliziten* Funktionszusammenhängen zu tun. Das heißt wir konnten für die abhängige Variable, zum Beispiel y , immer eine explizite Abhängigkeit von der unabhängigen Variable x angeben. In manchen Fällen ist es allerdings nicht möglich den Funktionszusammenhang zwischen y und x in die Form $y = y(x)$ zu bringen, sondern wir können nur einen *impliziten* Zusammenhang in der Form $f(x, y) = 0$ angeben. Die Frage, die wir in diesem Kapitel beantworten werden, wie man Ableitungen und totale Differenziale von solchen implizit gegebenen Funktionen bilden kann.

Beispiel 1: Wir beginnen mit einem Beispiel für eine implizite Funktion

$$f(x, y) = x^2 + y - \sin(xy) = 0.$$

Diese Gleichung können wir nicht nach y auflösen, wir können also keine explizite Form $y = y(x)$ angeben. Nichtsdestotrotz beschreibt $f(x, y) = 0$ eine Kurve in der xy -Ebene,³ für die es möglich sein sollte Tangenten zu bestimmen, das heißt Ableitungen zu bilden. Die Frage, die wir uns also stellen, ist: Wie ändert sich y , wenn wir x um ein dx ändern, wir suchen also $\frac{dy}{dx}$. Diesen Ausdruck können wir aber gewinnen, indem wir das totale Differenzial des impliziten Funktionszusammenhangs $f(x, y) = 0$ bilden:

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = [2x - y \cos(xy)] dx + [1 - x \cos(xy)] dy = 0.$$

Damit erhalten wir

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = -\frac{2x - y \cos(xy)}{1 - x \cos(xy)}.$$

und damit die gesuchte Änderung der Variable y in Bezug auf x . Im allgemeinen – wie auch in diesem Beispiel – erhalten wir das Ergebnis wieder in Form eines impliziten Funktionszusammenhangs.

³Für ein vorgegebenes x erhalten wir eine Gleichung für y , und damit können wir jedem x ein y zuordnen, wie wir es für Funktionen gewohnt sind.

Beispiel 2: Mithilfe des totalen Differenzials können wir auch Ableitungen von Funktionszusammenhängen bilden, die in Form mehrerer Gleichungen gegeben sind. Betrachten wir dazu wieder ein Beispiel:

$$\begin{aligned}\sqrt{s} + t^3 + u &= 0 \\ 2s + \frac{1}{\sqrt{t}} - e^u &= 0\end{aligned}$$

Wir haben es hier mit einem Gleichungssystem bestehend aus *zwei* Gleichungen mit den *drei* Variablen s , t , und u zu tun. Im Prinzip könnten wir also eine Variable, beispielsweise u eliminieren, und erhalten dann einen Zusammenhang zwischen den anderen beiden, also s und t , für die wir wiederum die implizite Ableitung $\frac{ds}{dt}$ berechnen könnten. Da es aber nicht immer einfach – oder überhaupt möglich – ist, diese Eliminierung durchzuführen, gehen wir einen anderen Weg, indem wir das totale Differenzial der obigen Gleichungen bestimmen:

$$\begin{aligned}\frac{1}{2\sqrt{s}}ds + 3t^2dt + du &= 0 \\ 2ds - \frac{1}{2\sqrt{t^3}}dt - e^u du &= 0\end{aligned}$$

Wir erhalten bei dieser Vorgehensweise immer ein *lineares* Gleichungssystem in den Differenzialen ds , dt , und du , das wir auf in jedem Fall lösen können. Wollen wir die Ableitung $\frac{ds}{dt}$ bilden, so eliminieren wir aus obiger Gleichung die Terme mit du , indem wir etwa die erste Gleichung mit e^u multiplizieren, und dann die Addition der beiden Gleichungen bilden

$$\left(\frac{e^u}{2\sqrt{s}} + 2\right) ds + \left(3t^2e^u - \frac{1}{2\sqrt{t^3}}\right) dt = 0$$

Damit erhalten wir für das gewünschte Ergebnis:

$$\frac{ds}{dt} = -\frac{3t^2e^u - \frac{1}{2\sqrt{t^3}}}{\frac{e^u}{2\sqrt{s}} + 2} = -\frac{3t^2\left(2s + \frac{1}{\sqrt{t}}\right) - \frac{1}{2\sqrt{t^3}}}{\frac{2s + \frac{1}{\sqrt{t}}}{2\sqrt{s}} + 2}$$

Beispiel 3: In einem letzten Beispiel zur impliziten Ableitung betrachten wir die Funktion

$$z = x^2 + xy,$$

wobei x und y wiederum implizite Funktionen der Variablen s und t sein sollen, in der Art

$$f_1(x, y, s, t) = x + y^2 - s - t = 0, \quad f_2(x, y, s, t) = x^2 - y - s^2 = 0$$

Was wir suchen, sind die partiellen Ableitungen $\frac{\partial z}{\partial s}$ und $\frac{\partial z}{\partial t}$. Wir gehen dabei so vor, dass wir zunächst das totale Differenzial von z anschreiben

$$dz = (2x + y)dx + xdy$$

und anschließend die Differenziale dx und dy mithilfe des impliziten Zusammenhangs zwischen x und y bzw. s und t durch die Differenziale ds und dt ausdrücken. Dazu bilden wir das totale Differenzial von f_1 und f_2

$$df_1 = dx + 2ydy - ds - dt = 0, \quad df_2 = 2xdx - dy - 2sds = 0$$

und berechnen aus dem linearen Gleichungssystem die unbekanntenen Differenziale dx und dy :

$$\begin{aligned} dx &= \frac{1}{1 + 4xy} [(1 + 4sy)ds + dt] \\ dy &= \frac{1}{1 + 4xy} [(2x - 2s)ds + 2xdt] \end{aligned}$$

Einsetzen der Differenziale dx und dy in das totale Differenzial dz ergibt

$$dz = \left[\frac{2x + y}{1 + 4xy} (1 + 4ys) + \frac{x}{1 + 4xy} (2x - 2s) \right] ds + \left[\frac{2x + y}{1 + 4xy} + \frac{2x^2}{1 + 4xy} \right] dt = \frac{\partial z}{\partial s} ds + \frac{\partial z}{\partial t} dt,$$

wodurch wir die das totale Differenzial von z nun durch ds und dt ausgedrückt haben. Die gesuchten partiellen Ableitungen können wir schließlich mit den Ausdrücken in eckigen Klammern identifizieren.

4.4 Extremwertwertberechnungen

Ob und vor allem an welcher Stelle Funktionen Extremwerte, das heißt *Maxima* oder *Minima* aufweisen, ist eine sehr häufige Problemstellung, die bei sehr vielen Fragestellungen in der Physik auftritt. Nehmen wir etwa als zu untersuchende Funktion die potentielle Energie eines Teilchens als Funktion seiner Position, dann wird das Teilchen – sofern wir es sich selbst überlassen – die Position einnehmen, an der seine potentielle Energie ein *Minimum* annimmt. An diesem Punkt angelangt, kann das Teilchen in Ruhe verharren, man sagt es hat eine *stationäre* Position erreicht, die es nur durch Zufuhr von Energie wieder verlassen kann. Ein solcher stationärer Zustand muss nicht notwendigerweise, der energetisch tiefstmögliche auf einer globalen Skala sein, sondern wir im allgemeinen nur ein *lokales Minimum* darstellen.

In diesem Kapitel werden wir uns damit beschäftigen, wie wir solche lokalen Extremstellen von Funktionen auffinden können, und wodurch wir sie charakterisieren können. Zunächst beschäftigen wir uns mit Funktionen einer Veränderlichen, also $f(x)$, bevor wir uns auch mehrdimensionaler Funktionen

annehmen.

4.4.1 Funktionen einer Veränderlichen

Zunächst geben wir eine Definition an, was wir unter einem *globalen* bzw. einem *lokalen* Maximum (Minimum) einer Funktion verstehen.

Definition. Sei eine Funktion f definiert auf $f : \mathbb{D} \subseteq \mathbb{R} \mapsto \mathbb{R}$, dann nennen wir x_0 ein *globales* Maximum (bzw. Minimum), wenn $\forall x \in \mathbb{D}$ gilt, dass $f(x_0) \geq f(x)$ (bzw. $f(x_0) \leq f(x)$).

Wir nennen x_0 ein *lokales* Maximum (bzw. Minimum), wenn ein $\varepsilon > 0$ existiert, sodass für $\forall x \in K(x_0, \varepsilon)$, also für alle x innerhalb der offenen ε -Kugel um x_0 , gilt, dass $f(x_0) \geq f(x)$ (bzw. $f(x_0) \leq f(x)$).

Die Suche nach dem globalen Maximum (Minimum) stellt sich oft als sehr schwierig heraus, vor allem für mehrdimensionale Probleme, für die keine einfachen Algorithmen existieren. Für die Existenz von *lokalen* Extrema können wir hingegen folgendes (bereits aus der Schule bekannte) hinreichende Kriterium angeben.

Kriterium von Fermat. Liegt an der Stelle x_0 ein lokales Extremum vor, so gilt $f'(x_0) = 0$.

Es ist zu beachten, dass das Verschwinden der 1. Ableitung nur eine notwendige Bedingung für die Existenz eines Extremums ist, aber keine hinreichende! Das heißt, der Umkehrschluss gilt nicht, und aus $f'(x_0) = 0$ folgt *nicht*, dass x_0 ein Extremum ist. Um die Frage zu klären, ob eine Funktion mit der Eigenschaft $f'(x_0)$ an der Stelle x_0 tatsächlich ein Extremum aufweist, und wenn ja welcher Art es ist, müssen wir höhere Ableitungen der Funktion f betrachten. Dazu zwei Beispiele.

Beispiel. Wir betrachten die zwei Funktionen $f_1(x) = x^6$ und $f_2(x) = x^7$. Wie Sie sich leicht überzeugen können, besitzt die Funktion $f_1(x) = x^6$ ein lokales Minimum an der Stelle $x_0 = 0$, während die Funktion $f_2(x) = x^7$ weder ein lokales Maximum noch ein Minimum (also kein Extremum) besitzt, sondern an der Stelle $x_0 = 0$ einen Wendepunkt aufweist. Für beide Funktionen verschwinden jedoch die 1. Ableitungen, also $f_1'(x) = 6x^5$ und $f_2'(x) = 7x^6$ und somit $f_1'(0) = 0$ und $f_2'(0) = 0$. Bei diesen Beispielen bringt uns auch die Berechnung der 2. Ableitungen nicht weiter, $f_1''(x) = 30x^4$ und $f_2''(x) = 42x^5$, da sie an der Stelle $x_0 = 0$ beide Null sind: $f_1''(0) = 0$ und $f_2''(0) = 0$. Anhand dieses Beispiels können wir allerdings folgendes Kriterium für die Existenz eines Extremums ableiten.

Kriterium für die Existenz eines Extremums. Sei $f(x)$ eine in dem Intervall (a, b) n -mal stetig differenzierbare Funktion deren 1. Ableitung an der Stelle $x_0 \in (a, b)$ verschwindet, also $f'(x_0) = 0$, und es gilt

$$f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0 \quad \text{und} \quad f^{(n)}(x_0) \neq 0,$$

dann hat f kein Extremum an x_0 , wenn n eine ungerade Zahl ist bzw. es existiert Extremum an der Stelle x_0 , wenn n eine gerade Zahl darstellt. Für letzteren Fall ist x_0 ein Minimum, wenn $f^{(n)}(x_0) > 0$, und x_0 ein Maximum, wenn $f^{(n)}(x_0) < 0$ gilt.

Wir definieren zwei weitere Begriffe, die in Zusammenhang mit Extrema häufig diskutiert werden.

Konvexheit, Konkavheit. Sei eine Funktion f definiert auf $f : \mathbb{D} \subseteq \mathbb{R} \mapsto \mathbb{R}$, dann nennen wir die Funktion *konvex* bzw. *konkav* auf dem Intervall $\mathbb{I} \subseteq \mathbb{D}$, wenn $\forall x_1 < x_2 \in \mathbb{I}$ gilt:

$$\text{konvex:} \quad f(x) \leq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \quad \forall x \in (x_1, x_2)$$

$$\text{konkav:} \quad f(x) \geq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \quad \forall x \in (x_1, x_2)$$

Anschaulich bedeutet diese Definition, dass eine *konvexe* Kurve *unter* der Sekante durch die Punkte x_1 und x_2 liegt, während eine *konkave* Kurve *oberhalb* dieser Sekante liegt. Für Funktionen, die in einem Intervall (a, b) zweimal stetig differenzierbar sind, gilt insbesondere folgendes Kriterium

$$f''(x) \geq 0 \quad \text{auf } (a, b) \iff f \quad \text{ist konvex auf } (a, b)$$

$$f''(x) \leq 0 \quad \text{auf } (a, b) \iff f \quad \text{ist konkav auf } (a, b)$$

Beispiel. Wir untersuchen die zwei Funktionen $f_1(x) = x + 2 \cos x$ und $f_2(x) = x + \cos x$ definiert auf $\mathbb{R} \mapsto \mathbb{R}$ auf lokale Extrema.

Zwei Hinweise: Das Verschwinden der 1. Ableitung weist nur auf *lokale* Extremstellen hin. Etwaige *globale* Minima oder Maxima, die an den Rändern des Definitionsbereichs liegen können, werden durch den oben beschriebenen Algorithmus nicht erfasst. Davon können wir uns ganz einfach überzeugen, wenn wir eine Funktion $f : [-1, 1] \mapsto \mathbb{R}$ definiert durch $f(x) = 1 + x^2$ betrachten. Diese Funktion hat offensichtlich ein lokales Minimum an der Stelle $x = 0$, weist aber zusätzlich Maxima an den Rändern des Intervalls bei $x = -1$ und $x = +1$ auf.

Ebenso können wir die Forderung $f'(x) = 0$ zum Auffinden von Minima nicht auf Funktionen anwenden, die nicht stetig differenzierbar sind. Beispielsweise weist die Betragsfunktion $|x| : \mathbb{R} \mapsto \mathbb{R}$ an der Stelle $x = 0$ ein Minimum auf, wohingegen ihre Ableitung auf ganz \mathbb{R} nicht verschwindet. Dieses Minimum "entgeht" unserem Algorithmus, weil die 1. Ableitung nicht überall (an $x = 0$) definiert ist.

4.4.2 Funktionen mehrerer Veränderlicher

Für eine Funktion $f(x, y)$ von zwei Variablen, $f : \mathbb{R}^2 \mapsto \mathbb{R}$, erhalten wir einen Extrempunkt, wenn die Tangentialebene horizontal verläuft. Da wir die Tangentialebene mithilfe des totalen Differenzials

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

ausdrücken können, bedeutet das, dass sowohl die partielle Ableitung nach x als auch die partielle Ableitung nach y verschwinden muss. Wir können diesen Sachverhalt leicht auf mehrere Dimensionen erweitern und formulieren folgendes Kriterium für die Existenz eines Extremums.

Kriterium für das Vorliegen eines Extremums. Hat eine Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}$ an der Stelle $\vec{a} = (a_1, a_2, \dots, a_n)$ ein lokales Extremum, dann gilt dass der Gradient der Funktion an der Stelle \vec{a} identisch ist mit dem Nullvektor.

$$\vec{\nabla} f(\vec{x}) \Big|_{\vec{x}=\vec{a}} = \vec{0}$$

Wiederum gilt es zu beachten, dass das Verschwinden des Gradienten nur eine notwendige Bedingung für die Existenz eines Extremums ist, aber keine hinreichende. Das heißt, der Umkehrschluss gilt nicht, und aus $\vec{\nabla} f(\vec{x}) = \vec{0}$ folgt *nicht*, dass \vec{x} ein Extremum aufweist. Um die Frage zu klären, ob eine Funktion mit der Eigenschaft $\vec{\nabla} f(\vec{x}) = \vec{0}$ tatsächlich ein Extremum aufweist, und wenn ja welcher Art es ist, müssen wir höhere, partielle Ableitungen der Funktion f betrachten.

Beispiele in 2D. Wir betrachten die drei Funktionen $f_1(x, y) = x^2 + y^2$, $f_2(x, y) = x^2 - y^2$, und $f_3(x, y) = -x^2 - y^2$ und berechnen jeweils den Gradienten.

$$\vec{\nabla} f_1 = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \quad \vec{\nabla} f_2 = \begin{pmatrix} 2x \\ -2y \end{pmatrix}, \quad \vec{\nabla} f_3 = \begin{pmatrix} -2x \\ -2y \end{pmatrix}.$$

An der Stelle $(x, y) = (0, 0)$ ist der Gradientenvektor aller drei Funktionen der Nullvektor. Um zu entscheiden, ob ein Maximum, ein Minimum, oder ein Sattelpunkt vorliegt berechnen wir alle 2.

partiellen Ableitungen f_{xx} , f_{yy} , und f_{xy} , und stellen damit folgende Determinante auf:

$$D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - f_{xy}f_{yx} = f_{xx}f_{yy} - f_{xy}^2$$

Für die drei Beispielfunktionen erhalten wir somit.

$$D_1 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = +4, \quad D_2 = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4, \quad D_3 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = +4,$$

Das Vorzeichen der Determinante D sowie die Vorzeichen der partiellen Ableitungen f_{xx} und f_{yy} geben nun Auskunft über die Art der Extremstelle beziehungsweise über das Vorliegen eines Sattelpunktes.

Maximum für $D > 0, f_{xx} < 0, f_{yy} < 0$

Minimum für $D > 0, f_{xx} > 0, f_{yy} > 0$

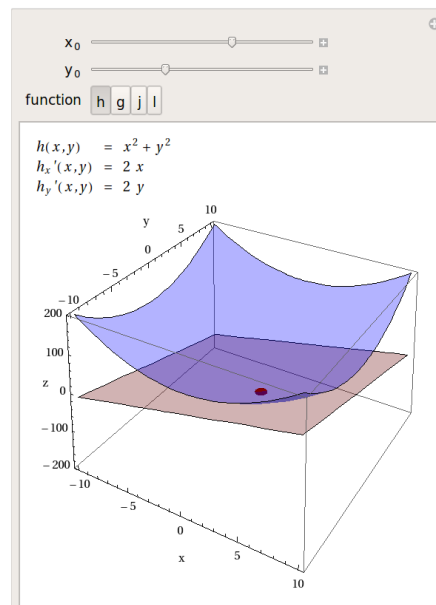
Sattelpunkt für $D < 0$

keine Aussage möglich für $D = 0$

Im Fall $D = 0$ muss die Art der Extremstelle mit Hilfe von höheren Ableitungen genauer untersucht werden.

CDF 28. Extrema in 2D <https://moodle.uni-graz.at/course/view.php?id=95124>

Tangentialebene an Funktionen mit lokalem Minimum (Maximum) oder Sattelpunkt.



Erweiterung auf n -Dimensionen. Um die Art von Extremstellen für n -dimensionale Funktionen f zu untersuchen, erweitern wir das oben vorgestellte Konzept der Determinanten und definieren die sogenannte **Hesse-Matrix**. Das ist eine $n \times n$ Matrix, die alle zweiten partiellen Ableitungen der Funktion f in folgender Form beinhaltet

$$H_f(\vec{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} \equiv \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

Zur Klassifizierung der Extrema wird der Begriff der *Definitheit* einer Matrix herangezogen (vgl. Vorlesung Lineare Algebra). Und zwar gilt folgender Sachverhalt.

Extrema und Hesse-Matrix. Verschwindet der Gradient der Funktion f an einer Stelle \vec{x}_0 , dann

- hat f ein lokales *Minimum* an der Stelle \vec{x}_0 , wenn $H_f(\vec{x}_0)$ *positiv definit* ist.
- hat f ein lokales *Maximum* an der Stelle \vec{x}_0 , wenn $H_f(\vec{x}_0)$ *negativ definit* ist.
- hat f einen Sattelpunkt an der Stelle \vec{x}_0 , wenn $H_f(\vec{x}_0)$ *indefinit* ist.

Um herauszufinden, ob eine Matrix positiv, oder negativ definit, oder indefinit ist, existieren mehrere Verfahren, die in der Vorlesung über Lineare Algebra besprochen werden. Erwähnt sei an dieser Stelle nur ein Kriterium, das Bezug nimmt auf die *Eigenwerte* der Matrix. Da es sich bei der Hesse-Matrix um eine symmetrische Matrix handelt – die gemischten Ableitungen sind für stetig differenzierbare Funktionen nach dem Satz von Schwarz gleich, sind alle ihre Eigenwerte reelle Zahlen. Eine Matrix nennt man nun *positiv definit*, wenn alle ihre Eigenwerte positiv sind, die Matrix ist *negativ definit*, wenn alle ihre Eigenwerte negativ sind, und sie ist *indefinit*, wenn sie sowohl positive als auch negative Eigenwerte aufweist.

Beispiel. Wir berechnen die Extremstellen der Funktion $f(x, y) = x^3 + y^3 - 3xy$ und untersuchen die Art der Extrempunkte, indem wir die Hesse-Matrix und deren Eigenwerte berechnen. Zunächst berechnen wir den Gradienten und die Hesse-Matrix der Funktion f

$$\vec{\nabla} f(x, y) = \begin{pmatrix} 3x^2 - 3y \\ 3y^2 - 3x \end{pmatrix}, \quad H_f(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

Setzen wir den Gradienten gleich $\vec{0}$, so finden wir die zwei reellen Lösungen

$$(x_1, y_1) = (0, 0) \quad \text{und} \quad (x_2, y_2) = (1, 1).$$

Nun müssen wir überprüfen, welcher Art die Extrema sind bzw. ob es sich um einen Sattelpunkt handelt. Dazu setzen wir die gefundenen Lösungen von $\vec{\nabla}f = \vec{0}$ in die allgemeine Form der Hesse-Matrix ein

$$H_f(0, 0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \quad H_f(1, 1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}.$$

Wir berechnen nun die Eigenwerte λ dieser zwei Matrizen. Zunächst für $H_f(0, 0)$:

$$\begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \iff \begin{pmatrix} -\lambda & -3 \\ -3 & -\lambda \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dieses lineare *homogene* Gleichungssystem besitzt eine nicht-triviale Lösung für den Fall, dass die Determinante der Matrix verschwindet, also

$$\begin{vmatrix} -\lambda & -3 \\ -3 & -\lambda \end{vmatrix} = 0 \iff \lambda^2 - 9 = 0 \Rightarrow \lambda = \pm 3$$

Das bedeutet also, dass die Hesse-Matrix an dem Punkt $(0, 0)$ sowohl positive als auch negative Eigenwerte besitzt, wodurch die Matrix *indefinit* ist. Dementsprechend handelt es sich bei dem Punkt $(0, 0)$ also um kein lokales Maximum oder Minimum, sondern um einen Sattelpunkt. Betrachten wir nun die Hesse-Matrix an dem zweiten Punkt $(1, 1)$ und berechnen wiederum ihre Eigenwerte aus dem charakteristischen Polynom der Determinante

$$\begin{vmatrix} 6 - \lambda & -3 \\ -3 & 6 - \lambda \end{vmatrix} = 0 \iff (6 - \lambda)^2 - 9 = 0 \Rightarrow \lambda_{1,2} = 6 \pm 3,$$

so finden wir, dass beide Eigenwerte, $\lambda_1 = 3$ und $\lambda_2 = 9$, positiv sind. Damit ist die Matrix positiv definit, und es handelt sich bei dem Punkt $(1, 1)$ um ein *lokales Minimum*. Wir berechnen abschließend noch die dazugehörigen *Eigenvektoren*. Setzen wir $\lambda_1 = 3$ bzw. $\lambda_2 = 9$ in die Eigenwertgleichung ein, so finden wir für die normierten Eigenvektoren

$$\vec{v}^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{v}^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Die Bedeutung dieser Eigenvektoren ist, dass sie die *Richtungen* in der xy -Ebene angeben, in denen die Funktion am geringsten (\vec{v}^1 mit $\lambda_1 = 3$) bzw. am stärksten (\vec{v}^2 mit $\lambda_2 = 9$) *gekrümmt* ist. Die

Eigenwerte der Hesse-Matrix geben also die Hauptkrümmungen der Funktion an.

Zum Abschluss dieses Kapitels sei noch erwähnt, dass wir mithilfe des Gradienten und der Hesse-Matrix einer Funktion auch die Taylor-Reihen Entwicklung einer Funktion bis inklusive zur 2. Ordnung wir folgt anschreiben können.

$$f(\vec{x}) = f(\vec{x}_0) + \left(\vec{\nabla} f(\vec{x}_0) \right) (\vec{x} - \vec{x}_0) + \frac{1}{2} (\vec{x} - \vec{x}_0) \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \mathcal{O}(|\vec{x} - \vec{x}_0|^3) \quad (4.15)$$

Handelt es bei dem Punkt \vec{x}_0 um ein lokales Extremum, verschwindet also der Gradient an dieser Stelle, so können wir also die Funktion f in der Umgebung dieses Punktes \vec{x}_0 bis zur quadratischen Ordnung annähern als:

$$f(\vec{x}) = f(\vec{x}_0) + \frac{1}{2} (\vec{x} - \vec{x}_0) \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \dots$$

Für die Funktion $f(x, y) = x^3 + y^3 - 3xy$ aus unserem obigen Beispiel bedeutet das für die Umgebung um den Punkt $(0, 0)$

$$f(x, y) = f(0, 0) + \frac{1}{2} (x, \quad y) \cdot \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = -3xy + \dots$$

Die Taylor-Reihe um das lokale Minimum an der Stelle $(1, 1)$ hat dann die Form

$$\begin{aligned} f(x, y) &= f(1, 1) + \frac{1}{2} (x - 1, \quad y - 1) \cdot \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix} \cdot \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix} \\ &= -1 + 3(x - 1)^2 - 3(x - 1)(y - 1) + 3(y - 1)^2 + \dots \end{aligned}$$

4.5 Extremwertwertberechnungen mit Nebenbedingungen

Sehr häufig tritt die Problemstellung auf, dass das Minimum (oder Maximum) einer Funktion gesucht wird, es aber zusätzliche *Nebenbedingungen* zu erfüllen sind. Beispielsweise, Sie suchen den kürzesten Abstand eines Punktes $(x, y, z) \in \mathbb{R}^3$ vom Ursprung. Das heißt wir wollen die Funktion $d(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ minimieren, nehmen aber als Nebenbedingung an, dass sich der Punkt (x, y, z) auf einer Ebene im Raum befindet. Wir erhalten als Ergebnis also den kürzesten Abstand der Ebene vom Ursprung. Als ein anderes Beispiel, das ebenfalls auf eine Extremwertaufgabe mit Nebenbedingungen führt, betrachten wir den Fall, dass wir aus einer Metallkugel den volumensgrößten Quader bestimmen wollen, der in der gegebenen Kugel Platz findet. Um solche und ähnliche Extremwertaufgaben zu lösen, lernen wir zwei Verfahren kennen, nämlich das *Eliminationsverfahren* und das Verfahren der *Lagrangeschen Multiplikatoren*.

4.5.1 Elimination

Wir beginnen mit einem konkreten Beispiel und behandeln die oben erwähnte Fragestellung: Wie bestimmt man den kürzesten Abstand einer Ebene zum Ursprung? Nehmen wir die Ebene

$$x - 2y - 2z = 3,$$

die also unsere *Nebenbedingung* darstellt. Die Funktion, die es zu minimieren gibt, ist die Abstandsfunktion

$$d(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

Es ist klar, dass das lokale (in diesem Fall auch das globale Minimum) dieser Funktion *ohne* Nebenbedingung der Ursprung des Koordinatensystems $(0, 0, 0)$ ist. Um die Nebenbedingung zu berücksichtigen, können wir eine der Variablen x , y , oder z mithilfe der Nebenbedingung eliminieren. Entscheiden wir uns für $x = 3 + 2y + 2z$ und setzen dies in die Abstandsfunktion ein, so erhalten wir

$$d(y, z) = \sqrt{(3 + 2y + 2z)^2 + y^2 + z^2}$$

Aufgrund der Elimination hängt die Abstandsfunktion nur mehr von *zwei* unabhängigen Variablen y und z ab. Aus dem vorangegangenen Kapitel wissen wir, wie wir die Extremstellen einer solchen Funktion finden können. Wir müssen fordern, dass die partiellen Ableitungen nach y und z verschwinden, also

$$\begin{aligned} \frac{\partial d}{\partial y} &= \frac{4(3 + 2y + 2z) + 2y}{2\sqrt{(3 + 2y + 2z)^2 + y^2 + z^2}} = 0 \\ \frac{\partial d}{\partial z} &= \frac{4(3 + 2y + 2z) + 2z}{2\sqrt{(3 + 2y + 2z)^2 + y^2 + z^2}} = 0 \end{aligned}$$

Wir erhalten ein lineares Gleichungssystem in den Unbekannten y und z , für das wir die Lösung $y = -\frac{2}{3}$ und $z = -\frac{2}{3}$ finden. Eingesetzt in die Ebenengleichung (die Nebenbedingung) erhalten wir noch $x = \frac{1}{3}$. Damit erhalten wir für den minimalen Abstand der Ebene vom Ursprung

$$d = \sqrt{\left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(-\frac{2}{3}\right)^2} = 1$$

Streng genommen müssten wir an dieser Stelle auch überprüfen, ob es sich tatsächlich um ein Minimum handelt, indem wir die Hesse-Matrix an dem Punkt $(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3})$ aufstellen, also die zweiten Ableitungen der Funktion $d(y, z)$ berechnen (Übung). Des weiteren können wir uns auch leicht davon überzeugen, dass wir zu demselben Ergebnis gelangt wären, hätten wir nicht x , sondern etwa y oder z aus der Abstandsfunktion eliminiert.

4.5.2 Lagrangesche Multiplikatoren

Das Eliminationsverfahren ist immer dann möglich, wenn die Nebenbedingung es zulässt, dass eine Variable *explizit* durch die anderen ausgedrückt werden kann, und dann in die Funktion, die es zu minimieren gilt, eingesetzt werden kann. Wenn die Nebenbedingung allerdings in Form einer *impliziten* Funktion gegeben ist, dann ist das Eliminationsverfahren nicht anwendbar. In einem solchen Fall kann man das Verfahren der *Lagrangeschen Multiplikatoren* anwenden. Wir werden das Verfahren zunächst ableiten. Danach wenden wir das Verfahren auf das obige Beispiel mit minimalen Anstand der Ebene vom Ursprung an. Schließlich erläutern wir das Verfahren auch anhand eines etwas komplizierten Beispiels.

Ableitung des Verfahrens der Lagrangeschen Multiplikatoren. Wir betrachten eine Funktion $f(x, y)$, deren Extremwerte wir berechnen wollen. Die Nebenbedingung bringen wir in die Form $\phi(x, y) = 0$. Für die Extremwerte der Funktion f muss ihr totales Differenzial df verschwinden

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0.$$

Die Differenziale dx und dy sind aber nicht unabhängig voneinander, sondern über die Nebenbedingung $\phi(x, y) = 0$ miteinander verknüpft. Das sehen wir unmittelbar, wenn wir das totale Differenzial der Nebenbedingung bilden

$$d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy = 0.$$

Wie wir auch schon bei der Bildung der *impliziten* Ableitung gesehen haben, können wir somit das Differenzial dy durch dx ausdrücken

$$dy = -\frac{\frac{\partial \phi}{\partial x}}{\frac{\partial \phi}{\partial y}} dx.$$

Eingesetzt in das totale Differenzial df finden wir

$$\left(\frac{\partial f}{\partial x} - \frac{\frac{\partial \phi}{\partial x}}{\frac{\partial \phi}{\partial y}} \frac{\partial f}{\partial y} \right) dx = 0$$

Da dx beliebig ist, muss in obiger Gleichung der Ausdruck in runden Klammern Null sein. Zusammen mit der Nebenbedingung $\phi(x, y) = 0$ haben wir dann ein System von 2 Gleichungen, das wir für die 2 Unbekannten x und y lösen können, und somit haben wir das Extremalproblem mit Nebenbedingung

gelöst.

$$\begin{aligned}\frac{\partial f}{\partial x} - \frac{\frac{\partial \phi}{\partial x} \partial f}{\frac{\partial \phi}{\partial y} \partial y} &= 0 \\ \phi(x, y) &= 0\end{aligned}$$

Um diese Methode leichter auf mehrere Veränderliche erweitern zu können, und auch um ein Verfahren zu finden, das wir uns leichter einprägen können, formen wir obiges Gleichungssystem in folgender Weise um. Wir führen eine *neue* Variable λ ein, die wir den *Lagrangeschen Multiplikator* nennen.

$$\lambda = -\frac{\frac{\partial f}{\partial y}}{\frac{\partial \phi}{\partial y}}$$

Mithilfe von λ können wir das oben abgeleitete System von 2 Gleichungen mit den Unbekannten x und y nun auf ein System von *drei* Gleichungen mit den *drei* Unbekannten x , y , und λ umschreiben.

$$\begin{aligned}\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} &= 0 \\ \phi(x, y) &= 0.\end{aligned}$$

Was haben wir jetzt erreicht, außer ein Problem mit ursprünglich 2 Unbekannten auf ein scheinbar komplizierteres Problem mit 3 Unbekannten umzuschreiben? Zum einen interessiert uns der Wert des Lagrangeschen Multiplikator λ meist nicht, das heißt wir brauchen ihn nicht auszurechnen, zum anderen können wir das obige System von drei Gleichungen sehr einfach aufstellen. Dazu betrachten wir die neue Funktion

$$F(x, y, \lambda) = f(x, y) + \lambda \phi(x, y).$$

Fordern wir von dieser Funktion, die ja die ursprüngliche zu minimierende Funktion $f(x, y)$ und die Nebenbedingung $\phi(x, y)$ multipliziert mit dem Lagrangeschen Multiplikator λ enthält, dass ihre partiellen Ableitungen verschwinden, dann erhalten genau dasselbe System von Gleichungen!

$$\begin{aligned}\frac{\partial F}{\partial x} = 0 &\Leftrightarrow \frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0 \\ \frac{\partial F}{\partial y} = 0 &\Leftrightarrow \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0 \\ \frac{\partial F}{\partial \lambda} = 0 &\Leftrightarrow \phi(x, y) = 0.\end{aligned}$$

Der Vorteil des Verfahrens der Lagrangeschen Multiplikatoren ist, dass es

- leicht zu merken ist,
- leicht auf mehrere Variablen anwendbar ist,
- leicht auf mehrere Nebenbedingungen erweiterbar ist,
- auch für Nebenbedingungen in implizierter Form geeignet ist.

Wir bemerken also, dass wir mithilfe der Methode des Lagrangeschen Multiplikators die Extremwertaufgabe mit Nebenbedingung in ein einfaches Extremwertproblem transformiert haben, das allerdings eine Variable mehr – nämlich den Lagrangeschen Multiplikator – umfasst als das ursprüngliche Problem.

Beispiel 1. Als erste Anwendung berechnen wir nun den minimalen Abstand der Ebene $x - 2y - 2z = 3$ vom Ursprung mithilfe der Methode der Lagrangeschen Multiplikatoren. Die Extremalbedingung (Abstandsfunktion) und die Nebenbedingung (Ebenengleichung) lautet in diesem Fall

$$\begin{aligned} f(x, y, z) &= x^2 + y^2 + z^2 \\ \phi(x, y, z) &= x - 2y - 2z - 3 = 0. \end{aligned}$$

Nach obiger Vorschrift bilden wir also die Funktion $F(x, y, z, \lambda)$ in der Art

$$F(x, y, z, \lambda) = f(x, y, z) + \lambda\phi(x, y, z) = x^2 + y^2 + z^2 + \lambda(x - 2y - 2z - 3)$$

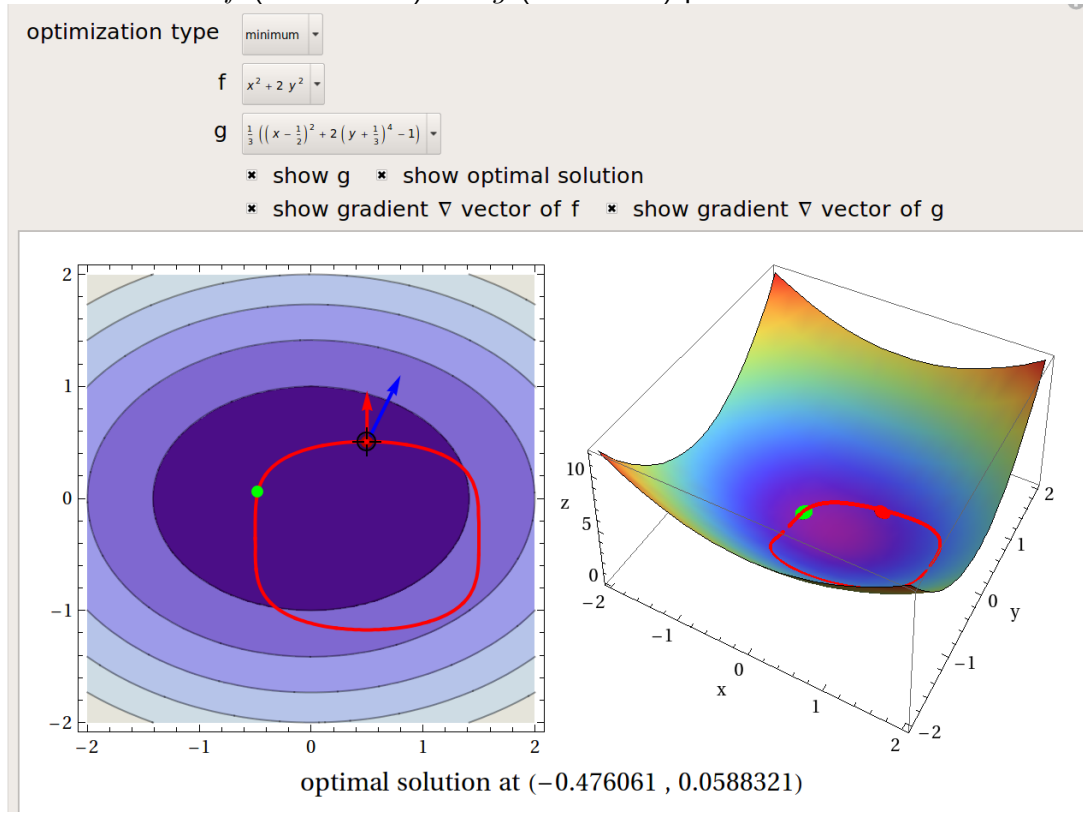
und setzen ihren Gradienten gleich Null:

$$\begin{aligned} \frac{\partial F}{\partial x} = 0 &\Leftrightarrow 2x + \lambda = 0 \\ \frac{\partial F}{\partial y} = 0 &\Leftrightarrow 2y - 2\lambda = 0 \\ \frac{\partial F}{\partial z} = 0 &\Leftrightarrow 2z - 2\lambda = 0 \\ \frac{\partial F}{\partial \lambda} = 0 &\Leftrightarrow x - 2y - 2z - 3 = 0 \end{aligned}$$

Dieses Gleichungssystem können wir am einfachsten auflösen, indem wir aus dem ersten drei Gleichungen ablesen, dass $x = -\frac{\lambda}{2}$, $y = \lambda$, und $z = \lambda$. Eingesetzt in die vierte Gleichung finden wir für $\lambda = -\frac{2}{3}$. Damit erhalten wir das bereits aus dem Eliminationsverfahren bekannte Ergebnis $x = \frac{1}{3}$, $y = -\frac{2}{3}$, und $z = -\frac{2}{3}$. Da Abstand ist entsprechend $d = 1$.

CDF 29. Lagrange Multiplikator [LagrangeMultiplikator2.cdf](#)

Diese CDF-Anwendung veranschaulicht, dass für die gesuchte Lösung gilt: $\nabla f = \lambda \nabla g$, wobei f die zu optimierende Funktion und g Nebenbedingung ist. Das bedeutet also, dass im gesuchten Minimum (grüner Punkt) die Gradienten von f (blauer Pfeil) und g (roter Pfeil) parallel zueinander stehen.



Lagrangesche Multiplikatoren mit mehreren Nebenbedingungen. Analog zur Herleitung des Verfahrens des Lagrangeschen Multiplikators lässt sich leicht zeigen, wie das Verfahren auf *mehrere Nebenbedingungen* erweitert werden kann. Wollen wir eine Funktion f minimieren, die von n Variablen x_1, x_2, \dots, x_n abhängt, und gelten die m Nebenbedingungen $\phi_1(x_1, x_2, \dots, x_n) = 0$, $\phi_2(x_1, x_2, \dots, x_n) = 0$, ..., $\phi_m(x_1, x_2, \dots, x_n) = 0$, dann erhalten wir die Extremstellen als stationäre Punkte der folgenden Funktion:

$$F(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) = f(x_1, x_2, \dots, x_n) + \lambda_1 \phi_1 + \dots + \lambda_m \phi_m.$$

Wir müssen also die folgenden $n+m$ Gleichungen für die Unbekannten x_1, x_2, \dots, x_n und $\lambda_1, \lambda_2, \dots, \lambda_m$ lösen

$$\begin{aligned}\frac{\partial F}{\partial x_i} &= 0 \quad \text{für } i = 1, 2, \dots, n \\ \frac{\partial F}{\partial \lambda_j} &= 0 \quad \text{für } j = 1, 2, \dots, m\end{aligned}$$

Beispiel 2. Es soll wiederum der kürzeste Abstand vom Nullpunkt berechnet werden, dieses mal allerdings zur Schnittkurve folgender beider Flächen: $xy = 1$ und $x + z = 0$. Das heißt, wir wollen die Funktion $f(x, y, z) = x^2 + y^2 + z^2$ minimieren, und haben *zwei* Nebenbedingungen, nämlich $\phi_1(x, y, z) = xy - 1 = 0$ und $\phi_2(x, y, z) = x + z = 0$. Wir suchen also nach einem stationären Punkt der Funktion

$$F(x, y, z, \lambda, \mu) = x^2 + y^2 + z^2 + \lambda(xy - 1) + \mu(x + z).$$

Die Bedingung, dass der Gradient von F verschwinden muss, führt auf folgende $3+2 = 5$ Gleichungen

$$\begin{aligned}\frac{\partial F}{\partial x} = 0 &\Leftrightarrow 2x + \lambda y + \mu = 0 \\ \frac{\partial F}{\partial y} = 0 &\Leftrightarrow 2y + \lambda x = 0 \\ \frac{\partial F}{\partial z} = 0 &\Leftrightarrow 2z + \mu = 0 \\ \frac{\partial F}{\partial \lambda} = 0 &\Leftrightarrow xy - 1 = 0 \\ \frac{\partial F}{\partial \mu} = 0 &\Leftrightarrow x + z = 0\end{aligned}$$

Wir eliminieren mithilfe der ersten beiden Gleichungen den Lagrangeparameter λ , und drücken dann x , y , und z mithilfe der Gleichungen 3–5 durch den zweiten Lagrangeparameter μ aus. Eingesetzt in die Kombination aus Gleichung 1 und 2 erhalten wir $\mu = \pm\sqrt[4]{8}$. Damit erhalten wir für die Variablen x , y , und z

$$x = \pm \frac{1}{\sqrt[4]{2}}, \quad y = \pm \sqrt[4]{2}, \quad z = \mp \frac{1}{\sqrt[4]{2}}.$$

Das Abstandsquadrat $f(x, y, z)$ erhalten wird dann zu $x^2 + y^2 + z^2 = 2\sqrt{2}$.