

Skript zur Vorlesung

**Vorkurs: Mathematische
Methoden der Physik**

von Volker Meden

gehalten im Wintersemester 2005/2006
an der Universität Göttingen

7. Oktober 2005

Inhaltsverzeichnis

1	Einführung	5
2	Differentialrechnung einer Veränderlichen	7
3	Integralrechnung einer Veränderlichen	23
4	Logarithmus- und Exponentialfunktion	31
5	Potenzreihen und Taylorentwicklung	37
6	Hyperbolische und trigonometrische Funktionen	45
7	Komplexe Zahlen	51
8	Vektorrechnung	57
9	Vektorwertige Funktionen	67
10	Funktionen mehrerer Veränderlicher	69
11	Matrizen	77
12	Das Eigenwertproblem und die Hauptachsentransformation	83

Kapitel 1

Einführung

Im Laufe ihres Studiums werden ihnen verschiedene Arten des Umgangs mit Mathematik begegnen, von ziemlich abstrakt bis sehr anschaulich. Der Schwerpunkt des Vorkurses liegt dabei weniger auf formaler Strenge in den Herleitungen, sondern auf einer relativ anschaulichen Darstellung und der Anwendung des Gelernten in physikalischen Fragestellungen. Dabei werden wir was die Vorbedingungen angeht nicht “bei Null anfangen”, sondern setzen ein gewisses (anschauliches) Vorwissen voraus. Zum Beispiel werden wir nicht näher auf den Stetigkeitsbegriff eingehen, sondern gehen davon aus, daß sie ein intuitives Verständnis davon haben, was es heißt, daß eine Kurve stetig ist (ohne Sprünge). Auf saubere Beweise von “Sätzen” - also Zusammenhängen - die ihnen zum Teil sicherlich bereits aus der Schule bekannt sind, und eine abstrakte Formulierung der Konzepte, wie sie sie in den Vorlesungen *Diff. I und II* und *AGLA I* kennenlernen werden, verzichten wir hier bewußt. Die Konzepte sollen vielmehr anhand einfacher Beispiele erklärt und geübt werden.

Ziel des Vorkurses ist es, die zum Verständnis der Vorlesung *Physik I* benötigten mathematischen Methoden bereitzustellen. Aus ihrer “Schulphysik” ist ihnen sicherlich bekannt, daß man zur Beschreibung der Dynamik eines Teilchens in einem Kraftfeld - also in der Mechanik, die den Hauptbestandteil der *Physik I* Vorlesung bildet - Analysis (Differential- und Integralrechnung) und lineare Algebra (Vektorrechnung) benötigt. Hier einige Beispiele aus der Mechanik, in denen die unter diesen Oberbegriffen zusammengefaßten mathematischen Methoden eine wichtige Rolle spielen:

- Bahnkurve - Vektor, Vektoraddition, Funktion: $\vec{x}(t) = x_1(t)\vec{e}_1 + x_2(t)\vec{e}_2 + x_3(t)\vec{e}_3$
- Geschwindigkeit - Differentiation (eindimensional) $v(t) = \frac{dx(t)}{dt} = \dot{x}(t)$
- Geschwindigkeitsvektor - Differentiation eines Vektors: $\vec{v}(t) = \frac{d\vec{x}(t)}{dt} = \dot{\vec{x}}(t)$
- Newtonsche Gleichung - zweite Ableitung, Differentialgleichung, vektorwertige Funktion eines Vektor, zusammengesetzte Funktion: $m\ddot{\vec{x}}(t) = \vec{F}(\vec{x}(t))$

z.B. $m\ddot{x}(t) = -Kx(t)$ (Hookesches Gesetz)

- Arbeit - Integration: $W(x_1 \rightarrow x_2) = \int_{x_1}^{x_2} F(x)dx$
- Leistung - Skalarprodukt: $P = \vec{F} \cdot \vec{v}$
- Drehimpuls - Vektorprodukt: $\vec{l} = \vec{x} \times \vec{p}$
- Potential - partielle Ableitung, Gradient (Differentialrechnung mehrerer Veränderlicher - *Diff. II*): $\vec{F}(\vec{x}) = -\vec{\nabla}V(\vec{x})$
- Rotationsenergie - Summenzeichen, Matrix (Tensor), Hauptachsentransformation: $E_{\text{rot}} = \frac{1}{2} \sum_{i,j=1}^3 \Omega_i I_{i,j} \Omega_j$

Am Ende des Vorkurses sollten sie an den Umgang mit einer Vielzahl dieser Begriffe und Rechenprozeduren gewöhnt sein. Aus Zeitgründen werden wir es nicht schaffen, genauer auf Funktionen *mehrerer Veränderlicher* einzugehen, so daß auch die Begriffe *partielle Ableitung* und *Gradient* nur sehr knapp und beispielhaft behandelt werden können. Weiterhin werden wir nur Zeit haben, den Begriff der Matrix anhand von Beispielen zu behandeln. Bei einem handelt es sich um die in der Mechanik wichtige Drehung eines Koordinatensystems in ein anderes (Koordinatentransformation). Im Vorkurs werden wir weiterhin wichtige Eigenschaften spezieller Funktionen, z.B. der

- Exponentialfunktion: $\exp(x) = e^x$; $\exp(x + y) = \exp(x) \exp(y)$
- Logarithmusfunktion: $\ln(x)$; $\ln(xy) = \ln(x) + \ln(y)$

diskutieren, Reihendarstellungen von Funktionen kennenlernen

- Exponentialfunktion: $\exp(x) = \sum_{n=0}^{\infty} x^n/n!$

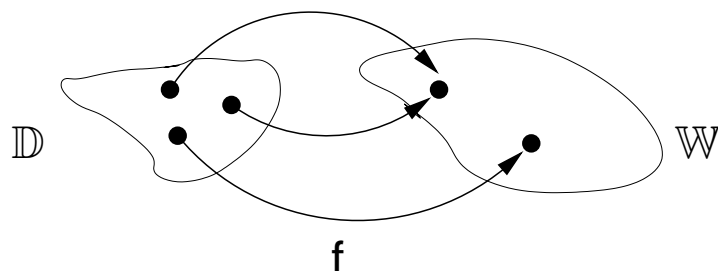
und komplexe Zahlen einführen

- $i = \sqrt{-1}$; $e^{ix} = \cos x + i \sin x$.

Kapitel 2

Differentialrechnung einer Veränderlichen

Eine *Funktion* $y = f(x)$ mit $x \in \mathbb{D}$ und $y \in \mathbb{W}$ ist eine spezielle *Abbildung*, bei der den Elementen des *Definitionsbereichs* \mathbb{D} - der z.B. eine Teilmenge der *reellen Zahlen* \mathbb{R} ist (ein oder mehrere Intervalle) - *eindeutig* ein Element des *Wertebereichs* \mathbb{W} - der eine ebensolche Teilmenge darstellt - zugeordnet wird.



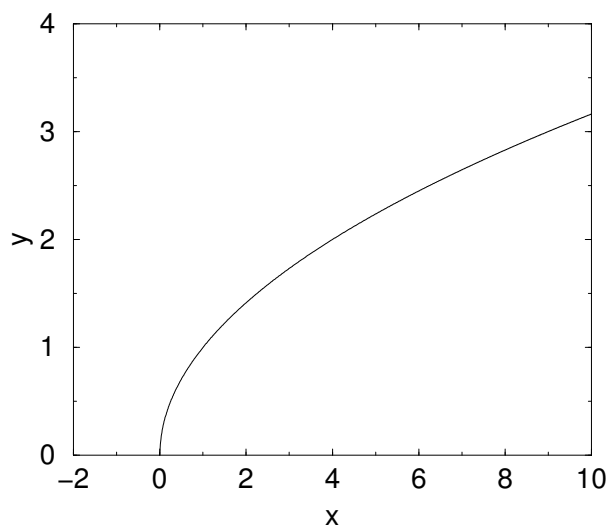
Um Intervalle als Teilmengen von \mathbb{R} zu bezeichnen, verwendet man die Notation

- (a, b) : $\{x \in \mathbb{R} | a < x < b\}$
- $[a, b]$: $\{x \in \mathbb{R} | a \leq x \leq b\}$
- $(a, b]$: $\{x \in \mathbb{R} | a < x \leq b\}$
- $[a, b)$: $\{x \in \mathbb{R} | a \leq x < b\}$

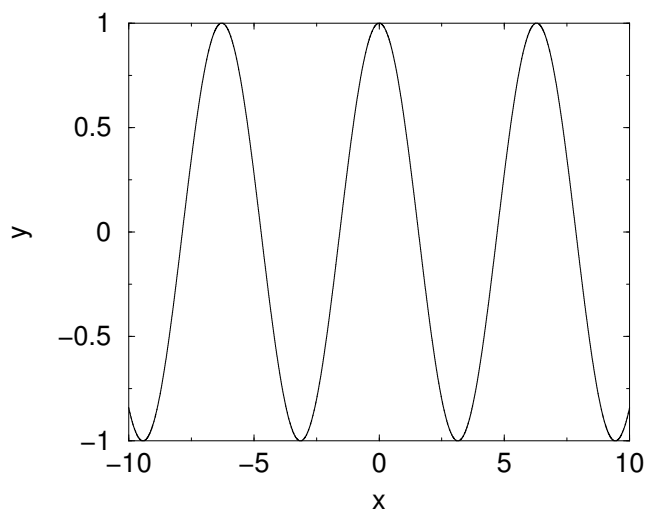
Häufig zeichnet man den *Graphen* der Funktion $f(x)$, welcher durch alle Punktpaare $x, f(x)$ mit $x \in \mathbb{D}$ gebildet wird. In der Physik ist es üblich, $y = f(x)$ zu schreiben. Betrachten wir zwei einfache Beispiele

$$y = \sqrt{x} \quad : \quad \mathbb{D} = [0, \infty) \quad , \quad \mathbb{W} = [0, \infty)$$

8 KAPITEL 2. DIFFERENTIALRECHNUNG EINER VERÄNDERLICHEN



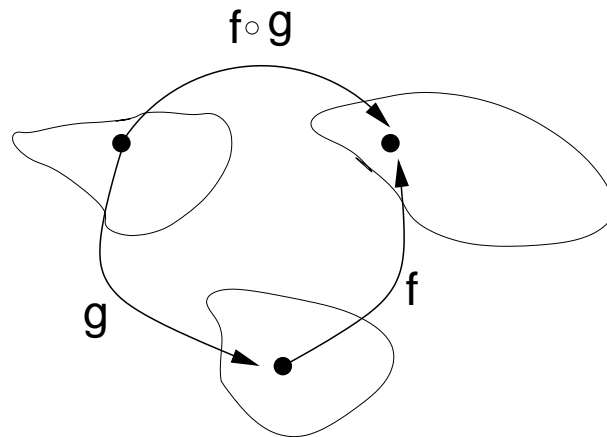
$$y = \cos x : \mathbb{D} = \mathbb{R} = (-\infty, \infty) , \mathbb{W} = [-1, 1]$$



wobei wir immer den größtmöglichen - also den “natürlichen” - Definitionsbereich gewählt haben.

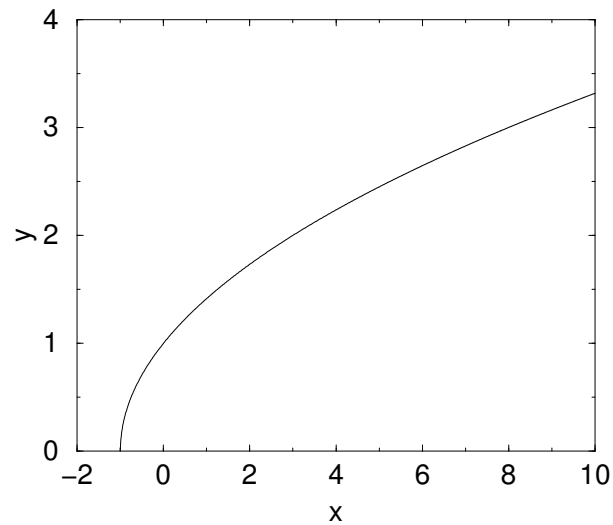
Man kann nun Summen, Differenzen, Produkte und Quotienten von Funktionen $f(x)$, $g(x)$, $h(x)$, ... bilden, wobei man sich in jedem speziellen Fall Gedanken über den Definitions- und Wertebereich der sich jeweils ergebenden Funktionen machen muß. Eine wichtige Rolle spielen auch *zusammengesetzte Funktionen*

$$(f \circ g)(x) = f(g(x))$$

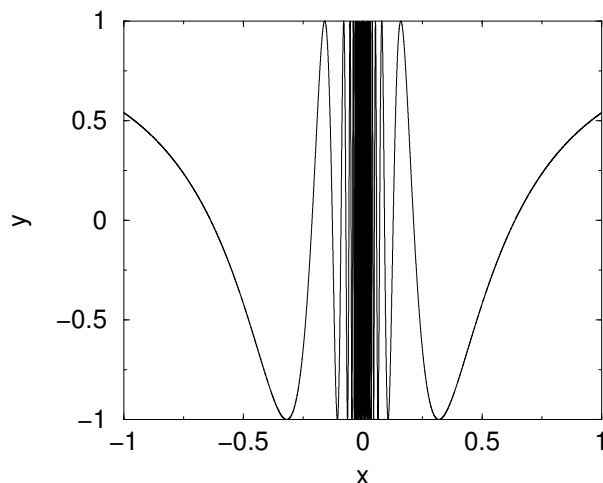


Der Definitionsbereich von $f \circ g$ ergibt sich aus den $x \in \mathbb{D}_g$, für die $g(x) \in \mathbb{D}_f$ ist. Als Beispiele betrachten wir

$$f(x) = \sqrt{x}; \quad g(x) = x + 1 \Rightarrow (f \circ g)(x) = f(g(x)) = \sqrt{x+1}; \quad \mathbb{D}_{f \circ g} = [-1, \infty)$$

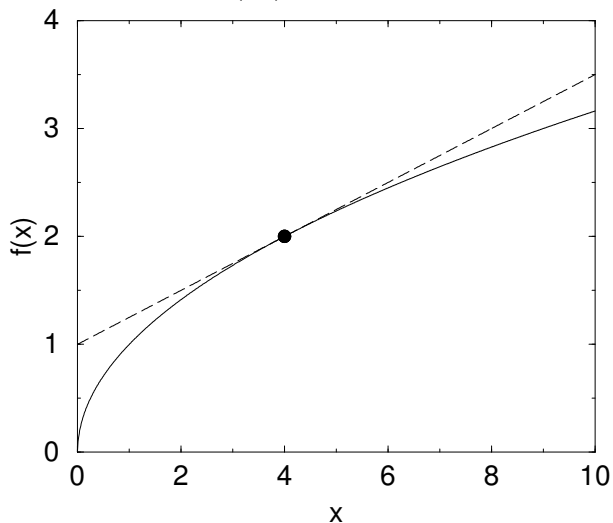


$$f(x) = \cos x; \quad g(x) = \frac{1}{x} \Rightarrow (f \circ g)(x) = f(g(x)) = \cos(1/x); \quad \mathbb{D}_{f \circ g} = \mathbb{R} \setminus \{0\}$$



Die zusammengesetzte Funktion im zweiten Beispiel ist eine “relativ verrückte” Funktion, da der Graph für $|x| \rightarrow 0$ immer schneller zwischen -1 und 1 hin und her “zappelt”.

Man bezeichnet eine Funktion f als *stetig* im Punkt x_0 , wenn der zugehörige Graph bei x_0 keine Sprünge hat ($\lim_{h \rightarrow 0} f(x_0+h) = f(x_0)$) und der Funktionswert bei Annäherung an x_0 kleiner Unendlich ist ($\lim_{h \rightarrow 0} f(x_0+h) < \infty$), d.h. f bei x_0 nicht *divergiert*. Eine formale Definition der Begriffe Stetigkeit und des oben stillschweigend benutzen *Grenzwertes* überlassen wir der *Diff. I* Vorlesung und gehen gleich zum Begriff der *Ableitung* über. Dazu nehmen wir an, daß die Funktion f in einer *Umgebung* (einem Intervall) um x_0 stetig ist. In der Schule wird die Ableitung meist mit Hilfe der *Sekante* durch zwei Punktepaare $x_1, f(x_1)$ und $x_2, f(x_2)$ mit $x_1 < x_0 < x_2$ definiert, die im Limes $x_1 \rightarrow x_0$ und $x_2 \rightarrow x_0$ in die *Tangente* am Punktepaar $x_0, f(x_0)$ übergeht.



Etwas formaler gilt

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

vorausgesetzt, daß dieser Grenzwert existiert. Anschaulich ist $f'(x_0)$ die *Steigung* der Tangente im Punkt x_0 . Falls $f'(x_0)$ existiert, so ist die Funktion im Punkt x_0 durch die Tangente *linear approximierbar*, d.h. näherungsweise durch eine *lineare* Funktion der Form $g(x) = a + bx$ beschreibbar (in der obigen Abbildung z.B. $a = 1$ und $b = 1/4$), und f ist in x_0 *differenzierbar*. Eine Funktion ist in x_0 z.B. dann nicht differenzierbar, wenn sie in diesem Punkt einen Knick hat. Wie wir später noch genauer diskutieren werden, spielt eine näherungsweise Beschreibung komplizierter Funktionen in der Physik eine sehr wichtige Rolle. Betrachten wir nun einige Beispiele, wie der Grenzprozeß $h \rightarrow 0$ ausgeführt werden kann und welche Ableitungen f' sich ergeben. Für $f(x) = \sqrt{x}$ mit $x > 0$ und $x + h > 0$ gilt

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{\sqrt{x+h} - \sqrt{x}}{h} \\ &= \frac{(\sqrt{x+h} - \sqrt{x})(\sqrt{x+h} + \sqrt{x})}{h(\sqrt{x+h} + \sqrt{x})} \\ &= \frac{h}{h(\sqrt{x+h} + \sqrt{x})} \\ &= \frac{1}{\sqrt{x+h} + \sqrt{x}}. \end{aligned}$$

Jetzt kann der Limes $h \rightarrow 0$ einfach durch Nullsetzen von h ausgeführt werden, und man erhält

$$\sqrt{x}' = \frac{1}{2\sqrt{x}} \quad \text{oder} \quad (x^{1/2})' = \frac{1}{2}x^{-1/2}.$$

An der Rechnung erkennen wir auch, daß \sqrt{x} bei $x = 0$ nicht differenzierbar ist, denn die Steigung der Tangente wird für $x \rightarrow 0$ immer größer und divergiert im Punkt $x = 0$. Als nächstes betrachten wir $f(x) = \sin x$. Um f' zu bestimmen, benötigen wir die *Additionstheoreme* der *trigonometrischen Funktionen* (siehe Übungen):

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y \quad (2.1)$$

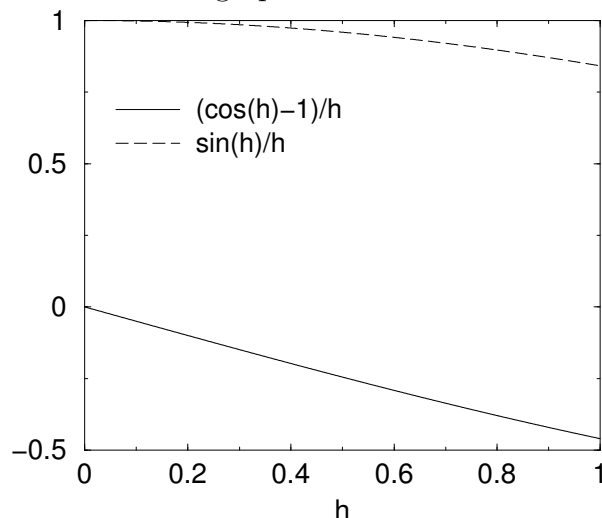
$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y \quad (2.2)$$

Mit Hilfe von Gl. (2.1) folgt:

$$\begin{aligned} (\sin x)' &= \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h \rightarrow 0} \left[\sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h} \right] \end{aligned}$$

12 KAPITEL 2. DIFFERENTIALRECHNUNG EINER VERÄNDERLICHEN

Wir haben nun noch die beiden Grenzwerte $\lim_{h \rightarrow 0} \frac{\cos h - 1}{h}$ und $\lim_{h \rightarrow 0} \frac{\sin h}{h}$ zu untersuchen. Das wollen wir hier graphisch tun¹



Also liefert der erste Grenzwert 0 und der zweite 1. Damit ergibt sich

$$(\sin x)' = \cos x.$$

Analog zeigt man (siehe Übungen):

$$(\cos x)' = -\sin x.$$

Wichtig - und ihnen sicherlich aus der Schule bereits bekannt - ist auch die Ableitung der Funktion $f_n(x) = x^n$ mit $n \in \mathbb{N}_0$.² Sie bestimmt sich aus

$$f'_n(x) = \lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h}.$$

Um den Limes $h \rightarrow 0$ ausführen zu können, müssen wir den Ausdruck $(x+h)^n$ *ausmultiplizieren*. Wir benötigen also

$$\begin{aligned} (x+h)^n &= \underbrace{(x+h)(x+h)\dots(x+h)}_{n\text{-mal}} \\ &= x^n + \frac{n}{1}x^{n-1}h + \frac{n(n-1)}{2 \cdot 1}x^{n-2}h^2 + \frac{n(n-1)(n-2)}{3 \cdot 2 \cdot 1}x^{n-3}h^3 + \dots + h^n. \end{aligned}$$

Das *kombinatorische* Problem, den Vorfaktor des Terms $x^{n-m}h^m$ mit $m \leq n$ zu bestimmen, führt uns auf die Begriffe *Fakultät* und *Binomialkoeffizient*.

Die Fakultät ist definiert als

$$n! = n(n-1)(n-2)(n-3) \cdot \dots \cdot 1$$

¹Mit Hilfe der Reihendarstellung der sin und cos Funktionen werden wir diese Grenzwerte später auch analytisch recht einfach bestimmen können (siehe Übungen).

² \mathbb{N} bezeichnet die *natürlichen* Zahlen $1, 2, 3, \dots$ und $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

wobei $n \in \mathbb{N}$ gilt. Ergänzt wird diese Definition durch $0! = 1$. Die Zahl der Möglichkeiten, n verschiedene Objekte anzuordnen, ist durch $n!$ gegeben. Zum Beispiel gibt es $4! = 24$ Möglichkeiten, die vier Buchstaben A, B, C, D anzuordnen: 4 Möglichkeiten für den ersten Buchstaben * 3 Möglichkeiten für den zweiten * 2 Möglichkeiten für den dritten * 1 Möglichkeit für den vierten. Abstrakter formuliert heißt das: Die Zahl der *Permutationen* von n verschiedenen Objekten ist durch $n!$ gegeben. Um das obige kombinatorische Problem “Wieviele Möglichkeiten gibt es in den n -Faktoren $(x + h)$, $(n - m)$ x 'e und m h 's zu finden” zu lösen, müssen wir einen Schritt weitergehen. Die m h 's kann man auf $\frac{n(n-1)(n-2)\dots(n-m+1)}{m!}$ Arten erhalten:

- n Möglichkeiten für das erste h
- $n - 1$ Möglichkeiten für das zweite h
- ...
- ...
- $n - m + 1$ Möglichkeiten für das m -te h

Dabei hat man jetzt jedoch die $m!$ Vertauschungen der Reihenfolge der Wahl als eigenständig gezählt, so daß man noch durch $m!$ dividieren muß. Den sich ergebenden Faktor bezeichnet man als Binomialkoeffizienten

$$\binom{n}{m} = \frac{n(n-1)(n-2)\dots(n-m+1)}{m!} = \frac{n!}{(n-m)!m!}.$$

Der Binomialkoeffizient gibt z.B. auch die Zahl der Kombinationen im Lotto “6 aus 49”: $\binom{49}{6} = 13983816$. Damit ergibt sich für $(x + h)^n$:

$$\begin{aligned} (x + h)^n &= \binom{n}{0}x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \dots + \binom{n}{n}h^n \\ &= \sum_{m=0}^n \binom{n}{m}x^{n-m}h^m. \end{aligned}$$

Wir haben in diesem Ausdruck das *Summenzeichen*

$$\sum_{m=0}^n = \text{Summe über } m \text{ von } 0 \text{ bis } n$$

eingeführt.

14 KAPITEL 2. DIFFERENTIALRECHNUNG EINER VERÄNDERLICHEN

Nach diesem Einschub kommen wir auf das Problem $(x^n)' = \dots$ zurück. Es ergibt sich

$$\begin{aligned} (x^n)' &= \lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h \rightarrow 0} \frac{x^n + \frac{n}{1}x^{n-1}h + \frac{n(n-1)}{2 \cdot 1}x^{n-2}h^2 + \frac{n(n-1)(n-2)}{3 \cdot 2 \cdot 1}x^{n-3}h^3 + \dots + h^n - x^n}{h} \\ &= nx^{n-1}, \end{aligned} \tag{2.3}$$

wobei der Limes $h \rightarrow 0$ jetzt durch Nullsetzen von h ausgeführt werden kann. Über einige Umwege haben wir so das ihnen bekannte Ergebnis erhalten.

Es gibt mehrere wichtige Rechenregeln, die einem helfen, Ableitungen von Funktionen zu bestimmen.³ Seien f und g differenzierbare Funktionen, so gilt

- $(f \pm g)' = f' \pm g'$
- Produktregel: $(f \cdot g)' = f' \cdot g + f \cdot g'$
- Quotientenregel: $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$

Den Beweis dieser Regeln verschieben wir auf die Übungen. Mit Hilfe der Quotientenregel läßt sich das Ergebnis Gl. (2.3) auf negative ganze Zahlen $n = -|m|$ erweitern (siehe Übungen), so daß jetzt für alle *ganzen* Zahlen $n \in \mathbb{Z}$, $(x^n)' = nx^{n-1}$ gilt. Genauer wollen wir hier die *Kettenregel* betrachten, die angibt, wie man zusammengesetzte Funktionen differenziert. Für die differenzierbaren Funktionen f und g müssen wir den Limes

$$\lim_{h \rightarrow 0} \frac{f(g(x+h)) - f(g(x))}{h} = \lim_{h \rightarrow 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \frac{g(x+h) - g(x)}{h}$$

ausführen, wobei wir angenommen haben, daß $g(x+h) \neq g(x)$ gilt, $g(x)$ also keine konstante Funktion ist - in welchem Fall $f(g(x+h)) - f(g(x)) = 0$ wäre und damit $(f \circ g)' = 0$ gelten würde. Mit der Abkürzung $\tilde{h} = g(x+h) - g(x)$ gilt

$$\lim_{h \rightarrow 0} \frac{f(g(x+h)) - f(g(x))}{h} = \lim_{h \rightarrow 0} \frac{f(g(x) + \tilde{h}) - f(g(x))}{\tilde{h}} \frac{g(x+h) - g(x)}{h}.$$

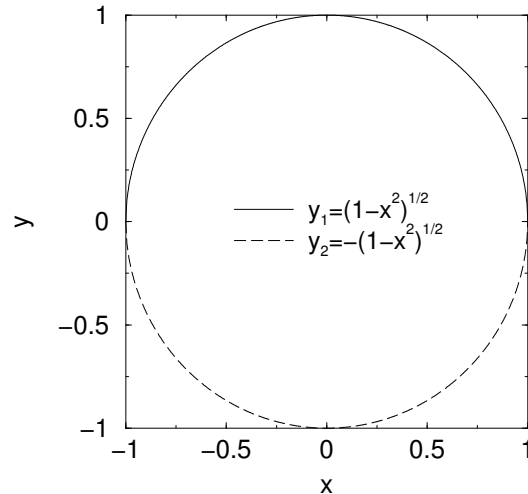
Wegen der Stetigkeit von g geht mit h auch \tilde{h} gegen 0, so daß

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

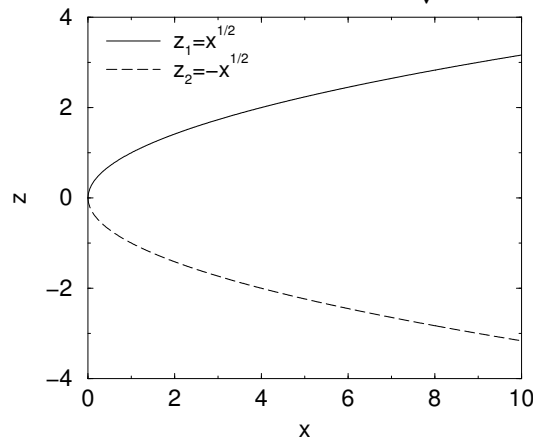
folgt. Eine "saubere" Ausformulierung des Beweises werden sie in *Diff. I* kennenlernen.

³Zum Beispiel läßt sich die Ableitung der *Tangensfunktion* $\tan x = \frac{\sin x}{\cos x}$ leicht mit Hilfe der Quotientenregel bestimmen (siehe später).

Als nächstes werden wir das Konzept der *impliziten Funktion* diskutieren, das für sich genommen wichtig ist und oft auch bei der Bestimmung von Ableitungen hilfreich sein kann. Durch eine Gleichung der Form $F(x, y) = 0$ wird im allgemeinen eine Kurve in der x - y -Ebene festgelegt. Zum Beispiel beschreibt $F(x, y) = x^2 + y^2 - 1 = 0$ einen Kreis mit Radius 1, den man als *Einheitskreis* bezeichnet. Löst man die Gleichung $x^2 + y^2 - 1 = 0$ nach y auf, so ergeben sich *zwei* Funktionen $y = \pm\sqrt{1 - x^2}$, wobei $x \in [-1, 1] = \mathbb{D}$ zu wählen ist.



Als weiteres Beispiel betrachten wir die Gleichung $G(x, z) = z^2 - x = 0$. Aufgelöst nach z ergeben sich wieder zwei Funktionen $z = \pm\sqrt{x}$ mit $\mathbb{D} = [0, \infty)$.



Die Ableitung des oberen Parabelastes kennen wir bereits:⁴ $\frac{dz_1}{dx} = \frac{1}{2\sqrt{x}}$. Man kann die Ableitung aber auch ohne Auflösen nach z erhalten, wenn man die Methode

⁴Wir schreiben hier zum ersten mal statt z_1' für die Ableitung der Funktion z_1 nach der Variablen x , $\frac{dz_1}{dx}$. Wir werden diese Bezeichnung für die Ableitung im folgenden häufiger verwenden. Es gilt zu beachten, daß $\frac{dz_1}{dx}$ kein "Bruch" im eigentlichen Sinne ist. Man darf also in der Gleichung $\frac{dz_1}{dx} = \frac{1}{2\sqrt{x}}$ *nicht* beide Seiten mit dx multiplizieren, da dx für sich genommen erst dann einen Sinn macht, wenn man den Umgang mit Differentialformen gelernt hat.

16 KAPITEL 2. DIFFERENTIALRECHNUNG EINER VERÄNDERLICHEN

der *impliziten Differentiation* verwendet:

$$\begin{aligned} z^2 &= x & \Big| & \frac{d}{dx} \\ 2z \frac{dz}{dx} &= 1 \\ \frac{dz}{dx} &= \frac{1}{2z}. \end{aligned}$$

Mit Hilfe dieses Ausdrucks erhalten wir auch gleich die Ableitungen für die beiden Funktionen:

$$\frac{dz_1}{dx} = \frac{1}{2\sqrt{x}} \quad ; \quad \frac{dz_2}{dx} = -\frac{1}{2\sqrt{x}}.$$

Besonders hilfreich ist die Differentiation impliziter Funktionen, wenn es nicht möglich ist, die Gleichung $F(x, y) = 0$ nach y aufzulösen. Zum Beispiel $F(x, y) = x^3 + y^3 - 9xy$:

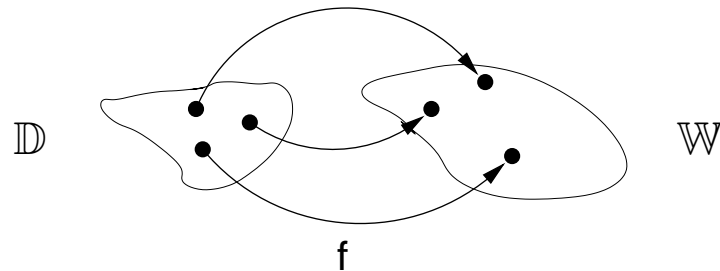
$$\begin{aligned} x^3 + y^3 - 9xy &= 0 & \Big| & \frac{d}{dx} \\ 3x^2 + 3y^2 \frac{dy}{dx} - 9y - 9x \frac{dy}{dx} &= 0 \\ (3y^2 - 9x) y' &= 9y - 3x^2 \\ y' &= \frac{9y - 3x^2}{3y^2 - 9x}. \end{aligned}$$

Wir können die implizite Differentiation auch verwenden, um die Regel $(x^n)' = nx^{n-1}$ (siehe Gl. (2.3)) auf *rationale* $\alpha = m/n$ mit $m, n \in \mathbb{Z}$ zu verallgemeinern. Die Menge der rationalen Zahlen wird oft mit \mathbb{Q} bezeichnet. Für $x > 0$ betrachten wir $y = x^{m/n}$, also $y^n = x^m$. Damit ergibt sich

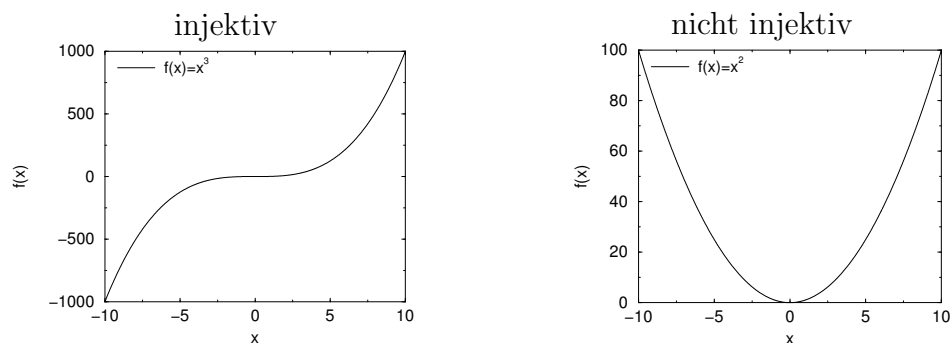
$$\begin{aligned} y^n &= x^m & \Big| & \frac{d}{dx} \\ ny^{n-1} \frac{dy}{dx} &= mx^{m-1} \\ y' &= \frac{m x^{m-1}}{n y^{n-1}} \\ y' &= \frac{m x^{m-1}}{n (x^{m/n})^{n-1}} \\ y' &= \frac{m x^{m-1}}{n x^{m-m/n}} \\ y' &= \frac{m}{n} x^{m/n-1}, \end{aligned}$$

also für rationale α : $(x^\alpha)' = \alpha x^{\alpha-1}$.

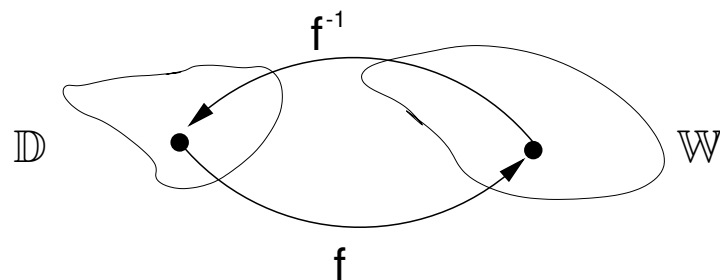
Aus der Schule kennen sie sicherlich das Konzept der *Umkehrfunktion*. Um diese einzuführen, müssen wir zunächst einen weiteren Begriff definieren: Eine Funktion f (allgemeiner eine Abbildung) heißt *injektiv* oder *eineindeutig*, falls für alle beliebigen Paare $x_1 \neq x_2$ aus \mathbb{D} auch $f(x_1) \neq f(x_2)$. Nur wenn diese Eigenschaft erfüllt ist, kann man eine Funktion umkehren.



Dazu zwei Beispiele:



Für eine injektive Funktion kann man die Umkehrfunktion f^{-1} definieren.



Für diese gilt dann für alle $x \in \mathbb{D}_f$

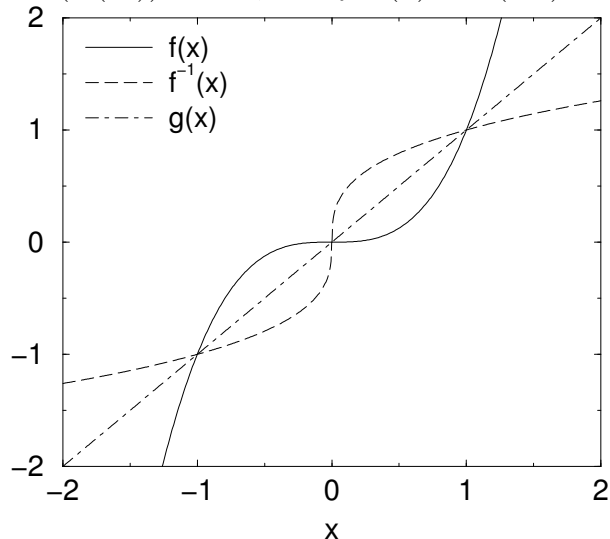
$$f^{-1}(f(x)) = x, \quad (2.4)$$

und für alle $x \in \mathbb{W}_f = \mathbb{D}_{f^{-1}}$

$$f(f^{-1}(x)) = x.$$

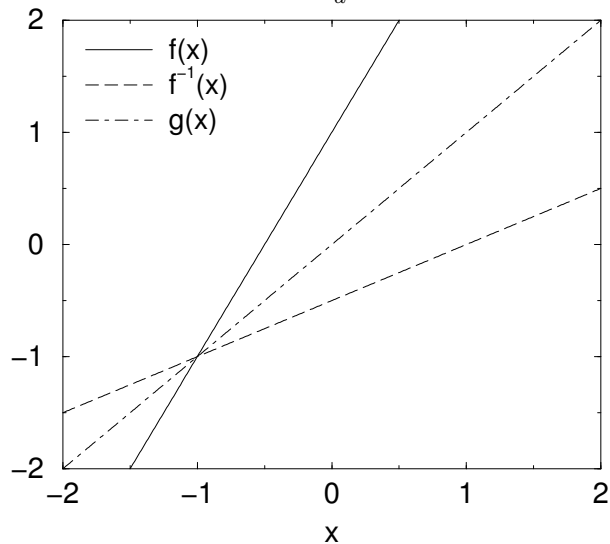
Wir betrachten zunächst einige einfache Beispiele zur Konstruktion von f^{-1} und zur Berechnung von $\frac{df^{-1}}{dx} = (f^{-1})' = f^{-1}'$ bei gegebenem f .

- $f(x) = x^3$: Da für $x \geq 0$, $(x^3)^{1/3} = x$ gilt erhält man $f^{-1}(x) = x^{1/3}$, und für $x < 0$ gilt $-((-x^3)^{1/3}) = x$, also $f^{-1}(x) = -(-x)^{1/3}$.



Die Ableitung von f^{-1} für $x > 0$ ergibt sich zu $f^{-1'}(x) = \frac{1}{3}x^{-2/3}$. Für $x < 0$ folgt mit Hilfe der Kettenregel $f^{-1'}(x) = \frac{1}{3}(-x)^{-2/3}$. Damit ist, wie man auch der Abbildung entnimmt, der Wert der Ableitung von f^{-1} bei x gleich dem bei $-x$.

- $f(x) = ax + b$, mit $a = 2$ und $b = 1$. Zur Bestimmung von f^{-1} lösen wir die Gleichung $y = ax + b$ nach x auf $x = \frac{1}{a}(y - b)$ und erhalten f^{-1} durch Vertauschen von x und y : $f^{-1}(x) = \frac{1}{a}(x - b)$.



Es gilt $f^{-1'}(x) = \frac{1}{a}$. Da $f'(x) = a$, finden wir, daß die Steigung der Inversen der linearen Funktion durch das Inverse der Steigung der linearen Funktion

selbst gegeben ist. Dieses ist die zentrale Beobachtung, mit deren Hilfe wir $f^{-1}'(x)$ allgemein bestimmen können.

Aus den obigen Abbildungen erkennt man eine allgemeingültige graphische Konstruktionsvorschrift, wie sich f^{-1} aus f bestimmen läßt: Durch Spiegelung an der *Winkelhalbierenden* $g(x) = x$. Aus dem zweiten Beispiel ergibt sich sofort auch ein Konstruktionsprinzip für f^{-1} : Wenn die Tangente an f durch $x_0, f(x_0)$ die Steigung $f'(x_0)$ hat, so hat die Tangente an f^{-1} durch $f(x_0), f^{-1}(f(x_0))$ die Steigung $1/f'(x_0)$, so daß allgemein

$$f^{-1}'(f(x_0)) = \frac{1}{f'(x_0)} \quad (2.5)$$

gilt. Weniger “konstruktiv” läßt sich Gl. (2.5) mit Hilfe der Definitionsgleichung Gl. (2.4) und der impliziten Differentiation herleiten (siehe Übungen). Mit Hilfe der Gl. (2.5) läßt sich z.B. die Ableitung der arctan-Funktion leicht bestimmen. Mit $f(x) = \tan x = \frac{\sin x}{\cos x}$ folgt nach der Quotientenregel

$$f'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

Damit gilt also

$$\frac{1}{f'(x_0)} = \cos^2 x_0 = \frac{\cos^2 x_0}{\sin^2 x_0 + \cos^2 x_0} = \frac{1}{1 + \tan^2 x_0},$$

so daß sich nach Gl. (2.5)

$$\arctan'(\tan x_0) = \frac{1}{1 + \tan^2 x_0}$$

und damit

$$(\arctan x)' = \frac{1}{1 + x^2}$$

ergibt.

Wie wir bereits in der Einleitung gesehen haben, benötigt man in der Mechanik auch *höhere Ableitungen* $f^{(n)}(x)$ einer Funktion f . Diese sind für $n = 2, 3, 4, \dots$ rekursiv definiert

$$\frac{d^n f(x)}{dx^n} = \frac{d}{dx} \left(\frac{d^{n-1} f(x)}{dx^{n-1}} \right).$$

Für “kleine” n verwendet man auch die “ n -Striche” (bzw. “ n -Punkte”) Notation f'', f''', \dots

Die trigonometrischen Funktionen haben dabei die interessante Eigenschaft

$$\begin{aligned}(\sin x)'' &= (\cos x)' = -\sin x, \\(\cos x)'' &= (-\sin x)' = -\cos x.\end{aligned}$$

Damit gilt für die Funktion $y(x) = a \sin x + b \cos x$, $y''(x) = -(a \sin x + b \cos x) = -y(x)$. Umgeformt folgt

$$y''(x) + y(x) = 0. \quad (2.6)$$

Diese Gleichung führt uns auf ein sehr wichtiges weiteres Konzept: Die *Differentialgleichung*. Eine Differentialgleichung ist eine Gleichung, in der die Funktion $y(x)$, ihre Ableitungen $y'(x), y''(x), \dots, y^{(n)}(x)$ und x selbst auftreten. Ist die n -te Ableitung die höchste vorkommende, so spricht man von einer *Differentialgleichung n -ter Ordnung*. Gl. (2.6) ist also eine Differentialgleichung zweiter Ordnung. Im allgemeinen ist es sehr kompliziert und oft unmöglich, die Lösung einer Differentialgleichung geschlossen (analytisch) anzugeben. Differentialgleichungen spielen in der Physik eine sehr wichtige Rolle. Zum Beispiel bildet die (eindimensionale) Newtonsche Gleichung $m\ddot{x}(t) = F(x(t))$ eine Differentialgleichung zweiter Ordnung. Ist nun die Kraft F proportional der Auslenkung x einer Masse m - was einer näherungsweisen Beschreibung der Rückstellkraft einer Feder entspricht - so ergibt sich das Hookesche Gesetz $m\ddot{x}(t) = -Kx(t)$ (K heißt die Rückstellkonstante), oder

$$\ddot{x}(t) + \frac{K}{m}x(t) = 0, \quad (2.7)$$

was bis auf einen Faktor der Gleichung (2.6) entspricht. Um eine Lösung für Gl. (2.7) zu finden, müssen wir die Lösung von Gl. (2.6) noch leicht modifizieren. Dazu berechnen wir zunächst die ersten und zweiten Ableitungen von $f(t) = \sin(\omega t)$ und $g(t) = \cos(\omega t)$ mit Hilfe der Kettenregel

$$\begin{aligned}\dot{f}(t) &= \omega \cos(\omega t) \quad ; \quad \ddot{f}(t) = -\omega^2 \sin(\omega t) = -\omega^2 f(t) \\ \dot{g}(t) &= -\omega \sin(\omega t) \quad ; \quad \ddot{g}(t) = -\omega^2 \cos(\omega t) = -\omega^2 g(t).\end{aligned}$$

Bilden wir jetzt wieder die *Linearkombination* $x(t) = af(t) + bg(t)$, so ist diese Funktion $x(t)$ eine Lösung von Gl. (2.7), falls $\omega = \sqrt{\frac{K}{m}}$. Im Problem der Bewegung eines Teilchens der Masse m , welches mit einer Feder mit der Federkonstanten K verbunden ist, muß die Lösung $x(t)$ jetzt noch zwei *Anfangsbedingungen* (oder *Randbedingungen*) erfüllen. Zur Zeit $t = 0$ soll das Teilchen am festgelegten Ort x_0 sein, so daß $b = x_0$ zu gelten hat, und es soll die Anfangsgeschwindigkeit v_0 haben, woraus $a = v_0/\omega = v_0\sqrt{\frac{m}{K}}$ folgt. Die Lösung der Newtonschen Gleichung lautet für diesen Fall somit

$$x(t) = x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t) \quad \text{mit} \quad \omega = \sqrt{\frac{K}{m}}.$$

Eine genauere Diskussion der sich ergebenden *oszillierenden* Bewegung verschieben wir auf die *Physik I* Vorlesung.

Zum Abschluß dieses Kapitels zwei wichtige Hinweise: Die vielen oben diskutierten Rechenricks sind bei der Bestimmung der Ableitung einer Funktion sehr hilfreich. Gerade für Physiker ist es jedoch auch legitim, sich durch graphische Darstellung der zu differenzierenden Funktion $f(x)$ ein anschauliches Bild von der Ableitung $f'(x)$ als Steigung von $f(x)$ im Punkt x zu verschaffen - und sei es nur als Konsistenzcheck des formal gewonnenen Ergebnisses. Dazu untersucht man für die Funktion f am besten Fragen der Art: Wo ist die Steigung groß, wo klein? Ist sie positiv oder negativ? Wo findet ein Vorzeichenwechsel der Steigung statt? In der Physik kennt man die zu differenzierende Funktion oft nicht *analytisch* - z.B. wenn man eine (eindimensionale) Bahnkurve eines Teilchens $x(t)$ experimentell bestimmt hat und an der Geschwindigkeit des Teilchens $v(t)$ interessiert ist. In diesem Fall muß die Differentiation *numerisch* ausgeführt werden. Leider ist die numerische Differentiation kein sehr "stabiles" Verfahren, was heißt, daß sich kleine Meßfehler bei der numerischen Differentiation verstärken können (siehe spätere Mathematikvorlesungen). Die einfachste Form der numerischen Differentiation kennen sie sicherlich bereits aus der "Schulphysik". Nehmen wir z.B. an, der Ort x eines Teilchens wurde zu den Zeiten $t_n = t_0 + nh$, mit $n \in \mathbb{N}_0$ und festem h , bestimmt. In diesem Fall können wir die Geschwindigkeit zur Zeit t_n , $v(t_n)$, näherungsweise zu $v(t_n) = \frac{x(t_{n+1}) - x(t_n)}{h}$ (Vorwärtsdifferenzenquotient) bestimmen. Genauso legitim wäre eine Zuordnung $v(t_n) = \frac{x(t_n) - x(t_{n-1})}{h}$ (Rückwärtsdifferenzenquotient). Eine glattere Kurve für die Geschwindigkeit liefert allerdings oft die Zuordnung $v(t_n) = \frac{x(t_{n+1}) - x(t_{n-1})}{2h}$ (zentrierter Differenzenquotient), bei der die Steigung im Punktepaar $t_n, x(t_n)$ gleich der Steigung der Sekante durch $t_{n-1}, x(t_{n-1})$ und $t_{n+1}, x(t_{n+1})$ gesetzt wird.

Kapitel 3

Integralrechnung einer Veränderlichen

Bereits gegen Ende des letzten Kapitels haben wir Differentialgleichungen betrachtet. Wir wollen nun die einfachste Differentialgleichung erster Ordnung

$$y'(x) = f(x)$$

untersuchen. Wir suchen also zu einer gegebenen Ableitung $y'(x) = f(x)$ eine zugehörige Funktion $y(x) = F(x)$, so daß $F'(x) = f(x)$ gilt. Man nennt eine Lösung der Differentialgleichung, also die Funktion $F(x)$, eine *Stammfunktion* von $f(x)$. Das sich stellende Problem ist die *Umkehrung der Ableitung*. Bereits im letzten Kapitel haben wir darauf hingewiesen, daß es im allgemeinen kein "Rezept" gibt, die Lösung einer Differentialgleichung zu bestimmen. Gleiches gilt für das Finden einer Stammfunktion. Einfache Klassen von Stammfunktionen erhält man durch das Umdrehen der Ableitungsrelationen, z.B. $(x^{n+1})' = (n+1)x^n$, so daß $F_c(x) = \frac{1}{n+1}x^{n+1} + c$, mit der beliebigen reellen Konstanten c , für $n+1 \neq 0$ eine Stammfunktion von $f(x) = x^n$ ist. Hier eine kleine Tabelle von Stammfunktionen

Funktion f	Stammfunktion F
$x^n ; n \neq -1$	$\frac{1}{n+1}x^{n+1} + c$
$\sin(ax) ; a \neq 0$	$-\frac{1}{a}\cos(ax) + c$
$\cos(ax) ; a \neq 0$	$\frac{1}{a}\sin(ax) + c$
$\frac{1}{1+x^2}$	$\arctan x + c$

Aus der Definition der Stammfunktion ergibt sich eine offensichtliche Rechenregel: Ist F Stammfunktion von f und G Stammfunktion von g , so ist $aF(x) + bG(x)$ Stammfunktion von $af(x) + bg(x)$. Zum Beispiel können wir mit

Hilfe dieser Regel eine Stammfunktion von $f(x) = \sin^2 x$ bestimmen

$$\begin{aligned} f(x) &= \sin^2 x = \frac{1}{2} - \frac{1}{2} \cos(2x) \\ \Rightarrow F(x) &= \frac{1}{2}x - \frac{1}{4} \sin(2x) + c. \end{aligned}$$

Zunächst einmal wollen wir noch die ihnen wahrscheinlich bekannte übliche Schreibweise (*unbestimmtes Integral*)

$$F(x) = \int f(x) dx \quad (3.1)$$

für eine Stammfunktion F zur Funktion f vermeiden. Sie wäre an dieser Stelle eher verwirrend, denn speziell ist nicht klar, was der Ausdruck dx hier zu bedeuten hat. Stattdessen schreiben wir

$$F(x) = \text{SF}[f(x)].$$

Als physikalisches Beispiel betrachten wir den freien, senkrechten Fall eines Steins. Die Newtonsche Gleichung für die zeitabhängige Geschwindigkeit $v(t)$ (nach unten) lautet in diesem Fall

$$\frac{dv(t)}{dt} = g \quad ; \quad \text{mit der Erdbeschleunigung } g \approx 9,81 \text{ m/s}^2.$$

Sie stellt damit eine Differentialgleichung von dem soeben diskutierten Typ dar, wobei die Funktion $f(t)$ durch die konstante Funktion $f(t) = g$ gegeben ist. Eine Stammfunktion zu $f(t) = g$ wird durch $v(t) = gt + c$ gegeben. Setzt man nun $t = 0$, so folgt $v(0) = c$, mit der Anfangsgeschwindigkeit $v(0)$, so daß wir bei vorgegebener Anfangsgeschwindigkeit auch $v(t) = v(0) + gt$ schreiben können.

Aus den Umkehrungen der Produkt- und der Kettenregel für die Differentiation ergeben sich zwei wichtige Tricks zur Bestimmung der Stammfunktion.

- Die Produktregel lautet

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

woraus sich

$$\text{SF}[f'(x)g(x) + f(x)g'(x)] = f(x)g(x)$$

oder

$$\text{SF}[f'(x)g(x)] = f(x)g(x) - \text{SF}[f(x)g'(x)] \quad (3.2)$$

ergibt. Kennt man also eine Stammfunktion von $f(x)g'(x)$ (oder $f'(x)g(x)$), so folgt eine von $f'(x)g(x)$ (oder $f(x)g'(x)$) mit Hilfe von Gl. (3.2). Man

nennt diesen Rechenrick *partielle Integration* - eine Bezeichnung, deren Ursprung erst später klar werden wird. Als Beispiel betrachten wir

$$\begin{aligned} f(x) = \sin x \quad ; \quad g(x) = x &\Rightarrow f'(x) = \cos x \quad ; \quad g'(x) = 1 \\ \Rightarrow \text{SF} [x \cos x] = x \sin x - \text{SF} [\sin x] &= x \sin x + \cos x + c \end{aligned}$$

- Sei $F(x)$ eine Stammfunktion von $f(x)$ und $g(x)$ eine differenzierbare Funktion, so lautet die Kettenregel

$$\frac{d}{dx} F(g(x)) = F'(g(x))g'(x) = f(g(x))g'(x).$$

Damit gilt die *Substitutionsregel*

$$\text{SF} [f(g(x))g'(x)] = F(g(x)). \quad (3.3)$$

Zum Beispiel

$$\begin{aligned} f(x) = \sin x \quad ; \quad g(x) = x^2 &\Rightarrow F(x) = -\cos x \quad ; \quad g'(x) = 2x \\ \Rightarrow \text{SF} [2x \sin(x^2)] &= -\cos(x^2) + c \end{aligned}$$

Die beiden Beispiele zeigen, wie man die Liste der Stammfunktionen mit Hilfe dieser Tricks beliebig verlängern kann. In der Praxis ergibt sich jedoch eine andere Situation. Eine Funktion ist gegeben, und man muß sehen, ob die obigen Rechenregeln bei der Suche nach einer Stammfunktion weiterhelfen. Betrachten wir dazu ein Beispiel: Wir wollen eine Stammfunktion zur Funktion $h(x) = x\sqrt{1-x}$ bestimmen. Es ist klar, daß der Faktor x durch Differenzieren einfacher wird. Eine Stammfunktion des Faktors $\sqrt{1-x}$ läßt sich leicht mit Hilfe der Substitutionsregel bestimmen:

$$\begin{aligned} f(x) = -x^{1/2} &\Rightarrow F(x) = -\frac{2}{3}x^{3/2} \\ g(x) = 1-x &\Rightarrow f(g(x))g'(x) = -(1-x)^{1/2}(-1) = (1-x)^{1/2} \\ &\Rightarrow \text{SF} [\sqrt{1-x}] = -\frac{2}{3}(1-x)^{3/2} + c \end{aligned}$$

Um gleich die partielle Integration anwenden zu können, benötigen wir auch eine Stammfunktion von $-\frac{2}{3}(1-x)^{3/2}$

$$\begin{aligned} f(x) = \frac{2}{3}x^{3/2} &\Rightarrow F(x) = \frac{2}{3} \cdot \frac{2}{5}x^{5/2} \\ g(x) = 1-x &\Rightarrow f(g(x))g'(x) = \frac{2}{3}(1-x)^{3/2}(-1) = -\frac{2}{3}(1-x)^{3/2} \\ &\Rightarrow \text{SF} \left[-\frac{2}{3}(1-x)^{3/2} \right] = \frac{2}{3} \cdot \frac{2}{5}(1-x)^{5/2} + c \end{aligned}$$

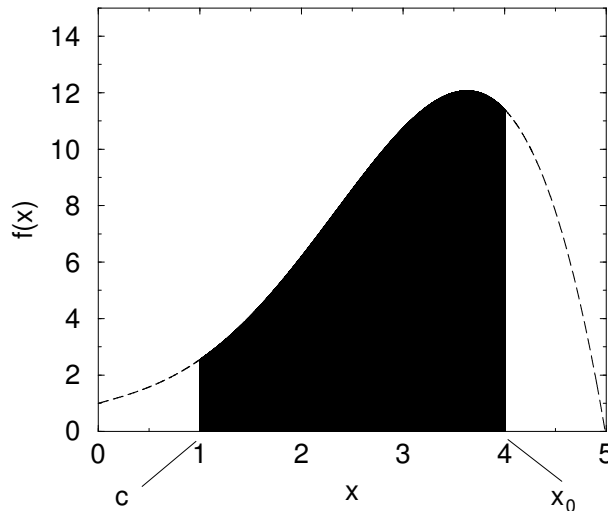
Damit können wir jetzt $\text{SF}[x\sqrt{1-x}]$ mit partieller Integration berechnen:

$$\begin{aligned} f'(x) &= \sqrt{1-x} \quad ; \quad g(x) = x \\ \Rightarrow \text{SF}[x\sqrt{1-x}] &= -\frac{2}{3}x(1-x)^{3/2} - \frac{2}{3} \frac{2}{5}(1-x)^{5/2} + c \end{aligned}$$

Von der Richtigkeit des Ergebnisses überzeugt man sich am besten immer durch Differenzieren.

Auch wenn dieses Kapitel mit “Integralrechnung” überschrieben ist, haben wir bis hierher nur von der Umkehrung der Differentiation und nicht von der Integration geredet. Wir wollen nun die Integralrechnung über den Zusammenhang zwischen der Stammfunktion und der *Flächenberechnung* einführen. Für eine stetige Funktion $f(x)$ wollen wir die dunkel dargestellte *Fläche* für festes c als Funktion von x_0 betrachten und bezeichnen sie mit $\tilde{F}(x_0)$:

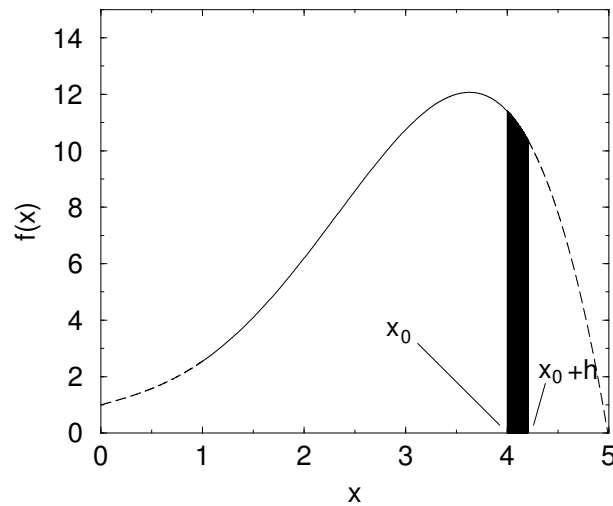
$\tilde{F}(x_0) =$ Fläche zwischen $f(x)$ und der x -Achse von c bis x_0



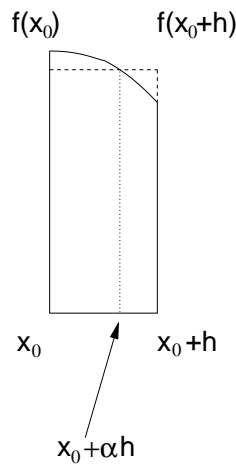
Wir untersuchen nun die Ableitung dieser Funktion $\tilde{F}(x_0)$

$$\tilde{F}'(x_0) = \lim_{h \rightarrow 0} \frac{\tilde{F}(x_0 + h) - \tilde{F}(x_0)}{h}.$$

Aufgrund der Definition von $\tilde{F}(x_0)$, ist $\tilde{F}(x_0 + h) - \tilde{F}(x_0)$ die Fläche zwischen $f(x)$ und der x -Achse von x_0 bis $x_0 + h$.



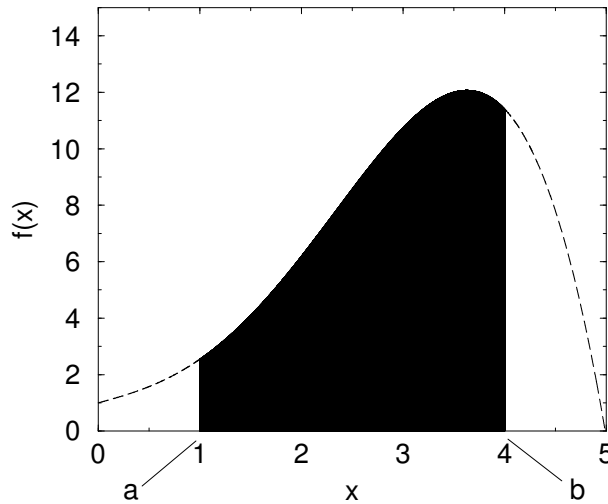
Vergrößert ergibt sich folgende Situation



Wie man in der Skizze erkennt, gibt es ein $\alpha \in \mathbb{R}$ mit $0 < \alpha \leq 1$, so daß die Fläche $\tilde{F}(x_0+h) - \tilde{F}(x_0)$ gleich der Fläche $hf(x_0 + \alpha h)$ eines Rechtecks ist. Dabei kennen wir α nicht, was aber im folgenden keine Rolle spielen wird. Wir haben damit

$$\begin{aligned}
 \tilde{F}'(x_0) &= \lim_{h \rightarrow 0} \frac{\tilde{F}(x_0+h) - \tilde{F}(x_0)}{h} \\
 &= \lim_{h \rightarrow 0} \frac{hf(x_0 + \alpha h)}{h} \\
 &= f(x_0).
 \end{aligned}$$

Es ergibt sich also $\tilde{F}'(x_0) = f(x_0)$, was bedeutet, daß \tilde{F} eine Stammfunktion von f ist. Wir schreiben damit statt \tilde{F} jetzt F . Als nächstes betrachten wir die Fläche zwischen $f(x)$ und der x -Achse von einem festen a bis zu einem festen b .



Die übliche Bezeichnung für diese Fläche ist

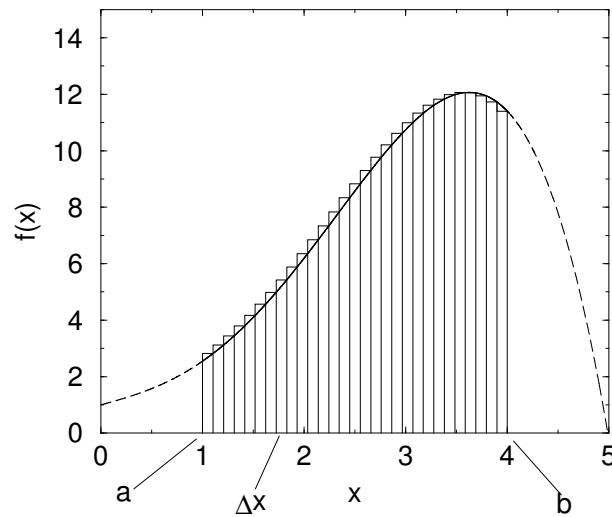
$$\int_a^b f(x) dx = \text{Fläche zwischen } f(x) \text{ und der } x\text{-Achse von } a \text{ bis } b,$$

so daß wir

$$\int_a^b f(x) dx = F(b) - F(a) \quad (3.4)$$

schreiben können. Die beliebige Konstante, die wir bei der Bestimmung einer Stammfunktion F immer hinzuaddieren mußten, fällt auf der rechten Seite dieser Gleichung dabei heraus. Man bezeichnet den Zusammenhang in Gl. (3.4) als *Hauptsatz der Differential- und Integralrechnung*. Er stellt einen Zusammenhang zwischen dem Problem der Bestimmung einer Stammfunktion und dem Problem der Flächenberechnung her. Man bezeichnet die linke Seite von Gl. (3.4) auch als *bestimmtes Integral*.

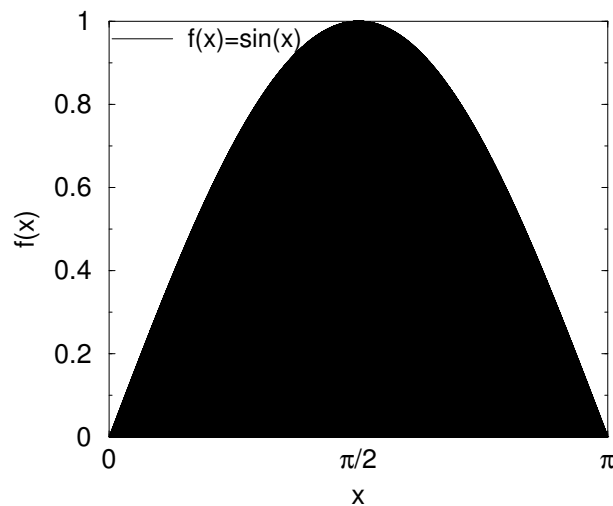
Bei den obigen Überlegungen haben wir uns keine Gedanken darüber gemacht, wie Flächen mit krummlinigen Begrenzungen (durch die Funktion $f(x)$) überhaupt sauber definiert sind. Als Grenzprozedur läßt sich das z.B. mit Hilfe der *Riemann-Summe* erreichen



$$\int_a^b f(x) dx = \lim_{N \rightarrow \infty} \sum_{i=1}^N f(a + i\Delta x) \Delta x \quad ; \quad \Delta x = \frac{b-a}{N}$$

Diese Formel gibt auch eine Möglichkeit ein bestimmtes Integral numerisch zu berechnen. Es gibt aber weit geschicktere numerische Berechnungsmöglichkeiten als die obige *Rechteckregel*, wie z.B. die *Trapezregel* oder noch besser die *Simpson-Regel*.

Als ein Beispiel der Flächenmessung (bestimmte Integration) betrachten wir $f(x) = \sin x$



$$\int_0^{\pi} \sin x dx = -\cos \pi - (-\cos 0) = -\cos x \Big|_0^{\pi} = 2$$

Es soll hier noch angemerkt werden, daß die Bezeichnung der Integrationsvariablen willkürlich ist, also

$$\int_a^b f(x) dx = \int_a^b f(u) du$$

gilt.

Jetzt wollen wir die Regeln zur Bestimmung der Stammfunktion so umformulieren, daß wir mit ihrer Hilfe bestimmte Integrale berechnen können. Betrachten wir zunächst die Substitutionsregel. Aus SF $[f(g(x))g'(x)] = F(g(x))$ folgt

$$\begin{aligned} \int_a^b f(g(x))g'(x) dx &= F(g(b)) - F(g(a)) \\ &= \int_{g(a)}^{g(b)} f(u) du, \end{aligned} \quad (3.5)$$

indem wir zweimal den Hauptsatz anwenden. Um den Anschluß an Lehrbücher zu finden, diskutieren wir diese Regel auch in der üblicheren Formulierung, die sich gut als Merkgel eignet. Dabei führt man die "Substitution"

$$\begin{aligned} u = g(x) &\Rightarrow \frac{du}{dx} = g'(x) \\ \Rightarrow f(g(x))g'(x) dx &= f(u) \frac{du}{dx} dx = f(u) du \end{aligned}$$

aus. Das letzte Gleichheitszeichen steht dabei in Anführungszeichen, damit man nicht vergißt, daß auch die Integrationsgrenzen verändert werden müssen (siehe Gl. (3.5)). In unserer ursprünglichen Herleitung mußten wir jedoch keine *infinitesimalen* Größen kürzen, was ein klarer Vorteil ist. Der Umgang mit diesen Größen kann formalisiert werden, wenn man den Begriff der Differentialform einführt, wozu wir jedoch hier keine Zeit haben. Dieser Begriff wird speziell in der Thermodynamik gebraucht.

Die Produktregel übersetzt sich wie folgt auf bestimmte Integrale: Wegen

$$\int_a^b \frac{dG(x)}{dx} dx = G(b) - G(a) = G(x)|_a^b$$

folgt durch Integration von $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$

$$\int_a^b f'(x)g(x) dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x) dx, \quad (3.6)$$

was die Bezeichnung partielle Integration erklärt. Beispiele werden sie in den Übungen rechnen.

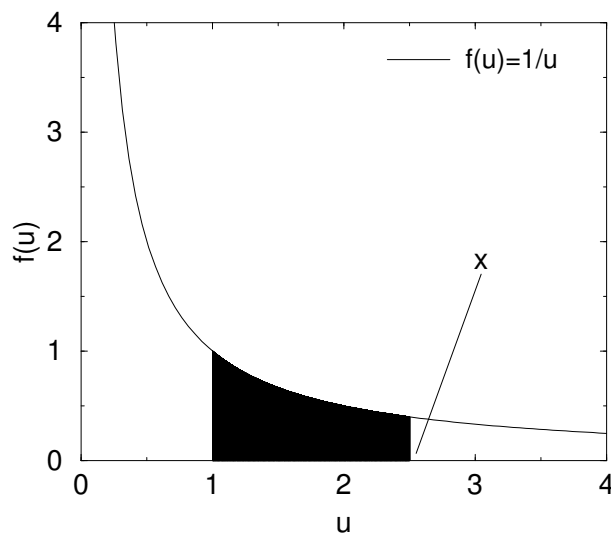
Kapitel 4

Logarithmus- und Exponentialfunktion

Wir wollen als nächstes eine über ein Integral definierte Funktion

$$L(x) = \int_1^x \frac{1}{u} du \quad , \quad x > 0 \quad (4.1)$$

betrachten.



Nach der Definition gilt $L(1) = 0$ und $L(x) > 0$ für $x > 1$. Für den Fall $0 < x < 1$ ist das Integral in Gl. (4.1) genauso wie in Gl. (3.4) definiert, was, da hier $f(u) > 0$ für alle $u > 0$ gilt, dazu führt, daß die Fläche zwischen $f(u)$ und der u -Achse von 1 bis x , also $L(x)$, *negativ* ist. Es gilt also $L(x) < 0$ für $0 < x < 1$. $L(x)$ ist

weiterhin *monoton wachsend*. Für $a > 0$ betrachten wir

$$\begin{aligned} L(ax) &= \int_1^{ax} \frac{1}{u} du \\ &= \int_1^a \frac{1}{u} du + \int_a^{ax} \frac{1}{u} du \\ &= L(a) + \int_a^{ax} \frac{1}{u} du. \end{aligned}$$

Im verbleibenden Integral verwenden wir die Substitution

$$\begin{aligned} \int_a^{ax} \frac{1}{u} du &= \int_a^{ax} \frac{a}{u} \frac{1}{a} du \\ &= \int_a^{ax} f(g(u))g'(u) du \end{aligned}$$

mit $f(v) = 1/v$ und $g(v) = v/a$. Nach Gl. (3.5) gilt damit

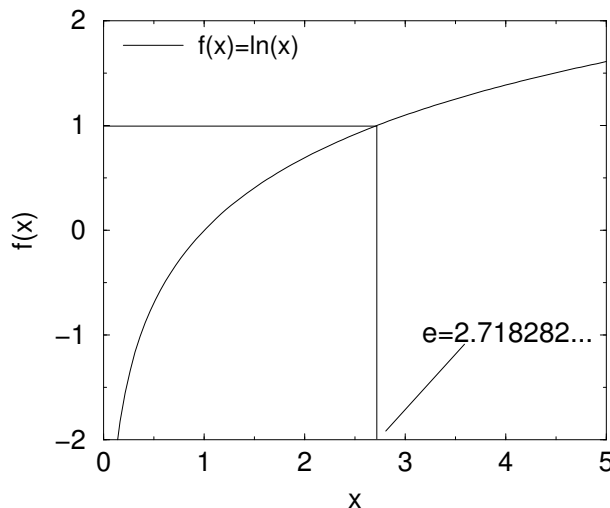
$$\begin{aligned} L(ax) &= L(a) + \int_1^x \frac{1}{t} dt \\ &= L(a) + L(x), \end{aligned} \tag{4.2}$$

also z.B. für $n \in \mathbb{N}$, $L(x^n) = nL(x)$. Gl. (4.2) ist die bekannte *Funktionalrelation* für die *Logarithmusfunktionen*. Die durch Gl. (4.1) definierte Funktion wird als *natürlicher Logarithmus* bezeichnet, und wir schreiben ab jetzt $\ln(x)$ statt $L(x)$

$$\ln(x) = \int_1^x \frac{1}{t} dt, \quad x > 0.$$

Man definiert eine wichtige Zahl e - die *Eulersche Zahl* - über die Gleichung

$$\ln(e) = 1.$$

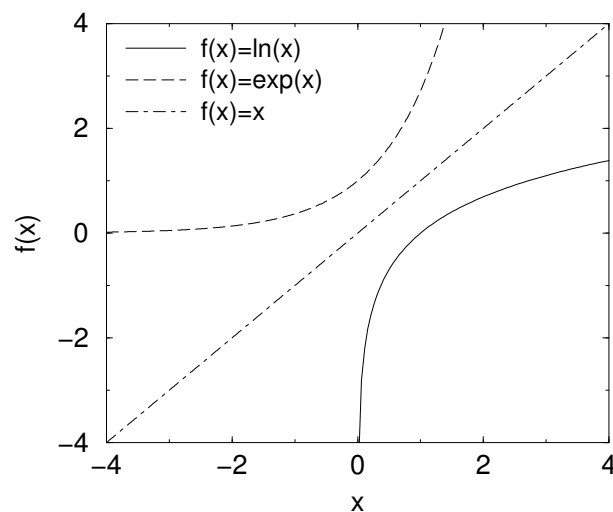


Aus der Skizze sieht man, daß e in der Nähe von 2.7 liegt. Genauer gilt $e = 2.718282\dots$. Nach dem Hauptsatz folgt $(\ln(x))' = 1/x$. Für $x \rightarrow \infty$ strebt auch $\ln(x)$ gegen ∞ , was man z.B. wie folgt sehen kann: Sei $n \in \mathbb{N}$; dann gilt

$$\ln(e^n x) = \ln(e^n) + \ln(x) = n \ln(e) + \ln(x) = n + \ln(x).$$

Für $n \rightarrow \infty$ streben das Argument $e^n x$ und der Funktionswert gegen ∞ .

Die Umkehrfunktion $\ln^{-1}(x)$ des natürlichen Logarithmus - die Logarithmusfunktion ist injektiv - bezeichnet man als *Exponentialfunktion* und kürzt sie $\exp(x)$ ab. Sie hat den Definitionsbereich $D_{\exp} = \mathbb{R}$.



Es gilt also

$$\begin{aligned} \exp(\ln(x)) &= x \quad \text{für } x \in \{x \in \mathbb{R} \mid x > 0\} = \mathbb{D}_{\ln} = \mathbb{W}_{\exp} \\ \ln(\exp(x)) &= x \quad \text{für } x \in \mathbb{R} = \mathbb{D}_{\exp} = \mathbb{W}_{\ln}. \end{aligned}$$

Aus der Funktionalgleichung des Logarithmus wollen wir jetzt eine entsprechende Gleichung für $\exp(x)$ herleiten. Für $a, b > 0$ folgt mit $a = \exp(\ln(a))$ und $b = \exp(\ln(b))$

$$\exp(\ln(ab)) = ab = \exp(\ln(a)) \exp(\ln(b)).$$

Die linke Seite dieser Gleichung kann nach Gl. (4.2) zu

$$\exp(\ln(ab)) = \exp(\ln(a) + \ln(b))$$

umgeformt werden. Damit gilt dann

$$\exp(\ln(a) + \ln(b)) = \exp(\ln(a)) \exp(\ln(b))$$

oder, da $a, b > 0$ beliebig waren,

$$\exp(x + y) = \exp(x) \exp(y). \quad (4.3)$$

Diese Funktionalrelation sollte ihnen von den *Potenzgesetzen* bekannt vorkommen: $x^n x^m = x^{n+m}$. In dieser Vorlesung haben wir sie bereits mehrfach benutzt. Aus Gl. (4.3) folgt für $x = y = 0$, daß $\exp(0) = 1$ gilt (was natürlich auch aus der Definition von $\exp(x)$ als Umkehrfunktion von $\ln(x)$ und $\ln(1) = 0$ folgt), und für $y = -x$ ergibt sich damit

$$\exp(-x) = \frac{1}{\exp(x)}. \quad (4.4)$$

Eine wiederholte Anwendung von Gl. (4.3) liefert

$$\exp(x_1 + x_2 + \dots + x_n) = \exp(x_1) \exp(x_2) \dots \exp(x_n). \quad (4.5)$$

Aus $e = \exp(\ln(e))$ und $\ln(e) = 1$ folgt $\exp(1) = e$, sowie aus Gl. (4.5) mit $x_i = 1$ für alle $i = 1, 2, \dots, n$

$$\exp(n) = e^n.$$

Mit Gl. (4.4) gilt dann auch

$$\exp(-n) = \frac{1}{e^n} = e^{-n},$$

d.h. für alle $n \in \mathbb{Z}$ ist die Exponentialfunktion die Potenzfunktion zur Basis e . Diese Aussage gilt auch für rationale $\alpha \in \mathbb{Q}$ mit $\alpha = n/m$

$$\begin{aligned} \left(\exp\left(\frac{n}{m}\right)\right)^m &= \underbrace{\exp\left(\frac{n}{m}\right) \cdot \exp\left(\frac{n}{m}\right) \cdot \dots \cdot \exp\left(\frac{n}{m}\right)}_{m\text{-mal}} \\ &= \exp\left(\frac{n}{m} + \frac{n}{m} + \dots + \frac{n}{m}\right) \\ &= \exp(n) = e^n, \end{aligned}$$

wobei wir Gl. (4.5) verwendet haben. Ziehen wir jetzt die m -te Wurzel, so folgt

$$\exp\left(\frac{n}{m}\right) = e^{n/m}.$$

Da man jeder Zahl aus \mathbb{R} mit einem $\alpha \in \mathbb{Q}$ beliebig nahe kommen kann (siehe *Diff. I* Vorlesung), gilt für alle $x \in \mathbb{R}$

$$\exp(x) = e^x.$$

Aus der Schule sind sie es wahrscheinlich gewohnt, daß beide Ausdrücke (Exponentialfunktion und Potenz zur Basis e) von Anfang an synonym verwendet werden, was wir im folgenden ebenfalls tun werden.

Die Ableitung der Exponentialfunktion können wir mit Hilfe von Gl. (2.5), also $f^{-1}'(f(x)) = \frac{1}{f'(x)}$, für $f(x) = \ln(x)$, d.h. $f'(x) = 1/x$ bestimmen:

$$\exp'(\ln x) = x \quad \Rightarrow \quad \exp'(z) = \exp(z)$$

Damit erfüllt die Exponentialfunktion $y(x) = \exp(x)$ die in der Physik sehr wichtige Differentialgleichung

$$y'(x) = y(x), \tag{4.6}$$

die man als *Wachstumsgleichung* bezeichnen kann.

Kapitel 5

Potenzreihen und Taylorentwicklung

Nach einem kleinen Einschub, werden wir auf Gl. (4.6) zurückkommen und mit ihrer Hilfe die *Reihendarstellung* der Exponentialfunktion herleiten. Zunächst soll aber erklärt werden, was eine *Potenzreihe* ist. Dazu betrachten wir die Funktionen ($N \in \mathbb{N}_0, x \in \mathbb{R}$)

$$S_N(x) = 1 + x + x^2 + x^3 + \dots + x^N = \sum_{n=0}^N x^n.$$

Indem wir diese Gleichung mit x multiplizieren und mit $+1-1$ erweitern, erhalten wir einen geschlossenen Ausdruck für $S_N(x)$

$$\begin{aligned} xS_N(x) &= 1 + x + x^2 + \dots + x^N + x^{N+1} - 1 \\ &= S_N(x) + x^{N+1} - 1 \\ \Rightarrow (1-x)S_N(x) &= 1 - x^{N+1} \\ \Rightarrow S_N(x) &= \frac{1 - x^{N+1}}{1 - x}. \end{aligned}$$

Für $|x| < 1$ geht der Term x^{N+1} im Zähler für $N \rightarrow \infty$ gegen Null, d.h.

$$\lim_{N \rightarrow \infty} S_N(x) = \frac{1}{1-x} \quad \text{für } |x| < 1.$$

Man kann die Funktion $1/(1-x)$ für $|x| < 1$ also auch als *unendliche Reihe* (auch Potenzreihe genannt)

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad \text{für } |x| < 1$$

schreiben. Eine allgemeine Potenzreihe hat die Form

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

mit $a_n \in \mathbb{R}$. Für die gerade diskutierte spezielle so genannte *geometrische Reihe* gilt $a_n = 1$ für alle $n \in \mathbb{N}_0$. Die Frage für welche x eine unendliche Reihe überhaupt auf einen endlichen Wert führt, d.h. *konvergiert*, werden wir hier nur am Rande betrachten. Sehr genau werden sie diese Frage in der *Diff. I* Vorlesung untersuchen.

Fast alle (in der Physik) wichtigen Funktionen lassen sich durch Potenzreihen darstellen, so z.B. auch die Exponentialfunktion. Um für diesen Fall die *Entwicklungskoeffizienten* a_n zu bestimmen betrachten wir die Differentialgleichung (4.6) $y(x) = y'(x)$, von der wir bereits wissen, daß sie von der Exponentialfunktion gelöst wird. Wir machen den *Ansatz*¹

$$y(x) = \sum_{n=0}^{\infty} a_n x^n \quad \Rightarrow \quad y'(x) = \sum_{n=1}^{\infty} a_n n x^{n-1} .$$

Dabei haben wir die unendliche Reihe summandenweise differenziert, was nur dann erlaubt ist, wenn die Reihe bestimmte Konvergenzeigenschaften hat (siehe die *Diff. I* Vorlesung). Umbenennen der Summationsvariablen $n = \tilde{n} + 1$ in $y'(x)$ liefert

$$y'(x) = \sum_{\tilde{n}=0}^{\infty} a_{\tilde{n}+1} (\tilde{n} + 1) x^{\tilde{n}} = \sum_{n=0}^{\infty} a_{n+1} (n + 1) x^n .$$

Bildet man jetzt $y'(x) - y(x)$, was nach der Differentialgleichung gleich Null ist, so folgt²

$$\sum_{n=0}^{\infty} ((n + 1)a_{n+1} - a_n) x^n = 0$$

eine Gleichung, die sicherlich dann erfüllt ist, wenn

$$(n + 1)a_{n+1} - a_n = 0 \quad \text{für alle } n \in \mathbb{N}_0 \quad (5.1)$$

gilt. Daß dieses Kriterium nicht nur *hinreichend*, sondern auch *notwendig* ist, werden sie in der *Diff. I* Vorlesung beweisen (Identitätssatz für Potenzreihen).

¹Das Lösen von Differentialgleichungen durch einen Potenzreihenansatz ist ein häufig verwendetes Verfahren.

²Wir haben hier die beiden Summen unter ein Summenzeichen geschrieben. Auch diese Operation ist nur dann zulässig, wenn die unendliche Reihe bestimmte Konvergenzeigenschaften erfüllt, die sie in der *Diff. I* Vorlesung kennenlernen werden.

Gl. (5.1) kann nun wie folgt umgeformt werden

$$\begin{aligned}
 a_{n+1} &= \frac{1}{n+1} a_n \\
 &= \frac{1}{n+1} \frac{1}{n} a_{n-1} \\
 &= \frac{1}{n+1} \frac{1}{n} \frac{1}{n-1} a_{n-2} \\
 &= \dots \\
 &= \frac{1}{(n+1)!} a_0.
 \end{aligned}$$

Die Lösung von $y'(x) = y(x)$ lautet damit

$$y(x) = a_0 \sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

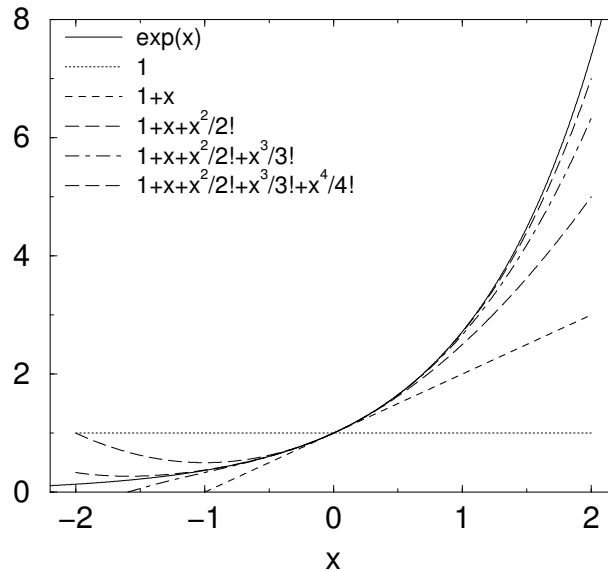
mit der zunächst noch freien Konstanten $a_0 \in \mathbb{R}$, woraus $y(0) = a_0$ folgt.³ Von der Exponentialfunktion wissen wir jedoch bereits, daß $\exp(0) = 1$ gilt. Damit hat dann also die Exponentialfunktion die Potenzreihenentwicklung

$$\exp(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \quad (5.2)$$

Diese Reihe *konvergiert* für alle $x \in \mathbb{R}$ (im Gegensatz zur geometrischen Reihe, die nur für $|x| < 1$ konvergiert), was sie in *Diff. I* beweisen werden.

Wenn man eine Reihendarstellung für eine Funktion kennt, dann ist es interessant zu untersuchen, ob man die “komplizierte” Funktion (also z.B. die Exponentialfunktion), nicht bereits durch ein paar wenige “einfache” Summanden der Potenzreihe gut approximieren (näheren) kann. Die folgende Abbildung zeigt dazu ein Beispiel.

³Dies ergibt sich da $0^0 = 1$ gilt.



Man erkennt, daß für $-1 \leq x \leq 1$ bereits die ersten fünf Summanden eine sehr gute Näherung für die Exponentialfunktion liefern. In dem betrachteten Beispiel ist es tatsächlich so daß die *Polynome*

$$g_N(x) = \sum_{n=0}^N \frac{x^n}{n!}$$

für wachsendes N eine systematische Verbesserung der Approximation liefern - speziell in der Nähe von $x = 0$.

Schaut man nur lange genug auf diese Abbildung, so fällt einem auf, daß die Übereinstimmung der Funktionen in der Nähe von Null etwas mit der Gleichheit von Ableitungen von $\exp(x)$ und den $g_N(x)$ bei $x = 0$ zu tun hat. Für die Exponentialfunktion $f(x) = \exp(x)$ gilt

$$f^{(k)}(0) = 1 \quad \text{da } f^{(k)}(x) = \exp(x).$$

Für die Polynome $g_N(x)$ ergibt sich

$$g_N^{(k)}(0) = \begin{cases} 1 & \text{für } k \leq N \\ 0 & \text{für } k > N, \end{cases}$$

so, daß für $g_N(x)$ die ersten N Ableitungen bei $x = 0$ mit denen der Exponentialfunktion übereinstimmen.

Diese Beobachtung können wir nun zu einem Konstruktionsprinzip zur Bestimmung einer polynomialen Näherung von Funktionen machen. Bei der Diskussion der Ableitung als lineare Approximierbarkeit haben wir es in gewissem Sinne bereits kennengelernt: Die Gleichung für die Tangente an eine Funktion ist das Polynom ersten Grades $g_1(x)$ und liefert eine Näherung für die Funktion.

Man nennt ein Polynom vom Grad n , wenn der Summand mit der höchsten auftretenden Potenz von x derjenige proportional zu x^n ist. Betrachten wir also die allgemeine Funktion $f(x)$. Wir wollen nun die Koeffizienten a_n des Polynoms

$$g_N(x) = a_0 + a_1x + a_2x^2 + \dots + a_Nx^N$$

so bestimmen, daß die ersten N Ableitungen von $g_N(x)$ mit denen der Funktion $f(x)$ bei $x = 0$ übereinstimmen. Betrachten wir ein $m \leq N$, so gilt

$$g_N^{(m)}(0) = a_m m! \quad .$$

Damit ist klar, daß wir, um die Forderung zu erfüllen,

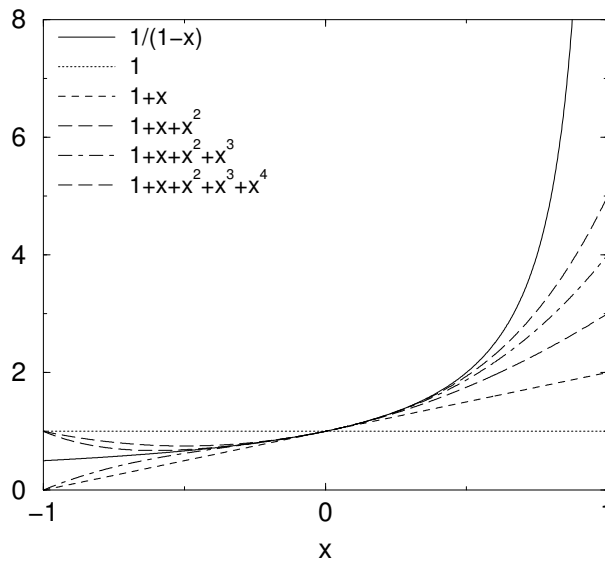
$$a_n = \frac{1}{n!} f^{(n)}(0) \quad (5.3)$$

wählen müssen. Es ergibt sich somit

$$g_N(x) = f(0) + \frac{1}{1!} f'(0)x + \frac{1}{2!} f''(0)x^2 + \dots + \frac{1}{N!} f^{(N)}(0)x^N. \quad (5.4)$$

Man nennt diese Näherung von $f(x)$ durch $g_N(x)$ die *Taylorentwicklung* um den Punkt $x = 0$. In der Übungsaufgabe 16 haben sie sie bereits kennengelernt und $g_3(x)$ für zwei Beispielfunktionen berechnet.

- $f(x) = \frac{1}{1-x}$: $g_3(x) = 1 + x + x^2 + x^3$



Die sich ergebenden Terme sind aber gerade die ersten Summanden der geometrischen Reihe, so daß wir wissen, daß die Taylorentwicklung für immer größere N und $|x| < 1$ eine immer bessere Näherung liefert. Für $N \rightarrow \infty$ spricht man dann auch von der *Taylorreihe*. Für dieses Beispiel konvergiert die über die Ableitungen von f definierte Taylorreihe gegen die Funktion f , was jedoch nicht immer der Fall sein muß (wie sie in späteren Vorlesungen genauer untersuchen werden).

$$\bullet f(x) = \sqrt{1+x} \quad : \quad g_3(x) = 1 + \frac{1}{2}x - \frac{1}{2} \frac{1}{4}x^2 + \frac{1}{2} \frac{1}{4} \frac{3}{6}x^3$$

Als weiteres Beispiel wollen wir hier die Funktion $f(x) = \ln(x+1)$ um $x=0$ Taylor-entwickeln.⁴ Nach Gl. (5.4) benötigen wir die Ableitungen von $f(x)$ bei $x=0$:

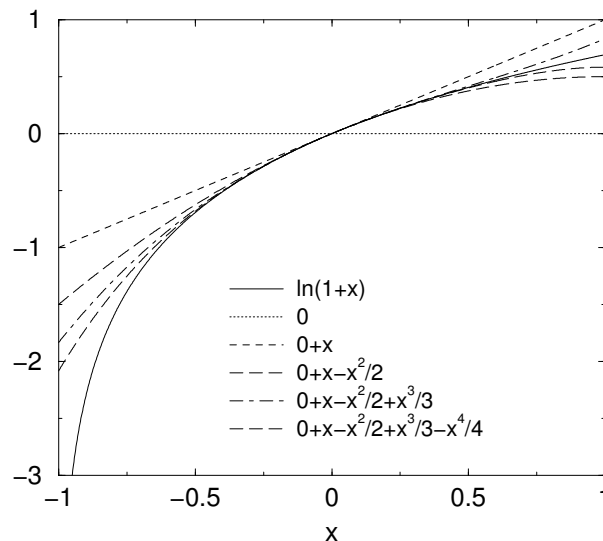
$$\begin{aligned} f(x) &= \ln(x+1) \quad \Rightarrow \quad f(0) = \ln(1) = 0 \\ f'(x) &= \frac{1}{x+1} \quad \Rightarrow \quad f'(0) = 1 = (-1)^0 0! \\ f''(x) &= (-1) \frac{1}{(x+1)^2} \quad \Rightarrow \quad f''(0) = -1 = (-1)^1 1! \\ f'''(x) &= (-1)(-2) \frac{1}{(x+1)^3} \quad \Rightarrow \quad f'''(0) = 2 = (-1)^2 2! \\ f^{(4)}(x) &= (-1)(-2)(-3) \frac{1}{(x+1)^4} \quad \Rightarrow \quad f^{(4)}(0) = -6 = (-1)^3 3! \\ &= \dots \\ f^{(N)}(x) &= (-1)^{N-1} (N-1)! \frac{1}{(x+1)^N} \quad \Rightarrow \quad f^{(N)}(0) = (-1)^{N-1} (N-1)! \end{aligned}$$

Die Koeffizienten der Taylorentwicklung ergeben sich dann nach Gl. (5.3) zu

$$a_n = \frac{1}{n!} f^{(n)}(0) = \frac{1}{n!} (-1)^{n-1} (n-1)! = (-1)^{n-1} \frac{1}{n},$$

also gilt

$$g_N(x) = \sum_{n=1}^N (-1)^{n-1} \frac{1}{n} x^n.$$



⁴Da $\ln(x)$ bei $x=0$ divergiert - gegen $-\infty$ geht - untersuchen wir die Funktion $\ln(x+1)$.

Bisher haben wir die Funktion f immer um den Punkt $x = 0$ entwickelt, d.h. die ersten N Ableitungen von g_N und f stimmen im Punkt $x = 0$ überein. Möchte man jedoch die Funktion f in der Umgebung eines beliebigen Punktes x_0 näherungsweise durch ein Polynom g_N beschreiben, so bietet es sich an die a_n so zu bestimmen, daß die Ableitungen bei x_0 übereinstimmen

$$g_N(x) = f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \dots + \frac{1}{N!}f^{(N)}(x_0)(x - x_0)^N.$$

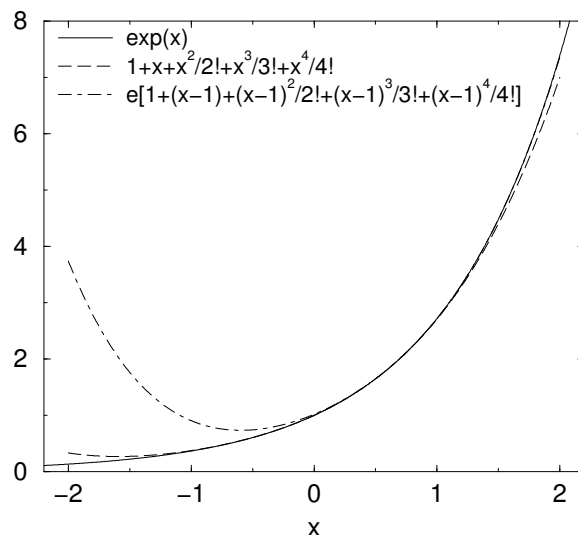
Als Beispiel betrachten wir wieder die Exponentialfunktion $f(x) = \exp(x)$ jetzt aber bei $x_0 = 1$. Es gilt

$$f^{(k)}(1) = e \quad \text{da } f^{(k)}(x) = \exp(x),$$

so daß sich

$$g_N(x) = e \left(1 + \frac{1}{1!}(x - 1) + \frac{1}{2!}(x - 1)^2 + \dots + \frac{1}{N!}(x - 1)^N \right)$$

ergibt. Die Abbildung macht deutlich, daß $g_4(x)$ mit den Koeffizienten der Entwicklung um den Punkt $x = 1$ in der Nähe von $x = 1$ eine wesentlich bessere Näherung an die Exponentialfunktion liefert als das $g_4(x)$ mit den aus der Entwicklung um $x = 0$ bestimmten a_n .



Wie in diesem Beispiel lassen sich auch andere Funktionen um Punkte x_0 Taylorentwickeln.

Kapitel 6

Hyperbolische und trigonometrische Funktionen als Potenzreihen: die imaginäre Einheit i

Die so genannten *Hyperbelfunktionen* Sinushyperbolikus $\sinh x$ und Kosinushyperbolikus $\cosh x$ sind über die Exponentialfunktion definiert

$$\sinh x = \frac{e^x - e^{-x}}{2} \quad ; \quad \cosh x = \frac{e^x + e^{-x}}{2}. \quad (6.1)$$

Drücken wir die Exponentialfunktion nun durch ihre Potenzreihendarstellung Gl. (5.2) aus, so ergibt sich

$$\begin{aligned} \cosh x &= \frac{e^x + e^{-x}}{2} \\ &= \frac{1}{2} \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} + \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} \right) \\ &= \frac{1}{2} \sum_{n=0}^{\infty} (1 + (-1)^n) \frac{x^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, \end{aligned} \quad (6.2)$$

da

$$\frac{1 + (-1)^n}{2} = \begin{cases} 1 & \text{für } n \text{ gerade} \\ 0 & \text{für } n \text{ ungerade.} \end{cases}$$

Analog folgt

$$\begin{aligned}
 \sinh x &= \frac{e^x - e^{-x}}{2} \\
 &= \frac{1}{2} \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} - \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} \right) \\
 &= \frac{1}{2} \sum_{n=0}^{\infty} (1 - (-1)^n) \frac{x^n}{n!} \\
 &= \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}. \tag{6.3}
 \end{aligned}$$

Die Ableitungen von $\sinh x$ und $\cosh x$ bestimmt man mit Hilfe der Definitionsgleichung (6.1) oder durch gliedweises Differenzieren der Reihendarstellungen:

$$\sinh' x = \left(\frac{e^x - e^{-x}}{2} \right)' = \frac{e^x + e^{-x}}{2} = \cosh x$$

oder

$$\sinh' x = \left(\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \right)' = \sum_{n=0}^{\infty} \frac{(2n+1)}{(2n+1)} \frac{x^{2n}}{(2n)!} = \cosh x.$$

Analog folgt

$$\cosh' x = \sinh x.$$

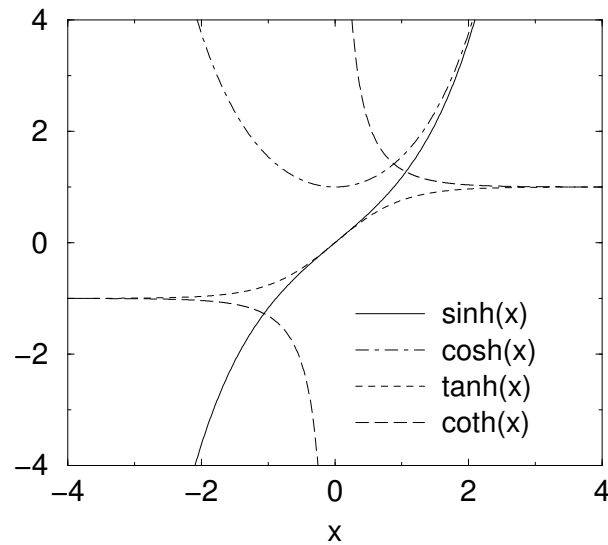
Wie bei den trigonometrischen Funktionen definiert man auch noch den Tangenshyperbolicus

$$\tanh x = \frac{\sinh x}{\cosh x}$$

und den Kotangenshyperbolicus

$$\coth x = \frac{\cosh x}{\sinh x}.$$

In der folgenden Abbildung sind die vier hyperbolischen Funktionen dargestellt.



Um jetzt auch die trigonometrischen Funktionen als Potenzreihen darzustellen, betrachten wir zunächst einmal die Differentialgleichung zweiter Ordnung, die $y(x) = e^{ax} = \exp(ax)$ erfüllt. Es gilt

$$(e^{ax})' = ae^{ax} \quad \Rightarrow \quad (e^{ax})'' = a^2 e^{ax},$$

so daß

$$y''(x) - a^2 y(x) = 0$$

folgt. Im Anschluß an Gl. (2.6) (siehe auch Gl. (2.7)) hatten wir gezeigt, daß $y(x) = b \sin(ax) + c \cos(ax)$, mit $a, b, c \in \mathbb{R}$ die Differentialgleichung

$$y''(x) + a^2 y(x) = 0$$

erfüllt. Erweitern wir jetzt die reellen Zahlen um die *imaginäre Einheit* i mit der Eigenschaft

$$i^2 = -1$$

so erfüllt

$$e^{iax} = \sum_{n=0}^{\infty} \frac{(iax)^n}{n!} \quad (6.4)$$

die Differentialgleichung

$$y''(x) - (ia)^2 y(x) = y''(x) - i^2 a^2 y(x) = y''(x) + a^2 y(x) = 0$$

deren Lösung wir bisher als $y(x) = b \sin(ax) + c \cos(ax)$ geschrieben hatten. Dabei haben wir i bezüglich der Rechenregeln wie eine gewöhnliche Zahl behandelt, also

z.B. $(ia)^2 = i^2 a^2$, dann aber verwendet, daß $i^2 = -1$ gilt. Betrachten wir nun den Spezialfall $a = 1$. Für die Lösung der Differentialgleichung $y''(x) + y(x) = 0$ in der bisherigen Form können wir dann $y(x) = c \cos x + b \sin x$ schreiben, wobei die Randbedingungen (bzw. Anfangsbedingungen) durch $y(0) = c$ und $y'(0) = b$ gegeben sind. Für die Lösung in der neuen Form e^{ix} ergeben sich die Randbedingungen

$$e^{ix}|_{x=0} = e^{i0} = e^0 = 1 \quad \text{und} \quad (e^{ix})'|_{x=0} = i e^{ix}|_{x=0} = i.$$

Sollen beide Lösungen die gleiche Funktion darstellen, so müssen nicht nur die Differentialgleichungen identisch sein, sondern auch die Randbedingungen. Setzen wir also $c = 1$ und $b = i$, so folgt

$$e^{ix} = \cos x + i \sin x. \quad (6.5)$$

Eine Größe der Form $z = a + ib$ mit $a, b \in \mathbb{R}$ und der imaginären Einheit i nennt man *komplexe Zahl*. Man bezeichnet die Menge der komplexen Zahlen mit \mathbb{C} .

Bevor wir tiefer in dieses Konzept eindringen, wollen wir aus Gl. (6.5) noch die Potenzreihendarstellungen von $\cos x$ und $\sin x$ extrahieren. Dazu betrachten wir Potenzen von i :

$$i^0 = 1; \quad i^1 = i; \quad i^2 = -1; \quad i^3 = i^2 i = (-1)i = -i; \quad i^4 = i^2 i^2 = (-1)(-1) = 1$$

Für $n \in \mathbb{N}_0$ haben wir somit

$$i^{2n} = (i^2)^n = (-1)^n \quad ; \quad i^{2n+1} = i^{2n} i = (-1)^n i.$$

Diese Eigenschaften nutzen wir nun aus, um Gl. (6.4) für den Spezialfall $a = 1$ umzuschreiben

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(ix)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(ix)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}. \end{aligned}$$

Ein Vergleich mit Gl. (6.5) liefert dann

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad (6.6)$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}. \quad (6.7)$$

Die Namensgebung der hyperbolischen Funktionen wird nun durch einen Vergleich der Gln. (6.6) und (6.7) mit den Gln. (6.2) und (6.3) klar. Daß es sich bei den durch die Gln. (6.6) und (6.7) gegebenen Funktionen um die in der Schule mit Hilfe rechtwinkliger Dreiecke definierten Funktionen handelt, ist jedoch nicht vollständig offensichtlich. Was aber klar ist:

- Da nur gerade bzw. ungerade Potenzen von x in $\cos x$ bzw. $\sin x$ auftreten, ist $\cos x$ eine *gerade* Funktion und $\sin x$ eine *ungerade*.
- $\sin 0 = 0$ und $\cos 0 = 1$.
- Gliedweise Differentiation liefert

$$\begin{aligned}\sin' x &= \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \right)' \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)}{(2n+1)} \frac{x^{2n}}{(2n)!} \\ &= \cos x\end{aligned}$$

und analog $\cos' x = -\sin x$.

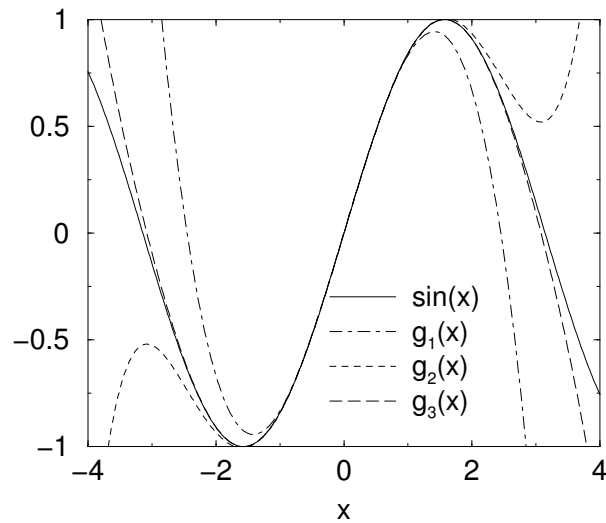
Alle diese Eigenschaften gelten analog für die ihnen bekannten trigonometrischen Funktionen, und sie machen daher plausibel, daß es sich tatsächlich um diese Funktionen handelt.

Für das Beispiel der Sinusfunktion wollen wir jetzt noch kurz untersuchen, wie gut sich $\sin x$ durch einige wenige Terme der Potenzreihe¹

$$g_N(x) = \sum_{n=0}^N (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

approximieren läßt.

¹Die Potenzreihen des \sin und \cos entsprechen, wie in den Beispielen im letzten Kapitel, wieder der Taylorreihe für eine Entwicklung der Funktion um den Punkt $x = 0$.



Aus der Abbildung wird klar, daß für hinreichend kleine $|x|$ bereits die ersten paar Terme eine sehr gute Näherung liefern.² Die Entwicklung der trigonometrischen Funktionen in niedrigen Ordnungen spielt in der Physik eine große Rolle und wird ihnen in der *Physik I* Vorlesung spätestens bei der Diskussion der Schwingungen (mit kleiner Amplitude) eines Fadenpendels begegnen.

²Die Abbildung verdeutlicht auch noch einmal, daß es sich bei der durch die Potenzreihe definierten Funktion tatsächlich um die geometrisch definierte Sinusfunktion handelt.

Kapitel 7

Komplexe Zahlen

Anlaß zur Erweiterung der Menge der reellen Zahlen \mathbb{R} war die Beobachtung, daß viele quadratische, kubische und quartische Gleichungen, z.B. die einfache quadratische Gleichung $x^2 + 2 = 0$, keine reellen Lösungen besitzen. Veröffentlicht wurde eine solche Erweiterung erstmalig 1545 durch Gerolamo Cardano. Seit 1539 wurde die den komplexen Zahlen zugrundeliegende Idee zur Lösung kubischer Gleichungen jedoch in speziellen Fällen bereits von einem gewissen Tartaglia “kommerziell” genutzt - der seinen “Trick” zur Lösung jedoch geheimgehalten hatte. Nachdem er Cardano einen “Geheimhaltungsschwur” hatte schwören lassen, verriet Tartaglia selbigem seine Methode. Dieser erweiterte sie und veröffentlichte seine Ergebnisse dennoch. Weiteres dazu finden sie auf folgender Internetseite

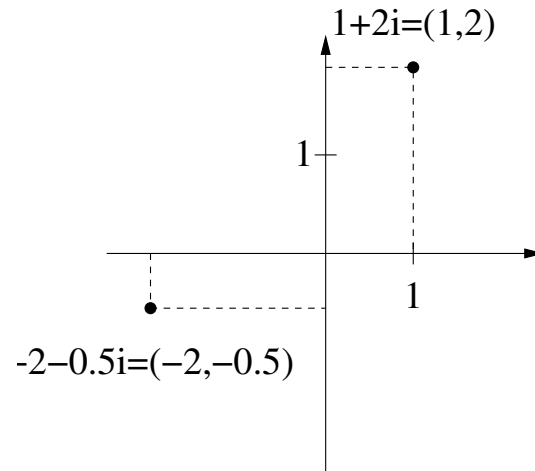
<http://www-history.mcs.st-and.ac.uk/~history/Mathematicians/Cardan.html>

und den zugehörigen Links.

Die fundamentale Beobachtung ist, daß man mit der scheinbar sinnlosen Lösung der obigen quadratischen Gleichung $x_{1/2} = \pm\sqrt{-2}$ so weiter rechnen kann, als wenn es sich um eine reelle Zahl handeln würde. Zum Beispiel gilt $x_1 + x_2 = 0$ und $x_1x_2 = 2$. Bereits im letzten Kapitel hatten wir die übliche Abkürzung $i = \sqrt{-1}$ eingeführt. Für zwei gegebene reelle Zahlen $a, b \in \mathbb{R}$ nennt man das zunächst formale Konstrukt $z = a + ib$ eine *komplexe Zahl*. Dem reellen Paar (a, b) wird eine komplexe Zahl zugeordnet

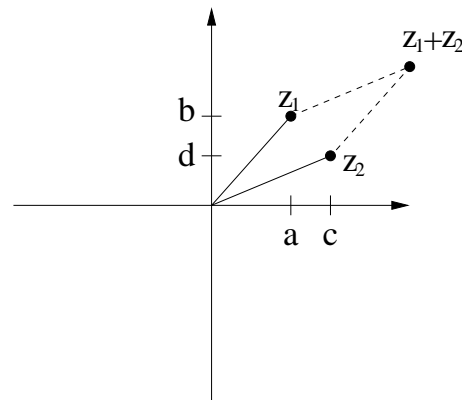
$$(a, b) \rightarrow a + ib.$$

Es gilt also z.B. $(0, 1) \rightarrow i$ oder $(2, +1/2) \rightarrow 2 + i 0.5$. Wie bereits erwähnt bezeichnet man die Menge der komplexen Zahlen, also die Paare (a, b) mit \mathbb{C} . Oft werden die komplexen Zahlen in der sogenannten *komplexen (oder Gaußschen) Zahlenebene* graphisch dargestellt. Die erste Komponente der komplexen Zahl nennt man Realteil ($a = \operatorname{Re}[z]$) und die zweite Imaginärteil ($b = \operatorname{Im}[z]$).



Für diese Zahlenpaare definiert man nun Verknüpfungen, so daß man mit ihnen rechnen kann wie mit reellen Zahlen.

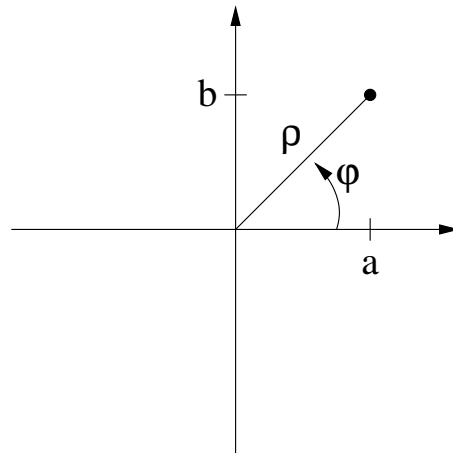
- **Addition:** $z_1 + z_2 = (a + ib) + (c + id) = a + c + ib + id = a + c + i(b + d)$. Für die Paare (a, b) und (c, d) also $(a, b) + (c, d) = (a + c, b + d)$. Dargestellt in der Gaußschen Zahlenebene entspricht die komplexe Addition der üblichen *Vektoraddition* (siehe später).



- **Multiplikation:** $z_1 z_2 = (a + ib)(c + id) = ac + ibc + iad + i^2 bd = ac - bd + i(bc + ad)$. Für die Paare (a, b) und (c, d) also $(a, b)(c, d) = (ac - bd, bc + ad)$. Es gilt damit

$$\begin{aligned} \operatorname{Re} [(a + ib)(c + id)] &= ac - bd \\ \operatorname{Im} [(a + ib)(c + id)] &= bc + ad. \end{aligned}$$

Um diese Rechenregel in der komplexen Zahlenebene darzustellen, ist es nützlich, die Punkte in *Polarkoordinaten* anzugeben.



Die Länge des Vektors vom Ursprung $(0,0)$ zum Punkt (a,b) nennt man den *Absolutbetrag* $|z|$ der komplexen Zahl z . Man schreibt

$$\rho = |z| = |a + ib| = \sqrt{a^2 + b^2}.$$

Den Winkel φ , der nur bis auf ganzzahlige Vielfache von 2π definiert ist, bezeichnet man oft als *Arcus* und schreibt

$$\varphi = \text{arc}(z) = \text{arc}(a + ib).$$

Aus der Skizze ist klar, daß $a = \rho \cos \varphi$ und $b = \rho \sin \varphi$, und für $a \neq 0$ gilt

$$\varphi = \arctan \left(\frac{b}{a} \right).$$

Damit können wir z zu

$$z = a + ib = \rho \cos \varphi + i\rho \sin \varphi = \rho e^{i\varphi}$$

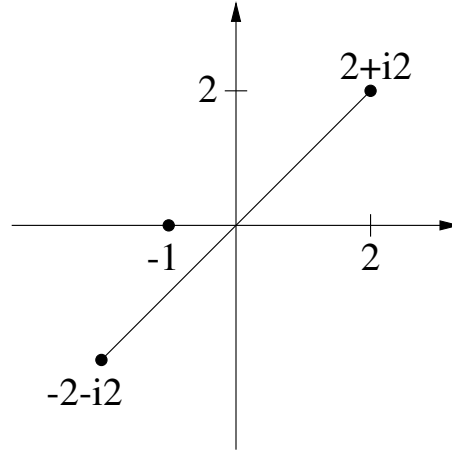
umschreiben, wobei wir Gl. (6.5) verwendet haben. Betrachten wir als Beispiel die komplexe Zahl $z = 2 + i2$. Dann gilt $\rho = \sqrt{4 + 4} = 2\sqrt{2}$ und $\varphi = \arctan \left(\frac{2}{2} \right) = \arctan 1 = \pi/4$. Also $2 + i2 = 2\sqrt{2}e^{i\pi/4}$.

Mit Hilfe dieser Einsichten können wir die Multiplikation zweier komplexer Zahlen in der Polardarstellung ausführen

$$z_1 z_2 = \rho_1 e^{i\varphi_1} \rho_2 e^{i\varphi_2} = \rho_1 \rho_2 e^{i(\varphi_1 + \varphi_2)}$$

und geometrisch interpretieren: Die Absolutbeträge werden multipliziert und die Polarwinkel addiert. Als Beispiel betrachten wir die Multiplikation von $z_1 = 2 + i2$ mit $z_2 = -1 + i0 = -1$, d.h. der rein reellen Zahl -1 - wir nehmen also das negative von z_1 : Wir haben $c = -1$ und $d = 0$. Damit folgt nach oben $z_1 z_2 = ac - bd + i(bc + ad) = -a - ib = -2 - i2$. In

Polardarstellung gilt $z_1 = 2\sqrt{2}e^{i\pi/4}$, $z_2 = e^{i\pi}$ und damit $z_1 z_2 = 2\sqrt{2}e^{i\pi 5/4}$ oder graphisch:



Die Subtraktion zweier komplexer Zahlen ergibt sich dann aus der Zusammensetzung der Regeln für die Multiplikation mit -1 und der Addition.

- **Division:** Bevor wir die Division zweier komplexer Zahlen definieren wollen, ist es sinnvoll, den Begriff der zu z *konjugiert komplexen Zahl* z^* (oft auch \bar{z}) einzuführen. Sie ist wie folgt gegeben

$$z = a + ib \quad \Rightarrow \quad z^* = a - ib,$$

d.h. der Imaginärteil von z geht in $-\text{Im}[z]$ über. Damit läßt sich der Absolutbetrag von z , $|z|$, leicht durch z und z^* ausdrücken. Aus $zz^* = (a + ib)(a - ib) = a^2 + b^2$ folgt

$$|z| = (zz^*)^{1/2}.$$

Nun zur Definition der Division. Sei dazu $z_2 = c + id \neq 0 + i0$. Dann betrachten wir¹

$$\frac{z_1}{z_2} = \frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{ac + bd + i(bc - ad)}{c^2 + d^2},$$

d.h. wir haben

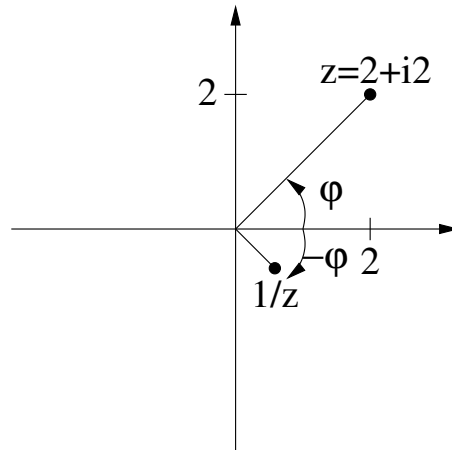
$$\text{Re} \left[\frac{a + ib}{c + id} \right] = \frac{ac + bd}{c^2 + d^2} \quad ; \quad \text{Im} \left[\frac{a + ib}{c + id} \right] = \frac{bc - ad}{c^2 + d^2},$$

und der “Trick” zur Berechnung war, mit dem konjugiert komplexen des Nenners zu erweitern. In Polarkoordinaten gilt

$$\frac{z_1}{z_2} = \frac{\rho_1 e^{i\varphi_1}}{\rho_2 e^{i\varphi_2}} = \frac{\rho_1}{\rho_2} e^{i(\varphi_1 - \varphi_2)},$$

¹Es kann aber durchaus $c = 0$ sein, wenn gleichzeitig $d \neq 0$ gilt, bzw. $d = 0$, wenn gleichzeitig $c \neq 0$.

da $\frac{1}{e^{ix}} = e^{-ix}$. Als Spezialfall betrachten wir $\frac{1}{z} = \frac{z^*}{zz^*} = \frac{z^*}{|z|^2}$, oder in Polarkoordinaten $\frac{1}{z} = \frac{1}{\rho}e^{-i\varphi}$ und graphisch



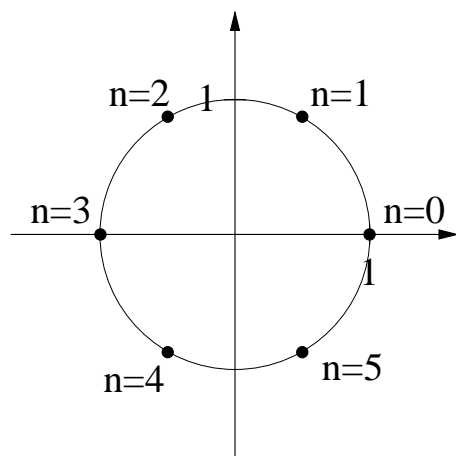
Wie für reelle Zahlen kann man nun Funktionen von komplexen Zahlen bilden - deren Funktionswert dann im allgemeinen wieder eine komplexe Zahl ist. Bereits mehrfach ist uns das Beispiel $\exp(ix)$, mit $x \in \mathbb{R}$, also die Exponentialfunktion mit *rein imaginärem* Argument begegnet. Man kann aber auch allgemeiner

$$f(z) = \exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = \exp(a + ib) = \exp(a) \exp(ib)$$

definieren, oder z.B. $\sin z$ und $\cos z$. Solche Funktionen und ihre komplexe Differentiation (nach z) und Integration werden im Rahmen der *Funktionentheorie* untersucht. Ein Besuch der Mathematikvorlesung über Funktionentheorie empfiehlt sich für jeden Physikstudenten (z.B. im dritten Studiensemester).

Manchmal dienen die komplexen Zahlen nur als Rechentrick - z.B. beim Integrieren - jedoch vereinfachen sich im Rahmen der komplexen Zahlen auch viele mathematische Aussagen, so daß sie ebenfalls von struktureller Bedeutung sind. Zum Beispiel gilt der hier nicht bewiesene *Fundamentalsatz der Algebra*: Sei $P_N(z) = a_0 + a_1z + a_2z^2 + \dots + a_Nz^N$ ein Polynom N -ten Grades, so hat es in der komplexen Ebene N Nullstellen z_i (mit $P_N(z_i) = 0$), wobei n -fache Nullstellen n -fach gezählt werden. Wir betrachten dazu ein Beispiel: Die reelle Gleichung $x^6 - 1 = 0$ hat die beiden reellen Nullstellen $x_{1/2} = \pm 1$. Die komplexe Gleichung $z^6 - 1 = 0$ hat dagegen vier weitere komplexe Nullstellen. Diese wollen wir jetzt finden. Mit $z^6 = 1$ gilt auch $(z^*)^6 = 1$, oder nach Multiplikation beider Gleichungen $(zz^*)^6 = 1$, d.h. $\rho^{12} = 1$. Da $\rho \geq 0$ reell ist, folgt $\rho = 1$. Damit haben die zu bestimmenden Nullstellen die Form $z_n = e^{i\varphi_n}$ mit $\varphi_n \in \mathbb{R}$. Wir betrachten dann $z_n^6 = (e^{i\varphi_n})^6 = e^{i6\varphi_n} = 1$. Da $e^{i2\pi m} = 1$ mit $m \in \mathbb{Z}$ gilt (siehe komplexe Zahlenebene), folgt durch Gleichsetzen $6\varphi_n = 2\pi n$ und damit $\varphi_n = n\frac{2\pi}{6}$ mit $n = 0, 1, 2, 3, 4, 5$.² Die zugehörigen z_n bezeichnet man als 6-te Einheitswurzeln.

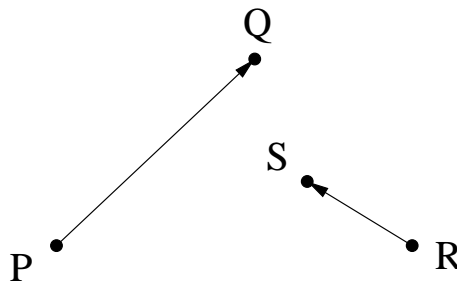
²Für $n = 6$ ergibt sich dieselbe Lösung wie für $n = 0$, für $n = 7$ dieselbe wie für $n = 1$, usw..



Kapitel 8

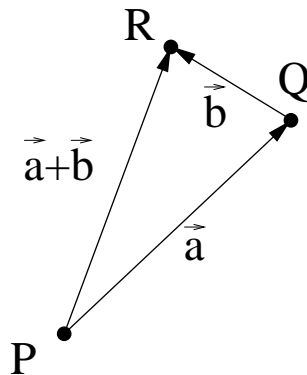
Vektorrechnung

Wir werden den Vektorbegriff anschaulich einführen und beschränken uns zunächst auf den zweidimensionalen euklidischen Raum. Die Elemente dieses Raumes sind Punkte P, Q, R, S, \dots . Geordnete Paare (P, Q) von Punkten werden mit einer Linie mit Pfeil verbunden, der vom Angriffspunkt P zum Zielpunkt Q zeigt: \vec{PQ} . Zwei solcher Pfeile sind *äquivalent*, wenn sie gleichsinnig parallel sind und gleiche Länge haben. Ein einzelner Pfeil einer Klasse äquivalenter Pfeile heißt *Repräsentant* des *Vektors* bzw. kurz *Vektor*. Vektoren werden wir mit lateinischen Buchstaben mit einem Pfeil \vec{a}, \vec{b}, \dots bezeichnen.



Diese Vektoren wollen wir nun z.B. durch Addition miteinander verknüpfen und durch die Multiplikation eines Vektors mit einer reellen (oder komplexen) Zahl - einem *Skalar* - einen Vektor auf einen anderen abbilden. Die Addition zweier Vektoren $\vec{a} + \vec{b}$ ist wie folgt definiert: Man wählt aus der Klasse \vec{a} einen beliebigen Pfeil (P, Q) und aus der Klasse \vec{b} den Pfeil (Q, R) mit dem Angriffspunkt Q und definiert

$$\vec{a} + \vec{b} = \vec{PQ} + \vec{QR} = \vec{PR}$$



Die Addition von Vektoren

- ist *kommutativ*: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$
- ist *assoziativ*: $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$
- hat ein *neutrales Element*, den Nullvektor $\vec{0}$, d.h. die Klasse von Pfeilen bei denen Angriffs- und Zielpunkt zusammenfallen und für den $\vec{a} + \vec{0} = \vec{a}$ gilt.
- Es ist anschaulich klar, daß es zu jedem Vektor \vec{a} einen Vektor \vec{b} gibt, so daß $\vec{a} + \vec{b} = \vec{0}$. Wir bezeichnen den Vektor \vec{b} auch mit $-\vec{a}$.

Allgemein nennt man die algebraische Struktur bestehend aus einer Menge G (hier die anschaulich definierten Vektoren) und einer Verknüpfungsvorschrift der Elemente von G (hier die ebenfalls anschaulich definierte und mit $+$ bezeichnete Addition zweier Vektoren) welche den obigen vier Eigenschaften genügt, eine *abelsche Gruppe*.

Kehren wir von dieser abstrakten Sichtweise zu den anschaulich definierten Vektoren in der zweidimensionalen Ebene zurück. Die Multiplikation eines Vektors mit einer reellen Zahl soll wie folgt definiert sein: Die Pfeile $\lambda\vec{a}$ sind für $\lambda > 0$ gleichsinnig parallel und für $\lambda < 0$ gegensinnig parallel zu den Pfeilen der Klasse \vec{a} . Sie haben die Länge $|\lambda| \cdot \text{Länge von } \vec{a}$. Für $\lambda = 0$ gilt $\lambda\vec{a} = \vec{0}$. Aus dieser Definition ergeben sich sofort die folgenden Eigenschaften für die Multiplikation von Vektoren mit Skalaren ($\lambda, \mu \in \mathbb{R}$):

- $1\vec{a} = \vec{a}$
- $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$
- $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$
- $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$

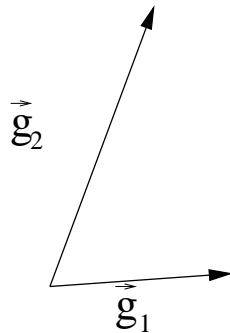
Wiederum abstrakt betrachtet nennt man eine abelsche Gruppe für die man eine Multiplikation mit einer reellen Zahl definieren kann¹ welche die gerade diskutierten Eigenschaften hat einen *reellen Vektorraum*.² Den reellen Vektorraum der oben anschaulich definierten Pfeile in zwei Raumdimensionen (zweidimensionale Ebene) bezeichnet man mit \mathbb{R}^2 . Analog kann man Vektorräume in d Dimensionen definieren. Sie werden allgemein mit \mathbb{R}^d bezeichnet.³

Um die Allgemeingültigkeit der Konzepte und Begriffe anzudeuten sind wir mehrfach vom exemplarischen Beispiel der Vektoren in der zweidimensionalen Ebene zu der abstrakten Sichtweise übergegangen. Auch in der Physik - nicht nur der Mathematik - spielen neben dem Vektorraum \mathbb{R}^d (bzw. \mathbb{C}^d) andere Vektorräume eine große Rolle. Zum Verständnis des Vorlesung *Physik I* benötigen sie jedoch nur Kenntnisse des \mathbb{R}^d , so daß wir unsere weiteren Überlegungen auf diesen beschränken werden.

Eine Anzahl von m Vektoren $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_m$ aus dem \mathbb{R}^d heißt *linear unabhängig*, wenn

$$\lambda_1 \vec{g}_1 + \lambda_2 \vec{g}_2 + \dots + \lambda_m \vec{g}_m = \vec{0}$$

nur dann gilt, wenn alle Koeffizienten λ_i gleich Null sind. Anderenfalls nennt man sie *linear abhängig*. Im \mathbb{R}^d gibt es (maximal) d linear unabhängige Vektoren, nimmt man jedoch einen weiteren Vektor hinzu, so sind die dann gegebenen $d + 1$ Vektoren linear abhängig. Zum Beispiel in $d = 2$ gilt:



\vec{g}_1 und \vec{g}_2 sind linear unabhängig, aber jeder beliebige dritte Vektor \vec{g}_3 läßt sich als *Linearkombination* von \vec{g}_1 und \vec{g}_2 schreiben

$$\vec{g}_3 = \lambda_1 \vec{g}_1 + \lambda_2 \vec{g}_2$$

mit zu bestimmenden Skalaren $\lambda_1, \lambda_2 \in \mathbb{R}$. Man bezeichnet einen Satz $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_d$ von d linear unabhängigen Vektoren im \mathbb{R}^d als *Basis*.

¹Das Ergebnis der Multiplikation eines Elementes der abelschen Gruppe mit der reellen Zahl muß, wie im obigen Beispiel, wieder ein Gruppenelement sein.

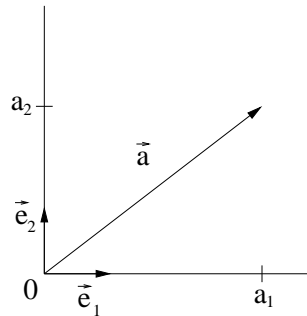
²Analog gibt es auch *komplexe Vektorräume*. In diesem Fall ist das Skalar eine komplexe Zahl aus \mathbb{C} .

³Ist die Multiplikation mit einem Skalar bezüglich der komplexen Zahlen definiert, so bezeichnet man die Vektorräume mit \mathbb{C}^d .

Der Einfachheit halber werden wir hier gleich mit *Orthogonalbasen* im \mathbb{R}^d (hier speziell im \mathbb{R}^2) arbeiten,⁴ genauer mit *Orthonormalbasen*: Die Basisvektoren \vec{e}_1 und \vec{e}_2 haben die Länge 1⁵ und stehen *senkrecht* aufeinander. Jeder Vektor \vec{a} hat dann die Zerlegung

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2.$$

Die a_i nennt man *Vektorkomponenten* zur vorgegebenen Basis \vec{e}_i .



Wir haben bei der obigen Konstruktion einen festen Punkt 0 als Anfangspunkt oder *Ursprung* ausgezeichnet. Damit ist dann jeder Punkt P durch den Vektor $\vec{x} = 0\vec{P}$ bestimmt. Man nennt diesen Vektor \vec{x} den *Ortsvektor* von P bezüglich des Anfangspunktes 0. In Komponenten kann dieser Vektor als $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$ geschrieben werden, und jeder Punkt P ist damit umkehrbar eindeutig durch die Zahlen (x_1, x_2) festgelegt. Sie heißen die *Koordinaten* des Punktes P bezüglich des *Koordinatensystems* $[0; \vec{e}_1, \vec{e}_2]$. Arbeitet man immer mit demselben Koordinatensystem, so bietet sich die Notation

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

an. Diese kennen sie sicherlich auch aus der Schule. Wir werden diese Notation hier nicht verwenden, da in der Physik häufig ein *Basiswechsel* notwendig ist, um die physikalischen Phänomene besser zu verstehen (z.B. den Basiswechsel in ein beschleunigtes Koordinaten- bzw. Bezugssystem). Wir schreiben manchmal bei vorgegebenem Koordinatensystem

$$[0; \vec{e}_1, \vec{e}_2] : \vec{x} \doteq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

wobei \doteq als “wird dargestellt durch” zu lesen ist. Zum Beispiel schreiben wir statt

$$\vec{a} + \vec{b} = a_1\vec{e}_1 + a_2\vec{e}_2 + b_1\vec{e}_1 + b_2\vec{e}_2 = (a_1 + b_1)\vec{e}_1 + (a_2 + b_2)\vec{e}_2$$

⁴Man kann auch allgemeiner mit nicht-orthogonalen Basisvektoren arbeiten.

⁵Zur genauen Definition des Begriffs “Länge” siehe unten.

einfach

$$\vec{a} + \vec{b} \doteq \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}.$$

Die Länge $|\vec{a}|$ eines Vektors ergibt sich nach obiger Skizze und dem Satz von Pythagoras zu

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2} = a.$$

Die Länge hat aber auch eine geometrische Bedeutung, die unabhängig von der Wahl des Koordinatensystems ist. Unabhängig von der Wahl des Koordinatensystems definieren wir das *Skalarprodukt* zweier Vektoren

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \varphi(\vec{a}, \vec{b}),$$

wobei $\varphi(\vec{a}, \vec{b})$ der von \vec{a} und \vec{b} eingeschlossene Winkel ist, wenn \vec{a} und \vec{b} denselben Angriffspunkt haben. Da die Kosinusfunktion gerade ist, gilt offensichtlich das *Kommutativgesetz* $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$. Um zu zeigen, daß das *Distributivgesetz* gilt, berechnen wir das Skalarprodukt in einer festen Orthonormalbasis:

$$\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 \quad \Rightarrow \quad a_1 = |\vec{a}| \cos \varphi_a, \quad a_2 = |\vec{a}| \sin \varphi_a,$$

wobei φ_a der Winkel zwischen der \vec{e}_1 -Richtung und \vec{a} ist. Analog gilt

$$\vec{b} = b_1 \vec{e}_1 + b_2 \vec{e}_2 \quad \Rightarrow \quad b_1 = |\vec{b}| \cos \varphi_b, \quad b_2 = |\vec{b}| \sin \varphi_b,$$

und damit

$$\begin{aligned} \vec{a} \cdot \vec{b} &= |\vec{a}||\vec{b}| \cos(\varphi_b - \varphi_a) \\ &= |\vec{a}||\vec{b}| [\cos \varphi_b \cos \varphi_a + \sin \varphi_b \sin \varphi_a] \\ &= a_1 b_1 + a_2 b_2, \end{aligned}$$

wie sie es wahrscheinlich aus der Schule kennen. Ist dann aber $\vec{b} = \vec{c} + \vec{d}$, so folgt

$$\begin{aligned} \vec{a} \cdot (\vec{c} + \vec{d}) &= a_1(c_1 + d_1) + a_2(c_2 + d_2) \\ &= a_1 c_1 + a_2 c_2 + a_1 d_1 + a_2 d_2 \\ &= \vec{a} \cdot \vec{c} + \vec{a} \cdot \vec{d}, \end{aligned}$$

d.h. das Distributivgesetz. Mit Hilfe dieses Gesetzes kann man $\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2$ auch anders verstehen: Wegen $\vec{e}_1 \cdot \vec{e}_1 = 1 = \vec{e}_2 \cdot \vec{e}_2$ und $\vec{e}_1 \cdot \vec{e}_2 = 0$ - da $\varphi(\vec{e}_1, \vec{e}_2) = \pi/2$ und $\cos(\pi/2) = 0$ - d.h.

$$\vec{e}_i \cdot \vec{e}_j = \delta_{i,j} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad (8.1)$$

folgt

$$\vec{a} \cdot \vec{b} = (a_1 \vec{e}_1 + a_2 \vec{e}_2) \cdot (b_1 \vec{e}_1 + b_2 \vec{e}_2) = a_1 b_1 + a_2 b_2.$$

Das soeben eingeführte und definierte $\delta_{i,j}$ bezeichnet man als Kronecker-Symbol bzw. Kronecker-Delta. Über das Skalarprodukt läßt sich die Länge $|\vec{a}|$ eines Vektors damit auch zu

$$|\vec{a}| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{a_1^2 + a_2^2}$$

schreiben.

Mit Hilfe des Distributivgesetzes beweist man sehr leicht den aus der Schule bekannten *Kosinussatz*

$$\begin{aligned} \vec{c} \cdot \vec{c} &= (\vec{b} - \vec{a}) \cdot (\vec{b} - \vec{a}) \\ &= \vec{a} \cdot \vec{a} + \vec{b} \cdot \vec{b} - 2\vec{a} \cdot \vec{b}, \end{aligned}$$

also

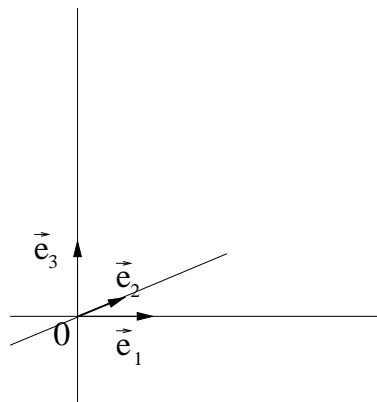
$$|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}||\vec{b}| \cos \varphi(\vec{a}, \vec{b}).$$

Wir können jetzt auch andersherum vorgehen, den Kosinussatz (weil aus der Schule bekannt) voraussetzen und einen Ausdruck für das Skalarprodukt herleiten:

$$\begin{aligned} |\vec{a}||\vec{b}| \cos \varphi(\vec{a}, \vec{b}) &= \frac{|\vec{a}|^2 + |\vec{b}|^2 - |\vec{c}|^2}{2} \\ &= \frac{a_1^2 + a_2^2 + b_1^2 + b_2^2 - [(b_1 - a_1)^2 + (b_2 - a_2)^2]}{2} \\ &= a_1 b_1 + a_2 b_2. \end{aligned}$$

Wir erhalten so also wieder das bereits bekannte Ergebnis.

Bisher haben wir meistens im zweidimensionalen Raum (der Ebene) argumentiert. Wir werden jetzt das Skalarprodukt auf den dreidimensionalen Fall - der in der Physik eine zentrale Rolle spielt - erweitern. Dazu betrachten wir zunächst eine Orthonormalbasis in $d = 3$.



Das durch die gezeigten Basisvektoren aufgespannte Koordinatensystem bezeichnet man als *Rechtssystem*, was man sich wie folgt erklären und merken kann: Zeigt der Daumen der rechten Hand in Richtung \vec{e}_1 , der Zeigfinger der rechten Hand in Richtung \vec{e}_2 und der Mittelfinger der rechten Hand in Richtung \vec{e}_3 , so handelt es sich um ein Rechtssystem. Macht man dasselbe mit der linken Hand, handelt es sich um ein Linkssystem. Ein Vektor \vec{a} wird in dieser Basis dann durch die Komponenten

$$\vec{a} \doteq \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

dargestellt, und durch zweimaliges Anwenden des Satzes von Pythagoras sieht man leicht, daß

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

gilt. Wir definieren dann das Skalarprodukt in $d = 3$ zu

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \varphi(\vec{a}, \vec{b}).$$

Um eine Komponentenschreibweise zu finden, kann man wieder den Kosinussatz verwenden, da \vec{a} , \vec{b} und $\vec{c} = \vec{b} - \vec{a}$ in einer Ebene des dreidimensionalen Raumes liegen

$$\begin{aligned} \vec{a} \cdot \vec{b} &= |\vec{a}||\vec{b}| \cos \varphi(\vec{a}, \vec{b}) \\ &= \frac{|\vec{a}|^2 + |\vec{b}|^2 - |\vec{c}|^2}{2} \\ &= \frac{a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - [(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2]}{2} \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3, \end{aligned}$$

was für jede beliebige Orthonormalbasis gilt. Völlig analog gilt in d -Dimensionen

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \dots + a_d b_d.$$

Wie in $d = 2$ ergibt sich damit sofort wieder das Distributivgesetz

$$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}.$$

Aus der Definition des Skalarprodukts läßt sich der Winkel, den zwei beliebige Vektoren (in $d = 3$) einschließen, bestimmen:

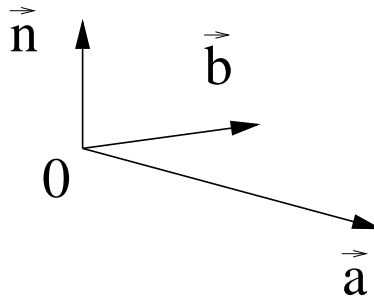
$$\cos \varphi(\vec{a}, \vec{b}) = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{b_1^2 + b_2^2 + b_3^2}}.$$

Zwei Vektoren stehen also senkrecht aufeinander (sind orthogonal zueinander), wenn $\vec{a} \cdot \vec{b} = 0$ gilt.

Für zwei allgemeine dreidimensionale Vektoren definieren wir nun das *Vektorprodukt* (oft auch *Kreuzprodukt* oder *äußeres Produkt* genannt) wie folgt

$$\vec{a} \times \vec{b} = \vec{n} |\vec{a}| |\vec{b}| \sin \tilde{\varphi}(\vec{a}, \vec{b}),$$

wobei \vec{n} ein Vektor der Länge 1 ist und auf \vec{a} und \vec{b} senkrecht steht. Die Richtung von \vec{n} ist so zu wählen, daß \vec{a} , \vec{b} und \vec{n} ein Rechtssystem bilden. $\tilde{\varphi}(\vec{a}, \vec{b})$ bezeichnet dabei den Betrag des *kleineren der beiden Winkel* zwischen \vec{a} und \vec{b} , so daß $0 < \tilde{\varphi}(\vec{a}, \vec{b}) < \pi$ und damit $\sin \tilde{\varphi}(\vec{a}, \vec{b}) > 0$ gilt. Die Richtung von \vec{n} läßt sich auch wie folgt bestimmen: Zeigen alle Finger der rechten Hand mit Ausnahme des Daumen in die Richtung der Drehung von \vec{a} auf \vec{b} mit dem kleineren der beiden Winkel, so zeigt der abgespreizte Daumen in die Richtung von \vec{n} (“Daumen-Regel der rechten Hand”).



Nach der Definition erfüllt das Vektorprodukt die folgenden Relationen

$$\begin{aligned} \vec{a} \times \vec{b} &= -\vec{b} \times \vec{a} \\ (\lambda \vec{a}) \times (\mu \vec{b}) &= (\lambda \mu) (\vec{a} \times \vec{b}), \quad \lambda, \mu \in \mathbb{R}. \end{aligned} \tag{8.2}$$

Aus der Definition ergibt sich ebenfalls

$$\vec{a} \times \vec{a} = \vec{0}. \tag{8.3}$$

Das Vektorprodukt ist distributiv

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c},$$

was wir hier jedoch aus Zeitgründen nicht beweisen wollen.⁶ Um die Komponenten von $\vec{a} \times \vec{b}$ bezüglich einer Orthonormalbasis angeben zu können, wollen wir

⁶Für einen Beweis siehe z.B. S. Großmann, “Mathematischer Einführungskurs für die Physik”, Teubner Verlag, Stuttgart 1984, Seite 70.

zunächst betrachten, was sich nach der Definition des Vektorproduktes für $\vec{e}_i \times \vec{e}_j$ mit $i, j = 1, 2, 3$ ergibt:

$$\begin{aligned}\vec{e}_1 \times \vec{e}_2 &= \vec{e}_3 \\ \vec{e}_2 \times \vec{e}_3 &= \vec{e}_1 \\ \vec{e}_1 \times \vec{e}_3 &= -\vec{e}_2.\end{aligned}$$

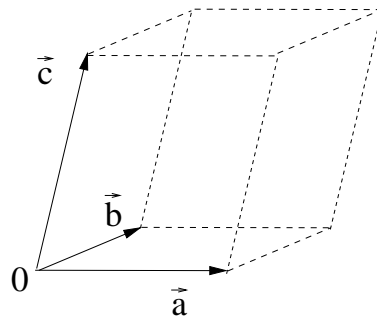
Die übrigen Terme ergeben sich mit Hilfe der Gln. (8.2) und (8.3). Man kann diese Gleichungen zu

$$\vec{e}_i \times \vec{e}_j = \vec{e}_k$$

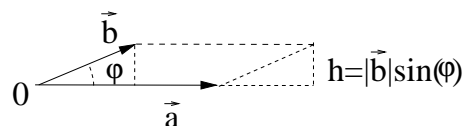
zusammenfassen, wobei (i, j, k) eine *zyklische Vertauschung* von $(1, 2, 3)$ sein muß - also $(1, 2, 3)$, $(2, 3, 1)$ oder $(3, 1, 2)$. Für die Komponenten bezüglich der Basis $[\vec{e}_1, \vec{e}_2, \vec{e}_3]$ folgt damit

$$\begin{aligned}\vec{a} \times \vec{b} &= (a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3) \times (b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3) \\ &= (a_1b_2 - a_2b_1)(\vec{e}_1 \times \vec{e}_2) + (a_1b_3 - a_3b_1)(\vec{e}_1 \times \vec{e}_3) + (a_2b_3 - a_3b_2)(\vec{e}_2 \times \vec{e}_3) \\ &= (a_2b_3 - a_3b_2)\vec{e}_1 + (a_3b_1 - a_1b_3)\vec{e}_2 + (a_1b_2 - a_2b_1)\vec{e}_3,\end{aligned}$$

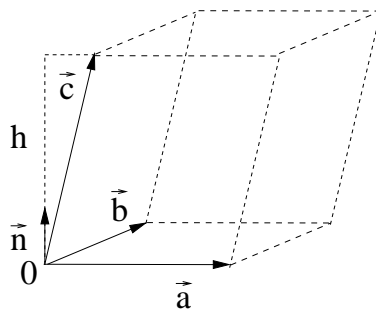
eine Regel, die sie sicherlich aus der Schule kennen. In der Mechanik wird das Kreuzprodukt z.B. in der Definition der Begriffe *Drehimpuls* und *Drehmoment* benötigt. Als mathematisches Beispiel für die Anwendung des Vektorproduktes wollen wir hier die Berechnung des Volumens eines durch drei nicht *koplanare* Vektoren \vec{a} , \vec{b} und \vec{c} aufgespannten Körpers betrachten.



Betrachten wir dazu zunächst die durch \vec{a} und \vec{b} aufgespannte Fläche. Sie ist durch Grundlinie * Höhe gegeben.



Das ist aber gerade der Betrag von $\vec{a} \times \vec{b}$. Die Richtung \vec{n} von $\vec{a} \times \vec{b}$ ist in der folgenden Skizze eingezeichnet. Das Volumen in der obigen Skizze ergibt sich dann zu Grundfläche * Höhe.



Die Höhe h ist aber gerade durch $\vec{n} \cdot \vec{c}$ gegeben, so daß insgesamt

$$V(\vec{a}, \vec{b}, \vec{c}) = (\vec{a} \times \vec{b}) \cdot \vec{c}$$

folgt. Man nennt $V(\vec{a}, \vec{b}, \vec{c})$ auch das *Spatprodukt*.

Kapitel 9

Vektorwertige Funktionen

Wie in der Einleitung diskutiert benötigt man in der Mechanik *vektorwertige Funktionen*, d.h. Abbildungen, bei denen einem Element aus \mathbb{R} eindeutig ein (meist dreidimensionaler) Vektor zugeordnet wird. Die Bahnkurve $\vec{x}(t)$ eines Teilchens, d.h. der zeitliche Verlauf des Ortes des Teilchens im dreidimensionalen Raum, ist ein Beispiel für solch eine Funktion. Legen wir eine Basis $[\vec{e}_1, \vec{e}_2, \vec{e}_3]$ fest (hier wieder eine Orthonormalbasis), so können wir jede vektorwertige Funktion $\vec{f}(x)$ als

$$\vec{f}(x) = f_1(x)\vec{e}_1 + f_2(x)\vec{e}_2 + f_3(x)\vec{e}_3$$

schreiben. Wir definieren zunächst die Stetigkeit einer vektorwertigen Funktion, indem wir die Stetigkeit aller ihrer Komponenten $f_i(x)$ fordern. Die Ableitung einer vektorwertigen Funktion liefert wieder eine solche und ist definiert als

$$\vec{f}'(x) = \lim_{h \rightarrow 0} \frac{\vec{f}(x+h) - \vec{f}(x)}{h}.$$

Drücken wir \vec{f} erneut durch die Komponenten aus, so folgt

$$\begin{aligned} \vec{f}'(x) &= \lim_{h \rightarrow 0} \left[\frac{f_1(x+h) - f_1(x)}{h} \vec{e}_1 + \frac{f_2(x+h) - f_2(x)}{h} \vec{e}_2 \right. \\ &\quad \left. + \frac{f_3(x+h) - f_3(x)}{h} \vec{e}_3 \right] \\ &= f_1'(x)\vec{e}_1 + f_2'(x)\vec{e}_2 + f_3'(x)\vec{e}_3, \end{aligned}$$

d.h. \vec{f}' ergibt sich durch komponentenweise Differentiation. Aus dieser Einsicht und den Differentiationsregeln für gewöhnliche Funktionen ergeben sich die folgenden Gesetze für vektorwertige Funktionen $\vec{a}(x)$, $\vec{b}(x)$ und die skalare Funktion $g(x)$:

- $(\vec{a}(x) + \vec{b}(x))' = \vec{a}'(x) + \vec{b}'(x)$

- $\left(\vec{a}(x) \cdot \vec{b}(x)\right)' = \vec{a}'(x) \cdot \vec{b}(x) + \vec{a}(x) \cdot \vec{b}'(x)$ (Produktregel)
- $\left(\vec{a}(x) \times \vec{b}(x)\right)' = \vec{a}'(x) \times \vec{b}(x) + \vec{a}(x) \times \vec{b}'(x)$ (Produktregel)
- $(g(x)\vec{a}(x))' = g'(x)\vec{a}(x) + g(x)\vec{a}'(x)$ (Produktregel)
- $(\vec{a}(g(x)))' = \vec{a}'(g(x))g'(x)$ (Kettenregel)

Wie für gewöhnliche Funktionen lassen sich höhere Ableitungen vektorwertiger Funktionen rekursiv definieren

$$\frac{d^n \vec{f}(x)}{dx^n} = \frac{d}{dx} \left(\frac{d^{n-1} \vec{f}(x)}{dx^{n-1}} \right).$$

Eine weitere Diskussion vektorwertiger Funktionen verschieben wir auf die Übungen.

Das Ableiten einer vektorwertigen Funktion darf nicht mit dem Ableiten einer Funktion, die von einem Vektor abhängt - und selbst wieder vektorwertig sein kann (Felder) - nach eben diesem Vektor, also dem Bilden des Gradienten, der Divergenz oder der Rotation, verwechselt werden. Auf diese werden wir im folgenden eingehen.

Kapitel 10

Funktionen mehrerer Veränderlicher

Auch wenn es in der Mechanik oft gelingt, durch geschickte Wahl der Koordinaten (Basisvektoren) Probleme so zu vereinfachen, daß sie von nur einer Orts- (oder auch Winkel-) Koordinate abhängen, so ist es zum allgemeinen Verständnis doch unvermeidbar, Funktionen mehrerer Veränderlicher - also z.B. der drei Koordinaten x_1, x_2, x_3 bezüglich einer fest gewählten Orthonormalbasis - einzuführen. Im allgemeinen kann die Kraft $\vec{F}(\vec{x})$, die auf ein Teilchen am Punkt \vec{x} wirkt, von diesen drei Koordinaten abhängen. Weiterhin ist \vec{F} selbst ein Vektor. Die Kraft ist also eine vektorwertige Funktion eines Vektors. Betrachten wir als Beispiel die ihnen sicherlich aus der Schule bekannte Gravitationskraft zwischen zwei Punkt-Teilchen der Massen M und m . Wir wählen ein Koordinatensystem mit einer orthonormalen Basis und dem Ursprung in der Masse M . Die Kraft auf das sich am Ort \vec{x} befindliche Teilchen der Masse m ist dann

$$\vec{F}(\vec{x}) = -\gamma \frac{Mm}{x_1^2 + x_2^2 + x_3^2} \frac{\vec{x}}{\sqrt{x_1^2 + x_2^2 + x_3^2}},$$

mit der Gravitationskonstanten γ . Den Vektor $\vec{x}/|\vec{x}|$ bezeichnet man auch als *Richtungsvektor*. Er hat die Länge 1. Die Kraft auf die Masse m an der Position \vec{x} zeigt damit immer ins Zentrum des Koordinatensystems - also auf die Masse M - und ihr Betrag fällt wie $1/|\vec{x}|^2$ mit dem Abstand vom Ursprung ab. Die Funktion \vec{F} bildet damit das Zahlentripel x_1, x_2, x_3 aus dem Raum $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$ auf ein Element F_1, F_2, F_3 aus eben diesem Raum ab. Bei dieser abstrakten Betrachtung haben wir die Dimensionen der physikalischen Meßgrößen Ort und Kraft vernachlässigt. Allgemein definiert man die Räume \mathbb{R}^n gemäß

$$\mathbb{R}^n = \{x_1, x_2, \dots, x_n | x_i \in \mathbb{R} \text{ für } i = 1, 2, \dots, n\} .$$

Als einfacheres Beispiel betrachten wir eine sehr dünne, "zweidimensionale" Metallplatte, die an verschiedenen Stellen beheizt wird. Damit ergibt sich zu

jedem festen Zeitpunkt t eine mit der Position¹

$$\vec{x} \doteq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

auf der Platte variierende Temperatur $T(\vec{x})$. Im Gegensatz zum Beispiel der Gravitationskraft ist die Funktion selbst in diesem Fall ein Skalar. Abstrakt betrachtet und für eine unendlich ausgedehnte Platte ist die Funktion T eine Abbildung vom \mathbb{R}^2 auf eine Teilmenge \mathbb{W} - den Wertebereich - der reellen Zahlen \mathbb{R} . Dabei haben wir erneut die Dimensionen des Ortes und der Temperatur vernachlässigt. Da es (in der Kelvin-Skala) keine negativen Temperaturen gibt, kommt als Wertebereich nur eine Teilmenge der positiven reellen Zahlen in Frage. Ist die Platte endlich, müssen wir auch einen Definitionsbereich \mathbb{D} der Funktion festlegen, bei dem es sich um eine Teilmenge von Zahlenpaaren aus den reellen Zahlen handelt. Für eine quadratische Platte der Kantenlänge $2a$ und dem Nullpunkt des Koordinatensystems im Zentrum gilt dann z.B.

$$\mathbb{D} = \{(x_1, x_2) \in \mathbb{R}^2 \mid -a \leq x_1 \leq a \text{ und } -a \leq x_2 \leq a\}.$$

Mißt man die lokale Temperatur der Platte nicht nur zu einer festen Zeit, sondern zu verschiedenen Zeiten t , so stellt die Temperatur eine Funktion von drei Veränderlichen dar. Dabei haben nur die beiden Ortskoordinaten x_1, x_2 die Dimension Länge, die dritte Variable jedoch die Dimension Zeit.

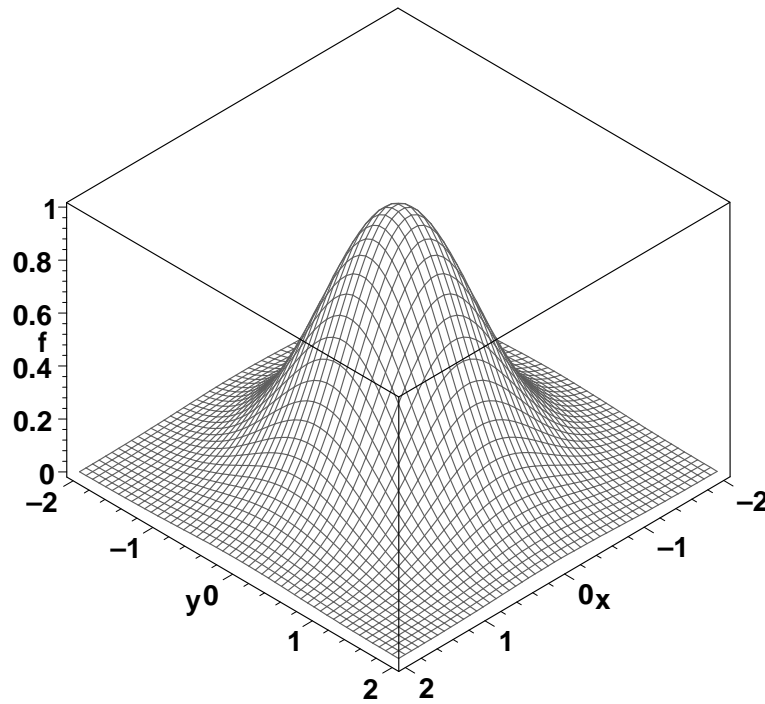
Ob es sich bei der Funktion um einen Vektor oder ein Skalar handelt, macht bei denen uns im folgenden interessierenden mathematischen Fragestellungen keinen Unterschied, und wir werden uns daher auf skalare Funktionen beschränken. Die gewonnenen Ergebnisse gelten im Falle von vektorwertigen Funktionen für jede ihrer Komponenten. Da sich weiterhin Funktionen zweier Veränderlicher gut graphisch darstellen lassen, werden wir zusätzlich meist solche Funktionen betrachten.

Nach dieser physikalischen Motivation betrachten wir zunächst einmal die Graphen dreier Beispielfunktionen.²

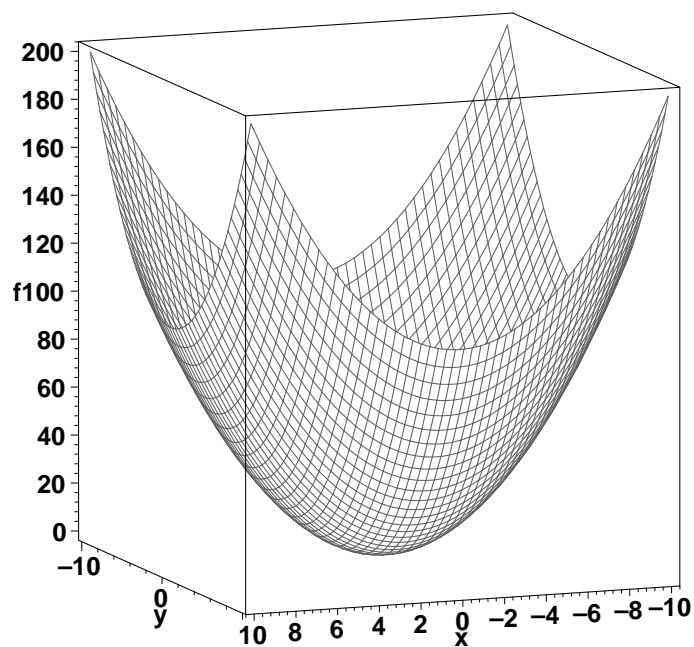
¹Wir verwenden wieder die “wird dargestellt durch” Notation, d.h. wir haben ein Koordinatensystem vorgegeben.

²Bisher haben wir immer $f(\vec{x})$ geschrieben. Alternativ kann man auch die Komponenten als Argumente explizit angeben: $f(x_1, x_2, \dots, x_n)$. Im folgenden sind x und y die Komponenten.

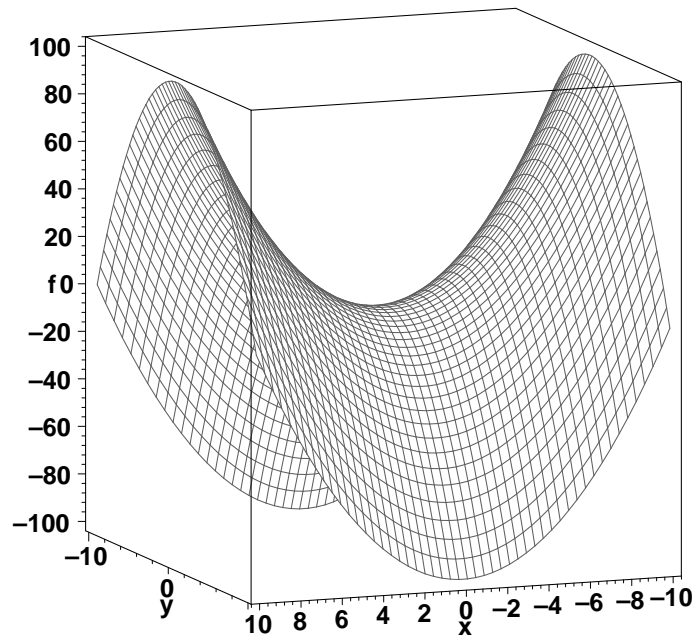
$$f(x, y) = \exp(-x^2 - y^2) : \mathbb{D} = \mathbb{R}^2, \mathbb{W} = (0, 1]$$



$$f(x, y) = x^2 + y^2 : \mathbb{D} = \mathbb{R}^2, \mathbb{W} = [0, \infty)$$



$$f(x, y) = x^2 - y^2 \quad : \quad \mathbb{D} = \mathbb{R}^2 \quad , \quad \mathbb{W} = \mathbb{R}$$



Die drei Beispielfunktionen haben unterschiedliche Eigenschaften. So fällt die erste, wenn man im Ursprung der x - y -Ebene startet, in alle Richtungen hin ab, die zweite steigt an, und bei der dritten hängt es davon ab, in welche Richtung man fortschreitet, ob die Funktion ansteigt oder abfällt. Auch sind die ersten beiden Funktionen symmetrisch bei einer Rotation mit beliebigem Winkel um die f -Achse, d.h., bei einer Rotation (mit beliebigem Winkel) um diese Achse geht der Graph der Funktion in sich selbst über. Bei der dritten Funktion ist das nicht der Fall. Dies liegt daran, daß die ersten beiden Funktionen nur von $x^2 + y^2$, d.h. dem Abstand (zum Quadrat) vom Ursprung der x - y -Ebene abhängen. Dies gilt so auch für die oben betrachteten Komponenten der Gravitationskraft. Die dritte Beispielfunktion erinnert an einen Sattel. Man nennt den Punkt $x = 0$, $y = 0$ daher auch einen *Sattelpunkt*. Wie wir für Funktionen einer Variablen gelernt haben, gibt uns die Ableitung einer Funktion an, wie sich der Funktionswert ändert, wenn wir von einer Stelle x nach $x + h$, für $h \rightarrow 0$, weitergehen. Wie wir gerade gelernt haben, kann die Änderung im Falle einer Funktion mehrerer Veränderlicher davon abhängen, in welche Richtung man weitergeht. Zur Beschreibung der Änderung benötigen wir somit eine vektorielle Größe. Diese wollen wir im folgenden wenig formal "herleiten".

Betrachten wir eine Funktion $f(x, y)$, so können wir uns fragen, wie sich der Funktionswert ändert, wenn wir ausgehend von x, y in die x -Richtung fortschreiten. Die Steigung an der Stelle x, y in diese Richtung wird uns durch die Ableitung von f nach x bei festem y gegeben, wenn wir als Argumente der abgeleiteten

Funktion wieder x, y einsetzen. Man nennt diese Ableitung die *partielle Ableitung* nach x und schreibt

$$\frac{\partial f(x, y)}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

Als Beispiel bilden wir die partielle Ableitung unserer ersten Beispielfunktion nach x

$$\frac{\partial f(x, y)}{\partial x} = -2x \exp(-x^2 - y^2) \quad (10.1)$$

und auch gleich die nach y

$$\frac{\partial f(x, y)}{\partial y} = -2y \exp(-x^2 - y^2). \quad (10.2)$$

Dabei haben wir stillschweigend angenommen, daß die Funktion f hinreichend harmlos ist, so daß die entsprechenden Ableitungen existieren - wie es in unserem Beispiel natürlich der Fall ist. Betrachten wir ein weiteres Beispiel

$$\frac{\partial}{\partial x_1} (x_1^2 + 2x_1x_2^3 + 5) = 2x_1 + 2x_2^3,$$

in dem die beiden Variablen mal nicht mit x und y bezeichnet sind. Die Ableitungsregeln für Funktionen einer Veränderlichen (Produktregel, Quotientenregel und Kettenregel) gelten auch für die partielle Ableitung, da es sich ja bei ihr um eine gewöhnliche Ableitung nach einer Variablen bei festgehaltenen anderen handelt.

Für die folgende Überlegung nehmen wir an, daß wir eine hinreichend glatte Funktion f haben - so daß alle Ableitungen, die wir bilden wollen, auch existieren - die von den n Variablen x_1, x_2, \dots, x_n abhängt. Wie für Funktionen einer Veränderlichen kann man partielle Ableitungen *höherer Ordnung* definieren. Bei partiellen Ableitungen gibt es aber nun die zusätzliche Option nicht nur zweimal nach x_i zu differenzieren, sondern erst nach x_i und dann nach x_j mit $j \neq i$

$$\frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f(x_1, x_2, \dots, x_n) = \frac{\partial^2 f(x_1, x_2, \dots, x_n)}{\partial x_j \partial x_i}.$$

Diese Idee läßt sich beliebig in höhere Ordnungen fortsetzen.

Kommen wir nun zurück zu der Frage, einen Vektor zu konstruieren, der die Information über die Steigung einer Funktion mehrerer Veränderlicher enthält. Betrachten wir zunächst unsere erste Beispielfunktion $f(x, y) = \exp(-x^2 - y^2)$. Mit den Gln. (10.1) und (10.2) wissen wir bereits, wie sich die Funktion ändert, wenn wir in x - bzw. y -Richtung fortschreiten. Das legt nahe, daß der Vektor

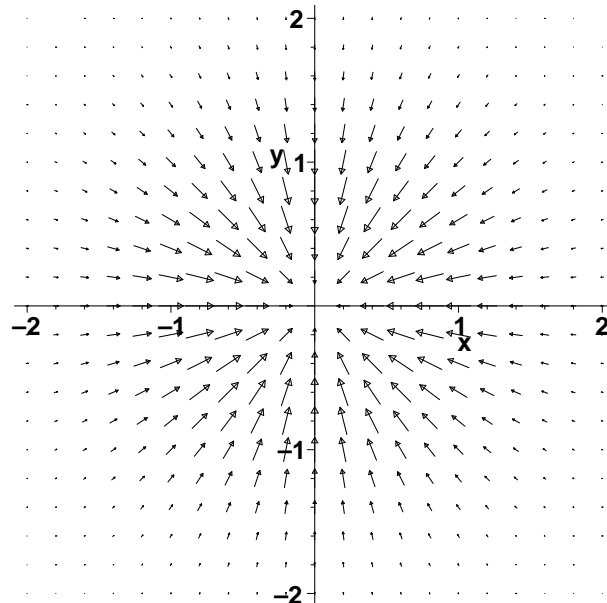
$$\text{grad } f(x, y) = \vec{\nabla} f(x, y) \doteq \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} = \begin{pmatrix} -2x \exp(-x^2 - y^2) \\ -2y \exp(-x^2 - y^2) \end{pmatrix}$$

der von uns gesuchte “Steigungsvektor” ist. Das Symbol grad steht für *Gradient*. Die Notationen $\text{grad } f$ und $\vec{\nabla} f$ sind beide gebräuchlich. Wir werden hier im folgenden die Notation $\vec{\nabla} f$ verwenden. Die Verallgemeinerung auf eine Funktion von n Veränderlichen ist offensichtlich

$$\vec{\nabla} f \doteq \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

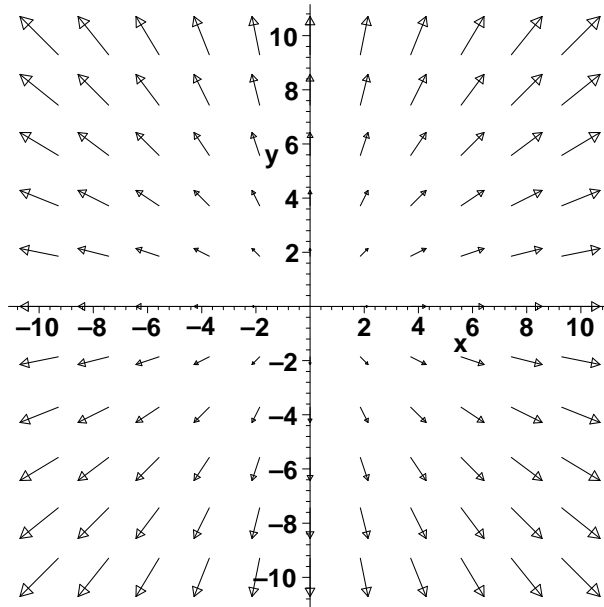
Man nennt die Menge der Vektoren $\vec{\nabla} f$ auch ein *Gradientenfeld*. Schauen wir uns nun die Gradientenfelder für unsere drei Beispielfunktionen an. Für die erste haben wir $\vec{\nabla} f$ soeben berechnet. Wir können das *Vektorfeld* nun graphisch darstellen, indem wir uns in der x - y -Ebene ein diskretes Gitter von Punkten vorgeben und an jedem Gitterpunkt einen Pfeil einzeichnen, wobei der Mittelpunkt des Pfeils auf dem Gitterpunkt liegt. Die Richtung des Pfeils gibt die Richtung von $\vec{\nabla} f$ an der Stelle \vec{x} an, und die Länge ist ein Maß für den Betrag von $\vec{\nabla} f$ an dieser Stelle. Im Beispiel haben wir das oben gezeigte Gebiet $-2 \leq x \leq 2$, $-2 \leq y \leq 2$ in ein 21×21 Gitter eingeteilt, d.h. die Mittelpunkte der Pfeile liegen bei $x_n = n/5$ und $y_n = n/5$ mit $n = -10, -9, \dots, 9, 10$.

$$\vec{\nabla} f(x, y) \doteq \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} = \begin{pmatrix} -2x \exp(-x^2 - y^2) \\ -2y \exp(-x^2 - y^2) \end{pmatrix}$$



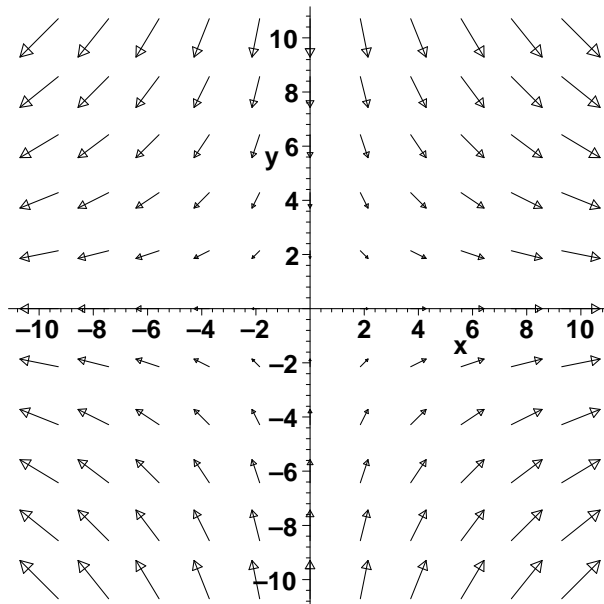
Das Gradientenfeld der zweiten Beispielfunktion $f(x, y) = x^2 + y^2$ ist

$$\vec{\nabla} f(x, y) \doteq \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$



und das der dritten $f(x, y) = x^2 - y^2$

$$\vec{\nabla} f(x, y) \doteq \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x \\ -2y \end{pmatrix}.$$



Die oben diskutierten Eigenschaften der Beispielfunktionen lassen sich auch in den zugehörigen Gradientenfeldern leicht erkennen.

Sind wir daran interessiert, wie groß die Steigung an einem Punkt x, y in eine durch einen Vektor \vec{n} der Länge 1 gegebene Richtung ist, so müssen wir das Skalarprodukt

$$\vec{n} \cdot \vec{\nabla} f(x, y)$$

bilden. Für eine Funktion von n Veränderlichen ist die Steigung im Punkt x_1, x_2, \dots, x_n bei Fortschreiten in Richtung des Vektors \vec{n} , mit $|\vec{n}| = 1$, durch

$$\vec{n} \cdot \vec{\nabla} f(x_1, x_2, \dots, x_n)$$

gegeben.

Ein in diesem Kontext wichtiger Begriff ist der der *Äquipotentiallinie*.³ Das sind Kurven in der x, y -Ebene entlang derer sich $f(x, y)$ nicht ändert. Ist so eine Äquipotentiallinie durch einen vom Ort x, y abhängigen Tangentialvektor an der Kurve $\vec{n}(x, y)$ der Länge 1 gegeben, so gilt damit $\vec{n}(x, y) \cdot \vec{\nabla} f(x, y) = 0$. Für die ersten beiden Beispielfunktionen sind die Äquipotentiallinien *Kreise* um den Ursprung, die durch $x^2 + y^2 = \text{const.}$ gegeben sind, für die dritte sind es *Hyperbeln* mit $x^2 - y^2 = \text{const.}$.

Da sich Gradientenfelder im physikalisch relevanteren dreidimensionalen Raum nicht so schön graphisch darstellen lassen, wollen wir hier nicht weiter auf diesen Fall eingehen. Wir wollen nur anmerken, daß das Konzept des Gradienten in der Physik z.B. bei der Bestimmung eines Kraftfeldes aus einem *Potential* $V(\vec{x})$ (bzw. aus der potentiellen Energie)

$$\vec{F}(\vec{x}) = -\vec{\nabla} V(\vec{x})$$

eine zentrale Rolle spielt. So läßt sich die am Anfang dieses Kapitels diskutierte Gravitationskraft durch Bilden des negativen des Gradienten aus dem Potential

$$V(\vec{x}) = -\gamma \frac{Mm}{|\vec{x}|}$$

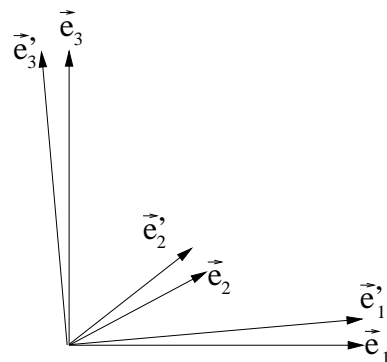
bestimmen. Mit dieser Bemerkung wollen wir das Kapitel beenden und zum Konzept der Matrix übergehen.

³Wir beschränken uns hier auf das Beispiel der Funktionen von zwei Veränderlichen.

Kapitel 11

Matrizen

Wir wollen den Begriff der Matrix hier anhand einer speziellen Klasse von Matrizen - denjenigen die Drehungen beschreiben - einführen. Diese spielen in der Mechanik eine zentrale Rolle. Wir betrachten dazu die Situation, daß wir eine orthonormale Basis $B = [\vec{e}_1, \vec{e}_2, \vec{e}_3]$ auf eine andere $B' = [\vec{e}'_1, \vec{e}'_2, \vec{e}'_3]$ rotieren möchten.



Einen beliebigen Vektor \vec{x} können wir dann in beiden Basen ausdrücken

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 = x'_1\vec{e}'_1 + x'_2\vec{e}'_2 + x'_3\vec{e}'_3. \quad (11.1)$$

Um die x'_i , mit $i = 1, 2, 3$ aus den x_j , mit $j = 1, 2, 3$ zu berechnen, drücken wir zunächst die neuen Basisvektoren \vec{e}'_i durch die alten \vec{e}_j aus

$$\vec{e}'_i = a_{i,1}\vec{e}_1 + a_{i,2}\vec{e}_2 + a_{i,3}\vec{e}_3. \quad (11.2)$$

Die sich ergebenden 3×3 Zahlen $a_{i,j}$ ordnen wir in ein 3×3 Schema, in dem der erste Index die Zeile angibt und der zweite die Spalte

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

welches wir mit \mathbf{A} bezeichnen und *Matrix* nennen. Die $a_{i,j}$ nennen wir *Matrixelemente*. Die Bedeutung der $a_{i,j}$ erkennen wir, indem wir Gl. (11.2) mit \vec{e}_j multiplizieren und Gl. (8.1) verwenden

$$a_{i,j} = \vec{e}'_i \cdot \vec{e}_j.$$

Da die Länge von \vec{e}'_i und \vec{e}_j Eins ist, ergibt sich nach der Definition des Skalarprodukts

$$a_{i,j} = \cos \left(\varphi \left(\vec{e}'_i, \vec{e}_j \right) \right).$$

Die Matrix \mathbf{A} beschreibt also alle Winkel der neuen Basis B' in Bezug auf B . Da $|\vec{e}'_i| = 1$ gilt, folgt aus Gl. (11.2)

$$\vec{e}'_i \cdot \vec{e}'_i = 1 = \sum_{j=1}^3 a_{i,j}^2 \quad (11.3)$$

für $i = 1, 2, 3$. Weiterhin sind die \vec{e}'_i paarweise senkrecht zueinander ($i \neq j$)

$$\vec{e}'_i \cdot \vec{e}'_j = 0 = \sum_{k=1}^3 a_{i,k} a_{j,k}. \quad (11.4)$$

Die Eigenschaften der Gln. (11.3) und (11.4) sind speziell für die eine Rotation darstellende Matrix. Für allgemeine Matrizen sind die *Matrixelemente* $a_{i,j}$ beliebige reelle (oder komplexe) Zahlen. Die Gln. (11.3) und (11.4) kann man zu

$$\sum_{k=1}^3 a_{i,k} a_{j,k} = \delta_{i,j} \quad (11.5)$$

zusammenfassen. Oft schreibt man diese Gleichung einfach als $a_{i,k} a_{j,k} = \delta_{i,j}$ und verwendet die *Einsteinsche Summationskonvention*, die vorschreibt, daß über alle auf einer Seite der Gleichung *doppelt* auftretenden Indizes summiert wird. Um zu sehen, wie man mit Hilfe der Matrix \mathbf{A} die x'_i in Gl. (11.1) aus den x_j berechnen kann, bilden wir auf beiden Seiten der Gleichung das Skalarprodukt mit \vec{e}'_1 . So ergibt sich

$$x'_1 = x_1 \vec{e}'_1 \cdot \vec{e}_1 + x_2 \vec{e}'_1 \cdot \vec{e}_2 + x_3 \vec{e}'_1 \cdot \vec{e}_3 = \sum_{j=1}^3 a_{1,j} x_j.$$

Völlig analog kann man x'_2 und x'_3 bestimmen. Insgesamt gilt

$$x'_i = \sum_{j=1}^3 a_{i,j} x_j.$$

Mit Hilfe der Matrix können wir also die neuen Vektorkomponenten aus den alten berechnen. Eine Matrix beschreibt allgemein eine Abbildung eines Vektors auf einen anderen. Die Klasse von Matrizen mit der Eigenschaft der Gl. (11.5) bezeichnet man als *orthogonale* Matrizen. Sie repräsentieren Drehungen.

Will man zwei Rotationen - gegeben durch die Matrizen \mathbf{A} und \mathbf{B} - (allgemeiner zwei Abbildungen) hintereinander ausführen, so kann man die Matrix \mathbf{C} , die beide Transformationen in einem Schritt vermittelt, wie folgt bestimmen

$$x''_i = \sum_{j=1}^3 b_{i,j} x'_j = \sum_{j=1}^3 b_{i,j} \sum_{k=1}^3 a_{j,k} x_k = \sum_{k=1}^3 c_{i,k} x_k,$$

mit

$$c_{i,k} = \sum_{j=1}^3 b_{i,j} a_{j,k}.$$

Kurz schreibt man $\mathbf{C} = \mathbf{BA}$. Man multipliziert die Matrizen \mathbf{B} und \mathbf{A} miteinander.

Will man von den neuen Koordinaten zurück auf die alten transformieren, so muß gelten

$$c_{i,k} = \sum_{j=1}^3 b_{i,j} a_{j,k} = \delta_{i,k},$$

damit $x''_i = x_i$ folgt. Man bezeichnet \mathbf{B} in diesem Fall als das *Inverse* von \mathbf{A} und schreibt $\mathbf{B} = \mathbf{A}^{-1}$. Mit Hilfe der Gl. (11.5) sieht man, daß sich für orthogonale Matrizen die Matrixelemente $\tilde{a}_{i,j}$ der inversen Matrix \mathbf{A}^{-1} aus den $a_{i,j}$ durch Vertauschen der Zeilen- und Spaltenindizes ergeben $\tilde{a}_{i,j} = a_{j,i}$. Allgemein nennt man die sich aus \mathbf{A} durch Vertauschen der Zeilen- und Spaltenindizes ergebende Matrix die zu \mathbf{A} *transponierte Matrix* und bezeichnet sie mit \mathbf{A}^T . Für orthogonale Matrizen gilt also $\mathbf{A}^T = \mathbf{A}^{-1}$.

Nach dieser Einführung des Matrix-Begriffs anhand eines in der Physik wichtigen Beispiels, wollen wir noch kurz auf ein paar allgemeine Gesichtspunkte eingehen. Allgemein nennen wir ein Schema mit n Zeilen und m Spalten

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \dots & \dots & \dots & \dots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix}$$

mit $a_{i,j} \in \mathbb{R}$ (bzw. $\in \mathbb{C}$) eine $n \times m$ -Matrix. Der erste Index ist der Zeilenindex, der zweite der Spaltenindex. Bei der Multiplikation einer Matrix mit einer reellen bzw. komplexen Zahl λ werden einfach *alle* Matrixelemente $a_{i,j}$ mit λ multipliziert. Für $\mathbf{B} = \lambda \mathbf{A}$ gilt also,

$$b_{i,j} = \lambda a_{i,j} \quad \text{für } i = 1, 2, \dots, n \text{ und } j = 1, 2, \dots, m .$$

Nur für den Fall, daß zwei Matrizen die gleiche Anzahl von Spalten und Zeilen haben kann die Addition zweier Matrizen \mathbf{A} und \mathbf{B} sinnvoll definiert werden. Sie ist komponentenweise definiert, d.h. es gilt für $\mathbf{C} = \mathbf{A} + \mathbf{B}$

$$c_{i,j} = a_{i,j} + b_{i,j} \quad \text{für } i = 1, 2, \dots, n \text{ und } j = 1, 2, \dots, m .$$

Eine Allgemeine $n \times m$ -Matrix beschreibt eine *lineare Abbildung* vom \mathbb{R}^m in den \mathbb{R}^n (bzw. vom \mathbb{C}^m in den \mathbb{C}^n).¹ Eine Abbildung A vom \mathbb{R}^m in den \mathbb{R}^n heißt linear, wenn sie die folgenden Eigenschaften für alle $\vec{x}, \vec{y} \in \mathbb{R}^m$ und $\lambda \in \mathbb{R}$ erfüllt:

$$\begin{aligned} A(\vec{x} + \vec{y}) &= A(\vec{x}) + A(\vec{y}) \\ A(\lambda \vec{x}) &= \lambda A(\vec{x}) . \end{aligned}$$

Um das *Bild* $A(\vec{x})$ eines Vektors $\vec{x} \in \mathbb{R}^m$ zu bestimmen, kann man die Matrix \mathbf{A} auf den Vektor \vec{x} anwenden

$$\vec{y} = \mathbf{A}\vec{x} , \tag{11.6}$$

mit $\vec{y} \in \mathbb{R}^n$. Komponentenweise schreibt man diese Gleichung wie folgt:

$$y_i = \sum_{j=1}^m a_{i,j} x_j \quad \text{für } i = 1, 2, \dots, n .$$

Eine Gleichung dieser Form haben sie wahrscheinlich in der Schule im Zusammenhang mit dem Problem von “ n linearen Gleichungen zur Bestimmung von m Unbekannten x_j , mit $j = 1, 2, \dots, m$ ” kennengelernt. In diesem Kontext ergibt sich das zur Abbildungsidee umgekehrte Problem: Die linken Seiten y_i sind vorgegeben und man möchte die x_j bestimmen, die den Satz von Gleichungen erfüllt. Diese Überlegung führt uns auf das Konzept der Inversen \mathbf{A}^{-1} einer Matrix \mathbf{A} . Bevor wir dieses Konzept allgemein genauer beleuchten, müssen wir die Multiplikation zweier Matrixen definieren. Da es sich bei $n \times m$ -Matrizen um lineare Abbildungen vom \mathbb{R}^m in den \mathbb{R}^n handelt soll die Matrix $\mathbf{C} = \mathbf{B}\mathbf{A}$ der Abbildung C entsprechen, die sich beim Hintereinanderausführen von A (zuerst) und B (als zweites) ergibt. Damit ist unter anderem klar, daß die Multiplikation nur für zwei Matrizen sinnvoll definiert werden kann, für die die Zahl der Zeilen

¹Von jetzt an, werden wir uns auf reelle Matrizen beschränken. Das folgende läßt sich aber leicht auf den Fall komplexer Matrizen verallgemeinern.

der zuerst anzuwendenden Matrix \mathbf{A} gleich der Zahl der Spalten der als zweites anzuwendenden Matrix \mathbf{B} ist. Die Matrix \mathbf{A} beschreibt damit eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^n , die Matrix \mathbf{B} eine vom \mathbb{R}^n in den \mathbb{R}^l , so daß \mathbf{C} eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^l beschreibt. Völlig analog zu den Überlegungen im obigen Beispiel ergibt sich unter Zuhilfenahme von Gl. (11.6) für die Matrixelemente von \mathbf{C}

$$c_{i,k} = \sum_{j=1}^n b_{i,j} a_{j,k} \quad \text{für } i = 1, 2, \dots, l \text{ und } k = 1, 2, \dots, m .$$

Im Gegensatz zur Multiplikation von reellen Zahlen ist damit klar, daß im allgemeinen \mathbf{BA} *nicht* gleich² \mathbf{AB} ist, da letzteres nur für $l = m$ überhaupt definiert ist. Aber selbst, wenn beide Produkte definiert sind (z.B. im physikalisch wichtigen Fall zweier $n \times n$ -Matrizen), gilt im allgemeinen $\mathbf{BA} \neq \mathbf{AB}$! Die Matrixmultiplikation ist im allgemeinen also nicht kommutativ. Ein Beispiel dazu werden sie in den Übungen rechnen. Im Gegensatz dazu gelten für die Matrixmultiplikation das Assoziativgesetz $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ und das Distributivgesetz $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$.

Betrachten wir im folgenden nur noch den Fall von $n \times n$ -Matrizen. Das Inverse \mathbf{A}^{-1} der Matrix \mathbf{A} ist definiert durch

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{1} .$$

Die *Einheitsmatrix* $\mathbf{1}$ hat die Matrixelemente $\delta_{i,j}$. Diese Gleichung erlaubt es \mathbf{A}^{-1} eindeutig zu berechnen, sofern es \mathbf{A}^{-1} überhaupt gibt. Ein einfaches Kriterium um festzustellen, ob \mathbf{A}^{-1} existiert, werden sie sehr bald in der *AGLA I* Vorlesung beweisen. Dort werden sie im Falle der Existenz auch ein systematisches Verfahren zum Bestimmen von \mathbf{A}^{-1} kennenlernen. Wie man leicht nachrechnet hat die Einheitsmatrix $\mathbf{1}$ die Eigenschaft $\mathbf{A}\mathbf{1} = \mathbf{A}$. Die *Nullmatrix* $\mathbf{0}$ ist dadurch definiert, daß alle Matrixelemente Null sind.³ Damit gilt $\mathbf{A}\mathbf{0} = \mathbf{0}$ für alle Matrizen \mathbf{A} . Nehmen wir an, daß das Inverse von \mathbf{A} im oben angesprochenen Beispiel des Problems "Bestimmung von n Unbekannten x_i aus n linearen Gleichungen" (die y_i sind bekannt)

$$y_i = \sum_{j=1}^n a_{i,j} x_j \quad \text{für } i = 1, 2, \dots, n .$$

existiert. In diesem Fall ergeben sich die Unbekannten x_i *eindeutig* aus

$$x_i = \sum_{j=1}^n \tilde{a}_{i,j} y_j \quad \text{für } i = 1, 2, \dots, n ,$$

²Zwei Matrizen sind gleich, wenn alle ihre Matrixelemente gleich sind.

³Durch diese Definition ist auch im allgemeinen Fall der $n \times m$ -Matrix eine Nullmatrix festgelegt.

wobei die $\tilde{a}_{i,j}$ die Matrixelemente von \mathbf{A}^{-1} bezeichnen. Für das Bilden der inversen Matrix gilt⁴

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} ,$$

wie sich aus $\mathbf{X}(\mathbf{AB}) = \mathbf{1}$ durch Multiplikation zuerst mit \mathbf{B}^{-1} und anschließender Multiplikation mit \mathbf{A}^{-1} jeweils von rechts auf beiden Seiten der Gleichung ergibt.

Im Kapitel 12 werden wir das so genannte Eigenwertproblem einer speziellen Klasse von Matrizen untersuchen.

⁴Wir nehmen an, daß alle auftretenden inversen Matrizen auch existieren.

Kapitel 12

Das Eigenwertproblem und die Hauptachsentransformation

In der Mechanik starrer Körper spielen das Eigenwertproblem und die Hauptachsentransformation eine große Rolle. Wir werden das Eigenwertproblem anhand eines einfachen Beispiels einführen und uns hier auf reelle Matrizen, also den \mathbb{R}^n , beschränken. Wir betrachten die (diagonale) 3×3 -Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix} .$$

Sie hat die Eigenschaft, daß

$$\begin{aligned} \mathbf{A} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \\ \mathbf{A} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 4 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} , \\ \mathbf{A} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} , \end{aligned}$$

d.h. die Basisvektoren \vec{e}_j reproduzieren sich bis auf einen Faktor, wenn man \mathbf{A} auf sie anwendet. Allgemein bezeichnet man einen Vektor \vec{v} der sich bei Anwenden einer Matrix \mathbf{A} bis auf einen Faktor $\lambda \in \mathbb{R}$ reproduziert, d.h. für den

$$\mathbf{A}\vec{v} = \lambda\vec{v}$$

gilt, als einen *Eigenvektor* von \mathbf{A} zum *Eigenwert* λ . Also sind die Basisvektoren \vec{e}_j Eigenvektoren der obigen (diagonalen) Matrix \mathbf{A} zu den Eigenwerten 1, 4 und

–2. Es stellen sich dann zwei Fragen: 1) Warum ist es interessant Eigenwerte und Eigenvektoren einer Matrix zu kennen? 2) Wie bestimmt man λ und \vec{v} ? Diese werden wir in diesem letzten Kapitel der Vorlesung beantworten, wobei wir uns primär auf Matrizen auf dem \mathbb{R}^2 und dem \mathbb{R}^3 beschränken werden.

Bevor wir dazu kommen, wollen wir jedoch ein weitere Eigenschaften von Eigenvektoren kennenlernen, eine alternative Notation für das Skalarprodukt einführen und eine Klasse von Matrizen charakterisieren auf deren Eigenwertproblem wir uns dann im folgenden beschränken werden

Nehmen wir an, daß wir einen Eigenvektor \vec{v} von \mathbf{A} zum Eigenwert λ kennen. Dann ist auch $c\vec{v}$, mit $c \in \mathbb{R} \setminus \{0\}$ ein Eigenvektor zum Eigenwert λ .¹ Das sieht man sehr leicht

$$\mathbf{A}(c\vec{v}) = c\mathbf{A}\vec{v} = c\lambda\vec{v} = \lambda(c\vec{v}) .$$

Somit ist ein Eigenvektor nur bis auf einen Faktor eindeutig definiert. Daher ist es sinnvoll einen auf 1 normierten Eigenvektor zum Eigenwert λ einzuführen: $|\vec{v}| = 1$. Von jetzt an gehen wir davon aus, daß die Eigenvektoren die wir betrachten auf 1 normiert sind.

Weiter oben haben wir den Begriff der transponierten Matrix eingeführt

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \dots & \dots & \dots & \dots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} \Rightarrow \mathbf{A}^T = \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n,2} \\ \dots & \dots & \dots & \dots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \dots & \dots & \dots & \dots \\ a_{1,m} & a_{2,m} & \dots & a_{n,m} \end{pmatrix} .$$

Wir werden die wichtige Relation

$$\mathbf{C} = \mathbf{A}\mathbf{B} \Rightarrow \mathbf{C}^T = \mathbf{B}^T\mathbf{A}^T$$

benötigen, die sich wie folgt ergibt:

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^m a_{i,k}b_{k,j} \\ \Rightarrow \tilde{c}_{i,j} &= c_{j,i} = \sum_{k=1}^m a_{j,k}b_{k,i} = \sum_{k=1}^m b_{k,i}a_{j,k} = \sum_{k=1}^m \tilde{b}_{i,k}\tilde{a}_{k,j} , \end{aligned}$$

wobei die mit einer Schlange bezeichneten Matrixelemente, diejenigen der transponierten Matrix sind.

¹Zwar kann 0 ein Eigenwert einer Matrix sein, jedoch niemals der Nullvektor ein Eigenvektor. Formal hat der Nullvektor immer die obige definierende Eigenschaft eines Eigenvektors, jedoch wäre der Eigenwert völlig unspezifisch da jede reelle Zahl μ , die Relation $\mathbf{A}\vec{0} = \mu\vec{0}$ erfüllt.

Eine Matrix mit der Eigenschaft $\mathbf{A} = \mathbf{A}^T$ bezeichnet man als *symmetrische Matrix*. Beispiele sind die Matrizen

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 3 & \frac{1}{7} & \pi \\ -1 & \pi & 0 \end{pmatrix}$$

und

$$\mathbf{B} = \begin{pmatrix} \cos(\varphi) & \sin(\theta) \\ \sin(\theta) & \sin(\varphi) \end{pmatrix}.$$

Wir werden hier *nur das Eigenwertproblem symmetrischer Matrizen betrachten*. Der Trägheitstensor eines starren Körpers im dreidimensionalen Raum ist eine solche symmetrische 3×3 -Matrix. Es ist wichtig zu erkennen, daß nur quadratische $n \times n$ -Matrizen symmetrisch sein können.

Um später einfacher vorgehen zu können, wollen wir eine alternative Notation einführen und das Skalarprodukt zweier Vektoren als Matrixmultiplikation eines transponierten Vektors und eines Vektors schreiben. Mit

$$\vec{x} \doteq \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \Rightarrow \vec{x}^T \doteq (x_1, x_2, x_3).$$

Damit können wir schreiben

$$\vec{x} \cdot \vec{y} = x_1 y_1 + \dots + x_3 y_3 \doteq (x_1, x_2, x_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \doteq \vec{x}^T \vec{y}.$$

Es gilt ebenfalls

$$\begin{aligned} \vec{y} = \mathbf{A}\vec{x} &\Rightarrow \vec{y}^T = (\mathbf{A}\vec{x})^T \doteq (a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3, \dots, a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3) \\ &= (x_1, x_2, x_3) \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{pmatrix} \doteq \vec{x}^T \mathbf{A}^T. \end{aligned}$$

Wir betrachten nun das Eigenwertproblem einer symmetrischen Matrix $\mathbf{A}^T = \mathbf{A}$. Seien $\vec{v}_i \neq \vec{v}_j$ zwei normierte Eigenvektoren zu den Eigenwerten λ_i und λ_j . Damit gilt

$$\vec{v}_i^T \mathbf{A} \vec{v}_j = \vec{v}_i^T \lambda_j \vec{v}_j = \lambda_j (\vec{v}_i \cdot \vec{v}_j).$$

Die linke Seite dieser Gleichung ist aber auch gleich

$$\vec{v}_i^T \mathbf{A} \vec{v}_j = (\mathbf{A} \vec{v}_i)^T \vec{v}_j = (\lambda_i \vec{v}_i)^T \vec{v}_j = \lambda_i (\vec{v}_i \cdot \vec{v}_j).$$

Aus diesen beiden Gleichungen folgt

$$(\lambda_j - \lambda_i) (\vec{v}_i \cdot \vec{v}_j) = 0.$$

Damit gibt es zwei Möglichkeiten:

1. Für $\lambda_j \neq \lambda_i$ folgt daß \vec{v}_j und \vec{v}_i senkrecht aufeinander stehen.
2. Für $\lambda_j = \lambda_i = \lambda$ spannen \vec{v}_i und \vec{v}_j eine zweidimensionale Ebene \mathcal{E} auf. Jeder Vektor $\vec{x} \in \mathcal{E}$ läßt sich als

$$\vec{x} = a\vec{v}_i + b\vec{v}_j; \quad a, b \in \mathbb{R}$$

schreiben. Daraus ergibt sich sofort, daß auch \vec{x} Eigenvektor zum Eigenwert λ ist, da

$$\mathbf{A}\vec{x} = \mathbf{A}(a\vec{v}_i + b\vec{v}_j) = a\mathbf{A}\vec{v}_i + b\mathbf{A}\vec{v}_j = \lambda(a\vec{v}_i + b\vec{v}_j) = \lambda\vec{x}.$$

Somit sind alle Vektoren in \mathcal{E} Eigenvektoren und wir können zwei Vektoren auswählen, die senkrecht aufeinander stehen.

Aus 1. und 2. ergibt sich, daß es zu einer symmetrischen $n \times n$ -Matrix einen Satz auf 1 normierter Eigenvektoren gibt, die paarweise senkrecht aufeinander stehen und die wir jetzt wieder mit \vec{v}_i , $i = 1, 2, \dots, n$, bezeichnen: $|\vec{v}_i| = 1$; $\vec{v}_i \perp \vec{v}_j$ für $i \neq j$. Damit ist $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ eine orthonormale Basis. Bei festgehaltenem Ursprung geht $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ durch eine Rotation aus $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ hervor. Betrachten wir wieder den Spezialfall einer 3×3 -Matrix \mathbf{A} und die dazugehörige Matrix

$$\mathbf{S} = \begin{pmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{pmatrix} = \begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \end{pmatrix}$$

die aus den zueinander senkrechten, normierten Eigenvektoren von \mathbf{A} gebildet ist. Es gilt

$$\mathbf{S}^T = \begin{pmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{pmatrix} = (\vec{v}_1, \vec{v}_2, \vec{v}_3).$$

Nach unseren obigen Überlegungen und denen des letzten Kapitels beschreibt \mathbf{S} eine *Rotation* und ist somit eine orthogonale Matrix mit $\mathbf{S}^T = \mathbf{S}^{-1}$. Wir betrachten nun

$$\begin{aligned} \mathbf{S}\mathbf{A}\mathbf{S}^T &= \begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \end{pmatrix} \mathbf{A}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = \begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \end{pmatrix} (\lambda_1 \vec{v}_1, \lambda_2 \vec{v}_2, \lambda_3 \vec{v}_3) \\ &= \begin{pmatrix} \lambda_1 \vec{v}_1 \cdot \vec{v}_1 & \lambda_2 \vec{v}_1 \cdot \vec{v}_2 & \lambda_3 \vec{v}_1 \cdot \vec{v}_3 \\ \lambda_1 \vec{v}_2 \cdot \vec{v}_1 & \lambda_2 \vec{v}_2 \cdot \vec{v}_2 & \lambda_3 \vec{v}_2 \cdot \vec{v}_3 \\ \lambda_1 \vec{v}_3 \cdot \vec{v}_1 & \lambda_2 \vec{v}_3 \cdot \vec{v}_2 & \lambda_3 \vec{v}_3 \cdot \vec{v}_3 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \end{aligned}$$

Durch das Anwenden von \mathbf{S} von links und \mathbf{S}^T von rechts, haben wir die Matrix \mathbf{A} somit auf Diagonalform gebracht. Die Einträge auf der Diagonalen sind die Eigenwerte. Dies ist die *Hauptachsentransformation*. Wie sie in der Mechanik lernen werden ist es sehr vorteilhaft, z.B. den Trägheitstensor eines starren Körpers mit Hilfe der Hauptachsentransformation auf Diagonalform zu bringen. Dazu benötigen wir natürlich die Matrix S , d.h. die Eigenvektoren der symmetrischen Matrix \mathbf{A} . Wir wollen daher als nächstes beschreiben, wie man die \vec{v}_i und λ_i bestimmt.

Um einen Weg zur Bestimmung von \vec{v}_i und λ_i zu finden betrachten wir die Definitionsgleichung.

$$\begin{aligned} \mathbf{A}\vec{v}_i &= \lambda_i\vec{v}_i = \lambda_i\mathbf{1}\vec{v}_i \\ \Rightarrow (\mathbf{A} - \lambda_i\mathbf{1})\vec{v}_i &= 0. \end{aligned}$$

Die Frage die sich somit zur Bestimmung von λ_i stellt ist: Für welche λ_i gibt es eine nicht-triviale Lösung $\vec{v}_i \neq 0$ der Gleichung in der letzten Zeile?² Wir haben also das Problem vorliegen: $\mathbf{B}\vec{x} = 0$; unter welcher Bedingung an \mathbf{B} gibt es Lösung mit $\vec{x} \neq 0$. Wir führen Vektoren \vec{b}_i ein, so daß

$$\mathbf{B} = \begin{pmatrix} \vec{b}_1^T \\ \vec{b}_2^T \\ \vec{b}_3^T \end{pmatrix},$$

woraus sich für die uns interessierende Gleichung

$$\mathbf{B}\vec{x} = \begin{pmatrix} \vec{b}_1 \cdot \vec{x} \\ \vec{b}_2 \cdot \vec{x} \\ \vec{b}_3 \cdot \vec{x} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ergibt. Die Frage ist also, ob es einen Vektor $\vec{x} \neq 0$ gibt, der senkrecht auf allen drei Vektoren \vec{b}_1 , \vec{b}_2 und \vec{b}_3 steht. Es gibt im \mathbb{R}^3 solch einen Vektor falls die \vec{b}_1 , \vec{b}_2 und \vec{b}_3 in einer Ebene liegen (d.h. koplanar sind). Also gibt es eine nicht-triviale Lösung wenn das durch die drei Vektoren \vec{b}_i aufgespannte dreidimensionale Volumen, d.h. das Spatprodukt der drei Vektoren \vec{b}_i , verschwindet:³

$$V(\vec{b}_1, \vec{b}_2, \vec{b}_3) = (\vec{b}_1 \times \vec{b}_2) \cdot \vec{b}_3 = 0.$$

In diesem Kontext nennt man $V(\vec{b}_1, \vec{b}_2, \vec{b}_3)$ die *Determinante* der 3×3 -Matrix \mathbf{B} , die man mit $\det \mathbf{B}$ bezeichnet.

Die Determinante berechnet sich dann wie folgt:

$$\begin{aligned} \det \mathbf{B} &= (\vec{b}_1 \times \vec{b}_2) \cdot \vec{b}_3 = \begin{pmatrix} b_{1,2}b_{2,3} - b_{1,3}b_{2,2} \\ b_{1,3}b_{2,1} - b_{1,1}b_{2,3} \\ b_{1,1}b_{2,2} - b_{1,2}b_{2,1} \end{pmatrix} \cdot \begin{pmatrix} b_{3,1} \\ b_{3,2} \\ b_{3,3} \end{pmatrix} \\ &= b_{1,2}b_{2,3}b_{3,1} - b_{1,3}b_{2,2}b_{3,1} + b_{1,3}b_{2,1}b_{3,2} \\ &\quad - b_{1,1}b_{2,3}b_{3,2} + b_{1,1}b_{2,2}b_{3,3} - b_{1,2}b_{2,1}b_{3,3}. \end{aligned}$$

²Zur Erinnerung: Der Nullvektor kann nicht Eigenvektor sein.

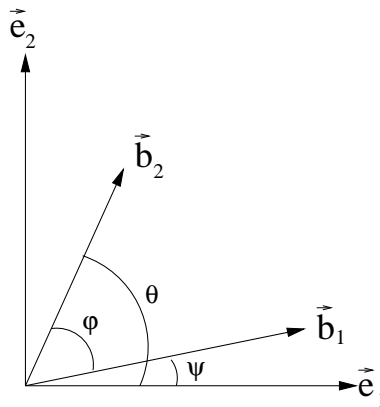
³Eine zweidimensionale Ebene im dreidimensionalen Raum hat Volumen 0.

Es gibt für diesen etwas unübersichtlichen Ausdruck eine einfache Merkregel, die in der Vorlesung angegeben wird.

Eine völlig analoge Argumentation kann für 2×2 -Matrizen geführt werden. Die Determinante einer 2×2 -Matrix ist das durch die Vektoren \vec{b}_1 und \vec{b}_2 aufgespannte zweidimensionale Volumen (also die Fläche), die sich gemäß

$$V(\vec{b}_1, \vec{b}_2) = |\vec{b}_1| |\vec{b}_2| \sin(\varphi) ,$$

mit dem von \vec{b}_1 und \vec{b}_2 eingeschlossenen Winkel φ , berechnet. Mit Hilfe der folgenden Skizze können wir diesen Ausdruck umformulieren, so daß sich die Determinante aus den Matrixelementen ergibt.



Es gilt

$$\begin{aligned} \det \mathbf{B} &= V(\vec{b}_1, \vec{b}_2) = |\vec{b}_1| |\vec{b}_2| \sin(\theta - \psi) = |\vec{b}_1| |\vec{b}_2| [\sin(\theta) \cos(\psi) - \cos(\theta) \sin(\psi)] \\ &= b_{1,1} b_{2,2} - b_{1,2} b_{2,1} . \end{aligned}$$

Wie man die Determinante für eine allgemeine $n \times n$ -Matrix definiert werden sie in der Vorlesung *AGLA I* kennen lernen.

Bezüglich der Bestimmung der Eigenwerte zu einer gegebenen Matrix \mathbf{A} stellt sich also die Frage, für welche λ

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

gilt. Nach den obigen Regeln zur Bestimmung der Determinanten von 2×2 - und 3×3 -Matrizen ist klar, daß es sich bei der linken Seite dieser Gleichung um ein Polynom zweiten bzw. dritten Grades in λ handelt. Man nennt es auch das *charakteristische Polynom*. Die Eigenwerte sind somit die *Nullstellen* des charakteristischen Polynoms.

Betrachten wir nun abschließend ein Beispiel was auch zeigt, wie man bei bekanntem Eigenwert einen Eigenvektor bestimmt. Weitere Beispiele werden sie in den Übungen rechnen. Wir betrachten die symmetrische 2×2 -Matrix

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

Das charakteristische Polynom ergibt sich zu

$$\det(\mathbf{A} - \lambda \mathbf{1}) = \det \begin{pmatrix} 1 - \lambda & -1 \\ -1 & 3 - \lambda \end{pmatrix} = (1 - \lambda)(3 - \lambda) - 1 = \lambda^2 - 4\lambda + 2,$$

mit den Nullstellen

$$\lambda_{1/2} = 2 \pm \sqrt{2}.$$

Die Matrix \mathbf{A} hat somit die Eigenwerte $\lambda_{1/2} = 2 \pm \sqrt{2}$. Der Eigenvektor \vec{v}_1 zu λ_1 ergibt sich als der Vektor, der die Gleichung

$$\begin{aligned} (\mathbf{A} - \lambda_1 \mathbf{1}) \vec{v}_1 &= 0 \\ \Rightarrow \begin{pmatrix} 1 - 2 - \sqrt{2} & -1 \\ -1 & 3 - 2 - \sqrt{2} \end{pmatrix} \begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

erfüllt. Die erste Komponente dieser Gleichung liefert

$$\begin{aligned} (-1 - \sqrt{2})v_{1,1} - v_{1,2} &= 0 \\ \Rightarrow v_{1,2} &= -(1 + \sqrt{2})v_{1,1}. \end{aligned}$$

Die sich aus der zweiten Komponente ergebende Gleichung liefert die gleiche Bedingung, was man daran sieht, daß sie mit der ersten Gleichung konsistent ist

$$-v_{1,1} + (1 - \sqrt{2})v_{1,2} = -v_{1,1} - (1 - \sqrt{2})(1 + \sqrt{2})v_{1,1} = -v_{1,1} + v_{1,1} = 0.$$

Also folgt für den Eigenvektor zum Eigenwert λ_1 , daß er die Form

$$\vec{v}_1 = \begin{pmatrix} x \\ -(1 + \sqrt{2})x \end{pmatrix},$$

mit $x \in \mathbb{R}$ hat. Das x bestimmen wir nun so, daß \vec{v}_1 auf 1 normiert ist

$$|\vec{v}_1| = \sqrt{x^2 + (1 + \sqrt{2})^2 x^2} = |x| \sqrt{4 + 2\sqrt{2}}$$

und damit

$$x = \frac{1}{\sqrt{4 + 2\sqrt{2}}},$$

wobei wir uns für das positive x entschieden haben. In den Übungen werden sie dieses Beispiel vervollständigen, d.h. \vec{v}_2 berechnen und explizit zeigen, daß die bestimmten λ_i und \vec{v}_i , mit $i = 1, 2$, die Definitionsgleichung für Eigenwerte und Eigenvektoren erfüllen.