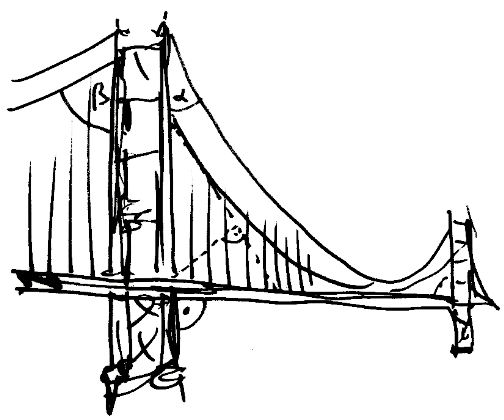
Skript zum Brückenkurs

Gehalten von Jens Eisert an der Freien Universität Berlin

Wintersemester 2011/2012

Version des Skriptes: 0.9



Inhaltsverzeichnis

Kapitel 1

Einleitung

1.0.1 Was der Kurs leisten soll und was nicht

Willkommen in der Welt der Naturwissenschaften und der Mathematik! Ohne kitschig klingen zu wollen, wird hier sicher eine Reise beginnen, die gewissermaßen in faszinierendes Neuland führt. Entgegen mancher landläufiger Meinung haben weder die Naturwissenschaften noch die Mathematik allzuviel mit sturen Formeln und sprödem Formalismus zu tun, sondern viel mit Kreativität, Ideen, und auch in einem Sinn Schönheit. Auch wenn die Endprodukte sehr formalisiert aussehen, stehen doch fast immer intuitive Einsichten und Querdenkereien am Anfang, und eine Menge interessante Tüftelei dazwischen.

Im Schlechten werden wir davon in diesem Kurs nicht allzuviel kennenlernen, denn wir sind hier noch gewissermaßen "vor" dem Studium. Der Kurs soll vor allem drei Dinge tun und leisten:

- Es gibt kein Leugnen, dass gerade die ersten Semester des Studiums die eigene Zeitplanung recht stark beanspruchen werden (aber keine Angst!). Es besteht auch ein wenig das Dilemma, dass die begleitenden Mathematikvorlesungen die Dinge "richtig" machen, während das Hauptfach voranprescht und viele Methoden als bekannt voraussetzt. Diese Methoden mögen bekannt sein oder auch nicht, und das kann auch viel damit zusammenhängen, in welchem Bundesland man Abitur gemacht hat oder ob man einen Leistungskurs in der Mathematik besucht hat oder auch nicht. Der Kurs soll die wichtigsten Methoden zusammenfassen, diese etwas weiter treiben, und alle in einer recht kohärenten Weise in Schwung bringen.

- Ein ganz klein wenig sollte, im Guten, durchschimmern, vor allem am Schluss, warum das alles so furchtbar spannend ist. Wenn sich das erst einmal nicht so anfühlen sollte: Keine Sorge, das wird schon (und das kann vor allem auch an mir liegen).

- Der Kurs soll Spaß machen. So ernst ist also alles nicht. Daran sollte nicht zuletzt auch die Übung ihren Teil haben, denn erst in der Übung kann man das Material wirklich verstehen. Vielleicht lernt man in der Übung ja auch ein paar nette und interessante Leute kennen.

Apropos Übung: Der Übungsbetrieb wird von ganz hochkarätigen Postdocs, Doktoranden und -innen, und Studierenden geleitet (vor allem aus meiner Gruppe). Dies sind alles fantastische Leute, und die Übungszettel sind (von ihnen) wohlüberlegt. Wir haben uns auch eng abgesprochen – aber auch nicht zu eng. Es kann sein, dass die Übungszettel nicht bis aufs Komma mit dem Skript übereinstimmen (den ich geschrieben habe, ein klein wenig hier und da von netten Kollegen klauend). Wer da eine gewisse Inkohärenz sieht, dem sei gesagt: Auch das ist Uni! An einer Universität arbeiten sehr engagierte Leute, die in ihrem Fach aufgehen. Aber Eigeninitiative zählt oft mehr als in der Schule, Dinge sind etwas weniger "verschulter" vorbereitet, eigene Ideen zählen mehr. Das sollte man zunächst einmal als eine gute Nachricht auffassen. Falls etwas zu inkohärent sein sollte, sind wir aber für Kritik sehr offen.

1.0.2 Eine kleine Gebrauchsanweisung

Hier eine kleine Gebrauchsanweisung, wie der Skript zu lesen ist. Erst einmal ist er zu lesen (das eine oder andere wissenswerte steht schon drin).

> Solche Kästen enthalten Dinge, die sehr wichtig sind: Die Aussagen, die man nach zehn Minuten Nachtschlaf oder zwei Bier noch selbstsicher vortragen können sollte.

> Beispiele sehen so aus. Wir werden uns viele Beispiele ansehen, denn erst in ihnen werden viele Dinge klarwerden.

Einige interessante Dinge, die aber für das Verständnis so sehr wichtig nicht sind, sind grau markiert.

1.0.3 Literaturliste

Es gibt jede Menge Bücher zu Brückenkursen über mathematische Methoden. Die Regel gilt, dass jedes dieser Bücher seinen Dienst tut. Einige Dinge sind so kanonisch, dass viele Bücher diesbezüglich sehr ähnliche Darstellungen wählen werden. Hier aber zwei Beispiele:

- K. Bosch, *Brückenkurs Mathematik* (Oldenbourg, 2004).
- G. Walz, *Brückenkurs Mathematik* (Spektrum, 2005).

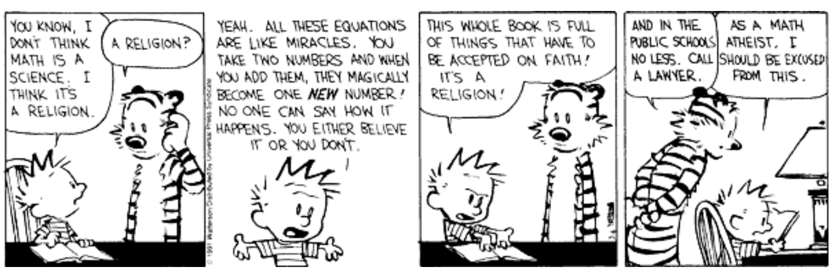
Vor allem ersteres scheint empfehlenswert. Wer ein wenig "vorlesen" möchte fürs Studium, dem oder der seien die folgenden Bücher empfohlen:

- M. Barner, F. Flohr, *Analysis I* (de Gruyter, 1991).
- R. Wüst, *Mathematik für Physiker und Mathematiker 1* (Wiley-VCH, 2009).
- O. Forster, *Analysis 1* (Vieweg und Teubner, 2008).

Kapitel 2

Rechenregeln und Funktionen

2.1 Rechenregeln



©Bill Watterson

Wir beginnen den Kurs mit einer Auffrischung einiger elementarer Rechenregeln, wie dem Bruchrechnen, der Potenzrechnung und den binomischen Formeln. Diese Dinge sind sicher aus der Schule bekannt, aber ein kurzer Blick auf diese Dinge kann nicht schaden. Mit der Beweisidee der vollständigen Induktion werden wir vielleicht auch eine neue Idee kennenlernen.

2.1.1 Bruch- und Potenzrechnung

Für die *Bruchrechnung* gelten die folgenden Regeln.

Bruchrechnung:

$$\frac{x}{y}+\frac{u}{v} = \frac{xv+uy}{yv}, \tag{2.1}$$

$$\frac{x}{y}\frac{u}{v} = \frac{xu}{yv}, \tag{2.2}$$

$$\frac{\frac{x}{y}}{\frac{u}{v}} = \frac{x}{y}\frac{v}{u} = \frac{xv}{yu}. \tag{2.3}$$

Für die *Potenzrechnung* sind die folgenden Regeln hilfreich.

Potenzrechnung:

$$x^{a+b} = x^a x^b, \tag{2.4}$$

$$(xy)^a = x^a y^a, \tag{2.5}$$

$$(x^a)^b = x^{ab}, \tag{2.6}$$

$$x^{1/n} = = \sqrt[n]{x}, \tag{2.7}$$

$$x^{-a} = \frac{1}{x^a}, \tag{2.8}$$

$$x^0 = 1. \tag{2.9}$$

Natürlich muss $x \neq 0$ gelten, wenn wir durch x teilen wollen; aber wir wollen die Notation an dieser Stelle nicht unnötig kompliziert machen. Überhaupt sind die obigen Aussagen stets gemeint als "für alle $x, y, u, v \in \mathbb{R}$", und eigentlich sollte man dies stets dazuschreiben. Man sollte sich dessen sehr bewusst machen, und in der Tat wird – zum Glück! – im Mathematikkurs ein genauerer Ton herrschen. Hier wollen wir aber pragmatisch sein und uns der Mathematik eher als einem "Handwerk" bedienen.

Aus diesen Regeln ergeben sich unmittelbar eine Reihe anderer Regeln, wie etwa für $c \neq 0$,

$$\left(\frac{x}{y}\right)\left(\frac{u}{v}\right)^{-1} = \frac{xv}{yu}. \tag{2.10}$$

2.1.2 Summen- und Produktzeichen

Es ist sehr nützlich, für Summen und Produkte über viele Elemente ein eigenes Symbol einzuführen. Ohne diese Symbole könnte man bei vielen Summen, die in der Praxis oft vorkommen, kaum die Übersicht wahren.

Summen- und Produktzeichen:

$$a_1 + \cdots + a_n = \sum_{j=1}^{n} a_j, \tag{2.11}$$

$$a_1 \cdot \ \ldots \ \cdot a_n = \prod_{j=1}^{n} a_j. \tag{2.12}$$

Natürlich kann der Index jederzeit umbenannt werden, also

$$\sum_{j=1}^{n} a_j = \sum_{k=1}^{n} a_k, \tag{2.13}$$

und es gilt etwa

$$\sum_{j=1}^{n} a_j + a_{n+1} = \sum_{j=1}^{n+1} a_j, \tag{2.14}$$

$$\sum_{j=1}^{n} a_j + \sum_{j=n+1}^{m} a_j = \sum_{j=1}^{m} a_j. \tag{2.15}$$

Insbesondere solche Summenzeichen werden wir von nun an sehr häufig sehen, ab und an auch die Produktsymbole.

2.1.3 Binomische Formeln

Aus der Schule bekannt sind auch die binomischen Formeln – oder zumindest die ersten drei binomischen Formeln, nicht vielleicht die allgemeine binomische Formel.

Binomische Formeln:

$$(x+y)^2 = x^2 + 2xy + y^2, \tag{2.16}$$

$$(x-y)^2 = x^2 - 2xy + y^2, \tag{2.17}$$

$$(x+y)(x-y) = x^2 - y^2, \tag{2.18}$$

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}. \tag{2.19}$$

Die ersten drei Formeln werden auch als *erste*, *zweite* und *dritte binomische Formel* bezeichnet. Hierbei ist der *Binomialkoeffizient definiert als*

$$\binom{n}{j} = \frac{n!}{j!(n-j)!}, \quad n! = \prod_{k=1}^{n} k. \tag{2.20}$$

Wir wollen uns die Mühe machen, kurz letztere Formel zu beweisen. Wir werden uns in einer eigenen Stunde elementaren Beweisideen widmen – und in dieser Einführung Beweise oft auch einfach weglassen. An dieser Stelle werden wir nur eine einfache Technik verwenden, die *vollständige Induktion* genannt wird. Wir stellen die Idee hier erst kurz vor, und wenden sie dann an, um die allgemeine binomische Formel zu beweisen.

Vollständige Induktion: Um zu zeigen, dass ein Satz für alle natürlichen Zahlen $n \geq m$ gilt, genügt es zu zeigen,

- dass er für $n = m$ gilt und
- dass aus der Gültigkeit des Satzes für eine Zahl $n \geq m$ auch die Gültigkeit für $n + 1$ folgt.

Den Schluss von n zu $n + 1$ wird auch als *Induktionsschritt* bezeichnet, der erste Schritt als *Induktionsanfang*.

Der Induktionsanfang ist für die binomische Formel für $n = 1$ offensichtlich: Wir wissen ja schon, dass

$$x + y = \binom{1}{0} x^1 y^0 + \binom{1}{1} x^0 y^1 \tag{2.21}$$

gilt. Etwas komplizierter ist der Induktionsschritt. Wir nehmen also an, dass die Formel (2.19) für n bereits gilt. Nun wollen wir schliessen, dass diese Formel auch für $n + 1$ richtig ist. Wir betrachten also

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n (x+y) \\
&= \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j} (x+y) \\
&= \sum_{j=0}^{n} \binom{n}{j} \left(x^{j+1} y^{n-j} + x^j y^{n-j+1} \right) \\
&= \sum_{j=1}^{n+1} \binom{n}{j-1} x^j y^{n-j+1} + \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j+1}.
\end{aligned} \tag{2.22}$$

Im letzten Schritt haben wir die Variable j um eins verschoben. Also erhalten wir

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n (x+y) \qquad (2.23) \\
&= \binom{n}{n} x^{n+1} + \binom{n}{0} y^{n+1} + \sum_{j=1}^{n} \left(\binom{n}{j-1} + \binom{n}{j} \right) x^j y^{n+1-j}.
\end{aligned}$$

An diesem Punkt verwenden wir die folgende nützliche Identität

$$\begin{aligned}\binom{n}{j-1}+\binom{n}{j} &= \frac{n!}{(j-1)!(n-j+1)!}+\frac{n!}{j!(n-j)!}\\ &= \frac{jn!+(n-j+1)n!}{j!(n-j+1)!}\\ &= \frac{(n+1)n!}{j!(n+1-j)!}=\frac{(n+1)!}{j!(n+1-j)!}\\ &= \binom{n+1}{j}. \qquad (2.24)\end{aligned}$$

Also erhalten wir

$$\begin{aligned}(x+y)^{n+1} &= \binom{n+1}{n+1}x^{n+1}+\binom{n+1}{0}y^{n+1}+\sum_{j=1}^{n}\binom{n+1}{j}x^jy^{n+1-j}\\ &= \sum_{j=0}^{n+1}\binom{n}{j}x^jy^{n+1-j}. \qquad (2.25)\end{aligned}$$

Damit ist der Induktionsschritt vollendet, und wir haben gezeigt, dass die allgemeine binomische Formel für alle natürlichen Zahlen n richtig ist.

Übrigens ist es interessant zu sehen, dass die Binomialkoeffizienten

$$\binom{n}{k}=\frac{n!}{k!(n-k)!} \qquad (2.26)$$

für kleine n und k aus folgender Pyramide gewonnen werden können:

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \end{array} \qquad (2.27)$$

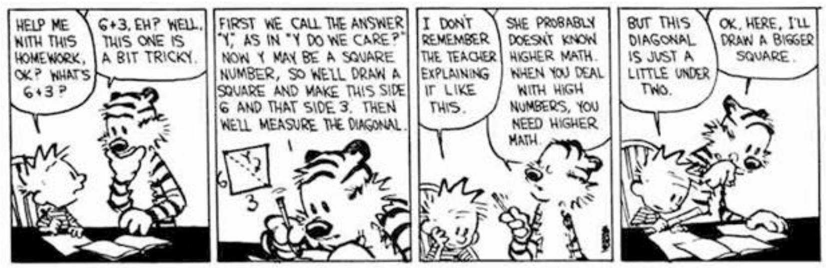
n legt die Zeile fest, beginnend mit $n=0$, j wird innerhalb der Spalten hochgezählt, wiederum mit $j=0$ beginnend. Jeder Eintrag in dieser Pyramide ergibt sich aus der Summe der beiden Zahlen in der vorhergehenden Zeile. Eine interessante Übung ist es, sich zu überlegen, wieso der Zusammenhang zwischen dieser Pyramide und den Binomialkoeffizienten besteht.

2.1.4 Geometrische Reihe

In ähnlicher Weise lässt sich auch die Richtigkeit der Formel für die *geometrische Reihe*

$$\sum_{j=0}^{n}q^k=\frac{1-q^{n+1}}{1-q} \qquad (2.28)$$

beweisen. Diese Aussage werden wir hier aber nicht ausführlich zeigen.



2.1.5 Quadratische Gleichungen

Eine quadratische Gleichung einer Variabler x ist eine Gleichung von der Form

$$x^2 + px + q = 0. \tag{2.29}$$

Hierfür gibt es eine einfache Lösungsformel.

Lösungsformel für quadratische Gleichungen: Die Lösungen von Gl. (2.29) sind gegeben durch

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \tag{2.30}$$

Eigentlich heißt eine solche quadratische Gleichung eine *quadratische Gleichung in Normalform*. Etwas allgemeiner haben wir

$$ax^2 + bx + c = 0, \tag{2.31}$$

mit der Lösungsformel

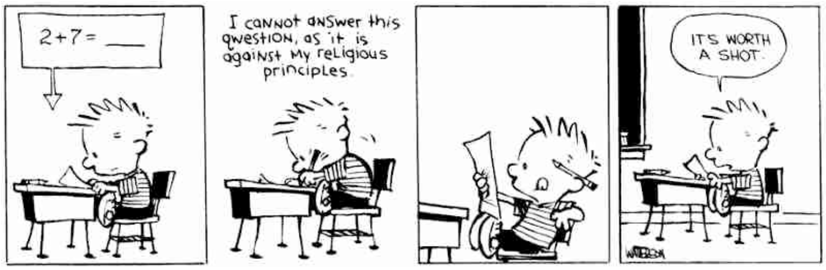
$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{2.32}$$

Diese Formel wird zuweilen auch "Mitternachtsformel" genannt, wohl weil man sie selbst in der finstersten Mitternachtsstunde noch fehlerfrei aufsagen können sollte. Falls $a \neq 0$, können wir aber ohne weiteres durch a dividieren und erhalten so

$$p = \frac{b}{a},\ q = \frac{c}{a}. \tag{2.33}$$

Übrigens gibt es auch für kubische Gleichungen auch noch eine Lösungsformel, die aber deutlich komplizierter aussieht. Für Gleichungen vierter Ordnung auch noch, wenngleich die Form dann nicht mehr sehr schön ist. Für allgemeine Gleichungen höheren Grades kann man zeigen – eine interessante Beobachtung – dass es keine geschlossene Lösungsformel mehr gibt.

2.2 Funktionen



©Bill Watterson

2.2.1 Definition einer Funktion

Die Wichtigkeit des Konzeptes der Funktion kann man nicht überbewerten. Insbesondere ist sie seit Newton aus der Physik nicht mehr wegzudenken. Funktionen sind Abbildungen: Gegeben seien zwei Mengen A und B, dann sei jedem Element x aus A ein Element $f(x)$ aus B zugeordnet. Man sagt, dass damit eine Abbildung oder eine *Funktion* f von A nach B gegeben ist. In der Mathematik schreibt man auch

$$f : A \to B, \tag{2.34}$$

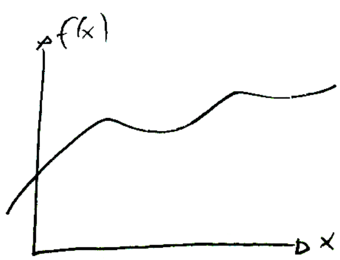
und

$$x \mapsto f(x) \tag{2.35}$$

für ein $x \in A$, oder auch

$$y = f(x). \tag{2.36}$$

Wir werden alle diese Notationen verwenden.



A heißt dann die *Definitionsmenge*, B die *Zielmenge* von f. $f(x)$ ist das *Bild* von x – oder das *Bild von x unter f*, wenn man ganz sicher sein will. Etwas pedantisch ist auch f die Funktion und $f(x)$ ein Element aus B nämlich das Bild von x, aber diesen kleinen Lapsus an Notation sieht man in der (nicht-mathematischen) Literatur recht oft. Wir werden aber versuchen, in diesem Skript pragmatisch zu bleiben und die Notation so wenig "mathematisiert" wie möglich zu halten – eine gründliche Einführung wird dann der erste Mathematikkurs geben. Also stürzen wir uns gleich in die Anwendungen. Wir beginnen mit einer Reihe von einfachen Funktionen. Die einfachste Funktion, die kein eigenes Kapitel verdient, ist die *identische Abbildung*: id. Sie bildet jedes Element auf sich selbst ab:

$$\mathrm{id} : x \mapsto x. \tag{2.37}$$

2.2.2 Absoluter Betrag und Quadratwurzel

Die Betragsfunktion $x \mapsto |x|$ wird uns oft begegnen.

Absoluter Betrag:

$$|x| = \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases} \tag{2.38}$$

Es gilt

$$\begin{aligned} |x| &= 0 \Leftrightarrow x = 0, && (2.39) \\ |xy| &= |x|\,|y|, && (2.40) \\ |x+y| &\leq |x| + |y|. && (2.41) \end{aligned}$$

Letztere Ungleichung wird *Dreiecksungleichung* genannt. Aus der Dreiecksungleichung folgen auch andere nützliche Regeln, etwa gilt stets

$$||x| - |y|| \leq |x + y|. \tag{2.42}$$

Die *Quadratwurzelfunktion* $x \mapsto \sqrt{x}$ ist nur für positive x erklärt.

2.2.3 Exponentialfunktion

Eine der wichtigsten Funktionen ist die *Exponentialfunktion*. Sie wird von nun an unser vertrauter Begleiter werden. An dieser Stelle können wir nur betonen, wie zentral die Exponentialfunktion in so vielen Gebieten der Mathematik, den Natur- und Wirtschaftswissenschaften ist – wir werden aber zunehmend sehen, warum dies so ist. Offensichtlich sind aber Begriffe wie "exponentielles Wachstum" in der Wirtschaftswissenschaft ubiquitär, die Exponentialfunktion kommt in der Gaußformel vor oder in der Eulerformel.

Hier wollen wir aber erst einmal pragmatisch bleiben und uns mit den wichtigsten Eigenschaften der Exponentialfunktion

$$x \mapsto e^x \tag{2.43}$$

beschäftigen. Diese Funktion enthält die reelle Zahl

$$e = 2.7182818\ldots. \tag{2.44}$$

Die Punkte sollen andeuten, dass mit diesen Stellen die Zahl keineswegs exakt definiert ist, sondern man unendlich viele Stellen verschieden von Null braucht, um sie im Dezimalsystem darzustellen. Eine Art, die Zahl zu definieren, besteht in der folgenden Formel:

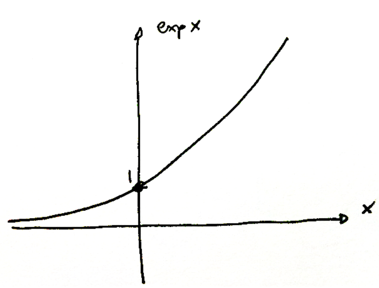
$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n . \tag{2.45}$$

Die Zahl e spielt eine Rolle in der Mathematik die in ihrer Wichtigkeit mit der von π vergleichbar ist (in der Tat enthält die Eulerformel beide diese Konstanten). Eine gute Aufgabe ist es, den Ausdruck in Gl. (2.45) mit Hilfe der binomischen Formeln auszudücken: Schon wenige Terme liefern dann eine recht genaue Approximation der Zahl e.

Eine wichtige Rechenegel für die Exponentialfunktion kennen wir schon aus der Potenzrechnung: In der Tat gilt

$$e^{a+b} = e^a e^b . \tag{2.46}$$

Wir werden uns nun ein kurzes erstes Beispiel ansehen, das dokumentiert, warum die Exponentialfunktion so wichtig ist: Prozesse, in denen das Wachstum einer Größe zu einer Zeit linear von dieser Größe abhängt, zeigen exponentielles Wachstum. Ein eindrucksvoller Prozess dieser Art zeigt sich in der Lösung des *Zinseszins-Problems*:



Gegeben sei ein Anfangskapital K_0. Wenn p Prozent Zinsen ausgezahlt werden, also der Geldwert nach jedem Jahr mit $a = p/100$ multipliziert und zusätzlich ausgezahlt wird, erhält man nach n Jahren also den Betrag

$$K_n = K_0 \prod_{j=1}^{k} (1+a) = K_0(1+a)^n. \tag{2.47}$$

Wenn man nun aber n Mal im Jahr p/n Prozent Zinsen ausgezahlt bekommt, beträgt das Kapital nach m Jahren

$$K_m = K_0 \left(1 + \frac{a}{n}\right)^{nm}. \tag{2.48}$$

Im Grenzfall, in dem man "kontinuierlich" Zinsen ausbezahlt bekommt, erhält man

$$\begin{aligned} K_m &= K_0 \lim_{n\to\infty} \left(1+\frac{a}{n}\right)^{nm} \\ &= K_0 \lim_{n\to\infty} \left(1+\frac{am}{n}\right)^{n} \\ &= K_0 e^{am}. \end{aligned} \tag{2.49}$$

Wir haben es also mit exponentiellem Wachstum zu tun. Nicht vergessen sollten wir allerdings, dass andere Prozesse, wie etwa die Inflation, auch gleichzeitig exponentiell in der Zeit verlaufen. Auch Ideen des "Wirtschaftswachstums" beziehen sich auf solche exponentiellen Prozesse.

Hilfreich sind auch die folgenden unteren und oberen Schranken:

$$1 + x \leq e^x, \; x \in \mathbb{R}, \tag{2.50}$$

$$e^x \leq \frac{1}{1-x}, \; x < 1. \tag{2.51}$$

Wir werden bald sehen, wie man solche Schranken zeigen kann.

2.2.4 Umkehrfunktion

Bevor wir uns weitere konkrete Funktionen ansehen, werden wir uns kurz mit dem Konzept der *Umkehrfunktion* vertraut machen. Zunächst einmal bezeichnet

$$f^{-1}(C) \subset A \tag{2.52}$$

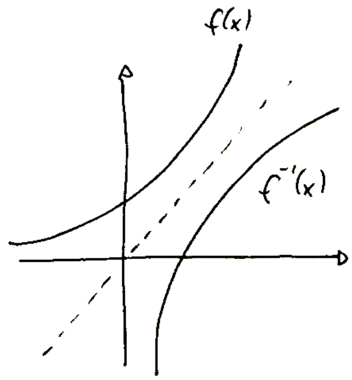
einer Teilmenge $C \subset B$ das *Urbild* von C unter f. Dies ist die Menge aller $x \in A$, für die $f(x) \in C$ liegt.

Dies heißt aber noch nicht, dass es eine Abbildung f^{-1} von B nach A gibt: Denn durchaus könnten selbst in dem Fall, in dem B nur ein einziges Element enthält, zu $f^{-1}(D)$ viele Elemente in A gehören. Eine solche Konstruktion macht erst dann Sinn, wenn zu jedem Element in A auch nur eines in B

und umgekehrt eines in B auch eines in A gehört. Solche Funktionen heißen injektiv. Genauer heißt eine Funktion *injektiv*, wenn aus

$$f(x) = f(y) \tag{2.53}$$

folgt, dass $x = y$. Für solche injektiven Funktionen kann man die *Umkehrabbildung* f^{-1} erklären. Diese hat eine anschauliche Interpretation: Da sie die Elemente "umgekehrt" abbildet, sieht ihr Graph so aus, als hätte man ihn an der Hauptdiagonalen gespiegelt. Die Idee der Umkehrfunktion mag etwas abstrakt klingen, es ist aber ein sehr nützliches Konzept.



2.2.5 Logarithmusfunktion

Eine wichtige Funktion ist die *Logarithmusfunktion*, die x auf $\log(x)$ abbildet. Die Rechenregeln sind wie folgt:

Rechenregeln des Logarithmus:

$$\log(xy) = \log(x) + \log(y), \tag{2.54}$$

$$\log(x^a) = a\log(x), \tag{2.55}$$

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y). \tag{2.56}$$

In der Tat ist die Logarithmusfunktion die Umkehrfunktion der e-Funktion – es ist auch einfach zu sehen, dass die Exponentialfunktion injektiv ist. Also folgt aus

$$y = e^x, \tag{2.57}$$

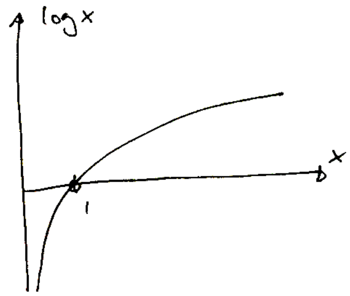
dass

$$x = \log y. \tag{2.58}$$

Wir wollen auch den *Logarithmus zu einer Basis* a kennenlernen: Dieser ist gegeben durch

$$\log_a(x) = \frac{\log(x)}{\log(a)}. \tag{2.59}$$

Dieser ist die Umkehrfunktion von $x \mapsto a^x$.



2.2.6 Verkettung von Funktionen

Eine *Verkettung von Funktionen* ist nichts anderes als eine Hintereinanderausführung. Wir ersparen uns hier das Kleingedruckte über Definitionsbereiche und stellen einfach fest, dass für Funktionen f und g durch $f \circ g$, definiert durch

$$x \mapsto f(g(x)), \tag{2.60}$$

wieder eine Funktion gegeben ist. Eine wichtige Eigenschaft der Verkettung ist die *Assoziativität*: Für beliebige Funktionen gilt

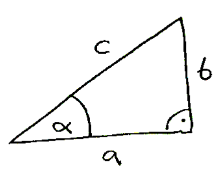
$$f \circ (g \circ h) = (f \circ g) \circ h. \tag{2.61}$$

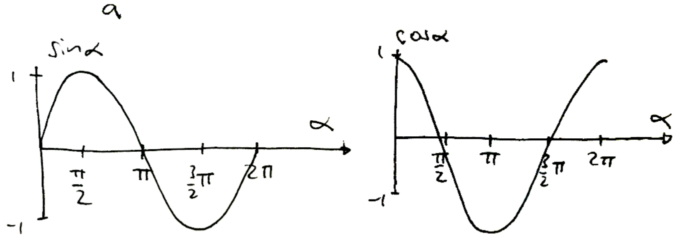
Natürlich gilt auch, wenn man ein wenig aufpasst wegen der Definitionsbereiche,

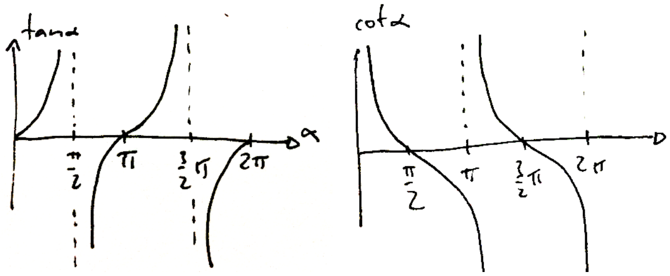
$$f^{-1} \circ f = f \circ f^{-1} = \mathrm{id}. \tag{2.62}$$

2.2.7 Trigonometrische Funktionen

Die trigonometrischen Funktionen führen wir an dieser Stelle am besten durch deren anschaulicher Rolle im rechtwinkligen Dreieck ein.







Trigonometrische Funktionen:

$$\sin\alpha = \frac{b}{c}, \quad (2.63)$$

$$\cos\alpha = \frac{a}{c}, \quad (2.64)$$

$$\tan\alpha = \frac{b}{a} = \frac{\sin\alpha}{\cos\alpha}, \quad (2.65)$$

$$\cot\alpha = \frac{a}{b} = \frac{\cos\alpha}{\sin\alpha}. \quad (2.66)$$

Hierbei werden c als *Hypothenuse*, a als *Ankathete* und b als *Gegenkathete* bezeichnet.

Wie wir wissen, gilt auch in einem rechtwinkigen Dreieck der Zusammenhang $a^2 + b^2 = c^2$. Hieraus folgt sofort der Zusammenhang

$$\left(\frac{a}{c}\right)^2 + \left(\frac{b}{c}\right)^2 = 1, \quad (2.67)$$

also

$$\sin^2\alpha + \cos^2\alpha = 1. \quad (2.68)$$

Es sollte noch darauf hingewiesen werden, dass wir Winkel stets im Bogenmaß auf dem Einheitskreis angeben.

Bogenmaß:

$$\alpha = 0 \leftrightarrow 0, \quad (2.69)$$

$$\alpha = \pi/2 \leftrightarrow 90^\circ, \quad (2.70)$$

$$\alpha = \pi \leftrightarrow 180^\circ, \quad (2.71)$$

$$\alpha = 2\pi \leftrightarrow 360^\circ. \quad (2.72)$$

Somit sind auch $\sin 0 = \sin \pi = \sin(2\pi) = 0$, sowie $\cos 0 = \cos(2\pi) = 1$ und $\cos \pi = -1$.

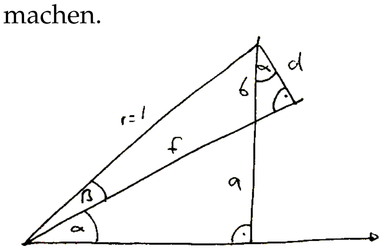
Es gibt eine Vielzahl von Zusammenhängen zwischen trigonometrischen Funktionen – dies liegt an deren enger Verbindung zu Exponentialfunktionen, wie wir sehen werden. Die wohl wichtigsten Identitäten dieser Art werden als *Additionstheoreme* bezeichnet:

Additionstheoreme:

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta, \tag{2.73}$$

$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta. \tag{2.74}$$

Anstatt diese Gleichungen zu beweisen, werden wir sie uns graphisch plausibel machen.



Aus den Additionstheorem folgen auch unmittelbar eine handvoll anderer Relationen, wie etwa

$$\sin 2\alpha = 2\sin\alpha\cos\alpha, \tag{2.75}$$

$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha, \tag{2.76}$$

oder

$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1}{2}(1-\cos\alpha)}, \tag{2.77}$$

$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1}{2}(1+\cos\alpha)}, \tag{2.78}$$

wobei das Vorzeichen je nach Quadrant des Winkels α zu wählen ist.

2.2.8 Hyperbelfunktionen

Die *Hyperbelfunktionen* sind definiert durch

$$\cosh x = \frac{e^x + e^{-x}}{2}, \tag{2.79}$$

$$\sinh x = \frac{e^x - e^{-x}}{2}. \tag{2.80}$$

Diese Funktionen werden als *Cosinus hyperbolicus* und *Sinus hyperbolicus* bezeichnet. Für sie gilt

$$\cosh^2 x - \sinh^2 x = 1, \tag{2.81}$$

da

$$\cosh^2 x - \sinh^2 x = \frac{1}{4}\left((e^x + e^{-x})^2 - (e^x - e^{-x})^2\right) = 1. \tag{2.82}$$

Ein Punkt $(\cosh x, \sinh x)$ lieft also auf der *Einheitshyperbel*

$$\{(x, y) : x^2 - y^2 = 1\}, \tag{2.83}$$

was den Namen der Funktionen nahelegt.

2.2.9 Ein kleines Rätsel von 1778

Was ist eigentlich

$$y = \sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\sqrt{2}\cdots}}} \quad ? \tag{2.84}$$

Diese Frage stellte sich schon Euler im Jahre 1778, also ist dies ein durchaus klassisches Problem.

Hier die Lösung: Zunächst einmal ergibt der Ausdruck präzise geschrieben nur Sinn als den Grenzwert der Folge

$$\sqrt{2},\ \sqrt{2}^{\sqrt{2}},\ \sqrt{2}^{\sqrt{2}^{\sqrt{2}}}, \ldots. \tag{2.85}$$

Wir nennen die Elemente dieser Folge $x_1, x_2, x_3, \ldots$. Per vollständiger Induktion, wie oben, kann man nun zeigen, dass

$$1 < x_j < x_{j+1} \tag{2.86}$$

ist für jedes j, also wachsen die Werte der Folge an (man nennt dies streng monoton steigend): In der Tat ist der Induktionsschritt einfach, denn es ist

$$x_{j+2} = \sqrt{2}^{x_{j+1}} > \sqrt{2}^{x_j} = x_{j+1}. \tag{2.87}$$

Nun wollen wir noch eine *obere Schranke* für alle Folgenglieder finden – also einen Wert, der von keinem Element überschritten wird. Wiederum ist dies einfach: Wenn wir im Ausdruck von einem s_j das oberste $\sqrt{2}$ durch den größeren Wert 2 ersetzen, erhalten wir für den gesamten Ausdruck 2. Also ist 2 eine obere Schranke. Da die Folge ansteigend ist und es eine obere Schranke gibt, muss die Folge konvergieren. Wir nennen den Grenzwert y.

Dieser muss nun

$$\sqrt{2}^y = y \tag{2.88}$$

erfüllen. Ebenso können wir die Lösung von

$$y \log \sqrt{2} = \log y \tag{2.89}$$

suchen, indem wir Schnittpunkte identifizieren. Dies ist einfach: Wir finden zwei Punkte, $y = 2$ und $y = 4$. $y = 4$ kann aber keine Lösung sein, das widerspricht der oberen Schranke. Also ist

$$y = 2. \tag{2.90}$$

So ergibt auch $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$ Sinn, aber nun reicht es aber... :)

Kapitel 3

Differentialrechnung

Wir werden uns heute dem Thema der Diffentialrechnung widmen. Dies wird uns morgen auch noch beschäftigen – und doch werden wir damit nur die Spitze des Eisbergs sehen. Dieses Kapitel wird auch sehr verschieden sein von dem Kapitel über Differentialrechnung im ersten Mathematikkurs: Hier geht es ganz pragmatisch um Regeln des Ableitens von Funktionen. Dass es da eine Menge – oft spannendes – Kleingedrucktes und mathematische Feinsinnigkeiten gibt, soll uns nun nicht allzusehr von unserem wichtigen Vorhaben ablenken. Das heißt natürlich zwingend, dass wir hier keinen "rigorosen" Standpunkt vertreten können. Die wichtigsten Regeln und Ideen sollten aber dennoch klarwerden, ohne dass dadurch der Mathematikkurs langweilig wird.

3.1 Kurvendiskussion und Konzept der Ableitung

3.1.1 Stetige Funktionen

Wir betrachten Stetigkeit von Funktionen $f : A \to B$. Eine Funktion Funktion heißt *stetig*, wenn man für einen beliebig kleinen Wert $\epsilon > 0$ auch ein $\delta > 0$ finden kann, so dass aus

$$|x - y| < \delta \tag{3.1}$$

auch folgt, dass

$$|f(x) - f(y)| < \epsilon. \tag{3.2}$$

Funktionswerte von nahe beieinander liegenden Punkten sind also wiederum nahe. Oder, anders gesagt, der Graph einer stetigen Funktion ist einer, den "man mit dem Stift zeichnen kann, ohne den Stift abzusetzen". Stetige Funktionen sind also solche, die keine Sprünge haben oder machen keine ähnlichen "Dummheiten". Unstetige Funktionen sind etwa die *Sprungfunktion*

$$|x| = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases} \tag{3.3}$$

Am Punkt $x = 0$ "springt" sie, ist also unstetig. Eine weitere unstetige Funktion ist

$$|x| = \begin{cases} 0, & x = 0, \\ \sin \frac{1}{x}, & x \neq 0. \end{cases} \tag{3.4}$$

Stetige Funktionen sind also in einem Sinne "glatte Funktionen".

3.1.2 Begriff der Steigung

Wir betrachten zunächst den Begriff der Steigung von Geraden. In einer linearen Funktion

$$y = ax + b, \tag{3.5}$$

die eine Gerade beschreibt, ist die Steigung a gerade der Vorfaktor des linearen Terms.

Eine naheliegende Anwendung ist die Berechung des Ortes bei einer geradlinigen, gleichförmigen Bewegung. Der Ort ist hier durch

$$s(t) = vt + s(0) \tag{3.6}$$

gegeben, also ist v gerade die *Geschwindigkeit* der Bewegung.

3.1.3 Ableitung

Der Begriff der Steigung ergibt aber auch Sinn für andere Funktionen – sofern sie stetig sind! In Punkten, in denen eine Funktion nicht stetig ist, kann man die Frage der Steigung also gar nicht erst stellen. Wenn aber eine Funktion in einem Punkt stetig ist, und auch eine *Tangente* besitzt , dann kann man die Steigung beliebiger Funktionen in einem Punkt auf die von Geraden zurückführen: Die Steigung der Funktion f im Punkt x is die Steigung der Tangente in diesem Punkt. Diese Steigung wird *Ableitung* von f im Punkt x bezeichnet.

Ableitung:

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \tag{3.7}$$

Höhere Ableitungen sind Ableitungen der Ableitung, wir können also f'' ebenso definieren. Damit die Ableitung Sinn ergibt, muss die Funktion in dem Punkt *differenzierbar* sein, also muss es eine Tangente geben. Dies kann aber nur der Fall sein, wenn eine Funktion stetig in diesem Punkt ist. Differenzierbarkeit ist also eine stärkere Forderung als Stetigkeit.

3.1.4 Elemente der Kurvendiskussion

Nun scheint ein guter Moment, uns Elementen der Kurvendiskussion zuzuwenden.

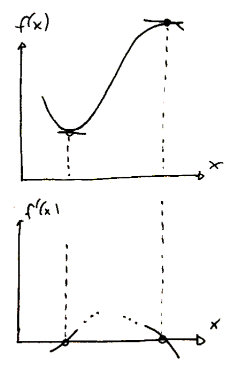
Nullstellen und Singularitäten: *Nullstellen* einer Funktion f sind x, für die

$$f(x) = 0 \tag{3.8}$$

gilt. Eine *Singularität* ist ein Punkt x mit

$$\lim_{y \to x} f(y) = \pm\infty. \tag{3.9}$$

Bei letzteren Grenzwert können Feinheiten der Art des Limes eine Rolle spielen.



Maxima und Minima: Ein *lokales Maximum* einer Funktion ist ein Punkt x, so dass $f(x)$ größer ist als alle anderen Funktionswerte in einem Intervall, das x einschließt. Ebenso kann man ein *lokales Minimum* definieren. Ist eine Funktion differenzierbar, so gilt für ein lokales Maximum sowie für ein Minimum

$$f'(x) = 0. \tag{3.10}$$

Ist sie zweimal differenzierbar, so entspricht

$$f''(x) < 0 \tag{3.11}$$

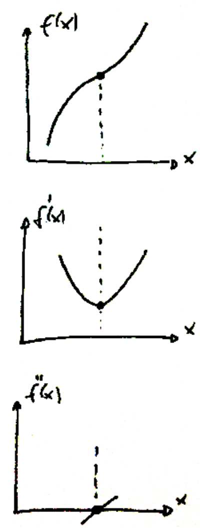
einem lokalen Maximum und

$$f''(x) > 0 \tag{3.12}$$

einem lokalen Maximum. Ein *globales Maximum* einer Funktion ist ein Punkt x, so dass

$$f(y) \leq f(x) \tag{3.13}$$

für alle y aus dem Definitionsbereich gilt.



Wendepunkte: Ein *Wendepunkt* einer zweimal differenzierbaren Funktion ist ein Punkt, für den

$$f''(x) = 0 \tag{3.14}$$

ist.

Schließlich erwähnen wir noch das *Verhalten der Funktion im Unendlichen* : Es ist der Wert von $\lim_{x\to\infty} f(x)$ oder $\lim_{x\to-\infty} f(x)$.

3.2 Anwendungen und Regeln

3.2.1 Ableitungen wichtiger Funktionen

Die einfachste Ableitung ist die einer linearen Funktion, es ist gewissermaßen die Definition selbst: Für

$$f(x) = ax + b \tag{3.15}$$

ist

$$f'(x) = \lim_{\Delta x\to 0} \frac{a(x+\Delta x) + b - (ax+b)}{\Delta x} = \lim_{\Delta x\to 0} a = a. \tag{3.16}$$

Etwas interessantere Beispiele sind die folgenden:

Quadratische Funktion:

$$\begin{aligned} f(x) &= x^2, && (3.17)\\ f'(x) &= \lim_{\Delta x\to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x\to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} \\ &= \lim_{\Delta x\to 0} (2x + \Delta x) = 2x. && (3.18) \end{aligned}$$

Potenz n-ten Grades:

$$\begin{aligned} f(x) &= x^n, && (3.19)\\ f'(x) &= \lim_{\Delta x\to 0} \frac{(x+\Delta x)^a - x^a}{\Delta x} = \lim_{\Delta x\to 0} \sum_{j=1}^{n} \binom{n}{j} (\Delta x)^j x^{n-j} \frac{1}{\Delta x} \\ &= \binom{n}{1} x^{n-1} = nx^{n-1}. && (3.20) \end{aligned}$$

Sinusfunktion:

$$\begin{aligned} f(x) &= \sin x, && (3.21)\\ f'(x) &= \lim_{\Delta x\to 0}\frac{\sin(x+\Delta x)-\sin x}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{\sin(x)\cos(\Delta x)+\cos(x)\sin(\Delta x)-\sin(x)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\left(\cos x\frac{\sin(\Delta x)}{\Delta x}+\sin(x)\frac{\cos(\Delta x)-1}{\Delta x}\right). && (3.22)\end{aligned}$$

Nun wissen wir, dass für kleines x etwa $\sin x \approx x$ gilt. Genauer ist

$$\lim_{\Delta x\to 0}\frac{\sin\Delta x}{\Delta x} = 1. \quad (3.23)$$

Außerdem gilt

$$\lim_{\Delta x\to 0}\frac{\cos(\Delta x)-1}{\Delta x} = \lim_{\Delta x\to 0}\frac{-2\sin^2(\frac{\Delta x}{2})}{\Delta x} = \lim_{\Delta x\to 0} -\frac{(\Delta x)^2}{2\Delta x} = 0. \quad (3.24)$$

Also ist

$$f'(x) = \cos x. \quad (3.25)$$

Cosinusfunktion: In ganz ähnlicher Weise können wir schließen, dass

$$\begin{aligned} f(x) &= \cos x, && (3.26)\\ f'(x) &= -\sin x. && (3.27)\end{aligned}$$

Exponentialfunktion:

$$\begin{aligned} f(x) &= e^x, && (3.28)\\ f'(x) &= \lim_{\Delta x\to 0}\frac{e^{x+\Delta x}-e^x}{\Delta x} = e^x\lim_{\Delta x\to 0}\frac{e^{\Delta x}-1}{\Delta x} = e^x. && (3.29)\end{aligned}$$

Logarithmusfunktion:

$$\begin{aligned} f(x) &= \log x, && (3.30)\\ f'(x) &= \lim_{\Delta x\to 0}\frac{\log(x+\Delta x)-\log x}{\Delta x} \\ &= \lim_{\Delta x\to 0}\frac{\log(1+\frac{\Delta x}{x})}{\Delta x} = \lim_{\Delta x\to 0}\frac{\frac{\Delta x}{x}}{\Delta x} = \frac{1}{x}, && (3.31)\end{aligned}$$

wobei x natürlich nicht 0 sein darf.

3.2.2 Ableitungsregeln

Diese Regeln muß man bald nicht im einzelnen herleiten, sondern man lernt sie gut kennen. Außerdem gibt es Tabellen, die einem die Arbeit abnehmen. Keine noch so umfangreiche Tabelle kann aber alle Ableitungen von sinnvollen Funktionen zusammenfassen. Also braucht man Regeln, um mit allgemeinen Ausdrücken mit handhabbarem Aufwand fertigzuwerden (sofern man nicht gerade Mathematica zur Hand hat :)). Es helfen bei Ableitungen zusammengesetzter Funktionen die folgenden Ableitungsregeln.

Summenregel: Die Ableitung einer Summe ist die Summe der Ableitungen: Aus

$$h = f + g, \tag{3.32}$$

folgt

$$h' = f' + g'. \tag{3.33}$$

Diese Regel zu beweisen ist sehr einfach:

$$\begin{aligned} h'(x) &= \lim_{\Delta x\to 0}\frac{h(x+\Delta x)-h(x)}{\Delta x} && (3.34)\\ &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)+g(x+\Delta x)-g(x)}{\Delta x} = f'(x)+g'(x). \end{aligned}$$

Produktregel: Aus

$$h = fg, \tag{3.35}$$

folgt

$$h' = fg' + f'g. \tag{3.36}$$

Wiederum können wir die Regel an dieser Stelle beweisen.

$$
\begin{aligned}
h'(x) &= \lim_{\Delta x\to 0}\frac{h(x+\Delta x)-h(x)}{\Delta x} && (3.37)\\
&= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)g(x+\Delta x)-f(x)g(x)}{\Delta x} && (3.38)\\
&= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)g(x+\Delta x)-f(x+\Delta x)g(x)+f(x+\Delta x)g(x)-f(x)g(x)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\left(f(x+\Delta x)\frac{g(x+\Delta x)-g(x)}{\Delta x}+\frac{f(x+\Delta x)-f(x)}{\Delta x}g(x)\right)\\
&= f(x)g'(x)+f'(x)g(x). && (3.39)
\end{aligned}
$$

Hier lohnt es sich, einige wenige Beispiele anzusehen.

> Für $h(x) = x\cos x$ ist
> $$h'(x) = -x\sin x + \cos x. \tag{3.40}$$

> Und für $h(x) = \sin^2(x)$ ist
> $$h'(x) = \cos x \sin x + \sin x \cos x = 2\sin x \cos x. \tag{3.41}$$

Wir erinnern uns, dass die Verkettung zweiter Funktionen mit dem Symbol $\circ$ geschrieben werden kann. Also ist

$$(f\circ g)(x) = f(g(x)). \tag{3.42}$$

Kettenregel: Aus
$$h = f\circ g, \tag{3.43}$$
folgt nun
$$h' = (f'\circ g)g', \tag{3.44}$$
also
$$h'(x) = f'(g(x))\cdot g'(x). \tag{3.45}$$

Wieder können wir diese Regel direkt unter Verwendung der Definition der

Ableitung beweisen:

$$\begin{aligned} h'(x) &= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x} && (3.46) \\ &= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \frac{g(x+\Delta x) - g(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x} \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} \\ &= f'(g(x)) \cdot g'(x). && (3.47) \end{aligned}$$

Hier wieder einige Beispiele.

Für $h(x) = \cos(\sin x)$ folgt

$$h'(x) = -\sin(\sin x) \cdot \cos x, \tag{3.48}$$

für $h(x) = \sin(x^2)$

$$h'(x) = \cos(x^2) \cdot 2x, \tag{3.49}$$

und für $h(x) = \sin^2 x$, wie wir ja schon wissen,

$$h'(x) = 2 \sin x \cos x. \tag{3.50}$$

Die letzte Regel, die wir hier prominent herausstellen wollen, ist die *Quotientenregel*

Quotientenregel: Aus

$$h = \frac{f}{g} \tag{3.51}$$

folgt

$$h' = \frac{f'g - fg'}{g^2}. \tag{3.52}$$

Diese Regel können wir direkt aus den vorangegangenen Regeln herleiten: Es ist ja

$$h = \frac{f}{g} = f \cdot g^{-1}, \tag{3.53}$$

also können wir die Produktregel und die Kettenregel anwenden. Wir wissen ja schon, dass

$$h' = f(g^{-1})' + f'g^{-1}. \tag{3.54}$$

Wir müssen uns also überlegen, was $(g^{-1})'$ ist. Aus der Kettenregel erhalten wir aber

$$(g^{-1})' = -\frac{g'}{g^2}. \tag{3.55}$$

Also

$$h' = \frac{f'g - fg'}{g^2}. \tag{3.56}$$

3.2.3 Ableitung der Umkehrfunktion

Für die Ableitung einer Umkehrfunktion haben wir eigentlich schon eine Regel: Sie kann aus der Kettenregel gewonnen werden. Wir wissen, dass für eine Funktion f und deren Umkehrfunktion gilt

$$f^{-1} \circ f = f \circ f^{-1} = \mathrm{id}, \tag{3.57}$$

was meint, dass

$$f(f^{-1}(x)) = x. \tag{3.58}$$

Also ist

$$f'(f^{-1}(x))(f^{-1}(x))' = 1, \tag{3.59}$$

und damit gilt die folgende Regel.

Ableitung der Umkehrfunktion:

$$(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}. \tag{3.60}$$

Wir betrachten einige Beispiele:

> Die Umkehrfunktion von f definiert durch $f(x) = e^x$ ist f^{-1} mit $f^{-1}(x) = \log x$. Also ist
>
> $$(\log(x))' = \frac{1}{e^{\log x}} = \frac{1}{x}. \tag{3.61}$$

> Wenn $f(x) = x^2$ ist $f^{-1}(x) = \sqrt{x}$. Nun folgt somit
>
> $$(\sqrt{x})' = \frac{1}{2\sqrt{x}}. \tag{3.62}$$

> Für $f(x) = \sin x$ ist $f^{-1}(x) = \arcsin x$, also
>
> $$(\arcsin x)' = \frac{1}{\cos(\arcsin x)}. \tag{3.63}$$

Für
$$f(x) = \tan x = \frac{\sin x}{\cos x} \tag{3.64}$$
ist $f^{-1} = \arctan x$, somit gilt
$$(\arctan x)' = \cos^2(\arctan x) = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}. \tag{3.65}$$

Schließlich ist die Umkehrfunktion von $\sinh$ die Funktion arsinh, und es gilt
$$(\operatorname{arsinh} x)' = \frac{1}{\sqrt{1 + x^2}}. \tag{3.66}$$
Die Umkehrfunktion von $\cosh$ heißt arcosh und
$$(\operatorname{arcosh} x)' = \frac{1}{\sqrt{x^2 - 1}}. \tag{3.67}$$

Am Rande bemerkt gelten für arsinh allerlei Kuriositäten. Es gilt etwa
$$\operatorname{arsinh} x = \log(x + \sqrt{x^2 + 1}). \tag{3.68}$$
Warum?

3.2.4 Regel von L'Hospital

Wir werden nun die Regel von L'Hospital kennenlernen, die wir ohne präzisen Beweis formulieren werden. Wir werden auch allerlei Kleingedrucktes über Definitionsbereiche unter den Tisch kehren, und vielmehr sie als praktische Regel formulieren:

Regel von L'Hospital: Es gelten
$$\lim_{x \to p} f(x) = 0, \; \lim_{x \to p} g(x) = 0 \tag{3.69}$$
oder
$$\lim_{x \to p} f(x) = \infty, \; \lim_{x \to p} g(x) = \infty. \tag{3.70}$$
(Unter passenden Differenzierbarkeitsvoraussetzungen) gilt dann
$$\lim_{x \to p} \frac{f'(x)}{g'(x)} = \lim_{x \to p} \frac{f(x)}{g(x)}. \tag{3.71}$$

Das erstaunliche dieser Aussage ist also, dass man dem unbestimmten Ausdruck des Typs

$$„\frac{f(p)}{g(p)} = \frac{0}{0}" \tag{3.72}$$

Bedeutung geben kann. Wenn wir sie auch nicht wirklich beweisen können, können wir diesen Satz zumindest plausibel machen, mit etwas, was *Mittelwertsatz* genannt wird. Falls

$$\lim_{x\to p} f(x) = 0, \ \lim_{x\to p} g(x) = 0 \tag{3.73}$$

gilt, kann man $f(p) = g(p) = 0$ setzen und erhält so eine stetige Funktion. Es gibt nun ein ξ, das zwischen x und p liegt, so dass

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(p)}{g(x) - g(p)} = \frac{f'(\xi)}{g'(\xi)}, \tag{3.74}$$

woraus obige Aussage für $x \to p$ folgt.

> In der Tat ist diese Aussage sehr hilfreich. Wir können damit etwa zeigen, dass für ein $x > 0$
>
> $$\lim_{y\to\infty} \left(1 + \frac{x}{y}\right)^y = e^x \tag{3.75}$$
>
> gilt. Diesen Ausdruck haben wir schon als Gl. (2.45) kennengelernt. Nun ist
>
> $$\left(1 + \frac{x}{y}\right)^y = e^{y \log(1+x/y)}, \tag{3.76}$$
>
> also müssen wir nur zeigen, dass
>
> $$\lim_{y\to\infty} y \log(1 + x/y) = x. \tag{3.77}$$
>
> Hierfür ersetzen wir x durch $1/z$ und betrachten
>
> $$\lim_{z\to 0} \frac{\log(1 + xz)}{z}. \tag{3.78}$$
>
> Nach der Regel von L'Hospital ist dieser Grenzwert gegeben durch
>
> $$\lim_{z\to 0} \frac{\frac{1}{1+xz} x}{1} = x. \tag{3.79}$$
>
> Damit ist die Aussage gezeigt.

3.2.5 Taylor-Reihen

Nun wollen wir noch kurz das Konzept der *Taylor-Reihe* andeuten. Dies ist ein wichtiges Konzept, das in vielen Wissenschaften hilfreich ist. Im Grunde haben wir einfache Instanzen der Taylor-Reihe schon kennengelernt, ohne dies zu merken. Etwa ist

$$e^x \approx 1 + x \tag{3.80}$$

für x sehr nahe bei 0. Oder wir haben gesehen, dass wiederum für x sehr nahe bei 0

$$\sin x \approx x. \tag{3.81}$$

Dies ist keine Überraschung: Denn der Funktionswert der Sinusfunktion bei $x = 0$ ist 0, und die erste Ableitung nimmt den Wert $+1$ an. Also ist die Näherung der Sinusfunktion Gl. (3.81) gerade die Näherung durch die Tangente in $x = 0$.

Also nochmal: Wir nähern eine Funktion mit Hilfe des Funktionswertes und der ersten Ableitung. So kann man die Frage stellen, ob man mit Hilfe der höheren Ableitungen weitere Korrekturen erreichen kann. Man könnte gar davon träumen, dass man mit Hilfe der zweiten Ableitung die richtige Krümmung approximieren kann. Dann die Änderung der Krümmung, und so weiter, so dass man zuletzt den Graphen beliebig genau approximieren kann.

In der Tat ist dies möglich, und die entsprechende Reihe, die dies leistet, heißt Taylorreihe. Wir werden sie nicht beweisen können, aber formulieren können wir sie schon.

Taylorreihe: Für eine Funktion f (die eine Reihe technischer Bedingungen erfüllen muss, insbesondere muss sie unendlich oft differenzierbar sein), heißt

$$\sum_{n=0}^{\infty} \frac{f^{(n)}}{n!}(x_0)(x - x_0)^n = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \dots \tag{3.82}$$

die *Taylorreihe* von f im Punkt x_0.

Zunächst einmal ist die Taylorreihe eine unendliche Reihe, man kann sie aber auch abbrechen und man kann schreiben für ein natürliches m

$$f(x) = \sum_{n=0}^{m} \frac{f^{(n)}}{n!}(x_0)(x - x_0)^n + r(x). \tag{3.83}$$

Der Witz besteht darin, dass bis in der Ordnung m dies die "beste Approximation" der Funktion ist, also gilt für das *Restglied* oder den *Fehler*

$$\lim_{x \to x_0} \frac{r(x)}{(x - x_0)^r}. \tag{3.84}$$

Bevor wir es aber nun wirklich zu weit treiben, hier ein kleines Beispiel.

Wir betrachten die Taylorreihe der Expontialfunktion um $x_0 = 0$: Wir wissen schon, dass

$$e^x \approx 1 + x. \tag{3.85}$$

Der nächste Korrekturterm lautet

$$e^x \approx 1 + x + \frac{1}{2}x^2. \tag{3.86}$$

In der Tat können wir hier alle Ableitungen bestimmen (wegen der einfachen Form der Ableitung der Exponentialfunktion) und erhalten

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{3.87}$$

Diese Reihe heißt *Exponentialreihe*.

Kapitel 4

Integralrechnung

©Bill Watterson

Der Integralrechnung liegt die Motivation zugrunde, die Fläche unter Graphen berechnen zu wollen. In der Tat stellt sich die Integralrechnung als ein ganz zentrales und immens nützliches Instrument dar. Wir werden auch sehen – oder vielmehr aus der Schule auffrischen – dass die Integration gewissermaßen die Umkehroperation der Differentialrechnung ist, was uns sofort gewisse Einblicke erlaubt. Wir werden aber auch sehen, dass, während man bei der Differentiation weitgehend stur gewissen Rezepten und Regeln folgen kann, die Integration mehr Kreativität erfordert. Einige wenige Regeln finden wir aber auch, die wir zusammenstellen wollen.

4.1 Konzept eines Integrals

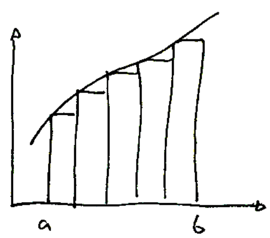
4.1.1 Riemannsches Integral

Riemannsches Integral:

$$\int_a^b dx f(x) = \lim_{n\to\infty} \sum_{j=0}^{n} f(x_j)(x_{j+1} - x_j), \tag{4.1}$$

wobei

$$x_j = \frac{b-a}{n} j + a. \tag{4.2}$$



Dies ist der Ausdruck für das *bestimmte Integral*. Die Fläche eingeschlossen von dem Graphen von f wird also immer besser approximiert durch stückweise konstante Funktionen. Es gelten die folgenden Regeln:

Rechenregeln für Integrale:

$$\int_a^b dx f(x) + \int_b^c dx f(x) = \int_a^c f(x), \tag{4.3}$$

$$\int_a^b dx f(x) = -\int_b^a dx f(x). \tag{4.4}$$

Die Integration kann als Umkehrfunktion der Differentiation angesehen werden. Genauer gilt der folgende wichtige Satz. Dieses ist von so zentraler Bedeutung, dass er sogar "Hauptsatz" genannt wird.

Hauptsatz der Differential- und Integralrechnung: Sei die Stammfunktion F gegeben durch
$$F(x) = \int_a^x dy f(y) + C \tag{4.5}$$
für eine Konstante $C \in \mathbb{R}$. Dann gilt
$$F'(x) = f(x). \tag{4.6}$$

Der Beweis ist elementar, und wir können ihn hier vorführen. Es gilt

$$\begin{aligned} F'(x) &= \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{1}{\Delta x} \left(\int_a^{x+\Delta x} dy f(y) + C - \int_a^x dy f(y) - C \right) \\ &= \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} dy f(y) = f(x). \end{aligned} \tag{4.7}$$

Noch eine Bemerkung zur Nomenklatur: Ein Ausdruck der Form

$$\int_a^b dx f(x) \tag{4.8}$$

wird *bestimmtes Integral* genannt, ein Ausdruck

$$F(x) = \int_0^x dy f(y) \tag{4.9}$$

unbestimmtes Integral.. Es gilt

$$\int_a^b dx f(x) = F(b) - F(a). \tag{4.10}$$

4.1.2 Erste Anwendungen

Der Hauptsatz ist nicht nur von fundamentaler Bedeutung: Er ist auch praktisch sehr hilfreich. Denn immer wenn wir die Ableitung einer Funktion kennen, kennen wir auch ein Paar von Funktionen, von denen eine das unbestimmte Integral der anderen ist. Wir können basierend auf dieser Einsicht bereits einige Integrale formulieren. Die Notation einschließlich einer Konstanten ist üblich.

Einige unbestimmte Integrale:

$$\int dx x^n = \frac{1}{1+n} x^{n+1} + C,\ n \neq 1, \quad (4.11)$$

$$\int dx \sin x = -\cos x + C, \quad (4.12)$$

$$\int dx \cos x = \sin x + C, \quad (4.13)$$

$$\int dx e^x = e^x + C, \quad (4.14)$$

$$\int dx \frac{1}{x} = \log x + C, \quad (4.15)$$

$$\int dx \frac{1}{\sqrt{1+x^2}} = \operatorname{arsinh} x + C, \quad (4.16)$$

$$\int dx \frac{1}{\sqrt{x^2-1}} = \operatorname{arcosh} x + C, \quad (4.17)$$

$$\int dx \frac{1}{1+x^2} = \arctan x + C. \quad (4.18)$$

Natürlich haben sich Leute die Mühe gemacht, Integrationstafeln zusammenzustellen. Keine noch so elaborierte Tafel wird aber jeden Fall abdecken können, und in der Tat erfordert Integration oft etwas Geschick und Erfahrung.

4.2 Integrationsregeln

4.2.1 Partielle Integration

Aus der Produktregel der Differentiation folgt die Regel der partiellen Integration:

Partielle Integration:

$$\int_a^b dx g'(x) f(x) = f(x) g(x)|_a^b - \int_a^b dx g(x) f'(x). \quad (4.19)$$

Hierbei ist $f(x)g(x)|_a^b = f(b)g(b) - f(a)g(a)$. Der Beweis ist einfach:

$$\begin{aligned} \int_a^b dx g'(x) f(x) &= \int_a^b dx \left((g(x)f(x))' - g(x)f'(x)\right) \\ &= f(x)g(x)|_a^b - \int_a^b dx g(x) f'(x). \end{aligned} \quad (4.20)$$

Diese Regel ist dann besonders hilfreich, wenn man Teile einer Funktion schon als Ableitung identifiziert, deren unbestimmtes Integral man kennt.

Beispielsweise ist, wieder in der verkürzten Notation

$$\int dx x e^x = \int dx x (e^x)' = x e^x - \int dx e^x = x e^x - e^x + C. \tag{4.21}$$

Oder

$$\begin{aligned} \int dx e^x \sin x &= \int dx (e^x)' \sin x = e^x \sin x - \int dx e^x \cos x \\ &= e^x \sin x - \int dx (e^x)' \cos x = e^x (\sin x - \cos x) - \int dx e^x \sin x. \end{aligned} \tag{4.22}$$

Also ist

$$\int dx e^x \sin x = \frac{1}{2} e^x (\sin x - \cos x) + C. \tag{4.23}$$

Weil wir das unbestimmte Integral der Sinusfunktion kennen, finden wir auch

$$\begin{aligned} \int dx \sin^2 x &= -\int dx \sin x (\cos x)' \\ &= -\sin x \cos x + \int dx \cos^2 x \\ &= -\sin x \cos x + \int dx (1 - \sin^2 x) \\ &= -\sin x \cos x + x - \int dx \sin^2 x, \end{aligned} \tag{4.24}$$

also

$$\int dx \sin^2 x = \frac{1}{2}(x - \sin x \cos x) + C. \tag{4.25}$$

Ebenso erhält man

$$\int dx \cos^2 x = \frac{1}{2}(x + \sin x \cos x) + C. \tag{4.26}$$

4.2.2 Integration durch Substitution

Aus der Kettenregel der Differentiation können wir auch eine Integrationsregel ableiten: Die Regel der Integration durch Substitution:

Integration durch Substitution:

$$\int_a^b dx g'(x) f(g(x)) = \int_{g(a)}^{g(b)} dy f(y), \tag{4.27}$$

$$\tag{4.28}$$

Wenn F eine Stammfunktion zu f ist, kann man auch schreiben

$$\int dx g'(x) f(g(x)) = F(g(x)). \tag{4.29}$$

Der Beweis ist wiederum nicht sehr schwierig:

$$\begin{aligned}
\int_a^b dx g'(x)(f \circ g)(x)) &= \int_a^b dx g'(x)(F' \circ g)(x) \\
&= \int_a^b dx \, ((F \circ g)(x))' = (F \circ g)(b) - (F \circ g)(a) \\
&= \int_{g(a)}^{g(b)} dy f(y).
\end{aligned} \tag{4.30}$$

Wir sehen also, dass auch diese Regel aus dem Hauptsatz folgt. Hier nun einige wenige Beispiele, hier etwas verkürzt dargestellt:

$$\int dy \sin(ay) = \int dx \frac{1}{a} \sin x = -\frac{1}{a} \cos x + C = -\frac{1}{a} \cos(ay) + C. \tag{4.31}$$

Setze $y = x/a$, also $y' = 1/a$.

$$\int dx \tan x = \int dx \frac{\sin x}{\cos x} = -\int^{\cos x} dg \frac{1}{g} + C = -\log(\cos x) + C, \tag{4.32}$$

wobei $g(x) = \cos x$, $g'(x) = -\sin x$.

$$\int dx \frac{\log x}{x} = \int^{\log x} dg g = \frac{1}{2} \log^2 x + C, \tag{4.33}$$

also $g(x) = \log x$ und $g'(x) = 1/x$.

$$\int dx \frac{1}{\cosh x} = 2\int dx \frac{1}{e^x + e^{-x}} = 2\int \frac{dy}{y}\frac{1}{y+1/y} = 2\int dy \frac{1}{1+y^2}$$
$$= 2\arctan y = 2\arctan e^x. \tag{4.34}$$

Hier ist $x = \log y$, also $e^x = y$, und damit $x' = 1/y$.

4.2.3 Partialbruchzerlegung

Manchmal kann auch die *Partialbruchzerlegung* helfen, die wir hier nur andeuten: Etwa kann man wie folgt integrieren:

$$\int dx \frac{1}{x^2 + a^2} = \int dx \frac{1}{(x-a)(x+a)} = \frac{1}{2a}\int dx \left(\frac{1}{x-a} - \frac{1}{x+a}\right)$$
$$= \frac{1}{2a}\left(\log(x-a) - \log(x+a)\right) = \frac{1}{2a}\log\frac{x-a}{x+a}. \tag{4.35}$$

4.3 Flächeninhalte und Volumina

Ein Integral beschreibt die Fläche unter dem Graphen einer Funktion. Insofern kann es kaum überraschen, dass die Integration ein nützliches Werkzeug ist, um Flächen- oder Volumeninhalte von geometrischen Objekten zu berechnen. Ein einfaches naheliegendes Beispiel ist der Flächeninhalt einer Kreisfläche mit Radius r. Wir können die Einheitssphäre durch

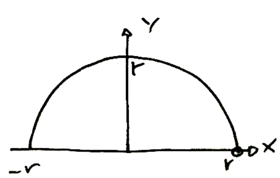
$$f(x) = \pm\sqrt{r^2 - x^2} \tag{4.36}$$

für $x \in [-r, r]$ parametrisieren ($[-r, r]$ bezeichnet hier das *abgeschlossene Intervall* von $-r$ bis r). Der Flächeninhalt ist also gerade gegeben durch

$$F = 2\int_{-r}^{r} dx \sqrt{r^2 - x^2} = 2\int_{-\pi/2}^{\pi/2} d\alpha \cos\alpha \sqrt{r^2 - r^2\sin^2\alpha}$$
$$= 2r^2 \int_{-\pi/2}^{\pi/2} d\alpha \cos\alpha \sqrt{1 - \sin^2\alpha}. \tag{4.37}$$

Mit $x = r\sin\alpha$ folgt $x' = r\cos\alpha$. Also ist

$$F = 2r^2 \int_{-\pi/2}^{\pi/2} d\alpha \cos^2\alpha = 2r^2 \left(\frac{1}{2}(\alpha + \sin\alpha\cos\alpha)\right)_{-\pi/2}^{\pi/2} = \pi r^2. \tag{4.38}$$



Ebenso können wir das *Volumen einer Kugel* berechnen: Wir können schon verwenden, dass wir den Flächenihalt einer Kreisfläche schon kennen. So erhalten wir

$$V = \pi \int_{-r}^{r} (r^2 - x^2) dx = \frac{4}{3}\pi r^3. \tag{4.39}$$

4.4 Länge von parametrisierten Kurven

Nun wollen wir uns zuletzt eine wichtige andere Anwendung der Integralrechnung ansehen, nämlich die Berechnung der Länge von Kurven. Dies ist eigentlich ein Thema, das in die elementare Differentialgeometrie gehört. In einem einfachen Beispiel können wir uns dieses aber nun schon vorknöpfen:

Wie betrachten den Graph der Funktion f. Betrachten wir für ein x_0 und ein $\Delta x > 0$ die Länge Δs der Hypothenuse des Dreiecks, das von $(x_0, f(x_0))$, $(x_0 + \Delta x, f(x_0))$, $(x_0 + \Delta x, f(x_0 + \Delta x))$. Diese Länge ist offensichtlich gegeben durch

$$\Delta s = \sqrt{\Delta x^2 + (f(x_0 + \Delta) - f(x_0))^2}. \tag{4.40}$$

Im Limes kleiner Δx erhalten wir so für die Kurvenlänge des Graphen zwischen a und b

$$L = \int_a^b dx \sqrt{1 + (f'(x))^2}. \tag{4.41}$$

4.5 Noch ein kleines Rätsel

Diese Situation kennt fast jeder: Man setzt sich in großer Runde an einen runden Tisch, und findet jeweils zwischen den Gedecken einen kleinen Teller vor mit einer Serviette. Nun ist stets nicht klar, ob man die Serviette zur Linken oder zur Rechten nehmen soll. Wir nehmen also an, eine Reihe von Mathematikern (und -innen) setzen sich an einen großen rechteckigen Tisch. Wenn beide Servietten da sind, nehmen sie zufällig entweder die rechte oder die linke. Wenn die Sitzplätze in zufälliger Ordnung besetzt werden, wie groß ist – im asymptotischen Limes eines sehr großen Tisches – der relative Anteil derjenigen, die ohne Serviette auskommen müssen?

Hier wieder die Lösung. Wir beschreiben den Tisch als ein Intervall $[0, T]$. Die "Greifzeit" für jeden Gast ist eine unabhängig, gleichverteilte Zufallsvariable in $[0, 1]$. Die Wahrscheinlichkeit, dass ein Gast zur Zeit t die rechte Serviette nicht mehr vorfindet, sei mit $p(t)$ bezeichnet. Dies kann deswegen passieren, weil der rechte Nachbar (die Nachbarin) sich vorher entschieden hat, seine linke Serviette zu nehmen, oder er (sie) keine Wahl hatte, weil seinerseits (ihrerseits) die rechte Serviette fehlte. Also ist

$$p(t) = \frac{1}{2}t + \frac{1}{2}\int_0^t p(s) ds. \tag{4.42}$$

Differenzierung nach t ergibt

$$p'(t) = \frac{1}{2} + \frac{1}{2}p, \tag{4.43}$$

und damit

$$2\log(1+p) = t + C, \tag{4.44}$$

aber die Konstante ist $C = 0$, da $p(0) = 0$. Daher ist

$$p(t) = e^{t/2} - 1. \tag{4.45}$$

Natürlich ist die Wahrscheinlichkeit, dass zu einer Zeit t die linke Serviette fehlt, die gleiche, und für eine Zeit t sind die beiden Ereignisse unabhängig. Die Wahrscheinlichkeit, dass ein Gast zur Zeit t ohne Serviette ausgeht, ist also

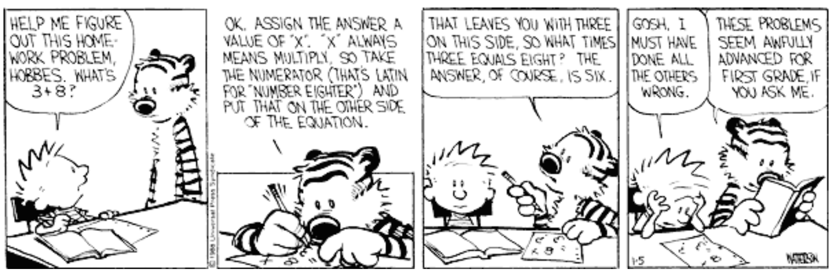
$$p(t)^2 = (e^{t/2} - 1). \tag{4.46}$$

Die Mittelung über den ganzen Prozess des Setzens ergibt

$$\int_0^1 (e^{t/2} - 1)^2 dt = (2 - \sqrt{e})^2. \tag{4.47}$$

Kapitel 5

Vektoren und Matrizen



©Bill Watterson

Viele Größen in den Naturwissenschaften, insbesondere der Physik, sind nicht durch eine einzige Zahl gekennzeichnet, sondern durch Vektoren. Etwa ist der Ort eines Teilchens im Dreidimensionalen durch die Koordinaten in den drei Raumrichtungen beschrieben, ebenso die Geschwindigkeit und die Beschleunigung. Es ergibt viel Sinn, in einer solchen Situation nicht die Koordinaten einzeln zu betrachten, sondern eben Vektoren als eigene Objekte zu verstehen: Etwa ergeben sich so klare Regeln, wie man Koordinaten vom Übergang von einem Koordinatensystem beschreibt. Am Ende des heutigen und morgigen Tages sollte klargeworden sein, wie sehr hilfreich die Objekte von Vektoren und Matrizen sind, auch wenn es sich letztlich nur um eine Art handelt, reelle Zahlen auf geschickte Art zu buchhalten.

5.1 Vektoren und das Skalarprodukt

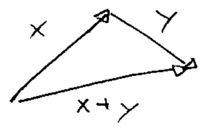
5.1.1 Vektoren

Wir werden *Vektoren* mit fettgedruckten Buchstaben bezeichnen, wie etwa $\mathbf{x}$. Zweidimensionale Vektoren kann man in der Ebene durch gerichtete Pfeile repräsentieren. So dargestellte Vektoren sind frei verschiebbar. Die Addition von

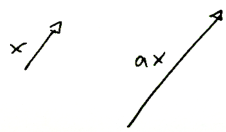
Vektoren

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \tag{5.1}$$

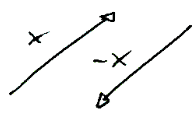
ergibt sich wie folgt:



Wir können Vektoren auch mit einer Zahl multiplizieren, und erhalten für ein $a > 0$ und etwa den Vektor $a\mathbf{x}$.



Für eine negative Zahl $a < 0$ ist $a\mathbf{x}$ auch erklärt: Dann ist der Vektor andersherum orientiert. Also ergibt die Subtraktion zweier Vektoren auch Sinn.



In aller Regel ist es hilfreich, ein *Koordinatensystem* von *Einheitsvektoren* zu wählen. Im $\mathbb{R}^n$, also für n Dimensionen, kann man etwa die Einheitsvektoren $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in Richtung der Koordinatenachsen wählen. Im Dreidimensionalen sind das gerade die Einheitsvektoren in die drei Raumrichtungen, im Zweidimensionalen zwei Einheitsvektoren in der Ebene. In einer Zerlegung

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \tag{5.2}$$

heißen $x_1, \ldots, x_n$ die *Komponenten* des Vektors. In einer etwas ungenauen Notation – denn streng genommen existiert der Vektor auch ohne seine Koordinatendarstellung – identifiziert man oft den Vektor und seine Koordinaten, und schreibt

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}. \tag{5.3}$$

So sind Vektoren also Elemente aus dem $\mathbb{R}^n$. Für die Vektoraddition

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \tag{5.4}$$

erhält man für die Komponenten

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}. \tag{5.5}$$

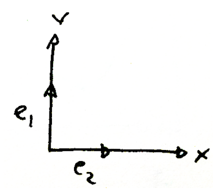
Dies folgt aus der Beobachtung, dass sich die Koordinaten der rechten Seite von Gl. (5.4) als

$$\sum_{j=1}^{n} x_j \mathbf{e}_j + \sum_{j=1}^{n} y_j \mathbf{e}_j = \sum_{j=1}^{n} (x_j + y_j) \mathbf{e}_j \tag{5.6}$$

ergeben. Ebenso ist

$$\alpha \mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \dots \\ \alpha x_n \end{bmatrix}. \tag{5.7}$$

Bei einem Wechsel von einem Koordinatensystem in ein anderes ändern sich auch dessen Koordinaten. Wir werden gleich darauf zurückkommen.



Schließlich sei noch der *Nullvektor* erwähnt, er hat die Komponenten

$$\mathbf{0} = \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix}. \tag{5.8}$$

5.1.2 Skalarprodukt

Skalarprodukt: Das *Skalarprodukt* von zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ist über Komponenten definiert als

$$\mathbf{x} \cdot \mathbf{y} = \sum_{j=1}^{n} x_j y_j. \tag{5.9}$$

Die *Länge eines Vektors* (eine Vektornorm) ergibt sich in Koordinaten als

$$\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x} = \sum_{j=1}^{n} |x_j|^2. \tag{5.10}$$

Wenn α den Winkel zwischen zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$ bezeichnet, kann man das Skalarprodukt auch koordinatenfrei als

$$\mathbf{x} \cdot \mathbf{y} = \|x\| \, \|y\| \cos \alpha \tag{5.11}$$

schreiben. Sind $\mathbf{x}$ und $\mathbf{y}$ senkrecht aufeinander, sind sie *orthogonal* gilt also

$$\mathbf{x} \cdot \mathbf{y} = 0 \tag{5.12}$$

Sind $\mathbf{x}$ und $\mathbf{y}$ dagegen parallel, so ist

$$\mathbf{x} \cdot \mathbf{y} = \|x\| \, \|y\|. \tag{5.13}$$

Offensichtlich ist das Skalarprodukt *kommutativ*, also gilt für beliebige Vektoren

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}. \tag{5.14}$$

Es gilt auch das *Assoziativgesetz für die Multiplikation mit Skalaren*, , was meint, dass für beliebige Vektoren

$$(\alpha\mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y}), \tag{5.15}$$

wie man unmittelbar mit Hilfe der Koordinatenschreibweise zeigen kann. Schließlich gilt auch das *Distributivgesetz*, also ist für drei Vektoren $\mathbf{x}$, $\mathbf{y}$ und $\mathbf{z}$ stets

$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}. \tag{5.16}$$

Am Rande sei bemerkt, dass dies bei weitem nicht das einzige Skalarprodukt ist. Im Allgemeinen ist jedes Produkt, dass bestimmte Bedingungen erfüllt (es ist eine positiv definite hermitesche Sesquilinearform) ein Skalarprodukt.

5.1.3 Eigenschaften der Länge als Vektornorm

Die Länge eines Vektors, wie jede Norm, erfüllt die sogenannte *Dreiecksungleichung*,

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{x}\| \tag{5.17}$$

für beliebige Vektoren $\mathbf{x}$ und $\mathbf{x}$; Der Name wird durch die geometrische Interpretation unmittelbar plausibel. Ebenso gilt

$$\|\alpha\mathbf{x}\| = \alpha\|\mathbf{x}\| \tag{5.18}$$

für reelles α und einen Vektor $\mathbf{x}$.

Eine sehr oft verwendete Ungleichung die Länge eines Vektors betreffend ist die *Cauchy-Schwarz-Ungleichung*:

$$|\mathbf{x} \cdot \mathbf{y}|^2 \leq \|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2. \tag{5.19}$$

5.1.4 Kreuzprodukt

Als vorläufig letztes Produkt wollen wir kurz das *Kreuzprodukt* kennenlernen: Für zwei dreidimensionale Vektoren gilt

$$\mathbf{x} \times \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \tag{5.20}$$

Das Kreuzprodukt dieser beiden Vektoren $\mathbf{x}$ und $\mathbf{y}$ ist also ein Vektor, der orthogonal zu der von den beiden Vektoren aufgespannten Ebene steht und mit ihnen ein sogenanntes Rechtssystem bildet. Die Länge dieses Vektors ist gleich dem Flächeninhalt des Parallelogramms, das von den Vektoren $\mathbf{x}$ und $\mathbf{y}$ aufgespannt wird.

5.2 Matrizen und lineare Gleichungssysteme

5.2.1 Matrizen

Wir werden nun Matrizen im Kontext linearer Gleichungssysteme kennenlernen. Es wird aber bald klarwerden, dass Matrizen auch in vielen anderen Zusammenhängen enorm wichtig sind.

> Ein Beispiel für ein lineares Gleichungssystem ist etwa das folgende:
>
> $$2x + 4y = 2, \tag{5.21}$$
> $$x - y = 4, \tag{5.22}$$
>
> zu finden ist eine Lösung für x und y.

Allgemeiner ist ein *lineares Gleichungssystem* für $\mathbf{x} \in \mathbb{R}^n$ ist eines von der folgenden Form:

Lineares Gleichungssystem:

$$A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = b_1, \tag{5.23}$$
$$A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n = b_2, \tag{5.24}$$
$$\vdots \tag{5.25}$$
$$A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n = b_m. \tag{5.26}$$

Dies kann man mit Hilfe einer Matrix A auch wie folgt schreiben:

$$A\mathbf{x} = \mathbf{b}, \tag{5.27}$$

in Komponenten

$$A \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}, \tag{5.28}$$

also

$$(A\mathbf{x})_j = \sum_{k=1}^{m} A_{j,k} x_k. \tag{5.29}$$

Das Produkt einer Matrix A mit einem Vektor $\mathbf{x}$ ist eben gerade gegeben durch die linke Seite der Gleichungen im Kasten gegeben.

Eine Matrix ist also einfach eine Anordnung von Zahlen, hier reellen Zahlen. Eine *quadratische Matrix* $A \in \mathbb{R}^{n\times n}$ hat Komponenten $A_{j,k}$ mit $j, k = 1, \dots, n$. Allgemeiner sind aber auch Matrizen $A \in \mathbb{R}^{m\times n}$ mit m verschieden von n definiert, sie werden oft *rechteckige Matrizen* genannt.

5.2.2 Gauss-Verfahren zur Lösung linearer Gleichungssysteme

Das Gauss-Verfahren ist ein – wie der Name schon suggeriert, altes – Verfahren, die Lösung eines solchen linearen Gleichungssystems zu finden. Die Grundidee besteht darin, dass wir beliebig Zeilen von linearen Gleichungssystemen linear kombinieren können, ohne die Lösung zu ändern. Wir beschreiben das Gauss-Verfahren erst an zwei Beispielen:

Wir betrachten ein Gleichungssystem

$$\begin{aligned} 4x_1 - 2x_2 &= 1, && (5.30) \\ x_1 + x_2 &= 2, && (5.31) \end{aligned}$$

also

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \tag{5.32}$$

Wenn wir $1/4$ der ersten Gleichung von der ersten abziehen, erhalten wir das Gleichungssystem

$$\begin{bmatrix} 4 & -2 \\ 0 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{7}{4} \end{bmatrix}. \tag{5.33}$$

Wir haben also erreicht, dass die Matrix eine 0 in der linken unteren Ecke hat. So können wir nach x_2 auflösen und finden

$$x_2 = \frac{7}{6}. \tag{5.34}$$

Da wir x_2 nun kennen, können wir auch nach x_1 auflösen und erhalten

$$x_1 = \frac{19}{12}. \tag{5.35}$$

Hier noch ein Beispiel:

$$\begin{aligned} 3x_1 + 4x_2 - 2x_3 &= 4, && (5.36)\\ x_1 - 2x_2 + 3x_3 &= 9, && (5.37)\\ -2x_1 + x_2 - x_3 &= -6, && (5.38) \end{aligned}$$

also

$$\begin{bmatrix} 3 & 4 & -2 \\ 1 & -2 & 3 \\ -2 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ -6 \end{bmatrix}. \tag{5.39}$$

Wir ersetzen nun die zweite Zeile durch -3 mal die zweite Zeile plus einmal die erste; ebenso die dritte durch zweimal die erste und dreimal die dritte. So erhalten wir

$$\begin{bmatrix} 3 & 4 & -2 \\ 0 & 10 & -11 \\ 0 & 11 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -23 \\ -10 \end{bmatrix}. \tag{5.40}$$

Nun wird die dritte Zeile durch 11 mal die zweite und 10 mal die dritte ersetzt,

$$\begin{bmatrix} 3 & 4 & -2 \\ 0 & 10 & -11 \\ 0 & 0 & -51 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -23 \\ -153 \end{bmatrix}. \tag{5.41}$$

Nun können wir wieder sequentiell auflösen und erhalten

$$x_3 = 3,\ x_2 = 1,\ x_1 = 2. \tag{5.42}$$

Nun ist jedes weitere Wort zuviel: Die Idee besteht offensichtlich darin, die Matrix in sogenannte Stufenform zu bringen.

- Erhalten wir eine rechte obere Dreiecksmatrix, so gibt es genau eine Lösung, die wir erhalten, indem wir die entstehenden Gleichungen sequentiell auflösen.

- Falls wir Zeilen erhalten, die ausschließlich Nulleinträge haben, können dagegen zwei Fälle auftreten. Wenn die rechte Seite dieser Zeile aus einer Zahl verschieden von 0 besteht, gibt es keine Lösung.

- Ist diese Zahl auch gleich 0, so gibt es unendlich viele Lösungen.

Weitere Fälle gibt es nicht.

5.2.3 Koordinatentransformationen

Hier wollen wir noch einmal kurz etwas genauer auf die Idee der Koordinatentransformation eingehen. Wie gesagt, während ein Vektor auch unabhängig von seinen Koordinaten definiert ist, werden sich seine Koordinaten sich beim übergang von einem Koordinatensystem in ein anderes ändern. Unter Koordinatentransformationen versteht man Abbildungen $\mathbf{x} \mapsto f(\mathbf{x})$, bei denen f den Übergang des Koordinatensystems beschreibt.

Typischerweise sind dies entweder lineare oder affine Abbildungen:

Lineare und affine Abbildungen: Abbildungen von der Form

$$\mathbf{x} \mapsto A\mathbf{x}, \tag{5.43}$$

heißen *lineare Abbildungen*, solche von der Form

$$\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b} \tag{5.44}$$

affine Abbildungen.

Die wohl wichtigste lineare Abbildung ist die der *Drehung*: Sie beschreibt, wie die Koordinaten in einem gedrehten System aussehen. Im $\mathbb{R}^2$ ist eine solche Drehung gerade beschrieben durch die folgende Transformation. Diese hat ihren eigenen Kasten verdient:

Drehmatrix:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{5.45}$$

dabei ist α der *Drehungswinkel*.

Man kann sich aber auch *Streckungen* vorstellen,

$$\mathbf{x} \mapsto t\mathbf{x} \tag{5.46}$$

mit einem $t > 0$. Verschiebungen sind affine Abbildungen, Drehungen gefolgt von Verschiebungen auch. Oft hilfreich sind aber auch Kugelkoordinaten und Polarkoordinaten, in einer Koordinatentransformation, die nicht affin ist. Darauf wollen wir hier aber nicht eingehen.

5.3 Elemente der Geometrie

5.3.1 Geraden

Wir wollen nun einfache geometrische Objekte betrachten.

Gerade: Paremetrisierung einer Geraden

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{a} + t\mathbf{b},\ t \in \mathbb{R}\} \tag{5.47}$$

Im $\mathbb{R}^2$, in dem wir für dieses Kapitel die Koordinaten mit x und y bezeichnen wollen, sind also alle Punkte der Geraden von der Form

$$\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} a_x \\ a_y \end{array}\right] + t \left[\begin{array}{c} b_x \\ b_y \end{array}\right]. \tag{5.48}$$

In der Tat ergibt sich so die Steigung der Geraden durch b_y/b_x. Wir können die Parameterdarstellung und die funktionale Form in einen einfachen Bezug setzen: Wir haben

$$y = a_y + \frac{t - a_x}{b_x} b_y = a_y + \frac{b_y}{b_x}(x - a_x). \tag{5.49}$$

Eine *Normale* auf eine solche Gerade im $\mathbb{R}^2$ ist bestimmt durch einen Vektor $\mathbf{n}$ mit

$$\mathbf{n} \cdot \mathbf{b} = 0, \tag{5.50}$$

$\|\mathbf{n}\| = 1$. Dies führt sofort zu

$$n_y = -\frac{n_x b_x}{b_y}. \tag{5.51}$$

Um noch Normierung zu erreichen,

$$\mathbf{n} = \frac{1}{\|\mathbf{b}\|} \left[\begin{array}{c} b_y \\ -b_x \end{array}\right]. \tag{5.52}$$

Wir können auch den *minimalen Abstand von Geraden vom Ursprung* im $\mathbb{R}^2$ berechnen: Wir müssen ja nur

$$\|\mathbf{x}\|^2 = \|\mathbf{a} + t\mathbf{b}\|^2 \tag{5.53}$$

in t minimieren. Wir suchen lokale Extrema durch Nullsetzen der Ableitung, also

$$2t\|\mathbf{b}\|^2 + 2\mathbf{a} \cdot \mathbf{b} = 0. \tag{5.54}$$

Also ist

$$t = -\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}, \tag{5.55}$$

und in der Tat ist dies ein Minimum.

Schließlich sei noch die *Normalform* von Geraden im $\mathbb{R}^2$ erwähnt:

$$\left\{\mathbf{x} \in \mathbb{R}^2 : \mathbf{n}(\mathbf{x} - \mathbf{x}_0) = 0\right\}. \tag{5.56}$$

Hierbei ist $\mathbf{x}_0$ ein beliebiger Punkt auf der Geraden.

5.3.2 Ebenen

Eine Ebene ist bestimmt durch zwei linear unabhängige Vektoren $\mathbf{a}$ und $\mathbf{b}$ und einen Punkt der Ebene $\mathbf{x}_0$. Vektoren heißen *linear unabhängig*, wenn sich der Nullvektor nur durch eine Linearkombination der Vektoren erzeugen lässt, in der alle Koeffizienten der Kombinaten den Wert 0 annehmen. Zwei Vektoren verschieden vom Nullvektor sind genau dann linear unabhängig, wenn sie nicht parallel sind. Also ist eine *Ebene* im $\mathbb{R}^n$ parametrisiert durch

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{x}_0 + s\mathbf{a} + t\mathbf{b},\ \ s, t \in \mathbb{R}\} \tag{5.57}$$

Wie erhalten wir nun im $\mathbb{R}^3$ die *Normale auf die Ebene*? Gesucht ist ja ein Vektor $\mathbf{n} \in \mathbb{R}^3$ mit $\|\mathbf{n}\| = 1$ und

$$\mathbf{n} \cdot \mathbf{a} = \mathbf{n} \cdot \mathbf{b} = 0. \tag{5.58}$$

Diese Orthogonalitätsrelationen führen auf

$$a_1 n_1 + a_2 n_2 + a_3 n_3 = 0, \tag{5.59}$$

$$b_1 n_1 + b_2 n_2 + b_3 n_3 = 0. \tag{5.60}$$

Multiplikation der ersten Gleichung mit b_2 und der zweiten mit a_1 ergibt

$$b_2(a_1 n_1 + a_2 n_2 + a_3 n_3) = 0, \tag{5.61}$$

$$a_1(b_1 n_1 + b_2 n_2 + b_3 n_3) = 0, \tag{5.62}$$

und also auch

$$(a_1 b_3 - a_3 b_1) n_1 + (a_2 b_3 - a_3 b_2) n_2 = 0. \tag{5.63}$$

Au ähnliche Art kann man auch n_2 eliminieren. So wird man auf den folgenden Normalenvektor geführt:

$$\mathbf{n} = \frac{n_1}{a_2 b_3 - a_3 b_2} \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_3 - a_3 b_2 \end{bmatrix}. \tag{5.64}$$

Dies ist aber nichts anderes als das normalisierte Kreuzprodukt! Wir erhalten also

$$\mathbf{n} = \frac{\mathbf{a} \times \mathbf{b}}{\|\mathbf{a} \times \mathbf{b}\|}. \tag{5.65}$$

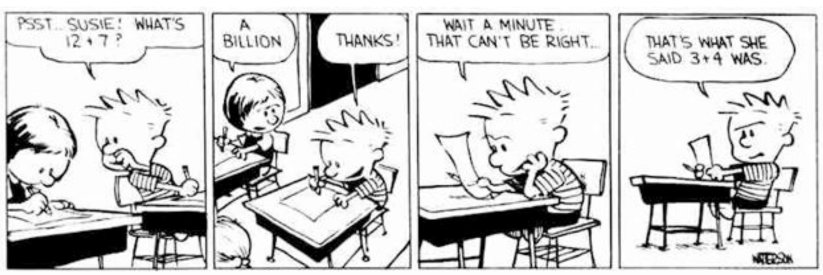
Ein Vektor $\mathbf{a} \times \mathbf{b}$ ist also normal zu $\mathbf{a}$ und $\mathbf{b}$. Wir können auch die Länge des Vektors $\mathbf{a} \times \mathbf{b}$ verstehen: Wie eine kurze Rechnung zeigt, ist nämlich

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\,\|\mathbf{b}\| \sin \alpha, \tag{5.66}$$

wobei α der Winkel zwischen den Vektoren $\mathbf{a}$ und $\mathbf{b}$ ist.

Kapitel 6

Lineare Algebra



In diesem Kapitel wollen wir uns Elementen der linearen Algebra zuwenden. Wieder geht es um Vektoren und Matrizen, aber in einer etwas fortgeschrittenen Weise. Diese Ideen werden sich für sehr viele Anwendungen als äußerst hilfreich erweisen – allerdings wird die Signifikanz mancher Konzepte vielleicht erst im eigentlichen Studium klar.

6.1 Elementares zu Matrizen

6.1.1 Matrixprodukt

Man kann zwei Matrizen miteinander multiplizieren. Dabei ist nicht das komponentenweise Produkt gemeint (wobei es das auch gibt, das heißt dann Hadamardprodukt). Sondern das folgende *Matrixprodukt*

Matrixprodukt: Für zwei Matrizen A und B ist das Produkt $C = AB$ erklärt durch

$$C_{j,k} = \sum_{l} A_{j,l} B_{l,k}. \tag{6.1}$$

6.1.2 Invertierbare Matrizen

Wir betrachten quadratische Matrizen $A \in \mathbb{R}^{n\times n}$. Eine besonders wichtige quadratische Matrix ist die *Einheitsmatrix* $\mathbb{1}$, mit Komponenten

$$\mathbb{1}_{j,k} = \delta_{j,k}. \tag{6.2}$$

Natürlich gilt für jeden Vektor

$$\mathbb{1}\mathbf{x} = \mathbf{x}, \tag{6.3}$$

was den Namen rechtfertigt. Eine Matrix heißt *invertierbar*, wenn das Gleichungssystem

$$A\mathbf{x} = 0 \tag{6.4}$$

folgt, dass $\mathbf{x} = 0$. Dann ist das *Inverse* A^{-1} definiert.

Eigenschaften der Inversen einer Matrix:

$$A^{-1}A = AA^{-1} = \mathbb{1}. \tag{6.5}$$

Hier ist natürlich das Matrixprodukt gemeint. Mit Hilfe der Inversen können wir lineare Gleichungssysteme sehr einfach lösen: Hat im linearen Gleichungssystem

$$A\mathbf{x} = \mathbf{b}, \tag{6.6}$$

die quadratische Matrix A ein Inverses, dann gilt

$$\mathbf{x} = A^{-1}\mathbf{b}. \tag{6.7}$$

6.1.3 Transponierte einer Matrix

Die *Transponierte* A^T einer Matrix A ist eine einfache Matrix: Sie hat die gleichen Komponenten, die nur an der Hauptdiagonalen gespiegelt sind: Also ist

$$(A^T)_{j,k} = A_{k,j}. \tag{6.8}$$

Eine Matrix, für die

$$A = A^T \tag{6.9}$$

gilt, heißt *symmetrische Matrix*. Die Transponierte erfüllt auch

$$(AB)^T = B^T A^T, \tag{6.10}$$

für beliebige Matrizen, wie man unter Zuhilfenahme der Definition des Matrixproduktes direkt ausrechnen kann:

$$((AB)^T)_{j,k} = (AB)_{k,j} = \sum_l A_{k,l}B_{l,j} = \sum_l (B^T)_{j,l}(A^T)_{l,k} = (B^T A^T)_{j,k}. \tag{6.11}$$

6.1.4 Determinante und Spur

Eine *Determinante* ist eine spezielle Funktion, die quadratischen Matrizen eine Zahl zuordnet. Anhand dieser Zahl lassen sich wichtige Eigenschaften der Matrix ablesen. Wir wollen sie hier nur für Matrizen im $\mathbb{R}^{2\times 2}$ und $\mathbb{R}^{3\times 3}$ ansehen, aber es sollte klar sein, dass eine Determinante für jede quadratische Matrix definiert werden kann.

Determinanten von 2×2 und 3×3-Matrizen:

$$\det \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} = A_{1,1}A_{2,2} - A_{2,1}A_{1,2}, \tag{6.12}$$

$$\begin{aligned} \det \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \; r \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} &= A_{1,1}A_{2,2}A_{3,3} + A_{1,2}A_{2,3}A_{3,1} \\ &+ A_{1,3}A_{2,1}A_{3,2} - A_{1,3}A_{2,2}A_{3,1} \\ &- A_{1,2}A_{2,1}A_{3,3} - A_{1,1}A_{2,3}A_{3,2}. \end{aligned} \tag{6.13}$$

Dass hier alle Permutationen vorkommen, ist kein Zufall – so ist für größere Matrizen die Determinante gerade definiert. Etwa kann man an der Determinante ablesen, ob eine quadratische Matrix invertierbar ist. Es gilt nämlich, dass

$$\det = 0 \tag{6.14}$$

genau dann, wenn A invertierbar ist. Übrigens gilt offensichtlich

$$\det A = \det A^T. \tag{6.15}$$

Für Determinanten gilt der *Determinantenmultiplikationssatz*, dass für beliebige Matrizen A und B

$$\det(AB) = \det A \det B \tag{6.16}$$

gilt.

Eine weitere wichtige Größe ist die *Spur* einer quadratischen Matrix $A \in \mathbb{R}^{n\times n}$: Sie ist die Summe aller Hauptdiagonalelemente der Matrix. Also

$$\mathrm{Sp}(A) = \sum_{j=1}^{n} A_{j,j}. \tag{6.17}$$

Nach Definition gilt

$$\mathrm{Sp}(A) = \mathrm{Sp}(A^T). \tag{6.18}$$

Die Spur hat die Eigenschaft, dass man die Matrizen in ihrem Argument *zyklisch* vertauschen kann. Es gilt also für beliebige Matrizen A, B, und C,

$$\mathrm{Sp}(ABC) = \mathrm{Sp}(CAB). \tag{6.19}$$

6.2 Eigenwerte und Eigenvektoren

6.2.1 Definition von Eigenwerten und Eigenvektoren

Die Nützlichkeit von Eigenwerten ist eigentlich nur zu unterschätzen. Das erste Semester wird schon mit einer Vielzahl von Anwendungen aufwarten, und diese Inflation wird sich im Verlaufe des Studiums nicht unbedingt ändern. Das Konzept an sich ist einfach:

Eigenwerte und Eigenvektoren: Gilt für eine Matrix A

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{6.20}$$

so heißt $\lambda \in \mathbb{R}$ *Eigenwert* und $\mathbf{x} \neq \mathbf{0}$ *Eigenvektor*.

"Eigen" ist hier übrigens kein Name – auch wenn in der englischsprachigen Literatur nicht ganz richtig von einem eigenvalue gesprochen wird, sondern es handelt sich gewissermaßen um den Selbstwert (ok, der Begriff ist schon belegt) und den Selbstvektor.

6.2.2 Spektrum von Matrizen

Wie findet man Eigenvektoren? Offensichtlich gilt für jeden Eigenwert und jeden Eigenvektor

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}, \tag{6.21}$$

also

$$(A - \lambda\mathbb{1})\mathbf{x} = \mathbf{0}. \tag{6.22}$$

Diese Gleichung definiert die Eigenwerte und ist in $\mathbf{x}$ ein lineares (offensichtlich) und homogenes (meint, die rechte Seite ist gleich Null) Gleichungssystem dar. Da $\mathbf{x} \neq \mathbf{0}$ vorausgesetzt wird, gibt es genau dann eine Lösung, wenn

$$\det(A - \lambda\mathbb{1}) = 0. \tag{6.23}$$

In λ ist die ein Polynom, und dessen Nullstellen sind gerade die Eigenwerte von A. Wir haben also das Problem vom Auffinden von Eigenwerten auf die Lösung eines Polynoms zurückgeführt! Dieses Polynom heißt *charakteristisches Polynom*. Hieraus folgt auch, dass eine $n \times n$-Matrix höchstens n verschiedene Eigenwerte haben kann.

Wir betrachten ein Beispiel:

$$A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & -1 & 1 \\ 2 & -1 & 3 \end{bmatrix}. \tag{6.24}$$

Also ist $A - \lambda\mathbb{1}$ die Matrix

$$A - \lambda\mathbb{1} = \begin{bmatrix} -\lambda & 2 & -1 \\ 2 & -1-\lambda & 1 \\ 2 & -1 & 3-\lambda \end{bmatrix}. \tag{6.25}$$

Wir berechnen die Determinante,

$$\begin{aligned} \det(A - \lambda\mathbb{1}) &= -\lambda(-1-\lambda)(3-\lambda) + 4 + 2 - (2\lambda + 2 + \lambda + 12 - 4\lambda) \\ &- \lambda^3 + 2\lambda^2 + 4\lambda - 8 - (\lambda - 2)(\lambda - 2)(\lambda + 2). \end{aligned} \tag{6.26}$$

Wir müssen nun die Nullstellen dieses Polynoms finden, man erhält die drei Nullstellen

$$\lambda_{1,2} = 2,\ \lambda_3 = -2. \tag{6.27}$$

Dies sind die Eigenwerte der Matrix A.

Hier sehen wir schon, dass bestimmte Eigenwerte mehrfach vorkommen können. Es können sogar alle Eigenwerte gleich sein: Die Einheitsmatrix $\mathbb{1}$ hat offensichtlich nur einen Eigenwert, nämlich 1, der n mal vorkommt, da ja

$$\mathbb{1} - 1\mathbb{1} = 0. \tag{6.28}$$

Die Vielfachheit eines Eigenwertes als Nullstelle des charakteristischen Polynoms heißt *algebraische Vielfachheit*. Die algebraische Vielfachheit des Eigenwertes 2 im obigen Beispiel ist also 2. Die Menge aller Eigenwerte wird als *Spektrum* bezeichnet, und man kennzeichnet Eigenwerte einer $n \times n$-Matrix meistens mit $\lambda_1, \ldots, \lambda_n$. Für symmetrische Matrizen kann man eine Basis angeben, die aus orthogonalen normierten Eigenvektoren $\mathbf{x}_1, \ldots, \mathbf{x}_n$ besteht, also gilt

$$\mathbf{x}_j \cdot \mathbf{x}_k = \delta_{j,k}. \tag{6.29}$$

6.2.3 Eigenräume

Für einen festen Eigenwert λ wird der Eigenraum der Lösungen

$$(A - \lambda\mathbb{1})\mathbf{x} = \mathbf{0} \tag{6.30}$$

als *Eigenraum* zum Eigenwert λ bezeichnet.

Wir betrachten dazu ein Beispiel, nämlich das gleiche wie das obige: Wir betrachten von

$$A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & -1 & 1 \\ 2 & -1 & 3 \end{bmatrix} \tag{6.31}$$

den Eigenraum zum Eigenwert 2. Hierfür muß man ein lineares Gleichungssystem lösen,

$$\left(\begin{bmatrix} 0 & 2 & -1 \\ 2 & -1 & 1 \\ 2 & -1 & 3 \end{bmatrix} - 2\mathbb{1} \right) \mathbf{x} = \begin{bmatrix} -2 & 2 & -1 \\ 2 & -3 & 1 \\ 2 & -1 & 1 \end{bmatrix} \mathbf{x} = \mathbf{0}. \tag{6.32}$$

Bringt man diese Matrix auf Stufenform, findet man

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}. \tag{6.33}$$

Die gesuchten Eigenvektoren sind also alle Vielfachen des Vektors

$$\mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \tag{6.34}$$

verschieden vom Nullvektor (denn der ist nicht erlaubt). Der Eigenraum zum Eigenwert 2 ist also eindimensional.

Dieser Eigenraum muß im allgemeinen aber nicht eindimensional ein. Die Dimension des Raumes der Eigenvektoren zu einem gegebenen Eigenwert heißt *geometrische Vielfachheit*. Es tritt hier sogar die etwas kuriose Situation auf, dass die algebraische Vielfachheit des Eigenwertes 2 gleich 2 ist, aber die geometrische 1 ist. Es ist aber allgemein so, dass die geometrische Vielfachheit höchstens gleich der algebraischen Vielfachheit ist. Den Beweis dieser Aussage wollen wir uns aber für das erste Semester aufheben.

Schließlich wollen wir nicht unerwähnt lassen, dass die Eigenwerte und Eigenvektoren manchmal auch als Rechtseigenwerte und Rechtseigenvektoren bezeichnet werden, in Abgrenzung von *Linkseigenwerten* und *Linkseigenvektoren*. Jene sind bestimmt durch die Gleichung

$$\mathbf{x}^T A = \lambda \mathbf{x}^T, \tag{6.35}$$

mit Hilfe der Transponierten.

6.2.4 Eigenwertzerlegung

Nun wollen wir die Eigenwerte und Eigenvektoren nutzen, um eine symmetrische Matrix A in eine sehr nützliche Form zu bringen. Wir bezeichnen die Eigenwerte von A mit $\lambda_1, \dots, \lambda_n$ und die normierten Eigenvektoren, die eine orthogonale Basis bilden, mit $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Wir wollen nun eine neue Matrix generieren, die nur aus den normierten Eigenvektoren als Spalten besteht:

$$U = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{array}\right]. \tag{6.36}$$

Dies ergibt wieder eine quadratische Matrix U. Nach Konstruktion gilt das folgende:

Diagonalisierung von symmetrischen Matrizen:

$$A = UDU^T. \tag{6.37}$$

Hierbei ist D die Diagonalmatrix

$$D = \left[\begin{array}{cccc} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{array}\right]. \tag{6.38}$$

Man sagt, dass man die symmetrische Matrix A “diagonalisiert” hat. Auch nach Konstruktion, denn die Eigenvektoren sind so gewählt, dass sie eine orthonormale Basis bilden, gilt auch

$$UU^T = U^T U = \mathbb{1}. \tag{6.39}$$

Matrizen, die diese Eigenschaft haben, heißen *orthogonale Matrizen*.

Orthogonale Matrizen: Orthogonale Matrizen sind solche quadratische Matrizen, die

$$UU^T = U^T U = \mathbb{1}. \tag{6.40}$$

erfüllen.

Nun wollen wir zuletzt noch die Determinante und die Spur über die Eigenwerte ausdrücken. Wir wissen ja schon, dass die Spur invariant unter zyklischen Vertauschungen ist. Also ist

$$\begin{aligned} \mathrm{Sp}(A) &= \mathrm{Sp}(UDU^T) = \mathrm{Sp}(U^T UD) \\ &= \mathrm{Sp}(\mathbb{1}D) = \mathrm{Sp}(D) = \sum_{j=1}^{n} \lambda_j. \end{aligned} \tag{6.41}$$

Also ist die Spur gerade die Summe der Eigenwerte. Ähnlich können wir für die Determinante vorgehen: Denn es ist

$$\det(A) = \det(UDU^T) = \det(U^T UD) = \det(D) = \prod_{j=1}^{n} \lambda_j. \tag{6.42}$$

Letztere Aussage gilt immer, auch wenn wir sie bisher nur für 2×2 und 3×3-Matrizen überprüfen können. Das heißt natürlich auch, dass wir die Eigenwerte einer 2×2-Matrix nur aus der Spur und der Determinante bestimmen können. Denn dann ist

$$\det(A) = \lambda_1 \lambda_2, \ \mathrm{Sp}(A) = \lambda_1 + \lambda_2, \tag{6.43}$$

zwei Gleichungen, die wir ohne weiteres nach λ_1 und λ_2 auflösen können.

6.2.5 Funktionen von Matrizen

Wir wissen noch etwas: Das Inverse einer Matrix A können wir als

$$A^{-1} = UD^{-1}U^T \tag{6.44}$$

schreiben (warum? Bedenke, dass das Inverse, falls es existiert, eindeutig ist). Also kennen wir auch die Eigenwerte von A^{-1}: Sie sind gegeben durch

$$\lambda_1^{-1}, \dots, \lambda_n^{-1}, \tag{6.45}$$

wenn $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind. Die Eigenvektoren sind identisch.

Diese Idee kann man noch etwas weiter treiben. In der Tat kann man für beliebige Funktionen f die Matrix $f(A)$ definieren als

$$f(A) = Uf(D)U^T. \tag{6.46}$$

In der Tat ist ja auch $A^{-1} = UD^{-1}U^T$.

6.2.6 Positive Matrizen

Eine Matrix, deren Eigenwerte positiv (oder genauer nichtnegativ) sind, heißt *positive Matrix*. Dies heißt nicht, dass die Einträge positiv (oder genauer nichtnegativ) sein müssen. Ein Beispiel ist das folgende: Die Eigenwerte von

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{6.47}$$

ist, wie wir nach unserem Trick über Determinanten und Spuren sofort sehen, $\lambda_1 = 0$ und $\lambda_2 = 2$. Also ist A eine positive Matrix. Die Elemente sind aber nicht positiv. Nur die Hauptdiagonalelemente sind immer nichtnegativ bei positiven Matrizen. Warum das so ist, werden wir hier aber nicht sehen können.

Nun aber noch eine Einsicht, die wir uns gründlicher vorknöpfen können: Für eine beliebige quadratische (genauer für jede Matrix, aber dies ist ein anderer Punkt) Matrix A ist

$$B = A^T A \tag{6.48}$$

immer positiv. Warum? Nun, wir können A als $A = UDU^T$ schreiben, also ist

$$B = UDU^T UDU^T = UD^2U^T. \tag{6.49}$$

Die Eigenwerte von UD^2U^T sind die Nullstellen von $\det(UD^2U^T - \lambda\mathbb{1})$, was wiederum die Nullstellen von $\det(D^2 - \lambda\mathbb{1})$ sind, weil $UU^T = \mathbb{1}$. Diese Nullstellen sind aber einfach

$$\lambda_1^2, \ldots, \lambda_n^2, \tag{6.50}$$

alles nichtnegative Zahlen.

6.3 Singulärwerte

Nun treiben wir es auf die Spitze und wollen sogar Matrizen, die nicht symmetrisch sind, in Diagonalform bringen. Der Beweis ist etwas kompliziert, aber wir wollen es einmal versuchen. Wir wissen, dass A^TA eine positive Matrix ist, und als eine solche, ist sie auch symmetrisch, denn es ist

$$(A^TA)^T = A^TA. \tag{6.51}$$

Sie hat also nur nichtnegative Eigenwerte. Wir nehmen nun an, dass alle Eigenwerte von A^TA strikt positiv sind – sonst kann man ganz ähnlich argumentieren, muss nur mit einer kleinen Komplikation klarwerden, die so interessant aber nicht ist. Also kann man eine orthogonale Matrix V und eine Diagonalmatrix mit strikt positiven Einträgen E finden, so dass

$$VA^TAV^T = E. \tag{6.52}$$

Wir nennen $D = E^{1/2}$. Wir wählen nun

$$U = AVD^{-1}. \tag{6.53}$$

Es gilt nun

$$UDV^T = AVDD^{-1}V^T = AVV^T = A. \tag{6.54}$$

Weiterhin ist

$$\begin{aligned} UU^T &= AVD^{-1}(AVD^{-1})^T = AVD^{-2}V^TA^T \\ &= A(A^TA)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = \mathbb{1}\mathbb{1} = \mathbb{1}, \end{aligned} \tag{6.55}$$

und

$$\begin{aligned} U^TU &= D^{-1}V^TA^TAVD^{-1} = D^{-1}V^T(VEV^T)VD^{-1} \\ &= D^{-1}EVD^{-1} = \mathbb{1}. \end{aligned} \tag{6.56}$$

U ist also auch orthogonal. Wir fassen also zusammen (etwas allgemeiner, denn wir brauchen die strikte Positivität nicht):

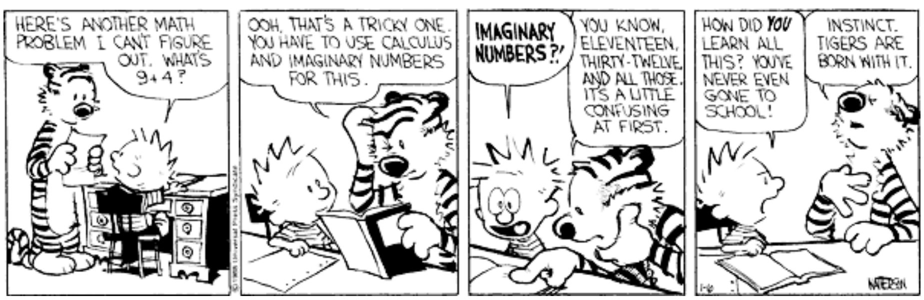
Singulärwertzerlegung: Für jede (reelle) Matrix A kann man orthogonale Matrizen U und V und nichtnegative Diagonalmatrizen D finden, so dass

$$A = UDV^T. \tag{6.57}$$

Wenn man von links und rechts also mit verschiedenen orthogonalen Matrizen multipliziert, kann man sogar nichtsymmetrische Matrizen in Diagonalform bringen.

Kapitel 7

Komplexe Zahlen



©Bill Watterson

Über den reellen Zahlen hat die Gleichung

$$x^2 = -1 \tag{7.1}$$

offensichtlich keine Lösung. Die komplexen Zahlen erweitern die reellen Zahlen in einer Weise, dass diese Gleichung lösbar wird. Dies geschieht durch die Einführung einer neuen Zahl i mit der Eigenschaft, dass

$$i^2 = -1. \tag{7.2}$$

Diese Zahl wird als *imaginäre Einheit* bezeichnet. Die Idee, mit Zahlen zu rechnen, deren Quadrat eine negative Zahl ist, ist so neu nicht – bereits im 16. Jahrhundert wurde sie verwendet; die Bezeichnung i für die imaginäre Einheit geht auf den Königsberger Mathematiker Euler im 18. Jahrhundert zurück.

7.1 Grundlegendes

7.1.1 Definition komplexer Zahlen

Komplexe Zahlen:

$$\mathbb{C} = \{x + iy : \, x, y \in \mathbb{R}\} \tag{7.3}$$

heißt Menge der komplexen Zahlen.

Die Addition und Multiplikation von komplexen Zahlen ist erklärt wie für reelle Zahlen mit i als einem "Parameter", unter Verwendung von $i^2 = -1$.

Beispiele sind wie folgt:

$$(3 + 3i) + (2 - i) = 3 + 2 + (3 - 1)i = 5 + 2i, \tag{7.4}$$
$$(7 + 2i) - i = 7 + (2 - 1)i = 7 + i. \tag{7.5}$$

Imaginär- und Realteil, Konjugation und Betrag: Von einer komplexen Zahl $z = x + iy$ heißt x *Realteil* und y *Imaginärteil*.

$$z^* = x - iy \tag{7.6}$$

heißt die zu z *konjugierte* Zahl. Die reelle Zahl

$$|z| = \sqrt{x^2 + y^2} \tag{7.7}$$

ist der *Betrag* von z.

Hier fällt uns natürlich gleich eines auf: Jede Zahl aus $\mathbb{C}$ kann mit einem Punkt im $\mathbb{R}^2$ identifiziert werden, denn beide sind durch zwei reelle Zahlen beschrieben. Der Realteil entspricht der "x-Achse" (reelle Achse) und der Imaginärteil der "y-Achse" (imaginäre Achse). In der Tat ist der Betrag $|z|$ einer komplexen Zahl ja gerade die Länge des Vektors, wenn wir z als Vektor auffassen würden.

BILD

7.1.2 Rechenregeln komplexer Zahlen

Dennoch sind die komplexen Zahlen ein sehr elegantes hilfreiches Konzept und durch ihre Rechenregeln "mehr", als stets nur zwei reelle Zahlen strukturlos herumzuschleppen. Wir fassen die wichtigsten Eigenschaften der Rechnung mit komplexen Zahlen zusammen:

Eigenschaften komplexer Zahlen: Sei $z = x + iy$ und $c = a + ib$. Dann ist

$$\begin{aligned} z + c &= x + a + (y + b)i, &(7.8)\\ zc &= (x + iy)(a + ib) = xa - yb + i(xb + ya), &(7.9)\\ |z|^2 &= zz^*, &(7.10)\\ (z + c)^* &= z^* + c^*, &(7.11)\\ |zc| &= |z|\,|c|, &(7.12)\\ |z + c| &\leq |z| + |c|. &(7.13) \end{aligned}$$

Manche dieser Gleichungen und Ungleichungen kennen wir schon: Etwa ist die letzte Zeile gerade die *Dreiecksungleichung*. Diese Aussagen kann man alle unmittelbar aus der Definition ableiten.

Beispielsweise ist

$$\begin{aligned} |z|^2 &= x^2 + y^2 = x^2 - (-1)y^2 = x^2 - i^2y^2 = (x + iy)(x - iy) \\ &= zz^*. \end{aligned} \tag{7.14}$$

Was wir noch nicht explizit erklärt haben, ist die *Division* zweier komplexer Zahlen: Die entsprechende Regel folgt aber aus derjenigen für die Multiplikation: Für $z = x + iy$ und $c = a + ib \neq 0$ ist

$$\frac{x + iy}{a + ib} = \frac{(x + iy)(a - ib)}{(a + ib)(a - ib)} = \frac{xa + yb}{a^2 + b^2} + i\frac{ya - xb}{a^2 + b^2}. \tag{7.15}$$

7.1.3 Fundamentalsatz der Algebra

Eine wichtige Eigenschaft wollen wir nur formulieren: Den Beweis werden wir auf das Studium verschieben müssen:

Fundamentalsatz der Algebra: Über $\mathbb{C}$ besitzt jedes Polynom

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 \tag{7.16}$$

mit $a_n \neq 0$ genau n Nullstellen.

Damit ist auch klar, dass jede $n \times n$-Matrix n Eigenwerte hat: Nur sind eben nicht alle Eigenwerte reell, sondern manche können komplex sein. Für reelle symmetrische Matrizen kann man sich aber sicher sein, dass alle Eigenwerte reell sind.

7.2 Komplexe Exponentialfunktion

7.2.1 Eulersche Formel

Die Eigenschaften von Funktionen mit komplexen Argumenten sind Thema der Funktionentheorie, ein Thema, das hier viel zu weit gehen würde. Für eine Funktion wollen wir uns aber ansehen, was passiert, wenn wir ihr ein komplexes Argument geben: Dies ist die Exponentialfunktion. Für komplexe Zahlen, also die Abbildung der *komplexen Exponentialfunktion*

$$z \mapsto e^z \tag{7.17}$$

gilt nach wie vor

$$e^{c+z} = e^c e^z \tag{7.18}$$

für $z, c \in \mathbb{C}$. Es gilt auch der folgende Satz:

Eulersche Formel: Für $x \in \mathbb{R}$ gilt

$$e^{ix} = \cos x + i \sin x. \tag{7.19}$$

Wenn wir komplexe Zahlen in der Ebene $\mathbb{R}^2$ repräsentieren, liegen alle

$$\{e^{ix} : x \in \mathbb{R}\} \tag{7.20}$$

auf einem Einheitskreis. Etwa sind

$$\begin{aligned} e^0 &= 1, && (7.21)\\ e^{i\pi/2} &= i, && (7.22)\\ e^{i\pi} &= -1, && (7.23)\\ e^{2\pi i} &= 1. && (7.24) \end{aligned}$$

Die Eulersche Formel wurde, wenig überraschend, erstmals von Leonhard Euler veröffentlicht, und zwar 1748 in der zweibändigen *Introductio in analysin infinitorum*. Ein Spezialfall ist offensichtlich

$$e^{i\pi} + 1 = 0. \tag{7.25}$$

Diese Gleichung enthält gleich vier der wichtigsten Konstanten der Mathematik:

- Die Kreiszahl π,
- die reelle Einheit 1,
- die Null,

- die imaginäre Einheit i.

Übrigens erlaubt die Eulersche Formel eine Herleitung der Additionstheoreme,

$$e^{i(x+y)} = \cos(x+y) + i\sin(x+y), \tag{7.26}$$

für reelles x und y.

7.2.2 Polardarstellung

Aus der Eulerschen Formel folgt auch, dass sich jedes $z \in \mathbb{C}$ außer der 0 eindeutig beschreiben läßt durch den Abstand r zum Ursprung und dem Winkel α zur reellen Achse:

$$z = r\cos\alpha + ir\sin\alpha = r(\cos\alpha + i\sin\alpha) = re^{i\alpha}. \tag{7.27}$$

Hieraus folgt die *Polardarstellung*

Polardarstellung: Zu jedem $z \in \mathbb{C}$ gibt es eine Polardarstellung:

$$z = re^{i\alpha} \tag{7.28}$$

mit $r \geq 0$ und $\alpha \in \mathbb{R}$.

Wenn $z = x + ib$ nicht reell ist, so ist

$$e^z = e^x(\cos b + i\sin b). \tag{7.29}$$

Kapitel 8

Differentialgleichungen

Mit einer gewissen Rechtfertigung könnte man die Physik als die Lehre der Differentialgleichungen mit Anwendungen in der Naturbeschreibung bezeichnen – seit Newton sind Differentialgleichungen aus der Physik nicht mehr wegzudenken. Dies würde deren Status nicht allzusehr überhöhen. Aber auch in vielen anderen Natur- und auch Wirtschaftswissenschaften sind sie von zentraler Bedeutung. Wir werden hier nur die Spitze des Eisbergs streifen, aber dennoch im Groben sehen, worum es hier geht. Die Wichtigkeit von Differentialgleichungen ist naheliegend: Oft hängt die Änderungsgrad von Größen von einfachen Funktionen dieser Größen selbst ab. Wenn wir dann ihre Entwicklung in der Zeit vorhersagen wollen, müßen wir eine Differentialgleichung lösen.

8.1 Gewöhnliche Differentialgleichungen

8.1.1 Radioaktiver Zerfall

Wir beginnen mit einer Reihe von Beispielen. Ein Beispiel einer Differentialgleichung sahen wir schon, als wir die kontinuierliche Verzinsung diskutierten. Wir wollen und das Problem noch einmal kurz in Erinnerung rufen, nun in der Lesart eines radioaktiven Zerfalls. Eine Substanz zerfalle in der Zeit mit einer

Zerfallskonstanten von λ. Wenn Funktionen von der Zeit abhängen, schreibt man Ableitungen zumeist auch mit einem Punkt $\dot{N}$ für die erste Ableitung und $\ddot{N}$ für die zweite. Zu einer gegebenen Zeit ist die Änderung der Anzahl der Atome N, aufgefasst als Funktion von $t \geq 0$, also gegeben durch

$$\dot{N}(t) = -\lambda N(t). \tag{8.1}$$

Offensichtlich ist die Rate des Zerfalls gerade proportional zu der Anzahl der Atome zu einer Zeit. Dies ist ein Problem einer Differentialgleichung: Also einer linearen Gleichung, die Funktionen und deren Ableitungen enthält. Wir kennen auch deren Lösung schon: Lautet die anfängliche Zahl der Atome $N(0) = N_0$ – dies ist die Anfangsbedingung – so lautet die Lösung der Differentialgleichung

$$N(t) = N_0 e^{-\lambda t}. \tag{8.2}$$

Wir haben also das *exponentielle Zerfallsgesetz* hergeleitet.

8.1.2 Bewegung eines Teilchens

Wir betrachten die Bewegung eines Teilchens im Gravitationsfeld. Von der Luftreibung wollen wir abstrahieren (was wir nicht müßten, warum wird das eigentlich immer gemacht?). Außerdem wollen wir nur die Bewegung in der z-Achse diskutieren, Bewegungen in die anderen Raumrichtungen sind durch das Gravitationsfeld ohnehin unverändert. Wenn das Teilchen die Masse m hat, und g die Gravitationskonstante ist, ist die Änderung der Geschwindigkeit v bestimmt durch

$$m\dot{v}(t) = -mg. \tag{8.3}$$

Diese Gleichung läßt sich integrieren und man findet

$$v(t) = -gt + v_0, \tag{8.4}$$

mit der Anfangsbedingung $v(0) = v_0$. Als Differentialgleichung für den Ort erhalten wir

$$m\ddot{x}(t) = -mg. \tag{8.5}$$

Nach zweimaliger Integration ergibt sich so

$$x(t) = -\frac{1}{2}gt^2 + v_0 t + s_0, \tag{8.6}$$

falls auch $x(0) = x_0$. Dies ist das Weg-Zeit Gesetz, als quadratische Abhängigkeit, denn das Teilchen wird ja immer schneller fallen.

8.1.3 Bewegung mit Reibung

Wenn wir nun einmal Reibung nicht vernachlässigen möchten, können wir die Reibungskraft als linear in der Geschwindigkeit modellieren (dies ist eine gute

Näherung, nur die Luftreibung bei schnellen Autofahrten ist eher quadratisch), erhalten wir

$$m\dot{v}(t) = -\beta v(t). \tag{8.7}$$

Diese Differentialgleichung kennen wir schon: Wir erhalten eine exponentiell gedämpfte Bewegung. Stellen wir die Differentialgleichung nicht für die Geschwindigkeit, sondern für den Ort auf, finden wir

$$m\ddot{x}(t) = -\beta v(t). \tag{8.8}$$

Nun haben wir aber ein Problem: Die rechte Seite ist keine Konstante, sondern sie hängt ihrerseits wieder von der Zeit ab. Wir haben also sowohl erste $v = \dot{x}$ als auch zweite Ableitungen in der Differentialgleichung. Genauer sind beide Ableitungen verknüpft, ohne die Funktion selber zu enthalten. Zur Lösung können wir $\dot{x} = v$ substituieren und in einem zweiten Schritt die Substitution durch Integration wieder rückgängig machen:

$$\dot{x}(t) = v_0 e^{-\beta t}, \tag{8.9}$$

und somit

$$x(t) = \int_c^t v_0 e^{-\beta s} ds = -\beta v_0 e^{-\beta t} + x_0, \tag{8.10}$$

wobei die Konstante c gerade so gewählt ist, dass die Anfangsbedingung $x(0) = x_0$ erfüllt wird.

8.1.4 Struktur von gewöhnlichen Differentialgleichungen

Gewöhnliche Differentialgleichungen: Eine Gleichung, in der Ableitungen einer unbekannten Funktion bis zur n-ten Ordnung auftreten, heißt eine gewöhnliche Differentialgleichung n-ter Ordnung.

Die Ordnung einer Differentialgleichung ist also bestimmt durch die höchste auftretende Ableitung. In *impliziter Form* kann man eine solche Gleichung schreiben als

$$F(t, x, \ldots x, \ddot{x}, \ldots, x^{(n)}) = 0, \tag{8.11}$$

wenn man die Gleichung nach der höchsten Ableitung auflösen, erhält man die *explizite Form*

$$x^{(n)} = f(t, x, \ldots x, \ddot{x}, \ldots, x^{(n-1)}). \tag{8.12}$$

Natürlich ist eine Funktion $t \mapsto x(t)$ eine Lösung, wenn sie mit ihren Ableitungen die Differentialgleichung identisch erfüllt.

8.2 Lineare Differentialgleichungen

Differentialgleichungen heißen *linear*, wenn die implizite und explizite Form lineare Gleichungen der Funktion und in allen ihren Ableitungen sind. Wir wollen uns in diesem Einblick ausschließlich mit solche linearen Differentialgleichungen beschäftigen.

Lineare Differentialgleichungen: Lineare Differentialgleichungen sind von der Form

$$x^{(n)} = f(x, \dots x, \ddot{x}, \dots, x^{(n-1)}) + g(t), \tag{8.13}$$

wobei f eine lineare Funktion ist.

Die Funktion g wird als *Inhomogenität* bezeichnet. Ist sie gleich Null, heißt die Differentialgleichung *homogen*.

Inhomogene Differentialgleichungen: Die Lösung einer inhomogenen Differentialgleichung setzt sich zusammen aus der allgemeinen Lösung der entsprechenden homogenen Differentialgleichung und einer speziellen Lösung der inhomogenen Differentialgleichung.

Also ist das Lösungsverfahren immer zweistufig:

- Man sucht die Lösung der homogenene Differentialgleichung. Für solche erster und zweiter Ordnung gibt es Standardverfahren, sonst benötigt man etwas mehr Kreativität.
- Dann identifiziert man eine spezielle Lösung der inhomogenen Differentialgleichung.

8.2.1 Lineare Differentialgleichungen erster Ordnung

Lineare Differentialgleichungen erster Ordnung:

$$\dot{x}(t) = a(t)x(t) + g(t). \tag{8.14}$$

Wir lösen erst die homogene Gleichung, $\dot{x}(t) = a(t)x(t)$, also

$$a(t) = \frac{\dot{x}(t)}{x(t)}, \tag{8.15}$$

oder

$$\int_0^t a(s)ds = \int_0^t \frac{\dot{x}(s)}{x(s)} ds = \log x(t) + c \tag{8.16}$$

für eine passende Konstante. Also

$$x(t) = \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds + c\right) \tag{8.17}$$

und alle Lösungen sind von der Form

$$x(t) = \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right) + d, \tag{8.18}$$

mit $d \in \mathbb{R}$.

Nun benötigen wir noch *eine* Lösung der inhomogenen Gleichung. Wie wir die finden, ist uns überlassen. Wir setzen an

$$x(t) = h(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right), \tag{8.19}$$

mit einer Funktion h, die wir noch bestimmen wollen. Die Kettenregel ergibt für die Ableitung

$$\dot{x}(t) = h(t)a(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right) + \dot{h}(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right). \tag{8.20}$$

Einsetzen in die Differentialgleichung ergibt

$$\begin{aligned} h(t)a(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right) &+ \dot{h}(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right) \\ &= h(t)a(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right) + g(t), \end{aligned} \tag{8.21}$$

nud damit für h

$$\dot{h}(t) = g(t) \exp\left(\int_0^t \frac{\dot{x}(s)}{x(s)} ds\right), \tag{8.22}$$

oder

$$h(t) = \int_0^t dr g(r) \exp\left(\int_0^r \frac{\dot{x}(s)}{x(s)} ds\right) + c_0, \tag{8.23}$$

mit einer passenden Konstanten c_0. Wir erhalten also schließlich die Lösung der inhomogenen Differentialgleichung

$$x(t) = h(t) \exp\left(\int_0^r \frac{\dot{x}(s)}{x(s)} ds\right), \tag{8.24}$$

mit einer Konstanten c_0, die aus der Anfangsbedingung $x(0)$ zu bestimmen ist.

Dies sieht zugegebenermaßen schrecklich aus, aber ist halb so wild, weil die Integrale typischerweise gelöst werden können.

8.2.2 Ein Beispiel einer Differentialgleichung erster Ordnung

Hier ein Beispiel, das dieses Verfahren anwendet:

$$ax(t) = \dot{x}(t) + \cos(\omega t). \tag{8.25}$$

Eine spezielle Lösung der inhomogenen Gleichung ist

$$x(t) = h(t)e^{at}, \tag{8.26}$$

ein Ausdruck, den wir ableiten können

$$\dot{x}(t) = \dot{h}(t)e^{at} + h(t)ae^{at} \tag{8.27}$$

und in die Differentialgleichung einsetzen,

$$(\dot{h}(t)e^{at} + ac(t)e^{at}) + \cos(\omega t) = ac(t)e^{at}. \tag{8.28}$$

Die Differentialgleichung für h ist somit

$$\dot{h}(t) = -\cos(\omega t)e^{-at}. \tag{8.29}$$

Zweimalige partielle Integration liefert

$$h(t) = -\frac{e^{-at}}{a^2+\omega^2}(-a\cos(\omega t) + \omega\sin(\omega t)) + c_0, \tag{8.30}$$

und also die gesamte Lösung

$$x(t) = h(t)e^{at}. \tag{8.31}$$

Nicht schön, aber auch nicht fürchterlich.

8.2.3 Lineare Differentialgleichungen zweiter Ordnung

Nun wollen wir noch kurz Differentialgleichungen zweiter Ordnung anreissen: Diese sind wichtig zur Beschreibung harmonischer Bewegungen, wie die eines Pendels. Erst einmal formulieren wir die allgemeine Form der Differentialgleichung:

Lineare Differentialgleichungen zweiter Ordnung:

$$\ddot{x}(t) + a(t)\dot{x}(t) + b(t)x(t) + g(t) = 0. \tag{8.32}$$

Natürlich ist die homogene Gleichung die, für die $g = 0$ ist. Bei der Diskussion der homogenen Gleichung wollen wir nun annehmen, dass die Koeffizienten a und b konstant sind: Dadurch können wir uns die unübersichtlich aussehenden obigen Integrale ersparen; es sollte aber klar sein, dass dies strukturell keinen besonders großen Unterschied macht. Wir betrachten also

$$\ddot{x}(t) + a\dot{x}(t) + bx(t) = 0. \tag{8.33}$$

Diese Gleichung läßt sich mit Hilfe eines Exponentialansatzes

$$x(t) = e^{\lambda t} \tag{8.34}$$

lösen. Wir beobachten drei Dinge:

- Ist x_1 als Funktion von t eine Lösung, so ist cx_1 für ein beliebiges $c \neq 0$ auch eine Lösung.
- Sind x_1 und x_2 Lösungen, so ist $c_1x_1 + c_2x_2$ auch eine Lösung. Genauer ist die allgemeine Lösung x eine Linearkombination

  $$x(t) = c_1x_1(t) + c_2x_2(t) \tag{8.35}$$

  zweier linear unabhängiger Lösungen.
- Ist $x = a + ib$ eine Lösung, so auch der Realteil a und der Imaginärteil b.

Nun suchen wir die Lösungen der inhomogenen Gleichung. Wir leiten den Exponentialansatz ab

$$\dot{x}(t) = \lambda e^{\lambda t} = \lambda x(t), \tag{8.36}$$

$$\ddot{x}(t) = \lambda^2 x(t) \tag{8.37}$$

und setzen ihn in die Differentialgleichung ein:

$$\lambda^2 x(t) + a\lambda x(t) + bx(t) = 0. \tag{8.38}$$

Die Lösung $x(t) = 0$ für alle Zeiten wollen wir ausschließen, also muss

$$\lambda^2 + a\lambda + b = 0 \tag{8.39}$$

gelten. Dies ist eine quadratische Gleichung für λ. Die Lösungen sind

$$\lambda_{1,2} = -\frac{a}{2} \pm \sqrt{\left(\frac{a}{2}\right)^2 - b}. \tag{8.40}$$

- Ist der Radikant negativ, dann sind die Lösungen für λ komplex. Physikalisch ergibt nur der Realteil der Lösung Sinn.
- Der Radikant ist positiv, dann sind die Lösungen reell.
- der Radikant ist 0. Dann gibt es nur eine Lösung. Die andere linear unabhängige Lösung ergibt sich aus $te^{\lambda t}$.

Die allgemeine Lösung der inhomogenen Differentialgleichung ergibt sich in jedem Fall durch

$$x(t) = c_1e^{\lambda_1 t} + c_2e^{\lambda_2 t}. \tag{8.41}$$

Die Lösung der inhomogenen Gleichung dann dann aus diesen Lösungen und einer speziellen Lösung des inhomogenen Problems gewonnen werden.

8.2.4 Beispiel des gedämpften Pendels

Das Federpendel mit Masse m und Rückstellkraft $-kx$ wird durch die Bewegungsgleichung

$$\ddot{x}(t) + \omega_0^2 x(t) = 0 \tag{8.42}$$

beschrieben, wobei $\omega^2 = k/m > 0$. Eine solche Differentialgleichung mit positiven Koeffizienten führt immer auf eine allgemeine Lösung der Form

$$x(t) = c_1 \sin(\omega_0 t) + c_2 \cos(\omega_0 t). \tag{8.43}$$

Da die Differentialgleichung bereits homogen ist, ist dies die Lösung der Differentialgleichung. Dies ist die *harmonische Schwingung* mit Frequenz ω_0.

Wenn zusätzlich zu der Federkraft eine Reibungskraft $-\beta v = -\beta \dot{x}$ wirkt, erhalten wir

$$\ddot{x}(t) + 2\gamma_0 \dot{x}(t) + \omega^2 x(t) = 0, \tag{8.44}$$

wobei $2\gamma = \beta/m > 0$. Hier tritt die Situation auf, dass wir verschiedene Fälle unterscheiden müssen:

- Ist der Radikant positiv, erhalten wir die Lösungen

  $$x(t) = e^{-\lambda t}(c_1 \cos(\omega t) + c_2 \sin(\omega t)), \tag{8.45}$$

  mit $\omega^2 = \omega_0^2 - \gamma^2$. Dies ist der Fall der *gedämpften Schwingung*.

- Im Fall eines negativen Radikanten werden wir auf

  $$x(t) = e^{-\lambda t}(c_1 e^{kt} + c_2 e^{-kt}). \tag{8.46}$$

  geführt, wobei das reelle $k^2 = \gamma^2 - \omega_0^2$. Dies ist der Kriechfall, der Pendel schwingt nicht, sondern "kriecht" zur Ausgangslage zurück, in einer exponentiell gedämpften Form.

- Schließlich betrachten wir noch kurz den Fall, an dem der Radikant verschwindet. Dies ist der *aperiodische Grenzfall*. Dies beschreibt die Situation, in der $\gamma = \omega_0$. Hier ergibt sich eine Lösung von der Form

  $$x(t) = v_0 t e^{-\gamma t}. \tag{8.47}$$

8.2.5 Beispiel zweier gekoppelter Pendel

Schließlich wollen wir den Fall zweier gekoppelter Pendel diskutieren. Hier sollte klarwerden, warum Eigenwerte und Eigenvektoren so nützlich sind, ebenso wie komplexe Zahlen. Auch werden wir Koordinatentransformationen wiedersehen. Eigentlich ist in diesem Beispiel fastalles vereint, was wir bisher kennengelernt haben. Die Bewegungsgleichungen sind schnell aufgestellt:

$$\ddot{x}_1(t) + \omega_0^2 x_1(t) + \xi(x_1(t) - x_2(t)) = 0, \tag{8.48}$$
$$\ddot{x}_2(t) + \omega_0^2 x_2(t) + \xi(x_2(t) - x_1(t)) = 0. \tag{8.49}$$

Für $\xi = 0$ hätten wir keine Kopplung, und die beiden Pendel würden für sich schwingen. Im Fall einer Kopplung ist die Situation aber komplizierter: Wir können keine Gleichung für x_1 lösen, weil die ja von x_2 abhängt. Ebenso hängt die für x_2 von x_1 ab.

Wir schreiben diese lineare Gleichung zunächst als Matrixgleichung:

$$\begin{bmatrix} \ddot{x}_1(t) \\ \ddot{x}_2(t) \end{bmatrix} = \begin{bmatrix} \omega_0^2 + \xi & -\xi \\ -\xi & \omega_0^2 + \xi \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = 0. \tag{8.50}$$

Die Kopplung manifestiert sich hier in der Beobachtung, dass die Matrix nicht diagonal ist. Aber wir wissen, wie wir Matrizen diagonalisieren können! Es gibt eine orthogonale Matrix U und eine diagonale Matrix D, so dass

$$\begin{bmatrix} \omega_0^2 + \xi & -\xi \\ -\xi & \omega_0^2 + \xi \end{bmatrix} = UDU^T. \tag{8.51}$$

D enthält auf der Hauptdiagonalen die Eigenwerte obiger Matrix, und man findet schnell

$$D = \begin{bmatrix} \omega_0^2 & 0 \\ 0 & \omega_0^2 + 2\xi \end{bmatrix}. \tag{8.52}$$

Also können wir ebenso schreiben

$$\begin{bmatrix} \ddot{x}_1(t) \\ \ddot{x}_2(t) \end{bmatrix} = UDU^T \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = 0, \tag{8.53}$$

oder, mit den neuen Koordinaten

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = U^T \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{8.54}$$

auch

$$\begin{bmatrix} \ddot{y}_1(t) \\ \ddot{y}_2(t) \end{bmatrix} = D \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = 0. \tag{8.55}$$

Dies sind aber zwei entkoppelte Gleichungen, die wir gerade so lösen können, wie oben beschrieben! Die Eigentwertzerlegung “zerlegt” das gekoppelte Problem also in zwei ungekoppelte.

Mehr noch: Wir finden auch einfach die Eigenvektoren, und damit U: Es sind

$$\begin{bmatrix} \omega_0^2 + \xi & -\xi \\ -\xi & \omega_0^2 + \xi \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \omega_0^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{8.56}$$

und

$$\begin{bmatrix} \omega_0^2 + \xi & -\xi \\ -\xi & \omega_0^2 + \xi \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = (\omega_0^2 + 2\xi) \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{8.57}$$

Also ist U gerade die Drehmatrix (schon wieder ein alter Bekannter!)

$$U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} / \sqrt{2}. \tag{8.58}$$

Diese Eigenvektoren kennzeichnen auch die beiden *Normalschwingungen*:

- In einer Situation schwingen die Pendel gegenläufig, mit Frequenz $\sqrt{\omega_0^2 + 2\xi}$.
- In der anderen miteinander mit Frequenz ω_0.

Wir haben hier also verstanden, wie man eine recht komplexe Situation mit Hilfe von Eigenvektoren und Eigenwerten in eine sehr einfache überführen kann. Die Interpretation ist auch anschaulich: Die obige Matrix kennzeichnet die Wechselwirkungssituation, und die Eigenvektoren bestimmen gerade die Eigen-, also die "Selbstschwingungen".

8.2.6 Partielle Differentialgleichungen

Die Gliederung des Skriptes sieht vor, hier ein paar Worte zu partiellen Differentialgleichungen zu verlieren. Es wird bei ein paar Worten bleiben: Wir werden sie nicht behandeln. Partielle Differentialgleichungen unterscheiden sich von gewöhnlichen, dass sie partielle Ableitungen nach mehreren Variablen enthalten. Sie sind sehr wichtig als Wellengleichungen, in der Quantenmechanik, und in der Strömungsmechanik. Dies werden wir uns aber wirklich für das eigentliche Studium aufheben.

Kapitel 9

Beweisideen

9.0.7 Reductio ad absurdum

9.0.8 $\sqrt{2}$ ist keine rationale Zahl

9.0.9 Fibonacci-Zahlen und der goldene Schnitt

9.0.10 Kuhwiegen