

Vorkurs Mathematik für Naturwissenschaftler

Wintersemester 2021/22, Tübingen

Sascha Eichmann

14. Oktober 2021

Inhaltsverzeichnis

1 Grundlagen der Arithmetik	4
1.1 Zahlen	4
1.2 Formales Rechnen mit Variablen	7
1.3 Summen und Produktzeichen	10
1.4 Potenzen und Wurzeln	12
1.5 Binomische Formeln und quadratische Gleichungen	13
1.6 Logarithmen	15
1.7 Größenbeziehungen	17
2 Mengenlehre	22
2.1 Was ist eine Menge?	22
2.2 Mengenoperationen	23
2.3 Tupel und das Kartesische Produkt	26
3 Abbildungen	28
3.1 Graphen von Abbildungen	29
3.2 Trigonometrische Funktionen	30
3.3 Grenzwerte und Stetigkeit	32
4 Differenzialrechnung	35
4.1 Was ist eine Ableitung?	35
4.2 Rechengesetze für Ableitungen	37
4.3 Extremwerte	40
5 Integralrechnung	43
5.1 Die Fläche unter einem Graph	43
5.2 Hauptsatz der Differential- und Integralrechnung	45
5.3 Tricks zur Berechnung von Integralen	48
6 Wahrscheinlichkeitsrechnung	51
6.1 Endlich viele mögliche Ereignisse, Kolmogorov Axiome	51
6.2 Modellierung, die Regel von Laplace	55
6.3 Erwartungswert und Varianz	56
6.4 Ausblick: \mathbb{R} als Menge der mögliche Ereignisse	58
A Partialbruchzerlegung zum Integrieren	61

B Differentialgleichungen	64
C Griechische Buchstaben	69
Literatur	70
Index	71

Dieses Skript basiert auf den in den vergangenen Jahren gehaltenen Vorkursen von Elmar Teuff, siehe [4], dem Buch [1], welches gut geeignet zum Einstieg in die Mathematik für Ingenieure ist, einigen Vorlesungsnotizen aus Halle [2] und dem Buch von Walter [5] über Analysis. Letzteres ist allerdings besser geeignet als Einstieg in die reine Mathematik, als in die Ingenieursmathematik.

Dank möchte ich an dieser Stelle an Julia Kientzle und Silvio Weging aussprechen, die mir bei der Ausarbeitung dieses Skriptes mit der hervorragender konstruktiver Kritik geholfen haben.

1 Grundlagen der Arithmetik

1.1 Zahlen

Wir entwickeln in diesem ersten Schritt in einer eher informellen Weise, die Zahlen mit denen wir arbeiten werden.

Zuerst stellen wir uns dabei auf einen einfachen Standpunkt und stellen fest, dass das Wort Zahl mit mit Verb zählen zusammenhängt. Also wollen wir Gegenstände oder ähnliche Dinge zählen. Hier am Beispiel von Äpfeln:

1 Apfel, 2 Äpfel, , 3 Äpfel,

Die Zahlen 1, 2, 3, 4, . . . nennen wir deswegen *natürliche Zahlen*. Für die Gesamtheit der natürlichen Zahlen benutzen wir das Symbol \mathbb{N} . (Achtung, laut DIN-Norm 5473 gehört die 0 zu den natürlichen Zahlen, hier ist diese ausgeschlossen. Dies ist in der mathematischen Literatur *nicht* einheitlich)

Was können wir nun mit natürlichen Zahlen tun? Wir können bereits gezählte Größen zusammenrechnen!

Wir nehmen nun 2 Haufen von Äpfeln, einen Haufen mit 5 Äpfeln und einen Haufen mit 10 Äpfeln. Wie viele Äpfel haben wir insgesamt?

5 Äpfel und 10 Äpfel macht insgesamt 15 Äpfel.

Für diese Aktion führen wir folgendens Zeichen ein $+$ und nennen es addieren. Etwas abstrahiert haben wir also

$$5 + 10 = 15.$$

Was passiert nun, wenn wir Äpfel entfernen, sie z.B. essen oder verschenken? Angenommen wir haben 10 Äpfel und geben davon 6 weg, wie viele Äpfel bleiben übrig?

10 Äpfel weniger 6 Äpfel, ergeben 4 Äpfel.

Dafür nutzen wir das Zeichen $-$ und nennen es subtrahieren. Also

$$10 - 6 = 4.$$

Was passiert aber wenn wir mehr Äpfel weggeben, als wir haben. Z.B. wenn wir Schulden machen wollen? Negative Äpfel existieren natürlich nicht, aber mit dem Minuszeichen ($-$) können wir trotzdem Buch darüber führen. So kommen wir zum Begriff der *ganzen Zahlen*:

. . . , -3, -2, -1, 0, 1, 2, 3,

Für die Gesamtheit der ganzen Zahlen nutzen wir das Symbol \mathbb{Z} . Nun können wir also addieren und subtrahieren nach Herzenslust.

Nun sind wir Menschen und deswegen versuchen wir unsere Rechnungen möglichst kurz zu halten. Wenn wir z.B. 10 Haufen Äpfel mit jeweils genau 3 Äpfel haben, wie können wir so unsere Gesamtheit an Äpfel ausrechnen? Bis dato müssen wir folgendes machen

$$3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 = 30.$$

Kürzen wir dies ab, indem wir ein neues Symbol einführen, den Malpunkt \cdot und nennen es Multiplikation. Damit schreiben wir

$$3 \cdot 10 = \text{die } 3 \text{ } 10 \text{ mal aufaddieren} = 30.$$

Was passiert nun, wenn wir das mit negativen Zahlen machen? Wie folgt: Ist genau eine von den beiden beteiligten Zahlen negativ, so wird das Ergebnis negativ. Sind beide Zahlen positiv oder negativ, so wird das Ergebnis positiv. Im Beispiel:

$$\begin{aligned} 3 \cdot 5 &= 15 \\ (-3) \cdot 5 &= -15 \\ (-3) \cdot (-5) &= 15 \\ 3 \cdot (-5) &= -15. \end{aligned}$$

Wie interagiert nun die Multiplikation mit der Addition? Im Beispiel:

$$4 \cdot 3 + 7 = 3 + 3 + 3 + 3 + 7 = 19,$$

d.h. wir führen die Punktrechnung vor der Strichrechnung durch. Was ist aber wenn wir es andersherum erzwingen wollen? Dann setzen wir Klammern:

$$4 \cdot (3 + 7) = 4 \cdot 10 = 40.$$

Wir können damit sogar ausmultiplizieren:

$$4 \cdot (3 + 7) = 4 \cdot 3 + 4 \cdot 7 = 12 + 28 = 40.$$

Nachdem wir uns nun die Multiplikation etwas genauer angesehen haben, stellt sich jetzt nun die Frage, können wir diese rückgängig machen? Anders ausgedrückt, wenn wir wissen, dass z.B. 5 mal eine unbekannte Zahl 15 ergibt, wie groß ist diese unbekannte Zahl? Wir können etwas herumprobieren und sehen dann, dass 3 das Ergebnis ist, also

$$5 \cdot ? = 15 \text{ also mit Überlegen } 5 \cdot 3 = 15.$$

Wir formalisieren dies nun mit dem Bruchstrich und nennen die Operation Division:

$$\frac{15}{5} = 3.$$

Was passiert aber nun, wenn wir keine Lösung unseres Problems finden? Z.B. wenn wir uns fragen mit welcher Zahl wir die 2 multiplizieren müssten, um zur 1 zu kommen? In der Sprache der Äpfel: Was passiert wenn ich meinen Apfel in zwei Teile aufteile? Ich bekomme 2 halbe Äpfel. Dafür schreiben wir einfach

$$\frac{1}{2}$$

Diesen Ausdruck nennen wir *Bruch*.

Damit erreichen wir wieder einen neuen Zahlbereich! Wir nennen die Gesamtheit aller Brüche die *rationalen Zahlen* und nutzen das Symbol \mathbb{Q} .

Hier müssen wir nun etwas Vorsicht walten lassen, denn solche Zahlen sind in ihrer Darstellung nicht eindeutig. Z.B. haben wir

$$\frac{4}{8} = \frac{1}{2} = \frac{4 \cdot 1}{4 \cdot 2}.$$

D.h. immer wenn im Zähler (der obere Teil des Bruches) und im Nenner (der untere Teil des Bruches) dieselbe Zahl als Faktor drinsteht, können wir diese wegstreichen! Diese Operation nennt man *Kürzen* eines Bruches.

Wie aber rechnen wir nun mit solchen Brüchen? Wir wollen wieder addieren, subtrahieren, multiplizieren und dividieren:

Überlegen wir uns zuerst die Addition: Wenn wir 2 Brüche haben, dann drücken diese jeweils gewisse Anteile aus. Z.B. $\frac{3}{5}$ gibt an, dass wir 3 von 5 Anteilen haben. Wollen wir nun zwei Brüche addieren, müssen wir erst zu einer gemeinsamen Gesamtanzahl an Anteilen kommen. Nehmen wir beispielhaft die Brüche $\frac{3}{5}$ und $\frac{7}{8}$. Wir können nun das kürzen von oben rückwärts durchführen (diese Operation nennt man *erweitern*) und so zu einer Gesamtzahl an Anteilen kommen:

$$\frac{3}{5} = \frac{8 \cdot 3}{8 \cdot 5} = \frac{24}{40} \quad \text{und} \quad \frac{7}{8} = \frac{5 \cdot 7}{5 \cdot 8} = \frac{35}{40}.$$

Hier haben wir jeweils mit dem anderen Nenner erweitert. Dies klappt immer! Nun können wir die Anteile einfach zusammenzählen und erhalten

$$\frac{3}{5} + \frac{7}{8} = \frac{24}{40} + \frac{35}{40} = \frac{24 + 35}{40} = \frac{59}{40}.$$

Wir fassen zusammen: Beim addieren von Brüchen berechnen wir zuerst die Gesamtanzahl an beteiligten Anteilen und erweitern die Brüche entsprechend. Diesen Vorgang nennt man *Hauptnenner* bilden. Dann addieren wir die resultierenden Zähler.

Der Hauptnenner selber ist das sogenannte *kleinste gemeinsame Vielfache* der beteiligten Nenner. In diesem Sinne ist damit die kleinste mögliche Gesamtzahl der beteiligten Anteile gemeint.

Bei der Subtraktion können wir fast genauso vorgehen. Wir bilden zuerst den Hauptnenner und ziehen aber nun die resultierenden Zähler voneinander ab. Im Beispiel:

$$\frac{3}{5} - \frac{7}{8} = \frac{24}{40} - \frac{35}{40} = \frac{24 - 35}{40} = \frac{-11}{40} = -\frac{11}{40}.$$

Kommen wir nun zur Multiplikation. Hier ist das Vorgehen zum Glück einfacher. Wenn man zwei Brüche miteinander multiplizieren möchte, multipliziert man den Zähler und den Nenner miteinander. Z.B.:

$$\frac{2}{5} \cdot \frac{8}{3} = \frac{2 \cdot 8}{5 \cdot 3} = \frac{16}{15}.$$

Wie dividieren wir nun? Ganz einfach, wir vertauschen Nenner mit Zähler des Bruches, durch den geteilt wird und multiplizieren dann damit. Anders ausgedrückt man bildet den *Kehrwert* und multipliziert damit. Beispiel:

$$\frac{5}{7} \div \frac{2}{9} = \frac{5}{7} \cdot \frac{9}{2} = \frac{5 \cdot 9}{7 \cdot 2} = \frac{45}{14}.$$

Dies passt übrigens zur 'normalen' Division mit ganzen Zahlen. Wollen wir zum Beispiel 18 durch 6 teilen, so erhalten wir

$$18 \cdot \frac{1}{6} = \frac{18}{1} \cdot \frac{1}{6} = \frac{18}{6} = \frac{3}{1} = 3.$$

Wir haben hier also nichts anderes gemacht als die Division durch $\frac{6}{1}$ durch die Multiplikation mit dem Kehrwert durchzuführen und zu kürzen.

Eine rationale Zahl kann man auch in einer Kommadarstellung schreiben, z.B.

$$\frac{5}{2} = 2,5.$$

Dabei kann es passieren, dass unendlich viele Kommastellen auftreten:

$$\frac{1}{3} = 0,3333\dots$$

Die 3 wiederholt sich nun. Was passiert wenn diese sich nicht mehr wiederholen bzw. die Zahlen keinem Muster mehr folgen, aber dennoch die Kommazahlen sich bis ins unendliche fortsetzen? Dann erreichen wir einen neuen Zahlbereich, denn solche Zahlen können wir nicht mehr als den Bruch zweier ganzer Zahlen darstellen. Diesen neuen Zahlbereich nennen wir die *reellen Zahlen*. Das Symbol dafür ist \mathbb{R} . Innerhalb der reellen Zahlen können wir alle Operationen, die wir besprochen haben durchführen, d.h. Addition, Subtraktion, Multiplikation und Division.

Berühmte reelle Zahlen, die keine rationalen Zahlen sind, sind z.B. $\sqrt{2}$ oder die Kreiszahl π . Wie alle unsere Zahlbereiche in einander liegen, verdeutlicht Bild 1.1.

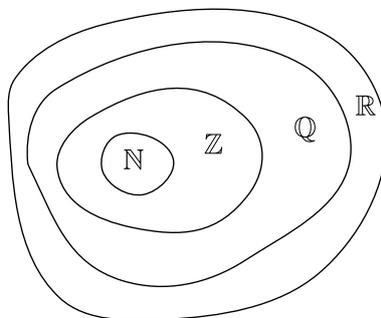


Abbildung 1.1: Diagramm wie die verschiedenen Zahlbereiche in einander liegen.

1.2 Formales Rechnen mit Variablen

In diesem Abschnitt werden wir eine formalere Sprache einführen und die Ideen aus Abschnitt 1.1 in dieser formulieren. Zuerst wenden wir uns dem Begriff der Variable zu:

Dazu zurück zur Division. Dort hatten wir uns gefragt welche Zahl man mit z.B. 5 multiplizieren muss, um zur 10 zu kommen. Diese (noch) unbekannte Zahl bezeichnen wir mit einem Buchstaben, z.B. x . Dann lautet die Frage also, welche reelle Zahl x erfüllt

$$5 \cdot x = 10.$$

In diesem Fall wäre $x = 2$ die Lösung. Wir können aber solche Buchstaben nicht nur als Platzhalter für eine Lösung verwenden, sondern auch als eine variable Zahl (deswegen der Name Variable). Falls wir zum Beispiel mehrere Unbekannte in einer Gleichung haben, können wir uns fragen was bei der einen Variable herauskommt, wenn wir die andere auf einen bestimmten Wert festlegen. Dazu ein Beispiel: Nehmen wir uns zwei reelle Variablen x und y her und diese sollen folgende Gleichung erfüllen:

$$\frac{2 \cdot x}{y} = 10.$$

Falls wir nun x festlegen möchten, wäre y unsere zu suchende Unbekannte. Wir können die Gleichung oben umstellen, indem wir zuerst mit y auf beiden Seiten multiplizieren:

$$\begin{aligned} \frac{2 \cdot x}{y} = 10 & \quad | \cdot y \\ \Rightarrow 2 \cdot x = 10 \cdot y & \quad | \cdot \frac{1}{10} \\ \Rightarrow \frac{2 \cdot x}{10} = y & \Rightarrow \frac{x}{5} = y. \end{aligned}$$

Die Pfeile \Rightarrow stehen dabei für eine logische Folgerung, d.h. in Worten, dass wir z.B. aus der ersten Gleichung die zweite folgern, indem wir mit y multiplizieren. Wir können unser Ergebnis sogar grafisch in einem kartesischen Koordinatensystem darstellen! Siehe dazu Bild 1.2.

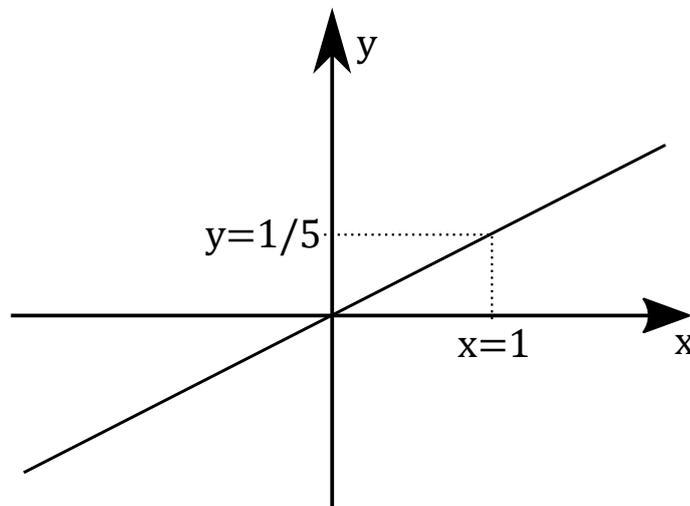


Abbildung 1.2: Skizze von $y = \frac{x}{5}$.

Nun können wir x beliebig wählen und erhalten so ein y . Z.B. $x = 1$ ergibt

$$y = \frac{2 \cdot x}{10} = \frac{2 \cdot 1}{10} = \frac{2}{10} = \frac{1}{5}.$$

Man sagt, wir haben die Gleichung nach y aufgelöst. Nun können wir das Ganze

aber auch herumdrehen und nach x auflösen:

$$\begin{aligned}\frac{2 \cdot x}{y} &= 10 & | \cdot y \\ \Rightarrow 2 \cdot x &= 10 \cdot y & | \cdot \frac{1}{2} \\ \Rightarrow x &= \frac{10 \cdot y}{2} = \frac{5 \cdot y}{1} = 5 \cdot y.\end{aligned}$$

Hier können wir nun y beliebig aus den reellen Zahlen wählen, oder? Nein, denn wenn wir die ursprüngliche Gleichung $\frac{2 \cdot x}{y} = 10$ lösen wollen, dürfen wir $y = 0$ nicht wählen, denn sonst würden wir durch 0 teilen.

Mit diesem Formalismus können wir nun die eher informell gestalteten Rechenregel aus Abschnitt 1.1 rekapitulieren: Dafür nehmen wir uns drei reelle Zahlen x, y, z . Für die Addition haben wir folgende zentrale Rechenregeln:

$$\begin{aligned}x + y &= y + x, \text{ kommutativ} \\ (x + y) + z &= x + (y + z), \text{ assoziativ.}\end{aligned}$$

Bei der Multiplikation sieht dies genauso aus:

$$\begin{aligned}x \cdot y &= y \cdot x, \text{ kommutativ} \\ (x \cdot y) \cdot z &= x \cdot (y \cdot z), \text{ assoziativ.}\end{aligned}$$

Spannender ist da schon eher die Interaktion von Multiplikation und Addition. Die folgende Regel nennt man Distributivgesetz oder informeller ausmultiplizieren/ausklammern:

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Wie verhält sich das nun, wenn man statt x ebenfalls eine Summe nimmt? Sei dafür w eine weitere reelle Zahl. Dann müssen wir nur konsequent ausmultiplizieren:

$$\begin{aligned}(x + y) \cdot (z + w) &= (x + y) \cdot z + (x + y) \cdot w = z \cdot (x + y) + w \cdot (x + y) \\ &= z \cdot x + z \cdot y + w \cdot x + w \cdot y.\end{aligned}$$

So bekommt man auch das Verhalten von negativen Summen, Stichwort *Minusklammer*:

$$-(x + y) = (-1) \cdot (x + y) = (-1) \cdot x + (-1) \cdot y = -x - y.$$

Beim Dividieren gehen wir wieder zu den Brüchen. Hier muss man nun aufpassen nicht durch 0 zu teilen, d.h. wir fordern zusätzlich, dass $y, w \neq 0$. Dann sind unsere Rechenregeln wie folgt

$$\begin{aligned}\frac{x}{y} + \frac{z}{w} &= \frac{x \cdot w + z \cdot y}{y \cdot w}, \text{ Addition von Brüchen} \\ \frac{x}{y} \cdot \frac{z}{w} &= \frac{x \cdot z}{y \cdot w}, \text{ Multiplikation von Brüchen.}\end{aligned}$$

Zum Schluss dieses Abschnitts besprechen wir noch kurz *Äquivalenzumformungen* von Gleichungen. Zum Beispiel nehmen wir uns eine Gleichung wie die folgende her

$$5 \cdot \frac{x}{3} + 2y = 10.$$

Dann können wir verschiedene Operationen durchführen um z.B. nach x oder y aufzulösen. Wir können beliebige Zahlen auf beide Seiten subtrahieren, z.B.

$$\begin{aligned} 5 \cdot \frac{x}{3} + 2 \cdot y = 10 & \quad | - 2 \cdot y \\ \Leftrightarrow 5 \cdot \frac{x}{3} = 10 - 2 \cdot y. \end{aligned}$$

Da wir die Operation rückgängig machen könnten, indem wir $2 \cdot y$ auf beide Seiten addieren, verlieren wir keine Informationen! Immer wenn dies der Fall ist, sprechen wir von einer sogenannten Äquivalenzumformung und schreiben deswegen den Äquivalenzpfeil \Leftrightarrow .

Ein Beispiel für eine reine Folgerung, wäre es, wenn wir eine Gleichung 'quadrieren' würden. Nehmen wir dafür $x = -5$. Wenn wir diese Gleichung mit x durchmultiplizieren, können wir x auf der rechten Seite mit -5 ersetzen und erhalten:

$$\begin{aligned} x = -5 & \quad | \cdot x \\ \Rightarrow x \cdot x = -5 \cdot x = (-5) \cdot (-5) = 25. \end{aligned}$$

Die Gleichung $x \cdot x = 25$ wird von $x = 5$ und $x = -5$ erfüllt, d.h. wir können nicht mehr auf die ursprüngliche Gleichung schließen! Deswegen nutzen wir hier nur den Folgerungspfeil \Rightarrow .

1.3 Summen und Produktzeichen

Hier wollen wir uns einer Notation widmen, wenn wir viele Summanden oder Faktoren haben. Z.B.

$$1 + 2 + 3 + \dots + n$$

Diese Art von Ausdrücken wollen wir nun kompakter schreiben. Seien dazu a_1, \dots, a_n reelle Zahlen. Dann setzen wir

$$\sum_{i=1}^n a_i := a_1 + \dots + a_n.$$

Im unteren Teil Index des Summenzeichens Σ steht dabei der Startindex, während im oberen Bereich der Endindex steht.

Im obigen Beispiel $1 + 2 + \dots + n$ gälte $a_i = i$, d.h. dies hat folgende Gestalt

$$1 + 2 + \dots + n = \sum_{i=1}^n i.$$

Für diese Notation haben wir wieder ein Distributivgesetz. Sei also zusätzlich x eine reelle Zahl, dann gilt

$$x \cdot \sum_{i=1}^n a_i = \sum_{i=1}^n x \cdot a_i.$$

Im obigen Beispiel mit $x = 2$:

$$2(1 + 2 + \dots + n) = 2 \sum_{i=1}^n i = \sum_{i=1}^n 2i = 2 + 4 + 6 + \dots + 2 \cdot n.$$

Ähnliches können wir auch für Produkte machen, z.B.

$$2 \cdot 4 \cdot 6 \dots \cdot 2n$$

Seien also wie vorher a_1, \dots, a_n reelle Zahlen. Dann definieren wir

$$\prod_{i=1}^n a_i := a_1 \cdot \dots \cdot a_n.$$

Wie auch beim Summenzeichen Σ steht hier der untere Index für den Anfang und der obere Index für das Ende des Produkts.

Im Beispiel $2 \cdot 4 \cdot 6 \dots \cdot 2n$ wären die $a_i = 2 \cdot i$. Mit dem Produktzeichen Π gilt dann also

$$2 \cdot 4 \cdot \dots \cdot 2n = \prod_{i=1}^n 2 \cdot i.$$

Unter Umständen kann es sein, dass man bei dem Summen (oder Produktzeichen) einen anderen Startindex (oder Endindex) schreiben möchte, um sich z.B. einen gewissen Zusammenhang zu verdeutlichen. Dies können wir mit dem sogenannten *Indexshift* angehen: Gehen wir dafür wieder von

$$\sum_{i=0}^n a_i$$

aus. Wenn ich nun bei 5 starten möchte, schiften wir den Index, indem wir einen neuen Index $k = i + 5$ einführen. Bei Start $i = 0$ hätten wir damit $k_{start} = 5$, also unser Ziel. Wie transformieren sich nun die anderen beteiligten Größen? Beim Index von a_i können wir nach k auflösen und erhalten $i = k - 5$. Beim Zielindex setzen wir für $i = n$ ein und erhalten $k_{ende} = n + 5$. Insgesamt erhalten wir also

$$\sum_{i=0}^n a_i = \sum_{k=5}^{n+5} a_{k-5}.$$

Nun wäre es lästig für jeden solcher Indexshifts einen neuen Index einzuführen, also nutzen wir einfach den alten weiter und erhalten so

$$\sum_{i=0}^n a_i = \sum_{i=5}^{n+5} a_{i-5}.$$

Eine weitere interessante Manipulation eines Summenzeichens ist die *Umkehrung der Summationsreihenfolge*. Schließlich wissen wir, dass für alle reellen x, y gilt $x + y = y + x$. Dies verallgemeinern wir nun auf das Summenzeichen: Hierbei müssen wir den Index beim a_i korrigieren und nun von oben herab zählen:

$$\sum_{i=0}^n a_i = \sum_{k=0}^n a_{n-k}.$$

Alle Summanden bleiben erhalten, wir haben lediglich die Reihenfolge vertauscht. Nun können wir wieder k mit i ersetzen und erhalten

$$\sum_{i=0}^n a_i = \sum_{i=0}^n a_{n-i}.$$

1.4 Potenzen und Wurzeln

Wir haben bereits die Summe von vielen gleichen Summanden abgekürzt durch Multiplikation. Dasselbe tun wir nun mit einem Produkt vieler gleicher Faktoren. Sei dafür x eine reelle Zahl und n eine natürliche Zahl.

$$\underbrace{x \cdot \dots \cdot x}_{n\text{-mal}} =: x^n.$$

Der Zusammenhang zum Produktzeichen ist wie folgt:

$$\prod_{i=1}^n x = \underbrace{x \cdot \dots \cdot x}_{n\text{-mal}} = x^n,$$

man beachte dabei, dass die Variable im Produktzeichen nicht von i abhängt. Insgesamt nennt man dies Potenzieren und sagt dafür x hoch n . Ein Beispiel dafür mit $x = 3$ und $n = 4$:

$$3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = 81.$$

Wir setzen die Potenz hoch 0 fest durch

$$x^0 = 1.$$

und falls der Exponent negativ ist, wird das x umgekehrt:

$$x^{-n} = \frac{1}{x^n}.$$

Die typischen Rechengesetze lauten nun

$$x^n \cdot x^m = \underbrace{x \cdot \dots \cdot x}_{n\text{-mal}} \cdot \underbrace{x \cdot \dots \cdot x}_{m\text{-mal}} = \underbrace{x \cdot \dots \cdot x}_{n+m\text{-mal}} = x^{n+m}$$

An diesem Gesetz erkennt man auch, warum man $x^0 = 1$ und $x^{-n} = \frac{1}{x^n}$ gesetzt hat:

$$1 = \frac{x^n}{x^n} = x^n \cdot x^{-n} = x^{n-n} = x^0.$$

Falls der Exponent gleich ist, bekommen wir folgende Gesetzmäßigkeit

$$x^n \cdot y^n = \underbrace{x \cdot \dots \cdot x}_{n\text{-mal}} \cdot \underbrace{y \cdot \dots \cdot y}_{n\text{-mal}} = \underbrace{x \cdot y \cdot \dots \cdot x \cdot y}_{n\text{-mal}} = (x \cdot y)^n$$

und für Potenzen von Potenzen gilt

$$(x^n)^m = \underbrace{x^n \cdot \dots \cdot x^n}_{m\text{-mal}} = x^{\sum_{i=1}^m n} = x^{n \cdot m}$$

Vorsicht hier beim Klammern, die folgenden Ausdrücke sind im Allgemeinen nicht gleich:

$$x^{(n^m)} \neq (x^n)^m.$$

Dafür ein einfaches Beispiel: $x = 2$, $n = 2$, $m = 3$:

$$2^{(2^3)} = 2^8 = 256$$

$$(2^2)^3 = 4^3 = 64.$$

Falls man sich nicht mehr sicher ist, welche Klammer für unser Potenzrechengesetz korrekt ist, so bietet es sich an, schnell ein Beispiel zu rechnen!

Wie auch beim Addieren und Multiplizieren wollen wir diese Operation umkehren. Nehmen wir uns also ein gegebene natürliche Zahl n und ein nichtnegatives reelles y . Wir suchen nun die Zahl x , die folgende Gleichung erfüllt:

$$x^n = y.$$

Für die *nichtnegative* Lösung dieses Problems nutzen wir folgende Notationen:

$$x = \sqrt[n]{y} = y^{\frac{1}{n}}.$$

Machen wir ein einfaches Beispiel dazu: Sei $n = 3$, und $y = 8$. D.h. um $\sqrt[3]{8}$ zu berechnen, fragen wir uns welche Zahl x erfüllt $x^3 = 8$. Nach etwas herumprobieren kommen wir auf die Zahl $x = 2$, denn $2^3 = 2 \cdot 2 \cdot 2 = 4 \cdot 2 = 8$.

Warum schreiben wir $y^{\frac{1}{n}}$? Mit dieser Notation sind die obigen Rechengesetze für Potenzen weiter vollständig gültig! Damit ist dann auch klar wie eine Potenz zu einer rationalen Zahl definiert ist:

$$x^{\frac{n}{m}} = (x^n)^{\frac{1}{m}} = (x^{\frac{1}{m}})^n$$

Über Approximationen (sei hier nur erwähnt, dies wird im ersten Semester genauer beleuchtet) können wir die Potenz auch für reelle Exponenten einführen. Alle Rechengesetze bleiben dabei erhalten!

1.5 Binomische Formeln und quadratische Gleichungen

Als Anwendung von Abschnitt 1.4 wollen wir nun die binomischen Formeln herleiten: Seien dafür a, b reelle Zahlen. Dann gilt

$$\begin{aligned} (a + b)^2 &= (a + b) \cdot (a + b) = a \cdot (a + b) + b \cdot (a + b) = a^2 + a \cdot b + b \cdot a + b^2 \\ &= a^2 + 2 \cdot a \cdot b + b^2 \end{aligned}$$

Dies nennen wir erste binomische Formel. Was passiert wenn wir subtrahieren, anstatt zu addieren?

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

So erhalten wir die zweite binomische Formel. Die dritte binomische Formel funktioniert wie folgt

$$(a + b) \cdot (a - b) = a \cdot (a - b) + b \cdot (a - b) = a^2 - a \cdot b + b \cdot a - b^2 = a^2 - b^2.$$

Zusammengefasst also

$$(a + b)^2 = a^2 + 2 \cdot a \cdot b + b^2, \text{ erste binomische Formel}$$

$$(a - b)^2 = a^2 - 2 \cdot a \cdot b + b^2, \text{ zweite binomische Formel}$$

$$(a + b) \cdot (a - b) = a^2 - b^2, \text{ dritte binomische Formel.}$$

Dies können wir nun nutzen um sogenannte quadratische Gleichungen zu lösen. Das gestellte Problem lautet: Suche eine reelle Zahl x , welche für gegebene reelle Zahlen p, q die Gleichung

$$x^2 + p \cdot x + q = 0$$

löst. Wir gehen hier mit der sogenannten *quadratischen Ergänzung* vor, d.h. unser Ziel ist es zur ersten binomischen Formel zu kommen (im Allgemeinen nennt man solche Tricks 'nährhafte Nullergänzungen'):

$$\begin{aligned} 0 &= x^2 + p \cdot x + q = x^2 + 2 \cdot \frac{p}{2} \cdot x + \left(\frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2 \\ &= \left(x + \frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2 \\ \Leftrightarrow \left(x + \frac{p}{2}\right)^2 &= \left(\frac{p}{2}\right)^2 - q \end{aligned}$$

Wir haben bereits an einem Beispiel am Ende von Abschnitt 1.2 gesehen, dass beim Wurzelziehen wir es mit einer positiven, sowie negativen Lösung zu tun bekommen könnten:

$$\begin{aligned} \Leftrightarrow x + \frac{p}{2} &= \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} \\ \Leftrightarrow x &= -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} \end{aligned}$$

Diese Formel wird *pq-Formel*, oder *Mitternachtsformel* genannt. Eine interessante Beobachtung liefert die sogenannte *Diskriminante* $\left(\frac{p}{2}\right)^2 - q$. Ist diese negativ, besitzt die Gleichung keine reelle Lösung, ist sie positiv, so besitzt sie zwei unterschiedliche Lösungen und falls sie gleich 0 ist, so gibt es genau eine Lösung.

Der allgemeinere Fall der Gleichung

$$a \cdot x^2 + b \cdot x + c = 0$$

mit reellen a, b, c kann auf den Fall der *pq-Formel* zurückgeführt werden, sofern $a \neq 0$ ist (der Fall $a = 0$ ist sowieso deutlich einfacher). In diesem Fall können wir durch a die Gleichung teilen und setzen $p = \frac{b}{a}$ und $q = \frac{c}{a}$ und erhalten

$$0 = x^2 + \frac{b}{a} \cdot x + \frac{c}{a} = x^2 + p \cdot x + q.$$

Darauf wenden wir nun unsere *pq-Formel* an und dies ergibt mit etwas erweitern und Wurzelziehen

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4 \cdot a^2} - \frac{c}{a}} = \frac{-b \pm \sqrt{b^2 - 4 \cdot a \cdot c}}{2 \cdot a}.$$

Diese Formel wird ebenfalls *Mitternachtsformel* genannt.

1.6 Logarithmen

Wir haben in Abschnitt 1.4 das Potenzieren invertiert, d.h. wir haben uns gefragt bei gegebenen Zahlen x und y , welcher Zahl a die Gleichung

$$a^x = y$$

erfüllt. Hier verändern wir die Frage und suchen nun stattdessen die Zahl x bei gegebenen positiven a und y . Die Lösung dazu nennen wir Logarithmen, in abstrakter Schreibweise:

$$\log_a y = x.$$

Den Wert a nennen wir dabei Basis. Anders ausgedrückt, gilt immer

$$a^{\log_a y} = y. \quad (1.1)$$

Machen wir ein schnelles Beispiel dazu: Was ist das Ergebnis von $\log_2 4$? Um das zu lösen, fragen wir uns also welche Zahl x löst

$$2^x = 4.$$

Nach etwas herumprobieren, stellen wir fest, dass das Ergebnis $x = 2$ ist. Also gilt

$$\log_2 4 = 2.$$

Eine besondere Bedeutung kommt nun der Eulerschen Zahl $e = 2.71\dots$ bei. Falls diese als Basis verwendet wird, sprechen wir von dem *natürlichen Logarithmus* und schreiben oftmals

$$\ln y := \log_e y.$$

Aber Vorsicht, in der Literatur kann es auch vorkommen, dass lediglich die Basis weggelassen wird. Im Zweifel sollte vorher überprüft werden, welche Basis nun benutzt wird, denn es kann auch vorkommen, dass damit der Logarithmus zur Basis 10 gemeint ist, d.h. $\log_{10} x = \log x$.

Kommen wir nun zu den wichtigen Rechenregel für Logarithmen: Diese lassen sich alle aus der obigen Definition bzw. Gleichung 1.1 herleiten. Wir fragen uns zuerst, was die Summe zweier Logarithmen wieder ergibt:

$$\log_a(x) + \log_a(y) = \log_a(x \cdot y).$$

Warum ist dies nun der Fall? Schreiben wir alles in Potenzform um und nutzen unsere Potenzrechengesetze wird es klarer:

$$a^{\log_a(x \cdot y)} = x \cdot y = a^{\log_a x} \cdot a^{\log_a y} = a^{\log_a x + \log_a y}.$$

Ein häufiger Fehler ist es nun zu verwechseln wo summiert und wo multipliziert wird. Es gilt im Allgemeinen nicht:

$$\log_a(x) \cdot \log_a(y) = \log_a(x + y).$$

Eine Überprüfung an einem Beispiel zeigt schnell, welches Rechengesetz nun richtig ist. Wir wählen $a = 2$, $x = 2$ und $y = 2$. Denn hier gilt

$$\log_2(2 + 2) = \log_2(4) = 2 \neq 1 = 1 \cdot 1 = \log_2(2) \cdot \log_2(2).$$

Aber was können wir denn nun erwarten, sollten wir mit eine Zahl mit einem Logarithmus multiplizieren?

$$y \cdot \log_a x = ?$$

Dafür nutzen wir wieder unsere Potenzrechengesetze und nehmen den linken Ausdruck in den Exponenten von a :

$$a^{y \log_a x} = (a^{\log_a x})^y = x^y = a^{\log_a(x^y)}.$$

Also gilt

$$y \cdot \log_a x = \log_a(x^y).$$

Nun können wir uns überlegen, was passiert, wenn wir in unserer ersten Regel dividieren bzw. subtrahieren:

$$\log_a \left(\frac{x}{y} \right) = \log_a(xy^{-1}) = \log_a(x) + \log_a(y^{-1}) = \log_a(x) - \log_a(y).$$

Eine wichtige Frage ist aber nun noch offen geblieben. Wie verhalten sich Logarithmen zu verschiedenen Basen? Setzen wir dafür $y = \log_b x$. Wir erhalten:

$$\begin{aligned} b^y &= x \\ \Rightarrow \log_a(b^y) &= \log_a x \\ \Rightarrow y \log_a b &= \log_a x \\ \Rightarrow \log_b x &= y = \frac{\log_a x}{\log_a b}. \end{aligned}$$

So können wir zu beliebigen Basen wechseln!

Fassen wir unsere Rechenregel zusammenfassen:

$$\log_a(x \cdot y) = \log_a x + \log_a y \tag{1.2}$$

$$\log_a(x^y) = y \cdot \log_a(x) \tag{1.3}$$

$$\log_a \left(\frac{x}{y} \right) = \log_a x - \log_a(y) \tag{1.4}$$

$$\log_a x = \frac{\log_b x}{\log_b a} \tag{1.5}$$

Die letzte Regel hat eine weitere interessante Anwendung. Oben hatten wir uns bereits überlegt, dass $\log_2 8 = 3$. Damit können wir nun wie folgt arbeiten um $\log_8 2$ zu berechnen, indem wir in der letzten Regel $a = 8$, $x = b = 2$ setzen:

$$\log_8 2 = \frac{\log_2 2}{\log_2 8} = \frac{1}{\log_2 8} = \frac{1}{3}.$$

Schauen wir uns noch den typischen Verlauf des Ausdrucks $\log_a x$ an, so dass wir das x als Variable betrachten. Zuerst beobachten wir mit $a^0 = 1$, dass $\log_a 1 = 0$ ist. Dies gilt unabhängig von der Basis! Weiter haben wir damit

$$\log_a \left(\frac{1}{x} \right) = \log_a 1 - \log_a(x) = -\log_a(x),$$

also ändert sich das Vorzeichen bei $x = 1$. Für $a = 2$ sieht der typische Verlauf nun aus wie in Abbildung 1.3.

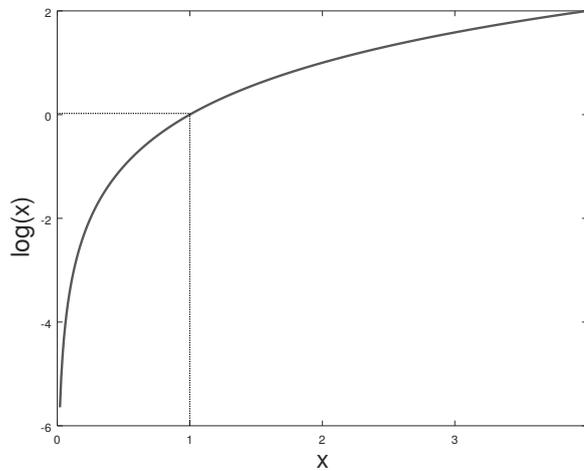


Abbildung 1.3: Skizze von $\log_2 x$.

1.7 Größenbeziehungen

Wir führen hier die Symbole $<$, \leq , $>$ und \geq ein. Diese sortieren reelle Zahlen der Größe nach. Zum Beispiel ist 5 echt größer als 2. In dem Fall schreiben wir

$$5 > 2.$$

Wir können dies auch andersherum interpretieren und sagen 2 ist echt kleiner als 5. Dann schreiben wir

$$2 < 5.$$

Falls wir uns nicht sicher sind, ob die beteiligten Größen nicht vielleicht sogar gleich sein könnten (z.B. bei Variablen) nutzen wir die Symbole \leq (kleiner gleich) und \geq (größer gleich). Geometrisch kann man sich dies am Zahlenstrang verdeutlichen, siehe Abbildung 1.4 Nun wollen wir natürlich Ungleichungen ma-

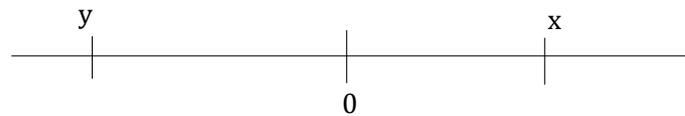


Abbildung 1.4: Zahlenstrang mit $y < x$.

nipulieren, z.B. mit Zahlen multiplizieren, den Kehrwert nehmen etc. Wie die resultierende Ungleichung aussieht kann man sich am Graphen der Operation merken. Dazu ein Beispiel: Was passiert, wenn wir die Ungleichung $y < x$ mit -1 multiplizieren? Der Verlauf von $-x$ über der x -Achse ist *strikt fallend* (siehe Abbildung 1.5), also erhalten wir

$$-y > -x$$

Was passiert wenn wir beide Seite der Ungleichung invertieren, d.h. den Keh-

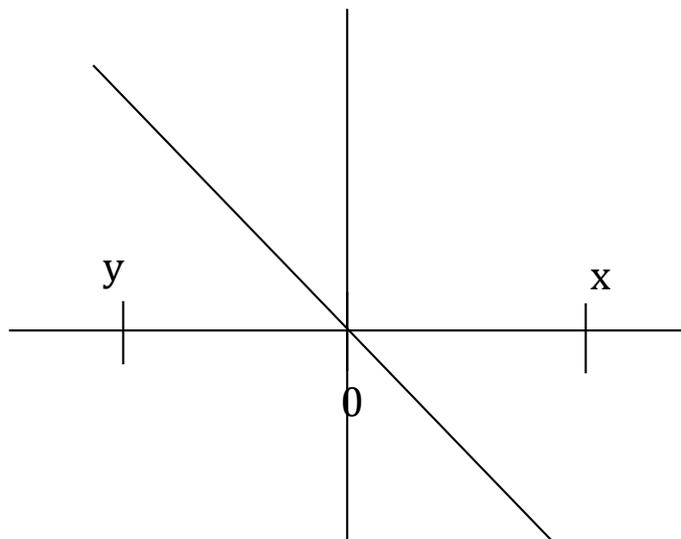


Abbildung 1.5: Graph von $-x$.

wert bilden? Die zugehörige Operation ist $\frac{1}{x}$ und die dazugehörige Skizze 1.6 verdeutlicht, dass sich das Ungleichungszeichen umkehrt. Also erhalten wir

$$\frac{1}{y} > \frac{1}{x}$$

wenn wir $y < x$ voraussetzen.

Die Faustregel lautet also wie folgt:

Die Ungleichung bleibt erhalten wenn der zugehörige Graph strikt wachsend ist.

Die Ungleichung dreht sich um wenn der zugehörige Graph strikt fallend ist.

Ein Beispiel für den Erhalt der Ungleichung ist z.B. das Wurzelziehen nichtnegativer Zahlen. Dort haben wir

$$0 \leq y < x \Rightarrow 0 \leq \sqrt{y} < \sqrt{x}$$

Folgende Tabelle fasst das Verhalten für $y < x$ zusammen:

aus $y < x$ folgt:

$$-y > -x$$

$$x, y \neq 0 : \frac{1}{y} > \frac{1}{x}$$

$$0 \leq x, y : \sqrt{y} < \sqrt{x}$$

$$0 < x, y : \log_2(y) < \log_2(x).$$

Ein wichtiger Merksatz dabei ist der folgende Zusammenhang spezialisiert auf die Multiplikation/Kehrwertbildung:

Bei Multiplikation mit negativen Zahlen sowie Kehrwertbildung kehrt sich das

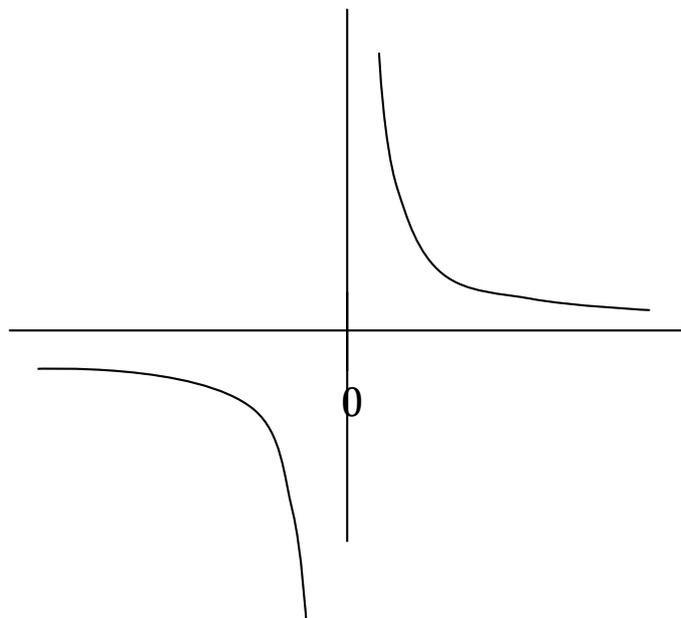


Abbildung 1.6: Graph von $\frac{1}{x}$.

Ungleichungszeichen um!

Eine wichtige Anwendung für solche Ungleichungen ist der *Betrag einer Zahl*. Mit dem Betrag einer Zahl wollen wir den Abstand zur 0 auf dem Zahlenstrahl ausrechnen. Dies geht nun wie folgt

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

In Worten ausgedrückt: Ist eine Zahl negativ, so streicht der Betrag das Minuszeichen. Ist eine Zahl nichtnegativ, so beeinflusst der Betrag diese Zahl nicht. Insbesondere gilt für alle reellen Zahlen x

$$|x| \geq 0.$$

Was passiert nun, wenn ich zwei Zahlen addiere und danach den Betrag ausrechne? Wie verhalten sich dazu die Beträge der ursprünglichen Summanden? Machen wir ein Beispiel:

$$|-6 + 5| = |-1| = 1, \quad |-6| + |5| = 6 + 5 = 11.$$

Es scheint also so zu sein, dass eine Ungleichung herrscht. Dies ist korrekt und zwar die sogenannte *Dreiecksungleichung*:

$$|x + y| \leq |x| + |y|.$$

Eine weitere Möglichkeit den Betrag auszudrücken ist über das Quadrat und anschließendes Wurzelziehen (Wir hatten festgelegt, dass Wurzeln immer nichtnegativ sein sollen!):

$$|x| = \sqrt{x^2}.$$

Das Quadrieren hatten wir im Zusammenhang mit Ungleichungen noch gar nicht betrachtet: Der Graph zu x^2 sieht aus wie in Abbildung 1.7. Dieser ist

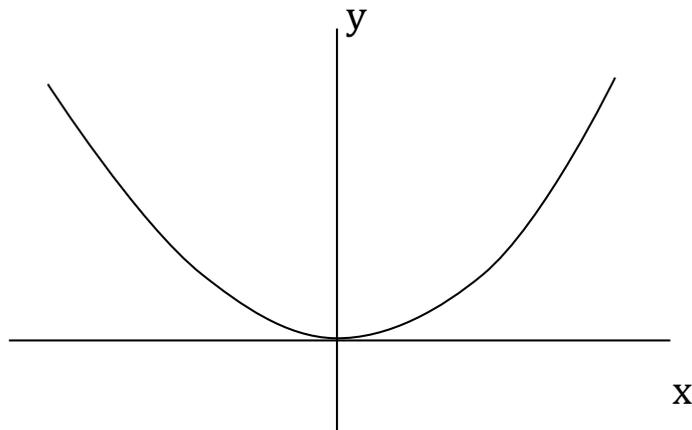


Abbildung 1.7: Graph von x^2 .

strikt fallend bis zur 0 und steigt danach. D.h. für beliebige Zahlen x, y mit $y < x$ können wir ohne weitere Informationen keine Aussage dazu machen ob x^2 oder y^2 größer ist. Z.B. $y = -5$ und $x = 1$ ergibt $y^2 = 25 > 1 = x^2$, aber $y = 2$ und $x = 5$ ergibt $y^2 = 4 < 25 = 5^2 = x^2$. Wir können aber eine Aussage machen, wenn die Vorzeichen von x und y gleich sind:

$$\begin{aligned} y < x \leq 0 &\Rightarrow y^2 > x^2 \\ 0 \leq y < x &\Rightarrow y^2 < x^2. \end{aligned}$$

Außerdem ist das Vorzeichen des Quadrates immer gleich, d.h. wir haben für alle reellen Zahlen x

$$0 \leq x^2.$$

Solch ein Verhalten gilt dabei prinzipiell auch für den Betrag.

Um unsere Ungleichungsmethoden zu verdeutlichen, lösen wir nun eine Ungleichung im Beispiel. Wir fragen uns welche reellen x erfüllen folgende Ungleichung?

$$(-2) \cdot 3^x < -6$$

Wir formen um, indem wir zuerst mit $-\frac{1}{2}$ durchmultiplizieren (Ungleichung kehrt sich um) und danach \log_3 (Ungleichung kehrt sich nicht um) anwenden:

$$\begin{aligned} (-2) \cdot 3^x &< -6 \\ \Leftrightarrow 3^x &> 3 \\ \Leftrightarrow \log_3(3^x) &> \log_3 3 \\ \Leftrightarrow x &> \log_3 3 = 1. \end{aligned}$$

Also erfüllen alle $x > 1$ die obige Ungleichung.

Häufig möchten wir auch Brüche vergleichen. Dort ist es nicht immer so klar, welcher von zwei Brüchen größer als ein anderer ist, falls man keinen Taschenrechner zur Hand hat. Folgendes Verfahren hilft aber: Seien also a, b, c, d

natürliche Zahlen ungleich null. Wie können wir also feststellen ob $\frac{a}{b}$ oder $\frac{c}{d}$ größer ist? Wir formulieren das Problem äquivalent um, um nur noch natürliche Zahlen zu haben:

$$\frac{a}{b} \leq (\geq) \frac{c}{d} \mid \cdot b, \cdot d \\ \Leftrightarrow a \cdot d \leq (\geq) c \cdot b.$$

Die Vergleichszeichen \leq bzw. \geq drehen sich nicht um, denn alle beteiligten Zahlen sind positiv. Ein Beispiel:

$$\frac{3}{8} < \text{oder} > \frac{7}{15}?$$

Wir multiplizieren mit den Nennern durch und erhalten:

$$\frac{3}{8} \cdot 8 \cdot 15 < \text{oder} > \frac{7}{15} \cdot 8 \cdot 15 \\ \Leftrightarrow 3 \cdot 15 < \text{oder} > 7 \cdot 8 \\ \Leftrightarrow 45 < \text{oder} > 56$$

Da $45 < 56$ gilt also $\frac{3}{8} < \frac{7}{15}$.

2 Mengenlehre

2.1 Was ist eine Menge?

Wir werden hier die Grundzüge der Mengenlehre wiederholen. Was ist eine Menge? Eine Menge ist einfach ausgedrückt eine Sammlung von Objekten, z.B. die Menge aller Lösungen einer quadratischen Gleichung, siehe Abschnitt 1.5. Diese Objekte nennen wir *Elemente*. Zur Notation: Sei A eine Menge und a ein Element dieser Menge, d.h. in A enthalten. Dann schreiben wir

$$a \in A.$$

Falls das Element b nicht zu A gehört, so schreiben wir

$$b \notin A.$$

Dies können wir in einem sogenannten *Mengendiagramm* verdeutlichen, siehe Bild 2.1.

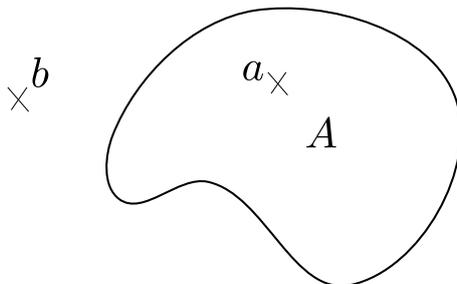


Abbildung 2.1: Beispiel eines Mengendiagramms.

Ein paar explizitere Beispiele dazu: Unsere Zahlenbereiche aus Abschnitt 1 sind Mengen, d.h. die natürlichen (\mathbb{N}), ganzen (\mathbb{Z}), rationalen (\mathbb{Q}) und reellen (\mathbb{R}) Zahlen sind jeweils Mengen. Dann gilt, da 1 eine natürliche Zahl ist

$$1 \in \mathbb{N}.$$

Weiter gilt z.B., dass echte Brüche keine ganzen Zahlen sind, z.B.:

$$\frac{1}{5} \notin \mathbb{Z} \text{ aber } \frac{1}{5} \in \mathbb{Q}.$$

Ein komplexeres Beispiel wäre $\sqrt{2} \notin \mathbb{Q}$. Dies werden wir hier aber aus Zeitgründen nicht genauer behandeln. Schauen wir uns unsere Zahlbereiche genauer an, so scheint es eine Hierarchie zu geben. Die natürlichen Zahlen sind in den ganzen Zahlen enthalten und so weiter (siehe auch Bild 1.1). Dieses Konzept nennt man *Teilmengenbeziehung*. Allgemeiner sagen wir, dass eine Menge A *Teilmenge* einer Menge B ist, falls alle Elemente von A auch in B enthalten sind. Etwas kürzer:

Für alle $a \in A$ gilt $a \in B$.

In diesem Falle schreiben wir

$$A \subset B$$

oder

$$B \supset A.$$

Im Beispiel unserer Zahlbereiche sieht dies nun wie folgt aus:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Wir können auch Mengen angeben, indem alle Elemente dieser Menge hinschreiben. Falls wir z.B. in der Menge A die Zahlen 1, 2, 3 zusammenfassen wollen, schreiben wir mit geschweiften Klammern

$$A := \{1, 2, 3\}.$$

Dieses Konzept können wir erweitern, indem wir Operationen und Einschränkungen in eine solche Mengenklammer setzen. Falls wir z.B. alle rationalen Zahlen echt größer als 5 in der Menge A zusammenfassen wollen, so schreiben wir

$$A := \{x \in \mathbb{Q} \mid x > 5\}.$$

Die Notation ist dabei wie folgt: Erst geben wir eine Grundmenge an (in diesem Fall \mathbb{Q}) und geben danach getrennt durch $|$ eine Reihe von Bedingungen an, die ebenfalls erfüllt sein sollen, (in diesem Fall $x > 5$). Es kann sein, dass in gewisser Literatur : statt $|$ verwendet wird.

Falls wir nun zwei Mengen A und B haben, welche unterschiedlich beschrieben werden, wir aber die Vermutung haben, dass diese gleich sind, so können wir dies über das Teilmengenkriterium überprüfen:

$$A = B \Leftrightarrow A \subset B \text{ und } B \subset A$$

Dazu ein Beispiel:

$$A := \{x \in \mathbb{R} \mid x > 5\} \text{ und } B := \{x \in \mathbb{R} \mid 2 \cdot x > 10\}.$$

Falls also $x \in A$, so gilt für dieses x die Ungleichung $x > 5$. Multiplizieren wir diese Ungleichung mit 2 erhalten wir $2 \cdot x > 10$. Also gilt $x \in B$. Da x beliebig war, gilt also $A \subset B$. Falls andersherum $x \in B$ so haben wir $2 \cdot x > 10$. Dividieren mit 2 gibt uns $x > 5$ und damit $x \in A$. Also $B \subset A$. Insgesamt also $A = B$.

Zuletzt in diesem Abschnitt führen wir noch den Begriff der *leeren Menge* ein. Dies ist die Menge, die keine Elemente enthält. Wir bezeichnen sie mit

$$\emptyset.$$

Da sie keine Elemente besitzt, ist sie eine Teilmenge zu allen Mengen, d.h. für allen Mengen A gilt

$$\emptyset \subset A.$$

2.2 Mengenoperationen

In diesem Abschnitt führen wir Operationen mit Mengen ein, um neue Mengen zu erzeugen. Dafür nehmen wir uns zwei Mengen A und B her. Der *Schnitt* dieser beiden Mengen ist die Menge, die die Elemente enthält, die gleichzeitig in A und B enthalten sind (siehe Skizze 2.2). In Formel schreiben wir für die

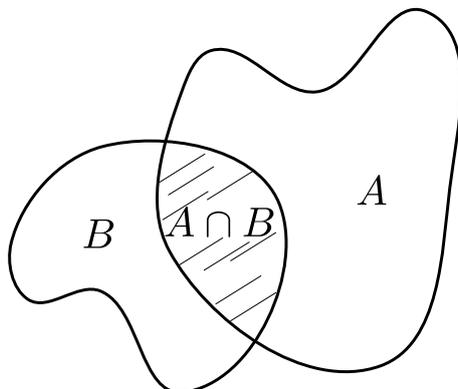


Abbildung 2.2: Schnitt von zwei Mengen.

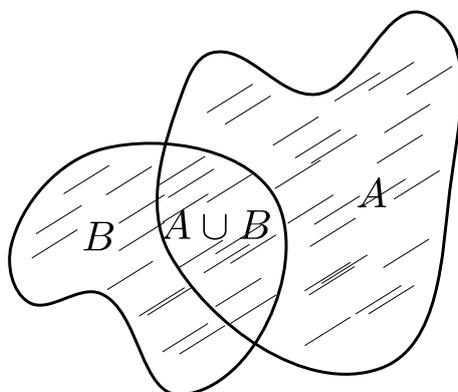


Abbildung 2.3: Vereinigung von zwei Mengen.

Schnittmenge

$$A \cap B := \{x \mid x \in A \text{ und } x \in B\}.$$

Die *Vereinigung* von A und B hingegen beschreibt die Menge aller Elemente die in A oder B liegen (nicht exklusiv!), siehe Skizze 2.3. In Formeln:

$$A \cup B := \{x \mid x \in A \text{ oder } x \in B\}.$$

Als nächstes widmen wir uns der *Differenzmenge*. Diese nimmt alle Elemente aus einer Menge hinaus, welche in einer anderen Menge liegen, siehe Skizze 2.4. In Formeln bedeutet dies

$$A \setminus B := \{a \in A \mid a \notin B\}.$$

Machen wir zu diesen drei Operationen ein schnelles Beispiel. Sei dafür

$$A := \{1, 2, 3, 4\}, \quad B := \{1, 42, 7\}.$$

Dann gilt

$$A \cap B = \{1\}, \quad A \cup B = \{1, 2, 3, 4, 7, 42\} \text{ und } A \setminus B = \{2, 3, 4\}.$$

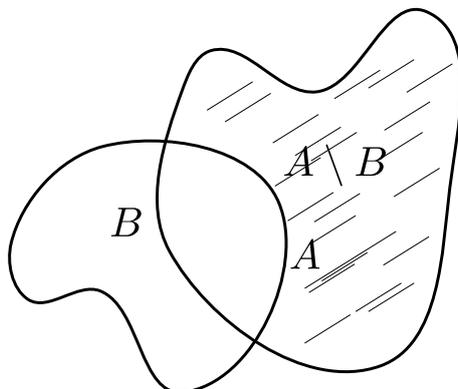


Abbildung 2.4: Differenz von zwei Mengen.

Falls es eine eindeutige Obermenge M gibt bzgl. der man arbeitet (d.h. $A \subset M$), so können wir auch das sogenannte *Komplement* einführen. Dies sind alle diese Elemente aus M , die nicht in der Menge A enthalten sind, siehe Skizze 2.5. In Formeln

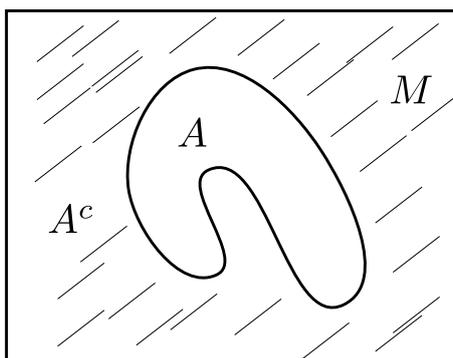


Abbildung 2.5: Komplement von zwei Mengen.

$$A^c := M \setminus A = \{x \in M \mid x \notin A\}.$$

Dazu ein Beispiel aus unseren Zahlbereichen:

$$A := \{x \in \mathbb{R} \mid x > 5\}.$$

Hier bilden wir nun das Komplement bzgl. der Obermenge \mathbb{R} und erhalten

$$A^c = \{x \in \mathbb{R} \mid x \leq 5\}.$$

Bevor wir mit der Theorie im nächsten Abschnitt weitermachen, führen wir noch *Intervalle* als Teilmenge der reellen Zahlen ein. Was ist ein Intervall? Dies ist die Menge aller reellen Zahlen, welche zwischen zwei gegebenen Zahlen liegen.

Seien also $a, b \in \mathbb{R}$ mit $a < b$. Dann setzen wir

$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$ abgeschlossenes Intervall

$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$ offenes Intervall

$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$ halboffenes Intervall

$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$ halboffenes Intervall

siehe auch Skizze 2.6. D.h. wir nutzen eine runde Klammer, soll die Grenze

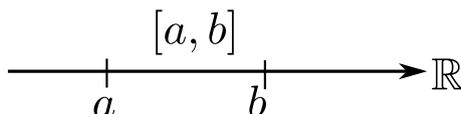


Abbildung 2.6: Ein abgeschlossenes Intervall.

nicht enthalten sein, und eine eckige Klammer, sollte diese enthalten sein. Als Anwendung können wir nun den Schnitt zweier Intervalle ausrechnen: Sei also $a < b$ und $c < d$. Dann gilt

$$[a, b] \cap [c, d] = \{x \in \mathbb{R} \mid a \leq x \leq b, c \leq x \leq d\}.$$

Wir führen nun beide Ungleichungen zusammen: $x \in [a, b] \cap [c, d]$ genau dann wenn

$$\max(a, c) \leq x \leq \min(b, d).$$

Siehe auch Skizze 2.7 Also gilt

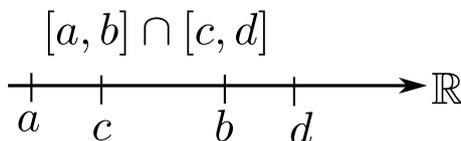


Abbildung 2.7: Schnitt zweier Intervalle.

$$[a, b] \cap [c, d] = [\max(a, c), \min(b, d)].$$

Wir können auch unendlich große Intervalle betrachten, indem wir die Symbole für unendlich ∞ und negativ unendlich $-\infty$ nutzen, z.B. für ein $a \in \mathbb{R}$

$$(a, \infty) := \{x \in \mathbb{R} \mid a < x\}$$

$$(-\infty, a) := \{x \in \mathbb{R} \mid a > x\}.$$

2.3 Tupel und das Kartesische Produkt

Mengen sind eine Sammlung ungeordneter Objekte. Möchte ich nun aber eine Ordnung mit möglichen Mehrfachnennungen benutzen, so verwendet man ein Tupel. Wie sieht das aus? Nehmen wir uns z.B. $a_1, \dots, a_n \in \mathbb{R}$ reelle Zahlen her. Dann können wir diese in einem n -Tupel anordnen. Die Notation sieht folgendermaßen aus:

$$(a_1, a_2, \dots, a_n).$$

Wir schreiben die Objekte also hintereinander weg, getrennt durch Kommata und begrenzt durch runde Klammern. Verwendet man ausschließlich Zahlen, so spricht man auch gerne von Vektoren. Etwas Vorsicht ist mit 2-Tupeln geboten. diese können mit offenen Intervallen verwechselt werden. Hier entscheidet der Kontext, welche Art von Objekt gemeint ist.

Nun können wir natürlich auch Mengen von Tupeln betrachten und sogar solche Mengen 'ohne' Tupel erzeugen. Dafür verwenden wir das kartesische Produkt. Formal: Seien A, B Mengen. Dann definieren wir das kartesische Produkt durch

$$A \times B := \{(a, b) \mid a \in A \text{ und } b \in B\}.$$

Damit können wir zum Beispiel *Rechtecke* mathematisch beschreiben, indem wir das kartesische Produkt von Intervallen betrachten: Sei dafür $a < b$ und $c < d$. Dann beschreibt

$$[a, b] \times [c, d] = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid a \leq x \leq b \text{ und } c \leq y \leq d\},$$

ein Rechteck im zweidimensionalen kartesischen Koordinatensystem, siehe Skizze 2.8. Bei diesem Beispiel fällt auf, dass wir generell höhere *Dimensionen* mit-

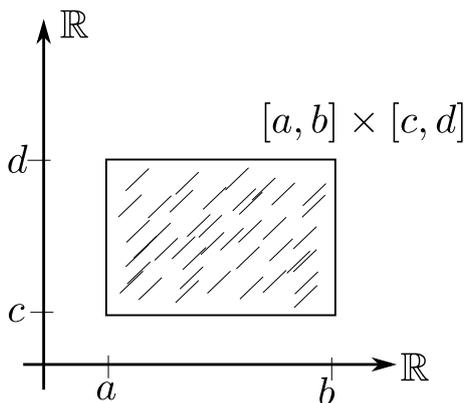


Abbildung 2.8: Rechteck durch kartesisches Produkt.

hilfe dieses kartesischen Produktes beschreiben können. Der zweidimensionale Raum wäre

$$\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$$

und der dreidimensionale Raum

$$\mathbb{R}^3 := \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

Dies kann man beliebig weiterspinnen und so zu n -dimensionalen Räumen kommen

$$\mathbb{R}^n := \mathbb{R} \times \dots \times \mathbb{R}.$$

3 Abbildungen

In Abschnitt 1 war öfter von Operationen wie der Anwendung des Logarithmusses die Rede. Nun wollen wir dieses Konzept zu sogenannten Abbildungen/Funktionen formalisieren. Seien dafür A und B Mengen. Eine Abbildung von A nach B ist eine Zuordnung, welche jedem Element aus A genau ein Element aus B zuordnet. Falls wir diese Abbildung f nennen, so schreiben wir

$$f : A \rightarrow B.$$

Hierbei kann man sich f als eine Art Rechenvorschrift vorstellen, die ein Element $a \in A$ verarbeitet und genau ein Element $f(a) \in B$ produziert. Ein Beispiel dazu: Sei $A = \mathbb{R} \setminus \{0\}$ und $B = \mathbb{R}$. Wir definieren als Rechenvorschrift ein $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ durch

$$f(x) = \frac{1}{x}.$$

Man kann dies notationell folgendermaßen zusammenfassen

$$\begin{aligned} f : \mathbb{R} \setminus \{0\} &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x}. \end{aligned}$$

Man beachte, dass wir die 0 hier explizit herausnehmen müssen, denn die Division durch 0 ist nicht erklärt.

Wichtig hierbei ist nun, dass nicht mehrere Ergebnisse herauskommen können und dass jedes Element aus A wirklich verarbeitet werden kann. Das Konzept, dass es sich bei einer Abbildung wirklich um eine Abbildung handelt, nennt man *Wohldefiniertheit* der Abbildung.

Weiter führen wir für $f : A \rightarrow B$ folgende Bezeichnungen ein:

$$\begin{aligned} A &\text{ heißt Definitionsbereich} \\ B &\text{ heißt Wertebereich} \end{aligned}$$

Abbildungen können miteinander interagieren. Zum Beispiel kann es sein, dass man Abbildungen hintereinander ausführen möchte. Die Notation ist dazu wie folgt: Sei $f : A \rightarrow B$ und $g : B \rightarrow C$. Dann definieren wir die *Komposition* durch

$$\begin{aligned} g \circ f &: A \rightarrow C \\ x &\mapsto g(f(x)). \end{aligned}$$

Ein Beispiel dazu:

$$\begin{aligned} f &: (0, \infty) \rightarrow \mathbb{R} \\ x &\mapsto x^2. \\ g &: (0, \infty) \rightarrow \mathbb{R} \\ x &\mapsto \sqrt{x}. \end{aligned}$$

Dann gilt

$$\begin{aligned} g \circ f &: (0, \infty) \rightarrow (0, \infty) \\ x &\mapsto g(f(x)) = \sqrt{x^2} = |x| = x. \end{aligned}$$

denn $x > 0$. Also kommt am Ende die Identität heraus. In diesem Fall spricht man davon, dass g eine *Umkehrabbildung* f besitzt. Notationell:

$$g^{-1} = f.$$

Falls wir mit reellwertigen Abbildungen arbeiten, können wir diese sogar multiplizieren und addieren: Sei dafür $f, g : A \rightarrow \mathbb{R}$ gegeben. Dann definieren wir

$$\begin{aligned} f \cdot g &: A \rightarrow \mathbb{R} \\ x &\mapsto f(x) \cdot g(x) \end{aligned}$$

und

$$\begin{aligned} f + g &: A \rightarrow \mathbb{R} \\ x &\mapsto f(x) + g(x). \end{aligned}$$

Zum Beispiel sei $f(x) = \log_2 x$ und $g(x) = x^2$. Dann haben wir

$$(f \cdot g + g)(x) = f(x) \cdot g(x) + g(x) = (\log_2 x) \cdot x^2 + x^2$$

Unsere Beispiele für Abbildungen geben bis dato immer nur Rechenvorschriften für einen großen Teil der reellen Zahl. Was ist, wenn wir diese nun an einzelnen Punkten oder Teilmengen von \mathbb{R} abändern möchten? Dazu führen wir folgende Notation für *abschnittsweise* definierte Abbildungen ein. Im Beispiel

$$x \mapsto \begin{cases} x^2, & x \leq 0 \\ \log_2 x, & x > 0. \end{cases}$$

D.h. Wir setzen untereinander Rechenvorschriften, welche in verschiedenen Teilmengen gelten sollen.

3.1 Graphen von Abbildungen

Hier stellen wir ein wichtiges geometrisches Hilfsmittel vor, um Abbildungen zu untersuchen. Diese haben wir bereits informell in Abschnitt 1 genutzt, und zwar den Graph einer Abbildung. Für eine Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ nutzen wir dafür das kartesische Koordinatensystem. Auf der x -Achse tragen wir den Eingabewert $x \in \mathbb{R}$ ab, während dazu auf der y -Achse der Wert $f(x)$ abgetragen wird, siehe z.B. Abbildung 1.7. Für allgemeine Abbildungen $f : A \rightarrow B$ können wir ebenfalls formal einen Graphen definieren und zwar als Menge!

$$\text{graph}(f) := \{(x, y) \in A \times B \mid f(x) = y\}$$

Damit können wir sehr allgemein den Punkt beschreiben, in denen zwei Abbildungen gleich sind oder in anderen Worten wir können die Schnittpunkte zweier Graphen beschreiben. Als Beispiel möchten wir feststellen wann $f(x) = x^2$ und $g(x) = x$ übereinstimmen, d.h. wir suchen die $x \in \mathbb{R}$ in denen $f(x) = x^2 = x = g(x)$ gilt. Diese Punkte können wir auch folgendermaßen beschreiben:

$$\begin{aligned} \text{graph}(f) \cap \text{graph}(g) &= \{(x, y) \in \mathbb{R}^2 \mid y = f(x) \text{ und } y = g(x)\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid y = g(x) = f(x)\} = \{(x, y) \in \mathbb{R}^2 \mid y = x^2 = x\} = \{(0, 0), (1, 1)\}. \end{aligned}$$

Nullstellen von reellwertigen Abbildungen f können so zum Beispiel folgendermaßen notiert und interpretiert werden

$$\text{graph}(f) \cap \text{graph}(x \mapsto 0).$$

In geometrischen Worten bedeutet dies, dass eine *Nullstelle* einer Abbildung einem Schnittpunkt des zugehörigen Graphens mit der x -Achse entspricht.

Der Graph bietet auch die Möglichkeit an, geometrisch *Umkehrabbildungen* zu finden. Für eine Abbildung $f : \mathbb{R} \supset D \rightarrow \mathbb{R}$ sucht man üblicherweise die Umkehrabbildung algebraisch, indem man für alle $x \in D$ die Gleichung $f(x) = y$ nach x auflöst. Hier gehen wir nun anders vor. Wir spiegeln den Graphen von f an dem Graphen der Identität, d.h. an $\text{graph}(x \mapsto x) \subset \mathbb{R}^2$. Falls die dabei entstandene Menge wieder ein Graph einer Abbildung ist, so ist diese Abbildung die Umkehrabbildung von f . Abbildung 3.1 demonstriert dies an dem Beispiel des natürlichen Logarithmusses $\ln : (0, \infty) \rightarrow \mathbb{R}$. Die Umkehrabbildung $\ln^{-1}(x) = e^x$ wird durch die Exponentialabbildung gegeben.

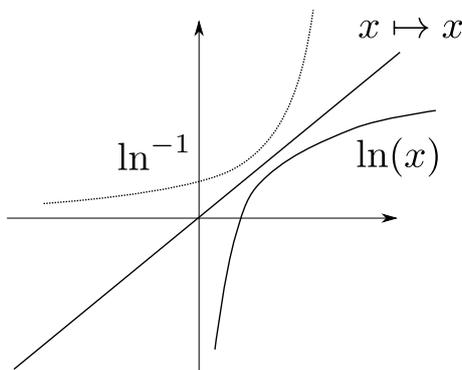


Abbildung 3.1: Inverse Abbildung des \ln durch Spiegeln.

3.2 Trigonometrische Funktionen

Als Beispiel für eine komplexere Abbildung und als Wiederholung führen wir hier trigonometrische Funktionen, d.h. \sin , \cos und \tan ein. Wir erklären diese über ein *rechtwinkliges Dreieck* im Einheitskreis. Für einen Winkel α erklären wir den $\sin(\alpha)$ und $\cos(\alpha)$ über Abbildung 3.2 Die Abbildung \tan definieren wir nun durch

$$\tan(\alpha) = \frac{\sin(\alpha)}{\cos(\alpha)}.$$

Üblicherweise werden Winkel in Grad gemessen. Für spätere analytische Zwecke ist allerdings das *Bogenmaß* (auch Bogenlänge genannt) als Maßeinheit günstiger. Die Umrechnungsformel lautet wie folgt

$$\text{Winkel in Bogenmaß} = \pi \cdot \frac{\text{Winkel in Grad}}{180}.$$

In Zukunft gehen wir immer davon aus, Winkel im Bogenmaß zu behandeln. Zum Beispiel in der folgenden Skizze 3.3 des Graphen vom Sinus: Anhand der

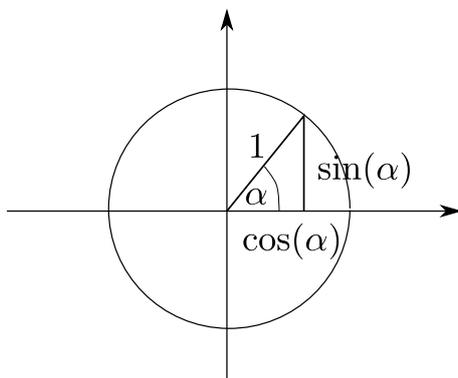


Abbildung 3.2: Definition des Sinus und des Kosinus am Einheitskreis.

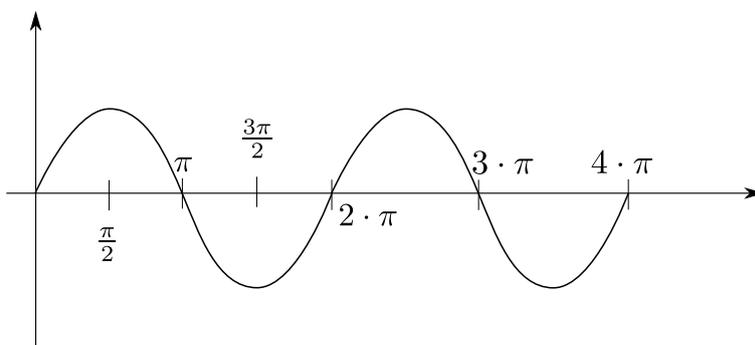


Abbildung 3.3: Graph von sin.

Definition über die Abbildung 3.2 sehen wir, dass sin periodisch ist, d.h. sich nach einer bestimmten Zeit wiederholt. Diese Periode ist genau $2 \cdot \pi$. In Formeln bedeutet dies

$$\text{Für alle } x \in \mathbb{R} \text{ gilt } \sin(x + 2 \cdot \pi) = \sin(x).$$

Dieses Konzept kann man verallgemeinern. Wir sagen eine Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ ist periodisch mit *Periode* $L > 0$, falls für alle $x \in \mathbb{R}$ gilt

$$f(x + L) = f(x).$$

Im Beispiel des Sinus und des Cosinus gilt für die Periode $L = 2 \cdot \pi$.

Mithilfe unserer Spiegelungsmethode aus Abschnitt 3.1 können wir den sin invertieren. Spiegeln wir den Graphen an erhalten wir aber Abbildung 3.4. Dort werden einem x -Wert unendlich viele y -Werte zugeordnet. Also ist dies streng genommen keine Abbildung. Die Idee ist es nun, den Graphen nach oben bzw. unten abzuschneiden, siehe Abbildung 3.5. Damit erhalten wir den sogenannten *Arkussinus*. Dafür schreiben wir

$$\sin^{-1} = \arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Analog können wir vorgehen um die Umkehrabbildungen vom Kosinus und Tan-

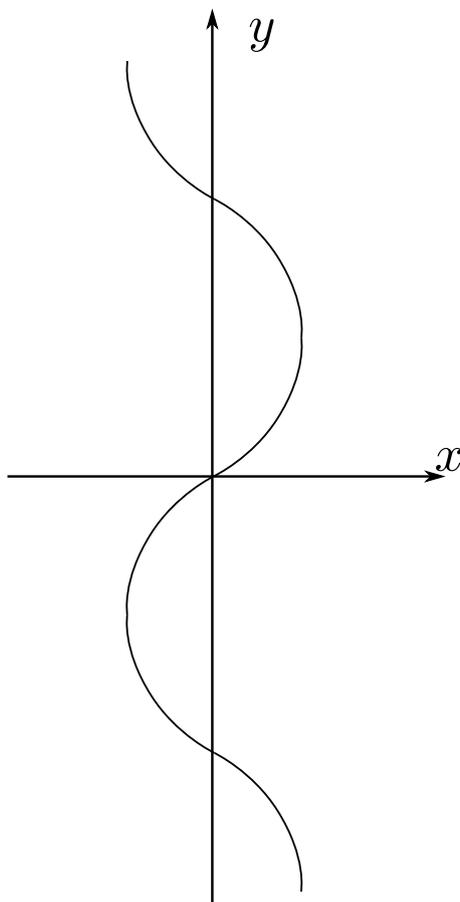


Abbildung 3.4: Naive Spiegelung des Graphen vom Sinus.

gens zu bestimmen. Wir erhalten:

$$\cos^{-1} = \arccos : [-1, 1] \rightarrow [0, \pi]$$

$$\tan^{-1} = \arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Die Bezeichnung für \arccos lautet *Arkuscosinus* und für \arctan *Arkustangens*.

3.3 Grenzwerte und Stetigkeit

In diesem Abschnitt werden wir auf eine informelle Weise Stetigkeit einführen. Formal korrekt wird dies dann in den Einführungsveranstaltungen im ersten Semester geschehen. Zuerst schauen wir uns an, was ein Grenzwert ist: Sei dafür $f : \mathbb{R} \supset D \rightarrow \mathbb{R}$ und $x_0 \in \mathbb{R}$. Falls wir ein $y \in \mathbb{R}$ finden, sodass für $x \in D$ beliebig nahe an x_0 gilt, dass $f(x)$ beliebig nahe an y liegt (siehe auch Skizze 3.6), so schreiben wir

$$\lim_{x \rightarrow x_0} f(x) = y.$$

In anderen Worten: Falls x zu x_0 läuft, so approximiert $f(x)$ den Wert y .

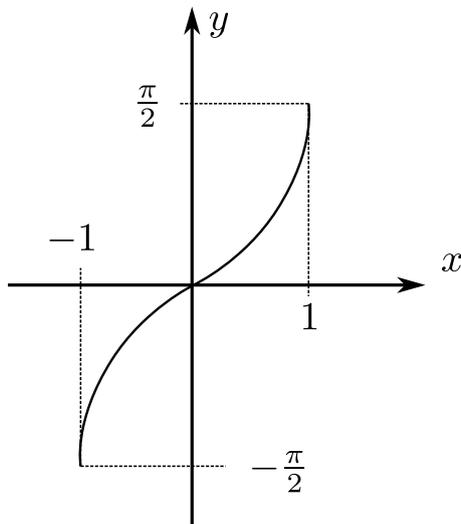


Abbildung 3.5: Graph von arcsin.

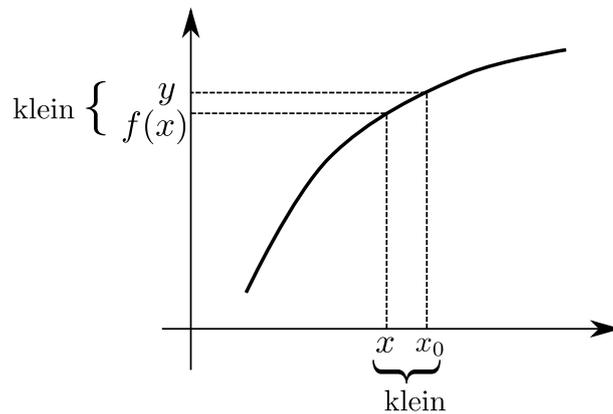


Abbildung 3.6: Definition des Grenzwertes einer Abbildung.

Machen wir ein Beispiel dazu:

$$f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$$

$$x \mapsto \frac{x^2 - 1}{x - 1}.$$

Wir betrachten $x_0 = 1$ und schauen uns an, wie f in der Nähe von $x_0 = 1$ aussieht:

$$f(x) = \frac{x^2 - 1}{x - 1} = \frac{(x - 1) \cdot (x + 1)}{x - 1} = x + 1.$$

Also erhalten wir

$$\lim_{x \rightarrow 1} f(x) = \lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = \lim_{x \rightarrow 1} x + 1 = 2.$$

Damit können wir nun im Allgemeinen erklären, was es für $f : D \rightarrow \mathbb{R}$ bedeutet in einem Punkt stetig zu sein: Sei dafür zusätzlich $x_0 \in D$ (dies war vorher nicht nötig!). Wir sagen f ist stetig in x_0 , falls

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Wir sagen weiter, dass f selbst stetig ist, falls f in allen $x_0 \in D$ stetig ist.

Welche Konsequenzen hat dies? Falls eine Abbildung f stetig ist, so können wir dessen Graphen zeichnen, ohne den Stift vom Papier zu nehmen! Abbildung 3.7 gibt dafür eine Skizze für eine nichtstetige Abbildung. Hier sehen wir einen sogenannten Sprung.

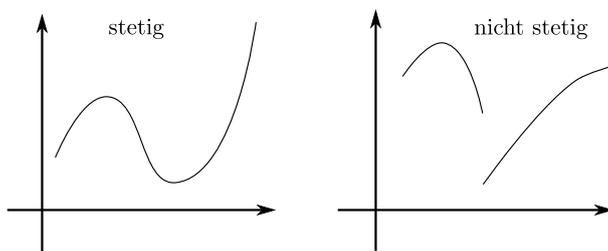


Abbildung 3.7: Vergleich von stetiger und nichtstetiger Abbildung.

Vorsicht ist geboten bei der Art von Unstetigkeitsstellen. Diese können sehr befremdlich werden. Ein klassisches Beispiel einer solchen Abbildung wäre

$$x \mapsto \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0. \end{cases}$$

Diese ist bei $x_0 = 0$ nicht stetig. Eine Skizze dieser Abbildung in der Nähe von $x = 0$ fällt bereits schwer. Einzusehen, dass diese Abbildung nicht stetig ist, ist mit abstrakteren Mitteln allerdings deutlich leichter. Aus Zeitgründen verzichten wir hier deshalb darauf.

Kommen wir zu eher positiven Aussagen in dieser Richtung: Addition, Multiplikation und Komposition von stetigen Abbildungen generiert wieder eine stetige Abbildung. Dies bedeutet, dass z.B. *Polynome*

$$P : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \sum_{i=0}^n \alpha_i x^i$$

mit $\alpha_i \in \mathbb{R}$ wieder stetig sind. Gleichzeitig sind die in Abschnitt 1 besprochenen Operationen ebenfalls stetig, d.h. Potenzieren, Wurzel ziehen, multiplikativ invertieren, Logarithmen bilden,...

4 Differenzialrechnung

4.1 Was ist eine Ableitung?

Hier wollen wir uns nun die Steigung von Abbildungen ansehen, bzw. deren Wachstumsrate. Wachstumsrate ist die Veränderung einer Abbildung im Verhältnis zur Veränderung der Eingabedaten. Für *Geraden* haben wir eine gute geometrische Vorstellung dafür. Eine Gerade ist eine Abbildung

$$G : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto m \cdot x + b$$

mit $m, b \in \mathbb{R}$. In Abbildung 4.1 wird ein *Steigungsdreieck* eingezeichnet. Dort ist die Veränderung in x -Richtung, also der Eingabewerte $x - x_0$. In y -Richtung, also

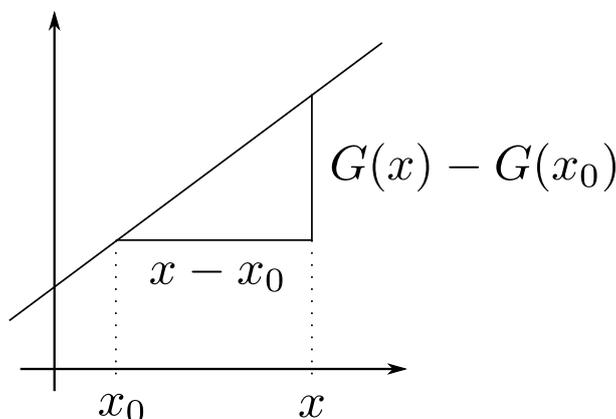


Abbildung 4.1: Steigung einer Geraden.

die Veränderung der Funktionswerte, handelt es sich dabei um $G(x) - G(x_0)$. Setzen wir diese ins Verhältnis, so erhalten wir

$$\frac{G(x) - G(x_0)}{x - x_0} = \frac{m \cdot x - b - (m \cdot x_0 - b)}{x - x_0} = \frac{m \cdot (x - x_0)}{x - x_0} = m.$$

Dieser Wert ist unabhängig von x und x_0 ! Wir nennen deswegen auch m die *Steigung* der Geraden G . Wenn wir dies nun für beliebigere Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ tun wollen, so wird dies etwas komplizierter. Schließlich ändert sich das Verhältnis $\frac{f(x) - f(x_0)}{x - x_0}$ abhängig davon, wo x und x_0 sind. Beide Werte können weit auseinander liegen und dann bleibt keine gute qualitative Aussage über f selber mehr übrig. Besser wäre eine rein lokale Definition. Die Idee ist es nun, statt durch $\frac{f(x) - f(x_0)}{x - x_0}$ eine *Sekante* (d.h. eine gerade Linie, die durch 2 Punkte des Graphen geht) eine *Tangente* (d.h. eine gerade Linie, die den Graphen einmal berührt) zu benutzen. Abbildung 4.2 gibt uns eine Skizze der Situation. Nun ist die Idee die Steigung dieser Tangente an $(x_0, f(x_0))$ als Definition der Steigung von f in x_0 zu setzen. Wie rechnen wir diese nun aus? Über Approximation der Sekantensteigungen! Wir sagen also eine Abbildung f ist in x_0 differenzierbar, falls folgender Grenzwert existiert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

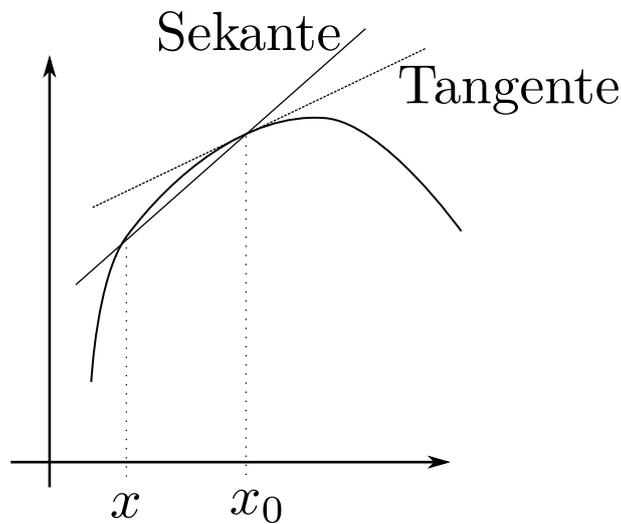


Abbildung 4.2: Steigung einer Geraden.

Diesen Ausdruck nennt man auch *Differenzialquotient*. Falls dieser existiert, schreiben wir

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

In allen Punkten in denen der Grenzwert existiert, können wir so eine Abbildung f' definieren. Diese nennen wir Ableitung von f . Zur Notation: Es ist durchaus üblich eine etwas andere Notation für die Ableitung zu nutzen, die das Steigungsdreieck widerspiegelt. Diese lautet wie folgt

$$\frac{df}{dx} = f'.$$

Das d deutet an, dass es sich um 'infinitesimal' kleine Änderungen von f bzw. x handelt, die danach durcheinander geteilt werden.

Machen wir ein Beispiel zur Definition der Ableitung: Sei $f(x) = x^2$. Dann gilt mithilfe der dritten binomischen Formel

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \rightarrow x_0} \frac{(x - x_0) \cdot (x + x_0)}{x - x_0} = \lim_{x \rightarrow x_0} x + x_0 = 2 \cdot x_0.$$

In diesem Falle ist die Ableitung also

$$f'(x) = 2 \cdot x.$$

Machen wir noch ein typisches Beispiel für eine nichtdifferenzierbare Funktion: $f(x) = |x|$ ist in $x_0 = 0$ nicht differenzierbar, denn

$$\lim_{x \rightarrow 0, x > 0} \frac{|x| - |0|}{x - 0} = \lim_{x \rightarrow 0, x > 0} \frac{x}{x} = 1$$

und

$$\lim_{x \rightarrow 0, x < 0} \frac{|x| - |0|}{x - 0} = \lim_{x \rightarrow 0, x < 0} \frac{-x}{x} = -1.$$

D.h. falls wir uns von rechts der Null nähern, erhalten wir einen anderen Grenzwert als wenn wir uns von links nähern. Eine Skizze des Graphen klärt die Situation ebenfalls, siehe Abbildung 4.3, denn in $x = 0$ gibt es eine Spitze im

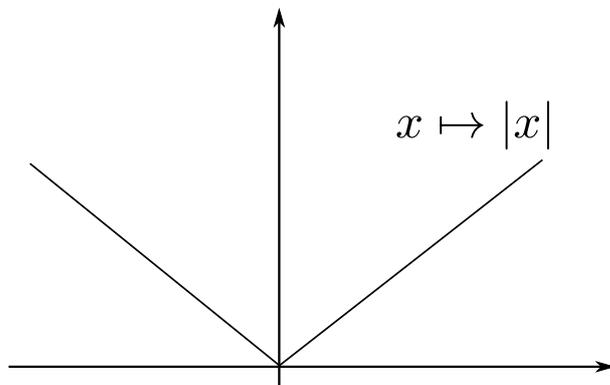


Abbildung 4.3: Skizze des Graphen von $|x|$.

Graphen und wir können dort keine eindeutige Tangente mehr anlegen.

4.2 Rechengesetze für Ableitungen

Hier wollen wir uns um das praktische Ausrechnen von Ableitungen kümmern. Wir werden eine Reihe von Rechenregeln erarbeiten und eine Sammlung von Ableitungen von Standardfunktionen geben. Zuerst zu den Rechenregeln: Ableiten ist linear. Dies bedeutet für f, g differenzierbar und $\alpha, \beta \in \mathbb{R}$ gilt, dass

$$x \mapsto \alpha f(x) + \beta g(x)$$

ist differenzierbar und es gilt

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

Begründen wir diese Gleichung anhand der Definition:

$$\begin{aligned} & \lim_{x \rightarrow x_0} \frac{\alpha f(x) + \beta g(x) - \alpha f(x_0) - \beta g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \alpha \frac{f(x) - f(x_0)}{x - x_0} + \beta \frac{g(x) - g(x_0)}{x - x_0} = \alpha f'(x_0) + \beta g'(x_0). \end{aligned}$$

Wie verhält es sich mit der Multiplikation von zwei differenzierbaren Abbildungen? wir leiten nun eine entsprechende Formel her. Seien dafür wieder f und g differenzierbare Abbildungen. Hier nutzen wir die Idee der sogenannten nahrhaften 0-Ergänzung.

$$\begin{aligned} & \lim_{x \rightarrow x_0} \frac{f(x) \cdot g(x) - f(x_0) \cdot g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{g(x)(f(x) - f(x_0)) + f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \lim_{x \rightarrow x_0} g(x) \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0} f(x_0) \frac{g(x) - g(x_0)}{x - x_0} = f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Hier haben wir benutzt, dass differenzierbare Abbildungen stetig sind, um $\lim_{x \rightarrow x_0} g(x) = g(x_0)$ zu schließen. Dies verwenden wir hier ohne Beweis. Wir haben also folgende Regel gefunden

$$(f \cdot g)' = f' \cdot g + f \cdot g',$$

genannt *Produktregel*.

Nun betrachten wir die Komposition von differenzierbaren Abbildungen. Wir leiten dies ähnlich wie die Produktregel her

$$\begin{aligned} & \lim_{x \rightarrow x_0} \frac{f(g(x)) - f(g(x_0))}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(g(x)) - f(g(x_0))}{x - x_0} \frac{g(x) - g(x_0)}{g(x) - g(x_0)} \\ &= \lim_{x \rightarrow x_0} \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)} \frac{g(x) - g(x_0)}{x - x_0} \\ &= \lim_{y \rightarrow y_0 = g(x_0)} \frac{f(y) - f(y_0)}{y - y_0} \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(g(x_0))g'(x_0). \end{aligned}$$

Die sogenannte *Kettenregel* hat also folgende Form:

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

In Worten als Merksatz zusammengefasst gilt also: Äußere Ableitung mal innere Ableitung.

Aus der Kettenregel können wir eine Formel für die Ableitung einer Umkehrabbildung finden. Wir wissen für ein differenzierbares f mit Umkehrabbildung f^{-1} gilt

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

Differenzieren wir diese Gleichung erhalten wir mithilfe der Kettenregel

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

Umstellen liefert

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

Wir können unsere Produktregel mit der Kettenregel kombinieren und so sogar Brüche ableiten (Man beachte hier, dass der Nenner, d.h. g in dem Punkt x in dem wir rechnen ungleich 0 sein muss):

$$\begin{aligned} \left(\frac{f}{g}\right)' &= \left(f \cdot \frac{1}{g}\right)' = \frac{f'}{g} + f(g^{-1})' \\ &= \frac{f'}{g} + f \cdot g' \cdot (-1) \cdot g^{-2} = \frac{f' \cdot g}{g^2} - \frac{f \cdot g'}{g^2} \\ &= \frac{f' \cdot g - f \cdot g'}{g^2}. \end{aligned}$$

Fassen wir also unsere Rechenregeln noch einmal zusammen:

$$(\alpha f + \beta g)' = \alpha f' + \beta g' \quad (\text{Linearität}) \quad (4.1)$$

$$(f \cdot g)' = f' \cdot g + f \cdot g' \quad (\text{Produktregel}) \quad (4.2)$$

$$(f \circ g)' = (f' \circ g) \cdot g' \quad (\text{Kettenregel}) \quad (4.3)$$

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}} \quad (\text{Umkehrabbildung}) \quad (4.4)$$

$$\left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2} \quad (\text{Quotientenregel}). \quad (4.5)$$

Aus praktischen Gründen sammeln wir nun Ableitungen für Abbildungen, die wir bereits in den vorherigen Abschnitten behandelt hatten:

Abbildung	Ableitung
$x \mapsto x^\alpha$	$x \mapsto \alpha \cdot x^{\alpha-1}$
$x \mapsto e^x$	$x \mapsto e^x$
$x \mapsto \ln(x)$	$x \mapsto \frac{1}{x}$
$x \mapsto \sin(x)$	$x \mapsto \cos(x)$
$x \mapsto \cos(x)$	$x \mapsto -\sin(x)$

Machen wir ein Beispiel zu den Rechenregeln. Wir setzen

$$f(x) = x \cdot \sin(x^2) + e^x + 5.$$

Dann gilt

$$\begin{aligned} f'(x) &= (x \cdot \sin(x^2) + e^x + 5)' \stackrel{\text{Linear}}{=} (x \cdot \sin(x^2))' + (e^x)' + (5)' \\ &\stackrel{\text{Prod.regel}}{=} 1 \cdot \sin(x^2) + x \cdot (\sin(x^2))' + e^x + 0 \\ &\stackrel{\text{Kettenregel}}{=} \sin(x^2) + x \cdot (2 \cdot x) \cos(x^2) + e^x \\ &= \sin(x^2) + 2x^2 \cos(x^2) + e^x. \end{aligned}$$

Für die Praxis machen wir nun eine schnelle Anwendung dieser Rechenregeln für Grenzwerte. Sei dafür I ein offenes Intervall und $x_0 \in I$. Seien weiter $f, g : I \rightarrow \mathbb{R}$ zwei stetig differenzierbare Abbildungen mit $f(x_0) = g(x_0) = 0$. Die Frage ist nun, welches Ergebnis bei

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$$

herauskommt. Intuitiv fragen wir uns also, welche der beiden Abbildungen schneller gegen 0 läuft. Die 'Geschwindigkeit' wird aber in der Ableitung kodiert. Deswegen wundert es nicht, dass die sogenannte *L'Hospital'sche* Regel gilt (Wir verzichten auf einen Beweis und verweisen stattdessen auf z.B. [5, Satz 10.11]):

Falls im obigen Falle zusätzlich gilt, dass $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ existiert, so haben wir

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Man beachte, dass diese Aussage auch gilt, falls $x_0 = \infty$ oder $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = \infty$.

Nun bekommen wir auch nichtlineare Phänomäne gut in den Griff. Z.B. können wir so folgenden Grenzwert schnell ausrechnen:

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{(\sin(x))'}{x'} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = \frac{1}{1} = 1.$$

4.3 Extremwerte

In diesem Abschnitt wollen wir uns Extremwerte von Abbildungen ansehen. Präziser beschäftigen wir uns hier mit lokalen Maximal/Minimalstellen einer Abbildung $f : (a, b) \rightarrow \mathbb{R}$. Dies ist informell gesagt ein Punkt im Definitionsbereich, sodass in der Nähe dieses Punktes der zugehörige Funktionswert Maximal/Minimal ist. Formeller heißt dies: Wir nennen $x_0 \in (a, b)$ ein *lokales Maximum (Minimum)*, falls in einer Umgebung $U := (x_0 - \delta, x_0 + \delta) \subset (a, b)$ (für $\delta > 0$ klein genug) von x_0 gilt

$$\text{Für alle } x \in U = (x_0 - \delta, x_0 + \delta) \text{ gilt } f(x_0) \geq (\leq) f(x).$$

Abbildung 4.4 gibt eine kleine Sammlung solcher möglichen Extremstellen.

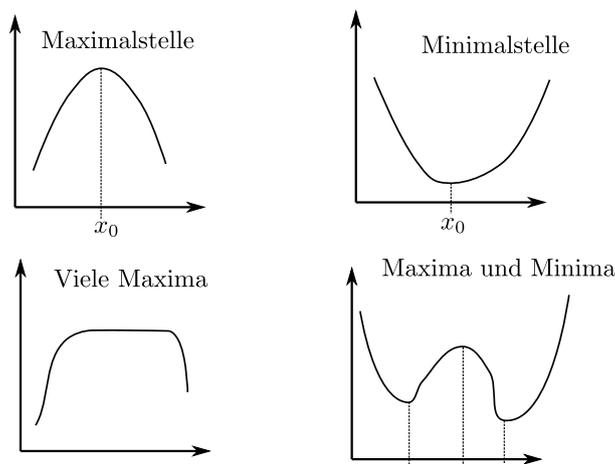


Abbildung 4.4: Eine Sammlung möglicher Extrema.

Bei Betrachtung dieser Sammlung machen wir folgende Beobachtung: Falls die Abbildung differenzierbar ist, so scheint die Ableitung, d.h. die Tangente, an der Maximalstelle parallel zur x -Achse zu verlaufen. In anderen Worten, dort scheint $f'(x_0) = 0$ zu gelten, siehe Abbildung 4.5

D.h. die Logik ist wie folgt:

$$x_0 \text{ Maximum/Minimum von } f \Rightarrow f'(x_0) = 0. \quad (4.6)$$

Wir setzen also voraus, ein Extremstelle zu haben und können dann daraus folgern, dass dort die Ableitung verschwindet. So ein logisches Verhalten nennt man notwendig. D.h. dass die Ableitung verschwindet ist notwendig für die Existenz einer Extremalstelle. Deswegen nennt man (4.6) gerne auch *notwendiges*

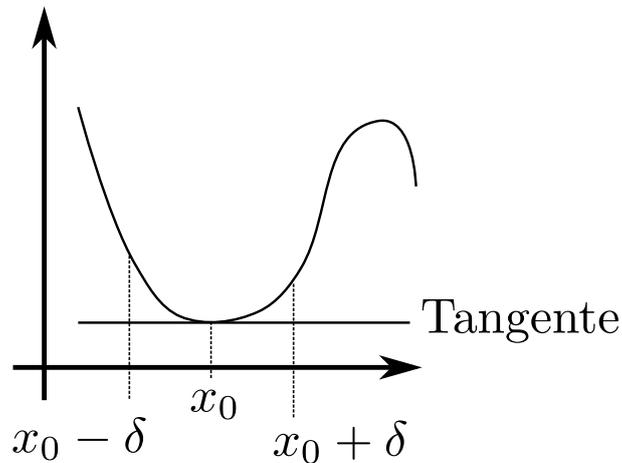


Abbildung 4.5: Ein notwendiges Kriterium einer Extremalstelle.

Kriterium. Damit können wir effektiv Kandidaten für eine Extremstelle berechnen.

Leider gibt es auch Abbildungen, die eine Nullstelle in der Ableitung haben, aber dort kein lokales Extremum aufweisen. Z.B. $f(x) = x^3$. Dafür gilt $f'(x) = 3 \cdot x^2$ und damit

$$0 = f'(x_0) = 3 \cdot x_0^2 \Rightarrow x_0 = 0.$$

Eine Skizze der Abbildung (siehe Abbildung 4.6) verdeutlicht aber schnell, dass es sich nicht um eine Extremstelle handelt.

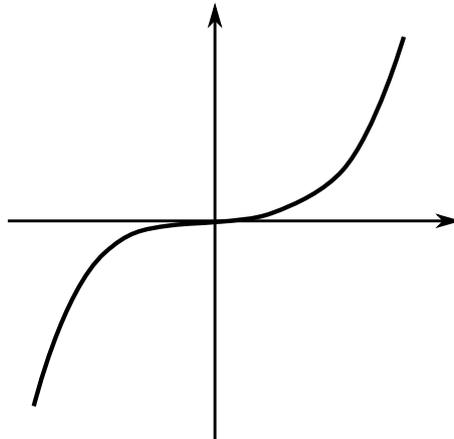


Abbildung 4.6: Eine Skizze von $x \mapsto x^3$.

Schauen wir uns das Verhalten einer typischen Minimalstelle etwas genauer an (siehe Abbildung 4.7). Vor dem Minimum fällt der Graph, d.h. die Ableitung dort ist negativ. Nach dem Minimum steigt der Graph wieder an, d.h. nun ist die Ableitung positiv. Insgesamt scheint die Ableitung selber also anzusteigen! In der Tat ist das genau der Fall. Wie können wir nun prägnant ausdrücken,

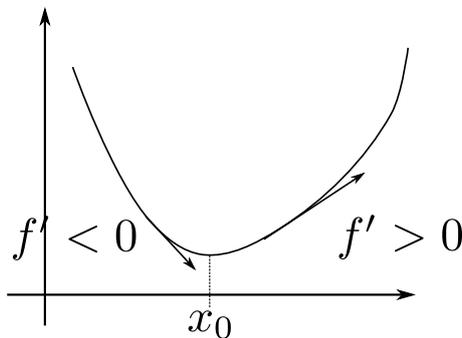


Abbildung 4.7: Ableitung steigt in der Nähe eines Minimums.

dass eine Abbildung monoton steigt? Wieder über die Ableitung, d.h. falls f zweimal stetig differenzierbar ist, so wäre ein *hinreichendes Kriterium* für ein Minimum $f''(x_0) > 0$. Fassen wir die Strategie zur Bestimmung von Extremalstellen zusammen:

- 1) Finden von Kandidaten mittels des notwendigen Kriteriums $f'(x_0) = 0$.
- 2) Entscheiden um was für ein Extremum es sich handelt:
 - (i) Falls $f''(x_0) > 0$, so ist x_0 ein Minimum.
 - (ii) Falls $f''(x_0) < 0$, so ist x_0 ein Maximum.
 - (iii) Falls $f''(x_0) = 0$, so ist keine Aussage möglich.

Punkt (iii) unterstreichen wir noch mit einem Beispiel: $f(x) = x^4$ besitzt in 0 ein Minimum, denn $f(0) = 0$ und $f(x) \geq 0$ für alle $x \in \mathbb{R}$. Außerdem gilt $f'(x) = 4 \cdot x^3$ und $f''(x) = 12 \cdot x^2$. Damit gilt $f'(0) = 0$ und $f''(0) = 0$. Betrachten wir $g(x) = -x^4$ bekommen wir dasselbe Ergebnis für die Ableitung bei 0 aber g besitzt ein Maximum. Im Zweifel muss man dann versuchen ein Vorzeichenwechsel in der Ableitung zu zeigen oder andere Tricks nutzen.

Nun rechnen wir noch ein Beispiel, um die eben gezeigten Methoden zu verdeutlichen: Wir setzen

$$f(x) = x^3 - x^2 + 5.$$

Dann gilt für die Ableitungen

$$f'(x) = 3 \cdot x^2 - 2 \cdot x \text{ und } f''(x) = 6 \cdot x - 2.$$

Dann bekommen wir als notwendiges Kriterium für die Existenz eines Extremums

$$0 = f'(x) = 3 \cdot x^2 - 2 \cdot x \Rightarrow 0 = x \cdot (3 \cdot x - 2).$$

Also sind die Kandidaten

$$x_1 = 0 \text{ und } x_2 = \frac{2}{3}.$$

Setzen wir diese nun in die zweite Ableitung ein, erhalten wir

$$f''(x_1) = -2 \quad f''(x_2) = 2.$$

Also haben wir bei x_1 ein lokales Maximum und bei x_2 ein lokales Minimum.

5 Integralrechnung

5.1 Die Fläche unter einem Graph

Hier wollen wir den Integralbegriff als Flächenmessung einführen. Wir nehmen uns dafür eine Abbildung $f : [a, b] \rightarrow \mathbb{R}$ und wollen nun den Flächeninhalt unter dem Graphen von f berechnen, d.h. die Fläche, die vom Graphen von f und der x -Achse eingeschlossen wird, siehe Abbildung 5.1.

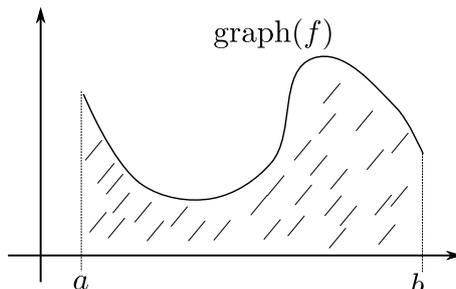


Abbildung 5.1: Fläche unter einem Graphen.

Nun ist die Frage, wie wir diese Fläche berechnen. Dazu fragen wir uns zuerst, welche Flächeninhalte wir den berechnen können. Und zwar die von *Rechtecken*! Die Idee ist es nun, die Fläche unter dem Graphen durch Rechtecke zu approximieren, siehe Abbildung 5.2. Wir unterteilen nun das Intervall $[a, b]$

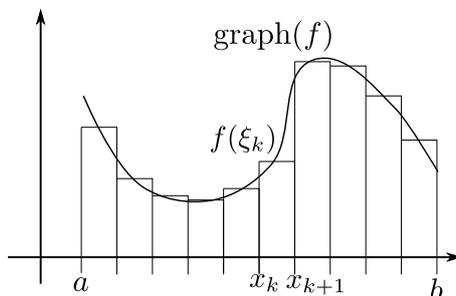


Abbildung 5.2: Fläche unter einem Graphen.

in $n \in \mathbb{N}$ gleichlange Teilintervalle $[x_k, x_{k+1}]$ (Das zugehörige Fachwort heißt *äquidistante Zerlegung*). Die Formel dafür lautet

$$x_k = a + \frac{b-a}{n} \cdot k, \quad k = 0, 1, \dots, n.$$

Dann nehmen wir uns ein $\xi_k \in [x_k, x_{k+1}]$ und nehmen als Höhe unseres Rechtecks $f(\xi_k)$. Dann ist die Approximation der Fläche durch unsere Rechtecke gegeben als

$$S_n(f, a, b) := \sum_{k=0}^{n-1} (\text{Länge von } [x_k, x_{k+1}]) \cdot f(\xi_k).$$

Berechnen wir also die Länge dieses Intervalls:

$$x_{k+1} - x_k = a + \frac{b-a}{n} \cdot k + 1 - \left(a + \frac{b-a}{n} \cdot k \right) = \frac{b-a}{n}.$$

also erhalten wir

$$S_n(f, a, b) = \sum_{k=0}^{n-1} \frac{b-a}{n} f(\xi_k).$$

Dieses S_n nennt man *Riemannsche Summe*. Die Hoffnung ist nun, dass dieser Ausdruck unabhängig von der Wahl von ξ_k immer zu demselben Ergebnis konvergiert, wenn wir $n \rightarrow \infty$ betrachten. Im Allgemeinen kann man dies nicht erwarten, sollte $f : [a, b] \rightarrow \mathbb{R}$ aber stetig sein, so kann man dies beweisen (siehe z.B. [5, Theorem 9.6]). Aus Zeitgründen werden wir hier aber darauf verzichten. Falls dies aber so ist, definieren wir unseren Integralbegriff als

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} S_n(f, a, b).$$

Wir machen nun ein Beispiel: $f(x) := 3 \cdot x^2$ auf $[a, b] := [0, 1]$, um die Komplexität dieses Problems zu verdeutlichen. Wir wählen die Zerlegung des Intervalls $x_k = \frac{k}{n}$ und werten f an der linken Seite des Intervalls aus. Dann erhalten wir

$$\begin{aligned} S_n(f, 0, 1) &= \sum_{k=0}^{n-1} \frac{1}{n} \cdot 3 \cdot \left(0 + \frac{1-0}{n} \cdot k \right)^2 \\ &= \sum_{k=0}^{n-1} \frac{3}{n^3} \cdot k^2 = \frac{3}{n^3} \cdot \sum_{k=0}^{n-1} k^2. \end{aligned}$$

Um den Wert also ausrechnen zu können, müssten wir $\sum_{k=0}^{n-1} k^2$ berechnen. Wir schauen in eine unserer Quellen ([1, Seite 86]) und bekommen

$$\sum_{k=0}^{n-1} k^2 = \frac{(n-1) \cdot n \cdot (2n-1)}{6}.$$

Also bekommen wir

$$\begin{aligned} S_n(f, 0, 1) &= \frac{3}{n^3} \cdot \frac{(n-1) \cdot n \cdot (2n-1)}{6} \\ &= \frac{3}{6 \cdot n^3} \cdot (2n^3 - 3n^2 + n) = \frac{1}{2} \left(2 - \frac{3}{n} + \frac{1}{n^2} \right) \rightarrow \frac{1}{2} \cdot 2 = 1 \end{aligned}$$

Also haben wir

$$\int_a^b 3 \cdot x^2 dx = 1.$$

Diese Methode kann aber bei komplexeren f sehr schwer zu benutzen sein. Sie ist aber für numerische Zwecke (d.h. Flächenberechnung mit dem Computer) sehr gut geeignet. Für analytisch exakte Lösungen zu dem Problem besprechen wir in Abschnitt 5.2 eine einfacheres Vorgehen.

Wir führen nun aber zuerst einige wichtige allgemeine Eigenschaften des Integrals aus:

1. Was passiert, wenn wir mehrere Funktionen addieren und diese dann integrieren wollen? Sei dafür $\beta, \gamma \in \mathbb{R}$ und $f, g : [a, b] \rightarrow \mathbb{R}$. Dann gilt für die Riemannsche Summe

$$\begin{aligned} S_n(\beta \cdot f + \gamma \cdot g, a, b) &= \sum_{k=0}^{n-1} \frac{b-a}{n} (\beta f(\xi_k) + \gamma g(\xi_k)) \\ &= \beta \cdot \sum_{k=0}^{n-1} \frac{b-a}{n} f(\xi_k) + \gamma \cdot \sum_{k=0}^{n-1} \frac{b-a}{n} g(\xi_k) \\ &= \beta S_n(f, a, b) + \gamma S_n(g, a, b). \end{aligned}$$

Insgesamt erhalten wir also wenn wir $n \rightarrow \infty$ betrachten

$$\int_a^b \beta f(x) + \gamma g(x) dx = \beta \int_a^b f(x) dx + \gamma \int_a^b g(x) dx.$$

Man nennt diese Eigenschaft Linearität. In anderen Worten: Falls man eine Summe integriert, integriert man einfach die zugehörigen Summanden.

2. Unter Umständen kann es sein, dass es günstiger ist in Abschnitten zu integrieren. In der Tat gilt für alle $c \in [a, b]$

$$\int_a^c f(x) dx + \int_c^b f(x) dx = \int_a^b f(x) dx,$$

denn die zugehörige Fläche wird nun in zwei getrennten Teilen berechnet. Abbildung 5.3 verdeutlicht diesen Zusammenhang.

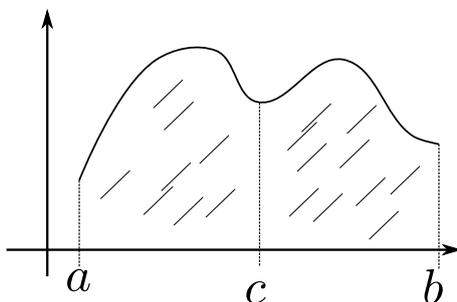


Abbildung 5.3: Integrieren in Abschnitten.

3. Unsere Definition einer Fläche besitzt ein Vorzeichen! Falls $f(\xi_k) < 0$ ist, so bekommen unsere Rechtecke eine negative Fläche zugeteilt, siehe Abbildung 5.4. Dies führt dazu, dass sich negative und positive Anteile gegenseitig heben können. Möchte man nun den wahren Flächeninhalt berechnen, so integriert man entweder von Nullstelle zu Nullstelle von f und korrigiert die Vorzeichen händisch, oder man integriert $x \mapsto |f(x)|$.

5.2 Hauptsatz der Differential- und Integralrechnung

Im vorherigen Abschnitt 5.1 haben wir im Beispiel von $\int_0^1 3 \cdot x^2 dx$ gesehen, dass das Auswerten von Integralen anhand der Definition sehr mühsam werden kann.

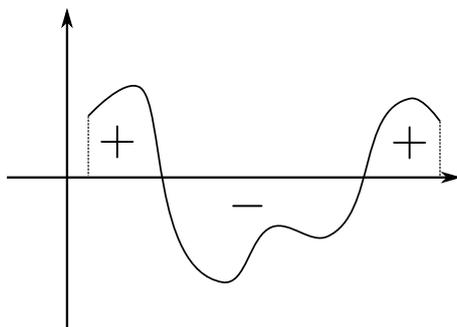


Abbildung 5.4: Negativer Flächeninhalt.

Hier werden wir nun einen einfacheren Weg vorstellen. Dazu nehmen wir uns ein abschnittsweise konstantes $f : [a, b] \rightarrow \mathbb{R}$ her und untersuchen folgende Abbildung:

$$F(x) := \int_a^x f(t) dt.$$

Man beachte in der Definition, dass hier eine andere Integrationsvariable benutzt wurde, um ein Verwechseln mit x zu vermeiden. Nun ist die Frage, ob wir einen Zusammenhang zwischen F und f herstellen können. Schauen wir uns dazu einen Graph beider Abbildungen in Abbildung 5.5 an, in dem Fall, dass f abschnittsweise konstant ist. Der Graph von F scheint schneller zu steigen, um

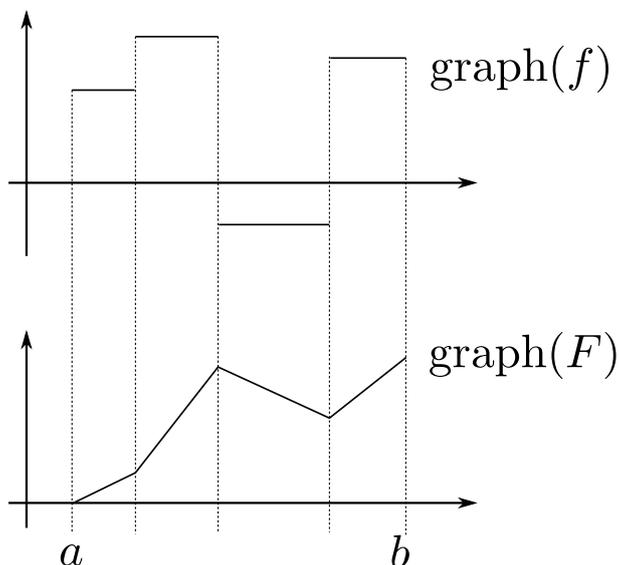


Abbildung 5.5: Vergleich von Abbildung zum Integralwert.

so größer f ist. Dies lässt einen Zusammenhang zwischen der Ableitung von F und selber vermuten. In der Tat ist dies der Fall. Die genaue Aussage ist der Hauptsatz der Differential- und Integralrechnung (kurz HDI):

Theorem (Hauptsatz der Differential- und Integralrechnung). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist die Abbildung $F : [a, b] \rightarrow \mathbb{R}$ mit

$$F(x) = \int_a^x f(t) dt$$

stetig und auf (a, b) differenzierbar. Weiter gilt für alle $x \in (a, b)$

$$F'(x) = f(x).$$

Einen Beweis werden hier aus Zeitgründen nicht führen, kann aber z.B. in [1, §11.2] nachgelesen werden. Wir führen einen Begriff für solche F ein. Wir nennen eine Abbildung $F : [a, b] \rightarrow \mathbb{R}$ *Stammfunktion* von $f : [a, b] \rightarrow \mathbb{R}$, falls für alle $x \in (a, b)$ die Abbildung F in x differenzierbar ist und es gilt

$$F'(x) = f(x).$$

Stammfunktionen bzgl. einer festen Abbildung f können sich nur um eine Konstante unterscheiden. Für eine Begründung seien F_1 und F_2 zwei Stammfunktionen von f . Dann gilt für alle $x \in (a, b)$

$$(F_1 - F_2)'(x) = F_1'(x) - F_2'(x) = f(x) - f(x) = 0$$

Also muss $F_1 - F_2$ eine konstante Abbildung sein.

Dies bedeutet insbesondere für unsere bereits gefundene Stammfunktion $F(x) = \int_a^x f(t) dt$ im Vergleich zu jeder anderen Stammfunktion \tilde{F} , dass wir eine Konstante $C \in \mathbb{R}$ finden mit $F - \tilde{F} = C$. Damit gilt

$$\int_a^b f(t) dt = F(b) - F(a) = \tilde{F}(b) + C - (\tilde{F}(a) + C) = \tilde{F}(b) - \tilde{F}(a).$$

Insgesamt erhalten wir so eine zuverlässige Methode, um Integrale für eine stetige Abbildung $f : [a, b] \rightarrow \mathbb{R}$ zu berechnen:

1. Finde Stammfunktion $F : [a, b] \rightarrow \mathbb{R}$ von f .
2. Berechne

$$F(b) - F(a) = \int_a^b f(t) dt.$$

Betrachten wir unser Beispiel aus Abschnitt 5.1 noch einmal: $f(x) = 3 \cdot x^2$ auf $[0, 1]$. Eine Stammfunktion zu f ist $F(x) = x^3$. Dann gilt

$$\int_0^1 3 \cdot x^2 dx = F(1) - F(0) = 1^3 - 0^3 = 1.$$

Dies ging deutlich schneller als zuvor!

Wir führen noch etwas Notation ein. Um die Differenz der Stammfunktionen etwas abzukürzen nutzt man auch gerne folgende Notationen:

$$[F(x)]_a^b := F(x)|_a^b := F(b) - F(a).$$

In unserem Beispiel hätte dies folgende Form:

$$\int_0^1 3 \cdot x^2 dx = [x^3]_0^1 = x^3|_0^1 = 1.$$

Welche von beiden Notationen nun genutzt wird, ist Geschmackssache.

Nun noch etwas Notation zu Stammfunktionen. Lässt man beim Integral die Grenzen weg, d.h.

$$\int f(x) dx$$

so meint man damit irgendeine Stammfunktion von f . Ein solcher Ausdruck wird auch *unbestimmtes Integral* genannt. Deswegen nennt man finden einer Stammfunktion von f auch integrieren von f .

Im letzten Teil dieses Abschnittes wollen wir einige Stammfunktionen sammeln:

Abbildung	Stammfunktion
$x \mapsto x^\alpha, \alpha \neq -1$	$x \mapsto \frac{1}{\alpha+1} \cdot x^{\alpha+1}$
$x \mapsto \frac{1}{x}$	$x \mapsto \ln(x)$
$x \mapsto e^x$	$x \mapsto e^x$
$x \mapsto \ln(x)$	$x \mapsto x \cdot \ln(x) - x$
$x \mapsto \sin(x)$	$x \mapsto -\cos(x)$
$x \mapsto \cos(x)$	$x \mapsto \sin(x)$
$x \mapsto \tan(x)$	$x \mapsto \ln(\cos(x))$

5.3 Tricks zur Berechnung von Integralen

Hier wenden wir uns der effektiven Berechnung von Stammfunktionen zu. Leider ist dies nicht immer möglich in dem Sinne, dass am Ende eine Abbildung bestehend aus grundlegenden Funktionen herauskommt. Ein solches Beispiel wäre $\int e^{-x^2} dx$.

Allerdings gibt es einige Methoden und Tricks, die in gewissen Fällen weiterhelfen. Diese wollen wir uns hier genauer ansehen. Essentiell wenden wir dabei den Hauptsatz der Differential- und Integralrechnung an und kombinieren ihn mit unseren Ableitungsregeln aus Abschnitt 4.2.

Wir starten mit einer einfachen Anwendung der Kettenregel, der sogenannten *logarithmische Integration*. Sei dafür ein $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Wir betrachten nun das Problem

$$\int \frac{f'(x)}{f(x)} dx = ?.$$

Mittels der Kettenregel bekommen wir

$$(\ln(|f(x)|))' = \begin{cases} \frac{f'(x)}{f(x)}, & f(x) > 0 \\ \frac{-f'(x)}{-f(x)}, & f(x) < 0 \end{cases}$$

Dies liefert direkt eine Stammfunktion:

$$\int \frac{f'(x)}{f(x)} dx = \ln(|f(x)|).$$

Machen wir ein Beispiel dazu:

$$\int \frac{\cos(x)}{\sin(x)} dx = ?.$$

In diesem Fall setzen wir $f(x) = \sin(x)$. Da $f'(x) = \cos(x)$ erhalten wir

$$\int \frac{\cos(x)}{\sin(x)} dx = \int \frac{f'(x)}{f(x)} dx = \ln |f(x)| = \ln |\sin(x)|$$

Wir schauen uns nun eine weitere Methode an, die *partielle Integration*: Dafür benötigen wir zwei stetig differenzierbare Abbildungen $f, g : [a, b] \rightarrow \mathbb{R}$. Mit der Produktregel für Ableitungen (siehe Abschnitt 4.2) gilt

$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

D.h. eine Stammfunktion von $f' \cdot g + f \cdot g'$ ist $f \cdot g$. Nach dem Hauptsatz der Differential- und Integralrechnung haben wir also

$$[f(x) \cdot g(x)]_a^b = \int_a^b (f \cdot g)'(x) dx = \int_a^b f'(x) \cdot g(x) + f(x) \cdot g'(x) dx.$$

Umstellen liefert die Formel der partiellen Integration:

$$\int_a^b f'(x) \cdot g(x) dx = [f(x) \cdot g(x)]_a^b - \int_a^b f(x) \cdot g'(x) dx.$$

Machen wir ein Beispiel dazu. Wir werden

$$\int_0^\pi x \cdot \sin(x) dx = ?$$

berechnen. Dazu setzen wir $f'(x) = \sin(x)$ und $g(x) = x$. Dann gilt mithilfe der partiellen Integration

$$\begin{aligned} \int_0^\pi x \cdot \sin(x) dx &= [-x \cdot \cos(x)]_0^\pi - \int_0^\pi -\cos(x) dx \\ &= \pi + [\sin(x)]_0^\pi = \pi. \end{aligned}$$

Diese Methode ist also besonders gut geeignet, wenn man ein Produkt hat, dessen Faktoren man einzeln integrieren kann.

Kommen wir nun zur *Substitutionsregel*. Diese transformiert die zu integrierende Abbildung und ist im Prinzip eine Umkehrung der Kettenregel. Sei dafür $g : [a, b] \rightarrow \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig differenzierbar. Wir setzen $F'(x) = f(x)$ (also F als Stammfunktion von f). Dann gilt nach Kettenregel

$$(F \circ g)'(t) = f(g(t)) \cdot g'(t).$$

Also erhalten wir mit dem Hauptsatz

$$\begin{aligned} \int_a^b f(g(t)) \cdot g'(t) dt &= \int_a^b (F \circ g)'(t) dt \\ &= [(F \circ g)(t)]_a^b = F(g(b)) - F(g(a)) \\ &= \int_{g(a)}^{g(b)} f(x) dx. \end{aligned}$$

Die Substitutionsformel lautet also

$$\int_a^b f(g(t)) \cdot g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx.$$

Mithilfe der Schreibweise der Ableitung mit Differentialen lässt diese sich leicht merken: Wir schreiben $x = g(t)$. Dann leiten wir nach t ab und erhalten

$$\frac{dx}{dt} = g'(t).$$

Formal multiplizieren wir das dt auf die andere Seite und erhalten

$$dx = g'(t)dt.$$

Nun können wir die Symbole x und dx im Integral ersetzen

$$\int f(x) dx = \int f(g(t))g'(t) dt$$

und erhalten so die Substitutionsformel für unbestimmte Integrale.

Nun rechnen wir ein Beispiel um die Methode zu verdeutlichen:

$$\int_1^e \frac{1}{x \cdot (\ln(x) + 1)} dx = ?$$

Wir berechnen zuerst eine Stammfunktion mittels Substitution. Dafür setzen wir

$$t = \ln(x) \text{ bzw. } e^t = x.$$

Dann erhalten wir

$$\frac{dx}{dt} = e^t \text{ bzw. } dx = e^t dt.$$

Damit wird aus dem unbestimmten Integral

$$\int \frac{1}{x \cdot (\ln(x) + 1)} dx = \int \frac{e^t}{e^t \cdot (t + 1)} dt = \int \frac{1}{t + 1} dt = \ln |t + 1|$$

Nun müssen wir natürlich noch $t = \ln(x)$ resubstituieren und erhalten also

$$\int \frac{1}{x \cdot (\ln(x) + 1)} dx = \ln |\ln(x) + 1|.$$

Damit bekommen wir

$$\int_1^e \frac{1}{x \cdot (\ln(x) + 1)} dx = [\ln |\ln(x) + 1|]_1^e = \ln(2) - \ln(1) = \ln(2).$$

Mit dem Trick der Resubstitution müssen wir uns auch keine Sorge um die transformierten Grenzen des Integrals machen. Diese kommen am Ende korrekt heraus.

Es gibt noch viele weitere Tricks um Integrale analytisch zu berechnen. Im Anhang A finden Sie zu Beispiel noch die Partialbruchzerlegung, welche das Integrieren von rationalen Funktionen ermöglicht.

6 Wahrscheinlichkeitsrechnung

Hier werden wir einige Grundbegriffe aus der Wahrscheinlichkeitstheorie durchgehen. Wir werden diese aber nicht in voller Allgemeinheit besprechen, sondern uns der Einfachheit halber dabei auf endliche Ereignismengen beschränken, siehe Abschnitt 6.1. Auch werden wir aus Zeitgründen nicht auf alle üblichen Begriffe eingehen können.

Abschließend werden wir in Abschnitt 6.4 noch eine Verallgemeinerung zu \mathbb{R} als Ereignismenge angeben. Dies dient aber nur als Ausblick und soll keine vollwertige Diskussion werden.

Die allgemeine Formulierung der Wahrscheinlichkeitstheorie kann z.B. in [1, §37] nachgelesen werden.

6.1 Endlich viele mögliche Ereignisse, Kolmogorov Axiome

Die Wahrscheinlichkeitstheorie in der Anwendung ist immer ein Modell der Realität. Dieses Modell soll aber gewissen mathematischen Gesetzmäßigkeiten unterliegen, die eine gute 'Wahrscheinlichkeit' ausmacht. Wir werden in diesem Abschnitt diese Gesetzmäßigkeiten erarbeiten und danach in Abschnitt 6.2 an einigen Beispielen die dazugehörige Modellierung betrachten. Die Idee dabei ist es Modell und Theorie möglichst voneinander zu trennen.

In ihrer vollen Allgemeinheit können wir die Wahrscheinlichkeitstheorie aus Zeitgründen nicht entwickeln, weswegen wir uns auf *endliche Ereignismengen* beschränken werden. Eine Ereignismenge ist dabei eine Menge, in der alle möglichen Ergebnisse eines zufälligen Ereignisses zusammengefasst sind. Als Beispiel wählen wir das einmalige Würfeln mit einem 6-seitigen Würfel. Als Ergebnis können die Zahlen 1 bis 6 auftreten, weswegen die zugehörige Ereignismenge X

$$X = \{1, 2, 3, 4, 5, 6\}$$

ist. D.h. dass wir uns in diesem Abschnitt auf zufällige Ereignisse beschränken wollen, bei denen lediglich endlich viele Ergebnisse auftreten können.

Um nun mathematisch die Wahrscheinlichkeitstheorie aufbauen zu können benötigen wir zuerst einen neuen Begriff aus der Mengenlehre, und zwar die *Potenzmenge*. Sei dafür X eine Menge. Die Potenzmenge von X ist die Menge aller Teilmengen von X und wird mit $\text{Pot}(X)$ notiert. Also gilt etwas abstrakter

$$\text{Pot}(X) = \{A \mid A \subset X\}.$$

Für ein schnelles Beispiel dazu machen wir für die Menge $X = \{1, 2, 3\}$. Dann erhalten wir für die Potenzmenge von diesem X

$$\text{Pot}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Wir führen den Begriff der Potenzmenge ein, um z.B. ein 'oder' Ereignis beschreiben zu können. Um beim Beispiel des 6-seitigen Würfels zu bleiben, wäre das Ereignis

würfel eine 1 oder 2

durch die Menge

$$\{1, 2\} \in \text{P}(\{1, 2, 3, 4, 5, 6\})$$

kodiert.

Nun definieren wir den Begriff eines *Wahrscheinlichkeitsmaßes*, welches jedem Element einer Potenzmenge eine Wahrscheinlichkeit zuordnen soll. Abstrakt nennen wir eine Abbildung $P : \text{Pot}(X) \rightarrow \mathbb{R}$ für eine endliche Menge X ein Wahrscheinlichkeitsmaß, falls folgende Eigenschaften (genannt die *Axiome von Kolmogorov*) erfüllt werden

1. Für alle $A \in \text{Pot}(X)$ gilt $0 \leq P(A) \leq 1$ (Nichtnegativität)
2. $P(X) = 1$ (Normierung)
3. Für $A, B \in \text{Pot}(X)$ mit $A \cap B = \emptyset$ gilt
 $P(A \cup B) = P(A) + P(B)$ (Additivität).

Die realweltliche Interpretation ist dabei wie folgt: Die Zahl die bei der Anwendung von P herauskommt entspricht unserer Wahrscheinlichkeit. Umso höher, desto wahrscheinlicher. Bei $P(A) = 1$ muss das Ereignis A auftreten und bei $P(A) = 0$ ist die Wahrscheinlichkeit 0, dass es eintritt.

Gehen wir wieder zu unserem Beispiel des 6-seitigen Würfels. Nehmen wir an dieser sei fair, d.h. jede Würfelseite hätte die gleiche Wahrscheinlichkeit aufzutreten. Dann gilt für einelementige Mengen, d.h. $x \in \{1, 2, 3, 4, 5, 6\}$ also

$$P(\{x\}) = \frac{1}{6}.$$

Die Additivitätseigenschaft (Axiom 3) liefert nun also für eine beliebige Menge $A \in \text{Pot}(\{1, 2, 3, 4, 5, 6\})$, welche wir durch $A = \{a_1, \dots, a_k\}$ ausdrücken also

$$P(A) = P(\{a_1, \dots, a_{k-1}\}) + P(\{a_k\}) = \dots = \sum_{j=1}^k P(\{a_j\}) = \sum_{j=1}^k \frac{1}{6} = \frac{k}{6}.$$

Hier ist k die Anzahl der Elemente der Menge A (diese notiert man gerne mit $\#A$ oder $|A|$) Für $A = \{1, 2\}$ haben wir also

$$P(\{1, 2\}) = \frac{2}{6} = \frac{1}{3}.$$

Dementsprechend ist die Wahrscheinlichkeit eine 1 oder 2 zu würfeln $\frac{1}{3}$. Aus diesem Beispiel sieht man außerdem sofort, dass wenn die Wahrscheinlichkeit für 1-Elementige Mengen festgelegt ist, so ist diese auch für alle Teilmengen $A \subset X$ festgelegt. Dafür darf aber X nur endliche viele Elemente besitzen.

Nun widmen wir uns wichtigen Folgerungen aus diesen Axiomen. Bei einer Entwicklung der Theorie auf beliebigen Mengen X bleiben diese immer noch korrekt. Zuerst widmen wir uns der Komplementbildung. Sei dafür $A \subset X$. Da $A \cap A^c = \emptyset$ liefert uns die Normierung und die Additivität

$$1 = P(X) = P(A \cup A^c) = P(A) + P(A^c).$$

Nach umstellen erhalten wir also

$$P(A^c) = 1 - P(A).$$

Insbesondere für $A = X$ erhalten wir

$$P(\emptyset) = P(X^c) = 1 - P(X) = 1 - 1 = 0.$$

Diese Gesetzmäßigkeit erlaubt das schnelle Berechnen von Wahrscheinlichkeiten von komplementären Ereignissen. Bei Würfelwurf mit 6-Seiten z.B. wollen wir die Wahrscheinlichkeit berechnen, dass der Würfel nicht die 1 anzeigt. Dies entspricht der Menge $A = X \setminus \{1\}$. Damit erhalten wir also

$$\text{Wahrscheinlichkeit 1 nicht zu würfeln} = P(A) = 1 - P(A^c) = 1 - P(\{1\}) = \frac{5}{6}.$$

Kommen wir nun zur Monotonie. Sei dafür $A, B \in \text{Pot}(X)$ mit $A \subset B$. Dann gilt insbesondere $B = A \cup (A^c \cap B)$. Da $A \cap (A^c \cap B) = \emptyset$ liefert die Additivität

$$P(B) = P(A) + P(A^c \cap B).$$

Das Nichtnegativitätsaxiom gibt uns $P(A^c \cap B) \geq 0$, also erhalten wir

$$P(B) \geq P(A).$$

Die Interpretation davon ist, dass falls wir mehr Möglichkeiten für unser Ereignis zulassen, dann wird die Wahrscheinlichkeit für dieses Ereignis auch größer.

Kommen wir nun zu einer Formel um die Wahrscheinlichkeit der Vereinigung zweier beliebiger Mengen zu berechnen. Sei dafür $A, B \in \text{Pot}(X)$. Dann gilt (vgl. Abbildung 6.1)

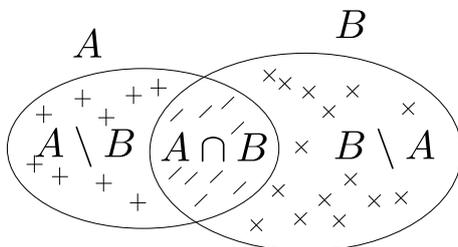


Abbildung 6.1: Vereinigung von A und B ausgedrückt mithilfe von Schnitten und Differenzmengen.

$$A \cup B = (A \cap B) \cup (A \setminus B) \cup (B \setminus A).$$

Da diese Mengen auf der rechten Seite paarweise disjunkt sind, erhalten wir mithilfe der Additivität

$$P(A \cup B) = P(A \cap B) + P(A \setminus B) + P(B \setminus A).$$

Außerdem haben wir aber auch $A = (A \cap B) \cup (A \setminus B)$ und wir erhalten damit

$$P(A) = P(A \cap B) + P(A \setminus B).$$

Vertauschen wir A und B in dieser Überlegung, erhalten wir außerdem

$$P(B) = P(A \cap B) + P(B \setminus A).$$

Setzen wir dies in unsere erste Überlegung ein, so erhalten wir

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(A \setminus B) + P(B \setminus A) \\ &= P(A \cap B) + P(A) - P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

Mit unserem Nichtnegativitätsaxiom erhalten wir nun sogar

$$P(A \cup B) \leq P(A) + P(B).$$

Fassen wir unsere Ergebnisse zusammen:

1. $P(A^c) = 1 - P(A)$ (Komplementbildung)
2. $P(\emptyset) = 0$ (leere Menge)
3. $A \subset B \Rightarrow P(A) \leq P(B)$ (Monotonie)
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Vereinigung)

Nun wollen wir uns noch der Frage stellen, was passiert, wenn wir zwei (oder mehrere) von einander unabhängige Wahrscheinlichkeitsexperimente hintereinandersetzen. Dann möchten wir natürlich eine Art kombinierte Wahrscheinlichkeit berechnen können. Abstrakt gesehen steckt ein *Satz von Carathéodory* dahinter, welchen wir hier kurz vorstellen wollen.

Zuerst ein kleines Beispiel dazu: Wir führen zuerst einen Würfelwurf mit einem sechseitigen Würfel durch und danach einen Wurf mit einer Münze. Für den Würfelwurf nennen wir die Ereignismenge X_W und für den Münzwurf X_M . Diese sind dementsprechend

$$X_W = \{1, 2, 3, 4, 5, 6\} \quad X_M = \{\text{Kopf}, \text{Zahl}\}.$$

Beim Hintereinanderausführen der Experimente gelangen wir so zu der ganzen Ereignismenge

$$X = X_W \times X_M.$$

Nun wollen wir ein Wahrscheinlichkeitsmass P für X festlegen, welches konsistent zu den Wahrscheinlichkeiten auf X_M und X_W ist. Dies ist Gegenstand des Satz von Carathéodory:

Theorem (Satz von Carathéodory). *Seien X_1 und X_2 endliche Mengen und P_1 bzw. P_2 Wahrscheinlichkeitsmaße auf X_1 bzw. X_2 . Dann gibt es ein eindeutiges Wahrscheinlichkeitsmaß P auf $X_1 \times X_2$, welches für alle $A \subset X_1$ und alle $B \subset X_2$ erfüllt*

$$P(A \times B) = P_1(A) \cdot P_2(B).$$

Auf einen Beweis werden wir hier aus Zeitgründen verzichten. Kehren wir lieber zu unserem Beispiel des Würfelwurfs und des nachfolgenden Münzwurfs zurück. Wir legen die Wahrscheinlichkeitsmaße P_W (Für den Würfelwurf) und P_M (Für den Münzwurf) fest durch

$$P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}, \quad P_M(\{\text{Kopf}\}) = P_M(\{\text{Zahl}\}) = \frac{1}{2}.$$

Dann gibt uns der Satz von Carathéodory nun ein Wahrscheinlichkeitsmaß P auf $X_W \times X_M$. Wollen wir nun z.B. die Wahrscheinlichkeit des Ereignisses berechnen zuerst eine 2 zu würfeln und danach beim Münzwurf Zahl zu erhalten, so ist entspricht dieses Ereignis der Menge

$$\{(2, \text{Zahl})\}.$$

Die Wahrscheinlichkeit nach dem Satz von Carathéodory ist

$$P(\{(2, \text{Zahl})\}) = P(\{2\} \times \{\text{Zahl}\}) = P_W(\{2\}) \cdot P_M(\{\text{Zahl}\}) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}.$$

6.2 Modellierung, die Regel von Laplace

Im vorherigen Abschnitt haben wir uns um den mathematischen Hintergrund gekümmert, welche Eigenschaften ein Wahrscheinlichkeitsmaß haben soll. Nun stellen wir uns die Frage, wie wir zu einem solchen Wahrscheinlichkeitsmaß kommen. Diesen Schritt nennt man Modellierung des Experimentes. Dabei gibt es streng genommen keine richtige oder falsche Modellierung, lediglich Modelle, die die Realität mehr oder weniger gut abbilden. In der Praxis kann man z.B. das Wahrscheinlichkeitsexperiment oft wiederholen um mit statistischen Methoden eine Einschätzung des Wahrscheinlichkeitsmaßes zu erhalten. Es gibt allerdings auch einige Heuristika, der man sich gerne bedient. Wir wollen hier eine davon vorstellen und zwar die *Regel von Laplace*:

Sei also X eine Menge der möglichen (endlichen) Ereignisse und $A \subset X$. Dann setzen wir heuristisch

$$P(A) := \frac{\text{Anzahl für } A \text{ günstige Ereignisse}}{\text{Anzahl der gesamt möglichen Ereignisse}}.$$

Aus der Einleitung in Abschnitt 6.1 ist klar, dass wir dies nur für die sogenannten *Elementarereignisse* tun müssen. $A \subset X$ nennt man Elementarereignis, falls A nur genau ein Element enthält. Ein einfaches Beispiel dazu hatten wir bereits in Abschnitt 6.1 beim einfachen Wurf eines Würfels gemacht.

Wir machen ein Beispiel zur Regel von Laplace: Wir nehmen nun einen Würfel mit 6 Seiten und würfeln ihn zweimal hintereinander. Als Ereignis betrachten wir die Summe der gewürfelten Augen der beiden Würfel, d.h. zeigt der erste Wurf 3 an und der zweite 4, so ist das Ergebnis $3 + 4 = 7$.

Also erhalten wir als Ereignismenge X die Zahlen von 2 bis 12, also

$$X := \{2, 3, \dots, 11, 12\}.$$

Für die Modellierung beschreiben wir das Ergebnis als Tupel der Länge 2, wir setzen also

$$Y := \{1, \dots, 6\} \times \{1, \dots, 6\} = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\}.$$

Die Interpretation ist dabei, dass in der ersten Komponente das Ergebnis des ersten Wurfes und in der zweiten Komponente das Ergebnis des zweiten Wurfes steht. Wenn wir nun die Anzahl der gesamt möglichen Ereignisse zählen, so müssen wir die Anzahl der Elemente in Y berechnen. Dazu halten wir die erste Komponente fest und zählen wie viele Tupel es geben kann. Falls die erste Komponente der 1 entspricht, erhalten wir

$$(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6) \in Y,$$

also 6 Tupel. Dies können wir für die anderen ersten Komponenten wiederholen (also noch 5 mal) und wir erhalten pro fixer erster Komponente 6 Tupel. Da wir die erste Komponente insgesamt nur 6 mal festhalten können, ist die Anzahl der Elemente in Y

$$|Y| = \#Y = 6 \cdot 6 = 36.$$

Gehen wir nun der Wahrscheinlichkeit für $3 \in X$ nach. Es gibt 2 Tupel, die die 3 erzeugen können, nämlich

$$3 = 1 + 2 = 2 + 1 \text{ also } (1, 2), (2, 1).$$

Also ist die Anzahl der günstigen Elemente also 2. Insgesamt gibt uns die Regel von Laplace also

$$P(\{3\}) = \frac{2}{36} = \frac{1}{18}.$$

Wiederholen wir dies für die anderen Elementarereignisse:

$$\begin{aligned} P(\{2\}) &= \frac{1}{36}, \text{ mit } (1, 1) \\ P(\{3\}) &= \frac{2}{36} = \frac{1}{18}, \text{ mit } (1, 2), (2, 1) \\ P(\{4\}) &= \frac{3}{36} = \frac{1}{12}, \text{ mit } (1, 3), (3, 1), (2, 2) \\ P(\{5\}) &= \frac{4}{36} = \frac{1}{9}, \text{ mit } (1, 4), (4, 1), (2, 3), (3, 2) \\ P(\{6\}) &= \frac{5}{36}, \text{ mit } (1, 5), (5, 1), (2, 4), (4, 2), (3, 3) \\ P(\{7\}) &= \frac{6}{36} = \frac{1}{6}, \text{ mit } (1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3) \\ P(\{8\}) &= \frac{5}{36}, \text{ mit } (2, 6), (6, 2), (3, 5), (5, 3), (4, 4) \\ P(\{9\}) &= \frac{4}{36} = \frac{1}{9}, \text{ mit } (3, 6), (6, 3), (4, 5), (5, 4) \\ P(\{10\}) &= \frac{3}{36} = \frac{1}{12}, \text{ mit } (4, 6), (6, 4), (5, 5) \\ P(\{11\}) &= \frac{2}{36} = \frac{1}{18}, \text{ mit } (5, 6), (6, 5) \\ P(\{12\}) &= \frac{1}{36}, \text{ mit } (6, 6) \end{aligned}$$

6.3 Erwartungswert und Varianz

In diesem Abschnitt geht es darum wie wir Vorhersagen machen können, sollte ein Experiment, für das wir das Wahrscheinlichkeitsmaß kennen, oft wiederholt wird. Ein Anwendung davon ist zum Beispiel Glücksspiel. Der Betreiber des Glücksspiel möchte beim Design seines Spieles natürlich im Schnitt Gewinn machen, dennoch sollte es einen gewissen Reiz für einen Spieler geben, an dem Spiel teilzunehmen.

Um diese Frage zu beantworten führen wir erst einmal etwas handlichere Notation für Summen ein. Sei X eine endliche Menge mit k Elementen, d.h. wir können die Elemente in X ohne Einschränkung durchnummerieren und setzen

$$X = \{x_1, x_2, \dots, x_k\}.$$

Weiterhin möchten wir jedem Element von X einen numerischen Wert zuweisen. Dafür sei $f : X \rightarrow \mathbb{R}$ eine Abbildung. Nun möchten wir eine Summe von f über X erklären. Wegen des Kommutativgesetzes ist folgende Setzung unabhängig davon wie wir X durchnummeriert haben:

$$\sum_{x \in X} f(x) := \sum_{j=1}^k f(x_j).$$

Nun beantworten wir die Eingangs gestellten Fragen: Sei X eine endliche Ereignismenge und P ein Wahrscheinlichkeitsmaß darauf. Weiter soll $f : X \rightarrow \mathbb{R}$ eine numerische Bewertung der Elementarereignisse sein. In der Praxis nennt man ein solches f auch *Zufallsvariable* (Oft wird diese mit X oder x notiert, dies vermeiden wir aber um mit der Ereignismenge X nicht in Konflikt zu geraten). Nun definieren wir den *Erwartungswert* von f durch

$$E(f) := \sum_{x \in X} f(x) \cdot P(\{x\}).$$

Die Interpretation des Erwartungswertes ist, dass bei einer hohen Wiederholung des Wahrscheinlichkeitsexperimentes im Schnitt dieser Wert bei der Auswertung mit f ergibt. Machen wir ein Beispiel:

Wir spielen ein einfaches Glücksspiel: Bei einem Würfelwurf mit einem 6 seitigen Würfel soll der Spieler seinen Einsatz von 1 Euro bei einer 1, 2 oder 3 verlieren. Bei 4 behält er seinen Einsatz und bei der 5 und 6 bekommt er 1 Euro ausbezahlt und darf seinen Einsatz behalten. In diesem Fall ist die Ereignismenge also $X = \{1, 2, 3, 4, 5, 6\}$ und das Wahrscheinlichkeitsmaß wird durch $P(A) = \frac{|A|}{6}$ gegeben wobei $A \subset X$ und $|A|$ die Anzahl der Elemente von A ist. Die Bewertung f führen wir aus der Sicht des Spielers wie folgt ein:

$$\begin{aligned} f(1) = f(2) = f(3) &= -1 \\ f(4) &= 0 \\ f(5) = f(6) &= 1 \end{aligned}$$

Dann ist der Erwartungswert

$$\begin{aligned} E(f) &= \sum_{x \in X} f(x)P(\{x\}) \\ &= f(1) \cdot \frac{1}{6} + f(2) \cdot \frac{1}{6} + f(3) \cdot \frac{1}{6} + f(4) \cdot \frac{1}{6} + f(5) \cdot \frac{1}{6} + f(6) \cdot \frac{1}{6} \\ &= -\frac{3}{6} + \frac{2}{6} = -\frac{1}{6}. \end{aligned}$$

Im Schnitt würde der Spieler also $\frac{1}{6}$ Euro verlieren, die Bank macht also im Schnitt ein Gewinngeschäft.

Nun widmen wir uns der Eingangs zweiten Frage: Wie sieht es mit dem Anreiz für den Spieler aus? Dies wollen wir messen, indem wir uns überlegen wie weit denn im Schnitt das Ergebnis vom Erwartungswert entfernt liegt. Umso größer dieser Abstand ist, desto weniger bemerkt der Spieler, dass er auf lange Sicht Verlust macht. Schließlich spielt dieser üblicherweise nur wenige Runden.

Kommen wir nun zur Definition: Sei dafür X wieder eine endliche Ereignismenge, P ein Wahrscheinlichkeitsmaß auf X und $f : X \rightarrow \mathbb{R}$ eine Zufallsvariable. Dann definieren wir die *Varianz* durch

$$\text{Var}(f) := E((f - E(f))^2) = \sum_{x \in X} (f(x) - E(f))^2 \cdot P(\{x\}).$$

Wegen des Quadrates innerhalb des Erwartungswertes wird die Varianz üblicherweise nicht in derselben Größenordnung wie der Erwartungswert liegen. Um allerdings dort grob zu bleiben, führen wir noch die *Standardabweichung* ein

$$\sigma(f) := \sqrt{\text{Var}(f)}.$$

Berechnen wir für unser Beispiel von oben die Varianz und Standardabweichung:

$$\begin{aligned} \text{Var}(f) &= \sum_{x \in X} (f(x) - E(f))^2 \cdot P(\{x\}) \\ &= \left(f(1) - \left(-\frac{1}{6}\right)\right)^2 P(\{1\}) + \left(f(2) - \left(-\frac{1}{6}\right)\right)^2 P(\{2\}) \\ &\quad + \left(f(3) - \left(-\frac{1}{6}\right)\right)^2 P(\{3\}) + \left(f(4) - \left(-\frac{1}{6}\right)\right)^2 P(\{4\}) \\ &\quad + \left(f(5) - \left(-\frac{1}{6}\right)\right)^2 P(\{5\}) + \left(f(6) - \left(-\frac{1}{6}\right)\right)^2 P(\{6\}) \\ &= \left(-1 + \frac{1}{6}\right)^2 \frac{3}{6} + \left(0 + \frac{1}{6}\right)^2 \cdot \frac{1}{6} + \left(1 + \frac{1}{6}\right)^2 \frac{2}{6} \\ &= \frac{5^2 \cdot 3}{6^2 \cdot 6} + \frac{7^2 \cdot 2}{6^2 \cdot 6} + \frac{1}{6^2 \cdot 6} = \frac{75 + 98 + 1}{36 \cdot 6} = \frac{174}{216} = \frac{29}{36} \approx 0,801. \end{aligned}$$

Mithilfe eines Taschenrechners kann man nun noch die Standardabweichung berechnen und erhält

$$\sigma(f) = \sqrt{\text{Var}(f)} = \sqrt{\frac{29}{36}} = \frac{\sqrt{29}}{6} \approx 0,898.$$

Also streuen wir um den Erwartungswert von $-\frac{1}{6}$ im Schnitt um 0,898 pro Würfelwurf herum.

Hierbei ermöglicht die Varianz bzw. Standardabweichung eine Einschätzung darüber, ob wir bei einem einmal durchgeführten Experiment nahe am Erwartungswert landen. Sollte also die Varianz genügend klein sein, so können wir im Einzelfall erwarten nahe beim Mittelwert zu landen.

6.4 Ausblick: \mathbb{R} als Menge der mögliche Ereignisse

Bis dato haben wir uns nur für den Fall interessiert, dass die Ereignismenge X eine endliche Menge ist. Dies liegt in erster Linie daran, dass ein großer Teil der für den allgemeinen Fall benötigten Mathematik über diesen Vorkurs hinausgeht. Andererseits sind unendlich große Mengen für die Modellierung natürlich sehr von Interesse. Ein einfaches Beispiel wäre das Werfen eines Dartpfeils auf eine Dartscheibe und das berechnen der Wahrscheinlichkeit ob der Pfeil in einem bestimmten Teil der Scheibe landet. Die Scheibe selber können wir als 2

dimensionale Fläche $S \subset \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ betrachten, sodass deren Flächeninhalt $area(S) > 0$ ist. Sei nun $A \subset S$ eine Menge von Punkten, deren Wahrscheinlichkeit wir berechnen wollen, dass diese beim Wurf getroffen werden. Eine sehr einfache Modellierung wäre nun

$$P(A) := \frac{area(A)}{area(S)}.$$

Dafür müssten wir allerdings die Flächenmessung mathematisch entwickeln. Dies würde den Rahmen dieser Veranstaltung bei weitem sprengen, denn für sehr komplizierte A ist es nicht trivial einen sinnvollen Flächenbegriff zu definieren. Die dahinterliegende Theorie nennt man Maßtheorie und diese liegt ebenfalls im Hintergrund der modernen Stochastik.

Um allerdings dennoch eine Idee davon zu vermitteln, wie man zumindestens auf \mathbb{R} einen Begriff entwickeln kann, nutzen wir die Idee der Dichte:

Wir nennen ein stetiges $h : \mathbb{R} \rightarrow [0, \infty)$ eine *Wahrscheinlichkeitsdichte*, falls

$$\int_{\mathbb{R}} h(x) dx = \lim_{K, L \rightarrow \infty} \int_{-K}^L h(x) dx = 1.$$

Dann definieren wir für ein Intervall $[a, b] \subset \mathbb{R}$ das Wahrscheinlichkeitsmaß assoziiert zu h durch

$$P([a, b]) := \int_a^b h(x) dx.$$

So können wir zumindestens für Intervalle einen Wahrscheinlichkeitsbegriff festlegen.

Ein berühmtes Beispiel für eine solche Dichte ist die Dichte der sogenannte *Normalverteilung*:

$$N_{\mu, \sigma}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

mit Parametern $\mu \in \mathbb{R}$, $\sigma \geq 0$. Diese Dichte kann nicht explizit integriert werden, d.h. sie besitzt keine explizit hinschreibbare Stammfunktion. Will man nun Wahrscheinlichkeiten mit ihr berechnen, so bemüht man entweder einen Computer oder Tabellen, in denen Werte berechnet wurden.

Zurück zur Theorie: Die in Abschnitt 6.1 entwickelten Gesetzmäßigkeiten bleiben im Prinzip erhalten. Der einzige Knackpunkt ist, dass die Potenzmenge als hinterliegendes Mengensystem nicht mehr funktioniert, sondern man nur von durch Intervallen 'generierbare' Mengen eine Wahrscheinlichkeit berechnen kann. Was dies genau bedeutet ist für die praktische Anwendung allerdings eher von untergeordneter Bedeutung.

Wir können nun auch einen Erwartungswert und eine Varianz definieren:

$$E(h) := \int_{\mathbb{R}} h(x) \cdot x dx$$

$$Var(h) := E((h - E(h))^2) = \int_{\mathbb{R}} (h(x) - E(h))^2 \cdot x dx,$$

mit denselben Interpretationen wie in Abschnitt 6.3. Bei der Normalverteilung gilt dann

$$\begin{aligned} E(N_{\mu, \sigma}) &= \mu \\ Var(N_{\mu, \sigma}) &= \sigma^2. \end{aligned}$$

Eine Idee für die Modellierung mithilfe der Normalverteilung ist es nun, viele realweltliche Experimente zu veranstalten und viel zu messen. Dann nimmt man einen Durchschnitt der Messdaten und setzt diesen Wert als μ , d.h. dem Erwartungswert.

Damit schließen wir die Entwicklung der Wahrscheinlichkeitstheorie. Weiteres Lesematerial zu solchen Dichten und deren Nutzen in der Statistik ist z.B. [1, §39].

A Partialbruchzerlegung zum Integrieren

Hier besprechen wir noch eine Methode um eine rationale Funktion zu integrieren, die sogenannte *Partialbruchzerlegung*. Wir werden aus Platzgründen diese nicht in voller Allgemeinheit besprechen, Details können in [1, §12.4] nachgelesen werden.

Eine rationale Funktion ist ein Bruch aus zwei Polynomen, nennen wir diese P und Q . D.h. wir wollen die Abbildung

$$f(x) = \frac{P(x)}{Q(x)}$$

integrieren. P und Q haben damit folgende Form

$$P(x) = \sum_{i=0}^{n_P} p_i x^i$$

$$Q(x) = \sum_{i=0}^{n_Q} q_i x^i$$

mit $n_P, n_Q \in \mathbb{N}$ sowie $p_i, q_i \in \mathbb{R}$ mit $p_{n_P} \neq 0$ und $q_{n_Q} \neq 0$. n_P bzw. n_Q heißen *Grad des Polynoms P bzw. Q* . Die Methode der Partialbruchzerlegung funktioniert nur, falls $n_P < n_Q$, ansonsten müssten wir vorher eine Polynomdivision (siehe z.B. [1, S. 114]) durchführen. Dann nehmen wir weiter an, dass Q nur einfache reelle Nullstellen besitzt, d.h. wir finden $x_1, \dots, x_{n_Q} \in \mathbb{R}$ paarweise verschieden (d.h. alle Nullstellen sind unterscheidlich) und ein $\alpha \in \mathbb{R} \setminus \{0\}$, sodass gilt

$$Q(x) = \alpha \prod_{i=0}^{n_Q} (x - x_i).$$

Anders ausgedrückt haben wir Q als Produkt aus *Linearfaktoren* geschrieben. Dann machen wir folgenden Ansatz:

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \dots + \frac{A_{n_Q}}{x - x_{n_Q}}$$

mit $A_1, \dots, A_{n_Q} \in \mathbb{R}$. Man kann zeigen, dass unter den gegebenen Voraussetzungen dies immer funktioniert (d.h. $n_P < n_Q$ und die Nullstellen x_i paarweise verschieden). Nun müssen wir diese unbekanntenen Koeffizienten A_i berechnen. Dazu bringen wir die rechte Seite der Gleichung auf einen einzigen Bruch:

$$\begin{aligned} \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \dots + \frac{A_{n_Q}}{x - x_{n_Q}} &= \frac{\alpha \sum_{j=1}^{n_Q} A_j \prod_{i=1, i \neq j}^{n_Q} (x - x_i)}{\alpha \prod_{i=0}^{n_Q} (x - x_i)} \\ &= \frac{\alpha \sum_{j=1}^{n_Q} A_j \prod_{i=1, i \neq j}^{n_Q} (x - x_i)}{Q(x)}. \end{aligned}$$

Nun können wir den Zähler ausmultiplizieren und die Koeffizienten des resultierenden Polynoms mit denen von P vergleichen. So erhalten wir ein lineares Gleichungssystem für die A_i , welches wir lösen können.

Nun können wir das Integral mithilfe des natürlichen Logarithmusses berechnen:

$$\begin{aligned} \int \frac{P(x)}{Q(x)} dx &= \int \sum_{i=1}^{n_Q} \frac{A_i}{x - x_i} dx = \sum_{i=1}^{n_Q} \int \frac{A_i}{x - x_i} dx \\ &= \sum_{i=1}^{n_Q} A_i \ln(|x - x_i|). \end{aligned}$$

Fassen wir zusammen:

- 1) Mithilfe einer Polynomdivision führen wir das Problem darauf zurück, dass der Grad im Nenner (von Q) echt größer als der Grad im Zähler (von P) ist.
- 2) Man berechne die Nullstellen x_i von Q und stelle sicher, dass diese paarweise verschieden sind und es wirklich n_Q Stück sind.
- 3) Man führe eine Partialbruchzerlegung von $\frac{P}{Q}$ mithilfe des Ansatzes

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{n_Q} \frac{A_i}{x - x_i}$$

durch.

- 4) Man berechne die A_i mithilfe eines Koeffizientenvergleichs.
- 5) Man integriert die einzelnen Summanden mithilfe des natürlichen Logarithmusses.

Machen wir ein Beispiel:

$$f(x) = \frac{x}{(x-1) \cdot (x-3)}.$$

Hier gilt $P(x) = x$ und $Q(x) = (x-1) \cdot (x-3)$, also bereits in Linearfaktoren zerlegt. Die Nullstellen von Q sind hier also $x_1 = 1$ und $x_2 = 3$. Diese sind verschieden, also können wir weiter nach unserem Kochrezept verfahren und machen nun einen Ansatz

$$\begin{aligned} \frac{x}{(x-1) \cdot (x-3)} &= \frac{A_1}{x-1} + \frac{A_2}{x-3} = \frac{A_1 \cdot (x-3) + A_2 \cdot (x-1)}{(x-1) \cdot (x-3)} \\ &= \frac{(A_1 + A_2) \cdot x - 3 \cdot A_1 - 1 \cdot A_2}{(x-1) \cdot (x-3)}. \end{aligned}$$

Vergleichen wir die Koeffizienten, erhalten wir

$$\begin{aligned} A_1 + A_2 &= 1 \\ -3 \cdot A_1 - A_2 &= 0. \end{aligned}$$

Die Lösung dafür lautet

$$A_1 = -\frac{1}{2} \text{ und } A_2 = \frac{3}{2}.$$

Damit gilt

$$\begin{aligned}\int \frac{x}{(x-1) \cdot (x-3)} dx &= \int -\frac{1}{2 \cdot (x-1)} + \frac{3}{2 \cdot (x-3)} dx \\ &= -\frac{1}{2} \ln(|x-1|) + \frac{3}{2} \ln(|x-3|).\end{aligned}$$

Für den Fall, dass die Nullstellen nicht paarweise disjunkt sind, oder weniger als n_Q vorhanden sind (wenn diese nicht mehr reellwertig sondern komplexwertig werden), so kann man immer noch arbeiten. Wir verweisen hier wieder aus Zeitgründen lediglich auf [1, §12.4].

B Differentialgleichungen

Wir werden in diesem Abschnitt einen kleinen Ausblick in die Welt der Differentialgleichungen wagen. Aus Zeitgründen können wir natürlich nicht in die Tiefe gehen, wir verweisen dafür auf [1, §13]. Differentialgleichungen spielen eine wichtige Rolle in der Modellierung von Phänomenen in der Natur, unsere Gesellschaft usw.

Was ist nun eine Differentialgleichung? Eine Differentialgleichung beschreibt einen Zusammenhang einer Abbildung zu ihren Ableitungen. Ein einfaches Beispiel wäre

$$y'(x) = y(x).$$

Eine Lösung einer solchen Differentialgleichung ist also eine *Abbildung*. Im obigen Beispiel können wir durch raten eine Lösung bekommen:

$$y(x) = e^x.$$

Wir können diese Idee modifizieren und erkennen, dass die Gleichung unendlich viele Lösungen besitzt. Die allgemeine Lösung lautet

$$y(x) = y_0 \cdot e^{x-x_0} \text{ mit } y_0, x_0 \in \mathbb{R}.$$

Wir haben also gewisse Freiheiten in der Lösungswahl. Diese können wir fixieren, indem wir sogenannte *Anfangsdaten* zusätzlich an die Lösung der Differentialgleichung fordern. Das Problem lautet nun

$$\begin{cases} y' = y \\ y(x_0) = y_0. \end{cases}$$

Diese Art von Problem nennt man *Anfangswertproblem*.

Kurz zur Terminologie: Die Ordnung einer Differentialgleichung ist die höchste beteiligte Ableitungsordnung. In unserem Beispiel wäre dies also 1.

Kommen wir nun zu einer Methode um eine gewisse Klasse von Anfangswertproblemen zu lösen, die sogenannte *Trennung der Variablen* (kurz T.d.V.). Diese ist wirksam, sobald die rechte Seite (die ohne Ableitungen) einer Differentialgleichung eine Produktstruktur besitzt, d.h. die Terme in der gesuchten Abbildung y und die Terme mit den Eingabewerten x multiplikativ getrennt ist. Formal bedeutet dies, dass die Gleichung von der Form

$$y' = f(y) \cdot h(x)$$

sein muss. Betrachten wir dazu ein Beispiel:

$$\begin{cases} y' = y^2 \cdot x \\ y(x_0) = y_0 \end{cases}$$

Hier ist $f(y) = y^2$ und $h(x) = x$. Wir bringen alle Terme in denen ein y enthalten ist auf die Seite vom y' und alle Terme mit x auf die andere Seite. Wir erhalten

$$\frac{y'(x)}{(y(x))^2} = x.$$

Wir integrieren nun diese Gleichung von x_0 bis x und bekommen

$$\int_{x_0}^x \frac{y'(t)}{(y(t))^2} dt = \int_{x_0}^x t dt.$$

Um die linke Seite auszurechnen benutzen wir die Substitutionsformel. Wir setzen $z := y(t)$. Dann gilt $t = y^{-1}(z)$. Also gilt mithilfe der Ableitung der Umkehrabbildung

$$\frac{dt}{dz}(z) = \frac{1}{y'(y^{-1}(z))}.$$

Die Grenzen transformieren sich zu $y(x_0) = y_0$ und $y(x)$:

$$\begin{aligned} \int_{x_0}^x \frac{y'(t)}{(y(t))^2} dt &= \int_{y_0}^{y(x)} \frac{y'(y^{-1}(z))}{(y(y^{-1}(z)))^2} \cdot \frac{1}{y'(y^{-1}(z))} dz = \int_{y_0}^{y(x)} \frac{1}{z^2} dz \\ &= -\frac{1}{z} \Big|_{y_0}^{y(x)} = \frac{1}{y_0} - \frac{1}{y(x)} \end{aligned}$$

Also erhalten wir insgesamt

$$\frac{1}{y_0} - \frac{1}{y(x)} = \int_{x_0}^x t dt = \frac{x}{2} - \frac{x_0}{2}.$$

Auflösen nach $y(x)$ liefert

$$y(x) = \frac{1}{\frac{1}{y_0} + \frac{x_0 - x}{2}}.$$

Wir können prüfen ob wir korrekt gerechnet haben, indem wir dies in unser gegebenes Anfangswertproblem einsetzen. Also löst das gefundene y wirklich unser Problem.

Betrachten wir nun noch eine Gleichung, welche in der Physik gerne zum modellieren von Schwingungen benutzt wird. Ein Beispiel einer solchen Gleichung von Ordnung 2 wäre.

$$y'' + y' - 2 \cdot y = 0.$$

Solche Gleichungen behandelt man mit einem sogenannten *Exponentialansatz*:

$$y(x) = e^{\alpha \cdot x}$$

Setzen wir diesen Ansatz in unsere Gleichung ein, erhalten wir

$$0 = \alpha^2 \cdot e^{\alpha \cdot x} + \alpha \cdot e^{\alpha \cdot x} - 2 \cdot e^{\alpha \cdot x} = (\alpha^2 + \alpha - 2) \cdot e^{\alpha \cdot x}.$$

Da $e^{\alpha \cdot x} \neq 0$, muss also $\alpha^2 + \alpha - 2 = 0$ sein. Mithilfe der pq -Formel (siehe Abschnitt 1.6) können wir dies lösen und erhalten

$$\alpha_{1/2} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = \frac{-1 \pm \sqrt{9}}{2} = \frac{-1 \pm 3}{2}.$$

Also sind die Lösungen $\alpha_1 = 1$ und $\alpha_2 = -2$. Lösungen sind also $x \mapsto e^x$ und $x \mapsto e^{(-2) \cdot x}$. Diese können wir noch linear kombinieren und man kann zeigen, dass alle Lösungen der Gleichung die folgende Gestalt haben müssen

$$x \mapsto \beta \cdot e^x + \gamma \cdot e^{(-2) \cdot x} \text{ mit } \beta, \gamma \in \mathbb{R}.$$

In der Praxis wird es vorkommen, dass die pq -Formel keine Lösung liefert. Dann muss man auf sogenannte komplexe Lösungen (siehe z.B. [1, §5]) ausweichen. Das prinzipielle Verfahren ist aber dasselbe.

Nun haben wir Methoden kennengelernt Differentialgleichungen zu lösen. In der Praxis sind diese aber so kompliziert, dass wir keine explizite Lösung mehr erwarten können. In diesem Fall bedient man sich gerne der Computerunterstützung um eine numerische Lösung zu generieren. Aus gegeben Anlass betrachten wir nun ein solches Beispiel. Die folgende Differentialgleichung ist das *SIR-Modell*, welches 1927 von Kermack und McKendrick entwickelt wurde um Ausbrüche von Krankheiten zu modellieren (siehe [3] für genauere Referenzen, sowie weitergehende Modelle).

$$\begin{aligned}S'(t) &= -\beta \cdot \frac{S(t)}{N} \cdot I(t) \\I'(t) &= \beta \cdot \frac{S(t)}{N} \cdot I(t) - \gamma \cdot I(t) \\R'(t) &= \gamma \cdot I(t).\end{aligned}$$

Dabei sind

- S : Susceptibles, d.h. Anzahl Personen die sich noch infizieren können,
- I : Infected, d.h. Anzahl infizierter Personen, die die susceptibles anstecken können,
- R : Removed, d.h. Anzahl der Personen, die immunisiert bzw. verstorben sind,
- N : Gesamtzahl der Bevölkerung, also $N = S + I + R$,
- β : Entspricht der Wahrscheinlichkeit, dass falls ein Infizierter einen Susceptible trifft, diesen auch ansteckt,
- γ : Entspricht der Rate, mit der Infizierte geheilt werden oder versterben.
- t : time, d.h. die Zeit.

Dieses Modell gibt keine Vorhersage über die Tödlichkeit einer Krankheit, denn in R sind sowohl geheilte Patienten auch als auch an der Krankheit verstorbene Personen enthalten. Diese Zahl ist die Gesamtheit der Personen, die die Krankheit nicht mehr weitergeben können.

Der Ausdruck $\beta \cdot \frac{S(t)}{N} \cdot I(t)$ gibt an, wie wahrscheinlich es ist, dass eine infizierte Person eine ansteckbare Person trifft und diese auch tatsächlich infiziert.

$\gamma \cdot I(t)$ entspricht nun der Rate mit der Infizierte zu den Removed übergehen, d.h. sich immunisieren oder versterben.

Eine wichtige Kenngröße des Systems ist die Reproduktionszahl R_0 :

$$R_0 := \frac{\beta}{\gamma},$$

denn diese ist unabhängig von der Zeiteinheit in der β und γ gemessen werden. Damit gibt diese also Aufschluss über die Gefährlichkeit der zu modellierenden Krankheit. Desto größer R_0 ist, umso mehr Menschen werden sich infizieren.

Diese Differentialgleichung unter Anfangsdaten (d.h. $S(0)$, $R(0)$ und $I(0)$ vorgegeben) exakt zu lösen ist leider nicht möglich. Wir können allerdings einige Eigenschaften feststellen. Z.B. ist $S + I + R$ konstant:

$$(S + I + R)' = S' + I' + R' = -\beta \cdot \frac{S(t)}{N} \cdot I(t) + \beta \cdot \frac{S(t)}{N} \cdot I(t) - \gamma \cdot I(t) + \gamma \cdot I(t) = 0$$

Damit gilt für alle Zeiten t , dass $N = S(0) + I(0) + R(0) = S(t) + I(t) + R(t)$, also bleibt die Gesamtzahl der Bevölkerung gleich, d.h. z.B. dass dort auch weiterhin die Verstorbenen mitgezählt werden. Weitere analytische Eigenschaften werden in [3, §2] hergeleitet)

Um nun doch zu quantifizierbaren Ergebnissen zu kommen, bedient man sich der Computerunterstützung und löst das Problem *numerisch*, also als Approximation mit dem Computer. Auf der Website der Vorlesung finden Sie ein kleines Computerprogramm mit der die folgenden Bilder erstellt wurden.

Diese Bilder (d.h. B.1,B.2, B.3 und B.4) wurden mit denselben Anfangsdaten erzeugt, d.h.

$$S(0) = 100000, I(0) = 50, R(0) = 0$$

und es wurden lediglich die Parameter β und γ geändert.

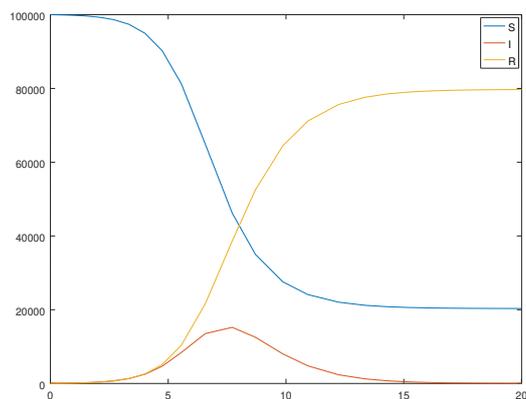


Abbildung B.1: SIR Modell mit Parametern $\beta = 2, \gamma = 1$.

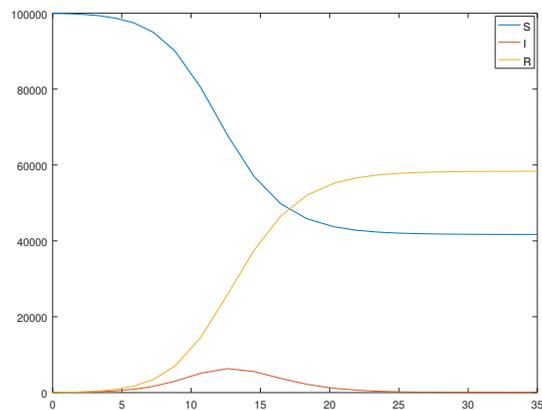


Abbildung B.2: SIR Modell mit Parametern $\beta = 1.5, \gamma = 1$.

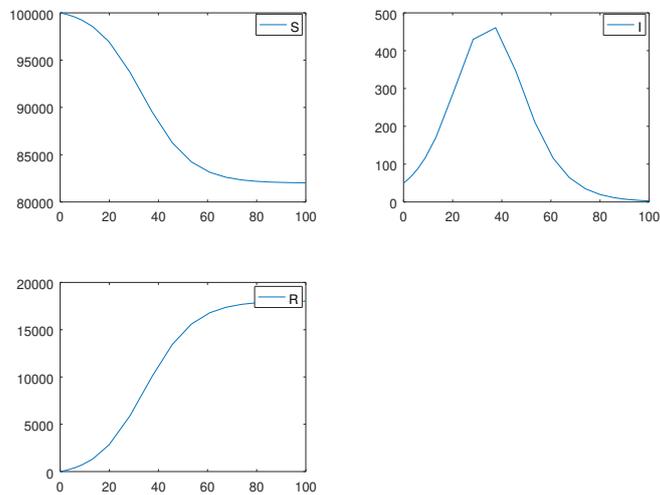


Abbildung B.3: SIR Modell mit Parametern $\beta = 1.1$, $\gamma = 1$.

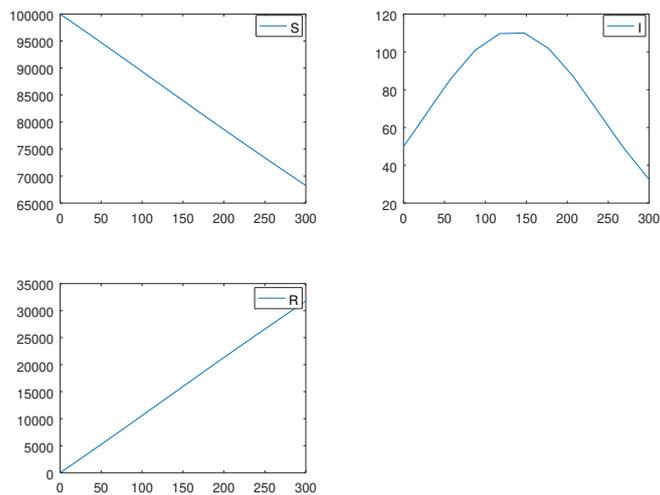


Abbildung B.4: SIR Modell mit Parametern $\beta = 0.7$, $\gamma = 1$.

C Griechische Buchstaben

Großbuchstaben	Kleinbuchstaben	Name
A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ϵ, ε	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ, ϑ	Theta
I	ι	Iota
K	κ, \varkappa	Kappa
Λ	λ	Lambda
M	μ	My
N	ν	Ny
Ξ	ξ	Xi
O	\omicron	Omikron
Π	π, ϖ	Pi
P	ρ, ϱ	Rho
Σ	σ, ς	Sigma
T	τ	Tau
Υ	υ	Ypsilon
Φ	ϕ, φ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Literatur

- [1] T. Arens, F. Hettlich, C. Karpfinger, U. Kockelkorn, K. Lichtenegger, and H. Stachel. *Mathematik*. Springer, 4. edition, 2018.
- [2] A. Berger and S. Rechner. Vorkurs Informatik. Martin-Luther-Universität Halle-Wittenberg, Wintersemester 2019/20. Vorlesungsnotizen und Übungsaufgaben.
- [3] R. Schaback. On COVID-19 Modelling. *Jahresber Dtsch Math-Ver*, 122:167–205, 2020. <https://doi.org/10.1365/s13291-020-00219-9>.
- [4] E. Teufl. Vorkurs Mathematik. gehalten Wintersemester 2019/20 in Tübingen.
- [5] W. Walter. *Analysis 1*. Springer, 7. edition, 2003.

Index

- SIR*-Modell, 66
- pq*-Formel, 14
- Äquivalenzumformung, 10
- äquidistante Zerlegung, 43
- Abbildung
 - abschnittsweise definiert, 29
 - Definitionsbereich, 28
 - Komposition, 28
 - Nullstelle, 30
 - Periode, 31
 - Umkehrabbildung, 29, 30
 - Wertebereich, 28
 - wohldefiniert, 28
- Anfangsdaten, 64
- Anfangswertproblem, 64
- Arkussinus, 32
- Arkussinus, 31
- Arkustangens, 32
- Axiome von Kolmogorov, 52
- Betrag, 19
- Bogenmaß, 30
- Bruch, 5
 - erweitern, 6
 - Hauptnenner, 6
 - kürzen, 6
 - Kehrwert, 6
- Differenzialquotient, 36
- Dimension, 27
- Diskriminante, 14
- Dreiecksungleichung, 19
- Elementarereignis, 55
- Ereignismenge, 51
- Erwartungswert, 57
- Exponentialansatz, 65
- Extremum
 - hinreichendes Kriterium, 42
 - lokal, 40
 - notwendiges Kriterium, 41
- ganze Zahlen, 4
- Gerade, 35
 - Steigung, 35
- Grad eines Polynoms, 61
- Integral
 - unbestimmt, 48
- Intervall, 25
- Kettenregel, 38
- kleinste gemeinsame Vielfache, 6
- L'Hospital, 39
- Linearfaktor, 61
- logarithmische Integration, 48
- Menge
 - Differenz, 24
 - Element, 22
 - Komplement, 25
 - leer, 23
 - Schnitt, 23
 - Teilmenge, 22
 - Vereinigung, 24
- Mengendiagramm, 22
- Minusklammer, 9
- Mitternachtsformel, 14
- Monotonie, 17
- natürliche Zahlen, 4
- natürlicher Logarithmus, 15
- Normalverteilung, 59
- numerisch, 67
- Partialbruchzerlegung, 61
- partielle Integration, 49
- Polynome, 34
- Potenzmenge, 51
- Produktregel, 38
- quadratische Ergänzung, 14
- rationale Zahlen, 5
- Rechengesetze
 - Ableitung, 39
 - Brüche, 9
 - Integral, 44
 - Logarithmen, 16
 - Potenzen, 12
 - Ungleichungen, 18
 - Zahlen, 9
- Rechteck, 27, 43
- rechtwinkliges Dreieck, 30

reelle Zahlen, 7
Regel von Laplace, 55
Riemannsche Summe, 44

Satz von Carathéodory, 54
Sekante, 35
Stammfunktion, 47
Standardabweichung, 58
Steigungsdreieck, 35
Substitutionsregel, 49
Summenzeichen
 Indexschift, 11
 Umkehrung der Summationsreihen-
 folge, 11

Tangente, 35
Trennung der Variablen, 64

Varianz, 58

Wahrscheinlichkeitsdichte, 59
Wahrscheinlichkeitsmaß, 52

Zufallsvariable, 57