ANALYSIS 1

Wolfgang Soergel

31. Januar 2012

In den ersten beiden Abschnitten habe ich Notationen und Begriffsbildungen zusammengefaßt, von denen ich mir vorstelle, daß sie zu Beginn des Studiums in enger Abstimmung zwischen den beiden Grundvorlesungen erklärt werden könnten. Ich habe mir große Mühe gegeben, an jeder Stelle nur soviel Begrifflichkeit einzuführen, wie gerade eben nötig ist, um einerseits einen glatten und transparenten Fluß der Argumentation zu ermöglichen und andererseits regelmäßig motivierende Anwendungen geben zu können. Sowohl die Anwendungen als auch der durchsichtige Aufbau der Theorie erfordern jedoch in meinen Augen eine nach Möglichkeit koordinatenfreie Darstellung und damit den Aufbau der zugehörigen Begrifflichkeit, so daß am Schluß im Vergleich zu anderen Texten doch eher mehr abstrakte Konzepte eingeführt werden. Wie Sie noch an verschiedenen Stellen merken werden, will ich Sie eben am liebsten davon überzeugen, daß das Abstrakte das eigentlich Konkrete ist!

Inhaltsverzeichnis

1 Einstimmung

1.1 Vollständige Induktion und binomische Formel

Satz 1.1.1. *Für jede natürliche Zahl* $n \geq 1$ *gilt* $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$.

Beweis. Bei diesem Beweis sollen Sie gleichzeitig das Beweisprinzip der **vollständigen Induktion** lernen. Wir bezeichnen mit $A(n)$ die Aussage, daß die Formel im Satz für ein gegebenes n gilt, und zeigen:

Induktionsbasis: Die Aussage $A(1)$ ist richtig. In der Tat gilt die Formel $1 = \frac{1(1+1)}{2}$.

Induktionsschritt: Aus $A(n)$ folgt $A(n+1)$. In der Tat, unter der Annahme, daß unsere Formel für ein gegebenes n gilt, der sogenannten **Induktionsannahme** oder **Induktionsvoraussetzung**, rechnen wir

$$\begin{aligned} 1 + 2 + \ldots + n + (n+1) &= \tfrac{n(n+1)}{2} + \tfrac{2(n+1)}{2} \\ &= \tfrac{(n+2)(n+1)}{2} \\ &= \tfrac{(n+1)((n+1)+1)}{2} \end{aligned}$$

und folgern so, daß die Formel auch für $n+1$ gilt.

Es ist damit klar, daß unsere Aussage $A(n)$ richtig ist alias daß unsere Formel gilt für alle $n = 1, 2, 3, \ldots$. $\square$

1.1.2. Das Zeichen $\square$ deutet in diesem Text das Ende eines Beweises an und ist in der neueren Literatur weit verbreitet. Buchstaben in Formeln werden in der Mathematik üblicherweise kursiv notiert, so wie etwa das n oder auch das A im vorhergehenden Beweis. Nur Buchstaben oder Buchstabenkombinationen, die stets dasselbe bedeuten sollen, schreibt man nicht kursiv, wie etwa $\sin$ für den Sinus oder $\log$ für den Logarithmus. Diese Konvention steht in gewissem Widerspruch zur in der Physik üblichen Konvention, Abkürzungen für Einheiten kursiv zu setzen, wie etwa m für "Meter".

1.1.3. Der vorhergehende Beweis stützt sich auf unser intuitives Verständnis der natürlichen Zahlen. Man kann das Konzept der natürlichen Zahlen auch formal einführen und so die natürlichen Zahlen in gewisser Weise "besser" verstehen. Das mögen Sie in der Logik lernen. Das Wort "Induktion" meint eigentlich "Hervorrufen", so wie etwa das Betrachten einer Wurst die Ausschüttung von Spucke induziert alias uns den Mund wässrig macht. Im Zussammenhang der vollständigen Induktion ist es dahingehend zu verstehen, daß die Richtigkeit unserer Aussage $A(0)$ die Richtigkeit von $A(1)$ induziert, die Richtigkeit von $A(1)$ hinwiederum die Richtigkeit von $A(2)$, die Richtigkeit von $A(2)$ die Richtigkeit von $A(3)$, und immer so weiter.

1.1.4. Es herrscht keine Einigkeit in der Frage, ob man die Null eine natürliche Zahl nennen soll. In diesem Text ist stets die Null mit gemeint, wenn von natürlichen Zahlen die Rede ist. Wollen wir die Null dennoch ausschließen, so sprechen wir wie oben von einer "natürlichen Zahl $n \geq 1$".

1.1.5. Ich will kurz begründen, warum es mir natürlich scheint, auch die Null eine natürliche Zahl zu nennen: Hat bildlich gesprochen jedes Kind einer Klasse einen Korb mit Äpfeln vor sich und soll seine Äpfel zählen, so kann es ja durchaus vorkommen, daß in seinem Korb gar kein Apfel liegt, weil es zum Beispiel alle seine Äpfel bereits gegessen hat. In der Begrifflichkeit der Mengenlehre ausgedrückt, die wir in 2.1 einführen werden, muß man die leere Menge endlich nennen, wenn man erreichen will, daß jede Teilmenge einer endlichen Menge wieder endlich ist. Will man dann zusätzlich erreichen, daß die Kardinalität jeder endlichen Menge eine natürliche Zahl ist, so darf man die Null nicht aus den natürlichen Zahlen herauslassen.

1.1.6. Man kann sich den Satz anschaulich klar machen als eine Formel für die Fläche eines Querschnitts für eine Treppe der Länge n mit Stufenabstand und Stufenhöhe eins. In der Tat bedeckt ein derartiger Querschnitt ja offensichtlich ein halbes Quadrat der Kantenlänge n nebst n halben Quadraten der Kantenlänge Eins. Ein weiterer Beweis geht so:

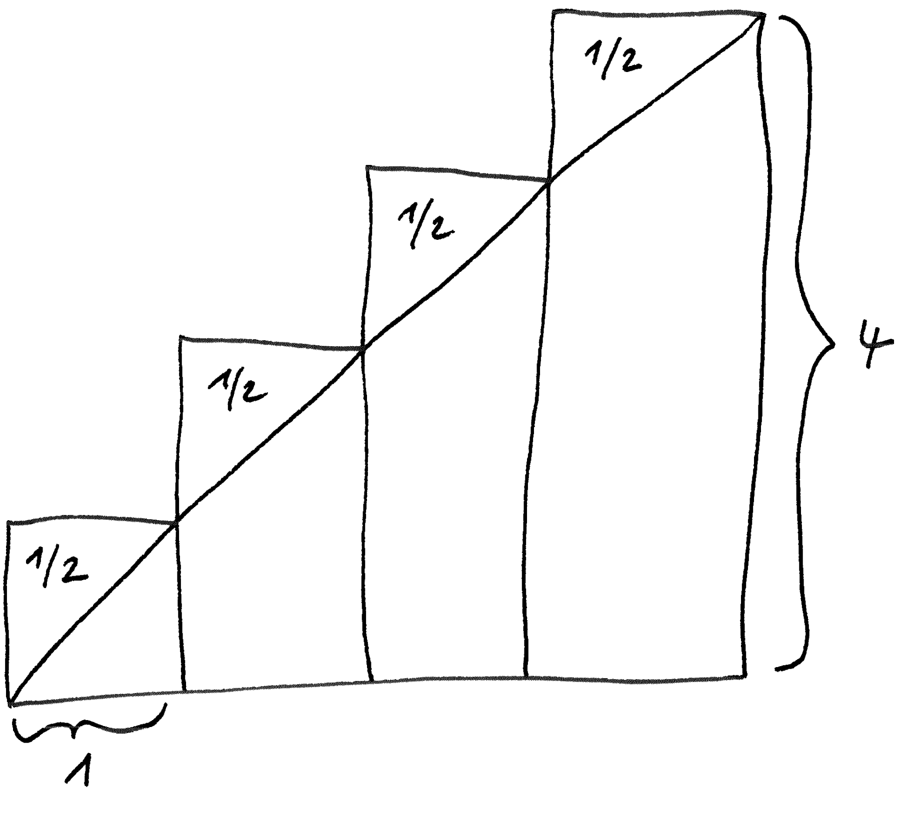
$$\begin{array}{rcccccc} 1+2+\ldots+n & = & 1/2 & + & 2/2 & +\ldots & +\,n/2 \\ & & +\,n/2 & + & (n-1)/2 & +\ldots & +\,1/2 \\ & = & \frac{n+1}{2} & + & \frac{n+1}{2} & +\ldots & +\,\frac{n+1}{2} \\ & = & n(n+1)/2 & & & & \end{array}$$

Ich will diesen Beweis benutzen, um eine neue Notation einzuführen.

Definition 1.1.7. Gegeben $a_1, a_2, \ldots, a_n$ schreiben wir

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \ldots + a_n$$

Das Symbol $\sum$ ist ein großes griechisches S und steht für "Summe". Das Symbol $:=$ deutet an, daß die Bedeutung der Symbole auf der doppelpunktbehafteten Seite des Gleichheitszeichens durch den Ausdruck auf der anderen Seite unseres Gleichheitszeichens definiert ist. Im obigen und ähnlichen Zusammenhängen heißen $a_1, \ldots, a_n$ die **Summanden** und i der **Laufindex**, da er eben etwa in unserem Fall von 1 bis n läuft und anzeigt alias "indiziert", welcher Summand gemeint ist.



Die Gesamtfläche dieses Treppenquerschnitts ist offensichtlich
$4^2/2 + 4/2 = 4 \cdot 5/2$

1.1.8. Das Wort “Definition” kommt aus dem Lateinischen und bedeutet “Abgrenzung”. In Definitionen versuchen wir, die Bedeutung von Symbolen und Begriffen so klar wie möglich festzulegen. Sie werden merken, daß man in der Mathematik die Angwohnheit hat, in Definitionen Worte der Umgangssprache wie Menge, Gruppe, Körper, Unterkörper, Abbildung etc. “umzuwidmen” und ihnen ganz spezielle und meist nur noch entfernt mit der umgangssprachlichen Bedeutung verwandte Bedeutungen zu geben. In mathematischen Texten sind dann durchgehend diese umgewidmeten Bedeutungen gemeint. In dieser Weise baut die Mathematik also wirklich ihre eigene Sprache auf, bei der jedoch die Grammatik und auch nicht ganz wenige Wörter doch wieder von den uns geläufigen Sprachen übernommen werden. Das muß insbesondere für den Anfänger verwirrend sein, der sich auch bei ganz harmlos daherkommenden Wörtern stets wird fragen müssen, ob sie denn nun umgangssprachlich gemeint sind oder vielmehr bereits durch eine Definition festgelegt wurden. Um hier zu helfen, habe ich mir große Mühe mit dem Index (ganz am Schluß dieses Skriptums, bitte nicht als Laufindex mißverstehen) gegeben, in dem alle an verschiedenen Stellen eingeführten oder umgewidmeten und dort fett gedruckten Begriffe verzeichnet sein sollten.

Beispiel 1.1.9. In der $\sum$-Notation liest sich der in 1.1.6 gegebene Beweis so:

$$\begin{array}{rcl}
\sum_{i=1}^n i & = & \sum_{i=1}^n \frac{i}{2} + \sum_{i=1}^n \frac{i}{2} \\
& & \text{und nach Indexwechsel } i = n+1-k \text{ hinten} \\
& = & \sum_{i=1}^n \frac{i}{2} + \sum_{k=1}^n \frac{n+1-k}{2} \\
& & \text{dann mache } k \text{ zu } i \text{ in der zweiten Summe} \\
& = & \sum_{i=1}^n \frac{i}{2} + \sum_{i=1}^n \frac{n+1-i}{2} \\
& & \text{und nach Zusammenfassen beider Summen} \\
& = & \sum_{i=1}^n \frac{n+1}{2} \\
& & \text{ergibt sich offensichtlich} \\
& = & n(\frac{n+1}{2})
\end{array}$$

Beispiel 1.1.10. Ein anderer Beweis derselben Formel kann auch durch die folgende von der Mitte ausgehend zu entwickelnde Gleichungskette gegeben werden:

$$(n+1)^2 = \sum_{i=0}^n (i+1)^2 - i^2 = \sum_{i=0}^n 2i+1 = 2\sum_{i=0}^n i + \sum_{i=0}^n 1 = n+1+2\sum_{i=0}^n i$$

Definition 1.1.11. In einer ähnlichen Bedeutung wie $\sum$ verwendet man auch das Symbol $\prod$, ein großes griechisches P, für “Produkt” und schreibt

$$\prod_{i=1}^n a_i := a_1 a_2 \dots a_n$$

Die $a_1, \ldots, a_n$ heißen in diesem und ähnlichen Zusammenhängen die **Faktoren** des Produkts.

Definition 1.1.12. Für jede natürliche Zahl $n \geq 1$ definieren wir die Zahl $n!$ (sprich: n **Fakultät**) durch die Formel

$$n! := 1 \cdot 2 \cdot \ldots \cdot n = \prod_{i=1}^{n} i$$

Wir treffen zusätzlich die Vereinbarung $0! := 1$ und haben also $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$ und so weiter.

Ergänzung 1.1.13. Wir werden in Zukunft noch öfter Produkte mit überhaupt keinem Faktor zu betrachten haben und vereinbaren deshalb gleich hier schon, daß Produkten, bei denen die obere Grenze des Laufindex um Eins kleiner ist als seine untere Grenze, der Wert 1 zugewiesen werden soll, also etwa $1 = \prod_{i=1}^{0} i$. Ebenso vereinbaren wir auch, daß Summen, bei denen die obere Grenze des Laufindex um Eins kleiner ist als seine untere Grenze, der Wert 0 zugewiesen werden soll, so daß wir in Erweiterung unserer Formel 1.1.1 etwa schreiben könnten $0 = \sum_{i=1}^{0} i$. Der Sinn dieser Erweiterungen zeigt sich darin, daß damit Formeln wie $\sum_{i=k}^{l} a_i = \sum_{i=k}^{m} a_i + \sum_{i=m+1}^{l} a_i$ auch für $m = k - 1$ richtig bleiben. Man mag sogar noch weiter gehen und die Definition von Summen auf beliebige untere und obere Grenzen so erweitern, daß diese Formeln richtig bleiben. In dieser Allgemeinheit ist die fragliche Notation jedoch nur beim kontinuierlichen Analogon $\int$ des Summenzeichens üblich, wie in 6.5.10 ausgeführt wird.

Satz 1.1.14 (Bedeutung der Fakultät). *Es gibt genau $n!$ Möglichkeiten, n voneinander verschiedene Objekte in eine Reihenfolge zu bringen.*

Beispiel 1.1.15. Es gibt genau $3! = 6$ Möglichkeiten, die drei Buchstaben a, b und c in eine Reihenfolge zu bringen, nämlich

$$\begin{array}{ccc} abc & bac & cab \\ acb & bca & cba \end{array}$$

In gewisser Weise stimmt unser Satz sogar für $n = 0$: In der Terminologie, die wir in 4.2 einführen, gibt es in der Tat genau eine Anordnung der leeren Menge.

Beweis. Hat man n voneinander verschiedene Objekte, so hat man n Möglichkeiten, ein Erstes auszusuchen, dann $(n - 1)$ Möglichkeiten, ein Zweites auszusuchen und so weiter, bis schließlich nur noch eine Möglichkeit bleibt, ein Letztes auszusuchen. Insgesamt haben wir also in der Tat wie behauptet $n!$ mögliche Reihenfolgen. □

Definition 1.1.16. Wir definieren für beliebiges n und jede natürliche Zahl k die **Binomialkoeffizienten** $\binom{n}{k}$ (sprich: n **über** k) durch die Regeln

$$\binom{n}{k} := \prod_{j=0}^{k-1} \frac{n-j}{k-j} = \frac{n(n-1)\dots \quad (n-k+1)}{k(k-1)\dots \qquad 1} \text{ für } k \geq 1 \text{ und } \binom{n}{0} := 1.$$

Der Sonderfall $k = 0$ wird im Übrigen auch durch unsere allgemeine Formel gedeckt, wenn wir unsere Konvention 1.1.13 beherzigen. Im Lichte des folgenden Satzes schlage ich vor, die Binomialkoeffizienten $\binom{n}{k}$ statt "n über k" inhaltsreicher "k aus n" zu sprechen.

1.1.17. Die Bezeichnung als Binomialkoeffizienten leitet sich von dem Auftreten dieser Zahlen als Koeffizienten in der "binomischen Formel" 1.1.22 ab.

Satz 1.1.18 (Bedeutung der Binomialkoeffizienten). *Gegeben natürliche Zahlen n und k gibt es genau $\binom{n}{k}$ Möglichkeiten, aus n voneinander verschiedenen Objekten k Objekte auszuwählen.*

Beispiel 1.1.19. Es gibt genau $\binom{4}{2} = \frac{4 \cdot 3}{2 \cdot 1} = 6$ Möglichkeiten, aus den vier Buchstaben a, b, c, d zwei auszuwählen, nämlich

$$\begin{array}{lll} a, b & b, c & c, d \\ a, c & b, d & \\ a, d & & \end{array}$$

Beweis. Wir haben n Möglichkeiten, ein erstes Objekt auszuwählen, dann $n - 1$ Möglichkeiten, ein zweites Objekt auszuwählen, und so weiter, also insgesamt $n(n-1)\dots(n-k+1)$ Möglichkeiten, k Objekte *der Reihe nach* auszuwählen. Auf die Reihenfolge, in der wir ausgewählt haben, kommt es uns aber gar nicht an, jeweils genau $k!$ von unseren $n(n-1)\dots(n-k+1)$ Möglichkeiten führen also nach 1.1.14 zur Auswahl derselben k Objekte. Man bemerke, daß unser Satz auch im Extremfall $k = 0$ noch stimmt, wenn wir ihn geeignet interpretieren: In der Terminologie, die wir gleich einführen werden, besitzt in der Tat jede Menge genau eine nullelementige Teilmenge, nämlich die leere Menge. □

1.1.20. Offensichtlich gilt für alle natürlichen Zahlen n mit $n \geq k$ die Formel

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

Das folgt einerseits sofort aus der formalen Definition und ist andererseits auch klar nach der oben erklärten Bedeutung der Binomialkoeffizienten: Wenn wir aus

n Objekten k Objekte auswählen, so bleiben $n-k$ Objekte übrig. Es gibt demnach gleichviele Möglichkeiten, k Objekte auszuwählen, wie es Möglichkeiten gibt, $n-k$ Objekte auszuwählen. Wir haben weiter $\binom{n}{n} = \binom{n}{0} = 1$ für jede natürliche Zahl $n \geq 0$ sowie $\binom{n}{1} = \binom{n}{n-1} = n$ für jede natürliche Zahl $n \geq 1$.

Definition 1.1.21. Wie in der Schule setzen wir $a^k := \prod_{i=1}^{k} a$, in Worten ist also gemeint "das Produkt von k-mal dem Faktor a", und verstehen im Lichte von 1.1.13 insbesondere $a^0 = 1$.

Satz 1.1.22. *Für jede natürliche Zahl n gilt die* **binomische Formel**

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

1.1.23. Man beachte, wie wichtig unsere Konvention $a^0 = 1$ und insbesondere auch $0^0 = 1$ für die Gültigkeit dieser Formel ist.

1.1.24. Die Bezeichung "binomische Formel" leitet sich ab von der Vorsilbe "bi" für Zwei, wie etwa in englisch "bicycle" für "Zweirad" alias "Fahrrad", und dem lateinischen Wort "nomen" für "Namen". Mit den beiden "Namen" sind hier a und b gemeint. Mehr dazu wird in 6.1.20 erklärt.

Erster Beweis. Beim Ausmultiplizieren erhalten wir so oft $a^k b^{n-k}$, wie es Möglichkeiten gibt, aus unseren n Faktoren $(a+b)$ die k Faktoren auszusuchen, "in denen wir beim Ausmultiplizieren das b nehmen". Dieses Argument werden wir in 2.1.19 noch besser formulieren. □

Zweiter Beweis. Dieser Beweis ist eine ausgezeichnete Übung im Umgang mit unseren Symbolen und mit der vollständigen Induktion. Er scheint mir jedoch auch in einer für Beweise durch vollständige Induktion typischen Weise wenig durchsichtig. Zunächst prüfen wir für beliebiges n und jede natürliche Zahl $k \geq 1$ die Formel

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

durch explizites Nachrechnen. Dann geben wir unserer Formel im Satz den Namen $A(n)$ und prüfen die Formel $A(0)$ und zur Sicherheit auch noch $A(1)$ durch Hinsehen. Schließlich gilt es, den Induktionsschritt durchzuführen, als da heißt,

$A(n+1)$ aus $A(n)$ zu folgern. Dazu rechnen wir

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&\quad \text{und mit der Induktionsvoraussetzung} \\
&= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\
&\quad \text{und durch Ausmultiplizieren} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&\quad \text{und Indexwechsel } k = i-1 \text{ in der ersten Summe} \\
&= \sum_{i=1}^{n+1}\binom{n}{i-1}a^i b^{n-i+1} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&\quad \text{dann mit } k \text{ statt } i \text{ und Absondern von Summanden} \\
&= a^{n+1}b^0 + \sum_{k=1}^{n}\binom{n}{k-1}a^k b^{n-k+1} + \\
&\qquad + \sum_{k=1}^{n}\binom{n}{k}\, a^k b^{n-k+1} + a^0 b^{n+1} \\
&\quad \text{und nach Zusammenfassen der mittleren Summen} \\
&= a^{n+1}b^0 + \sum_{k=1}^{n}\binom{n+1}{k}a^k b^{n-k+1} + a^0 b^{n+1} \\
&\quad \text{und Einbeziehen der abgesonderten Summanden} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^k b^{n+1-k}
\end{aligned}$$

und folgern so tatsächlich $A(n+1)$ aus $A(n)$. □

1.1.25. Die Formel $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ für $k \geq 1$ kann man zur effektiven Berechnung der Binomialkoeffizienten mit dem sogenannten **Pascal'schen Dreieck** benutzen: Im Schema

$$\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}$$

seien die Einsen an den Rändern vorgegeben und eine Zahl in der Mitte berechne sich als die Summe ihrer beiden oberen "Nachbarn". Dann stehen in der $(n+1)$-ten Zeile der Reihe nach die Binomialkoeffizienten $\binom{n}{0} = 1, \binom{n}{1} = n, \ldots$ bis $\binom{n}{n-1} = n, \binom{n}{n} = 1$. Wir haben also zum Beispiel

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Übung 1.1.26. Man finde und beweise eine Formel für $\sum_{i=1}^{n} i^2$. Hinweis: Man suche zunächst eine Formel für $\sum_{i=1}^{n} i^3 - (i-1)^3$ und beachte $i^3 - (i-1)^3 = 3i^2 - 3i + 1$.

1.2 Fibonacci-Folge und Vektorraumbegriff

Beispiel 1.2.1. Die **Fibonacci-Folge**

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\dots$$

entsteht, indem man mit $f_0 = 0$ und $f_1 = 1$ beginnt und dann jedes weitere Folgenglied als die Summe seiner beiden Vorgänger bildet. Wir suchen nun für die Glieder f_i dieser Folge eine geschlossene Darstellung. Dazu vereinbaren wir, daß wir Folgen $x_0, x_1, x_2, \dots$ mit der Eigenschaft $x_n = x_{n-1} + x_{n-2}$ für $n = 2, 3, 4, \dots$ **Folgen vom Fibonacci-Typ** nennen wollen. Kennen wir die beiden ersten Glieder einer Folge vom Fibonacci-Typ, so liegt natürlich bereits die gesamte Folge fest. Nun bemerken wir, daß für jede Folge $x_0, x_1, x_2, \dots$ vom Fibonacci-Typ und jedes α auch die Folge $\alpha x_0, \alpha x_1, \alpha x_2, \dots$ vom Fibonacci-Typ ist, und daß für jede weitere Folge $y_0, y_1, y_2, \dots$ vom Fibonacci-Typ auch die gliedweise Summe $(x_0 + y_0), (x_1 + y_1), (x_2 + y_2), \dots$ eine Folge vom Fibonacci-Typ ist. Der Trick ist dann, danach zu fragen, für welche β die Folge $x_i = \beta^i$ vom Fibonacci-Typ ist. Das ist ja offensichtlich genau dann der Fall, wenn gilt $\beta^2 = \beta + 1$, als da heißt für $\beta_\pm = \frac{1}{2}(1 \pm \sqrt{5})$. Für beliebige c, d ist mithin die Folge

$$x_i = c\beta_+^i + d\beta_-^i$$

vom Fibonacci-Typ, und wenn wir c und d bestimmen mit $x_0 = 0$ und $x_1 = 1$, so ergibt sich eine explizite Darstellung unserer Fibonacci-Folge. Wir suchen also c und d mit

$$\begin{array}{rcl} 0 & = & c + d \\ 1 & = & c\left(\frac{1}{2}(1+\sqrt{5})\right) + d\left(\frac{1}{2}(1-\sqrt{5})\right) \end{array}$$

und folgern leicht $c = -d$ und $1 = c\sqrt{5}$ alias $c = 1/\sqrt{5} = -d$. Damit ergibt sich schließlich für unsere ursprüngliche Fibonacci-Folge die explizite Darstellung

$$f_i = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^i - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^i$$

Im übrigen ist der zweite Summand hier immer kleiner als $1/2$, so daß wir f_i auch beschreiben können als diejenige ganze Zahl, die am nächstem am ersten Summanden liegt. Es wäre rückblickend natürlich ein Leichtes gewesen, diese Formel einfach zu "raten" um sie dann mit vollständiger Induktion 1.1.1 zu beweisen. Diese Art mathematischer Zaubertricks halte ich jedoch für unehrenhaft. Ich werde deshalb stets nach Kräften versuchen, das Tricksen zu vermeiden, auch wenn die Beweise dadurch manchmal etwas länger werden sollten. Eine Möglichkeit, auch den letzten verbleibenden Trick aus den vorhergehenden Überlegungen zu eliminieren, zeigt **??**. Die bei unserer Lösung auftretende reelle Zahl $\frac{1}{2}(1+\sqrt{5})$

ist im Übrigen auch bekannt als “goldener Schnitt” aus Gründen, die in nebenstehendem Bild diskutiert werden. In 5.3.4 dürfen Sie dann zur Übung zeigen, daß der Quotient zweier aufeinanderfolgender Fibonacci-Zahlen gegen den goldenen Schnitt strebt, daß also genauer und in Formeln für unsere Fibonacci-Folge $f_0, f_1, f_2, \ldots$ von oben gilt

$$\lim_{i\to\infty} \frac{f_{i+1}}{f_i} = \frac{1+\sqrt{5}}{2}$$

Übung 1.2.2. Kann man für jede Folge $x_0, x_1, \ldots$ vom Fibonacci-Typ Zahlen c, d finden mit $x_i = c\beta_+^i + d\beta_-^i$ für alle i? Finden Sie eine geschlossene Darstellung für die Glieder der Folge, die mit $0, 0, 1$ beginnt und dem Bildungsgesetz $x_n = 2x_{n-1} + x_{n-2} - 2x_{n-3}$ gehorcht.

Beispiel 1.2.3. Ein System von mehreren Gleichungen, in denen dieselben Unbekannten auftauchen, nennt man auch ein **Gleichungssystem**. Wir betrachten ein “homogenes lineares” Gleichungssystem alias ein Gleichungssystem der Gestalt

$$\begin{array}{lclcl} \alpha_{11}x_1 + \alpha_{12}x_2 + & \ldots & +\alpha_{1m}x_m & = & 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + & \ldots & +\alpha_{2m}x_m & = & 0 \\ & \vdots & & & \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + & \ldots & +\alpha_{nm}x_m & = & 0 \end{array}$$

Wie man zu vorgegebenen $\alpha_{i,j}$ für $1 \le i \le n$ und $1 \le j \le m$ die Menge L aller Lösungen $(x_1, \ldots, x_m)$ ermittelt, sollen sie später in dieser Vorlesung lernen. Zwei Dinge aber sind a priori klar:

1. Sind $(x_1, \ldots, x_m)$ und $(x_1', \ldots, x_m')$ Lösungen, so ist auch ihre komponentenweise Summe $(x_1 + x_1', \ldots, x_m + x_m')$ eine Lösung;

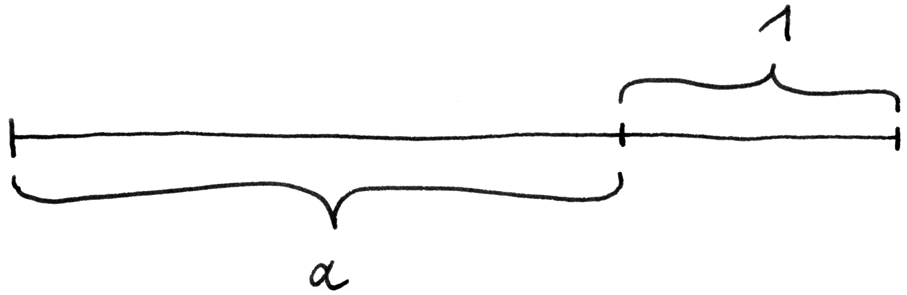
2. Ist $(x_1, \ldots, x_m)$ eine Lösung und α eine reelle Zahl, so ist auch das komponentenweise Produkt $(\alpha x_1, \ldots, \alpha x_m)$ eine Lösung.

Beispiel 1.2.4. Wir betrachten die Menge aller Funktionen $f : \mathbb{R} \to \mathbb{R}$, die zweimal differenzierbar sind und der Differentialgleichung

$$f'' = -f$$

genügen. Lösungen sind zum Beispiel die Funktionen $\sin, \cos$, die Nullfunktion oder auch die Funktionen $f(x) = \sin(x+a)$ für konstantes a. Wie man die Menge L aller Lösungen beschreiben kann, sollen Sie nicht hier lernen. Zwei Dinge aber sind a priori klar:

1. Mit f und g ist auch die Funktion $f + g$ eine Lösung;



Der **goldene Schnitt** ist das Verhältnis, in dem eine Strecke geteilt werden muß, damit das Verhältnis vom größeren zum kleineren Stück gleich dem Verhältnis des Ganzen zum größeren Stück ist, also die positive Lösung der Gleichung $a/1 = (1 + a)/a$ alias $a^2 - a - 1 = 0$, also $a = (1 + \sqrt{5})/2$.

2. Ist f eine Lösung und α eine reelle Zahl, so ist auch αf eine Lösung.

Beispiel 1.2.5. Wir betrachten die Gesamtheit aller Parallelverschiebungen der Tafelebene. Graphisch stellen wir solch eine Parallelverschiebung dar durch einen Pfeil von irgendeinem Punkt zu seinem Bild unter der Verschiebung. Im nebenstehenden Bild stellen etwa alle gepunktelten Pfeile dieselbe Parallelverschiebung dar. Was für ein Ding diese Gesamtheit P aller Parallelverschiebungen eigentlich ist, scheint mir recht undurchsichtig, aber einiges ist a priori klar:

1. Sind p und q Parallelverschiebungen, so ist auch ihre "Hintereinanderausführung" $p \circ q$, sprich "p nach q", eine Parallelverschiebung.

2. Ist α eine reelle Zahl und p eine Parallelverschiebung, so können wir eine neue Parallelverschiebung αp bilden, das "α-fache von p". Bei negativen Vielfachen vereinbaren wir hierzu, daß eine entsprechende Verschiebung in die Gegenrichtung gemeint ist.

3. Führen wir eine neue Notation ein und schreiben für die Hintereinanderausführung $p \dot{+} q := p \circ q$, so gelten für beliebige Parallelverschiebungen p, q, r der Tafelebene und beliebige reelle Zahlen α, β die Formeln

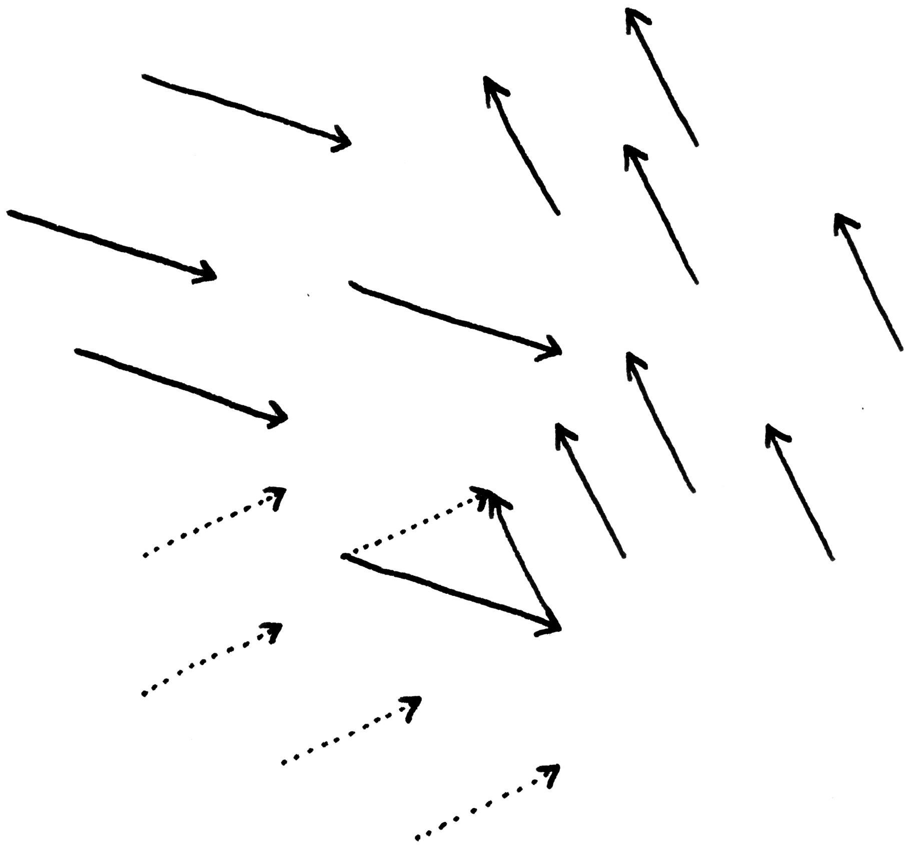
$$\begin{array}{ccc} (p \dot{+} q) \dot{+} r & = & p \dot{+} (q \dot{+} r) \\ p \dot{+} q & = & q \dot{+} p \\ \alpha(\beta p) & = & (\alpha\beta)p \\ (\alpha + \beta)p & = & (\alpha p) \dot{+} (\beta p) \\ \alpha(p \dot{+} q) & = & (\alpha p) \dot{+} (\alpha q) \end{array}$$

Will man sich die Gesamtheit aller Parallelverschiebungen der Tafelebene anschaulich machen, so tut man im Übrigen gut daran, einen Punkt als "Ursprung" auszuzeichnen und jede Parallelverschiebung mit dem Punkt der Tafelebene zu identifizieren, auf den unsere Parallelverschiebung diesen Ursprung abbildet.

Beispiel 1.2.6. Analoges gilt für die Gesamtheit der Parallelverschiebung des Raums unserer Anschauung und auch für die Gesamtheit aller Verschiebungen einer Geraden und, mit noch mehr Mut, für die Gesamtheit aller Zeitspannen.

1.2.7. Die Formeln unserer kleinen Formelsammlung von 1.2.5.3 gelten ganz genauso auch für die Lösungsmenge unserer Differentialgleichung $f'' = -f$, wenn wir $f \dot{+} g := f + g$ verstehen, für die Lösungsmenge unseres linearen Gleichungssystems, wenn wir

$$(x_1, \ldots, x_m) \dot{+} (x'_1, \ldots, x'_m) := (x_1 + x'_1, \ldots, x_m + x'_m)$$



Die Hintereinanderausführung der beiden Parallelverschiebungen der Tafel- oder hier vielmehr der Papierebene, die durch die durchgezogenen Pfeile dargestellt werden, wird die durch die gepunktelten Feile dargestellt.

als “komponentenweise Addition” verstehen, und für die Menge aller Folgen vom Fibonacci-Typ, wenn wir ähnlich die Summe $+$ zweier Folgen erklären. Ein wesentliches Ziel der folgenden Vorlesungen über lineare Algebra ist es, einen abstrakten Formalismus aufzubauen, dem sich alle diese Beispiele unterordnen. Dadurch soll zweierlei erreicht werden:

1. Unser abstrakter Formalismus soll uns dazu verhelfen, die uns als Augentieren und Nachkommen von Ästehüpfern angeborene räumliche Anschauung nutzbar zu machen zum Verständnis der bis jetzt gegebenen Beispiele und der vielen weiteren Beispiele von Vektorräumen, denen Sie im Verlauf Ihres Studiums noch begegnen werden. So werden sie etwa lernen, daß man sich die Menge aller Folgen vom Fibonacci-Typ durchaus als Ebene vorstellen darf und die Menge aller Folgen mit vorgegebenem Folgenglied an einer vorgegebenen Stelle als eine Gerade in dieser Ebene. Suchen wir also alle Folgen vom Fibonacci-Typ mit zwei vorgegebenen Folgengliedern, so werden wir im allgemeinen genau eine derartige Lösung finden, da sich eben zwei Geraden aus einer Ebene im allgemeinen in genau einem Punkt schneiden. In diesem Licht betrachtet soll der abstrakte Formalismus uns also helfen, a priori unanschauliche Fragestellungen der Anschauung zugänglich zu machen. Ich denke, diese Nähe zur Anschauung ist auch der Grund dafür, daß die lineare Algebra meist an den Anfang des Studiums gestellt wird: Von der Schwierigkeit des Formalismus her gesehen gehört sie nämlich keineswegs zu den einfachsten Gebieten der Mathematik, hier würde ich eher an Gruppentheorie oder Graphentheorie oder dergleichen denken.

2. Unser abstrakter Formalismus soll so unmißverständlich sein und seine Spielregeln so klar, daß Sie in die Lage versetzt werden, alles nachzuvollziehen und mir im Prinzip und vermutlich auch in der Realität Fehler nachzuweisen. Schwammige Begriffe wie “Tafelebene” oder “Parallelverschiebung des Raums” haben in einem solchen Formalismus keinen Platz mehr. In diesem Licht betrachtet verfolgen wir mit dem Aufbau des abstrakten Formalismus auch das Ziel einer großen Vereinfachung durch die Reduktion auf die Beschreibung einiger weniger Aspekte der uns umgebenden in ihrer Komplexität kaum präzise faßbaren Wirklichkeit.

Die lineare Algebra hat in meinen Augen drei wesentliche Aspekte: Einen **geometrischen Aspekt**, wie ihn das Beispiel 1.2.5 der Gesamtheit aller Parallelverschiebungen illustriert; einen **algorithmischen Aspekt**, unter den ich das Beispiel 1.2.3 der Lösungsmenge eines linearen Gleichungssystems und insbesondere explizite Verfahren zur Bestimmung dieser Lösungsmenge einordnen würde; und einen **abstrakt-algebraischen Aspekt**, eine Art gedankliches Skelett, das Algorithmik und Geometrie verbindet und Brücken zu vielen weiteren Anwendungen schafft, die man dann auch als das Fleisch auf diesem Gerippe ansehen mag. Ich will im weiteren Verlauf dieser Vorlesungen zur linearen Algebra versuchen, diese drei Aspekte zu einer Einheit zu fügen. Ich hoffe, daß Sie dadurch in die Lage versetzt werden, eine Vielzahl von Problemen mit den verbundenen Kräften Ihrer räumlichen Anschauung, Ihrer algorithmischen Rechenfähigkeiten und Ihres abstrakt-logischen Denkens anzugehen. Als Motivation für den weiteren Fortgang der Vorlesungen über lineare Algebra beschreibe ich nun das "Rückgrat unseres Skeletts" und formuliere ohne Rücksicht auf noch unbekannte Begriffe und Notationen die abstrakte Definition eines reellen Vektorraums.

Definition 1.2.8. Ein **reeller Vektorraum** ist ein Tripel bestehend aus den folgenden drei Dingen:

1. Einer Menge V;
2. Einer Verknüpfung $V \times V \to V$, $(v, w) \mapsto v \dot{+} w$, die V zu einer abelschen Gruppe macht;
3. Einer Abbildung $\mathbb{R} \times V \to V$, $(\alpha, v) \mapsto \alpha v$,

derart, daß für alle $\alpha, \beta \in \mathbb{R}$ und alle $v, w \in V$ gilt:

$$\begin{array}{rcl} \alpha(\beta v) & = & (\alpha\beta)v \\ (\alpha + \beta)v & = & (\alpha v) \dot{+} (\beta v) \\ \alpha(v \dot{+} w) & = & (\alpha v) \dot{+} (\alpha w) \\ 1v & = & v \end{array}$$

Hier ist nun viel zu klären: Was ist eine Menge? Eine Verknüpfung? Eine abelsche Gruppe? Eine Abbildung? Was bedeuten die Symbole $\times$, $\to$, $\mapsto$, $\in$, $\mathbb{R}$? Wir beginnen in der nächsten Vorlesung mit der Klärung dieser Begriffe und Notationen.

1.2.9. Bereits hier will ich jedoch die Symbole α und β erklären: Sie heißen "Alpha" und "Beta" und sind die beiden ersten Buchstaben des griechischen Alphabets, das ja auch nach ihnen benannt ist. Bei der Darstellung von Mathematik hilft es, viele verschiedene Symbole und Symbolfamilien zur Verfügung zu haben. Insbesondere werden die griechischen Buchstaben oft und gerne verwendet. Ich schreibe deshalb hier zum Nachschlagen einmal das griechische Alphabet auf. In der ersten Spalte stehen der Reihe nach die griechischen Kleinbuchstaben, dahinter die zugehörigen Großbuchstaben, dann ihr lateinisches Analogon soweit vorhanden, und schließlich, wie man diesen griechischen Buchstaben auf Deutsch benennt und spricht.

α	A	a	alpha
β	B	b	beta
γ	Γ	g	gamma
δ	Δ	d	delta
ϵ, ε	E	e	epsilon
ζ	Z	z	zeta
η	H	ä	eta
θ, ϑ	Θ	th	theta
ι	I	i	iota
κ	K	k	kappa
λ	Λ	l	lambda
μ	M	m	my, sprich "mü"
ν	N	n	ny, sprich "nü"
ξ	Ξ	x	xi
o	O	o	omikron
π	Π	p	pi
ρ, ϱ	P	r	rho
σ, ς	Σ	s	sigma
τ	T	t	tau
υ	Υ	y	ypsilon
ϕ, φ	Φ	f	phi
χ	X	ch	chi
ψ	Ψ	ps	psi
ω	Ω	oh	omega

2 Naive Mengenlehre und Kombinatorik

2.1 Mengen

2.1.1. Beim Arbeiten mit reellen Zahlen oder räumlichen Gebilden reicht auf der Schule ein intuitives Verständnis meist aus, und wenn die Intuition in die Irre führt, ist ein Lehrer zur Stelle. Wenn Sie jedoch selbst unterrichten oder etwas beweisen wollen, reicht dieses intuitive Verständnis nicht mehr aus. *Im folgenden werden deshalb zunächst der Begriff der reellen Zahlen und der Begriff des Raums zurückgeführt auf Grundbegriffe der Mengenlehre, den Begriff der rationalen Zahlen und elementare Logik.* Bei der Arbeit mit diesen Begriffen führt uns die Intuition nicht so leicht in die Irre, wir geben uns deshalb mit einem intuitiven Verständnis zufrieden und verweisen jeden, der es noch genauer wissen will, auf eine Vorlesung über Logik. Wir beginnen mit etwas naiver Mengenlehre, wie sie von Georg Cantor in den Jahren 1874-1897 begründet wurde, und von der der berühmte Mathematiker David Hilbert einmal sagte: "Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können". Natürlich gab es auch vor der Mengenlehre schon hoch entwickelte Mathematik, bei Carl Friedrich Gauß Tod 1855 gab es diese Theorie noch gar nicht und Fourier fand seine "Fourierentwicklung" sogar bereits zu Beginn des 19.-ten Jahrhunderts. Er behauptete auch gleich in seiner "Théorie analytique de la chaleur", daß sich jede beliebige periodische Funktion durch eine Fouriereihe darstellen lasse, aber diese Behauptung stieß bei anderen berühmten Mathematikern seiner Zeit auf Ablehnung und es entstand darüber ein heftiger Disput. Erst in besagtem "Paradies der Mengenlehre" konnten die Fourier's Behauptung zugrundeliegenden Begriffe soweit geklärt werden, daß dieser Disput nun endgültig beigelegt ist. Ähnlich verhält es sich auch mit vielen anderen Fragestellungen. Da die Mengenlehre darüber hinaus auch vom didaktischen Standpunkt aus eine äußerst klare und durchsichtige Darstellung mathematischer Sachverhalte ermöglicht, hat sie sich als Grundlage der höheren Mathematik und der Ausbildung von Mathematikern an Universitäten schnell durchgesetzt und ist nun weltweit ein wesentlicher Teil des "Alphabets der Sprache der Mathematiker". Man wird an Universitäten sogar geradezu dazu erzogen, geometrischen Argumenten keine Beweiskraft zuzugestehen, und ich halte das bei der Ausbildung von Mathematikern auch für angemessen. Bei der Mathematik-Ausbildung im allgemeinen scheint mir dieses Vorgehen dahingegen nicht zielführend: In diesem Kontext sollte man meines Erachtens nicht mit demselben Maß messen, auch ohne alle Mengenlehre geometrisch erklärte Begriffe wie Gerade und Kreis, Ebene und Raum, als wohldefinierte Objekte der Mathematik zulassen, und geometrischen Argumenten durchaus Beweiskraft zugestehen.

2.1.2. Im Wortlaut der ersten Zeilen des Artikels "Beiträge zur Begründung der transfiniten Mengenlehre (Erster Aufsatz)" von Georg Cantor, erschienen im Jahre 1895, hört sich die Definition einer Menge so an:

> Unter einer **Menge** verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objecten m unserer Anschauung oder unseres Denkens (welche die **Elemente** von M genannt werden) zu einem Ganzen.

Verbinden wir mit einer Menge eine geometrische Vorstellung, so nennen wir ihre Elemente auch **Punkte** und die Menge selbst einen **Raum**. Ein derartiges Herumgerede ist natürlich keine formale Definition und birgt auch verschiedene Fallstricke, vergleiche 2.1.21. Das Ziel dieser Vorlesung ist aber auch nicht eine formale Begründung der Mengenlehre, wie Sie sie später in der Logik kennenlernen können. Sie sollen vielmehr die Bedeutung dieser Worte intuitiv erfassen wie ein Kleinkind, das Sprechen lernt: Indem sie mir und anderen Mathematikern zuhören, wie wir mit diesen Worten sinnvolle Sätze bilden, uns nachahmen, und beobachten, welchen Effekt Sie damit hervorrufen. Unter anderem dazu sind die Übungsgruppen da.

Beispiele 2.1.3. Endliche Mengen gibt man oft durch eine vollständige Liste ihrer Elemente in geschweiften Klammern an, zum Beispiel in der Form $X = \{x_1, x_2, \ldots, x_n\}$. Diese geschweiften Klammern heißen auch **Mengenklammern**. Die Elemente dürfen mehrfach genannt werden und es kommt nicht auf die Reihenfolge an, in der sie genannt werden. So haben wir also $\{1, 1, 2\} = \{2, 1\}$. Die Aussage "x ist Element von X" wird mit $x \in X$ abgekürzt, ihre Verneinung "x ist nicht Element von X" mit $x \notin X$. Es gibt auch die sogenannte **leere Menge** $\emptyset = \{\ \}$, die gar kein Element enthält. Andere Beispiele sind die Menge der **natürlichen Zahlen** $\mathbb{N} = \{0, 1, 2, \ldots\}$, die Menge der **ganzen Zahlen** $\mathbb{Z} = \{0, 1, -1, 2, -2, \ldots\}$ und die Menge der **rationalen Zahlen** $\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z},\ q \neq 0\}$. Deren Name kommt von lateinisch "ratio" für "Verhältnis". Man beachte, daß wir auch hier Elemente mehrfach genannt haben, es gilt ja $p/q = p'/q'$ genau dann, wenn $pq' = p'q$. Auf deutsch bezeichnet man die rationalen Zahlen manchmal auch als **Bruchzahlen**.

Ergänzung 2.1.4. Das Gleichheitszeichen = scheint auf ein 1557 von Robert Recorde publiziertes Buch zurückzugehen und soll andeuten, daß das, was auf der linken und rechten Seite dieses Zeichens steht, so gleich ist wie die beiden Strichlein, die das uns heute so selbstverständliche Gleichheitszeichen bilden. Davor schrieb man statt einem Gleichheitszeichen meist ae für "äquivalent".

2.1.5. In Texten, in deren Konventionen die Null keine natürliche Zahl ist, verwendet man meist die abweichenden Notationen $\mathbb{N}$ für die Menge $\{1, 2, \ldots\}$ und $\mathbb{N}_0$ für die Menge $\{0, 1, 2, \ldots\}$. Die in diesem Text verwendete Notation $\mathbb{N} = \{0, 1, 2, \ldots\}$ stimmt mit der internationalen Norm ISO 31-11 überein.

Definition 2.1.6. Eine Menge Y heißt **Teilmenge** einer Menge X genau dann, wenn jedes Element von Y auch ein Element von X ist. Man schreibt dafür $Y \subset X$ oder $X \supset Y$. Zum Beispiel gilt stets $\emptyset \subset X$, und $\{x\} \subset X$ ist gleichbedeutend zu $x \in X$. Zwei Teilmengen einer gegebenen Menge, die kein gemeinsames Element haben, heißen **disjunkt**.

Bemerkung 2.1.7. Unsere Notation $\subset$ weicht ab von der internationalen Norm ISO 31-11, die statt unserem $\subset$ das Symbol $\subseteq$ vorschlägt. In den Konventionen ISO 31-11 hat das Symbol $\subset$ abweichend die Bedeutung einer **echten**, d.h. von der ganzen Menge verschiedenen Teilmenge, für die wir hinwiederum die Bezeichnung $\subsetneq$ verwenden werden. Meine Motivation für diese Abweichung ist, daß das Symbol für beliebige Teilmengen sehr häufig und das für echte Teilmengen nur sehr selten vorkommt. Die hier verwendete Notation ist auch weit verbreitet und schon sehr viel länger in Gebrauch, das Symbol $\subseteq$ ist eine vergleichsweise neue Konvention. Ich komme muß jedoch zugeben, daß die hier gewählte Notation mit den üblichen und auch in diesem Text verwendeten Notationen $<$ und $\leq$ nicht gut zusammenpaßt.

Definition 2.1.8. Wir vereinbaren, daß wir die leere Menge endlich nennen wollen, damit jede Teilmenge einer endlichen Menge auch wieder endlich ist. Die Zahl der Elemente einer endlichen Menge X nennen wir ihre **Kardinalität** oder **Mächtigkeit** und notieren sie $|X|$ oder $\operatorname{card}(X)$. In der Literatur findet man auch die Notation $\sharp X$. Ist X unendlich, so schreiben wir kurz $|X| = \infty$ und ignorieren in unserer Notation, daß auch unendliche Mengen "verschieden groß" sein können, für ein Beispiel siehe 5.3.6 und für eine genauere Diskussion des Begriffs der Kardinalität **??**. Für endliche Mengen X ist demnach ihre Kardinalität stets eine natürliche Zahl $|X| \in \mathbb{N}$ und $|X| = 0$ ist gleichbedeutend zu $X = \emptyset$.

Definition 2.1.9. Oft bildet man neue Mengen als Teilmengen bestehender Mengen und schreibt $Y = \{x \in X \mid x \text{ hat eine gewisse Eigenschaft}\}$. Zum Beispiel gilt $\mathbb{N} = \{a \in \mathbb{Z} \mid a \geq 0\}$ oder $\{0, 1\} = \{a \in \mathbb{N} \mid a^2 = a\}$.

2.1.10. Bereits an dieser Stelle ist unsere Notation nicht eindeutig: Ich wollte mit $\{0, 1\}$ die zweielementige Menge mit den beiden Elementen Null und Eins andeuten, das könnte jedoch auch als die Menge mit der Dezimalzahl $0, 1$ als einzigem Element interpretiert werden. Es wird noch oft vorkommen, daß sich die Bedeutung einer Formel erst aus dem Kontext erschließt. Im folgenden werden Kommas fast nie als Kommas einer Dezimalzahl zu verstehen sein.

Definition 2.1.11. Es ist auch erlaubt, die "Menge aller Teilmengen" einer gegebenen Menge X zu bilden. Sie heißt die **Potenzmenge** von X und wird mit $\mathcal{P}(X)$ bezeichnet.

2.1.12. Ist X eine endliche Menge, so ist auch ihre Potenzmenge endlich und es gilt $|\mathcal{P}(X)| = 2^{|X|}$. Für die drei-elementige Menge $X = \{1, 2, 3\}$ besteht zum Beispiel $\mathcal{P}(X)$ aus $2^3 = 8$ Elementen, genauer gilt

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

Definition 2.1.13. Gegeben zwei Mengen X, Y können wir auf verschiedene Arten neue Mengen bilden:

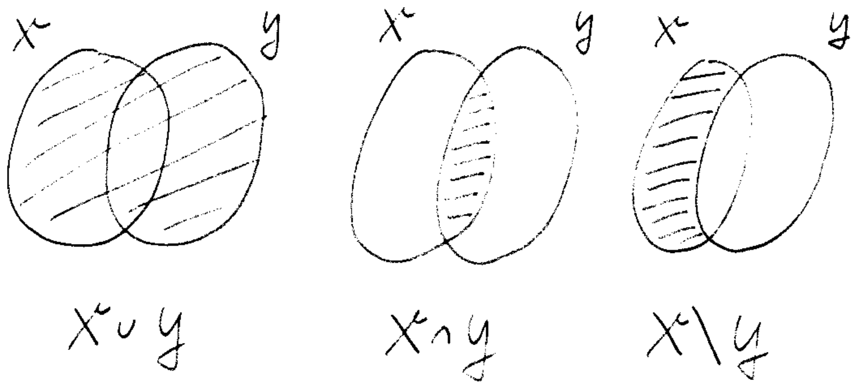
1. Die **Vereinigung** $X \cup Y := \{z \mid z \in X \text{ oder } z \in Y\}$, zum Beispiel ist $\{1, 2\} \cup \{2, 3\} = \{1, 2, 3\}$.

2. Den **Schnitt** $X \cap Y := \{z \mid z \in X \text{ und } z \in Y\}$, zum Beispiel ist $\{1, 2\} \cap \{2, 3\} = \{2\}$. Zwei Mengen sind also disjunkt genau dann, wenn ihr Schnitt die leere Menge ist.

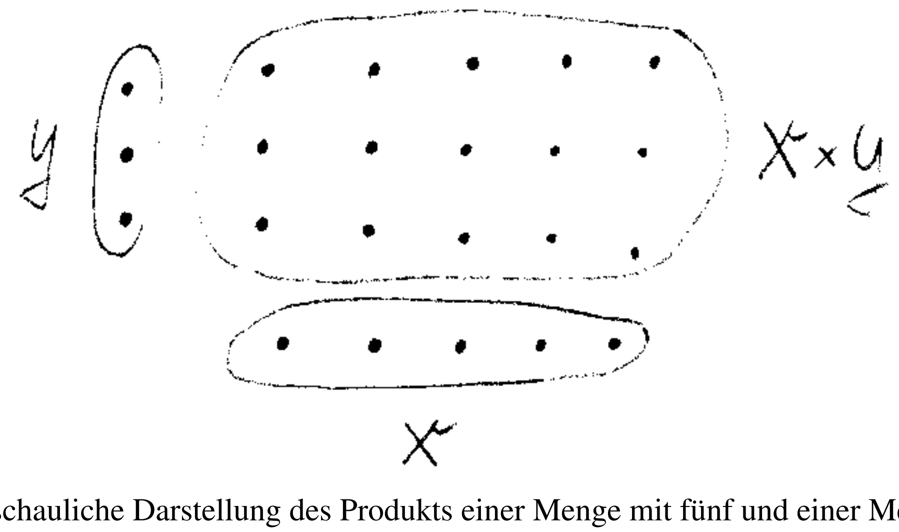
3. Die **Differenz** $X \backslash Y := \{z \in X \mid z \notin Y\}$, zum Beispiel haben wir $\{1, 2\} \backslash \{2, 3\} = \{1\}$. Man schreibt statt $X \backslash Y$ auch $X - Y$. Ist Y eine Teilmenge von X, so heißt $X \backslash Y$ das **Komplement** von Y in X.

4. Das **kartesische Produkt** $X \times Y := \{(x, y) \mid x \in X, \ y \in Y\}$, als da heißt die Menge aller geordneten Paare. Es gilt also $(x, y) = (x', y')$ genau dann, wenn gilt $x = x'$ und $y = y'$. Zum Beispiel haben wir $\{1, 2\} \times \{1, 2, 3\} = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\}$. Oft benutzt man für das kartesische Produkt $X \times X$ einer Menge X mit sich selbst die Abkürzung $X \times X = X^2$.

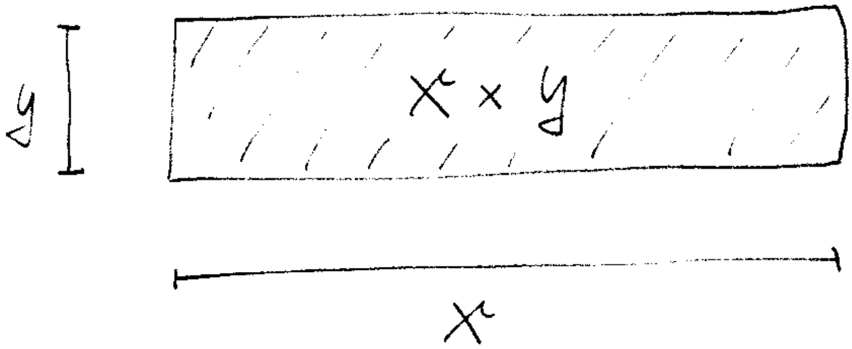
2.1.14. Wir werden in unserer naiven Mengenlehre die ersten drei Operationen nur auf Teilmengen einer gemeinsamen Obermenge anwenden, die uns in der einen oder anderen Weise bereits zur Verfügung steht. Die Potenzmenge und das kartesische Produkt dahingegen benutzen wir, um darüber hinaus neue Mengen zu erschaffen. Diese Konstruktionen erlauben es, im Rahmen der Mengenlehre so etwas wie Abstraktionen zu bilden: Wenn wir uns etwa die Menge T aller an mindestens einem Tag der Weltgeschichte lebenden oder gelebt habenden Tiere als eine Menge im Cantor'schen Sinne denken, so würden wir Konzepte wie "männlich" oder "Hund" oder "Fleischfresser" formal als Teilmengen dieser Menge definieren, d.h. als Elemente von $\mathcal{P}(T)$, und das Konzept "ist Kind von" als eine Teilmenge des kartesischen Produkts dieser Menge T mit sich selbst, also als ein Element von $\mathcal{P}(T \times T)$.



Eine gute Anschauung für die ersten drei Operationen liefern die sogenannten **van-de-Ven-Diagramme** wie sie die obenstehenden Bilder zeigen. Sie sind allerdings nicht zu genau zu hinterfragen, denn ob die Punkte auf einem Blatt Papier im Sinne von Cantor “bestimmte wohlunterschiedene Objekte unserer Anschauung” sind, scheint mir sehr fraglich. Wenn man jedoch jedes der schraffierten Gebiete im Bild auffasst als die Menge aller darin liegenden Kreuzungspunkte auf einem dazugedachten Millimeterpapier und keine dieser Kreuzungspunkte auf den Begrenzungslinien liegen, so können sie wohl schon als eine Menge im Cantor’schen Sinne angesehen werden.



Anschauliche Darstellung des Produkts einer Menge mit fünf und einer Menge mit drei Elementen. Hier wird ein Paar (x, y) dargestellt durch einen fetten Punkt, der über x und neben y liegt.



Dies Bild muß anders interpretiert werden als das Vorherige. Die Mengen X und Y sind nun zu verstehen als die Mengen der Punkte der vertikalen und horizontalen Geradensegmente und ein Punkt des Quadrats meint das Element $(x, y) \in X \times Y$, das in derselben Höhe wie $y \in Y$ senkrecht über $x \in X$ liegt.

2.1.15. Für das Rechnen mit Mengen überlegt man sich die folgenden Regeln:

$$\begin{array}{rcl} X \cap (Y \cap Z) & = & (X \cap Y) \cap Z \\ X \cup (Y \cup Z) & = & (X \cup Y) \cup Z \\[1ex] X \cup (Y \cap Z) & = & (X \cup Y) \cap (X \cup Z) \\ X \cap (Y \cup Z) & = & (X \cap Y) \cup (X \cap Z) \\[1ex] X \backslash (Y \cup Z) & = & (X \backslash Y) \cap (X \backslash Z) \\ X \backslash (Y \cap Z) & = & (X \backslash Y) \cup (X \backslash Z) \\[1ex] X \backslash (X \backslash Y) & = & X \cap Y \end{array}$$

Eine gute Anschauung für diese Regeln liefern die van-de-Ven-Diagramme, wie sie die nebenstehenden Bilder zeigen. Die vorletzte und vorvorletzte Gleichung faßt man auch unter der Bezeichnung **de Morgan'sche Regeln** zusammen.

Übung 2.1.16. Sind X und Y endliche Mengen, so gilt für die Kardinalitäten $|X \times Y| = |X| \cdot |Y|$ und $|X \cup Y| = |X \backslash Y| + |X \cap Y| + |Y \backslash X|$.

Satz 2.1.17 (Bedeutung der Binomialkoeffizienten). *Gegeben $n, k \in \mathbb{N}$ gibt der Binomialkoeffizient $\binom{n}{k}$ die Zahl der k-elementigen Teilmengen einer n-elementigen Menge an, in Formeln:*

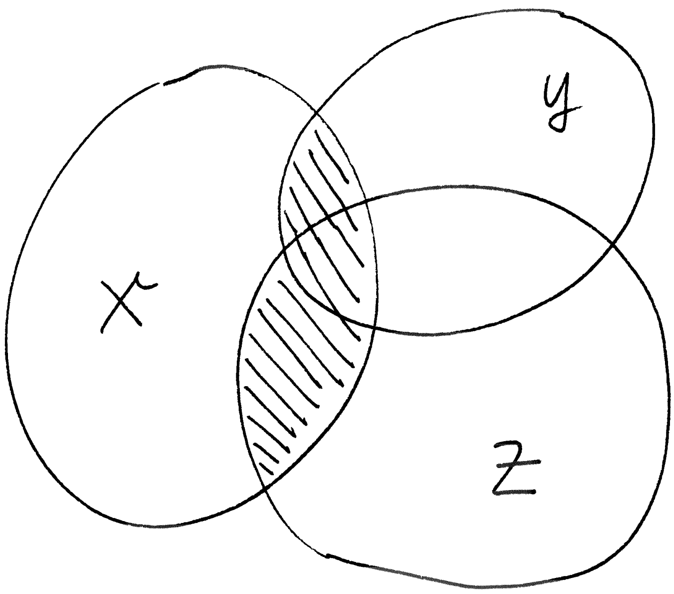
$$|X| = n \quad \textit{impliziert} \quad |\{Y \subset X \mid |Y| = k\}| = \binom{n}{k}$$

Beweis. Vollständige Induktion über n. Für $n = 0$ gilt die Aussage, denn eine nullelementige Menge hat genau eine k-elementige Teilmenge falls $k = 0$ und keine k-elementige Teilmenge falls $k \geq 1$. Nehmen wir nun an, die Aussage sei für ein n schon bewiesen. Eine $(n+1)$-elementige Menge X schreiben wir als $X = M \cup \{x\}$, wo M eine n-elementige Menge ist und $x \notin M$. Ist $k = 0$, so gibt es genau eine k-elementige Teilmenge von $M \cup \{x\}$, nämlich die leere Menge. Ist $k \geq 1$, so gibt es in $M \cup \{x\}$ nach Induktionsannahme genau $\binom{n}{k}$ k-elementige Teilmengen, die x nicht enthalten. Die k-elementigen Teilmengen dahingegen, die x enthalten, ergeben sich durch Hinzunehmen von x aus den $(k-1)$-elementigen Teilmengen von M, von denen es gerade $\binom{n}{k-1}$ gibt. Insgesamt hat $M \cup \{x\}$ damit also genau $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$ k-elementige Teilmengen. □

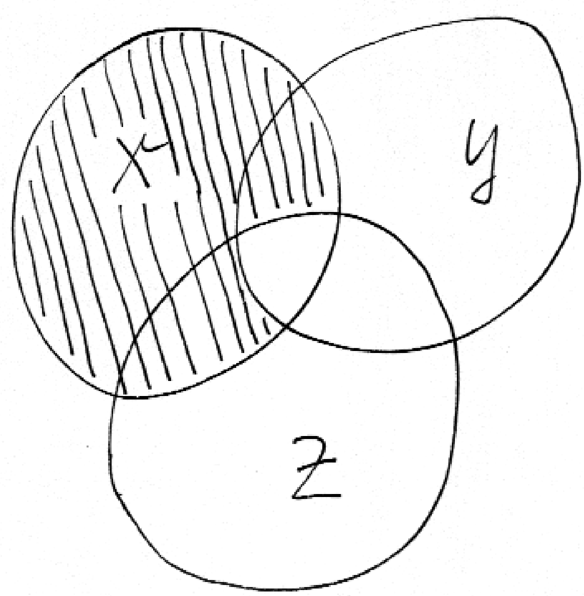
Bemerkung 2.1.18. Wieder scheint mir dieser Beweis in der für vollständige Induktion typischen Weise undurchsichtig. Ich ziehe deshalb den in 1.1.18 gegebenen weniger formellen Beweis vor. Man kann auch diesen Beweis formalisieren und verstehen als Spezialfall der sogenannten "Bahnformel" **??**, vergleiche **??**.

Bemerkung 2.1.19. Wir geben nun die versprochene präzise Formulierung unseres ersten Beweises der binomischen Formel 1.1.22. Wir rechnen dazu

$$(a+b)^n = \sum_{Y \subset \{1,2,\ldots,n\}} a^{|Y|} b^{n-|Y|}$$



$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$$



$$X \setminus (Y \cap Z) = (X \setminus Y) \cup (X \setminus Z)$$

wo die rechte Seite in Verallgemeinerung der in Abschnitt 1.1 eingeführten Notation bedeuten soll, daß wir für jede Teilmenge Y von $\{1, 2, \ldots, n\}$ den angegebenen Ausdruck $a^{|Y|}b^{n-|Y|}$ nehmen und alle diese Ausdrücke aufsummieren. Dann fassen wir gleiche Summanden zusammen und erhalten mit 2.1.17 die binomische Formel.

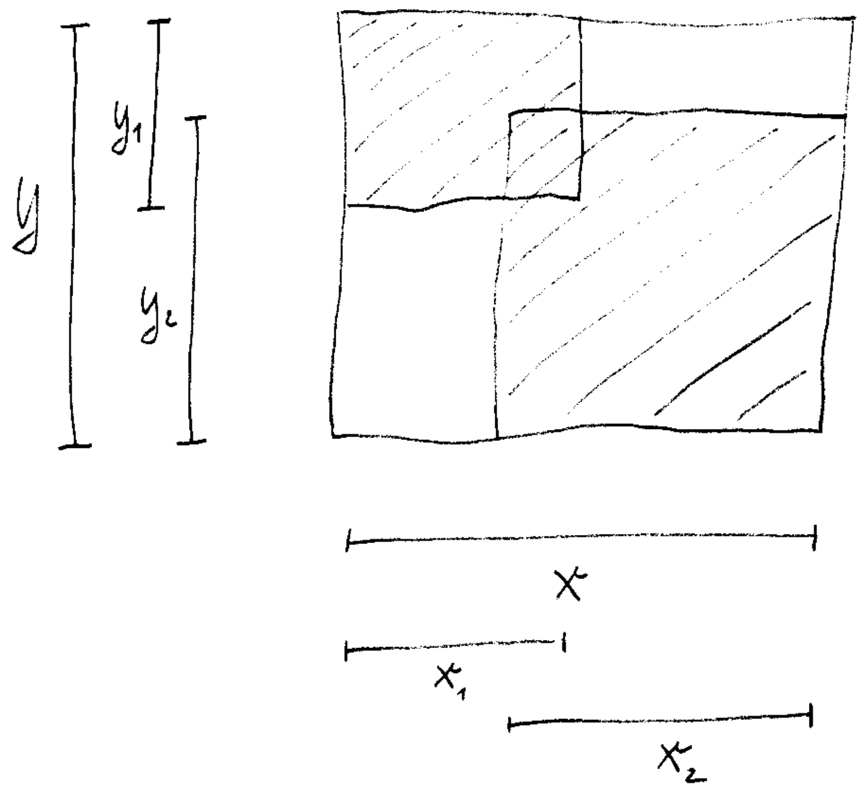
Ergänzende Übung 2.1.20. Es gilt $\sum_k \binom{n}{k} = 2^n$.

Ergänzung 2.1.21. Ich will nicht verschweigen, daß der in diesem Abschnitt dargestellte naive Zugang zur Mengenlehre durchaus begriffliche Schwierigkeiten mit sich bringt: Zum Beispiel darf die Gesamtheit $\mathcal{M}$ aller Mengen nicht als Menge angesehen werden, da wir sonst die "Menge aller Mengen, die sich nicht selbst als Element enthalten", gegeben durch die formelhafte Beschreibung $\mathcal{N} = \{A \in \mathcal{M} \mid A \notin A\}$, bilden könnten. Für diese Menge kann aber weder $\mathcal{N} \in \mathcal{N}$ noch $\mathcal{N} \notin \mathcal{N}$ gelten ... Diese Art von Schwierigkeiten kann erst ein formalerer Zugang klären und auflösen, bei dem man unsere naiven Vorstellungen durch Ketten von Zeichen aus einem wohlbestimmten endlichen Alphabet ersetzt und unsere Vorstellung von Wahrheit durch die Verifizierbarkeit vermittels rein algebraischer "erlaubter Manipulationen" solcher Zeichenketten, die in "Axiomen" festgelegt werden. Diese Verifikationen kann man dann durchaus auch einer Rechenmaschine überlassen, so daß wirklich auf "objektivem" Wege entschieden werden kann, ob ein "Beweis" für die "Richtigkeit" einer unserer Zeichenketten in einem vorgegebenen axiomatischen Rahmen stichhaltig ist. Allerdings kann in derartigen Systemen von einer Zeichenkette algorithmisch nur entschieden werden, ob sie eine "sinnvolle Aussage" ist, nicht aber, ob sie "bewiesen" werden kann. Noch viel stärker zeigt der Unvollständigkeitssatz von Gödel, daß es in einem derartigen axiomatischen Rahmen, sobald er reichhaltig genug ist für eine Beschreibung des Rechnens mit natürlichen Zahlen, stets sinnvolle Aussagen gibt derart, daß entweder sowohl die Aussage als auch ihre Verneinung oder aber weder die Aussage noch ihre Verneinung bewiesen werden können. Mit diesen und ähnlichen Fragestellungen beschäftigt sich die Logik.

2.1.22. Um mich nicht dem Vorwurf auszusetzen, während des Spiels die Spielregeln ändern zu wollen, sei bereits hier erwähnt, daß in **??** noch weitere wichtige Konstruktionen der Mengenlehre eingeführt werden, und daß in **??** einige weniger offensichtliche Folgerungen erläutert werden, die meines Erachtens bereits an den Rand dessen gehen, was man in unserem informellen Rahmen als Argumentation noch vertreten kann.

2.2 Abbildungen

Definition 2.2.1. Seien X, Y Mengen. Eine **Abbildung** $f : X \to Y$ ist eine Zuordnung, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet,



Aus $X = X_1 \cup X_2$ und $Y = Y_1 \cup Y_2$ folgt noch lange nicht
$X \times Y = (X_1 \times Y_1) \cup (X_2 \times Y_2)$

das **Bild** von x unter f, auch genannt der **Wert** von f an der Stelle x. Man spricht dann auch vom **Auswerten** der Funktion f an der Stelle x oder vom **Einsetzen** von x in f.

2.2.2. Wem das zu vage ist, der mag die alternative Definition vorziehen, nach der eine **Abbildung** $f : X \to Y$ eine Teilmenge $f \subset X \times Y$ ist derart, daß es für jedes $x \in X$ genau ein $y \in Y$ gibt mit $(x, y) \in f$. Dies eindeutig bestimmte y schreiben wir dann $f(x)$ und sind auf einem etwas formaleren Weg wieder am selben Punkt angelangt. In unseren Konventionen nennen wir besagte Teilmenge den **Graphen von** f und notieren sie mit dem Symbol Γ (sprich: Gamma), einem großen griechischen G, und schreiben

$$\Gamma(f) := \{(x, f(x)) \mid x \in X\} \subset X \times Y$$

Definition 2.2.3. Ist $f : X \to Y$ eine Abbildung, so nennen wir X ihren **Definitionsbereich** und Y ihren **Wertebereich**. Zwei Abbildungen nennen wir gleich genau dann, wenn sie denselben Definitionsbereich X, denselben Wertebereich Y und dieselbe Abbildungsvorschrift $f \subset X \times Y$ haben. Die Menge aller Abbildungen von X nach Y bezeichnen wir mit

$$\mathrm{Ens}(X, Y)$$

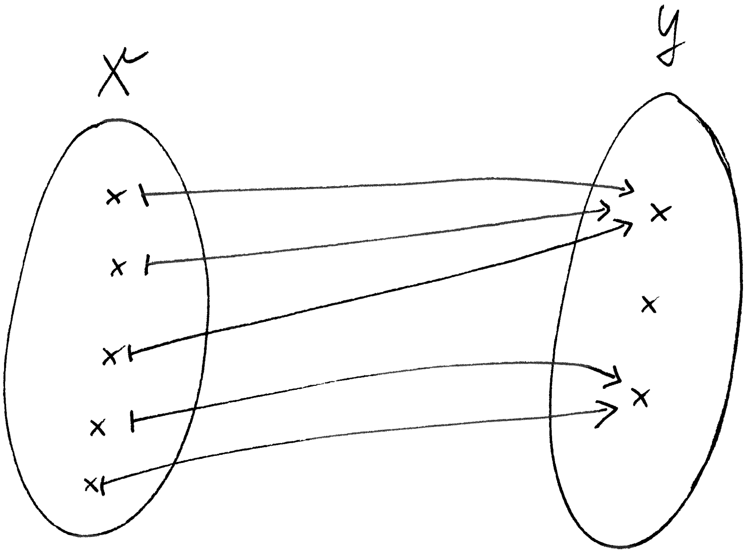
nach der französischen Übersetzung **ensemble** des deutschen Begriffs "Menge". Üblich ist auch die Notation Y^X.

Bemerkung 2.2.4. Noch gebräuchlicher ist die Bezeichnung $\mathrm{Abb}(X, Y)$ für die Menge aller Abbildungen von X nach Y. Ich will jedoch sehr viel später die "Kategorie aller Mengen" mit Ens bezeichnen und für je zwei Objekte X, Y einer Kategorie $\mathcal{C}$ die Menge aller "Morphismen" von X nach Y mit $\mathcal{C}(X, Y)$, und das erklärt dann erst vollständig die hier gewählte Bezeichnung für Mengen von Abbildungen.

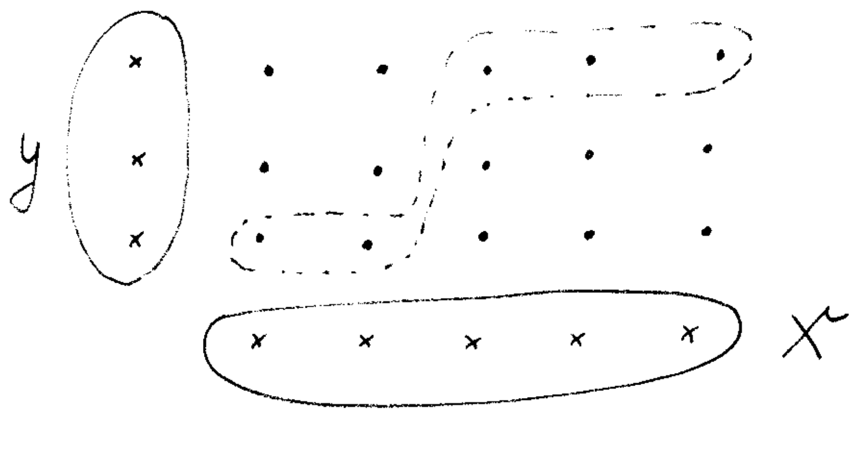
2.2.5. Wir notieren Abbildungen oft in der Form

$$\begin{array}{rccc} f: & X & \to & Y \\ & x & \mapsto & f(x) \end{array}$$

und in verschiedenen Verkürzungen dieser Notation. Zum Beispiel sprechen wir von "einer Abbildung $\mathbb{N} \to \mathbb{N}$ von der Menge der natürlichen Zahlen in sich selber" oder "der Abbildung $n \mapsto n^3$ von der Menge der natürlichen Zahlen in sich selber". Wir benutzen unsere zwei Arten von Pfeilen auch im allgemeinen in derselben Weise.



Eine Abbildung einer Menge mit fünf in eine mit drei Elementen



Der Graph der oben angegebenen Abbildung, wobei das X oben mit dem X hier identifiziert wurde durch "Umkippen nach Rechts"

Beispiel 2.2.6. Für jede Menge X haben wir die **identische Abbildung** oder **Identität**

$$\begin{array}{rccc} \mathrm{id} = \mathrm{id}_X : & X & \to & X \\ & x & \mapsto & x \end{array}$$

Ein konkreteres Beispiel für eine Abbildung ist das Quadrieren

$$\begin{array}{rccc} q : & \mathbb{Z} & \to & \mathbb{Z} \\ & n & \mapsto & n^2 \end{array}$$

Beispiel 2.2.7. Gegeben zwei Mengen X, Y erklärt man die sogenannten **Projektionsabbildungen** oder **Projektionen** $\mathrm{pr}_X : X \times Y \to X$ bzw. $\mathrm{pr}_Y : X \times Y \to Y$ durch die Vorschrift $(x, y) \mapsto x$ bzw. $(x, y) \mapsto y$. In manchen Zusammenhängen notiert man sie auch abweichend pr_1 und pr_2 für die "Projektion auf die erste bzw. zweite Komponente".

Definition 2.2.8. Ist $f : X \to Y$ eine Abbildung, so definieren wir ihr **Bild** oder genauer ihre **Bildmenge**, eine Teilmenge $\operatorname{im} f \subset Y$, durch

$$\operatorname{im} f := \{y \in Y \mid \text{Es gibt } x \in X \text{ mit } f(x) = y\}$$

für französisch und englisch **image**. Eine Abbildung, deren Bild aus höchstens einem Element besteht, nennen wir eine **konstante Abbildung**. Eine Abbildung, deren Bild aus genau einem Element besteht, nennen wir eine **einwertige Abbildung**. In anderen Worten ist eine einwertige Abbildung also eine konstante Abbildung mit nichtleerem Definitionsbereich.

Definition 2.2.9. Ist $f : X \to Y$ eine Abbildung und $A \subset X$ eine Teilmenge, so definieren wir das **Bild von A unter f**, eine Teilmenge $f(A) \subset Y$, durch

$$f(A) := \{y \in Y \mid \text{Es gibt } x \in A \text{ mit } f(x) = y\}$$

Beispiel 2.2.10. Per definitionem haben wir für eine Abbildung $f : X \to Y$ stets $f(X) = \operatorname{im} f$. Für unsere Abbildung $q : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n^2$ von eben könnten wir die Menge aller Quadratzahlen schreiben als

$$q(\mathbb{Z}) = \{a^2 \mid a \in \mathbb{Z}\}$$

Ebenso wäre $\{2a \mid a \in \mathbb{N}\}$ eine mögliche formelmäßige Darstellung der Menge aller geraden natürlichen Zahlen, und $\{ab \mid a, b \in \mathbb{N}, a \geq 2, b \geq 2\}$ wäre eine formelmäßige Darstellung der Menge aller natürlichen Zahlen, die nicht prim und auch nicht Null oder Eins sind.

2.2.11. Gegeben ein festes $c \in Y$ schreiben wir oft auch kurz c für die konstante Abbildung $X \to Y$, $x \mapsto c$ für alle $x \in X$. Damit verbunden ist die Hoffnung, daß aus dem Kontext klar wird, ob im Einzelfall die Abbildung $c : X \to Y$ oder das Element $c \in Y$ gemeint sind.

Definition 2.2.12. Ist $f : X \to Y$ eine Abbildung und $B \subset Y$ eine Teilmenge, so definieren wir ihr **Urbild**, eine Teilmenge von $f^{-1}(B) \subset X$, durch

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}$$

Besteht B nur aus einem Element x, so schreiben wir auch $f^{-1}(x)$ statt $f^{-1}(\{x\})$ und nennen diese Menge die **Faser von** f **über** x. Das Quadrieren q aus 2.2.10 hat etwa die Fasern $q^{-1}(1) = \{1, -1\}$ und $q^{-1}(-1) = \emptyset$.

Definition 2.2.13. Sind schließlich drei Mengen X, Y, Z gegeben und Abbildungen $f : X \to Y$ und $g : Y \to Z$, so definieren wir eine Abbildung $g \circ f : X \to Z$, die **Verknüpfung** der Abbildungen f und g, durch die Vorschrift

$$\begin{array}{rccl} g \circ f : & X & \to & Z \\ & x & \mapsto & g(f(x)) \end{array}$$

2.2.14. Die Notation $g \circ f$, sprich "g nach f", für "erst f, dann g" ist gewöhnungsbedürftig, erklärt sich aber durch die Formel $(g \circ f)(x) = g(f(x))$.

Beispiel 2.2.15. Betrachten wir zusätzlich zum Quadrieren $q : \mathbb{Z} \to \mathbb{Z}$ die Abbildung $t : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x+1$, so gilt $(q \circ t)(x) = (x+1)^2$ aber $(t \circ q)(x) = x^2 + 1$. Natürlich gilt $(g \circ f)(A) = g(f(A))$ für jede Teilmenge $A \subset X$ und umgekehrt auch $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$ für jede Teilmenge $C \subset Z$.

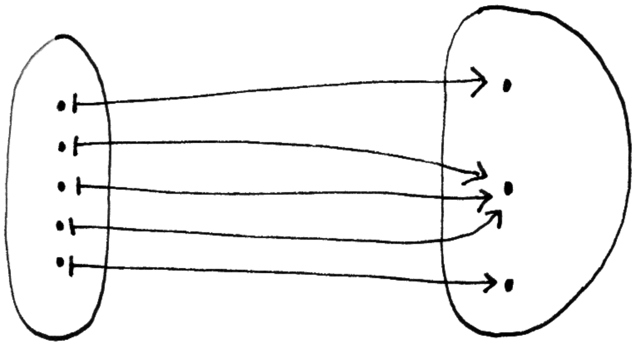
Ergänzende Übung 2.2.16. Sei X eine Menge und $f : X \to X$ eine Abbildung. Bezeichne $f^n = f \circ f \circ \ldots \circ f : X \to X$ das "n-malige Ausführen von f". Im Extremfall $n = 0$ verstehen wir $f^0 = \mathrm{id}$. Sei nun $C \subset X$ eine Teilmenge. Man zeige: Stimmen für ein $n \in \mathbb{N}$ die Bildmengen $f^n(C)$ und $f^{n+1}(C)$ überein, so gilt bereits $f^n(C) = f^{n+1}(C) = f^{n+2}(C) \ldots$ Stimmen für ein $n \in \mathbb{N}$ die Urbildmengen $(f^n)^{-1}(C)$ und $(f^{n+1})^{-1}(C)$ überein, so gilt bereits $(f^n)^{-1}(C) = (f^{n+1})^{-1}(C) = (f^{n+2})^{-1}(C) \ldots$

Definition 2.2.17. Sei $f : X \to Y$ eine Abbildung.

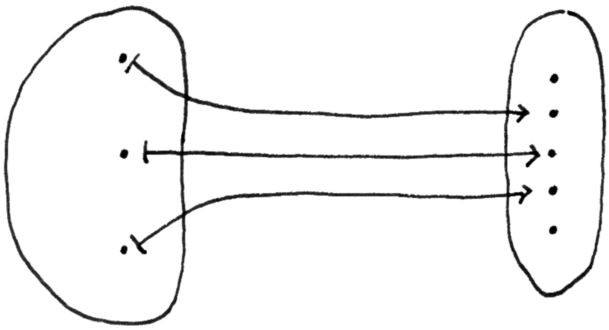
1. f heißt **injektiv** oder eine **Injektion** genau dann, wenn aus $x \neq x'$ folgt $f(x) \neq f(x')$. Gleichbedeutend ist die Forderung, daß es für jedes $y \in Y$ höchstens ein $x \in X$ gibt mit $f(x) = y$. Injektionen schreibt man oft $\hookrightarrow$.

2. f heißt **surjektiv** oder eine **Surjektion** genau dann, wenn es für jedes $y \in Y$ mindestens ein $x \in X$ gibt mit $f(x) = y$. Surjektionen schreibt man manchmal $\twoheadrightarrow$.

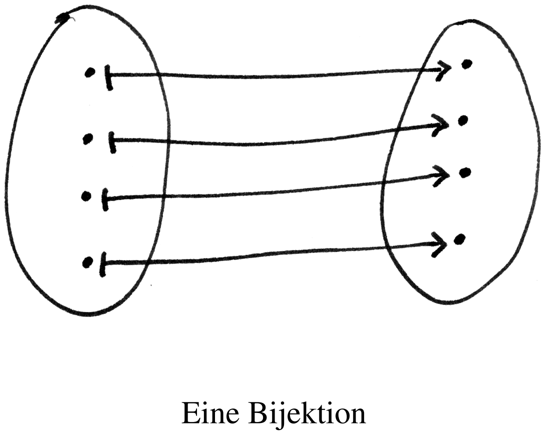
3. f heißt **bijektiv** oder eine **Bijektion** genau dann, wenn f injektiv und surjektiv ist. Gleichbedeutend ist die Forderung, daß es für jedes $y \in Y$ genau ein $x \in X$ gibt mit $f(x) = y$. Bijektionen schreibt man oft $\xrightarrow{\sim}$.



Eine Surjektion



Eine Injektion



Eine Bijektion

2.2.18. Ist $X \subset Y$ eine Teilmenge, so ist die **Einbettung** oder **Inklusion** $i : X \to Y$, $x \mapsto x$ stets injektiv. Ist $g : Y \to Z$ eine Abbildung und $X \subset Y$ eine Teilmenge, so nennen wir die Verknüpfung $g \circ i$ von g mit der Inklusion auch die **Einschränkung** von g auf X und notieren sie $g \circ i =: g|X = g|_X : X \to Z$. Oft bezeichnen wir eine Einschränkung aber auch einfach mit demselben Buchstaben g in der Hoffnung, daß der Leser aus dem Kontext erschließen kann, welche Abbildung genau gemeint ist.

2.2.19. Ist $f : X \to Y$ eine Abbildung, so ist die Abbildung $f : X \to f(X)$, $x \mapsto f(x)$ stets surjektiv. Der Leser möge entschuldigen, daß wir hier zwei verschiedene Abbildungen mit demselben Symbol f bezeichnet haben. Das wird noch öfter vorkommen. Überhaupt ignorieren wir, gegeben Mengen X, Y und eine Teilmenge $Z \subset Y$, im folgenden meist den Unterschied zwischen "Abbildungen von X nach Y, deren Bild in Z enthalten ist" und "Abbildungen von X nach Z".

Ergänzung 2.2.20. Eine Abbildung $f : X \to \mathcal{P}(X)$ von einer Menge in ihre Potenzmenge kann nie surjektiv sein. In der Tat, betrachten wir in X die Teilmenge $A = \{x \in X \mid x \notin f(x)\}$, so kann es kein $y \in X$ geben mit $f(y) = A$, denn für solch ein y hätten wir entweder $y \in A$ oder $y \notin A$, und aus $y \in A$ alias $y \in f(y)$ folgte $y \notin A$, wohingegen aus $y \notin A$ alias $y \notin f(y)$ folgte $y \in A$.

Beispiele 2.2.21. Unsere Abbildung $q : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n^2$ ist weder injektiv noch surjektiv. Die Identität $\mathrm{id} : X \to X$ ist stets bijektiv. Sind X und Y endliche Mengen, so gibt es genau dann eine Bijektion von X nach Y, wenn X und Y dieselbe Kardinalität haben, in Formeln $|X| = |Y|$.

Satz 2.2.22. *Seien $f, f_1 : X \to Y$ und $g, g_1 : Y \to Z$ Abbildungen.*

1. *Ist $g \circ f$ injektiv, so ist f injektiv.*
2. *Sind g und f injektiv, so auch $g \circ f$.*
3. *Genau dann ist g injektiv, wenn aus $g \circ f = g \circ f_1$ schon folgt $f = f_1$.*

Beweis. Übung. Besonders elegant ist es, zunächst die letzte Aussage zu zeigen, und dann die vorderen Aussagen ohne weitere Betrachtung von Elementen zu folgern. □

Satz 2.2.23. *Seien $f, f_1 : X \to Y$ und $g, g_1 : Y \to Z$ Abbildungen.*

1. *Ist $g \circ f$ surjektiv, so ist g surjektiv.*
2. *Sind g und f surjektiv, so auch $g \circ f$.*
3. *Genau dann ist f surjektiv, wenn aus $g \circ f = g_1 \circ f$ schon folgt $g = g_1$.*

Beweis. Übung. Besonders elegant ist es, zunächst die letzte Aussage zu zeigen, und dann die vorderen Aussagen ohne weitere Betrachtung von Elementen zu folgern. □

2.2.24. Ist $f : X \to Y$ eine bijektive Abbildung, so ist offensichtlich die Menge $\{(f(x), x) \in Y \times X \mid x \in X\}$ im Sinne von 2.2.2 eine Abbildung oder, vielleicht klarer, der Graph einer Abbildung $Y \to X$. Diese Abbildung in die Gegenrichtung heißt die **Umkehrabbildung** oder **Umkehrfunktion** auch die **inverse Abbildung** zu f und wird mit $f^{-1} : Y \to X$ bezeichnet. Offensichtlich ist mit f auch f^{-1} eine Bijektion.

Beispiel 2.2.25. Die Umkehrabbildung unserer Bijektion $t : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x + 1$ ist die Abbildung $\mathbb{Z} \to \mathbb{Z}$, $x \mapsto x - 1$.

Übung 2.2.26. Gegeben eine Bijektion $f : X \to Y$ ist $g = f^{-1}$ die einzige Abbildung $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$. Ebenso ist auch $h = f^{-1}$ die einzige Abbildung $h : Y \to X$ mit $h \circ f = \mathrm{id}_X$.

2.2.27. Gegeben drei Mengen X, Y, Z haben wir eine offensichtliche Bijektion

$$\mathrm{Ens}(X \times Y, Z) \xrightarrow{\sim} \mathrm{Ens}(X, \mathrm{Ens}(Y, Z))$$

Etwas vage formuliert ist also eine Abbildung $X \times Y \to Z$ von einem kartesischen Produkt $X \times Y$ in eine weitere Menge Z dasselbe wie eine Abbildung, die jedem $x \in X$ eine Abbildung $Y \to Z$ zuordnet, und symmetrisch natürlich auch dasselbe wie eine Abbildung, die jedem $y \in Y$ eine Abbildung $X \to Z$ zuordnet. In der exponentiellen Notation liest sich das ganz suggestiv als kanonische Bijektion $Z^{(X \times Y)} \xrightarrow{\sim} (Z^X)^Y$. In diesem Sinne sind also die in der Schule derzeit so beliebten “Funktionen mit Parameter” nichts anderes als “Funktionen von zwei Variablen”.

Satz 2.2.28 (Bedeutung der Fakultät). *Sind X und Y zwei Mengen mit je n Elementen, so gibt es genau $n!$ bijektive Abbildungen $f : X \to Y$.*

Beweis. Sei $X = \{x_1, \ldots, x_n\}$. Wir haben n Möglichkeiten, ein Bild für x_1 auszusuchen, dann noch $(n - 1)$ Möglichkeiten, ein Bild für x_2 auszusuchen, und so weiter, bis schließlich nur noch 1 Element von Y als mögliches Bild von x_n in Frage kommt. Insgesamt gibt es also $n(n-1) \cdots 1 = n!$ Möglichkeiten für f. Da wir $0! = 1$ vereinbart hatten, stimmt unser Satz auch für $n = 0$. □

Ergänzende Übung 2.2.29. Seien X, Y endliche Mengen. So gibt es genau $|Y|^{|X|}$ Abbildungen von X nach Y, und unter diesen Abbildungen sind genau $|Y|(|Y| - 1)(|Y| - 2) \ldots (|Y| - |X| + 1)$ Injektionen.

Ergänzende Übung 2.2.30. Sei X eine Menge mit n Elementen und seien natürliche Zahlen $\alpha_1, \ldots, \alpha_r \in \mathbb{N}$ gegeben mit $n = \alpha_1 + \ldots + \alpha_r$. Man zeige: Es gibt genau $n!/(\alpha_1! \cdots \alpha_r!)$ Abbildungen $f : X \to \{1, \ldots, r\}$, die jedes i genau α_i-mal als Wert annehmen, in Formeln

$$\frac{n!}{\alpha_1! \cdots \alpha_r!} = \operatorname{card}\{f \mid |f^{-1}(i)| = \alpha_i \text{ für } i = 1, \ldots r\}$$

Ergänzung 2.2.31. Manche Autoren bezeichnen die Zahlen aus der vorherigen Übung 2.2.30 auch als **Multinomialkoeffizienten** und verwenden die Notation

$$\frac{n!}{\alpha_1! \cdots \alpha_r!} =: \binom{n}{\alpha_1; \ldots; \alpha_r}$$

Mich überzeugt diese Notation nicht, da sie im Gegensatz zu unserer Notation für die Binomialkoeffizienten recht eigentlich nichts kürzer macht.

Ergänzende Übung 2.2.32. Man zeige die Formel

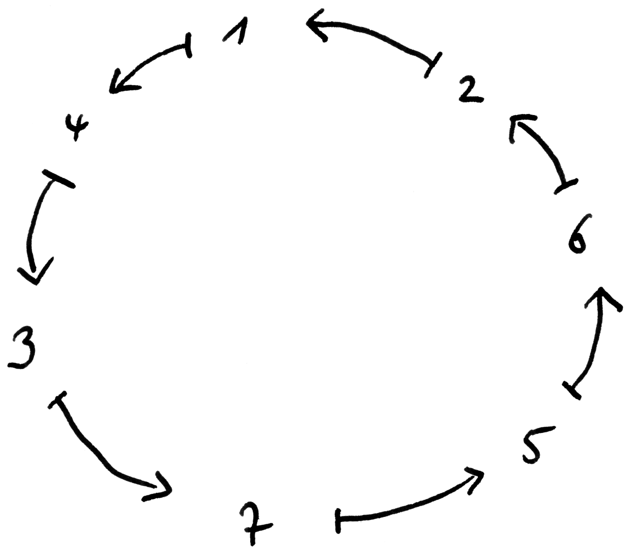
$$(x_1 + \ldots + x_r)^n = \sum_{\alpha_1 + \ldots + \alpha_r = n} \frac{n!}{\alpha_1! \cdots \alpha_r!} x_1^{\alpha_1} \cdots x_r^{\alpha_r}$$

Hier ist zu verstehen, daß wir für alle $\alpha_1, \ldots, \alpha_r \in \mathbb{N}$ mit $\alpha_1 + \ldots + \alpha_r = n$ den angegebenen Ausdruck nehmen und alle diese Ausdrücke aufsummieren.

Ergänzende Übung 2.2.33. Eine **zyklische Anordnung** einer endlichen Menge M ist eine Abbildung $z : M \to M$ derart, daß wir durch mehrmaliges Anwenden von z auf ein beliebiges Element $x \in M$ jedes Element $y \in M$ erhalten können. Man zeige, daß es auf einer n-elementigen Menge mit $n \geq 1$ genau $(n-1)!$ zyklische Anordnungen gibt. Die Terminologie "zyklische Anordnung" macht mich nicht besonders glücklich, da unsere Struktur nun beim besten Willen keine Anordnung im Sinne von 4.2 ist. Andererseits ist aber das Angeben einer Anordnung auf einer endlichen Menge M schon auch etwas Ähnliches.

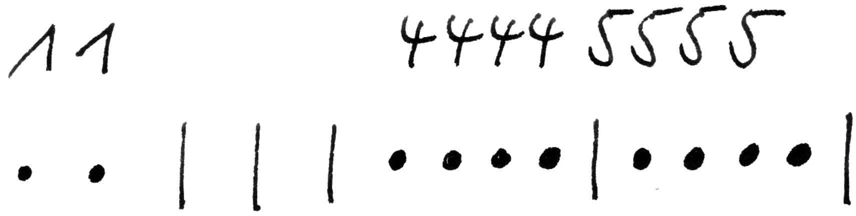
Ergänzende Übung 2.2.34. Sei X eine Menge mit $n \geq 1$ Elementen und sei m eine natürliche Zahl. Man zeige, daß es genau $\binom{n+m-1}{n-1}$ Abbildungen $f : X \to \mathbb{N}$ gibt mit $\sum_{x \in X} f(x) = m$. Hinweis: Man denke sich $X = \{1, 2, \ldots, n\}$ und veranschauliche sich dann f als eine Folge auf $f(1)$ Punkten gefolgt von einem Strich gefolgt von $f(2)$ Punkten gefolgt von einem Strich und so weiter, insgesamt also eine Folge aus $n + m - 1$ Symbolen, davon m Punkten und $n - 1$ Strichen.

Ergänzung 2.2.35. Gegeben eine Menge X mag man sich eine Abbildung $X \to \mathbb{N}$ veranschaulichen als eine "Menge von Elementen von X, in der jedes Element mit einer wohlbestimmten Vielfachheit vorkommt". Aufgrund dieser Vorstellung nennt man eine Abbildung $X \to \mathbb{N}$ auch eine **Multimenge** von Elementen von



Versuch der graphischen Darstellung einer zyklischen Anordnung auf der Menge $\{1, 2, \ldots, 7\}$. Die Pfeile $\mapsto$ sollen jeweils den Effekt der Abbildung z veranschaulichen.

x	1	2	3	4	5	6
$f(x)$	2	0	0	4	4	0



Eine Abbildung $f : \{1, 2, \ldots, n\} \to \mathbb{N}$ im Fall $n = 6$ mit Wertesumme $m = 10$ und die Veranschaulichung nach der Vorschrift aus Übung 2.2.34 als Folge bestehend aus m Punkten und $n - 1$ Strichen.

X. Unter der Kardinalität einer Multimenge verstehen wir die Summe über alle Werte der entsprechenden Abbildung, aufgefaßt als ein Element von $\mathbb{N} \sqcup \{\infty\}$. Ich notiere eine Multimenge mit verdoppelten Mengenklammern, so wäre etwa $\{\{5, 5, 5, 7, 7, 1\}\}$ die hoffentlich offensichtliche Multimenge von natürlichen Zahlen der Kardinalität 6.

2.3 Logische Symbole und Konventionen

2.3.1. In der Mathematik meint **oder** immer, daß auch beides erlaubt ist. Wir haben diese Konvention schon benutzt bei der Definition der Vereinigung wenn wir schreiben $X \cup Y = \{z \mid z \in X \text{ oder } z \in Y\}$, zum Beispiel haben wir $\{1, 2\} \cup \{2, 3\} = \{1, 2, 3\}$.

2.3.2. Sagt man der Mathematik, es gebe ein Objekt mit diesen und jenen Eigenschaften, so ist stets gemeint, daß es *mindestens ein* derartiges Objekt geben soll. Hätten wir diese Sprachregelung rechtzeitig vereinbart, so hätten wir zum Beispiel das Wörtchen “mindestens” in Teil 2 von 2.2.17 bereits weglassen können. Sagt ihnen also ein Mathematiker, er habe einen Bruder, so kann es auch durchaus sein, daß er noch weitere Brüder hat! Will man in der Mathematik Existenz und Eindeutigkeit gleichzeitig ausdrücken, so sagt man, es gebe **genau ein** Objekt mit diesen und jenen Eigenschaften. Sagt ihnen also ein Mathematiker, er habe genau einen Bruder, so können sie sicher sein, daß er nicht noch weitere Brüder hat.

2.3.3. Die folgenden Abkürzungen erweisen sich als bequem und werden recht häufig verwendet:

$\forall$	für alle (ein umgedrehtes A wie “alle”)
$\exists$	es gibt (ein umgedrehtes E wie “existiert”)
$\exists!$	es gibt genau ein
$\ldots \Rightarrow \cdots$	aus … folgt $\cdots$
$\ldots \Leftarrow \cdots$	… folgt aus $\cdots$
$\ldots \Leftrightarrow \cdots$	… ist gleichbedeutend zu $\cdots$

Ist zum Beispiel $f : X \to Y$ eine Abbildung, so können wir unsere Definitionen injektiv, surjektiv, und bijektiv etwas formaler so schreiben:

$$\begin{array}{lcl} f \text{ injektiv} & \Leftrightarrow & ((f(x) = f(z)) \Rightarrow (x = z)) \\ f \text{ surjektiv} & \Leftrightarrow & \forall y \in Y \exists x \in X \text{ mit } f(x) = y \\ f \text{ bijektiv} & \Leftrightarrow & \forall y \in Y \exists! x \in X \text{ mit } f(x) = y \end{array}$$

Ergänzung 2.3.4. In den Zeiten des Bleisatzes war es nicht einfach, neue Symbole in Zeitschriften gedruckt zu kriegen. Irgendwelche Buchstaben verdreht zu setzen, war aber unproblematisch. So entstanden die Symbole $\forall$ und $\exists$.

2.3.5. Bei den “für alle” und “es gibt” kommt es in der Mathematik, anders als in der weniger präzisen Umgangssprache, entscheidend auf die Reihenfolge an. Man betrachte zum Beispiel die beiden folgenden Aussagen:

> “Für alle $n \in \mathbb{N}$ gibt es $m \in \mathbb{N}$ so daß gilt $m \geq n$”
>
> “Es gibt $m \in \mathbb{N}$ so daß für alle $n \in \mathbb{N}$ gilt $m \geq n$”

Offensichtlich ist die Erste richtig, die Zweite aber falsch. Weiter mache man sich klar, daß die “für alle” und “es gibt” bei Verneinung vertauscht werden. Äquivalent sind zum Beispiel die beiden folgenden Aussagen

> “Es gibt kein $n \in \mathbb{N}$ mit $n^2 = 2$”
>
> “Für alle $n \in \mathbb{N}$ gilt nicht $n^2 = 2$”

2.3.6. Wollen wir zeigen, daß aus einer Aussage A eine andere Aussage B folgt, so können wir ebensogut zeigen: Gilt B nicht, so gilt auch A nicht. In formelhafter Schreibweise haben wir also

$$(A \Rightarrow B) \Leftrightarrow ((\text{nicht } B) \Rightarrow (\text{nicht } A))$$

Wollen wir zum Beispiel zeigen $(g \circ f \text{ surjektiv}) \Rightarrow (g \text{ surjektiv})$, so reicht es, wenn wir uns überlegen: Ist g nicht surjektiv, so ist $g \circ f$ erst recht nicht surjektiv.

2.3.7. In der Literatur findet man oft die Abkürzung **oBdA** für “ohne Beschränkung der Allgemeinheit”.

3 Algebraische Grundbegriffe

Auf der Schule versteht man unter einer “reellen Zahl” meist einen unendlichen Dezimalbruch, wobei man noch aufpassen muß, daß durchaus verschiedene unendliche Dezimalbrüche dieselbe reelle Zahl darstellen können, zum Beispiel gilt in den reellen Zahlen ja

$$0,99999\ldots = 1,00000\ldots$$

Diese reellen Zahlen werden dann addiert, subtrahiert, multipliziert und dividiert ohne tiefes Nachdenken darüber, wie man denn zum Beispiel mit den eventuell unendlich vielen Überträgen bei der Addition und Subtraktion umgehen soll, und warum dann Formeln wie $(a+b)-c = a+(b-c)$ wirklich gelten, zum Beispiel für $a = b = c = 0,999\ldots$. Dieses tiefe Nachdenken wollen wir im Folgenden vom Rest der Vorlesung abkoppeln und müssen dazu sehr präzise formulieren, welche Eigenschaften für die Addition, Multiplikation und Anordnung in “unseren” reellen Zahlen gelten sollen: In der Terminologie, die in den folgenden Abschnitten eingeführt wird, werden wir die reellen Zahlen charakterisieren als einen angeordneten Körper, in dem jede nichtleere Teilmenge mit einer unteren Schranke sogar eine größte untere Schranke besitzt. Von dieser Charakterisierung ausgehend erklären wir dann, welche reelle Zahl ein gegebener unendlicher Dezimalbruch darstellt, und errichten das Gebäude der Analysis. In demselben Begriffsgebäude modellieren wir auch den Anschauungsraum, vergleiche 1.2.8 oder besser **??** und **??**. Um diese Charakterisierungen und Modellierungen verständlich zu machen, führen wir zunächst einige grundlegende algebraische Konzepte ein, die Ihnen im weiteren Studium der Mathematik noch oft begegnen werden.

3.1 Mengen mit Verknüpfung

Definition 3.1.1. Eine **Verknüpfung** $\top$ **auf einer Menge** A ist eine Abbildung

$$\begin{array}{rccc} \top: & A \times A & \to & A \\ & (a,b) & \mapsto & a \top b \end{array}$$

die jedem geordneten Paar (a, b) mit $a, b \in A$ ein weiteres Element $(a \top b) \in A$ zuordnet.

Beispiele 3.1.2. 1. Die Addition von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} +: & \mathbb{Z} \times \mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a+b \end{array}$$

	0	1	2	3	4
0	0	0	0	0	0
1	0	1	1	1	1
2	0	1	2	2	2
3	0	1	2	3	3
4	0	1	2	3	4

Man kann Verknüpfungen auf endlichen Mengen darstellen durch ihre **Verknüpfungstafel**. Hier habe ich etwa die Verknüpfungstafel der Verknüpfung $\min$ auf der Menge $\{0, 1, 2, 3, 4\}$ angegeben. Eigentlich muß man sich dazu einigen, ob im Kästchen aus der Spalte a und der Zeile b nun $a \top b$ oder vielmehr $b \top a$ stehen soll, aber bei einer kommutativen Verknüpfung wie $\min$ kommt es darauf zum Glück nicht an.

und	Wahr	Falsch
Wahr	Wahr	Falsch
Falsch	Falsch	Falsch

oder	Wahr	Falsch
Wahr	Wahr	Wahr
Falsch	Wahr	Falsch

Die Wahrheitstafeln für “und” und “oder”. Gemeint ist hier wie stets in der Mathematik das “nichtausschließende oder”. Sagen wir, es gelte A oder B, so ist insbesondere auch erlaubt, daß beides gilt. Bei der Wahrheitstafel für das “ausschließende oder” müßte oben links als Verknüpfung von “Wahr” mit “Wahr” ein “Falsch” stehen.

2. Die Multiplikation von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} \cdot : & \mathbb{Z}\times\mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a\cdot b \end{array}$$

3. Die Zuordnung $\min$, die jedem Paar von natürlichen Zahlen die kleinere zuordnet (wenn sie verschieden sind, man setzt sonst $\min(a,a) = a$), ist eine Verknüpfung

$$\begin{array}{rccc} \min : & \mathbb{N}\times\mathbb{N} & \to & \mathbb{N} \\ & (a,b) & \mapsto & \min(a,b) \end{array}$$

4. Sei X eine Menge. Wir kürzen die Menge $\mathrm{Ens}(X,X)$ aller Abbildungen von X in sich selber oft mit $\mathrm{Ens}(X,X) =: \mathrm{Ens}(X)$ ab. Die Verknüpfung von Abbildungen liefert eine Verknüpfung auf der Menge $\mathrm{Ens}(X)$ aller Abbildungen von X in sich selber

$$\begin{array}{rccc} \circ : & \mathrm{Ens}(X)\times\mathrm{Ens}(X) & \to & \mathrm{Ens}(X) \\ & (f,g) & \mapsto & f\circ g \end{array}$$

5. Die Subtraktion von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} - : & \mathbb{Z}\times\mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a-b \end{array}$$

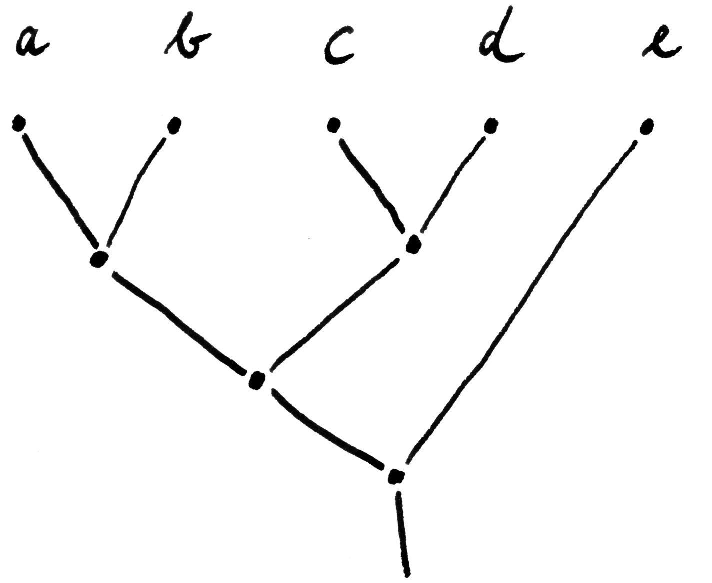
6. Jede Verknüpfung $\top$ auf einer Menge induziert eine Verknüpfung auf ihrer Potenzmenge vermittels der Vorschrift

$$U\top V = \{u\top v \mid u\in U,\ v\in V\}$$

7. Gegeben Mengen mit Verknüpfung $(A,\top)$ und $(B,\perp)$ erhalten wir die **komponentenweise Verknüpfung** auf ihrem Produkt $A\times B$ vermittels der Vorschrift $((a,b),(a',b')) \mapsto ((a\top a'),(b\perp b'))$.

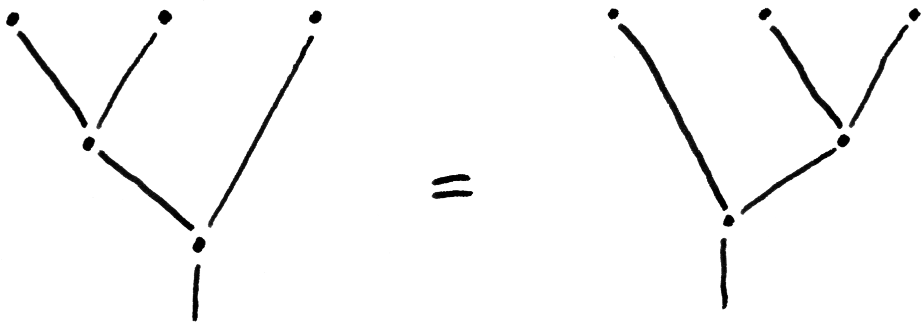
8. Die logischen Operationen "und", "oder", "impliziert" und dergleichen mehr können als Verknüpfungen auf der zweielementigen Menge $\{\text{Wahr}, \text{Falsch}\}$ aufgefaßt werden. Die zugehörigen Verknüpfungstabellen heißen **Wahrheitstafeln**. Bei einem formalen Zugang werden diese Tafeln, wie sie für "und" und "oder" auf der vorhergehenden Seite zu finden sind, sogar zur Definition der jeweiligen Begriffe.

Definition 3.1.3. Eine Verknüpfung $\top$ auf einer Menge A heißt **assoziativ** genau dann, wenn gilt $a\top(b\top c) = (a\top b)\top c \quad \forall a,b,c\in A$. Sie heißt **kommutativ** genau dann, wenn gilt $a\top b = b\top a \quad \forall a,b\in A$.



$$((a \top b) \top (c \top d)) \top e$$

Mögliche “Klammerungen” mag man sich graphisch wie oben angedeutet veranschaulichen. Die Assoziativität bedeutet dann graphisch so etwas wie



wobei das Gleichheitszeichen nur meint, daß beide Klammerungen stets dasselbe liefern, wenn wir oben drei Elemente unserer Menge mit Verknüpfung einfüllen. . .

Beispiele 3.1.4. Von unseren Beispielen sind die ersten drei assoziativ und kommutativ, das vierte ist assoziativ aber nicht kommutativ falls X mehr als ein Element hat, das fünfte ist weder assoziativ noch kommutativ.

3.1.5. Ist eine Verknüpfung assoziativ, so liefern auch ungeklammerte Ausdrücke der Form $a_1 \top a_2 \ldots \top a_n$ wohlbestimmte Elemente von A, das Resultat ist genauer unabhängig davon, wie wir die Klammern setzen. Um diese Erkenntnis zu formalisieren, vereinbaren wir für so einen Ausdruck die Interpretation

$$a_1 \top a_2 \ldots \top a_n := a_1 \top (a_2 \top (\ldots (a_{n-1} \top a_n) \ldots))$$

und zeigen

Lemma 3.1.6. *Gegeben* $(A, \top)$ *eine Menge mit einer assoziativen Verknüpfung und* $a_1, \ldots, a_n, b_1, \ldots, b_m \in A$ *gilt*

$$(a_1 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m) = a_1 \top \ldots \top a_n \top b_1 \top \ldots \top b_m$$

Beweis. Wir folgern aus dem Assoziativgesetz $(a_1 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m) = a_1 \top ((a_2 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m))$ und sind fertig mit vollständiger Induktion über n. □

3.1.7. Das Wort Lemma, im Plural Lemmata, kommt von griechisch $\lambda\alpha\mu\beta\alpha\nu\epsilon\iota\nu$ "nehmen" und bezeichnet in der Mathematik kleinere Resultate oder auch Zwischenschritte von größeren Beweisen, denen der Autor außerhalb ihres engeren Kontexts keine größere Bedeutung zumißt.

Ergänzung 3.1.8. Die Zahl der Möglichkeiten, einen Ausdruck in $n+1$ Faktoren so zu verklammern, daß in jedem Schritt nur die Verknüpfung von je zwei Elementen zu berechnen ist, heißt die **n-te Catalan-Zahl** und wird C_n notiert. Die ersten Catalan-Zahlen sind $C_0 = C_1 = 1$, $C_2 = 2$, $C_3 = 5$: Die fünf möglichen Verklammerungen von 4 Elementen sind etwa $(ab)(cd)$, $a(b(cd))$, $a((bc)d)$, $((ab)c)d$ und $(a(bc))d$. Im allgemeinen zeigen wir in 8.3.9, daß sich die Catalan-Zahlen durch die Binomial-Koeffizienten 1.1.16 ausdrücken lassen vermittels der amüsanten Formel

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1}$$

Definition 3.1.9. Sei $(A, \top)$ eine Menge mit Verknüpfung. Ist $n \in \{1, 2, \ldots\}$ eine von Null verschiedene natürliche Zahl und $a \in A$, so schreiben wir

$$\underbrace{a \top a \top \ldots \top a}_{n\text{-mal}} =: n^{\top} a$$

3.1.10. Ist m eine zweite von Null verschiedene natürliche Zahl, so erhalten wir für assoziative Verknüpfungen mithilfe unseres Lemmas 3.1.6 die Regeln $(n+m)^\top a = (n^\top a)\top(m^\top a)$ und $(nm)^\top a = n^\top(m^\top a)$. Ist unsere Verknüpfung auch noch kommutativ, so gilt zusätzlich $n^\top(a\top b) = (n^\top a)\top(n^\top b)$.

3.1.11. Sei $(A, \top)$ eine Menge mit Verknüpfung. Eine Teilmenge $B \subset A$ heißt **abgeschlossen unter der Verknüpfung** genau dann, wenn aus $a, b \in B$ folgt $a\top b \in B$. Natürlich ist in diesem Fall auch $(B, \top)$ eine Menge mit Verknüpfung, man spricht dann von der **auf** B **induzierten Verknüpfung**. Zum Beispiel ist $\mathbb{N} \subset \mathbb{Z}$ abgeschlossen unter der Addition, aber $\mathbb{Z}_{\neq 0} \subset \mathbb{Q}_{\neq 0}$ ist nicht abgeschlossen unter der durch die Division gegebenen Verknüpfung $(a, b) \mapsto a/b$.

Definition 3.1.12. Sei $(A, \top)$ eine Menge mit Verknüpfung. Ein Element $e \in A$ heißt **neutrales Element** von $(A, \top)$ genau dann, wenn gilt

$$e\top a = a\top e = a \quad \forall a \in A$$

3.1.13. In einer Menge mit Verknüpfung $(A, \top)$ kann es höchstens ein neutrales Element e geben, denn für jedes weitere Element e' mit $e'\top a = a\top e' = a \quad \forall a \in A$ haben wir $e' = e'\top e = e$. Wir dürfen also den bestimmten Artikel verwenden und in einer Menge mit Verknüpfung von *dem* neutralen Element reden.

Definition 3.1.14. Ein **Monoid** ist eine Menge mit einer assoziativen Verknüpfung, in der es ein neutrales Element gibt. Ist $(A, \top)$ ein Monoid, so erweitern wir unsere Notation $n^\top a$ aus 3.1.9 auf alle natürlichen Zahlen $n \in \mathbb{N}$, indem wir $0^\top a$ als das neutrale Element von A verstehen, für alle $a \in A$.

3.1.15. Das Wort "Monoid" ist wohl von griechisch "$\mu o\nu o\varsigma$" für "allein" abgeleitet: Ein Monoid besitzt nur eine einzige Verknüpfung.

3.1.16. In Monoiden gelten die Regeln 3.1.10 für alle $n, m \in \mathbb{N}$. Die natürlichen Zahlen bilden mit der Addition ein Monoid $(\mathbb{N}, +)$ mit neutralem Element 0. Sie bilden auch unter der Multiplikation ein Monoid $(\mathbb{N}, \cdot)$ mit neutralem Element 1. Für jede Menge X ist die Menge $\mathrm{Ens}(X)$ der Abbildungen von X in sich selbst ein Monoid. Die leere Menge ist *kein* Monoid, ihr fehlt das neutrale Element.

3.1.17. Gegeben eine Abbildung $I \to A$, $i \mapsto a_i$ von einer endlichen Menge in ein kommutatives additiv bzw. multiplikativ notiertes Monoid vereinbaren wir die Notationen

$$\sum_{i\in I} a_i \qquad \text{bzw.} \qquad \prod_{i\in I} a_i$$

für die "Verknüpfung aller a_i". Ist I die leere Menge, so vereinbaren wir, daß dieser Ausdruck das neutrale Element von A bedeuten möge, also 0 bzw. 1. Wir haben diese Notation bereits verwendet in 2.1.19, und für die konstante Abbildung $I \to \mathbb{N}$, $i \mapsto 1$ hätten wir zum Beispiel

$$\sum_{i \in I} 1 = |I|$$

Unsere Konvention 1.1.13 für mit einem Laufindex notierte Summen bzw. Produkte verwenden wir bei Monoiden analog.

Übung 3.1.18. Sei X eine Menge. Das Schneiden von Teilmengen ist eine Verknüpfung

$$\begin{array}{rccc} \cap : & \mathcal{P}(X) \times \mathcal{P}(X) & \to & \mathcal{P}(X) \\ & (A, B) & \mapsto & A \cap B \end{array}$$

auf der Potenzmenge. Dasselbe gilt für die Vereinigung und das Bilden der Differenzmenge. Welche dieser Verknüpfungen sind kommutativ oder assoziativ? Welche besitzen neutrale Elemente?

Ergänzende Übung 3.1.19. Man gebe die Wahrheitstafeln für $\Rightarrow$ und $\Leftrightarrow$ an. Bezeichne weiter $\neg : \{\text{Wahr}, \text{Falsch}\} \to \{\text{Wahr}, \text{Falsch}\}$ die "Verneinung". Man zeige, daß die Formel

$$(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$$

beim Einsetzen beliebiger Wahrheitswerte aus $\{\text{Wahr}, \text{Falsch}\}$ für A und B stets den Wert "Wahr" ausgibt, in Übereinstimmung mit unseren eher intuitiven Überlegungen in 2.3.6.

3.2 Gruppen

3.2.1. Ich empfehle, bei der Lektüre dieses Abschnitts die Tabelle auf Seite 51 gleich mitzulesen, die die Bedeutungen der nun folgenden Formalitäten in den zwei gebräuchlichsten Notationssystemen angibt. In diesen Notationssystemen sollten alle Formeln aus der Schulzeit vertraut sein. Wir erinnern uns an die Definition eines Monoids aus 3.1.14: Ein Monoid ist eine Menge mit einer assoziativen Verknüpfung, für die es in unserer Menge ein neutrales Element gibt.

Definition 3.2.2.

1. Ist $(A, \top)$ ein Monoid und $a \in A$ ein Element, so nennen wir ein weiteres Element $\bar{a} \in A$ **invers zu** a genau dann, wenn gilt $a \top \bar{a} = e = \bar{a} \top a$ für $e \in A$ das neutrale Element unseres Monoids. Ein Element, das ein Inverses besitzt, heißt **invertierbar**.
2. Eine **Gruppe** ist ein Monoid, in dem jedes Element ein Inverses besitzt.
3. Eine **kommutative Gruppe** oder **abelsche Gruppe** ist eine Gruppe, deren Verknüpfung kommutativ ist.

3.2.3. Der Begriff einer “Gruppe” wurde von Évariste Galois (1811-1832) in die Mathematik eingeführt. Er verwendet den Begriff “Gruppe von Transformationen” sowohl in der Bedeutung einer “Menge von bijektiven Selbstabbildungen einer gegebenen Menge” als auch in der Bedeutung einer “Menge von bijektiven Selbstabbildungen einer gegebenen Menge, die abgeschlossen ist unter Verknüpfung und Inversenbildung”, und die damit in der Tat ein Beispiel für eine Gruppe im Sinne der obigen Definition bildet. Unsere obige Definition 3.2.2 geht auf eine Arbeit von Arthur Cayley aus dem Jahre 1854 mit dem Titel “On the theory of groups as depending on the symbolic equation $\theta^n = 1$” zurück und wurde damit formuliert, bevor Cantor die Sprache der Mengenlehre entwickelte. Die Terminologie “abelsche Gruppe” wurde zu Ehren des norwegischen Mathematikers Niels Hendrik Abel eingeführt.

Lemma 3.2.4. *Jedes Element eines Monoids besitzt höchstens ein Inverses.*

Beweis. Aus $a\top\bar{a} = e = \bar{a}\top a$ und $a\top b = e = b\top a$ folgt durch Anwenden von $b\top$ auf die erste Gleichung mit dem Assoziativgesetz sofort $\bar{a} = b$. □

3.2.5. Wir dürfen also den bestimmten Artikel benutzen und von nun an von *dem* Inversen eines Elements eines Monoids und insbesondere auch einer Gruppe reden. Offensichtlich ist das Inverse des Inversen stets das ursprüngliche Element, in Formeln $\bar{\bar{a}} = a$.

Lemma 3.2.6. *Sind a und b Elemente einer Gruppe oder allgemeiner invertierbare Elemente eines Monoids, so wird das Inverse von $a\top b$ wird gegeben durch die Formel $\overline{(a\top b)} = \bar{b}\top\bar{a}$.*

Beweis. In der Tat rechnen wir schnell $(a\top b)\top(\bar{b}\top\bar{a}) = e = (\bar{b}\top\bar{a})\top(a\top b)$. Diese Formel ist auch aus dem täglichen Leben vertraut: Wenn man sich morgends zuerst die Strümpfe anzieht und dann die Schuhe, so muß man abends zuerst die Schuhe ausziehen und dann die Strümpfe. □

Beispiele 3.2.7. Von unseren Beispielen 3.1.2 für Verknüpfungen oben ist nur $(\mathbb{Z}, +)$ eine Gruppe, und diese Gruppe ist kommutativ. Ein anderes Beispiel für eine kommutative Gruppe ist die Menge $\mathbb{Q}\backslash\{0\}$ der von Null verschiedenen rationalen Zahlen mit der Multiplikation als Verknüpfung.

Übung 3.2.8. Die invertierbaren Elemente eines Monoids bilden stets eine Gruppe. Ein Element a eines Monoids A ist invertierbar genau dann, wenn es $b, c \in A$ gibt mit $b\top a = e = a\top c$ für e das neutrale Element.

Definition 3.2.9. Ist $(A, \top)$ eine Gruppe, so erweitern wir unsere Notation $n^\top a$ aus 3.1.14 auf alle $n \in \mathbb{Z}$, indem wir setzen $n^\top a = (-n)^\top \bar{a}$ für $n \in \{-1, -2, \dots\}$.

	123	213	312	321	132	231
123	123	213	312	321	132	231
213	213	123	321	312	231	132
312	312	132	231	213	321	123
321	321	231	132	123	312	213
132	132	312	213	231	123	321
231	231	321	123	132	213	312

Die Verknüpfungstafel der Gruppe aller Permutationen der Menge $\{1, 2, 3\}$. Eine solche Permutation σ habe ich dargestellt durch das geordnete Zahlentripel $\sigma(1)\sigma(2)\sigma(3)$, und im Kästchen aus der Zeile τ und der Spalte σ steht $\tau \circ \sigma$.

3.2.10. In einer Gruppe gelten offensichtlich sogar für alle ganzen Zahlen $n \in \mathbb{Z}$ die Regeln $(n+m)^{\top}a = (n^{\top}a)\top(m^{\top}a)$ und $(nm)^{\top}a = n^{\top}(m^{\top}a)$. Ist die Gruppe kommutativ, so gilt zusätzlich $n^{\top}(a\top b) = (n^{\top}a)\top(n^{\top}b)$ für alle $n \in \mathbb{Z}$.

3.2.11. Verknüpfungen werden meist additiv oder multiplikativ geschrieben, also $a + b$ oder $a \cdot b$, wobei die additive Schreibweise kommutativen Verknüpfungen vorbehalten ist und die Bruchnotation $1/a$ und b/a aus nebenstehender Tabelle kommutativen multiplikativ geschriebenen Monoiden, in denen b invertierbar ist. Bei additiv geschriebenen Gruppen bezeichnet man das Inverse von a meist als das **Negative** von a. Bei nichtkommutativen und multiplikativ notierten Gruppen oder Monoiden benutzt man für das Inverse von a nur die von der allgemeinen Notation a^n abgeleitete Notation a^{-1}. Die nebenstehende Tabelle faßt die üblichen Notationen für unsere abstrakten Begriffsbildungen in diesem Kontext zusammen und gibt unsere allgemeinen Resultate und Konventionen in diesen Notationen wieder. Diejenigen Formeln und Konventionen, die keine Inversen brauchen, benutzt man auch allgemeiner für beliebige Monoide. Für die Gruppe der invertierbaren Elemente eines multiplikativ notierten Monoids A verwenden wir die Notation $A^{\times}$. Zum Beispiel haben wir $\mathbb{Z}^{\times} = \{1, -1\}$.

Beispiel 3.2.12. Für jede Menge X ist die Menge aller Bijektionen von X auf sich selbst eine Gruppe, mit der Komposition von Abbildungen als Verknüpfung. Wir notieren diese Gruppe $\mathrm{Ens}^{\times}(X)$ in Übereinstimmung mit unserer Konvention 3.2.11, schließlich handelt es sich um die Gruppe der invertierbaren Elemente des Monoids $\mathrm{Ens}(X)$. Ihre Elemente heißen die **Permutationen von** X. Die Gruppe der Permutationen einer Menge X ist für $|X| > 2$ nicht kommutativ. Das Inverse einer Bijektion ist ihre Umkehrabbildung.

Übung 3.2.13. Sind a, b, c Elemente einer Gruppe, so folgt aus $a\top b = a\top c$ bereits $b = c$. Ebenso folgt auch aus $b\top a = c\top a$ bereits $b = c$. Salopp gesprochen dürfen wir also in einer Gruppe “kürzen”. Dasselbe gilt allgemeiner in einem beliebigen Monoid, wenn wir a invertierbar annehmen.

Ergänzende Übung 3.2.14. Sei A ein Monoid und e sein neutrales Element. Man zeige: Unser Monoid ist genau dann eine Gruppe, wenn es für jedes $a \in A$ ein $\bar{a} \in A$ gibt mit $\bar{a}\top a = e$, und dies Element $\bar{a}$ ist dann notwendig das Inverse von a in A. Noch Mutigere zeigen: Ist A eine Menge mit assoziativer Verknüpfung und existiert ein $e \in A$ mit $e\top a = a\ \forall a \in A$ sowie für jedes $a \in A$ ein $\bar{a} \in A$ mit $\bar{a}\top a = e$, so ist A eine Gruppe.

abstrakt	additiv	multiplikativ
$a \top b$	$a+b$	$a \cdot b$, $a \circ b$, ab
e	0	1
$\bar{b}$	$-b$	$1/b$
$a \top \bar{b}$	$a-b$	a/b
$n^\top a$	na	a^n
$e \top a = a \top e = a$	$0+a=a+0=a$	$1 \cdot a = a \cdot 1 = a$
$a \top \bar{a} = e$	$a-a=0$	$a/a=1$
$\bar{\bar{a}} = a$	$-(-a)=a$	$1/(1/a)=a$
$(-1)^\top a = \bar{a}$	$(-1)a=-a$	$a^{-1}=1/a$
$\overline{(a \top b)} = \bar{b} \top \bar{a}$	$-(a+b)=(-b)+(-a)$	$(ab)^{-1}=b^{-1}a^{-1}$, $1/ab=(1/b)(1/a)$
$\overline{(a \top \bar{b})} = b \top \bar{a}$	$-(a-b)=b-a$	$1/(a/b)=b/a$
$n^\top(m^\top a)=(nm)^\top a$	$n(ma)=(nm)a$	$(a^m)^n=a^{nm}$
$(m+n)^\top a = (m^\top a)\top(n^\top a)$	$(m+n)a=(ma)+(na)$	$a^{(m+n)}=(a^m)(a^n)$
$\overline{n^\top a} = (-n)^\top a$	$-(na)=(-n)a$	$(a^n)^{-1}=a^{-n}$
$0^\top a = e$	$0a=0$	$a^0=1$
$n^\top(a \top b)=(n^\top a)\top(n^\top b)$	$n(a+b)=(na)+(nb)$	$(ab)^n=(a^n)(b^n)$

Tabelle 1: Konventionen und Formeln in verschiedenen Notationssystemen. Bereits diese Tabelle muß mit einigen Hintergedanken gelesen werden, weil die Symbole $+, -, 0, 1$ darin in zweierlei Bedeutung vorkommen: Manchmal meinen sie konkrete Operationen in $\mathbb{Z}$ bzw. Elemente von $\mathbb{Z}$, manchmal stehen sie aber auch für Verknüpfung, Inversenbildung und neutrale Elemente in abstrakten Monoiden. Es scheint mir eine gute Übung, die Tabelle durchzugehen und allen Symbolen $0, 1$ einen Hut aufzusetzen, wenn sie nicht ganze Zahlen bedeuten.

3.3 Körper

Definition 3.3.1. Ein **Körper** $(K, +, \cdot)$ (englisch **field**, französisch **corps**) ist eine Menge K mit zwei assoziativen und kommutativen Verknüpfungen, der sogenannten **Addition** $+$ und **Multiplikation** $\cdot$ des Körpers, derart daß die folgenden drei Bedingungen erfüllt sind:

1. $(K, +)$ ist eine Gruppe, die **additive Gruppe** des Körpers.

2. Bezeichnet $0_K \in K$ das neutrale Element der Gruppe $(K, +)$, so folgt aus $a \neq 0_K \neq b$ schon $a \cdot b \neq 0_K$ und $(K \backslash \{0_K\}, \cdot)$ ist eine Gruppe, die **multiplikative Gruppe** des Körpers.

3. Es gilt das **Distributivgesetz**

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad \forall a, b, c \in K$$

Ergänzung 3.3.2. Der Begriff "Körper" ist in diesem Zusammenhang wohl zu verstehen als "besonders gut unter den verschiedensten Rechenoperationen abgeschlossener Zahlbereich", in Analogie zu geometrischen Körpern wie Kugeln oder Zylindern, die man entsprechend als "besonders gut in sich geschlossene Bereiche des Raums" ansehen könnte. Die Bezeichnung "Distributivgesetz" rührt daher, daß uns dies Gesetz erlaubt, beim Multiplizieren eines Elements mit einer Summe den "Faktor auf die Summanden zu verteilen".

3.3.3. Wenn wir mit Buchstaben rechnen, werden wir meist $a \cdot b = ab$ abkürzen. Zusätzlich vereinbaren wir zur Vermeidung von Klammern die Regel "Punkt vor Strich", so daß also zum Beispiel das Distributivgesetz kürzer in der Form $a(b + c) = ab + ac$ geschrieben werden kann. Die multiplikative Gruppe eines Körpers K notieren wir $K^\times = K \backslash \{0_K\}$ in Übereinstimmung mit unserer allgemeinen Notation 3.2.11, schließlich handelt es sich um die Menge der invertierbaren Elemente des multiplikativen Monoids K. Für das neutrale Element der Multiplikation vereinbaren wir die Bezeichnung $1_K \in K^\times$. Wir kürzen meist 0_K ab durch 0 und 1_K durch 1 in der Erwartung, daß man aus dem Kontext erschließt, ob mit 0 und 1 natürliche Zahlen oder Elemente eines speziellen Körpers gemeint sind. Meist kommt es darauf im Übrigen gar nicht an.

3.3.4. Für alle a, b in einem Körper und alle $n \geq 0$ gilt die binomische Formel

$$(a + b)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} a^\nu b^{n-\nu}$$

Um das einzusehen prüft man, daß wir bei der Herleitung nach 1.1.22 nur Körperaxiome verwandt haben. Man beachte hierbei unsere Konvention $0_K^0 = 1_K$

aus 3.1.14, angewandt auf das Monoid $(K, \cdot)$ in Verbindung mit der notationellen Konvention auf Seite 51. Die Multiplikation mit den Binomialkoeffizienten in dieser Formel ist zu verstehen als wiederholte Addition im Sinne der Bezeichnungskonvention na auf Seite 51, angewandt auf den Spezialfall der additiven Gruppe unseres Körpers.

Beispiele 3.3.5. Ein Beispiel für einen Körper ist der Körper der rationalen Zahlen $(\mathbb{Q}, +, \cdot)$. Ein anderes Beispiel ist der zweielementige Körper mit den durch die Axiome erzwungenen Rechenregeln, der fundamental ist in der Informatik. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ bilden keinen Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist, da es nämlich in $\mathbb{Z} \setminus \{0\}$ nur für 1 und -1 ein multiplikatives Inverses gibt. Es gibt keinen einelementigen Körper, da das Komplement seines Nullelements die leere Menge sein müßte: Dies Komplement kann dann aber unter der Multiplikation keine Gruppe sein, da es eben kein neutrales Element haben könnte.

Lemma 3.3.6 (Folgerungen aus den Körperaxiomen). *Sei K ein Körper. So gilt*

1. $a0 = 0 \quad \forall a \in K$.
2. $ab = 0 \Rightarrow a = 0$ *oder* $b = 0$.
3. $-a = (-1)a \quad \forall a \in K$.
4. $(-1)(-1) = 1$.
5. $(-a)(-b) = ab \quad \forall a, b \in K$.
6. $\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}$ *für alle* $a, c \in K$ *und* $b, d \in K^\times$.
7. $\frac{ac}{bc} = \frac{a}{b}$ *für alle* $a \in K$ *und* $b, c \in K^\times$.
8. $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$ *für alle* $a, c \in K$ *und* $b, d \in K^\times$.
9. $m(ab) = (ma)b$ *für alle* $m \in \mathbb{Z}$ *und* $a, b \in K$.

Beweis.

1. Man folgert das aus $a0 + a0 = a(0+0) = a0$ durch Hinzuaddieren von $-(a0)$ auf beiden Seiten.
2. In der Tat folgt aus $(a \neq 0$ und $b \neq 0)$ schon $(ab \neq 0)$ nach den Körperaxiomen.
3. In der Tat gilt $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$, und $-a$ ist ja gerade definiert als das eindeutig bestimmte Element von K so daß $a + (-a) = 0$.
4. In der Tat gilt nach dem Vorhergehenden $(-1)(-1) = -(-1) = 1$.
5. Um das nachzuweisen ersetzen wir einfach $(-a) = (-1)a$ und $(-b) = (-1)b$ und verwenden $(-1)(-1) = 1$.
6. Das ist klar.
7. Das ist klar.
8. Das wird bewiesen, indem man die Brüche auf einen Hauptnenner bringt und das Distributivgesetz anwendet.
9. Das folgt durch wiederholtes Anwenden des Distributivgesetzes.

$\square$

3.3.7. Die Frage, wie das Produkt zweier negativer Zahlen zu bilden sei, war lange umstritten. Mir scheint der vorhergehende Beweis das überzeugendste Argument für "Minus mal Minus gibt Plus": Es sagt salopp gesprochen, daß man diese Regel adoptieren muß, wenn man beim Rechnen das Ausklammern ohne alle Einschränkungen erlauben will.

Definition 3.3.8. Gegeben Mengen mit Verknüpfung $(A, \top)$ und $(B, \perp)$ verstehen wir unter einem **Homomorphismus** von A nach B eine Abbildung $\varphi : A \to B$ derart, daß gilt $\varphi(a \top a') = \varphi(a) \perp \varphi(a')$ für alle $a, a' \in A$. Sind unsere beiden Mengen mit Verknüpfung Monoide, so verstehen wir unter einem **Homomorphismus von Monoiden** einen Homomorphismus von Mengen mit Verknüpfung, der zusätzlich das neutrale Element auf das neutrale Element abbildet. Einen Monoidhomomorphismus zwischen zwei Gruppen nennen wir auch einen **Gruppenhomomorphismus**. Einen bijektiven Homomorphismus nennen wir oft auch einen **Isomorphismus**.

3.3.9. Jeder Homomorphismus φ von Mengen mit Verknüpfung zwischen zwei Gruppen ist bereits ein Gruppenhomomorphismus, bildet also das neutrale Element auf das neutrale Element ab: In der Tat folgt das aus $\varphi(e) = \varphi(e \cdot e) = \varphi(e) \cdot \varphi(e)$ durch Kürzen unmittelbar. Für einen Homomorphismus von Mengen mit Verknüpfung zwischen zwei Monoiden muß das jedoch nicht gelten, wie das Beispiel der Nullabbildung $(\mathbb{Z}, +) \to (\mathbb{Z}, \cdot)$ zeigt.

Vorschau 3.3.10. Den Begriff eines Isomorphismus haben wir hier etwas schlampig eingeführt: Im allgemeinen nennt man einen Homomorphismus ϕ nach **??** einen Isomorphismus genau dann, wenn es einen Homomorphismus ψ in die Gegenrichtung gibt derart, daß beide Kompositionen $\psi \circ \phi$ und $\phi \circ \psi$ die Identität sind. In den Fällen, die uns bis auf weiteres begegnen werden, ist jedoch diese "richtige" Definition zu der oben gegebenen schlampigen Definition äquivalent. Der erste Fall, in dem das nicht mehr richtig ist, wird Ihnen in diesen Vorlesungen in 9.5.21 begegnen: Eine bijektive stetige Abbildung von topologischen Räumen muß keineswegs ein Isomorphismus von topologischen Räumen sein alias eine stetige Umkehrabbildung besitzen.

3.3.11. Die Terminologie kommt von griechisch "$\mu o \rho \varphi \eta$" für "Gestalt" oder für uns besser "Struktur" und griechisch "$o \mu o \iota \varsigma$" für "gleich, ähnlich". Auf deutsch könnte man statt Homomorphismus auch "strukturerhaltende Abbildung" sagen. Das Wort "Isomorphismus" wird analog gebildet mit griechisch "$\iota \sigma o \varsigma$" für "gleich".

Übung 3.3.12. Gegeben Gruppen H und G bezeichne

$$\operatorname{Grp}(H, G)$$

die Menge aller Gruppenhomomorphismen von H nach G. Man zeige, daß für jede Gruppe G die Vorschrift $\varphi \mapsto \varphi(1)$ eine Bijektion $\operatorname{Grp}(\mathbb{Z}, G) \xrightarrow{\sim} G$ liefert.

Ein Gruppenhomomorphismus von der additiven Gruppe der ganzen Zahlen in irgendeine weitere Gruppe ist in Worten also festgelegt und festlegbar durch das Bild von Eins. Hinweis: Man erinnere 3.2.10. Man beachte, daß die 1 nicht das neutrale Element der Gruppe $\mathbb{Z}$ meint, die hier vielmehr als additive Gruppe zu verstehen ist.

3.3.13. Analoge Definitionen verwenden wir auch bei Mengen mit mehr als einer Verknüpfung. Zum Beispiel ist ein **Körperhomomorphismus** φ von einem Körper K in einen Körper L eine Abbildung $\varphi : K \to L$ derart, daß gilt $\varphi(a+b) = \varphi(a) + \varphi(b)$ und $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in K$ und $\varphi(1) = 1$. Die Bedingung $\varphi(1) = 1$ ist nur nötig, um den Fall der Nullabbildung auszuschließen. In anderen Worten mag man einen Körperhomomorphismus auch definieren als eine Abbildung, die sowohl für die Addition als auch für die Multiplikation ein Monoidhomomorphismus ist. Unter einem **Körperisomorphismus** verstehen wir wieder einen bijektiven Körperhomomorphismus.

Übung 3.3.14. Ist K ein Körper derart, daß es kein $x \in K$ gibt mit $x^2 = -1$, so kann man die Menge $K \times K = K^2$ zu einem Körper machen, indem man die Addition und Multiplikation definiert durch

$$\begin{array}{rcc} (a,b) + (c,d) & := & (a+c, b+d) \\ (a,b) \cdot (c,d) & := & (ac - bd, ad + bc) \end{array}$$

Die Abbildung $K \to K^2$, $a \mapsto (a, 0)$ ist dann ein Körperhomomorphismus. Kürzen wir $(a, 0)$ mit a ab und setzen $(0, 1) = \mathrm{i}$, so gilt $\mathrm{i}^2 = -1$ und $(a, b) = a + b\,\mathrm{i}$ und die Abbildung $a + b\,\mathrm{i} \mapsto a - b\,\mathrm{i}$ ist ein Körperisomorphismus $K^2 \xrightarrow{\sim} K^2$.

3.3.15. Auf die in der vorhergehenden Übung 3.3.14 erklärte Weise können wir etwa aus dem Körper $K = \mathbb{R}$ der "reellen Zahlen", sobald wir ihn kennengelernt haben, direkt den Körper $\mathbb{C}$ der **komplexen Zahlen** konstruieren. Unser Körperisomorphismus gegeben durch die Vorschrift $a + b\,\mathrm{i} \mapsto a - b\,\mathrm{i}$ heißt in diesem Fall die **komplexe Konjugation** und wird auch $z \mapsto \bar{z}$ notiert. In **??** diskutieren wir die komplexen Zahlen noch ausführlicher. Man beachte, wie mühelos das alles in der Sprache der Mengenlehre zu machen ist. Als die komplexen Zahlen erfunden wurden, gab es noch keine Mengenlehre und beim Rechnen beschränkte man sich auf das Rechnen mit "reellen" Zahlen, ja selbst das Multiplizieren zweier negativer Zahlen wurde als eine fragwürdige Operation angesehen, und das Ziehen einer Quadratwurzel aus einer negativen Zahl als eine rein imaginäre Operation. In gewisser Weise ist es das ja auch geblieben, aber die Mengenlehre liefert eben unserer Imagination eine wunderbar präzise Sprache, in der wir uns auch über imaginierte Dinge unmißverständlich austauschen können. Man kann dieselbe Konstruktion auch allgemeiner durchführen, wenn man statt -1 irgendein anderes Element eines Körpers K betrachtet, das kein Quadrat ist. Noch allgemeinere Konstruktionen zur “Adjunktion höherer Wurzeln” oder sogar der “Adjunktion von Nullstellen polynomialer Gleichungen” können sie in der Algebra kennenlernen, vergleiche etwa **??**.

Übung 3.3.16. Jeder Gruppenhomomorphismus $\varphi : G \to H$ vertauscht mit Inversenbildung, in Formeln $\varphi(a^{-1}) = (\varphi(a))^{-1}\ \forall a \in G$. Ein Körperhomomorphismus ist stets injektiv.

4 Die reellen Zahlen

Hier trennen sich nun die Wege der linearen Algebra, in der man sich zunächst nur auf die in Kapitel eingeführten algebraischen Konzepte stützt und bis auf weiteres mit beliebigen Körpern arbeiten kann, und der Analysis, für die das Wesen der reellen Zahlen grundlegend ist. Bei der Modellierung des Anschauungsraums spielen jedoch die reellen Zahlen wieder eine zentrale Rolle. Überspitzt könnte man sagen, daß man im Gegensatz zu früher, als die mathematische Modellierung der Ebene mithilfe der euklidischen Axiome an den Anfang gestellt wurde, seit dem Anfang des 20.-ten Jahrhunderts eher von der Modellierung der Gerade ausgeht, wie wir sie im folgenden kennenlernen werden.

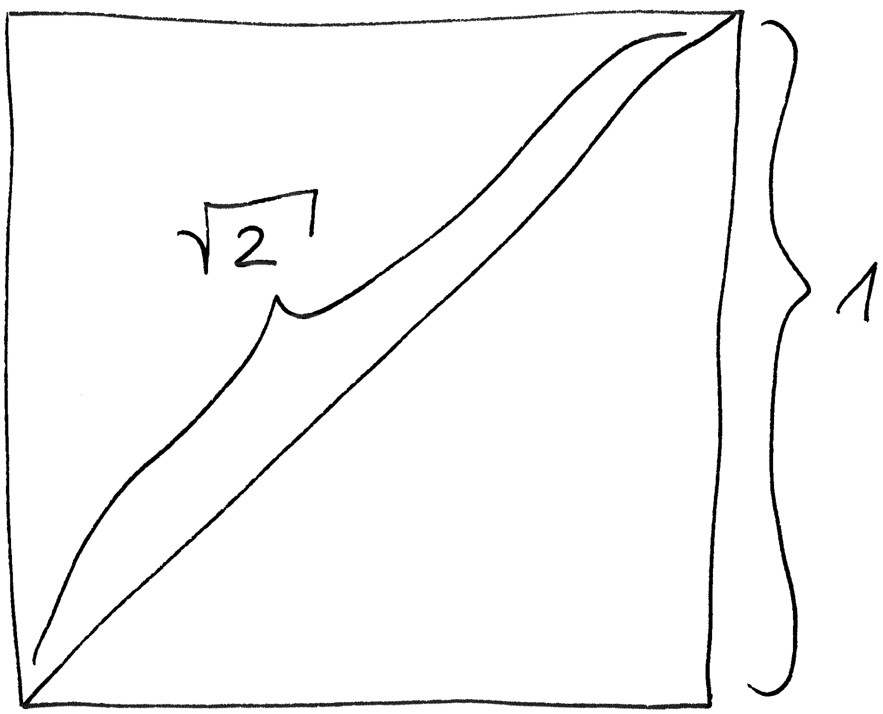
4.1 Wurzeln rationaler Zahlen

Satz 4.1.1. *Es gibt keine rationale Zahl $x \in \mathbb{Q}$ mit $x^2 = 2$.*

Bemerkung 4.1.2. Dieser Satz erklärt, warum wir uns mit den rationalen Zahlen nicht zufrieden geben. In der Tat suchen wir nach einem Zahlbereich, in dem jeder "anschaulichen Länge", wie zum Beispiel der Länge der Diagonale eines Quadrats der Kantenlänge Eins, auch tatsächlich eine Zahl entspricht. Wir zeigen in 5.3.2, daß im Zahlbereich der reellen Zahlen immerhin aus allen nichtnegativen Zahlen Quadratwurzeln gezogen werden können, und diskutieren in 5.4.1, wie man sogar unsere anschauliche Vorstellung von der Länge des Einheitskreises zur Definition einer reellen Zahl präzisieren kann.

Erster Beweis. Setzen wir die in **??** bewiesene Eindeutigkeit der Primfaktorzerlegung als bekannt voraus, so folgt unmittelbar, daß das Quadrat eines unkürzbaren Bruches mit Nenner $\neq \pm 1$ wieder ein unkürzbarer Bruch mit Nenner $\neq \pm 1$ ist. Aus $x \in \mathbb{Q} \setminus \mathbb{Z}$ folgt also $x^2 \in \mathbb{Q} \setminus \mathbb{Z}$. Gäbe es mithin eine rationale Zahl $x \in \mathbb{Q}$ mit $x^2 = 2$, so müßte x bereits selbst eine ganze Zahl sein. Offensichtlich gibt es jedoch keine ganze Zahl $x \in \mathbb{Z}$ mit $x^2 = 2$. □

Zweiter Beweis. Ohne die Eindeutigkeit der Primfaktorzerlegung als bekannt vorauszusetzen können wir in unserer speziellen Situation auch elementarer mit dem Primfaktor Zwei durch Widerspruch argumentieren: Nehmen wir an, wir fänden ganze Zahlen $p, q \in \mathbb{Z}$ mit $q \neq 0$ derart, daß $x = p/q$ ein unkürzbarer Bruch wäre mit $x^2 = 2$. Es folgte $p^2 = 2q^2$, also p^2 gerade, also p gerade, also p^2 durch 4 teilbar, also q^2 gerade, also q gerade. Dann wäre unser Bruch aber doch kürzbar gewesen, nämlich durch Zwei. □


√2
1

4.2 Ordnungen auf Mengen

Definition 4.2.1. Eine **Relation** R auf einer Menge X ist eine Teilmenge $R \subset X \times X$ des kartesischen Produkts von X mit sich selbst, also eine Menge von Paaren von Elementen aus X. Statt $(x, y) \in R$ schreiben wir meist xRy. Eine Relation R heißt eine **Ordnungsrelation** oder auch eine **partielle Ordnung** oder **Halbordnung** oder auch einfach nur eine **Ordnung** genau dann, wenn für alle $x, y, z \in X$ gilt:

1. **Transitivität:** $(xRy$ und $yRz) \Rightarrow xRz$;

2. **Antisymmetrie:** $(xRy$ und $yRx) \Rightarrow x = y$;

3. **Reflexivität:** Es gilt xRx für alle $x \in X$.

Auf Englisch benutzt man für eine partiell geordnete Menge alias einen "partially ordered set" auch oft die Abkürzung **poset**. Eine Ordnungsrelation heißt eine **Anordnung** oder auch eine **totale Ordnung** oder auch eine **lineare Ordnung** genau dann, wenn wir zusätzlich haben

4. **Totalität:** Für alle $x, y \in X$ gilt xRy oder yRx.

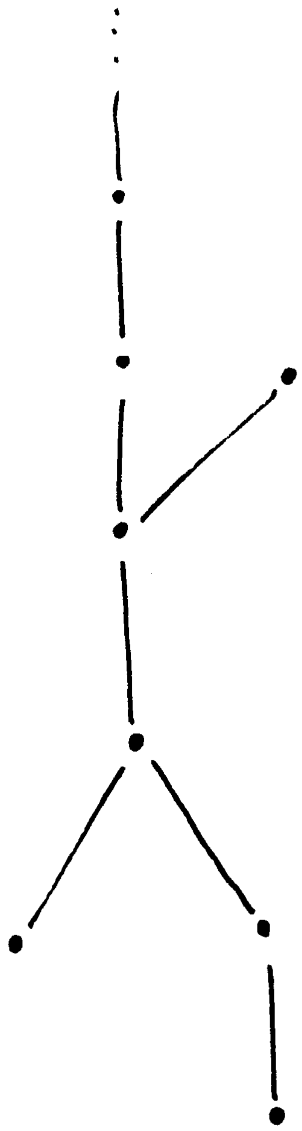
4.2.2. Ordnungsrelationen schreibt man meist $x \leq y$, statt $x \leq y$ schreibt man dann oft auch $y \geq x$. Weiter kürzt man ($x \leq y$ und $x \neq y$) ab mit $x < y$ und ebenso ($x \geq y$ und $x \neq y$) mit $x > y$. Auf jeder angeordneten Menge definieren wir Verknüpfungen $\max$ und $\min$ in offensichtlicher Verallgemeinerung von 3.1.2.3.

Definition 4.2.3. Sei $(Y, \leq)$ eine partiell geordnete Menge.

1. Ein Element $g \in Y$ heißt ein **größtes Element** von Y genau dann, wenn gilt $g \geq y \quad \forall y \in Y$. Ein Element $g \in Y$ heißt ein **maximales Element** von Y genau dann, wenn es kein $y \in Y$ gibt mit $y > g$.

2. Ein Element $k \in Y$ heißt ein **kleinstes Element** von Y genau dann, wenn gilt $k \leq y \quad \forall y \in Y$. Ein Element $k \in Y$ heißt ein **minimales Element** von Y genau dann, wenn es kein $y \in Y$ gibt mit $y < k$.

4.2.4. Jede partiell geordnete Menge besitzt höchstens ein größtes und höchstens ein kleinstes Element. Wir dürfen deshalb den bestimmten Artikel verwenden und von **dem** größten bzw. kleinsten Element reden. Besitzt eine partiell geordnete Menge ein größtes bzw. ein kleinstes Element, so ist dies auch ihr einziges maximales bzw. minimales Element. Sonst kann es jedoch maximale bzw. minimale Elemente in großer Zahl geben, zumindest dann, wenn unsere Ordnungsrelation keine Anordnung ist.



Eine partiell geordnete Menge mit zwei minimalen und einem maximalen Element, die weder ein kleinstes noch ein größtes Element besitzt. Die Darstellung ist in der Weise zu verstehen, daß die fetten Punkte die Elemente unserer Menge bedeuten und daß ein Element größer ist als ein anderers genau dann, wenn es von diesem “durch einen aufsteigenden Weg erreicht werden kann”.

Definition 4.2.5. Sei $(X, \leq)$ eine partiell geordnete Menge und $Y \subset X$ eine Teilmenge.

1. Ein Element $o \in X$ heißt eine **obere Schranke** von Y genau dann, wenn gilt $o \geq y \quad \forall y \in Y$.

2. Ein Element $u \in X$ heißt eine **untere Schranke** von Y genau dann, wenn gilt $u \leq y \quad \forall y \in Y$.

Definition 4.2.6. Sei $(X, \leq)$ eine partiell geordnete Menge und $Y \subset X$ eine Teilmenge.

1. Ein Element $s \in X$ heißt die **kleinste obere Schranke** oder das **Supremum**indexsup@$\sup$, Supremum von Y in X genau dann, wenn s das kleinste Element ist in der Menge $\{o \in X \mid o \text{ ist obere Schranke von } Y\}$. Wir schreiben dann $s = \sup Y$ oder genauer $s = \sup_X Y$.

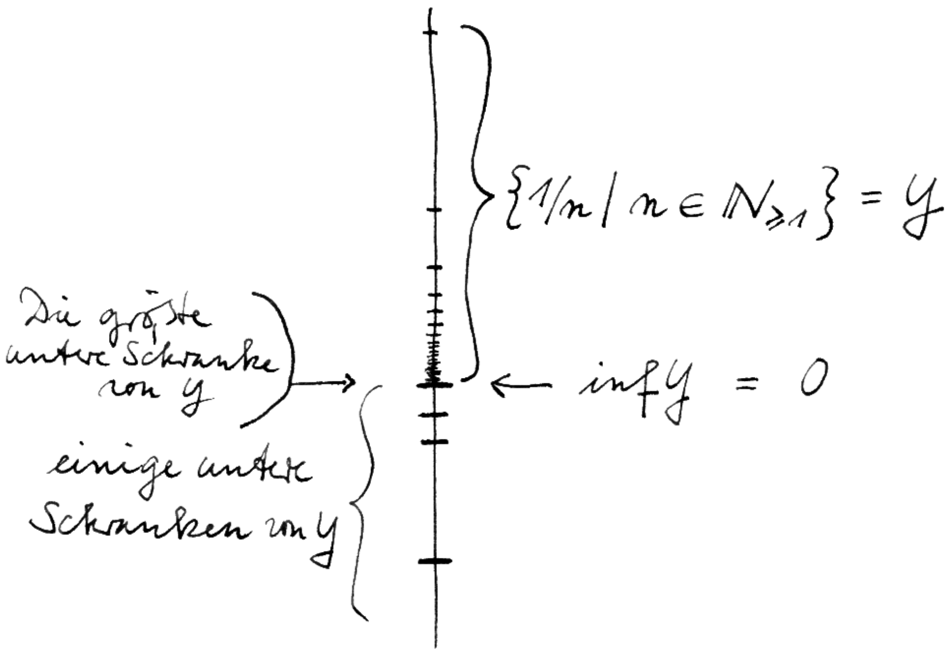
2. Ein Element $i \in X$ heißt die **größte untere Schranke** oder auch das **Infimum** von Y in X genau dann, wenn i das größte Element ist in der Menge $\{u \in X \mid u \text{ ist untere Schranke von } Y\}$. Wir schreiben dann $i = \inf Y$ oder genauer $i = \inf_X Y$.

Beispiel 4.2.7. Die Teilmenge $Y = \{q \in \mathbb{Q} \mid q < 1\} \subset \mathbb{Q}$ hat kein größtes Element, besitzt jedoch in $\mathbb{Q}$ eine kleinste obere Schranke, nämlich $\sup Y = 1$. Die Teilmenge $Z = \{q \in \mathbb{Q} \mid q \leq 1\} \subset \mathbb{Q}$ hat ein größtes Element, nämlich die 1, und das ist dann natürlich auch gleichzeitig ihre kleinste obere Schranke in $\mathbb{Q}$, also haben wir auch $\sup Z = 1$. Die Teilmenge $Y = \{q \in \mathbb{Q} \mid q < 1\} \subset \mathbb{Q}$ hat in $\mathbb{Q}$ keine untere Schranke und dann natürlich erst recht keine größte untere Schranke.

4.2.8. Besitzt eine Teilmenge $Y \subset X$ ein größtes Element $g \in Y$, so gilt $g = \sup Y$. Besitzt eine Teilmenge $Y \subset X$ ein kleinstes Element $k \in Y$, so gilt $k = \inf Y$. Sind Teilmengen $Z \subset Y \subset X$ gegeben und besitzen Z und Y ein Supremum in X, so gilt $\sup Z \leq \sup Y$.

Ergänzendes Beispiel 4.2.9. Auf der Potenzmenge einer beliebigen Menge ist die Inklusionsrelation eine partielle Ordnung. Bezüglich dieser Ordnung ist die Vereinigungsmenge im Sinne von **??** eines Mengensystems sein Supremum und die Schnittmenge sein Infimum.

Übung 4.2.10. Die Menge $\{x \in \mathbb{Q} \mid x^2 \leq 2\}$ besitzt in $\mathbb{Q}$ keine größte untere Schranke.



Die Menge $Y = \{1/n \mid n \in \mathbb{N}_{\geq 1}\}$ besitzt in $\mathbb{Q}$ eine größte untere Schranke, nämlich die Null, in Formeln $\inf Y = 0$. Sie besitzt in $\mathbb{Q}$ auch eine kleinste obere Schranke, nämlich die Eins, in Formeln $\sup Y = 1$.

4.3 Angeordnete Körper

Definition 4.3.1. Ein **angeordneter Körper** ist ein Körper $(K, +, \cdot)$ mit einer Anordnung $\leq$ derart, daß gilt

1. $x \leq y \;\Rightarrow\; x + z \leq y + z \quad \forall x, y, z \in K.$
2. $(0 \leq x$ und $0 \leq y) \;\Rightarrow\; 0 \leq xy \quad \forall x, y \in K.$

Die Elemente $x \in K$ mit $x > 0$ bzw. $x < 0$ nennt man **positiv** bzw. **negativ**. Die Elemente mit $x \geq 0$ bzw. $x \leq 0$ nennt man folgerichtig **nichtnegativ** bzw. **nichtpositiv**.

Lemma 4.3.2. *In jedem angeordneten Körper gilt:*

1. $(x \leq y$ *und* $a \leq b) \Rightarrow (x + a \leq y + b).$
2. $(x \leq y$ *und* $a \geq 0) \Rightarrow (ax \leq ay).$
3. $(0 \leq x < y$ *und* $0 \leq a < b) \Rightarrow (0 \leq xa < yb).$
4. $(x \leq y) \Rightarrow (-y \leq -x).$
5. $(x \geq y$ *und* $a \leq 0) \Rightarrow (ax \leq ay).$
6. $(x \neq 0) \Rightarrow (x^2 > 0).$
7. $1 > 0.$
8. $(x > 0) \Rightarrow (x^{-1} > 0).$
9. $(0 < x \leq y) \Rightarrow (0 < y^{-1} \leq x^{-1}).$

Beweis.

1. In der Tat folgt $x + a \leq y + a \leq y + b$.
2. In der Tat folgt $0 \leq y - x$, also $0 \leq a(y - x) = ay - ax$ und damit dann $ax \leq ay$.
3. In der Tat erhalten wir $0 \leq xa \leq ya < yb$.
4. Das folgt durch Addition von $(-y - x)$ auf beiden Seiten.
5. In der Tat folgern wir $x \geq y \Rightarrow (-a)x \geq (-a)y \Rightarrow ax \leq ay$.
6. In der Tat ist $x^2 = (-x)^2$ und $x > 0 \Leftrightarrow (-x) < 0$.
7. Das folgt aus $1 = 1^2 \neq 0$.

8. Das folgt durch Multiplikation mit $(x^{-1})^2$.

9. Das folgt durch Multiplikation mit $y^{-1}x^{-1}$.

$\square$

4.3.3. Schreiben wir zur besonderen Betonung wieder 0_K und 1_K, so gelten demnach in jedem angeordneten Körper K die Ungleichungen

$$\ldots < (-1_K) + (-1_K) < (-1_K) < 0_K < 1_K < 1_K + 1_K < \ldots$$

Insbesondere folgt aus $m1_K = n1_K$ für $m, n \in \mathbb{Z}$ schon $m = n$. Die Abbildung $\mathbb{Z} \to K$, $m \mapsto m1_K$ ist also eine Injektion. Das Bild dieser Injektion müßte wohl eigentlich einen eigenen Namen kriegen, zum Beispiel $\mathbb{Z}_K$, aber wir kürzen unsere Notation ab, bezeichnen dieses Bild auch mit $\mathbb{Z}$ und schreiben kürzer m statt $m1_K$. Weiter erhalten wir auch eine Injektion $\mathbb{Q} \to K$, $m/n \mapsto m1_K/n1_K$, die wir zum Beispiel $q \mapsto q_K$ notieren könnten. Wir sind auch hier etwas nachlässig, bezeichnen das Bild unserer Injektion $\mathbb{Q} \to K$ meist kurzerhand mit demselben Buchstaben $\mathbb{Q}$ statt genauer $\mathbb{Q}_K$ zu schreiben, und hängen auch den Elementen von $\mathbb{Q}$ meist keinen Index an, wenn wir eigentlich ihr Bild in K meinen.

Übung 4.3.4. Man zeige, daß für jeden angeordneten Körper K die in 4.3.3 definierte Abbildung $\mathbb{Q} \to K$ ein Körperhomomorphismus ist.

Definition 4.3.5. Für jeden angeordneten Körper K definieren wir eine Abbildung $K \to K$, $x \mapsto |x|$, den **Absolutbetrag**, durch die Vorschrift

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

4.3.6. Wir listen einige Eigenschaften des Absolutbetrags auf. Der Beweis der ersten vier sei dem Leser überlassen.

1. $|x| = 0 \Leftrightarrow x = 0$

2. $|-x| = |x|$

3. $|xy| = |x||y|$

4. $|x^{-1}| = |x|^{-1}$ falls $x \neq 0$

5. Es gilt die sogenannte **Dreiecksungleichung**

$$|x + y| \leq |x| + |y| \quad \forall x, y \in K$$

in Worten: Der Betrag einer Summe ist stets kleinergleich der Summe der Beträge der Summanden. In der Tat gilt ja $x+y \leq |x|+|y|$ und ebenso $-(x+y) \leq |x| + |y|$. Unsere Ungleichung heißt deshalb Dreiecksungleichung, weil sie in einem allgemeineren Kontext sagt, daß in einem Dreieck zwei Seiten zusammen stets länger sind als die dritte.

6. $||a| - |b|| \leq |a+b| \quad \forall a, b \in K.$

   In der Tat folgt aus der Dreiecksungleichung $|a| = |(a+b)+(-b)| \leq |a+b| + |-b| = |a+b| + |b|$, also $|a| - |b| \leq |a+b|$. Ebenso folgert man aber auch $|b| - |a| \leq |a+b|$.

Übung 4.3.7. In jedem angeordneten Körper gilt:

1. Aus $|x-a| \leq \eta$ und $|y-b| \leq \eta$ folgt $|(x+y)-(a+b)| \leq 2\eta$;
2. Aus $|x-a| \leq \eta \leq 1$ und $|y-b| \leq \eta \leq 1$ folgt $|xy-ab| \leq \eta(|b|+1+|a|)$;
3. Aus $|y-b| \leq \eta \leq |b|/2$ und $b \neq 0$ folgt $y \neq 0$ und $|1/y - 1/b| \leq 2\eta/|b|^2$.

Ergänzende Übung 4.3.8. In jedem angeordneten Körper gilt für $x \geq -1$ und $n \in \mathbb{N}$ die sogenannte **Bernoulli-Ungleichung** $(1+x)^n \geq 1+nx$. Hinweis: Vollständige Induktion.

4.4 Die reellen Zahlen

4.4.1. Der folgende Satz enthält diejenige Charakterisierung eines gewissen angeordneten Körpers von “reellen Zahlen”, auf der die ganze Vorlesung aufbaut. Er bildet auch einen wesentlichen Teil des Begriffsgebäudes, das es uns ermöglicht, unsere geometrischen Vorstellungen in der heute gebräuchlichen aus den Symbolen der Mengenlehre aufgebauten Sprache der höheren Mathematik wiederzufinden. Bereits im vierten Jahrhundert vor Christus erklärte der griechische Mathematiker Eudoxos eine Zahl als das Verhältnis zweier Längen und gab damit eine geometrische Beschreibung dessen, was wir heute “positive reelle Zahlen” nennen würden. Die logischen Feinheiten der Beziehung dieses “geometrischen” Zahlbegriffs zum “algorithmischen” Zahlbegriff, der vom Prozeß des Zählens herkommt, wurden erst nach und nach verstanden. Den folgenden Satz und seinen Beweis mag man als den Schlußpunkt dieser Entwicklung ansehen. Der Zwischenwertsatz 6.2.6 illustriert, wie gut unsere im Reich der abstrakten Logik und Mengenlehre konstruierten reellen Zahlen unsere geometrische Anschauung modellieren.

4.4.2. Der Beweis der Existenz eines angordneten Körpers mit den im Satz präzisierten Eigenschaften ist für das weitere Verständnis der Vorlesung belanglos. Ich gebe hier nur eine Beweisskizze, als da heißt einen Beweis für höhere Semester, um Sie zu überzeugen, daß wir nicht während des nächsten halben Jahres Folgerungen ziehen aus Grundannahmen, die überhaupt nie erfüllt sind. Ich rate dazu, bei der ersten Lektüre auf das genauere Studium des Existenzbeweises zu verzichten, der sich bis 4.4.8 hinzieht. Vom Beweis der Eindeutigkeit sind Teile durchaus auch für die Ziele dieser Vorlesung von Belang. Wir formulieren diese

Teile als die eigenständigen Aussagen 4.4.12 und 4.4.13. Der Rest wird dem Leser als Übung 4.4.18 überlassen, die wieder für höhere Semester gedacht ist.

Satz 4.4.3 (Charakterisierung der reellen Zahlen). *1. Es gibt einen angeordneten Körper $(\mathbb{R}, +, \cdot, \leq)$ derart, daß in der angeordneten Menge $\mathbb{R}$ jede nichtleere Teilmenge mit einer unteren Schranke auch eine größte untere Schranke besitzt.*

2. Solch ein angeordneter Körper ist im wesentlichen eindeutig bestimmt. Ist genauer $(\mathbb{R}', +, \cdot, \leq)$ ein weiterer derartiger angeordneter Körper, so gibt es genau einen Körperisomorphismus $\varphi : \mathbb{R} \xrightarrow{\sim} \mathbb{R}'$, und für diesen Körperisomorphismus gilt $\alpha \leq \beta \Leftrightarrow \varphi(\alpha) \leq \varphi(\beta)$.

4.4.4. Die im Satz gegebene Charakterisierung trifft auf den angeordneten Körper der rationalen Zahlen nicht zu. Zum Beispiel wissen wir nach 4.2.10, daß die Menge $\{x \in \mathbb{Q} \mid x^2 \leq 2\}$ in $\mathbb{Q}$ keine größte untere Schranke besitzt. Sie ist jedoch nicht leer und besitzt in $\mathbb{Q}$ durchaus untere Schranken, nur eben keine größte.

Übung 4.4.5. Man zeige, daß es für je zwei nichtleere Teilmengen $M, N \subset \mathbb{R}$ mit $x \leq y$ für alle $x \in M$ und $y \in N$ stets ein $a \in \mathbb{R}$ gibt mit $x \leq a \leq y$ für alle $x \in M$ und $y \in N$. Man zeige weiter, daß diese Bedingung an eine angeordnete Menge sogar gleichbedeutend ist zur Forderung, jede nichtleere Teilmenge mit einer unteren Schranke möge auch eine größte untere Schranke besitzen.

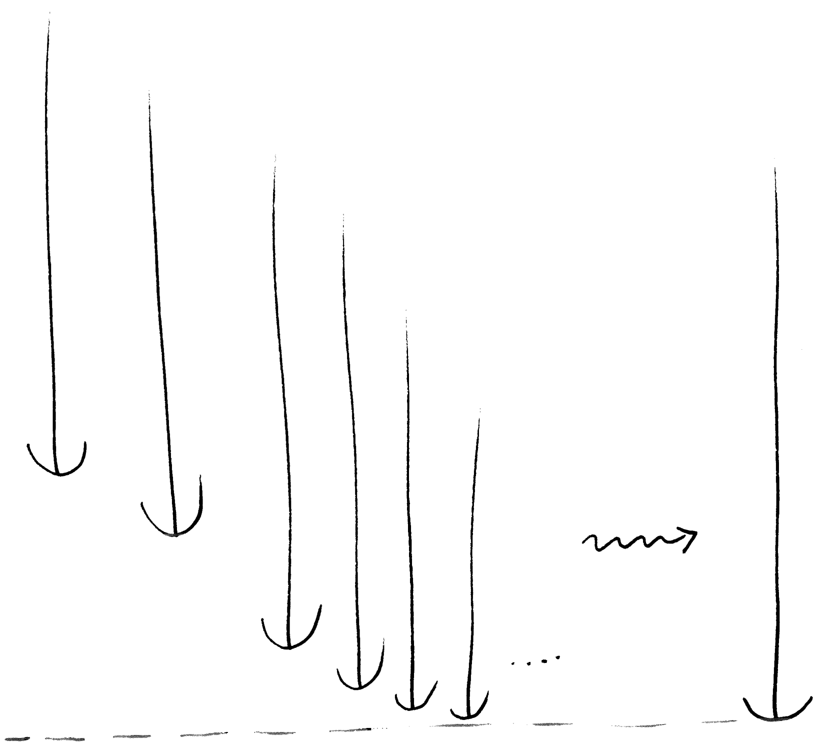
Beweis von 4.4.3.1. Wir konstruieren einen derartigen Körper $\mathbb{R} = \mathbb{R}_{\mathrm{D}}$ als eine Menge $\mathbb{R} \subset \mathcal{P}(\mathbb{Q})$ von Teilmengen der Menge $\mathbb{Q}$ aller rationalen Zahlen,

$$\mathbb{R} \quad := \quad \left\{ \alpha \subset \mathbb{Q} \;\middle|\; \begin{array}{l} \alpha \text{ ist nicht leer,} \\ \alpha \text{ hat eine untere Schranke,} \\ \alpha \text{ hat kein kleinstes Element,} \\ \alpha \text{ enthält mit } x \text{ auch jedes } y > x. \end{array} \right\}$$

Man nennt solch ein α einen **Dedekind'schen Schnitt** und bezeichnet die so konstruierte Menge $\mathbb{R}$ als das "Dedekind'sche Modell der reellen Zahlen". Auf unserer Menge $\mathbb{R}$ von Teilmengen von $\mathbb{Q}$ ist die Inklusionsrelation eine Anordnung und wir schreiben $\alpha \leq \beta$ statt $\alpha \supset \beta$. Ist $Y \subset \mathbb{R}$ eine nichtleere Teilmenge mit unterer Schranke, so liegt offensichtlich auch die Vereinigung

$$\bigcup_{\alpha \in Y} \alpha = \{q \in \mathbb{Q} \mid \text{Es gibt } \alpha \in Y \text{ mit } q \in \alpha\}$$

aller Teilmengen aus Y in $\mathbb{R}$ und ist das Infimum von Y. Damit haben wir bereits eine angeordnete Menge mit der geforderten Eigenschaft konstruiert. Wir müssen darauf nur noch eine Addition und eine Multiplikation erklären derart, daß unsere



Zum Infimum in $\mathbb{R}$

Struktur zu einem angeordneten Körper wird. Die Addition ist unproblematisch: Wir setzen

$$\alpha + \beta = \{x + y \mid x \in \alpha,\ y \in \beta\}$$

und prüfen mühelos, daß aus $\alpha, \beta \in \mathbb{R}$ schon folgt $\alpha + \beta \in \mathbb{R}$, daß $\mathbb{R}$ so zu einer kommutativen Gruppe wird mit neutralem Element $0_{\mathbb{R}} = \{x \in \mathbb{Q} \mid x > 0\}$, und daß gilt $\alpha \leq \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma$ für alle $\alpha, \beta, \gamma \in \mathbb{R}$. Wir erlauben uns nun die Abkürzung $0_{\mathbb{R}} = 0$. Die Multiplikation positiver Elemente ist ebenfalls unproblematisch: Für $\alpha, \beta \in \mathbb{R}$ mit $\alpha > 0$, $\beta > 0$ setzen wir

$$\alpha\beta = \{xy \mid x \in \alpha,\ y \in \beta\}$$

und prüfen mühelos, daß aus $\alpha, \beta \in \mathbb{R}_{>0}$ schon folgt $\alpha\beta \in \mathbb{R}_{>0}$ und daß $\mathbb{R}_{>0}$ so zu einer kommutativen Gruppe mit neutralem Element $1_{\mathbb{R}} = \{x \in \mathbb{Q} \mid x > 1\}$ wird. Das Distributivgesetz in $\mathbb{Q}$ impliziert mit diesen Definitionen auch sofort die Regel

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

für alle $\alpha, \beta, \gamma \in \mathbb{R}_{>0}$. Um unsere Multiplikation so auf ganz $\mathbb{R}$ auszudehnen, daß $\mathbb{R}$ ein angeordneter Körper wird, verwenden wir das anschließende technische Lemma 4.4.6. Der Beweis der in Teil 2 behaupteten Eindeutigkeit ist dann Übung 4.4.18. □

Lemma 4.4.6. *Sei $(R, +)$ eine kommutative Gruppe mit einer Anordnung $\leq$ derart, daß gilt $\alpha \leq \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma \quad \forall \alpha, \beta, \gamma \in R$. Sei auf $R_{>0}$ eine Verknüpfung $(\alpha, \beta) \mapsto \alpha\beta$ gegeben, die $R_{>0}$ zu einer kommutativen Gruppe macht. Gilt außerdem die Regel*

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma \quad \forall \alpha, \beta, \gamma \in R_{>0}$$

so gibt es genau eine Fortsetzung unserer Verknüpfung $(\alpha, \beta) \mapsto \alpha\beta$ auf ganz R derart, daß $(R, +, \cdot, \leq)$ ein angeordneter Körper wird.

Ergänzung 4.4.7. Dies Lemma oder eigentlich sein Beweis kann als die mathematische Begründung der Regel "Minus mal Minus gibt Plus" verstanden werden,.

Beweis. Wenn die Fortsetzung unserer Multiplikation auf ganz R das Distributivgesetz erfüllen soll, müssen wir notwendig setzen

$$\alpha\beta = \begin{cases} 0 & \alpha = 0 \text{ oder } \beta = 0; \\ -((-\alpha)\beta) & \alpha < 0,\ \beta > 0; \\ -(\alpha(-\beta)) & \alpha > 0,\ \beta < 0; \\ (-\alpha)(-\beta) & \alpha < 0,\ \beta < 0. \end{cases}$$

Es gilt nur noch, für diese Multiplikation die Körperaxiome nachzuweisen. Unsere Multiplikation auf R ist offensichtlich kommutativ und assoziativ und macht $R\backslash\{0\}$ zu einer Gruppe. Wir müssen also nur noch das Distributivgesetz

$$\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma \quad \forall\alpha,\beta,\gamma\in R$$

nachweisen. Ohne Beschränkung der Allgemeinheit dürfen wir hierbei $\alpha>0$ und $\beta+\gamma>0$ annehmen, und die einzigen nicht offensichtlichen Fälle sind dann $\beta>0,\ \gamma<0$ bzw. $\beta<0,\ \gamma>0$. Sei ohne Beschränkung der Allgemeinheit $\beta>0,\ \gamma<0$. Nach unseren Annahmen gilt ja die Regel

$$\alpha\beta=\alpha((\beta+\gamma)+(-\gamma))=\alpha(\beta+\gamma)+\alpha(-\gamma)$$

und daraus folgt sofort $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$ auch in diesem letzten Fall. □

Übung 4.4.8. Bezeichne $\mathbb{R}=\mathbb{R}_{\mathrm{D}}$ das Dedekind'sche Modell der reellen Zahlen. Man zeige, daß die durch die Struktur eines angeordneten Körpers auf $\mathbb{R}$ nach 4.3.3 definierte Abbildung $\mathbb{Q}\hookrightarrow\mathbb{R}$ auch beschrieben werden kann durch die Vorschrift $p\mapsto p_{\mathbb{R}}=\{q\in\mathbb{Q}\mid q>p\}$.

Definition 4.4.9. Wir wählen für den weiteren Verlauf der Vorlesung einen festen angeordneten Körper $(\mathbb{R},+,\cdot,\leq)$, in dem jede nichtleere Teilmenge mit einer unteren Schranke auch eine größte untere Schranke besitzt, erlauben uns wegen der in 4.4.3.2 formulierten "Eindeutigkeit bis auf eindeutigen Isomorphismus" den bestimmten Artikel und nennen ihn **den Körper $\mathbb{R}$ der reellen Zahlen**.

Übung 4.4.10. Man zeige: Jede nichtleere Teilmenge von $\mathbb{R}$ mit einer oberen Schranke hat auch eine kleinste obere Schranke.

Übung 4.4.11. Seien X und Y nichtleere nach oben beschränkte Teilmengen von $\mathbb{R}$. Bezeichnet $X+Y\subset\mathbb{R}$ die Menge $\{x+y\mid x\in X,\ y\in Y\}$, so zeige man $\sup(X+Y)=\sup X+\sup Y$.

Satz 4.4.12. *Die natürlichen Zahlen besitzen keine obere Schranke in den reellen Zahlen, d.h. für alle $x\in\mathbb{R}$ gibt es ein $n\in\mathbb{N}$ mit $n>x$.*

Beweis. Das ist evident für den angeordneten Körper $\mathbb{R}$, den wir im Beweis von 4.4.3 konstruiert haben. Da diese Konstruktion jedoch nicht ganz einfach war, zeigen wir auch, wie unser Satz direkt aus unserer Definition 4.4.9 abgeleitet werden kann. Dazu argumentieren wir durch Widerspruch: Hätte die Teilmenge $\mathbb{N}\subset\mathbb{R}$ eine obere Schranke, so hätte sie nach 4.4.10 auch eine kleinste obere Schranke a. Dann wäre aber $a-1<a$ keine obere Schranke von $\mathbb{N}$, also gäbe es $n\in\mathbb{N}$ mit $n>(a-1)$. Es folgte $(n+1)>a$, und a wäre doch keine obere Schranke von $\mathbb{N}$ gewesen. □

Korollar 4.4.13. *1. Unter jeden reellen Zahl findet man noch ganze Zahlen, in Formeln: Für alle $x \in \mathbb{R}$ gibt es $m \in \mathbb{Z}$ mit $m < x$.*

2. Für jede positive reelle Zahl $\varepsilon > 0$ gibt es eine positive natürliche Zahl $n \geq 1$ mit $0 < \frac{1}{n} < \varepsilon$.

3. Zwischen zwei verschiedenen reellen Zahlen liegt stets noch eine rationale Zahl, in Formeln: Gegeben $x < y$ in $\mathbb{R}$ gibt es $r \in \mathbb{Q}$ mit $x < r < y$.

Beweis. Die erste Aussage folgt aus 4.4.12. Um die Zweite zu zeigen, suche man $n > 1/\varepsilon$. Um die dritte Aussage zu zeigen, suchen wir zunächst $n \in \mathbb{N}$, $n \geq 1$ mit $0 < \frac{1}{n} < y - x$, also $1 < ny - nx$, das heißt $1 + nx < ny$. Nun gibt es $a, b \in \mathbb{Z}$ mit $a < ny < b$, also gibt es eine größte ganze Zahl m mit $m < ny$ und folglich $ny \leq m + 1$, woraus hinwiederum folgt $nx < m$ und dann $x < \frac{m}{n} < y$. □

4.4.14. Ein angeordneter Körper heißt **archimedisch angeordnet** genau dann, wenn es zu jedem Element x des Körpers eine natürliche Zahl $n \in \mathbb{N}$ gibt mit $n > x$. Das obige Korollar 4.4.13 gilt mit demselben Beweis für jeden archimedisch angeordneten Körper.

Definition 4.4.15. Mit einem endlichen Dezimalbruch wie $3,141$ bezeichnet man wie auf der Schule die rationale Zahl $3141/1000$. Die durch einen **unendlichen Dezimalbruch** wie $3,1415\ldots$ dargestellte reelle Zahl definieren wir als das Supremum der Menge aller ihrer endlichen Teilausdrücke bzw. das Infimum, wenn ein Minus davorsteht. Wir setzen also zum Beispiel

$$3,1415\ldots \quad = \quad \sup \left\{ \begin{array}{l} 3 \\ 3,1 \\ 3,14 \\ 3,141 \\ 3,1415 \\ \ldots \end{array} \right\}$$

wo wir die Elemente der Menge aller endlichen Teilausdrücke der Übersichtlichkeit halber untereinander geschrieben haben statt sie durch Kommata zu trennen und rationale Zahlen stillschweigend identifiziert haben mit ihren Bildern in $\mathbb{R}$.

Proposition 4.4.16. *1. Jede reelle Zahl läßt sich durch einen unendlichen Dezimalbruch darstellen.*

2. Genau dann stellen zwei verschiedene unendliche Dezimalbrüche dieselbe reelle Zahl dar, wenn es eine Stelle vor oder nach dem Komma gibt und eine von Neun verschiedene Ziffer z derart, daß die beiden Dezimalbrüche bis zu dieser Stelle übereinstimmen, ab dieser Stelle jedoch der eine die Form $z99999\ldots$ hat und der andere die Form $(z + 1)00000\ldots$.

Beweis. Für den ersten Teil reicht es zu zeigen, daß sich jede nichtnegative reelle Zahl $y \geq 0$ als ein unendlicher Dezimalbruch darstellen läßt. Nehmen wir zu jedem $s \in \mathbb{N}$ die größte reelle Zahl $r_s \leq y$ unter y mit höchstens s Stellen nach dem Komma, so gilt

$$y = \sup\{r_0, r_1, r_2, \ldots\}$$

In der Tat ist y eine obere Schranke dieser Menge, aber jede reelle Zahl $x < y$ ist keine obere Schranke dieser Menge: Nach 4.4.13 oder, genauer, seinem Beweis gibt es nämlich für jedes $x \in \mathbb{R}$ mit $x < y$ ein $s \in \mathbb{N}$ und ein $m \in \mathbb{Z}$ mit $x < m \cdot 10^{-s} < y$, als da heißt, es gibt ein s mit $x < r_s$. Den zweiten Teil überlassen wir dem Leser zur Übung. □

Ergänzung 4.4.17. Ich will die Gleichheit $1 = 0,999\ldots$ auch noch im Dedekind'schen Modell der reellen Zahlen erläutern und beginne dazu mit der offensichtlichen Gleichheit

$$\{q \in \mathbb{Q} \mid q > -1\} = \bigcup_{n \geq 1} \{q \in \mathbb{Q} \mid q > -0,\underbrace{999\ldots9}_{n}\}$$

von Teilmengen von $\mathbb{Q}$. Sie bedeutet nach 4.4.8 und der Beschreibung des Infimums im Dedekind'schen Modell der reellen Zahlen $\mathbb{R} = \mathbb{R}_{\mathrm{D}}$ genau die Gleichheit

$$-1_{\mathbb{R}} = \inf\{(-0,\underbrace{999\ldots9}_{n})_{\mathbb{R}} \mid n \geq 1\}$$

und mit unserer Konvention für die Interpretation unendlicher Dezimalbrüche 4.4.15 dann auch die Gleichheit von reellen Zahlen

$$-1 = -0,999\ldots$$

Jetzt gilt es nur noch, auf beiden Seiten das Negative zu nehmen. Ich bemerke weiter, daß es durchaus möglich ist, die Menge aller unendlichen Dezimalbrüche ohne alle Identifizierungen zu betrachten und mit der Struktur einer angeordneten Menge zu versehen, in der dann sogar jede nichtleere Teilmenge mit unterer Schranke eine größte untere Schranke hat. Es ist jedoch nicht möglich, die gewohnte Addition und Multiplikation von den endlichen Dezimalbrüchen so auf diese angeordnete Menge fortzusetzen, daß wir einen angeordneten Körper erhalten. Der naive Ansatz scheitert hier bereits daran, daß nicht klar ist, wie man mit den eventuell unendlich vielen Überträgen bei der Addition und Multiplikation umgehen soll.

Ergänzende Übung 4.4.18 (Für höhere Semester). Man zeige 4.4.3.2. Hinweis: Die gesuchte Bijektion φ kann zum Beispiel konstruiert werden durch die Vorschrift $\varphi(\alpha) = \inf\{q_{\mathbb{R}'} \mid q \in \mathbb{Q},\ q_{\mathbb{R}} > \alpha\}$. Man zeige allgemeiner auch, daß jeder

Körperhomomorphismus $\mathbb{R} \to \mathbb{R}$ die Identität ist. Hinweis: Nach 5.3.2 sind die nichtnegativen reellen Zahlen genau die Quadrate, jeder Körperhomomorphismus $\mathbb{R} \to \mathbb{R}$ erhält also die Anordnung. Andererseits aber muß er auf $\mathbb{Q}$ die Identität sein.

5 Folgen und Reihen

5.1 Konvergenz von Folgen

Definition 5.1.1. Eine Teilmenge einer Menge mit Ordnungsrelation heißt ein **Intervall** genau dann, wenn mit zwei beliebigen Punkten aus besagter Teilmenge auch jeder Punkt zwischen den beiden zu unserer Teilmenge gehört.

5.1.2. Ist $(X, \leq)$ eine Menge mit Ordnungsrelation, so heißt demnach in Formeln ausgedrückt eine Teilmenge $I \subset X$ ein Intervall oder genauer ein "Intervall in X" genau dann, wenn für beliebige $x, y, z \in X$ mit $x < y < z$ aus $x, z \in I$ folgt $y \in I$. Jeder Schnitt von Intervallen ist wieder ein Intervall.

Ergänzende Übung 5.1.3. Jede Teilmenge Y einer angeordneten Menge X ist die disjunkte Vereinigung aller maximalen in Y enthaltenen nichtleeren Intervalle von X. Hinweis: Hat ein System von Intervallen von X nichtleeren Schnitt, so ist auch seine Vereinigung wieder ein Intervall von X.

Definition 5.1.4. Wir erweitern die reellen Zahlen durch die zwei Punkte $-\infty$ und ∞ zu der in hoffentlich offensichtlicher Weise angeordneten Menge der sogenannten **erweiterten reellen Zahlen**

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$$

5.1.5. Jede Teilmenge von $\overline{\mathbb{R}}$ besitzt in $\overline{\mathbb{R}}$ ein Supremum und ein Infimum. Für ein Intervall $I \subset \overline{\mathbb{R}}$ mit Supremum $a = \sup I$ und Infimum $b = \inf I$ gibt es die Alternativen $a \in I$ oder $a \notin I$ und $b \in I$ oder $b \notin I$. Es gibt damit vier Typen von Intervallen in $\overline{\mathbb{R}}$, für die die beiden folgenden Notationen gebräuchlich sind:

$$\begin{array}{lclcl}
[a, b] & = & [a, b] & = & \{x \in \overline{\mathbb{R}} \mid a \leq x \leq b\} \\
]a, b[& = & (a, b) & = & \{x \in \overline{\mathbb{R}} \mid a < x < b\} \\
[a, b[& = & [a, b) & = & \{x \in \overline{\mathbb{R}} \mid a \leq x < b\} \\
]a, b] & = & (a, b] & = & \{x \in \overline{\mathbb{R}} \mid a < x \leq b\}
\end{array}$$

Wählen wir hier $a, b \in \overline{\mathbb{R}}$ beliebig mit $a < b$, so erhalten wir genau alle Intervalle in $\overline{\mathbb{R}}$ mit mehr als einem Element. Wir benutzen die eben erklärten Notationen jedoch auch im Fall $a \geq b$, sie bezeichnen dann manchmal eine einpunktige Menge und meist die leere Menge. Ich hoffe, daß der Leser aus dem Kontext erschließen kann, wann mit (a, b) ein Intervall gemeint ist und wann ein Paar aus $\mathbb{R}^2$. Ein Intervall in $\mathbb{R}$ nennen wir ein **reelles Intervall**.

5.1.6. Sind a und b konkrete Zahlen, etwa $a = 1$ und $b = 27$, so wäre zu allem Überfluß auch noch eine dritte Lesart von $(1, 27)$ als die in Klammern notierte Dezimalzahl 1,27 denkbar. Ich hoffe, daß der Leser aus dem Kontext erschließen kann, was jeweils gemeint ist. Wenn man genau hinguckt, sollte auch im letzteren Fall der Abstand nach dem Komma etwas kleiner sein.

5.1.7. Ein Intervall in $\overline{\mathbb{R}}$ heißt **kompakt** genau dann, wenn es eines unserer $[a, b]$ ist. Der Begriff "kompakt" wird in 6.4.1 auf beliebige Teilmengen von $\overline{\mathbb{R}}$ verallgemeinert. Ein reelles Intervall heißt **offen** genau dann, wenn es eines unserer Intervalle (a, b) ist. Der Begriff "offen" wird in 7.3.1 auf beliebige Teilmengen von $\overline{\mathbb{R}}$ verallgemeinert. Wir nennen ein reelles Intervall **halboffen** genau dann, wenn es nicht aus einem einzigen Punkt besteht, und verallgemeinern den Begriff "halboffen" in 7.1.1 auf beliebige Teilmengen von $\mathbb{R}$. In der in diesem Text verwendeten Terminologie sind mithin alle offenen Intervalle auch halboffen. In der Literatur wird der Begriff halboffen meist abweichend davon verwendet als Bezeichnung für reelle Intervalle, die weder offen noch kompakt sind.

Definition 5.1.8. Gegeben ein Punkt $x \in \overline{\mathbb{R}}$ vereinbaren wir nun, welche Teimengen von $\overline{\mathbb{R}}$ wir **Umgebungen von** x nennen wollen. Wir geben diese Definition separat für reelle Zahlen und für die beiden Punkte $\pm\infty$.

1. Gegeben $x \in \mathbb{R}$ heißt eine Teilmenge $W \subset \overline{\mathbb{R}}$ eine **Umgebung von** x genau dann, wenn sie ein Intervall (a, b) umfaßt mit $a < x < b$.

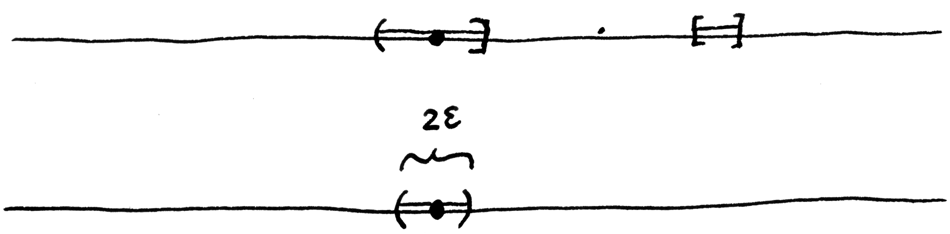
2. Für $x = \infty$ heißt eine Teilmenge $W \subset \overline{\mathbb{R}}$ eine **Umgebung von** x genau dann, wenn sie ein Intervall $(a, \infty]$ umfaßt mit $a < \infty$.

3. Für $x = -\infty$ heißt eine Teilmenge $W \subset \overline{\mathbb{R}}$ eine **Umgebung von** x genau dann, wenn sie ein Intervall $[-\infty, b)$ umfaßt mit $\infty < b$.

5.1.9. Gegeben $x \in \mathbb{R}$ und $\varepsilon > 0$ nennen wir das offene Intervall $(x-\varepsilon, x+\varepsilon)$ die ε-**Umgebung von** x. Eine Umgebung von $x \in \mathbb{R}$ können wir auch charakterisieren als eine Teilmenge $W \subset \overline{\mathbb{R}}$, die für mindestens ein reelles $\varepsilon > 0$ das Intervall $(x - \varepsilon, x + \varepsilon)$ umfaßt, oder äquivalent als eine Teilmenge, die für mindestens ein reelles $\varepsilon > 0$ das Intervall $[x - \varepsilon, x + \varepsilon]$ umfaßt. Eine der Motivationen für unsere großzügige Definition des Umgebungsbegriffs ist, daß er uns durch seine große Allgemeinheit dazu verhelfen soll, die Diskussion, ja die bloße Erwähnung derartiger Nebensächlichkeiten weitgehend zu vermeiden. Die Aufspaltung der Definition in drei Fälle ist natürlich nicht besonders befriedigend. Sie ermöglicht es uns jedoch im weiteren Verlauf, viele noch viel weiter gehende Fallunterscheidungen zu vermeiden.

Beispiel 5.1.10. Das kompakte Intervall $[0, 1]$ ist eine Umgebung von jedem Punkt aus dem offenen Intervall $(0, 1)$, aber von keinem anderen Punkt der erweiterten reellen Zahlengeraden. $\mathbb{Q}$ ist für keinen Punkt der erweiterten reellen Zahlengeraden eine Umgebung.

5.1.11. Offensichtlich besitzen je zwei verschiedene Punkte der erweiterten reellen Zahlengeraden zueinander disjunkte Umgebungen, und der Schnitt von je zwei Umgebungen ein- und desselben Punktes ist wieder eine Umgebung des besagten Punktes.



Versuch der graphischen Darstellung einer Umgebung sowie einer ε-Umgebung eines hier fett eingezeichneten Punktes der reellen Zahlengeraden. Die obere Umgebung besteht aus zwei halboffenen Intervallen und einem einzelnen Punkt.

Ergänzung 5.1.12. Bei der Vor- und Nachbereitung dieser Vorlesung ist mir erst richtig klar geworden, welch großer Teil der Diskussion der Begriffe Grenzwert und Stetigkeit im Rahmen der Analysis einer reellen Veränderlichen nur die Struktur der reellen Zahlen als angeordnete Menge betrifft. Die Resultate dieses Abschnitts mit Ausnahme der Beschreibung aller Intervalle 5.1.5 gelten im Übrigen mit unverändertem Beweis auch allgemeiner für einen beliebigen archimedisch angeordneten Körper und zu einem guten Teil sogar für einen beliebigen angeordneten Körper.

Definition 5.1.13. Eine Abbildung $\mathbb{N} \to X$, $n \mapsto x_n$ von den natürlichen Zahlen in eine Menge X nennen wir eine **Folge** in X. Wir schreiben eine Folge meist $(x_n)_{n\in\mathbb{N}}$ oder $x_0, x_1, x_2, \ldots$ oder auch einfach nur x_n. Die x_i heißen die **Folgenglieder**. Manchmal nennen wir allerdings auch Abbildungen Folgen, die erst ab $n = 1$ definiert sind.

Definition 5.1.14. Sagen wir, eine Aussage gelte für **fast alle Elemente** einer Menge, so soll das bedeuten, daß sie gilt für alle Elemente bis auf höchstens endlich viele Ausnahmen. Sagen wir, eine Aussage gelte für **fast alle Glieder** einer Folge, so soll das bedeuten, daß für fast alle Indizes n unsere Aussage für das n-te Folgenglied gilt.

Definition 5.1.15. Sei $x_0, x_1, \ldots$ eine Folge in $\overline{\mathbb{R}}$ und $x \in \overline{\mathbb{R}}$ ein Punkt. Wir sagen, **die Folge x_n konvergiere gegen** x genau dann, wenn jede Umgebung von x fast alle Glieder der Folge enthält. Wir schreiben in diesem Fall auch

$$\lim_{n\to\infty} x_n = x$$

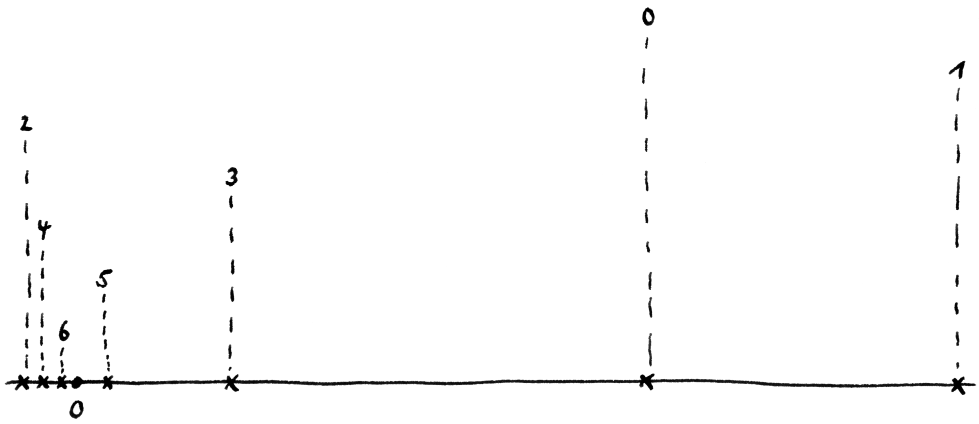
und nennen x einen **Grenzwert** oder lateinisierend **Limes** der Folge. Nach dem im folgenden bewiesenen Lemma 5.1.20 dürfen wir uns sogar den bestimmten Artikel erlauben und von *dem* Grenzwert reden.

Beispiel 5.1.16. Die **konstante Folge** $x_n = x\ \forall n$ konvergiert gegen x. In der Tat liegen bei dieser Folge in jeder Umgebung von x nicht nur fast alle, sondern sogar alle Folgenglieder.

Beispiel 5.1.17. Die Folge $x_n = n$ konvergiert gegen plus Unendlich, in Formeln

$$\lim_{n\to\infty} n = \infty$$

In der Tat umfaßt jede Umgebung U von ∞ per definitionem ein Intervall der Gestalt $(a, \infty]$ für $a \in \mathbb{R}$, und bereits jedem derartigen Intervall liegen nach 4.4.12 jeweils fast alle Folgenglieder.



Graphische Darstellung der Folge $x_n = 3^{3-n} - (-2)^{4-n}$, die gegen Null konvergiert, wie Sie bald werden zeigen können. Die Folgenglieder sind die kleinen Kreuzchen auf der reellen Achse, ihre Indizes tragen sie an unterschiedlich langen gestrichelt eingezeichneten Stangen.

Definition 5.1.18. Eine **Nullfolge** ist eine Folge, die gegen Null konvergiert.

Beispiel 5.1.19. Die Folge $x_n = 1/n$ ist eine Nullfolge, in Formeln

$$\lim_{n\to\infty} \frac{1}{n} = 0$$

In der Tat umfaßt jede Umgebung U von 0 per definitionem ein Intervall der Gestalt $(-\varepsilon, \varepsilon)$ für $\varepsilon > 0$, und bereits in jedem derartigen Intervall liegen alle Folgenglieder mit $n > (1/\varepsilon)$, nach 4.4.12 also jeweils fast alle Folgenglieder.

Lemma 5.1.20 (Eindeutigkeit des Grenzwerts). *Ein- und dieselbe Folge kann nicht gegen zwei verschiedene Punkte konvergieren, in Formeln*

$$(\lim_{n\to\infty} x_n = x \textit{ und } \lim_{n\to\infty} x_n = y) \Rightarrow (x = y)$$

Beweis. Durch Widerspruch. Sind unsere Punkte x und y verschieden, so besitzen sie auch disjunkte Umgebungen U und V. Dann können aber von unseren unendlich vielen Folgengliedern nicht fast alle in der Umgebung U von x und fast alle in der Umgebung V von y liegen, also kann unsere Folge nicht gleichzeitig gegen x und gegen y konvergieren. □

5.1.21. Man beachte, wie unser Umgebungsbegriff uns hier dabei hilft, den Beweis kurz und prägnant zu halten und die Diskussion von Sonderfällen für $\{x, y\} \cap \{-\infty, \infty\} \neq \emptyset$ zu vermeiden.

Definition 5.1.22. Unter einer **Umgebungsbasis** eines Punktes versteht man ein System alias eine Menge von Umgebungen besagten Punktes derart, daß jede Umgebung unseres Punktes mindestens eine Umgebung unseres Systems umfaßt.

Beispiele 5.1.23. Die ε-Umgebungen eines Punktes $x \in \mathbb{R}$ bilden eine Umgebungsbasis von x, desgleichen aber auch alle Intervalle $[x - 3\varepsilon, x + 4\varepsilon)$ mit $\varepsilon > 0$ oder alle Intervalle $[x - 1/n, x + 1/n]$ mit $n \in \mathbb{N}_{\geq 1}$. Eine Umgebungsbasis von ∞ bilden etwa die Intervalle $[K, \infty]$ mit $K \in \mathbb{R}$ oder auch mit $K \in \mathbb{N}$.

5.1.24. Um die Konvergenz einer Folge gegen einen Punkt nachzuweisen, müssen wir offensichtlich nur für alle Umgebungen aus einer Umgebungsbasis prüfen, daß fast alle Folgenglieder darin liegen. Konvergenz gegen einen Punkt $x \in \mathbb{R}$ etwa ist also gleichbedeutend dazu, daß für jedes $\varepsilon > 0$ die ε-Umgebung von x fast alle Glieder der Folge enthält, im Fall einer reellen Folge (x_n) also dazu, daß es

$$\text{für jedes } \varepsilon > 0 \text{ ein } N = N_\varepsilon \in \mathbb{N} \text{ gibt mit } n \geq N \Rightarrow |x_n - x| < \varepsilon.$$

Dahingegen ist $\lim_{n\to\infty} x_n = \infty$ gleichbedeutend dazu, daß für jedes $K \in \mathbb{N}$ fast alle Folgenglieder oberhalb von K liegen.

5.1.25. Wenn Sie in der Analysis die Formulierung "für alle $\varepsilon > 0$ gilt was auch immer" antreffen, so dürfen Sie erwarten, daß dieses "was auch immer", wenn es denn für ein gegebenes $\varepsilon > 0$ gilt, für alle größeren $\varepsilon > 0$ eh gilt. Salopp gesprochen besteht also die unausgesprochene Übereinkunft, durch die Verwendung des Buchstabens ε das anzudeuten, was man umgangssprachlich vielleicht mit "für jedes auch noch so kleine $\varepsilon > 0$" ausdrücken würde. Sie müssen nur einmal versuchen, beim Vorrechnen einer Übungsaufgabe statt ε den Buchstaben M zu verwenden: Auch wenn formal alles richtig sein sollte, wird Ihr Tutor deutlich länger darüber nachdenken müssen, ob Ihre Formulierung auch wirklich stimmt! "Sei $\varepsilon < 0$" schließlich ist ein mathematischer Witz.

5.1.26. Ich will versuchen, in der Vorlesung einem Farbencode zu folgen, nach dem vorgegebene Umgebungen von Grenzwerten und dergleichen in gelber Farbe dargestellt werden, dazu zu findende N und dergleichen dahingegen in roter Farbe.

5.1.27. Mit unserer Konvention für die "Konvergenz gegen $\pm\infty$" bewegen wir uns zwar im Rahmen des allgemeinen Begriffs der Konvergenz in "topologischen Räumen" 9.6.1, aber außerhalb der in der einführenden Literatur zur Analysis üblichen Konventionen, in denen die Terminologie **bestimmte Divergenz gegen** $\pm\infty$ verwendet wird. Üblicherweise bleibt in anderen Worten der Begriff der konvergenten Folge reserviert für Folgen, die gegen eine reelle Zahl konvergieren. Wir nennen solche Folgen **reell konvergent**. Falls eine Folge nicht konvergiert, auch nicht gegen ∞ oder $-\infty$, so nennt man sie **unbestimmt divergent**. Wir verlieren mit unserer Terminologie zwar etwas an Kohärenz, da wir im weiteren "Reihen" aus wieder anderen Gründen nur dann konvergent nennen werden, wenn die Folge ihrer Partialsummeen *reell* konvergent ist. Das schien mir jedoch ein kleineres Übel, als es eine unnötig einschränkende oder in Fälle aufspaltende Formulierung von Aussagen wie 5.1.31 oder 5.2.6 wäre.

Proposition 5.1.28. *Für jede Folge x_n von Null verschiedener reeller Zahlen gilt*

$$\lim_{n\to\infty} x_n = 0 \Leftrightarrow \lim_{n\to\infty} |x_n^{-1}| = \infty$$

Beweis. Für alle $K > 0$ gilt $|x_n^{-1}| \in (K, \infty] \Leftrightarrow x_n \in (-K^{-1}, K^{-1})$. Gilt also die rechte Seite bei vorgegebenem $K > 0$ für fast alle Folgenglieder, so auch die Linke. Ebenso gilt für alle $\varepsilon > 0$ offensichtlich $x_n \in (-\varepsilon, \varepsilon) \Leftrightarrow |x_n^{-1}| \in (\varepsilon^{-1}, \infty]$. Gilt also die rechte Seite bei vorgegebenem $\varepsilon > 0$ für fast alle Folgenglieder, so auch die Linke. □

5.1.29. Vereinbaren wir $1/|\infty| = 1/|-\infty| = 0$ und $|1/0| = \infty$, so gilt diese Proposition mit demselben Beweis sogar für jede Folge in $\overline{\mathbb{R}}$.

Proposition 5.1.30. *Die folgende Tabelle beschreibt das Konvergenzverhalten der Folge $(x^n)_{n\in\mathbb{N}}$ der Potenzen von x in Abhängigkeit von x:*

$$\begin{array}{rl} x > 1 & \lim_{n\to\infty} x^n = \infty; \\ x = 1 & \lim_{n\to\infty} x^n = 1; \\ |x| < 1 & \lim_{n\to\infty} x^n = 0; \\ x \le -1 & \textit{Die Folge } x^n \textit{ divergiert unbestimmt.} \end{array}$$

Beweis. Im Fall $x > 1$ schreiben wir $x = 1 + y$ mit $y > 0$ und erhalten mit der binomischen Formel

$$x^n = (1+y)^n \geq 1 + ny$$

Aber natürlich gilt $1 + ny \geq K$ genau dann, wenn gilt $n \geq (K-1)/y$, und das gilt bei festem K für fast alle n. Im Fall $x = 1$ ist die Folge konstant 1 und es ist nichts zu zeigen. Falls $0 < |x| < 1$ gilt nach dem Vorhergehenden $\lim_{n\to\infty} |1/x^n| = \infty$ und daraus folgt mit Proposition 5.1.28 $\lim_{n\to\infty} x^n = 0$. Für $x = 0$ gilt das natürlich eh. Im Fall $x \leq -1$ gilt $|x^n - x^{n+1}| \geq 2$ für alle n. Also kann die Folge nicht gegen eine reelle Zahl a konvergieren, denn dann müßte gelten $|a - x^n| < 1$ für fast alle n und dann nach der Dreiecksungleichung $|x^n - x^{n+1}| < 2$ für fast alle n. Die Folge kann in diesem Fall aber auch nicht gegen $\pm\infty$ konvergieren, da die Folgenglieder immer abwechselnd positiv und negativ sind. □

Lemma 5.1.31 (Quetschlemma). *Sind in $\overline{\mathbb{R}}$ drei Folgen a_n, b_n, c_n gegeben mit $a_n \leq b_n \leq c_n$ für alle n, und konvergieren a_n und c_n gegen denselben Grenzwert, so konvergiert auch b_n gegen diesen Grenzwert.*

Beweis. Das folgt aus den Definitionen mit der Erkenntnis, daß sich jede Umgebung eines Punktes verkleinern läßt zu einer Umgebung desselben Punktes, die ein Intervall ist. Es reicht ja, für jede solche "Intervallumgebung" I des gemeinsamen Grenzwerts von a_n und c_n zu zeigen, daß fast alle b_n darinliegen. Das ist aber klar, da fast alle a_n und fast alle c_n darinliegen. □

Beispiel 5.1.32. Konvergiert eine Folge reeller Zahlen a_n gegen ∞, so konvergiert jede Folge reeller Zahlen b_n mit $a_n \leq b_n$ für alle n auch gegen ∞. Das folgt zum Beispiel, indem wir als c_n die konstante Folge ∞ nehmen und das Quetschlemma anwenden.

Lemma 5.1.33 (Erhaltung von Ungleichungen). *Seien a_n, b_n Folgen in $\overline{\mathbb{R}}$ mit Grenzwerten $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Gilt $a_n \leq b_n$ für alle n, so folgt $a \leq b$.*

Beweis. Wäre hier $b < a$, so fänden wir k mit $b < k < a$. Dann wäre $[-\infty, k)$ eine Umgebung von b und $(k, \infty]$ eine Umgebung von a. Fast alle a_n müssten also in $(k, \infty]$ liegen und fast alle b_n in $[-\infty, k)$ und es folgte $a_n > b_n$ für fast alle n im Widerspruch zu unserer Annahme. □

Satz 5.1.34 (Rechenregeln für Grenzwerte). *Seien a_n, b_n Folgen reeller Zahlen mit reellen Grenzwerten $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$.*

1. *Die Summe bzw. das Produkt unserer Folgen konvergieren gegen die Summe bzw. das Produkt ihrer Grenzwerte, in Formeln*

$$\begin{array}{lcl} \lim_{n\to\infty}(a_n + b_n) & = & a + b \\ \lim_{n\to\infty}(a_n b_n) & = & ab \end{array}$$

2. *Sind alle Glieder der Folge b_n sowie ihr Grenzwert b von Null verschieden, so gilt für die Folge der Kehrwerte*

$$\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b}$$

Beweis. Wir beginnen den Beweis mit einem Lemma.

Lemma 5.1.35. 1. *Gegeben $a, b \in \mathbb{R}$ und eine Umgebung W von $a + b$ gibt es Umgebungen U von a und V von b mit $U + V \subset W$.*

2. *Gegeben $a, b \in \mathbb{R}$ und eine Umgebung W von $a \cdot b$ gibt es Umgebungen U von a und V von b mit $U \cdot V \subset W$.*

3. *Gegeben $b \in \mathbb{R}^\times$ und eine Umgebung W von b^{-1} gibt es eine Umgebung $V \subset \mathbb{R}^\times$ von b mit $V^{-1} \subset W$.*

5.1.36. In Erinnerung an 3.1.2.6 verstehen wir hier $U + V = \{x + y \mid x \in U,\ y \in V\}$ und $U \cdot V = \{x \cdot y \mid x \in U,\ y \in V\}$. In Anlehnung an 2.2.8 verstehen wir weiter $V^{-1} = \{x^{-1} \mid x \in V\}$.

Beweis des Lemmas. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, daß W sogar eine ε-Umgebung von $a + b$ ist. Nehmen wir dann für U bzw. V die $\varepsilon/2$-Umgebung von a bzw. b, so gilt in der Tat $U + V \subset W$. Für die zweite Formel beginnen wir mit der Abschätzung

$$\begin{array}{lcl} |xy - ab| & = & |(x - a)y + a(y - b)| \\ & \leq & |x - a||y| + |a||y - b| \end{array}$$

Aus den beiden Ungleichungen $|x - a| < \eta$ und $|y - b| < \eta$ folgt zunächst $|y| < |b| + \eta$ und dann

$$|xy - ab| \leq \eta(|b| + \eta + |a|)$$

Ohne Beschränkung der Allgemeinheit dürfen wir nun wieder annehmen, daß W eine ε-Umgebung von $a \cdot b$ ist. Wählen wir dann ein $\eta \in (0,1)$ mit $\varepsilon > \eta(|b|+1+|a|)$ und nehmen als U bzw. V die η-Umgebungen von a bzw. b, so gilt folglich in der Tat $U \cdot V \subset W$. Um die letzte Aussage zu zeigen, nehmen wir der Einfachkeit halber $b > 0$ an. Ohne Beschränkung der Allgemeinheit dürfen wir nun $W = (a, d)$ mit $0 < a < b < d < \infty$ annehmen, und dann betrachten wir schlicht $V = (d^{-1}, a^{-1})$ und sind fertig. □

Jetzt zeigen wir den Satz. Wir müssen für jede Umgebung W von $a+b$ zeigen, daß fast alle Glieder der Folge $a_n + b_n$ darinliegen. Nach dem vorhergehenden Lemma 5.1.35 finden wir jedoch Umgebungen U von a und V von b mit $U + V \subset W$. Da nach Annahme fast alle Glieder der ersten Folge in U liegen und fast alle Glieder der zweiten Folge in V, liegen damit in der Tat fast alle Glieder der Folge $a_n + b_n$ in W. Ganz genauso folgt aus den anderen Teilen von Lemma 5.1.35, daß der Grenzwert des Produktes zweier Folgen das Produkt der Grenzwerte ist, und ähnlich aber einfacher, daß der Grenzwert der Kehrwerte der Kehrtwert des Grenzwerts ist. □

Beispiel 5.1.37.

$$\lim_{n\to\infty} \frac{5n^3+n}{3n^3+n^2} = \lim_{n\to\infty} \frac{5+(1/n)^2}{3+(1/n)} = \frac{5+0^2}{3+0} = \frac{5}{3}$$

5.1.38. Die Addition und die Multiplikation $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ lassen sich nicht so zu Abbildungen von ganz $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ nach $\overline{\mathbb{R}}$ fortsetzen, daß die ersten beiden Teile von 5.1.35 entsprechend gelten. Alle derartigen Fortsetzungen auf Teilmengen von $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ stimmen jedoch auf dem Schnitt der jeweiligen Definitionsbereiche überein, so daß es sowohl für die Addition als auch für die Multiplikation jeweils eine größtmögliche "sinnvolle Fortsetzung" gibt, die wir im Rahmen der Topologie **??** als die größtmögliche "stetige Fortsetzung" werden verstehen können. Wir beschreiben diese Fortsetzungen von Addition und Multiplikation durch Abbildungen $+, \cdot : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \overline{\mathbb{R}} \cup \{*\}$ mit einem eigenen Symbol $*$ für "nicht sinnvoll in $\overline{\mathbb{R}}$ zu definieren". Unsere Fortsetzungen werden mit dieser Konvention gegeben durch die Formeln

$$\begin{array}{rcrcrl}
a+\infty & = & \infty+a & = & \infty & \forall a \in \mathbb{R} \cup \{\infty\} \\
a+(-\infty) & = & -\infty+a & = & -\infty & \forall a \in \mathbb{R} \cup \{-\infty\} \\
\infty+(-\infty) & = & -\infty+\infty & = & * &
\end{array}$$

und

$$\begin{array}{rclcll} a\infty & = & \infty a & = & \infty & \forall a \in \overline{\mathbb{R}}, a > 0 \\ a\infty & = & \infty a & = & -\infty & \forall a \in \overline{\mathbb{R}}, a < 0 \\ a(-\infty) & = & (-\infty)a & = & -\infty & \forall a \in \overline{\mathbb{R}}, a > 0 \\ a(-\infty) & = & (-\infty)a & = & \infty & \forall a \in \overline{\mathbb{R}}, a < 0 \\ 0\infty & = & \infty 0 & = & * & \\ 0(-\infty) & = & (-\infty)0 & = & * & \end{array}$$

Übung 5.1.39. Man zeige: Die Regeln 5.1.34.1 zum Vertauschen von Grenzwertbildung mit Addition und Multiplikation gelten auch noch, wenn wir $a, b \in \overline{\mathbb{R}}$ zulassen und $a + b$ beziehungsweise $a \cdot b$ sinnvoll definiert sind im Sinne der vorhergehenden Bemerkung 5.1.38.

Übung 5.1.40. Für jedes $x \in \overline{\mathbb{R}}$ gibt es eine absteigende Folge von Umgebungen $U_0 \supset U_1 \supset U_2 \supset \dots$ derart, daß jede Umgebung von x fast alle der U_n umfaßt.

Übung 5.1.41. Ist $A \subset \overline{\mathbb{R}}$ eine nichtleere Teilmenge, so ist $\sup A$ das größte Element in der Menge G aller Punkte aus den erweiterten reellen Zahlen, die Grenzwerte von Folgen aus A sind.

Übung 5.1.42. Aus $\lim_{n\to\infty} a_n = a$ folgt $\lim_{n\to\infty} |a_n| = |a|$. Umgekehrt folgt aus $\lim_{n\to\infty} |a_n| = 0$ bereits $\lim_{n\to\infty} a_n = 0$.

Übung 5.1.43. Ist (a_n) eine Folge reeller Zahlen, die gegen eine reelle Zahl konvergiert, so gilt $\lim_{n\to\infty}(a_{n+1} - a_n) = 0$.

5.2 Vollständigkeit der reellen Zahlen

Definition 5.2.1. Eine Menge von reellen Zahlen heißt **beschränkt** genau dann, wenn sie in $\mathbb{R}$ eine obere und eine untere Schranke besitzt. Eine Folge reeller Zahlen heißt beschränkt genau dann, wenn die Menge der Folgenglieder beschränkt ist.

Lemma 5.2.2. *Jede reell konvergente Folge von reellen Zahlen ist beschränkt.*

Beweis. Ist $x \in \mathbb{R}$ der Grenzwert unserer Folge, so liegen fast alle Folgenglieder in $[x - 1, x + 1]$. Die endlich vielen Ausnahmen können wir durch hinreichend große Schranken auch noch einfangen. □

Definition 5.2.3. Eine Folge x_n in einer Menge mit Ordnungsrelation heißt
monoton wachsend genau dann, wenn gilt $x_0 \leq x_1 \leq x_2 \leq \dots$
streng monoton wachsend genau dann, wenn gilt $x_0 < x_1 < x_2 < \dots$
monoton fallend genau dann, wenn gilt $x_0 \geq x_1 \geq x_2 \geq \dots$
streng monoton fallend genau dann, wenn gilt $x_0 > x_1 > x_2 > \dots$

Eine Folge heißt **monoton** genau dann, wenn sie monoton wächst oder monoton fällt. Eine Folge heißt **streng monoton** genau dann, wenn sie streng monoton wächst oder streng monoton fällt. Diese Begriffe werden in 6.2.1 auf Abbildungen zwischen beliebigen angeordneten Mengen verallgemeinert.

Lemma 5.2.4. *Jede monoton wachsende Folge in $\overline{\mathbb{R}}$ konvergiert gegen das Supremum der Menge ihrer Folgenglieder. Jede monoton fallende Folge in $\overline{\mathbb{R}}$ konvergiert gegen das Infimum der Menge ihrer Folgenglieder. Jede monotone beschränkte Folge von reellen Zahlen konvergiert gegen eine reelle Zahl.*

Beweis. Wir zeigen nur die erste Aussage, die Zweite zeigt man analog und die Dritte ist eine offensichtliche Konsequenz. Sei in der Tat s das Supremum alias die kleinste obere Schranke der Menge aller Folgenglieder. Kein p mit $p < s$ ist dann eine obere Schranke der Menge aller Folgenglieder, folglich liegen für jedes $p < s$ ein und damit wegen der Monotonie fast alle x_n in $(p, s]$. Damit liegen in jeder Umgebung von s fast alle Folgenglieder. □

Definition 5.2.5. Sei $x_0, x_1, x_2, \ldots$ eine Folge. Sind $0 \leq n_0 < n_1 < n_2 < \ldots$ natürliche Zahlen, so nennen wir die Folge $(x_{n_k})_{k \in \mathbb{N}}$ mit Gliedern

$$x_{n_0}, x_{n_1}, x_{n_2}, \ldots$$

eine **Teilfolge** der Folge x_n. Schreiben wir eine Folge als eine Abbildung $x : \mathbb{N} \to X$, $x \mapsto x(n) = x_n$, so ist eine Teilfolge von x demnach eine Abbildung der Gestalt $x \circ f$ für eine streng monoton wachsende Folge $f : \mathbb{N} \to \mathbb{N}$.

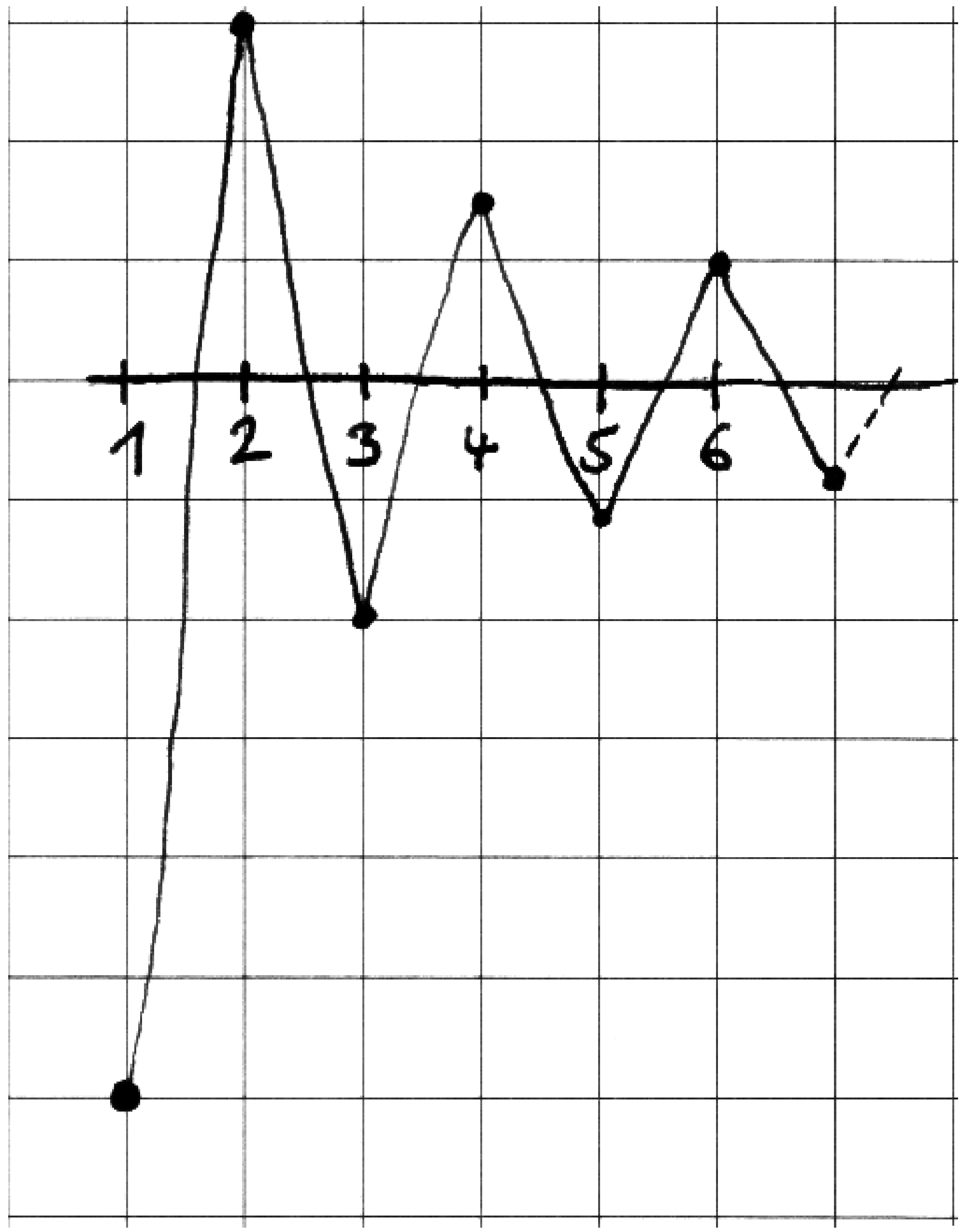
Lemma 5.2.6. *Jede Teilfolge einer konvergenten Folge konvergiert, und zwar gegen denselben Grenzwert wie die ursprüngliche Folge.*

Beweis. Das ist klar nach den Definitionen. □

Lemma 5.2.7. *Jede Folge in einer angeordneten Menge besitzt eine monotone Teilfolge.*

Bemerkung 5.2.8. Hier mögen Sie sich an unsere Sprachregelung 2.3.2 erinnern. Gemeint ist demnach: Jede Folge in einer angeordneten Menge besitzt mindestens eine monotone Teilfolge.

Beweis. Wir nennen ein Folgenglied x_n oder präziser seinen Index n einen "Aussichtspunkt" der Folge genau dann, wenn alle späteren Folgenglieder kleiner sind, in Formeln $x_n > x_m$ für alle $m > n$. Besitzt unsere Folge unendlich viele Aussichtspunkte, so bilden diese eine streng monoton fallende Teilfolge. Sonst gibt es einen letzten Aussichtspunkt x_n. Dann finden wir aber eine monoton wachsende Teilfolge, die mit x_{n+1} beginnt, denn ab dem Index $n+1$ kommt dann nach jedem Folgenglied noch ein anderes, das mindestens ebenso groß ist. □



Bei der Folge $(-1)^n 6/n$ ist jedes zweite Folgenglied ein Ausichtspunkt im Sinne des Beweises von Lemma 5.2.7.

Satz 5.2.9 (Bolzano-Weierstraß). *Jede Folge in $\overline{\mathbb{R}}$ besitzt eine in $\overline{\mathbb{R}}$ konvergente Teilfolge. Jede beschränkte Folge von reellen Zahlen besitzt eine reell konvergente Teilfolge.*

Beweis. Jede Folge in $\overline{\mathbb{R}}$ besitzt nach 5.2.7 eine monotone Teilfolge, und diese ist nach 5.2.4 konvergent in $\overline{\mathbb{R}}$. Ist unsere Folge beschränkt, so ist auch jede solche Teilfolge beschränkt und konvergiert folglich gegen eine reelle Zahl. □

Definition 5.2.10. Eine Folge $(x_n)_{n\in\mathbb{N}}$ von reellen Zahlen heißt eine **Cauchy-Folge** genau dann, wenn es für jedes $\varepsilon > 0$ ein $N = N_\varepsilon \in \mathbb{N}$ gibt derart, daß gilt $|x_n - x_m| < \varepsilon$ falls $n, m \geq N$. Analog erklärt man Cauchy-Folgen in beliebigen angeordneten Körpern.

Satz 5.2.11. *Eine Folge reeller Zahlen konvergiert gegen eine reelle Zahl genau dann, wenn sie eine Cauchy-Folge ist.*

Beweis. Daß jede reell konvergente Folge Cauchy sein muß, ist leicht zu sehen: Aus $\lim_{n\to\infty} x_n = x$ folgt, daß es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt mit $|x_n - x| < \varepsilon/2$ für $n \geq N$. Daraus folgt dann $|x_n - x_m| < \varepsilon$ für $n, m \geq N$. Wir zeigen nun umgekehrt, daß auch jede Cauchy-Folge gegen eine reelle Zahl konvergiert. Eine Cauchy-Folge x_n ist sicher beschränkt, denn wählen wir für $\varepsilon = 1$ ein $N = N_\varepsilon$, so liegen fast alle Folgenglieder im Intervall $(x_N - 1, x_N + 1)$, und die endlich vielen Ausnahmen können wir durch eine hinreichend große Schranke auch noch einfangen. Unsere Cauchy-Folge besitzt daher nach 5.2.9 eine reell konvergente Teilfolge x_{n_k}, sagen wir $\lim_{k\to\infty} x_{n_k} = x$. Wir behaupten, daß dann auch die Folge x_n selbst gegen x konvergiert. In der Tat gibt es für alle $\varepsilon > 0$ ein N_ε mit $n, m \geq N_\varepsilon \Rightarrow |x_n - x_m| < \varepsilon$. Aus $n \geq N_\varepsilon$ folgt damit insbesondere $|x_n - x_{n_k}| < \varepsilon$ für fast alle k und dann im Grenzwert $|x_n - x| \leq \varepsilon$, da ja die Ungleichungen $-\varepsilon \leq x_n - x_{n_k} \leq \varepsilon$ bestehen bleiben beim Grenzübergang $k \to \infty$. □

5.2.12. Ein angeordneter Körper, in dem jede Cauchy-Folge konvergiert, heißt **vollständig**. Dieser Begriff ist Teil einer alternativen Charakterisierung der reellen Zahlen, die wir hier als Übung formulieren.

Übung 5.2.13. Gegeben ein angeordneter Körper sind gleichbedeutend: (1) Jede nichtleere Teilmenge mit einer unteren Schranke besitzt eine größte untere Schranke und (2) Der Körper ist archimedisch angeordnet und vollständig.

Übung 5.2.14 (**Intervallschachtelungsprinzip**). Gegeben eine absteigende Folge von nichtleeren kompakten Intervallen $I_0 \supset I_1 \supset I_2 \ldots$ ist auch ihr Schnitt $\bigcap_{\nu\in\mathbb{N}} I_\nu$ nicht leer.

Übung 5.2.15. Konvergiert eine Teilfolge einer Cauchyfolge, so konvergiert bereits die ganze Cauchyfolge, und zwar gegen denselben Grenzwert.

5.3 Vergleich von $\mathbb{Q}$ und $\mathbb{R}$

5.3.1. Wir erinnern daran, daß es nach 4.1.1 keine rationale Zahl mit Quadrat Zwei gibt. Im Gegensatz dazu zeigen wir nun, daß es in den reellen Zahlen zu jeder nichtnegativen reellen Zahl eine Quadratwurzel gibt.

Satz 5.3.2. *Für jede nichtnegative reelle Zahl $a \geq 0$ gibt es genau eine nichtnegative reelle Zahl $x \geq 0$ mit $x^2 = a$. Man bezeichnet dies x mit $\sqrt{a}$ und nennt es die* **Wurzel** *oder genauer* **Quadratwurzel** *von a.*

Beweis. Haben wir einmal eine Lösung $X = b$ der Gleichung $X^2 - a = 0$ gefunden, so gilt $X^2 - a = (X + b)(X - b)$. Folglich ist $X = -b$ dann die einzige andere Lösung, denn ein Produkt in einem Körper ist per definitionem nur dann Null, wenn einer der Faktoren Null ist. Die Eindeutigkeit der nichtnegativen Lösung ist damit klar und nur die Existenz muß noch gezeigt werden. Wir konstruieren dazu eine Folge und beginnen mit $x_0 = \max(1, a)$. Dann gilt sicherlich schon einmal $x_0^2 \geq a$. Gegeben $x_n > 0$ mit $x_n^2 \geq a$ machen wir den Ansatz $(x_n - \varepsilon)^2 = a$ und erhalten $\varepsilon = (x_n^2 + \varepsilon^2 - a)/2x_n$. Nun vernachlässigen wir $\varepsilon^2/2x_n$, vergessen unseren Ansatz und setzen

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{x_n^2 + a}{2x_n}$$

Man sieht sofort, daß aus $x_n^2 \geq a$ und $x_n > 0$ folgt $x_{n+1}^2 \geq a$ und $x_n \geq x_{n+1} > 0$. Da die Folge der x_n monoton fällt und durch Null nach unten beschränkt ist, besitzt sie einen Grenzwert $x \geq 0$. Aus der Gleichung

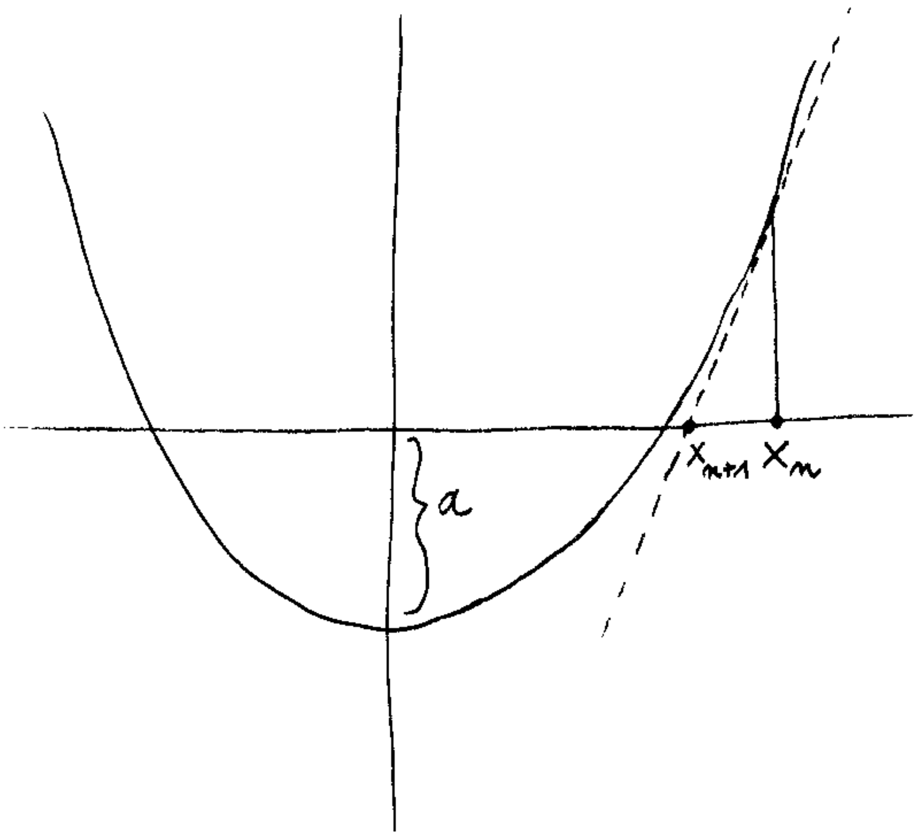
$$2x_n x_{n+1} = x_n^2 + a$$

folgt dann $x^2 = a$ durch Übergang zum Grenzwert für $n \to \infty$ auf beiden Seiten. □

5.3.3. Die Ungleichungen $a/x_n \leq \sqrt{a} < x_n$ erlauben uns sogar abzuschätzen, wie gut unsere Approximation x_n mindestens sein muß. Machen wir für den Fehler den Ansatz $x_n = \sqrt{a}(1 + f_n)$, so ergibt sich mit etwas Rechnen

$$f_{n+1} = \frac{f_n^2}{2(1 + f_n)}$$

und indem wir im Nenner die 1 beziehungsweise f_n verkleinern zu Null erhalten wir die Abschätzung $f_{n+1} \leq \frac{1}{2}\min(f_n, f_n^2)$. Sobald also x_n so nah bei $\sqrt{a}$ ist, daß gilt $f_n < 1$, "verdoppelt sich die Anzahl der richtigen Stellen beim Übergang von x_n zu x_{n+1}". Man spricht unter diesen Umständen auch von **quadratischer**



Graphische Darstellung unserer induktiven Formel für die Glieder der Folge x_n mit Grenzwert $\sqrt{a}$ aus dem Beweis von 5.3.2

Konvergenz. Anschaulich erhält man x_{n+1}, indem man von x_n senkrecht hochgeht zum Graph der Funktion $y = x^2 - a$ und dann auf der Tangente an diesen Graphen wieder herunter auf die x-Achse. Es ist damit auch anschaulich klar, daß unser Verfahren sehr schnell konvergieren sollte. Dieses Verfahren kann auch zur Bestimmung der Nullstellen allgemeinerer Funktionen anwenden. Es heißt das **Newton-Verfahren**.

Übung 5.3.4. Wir erinnern die Fibonacci-Folge und den goldenen Schnitt aus 1.2.1. Man zeige, daß der Quotient zweier aufeinanderfolgender Fibonacci-Zahlen gegen den goldenen Schnitt strebt, daß also in Formeln gilt

$$\lim_{\overrightarrow{i\to\infty}} \frac{x_{i+1}}{x_i} = \frac{1+\sqrt{5}}{2}$$

für unsere Fibonacci-Folge $x_0, x_1, x_2, \ldots$ aus 1.2.1.

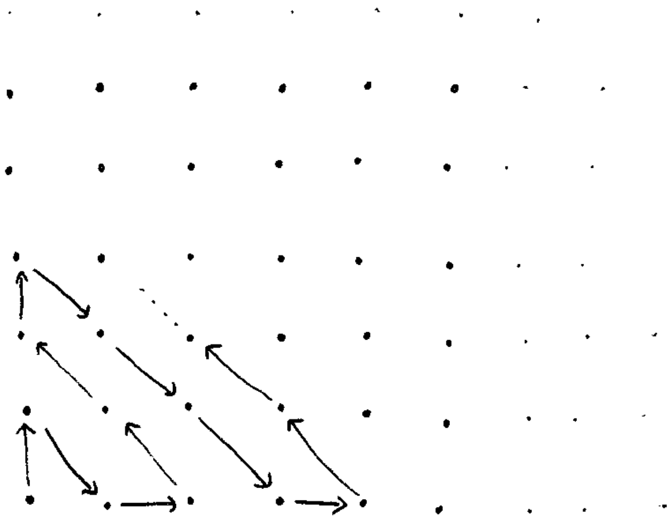
Definition 5.3.5. Eine Menge heißt **abzählbar** genau dann, wenn es eine Bijektion unserer Menge mit einer Teilmenge der Menge $\mathbb{N}$ aller natürlichen Zahlen gibt. Gleichbedeutend können wir auch fordern, daß unsere Menge entweder leer ist oder es eine Surjektion von $\mathbb{N}$ darauf gibt. Eine Menge heißt **abzählbar unendlich** genau dann, wenn sie abzählbar aber nicht endlich ist. Eine Menge heißt **überabzählbar** genau dann, wenn sie nicht abzählbar ist.

Satz 5.3.6. *1. Es gibt eine Bijektion $\mathbb{N} \xrightarrow{\sim} \mathbb{Q}$, in Worten: Die Menge der rationalen Zahlen ist* **abzählbar unendlich**.

2. Es gibt keine Surjektion $\mathbb{N} \twoheadrightarrow \mathbb{R}$, in Worten: Die Menge der reellen Zahlen ist **überabzählbar**.

Beweis. 1. Für jede natürliche Zahl N gibt es nur endlich viele Brüche $p/q \in \mathbb{Q}$ mit $p, q \in \mathbb{Z}$, $q \neq 0$ und $|p| \leq N$, $|q| \leq N$. Wir beginnen unser Abzählen von $\mathbb{Q}$ mit den Brüchen für $N = 1$, dann nehmen wir die Brüche hinzu mit $N = 2$, und indem wir so weitermachen zählen wir ganz $\mathbb{Q}$ ab.

2. Hierzu verwenden wir das **Cantor'sche Diagonalverfahren**. Man beachte zunächst, daß ein unendlicher Dezimalbruch, in dem die Ziffern Null und Neun nicht vorkommen, nur dann dieselbe reelle Zahl darstellt wie ein beliebiger anderer unendlicher Dezimalbruch, wenn die beiden in jeder Stelle übereinstimmen. Wir nehmen nun eine beliebige Abbildung $\mathbb{N} \to \mathbb{R}$, $i \mapsto r_i$, und zeigen, daß sie keine Surjektion sein kann. Wir schreiben dazu jedes r_i als unendlichen Dezimalbruch. Dann finden wir einen unendlichen Dezimalbruch r, bei dem die Ziffern Null und Neun nicht vorkommen und so, daß r an der i-ten Stelle nach dem Komma verschieden ist von r_i und an der ersten Stelle vor dem Komma von r_0. Dies r ist dann verschieden von allen r_i und unsere Abbildung $i \mapsto r_i$ kann keine Surjektion gewesen sein. □



$\mathbb{N} \times \mathbb{N}$ ist abzählbar und damit natürlich auch allgemeiner das Produkt von je zwei und dann auch von endlich vielen abzählbaren Mengen.

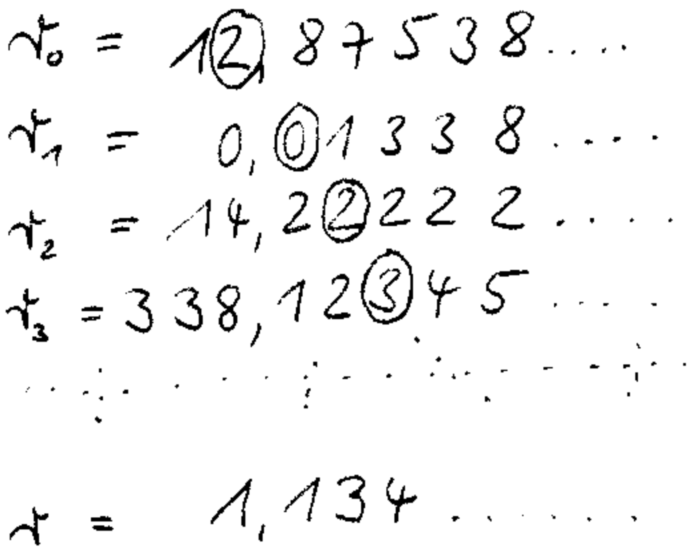


Illustration zum Cantor'schen Diagonalverfahren. Ähnlich zeigt man, daß die Menge $\mathrm{Ens}(\mathbb{N}, E)$ aller Abbildungen von $\mathbb{N}$ in eine Menge E mit mindestens zwei Elementen nicht abzählbar ist.

Bemerkung 5.3.7. Man kann sich fragen, ob jede Teilmenge der reellen Zahlen entweder abzählbar ist oder in Bijektion zu den reellen Zahlen selber. Die schon auf Cantor zurückgehende Vermutung, das könnte gelten, heißt die **Kontinuumshypothese**. Sie wurde 1963 von Paul Cohen in sehr merkwürdiger Weise geklärt: Er zeigte, daß unsere Frage in dem axiomatischen Rahmen, in dem man die Mengenlehre üblicherweise formalisiert, nicht entscheidbar ist. Cohen wurde für diese Leistung auf dem internationalen Mathematikerkongress 1966 mit der Fields-Medaille ausgezeichnet. Die Kontinuumshypothese ist übrigends die erste Frage einer berühmten Liste von 23 Fragestellungen, den sogenannten **Hilbert'schen Problemen**, die David Hilbert in seiner Ansprache auf dem internationalen Mathematikerkongress 1900 in Paris vorstellte in seinem Bemühen, "aus verschiedenen mathematischen Disziplinen einzelne bestimmte Probleme zu nennen, von deren Behandlung eine Förderung der Wissenschaft sich erwarten läßt".

5.4 Die Kreiszahl π

5.4.1. Bekanntlich bezeichnet π, ein kleines griechisches P für Perimeter, das Verhältnis zwischen Umfang und Durchmesser eines Kreises. Um diese Anschauung zu formalisieren zur Definition einer reellen Zahl im Sinne von 4.4.9 gehen wir aus von der anschaulichen Bedeutung von π als Länge des Halbkreises H mit Radius Eins

$$H = \{(a, b) \in \mathbb{R}^2 \mid a^2 + b^2 = 1,\ b \geq 0\}$$

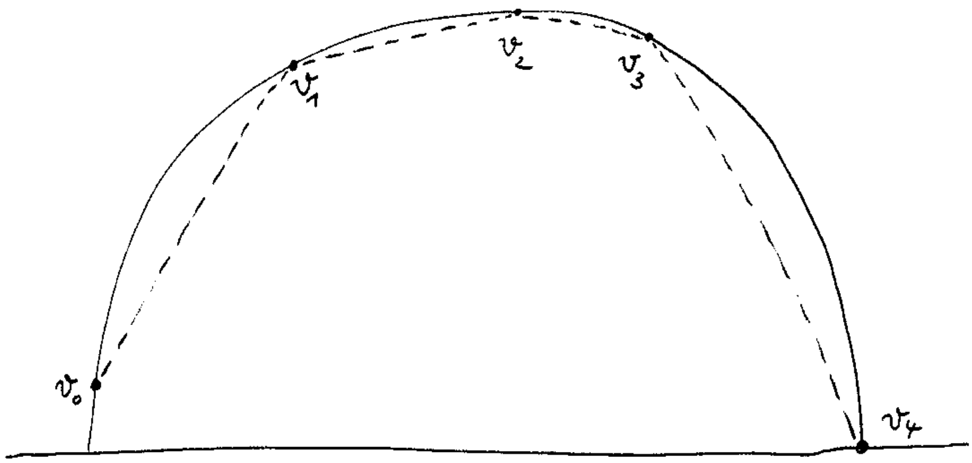
Seien $x, y : \mathbb{R}^2 \to \mathbb{R}$ die beiden Abbildungen, die jedem Punkt der Ebene seine erste bzw. zweite Koordinate zuordnen, also $v = (x(v), y(v)) \quad \forall v \in \mathbb{R}^2$. Die Distanz $d(v, w) \in \mathbb{R}$ zwischen zwei Punkten $v, w \in \mathbb{R}^2$ der Ebene erklären wir in Erinnerung an den Satz des Pythagoras durch die Formel

$$d(v, w) := \sqrt{(x(v) - x(w))^2 + (y(v) - y(w))^2}$$

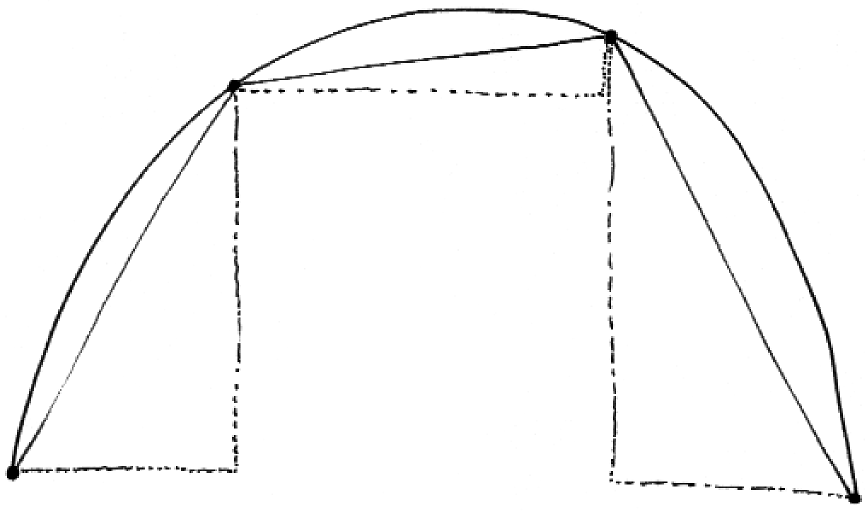
Die **Kreiszahl** $\pi \in \mathbb{R}$ definieren wir dann als das Supremum über die "Längen aller in unseren Halbkreis H einbeschriebenen Polygonzüge", in Formeln

$$\pi := \sup \left\{ \sum_{i=1}^{n} d(v_{i-1}, v_i) \;\middle|\; \begin{array}{l} n \in \mathbb{N},\ v_0, v_1, \ldots, v_n \in H, \\ x(v_0) < x(v_1) < \ldots < x(v_n) \end{array} \right\}$$

Mithilfe der Abschätzung $\sqrt{a^2 + b^2} \leq |a| + |b|$ erkennt man, daß die Zahl 4 eine obere Schranke ist für unsere Menge von Längen von Polygonzügen, mithin haben wir hier in der Tat eine reelle Zahl $\pi \in \mathbb{R}$ definiert. Wir werden in 10.6.19 sehen, wie man diese Zahl im Prinzip bis zu einer beliebig vorgegebenen Stelle nach dem Komma berechnen kann. Die Definition selbst ist sehr einfach. Ich habe sie nur deshalb nicht gleich im Zusammenhang mit der Definition der reellen



Ein einbeschriebener Polygonzug



Diese Abbildung soll veranschaulichen, warum 4 eine obere Schranke für die Längen einbeschriebener Polygonzüge ist: Die horizontalen Stücke und die vertikalen Stücke haben jeweils zusammengenommen eine Gesamtlänge ≤ 2.

Zahlen gegeben, weil sie die Existenz von Quadratwurzeln benötigt, die erst in 5.3.2 gezeigt wurde.

Ergänzung 5.4.2. Die Zahl π ist nicht rational, in Formeln $\pi \notin \mathbb{Q}$, wie Lambert bereits 1766 zeigen konnte. Anders ausgedrückt läßt sich π nicht durch einen periodischen Dezimalbruch darstellen. Wir geben einen Beweis in 10.6.7. Unsere Kreiszahl π ist noch nicht einmal **algebraisch**, als da heißt Nullstelle eines nichttrivialen Polynoms mit rationalen Koeffizienten, d.h. es gilt keine Gleichung der Gestalt

$$\pi^n + q_{n-1}\pi^{n-1} + \ldots + q_1\pi + q_0 = 0 \quad \text{mit } q_{n-1}, \ldots, q_0 \in \mathbb{Q} \text{ und } n \geq 1.$$

Reelle Zahlen, die nicht algebraisch sind, heißen **transzendent**, lateinisch für "überschreitend", da ihre Behandlung "die Grenzen der Algebra überschreitet". Die Transzendenz von π wurde 1882 von Lindemann in Freiburg bewiesen. Seine Büste steht im vierten Stock des Mathematischen Instituts. Er war übrigends Hilbert's Doktorvater.

5.5 Grenzwerte von Reihen

Definition 5.5.1. Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge reeller Zahlen. Der Ausdruck

$$\sum_{k=0}^{\infty} a_k$$

bezeichnet die Folge der **Partialsummen** $s_n = \sum_{k=0}^{n} a_k$ und, falls die Folge dieser Partialsummen konvergiert, auch ihren Grenzwert $\lim_{n\to\infty} s_n = s$. Wir sagen dann, die **Reihe** $\sum_{k=0}^{\infty} a_k$ **konvergiere gegen** s. Nennen wir eine Reihe **konvergent**, so meinen wir stets, daß unsere Reihe gegen eine reelle Zahl konvergiert und nicht etwa gegen $\pm\infty$. Die a_k heißen die **Reihenglieder**.

5.5.2. Es spielt für das Konvergenzverhalten einer Reihe keine Rolle, wenn wir *endlich viele* ihrer Glieder abändern. Das beinflußt nur den Grenzwert und ändert ihn eben um die Summe unserer endlich vielen Änderungen.

Bemerkung 5.5.3. Es wäre terminologisch kohärenter gewesen, wie bei Folgen auch bei Reihen von "reell konvergenten Reihen" zu sprechen. Das schien mir jedoch ungeschickt, da man den Begriff dann nicht als Verb verwenden kann: "Die Reihe reell-konvergiert" klingt einfach zu holprig, und Sprechweisen wie "die Reihe konvergiert absolut" sind oft praktisch.

Beispiel 5.5.4. Dies Beispiel illustriert den oft nürtzlichen **Teleskopsummentrick**:

$$\begin{aligned}
\sum_{k=1}^{\infty} \tfrac{1}{k(k+1)} &= \lim_{n\to\infty} \sum_{k=1}^{n} \left(\tfrac{1}{k} - \tfrac{1}{k+1}\right) \\
&= \lim_{n\to\infty} \left(1 - \tfrac{1}{n+1}\right) \\
&= 1
\end{aligned}$$

Satz 5.5.5 (Geometrische Reihe). *Sei* $|x| < 1$. *So gilt*

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

Beweis. Sicher gilt $(1-x)(1+x+\ldots+x^n) = 1-x^{n+1}$, die Partialsummen unserer Reihe ergeben sich also zu

$$1+x+\ldots+x^n = \frac{1-x^{n+1}}{1-x}$$

und streben für $n \to \infty$ wie gewünscht gegen $\frac{1}{1-x}$. □

Beispiel 5.5.6. Es gilt $1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots = 2$ und

$$0{,}999\ldots = \frac{9}{10}\sum_{k=0}^{\infty}\frac{1}{10^k} = 1$$

Übung 5.5.7. Genau dann läßt sich eine reelle Zahl durch einen periodischen Dezimalbruch darstellen, wenn sie rational ist.

Satz 5.5.8. *Sind* $\sum a_k$ *und* $\sum b_k$ *konvergente Reihen, so konvergieren auch die Reihen* $\sum(a_k+b_k)$ *und* $\sum \lambda a_k$ *und es gilt:*

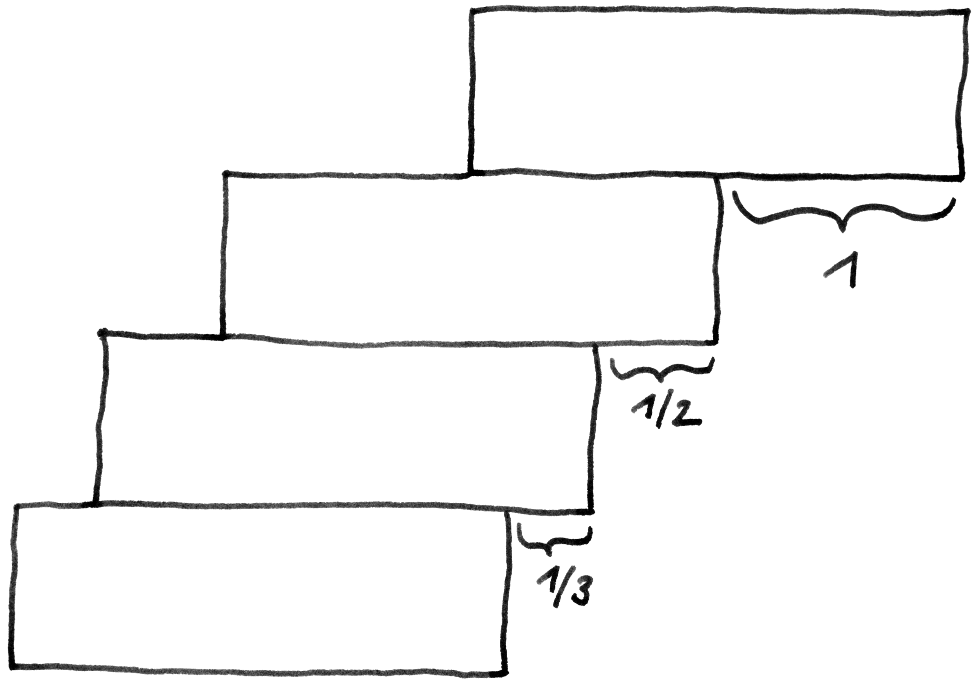
$$\begin{array}{rcl} \sum(a_k+b_k) & = & \sum a_k + \sum b_k \\ \sum \lambda a_k & = & \lambda \sum a_k \end{array}$$

Beweis. Das folgt sofort, wenn man die entsprechenden Aussagen für Folgen 5.1.34 auf die Folgen der Partialsummen anwendet. □

5.5.9. Eine Reihe kann nur dann konvergieren, wenn die Folge der Reihenglieder gegen Null strebt. In der Tat folgt das sofort, wenn wir 5.1.43 auf die Folge der Partialsummen anwenden.

Lemma 5.5.10. *Eine Reihe, die aus nichtnegativen Gliedern besteht, konvergiert genau dann, wenn die Folge ihrer Partialsummen beschränkt ist.*

Beweis. Ist kein Reihenglied negativ, so wächst die Folge der Partialsummen monoton. Ist diese Folge auch noch beschränkt, so muß sie nach 5.2.4 reell konvergent sein. Die Umkehrung ist eh klar. □



Die Divergenz der harmonischen Reihe 5.5.11 zeigt, daß man mit hinreichend vielen identischen Bauklötzen einen beliebig weit neben seinem Grundklotz endenden Turm bauen kann. Obiges Bild zeigt etwa, wie weit man mit vier Klötzen so gerade eben mal kommen kann.

Beispiel 5.5.11. Die **harmonische Reihe** $\sum_{k=1}^{\infty} \frac{1}{k}$ konvergiert nicht, da ja gilt

$$\begin{array}{rcl} \frac{1}{2} & \geq & \frac{1}{2} \\ \frac{1}{3} + \frac{1}{4} & \geq & \frac{1}{2} \\ \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} & \geq & \frac{1}{2} \\ \text{und} & \text{so} & \text{weiter.} \end{array}$$

Jedoch konvergieren die Reihen $\sum_{k=1}^{\infty} \frac{1}{k^s}$ für $s = 2, 3, 4, \ldots$, da für jede dieser Reihen die Folge der Partialsummen beschränkt ist durch $1 + \sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 2$.

Vorschau 5.5.12. In der Funktionentheorie können Sie lernen, daß diese Reihen sogar eine außerordentlich interessante Funktion $\zeta(s)$ definieren, die sogenannte **Riemann'sche ζ-Funktion**. Wir werden in **??** zeigen, daß zum Beispiel gilt $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$, $\zeta(6) = \frac{\pi^6}{945}$ und nach **??** haben wir sogar ganz allgemein $\zeta(2n) \in \mathbb{Q}\pi^{2n}$ für beliebige natürliche Zahlen $n \geq 1$. Alle diese Formeln sind berühmte Resultate des 1707 in Basel geborenen Mathematikers **Leonhard Euler**. Als Übung 7.3.24 werden Sie im übrigen zeigen, daß sogar die Reihe der Kehrwerte aller Primzahlen bereits divergiert. Für diejenigen unter Ihnen, die die komplexen Zahlen bereits kennen, sei erwähnt, daß es mit etwas größerem Aufwand sogar gelingt, $\zeta(s)$ zu definieren für jede komplexe Zahl $s \neq 1$, vergleiche etwa **??**. Die vielleicht berühmteste Vermutung der Mathematik, die sogenannte **Riemann'sche Vermutung** besagt, daß alle Nullstellen der Riemann'schen ζ-Funktion, die nicht auf der reellen Achse liegen, Realteil $1/2$ haben müssen. Ein Beweis dieser Vermutung hätte weitreichende Konsequenzen für unser Verständnis der Verteilung der Primzahlen, wie der Beweis des Primzahlsatzes **??** illustriert. Die Riemann'sche Vermutung ist übrigends der Kern des **achten Hilbert'schen Problems**.

Definition 5.5.13. Wir sagen, eine Reihe $\sum_{k=0}^{\infty} a_k$ **konvergiere absolut** genau dann, wenn die Reihe der Absolutbeträge ihrer Reihenglieder konvergiert, in Formeln $\sum_{k=0}^{\infty} |a_k| < \infty$.

Beispiel 5.5.14. Die sogenannte **alternierende harmonische Reihe**

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

konvergiert, aber nicht absolut. Daß die Reihe nicht absolut konvergiert, hatten wir schon in 5.5.11 gesehen. Um zu zeigen, daß unsere Reihe dennoch konvergiert, beachten wir, daß für die Folge s_n der Partialsummen gilt

$$s_2 \leq s_4 \leq s_6 \leq \ldots s_5 \leq s_3 \leq s_1$$

Folglich existiert $S = \sup\{s_2, s_4, \ldots\}$. Da aber gilt $s_{2k} \leq S \leq s_{2k+1}$ für alle k erhalten wir $|S - s_n| \leq \frac{1}{n}$ und folglich $\lim_{n\to\infty} s_n = S$. Wir werden in 8.4.1 sehen, daß genauer gilt $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \log 2$.

Übung 5.5.15. Man zeige das **Leibniz'sche Konvergenzkriterium**: Ist a_k eine monoton fallende Nullfolge, so konvergiert die Reihe $\sum_{k=0}^{\infty}(-1)^k a_k$.

Satz 5.5.16. *Jede absolut konvergente Reihe konvergiert.*

Beweis. Sei $\sum_{k=0}^{\infty} a_k$ unsere absolut konvergente Reihe. Seien

$$s_n = \sum_{k=0}^{n} a_k, \quad S_n = \sum_{k=0}^{n} |a_k|$$

die Partialsummen der Reihe selbst und der Reihe der Absolutbeträge. Nach Annahme konvergiert die Folge der S_n in $\mathbb{R}$ und ist also eine Cauchy-Folge. Da aber gilt $|s_n - s_m| = |\sum_{k=m+1}^{n} a_k| \leq \sum_{k=m+1}^{n} |a_k| = S_n - S_m \quad \forall n > m$, ist dann auch s_n eine Cauchy-Folge und konvergiert in $\mathbb{R}$ nach 5.2.11. □

Satz 5.5.17 (Umordnungssatz). *Ist $\sum_{k=0}^{\infty} a_k$ eine absolut konvergente Reihe und $u : \mathbb{N} \to \mathbb{N}$ eine Bijektion, so ist auch $\sum_{k=0}^{\infty} a_{u(k)}$ eine absolut konvergente Reihe und es gilt*

$$\sum_{k=0}^{\infty} a_{u(k)} = \sum_{k=0}^{\infty} a_k$$

Beweis. Da $\sum |a_k|$ konvergiert, finden wir sicher für jedes $\varepsilon > 0$ ein N mit $\sum_{k=N+1}^{\infty} |a_k| \leq \varepsilon$. Ist M so groß, daß gilt $u(\{1, \ldots, M\}) \supset \{1, \ldots, N\}$, so erhalten wir daraus für alle $n \geq N$ die Abschätzung

$$\left| \sum_{k=0}^{M} a_{u(k)} - \sum_{k=0}^{n} a_k \right| \leq \varepsilon$$

Diese Abschätzung gilt nach 5.1.33 und 5.1.42 dann auch im Grenzwert $n \to \infty$ und zeigt, daß die Folge der Partialsummen der umgeordneten Reihe konvergiert und denselben Grenzwert hat wie die Folge der Partialsummen der ursprünglichen Reihe. Wenden wir diese Erkenntnis an auf die Reihe der Absolutbeträge, so folgt auch die absolute Konvergenz der umgeordneten Reihe. □

Ergänzung 5.5.18. Ist $\sum a_k$ eine konvergente Reihe reeller Zahlen, die *nicht* absolut konvergiert, so gibt es für jedes $x \in \mathbb{R}$ eine Umordnung $u : \mathbb{N} \xrightarrow{\sim} \mathbb{N}$ mit $\sum_{k=0}^{\infty} a_{u(k)} = x$. In der Tat divergieren in diesem Fall die Reihen ihrer positiven und ihrer negativen Terme jeweils für sich genommen. Die Strategie ist nun, erst nur positive Reihenglieder zu nehmen, bis man oberhalb von x ist, dann nur negative, bis man wieder drunterrutscht, und immer so weiter.

Ergänzende Übung 5.5.19. Ist $\sum a_k$ eine konvergente Reihe reeller Zahlen und $u : \mathbb{N} \xrightarrow{\sim} \mathbb{N}$ eine Umordnung mit der Eigenschaft, daß $|u(k) - k|$ beschränkt ist, so konvergiert auch die umgeordnete Reihe $\sum a_{u(k)}$ und zwar gegen denselben Grenzwert.

Proposition 5.5.20 (Majorantenkriterium). *Sei $\sum a_k$ eine Reihe. Gibt es für unsere Reihe eine konvergente* **Majorante***, als da heißt eine konvergente Reihe $\sum b_k$ mit $|a_k| \leq b_k$ für fast alle k, so konvergiert unsere Reihe $\sum a_k$ absolut.*

Beweis. Aus 5.5.10 folgt in der Tat die Konvergenz der Reihe $\sum |a_k|$. □

Korollar 5.5.21 (Quotientenkriterium). *Sei $\sum a_k$ eine Reihe mit nichtverschwindenden Gliedern. Gibt es $\theta < 1$ mit $|a_{k+1}/a_k| < \theta$ für alle k, so konvergiert die Reihe $\sum a_k$ absolut.*

5.5.22. Bei diesem Kriterium ist wesentlich, daß θ nicht von k abhängt, die Ungleichungen $|a_{k+1}/a_k| < 1$ gelten ja auch für die divergente harmonische Reihe. Es gibt jedoch auch Reihen wie $\sum \frac{1}{k^2}$, die absolut konvergieren, obwohl sie unser Kriterium nicht dazu zwingt.

Beweis. Aus der Annahme folgt $|a_k| \leq |a_0|\theta^k$ für alle k, mithin ist die nach 5.5.5 konvergente Reihe $\sum |a_0|\theta^k$ eine Majorante unserer Reihe. □

Korollar 5.5.23. *Sei $\sum a_k$ eine Reihe mit nichtverschwindenden Gliedern. Gilt $\lim_{k\to\infty} |a_{k+1}/a_k| < 1$, so konvergiert die Reihe $\sum a_k$ absolut.*

Beweis. Klar nach dem Quotientenkriterium 5.5.21. □

Definition 5.5.24. Gegeben eine beliebige Menge I und eine Abbildung $I \to \mathbb{R}$, $i \mapsto a_i$ sagt man, die Familie der a_i sei **summierbar mit Summe** s und schreibt

$$\sum_{i\in I} a_i = s$$

als Abkürzung für die Aussage: Für jede Umgebung U von s gibt es eine endliche Teilmenge $I_U \subset I$ derart, daß für jedes endliche J mit $I_U \subset J \subset I$ gilt

$$\sum_{i\in J} a_i \in U$$

Ergänzung 5.5.25. Mir gefällt diese Definition besonders gut, da darin von einer Reihenfolge der Summanden erst gar nicht die Rede ist. In der folgenden Übung dürfen Sie zeigen, daß die in der vorhergehenden Definition erklärte Summierbarkeit im wesentlichen gleichbedeutend zu absoluter Konvergenz ist. Später, wenn wir auch in Vektorräumen summieren, erweist sich jedoch das Analogon 10.5.11 der Summierbarkeit als der nützlichere Begriff.

Übung 5.5.26. Man zeige, daß in einer summierbaren Familie von reellen Zahlen nur für höchstens abzählbar viele Indizes $i \in I$ das entsprechende a_i von Null verschieden sein kann. Hinweis: Sonst gäbe es ein $n \geq 1$ derart, daß für unendlich viele i gälte $|a_i| > 1/n$. Man zeige dann weiter, eine Familie von reellen Zahlen summierbar ist genau dann, wenn für eine und jede Abzählung ihrer von Null verschiedenen Glieder die so entstehende Reihe absolut konvergiert, und daß dann die Summe unserer Familie der Grenzwert der entsprechenden Reihe ist.

Ergänzende Übung 5.5.27. Gegeben eine summierbare Familie von reellen Zahlen $(a_i)_{i\in I}$ zeige man, daß auch für eine beliebige Teilmenge $J \subset I$ die Familie $(a_i)_{i\in J}$ summierbar ist, und daß für eine beliebige Zerlegung $I = \bigsqcup_{k\in K} I(k)$ von I in eine Vereinigung von paarweise disjunkten Teilmengen $I(k)$ gilt $\sum_{i\in I} a_i = \sum_{k\in K}(\sum_{i\in I(k)} a_i)$. Weiter zeige man für jede aufsteigende Familie von Teilmengen $I_0 \subset I_1 \subset \ldots$ mit Vereinigung I die Formel $\sum_{i\in I} a_i = \lim_{n\to\infty} \sum_{i\in I_n} a_i$. Diese Aussagen sind im übrigen Spezialfälle des Satzes von Fubini 9.6.18 und des Satzes über dominierte Konvergenz 9.5.10 aus der Theorie des Lebesgue-Integrals.

5.6 Wachstum und Zerfall

Definition 5.6.1. Wir setzen

$$\exp(x) := \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots$$

und erhalten so eine Abbildung $\exp : \mathbb{R} \to \mathbb{R}$, die **Exponentialfunktion**.

5.6.2. Die fragliche Reihe konvergiert für alle $x \in \mathbb{R}$ nach dem Quotientenkriterium oder genauer seinem Korollar 5.5.23 und sie konvergiert sogar außerordentlich schnell. Von einem formalen Standpunkt aus betrachtet ist unsere Definition also völlig unproblematisch und von einem rechentechnischen Standpunkt aus betrachtet ist sie sogar ziemlich geschickt. Sie hat nur den Nachteil, daß aus der Definition heraus weder klar wird, warum gerade diese Funktion den Namen Exponentialfunktion verdienen sollte, noch warum sie überhaupt von Interesse ist. Ich erläutere das in den gleich anschließenden Bemerkungen.

Proposition 5.6.3. *Es gilt* $\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$.

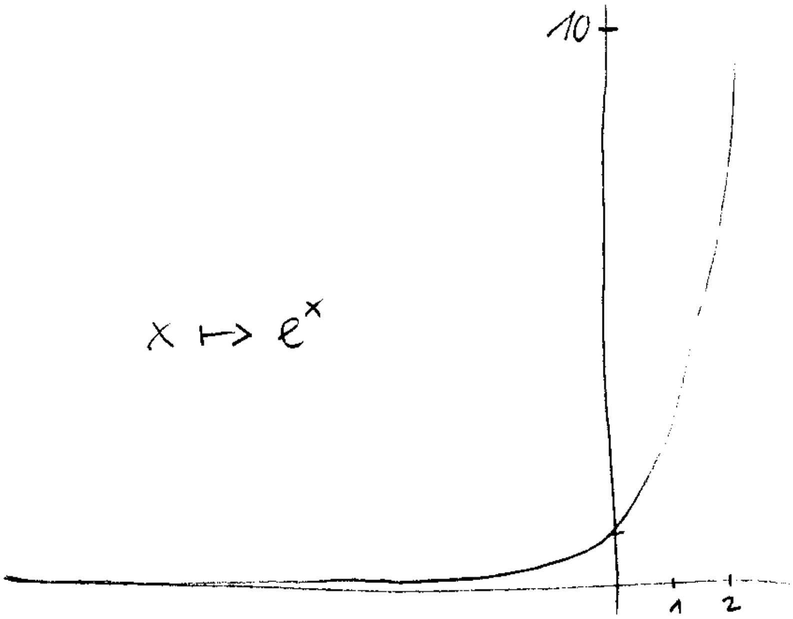
5.6.4. Die Proposition kann man dahingehend interpretieren, daß $\exp(x)$ das Kapital ist, das in x Jahren aus einem Euro entsteht bei einer “kontinuierlichen Verzinsung mit einem Zinssatz von 100%”. Legen wir das Geld zum Beispiel für ein Jahr an, so haben wir bei jährlicher Verzinsung am Ende des Jahres zwei Euro auf dem Konto. Bei monatlicher Verzinsung ergeben sich mit Zinseszinsen schon

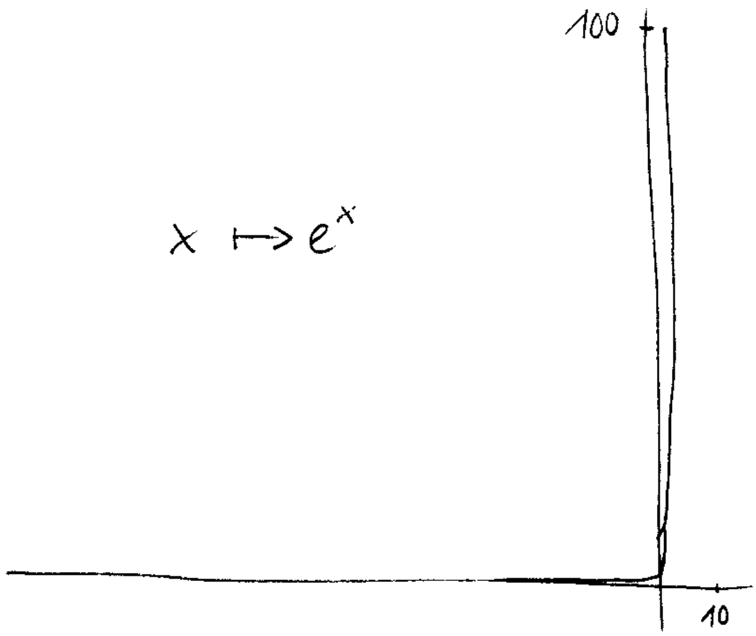
$(1 + \frac{1}{12})^{12}$ Euro, und bei kontinuierlicher Verzinsung $\mathrm{e} = \exp(1) = 2{,}781\ldots$ Euro. Man nennt $\mathrm{e} = \exp(1)$ die **Euler'sche Zahl**. In der Schule haben Sie möglicherweise e^x statt $\exp(x)$ geschrieben, aber wir erlauben uns das erst ab 6.2.15, wo wir für beliebiges $a > 0$ die Abbildung $\mathbb{Z} \to \mathbb{R}$, $n \mapsto a^n$ zu einer Abbildung $\mathbb{R} \to \mathbb{R}$, $b \mapsto a^b$ fortsetzen und zwar nach 6.3.28 auf die einzig mögliche Weise, bei der die Funktion $b \mapsto a^b$ "monoton" ist im Sinne von 6.2.1 und die "Funktionalgleichung" $a^{b+c} = a^b a^c$ erfüllt.

5.6.5. In einem Ausdruck der Gestalt a^b nennt man a die **Basis** und b den **Exponenten**, weil er eben exponiert oben an die Basis geschrieben wird. Daher rührt auch die Bezeichnung als "Exponentialfunktion". Ich würde unsere Funktion viel lieber ihrer Natur nach die "Funktion des natürlichen Wachstums" oder die "Wachstumsfunktion" nennen, aber die aus der Schreibweise abgeleitete Bezeichnung hat sich nun einmal durchgesetzt, mag sie auch aus historischen Zufällen entstanden sein: Hätte sich für die Bezeichnung des Quadrats einer Zahl a statt der Notation a^2 die Notation a_2 eingebürgert, so würde in Anbetracht dieses Schemas der Begriffsbildung unsere Exponentialfunktion heute vielleicht "Pedestalfunktion" heißen. . .

5.6.6. Eine infinitesimale Formulierung der in 5.6.4 erläuterten Bedeutung der Exponentialfunktion gibt Korollar 7.3.12, in dem die Exponentialfunktion charakterisiert wird als die eindeutig bestimmte differenzierbare Funktion von den reellen Zahlen in sich selber, die mit ihrer eigenen Ableitung übereinstimmt und bei Null den Wert Eins annimmt. Gehen wir von dieser Charakterisierung aus, so führt uns der Formalismus der Taylorreihen 8.2.2 ganz natürlich zu der Reihe, die wir in 5.6.1 haben vom Himmel fallen lassen, um möglichst schnell erste substanzielle Anwendungen unserer Betrachtungen zu Folgen und Reihen geben zu können. Eigentlich will ich es ja nach Möglichkeit vermeiden, Formeln vom Himmel fallen zu lassen, aber in diesem Fall schienen mir doch die didaktischen Vorteile der dadurch ermöglichten frühzeitigen Einführung dieser außerordentlich wichtigen Funktion zu überwiegen.

5.6.7. Die Exponentialfunktion wächst ungeheuer schnell. Eine Vorstellung davon mag die Erkenntnis 10.3.10 geben, nach der der Graph der Funktion $(\exp(x) + \exp(-x))/2$, in einem jeweils der speziellen Situation angepaßten Maßstab auf die Wand gemalt, genau die Gestalt einer zwischen zwei Nägeln durchhängenden Kette hat. Ist die Kette zwanzigmal so lang ist wie der Abstand der beiden Nägel, so stellt sie unsere Funktion in etwa auf dem Intervall von $-2{,}3$ bis $2{,}3$ dar. Ist die Kette zweihundertmal so lang wie der Abstand der beiden Nägel, so erhalten wir unsere Funktion in etwa auf dem Intervall von $-4{,}6$ bis $4{,}6$. Und eine Zwei-Meter-Kette hängt zwischen zwei im Abstand von einem knappen Zentimeter eingeschlagenen Nägeln schon recht steil!





Der Graph der Exponentialfunktion in zwei Maßstäben. Man erkennt unschwer, daß ein konstantes Wirtschaftswachstum über längere Zeiträume in einer Katastrophe enden muß. In derselben Weise entwickelt sich im Übrigen auch die Geschwindigkeit einer Vorlesung unter der Annahme, daß die Stoffmenge, die in einer Stunde vermittelt werden kann, proportional ist zur Menge des Stoffes, den die Zuhörer bereits kennen. . .

Beweis. Mit der binomischen Formel 1.1.22 ergibt sich

$$\left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{x}{n}\right)^k = \sum_{k=0}^{\infty} \frac{x^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n\ \cdot\ n\cdot\quad \dots\quad \cdot\ n}$$

Für beliebige $M, n \in \mathbb{N}$ mit $n \geq 1$ gilt also

$$\begin{aligned} \left|\exp(x) - \left(1+\tfrac{x}{n}\right)^n\right| \leq\ & \left|\exp(x) - \textstyle\sum_{k=0}^{M} \frac{x^k}{k!}\right| \\ & + \left|\textstyle\sum_{k=0}^{M} \frac{x^k}{k!} - \sum_{k=0}^{M} \frac{x^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n\cdot n\cdot\ \dots\ \cdot n}\right| \\ & + \left|\textstyle\sum_{k=M+1}^{\infty} \frac{x^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n\cdot n\cdot\ \dots\ \cdot n}\right| \end{aligned}$$

Da die Exponentialreihe für vorgegebenes x absolut konvergiert, gibt es für jedes $\varepsilon > 0$ ein $M = M_\varepsilon$ derart, daß für jedes n der erste und der letzte Term rechts beschränkt sind durch ε. Für dies feste M geht der mittlere Term bei $n \to \infty$ gegen Null, es gibt also $N = N_\varepsilon$ derart, daß er für dies feste M kleiner wird als ε falls $n \geq N$. Damit gilt $\left|\exp(x) - \left(1+\frac{x}{n}\right)^n\right| \leq 3\varepsilon$ falls $n \geq N$. □

Satz 5.6.8 (Funktionalgleichung der Exponentialfunktion). *Die Exponentialfunktion ist ein Gruppenhomomorphismus von der additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der von Null verschiedenen reellen Zahlen. Anders gesagt, für alle $x, y \in \mathbb{R}$ gilt $\exp(x) \neq 0 \neq \exp(y)$ und*

$$\exp(x+y) = \exp(x)\exp(y)$$

5.6.9. Stellen wir uns $\exp(x)$ vor als das Vermögen, daß in x Jahren aus einem Euro entsteht bei kontinuierlicher Verzinsung mit 100%, so erhalten wir offensichtlich gleichviel, ob wir unser Kapital $\exp(x)$ nach x Jahren gleich wieder für y Jahre anlegen, oder ob wir unseren Euro gleich von Anfang an $x + y$ Jahre arbeiten lassen. Das ist die Bedeutung der Funktionalgleichung. In 6.3.28 werden Sie zeigen, daß die Gruppenhomomorphismen $\varphi : \mathbb{R} \to \mathbb{R}^\times$ mit der Eigenschaft $x < y \Rightarrow \varphi(x) < \varphi(y)$ genau die Abbildungen $\varphi(x) = \exp(ax)$ sind mit $a > 0$. In 6.2.11 werden wir zeigen, daß die Exponentialfunktion sogar einen Isomorphismus zwischen der additiven Gruppe der reellen Zahlen und der multiplikativen Gruppe der positiven reellen Zahlen liefert.

5.6.10. Der gleich folgende Beweis der Funktionalgleichung gefällt mir nicht besonders. Ein anderer aber in seiner Weise auch etwas verwickelter Beweis wird in 5.6.17 vorgestellt. Ein mehr konzeptueller Zugang zur Exponentialfunktion und ihrer Funktionalgleichung wird in 7.6.10 und 7.6.11 skizziert. Er benötigt jedoch Hilfsmittel, die uns hier noch nicht zur Verfügung stehen, und er läßt auch nicht so einfach auf den Fall von Matrizen zu verallgemeinern, der für uns bei der Diskussion von Sinus und Cosinus eine wesentliche Rolle spielen wird. Wir schicken dem eigentlichen Beweis einige allgemeine Betrachtungen voraus.

Satz 5.6.11 (Produkt von Reihen). *Sind $\sum_{i=0}^{\infty} a_i$ und $\sum_{j=0}^{\infty} b_j$ absolut konvergente Reihen, so konvergiert auch die Summe der Produkte $a_i b_j$ für $(i,j) \in \mathbb{N} \times \mathbb{N}$ im Sinne von 5.5.24 und es gilt*

$$\sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} a_i b_j = \left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right)$$

Beweis. Es reicht zu zeigen, daß für irgendeine Bijektion $w : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, $k \mapsto (u(k), v(k))$ die Reihe $\sum_{k=0}^{\infty} a_{u(k)} b_{v(k)}$ absolut konvergiert und daß gilt

$$\sum_{k=0}^{\infty} a_{u(k)} b_{v(k)} = \left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right)$$

Natürlich haben wir

$$\sum_{k=0}^{n} |a_{u(k)} b_{v(k)}| \leq \left(\sum_{i=0}^{N} |a_i|\right)\left(\sum_{j=0}^{M} |b_j|\right)$$

falls gilt $N \geq \max(u(0), \ldots, u(n))$ und $M \geq \max(v(0), \ldots, v(n))$. Damit konvergiert unsere Reihe $\sum a_{u(k)} b_{v(k)}$ absolut, und nach dem Umordnungssatz 5.5.17 können wir die Bijektion $w : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ nehmen, die wir wollen, um den Grenzwert zu bestimmen. Jetzt wählen wir unsere Bijektion $w = (u, v)$ so, daß sie Bijektionen

$$\{0, \ldots, n^2 - 1\} \xrightarrow{\sim} \{0, \ldots, n-1\} \times \{0, \ldots, n-1\}$$

induziert, und erhalten Partialsummen

$$\sum_{k=0}^{n^2-1} a_{u(k)} b_{v(k)} = \left(\sum_{i=0}^{n-1} a_i\right)\left(\sum_{j=0}^{n-1} b_j\right)$$

Der Übergang zum Grenzwert $n \to \infty$ zeigt dann die Behauptung. □

5.6.12. Das Ende des Beweises hätte sehr viel besser ausgesehen, wenn wir unsere Reihen mit dem Index $k = 1$ beginnen ließen, aber so sieht der Anfang des Beweises natürlicher aus. In Wirklichkeit zeigen wir eh für beliebige im Sinne von 5.5.24 summierbare Familien reeller Zahlen $(a_i)_{i\in I}$ und $(b_j)_{j\in J}$, daß auch die Familie aller Produkte $(a_i b_j)_{(i,j)\in I\times J}$ summierbar ist und daß gilt

$$\left(\sum_{i\in I} a_i\right)\left(\sum_{j\in J} b_j\right) = \sum_{(i,j)\in I\times J} a_i b_j$$

Beweis der Funktionalgleichung 5.6.8. Wir rechnen

$$\begin{array}{rcl} \exp(x+y) & = & \sum_{k=0}^{\infty} \frac{1}{k!}(x+y)^k \\ & = & \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{i+j=k} \frac{k!}{i!j!} x^i y^j \right) \\ & = & \sum_{k=0}^{\infty} \left(\sum_{i+j=k} \frac{x^i}{i!} \frac{x^j}{j!} \right) \end{array}$$

wo wir im zweiten Schritt die binomische Formel verwenden und der Index $i+j=k$ an einer Summe bedeutet, daß wir über alle Paare $(i,j) \in \mathbb{N} \times \mathbb{N}$ mit $i+j=k$ summieren. Andererseits erhalten wir mit unserem Satz über das Produkt von Reihen bei einer geeigneten Wahl der Bijektion $w : \mathbb{N} \xrightarrow{\sim} \mathbb{N} \times \mathbb{N}$ und nach Übergang zu einer Teilfolge der Folge der Partialsummen auch

$$\begin{array}{rcl} \exp(x)\exp(y) & = & \left(\sum_{i=0}^{\infty} \frac{x^i}{i!} \right) \left(\sum_{j=0}^{\infty} \frac{y^j}{j!} \right) \\ & = & \sum_{k=0}^{\infty} \left(\sum_{i+j=k} \frac{x^i}{i!} \frac{y^j}{j!} \right) \end{array}$$

Da sicher gilt $\exp(0) = 1$, folgt $\exp(x)\exp(-x) = \exp(x + (-x)) = 1$ für alle $x \in \mathbb{R}$ und mithin $\exp(x) \neq 0$ für alle $x \in \mathbb{R}$. $\square$

Übung 5.6.13. Man folgere aus der Funktionalgleichung der Exponentialfunktion 5.6.8 die Formeln $\exp(-x) = \exp(x)^{-1}$, $\exp(x) > 0 \quad \forall x \in \mathbb{R}$, $\exp(n) = \mathrm{e}^n$, $\exp(nx) = (\exp x)^n \ \forall n \in \mathbb{Z}$ sowie $\exp(x/2) = \sqrt{\exp(x)}$.

Ergänzende Übung 5.6.14. Der Übersichtlichkeit halber kürzen wir hier im Vorgriff auf 6.2.15 schon $\exp(x) = \mathrm{e}^x$ ab. Man zeige, daß für alle $i, N \in \mathbb{N}$ gilt

$$\lim_{n\to\infty} \binom{nN}{i} \left(\frac{1}{n}\right)^i \left(1 - \frac{1}{n}\right)^{Nn-i} = \frac{N^i \,\mathrm{e}^{-N}}{i!}$$

Dieses Resultat ist in der Stochastik wichtig, wie ich im folgenden ausführen will. Gegeben $\lambda \in \mathbb{R}$ heißt die Funktion $i \mapsto \lambda^i \,\mathrm{e}^{-\lambda}/i!$ ganz allgemein die **Poisson-Verteilung** mit Parameter λ. Sie hat die folgende Bedeutung: Knetet man in einen großen Teig genau nN Rosinen ein und teilt ihn dann in n Rosinenbrötchen, so ist $\left(\frac{1}{n}\right)^i \left(1 - \frac{1}{n}\right)^{Nn-i}$ die Wahrscheinlichkeit, daß i vorgegebene Rosinen in einem fest gewählten Brötchen landen und die restlichen Rosinen in den anderen Brötchen. Mithin ist $\binom{nN}{i} \left(\frac{1}{n}\right)^i \left(1 - \frac{1}{n}\right)^{Nn-i}$ die Wahrscheinlichkeit, daß in einem fest gewählten Brötchen genau i Rosinen landen. Ist unser Brötchen klein im Vergleich zum ganzen Teig, so liegt diese Wahrscheinlichkeit also nahe bei $N^i \,\mathrm{e}^{-N}/i!$ oder allgemeiner bei $\lambda^i \,\mathrm{e}^{-\lambda}/i!$ mit λ der durchschnittlichen Zahl von Rosinen in dem Teigvolumen, das man für ein Brötchen braucht. Genau genommen stimmt das allerdings nur für punktförmige Rosinen, denn sonst liefert die Größe des Brötchens eine obere Schranke für die möglichen Anzahlen der darin verbackenen Rosinen.

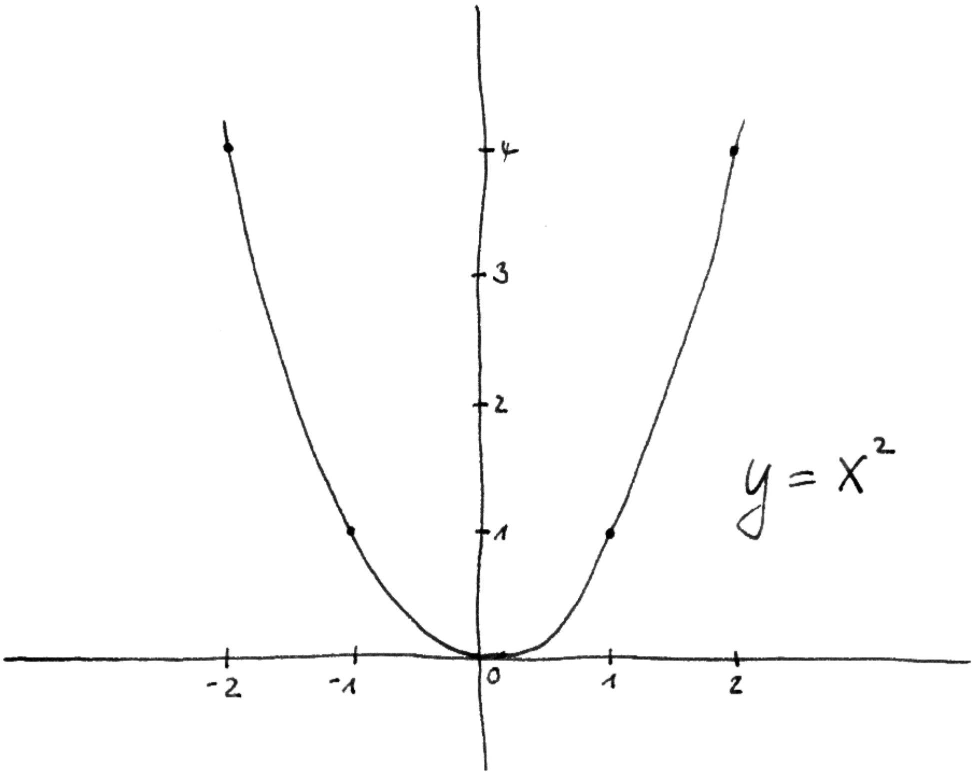
Ergänzende Übung 5.6.15. Man berechne die Euler'sche Zahl e bis auf 5 sichere Stellen hinter dem Komma.

Ergänzende Übung 5.6.16. Die Euler'sche Zahl e ist nicht rational. Man zeige dies, indem man von ihrer Darstellung als Reihe ausgeht und durch geeignete Abschätzungen nachweist, daß $q!\mathrm{e}$ für $q \in \mathbb{N}$ mit $q \geq 2$ nie eine ganze Zahl sein kann.

Ergänzende Übung 5.6.17. In dieser Übung sollen Sie einen anderen Zugang zur Funktionalgleichung der Exponentialfunktion ausarbeiten, den ich in einer Arbeit von Martin Kneser kennengelernt habe: Man zeige dazu in Verallgemeinerung von 5.6.3, daß für jede reelle Zahl $x \in \mathbb{R}$ aus $\lim_{n\to\infty} x_n = x$ folgt $\lim_{n\to\infty} \left(1 + \frac{x_n}{n}\right)^n = \exp(x)$. Mithilfe der Identität

$$\left(1 + \frac{x}{n}\right)\left(1 + \frac{y}{n}\right) = \left(1 + \frac{x + y + (xy/n)}{n}\right)$$

folgere man dann die Funktionalgleichung.



Die Parabel $y = x^2$ ist der Graph der Funktion $f(x) = x^2$.

6 Stetigkeit

6.1 Definition und erste Beispiele

6.1.1. Abbildungen mit Werten in irgendeiner Art von Zahlen nennen wir **Funktionen**. Wir erlauben hier auch Werte in $\overline{\mathbb{R}}$. Wollen wir besonders betonen, daß nur reelle Zahlen als Werte angenommen werden, so sprechen wir von **reellwertigen Funktionen**. Reellwertige Funktionen auf der reellen Zahlengeraden kann man sich auf mindestens vier verschiedene Arten vorstellen:

1. In der Schule ist es üblich, eine Funktion $f : \mathbb{R} \to \mathbb{R}$ durch ihren Graphen $\Gamma(f) = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}$ zu veranschaulichen, also durch eine Teilmenge der Ebene $\mathbb{R}^2$.

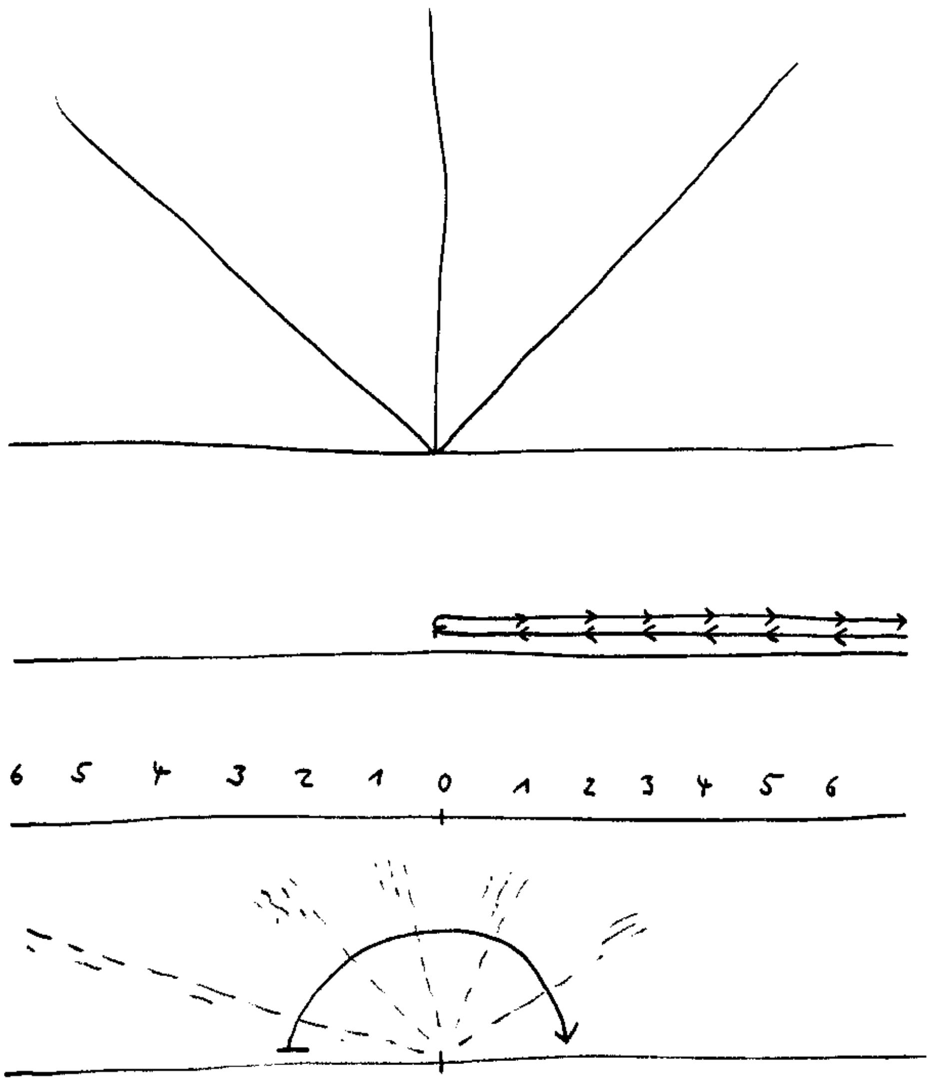
2. In der Physik ist es üblich, sich eine Abbildung $f : \mathbb{R} \to X$, $t \mapsto f(t)$ von $\mathbb{R}$ in irgendeine Menge X vorzustellen als ein Teilchen, das sich "im Raum X bewegt und sich zur Zeit t (für lateinisch "tempus") am Punkt $f(t)$ befindet". In unserem Fall hätten wir uns also ein Teilchen vorzustellen, das sich auf der Zahlengerade $X = \mathbb{R}$ bewegt.

3. Eine reellwertige Funktion auf einer beliebigen Menge kann man sich als eine Temperaturverteilung auf besagter Menge vorstellen, im vorliegenden Fall also als eine Temperaturverteilung auf der reellen Zahlengeraden.

4. In der Mathematik ist es auch nützlich, sich eine Funktion $f : \mathbb{R} \to \mathbb{R}$ wirklich als Abbildung der Zahlengerade auf sich selber vorzustellen. Als Beispiel betrachten wir den Absolutbetrag, der als Abbildung aufgefaßt den negativen Teil der Zahlengerade auf den positiven Teil herüberklappt.

6.1.2. Beliebige Abbildungen $\mathbb{R} \to \mathbb{R}$ können wild aussehen, man denke nur etwa an die Abbildung, die jeder rationalen Zahl den Betrag ihres Nenners nach vollständigem Kürzen zuordnet und jeder irrationalen Zahl ihre fünfte Nachkommastelle. Wir führen nun die Klasse der "stetigen Funktionen" ein und zeigen insbesondere, daß für stetige auf einem Intervall definierte und injektive Funktionen auch ihr Bild ein Intervall ist und die Umkehrfunktion stetig. Das liefert uns dann viele neue Funktionen als Umkehrfunktionen bereits bekannter Funktionen. Anschaulich ist eine reellwertige Funktion auf einem reellen Intervall stetig genau dann, wenn man "ihren Graphen zeichnen kann ohne den Stift abzusetzen". Diese Anschauung werden wir im Folgenden präzisieren. Wir erinnern an den Umgebungsbegriff aus 5.1.8.

Definition 6.1.3. Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $f : D \to \overline{\mathbb{R}}$ eine Funktion. Gegeben ein Punkt $p \in D$ heißt unsere Funktion **stetig bei** p genau dann, wenn es



Vier verschiedene Anschauungen für eine reellwertige Funktion einer reellen Veränderlichen am Beispiel des Absolutbetrags $x \mapsto |x|$

für jede Umgebung U von $f(p)$ eine Umgebung U' von p gibt mit $f(U' \cap D) \subset U$. Unsere Abbildung heißt **stetig** (englisch **continuous**, französisch **continue**) genau dann, wenn sie stetig ist bei jedem Punkt $p \in D$.

6.1.4. Wir vereinbaren diese Definition gleich im Fall einer Teilmenge $D \subset \overline{\mathbb{R}}$ und einer Funktion $f : D \to \overline{\mathbb{R}}$, weil so der Zusammenhang mit dem Grenzwertbegriff in 6.3 in voller Allgemeinheit dargestellt werden kann. Wie bei Folgen reicht es auch hier, die Existenz von U' für alle Umgebungen U eines Fundamentalsystems von Umgebungen von $f(p)$ zu zeigen.

Beispiele 6.1.5. Für alle $D \subset \overline{\mathbb{R}}$ ist die Einbettung $i : D \hookrightarrow \overline{\mathbb{R}}$, $x \mapsto x$ stetig. Der Absolutbetrag $\mathrm{abs} : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ ist stetig. Für jedes $c \in \overline{\mathbb{R}}$ ist die konstante Funktion $c : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, $x \mapsto c$ stetig. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = 1$ für $x \geq 0$ und $f(x) = 0$ für $x < 0$ ist nicht stetig bei $p = 0$, ihre Einschränkung auf $D = \mathbb{R}^\times$ ist jedoch stetig. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$ für $x \in \mathbb{Q}$ und $f(x) = 0$ für $x \notin \mathbb{Q}$ ist nur an der Stelle $p = 0$ stetig.

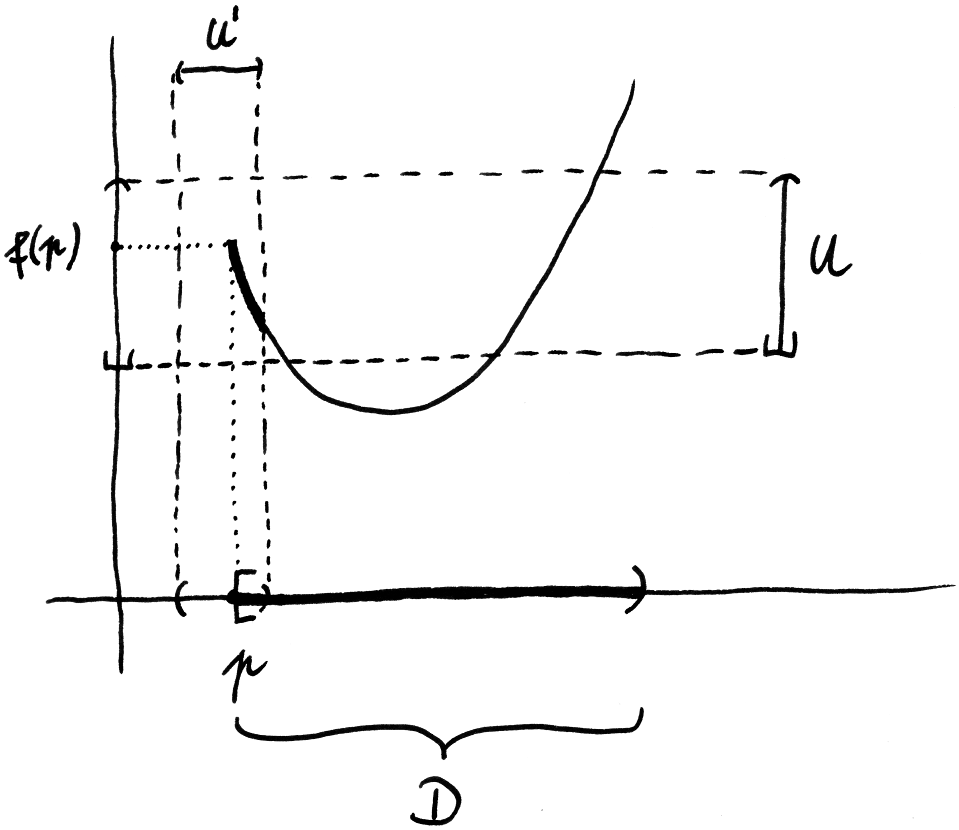
Satz 6.1.6 (Die Verknüpfung stetiger Funktionen ist stetig). *Seien genauer $D, E \subset \overline{\mathbb{R}}$ Teilmengen, $f : D \to \overline{\mathbb{R}}$, $g : E \to \overline{\mathbb{R}}$ Funktionen, und es gelte $f(D) \subset E$. Ist f stetig bei einem Punkt $p \in D$ und g stetig bei seinem Bild $f(p)$, so ist auch $g \circ f : D \to \overline{\mathbb{R}}$, $x \mapsto g(f(x))$ stetig bei p.*

6.1.7. Hier machen wir unsere Ankündigung aus 2.2.19 wahr. Genau genommen ist die Notation $g \circ f$ nämlich nicht korrekt: Wir müßten eigentlich erst eine Abbildung $\tilde{f} : D \to E$ definieren durch $\tilde{f}(x) = f(x)$ für alle $x \in D$ und dann die Abbildung $g \circ \tilde{f}$ betrachten. Das ist nun jedoch meiner Ansicht nach ein Fall, in dem größere Präzision nicht mehr zur besseren Verständlichkeit beiträgt.

Beweis. Da g stetig ist bei $f(p)$, finden wir für jede Umgebung U von $g(f(p))$ eine Umgebung U' von $f(p)$ mit $g(U' \cap E) \subset U$. Da f stetig ist bei p, finden wir für diese Umgebung U' von $f(p)$ eine Umgebung U'' von p mit $f(U'' \cap D) \subset U'$. Damit finden wir in der Tat für jede Umgebung U von $(g \circ f)(p)$ eine Umgebung U'' von p mit $(g \circ f)(U'' \cap D) \subset U$. □

6.1.8. Einschränkungen stetiger Funktionen sind stetig. Ist genauer $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $f : D \to \overline{\mathbb{R}}$ stetig bei $p \in D$ und $E \subset D$ eine Teilmenge mit $p \in E$, so ist auch die Einschränkung $f|_E : E \to \overline{\mathbb{R}}$ stetig bei p. Das folgt einerseits sofort aus der Definition und andererseits auch aus dem vorhergehenden Satz 2.2.19, indem wir die Einschränkung als Verknüpfung mit der nach 6.1.5 stetigen Einbettung von E schreiben.

6.1.9 (**Stetigkeit ist eine lokale Eigenschaft**). Wird eine Funktion stetig bei einem Punkt nach Einschränkung auf eine Umgebung des besagten Punktes, so war sie dort schon selbst stetig. Ist genauer und in Formeln $D \subset \overline{\mathbb{R}}$ eine Teilmenge



In diesem Bild habe ich für *eine* Umgebung U von $f(p)$ mal eine mögliche Umgebung U' von p eingezeichnet. Stetigkeit bei p bedeutet jedoch sehr viel stärker, daß wir für *jede* Umgebung U von $f(p)$ eine in der in 6.1.3 präzisierten Weise mögliche Umgebung U' von p finden können. Fett eingezeichnet ist auf der x-Achse der Definitionsbereich D unserer Funktion f, der Punkt p ist sein kleinstes Element. Auf dem Graphen von f habe ich den Teil über $U' \cap D$ fett eingezeichnet, damit man gut sehen kann, daß in der Tat gilt $f(U' \cap D) \subset U$. Unsere eigentlichen U und U' sind die Projektionen der so bezeichneten "freischwebenden Intervalle" auf die jeweiligen Koordinatenachsen, was ich versucht habe, durch die gestrichelten Linien anzudeuten.

und $f : D \to \overline{\mathbb{R}}$ eine Funktion und $p \in D$ ein Punkt und gibt es eine Umgebung U von p derart, daß die Einschränkung von f auf $D \cap U$ stetig ist bei p, so ist auch $f : D \to \overline{\mathbb{R}}$ bereits stetig bei p. Um das zum Ausdruck zu bringen, sagt man dann etwas vage, die "Stetigkeit sei eine lokale Eigenschaft".

Lemma 6.1.10 (**ε-δ-Kriterium für die Stetigkeit**)**.** *Sei $D \subset \mathbb{R}$ eine Teilmenge, $f : D \to \mathbb{R}$ eine reellwertige Funktion und $p \in D$ ein Punkt. Genau dann ist f stetig bei p, wenn es für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt derart, daß für alle $x \in D$ mit $|x - p| < \delta$ gilt $|f(x) - f(p)| < \varepsilon$.*

Beweis. Das folgt sofort aus unserer Definition der Stetigkeit 6.1.3 und des Umgebungsbegriffs 5.1.8. □

Übung 6.1.11. Seien $I, J \subset \overline{\mathbb{R}}$ Intervalle mit nichtleerem Schnitt und sei $f : (I \cup J) \to \overline{\mathbb{R}}$ eine Funktion. Man zeige: Sind die Einschränkungen $f|_I$ und $f|_J$ stetig, so ist auch f selbst stetig. Im übrigen wird sich der "schwierige" Fall dieser Übung als Spezialfall von 9.5.32 erweisen.

Lemma 6.1.12. *Die Funktion $\mathbb{R}^\times \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ ist stetig.*

Beweis. Das ist gerade die Aussage von 5.1.35.3. □

Lemma 6.1.13. *Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist stetig.*

6.1.14. Die Stetigkeit der Exponentialfunktion können wir später auch aus dem allgemeinen Satz 8.1.5 folgern: Potenzreihen stellen auf ihrem Konvergenzbereich immer stetige Funktionen dar.

Beweis. Wir wenden das ε-δ-Kriterium 6.1.10 an. Mit der Funktionalgleichung finden wir

$$|\exp(x) - \exp(p)| = |\exp(p)| \cdot |\exp(x - p) - \exp(0)|$$

Nun beachten wir, daß für $|y| \le 1$ gilt

$$|\exp(y) - \exp(0)| = |y| \cdot \left| \sum_{i=1}^{\infty} \frac{y^{i-1}}{i!} \right| \le |y| \cdot \sum_{i=1}^{\infty} \frac{|y|^{i-1}}{(i-1)!} \le |y| \exp(1)$$

wo wir im zweiten Schritt die Dreiecksungleichung sowie die Erhaltung von Ungleichungen im Grenzwert verwenden und die Nenner verkleinern. Aus $|x - p| \le 1$ folgt also $|\exp(x) - \exp(p)| \le \exp(p)|x - p|\,\mathrm{e}$ und für gegebenes $\varepsilon > 0$ können wir mithin $\delta = \inf\{1, \varepsilon/(\exp(p)\,\mathrm{e})\}$ nehmen. □

Definition 6.1.15. Gegeben reellwertige Funktionen $f, g : D \to \mathbb{R}$ definieren wir die Funktionen $f + g, fg : D \to \mathbb{R}$ durch $(f + g)(x) = f(x) + g(x)$, $(fg)(x) = f(x) \cdot g(x)\ \forall x \in D$.

Satz 6.1.16. Summe und Produkt stetiger Funktionen sind stetig. *Ist genauer $D \subset \overline{\mathbb{R}}$ gegeben und $p \in D$ ein Punkt und sind $f, g : D \to \mathbb{R}$ stetig bei p, so sind auch $f + g$ und fg stetig bei p.*

Bemerkung 6.1.17. Nehmen f und g allgemeiner Werte in $\overline{\mathbb{R}}$ an und sind an jeder Stelle $x \in D$ die Summe $f(x) + g(x)$ bzw. das Produkt $f(x) \cdot g(x)$ sinnvoll definiert im Sinne von 5.1.38, so gilt der Satz entsprechend mit fast demselben Beweis.

Beweis. Wir zeigen das nur für die Multiplikation, der Fall der Addition geht analog. Ist W eine Umgebung von $f(p)g(p)$, so gibt es nach 5.1.35 Umgebungen U von $f(p)$ und V von $g(p)$ mit $U \cdot V \subset W$. Da sowohl f als auch g stetig sind bei p, gibt es weiter Umgebungen U' und V' von p mit $f(U' \cap D) \subset U$ und $g(V' \cap D) \subset V$. Ihr Schnitt $U' \cap V'$ ist dann eine Umgebung W' von p mit $(fg)(W' \cap D) \subset W$. □

Definition 6.1.18. Wir nennen eine Funktion $\mathbb{R} \to \mathbb{R}$ eine **Polynomfunktion** genau dann, wenn es $a_0, a_1, \ldots, a_n \in \mathbb{R}$ gibt derart, daß unsere Funktion gegeben wird durch die Abbildungsvorschrift $x \mapsto a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$.

6.1.19. Jede Folge reeller Zahlen $(a_n)_{n \in \mathbb{N}}$, in der fast alle Folgenglieder Null sind, liefert uns eine Polynomfunktion. Nach **??** liefern verschiedene Folgen auch stets verschiedene Funktionen. Im Licht dieser Tatsache werden wir unsere Polynomfunktionen meist kürzer Polynome nennen, obwohl eigentlich der Begriff Polynom eher den formalen Ausdruck meint. Diese Unterscheidung ist aber erst in der Algebra wirklich von Belang, wenn man zum Beispiel mit endlichen Körpern arbeitet, vergleiche **??**.

Bemerkung 6.1.20. Das Wort "Polynom" kommt von der griechischen Vorsilbe "poly" für "viele" und dem lateinischen Wort "nomen" für "Namen". Allgemeiner betrachtet man auch Polynome in mehreren Veränderlichen und meint damit Ausdrücke wie etwa $xyz + 7x^2y^4 - 12z + 1$. Dieses Polynom ist die Summe der vier **Monome** xyz, $7x^2y^4$, $-12z$ und 1, wobei das Wort "Monom" diesmal mit der griechischen Vorsilbe "mono" für "allein" gebildet ist. Einen Ausdruck wie $xyz + 7x^2y^4$ würde man als **Binom** bezeichnen, diesmal mit der griechischen Vorsilbe "bi" für "Zwei". Ein anderes Binom wäre der Ausdruck $(x + y)$, dessen Potenzen die "binomische" Formel 1.1.22 explizit angibt. Es sollte klar sein, wie man aus unserer binomischen Formel auch Formeln für die Potenzen eines beliebigen Binoms erhält.

Korollar 6.1.21. *Polynomfunktionen sind stetig.*

Beweis. Das folgt induktiv aus dem vorhergehenden Satz 6.1.16. □

Korollar 6.1.22 (Quotienten stetiger Funktionen sind stetig). *Ist genauer $D \subset \overline{\mathbb{R}}$ eine Teilmenge und sind $f, g : D \to \mathbb{R}$ stetig und hat g keine Nullstelle in D, so ist auch die Funktion $f/g : D \to \mathbb{R}$, $x \mapsto f(x)/g(x)$ stetig.*

Beweis. Bezeichne $i : \mathbb{R}^\times \to \mathbb{R}$ die Funktion $x \mapsto 1/x$. Sie ist stetig nach 6.1.12. Also ist auch $i \circ g : D \to \mathbb{R},\ x \mapsto 1/g(x)$ stetig, und dann auch das Produkt dieser Funktion mit der stetigen Funktion f. □

6.1.23. Eine Funktion, die sich als der Quotient eines Polynoms durch ein von Null verschiedenes Polynom darstellen läßt, heißt eine **rationale Funktion**. So eine Funktion ist a priori natürlich nur da definiert, wo der Nenner nicht verschwindet, und ist nach dem vorhergehenden Korollar auf dem Komplement der Nullstellenmenge ihres Nenners stetig. Betrachten wir unsere rationale Funktion jedoch nicht als Abbildung, sondern als formalen Ausdruck, so verstehen wir unter ihrem Definitionsbereich die etwas größere Menge, auf der "nach maximalem Kürzen" der Nenner keine Nullstellen hat, vergleiche **??**. Man beachte jedoch, daß die hier gegebene etwas unscharfe Formulierung "nach maximalem Kürzen" eigentlich erst in **??** gerechtfertigt wird, wo wir die Eindeutigkeit der "Primfaktorzerlegung" in Polynomringen diskutieren. Bleibt auch nach maximalem Kürzen noch ein nichtkonstantes Polynom im Nenner stehen, so spricht man von einer **gebrochen rationalen Funktion**.

6.2 Umkehrfunktionen und Zwischenwertsatz

Definition 6.2.1. Eine Abbildung f zwischen angeordneten Mengen heißt

monoton wachsend,	wenn gilt	$x \leq y$	$\Rightarrow$	$f(x) \leq f(y)$
streng monoton wachsend,	wenn gilt	$x < y$	$\Rightarrow$	$f(x) < f(y)$
monoton fallend,	wenn gilt	$x \leq y$	$\Rightarrow$	$f(x) \geq f(y)$
streng monoton fallend,	wenn gilt	$x < y$	$\Rightarrow$	$f(x) > f(y)$

Eine Abbildung heißt **(streng) monoton** genau dann, wenn sie (streng) monoton wachsend oder fallend ist. Das verallgemeinert unsere Begriffe für Folgen aus 5.2.3.

Proposition 6.2.2. *Ist $I \subset \overline{\mathbb{R}}$ ein Intervall und $f : I \to \overline{\mathbb{R}}$ streng monoton, so ist die Umkehrfunktion $f^{-1} : f(I) \to \overline{\mathbb{R}}$ stetig.*

6.2.3. Wieder ist unsere Notation nicht ganz korrekt: Wir müßten genau genommen eigentlich die Bijektion $\tilde{f} : I \xrightarrow{\sim} f(I), x \mapsto f(x)$ betrachten, dazu die inverse Abbildung $\tilde{f}^{-1} : f(I) \xrightarrow{\sim} I$ nehmen, und unser $f^{-1} : f(I) \to \overline{\mathbb{R}}$ definieren als die Verknüpfung von $\tilde{f}^{-1}$ mit der Einbettung $I \hookrightarrow \overline{\mathbb{R}}$.

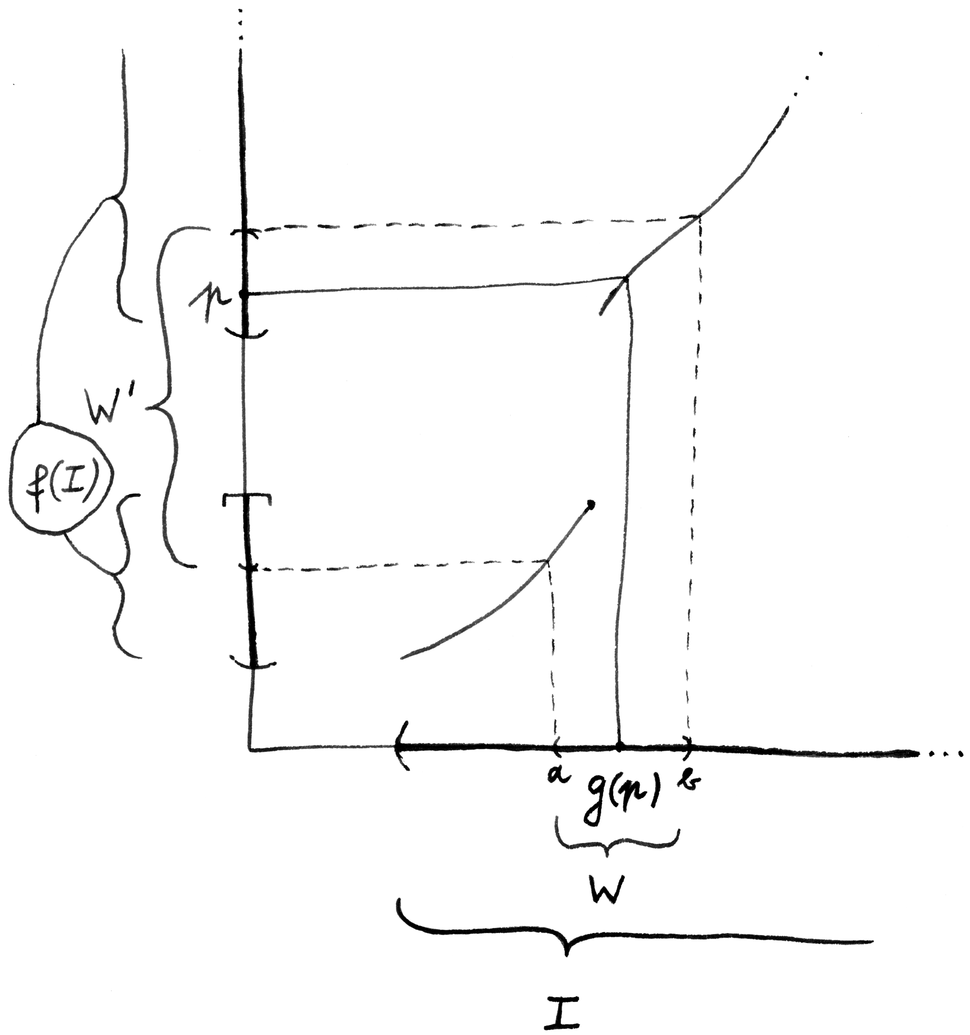


Illustration zum Beweis von 6.2.2. Statt U und U' heißen unsere Umgebungen in diesem Bild allerdings W und W'. Man beachte, wie sich, salopp gesprochen, "Unstetigkeitsstellen der Ausgangsfunktion in Definitionslücken der Umkehrfunktion verwandeln".

6.2.4. Die Umkehrfunktion f^{-1} darf nicht verwechselt werden mit der Funktion $x \mapsto 1/f(x)$: Ist zum Beispiel $f : (0, \infty) \to \mathbb{R}$ gegeben durch $f(x) = x^2$, so haben wir $1/f(x) = x^{-2}$, aber die Umkehrabbildung ist $f^{-1}(y) = \sqrt{y}$. Die Notation ist hier leider nicht ganz eindeutig. Oft muß man aus dem Kontext erschließen, ob mit f^{-1} die Umkehrfunktion von f oder vielmehr die "Kehrwertfunktion" $x \mapsto 1/f(x)$ gemeint ist.

Beweis. Sei ohne Beschränkung der Allgemeinheit f streng monoton wachsend. Bezeichne $g : f(I) \to \overline{\mathbb{R}}$ die Umkehrfunktion, die unter diesen Umständen auch streng monoton wachsen muß. Gegeben $p \in f(I)$ müssen wir für jede Umgebung U von $g(p)$ eine Umgebung U' von p finden mit der Eigenschaft $g(U' \cap f(I)) \subset U$. Wir unterscheiden drei Fälle: Ist das Definitionsintervall I unserer Ausgangsfunktion f eine Umgebung von $g(p)$, so umfaßt jede Umgebung U von $g(p)$ ein Intervall der Gestalt (a, b) mit $a < g(p) < b$ und $a, b \in I$ und wir können schlicht $U' = (f(a), f(b))$ nehmen. Besteht I nur aus einem einzigen Punkt, so ist eh alles klar. Besteht schließlich I aus mehr als einem Punkt und ist $g(p)$ eine seiner Grenzen, ohne Beschränkung der Allgemeinheit sein Supremum, so umfaßt jede Umgebung U von $g(p)$ ein Intervall der Gestalt $(a, g(p)]$ mit $a \in I$ und wir können $U' = (f(a), \infty]$ nehmen. □

Beispiel 6.2.5. Wenden wir unseren Satz an auf die Funktion $f : [0, \infty) \to \mathbb{R}$, $x \mapsto x^2$, so finden wir insbesondere, daß das Wurzelziehen

$$[0, \infty) \to \mathbb{R},\ x \mapsto \sqrt{x}$$

eine stetige Funktion ist. Da sich die Funktion $x \mapsto x^2$ zu einer streng monotonen Bijektion $[0, \infty] \xrightarrow{\sim} [0, \infty]$ fortsetzen läßt durch die Vorschrift $\infty \mapsto \infty$, läßt sich nach unserem Satz die Wurzelfunktion durch die Vorschrift $\infty \mapsto \infty$ zu einer stetigen Bijektion $[0, \infty] \xrightarrow{\sim} [0, \infty]$ fortsetzen. Analoges gilt auch für höhere Wurzeln, nur haben wir bis jetzt noch nicht bewiesen, daß wir solche höheren Wurzeln auch tatsächlich aus jeder nichtnegativen reellen Zahl ziehen können. Das folgt jedoch aus den Sätzen, die wir im Anschluß behandeln.

Satz 6.2.6 (Zwischenwertsatz). *Für $a \leq b$ aus $\overline{\mathbb{R}}$ nimmt eine stetige Funktion $f : [a, b] \to \overline{\mathbb{R}}$ jeden Wert zwischen $f(a)$ und $f(b)$ an.*

Erster Beweis. Wir dürfen ohne Beschränkung der Allgemeinheit $f(a) \leq f(b)$ annehmen. Gegeben $z \in [f(a), f(b)]$ suchen wir $p \in [a, b]$ mit $f(p) = z$. Wir betrachten dazu

$$p = \sup\{x \in [a, b] \mid f(x) \leq z\}$$

und behaupten $f(p) = z$. Um das zu zeigen führen wir die Annahmen $f(p) < z$ und $z < f(p)$ beide zum Widerspruch: Aus $f(p) < z$ folgte zunächst $p < b$, und

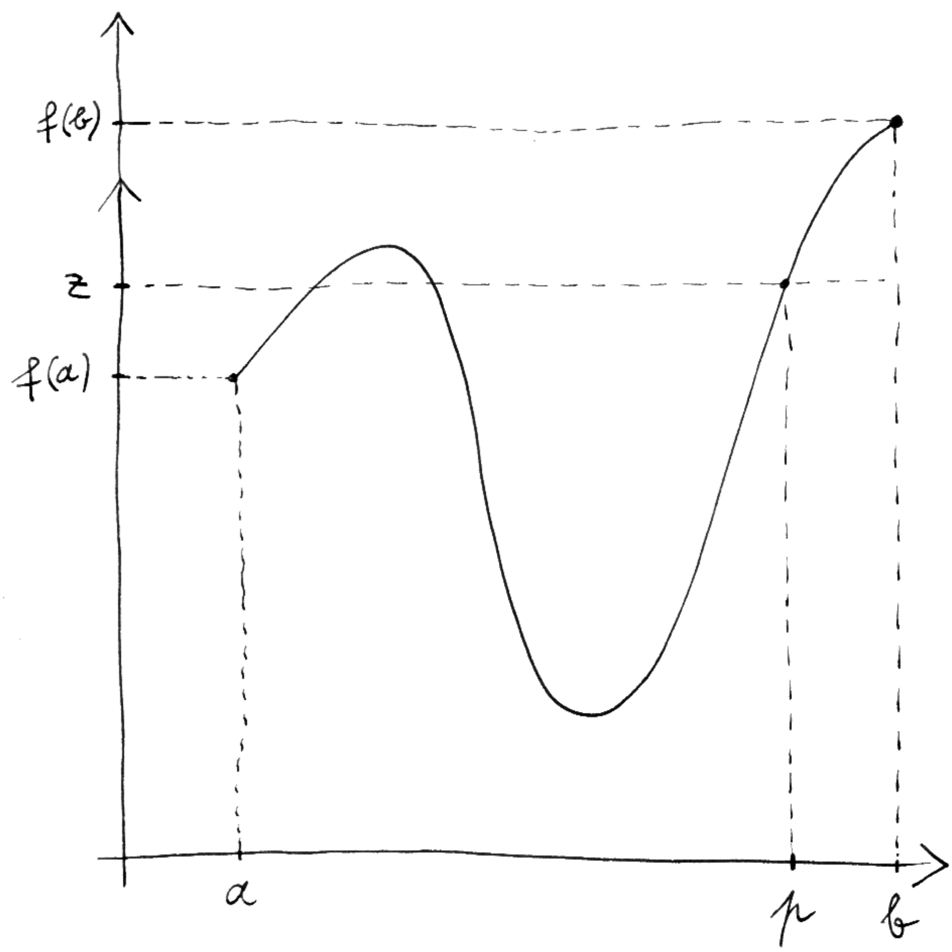


Illustration zum Zwischenwertsatz. Im vorliegenden Fall wir der gegebene Zwischenwert z sogar dreimal als Funktionswert angenommen. Unser erster Beweis führt stets zum hier eingezeichneten größten $p \in [a, b]$, an dem z als Funktionswert angenommen wird.

dann gäbe es aufgrund der Stetigkeit ein p' mit $p < p' < b$ und $f(p') \leq z$ und p wäre gar keine obere Schranke unserer Menge gewesen. Aus $z < f(p)$ folgte zunächst $a < p$, und dann gäbe es aufgrund der Stetigkeit ein p' mit $a < p' < p$ und $z < f(x)$ für alle $x \in [p', p]$. Also wäre auch p' schon eine obere Schranke unserer Menge und p könnte nicht ihre kleinste obere Schranke gewesen sein. □

Zweiter Beweis für den Fall $a, b \in \mathbb{R}$. Wir dürfen ohne Beschränkung der Allgemeinheit $f(a) \leq f(b)$ annehmen. Gegeben $z \in [f(a), f(b)]$ suchen wir $p \in [a, b]$ mit $f(p) = z$. Dazu nehmen wir den Mittelpunkt m_0 unseres Intervalls her und werten unsere Funktion dort aus. Gilt $f(m_0) \geq z$, so setzen wir $a_1 = a$ und $b_1 = m_0$. Sonst setzen wir $a_1 = m_0$ und $b_1 = b$. In jedem Fall gilt $f(a_1) \leq z \leq f(b_1)$. Anschließend nehmen wir den Mittelpunkt m_1 des nur noch halb so großen Intervalls $[a_1, b_1]$ her und verfahren genauso. Auf diese Weise erhalten wir eine monoton wachsende Folge $a = a_0, a_1, \ldots$ und eine monoton fallende Folge $b = b_0, b_1, \ldots$ mit $b_n - a_n = 2^{-n}(b - a)$ und $f(a_n) \leq z \leq f(b_n)$ und für alle n. Unsere beiden Folgen müssen also konvergieren, und zwar gegen denselben Grenzwert $p \in [a, b]$. Aus der Stetigkeit von f und der Erhaltung von Ungleichungen im Grenzwert nach 6.3.18 folgt dann

$$f(p) = \lim_{n\to\infty} f(a_n) \quad \leq \quad z \quad \leq \quad \lim_{n\to\infty} f(b_n) = f(p)$$

und damit $z = f(p)$ wie gewünscht. Dieser Beweis hat den Nachteil, nur für $a, b \in \mathbb{R}$ zu funktionieren und bereits im Vorgriff die in diesem Text erst in 6.3.18 bewiesenen Sätze über die Vertauschbarkeit von Grenzwertbildung und dem Anwenden stetiger Funktionen zu verwenden. Dafür hat er den Vorteil, einen auch in der Praxis gangbaren Lösungsalgorithmus zu beschreiben, das sogenannte **Intervallhalbierungsverfahren**. □

Korollar 6.2.7 (Abstrakter Zwischenwertsatz). *Das Bild eines Intervalls unter einer stetigen Funktion ist ein Intervall. Ist also in Formeln $I \subset \overline{\mathbb{R}}$ ein Intervall und $f : I \to \overline{\mathbb{R}}$ stetig, so ist auch $f(I)$ ein Intervall.*

Beweis. Das folgt sofort aus 6.2.6. □

Satz 6.2.8 (über die Umkehrfunktion). *Ist $I \subset \overline{\mathbb{R}}$ ein Intervall und $f : I \to \overline{\mathbb{R}}$ streng monoton und stetig, so ist auch $f(I) \subset \overline{\mathbb{R}}$ ein Intervall und die Umkehrfunktion $f^{-1} : f(I) \to \overline{\mathbb{R}}$ ist streng monoton und stetig.*

Beweis. Dieser Satz ist nur die Zusammenfassung des abstrakten Zwischenwertsatzes 6.2.7 mit der Proposition 6.2.2 über die Stetigkeit der Umkehrfunktion. □

Definition 6.2.9. Für $q \in \mathbb{N}$, $q \geq 1$ definieren wir die **q-te Wurzel**

$$\sqrt[q]{\ } : [0, \infty) \to \mathbb{R}$$

als die Umkehrfunktion zur q-ten Potenzfunktion $x \mapsto x^q$. Nach 6.2.2 ist $x \mapsto \sqrt[q]{x}$ stetig.

Ergänzende Übung 6.2.10. Sei $\sum a_k$ eine Reihe. Man zeige das **Wurzelkriterium**: Gilt $\lim_{k\to\infty} \sqrt[k]{|a_k|} < 1$, so konvergiert die Reihe $\sum a_k$ absolut. Hinweis: Analog zum Beweis des Quotientenkriteriums. Meiner Erfahrung nach ist dies Kriterium in der Praxis selten von Nutzen, da es meist auf schwer zu bestimmende Grenzwerte führt.

Definition 6.2.11. Aus dem Zwischenwertsatz folgt $\exp(\mathbb{R}) = (0, \infty)$, denn wir haben offensichtlich $\lim_{n\to\infty} \exp(n) = \infty$ und damit $\lim_{n\to-\infty} \exp(n) = 0$ nach 5.1.28. Wir können also den **(natürlichen) Logarithmus** einführen als die Umkehrfunktion

$$\log : (0, \infty) \to \mathbb{R}$$

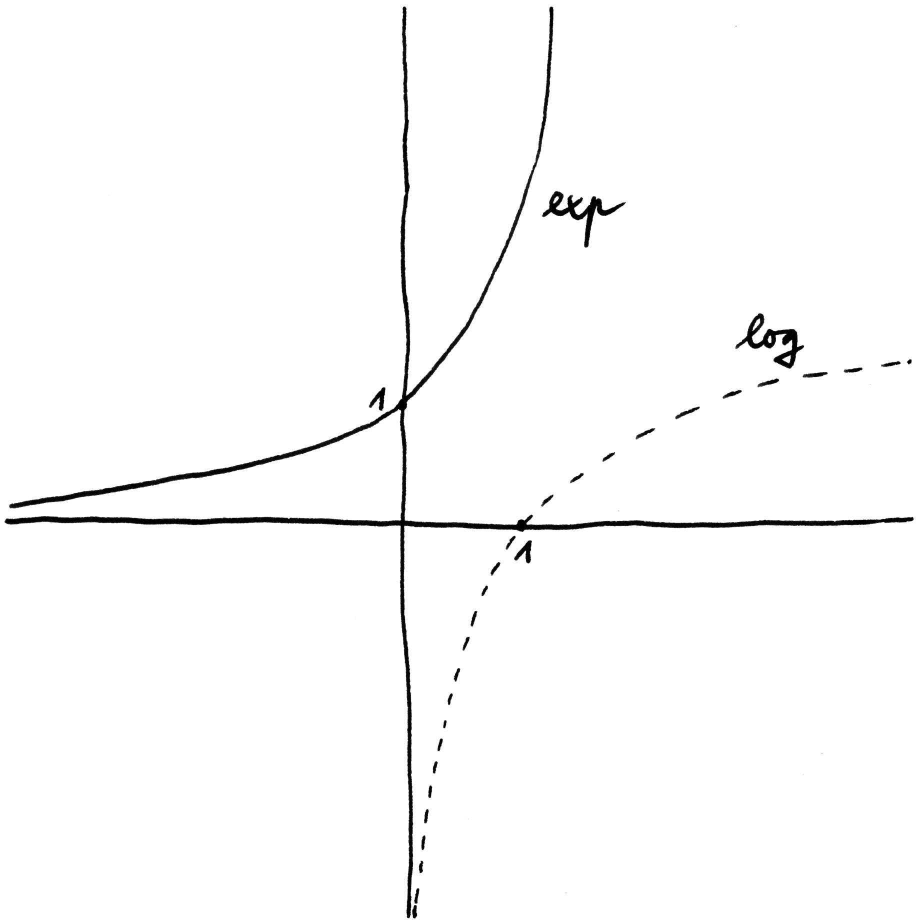
der Exponentialfunktion, $\log(\exp(x)) = x$, und erhalten aus Satz 6.2.8 die Stetigkeit des Logarithmus. In der französischen Literatur bezeichnet man diese Funktion auch als **logarithme népérien** in Erinnerung an den schottischen Mathematiker John Napier, der die ersten Logarithmentafeln aufstellte.

6.2.12. Die Exponentialfunktion liefert nach 5.6.8 und 6.2.11 einen Isomomorphismus zwischen der additiven Gruppe der reellen Zahlen und der multiplikativen Gruppe aller positiven reellen Zahlen. Daraus folgt sofort

$$\log(xy) = \log x + \log y \text{ und } \log(\mathrm{e}) = 1$$

Übung 6.2.13. Man folgere $\log(1) = 0$, $\log(x^{-1}) = -\log(x)$, $\log(x^n) = n\log(x)$ und $\log(\sqrt[q]{x}) = \frac{1}{q}\log(x)$.

6.2.14. Die Notation “log” ist leider nicht universell. Auf vielen Taschenrechnern und auch in älteren Büchern wird unsere Funktion “log” notiert als “ln” für “logarithmus naturalis” oder “logarithme Néperien” und das Kürzel “log” steht für den “Logarithmus zur Basis 10”, den wir in 6.2.17 einführen und mit $\log_{10}$ bezeichnen werden. Der Logarithmus zur Basis 10 wird in manchen Quellen auch der “Brigg’sche Logarithmus” genannt und mit “lg” bezeichnet. Die Norm ISO 31-11 empfiehlt die Notationen “ln” und “lg”, wir verwenden jedoch $\log$ statt $\ln$, weil das in der reinen Mathematik so üblich ist und wir damit der Konvention folgen, spezielle Funktionen nach Möglichkeit mit Kürzeln aus drei Buchstaben zu notieren. “Logarithmus” ist das griechische Wort für “Rechnung”, und für das Rechnen waren die Logarithmentafeln, in denen die Werte des Logarithmus zur Basis Zehn aufgelistet wurden, auch außerordentlich praktisch: Mit ihrer Hilfe konnte man nämlich Divisionen in Subtraktionen verwandeln und Wurzelziehen in Divisionen, wie wir gleich näher ausführen. Die Entdeckung der Logarithmen und die ersten Logarithmentafeln von Napier bedeuteten für die damalige Wissenschaft eine ungeheure Arbeitserleichterung und wurden begeistert begrüßt.



Die Graphen von Logarithmus und Exponentialfunktion gehen wie immer bei Umkehrfunktionen durch Spiegelung an der Hauptdiagonalen $x = y$ auseinander hervor.

Definition 6.2.15 (Allgemeine Potenzen). Gegeben $a, x \in \mathbb{R}$ mit $a > 0$ setzen wir

$$a^x = \exp(x \log a)$$

6.2.16. Man prüft ohne Schwierigkeiten die Formeln $a^0 = 1$, $a^1 = a$ und $a^{x+y} = a^x a^y$ und folgert insbesondere, daß im Fall $a > 0$ und $n \in \mathbb{Z}$ das hier definierte a^n übereinstimmt mit unserem a^n aus der Tabelle am Ende von Abschnitt 3.2.11. Für beliebige $a, b > 0$, $x, y \in \mathbb{R}$ und $q \in \mathbb{N}$, $q \geq 1$ prüft man leicht die Rechenregeln $a^{xy} = (a^x)^y$, $(ab)^x = a^x b^x$, $\sqrt[q]{a} = a^{1/q}$ und $\log(a^x) = x \log a$. Ist speziell $a = \mathrm{e}$ die Euler'sche Zahl, so gilt $\log \mathrm{e} = 1$ und folglich $\exp(x) = \mathrm{e}^x$.

Definition 6.2.17. Gegeben $a > 0$ nennt man die Umkehrabbildung der Abbildung $x \mapsto a^x$ auch den **Logarithmus zur Basis** a

$$\log_a : (0, \infty) \to \mathbb{R}$$

6.2.18. Der natürliche Logarithmus ist also der Logarithmus mit der Euler'schen Zahl e als Basis, in Formeln $\log = \log_{\mathrm{e}}$. Der Logarithmus zur Basis a läßt sich durch den natürlichen Logarithmus ausdrücken vermittels der Formel $\log_a x = \frac{\log x}{\log a}$. Man kommt deshalb meist mit dem natürlichen Logarithmus aus. Die Notation $\log_a$ ist konform mit der Norm ISO 31-11, in der zusätzlich auch noch die Abkürzung $\log_2 = \mathrm{lb}$ für den **binären Logarithmus** empfohlen wird.

Ergänzende Übung 6.2.19. Seien $a \leq b$ aus $\overline{\mathbb{R}}$ gegeben und sei $f : [a, b] \to \mathbb{R}$ stetig mit $f(b) = 0$. Man zeige, daß dann f eine kleinste Nullstelle in $[a, b]$ hat.

Übung 6.2.20. Jede Polynomfunktion $x \mapsto a_n x^n + \ldots + a_0$ mit $a_n \neq 0$ und n ungerade besitzt mindestens eine reelle Nullstelle. Ist $a_n > 0$ und $a_0 > 0$, so besitzt sie sogar mindestens eine negative Nullstelle.

Übung 6.2.21. Man zeige für $a > 0$ die Identität $\lim_{n\to\infty} n(1 - \sqrt[n]{a}) = \log a$.

Beispiel 6.2.22. Bei einem radiokativen Material wird nach einem Tag nur noch 90% der Strahlungsaktivität gemessen. Wie lange dauert es, bis die Aktivität auf die Hälfte abgeklungen ist? Nun, halten wir einen Referenzzeitpunkt beliebig fest, so wird die Aktivität nach Vergehen einer Zeitspanne τ gemäß Überlegungen wie in 5.6.4 gegeben durch eine Formel der Gestalt $M\,\mathrm{e}^{c\tau}$ mit unbekannten M und c. Wir kürzen die Zeiteinheit "Stunde" ab mit dem Buchstaben h für lateinisch "hora". Formal ist h eine Basis des in 1.2.6 angedachten eindimensionalen reellen Vektorraums "aller Zeitspannen", und formal ist auch M ein Element eines gedachten eindimensionalen reellen Vektorraums "aller Strahlungsaktivitäten" und c liegt im Raum der Linearformen auf dem Raum aller Zeitspannen, in dem wir mit h^{-1} dasjenige Element bezeichnen werden, das auf h den Wert Eins annimmt. So formal will ich hier aber eigentlich gar nicht werden und schreibe das Auswerten solch einer Linearform auf einer Zeitspanne schlicht als Produkt.

Für $\tau = 24\,\mathrm{h}$ wissen wir nun nach unserer Messung $M\,\mathrm{e}^{c\cdot 24\,\mathrm{h}} = \frac{90}{100}M$ und damit $c = ((\log 9/10)/24)\,\mathrm{h}^{-1}$. Bezeichnet t die Zeitspanne, nach der die Aktivität auf die Hälfte abgeklungen ist, so haben wir also $M\,\mathrm{e}^{ct} = M/2$ alias $ct = -\log 2$ und damit

$$t = -\frac{\log 2}{c} = \frac{-\log(2)\cdot 24}{\log(9/10)}\,\mathrm{h}$$

Ergänzende Übung 6.2.23. Das eingestrichene A liegt bei 440 Herz. Bei wieviel Herz etwa liegt das eingestrichene F? Hinweis: Die Lösung dieser Aufgabe benötigt zusätzlich zu mathematischen Kenntnissen auch physikalische und musikalische Vorkenntnisse.

Ergänzung 6.2.24. Versteht man eine positive reelle Zahl als das Verhältnis zweier gleichartiger Größen, so sagt man für den Zehnerlogarithmus besagter Zahl auch, er "drücke das Verhältnis in **Bel** aus". Diese Sprechweise ehrt Alexander Graham Bell, der das Telephon einführte. So könnte man etwa sagen, das Verhältnis von Gramm zu Kilo betrage 3 Bel, statt von einem Verhältnis von Eins zu Tausend alias 1 zu 10^3 zu reden. Das Verhältnis von Kilo zu Tonne beträgt natürlich auch 3 Bel, und das Verhältnis von Gramm zu Tonne folglich 6 Bel. Häufig redet man auch von **Dezibel** alias "zehntel Bel". Das ist jedoch "multiplikativ" zu verstehen, ein Verhältnis von 1 Dezibel meint also ein Verhältnis von Eins zu $\sqrt[10]{10} \approx 1{,}26$.

Ergänzung 6.2.25. Physikalisch beschreibt man eine Lautstärke durch die Leistung, die sie etwa an einer Membran verrichtet. Gibt man eine Lautstärke in Dezibel an, so ist allerdings das Verhältnis des Quadrats dieser Leistung zum Quadrat der Leistung einer Standardlautstärke gemeint. Diese Standardlautstärke ist so vereinbart, daß sie etwa bei der Hörschwelle eines menschliches Ohrs liegt. Eine Lautstärke von Null Dezibel bedeutet also, daß ein gesunder Mensch das Geräusch so gerade noch hören kann, und jede Erhöhung einer Lautstärke um zwanzig Dezibel bedeutet, daß das Ohr zehnmal soviel Leistung aufnehmen wird. Bei einer Erhöhung um zehn Dezibel wird das Ohr also $\sqrt{10} \approx 3$ mal soviel Leistung aufnehmen.

6.3 Grenzwerte von Funktionen

Definition 6.3.1. Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge. Ein Punkt $p \in \overline{\mathbb{R}}$ heißt ein **Häufungspunkt von** D **in** $\overline{\mathbb{R}}$ genau dann, wenn jede Umgebung von p mindestens einen von p verschiedenen Punkt mit D gemeinsam hat. Diejenigen Punkte von D, die keine Häufungspunkte von D sind, nennt man **isolierte Punkte von** D.

Ergänzung 6.3.2. Ich finde es verwirrend, daß ein Häufungpunkt von D nicht notwendig ein Punkt von D zu sein braucht. Manche Autoren erklären zusätzlich noch die "Häufungspunkte einer Folge" als die Punkte aus $\overline{\mathbb{R}}$ mit der Eigenschaft, daß in jeder ihrer Umgebungen unendlich viele Folgenglieder liegen.

Beispiele 6.3.3. Die Menge der ganzen Zahlen $\mathbb{Z}$ besteht aus isolierten Punkten und hat in $\mathbb{R}$ keine Häufungspunkte. In $\overline{\mathbb{R}}$ hat sie jedoch die beiden Häufungspunkte $\pm\infty$. In einem halboffenen, d.h. nicht aus einem einzigen Punkt bestehenden Intervall von $\overline{\mathbb{R}}$ ist jeder Punkt ein Häufungspunkt.

6.3.4. Wir vereinbaren für die Differenz einer Menge X und einer einpunktigen Menge $\{p\}$ die abkürzende Schreibweise $X\backslash\{p\} = X\backslash p$.

Übung 6.3.5. Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge. Genau dann ist $p \in \overline{\mathbb{R}}$ Häufungspunkt von D, wenn es eine Folge reeller Zahlen x_n in $D\backslash p$ gibt mit $\lim_{n\to\infty} x_n = p$.

Definition 6.3.6 (Grenzwerte von Funktionen). Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge, $p \in \overline{\mathbb{R}}$ ein Häufungspunkt von D und $f : D\backslash p \to \overline{\mathbb{R}}$ eine Funktion. Sei b ein weiterer Punkt aus $\overline{\mathbb{R}}$. Wir sagen, $f(x)$ **strebt gegen** b **für** $x \to p$ und schreiben

$$\lim_{x\to p} f(x) = b$$

genau dann, wenn es für jede Umgebung W des Grenzwerts b eine Umgebung W' des Punktes p gibt mit $f(W' \cap D\backslash p) \subset W$.

6.3.7. Man beachte, wie elegant es gelingt, hier die Fälle $\pm\infty$ mithilfe des Umgebungsbegriffs einzubinden. Natürlich reicht es auch hier wieder, die Existenz von W' für jedes W aus einem Fundamentalsystem von Umgebungen des Grenzwerts b nachzuweisen. Ist f auf ganz D definiert, so meinen wir mit obiger Notation stets implizit den Grenzwert der Einschränkung von f auf $D\backslash p$. Der Grund dafür, daß wir Grenzwerte nur an Häufungspunkten erklären, ist das folgende Lemma.

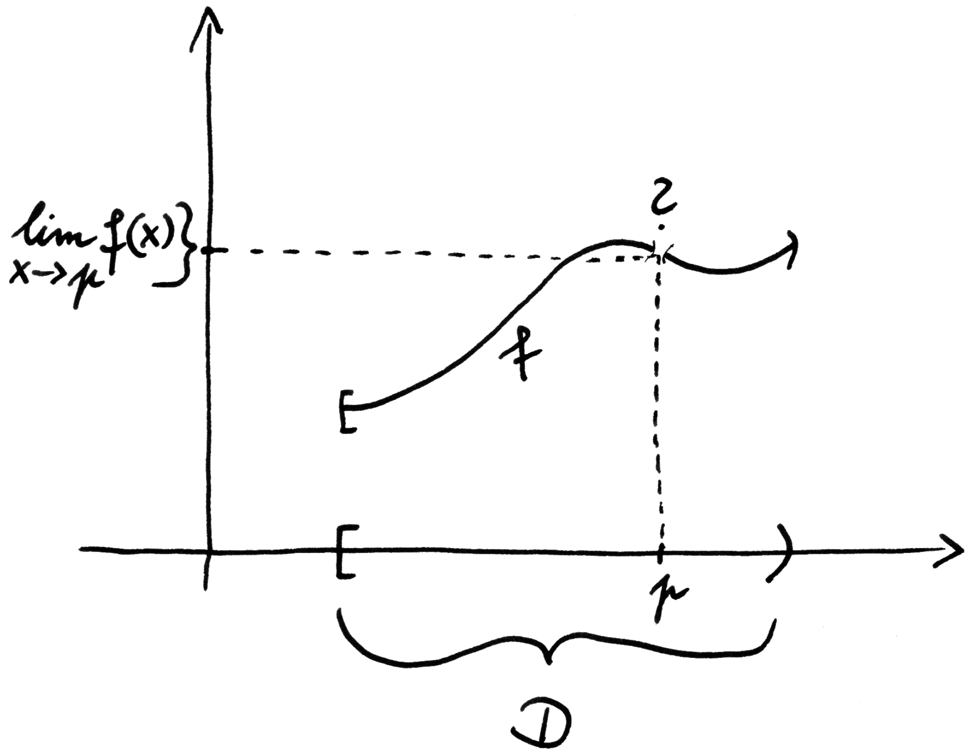
Beispiele 6.3.8. Es gilt $\lim_{x\to\infty} x = \infty$, wir können ja in diesem Fall schlicht $W' = W$ nehmen. Für eine konstante Funktion $f(x) = c$ gilt $\lim_{x\to p} c = c$, hier können wir für jedes W einfach $W' = \overline{\mathbb{R}}$ nehmen. Für $D = \mathbb{N}$ und $p = \infty$ spezialisiert der eben eingeführte Grenzwertbegriff zu unserem Grenzwertbegriff für Folgen aus 5.1.15.

Lemma 6.3.9 (Eindeutigkeit des Grenzwerts). *Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge, $p \in \overline{\mathbb{R}}$ ein Häufungspunkt von D und $f : D\backslash p \to \overline{\mathbb{R}}$ eine Funktion. Aus $\lim_{x\to p} f(x) = a$ und $\lim_{x\to p} f(x) = b$ folgt $a = b$.*

Beweis. Durch Widerspruch. Wäre $a \neq b$, so gäbe es Umgebungen V von a und W von b mit $V \cap W = \emptyset$. Wir fänden Umgebungen V' und W' von p mit $f(V' \cap D\backslash p) \subset V$ und $f(W' \cap D\backslash p) \subset W$, mithin gälte

$$f(V' \cap W' \cap D\backslash p) = \emptyset$$

Da aber $V' \cap W'$ eine Umgebung von p ist und p ein Häufungspunkt von D gilt notwendig $V' \cap W' \cap D\backslash p \neq \emptyset$. Dieser Widerspruch beendet den Beweis. $\square$



Der Grenzwert oder Limes einer Funktion mit einer “Fehlstelle” in ihrem Definitionsbereich ist, wenn er existiert, der Wert an der fraglichen Fehlstelle der einzig möglichen dort stetigen Fortsetzung.

6.3.10. Der Grenzwert einer Funktion für $x \to p$ hängt nur von ihrem Verhalten in einer Umgebung von p ab. Ist genauer $p \in \overline{\mathbb{R}}$ Häufungspunkt einer Teilmenge $D \subset \overline{\mathbb{R}}$ und $f : D\backslash p \to \overline{\mathbb{R}}$ eine Funktion und U eine Umgebung von p, so existiert $\lim_{x\to p} f$ genau dann, wenn $\lim_{x\to p} f|_{U\cap D\backslash p}$ existiert, und unter diesen Umständen stimmen die beiden Grenzwerte überein.

Proposition 6.3.11 (Grenzwerte und Stetigkeit). *Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $p \in D$ ein Häufungspunkt von D und $f : D\backslash p \to \overline{\mathbb{R}}$ eine Funktion. Genau dann gilt $\lim_{x\to p} f(x) = b$ für ein $b \in \overline{\mathbb{R}}$, wenn die Fortsetzung von f auf D durch $f(p) = b$ stetig ist bei p.*

Beweis. Das folgt sofort aus unseren Definitionen 6.1.3 und 6.3.6. □

6.3.12. In anderen Worten ist eine Funktion $f : D \to \overline{\mathbb{R}}$ stetig bei einem Häufungspunkt $p \in D$ von D genau dann, wenn gilt $\lim_{x\to p} f(x) = f(p)$. Salopp gesprochen verhält es sich also so, daß eine Funktion mit einer einpunktigen Definitionslücke, die ein Häufungspunkt ihres Definitionsbereiches ist, auf höchstens eine Weise stetig in diese Definitionslücke hinein fortgesetzt werden kann. Der Wert dieser stetigen Fortsetzung heißt dann der Grenzwert unserer Funktion an besagter Stelle.

Beispiel 6.3.13. Wir zeigen $\lim_{x\to\infty} 1/x = 0$. In der Tat, für jede Umgebung W von 0 gibt es $\varepsilon > 0$ mit $(-\varepsilon, \varepsilon) \subset W$, und nehmen wir als Umgebung W' von ∞ die Menge $W' = (1/\varepsilon, \infty]$, so gilt offensichtlich $x \in W' \Rightarrow 1/x \in W$. Alternativ können wir auch wie folgt argumentieren: Die Funktion $[0, \infty] \to [0, \infty]$ mit $x \mapsto 1/x$ für $0 < x < \infty$ und $0 \mapsto \infty$ und $\infty \mapsto 0$ ist streng monoton, deshalb ist nach 6.2.2 ihre Umkehrfunktion stetig. Die Umkehrfunktion fällt aber in diesem Fall mit der Funktion selbst zusammen. Also ist unsere Funktion stetig und mit 6.3.11 folgt $\lim_{x\to\infty} 1/x = 0$.

Übung 6.3.14 (**Rechenregeln für Grenzwerte**). Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $p \in D$ ein Häufungspunkt und seien $f, g : D\backslash p \to \mathbb{R}$ reellwertige Funktionen mit reellen Grenzwerten $\lim_{x\to p} f(x) = b$ und $\lim_{x\to p} g(x) = c$. So gilt $\lim_{x\to p}(f + g)(x) = b + c$ und $\lim_{x\to p}(fg)(x) = bc$.

Übung 6.3.15 (**Quetschlemma**). Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $p \in D$ ein Häufungspunkt und seien $f, g, h : D\backslash p \to \overline{\mathbb{R}}$ Funktionen mit $f(x) \leq g(x) \leq h(x)$ für alle $x \in D\backslash p$. So folgt aus $\lim_{x\to p} f(x) = b = \lim_{x\to p} h(x)$ schon $\lim_{x\to p} g(x) = b$.

Beispiel 6.3.16. Aus dem Quetschlemma 6.3.15 und der Darstellung von e^x durch die Exponentialreihe folgt $\lim_{x\to\infty} \mathrm{e}^x = \infty$. Aus 6.3.11 und 6.2.2 folgt dann $\lim_{y\to\infty} \log y = \infty$. Das Quetschlemma mit der Darstellung von e^x duch die Exponentialreihe liefert auch $\lim_{x\to\infty}(x/\mathrm{e}^x) = 0$ und durch Substitution $x = \log y$ und 6.3.11 und 6.1.6 folgt dann $\lim_{y\to\infty}(\log y/y) = 0$.

Übung 6.3.17. Man zeige $\lim_{x\to\infty}(x^a/b^x) = 0$ für alle $a \in \mathbb{R}$ und $b > 1$. Man zeige $\lim_{x\to\infty}(\log x/x^c) = 0$ für alle $c > 0$. Man zeige $\lim_{n\to\infty} \sqrt[n]{n} = 1$. Man zeige

$$\lim_{n\to\infty} \sqrt[n]{\frac{(5n)!}{(n!)^5}} = 5^5$$

6.3.18. Das Anwenden einer stetigen Funktion vertauscht mit der Grenzwertbildung. Ist genauer $E \subset \overline{\mathbb{R}}$ und ist $q \in E$ ein Häufungspunkt von E und $g : E \backslash q \to \overline{\mathbb{R}}$ eine Funktion mit Bild in $D \subset \overline{\mathbb{R}}$ und existiert $\lim_{x\to q} g(x) = p$ und liegt auch in D und ist $f : D \to \overline{\mathbb{R}}$ stetig bei p, so gilt

$$\lim_{x\to q} f(g(x)) = f\left(\lim_{x\to q} g(x)\right)$$

Wir erhalten diese Aussage mithilfe von 6.3.11 als direkte Konsequenz aus der Stetigkeit der Verknüpfung stetiger Funktionen 6.1.6. Speziell folgt für jede Funktion $f : D \to \overline{\mathbb{R}}$, die stetig ist bei einer Stelle $p \in D$, und jede Folge a_n in D mit $\lim_{n\to\infty} a_n = p$ bereits $\lim_{n\to\infty} f(a_n) = f(p)$.

6.3.19. Man folgert so zum Beispiel die Vertauschbarkeit 5.1.42 von Grenzwertbildung und Absolutbetrag aus der Stetigkeit des Absolutbetrags und die Vertauschbarkeit 5.1.34.2 von Grenzwertbildung mit Kehrwerten aus der Stetigkeit der Abbildung $x \mapsto 1/x$.

Beispiel 6.3.20. Für alle $a > 0$ folgt aus der Stetigkeit der Exponentialfunktion und indem man für ein logisch vollständiges Argument die folgende Gleichungskette von hinten nach vorne liest

$$\begin{array}{rcl} \lim_{n\to\infty} \sqrt[n]{a} & = & \lim_{n\to\infty} \exp\left(\frac{1}{n}\log a\right) \\ & = & \exp\left(\lim_{n\to\infty} \frac{1}{n}\log a\right) \\ & = & \exp(0) \\ & = & 1 \end{array}$$

6.3.21. Es gibt noch weitere Möglichkeiten, aus dem Zusammenhang zwischen Grenzwert und Stetigkeit 6.3.11 sowie der Stetigkeit der Verknüpfung stetiger Funktionen 6.1.6 ähnliche Aussagen abzuleiten. Zusammen mit der bereits als 6.3.18 ausführlicher besprochenen Aussage erhält man so insbesondere die drei Implikationen

$$\begin{array}{lcl} (\lim_{x\to q} g(x) = p \text{ und } f \text{ stetig bei } p) & \Rightarrow & \lim_{x\to q} f(g(x)) = f(p) \\ \left(g \text{ stetig bei } q \text{ und } \lim_{y\to g(q)} f(y) = b\right) & \Rightarrow & \lim_{x\to q} f(g(x)) = b \\ (\lim_{x\to q} g(x) = p \text{ und } \lim_{y\to p} f(y) = b) & \Rightarrow & \lim_{x\to q} f(g(x)) = b \end{array}$$

Übung 6.3.22 (**Erhaltung von Ungleichungen im Grenzwert**). Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge und $p \in D$ ein Häufungspunkt und seien $f, g : D \backslash p \to \overline{\mathbb{R}}$ zwei Funktionen, die Grenzwerte besitzen für $x \to p$. Gilt $f(x) \leq g(x)$ für alle $x \in D \backslash p$, so folgt $\lim_{x \to p} f(x) \leq \lim_{x \to p} g(x)$.

Übung 6.3.23. Gilt für eine durch ein Polynom vom Grad $\leq n$ gegebene Funktion $f(x) = a_n x^n + \ldots + a_1 x + a_0$ und irgendein reelles p die Formel $\lim_{x \to p} f(x)/(x - p)^n = 0$, so folgt $a_0 = a_1 = \ldots = a_n = 0$. Hinweis: Durch Verschieben kann man sich auf den Fall $p = 0$ zurückziehen.

Definition 6.3.24. Eine besondere Notation vereinbaren wir für den Fall einer Funktion $f : (a, b) \backslash p \to \overline{\mathbb{R}}$ mit $p \in (a, b)$, wenn wir den Grenzwert für $x \to p$ ihrer Restriktion auf (a, p) oder auf (p, b) untersuchen wollen. Wir sprechen dann vom **linkseitigen** bzw. vom **rechtseitigen Grenzwert** und notieren diese Grenzwerte

$$\lim_{x \to p} f|_{(a,p)} = \lim_{x \nearrow p} f(x) \quad \text{und} \quad \lim_{x \to p} f|_{(p,b)} = \lim_{x \searrow p} f(x)$$

In diesem Fall existiert der Grenzwert genau dann, wenn der linkseitige Grenzwert und der rechtseitige Grenzwert existieren und übereinstimmen, wie man leicht aus den Definitionen folgert.

Beispiel 6.3.25. Es gilt $\lim_{x \searrow 0} 1/x = \infty$ und $\lim_{x \nearrow 0} 1/x = -\infty$.

Satz 6.3.26 (Einparameteruntergruppen von $\mathbb{R}$). *Die stetigen Gruppenhomomorphismen $\mathbb{R} \to \mathbb{R}$ von der additiven Gruppe der reellen Zahlen in sich selber sind genau die Abbildungen $x \mapsto \lambda x$ für festes $\lambda \in \mathbb{R}$.*

Beweis. Sei $F : \mathbb{R} \to \mathbb{R}$ ein stetiger Gruppenhomomorphismus, als da heißt eine stetige Abbildung mit $F(x + y) = F(x) + F(y) \quad \forall x, y \in \mathbb{R}$. Es reicht zu zeigen, daß F die Gleichung $F(x) = xF(1)$ erfüllt. Auch ohne die Stetigkeit von F zu benutzen, folgern wir $F(q) = qF(1)$ zunächst für alle $q \in \mathbb{N}$, dann für alle $q \in \mathbb{Z}$, dann für alle $q \in \mathbb{Q}$. Um unsere Gleichung $F(x) = xF(1)$ sogar für alle $x \in \mathbb{R}$ zu zeigen, wählen wir eine Folge q_n von rationalen Zahlen mit $\lim_{n \to \infty} q_n = x$ und erhalten $F(x) = \lim_{n \to \infty} F(q_n) = \lim_{n \to \infty} q_n F(1) = xF(1)$, also $F(x) = \lambda x$ für $\lambda = F(1)$. □

Satz 6.3.27 (Einparameteruntergruppen von $\mathbb{R}^\times$). *Die stetigen Gruppenhomomorphismen $\mathbb{R} \to \mathbb{R}^\times$ von der additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der von Null verschiedenen reellen Zahlen sind genau die Abbildungen $x \mapsto a^x$ für festes $a \in \mathbb{R}_{>0}$.*

Beweis. Ist $G : \mathbb{R} \to \mathbb{R}^\times$ ein stetiger Gruppenhomomorphismus, als da heißt eine stetige Abbildung mit $G(x + y) = G(x)G(y) \quad \forall x, y \in \mathbb{R}$, so bildet G nach

3.3.16 notwendig das neutrale Element auf das neutrale Element ab, in Formeln $G(0) = 1$. Mit dem Zwischenwertsatz folgt $G(x) > 0 \quad \forall x \in \mathbb{R}$. Damit können wir den Gruppenhomomorphismus $F = \log \circ G : \mathbb{R} \to \mathbb{R}$ bilden und aus 6.3.26 folgt sofort $F(x) = xF(1)$, also $G(x) = \exp(F(x)) = \exp(xF(1)) = \exp(x \log G(1))$. □

Ergänzende Übung 6.3.28. Die monotonen Gruppenhomomorphismen $\mathbb{R} \to \mathbb{R}^\times$ von der additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der von Null verschiedenen reellen Zahlen sind genau die stetigen Gruppenhomomorphismen, also genau die Abbildungen $x \mapsto a^x$ für festes $a \in \mathbb{R}_{>0}$.

Satz 6.3.29 (Stetigkeit als Folgenstetigkeit). *Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge, $f : D \to \overline{\mathbb{R}}$ eine Funktion und $a \in D$ ein Punkt. So sind gleichbedeutend:*

1. *f ist stetig bei a.*
2. *Für jede Folge $a_0, a_1, \ldots$ von Punkten aus D mit $\lim_{n\to\infty} a_n = a$ gilt $\lim_{n\to\infty} f(a_n) = f(a)$.*

Bemerkung 6.3.30. Die in diesem Satz gegebene Charakterisierung der Stetigkeit wird vielfach sogar als Definition derselben gewählt. Das ist auch der Grund, warum ich den Satz bereits hier beweise, obwohl er im weiteren Verlauf der Vorlesung erst sehr viel später eine Rolle spielen wird.

Beweis. $1 \Rightarrow 2$ haben wir schon in 6.3.18 erledigt, wir konzentrieren uns deshalb auf $2 \Rightarrow 1$. Das zeigen wir durch Widerspruch. Für unser $a \in \overline{\mathbb{R}}$ finden wir nach 5.1.40 eine absteigende Folge von Umgebungen $V_0 \supset V_1 \supset V_2 \supset \ldots$ derart, daß jede Umgebung V von a fast alle V_n umfaßt. Ist f nicht stetig bei a, so gibt es eine Umgebung U von $f(a)$ derart, daß für kein n gilt $f(V_n \cap D) \subset U$. Für jedes n finden wir also $a_n \in V_n \cap D$ mit $f(a_n) \notin U$. Die a_n bilden dann eine Folge in D mit $\lim_{n\to\infty} a_n = a$, für die nicht gilt $\lim_{n\to\infty} f(a_n) = f(a)$. □

Ergänzende Übung 6.3.31. Man finde alle stetigen Funktionen $G : \mathbb{R}^\times \to \mathbb{R}$ mit $G(xy) = G(x) + G(y) \;\; \forall x, y \in \mathbb{R}^\times$, also alle stetigen Gruppenhomomorphismen von der multiplikativen Gruppe der von Null verschiedenen reellen Zahlen in die additive Gruppe aller reellen Zahlen.

Ergänzende Übung 6.3.32. Man zeige, daß die Aussage der vorhergehenden Sätze 6.3.26 und 6.3.27 sogar folgt, wenn wir von unseren Gruppenhomomorphismen nur die Stetigkeit bei Null fordern.

Ergänzende Übung 6.3.33. Sei $L \subset \mathbb{R}$ eine Untergruppe der additiven Gruppe der reellen Zahlen ohne reelle Häufungspunkte. So gibt es $\alpha \in \mathbb{R}$ mit $L = \mathbb{Z}\alpha$.

Übung 6.3.34. Sei $D \subset \overline{\mathbb{R}}$ eine Teilmenge, $p \in D$ ein Häufungspunkt und $f : D \backslash p \to \overline{\mathbb{R}}$ eine Funktion. So gilt $\lim_{x\to p} f(x) = b$ genau dann, wenn für jede Folge x_n in $D \backslash p$ mit $x_n \to p$ gilt $f(x_n) \to b$.

6.4 Stetige Funktionen auf Kompakta

Definition 6.4.1. Man nennt eine Teilmenge $K \subset \overline{\mathbb{R}}$ **kompakt** oder ein **Kompaktum** genau dann, wenn jede Folge in K eine Teilfolge besitzt, die gegen einen Punkt aus K konvergiert. Eine kompakte Teilmenge $K \subset \mathbb{R}$ heißt ein **reelles Kompaktum**.

6.4.2. Mit dem Satz von Bolzano-Weierstraß 5.2.9 sieht man leicht, daß die kompakten Intervalle in $\overline{\mathbb{R}}$ genau die Intervalle der Gestalt $[a, b]$ mit $a, b \in \overline{\mathbb{R}}$ sind.

Satz 6.4.3 (Extremwerte auf Kompakta). *Jede stetige Funktion auf einem nichtleeren Kompaktum nimmt das Supremum und das Infimum der Menge ihrer Funktionswerte als Funktionswert an.*

6.4.4. Ist in Formeln $K \neq \emptyset$ kompakt und $f : K \to \overline{\mathbb{R}}$ stetig, so gibt es demnach $p, q \in K$ mit $f(p) \leq f(x) \leq f(q) \quad \forall x \in K$. Ist die Funktion f reellwertig, so ist insbesondere ihr Bild beschränkt. Salopp gesprochen kann also eine stetige reellwertige Funktion auf einem Kompaktum nicht nach Unendlich streben.

Beweis. Da K nicht leer ist, finden wir eine Folge x_n in K mit $\lim_{n\to\infty} f(x_n) = \sup f(K)$. Diese Folge besitzt nun nach Annahme eine Teilfolge, die gegen einen Punkt q aus K konvergiert. Indem wir zu dieser Teilfolge übergehen dürfen wir sogar annehmen, unsere Folge sei selbst schon konvergent mit $\lim_{n\to\infty} x_n = q$. Mit 6.3.29 folgt dann

$$f(q) = \lim_{n\to\infty} f(x_n) = \sup f(K)$$

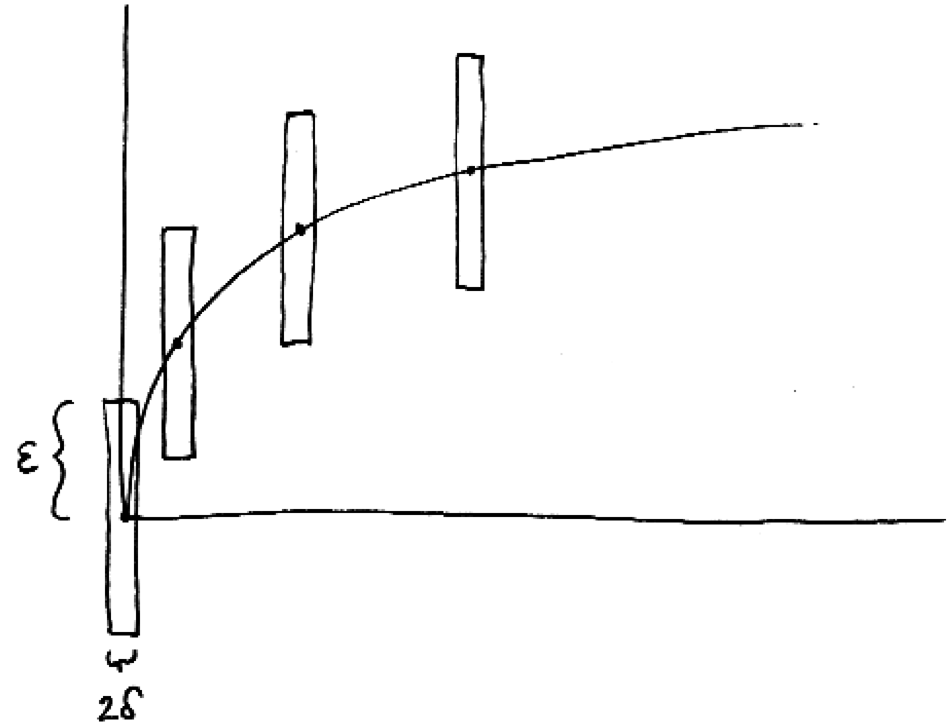
Die Existenz von $p \in K$ mit $f(p) = \inf f(K)$ zeigt man analog. □

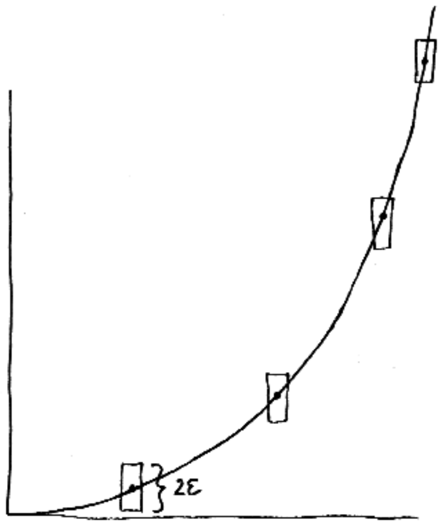
Übung 6.4.5. Jede endliche Vereinigung von Kompakta ist kompakt.

Übung 6.4.6. Das Bild eines Kompaktums unter einer stetigen Abbildung ist stets auch wieder kompakt.

Definition 6.4.7. Sei $D \subset \mathbb{R}$ eine Menge von reellen Zahlen. Eine reellwertige Funktion $f : D \to \mathbb{R}$ heißt **gleichmäßig stetig** genau dann, wenn folgende Aussage richtig ist: Für beliebiges $\varepsilon > 0$ gibt es $\delta > 0$ derart, daß für alle $x, y \in D$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.

6.4.8. Bei der Definition der gleichmäßigen Stetigkeit kommt es wesentlich auf den Definitionsbereich D an. Da wir Funktionen vielfach angeben, ohne ihren Definitionsbereich explizit festzulegen, ist es in diesem Zusammenhang oft sinnvoll, den jeweils gemeinten Definitionsbereich zu präzisieren. Dazu benutzen wir die Sprechweise f **ist gleichmäßig stetig auf** D.





Gleichmäßige Stetigkeit für stetige Funktionen auf Intervallen kann man sich anschaulich wie folgt denken: Für eine beliebig für ein Rechteck vorgegebene Höhe $2\varepsilon > 0$ findet man bei gleichmäßiger Stetigkeit immer eine Breite $2\delta > 0$ derart, daß an welchen Punkt des Graphen meiner Funktion ich das Zentrum meines Rechtecks auch verschiebe, der Graph das Rechteck nie durch die Ober- oder Unterkante verläßt. So ist etwa die Wurzelfunktion gleichmäßig stetig auf $\mathbb{R}_{\geq 0}$, die Quadratfunktion jedoch nicht.

6.4.9. Ich will nun den Unterschied zur Stetigkeit diskutieren. Eine reellwertige Funktion $f : D \to \mathbb{R}$ mit Definitionsbereich $D \subset \mathbb{R}$ heißt ja stetig bei $p \in D$ genau dann, wenn es für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, p) > 0$ gibt derart, daß für alle $x \in D$ mit $|x - p| < \delta(\varepsilon, p)$ gilt $|f(x) - f(p)| < \varepsilon$. Des weiteren heißt sie stetig, wenn sie an jeder Stelle $p \in D$ stetig ist. Gleichmäßige Stetigkeit bedeutet nun, daß für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon)$ gewählt werden kann, das es als $\delta(\varepsilon) = \delta(\varepsilon, p)$ für alle $p \in D$ gleichzeitig tut.

Beispiel 6.4.10. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist *nicht* gleichmäßig stetig auf $\mathbb{R}$, denn $|x^2 - y^2| = |x - y||x + y|$ kann auch für sehr kleines $|x - y|$ noch groß sein, wenn nur $|x + y|$ hinreichend groß ist. Die Einschränkung dieser Funktion auf ein beliebiges reelles Kompaktum ist aber daselbst gleichmäßig stetig nach dem anschließenden Satz.

Satz 6.4.11 (Gleichmäßige Stetigkeit auf Kompakta). *Jede reellwertige stetige Funktion auf einem reellen Kompaktum ist auf besagtem Kompaktum gleichmäßig stetig.*

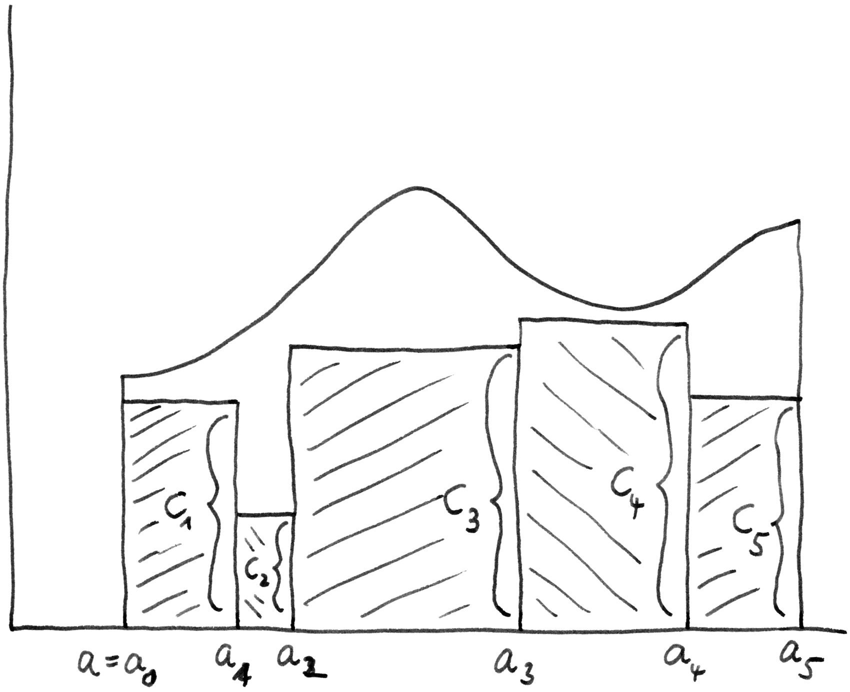
Beweis. Wir argumentieren durch Widerspruch und zeigen, daß eine Funktion auf einem reellen Kompaktum, die nicht gleichmäßig stetig ist, auch nicht stetig sein kann. Sei dazu $K \subset \mathbb{R}$ unser Kompaktum und $f : K \to \mathbb{R}$ unsere Funktion. Wäre f nicht gleichmäßig stetig, so gäbe es ein $\varepsilon > 0$, für das wir kein $\delta > 0$ finden könnten: Wir probieren alle $\delta = \frac{1}{n}$ und finden immer wieder $x_n, y_n \in K$ mit $|x_n - y_n| < \frac{1}{n}$, für die dennoch gilt $|f(x_n) - f(y_n)| \geq \varepsilon$. Gehen wir zu einer Teilfolge über, so dürfen wir ohne Beschränkung der Allgemeinheit annehmen, daß die Folge der x_n gegen einen Punkt von K konvergiert, in Formeln $\lim_{n\to\infty} x_n = x$ mit $x \in K$. Damit folgt natürlich auch $\lim_{n\to\infty} y_n = x$. Wäre nun f stetig bei x, so folgte

$$\lim_{n\to\infty} f(x_n) = f(x) = \lim_{n\to\infty} f(y_n)$$

und damit lägen notwendig fast alle $f(x_n)$ und fast alle $f(y_n)$ im Intervall $(f(x) - \varepsilon/2, f(x) + \varepsilon/2)$. Das steht jedoch im Widerspruch dazu, daß ja nach Konstruktion gilt $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle n. Wir haben also gezeigt, daß eine Funktion auf einem reellen Kompaktum, die nicht gleichmäßig stetig ist, auch nicht stetig sein kann. $\square$

6.5 Integration stetiger Funktionen

Definition 6.5.1. Sei $[a, b] \subset \mathbb{R}$ ein nichtleeres kompaktes reelles Intervall und $f : [a, b] \to \mathbb{R}$ eine stetige reellwertige Funktion. Wir definieren die Menge $I(f) \subset \mathbb{R}$ aller "naiven Integrale zu Treppen, die unter f liegen" durch



Die schraffierte Fläche stellt ein Element von $I(f)$ dar für die durch den geschwungenen Graphen dargestellte Funktion f.

$$I(f) := \left\{ \sum_{i=1}^{n} c_i(a_i - a_{i-1}) \;\middle|\; \begin{array}{l} \text{Alle möglichen Wahlen von } n \in \mathbb{N} \\ \text{und von Stellen } a = a_0 \le a_1 \le \ldots \le a_n = b \\ \text{und von Werten } c_1, \ldots, c_n \in \mathbb{R} \text{ derart,} \\ \text{daß gilt } \; f(x) \ge c_i \text{ für alle } x \in [a_{i-1}, a_i] \end{array} \right\}$$

Da f stetig ist, hat es nach 6.4.3 einen beschränkten Wertebereich, d.h. es gibt $m, M \in \mathbb{R}$ mit $m \le f(x) \le M \quad \forall x \in [a, b]$. Daraus folgt, daß $m(b-a)$ zu $I(f)$ gehört und daß $M(b-a)$ eine obere Schranke von $I(f)$ ist. Als nichtleere nach oben beschränkte Teilmenge von $\mathbb{R}$ hat $I(f)$ nach 4.4.10 ein Supremum in $\mathbb{R}$. Wir nennen dies Supremum das **Integral der Funktion f über das Intervall** $[a, b]$ und schreiben

$$\sup I(f) =: \int_a^b f(x) \;\mathrm{d}x = \int_a^b f = \int_{[a,b]} f$$

Auf die Bedeutung dieser Notationen gehen wir in 6.5.6 ein.

6.5.2. Anschaulich mißt das Integral von f die Fläche zwischen dem Graphen von f und der x-Achse, wobei Flächenstücke unterhalb der x-Achse negativ zu rechnen sind. Das Wort ist aus dem Lateinischen abgeleitet und bedeutet so etwas wie “Zusammenfassung”. Der folgende Satz listet einige Eigenschaften unseres Integrals auf. Man kann leicht zeigen, daß unser Integral sogar durch diese Eigenschaften charakterisiert wird.

Satz 6.5.3 (Integrationsregeln). *Sei $[a, b] \subset \mathbb{R}$ ein nichtleeres kompaktes Intervall.*

1. *Für die konstante Funktion mit dem Wert 1 gilt $\int_a^b 1 = b - a$.*

2. *Für $f : [a, b] \to \mathbb{R}$ stetig und $z \in [a, b]$ gilt $\int_a^b f = \int_a^z f + \int_z^b f$.*

3. *Seien $f, g : [a, b] \to \mathbb{R}$ stetig. Gilt $f(x) \le g(x) \quad \forall x \in [a, b]$, in Kurzschreibweise $f \le g$, so folgt $\int_a^b f \le \int_a^b g$.*

4. *Für $f, g : [a, b] \to \mathbb{R}$ stetig gilt $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.*

5. *Für $f : [a, b] \to \mathbb{R}$ stetig und $\lambda \in \mathbb{R}$ gilt $\int_a^b \lambda f = \lambda \int_a^b f$.*

6.5.4. Die beiden letzten Punkte bedeuten in der Sprache der linearen Algebra, daß das Integral eine Linearform auf dem reellen Vektorraum aller stetigen reellwertigen Funktionen auf unserem kompakten Intervall ist, als da heißt, eine lineare Abbildung in den Körper der reellen Zahlen.

Beweis. 1. Wir wissen ja schon, daß aus $m = 1 \leq f(x) \leq 1 = M$ folgt $b - a \leq \int_a^b f \leq b - a$.

2. Für Teilmengen $A, B \subset \mathbb{R}$ definiert man eine neue Teilmenge $A + B \subset \mathbb{R}$ durch die Vorschrift $A + B = \{x + y \mid x \in A,\ y \in B\}$. Offensichtlich gilt

$$I(f) = I(f|_{[a,z]}) + I(f|_{[z,b]})$$

Für beliebige nichtleere nach oben beschränkte Teilmengen $A, B \subset \mathbb{R}$ haben wir aber $\sup(A + B) = \sup A + \sup B$ nach Übung 4.4.11.

3. Aus $f \leq g$ folgt offensichtlich $I(f) \subset I(g)$.

4 & 5. Um die letzten beiden Aussagen zu zeigen, müssen wir etwas weiter ausholen. Für unsere stetige Funktion $f : [a, b] \to \mathbb{R}$ und beliebiges $r \in \mathbb{N}$, $r \geq 1$ unterteilen wir unser Intervall **äquidistant**, lateinisierend für "mit gleichen Abständen", durch

$$a = t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_r = b$$

Es gilt also $t_i = a + i(b - a)/r$. Wir definieren nun die **r-te Riemann-Summe** $S^r(f) \in \mathbb{R}$ durch die Vorschrift

$$S^r(f) := \sum_{i=1}^{r} f(t_i)(t_i - t_{i-1}) = \sum_{i=1}^{r} f(t_i)(\tfrac{b-a}{r})$$

In der anschließenden Proposition 6.5.5 werden wir zeigen

$$\int_a^b f = \lim_{r\to\infty} S^r(f)$$

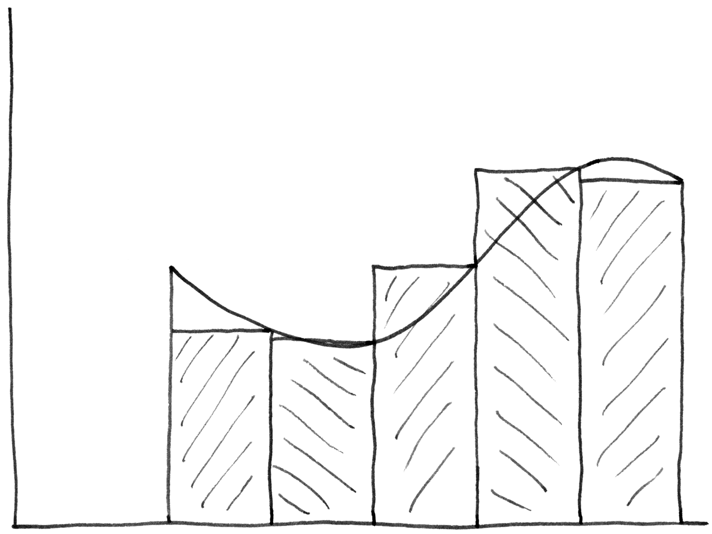
Damit erhalten wir dann sofort

$$\begin{array}{rcl} \int_a^b (f+g) & = & \lim\limits_{r\to\infty} S^r(f+g) \\ & = & \lim\limits_{r\to\infty} (S^r(f) + S^r(g)) \\ & = & \lim\limits_{r\to\infty} S^r f + \lim\limits_{r\to\infty} S^r g \\ & = & \int_a^b f + \int_a^b g \end{array}$$

und ähnlich folgt $\int_a^b \lambda f = \lambda \int_a^b f$. □

Proposition 6.5.5. *Für $a \leq b$ und $f : [a, b] \to \mathbb{R}$ stetig gilt $\int_a^b f = \lim\limits_{r\to\infty} S^r(f)$.*

Ergänzung 6.5.6. In der Notation $\int_a^b f(x)\,\mathrm{d}x$, die auf den Philosophen und Mathematiker Leibniz zurückgeht, bedeutet das Integralzeichen $\int$ ein *S* wie "Summe"



Die schraffierte Fläche stellt die fünfte Riemannsumme der durch den geschwungenen Graphen beschriebenen Funktion dar.

und $\mathrm{d}x$ meint die "Differenz im x-Wert". Manchmal verwendet man auch allgemeiner für eine weitere Funktion $g : [a, b] \to \mathbb{R}$ mit guten Eigenschaften, etwa für g die Differenz zweier monoton wachsender Funktionen, die Notation $\int f \,\mathrm{d}g$, und meint damit den Grenzwert der Summen $\sum_{i=1}^{r} f(t_i)(g(t_i) - g(t_{i-1}))$, dessen Existenz in 6.3.1 folgende in großer Allgemeinheit diskutiert wird.

6.5.7. Man mag versucht sein, die in Proposition 6.5.5 enthaltene Beschreibung gleich als Definition des Integrals zu nehmen. Ich rate davon jedoch ab, da die Existenz des fraglichen Grenzwerts nicht so leicht zu zeigen ist, und da auch die zweite unserer Integrationsregeln 6.5.3 aus dieser Definition sehr viel schlechter herzuleiten ist.

Beweis. Wir definieren zusätzlich zur r-ten Riemann-Summe die **r-ten Untersummen** und **Obersummen** durch

$$\underline{S}^r(f) := \sum_{i=1}^{r} (\inf f[t_{i-1}, t_i])(\tfrac{b-a}{r}) \qquad \text{und} \qquad \overline{S}^r(f) := \sum_{i=1}^{r} (\sup f[t_{i-1}, t_i])(\tfrac{b-a}{r})$$

und behaupten die Ungleichungen

$$\begin{array}{ccccc} \underline{S}^r(f) & \leq & S^r(f) & \leq & \overline{S}^r(f) \\ \underline{S}^r(f) & \leq & \int_a^b f & \leq & \overline{S}^r(f) \end{array}$$

Die erste Zeile ist offensichtlich. Die zweite Zeile erhalten wir zum Beispiel, indem wir aus den bereits bewiesenen Teilen 1 und 3 des Satzes die Ungleichungen

$$(\inf f[t_{i-1}, t_i])(\tfrac{b-a}{r}) \leq \int_{t_{i-1}}^{t_i} f \leq (\sup f[t_{i-1}, t_i])(\tfrac{b-a}{r})$$

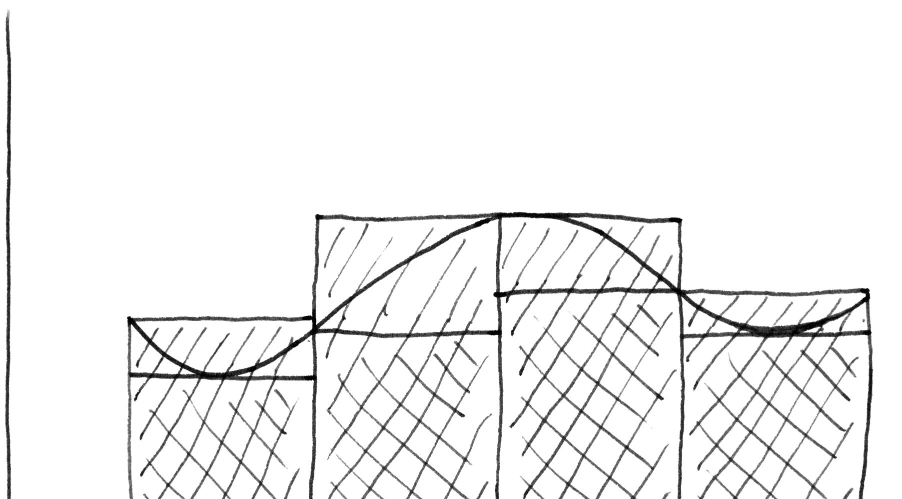
folgern, diese Ungleichungen aufsummieren, und die Mitte mit dem auch bereits bewiesenen Teil 2 des Satzes zu $\int_a^b f$ zusammenfassen. Damit sind beide Zeilen von Ungleichungen bewiesen. Nun ist f auf $[a, b]$ gleichmäßig stetig, für beliebiges $\varepsilon > 0$ existiert also $\delta = \delta_\varepsilon > 0$ mit $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Wählen wir N mit $(\frac{b-a}{N}) < \delta$ und nehmen dann $r \geq N$, so schwankt unsere Funktion f auf jedem Teilintervall $[t_{i-1}, t_i]$ höchstens um ε, mithin gilt $0 \leq \overline{S}^r(f) - \underline{S}^r(f) \leq \varepsilon(b - a)$ und damit $|S^r(f) - \int_a^b f| \leq \varepsilon(b - a)$. Die Proposition ist gezeigt. $\square$

Beispiel 6.5.8.
$$\begin{aligned} \int_0^b x \,\mathrm{d}x &= \lim_{r\to\infty} \left(\tfrac{b}{r} + \tfrac{2b}{r} + \cdots \tfrac{rb}{r}\right) \tfrac{b}{r} \\ &= \lim_{r\to\infty} \tfrac{r(r+1)}{2} \cdot \tfrac{b^2}{r^2} \qquad \text{nach 1.1.1} \\ &= \tfrac{b^2}{2} \end{aligned}$$

$$\sum_{i=1}^{n} f(a_i)\,(a_i - a_{i-1})$$

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

In der Notation $\int_a^b f(x)\,\mathrm{d}x$, die auf den Philosophen und Mathematiker Leibniz zurückgeht, bedeutet das Integralzeichen $\int$ ein *S* wie “Summe” und $\mathrm{d}x$ meint die “Differenz im x-Wert”.



Die schraffierte Fläche stellt die vierte Obersumme der durch den geschwungenen Graphen beschriebenen Funktion dar, der kreuzweise schraffierte Teil ihre Untersumme.

Lemma 6.5.9. *Für $a \leq b$ und $f : [a, b] \to \mathbb{R}$ stetig gilt die Abschätzung*

$$\left| \int_a^b f \right| \leq \int_a^b |f|$$

Beweis. Aus $-|f| \leq f \leq |f|$ folgt $-\int |f| \leq \int f \leq \int |f|$. $\square$

6.5.10. Ist $I \subset \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine stetige Funktion, und sind $a, b \in I$ gegeben mit $a > b$, so definieren wir $\int_a^b f(x)\,\mathrm{d}x$ durch die Vorschrift

$$\int_a^b f(x)\,\mathrm{d}x := -\int_b^a f(x)\,\mathrm{d}x$$

Mit dieser Konvention gilt dann für beliebige a, b, c in einem reellen Intervall I und jede stetige Funktion $f : I \to \mathbb{R}$ die Formel

$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x$$

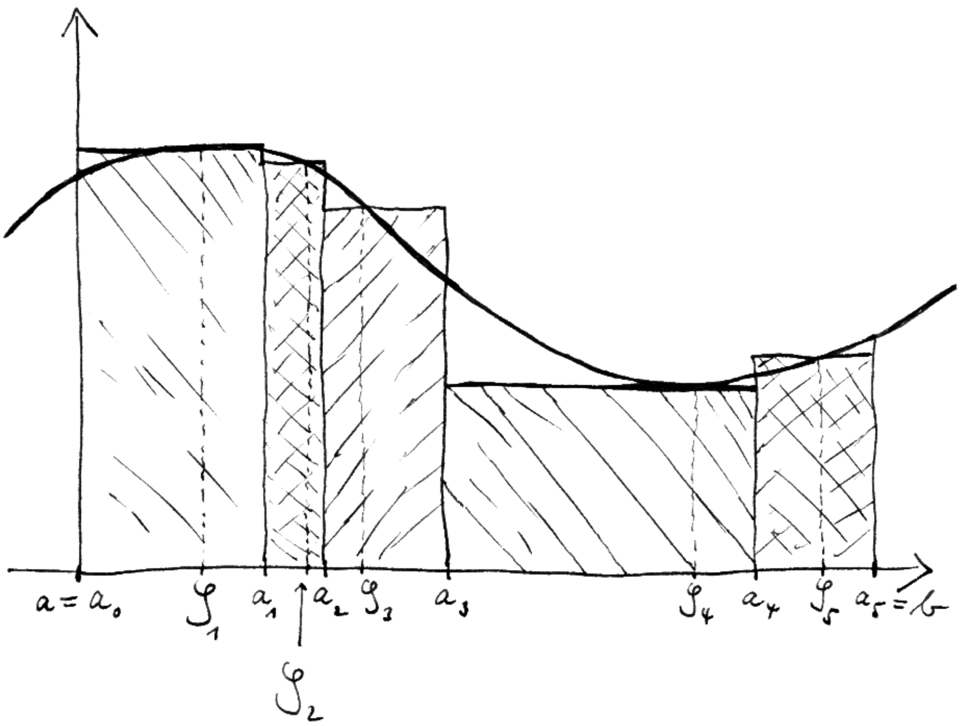
Ergänzende Übung 6.5.11. Sei $a = a_0 \leq a_1 \leq a_2 \leq \ldots \leq a_n = b$ eine **Unterteilung** des kompakten reellen Intervalls $[a, b]$. Gegeben eine Funktion $f : [a, b] \to \mathbb{R}$ bezeichnet man die Summen

$$\sum_{i=1}^{n} f(\zeta_i)(a_i - a_{i-1})$$

mit beliebigen $\zeta_i \in [a_{i-1}, a_i]$ als die **Riemann-Summen** von f zur vorgegebenen Unterteilung. Die maximale Länge $\sup\{a_i - a_{i-1} \mid 1 \leq i \leq n\}$ eines Teilintervalls heißt die **Feinheit** der Unterteilung. Man zeige: Ist $f : [a, b] \to \mathbb{R}$ stetig, so gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart, daß alle Riemann-Summen von f zu Unterteilungen der Feinheit $\leq \delta$ vom Integral $\int f$ einen Abstand $\leq \varepsilon$ haben.

Ergänzung 6.5.12. Allgemeiner heißt eine nicht notwendig stetige Funktion auf einem kompakten reellen Intervall $f : [a, b] \to \mathbb{R}$ **Riemann-integrierbar mit Integral** $\int f$ genau dann, wenn die Bedingung vom Ende der vorhergehenden Übung erfüllt ist, daß es nämlich zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, daß alle Riemann-Summen von f zu Unterteilungen der Feinheit $\leq \delta$ von $\int f$ einen Abstand $\leq \varepsilon$ haben. Wir werden jedoch in diesem Text den Begriff der Riemann-Integrierbarkeit vermeiden: Später wird eh das sehr viel stärkere Lebesgue-Integrals eingeführt, und bis dahin reicht unser Integrationsbegriff für stetige Funktionen aus.

Ergänzende Übung 6.5.13. Man zeige den **Mittelwertsatz der Integralrechnung**: Gegeben ein nichtleeres kompaktes Intervall $[a, b] \subset \mathbb{R}$ und $f : [a, b] \to \mathbb{R}$ stetig gibt es $\xi \in [a, b]$ mit $\int f = (b - a) f(\xi)$. Gegeben eine weitere stetige Funktion $g : [a, b] \to \mathbb{R}_{\geq 0}$ gibt es sogar $\xi \in [a, b]$ mit $\int fg = f(\xi) \int g$.



Darstellung einer Riemannsumme im Sinne von 6.5.11.

Ergänzende Übung 6.5.14. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt **gerade** genau dann, wenn gilt $f(-x) = f(x)$ für alle x, und **ungerade** genau dann, wenn gilt $f(-x) = -f(x)$ für alle x. Man zeige für jede ungerade stetige Funktion und alle reellen r die Formel $\int_{-r}^{r} f = 0$.

7 Differentiation und Integration

7.1 Differentiation

Definition 7.1.1. Wir nennen eine Teilmenge $D \subset \mathbb{R}$ **halboffen** genau dann, wenn sie mit jedem Punkt auch ein ganzes Intervall umfaßt, das besagten Punkt enthält und nicht nur aus diesem einen Punkt besteht.

7.1.2. Insbesondere ist also ein Intervall halboffen in diesem Sinne genau dann, wenn es nicht aus einem einzigen Punkt besteht. In der Literatur wird der Begriff "halboffen" meist abweichend verwendet für Intervalle, die weder offen noch kompakt sind, also für reelle Intervalle der Gestalt $(a, b]$ oder $[a, b)$. Bei uns heißen jedoch auch Intervalle der Gestalt $[a, b]$ mit $a < b$ halboffen, da sie eben halboffen sind als Teilmengen von $\mathbb{R}$ im Sinne der obigen Definition.

Definition 7.1.3. Sei $I \subset \mathbb{R}$ eine halboffene Teilmenge und $p \in I$ ein Punkt. Eine Funktion $f : I \to \mathbb{R}$ heißt **differenzierbar bei** p **mit Ableitung** $b \in \mathbb{R}$ genau dann, wenn gilt

$$\lim_{x \to p} \frac{f(x) - f(p)}{x - p} = b$$

Wir kürzen diese Aussage ab durch $f'(p) = b$ oder $\frac{\mathrm{d}f}{\mathrm{d}x}(p) = b$ oder, wenn die Funktion $x \mapsto f(x)$ durch einen größeren Ausdruck in x gegeben ist, durch

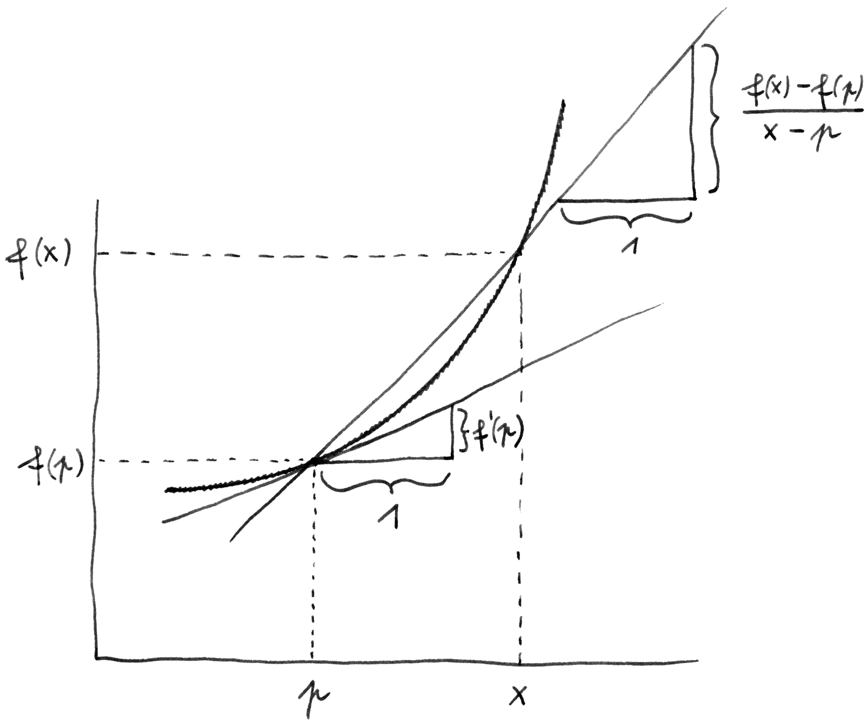
$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=p} f(x) = b$$

7.1.4. Jede nicht vertikale, d.h. nicht zur y-Achse parallele Gerade in der Ebene ist die Lösungsmenge genau einer Gleichung der Gestalt $y = a + bx$. Die reelle Zahl b heißt in diesem Fall die **Steigung** unserer Gerade. Anschaulich bedeutet der **Differenzenquotient**

$$\frac{f(x) - f(p)}{x - p}$$

die Steigung der Gerade durch die Punkte $(p, f(p))$ und $(x, f(x))$. Diese Gerade heißt auch eine **Sekante**, lateinisch für "Schneidende", da sie eben den Graphen unserer Funktion in den beiden besagten Punkten schneidet. Der Grenzwert $f'(p)$ der Sekantensteigungen bedeutet anschaulich die Steigung der **Tangente**, lateinisch für "Berührende", an den Graphen von f im Punkt $(p, f(p))$. Die Umkehrung dieser Anschauung liefert auch eine präzise Definition besagter Tangente als der Gerade durch den Punkt $(p, f(p))$ mit der Steigung $f'(p)$.

7.1.5. Wir geben noch zwei Umformulierungen der Definition der Differenzierbarkeit. Ist $I \subset \mathbb{R}$ eine halboffene Teilmenge und $p \in I$, so ist nach 6.3.11 eine



Eine Sekantensteigung und die Tangentensteigung der durch den gezahnten Graphen dargestellten Funktion

Funktion $f : I \to \mathbb{R}$ differenzierbar bei p mit Ableitung $f'(p) = b$ genau dann, wenn es eine Funktion $\varphi : I \to \mathbb{R}$ gibt, die stetig ist bei p mit Funktionswert $\varphi(p) = b$ derart, daß für alle $x \in I$ gilt

$$f(x) = f(p) + (x - p)\varphi(x)$$

Anschaulich bedeutet diese Forderung, daß die "Sekantensteigungsfunktion" $\varphi(x) = (f(x) - f(p))/(x - p)$ durch die Vorschrift $\varphi(p) = b$ stetig an die Stelle p fortgesetzt werden kann. In nochmals anderen Formeln ist unsere Funktion $f : I \to \mathbb{R}$ differenzierbar bei p mit Ableitung b genau dann, wenn gilt

$$f(p + h) = f(p) + bh + \varepsilon(h)h$$

für eine Funktion ε, die stetig ist bei Null und die dort den Wert Null annimmt. Hierbei ist zu verstehen, daß die Funktion ε definiert sein soll auf der Menge aller h mit $h + p \in I$. Diese Formulierung des Ableitungsbegriffs hat den Vorteil, besonders gut zum Ausdruck zu bringen, daß für festes p und kleines h der Ausdruck $f(p) + bh$ eine gute Approximation von $f(p + h)$ ist.

Beispiele 7.1.6. Eine konstante Funktion auf einer halboffenen Menge von reellen Zahlen ist bei jedem Punkt besagter Menge differenzierbar mit Ableitung Null. Die Funktion $\mathrm{id} : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$ hat bei jedem Punkt p die Ableitung $\mathrm{id}'(p) = \frac{\mathrm{d}x}{\mathrm{d}x}(p) = 1$.

Lemma 7.1.7. *Die Funktion $x \mapsto \frac{1}{x}$ ist differenzierbar bei jedem Punkt von $\mathbb{R}^\times$ und ihre Ableitung bei einer Stelle $p \in \mathbb{R}^\times$ ist $-\frac{1}{p^2}$.*

Beweis. Wir rechnen $\lim_{x\to p} \frac{\frac{1}{x} - \frac{1}{p}}{x-p} = \lim_{x\to p} \frac{-1}{xp} = -\frac{1}{p^2}$ nach 6.3.12. □

Lemma 7.1.8. *Sei $I \subset \mathbb{R}$ halboffen. Ist eine Funktion $f : I \to \mathbb{R}$ differenzierbar bei $p \in I$, so ist f stetig bei p.*

Beweis. Das folgt sofort aus der vorhergehenden Proposition 7.1.5. □

Definition 7.1.9. Ist $f : (a, b) \to \mathbb{R}$ definiert auf einem offenen Intervall um einen Punkt $p \in (a, b)$ und existieren die Grenzwerte

$$\lim_{x \nearrow p} \frac{f(x) - f(p)}{x - p} \quad \text{bzw.} \quad \lim_{x \searrow p} \frac{f(x) - f(p)}{x - p}$$

im Sinne von 6.3.24, so nennen wir sie die **linksseitige** bzw. die **rechtsseitige Ableitung** von f an der Stelle p.

Ergänzende Übung 7.1.10. Sei $f : (a, b) \to \mathbb{R}$ definiert auf einem offenen Intervall um einen Punkt $p \in (a, b)$. Genau dann ist f differenzierbar bei p, wenn dort die linksseitige und die rechtsseitige Ableitung existieren und übereinstimmen.

Beispiel 7.1.11. Man erkennt so zum Beispiel, daß der Absolutbetrag $\mathrm{abs} : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ nicht differenzierbar ist bei $p = 0$, da dort die linksseitige und die rechtsseitige Ableitung zwar existieren, aber nicht übereinstimmen.

7.2 Ableitungsregeln

Proposition 7.2.1. *Seien $I \subset \mathbb{R}$ halboffen und $f, g : I \to \mathbb{R}$ differenzierbar bei einem Punkt $p \in I$. So sind auch die Funktionen $f+g$ und fg differenzierbar bei p und es gilt*

$$(f+g)'(p) = f'(p) + g'(p) \quad \textit{und} \quad (fg)'(p) = f'(p)g(p) + f(p)g'(p)$$

Beweis. Wir schreiben wie in 7.1.5

$$\begin{array}{rcccccc} f(p+h) & = & f(p) & + & f'(p)h & + & \varepsilon(h)h \\ g(p+h) & = & g(p) & + & g'(p)h & + & \hat{\varepsilon}(h)h \end{array}$$

für Funktionen $\varepsilon, \hat{\varepsilon}$, die stetig sind bei Null die dort verschwinden, und erhalten durch Addieren bzw. Multiplizieren dieser Gleichungen

$$\begin{array}{rcl} (f+g)(p+h) & = & (f+g)(p) + [f'(p) + g'(p)]\, h \quad + [\varepsilon(h) + \hat{\varepsilon}(h)]\, h \\ (fg)(p+h) & = & (fg)(p) + [f'(p)g(p) + f(p)g'(p)]\, h \\ & & + [f'(p)g'(p)h + \varepsilon(h)g(p) + f(p)\hat{\varepsilon}(h) + h\varepsilon(h)\hat{\varepsilon}(h)]\, h \end{array}$$

Nach unseren Kenntnissen über stetige Funktionen steht aber in der zweiten eckigen Klammer auf der rechten Seite jeder dieser Gleichungen eine Funktion, die stetig ist bei $h = 0$ und die dort den Wert Null annimmt. □

Definition 7.2.2. Ist eine Funktion $f : I \to \mathbb{R}$ definiert auf einer halboffenen Teilmenge $I \subset \mathbb{R}$ und differenzierbar an jedem Punkt von I, so nennen wir f **differenzierbar auf** I und nennen die Funktion $f' : I \to \mathbb{R}$, $p \mapsto f'(p)$ ihre **Ableitung**.

Beispiel 7.2.3. Zum Beispiel ist $1/x$ differenzierbar auf $\mathbb{R}^\times$ mit Ableitung $-1/x^2$, und sind f und g differenzierbar, so auch $f+g$ und fg und für ihre Ableitungen gelten die **Summenregel** und die **Produktregel** oder **Leibniz-Regel**

$$(f+g)' = f' + g' \quad \text{und} \quad (fg)' = f'g + fg'$$

Korollar 7.2.4 (Ableiten ganzzahliger Potenzen). *Für alle $n \in \mathbb{Z}$ und unter der Voraussetzung $x \neq 0$ im Fall $n \leq 0$ ist die Ableitung der Funktion $x \mapsto x^n$ die Funktion $x \mapsto nx^{n-1}$.*

Beweis. Man zeigt das durch vollständige Induktion über n separat für $n \geq 0$ und $n \leq -1$. Im Fall $n = 0$ ist die Ableitung der konstanten Funktion $x^0 = 1$ natürlich überall definiert und Null, kann aber aber nur für $x \neq 0$ in der Form $0x^{-1}$ geschrieben werden. □

Satz 7.2.5 (Kettenregel). *Seien $I, J \subset \mathbb{R}$ halboffene Teilmengen und $f : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$ Funktionen und es gelte $f(I) \subset J$. Sei f differenzierbar bei p und g differenzierbar bei $f(p)$. So ist $g \circ f : I \to \mathbb{R}$ differenzierbar bei p mit Ableitung*

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

Beweis. Der besseren Übersichtlichkeit halber benutzen wir hier Großbuchstaben für die Ableitungen und setzen $f'(p) = L$ und $g'(f(p)) = M$. Wir haben

$$\begin{aligned} f(p+h) &= f(p) + Lh + \varepsilon(h)h \\ g(f(p)+k) &= g(f(p)) + Mk + \hat{\varepsilon}(k)k \end{aligned}$$

für Funktionen ε und $\hat{\varepsilon}$, die stetig sind bei Null und die dort verschwinden. Wir erhalten durch Einsetzen von $Lh + \varepsilon(h)h$ für k sofort

$$\begin{aligned} g(f(p+h)) &= g(f(p) + Lh + \varepsilon(h)h) \\ &= g(f(p)) + MLh + M\varepsilon(h)h + \hat{\varepsilon}(Lh + \varepsilon(h)h)(L + \varepsilon(h))h \end{aligned}$$

Es ist nun aber offensichtlich, daß sich hier die Summe der Terme ab dem dritten Summanden einschließlich in der Gestalt $\eta(h)h$ schreiben läßt für eine Funktion η, die stetig ist bei Null und die dort verschwindet, und wir erhalten

$$(g \circ f)(p+h) = (g \circ f)(p) + MLh + \eta(h)h$$ □

Proposition 7.2.6 (Quotientenregel). *Sei $I \subset \mathbb{R}$ eine halboffene Teilmenge, $f : I \to \mathbb{R}$ eine Funktion ohne Nullstelle und $p \in I$ ein Punkt.*

1. *Ist f differenzierbar bei p, so ist auch $1/f : I \to \mathbb{R}$, $x \mapsto 1/f(x)$ differenzierbar bei p und hat dort die Ableitung $-f'(p)/f(p)^2$.*

2. *Ist zusätzlich $g : I \to \mathbb{R}$ differenzierbar bei p, so ist auch g/f differenzierbar bei p mit Ableitung*

$$\left(\frac{g}{f}\right)'(p) = \frac{g'(p)f(p) - g(p)f'(p)}{f(p)^2}$$

Beweis. 1 folgt sofort aus 7.1.7 mit der Kettenregel. 2 folgt aus 1 mit der Produktregel. □

7.2.7. Ist $I \subset \mathbb{R}$ eine halboffene Teilmenge und sind $g, f : I \to \mathbb{R}$ differenzierbar und hat f keine Nullstelle auf I, so ist mithin auch g/f differenzierbar auf I mit Ableitung

$$\left(\frac{g}{f}\right)' = \frac{g'f - gf'}{f^2}$$

Lemma 7.2.8 (Ableitung der Exponentialfunktion). *Die Exponentialfunktion ist ihre eigene Ableitung, in Formeln* $\exp'(p) = \exp(p) \quad \forall p \in \mathbb{R}$.

Beweis. In 8.1.15 werden wir lernen, daß man "Potenzreihen gliedweise differenzieren darf". Da wir das aber bis jetzt noch nicht wissen, müssen wir etwas mehr arbeiten. Wir bestimmen zunächst die Ableitung der Exponentialfunktion an der Stelle $p = 0$ und erhalten

$$\begin{array}{rcl} \exp'(0) & = & \lim_{x\to 0} \frac{\exp(x)-\exp(0)}{x-0} \\ & = & \lim_{x\to 0} \sum_{i=1}^{\infty} \frac{x^{i-1}}{i!} \\ & = & \lim_{x\to 0} \left(1 + x \sum_{i=2}^{\infty} \frac{x^{i-2}}{i!}\right) \\ & = & 1 \end{array}$$

nach 6.3.15 und 6.3.14, da ja $\left|\sum_{i=2}^{\infty} \frac{x^{i-2}}{i!}\right|$ für $|x| \leq 1$ beschränkt ist durch die Euler'sche Zahl e. Um $\exp'(p)$ für beliebiges p zu bestimmen, rechnen wir

$$\begin{array}{rcl} \exp'(p) & = & \lim_{h\to 0} \frac{\exp(p+h)-\exp(p)}{h} \\ & = & \lim_{h\to 0} \frac{\exp(h)-1}{h} \exp(p) \\ & = & \exp(p) \end{array}$$

wo wir im letzten Schritt den schon behandelten Fall $p = 0$ verwenden und formal im ersten Schritt 6.3.21 benutzen, um den Grenzwert $x \to p$ in einen Grenzwert $h \to 0$ umzuformen. □

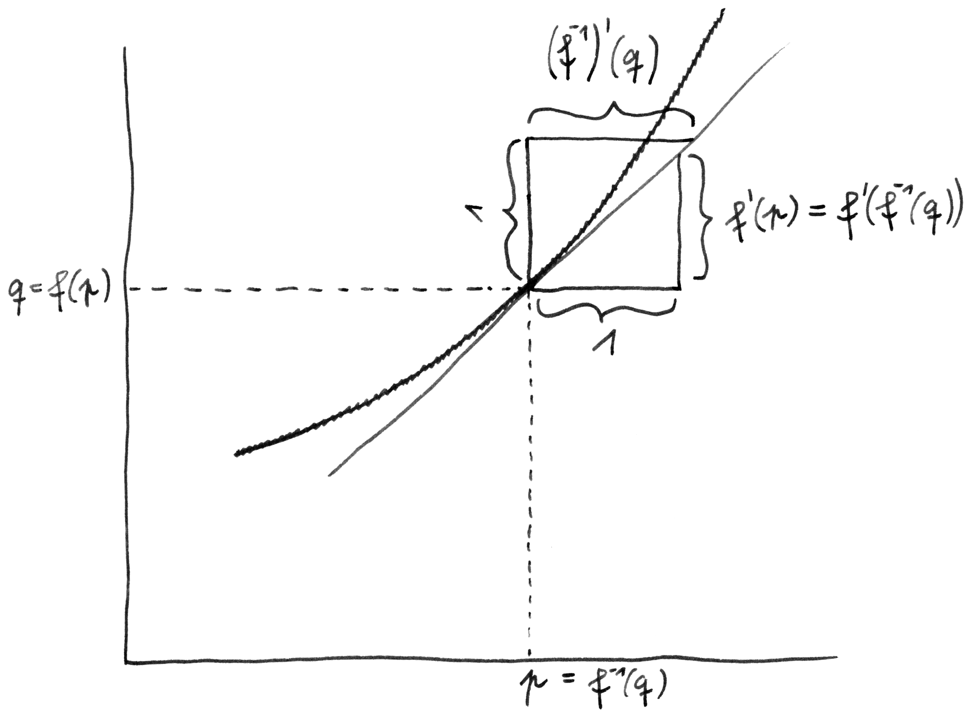
Satz 7.2.9 (Ableitung von Umkehrfunktionen). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und sei $f : I \to \mathbb{R}$ streng monoton, stetig auf I und differenzierbar bei $p \in I$ mit Ableitung $f'(p) \neq 0$. So ist auch die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ differenzierbar bei $q = f(p)$ mit Ableitung*

$$(f^{-1})'(q) = 1/f'(f^{-1}(q))$$

Beweis. Nach unseren Annahmen gibt es eine stetige Funktion ohne Nullstelle $\varphi : I \to \mathbb{R}$ mit $f(x) - f(p) = (x-p)\varphi(x)$ und $\varphi(p) = f'(p)$. Setzen wir hier $x = f^{-1}(y)$, so ist $\psi = 1/(\varphi \circ f^{-1}) : f(I) \to \mathbb{R}$ eine stetige Funktion mit $(y-q)\psi(y) = f^{-1}(y) - f^{-1}(q)$ und $\psi(q) = 1/f'(p)$. □

Beispiel 7.2.10. Die **Ableitung des Logarithmus** ist mithin

$$\log'(q) = \frac{1}{\exp(\log q)} = \frac{1}{q}$$



Veranschaulichung der Formel 7.2.9 für die Ableitung von Umkehrfunktionen

Damit ergibt sich für alle $a \in \mathbb{R}$ die **Ableitung der allgemeinen Potenzen** $(0, \infty) \to \mathbb{R}$, $x \mapsto x^a$ zu $x \mapsto ax^{a-1}$. In der Tat, nach Definition gilt ja $x^a = \exp(a \log x)$, die Ableitung wird also $a\frac{1}{x}\exp(a \log x) = ax^{a-1}$.

Lemma 7.2.11. *Die Funktion* $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) \quad = \quad \begin{cases} \mathrm{e}^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

ist beliebig oft differenzierbar.

Beweis. Wir betrachten allgemeiner für alle $n \in \mathbb{N}$ die Funktion $f_n : \mathbb{R} \to \mathbb{R}$,

$$f_n(x) \quad = \quad \begin{cases} x^{-n}\,\mathrm{e}^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

und zeigen, daß sie differenzierbar ist mit Ableitung $f_n' = -nf_{n+1} + f_{n+2}$. Damit sind wir dann natürlich fertig. Das einzige Problem ist die Ableitung an der Stelle $p = 0$, wo wir nachweisen müssen, daß die Sekantensteigungen “von rechts” auch gegen Null streben, daß also für alle $n \in \mathbb{N}$ gilt

$$\lim_{x \searrow 0} x^{-n-1}\,\mathrm{e}^{-1/x} = 0$$

Nun wissen wir aber nach der Definition von $\exp$, daß für jedes $m \in \mathbb{N}$ und $x > 0$ gilt $\exp(x) > x^m/m!$. Für jedes $n \in \mathbb{N}$ und $x > 0$ gilt also $\exp(1/x) > x^{-n-2}/(n+2)!$ und wir folgern $0 < x^{-n-1}\,\mathrm{e}^{-1/x} < (n+2)!\,x$ für $x > 0$. □

7.3 Folgerungen aus Eigenschaften der Ableitung

Definition 7.3.1. Eine Teilmenge der reellen Zahlen oder auch von $\overline{\mathbb{R}}$ heißt **offen** genau dann, wenn sie für jeden ihrer Punkte eine Umgebung ist. In Formeln ist demnach eine Teilmenge $U \subset \mathbb{R}$ offen genau dann, wenn es für jeden Punkt $p \in U$ ein $\varepsilon > 0$ gibt mit $(p - \varepsilon, p + \varepsilon) \subset U$.

Satz 7.3.2 (Notwendige Bedingung für ein Extremum). *Nimmt eine reellwertige Funktion, die auf einer offenen Teilmenge der reellen Zahlen definiert ist, an einem Punkt dieser offenen Teilmenge ihr Maximum oder ihr Minimum an und ist sie dort auch differenzierbar, so verschwindet an diesem Punkt ihre Ableitung.*

7.3.3. Die Bedingung, unsere Teilmenge sei offen, ist an dieser Stelle wesentlich: Gegeben reelle Zahlen $a < b$ nimmt etwa die Funktion $x \mapsto x$ auf dem Intervall $[a, b]$ ihr Minimum bei a und ihr Maximum bei b an, aber die Ableitung unserer Funktion verschwindet weder bei a noch bei b. Für Minima oder Maxima einer differenzierbaren Funktion $f : [a, b] \to \mathbb{R}$ kommen ganz allgemein nach unserem Satz nur in Frage: Einerseits Endpunkte des Intervalls, und andererseits die Punkte im Innern des Intervalls, an denen die Ableitung verschwindet. Mit etwas Glück können wir unter diesen Punkten dann durch Ausprobieren herauskriegen, wo das Minimum und das Maximum wirklich angenommen werden.

Beweis. Bezeichne $U \subset \mathbb{R}$ unsere offene Teilmenge und $f : U \to \mathbb{R}$ unsere Funktion. Nimmt f ein Maximum an bei $p \in U$, so gilt für die Sekantensteigungsfunktion $\varphi(x) = \frac{f(x)-f(p)}{x-p}$ offensichtlich $\varphi(x) \geq 0$ für $x < p$ und $\varphi(x) \leq 0$ für $x > p$. Wenn der Grenzwert der Sekantensteigungen existiert, so folgt mit 6.3.22 notwendig $0 \leq \lim_{x \nearrow p} \varphi(x) = \lim_{x \to p} \varphi(x) = \lim_{x \searrow p} \varphi(x) \leq 0$ und damit ist dann dieser Grenzwert Null. Nimmt f ein Minimum an bei p, so argumentiert man analog. $\square$

Übung 7.3.4. An welchen Stellen nimmt die Funktion $[-1, 2] \mapsto \mathbb{R}$ gegeben durch $x \mapsto |2 - x^2|$ ihr Minimum und Maximum an?

Beispiel 7.3.5. Das **Brechungsgesetz** behauptet, daß das Verhältnis vom Sinus des Eintrittswinkels zum Sinus des Austrittswinkels eines Lichtstrahls beim Übergang zwischen zwei Medien, sagen wir Luft und Wasser, konstant ist. Wir leiten es nun ab aus dem sogenannten **Fresnel'schen Prinzip**, nach dem ein Lichtstrahl "stets den schnellsten Weg nimmt". Ist sagen wir die Lichtgeschwindigkeit in Wasser das γ-fache der Lichtgeschwindigkeit in Luft, so sollte nach diesem Prinzip der Lichtstrahl mit den Bezeichnungen aus nebenstehendem Bild bei dem x in das Wasser eintauchen, für das der Ausdruck

$$\sqrt{a^2 + x^2} + \gamma\sqrt{b^2 + (L - x)^2}$$

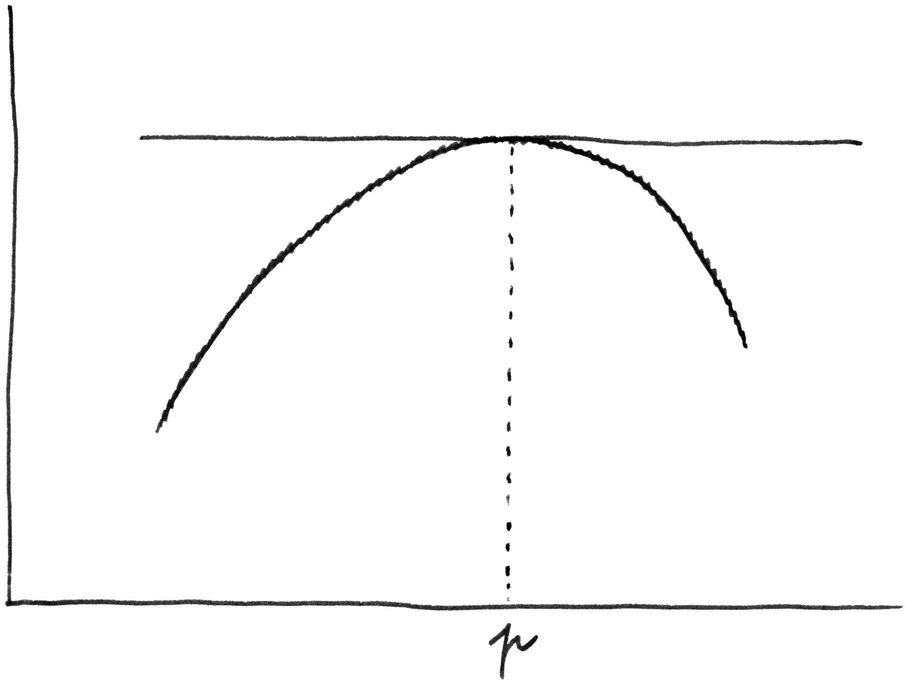
minimal wird. Ableiten liefert dafür die notwendige Bedingung

$$\frac{2x}{2\sqrt{a^2 + x^2}} + \gamma \frac{-2(L - x)}{2\sqrt{b^2 + (L - x)^2}} = 0$$

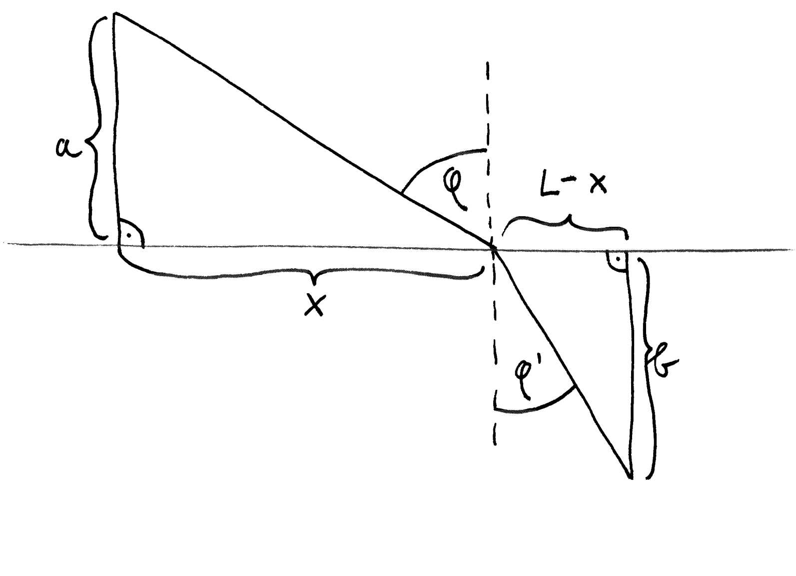
und damit steht das Brechnungsgesetz $\sin \varphi = \gamma \sin \varphi'$ auch schon da.

Ergänzende Übung 7.3.6. Bei welchem Verhältnis zwischen Durchmesser und Höhe umfaßt eine Konservendose mit fest vorgegebener Oberfläche das größtmögliche Volumen?

Satz 7.3.7 (von Rolle). *Seien $a < b$ aus $\overline{\mathbb{R}}$ gegeben und sei $f : [a, b] \to \mathbb{R}$ stetig auf dem kompakten Intervall $[a, b]$ und differenzierbar auf dem offenen Intervall (a, b). Gilt dann $f(a) = f(b)$, so gibt es $p \in (a, b)$ mit $f'(p) = 0$.*



Veranschaulichung der notwendigen Bedingung für ein Maximum 7.3.2



Zum Brechungsgesetz

Beweis. Nach 6.4.3 gibt es Punkte $p, q \in [a, b]$, an denen f sein Maximum und sein Minimum annimmt. Liegt einer dieser Punkte im Innern (a, b) unseres Intervalls, so verschwindet dort die Ableitung nach dem vorhergehenden Satz und wir sind fertig. Nimmt f sein Maximum und sein Minimum auf dem Rand des Intervalls an, so ist die Funktion f wegen unserer Annahme $f(a) = f(b)$ konstant und wir sind auch fertig. □

Korollar 7.3.8 (Mittelwertsatz). *Seien $a < b$ aus $\mathbb{R}$ gegeben und sei $f : [a, b] \to \mathbb{R}$ stetig auf dem ganzen kompakten Intervall $[a, b]$ und differenzierbar auf dem offenen Intervall (a, b). So gibt es $p \in (a, b)$ mit*

$$f'(p) = \frac{f(b) - f(a)}{b - a}$$

Beweis. Man wende den vorhergehenden Satz von Rolle 7.3.7 an auf die Funktion $g : [a, b] \to \mathbb{R}$, $g(x) = f(x) - f(a) - (x - a)\frac{f(b)-f(a)}{b-a}$, die aus f entsteht durch “Subtraktion der Sekanten”. □

Übung 7.3.9. Eine differenzierbare Funktion auf einem halboffenen Intervall, deren Ableitung beschränkt ist, ist gleichmäßig stetig.

Übung 7.3.10. Gegeben eine differenzierbare Funktion auf einem halboffenen Intervall ist das Bild des fraglichen Intervalls unter der Ableitung unserer Funktion wieder ein Intervall. Hinweis: Mittelwertsatz. Man beachte, daß die Stetigkeit der Ableitung nicht vorausgesetzt wird.

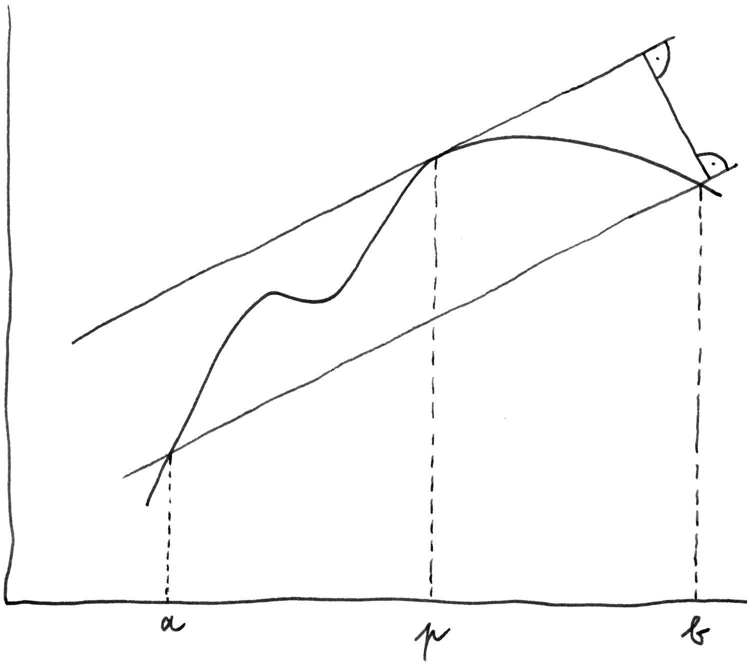
Satz 7.3.11 (Erste Ableitung und Monotonie). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und $f : I \to \mathbb{R}$ eine differenzierbare Funktion. So gilt*

$$\begin{array}{rcl} f' > 0 & \Rightarrow & f \textit{ wächst streng monoton} \\ f' \geq 0 & \Leftrightarrow & f \textit{ wächst monoton} \\ f' < 0 & \Rightarrow & f \textit{ fällt streng monoton} \\ f' \leq 0 & \Leftrightarrow & f \textit{ fällt monoton} \\ f' = 0 & \Leftrightarrow & f \textit{ ist konstant} \end{array}$$

Beweis. Es reicht, die beiden ersten Aussagen zu zeigen. Wächst f nicht streng monoton, so gibt es $a < b$ mit $f(a) \geq f(b)$ und nach dem Mittelwertsatz finden wir $p \in (a, b)$ mit

$$f'(p) = \frac{f(a) - f(b)}{a - b} \leq 0$$

Wächst f nicht monoton, so finden wir in derselben Weise $p \in I$ mit $f'(p) < 0$. Das zeigt schon mal $\Rightarrow$. Umgekehrt folgt aus f monoton wachsend, daß alle Sekantensteigungen nichtnegativ sind, und damit auch alle Grenzwerte von Sekantensteigungen. □



Veranschaulichung des Mittelwertsatzes 7.3.8

Korollar 7.3.12 (Funktionen, die ihre eigene Ableitung sind). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall. Genau dann stimmt eine differenzierbare Funktion $f : I \to \mathbb{R}$ überein mit ihrer eigenen Ableitung, wenn sie ein Vielfaches der Exponentialfunktion ist, in Formeln*

$$f' = f \quad \Leftrightarrow \quad \exists c \in \mathbb{R} \textit{ mit } f(x) = c \exp(x)$$

7.3.13. Eine differenzierbare Funktion $f : \mathbb{R}^\times \to \mathbb{R}$ mit $f' = f$ muß keineswegs ein Vielfaches der Exponentialfunktion sein. Zum Beispiel wäre die Funktion f mit $f(x) = 0$ für $x < 0$ und $f(x) = 5 \exp(x)$ für $x > 0$ auch eine Möglichkeit. Aber gut, $\mathbb{R}^\times$ ist ja auch kein Intervall.

Beweis. Bezeichne $I \subset \mathbb{R}$ unser halboffenes Intervall und $f : I \to \mathbb{R}$ unsere differenzierbare Funktion mit $f = f'$. Die Ableitung der Funktion $f(x)\,\mathrm{e}^{-x}$ ergibt sich mit der Produktregel zu $f'(x)\,\mathrm{e}^{-x} - f(x)\,\mathrm{e}^{-x} = 0$, mithin ist die Funktion $f(x)\,\mathrm{e}^{-x}$ konstant, sagen wir mit einzigem Funktionswert c, und wir folgern $f(x) = c\,\mathrm{e}^x \;\; \forall x \in I$. □

Übung 7.3.14. Sei $\alpha \in \mathbb{R}$ gegeben. Ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar und löst die Differentialgleichung $f' = \alpha f$, so gilt $f(x) = f(0)\,\mathrm{e}^{\alpha x}$ für alle $x \in \mathbb{R}$.

Satz 7.3.15 (Hinreichende Bedingung für ein lokales Extremum). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall, $f : I \to \mathbb{R}$ differenzierbar und $p \in I$ ein Punkt mit $f'(p) = 0$. Sei die Ableitung f' von f differenzierbar bei p.*

1. *Gilt $f''(p) > 0$, so besitzt f ein* **isoliertes lokales Minimum** *bei p, als da heißt, es gibt $r > 0$ derart, daß gilt $f(q) > f(p)$ für alle $q \in I$ mit $0 < |q - p| < r$.*

2. *Gilt $f''(p) < 0$, so besitzt f ein* **isoliertes lokales Maximum** *bei p.*

Beweis. Wir schreiben

$$\begin{array}{rcl} f'(x) & = & f'(p) + (x-p)\varphi(x) \\ & = & (x-p)\varphi(x) \end{array}$$

mit φ stetig in p und $\varphi(p) = f''(p) > 0$. So gibt also $r > 0$ mit $\varphi(q) > 0$ für $q \in I \cap (p - r, p + r)$, und wir folgern $f'(q) < 0$ für $q \in I \cap (p - r, p)$ und $f'(q) > 0$ für $q \in I \cap (p, p + r)$. Unseren Funktion f fällt also streng monoton auf $I \cap (p - r, p)$ und wächst streng monoton auf $I \cap (p, p + r)$. Der andere Fall $f''(p) < 0$ wird analog behandelt. □

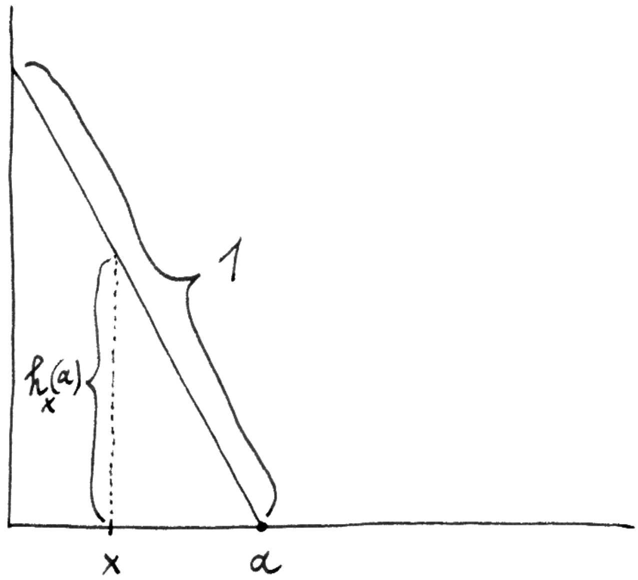


Illustration zu Übung 7.3.16.

Ergänzende Übung 7.3.16. Man zeige, daß ein Punkt $(x, y) \in (0, 1)^2$ genau dann auf einem Geradensegment der Länge Eins mit einem Ende auf der x-Achse und dem anderen Ende auf der y-Achse liegt, wenn gilt $y^{2/3}+x^{2/3} \leq 1$. Hinweis: Man halte x fest und berechne für alle $a \in [x, 1]$ die Höhe $h_x(a)$ an der Stelle x eines Brettes der Länge 1, das bei a auf der x-Achse steht und an die y-Achse angelehnt ist. Dann bestimme man das Maximum dieser Höhen bei festem x und variablem α.

Definition 7.3.17. Wir nennen eine Funktion $f : I \to \mathbb{R}$ auf einem reellen Intervall I **konvex** bzw. **konkav** genau dann, wenn ihr Graph unter bzw. über jeder seiner Sekanten liegt, wenn also in Formeln für alle $x < y < z$ aus I gilt

$$\frac{f(x) - f(y)}{x - y} \leq \frac{f(y) - f(z)}{y - z}$$

bzw. $\geq$ für konkave Funktionen.

Übung 7.3.18. Man zeige, daß diese Bedingung gleichbedeutend ist zu

$$f(tx + (1 - t)z) \leq tf(x) + (1 - t)f(z) \;\; \forall t \in [0, 1]$$

Satz 7.3.19 (Zweite Ableitung und Konvexität). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und sei $f : I \to \mathbb{R}$ eine zweimal differenzierbare Funktion. So gilt*

$$\begin{array}{lcl} f \text{ ist konvex} & \Leftrightarrow & f''(x) \geq 0 \quad \forall x \in I \\ f \text{ ist konkav} & \Leftrightarrow & f''(x) \leq 0 \quad \forall x \in I \end{array}$$

Beweis. Wir zeigen nur die erste Aussage. Ist f nicht konvex, so gibt es x, y, z mit $x < y < z$ aber

$$\frac{f(x) - f(y)}{x - y} > \frac{f(y) - f(z)}{y - z}$$

Nach dem Mittelwertsatz finden wir dann aber $\xi < \zeta$ mit $f'(\xi) > f'(\zeta)$ und bei nochmaligem Anwenden η mit $f''(\eta) < 0$. Ist umgekehrt f konvex, so reicht es nach 7.3.11 zu zeigen, daß f' monoton wächst. Kürzen wir die Steigung der Sekante durch $(x, f(x))$ und $(y, f(y))$ ab mit $s_{xy} = \frac{f(x)-f(y)}{x-y}$, so impliziert die Konvexität die Ungleichungskette

$$s_{xy} \leq s_{xz} \leq s_{yz}$$

Hier ist $s_{xy} \leq s_{yz}$ eine direkte Konsequenz der Konvexität, und da sicher gilt $(x - y)s_{xy} + (y - z)s_{yz} = (x - z)s_{xz}$, liegt s_{xz} als ein "gewichtetes Mittel" zwischen s_{xy} und s_{xz}. Unsere Ungleichungskette schreiben wir nun aus zu

$$\frac{f(x) - f(y)}{x - y} \leq \frac{f(x) - f(z)}{x - z} \leq \frac{f(y) - f(z)}{y - z}$$

Links der Graph einer konvexen, rechts der einer konkaven Funktion

Die Sekantensteigungsfunktionen $y \mapsto \frac{f(x)-f(y)}{x-y}$ und $y \mapsto \frac{f(y)-f(z)}{y-z}$ wachsen insbesondere monoton auf $(x, z]$ bzw. $[x, z)$ und im Grenzwert folgt

$$f'(x) \leq \frac{f(x) - f(z)}{x - z} \leq f'(z) \qquad \square$$

Definition 7.3.20. Wir nennen eine Funktion $f : I \to \mathbb{R}$ auf einem reellen Intervall I **streng konvex** (bzw. **streng konkav**) genau dann, wenn ihr Graph echt unter (bzw. echt über) jeder Sekante liegt, wenn also in Formeln für alle $x < y < z$ aus I gilt

$$\frac{f(x) - f(y)}{x - y} < \frac{f(y) - f(z)}{y - z}$$

bzw. $>$ für streng konkave Funktionen.

Satz 7.3.21. *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und $f : I \to \mathbb{R}$ zweimal differenzierbar. So gilt*

$$\begin{array}{lcl} f \textit{ ist streng konvex} & \Leftarrow & f''(x) > 0 \quad \forall x \in I \\ f \textit{ ist streng konkav} & \Leftarrow & f''(x) < 0 \quad \forall x \in I \end{array}$$

Beweis. Übung. $\square$

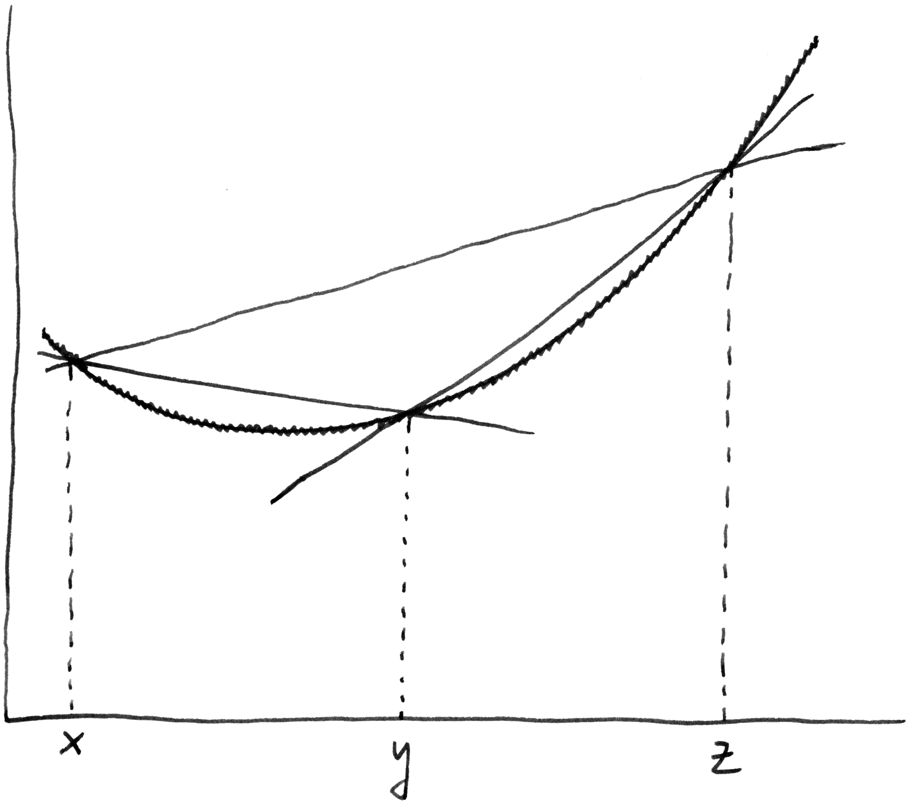
Übung 7.3.22. Gegeben reelle Zahlen $a, b \geq 0$ und $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ zeige man die **Young'sche Ungleichung** $ab \leq p^{-1}a^p + q^{-1}b^q$. Hinweis: Man gehe von der Konvexität der Exponentialfunktion aus.

Übung 7.3.23. Gegeben reelle Zahlen $a, b \geq 0$ und $p \geq 1$ zeige man die Ungleichung $(a + b)^p \leq 2^{p-1}(a^p + b^p)$. Hinweis: Man gehe von der Konvexität der Funktion $[0, \infty) \to \mathbb{R}$, $x \mapsto x^p$ aus.

Ergänzende Übung 7.3.24. Für $P \subset \mathbb{N}$ die Menge aller Primzahlen gilt

$$\sum_{p \in P} \frac{1}{p} = \infty$$

In der Tat folgt aus $\sum_{k=1}^{\infty} 1/k = \infty$ und der Eindeutigkeit der Primfaktorzerlegung und der Entwicklung in eine geometrische Reihe $(1 - 1/p)^{-1} = (1 + p^{-1} + p^{-2} \ldots)$, daß die Menge der Partialprodukte des unendlichen Produkts $\prod_{p \in P}(1 - 1/p)^{-1}$ nicht beschränkt sein kann. Nun wende man den Logarithmus an und schätze ab.



Veranschaulichung unserer Ungleichungskette für die Sekantensteigungen bei konvexen Funktionen aus dem Beweis von 7.3.19

7.4 Regeln von de l'Hospital

Satz 7.4.1 (Regeln von de l'Hospital). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und p ein Häufungspunkt von I in $\overline{\mathbb{R}}$. Seien $f, g : I \to \mathbb{R}$ differenzierbare Funktionen derart, daß gilt*

$$\lim_{x \to p} f(x) = \lim_{x \to p} g(x) = 0 \textit{ oder } \lim_{x \to p} |f(x)| = \lim_{x \to p} |g(x)| = \infty$$

Haben g und g' keine Nullstelle auf $I \backslash p$ und existiert der Grenzwert des Quotienten der Ableitungen $\lim_{x \to p}(f'(x)/g'(x))$ in $\overline{\mathbb{R}}$, so existiert auch der Grenzwert des Quotienten der Funktionen $\lim_{x \to p}(f(x)/g(x))$ in $\overline{\mathbb{R}}$ und es gilt

$$\lim_{x \to p} \left(\frac{f(x)}{g(x)} \right) = \lim_{x \to p} \left(\frac{f'(x)}{g'(x)} \right)$$

Beweis. Wir beginnen den Beweis mit folgendem

Lemma 7.4.2 (Allgemeiner Mittelwertsatz). *Seien $a < b$ in $\mathbb{R}$ gegeben und seien f, g stetige reellwertige Funktionen auf dem kompakten Intervall $[a, b]$, die differenzierbar sind auf dem offenen Intervall (a, b). So gibt es $\xi \in (a, b)$ mit*

$$f'(\xi)(g(a) - g(b)) = g'(\xi)(f(a) - f(b))$$

7.4.3. Ich kann für diesen Satz leider keine Anschauung anbieten. Verschwindet g' nirgends auf (a, b), so gilt $g(a) \neq g(b)$ nach dem Satz von Rolle 7.3.7 und wir können unsere Gleichung schreiben in der Form

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f'(\xi)}{g'(\xi)}$$

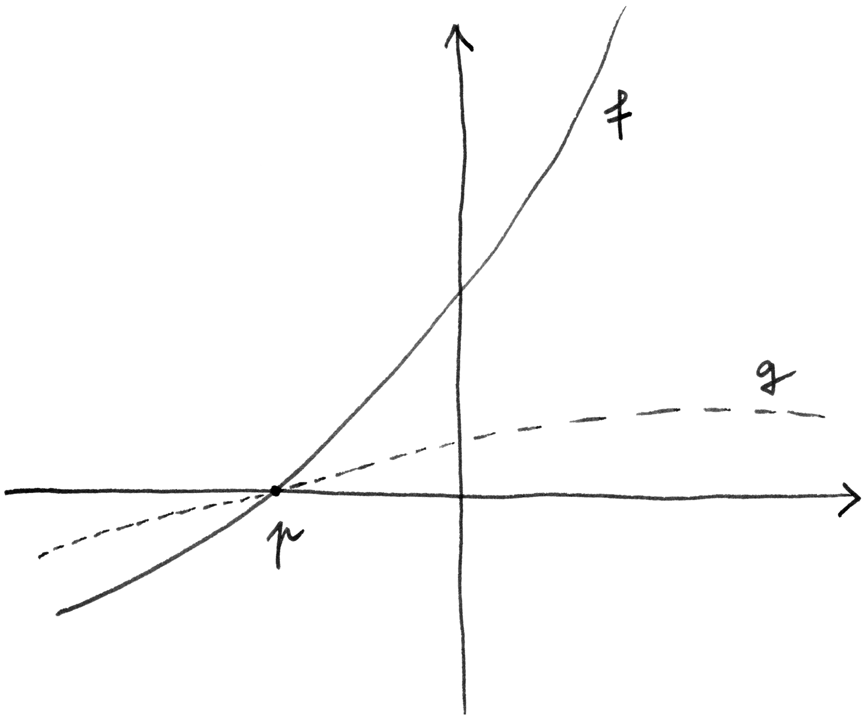
Ist also $g(x) = x$, so erhalten wir unseren Mittelwertsatz 7.3.8 als Spezialfall. Der allgemeine Mittelwertsatz wird nur beim Beweis der Regeln von de l'Hospital eine Rolle spielen, weshalb ich ihm auch nur den Status eines Lemmas eingeräumt habe.

Beweis. Man wende den Satz von Rolle 7.3.7 an auf die Funktion

$$F(x) = f(x)\,(g(a)) - g(b)) \; - g(x)\,(f(a) - f(b)) \qquad \square$$

Jetzt zeigen wir die Regeln von de l'Hospital. Wir dürfen ohne Beschränkung der Allgemeinheit $p \notin I$ annehmen, indem wir sonst I an der Stelle p in zwei Teile zerschneiden und den rechtsseitigen und linksseitigen Grenzwert bei p getrennt betrachten. Für jede Umgebung W des Grenzwerts

$$q = \lim_{x \to p} f'(x)/g'(x)$$



Veranschaulichung der Regel von de l'Hospital 7.4.1 im Fall $p \in \mathbb{R}$ unter der Annahme, daß beide Funktionen bei p nach Null streben. Es scheint mir anschaulich klar, daß der Grenzwert des Quotienten sich nicht ändert, wenn wir beide Funktionen durch ihre beste lineare Approximation bei p ersetzen, und das ist auch genau die anschauliche Bedeutung der Regel von de l'Hospital.

finden wir nun per definitionem eine Intervallumgebung V von p mit $\xi \in I \cap V \Rightarrow f'(\xi)/g'(\xi) \in W$. Für beliebige $a, b \in I \cap V$ mit $a \neq b$ gilt dann $g(a) \neq g(b)$, da nach Annahme die Ableitung von g keine Nullstelle auf $I \backslash p$ hat, und aus dem allgemeinen Mittelwertsatz folgt dann weiter

$$\frac{f(a) - f(b)}{g(a) - g(b)} \in W$$

Von nun an müssen wir die beiden Fälle im Satz getrennt weiterbehandeln. Zunächst nehmen wir an, es gelte $\lim_{x \to p} f(x) = \lim_{x \to p} g(x) = 0$. Ist W ein kompaktes Intervall, so folgt $f(a)/g(a) \in W$ sofort, indem wir a festhalten, b gegen p streben lassen und uns an die Erhaltung von Ungleichungen im Grenzwert erinnern. Die Behauptung im ersten Fall folgt dann, da für jeden Punkt seine kompakten Intervallumgebungen ein Fundamentalsystem von Umgebungen bilden. Jetzt behandeln wir noch den Fall $\lim_{x \to p} |f(x)| = \lim_{x \to p} |g(x)| = \infty$. In diesem Fall dürfen wir ohne Beschränkung der Allgemeinheit auch annehmen, daß f auf I keine Nullstelle hat, so daß gilt

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f(b)}{g(b)} \left(\frac{1 - f(a)/f(b)}{1 - g(a)/g(b)} \right)$$

Sei nun $a \in I \cap V$ fest gewählt. Für jedes und insbesondere auch für sehr nahe bei Eins liegendes $\alpha \in (0, 1)$ finden wir dann eine Umgebung U von p derart, daß gilt

$$b \in U \cap I \quad \Rightarrow \quad \frac{1 - f(a)/f(b)}{1 - g(a)/g(b)} \in (\alpha, \alpha^{-1})$$

Indem wir zusätzlich $U \subset V$ wählen, finden wir damit für jede Umgebung W von q und jedes $\alpha \in (0, 1)$ eine Umgebung U von p mit $b \in U \cap I \Rightarrow f(b)/g(b) \in (\alpha, \alpha^{-1})W$. Die Behauptung folgt nun im zweiten Fall, da es für jede Umgebung W' von q eine Umgebung W von q und ein $\alpha \in (0, 1)$ gibt mit $(\alpha, \alpha^{-1})W \subset W'$. □

Beispiel 7.4.4. Für $\lambda > 0$ gilt

$$\begin{array}{rcl} \lim_{x \to \infty} (\log x)/x^\lambda & = & \lim_{x \to \infty} (1/x)/\lambda x^{\lambda - 1} \\ & = & \lim_{x \to \infty} 1/\lambda x^\lambda \\ & = & 0 \end{array}$$

7.5 Zusammenhang zwischen Integral und Ableitung

Satz 7.5.1 (Hauptsatz der Differential- und Integralrechnung). *Ist $I \subset \mathbb{R}$ ein halboffenes Intervall, $f : I \to \mathbb{R}$ stetig und $a \in I$ ein Punkt, so ist die Funktion*

$$\begin{array}{rccl} F : & I & \to & \mathbb{R} \\ & x & \mapsto & \int_a^x f(t) \, \mathrm{d}t \end{array}$$

die einzige differenzierbare Funktion $F : I \to \mathbb{R}$ mit $F' = f$ und $F(a) = 0$.

7.5.2. Im Fall $x < a$ ist dies Integral unter Verwendung unserer Konvention 6.5.10 als $\int_a^x f(t)\ \mathrm{d}t = -\int_x^a f(t)\,\mathrm{d}t$ zu interpretieren.

Beweis. Für $p \in I$ rechnen wir

$$\lim_{x\to p} \frac{F(x)-F(p)}{x-p} = \lim_{x\to p} \frac{1}{x-p}\int_p^x f(t)\ \mathrm{d}t$$

Da nun f stetig ist, finden wir für beliebiges $\varepsilon > 0$ ein $\delta > 0$ derart, daß aus $|t-p| \leq \delta$ folgt $f(p)-\varepsilon \leq f(t) \leq f(p)+\varepsilon$. Aus $0 < |x-p| \leq \delta$ folgen also durch Bilden des Integrals und Teilen durch $(x-p)$ die Ungleichungen

$$f(p)-\varepsilon \ \leq\ \frac{1}{x-p}\int_p^x f(t)\ \mathrm{d}t \ \leq\ f(p)+\varepsilon$$

und damit ist gezeigt, daß der Ausdruck in der Mitte für $x \to p$ gegen $f(p)$ konvergiert. Es folgt $F'(p) = f(p)$ für alle $p \in I$. Ist $G : I \to \mathbb{R}$ auch differenzierbar mit $G' = f$, so verschwindet die Ableitung der Differenz $G - F$. Nach 7.3.11 ist diese Differenz also konstant, und haben wir dann auch noch $G(a) = 0$, so ist sie konstant Null und wir folgern $F = G$. □

Korollar 7.5.3 (Integrieren mit Stammfunktionen). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Ist $G : I \to \mathbb{R}$ eine* **Stammfunktion** *von f, als da heißt eine differenzierbare Funktion mit Ableitung $G' = f$, so gilt für alle $a, b \in I$ die Formel*

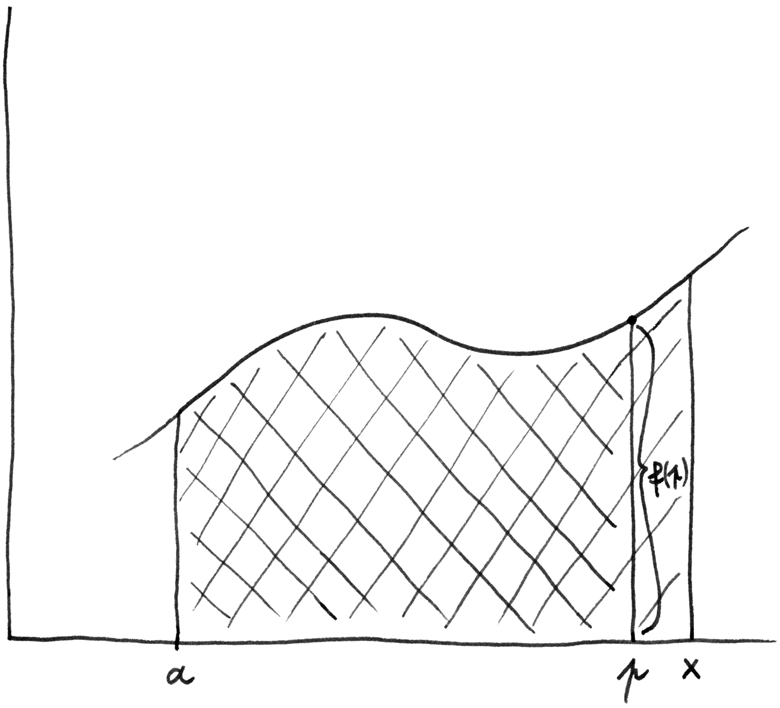
$$\int_a^b f(t)\ \mathrm{d}t = G(b) - G(a)$$

Beweis. Wir betrachten $F : I \to \mathbb{R}$ wie im Satz und folgern aus der Eindeutigkeitsaussage von dort $F(x) = G(x) - G(a)$ für alle $x \in [a, b]$. □

7.5.4. Ist G ein komplizierter Ausdruck, so ist es bequem und üblich, die Differenz $G(b) - G(a)$ mit $G(x)|_a^b$ abzukürzen. Man spricht einen solchen Ausdruck "G, ausgewertet zwischen den Grenzen a und b". Für a, b positiv ergibt sich zum Beispiel

$$\int_a^b \frac{1}{x}\ \mathrm{d}x = \log x|_a^b = \log b - \log a$$

Die Ableitung von $\mathbb{R}^\times \to \mathbb{R}$, $x \mapsto \log|x|$ ist im Übrigen $x \mapsto 1/x$, als da heißt, für a, b negativ würden wir $\log|b| - \log|a|$ erhalten. Über den Nullpunkt hinweg dürfen wir die Funktion $1/x$ aber natürlich trotzdem nicht in dieser Weise integrieren, und unser Korollar erlaubt das auch nicht, es trifft vielmehr nur Aussagen über die Integration von auf einem Intervall definierten stetigen Funktionen.



Veranschaulichung des Hauptsatzes der Differential- und Integralrechnung 7.5.1. Die kreuzweise schraffierte Fläche stellt $F(p)$ dar, die irgendwie schraffierte $F(x)$, die einfach schraffierte $F(x) - F(p)$. Ich hoffe, man sieht, daß die Fläche unter der Kurve beim Verschieben der oberen Grenze um so stärker wächst, je größer dort der Wert unserer Funktion ist. Das ist qualitativ ausgedrückt die anschauliche Bedeutung des Hauptsatzes.

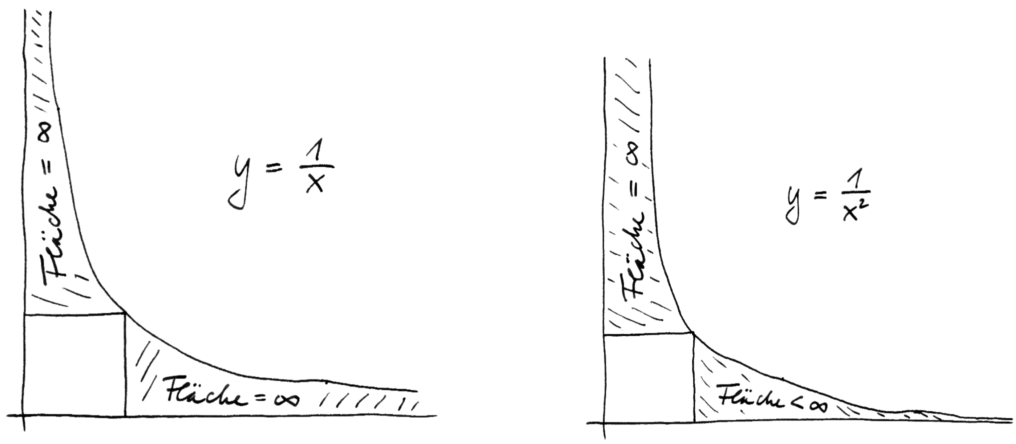


Illustration zu Übung 7.5.5.

Übung 7.5.5. Gegeben $\alpha \in \mathbb{R}$ zeige man, daß $\lim_{t\to\infty} \int_1^t x^\alpha \,\mathrm{d}x$ existiert in $\overline{\mathbb{R}}$ und daß dieser Grenzwert endlich ist genau dann, wenn gilt $\alpha < -1$. Des weiteren zeige man, daß $\lim_{\varepsilon\to 0} \int_\varepsilon^1 x^\alpha \,\mathrm{d}x$ existiert in $\overline{\mathbb{R}}$ und daß dieser Grenzwert endlich ist genau dann, wenn gilt $\alpha > -1$. Anschaulich gesprochen ist also die Hyperbel $x \mapsto (1/x)$ gerade der Grenzfall, in dem sowohl die Fläche zwischen Kurve und x-Achse ab jedem x-Wert als auch symmetrisch die Fläche zwischen Kurve und y-Achse ab jedem y-Wert unendlich groß sind.

Ergänzende Übung 7.5.6. Der **Integrallogarithmus** $\mathrm{Li} : (1,\infty) \to \mathbb{R}$ wird erklärt durch die Vorschrift

$$\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t}$$

Man zeige $\lim_{x\to\infty}(x \,\mathrm{Li}(x)/\log(x)) = 1$.

7.6 Integrationsregeln

Satz 7.6.1 (Integration durch Substitution). *Gegeben zwei reelle Zahlen $a < b$ und $g : [a,b] \to \mathbb{R}$ stetig differenzierbar und $f : g([a,b]) \to \mathbb{R}$ stetig gilt*

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(y)\,\mathrm{d}y$$

wobei im Fall $g(b) < g(a)$ das Integral rechts der Konvention 6.5.10 gemäß als das Negative des Integrals über das Intervall $[g(b), g(a)]$ zu verstehen ist.

7.6.2. Es gibt durchaus differenzierbare Abbildungen, deren Ableitung nicht stetig ist. Ein Beispiel findet man erst in 10.6.6, da wir vorher den Sinus noch nicht zur Verfügung haben, aber davon abgesehen kann es auch hier schon verstanden werden.

Beweis. Ist g konstant, so verschwinden beide Seiten und die Formel gilt. Ist sonst F eine Stammfunktion von f, so ist $F \circ g$ nach der Kettenregel 7.2.5 eine Stammfunktion von $t \mapsto f(g(t))g'(t)$, und berechnen wir beide Integrale mithilfe dieser Stammfunktionen, so ergibt sich auf beiden Seiten der Wert $F(g(b)) - F(g(a))$. □

7.6.3. Wächst g streng monoton, so kann man auch leicht einen anschaulicheren Beweis geben, indem man $a = a_0 < a_1 < \ldots < a_r = b$ äquidistant unterteilt und nach dem Mittelwertsatz $\xi_i \in (a_{i-1}, a_i)$ findet mit $g(a_i) - g(a_{i-1}) = g'(\xi_i)(a_i - a_{i-1})$. Damit gilt ja die Gleichheit von Riemannsummen

$$\sum_{i=1}^r f(g(\xi_i))g'(\xi_i)(a_i - a_{i-1}) = \sum_{i=1}^r f(g(\xi_i))(g(a_i) - g(a_{i-1}))$$

Mithilfe von 6.5.11 zeigt man dann, daß diese Gleichheit im Grenzübergang $r \to \infty$ gerade die Substitutionsregel liefert.

7.6.4. Als Gedächtnisstütze, die in 6.3 auch echte Bedeutung erhält, kann man sich merken, daß beim Substituieren, lateinisch ebenso wie "Prostituieren" für "Ersetzen", von y durch $g(x)$ nicht nur $f(y)$ durch $f(g(x))$ ersetzt werden muß, sondern auch $\mathrm{d}y$ durch $g'(x)\,\mathrm{d}x$, wie es die Notation $g'(x) = \frac{\mathrm{d}y}{\mathrm{d}x}$ suggeriert.

Beispiel 7.6.5. Indem wir uns das Integral erst etwas zurechtlegen, erhalten wir durch die Substitution $g(x) = x^2 = y$, $g'(x) = 2x\,\mathrm{d}x = \mathrm{d}y$ leicht

$$\int_a^b \frac{x\,\mathrm{d}x}{\sqrt{1+x^2}} = \frac{1}{2}\int_a^b \frac{2x\,\mathrm{d}x}{\sqrt{1+x^2}} = \frac{1}{2}\int_{a^2}^{b^2} \frac{\mathrm{d}y}{\sqrt{1+y}} = \sqrt{1+y}\,\Big|_{a^2}^{b^2} = \sqrt{1+x^2}\,\Big|_a^b$$

und als eine Stammfunktion des ursprünglichen Integranden ergibt sich die Funktion $\sqrt{1+x^2}$.

7.6.6. In Anwendungen wird man oft mit einer umkehrbaren Abbildung g arbeiten und die Regel in der Form

$$\int_\alpha^\beta f(y)\,\mathrm{d}y = \int_{g^{-1}(\alpha)}^{g^{-1}(\beta)} f(g(x))g'(x)\,\mathrm{d}x$$

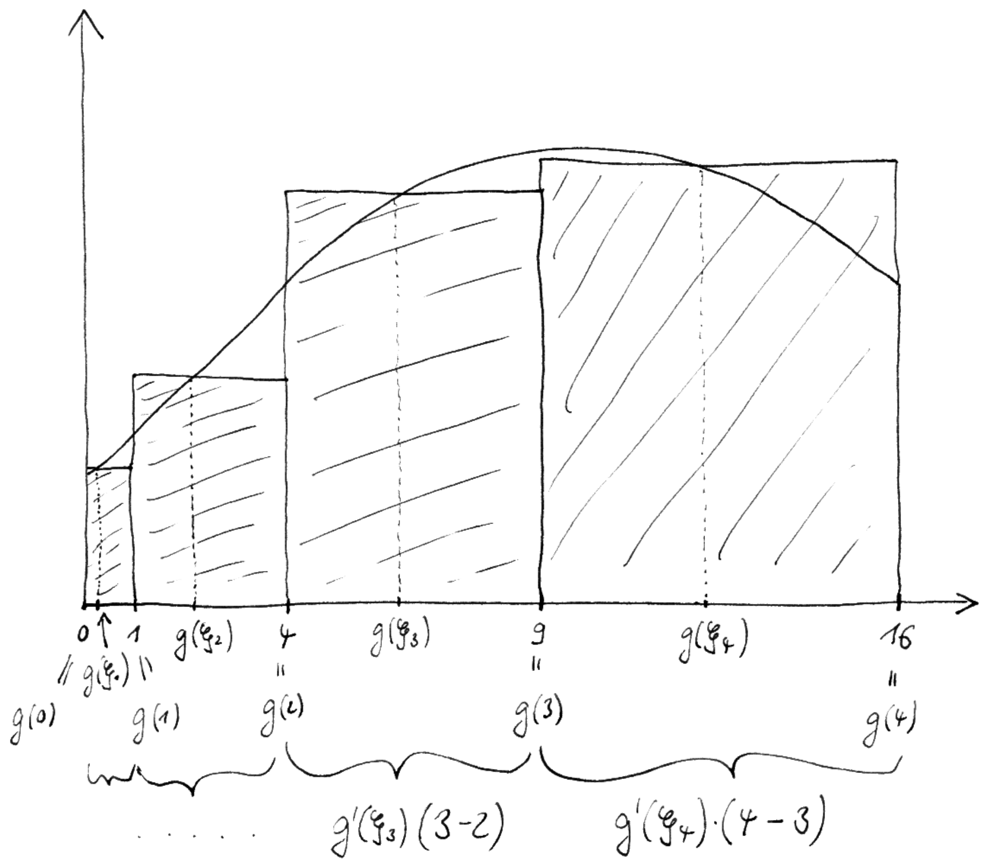
verwenden. Auch wenn g nicht umkehrbar ist, kann man in dieser Weise vorgehen, statt $g^{-1}(\alpha)$ und $g^{-1}(\beta)$ darf man dann eben irgendwelche a, b wählen derart, daß g auf ganz $[a, b]$ stetig differenzierbar ist und daß gilt $\alpha = g(a)$ und $\beta = g(b)$. Meist arbeiten wir sogar ohne explizite Erwähnung der Grenzen: Suchen wir nur eine Stammfunktion, so kommt es auf die untere Grenze eh nicht an und die obere Grenze wird mitgedacht und nach gelungener Integration wieder "zurücksubstituiert", wie es unsere Formel fordert.

Beispiel 7.6.7. Wir bestimmen eine Stammfunktion von $f(y) = y\sqrt{1-y}$ durch die Substitution $\sqrt{1-y} = x$, $y = 1 - x^2 = g(x)$, $\mathrm{d}y = -2x\,\mathrm{d}x$ zu

$$\begin{array}{rcl}
\int y\sqrt{1-y}\ \mathrm{d}y & = & \int (1-x^2)x(-2x)\ \mathrm{d}x \\
& = & \int (2x^4 - 2x^2)\ \mathrm{d}x \\
& = & \frac{2}{5}x^5 - \frac{2}{3}x^3 \\
& = & \frac{2}{5}(1-y)^2\sqrt{1-y} - \frac{2}{3}(1-y)\sqrt{1-y}
\end{array}$$

Satz 7.6.8 (Partielle Integration). *Gegeben reelle Zahlen $a < b$ und zwei stetig differenzierbare Funktionen $f, g : [a, b] \to \mathbb{R}$ gilt*

$$\int_a^b fg' = fg|_a^b - \int_a^b f'g$$



Anschauliche Bedeutung der Substitutionsregel nach 7.6.3 im Fall $g(x) = x^2$. Die schraffierte Fläche stellt in diesem Spezialfall für $r = 4$ und das Intervall $[a, b] = [0, 4]$ die Summe

$$\sum_{i=1}^{r} f(g(\xi_i))g'(\xi_1)(a_i - a_{i-1}) = \sum_{i=1}^{r} f(g(\xi_i))(g(a_i) - g(a_{i-1}))$$

dar. Die ξ_i sind mit dem Mittelwertsatz gerade so gewählt, daß gilt $g(a_i) - g(a_{i-1}) = g'(\xi_i)(a_i - a_{i-1})$. Ich finde, man sieht sehr gut, daß diese Summen gegen $\int_{g(a)}^{g(b)} f(y)\,\mathrm{d}y = \int_0^{16} f(y)\,\mathrm{d}y$ konvergieren.

Beweis. Nach der Produktregel gilt $fg' = (fg)' - f'g$ auf $[a, b]$, folglich stimmen auch die Integrale dieser Funktionen überein. $\square$

Beispiele 7.6.9.
$$\begin{aligned}\int x^2\,\mathrm{e}^x\;\mathrm{d}x &= x^2\;\mathrm{e}^x - \int 2x\,\mathrm{e}^x\;\mathrm{d}x\\ &= x^2\,\mathrm{e}^x - 2x\,\mathrm{e}^x + \int 2\,\mathrm{e}^x\;\mathrm{d}x\\ &= (x^2 - 2x + 2)\,\mathrm{e}^x\end{aligned}$$

$$\begin{aligned}\int \log x\;\mathrm{d}x &= x\log x - \int x\tfrac{1}{x}\;\mathrm{d}x\\ &= x\log x - x\end{aligned}$$

7.6.10. Ich erkläre nun noch einen alternativen Zugang zur Exponentialfunktion, der unsere Definition über eine a priori unmotivierte Reihe vermeidet. Suchen wir eine stetig differenzierbare Funktion $g : \mathbb{R} \to \mathbb{R}_{>0}$ mit $g' = g$ und $g(0) = 1$, so haben wir ja nach der Substitutionsregel

$$x = \int_0^x \mathrm{d}t = \int_0^x \frac{g'(t)}{g(t)}\,\mathrm{d}t = \int_1^{g(x)} \frac{1}{u}\,\mathrm{d}u$$

Man definiert also notgedrungen eine Funktion $\log : \mathbb{R}_{>0} \to \mathbb{R}$ durch die Vorschrift $\log(y) = \int_1^y \frac{1}{u}\,\mathrm{d}u$, und die gesuchte Funktion g muß deren Umkehrfunktion sein.

Ergänzende Übung 7.6.11. Wie könnte ein Autor, der den Zugang 7.6.10 zur Exponentialfunktion gewählt hat, die Funktionalgleichung 5.6.8 beweisen?

Ergänzende Übung 7.6.12. Man zeige durch Vergleich mit dem Integral der Funktion $x^{-\alpha}$, daß für jedes $\alpha > 1$ die Reihe $\sum_{k=1}^{\infty} k^{-\alpha}$ konvergiert.

7.7 Hyperbolische trigonometrische Funktionen

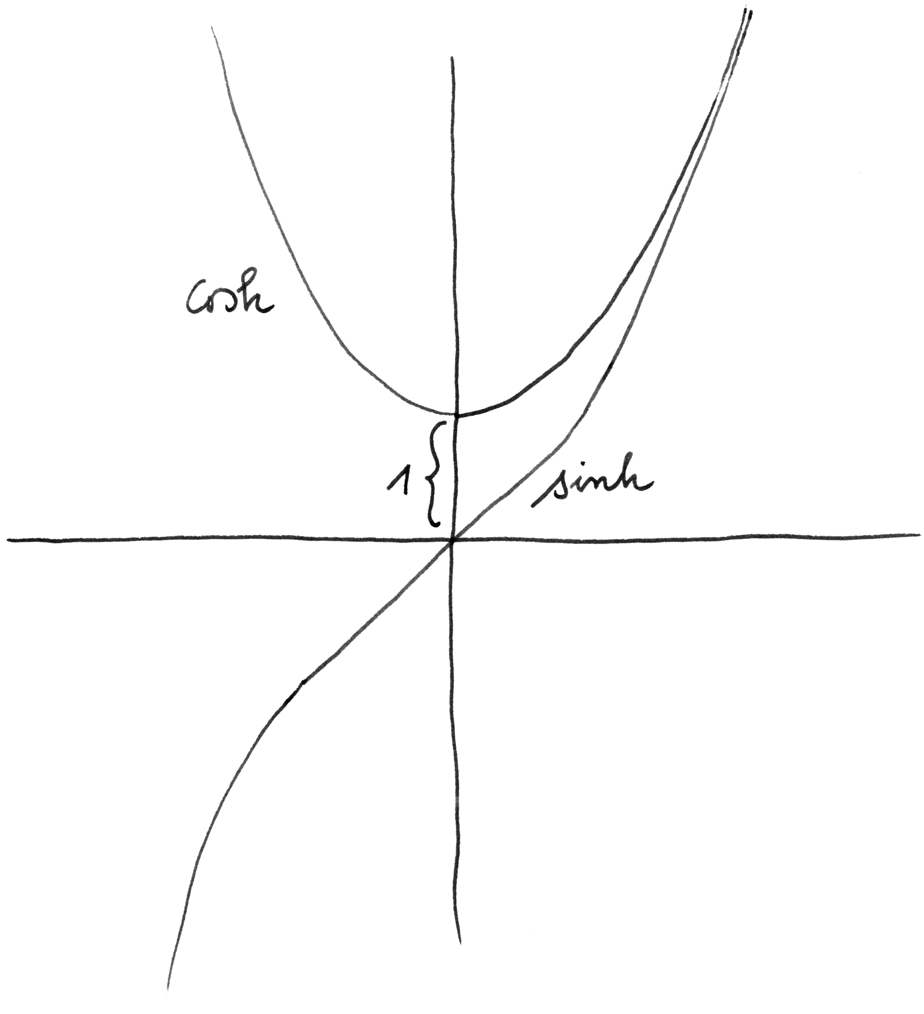
Definition 7.7.1. Der **Sinus hyperbolicus** und **Cosinus hyperbolicus** sind die Abbildungen $\sinh, \cosh : \mathbb{R} \to \mathbb{R}$, die gegeben werden durch die Formeln

$$\sinh x = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} \qquad \text{und} \qquad \cosh x = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}$$

7.7.2. Den Graphen des Cosinus hyperbolicus nennt man auch die **Kettenlinie**, weil er dieselbe Gestalt hat wie eine hängende Kette. Wir zeigen das in 10.3.10 im Anschluß an die Diskussion der Bogenlänge.

7.7.3. Offensichtlich gilt $\sinh(0) = 0$, $\cosh(0) = 1$, $\sinh(-x) = -\sinh(x)$, und $\cosh(-x) = \cosh(x)$, die Ableitungen unserer Funktionen sind

$$\sinh' = \cosh, \quad \cosh' = \sinh$$



Die Graphen von Sinus und Cosinus hyperbolicus es gelten $\cosh^2 - \sinh^2 = 1$ und die Additionstheoreme

$$\begin{aligned} \sinh(a+b) &= \sinh(a)\cosh(b) + \cosh(a)\sinh(b) \\ \cosh(a+b) &= \cosh(a)\cosh(b) + \sinh(a)\sinh(b) \end{aligned}$$

Die Funktion $\cosh$ nimmt bei $x = 0$ ihr Minimum an und $\sinh$ ist eine Bijektion $\sinh : \mathbb{R} \to \mathbb{R}$. Die inverse Abbildung nennt man **Area Sinus hyperbolicus** und bezeichnet sie mit $\operatorname{arsinh} : \mathbb{R} \to \mathbb{R}$. Sie läßt sich auch elementar ausdrücken als $\operatorname{arsinh}(x) = \log(x + \sqrt{x^2+1})$ und für die Ableitung von arsinh erhalten wir

$$\operatorname{arsinh}'(y) = \frac{1}{\sqrt{1+y^2}}$$

da ja gilt $\sinh'(x) = \cosh(x) = \sqrt{1 - \sinh^2(x)}$, $\sinh'(\operatorname{arsinh} y) = \sqrt{1+y^2}$. Ähnlich liefert $\cosh$ eine Bijektion $\cosh : [0, \infty) \to [1, \infty)$, die inverse Abbildung **Area Cosinus hyperbolicus** $\operatorname{arcosh} : [1, \infty) \to [0, \infty)$ kann geschrieben werden als $\operatorname{arcosh}(x) = \log(x + \sqrt{x^2 - 1})$ und ist differenzierbar auf $(1, \infty)$ mit der Ableitung

$$\operatorname{arcosh}'(y) = \frac{1}{\sqrt{y^2-1}}$$

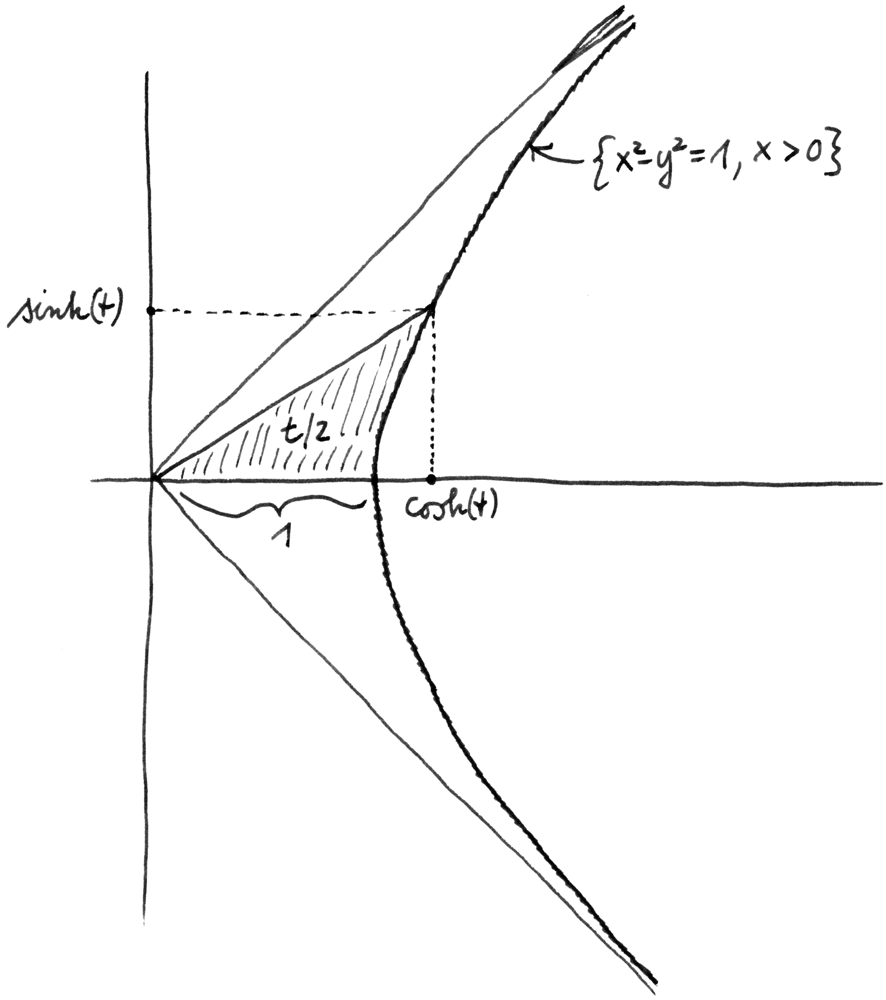
Eher selten benutzt man den **Secans hyperbolicus** $x \mapsto 1/\cosh(x)$, den **Cosecans hyperbolicus** $x \mapsto 1/\sinh(x)$ und den **Tangens hyperbolicus** $\tanh : x \mapsto \sinh(x)/\cosh(x)$ sowie seine Umkehrung $\operatorname{artanh} : (-1, 1) \xrightarrow{\sim} \mathbb{R}$.

7.7.4. Die Namen unserer Funktionen haben den folgenden Hintergrund: Für $t \in \mathbb{R}$ durchläuft der Punkt mit Koordinaten $(\cosh t, \sinh t)$ den Hyperbelast

$$\{(x, y) \in \mathbb{R}^2 \mid (x+y)(x-y) = 1, \quad x > 0\}$$

und zwar ist $|t/2|$ gerade die Fläche oder lateinisch “area”, die zwischen x-Achse, Hyperbel und dem Geradensegment von $(0, 0)$ nach $(\cosh t, \sinh t)$ eingeschlossen ist. Es ist eine ausgezeichnete Übung, diese Behauptung nachzurechnen. Man sieht so die Verwandschaft zu den üblichen trigonometrischen Funktionen, bei denen man nur die Hyperbel $x^2 - y^2 = 1$ zu ersetzen hat durch den Einheitskreis $x^2 + y^2 = 1$, wie wir später zeigen werden. Die Herkunft der Bezeichnung “Sinus” wird auch dort, genauer in 10.4.4 erklärt. Die Bezeichnung **Trigonometrie** bedeutet übrigends “Dreiecksmessung”, die Wurzel “gon” taucht auch im deutschen Wort “Knie” auf und bedeutet im Griechischen sowohl “Knie” als auch im übertragenen Sinne “Winkel”.

7.7.5. Die Lösungsmengen in der Ebene $\mathbb{R}^2$ von Gleichungen der Gestalt $ax^2 + bxy + cy^2 = d$ mit $(a, b, c) \neq (0, 0, 0)$ heißen **ebene Quadriken** oder auch **Kegelschnitte**, da man sie erhalten kann als Schnitte räumlicher Ebenen mit dem



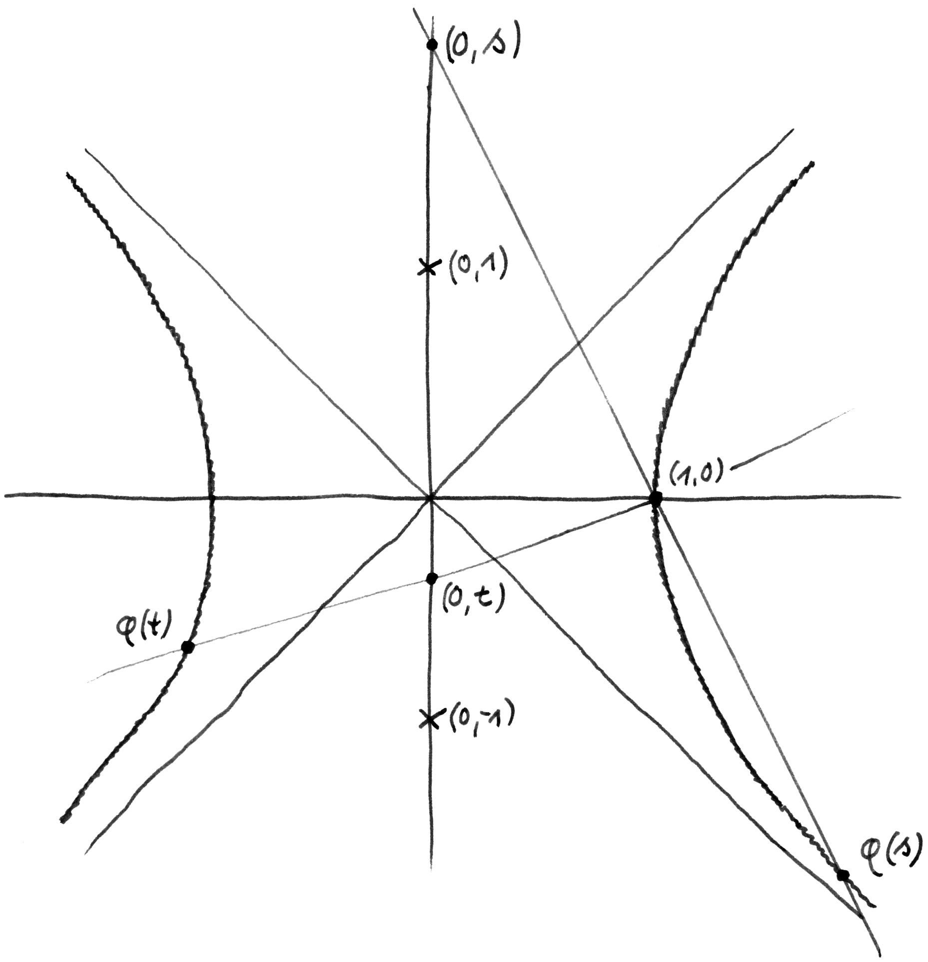
Die geometrische Bedeutung von Sinus und Cosinus hyperbolicus

Kegel $\{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 = z^2\}$ bei geeigneter orthonormaler Identifikation unserer räumlichen Ebenen mit dem $\mathbb{R}^2$. Jeder Kegelschnitt ist bis auf Drehung und Verschiebung eine **Ellipse** $\alpha x^2 + \beta y^2 = 1$ mit $\alpha, \beta > 0$, eine **Hyperbel** $xy = \gamma$ mit $\gamma > 0$, eine **Parabel** $x^2 = \delta y$ mit $\delta > 0$, ein Geradenkreuz, eine Gerade, ein Punkt oder die leere Menge. Die Bezeichnung "Parabel" kommt hier vom griechischen Wort für "Werfen". In der Tat beschreibt ein Wurfgeschoss unter Vernachlässigung des Luftwiderstands stets eine "parabolische" Bahn, vergleiche **??**.

Ergänzende Übung 7.7.6. Sei $H = \{(x,y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1\}$ die Hyperbel. Wir konstruieren eine Bijektion $\varphi : \mathbb{R}\backslash\{\pm 1\} \xrightarrow{\sim} H\backslash(1,0)$, indem wir jedem Punkt $t \in \mathbb{R}\backslash\{\pm 1\}$ den Schnittpunkt der Geraden durch $(0,t)$ und $(1,0)$ mit $H\backslash(1,0)$ zuordnen. Man prüfe, daß diese Abbildung gegeben wird durch

$$\varphi(t) = \left(\frac{t^2+1}{t^2-1}, \frac{2t}{t^2-1}\right)$$

Eine eng verwandte Parametrisierung des Einheitskreises werden wir in 10.6.17 besprechen.



Die geometrische Bedeutung der Abbildung φ aus 7.7.6

8 Potenzreihen und höhere Ableitungen

8.1 Funktionenfolgen und Potenzreihen

Satz 8.1.1. *Sei $(a_\nu)_{\nu\in\mathbb{N}}$ eine Folge reeller Zahlen. Konvergiert für eine reelle Zahl z die Reihe $\sum_{\nu=0}^\infty a_\nu z^\nu$, so konvergiert die Reihe $\sum_{\nu=0}^\infty a_\nu x^\nu$ absolut für alle reellen x mit $|x| < |z|$.*

Vorschau 8.1.2. Sobald Sie die komplexen Zahlen kennengelernt haben, werden Sie unschwer erkennen, daß dieser Satz ganz genauso und mit demselben Beweis auch gilt, wenn wir darin überall "reell" durch "komplex" ersetzen.

Beweis. Die Glieder einer konvergenten Reihe sind beschränkt, unter unseren Annahmen gibt es also eine endliche Schranke B mit $|a_\nu z^\nu| \leq B$ für alle ν. Aus $|x| < |z|$ folgt aus der Konvergenz der geometrischen Reihe 5.5.5 dann

$$\sum_{\nu=0}^\infty |a_\nu x^\nu| \leq \sum_{\nu=0}^\infty |a_\nu z^\nu|\, |x/z|^\nu \leq \sum_{\nu=0}^\infty B\, |x/z|^\nu < \infty \qquad \square$$

Definition 8.1.3. Ein Ausdruck der Gestalt $\sum_{\nu=0}^\infty a_\nu x^\nu$ heißt eine **Potenzreihe**. Eine Potenzreihe anzugeben bedeutet also nichts anderes, als die Folge $(a_\nu)_{\nu\in\mathbb{N}}$ ihrer Koeffizienten anzugeben. Der **Konvergenzradius** $r \in [0,\infty]$ einer Potenzreihe $\sum a_\nu x^\nu$ ist per definitionem

$$r = \sup\{\, |z| \ \mid \sum a_\nu z^\nu \text{ konvergiert}\}$$

Die Bezeichnung als Radius wird erst im Kontext komplexer Potenzreihen **??** verständlich. Nach 8.1.1 definiert jede Potenzreihe mit Konvergenzradius r vermittels der Abbildungsvorschrift $x \mapsto \sum_{\nu=0}^\infty a_\nu x^\nu$ eine reellwertige Funktion auf dem Intervall $(-r, r)$.

Übung 8.1.4. Ist $(a_\nu)_{\nu\in\mathbb{N}}$ eine Folge von von Null verschiedenen reellen Zahlen und existiert der Grenzwert $\lim_{\nu\to\infty} |a_\nu/a_{\nu+1}|$ in $[0,\infty]$, so ist dieser Grenzwert der Konvergenzradius der Potenzreihe $\sum_{\nu=0}^\infty a_\nu x^\nu$.

Satz 8.1.5 (über Potenzreihen). *Die durch eine Potenzreihe $\sum_{\nu=0}^\infty a_\nu x^\nu$ mit Konvergenzradius $r > 0$ definierte Funktion $s : (-r, r) \to \mathbb{R}$ ist stetig, ja sogar differenzierbar, und ihre Ableitung wird an jeder Stelle $x \in (-r, r)$ gegeben durch die Potenzreihe $s'(x) = \sum_{\nu=1}^\infty \nu a_\nu x^{\nu-1}$*

Beispiel 8.1.6. Dieser Satz liefert unmittelbar einen zweiten Beweis für die bereits in 7.2.8 bewiesene Tatsache, daß die Exponentialfunktion ihre eigene Ableitung ist.

8.1.7. Der Beweis dieses Satzes wird den größten Teil dieses Abschnitts einnehmen. Wir zeigen in 8.1.13, daß jede durch eine Potenzreihe definierte Funktion stetig ist, und zeigen die weitergehenden Aussagen in 8.1.15. Zunächst jedoch müssen wir einige technische Hilfsmittel bereitstellen.

Definition 8.1.8. Sei D eine Menge, $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : D \to \mathbb{R}$ und $f : D \to \mathbb{R}$ eine weitere Funktion.

1. Wir sagen, die Folge f_n **konvergiert punktweise** gegen die Funktion f genau dann, wenn für alle $x \in D$ gilt $\lim_{n\to\infty} f_n(x) = f(x)$.

2. Wir sagen, die Folge f_n **konvergiert gleichmäßig** gegen die Funktion f und schreiben $\lim_{n\to\infty} f_n = f$ genau dann, wenn es für beliebiges $\varepsilon > 0$ ein $N = N_\varepsilon \in \mathbb{N}$ gibt derart, daß für alle $n \geq N$ und alle $x \in D$ gilt $|f_n(x) - f(x)| < \varepsilon$.

8.1.9. Aus gleichmäßiger Konvergenz folgt sicher punktweise Konvergenz. Das Umgekehrte gilt nicht: Die Funktionenfolge $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$ konvergiert punktweise, aber nicht gleichmäßig gegen die Grenzfunktion $f : [0,1] \to \mathbb{R}$ mit $f(x) = 0$ für $x \neq 1$ und $f(1) = 1$.

Satz 8.1.10. *Konvergiert eine Folge von stetigen reellwertigen Funktionen gleichmäßig gegen eine Grenzfunktion, so ist auch diese Grenzfunktion stetig.*

Beweis. Sei $D \subset \mathbb{R}$ der gemeinsame Definitionsbereich und seien $f_n : D \to \mathbb{R}$ die Funktionen unserer gleichmäßig konvergenten Folge und $f : D \to \mathbb{R}$ ihre Grenzfunktion. Es reicht sicher, die Stetigkeit von f in jedem Punkt $p \in D$ zu zeigen. Sei also $\varepsilon > 0$ beliebig vorgegeben. Sicher finden wir ein n derart, daß für alle $x \in D$ gilt

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3}$$

Da f_n stetig ist in p, finden wir weiter $\delta > 0$ derart, daß für alle $x \in D$ mit $|x - p| < \delta$ gilt

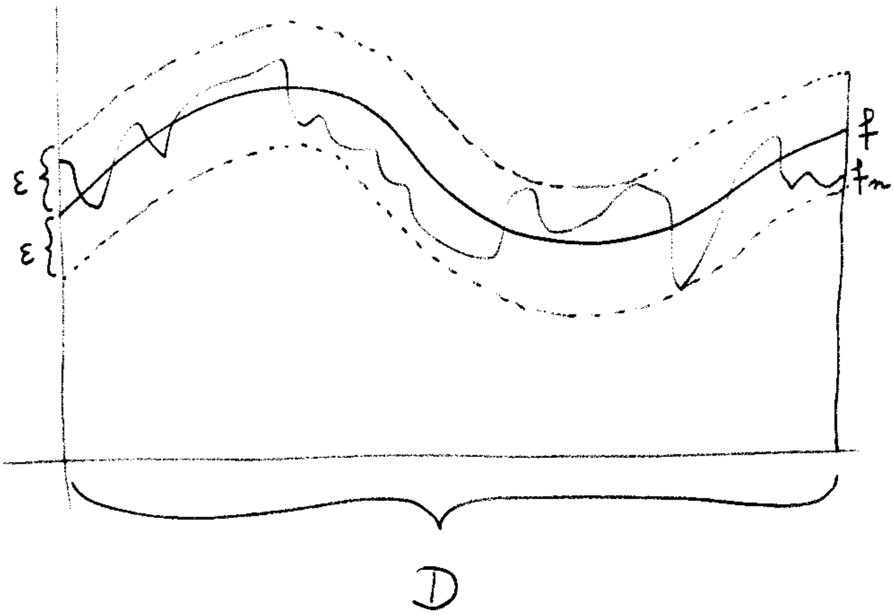
$$|f_n(x) - f_n(p)| < \frac{\varepsilon}{3}$$

Es folgt für alle $x \in D$ mit $|x - p| < \delta$ leicht

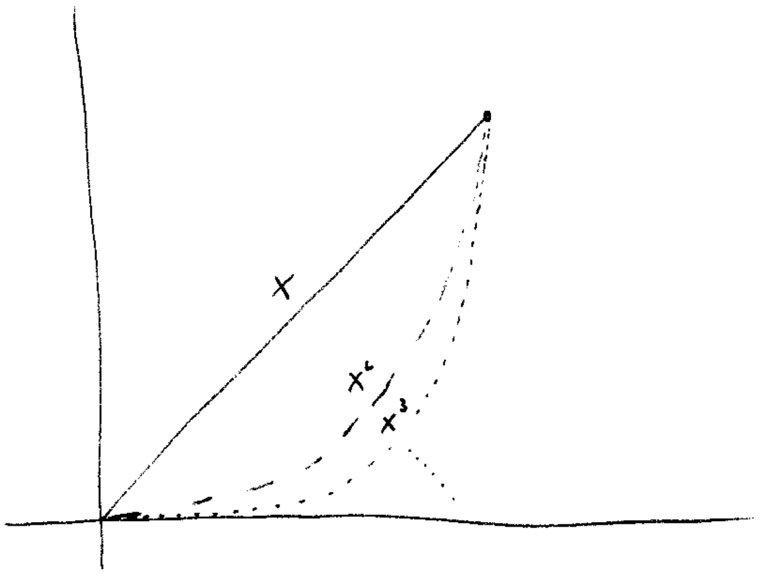
$$\begin{array}{rcl} |f(x) - f(p)| & \leq & |f(x) - f_n(x)| \\ & & +|f_n(x) - f_n(p)| \\ & & \quad +|f_n(p) - f(p)| \ < \ \varepsilon \end{array}$$

□

Proposition 8.1.11. *Ist $\sum a_\nu x^\nu$ eine Potenzreihe mit Konvergenzradius r, so konvergiert die Folge ihrer Partialsummen $s_n(x) = \sum_{\nu=0}^n a_\nu x^\nu$ für alle $\rho \in [0, r)$ gleichmäßig auf dem Kompaktum $D = [-\rho, \rho]$.*



Bei gleichmäßiger Konvergenz müssen für jedes $\varepsilon > 0$ fast alle f_n auf dem ganzen Definitionsbereich D im "ε-Schlauch um f" bleiben.



Die punktweise aber nicht gleichmäßig konvergente Funktionenfolge der Funktionen $x \mapsto x^n$ aus 8.1.9

8.1.12. Eine Potenzreihe mit Konvergenzradius r konvergiert im allgemeinen keineswegs gleichmäßig auf dem ganzen Intervall $(-r, r)$. Ein Gegenbeispiel wäre etwa die Exponentialfunktion, ein anderes die geometrische Reihe.

Beweis. Für alle $x \in [-\rho, \rho]$ gilt $|a_\nu x^\nu| \leq |a_\nu \rho^\nu|$ und folglich

$$|s_n(x) - s(x)| \leq \sum_{\nu=n+1}^{\infty} |a_\nu \rho^\nu|$$

Da $\sum_{\nu=0}^{\infty} |a_\nu \rho^\nu|$ konvergiert, gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\sum_{\nu=N+1}^{\infty} |a_\nu \rho^\nu| < \varepsilon$$

Für $n \geq N$ und beliebiges $x \in [-\rho, \rho]$ gilt dann $|s_n(x) - s(x)| < \varepsilon$. □

Korollar 8.1.13. *Die durch eine Potenzreihe mit positivem Konvergenzradius $r > 0$ auf $(-r, r)$ definierte Funktion ist stetig.*

Beweis. Nach Proposition 8.1.11 und Satz 8.1.10 ist unsere Funktion für alle $\rho \in [0, r)$ stetig auf $[-\rho, \rho]$ als gleichmäßiger Grenzwert stetiger Funktionen. Daraus folgt mühelos, daß sie stetig ist auf ganz $(-r, r)$. □

Satz 8.1.14. Das Integral eines gleichmäßigen Grenzwerts ist der Grenzwert der Integrale. *Sei genauer die Funktion $f : [a, b] \to \mathbb{R}$ gleichmäßiger Grenzwert einer Folge von stetigen Funktionen $f_n : [a, b] \to \mathbb{R}$. So gilt*

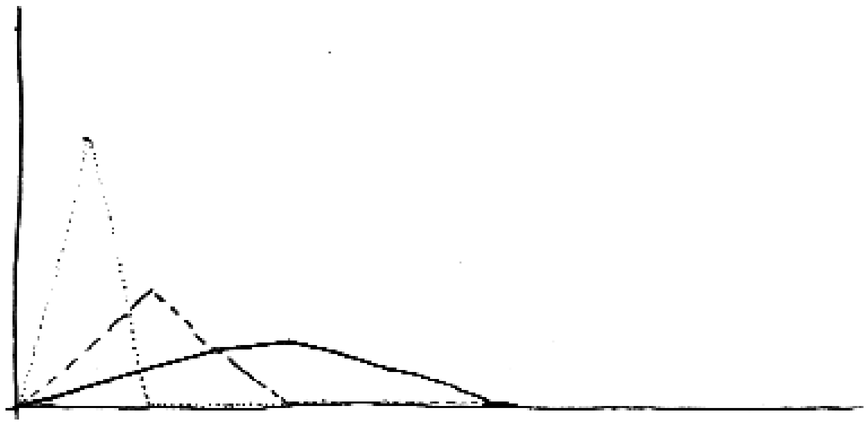
$$\int_a^b f = \lim_{n \to \infty} \int_a^b f_n$$

Beweis. Für alle $\varepsilon > 0$ gibt es $N \in \mathbb{N}$ derart, daß gilt $|f(x) - f_n(x)| < \varepsilon$ für alle $n \geq N$ und alle $x \in [a, b]$. Es folgt

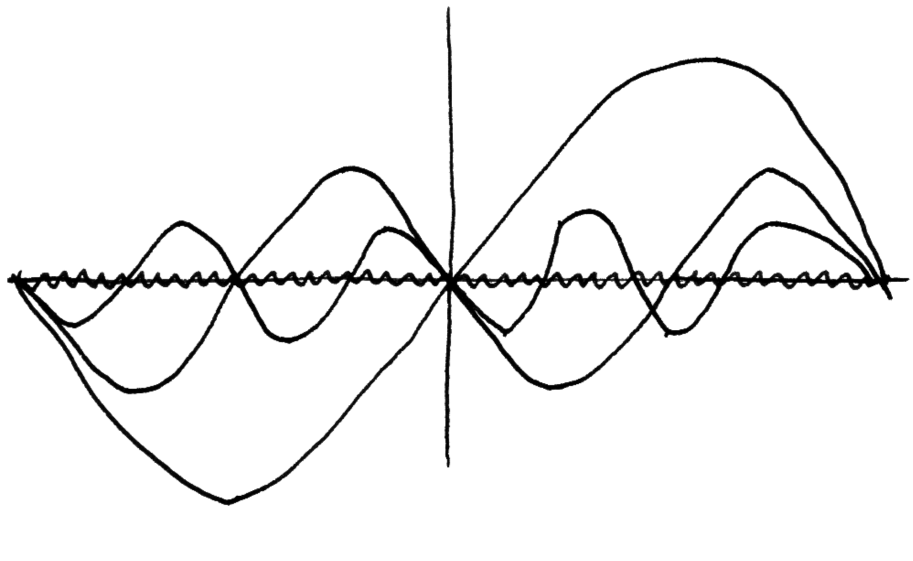
$$\left| \int_a^b f - \int_a^b f_n \right| = \left| \int_a^b (f - f_n) \right| \leq \int_a^b |f - f_n| \leq (b - a)\varepsilon$$

für alle $n \geq N$. □

Satz 8.1.15. Potenzreihen dürfen gliedweise integriert und differenziert werden. *Genauer gilt:*



Einige Glieder einer Funktionenfolge, die punktweise gegen die Nullfunktion konvergiert, deren Integrale jedoch nicht gegen das Integral der Nullfunktion konvergieren.



Einige Glieder einer Funktionenfolge vom Typ $f_n = \frac{1}{n}\sin(nx)$, die gleichmäßig gegen die Nullfunktion konvergiert, deren Ableitungen jedoch nicht gleichmäßig gegen die Ableitung der Nullfunktion konvergieren.

1. *Ist die Funktion $f : (-r, r) \to \mathbb{R}$ gegeben durch die Potenzreihe $f(x) = \sum_{\nu=0}^{\infty} a_\nu x^\nu$, so konvergiert auch die Potenzreihe*

$$\sum_{\nu=0}^{\infty} \tfrac{1}{\nu+1} a_\nu x^{\nu+1}$$

   auf $(-r, r)$, und zwar gegen eine Stammfunktion von f.

2. *Ist die Funktion $g : (-r, r) \to \mathbb{R}$ gegeben durch die Potenzreihe $g(x) = \sum_{\nu=0}^{\infty} b_\nu x^\nu$, so ist g differenzierbar und die Potenzreihe*

$$\sum_{\nu=1}^{\infty} \nu b_\nu x^{\nu-1}$$

   konvergiert auf $(-r, r)$ gegen die Ableitung g' unserer Funktion g.

Beweis. 1. Man wende den vorhergehenden Satz 8.1.14 auf die Folge $f_n(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$ der Partialsummen an.

2. Wir zeigen zunächst, daß die Potenzreihe $\sum_{\nu=1}^{\infty} \nu b_\nu x^{\nu-1}$ auch auf $(-r, r)$ konvergiert. Nach dem Quotientenkriterium konvergiert schon mal die Reihe $\sum_{\nu=1}^{\infty} \nu z^{\nu-1}$ für $|z| < 1$. Für $|x| < r$ wählen wir nun ein ρ mit $|x| < \rho < r$ und dazu eine Schranke B mit $|b_\nu \rho^{\nu-1}| \leq B$ für alle ν. Dann können wir abschätzen

$$|\nu b_\nu x^{\nu-1}| = \left|\nu b_\nu \rho^{\nu-1}(x/\rho)^{\nu-1}\right| \leq \nu B z^{\nu-1}$$

für $z = |x/\rho| < 1$. Nach dem Majorantenkriterium konvergiert also die Reihe $\sum_{\nu=1}^{\infty} \nu b_\nu x^{\nu-1}$. Dann wissen wir aber nach Teil 1, daß für die Funktion $f(x) = \sum_{\nu=1}^{\infty} \nu b_\nu x^{\nu-1}$ unser $g(x) = \sum_{\nu=0}^{\infty} b_\nu x^\nu$ eine Stammfunktion ist. □

Definition 8.1.16. Die **n-te Ableitung** einer Funktion f bezeichnen wir, falls sie existiert, mit $f^{(n)}$. Es ist also $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f''$ und allgemein $f^{(n+1)} = (f^{(n)})'$.

8.1.17. Ist eine Funktion $f : (-r, r) \to \mathbb{R}$ gegeben durch die für $x \in (-r, r)$ konvergente Potenzreihe $f(x) = \sum_{\nu=0}^{\infty} a_\nu x^\nu$, so folgt aus 8.1.15 sofort

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

Wenn also in anderen Worten eine fest vorgegebene Funktion f in einer Umgebung des Nullpunkts durch eine Potenzreihe dargestellt werden kann, so muß unsere Funktion dort beliebig oft differenzierbar sein und die fragliche Potenzreihe ist die Reihe

$$\sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} x^\nu$$

Diese Potenzreihe hinwiederum kann man ganz allgemein für jede in einer Umgebung des Nullpunkts beliebig oft differenzierbare Funktion erklären. Sie heißt dann die **Taylorreihe** besagter Funktion oder genauer ihre **Taylorreihe am Nullpunkt**, muß aber keineswegs positiven Konvergenzradius haben und muß, selbst wenn sie positiven Konvergenzradius hat, keineswegs gegen besagte Funktion konvergieren. Zum Beispiel hat nach 7.2.11 die Funktion

$$f(x) \quad = \quad \begin{cases} \mathrm{e}^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

im Nullpunkt die Ableitungen $f^{(\nu)}(0) = 0 \quad \forall \nu \geq 0$, aber es gilt dennoch $f(x) > 0$ für $x > 0$. Inwiefern die Taylorreihe dennoch unsere Funktion recht gut approximiert, erklären wir in 8.2.

Definition 8.1.18 (Verallgemeinerte Binomialkoeffizienten). Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ setzen wir

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k(k-1)\dots 1} \text{ falls } k \geq 1 \text{ und } \binom{\alpha}{0} = 1.$$

Proposition 8.1.19 (Binomische Reihe). *Für $|x| < 1$ und $\alpha \in \mathbb{R}$ gilt*

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

8.1.20. Sobald wir komplexe Zahlen und in 1.4.3 komplexe Exponenten eingeführt haben, wird der Leser erkennen, daß die Proposition mit demselben Beweis auch für $x, \alpha \in \mathbb{C}$ gilt, verleiche **??**. Aus unseren Überlegungen in 8.1.17 folgt auch unmittelbar, daß die Koeffizienten einer Potenzreihenentwicklung von $(1+x)^\alpha$, wenn es sie denn gibt, notwendig die verallgemeinerten Binomialkoeffizienten sein müssen. Daß es aber eine derartige Entwicklung auch wirklich gibt, muß noch gezeigt werden. Also an die Arbeit!

Beweis. Für $\alpha \in \mathbb{Z}_{\geq -1}$ kennen wir diese Formel schon: Im Fall $\alpha \in \mathbb{N}$ gilt $\binom{\alpha}{k} = 0$ falls $k > \alpha$ und wir erhalten einen Spezialfall der binomischen Formel 3.3.4. Im Fall $\alpha = -1$ gilt $\binom{\alpha}{k} = (-1)^k$ und wir erhalten die geometrische Reihe 5.5.5 für $-x$. Unter der Annahme $\alpha \notin \mathbb{N}$ sagt uns das Quotientenkriterium, daß die Potenzreihe rechts für $|x| < 1$ konvergiert, sagen wir gegen die Funktion $f : (-1, 1) \to \mathbb{R}$. Ich behaupte $(1 + x)f'(x) = \alpha f(x)$. Das prüft man durch gliedweises Differenzieren der binomischen Reihe mithilfe der Beziehung

$$\binom{\alpha-1}{k} + \binom{\alpha-1}{k-1} = \binom{\alpha}{k}$$

die durch eine kurze Rechnung gezeigt wird. Aus $(1+x)f'(x) = \alpha f(x)$ und $f(0) = 1$ folgt aber bereits $f(x) = (1{+}x)^\alpha$, denn setzt man $\varphi(x) = f(x)/(1+x)^\alpha$, so gilt $\varphi(0) = 1$ und

$$\varphi'(x) = \frac{f'(x)(1+x)^\alpha - \alpha(1+x)^{\alpha-1}f(x)}{(1+x)^{2\alpha}} = 0 \qquad \square$$

Beispiel 8.1.21. Die Darstellung

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad \forall x \in (-1,1)$$

ergibt sich, indem wir die geometrische Reihe $(1+x)^{-1} = 1 - x + x^2 - x^3 \pm \dots$ gliedweise integrieren. Ähnlich ergibt sich die Entwicklung von $\operatorname{arsinh}(x)$ in eine Potenzreihe, indem wir die binomische Reihe zu $(1+x^2)^{-1/2}$ gliedweise integrieren.

Beispiel 8.1.22. Um die fünfte Ableitung bei $x = 0$ von $(\mathrm{e}^{x^2}\sinh x)$ bestimmen, rechnen wir

$$\begin{aligned}
\mathrm{e}^{x^2} &= 1 + x^2 + \tfrac{x^4}{2} + \tfrac{x^6}{3!} + \dots \\
\sinh x &= x + \tfrac{x^3}{3!} + \tfrac{x^5}{5!} + \dots \\
\mathrm{e}^{x^2}\sinh x &= x + x^3(\tfrac{1}{3!} + 1) + x^5\left(\tfrac{1}{5!} + \tfrac{1}{3!} + \tfrac{1}{2}\right) + \dots
\end{aligned}$$

wo wir einmal unsere Erkenntnisse über das Produkt absolut konvergenter Reihen verwendet haben, und die gesuchte fünfte Ableitung bei $x = 0$ ergibt sich mit 8.1.17 zu

$$5!\left(\frac{1}{5!} + \frac{1}{3!} + \frac{1}{2}\right) = 1 + 20 + 60 = 81$$

8.1.23. Die Formel für die binomische Reihe kann umgeschrieben werden zur nun für alle $y \in (0,2)$ gültigen Darstellung

$$y^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} (y-1)^k$$

Gehört allgemeiner ein Punkt p zum Definitionsbereich einer Funktion f, so bezeichnen wir eine Darstellung von f der Gestalt

$$f(p+h) = \sum_{\nu=0}^{\infty} a_\nu h^\nu \qquad \text{alias} \qquad f(y) = \sum_{\nu=0}^{\infty} a_\nu (y-p)^\nu$$

als eine **Entwicklung von f in eine Potenzreihe um den Punkt** p. Mit denselben Argumenten wie zuvor gilt dann $a_\nu = f^{(\nu)}(p)/\nu!$. Ist allgemeiner f beliebig oft differenzierbar bei p, so erklären wir die **Taylorreihe von f zum Entwicklungspunkt** p als die Potenzreihe

$$\sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(p)}{\nu!} h^\nu$$

Auch wenn die Partialsummen dieser Reihe nicht gegen $f(p+h)$ konvergieren müssen, liefern sie doch die "bestmöglichen" Approximationen von $f(p+h)$ durch Polynome in h von einem vorgegebenen maximalen Grad, wie wir im folgenden Abschnitt 8.2 ausführen werden.

Ergänzende Übung 8.1.24. Eine Funktion, die für jeden Punkt ihres Definitionsbereichs in einer Umgebung besagten Punktes durch eine Potenzreihe dargestellt werden kann, heißt **analytisch**. Wir werden erst in **??** im Rahmen der Funktionentheorie zeigen, daß Potenzreihen analytische Funktionen liefern: Dort geht es mit den Tricks der Funktionentheorie sehr elegant, wir könnten es aber etwas weniger elegant auch hier schon zeigen. Man zeige: Stimmen zwei auf demselbem reellen Intervall definierte analytische Funktionen auf der Umgebung eines Punktes aus unserem Intervall überein, so sind sie gleich. Hinweis: Man betrachte das Supremum der Menge aller Punkte, an denen unsere beiden Funktionen übereinstimmen.

Ergänzende Übung 8.1.25. Wie lauten die ersten vier Koeffizienten der Potenzreihenentwicklung von $\sqrt{1+x}$? Wie lauten die ersten vier Koeffizienten der Potenzreihenentwicklung von $\sqrt{2+x}$? Gemeint ist jeweils die Entwicklung um $x = 0$.

Ergänzende Übung 8.1.26. Wir würfeln mit einem Würfel eine unendlich lange Zahlenreihe. Anschaulich ist klar, daß der durchschnittliche Abstand zwischen zwei aufeinanderfolgenden Einsen gerade Sechs sein muß. Betrachtet man nun die Wahrscheinlichkeiten, daß die nächste Eins beim nächsten Wurf, beim übernächsten Wurf etc. kommt, multipliziert sie jeweils mit Eins, Zwei etc. und summiert diese Produkte auf, so sollte sich auch dieser dieser durchschnittliche Abstand ergeben. Die Aufgabe ist nun, zu beweisen, daß die Reihe

$$\sum_{n=1}^{\infty} n \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}$$

auch tatsächlich gegen 6 konvergiert.

8.2 Taylorentwicklung

8.2.1. Um im Folgenden auch den Fall $n = 0$ zulassen zu dürfen, vereinbaren wir, daß "0-mal differenzierbar bei p" zu verstehen sein soll als "stetig bei p".

Satz 8.2.2 (Taylorentwicklung). *Sei $I \subset \mathbb{R}$ ein halboffenes Intervall, $p \in I$ ein Punkt und $f : I \to \mathbb{R}$ eine Funktion, deren $(n-1)$-te Ableitung $f^{(n-1)}$ auf ganz I existiert und deren n-te Ableitung $f^{(n)}(p)$ bei p existiert. So gilt:*

1. *Es gibt genau ein Polynom Q vom Grad $\leq n$ mit der Eigenschaft*

$$f(x) = Q(x) + (x-p)^n \varepsilon(x-p)$$

für eine Funktion ε mit $\lim_{h \to 0} \varepsilon(h) = 0$.

2. *Dieses Polynom kann auch charakterisiert werden als das eindeutig bestimmte Polynom Q vom Grad $\leq n$, dessen Ableitungen bei p bis zur n-ten Ableitung einschließlich mit den entsprechenden Ableitungen bei p unserer Funktion f übereinstimmen, in Formeln $f^{(\nu)}(p) = Q^{(\nu)}(p)$ für $0 \leq \nu \leq n$.*

8.2.3. Wir nennen Q das **Approximationspolynom bis zur Ordnung n an f bei p**. Der Graph des Approximationspolynoms bis zur ersten Ordnung heißt die **Tangente** an den Graphen von f im Punkt $(p, f(p))$, der Graph des Approximationspolynoms bis zur zweiten Ordnung die **Schmiegeparabel**.

Beweis. Sicher gibt es genau ein Polynom Q, das die Bedingung aus Teil 2 erfüllt, nämlich das Polynom

$$Q(x) = f(p) + f'(p)(x-p) + \frac{f^{(2)}(p)}{2!}(x-p)^2 + \ldots + \frac{f^{(n)}(p)}{n!}(x-p)^n$$

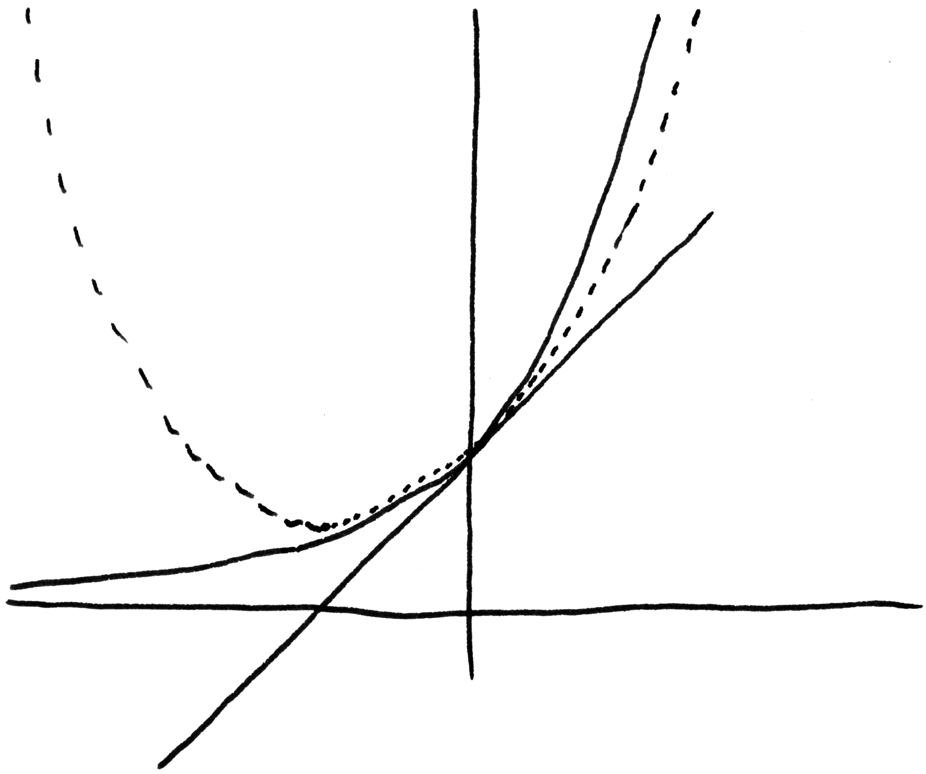
das aus den ersten $n+1$ Termen der Taylorreihe von f beim Entwicklungspunkt p besteht. Betrachten wir für dieses Polynom Q die Differenz $r = f - Q$, so verschwinden die ersten n Ableitungen von r bei $x = p$ und wir erhalten durch wiederholte Anwendung der Regeln von de l'Hospital 7.4.1 und die Definition von $r^{(n)}(p)$ in der Tat

$$\lim_{x \to p} \frac{r(x)}{(x-p)^n} = \ldots = \lim_{x \to p} \frac{r^{(n-1)}(x)}{n!\,(x-p)} = \frac{r^{(n)}(p)}{n!} = 0$$

Damit bleibt nur noch zu zeigen, daß kein anderes Polynom $\hat{Q}$ die Bedingung aus Teil 1 erfüllen kann. In der Tat folgt aber für zwei Polynome vom Grad $\leq n$ aus

$$\lim_{x \to p} \frac{\hat{Q}(x) - Q(x)}{(x-p)^n} = 0$$

nach 6.3.23 bereits $\hat{Q}(x) = Q(x)$. □



Eine unvollkommene Darstellung der Tangente $y = x + 1$ und Schmiegeparabel $y = x^2/2 + x + 1$ an den Graph der Exponentialfunktion im Punkt $(0, 1)$.

8.2.4. Wir können die Aussage des Satzes dahingehend umformulieren, daß gilt

$$f(p+h) = f(p) + f'(p)h + \frac{f^{(2)}(p)}{2!}h^2 + \ldots + \frac{f^{(n)}(p)}{n!}h^n + h^n\varepsilon(h)$$

für eine Funktion ε mit $\lim_{h\to 0}\varepsilon(h) = 0$. Schärfere Abschätzungen für das Restglied unter stärkeren Annahmen an die zu approximierende Funktion liefert die gleich folgende Proposition.

Proposition 8.2.5 (Restglieddarstellungen zur Taylorentwicklung). *Ist $I \subset \mathbb{R}$ ein halboffenes Intervall, $p \in I$ ein Punkt und $f : I \to \mathbb{R}$ eine Funktion, so gilt für alle $n \in \mathbb{N}$:*

Lagrange'sche Form des Restglieds: *Ist f sogar auf dem ganzen Intervall I und sogar $(n+1)$-mal differenzierbar, so gibt es für alle $x \in I$ ein ξ zwischen p und x mit*

$$f(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(p)}{\nu!}(x-p)^\nu + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-p)^{n+1}$$

Integraldarstellung des Restglieds: *Ist $f^{(n+1)}$ zusätzlich auch noch stetig auf ganz I, so gilt für alle $x \in I$ die Formel*

$$f(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(p)}{\nu!}(x-p)^\nu + \frac{1}{n!}\int_p^x (x-t)^n f^{(n+1)}(t)\,\mathrm{d}t$$

Beweis. Wir kürzen die Summe wie zuvor mit $\sum = Q(x)$ ab. Für den Rest $r(x) = f(x) - Q(x)$ verschwinden, wie wir bereits wissen, die ersten n Ableitungen bei $x = p$, und unter unseren zusätzlichen Voraussetzungen ist r sogar $(n+1)$-mal differenzierbar auf ganz I mit $r^{(n+1)} = f^{(n+1)}$. Um daraus die beiden Darstellungen des Restglieds zu erhalten, brauchen wir nur zwei einfache Rechnungen.

1. Mit dem verallgemeinerten Mittelwertsatz und der Erkenntnis, daß Nenner und Zähler bei $x = p$ verschwinden, erhalten wir unter der Annahme $p \neq x$ sofort

$$\frac{r(x)}{(x-p)^{n+1}} = \frac{r'(\xi_1)}{(n+1)(\xi_1-p)^n} = \frac{r''(\xi_2)}{n(n+1)(\xi_2-p)^{n-1}} = \ldots = \frac{r^{(n+1)}(\xi)}{(n+1)!}$$

2. Durch wiederholtes partielles Integrieren erhalten wir

$$\begin{aligned} r(x) = \textstyle\int_p^x r'(t)\,\mathrm{d}t = \int_p^x (x-t)r''(t)\,\mathrm{d}t &= \tfrac{1}{2}\textstyle\int_p^x (x-t)^2 r'''(t)\,\mathrm{d}t \\ &= \ldots \\ &= \tfrac{1}{n!}\textstyle\int_p^x (x-t)^n r^{(n+1)}(t)\,\mathrm{d}t \qquad \square \end{aligned}$$

8.2.6. Diese Abschätzungen liefern umgekehrt auch zwei alternative Beweise für den Satz über die Taylorentwicklung. Allerdings benötigen diese alternativen Beweise wesentlich stärkere Voraussetzungen als unser ursprünglicher Beweis.

Ergänzende Übung 8.2.7. Ist eine Funktion auf einem offenen Intervall n-mal stetig differenzierbar für $n \geq 1$ und verschwinden an einer Stelle p alle ihre höheren Ableitungen unterhalb der n-ten, so hat sie bei p ein isoliertes lokales Maximum bzw. Minimum, falls n gerade ist und die n-te Ableitung bei p negativ bzw. positiv, und kein lokales Extremum, falls n ungerade ist. Wem diese Übung zu einfach ist, der mag dasselbe zeigen unter der schwächeren Annahme, daß $f^{(n-1)}$ zwar bei p, aber nicht notwendig auf dem ganzen Intervall differenzierbar ist.

8.3 Rechnen mit Approximationen

Definition 8.3.1. Seien $f, g : D \to \mathbb{R}$ zwei auf einer Teilmenge $D \subset \mathbb{R}$ definierte Funktionen. Sei $p \in D$ ein Punkt und $n \in \mathbb{N}$ eine natürliche Zahl. Wir sagen, f und g **stimmen bei p überein bis zur Ordnung** n und schreiben

$$f \sim_p^n g \qquad \text{oder genauer} \qquad f(x) \sim_{x=p}^n g(x)$$

genau dann, wenn gilt $f(p+h) = g(p+h) + h^n \varepsilon(h)$ für eine Funktion ε, die stetig ist bei Null mit Funktionswert $\varepsilon(0) = 0$, und die eben definiert ist auf der Menge aller h mit $p + h \in D$.

8.3.2. Die Notation $f \sim_p^n g$ scheint mir bequem und suggestiv, sie ist jedoch unüblich. Häufig nennt man eine Funktion, die bei $x = 0$ mit der Nullfunktion übereinstimmt bis zur Ordnung n, auch ein **kleines o von** x^n und bezeichnet so eine Funktion mit $\mathrm{o}(x^n)$. In dieser Notation würde man statt $f \sim_p^n g$ schreiben $f(x) = g(x) + \mathrm{o}((x-p)^n)$.

8.3.3. Natürlich folgt aus $f \sim_p^n g$ und $g \sim_p^n h$ schon $f \sim_p^n h$. Sind P und Q Polynome vom Grad $\leq n$ und ist p ein Häufungspunkt von D, so folgt aus $P \sim_p^n Q$ schon $P = Q$. Der Satz über die Taylorentwicklung 8.2.2 liefert uns für eine n-mal stetig differenzierbare Funktion f auf einem halboffenen Intervall D das eindeutig bestimmte Polynom Q vom Grad $\leq n$, das bei p mit f übereinstimmt bis zur Ordnung n. Genauer besagt dieser Satz, daß dieses Polynom Q charakterisiert wird durch die Bedingungen $Q^{(\nu)}(p) = f^{(\nu)}(p)$ für $0 \leq \nu \leq n$.

Satz 8.3.4 (Rechnen mit Approximationen). *1. Seien $f, g : D \to \mathbb{R}$ zwei auf einer Teilmenge $D \subset \mathbb{R}$ definierte Funktionen. Sei $p \in D$ ein Punkt und seien P, Q Polynome mit $f \sim_p^n P$ und $g \sim_p^n Q$. So folgt*

$$f + g \sim_p^n P + Q \qquad \textit{und} \qquad fg \sim_p^n PQ$$

2. **(Höhere Kettenregel).** *Seien $f : D \to \mathbb{R}$ und $g : E \to \mathbb{R}$ auf Teilmengen $D, E \subset \mathbb{R}$ definierte Funktionen mit $f(D) \subset E$. Sei $p \in D$ ein Punkt und seien P, Q Polynome mit $f \sim_p^n P$ und $g \sim_{f(p)}^n Q$. So folgt*

$$g \circ f \sim_p^n Q \circ P$$

8.3.5. Im Fall $n = 0$ spezialisiert dieser Satz zur Stetigkeit von Summen und Produkten 6.1.16 sowie Verknüpfungen 6.1.6. Im Fall $n = 1$ spezialisiert er zur Summenregel 7.2.1, Produktregel 7.2.1 und Kettenregel 7.2.5.

Beweis. 1. Das bleibt dem Leser überlassen. Im Fall der Summe gilt das sogar für beliebige Funktionen P und Q, im Fall des Produkts reicht anstelle der Polynomialität die zusätzliche Annahme, daß P und Q in einer Umgebung von p beschränkt sind.

2. Wir zeigen zunächst $g \circ f \sim_p^n Q \circ f$ und dann $Q \circ f \sim_p^n Q \circ P$. Für die erste Aussage schreiben wir $g(y) = Q(y) + (y - f(p))^n \varepsilon(y - f(p))$ für ε stetig bei Null mit Funktionswert Null. Durch Einsetzen von $y = f(x)$ und Erweitern des letzten Terms mit $(x-p)^n$ erhalten wir

$$(g \circ f)(x) = (Q \circ f)(x) + (x-p)^n \left[\left(\frac{f(x) - f(p)}{x - p} \right)^n \varepsilon(f(x) - f(p)) \right]$$

für alle $x \neq p$. Im Fall $n \geq 1$ stimmt f bei p bis mindestens zur Ordnung 1 überein mit dem Polynom P, folglich ist f differenzierbar bei p und der Ausdruck in eckigen Klammern strebt für $x \to p$ gegen Null. Im Fall $n = 0$ stimmt f bei p bis zur Ordnung 0 überein mit dem Polynom P, also ist f zumindest stetig in p und der Ausdruck in eckigen Klammern strebt für $x \to p$ wieder gegen Null. Wir müssen also nur noch für jedes Polynom Q zeigen

$$Q \circ f \sim_p^n Q \circ P$$

Das folgt jedoch sofort aus Teil 1. □

Beispiel 8.3.6. Um die sechste Ableitung bei $x = 0$ von $1/\cosh(x)$ zu berechnen, erinnern wir uns an

$$\begin{array}{rcl} \cosh x & = & 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots \\ (1+y)^{-1} & = & 1 - y + y^2 - y^3 + y^4 - y^5 + y^6 \dots \end{array}$$

wo wir Gleichheitszeichen und Pünktchen geschrieben haben statt $\sim^6$ mit entsprechenden Spezifikationen. Mit unserer "höheren Kettenregel" 8.3.4 erhalten wir dann sofort

$$\begin{aligned} 1/\cosh(x) &= 1 - \left(\tfrac{x^2}{2} + \tfrac{x^4}{4!} + \tfrac{x^6}{6!}\right) \\ &\quad + \left(\tfrac{x^2}{2} + \tfrac{x^4}{4!} + \tfrac{x^6}{6!}\right)^2 \\ &\quad - \left(\tfrac{x^2}{2} + \tfrac{x^4}{4!} + \tfrac{x^6}{6!}\right)^3 + \ldots \end{aligned}$$

Als Koeffizient von x^6 ergibt sich

$$-\frac{1}{6!} + \frac{1}{4!} - \frac{1}{8} = \frac{1}{6!}(-1 + 30 - 90) = -\frac{61}{6!}$$

und die fragliche sechste Ableitung bei $x = 0$ ist mithin -61. Eine andere Möglichkeit wäre, das Approximationspolynom sechsten Grades an $1/\cosh(x)$ in $x = 0$ als $a_0 + a_1 x + \ldots + a_6 x^6$ anzusetzen und aus der "höheren Produktregel" die Gleichung

$$\left(1 + \frac{x^2}{2} + \frac{x^4}{4!} + \frac{x^6}{6!}\right)\left(a_0 + a_1 x + \ldots + a_6 x^6\right) = 1 + bx^8 + \ldots$$

zu folgern, die es uns hinwiederum erlaubt, induktiv die a_ν zu bestimmen. Diese Rechnung kann im vorliegenden Fall zusätzlich vereinfacht werden durch die Erkenntnis, daß eh gilt $0 = a_1 = a_3 = a_5 = \ldots$, da unsere Funktion nämlich gerade ist.

Übung 8.3.7. Gegeben zwei beliebig oft differenzierbare Funktionen auf einem Intervall ist die Taylorreihe ihrer Summe die Summe der Taylorreihen und die Taylorreihe des Produkts das Produkt der Taylorreihen. Hier verstehen wir Produkt und Summe von Potenzreihen im formalen Sinn, vergleiche **??**.

Ergänzende Übung 8.3.8. Man zeige, daß die Identitäten $\exp(\log(x+1)) = x+1$ und $\log((\mathrm{e}^x - 1) + 1) = x$ auch als formale Identitäten von Potenzreihen gelten, daß also etwa im ersten Fall für $k \geq 2$ gilt

$$\sum_{j(1)+\ldots+j(i)=k} \frac{1}{i!}\left(\frac{-(-1)^{j(1)}}{j(1)}\right)\cdots\left(\frac{-(-1)^{j(i)}}{j(i)}\right) = 0$$

wo die Summe über alle $i \geq 1$ und alle Abbildungen $j : \{1, \ldots, i\} \to \mathbb{N}_{\geq 1}$ läuft, bei denen k die Summe der Werte ist, wohingegen dieselbe Summe für $k = 1$ gerade den Wert Eins ergibt. Allgemeiner führe man aus, inwiefern die Taylorreihe der Verknüpfung zweier unendlich oft differenzierbarer Funktionen gerade die Verknüpfung ihrer Taylorreihen ist.

Ergänzung 8.3.9. Wir zeigen nun die in 3.1.8 versprochene Formel

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

für die ebendort eingeführten Catalan-Zahlen C_n. Diese Zahlen erfüllen offensichtlich die Rekursion

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$

Die **erzeugende Funktion** der Folge der Catalan-Zahlen alias die formale Potenzreihe $P = \sum_{n\geq 0} C_n x^n$ erfüllt demnach im Ring der formalen Laurentreihen aus **??** die Formel $xP^2 = P - 1$ alias $P^2 - \frac{1}{x}P + \frac{1}{x} = 0$. Damit folgt, immer im Ring der formalen Laurentreihen, eine Identität der Form

$$P = \frac{1}{2x} \pm \sqrt{\frac{1}{4x^2} - \frac{1}{x}} = \frac{1 \pm \sqrt{1-4x}}{2x}$$

falls die fragliche Wurzel im Ring der formalen Laurentreihen existieren sollte, wobei das Vorzeichen noch zu bestimmen ist. Nach 8.3.7 existiert diese Wurzel jedoch in der Tat und kann beschrieben werden durch die binomische Reihe

$$\sqrt{1-4x} = \sum_{n=0}^{\infty} \binom{1/2}{n} (-4x)^n$$

Deren konstanter Term ist Eins, so daß für unser Vorzeichen nur das Minus in Frage kommt. Damit muß notwendig gelten

$$P = \frac{1 - \sqrt{1-4x}}{2x}$$

Die Koeffizienten unserer binomischen Reihe lassen sich vereinfachen zu

$$\begin{aligned}
\binom{1/2}{n}(-1)^n 4^n &= \frac{1}{n!}\left(\frac{1}{2}\right)\left(\frac{-1}{2}\right)\left(\frac{-3}{2}\right)\cdots\left(\frac{-2n+3}{2}\right)(-1)^n 4^n \\
&= -\frac{1}{n!}(1\cdot 3\ldots(2n-3))2^n \\
&= \frac{-2(2(n-1))!}{(n-1)!(n-1)!}
\end{aligned}$$

wobei die letzte Identität nur für $n \geq 1$ stimmt. Setzen wir nun alles zusammen, so ergibt sich wie gewünscht

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

8.4 Der Abel'sche Grenzwertsatz*

8.4.1. In diesem Abschnitt will ich mein Versprechen einlösen und die Formel

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \log 2$$

zeigen. Wenn wir $x = 1$ in die Reihenentwicklung von $\log(1 + x)$ einsetzen dürften, so folgte das sofort. Die Schwierigkeit liegt darin, daß wir bisher nur für $|x| < 1$ nachgewiesen haben, daß unsere Potenzreihe aus 8.1.21 gegen $\log(1 + x)$ konvergiert. Der folgende Satz hilft uns, diese Schwierigkeit zu überwinden, und wird auch in 10.6.16 bei der Herleitung der wunderbaren Formel $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \ldots$ benötigt. Beide Formeln sehe aber mehr als schöne Blüten und weniger als zentrale Inhalte an, und davon abgesehen spielt der abelsche Grenzwersatz im weiteren Verlauf der Vorlesung keine Rolle.

Satz 8.4.2 (Abel'scher Grenzwertsatz). *Konvergiert eine Potenzreihe auch noch auf einem Randpunkt ihres Konvergenzbereichs, so stellt sie bis in diesen Randpunkt hinein eine stetige Funktion dar.*

Vorschau 8.4.3. Für diejenigen Leser, die bereits mit komplexen Potenzreihen vertraut sind, sei bemerkt, daß die entsprechende Aussage im Komplexen in dieser Form nicht mehr gilt. Mehr dazu finden Sie etwa in **??**.

Beweis. Wir dürfen ohne Beschränkung der Allgemeinheit annehmen, daß $x = 1$ der besagte Randpunkt des Konvergenzbereichs ist. Sei also $\sum_{k=0}^{\infty} a_k$ eine konvergente Reihe reeller Zahlen. Wir müssen zeigen, daß die Reihe $f(x) = \sum_{k=0}^{\infty} a_k x^k$ für $x \in [0, 1]$ eine stetige Funktion darstellt. Dazu schreiben wir die Differenzen der Partialsummen in der Form

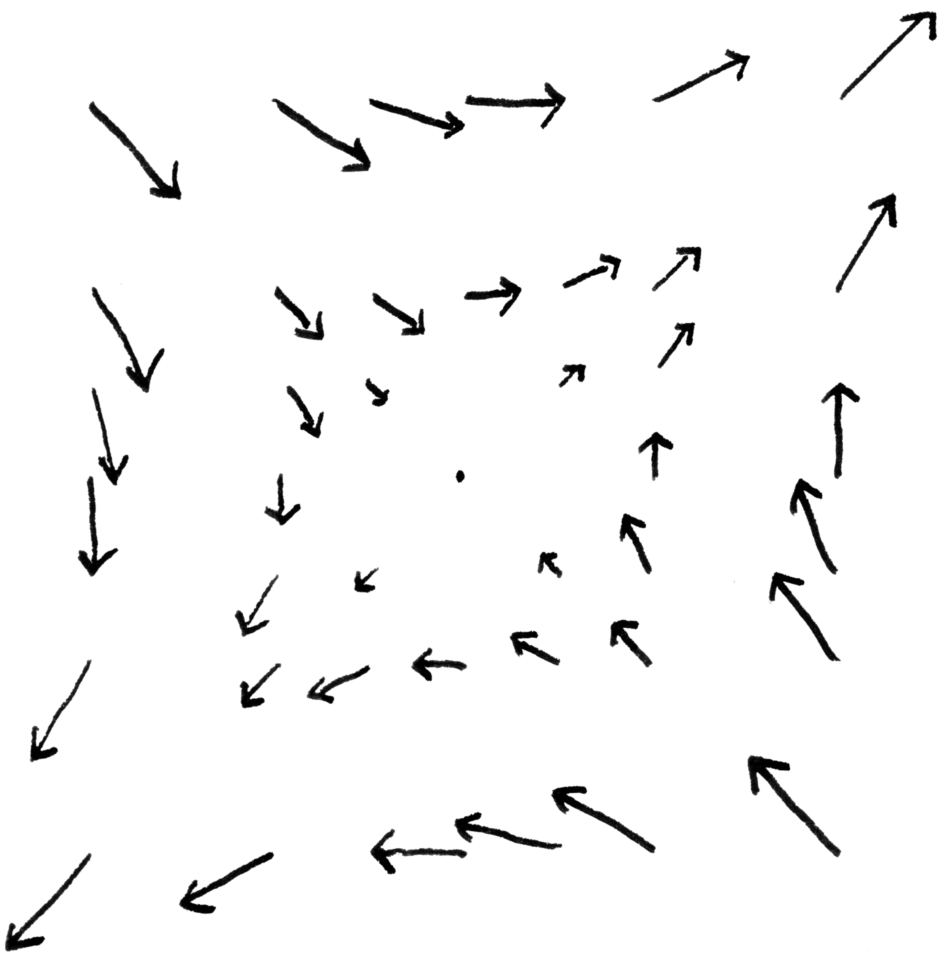
$$\begin{array}{rcrl}
\sum_{k=n}^{m} a_k x^k & = & (x^n - x^{n+1}) & (a_n) \\
 & + & (x^{n+1} - x^{n+2}) & (a_n + a_{n+1}) \\
 & + & (x^{n+2} - x^{n+3}) & (a_n + a_{n+1} + a_{n+2}) \\
 & & \ldots & \ldots \\
 & + & (x^{m-1} - x^m) & (a_n + \ldots\ldots\ldots + a_{m-1}) \\
 & + & x^m & (a_n + \ldots\ldots\ldots + a_{m-1} + a_m)
\end{array}$$

Für alle $\varepsilon > 0$ finden wir nun wegen der Konvergenz unserer Reihe ein N derart, daß für alle n, m mit $N \leq n \leq m$ gilt $|a_n + \ldots + a_m| \leq \varepsilon$. Für alle n, m mit $N \leq n \leq m$ und alle $x \in [0, 1]$ folgt daraus die Abschätzung

$$\left|\sum_{k=n}^{m} a_k x^k\right| \leq (x^n - x^m)\,\varepsilon + x^m \varepsilon \leq \varepsilon$$

Diese Abschätzung zeigt die gleichmäßige Konvergenz der Folge der Partialsummen auf $[0, 1]$ und damit die Stetigkeit der Grenzfunktion. □

Vorschau 8.4.4. Läßt sich die durch eine Potenzreihe im Inneren des Konvergenzintervalls definierte Funktion stetig auf einen Randpunkt fortsetzen, so muß die Potenzreihe an besagtem Randpunkt keineswegs konvergieren. Nehmen wir der Einfachkeit halber an, daß das Konvergenzintervall $(-1, 1)$ ist und der fragliche Randpunkt die 1, so folgt die Konvergenz von $\sum a_n$ jedoch aus der stetigen Fortsetzbarkeit der Funktion $\sum a_n x^n$ von $x \in [0, 1)$ auf $[0, 1]$ zusammen mit der **Tauber-Bedingung**, daß die Folge na_n betragsmäßig beschränkt sein möge. Unter der stärkeren Annahme $\lim_{n\to\infty} na_n = 0$ wurde das bereits von Tauber gezeigt.



Die Darstellung der Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (y/4, x/4))$ als ebenes Vektorfeld. Als Abbildung der Ebene auf sich selber beschreibt sie eine Stauchung um den Faktor 4 gefolgt von einer Spiegelung an der Hauptdiagonalen $y = x$.

9 Stetigkeit in mehreren Veränderlichen

9.1 Vorschläge zur Veranschaulichung

9.1.1. Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ schreiben wir in der Form

$$(x_1, \dots, x_n) \mapsto (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

oder abkürzend $f = (f_1, \dots, f_m)$. Man kann sich derartige Abbildungen auf die verschiedensten Arten vorstellen.

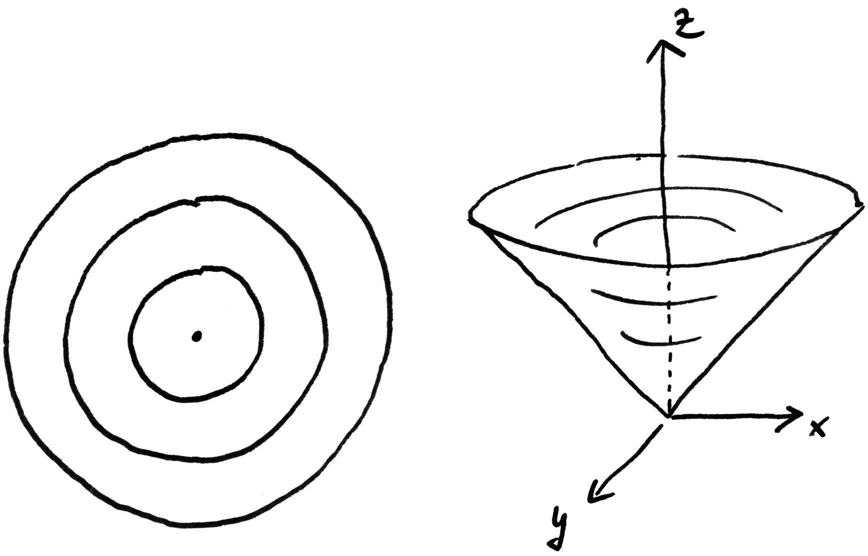
1. Den Fall $n = m = 1$ hatten wir schon in 6.1.1 ausführlich behandelt und sogar etwas allgemeiner mögliche Interpretationen einer Abbildung von $\mathbb{R}$ in einen beliebigen Raum bzw. von einem beliebigen Raum nach $\mathbb{R}$ besprochen: Erstere kann man sich etwa veranschaulichen als Beschreibung der Bewegung eines Teilchens in besagtem Raum, Letztere als eine Temperaturverteilung auf besagtem Raum.

2. Im Fall $n + m = 3$ kann man sich die Abbildung f ähnlich wie im Fall $n = m = 1$ gut durch ihren Graphen $\Gamma(f) = \{(x, f(x)) \mid x \in \mathbb{R}\} \subset \mathbb{R}^3$ bzw. $\Gamma(f) = \{(x, y, f(x, y)) \mid x, y \in \mathbb{R}\} \subset \mathbb{R}^3$ veranschaulichen. Der Graph einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ ist anschaulich eine hügelige Landschaft. Der Graph einer Abbildung $f : \mathbb{R} \to \mathbb{R}^2$ sieht aus wie ein Draht im $\mathbb{R}^3$ mit genau einem Punkt für jede vorgegebene x-Koordinate. Zum Beispiel ist der Graph jeder konstanten Abbildung $\mathbb{R} \to \mathbb{R}^2$ eine Parallele zur x-Achse und der Graph jeder konstanten Abbildung $\mathbb{R}^2 \to \mathbb{R}$ eine "vollständig platte Landschaft" alias eine zur (x, y)-Ebene parallele Ebene.

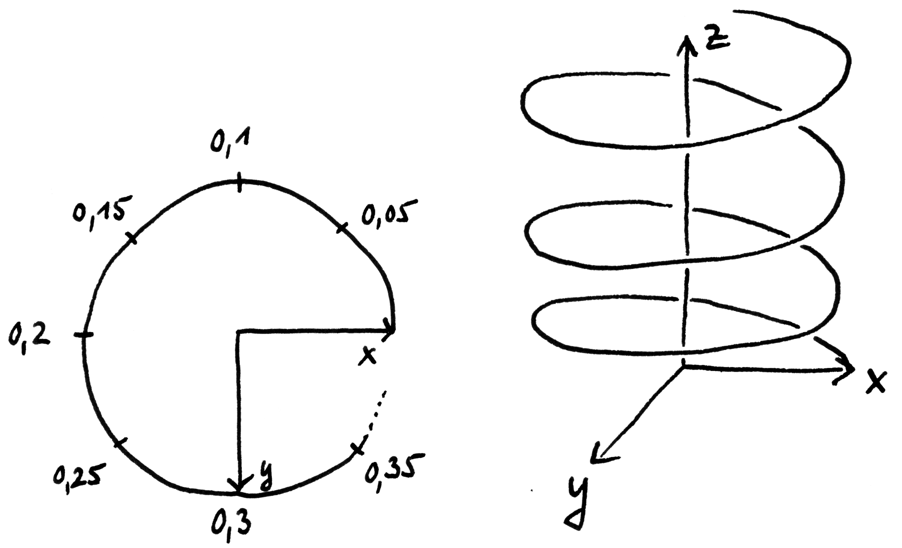
3. Eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ kann man auch gut graphisch darstellen, indem man auf der Ebene $\mathbb{R}^2$ die **Niveaulinien** einzeichnet, die im Bild der Hügellandschaft die Höhenlinien in einer Landkarte für unsere Landschaft wären, in Formeln die Mengen $\{(x, y) \mid f(x, y) = c\}$ für verschiedene, meist äquidistant gewählte $c \in \mathbb{R}$. Auch eine Funktion $f : \mathbb{R}^3 \to \mathbb{R}$ kann man sich noch gut mithilfe ihrer analog definierten **Niveauflächen** vorstellen, aber mit dem Zeichnen wird es dann schon schwierig.

4. Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$ kann man sich vorstellen als ein Vektorfeld, jedem Punkt $x \in \mathbb{R}^n$ wird ja in der Tat ein Vektor $f(x) \in \mathbb{R}^n$ zugeordnet.

5. Es ist auch oft nützlich, sich f wirklich als eine Abbildung vorzustellen. Die Abbildung $x \mapsto (x, x)$ ist in diesem Bild zum Beispiel die diagonale Einbettung der Zahlengerade in die Ebene, und $(x, y) \mapsto (y, x)$ ist die Spiegelung am Bild unserer diagonalen Einbettung.



Die Niveaulinien und der Graph der Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto \sqrt{x^2 + y^2}$. Der Graph dieser Funktion hat die Gestalt einer Eistüte mit dem Öffnungswinkel $90°$, die mit ihrer Spitze senkrecht auf den Ursprung in der xy-Ebene steht.



Die Darstellung als bewegtes Teilchen und der Graph der Funktion $f : \mathbb{R} \to \mathbb{R}^2$, $z \mapsto (\cos(2\pi z/(0,4)), \sin(2\pi z/(0,4)))$. Anders als im Text haben wir hier eine Funktion der z-Koordinate dargestellt.

9.1.2. Ich will den Begriff der Stetigkeit nun statt für Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$ gleich im allgemeineren Rahmen der sogenannten “metrischen Räume” diskutieren, bei dem die bisherige stark von Koordinaten abhängige Darstellung von einer mehr abstrakt-geometrischen Sichtweise abgelöst wird. Dieser koordinatenfreie Zugang benötigt zwar einen größeren begrifflichen Aufwand, aber ich denke, dieser Aufwand lohnt, und zwar aus den folgenden Gründen: Erstens umfaßt unser Rahmen so auch unendlichdimensionale Räume wie zum Beispiel die für die Quantenmechanik wichtigen Hilberträume. Zweitens treten in meinen Augen auch schon im endlichdimensionalen Kontext die Zusammenhänge bei einer koordinatenfreien Behandlung klarer hervor. Man kennt das aus der Physik: Rechnet man wie üblich mit Einheiten, die ja mathematisch schlicht Basen eindimensionaler Vektorräume sind, so treten auch die Konsistenz und der Sinn physikalischer Formeln viel klarer zu Tage, als wenn man mit bloßen Zahlen arbeitet.

9.2 Stetigkeit bei metrischen Räumen

Definition 9.2.1. Eine **Metrik** d auf einer Menge X ist eine Abbildung $d : X \times X \to \mathbb{R}_{\geq 0}$ derart, daß für alle $x, y, z \in X$ gilt:

1. $d(x, y) = 0 \Leftrightarrow x = y$;
2. $d(x, y) = d(y, x)$;
3. $d(x, y) \leq d(x, z) + d(z, y)$.

Ein **metrischer Raum** ist ein Paar $X = (X, d)$ bestehend aus einer Menge X und einer Metrik d auf X.

Beispiel 9.2.2. Der Buchstabe d steht in diesem Zusammenhang wohl für das Wort “Distanz”. Auf dem $\mathbb{R}^n$ liefert der übliche **euklidische Abstand** $d(x, y) = \sqrt{(x_1 - y_1)^2 + \ldots (x_n - y_n)^2}$ eine Metrik. Die Ungleichung aus der Definition einer Metrik wird in diesem Beispiel in **??** formal bewisen und bedeutet anschaulich, daß in einem Dreieck mit Seitenlängen a, b, c stets gilt $a \leq b + c$. Sie heißt deshalb auch ganz allgemein die **Dreiecksungleichung**.

Beispiel 9.2.3. Auf dem $\mathbb{R}^n$ ist auch der **Betragsabstand**

$$d(x, y) = \sup_{1 \leq i \leq n} |x_i - y_i|$$

eine Metrik. *Wenn nichts anderes gesagt ist, fassen wir den $\mathbb{R}^n$ stets auf als einen metrischen Raum mit dem Betragsabstand als Metrik.* Diese Metrik ist zwar weniger anschaulich ist als der euklidische Abstand, läßt sich aber einfacher handhaben.

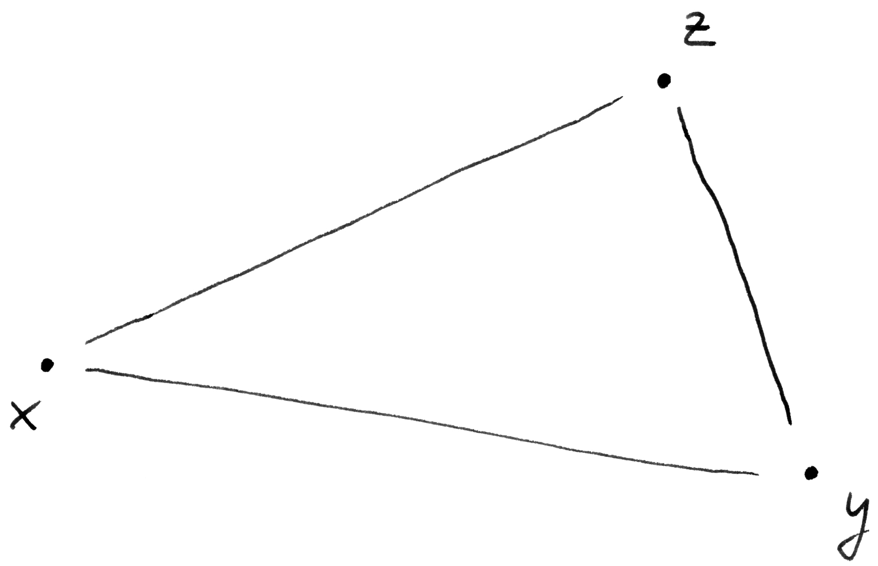


Illustration zur Dreiecksungleichung

Beispiel 9.2.4. Jede Teilmenge eines metrischen Raums ist mit der **induzierten Metrik** selbst ein metrischer Raum.

Definition 9.2.5. Sei X ein metrischer Raum. Für $x \in X$ und $\varepsilon > 0$ bezeichne

$$\mathrm{B}(x; \varepsilon) = \{z \in X \mid d(x, z) < \varepsilon\}$$

den ε**-Ball** um x oder auch die ε**-Kugel** um x oder auch die ε**-Umgebung** von x.

Beispiel 9.2.6. Für den euklidischen Abstand im $\mathbb{R}^3$ ist der Ball um x mit Radius ε tatsächlich ein Ball. Für den Betragsabstand hat $\mathrm{B}(x; \varepsilon)$ dahingegen die Gestalt eines Würfels mit Mittelpunkt x und Seitenlänge 2ε.

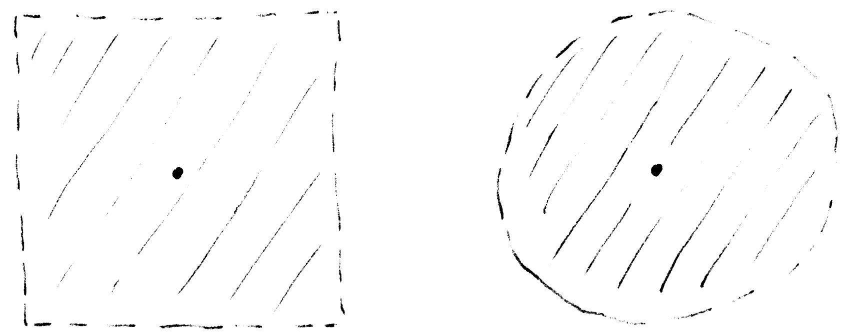
Definition 9.2.7. Unter einer **Umgebung** eines Punktes in einem metrischen Raum versteht man eine Teilmenge von besagtem Raum, die einen ganzen Ball um unseren Punkt umfaßt.

9.2.8. Die Umgebungen eines Punktes im $\mathbb{R}^n$ bezüglich der euklidischen Metrik sind dieselben wie seine Umgebungen bezüglich der Betragsmetrik, was man unschwer explizit prüft und was formal auch aus 9.9.21 folgen wird. Das ist der Grund dafür, daß wir im Folgenden Definitionen nach Möglichkeit mithilfe von Umgebungen formulieren, denn für so definierte Begriffe ist a priori klar, daß im Fall des $\mathbb{R}^n$ ihre Bedeutung nicht davon abhängt, ob wir mit dem euklidischen Abstand oder mit dem Betragsabstand arbeiten.

9.2.9. Der Schnitt von endlich vielen Umgebungen eines Punktes in einem metrischen Raum ist wieder eine Umgebung besagten Punktes. Je zwei verschiedene Punkte eines metrischen Raums besitzen disjunkte Umgebungen. Genauer sind für x, y mit $d(x, y) = r > 0$ die $(r/2)$-Bälle um x und y disjunkt. In der Tat folgte für z aus dem Schnitt ja mit Hilfe der Dreiecksungleichung $r = d(x, y) \leq d(x, z) + d(y, z) < r$, also kann es solch ein z nicht geben.

Definition 9.2.10. Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt **stetig im Punkt** $p \in X$ genau dann, wenn es für jede Umgebung U von $f(p)$ eine Umgebung U' von p gibt mit $f(U') \subset U$. Eine Abbildung zwischen metrischen Räumen heißt **stetig** genau dann, wenn sie stetig ist in jedem Punkt.

9.2.11. Unter einer **Umgebungsbasis** eines Punktes versteht man eine Menge von Umgebungen besagten Punktes derart, daß jede seiner Umgebungen mindestens eine Umgebung unseres Systems umfaßt. Zum Beispiel bilden alle ε-Umgebungen eines Punktes eine solche Umgebungsbasis. Um die Stetigkeit einer Abbildung f in einem Punkt p nachzuweisen, müssen wir offensichtlich nur für alle Mengen aus einer Umgebungsbasis seines Bildes $f(p)$ prüfen, daß es jeweils eine Umgebung von p gibt, die dahinein abgebildet wird. Gleichbedeutend zur Stetigkeit einer Abbildung f im Punkt p ist also insbesondere die Forderung, daß es für jedes $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ gibt derart, daß gilt $f(\mathrm{B}(p; \delta)) \subset \mathrm{B}(f(p); \varepsilon)$.



Bälle in der Ebene für den Betragsabstand und den euklidischen Abstand

Beispiel 9.2.12. Einfache Beispiele für stetige Abbildungen sind Einbettungen von einem Teilraum, konstante Abbildungen, oder auch die Projektion eines $\mathbb{R}^n$ auf eine Koordinate. In diesen Fällen können wir einfach $\delta = \varepsilon$ nehmen.

Beispiel 9.2.13. Als etwas kompliziertere Beispiele bemerken wir, daß **die Addition und die Multiplikation** $\mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x + y$ bzw. $(x, y) \mapsto xy$ **stetig sind.** Das ist im Wesentlichen die Aussage der ersten beiden Teile von Lemma 5.1.35.

Bemerkung 9.2.14. Es gibt Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ derart, daß sowohl $x \mapsto f(x, b)$ als auch $y \mapsto f(a, y)$ stetig sind für alle b bzw. alle a, daß aber dennoch die Funktion f selbst *nicht* stetig ist. Als Beispiel betrachte man die Funktion mit $(x, y) \mapsto xy/(x^2 + y^2)$ für $(x, y) \neq (0, 0)$ und $(0, 0) \mapsto 0$. Sie ist nicht stetig am Nullpunkt nach dem anschließenden Satz 9.2.15, da nämlich ihre Verknüpfung mit $\mathbb{R} \to \mathbb{R}^2$, $t \mapsto (t, t)$ nicht stetig ist bei $t = 0$. Die Stetigkeit von $t \mapsto (t, t)$ hinwiederum mag man aus der Komponentenregel 9.2.18 folgern. Der Anschauung mag die Erkenntnis helfen, daß unsere merkwürdige Funktion, wenn man vom Ursprung selbst einmal absieht, auf allen Geraden durch den Ursprung konstant ist. Auf den beiden Koordinatenachsen ist unsere Funktion konstant Null, auf allen anderen Geraden durch den Ursprung jedoch nimmt sie nur am Ursprung den Wert Null an und sonst konstant einen von Null verschiedenen Wert.

Satz 9.2.15. *Jede Verknüpfung stetiger Abbildungen ist stetig.*

Beweis. Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen zwischen metrischen Räumen und $p \in X$ ein Punkt. Wir zeigen genauer: Ist f stetig bei p und g stetig bei $f(p)$, so ist $(g \circ f)$ stetig bei p. Ist in der Tat g stetig bei $f(p)$, so finden wir für jede Umgebung U von $g(f(p))$ eine Umgebung U' von $f(p)$ mit $g(U') \subset U$. Ist zusätzlich f stetig ist bei p, finden wir für diese Umgebung U' von $f(p)$ weiter eine Umgebung U'' von p mit $f(U'') \subset U'$. Damit haben wir aber auch eine Umgebung U'' von p gefunden mit $(g \circ f)(U'') \subset U$. □

Definition 9.2.16. Gegeben metrische Räume (X_i, d_i) für $1 \leq i \leq n$ machen wir das Produkt $X = X_1 \times \ldots \times X_n$ zu einem metrischen Raum durch die **Produktmetrik**

$$d(x, y) = \sup_{1 \leq i \leq n} d_i(x_i, y_i)$$

für $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$.

Beispiel 9.2.17. Der Betragsabstand auf $\mathbb{R}^{n+m}$ ist die Produktmetrik zu den Betragsabständen auf $\mathbb{R}^n$ und $\mathbb{R}^m$.

Proposition 9.2.18 (Komponentenregel). *Seien Z und $X_1, \ldots, X_n$ metrische Räume und $f_i : Z \to X_i$ Abbildungen. Genau dann ist die Abbildung $f = (f_1, \ldots, f_n) : Z \to X_1 \times \ldots \times X_n$ stetig, wenn alle f_i stetig sind.*

9.2.19. Wenden wir diese Proposition an mit f der Identität auf einem Produkt, so impliziert die Stetigkeit der Identität, daß alle Projektionsabbildungen $\mathrm{pr}_i : X_1 \times \ldots \times X_n \to X_i$ stetig sein müssen.

Beweis. Da die Projektionen pr_i Abstände zwischen Punkten nie vergrößern, können wir ihre Stetigkeit direkt zeigen, "indem wir jeweils $\delta = \varepsilon$ nehmen". Ist f stetig, so sind folglich auch die $f_i = \mathrm{pr}_i \circ f$ stetig als Verknüpfungen stetiger Abbildungen. Sind umgekehrt alle f_i stetig in p, so gibt es für jedes $\varepsilon > 0$ gewisse δ_i mit $d(p, z) < \delta_i \Rightarrow d_i(f_i(p), f_i(z)) < \varepsilon$, wo d_i die Metrik auf X_i bezeichnet. Nehmen wir $\delta = \inf \delta_i$, so gilt

$$d(p, z) < \delta \Rightarrow d(f(p), f(z)) < \varepsilon$$

und das ist gleichbedeutend zu $f(\mathrm{B}(p; \delta)) \subset \mathrm{B}(f(p); \varepsilon)$. □

Korollar 9.2.20. *Ist X ein metrischer Raum und sind f, g stetige Abbildungen $X \to \mathbb{R}$, so sind auch $f + g$ und fg stetige Abbildungen $X \to \mathbb{R}$.*

Beweis. Wir schreiben $f+g$ bzw. fg als die Verknüpfung der nach 9.2.18 stetigen Abbildung $X \to \mathbb{R}^2$, $x \mapsto (f(x), g(x))$ mit der nach 9.2.13 stetigen Addition bzw. Multiplikation $\mathbb{R}^2 \to \mathbb{R}$. □

Beispiel 9.2.21. Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch die Vorschrift

$$(x, y) \mapsto (x \sinh(y), x^2 y^3)$$

ist stetig. In der Tat reicht es nach der Komponentenregel zu zeigen, daß ihre beiden Komponenten f_1 und f_2 stetig sind. Wir zeigen das nur für die erste Komponente und überlassen die Behandlung der zweiten Komponente dem Leser. Warum also ist die Abbildung $f_1 : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x \sinh(y)$ stetig? Nun, $(x, y) \mapsto x$ ist stetig nach 9.2.19 als Projektion auf eine Koordinate, $(x, y) \mapsto y$ desgleichen, $(x, y) \mapsto \sinh(y)$ dann auch als Verknüpfung stetiger Funktionen, und schließlich auch $(x, y) \mapsto x \sinh(y)$ als Produkt stetiger Funktionen.

Übung 9.2.22. Ist (X, d) ein metrischer Raum, so ist für alle $z \in X$ die Abbildung $X \to \mathbb{R}$, $x \mapsto d(x, z)$ stetig. Hinweis: Dreiecksungleichung. Ist allgemeiner $A \subset X$ eine nichtleere Teilmenge, so ist auch die Abbildung $d_A : X \to \mathbb{R}$ gegeben durch $d_A(x) = \inf\{d(x, a) \mid a \in A\}$ stetig.

Übung 9.2.23. Ist (X, d) ein metrischer Raum, so ist die Metrik stetig als Abbildung $d : X \times X \to \mathbb{R}$.

Übung 9.2.24. Wir versehen den Körper der komplexen Zahlen $\mathbb{C}$ mit der Metrik $d(z, w) = |z - w|$. Man zeige, daß dann die Addition und die Multiplikation stetige Abbildungen $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ sind und das Bilden des Inversen eine stetige Abbildung $\mathbb{C}^\times \to \mathbb{C}^\times$.

Ergänzende Übung 9.2.25. Man zeige, daß das Invertieren von Matrizen eine stetige Abbildung $\mathrm{GL}(n; \mathbb{C}) \to \mathrm{GL}(n; \mathbb{C})$ ist. Hinweis: Cramer'sche Regel **??**.

Übung 9.2.26. Sei $f : X \to Y$ eine Abbildung von metrischen Räumen, die Abstände nicht verkleinert, in Formeln $d(f(x), f(z)) \geq d(x, z)\ \forall x, z \in X$. Man zeige, daß f injektiv ist und $f^{-1} : f(X) \to X$ stetig.

9.3 Konvergenz von Folgen in metrischen Räumen

Definition 9.3.1. Sei $\mathbb{N} \to X$, $n \mapsto x_n$ eine Folge in einem metrischen Raum X und $x \in X$ ein Punkt. Wir sagen, die Folge x_n **strebt gegen** x oder **konvergiert gegen** x und nennen x den **Grenzwert der Folge** und schreiben

$$\lim_{n\to\infty} x_n = x$$

genau dann, wenn jede Umgebung von x fast alle Glieder unserer Folge enthält. Gleichbedeutend können wir ebensogut auch fordern, daß jeder Ball um x fast alle Glieder unserer Folge enthält.

9.3.2. Der Grenzwert einer Folge ist eindeutig, wenn er existiert. Das folgt wie im Beweis von 5.1.20 daraus, daß nach 9.2.9 je zwei verschiedene Punkte eines metrischen Raums disjunkte Umgebungen besitzen.

Übung 9.3.3. Sei (x_n, y_n) eine Folge im Produkt $X \times Y$ der metrischen Räume X und Y. Genau dann konvergiert unsere Folge gegen (x, y), wenn x_n gegen x konvergiert und y_n gegen y. Man formuliere und beweise auch die offensichtliche Verallgemeinerung auf beliebige endliche Produkte metrischer Räume.

Definition 9.3.4. Ein metrischer Raum heißt **beschränkt** genau dann, wenn es für die möglichen Abstände zwischen Punkten unseres Raums eine reelle obere Schranke gibt. Eine Abbildung in einen metrischen Raum heißt **beschränkt** genau dann, wenn ihr Bild beschränkt ist.

Beispiel 9.3.5. Sei D eine Menge und X ein metrischer Raum. Auf dem Raum $\mathrm{Ens}^{\mathrm{b}}(D, X)$ aller beschränkten Abbildungen $f : D \to X$ kann man eine Metrik erklären durch die Vorschrift

$$d(f, g) = \sup\{d(f(p), g(p)) \mid p \in D\}$$

im Fall $D \neq \emptyset$ und in offensichtlicher Weise im Fall $D = \emptyset$. Diese Metrik heißt die **Metrik der gleichmäßigen Konvergenz**. In der Tat ist für f_n, f im Funktionenraum $\mathrm{Ens}^{\mathrm{b}}(D, \mathbb{R})$ mit dieser Metrik die Konvergenz

$$\lim_{n\to\infty} f_n = f$$

gleichbedeutend dazu, daß die Abbildungen f_n im Sinne unserer Definition 8.1.8 gleichmäßig gegen die Abbildung f konvergieren.

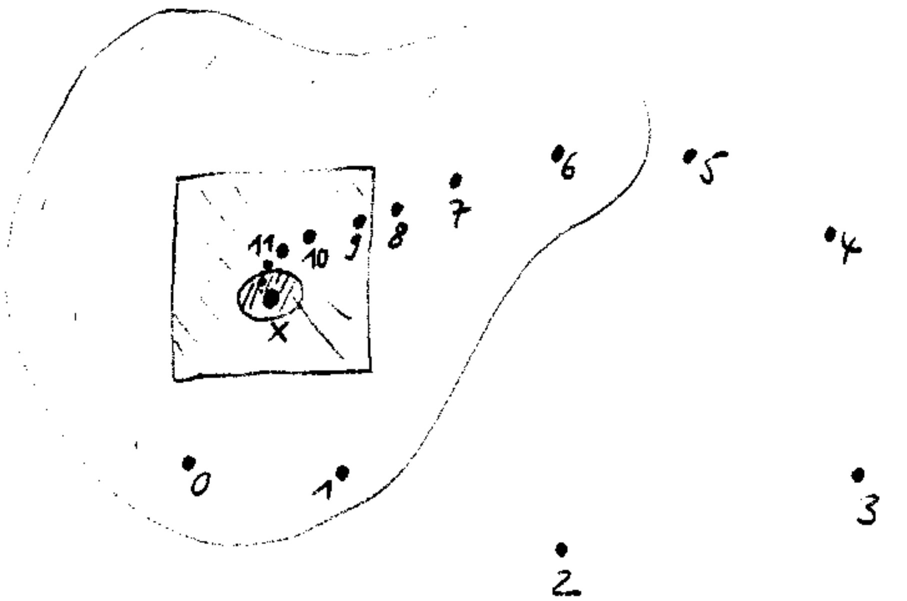


Illustration zur Konvergenz von Folgen. Eingezeichnet sind drei Umgebungen eines Punktes x, in jeder sollen fast alle Folgenglieder liegen.

Definition 9.3.6. Sei D eine Menge, X ein metrischer Raum, $f_n : D \to X$ eine Folge von Abbildungen und $f : D \to X$ eine weitere Abbildung.

1. Wir sagen, die Folge der f_n **konvergiere punktweise** gegen f genau dann, wenn gilt $\lim_{n\to\infty} f_n(p) = f(p)$ für alle Punkte $p \in D$.

2. Wir sagen, die Folge der f_n **konvergiere gleichmäßig** gegen f genau dann, wenn es für jedes $\varepsilon > 0$ ein $N = N_\varepsilon$ gibt mit

$$n \geq N \Rightarrow (d(f_n(p), f(p)) < \varepsilon \ \forall p \in D)$$

Übung 9.3.7. Für beschränkte Abbildungen von einer Menge D in einen metrischen Raum X ist auch in dieser Allgemeinheit gleichmäßige Konvergenz gleichbedeutend zur Konvergenz im Raum $\mathrm{Ens}^{\mathrm{b}}(D, X)$ mit seiner eben erklärten "Metrik der gleichmäßigen Konvergenz".

Übung 9.3.8. Für jede konvergente Folge in einem metrischen Raum ist die Menge der Folgenglieder beschränkt.

Übung 9.3.9 (**Stetigkeit als Folgenstetigkeit**). Sei $f : X \to Y$ eine Abbildung von metrischen Räumen. Genau dann ist f stetig in p, wenn für jede Folge x_n mit $\lim_{n\to\infty} x_n = p$ gilt $\lim_{n\to\infty} f(x_n) = f(p)$. Hinweis: 6.3.29.

9.4 Abgeschlossene und offene Teilmengen

Definition 9.4.1. Sei X ein metrischer Raum und $A \subset X$ eine Teilmenge. Ein Punkt $x \in X$ heißt ein **Berührungspunkt von** A genau dann, wenn es eine Folge in A gibt, die gegen x konvergiert. Eine Teilmenge $A \subset X$ heißt **abgeschlossen** oder präziser **abgeschlossen in** X genau dann, wenn sie alle ihre Berührungspunkte enthält, wenn sie also "abgeschlossen ist unter der Bildung von Grenzwerten". Statt "A ist eine abgeschlossene Teilmenge von X" schreiben wir kurz aber unüblich

$$A \underset{\mathrm{A}}{\subset} X$$

9.4.2. Wenn wir eine Menge einfach nur "abgeschlossen" nennen, so in der Hoffnung, dem Leser sei klar, in Bezug auf welchen größeren Raum X dies "abgeschlossen" gemeint ist. Ist X ein metrischer Raum und sind $U \subset Y \subset X$ Teilmengen, so meint $U \underset{\mathrm{A}}{\subset} Y$, daß U abgeschlossen ist als Teilmenge des Raums Y mit seiner induzierten Metrik.

Übung 9.4.3. Der Schnitt eines beliebigen Systems von abgeschlossenen Teilmengen eines metrischen Raums ist abgeschlossen. Die Vereinigung von endlich vielen abgeschlossenen Teilmengen eines metrischen Raums ist abgeschlossen. Die leere Menge und der ganze Raum sind abgeschlossen. Jede einpunktige und damit auch jede endliche Teilmenge eines metrischen Raums ist abgeschlossen. Jedes kompakte reelle Intervall ist abgeschlossen in $\mathbb{R}$.

Ergänzende Übung 9.4.4. Ein Abbildung zwischen metrischen Räumen ist stetig genau dann, wenn ihr Graph abgeschlossen ist im kartesischen Produkt unserer beiden Räume. Hinweis: 9.3.9

Ergänzende Übung 9.4.5. Jede abgeschlossene echte Untergruppe der reellen Zahlengeraden ist zyklisch, als da heißt von der Gestalt $\mathbb{Z}\alpha$ für ein $\alpha \in \mathbb{R}$. Hinweis: Ist $G \subset \mathbb{R}$ unsere Untergruppe, so betrachte man $\inf(G \cap \mathbb{R}_{>0})$.

Definition 9.4.6. Eine Teilmenge eines metrischen Raums heißt **offen** oder genauer **offen in unserem metrischen Raum** genau dann, wenn sie für jeden ihrer Punkte eine Umgebung ist, d.h. wenn sie mit jedem Punkt auch einen ganzen Ball um besagten Punkt enthält. Statt "U ist eine offene Teilmenge von X" schreiben wir kurz aber unüblich

$$U \Subset X$$

9.4.7. Wenn wir eine Menge einfach nur "offen" nennen, so in der Hoffnung, dem Leser sei klar, in Bezug auf welchen größeren Raum X dies "offen" gemeint ist. Ist X ein metrischer Raum und sind $U \subset Y \subset X$ Teilmengen, so meint $U \Subset Y$, daß U offen ist als Teilmenge des Raums Y mit seiner induzierten Metrik.

Übung 9.4.8. Der Schnitt von endlich vielen offenen Teilmengen eines metrischen Raums ist offen. Die Vereinigung eines beliebigen Systems von offenen Teilmengen eines metrischen Raums ist offen. Die leere Menge und der ganze Raum sind offen. In einem endlichen metrischen Raum ist jede Teilmenge offen und abgeschlossen. Die im Sinne unserer hier gegebenen Definition "offenen" Intervalle von $\mathbb{R}$ sind genau die Intervalle (a, b) für $a, b \in \overline{\mathbb{R}}$, d.h. unsere "offenen reellen Intervalle" aus 5.1.7.

Beispiele 9.4.9. In einem metrischen Raum ist ein Ball $\mathrm{B}(x; r)$ stets offen, denn für $z \in \mathrm{B}(x; r)$ gilt $d(x, z) < r$, also gibt es $\varepsilon > 0$ mit $d(x, z) < r - \varepsilon$, und dann haben wir aber $\mathrm{B}(z; \varepsilon) \subset \mathrm{B}(x; r)$ nach der Dreiecksungleichung. Insbesondere umfaßt jede Umgebung eines Punktes eine offene Umgebung desselben Punktes.

Satz 9.4.10. *Eine Teilmenge eines metrischen Raums ist abgeschlossen genau dann, wenn ihr Komplement offen ist.*

Beweis. Sei X unser metrischer Raum und $M \subset X$ eine Teilmenge. Ist M nicht abgeschlossen, so gibt es einen Punkt $p \in X \backslash M$, der Berührungspunkt von M ist, also $p = \lim_{n\to\infty} x_n$ mit $x_n \in M$. Dann kann es aber kein $\varepsilon > 0$ geben mit

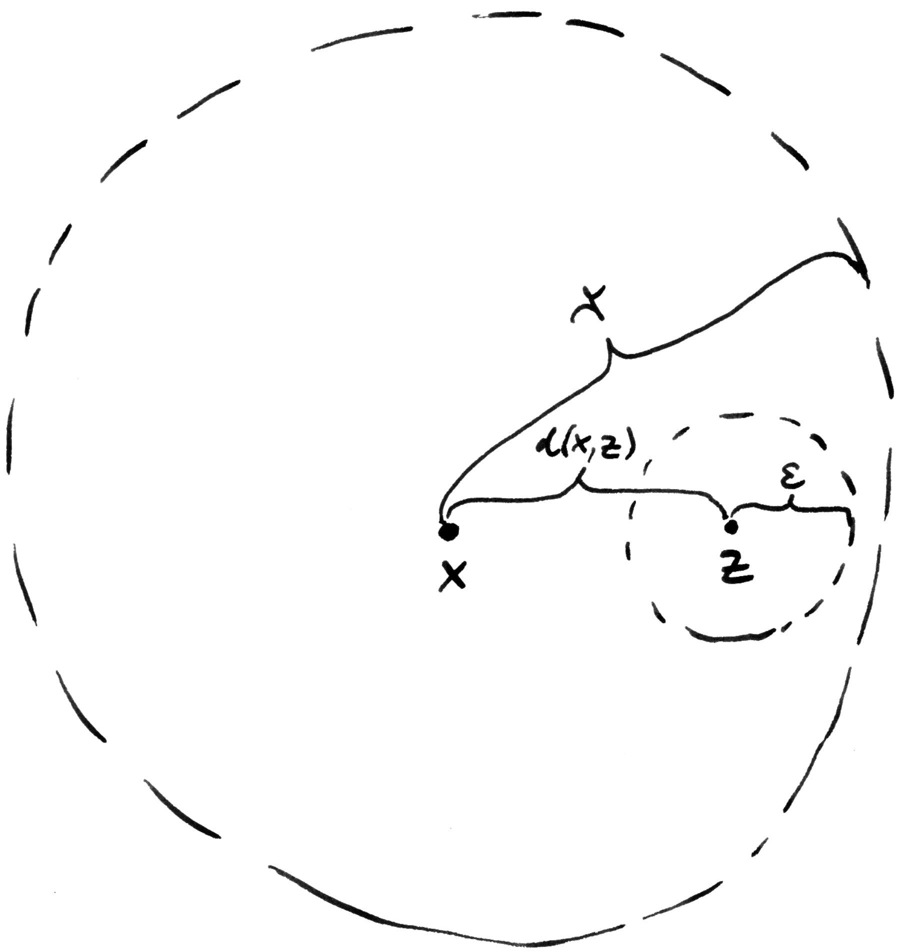


Illustration zu Beispiel 9.4.9: Ein Ball in einem metrischen Raum ist stets offen.

$\mathrm{B}(p;\varepsilon) \subset X\backslash M$, also ist $X\backslash M$ nicht offen. Ist $X\backslash M$ nicht offen, so gibt es einen Punkt $p \in X\backslash M$ derart, daß gilt

$$\mathrm{B}(p; 1/n) \cap M \neq \emptyset \quad \forall n \geq 1$$

Wählen wir jeweils einen Punkt $x_n \in \mathrm{B}(p; 1/n) \cap M$, so gilt $\lim_{n\to\infty} x_n = p$ und M ist nicht abgeschlossen. □

Übung 9.4.11. Nimmt man zu einer Teilmenge M eines metrischen Raums X alle ihre Berührungspunkte hinzu, so erhält man eine abgeschlossene Menge, genauer: Die kleinste abgeschlossene Menge, die M umfaßt. Diese Menge heißt auch der **Abschluß** von M in X und wird mit $\overline{M}$ bezeichnet.

Ergänzende Übung 9.4.12. Sei X ein metrischer Raum und seien $A, B \subset X$ disjunkte, abgeschlossene Teilmengen. So gibt es eine stetige Funktion $f : X \to [0, 1]$ mit $f|_A = 0$ und $f|_B = 1$. Hinweis: Man betrachte d_A wie in Übung 9.2.22 mache den Ansatz $f(z) = g(d_A(z), d_B(z))$ für geeignetes $g : \mathbb{R}^2\backslash 0 \to [0, 1]$.

Übung 9.4.13. Sei X ein metrischer Raum, $z \in X$ ein Punkt, $r \in \mathbb{R}$ eine reelle Zahl. So ist die Menge $\{x \in X \mid d(x, z) \leq r\}$ abgeschlossen.

Übung 9.4.14. Ist X ein metrischer Raum und $A \subset X$ eine Teilmenge, so kann ihr Abschluß $\bar{A}$ in der Notation von 9.2.22 beschrieben werden als die Menge $\bar{A} = \{x \in X \mid d(x, A) = 0\}$.

9.5 Topologische Räume

9.5.1. Der Begriff eines topologischen Raums wird erst sehr viel später unumgänglich werden, in der hier vorgesehenen Entwicklung der Theorie erst bei der Diskussion von abstrakten Mannigfaltigkeiten in **??** folgende. Daß ich ihn dennoch bereits hier einführe, hat mehrere Gründe. Zum Ersten erlaubt dieser Begriffsapparat eine große Vereinheitlichung: Zum Beispiel können in diesem Rahmen alle bisher betrachteten Grenzwertbegriffe unter einen Hut gebracht werden, wie in 9.6.9 ausgeführt wird. Zum Zweiten erlaubt er es, den Begriff der Stetigkeit im Kontext endlichdimensionaler reeller Vektorräume sehr unmittelbar zu fassen, indem man eben jeden derartigen Raum mit seiner "natürlichen Topologie" aus 9.9.22 versieht. Sobald das einmal getan ist, kann man auch in diesem Kontext mit Stetigkeit arbeiten, ohne Koordinaten einzuführen zu müssen. Und zum Dritten scheint es mir auch unabhängig davon wichtig, daß Sie beizeiten mit diesem Begriffsapparat vertraut werden, der die ganze Mathematik durchdringt: Um ihn an einfachen Beispielen einzuüben, will ich deshalb alles, was in dieser Vorlesung ohne große Umwege in der Allgemeinheit topologischer Räume formuliert und bewiesen werden kann, auch in diesem Rahmen formulieren und beweisen.

Vielfach werden die Aussagen und Beweise dadurch sogar einfacher, und ich denke, dieser Vorteil wiegt zum Teil bereits die zusätzlichen Schwierigkeiten auf, die durch das Erlernen dieses neuen Begriffsapparats und seiner Beziehungen zu den primären Zielen der Vorlesung sicher auch entstehen können.

9.5.2. Gegeben eine Menge X können wir die Menge $\mathcal{P}(X)$ aller Teilmengen von X bilden, die sogenannte Potenzmenge von X. Weil es mich verwirrt, über Mengen von Mengen zu reden, nenne ich wie in **??** Teilmengen von $\mathcal{P}(X)$ lieber **Systeme von Teilmengen von** X und spreche im folgenden von **Teilsystemen**, wenn ich Teilmengen solcher Mengensysteme meine.

Definition 9.5.3. Eine **Topologie** $\mathcal{T}$ auf einer Menge X ist ein System von Teilmengen $\mathcal{T} \subset \mathcal{P}(X)$, das stabil ist unter dem Bilden von endlichen Schnitten und beliebigen Vereinigungen. In Formeln ausgedrückt fordern wir von einer Topologie also

1. $U_1, \ldots, U_n \in \mathcal{T} \Rightarrow U_1 \cap \ldots \cap U_n \in \mathcal{T}$ für $n \geq 0$ und insbesondere auch $X \in \mathcal{T}$ als der Spezialfall $n = 0$. Gleichbedeutend dazu sind die beiden Forderungen $X \in \mathcal{T}$ und $U, V \in \mathcal{T} \Rightarrow U \cap V \in \mathcal{T}$;

2. $\mathcal{U} \subset \mathcal{T} \Rightarrow \bigcup_{U \in \mathcal{U}} U \in \mathcal{T}$ und damit insbesondere auch $\emptyset \in \mathcal{T}$, da ja das leere Mengensystem $\mathcal{U} = \emptyset$ in jedem Mengensystem enthalten ist.

Ein **topologischer Raum** ist ein Paar $(X, \mathcal{T})$ bestehend aus einer Menge mitsamt einer Topologie. Statt $U \in \mathcal{T}$ schreiben wir meist

$$U \Subset X$$

und nennen U eine **offene Teilmenge von** X. Die Notation $\Subset$ ist jedoch unüblich.

Definition 9.5.4. Sei X ein topologischer Raum und $p \in X$ ein Punkt. Eine Teilmenge $U \subset X$ heißt eine **Umgebung von** p genau dann, wenn es eine offene Menge $V \Subset X$ gibt mit $p \in V \subset U$.

Definition 9.5.5. Eine Abbildung $f : X \to Y$ zwischen topologischen Räumen heißt **stetig im Punkt** $p \in X$ genau dann, wenn es für jede Umgebung U von $f(p)$ eine Umgebung U' von p gibt mit $f(U') \subset U$. Eine Abbildung zwischen topologischen Räumen heißt **stetig** genau dann, wenn sie stetig ist in jedem Punkt.

Übung 9.5.6. Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen zwischen topologischen Räumen. Ist f stetig in $p \in X$ und g stetig in $f(p) \in Y$, so ist $g \circ f$ stetig in p. Hinweis: 9.2.15.

Beispiel 9.5.7. Für jeden metrischen Raum bildet das System seiner im Sinne von 9.4.6 offenen Teilmengen eine Topologie, die **metrische Topologie**. Wir fordern von einer Topologie *nicht*, daß ein beliebiger Schnitt offener Mengen stets wieder offen sein soll: Sonst müßten ja in unserem Beispiel der metrischen Räume alle einpunktigen Mengen offen sein, als Schnitte immer kleinerer Bälle. Da nach 9.4.9 Bälle in metrischen Räumen stets offen sind, ist in metrischen Räumen eine Umgebung eines Punktes im topologischen Sinne 9.5.4 dasselbe wie eine Umgebung im metrischen Sinne 9.2.7. Insbesondere ist eine Abbildung zwischen metrischen Räumen "topologisch stetig" im Sinne der obigen Definition 9.5.5 genau dann, wenn sie "metrisch stetig" ist im Sinne unserer Definition 9.2.10.

Beispiel 9.5.8. Auf unserer erweiterten reellen Zahlengeraden $\overline{\mathbb{R}}$ erklären wir eine Topologie durch die Vorschrift, daß eine Menge $U \subset \overline{\mathbb{R}}$ offen sein möge genau dann, wenn sie für jedes ihrer Elemente eine Umgebung im Sinne von 5.1.8 ist. Unsere Umgebungen im Sinne von 5.1.8 sind dann auch genau die Umgebungen für diese Topologie im Sinne von 9.5.4, und eine Abbildung $\overline{\mathbb{R}} \to \overline{\mathbb{R}}$ ist offensichtlich topologisch stetig im Sinne der obigen Definition 9.5.5 genau dann, wenn sie stetig ist im Sinne unserer Definition 6.1.3. Um die Beziehung zu unserem Stetigkeitsbegriff 6.1.3 für Abbildungen von einer Teilmenge $D \subset \overline{\mathbb{R}}$ nach $\overline{\mathbb{R}}$ zu klären, vereinbaren wir zunächst, in welcher Weise Teilmengen topologischer Räume mit einer Topologie versehen werden sollen.

Definition 9.5.9. Ist X ein topologischer Raum und $Y \subset X$ eine Teilmenge, so erklärt man die **induzierte Topologie** oder **Spurtopologie** auf Y durch die Vorschrift

$$U \circledcirc Y \Leftrightarrow \exists V \circledcirc X \text{ mit } U = V \cap Y$$

In Worten ist also eine Teilmenge von Y offen für die induzierte Topologie genau dann, wenn sie der Schnitt von Y mit einer offenen Teilmenge von X ist. Ab jetzt fassen wir stillschweigend jede Teilmenge Y eines topologischen Raums X auf als topologischen Raum mit der induzierten Topologie.

Übung 9.5.10. Man zeige, daß das in 9.5.9 beschriebene Mengensystem auf einer Teilmenge eines topologischen Raums in der Tat eine Topologie auf besagter Teilmenge liefert, und daß die Einbettungsabbildung stetig ist.

9.5.11. Wenn wir eine Menge einfach nur "offen" nennen, so in der Hoffnung, dem Leser sei klar, in Bezug auf welchen größeren Raum X dies "offen" gemeint ist. Ist X ein topologischer Raum und sind $M \subset Y \subset X$ Teilmengen, so meint $M \circledcirc Y$, daß M offen ist als Teilmenge des Raums Y mit seiner induzierten Topologie.

Übung 9.5.12. Man zeige, daß auf einer Teilmenge eines metrischen Raums die Spurtopologie zur metrischen Topologie mit der Topologie zur induzierten Metrik übereinstimmt.

Beispiel 9.5.13. Gegeben eine Teilmenge $D \subset \overline{\mathbb{R}}$ und eine Abbildung $f : D \to \overline{\mathbb{R}}$ ist f stetig an einer Stelle $p \in D$ im Sinne von 6.1.3 genau dann, wenn sie stetig ist bei p im topologischen Sinne für die auf D induzierte Topologie. Desgleichen ist unsere Abbildung stetig im Sinne von 6.1.3 genau dann, wenn sie stetig ist im topologischen Sinne.

Übung 9.5.14. Man zeige für jeden topologischen Raum: Der Schnitt von zwei Umgebungen eines Punktes ist wieder eine Umgebung besagten Punktes. Jede Umgebung eines Punktes kann verkleinert werden zu einer offenen Umgebung desselben Punktes.

Übung 9.5.15. Eine Teilmenge eines topologischen Raums ist offen genau dann, wenn sie für jeden ihrer Punkte eine Umgebung ist.

Übung 9.5.16. Sei X ein topologischer Raum und $U \Subset X$ eine offene Teilmenge. So ist eine Teilmenge $M \subset U$ offen in U genau dann, wenn sie offen ist in X. In Formeln gilt unter der Voraussetzung $U \Subset X$ für Teilmengen $M \subset U$ also $(M \Subset U \Leftrightarrow M \Subset X)$.

Satz 9.5.17 (Stetigkeit und Urbilder offener Mengen). *Eine Abbildung zwischen topologischen Räumen und insbesondere auch zwischen metrischen Räumen ist stetig genau dann, wenn darunter das Urbild jeder offenen Menge offen ist.*

Beweis. Sei $f : X \to Y$ stetig an jeder Stelle $p \in X$. Gegeben $U \Subset Y$ offen ist ja U Umgebung eines jeden seiner Punkte. Folglich gibt es für jede Stelle $p \in f^{-1}(U)$ eine Umgebung U'_p mit $f(U'_p) \subset U$. Diese U'_p können sogar offen gewählt werden, und damit ist $f^{-1}(U)$ offen als die Vereinigung aller U'_p mit $p \in f^{-1}(U)$. Ist umgekehrt $p \in X$ gegeben, so gibt es für jede Umgebung U von $f(p)$ eine offene, in U enthaltene Umgebung V von $f(p)$, und ist das Urbild jeder offenen Menge offen und $U' = f^{-1}(V)$ ist eine Umgebung von p mit $f(U') \subset U$. Ist also das Urbild jeder offenen Menge offen, so ist unsere Abbildung auch stetig an jeder Stelle p. □

9.5.18. Entwickelt man die Theorie der topologischen Räume ab initio, so wird man in der Regel die im vorhergehenden Satz enthaltene Charakterisierung wegen ihrer großen Eleganz gleich als Definition der Stetigkeit nehmen. Daß die Verknüpfung stetiger Abbildungen stetig ist, kann man von dieser Definition ausgehend sehr leicht und direkt einsehen, indem man beachtet, daß aus $f : X \to Y$ und $g : Y \to Z$ stetig folgt $V \Subset Z \Rightarrow g^{-1}(V) \Subset Y \Rightarrow f^{-1}(g^{-1}(V)) \Subset X$. Da nun gilt $f^{-1}(g^{-1}(V)) = (g \circ f)^{-1}(V)$, ist damit auch $(g \circ f)$ stetig.

Übung 9.5.19 (**Universelle Eigenschaft der induzierten Topologie**). Sei $f : X \to Y$ eine Abbildung zwischen topologischen Räumen und $Z \subset Y$ eine Teilmenge mit $f(X) \subset Z$. So ist f stetig genau dann, wenn die induzierte Abbildung

$f : X \to Z$ stetig ist für die auf Z induzierte Topologie. Analoges gilt für Stetigkeit in einem Punkt.

Beispiele 9.5.20. Es gibt auch Topologien, die unserer bis hierher entwickelten Anschauung eher ungewohnt sein mögen: Auf jeder Menge können wir etwa etwa die **Klumpentopologie** betrachten, die nur aus der ganzen Menge und der leeren Menge besteht, oder die **diskrete Topologie**, indem wir schlicht alle Teilmengen als offen ansehen. Einen topologischen Raum mit der diskreten Topologie nennen wir auch kurz einen **diskreten Raum**.

Beispiele 9.5.21. Jede konstante Abbildung ist stetig. Die Identität auf einem topologischen Raum ist immer stetig. Jede Abbildung in einen Raum mit der Klumpentopologie ist stetig. Jede Abbildung aus einem Raum mit der diskreten Topologie ist stetig.

Ergänzende Übung 9.5.22. Auf jeder Menge kann man die **koendliche Topologie** erklären durch die Vorschrift, daß außer der leeren Menge nur die Komplemente endlicher Mengen offen sein sollen.

Ergänzende Übung 9.5.23. Auf jeder partiell geordneten Menge kann man die **Ordnungstopologie**, auch genannt **Alexandroff-Topologie**, erklären durch die Vorschrift, daß genau die Teilmengen offen sein sollen, die mit einem Element auch jedes kleinere Element enthalten. Genau dann entsteht eine Topologie in dieser Weise aus einer partiellen Ordnung, wenn es für jedes Element eine kleinste offene Menge gibt, die es umfaßt.

Definition 9.5.24. Eine Teilmenge M eines topologischen Raums X heißt **abgeschlossen** oder präziser **abgeschlossen in** X und wir schreiben in Formeln $M \underset{A}{\subset} X$ genau dann, wenn ihr Komplement $X \backslash M$ offen ist.

9.5.25. Wenn wir eine Menge einfach nur "abgeschlossen" nennen, so in der Hoffnung, dem Leser sei klar, in Bezug auf welchen größeren Raum X dies "abgeschlossen" gemeint ist. Ist X ein topologischer Raum und sind $M \subset Y \subset X$ Teilmengen, so meint $M \underset{A}{\subset} Y$, daß M abgeschlossen ist als Teilmenge des Raums Y mit seiner induzierten Topologie 9.5.9. Die Terminologie kommt von dem Fall metrischer Räume her, in dem die Komplemente offener Mengen gerade diejenigen Teilmengen waren, die "abgeschlossen sind unter dem Bilden von Grenzwerten".

Übung 9.5.26. Gegeben ein topologischer Raum X mit einer Teilmenge Y zeige man: $A \underset{A}{\subset} Y \Leftrightarrow \exists B \underset{A}{\subset} X$ mit $A = B \cap Y$.

Übung 9.5.27. Sei X ein topologischer Raum und $A \underset{A}{\subset} X$ eine abgeschlossene Teilmenge. So ist eine Teilmenge $B \subset A$ abgeschlossen in A unter der Spurtopologie genau dann, wenn B abgeschlossen ist in X. In Formeln gilt unter der Voraussetzung $A \underset{A}{\subset} X$ für Teilmengen $B \subset A$ also $(B \underset{A}{\subset} A \Leftrightarrow B \underset{A}{\subset} X)$.

Lemma 9.5.28. *Jede endliche Vereinigung und beliebige Schnitte abgeschlossener Mengen sind abgeschlossen.*

Beweis. Das folgt sofort aus der Formel

$$X \backslash \bigcap_{M \in \mathcal{M}} M = \bigcup_{M \in \mathcal{M}} (X \backslash M)$$

die für jedes System $\mathcal{M} \subset \mathcal{P}(X)$ von Teilmengen einer Menge X gilt. □

Definition 9.5.29. Gegeben ein topologischer Raum X und eine Teilmenge $M \subset X$ gibt es stets eine kleinste abgeschlossene Teilmenge von X, die M umfaßt, nämlich den Schnitt über alle abgeschlossenen Teilmengen von X, die M umfassen. Wir notieren sie $\mathrm{Cl}_X(M) = \mathrm{Cl}(M) = \overline{M}$ und nennen sie den **Abschluß von** M oder genauer den **Abschluß von** M **in** X.

9.5.30. Eine Abbildung ist stetig genau dann, wenn darunter das Urbild jeder abgeschlossenen Menge abgeschlossen ist: Das folgt unmittelbar aus Satz 9.5.17, da das Urbild des Komplements einer Menge stets das Komplement ihres Urbilds ist.

Beispiel 9.5.31. Wir geben einen neuen Beweis für die Erhaltung von Ungleichungen im Grenzwert 5.1.33, der zwar nur für reelle Folgen mit reellen Grenzwerten funktioniert, aber dafür viele Möglichkeiten der Verallgemeinerung aufzeigt. Zunächst ist die Menge $H = \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$ abgeschlossen in $\mathbb{R}^2$ nach 9.5.30 als Urbild der abgeschlossenen Menge $[0, \infty) \underset{\mathcal{A}}{\subset} \mathbb{R}$ unter der stetigen Abbildung $(x, y) \mapsto y - x$. Ist (x_n, y_n) eine konvergente Folge in H, so liegt mithin auch ihr Grenzwert in H, und das bedeutet gerade die Erhaltung von Ungleichungen im Grenzwert.

Proposition 9.5.32. *Sei $f : X \to Y$ eine Abbildung topologischer Räume.*

1. *Sei $\mathcal{U}$ eine* **offene Überdeckung** *von X, d.h. ein System offener Teilmengen mit $X = \bigcup_{U \in \mathcal{U}} U$. So ist f stetig genau dann, wenn $f|_U$ stetig ist für alle $U \in \mathcal{U}$.*

2. *Sei X überdeckt von* endlich vielen *abgeschlossenen Teilmengen, in Formeln $A_1, \ldots, A_n \underset{\mathcal{A}}{\subset} X$ und $X = \bigcup_{i=1}^n A_i$. So ist f stetig genau dann, wenn $f|_{A_i}$ stetig ist für alle $i = 1, \ldots n$.*

Beweis. Ist f stetig, so sind alle $f|_U$ stetig als Verknüpfung von f mit der stetigen Inklusion $U \hookrightarrow X$. Sind andererseits alle $f|_U$ stetig, so ist für alle $W \overset{\circ}{\subset} Y$ und alle $U \in \mathcal{U}$ das Urbild $f^{-1}(W) \cap U$ offen in U, nach 9.5.16 ist also $f^{-1}(W) \cap U$ sogar offen in X, und damit ist dann natürlich auch $f^{-1}(W) = \bigcup_{U \in \mathcal{U}} f^{-1}(W) \cap$

U offen in X als Vereinigung offener Mengen. Mithin ist f stetig. Teil 2 zeigt man ähnlich: Nach 9.5.30 muß nur gezeigt werden, daß für jede abgeschlossene Teilmenge $B \underset{\mathrm{A}}{\subset} Y$ von Y ihr Urbild $f^{-1}(B)$ abgeschlossen ist in X. Da aber gilt $f^{-1}(B) = f_1^{-1}(B) \cup \ldots \cup f_n^{-1}(B)$ und $f_i^{-1}(B) \underset{\mathrm{A}}{\subset} A_i$ nach Annahme folgt die Proposition aus 9.5.27 und den Definitionen. □

Übung 9.5.33. Seien X ein topologischer Raum und Y, Z metrische Räume. Man zeige, daß eine Abbildung $(f, g) : X \to Y \times Z$ stetig ist genau dann, wenn f und g stetig sind. Man zeige, daß Produkt und Summe von stetigen reellwertigen Funktionen auf einem topologischen Raum wieder stetig sind.

Vorschau 9.5.34. In **??** werden wir erklären, wie man ganz allgemein das Produkt topologischer Räume so mit einer Topologie versehen kann, daß das Analogon der vorhergehenden Übung auch für beliebige topologische Räume Y, Z gilt.

9.6 Grenzwerte in topologischen Räumen

Definition 9.6.1. Sei $\mathbb{N} \to Y$, $n \mapsto y_n$ eine Folge in einem topologischen Raum Y und $b \in Y$ ein Punkt. Wir sagen, die Folge y_n **strebt gegen** b oder **konvergiert gegen** b und nennen b einen **Grenzwert der Folge** und schreiben

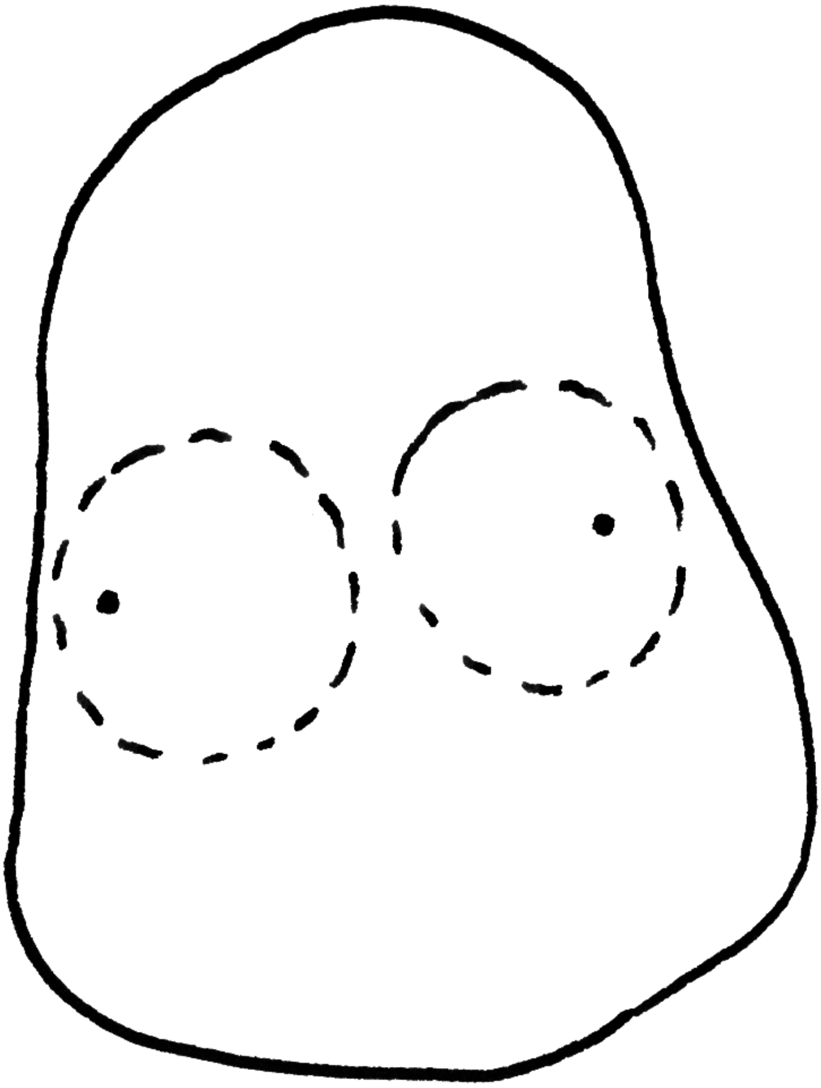
$$\lim_{n \to \infty} y_n = b$$

genau dann, wenn jede Umgebung von b fast alle Glieder unserer Folge enthält. Gleichbedeutend können wir ebensogut auch fordern, daß jede offene Menge um b fast alle Glieder unserer Folge enthält.

9.6.2. In dieser Allgemeinheit ist der Grenzwertbegriff nur noch eingeschränkt sinnvoll, da der Grenzwert einer Folge nicht mehr eindeutig zu sein braucht. Fordern wir jedoch von unserem topologischen Raum die sogenannte **Hausdorff-Eigenschaft**, daß je zwei verschiedene Punkte darin disjunkte Umgebungen besitzen, so ist der Grenzwert einer Folge eindeutig, wenn er existiert. Ein topologischer Raum mit der Hausdorff-Eigenschaft heißt ein **Hausdorff-Raum**.

Vorschau 9.6.3. Für allgemeine topologische Räume ist es nicht mehr richtig, daß jede unter der Bildung von Grenzwerten von Folgen abgeschlossene Teilmenge auch tatsächlich abgeschlossen ist. Ein Gegenbeispiel gebe ich in **??**, eine Zusatzbedingung, unter der das doch wieder gilt, in **??**.

Übung 9.6.4. Konvergiert eine Folge von stetigen Funktionen von einem topologischen Raum in einen metrischen Raum gleichmäßig, so ist auch die Grenzfunktion stetig. Hinweis: Man kopiere den Beweis von 8.1.10.



In einem Hausdorffraum haben je zwei verschiedene Punkte disjunkte Umgebungen.

Definition 9.6.5. Ein Punkt eines topologischen Raums heißt ein **Häufungspunkt** genau dann, wenn jede seiner Umgebungen auch noch andere Punkte unseres Raums enthält, wenn also in anderen Worten die nur aus besagtem Punkt bestehende Menge nicht offen ist. Ist $D \subset X$ eine Teilmenge eines topologischen Raums und $p \in X$ ein Häufungspunkt von $D \cup \{p\}$ im Sinne der vorhergehenden Definition, so sagen wir auch, p sei ein ein **Häufungspunkt von** D **in** X.

Ergänzung 9.6.6. Manche Autoren erklären auch noch die "Häufungspunkte einer Folge in einem topologischen Raum X" als die Punkte von X mit der Eigenschaft, daß in jeder ihrer Umgebungen unendlich viele Folgenglieder liegen.

Übung 9.6.7. Genau dann ist p Häufungspunkt des metrischen Raums X, wenn es eine Folge x_n in $X \backslash p$ gibt mit $\lim_{n\to\infty} x_n = p$.

Definition 9.6.8 (Grenzwerte von Abbildungen). Seien X, Y topologische Räume, $p \in X$ ein Häufungspunkt und $f : X \backslash p \to Y$ eine Abbildung. Sei weiter b ein Punkt aus Y. Wir sagen, $f(x)$ **strebt gegen** b **für** $x \to p$ und schreiben

$$\lim_{x\to p} f(x) = b$$

genau dann, wenn es für jede Umgebung W des Grenzwerts b eine Umgebung W' des Punktes p gibt mit $f(W' \backslash p) \subset W$.

9.6.9. Die vorstehende Definition verallgemeinert alle bisher betrachteten Grenzwertbegriffe: Den Grenzwertbegriff für Folgen in den erweiterten reellen Zahlen nach 5.1.15, für Abbildungen zwischen Teilmengen der erweiterten reellen Zahlen 6.3.6, für Folgen in metrischen Räumen 9.3.1 und auch für Folgen in topologischen Räumen 9.6.1. Der Fall von Folgen ist jeweils der Spezialfall, in dem wir $X = \mathbb{N} \sqcup \{\infty\}$ mit der von $\overline{\mathbb{R}}$ induzierten Topologie versehen und darin den Häufungspunkt $p = \infty$ betrachten. Ich habe den topologischen Raum bei der Definition der Folgenkonvergenz 9.6.1 nur deshalb etwas ungewöhnlich mit Y bezeichnet, um deutlich zu machen, inwiefern es sich dabei um einem Spezialfall unseres allgemeinen Grenzwertbegriffs 9.6.8 handelt.

Übung 9.6.10. Auch in der Allgemeinheit von 9.6.8 ist der Grenzwert eindeutig, wenn er existiert und der Wertebereich Hausdorff ist.

Übung 9.6.11. Seien X, Y topologische Räume, $p \in X$ ein Häufungspunkt und $f : X \to Y$ eine Abbildung. Man zeige, daß auch in dieser Allgemeinheit $\lim_{x\to p} f(x) = f(p)$ gleichbedeutend ist zur Stetigkeit von f bei p. Man diskutiere des weiteren Analoga zu 6.3.21.

Übung 9.6.12. Seien X ein topologischer Raum, $p \in X$ ein Häufungspunkt und $f_i : X \backslash p \to Y_i$ Abbildungen in metrische Räume für $1 \le i \le n$. Sei $Y = Y_1 \times \ldots \times Y_n$ das Produkt und $f = (f_1, \ldots, f_n) : X \to Y$. So ist $\lim_{x\to p} f(x) = b$ für $b = (b_1, \ldots, b_n)$ gleichbedeutend zu $\lim_{x\to p} f_i(x) = b_i \ \forall i$.

Übung 9.6.13 (**Quetschlemma**). Seien X ein topologischer Raum, $p \in X$ ein Häufungspunkt und $f, g, h : X \backslash p \to \mathbb{R}$ Funktionen mit der Eigenschaft $f(x) \leq g(x) \leq h(x)\ \forall x \in X \backslash p$. So folgt aus $\lim_{x \to p} f(x) = b = \lim_{x \to p} h(x)$ bereits $\lim_{x \to p} g(x) = b$.

Übung 9.6.14. Im Fall einer Abbildung in einen metrischen Raum mit Metrik d ist $\lim_{x \to p} f(x) = y$ gleichbedeutend zu $\lim_{x \to p} d(f(x), y) = 0$.

Ergänzung 9.6.15. Gegeben Abbildungen $f : X \to Y$, $g : X \to Z$ von topologischen Räumen mit Y Hausdorff und $x \in X$, $y \in Y$, $z \in Z$ findet man manchmal die Notation

$$\lim_{g(x) \to z} f(x) = y$$

Das soll dann bedeuten, daß es für jede Umgebung U von y eine Umgebung V von z gibt mit $g(x) \in (V \backslash z) \Rightarrow f(x) \in U$. Zum Beispiel gilt für jedes nicht konstante Polynom $P : \mathbb{C} \to \mathbb{C}$ die Formel $\lim_{|z| \to \infty} |P(z)| = \infty$.

Ergänzung 9.6.16. Um durchgehend mit topologischen Räumen arbeiten zu können, müßten wir nun erklären, wie das Produkt zweier topologischer Räume mit einer Topologie zu versehen ist, und allerhand Eigenschaften wie etwa das Analogon der Komponentenregel prüfen. Das alles werde ich voerest vermeiden und erst in **??** diskutieren, weil ich fürchte, den Sinn dieser Abstraktionen hier noch nicht ausreichend begründen zu können.

9.7 Kompakte metrische Räume

Definition 9.7.1. Ein metrischer Raum heißt **kompakt** oder ausführlicher **folgenkompakt** genau dann, wenn jede Folge in unserem Raum eine konvergente Teilfolge besitzt.

9.7.2. Eine Teilmenge A eines metrischen Raums nennen wir kompakt oder auch ein **Kompaktum** genau dann, wenn sie kompakt ist als metrischer Raum mit der induzierten Metrik, wenn also jede Folge in A eine Teilfolge besitzt, die *gegen einen Punkt aus* A konvergiert.

Übung 9.7.3. Endliche Vereinigungen kompakter Teilmengen eines metrischen Raums sind stets wieder kompakt.

9.7.4. Jeder kompakte metrische Raum ist beschränkt. Ist in der Tat ein Raum nicht beschränkt, so finden wir darin eine Folge x_n mit $d(x_0, x_n) \geq n$, und diese Folge kann nach 9.3.8 keine konvergente Teilfolge haben.

Proposition 9.7.5. *Jedes endliche Produkt von kompakten metrischen Räumen ist kompakt.*

Beweis. Sei $X = X_1 \times \ldots \times X_n$ mit kompakten X_i. Sei eine Folge in X gegeben. Da X_1 kompakt ist, finden wir eine Teilfolge unserer Folge, die in der ersten Koordinate konvergiert. Da auch X_2 kompakt ist, finden wir von dieser Teilfolge hinwiederum eine Teilfolge, die auch in der zweiten Koordinate konvergiert. Indem wir so weitermachen, finden wir schließlich eine Teilfolge, die in jeder Koordinate konvergiert. Diese Teilfolge konvergiert dann nach 9.3.3 auch in X. □

Lemma 9.7.6. *Eine kompakte Teilmenge eines metrischen Raums ist stets abgeschlossen.*

Beweis. Sei X unser Raum und $A \subset X$ unsere Teilmenge. Ist A nicht abgeschlossen, so gibt es eine Folge in A, die gegen einen Punkt aus $X \backslash A$ konvergiert. Solch eine Folge kann aber unmöglich eine Teilfolge haben, die gegen einen Punkt aus A konvergiert. □

Übung 9.7.7. Eine Teilmenge eines kompakten metrischen Raums ist kompakt genau dann, wenn sie abgeschlossen ist.

Satz 9.7.8 (Heine-Borel). *Eine Teilmenge des $\mathbb{R}^n$ ist kompakt genau dann, wenn sie beschränkt und abgeschlossen ist.*

Beweis. Nach 9.7.4 und 9.7.6 ist eine kompakte Teilmenge eines metrischen Raums stets beschränkt und abgeschlossen. In der anderen Richtung wissen wir schon, daß für jedes $k \geq 0$ das Intervall $[-k, k]$ kompakt ist. Falls eine Teilmenge $A \subset \mathbb{R}^n$ beschränkt ist, finden wir ein k mit $A \subset [-k, k]^n$. Nach 9.7.5 ist nun $[-k, k]^n$ kompakt, und als abgeschlossene Teilmenge eines kompakten Raums ist nach 9.7.3 dann auch A selbst kompakt. □

Beispiel 9.7.9. Die Menge $[0, 1] \cap \mathbb{Q}$ ist abgeschlossen in $\mathbb{Q}$ und beschränkt, ist aber nicht kompakt für die induzierte Metrik.

Proposition 9.7.10. *Unter einer stetigen Abbildung metrischer Räume werden Kompakta stets auf Kompakta abgebildet.*

Beweis. Sei $f : X \to Y$ unsere stetige Abbildung und $A \subset X$ ein Kompaktum. Ist y_n eine Folge in $f(A)$, so finden wir eine Folge x_n in A mit $f(x_n) = y_n$. Falls A kompakt ist, besitzt die Folge x_n eine Teilfolge x_{n_k}, die gegen einen Punkt $x \in A$ konvergiert. Dann ist y_{n_k} nach 9.3.9 eine Teilfolge der Folge y_n, die gegen einen Punkt von $f(A)$ konvergiert, nämlich gegen $f(x)$. □

Korollar 9.7.11. *Jede stetige reellwertige Funktion auf einem kompakten metrischen Raum ist beschränkt und nimmt, wenn unser Raum nicht leer ist, das Supremum und das Infimum der Menge ihrer Funktionswerte an.*

9.7.12. Ist also in Formeln X ein nichtleerer kompakter Raum und $f : X \to \mathbb{R}$ stetig, so gibt es $p, q \in X$ mit $f(p) \leq f(x) \leq f(q) \;\; \forall x \in X$.

Beweis. Nach 9.7.10 ist $f(X) \subset \mathbb{R}$ kompakt, also beschränkt und abgeschlossen. Aus $X \neq \emptyset$ folgt weiter $f(X) \neq \emptyset$. Damit besitzt $f(X)$ ein Supremum und ein Infimum in $\mathbb{R}$. Da aber $f(X)$ kompakt, also abgeschlossen ist, folgt $\sup f(X) \in f(X)$ und $\inf f(X) \in f(X)$. Es gibt in anderen Worten $p, q \in X$ mit $\sup f(X) = f(p)$ und $\inf f(X) = f(q)$. □

Definition 9.7.13. Eine stetige Abbildung von metrischen Räumen heißt **gleichmäßig stetig** genau dann, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt derart, daß gilt

$$d(x, y) < \delta \;\;\Rightarrow\;\; d(f(x), f(y)) < \varepsilon$$

Satz 9.7.14 (Gleichmäßige Stetigkeit auf Kompakta). *Jede stetige Abbildung von einem kompakten metrischen Raum in einen weiteren metrischen Raum ist gleichmäßig stetig.*

Beweis. Mutatis mutandis zeigt das der Beweis von Satz 6.4.11. □

Ergänzende Übung 9.7.15. Ist in einem metrischen Raum eine abzählbare Familie kompakter Teilmengen $(K_n)_{n\in\mathbb{N}}$ gegeben mit leerem Schnitt $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$, so gibt es schon ein N mit $K_0 \cap \ldots \cap K_N = \emptyset$. Das wird verallgemeinert auf den Fall beliebiger Familien in 9.10.8.

Ergänzende Übung 9.7.16. Sei (X, d) ein metrischer Raum, $K \subset X$ kompakt und $A \subset X$ abgeschlossen mit $A \cap K = \emptyset$. So gibt es $\delta > 0$ mit $d(x, y) \geq \delta$ für alle $x \in A$, $y \in K$.

Ergänzende Übung 9.7.17. Man zeige, daß man auf dem Raum $\operatorname{Ens}(\mathbb{N}, \{W, Z\})$ aller Folgen in der zweielementigen Menge $\{W, Z\}$ eine Metrik erklären kann durch die Vorschrift $d(\omega, \eta) = 2^{-n}$ für $n \in \mathbb{N}$ die kleinste Zahl mit $\omega(n) \neq \eta(n)$, bzw. $d(\omega, \eta) = 0$ falls $\omega = \eta$. Man zeige weiter, daß der so gebildete metrische Raum kompakt ist. Nebenbei bemerkt denke ich hier bei W an "Wappen" und bei Z an "Zahl" und bei der Übung an Anwendungen in der Wahrscheinlichkeitstheorie.

9.8 Affine Räume

9.8.1. Dieser Abschnitt ist ein Auszug aus Abschnitt **??** der linearen Algebra. Ich habe ihn hier nur eingefügt, um Unklarheiten zu vermeiden, was die im weiteren verwendeten Notationen und Begriffsbildungen angeht.

Definition 9.8.2. Ein **affiner Raum** oder kurz **Raum** über einem Körper k ist ein Tripel

$$E = (E, \vec{E}, a)$$

bestehend aus einer nichtleeren Menge E, einer abelschen Gruppe $\vec{E} \subset \mathrm{Ens}^{\times} E$ von Permutationen von E, von der man fordert, daß für alle $p \in E$ das Anwenden auf p eine Bijektion $\vec{E} \xrightarrow{\sim} E$ besagter Gruppe mit unserem Raum liefert, sowie einer Abbildung $a : k \times \vec{E} \to \vec{E}$, die die abelsche Gruppe $\vec{E}$ zu einem k-Vektorraum macht. Die Elemente von $\vec{E}$ heißen die **Translationen** oder **Richtungsvektoren** unseres affinen Raums und den Vektorraum $\vec{E}$ selbst nennen wir den **Richtungsraum** unseres affinen Raums E. Die Operation von k auf $\vec{E}$ mag man die **Reskalierung von Translationen** nennen. Unter der **Dimension** unseres affinen Raums verstehen wir die Dimension seines Richtungsraums. Das Resultat der Operation von $\vec{v} \in \vec{E}$ auf $p \in E$ notieren wir $\vec{v} + p := \vec{v}(p)$ oder manchmal auch $p + \vec{v}$.

9.8.3. Die eben eingeführte Notation für den Richtungsraum eines affinen Raums steht leider in Konflikt mit der Notation aus 10.4.7, nach der mit Pfeilen versehene Mannigfaltigkeiten orientierte Mannigfaltigkeiten andeuten sollen. Was jeweils gemeint ist, muß der Leser aus dem Kontext erschließen.

9.8.4. Ist E ein affiner Raum, so liefert nach Annahme für jedes $p \in E$ die Operation eine Bijektion $\vec{E} \xrightarrow{\sim} E$, $\vec{u} \mapsto \vec{u} + p$ und es gilt $\vec{0} + p = p$ sowie $\vec{u} + (\vec{v} + p) = (\vec{u} + \vec{v}) + p$ für alle $\vec{u}, \vec{v} \in \vec{E}$ und $p \in E$. Flapsig gesprochen ist also ein affiner Raum ein “Vektorraum, bei dem man den Ursprung vergessen hat”. Gegeben $p, q \in E$ definieren wir $p - q$ als den Richtungsvektor $\vec{u} \in \vec{E}$ mit $p = \vec{u} + q$.

Beispiel 9.8.5. Jeder Vektorraum ist in offensichtlicher Weise auch ein affiner Raum. Es scheint mir besonders sinnfällig, den uns umgebenden Raum mathematisch als dreidimensionalen reellen affinen Raum zu modellieren: Hierbei denkt man sich $\vec{E}$ als die Gruppe aller “Parallelverschiebungen”. Ähnlich mag man die Zeit modellieren als einen eindimensionalen reellen affinen Raum. Die leere Menge kann in meinen Konventionen nie ein affiner Raum sein, es gibt hierzu jedoch auch andere Konventionen.

9.8.6. Meist findet man die begriffliche Variante eines **affinen Raums über einem vorgegebenen Vektorraum**: Darunter versteht man dann eine Menge E mit einer freien transitiven Operation des vorgegebenen Vektorraums. Ich ziehe die oben gegebene Variante vor, da sie jeden Bezug auf einen vorgegebenen Vektorraum vermeidet und den Anschauungsraum meines Erachtens besser modelliert.

Definition 9.8.7. Eine Abbildung $\varphi : E \to E'$ zwischen affinen Räumen heißt eine **affine Abbildung** genau dann, wenn es eine lineare Abbildung zwischen den zugehörigen Richtungräumen $\vec{\varphi} : \vec{E} \to \vec{E}'$ gibt mit

$$\varphi(p) - \varphi(q) = \vec{\varphi}(p - q) \ \ \forall p, q \in E$$

Diese lineare Abbildung $\vec{\varphi}$ ist dann durch φ eindeutig bestimmt und heißt der **lineare Anteil** unserer affinen Abbildung.

Übung 9.8.8. Die Verknüpfung affiner Abbildungen ist affin und der lineare Anteil einer Verknüpfung affiner Abbildungen ist die Verknüpfung ihrer linearen Anteile.

9.9 Normierte Räume

9.9.1. Unter einem **reellen Vektorraum** bzw. einem **reellen Raum** verstehen wir einen Vektorraum bzw. einen affinen Raum über dem Körper der reellen Zahlen. Wollen wir einen reellen Vektorraum bzw. affinen Raum mit einer Metrik versehen, so reicht es, wenn wir jedem seiner Vektoren bzw. Richtungsvektoren in geeigneter Weise eine "Länge" zuordnen. Einen solchen abstrakten Längenbegriff für die Vektoren eines Vektorraums nennt man eine "Norm". Die Details folgen.

Definition 9.9.2. Sei V ein reeller Vektorraum. Eine **Norm** auf V ist eine Abbildung $\| \ \| : V \to \mathbb{R}_{\geq 0}$, $v \mapsto \|v\|$ derart, daß gilt:

1. $\|\lambda v\| = |\lambda| \cdot \|v\| \quad \forall v \in V, \lambda \in \mathbb{R};$
2. $\|v\| = 0 \Leftrightarrow v = 0;$
3. $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V.$

Unter einem **normierten Vektorraum** versteht man ein Paar $(V, \| \ \|)$ bestehend aus einem Vektorraum V und einer Norm $\| \ \|$ auf V.

Ergänzung 9.9.3. Für Leser, die schon mit komplexen Zahlen vertraut sind, sei noch erwähnt, daß man von einer Norm auf einem komplexen Vektorraum stärker fordert, daß die erste Bedingung sogar für alle $\lambda \in \mathbb{C}$ gelten soll, wobei $|\lambda|$ als die "Norm der komplexen Zahl λ" im Sinne von **??** zu verstehen ist.

9.9.4. Jeder normierte Vektorraum wird ein metrischer Raum vermittels der **durch die Norm induzierten Metrik**

$$d(v, w) = \|v - w\|$$

Zum Beispiel gehört unser Betragsabstand auf dem $\mathbb{R}^n$ zur Maximumsnorm. Wir dürfen damit in normierten Vektorräumen über Stetigkeit und Konvergenz von Folgen reden. Allgemeiner verstehen wir unter einem **normierten affinen Raum** einen reellen (oder komplexen) affinen Raum im Sinne von 9.8.2, dessen Richtungsraum mit einer Norm versehen ist. Auch jeder normierte affine Raum trägt eine natürliche Metrik, die durch dieselbe Formel beschrieben wird. Reden wir ohne nähere Spezifikation von einem **normierten Raum**, so meinen wir einen normierten affinen Raum. Leser, die mit dem Begriff eines affinen Raums noch nicht vertraut sind, mögen sich aber auch einen normierten Vektorraum denken.

Übung 9.9.5. Für je zwei Vektoren v, w eines normierten Vektorraums gilt $\|v + w\| \geq |\|v\| - \|w\||$.

Beispiel 9.9.6. Mit $v \mapsto \|v\|$ ist auch $v \mapsto \alpha\|v\|$ eine Norm, für jedes $\alpha > 0$. Auf dem Nullraum gibt es nur eine Norm, die eben den Nullvektor auf Null wirft.

Beispiel 9.9.7. Auf dem $\mathbb{R}^n$ definiert man die **euklidische Norm** von $v = (v_1, \ldots, v_n)$ durch $\|v\| = \|v\|_2 = \sqrt{\langle v, v\rangle} = \sqrt{v_1^2 + \ldots + v_n^2}$. Wie man formal zeigt, daß das tatsächlich eine Norm ist, diskutieren wir in **??**.

Beispiel 9.9.8. Auf dem $\mathbb{R}^n$ für $n > 0$ definiert man die **Maximumsnorm** von $v = (v_1, \ldots, v_n)$ durch $|v| = \|v\|_\infty = \max(|v_1|, \ldots, |v_n|)$. Auf dem Raum $V = \operatorname{Ens}^{\mathrm{b}}(D, \mathbb{R})$ aller beschränkten reellwertigen Funktionen auf einer Menge D haben wir die **Supremumsnorm**, gegeben für $D \neq \emptyset$ durch

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in D\}$$

und im Fall $D = \emptyset$ als die einzig mögliche Norm auf dem Nullraum. Für eine endliche Menge D mit n Punkten erhalten wir unsere Maximumsnorm auf dem $\mathbb{R}^n$ als Spezialfall der Supremumsnorm. Noch allgemeiner definieren wir für jeden normierten Vektorraum $(W, |\ |)$ auf dem Raum $V = \operatorname{Ens}^{\mathrm{b}}(D, W)$ aller beschränkten Abbildungen von D nach W die Supremumsnorm durch $\|f\|_\infty = \sup\{|f(x)| \mid x \in D\}$ im Fall $D \neq \emptyset$ und im Fall $D = \emptyset$ als die einzig mögliche Norm auf dem Nullraum. Die zu unserer Supremumsnorm gehörige Metrik ist in allen diesen Fällen die Metrik der gleichmäßigen Konvergenz.

Beispiel 9.9.9. Sind $V_1, \ldots, V_n$ normierte Vektorräume, so erklären wir die **Produktnorm** auf ihrem Produkt $V_1 \times \ldots \times V_n$ für $n > 0$ durch die Vorschrift $\|(v_1, \ldots, v_n)\| = \sup \|v_i\|$. Offensichtlich induziert die Produktnorm die Produktmetrik.

Übung 9.9.10. Gegeben ein normierter Vektorraum $(V, \|\ \|)$ sind die folgenden Abbildungen stetig: Die Norm $\|\ \| : V \to \mathbb{R}$, die Addition $V \times V \to V$, und die Multiplikation mit Skalaren $\mathbb{R} \times V \to V$. Ist unsere Norm die euklidische Norm zu einem Skalarprodukt $V \times V \to \mathbb{R}$, so ist auch dies Skalarprodukt stetig. Leser, die bereits mit komplexen Zahlen vertraut sind, zeigen Analoges auch für komplexe Vektorräume.

Ergänzende Übung 9.9.11 (**Einparameteruntergruppen normierter Vektorräume**). Die stetigen Gruppenhomomorphismen von der additive Gruppe der reellen Zahlen in einen normierten Vektorraum sind genau die linearen Abbildungen. Hinweis: 6.3.26.

Ergänzende Übung 9.9.12. In einem normierten reellen Vektorraum ist jede nichtleere offene Teilmenge bereits ein Erzeugendensystem.

Satz 9.9.13 (Stetigkeit linearer Abbildungen). *Eine lineare Abbildung $f : V \to W$ zwischen normierten Vektorräumen ist stetig genau dann, wenn es eine Konstante $C \geq 0$ gibt derart, daß gilt*

$$\|f(v)\| \leq C\|v\| \quad \forall v \in V$$

9.9.14. Sie werden in 9.9.23 zeigen, daß lineare Abbildungen von einem *endlichdimensionalen* normierten Vektorraum in einen beliebigen weiteren normierten Vektorraum immer stetig sind.

Beweis. Ist f stetig, so gibt es $\delta > 0$ mit $\|v - 0\| \leq \delta \Rightarrow \|f(v) - f(0)\| \leq 1$. Setzen wir $C = 1/\delta$, so folgt $\|f(v)\| \leq C\|v\|$ zunächst für alle Vektoren v der Norm $\|v\| = \delta$ und dann durch Multiplikation mit Skalaren für alle $v \in V$. Gibt es umgekehrt ein $C > 0$ mit $\|f(v)\| \leq C\|v\| \quad \forall v \in V$, so finden wir für alle $\varepsilon > 0$ ein $\delta = \varepsilon/C > 0$ so daß gilt

$$\|v - w\| \leq \delta \quad \Rightarrow \quad \|f(v) - f(w)\| = \|f(v - w)\| \leq C\delta = \varepsilon$$ □

Übung 9.9.15. Man zeige: Jede stetige lineare Abbildung zwischen normierten Vektorräumen ist gleichmäßig stetig.

Übung 9.9.16. Die Menge aller stetigen reellwertigen Funktionen auf einem Raum X notiere ich $\mathcal{C}(X, \mathbb{R})$. Das $\mathcal{C}$ steht hier für englisch "continous" und französisch "continu". Man zeige: Versehen wir die Menge $\mathcal{C}([a, b], \mathbb{R})$ aller stetigen reellwertigen Funktionen auf einem kompakten reellen Intervall $[a, b]$ mit der Supremumsnorm, so wird das Integral $f \mapsto \int_a^b f(t)\,\mathrm{d}t$ eine stetige Abbildung $\mathcal{C}([a, b], \mathbb{R}) \to \mathbb{R}$.

Übung 9.9.17. Bezeichnet $\mathcal{C}^1([a, b], \mathbb{R}) \subset \mathcal{C}([a, b], \mathbb{R})$ den Teilraum der einmal stetig differenzierbaren Funktionen, so ist das Ableiten $f \mapsto f'$ *keine* stetige Abbildung $\mathcal{C}^1([a, b], \mathbb{R}) \to \mathcal{C}([a, b], \mathbb{R})$.

Übung 9.9.18. Seien U, V, W normierte Vektorräume. Eine bilineare Abbildung $F : U \times V \to W$ ist stetig genau dann, wenn es eine Konstante $C > 0$ gibt mit $\|F(u, v)\| \leq C\|u\|\|v\|$. Man formuliere und beweise die analoge Aussage auch für multilineare Abbildungen.

Übung 9.9.19. Gegeben eine Menge D und ein normierter Vektorraum V erkläre man auf dem Raum $\mathrm{Ens}^{\mathrm{b}}(D, V)$ der beschränkten Abbildungen $D \to V$ eine Norm derart, daß die zugehörige Metrik die Metrik der gleichmäßigen Konvergenz aus 9.3.5 wird.

Definition 9.9.20. Zwei Normen $\| \ \|, | \ |$ auf einem reellen Vektorraum V heißen **äquivalent** genau dann, wenn es positive Konstanten $c, C > 0$ gibt derart, daß gilt

$$\|v\| \leq C|v| \text{ und } |v| \leq c\|v\| \quad \forall v \in V$$

Satz 9.9.21 (Äquivalenz von Normen). *Auf einem endlichdimensionalen reellen Vektorraum sind je zwei Normen äquivalent.*

Beweis. Wir dürfen ohne Beschränkung der Allgemeinheit annehmen, daß V der $\mathbb{R}^n$ ist mit $n \geq 1$ und daß eine unserer Normen die Maximumsnorm $|v|$ ist. Sei $\| \ \|$ eine zweite Norm. Bezeichnet $\mathrm{e}_1, \ldots, \mathrm{e}_n$ die Standardbasis des $\mathbb{R}^n$ und ist $v = v_1 \mathrm{e}_1 + \ldots + v_n \mathrm{e}_n$, so haben wir

$$\begin{array}{rcl} \|v\| & = & \|v_1 \mathrm{e}_1 + \ldots + v_n \mathrm{e}_n\| \\ & \leq & |v_1| \cdot \|\mathrm{e}_1\| + \ldots + |v_n| \cdot \|\mathrm{e}_n\| \\ & \leq & |v| \cdot C \end{array}$$

mit $C = \|\mathrm{e}_1\| + \ldots + \|\mathrm{e}_n\|$. Insbesondere folgern wir, daß $\| \ \| : \mathbb{R}^n \to \mathbb{R}$ stetig ist für die durch die Maximumsnorm $| \ |$ gegebene Metrik auf $\mathbb{R}^n$, aus $d(x, y) = |x - y| < \varepsilon / C$ folgt nämlich $|\|x\| - \|y\|| \leq \|x - y\| < \varepsilon$. Nun ist aber die Oberfläche des Hyperkubus

$$K = \{v \in \mathbb{R}^n \mid |v| = 1\}$$

$| \ |$-kompakt nach 9.7.8 und nicht leer falls gilt $n \geq 1$. Nach 9.7.11 nimmt folglich die Funktion $\| \ \|$ auf K ein Minimum a an, und da K nicht den Nullvektor enthält, ist dies Minimum notwendig positiv, $a > 0$. Wir folgern zunächst einmal $a|v| \leq \|v\|$ für alle $v \in K$, dann gilt aber natürlich auch $a|\lambda v| \leq \|\lambda v\|$ für alle $\lambda \in \mathbb{R}$ und $v \in K$, also $a|w| \leq \|w\| \quad \forall w \in \mathbb{R}^n$. Mit $c = 1/a$ gilt also $|w| \leq c\|w\| \quad \forall w \in \mathbb{R}^n$. □

Variante zum Schluß des vorhergehenden Beweises. Statt hier mit Kompaktheit zu argumentieren, kann man alternativ auch mit Induktion über n und "Vollständigkeit" argumentieren, was uns insbesondere beim Beweis der Verallgemeinerung **??** helfen wird. Die Argumentation verläuft dann wie folgt: Wir betrachten die affinen Hyperebenen $H_i = \{x \mid x_i = 1\}$. Aus der Induktionsannahme können wir durch Widerspruch folgern, daß es positive Konstanten $a_i > 0$ gibt mit

$$a_i \leq \|w\| \quad \forall w \in H_i$$

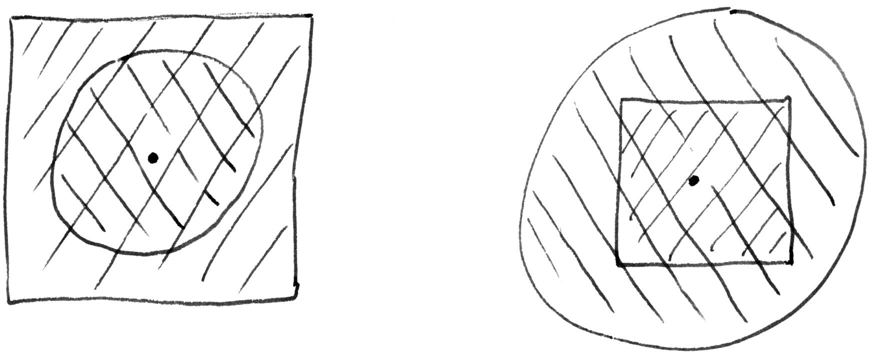


Illustration zur Äquivalenz von Normen am Beispiel der Betragsnorm und der euklidischen Norm auf dem $\mathbb{R}^2$.

In der Tat gäbe es sonst in H_i eine Folge w_ν mit $\|w_\nu\| \to 0$ für $\nu \to \infty$. Diese Folge wäre im Sinne von 10.5.1 eine Cauchy-Folge für die von $\| \; \|$ auf H_i induzierte Metrik. Dann wäre sie aber wegen der Äquivalenz der Normen nach der Induktionsannahme auch eine Cauchy-Folge für die von der Maximumsnorm $| \; |$ auf H_i induzierte Metrik und müßte nach 10.5.2 konvergieren gegen einen Punkt $w \in H_i$ mit $\|w\| = 0$. Widerspruch! Nun gibt es für $v \in \mathbb{R}^n \backslash 0$ stets $\lambda \in \mathbb{R}$ mit $|\lambda| = |v|$ derart, daß $\lambda^{-1}v$ in einer der affinen Hyperebenen H_i liegt. Mit $a = \inf(a_i)$ folgt $a \leq \|\lambda^{-1}v\|$ und $|v| \leq c\|v\|$ für $c = 1/a$. □

9.9.22. Wir nennen eine Teilmenge eines endlichdimensionalen reellen Raums **offen** bzw. **abgeschlossen** genau dann, wenn sie offen ist für die von irgendeiner Norm auf seinem Richtungsraum induzierte Metrik. Nach unserem Satz 9.9.21 über die Äquivalenz von Normen ist sie dann notwendig offen für jede von einer Norm induzierte Metrik. Die so erklärten offenen Teilmengen bilden die sogenannte **natürliche Topologie** auf unserem endlichdimensionalen reellen Raum.

Übung 9.9.23. Jede lineare Abbildung von einem endlichdimensionalen Vektorraum in einen normierten Vektorraum W ist stetig. Sind allgemeiner endlichdimensionale Vektorräume $V_1, \ldots, V_n$ gegeben, so ist jede multilineare Abbildung $V_1 \times \ldots \times V_n \to W$ stetig.

Definition 9.9.24. Ist $f : V \to W$ eine stetige lineare Abbildung normierter Vektorräume, so heißt die kleinstmögliche Konstante $C \geq 0$ wie in 9.9.13 auch die **Operatornorm** $\|f\|$ von f, in Formeln

$$\|f\| = \sup\{\|f(v)\| \mid \|v\| \leq 1\}$$

Übung 9.9.25. Sind $f : V \to W$ und $g : W \to X$ stetige Abbildungen zwischen normierten Vektorräumen, so gilt $\|g \circ f\| \leq \|g\|\|f\|$.

9.9.26. Die stetigen linearen Abbildungen zwischen normierten Vektorräumen V, W nennt man auch **beschränkte Operatoren**, da sie nach 9.9.13 genau die linearen Abbildungen sind, die den Einheitsball auf eine beschränkte Menge abbilden. Ich notiere die Menge aller solchen Abbildungen $\mathcal{B}(V, W)$ oder auch $\mathcal{B}_\mathbb{R}(V, W)$, wenn ich besonders betonen will, daß reell-lineare Abbildungen gemeint sind und nicht etwa "komplex-lineare" Abbildungen, wie wir sie später für gewöhnlich betrachten werden. Ich werde die Notation $\mathcal{B}$ benutzen, die Terminologie jedoch vermeiden und nach Möglichkeit von **stetigen Operatoren** reden, da diese ja keineswegs beschränkte Abbildungen im Sinne von 9.3.4 zu sein brauchen.

Übung 9.9.27. Der Raum $\mathcal{B}(V, W)$ aller stetigen linearen Abbildungen zwischen normierten Vektorräumen V, W ist ein Untervektorraum im Raum $\mathrm{Hom}(V, W)$ aller linearen Abbildungen von V nach W, und die in 9.9.24 eingeführte Abbildung $f \mapsto \|f\|$ ist eine Norm auf $\mathcal{B}(V, W)$.

Ergänzende Übung 9.9.28. Sind normierte Vektorräume $V_1, \ldots, V_n$ und W gegeben und ist $f : V_1 \times \ldots \times V_n \to W$ eine stetige multilineare Abbildung, so heißt die kleinstmögliche Konstante $C \geq 0$ wie in 9.9.18 die **Norm** von f und wird notiert

$$\|f\| = \sup\{\|f(v_1, \ldots, v_n)\| \mid \|v_i\| \leq 1\}$$

Man zeige, daß wir so eine Norm auf dem Vektorraum $\mathcal{B}(V_1, \ldots, V_n; W)$ aller stetigen multilinearen Abbildungen erhalten. Weiter zeige man: Die offensichtliche Abbildung liefert einen Isomorphismus von normierten Räumen

$$\mathcal{B}(V_1, \mathcal{B}(V_2, \ldots, V_n; W)) \xrightarrow{\sim} \mathcal{B}(V_1, \ldots, V_n; W)$$

9.10 Überdeckungen kompakter metrischer Räume

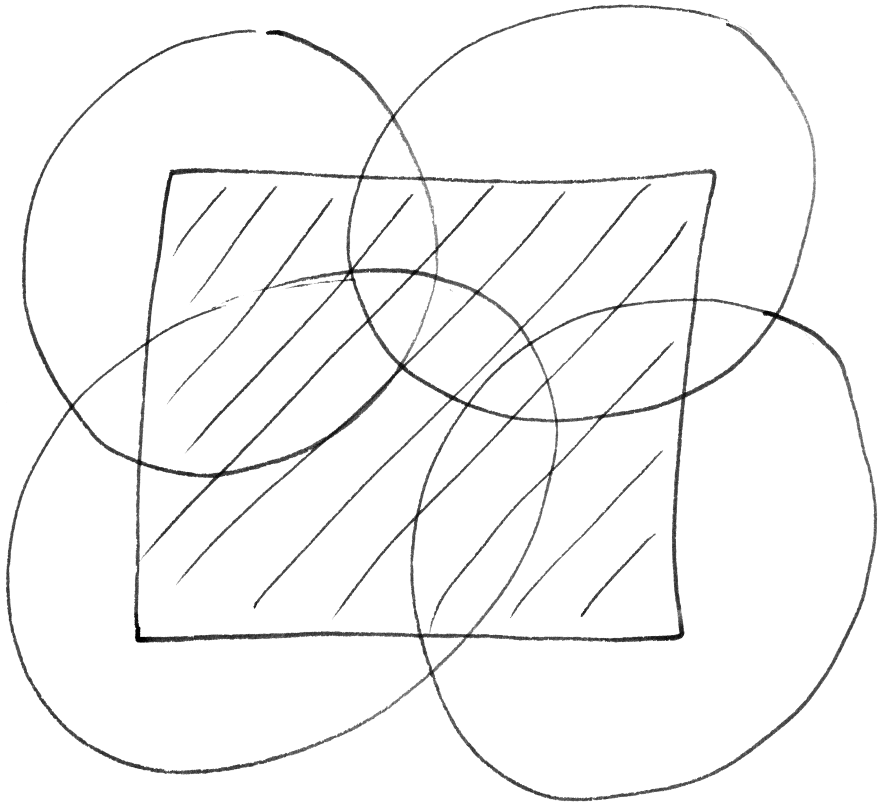
Definition 9.10.1. Sei X eine Menge. Unter einer **Überdeckung** von X versteht man ein System $\mathcal{U} \subset \mathcal{P}(X)$ von Teilmengen von X mit Vereinigung X, in Formeln ausgedrückt $X = \bigcup_{U \in \mathcal{U}} U$. Unter einer **Teilüberdeckung** einer Überdeckung $\mathcal{U}$ versteht man ein Teilsystem $\mathcal{V} \subset \mathcal{U}$, das auch selbst schon eine Überdeckung ist.

Definition 9.10.2. Unter einer **offenen Überdeckung** eines metrischen Raums oder allgemeiner eines topologischen Raums versteht man eine Überdeckung, die aus offenen Teilmengen besteht.

Satz 9.10.3 (Kompaktheit und offene Mengen). *Ein metrischer Raum ist folgenkompakt genau dann, wenn jede offene Überdeckung unseres Raums eine endliche Teilüberdeckung besitzt.*

9.10.4. Ich hoffe, daß Sie im weiteren Verlauf dieser Vorlesung noch sehen werden, wie wichtig diese Charakterisierung der Kompaktheit ist. Im Kontext topologischer Räume wird Satz 9.10.3 die Definition der Kompaktheit **??**: Ein topologischer Raum heißt **kompakt** oder manchmal auch ausführlicher **überdeckungskompakt** genau dann, wenn jede offene Überdeckung unseres Raums eine endliche Teilüberdeckung besitzt. In der französischen Literatur bezeichnet man diese Eigenschaft eines topologischen Raums meist abweichend als **quasikompakt** und fordert von einem kompakten Raum zusätzlich die Hausdorff-Eigenschaft. Topologische Räume mit der Eigenschaft, daß jede Folge eine konvergente Teilfolge besitzt, heißen dahingegen **folgenkompakt**.

Ergänzung 9.10.5. Ein Beispiel für einen überdeckungskompakten aber nicht folgenkompakten topologischen Raum finden Sie in **??**, ein Beispiel für einen folgenkompakten aber nicht überdeckungskompakten topologischen Raum in **??**. Besitzt ein überdeckungskompakter topologischer Raum die zusätzliche Eigenschaft, daß



Eine Überdeckung eines Quadrats durch vier Kreisscheiben man für jeden seiner Punkte eine Folge von Umgebungen derart finden kann, daß jede seiner Umgebungen mindestens eine Umgebung dieser Folge umfaßt, so ist er auch folgenkompakt mit demselben Argument, wie wir es im Beweis des Satzes verwenden.

9.10.6. Sei X eine Menge. Unter einer **Überdeckung** einer Teilmenge $Y \subset X$ durch Teilmengen von X versteht man ein Mengensystem $\mathcal{U} \subset \mathcal{P}(X)$ mit $Y \subset \bigcup_{U \in \mathcal{U}} U$. Man folgert leicht aus 9.10.3, daß eine Teilmenge Y eines metrischen Raums X kompakt ist genau dann, wenn jede Überdeckung von Y durch offene Teilmengen von X eine endliche Teilüberdeckung besitzt.

Beweis. Sei X ein metrischer Raum. Ist X nicht kompakt, so finden wir in X eine Folge ohne konvergente Teilfolge. Dann besitzt jeder Punkt von X eine offene Umgebung, die nur endlich viele Folgenglieder enthält, und alle diese offenen Umgebungen bilden eine offene Überdeckung von X ohne endliche Teilüberdeckung. Das zeigt die eine Richtung. Den Beweis der anderen Richtung beginnen wir mit einem Lemma, das auch für sich genommen oft hilfreich ist.

Lemma 9.10.7 (Überdeckungssatz von Lebesgue). *Ist X ein kompakter metrischer Raum und $\mathcal{U}$ eine offene Überdeckung von X, so gibt es ein $\varepsilon > 0$ derart, daß für alle $x \in X$ der ε-Ball $\mathrm{B}(x;\varepsilon)$ um x ganz in einer der überdeckenden offenen Mengen $U \in \mathcal{U}$ enthalten ist.*

Erster Beweis. Gäbe es kein solches $\varepsilon > 0$, so könnten wir für jedes $n \in \mathbb{N}_{\geq 1}$ einen Punkt $x_n \in X$ finden derart, daß $\mathrm{B}(x_n; 1/n)$ in keinem $U \in \mathcal{U}$ enthalten wäre. Durch Übergang zu einer Teilfolge könnten wir ohne Beschränkung der Allgemeinheit zusätzlich annehmen, daß die Folge der x_n konvergiert, etwa gegen $x \in X$. Nun finden wir jedoch ein $U \in \mathcal{U}$ mit $x \in U$ und dazu $\rho > 0$ mit $\mathrm{B}(x;\rho) \subset U$ und dazu N mit $d(x_N, x) < \rho/2$ und $1/N < \rho/2$, und dann gälte $\mathrm{B}(x_N; 1/N) \subset \mathrm{B}(x_N; \rho/2) \subset \mathrm{B}(x;\rho) \subset U$ im Widerspruch zur Wahl der x_n. □

Zweiter Beweis. Man betrachte die Funktion $f : X \to \mathbb{R}$ gegeben durch die Vorschrift

$$f(x) = \sup\{r \leq 1 \mid \text{Es gibt } U \in \mathcal{U} \text{ mit } \mathrm{B}(x;r) \subset U\}$$

Die Dreiecksungleichung liefert $|f(x) - f(y)| \leq d(x,y)$, insbesondere ist f stetig. Ohne Beschränkung der Allgemeinheit dürfen wir $X \neq \emptyset$ annehmen. Dann nimmt f nach 9.7.11 sein Minimum an, und dies Minimum ist ein mögliches $\varepsilon > 0$. □

Um die andere Implikation im Satz zu zeigen sei nun X kompakt und $\mathcal{U}$ eine offene Überdeckung von X. Es gilt zu zeigen, daß sie eine endliche Teilüberdeckung besitzt. Wählen wir zu unserer Überdeckung $\mathcal{U}$ ein ε wie im Überdeckungssatz

9.10.7, so reicht es auch zu zeigen, daß es eine endliche Teilmenge $E \subset X$ gibt mit

$$X = \bigcup_{x \in E} \mathrm{B}(x; \varepsilon)$$

In der Tat liegt ja der ε-Ball $\mathrm{B}(x;\varepsilon)$ um ein beliebiges $x \in X$ nach Wahl von ε schon in einem der $U \in \mathcal{U}$. Gäbe es aber für ein $\varepsilon > 0$ keine endliche Überdeckung von X durch ε-Bälle, so könnten wir induktiv eine Folge $(x_n)_{n\in\mathbb{N}}$ konstruieren mit $x_n \notin \bigcup_{0\le\nu<n} \mathrm{B}(x_\nu;\varepsilon)$ für alle n, also $d(x_n, x_m) \geq \varepsilon$ für $n \neq m$, und diese Folge könnte keine konvergente Teilfolge haben, im Widerspruch zur Annahme. □

Übung 9.10.8. Ist in einem kompakten topologischen Raum X ein System abgeschlossener Teilmengen $\mathcal{K} \subset \mathcal{P}(X)$ mit leerem Schnitt $\bigcap_{K\in\mathcal{K}} K = \emptyset$ gegeben, so gibt es bereits ein endliches Teilsystem $\mathcal{E} \subset \mathcal{K}$ mit $\bigcap_{K\in\mathcal{E}} K = \emptyset$.

Ergänzende Übung 9.10.9 (**Satz von Dini**). Eine monoton wachsende Folge stetiger reellwertiger Funktionen auf einem kompakten Raum, die punktweise gegen eine stetige Funktion konvergiert, konvergiert sogar gleichmäßig. Hinweis: 9.10.8.

Ergänzende Übung 9.10.10. Man zeige, daß das Bild eines kompakten topologischen Raums unter einer stetigen Abbildung kompakt ist für die Spurtopologie. Insbesondere ist jede stetige reellwertige Funktion auf einem kompakten topologischen Raum beschränkt.

9.11 Integrale mit Parametern

Satz 9.11.1 (über Integrale mit Parametern). *Gegeben ein metrischer Raum X und eine stetige Funktion $f : X \times [a, b] \to \mathbb{R}$ ist auch die Funktion $X \to \mathbb{R}$, $x \mapsto \int_a^b f(x,t)\,\mathrm{d}t$ stetig.*

9.11.2. Ich zeige diesen Satz in großer Allgemeinheit als Anwendung unserer neuen Charakterisierung der Kompaktheit und als Illustration für die Kraft der allgemeinen Theorie metrischer Räume. Ist X offen oder abgeschlossen in einem $\mathbb{R}^n$, so kann man auch elementarer mit der gleichmäßigen Stetigkeit argumentieren. Diesen Beweis gebe ich als Alternative auch noch an.

Beweis. Versehen wir den Raum $\mathcal{C}([a, b], \mathbb{R})$ aller stetigen reellwertigen Funktionen auf $[a, b]$ mit der Supremumsnorm, so ist nach dem gleich folgenden Satz 9.11.4 die von f induzierte Abbildung $\tilde{f} : X \to \mathcal{C}([a, b], \mathbb{R})$, $x \mapsto f(x, \ )$ stetig. Nach Übung 9.9.16 ist weiter das Integrieren $\int : \mathcal{C}([a, b], \mathbb{R}) \to \mathbb{R}$ stetig. Damit ist unsere Abbildung $\int \circ \tilde{f} : X \to \mathbb{R}$ stetig als eine Verknüpfung stetiger Abbildungen. □

Alternativer Beweis. Ist X offen oder abgeschlossen in einem $\mathbb{R}^n$, so kann man auch elementarer argumentieren. Zunächst reicht es ja, die Stetigkeit an jeder Stelle $x \in X$ nachzuweisen. Mit dieser Überlegung können wir uns leicht auf den Fall zurückziehen, daß X kompakt ist. Dann ist aber auch $X \times [a, b]$ kompakt und nach 9.7.14 ist f dort gleichmäßig stetig. Für alle $\varepsilon > 0$ gibt es insbesondere $\delta > 0$ mit

$$|x - y| < \delta \;\Rightarrow\; |f(x, t) - f(y, t)| < \varepsilon \text{ für alle } t \in [a, b].$$

Aus $|x - y| < \delta$ folgt mithin

$$\left| \int_a^b f(x, t)\,\mathrm{d}t - \int_a^b f(y, t)\,\mathrm{d}t \right| \leq \int_a^b |f(x, t) - f(y, t)|\,\mathrm{d}t \leq (b - a)\varepsilon$$

und das zeigt die Behauptung. □

9.11.3. Den Raum aller stetigen Abbildungen von einem kompakten Raum X in einen metrischen Raum Y, versehen mit der Metrik der gleichmäßigen Konvergenz, wird $\mathcal{C}(X, Y)$ notiert. Das $\mathcal{C}$ steht hier für englisch “continous” und französisch “continu”.

Satz 9.11.4 (Stetige Abbildungen in Abbildungsräume). *Seien X, Y und K metrische Räume. Ist K kompakt, so ist eine Abbildung $f : X \times K \to Y$ stetig genau dann, wenn die induzierte Abbildung $\tilde{f} : X \to \mathcal{C}(K, Y)$ stetig ist für die Metrik der gleichmäßigen Konvergenz auf $\mathcal{C}(K, Y)$.*

Beweis. Daß aus der Stetigkeit von $\tilde{f}$ die Stetigkeit von f folgt, sieht man ohne weitere Schwierigkeiten. Wir zeigen nun die andere Richtung und müssen die Stetigkeit von $\tilde{f}$ an jeder Stelle $p \in X$ nachweisen. Sei diese Stelle p ab jetzt fest gewählt und sei $\varepsilon > 0$ gegeben. Aufgrund der Stetigkeit von f gibt es für jedes $s \in K$ ein $\delta_s > 0$ mit

$$f\,\mathrm{B}((p, s); \delta_s) \subset \mathrm{B}(f(p, s); \varepsilon)$$

Nun gilt für unsere Metrik auf $X \times K$ ja $\mathrm{B}((p, s); \delta) = \mathrm{B}(p; \delta) \times \mathrm{B}(s; \delta)$ und nach 9.10.3 gibt es eine endliche Teilmenge $E \subset K$ mit $K \subset \bigcup_{s \in E} \mathrm{B}(s; \delta_s)$. Für $\eta = \min_{s \in E} \delta_s$ behaupten wir dann

$$x \in \mathrm{B}(p; \eta) \;\Rightarrow\; d(f(x, t), f(p, t)) < 2\varepsilon \;\; \forall t \in K$$

In der Tat finden wir für jedes $t \in K$ ein $s \in E$ mit $t \in \mathrm{B}(s; \delta_s)$ und für dies s liegen (p, t) und (x, t) beide in $\mathrm{B}((p, s); \delta_s)$. Damit ist die Stetigkeit von $\tilde{f}$ bei p gezeigt. □

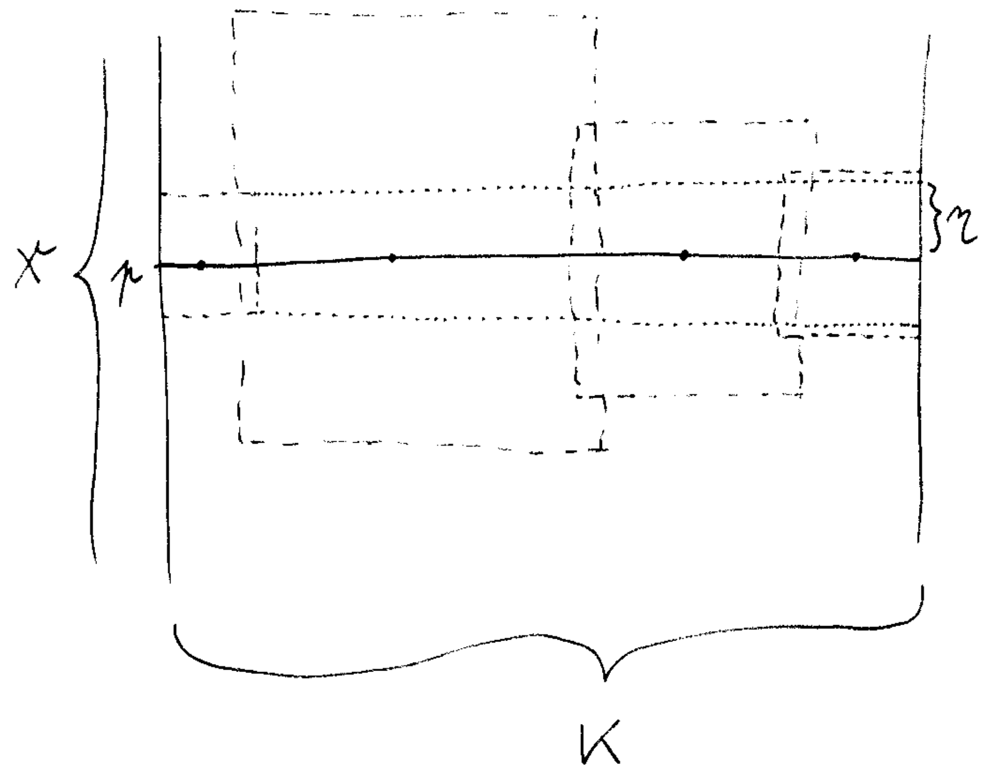


Illustration zum Beweis von Satz 9.11.4. Die mit gestrichelten Rändern eingezeichneten Quadrate sind so gewählt, daß unsere Abbildung f auf jedem Quadrat höchstens um den Abstand ε von ihrem Wert im Zentrum des jeweiligen Quadrats abweicht. Die gepunktelten Linien begrenzen einen Streifen der Breite 2η, in dem unsere Funktion auf jeder Vertikalen höchstens um 2ε von ihrem Wert am Schnittpunkt der besagten Vertikalen mit der fett eingezeichneten Horizontalen abweicht.

10 Raumwertige Funktionen

10.1 Bogenlänge in metrischen Räumen

Definition 10.1.1. Gegeben ein nichtleeres Intervall $I \subset \mathbb{R}$, ein metrischer Raum (X, d) und eine Abbildung $\gamma : I \to X$ definieren wir die **Länge** $L(\gamma) \in \overline{\mathbb{R}}$ von γ als das Supremum über "die Längen aller einbeschriebenen Polygonzüge", in Formeln

$$L(\gamma) = \sup \left\{ \sum_{i=1}^{n} d(\gamma(t_i), \gamma(t_{i-1})) \middle| t_0, \ldots, t_n \in I,\ t_0 \leq \ldots \leq t_n \right\}$$

Man spricht in diesem Zusammenhang meist von **Bogenlänge**.

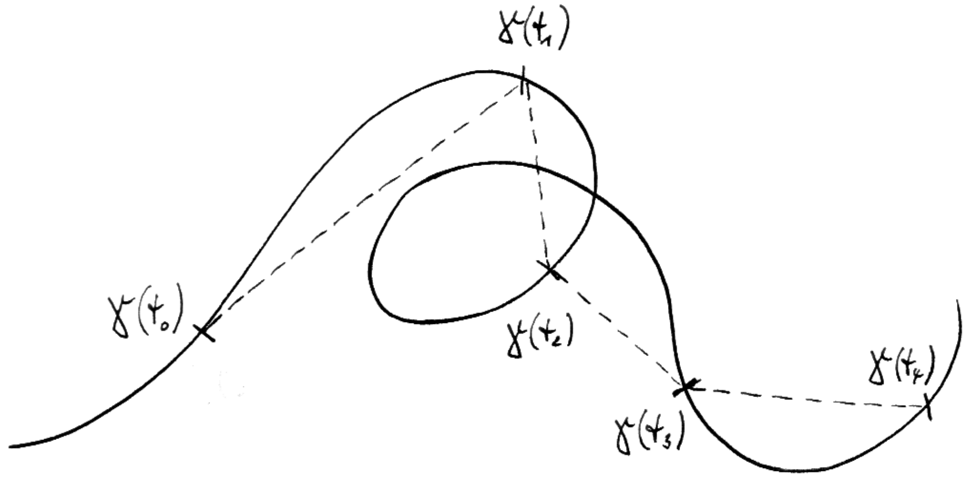
10.1.2. Diese Definition liefert uns sogar einen Längenbegriff für eine Abbildung von einer beliebigen angeordneten Menge in einen metrischen Raum. Eine stetige Abbildung von einem nichtleeren kompakten halboffenen reellen Intervall in einen metrischen oder allgemeiner topologischen Raum nennen wir einen **Weg** in unserem Raum. Wir interessieren uns besonders für die Länge von Wegen im $\mathbb{R}^k$ und verstehen in diesem Zusammenhang die Länge stets in Bezug auf die euklidische Metrik.

10.1.3. Offensichtlich ist unsere Bogenlänge "invariant unter Reparametrisierung", genauer haben wir für jede monotone Surjektion $\psi : J \to I$ notwendig $L(\gamma \circ \psi) = L(\gamma)$. Unsere Definition der Kreiszahl π aus 5.4.1 können wir schreiben als $\pi = L(\gamma)$ für $\gamma : [-1, 1] \to \mathbb{R}^2$, $x \mapsto (x, \sqrt{1 - x^2})$.

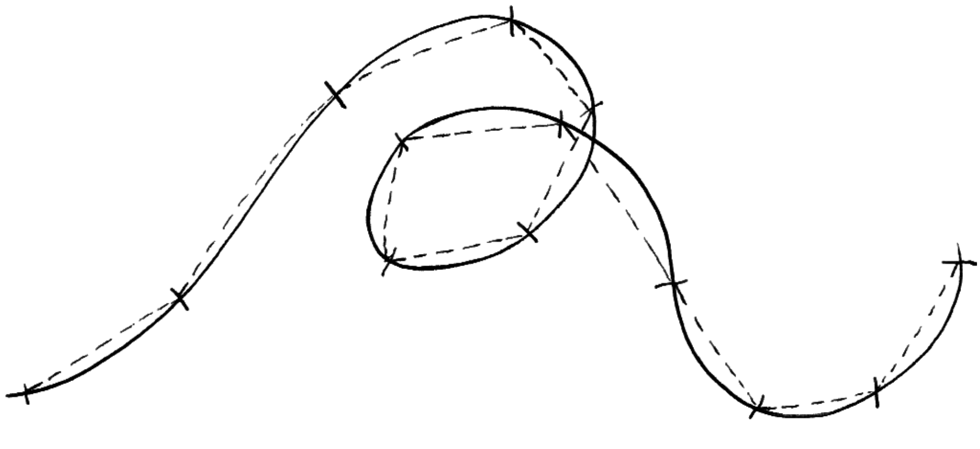
Übung 10.1.4. Gegeben $s \in \mathbb{R}$ bezeichne $(s\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ die Multiplikation mit s. Sei $\gamma : I \to \mathbb{R}^n$ eine Abbildung von einem Intervall nach $\mathbb{R}^n$. Man zeige $L((s\cdot) \circ \gamma) = |s| L(\gamma)$ für $s \neq 0$. Ebenso zeige man $L(A \circ \gamma) = L(\gamma)$ für jede orthogonale, d.h. abstandserhaltende Abbildung $A : \mathbb{R}^n \to \mathbb{R}^n$.

Übung 10.1.5. Ist $\gamma : [a, b] \to \mathbb{R}^n$ eine Abbildung, so gilt $L(\gamma) \geq \|\gamma(a) - \gamma(b)\|$ und Gleichheit haben wir genau dann, wenn γ aus dem Weg $\phi : [0, 1] \to \mathbb{R}^n$, $t \to t\gamma(a) + (1-t)\gamma(b)$ "entsteht durch monotone Umparametrisierung", genauer: Wenn es $\psi : [a, b] \to [0, 1]$ monoton gibt mit $0, 1 \in \psi([a, b])$ und $\gamma = \phi \circ \psi$.

10.1.6. Um Bogenlängen zu berechnen benutzt man meist die Darstellung als Integral 10.3.2. Sie verwendet den Begriff der Ableitung 10.2.1 von Funktionen einer reellen Veränderlichen mit Werten in affinen Räumen, mit dem wir uns nun beschäftigen werden.



Eine Approximation eines Weges durch einen Polygonzug



Eine bessere Approximation durch einen Polygonzug

10.2 Ableiten von raumwertigen Funktionen

Definition 10.2.1. Seien $\gamma : I \to X$ eine Abbildung von einer halboffenen Teilmenge $I \subset \mathbb{R}$ in einen normierten Raum X und $p \in I$ ein Punkt von I. Wir nennen γ **differenzierbar bei** p genau dann, wenn der Grenzwert $\lim_{t\to 0}(\gamma(p+t)-\gamma(p))/t$ in $\vec{X}$ existiert im Sinne von 9.6.8. In diesem Fall nennen wir besagten Grenzwert die **Ableitung von** γ **bei** p und notieren diesen Vektor

$$\gamma'(p) := \lim_{q\to p} \frac{\gamma(q)-\gamma(p)}{q-p} = \lim_{t\to 0} \frac{\gamma(p+t)-\gamma(p)}{t}$$

Ist γ differenzierbar an allen Stellen $p \in I$, so nennen wir γ **differenzierbar** oder genauer **differenzierbar auf** I.

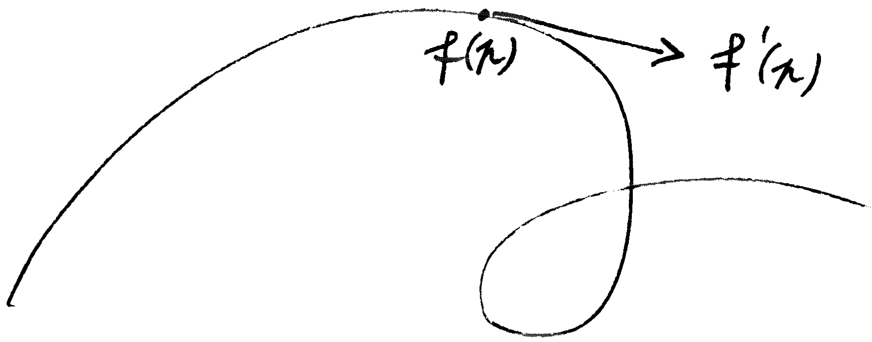
10.2.2. Ich denke mir eine Abbildung von einer halboffenen Teilmenge $I \subset \mathbb{R}$ in einen normierten Raum X auch gerne als Beschreibung eines Teilchens, das sich in X bewegt, und denke mir also I als ein Zeitintervall. Dann nenne ich $\gamma'(p)$ auch die “Geschwindigkeit” oder genauer den “Geschwindigkeitsvektor” von γ zum Zeitpunkt p und schreibe sogar manchmal $\dot{\gamma}$ statt γ'. In physikalischen Zusammenhängen verwende ich diese Begriffe jedoch präziser nur für Funktionen auf einer halboffenen Teilmenge $I \subset \mathbb{T}$ unseres mathematischen Modells der Zeit aus **??**, vergleiche **??**, denn eine physikalische Geschwindigkeit darf ja nicht die Einheit einer Länge haben. Wie man für Abbildungen von beliebigen eindimensionalen reellen Räumen in normierte reelle Räume die Ableitung definiert, besprechen wir in 4.2.2 und 4.2.14.

10.2.3. Formal folgt die in der Definition implizit behauptete Gleichheit der beiden Grenzwerte aus dem Analogon der zweiten Aussage von 6.3.21, die sich wie in 9.6.11 kurz erwähnt mitsamt ihrem Beweis ohne weitere Schwierigkeiten auf den Fall von Grenzwerten bei metrischen oder sogar topologischen Räumen verallgemeinern läßt.

10.2.4. Ich lege hier die Begrifflichkeit normierter affiner Räume im Sinne von 9.8.2 zugrunde. Der Leser mag sich stattdessen auch normierte Vektorräume oder sogar den $\mathbb{R}^n$ denken. Die gewählte Allgemeinheit modelliert jedoch meines Erachtens besser unsere Anschauung bewegter Teilchen, etwa im uns umgebenden Raum oder auch auf der Tafelebene. Des weiteren hoffe ich, daß die begriffliche Trennung von Punkten einerseits und Richtungsvektoren andererseits auch das Verständnis fördern mag.

Übung 10.2.5. Auch für Abbildungen halboffener Teilmengen von $\mathbb{R}$ in normierte Räume folgt aus der Differenzierbarkeit bereits die Stetigkeit.

Übung 10.2.6. Sei $\gamma : I \to X$ eine Abbildung von einer halboffenen Teilmenge $I \subset \mathbb{R}$ in einen normierten Raum X und sei $L : X \to Y$ eine stetige affine



Der Geschwindigkeitsvektor ist stets tangential an die Bahnkurve. Seine Länge hängt jedoch von der Anzeige des Tachometers ab, wenn wir uns hier mal ein Auto denken, das auf einem Fußballfeld herumkurvt. Dieser Aspekt ist in einem Bild leider schwer darzustellen.

Abbildung in einen weiteren normierten Raum Y. Ist γ differenzierbar an einer Stelle $p \in I$, so ist auch $L \circ \gamma$ differenzierbar bei p und es gilt

$$(L \circ \gamma)'(p) = \vec{L}(\gamma'(p))$$

Das zeigt insbesondere, daß unsere Ableitung sich nicht ändert, wenn wir zu einer anderen aber äquivalenten Norm auf X übergehen. Später wird sich diese Aussage als Spezialfall der Kettenregel in mehreren Veränderlichen 4.3.1 erweisen.

10.2.7. Im Spezialfall $X = \mathbb{R}$ sind unsere Definitionen identisch zu unseren bisherigen Definitionen für reellwertige Funktionen. Was im Fall $X = \mathbb{R}^m$ passiert, zeigt das folgende Lemma.

Lemma 10.2.8 (Komponentenregel). *Sei $X = X_1 \times \ldots \times X_m$ ein Produkt normierter Räume, $I \subset \mathbb{R}$ eine halboffene Teilmenge, $\gamma = (\gamma_1, \ldots, \gamma_m) : I \to X$ eine Abbildung und $p \in I$ ein Punkt. Genau dann ist γ differenzierbar bei p, wenn alle γ_j differenzierbar sind bei p, und dann gilt*

$$\gamma'(p) = (\gamma_1'(p), \ldots, \gamma_m'(p))$$

Beweis. Das folgt aus 9.6.12 und sei dem Leser überlassen. Man beachte, daß wir bereits bei der Formulierung die kanonische Identifikation zwischen dem Richtungsraum eines Produkts und dem Produkt der Richtungsräume der Faktoren **??** verwendet haben. □

10.2.9. Wie in **??** heißt eine Teilmenge eines reellen Raums **konvex** genau dann, wenn sie mit je zwei Punkten auch das ganze die beiden Punkte verbindende Geradensegment enthält.

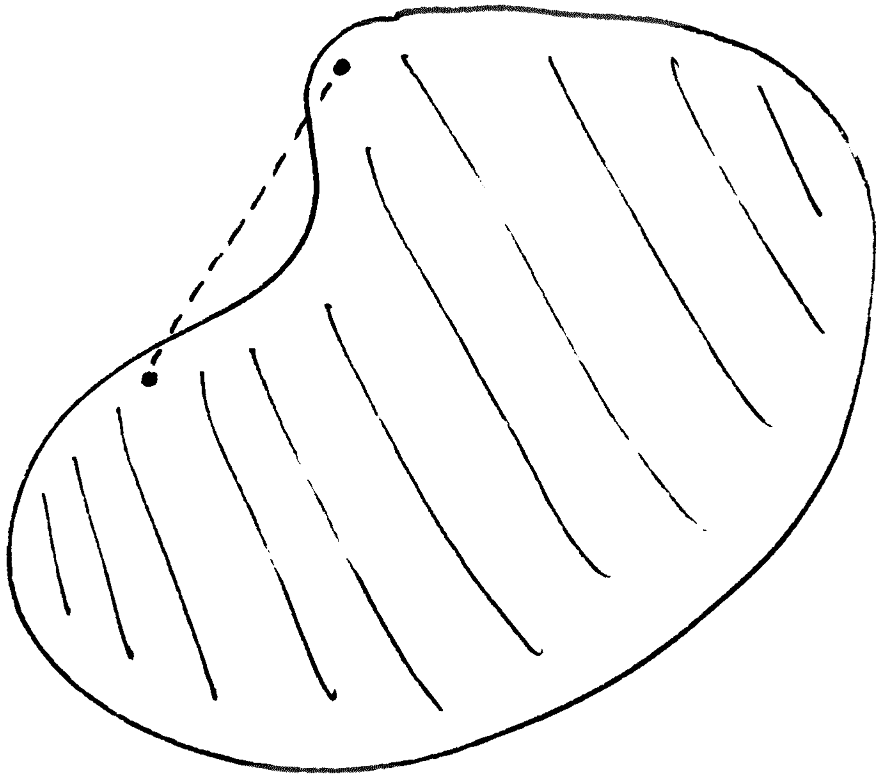
Übung 10.2.10. In einem normierten Raum ist jeder Ball konvex.

Satz 10.2.11 (Schrankensatz). *Seien X ein normierter Raum, $a < b$ reelle Zahlen, und $\gamma : [a, b] \to X$ eine differenzierbare Abbildung. Ist $C \subset \vec{X}$ eine offene oder abgeschlossene konvexe Teilmenge und gilt $\gamma'(t) \in C$ für alle $t \in [a, b]$, so folgt*

$$\gamma(b) - \gamma(a) \in (b - a)C$$

10.2.12. Man folgert leicht eine Variante, die auch $a \geq b$ erlaubt: Ist $I \subset \mathbb{R}$ ein halboffenes Intervall, $\gamma : I \to X$ eine differenzierbare Abbildung und gilt $\gamma'(t) \in C$ für alle $t \in I$, so folgt $\gamma(b) - \gamma(a) \in (b - a)C \quad \forall a, b \in I$.

Beispiel 10.2.13. Für dieses Beispiel identifizieren wir implizit die Zeitachse $\mathbb{T}$ mit der reellen Zahlengerade $\mathbb{R}$ derart, daß jeder Stunde ein Intervall der Länge Eins entspricht und damit der Zeitspanne $\mathrm{h} \in \vec{\mathbb{T}}$ der Richtungsvektor $1 \in \vec{\mathbb{R}} = \mathbb{R}$.



Eine nicht konvexe Teilmenge der Ebene

Das bedeutet insbesondere, daß wir implizit auch vektorielle Geschwindigkeiten mit Richtungsvektoren identifizieren dürfen. Ist dann C eine Kreisscheibe im Richtungsraum der Anschauungsebene mit Radius 20km/h, so können wir den Inhalt des Satzes interpretieren wie folgt: Fahren wir mit einem Geländewagen um 14:00 an einem Parkplatz los und kurven durch die Gegend und der Tacho zeigt nie mehr als 20km/h an, so sind wir um 17:00 höchstens 60km von unserem ursprünglichen Parkplatz entfernt. Besteht C dahingegen aus einem einzigen Punkt, der sagen wir die Geschwindigkeit von 20km/h in einer festen Richtung bedeutet, so besagt unser Satz: Fahren wir konstant mit 20km/h in diese Richtung, so haben wir um 17:00 genau 60km in besagte Richtung zurückgelegt. Im übrigen wird in 4.2.14 erklärt, wie man auch mit "echten" Geschwindigkeiten formal korrekt arbeiten kann.

10.2.14. Der Satz folgt im Fall $X = \mathbb{R}$ leicht aus unserem bisherigen Mittelwertsatz 7.3.8 und er spielt auch im allgemeinen eine ähnliche Rolle, indem er es erlaubt, "den von einem Teilchen in einem Zeitintervall $[a, b]$ gewonnenen Abstand von seinem Ausgangspunkt aus der Kenntnis seiner lokalen Geschwindigkeiten abzuschätzen". Jedoch kann man für höherdimensionales X im allgemeinen keinen Zeitpunkt mehr finden, zu dem das Teilchen "mittlere Geschwindigkeit" hätte, d.h. es gibt für höherdimensionales X im allgemeinen keinen Zeitpunkt $\xi \in [a, b]$ mit $\gamma(b) - \gamma(a) = (b - a)\gamma'(\xi)$. Man stelle ich etwa vor, daß unser Geländewagen ein Rundtour fährt, bei der er zu keiner Zeit die Geschwindigkeit Null hat. Ich bin deshalb von der in der älteren Literatur üblichen Bezeichnung als "Mittelwertsatz in mehreren Veränderlichen" nicht vollständig befriedigt. Oft wird auch nur der Fall betrachtet, daß C ein offener Ball oder auch ein abgeschlossener Ball mit Zentrum im Ursprung ist: Aus $\|\gamma'(t)\| \leq K \ \forall t \in [a, b]$ folgt so etwa $\|\gamma(b) - \gamma(a)\| \leq (b - a)K$.

10.2.15. Offensichtlich ist eine Teilmenge C eines reellen Vektorraums genau dann konvex, wenn für beliebige reelle $s, t \geq 0$ gilt $sC + tC = (s + t)C$. Das zeigt, daß in unserem Satz die Aussage für das ganze Intervall folgt, wenn wir sie für alle Stücke einer Zerlegung in Teilintervalle zeigen können.

Erster Beweis. Ist C abgeschlossen, so schreiben wir C als den Schnitt der offenen konvexen Mengen $C + \mathrm{B}(0; \eta)$. Wir dürfen also ohne Beschränkung der Allgemeinheit C offen annehmen. Wir betrachten nun

$$s = \sup\{q \in [a, b] \mid \gamma(x) - \gamma(a) \in (x - a)C \ \forall x \in [a, q]\}$$

und zeigen zunächst $s = b$. Für alle $p \in [a, b]$ finden wir ja eine offene Umgebung $U_p \Subset [a, b]$ mit

$$\frac{\gamma(q) - \gamma(p)}{q - p} \in C \text{ für alle } q \in U_p \backslash p$$

Insbesondere folgern wir $s > a$ und müssen nur noch die Annahme $s < b$ zum Widerspruch führen. Aber wäre $s < b$, so fänden wir $\varepsilon > 0$ mit $[s-\varepsilon, s+\varepsilon] \subset U_s$ und die Aussage des Satzes gälte für die Einschränkung von γ auf die Intervalle $[a, s-\varepsilon]$, $[s-\varepsilon, s]$ und $[s, s+\varepsilon]$. Daraus folgte jedoch mit 10.2.15 die Aussage des Satzes für das Intervall $[a, s+\varepsilon]$ im Widerspruch zur Wahl von s. Mithin haben wir $s = b$. Da es aber mit denselben Argumenten auch ein $\eta > 0$ gibt derart, daß die Aussage des Satzes für die Einschränkung von γ auf $[b-\eta, b]$ gilt, folgt die Aussage des Satzes für das ganze Intervall $[a, b]$. □

Zweiter Beweis. Wir beginnen wie beim ersten Beweis und finden Umgebungen U_p wie dort, die wir sogar als Schnitte mit $[a, b]$ von offenen Bällen $\mathrm{B}(p; \varepsilon_p)$ annehmen dürfen. Da $[a, b]$ kompakt ist, wird es nach 9.10.3 überdeckt durch endlich viele solcher Umgebungen U_p. Seien nun $a = p_0 < p_1 < p_2 < \ldots < p_r = b$ die Elemente einer kleinstmöglichen Menge von Punkten, die a und b enthält und für die die zugehörigen Umgebungen $[a, b]$ überdecken. Es ist dann leicht zu sehen, daß wir Zwischenpunkte $q_i \in (p_{i-1}, p_i)$ finden können derart, daß auf jedem Teilintervall der so entstehenden Unterteilung von $[a, b]$ in $2r$ Teilintervalle die Folgerung unseres Mittelwertsatzes gilt. Mithin gilt sie auch für das ganze Intervall $[a, b]$. □

Übung 10.2.16. Sei $I \subset \mathbb{R}$ eine halboffene Teilmenge und seien $A : I \to \mathrm{M}(n \times m; \mathbb{R})$ und $B : I \to \mathrm{M}(m \times k; \mathbb{R})$ zwei differenzierbare matrixwertige Funktionen. So ist auch das Produkt $AB : t \mapsto A(t)B(t)$ differenzierbar und die Geschwindigkeit $(AB)'$ der Produktfunktion $AB : I \to \mathrm{M}(n \times k; \mathbb{R})$ wird gegeben durch die Formel

$$(AB)' = A'B + AB'$$

Übung 10.2.17. Man formuliere und zeige die Summenregel für vektorwertige Funktionen.

Übung 10.2.18. Man zeige für komplexwertige Funktionen einer reellen Veränderlichen die Summenregel $(f+g)' = f'+g'$, die Produktregel $(fg)' = f'g+fg'$ und die Regel für die Ableitung des Kehrwerts $(1/f)' = -f'/f^2$. In 1.5.7 und 1.5.14 werden wir das sogar für komplexwertige Funktionen einer komplexen Veränderlichen zeigen.

Übung 10.2.19. Für jedes $\lambda \in \mathbb{C}$ und $m \in \mathbb{Z}$ ist die Abbildung $\mathbb{R} \to \mathbb{C}$ bzw. $\mathbb{R}\backslash\lambda \to \mathbb{C}$, $t \mapsto (t-\lambda)^m$ differenzierbar mit Ableitung $t \mapsto m(t-\lambda)^{m-1}$. Später wird das mit 1.5.10 und 1.5.12 ganz schnell gehen.

Ergänzende Übung 10.2.20. Man zeige, daß für jede stetig differenzierbare Abbildung $\gamma : I \to X$ von einer halboffenen Teilmenge $I \subset \mathbb{R}$ in einen normierten reellen Raum X die "Tangenten-Sekanten-Abbildung"

$$\begin{array}{rccl} \phi : & I^2 & \to & V \\ & (s,t) & \mapsto & \begin{cases} \frac{\gamma(s)-\gamma(t)}{s-t} & s \neq t; \\ \gamma'(s)=\gamma'(t) & s = t, \end{cases} \end{array}$$

stetig ist. Hinweis: Schrankensatz 10.2.11.

10.3 Die Bogenlänge als Integral

Definition 10.3.1. Gegeben ein differenzierbarer Weg in einem normierten Raum erklären wir seine **absolute Geschwindigkeit** zu einem gegebenen Zeitpunkt als die Norm des Geschwindigkeitsvektors.

Satz 10.3.2 (Bogenlänge als Integral). *Die Länge eines stetig differenzierbaren Weges $\gamma : [a,b] \to X$ in einem normierten reellen Raum X stimmt überein mit dem Integral über seine absolute Geschwindigkeit, in Formeln*

$$L(\gamma) = \int_a^b \|\gamma'(t)\| \, \mathrm{d}t$$

Beweis. Sei $\varepsilon > 0$ beliebig. Da γ' nach 9.7.14 gleichmäßig stetig ist auf $[a,b]$, finden wir ein $\delta > 0$ mit $\|\gamma'(x) - \gamma'(y)\| < \varepsilon$ falls $|x-y| \leq \delta$. Gegeben eine Unterteilung $a = a_0 \leq a_1 \leq \ldots \leq a_r = b$ einer Feinheit $\leq \delta$ folgern wir aus dem Schrankensatz 10.2.11 dann

$$\gamma(a_i) - \gamma(a_{i-1}) \in (a_i - a_{i-1})(\gamma'(a_i) + \mathrm{B}(0;\varepsilon))$$

und insbesondere $\|\gamma(a_i) - \gamma(a_{i-1})\| \in (a_i - a_{i-1})\|\gamma'(a_i)\| + (a_i - a_{i-1})[-\varepsilon, \varepsilon]$. Durch Aufsummieren folgt

$$\left| \sum_{i=1}^r \|\gamma(a_i) - \gamma(a_{i-1})\| \quad - \sum_{i=1}^r \|\gamma'(a_i)\|(a_i - a_{i-1}) \right| \leq (b-a)\varepsilon$$

für jede Unterteilung der Feinheit $\leq \delta$. Das zeigt schon $L(\gamma) < \infty$. Nach 6.5.11 können wir weiter δ sogar so klein wählen, daß in unserer Differenz die rechte Summe zusätzlich einen Abstand $\leq \varepsilon$ hat vom Integral $\int \|\gamma'\|$ für jede Unterteilung der Feinheit $\leq \delta$. Da aber die Länge approximierender Polygonzüge beim Hinzufügen von Zwischenpunkten nur größer werden kann, finden wir eine Unterteilung von dieser Feinheit, für die die linke Summe von $L(\gamma)$ einen Abstand $\leq \varepsilon$ hat. Zusammen erhalten wir

$$\left| L(\gamma) - \int \|\gamma'\| \right| \leq (b-a+2)\varepsilon$$

Da das für alle $\varepsilon > 0$ gilt, folgt $L(\gamma) = \int \|\gamma'\|$ wie gewünscht. □

Ergänzung 10.3.3. Ein Weg in einem metrischen Raum heißt **rektifizierbar** genau dann, wenn er endliche Länge hat. Ein stetig differenzierbarer Weg in einem normierten Raum ist also insbesondere stets rektifizierbar.

Übung 10.3.4. Eine Abbildung von einem halboffenen reellen Intervall in einen metrischen Raum heißt **nach der Bogenlänge parametrisierend** genau dann, wenn ihre Restriktion auf jedes nichtleere halboffene kompakte Teilintervall dieselbe Länge hat wie das Teilintervall selber. Man zeige, daß eine stetig differenzierbarere Abbildung in einen normierten Raum genau dann nach der Bogenlänge parametrisierend ist, wenn die zugehörige absolute Geschwindigkeit konstant Eins ist.

Übung 10.3.5. Man zeige, daß sich jede stetig differenzierbare Abbildung von einem halboffenen reellen Intervall nach $\mathbb{R}^n$ mit nirgends verschwindender Geschwindigkeit "nach der Bogenlänge parametrisieren" läßt, daß es genauer für solch eine Abbildung $\gamma : I \to X$ stets eine stetig differenzierbare Bijektion $\psi : J \xrightarrow{\sim} I$ gibt derart, daß $\gamma \circ \psi$ nach der Bogenlänge parametrisierend ist. Das gilt auch für Wege in beliebigen normierten Vektorräumen, nur benötigt man zum Argumentieren in dieser Allgemeinheit die Kettenregel 4.3.1, die uns hier noch nicht zur Verfügung steht.

Übung 10.3.6. Gegeben ein stetig differenzierbarer Weg $\gamma : [a, b] \to X$ in einem normierten Raum X und eine stetige Funktion $f : \gamma([a, b]) \to \mathbb{R}$ definiert man das **Kurvenintegral** von f längs γ als die reelle Zahl

$$\int_\gamma f = \int_a^b f(\gamma(t))\ \|\gamma'(t)\|\ \mathrm{d}t$$

Man zeige, daß das Kurvenintegral unabhängig ist von der Parametrisierung und daß es mit denselben Notationen wie oben geschrieben werden kann als der Grenzwert der Riemannsummen

$$S_\gamma^r(f) = \sum_{i=1}^{r} f(\gamma(a_i))\ \|\gamma(a_i) - \gamma(a_{i-1})\|$$

Als Kür definiere man allgemeiner das Kurvenintegral längs eines beliebigen rektifizierbaren Weges $\gamma : [a, b] \to X$ in einem metrischen Raum X.

10.3.7 (**Anschauung zum Kurvenintegral**). Die Länge eines stetig differenzierbaren Weges ist in dieser Terminologie das Kurvenintegral der konstanten Funktion Eins längs unseres Weges. Der Schwerpunkt eines durch eine Abbildung $\gamma : [a, b] \to \mathbb{R}^3$ beschriebenen homogenen gebogenen Drahtes hat als Koordinaten die Integrale der Koordinatenfunktionen x, y, z längs γ dividiert durch die Länge unseres Weges. Stellen wir uns allgemeiner eine Erdwärmeanlage vor, bei der kaltes Wasser in einem Rohr durch heißes Gestein gepumpt wird um am Ende immer noch vergleichsweise kalt aber doch etwas wärmer herauszukommen, und beschreibt γ unser Rohr und f die Wärme der Erde an den jeweiligen Stellen, so würde unser Kurvenintegral nach Einfügen der entsprechenden physikalischen Konstanten die Temperaturdifferenz zwischen eintretendem und austretendem Wasser beschreiben.

10.3.8. Das hier definierte Kurvenintegral wird oft auch als "Wegintegral" bezeichnet. Ich will den Begriff des Wegintegrals jedoch für eine andere Konstruktion reservieren, die in 6.3 besprochen werden wird.

Übung 10.3.9. Gegeben eine stetig differenzierbare Funktion $f : [a, b] \to \mathbb{R}$ wird die Länge ihres Graphen, d.h. die Länge des Weges $\gamma : [a, b] \to \mathbb{R}^2$, $t \mapsto (t, f(t))$ gegeben durch das Integral $L(\gamma) = \int_a^b \sqrt{1 + f'(t)^2}\,\mathrm{d}t$.

10.3.10 (**Gestalt einer hängenden Kette**). Wir gehen hier davon aus, daß die Gestalt einer hängenden Kette durch den Graphen einer stetig differenzierbaren Funktion $f : \mathbb{R} \to \mathbb{R}$ beschrieben wird, und wollen im Folgenden zeigen, daß diese Funktion im Wesentlichen der Cosinus hyperbolicus sein muß. Auf das Kettensegment über einem kompakten Intervall $[a, b]$ wirken die Zugkraft in der Kette von beiden Seiten sowie die Schwerkraft. Bezeichnet L_a^b die Länge des besagten Kettensegments und $v_x = (1, f'(x))$ den Tangentenvektor an unsere Kurve bei $(x, f(x))$ mit 1 als erster Komponente, so bedeutet das Kräftegleichgewicht die vektorielle Gleichung

$$0 = -c_a v_a + c_b v_b - D(0, L_a^b)$$

für geeignete positive Zahlen c_a, c_b und eine positive Konstante D, die von den physikalischen Konstanten unseres Problems abhängen. Durch Betrachtung der ersten Komponenten liefert unsere vektorielle Gleichung für das Kräftegleichgewicht erst einmal $c_a = c_b = c$ und durch Betrachtung der zweiten Komponenten dann

$$cf'(a) - cf'(b) = -DL_a^b = -D \int_a^b \sqrt{1 + f'(x)^2}\,\mathrm{d}x$$

Folglich erfüllt unsere Funktion eine Differentialgleichung der Gestalt

$$f'(a) - f'(b) = -k \int_a^b \sqrt{1 + f'(x)^2}\,\mathrm{d}x$$

für positives $k = D/c$, mithin gilt $f''(x) = k\sqrt{1 + f'(x)^2}$ woraus wir folgern

$$\int_a^b \frac{f''(x)\,\mathrm{d}x}{k\sqrt{1 + f'(x)^2}} = b - a$$

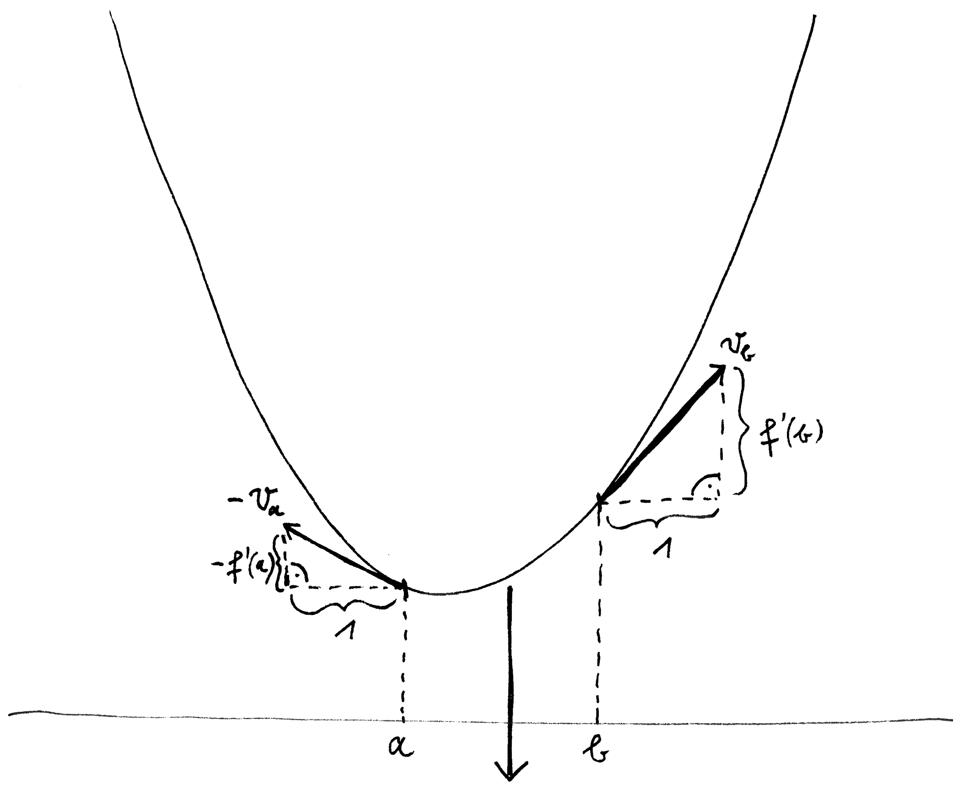


Illustration zur hängenden Kette

und mit der Substitution $f'(x) = y$, $f''(x)\,\mathrm{d}x = \mathrm{d}y$ weiter

$$\int_{f'(a)}^{f'(b)} \frac{\mathrm{d}y}{k\sqrt{1+y^2}} = b - a$$

Dies Integral lösen wir durch die Substitution $y = \sinh t$, $\mathrm{d}y = \cosh t\,\mathrm{d}t$ und erhalten als Stammfunktion für den Integranden $\frac{1}{k}\operatorname{arsinh} y$. Damit ergibt sich $\frac{1}{k}\operatorname{arsinh} f'(b) = b + m$ für eine weitere Konstante m und so $f'(b) = \sinh\left(\frac{b+m}{k}\right)$ und damit schließlich

$$f(b) = k\cosh\left(\frac{b+m}{k}\right) + h$$

für geeignete Konstanten k, m und h. Hier beschreibt k, wie "steil" die Kette hängt, m ist das Negative der x-Koordinate der Stelle kleinster Höhe, und h beschreibt, wie hoch unsere Kette hängt.

10.4 Definition von Sinus und Cosinus

Satz 10.4.1. *Es gibt genau eine stetig differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^2$ mit $\|\gamma(t)\| = \|\gamma'(t)\| = 1$ für alle $t \in \mathbb{R}$, $\gamma_2'(0) > 0$ und $\gamma(0) = (1, 0)$.*

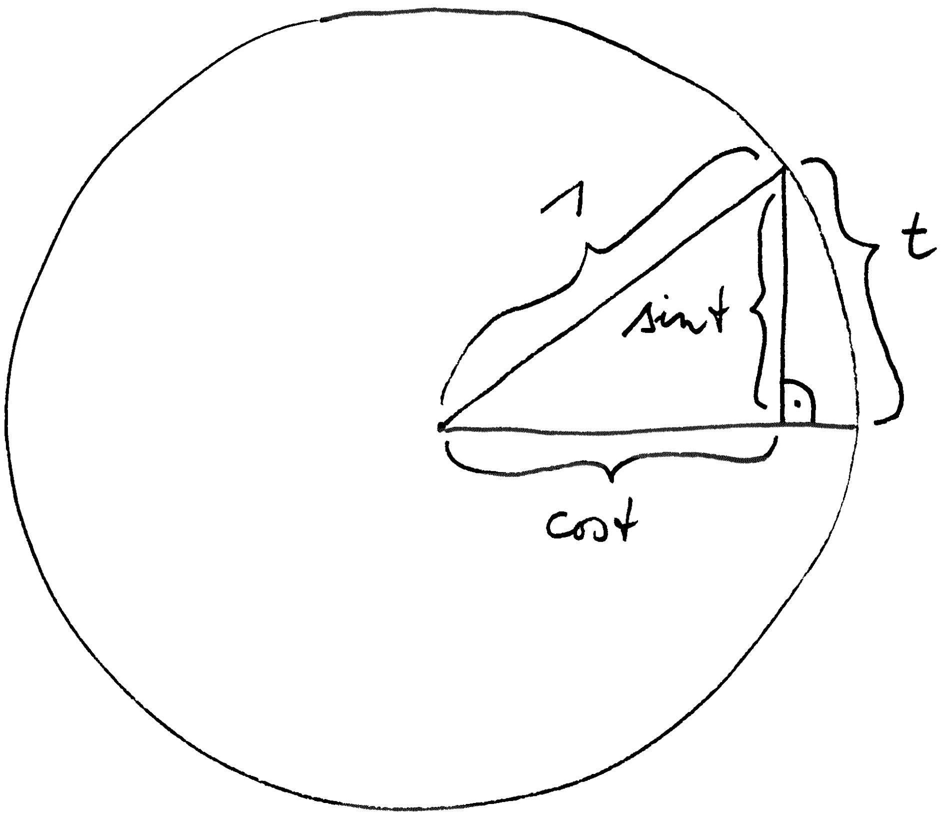
10.4.2. Diese Bedingungen bedeuten anschaulich, daß $\gamma = (\gamma_1, \gamma_2)$ die Bewegung eines punktförmigen Teilchens in der Ebene $\mathbb{R}^2$ beschreibt, das auf dem Einheitskreis $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ mit konstanter absoluter Geschwindigkeit 1 im Gegenuhrzeigersinn umläuft und sich zum Zeitpunkt $t = 0$ an der Stelle $(1, 0)$ befindet. Mithilfe von 7.3.10 können Sie zur Übung sogar zeigen, daß der Satz auch dann noch gilt, wenn wir von unserer Abbildung statt der stetigen Differenzierbarkeit nur die Differenzierbarkeit fordern.

Definition 10.4.3. Wir nennen die beiden Komponenten der Abbildung γ aus dem vorhergehenden Satz 10.4.1 den **Cosinus** und den **Sinus** und notieren sie

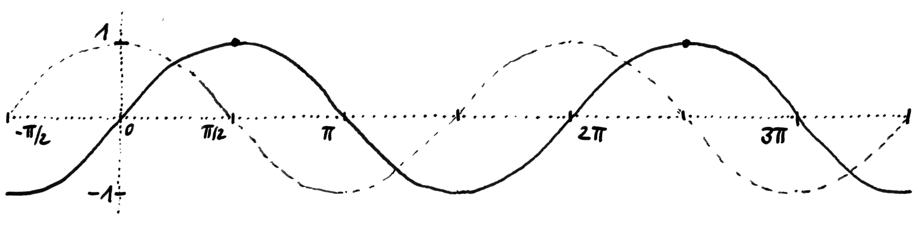
$$\cos, \sin : \mathbb{R} \to \mathbb{R}$$

In Formeln sind die Funktionen Sinus und Cosinus also definiert durch die Gleichung $\gamma(t) = (\cos t, \sin t)$ mit γ unserer eindeutig bestimmten Abbildung aus 10.4.1.

Bemerkung 10.4.4. Auf Taschenrechnern muß man, um die hier definierten Funktionen $\sin$ und $\cos$ zu erhalten, meist noch spezifizieren, daß die Eingabe im Bogenmaß, auf englisch "Radians" oder abgekürzt "rad", zu verstehen sein soll. Auf lateinisch bedeutet Sinus übrigends "Bodenwelle" und "Busen", wortverwandt ist französisch "le sein".



Sinus und Cosinus am Einheitskreis



Die Graphen vom Sinus als durchgezogene Linie und vom Cosinus als gestrichelte Linie. Der Sinus hat seine Maxima genau an den Stellen, an denen der Cosinus verschwindet und eine negative Ableitung besitzt, vergleiche 7.3.15.

10.4.5. Im Vorgriff auf **??** erkläre ich bereits hier für Vektoren $v, w \in \mathbb{R}^n$ ihr **Skalarprodukt** $\langle v, w\rangle \in \mathbb{R}$ durch $\langle v, w\rangle = v_1 w_1 + \ldots + v_n w_n$ für $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$. Die anschauliche Bedeutung wird in **??** erläutert. Offensichtlich gilt $\|v\|^2 = \langle v, v\rangle$ für die euklidische Norm im Sinne von 9.9.7, und jedenfalls im $\mathbb{R}^2$ gilt $\langle v, w\rangle = 0$ offensichtlich genau dann, wenn v und w anschaulich aufeinander senkrecht stehen oder einer der beiden Vektoren Null ist.

Beweis von 10.4.1. Wir zeigen hier nur, daß eine stetig differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^2$ die Bedingungen des Satzes erfüllt genau dann, wenn gilt

$$\gamma(0) = (1, 0),\ \gamma_1' = -\gamma_2 \text{ und } \gamma_2' = \gamma_1.$$

Damit folgt unser Satz 10.4.1 dann aus dem allgemeinen Satz 10.4.9, den wir im Anschluß beweisen. Zunächst ist für eine differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^n$ die Länge $\|\gamma(t)\|$ des Ortsvektors konstant genau dann, wenn ihr Quadrat $\langle\gamma(t), \gamma(t)\rangle$ konstant ist genau dann, wenn dessen Ableitung $2\langle\gamma(t), \gamma'(t)\rangle$ verschwindet genau dann, wenn zu jedem Zeitpunkt t der Geschwindigkeitsvektor $\gamma'(t)$ senkrecht steht auf dem Ortsvektor $\gamma(t)$. Die einzigen auf $(a, b) \in \mathbb{R}^2$ senkrechten Vektoren derselben Länge wie (a, b) sind nun aber $(-b, a)$ und $(b, -a)$. Für eine differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^2$ sind also $\|\gamma(t)\|$ und $\|\gamma'(t)\|$ konstant von derselben Länge genau dann, wenn zu jedem Zeitpunkt t gilt $\gamma'(t) = \pm(-\gamma_2(t), \gamma_1(t))$, wobei das Vorzeichen im Prinzip noch von t abhängen kann. Fordern wir allerdings die Stetigkeit der Ableitung, so muß dieses Vorzeichen konstant sein, und unter der zusätzlichen Bedingung $\gamma(0) = (1, 0)$ ist unser Vorzeichen ein $+$ genau dann, wenn gilt $\gamma_2'(0) > 0$. □

10.4.6. Man kann die Bedingungen $\gamma_1' = -\gamma_2$ und $\gamma_2' = \gamma_1$ zusammenfassen zur Matrix-Gleichung

$$\begin{pmatrix} \gamma_1' \\ \gamma_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

Wir zeigen nun ganz allgemein, wie man Systeme von Differentialgleichungen dieser Art löst. Genauer bestimmen wir für eine gegebene quadratische Matrix $M \in \mathrm{M}(n \times n; \mathbb{R})$ alle differenzierbaren Abbildungen $\gamma : \mathbb{R} \to \mathbb{R}^n$ mit

$$\gamma'(t) = M\gamma(t) \quad \forall t \in \mathbb{R}$$

Bei dieser Schreibweise fassen wir implizit die Elemente des $\mathbb{R}^n$ als Spaltenvektoren auf, also $\gamma = (\gamma_1, \ldots, \gamma_n)^\top$, wo der obere Index $\top$ unsere Zeilenmatrix in eine Spaltenmatrix transponiert. Man nennt so eine Gleichung auch ein **homogenes System von linearen Differentialgleichungen mit konstanten Koeffizienten**. Die Spezifikation "mit konstanten Koeffizienten" grenzt unsere Gleichung ab von dem noch allgemeineren Fall, bei dem auch die Matrix M noch von t abhängt. Die Spezifikation "homogen" grenzt es ab vom allgemeineren Fall einer

Gleichung der Gestalt $\gamma'(t) = M\gamma(t) + f(t)$ für eine zusätzlich gegebene vektorwertige Funktion f, den wir in 2.4.1 diskutieren. Anschaulich gesprochen geben wir uns auf dem $\mathbb{R}^n$ das sehr spezielle Vektorfeld $x \mapsto Mx$ vor und interessieren uns für die Bahnen solcher Teilchen, die bei $x \in \mathbb{R}^n$ jeweils die Geschwindigkeit Mx haben.

10.4.7. Im Fall $n = 1$ hat M genau einen Eintrag $a \in \mathbb{R}$, und wir hatten schon in 7.3.14 gesehen, daß alle Lösungen der Differentialgleichung $\gamma' = a\gamma$ die Form $\gamma(t) = c\exp(at)$ haben. Im Allgemeinen definieren wir die **Exponentialfunktion auf Matrizen** durch die Vorschrift

$$\begin{array}{rccl} \exp: & \mathrm{M}(n \times n; \mathbb{R}) & \to & \mathrm{M}(n \times n; \mathbb{R}) \\ & M & \mapsto & \sum_{k=0}^{\infty} \frac{1}{k!} M^k = I + M + \frac{1}{2}M^2 + \frac{1}{6}M^3 + \dots \end{array}$$

Hier bedeutet $M^0 = I$ nach unserer Konvention 3.1.14 die Einheitsmatrix und unsere unendliche Reihe ist zu verstehen als der Grenzwert der Folge ihrer Partialsummen. Es ist nur noch zu zeigen, daß diese Grenzwerte existieren. Bezeichnen wir dazu für eine quadratische Matrix $M \in \mathrm{M}(n \times n; \mathbb{R})$ mit $|M|$ das Maximum der Absolutbeträge ihrer Einträge, so gilt offensichtlich $|MB| \leq n|M||B|$, also $|M^k| \leq (n|M|)^k$, und dann zeigt die Konvergenz der Exponentialreihe zu $(n|M|)$ schon die absolute Konvergenz aller Reihen von Matrixeinträgen in der Exponentialreihe zu M.

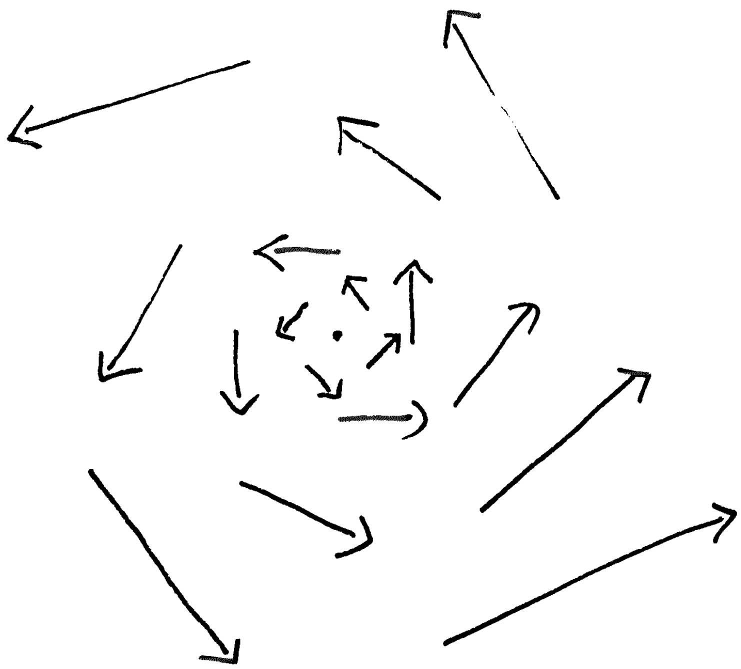
10.4.8. Die Stetigkeit von $\exp : \mathrm{M}(n \times n; \mathbb{R}) \to \mathrm{M}(n \times n; \mathbb{R})$ dürfen Sie in größerer Allgemeinheit als Übung 10.5.27 selbst beweisen.

Satz 10.4.9 (Lineare Differentialgleichungen). *Ist $M \in \mathrm{M}(n \times n; \mathbb{R})$ eine quadratische Matrix und $c \in \mathbb{R}^n$ ein Spaltenvektor, so gibt es genau eine differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^n$ mit Anfangswert $\gamma(0) = c$ derart, daß gilt $\gamma'(t) = M\gamma(t)$ für alle $t \in \mathbb{R}$, und diese Abbildung wird gegeben durch die Vorschrift*

$$\gamma(t) = \exp(tM)c$$

10.4.10. Es ist durchaus möglich, mithilfe dieses Satzes auch ganz konkrete Differentialgleichungen ganz konkret zu lösen. Wir gehen darauf in Abschnitt 2 näher ein.

Beweis. Wir behaupten zunächst, daß die Abbildung $g : \mathbb{R} \to \mathrm{M}(n \times n; \mathbb{R})$, $t \mapsto \exp(tM)$ differenzierbar ist mit der Ableitung $g'(t) = M\exp(tM)$. In der Tat wissen wir nach 8.1.15, daß man Potenzreihen gliedweise differenzieren darf, und unsere Formel ergibt sich, wenn wir diese Erkenntnis anwenden auf alle Einträge unserer Matrix. Nach Lemma 10.2.6 ist nun auch die Abbildung $\gamma : \mathbb{R} \to \mathbb{R}^n$, $t \mapsto \exp(tM)c$ differenzierbar mit Ableitung $\gamma'(t) = M\exp(tM)c = M\gamma(t)$, und die Bedingung $\gamma(0) = c$ ist offensichtlich. Unsere Funktion ist damit eine



Das ebene Vektorfeld $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

Lösung der Differentialgleichung mit dem vorgegebenen Anfangswert. Ist umgekehrt $\gamma(t)$ eine beliebige Lösung unserer Differentialgleichung $\gamma' = M\gamma$, so berechnen wir die Ableitung der Funktion $t \mapsto h(t) = \exp(-tM)\gamma(t)$ mithilfe der matrixwertigen Produktregel 10.2.16 und erhalten

$$h'(t) = -M\exp(-tM)\gamma(t) + \exp(-tM)\gamma'(t) = 0$$

Die Funktion $h(t) = \exp(-tM)\gamma(t)$ ist also konstant mit Wert $\gamma(0)$ und mit dem anschließenden Lemma 10.4.11 folgt $\gamma(t) = \exp(tM)\gamma(0)$. □

Lemma 10.4.11. *Die Exponentialabbildung wirft die Null auf die Identität, und sind A, B zwei kommutierende quadratische Matrizen, in Formeln $AB = BA$, so gilt*

$$\exp(A + B) = (\exp A)(\exp B)$$

10.4.12. Insbesondere folgt $\exp(-A) = (\exp A)^{-1}$. Die Exponentialabbildung ist mithin eine Abbildung von der Menge aller quadratischen Matrizen in die Menge aller invertierbaren quadratischen Matrizen

$$\exp : \mathrm{M}(n \times n; \mathbb{R}) \to \mathrm{GL}(n; \mathbb{R})$$

Für den Beweis des Lemmas geben wir zunächst nur eine Skizze, die dann im anschließenden Abschnitt ausgemalt wird.

Beweisskizze. Genau wie bei der Diskussion des Produkts absolut konvergenter Reihen in 5.6.11 zeigt man

$$(\exp A)(\exp B) = \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} \frac{A^i B^j}{i!j!}$$

Dann faßt man mithilfe von 10.5.20 die Terme mit $i + j = k$ zusammen und landet wegen $AB = BA$ wie beim Beweis der Funktionalgleichung der Exponentialfunktion bei der Reihe für $\exp(A + B)$. Um das alles formal zu rechtfertigen, kann man mit den einzelnen Matrixeinträgen argumentieren und sich so auf unsere Resultate über Reihen reeller Zahlen zurückziehen. Ich will aber stattdessen diese Schwierigkeit als Motivation nutzen und gleich im nächsten Abschnitt 10.5 eine allgemeine Begrifflichkeit entwickeln, in der dieser Beweis einfach und natürlich wird und die auch darüber hinaus von Nutzen ist. □

Übung 10.4.13. Ist $M \in \mathrm{M}(n \times n; \mathbb{R})$ eine quadratische Matrix und $I \subset \mathbb{R}$ ein halboffenes Intervall, so bildet die Menge aller differenzierbaren Abbildungen $\gamma : I \to \mathbb{R}^n$ mit $\gamma'(t) = M\gamma(t)$ für alle $t \in I$ einen Untervektorraum L im Vektorraum $\mathrm{Ens}(I, \mathbb{R}^n)$ aller Abbildungen $I \to \mathbb{R}^n$, den **Lösungsraum** unserer Differentialgleichung, und das Auswerten an einer beliebigen Stelle $t_0 \in I$ definiert einen Vektorraumisomorphismus $L \xrightarrow{\sim} \mathbb{R}^n$, $\gamma \mapsto \gamma(t_0)$, den **Anfangswertisomorphismus**.

Übung 10.4.14. Gegeben eine Diagonalmatrix $M = \operatorname{diag}(a_1, \ldots, a_n)$ haben wir $\exp(M) = \operatorname{diag}(\mathrm{e}^{a_1}, \ldots, \mathrm{e}^{a_n})$. Analoges gilt allgemeiner auch für blockdiagonale Matrizen.

10.5 Vollständigkeit und Exponential von Matrizen

Definition 10.5.1. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum heißt eine **Cauchy-Folge** genau dann, wenn es für jedes $\varepsilon > 0$ ein $N = N_\varepsilon$ gibt derart, daß gilt

$$n, m > N \;\Rightarrow\; d(x_n, x_m) \leq \varepsilon$$

Ein metrischer Raum X heißt **vollständig** genau dann, wenn jede Cauchy-Folge in X konvergiert.

Beispiele 10.5.2. Die Zahlengerade $\mathbb{R}$ ist vollständig nach 5.2.11. Weiter ist offensichtlich jede abgeschlossene Teilmenge eines vollständigen Raums vollständig. Darüber hinaus ist auch jedes endliche Produkt vollständiger metrischer Räume vollständig. Insbesondere ist der $\mathbb{R}^n$ vollständig für den Betragsabstand im Sinne von 9.2.3. Dahingegen ist $X = \mathbb{Q}$ mit dem Betragsabstand kein vollständiger metrischer Raum, und auch wenn wir aus der Zahlengerade einen Punkt entfernen, erhalten wir bereits einen unvollständigen metrischen Raum.

Übung 10.5.3. Jeder kompakte metrische Raum ist vollständig. Jede vollständige Teilmenge eines metrischen Raums ist abgeschlossen.

Übung 10.5.4. Das Produkt zweier vollständiger metrischer Räume ist stets wieder vollständig.

Übung 10.5.5. Jede gleichmäßig stetige Abbildung $f : A \to Y$ von einer Teilmenge A eines metrischen Raums X in einen vollständigen metrischen Raum Y kann auf genau eine Weise zu einer stetigen Abbildung $\bar{A} \to Y$ auf den Abschluß von A in X fortgesetzt werden. Vergleiche auch **??**.

Übung 10.5.6. Konvergiert eine Teilfolge einer Cauchyfolge, so konvergiert bereits die ganze Cauchyfolge, und zwar gegen denselben Grenzwert.

Definition 10.5.7. Unter einem **Banach-Raum** oder genauer einem **reellen Banach-Raum** versteht man einen vollständigen normierten reellen Vektorraum. Sobald wir die komplexen Zahlen kennengelernt haben, werden wir auch und sogar überwiegend mit komplexen Banachräumen arbeiten.

Lemma 10.5.8. *Jeder endlichdimensionale normierte reelle Vektorraum ist vollständig, in anderen Worten also ein Banachraum.*

Beweis. Wir wählen irgendeinen Vektorraumisomorphismus mit dem $\mathbb{R}^n$. Die so induzierte Norm auf dem $\mathbb{R}^n$ ist nach 9.9.21 äquivalent zur Maximumsnorm und liefert also dieselben Cauchyfolgen und dieselben Grenzwerte von Folgen. Die Maximumsnorm auf dem $\mathbb{R}^n$ hinwiederum führt zum Betragsabstand, und für diese Metrik wissen wir seit 10.5.2, daß sie den $\mathbb{R}^n$ zu einem vollständigen metrischen Raum macht. □

Übung 10.5.9. Seien V, W normierte Vektorräume. Ist W vollständig, so ist auch der Raum $\mathcal{B}(V, W)$ der stetigen linearen Abbildungen von V nach W aus 9.9.27 vollständig.

Übung 10.5.10. Ist V ein Banachraum und D eine Menge, so ist auch der Vektorraum $\mathrm{Ens}^{\mathrm{b}}(D, V)$ aus 9.9.8 aller beschränkten Abbildungen von D nach V mit seiner Supremumsnorm vollständig.

Definition 10.5.11. Gegeben ein normierter Vektorraum V und eine Familie $(v_i)_{i\in I}$ von Vektoren aus V sagt man, die Familie der v_i sei **summierbar mit Summe** $s \in V$ und schreibt

$$\sum_{i\in I} v_i = s$$

als Abkürzung für die Aussage, daß es für jede Umgebung U von s eine endliche Teilmenge $I_U \subset I$ gibt derart, daß für jede endliche Obermenge J von I_U in I gilt

$$\sum_{i\in J} v_i \in U$$

Man sieht leicht, daß die Summe einer summierbaren Familie stets eindeutig bestimmt ist. Dieselbe Definition verwenden wir später allgemeiner für beliebige "abelsche Hausdorff'sche topologische Gruppen".

Übung 10.5.12. Eine abzählbare Familie ist summierbar genau dann, wenn für jede Abzählung die Folge der Partialsummen konvergiert und für je zwei Abzählungen die entsprechenden Grenzwerte übereinstimmen.

Übung 10.5.13. Gegeben ein normierter Vektorraum V und eine summierbare Familie $(v_i)_{i\in I}$ von Vektoren von V und eine stetige lineare Abbildung L von V in einen weiteren normierten Vektorraum ist auch die Bildfamilie summierbar und es gilt

$$\sum_{i\in I} L(v_i) = L\left(\sum_{i\in I} v_i\right)$$

Analoges gilt auch allgemeiner für beliebige "abelsche Hausdorff'sche topologische Gruppen".

Ergänzende Übung 10.5.14. Gegeben normierte Vektorräume V, W, X und eine stetige bilineare Abbildung $b : V \times W \to X$ und summierbare Familien $(v_i)_{i\in I}$ von Vektoren von V und $(w_i)_{j\in J}$ von Vektoren von W ist auch die durch $I \times J$ indizierte Familie der $b(v_i, w_j)$ summierbar und es gilt

$$\sum_{(i,j)\in I\times J} b(v_i, w_j) = b\left(\sum_{i\in I} v_i, \sum_{j\in J} w_j\right)$$

Analoges gilt auch allgemeiner für beliebige "abelsche Hausdorff'sche topologische Gruppen".

Definition 10.5.15. Eine Familie $(v_i)_{i\in I}$ von Vektoren in einem normierten Vektorraum heißt **absolut summierbar** genau dann, wenn die Familie ihrer Normen $(\|v_i\|)_{i\in I}$ summierbar ist.

Lemma 10.5.16. *In einem Banachraum ist jede absolut summierbare Familie summierbar und die Norm der Summe kann nach oben abgeschätzt werden durch die Summe der Normen.*

Beweis. Nach 5.5.26 sind bei einer absolut summierbaren Familie höchstens abzählbar viele Vektoren von Null verschieden, so daß wir uns auf Familien beschränken dürfen, die durch $\mathbb{N}$ indiziert sind. Sei also $(v_k)_{k\in\mathbb{N}}$ unsere Familie. Die Partialsummen $s_n = \sum_{k=0}^{n} v_k$ bilden eine Cauchy-Folge, da für $m \geq n$ ja gilt

$$\|s_n - s_m\| = \left\|\sum_{k=n+1}^{m} v_k\right\| \leq \sum_{k=n+1}^{m} \|v_k\| \leq \sum_{k=n+1}^{\infty} \|v_k\|$$

und das wird für hinreichend großes n beliebig klein. Mithin konvergiert die Folge der Partialsummen gegen einen Grenzwert s. Den Nachweis, daß dieser Grenzwert auch die Summe im Sinne der Definition 10.5.11 sein muß, überlasse ich dem Leser. □

Ergänzung 10.5.17. In 5.5.26 hatten wir gesehen, daß jede summierbare Familie reeller Zahlen absolut summierbar ist. Dasselbe gilt für summierbare Familien in endlichdimensionalen normierten Räumen. In beliebigen normierten Räumen gilt es jedoch nicht mehr, ein typisches Gegenbeispiel ist etwa die "Konvergenz im quadratischen Mittel" in 3.3.5 oder allgemeiner in **??**.

Übung 10.5.18. Man zeige, daß eine summierbare Familie in einem Banachraum höchstens abzählbar viele von Null verschiedene Summanden haben kann.

Ergänzung 10.5.19. In allgemeinen "Hausdorff'schen topologischen Vektorräumen" kann es auch summierbare Familien mit überzählbar vielen von Null verschiedenen Summanden geben. Ist zum Beispiel X eine überzählbare Menge mit ihrer diskreten Topologie und $\mathcal{C}(X, \mathbb{R})$ der Raum der reellwertigen Funktionen auf X mit seiner kompakt-offenen Topologie, so ist die Familie der charakteristischen Funktionen aller Punkte von X summierbar mit der konstanten Funktion Eins als Summe.

Ergänzende Übung 10.5.20. Gegeben eine summierbare Familie $(v_i)_{i\in I}$ in einem Banachraum zeige man, daß auch jede Teilfamilie summierbar ist und daß für eine beliebig vorgegebene Zerlegung $I = \bigsqcup_{k\in K} I(k)$ von I in eine Vereinigung von paarweise disjunkten Teilmengen $I(k)$ gilt

$$\sum_{i\in I} v_i = \sum_{k\in K} \left(\sum_{i\in I(k)} v_i \right)$$

Hinweis: Man beginne mit dem Fall, daß K endlich ist. Die Aussage gilt allgemeiner für jede vollständige Hausdorff'sche abelsche topologische Gruppe.

Definition 10.5.21. Ist V ein endlichdimensionaler reeller Vektorraum und $A : V \to V$ eine lineare Abbildung, so definieren wir eine weitere lineare Abbildung $\exp(A) : V \to V$ als den Grenzwert der sogenannten **Exponentialreihe**

$$\exp(A) = \sum_{k\in\mathbb{N}} \frac{A^k}{k!}$$

10.5.22. Wählen wir eine Norm auf V und versehen den Raum $\operatorname{End} V$ aller Endomorphismen von V mit der Operatornorm, so gilt offensichtlich $\|A^k\| \leq \|A\|^k$ und unsere Familie ist summierbar nach 10.5.16, da sie nämlich absolut summierbar ist bezüglich dieser und dann bezüglich jeder Norm. Für eine Operatornorm wie eben erhält man zusätzlich die Abschätzung $\|\exp A\| \leq \exp \|A\|$.

Ergänzung 10.5.23. Ist allgemeiner V ein Banachraum und $A : V \to V$ eine stetige lineare Abbildung, so kann man in derselben Weise eine stetige lineare Abbildung $\exp(A) : V \to V$ erklären. Der Grenzwert ist in diesem Fall im Banachraum $\mathcal{B}(V) := \mathcal{B}(V, V)$ aller stetigen linearen Abbildungen von V in sich selbst aus 10.5.9 zu bilden. Die im Folgenden bewiesenen Aussagen verallgemeinern sich ohne Schwierigkeiten auf diesen Fall. Er ist für die Quantenmechanik fundamental, denn die zeitliche Entwicklung eines quantenmechanischen Systems mit Hamiltonoperator H wird dadurch beschrieben, daß ein Zustand ψ in der Zeitspanne t in den Zustand $\exp(\mathrm{i}\, tH)\psi$ übergeht.

Lemma 10.5.24. *Die Exponentialabbildung wirft die Null auf die Identität, und sind A, B zwei kommutierende Endomorphismen, gilt also in Formeln $AB = BA$, so folgt*

$$\exp(A + B) = (\exp A)(\exp B)$$

10.5.25. Insbesondere folgt $\exp(-A) = (\exp A)^{-1}$, die Exponentialabbildung ist mithin eine Abbildung von der Menge der Endomorphismen in die Menge der Automorphismen $\exp : \operatorname{End} V \to \operatorname{Aut} V$. Die Aussage des Lemmas gilt ganz allgemein für beliebige stetige Endomorphismen von Banachräumen.

Beweis. Genau wie bei der Diskussion des Produkts absolut konvergenter Reihen in 5.6.11 zeigt man zunächst

$$(\exp A)(\exp B) = \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} \frac{A^i B^j}{i!j!}$$

Dann faßt man mithilfe von 10.5.20 die Terme mit $i + j = k$ zusammen und landet wegen $AB = BA$ wie beim Beweis der Funktionalgleichung der Exponentialfunktion 5.6.8 bei der Reihe für $\exp(A + B)$. $\square$

Ergänzung 10.5.26. Es gilt auch eine koordinatenfreie Variante von 10.4.9. Ist genauer V ein Banachraum und $A : V \to V$ eine stetige lineare Abbildung und $c \in V$ ein Vektor, so gibt es genau eine differenzierbare Abbildung $\gamma : \mathbb{R} \to V$ mit $\gamma(0) = c$ und $\gamma'(t) = A\gamma(t) \quad \forall t \in \mathbb{R}$, und diese Abbildung wird gegeben durch die Formel

$$\gamma(t) = \exp(tA)c$$

Der Beweis verläuft völlig analog zum Beweis von 10.4.9. Problematisch ist nur, daß die Produktregel in der benötigten Allgemeinheit erst in 4.4.5 zur Verfügung gestellt wird.

Ergänzende Übung 10.5.27. Für jeden Banachraum V ist $\exp : \mathcal{B}(V) \to \mathcal{B}(V)$ stetig. Hinweis: 9.6.4.

Übung 10.5.28. Sind A, B stetige Endomorphismen von Banachräumen V, W und ist $P : W \to V$ stetig linear mit $AP = PB$, so gilt $(\exp A)P = P(\exp B)$. Ist insbesondere P invertierbar, so gilt $\exp(PAP^{-1}) = P(\exp A)P^{-1}$.

Übung 10.5.29. Gegeben eine Menge D und ein vollständiger metrischer Raum Y ist auch der Raum $\operatorname{Ens}^{\mathrm{b}}(D, Y)$ aller beschränkten Abbildungen von D nach Y vollständig für die Metrik der gleichmäßigen Konvergenz 9.3.5.

Übung 10.5.30. Gegeben ein topologischer Raum D und ein vollständiger metrischer Raum Y ist auch der Raum $\mathcal{C}_{\mathrm{b}}(D, Y)$ aller stetigen beschränkten Abbildungen von D nach Y vollständig für die Metrik der gleichmäßigen Konvergenz 9.3.5. Hinweis: Man verwende 9.6.4.

Übung 10.5.31. Gegeben ein halboffenes kompaktes Intervall $I \subset \mathbb{R}$ und ein endlichdimensionaler reeller Vektorraum V ist auch der Raum $\mathcal{C}^1(I, V)$ aller einmal stetig differenzierbaren Abbildungen von I nach V vollständig für die Norm $\|\gamma\|_\infty + \|\gamma'\|_\infty$ der gleichmäßigen Konvergenz der Funktion und ihrer ersten Ableitung. Hinweis: Man verallgemeinere 9.6.4 und verwende 8.1.14.

Ergänzende Übung 10.5.32. Es gibt eine **stetige Surjektion vom Einheitsintervall** $[0,1]$ **auf das Einheitsquadrat** $[0,1]^2$. Um diese auf den ersten Blick verblüffende Tatsache einzusehen, unterteile man das Einheitsintervall in neun gleiche Abschnitte und das Einheitsquadrat in vier gleiche Quadrate und wähle irgendeinen Weg $[0,1] \to [0,1]^2$, der den $2i$-ten Abschnitt in das i-te Quadrat abbildet, für irgendeine Nummerierung der vier Quadrate. Dann unterteile man die $2i$-ten Abschnitte von eben jeweils in neun gleiche Unterabschnitte und die vier Quadrate von eben jeweils in vier gleiche Unterquadrate und ändere den Weg von eben auf den $2i$-ten Abschnitten von eben so ab, daß sie immer noch im i-ten Quadrat landen und zusätzlich die $2j$-ten Unterabschnitte des $2i$-ten Abschnitts im j-ten Unterquadrat des i-ten Quadrats landen, für irgendeine Nummerierung dieser Unterquadrate. Indem man immer so weitermacht, erhält man eine gleichmäßig konvergente Folge von Abbildungen. Der Grenzwert dieser Folge ist die gesuchte Surjektion.

10.6 Eigenschaften von Sinus und Cosinus

10.6.1. Aus der Definition 10.4.3 und der nachfolgenden Diskussion wissen wir bereits, daß $\sin$ und $\cos$ differenzierbare Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ sind mit $\sin' = \cos$, $\cos' = -\sin$, $\sin(0) = 0$ und $\cos(0) = 1$. Wir wissen weiter, daß gilt $\sin^2 + \cos^2 = 1$, wo wir $(\sin t)^2 = \sin^2 t$ und $(\cos t)^2 = \cos^2 t$ abgekürzt haben. Diese Abkürzungen sind auch üblich für alle anderen trigonometrischen bzw. hyperbolischen trigonometrischen Funktionen, denn das spart Klammern und die alternativ möglichen Bedeutungen $\sin^2 t = \sin(\sin t)$ etc. kommen nie vor.

Satz 10.6.2 (Reihenentwicklung von Sinus und Cosinus). *Unsere Funktionen* $\cos$ *und* $\sin$ *werden dargestellt durch die absolut konvergenten Reihen*

$$\begin{array}{rclcl} \cos t &=& \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} &=& 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots \\ \sin t &=& \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} &=& t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots \end{array}$$

und für alle $t \in \mathbb{R}$ *gilt* $\cos(-t) = \cos t$ *sowie* $\sin(-t) = -\sin t$.

Beweis. Die Reihen ergeben sich aus der Darstellung

$$\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \exp \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

und der zweite Punkt folgt aus den Reihendarstellungen. □

Lemma 10.6.3. *Für alle* $t \in \mathbb{R}$ *gilt*

$$\exp \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

Beweis. Wir müssen nur noch die Gleichheit der zweiten Spalten zeigen, also

$$\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = \exp \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Das folgt aber mit 10.4.9 daraus, daß das Paar von Funktionen $(-\sin t, \cos t)$ auch unserem System von Differentialgleichungen 10.4.6 genügt und darüber hinaus den richtigen Anfangswert hat. □

Proposition 10.6.4 (Additionsformeln). *Für alle $a, b \in \mathbb{R}$ gilt*

$$\begin{array}{rcrcl} \cos(a+b) & = & \cos a \cos b & - & \sin a \sin b \\ \sin(a+b) & = & \cos a \sin b & + & \sin a \cos b \end{array}$$

Beweis. Das folgt mit 10.6.3 aus der Matrixgleichung

$$\exp \begin{pmatrix} 0 & -(a+b) \\ (a+b) & 0 \end{pmatrix} = \exp \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \exp \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$$

die wir ihrerseits aus 10.5.24 folgern. □

Satz 10.6.5 (Nullstellen von Sinus und Cosinus). 1. *Die Nullstellen des Sinus sind genau die ganzzahligen Vielfachen von π, in Formeln $\{t \in \mathbb{R} \mid \sin t = 0\} = \{n\pi \mid n \in \mathbb{Z}\}$.*

2. *Die Nullstellen des Cosinus sind genau die halbzahligen Vielfachen von π, in Formeln $\{t \in \mathbb{R} \mid \cos t = 0\} = \{\frac{\pi}{2} + n\pi \mid n \in \mathbb{Z}\}$.*

Beweis. Zunächst einmal zeigen wir durch Widerspruch, daß der Cosinus überhaupt positive Nullstellen hat. Sicher gilt $\cos(0) = 1 > 0$. Hätte der Cosinus keine positive Nullstelle, so müßte $\cos t$ positiv bleiben für alle $t \geq 0$. Dann müßte nach 7.3.11 also der Sinus streng monoton wachsen auf $[0, \infty)$ und damit müßte

$$\int_0^x \sin t \, \mathrm{d}t = \cos 0 - \cos x$$

für $x \to \infty$ über alle Grenzen wachsen. Das ist aber unmöglich, denn es gilt $\cos^2 \leq \cos^2 + \sin^2 = 1$. Folglich hat der Cosinus eine kleinste positive Nullstelle

$$a = \inf\{t \in \mathbb{R} \mid t > 0 \text{ und } \cos t = 0\}$$

Wir zeigen als nächstes, daß gilt $a = \pi/2$. Der Cosinus ist natürlich positiv auf $(-a, a)$, folglich ist der Sinus streng monoton auf $[-a, a]$, und aus $\cos^2 + \sin^2 = 1$ und $\sin'(0) > 0$ folgt $\sin(-a) = -1$, $\sin(a) = 1$. Der Sinus definiert also eine Bijektion

$$\sin : [-a, a] \xrightarrow{\sim} [-1, 1]$$

Da sich die Länge eines Weges nach 10.1.3 unter surjektiver monotoner Umparametrisierung nicht ändert, ist nun unser in 5.4.1 und in 10.1.3 definiertes π auch die Länge des Weges

$$\begin{array}{rccc} \gamma : & [-a, a] & \to & \mathbb{R}^2 \\ & t & \mapsto & (\sin t, \cos t) \end{array}$$

und wir erhalten

$$\pi = L(\gamma) = \int_{-a}^{a} \sqrt{\cos^2 t + \sin^2 t}\ \mathrm{d}t = \int_{-a}^{a} 1\,\mathrm{d}t = 2a$$

Damit ist in der Tat $\pi/2$ die kleinste positive Nullstelle des Cosinus und wir erhalten gleichzeitig $\sin(\pi/2) = 1$. Mit den Additionsformeln 10.6.4 folgern wir $\sin(x + \frac{\pi}{2}) = \cos x$, $\cos(x + \frac{\pi}{2}) = -\sin x$, insbesondere ergibt sich $\sin(x + \pi) = -\sin x$ und $\cos(x + \pi) = -\cos x$ und damit $\sin(n\pi) = 0 \quad \forall n \in \mathbb{Z}$. Hätte der Sinus noch eine weitere Nullstelle, so fänden wir mithilfe der Formel $\sin(x+\pi) = -\sin x$ auch eine Nullstelle des Sinus in $(0, \pi)$ und damit eine Nullstelle des Cosinus in $(-\pi/2, \pi/2)$. Da gilt $\cos t = \cos(-t)$ hätten wir dann sogar eine Nullstelle des Cosinus in $[0, \pi/2)$, und das ist unmöglich. Also haben der Sinus und dann auch der Cosinus genau die im Satz behaupteten Nullstellen. □

Ergänzende Übung 10.6.6. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^2 \sin(x^{-1})$ für $x \neq 0$ und $f(0) = 0$ ist differenzierbar auf $\mathbb{R}$, aber ihre Ableitung ist nicht stetig beim Nullpunkt.

Satz 10.6.7. *Die Kreiszahl π ist nicht rational, in Formeln $\pi \notin \mathbb{Q}$.*

10.6.8. Das wurde bereits 1766 von Johann Heinrich Lambert gezeigt. Der Beweis der Transzendenz von π ist schwieriger, mir gefällt die Darstellung in [**?**]. Der hier gegebene Beweis der Irrationalität wirkt auf mich wie Zauberei. Ich folge der Darstellung von Stewart [**?**].

Beweis. Man betrachte für reelles $\alpha \neq 0$ und natürliches n das Integral

$$I_n = I_n(\alpha) = \int_{-1}^{1} (1 - x^2)^n \cos(\alpha x)\,\mathrm{d}x$$

Partielles Integrieren liefert $\alpha^2 I_n = 2n(2n-1)I_{n-1} - 4n(n-1)I_{n-2}$ für $n \geq 2$. Vollständige Induktion zeigt dann

$$\alpha^{2n+1} I_n = n!(P_n(\alpha) \sin \alpha + Q_n(\alpha) \cos \alpha)$$

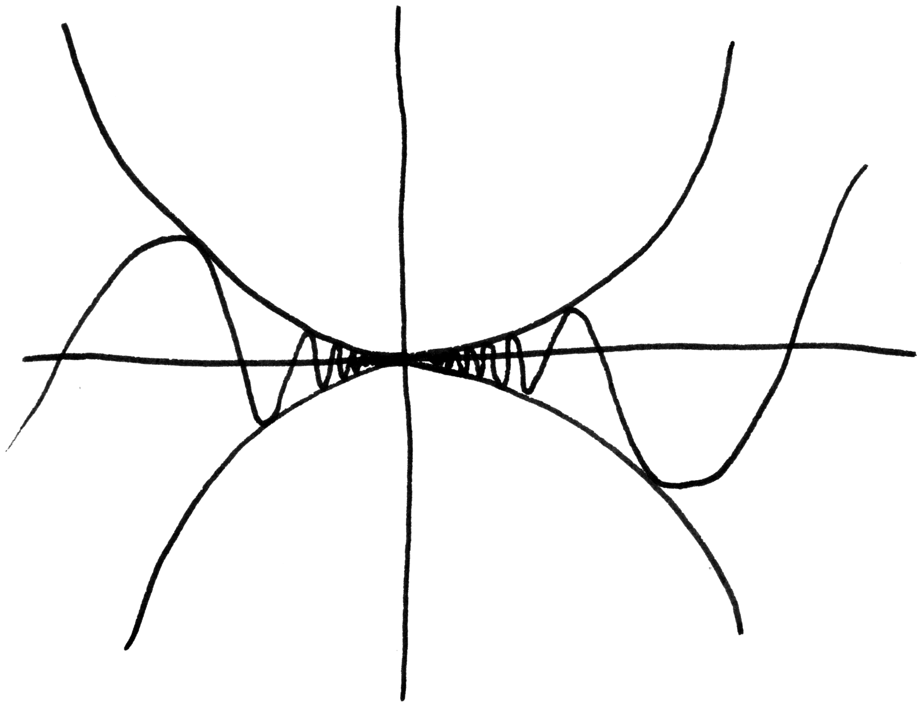


Illustration zu 10.6.6. Der Graph der Funktion ist zwischen zwei parabolischen Backen eingezwängt und hat deshalb am Ursprung Null Grenzwert der Sekantensteigungen, wird aber nah vom Ursprung beliebig steil.

für $P_n, Q_n \in \mathbb{Z}[X]$ Polynome vom Grad $\leq 2n$. Wäre nun $\pi = a/b$ mit $a, b \in \mathbb{Z}$ und setzen wir oben $\alpha = \pi$ ein, so ergäbe sich, daß

$$\frac{a^{2n+1}}{n!} I_n(\pi)$$

für alle $n \in \mathbb{N}$ eine ganze Zahl sein muß im Widerspruch dazu, daß dieser Ausdruck für alle n von Null verschieden ist und für $n \to \infty$ gegen Null strebt. $\square$

Übung 10.6.9. Man führe die partiellen Integrationen des vorhergehenden Beweises aus und prüfe die Induktionsbasis, als da heißt die Fälle $n = 0, 1$.

Satz 10.6.10 (Fläche des Einheitskreises). *Es gilt* $\pi/2 = \int_{-1}^{1} \sqrt{1-t^2}\ \mathrm{d}t$, *anschaulich gesprochen ist also* π *die Fläche des Einheitskreises.*

Beweis. Wir substituieren $t = \sin x$, $\mathrm{d}t = \cos x\ \mathrm{d}x$ und erhalten

$$\int_{-1}^{1} \sqrt{1-t^2}\ \mathrm{d}t = \int_{-\pi/2}^{\pi/2} \cos^2 x\ \mathrm{d}x$$

Mithilfe der Formel $\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos 2x$, die ihrerseits aus den Additionsformel folgt, ergibt sich unser Integral mühelos zu

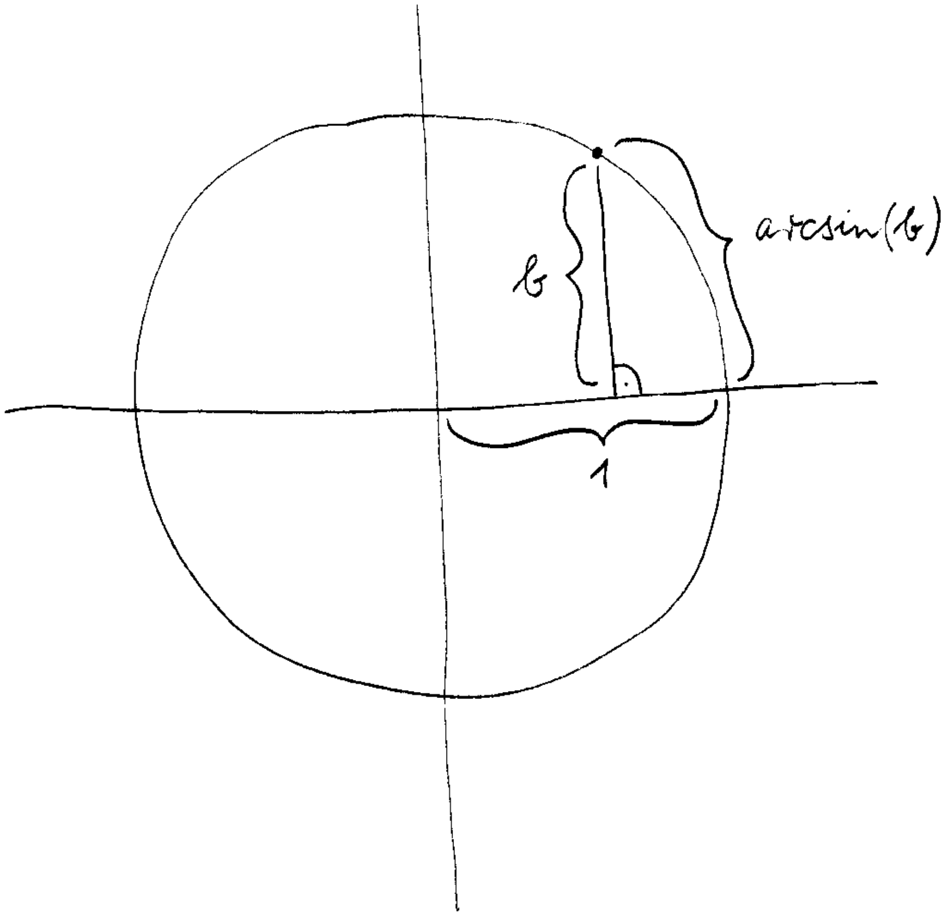
$$\frac{1}{2}x + \frac{1}{4}\sin 2x \Big|_{-\pi/2}^{\pi/2} = \frac{\pi}{2}$$ $\square$

Definition 10.6.11. Der Sinus wächst streng monoton auf $[-\pi/2, \pi/2]$ und definiert folglich eine Bijektion $\sin : [-\pi/2, \pi/2] \to [-1, 1]$, deren Umkehrabbildung man auch den **Arcussinus** nennt und notiert als

$$\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$$

10.6.12. Die Bezeichnung "arcussinus" kommt von lateinisch "arcus" für "Bogen". In der Tat bedeutet $\arcsin b$ für $b \in [0, 1]$ die Länge des Kreisbogens, der vom Punkt $(1, 0)$ bis zum Punkt $(\sqrt{1-b^2}, b)$ der Höhe b auf dem Einheitskreis reicht, wie der Leser zur Übung nachrechnen mag. Der Arcussinus ist nach 7.2.9 differenzierbar auf $(-1, 1)$ und seine Ableitung ergibt sich mit unserer Regel für die Ableitung einer Umkehrfunktion zu

$$\begin{aligned} \arcsin'(x) &= 1/(\cos(\arcsin x)) \\ &= 1/\sqrt{1-\sin^2(\arcsin x)} \\ &= 1/\sqrt{1-x^2} \end{aligned}$$



Der Arcussinus

Damit ergibt sich die Reihenentwicklung

$$\arcsin(x) = x + \frac{1 \cdot x^3}{2 \cdot 3} + \frac{1 \cdot 3 \cdot x^5}{2 \cdot 4 \cdot 5} + \dots \quad \forall x \in (-1, 1)$$

In der Tat gilt $\arcsin(x) = \int_0^x (1-t^2)^{-1/2} \, \mathrm{d}t$, und nach der binomischen Reihe 8.1.19 haben wir

$$(1-t^2)^{-(1/2)} = \sum_{k=0}^{\infty} \binom{-1/2}{k} (-t^2)^k = 1 + \frac{1}{2}\, t^2 + \frac{1 \cdot 3}{2 \cdot 4}\, t^4 + \dots$$

10.6.13. Analog fällt der Cosinus streng monoton auf dem Intervall $[0, \pi]$ und definiert folglich eine Bijektion $\cos : [0, \pi] \xrightarrow{\sim} [-1, 1]$, deren Umkehrabbildung der **Arcuscosinus** heißt und $\arccos$ notiert.

Definition 10.6.14. Für alle $x \in \mathbb{R}$ mit $\cos x \neq 0$ definieren wir den **Tangens** von x durch

$$\tan x = \frac{\sin x}{\cos x}$$

10.6.15. Anschaulich bedeutet $\tan(x)$ für $x \in (0, \pi/2)$ die Höhe, in der der Strahl durch den Nullpunkt und den Punkt des Einheitskreises, der mit dem Punkt $(1, 0)$ ein Kreissegment der Länge x begrenzt, die Tangente an unseren Einheitskreis im Punkt $(1, 0)$ trifft. Man benutzt auch den **Cotangens** $\cot(x) = \cos(x)/\sin(x)$ und eher selten den **Secans** $\sec(x) = 1/\cos(x)$ und **Cosecans** $\operatorname{cosec}(x) = 1/\sin(x)$.

10.6.16. Die Ableitung des Tangens ist $\tan'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2(x)$, insbesondere ist der Tangens streng monoton wachsend auf $(-\pi/2, \pi/2)$. Da er an den Grenzen sogar gegen $\pm\infty$ strebt, liefert der Tangens eine Bijektion $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$, und wir können die Umkehrfunktion **Arcustangens**

$$\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$$

betrachten. Die Ableitung von $\arctan$ ergibt sich mit 7.2.9 zu

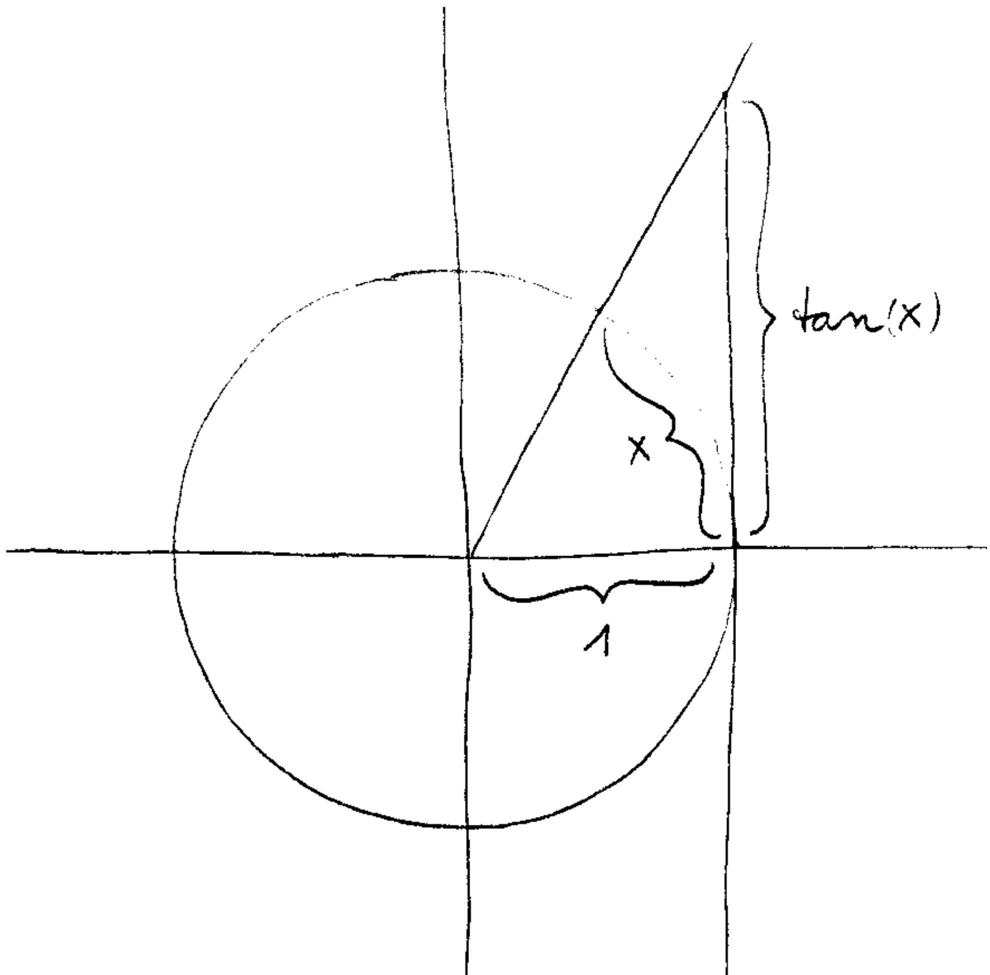
$$\arctan'(t) = \frac{1}{1+t^2}$$

Damit erhalten wir durch gliedweises Integrieren der geometrischen Reihe 5.5.5 für den Arcustangens für $|t| < 1$ die Reihenentwicklung

$$\arctan(t) = t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} \dots$$

und mit dem Abel'schen Grenzwertsatz 8.4.2 ergibt sich

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \dots$$



Der Tangens

Es scheint, daß diese Formel bereits in dem 1530 erschienenen Analysis-Buch "Ganita Yuktibhasa" des Autors Jyesthadeva zu finden ist, eines Mathematikers aus Kerala in Indien, und daß sie auf den indischen Mathematiker Madhava zurückgeht. Außerdem erhalten wir so auch die bemerkenswerte Identität

$$\lim_{x \to \infty} \int_{-x}^{x} \frac{1}{1+t^2} \;\mathrm{d}t = \pi$$

Übung 10.6.17. Sei $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ der Einheitskreis. Wir konstruieren eine Bijektion $\gamma : \mathbb{R} \xrightarrow{\sim} S^1 \setminus (-1, 0)$, indem wir jedem Punkt $t \in \mathbb{R}$ den Schnittpunkt der Gerade durch $(-1, 0)$ und $(0, t)$ mit $S^1 \setminus (-1, 0)$ zuordnen. Man prüfe, daß diese Abbildung gegeben wird durch

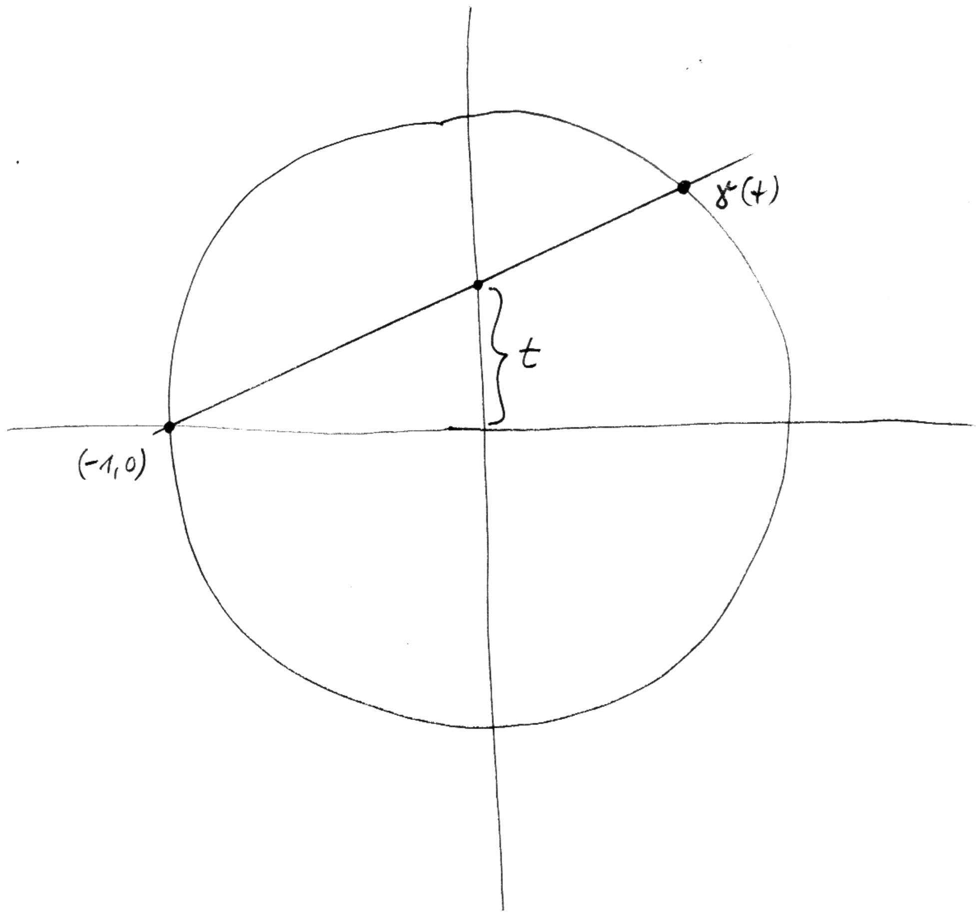
$$\gamma(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right)$$

Man prüfe $\|\gamma'(t)\| = 2/(1+t^2)$ und interpretiere die vorstehende bemerkenswerte Formel. Der Punkt $(\cos\tau, \sin\tau)$ für $\tau \in (-\pi, \pi)$ wird hierbei übrigends parametrisiert durch $t = \tan(\tau/2)$, wie man durch Rechnung oder elementargeometrische Überlegungen prüft. Man beachte auch die Ähnlichkeit zur Parametrisierung der Hyperbel 7.7.6.

Ergänzung 10.6.18. Die Abbildung γ aus der vorstehenden Übung liefert im Übrigen auch eine Bijektion von $\mathbb{Q}$ auf die Punkte von $S^1 \setminus (-1, 0)$ mit rationalen Koordinaten. Diese Bijektion ist äußerst hilfreich bei der Bestimmung aller **pythagoreischen Zahlentripel**, d.h. aller Tripel a, b, c von natürlichen Zahlen mit $a^2 + b^2 = c^2$. Die Abbildung γ aus der vorstehenden Übung liefert allgemeiner sogar für jeden Körper k einer von Zwei verschiedenen Charakteristik eine Bijektion von k auf das Komplement des Punktes $(-1, 0)$ in der Lösungsmenge der Gleichung $x^2 + y^2 = 1$ in der Ebene $k^2 = k \times k$. In der algebraischen Geometrie können Sie dann lernen, wie man das zu einer Bijektion von $\mathbb{P}^1 k$ mit der Quadrik $Q \subset \mathbb{P}^2 k$ erweitert, die durch die homogenisierte Gleichung $x^2 + y^2 = z^2$ definiert wird.

Ergänzende Übung 10.6.19. Mithilfe der Relation $\tan(\pi/6) = 1/2$ berechne man π auf drei sichere Stellen hinter dem Komma. Es gibt im übrigen wesentlich effizientere Verfahren zur Berechnung von π, vergleiche [**?**].

Übung 10.6.20. Man finde Stammfunktionen zu den Kehrwerten quadratischer Polynome, also zu Funktionen der Gestalt $x \mapsto (x^2 + ax + b)^{-1}$. Hinweis: Hat das fragliche quadratische Polynom zwei verschiedene reelle Nullstellen $\lambda \neq \mu$, so kann man unsere Funktion in der Gestalt $\alpha/(x - \lambda) + \beta/(x - \mu)$ schreiben. Sonst bringe man sie in die Form $((x + a/2)^2 + d)^{-1}$ mit $d \geq 0$ und erinnere sich an $\arctan'(t) = 1/(1 + t^2)$.



Die Abbildung γ aus 10.6.17

Übung 10.6.21. Man finde eine Stammfunktion für den Arcustangens. Hinweis: Man wende auf das Produkt $1 \cdot \arctan$ partielle Integration an.

Index