

**Einführung in die
Differential- und Integralrechnung**

für Studierende des Köln-Kollegs
im Leistungskurs des 3./4. Semesters

Unterrichtsbegleitende Skripten sowie
Übungen mit ausführlichen Lösungen

Norbert Klingen

Köln 2003

Inhalt

I. Zahlfolgen und der Grenzwertbegriff

§1 Definitionen und grundlegende Eigenschaften	1
a. Folgen.	1
b. Explizite und rekursive Definitionen.	1
c. Vollständige Induktion.	3
d. Monotonie und Beschränktheit.	4
e. Geometrische Folgen.	6
§2 Konvergenz	7
a. Definition des Konvergenzbegriffs.	7
b. Grenzwertsätze.	8
c. Anwendungen der Grenzwertsätze.	11
d. Konvergenznachweis ohne Kenntnis des Grenzwerts.	12
e. Anwendung: Das Heron'sche Verfahren.	13

II. Rationale Funktionen

§3 Ganz-rationale Funktionen und ihre Nullstellen	15
a. Polynomterme und ganz-rationale Funktionen.	15
b. Nullstellen und Linearfaktoren.	16
c. Rationale Nullstellen ganz-rationaler Funktionen.	20
d. Zerlegung in Linearfaktoren.	21
e. Nullstellenordnung und Vorzeichenwechsel.	21
f. Grad und Graph.	23
g. Grad und Schnittpunkte.	23
§4 Nullstellen und Stetigkeit	24
a. Approximation von Nullstellen.	24
b. Stetigkeit.	25
§5 Rationale Funktionen	27
a. Definitionslücken und Vorzeichenwechsel.	27
b. Grenzwerte von Funktionen.	28
c. Polstellen und hebbare Definitionslücken.	29
d. Das Verhalten im Unendlichen und Asymptoten.	32

III. Differentialrechnung

§6 Der Ableitungsbegriff	34
a. Sekantensteigung, Differenzenquotient, Ableitung, Tangente.	34
b. Differenzierbarkeit und Stetigkeit.	36
c. Die Ableitung der Potenzfunktionen.	38
d. Erste Ableitungsregeln.	40
§7 Monotonie und Extrema	41
a. Extremstellen und stationäre Stellen.	41
b. Monotonieintervalle.	42
c. Der Monotoniesatz.	44
d. Beweise.	44
§8 Höhere Ableitungen	46
a. Krümmung und Wendestellen.	46
b. Vorzeichenwechsel und Ableitungen.	47
c. Hinreichende Kriterien für Extrem-/ Wendestellen mittels höherer Ableitungen.	48

§9 Weitere Ableitungsregeln	48
a. Produkt- und Quotientenregel.....	48
b. Kettenregel.....	50

IV. Transzendente Funktionen

§10 Exponential- und Logarithmusfunktionen	52
a. Die Exponentialfunktionen.....	52
b. Differenzierbarkeit der Exponentialfunktionen.....	54
c. Logarithmusfunktionen.....	54
§11 Der natürliche Logarithmus und die Eulersche Zahl e.	57
a. Der natürliche Logarithmus.....	57
b. Die e -Funktion.....	58
c. Die Regeln von de l'Hospital.....	59
§12 Die trigonometrischen Funktionen	62
a. Definition am Einheitskreis.....	62
b. Additionstheoreme und Ableitung der trigonometrischen Funktionen.....	63
c. Die Schwingungsdifferentialgleichung.....	65

V. Integralrechnung

§13 Flächeninhalt und Integral	67
a. Grundprinzipien der Flächenberechnung.....	67
b. Intervallzerlegungen.....	67
c. Ober- und Untersummen.....	68
d. Das Integral.....	69
e. Integrierbare Funktionen.....	71
§14 Die Berechnung von Integralen und der Hauptsatz	72
a. Berechnung mittels Ober-/Untersummen.....	72
b. Erste Integrationsregeln.....	75
c. Integralfunktionen und der Hauptsatz.....	77
d. Stammfunktionen und die Integralformel.....	78
e. Bestimmung von Stammfunktionen.....	79
§15 Fläche, Volumen, Bogenlänge	81
a. Flächen zwischen Graph und x -Achse.....	81
b. Flächen zwischen Graphen.....	82
c. Volumina von Rotationskörpern.....	84
d. Die Bogenlänge.....	84
§16 Ergänzungen	86
a. Substitution.....	86
b. Partielle Integration.....	88
c. Anwendung: Die Kreiszahl π	88
d. Flächeninhalt und Bogenlänge von Kreisen.....	91

I. Zahlfolgen und der Grenzwertbegriff

§1 Definitionen und grundlegende Eigenschaften

a. Folgen. Bei zahlreichen Gelegenheiten werden irgendwelche Objekte nummeriert und listenmäßig erfasst. Zum Beispiel die Namensliste Ihres Leistungskurses Mathematik auf der regelmäßig umlaufenden Anwesenheitsliste:

1. Albrecht, Felix
2. Dubois, Jeannette
3. Gafert, Daniel
4. Glaubitz, Dominique
5. Hira, Brinder Singh
6. Kazemi-Tabar, Marcel
7. Klapproth, Oliver
8. Naumova, Natalia
9. Reichert, Tobias
10. Rieger, Thomas
11. Vesen, Daniel
12. Werner, Enrico
13. Wolff, Esther

Durch eine solche Liste wird eine *Zuordnung* hergestellt: $1 \mapsto$ Albrecht, $2 \mapsto$ Dubois, $3 \mapsto$ Gafert, ... (Lesen Sie ' \mapsto ' als 'wird zugeordnet' oder 'wird abgebildet auf'.) Da den Zahlen jeweils genau ein Name zugeordnet wird, liegt eine *Funktion* vor, deren *Definitionsbereich* eine Menge von natürlichen Zahlen ist. Eine Funktion dieser Art nennt man eine *Folge*. Bei dem genannten Beispiel ist der Definitionsbereich die Menge $\{1, \dots, 13\}$, da Ihr Leistungskurs 13 Teilnehmer hat. Der Definitionsbereich ist somit eine endliche Menge. Man spricht daher hier auch von einer *endlichen* Folge.

Von besonderer Bedeutung sind solche Folgen, deren *Werte* Zahlen sind. Sie nennt man *Zahlfolgen*. Willkürliche Beispiele sind etwa 2, 4, 6, 8 ... oder 1, 4, 9, 16 ... oder 2, 3, 5, 7, 11, 13 ... Wir werden uns dabei im Folgenden vorwiegend für *unendliche* Zahlfolgen interessieren, und sprechen dann einfach von Folgen:

(1.1) Definition: Eine (unendliche) *Folge* ist eine Funktion, deren Definitionsbereich die Menge der natürlichen Zahlen \mathbb{N} ist und deren Werte (beliebige reelle) Zahlen sind.

Für Zahlfolgen haben sich besondere Zeichen eingebürgert: Den Funktionsterm $f(n)$ bezeichnet man bei Folgen oft in der *Indexschreibweise* mit a_n und nennt ihn auch das n -te Folgenglied. Die gesamte Folge wird mit $(a_n)_{n \in \mathbb{N}}$ oder auch kurz mit (a_n) bezeichnet. Manchmal ist es nützlich, die Numerierung der Folgenglieder schon mit der 0 beginnen zu lassen; dann ist $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ der Definitionsbereich und man schreibt für die Folge $(a_n)_{n \in \mathbb{N}_0}$.

Übung: Lehrbuch S. 68, Nr. 3,4.

b. Explizite und rekursive Definitionen. In der letzten Übungsaufgabe Nr. 4 werden die Folgen nicht *explizit* durch Angabe eines Terms für a_n definiert, sondern durch eine sogenannte *Rekursion*. Diese besteht aus zwei Teilen: der Definition des Anfangsgliedes a_1 und einer Vorschrift, mit der man aus dem Folgenglied a_n das nachfolgende a_{n+1} berechnet. Siehe Übungsaufgaben (1).

Um die Übereinstimmung einer explizit und einer rekursiv definierten Folge zu zeigen, geht man wie folgt vor:

1. Man zeigt, dass bei beiden Definitionen *das jeweils erste Folgenglied* übereinstimmt.
2. Sodann zeigt man, dass die explizite Definition die rekursive *erfüllt*. Dazu setzt man die explizite Definition in die rekursive ein und muss dann zeigen, dass die entstehende Beziehung für alle n gültig ist.

Wir wollen dies einmal an den Beispielen der Übungen (1) durchführen.

1. Beispiel: Übereinstimmung der expliziten Definition 5) mit der rekursiven Definition 4):

1. Bei beiden Definitionen ergibt sich $a_1 = 1$.
2. Wir setzen die explizite Definition (siehe 5))

$$a_n = \frac{n(n+1)}{2} \quad (1)$$

in die rekursive Definition (siehe 4))

$$a_n = a_{n-1} + n \quad (2)$$

ein. Dazu müssen wir die explizite Form für a_n und a_{n-1} zur Verfügung haben. a_n ist in (1) gegeben. Wir erhalten daraus a_{n-1} , wenn wir in (1) *überall* n durch $n-1$ ersetzen. (Dies ist zulässig, da ja (1) für alle natürlichen Zahlen gültig ist.)

$$a_{n-1} = \frac{(n-1)(n-1+1)}{2} = \frac{(n-1)n}{2}. \quad (1')$$

Setzt man (1) und (1') in (2) ein, so erhält man die Formel

$$\frac{n(n+1)}{2} = \frac{(n-1)n}{2} + n. \quad (3)$$

Wir müssen nun zeigen, dass diese Gleichung (3) über der Grundmenge \mathbb{N} allgemeingültig ist, d. h. dass für alle $n \in \mathbb{N}$ wahr ist. Dazu formen wir (3) äquivalent um:

$$\begin{aligned} \frac{n(n+1)}{2} = \frac{(n-1)n}{2} + n &\iff n(n+1) = (n-1)n + 2n \\ &\iff n^2 + n = n^2 - n + 2n \iff n^2 + n = n^2 + n. \end{aligned}$$

Die letzte der Gleichungen ist offenbar allgemeingültig, da die Terme auf beiden Seiten identisch sind. Damit ist auch (3) allgemeingültig und wir haben gezeigt: Die durch (1) explizit definierte Folge erfüllt die rekursive Definition (2); beide Folgen sind also identisch.

2. Beispiel: Die Übereinstimmung der Folgen 7) und 8) der Übungen (1).

1. Das erste Folgenglied ist in beiden Fällen $a_1 = 1$.

2. Die explizite Form von 7) ist

$$a_n = n^2. \quad (1)$$

Dies ergibt für $n-1$

$$a_{n-1} = (n-1)^2. \quad (1')$$

Setzt man dies in die rekursive Definition (siehe 8))

$$a_n = a_{n-1} + 2n - 1 \quad (2)$$

ein, so ergibt sich die Gleichung

$$n^2 = (n-1)^2 + 2n - 1. \quad (3)$$

Wir müssen (3) also zeigen, dass diese Gleichung allgemeingültig ist.

$$n^2 = (n-1)^2 + 2n - 1 \iff n^2 = n^2 - 2n + 1 + 2n - 1 \iff n^2 = n^2.$$

Wieder erhalten wir durch einfache Äquivalenzumformungen, dass die Gleichung (3) allgemeingültig ist, also die explizite Definition (1) und die rekursive Definition (2) dieselbe Folge definieren.

Durch den Nachweis der Übereinstimmung der Folgen 4),5) bzw. 7),8) haben wir die folgenden interessanten ‘Summationsformeln’ bewiesen:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2} \quad \text{und} \quad 1 + 3 + 5 + \dots + (2n-1) = n^2$$

c. Vollständige Induktion. Die obigen Überlegungen sind Spezialfälle des sog. Beweisprinzips der *vollständigen Induktion*. Diese ist ein wichtiges Mittel, Formeln (oder allgemein Aussageformen) für alle natürliche Zahlen zu beweisen. Sie steht im Kontrast zum Prinzip der (unvollständigen) Induktion, bei dem aufgrund von *Einzelfällen* auf eine allgemeine Gesetzmäßigkeit geschlossen wird. Dieses in allen Naturwissenschaften verwendete erkenntnistheoretische Prinzip ist keine *Beweismethode* im mathematisch-logischen Sinne.

Es sei $A(n)$ eine zu beweisende Formel (oder Aussageform) mit der Variablen n für beliebige natürliche Zahlen. Beispiele für derartige Formeln sind etwa

$$\begin{aligned} A(n) : \quad & 1 + 2 + \dots + n = \frac{n(n+1)}{2}, \\ B(n) : \quad & 1 + 3 + 5 + \dots + (2n-1) = n^2, \\ C(n) : \quad & 1^2 + 2^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1), \\ D(n) : \quad & 1^3 + 2^3 + \dots + n^3 = \frac{n^2(n+1)^2}{4}. \end{aligned}$$

Will man derartige Formeln für alle n als richtig nachweisen, so kann man dies oft nicht *direkt* auf einen Schlag schaffen, vielmehr beweist man die Aussagen der Reihe nach für $n = 1, 2, 3, \dots$. Da man aber nicht unendlich viele Einzelbeweise führen kann, benutzt man hier das Prinzip der vollständigen Induktion, das folgendermaßen lautet:

(1.2) Beweisprinzip durch vollständige Induktion:

Zum Beweis einer Aussageform $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$ genügt der Nachweis folgender Tatsachen:

Induktionsanfang: $A(n)$ gilt für $n = 1$.

Induktionsschritt für beliebiges n :

Aus der Gültigkeit von $A(n)$ folgt auch die Gültigkeit von $A(n+1)$.

Man erkennt den einleuchtenden Grundgedanken: Wenn $A(1)$ bewiesen ist (Induktionsanfang), muss nach dem Induktionsschritt auch die nächste Aussage $A(2)$ gelten, daraus folgt wieder nach dem Induktionsschritt die Gültigkeit von $A(3)$, dann $A(4)$, $A(5)$ usw.

Der Vorteil des Induktionsprinzips ist die Tatsache, dass man beim Nachweis von $A(n+1)$ die Gültigkeit von $A(n)$ *voraussetzen* kann. Man hat also scheinbar zusätzliche Information gewonnen. Wir wollen dies an den oben angegebenen Aussageformen exemplarisch verdeutlichen.
Nachweis von $A(n)$ für alle n :

Induktionsanfang $n = 1$: Setzt man $n = 1$, so reduziert sich die Summe auf der linken Seite von $A(n)$ auf *einen* Summanden 1, und $A(1)$ lautet dann $1 = \frac{1(1+1)}{2}$, was ganz offensichtlich wahr ist. Damit ist der Induktionsanfang bewiesen.

Induktionsschritt $n \rightarrow n+1$: Wir fixieren ein $n \in \mathbb{N}$ und gehen davon aus, dass die Formel

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

gilt. Wir müssen nun die Gültigkeit von

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

nachweisen. (Man nennt $A(n+1)$ die Induktions*behauptung* und $A(n)$ die Induktions*voraussetzung*.) Wir formen nun $A(n+1)$ äquivalent um, wobei wir $A(n)$ benutzen dürfen:

$$\begin{aligned} A(n+1) &\iff (1+2+\dots+n) + (n+1) = \frac{(n+1)(n+2)}{2} \\ &\stackrel{A(n)}{\iff} \frac{n(n+1)}{2} + n+1 = \frac{(n+1)(n+2)}{2} \\ &\iff n(n+1) + 2(n+1) = (n+1)(n+2) \\ &\iff (n+2)(n+1) = (n+1)(n+2) \end{aligned}$$

(Begründen Sie alle Umformungsschritte!) Die letzte Gleichung ist offenbar allgemeingültig, womit der Induktionsschritt bewiesen ist. Nach dem Prinzip der vollständigen Induktion ist $A(n)$ eine über \mathbb{N} *allgemeingültige* Aussage.

Nachweis von $B(n)$ für alle n :

Induktionsanfang $n=1$: Setzt man $n=1$, so reduziert sich die Summe auf der linken Seite von $B(n)$ auf *einen* Summanden 1, und $B(1)$ lautet dann $1 = 1^2$, was ganz offensichtlich wahr ist. Damit ist der Induktionsanfang bewiesen.

Induktionsschritt $n \rightarrow n+1$: Wir fixieren ein $n \in \mathbb{N}$ und wollen

$$B(n+1) : \quad 1 + 3 + \dots + (2n-1) + (2(n+1)-1) = (n+1)^2$$

beweisen, wobei wir

$$B(n) : \quad 1 + 3 + \dots + (2n-1) = n^2$$

voraussetzen können. Wieder formen wir $B(n+1)$ mittels $B(n)$ äquivalent um

$$B(n+1) \stackrel{B(n)}{\iff} n^2 + (2n+1) = (n+1)^2$$

und erkennen unmittelbar, dass diese Aussage allgemeingültig ist (binomische Formel!).

Übung: Beweisen Sie, dass auch die Formeln $C(n)$ und $D(n)$ allgemeingültig sind. Welche erstaunliche Beziehung ergibt sich aus der Kombination von $A(n)$ mit $D(n)$?

d. Monotonie und Beschränktheit. Bei der Untersuchung der Folgen von Übungen (1) haben wir uns auf die folgenden Eigenschaften konzentriert:

(1.3) Definition: Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt

... *alternierend*, wenn die Folgenglieder ständig ihr Vorzeichen wechseln.

... *nach oben beschränkt*, wenn es eine Zahl S gibt mit $a_n \leq S$ für alle $n \in \mathbb{N}$. Eine solche Zahl S nennt man eine *obere Schranke*. Entsprechend definiert man *untere Schranke* und *nach unten beschränkt*.

... *beschränkt*, wenn sie nach oben *und* nach unten beschränkt ist.

... *monoton steigend*, wenn für alle $n \in \mathbb{N}$ gilt $a_n \leq a_{n+1}$. Entsprechend definiert man *monoton fallend*.

... *monoton*, wenn sie monoton steigend *oder* monoton fallend ist.

Zum *Nachweis* dieser Eigenschaften für konkrete Folgen geht man etwa wie folgt vor. Untersucht man eine Folge auf Monotonie, etwa ob sie monoton steigt, so hat man zu zeigen, dass die Ungleichung

$$a_n \leq a_{n+1}$$

allgemeingültig ist, d. h. ihre Lösungsmenge ganz \mathbb{N} ist. Man hat also eine *Ungleichung zu lösen*. Wir führen dies einmal für die Folge

$$a_n = \frac{3n-5}{2n+7}$$

durch. Aufgrund der Berechnung der ersten Folgenglieder kann man vermuten, dass die Folge monoton steigt. Wir untersuchen also die Ungleichung $a_n \leq a_{n+1}$. Dazu bestimmen wir

zunächst eine Formel für a_{n+1} , indem wir in der gegebenen Formel für a_n die Variable n durch $n + 1$ ersetzen:

$$a_n = \frac{3n - 5}{2n + 7} \implies a_{n+1} = \frac{3(n+1) - 5}{2(n+1) + 7} = \frac{3n - 2}{2n + 9}.$$

Dann gilt

$$\begin{aligned} & a_n \leq a_{n+1} \\ \iff & \frac{3n - 5}{2n + 7} \leq \frac{3n - 2}{2n + 9} \quad | \cdot (2n + 7)(2n + 9) > 0 \\ \iff & (3n - 5)(2n + 9) \leq (3n - 2)(2n + 7) \\ \iff & 6n^2 + 17n - 45 \leq 6n^2 + 17n - 14 \quad | -6n^2 - 17n \\ \iff & -45 \leq -14. \end{aligned}$$

(Wiederholen Sie zur Übung die für Ungleichungen gültigen Äquivalenzumformungen.) Nun ist die letzte Ungleichung für alle natürlichen Zahlen n richtig, also $a_n \leq a_{n+1}$ für alle n , was zu zeigen war.

Bei Monotonienachweisen rekursiv definierter Folgen beachte man, dass wegen $a_n \leq a_{n+1}$ $\iff a_{n+1} - a_n \geq 0$ gilt:

$$(a_n)_{n \in \mathbb{N}} \begin{cases} \text{monoton steigend} \\ \text{monoton fallend} \end{cases} \iff a_{n+1} - a_n \begin{cases} \geq 0 \\ \leq 0 \end{cases} \text{ für alle } n \in \mathbb{N}.$$

In den Beispielen 4) und 8) ist aufgrund der rekursiven Definition die Differenz zweier aufeinanderfolgender Glieder explizit angegeben: in 4) gilt $a_n - a_{n-1} = n \geq 0$ und in 8) $a_n - a_{n-1} = 2n - 1 \geq 0$. Beide Folgen sind also monoton wachsend.

Wir wollen uns nun dem *Nachweis der Beschränktheit* von Folgen zuwenden. Auch dies führt auf die Lösung von Ungleichungen. Jedoch muss man *zunächst einen Kandidaten für eine obere/untere Schranke finden!* Hat man diesen jedoch gefunden, so muss man zum *Nachweis* der Beschränktheit (etwa nach oben) lediglich zeigen, dass die folgende Ungleichung für alle $n \in \mathbb{N}$ gültig ist:

$$a_n \leq S.$$

Wir wollen dieses Vorgehen wieder an der schon oben betrachteten Folge

$$a_n = \frac{3n - 5}{2n + 7}$$

erläutern. Aufgrund der Berechnung der ersten Folgenglieder *vermuten* wir, dass diese Folge nach oben durch $S = 2$ beschränkt sein *könnte*. Zum *Nachweis* müssen wir die folgende Ungleichung als allgemeingültig nachweisen:

$$\begin{aligned} & a_n \leq 2 \\ \iff & \frac{3n - 5}{2n + 7} \leq 2 \quad | \cdot (2n + 7) > 0 \\ \iff & 3n - 5 \leq 4n + 14 \quad | -3n - 14 \\ \iff & -19 \leq n \end{aligned}$$

Wieder ist die letzte Ungleichung über der Grundmenge \mathbb{N} allgemeingültig und die Behauptung ist bewiesen. (Beweisen Sie zur Übung auf gleichem Wege, dass sogar $3/2$ eine obere Schranke ist.)

Man kann diese Methode der Rückführung auf zu lösende Ungleichungen auch benutzen, um evtl. eine Schranke zu finden. Auch dazu ein Beispiel:

$$a_n = \frac{9n + 14}{4n + 12}.$$

Wir suchen eine obere Schranke S , d. h. eine Zahl S , für die die Ungleichung $a_n \leq S$ allgemeingültig ist. Dies bedeutet, dass wir die Ungleichung $a_n \leq S$ für beliebiges S lösen müssen. Man hat es daher mit einer Ungleichung mit einem *Parameter* (nämlich S) zu tun. (Die Unbekannte ist nach wie vor n .) Mittels Äquivalenzumformungen erhalten wir:

$$\begin{aligned} & a_n \leq S \\ \iff & \frac{9n+14}{4n+12} \leq S & | \cdot(4n+12) > 0 \\ \iff & 9n+14 \leq S \cdot (4n+12) & | -9n - 12S \\ \iff & 14 - 12S \leq (4S - 9)n \end{aligned}$$

Wir erkennen an der letzten Ungleichung: Wählt man man S so, dass $4S - 9 > 0$ ist (zum Beispiel $S = 3$), so wird die rechte Seite positiv, die linke dagegen negativ, die Ungleichung also für jedes $n \in \mathbb{N}$ gültig. Damit ist $S = 3$ eine obere Schranke für die Folge (a_n) .

Man kann durch *systematische* Lösung der letzten Ungleichung mit allen notwendigen Fallunterscheidungen (vgl. Skript Einführungsphase, Abschnitt 4.d.) zeigen, dass $S = 9/4$ die *kleinste obere Schranke* ist.

e. Geometrische Folgen. Unter den Beispielen der Übungsaufgaben (1) kamen einige miteinander verwandte Folgen vor, die Folgen 1), 6) und 11).

(1.4) Definition: Eine *geometrische Folge* ist eine Folge (a_n) , bei der jedes Folgenglied durch Multiplikation mit einer festen Zahl $q \neq 0$ aus dem vorangehenden entsteht:

$$a_n = q \cdot a_{n-1} \text{ für alle } n \in \mathbb{N}. \quad (1)$$

Bei geometrischen Folgen nummeriert man (üblicher- und sinnvollerweise) bei $n = 0$ beginnend. Eine geometrische Folge ist also durch ihr Anfangsglied a_0 und den Quotienten q eindeutig bestimmt. Man kann sie explizit beschreiben durch

$$a_n = a_0 \cdot q^n \text{ für alle } n \in \mathbb{N}_0. \quad (2)$$

Übung: Bestimmen Sie q für die Beispielfolgen 1), 6) und 11). (Antwort: 1) $q = -1$, 6) $q = 3$, 11) $q = -1/3$.) Vergleichen Sie die auf dem Lösungsblatt angegebene explizite Form mit der in Formel (2).

Übungen: Lehrbuch, S. 71, Nr. 6,7,8,9.

Geometrische Folgen sind in vielen Anwendungen von Bedeutung: Verzinsung, Wachstum, radioaktiver Zerfall. In all diesen Fällen ist die betrachtete Größe ein festes Vielfaches der vorherigen. Siehe etwa Lehrbuch, S. 70, Nr. 2.

Übungen: Lehrbuch S. 71, Nr. 10–12.

Wir wollen nun die geometrischen Folgen hinsichtlich der in (1.3) definierten Eigenschaften untersuchen. Wir stellen die Eigenschaften nur zusammen; ihr Nachweis ist in keinem Falle schwierig.

(1.5) Eigenschaften geometrischer Folgen: Wir betrachten eine geometrische Folge (a_n) mit Anfangsglied a_0 und Quotient q , also $a_n = a_{n-1} \cdot q = a_0 \cdot q^n$.

a) Ist $q < 0$, so ist (a_n) alternierend.

b) Ist $q > 1$ und $a_0 > 0$, so ist (a_n) monoton wachsend und unbeschränkt.

c) Ist $0 < q < 1$ und $a_0 > 0$, so ist (a_n) monoton fallend und beschränkt; 0 ist eine untere und a_0 eine obere Schranke.

d) Ist $q < 0$, so sind die Aussagen von b), c) für die Folge $|a_n|$ (mit Quotient $|q|$) gültig.

e) Ist $a_0 < 0$, so gelten die Aussagen von b) und c), wenn man 'oben und unten' vertauscht. (Präzisieren Sie dies.)

§2 Konvergenz

a. Definition des Konvergenzbegriffs. Wir kommen nun zu dem *fundamentalen Begriff* der ganzen Analysis, dem *Konvergenzbegriff*. Wir behandeln ihn hier zunächst für Folgen. Bei den Beispielen der Übungsaufgaben (1) kamen Folgen vor (Nr. 2,10,11), bei denen sich die Folgenglieder mit fortschreitendem n immer geringfügiger änderten und sich einer Zahl annäherten. In den Beispielen 2) und 11) näherten sich die Folgenglieder immer mehr der Zahl 0, während im Beispiel 10) zwar die Folgenglieder sich immer weniger änderten und auch einem Zahlwert zuzustreben schienen, es aber nicht ersichtlich war, welcher Zahl. (Wir werden später darauf zurückkommen.)

Wir wollen zunächst dieses von Ihnen allen beobachtete Phänomen der Annäherung einer Folge an eine Zahl mathematisch präzisieren. Der mathematische Fachausdruck dafür ist ‘*Konvergenz*’. In einem ersten Schritt sprechen wir von Konvergenz einer Zahlfolge $(a_n)_{n \in \mathbb{N}}$ gegen einen Grenzwert a , wenn die Zahlen a_n der Zahl a *beliebig nahe kommen*.

Wie kann man nun dieses ‘*nahe kommen*’ präzisieren? Die Nähe zur Zahl a beschreibt man durch den Abstand von ihr. Man gibt also eine positive Zahl ε vor (etwa $1/10$, $1/100$,...) und verlangt, dass die Folgenglieder a_n von a weniger als ε Abstand haben:

$$|a_n - a| < \varepsilon \quad \text{bzw.} \quad a - \varepsilon < a_n < a + \varepsilon. \quad (*)$$

Aber für welche Folgenglieder soll dies gelten? Wir haben in Beispielen gesehen, dass bei kleinem ε natürlich nicht alle Folgenglieder so nahe bei a liegen. Auf der anderen Seite genügt es aber auch nicht, dass nur eines oder einige der a_n diese Abschätzung (*) erfüllen. Z. B. bei der Folge 1) auf dem ersten Übungsblatt kommt jedes zweite Folgenglied der Zahl 10 natürlich sehr nahe (ist eben mit ihr identisch), aber andererseits gibt es auch immer wieder Folgenglieder, die von 10 weit entfernt sind! Man verlangt daher, dass die Abschätzung (*) *schließlich* in \mathbb{N} gilt. Darunter verstehen wir, dass sie *von einer Nummer ab* gilt:

‘schließlich’ = ‘von einer gewissen Nummer ab’

Dies bedeutet: Es gibt eine Zahl n_0 , von der ab die Abschätzung (*) *für alle größeren* $n \in \mathbb{N}$ gilt:

$$\begin{aligned} |a_n - a| < \varepsilon \text{ gilt schließlich in } \mathbb{N} \\ \iff \text{Es gibt ein } n_0 \in \mathbb{N} \text{ mit: } |a_n - a| < \varepsilon \text{ für alle } n \in \mathbb{N}, n \geq n_0. \end{aligned}$$

Es darf also nur *endlich viele* $n \in \mathbb{N}$ geben, für die (*) falsch ist. Man sagt dann, dass (*) für *fast alle* $n \in \mathbb{N}$ gilt:

‘für fast alle’ = ‘für alle bis auf endlich viele Ausnahmen’

(Im Bereich der natürlichen Zahlen bedeutet ‘für fast alle’ dasselbe wie ‘schließlich’, nicht jedoch im Bereich der reellen Zahlen.)

Damit haben wir beschrieben, was wir unter *in die Nähe kommen* verstehen wollen. Um jedoch zu erfassen, dass die Folgenglieder a_n dem Grenzwert a *beliebig nahe kommen*, müssen wir dies alles **für jedes** $\varepsilon > 0$ fordern! Die obigen Bedingungen nur für ein ε oder auch für endlich viele zu fordern, und seien sie auch noch so klein, reicht nicht aus, um den Konvergenzbegriff korrekt zu treffen. Entscheidend in der nachfolgenden Definition der Konvergenz ist daher die Forderung ‘für jedes $\varepsilon > 0$ ’! Damit kommen wir zu der folgenden fundamentalen

(2.1) Definition der Konvergenz: Eine Zahlfolge $(a_n)_{n \in \mathbb{N}}$ *konvergiert* gegen eine Zahl $a \in \mathbb{R}$, wenn für *jede* positive reelle Zahl $\varepsilon > 0$ die Abschätzung

$$|a_n - a| < \varepsilon \quad \text{bzw.} \quad a - \varepsilon < a_n < a + \varepsilon$$

in \mathbb{N} *schließlich* gilt, d. h. wenn sie von einer gewissen Nummer n_0 ab für *alle größeren* $n \in \mathbb{N}$ gilt. Ist dies der Fall, so schreibt man auch $a_n \rightarrow a$ ($n \rightarrow \infty$) (lesen Sie: a_n *strebt gegen* a für n gegen Unendlich) und nennt a auch *Grenzwert* der Folge (a_n) . Man benutzt dafür das Symbol

$$a = \lim_{n \rightarrow \infty} a_n.$$

(Lesen Sie: ‘Limes a_n für n gegen Unendlich’). Dieses Symbol ist *nur definiert, wenn die Folge (a_n) konvergiert!*

Auf der Basis dieser Definition kann man nun Konvergenznachweise führen, vorausgesetzt man hat einen ‘Kandidaten’ a für den Grenzwert. Man muss dann für beliebiges $\varepsilon > 0$ die Ungleichung $|a_n - a| < \varepsilon$ bzgl. $n \in \mathbb{N}$ untersuchen. Anders als bei den Monotonie- und Beschränktheitsnachweisen muss man hier aber nicht zeigen, dass diese Ungleichung allgemeingültig, die Lösungsmenge also ganz \mathbb{N} ist. Es genügt vielmehr der Nachweis, dass von einer geeigneten (von ε abhängigen) Zahl n_0 ab alle weiteren Zahlen $n \geq n_0$ zur Lösungsmenge gehören, d. h. dass die Abschätzung *schließlich* gilt.

Wir führen dies einmal für die Folge $a_n = \frac{3n-5}{2n+7}$ durch. Wir zeigen, dass $a = \frac{3}{2}$ Grenzwert dieser Folge ist. Dazu muss man für ein *beliebiges* $\varepsilon > 0$ die folgende Ungleichung untersuchen:

$$\begin{aligned} |a_n - a| < \varepsilon &\iff \left| \frac{3n-5}{2n+7} - \frac{3}{2} \right| < \varepsilon \iff \frac{|(3n-5) \cdot 2 - 3 \cdot (2n+7)|}{|2(2n+7)|} < \varepsilon \\ &\iff |6n - 10 - 6n - 21| < \varepsilon(4n + 14) \iff 31 < 4\varepsilon n + 14\varepsilon \\ &\iff 31 - 14\varepsilon < 4\varepsilon n \iff \frac{31 - 14\varepsilon}{4\varepsilon} < n \end{aligned}$$

Die letzte Ungleichung ist in \mathbb{N} *schließlich* gültig, nämlich von der kleinsten natürlichen Zahl n_0 ab, die oberhalb der reellen Zahl $\frac{31-14\varepsilon}{4\varepsilon}$ liegt. Da dies für jedes positive ε gilt, ist der Konvergenznachweis erbracht.

Für $\varepsilon = 0,1$ z. B. ergibt sich die Bedingung $n > \frac{31-1,4}{0,4} = 74$; also unterscheiden sich von der Nummer 75 ab alle weiteren Folgenglieder a_n vom Grenzwert $a = \frac{3}{2}$ um weniger als 0,1.

Weitere Beispiele finden sich in den Übungen (3). Dieses Verfahren des Konvergenznachweises ist mühselig und setzt außerdem voraus, den richtigen Grenzwert vorher zu kennen. Wir wollen nun im folgenden einen systematischen Weg zum Grenzwert samt Konvergenznachweis entwickeln.

Bevor wir zu den wichtigen Grenzwertsätzen kommen, noch einige weitere einfache, aber nützliche Beobachtungen. Dabei verstehen wir unter einer *Nullfolge* eine Folge, die den Grenzwert 0 hat.

- (2.2) Bemerkung:** a) (a_n) ist genau dann eine Nullfolge, wenn $(|a_n|)$ eine Nullfolge ist.
 b) Eine Folge (a_n) konvergiert gegen a genau dann, wenn die Folge $(a_n - a)$ eine Nullfolge ist.
 c) Konvergente Folgen sind beschränkt.

a) und b) sind unmittelbar einsichtig. In a) ist in beiden Fällen dieselbe Ungleichung $|a_n| < \varepsilon$ zu untersuchen. Ebenso in b): Der Nachweis, dass $(a_n - a)$ eine Nullfolge ist, führt auf die Ungleichung $|(a_n - a) - 0| < \varepsilon$, und die Untersuchung der Konvergenz der Folge (a_n) gegen den Grenzwert a führt auf dieselbe Ungleichung $|a_n - a| < \varepsilon$. Also stimmen auch die Lösungsmengen überein. Umfasst also die eine einen Restabschnitt natürlicher Zahlen, so auch die andere – und umgekehrt. Damit ist b) bewiesen.

Für c) argumentiert man so: Ist die Folge (a_n) konvergent, so besitzt sie einen Grenzwert a . Für $\varepsilon = 1$ erhält man dann aufgrund der Konvergenzdefinition die Existenz einer Zahl $n_0 \in \mathbb{N}$ mit

$$a - 1 < a_n < a + 1 \text{ für } n \geq n_0.$$

Damit hat man eine obere ($a + 1$) und eine untere ($a - 1$) Schranke für fast alle Folgenglieder gefunden. Man muss nun nur noch die endlich vielen Ausnahmen a_1, \dots, a_{n_0-1} beachten. Man wählt daher S_o größer als $a + 1$ und größer als die endlich vielen Ausnahmen a_1, \dots, a_{n_0-1} . Dann gilt die Abschätzung $a_n < S_o$ ohne jede Ausnahme für *alle* $n \in \mathbb{N}$: S_o ist eine obere Schranke für die Folge (a_n) . Ebenso findet man eine untere Schranke S_u , indem man eine Zahl kleiner als $a - 1$ und kleiner als die endlich vielen Ausnahmen a_1, \dots, a_{n_0-1} wählt. (Da jeweils nur *endlich viele* Folgenglieder beachtet werden müssen, ist eine solche Wahl immer möglich.)

b. Grenzwertsätze. Die Konvergenznachweise auf der Basis der Definition sind auf die Dauer mühselig. Man stellt fest, dass man dabei immer dieselben Überlegungen anzustellen hat.

Es ist daher sinnvoll, die allgemeinen Gesetzmäßigkeiten dahinter zu erforschen. Dazu gehören die folgenden Grenzwertsätze, die es ermöglichen, aus bekannten Konvergenzaussagen neue zu gewinnen. Ein besonders einsichtiges Beispiel ist der

(2.3) Schachtelungssatz: Sind zwei Folgen (a_n) und (b_n) konvergent mit demselben Grenzwert a und ist (c_n) eine Folge mit $a_n \leq c_n \leq b_n$ für (fast) alle n , so ist auch (c_n) konvergent gegen denselben Grenzwert a .

Beweis: Ist $\varepsilon > 0$ beliebig vorgegeben, so liegen in der ε -Umgebung $]a - \varepsilon, a + \varepsilon[$ von a fast alle Folgenglieder a_n und fast alle Folgenglieder b_n . Dann gilt aber auch für fast alle $n \in \mathbb{N}$, dass a_n und b_n in diesem Intervall liegen¹⁾. Da c_n dazwischen liegt, muss c_n ebenfalls in dem Intervall $]a - \varepsilon, a + \varepsilon[$ liegen. Dies gilt für fast alle n , womit alles gezeigt ist.

Ein etwas mehr algebraischer Beweis: Gemäß der Definition müssen wir zeigen: Für jedes beliebige $\varepsilon > 0$ gilt $a - \varepsilon < c_n < a + \varepsilon$ für fast alle $n \in \mathbb{N}$.

Da (a_n) gegen a konvergiert, gibt es (gemäß derselben Definition) eine Zahl $n_0 \in \mathbb{N}$ mit $a - \varepsilon < a_n < a + \varepsilon$ für alle $n \geq n_0$. Wegen der Konvergenz von (b_n) gegen denselben Grenzwert gibt es auch eine (evtl. andere) Zahl $n_1 \in \mathbb{N}$ mit $a - \varepsilon < b_n < a + \varepsilon$ für alle $n \geq n_1$. Ist n_2 die größere der beiden Zahlen n_0 und n_1 , so gelten für alle $n \geq n_2$ beide Abschätzungen, insbesondere die folgenden Teilaussagen davon:

$$a - \varepsilon < a_n \quad \text{und} \quad b_n < a + \varepsilon \quad \text{für alle } n \geq n_2.$$

Wegen $a_n \leq c_n \leq b_n$ folgt dann natürlich

$$a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon, \quad \text{also } a - \varepsilon < c_n < a + \varepsilon \quad \text{für alle } n \geq n_2.$$

Damit ist die Konvergenz von (c_n) gegen a bewiesen.

(2.4) Grenzwertsätze: Die Folgen (a_n) und (b_n) seien konvergent mit Grenzwert a bzw. b . Dann gelten die folgenden Aussagen:

a) (Summe) Die Summen-/Differenzfolge $(a_n \pm b_n)$ ist ebenfalls konvergent mit dem Grenzwert $a \pm b$.

b) (Produkt) Die Produktfolge $(a_n b_n)$ ist ebenfalls konvergent mit dem Grenzwert ab .

c) (Quotient) Wenn alle Folgenglieder $b_n \neq 0$ sind und außerdem der Grenzwert $b \neq 0$ ist (!), dann ist auch die Quotientenfolge

$$\frac{a_n}{b_n} \quad \text{konvergent gegen den Grenzwert } \frac{a}{b}.$$

Man kann diese Aussagen unmittelbar auf der Basis der Definition beweisen. Wir wollen hier den Beweisgang in Teilschritte zerlegen, die einzeln genommen leichter zu beweisen sind, und die zusammen die obigen Grenzwertsätze ergeben. Dabei ist die Herleitung der Grenzwertsätze aus diesen Teilaussagen typisch für das Vorgehen bei derartigen Problemen. Überdies ist die nachfolgende Teilaussage 4. auch von eigenständigem Interesse, während die anderen Aussagen Spezialfälle der Grenzwertsätze sind. Wir zeigen der Reihe nach:

1. Summen von Nullfolgen sind Nullfolgen.
2. Konstante Vielfache $(c \cdot a_n)$ von Nullfolgen (a_n) sind wieder Nullfolgen.
3. Summensatz a).
4. Das Produkt aus einer Nullfolge und einer beschränkten Folge ist wieder eine Nullfolge.
5. Produktsatz b).
6. Konvergiert $b_n \neq 0$ gegen $b \neq 0$, so ist die Folge der Kehrwerte $(1/b_n)$ beschränkt.
7. Quotientensatz c).

Ad 1. Wir müssen die Ungleichung(skette)

$$-\varepsilon < a_n + b_n < \varepsilon \tag{1}$$

¹⁾ Hier ist eine kleine Schwierigkeit in der Formulierung 'fast alle' verborgen; siehe die später folgende Definition von n_2 .

untersuchen und zeigen, dass sie für fast alle $n \in \mathbb{N}$ gilt. Diese Abschätzung ist sicherlich dann erfüllt, wenn die beiden Abschätzungen

$$-\frac{\varepsilon}{2} < a_n < \frac{\varepsilon}{2} \quad \text{und} \quad -\frac{\varepsilon}{2} < b_n < \frac{\varepsilon}{2}$$

zugleich wahr sind (man addiere einfach). Da $\varepsilon/2$ positiv ist und (a_n) eine Nullfolge ist, gibt es eine Zahl $n_a \in \mathbb{N}$, so dass für $n \geq n_a$ die erste Ungleichung wahr ist. Genauso gibt es eine Zahl $n_b \in \mathbb{N}$, von der ab die zweite Ungleichung wahr ist. Von der größeren der beiden Zahlen n_a, n_b ab sind dann aber *beide Ungleichungen zusammen* wahr. Für fast alle n sind also *beide* wahr, woraus dann die gewünschte Abschätzung (1) folgt.

Ad 2. Man muss die Ungleichung $|ca_n| < \varepsilon$ untersuchen und zeigen, dass sie für fast alle $n \in \mathbb{N}$ erfüllt ist. Ist $c = 0$, so lautet die Ungleichung einfach $0 < \varepsilon$ und sie ist sogar für alle $n \in \mathbb{N}$ erfüllt. Ist $c \neq 0$, so gilt

$$|ca_n| < \varepsilon \iff |a_n| < \frac{\varepsilon}{|c|}.$$

Da (a_n) aber eine Nullfolge ist, ist die rechte Ungleichung für fast alle $n \in \mathbb{N}$ erfüllt. Gleiches gilt dann auch für die linke Ungleichung.

Ad 3. Da (a_n) gegen a konvergiert, ist $(a_n - a)$ eine Nullfolge; ebenso ist $(b_n - b)$ eine Nullfolge. Dann ist nach 2. auch $(a_n - a) + (b_n - b) = (a_n + b_n) - (a + b)$ eine Nullfolge. Das bedeutet aber (siehe (2.2) b)), dass $(a_n + b_n)$ gegen $a + b$ konvergiert. Die Aussage über die Differenz ergibt sich ebenfalls aus 2., denn $(a_n - b_n) - (a - b) = (a_n - a) + (b - b_n)$ ist ebenfalls eine Summe von Nullfolgen.

Ad 4. Ist (a_n) eine Nullfolge und (b_n) beschränkt, etwa $|b_n| \leq M$ für alle n , so gilt für alle n

$$0 \leq |a_n b_n| \leq |a_n| \cdot M.$$

Nun ist $(|a_n|)$ und dann gemäß 2. auch $(M \cdot |a_n|)$ eine Nullfolge. Nach dem Schachtelungssatz ist dann auch die zwischen die Nullfolgen 0 und $(M \cdot |a_n|)$ eingezwängte Folge $(|a_n b_n|)$ eine Nullfolge.

Ad 5. Wir müssen zeigen, dass die Folge $(a_n b_n - ab)$ eine Nullfolge ist. Es gilt

$$a_n b_n - ab = a_n b_n - ab_n + ab_n - ab = (a_n - a)b_n + a(b_n - b). \quad (2)$$

In der letzten Formel ist $(a_n - a)$ eine Nullfolge, (b_n) als konvergente Folge beschränkt, also nach 4. auch $((a_n - a)b_n)$ eine Nullfolge. Noch einfacher (mittels 2.) erkennt man, dass auch der zweite Bestandteil $a(b_n - b)$ eine Nullfolge ist. Nach 3. ist dann auch die Summenfolge eine Nullfolge, was zu beweisen war.

Ad 6. Gesucht ist eine obere Schranke S für $|1/b_n|$. S muss dann notwendig positiv sein. Für $S > 0$ gilt jedoch

$$\left| \frac{1}{b_n} \right| \leq S \iff \frac{1}{S} \leq |b_n|.$$

Wir müssen also für die Folge $|b_n|$ eine untere Schranke finden, *die positiv ist*. (Die offensichtliche untere Schranke 0 für $|b_n|$ nutzt nichts, weil dafür die obige Äquivalenz nicht gilt.)

Da $b \neq 0$ ist, gilt entweder $b > 0$ oder $b < 0$. Wir behandeln zunächst den Fall $b > 0$. Gemäß der Konvergenzdefinition folgt für jedes $\varepsilon > 0$:

$$b - \varepsilon < b_n < b + \varepsilon \text{ für fast alle } n \in \mathbb{N}.$$

Da $b > 0$ ist, ist auch $b - \varepsilon > 0$, wenn man nur ε klein genug wählt, etwa $\varepsilon = b/2$. Dann ist $b - \varepsilon = b/2 > 0$. Man erhält so die Abschätzung

$$b/2 < b_n \text{ für fast alle } n \in \mathbb{N}. \quad (3)$$

Damit haben wir gezeigt:

Hat eine Folge einen positiven Grenzwert b , so sind fast alle Folgenglieder größer als $b/2$, insbesondere also positiv.

Genauso folgt

Hat eine Folge einen negativen Grenzwert b , so sind fast alle Folgenglieder kleiner als $b/2$, insbesondere also negativ.

Diese zweite Aussage ergibt sich aus der ersten, indem man von der Folge (b_n) mit negativem Grenzwert $b < 0$ zur Folge $(-b_n)$ übergeht, die den positiven Grenzwert $-b$ hat. Folglich gilt nach der bereits bewiesenen ersten Aussage für fast alle n : $-b_n \geq -b/2$ bzw. äquivalent $b_n \leq b/2$ wie behauptet.

In beiden Fällen erhält man so:

$$|b_n| > \frac{|b|}{2} \text{ für fast alle } n.$$

Die endlich vielen Ausnahmen $|b_n|$ von dieser Abschätzung sind positive Zahlen ($b_n \neq 0$ nach Voraussetzung), die unterhalb von $|b|/2$ liegen. Wählt man unter ihnen die kleinste, so erhält man dadurch eine positive untere Schranke für *alle* $|b_n|$. Der Kehrwert davon ist dann die gesuchte obere Schranke für $(1/|b_n|)$.

Ad 7. Mit dem Beweis von 6. haben wir den entscheidenden Schritt für 7. bereits getan. Wegen des Grenzwertsatzes für Produkte genügt es zu zeigen, dass unter den angegebenen Voraussetzungen die Folge der Kehrwerte $1/b_n$ gegen $1/b$ konvergiert, d. h.

$$\frac{1}{b_n} - \frac{1}{b} \text{ ist eine Nullfolge.}$$

Gemäß den Bruchrechenregeln gilt

$$\frac{1}{b_n} - \frac{1}{b} = \frac{b - b_n}{bb_n} = (b - b_n) \cdot \frac{1}{b} \cdot \frac{1}{b_n}.$$

Nun ist $(b - b_n)$ nach Voraussetzung eine Nullfolge; nach Multiplikation mit der festen Zahl $1/b$ bleibt dies eine Nullfolge (siehe 2.) und nach Multiplikation mit der beschränkten Folge $1/b_n$ (siehe 6.) erhält man (siehe 3.) wiederum eine Nullfolge; was zu beweisen war.

c. Anwendungen der Grenzwertsätze. Mit Hilfe der Grenzwertsätze kann man viele Konvergenzuntersuchungen samt Bestimmung des Grenzwertes zurückführen auf die folgende einfache, aber dennoch wichtige Grundtatsache:

$$\boxed{\lim_{n \rightarrow \infty} \frac{1}{n} = 0.}$$

Wegen der fundamentalen Bedeutung dieser Aussage hier der kurze Beweis:

$$\left| \frac{1}{n} \right| < \varepsilon \iff \frac{1}{n} < \varepsilon \iff \frac{1}{\varepsilon} < n.$$

Man sieht, dass die letzte Ungleichung in \mathbb{N} *schließlich* gültig ist, nämlich von der kleinsten natürlichen Zahl n_0 ab, die größer ist als die reelle Zahl $\frac{1}{\varepsilon}$.

Wir wollen nun das typische Vorgehen an dem schon früher betrachteten Beispiel

$$a_n = \frac{3n - 5}{2n + 7}$$

erläutern. Insbesondere wollen wir hier zeigen, wie man auf den früheren Ansatz $a = 3/2$ für den Grenzwert *zwangsläufig* stößt, wenn man die Grenzwertsätze zur Verfügung hat.

Wir formen das Folgenglied um, indem wir mit n kürzen:

$$a_n = \frac{3n - 5}{2n + 7} = \frac{3 - \frac{5}{n}}{2 + \frac{7}{n}}.$$

Nun betrachtet man den Aufbau dieses Terms und untersucht schrittweise auf Konvergenz. Wie erwähnt, ist Folge $1/n$ eine Nullfolge. Aus den Grenzwertsätzen folgt, dass dann auch $\frac{5}{n} = 5 \cdot \frac{1}{n}$ eine Nullfolge ist und daher der Zähler $3 - \frac{5}{n}$ gegen 3 konvergiert. Genauso folgert man aus den Grenzwertsätzen, dass die Nennerfolge gegen 2 konvergiert. *Da der Grenzwert der Nennerfolge $\neq 0$ ist*, konvergiert die Quotientfolge gegen $3/2$, und der Konvergenzbeweis (samt Bestimmung des Grenzwertes) ist vollständig.

Dieses Verfahren kann man auch für kompliziertere Folgen anwenden (bei denen unser bisheriges Vorgehen nicht mehr zum Erfolg führen würde).

$$a_n = \frac{2n^3 - 4n + 1}{-n^3 + 3n^2 - n} = \frac{2 - \frac{4}{n^2} + \frac{1}{n^3}}{-1 + \frac{3}{n} - \frac{1}{n^2}} \text{ konvergiert gegen } \frac{2}{-1} = -2.$$

$$a_n = \frac{3n^4 + 3n - 2}{2n^5 + 3n^3 - 2n} = \frac{\frac{3}{n} + \frac{3}{n^4} - \frac{2}{n^5}}{2 + \frac{3}{n^2} - \frac{2}{n^4}} \text{ konvergiert gegen } \frac{0}{2} = 0.$$

Die nächste Folge hingegen ist *nicht* konvergent, vielmehr gilt

$$a_n = \frac{2n^5 + 3n^2 - 1}{3n^4 - 2n} = n \cdot \frac{2 + \frac{3}{n^3} - \frac{1}{n^5}}{3 - \frac{2}{n^3}} \text{ wächst über jede Schranke.}$$

Man argumentiert dabei etwa folgendermaßen. Indem man in Zähler und Nenner die jeweils höchste Potenz von n ausklammert, erhält man die angegebene zweite Darstellung von a_n . Der zweite Faktor darin ist eine konvergente Folge mit positivem Grenzwert ($2/3$), also *schließlich* nach unten durch eine *positive* Zahl S beschränkt, etwa $S = 1/2$ (oder auch jede andere Zahl unterhalb $2/3$). (Siehe dazu Schritt 6. im Beweis der Grenzwertsätze). Damit gilt dann folgende Abschätzung für fast alle n :

$$a_n = n \cdot \frac{2 + \frac{3}{n^3} - \frac{1}{n^5}}{3 - \frac{2}{n^3}} \geq n \cdot S = \frac{n}{2}.$$

Dies bedeutet aber, dass a_n über alle Grenzen wächst, insbesondere also unbeschränkt und nicht konvergent ist.

Übung: Versuchen Sie aus diesen Beispielen eine allgemeine Gesetzmäßigkeit zu entnehmen und formulieren Sie diese.

Weitere Anwendungsmöglichkeiten für die Grenzwertsätze ist die Bestimmung von Grenzwerten rekursiv definierter Folgen (siehe am Ende des nachfolgenden Abschnitts e.) Dies setzt jedoch voraus, dass die Konvergenz anderweitig nachgewiesen wird – ohne Kenntnis des Grenzwertes!

d. Konvergenznachweis ohne Kenntnis des Grenzwertes. Unsere bisherigen Methoden für Konvergenznachweise versagen bei der noch ausstehenden Folge 10). Dies liegt nicht zuletzt daran, dass wir keinen Kandidaten für den möglichen Grenzwert haben. Dennoch schien klar, dass die Folge 10) konvergiert. Aber ein Beweis steht noch aus. Wir zeigen allgemein den folgenden wichtigen Satz, der mit der sog. Vollständigkeit des reellen Zahlbereichs \mathbb{R} gleichbedeutend ist (siehe den nachfolgenden Beweis).

(2.5) Monotoniekriterium: *Eine monoton wachsende, nach oben beschränkte Folge reeller Zahlen besitzt einen Grenzwert in \mathbb{R} .*

(Entsprechendes gilt für monoton fallende Folgen. Formulieren Sie dies und folgern Sie es aus der hier formulierten Aussage.)

Beweis: Es sei also a_n eine monoton wachsende Folge und S eine obere Schranke für a_n . Also gilt für alle $n \in \mathbb{N}$

$$a_1 \leq a_2 \leq \dots \leq a_n \leq a_{n+1} \leq \dots \leq S.$$

Es liegen also alle Folgenglieder zwischen a_1 und S . Wir unterteilen nun das Intervall $I_1 = [a_1, S]$ in zwei Hälften $[a_1, S_1]$ und $[S_1, S]$. Nun sind zwei Fälle möglich:

- Entweder liegen alle Folgenglieder im unteren Intervall $[a_1, S_1]$,
- oder es liegt wenigstens ein Folgenglied, wegen der Monotonie dann aber auch alle nachfolgenden, im rechten Intervall $[S_1, S]$.

In einem der beiden Teilintervalle liegen also fast alle Folgenglieder; sei dieses halb so breite Intervall mit I_2 bezeichnet. Nun kann man diesen Prozess mit I_2 statt I_1 wiederholen. Man erhält so eine Folge von ineinander liegenden Intervallen $I_1 \supset I_2 \supset I_3 \supset \dots$, deren Breiten sich ständig halbieren und daher eine Nullfolge bilden. Dies ist nichts anderes als eine *Intervallschachtelung*, und gemäß der *Vollständigkeit* des reellen Zahlbereichs gibt es genau eine reelle Zahl, die zu allen diesen Intervallen gehört. Sei a diese Zahl.

Die so gefundene Zahl a ist Grenzwert der Folge. Zum Beweis sei $\varepsilon > 0$ beliebig vorgegeben. Da die Intervallbreiten eine Nullfolge bilden, gibt es ein Intervall I_k , dessen Breite kleiner als ε ist. In diesem Intervall liegen sowohl a als auch fast alle Folgenglieder a_n , also muss der Abstand zwischen a und a_n kleiner ε sein. Damit gilt für fast alle n die Abschätzung $|a - a_n| < \varepsilon$. Da dies für jedes ε gilt, ist a Grenzwert der Folge a_n .

Wann immer man Konvergenz nachweisen will, ohne dass der Grenzwert bekannt ist, benutzt man dieses Monotoniekriterium.

e. Anwendung: Das Heron'sche Verfahren. Für eine beliebige Zahl $c > 0$ definieren wir die folgende Rekursion

$$a_{n+1} = \frac{1}{2} \left(a_n + \frac{c}{a_n} \right). \quad (*)$$

Die Folge (10) ist ein Spezialfall davon mit $c = 3$. Wir wollen nun Folgen mit dieser Rekursion untersuchen. Das Folgenglied a_1 lassen wir dabei zunächst noch offen, wir verlangen nur, dass $a_1 > 0$ ist. (Ein möglicher Startwert wäre $a_1 = c$. Bei Folge (10) hatten wir $a_1 = 3/2$ gewählt.)

Für die Untersuchung der Konvergenz bei Folgen mit der Rekursion (*) können wir unsere früheren Methoden nicht anwenden: Das Vorgehen entsprechend der Definition (Abschnitt a.) ist nicht möglich, weil wir keinen Kandidaten für den Grenzwert haben, aber auch die Anwendung der Grenzwertsätze (Abschnitt c.) führt nicht zum Ziel, weil der Folgenterm nicht explizit vorliegt. Wir können ihn deshalb nicht *so umformen, dass er aus Teiltermen aufgebaut ist, deren Grenzwerte bekannt sind*. Wir werden daher hier unser Monotoniekriterium anwenden.

Eine Berechnung der ersten Folgenglieder der Folge (10) ergibt näherungsweise

$$a_2 \approx 1,75 ; a_3 \approx 1,732142857 ; a_4 \approx 1,732050810 ; a_5 \approx 1,732050807 .$$

Dies lässt vermuten, dass die Folge vom zweiten Glied ab monoton fällt! Wenn wir diese Vermutung beweisen und außerdem noch zeigen können, dass die Folge nach unten beschränkt ist, dann folgt aus dem Monotoniekriterium die Konvergenz der Folge. Wir führen dies nicht nur für die Folge (10), sondern allgemein für (*) mit beliebigem $c > 0$ und $a_1 > 0$ durch.

Beschränktheit: Wir beweisen, dass alle Folgenglieder positiv sind:

$$a_n > 0 \text{ für alle } n \in \mathbb{N} .$$

Aus der Rekursion

$$a_{n+1} = \frac{1}{2} \left(a_n + \frac{c}{a_n} \right)$$

entnimmt man aber sofort: Ist $a_n > 0$, so ist auch $c/a_n > 0$, also $a_n + c/a_n > 0$ und damit $a_{n+1} > 0$. Das bedeutet: Ist ein Folgenglied positiv, so auch das nächste! Da nach Voraussetzung $a_1 > 0$ ist, ist dann auch $a_2 > 0$, dann also auch $a_3 > 0$, usw. Mithin sind *alle* Folgenglieder positiv; 0 ist eine untere Schranke der Folge.

Monotonie: Um zu untersuchen, ob die Folge (a_n) monoton fällt, muss man die Abschätzung $a_{n+1} \leq a_n$ studieren:

$$\begin{aligned} a_{n+1} \leq a_n &\iff \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \leq a_n \quad | \cdot 2a_n (> 0, \text{ siehe oben!}) \\ &\iff a_n^2 + c \leq 2a_n^2 \iff c \leq a_n^2 \end{aligned} \quad (1)$$

Dies zeigt: Ist a_n^2 größer-gleich c , so ist das nächste Folgenglied a_{n+1} kleiner-gleich a_n . Wir müssen nun also untersuchen, ob bzw. wann $a_n^2 \geq c$ ist. Über a_1 wissen wir fast nichts, weil wir nicht viel vorausgesetzt haben. Aber über die folgenden Glieder haben wir aufgrund der Rekursion mehr Informationen. Wir werden also untersuchen, ob bzw. ab wann $a_{n+1}^2 \geq c$ gilt:

$$\begin{aligned} a_{n+1}^2 \geq c &\iff \frac{1}{4}\left(a_n + \frac{c}{a_n}\right)^2 \geq c \\ &\iff a_n^2 + 2c + \frac{c^2}{a_n^2} \geq 4c \\ &\iff a_n^2 - 2c + \frac{c^2}{a_n^2} \geq 0 \\ &\iff \left(a_n - \frac{c}{a_n}\right)^2 \geq 0. \end{aligned} \quad (2)$$

Da ein Quadrat immer ≥ 0 ist, ist die letzte Abschätzung immer wahr, also auch die erste: Für alle $n \in \mathbb{N}$ gilt $a_{n+1}^2 \geq c$. Wenn n alle natürlichen Zahlen durchläuft, dann durchläuft a_{n+1} alle Folgenglieder a_2, a_3, \dots vom *zweiten* ab. Gemäß (2) bedeutet dies: *Vom zweiten Glied ab ist die Folge (a_n) monoton fallend.* Dies entspricht genau unserer Beobachtung bei der Folge 10)!

Bestimmung des Grenzwertes: Wir haben damit bewiesen: Folgen mit der Rekursion (*) sind konvergent; sie besitzen also einen Grenzwert, den wir a nennen wollen:

$$\lim_{n \rightarrow \infty} a_n = a.$$

Dieser Grenzwert ist zunächst nicht näher bekannt. Jedoch können wir aus (2) entnehmen $a_{n+1}^2 \geq c$, also $|a_{n+1}| \geq \sqrt{c}$, und damit wegen $a_{n+1} > 0$ sogar $a_{n+1} \geq \sqrt{c}$. Damit sind fast alle Folgenglieder (nämlich alle vom zweiten ab) $\geq \sqrt{c}$, also muss auch der Grenzwert $a \geq \sqrt{c}$ sein! (Warum eigentlich?) Insbesondere gilt $a > 0$.

Wir wollen nun mit Hilfe der Grenzwertsätze zeigen, dass $a = \sqrt{c}$ gelten muss. Wegen $a_n \rightarrow a$ und $a \neq 0$ folgt aus den Grenzwertsätzen

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \text{ konvergiert gegen } \frac{1}{2}\left(a + \frac{c}{a}\right).$$

Nun ist aber die Folge $(a_{n+1})_{n \in \mathbb{N}}$ nichts anderes als die Folge a_2, a_3, a_4, \dots . Diese hat natürlich auch den Grenzwert a , also gilt

$$a = \frac{1}{2}\left(a + \frac{c}{a}\right).$$

Wir lösen diese Gleichung:

$$a = \frac{1}{2}\left(a + \frac{c}{a}\right) \iff 2a^2 = a^2 + c \iff a^2 = c \iff |a| = \sqrt{c}.$$

Als Grenzwert a kommen also nur die Zahlen $\pm\sqrt{c}$ in Frage; da $a > 0$ ist, folgt notwendig: $a = \sqrt{c}$. Wir haben damit bewiesen:

(2.6) Heron's Approximation für \sqrt{c} : Ist $c > 0$ eine beliebige Zahl und definieren wir die Folge (a_n) bei beliebigem $a_1 > 0$ rekursiv durch

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

so konvergiert diese Folge gegen \sqrt{c} .

Man hat so eine (sehr schnelle) Berechnungsmethode für Näherungswerte für Quadratwurzeln. Sie geht zurück auf den Mathematiker Heron von Alexandrien (ca. 60 n. Chr.) Die Konvergenz ist um so schneller, je näher der Startwert a_1 bei \sqrt{c} liegt.

Das obige Vorgehen zur Bestimmung des Grenzwertes kann man sehr oft bei rekursiv definierten Folgen anwenden. Es eignet sich aber nicht zum Konvergenznachweis, sondern nur zur Bestimmung des *möglichen* Grenzwertes – für den Fall, dass Konvergenz vorliegt. Diese muss man unabhängig davon nachweisen, etwa mit dem Monotoniekriterium. Vorsicht! Die Tatsache, dass man auf diese Weise einen Kandidaten für den Grenzwert gefunden hat, bedeutet nicht, dass Konvergenz vorliegt. [Untersuchen Sie beispielsweise die Rekursion $a_{n+1} = 1 - a_n + \frac{1}{a_n}$ mit verschiedenen Startwerten a_1 .]

II. Rationale Funktionen

§3 Ganz-rationale Funktionen und ihre Nullstellen

a. Polynomterme und ganz-rationale Funktionen. Wir erinnern an den Funktionsbegriff der Mathematik:

(3.1) Definition: Es sei $D \subset \mathbb{R}$ eine Teilmenge der reellen Zahlen \mathbb{R} . Eine *Funktion* $f : D \rightarrow \mathbb{R}$ (lesen Sie: ‘ f von D in \mathbb{R} ’) ist eine Zuordnung, die jeder Zahl $r \in D$ eine *eindeutig* bestimmte Zahl $f(r) \in \mathbb{R}$ zuordnet. Man nennt $f(r)$ (lies ‘ f von r ’) den *Funktionswert* von f an der *Stelle* r . D heißt *Definitionsbereich* von f , \mathbb{R} ist der *Zielbereich*.

Eine häufig genutzte Möglichkeit, eine Funktion f zu definieren, ist die Angabe eines *Funktionsterms* $f(x)$, d. h. eines sinnvollen ‘Rechenausdrucks’ mit *einer* Variablen x . Zum Beispiel:

$$f(x) = x^3 + 4x, \quad g(x) = \frac{1}{x} + x^7, \quad h(x) = x^2 + 4,2 \cdot x - \pi.$$

Ein solcher Term bestimmt eine Funktion durch die Vorschrift: Man setze eine reelle Zahl r in den Term ein, berechne den Wert und definiere dann das Ergebnis als Funktionswert $f(r)$ an der Stelle r . Die Funktion ist dann für alle die reellen Zahlen definiert, für die die Einsetzung in den Funktionsterm sinnvoll ist. So ist z. B. die Einsetzung $g(0)$ nicht definiert, da $\frac{1}{0}$ nicht definiert ist. Ist $f(x)$ ein Funktionsterm, so ist

$$D(f) = \{r \in \mathbb{R} \mid \text{Die Einsetzung } f(r) \text{ ist sinnvoll definiert.}\}$$

der *natürliche* oder *maximale* Definitionsbereich von f . In unseren Beispielen:

$$D(f) = D(h) = \mathbb{R}, \quad D(g) = \mathbb{R} \setminus \{0\}.$$

Ist f durch einen Funktionsterm $f(x)$ gegeben, so schreibt man für die Funktion f auch $x \mapsto f(x)$ (lesen Sie: ‘ x geht über in $f(x)$ ’ oder ‘ x wird abgebildet auf $f(x)$ ’).

Anmerkung: 1) Funktionen können durch sehr unterschiedliche Zuordnungsvorschriften definiert sein. Die entstehenden Funktionen sind jedoch *identisch*, wenn sie denselben Definitionsbereich D und an *allen* Stellen $r \in D$ *dieselben* Funktionswerte haben:

$$f = g \iff f(r) = g(r) \quad \text{für alle } r \in D(f) = D(g)$$

2) Der Funktionsbegriff in der Mathematik ist allgemeiner als in (3.1) formuliert. Wir beschränken uns hier auf die oben definierten *reellwertigen* Funktionen, die auf Teilen der reellen Zahlengeraden definiert sind.

3) In einem Funktionsterm kann die Variable x sogar fehlen: $f(x) = 3$. Dann wird dadurch eine *konstante* Funktion definiert, die jeder reellen Zahl denselben Wert 3 zuordnet.

4) Typische, nicht durch einen einzelnen Funktionsterm beschriebene Funktionen sind die ‘abschnittsweise’ definierten Funktionen. Z. B. die Funktion f gegeben durch $x \mapsto f(x)$ mit

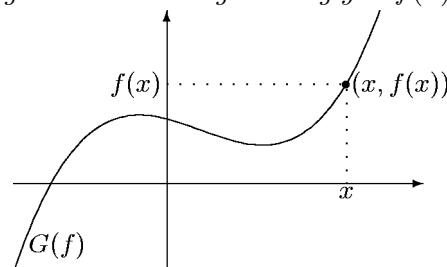
$$f(x) = \begin{cases} x^2 + x + 1 & \text{falls } -1 \leq x \leq 1, \\ 3x & \text{falls } x > 1, \\ -3x & \text{falls } x < -1. \end{cases}$$

Um sich eine Vorstellung von einer Funktion $f : D \rightarrow \mathbb{R}$ zu machen, stellt man sie durch ihren *Graphen* $G(f)$ in einem Koordinatenkreuz dar:

$$G(f) = \{ (x, y) \in \mathbb{R} \times \mathbb{R} \mid x \in D, y = f(x) \}.$$

Der Graph ist also nichts anderes als die *Lösungsmenge* der *Funktionsgleichung* $y = f(x)$.

Für jede reelle Zahl x im Definitionsbereich markiert man (siehe nebenstehende Skizze) im Koordinatenkreuz den Punkt $(x, f(x))$, dessen x -Koordinate x und dessen y -Koordinate der zugehörige Funktionswert $f(x)$ ist. Die Menge aller dieser Punkte in der x - y -Ebene ist der Graph von f .



Der Graph charakterisiert die Funktion vollständig, denn zwei Funktionen stimmen dann und nur dann überein, wenn ihre Graphen übereinstimmen:

$$f = g \iff G(f) = G(g).$$

Im Folgenden wollen wir uns zunächst mit der wichtigen Funktionsklasse der *ganz-rationalen* Funktionen auseinandersetzen. Das sind die Funktionen, die sich durch einen sog. *Polynomterm* definieren lassen:

(3.2) Definition: a) Ein *Polynomterm* (oder einfach ein *Polynom*) in der Variablen x ist ein Term der Form

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \quad (a_n \neq 0) \quad (*)$$

mit $n \in \mathbb{N}_0$, reellen Zahlen $a_n, \dots, a_0 \in \mathbb{R}$ und $a_n \neq 0$. Ein Polynomterm ist also eine Summe von Vielfachen von Potenzen der Variablen x .

b) Die Zahl a_k nennt man den *Koeffizienten* (Vorfaktor) von x^k ($k = 0, 1, 2, \dots, n$). Den im Polynomterm (*) auftretenden höchsten Exponenten n von x definiert man als den *Grad* des Polynomterms, $a_n x^n$ nennt man den *führenden Term* und a_n den *führenden Koeffizienten*. Ist der führende Koeffizient $a_n = 1$, so heißt der Polynomterm *normiert*.

c) Zusätzlich wollen wir noch den Term ‘0’ zu den Polynomtermen rechnen. Diesem ordnet man keinen Grad zu!

d) Unter einer *ganz-rationalen* Funktion versteht man eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die durch einen Polynomterm $f(x)$ als Funktionsterm definiert werden kann. Wir werden weiter unten sehen, dass es nur *einen* Polynomterm geben kann, der f definiert. Man kann daher die Begriffe *führender Term*, *führender Koeffizient* und *Grad* auf ganz-rationale Funktion f übertragen.

Wir bemerken: Polynomterme vom Grade 0 führen zu den *konstanten* Funktionen $f(x) = a_0$ mit $a_0 \neq 0$. Der Term ‘0’ liefert die sog. *Nullfunktion* $f(x) = 0$, d. i. die konstante Funktion mit dem Wert 0. Die nicht-konstanten ganz-rationalen Funktionen sind genau die vom Grad ≥ 1 .

b. Nullstellen und Linearfaktoren. Bitte beachten Sie, dass eine ganz-rationale Funktion auch durch ganz anders gestaltete Funktionsterme gegeben sein kann. So ist etwa durch den Funktionsterm

$$f(x) = (x - 1)(x + 2)(x - 3)$$

eine ganz-rationale Funktion f gegeben, da man diesen Term durch ‘Ausmultiplizieren’ auf die in (3.2) geforderte Form eines Polynomterms bringen kann. (Geben Sie für dieses Beispiel den Grad n und die Koeffizienten a_0, \dots, a_n an.)

Die nicht ausmultiplizierte, *faktorierte* Form des Funktionsterms hat aber den großen Vorteil, dass man aus ihr unmittelbar die Nullstellen der Funktion f ablesen kann. Da ein Produkt nur dann 0 sein kann, wenn einer der Faktoren 0 ist, gilt:

$$\begin{aligned} f(x) = 0 &\iff (x-1)(x+2)(x-3) = 0 \\ &\iff x-1 = 0 \vee x+2 = 0 \vee x-3 = 0 \\ &\iff x = 1 \vee x = -2 \vee x = 3 \end{aligned}$$

Ist also eine ganzrationale Funktion f nicht in dieser Form gegeben, sondern etwa in der Standardform als Polynomterm

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

so muss man versuchen, den Funktionsterm in die Form

$$f(x) = (x-x_1) \cdot (x-x_2) \cdot \dots \cdot (x-x_n) \quad (*)$$

zu bringen. Man versucht also, $f(x)$ in Linearfaktoren zu zerlegen. Das entscheidende Hilfsmittel dabei ist die

Polynomdivision: Wir gehen aus von zwei Polynomtermen, keiner davon der Term ‘0’, etwa

$$f(x) = 3x^5 + 4x^4 - x^2 + 3x - 1 \quad \text{und} \quad g(x) = x^2 + 2x - 3.$$

Wie von Polynomtermen gefordert, sind sie nach fallenden x -Potenzen sortiert. Wir wollen versuchen, $f(x)$ durch $g(x)$ zu dividieren, und gehen dabei ähnlich vor wie bei der schriftlichen Division mehrstelliger natürlicher Zahlen. Wir beachten zunächst nur die *führenden* Terme in $f(x)$ bzw. $g(x)$ und dividieren diese:

$$3x^5 : x^2 = 3x^3.$$

Dann multiplizieren wir den Divisor $g(x)$ mit $3x^3$ und subtrahieren dies vom Dividenten $f(x)$:

$$\begin{array}{r} (3x^5 + 4x^4 \quad -x^2 + 3x - 1) : (x^2 + 2x - 3) = 3x^3 + \dots \\ -(3x^5 + 6x^4 - 9x^3) \\ \hline -2x^4 + 9x^3 - x^2 + 3x - 1 \end{array}$$

Mit dem entstehenden Restpolynom (wir nennen es $f_1(x)$)

$$f_1(x) = f(x) - 3x^3 \cdot g(x) = -2x^4 + 9x^3 - x^2 + 3x - 1$$

arbeitet man nun an Stelle von $f(x)$ weiter: Wieder dividiert man den führenden Term (jetzt $-2x^4$) durch den führenden Term x^2 von $g(x)$ und erhält als Ergebnis $-2x^2$. Damit multipliziert man den Divisor $g(x)$ und subtrahiert das Produkt von $f_1(x)$:

$$\begin{array}{r} (3x^5 + 4x^4 \quad -x^2 + 3x - 1) : (x^2 + 2x - 3) = 3x^3 - 2x^2 + \dots \\ -(3x^5 + 6x^4 - 9x^3) \\ \hline -2x^4 + 9x^3 - x^2 + 3x - 1 \\ -(-2x^4 - 4x^3 + 6x^2) \\ \hline 13x^3 - 7x^2 + 3x - 1 \end{array}$$

So entsteht das nächste Polynom

$$f_2(x) = f_1(x) - (-2x^2) \cdot g(x) = f(x) - (3x^3 - 2x^2) \cdot g(x) = 13x^3 - 7x^2 + 3x - 1.$$

Man verfährt nun weiter wie oben beschrieben bis

$$\begin{array}{r}
 (3x^5 + 4x^4 \quad -x^2 + 3x - 1) : (x^2 + 2x - 3) = 3x^3 - 2x^2 + 13x - 33 + \dots \\
 -(\underline{3x^5 + 6x^4 - 9x^3}) \\
 \quad -2x^4 + 9x^3 \quad -x^2 + 3x - 1 \\
 -(\underline{-2x^4 - 4x^3 + 6x^2}) \\
 \quad \quad 13x^3 - 7x^2 + 3x - 1 \\
 -(\underline{13x^3 + 26x^2 - 39x}) \\
 \quad \quad \quad -33x^2 + 42x - 1 \\
 -(\underline{-33x^2 - 66x + 99}) \\
 \quad \quad \quad \quad 108x - 100
 \end{array}$$

An dieser Stelle kann man das Verfahren nicht mehr weiterführen, da das *Restpolynom*

$$r(x) = 108x - 100$$

als höchste x -Potenz nur x^1 enthält, die nicht mehr durch x^2 teilbar ist:

- Der Grad des Restpolynoms $r(x)$ ist *kleiner* als der Grad des Divisors $g(x)$.

Dieses ‘Restpolynom’ $r(x)$ ist entstanden, indem wir sukzessive gewisse Vielfache des Divisors $g(x)$ von $f(x)$ subtrahiert haben, nämlich

$$r(x) = f(x) - (3x^3 - 2x^2 + 13x - 33) \cdot g(x),$$

wobei der Faktor vor $g(x)$ (wir wollen ihn $q(x)$ nennen)

$$q(x) = 3x^3 - 2x^2 + 13x - 33$$

der bei der obigen Rechnung gefundene ‘Quotient’ ist. Unsere Polynomdivision hat also zu den beiden Polynomtermen $f(x)$ und $g(x)$ zwei Polynomterme $q(x)$ und $r(x)$ geliefert mit den folgenden beiden Eigenschaften:

$$f(x) = q(x) \cdot g(x) + r(x) \quad (*)$$

und

$$\text{Grad von } r(x) < \text{Grad von } g(x). \quad (**)$$

Dieses Verfahren der Polynomdivision mit Rest ist allgemein anwendbar. Zwar war in obigem Beispiel der Polynomterm $g(x)$ ‘normiert’, d. h. der führende Koeffizient von $g(x)$ war 1. Dies hat die Rechnung erleichtert (es traten keine Brüche auf), aber das Verfahren ist auch anwendbar, wenn etwa der führende Term in $g(x)$ die Form $2x^2$ o. ä. gehabt hätte. Man hätte dann immer durch diesen Term (einschließlich des Koeffizienten) dividieren müssen. Das Verfahren ist also auf alle Polynomterme $g(x)$ der allgemeinen Form

$$g(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

mit $a_n \neq 0$ anwendbar. Damit ist für $g(x)$ lediglich das Nullpolynom ‘0’ ausgeschlossen. Wir fassen nun die Überlegungen zusammen in dem folgenden

(3.3) Satz: (Polynomdivision) *Zu je zwei Polynomtermen $f(x)$ und $g(x)$, $g(x)$ sei nicht der Term ‘0’, gibt es Polynome $q(x)$ und $r(x)$ mit den Eigenschaften*

$$f(x) = q(x) \cdot g(x) + r(x) \quad (*)$$

und

$$r(x) \text{ ist der Term } 0 \text{ oder Grad von } r(x) < \text{Grad von } g(x). \quad (**)$$

Man berechnet diese Polynome $q(x)$ und $r(x)$ mit Hilfe des oben beschriebenen Verfahrens.

[Es sei angemerkt, dass die Polynome $q(x)$ und $r(x)$ durch die beiden Forderungen (*) und (**) *eindeutig* bestimmt sind.]

Wir benutzen nun diese Polynomdivision, um zu zeigen:

(3.4) Satz: Gegeben sei eine ganzrationale Funktion f mit dem Funktionsterm $f(x)$. Ist a eine Nullstelle von f , d. h. $f(a) = 0$, so kann man aus dem Polynomterm $f(x)$ den Linearfaktor $(x - a)$ multiplikativ abspalten. Präziser: Es gibt einen Polynomterm $h(x)$ mit

$$f(x) = (x - a) \cdot h(x).$$

Zum *Beweis* dieses Satzes und zur *Berechnung* von $h(x)$ dividiert man $f(x)$ durch $(x - a)$ gemäß (3.3). Dies ergibt

$$f(x) = (x - a) \cdot h(x) + r(x),$$

wobei das Restpolynom $r(x)$ entweder 0 ist oder einen Grad < 1 (=Grad von $(x - a)$), also den Grad 0 hat. Folglich ist $r(x)$ eine Konstante c_0 und damit

$$f(x) = (x - a) \cdot h(x) + c_0.$$

Setzt man nun in diese Termgleichung a ein, so folgt

$$0 = f(a) = 0 \cdot h(a) + c_0 = c_0.$$

Die Konstante c_0 ist also 0 und es folgt wie behauptet

$$f(x) = (x - a) \cdot h(x).$$

Hat man eine ganz-rationale Funktion f in dieser Weise zerlegt, so ist durch $h(x)$ wieder eine ganz-rationale Funktion gegeben, deren Grad um 1 kleiner ist als der von f . Ist nun a zufällig auch Nullstelle von h , so kann man wieder gemäß (3.4) den Linearfaktor $(x - a)$ abspalten. Wiederholt man dies solange wie möglich, so erhält man schließlich

$$f(x) = (x - a)^k \cdot g(x) \quad \text{mit } k \in \mathbb{N}, g(a) \neq 0. \quad (+)$$

Diese Zerlegung ist eindeutig; k gibt an, wie oft man den Linearfaktor $(x - a)$ aus dem Term $f(x)$ abspalten kann. Man nennt k die Ordnung der Nullstelle a von f :

(3.5) Definition: Ist a Nullstelle einer ganz-rationalen Funktion f , so nennt man die Zahl k in (+) die *Ordnung* oder *Vielfachheit* der Nullstelle a von f .

Ist eine Zerlegung der Form (+) gegeben, so ist der Grad der ganz-rationalen Funktion g um k kleiner als der von f . Außerdem sind die Nullstellen von f gerade gegeben durch a und alle Nullstellen von g . Damit ist die Nullstellenberechnung von f auf die von g zurückgeführt, und wegen des geringeren Grades von g im Prinzip vereinfacht. Nun kann man die Abspaltung von Nullstellen für das Polynom $g(x)$ statt $f(x)$ erneut durchführen. Dies kann man solange wiederholen, bis das entstandene Polynom $g(x)$ überhaupt keine Nullstellen mehr hat:

(3.6) Satz: a) Hat eine ganz-rationale Funktion f vom Grade n die verschiedenen Nullstellen x_1, x_2, \dots, x_m mit den Ordnungen k_1, \dots, k_m , so kann der Funktionsterm $f(x)$ dargestellt werden als

$$f(x) = (x - x_1)^{k_1} \cdot (x - x_2)^{k_2} \cdot \dots \cdot (x - x_m)^{k_m} \cdot g(x),$$

wobei $g(x)$ eine ganz-rationale Funktion g beschreibt, die bei x_1, \dots, x_m keine Nullstelle mehr hat. g hat den Grad $n - (k_1 + \dots + k_m)$.

b) Hat g überhaupt keine Nullstellen mehr auf \mathbb{R} , so sind x_1, \dots, x_m sämtliche Nullstellen von f – und umgekehrt.

c) Durch Vergleich der Grade ergibt sich:

Die Summe der Nullstellenordnungen einer ganz-rationalen Funktion ist höchstens gleich dem Grad n . Insbesondere hat eine ganz-rationale Funktion f vom Grad n höchstens n verschiedene

Nullstellen.

Daraus ergibt sich dann auch, dass verschiedene Polynomterme auch verschiedene ganz-rationale Funktionen definieren.

Nullstellenfreie ganz-rationale Funktionen sind z. B. die konstanten Funktionen: $g(x) = c$, $c \neq 0$, aber auch Funktionen höheren Grades können nullstellenfrei sein, etwa $g(x) = x^2 + 1$ o. ä. Ist $g(x) = c$ konstant, so erhält man für f eine Beschreibung des Funktionsterms in der Gestalt

$$f(x) = c \cdot (x - x_1)^{k_1} \cdot (x - x_2)^{k_2} \cdot \dots \cdot (x - x_m)^{k_m}$$

mit den verschiedenen Nullstellen x_1, x_2, \dots, x_m . Dabei ist der Faktor c offensichtlich gerade der höchste Koeffizient a_n von $f(x)$.

c. Rationale Nullstellen ganz-rationaler Funktionen. Die bisherigen Überlegungen hängen entscheidend davon ab, dass man zu gegebenem Polynomterm $f(x)$ eine Nullstelle findet. Nun kennen Sie für quadratische Polynomterme $x^2 + px + q$ ein Verfahren zur Berechnung der Nullstellen, die sog. p, q -Formel. Für Polynomterme höheren Grades gibt es eine solche Auflösungsformel i. a. *nicht*. Genauer gilt: Für Grad 3 und 4 gibt es zwar (komplizierte) Auflösungsformeln, ist jedoch der Grad ≥ 5 , so gibt es *nachweislich* keine allgemeingültigen Auflösungsformeln.

Beschränkt man sich jedoch auf ganz-rationale Funktionen, deren Koeffizienten rationale Zahlen sind — wie dies in unserem Unterricht fast ausschließlich der Fall sein wird —, so kann man zumindest *die* Nullstellen bestimmen, die *rationale* Zahlen sind. Sind die Koeffizienten a_i in dem Polynomterm

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

rationale Zahlen, so kann man durch Multiplikation mit dem Hauptnenner die Koeffizienten ganzzahlig machen, ohne dass sich dabei die Nullstellen ändern. Wir werden also im folgenden nur Polynomterme mit Koeffizienten $a_0, \dots, a_n \in \mathbb{Z}$ betrachten! Dann gilt der folgende

(3.7) Satz: (Rationale Nullstellen ganz-rationaler Funktionen) Gegeben sei ein Polynomterm

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

mit Koeffizienten $a_0, \dots, a_n \in \mathbb{Z} (!)$. Es sei $\frac{b}{c} \in \mathbb{Q}$ eine rationale Nullstelle von f mit zueinander teilerfremden Zähler $b \in \mathbb{Z}$ und Nenner $c \in \mathbb{N}$. Dann gilt:

$$b \text{ teilt } a_0 \quad \text{und} \quad c \text{ teilt } a_n.$$

Als rationale Nullstellen von f kommen also nur Zahlen in Frage, deren Zähler ein Teiler von a_0 und deren Nenner ein Teiler von a_n ist! Diese kann man bestimmen, und dann durch Einsetzen feststellen, welche davon Nullstellen sind. Hat man eine Nullstelle gefunden, so spaltet man den entsprechenden Linearfaktor aus $f(x)$ ab und arbeitet mit dem verbleibenden Polynomterm kleineren Grades weiter.

Als Spezialfälle seien erwähnt:

- 1) Als ganzzahlige Nullstellen b (Nenner $c = 1!$) kommen nur Teiler von a_0 in Frage!
- 2) Ist $a_n = 1$, so muss der Nenner $c = 1$ sein; mögliche rationale Nullstellen b/c sind also notwendig ganzzahlig.

Zum *Beweis* von (3.7): Es sei — wie angegeben — $0 = f(\frac{b}{c})$, also

$$0 = a_n \frac{b^n}{c^n} + a_{n-1} \frac{b^{n-1}}{c^{n-1}} + \dots + a_1 \frac{b}{c} + a_0.$$

Nach Multiplikation mit c^n erhält man:

$$0 = a_n b^n + a_{n-1} b^{n-1} c + a_{n-2} b^{n-2} c^2 + \dots + a_1 b c^{n-1} + a_0 c^n. \quad (*)$$

Löst man dies nach $-a_0c^n$ auf, so erhält man:

$$\begin{aligned} -a_0c^n &= a_nb^n + a_{n-1}b^{n-1}c + \dots + a_1bc^{n-1} \\ &= b \cdot (a_nb^{n-1} + a_{n-1}b^{n-2}c + \dots + a_1c^{n-1}). \end{aligned}$$

Da der Ausdruck in Klammern eine ganze Zahl ist (die Koeffizienten a_i liegen in \mathbb{Z} (!) und b, c ebenfalls), ist b ein Teiler von $-a_0c^n$. Da b und c aber teilerfremd sind, muss b ein Teiler von a_0 sein.

Löst man nun (*) nach $-a_nb^n$ auf, so erhält man entsprechend aus

$$a_nb^n = c(a_{n-1}b^{n-1} + a_{n-2}b^{n-2}c + \dots + a_1bc^{n-2} + a_0c^{n-1}),$$

dass c ein Teiler von $-a_nb^n$ sein muss. Wie oben folgt wieder: c ist Teiler von a_n . Damit ist Satz (3.7) bewiesen.

d. Zerlegung in Linearfaktoren. Wir gehen aus von einem Polynomterm $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$. Gesucht sind alle Nullstellen und damit verbunden eine möglichst weitgehende Zerlegung des Terms $f(x)$ in Linearfaktoren. Beachten Sie, dass jede *Faktorisierung* zugleich eine *Aufspaltung* des Problems in zwei Teilprobleme liefert, auf die man alle Methoden erneut anwenden kann. Wir wollen nun alle uns zur Verfügung stehenden Mittel zusammenstellen:

- 1) Faktorisieren durch *Ausklammern* oder *binomische Formeln*.
- 2) Für *quadratische* Gleichungen wende man die p, q -Formel oder quadratische Ergänzung zur Berechnung aller Nullstellen an. Der Satz von Vieta liefert auch eine Faktorisierung: Sind x_1, x_2 sämtliche Lösungen von $x^2 + px + q = 0$, so gilt $x^2 + px + q = (x - x_1)(x - x_2)$ (siehe Einführungsskript, S. 52).
- 3) Für Gleichungen höheren Grades stelle man zunächst fest, ob *Substitution* möglich ist. Dies ist dann der Fall, wenn *alle* im Term $f(x)$ auftretenden *Exponenten* von x Vielfache einer festen Zahl $k \geq 2$ sind, wie etwa in $x^6 + 2x^3 + 5 = 0$ ($k = 3$) oder $3x^6 - 2x^4 - 3x^2 + 1 = 0$ ($k = 2$). Man ersetzt (*substituiert*) dann nämlich x^k durch eine neue Variable z und erhält so eine Gleichung niederen Grades in der neuen Variablen z ($z^2 + 2z + 5 = 0$ im ersten und $3z^3 - 2z^2 - 3z + 1 = 0$ im zweiten Fall). Diese löse man unter Verwendung aller zur Verfügung stehenden Hilfsmittel; dies liefert dann auch eine Faktorisierung der Terme mit der Variablen z . Am Schluss muss man dann noch die *Substitution rückgängig* machen, indem man für jede gefundene Lösung z die Gleichung $x^k = z$ (durch Wurzelziehen) löst bzw. in allen Faktorisierungen die Variable z wieder durch x^k ersetzt.
- 4) Man suche unter Beachtung von Satz (3.7) eine *rationale* Nullstelle a und spalte ggf. durch Polynomdivision den zugehörigen Linearfaktor $x - a$ ab.

Das hier skizzierte Verfahren bricht ab, wenn

- a) alle Nullstellen gefunden sind (!), oder
- b) eine Gleichung mindestens dritten Grades vorliegt, die
 - 1) nicht durch Substitution vereinfacht werden kann und
 - 2) keine *rationale* Nullstelle mehr hat.

e. Nullstellenordnung und Vorzeichenwechsel. Gelingt bei diesem Verfahren eine Bestimmung *aller* Nullstellen der ganzrationalen Funktion und damit eine *faktorierte* Form des Funktionsterms $f(x)$ wie in Satz (3.6), so kann man daraus nicht nur unmittelbar die Nullstellen der Funktion f ablesen, sondern auch die Vorzeichenverteilung für die Funktionswerte $f(x)$. Gehen wir einmal von folgendem Beispiel

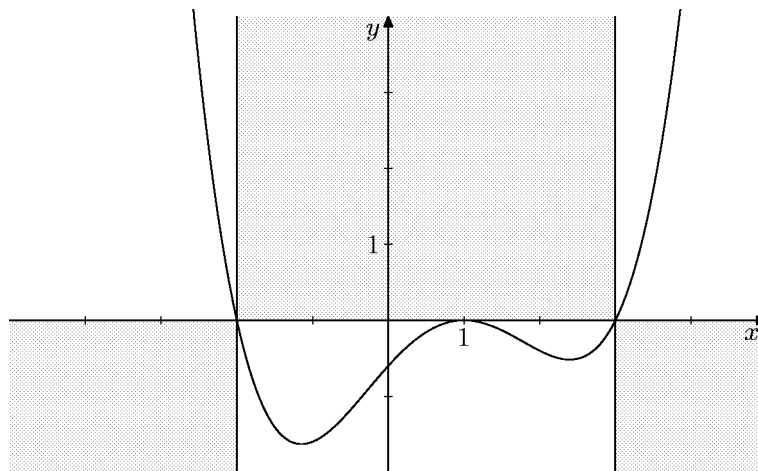
$$f(x) = (x - 1)^2(x + 2)(x - 3). \quad (*)$$

aus. Die Nullstellen von f sind offensichtlich $+1$ (von 2. Ordnung) und $-2, +3$ jeweils von erster Ordnung. Das Vorzeichen von $f(x)$ wird nun durch das Vorzeichen der einzelnen Linearfaktoren bestimmt: Ist eine *gerade* Anzahl von Linearfaktoren negativ, so ist das Produkt (und damit der

Funktionswert $f(x)$ positiv; ist eine *ungerade* Anzahl von Linearfaktoren negativ, so ist auch der Funktionswert $f(x)$ negativ. Nun ist aber das Vorzeichen eines einzelnen Linearfaktors $x - a$ für verschiedene Werte von x leicht abzulesen:

$$x - a \begin{cases} > 0 & \text{für } x > a, \\ < 0 & \text{für } x < a. \end{cases}$$

So sind für reelle Zahlen x größer als alle auftretenden Nullstellen, d. h. für $x > 3$, alle Linearfaktoren in (*) positiv ($x + 2 > 5$, $x - 1 > 2$, $x - 3 > 0$), das Produkt also auch. Liegt x zwischen der zweiten und dritten Nullstelle, also $1 < x < 3$, so ist nur der letzte Linearfaktor $x - 3$ negativ, alle anderen sind positiv, also das Produkt $f(x)$ negativ. Genauso untersucht man die anderen Bereiche zwischen den Nullstellen. Auf diese Art und Weise erhält man einen ersten groben Überblick über den Verlauf der Funktion f und kann eine vorläufige Skizze des Graphen entwerfen. Der in der nachfolgenden Skizze angedeutete Graph ist ein *denkbarer* Verlauf, er kann in den Einzelheiten durchaus anders aussehen (siehe die spätere Diskussion in Kapitel II). Durch die obigen Überlegungen sind jedoch gewisse Bereiche der Koordinatenebene bestimmt, in denen der Graph *mit Sicherheit nicht* verlaufen kann. Diese sind in der nachfolgenden Skizze schraffiert.

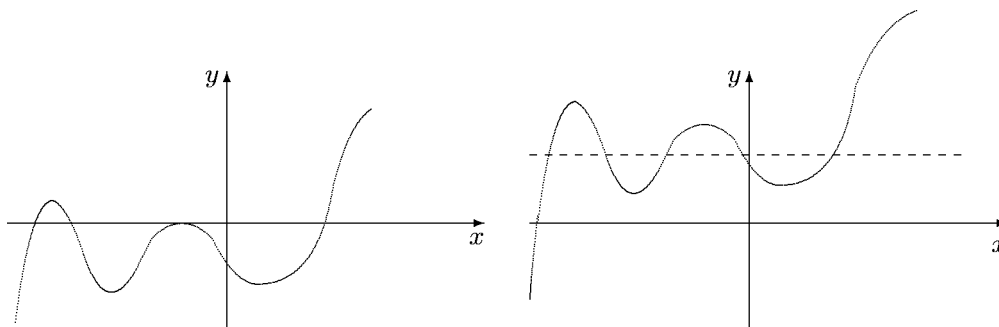


Wir wollen an diesem Beispiel auch das wichtige Phänomen des *Vorzeichenwechsels* an einer Nullstelle studieren. Die Funktionswerte $f(x)$ ändern offenbar bei den Nullstellen $x = -2$ und $x = +3$ ihr Vorzeichen, bei $+1$ hingegen nicht. Betrachten wir einmal die Nullstelle $+3$: Hier ändert der Linearfaktor $x - 3$ sein Vorzeichen, für $x < 3$ hat er negative Werte und für $x > 3$ positive. Da die anderen Linearfaktoren bei $+3$ ihr Vorzeichen nicht ändern, ergibt sich insgesamt ein Vorzeichenwechsel für $f(x)$. Anders bei $+1$: Hier ändert zwar der Linearfaktor $x - 1$ ebenfalls sein Vorzeichen, sein Quadrat $(x - 1)^2$ aber natürlich nicht. Es ergibt sich also *kein* Vorzeichenwechsel für $f(x)$. Wir erkennen so einen Zusammenhang zwischen Vorzeichenwechsel an einer Nullstelle a und der Häufigkeit, mit der der zugehörige Linearfaktor $x - a$ in der Zerlegung von $f(x)$ vorkommt. Ein Vorzeichenwechsel bei a liegt dann vor, wenn der Linearfaktor $x - a$ in *ungerader* Anzahl in der Produktzerlegung von $f(x)$ vorkommt. Kommt er dagegen in gerader Anzahl vor, so kann sich das Vorzeichen nicht ändern. Diese Häufigkeit eines Linearfaktors $x - a$ in der Zerlegung von $f(x)$ ist nun nichts anderes als die Ordnung der entsprechenden Nullstelle a . Wir halten fest:

(3.8) Satz: *Eine ganz-rationale Funktion f hat genau an den Stellen einen Vorzeichenwechsel, wo sie eine Nullstelle ungerader Ordnung hat.*

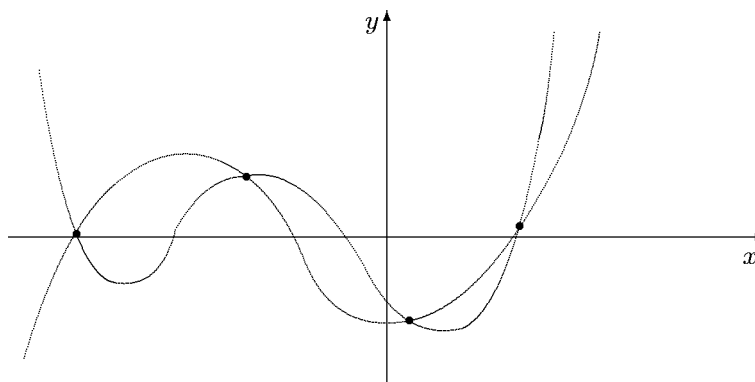
Wir haben uns dieses Ergebnis klar gemacht für ganz-rationale Funktionen, deren Funktionsterm $f(x)$ sich *vollständig* in Linearfaktoren zerlegen lässt. Der Satz gilt aber in der formulierten Allgemeinheit: Wir zerlegen $f(x) = (x - a)^k \cdot g(x)$, wobei k die Nullstellenordnung von f an der Stelle a ist. Das heißt $g(a) \neq 0$. Daraus werden wir in §4 folgern, dass die ganzrationale Funktion g ‘in der Nähe’ von a ihr Vorzeichen *nicht* ändert. Ob also f bei a einen Vorzeichenwechsel hat, liegt ausschließlich am Verhalten von $(x - a)^k$, und damit daran, ob k ungerade oder gerade ist.

f. Grad und Graph. Ist der Graph einer ganz-rationalen Funktion gegeben, so kann man aus der Zahl der Nullstellen entnehmen, wie groß der Grad mindestens sein muss. In dem linken Beispiel hat die Funktion f vier Nullstellen, von denen mindestens eine doppelt ist, da dort kein Vorzeichenwechsel vorliegt. Also muss (wieder nach Satz (3.6)) f mindestens den Grad $1+1+1+2=5$ haben.



Diese Überlegungen kann man auch auf den rechten Graphen anwenden, der nur eine Nullstelle hat. Verschiebt man den Graphen geeignet in y -Richtung, so erhält man 5 Nullstellen. Bei einer solchen Verschiebung verändert sich lediglich das absolute Glied a_0 des Funktionsterms, aber nicht der Grad. Man kann nun die obigen Überlegungen auf den verschobenen Graphen anwenden und so feststellen, dass der Grad mindestens 5 sein muss.

g. Grad und Schnittpunkte. Die Schnittstellen zweier Graphen $G(g)$ und $G(h)$ sind die Nullstellen der Differenzfunktion $f = g - h$. Sind g und h ganzrationale Funktionen,



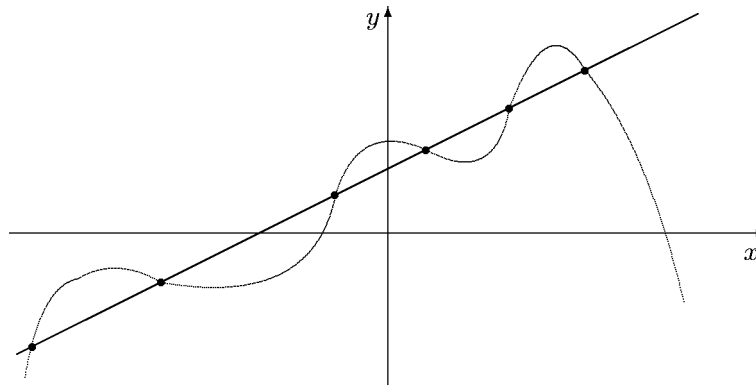
so ist f ebenfalls ganzrational und der Grad von f höchstens so groß ist, wie der größere der beiden Grade von g und h . Es gibt also höchstens so viele Lösungen, und damit Schnittpunkte, wie der größere der beiden Grade von g und h angibt! Die obige Skizze zeigt ein Beispiel mit den Graphen zweier ganz-rationaler Funktionen vom Grade 3 bzw. 4.

Betrachten wir nun einmal den Spezialfall, dass eine der Kurven eine Gerade ist, also zu einer ganzrationalen Funktion vom Grad ≤ 1 gehört. Dann ist die Anzahl der Schnittpunkte höchstens so groß wie der Grad der anderen Kurve. Wir erhalten so aus obigen Überlegungen:

Schneidet eine beliebige Gerade den Graphen einer ganz-rationalen Funktion g in $r \geq 2$ Punkten, so muss g mindestens den Grad r haben.

Dieses Kriterium verallgemeinert unsere Überlegungen aus f.: Dort hatten wir gesehen, dass der Grad mindestens so groß ist wie die Zahl der Nullstellen. Die Zahl der Nullstellen ist aber nichts anderes als die Anzahl der Schnittpunkte des Graphen der Funktion mit der x -Achse. In dem jetzt formulierten Kriterium kann man nun statt der x -Achse *jede beliebige Gerade* betrachten, die den Graphen in mindestens 2 Punkten schneidet: Die Anzahl der Schnittpunkte gibt dann die Mindestgröße des Grades an.

In dem folgenden Beispiel ist eine Gerade eingezeichnet, die den Graphen $G(g)$ in 6 Punkten schneidet. Also muss der Grad von g mindestens 6 sein.



§4 Nullstellen und Stetigkeit

a. Approximation von Nullstellen. Wir haben im vorangehenden Paragraphen einige Methoden zur exakten Lösung von Polynomgleichungen kennengelernt (Bestimmung der rationalen Nullstellen mittels Satz 1.7, Abspaltung des entsprechenden Linearfaktors durch Polynomdivision, Substitution in geeigneten Fällen). Diese Verfahren sind aber nicht immer anwendbar.

Dagegen ist die nun folgende Methode zwar immer anwendbar, aber rechenintensiv (man kommt kaum ohne Taschenrechner aus) und sie führt nur zur näherungsweise Bestimmung gewisser Nullstellen. Man geht davon aus, dass die Funktion an zwei Stellen $a_1 < b_1$ Werte mit unterschiedlichem Vorzeichen hat; etwa:

$$f(a_1) < 0 \text{ und } f(b_1) > 0.$$

(Solche Stellen kann man etwa bei der Erstellung einer Wertetabelle finden.) Man schließt dann, dass zwischen a_1 und b_1 eine Nullstelle liegen muss. *Dies ist richtig für ganzrationale Funktionen, aber nicht allgemein!* Wir wollen uns nun überlegen, warum dies so ist, und warum dies nicht in jedem Falle stimmt. Man versucht die vermutete Nullstelle zwischen a_1 und b_1 einzugrenzen, indem man den Funktionswert von f am Mittelpunkt c zwischen a_1 und b_1 ausrechnet. Es gibt drei Möglichkeiten:

1. $f(c) = 0$. Dann haben wir eine Nullstelle gefunden, und das Verfahren ist beendet.
2. $f(c) > 0$. Dann liegt zwischen a_1 und c ein Vorzeichenwechsel vor, also suchen wir in diesem Teilintervall (es ist halb so lang wie das Ausgangsintervall) weiter.
3. $f(c) < 0$. Dann liegt zwischen c und b_1 ein Vorzeichenwechsel vor und wir suchen in diesem Teilintervall weiter.

Mit dem neuen Teilintervall wiederholen wir unser Verfahren.

Wenn wir also nicht direkt auf eine Nullstelle stoßen, finden wir schließlich eine Folge ineinandergeschachtelter Intervalle

$$[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \dots \quad (1)$$

mit

$$f(a_n) < 0 \text{ und } f(b_n) > 0 \quad \text{für alle } n. \quad (2)$$

Wir betrachten nun die beiden ‘Eckpunkt’-Folgen (a_n) und (b_n) . Nach (1) sind beide Folgen beschränkt und (a_n) ist monoton wachsend, während (b_n) monoton fällt. Beide Folgen haben also einen Grenzwert (Monotoniekriterium). Da die Intervalllängen $l_n = b_n - a_n$ sich bei jedem Schritt halbieren, bilden sie eine Nullfolge: $a_n - b_n \rightarrow 0$. Damit müssen die Grenzwerte von (a_n) und (b_n) übereinstimmen. Nennen wir den gemeinsamen Grenzwert a :

$$\lim_{n \rightarrow \infty} a_n = a = \lim_{n \rightarrow \infty} b_n. \quad (3)$$

Wir behaupten nun, dass a eine Nullstelle von f ist, *vorausgesetzt die beiden folgenden Bedingungen sind erfüllt:*

A) Die Funktion f ist an der Stelle a definiert!

B) Die Folgen $f(a_n)$ und $f(b_n)$ der Funktionswerte konvergieren, und zwar gegen $f(a)$!

Beide Annahmen sind für ganzrationale Funktionen erfüllt: A) ist überhaupt kein Problem, weil ganzrationale Funktionen auf ganz \mathbb{R} definiert sind. B) ist eine Folge der Grenzwertsätze: Ist etwa $f(x) = x^3 - 2x + 1$, so ist $f(a_n) = a_n^3 - 2a_n + 1$, und wegen $a_n \rightarrow a$ konvergiert dies gemäß den Grenzwertsätzen gegen $a^3 - 2a + 1 = f(a)$. (Genauso argumentiert man für eine allgemeine ganzrationale Funktion $f(x) = \alpha_k x^k + \alpha_{k-1} x^{k-1} + \dots + \alpha_1 x + \alpha_0$.)

Mittels A) und B) schließen wir folgendermaßen weiter: Gemäß (2) ist $f(a_n) < 0$. Daher kann der Grenzwert dieser Folge (er existiert nach B) und ist gleich $f(a)$) nicht positiv sein: $f(a) \leq 0$. Indem man die Folge $f(b_n)$ betrachtet, folgt: $f(a) \geq 0$. Beides zusammen ergibt: $f(a) = 0$.

(4.1) Satz: *Ist f eine ganzrationale Funktion und sind a, b zwei Stellen, an denen f unterschiedliches Vorzeichen hat: $f(a) < 0$ und $f(b) > 0$, oder $f(a) > 0$ und $f(b) < 0$, so liegt zwischen a und b eine Nullstelle von f .*

Mittels des Halbierungsverfahrens kann man beliebig kleine Intervalle finden, in denen eine Nullstelle von f liegen muss. Man kann so Nullstellen von f beliebig genau approximieren.

Wir halten als wichtige Folgerung fest:

(4.2) Folgerung: *Hat eine ganzrationale Funktion keine Nullstelle, so haben alle ihre Werte dasselbe Vorzeichen! Ganzrationale Funktionen können nur an ihren Nullstellen ihr Vorzeichen wechseln!*

b. Stetigkeit. Die Gültigkeit des Satzes (4.1) ist nicht auf die ganzrationalen Funktionen beschränkt. Wir haben gesehen, dass es zum Beweis des Satzes (4.1) nur auf die Eigenschaften A) und B) ankommt. Wir haben im Unterricht aber auch an Beispielen gesehen, dass man weder auf die Forderung A) noch auf B) verzichten kann. (Bereits zur Formulierung von B) benötigt man A).)

Um die Eigenschaft A) zu sichern, müssen wir verlangen, dass die Funktion in dem Intervall $[a_1, b_1]$, mit dem wir starten, keine Definitionslücken hat; sie muss an *jeder* Zwischenstelle definiert sein. Um B) zu sichern, verlangen wir, dass die Funktion f für jede Stelle a in ihrem Definitionsbereich $D(f)$ die folgende Eigenschaft hat:

(S) *Für jede Folge $a_n \in D(f)$, die gegen a konvergiert, konvergiert die Folge $f(a_n)$ gegen $f(a)$.* Dies enthält *zwei* Forderungen: 1. Die Folgen $f(a_n)$ müssen konvergieren und 2. der Grenzwert muss gerade der Funktionswert $f(a)$ sein.

(4.3) Definition: a) Eine Funktion f heißt *stetig an einer Stelle a ihres Definitionsbereiches* genau dann, wenn die Bedingung S erfüllt ist.

b) Eine Funktion f heißt *stetig*, wenn sie an *allen* Stellen ihres Definitionsbereiches stetig ist.

Mit diesem Begriff können wir die Ergebnisse der Überlegungen in a. wie folgt formulieren:

(*) *Sind $a < b$ zwei Stellen, an denen eine Funktion f Werte mit unterschiedlichem Vorzeichen hat, und ist f über dem ganzen Intervall $[a, b]$ definiert und stetig, so liegt zwischen a und b eine Nullstelle von f .*

Man kann dieses Ergebnis sogar ausbauen zu dem folgenden wichtigen

(4.4) Satz: (Zwischenwertsatz) *Ist die Funktion f über einem Intervall $[a, b]$ definiert und dort stetig, so gilt: Jeder Wert d zwischen $f(a)$ und $f(b)$ ($f(a) \leq d \leq f(b)$ bzw. $f(a) \geq d \geq f(b)$) wird an einer Stelle $c \in [a, b]$ als Funktionswert angenommen:*

$$d = f(c) \text{ für ein (geeignetes) } c \in [a, b].$$

(4.4) lässt sich völlig analog wie (4.1) beweisen; man muss lediglich 0 durch d ersetzen. Man kann aber auch (4.4) unmittelbar aus (*) herleiten, ohne den gesamten Beweis zu modifizieren. Dazu geht man von der gegebenen Funktion f zur Funktion g über, die definiert ist durch $g(x) = f(x) - d$. Diese hat genau dort ihre Nullstellen, wo die Funktion f den Wert d annimmt. Außerdem erfüllt sie die Voraussetzungen von (*), denn aufgrund der Grenzwertsätze sind Summen, Differenzen, Produkte und Quotienten stetiger Funktionen wieder stetig! (Siehe Übungen (6) bzw. §5.) Aus (*) folgt dann die gewünschte Behauptung. (Führen Sie einen der beiden angedeuteten Beweise zur Übung selbständig durch!)

Anschaulich bedeutet dieser Zwischenwertsatz, dass es im Graphen von f keine Lücken gibt! Dies ist die graphisch anschauliche Beschreibung der Stetigkeit: *Über einem zusammenhängenden Stück des Definitionsbereiches ist der Graph einer stetigen Funktion ebenfalls zusammenhängend.*

Wir bemerken, dass Satz (4.1) ein Spezialfall des Zwischenwertsatzes (4.4) ist, indem man nämlich $d = 0$ wählt und beachtet, dass aufgrund der Grenzwertsätze alle ganzrationalen Funktionen stetig sind.

(4.5) Satz: a) *Summe, Differenz, Produkt und Quotient stetiger Funktionen sind stetig. Insbesondere:*

b) *Alle ganzrationalen Funktionen sowie deren Quotienten, die rationalen Funktionen (siehe §5), sind stetig.*

c) *Die Wurzelfunktion $f(x) = \sqrt{x}$ ist stetig.*

d) *Die Betragsfunktion $f(x) = |x|$ ist stetig.*

Beweis: a) Wir argumentieren nur für den Quotienten (die anderen Fälle sind noch einfacher): Es seien also g und h zwei Funktionen, die an einer Stelle a definiert und stetig sind. Wir wollen zeigen, dass dann auch die Quotientenfunktion $f = \frac{g}{h}$ an der Stelle a stetig ist, wenn sie dort definiert ist. Nun ist $f(x) = \frac{g(x)}{h(x)}$ gerade dort definiert, wo der Nenner $h(x)$ nicht den Wert 0 annimmt. Ist also a eine Stelle im Definitionsbereich von f , so gilt $h(a) \neq 0$.

Gemäß Definition der Stetigkeit müssen wir für eine beliebige Folge $x_n \in \mathcal{D}(f)$ mit Grenzwert a die Folge der Funktionswerte

$$f(x_n) = \frac{g(x_n)}{h(x_n)}$$

auf Konvergenz untersuchen. Da g bei a stetig ist und x_n gegen a konvergiert, muss $g(x_n)$ gegen $g(a)$ konvergieren. Entsprechend gilt auch $h(x_n) \rightarrow h(a)$:

$$g(x_n) \rightarrow g(a) \quad \text{und} \quad h(x_n) \rightarrow h(a). \quad (*)$$

Hierauf wenden wir nun den Grenzwertsatz für Quotienten an (siehe I Satz (4.4) c)). Wegen $x_n \in \mathcal{D}(f)$ gilt $h(x_n) \neq 0$ und wegen $a \in \mathcal{D}(f)$ entsprechend $h(a) \neq 0$:

$$h(x_n) \neq 0 \quad \text{und} \quad h(a) \neq 0. \quad (**)$$

Damit sind die Voraussetzungen des Grenzwertsatzes für Quotienten erfüllt und es folgt

$$f(x_n) = \frac{g(x_n)}{h(x_n)} \rightarrow \frac{g(a)}{h(a)} = f(a) \quad \text{für } n \rightarrow \infty.$$

Da diese Überlegungen für *alle* derartigen Folgen x_n gilt, ist die Funktion f an der Stelle $a \in \mathcal{D}(f)$ stetig.

b) Für die Funktion $f(x) = x$ ist die Stetigkeitsforderung ganz selbstverständlich erfüllt. (Wenn x_n gegen a konvergiert, konvergiert $f(x_n) = x_n$ eben gegen $f(a) = a$.)

Genauso selbstverständlich sind konstante Funktionen ($f(x) = c$) stetig.

Da die ganzrationalen Funktionen mittels Addition, Subtraktion und Multiplikation aus Konstanten und $f(x) = x$ aufgebaut werden, müssen sie nach a) stetig sein. Die rationalen Funktionen sind Quotienten ganzrationaler, und daher ebenfalls stetig (an jeder Stelle ihres Definitionsbereiches).

c) folgt aus dem in den Übungen bewiesenen Grenzwertsatz für Quadratwurzeln: Sei $f(x) = \sqrt{x}$ mit $\mathcal{D}_f = \mathbb{R}^{\geq 0} = [0, \infty[$. Sei $a \in \mathcal{D}_f$ und $x_n \in \mathcal{D}_f$ eine Folge mit $x_n \rightarrow a$. Wegen $x_n \in \mathcal{D}_f$ gilt $x_n \geq 0$. Dann besagt der Grenzwertsatz für Wurzeln:

$$x_n \geq 0, \quad x_n \rightarrow a \implies \sqrt{x_n} \rightarrow \sqrt{a}.$$

Also gilt $f(x_n) \rightarrow f(a)$ und die Stetigkeit von $f(x) = \sqrt{x}$ ist gezeigt.

d) Für $x > 0$ ist $f(x) = |x| = x$ identisch mit der ganzrationalen Funktion $g(x) = x$, also stetig. Entsprechend ist für $x < 0$ $f(x) = -x$ identisch mit der ganzrationalen Funktion $h(x) = -x$ und dort ebenfalls stetig. Man muss nun noch die Stetigkeit an der *Nahtstelle* $x = 0$ untersuchen. Dazu betrachten wir eine *beliebige* Folge a_n die gegen 0 konvergiert. Wir müssen zeigen, dass dann auch $f(a_n) = |a_n|$ gegen 0 konvergiert. Dies haben wir bereits früher gezeigt: Eine Folge a_n ist genau dann eine Nullfolge, wenn die Folge der Beträge $|a_n|$ gegen 0 konvergiert.

§5 Rationale Funktionen

a. Definitionslücken und Vorzeichenwechsel. In diesem Paragraphen wollen wir nun über die bisherige Funktionsklasse der ganz-rationalen Funktionen hinausgehen und die umfassendere Klasse der *rationalen Funktionen* kennenlernen und genauer studieren. So wie man die rationalen Zahlen definiert hat als die Quotienten ganzer Zahlen (mit von 0 verschiedenem Nenner), so definieren wir nun:

(5.1) Definition: Eine *rationale Funktion* f ist der Quotient zweier ganz-rationaler Funktionen, wobei die Nennerfunktion nicht die Nullfunktion ist:

$$f(x) = \frac{g(x)}{h(x)}.$$

Dabei sind $g(x)$ und $h(x)$ Funktionsterme beliebiger ganz-rationaler Funktionen, für $h(x)$ ist lediglich der Term '0' ausgeschlossen.

Anmerkungen:

1) Unsere bisherigen ganz-rationalen Funktionen sind natürlich auch rationale Funktionen im Sinne dieser Definition, da man als Nenner die konstante Funktion h vom Wert 1 ($h(x) = 1$) wählen kann.

2) Rationale Funktionen, die nicht ganz-rational sind, nennt man *gebrochen rational*.

3) Anders als die ganz-rationalen Funktionen sind rationale Funktionen im allgemeinen nicht auf ganz \mathbb{R} definiert, sie haben *Definitionslücken*. Diese sind gegeben durch die Nullstellen des Nenners $h(x)$: An Stellen $a \in \mathbb{R}$ mit $h(a) = 0$ ist $f(a) = \frac{g(a)}{h(a)}$ *nicht definiert*, die Funktion f besitzt dort eine Definitionslücke. Der Definitionsbereich einer rationalen Funktion f mit Nennerterm $h(x)$ besteht also aus allen reellen Zahlen mit Ausnahme der Nullstellen des Nenners:

$$D(f) = \mathbb{R} \setminus \{a \in \mathbb{R} \mid h(a) = 0\}.$$

f besitzt also nur endlich viele Definitionslücken; ihre Anzahl ist höchstens so groß wie der Grad des Nenners $h(x)$.

Betrachten wir einmal als Beispiel die rationale Funktion

$$f(x) = \frac{x+1}{x^2-4}.$$

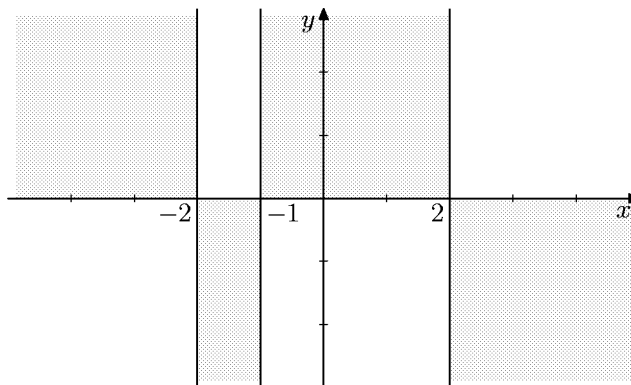
Die Definitionslücken sind leicht bestimmt, sie sind die Nullstellen des Nenners $x^2 - 4 = (x - 2)(x + 2)$, also ± 2 . Der Definitionsbereich von f ist also $D(f) = \mathbb{R} \setminus \{-2, +2\}$.

Wir wollen nun in ähnlicher Weise wie bei den ganzrationalen Funktionen die Vorzeichenverteilung bestimmen. Dies geht mit denselben Methoden, da auch hier der Funktionsterm aus Linearfaktoren aufgebaut ist:

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

Zwar kommen jetzt auch Linearfaktoren im Nenner vor, aber auf das Vorzeichen des Bruches haben diese denselben Einfluss wie Linearfaktoren im Zähler. Dies bedeutet, dass die Funktion f nicht nur an ihrer Nullstelle -1 das Vorzeichen wechselt, sondern auch an den beiden Definitionslücken ± 2 .

Da für x -Werte größer als $+2$ (d. h. größer als alle Nullstellen und alle Lücken) sämtliche Linearfaktoren positive Werte annehmen, hat $f(x)$ dort positives Vorzeichen. Damit können wir bereits große Bereiche der Koordinatenebene schraffieren, in denen der Graph von f *nicht* verlaufen kann:



Wie wir gesehen haben, können gebrochen-rationale Funktionen ihr Vorzeichen nicht nur an Nullstellen, sondern auch an Definitionslücken wechseln: Als mögliche Stellen für Vorzeichenwechsel kommen also die Nullstellen des Zählers *und* die Nullstellen *des Nenners* in Frage. Nun wechselt ein Quotient sein Vorzeichen aber nur dann, wenn der Zähler oder der Nenner, *aber nicht beide* (!) ihr Vorzeichen wechseln. Damit erhalten wir

(5.2) Satz: a) *Rationale Funktionen können ihr Vorzeichen nur an ihren Nullstellen oder ihren Definitionslücken wechseln.*

b) *Eine rationale Funktion f wechselt bei a ihr Vorzeichen genau dann, wenn dort der Zähler oder der Nenner, aber nicht beide, eine Nullstelle ungerader Ordnung haben.*

b. Grenzwerte von Funktionen. Aber dieses Ergebnis ermöglicht es noch nicht, einen einigermaßen realistischen Verlauf des Graphen von f zu skizzieren, weil — anders als bei ganz-rationalen Funktionen — der Graph von f wegen der Definitionslücken bei ± 2 dort nicht die x -Achse schneiden kann. Wir müssen also den Funktionsverlauf in der Nähe der Lücken genauer studieren. Das gleiche gilt für den Verlauf des Graphen ‘für $x \rightarrow \pm\infty$ ’, weil auch hier neue Phänomene auftreten.

Um das Verhalten in der Nähe von Lücken zu erfassen, untersuchen wir die Funktionswerte an Stellen $x \neq a$ ‘in der Nähe’ der Lücke a (in unserem obigen Beispiel $a = 2$ oder $a = -2$) und ‘lassen dann x immer näher auf a zulaufen’. Das bedeutet, wir untersuchen den *Grenzwert der Funktion bei a* . Darunter verstehen wir das Folgende: Wir betrachten beliebige Folgen x_n mit $x_n \neq a$, die gegen a konvergieren, und untersuchen die Funktionswertfolge $f(x_n)$ auf Konvergenz. Hat für *jede* derartige Folge x_n diese neue Folge $f(x_n)$ einen Grenzwert, und zwar immer *denselben*, etwa b , so sagen wir: Die Funktion f konvergiert für $x \rightarrow a$ gegen den Wert b :

(5.3) Definition: *Es sei f eine Funktion und $a \in \mathbb{R}$. a braucht nicht zum Definitionsbereich von f zu gehören, aber f muss in einer Umgebung von a definiert sein. Dann definieren wir: f konvergiert für $x \rightarrow a$ gegen b , wenn für jede Folge $x_n \in \mathcal{D}(f)$ mit $x_n \neq a$ und $x_n \rightarrow a$ die Folge $f(x_n)$ der Funktionswerte konvergiert, und zwar immer gegen denselben Grenzwert b :*

$$\lim_{n \rightarrow \infty} x_n = a \implies \lim_{n \rightarrow \infty} f(x_n) = b. \quad (*)$$

Man nennt dann b den Grenzwert von f an der Stelle a und bezeichnet ihn mit $\lim_{x \rightarrow a} f(x) = b$.

$$\lim_{x \rightarrow a} f(x) = b \iff \boxed{\text{Für jede Folge } x_n \in \mathcal{D}(f) \text{ mit } x_n \neq a \text{ und } x_n \rightarrow a \text{ gilt } f(x_n) \rightarrow b.}$$

Lässt man nur Folgen x_n zu mit $x_n > a$ (x_n konvergiert von oben gegen a), so spricht man vom *oberen oder rechtsseitigen Grenzwert* und bezeichnet ihn mit

$$\lim_{x \searrow a} f(x) \quad \text{oder} \quad \lim_{x \rightarrow a^+} f(x).$$

Entsprechend definiert man den unteren oder linksseitigen Grenzwert

$$\lim_{x \nearrow a} f(x) \quad \text{oder} \quad \lim_{x \rightarrow a^-} f(x).$$

Wir betonen nochmals, dass die Forderung (*) für jede Folge x_n gestellt wird, und der Grenzwert b der Funktionswertfolge $f(x_n)$ dabei immer derselbe sein muss. Erst dann können wir vom Grenzwert $\lim_{x \rightarrow a} f(x) = b$ sprechen.

Mit Hilfe des Grenzwertbegriffes für Funktionen kann man die Stetigkeit (siehe §4) folgendermaßen beschreiben:

(5.4) Bemerkung: Eine Funktion f ist an einer Stelle $a \in \mathbb{D}(f)$ stetig, wenn f bei a einen Grenzwert besitzt und dieser gerade $f(a)$ ist. In Kurzform:

$$f \text{ stetig bei } a \iff \lim_{x \rightarrow a} f(x) = f(a).$$

Noch stärker verkürzt bedeutet die Stetigkeit bei a :

Grenzwert bei a gleich Funktionswert bei a .

Wir weisen auch darauf hin, dass in Definition (5.3) bei den Folgen x_n der Wert a ausgeschlossen wird. Wenn a nicht zum Definitionsbereich von f gehört, ist dies automatisch der Fall. Aber auch wenn a zu $\mathbb{D}(f)$ gehört, wird dies verlangt. Um die Bedeutung dieser Bedingung zu erkennen, betrachten wir einmal die Funktion $f(x) = |\text{sign}(x)|$ aus den Übungen (6). Nach Definition gilt dann

$$f(x) = |\text{sign}(x)| = \begin{cases} +1 & \text{falls } x \neq 0, \\ 0 & \text{falls } x = 0. \end{cases}$$

Zur Untersuchung des Grenzwertes von f bei $a = 0$ müssen wir eine beliebige Folge x_n betrachten mit $x_n \neq a = 0$ und $x_n \rightarrow a = 0$. Dann gilt (wegen $x_n \neq 0!$) stets $f(x_n) = +1$ und der Grenzwert dieser Folge $f(x_n)$ ist dann natürlich ebenfalls $+1$. Da dies für jede derartige Folge x_n gilt, erhält man:

f besitzt an der Stelle $a = 0$ den Grenzwert $+1$:

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} |\text{sign}(x)| = +1.$$

Da dieser Grenzwert aber nicht der Funktionswert $f(0) = |\text{sign}(0)| = 0$ ist, ist die Funktion f an der Stelle $a = 0$ unstetig. Vergleichen Sie diese Ergebnisse mit Ihrer Skizze und überzeugen Sie sich daran, dass es sinnvoll ist, hier vom Grenzwert $+1$ an der Stelle 0 zu sprechen.

[Übung: Untersuchen Sie die Funktion sign auf Konvergenz an der Stelle 0 .]

c. Polstellen und hebbare Definitionslücken. Um das unterschiedliche Verhalten rationaler Funktionen an Lücken zu zeigen, betrachten einmal die folgenden drei Beispiele:

$$f_1(x) = \frac{x+1}{x-1}, \quad f_2(x) = \frac{x+1}{x^2-2x+1}, \quad f_3(x) = \frac{x^2+x-2}{x-1}.$$

Alle drei Funktionen haben lediglich bei $a = 1$ eine Definitionslücke, sie haben also den Definitionsbereich $\mathbb{R} \setminus \{1\}$. Ihr unterschiedliches Verhalten an dieser Lücke erkennen wir durch Untersuchung der Grenzwerte $\lim_{x \rightarrow 1} f_i(x)$.

Wir betrachten also Folgen $x_n \neq 1$, die gegen 1 konvergieren, und untersuchen die Folgen der Funktionswerte:

$$f_1(x_n) = \frac{x_n+1}{x_n-1}, \quad f_2(x_n) = \frac{x_n+1}{x_n^2-2x_n+1}, \quad f_3(x_n) = \frac{x_n^2+x_n-2}{x_n-1}.$$

Für die speziellen Folgen $x_n = 0,9; 0,99; 0,999; \dots$ bzw. $z_n = 1,1; 1,01; 1,001; \dots$ ergeben sich dann folgende Werte

x	0,9	0,99	0,999	...	1,001	1,01	1,1
$f_1(x)$	-19	-199	-1999	...	2001	201	21
$f_2(x)$	190	19900	1999000	...	2001000	20100	210
$f_3(x)$	2,9	2,99	2,999	...	3,001	3,01	3,1

Das durch diese Werte zu vermutende Konvergenzverhalten gilt tatsächlich für *jede* entsprechende Folge x_n bzw. z_n und führt zu folgenden Aussagen:

$$\begin{aligned} \lim_{x \nearrow 1} f_1(x) &= -\infty, & \lim_{x \searrow 1} f_1(x) &= +\infty, \\ \lim_{x \rightarrow 1} f_2(x) &= +\infty, \\ \lim_{x \rightarrow 1} f_3(x) &= 3. \end{aligned}$$

[Anmerkung: Man sagt für eine Folge a_n (und sinngemäß für eine Funktion), sie ‘konvergiere’ gegen $+\infty$, wenn sie über alle Grenzen wächst, und das soll heißen, wenn für jede reelle Zahl S die Folgenglieder a_n *schließlich* oberhalb von S liegen. Entsprechend definiert man die ‘Konvergenz’ gegen $-\infty$. Diese Sprechweisen sind suggestiv und nützlich, aber es ist **Vorsicht** geboten. Da $\pm\infty$ *keine reelle Zahl* ist, sind solche Folgen nicht konvergent! (Daher die Anführungszeichen in obiger Formulierung.) Die wichtigsten Beispiele für solche gegen $+\infty$ ‘konvergenten’ Folgen, sind Kehrwerte von positiven Nullfolgen.]

Wir wollen nun die obigen Konvergenzaussagen herleiten. Wegen $x_n \rightarrow 1$ konvergieren alle drei Nennerfolgen gegen 0, da alle drei Nennerfunktionen stetig sind und bei 1 eine Nullstelle haben. Genauso stellt man für die Zählerfolgen fest, dass sie in den entsprechenden Fällen gegen 2, 2, 0 konvergieren. Außerdem beachten wir, dass das Vorzeichen der Nennerfolgen in $f_1(x_n)$ und $f_3(x_n)$ davon abhängt, ob $x_n < 1$ oder $x_n > 1$ ist; bei $f_2(x_n)$ hingegen ist die Nennerfolge $x_n^2 - 2x_n + 1 = (x_n - 1)^2$ stets positiv.

Daher können wir zunächst die obigen Grenzwertaussagen für f_1 und f_2 folgern. Für f_3 hingegen ist das Grenzverhalten noch unklar, da Zähler- und Nennerfolge gegen 0 streben. Diese letztgenannte Tatsache folgt wegen der Stetigkeit von Zähler und Nenner daraus, dass Zähler und Nenner von f_3 an der Stelle $a = 1$ *beide* eine Nullstelle haben. Man kann daher aus beiden Termen den Linearfaktor $x - a$ abspalten und dann kürzen. Man erhält so die folgende Beschreibung von $f_3(x)$:

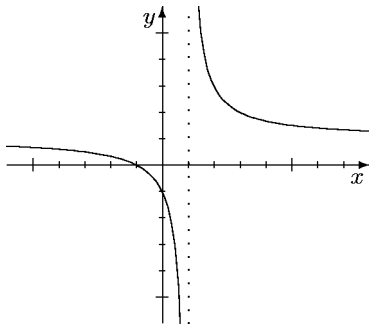
$$f_3(x) = \frac{x^2 + x - 2}{x - 1} = \frac{(x - 1)(x + 2)}{x - 1} = x + 2 \text{ für } x \neq 1.$$

(Die oben berechneten speziellen Werte zeigen in der Tat $f_3(x_n) = x_n + 2$.) Diese Beschreibung ermöglicht nun die Grenzwertbestimmung auch für $f_3(x_n)$ bei beliebiger Folge $x_n \rightarrow 1$ ($x_n \neq 1$):

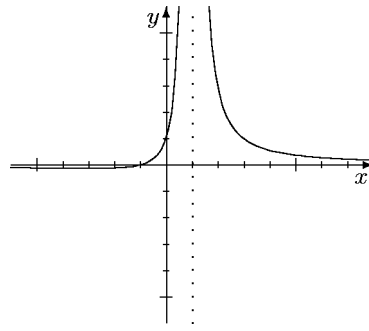
$$f_3(x_n) = x_n + 2 \rightarrow 1 + 2 = 3.$$

Diese Berechnung der Grenzwerte der Funktionen f_i bei $a = 1$ führt zu den untenstehenden Skizzen in der Nähe der Lücke $a = 1$.

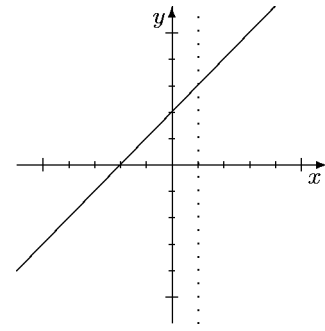
Wir erkennen deutlich das unterschiedliche Verhalten der drei Funktionen bei ihren Lücken bedingt durch die unterschiedlichen Grenzwerte an der Stelle $a = 1$. Im letzten Fall existiert eine reelle Zahl als Grenzwert, und der Graph von f_3 lässt sich durch Einsetzen eines geeigneten Lückenwertes (hier +2) stetig komplettieren; wir sprechen daher von einer *hebbaren Lücke*. In den beiden ersten Fällen ist dies nicht möglich; wir sprechen dann von einem Pol, und zwar bei f_1 von einem Pol *mit* und bei f_2 *ohne Vorzeichenwechsel*. Wir definieren also:



Graph von f_1



Graph von f_2



Graph von f_3

(5.5) Definition: Sei f eine rationale Funktion und a eine Definitionslücke.

a) Wir nennen a eine *hebbare Lücke* von f , wenn f bei a eine reelle Zahl als Grenzwert hat:

$$\lim_{x \rightarrow a} f(x) \text{ existiert in } \mathbb{R}.$$

Wir nennen diesen Grenzwert dann den Lückenwert von f bei a .

b) Wir nennen a einen *Pol*, wenn gilt

$$\lim_{x \nearrow a} f(x) = \pm\infty, \quad \lim_{x \searrow a} f(x) = \pm\infty.$$

Wenn dabei rechts- und linksseitiger Grenzwert unterschiedliches Vorzeichen haben, spricht man von einem Pol *mit*, andernfalls *ohne Vorzeichenwechsel*.

Worauf beruhte nun dieses unterschiedliche Verhalten der drei Funktionen. Ursache für die *Pole* bei f_1 und f_2 ist die Tatsache, dass a als Lücke Nullstelle des Nenners, *aber nicht des Zählers* ist. Bei f_3 hingegen ist a Nullstelle von Zähler und Nenner, und deshalb Zähler- und Nennerterm durch $x - a$ kürzbar. Kürzt man durch $x - a$ *soweit wie möglich*, so erhält man eine *Ersatzfunktion* \tilde{f}_3 mit dem Funktionsterm $\tilde{f}_3(x) = x + 1$, die mit f_3 an allen Stellen $x \neq 1$ übereinstimmt. Darüberhinaus hat \tilde{f}_3 bei $x = a = 1$ *keine* Definitionslücke mehr. Man kann daher den Grenzwert leicht bestimmen:

$$\lim_{x \rightarrow 1} f_3(x) = \lim_{x \rightarrow 1} \tilde{f}_3(x) = \tilde{f}_3(1) = 2,$$

wobei die vorletzte Gleichung aus der Stetigkeit von \tilde{f}_3 an der Stelle $a = 1$ folgt. Bei einer hebbaren Lücke a ist der Lückenwert gerade der Wert der Ersatzfunktion an der Lücke. Wir fassen unsere bisherige Diskussion zusammen:

(5.6) Fazit: a) Ist f eine rationale Funktion, $f(x) = \frac{g(x)}{h(x)}$ ihr Funktionsterm mit ganzrationalem Zähler $g(x)$ und Nenner $h(x)$, so sind die Nullstellen des Nenners $h(x)$ die Definitionslücken von f .

b) Ist eine solche Nullstelle a des Nenners keine Nullstelle des Zählers, so ist a Polstelle von f .

c) Ist dagegen $a \in \mathbb{R}$ Nullstelle des Nenners $h(x)$ und des Zählers $g(x)$, so spalte man in beiden den Linearfaktor $x - a$ so oft wie möglich ab (durch Polynomdivision) und kürze. Man erhält so eine Ersatzfunktion \tilde{f} , die mit der gegebenen Funktion außerhalb a übereinstimmt, aber evtl. zusätzlich auch bei a definiert ist. Man arbeite nun mit der Ersatzfunktion weiter.

d) Hat auch diese Ersatzfunktion \tilde{f} bei a eine Lücke, so ist a eine Polstelle von \tilde{f} und f . Ist dagegen a keine Lücke der Ersatzfunktion, so ist a hebbare Lücke von f . Dies ist genau dann der Fall, wenn der Zähler $g(x)$ der gegebenen Funktion f bei a eine Nullstellenordnung hatte, die mindestens so groß war wie die Nullstellenordnung des Nenners $h(x)$ bei a . [Dann werden nämlich im Nenner $h(x)$ alle auftretenden Linearfaktoren $x - a$ gekürzt und danach ist a keine Nullstelle des Nenners mehr, also keine Definitionslücke der Ersatzfunktion.]

e) Eine Lücke einer rationalen Funktion ist entweder hebbar oder ein Pol.

d. Das Verhalten im Unendlichen und Asymptoten. Neben dem Verhalten bei Lücken und der Vorzeichenverteilung interessiert uns nun noch das Verhalten rationaler Funktionen ‘im Unendlichen’. Während bei ganz-rationalen Funktionen vom Grad ≥ 1 für $x \rightarrow \pm\infty$ die Funktionswerte ebenfalls gegen $+\infty$ oder $-\infty$ streben, können bei gebrochen rationalen Funktionen noch andere Phänomene auftreten. Etwa bei den in c. betrachteten Beispielen f_1 , f_2 bzw. f_3 stellt man fest:

$$\lim_{x \rightarrow \infty} f_1(x) = 1, \quad \lim_{x \rightarrow \infty} f_2(x) = 0, \quad \lim_{x \rightarrow \infty} f_3(x) = \infty.$$

Dies, und den folgenden allgemeinen Satz, erkennt man mit den Methoden aus Kap. I, §2 c (Ausklammern der höchsten x -Potenz in Zähler und Nenner):

(5.7) Satz: Sei f eine rationale Funktion mit dem Funktionsterm

$$f(x) = \frac{a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \dots + b_1 x + b_0}$$

und $a_n \neq 0$, $b_m \neq 0$. (Also sind n der Zählergrad und m der Nennergrad von f , a_n der höchste Koeffizient im Zähler und b_m der höchste Koeffizient im Nenner.)

a) Ist der Zählergrad n kleiner als der Nennergrad m , $n < m$, so gilt

$$f(x) \rightarrow 0 \text{ für } x \rightarrow \pm\infty,$$

und wir sagen: Die x -Achse ist Asymptote (Schmiegegerade) für f .

b) Stimmen Zähler- und Nennergrad überein, $n = m$, so gilt

$$f(x) \rightarrow \frac{a_n}{b_m} \text{ für } x \rightarrow \pm\infty,$$

und wir sagen: Die Parallele zur x -Achse mit der Gleichung $y = a_n/b_m$ ist Asymptote für f .

c) Ist der Zählergrad n größer als der Nennergrad m , $n > m$, so gilt

$$\lim_{x \rightarrow \infty} f(x) = \begin{cases} +\infty & \text{für } \frac{a_n}{b_m} > 0, \\ -\infty & \text{für } \frac{a_n}{b_m} < 0, \end{cases}.$$

Ist $n - m$ gerade, so gilt für $x \rightarrow -\infty$ dasselbe, während bei $n - m$ ungerade sich die Vorzeichen umkehren:

$$\lim_{x \rightarrow -\infty} f(x) = \begin{cases} + \lim_{x \rightarrow \infty} f(x) & \text{für } n - m \text{ gerade,} \\ - \lim_{x \rightarrow \infty} f(x) & \text{für } n - m \text{ ungerade.} \end{cases}$$

Beweis: Wir klammern im Zähler x^n und im Nenner x^m aus und erhalten

$$f(x) = \frac{x^n}{x^m} \cdot \frac{a_n + a_{n-1}x^{-1} + \dots + a_1x^{1-n} + a_0x^{-n}}{b_m + b_{m-1}x^{-1} + \dots + b_1x^{1-m} + b_0x^{-m}}. \quad (*)$$

Aus den Grenzwertsätzen erhält man wegen $b_m \neq 0$: Der zweite Bruch konvergiert gegen a_n/b_m .

Das Verhalten des ersten Faktors x^n/x^m hängt nur von n und m ab:

Ist $\underline{n = m}$, so ist dieser Faktor 1, und $f(x)$ strebt gegen a_n/b_m , wie in b) behauptet.

Ist $\underline{n < m}$, so ist $m - n > 0$ eine natürliche Zahl und es gilt

$$\frac{x^n}{x^m} = \frac{1}{x^{m-n}} \rightarrow 0 \text{ für } x \rightarrow \pm\infty.$$

Aus (*) ergibt sich die Behauptung von Teil a).

Ist schließlich $\underline{n > m}$, so ist $n - m > 0$ eine natürliche Zahl und daher

$$\lim_{x \rightarrow \infty} \frac{x^n}{x^m} = \lim_{x \rightarrow \infty} x^{n-m} = \infty.$$

Da a_n/b_m nicht Null ist, folgt daraus gemäß (*) die erste Aussage von c). Für $x \rightarrow -\infty$ gilt

$$\lim_{x \rightarrow -\infty} x^{n-m} = \begin{cases} \infty & \text{für } n-m \text{ gerade,} \\ -\infty & \text{für } n-m \text{ ungerade,} \end{cases}$$

und die zweite Behauptung von c) folgt.

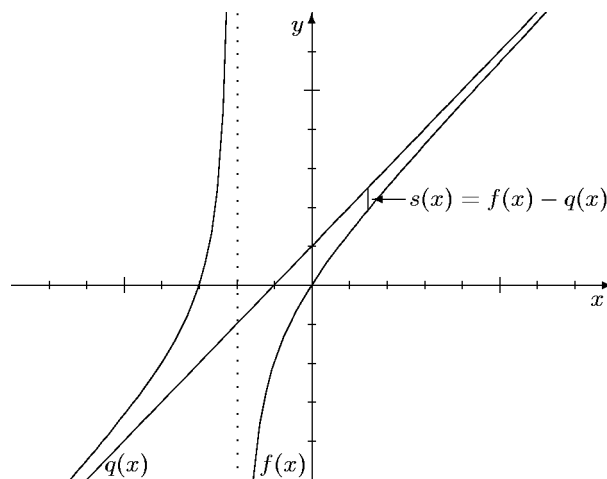
Wir wollen nun den Fall c) (Zählergrad größer als Nennergrad) noch etwas genauer untersuchen, und zwar mit Hilfe der Polynomdivision: Ist f eine rationale Funktion, so kann man den Zähler mit Rest durch den Nenner dividieren. Betrachten wir einmal das Beispiel

$$f(x) = \frac{x^2 + 3x}{x + 2}.$$

Dividiert man den Zähler $g(x) = x^2 + 3x$ durch den Nenner $h(x) = x + 2$, so erhält man als Quotient $q(x) = x + 1$ und als Rest $r(x) = -2$, also

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

Damit haben wir $f(x)$ dargestellt als Summe aus einem Polynomterm $q(x) = x + 1$ und einem gebrochen-rationalen Term $s(x) = -2/(x + 2)$, bei dem der Zählergrad echt kleiner ist als der Nennergrad. Gemäß Satz (5.7) a) konvergiert $s(x)$ gegen 0 für $x \rightarrow \pm\infty$. Wegen $f(x) = x + 1 + s(x)$ bedeutet dies, dass der Abstand zwischen $f(x)$ und $x + 1$ gegen 0 strebt: Der Graph



von f schmiegt sich immer mehr der Geraden mit der Gleichung $y = x + 1$ an; f hat damit eine Asymptote, ihre Gleichung ist gegeben durch $y = x + 1$.

Diese Überlegungen kann man allgemein für beliebige rationale Funktionen f durchführen, deren Zählergrad größer ist als der Nennergrad: Mittels Polynomdivision erhält man eine Darstellung von $f(x) = \frac{g(x)}{h(x)}$ in der Form

$$f(x) = q(x) + s(x), \quad q(x) \text{ ganz-rational,} \quad s(x) \rightarrow 0 \text{ für } x \rightarrow \pm\infty.$$

Damit schmiegt sich der Graph von f dem Graphen der ganz-rationalen Funktion q immer mehr an. Auf diese Weise hat man eine weitere Methode, mit der man das Verhalten einer rationalen Funktion f im Unendlichen bestimmen kann: Man dividiert den Zähler von f durch den Nenner mit Rest und untersucht den gefundenen Quotienten $q(x)$. Das Verhalten der rationalen Funktion f im Unendlichen stimmt dann mit dem von q überein. Da q ganz-rational ist, ist letzteres bekannt. (q hat den Grad $n - m$ und den höchsten Koeffizienten a_n/b_m .)

Diese Überlegungen zeigen außerdem, dass auch im Falle c) Asymptoten existieren können. Ist nämlich der Zählergrad von f nur um 1 größer als der Nennergrad, so ist $q(x)$ vom Grad

1 und der Graph von q eine Gerade mit dem Anstieg a_n/b_m , an die sich der Graph von f 'anschmiegt'.

Wir können nun unsere anschauliche Vorstellung einer Schmiegegeraden in der folgenden Definition festhalten und Satz (5.7) komplettieren:

(5.8) Definition: Sei f eine rationale Funktion. Eine *Asymptote* für f ist eine Gerade, die Graph einer linearen Funktion q ist mit der Eigenschaft: Der Abstand zwischen $f(x)$ und $q(x)$ konvergiert gegen 0 für $x \rightarrow \pm\infty$. In Formeln:

$$f(x) - q(x) \rightarrow 0 \text{ für } x \rightarrow \pm\infty.$$

Die in (5.7) a), b) genannten Asymptoten sind tatsächlich Asymptoten im Sinne dieser allgemeinen Definition. Außerdem erhalten wir die folgende Komplettierung von Satz (5.7):

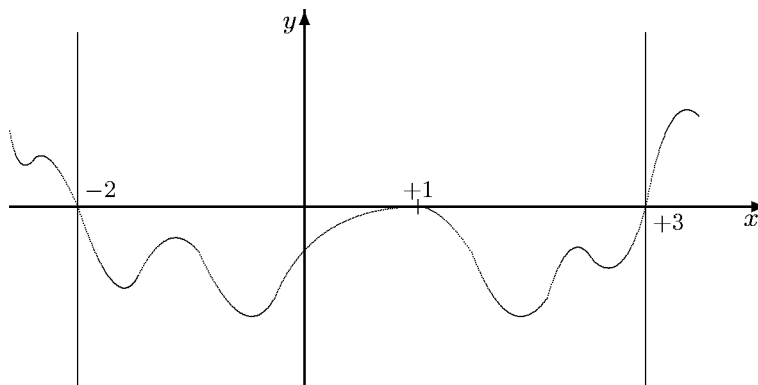
(5.9) Satz: Es sei f eine rationale Funktion.

- a) f besitzt genau dann eine Asymptote, wenn der Zählergrad höchstens um 1 größer ist als der Nennergrad.
- b) Ist der Zählergrad echt kleiner als der Nennergrad, so ist die x -Achse Asymptote für f .
- c) Ist der Zählergrad gleich dem Nennergrad, so ist die Parallele zur x -Achse mit der Gleichung $y = a_n/b_m$ Asymptote für f ; dabei sind a_n bzw. b_m die Koeffizienten der jeweils höchsten x -Potenz im Zähler bzw. Nenner.
- c) Ist der Zählergrad um 1 größer als der Nennergrad, so erhält man die Gleichung der Asymptote, indem man den Zähler $g(x)$ von $f(x)$ durch den Nenner $h(x)$ mit Rest dividiert: $g(x) : h(x) = q(x)$ Rest $r(x)$. $y = q(x)$ ist dann die Gleichung der Asymptoten. Ihr Anstieg ist das Verhältnis a_n/b_m der höchsten Koeffizienten von Zähler und Nenner von $f(x)$.
- d) Ist der Zählergrad um mehr als 1 größer als der Nennergrad, so erhält man durch den Graphen von q keine Schmiegegerade, sondern eine Schmiegeparabel (vom Grade $n - m$), also den Graph einer ganz-rationalen Funktion, dem sich der Graph der rationalen Funktion für $x \rightarrow \pm\infty$ anschmiegt.

III. Differentialrechnung

§6 Der Ableitungsbegriff

Unsere bisherigen Diskussionen in Kapitel II ermöglichen nur die Bestimmung der Nullstellen und Lücken rationaler Funktionen und des *Vorzeichens der Funktionswerte*. Welchen *Verlauf* jedoch die Funktion dort hat, kann aus den bisherigen Überlegungen nicht entnommen werden. So lassen die Berechnungen des Beispiels (*) auf S. 21f. nicht nur den dort skizzierten Verlauf, sondern etwa auch den folgenden zu:



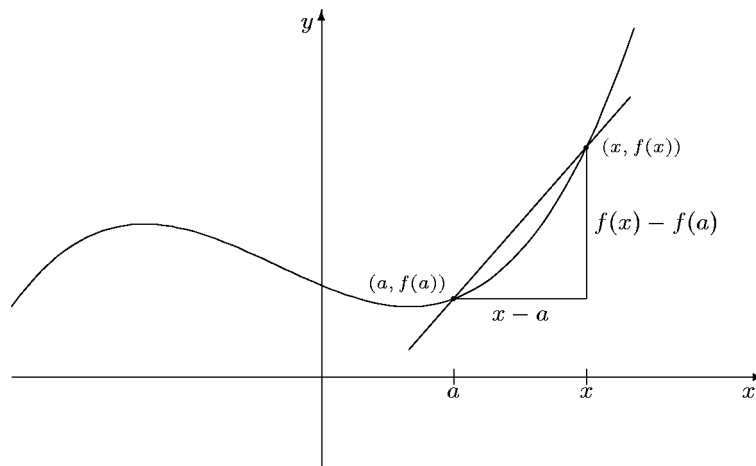
Um nun den Verlauf zwischen zwei Nullstellen bzw. Lücken genauer zu erfassen, wollen wir den Begriff des *Anstiegs* des Funktionsgraphen präzisieren. Wenn uns dies gelingt und wir entscheiden können, in welchen Bereichen der Funktionsgraph *steigt* und wo er *fällt*, so können wir den Verlauf des Graphen schon recht genau bestimmen.

a. Sekantensteigung, Differenzenquotient, Ableitung, Tangente. Der Begriff des *Anstiegs* eines Funktionsgraphen ist für lineare Funktionen bekannt. Eine lineare Funktion, d. i.

eine ganz-rationale Funktion vom Grade ≤ 1 , ist gegeben durch einen Funktionsterm der Form $f(x) = mx + b$ mit reellen Zahlen $m, b \in \mathbb{R}$. Ihr Graph ist eine Gerade mit dem *y-Achsenabschnitt* b und dem *Anstieg* m . Dieser ergibt sich aus den Koordinaten zweier verschiedener Punkte $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ auf der Geraden durch

$$\text{Geradenanstieg } m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Will man nun für einen beliebigen Funktionsgraphen den Anstieg bestimmen, so hat man zunächst das Problem, dass der Anstieg an verschiedenen Stellen verschieden ist. Wir müssen also zunächst eine *Stelle* a fixieren und untersuchen den Graphen in der Nähe des Punktes $P = (a, f(a))$. Dazu wählen wir in der Nähe von P einen weiteren Punkt $Q = (x, f(x))$ auf $G(f)$ ($x \neq a$) und verbinden beide durch eine Gerade. Eine solche *Gerade durch zwei Punkte des Graphen* $G(f)$ nennt man eine *Sekante*.



Der Anstieg dieser Sekante ist gegeben durch

$$D(x) = \frac{f(x) - f(a)}{x - a}.$$

Man nennt diesen *Sekantenanstieg* wegen der Form des Terms auch *Differenzenquotient*. Dieser ist (bei fixierter Stelle a) von x abhängig: Er stellt eine *Funktion* von x dar. Ändert man x , so ändert sich in der Regel der Sekantenanstieg. Etwa für die Funktion f gegeben durch $f(x) = x^2$ und die Stelle a ist der Differenzenquotient dann gegeben durch

$$D(x) = \frac{f(x) - f(a)}{x - a} = \frac{x^2 - a^2}{x - a}.$$

Will man den Anstieg von $G(f)$ an der Stelle a , d. h. beim Punkt $P = (a, f(a))$ erfassen, so nähert man sich mit dem Punkt $Q = (x, f(x))$ immer mehr P , d. h. x nähert sich immer mehr a . Man untersucht dann das Verhalten der Sekantenanstiege $D(x)$ bei Annäherung von x an a , d. h. man führt den Grenzübergang $x \rightarrow a$ durch. Existiert dann der Grenzwert von $D(x)$ für $x \rightarrow a$, so definiert man diesen als *Anstieg* von f an der Stelle a und bezeichnet ihn mit $f'(a)$:

(6.1) Definition: Es sei f eine Funktion und a eine Stelle im Definitionsbereich. Dann definiert man:

a) f ist an der Stelle a differenzierbar, wenn der Grenzwert

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

in \mathbb{R} existiert.

b) Ist dies der Fall, so nennen wir den Grenzwert

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

den *Anstieg von f an der Stelle a* .

c) Indem man jeder Stelle a , an der f differenzierbar ist, den Anstieg $f'(a)$ zuordnet

$$a \mapsto f'(a)$$

erhält man eine neue Funktion f' , die sog. *Ableitung*(sfunktion). Ihr Definitionsbereich ist die Menge

$$\mathcal{D}(f') = \{a \in \mathcal{D}(f) \mid f \text{ ist bei } a \text{ differenzierbar}\}.$$

d) Man sagt f ist *differenzierbar*, wenn f an allen Stellen ihres Definitionsbereiches differenzierbar ist.

Nachdem wir den Begriff des Anstiegs einer Funktion f an einer Stelle a definiert haben, können wir nun auch den Begriff der Tangente präzise fassen:

(6.2) Definition: Gegeben ist eine Funktion f , die an der Stelle a differenzierbar ist, für die also der Anstieg $f'(a)$ definiert ist. Dann versteht man unter der *Tangente* an den Graphen von f im Punkte $(a, f(a))$ die *Gerade* durch den Punkt $(a, f(a))$, die dieselbe Steigung hat wie f an der Stelle a , deren Steigung also genau $f'(a)$ ist.

Die Tangente an einen Graphen G in einem Punkt $P \in G$ ist die Gerade durch den Punkt P , die denselben Anstieg hat wie der Graph in diesem Punkt.

(6.3) Bemerkung: Die Tangente an den Graphen von f an der Stelle a , das soll heißen im Punkte $(a, f(a))$, ist gegeben durch den Graphen der linearen Funktion

$$t(x) = f(a) + f'(a)(x - a).$$

Wir nennen t die *Tangentenfunktion* zu f an der Stelle a

Beweis: Als Gerade ist die Tangente Graph einer linearen Funktion $t(x) = mx + b$. Dabei ist der Anstieg m festgelegt; er soll gerade der Anstieg $f'(a)$ der Funktion f an der Stelle a sein: $m = f'(a)$. Damit lautet die Tangentenfunktion $t(x) = f'(a)x + b$. Man muss nun b bestimmen. Dazu verwenden wir die zweite Forderung bei der Tangentendefinition: Die Tangente soll durch den Punkt $P = (a, f(a))$ verlaufen; das bedeutet, dass P auf dem Graphen von t liegen, also $t(a) = f(a)$ sein muss:

$$f(a) = t(a) = f'(a) \cdot a + b, \quad \text{also } b = f(a) - f'(a) \cdot a.$$

Damit erhält man — wie behauptet — die Tangentenfunktion

$$t(x) = f'(a)x + f(a) - f'(a) \cdot a = f(a) + f'(a)(x - a).$$

b. Differenzierbarkeit und Stetigkeit. Man beachte, dass die Existenz des Anstiegs $f'(a)$ bzw. einer Tangente an einer Stelle a nicht selbstverständlich ist. Zunächst einmal kann eine Funktion nicht differenzierbar sein, wenn sie nicht wenigstens stetig ist:

Differenzierbarkeit hat Stetigkeit zur Folge!

(6.4) Satz: Ist eine Funktion f an einer Stelle a differenzierbar, so ist sie dort auch stetig.

Beweis: Wir müssen zeigen:

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \text{ existiert in } \mathbb{R} \implies \lim_{x \rightarrow a} f(x) = f(a).$$

Wir setzen voraus, dass der Grenzwert des Differenzenquotienten existiert, und betrachten folgende Darstellung für $f(x)$:

$$f(x) = f(a) + \frac{f(x) - f(a)}{x - a} \cdot (x - a) \quad \text{gültig für } x \neq a.$$

Aufgrund der Grenzwertsätze ergibt sich daraus

$$f(x) = f(a) + \underbrace{\frac{f(x) - f(a)}{x - a}}_{\rightarrow f'(a)} \cdot \underbrace{(x - a)}_{\rightarrow 0} \rightarrow f(a) + f'(a) \cdot 0 = f(a) \quad \text{für } x \rightarrow a,$$

was zu beweisen war.

Aber auch stetige Funktionen sind nicht notwendig differenzierbar. Typisches Beispiel ist die Betragsfunktion:

(6.5) Beispiel: Die Betragsfunktion $f(x) = |x|$ ist an der Stelle $a = 0$ nicht differenzierbar; sie besitzt an dieser Stelle keine eindeutige Tangente.

Zur Begründung betrachte man einmal die Funktion f nur über dem Intervall $[0, \infty[$. Dort ist $f(x) = x$, der Graph in diesem Bereich eine Gerade mit Anstieg 1. Also

$$\lim_{x \searrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \searrow 0} \frac{x}{x} = 1.$$

Betrachtet man jedoch den Bereich $] - \infty, 0]$, so gilt dort $f(x) = -x$ und man erhält als Anstiegswert -1 :

$$\lim_{x \nearrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \nearrow 0} \frac{-x}{x} = -1.$$

Wenn aber die beiden *einseitigen* Grenzwerte von $\frac{f(x) - f(a)}{x - a}$ nicht übereinstimmen, kann der gesuchte (beidseitige) Grenzwert $\lim_{x \rightarrow 0} \dots$ nicht existieren: $f(x) = |x|$ ist an der Stelle 0 nicht differenzierbar. Geometrisch erkennt man dies an dem *Knick*, den der Graph dort hat.

Anschaulich gesprochen liegt Differenzierbarkeit vor, wenn der Graph *glatt* ist und keinen *Knick* hat (wie bei der Betragsfunktion an der Stelle 0).

An allen anderen Stellen ist die Betragsfunktion jedoch differenzierbar. (Bestimmen Sie zur Übung die Ableitungsfunktion f' samt ihrem Definitionsbereich!)

Als Gegenstück zum obigen Beispiel wollen wir nun zeigen, dass alle rationalen Funktionen differenzierbar sind:

(6.6) Satz: Rationale Funktionen f sind an allen Stellen ihres Definitionsbereiches differenzierbar; es existiert daher die Ableitungsfunktion f' und diese hat denselben Definitionsbereich wie f .

Beweis: Sei f eine rationale Funktion und a eine Stelle im Definitionsbereich von f . Wir untersuchen den Differenzenquotienten

$$D(x) = \frac{f(x) - f(a)}{x - a}.$$

Da f eine rationale Funktion ist, ist $f(x) - f(a)$ und dann auch $D(x)$ rational. Wegen des Nenners $x - a$ ist D an der Stelle a nicht definiert, a ist eine Lücke von D . Der Zähler $f(x) - f(a)$ ist jedoch bei a definiert und hat dort den Wert $f(a) - f(a) = 0$; es lässt sich also im Zähler der Linearfaktor $x - a$ abspalten: $f(x) - f(a) = (x - a)q(x)$, wobei q eine bei a definierte (!) rationale Funktion ist. Damit lässt sich in $D(x)$ der Linearfaktor kürzen und man erhält

$$D(x) = \frac{(x - a)q(x)}{x - a} = q(x) \quad \text{für } x \neq a.$$

Da q bei a definiert und als rationale Funktion stetig ist, existiert

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} q(x) = q(a) \in \mathbb{R}.$$

Dies bedeutet: f ist an der Stelle a differenzierbar. (Beachten Sie: q ist nichts anderes als die Ersatzfunktion \tilde{D} und die Lücke a von D hat sich als hebbare Lücke von D erwiesen.)

In den Übungen haben Sie Gelegenheit, diesen Beweis an konkreten Beispielen nachzuarbeiten und dabei dann *explizite Formeln* für $f'(a)$ herzuleiten (siehe auch Abschnitt c.).

c. Die Ableitung der Potenzfunktionen. Wir wollen hier nun allgemein für *beliebige* Potenzfunktionen die Ableitungsfunktion f' *explizit* bestimmen.

(6.7) Satz: Die Potenzfunktionen f gegeben durch $f(x) = x^n$ ($n \in \mathbb{N}$) sind differenzierbar und ihre Ableitungsfunktion f' ist gegeben durch:

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

Beweis: Wir betrachten eine beliebige Stelle a und untersuchen den Differenzenquotienten

$$D(x) = \frac{x^n - a^n}{x - a}.$$

Dieser lässt sich folgendermaßen darstellen

$$D(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + ax^{n-2} + a^2x^{n-3} + \dots + a^{n-2}x + a^{n-1}. \quad (*)$$

Diese Formel haben Sie in Spezialfällen in den Übungen selbst ermittelt. Man beweist (*) am besten durch Ausmultiplizieren

$$\begin{aligned} (x - a) & (x^{n-1} + ax^{n-2} + \dots + a^{n-2}x + a^{n-1}) \\ = & \begin{array}{cccccccc} x^n & + & ax^{n-1} & + & \dots & + & a^{n-2}x^2 & + & a^{n-1}x \\ & & - & ax^{n-1} & - & a^2x^{n-2} & - & \dots & - & a^{n-1}x & - & a^n \end{array} \\ = & \begin{array}{cccccccc} x^n & & & & & & & & & & & - & a^n \end{array} \end{aligned}$$

Folglich erhält man für $f(x) = x^n$ als Ableitungswert

$$\begin{aligned} f'(a) &= \lim_{x \rightarrow a} \frac{x^n - a^n}{x - a} = \lim_{x \rightarrow a} (x^{n-1} + ax^{n-2} + a^2x^{n-3} + \dots + a^{n-2}x + a^{n-1}) \\ &= a^{n-1} + aa^{n-2} + a^2a^{n-3} + \dots + a^{n-2}a + a^{n-1} \\ &= n \cdot a^{n-1}. \end{aligned}$$

Dies gilt für alle $a \in \mathbb{R} = \mathcal{D}(f)$ und die Behauptung ist bewiesen.

Bei richtiger Interpretation ist Satz (6.7) auch richtig für $n = 0$. Für $n = 0$ erhält man die konstante Funktion $f(x) = x^0 = 1$. Ihr Graph ist eine Gerade parallel zur x -Achse, ihre Ableitung also an allen Stellen 0. Allgemein gilt:

(6.8) Bemerkung: Konstante Funktionen $f(x) = c$ ($c \in \mathbb{R}$) haben die Ableitung $f'(x) = 0$.

$$f(x) = c \text{ konstant} \implies f'(x) = 0.$$

Wir wollen nun Satz (6.7) sogar ausdehnen auf negative Exponenten, d. h. auf die Kehrwerte $f(x) = \frac{1}{x^n} = x^{-n}$ von Potenzfunktionen. Wieder müssen wir den Differenzenquotienten von f

untersuchen. Sei also a eine Stelle des Definitionsbereiches von f , d. h. $a \neq 0$. Dann gilt für den Differenzenquotienten (wieder unter Benutzung der oben hergeleiteten Formel (*))

$$\begin{aligned}\frac{\frac{1}{x^n} - \frac{1}{a^n}}{x - a} &= \frac{a^n - x^n}{x^n a^n (x - a)} = \left(-\frac{1}{x^n a^n}\right) \frac{x^n - a^n}{x - a} \\ &= \left(-\frac{1}{x^n a^n}\right) (x^{n-1} + ax^{n-2} + a^2 x^{n-3} + \dots + a^{n-2} x + a^{n-1}),\end{aligned}$$

so dass man den Grenzwert bestimmen kann:

$$\begin{aligned}f'(a) &= \lim_{x \rightarrow a} \left[\left(-\frac{1}{x^n a^n}\right) (x^{n-1} + ax^{n-2} + a^2 x^{n-3} + \dots + a^{n-2} x + a^{n-1}) \right] \\ &= \left(-\frac{1}{a^n a^n}\right) (a^{n-1} + aa^{n-2} + a^2 a^{n-3} + \dots + a^{n-2} a + a^{n-1}) \\ &= -\frac{na^{n-1}}{a^{2n}} = -\frac{n}{a^{n+1}}.\end{aligned}$$

Beschreibt man nun die Ausgangs- und Ableitungsfunktion nicht durch Brüche, sondern durch Potenzen mit *negativen* Exponenten, so erhält man

$$f(x) = x^{-n} \implies f'(x) = -nx^{-n-1}$$

und erkennt, dass Satz (6.7) in der angegebenen Form auch richtig ist für negative Exponenten:

(6.7') Satz: Für jede ganze Zahl $n \in \mathbb{Z}$ ist die Potenzfunktion f , gegeben durch $f(x) = x^n$, differenzierbar und ihre Ableitungsfunktion f' ist gegeben durch

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

Wir erwähnen bereits hier, dass dieser Satz sogar für *gebrochene* Exponenten im wesentlichen richtig bleibt. Wir werden dies aber erst später beweisen. Setzt man in (6.7') $n = 1/2$, so erhält man für $f(x) = \sqrt{x} = x^{1/2}$ die Ableitung $f'(x) = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}$. Für diese Aussage wollen wir ad hoc einen Beweis geben und zugleich auf die dabei auftretende Problematik hinweisen:

(6.9) Satz: Die Wurzelfunktion $f(x) = \sqrt{x}$ ist an allen Stellen $a > 0$ differenzierbar, und die Ableitungsfunktion ist gegeben durch

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

An der Stelle $a = 0$ ist die Wurzelfunktion *nicht* differenzierbar.

Zum *Beweis* untersuchen wir wieder den Differenzenquotienten:

$$\frac{\sqrt{x} - \sqrt{a}}{x - a} = \frac{\sqrt{x} - \sqrt{a}}{(\sqrt{x})^2 - (\sqrt{a})^2} = \frac{\sqrt{x} - \sqrt{a}}{(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a})} = \frac{1}{\sqrt{x} + \sqrt{a}}.$$

Damit können wir an Stellen $a > 0$ den Grenzübergang $x \rightarrow a$ durchführen:

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

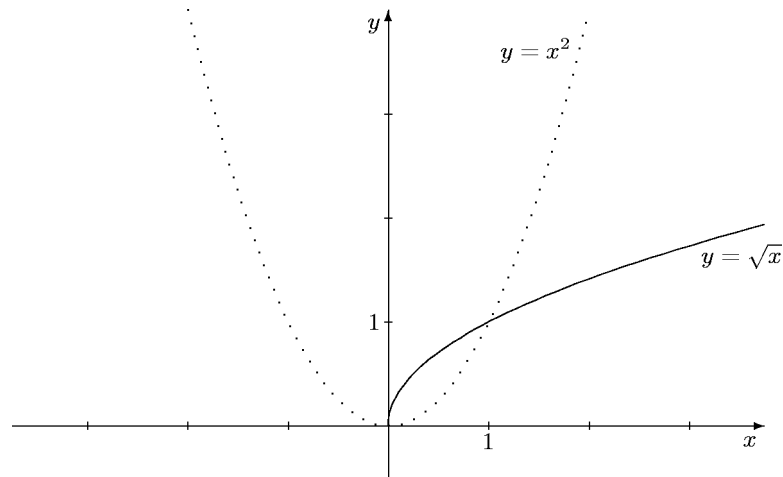
Beachten Sie, dass wir beim vorletzten Gleichheitszeichen die Stetigkeit der Wurzelfunktion benutzt haben sowie die Tatsache, dass $a \neq 0$ ist.

Für die Stelle $a = 0$ muss man den Differenzenquotienten

$$\frac{\sqrt{x} - \sqrt{0}}{x - 0} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}}$$

untersuchen. Für $x \rightarrow 0$ hat aber $\frac{1}{\sqrt{x}}$ keinen Grenzwert in \mathbb{R} , sondern wächst über alle Grenzen: $f'(0)$ existiert nicht!

Diese Resultate lassen sich auch deutlich am Verlauf des Graphen der Wurzelfunktion erkennen. Dieser Graph ist die obere Hälfte der an der Winkelhalbierenden des I./III. Quadranten gespiegelten Normalparabel mit Scheitel $(0, 0)$ (siehe Skizze). Die Normalparabel hat in ihrem



Scheitelpunkt die x -Achse als Tangente. Bei der Spiegelung geht diese in die y -Achse über. Die y -Achse berührt also im Koordinatenursprung den Graphen der Wurzelfunktion. Für die y -Achse ist aber keine Steigung definiert. Dies entspricht genau dem obigen Resultat, dass für die Wurzelfunktion an der Stelle 0 der Anstieg $f'(0)$ nicht existiert.

d. Erste Ableitungsregeln. Aufgrund von Satz (6.7') kennen wir nun erste Formeln für die Ableitung der Potenzfunktionen $f(x) = x^n$ mit $n \in \mathbb{Z}$. Wir wollen nun versuchen allgemein ganzrationale Funktionen f abzuleiten. Wichtiges Hilfsmittel sind dafür die folgenden Ableitungsregeln.

(6.10) Satz: (Erste Ableitungsregeln)

a) (Faktorregel) Ist eine Funktion g an einer Stelle $a \in \mathbb{D}_g$ differenzierbar, so ist für jede reelle Zahl $c \in \mathbb{R}$ auch die Funktion $f = cg$, die durch den Funktionsterm $f(x) = c \cdot g(x)$ gegeben ist, bei a differenzierbar, und es gilt

$$f'(a) = (cg)'(a) = c \cdot g'(a).$$

b) (Summenregel) Sind g und h an einer Stelle a differenzierbare Funktionen, so ist auch ihre Summenfunktion $f = g + h$, gegeben durch den Funktionsterm $f(x) = g(x) + h(x)$, bei a differenzierbar. Es gilt dann

$$f'(a) = (g + h)'(a) = g'(a) + h'(a).$$

Beweis: a) Wir müssen den Differenzenquotienten von $f = c \cdot g$ zur Stelle a untersuchen. Dieser ist gegeben durch

$$\frac{f(x) - f(a)}{x - a} = \frac{c \cdot g(x) - c \cdot g(a)}{x - a} = c \cdot \frac{g(x) - g(a)}{x - a}.$$

Der Differenzenquotient von $f = cg$ ist also das c -fache des Differenzenquotienten von g . Nun ist g bei a differenzierbar; also existiert der Grenzwert

$$g'(a) = \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a}.$$

Dann existiert aufgrund unserer Grenzwertsätze auch der Grenzwert

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} c \cdot \frac{g(x) - g(a)}{x - a} = c \cdot g'(a).$$

Ad b): Wieder betrachten wir den Differenzenquotienten von f :

$$\begin{aligned} \frac{f(x) - f(a)}{x - a} &= \frac{(g(x) + h(x)) - (g(a) + h(a))}{x - a} = \frac{g(x) + h(x) - g(a) - h(a)}{x - a} \\ &= \underbrace{\frac{g(x) - g(a)}{x - a}}_{\rightarrow g'(a)} + \underbrace{\frac{h(x) - h(a)}{x - a}}_{\rightarrow h'(a)}. \end{aligned}$$

Der Differenzenquotient der Summenfunktion $f = g + h$ ist also die Summe der Differenzenquotienten von g und von h , und diese konvergieren gegen die jeweiligen Ableitungswerte. Gemäß den Grenzwertsätzen erhält man daraus dann

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

Damit ist Teil b) bewiesen.

Achtung: Die Ableitungsregel für die Produktfunktion $g \cdot h$ werden wir erst später kennenlernen. Sie ist komplizierter als die Regel (6.10),b) für $g + h$: *Die Ableitung der Produktfunktion $g \cdot h$ ist nicht das Produkt der Ableitungen!* Dies kann man leicht an den uns schon bekannten Ableitungsregeln für die Potenzfunktionen ablesen: Betrachten wir $g(x) = x^2$ und $h(x) = x^3$. Dann ist $g'(x) \cdot h'(x) = 2x \cdot 3x^2 = 6x^3$ offensichtlich **nicht** die Ableitung von $g(x) \cdot h(x) = x^5$.

Beliebige ganzrationale Funktionen werden durch Addition und Multiplikation mit Zahlen aus den Potenzfunktionen $f(x) = x^n$ aufgebaut. Dabei sind für n beliebige natürliche Zahlen oder 0 zulässig. Gemäß (6.7) und (6.10) können wir daher alle ganz-rationalen Funktionen ableiten und erhalten den folgenden

(6.11) Satz: *Ganzrationale Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ sind differenzierbar. Ist die Funktion f durch den Funktionsterm*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$$

gegeben, so ist die Ableitung $f' : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + 2 a_2 x + a_1.$$

Insbesondere sehen wir, dass die Ableitungsfunktion wieder eine ganz-rationale Funktion ist, deren Grad um 1 kleiner ist als der von f , wenn letzterer ≥ 1 ist.

Wir müssen lediglich noch die letzte Aussage begründen. f hat den Grad n , wenn $a_n \neq 0$ ist. Ist außerdem $n \geq 1$, also $n \neq 0$, so ist natürlich auch $n a_n \neq 0$ und f' hat den Grad $n - 1$.

§7 Monotonie und Extrema

a. Extremstellen und stationäre Stellen. Nachdem wir nun in der Lage sind, beliebige ganz-rationale Funktionen abzuleiten, wollen wir uns jetzt wieder der Frage zuwenden, was wir aus der Kenntnis der Ableitung f' über die Funktion f selbst erfahren können.

(7.1) Definition: Sei f eine Funktion und a eine Stelle im Definitionsbereich von f .

a) Wir sagen: f hat bei a ein (lokales) *Maximum*, wenn in einer (kleinen) Umgebung von a $f(a) \geq f(x)$ gilt, d.h. genauer: es gilt $f(x) \leq f(a)$ für alle $x \in \mathcal{D}_f$, $|x - a| < \varepsilon$ mit einem (kleinen) $\varepsilon > 0$.

b) Entsprechend spricht man von einem (lokalen) *Minimum*, wenn $f(a) \leq f(x)$ in einer Umgebung von a gilt.

c) f hat bei a ein *Extremum*, wenn ein Maximum oder Minimum vorliegt.

Wir nennen a entsprechend *Maximums-*, *Minimums-* oder einfach *Extremstelle*, und den zugehörigen Punkt $(a, f(a))$ des Graphen von f nennen wir dementsprechend *Hoch-*, *Tief-* bzw. *Extrempunkt*.

(7.2) Satz: (Notwendiges Extremstellenkriterium)

Es sei f eine Funktion und $a \in \mathbb{R}$.

a) Ist a eine Extremstelle von f im Innern des Definitionsbereiches und ist f bei a differenzierbar, so ist a eine Nullstelle von f' .

b) Aber **Achtung!** Nullstellen von f' sind nicht unbedingt auch Extremstellen!

Der Beweis von a) ergibt sich unmittelbar aus der Definition von $f'(a)$. Wir nehmen einmal an, dass $f'(a) > 0$ ist. Dann muss auch der Differenzenquotient $\frac{f(x)-f(a)}{x-a}$, der sich dem positiven Wert $f'(a)$ ja immer mehr annähert, schließlich für x nahe bei a ebenfalls positiv sein:

$$\frac{f(x) - f(a)}{x - a} > 0 \quad \text{für } x \in \mathcal{D}_f \text{ hinreichend nahe bei } a.$$

Dies bedeutet, dass der Zähler $f(x) - f(a)$ und der Nenner $x - a$ in der Nähe von a stets dasselbe Vorzeichen haben. Ist also $x \in \mathcal{D}_f$ und $x < a$, so ist $f(x) < f(a)$, und ist $x > a$, so ist $f(x) > f(a)$. Da a im Innern des Definitionsbereiches liegt, gibt es Stellen $x \in \mathcal{D}_f$ mit $x < a$ und solche mit $x > a$. Damit erkennt man, dass für x in der Nähe von a sowohl Werte $f(x)$ ober- wie unterhalb von $f(a)$ auftreten; a kann also keine Extremstelle sein. Genauso schließt man für $f'(a) < 0$ und erhält insgesamt: Ist $f'(a) \neq 0$, so ist a keine Extremstelle von f , womit Satz (7.2) a) bewiesen ist.

Der Beweis von (7.2) zeigt, dass an Stellen a mit $f'(a) > 0$ die Funktion f wächst, während sie an Stellen a mit $f'(a) < 0$ fällt. Ist dagegen $f'(a) = 0$, so kann man nichts Allgemeingültiges über Wachstum oder Abnahme sagen. Man nennt daher die Nullstellen von f' auch *stationäre Stellen* von f :

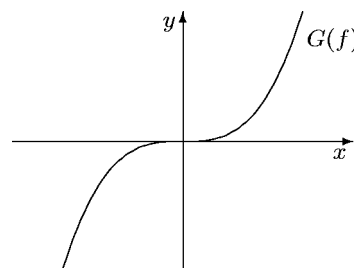
Stationäre Stelle von $f =$ Nullstelle von f' .

Den zugehörigen Punkt $(a, f(a))$ auf dem Graphen von f nennen wir entsprechend einen *stationären Punkt*. Die stationären Punkte eines Graphen sind also die Punkte, an denen die Tangente *waagrecht* verläuft. Damit kann man Satz (7.2) a) grob so formulieren:

Innere Extremstellen sind notwendig stationäre Stellen.

b) Aber nicht jede stationäre Stelle ist auch eine Extremstelle. Dies zeigt der nebenstehend skizzierte Graph der Funktion $f(x) = x^3$: An der Stelle $a = 0$ gilt $f'(a) = 3a^2 = 0$, also ist a stationäre Stelle von f . Aber der Punkt $(0, 0)$ ist kein Extrempunkt von f .

Man nennt einen solchen Punkt einen *Sattelpunkt*.



Um also die Extremstellen einer differenzierbaren Funktion zu ermitteln, muss man zunächst die stationären Stellen bestimmen. Dies sind gerade die Nullstellen der Ableitung f' . Ob aber bei einer Nullstelle a der Ableitung f' tatsächlich ein Extremum von f vorliegt, und welcher Art es ist (Maximum oder Minimum), hängt nun davon ab, welches Verhalten f *vor* und *hinter* der Stelle a hat. Dies wollen wir im nächsten Abschnitt genauer untersuchen.

b. Monotonieintervalle

(7.3) Definition: (Monotonie) Es sei f eine Funktion und I ein Intervall im Definitionsbereich.

a) Wir nennen f *monoton wachsend über I* , wenn für alle $x_1, x_2 \in I$ gilt:

$$x_1 < x_2 \implies f(x_1) \leq f(x_2).$$

Gilt in den Abschätzungen der Funktionswerte sogar $f(x_1) < f(x_2)$, so nennt man f *streng monoton wachsend*.

b) Entsprechend definiert man *monoton fallend über I* durch

$$x_1 < x_2 \implies f(x_1) \geq f(x_2).$$

c) Wir nennen eine Funktion f über I *monoton*, wenn sie entweder über ganz I monoton wachsend oder über ganz I monoton fallend ist.

Beweisen Sie zur Übung:

f ist (streng) monoton wachsend $\iff -f$ ist (streng) monoton fallend.

Diese Tatsache ermöglicht es, viele Aussagen über monotonen Wachstum unmittelbar auf monoton fallende Funktionen zu übertragen, ohne einen zweiten Beweis führen zu müssen.

Das entscheidende Mittel zur Untersuchung der Monotonie von Funktionen ist der folgende

(7.4) Satz: (Schwacher Monotoniesatz) *Ist f eine über einem Intervall I differenzierbare Funktion und hat die Ableitung f' über I nur strikt positive Werte: $f'(x) > 0$, so ist f über I streng monoton wachsend.*

Bevor wir diesen Satz beweisen, wollen wir seine Bedeutung für die Funktionsuntersuchungen darstellen. Aufgrund dieses Satzes ist die Untersuchung der *Monotonie* einer Funktion f gleichbedeutend mit der Untersuchung der *Vorzeichenverteilung* ihrer *Ableitungsfunktion f'* !

Monotonieuntersuchung von $f \leftrightarrow$ Vorzeichenuntersuchung von f' .

Wir können also unsere Überlegungen aus Kapitel II auf die Ableitung f' anwenden und so die Monotonieintervalle der Ausgangsfunktion f bestimmen. Man geht dazu folgendermaßen vor:

- 1) Die Ableitung f' berechnen.
- 2) Für die Ableitung f' die Nullstellen und die Vorzeichenverteilung bestimmen. (Für rationale Funktionen wie in Kapitel II, §3 ausführlich diskutiert.)
- 3) Gilt in einem Intervall I stets $f'(x) > 0$, so ist f in diesem Bereich streng monoton wachsend, während bei $f'(x) < 0$ die Funktionswerte von f über I strikt abnehmen müssen.

Mit Hilfe der Monotonieintervalle kann man nun auch die Extremstellen ermitteln. Nach Satz (7.2) kommen nur Nullstellen von f' als innere Extremstellen in Frage. Der folgende Satz gibt ein *hinreichendes* Kriterium dafür, wann eine stationäre Stelle tatsächlich Extremstelle ist.

(7.5) Satz: (Hinreichendes Extremstellenkriterium)

Sei f differenzierbar und a eine Stelle im Innern des Definitionsbereiches.

a) Hat f' bei a eine Nullstelle mit Vorzeichenwechsel, so besitzt f bei a ein Extremum. Genauer gilt:

Wechselt f' bei a sein Vorzeichen von '−' zu '+', so hat f bei a ein Minimum, wechselt f' bei a sein Vorzeichen von '+' zu '−', so hat f bei a ein Maximum.

b) Ist a zwar eine Nullstelle von f' , aber die Werte von f' in der Nähe einheitlich positiv oder einheitlich negativ (liegt also kein Vorzeichenwechsel vor), so hat f bei a auch kein Extremum, sondern einen sog. *Sattelpunkt*.

Beweis: a) folgt aus Satz (7.4): Gilt 'kurz vor' a , d. h. in einem kleinen Intervall $]b, a[$ (mit $b < a$), $f'(x) < 0$, so ist f in diesem Bereich gemäß Satz (7.4) streng monoton fallend, also $f(a)$ der kleinste Wert von f in diesem Bereich; gilt 'kurz nach' a , d. h. in einem kleinen Intervall $]a, c[$ (mit $c > a$), $f'(x) > 0$, so ist die Funktion f in diesem Bereich streng monoton steigend, also auch hier $f(a)$ der kleinste Wert. Damit hat man eine kleine 'Umgebung' $]b, c[$ von a gefunden, in der $f(a)$ der kleinste auftretende Wert ist. Das heißt aber nichts anderes, als dass f bei a ein lokales Minimum hat. Genauso argumentiert man für den anderen Vorzeichenwechsel und für c).

Anmerkung: Sieht man einmal von den Funktionen ab, deren Ableitungsfunktion f' überall den Wert 0 annimmt, so haben alle differenzierbaren Funktionen, die Ihnen im Mathematikunterricht begegnen werden, die folgende Eigenschaft: In der Nähe einer Nullstelle a von f' liegt

keine weitere Nullstelle von f' (die Nullstellen von f' liegen *isoliert*) und einer der in b) oder c) genannten Fälle liegt vor. Daher können wir für *differenzierbare* Funktionen f , deren Ableitung f' nur *isolierte* Nullstellen haben, Satz (7.5) kurz so formulieren:

Innere Extremstellen von $f =$ Nullstellen von f' mit Vorzeichenwechsel.

c. Der Monotoniesatz. Die auf diese Weise nicht erfassten Funktionen f mit Ableitungsfunktion $f'(x) = 0$ sind aber wohlbekannt, denn es gilt der folgende wichtige

(7.7) Satz: *Ist I ein Intervall (!) und f eine auf I differenzierbare Funktion, so sind folgende Aussagen äquivalent:*

- i) $f'(x) = 0$ für alle $x \in I$.
- ii) f ist auf I konstant: $f(x) = c$ für alle $x \in I$.

Dass aus ii) die Aussage i) folgt, war bereits in (6.8) bemerkt worden; Die umgekehrte Richtung i) \implies ii) ist der entscheidende Schritt. Dieser scheint anschaulich genauso klar, aber **Vorsicht:** Die Stärke des Satzes liegt gerade darin, dass er für *jede beliebige* Funktion gilt, insbesondere für die, für die wir eben noch keine *Anschauung* haben! Daher muss auch der Beweis frei von der Anschauung durchgeführt werden.

Für diesen Beweis benötigt man den *Monotoniesatz*. Hierbei genügt aber *nicht* der *schwache* Monotoniesatz (7.4). Vielmehr muss dieser verschärft werden zum vollen Monotoniesatz, wie er nachfolgend formuliert ist.

(7.8) Satz: (Monotoniesatz) *Gegeben sei eine Funktion f , die über einem Intervall I differenzierbar ist. Dann gilt die folgende Äquivalenz:*

$$f'(x) \geq 0 \text{ für alle } x \in I \iff f \text{ ist monoton wachsend über } I.$$

Entsprechendes gilt für $f'(x) \leq 0$ und monoton fallend.

Die Verschärfung gegenüber dem schwachen Monotoniesatz besteht darin, dass man als Voraussetzung nur benötigt, dass $f'(x)$ stets größer *oder gleich* 0 ist. Man kann dann zwar nicht mehr die strenge Monotonie folgern, wohl aber die Monotonie. Umgekehrt kann aus der Monotonie gefolgert werden, dass $f'(x) \geq 0$ sein muss. Man beachte, dass auch bei strenger Monotonie nicht überall $f'(x) > 0$ gelten muss. Gegenbeispiel ist wieder die Funktion $f(x) = x^3$, die über ganz \mathbb{R} streng monoton wächst, aber ihre Ableitung $f'(x) = 3x^2$ nimmt dennoch an einer Stelle den Wert 0 an.

Mit Hilfe des Monotoniesatzes (7.8) wird der *Beweis* von (7.7) sehr einfach: Ist $f'(x) = 0$ auf ganz I , so ist natürlich $f'(x) \geq 0$ und $f'(x) \leq 0$ auf ganz I , so dass nach dem Monotoniesatz für alle $b \geq a$ sowohl $f(b) \geq f(a)$ als auch $f(b) \leq f(a)$, also $f(b) = f(a)$ gilt. Damit ist f konstant.

d. Beweise.

(7.10) Bemerkung: *a) Eine Funktion f ist genau dann über einem Intervall I monoton wachsend, wenn alle Differenzenquotienten ≥ 0 sind:*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0 \text{ für alle } x_1, x_2 \in I, x_1 \neq x_2.$$

Und f ist streng monoton wachsend, wenn alle Differenzenquotienten > 0 sind.

b) Entsprechend ist f über I monoton fallend, wenn alle Differenzenquotienten ≤ 0 sind bzw. streng monoton fallend, wenn alle Differenzenquotienten < 0 sind.

Wir beweisen nur eine der vier Aussagen. Die übrigen Beweise verlaufen analog. Wir haben folgende Äquivalenzen:

$$\begin{aligned} & f \text{ ist streng monoton wachsend über } I \\ \iff & f(x_2) > f(x_1) \text{ für alle } x_1, x_2 \in I, x_2 > x_1 \\ \iff & f(x_2) - f(x_1) \text{ und } x_2 - x_1 \text{ haben gleiches Vorzeichen für alle } x_1, x_2 \in I \\ \iff & \frac{f(x_2) - f(x_1)}{x_2 - x_1} > 0 \text{ für alle } x_1, x_2 \in I, x_1 \neq x_2 \end{aligned}$$

Zum Beweis des schwachen Monotoniesatzes (7.4) müssen wir gemäß (7.10) zeigen:

Gilt auf einem Intervall I überall $f'(x) > 0$, so gilt für alle Differenzenquotienten $\frac{f(b)-f(a)}{b-a} > 0$ ($a, b \in I, a \neq b$).

Oder anders gewendet

Ist einer der Differenzenquotienten ≤ 0 , so muss an wenigstens einer Stelle z des Intervalls der Ableitungswert $f'(z) \leq 0$ sein.

Für den Beweis ist entscheidend, dass wir ein Intervall zugrundegelegt haben, und wir werden wesentlich die *Vollständigkeit* der reellen Zahlen \mathbb{R} benutzen!

Wir gehen also davon aus, dass wenigstens ein Differenzenquotient ≤ 0 ist. Es existieren daher $a, b \in I, a < b$ mit $\frac{f(b)-f(a)}{b-a} \leq 0$, also $f(b) \leq f(a)$. Wir betrachten nun den Mittelpunkt c zwischen a und b . Da I ein Intervall ist, liegt c in I (!) und $f(c)$ ist definiert. Nun gibt es 2 Möglichkeiten:

Entweder 1) $f(a) \geq f(c)$: Dann ist $a < c$ und $f(a) \geq f(c)$.

oder 2) $f(c) > f(a)$: Dann ist $c < b$ und $f(c) > f(a) \geq f(b)$.

In jedem Falle haben wir im Intervall $[a, b]$ ein halb so langes Intervall $[a_1, b_1]$ gefunden (nämlich $[a, c]$ oder $[c, b]$), bei dem wiederum $f(a_1) \geq f(b_1)$ gilt. Wir können nun dieses Verfahren der Intervallteilung fortsetzen und erhalten eine *Intervallschachtelung*

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \dots,$$

wobei stets $f(a_n) \geq f(b_n)$ ist. Aufgrund der Vollständigkeit von \mathbb{R} enthält diese Intervallschachtelung genau eine reelle Zahl, die wir z nennen wollen: $a_n \leq z \leq b_n$ für alle n und z ist der Grenzwert der Folgen (a_n) und (b_n) .

Wir wollen daraus nun *eine* Folge c_n konstruieren, die gegen z konvergiert und für die alle Differenzenquotienten

$$\frac{f(c_n) - f(z)}{c_n - z} \leq 0 \quad (*)$$

sind. Wir definieren das Folgenglied c_n als a_n oder b_n , jenachdem welcher der folgenden Fälle vorliegt:

1) $z = a_n$: Setzen wir dann $c_n = b_n$, so gilt $f(c_n) - f(z) = f(b_n) - f(a_n) \leq 0$ und $c_n - z = b_n - a_n > 0$, also ist (*) erfüllt.

2) $z = b_n$: Hier setzen wir $c_n = a_n$, und wieder gilt (*).

3) $f(a_n) \geq f(z)$ und $a_n < z$: Dann gilt $\frac{f(a_n) - f(z)}{a_n - z} < 0$, und wir setzen $c_n = a_n$.

4) $f(z) \geq f(b_n)$ und $z < b_n$: Dann ist $\frac{f(b_n) - f(z)}{b_n - z} < 0$, und wir setzen $c_n = b_n$.

Wegen $f(a_n) \geq f(b_n)$ und $a_n < b_n$ muss einer der 4 Fälle eintreten. Wir erhalten so eine Folge c_n , die wie a_n und b_n ebenfalls gegen z konvergiert, und für die alle Differenzenquotienten

$$\frac{f(c_n) - f(z)}{c_n - z} \leq 0$$

sind. Wegen $c_n \rightarrow z$ folgt dann durch Grenzübergang

$$f'(z) = \lim_{n \rightarrow \infty} \frac{f(c_n) - f(z)}{c_n - z} \leq 0.$$

Damit ist eine Stelle $z \in I$ gefunden mit $f'(z) \leq 0$, und der schwache Monotoniesatz ist bewiesen.

Wir kommen nun zum Beweis des Monotoniesatzes (7.8): Die Folgerung ii) \Rightarrow i) beweist man mittels Bemerkung (7.10): Ist f monoton wachsend, so sind für alle $x \neq a$ die Differenzenquotienten

$$\frac{f(x) - f(a)}{x - a} \geq 0.$$

Damit muss auch der Ableitungswert $f'(a)$, gegen den diese Differenzenquotienten konvergieren (für $x \rightarrow a$), ebenfalls ≥ 0 sein.

Die umgekehrte Richtung i) \Rightarrow ii) folgern wir aus dem schwachen Monotoniesatz. Sei also i) erfüllt, d. h. $f'(x) \geq 0$ für alle $x \in I$. Dann gilt für jede negative Zahl $m < 0$ natürlich $f'(x) > m$ ($x \in I$). Wir fixieren $m < 0$ und ein beliebiges $a \in I$ und betrachten die

$$\text{Hilfsfunktion } g(x) = f(x) - f(a) - m(x - a) :$$

Nun gilt nach den Ableitungsregeln $g'(x) = f'(x) - m$ für alle $x \in I$. Wegen $f'(x) > m$ ergibt sich dann für die Hilfsfunktion g

$$g'(x) > 0 \quad \text{für alle } x \in I .$$

Man kann also auf g den schwachen Monotoniesatz (7.4) anwenden und erhält: g ist über I streng monoton wachsend, also gilt für alle $b \in I, b \neq a$

$$0 < \frac{g(b) - g(a)}{b - a} = \frac{f(b) - f(a) - m(b - a)}{b - a} = \frac{f(b) - f(a)}{b - a} - m .$$

(Beachten Sie $g(a) = 0$.) Wir erhalten also für *jede negative* Zahl m :

$$m < \frac{f(b) - f(a)}{b - a} .$$

Dies bedeutet, dass die Zahl $\frac{f(b)-f(a)}{b-a}$ oberhalb *aller negativen* Zahlen m liegt, also selbst *nicht* negativ sein kann. Das heißt:

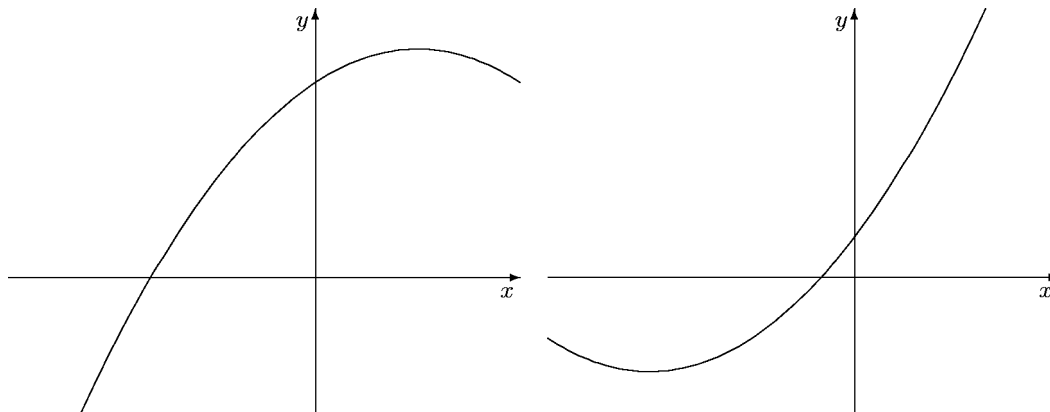
$$\frac{f(b) - f(a)}{b - a} \geq 0 .$$

Da diese Überlegungen für jedes $a \in I$ und für jedes $b \in I, b \neq a$ gelten, ist damit gemäß Bemerkung (7.10) der Monotoniesatz bewiesen.

§8 Höhere Ableitungen

Wir haben im vorhergehenden Paragraphen gesehen, dass man das Monotonieverhalten einer differenzierbaren Funktion f mit Hilfe ihrer Ableitung f' studieren kann. Und zwar ist das *Monotonieverhalten der Funktion f* durch die *Vorzeichenverteilung der Ableitungsfunktion f'* bestimmt. Insbesondere kann man dadurch die *Extremstellen der Funktion f* als die *Vorzeichenwechselstellen der Ableitungsfunktion f'* charakterisieren. Wir wollen uns nun mit weiteren Charakteristika des Graphen $G(f)$ einer Funktion f beschäftigen:

a. Krümmung und Wendestellen. Wir wollen den Begriff der *Krümmung* eines Funktionsgraphen mathematisch präzisieren. Wir veranschaulichen uns dazu einmal verschieden gekrümmte Funktionsgraphen:

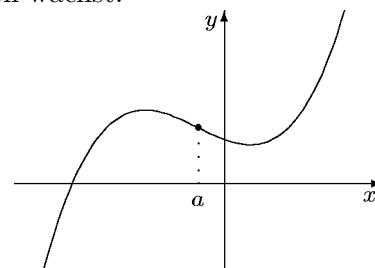


Der erste Graph hat offenbar eine *Rechtskrümmung*, während der zweite *linksgekrümmt* ist. Wenn man nun einmal die Änderung des Anstiegs bei beiden Funktionen vergleicht, so erkennt

man: Bei dem rechtsgekrümmten Graphen wird (mit zunehmendem x -Wert) der Anstieg immer geringer, schließlich sogar negativ! Beim linksgekrümmten zweiten Graphen hingegen wächst der Anstieg (bei zunehmender x -Koordinate). Dies führt zu der folgenden mathematisch präzisen Definition des Krümmungsbegriffes:

(8.1) Definition: a) Wir nennen eine differenzierbare Funktion f über einem Teilintervall ihres Definitionsbereiches *rechtsgekrümmt*, wenn ihre Ableitung f' über diesem Bereich monoton fällt, und *linksgekrümmt*, wenn f' in diesem Bereich monoton wächst.

b) Unter einer *Wendestelle* einer Funktion f versteht man solche Stellen a im Definitionsbereich, an denen die Funktion ihr Krümmungsverhalten ändert, das heißt: In einem kleinen Bereich 'vor' a ist f rechts- und kurz 'nach' a linksgekrümmt, oder umgekehrt.



Wir halten fest: Das *Krümmungsverhalten einer Funktion f* ist durch das *Monotonieverhalten ihrer Ableitung f'* bestimmt. Dementsprechend sind *Wendestellen einer Funktion f* solche Stellen a , an denen die Ableitung f' ihr Monotonieverhalten ändert, also *Extremstellen der Ableitung f'* .

Wendestellen von $f = \text{Extremstellen von } f'$.

Ist nun die Funktion f' selbst auch differenzierbar, so sind gemäß §7b. die Extremstellen von f' gerade die Vorzeichenwechselstellen der nächsten Ableitung $(f')'$.

Die Funktion $f'' = (f')'$ ist die *zweite* Ableitung von f . Entsprechend definiert man die *höheren Ableitungsfunktionen* $f^{(3)} = f''' = (f'')'$, $f^{(4)} = (f^{(3)})'$ usw.

Formulieren wir die obigen Überlegungen mit Hilfe dieser höheren Ableitungen, so erhalten wir:

(8.2) Satz: a) Eine zweimal differenzierbare Funktion f kann nur (muss jedoch nicht!) an solchen Stellen a eine Wendestelle haben, an denen die zweite Ableitung f'' eine Nullstelle hat: $f''(a) = 0$.

b) Wechselt die zweite Ableitung f'' bei a ihr Vorzeichen, so besitzt die Ausgangsfunktion f bei a eine Wendestelle.

c) Ist a zwar eine Nullstelle von f'' , aber die Werte von f'' in der Nähe einheitlich ≥ 0 oder einheitlich ≤ 0 (liegt also kein Vorzeichenwechsel von f'' vor), so hat f bei a auch keine Wendestelle.

Für zweimal differenzierbare Funktionen f , deren zweite Ableitung nur isolierte Nullstellen hat, kann man kurz sagen:

Die Funktion f hat genau dort ihre Wendestellen, wo die Ableitung f' Extremstellen hat, bzw. wo die zweite Ableitung f'' eine Nullstelle mit Vorzeichenwechsel hat.

Wendestellen von $f = \text{Nullstellen von } f'' \text{ mit Vorzeichenwechsel.}$

b. Vorzeichenwechsel und Ableitungen. Ob eine rationale Funktion f an einer Nullstelle a einen Vorzeichenwechsel hat, kann man an der Nullstellenordnung k (siehe Kap. II, Definition (3.5), S. 19) der Funktion f an der Stelle a erkennen. Diese Nullstellenordnung kann man leicht ablesen, wenn man *alle* Nullstellen von f bestimmt und den Funktionsterm in Linearfaktoren zerlegt hat. Kennt man jedoch nur diese eine Nullstelle a , so muss man gemäß der Definition den Funktionsterm $f(x)$ von f so oft wie möglich durch den Linearfaktor $x - a$ dividieren (mittels Polynomdivision) und so die Nullstellenordnung ermitteln. Wir wollen nun noch eine andere Methode kennenlernen, bei der man mit Hilfe der Ableitung Vorzeichenwechsel nachweisen kann. Diese Methode erfasst aber *nur einfache Nullstellen!* Sie benutzt die folgende

(8.3) Bemerkung: Es sei f eine differenzierbare Funktion. Dann gilt:

a) Ist $f(a) = 0$ und $f'(a) \neq 0$, so hat f bei a einen Vorzeichenwechsel. Genauer:

b) Ist $f(a) = 0$ und $f'(a) > 0$, so hat f bei a einen Vorzeichenwechsel von '−' nach '+'. Ist $f(a) = 0$ und $f'(a) < 0$, so hat f bei a einen Vorzeichenwechsel von '+' nach '−'.

Achtung: Ist $f'(a) = 0$, so kann man nichts über einen evtl. Vorzeichenwechsel von f an der Stelle a sagen. Die Bedingung $f'(a) \neq 0$ in (8.3) ist nur eine *hinreichende* Bedingung für das Vorliegen eines Vorzeichenwechsels, sie ist nicht *notwendig*. Sie besagt gerade, dass die Nullstelle a von f von *erster* Ordnung ist. Vorzeichenwechsel von f an Nullstellen 3., 5. ... Ordnung werden also durch (8.3) *nicht erfasst*. [Wir werden in §9 sehen, dass man jedoch mittels höherer Ableitungen auch die genaue Nullstellenordnung von f bei a bestimmen kann.]

Als Begründung für (8.3) ist nur zu bemerken: Ist $f'(a) > 0$, so steigt f an der Stelle a , d. h. 'kurz vor' a sind die Werte $f(x)$ kleiner als $f(a) = 0$ und 'kurz hinter' a gilt $f(x) > f(a) = 0$ (Beweis von Satz (7.3)). Damit ist b) bewiesen. Für c) argumentiert man genauso.

c. Hinreichende Kriterien für Extrem-/ Wendestellen mittels höherer Ableitungen. Extremstellen sind 'Vorzeichenwechselstellen' der ersten Ableitung f' , während Wendestellen 'Vorzeichenwechselstellen' der zweiten Ableitung f'' sind (Sätze (7.4), (8.2)). Nun wissen wir nach Bemerkung (8.3), dass für eine Funktion bei a ein Vorzeichenwechsel sicherlich dann vorliegt, wenn die Funktion bei a den Wert 0 hat, ihre Ableitung aber nicht! Für die Anwendungen muss man nun aber genau im Auge behalten, für welche Funktion man einen Vorzeichenwechsel feststellen will; für f oder f' oder f'' !

Beachtet man dies und kombiniert man nun (8.3) mit den Kriterien für Extrema (Satz (7.4)) bzw. für Wendestellen (Satz (8.2)), so erhält man die folgenden, häufig benutzten Kriterien:

(8.4) Satz: *Es sei f eine 2- bzw. 3-mal differenzierbare Funktion und $a \in \mathbb{R}$. Dann gelten die folgenden Aussagen:*

E1) f hat bei a eine Extremstelle $\implies f'(a) = 0$.

E2) $f'(a) = 0 \wedge f''(a) > 0 \implies f$ hat bei a ein Minimum.

E3) $f'(a) = 0 \wedge f''(a) < 0 \implies f$ hat bei a ein Maximum.

W1) f hat bei a eine Wendestelle $\implies f''(a) = 0$.

W2) $f''(a) = 0 \wedge f'''(a) \neq 0 \implies f$ hat bei a eine Wendestelle.

E1) und W1) sind wörtliche Wiederholungen von (7.4),a) bzw. (8.2),a). Für W2) beachte man, dass aus der Voraussetzung $f''(a) = 0 \wedge f'''(a) \neq 0$ nach (8.3) a) (angewendet auf die Funktion f'' !) folgt: f'' hat bei a einen Vorzeichenwechsel. Aus (8.2) b) folgt dann die Behauptung von W2).

Genauso folgen E2) und E3) mittels (8.3) b), c) (angewendet auf f') und (7.4) b).

Zum Schluss einige kritische Bemerkungen zu den im Schulbereich häufig verwendeten Kriterien E2), E3), W2) von Satz (8.4):

1. Im Falle $f'(a) = f''(a) = 0$ erlaubt (8.4) keine Entscheidung über das Vorliegen eines Extremums bei f . Dagegen ist oft ohne großen Aufwand zu erkennen, ob f' bei a einen Vorzeichenwechsel hat. Dies gilt insbesondere dann, wenn man alle Extremstellen von f bestimmen will, also ohnehin alle Nullstellen von f' bestimmen muss.

2. Gleiches gilt im Falle $f''(a) = f'''(a) = 0$ für Wendestellen.

3. Die Berechnung von f''' ist für ganzrationale Funktionen problemlos. Für die gebrochenrationalen Funktionen dagegen kann die Berechnung einer dritten Ableitung bereits erhebliche Probleme bereiten. Dagegen erfordert die Bestimmung der Nullstellenordnung von f'' bei a einen wesentlich geringeren Aufwand. (Man braucht nur die Funktion im Zähler zu untersuchen, und diese ist *ganzrational*! Siehe dazu die Beispiele für Kurvendiskussionen rationaler Funktionen.)

Fazit: *In aller Regel ist die Untersuchung eines eventuellen Vorzeichenwechsels (bei f' bzw. f'') vorteilhafter als die unreflektierte Anwendung der Kriterien von Satz (8.4).*

§9 Weitere Ableitungsregeln

a. Produkt- und Quotientenregel. Bisher haben wir nur für Potenzfunktionen und darauf aufbauend mittels der Faktor- und Summenregel (Satz (6.6)) für alle ganz-rationale Funktionen die Ableitungen bestimmen können. Um alle rationalen Funktionen differenzieren zu können, benötigen wir Ableitungsregeln für Produkte und vor allem Quotienten von Funktionen.

(9.1) Satz: *Seien g und h zwei differenzierbare Funktionen.*

a) (Produktregel) *Ist $f = g \cdot h$ das Produkt von g und h , d. h. gegeben durch $f(x) = g(x) \cdot h(x)$,*

so ist f differenzierbar und die Ableitung berechnet sich als

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

b) (Quotientenregel) Ist $f = \frac{g}{h}$ der Quotient von g und h d. h. gegeben durch $f(x) = \frac{g(x)}{h(x)}$, so ist die Funktion f (an allen Stellen ihres Definitionsbereiches) differenzierbar und die Ableitung berechnet sich als

$$f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{(h(x))^2}.$$

Merkregeln:

Ableitung eines Produktes: 'Ableitung des ersten Faktors multipliziert mit zweitem Faktor plus erster Faktor multipliziert mit der Ableitung des zweiten Faktors.'

Ableitung eines Quotienten: 'Ableitung des Zählers multipliziert mit dem Nenner minus Zähler multipliziert mit der Ableitung des Nenners, und das Ganze dividiert durch das Quadrat des Nenners.'

Als wichtige Folgerung von b) halten wir fest, dass alle rationalen Funktionen (auf ihrem Definitionsbereich) differenzierbar sind.

Beweis: a) Zum Beweis muss man den Differenzenquotienten $\frac{f(x)-f(a)}{x-a}$ berechnen und den Grenzwert für $x \rightarrow a$ untersuchen. Nun gilt

$$\begin{aligned} \frac{f(x) - f(a)}{x - a} &= \frac{g(x)h(x) - g(a)h(a)}{x - a} \\ &= \frac{g(x)h(x) - g(a)h(x) + g(a)h(x) - g(a)h(a)}{x - a} \\ &= \frac{g(x)h(x) - g(a)h(x)}{x - a} + \frac{g(a)h(x) - g(a)h(a)}{x - a} \\ &= \frac{g(x) - g(a)}{x - a} \cdot h(x) + g(a) \cdot \frac{h(x) - h(a)}{x - a} \end{aligned}$$

Durch diese Umformungen haben wir den Differenzenquotienten von f durch die Differenzenquotienten von g und h ausgedrückt. Deren Verhalten für $x \rightarrow a$ ist bekannt, da g und h differenzierbar sind: Ihre Differenzenquotienten streben gegen den jeweiligen Ableitungswert $g'(a)$ bzw. $h'(a)$. Da h bei a stetig ist (Satz (6.4)), gilt $\lim_{x \rightarrow a} h(x) = h(a)$ und man erhält aus den Grenzwertsätzen insgesamt

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

Zu b): Es sei a eine Stelle des Definitionsbereichs von f , also mit $h(a) \neq 0$. Wieder formen wir zunächst den Differenzenquotienten von f so um, dass er durch die Differenzenquotienten von g und h ausgedrückt wird. Zunächst ist

$$\begin{aligned} f(x) - f(a) &= \frac{g(x)}{h(x)} - \frac{g(a)}{h(a)} = \frac{g(x)h(a) - g(a)h(x)}{h(x)h(a)} \\ &= \frac{1}{h(x)h(a)} \cdot (g(x)h(a) - g(a)h(x) + g(a)h(a) - g(a)h(x)) \\ &= \frac{1}{h(x)h(a)} \cdot ((g(x) - g(a))h(a) - g(a)(h(x) - h(a))), \end{aligned}$$

und nach Division durch $x - a$ erhält man

$$\frac{f(x) - f(a)}{x - a} = \frac{1}{h(x)h(a)} \cdot \left(\frac{g(x) - g(a)}{x - a} h(a) - g(a) \frac{h(x) - h(a)}{x - a} \right).$$

Da h an der Stelle a stetig ist (Satz (6.4)) und $h(a) \neq 0$ gilt, folgt $h(x) \rightarrow h(a) \neq 0$ und man erhält unter Anwendung der Grenzwertsätze

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \frac{1}{h(a)h(a)} \cdot (g'(a)h(a) - g(a)h'(a)),$$

was zu beweisen war.

b. Kettenregel. Eine weitere wichtige Regel betrifft die Ableitung von Funktionen, die durch sog. Verkettung entstehen:

(9.2) Satz: (Kettenregel) *Ist f die Verkettung zweier differenzierbarer Funktionen g und h , also gegeben durch $f(x) = g(h(x))$, d. h. durch Einsetzung von $h(x)$ in g , so ist auch f differenzierbar und die Ableitung berechnet sich als*

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

Merkregel: Ableitung der Einsetzung einer ('inneren') Funktion in eine ('äußere') Funktion: 'Äußere Funktion ableiten und innere einsetzen, mal innerer Ableitung'; noch prägnanter, aber auch unpräziser: 'Äußere Ableitung mal innerer Ableitung'.

Beweis: Wir wollen den Beweis nicht in voller Allgemeinheit führen. Zur Vereinfachung nehmen wir an, dass h bei a streng monoton ist, dass also $h(x) \neq h(a)$ angenommen werden kann. (Dies gilt für die schon früher betrachteten Funktionen mit isolierten Nullstellen der Ableitung.) Wir erhalten unter dieser Annahme

$$\frac{f(x) - f(a)}{x - a} = \frac{g(h(x)) - g(h(a))}{x - a} = \frac{g(h(x)) - g(h(a))}{h(x) - h(a)} \cdot \frac{h(x) - h(a)}{x - a}.$$

Für $x \rightarrow a$ konvergiert der zweite Faktor gegen $h'(a)$, während im ersten Faktor ein Ausdruck der Form $\frac{g(z) - g(b)}{z - b}$ steht mit $z = h(x)$ und $b = h(a)$. Wegen der Stetigkeit von h (Satz (6.4)) an der Stelle a folgt aus $x \rightarrow a$ sofort $z = h(x) \rightarrow h(a) = b$. Für $z \rightarrow b$ konvergiert dann aber der Differenzenquotient $\frac{g(z) - g(b)}{z - b}$ gegen die Ableitung $g'(b) = g'(h(a))$. Also:

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = g'(h(a)) \cdot h'(a).$$

(9.3) Folgerung: a) *Es sei $m \in \mathbb{Z}$ eine ganze Zahl, $m \neq 0$ und $f(x) = (x - a)^m g(x)$ mit einer differenzierbaren Funktion g . Dann ist f differenzierbar und es gilt $f'(x) = (x - a)^{m-1} (mg(x) + g'(x)(x - a))$.*

b) *Hat eine rationale Funktion f bei a eine Nullstelle der Ordnung $k \geq 2$, so hat f' bei a eine Nullstelle der Ordnung $k - 1$.*

c) *Hat f bei a eine einfache Nullstelle, so gilt $f'(a) \neq 0$.*

d) *Hat eine rationale Funktion f bei a eine Polstelle der Ordnung $k \geq 1$, so hat f' bei a eine Polstelle der Ordnung $k + 1$.*

Zum *Beweis* von a) folgern wir zunächst aus der Ketten- und allgemeinen Potenzregel (6.4'), dass $(x - a)^m$ als Ableitung die Funktion $m(x - a)^{m-1}$ hat (die Ableitung der inneren Funktion $x - a$ ist 1!). Mit der Produktregel folgt dann

$$f'(x) = m(x - a)^{m-1} \cdot g(x) + (x - a)^m \cdot g'(x) = (x - a)^{m-1} (mg(x) + (x - a)g'(x)).$$

Für den Beweis von b)/c) wiederhole man zunächst die Definition der Nullstellenordnung (II, §3, (3.5)). Ist a eine Nullstelle von f von der Ordnung k , so gilt gemäß dieser Definition

$f(x) = (x - a)^k g(x)$ mit $g(a) \neq 0$. (Andernfalls könnte man noch weitere Faktoren $(x - a)$ aus $f(x)$ abspalten.) Dann folgt nach a) $f'(x) = (x - a)^{k-1} \cdot h(x)$, wobei wir abkürzend $h(x) = kg(x) + (x - a)g'(x)$ schreiben. Wegen $h(a) = kg(a) + 0 \cdot g'(a) = kg(a) \neq 0$ hat dann für $k - 1 \geq 1$ die Funktion $f'(x) = (x - a)^{k-1} h(x)$ bei a eine Nullstelle der genauen Ordnung $k - 1$, während bei $k - 1 = 0$ folgt $f'(a) = h(a) \neq 0$.

d) Hat die rationale Funktion f bei a eine Polstelle der Ordnung k , so kann man f folgendermaßen darstellen:

$$f(x) = \frac{g(x)}{(x - a)^k}$$

mit einer rationalen Funktion g , die bei a nicht den Wert 0 annimmt (sonst könnte man $x - a$ abspalten, dann kürzen und die Polordnung wäre geringer!). Also gilt $f(x) = (x - a)^{-k} \cdot g(x)$ und daher (wie oben)

$$f'(x) = (x - a)^{-k-1}((-k)g(x) + (x - a)g'(x)) = \frac{h(x)}{(x - a)^{k+1}} \quad \text{mit } h(a) \neq 0.$$

Also hat f' bei a ebenfalls einen Pol, jetzt jedoch mit der Ordnung $k + 1$.

Als Konsequenz dieser Überlegungen erhalten wir die folgende Methode zur Bestimmung der Nullstellenordnung mittels höherer Ableitungen:

(9.5) Satz: *Hat f bei a eine Nullstelle, so ist die Nullstellenordnung von f bei a die kleinste Zahl $k \in \mathbb{N}$ mit $f^{(k)}(a) \neq 0$, oder anders formuliert: die Zahl $k \in \mathbb{N}$ mit*

$$0 = f(a) = f'(a) = f''(a) = \dots = f^{(k-1)}(a) \quad \text{und} \quad f^{(k)}(a) \neq 0. \quad (*)$$

Beweis: Zunächst folgt aus (9.3), b): Ist $f(a) = 0$ und $f'(a) \neq 0$, so kann f bei a keine Nullstelle einer Ordnung ≥ 2 haben, also muss die Nullstelle einfach sein. Mit (9.3) c) folgt, dass f bei a *genau dann* eine einfache Nullstelle hat, wenn $f(a) = 0$ und $f'(a) \neq 0$ ist. Damit ist (9.5) für $k = 1$ bewiesen.

Hat nun f bei a eine Nullstelle der Ordnung $k \geq 2$, so hat f' nach (9.3) b) bei a eine Nullstelle der Ordnung $k - 1$, f'' dann eine Nullstelle der Ordnung $k - 2$, usw. bis: $f^{(k-1)}$ hat bei a eine einfache Nullstelle, also $f^{(k)}(a) \neq 0$. Damit ist (*) gezeigt.

Warnung! (9.3) b) ist nicht umkehrbar: Hat f' bei a eine Nullstelle der Ordnung m , so folgt *nicht*, dass f bei a eine Nullstelle der Ordnung $m + 1$ hat. Vielmehr braucht f dort überhaupt keine Nullstelle zu haben! $f(x) = 3x^4$ und $g(x) = 3x^4 + 1$ bestimmen zwei verschiedene Funktionen mit derselben Ableitung $f'(x) = g'(x) = 12x^3$. Diese Ableitung hat bei 0 eine dreifache Nullstelle. Während zwar f bei 0 eine vierfache Nullstelle besitzt, hat g überhaupt keine Nullstelle, da alle Werte von g positiv sind. Hat f' bei a eine m -fache Nullstelle, so hat f bei a eine $(m + 1)$ -fache Nullstelle, *nur wenn a auch wirklich Nullstelle von f ist!*

IV. Transzendente Funktionen

Bisher haben wir die folgenden Funktionenklassen studiert: die *ganzrationalen* Funktionen, dann die *rationalen* (definiert als Quotienten ganz-rationaler) und schließlich noch Funktionen, die aus Wurzelfunktionen zusammengesetzt waren (sie bezeichnet man auch als *algebraische* Funktionen). Wir wollen nun wichtige Vertreter der Klasse der transzendenten (= nicht algebraischen) Funktionen kennenlernen: die Exponentialfunktionen, insbesondere die Euler'sche Funktion, sodann damit verbunden die Logarithmusfunktionen und schließlich die ebenso wichtige Gruppe der trigonometrischen Funktionen (Sinus, Kosinus, Tangens und deren Umkehrfunktion Arkustangens, etc.). Ihnen allen ist gemein, dass man sie nicht durch algebraische Rechenoperationen definieren kann; zu ihrer Definition gehört untrennbar der Grenzwertbegriff.

§10 Exponential- und Logarithmusfunktionen

a. Die Exponentialfunktionen. Es sei im Folgenden b eine *positive* reelle Zahl. Dann haben wir in der Einführungsphase Potenzen b^x von b für *beliebige reelle Exponenten* x definiert. Diese Definition erfolgte schrittweise:

1. $x = n \in \mathbb{N}_0$: Dann definierte¹⁾ man die Potenz b^n als das Produkt aus n gleichen Faktoren b

$$b^n = \underbrace{b \cdot \dots \cdot b}_{n\text{-mal}}$$

und setzte (aus guten Gründen) $b^0 = 1$.

Auf der Grundlage dieser Definition erkennt man unmittelbar die folgenden fundamentalen Potenzgesetze

$$\begin{aligned} (P1) \quad & b^{n+m} = b^n \cdot b^m, \\ (P2) \quad & (b^n)^m = b^{nm}, \\ (P3) \quad & \text{Für } b > 1: \quad n < m \iff b^n < b^m, \end{aligned}$$

gültig für alle natürlichen Zahlen $n, m \in \mathbb{N}_0$.

2. $x = n \in \mathbb{Z}$: Man erweiterte die obige Definition durch die folgende Festsetzung²⁾

$$b^{-n} = \frac{1}{b^n}$$

für $n \in \mathbb{N}$. Diese Definition war nicht willkürlich, sondern *zwangsläufig*, wenn die Eigenschaft (P1) gültig bleiben sollte. Denn gemäß (P1) muss gelten

$$b^{-n} \cdot b^n = b^{-n+n} = b^0 = 1, \text{ also } b^{-n} = \frac{1}{b^n}.$$

Man kann dann zeigen, dass mit dieser erweiterten Definition die Eigenschaften (P1) – (P3) gültig bleiben (für alle $n, m \in \mathbb{Z}$).

3. $x = r = \frac{n}{m} \in \mathbb{Q}$: Für beliebige rationale Exponenten $r = n/m$ mit $n \in \mathbb{Z}, m \in \mathbb{N}$ erweiterte man die Definition zu³⁾

$$b^{\frac{n}{m}} = \sqrt[m]{b^n}.$$

Auch diese Definition war *zwangsläufig*, wenn die Eigenschaft (P2) gültig bleiben sollte, denn gemäß (P2) muss gelten

$$\left(b^{\frac{n}{m}}\right)^m = b^{\frac{n}{m} \cdot m} = b^n, \text{ also } b^{\frac{n}{m}} = \sqrt[m]{b^n}.$$

Da b als positiv vorausgesetzt war, existiert die m -te Wurzel stets. Man kann dann wieder zeigen, dass mit dieser erweiterten Definition die Eigenschaften (P1) – (P3) für beliebige Exponenten in \mathbb{Q} gültig bleiben.

4. $x \in \mathbb{R}$: Eine reelle Zahl x ist gegeben durch eine Intervallschachtelung von rationalen Zahlen r_i, s_i :

$$[r_1, s_1] \supset [r_2, s_2] \supset [r_3, s_3] \supset \dots \ni x,$$

in der nur x enthalten ist. Wegen der Eigenschaft (P3) bilden dann auch die Potenzen von b (wir setzen im Moment $b > 1$ voraus)

$$[b^{r_1}, b^{s_1}] \supset [b^{r_2}, b^{s_2}] \supset \dots$$

¹⁾ Diese Definition machte noch für jede beliebige Basis $b \in \mathbb{R}$ Sinn.

²⁾ Wegen des Nenners b^n wird hier benötigt, dass $b \neq 0$ ist.

³⁾ Wegen der Wurzel $\sqrt[m]{b^n}$ wird hier benötigt, dass $b > 0$ ist.

eine ineinanderliegende Folge von Intervallen. Wenn die Eigenschaft (P3) gültig bleiben soll, muss die gesuchte Zahl b^x in der letztgenannten Intervallschachtelung liegen:

$$[b^{r_1}, b^{s_1}] \supset [b^{r_2}, b^{s_2}] \supset \dots \ni b^x.$$

Wegen der Vollständigkeit der reellen Zahlen, *existiert* eine solche Zahl. Man kann (und muss) nun zeigen, dass es *nur eine* Zahl in dieser Intervallschachtelung gibt, und definiert dann dadurch b^x . Außerdem zeigt man, dass wieder die drei fundamentalen Eigenschaften (P1) – (P3) gültig bleiben, jetzt für beliebige reelle Exponenten.

Den Fall $b < 1$ führt man mittels $b^x = (\frac{1}{b})^{-x}$ auf die Basis $\frac{1}{b} > 1$ zurück.

Unter Verwendung des uns jetzt zur Verfügung stehenden Grenzwertbegriffes kann man den vierten Schritt der Definition von b^x auch so beschreiben: Jede reelle Zahl x ist Grenzwert einer Folge von rationalen Zahlen r_n (etwa der endlichen Abschnitte r_n der Dezimalzahldarstellung von x). Dann ist b^x der Grenzwert der Folge der Potenzen b^{r_n} :

$$x = \lim_{n \rightarrow \infty} r_n \implies b^x = \lim_{n \rightarrow \infty} b^{r_n}.$$

Man sagt auch: Die auf \mathbb{Q} bereits definierte Funktion $f(x) = b^x$ wird *stetig* auf \mathbb{R} fortgesetzt.

Auf diese Weise ist nun für jede positive Basis $b > 0$ eine Funktion f definiert durch $f(x) = b^x$. Man nennt diese Funktion die *Exponentialfunktion zur Basis b* . (Im Kontrast zu den Potenzfunktionen ist hier die Basis fest und der Exponent variabel.)

Bei den nun folgenden Untersuchungen der Exponentialfunktionen wird man nicht ständig die obige mehrstufige Definition benötigen, vielmehr benutzt man nur die folgenden Tatsachen (aus denen sich aber, wie wir bereits gesehen haben, die frühere Definition zwangsläufig ergibt):

(1.1) Satz: Die Exponentialfunktion \exp_b zur Basis $b > 0$ ist definiert durch $\exp_b(x) = b^x$ und hat die folgenden Eigenschaften:

- a) $b^x > 0$ für alle $x \in \mathbb{R}$, d. h. \exp_b nimmt nur positive Werte an, hat insbesondere keine Nullstellen,
- b) $b^0 = 1$, $b^1 = b$,
- c) $b^{x+y} = b^x \cdot b^y$ für alle $x, y \in \mathbb{R}$,
- d) $(b^x)^y = b^{xy}$ für alle $x, y \in \mathbb{R}$,
- e) für $b > 1$ gilt: $x < y \iff b^x < b^y$,
d. h. \exp_b ist streng monoton wachsend,
- f) für $b > 1$ gilt:

$$\lim_{x \rightarrow \infty} b^x = \infty, \quad \lim_{x \rightarrow -\infty} b^x = 0.$$

Die Eigenschaften a), b), c), d) und e) ergeben sich aus der Definition sowie der Gültigkeit der Potenzgesetze (P1) – (P3). Für f) brauchen wir nur die erste Behauptung zu überprüfen, denn wenn $b^x \rightarrow \infty$ für $x \rightarrow \infty$ gilt, so folgt daraus $\frac{1}{b^x} \rightarrow 0$ für $x \rightarrow \infty$, also

$$\lim_{x \rightarrow -\infty} b^x = \lim_{x \rightarrow \infty} b^{-x} = \lim_{x \rightarrow \infty} \frac{1}{b^x} = 0.$$

Für die erste Behauptung in f) muss man wegen der Monotonie lediglich zeigen, dass bei $b > 1$ die Folge der Potenzen (b^n) unbeschränkt ist. Dies ist anschaulich plausibel. Man beweist es formal korrekt mit Hilfe der ‘Bernoullischen Ungleichung’ $b^n = (1 + (b - 1))^n \geq 1 + n(b - 1)$ für alle $n \in \mathbb{N}$. Für $b > 1$, also $b - 1 > 0$, wächst dann bereits die untere Schranke $1 + n(b - 1)$ über alle Grenzen, also erst recht b^n . Die Bernoullische Ungleichung selbst haben wir mittels ‘vollständiger Induktion’ bewiesen.

Übung: Formulieren Sie die zu e) und f) analogen Aussagen für $b < 1$ und folgern Sie sie aus Satz (1.1).

b. Differenzierbarkeit der Exponentialfunktionen. In diesem Abschnitt wollen wir die analytischen Eigenschaften der Exponentialfunktionen (wie Stetigkeit und Differenzierbarkeit) untersuchen. Wir betrachten unmittelbar die Differenzierbarkeit und studieren daher das Verhalten des Differenzenquotienten der Funktion $f(x) = b^x$ zur Stelle a :

$$\frac{f(x) - f(a)}{x - a} = \frac{b^x - b^a}{x - a}.$$

Setzt man darin $h = x - a$, also $x = a + h$, und benutzt die fundamentale Eigenschaft aller Exponentialfunktionen ($b^{x+y} = b^x \cdot b^y$), so erhält man

$$\frac{b^x - b^a}{x - a} = \frac{b^{a+h} - b^a}{h} = \frac{b^a b^h - b^a}{h} = b^a \cdot \frac{b^h - 1}{h}.$$

Wenn man nun den Grenzübergang $x \rightarrow a$ durchführt, so bedeutet dies dasselbe wie $h \rightarrow 0$. Da im letzten Term der Faktor b^a unabhängig von h ist, erhält man

$$f'(a) = \lim_{x \rightarrow a} \frac{b^x - b^a}{x - a} = b^a \cdot \lim_{h \rightarrow 0} \frac{b^h - 1}{h}. \quad (*)$$

Damit ist die Differenzierbarkeit der Exponentialfunktion $f(x) = b^x$ (unabhängig von der betrachteten Stelle a) auf die Existenz des einen Grenzwertes $\lim_{h \rightarrow 0} \frac{b^h - 1}{h}$ zurückgeführt. Den Nachweis der Existenz dieses Grenzwertes wollen wir auf einen späteren Abschnitt verschieben. Wir setzen im Folgenden voraus, dass dieser Grenzwert existiert, und bezeichnen ihn mit $L(b)$:

$$L(b) = \lim_{h \rightarrow 0} \frac{b^h - 1}{h}.$$

Die Bedeutung dieses Grenzwertes erkennt man, wenn man in obiger Formel (*) $a = 0$ setzt:

$$f'(0) = b^0 \cdot L(b) = L(b);$$

$L(b)$ ist also gerade der Ableitungswert der Exponentialfunktion $f = \exp_b$ an der Stelle 0. Damit haben wir gezeigt:

(1.2) Satz: Wenn die Exponentialfunktion \exp_b zu einer Basis $b > 0$ an der Stelle 0 differenzierbar ist, wenn also der Grenzwert

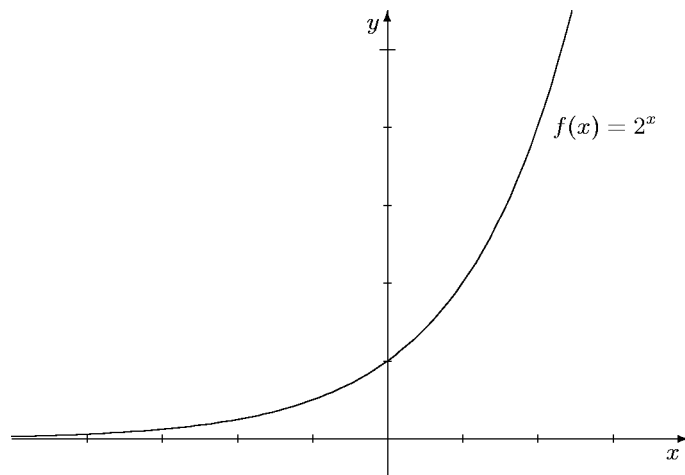
$$L(b) = \lim_{h \rightarrow 0} \frac{b^h - 1}{h}$$

existiert, dann ist \exp_b an allen Stellen $a \in \mathbb{R}$ differenzierbar und der Ableitungswert ist $\exp'_b(a) = L(b) \cdot \exp_b(a)$. Damit gilt also:

$$\exp'_b = L(b) \cdot \exp_b \quad \text{mit} \quad L(b) = \exp'_b(0).$$

Die Ableitung einer Exponentialfunktion ist ein festes Vielfaches der Exponentialfunktion selbst.

c. Logarithmusfunktionen. Wir betrachten im Folgenden der Einfachheit halber nur Basen $b > 1$ (alle Resultate gelten in analoger Form auch für $b < 1$, nicht aber für $b = 1$). Nach (1.1), e) wissen wir, dass die Exponentialfunktion streng monoton wächst. Außerdem wissen wir aufgrund der Grenzwertaussagen (1.1), f), dass b^x sowohl beliebig große positive Werte annimmt als auch positive Werte, die beliebig nahe bei 0 liegen. Dann wird auch jeder dazwischenliegende, also jeder positive Wert als Funktionswert b^x angenommen, denn nach Satz (1.2) (Existenz von $L(b)$ vorausgesetzt) ist \exp_b differenzierbar, also erst recht stetig (III, (4.1)). Der Graph weist also keine Lücken auf (Zwischenwertsatz II, (2.4)). Damit ergibt sich der typische Verlauf einer Exponentialfunktion mit $b > 1$, wie er in der folgenden Skizze für $b = 2$ dargestellt ist.



Es ist also *jede positive* Zahl y als Potenz b^x von b darstellbar: $y = b^x$. Und wegen (1.1) e) ist dabei der Exponent x eindeutig bestimmt; dieser ist der sogenannte *Logarithmus von y zur Basis b* :

Der Logarithmus $\log_b(y)$ einer positiven Zahl y zur Basis b ist der (eindeutig) bestimmte Exponent $x \in \mathbb{R}$ mit $b^x = y$:

$$x = \log_b(y) \iff y = b^x .$$

Ist also $y = b^x = \exp_b(x)$, so ist umgekehrt $x = \log_b(y)$; die Logarithmusfunktion \log_b ist die Umkehrfunktion zur Exponentialfunktion \exp_b .

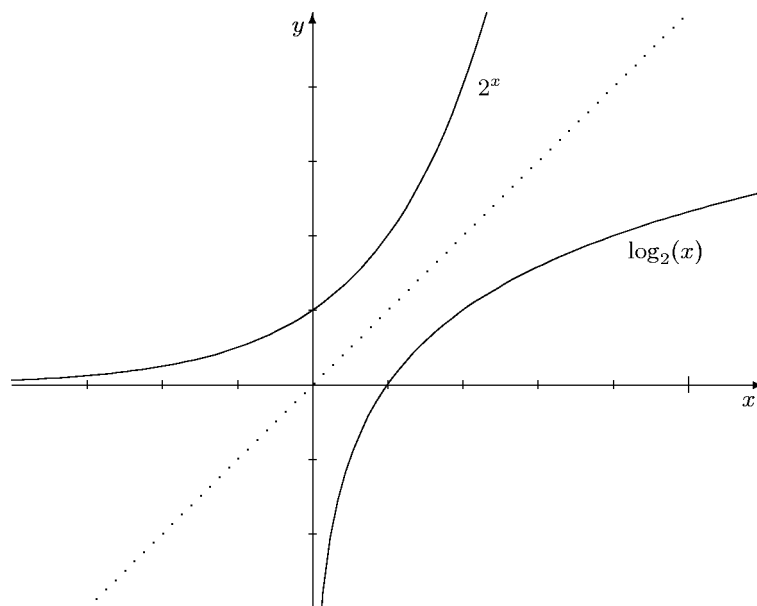
Aus den Eigenschaften (1.1) der Exponentialfunktionen entnimmt man daher unmittelbar entsprechende Eigenschaften für die Logarithmusfunktionen:

(1.3) Satz: Die Logarithmusfunktion \log_b zur Basis $b > 1$ hat die folgenden Eigenschaften:

- a) \log_b hat den Definitionsbereich $]0, \infty[$.
- b) $\log_b(1) = 0$, $\log_b(b) = 1$,
- c) $\log_b(xy) = \log_b(x) + \log_b(y)$ für $x, y > 0$,
- d) $\log_b(x^r) = r \cdot \log_b(x)$ für $x > 0$, $r \in \mathbb{R}$,
- e) für $x, y > 0$ gilt: $x < y \iff \log_b(x) < \log_b(y)$,
d. h. \log_b ist streng monoton wachsend,
- f)

$$\lim_{x \searrow 0} \log_b(x) = -\infty, \quad \lim_{x \rightarrow \infty} \log_b(x) = \infty .$$

Den Graphen der Logarithmusfunktionen gewinnt man aus dem Graphen der Exponentialfunktion, indem man die Koordinaten x und y vertauscht. Dies bedeutet, dass die beiden Graphen spiegelbildlich zueinander sind bzgl. der Winkelhalbierenden des I./III. Quadranten:



Aufgrund der engen Beziehung zwischen Exponential- und Logarithmusfunktion können wir nun auch letztere differenzieren. Dazu betrachten wir den Differenzenquotienten

$$\frac{\log_b(x) - \log_b(a)}{x - a}.$$

Wir führen vorübergehend folgende Bezeichnungen ein: $y = \log_b(x)$ und $c = \log_b(a)$. Dann gilt aufgrund der Definition des Logarithmus $x = b^y$ und $a = b^c$, und der obige Differenzenquotient nimmt folgende Gestalt an:

$$\frac{\log_b(x) - \log_b(a)}{x - a} = \frac{y - c}{b^y - b^c} = \frac{1}{\frac{b^y - b^c}{y - c}}. \quad (*)$$

Der Nenner dieses letzten Doppelbruches ist aber nichts anderes als der Differenzenquotient der Exponentialfunktion zur Basis b .

Wir müssen nun den Grenzübergang $x \rightarrow a$ durchführen und zeigen zunächst: Strebt eine Folge x_n gegen a , so strebt $y_n = \log_b(x_n)$ gegen $\log_b(a) = c$:

$$x_n \rightarrow a \implies y_n = \log_b(x_n) \rightarrow \log_b(a) = c.$$

(Wir zeigen also zunächst die *Stetigkeit* von \log_b und dann erst die Differenzierbarkeit.)

Sei also x_n eine beliebige Folge, die gegen a konvergiert. Um zu zeigen, dass y_n gegen c konvergiert, müssen wir zeigen, dass für jede positive Zahl $\varepsilon > 0$ die Abschätzung $c - \varepsilon < y_n < c + \varepsilon$ für fast alle n richtig ist (vgl. Definition der Konvergenz in I., §2, S. 7). Aufgrund der *strengen* Monotonie der Exponentialfunktionen gilt

$$b^{c-\varepsilon} < b^c < b^{c+\varepsilon}.$$

Da die Folge x_n gegen $a = b^c$ konvergiert, folgt daraus

$$b^{c-\varepsilon} < x_n < b^{c+\varepsilon} \text{ für fast alle } n.$$

Wendet man nun auf diese letzte Ungleichung die streng *monotone* Logarithmusfunktion \log_b an, so erhält man

$$\log_b(b^{c-\varepsilon}) < \log_b(x_n) < \log_b(b^{c+\varepsilon}) \text{ für fast alle } n.$$

Dies heißt aber nichts anderes als

$$c - \varepsilon < y_n < c + \varepsilon \quad \text{für fast alle } n.$$

Damit haben wir gezeigt: Strebt x gegen a , so strebt y gegen $\log_b(a) = c$. Dann strebt aber der Nenner in (*) (dies ist der Differenzenquotient der Exponentialfunktion \exp_b an der Stelle c) gegen den Ableitungswert $\exp'_b(c) = \exp_b(c) \cdot L(b)$. Da dieser Wert $\neq 0$ ist, erhalten wir aus dem Grenzwertsatz für Quotienten:

$$\lim_{x \rightarrow a} \frac{\log_b(x) - \log_b(a)}{x - a} = \frac{1}{\lim_{y \rightarrow c} \frac{by - bc}{y - c}} = \frac{1}{L(b) \cdot bc} = \frac{1}{L(b) \cdot a}.$$

Damit ist der folgende Satz gezeigt:

(1.4) Satz: Für alle $b > 1$ ist die Logarithmusfunktion \log_b auf ihrem Definitionsbereich $]0, \infty[$ differenzierbar und es gilt

$$\log'_b(x) = \frac{1}{L(b)} \cdot \frac{1}{x}.$$

§11 Der natürliche Logarithmus und die Eulersche Zahl e .

a. Der natürliche Logarithmus. Wir wollen nun die bei der Ableitung der Exponentialfunktionen aufgetretene Größe $L(b)$ (definiert für alle $b > 0$) genauer analysieren. Als Ableitungsübungen differenzieren wir einmal die Funktionen $f(x) = a^x b^x$ ($a, b > 0$) und $g(x) = a^{rx}$ ($a > 0, r \in \mathbb{R}$). Wir erhalten mit der Produkt- bzw. Kettenregel

$$\begin{aligned} f'(x) &= L(a)a^x \cdot b^x + a^x \cdot L(b)b^x = (L(a) + L(b)) \cdot a^x b^x, \text{ sowie} \\ g'(x) &= L(a)a^{rx} \cdot r = rL(a) \cdot a^{rx}. \end{aligned} \tag{1}$$

Nun kann man aufgrund der Potenzgesetze die Funktion f und g auch anders darstellen

$$f(x) = (ab)^x \quad \text{bzw.} \quad g(x) = (a^r)^x,$$

und erhält daraus (durch unmittelbare Anwendung der Ableitungsregel (1.2))

$$f'(x) = L(ab) \cdot (ab)^x \quad \text{und} \quad g'(x) = rL(a) \cdot (a^r)^x. \tag{2}$$

Ein Vergleich von (1) und (2) zeigt die folgenden bemerkenswerten Formeln

$$L(ab) = L(a) + L(b) \quad \text{und} \quad L(a^r) = r \cdot L(a) \quad \text{für alle } a, b > 0, r \in \mathbb{R}, \tag{3}$$

Diese beiden Eigenschaften haben wir als typische Eigenschaften einer Logarithmusfunktion kennengelernt (siehe (1.3), c), d)). Dies führt zu der

Vermutung: L ist eine Logarithmusfunktion.

Wenn man dies beweisen will, muss man eine Basis finden, zu der L die Logarithmusfunktion ist. Wir suchen also eine Zahl e mit

$$L = \log_e \quad \text{d. h.} \quad L(b) = \log_e(b) \quad \text{für alle } b > 0.$$

Für eine solche Zahl e , muss also gelten:

$$L(b) = \log_e(b) \iff e^{L(b)} = e^{\log_e(b)} = b \iff e = b^{1/L(b)}.$$

Dabei gilt die letzte Äquivalenzumformung falls $L(b) \neq 0$ ist. Ist dies aber der Fall, so ist durch die letzte Gleichung ein Kandidat für e gefunden:

$$e = b^{1/L(b)} \quad \text{für irgendein } b > 0 \text{ mit } L(b) \neq 0.$$

Ein solches b ist aber leicht zu finden: Wäre nämlich $L(b) = 0$, so bedeutet dies nach (1.2) $\exp_b'(x) = 0$ für alle x , also ist $\exp_b(x) = b^x$ konstant. Dies ist aber nur für $b = 1$ möglich. Für alle anderen $b \neq 1$ gilt $L(b) \neq 0$.

Wir definieren nun die Zahl e z. B. durch

$$e = 2^{1/L(2)}.$$

Dann gilt

$$L(e) = L(2^{1/L(2)}) \stackrel{(3)}{=} \frac{1}{L(2)} \cdot L(2) = 1.$$

Wegen $L(e) > 0$ ist e^x streng monoton wachsend (siehe (1.2)), also $e > 1$. Nun zeigen wir $L = \log_e$:

$$L(b) = \log_e(b) \iff L(e^{\log_e(b)}) = \log_e(b) \stackrel{(3)}{\iff} \log_e(b) \cdot L(e) = \log_e(b).$$

Wegen $L(e) = 1$ ist die letzte Gleichung wahr, also $L = \log_e$. Damit ist L eine Logarithmusfunktion, und wir definieren

(2.1) Definition: Die Funktion $L = \log_e$ wird *natürliche Logarithmusfunktion* genannt und mit \ln bezeichnet. Ihre Basis e heißt *Eulersche Zahl*. Die Exponentialfunktion $\exp_e(x) = e^x$ zur Basis e wird kurz *e-Funktion* genannt und einfach als *die Exponentialfunktion* \exp (ohne den Index e) bezeichnet.

Aus (1.4) erhalten wir

(2.2) Satz: Die natürliche Logarithmusfunktion \ln hat die Ableitungsfunktion

$$\ln'(x) = \frac{1}{x}.$$

Beweis: Nach (1.4) gilt wegen $L(e) = 1$

$$\ln'(x) = \log_e'(x) = \frac{1}{L(e)} \cdot \frac{1}{x} = \frac{1}{x}.$$

b. Die e-Funktion. Wir wollen nun die verschiedensten Charakterisierungen der Eulerschen Zahl e bzw. der e -Funktion zusammentragen:

(2.3) Satz: (Charakterisierungen von \ln und e)

- $\ln(a)$ ist der Anstieg der Exponentialfunktion \exp_a zur Basis a an der Stelle 0, d. h. $\ln(a) = \lim_{h \rightarrow 0} \frac{1}{h}(a^h - 1)$.
- e ist die positive reelle Zahl $\neq 1$ mit $\ln = \log_e$.
- e ist die Basis des natürlichen Logarithmus \ln .
- e ist die positive reelle Zahl mit $\ln(e) = 1$.
- e ist die positive reelle Zahl mit $(e^x)' = e^x$.
- e ist der Grenzwert der Folge $(1 + \frac{1}{n})^n$.

Von diesen Aussagen sind nur noch wenige zu begründen: a) ist die hier zugrundegelegte Definition des natürlichen Logarithmus. b), c) und d) sind Ausformulierungen der gewählten Definition der Eulerschen Zahl; beim Nachweis der Existenz von e haben wir gezeigt, dass es nur eine Zahl mit der Eigenschaft c) gibt. e) ergibt sich aus Satz (1.2) mittels a) und d):

$$(e^x)' \stackrel{(1.2)}{=} L(e) \cdot e^x \stackrel{a)}{=} \ln(e) \cdot e^x \stackrel{d)}{=} e^x.$$

Lediglich f) scheint eine völlig neue Beschreibung der Zahl e zu sein. Sie ist hier aufgenommen, weil sie oft als Definition gewählt wird. Zum Beweis von f) logarithmieren wir die zu untersuchende Folge $x_n = (1 + \frac{1}{n})^n$ und erhalten (wegen $\ln(1) = 0$)

$$y_n := \ln(x_n) = \ln((1 + \frac{1}{n})^n) = n \cdot \ln(1 + \frac{1}{n}) = \frac{\ln(1 + \frac{1}{n})}{\frac{1}{n}} = \frac{\ln(1 + \frac{1}{n}) - \ln(1)}{(1 + \frac{1}{n}) - 1}.$$

Diese letzte, scheinbar unnötig komplizierte Form von y_n ist offenbar der Differenzenquotient $\frac{\ln(x) - \ln(a)}{x - a}$ der Funktion \ln an der Stelle $a = 1$ mit dem x -Wert $x = 1 + \frac{1}{n}$. Für $n \rightarrow \infty$ gilt $x \rightarrow 1 = a$ und wir erhalten

$$\lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} \frac{\ln(1 + \frac{1}{n}) - \ln(1)}{(1 + \frac{1}{n}) - 1} = \lim_{x \rightarrow 1} \frac{\ln(x) - \ln(1)}{x - 1} = \ln'(1) \stackrel{(2.2)}{=} 1.$$

Aus $\ln(x_n) = y_n \rightarrow 1$ folgt wegen der Stetigkeit der e -Funktion

$$x_n = e^{y_n} \rightarrow e^1 = e,$$

wie in f) behauptet.

In diesem Satz fehlt eine wichtige Charakterisierung des natürlichen Logarithmus \ln als diejenige Stammfunktion von $f(x) = \frac{1}{x}$ über dem Intervall $]0, \infty[$, die bei 1 den Wert 0 hat (siehe später, Integralrechnung).

Die wesentlichste Eigenschaft der e -Funktion ist die Aussage e). Diese verschärfen wir zu dem folgenden Satz über die Lösungen einer fundamentalen Differentialgleichung:

(2.4) Satz: (Die Differentialgleichung $f' = k \cdot f$) *Es sei k eine reelle Zahl und f eine Lösung der Differentialgleichung $f' = k \cdot f$, d. h. f ist eine differenzierbare Funktion mit der Eigenschaft $f'(x) = k \cdot f(x)$ für alle $x \in \mathbb{R}$. Dann gilt:*

$$\text{Es gibt eine reelle Zahl } c \in \mathbb{R} \text{ mit: } f(x) = c \cdot e^{kx}.$$

Alle Lösungen der Differentialgleichung $f' = k \cdot f$ sind also Vielfache der Exponentialfunktion $e^{kx} = (e^k)^x$.

Dabei ist dann der Faktor c gerade der sog. Anfangswert $f(0)$ von f : $c = f(0)$.

Beweis: Wir betrachten die Quotientenfunktion

$$g(x) := \frac{f(x)}{e^{kx}}.$$

Da die e -Funktion nur positive Werte annimmt, ist g auf ganz \mathbb{R} definiert. Berechnet man (mit der Quotientenregel) die Ableitung von g so erhält man für alle x

$$g'(x) = \frac{f'(x) \cdot e^{kx} - f(x) \cdot ke^{kx}}{e^{2kx}} = \frac{f'(x) - f(x) \cdot k}{e^{kx}}.$$

Wegen $f'(x) = k \cdot f(x)$ ist der letzte Term 0, also $g'(x) = 0$ für alle x . Dann muss g aber konstant sein (III, Satz (2.7)). Also gibt es eine reelle Zahl c mit $g(x) = c$ für alle x . Das bedeutet:

$$f(x) = c \cdot e^{kx} \quad \text{für alle } x \in \mathbb{R}.$$

Setzt man in diese Gleichung $x = 0$ ein, so erhält man $f(0) = c$, wie behauptet.

Dieser Satz ist die entscheidende Grundlage für die universelle Bedeutung der Exponentialfunktion. Bei allen Naturvorgängen, bei denen die Änderungsgeschwindigkeit ($f'(x)$) einer Größe $f(x)$ proportional zur Größe $f(x)$ selbst ist, ergibt sich f als Vielfaches einer Exponentialfunktion. Beispiele sind radioaktive Zerfallsprozesse, elektrische Entladungsprozesse, aber auch Wachstumsprozesse in der Biologie oder Aufladungsvorgänge in der Elektrizitätslehre; überall werden die entscheidenden Größen durch die e -Funktion beherrscht.

Es ist schwer, die Bedeutung der e -Funktion inner- wie außerhalb der Mathematik zu überschätzen!

c. Die Regeln von de l'Hospital. Dies sind keine besonderen Regeln für Exponentialfunktionen, sondern sie stellen eine sehr allgemein anwendbare Methode zur Bestimmung von

Grenzwerten dar. Sie werden hier behandelt, da sie bei der Untersuchung transzendenter Funktionen häufig benutzt werden.

Es geht um die Bestimmung von Grenzwerten von Funktionen der Form $f(x) = \frac{g(x)}{h(x)}$. Sind g und h stetig und bei a definiert, so wissen wir aufgrund der Grenzwertsätze bereits:

$$h(a) \neq 0 \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = \frac{g(a)}{h(a)}, \quad (1)$$

$$h(a) = 0, g(a) \neq 0 \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = \pm\infty. \quad (2)$$

Offen ist der Fall $g(a) = h(a) = 0$. Welchen Wert dann $\lim_{x \rightarrow a} \frac{g(x)}{h(x)}$ annimmt, kann man in vielen Fällen mit Hilfe der Differentialrechnung untersuchen. Folgende Überlegungen enthalten die Grundidee der Regel von de l'Hospital im einfachsten Fall.

Unter der Voraussetzung $g(a) = h(a) = 0$ gilt:

$$\frac{g(x)}{h(x)} = \frac{g(x) - g(a)}{h(x) - h(a)} = \frac{g(x) - g(a)}{x - a} \cdot \frac{x - a}{h(x) - h(a)} = \frac{\frac{g(x) - g(a)}{x - a}}{\frac{h(x) - h(a)}{x - a}}.$$

Wir erhalten also einen Quotienten der Differenzenquotienten von g und h . Sind nun g und h differenzierbar bei a , so strebt der Zähler gegen $g'(a)$ und der Nenner gegen $h'(a)$. Der ganze Bruch strebt dann gegen $\frac{g'(a)}{h'(a)}$ – vorausgesetzt $h'(a) \neq 0$! Insgesamt ergibt sich so die folgende zusätzliche Regel für bei a differenzierbare g, h :

$$g(a) = h(a) = 0, h'(a) \neq 0 \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = \frac{g'(a)}{h'(a)}. \quad (3)$$

Anwendungsbeispiele:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{x^2 - 3x}{e^x - 1} &= -3 \\ \lim_{x \rightarrow -1} \frac{e^x - e^{-x}}{x^2 - 1} &= \frac{e^{-1} + e}{2e^{-1}} = \frac{1 + e^2}{2}. \end{aligned}$$

Nun kann es aber sein, dass auch $h'(a) = 0$ ist, wie in folgendem Beispiel:

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x^2}.$$

Darauf ist (3) dann nicht anwendbar. Wohl aber die folgende umfassende Form der Regel von de l'Hospital, deren Beweis jedoch nicht so einfach wie der obige ist.

1. Regel von de l'Hospital: (Grenzwerte vom Typ $\frac{0}{0}$) *Es sei a eine Stelle im oder am Rande des Definitionsbereichs von g und h . Beide Funktionen seien in einer Umgebung von a differenzierbar und dort gelte $h'(x) \neq 0$. Dann gilt:*

$$\lim_{x \rightarrow a} g(x) = \lim_{x \rightarrow a} h(x) = 0 \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = \lim_{x \rightarrow a} \frac{g'(x)}{h'(x)}, \quad (4)$$

vorausgesetzt der letztgenannte Grenzwert existiert in \mathbb{R} oder ist gleich $\pm\infty$.

Die Vorteile dieser allgemeinen Form sind

1. Anwendbarkeit auch bei $h'(a) = 0$,
2. rekursive Anwendbarkeit bei $g'(a) = h'(a) = 0$,
3. Anwendbarkeit, wenn g, h bei a nicht definiert sind.

Die nächsten drei Beispiele belegen jeden dieser Vorteile:

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{e^x - 1}{x^2} &= \lim_{x \rightarrow 0} \frac{e^x}{2x} \stackrel{(2)}{=} \infty, \\ \lim_{x \rightarrow 0} \frac{e^{2x} - 2e^x + 1}{x^2} &= \lim_{x \rightarrow 0} \frac{2e^{2x} - 2e^x}{2x} = \lim_{x \rightarrow 0} \frac{4e^{2x} - 2e^x}{2} = \frac{4 - 2}{2} = 1, \\ \lim_{x \rightarrow 0} \frac{\ln(x+1)}{1 - e^x} &= \lim_{x \rightarrow 0} \frac{\frac{1}{x+1}}{-e^x} = \lim_{x \rightarrow 0} \frac{1}{-(x+1)e^x} = -1.\end{aligned}$$

Die zweite Regel von de l'Hospital betrifft Grenzwerte vom sog. Typ $\frac{??}{\infty}$, d. h. $\lim_{x \rightarrow a} \frac{g(x)}{h(x)}$ im Falle $\lim_{x \rightarrow a} h(x) = \pm\infty$. Aufgrund der Grenzwertsätze ist uns hierzu bekannt:

$$\lim_{x \rightarrow a} h(x) = \pm\infty, \quad g \text{ bei } a \text{ beschränkt} \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = 0. \quad (5)$$

Mit Hilfe der Differentialrechnung kann man weitere Fälle behandeln:

2. Regel von de l'Hospital: (Grenzwerte vom Typ $\frac{??}{\infty}$) *Es sei a eine Stelle im oder am Rande des Definitionsbereichs von g und h . Beide Funktionen seien in einer Umgebung von a differenzierbar und dort gelte $h'(x) \neq 0$. Dann gilt:*

$$\lim_{x \rightarrow a} h(x) = \pm\infty \implies \lim_{x \rightarrow a} \frac{g(x)}{h(x)} = \lim_{x \rightarrow a} \frac{g'(x)}{h'(x)}, \quad (6)$$

vorausgesetzt der letztgenannte Grenzwert existiert in \mathbb{R} oder ist gleich $\pm\infty$.

Als wichtigstes Anwendungsbeispiel dafür formulieren wir:

Folgerung: *Die Exponentialfunktion wächst stärker als jede Potenz:*

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{x^k}{e^x} &= 0 \quad \text{für alle } k \geq 0, \\ \lim_{x \rightarrow \infty} \frac{h(x)}{e^x} &= 0 \quad \text{für jede ganzrationale Funktion } h.\end{aligned}$$

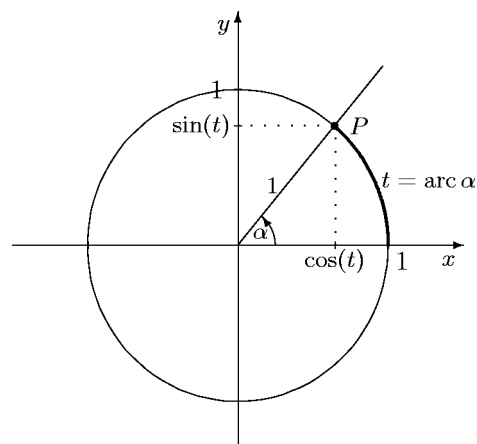
Entsprechend gilt: *Die Logarithmusfunktion \ln wächst schwächer als jede Potenz:*

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{\ln(x)}{x^k} &= 0 \quad \text{für alle } k > 0, \\ \lim_{x \rightarrow \infty} \frac{\ln(x)}{h(x)} &= 0 \quad \text{für jede nicht konstante ganzrationale Funktion } h.\end{aligned}$$

§12 Die trigonometrischen Funktionen

a. Definition am Einheitskreis. Wir betrachten den sog. Einheitskreis, den Kreis vom Radius 1, in einem kartesischen Koordinatensystem und beliebige Punkte P darauf. Diese können auf verschiedene Weisen beschrieben werden (siehe Skizze):

1. Durch den Winkel α zwischen positiver x -Achse und der Verbindung vom Zentrum zu P . Dieser Winkel wird mit einem Vorzeichen versehen, positiv bei Drehung gegen Uhrzeigersinn und negativ im umgekehrten Fall.
2. Durch die Maßzahl t der Länge des den Winkel beschreibenden Kreisbogens auf dem Einheitskreis, das sog. Bogenmaß $t = \text{arc } \alpha$ des Winkels, wieder mit dem entsprechenden Vorzeichen je nach Drehrichtung.
3. Durch die beiden Koordinaten von P . In Abhängigkeit vom Bogenmaß t erhalten die Koordinaten von P einen Namen: Die x -Koordinate ist der Cosinus $\cos(t)$, die y -Koordinate ist der Sinus $\sin(t)$.



Auf diese Weise sind die *trigonometrischen Funktionen* \sin und \cos für beliebiges $t \in \mathbb{R}$ definiert. Für $0 < t < \frac{\pi}{2}$, d. h. für Winkel α zwischen 0° und 90° erhält man aus dem Strahlensatz die bekannte Beschreibung von Sinus und Cosinus am rechtwinkligen Dreieck:

$$\sin(\alpha) = \frac{\text{Gegenkathete}}{\text{Hypotenuse}}, \quad \cos(\alpha) = \frac{\text{Ankathete}}{\text{Hypotenuse}},$$

wobei präziser formuliert jeweils die Länge der entsprechenden Dreiecksseiten gemeint ist. Diese Beschreibung ist aber nur für spitze Winkel richtig, während die obige Definition am Einheitskreis allgemeingültig ist.

Aus der Definition am Einheitskreis entnimmt man unmittelbar die folgenden grundlegenden Eigenschaften:

$$\boxed{-1 \leq \sin(t) \leq +1, \quad -1 \leq \cos(t) \leq +1} \quad (1)$$

Genauer ergibt sich aus dem Satz des Pythagoras:

$$\boxed{\sin^2(t) + \cos^2(t) = 1} \quad (2)$$

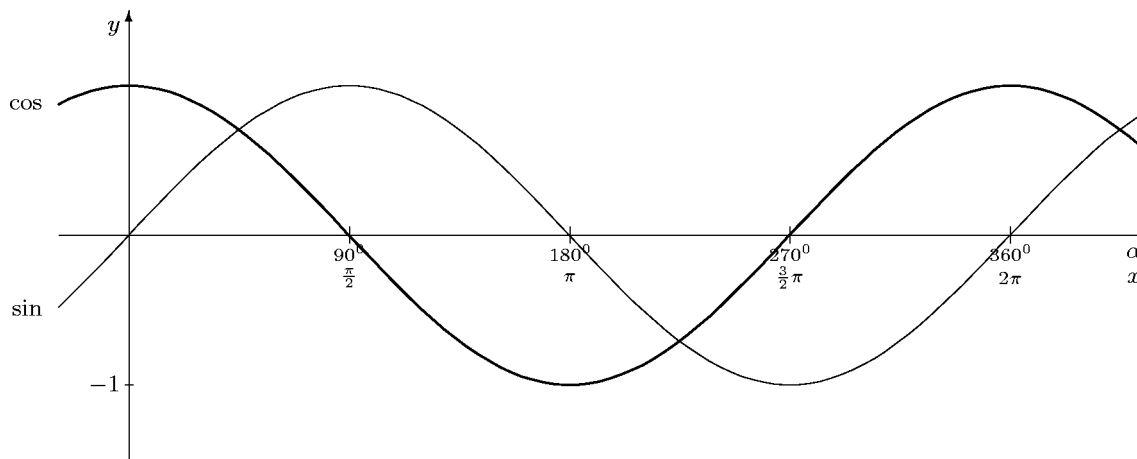
Weiter gilt:

$$\boxed{\sin \text{ und } \cos \text{ haben die Periode } 2\pi : \begin{cases} \sin(t + 2\pi) = \sin(t) \\ \cos(t + 2\pi) = \cos(t) \end{cases}} \quad (3)$$

Indem man den Punkt P mit dem an der x -Achse gespiegelten Punkt vergleicht, erkennt man:

$$\boxed{\begin{array}{ll} \cos(-t) = \cos(t): & \cos \text{ ist achsensymmetrisch,} \\ \sin(-t) = -\sin(t): & \sin \text{ ist punktsymmetrisch.} \end{array}} \quad (4)$$

Mit Hilfe einer genauen Zeichnung des Einheitskreises erhält man etwa den folgenden Verlauf der Graphen der beiden trigonometrischen Funktionen:



In dieser Skizze erkennt man eine Reihe weiterer Symmetrien, die sich alle aus entsprechenden Symmetrien am Einheitskreis ergeben. (Siehe Skript zur Einführungsphase.) Eine besonders wichtige wollen wir hier noch festhalten: Verschiebt man den Graphen der Cosinus-Funktion um $\frac{\pi}{2}$ nach rechts, so erhält man den Graphen der Sinus-Funktion: $\cos(t - \frac{\pi}{2}) = \sin(t)$. Umgekehrt ergibt dieselbe Verschiebung beim Sinus-Graphen den Graphen von $-\cos$: $\sin(t - \frac{\pi}{2}) = -\cos(t)$.

$$\boxed{\cos(t - \frac{\pi}{2}) = \sin(t), \quad \sin(t - \frac{\pi}{2}) = -\cos(t)} \quad (5)$$

b. Additionstheoreme und Ableitung der trigonometrischen Funktionen. Wir bemerken zunächst, dass aufgrund der obigen Definition die trigonometrischen Funktionen *stetig* sind: Nähert sich nämlich das Bogenmaß t einem festen Wert t_0 , so nähert sich der zugehörige Punkt $P = (\cos(t), \sin(t))$ auf dem Einheitskreis dem Punkt $(\cos(t_0), \sin(t_0))$, insbesondere also $\cos(t) \rightarrow \cos(t_0)$ und $\sin(t) \rightarrow \sin(t_0)$.

$$\boxed{\text{sin und cos sind stetig.}} \quad (6)$$

Wir wollen dies nun zur Differenzierbarkeit verschärfen. So wie die Ableitungsformel für die Exponentialfunktionen auf deren Funktionalgleichung $\exp(x + y) = \exp(x) \cdot \exp(y)$ beruhte, benötigen wir für die Ableitung der trigonometrischen Funktionen die

Additionstheoreme: Für beliebige $a, b \in \mathbb{R}$ gilt:

$$\begin{aligned} \cos(a + b) &= \cos(a) \cos(b) - \sin(a) \sin(b), & \cos(a - b) &= \cos(a) \cos(b) + \sin(a) \sin(b), \\ \sin(a + b) &= \sin(a) \cos(b) + \cos(a) \sin(b), & \sin(a - b) &= \sin(a) \cos(b) - \cos(a) \sin(b). \end{aligned}$$

Reduktion des Beweises:

1. Aufgrund der Symmetrien und Periodizitäten der trigonometrischen Funktionen genügt es diese Beziehungen für $0 \leq a, b \leq \pi/2$ zu beweisen.
2. Die jeweils in einer Zeile stehenden Formeln sind äquivalent (man ersetze b durch $-b$ und benutze die Symmetrien von sin und cos), so dass nur je eine bewiesen werden muss.
3. Wegen (5) ergeben sich die Formeln der zweiten Zeile aus denen der ersten.
4. Es genügt also der Beweis *einer* der 4 Formeln.

Einen geometrischen Beweis für eine dieser Formeln finden Sie im Lehrbuch Analysis 1, S. 58. Arbeiten Sie ihn gründlich durch.

Mit Hilfe der Additionstheoreme können wir die Differenzierbarkeit von sin und cos an einer beliebigen Stelle x zurückführen auf die Differenzierbarkeit an der Stelle 0: Setzt man in einem der Additionstheoreme $b = x - a$ so erhält man

$$\sin(x) = \sin(a) \cos(x - a) + \cos(a) \sin(x - a) \quad (A)$$

Sind nun \cos und \sin an der Stelle 0 differenzierbar, so sind nach der Kettenregel $\cos(x-a)$ und $\sin(x-a)$ an der Stelle $x=a$ differenzierbar. Mit Faktor- und Summenregel folgt dann aus (A), dass auch $\sin(x)$ an der Stelle a differenzierbar ist mit

$$\sin'(a) = \sin(a) \cos'(0) + \cos(a) \sin'(0) \quad \text{für jedes } a \in \mathbb{R}.$$

oder in vertrauterer Darstellung:

$$\sin'(x) = \sin(x) \cos'(0) + \cos(x) \sin'(0). \quad (7)$$

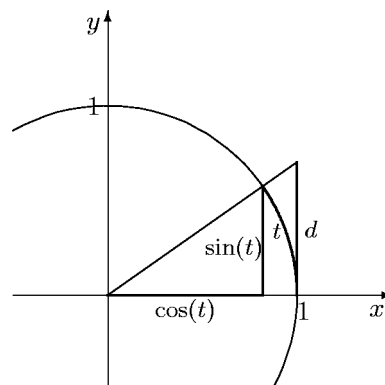
Damit braucht man nur noch die Differenzierbarkeit beider Funktionen an der Stelle 0 zu untersuchen. Wir zeigen nun, dass beide Ableitungswerte existieren, und zwar:

$$\sin'(0) = 1 \quad \text{und} \quad \cos'(0) = 0. \quad (8)$$

Die Bestimmung dieser Grenzwerte beruht auf der Abschätzung:

$$\cos(t) < \frac{\sin(t)}{t} < 1 \quad \text{für } t \neq 0. \quad (*)$$

Der nebenstehenden Skizze entnehmen wir: $\sin(t) < t < d$ (bei $t > 0$ und $t < \frac{\pi}{2}$). (Diese anschaulich einsichtigen Tatsachen ergeben sich durch einen Längen- ($\sin t < t$) und durch einen Flächenvergleich ($t < d$). Eine präzise Begründung erfordert genaue Begriffe der Bogenlänge von Kurven und des Flächeninhalts von krummlinig begrenzten Flächenstücken, die wir im Rahmen der Integralrechnung behandeln werden.) Mit Hilfe des Strahlensatzes erhalten wir für die Größe d : $\frac{d}{1} = \frac{\sin(t)}{\cos(t)}$. Insgesamt folgt also:



$$\sin(t) < t \quad \text{sowie} \quad t < \frac{\sin(t)}{\cos(t)}$$

Mit den üblichen äquivalenten Umformungen erhält man dann (*).

Im Fall $t < 0$ argumentiert man sinngemäß.

Wegen $\lim_{t \rightarrow 0} \cos(t) = \cos(0) = 1$ folgt aus (*) mit dem Schachtelungssatz die Existenz des Grenzwertes

$$\lim_{t \rightarrow 0} \frac{\sin(t)}{t} = 1.$$

Also ist \sin an der Stelle 0 differenzierbar mit $\sin'(0) = 1$.

Da $\cos(t)$ in einer Umgebung von 0 positiv ist, erhält man aus (2), dem Satz des Pythagoras, $\cos(t) = \sqrt{1 - \sin^2(t)}$. Da \sin an der Stelle 0 differenzierbar ist, erhalten wir mit bekannten Ableitungsregeln und der Kettenregel dann: \cos ist an der Stelle 0 differenzierbar und es gilt

$$\cos'(0) = \frac{1}{2\sqrt{1 - \sin^2(0)}} \cdot (-2 \sin(0) \sin'(0)) = 0.$$

Insgesamt erhalten wir nun aus (7) und (8): \sin ist überall differenzierbar und es gilt $\sin'(x) = \cos(x)$. In gleicher Weise untersucht man die Differenzierbarkeit von \cos . Insgesamt folgt damit der fundamentale

Satz: Die trigonometrischen Funktionen \sin und \cos sind differenzierbar, und es gilt:

$$\boxed{\sin' = \cos, \quad \cos' = -\sin.}$$

c. Die Schwingungsdifferentialgleichung. Die trigonometrischen Funktionen sind zunächst (schon dem Namen nach) mit geometrischen Fragen verknüpft. Ihre besondere universelle Bedeutung in der Physik und allgemein in den Naturwissenschaften verdanken sie aber einer anderen Tatsache. Aufgrund der obigen Ableitungsregeln erhalten wir $\sin'' = -\sin$ und $\cos'' = -\cos$. Die beiden Funktionen \sin und \cos sind also Lösungen der Gleichung $f'' = -f$. Man nennt dies eine *Differentialgleichung*, weil hier eine Beziehung zwischen f und deren Ableitungen hergestellt wird.

Das entscheidende Resultat ist nun: *Jede* Lösung dieser Differentialgleichung ist durch die beiden trigonometrischen Funktionen darstellbar. Dies hat zur Folge, dass überall, wo physikalische Größen der obigen (oder einer eng verwandten) Differentialgleichung genügen, notwendig die trigonometrischen Funktionen in Erscheinung treten. Die physikalischen Größen verhalten sich also periodisch, es liegen harmonische Schwingungen vor. Daher nennt man diese Differentialgleichung (und eng verwandte) auch *Schwingungsdifferentialgleichung*.

Satz: (Die Differentialgleichung $f'' = -f$)

a) Ist f eine zweimal differenzierbare Funktion mit der Eigenschaft $f'' = -f$, so gibt es geeignete Konstanten $a, b \in \mathbb{R}$ mit

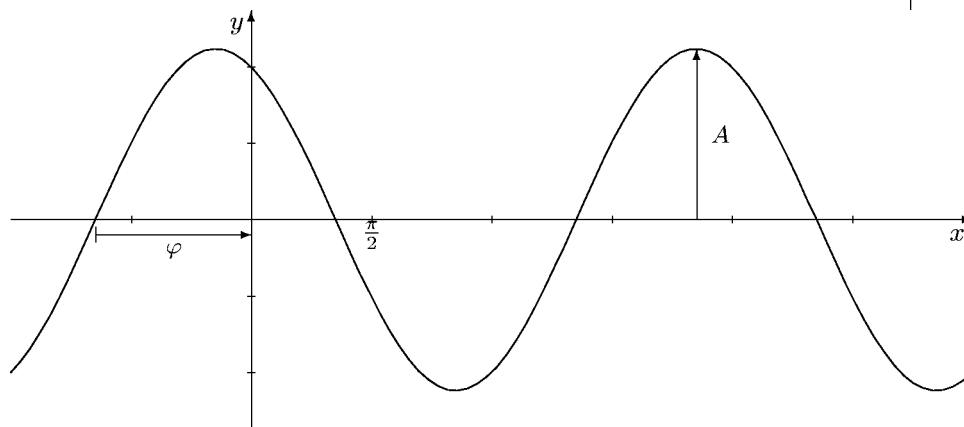
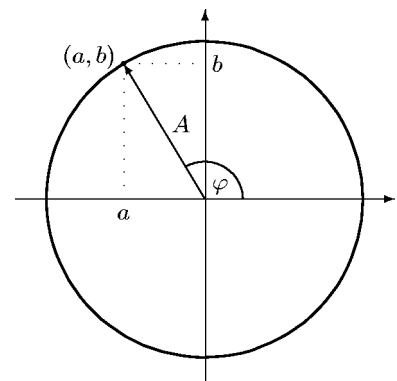
$$f(x) = a \sin(x) + b \cos(x).$$

Dabei sind $a = f'(0)$ und $b = f(0)$ die sog. Anfangswerte von f .

b) Jede solche Funktion lässt sich auch in folgender Form darstellen:

$$f(x) = A \sin(x + \varphi)$$

mit $A \geq 0$, $A^2 = a^2 + b^2$ und dem Winkel φ mit $A \cos \varphi = a$ und $A \sin \varphi = b$ (siehe nebenstehende Skizze, für $A \neq 0$ ist φ eindeutig im Bereich $] -\pi, \pi]$). Damit hat der Graph von f folgendes Aussehen:



A gibt also den größten Wert von f an (physikalisch: die Amplitude der Schwingung) und φ die sog. Phase (oder den Phasenwinkel) der Schwingung im Start(zeit)punkt.

Beweis: a) Gesucht sind Zahlen a, b mit $f(x) = a \sin(x) + b \cos(x)$. Dies ist nur *eine* Gleichung für *zwei* Unbekannte. Durch Ableiten erhält man jedoch eine zweite: $f'(x) = a \cos(x) - b \sin(x)$. Damit hat man ein System von 2 Gleichungen für 2 Unbekannte:

$$\begin{aligned} f(x) &= a \sin(x) + b \cos(x) \\ f'(x) &= a \cos(x) - b \sin(x) \end{aligned}$$

Diese löse man wie üblich nach a, b auf. Multiplikation der ersten Gleichung mit $\sin(x)$, der zweiten mit $\cos(x)$ und Addition ergibt

$$\begin{aligned} f(x) \sin(x) &= a \sin^2(x) + b \sin(x) \cos(x) \\ f'(x) \cos(x) &= a \cos^2(x) - b \sin(x) \cos(x) \\ \implies f(x) \sin(x) + f'(x) \cos(x) &= a(\sin^2(x) + \cos^2(x)) = a. \end{aligned}$$

Genauso erhält man

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

Man überprüft leicht, dass die so gefundenen Terme für a, b die behauptete Darstellung in a) liefern. Man muss aber noch zeigen, dass a, b tatsächlich Konstanten sind und nicht von x abhängen!

Erst hierfür wird die Voraussetzung an f benutzt: $f'' = -f$. Mit ihrer Hilfe zeigen wir, dass die für a, b gefundenen Funktionsterme die Ableitung 0 haben, also konstant sind:

$$\begin{aligned} (f(x) \sin(x) + f'(x) \cos(x))' &= f'(x) \sin(x) + f(x) \cos(x) + f''(x) \cos(x) + f'(x)(-\sin(x)) \\ &= f(x) \cos(x) + (-f(x)) \cos(x) = 0. \end{aligned}$$

Genauso zeigt man, dass b konstant ist. Setzt man in der Darstellung $f(x) = a \sin(x) + b \cos(x)$ $x = 0$ ein, so erhält man $f(0) = a \cdot 0 + b \cdot 1 = b$, und entsprechend gewinnt man aus der Darstellung für f' auch die letzte noch offene Behauptung von a): $f'(0) = a$.

Für die Darstellung b) benutzt man die Additionstheoreme:

$$A \sin(x + \varphi) = A \cdot (\sin(x) \cos(\varphi) + \cos(x) \sin(\varphi)) = A \cos(\varphi) \cdot \sin(x) + A \sin(\varphi) \cdot \cos(x).$$

Dies ist gleich $f(x) = a \sin(x) + b \cos(x)$, wenn folgendes gilt:

$$a = A \cos(\varphi) \quad \text{und} \quad b = A \sin(\varphi).$$

Der Satz des Pythagoras ergibt $a^2 + b^2 = A^2(\cos^2(\varphi) + \sin^2(\varphi)) = A^2$ und der Winkel φ ist dann bestimmt durch

$$\frac{a}{A} = \cos(\varphi), \quad \frac{b}{A} = \sin(\varphi), \quad \text{bzw.} \quad (a, b) = (A \cos(\varphi), A \sin(\varphi)),$$

wie in der Skizze dargestellt.

Anmerkung (nicht nur für Physiker): Die allgemeine Schwingungsdifferentialgleichung lautet $f'' = -kf$ mit *positivem* k . Setzt man $\omega = \sqrt{k}$, so erhält man die handlichere Form $f'' = -\omega^2 f$, die man durch den Ansatz $f(x) = g(\omega x)$ auf die vorangehende Differentialgleichung zurückführt:

Die Kettenregel zeigt $f'(x) = \omega g'(\omega x)$, $f''(x) = \omega^2 g''(\omega x)$, so dass aus $f'' = -\omega^2 f$ sofort folgt: $g''(\omega x) = -g(\omega x)$ für jedes x . Mithin ist g eine Lösung der Differentialgleichung $g'' = -g$. Nach dem vorangehenden Satz folgt also $g(x) = a \sin(x) + b \cos(x) = A \sin(x + \varphi)$. Damit ergibt sich der

Satz: Jede Lösung der sog. Schwingungsdifferentialgleichung $f'' = -\omega^2 f$ ist von der Form

$$f(x) = a \sin(\omega x) + b \cos(\omega x) = A \sin(\omega(x + x_0))$$

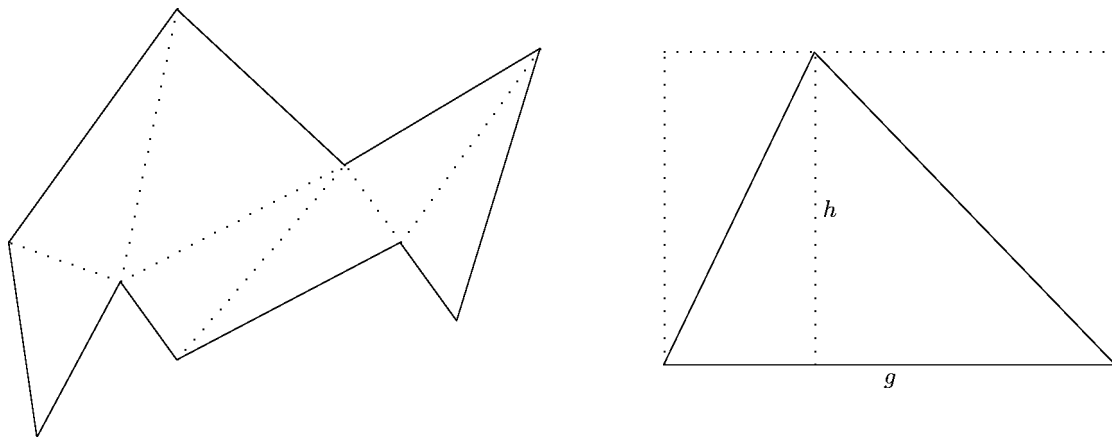
mit $f(0) = b$, $f'(0) = \omega a$, $A = \sqrt{a^2 + b^2}$ und $(a, b) = (A \cos(\omega x_0), A \sin(\omega x_0))$.

V. Integralrechnung

Gegenstand der Integralrechnung ist das uralte Problem der Berechnung von Flächeninhalten und Volumina. Aber auch Bogenlängen werden mit Hilfe der Integralrechnung ermittelt.

§13 Flächeninhalt und Integral

a. Grundprinzipien der Flächenberechnung. Die Flächenberechnung stellt kein großes Problem dar, solange die Flächenstücke geradlinig begrenzt sind: Man kann sie in Dreiecke ‘zerschneiden’ und Dreiecksflächen kann man nach der Formel *Fläche gleich Grundlinie mal*



Höhe durch 2 berechnen. Diese Formel wiederum erhält man dadurch, dass man ein beliebiges Dreieck mittels einer Höhe in zwei rechtwinklige Dreiecke zerlegt und diese durch geeignete Verdopplung zu Rechtecken ergänzt. Insgesamt erhält man so ein Gesamtrechteck, das doppelt so groß ist wie das gegebene Dreieck. Die Fläche dieses Rechtecks ist das Produkt $g \cdot h$ der Kantenlängen. Die Hälfte davon ist dann die Dreiecksfläche.

In diesen kurzen Bemerkungen sind schon die elementarsten Eigenschaften des Flächeninhalts angesprochen:

1. Rechtecksfläche = Produkt der Kantenlängen. (Dies ist im Kern die Definition des Flächeninhalts.)
2. Zerschneidet man Flächenstücke, so addieren sich die Einzelflächeninhalte zum Gesamtflächeninhalt.
3. Deckungsgleiche (kongruente) Flächenstücke haben denselben Flächeninhalt.

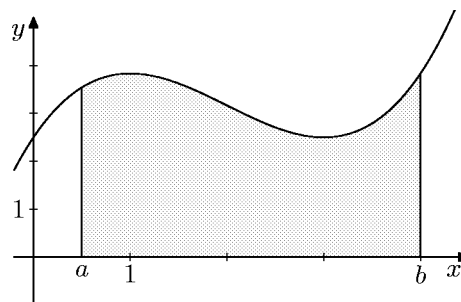
Schließlich wollen wir noch eine weitere, ebenso offensichtliche Eigenschaft formulieren, die aber für das folgende auch ebenso fundamental ist:

4. Liegt ein Flächenstück vollständig in einem anderen, so ist sein Flächeninhalt höchstens so groß wie der des umfassenderen Flächenstücks.

Diese vier Eigenschaften sind die Grundregeln über Flächeninhalte, auf denen die nachfolgenden Überlegungen basieren. Sie sind hier explizit formuliert, um die Fundamente deutlich offen zu legen. Zugleich sollen dies die einzigen Eigenschaften sein, die wir über Flächeninhalte verwenden werden.

b. Intervallzerlegungen. Das eigentliche Problem ist die Flächenberechnung für *krummlinig begrenzte* Flächenstücke. Durch geeignetes Zerschneiden kann man sich auf Flächenstücke beschränken, die wie skizziert zwischen dem Graphen einer Funktion und der x -Achse liegen und seitlich durch Parallelen zur y -Achse begrenzt sind. Mit a und b bezeichnen wir die Lage der seitlichen Begrenzungen und es sei $a < b$. Diese Bezeichnungen werden wir im folgenden stets verwenden.

Da man krummlinig begrenzte Flächenstücke nicht in Rechtecke (oder allgemeiner, in geradlinig begrenzte Flächenstücke) zerlegen kann, versucht man den Flächeninhalt zunächst *anzunähern*, indem man



das Flächenstück durch Rechtecke auszuschöpfen versucht und dann den Fehler untersucht. Diese Methode des *Ausschöpfens* wurde bereits von *Archimedes* zur Bestimmung der Kreisfläche und damit der Zahl π verwendet.

Das folgende, systematische Vorgehen geht auf den Mathematiker Bernhard Riemann (1826–1866) zurück. Gegeben ist eine Funktion f (deren Graph die *Randkurve* des Flächenstücks bildet) sowie ein Intervall $J = [a, b]$ ($a < b$), wodurch die anderen geradlinigen Berandungen gegeben sind (siehe die obige Skizze). Wir unterteilen das uns interessierende Flächenstück in n (schmale) Streifen, wobei n irgendeine natürliche Zahl ist. Eine solche Unterteilung von J wird festgelegt durch Auswahl von $n + 1$ Stellen in J :

$$a = x_0 < x_1 < \dots < x_k < \dots < x_{n-1} < x_n = b.$$

Ein solches System von Zahlen nennen wir eine *Zerlegung* \mathcal{Z} des Intervalls. Die dadurch entstehenden n Teilintervalle

$$J_k = [x_{k-1}, x_k] \quad (k = 1, \dots, n)$$

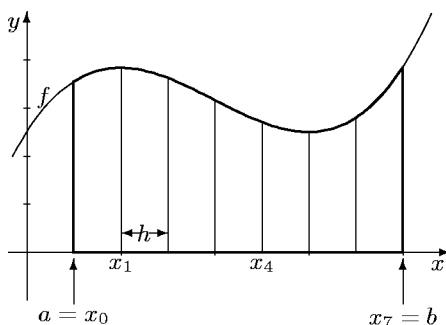
haben die jeweilige Breite $h_k = x_k - x_{k-1}$. Als Maß für die *Feinheit* der Zerlegung wählen wir die größte auftretende Breite h_k eines der Teilintervalle J_k , und bezeichnen sie mit $h(\mathcal{Z})$.

Von besonderem Interesse sind die *äquidistanten* Zerlegungen, bei denen alle h_k einander gleich sind: $h_k = h$. Für diese gilt dann bei n Streifen

$$\text{Streifenbreite: } h = \frac{b - a}{n},$$

$$\text{Zerlegungsstellen: } x_k = a + kh \quad (k = 0, \dots, n).$$

Die folgende Skizze veranschaulicht ein Beispiel einer solchen Zerlegung in gleichlange Teilintervalle. Dabei ist $a = \frac{1}{2}$, $b = 4$, $n = 7$ und $h = \frac{b-a}{n} = \frac{1}{2}$. Die interessierende Gesamtfläche ist mit dickerer Strichstärke umrandet.



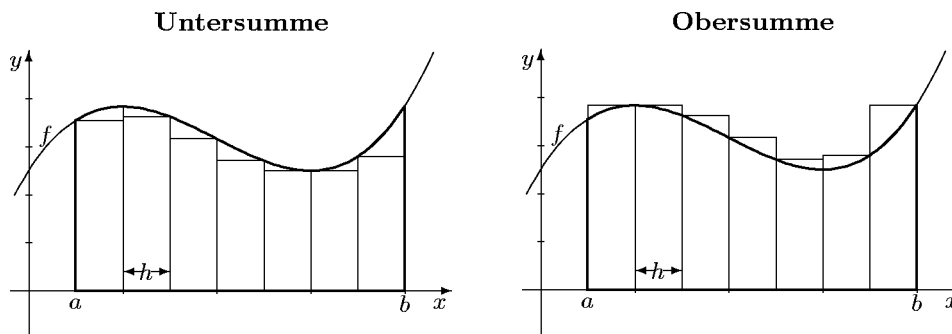
c. Ober- und Untersummen. Man versucht nun den Flächeninhalt einzuschachteln durch die sog. Ober- und Untersummen. Um diese zu definieren, muss die Funktion f über dem ganzen Intervall $J = [a, b]$ *definiert* und dort *beschränkt* sein, d. h. ihre *Werte* müssen zwischen einer oberen und einer unteren Schranke liegen. Wir betonen, dass die Funktion insbesondere auch an den Randstellen a und b definiert sein muss!

In unseren nachfolgenden Formulierungen werden wir jedoch der Einfachheit halber mehr voraussetzen: Der Graph soll über dem interessierenden Bereich $J = [a, b]$ keine Lücken aufweisen. Dies bedeutet, dass die Funktion f über dem Intervall $J = [a, b]$ *stetig* sein soll. Mit geringen Modifikationen bleiben die Aussagen allgemein gültig. Wo dies nicht der Fall ist, wird die Stetigkeit ausdrücklich gefordert.

Ebenfalls der Einfachheit halber beschränken wir uns im folgenden soweit wie möglich auf äquidistante Zerlegungen.

Untersummen: Man bildet aus jedem Streifen ein möglichst großes Rechteck, das ganz *unterhalb* des Graphen bleibt, so dass also sämtliche Rechtecke zusammen vollständig im dick

umrandeten Flächenstück enthalten sind (siehe Skizze). Die Höhe des jeweiligen Rechtecks wird



bestimmt durch den *kleinsten* Wert, den die Funktion f im Bereich des entsprechenden Teilintervalls $J_k = [x_{k-1}, x_k]$ annimmt. Wir bezeichnen diesen kleinsten Wert im k -ten Streifen mit $f(z_k)$. Dabei gibt z_k eine *Stelle* im Intervall $J_k = [x_{k-1}, x_k]$ an, an der die Funktion f diesen kleinsten Wert erreicht.

Die Summe der Flächeninhalte der so gebildeten n Rechtecke ist dann die sog. *Untersumme* $U_n(f)$ von f zu n gleich breiten Streifen:

$$U_n(f) = hf(z_1) + hf(z_2) + \dots + hf(z_n) = h \cdot (f(z_1) + f(z_2) + \dots + f(z_n)).$$

Sie ist also ein Produkt aus der Streifenbreite $h = \frac{b-a}{n}$ multipliziert mit einer Summe von n Funktionswerten, und zwar für jedes der n Teilintervalle der *kleinste Funktionswert*.

Obersummen: Hier geht man analog vor, nur bildet man nun aus jedem Streifen ein Rechteck, das den Bereich zwischen Graph und x -Achse vollständig *umfasst*, dessen obere Begrenzung also *oberhalb* des Graphen liegt (siehe obige Skizze rechts).

Die Höhe des jeweiligen Rechtecks wird bestimmt durch den *größten* Wert, den die Funktion f im Bereich des entsprechenden Teilintervalls $J_k = [x_{k-1}, x_k]$ annimmt. Wir bezeichnen diesen *größten* Wert im k -ten Streifen mit $f(Z_k)$. Dabei gibt Z_k eine *Stelle* im Intervall $J_k = [x_{k-1}, x_k]$ an, an der die Funktion f diesen größten Wert erreicht.

Die Summe der so gebildeten n Rechtecksflächen ist dann die sog. *Obersumme* $O_n(f)$ von f zu n gleichbreiten Streifen:

$$O_n(f) = hf(Z_1) + hf(Z_2) + \dots + hf(Z_n) = h \cdot (f(Z_1) + f(Z_2) + \dots + f(Z_n)).$$

Sie ist also wieder ein Produkt aus der Streifenbreite $h = \frac{b-a}{n}$ multipliziert mit einer Summe von ebenfalls n Funktionswerten, nur eben diesmal für jedes der n Teilintervalle der *größte Funktionswert*.

Aufgrund dieser Definition und den einleitend genannten Grundprinzipien jeglicher Flächenberechnung (insbesondere der so selbstverständlichen Eigenschaft 4.) gilt nun die folgende Beziehung für den gesuchten Flächeninhalt A des dick umrandeten Flächenstücks in obiger Skizze.

$$U_n(f) \leq A \leq O_m(f) \quad \text{für jegliche Zahlen } n, m. \tag{1}$$

Diese Beziehung bedeutet eine Einschachtelung des gesuchten Flächeninhalts A zwischen explizit berechenbare Werte. (Es sei angemerkt, dass diese Beziehung nicht nur für äquidistante Zerlegungen gilt: Untersummen sind generell höchstens und Obersummen mindestens so groß wie der Flächeninhalt A .)

d. Das Integral. Dass man nun aus dieser *Einschachtelung* des Flächenwertes eine präzise Bestimmung von A gewinnen kann, beruht auf den folgenden Überlegungen. Verfeinert man eine Intervallzerlegung Z etwa durch Halbierung aller Intervalle, so wächst die Untersumme und die Obersumme fällt. Zusammen mit (1) ergibt dies:

$$U_n(f) \leq U_{2n}(f) \leq A \leq O_{2n}(f) \leq O_n(f). \tag{2}$$

Begründung: Zerlegt man ein Intervall in zwei Teilintervalle, so müssen die Werte der Funktion auf den zwei neugebildeten Teilintervallen natürlich mindestens so groß sein wie der kleinste Wert für das gesamte Intervall; die Untersumme kann also nicht absinken. Entsprechend kann die Obersumme nicht anwachsen.

Anmerkung: Für diese Argumentation ist entscheidend, dass man von einer gegebenen Zerlegung zu einer *feineren* Zerlegung übergeht, indem man nur *neue* Stellen *hinzufügt*. Man kann auf diese Weise etwa U_n mit U_{2n} oder U_{3n} vergleichen, i. a. jedoch nicht U_n und U_{n+1} .

Betrachtet man nun von $U_1(f)$ ausgehend die durch Halbierung entstehende Untersummenfolge, so ist diese monoton wachsend:

$$U_1(f) \leq U_2(f) \leq U_4(f) \leq U_8(f) \leq \dots \leq U_{2^n}(f) \leq \dots$$

Da sie gemäß (1) außerdem auch noch beschränkt ist, muss sie konvergieren! Dasselbe gilt sinngemäß für die monoton fallende Obersummenfolge. Für die in \mathbb{R} existierenden Grenzwerte gilt dann (wiederum nach (1))

$$\lim_{n \rightarrow \infty} U_{2^n}(f) \leq A \leq \lim_{n \rightarrow \infty} O_{2^n}(f). \quad (3)$$

Man erkennt nun: Haben Unter- und Obersummenfolgen *denselben* Grenzwert, so ist dieser gerade der gesuchte Flächeninhalt A . Dies ist der Grund für folgende Definition:

Definition: Eine über einem Intervall J definierte und beschränkte Funktion f heißt über J *integrierbar*, wenn die durch fortgesetzte Halbierung entstehenden Ober- und Untersummenfolgen denselben Grenzwert haben:

$$f \text{ integrierbar} \iff \lim_{n \rightarrow \infty} U_{2^n}(f) = \lim_{n \rightarrow \infty} O_{2^n}(f).$$

Man definiert dann das *Integral* als den gemeinsamen Grenzwert:

$$\int_a^b f = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} U_{2^n}(f) = \lim_{n \rightarrow \infty} O_{2^n}(f).$$

Lesen Sie: *Integral der Funktion f in den Grenzen von a bis b* oder etwas kürzer *Integral von a bis b $f(x) dx$* .

Die Schreibweise $\int_a^b f(x) dx$ ist historisch gewachsen. Das Zeichen \int ist ein stilisiertes S , das an die Summenbildung erinnern soll, und das dx symbolisiert die in den Ober- und Untersummen auftretenden *Differenzen* $x_{x+1} - x_k$ von x -Werten. Für uns hat das dx keine eigenständige Bedeutung, es ist nur Bestandteil des Integralsymbols $\int_a^b \dots dx$.

Ohne den technisch etwas aufwendigeren Beweis vermerken wir, dass nicht nur die durch fortgesetzte Halbierung entstehende Untersummenfolge $U_{2^n}(f)$, sondern die gesamte Untersummenfolge $U_n(f)$ konvergiert (diese braucht allerdings nicht monoton zu sein), und zwar gegen denselben Grenzwert. Um dies zu zeigen, muss man verschiedene Intervallzerlegungen ‘ineinander mischen’ und daher beliebige Intervallzerlegungen (nicht nur äquidistante) betrachten. Der Beweis zeigt dann sogar, dass die Untersummenfolgen zu beliebigen Intervallzerlegungen \mathcal{Z}_n konvergieren, wenn nur deren maximale Intervallbreite $h(\mathcal{Z}_n)$ gegen 0 strebt. Der Grenzwert ist dann immer derselbe.

Gleiche Aussagen gelten für Obersummen. Wenn also für *eine* Zerlegungsfolge Ober- und Untersummen denselben Grenzwert haben, so gilt dies für *alle* Zerlegungsfolgen \mathcal{Z}_n :

Satz: Ist f im oben definierten Sinne über $J = [a, b]$ integrierbar, so gilt allgemein für die vollständige Unter- und Obersummenfolge:

$$\int_a^b f = \lim_{n \rightarrow \infty} U_n(f) = \lim_{n \rightarrow \infty} O_n(f).$$

Das Integral ist der gemeinsame Grenzwert der Ober- und der Untersummen.

Es gilt sogar noch mehr. In Verallgemeinerung der Ober- und Untersummen kann man auch beliebige Zwischenstellen $v_k \in J_k = [x_{k-1}, x_k]$ (nicht notwendig Stellen minimaler oder maximaler Werte) betrachten, etwa die jeweilige Mitte v_k des Intervalls J_k . Nach erfolgter Wahl der Zwischenstellen v_k definiert man dann (in Analogie zu Unter- und Obersummen) die zugehörige *Riemann-Summe*:

$$R_n(f) = f(v_1)(x_1 - x_0) + f(v_2)(x_2 - x_1) + \dots + f(v_n)(x_n - x_{n-1}).$$

Da die Funktionswerte $f(v_k)$ zwischen den minimalen und maximalen Werten von f in dem jeweiligen Intervall liegen, liegt die Riemannsumme $R_n(f)$ zwischen der Unter- und Obersumme der zugehörigen Intervallzerlegung: $U_n(f) \leq R_n(f) \leq O_n(f)$.

Ist nun die Funktion f über J integrierbar, so haben $U_n(f)$ und $O_n(f)$ denselben Grenzwert $\int_a^b f$, nach dem Schachtelungssatz muss dann auch die dazwischen liegende Riemann-Summe $R_n(f)$ gegen das Integral konvergieren und man erhält den folgenden umfassenden Satz:

Satz: Ist f im oben definierten Sinne über $J = [a, b]$ integrierbar, so sind alle Riemannsummen $R_n(f)$ (zu Zerlegungsfolgen \mathcal{Z}_n mit $h(\mathcal{Z}_n) \rightarrow 0$) konvergent mit dem Integral $\int_a^b f$ als gemeinsamem Grenzwert.

$$\int_a^b f = \lim_{n \rightarrow \infty} R_n(f).$$

Das Integral ist der gemeinsame Grenzwert aller Riemannsummen.

e. Integrierbare Funktionen. Wir wollen nun Beispiele für integrierbare Funktionen kennenlernen. Um eine Funktion f über einem Intervall $J = [a, b]$ integrieren zu können, muss zunächst einmal die Funktion an *allen* Stellen des Integrationsintervalles $J = [a, b]$ definiert sein, insbesondere auch an den Randstellen a und b ! Um die Ober- bzw. Untersummen bilden zu können, muss die Funktion über dem Intervall J *beschränkt* sein.

Da die Ober- bzw. Untersummenfolge bereits als konvergent nachgewiesen ist, ist die entscheidende Forderung bei der Integrierbarkeit die *Gleichheit* der beiden Grenzwerte. Dies kann man dann auch folgendermaßen charakterisieren:

Bemerkung: Eine beschränkte Funktion f über einem Intervall J ist genau dann *integrierbar*, wenn die Differenz von Ober- und Untersummen eine Nullfolge ist:

$$f \text{ integrierbar} \iff \lim_{n \rightarrow \infty} (O_n(f) - U_n(f)) = 0.$$

Eine erste, zugleich aber auch weitgehende Antwort auf die Frage nach Beispielen integrierbarer Funktionen ist der folgende Satz.

Satz: Jede über einem Intervall $J = [a, b]$ definierte und dort monotone Funktion f ist integrierbar.

Monotone Funktionen sind integrierbar.

Es gilt genauer für äquidistante Ober- und Untersummen:

$$O_n(f) - U_n(f) = \frac{b-a}{n} \cdot |f(b) - f(a)|. \quad (4)$$

Beweis: Es genügt den Zusatz zu beweisen, denn dann ist die Differenz von Ober- und Untersumme eine Nullfolge (n im Nenner, alle anderen Größen sind konstant!)

Sei für den Beweis von (4) die Funktion f über dem Intervall $J = [a, b]$ monoton wachsend. Dies bedeutet für jedes Teilintervall $J_k = [x_{k-1}, x_k]$ einer Zerlegung \mathcal{Z} , dass der kleinste Funktionswert am linken Rand x_{k-1} und der größte Wert am rechten Rand x_k des Teilintervalls J_k

erreicht wird. Also gilt mit den obigen Bezeichnungen $f(z_k) = f(x_{k-1})$ und $f(Z_k) = f(x_k)$ und folglich

$$U_n(f) = h(f(x_0) + \dots + f(x_{n-1})) \quad \text{und} \quad O_n(f) = h(f(x_1) + \dots + f(x_n)).$$

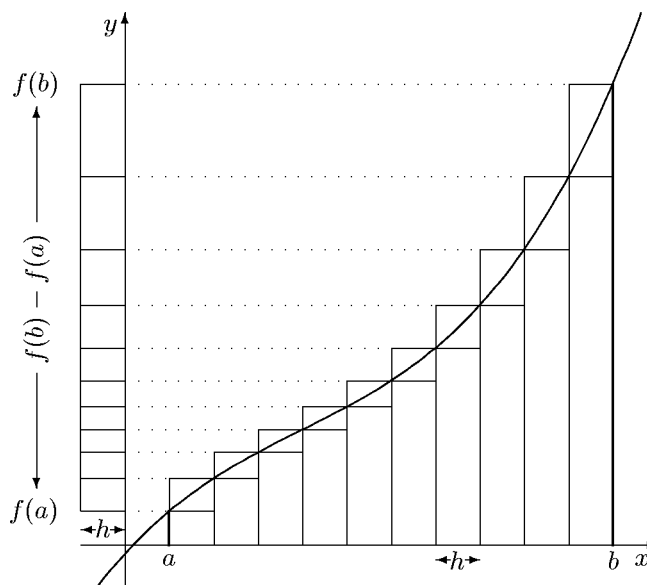
Bildet man die Differenz, so fallen die übereinstimmenden Summanden heraus und es ergibt sich

$$O_n(f) - U_n(f) = h \cdot (f(x_n) - f(x_0)) = h \cdot (f(b) - f(a)).$$

Die nebenstehende Skizze veranschaulicht das obige Resultat für monoton steigende Funktionen. Der Unterschied zwischen Ober- und Untersumme ist der Flächeninhalt der Rechtecke, die vom Graphen durchschnitten werden. Schiebt man diese Rechtecke (wie Bauklötze) zusammen, so erhält man das an der y -Achse skizzierte Rechteck. Dessen Breite ist die Streifenbreite h und die Höhe ist gerade $f(b) - f(a)$, so dass dieses Rechteck den Flächeninhalt $h \cdot (f(b) - f(a))$ hat.

Für eine monoton fallende Funktion erhält man entsprechend

$$O_n(f) - U_n(f) = h \cdot (f(a) - f(b)).$$



In jedem Falle erhält man die obige Formel und damit die Integrierbarkeit von f .

Neben den monotonen sind die stetigen Funktionen die zweite wichtige Klasse integrierbarer Funktionen.

Satz: Jede über einem Intervall $J = [a, b]$ definierte und dort stetige Funktion f ist über J integrierbar.

Für diesen Satz können wir den *Beweis* nur ein wenig plausibel machen. Streben in einer Zerlegungsfolge die Intervallbreiten $x_k - x_{k-1}$ gegen 0, so wird der Abstand der Stellen z_k und Z_k , die ja zwischen x_{k-1} und x_k liegen, ebenfalls gegen 0 streben. Da die Funktion f stetig ist, muss dann auch der Abstand $f(Z_k) - f(z_k)$ zwischen größtem und kleinstem Funktionswert in einem Teilintervall beliebig klein werden. Dies führt dazu, dass die Differenz von Ober- und Untersumme ebenfalls gegen 0 strebt.

Die Problematik dieses ‘Beweises’ liegt in den letzten Schritten, da nämlich zugleich die Anzahl der Summanden in Ober- und Untersumme unbegrenzt wächst. Um obigen ‘Beweis’ zu retten, muss man zeigen, dass man eine universelle, von k unabhängige Abschätzung für $f(Z_k) - f(z_k)$ finden kann. Dies ist der Fall!

§14 Die Berechnung von Integralen und der Hauptsatz

Wir wollen uns nun mit der exakten Berechnung von Integralen auseinandersetzen.

a. Berechnung mittels Ober-/Untersummen. Wenn wir auf der Basis der Definition Integrale berechnen, müssen wir Ober- oder Untersummen in Abhängigkeit von n bestimmen und dann den Grenzwert für $n \rightarrow \infty$ ermitteln. In einem ersten Beispiel wollen wir ein Integral mittels Ober- und Untersummen berechnen, dessen Wert wir aufgrund elementargeometrischer Überlegungen bereits kennen.

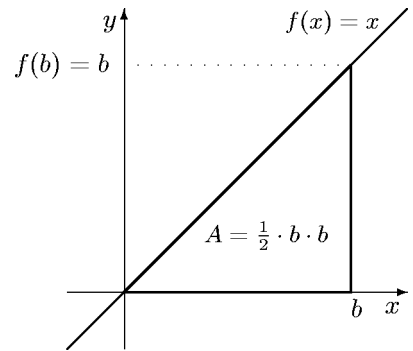
1. Beispiel: $f(x) = x$ und $0 = a < b$.

Gemäß der geometrischen Deutung des Integrals ist dieses Integral $\int_0^b f$ gerade der Flächeninhalt eines Dreiecks der Breite b und der Höhe $f(b) = b$ (siehe nebenstehende Skizze) und muss daher den Wert

$$\int_0^b x \, dx = \frac{1}{2}b^2$$

haben. Wir wollen dieses Integral nun einmal gemäß der Definition als Grenzwert von Obersummen berechnen.

Es ist also $a = 0$, $b > 0$. Die äquidistante Zerlegung von $J = [0, b]$ in n gleich breite Streifen hat dann die Daten (s. o.)



Streifenbreite: $h = \frac{b}{n}$,

Zerlegungsstellen: $x_k = kh = k \cdot \frac{b}{n}$ ($k = 0, \dots, n$).

Da die Funktion f monoton wächst, wird auf jedem Teilintervall $J_k = [x_{k-1}, x_k]$ der kleinste Wert am linken Rand x_{k-1} und der größte Wert am rechten Rand x_k angenommen. Also ist der größte Wert

$$f(Z_k) = f(x_k) = x_k = kh \quad (k = 1, \dots, n).$$

Damit ist die n -te Obersumme

$$\begin{aligned} O_n(f) &= h(f(Z_1) + \dots + f(Z_n)) = h(h + 2h + 3h + \dots + nh) \\ &= h^2(1 + 2 + 3 + \dots + n) = \frac{b^2}{n^2} \cdot (1 + 2 + 3 + \dots + n). \end{aligned}$$

Mit dieser Darstellung ist noch keine Berechnung des Grenzwertes der Obersummen möglich. Zwar konvergiert der erste Faktor b^2/n^2 gegen 0 für $n \rightarrow \infty$, aber zugleich wächst die Summe in der Klammer unbegrenzt.

An dieser Stelle benötigt man eine Summenformel für die in der Klammer auftretende Summe der ersten n natürlichen Zahlen. Dafür gilt

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Diese Formel haben wir bereits im Rahmen der ersten Übung des 3. Semesters bewiesen: Übereinstimmung der Folgen 5) und 9) von Übungsblatt (1). (Siehe auch Skript I §1 b., S. 2) Diese Formel kann man aber auch unmittelbar einsehen (Gauß-Anekdote!).

Mit dieser Formel ergibt sich nun

$$O_n(f) = \frac{b^2}{n^2} \cdot \frac{n(n+1)}{2} = \frac{b^2}{2} \cdot \frac{n+1}{n}.$$

Nach den Grenzwertsätzen gilt

$$\lim_{n \rightarrow \infty} \frac{n+1}{n} = 1.$$

(Wiederholen Sie die Berechnung der Grenzwerte im Unendlichen für rationale Funktionen!)
Damit erhalten wir insgesamt

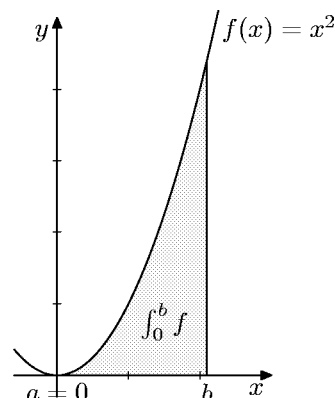
$$\int_0^b f = \lim_{n \rightarrow \infty} O_n(f) = \frac{b^2}{2} \cdot \lim_{n \rightarrow \infty} \frac{n+1}{n} = \frac{b^2}{2}.$$

Dieses Ergebnis ist im Einklang mit der elementargeometrischen Berechnung des Integrals.

2. Beispiel: $f(x) = x^2$ und $a = 0 < b$.

Hier berechnen wir erstmals den exakten Flächeninhalt eines krummlinig begrenzten Flächenstücks. Wir können die obigen Zerlegungsdaten $h = \frac{b}{n}$ und $x_k = k \cdot h$ weiterverwenden. Außerdem ist auch diese Funktion über dem Integrationsintervall $[0, b]$ monoton wachsend, also gilt wieder $f(Z_k) = f(x_k) = f(kh)$. Für $f(x) = x^2$ bedeutet dies dann $f(Z_k) = f(kh) = (kh)^2$, also

$$\begin{aligned} O_n(f) &= h \cdot (h^2 + (2h)^2 + \dots + (nh)^2) \\ &= h^3 \cdot (1^2 + 2^2 + 3^2 + \dots + n^2). \end{aligned}$$



Hier benötigt man nun eine Summationsformel für die Summe der ersten n Quadratzahlen:

$$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

(Siehe dazu auch Skript S. 3.) Mit dieser Formel ergibt sich dann

$$O_n(f) = \frac{b^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{b^3}{6} \cdot \frac{(n+1)(2n+1)}{n^2}.$$

Der letzte Bruch stellt in Abhängigkeit von n eine rationale Funktion dar mit gleichem Zähler- und Nennergrad. Also ist der Quotient der führenden Koeffizienten der Grenzwert für $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \frac{(n+1)(2n+1)}{n^2} = \frac{2}{1} = 2$$

und damit

$$\int_0^b f = \lim_{n \rightarrow \infty} O_n(f) = \frac{b^3}{6} \cdot 2 = \frac{1}{3}b^3.$$

Damit haben wir eine erste Flächenformel für ein krummlinig begrenztes Flächenstück:

$$\boxed{\int_0^b x^2 dx = \frac{1}{3}b^3.}$$

In gleicher Weise erhält man unter Verwendung der Summationsformel $\sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4}$

(siehe S. 3) die Integrationsformel $\int_0^b x^3 dx = \frac{1}{4}b^4$. Damit haben wir für die ersten Potenzfunktionen folgende Integrationsformeln für $b > 0$:

$$\boxed{\int_0^b 1 dx = b, \quad \int_0^b x dx = \frac{1}{2}b^2, \quad \int_0^b x^2 dx = \frac{1}{3}b^3, \quad \int_0^b x^3 dx = \frac{1}{4}b^4.}$$

(Dabei ergibt sich die erste Formel als Flächenformel für ein Rechteck der Breite b und der Höhe 1.)

b. Erste Integrationsregeln. Wir wollen in diesem Abschnitt einige allgemeine Integrationsregeln kennenlernen, mit deren Hilfe man Integrale zerlegen und auf einfachere zurückführen kann.

Satz: (Intervalladditivität) *Ist $a < b < c$, so ist jede Funktion f , die über den Intervallen $[a, b]$ und $[b, c]$ integrierbar ist, auch über $[a, c]$ integrierbar und es gilt*

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Die Formel ist für Funktionen mit positiven Werten geometrisch unmittelbar einsichtig. Will man den Satz für beliebige Funktionen und einschließlich der Integrierbarkeitsaussage beweisen, betrachtet man Zerlegungen der beiden Teilintervalle und fügt diese zu einer Zerlegung des Gesamtintervalls zusammen. Dabei ist man dann gezwungen nicht nur äquidistante, sondern allgemeinere Intervallzerlegungen zu betrachten.

Die obige Relation bleibt sogar bei *beliebiger* Lage der drei Integrationsgrenzen a, b, c gültig, wenn man folgende Vereinbarung trifft:

$$\int_a^b f = - \int_b^a f.$$

Diese Definition ist verständlich, wenn man einmal die Formel für äquidistante Untersummen betrachtet, in der der Faktor $h = \frac{b-a}{n}$ auftritt. Wenn hier nun die Rollen von a und b vertauscht werden, ändert sich das Vorzeichen. Wir werden später mit der Integralformel (S. 79) ein weiteres Argument dafür kennenlernen, dass diese Ausdehnung der Integraldefinition sinnvoll ist!

Die weiteren Integrationsregeln betreffen die Abhängigkeit des Integrals von der zu integrierenden Funktion (dem *Integranden*).

Satz: *Sind f, g zwei integrierbare Funktion über dem Intervall $[a, b]$ und ist $c \in \mathbb{R}$ eine Konstante, so sind auch cf und $f + g$ integrierbar und es gelten die*

$$\text{Faktorregel: } \int_a^b cf(x) dx = c \cdot \int_a^b f(x) dx,$$

$$\text{Summenregel: } \int_a^b (f(x) \pm g(x)) dx = \int_a^b f(x) dx \pm \int_a^b g(x) dx.$$

Um dieses allgemeine Resultat zu beweisen, gehen wir auf die Definition zurück und untersuchen die Ober- und Untersummen bzw. allgemein beliebige Riemannsummen. Wir berechnen für beliebige Intervallzerlegungen \mathcal{Z} und Zwischenstellen v_k die Riemannsummen:

$$\begin{aligned} R_n(cf) &= \sum_{k=1}^n cf(v_k)(x_k - x_{k-1}) = c \cdot \sum_{k=1}^n f(v_k)(x_k - x_{k-1}) = c \cdot R_n(f), \\ R_n(f \pm g) &= \sum_{k=1}^n (f(v_k) \pm g(v_k))(x_k - x_{k-1}) = \sum_{k=1}^n f(v_k)(x_k - x_{k-1}) \pm \sum_{k=1}^n g(v_k)(x_k - x_{k-1}) \\ &= R_n(f) \pm R_n(g). \end{aligned}$$

Eine Riemannsumme zu $f + g$ ist also die Summe der Riemannsummen zu f und zu g ; entsprechend für die Differenz und Vielfache.

Nun konvergieren für die integrierbaren Funktionen f und g alle Riemannsummen gegen das zugehörige Integral. Mit Hilfe der Grenzwertsätze folgt also aus den obigen Beziehungen sofort:

$$\lim_{n \rightarrow \infty} R_n(cf) = c \cdot \lim_{n \rightarrow \infty} R_n(f) = c \cdot \int_a^b f,$$

$$\lim_{n \rightarrow \infty} R_n(f \pm g) = \lim_{n \rightarrow \infty} R_n(f) \pm \lim_{n \rightarrow \infty} R_n(g) = \int_a^b f \pm \int_a^b g.$$

Also haben alle Riemannsummen $R_n(cf)$ denselben Grenzwert $c \cdot \int_a^b f$. Dies bedeutet, dass cf integrierbar ist mit

$$\int_a^b cf = \lim_{n \rightarrow \infty} R_n(cf) = c \cdot \int_a^b f.$$

Ebenso folgt die Summenregel.

Abschließend sei in diesem Abschnitt noch die folgende *Abschätzung* für Integrale genannt.

Satz: Sind $a < b$ und f, g integrierbar über dem Intervall $J = [a, b]$, so gilt:

$$f(x) \leq g(x) \text{ für } a \leq x \leq b \implies \int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Daraus erhält man (sogar für beliebige $a \neq b$):

$$s \leq f(x) \leq S \implies s \leq \frac{1}{b-a} \cdot \int_a^b f(x) dx \leq S.$$

Begründung: Mit den obigen Regeln erhält man

$$\int_a^b g - \int_a^b f = \int_a^b (g - f) = \int_a^b h.$$

Die Funktion $h = g - f$ nimmt über J keine negativen Werte an, also ist auch jede Untersumme von h größer oder gleich 0. Dann kann auch das Integral, das ja der Grenzwert der Untersummen ist, nicht negativ sein:

$$h(x) \geq 0 \text{ für alle } x \in J \implies \int_a^b h \geq 0.$$

Damit ist die erste Abschätzung bewiesen.

Wendet man diese erste Abschätzung auf $s \leq f(x) \leq S$ an, so erhält man für $a < b$:

$$s(b-a) = \int_a^b s dx \leq \int_a^b f(x) dx \leq \int_a^b S dx = S(b-a)$$

Nach Division durch $b-a$ ($> 0!$) ergibt sich die zweite Abschätzung. Ist nun $b < a$, so gilt entsprechend $s \leq \frac{1}{a-b} \int_b^a f \leq S$, woraus wegen

$$\frac{1}{a-b} \cdot \int_b^a f = \frac{(-1)}{b-a} \cdot (-1) \int_a^b f = \frac{1}{b-a} \int_a^b f$$

die Behauptung auch in diesem Fall folgt.

c. Integralfunktionen und der Hauptsatz. Dieser Satz ist das zentrale Resultat der Differential- und Integralrechnung. Er stellt eine überraschende Verbindung her zwischen zwei scheinbar völlig verschiedenen Fragestellungen und Konzepten. Auf der einen Seite die Differentialrechnung, die den Begriff der *Steigung* von Kurven und darauf aufbauend die Untersuchung von Funktionsverläufen zum Thema hat, und auf der anderen Seite die Integralrechnung, deren zentraler Begriff der *Flächeninhalt* ist.

Um den Hauptsatz formulieren zu können, betrachten wir das Integral $\int_a^b f$ einer Funktion f ebenfalls als Funktion, und zwar in Abhängigkeit von der oberen Grenze. Um zu betonen, dass wir die obere Grenze als Funktionsvariable betrachten, bezeichnen wir diese mit x . Man muss dann aber die Integrationsvariable anders benennen.

Definition: Ist f über einem Intervall J integrierbar und $a \in J$, so definiert man die Integralfunktion (zum Integranden f und Startwert a)

$$\text{Integralfunktion } I_a(x) = \int_a^x f = \int_a^x f(t) dt$$

Beachten Sie, dass diese Integralfunktion nicht nur für $x > a$, sondern auch für $x < a$ definiert ist: $\int_a^x f = -\int_x^a f$. Und für $x = a$ gilt

$$I_a(x) = \int_a^a f = 0.$$

Insbesondere haben Integralfunktionen also immer eine Nullstelle, und zwar beim Startwert a .

Der zentrale Hauptsatz der Differential- und Integralrechnung besagt nun:

Hauptsatz: Ist eine Funktion f stetig über einem Intervall J , so ist jede Integralfunktion I_a von f über dem Intervall J ($a \in J$) differenzierbar und ihre Ableitung ist die Ausgangsfunktion f :

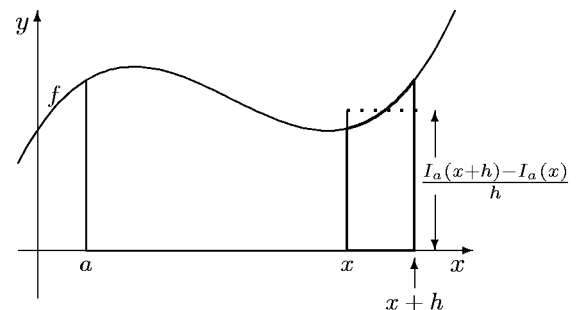
$$I_a'(x) = f(x) \quad \text{für alle } x \in J.$$

Beweis: Da eine Integralfunktion I_a durch einen neuen, komplizierten Prozess definiert ist, können wir zu ihrer Ableitung keine der bekannten Ableitungsregeln nutzen; wir müssen auf die Definition zurück: Die Ableitung ist Grenzwert der Differenzenquotienten. Dies bedeutet hier:

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

Um den Hauptsatz zu beweisen, muss man also zeigen, dass dieser Grenzwert für jedes x genau $f(x)$ ist.

Wir wollen vor dem eigentlichen Beweis die geometrische Grundidee in folgender spezieller Situation verdeutlichen: f habe nur positive Werte und es sei $a < x < x+h$. Dann kann man die Differenz $I_a(x+h) - I_a(x)$ geometrisch veranschaulichen als den Flächeninhalt des nebenstehend skizzierten Streifens. Den Differenzenquotient $\frac{I_a(x+h) - I_a(x)}{h}$ erhält man dann, indem man den Flächeninhalt dieses Streifens durch h , d. h. durch seine Breite dividiert. Dieser Quotient gibt also die Höhe eines flächengleichen Rechtecks an (siehe Skizze).



Wenn nun die Streifenbreite h gegen 0 strebt, muss wegen der Stetigkeit von f die Höhe des skizzierten Rechtecks gegen $f(x)$ streben:

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

Da diese geometrischen Überlegungen nur für eine spezielle Konstellation ($f(x) > 0$, $a < x < x+h$) gültig sind, wollen wir nun einen allgemeingültigen Beweis führen. Er basiert auf den Integrationsregeln aus dem vorangehenden Abschnitt. Aufgrund der Intervalladditivität gilt

$$I_a(x+h) - I_a(x) = \int_a^{x+h} f - \int_a^x f = \int_x^{x+h} f.$$

Also ist der Differenzenquotient darstellbar als

$$\frac{I_a(x+h) - I_a(x)}{h} = \frac{1}{h} \int_x^{x+h} f.$$

Ist nun $s = f(z)$ das Minimum der Funktionswerte von f über dem Intervall von x bis $x+h$ und $S = f(Z)$ das entsprechende Maximum ($x \leq z, Z \leq x+h$), so gilt gemäß der Abschätzungsformel (S. 76):

$$f(z) = s \leq \frac{1}{h} \int_x^{x+h} f \leq S = f(Z).$$

Wegen $x \leq z, Z \leq x+h$ müssen die Stellen z und Z gegen x streben, wenn h gegen 0 strebt. Wegen der Stetigkeit von f ergibt sich dann: $f(z) \rightarrow f(x)$ und $f(Z) \rightarrow f(x)$. Da also beide Schranken gegen denselben Grenzwert $f(x)$ streben, muss nach dem Schachtelungssatz (siehe S. 9) auch der dazwischen eingezwängte Differenzenquotient konvergieren, und sein Grenzwert ist dann ebenfalls $f(x)$:

$$\begin{array}{ccccc} f(z) & \leq & \frac{I_a(x+h) - I_a(x)}{h} & \leq & f(Z) \\ \text{für } h \rightarrow 0: & & \downarrow & & \downarrow \\ f(x) & \leq & I'_a(x) & \leq & f(x) \end{array}$$

Also $I'_a(x) = f(x)$ und der Hauptsatz ist bewiesen.

d. Stammfunktionen und die Integralformel. Eine wichtige Konsequenz des Hauptsatzes ist die folgende, äußerst effektive Methode zur Berechnung von Integralen mittels Stammfunktionen.

Definition: Unter einer Stammfunktion einer Funktion f versteht man eine differenzierbare Funktion F , deren Ableitung f ist:

$$F \text{ ist Stammfunktion von } f \iff F' = f.$$

Hinweis: Ist F eine Stammfunktion von f , so sind alle Funktionen $F(x) + c$ ($c \in \mathbb{R}$ eine Konstante) ebenfalls Stammfunktionen von f , denn die Ableitung einer Konstanten ist 0. Also hat eine Funktion f immer unendlich viele Stammfunktionen – oder gar keine. Letzteres kommt aber für die Ihnen bekannten Funktionen kaum vor, denn es gilt:

Folgerung aus dem Hauptsatz: Jede über einem Intervall J stetige Funktion besitzt eine Stammfunktion.

Dieses Resultat ist um so bemerkenswerter, als stetige Funktionen weit komplizierter sein können, als die Ihnen bisher bekannten Funktionen (die in der Regel differenzierbar sind).

Für das Folgende ist von fundamentaler Bedeutung, dass man für Funktionen über einem Intervall eine vollständige Übersicht über alle Stammfunktionen hat:

Satz: Sind F_1 und F_2 zwei über einem Intervall J definierte Stammfunktionen derselben Funktion f , so gibt es eine Konstante $c \in \mathbb{R}$ mit

$$F_1(x) = F_2(x) + c \quad \text{für alle } x \in J.$$

Der *Beweis* ergibt sich unmittelbar aus einem fundamentalen Satz der Differentialrechnung: Die Differenzfunktion $F_1 - F_2$ hat die Ableitung $F_1' - F_2' = f - f = 0$, ist also gemäß Satz (2.7) (S. 44) über dem Intervall (!) J konstant:

$$F_1(x) - F_2(x) = c \quad \text{für alle } x \in J.$$

Dies ist aber genau die Behauptung.

Kombiniert man nun dieses Ergebnis mit dem Hauptsatz, erhält man die folgende wichtige

Integralformel: *Es sei J ein Intervall und f eine stetige Funktion über J . Dann gilt für beliebige $a, b \in J$:*

$$\int_a^b f(x) dx = F(b) - F(a) \quad \text{für jede Stammfunktion } F \text{ von } f.$$

Beweis: Nach dem Hauptsatz ist I_a eine Stammfunktion von f über dem Intervall J . Andererseits ist nach Voraussetzung F ebenfalls eine Stammfunktion von f über demselben Intervall J . Also existiert eine Konstante $c \in \mathbb{R}$ mit der Eigenschaft

$$I_a(x) = \int_a^x f = F(x) + c \quad \text{für alle } x \in J. \quad (*)$$

Es gilt nun nur noch c , die sog. *Integrationskonstante* zu bestimmen. Dazu beachten wir, dass Integralfunktionen immer eine Nullstelle haben, nämlich

$$I_a(a) = \int_a^a f = 0.$$

Indem wir (*) für $x = a$ auswerten, erhalten wir

$$0 = F(a) + c, \quad \text{also } c = -F(a),$$

und damit

$$\int_a^x f = F(x) - F(a) \quad \text{für alle } x \in J.$$

Indem man $x = b$ setzt erhält man schließlich die behauptete Integralformel.

In den Übungen haben wir ein Vielzahl von Beispielen für die Effektivität dieser Integralformel kennengelernt. Bei der Anwendung der Integralformel muss man zunächst eine Stammfunktion F für den Integranden f finden und diese dann an den Integrationsgrenzen auswerten. Dabei haben wir die folgende Schreibweise verwendet:

$$\int_a^b f = \int_a^b f(x) dx = \left[F(x) \right]_a^b = F(b) - F(a)$$

e. Bestimmung von Stammfunktionen. Die Berechnung von Integralen auf diesem Wege steht und fällt mit der Kenntnis oder Ermittlung von Stammfunktionen zu gegebenen Funktionen. Da es sich dabei um den zur Ableitung umgekehrten Prozess handelt, kann man aus Ableitungsregeln immer auch Regeln zur Ermittlung von Stammfunktionen gewinnen. Erste Beispiele sind

Satz: (Stammfunktionen für Potenzfunktionen) Für $n \in \mathbb{R}$, $n \neq -1$ gilt:

$$F(x) = \frac{1}{n+1} x^{n+1} \text{ ist ein Stammfunktionsterm für } f(x) = x^n.$$

Satz: (Stammfunktionsregeln) Wir setzen voraus, dass F, G Stammfunktionen für f, g sind. Dann gelten die folgenden Regeln für Stammfunktionen:

a) Summenregel:

$$F \pm G \text{ ist Stammfunktion von } f \pm g.$$

b) Faktorregel: Für beliebige Konstanten $c \in \mathbb{R}$ gilt:

$$cF \text{ ist Stammfunktion von } cf.$$

c) Lineare Substitutionsregel: Für $m, n \in \mathbb{R}$, $m \neq 0$ gilt:

$$\frac{1}{m}F(mx + n) \text{ ist ein Stammfunktionsterm für } f(mx + n).$$

Aufgrund der Regeln a), b) kann man für alle ganzrationalen Funktionen unmittelbar eine Stammfunktion angeben und dann mit der Integralformel entsprechende Integrale problemlos berechnen:

Folgerung: Ist $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$ eine ganzrationale Funktion, so ist durch

$$F(x) = \frac{a_n}{n+1} x^{n+1} + \frac{a_{n-1}}{n} x^n + \dots + \frac{a_2}{3} x^3 + \frac{a_1}{2} x^2 + a_0 x$$

eine Stammfunktion von f gegeben.

Man *beweist* alle diese Behauptungen, indem man die vorgebliche Stammfunktion ableitet und feststellt, dass sich die gewünschte Funktion ergibt. Generell ist der *Nachweis*, dass eine *vorgegebene* Funktion tatsächlich Stammfunktion ist, einfach durch Ableiten zu führen (s. o.). Viel schwieriger ist es jedoch, eine Stammfunktion zu *ermitteln*. Anders als das Ableiten führt dies bereits bei rationalen Funktionen zu erheblichen Problemen. Während die Ableitung einer rationalen Funktion stets wieder eine rationale Funktion ist (Quotientenregel), gilt bei der Suche nach Stammfunktionen nichts Entsprechendes: Es gibt rationale Funktionen, die keine rationale Stammfunktion besitzen. Ganz konkret gilt:

Satz: Die Funktion $f(x) = \frac{1}{x}$ hat über dem Intervall $]0, \infty[$ als Stammfunktion $\ln(x)$, den natürlichen Logarithmus! Über dem gesamten Definitionsbereich $\mathbb{D}_f = \mathbb{R} \setminus \{0\}$ ist eine Stammfunktion gegeben durch

$$F(x) = \ln(|x|).$$

Beweis: Eine der fundamentalen Eigenschaften von \ln war die Ableitungsregel $\ln'(x) = \frac{1}{x}$. Also ist \ln (definiert über $]0, \infty[$) eine Stammfunktion von f . Für die weitere Behauptung muss man über dem anderen Teilintervall $] - \infty, 0[$ die Funktion

$$F(x) = \ln(|x|) = \ln(-x) \quad \text{für } x < 0$$

ableiten. Dies ergibt nach der Kettenregel:

$$F'(x) = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

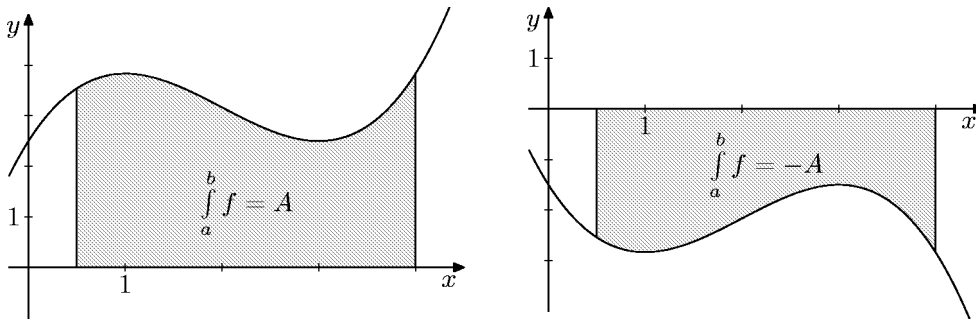
Also hat das angegebene F überall auf seinem Definitionsbereich $\mathbb{D}_F = \mathbb{R} \setminus \{0\}$ die gewünschte Ableitung $F'(x) = \frac{1}{x}$.

Warnung: Nicht alle Stammfunktionen von $f(x) = \frac{1}{x}$ sind von der Form $\ln(|x|) + c$ ($c \in \mathbb{R}$). Auch $G(x) = \ln(|x|) + \text{sign}(x)$ ist eine Stammfunktion von f , ohne dass $G(x) - F(x) = \text{sign } x$ über dem Definitionsbereich $D_f = D_F = \mathbb{R} \setminus \{0\}$ konstant ist. Die Ursache dafür ist, dass

hier der Definitionsbereich *kein Intervall* ist! (Der Satz über den Zusammenhang zwischen den verschiedenen Stammfunktionen einer Funktion (siehe S. 78) gilt nur über Intervallen!)

§15 Fläche, Volumen, Bogenlänge

a. Flächen zwischen Graph und x -Achse. Unsere Einführung des Integrals in §13 orientierte sich am Problem der Flächenberechnung. Dabei hatten wir die Fläche zwischen dem Graphen einer Funktion und der x -Achse in den Grenzen von a bis b betrachtet. Wir hatten gesehen (siehe S. 70), dass dieser Flächeninhalt gleich dem Integral $\int_a^b f$ ist. Allerdings hatten wir bei der Herleitung benutzt, dass $a < b$ ist und die Funktion f über dem Integrationsintervall $J = [a, b]$ ausschließlich Werte ≥ 0 hat (siehe nachfolgende Skizze links). Dadurch war sichergestellt, dass Unter- bzw. Obersummen als Summen von Rechtecksflächen interpretiert werden konnten und ihr gemeinsamer Grenzwert gleich dem Flächeninhalt A ist. Hat f *ausschließlich* Werte ≤ 0 , so sind für $a < b$ auch alle Ober- und Untersummen und das Integral $\int_a^b f \leq 0$. In diesem Falle gibt das Integral dann gerade das Negative des Flächeninhalts A an (siehe rechte Skizze).



Wir fassen zusammen:

Bemerkung: Haben die Funktionswerte von f im Integrationsintervall $J = [a, b]$ ein einheitliches Vorzeichen, so ist das Integral $\int_a^b f$ gleich dem Flächeninhalt A des Flächenstücks zwischen Graph und x -Achse in den Grenzen von a bis b , versehen mit dem entsprechenden Vorzeichen von f .

$$\int_a^b f = \begin{cases} +A & \text{falls } f(x) \geq 0 \text{ für alle } x \in J = [a, b], \\ -A & \text{falls } f(x) \leq 0 \text{ für alle } x \in J = [a, b]. \end{cases}$$

Warnung: Flächeninhalte sind immer positiv, Integrale können auch negativ sein!

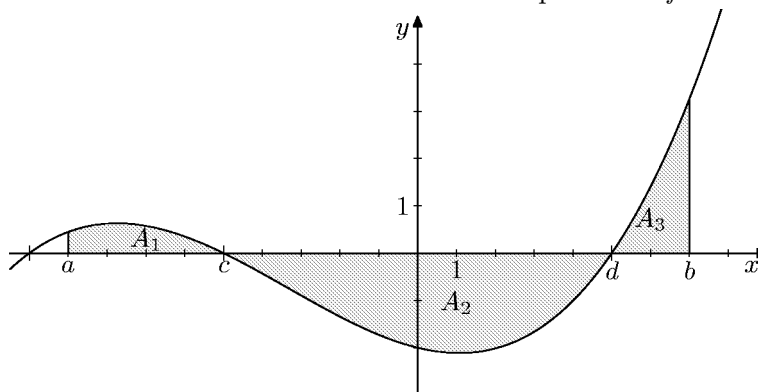
Wir können nun diese Tatsache umkehren zu einer Methode, Flächeninhalte durch Integrale zu berechnen. Für Funktionen, die im Integrationsintervall ihr Vorzeichen *nicht* ändern, ist der Flächeninhalt A bis auf das Vorzeichen gleich dem Integral, also:

Für Funktionen, die über einem Intervall $[a, b]$ ihr Vorzeichen *nicht* ändern, ist der Flächeninhalt A der Fläche zwischen dem Graphen von f und der x -Achse im Bereich zwischen a und b gleich dem Betrag des Integrals:

$$A = \left| \int_a^b f(x) dx \right| .$$

Will man für beliebige Funktionen diesen Flächeninhalt berechnen, so muss man das Intervall $[a, b]$ in einzelne Abschnitte unterteilen, in denen f das Vorzeichen nicht ändert (in der Skizze die Intervalle $[a, c]$, $[c, d]$ und $[d, b]$) und darauf dann die obige Formel anwenden. Die Grenzen dieser Abschnitte sind die Vorzeichenwechselstellen von f im Intervall $[a, b]$. Mit den Bezeichnungen

der Skizze ist dann der Flächeninhalt A zwischen dem Graphen von f und der x -Achse in den



Grenzen von a bis b

$$A = A_1 + A_2 + A_3 = \left| \int_a^c f \right| + \left| \int_c^d f \right| + \left| \int_d^b f \right|.$$

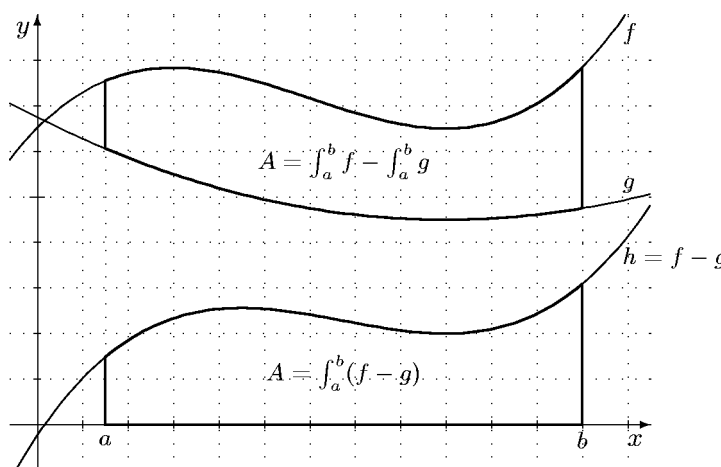
Wir fassen zusammen:

Um die Fläche zwischen dem Graphen einer Funktion f und der x -Achse in gegebenen Grenzen a, b zu bestimmen, muss man also:

1. Die Nullstellen von f mit VZW im Innern von $[a, b]$ ermitteln.
2. Von Nullstelle zu Nullstelle bzw. Rand integrieren.
3. Die Beträge der ermittelten Integrale addieren.

b. Flächen zwischen Graphen. Wir wollen nun allgemeiner Flächen zwischen zwei Graphen berechnen. Wir werden sehen, dass sich dieses Problem auf das soeben gelöste zurückführen lässt. Es gilt nämlich der folgende

Satz: Sind f und g zwei integrierbare Funktionen über einem Intervall $[a, b]$, so ist die Fläche zwischen ihren Graphen in den Grenzen von a bis b genauso groß wie die Fläche zwischen dem Graphen der Differenzfunktion $h = f - g$ und der x -Achse in denselben Grenzen.

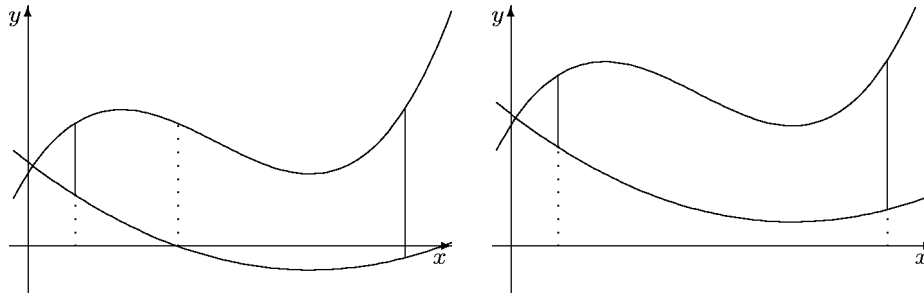


Beweis: Zunächst unterteilen wir das Integrationsintervall in Teilintervalle, in denen $f - g$ sein Vorzeichen nicht wechselt. In jedem dieser Abschnitte verläuft dann eine der Funktionen ständig oberhalb der anderen, etwa f oberhalb von g wie in obiger Skizze. Unter diesen

Umständen ist die gesuchte Fläche A zwischen beiden Graphen gerade die Differenz $\int_a^b f - \int_a^b g$ zweier Integrale und aufgrund der Summenregel für Integrale gilt dann:

$$A \stackrel{(*)}{=} \int_a^b f - \int_a^b g = \int_a^b (f - g) = \int_a^b h \quad \text{falls } f(x) \geq g(x) \text{ über } [a, b].$$

Dabei ist (*) geometrisch unmittelbar einsichtig, wenn – wie in obiger Skizze – g keine negativen Werte hat. Aber auch wenn g sein Vorzeichen ändert (siehe unten links), bleibt die Beziehung (*) gültig. Um dies zu erkennen, verschiebt man *beide* Funktionsgraphen gemeinsam so weit nach

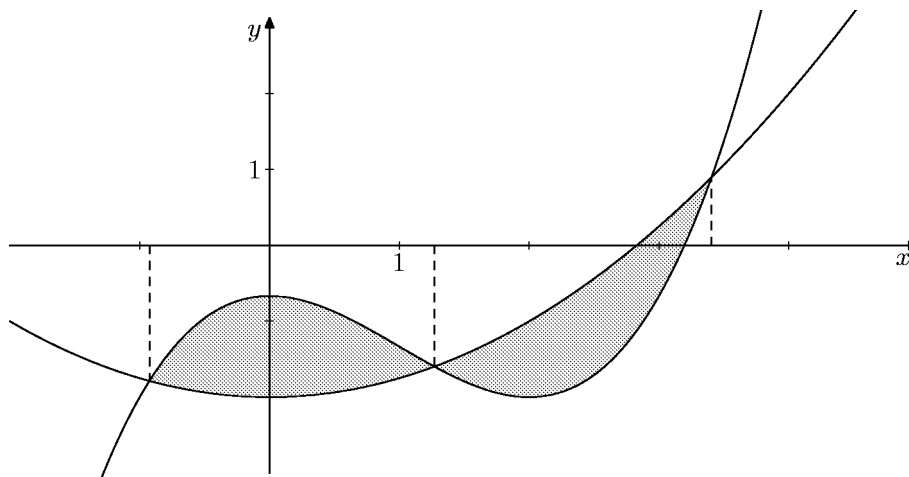


oben, dass im betrachteten Bereich beide Funktionen oberhalb der x -Achse verlaufen. Bei einer solchen Verschiebung ändern sich weder die eingeschlossene Fläche noch die Differenzfunktion $f - g$, also auch nicht deren Integral. Man kann daher stets die im rechten Bild skizzierte Situation erreichen, in der wir die obige Beziehung (*) und damit den Satz bewiesen hatten.

Durch diesen Satz ist die Berechnung der Fläche zwischen den Graphen zweier Funktionen f, g zurückgeführt auf die oben beschriebene Berechnung der Fläche zwischen einem Graphen (dem der Differenzfunktion $h = f - g$) und der x -Achse:

- Um die Fläche zwischen den Graphen zweier Funktionen f, g in vorgegebenen Grenzen a, b zu bestimmen, muss man also:
1. Die Differenzfunktion $h = f - g$ bilden.
 2. Für $h(!)$ die Nullstellen mit VZW im Bereich $[a, b]$ ermitteln.
 3. Die Funktion h von Nullstelle zu Nullstelle bzw. Rand integrieren.
 4. Die Beträge der ermittelten Integrale addieren.

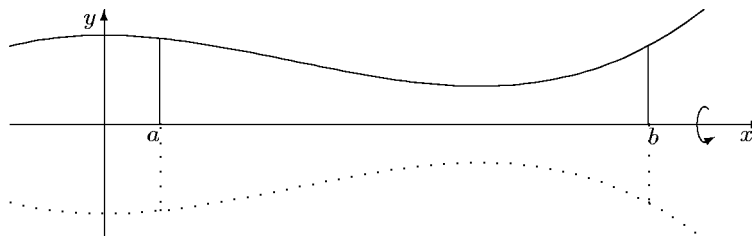
Schließlich wollen wir noch die Fläche ermitteln, die von zwei Graphen *eingeschlossen* wird. Darunter versteht man den Bereich, der *rundum* von den Graphen begrenzt wird, also den Bereich zwischen den beiden Graphen vom ersten bis zum letzten Schnittpunkt.



Um die von zwei Funktionsgraphen *eingeschlossene* Fläche zu bestimmen, muss man also:

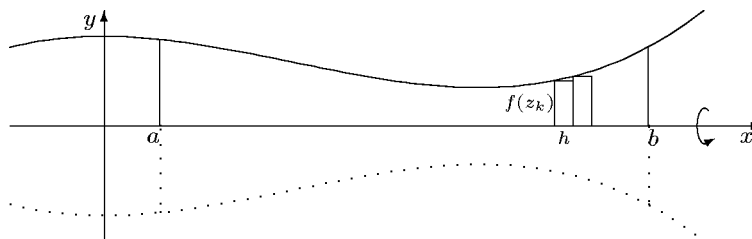
1. Die Differenzfunktion $h = f - g$ bilden.
2. Für h sämtliche Nullstellen ermitteln.
3. Von Nullstelle zu Nullstelle integrieren.
4. Die Beträge der ermittelten Integrale addieren.

c. Volumina von Rotationskörpern. Nicht nur Flächen, sondern auch Volumina kann man durch Integrale berechnen. Wir wollen hier sog. Rotationskörper betrachten. Sie entstehen



durch *Rotation* einer Fläche *um die x-Achse*. Wir wollen das Volumen des sich ergebenden Rotationskörpers bestimmen, wenn das rotierende Flächenstück in den Grenzen von a bis b durch den Graphen einer stetigen Funktion f begrenzt ist.

Um dieses Volumen zu bestimmen, schöpft man den Rotationskörper durch die Rotationskörper zu Untersummen aus. Deren Volumen ist berechenbar, denn es setzt sich aus lauter ‘Scheiben’ zusammen, deren Höhe gerade die Breite h der Intervalle ist und deren Radius der entsprechende Funktionswert $f(z_k)$ (z_k jeweils eine Stelle in einem der Streifen). Als Volumen



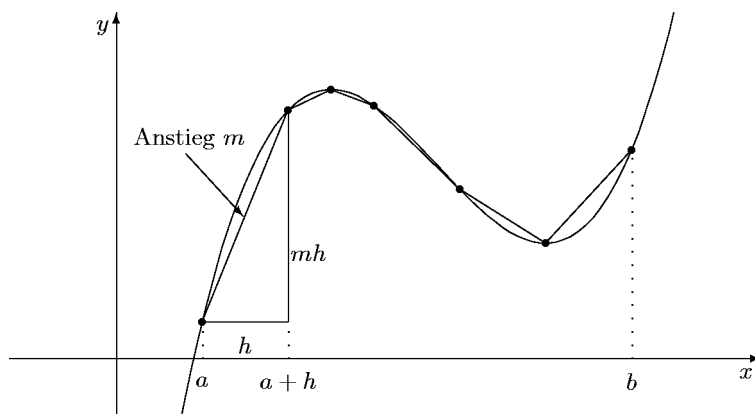
einer solchen Scheibe erhält man (unter Verwendung der Flächenformel für Kreise)

$$\pi \cdot f^2(z_k) \cdot h = \pi \cdot (f(z_k))^2 \cdot h.$$

Summiert man nun diese Volumina für sämtliche Streifen der Unterteilung auf, so erhält man gerade eine Untersumme für das Integral $\int_a^b \pi \cdot f^2(x) dx$. Für $n \rightarrow \infty$ bzw. $h \rightarrow 0$ konvergieren die Untersummen dann gegen dieses Integral, das damit das gesuchte Volumen darstellt:

Rotationsvolumen: $V = \pi \int_a^b f^2(x) dx.$

d. Die Bogenlänge. Aber auch zur Berechnung der Länge von Kurvenstücken kann man das Integral verwenden. Das Kurvenstück sei als Graph einer Funktion f in den Grenzen von a bis b gegeben. Zur Berechnung unterteilen wir es durch Zwischenpunkte und verbinden diese durch Geradenstücke (*Sehnen*), wie in der Skizze angegeben.



Die Länge einer einzelnen Sehne berechnen wir mit Hilfe des Satzes des Pythagoras wie folgt: Hat die waagerechte Kathete die Länge h , so hat die senkrechte die Länge $m \cdot h$, wenn m der *Anstieg der Sehne* ist. Also ist die Länge der Sehne

$$l_{\text{Sehne}} = \sqrt{h^2 + (mh)^2} = h \cdot \sqrt{1 + m^2}, \quad m \text{ der Sehnenanstieg.}$$

Man kann nun zeigen (siehe den nachfolgenden Mittelwertsatz), dass der Sehnenanstieg m im k -ten Streifen zugleich auch Tangentenanstieg von f an einer gewissen *Zwischenstelle* v_k in dem Streifen ist, also die Sehnenlänge folgende Form annimmt:

$$l_{\text{Sehne}} = h \cdot \sqrt{1 + (f'(v_k))^2}.$$

Summiert man nun alle diese Längen auf, so erhält man eine Riemannsumme R_n für die Funktion $\sqrt{1 + (f'(x))^2}$. Man definiert daher die Bogenlänge als den Grenzwert für $h \rightarrow 0$ dieser Riemannsummen. Dies ist aber gerade das zugehörige Integral (wenn es existiert):

Bogenlänge: $l = \int_a^b \sqrt{1 + (f'(x))^2} dx.$

Mittelwertsatz

...**der Integralrechnung**: Ist f eine stetige Funktion über $[a, b]$, so gilt

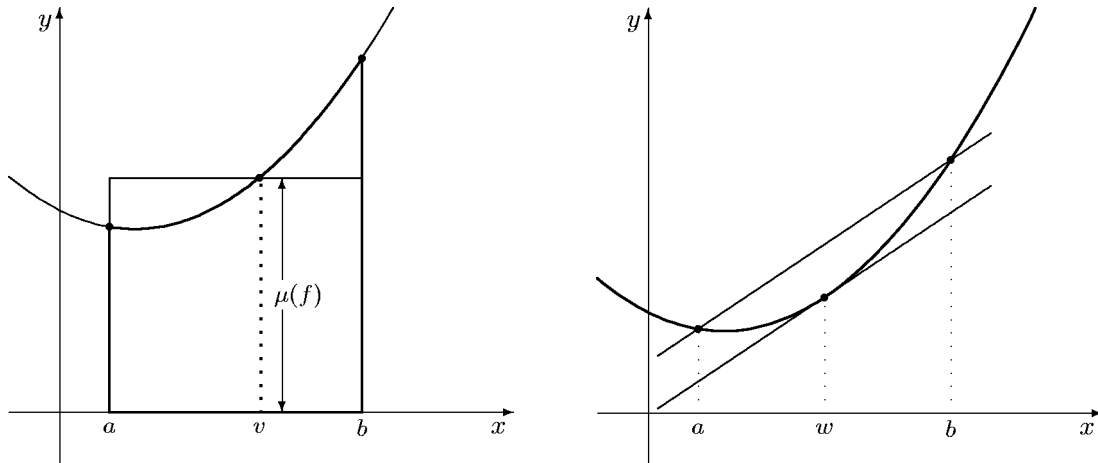
$$\int_a^b f(x) dx = f(v) \cdot (b - a) \quad \text{für ein passendes } v \in [a, b].$$

...**der Differentialrechnung**: Ist f eine Funktion mit stetiger Ableitung über $[a, b]$, so gilt:

$$\frac{f(b) - f(a)}{b - a} = f'(w) \quad \text{für ein passendes } w \in [a, b].$$

Die nachfolgenden Skizzen veranschaulichen beide Aussagen. Das Rechteck in der linken Skizze hat den Flächeninhalt $f(v) \cdot (b - a)$. Dieser stimmt mit dem Inhalt des dick umrandeten Flächenstücks unter dem Graphen von f überein. Man nennt den Wert $f(v)$ den *Mittelwert* von f über dem Intervall von a bis b :

Mittelwert: $\mu(f) = f(v) = \frac{1}{b - a} \int_a^b f$



Die rechte Skizze veranschaulicht den Mittelwertsatz der Differentialrechnung. Der Differenzenquotient $\frac{f(b)-f(a)}{b-a}$ stellt den Anstieg der Sekante dar. Diese stimmt mit einem Ableitungswert $f'(w)$, also einem Tangentenanstieg überein. Der Mittelwertsatz besagt also:

- Jeder Sekantenanstieg ist gleich dem Anstieg der Funktion an einer Zwischenstelle; oder mit anderen Worten:
- jede Sekante verläuft parallel zu einer Tangente an den Graphen an einer Zwischenstelle.

Beweis der Mittelwertsätze: a) Es sei $s = f(z)$ der kleinste und $S = f(Z)$ der größte Wert von f im Intervall $[a, b]$. Bei den Integralabschätzungen (siehe S. 76) hatten wir daraus bereits gefolgert

$$f(z) = s \leq \frac{1}{b-a} \int_a^b f \leq S = f(Z).$$

Da f stetig ist, muss nach dem Zwischenwertsatz (s. S. 25) auch der zwischen $f(z)$ und $f(Z)$ liegende Mittelwert $\mu(f)$ ein Funktionswert von f sein:

$$\mu(f) = \frac{1}{b-a} \int_a^b f = f(v) \quad \text{für ein geeignetes } v \in [a, b].$$

Dies ergibt den Mittelwertsatz der Integralrechnung.

b) Der zweite Mittelwertsatz wird mit dem Hauptsatz auf den ersten zurückgeführt. Die Funktion f ist eine Stammfunktion von f' , also folgt aus der Integralformel:

$$f(b) - f(a) = \int_a^b f'(x) dx$$

Damit ergibt sich nun unter Anwendung von a) auf f' :

$$\frac{f(b) - f(a)}{b-a} = \frac{1}{b-a} \cdot \int_a^b f'(x) dx = f'(w) \quad \text{für ein } w \in [a, b].$$

§16 Ergänzungen

In diesem Abschnitt wollen wir zwei weitere Integrationsregeln kennenlernen. Unser Ausgangspunkt ist die Tatsache, dass jede Ableitungsregel unmittelbar als eine Regel für Stammfunktionen verstanden werden kann.

a. Substitution. Oft entsteht eine Funktion aus einer einfacheren, indem man für die Funktionsvariable einen Term einsetzt (*substituiert*). So entsteht etwa e^{x^2+x} aus e^z durch die Substitution $z = x^2 + x$. Die Substitutionsregel gibt an, unter welchen Umständen und wie man

für solche komplizierteren Funktionen eine Stammfunktion finden kann. **Substitutionsregel:** Ist u eine differenzierbare Funktion und g eine stetige Funktion, so gilt:

$$\int_a^b g(u(x)) \cdot u'(x) dx = \int_{u(a)}^{u(b)} g(z) dz.$$

Beweis: Sei G eine Stammfunktion von g . Dann gilt gemäß der Kettenregel:

$$G(u(x))' = G'(u(x)) \cdot u'(x) = g(u(x)) \cdot u'(x).$$

Damit ist durch $G(u(x))$ also eine Stammfunktion für $g(u(x))u'(x)$ gegeben und mit der Integralformel folgt daraus:

$$\int_a^b g(u(x))u'(x)dx = \left[G(u(x)) \right]_a^b = G(u(b)) - G(u(a)) = \left[G(z) \right]_{u(a)}^{u(b)} = \int_{u(a)}^{u(b)} g(z) dz.$$

Dies ist aber genau die Behauptung.

Die *Anwendung* der Substitutionsregel ist leider nicht so einfach wie der obige Beweis. Geeignete Substitutionen $z = u(x)$ zu finden, bei denen ein kompliziertes Integral einfacher wird, ist eine Kunst und als solche Folge umfangreicher Übung und Erfahrung. Für uns soll zunächst das folgende Vorgehen ausreichen:

Will man die Substitutionsregel zur Berechnung eines Integrals $\int_a^b f$ verwenden, versucht man den Integranden $f(x)$ in der Form $f(x) = u'(x) \cdot g(u(x))$ darzustellen. Man sucht also in dem zu integrierenden Funktionsterm $f(x)$ einen Teilterm $u(x)$, dessen *Ableitung* $u'(x)$ ebenfalls darin auftritt. Durch den Vergleich $f(x) = g(u(x))u'(x)$ ermittle man g und für g dann eine Stammfunktion G .

Beispiel: $f(x) = xe^{x^2}$.

Mit $u(x) = x^2$, also $u'(x) = 2x$, nimmt dieser Funktionsterm die Form $f(x) = u'(x) \cdot \frac{1}{2}e^{u(x)}$ an. Durch Vergleich mit $f(x) = u'(x) \cdot g(u(x))$ kann man dann auch die Funktion g ablesen: $g(z) = \frac{1}{2}e^z$. Die Substitutionsregel liefert nun eine Stammfunktion für $u'(x)g(u(x))$, wenn eine Stammfunktion G für g bekannt ist. Durch die Substitutionsregel erfolgt eine Verlagerung des Problems von der Ausgangsfunktion $f(x) = xe^{x^2}$ auf die (einfachere) Funktion $g(z) = \frac{1}{2}e^z$, für die natürlich $G(z) = \frac{1}{2}e^z$ eine Stammfunktion ist. Damit ist $G(u(x)) = \frac{1}{2}e^{u(x)} = \frac{1}{2}e^{x^2}$ eine Stammfunktion der ursprünglich gegebenen Funktion f .

Tip: Wenn man mit Hilfe der Substitutionsregel eine Stammfunktion für einen Integranden gefunden hat, sollte man in jedem Falle dieses Ergebnis durch Ableiten überprüfen! Nach erfolgreicher Überprüfung ist die gesamte Herleitung unwichtig. Hier etwa: $\frac{1}{2}e^{x^2}$ hat als Ableitung gemäß der Kettenregel: $\frac{1}{2}e^{x^2} \cdot 2x = xe^{x^2}$.

Besonders wichtige Integraltypen, die mit der Substitutionsregel berechnet werden können, sind $\int \frac{u}{u'}$ und $\int uu'$. Konkrete Beispiele für diese Typen sind $\int \tan x dx = \int \frac{\sin x}{\cos x} dx$ und $\int \sin x \cdot \cos x dx$.

$$\ln(|u(x)|) \text{ ist eine Stammfunktion für } \frac{u'(x)}{u(x)}.$$

$$\frac{1}{2}u^2(x) \text{ ist eine Stammfunktion für } u(x) \cdot u'(x).$$

Zum *Nachweis* braucht man nur die angegebenen Stammfunktionen mit der Kettenregel abzuleiten.

Warnung: Wenn man die erste Regel zur Berechnung eines Integrals $\int_a^b \frac{u'(x)}{u(x)} dx$ verwenden will, beachte man, dass $\frac{u'}{u}$ über dem Integrationsintervall $[a, b]$ vollständig definiert sein muss,

insbesondere muss u dort überall differenzierbar sein und *es darf keine Nullstelle von u zwischen a und b liegen!* Erst wenn dies gesichert ist, gilt

$$\int_a^b \frac{u'(x)}{u(x)} = \ln(|u(b)|) - \ln(|u(a)|) = \ln \frac{|u(b)|}{|u(a)|} = \ln \left| \frac{u(b)}{u(a)} \right|.$$

Wenn aber u im Intervall $[a, b]$ keine Nullstelle hat, müssen $u(b)$ und $u(a)$ dasselbe Vorzeichen haben (u ist stetig) und folglich ist $\frac{u(b)}{u(a)}$ positiv, also $\left| \frac{u(b)}{u(a)} \right| = \frac{u(b)}{u(a)}$. Insgesamt folgt also:

$$u \text{ nullstellenfrei über } [a, b] \implies \int_a^b \frac{u'(x)}{u(x)} dx = \ln \frac{u(b)}{u(a)}.$$

b. Partielle Integration. Diese Regel ergibt sich aus der Produktregel für die Ableitung; statt partieller Integration spricht man daher auch von *Produktintegration*.

Partielle Integration: Sind u, v differenzierbare Funktionen über dem Intervall $[a, b]$, so gilt:

$$\int_a^b u'(x)v(x) dx = [u(x)v(x)]_a^b - \int_a^b u(x)v'(x) dx.$$

Diese Regel ergibt sich aus der Produktregel: Wegen $(u(x)v(x))' = u'(x)v(x) + u(x)v'(x)$ ist durch $u(x)v(x)$ eine Stammfunktion für $u'(x)v(x) + u(x)v'(x)$ gegeben, also

$$\int_a^b (u'(x)v(x) + u(x)v'(x)) dx = [u(x)v(x)]_a^b.$$

Mit der Summenregel für Integrale folgt hieraus die Behauptung.

Der Name *partielle* Integration erklärt sich daraus, dass diese Integrationsregel keine vollständige Berechnung des Integrals einer Produktfunktion zulässt, sondern nur eine partielle. Das Problem wird vom Integral der Funktion $u'v$ auf das Integral der Funktion uv' verlagert. Dennoch kann in vielen Fällen die partielle Integration erfolgreich zur vollständigen Berechnung von Integralen führen.

Die *Anwendung* der partiellen Integration ist nicht so problematisch wie bei der Substitution. Man muss den Integranden so in Faktoren zerlegen, dass man für einen Faktor (hier u' genannt) eine Stammfunktion (hier u) kennt. Man verlagert dann das Problem von der Berechnung von $\int u'v$ auf die Berechnung von $\int uv'$. Dieses Integral muss man dann weiter untersuchen. Zum Beispiel:

$$\begin{aligned} \int_a^b x \ln(x) dx &= \left[\frac{x^2}{2} \ln(x) \right]_a^b - \int_a^b \frac{x^2}{2} \cdot \frac{1}{x} dx = \left[\frac{x^2}{2} \ln(x) \right]_a^b - \int_a^b \frac{x}{2} dx \\ &= \left[\frac{x^2}{2} \ln(x) - \frac{x^2}{4} \right]_a^b. \end{aligned}$$

Hierbei war $u'(x) = x$ gewählt und $v(x) = \ln(x)$.

Auf die gleiche Weise lassen sich alle Integrale der Form $\int_a^b x^n \ln(x) dx$ mit Hilfe der partiellen Integration vollständig bestimmen: Man wähle jeweils $u'(x) = x^n$ und $v(x) = \ln(x)$. (Was ergibt sich für $n = 0$?)

c. Anwendung: Die Kreiszahl π . Neben der Eulerschen Zahl e ist die *Kreiszahl* π die zweite fundamentale transzendente Zahl. Diese haben wir definiert als das Bogenmaß des gestreckten Winkels, also als die Länge eines Halbkreisbogens vom Radius 1.

1. Geometrische Definition:

$$\text{Die reelle Zahl } \pi \text{ ist die Länge des Halbkreisbogens vom Radius 1.}$$

Daneben ist eine zweite Beschreibung der Zahl π möglich, die wir bereits bei der Berechnung von Rotationsvolumina benutzt haben:

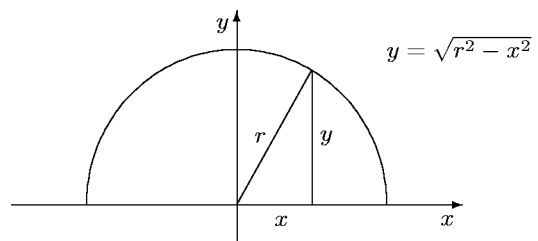
2. Geometrische Definition:

Die reelle Zahl π ist der Flächeninhalt des Kreises vom Radius 1.

Dies sind zwei völlig verschiedene Definitionen, und es ist auf den ersten Blick absolut nicht erkennbar, warum die Maßzahl des Flächeninhaltes des Einheitskreises exakt gleich der Maßzahl der Länge des Halbkreisbogens sein sollte. Aber dies ist doch der Fall, wie wir nun mit analytischen Mitteln (der Integralrechnung) zeigen wollen:

Beide geometrische Definitionen von π sind gleichwertig.

Dazu werden wir zunächst beide Definitionen in eine analytische Form bringen. Mit dem Satz des Pythagoras beschreiben wir den oberen Halbkreis als Funktionsgraphen: Für die Punkte (x, y) auf einem Kreis vom Radius r gilt $x^2 + y^2 = r^2$, und daher für $y \geq 0$ (oberer Halbkreis!) $y = \sqrt{r^2 - x^2}$. Damit ist der obere Halbkreis Graph der Funktion



$$f(x) = \sqrt{r^2 - x^2}.$$

Wir benötigen im folgenden $r = 1$.

Gemäß der ersten geometrischen Definition ist π die Bogenlänge des Graphen von f in den Grenzen von -1 bis 1 :

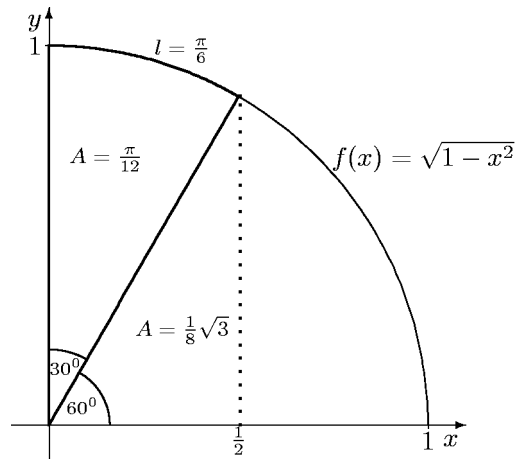
$$\pi = \int_{-1}^1 \sqrt{1 + f'(x)^2} dx = 2 \int_0^1 \sqrt{1 + f'(x)^2} dx.$$

Wir berechnen den Integranden:

$$\begin{aligned} f'(x) &= \frac{1}{2}(1 - x^2)^{-\frac{1}{2}} \cdot (-2x) = -\frac{x}{\sqrt{1 - x^2}}, \\ 1 + f'(x)^2 &= 1 + \frac{x^2}{1 - x^2} = \frac{1 - x^2 + x^2}{1 - x^2} = \frac{1}{1 - x^2}, \\ \sqrt{1 + f'(x)^2} &= \sqrt{\frac{1}{1 - x^2}} = \frac{1}{\sqrt{1 - x^2}}. \end{aligned}$$

Hier treten nun Probleme mit dem o. g. Integral auf: Der Integrand $\frac{1}{\sqrt{1-x^2}}$ ist an den Randstellen ± 1 *nicht definiert*, so dass das o. g. Integral im Sinne unserer Definition ebenfalls *nicht* existiert. (Solche Probleme sind der Grund für Erweiterungen des Integralbegriffs, die wir evtl. im 6. Semester thematisieren werden.)

Um eine auch mit unserem Integralbegriff sinnvolle analytische Beschreibung der Zahl π zu gewinnen, betrachten wir nur einen Kreissektor vom Winkel 30° . Die beiden geometrischen Definitionen besagen dann, dass dessen Bogenmaß $\pi/6$ und die Fläche $\frac{\pi}{12}$ ist.



Da $\cos(60^\circ) = \frac{1}{2}$ ist, können wir die Kreisbogenlänge $l = \frac{\pi}{6}$ als Integral in den Grenzen von 0 bis $\frac{1}{2}$ beschreiben:

1. analytische Definition:
$$\frac{\pi}{6} = \int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} dx$$

Auch die Fläche des Sektors kann durch ein Integral beschrieben werden, wenn man die Dreiecksfläche subtrahiert. Diese beträgt

$$A_{\text{Dreieck}} = \frac{1}{2} \cdot \frac{1}{2} \cdot f\left(\frac{1}{2}\right) = \frac{1}{4} \cdot \sqrt{\frac{3}{4}} = \frac{1}{8}\sqrt{3}.$$

Man erhält so die folgende

2. analytische Definition:
$$\frac{\pi}{12} = \int_0^{\frac{1}{2}} \sqrt{1-x^2} dx - \frac{1}{8}\sqrt{3}$$

Um zu zeigen, dass beide Definitionen von π gleichwertig sind, muss man folgende Gleichung nachweisen:

$$\int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} dx = 2 \int_0^{\frac{1}{2}} \sqrt{1-x^2} dx - \frac{1}{4}\sqrt{3}.$$

Beweis: Wir wenden die partielle Integration an auf

$$\int_0^{\frac{1}{2}} \sqrt{1-x^2} dx = \int_0^{\frac{1}{2}} 1 \cdot \sqrt{1-x^2} dx$$

mit $u'(x) = 1$, $v(x) = \sqrt{1-x^2}$. Wir erhalten so

$$\begin{aligned} \int_0^{\frac{1}{2}} \sqrt{1-x^2} dx &= \left[x \cdot \sqrt{1-x^2} \right]_0^{\frac{1}{2}} - \int_0^{\frac{1}{2}} x \cdot \frac{1}{2\sqrt{1-x^2}} \cdot (-2x) dx \\ &= \frac{1}{2}\sqrt{\frac{3}{4}} + \int_0^{\frac{1}{2}} \frac{x^2}{\sqrt{1-x^2}} dx \\ &= \frac{1}{2}\sqrt{\frac{3}{4}} + \int_0^{\frac{1}{2}} \frac{x^2-1}{\sqrt{1-x^2}} dx + \int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} dx \\ &= \frac{1}{4}\sqrt{3} - \int_0^{\frac{1}{2}} \sqrt{1-x^2} dx + \int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} dx. \end{aligned}$$

Im ersten Augenblick sieht es so aus, als hätten wir uns im Kreis gedreht: Ein Term auf der rechten Seite ist das Ausgangsintegral. Aber dies ist nur scheinbar eine Sackgasse: Man kann die entstandene Gleichung nach dem letzten Integral auflösen und erhält

$$2 \int_0^{\frac{1}{2}} \sqrt{1-x^2} dx - \frac{1}{4} \sqrt{3} = \int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} dx,$$

womit gerade die Behauptung bewiesen ist.

d. Flächeninhalt und Bogenlänge von Kreisen. Wir leiten nun mit Hilfe der Integralrechnung die bekannten Flächen- und Umfangsformeln für Kreise her. Da Fläche und Umfang des Einheitskreises gemäß c. als π bzw. 2π bekannt sind, genügt es, das folgende Verhalten von Fläche und Umfang bei Streckung (oder Stauchung) zu beweisen.

Streckung von Kreisen:

Die Fläche eines Kreises vom Radius r ist das r^2 -fache der Fläche des Einheitskreises, also $A = \pi r^2$.

Der Umfang eines Kreises vom Radius r ist das r -fache des Umfangs des Einheitskreises, also $U = 2\pi r$.

Dies bedeutet: Streckt (oder staucht) man einen Kreis um den Faktor r , so multipliziert sich der Umfang mit dem Streckungsfaktor r , während sich die Fläche mit dem *Quadrat* des Streckungsfaktors ändert.

Beweis: Wie in c) stellen wir die Fläche A_r eines Viertelkreises vom Radius r als Integral dar:

$$A_r = \int_0^r \sqrt{r^2 - x^2} dx = \int_0^r \sqrt{r^2 \left(1 - \frac{x^2}{r^2}\right)} dx = r \cdot \int_0^r \sqrt{1 - \left(\frac{x}{r}\right)^2} dx.$$

Auf das letzte Integral wenden wir lineare Substitution an, und zwar mit $u(x) = \frac{x}{r}$ und $g(z) = \sqrt{1 - z^2}$. Es ist $u'(x) = \frac{1}{r}$ konstant. Daher läßt sich die Substitutionsregel problemlos anwenden:

$$\begin{aligned} \int_0^r \sqrt{1 - \left(\frac{x}{r}\right)^2} dx &= \int_0^r g(u(x)) dx = r \int_0^r g(u(x)) u'(x) dx \\ &= r \int_{u(0)}^{u(r)} g(z) dz = r \int_0^1 \sqrt{1 - z^2} dz = r A_1. \end{aligned}$$

Insgesamt ergeben beide Formeln $A_r = r^2 \cdot A_1$, wie behauptet. Zusammen mit der Definition von π als Fläche des Einheitskreises erhalten wir die bekannte Flächeninhaltsformel.

Mit denselben Rechnungen wie in Abschnitt c. (führen Sie sie zur Übung durch!) beschreiben wir die Länge eines 30° -Kreisbogens vom Radius r als

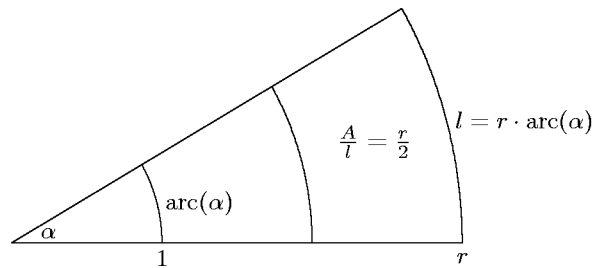
$$l_r = \int_0^{\frac{\pi}{2}} \frac{r}{\sqrt{r^2 - x^2}} dx.$$

‘Kürzt’ man den Integranden mit r und wendet dann wieder lineare Substitution an ($u(x) = \frac{x}{r}$), so erhält man

$$l_r = \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{1 - \left(\frac{x}{r}\right)^2}} dx = r \cdot \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{1 - x^2}} dx = r \cdot l_1.$$

Die allgemeine Flächen- und Umfangsformel zeigt, dass bei einem Kreis das Verhältnis von Fläche zum Umfang gleich $\frac{r}{2}$ ist. Dies überträgt sich dann sinngemäß auf *Kreis-sektoren*: Auch

hier ist das Verhältnis von Fläche zur Länge der Kreisbögen gleich $\frac{r}{2}$. Damit kann man die Fläche



von Kreissektoren, die man auch als *Kreisbogendreiecke* bezeichnet, wie folgt aus der Bogenlänge l (und damit aus dem Bogenmaß $\text{arc}(\alpha)$ des Zentralwinkels) berechnen:

$$A = \frac{r}{2} \cdot l = \frac{1}{2} r^2 \cdot \text{arc}(\alpha)$$

Die erste Form dieser Flächenformel besagt geometrisch gesprochen: Die Fläche eines Kreisbogendreiecks ist das Produkt aus ‘Grundlinien-’ und Bogenlänge geteilt durch 2. In dieser Form erinnert sie sehr an die übliche Flächenformel für Dreiecke, da der Kreisbogen senkrecht zur ‘Grundlinie’ steht, also die ‘Höhe’ im Kreisbogendreieck ist.