

# Differenzialquotient

10. August 2018

## Inhaltsverzeichnis

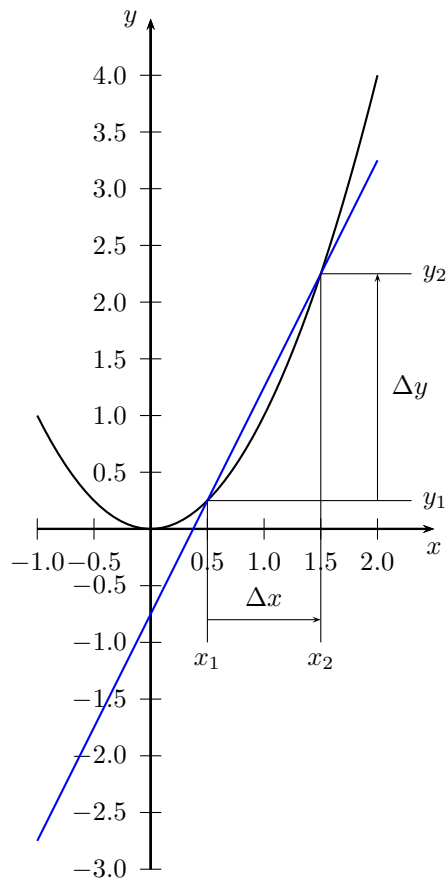
<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Mechanik . . . . .	3
1.2	Thermodynamik . . . . .	3
1.3	Elektrotechnik . . . . .	4
<b>2</b>	<b>Grenzwerte</b>	<b>6</b>
2.1	Differenzialquotient als Grenzwert von Differenzenquotienten . . . . .	6
2.2	Allgemeinerer Grenzwertbegriff . . . . .	8
2.3	Grenzwertsätze . . . . .	9
2.4	Stetigkeit . . . . .	10
<b>3</b>	<b>Polynome</b>	<b>11</b>
3.1	Monome und Polynome . . . . .	11
3.2	Polynominterpolation . . . . .	11
<b>4</b>	<b>Differenzialquotienten von Polynomen</b>	<b>13</b>
4.1	Differenzialquotienten von linearen Funktionen und Monomen . . . . .	13
4.2	Linearität von Differenzen- und Differenzialquotienten . . . . .	14
4.3	Differenzialquotienten von Polynomen . . . . .	15
<b>5</b>	<b>Numerische Lösung von Differenzialgleichungen</b>	<b>16</b>
<b>A</b>	<b>Verallgemeinerte Binomische Formel</b>	<b>19</b>

## 1 Einleitung

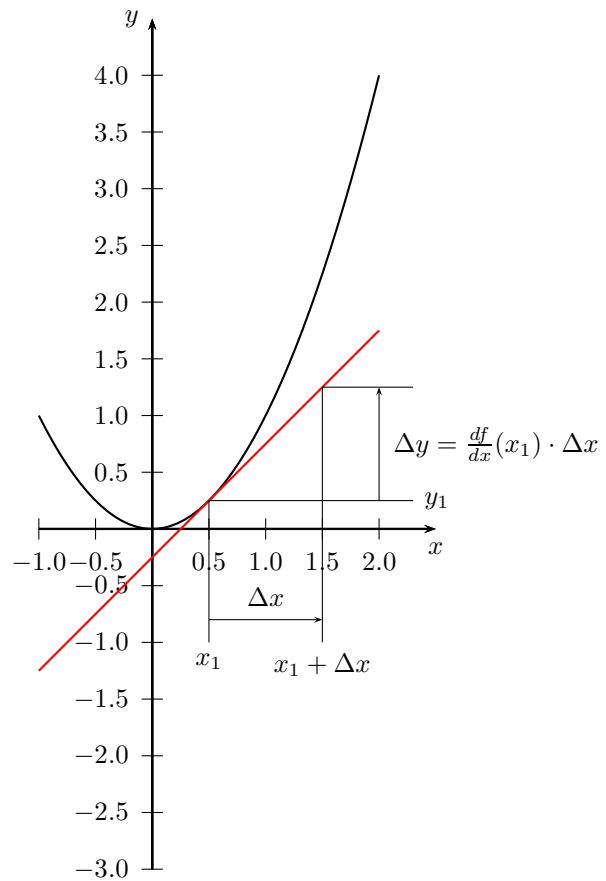
Die *Sekante*, die eine Kurve  $f(x)$  an zwei vorgegebenen Stellen  $x_1, x_2$  schneidet, ist die Gerade

$$g(x) := f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} \cdot (x - x_1). \quad (1)$$

Als *Tangente* an der Stelle  $x_1$  wird diejenige Gerade bezeichnet, die sich als Grenzgerade bei Annäherung von  $x_2$  an  $x_1$  ergibt. Im Rahmen der Einleitung wollen wir uns mit dieser etwas vagen Andeutung und dem folgenden Beispiel zufrieden geben. In Abschnitt 2 gehen wir dann näher darauf ein, wie die Annäherung von  $x_2$  an  $x_1$  gemeint ist.



Differenzenquotient



Differenzialquotient

Abbildung 1: Differenzenquotient als Anstieg der Sekanten und Differenzialquotient als Anstieg der Tangenten

In Abbildung 1 ist zweimal die quadratische Funktion

$$y = f(x) := x^2$$

dargestellt. Die Sekante, die den Graph von  $f$  an den Stellen  $x_1 = 0.5$  und  $x_2 = 1.5$  schneidet, ist blau eingezeichnet und rot die Tangente an der Stelle  $x_1$ .

Der Anstieg der Sekante ergibt sich aus dem Differenzenquotient

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Lässt man die Differenz  $\Delta x = x_2 - x_1$  betragsmäßig immer kleiner werden, nähert sich der Differenzenquotient immer mehr dem Anstieg der Tangenten, der als *Differenzialquotient*  $\frac{df}{dx}(x_1)$  an der Stelle  $x_1$  bezeichnet wird.

Gleichungen mit Differenzialquotienten heißen *Differenzialgleichungen*.

Interpretiert man die x-Achse als Zeit und die y-Achse als Weg, so ist der Differenzenquotient  $\frac{y_2 - y_1}{x_2 - x_1}$  die Durchschnittsgeschwindigkeit im Zeitintervall von  $x_1$  bis  $x_2$  und der Differenzialquotient  $\frac{df}{dx}(x_1)$  ist die Momentangeschwindigkeit zum Zeitpunkt  $x_1$ .

Fast alle physikalischen Erscheinungen lassen sich durch Differenzialgleichungen beschreiben. In den folgenden Unterabschnitten wollen wir uns einige wenige Beispiele anschauen.

## 1.1 Mechanik

Mit der Federkonstante  $k$  und der entspannten Länge  $l$  einer Feder ergibt sich die Federkraft aus der Gleichung

$$F(x) = k(l - x).$$

Die Beschleunigung  $a(t)$  eines Körpers mit der Masse  $m$  zur Zeit  $t$  ergibt sich mit der auf den Körper momentan einwirkenden Kraft  $F(t)$  aus der Gleichung

$$m \cdot a(t) = F(t).$$

Die Beschleunigung ist dabei als momentane Änderung der Geschwindigkeit  $v(t)$ , also als Differenzialquotient

$$a(t) = \frac{dv}{dt}(t)$$

zu interpretieren und die Geschwindigkeit ist die momentane Änderung

$$v(t) = \frac{dx}{dt}(t)$$

des Ortes  $x(t)$ . So ergibt sich für die Kombination des Körpers mit der Feder das Differenzialgleichungssystem

$$\begin{aligned} m \cdot \frac{dv}{dt}(t) &= k(l - x(t)), \\ \frac{dx}{dt}(t) &= v(t). \end{aligned}$$

## 1.2 Thermodynamik

Wir betrachten eine sehr lange Metallstange die wir mit einer Ortsvariable  $x$  koordinatisieren. Um das Beispiel einfach zu halten interessieren wir uns erst einmal nicht für die Enden der Stange und nehmen an, dass  $x$  von  $-\infty$  bis  $+\infty$  läuft.

Die Stange habe eine von der Koordinate  $x$  unabhängige Querschnittsfläche  $A$ , eine Dichte  $\rho$  und eine Wärmekapazität  $c$ .

Wir nehmen an, dass zum Anfangszeitpunkt  $t = 0$  der Temperaturverlauf  $T_0(x)$  gegeben ist und interessieren uns zu jedem Zeitpunkt  $t > 0$  für den Temperaturverlauf  $T(x, t)$  an allen Stellen  $x$  der Stange.

Der *Wärmestrom*  $\dot{q}(x, t)$  ist die Wärme, die pro Zeiteinheit durch die Querschnittsfläche  $A$  an der Stelle  $x$  in positiver  $x$ -Richtung strömt.

Die Wärme fließt immer von Punkten höherer Temperatur zu Punkten niedrigerer Temperatur. Je höher die Temperaturdifferenz ist, desto größer ist der Wärmestrom.

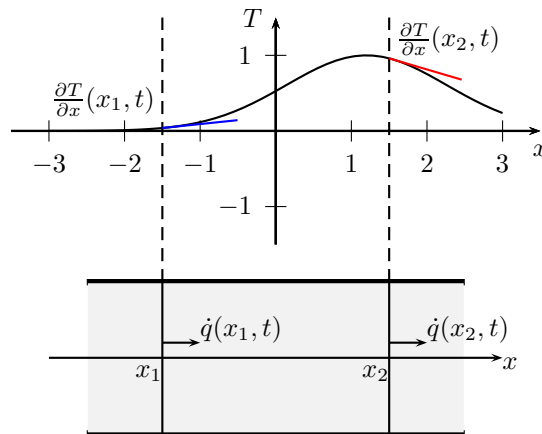


Abbildung 2: Wärmeleitung in einer unendlich langen Metallstange; Unten: Ausschnitt der Metallstange; Oben: Temperaturverlauf entlang des Ausschnitts zusammen mit Tangentenausschnitten für den Temperaturanstieg  $\frac{\partial T}{\partial x}(x, t)$  an den Stellen  $x = x_1$  und  $x = x_2$

Physikalische Messungen zeigen, dass der Wärmestrom direkt proportional zum Temperaturgefälle  $-\frac{\partial T}{\partial x}(x, t)$  ist. Der Differenzialquotient mit dem geschwungenen  $\partial$  heißt *partielle Ableitung* und bedeutet, dass alle Größen, nach denen nicht differenziert wird, während des Differenzierens konstant zu halten sind. Die Proportionalitätskonstante  $k$  ist nur vom Material und von der Querschnittsfläche abhängig. Wir erhalten also für den Wärmefluss die Gleichung

$$\dot{q}(x, t) = -k \cdot \frac{\partial T}{\partial x}(x, t). \quad (2)$$

Betrachten wir ein kurzes Stück der Stange, das von  $x_1$  bis  $x_2$  geht (mit  $x_1 < x_2$ , siehe Abbildung 2). Die Länge dieses Stücks ist  $\Delta x := x_2 - x_1$ .

Für eine Erhöhung der Temperatur  $\bar{T}(x_1, x_2, t)$  des Stückchens um  $\Delta \bar{T}$  wird die Wärmemenge

$$\Delta Q = \underbrace{A \cdot \Delta x}_{\text{Volumen}} \cdot \rho \cdot c \cdot \Delta \bar{T}(x_1, x_2, t)$$

benötigt. Wie in Abbildung 2 ersichtlich ist, fließt der Wärmestrom  $\dot{q}(x_1, t)$  in den Abschnitt  $[x_1, x_2]$  hinein, während  $\dot{q}(x_2, t)$  aus dem Abschnitt herausfließt. Für eine über das Stückchen gemittelte momentane Temperaturänderung  $\frac{\partial \bar{T}}{\partial t}(x_1, x_2, t)$  ist ein Wärmestrom

$$\dot{q}(x_1, t) - \dot{q}(x_2, t) = A \cdot \Delta x \cdot \rho \cdot c \cdot \frac{\partial \bar{T}}{\partial t}(x_1, x_2, t)$$

erforderlich. Division durch  $\Delta x$  und Berücksichtigung von  $x_2 = x_1 + \Delta x$  liefert

$$-\frac{\dot{q}(x_1 + \Delta x, t) - \dot{q}(x_1, t)}{\Delta x} = A \cdot \rho \cdot c \cdot \frac{\partial \bar{T}}{\partial t}(x_1, x_2, t).$$

Auf der linken Seite taucht der Differenzenquotient des Wärmeflusses bzgl.  $x$  auf. Lassen wir noch  $\Delta x$  gegen null streben, so erhalten wir

$$-\frac{\partial \dot{q}}{\partial x}(x_1, t) = A \cdot \rho \cdot c \cdot \frac{\partial T}{\partial t}(x_1, t). \quad (3)$$

Dabei haben wir berücksichtigt, dass die über das Intervall  $[x_1, x_2]$  gemittelte Temperatur  $\bar{T}(x_1, x_2, t)$  gegen die Temperatur  $T(x_1, t)$  an der Stelle  $x_1$  strebt, wenn  $x_2$  gegen  $x_1$  strebt.

Da wir in der Formel nicht mehr zwischen den zwei Koordinaten  $x_1, x_2$  unterscheiden müssen ( $x_2$  strebt ja gegen  $x_1$ ), können wir auch einfach wieder  $x$  schreiben und die zwei Gleichungen (2) und (3) zum folgenden System zusammenfassen:

$$\begin{aligned} \dot{q}(x, t) &= -k \frac{\partial T}{\partial x}(x, t), \\ -\frac{\partial \dot{q}}{\partial x}(x, t) &= A \cdot \rho \cdot c \cdot \frac{\partial T}{\partial t}(x, t). \end{aligned}$$

Dieses (partielle) Differenzialgleichungssystem beschreibt die Wärmeleitung in der Metallstange. Um dir noch einen Anfassunger für eine eventuelle Literaturrecherche zu geben sei hier ohne weitere Erläuterung erwähnt, dass die zwei partiellen Differenzialgleichungen oft zur *Wärmeleitungsgleichung*  $a \frac{\partial^2 T}{\partial x^2}(x, t) = \frac{\partial T}{\partial t}(x, t)$  mit  $a = \frac{k}{\rho c}$  zusammengefasst werden.

### 1.3 Elektrotechnik

Die Spannung  $U$  über einem Kondensator ist proportional zu der auf ihm gespeicherten Ladung  $Q$ . Der Proportionalitätsfaktor ist die Kapazität  $C$  und die Gleichung für die Spannung am Kondensator lautet

$$C \cdot U = Q \quad (4)$$

Die momentane Erhöhung der Ladung  $\frac{dQ}{dt}(t)$  auf dem Kondensator ist gleich dem Ladungsträgerzufluss, also dem Strom  $I(t)$ , der am positiven Anschluss in den Kondensator hineinfließt:

$$I(t) = \frac{dQ}{dt}(t) \quad (5)$$

In Abschnitt 4.2 sehen wir, dass konstante Linearfaktoren bei der Differenziation einfach herausgezogen werden können. So wird aus (4) die Gleichung

$$C \cdot \frac{dU}{dt}(t) = \frac{dQ}{dt}(t)$$

und mit (5)

$$C \cdot \frac{dU}{dt}(t) = I(t). \quad (6)$$

Betrachten wir nun den Stromkreis in Abbildung 3 mit einer Spannungsquelle  $B$  mit Quellspannung  $U_B$ , z.B. einer Batterie, einem Widerstand mit Widerstandswert  $R$  und einem Kondensator mit Kapazität  $C$ .

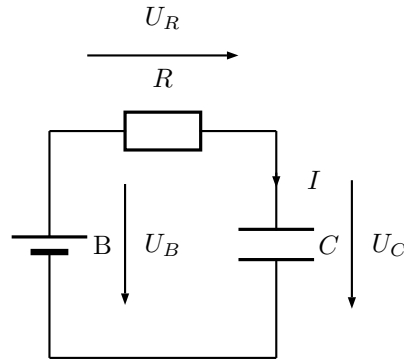


Abbildung 3: Stromkreis aus Spannungsquelle, Widerstand und Kondensator

Das Kirchhoffsche Stromgesetz gibt vor, dass durch alle drei Bauelemente der gleiche Strom, nämlich  $I$  fließt. Aus dem Kirchhoffschen Spannungsgesetz folgt die Maschengleichung

$$0 = -U_B + U_R(t) + U_C(t)$$

Mit dem Ohmschen Gesetz  $U_R = R \cdot I$  für den Widerstand erhält man daraus die Gleichung

$$0 = -U_B + R \cdot I(t) + U_C(t)$$

$$I(t) = \frac{U_B - U_C(t)}{R}$$

Einsetzen dieser Gleichung in die Strom-Spannungsrelation (6) für den Kondensator liefert

$$C \frac{dU_C}{dt}(t) = \frac{U_B - U_C(t)}{R}$$

$$\tau \frac{dU_C}{dt}(t) = U_B - U_C(t) \quad (7)$$

mit der *Zeitkonstante*  $\tau := RC$ . Die letzte Gleichung ist eine Differenzialgleichung für  $U_C(t)$ . Bei ihr hängt der Anstieg  $\frac{dU_C}{dt}(t)$  vom Momentanwert  $U_C(t)$  ab. Je weiter sich  $U_C(t)$  dem Endwert  $U_B$  nähert, desto kleiner wird der Anstieg (siehe Abbildung 4).

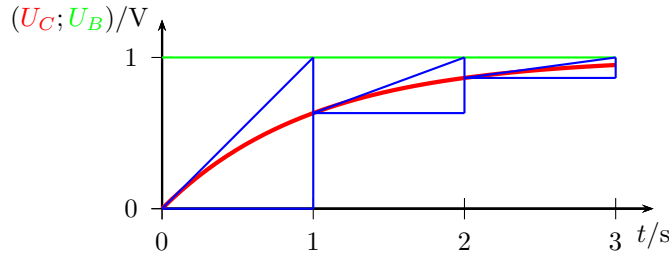


Abbildung 4: Zeitverlauf von  $U_C(t)$  (bei  $\tau = 1\text{s}$  und  $U_B = 1\text{V}$ ); In jedem Punkt von  $U_C$  ist der Anstieg der Kurve so groß wie der Abstand des Momentanwertes vom Endwert  $U_B$ . Zur Verdeutlichung sind drei Anstiegsdreiecke blau eingezeichnet.

## 2 Grenzwerte

### 2.1 Differenzialquotient als Grenzwert von Differenzenquotienten

Aus (1) kennen wir bereits die Gleichung für die Sekante, die an den Stellen  $x_1$  und  $x_2$  eine Kurve  $f(x)$  schneidet. Der Anstieg der Sekante ist

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad (8)$$

Außerdem wurde in Abschnitt 1 erwähnt, dass die Tangente die Grenzgerade bei Annäherung von  $x_2$  an  $x_1$  ist.

Den Anstieg  $\frac{df}{dx}(x)$  der Tangente an der Stelle  $x_1$  können wir jedoch nicht direkt aus Gleichung (8) berechnen, da für  $x_2 = x_1$  der Nenner null wird.

Finden wir stattdessen eine Zahl  $\frac{df}{dx}(x_1)$ , der sich der Sekantenanstieg (8) annähert, wenn  $x_2$  an  $x_1$  heranrückt? Wie ist dabei „annähern“ zu verstehen?

Wir können den Differenzenquotienten (8) nur an Stellen  $x_2$  ungleich  $x_1$  auswerten und müssen deshalb bei jeder konkreten Auswertung von  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  eine absolute Abweichung  $\left| \frac{df}{dx}(x_1) - \frac{\Delta f}{\Delta x}(x_1, x_2) \right|$  größer Null zulassen.

**Definition 1** Wir definieren eine Zahl  $\frac{df}{dx}(x_1)$  als Grenzwert von  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  bei Annäherung von  $x_2$  an  $x_1$ , falls für jede (beliebig knapp über Null liegende) Fehlertoleranz  $\varepsilon > 0$  die Abweichung  $\left| \frac{df}{dx}(x_1) - \frac{\Delta f}{\Delta x}(x_1, x_2) \right|$  nicht größer als  $\varepsilon$  wird, wenn wir uns auf Stellen  $x_2$  ungleich  $x_1$  beschränken, die einen (von  $\varepsilon$  abhängigen) Maximalabstand  $\delta > 0$  von  $x_1$  haben.

Als Beispiel ermitteln wir die Ableitung von  $f(x) = x^2$  an der Stelle  $x = x_1 := 0.5$ .

Für  $x_2 \neq x_1$  gilt

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \frac{x_2^2 - x_1^2}{x_2 - x_1} = \frac{(x_2 - x_1)(x_2 + x_1)}{x_2 - x_1} = x_2 + x_1 \quad (9)$$

Im Zähler kam die binomische Formel  $x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1)$  zum Einsatz. Bei der Rechnung hebt sich danach der Nenner  $x_2 - x_1$  heraus. Die rechte Seite ist frei von Divisionen, problemlos an der Stelle  $x_2 = x_1$  auswertbar und man erhält für sie an dieser Stelle  $2x_1$ .

Wir vermuten also, dass

$$\frac{df}{dx}(x_1) = 2x_1 \quad (10)$$

die Ableitung von  $f(x) = x^2$  an der Stelle  $x = x_1$  ist.

Wir prüfen ob  $\frac{df}{dx}(x_1) = 2x_1$  der obigen Definition für den Grenzwert von  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  bei Annäherung von  $x_2$  an  $x_1$  standhält.

Wir geben uns eine beliebig knapp über Null liegende Fehlertoleranz  $\varepsilon > 0$  vor und schauen, ob wir einen Maximalabstand  $\delta > 0$  finden, so dass für alle  $x_2$ , die keinen größeren Abstand von  $x_1$  haben, die also  $|x_2 - x_1| \leq \delta$  erfüllen, die Relation

$$\left| \frac{df}{dx}(x_1) - \frac{\Delta f}{\Delta x}(x_1, x_2) \right| \leq \varepsilon$$

erfüllt ist. Einsetzen von (9) und (10) liefert die Ungleichung

$$\left| \underbrace{\frac{df}{dx}(x_1)}_{\frac{df}{dx}(x_1)} - \underbrace{\frac{\Delta f}{\Delta x}(x_1, x_2)}_{\frac{\Delta f}{\Delta x}(x_1, x_2)} \right| \leq \varepsilon,$$

die nach Vereinfachung in die folgende Relation übergeht:

$$|x_1 - x_2| \leq \varepsilon$$

Wie wir an dieser Ungleichung sehen, können wir im Beispiel  $f(x) = x^2$  einfach den Maximalabstand  $\delta = \varepsilon$  nutzen. Somit ist die Voraussetzung aus obiger Definition erfüllt und  $\frac{df}{dx}(x_1) = 2x_1$  ist der Grenzwert von  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  bei Annäherung von  $x_2$  an  $x_1$ .

In Abbildung 5 sind zur Veranschaulichung bei  $f(x) = x^2$  die Grenzsekanten für die Fehlertoleranz  $\varepsilon = 0.5$  eingetragen. Bei dieser Fehlertoleranz ergibt sich ein möglicher Maximalabstand  $|x_2 - x_1| = 0.5$  und  $x_2$  kann im Intervall von  $x_2 = 0$  bis  $x_2 = 1$  gewählt werden.

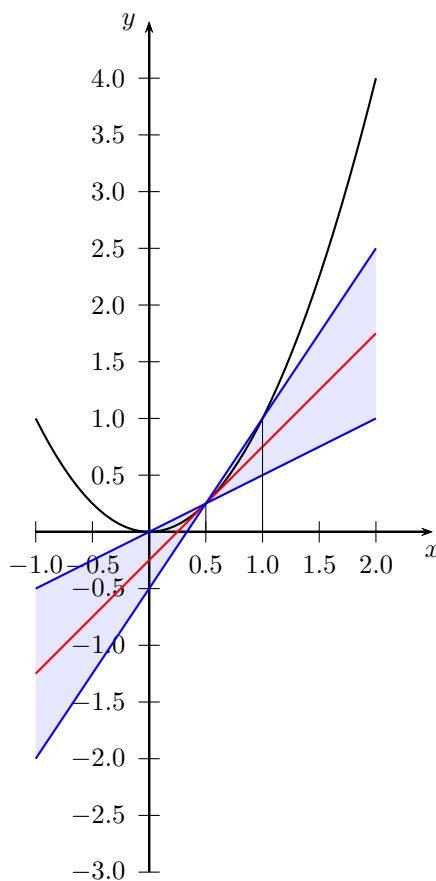


Abbildung 5: Rot: Tangente für  $f(x_1) = x_1^2$  im Punkt  $x_1 = 0.5$ ; Blau: Bereich der Sekanten mit einer Fehlertoleranz von  $\varepsilon = 0.5$  für die Abweichung des Differenzenquotienten  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  von der Tangente  $\frac{df}{dx}(x_1)$ ; Der Maximalabstand  $|x_2 - x_1| = 0.5$  erlaubt Sekanten mit  $x_2$  im Intervall von  $x_2 = 0$  bis  $x_2 = 1$ .

## 2.2 Allgemeinerer Grenzwertbegriff

Der Grenzwertbegriff ist nicht nur auf Differenzialquotienten anwendbar, sondern auf beliebige reelle Funktionen  $f(x)$ .

**Definition 2** Eine reelle Funktion  $f(x)$  mit einem reellen Argument  $x$  strebt bei Annäherung von  $x$  an eine Stelle  $X$  gegen einen Grenzwert  $F$ , falls es für jede positive Fehlerschranke  $\varepsilon > 0$  einen Maximalabstand  $\delta > 0$  gibt, so dass für alle von  $X$  verschiedene Argumente  $x$ , die nicht weiter als  $\delta$  von  $X$  entfernt sind, die Abweichung des Funktionswertes  $f(x)$  vom Grenzwert  $F$  nicht größer als die vorgegebene Fehlerschranke  $\varepsilon$  ist, d.h.,  $|f(x) - F| \leq \varepsilon$  gilt.

Als erstes Beispiel schauen wir uns den Grenzwert von  $f(x) := x \cdot \sin\left(\frac{1}{x}\right)$  bei Annäherung von  $x$  an 0 an. Aufgrund der Division durch  $x$  ist  $f(x)$  nicht an der Stelle  $x = 0$  auswertbar.

Jedoch wird der Betrag von  $\sin\left(\frac{1}{x}\right)$  nicht größer als 1. Damit ist der Betrag von  $f(x) = x \cdot \sin\left(\frac{1}{x}\right)$  durch  $|x|$  beschränkt. Da  $|x|$  bei Annäherung von  $x$  an 0 den Grenzwert 0 hat, erwarten wir, dass das bei  $|f(x)|$  auch der Fall ist. Geben wir uns eine beliebig knapp über 0 liegende Fehlerschranke  $\varepsilon$  vor und untersuchen, für welchen Maximalabstand  $\delta$  wir absichern können, dass die Abweichung  $|f(x)|$  nicht größer als  $\varepsilon$  wird:

$$\begin{aligned} |f(x) - F| &= \left| x \cdot \sin\left(\frac{1}{x}\right) - 0 \right| \\ &\leq |x| \left| \sin\left(\frac{1}{x}\right) \right| \\ &\leq |x| \quad (\text{unter Berücksichtigung von } \left| \sin\left(\frac{1}{x}\right) \right| \leq 1) \end{aligned}$$

Wir können also die maximale Abweichung  $\delta$  für  $|x|$  gleich der vorgegebenen Fehlerschranke  $\varepsilon$  wählen, damit diese nicht von  $|x \cdot \sin\left(\frac{1}{x}\right)|$  überschritten wird. Im linken Teil der Abbildung 6 ist als Beispiel die Fehlerschranke  $\varepsilon = 0.5$  vorgegeben und gezeigt, dass diese Fehlerschranke für  $x$ -Werte im Intervall von  $-0.5$  bis  $0.5$  nicht überschritten wird.

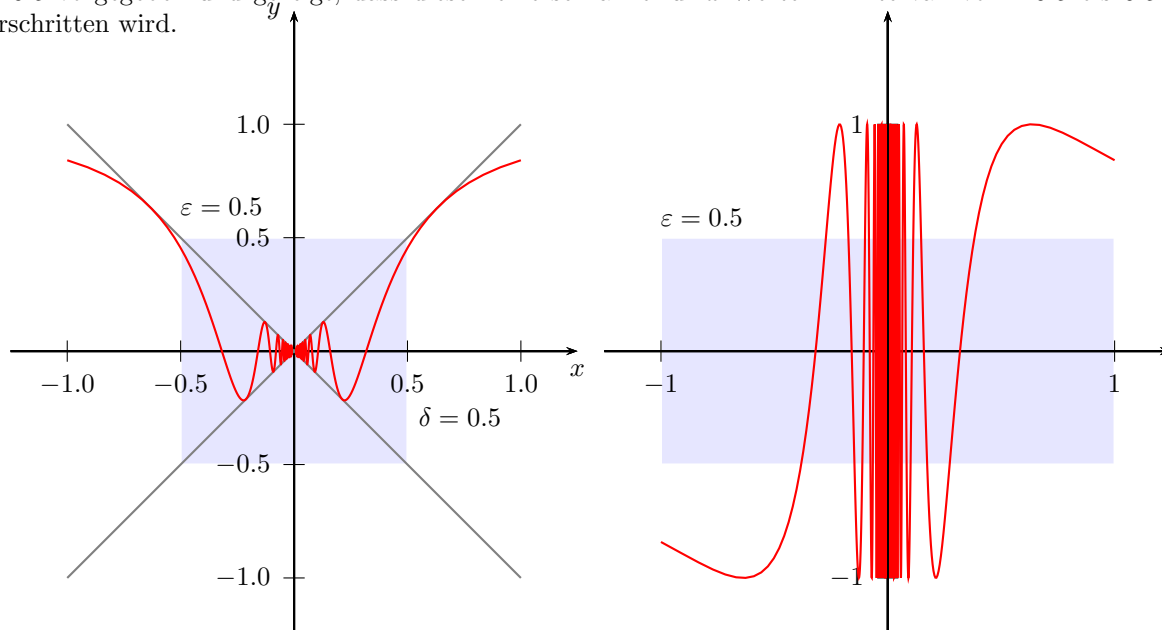


Abbildung 6: Links: Graph von  $f(x) = x \cdot \sin\left(\frac{1}{x}\right)$ ; Rechts: Graph von  $f(x) = \sin\left(\frac{1}{x}\right)$

Als zweites Beispiel schauen wir uns die Funktion  $f(x) = \sin\left(\frac{1}{x}\right)$  an, die ebenfalls bei 0 nicht auswertbar ist.

Egal wie knapp über Null wir die maximale Abweichung  $\delta$  wählen, in dem Intervall von 0 bis  $\delta$  liegen immer noch unendlich viele Minimalstellen  $\tilde{x}$  mit  $\sin(1/\tilde{x}) = -1$  und unendlich viele Maximalstellen  $\hat{x}$  mit  $\sin(1/\hat{x}) = 1$ . Zur Konstruktion jeweils einer dieser Stellen wählt einfach eine hinreichend große natürliche Zahl  $n$  mit  $\frac{1}{2\pi n} \leq \delta$  und nutzt  $\tilde{x} := \frac{1}{2\pi n + 3\pi/4}$  beziehungsweise  $\hat{x} := \frac{1}{2\pi n + \pi/4}$ . Alle größeren natürlichen Zahlen  $m > n$



liefern mit der selben Konstruktion weitere Minimalstellen und Maximalstellen für  $\sin(1/x)$  mit Abstand von Null, der kleiner als der vorgegebene Abstand  $\delta$  ist.

Jeder Versuch einen Grenzwert  $F$  für  $f$  bei Annäherung von  $x$  an 0 zu konstruieren scheitert, denn für diese Zahl  $F$  wäre der Fehler entweder bei einem lokalen Minimum oder bei einem lokalen Maximum mindestens 0.5 groß.

## 2.3 Grenzwertsätze

**Produkt** Konvergieren zwei Funktionen  $f(x)$  und  $g(x)$  bei Annäherung von  $x$  an eine Stelle  $X$  gegen Grenzwerte  $F$  beziehungsweise  $G$ , so konvergiert das Produkt  $f(x) \cdot g(x)$  gegen  $F \cdot G$ .

Wir geben uns zunächst eine Fehlerschranke  $\bar{\varepsilon} > 0$  für  $f$  und  $g$  als Variable vor. In Abhängigkeit von  $\bar{\varepsilon}$  finden wir eine Formel für eine Schranke  $\varepsilon$  der Abweichung des Produkts  $f(x)g(x)$  vom Grenzwertprodukt  $FG$ . Wir stellen diese Formel dann nach  $\bar{\varepsilon}$  um, so dass wir  $\varepsilon$  vorgeben können, daraus  $\bar{\varepsilon}$  berechnen können und mit  $\bar{\varepsilon}$  dann den maximalen Abstand  $\delta$  ermitteln.

Um die Voraussetzungen  $|f(x) - F| \leq \bar{\varepsilon}$  und  $|g(x) - G| \leq \bar{\varepsilon}$  anwenden zu können, addieren wir in  $|f(x)g(x) - FG|$  den Term  $f(x)G$  und ziehen ihn gleich wieder ab. Das ändert nichts am Wert unter dem Betragszeichen.

$$|f(x)g(x) - FG| = |f(x)g(x) - f(x)G + f(x)G - FG|$$

Durch Ausklammern von  $f(x)$  und  $G$  gewinnen wir die Terme  $f(x) - F$  und  $g(x) - G$ :

$$= |f(x) \cdot (g(x) - G) + (f(x) - F) \cdot G|$$

Jetzt verwenden wir die für Summen allgemeingültige Abschätzung  $|a+b| \leq |a|+|b|$  um den Betrag der Summe in eine Summe von Beträgen aufzusplitten.

$$\leq |f(x) \cdot (g(x) - G)| + |(f(x) - F) \cdot G|$$

Bei Beträgen von Produkten ist es egal, ob man zuerst multipliziert oder zuerst die Beträge bildet, d.h., es gilt allgemein  $|a \cdot b| = |a| \cdot |b|$ , was wir auch zur Umformung der letzten Formel ausnutzen.

$$= |f(x)| \cdot |g(x) - G| + |f(x) - F| \cdot |G|$$

Bis auf  $|f(x)|$  haben wir durch den bekannten Wert  $G$  und die vorgegebenen Fehlertoleranzen für  $|g(x) - G|$  und  $|f(x) - F|$  für alle Beträge in der letzten Formel Abschätzungen. Um auch  $|f(x)|$  abschätzen zu können, fügen wir unter dem Betragszeichen  $-F + F$  hinzu. In dem entstehenden Term  $|f(x) - F + F|$  können wir  $|f(x) - F|$  und  $|F|$  abschätzen.

$$\begin{aligned} &= |f(x) - F + F| \cdot |g(x) - G| + |f(x) - F| \cdot |G| \\ &= (|f(x) - F| + |F|) \cdot |g(x) - G| + |f(x) - F| \cdot |G| \\ &\leq (\bar{\varepsilon} + |F|) \cdot \bar{\varepsilon} + \bar{\varepsilon}|G| =: \varepsilon. \end{aligned}$$

Letztendlich erhalten wir eine Gleichung mit der wir aus der Fehlertoleranz  $\bar{\varepsilon}$  für  $f$  und  $g$  eine Fehlertoleranz  $\varepsilon$  für das Produkt  $f \cdot g$  berechnen können.

Wir können auch  $\varepsilon$  vorgeben und das nötige  $\bar{\varepsilon}$  daraus berechnen, wenn wir die Gleichung

$$(\bar{\varepsilon} + |F|) \cdot \bar{\varepsilon} + \bar{\varepsilon}|G| = \varepsilon$$

nach  $\bar{\varepsilon}$  auflösen:

$$\begin{aligned} \bar{\varepsilon}^2 + (|F| + |G|)\bar{\varepsilon} - \varepsilon &= 0 \\ \bar{\varepsilon} &= -\frac{|F| + |G|}{2} + \sqrt{\left(\frac{|F| + |G|}{2}\right)^2 + \varepsilon} \end{aligned}$$

Es kommt nur die positive Wurzel als Lösung infrage, da die negative Wurzel zu einer negativen Lösung führt, die sich nicht als Fehlertoleranz eignet.

Für  $\bar{\varepsilon}$  gibt es Maximalabstände  $\delta_f$  und  $\delta_g$ , so dass Toleranzschranke  $|f(x) - F| \leq \bar{\varepsilon}$  und  $|g(x) - G| \leq \bar{\varepsilon}$  eingehalten wird.

Wir nutzen den kleineren Maximalabstand  $\delta := \min(\delta_f, \delta_g)$  um die Fehlerschranke  $\bar{\varepsilon}$  für  $f$  und  $g$  zu erfüllen. Somit ist auch die Fehlerschranke  $|f(x)g(x) - FG| \leq \varepsilon$  erfüllt.

Zusammenfassend gesagt, finden wir zu einer vorgegebenen Fehlerschranke  $\varepsilon$  eine Maximalabweichung  $\delta$ , so dass für alle Argumente  $x$  mit  $|x - X| \leq \delta$  die Ungleichung  $|f(x)g(x) - FG| \leq \varepsilon$  erfüllt ist. Das Produkt  $f(x)g(x)$  hat also für die Annäherung von  $x$  an  $X$  den Grenzwert  $FG$ .

**Linearkombination** Seien  $\alpha$  und  $\beta$  zwei reelle Konstanten. Konvergieren zwei Funktionen  $f(x)$  und  $g(x)$  bei Annäherung von  $x$  an eine Stelle  $X$  gegen Grenzwerte  $F$  beziehungsweise  $G$ , so konvergiert auch  $\alpha \cdot f(x) + \beta \cdot g(x)$  gegen den Grenzwert  $\alpha \cdot F + \beta \cdot G$ .

Für den uninteressanten Fall  $\alpha = \beta = 0$  ist klar, dass der Grenzwert von  $\alpha f(x) + \beta g(x)$  Null ist. Nehmen wir jetzt also an, dass mindestens eine der Zahlen  $\alpha$  und  $\beta$  von Null verschieden ist. Wir gehen bei der Fehlerabschätzung für  $\alpha f(x) + \beta g(x)$  analog zu der beim Produkt vor:

$$\begin{aligned} |\alpha f(x) + \beta g(x) - (\alpha F + \beta G)| &= |\alpha(f(x) - F) + \beta(g(x) - G)| \\ &\leq |\alpha| \cdot |f(x) - F| + |\beta| \cdot |g(x) - G| \\ &\leq (|\alpha| + |\beta|)\bar{\varepsilon} =: \varepsilon \end{aligned}$$

Da wir den Trivialfall  $\alpha = \beta = 0$  ausgeschlossen haben, können wir nach  $\bar{\varepsilon}$  auflösen:

$$\bar{\varepsilon} = \frac{\varepsilon}{|\alpha| + |\beta|}$$

Die Fehlertoleranz  $\varepsilon$  für die Linearkombination lässt sich also in eine Fehlertoleranz  $\bar{\varepsilon}$  für  $f$  und  $g$  rückrechnen. Damit finden wir Maximalabweichungen  $\delta_f$  und  $\delta_g$ , für  $f$  beziehungsweise  $g$ , bei denen die Fehlertoleranz  $\bar{\varepsilon}$  jeweils eingehalten wird. Mit der Maximalabweichung  $\delta := \min(\delta_f, \delta_g)$  wird für  $f$  und  $g$  die Fehlertoleranz  $\bar{\varepsilon}$  eingehalten. Die Linearkombination erfüllt nach obiger Rechnung die Fehlertoleranz  $\varepsilon$ . Der Grenzwert der Linearkombination ist somit die Linearkombination der Grenzwerte  $F$  und  $G$ .

## 2.4 Stetigkeit

Für den Differenzenquotient  $\frac{\Delta f}{\Delta x}(x_1, x_2)$  der quadratischen Funktion  $f(x) = x^2$  haben wir in Abschnitt 2.1 die nennerfreie Berechnungsvorschrift  $\frac{\Delta f}{\Delta x}(x_1, x_2) = x_2 + x_1$  gefunden (siehe Gleichung (9)). Die Auswertung der nennerfreien rechten Seite  $x_2 + x_1$  an der Stelle  $x_2 = x_1$  hat uns zur Vermutung geführt, dass für die Ableitung  $\frac{df}{dx}(x_1) = 2x_1$  gilt.

In diesem Abschnitt sehen wir, dass  $g(x_2) = x_2 + x_1$  ein Beispiel für eine an der Stelle  $x_1$  stetige Funktion ist. Bei Funktionen mit dieser schönen Eigenschaft kann man den Grenzwert an der Stelle  $x_2 = x_1$  einfach durch Einsetzen von  $x_1$  berechnen.

**Definition 3** Sei  $f(x)$  eine reelle Funktion in Abhängigkeit eines reellen Arguments  $x$ . Die Funktion  $f(x)$  ist an einer Stelle  $X$  stetig, wenn der Grenzwert von  $f(x)$  bei Annäherung von  $x$  an  $X$  gleich  $f(X)$  ist.

Ist  $f(x)$  an allen Stellen  $X$  ihres Definitionsbereiches stetig, so wird  $f(x)$  einfach nur als stetig bezeichnet.

Lineare Funktionen

$$f(x) = p_0 + p_1 x$$

sind stetig. Um uns davon zu überzeugen müssen wir nach der Grenzwertdefinition zu jeder Fehlerschranke  $\varepsilon$  einen Maximalabstand  $\delta$  mit  $|f(x) - f(X)| \leq \varepsilon$  für alle  $x$  mit  $|x - X| \leq \delta$  finden. Ist  $p_1$  gleich null, so ist der Fehler  $|f(x) - f(X)|$  unabhängig von  $x$  und  $X$  gleich null, der Grenzwert ist also  $f(X)$ .

Für den Fall  $p_1 \neq 0$  nutzen wir  $\delta = \frac{\varepsilon}{|p_1|}$  als Maximalabstand für den die Fehlerschranke  $\varepsilon$  eingehalten wird:

$$|f(x) - f(X)| = |p_0 + p_1 x - (p_0 + p_1 X)| = |p_1| \cdot |x - X| \leq |p_1| \frac{\varepsilon}{|p_1|} = \varepsilon$$

Lineare Funktionen sind also stetig.

Seien zwei an der Stelle  $X$  stetige Funktionen  $f(x)$  und  $g(x)$  gegeben. Nach Abschnitt 2.3 ist der Grenzwert des Produkts  $f(x) \cdot g(x)$  gleich dem Produkt der Grenzwerte  $f(X) \cdot g(X)$  und der Grenzwert einer Linearkombination  $\alpha f(x) + \beta g(x)$  gleich der Linearkombination der Grenzwerte  $\alpha f(X) + \beta g(X)$ . Die Stetigkeit der Funktionen  $f$  und  $g$  überträgt sich also auf ihr Produkt und ihre Linearkombinationen.

Die Stetigkeit der Potenzfunktionen  $f(x) = x^n$  für  $n = 0, 1, 2, \dots$  ergibt sich daraus, dass diese als Produkte der linearen Funktion  $g(x) = x$  mit sich selber darstellbar sind.

## 3 Polynome

### 3.1 Monome und Polynome

Polynome sind wichtige Hilfsmittel in der Mathematik und Physik. Zum Beispiel kann man mit ihnen an Stützstellen vorgegebene Werte glatt interpolieren und Differentialgleichungen der Sorte, die wir im Einleitungsabschnitt kennengelernt haben, numerisch lösen.

Die Potenzfunktionen  $f(x) = x^n$  mit  $n = 1, 2, \dots$  werden auch als *Monome* bezeichnet. Linearkombination

$$f(x) = p_0 + p_1 \cdot x + p_2 \cdot x^2 + \dots + p_n \cdot x^n \quad (11)$$

von Monomen sind *Polynome*. Die Koeffizienten  $p_i$  mit  $i = 0, \dots, n$  sind dabei vorgegebene Konstanten.

Für Summen, wie in Gleichung (11) nutzen wir im Folgenden auch das Summenzeichen

$$f(x) = \sum_{i=0}^n p_i x^i \quad (12)$$

Obwohl  $x^0$  an der Stelle  $x = 0$  nicht auswertbar ist, wird formal vereinbart, dass bei Variablen wie  $x$  der Term  $x^0$  als 1 zu interpretieren ist, um die Kurzschreibweise (12) effizient einsetzen zu können.

Polynome ersten Grades, d.h., mit  $n = 1$  kennst du sicher, das sind linearen Funktionen mit Konstantanteil  $p_0$  und Anstieg  $p_1$ . Zum Beispiel ergibt sich mit den Koeffizienten  $p_0 = 1$  und  $p_1 = 2$  das Polynom  $p(x) = 1 + 2x$ , das im nächsten Abschnitt als Gleichung (14) auftaucht.

Auch quadratische Polynome, d.h., mit  $n = 2$  habt ihr sicher schon in der Schule gehabt. Beispielsweise ist  $p(x) = 1 + \frac{7}{2}x - \frac{3}{2}x^2$  das im nächsten Abschnitt berechnete Interpolationspolynom (16) mit  $p_0 = 1$ ,  $p_1 = \frac{7}{2}$  und  $p_2 = -\frac{3}{2}$ .

Polynomfunktionen  $p(x) = p_0 + p_1 \cdot x + p_2 \cdot x^2 + \dots + p_n x^n$  sind als Linearkombinationen von Potenzfunktionen ebenfalls stetig.

Der Grenzwert einer Polynomfunktion  $p(x)$  bei Annäherung von  $x$  an eine Stelle  $X$  kann also immer durch Auswertung  $p(X)$  des Polynoms an der Stelle  $X$  berechnet werden.

### 3.2 Polynominterpolation

Zur Erstellung eines Höhenprofils eines Berges wird an einer Reihe von Stellen  $x_0, x_1, \dots, x_n$  in horizontaler Richtung jeweils die Höhe  $y_0, y_1, \dots, y_n$  vermessen. Man hat nur einzelne Stellen und einzelne Höhenwerte, der Höhenverlauf des Berges ist jedoch eigentlich glatt.

Verbindet man die Punkte  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  durch Geradenstücke, so entstehen an den Verbindungspunkten  $(x_k, y_k)$  mit  $k = 1, \dots, n-1$  unerwünschte Knicke (siehe verbindender Polygonzug in Abbildung 7).

Statt der Verbindungsgeraden kann man auch eine glatte Funktion  $p(x)$  zur Verbindung der Punkte nutzen. Als Bedingung dafür, dass diese Funktion die Punkte  $(x_k, y_k)$  mit  $k = 0, \dots, n$  verbindet, muss  $p(x)$  an den Stellen  $x_k$  die Werte  $y_k$  annehmen:

$$p(x_k) = y_k$$

für  $k = 0, \dots, n$ . Nach der Konstruktion der Funktion  $p(x)$ , kann man diese auch an anderen Stellen als  $x = x_0, \dots, x_n$  auswerten. Die Konstruktion der Funktion  $p(x)$  und die Auswertung zwischen den vorgegebenen Stellen bezeichnet man als *Interpolation*.

Am Häufigsten nutzt man Polynome als glatte Funktion  $p(x)$  bei der Interpolation.

Wir schauen uns im Folgenden das leicht verständliche und wichtige Neville-Verfahren zur Polynominterpolation an. Dieses Verfahren bildet zum Beispiel auch die Grundlage für die bekanntere Newton-Interpolation, auf die wir hier jedoch nicht eingehen.

Beim Neville-Verfahren baut man das Interpolationspolynom rekursiv auf.

Sei  $p(x; x_0, \dots, x_{n-1})$  ein Polynom mit Höchstgrad  $(n-1)$ , das die Werte  $y_0, \dots, y_{n-1}$  an den Stellen  $x_0, \dots, x_{n-1}$  interpoliert und analog sei  $p(x; x_1, \dots, x_n)$  ein Polynom mit Höchstgrad  $(n-1)$ , das die Werte  $y_1, \dots, y_n$  an den Stellen  $x_1, \dots, x_n$  interpoliert.

Wir schauen uns die Eigenschaften des daraus konstruierten Polynoms

$$p(x; x_0, \dots, x_n) := \frac{p(x; x_0, \dots, x_{n-1}) \cdot (x_n - x) + p(x; x_1, \dots, x_n) \cdot (x - x_0)}{x_n - x_0}. \quad (13)$$

an:

- Das konstruierte Polynom hat Höchstgrad  $n$ , denn im Zähler kommen die Faktoren  $(x_n - x)$  und  $(x - x_0)$  hinzu, in denen jeweils einmal  $x$  vorkommt. Es kann jedoch sein, dass sich die entstehenden Terme von Grad  $n$  gerade aufheben. Deshalb kann man nur eine Aussage für den Höchstgrad treffen und nicht für den Polynomgrad selber.
- Das konstruierte Polynom interpoliert die Werte  $y_1, \dots, y_{n-1}$  an den Stellen  $x_1, \dots, x_{n-1}$ :  
Für  $k = 1, \dots, n-1$  haben  $p(x_k; x_0, \dots, x_{n-1})$  und  $p(x_k; x_1, \dots, x_n)$  den gleichen Wert  $y_k$  und man kann diesen Faktor ausklammern

$$p(x_i; x_0, \dots, x_n) = \frac{y_k \cdot ((x_n - x) + (x - x_0))}{x_n - x_0} = y_k$$

- Zusätzlich interpoliert das konstruierte Polynom den Wert  $y_0$  an der Stelle  $x_0$  und den Wert  $y_n$  an der Stelle  $x_n$ :

Bei  $x = x_0$  fällt der hintere Term des Zählers mit Faktor  $(x - x_0)$  weg. Übrig bleibt

$$p(x_0; x_0, \dots, x_n) = \frac{p(x_0, x_0, \dots, x_{n-1}) \cdot (x_n - x_0)}{x_n - x_0} = p(x_0, x_0, \dots, x_{n-1}) = y_0$$

Dabei wurde berücksichtigt, dass  $p(x; x_0, \dots, x_{n-1})$  an der Stelle  $x_0$  den Wert  $y_0$  interpoliert.

Analog fällt bei  $x = x_n$  der vordere Term mit Faktor  $(x_n - x)$  weg. Es ergibt sich wie im vorhergehenden Fall  $p(x_n; x_0, \dots, x_n) = p(x_n; x_1, \dots, x_n) = y_n$ .

Als Start für die rekursive Konstruktion der Interpolationspolynome mittels (13) kann man die konstanten Polynome  $p(x, x_0) = y_0$  nutzen.

**Beispiel** Wir konstruieren das Interpolationspolynom 2. Grades mit folgenden Interpolationspunkten:

$i$	$x_i$	$y_i$
0	0	1
1	1	3
2	2	2

Tabelle 1: Stützstellen und Werte für das Interpolationspolynom

Polynome vom Grad 0, die die Daten an jeweils einer Stelle interpolieren:

$$p(x; 0) = 1 \qquad p(x; 1) = 3 \qquad p(x; 2) = 2$$

Polynome vom Grad 1, die die Daten an jeweils zwei Stellen interpolieren:

$$p(x; 0, 1) = \frac{p(x; 0) \cdot (x_1 - x) + p(x; 1) \cdot (x - x_0)}{x_1 - x_0} = \frac{1 \cdot (1 - x) + 3 \cdot (x - 0)}{1 - 0} = 2x + 1 \quad (14)$$

$$p(x; 1, 2) = \frac{p(x; 1) \cdot (x_2 - x) + p(x; 2) \cdot (x - x_1)}{x_2 - x_1} = \frac{3 \cdot (2 - x) + 2 \cdot (x - 1)}{2 - 1} = -x + 4 \quad (15)$$

Polynom vom Grad 2, das den gesamten Datensatz interpoliert:

$$\begin{aligned} p(x; 0, 1, 2) &= \frac{p(x; x_0, x_1) \cdot (x_2 - x) + p(x; x_1, x_2) \cdot (x - x_0)}{x_2 - x_0} \\ &= \frac{(2x + 1) \cdot (2 - x) + (-x + 4) \cdot (x - 0)}{2 - 0} \\ &= -\frac{3}{2}x^2 + \frac{7}{2}x + 1 \end{aligned} \quad (16)$$

Das resultierende Polynom ist in Abbildung 7 dargestellt. Die Interpolationspunkte an den Stellen  $(x_i, y_i)$  sind durch kleine blaue Punkte gekennzeichnet.

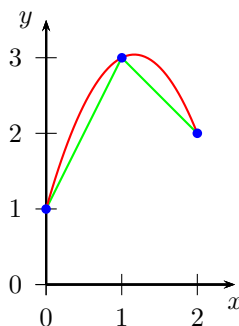


Abbildung 7: Verbindender Polygonzug (grün) und Interpolationspolynom (rot) für die blau dargestellten Punkte  $(x_0, y_0) = (0, 1)$ ,  $(x_1, y_1) = (1, 3)$  und  $(x_2, y_2) = (2, 2)$

## 4 Differenzialquotienten von Polynomen

In Abschnitt 3.1 haben wir Polynome

$$p(x) = \sum_{i=0}^n p_i x^i$$

als Linearkombinationen von Monomen  $x^0, x^1, \dots$  kennengelernt. Die Nützlichkeit von Polynomen für die Interpolation hat sich in Abschnitt 3.2 erwiesen.

Dort stand die Aufgabe, aus Stellen  $x_0, x_1, \dots, x_n$  mit zugehörigen Höhenangaben  $y_0, y_1, \dots, y_n$  ein glattes Höhenprofil eines Berges zu erstellen.

Ob man mit einem bestimmten Fahrzeugtyp den Berg überwinden kann, hängt von seiner maximalen Steigung ab.

Aus den diskreten Höhenmessungen kann man die Steigung des Berges approximieren, indem man den Differenzialquotient des mit Polynomen interpolierten Höhenprofils bestimmt. Die Berechnung des Differenzialquotients von Polynomen ist Inhalt dieses Hauptabschnittes.

### 4.1 Differenzialquotienten von linearen Funktionen und Monomen

In der Einleitung wurde der Differenzialquotient als Anstieg der Tangenten eingeführt. Lineare Funktionen beschreiben Geraden, bei denen der Anstieg konstant ist.

Alle Sekanten einer linearen Funktion stimmen mit der durch sie beschriebenen Gerade überein und haben den Anstieg dieser Gerade. Als Grenzwert haben auch alle Tangenten diesen Anstieg.

Für den formalen Test beschreiben wir die lineare Funktion mit zwei Konstanten  $p_0$  und  $p_1$  durch die Formel

$$f(x) = p_0 + p_1 x$$

Für beliebige voneinander verschiedene Stellen  $x_1, x_2$  ergibt sich der Sekantenanstieg gemäß

$$\begin{aligned} \frac{\Delta f}{\Delta x}(x_1, x_2) &= \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ &= \frac{(p_0 + p_1 x_2) - (p_0 + p_1 x_1)}{x_2 - x_1} \\ &= \frac{p_1(x_2 - x_1)}{x_2 - x_1} = p_1. \end{aligned} \tag{17}$$

Der Anstieg aller Sekanten ist also konstant gleich  $p_1$  und damit ist auch der Anstieg als Grenzwert des Sekantenanstiegs gleich  $p_1$ .

Aus den Differenzenquotienten der Potenzfunktion  $f(x) = x^n$  vom Grad  $n = 2, 3, \dots$

$$\begin{aligned}\frac{\Delta f}{\Delta x}(x_1, x_2) &= \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ &= \frac{x_2^n - x_1^n}{x_2 - x_1}\end{aligned}$$

erhält man mit der erweiterten binomischen Formel  $a^n - b^n = (a - b) \cdot \left(\sum_{i=0}^{n-1} a^i b^{n-1-i}\right)$  aus Anhang A wieder eine Form

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \frac{(x_2 - x_1) \cdot \left(\sum_{i=0}^{n-1} x_2^i x_1^{n-1-i}\right)}{x_2 - x_1},$$

aus der sich der Nenner herauskürzt:

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \sum_{i=0}^{n-1} x_2^i x_1^{n-1-i}$$

Die rechte Seite in der letzten Darstellung des Differenzenquotienten ist ein Polynom in der Variable  $x_2$  und damit an der Stelle  $x_2 = x_1$  stetig. Wir können den Differenzialquotient also durch Auswertung der rechten Seite an der Stelle  $x_2 = x_1$  berechnen:

$$\frac{df}{dx}(x_1) = \sum_{i=0}^{n-1} \underbrace{x_1^i x_1^{n-1-i}}_{n\text{-mal } x_1^{n-1}} \quad (18)$$

Zuletzt fassen wir die rechts stehenden Summenterme zusammen. Da wir jetzt nur noch ein Argument zu berücksichtigen haben, können wir dabei  $x_1$  auch in  $x$  umbenennen:

$$\frac{df}{dx}(x) = nx^{n-1} \quad (19)$$

Die letzte Gleichung ist streng genommen nur für  $n = 2, 3, \dots$  anwendbar, da sie für  $x = 0$  in den Fällen  $n = 0$  und  $n = 1$  auf die nicht definierten Terme  $0 \cdot \frac{1}{0}$  beziehungsweise  $1 \cdot 0^0$  führt. Das ist einer der Gründe dafür, dass wir lineare Funktionen vorher separat behandelt haben.

Jedoch wird, wie bei der Einführung des Summenzeichens in Abschnitt 3.1 zur Vereinfachung der Schreibweise zumindest im Fall  $n = 1$  der Term  $0^0$  als 1 interpretiert, so dass die Formel  $\frac{df}{dx}(x) = nx^{n-1}$  auch in diesem Fall gültig bleibt.

## 4.2 Linearität von Differenzen- und Differenzialquotienten

Der Differenzialquotient ist als auf Funktionen anzuwendende Operation *linear*. Diese Eigenschaft ist außerordentlich wichtig. Abschnitt 4.3 wird uns einen Eindruck davon vermitteln, wenn wir mit Hilfe der Linearität Differenzialquotienten von Polynomen auf Differenzialquotienten von Monomen zurückführen.

In diesem Abschnitt schauen wir uns zuerst an, was die Linearität des Differenzialquotienten bedeutet soll. Danach überzeugen wir uns davon, dass Differenzenquotienten linear sind und dass sich diese Eigenschaft durch Grenzwertbildung auf Differenzialquotienten überträgt.

Seien  $\alpha, \beta$  reelle Konstanten und  $f(x), g(x)$  reelle Funktionen, die an der Stelle  $x$  differenzierbar sind.

Dann ist die Linearkombination

$$f(x) = \alpha g(x) + \beta h(x)$$

der Funktionen ebenfalls differenzierbar und es gilt die Gleichung

$$\frac{df}{dx}(x) = \frac{d}{dx}(\alpha g(x) + \beta h(x)) = \alpha \frac{dg}{dx}(x) + \beta \frac{dh}{dx}(x),$$

das heißt, der Differenzialquotient ist *linear*.

Analog gelten für Differenzenquotienten der Funktion  $f$  mit zwei voneinander verschiedenen Stellen  $x_1, x_2$  die Gleichungen

$$\begin{aligned}\frac{\Delta f}{\Delta x}(x_1, x_2) &= \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ &= \frac{\alpha g(x_2) + \beta h(x_2) - (\alpha g(x_1) + \beta h(x_1))}{x_2 - x_1} \\ &= \alpha \frac{g(x_2) - g(x_1)}{x_2 - x_1} + \beta \frac{h(x_2) - h(x_1)}{x_2 - x_1} \\ &= \alpha \frac{\Delta}{\Delta x} g(x_1, x_2) + \beta \frac{\Delta}{\Delta x} h(x_1, x_2)\end{aligned}$$

Im Abschnitt 2.3 über Grenzwertsätze haben wir uns schon davon überzeugt, dass der Grenzwert einer Linearkombination gleich der Linearkombination von Grenzwerten ist. Bilden wir auf beiden Seiten der Gleichung

$$\frac{\Delta f}{\Delta x}(x_1, x_2) = \alpha \frac{\Delta g}{\Delta x}(x_1, x_2) + \beta \frac{\Delta h}{\Delta x}(x_1, x_2)$$

den Grenzwert für die Annäherung von  $x_2$  an  $x_1$ , so erhalten wir die Gleichung

$$\frac{df}{dx}(x_1) = \alpha \frac{dg}{dx}(x_1) + \beta \frac{dh}{dx}(x_1).$$

Somit hat sich die Linearität des Differenzialquotienten bestätigt.

### 4.3 Differenzialquotienten von Polynomen

In Abschnitt 3.1 wurden Monome als Potenzfunktionen  $x^n$  mit  $n = 0, 1, 2, \dots$  eingeführt, wobei  $x^0 = 1$  gesetzt wird unabhängig davon, dass  $x$  auch null werden kann. Polynome sind Linearkombinationen

$$p(x) = p_0 + p_1 x + \dots + p_n x^n$$

von Monomen mit konstanten Koeffizienten  $p_0, p_1, \dots, p_n$ . Mit dem Summenzeichen lautet dieselbe Gleichung

$$p(x) = \sum_{i=0}^n p_i x^i.$$

Für den Differenzialquotient von Monomen  $x^n$  mit  $n = 1, 2, \dots$  haben wir bereits in Abschnitt 4.3 die Gleichung  $\frac{d}{dx}(x^n) = n \cdot x^{n-1}$  hergeleitet. Dabei nutzen wir die Festlegung  $x^0 = 1$  unabhängig davon, dass  $x$  auch null werden kann.

In Abschnitt 4.2 haben wir uns davon überzeugt, dass der Differenzenquotient eine lineare Operation darstellt. Damit ist der Differenzialquotient einer Linearkombination differenzierbarer Funktionen gleich der Linearkombination der Differenzialquotienten der Funktionen.

Die Formel für die Ableitung von Polynomen ist jetzt nur noch ein Zusammensetzen all dieser Bausteine.

Da die Monome differenzierbar sind und die Polynome Linearkombinationen von Monomen, ist die Ableitung von Polynomen gerade die zugehörige Linearkombination der Differenzialquotienten der eingehenden Monome.

Gesagt, getan. Wir gehen von der Darstellung eines Polynoms mit dem Summenzeichen aus:

$$p(x) = \sum_{i=0}^n p_i x^i \tag{20}$$

Der Differenzialquotient des Konstantanteils  $p_0$  ist null. Das berücksichtigen wir, indem wir die Summe für den Differenzialquotient des Polynoms erst ab 1 laufen lassen. Die Differenzialquotienten der anderen Monome  $x^i$  mit  $i = 1, 2, \dots$  sind  $i \cdot x^{i-1}$ . Das haben wir uns eben in Abschnitt 4.1 angesehen.

Wenden wir das auf die verbleibenden Monome in Gleichung (20) an, erhalten wir die endgültige Formel

$$\frac{dp}{dx}(x) = \sum_{i=1}^n p_i i x^{i-1}$$

für die Ableitung des Polynoms  $p(x)$  aus Gleichung (20).

## 5 Numerische Lösung von Differenzialgleichungen

Zur Lösung von Differenzialgleichungen, wie wir sie im Einleitungsabschnitt kennengelernt haben, gibt es viele numerische Verfahren und ein Großteil dieser Verfahren nutzt Polynominterpolation.

Wir lösen die Differenzialgleichung

$$\frac{dy}{dx}(x) = \alpha y(x) \quad (21)$$

mit einem solchen Verfahren numerisch approximativ für kleine Beträge von  $x$ . Die Differenzialgleichung (21) passt zur Gleichung (7) für das elektrische Netzwerk in Abschnitt 1.3, wenn wir  $x = t$ ,  $y = U$ ,  $\alpha = -\frac{1}{\tau}$  und  $U_B = 0$  setzen. Sie beschreibt dann den Entladevorgang eines RC-Gliedes.

Wir nutzen bei dem numerischen Verfahren aus, dass sich der Betrag von  $x^n$  stark verkleinert, wenn der Betrag von  $x$  kleiner wird. Nutzen wir zum Beispiel  $x$  nur im Bereich von  $-0.1$  bis  $0.1$ , so ändert sich  $x^5$  nur im Bereich von  $-0.00001$  bis  $0.00001$ .

Davon lassen wir uns leiten und setzen für  $y(x)$  in Gleichung (21) den polynomialen Ansatz

$$y(x) = \sum_{i=0}^n p_i x^i \quad (22)$$

ein. Wir gleichen dann die bisher noch freien Koeffizienten  $p_i$  so ab, dass die Gleichung für so viele Potenzen niedrigen Grades erfüllt wird, wie es möglich ist. Den Fehler in den höheren Potenzen von  $x$  vernachlässigen wir. Dieser Fehler wird sehr klein, wenn die Exponenten der vernachlässigten Potenzen groß sind und wir den Betrag von  $x$  hinreichend klein wählen.

Einsetzen des Ansatzes (22) in Gleichung (21) liefert:

$$\frac{d}{dx} \left( \sum_{i=0}^n p_i x^i \right) = \alpha \left( \sum_{i=0}^n p_i x^i \right)$$

Auf der linken Seite wenden wir die Differenziation auf das Polynom an und erhalten

$$\sum_{i=1}^n p_i i x^{i-1} = \sum_{i=0}^n \alpha p_i x^i. \quad (23)$$

Hier sei noch einmal daran erinnert, dass beim Differenzieren der erste Summand  $p_0 x^0 = p_0$  von  $\sum_{i=0}^n p_k x^i$  herausfällt. Deshalb fängt die Summe erst bei Index 1 an. Für einen Vergleich der Koeffizienten zu den Potenzen von  $x$  ist es günstiger, wenn links und rechts die selben Potenzen von  $x$  stehen. Deshalb ersetzen wir auf der rechten Seite den Exponenten  $i$  durch einen neuen Index  $\bar{i} - 1$ . Die spezielle Wahl der Ersetzung erweist sich zuletzt als günstig. Aus  $\bar{i} - 1 = i$  ergibt sich  $\bar{i} = i + 1$ . Das benutzt man zum Beispiel bei der Berechnung des neuen Startindex  $\bar{i} = i + 1 = 0 + 1 = 1$  und Endindex  $\bar{i} = i + 1 = n + 1$  für die Summe auf der rechten Seite von (23). Nach der Indexersetzung hat Gleichung (23) insgesamt die Form

$$\sum_{i=1}^n p_i i x^{i-1} = \sum_{\bar{i}=1}^{n+1} \alpha p_{\bar{i}-1} x^{\bar{i}-1}.$$

Die Zählvariable  $\bar{i}$  ist der neue Summationsindex der Summe auf der rechten Seite, für den wir uns den Namen selber aussuchen können, solange er nicht mit bereits in der Summe vorkommenden Variablennamen kollidiert. Wir bezeichnen ihn wieder mit  $i$ , damit der Vergleich der in beiden Summen auftretenden Potenzen von  $x$  einfacher wird.

Bei der entstehenden Gleichung

$$\sum_{i=1}^n p_i i x^{i-1} = \sum_{i=1}^{n+1} \alpha p_{i-1} x^{i-1} \quad (24)$$

können wir die Koeffizienten der Potenzen  $x^{i-1}$  für  $i = 1, \dots, n$  abgleichen. Die Koeffizienten von  $x^{i-1}$  für  $i = 1, \dots, n$  auf der linken und rechten Seite der Gleichung sind  $p_i i$  beziehungsweise  $\alpha p_{i-1}$ . Die Gleichungen für deren Koeffizientenabgleich der Potenzen in (24) lauten

$$i p_i = \alpha p_{i-1} \quad \text{für } i = 1, \dots, n$$



Der Koeffizientenabgleich für die höchste Potenz  $x^n$  in Gleichung (24) liefert die Gleichung

$$0 = \alpha p_n$$

Damit erhält man insgesamt das Gleichungssystem

$$1 \cdot p_1 = \alpha p_0 \quad (25)$$

$$2 \cdot p_2 = \alpha p_1 \quad (26)$$

$$\vdots \quad (27)$$

$$n \cdot p_n = \alpha p_{n-1} \quad (28)$$

$$0 = \alpha p_n \quad (29)$$

für die Koeffizienten  $p_i$ .

Dieses Gleichungssystem ist von unten nach oben lösbar ( $p_n = 0 \Rightarrow p_{n-1} = 0 \Rightarrow \dots \Rightarrow p_0 = 0$ ) und hat nur noch die Trivallösung, dass alle Koeffizienten  $p_i$  null sind, also  $y(x) = 0$  gilt.

Um nichttriviale Lösungen zu erhalten, vernachlässigen wir, wie bereits einleitend angesprochen, den Koeffizientenvergleich für die höchste Potenz.

Das kann man zum Beispiel so interpretieren, dass man statt der Lösung für (21) die exakte Lösung der leicht gestörten Differenzialgleichung

$$\tau \frac{dy}{dx} = \alpha y - \alpha p_n x^n \quad (30)$$

ermittelt. Die Störung  $-\alpha p_n x^n$  der Differenzialgleichung bleibt klein, wenn wir uns nicht zu weit hinaus wagen. Das heißt, wenn wir die gestörte Differenzialgleichung nur für kleine Werte von  $x$  nutzen.

Der Koeffizientenvergleich für den Polynomansatz (22) bei (30) liefert anstelle von (29) die immer erfüllte Gleichung  $0 = 0$ , die aus dem System gestrichen werden kann.

Es bleiben die Gleichungen

$$ip_i = \alpha p_{i-1}$$

für  $i = 1, \dots, n$ , oder umgestellt nach  $p_i$  die Gleichungen

$$p_i = \alpha \frac{p_{i-1}}{i}. \quad (31)$$

In der letzten Form kann man die Gleichungen als rekursive Zuweisungskette interpretieren. Der Koeffizient  $p_i$  wird aus dem vorhergehenden  $p_{i-1}$  berechnet.

Der Koeffizient  $p_0$  bleibt dabei frei wählbar. Aus (22) erkennt man, dass  $p_0$  gerade der Wert von  $y(x)$  für  $x = 0$  ist. (Im Beispiel unseres Entladevorgangs ist das die Anfangsspannung des Kondensators.)

Wir können also zusätzlich

$$p_0 = y(0)$$

schreiben und den Anfangswert  $y(0)$  frei vorgeben.

Bei jedem Schritt  $i = 1, \dots, n$  des Systems (31) von Zuweisungen kommt ein Faktor  $\alpha$  hinzu und es wird durch den aktuellen Index  $i$  dividiert. So wird im Zähler  $p_0$  so oft mit  $\alpha$  multipliziert, wie die Zuweisung (31) angewandt wird und im Nenner entsteht das Produkt aller Indizes bis  $i$ . Auf diesem Weg erhält man für die Koeffizienten  $p_i$  mit  $i = 1, 2, \dots, n$  die Berechnungsvorschrift

$$p_i = \frac{\alpha^i y(0)}{1 \cdot 2 \cdot \dots \cdot (i-1) \cdot i}.$$

Das Produkt  $1 \cdot 2 \cdot \dots \cdot (i-1) \cdot i$  bezeichnet man auch als Fakultät von  $i$  und schreibt  $i! := 1 \cdot 2 \cdot \dots \cdot (i-1) \cdot i$ . Damit bekommen die Koeffizienten  $p_i$  für  $i = 1, \dots, n$  die Form

$$p_i = \frac{\alpha^i y(0)}{i!}. \quad (32)$$

Setzen wir diese Koeffizienten in den Ansatz (22) ein, ergibt sich die Polynomfunktion

$$y(x) = y(0) \cdot \sum_{i=0}^n \frac{1}{i!} (\alpha x)^i, \quad (33)$$

die als numerische Approximation der Lösung von (21) interpretiert wird.

In Abbildung 8 sind die exakte Lösung von (21) und die numerischen Approximationen (33) dargestellt. Außerdem ist die Abweichung der numerischen Approximation von der exakten Lösung zu sehen.

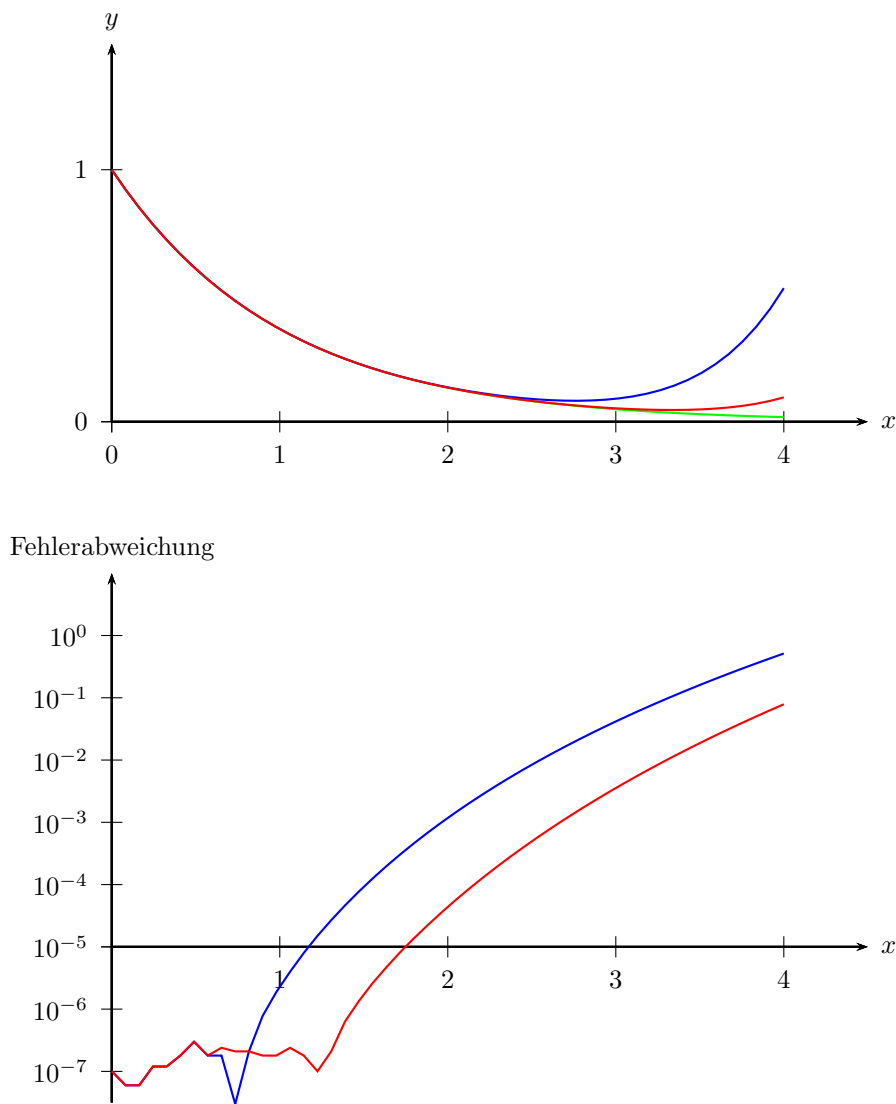


Abbildung 8: Approximation der Differentialgleichung  $\frac{dy}{dt} = -y$ ; Oben: Lösung und numerischen Approximationen; Unten: Abweichung der numerischen Approximationen von der exakten Lösung (im logarithmischen Maßstab); Grün: theoretische Lösung von (11); Blau: Numerische Approximation der Lösung mit  $n = 8$ ; Rot: Numerische Approximation der Lösung mit  $n = 10$ .

Das Rauschen auf einem Fehlerniveau von ungefähr  $10^{-7} \dots 10^{-8}$  für  $x \lesssim 1$  rührt von der Approximation reeller Zahlen durch Computerzahlen her. Es ist durch die Restriktionen der zur Programmierung verwendeten Sprache Postscript<sup>TM</sup> gegeben und damit unvermeidbar, solange man diese Programmiersprache benutzt. Andere Sprachen, wie zum Beispiel C++ bieten noch andere Datentypen zur Approximation von reellen Zahlen an, bei denen das unvermeidliche Fehlerniveau bei  $10^{-16}$  liegt.

Weiterhin ist auffällig, dass wir mit dem Maximalabstand  $\delta \approx 1.2$  bei  $n = 8$  für eine Fehlerschranke  $\varepsilon = 1e-5$  ziemlich weit mit unserer Approximation kommen.

Das liegt an den mit wachsendem  $n$  schnell kleiner werdenden Koeffizienten  $p_n = \frac{\alpha^n y(0)}{n!}$ . Die Fakultät  $n!$  wächst mit größer werdendem  $n$  schneller als jedes Polynom. Das lässt den Nenner schnell wachsen und den Zähler und den Bruch schnell kleiner werden.

Zum Abschluss soll noch erwähnt werden, dass man die Differenzialgleichung für größere Zeitintervalle lösen kann, indem man sie nur bis 1 löst, dort mit dem aktuellen Wert für  $y(1)$  neu startet, von 1 bis 2 löst und so weiter.

## A Verallgemeinerte Binomische Formel

Sicherlich habt ihr in der Schule bereits die binomische Formel  $(a-b) \cdot (a+b) = a^2 - b^2$  kennengelernt, von deren Gültigkeit man sich einfach durch Ausmultiplizieren überzeugt.

Das Abspalten des Faktors  $(a-b)$  geht auch bei Differenzen von höheren Potenzen als zwei. Genauer gilt für  $n = 2, 3, \dots$

$$a^n - b^n = (a-b) \cdot \left( \sum_{i=0}^{n-1} a^i b^{n-1-i} \right)$$

Wie bei der quadratischen Variante kann man sich von der Richtigkeit dieser Gleichung durch Ausmultiplizieren der rechten Seite überzeugen.

Das Distributivgesetz liefert zwei Summen.

$$(a-b) \cdot \left( \sum_{i=0}^{n-1} a^i b^{n-1-i} \right) = \sum_{i=0}^{n-1} a^{i+1} b^{n-1-i} - \sum_{i=0}^{n-1} a^i b^{n-i}$$

Beim Minuend kommt ein Faktor  $a$  hinzu und beim Subtrahend ein Faktor  $b$ .

Die Summenglieder des Subtrahenden heben die des Minuenden ab Index  $i = 1$  auf.

Das sieht man besonders einfach, wenn man durch die Indextransformation  $\bar{i} = i + 1$  die Potenzen im Minuenden an die im Subtrahenden angleicht

$$= \sum_{\bar{i}=1}^n a^{\bar{i}} b^{n-\bar{i}} - \sum_{i=0}^{n-1} a^i b^{n-i}$$

Nennen wir den Index  $\bar{i}$  in  $i$  um, erhalten wir mit

$$= \sum_{i=1}^n a^i b^{n-i} - \sum_{i=0}^{n-1} a^i b^{n-i}$$

zwei Summen als Minuend und Subtrahend, bei denen sich alle Summanden gegenseitig aufheben, bis auf diejenigen mit Index  $i = 0$  und  $i = n$ :

$$\begin{aligned} &= \underbrace{a^n}_{i=n} + \underbrace{\left( \sum_{i=1}^{n-1} a^i b^{n-i} - a^i b^{n-i} \right)}_{=0} - \underbrace{b^n}_{i=0} \\ &= a^n - b^n \end{aligned}$$

Übrig bleiben nur die zwei Summanden, die wir laut der verallgemeinerten binomischen Formel erwarten. Womit sich diese Formel bestätigt.