Ableitung

[Differentialrechnung]

Armin P. Barth

ETH zürich

Abstraktionen

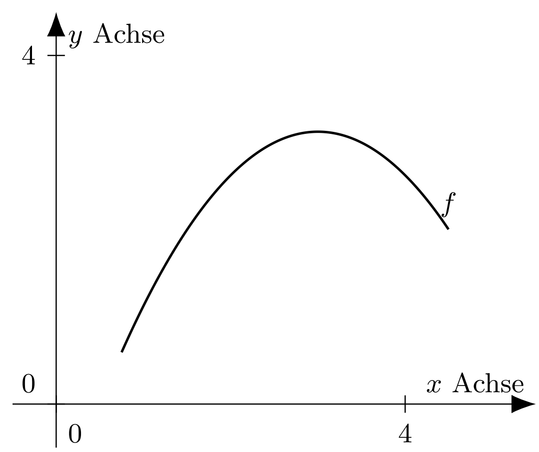
Es mag immer wieder erstaunen, dass die Mathematik in der Praxis so hervorragend funktioniert. Wir rechnen ja - genau genommen - nicht mit den realen Objekten. Wir ersetzen eine Kanonenkugel, einen Turmspringer oder gleich einen ganzen Planeten durch einen Massenpunkt, wir vereinfachen die Formen, idealisieren die Situationen und lassen viele Aspekte, die in Wirklichkeit auch eine Rolle spielen, ganz ausser Acht. Der Grund, weshalb trotzdem fast alles, was die Mathematik für die Praxis beisteuert, so tadellos funktioniert, ist, dass wir in den meisten Fällen *erfolgreich abstrahieren.*

Abstrahieren kommt aus dem Lateinischen und bedeutet *absondern, abziehen.* In der Mathematik versteht man unter Abstraktion (genauer: *Abstraktion der Identifizierung*) oftmals den Prozess des Abziehens gewisser Eigenschaften der untersuchten Objekte oder Prozesse und des gleichzeitigen Hervorhebens eines als besonders wichtig erachteten Merkmals. Die natürlichen Zahlen sind das Resultat einer solchen Abstraktion. Könnte man nämlich alle Zweiermengen (Mengen mit genau zwei Elementen) betrachten, also etwa die Menge der beiden Himmelskörper Sonne und Mond, die Menge der beiden Ohren am Kopf, die Menge der Begriffe „Mann“ und „Frau“, usw., und würde man alle Merkmale wegziehen, in denen sich einzelne dieser Mengen unterscheiden und dafür einzig das Merkmal „Anzahl“ hervorheben, so würde man zur abstrakten Idee der Zahl 2 gelangen. Unter diesem Gesichtspunkt können nun alle Zweiermengen identifiziert werden, und jede einzelne ist ein konkretes Beispiel zu dieser Zahl und veranschaulicht sie.

Genau das werden wir nun im grossen Stil umsetzen. Man trifft in der Praxis unzählige „Bewegungen“ an: die Bewegungen von Planeten oder Maschinenteilen, den Sprint einer Läuferin beim 100-Meter-Lauf, das Fahrverhalten eines Sportwagens, die Bewegung eines Echolotsignals oder eines Wasserpegels, die Art und Weise, wie sich die Konzentration eines Stoffes bei einer chemischen Reaktion „bewegt“, usw. Wenn wir alle spezifischen Merkmale dieser vielen Beispiele abziehen und einzig hervorheben, worum es im Kern geht, so können wir alle Beispiele identifizieren, und wir gelangen zur Idee der mathematischen Funktion (in einer Variablen) und zu Funktionswerten und Differenzen von Inputwerten und Funktionswerten und Quotienten solcher. Wir werden also über „reine Mathematik“ reden. Aber der enorme Vorteil wird sein, dass alle Erkenntnisse, die wir dabei gewinnen, für Myriaden von konkreten Alltagssituationen Gültigkeit haben werden.

Der Differenzenquotient

Gegeben sei also eine Funktion f, die in einem Intervall I der reellen Zahlen definiert ist.

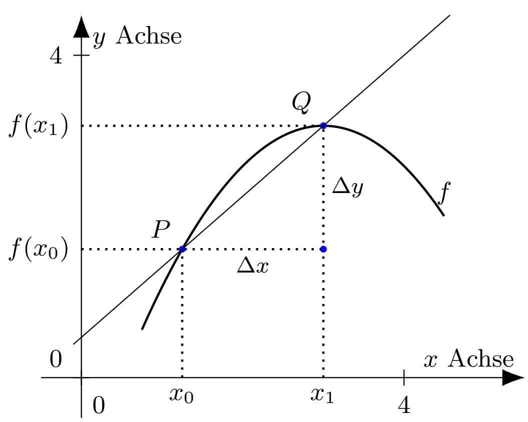


Wir können dabei an jede der ober erwähnten Bewegungen denken, dann könnte man sich unter x die Zeit und unter y die Position der Sprinterin, des Echolotsignals, die Höhe des Wasserpegels oder der Alkoholkonzentration oder auch die Geschwindigkeit des Sportwagens vorstellen, aber das ist alles einerlei. Im Kern haben wir immer eine Funktion vor uns, und über das Verhalten einer solchen Funktion werden wir nun neue Erkenntnisse gewinnen.

Zunächst betrachten wir eine beliebige Stelle $x_0 \in I$ und eine „nahe gelegene" Stelle $x_1 \in I$ $(x_1 \neq x_0)$. Die Funktionswerte $f(x_0)$ und $f(x_1)$ zu bestimmen, könnte also zum Beispiel bedeuten, zu zwei konkreten Zeitpunkten die Position der Läuferin während ihres 100-Meter-Sprintes zu betrachten. Aus der Kenntnis der beiden Stellen sowie der beiden Funktionswerte können wir sicher die Differenzen bilden:

$$\begin{aligned} \Delta x &= x_1 - x_0 \\ \Delta y &= f(x_1) - f(x_0). \end{aligned}$$

Dividiert man den Zuwachs in y-Richtung (Δy) durch den Zuwachs in x-Richtung (Δx), was erlaubt ist, da wir ja $x_1 \neq x_0$ voraussetzen, so erhält man – geometrisch gesprochen – die Steigung der Sekante durch die beiden Punkte $P(x_0, f(x_0))$ und $Q(x_1, f(x_1))$:



Wir sind uns aber bewusst, dass die geometrische Interpretation nur eine von vielen ist. Ebenso gut kann dieser Quotient die durchschnittliche Geschwindigkeit der Läuferin oder die durchschnittliche Abnahmerate der Alkoholkonzentration pro Zeiteinheit oder die durchschnittliche Beschleunigung des Sportwagens zwischen den Zeitpunkten x_0 und x_1 bedeuten; genau darin liegt ja die enorme Kraft der mathematischen Abstraktion.

Merke:

Zu einer gegebenen (über $I \subset \mathbb{R}$ definierten) Funktion f und zwei verschiedenen Stellen $x_0, x_1 \in I$ heisst der Term

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Differenzenquotient.

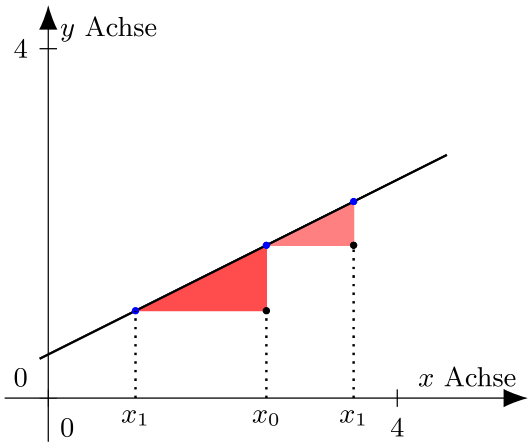
Er gibt die Steigung der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$ an.

Alternativ wird häufig x_1 als $x_0 + \Delta x$ oder als $x_0 + h$ bezeichnet. Dann nimmt der Differenzenquotient diese Gestalt an:

$$\begin{aligned} \frac{\Delta y}{\Delta x} &= \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \\ &= \frac{f(x_0 + h) - f(x_0)}{h}. \end{aligned}$$

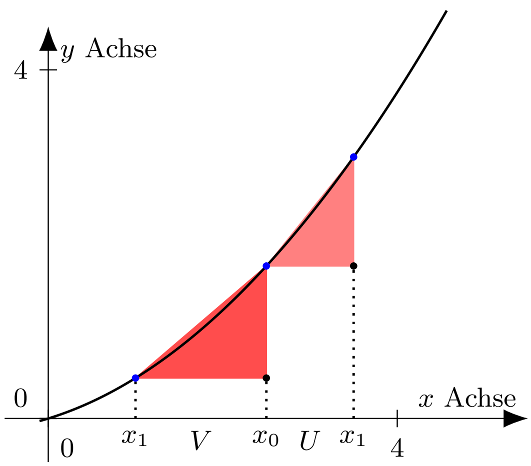
Links oder rechts?

Es muss betont werden, dass, im Gegensatz zu obiger Skizze, x_1 natürlich auch links von x_0 sein kann; wir haben ja einzig $x_1 \neq x_0$ verlangt. Kann das eine Rolle spielen? Betrachten wir einmal eine lineare und streng monoton wachsende Funktion und eine feste, aber beliebige Stelle x_0:



Wir wählen x_1 einmal links und einmal rechts von x_0. Für $x_0 < x_1$ liefert der Differenzenquotient einfach die Steigung der Geraden; sowohl Zähler als auch Nenner sind positiv. Für $x_1 < x_0$ sind Zähler und Nenner beide negativ, das Verhältnis liefert aber wieder denselben Wert, nämlich die Steigung der Geraden.

Betrachten wir als zweites Beispiel einen Ort-versus-Zeit-Graphen, der lokal vielleicht den Anfang eines Sprints darstellen kann:

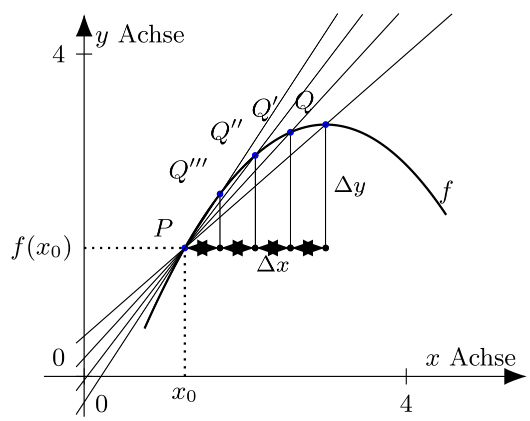


Für $x_0 < x_1$ liefert der Differenzenquotient die durchschnittliche Geschwindigkeit im Zeitintervall U, für $x_1 < x_0$ dagegen die durchschnittliche Geschwindigkeit im Zeitintervall V. Diese Werte können (und werden meist) verschieden sein. Da die Läuferin anfangs immer schneller rennt, ist hier die durchschnittliche Geschwindigkeit im Intervall V sicher kleiner als diejenige im Intervall U.

Wir müssen uns also einprägen, dass es, wenn man Sehnensteigungen ausrechnet, durchaus einen Unterschied macht, ob man x_1 links oder rechts von x_0 wählt. Das wird uns in Kürze noch intensiver beschäftigen.

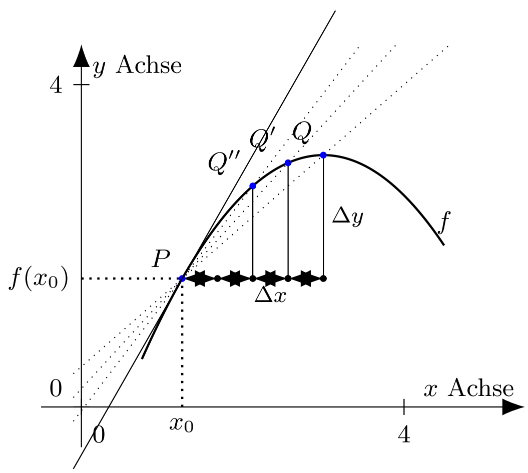
Die Ableitung

Der Differenzenquotient gibt nur die *durchschnittliche* Änderungsrate der Funktion im Intervall Δx an. Oftmals ist man aber interessiert an der *momentanen*. Um diese formal und rechnerisch in den Griff zu bekommen, bedienen wir uns einer ebenso einfachen wie überzeugenden Idee: Wir lassen Δx immer kleiner werden und untersuchen das Trendverhalten des Differenzenquotienten für $\Delta x \to 0$ bzw. $x_1 \to x_0$ bzw. $h \to 0$.



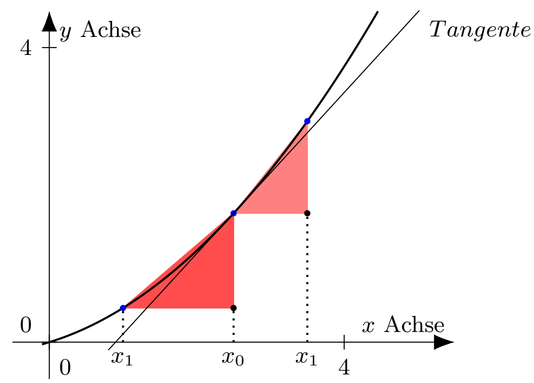
Dieses Vorgehen bedeutet zum Beispiel, die durchschnittliche Geschwindigkeit in einem immer kleineren Zeitintervall Δx zu bestimmen.

Es ist plausibel, dass man dadurch rechnerisch immer näher zu der momentanen Geschwindigkeit an der Stelle x_0 gelangt. Die momentane Geschwindigkeit eines bewegten Objektes zum Zeitpunkt $x_0 = 2$ wird sich ja in der Regel kaum stark unterscheiden von der durchschnittlichen zwischen den Zeitpunkten $x_0 = 2$ und $x_1 = 2.01$, egal, wie wechselhaft das Geschwindigkeitsverhalten des Objektes auch sein mag. Wir können daher sagen, dass wir die momentane Geschwindigkeit zum Zeitpunkt x_0 umso genauer erhalten, je kleiner wir im Differenzenquotienten Δx (bzw. h) wählen, beziehungsweise je näher x_1 bei x_0 liegt.

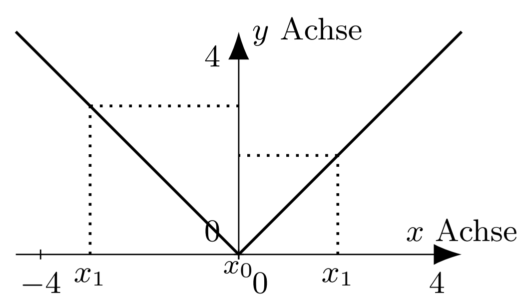


Wieder allgemein (also losgelöst von möglichen konkreten Bedeutungen der Grössen x und y) gesprochen, heisst das, dass wir in den meisten Fällen erwarten dürfen, dass der Differenzenquotient für $\Delta x \to 0$ konvergiert und dass der Grenzwert die *momentane Änderungsrate der Funktion an der Stelle* x_0 angibt oder – geometrisch gesprochen – die Steigung der Tangente an den Graphen im Punkt P. Die Grenzlage der Sekante ist die Tangente, und deren Steigung werden wir finden, wenn wir den erwähnten Grenzwertprozess durchführen. Und eben diese Steigung kann so viele konkrete, praktische, nützliche Bedeutungen haben.

Werden wir immer denselben Grenzwert erhalten, ob wir uns der Stelle x_0 von links oder von rechts nähern? Rücken wir ein weiteres Mal die oben schon verwendete Sprint-Funktion ins Blickfeld:



In diesem Beispiel liefert der Differenzenquotient für $x_1 < x_0$ stets kleinere Werte als für $x_1 > x_0$, da, wie gesagt, die Läuferin anfangs immer schneller rennt. Der Grenzwert für $x_1 \to x_0$ wird aber in beiden Fällen derselbe sein, nämlich die Steigung der Tangente an den Graphen an der Stelle x_0 beziehungsweise die Momentangeschwindigkeit der Läuferin zu diesem Zeitpunkt.



Tatsächlich gibt es, wenn auch selten, Fälle, bei denen es einen Unterschied macht, ob man sich bei der Grenzwertbildung von rechts oder von links der Stelle x_0 nähert. Betrachten wir dazu die Funktion $f(x) = |x|$ über $I = \mathbb{R}$:

An der Stelle $x_0 = 0$ ist

$$\lim_{x_1 \to 0} \left(\frac{f(x_1) - 0}{x_1 - 0} \right) = \lim_{x_1 \to 0} \left(\frac{|x_1|}{x_1} \right).$$

Für Stellen $x_1 > x_0$ ist dieser Grenzwert gleich 1. Ist aber $x_1 < x_0$ (das heisst nähern wir uns der Stelle 0 von links), dann ist dieser Grenzwert gleich -1. Der Grenzwert ist also nicht eindeutig. Das ist auch geometrisch klar: Will man an der Stelle $(0, 0)$ eine Tangente an den Graphen heften, so wird man sich wie Buridans Esel fühlen und sich zwischen den beiden gleichwertigen Möglichkeiten nicht entscheiden können.

Freilich sagen wir in einem solchen Fall, dass der Grenzwert für $x_1 \to x_0$ *nicht existiert*, denn ein Grenzwert muss ja, damit man ihn so nennt, eindeutig sein. Glücklicherweise führt bei den meisten Funktionen und an den meisten Stellen x_0 der Grenzwertprozess $x_1 \to x_0$ zum selben Wert, ob man sich von links oder von rechts der besagten Stelle nähert. Und dann liefert die Grenzwertbildung die Tangentensteigung, und diese Zahl nennen wir *erste Ableitung* der Funktion an der Stelle x_0. Die folgende Definition enthält alle wesentlichen Details.

Definition:
Existiert der Grenzwert

$$\lim_{\substack{x_1 \to x_0 \\ x_1 \neq x_0}} \frac{f(x_1) - f(x_0)}{x_1 - x_0} \qquad \textit{bzw.}$$

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \qquad \textit{bzw.}$$

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

so heisst er ***Differentialquotient*** *oder die* ***1. Ableitung der Funktion an der Stelle*** *x_0, und er wird mit*

$$f'(x_0) \qquad \textit{(lies: „}f\textit{ Strich von }x_0\textit{“)}$$

oder

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$$

bezeichnet. Die Funktion heisst dann ***differenzierbar an der Stelle*** *x_0. (Um die Notation nicht zu überfrachten, lassen wir den Zusatz $x_1 \neq x_0$ künftig weg.)*

Die 1. Ableitung der Funktion an der Stelle x_0 misst die Steigung der Tangente an den Graphen im Punkt $(x_0, f(x_0))$, also die momentane Änderungsrate.

Die Funktion, die jeder Stelle $x \in I$ den Wert $f'(x)$ zuordnet, heisst ***1. Ableitung (Differentiation) der Funktion f nach*** *x, und sie wird mit*

$$f' \quad \textit{oder} \quad \frac{\mathrm{d}f}{\mathrm{d}x}$$

bezeichnet.

Bildet man zu einer gegebenen Funktion deren 1. Ableitung, so ***leitet man die Funktion ab*** *oder* ***differenziert*** *sie.*

Da die 1. Ableitung einer gegebenen Funktion f in der Variablen x selber auch wieder eine Funktion in der Variablen x ist, kann man auch diese Funktion wieder ableiten und gelangt so zur **2. Ableitung** der Funktion f, in Zeichen:

$$f'' \quad \text{oder} \quad f^{(2)} \quad \text{oder} \quad \frac{d^2 f}{d^2 x}.$$

Die n-te Ableitung ($n \in \mathbb{N}$) der Funktion f bezeichnet man mit

$$f^{(n)} \quad \text{oder} \quad \frac{d^n f}{d^n x}.$$

Drei Beispiele

Erinnern wir uns erneut an den Sprint der Läuferin. Und nehmen wir an, dass der Anfang ihrer Bewegung durch die Funktion $f(x) = 0.25 \cdot x^2$ gegeben ist, in der der Zeit die Position zugeordnet wird. Fixiert man etwa den Zeitpunkt $x_0 = 2$ und wählt $x_1 = 3$, so findet man

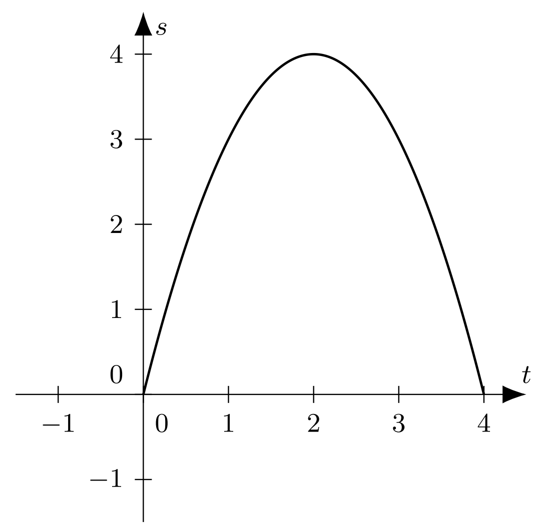
$$\begin{aligned}\frac{\Delta y}{\Delta x} &= \frac{f(x_1) - f(x_0)}{\Delta x} \\ &= \frac{0.25 \cdot 3^2 - 0.25 \cdot 2^2}{1} \\ &= 1.25\end{aligned}$$

Sie rennt also zwischen den Zeitpunkten 2 und 3 mit einer durchschnittlichen Geschwindigkeit von 1.25 LE pro ZE. Wählt man $x_1 = 2.001$, so findet man noch immer eine durchschnittliche Geschwindigkeit und zwar 1.00025 LE pro ZE. Und für $\Delta x \to 0$ beziehungsweise $x_1 \to 2$ wird man feststellen, dass $f'(2)$ definiert ist und den Wert 1 hat. Dabei ist es einerlei, ob man sich von links oder von rechts der Stelle 2 nähert. Die Bedeutung dieser Zahl 1 ist nun eine mehrfache: Es ist die Steigung der Tangente an den Graphen im Punkt $(2, f(2))$, und es ist gleichzeitig die momentane Geschwindigkeit des bewegten Objektes zum Zeitpunkt 2, je nachdem, worauf wir unser Augenmerk richten wollen.

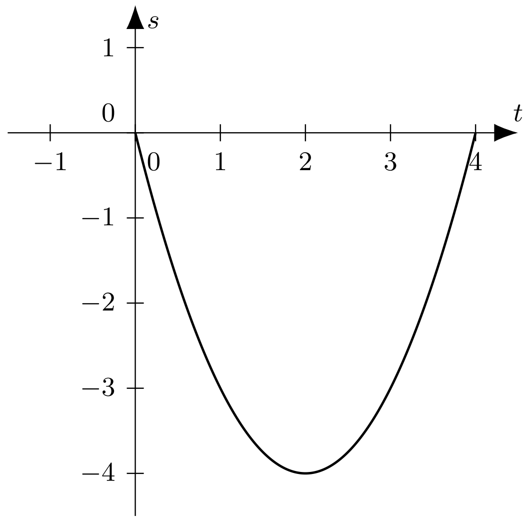
Angenommen, ein Objekt bewegt sich gemäss der Ort-Zeit-Funktion $f(x) = -10x + 90$. Fixieren wir $x_0 = 3$ und wählen etwa $x_1 = 5$, so erhalten wir

$$\begin{aligned}\frac{\Delta y}{\Delta x} &= \frac{(-10 \cdot 5 + 90) - (-10 \cdot 3 + 90)}{2} \\ &= -10.\end{aligned}$$

Eine negative Geschwindigkeit? Kann das sein? Nun ja, klar, wenn sich das Objekt rückwärts bewegt (relativ zu unserer gewählten Skala), dann wird Δy automatisch negativ sein. Daran dürfen wir uns nicht stören, im Gegenteil: Dass eine Geschwindigkeit positiv oder negativ sein kann, gibt uns einen weiteren wertvollen Hinweis über die Bewegung, nämlich über ihre Richtung. Bewegt sich das Objekt zum Beispiel gemäss



so bewegt es sich erst vorwärts und dann rückwärts, das heisst, die Geschwindigkeit (und damit die erste Ableitung) wird erst positiv und dann negativ sein. Bewegt es sich aber gemäss



so bewegt es sich erst rückwärts und dann vorwärts, das heisst, die Geschwindigkeit (und damit die erste Ableitung) wird erst negativ und dann positiv sein.

Im Alltag benutzt man kaum je negative Geschwindigkeiten, dort meint man eigentlich immer das Tempo (Schnelligkeit), wenn man über Geschwindigkeit spricht. In der Mathematik und den Naturwissenschaften hat Geschwindigkeit aber immer eine Richtung. Analog dazu wird auch in der englischen Sprache *speed* benutzt, wenn man nur über das Tempo spricht, und *velocity*, wenn die Geschwindigkeit eine Richtung hat.)

Die Funktion $f(x) = |x|$ ist an der Stelle $x_0 = 0$ nicht differenzierbar, weil der Grenzwert für $x_1 \to 0$ nicht existiert. An jeder anderen Stelle ist sie aber durchaus differenzierbar, denn an jeder anderen Stelle $x_0 \neq 0$ hat die Tangente an den Graphen entweder Steigung 1 oder Steigung -1, so dass wir also sagen können:

$$f'(x_0) = \begin{cases} 1 & \text{, falls } x_0 > 0 \\ -1 & \text{, falls } x_0 < 0. \end{cases}$$

Am Ende dieses Textes haben wir nun immerhin verstanden, dass die durchschnittliche Geschwindigkeit (oder allgemeiner: die durchschnittliche Änderungsrate pro ZE) durch eine Sekantensteigung und dass die momentane Geschwindigkeit (oder allgemeiner: die momentane Änderungsrate an einer Stelle) durch eine Tangentensteigung erfasst werden kann, und wir haben alle diese Konzepte formalisiert. Aber wir haben noch keine Ahnung, wie nun im konkreten Fall eine solche Tangentensteigung berechnet werden kann – oder wir unterschätzen vielleicht die Schwierigkeiten, die uns dabei noch begegnen werden.

Es gibt also noch viel zu tun: Wir werden lernen, wie Ableitungen skizziert und berechnet werden können, und wir werden eine riesige Zahl von Anwendungen erleben. Das alles wird uns ein immer klareres und beeindruckenderes Bild von dem ungeheuren Potenzial der Differentialrechnung geben.