Die Ableitung auf mehr als einem Weg

Reinhard Oldenburg, Dilan Bacaru, Sabrina Scheffler, Adrian Schlotterer

Angaben zur Veröffentlichung / Publication details:

Oldenburg, Reinhard, Dilan Bacaru, Sabrina Scheffler, and Adrian Schlotterer. 2018. "Die Ableitung auf mehr als einem Weg." Der Mathematikunterricht: MU - Beiträge zu seiner fachlichen und fachdidaktischen Gestaltung 64 (3): 3–15.

REINHARD OLDENBURG, DILAN BACARU, SABRINA SCHEFFLER ADRIAN SCHLOTTERER

Die Ableitung auf mehr als einem Weg

Die Ableitung ist ein zentraler Begriff der Analysis und lässt sich auf verschiedene Weisen definieren. Der Aufsatz untersucht einige davon in Hinblick auf ihre didaktische Eignung und Funktion.

Eine grobe Einteilung der Definitionsmöglichkeiten der Ableitung einer reell wertigen Funktion wird in (Greefrath et al. 2016) gegeben. Dort werden die beiden Aspekte der Ableitung als Grenzwert des Differenzenquotienten und als lokale Linearisierung erläutert. Die Umsetzung dieser Aspekte in konkrete Definitionen lässt noch allerlei Freiheiten. In diesem Aufsatz werden einige bekannte, weniger bekannte und neue Varianten der Ableitung vorgestellt.

1 Lokale Linearisierung

Wenn man mit einem Funktionsplotter auf dem Graphen einer beliebig komplizierten, aber differenzierbaren Funktion einen Punkt markiert und dessen Umgebung durch hineinzoomen vergrößert, landet man am Ende bei einem Bild, das von dem einer Geraden nicht zu unterscheiden ist. Dieses Phänomen ist sowohl eindrücklich als auch für die Anwendungen der Ableitung extrem wichtig. Dennoch spielt es bisher im Unterricht nur eine marginale Rolle. Henning (2018, in diesem Heft) zeigt einen erprobten Unterrichtsgang auf, der diese Idee ins Zentrum stellt. Hier werden einige andere Ideen dargestellt, gewissermaßen als Steinbruch für mögliche Unterrichtsversuche.

1.1 Lokale Linearisierung: „Stetige Steigung“

Auf den Mathematiker Caratheodory geht die Idee zurück, die Ableitung nicht aus dem Differenzenquotienten zu entwickeln, sondern eine Form zu wählen, die ohne Quotienten auskommt. Das vereinfacht einige Rechnungen und Beweise. Dies wurde kürzlich von Range wieder diskutiert (Range 2016). Eine didaktische Reflexion steht jedoch aus.

Die Grundidee

Caratheodory hat in seinem Buch zur Funktionentheorie (1961) eine Definition der Differenzierbarkeit gegeben, die ohne expliziten Grenzwert und ohne Differenzenquotient auskommt. Motivation kann aus der Betrachtung linearer Funktionen kommen, für die gilt, dass Änderungen der unabhängigen und der abhängigen Größe proportional sind, d.h. der Quotient $q = \frac{f(x)-f(x_0)}{x-x_0}$ ist konstant und kann als Steigung interpretiert werden. Es ist dann $f(x) - f(x_0) = q \cdot (x - x_0), \Delta y = q \cdot \Delta x$. Wenn die Funktion nicht linear ist, ist der Wert des Quotienten nicht konstant. Man kann daher auf die Idee kommen, an dieser Stelle eine Funktion zu verwenden. Wenn der Quotient nicht zu wild von der Stelle abhängt, sollte diese Funktion stetig sein. Genau das ist die Idee der Definition von Caratheodory, die leicht umformuliert lautet, dass man eine Funktion f in $x_0$differenzierbar nennt, wenn es eine stetige Funktion q auf einer offenen Umgebung von x_0 gibt, so dass $f(x) - f(x_0) = q(x) \cdot (x - x_0)$ auf der Umgebung gilt. Dann nennt man $q(x_0)$ die Ableitung an der Stelle x_0 und bezeichnet diese Zahl als $f'(x_0)$.

Diese Definition ist kurz und elegant, aber es ist auch klar, wo sie die Schwierigkeit versteckt: Satt der üblichen Grenzwertberechnung ist zu prüfen, ob es eine solche Funktion gibt. Wie man das macht, ist nicht offensichtlich. Immerhin ist aber für $x \neq x_0$ der Wert $q(x) = \frac{f(x)-f(x_0)}{x-x_0}$ eindeutig bestimmt, so dass die Frage nur ist, ob $q(x_0)$ so fest gesetzt werden kann, dass q in x_0 stetig ist. Dieser Funktionswert ergibt sich, wenn er denn existiert, aus dem Grenzwert für $x \to x_0$ von $q(x)$ für $x \neq x_0$. Dies schließt die Definition an die übliche an.

Die Funktion q hängt von x_0a b, es wäre deswegen angebracht, q_{x_0} zu schreiben, aber wir werden später einen Index an q noch für andere Zwecke benötigen.

Ein erster didaktischer Gewinn ist offensichtlich: Man muss nicht mit rationalen Termen arbeiten: Multiplizieren ist einfacher als Dividieren. Ein zweiter Gewinn ist von fragwürdiger Qualität: Der Grenzwertbegriff wird nicht explizit benötigt, aber evtl. möchte man dessen Bedeutung betonen statt ihn klein zu reden, und außerdem wird – als Ersatz für den Grenzwert – der Stetigkeitsbegriff benötigt.

Auf Ebene der Grundvorstellungen erkennt man neben der lokalen Linearität (linear wäre die Funktion, wenn q konstant wäre, da q immerhin stetig ist, ändern sich die Werte in der Umgebung nicht sprunghaft, lokal ist die Funktion also fast linear) sehr gut die des Verstärkungsfaktors: q ist eben der Faktor, der eine Veränderung $x - x_0$ verstärkt.

Stetigkeit

Wegen der Bedeutung der Stetigkeit in der Definition soll dieser Begriff etwas ausführlicher diskutiert werden. Die klassische intuitive Vorstellung von stetigen Funktionen ist die einer Funktion, deren Graph in einem Zug durchgezeichnet werden kann. Diese Vorstellung ist zwar am Rande fachlich unscharf, aber auf Schulniveau leicht zu erklären und kann benutzt werden, um stetige und unstetige Funktionen zu erkennen. Noch etwas enger ist die Auffassung, dass eine stetige Funktion eine ist, die keine Sprünge macht. Dies ist formal noch nicht sauber, aber diese Begriffsfassung reicht aus, um zu verstehen, dass Polynomfunktionen stetig sind und dass die üblichen Verknüpfungen stetiger Funktionen wieder stetig sind. Wenn man möchte, kann man die Vorstellung sogar zu einer fachlich korrekten Definition ausbauen: Eine Funktion f ist in x_0 stetig, falls ihre Funktionswerte in einer Umgebung von x_0 nicht stark von $f(x_0)$ abweichen, d.h. dass $|f(x) - f(x_0)| < \epsilon$ für jedes beliebig kleine ϵ, wenn nur x hinreichend dicht bei x_0 liegt.

Motivation aus Anwendungen

Kein anderer Begriff steht historisch und didaktisch so beispielhaft für Anwendungen der Differentiation wie die Geschwindigkeit. Üblicherweise wird sie als Verhältnis von Weg und Zeit definiert: $v = \frac{s}{t}$. Äquivalent dazu, aber evtl. etwas ungewohnt, ist es, zu sagen, dass die Geschwindigkeit der Faktor ist, der Zeit in Weg übersetzt: $s = v \cdot t$. Diese Formulierung hat aber durchaus Vorteile. Jetzt wird eine Bewegung zwischen den Zeitpunkten t_1, t_2 betrachtet, bei der die Position durch eine Funktion y gegeben ist. Dann ist also $s = y(t_2) -$

$y(t_1), t = t_2 - t_1$ und damit $v = \frac{y(t_2)-y(t_1)}{t_2-t_1}, y(t_2) - y(t_1) = v \cdot (t_2 - t_1)$. Die multiplikative Form ist auch noch für $t_1 = t_2$ sinnvoll. Bei beiden Formen ist es im Übrigen egal ob $t_1 < t_2$ oder $t_1 > t_2$, man erhält jeweils die mittlere Geschwindigkeit $v(t_{1,} t_2)$ in diesem Intervall. Intuitiv ist klar, dass man t_1 festhalten und t_2 gedanklich auf der Zeitachse an t_1 vorbeischieben kann, ohne dass bei der Geschwindigkeit ein Sprung auftritt, so dass auch $v(t_1, t_1)$ einen bestimmten Wert hat.
Ist etwa (als willkürliches Beispiel) $y(t) = t^2, t_1 = 1$, so findet man $y(t_2) - y(t_1) = t_2^2 - 1 = (t_2 + 1) \cdot (t_2 - 1) = v \cdot (t_2 - 1)$, also ist $v = t_1 + 1$ (hier wird alles ohne Einheiten gerechnet, Physiker mögen sie sich hinzudenken, sodass die Formeln zu $y(t) = a \cdot t^2, t_1 = 1s, v = a \cdot (t_1 + 1s)$) werden. Dann kann man auch die Momentangeschwindigkeit in $t_1 = t_2 = 1$ bestimmen: $v(1,1) = 2$.

Geometrisch-Algebraische Motivation

Als Beispielsfunktion dient hier $f(x) = x^2$ und man sucht nach einer Berechnungsmöglichkeit für Sekanten. Die Sekantensteigung setzt Änderungen in Beziehung. Für die Änderung der abhängigen Größe zwischen x, x_0 gilt $\Delta y = f(x) - f(x_0) = x^2 - x_0^2 = (x + x_0) \cdot (x - x_0) = (x + x_0) \cdot \Delta x$. Dies kann als Prototyp der oben verwendeten Struktur gelten.

Überblick über die Entwicklung der Analysis

Dieser Abschnitt entwickelt die Theorie in extrem knapper Form. Es geht darum, zu zeigen, wie Beweise laufen können, die Umsetzung in den Unterricht bleibt dem Leser überlassen.
Ausgehend von der Definition, dass die Funktion f in x_0 differenzierbar ist, falls $f(x) - f(x_0) = q(x) \cdot (x - x_0)$ mit einer stetigen (von x_0 abhängenden) Funktion q, definiert man die Ableitung an der Stelle x_0 als $f'(x_0) := q(x_0)$
Daraus ergeben sich relativ leicht viele Sätze:
Aus der Differenzierbarkeit von f folgt Stetigkeit, da die rechte Seite der definierenden Gleichung stetig ist.
Die Summe $u + v$ differenzierbarer Funktionen ist differenzierbar: $u(x) - u(x_0) = q_u(x) \cdot (x - x_0)$ und $v(x) - v(x_0) = q_v(x) \cdot (x - x_0)$ werden addiert zu
$u(x) + v(x) - u(x_0) - v(x_0) = q_u(x) \cdot (x - x_0) + q_v(x) \cdot (x - x_0) = (q_u(x) + q_v(x)) \cdot (x - x_0)$.
Also folgt $(u + v)'(x_0) = u'(x_0) + v'(x_0)$.
Die Konstantenregel folgt analog.
Für die quadratische Funktion $f(x) = x^2$ gilt: $x^2 - x_0^2 = (x + x_0) \cdot (x - x_0)$, also ist $q(x) = x + x_0$ und die Ableitung ist $f'(x_0) = q(x_0) = 2x_0$
Bei höheren Potenzen $f(x) = x^n$ gilt:
$x^n - x_0^n = (x^{n-1} + x^{n-2}x_0 + x^{n-3}x_0^2 + \ldots x_0^{n-1}) \cdot (x - x_0)$
also ist $q(x) = x^{n-1} + x^{n-2}x_0 + x^{n-3}x_0^2 + \ldots x_0^{n-1}, f'(x_0) = q(x_0) = n \cdot x^{n-1}$.
Für den Kehrwert rechnet man: $\frac{1}{x} - \frac{1}{x_0} = \frac{x_0 - x}{x x_0}$, also $q(x) = \frac{-1}{x x_0}$ und die Ableitung ist $\frac{-1}{x_0^2}$
Zur Produktregel:

$$u(x) \cdot v(x) - u(x_0) \cdot v(x_0) =$$
$$(u(x_0) + q_u(x) \cdot (x - x_0)) \cdot (v(x_0) + q_v(x) \cdot (x - x_0)) - u(x_0) \cdot v(x_0) =$$
$$u(x_0) \cdot q_v(x) \cdot (x - x_0) + q_u(x) \cdot (x - x_0) \cdot v(x_0) + q_u(x) \cdot q_v(x) \cdot (x - x_0)^2$$

Also ist für die Produktfunktion $q(x) = u(x_0) \cdot q_v(x) + q_u(x) \cdot v(x_0) + q_u(x) \cdot q_v(x) \cdot (x - x_0)$ und die Ableitung wie erwartet: $(uv)'(x_0) = q(x_0) = u(x_0) \cdot q_v(x_0) + q_u(x_0) \cdot v(x_0) = u(x_0) \cdot v'(x_0) + u'(x_0) \cdot v(x_0)$
Besonders einfach wird der Beweis der Kettenregel:
$u(v(x)) - u(v(x_0)) = q_u(v(x)) \cdot (v(x) - v(x_0)) = q_u(v(x)) \cdot q_v(x) \cdot (x - x_0)$, also ist $(u \circ v)'(x_0) = u'(v(x_0)) \cdot v'(x_0)$.
Die Quotientenregel erhält man entweder über Produkt-, Ketten- und Kehrwertregel, oder mit etwas Rechenaufwand:

$$\frac{u(x)}{v(x)} - \frac{u(x_0)}{v(x_0)} = \frac{u(x) \cdot v(x_0) - u(x_0) \cdot v(x)}{v(x) \cdot v(x_0)} =$$

$$\frac{(u(x_0) + q_u(x) \cdot (x - x_0)) \cdot v(x_0) - u(x_0) \cdot (v(x_0) + q_u(x) \cdot (x - x_0))}{v(x) \cdot v(x_0)} =$$

$$\frac{q_u(x) \cdot (x - x_0) \cdot v(x_0) - u(x_0) \cdot q_u(x) \cdot (x - x_0)}{v(x) \cdot v(x_0)} = \frac{q_u(x) \cdot v(x_0) - u(x_0) \cdot q_u(x)}{v(x) \cdot v(x_0)} \cdot (x - x_0)$$

Bei allen bisherigen Rechnungen waren keine Tricks nötig, die nicht algebraisch naheliegen. Bei der Wurzel geht es wohl aber nur mit Trick: $\sqrt{x} - \sqrt{x_0} = \frac{(\sqrt{x}-\sqrt{x_0})\cdot(\sqrt{x}+\sqrt{x_0})}{\sqrt{x}+\sqrt{x_0}} = \frac{x-x_0}{\sqrt{x}+\sqrt{x_0}}$, also ist hier $q(x) = \frac{1}{\sqrt{x}+\sqrt{x_0}}$ und die Ableitung $(\sqrt{x_0})' = q(x_0) = \frac{1}{2\sqrt{x_0}}$.
Auch bei transzendenten Funktion ergibt sich kein Vorteil:

$$\sin(x) - \sin(x_0) = 2\cos(\frac{x}{2} + \frac{x_0}{2}) \cdot \sin(\frac{x}{2} - \frac{x_0}{2}) = \cos(\frac{x}{2} + \frac{x_0}{2}) \cdot \frac{\sin(\frac{x}{2} - \frac{x_0}{2})}{\frac{x - x_0}{2}} \cdot (x - x_0)$$

Man sieht, dass hier immer noch der Grenzwert von $\frac{sin(x)}{x}$ für $x \to 0$ entscheidend ist und begründet werden muss. Man kann aber wie in (Greefrath et al, S. 188) argumentieren.
Auch bei den anderen transzendenten Funktionen ergibt sich keine Änderung.

Nach den üblichen Kalkülregeln werden nun einige Anwendungen der Differentiation besprochen. Die Monotoniesätze und die Sätze zu lokalen Extrema sind Anwendungen der Ableitung, die selbst wieder eine bekannte Fülle von weiteren Anwendungen auch in der realen Welt nach sich ziehen.
In einem lokalen Maximum x_0 ist die Ableitung 0: Wäre $q(x_0) > 0$, dann auch $q(x) > 0$ zumindest in einer gewissen Umgebung von x_0 (da q stetig) und deswegen ist für $x - x_0 > 0$ auch $f(x) - f(x_0) > 0$, also $f(x) > f(x_0)$, also wäre x_0 keine Maximalstelle. Analog argumentiert man bei $q(x_0) < 0$.

Ganz einfach argumentiert man bei der Monotonie. Ist f monoton wachsend, muss offensichtlich $q(x), x \neq x_0$ nicht-negative Werte annehmen, wegen der Stetigkeit kann auch in x_0 kein negativer Wert angenommen werden. Für die umgekehrte Richtung notieren wir den Entwicklungspunkt, d. h: $f(x) - f(x_0) = q_{x_0}(x) \cdot (x - x_0)$. Ist die Ableitung positiv, also $\forall x_0 \in [a, b]$: $f'(x_0) = q_{x_0}(x_0) > 0$, dann gibt es um jedes x_0 eine ganze Umgebung U_{x_0}, in der $q_{x_0}(x) > 0, x \in U_{x_0}$. In dieser Umgebung ist dann f streng monoton steigend. Da das für

alle x_0 gilt, ist die Funktion auf dem ganzen Intervall streng monoton wachsend. Die Komplexität dieses Argumentes stellt leider allenfalls einen kleinen Vorteil gegenüber dem üblichen Vorgehen dar.
Der Zugang kann an mehrere Grundvorstellungen zur Ableitung anknüpfen: Da q stetig ist, ist für $x \approx x_0$ auch $q(x) \approx q(x_0)$, also $f(x) - f(x_0) \approx f'(x_0) \cdot (x - x_0)$. Die Ableitung ist also ein Verstärkungsfaktor für kleine Änderungen und $f(x)$ kann linear approximiert werden: $f(x) \approx f(x_0) + f'(x_0) \cdot (x - x_0)$ (vgl. Greefrath et al. Abschnitt 4.2.4). Mit der in der Physik üblichen Notation für Änderungen schreibt man: $\Delta x = x - x_{0,} \Delta y = f(x) - f(x_0) = q(x) \cdot \Delta x$ und wenn man $q(x)$ durch $q(x_0)$ nähert, erhält man eine lineare Prognose für die Änderung, die man als Differential bezeichnet: $dy = q(x_0) \cdot dx$, vgl. (Oldenburg 2016). Wegen der Stetigkeit von q ist für kleine dx approximativ $dy \approx \Delta y$.

Integral

Dieser Abschnitt skizziert kurz, wie man zur Integration als Rekonstruktion in diesem Kontext kommen kann. Ausgangspunkt ist der Wunsch, die Änderung $f(x) - f(x_0)$ aus den Änderungsfaktoren auszurechnen. Die Beziehung $f(x) - f(x_0) = q(x) \cdot (x - x_0)$ gilt zwar allgemein, aber q an beliebigen Stellen auszurechnen ist nicht einfacher, als f an beliebigen Stellen zu ermitteln. Es hilft hier aber die Idee der lokalen linearen Approximierbarkeit, d. h. wenn x, x_0 nahe beieinander sind, ist der Wert von q in $f(x) - f(x_0) = q(x) \cdot (x - x_0)$ nahe bei $q(x_0)$. Deswegen zerkleinert man das Intervall $[x_0, x]$, sodass man auf den Teilintervallen q durch die Werte der Ableitungsfunktion nähern kann.
Da die Funktion q für verschiedenen Stellen benötigt wird, soll die Stelle hier als Index notiert werden, d.h.: $f(x) - f(x_0) = q_{x_0}(x) \cdot (x - x_0)$. Für drei Stellen x_0, x_1, x_2 ergibt sich:
$f(x_1) - f(x_0) = q_{x_0}(x_1) \cdot (x_1 - x_0)$ und $f(x_2) - f(x_1) = q_{x_1}(x_2) \cdot (x_2 - x_1)$.
Die Summe ist $f(x_2) - f(x_0) = q_{x_0}(x_1) \cdot (x_1 - x_0) + q_{x_1}(x_2) \cdot (x_2 - x_1)$. Man iteriert weiter:
$f(x_n) - f(x_0) = \sum_{i=0}^{n-1} q_{x_i}(x_{i+1}) \cdot (x_{i+1} - x_i) \approx \sum_{i=0}^{n-1} q_{x_i}(x_i) \cdot (x_{i+1} - x_i) = \sum_{i=0}^{n-1} f'(x_i) \cdot (x_{i+1} - x_i)$.
Der Näherungsfehler wird umso kleiner, je näher x_i, x_{i+1}zusammen liegen. Für $x_i = a + \frac{b-a}{n} \cdot i$ (also $x_0 = a, x_n = b$) ergibt sich die Näherung $f(b) - f(a) \approx \sum_{i=0}^{n-1} f'(x_i) \cdot \frac{b-a}{n}$ und diese wird offensichtlich besser, wenn man $n \to \infty$ betrachtet. Den Grenzwert der rechten Seite definiert man als Integral von f' (auch wenn f' nicht die Ableitung einer bekannten Funktion sein sollte - der Grenzwert existiert zumindest für beliebige stetige Funktionen). Wenn der Approximationsfehler im Grenzwert gegen 0 geht[1], hat man den Hauptsatz gefunden.

Fazit

[1]Dies ist einfach zu begründen: Der Approximationsfehler ist $\sum_{i=0}^{n-1}(q_{x_i}(x_{i+1}) - q_{x_i}(x_i)) \cdot \frac{b-a}{n}$ Zwar wächst die Zahl der Summanden mit n, aber der zweite Faktor ist proportional $1/n$, so dass sich dies kompensiert. Da der erste Faktor wegen der Stetigkeit gegen Null geht, gilt das auch für die Summe – der Grenzwert ist also 0.

Die Ableitungsdefinition als Caratheodory hat didaktisches Potenzial, auch wenn sie kein Wundermittel ist. Besonders hervorzuheben ist, dass von ihr aus sowohl die klassische Definition über den Grenzwert als auch die über die lineare Approximierbarkeit sich gut erschließt. Damit ist sie eine gute Basis, um alle Grundvorstellungen zur Ableitung aufzubauen.

1.2 Lokale Linearisierung: Das aktive Funktionenmikroskop

Differenzierbare Funktionen erscheinen lokal linear. Im sogenannten Funktionsmikroskop (auch Funktionenlupe) erscheint bei ausreichender Vergrößerung einer Stelle des Funktionsgraphens nahezu eine Gerade. Durch Zoomen in einem Funktionsplotter (z.B. GeoGebra) können Lernende dieses Phänomen erfahren, es scheint, dass dies keinen rechnerischen Weg eröffnet. Dem ist aber nicht so. Der Trick ist, dass man statt des passiven Zoomens (das Objekt bleibt unverändert, man betrachtet es nur in anderem Maßstab), aktiv zoomt, den Funktionsgraphen also zentrisch streckt.
Beim Funktionsmikroskop wird der Graph einer festen Funktion in verschiedenen Vergrößerungsstufen betrachtet. Diesen Gesichtspunkt dreht man jetzt um: Zur gegebenen Funktion f definiert man eine neue Funktion f_k, deren Graph durch zentrische Streckung mit Faktor k aus dem der Ausgangsfunktion entsteht.
<u>1. Zentrische Streckung in Koordinaten realisieren.</u>
Zentrische Streckungen am Ursprung sind einfach in Koordinaten zu beschreiben: $(x, y) \rightarrow (kx, ky)$. Etwas mühsamer ist es, um einen beliebigen Punkt (x_0, y_0) zu strecken. Dazu verschiebt man zunächst so, dass das Zentrum im Ursprung liegt $(x, y) \rightarrow (x - x_0, y - y_0)$, streckt $\rightarrow (k \cdot (x - x_0), k \cdot (y - y_0))$ und schiebt danach zurück $\rightarrow (x_0 + k \cdot (x - x_0), y_0 + k \cdot (y - y_0))$. Dass das das Richtige macht, lässt sich auch direkt in Geogebra studieren; Man erzeugt einen Punkt Z als Zentrum und einen abzubildenden Punkt A, sowie eine Zahl k als Streckfaktor. Der gestreckte Punkt wird dann in GeoGebra-Syntax als (x[Z]+k*(x[A]-x[Z]),y[Z]+k*(y[A]-y[Z])) definiert.
<u>2. Eine „gestreckte“ Funktion definieren.</u>
Es sei f eine Funktion. Man hält einen Punkt ihres Graphen $(x_0, y_0), y_0 = f(x_0)$ fest. Der Punkt $(x, f(x))$ des Graphen von f wird nun zentrisch mit diesem Zentrum gestreckt und liefert: $(x', y') := (x_0 + k \cdot (x - x_0), f(x_0) + k \cdot (f(x) - f(x_0)))$. Dieser Punkt soll auf dem Graphen der „gestreckten“ Funktion liegen, die definiert ist durch $y' = f_k(x')$. Um den Funktionsterm von f_k zu erhalten, muss noch das x durch $x = \frac{x' - x_0}{k} + x_0$ ersetzt werden, es entsteht: $y' = f_k(x') = f(x_0) + k \cdot (f(\frac{x' - x_0}{k} + x_0) - f(x_0))$
<u>3. Grenzwertbetrachtung</u>
Letzter Schritt ist die Betrachtung des Verhaltens der gestreckten Funktion für den Fall, dass der Streckfaktor gegen unendlich strebt. Dies ist allgemein nicht mehr so übersichtlich umsetzbar – deswegen eine Betrachtung am Beispiel: $y = f(x) = x^2$. Dann sind $x' = x_0 + k \cdot (x - x_0), y' = x_0^2 + k \cdot (x^2 - x_0^2)$. Der Zusammenhang zwischen x', y'ist:
$y' = x_0^2 + k \cdot ((\frac{x' - x_0}{k} + x_0)^2 - x_0^2) = 2x'x_0 + \frac{(x' - x_0)^2}{k}$.
Für große Vergrößerungsfaktoren k wird der zweite Summand sehr klein, sodass der Zusammenhang zwischen x', y' nahezu linear ist – mit Steigung $2x_0$.

Dass bei diesem Prozess die gewöhnliche Ableitung entsteht, kann leicht nachgerechnet werden. Dazu ersetzt man den Grenzwert $k \to \infty$ durch $h := \frac{x\prime - x_0}{k} \to 0$. Damit ist $f_k(x') = f(x_0) + \frac{1}{h} \cdot (f(x_0 + h) - f(x_0)) \cdot (x' - x_0)$ und der Grenzwert ist offensichtlich.
Zentrische Streckungen sind besonders einfach, wenn das Zentrum im Ursprung liegt. Man kann daher den Graphen entsprechend verschieben, so dass also $f(0) = 0$ ist. Dann gilt:
$x' = kx, y' = ky$ und $y' = ky = k \cdot f(x) = k \cdot f\left(\frac{x\prime}{k}\right)$
Differenzierbarkeit bedeutet, dass für sehr große k der Zusammenhang *y*-*x* proportional ist, also $\frac{y'}{x'}$ eine Konstante: $m = \lim\limits_{k\to\infty} \frac{y\prime}{x\prime} = \lim\limits_{k\to\infty} \frac{k\cdot f(x\prime/k)}{x\prime} = \lim\limits_{k\to\infty} \frac{f\left(\frac{x\prime}{k}\right) - f(0)}{\frac{x\prime}{k} - 0}$
Beispiel. Ableitung von $f(x) = x^3$ in 1. verschobene Funktion: $g(x) = (x+1)^3 - 1$.
$y' = k \cdot \left(\left(\frac{x\prime}{k} + 1\right)^3 - 1\right) = 3x' + \frac{3x\prime^2}{k} + \frac{x\prime^3}{k^2} \cong 3x'$, also $g'(0) = 3, f'(1) = 3$
Weiteres Beispiel. Ableitung von $f(x) = \sqrt{x}$ in 1. verschobene Funktion: $g(x) = \sqrt{x+1} - 1$.
$y' = k \cdot \left(\sqrt{\frac{x\prime}{k} + 1} - 1\right) = k\left(\sqrt{\frac{x\prime}{k} + 1} - 1\right)\frac{\sqrt{\frac{x\prime}{k}+1}+1}{\sqrt{\frac{x\prime}{k}+1}+1} = k\left(\frac{x\prime}{k} + 1 - 1\right)\frac{1}{\sqrt{\frac{x\prime}{k}+1}+1} = x'\frac{1}{\sqrt{\frac{x\prime}{k}+1}+1} \cong \frac{x\prime}{2}$, also $g'(0) = 1/2, f'(1) = 1/2$
Selbstverständlich sollte man das Zoomen des Funktionsgraphen mit dynamischer Geometrie visualisieren (siehe Abb.1).

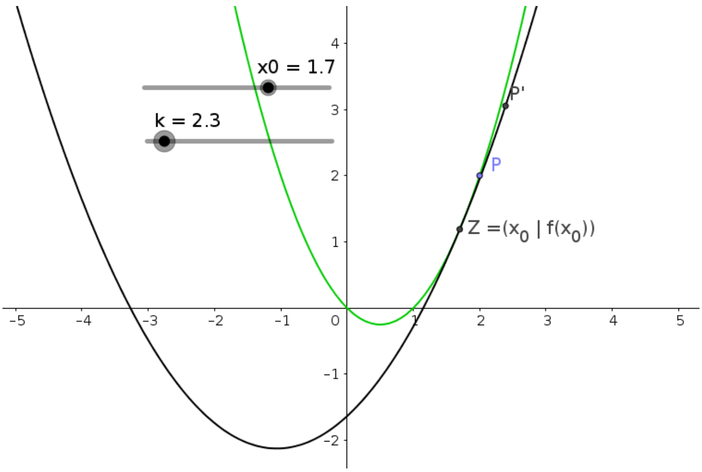

Abb. 1: Die Ausgangsparabel wird um den Punkt Z zentrisch gestreckt. Je größer man den Streckfaktor *k* wählt, desto mehr ähnelt der gestreckte Graph einer Geraden.

Fazit

Die aktive Funktionslupe ermöglicht den Brückenschlag von der intuitiven Idee, dass der Graph lokal gerade erscheint, hin zu einer quantitativen Erfassung des Phänomens. Der entscheidende Schritt ist, dass der Vergrößerungsfaktor quantifiziert werden kann, während der

Zoom-Faktor der Lupe nicht quantifiziert wird. Die Brücke ist allerdings noch nicht ideal. Sie führt auf Rechnungen, die einiges algebraisches Gespür erfordern.

1.3 Lokale Linearisierung: Approximationsgüte

Donald Knuth hat in einem nicht mehr online verfügbaren Leserbrief für den folgenden Weg geworben: Als erstes führt man die sogenannte A-Schreibweise ein, um eine beliebige Zahl zu bezeichnen, die vom Betrag her aber nicht größer als die angegebene sein darf, d.h. $A(5)$ steht für eine beliebige Zahl, deren Betrag höchstens 5 ist. Damit kann man Fehler angeben, z.B. $\pi = 3{,}14 + A(0.005)$, oder Abschätzungen angeben, z.B. $10^{A(2)} = A(200)$. Es gelten naheliegende Rechenregeln: $x, y \geq 0$: $A(x) + A(y) = A(x + y), A(x) \cdot A(y) = A(x \cdot y)$.
Ein paar weitere Beispiele: $\sin(x) = A(1); A(x) = x; A(x) = x \cdot A(1)$.
Das A fungiert gewissermaßen wie eine Gummizahl, die – im gesetzten Rahmen – immer passend gewählt ist.
Wenn es nicht auf die absolute Größe des Fehlers ankommt, sondern nur, wie schnell dieser wächst, kann man eine noch unspezifischere „Gummistelle" einführen: Es soll $O(x)$ stehen für $C \cdot A(x)$ mit einer beliebigen Zahl C. Dann ist z. B. klar, dass $5x^2 + 3x + 9 = O(x^2)$.
Anschaulich gesprochen bedeutet $f(x) = O(g(x))$, dass f nicht wesentlich schneller wächst als g, also höchstens um einen bestimmten Faktor. Anders ausgedrückt: Die Wertemenge von $\frac{f(x)}{g(x)}$ muss beschränkt bleiben – aber diese formalere Fassung braucht man hier gar nicht. Es gilt $O(f + g) = O(f) + O(g)$, und weil es auf Konstanten nicht ankommt, gilt z. B. $O(x^2) + O(3x^2) + O(x) = O(x^2)$.
Eine verschärfte Ableitungsdefinition wäre die folgende: $f'(x)$ ist eine Zahl, so dass $f(x + \epsilon) = f(x) + f'(x) \cdot \epsilon + O(\epsilon^2)$.
Diese Fassung erbt fast alle Vorteile der anderen Zugänge über lokale Linearisierung, wie sie sich auch bei Caratheodory gezeigt haben, d. h. Polynome lassen sich einfach ableiten, z.B. für $f(x) = x^2$: $f(x + \epsilon) = (x + \epsilon)^2 = x^2 + 2x\epsilon + O(\epsilon^2)$. Ketten- und Produktregel lassen sich einfach herleiten.
Beispiel: $g\big(f(x + \epsilon)\big) = g\big(f(x) + f'(x) \cdot \epsilon + O(\epsilon^2)\big) = g\big(f(x)\big) + g'\big(f(x)\big) \cdot \big(f'(x) \cdot \epsilon + O(\epsilon^2)\big) + O\left(\big(f'(x) \cdot \epsilon + O(\epsilon^2)\big)^2\right) = g\big(f(x)\big) + g'\big(f(x)\big) \cdot f'(x) \cdot \epsilon + O(\epsilon^2)$
Später, wenn der Hauptsatz der Differential- und Integralrechnung bekannt ist, kann die Bedeutung der zweiten Ableitung im Kontext der lokalen Linearisierung erarbeitet werden (Taylor-Approximation):

$$f(x + \epsilon) - f(x) = \int_0^\epsilon f'(x + t)dt = \int_0^\epsilon (f'(x) + f''(x) \cdot t + O(t^2))dt$$
$$= f'(x) + f''(x) \cdot \frac{\epsilon^2}{2} + O(\epsilon^3)$$

2 Differentiale

Die Schreibweise $f'(x) = \frac{dy}{dx}$ ist oft sehr suggestiv und intuitiv. Insbesondere beim Modellieren (dx ist ein kleiner Zuwachs von x) und bei formalen Rechnungen (etwa Kettenregel oder Substitution) sind Differentiale praktisch. In der Schule werden sie aber i.d.R. nicht als eigenständige Objekte eingeführt, sondern $\frac{dy}{dx}$ wird nicht als Bruch, sondern als rein formale Schreibweise bezeichnet. Wenn man den Differentialen dx, dy einzeln Sinn geben will, gibt es zwei Strategien: Sie können infinitesimal kleine Größen sein (Nichtstandardanalysis) oder endliche Größen, die sich auf die Linearisierung des funktionalen Zusammenhangs beziehen. Im zweiten Sinne wurde in (Oldenburg 2016a) ein Weg skizziert, Differentiale einzuführen. Mittlerweile wurde dieser Weg in der Schule erprobt und soll hier skizziert werden.

Zunächst wurden anhand realistischer Daten verschiedene Wachstumsprozesse betrachtet (linear, exponentiell, logistisch), die einen Zusammenhang $y = f(x)$ darstellen und es wurde versucht, aus den vorhandenen Daten (Wertetabelle und/oder Graph) Prognosen für Änderungen Δy zu erstellen, wenn sich die Unabhängige um Δx ändert. Hier konnte festgehalten werden, dass lokale Prognosen linear einfach sind. Eine lineare Fortsetzung eines Trends ist ein besonders einfaches Modell, leicht zu rechnen und lokal plausibel. Damit kann definiert werden: Das Differential dx einer Größe x ist die bei lokaler Betrachtung „bestmögliche" lineare Prognose der Änderung Δx. Für kleine Änderungen erwartet man, dass $\Delta x \approx dx$. Die Summenregel $d(u + v) = du + dv$ erarbeitet sich ganz natürlich. Verblüffend unproblematisch war auch die Bearbeitung der Produktregel: Bei der Änderung $u \to u + du, v \to v + dv$ ändert sich das Produkt zwar um $u \cdot dv + v \cdot du + du \cdot dv$, wie man leicht rechnet, aber dass der letzte Term besonders klein ist und als nichtlinearer Term die Struktur kompliziert macht, hat die Lernenden sofort überzeugt, im Sinne einer Modellbildung (lineares Modell) diesen Term wegzulassen. Es konnten damit dann auch Potenzen und Wurzeln abgeleitet werden und die Kettenregel ergab sich einfach aus dem Rechnen. Selbstverständlich ist eine spätere Exaktifizierung mit Grenzwerten sinnvoll, aber für das Modellieren und Argumentieren nicht notwendig. Wer das in (Oldenburg 2016a) skizzierte Programm im Unterricht durchführen möchte, kann Arbeitsblätter des Versuchs beim ersten Autor anfordern. Im ganzen Unterrichtsversuch gab es nur eine kritische Stelle: Beim Bestimmen von Tangenten an $y = x^2$ haben die Lernenden gerechnet $dy = 2xdx$ und daraus die Tangentensteigung $\frac{dy}{dx} = 2x$ bestimmt. Für die Stelle $x = 1$ wurde damit die Steigung in (1,1) berechnet und der Graph zeigte, wie schön das passt. Allerdings wandte eine Schülerin ein: „Wenn wir $x = 1$ setzen, dann ist x konstant, also $dx = 0$ und deswegen auch $dy = 0$". Eine angemessene Antwort ist leider nicht ganz einfach: Der Operator d wirkt nicht auf Zahlen, sondern auf Terme. Die Variablen mit Zahlen zu belegen ist eine spätere Entscheidung. So wie man beim Programmieren zwischen Programmierzeit und Laufzeit, oder bei PowerPoint zwischen Designzeit und Vorführungszeit unterscheidet, so sollte man auch in der Mathematik unterscheiden, zwischen Kalkülzeit und Zahlenrechenzeit. Mehr dazu findet sich in (Oldenburg, 2016b).

3 Differenzenquotient

Bisher dominieren in der Schule Zugänge zur Ableitung, bei der der Differenzenquotient im Fokus steht. Auch hier gibt es viele Variationen, die neue Optionen eröffnen.

3.1: Differenzenquotient: Term-Variationen

Bereits früher (Oldenburg 2005) wurde gezeigt, dass es zum Differenzenquotienten $\lim\limits_{h\to 0}\frac{f(x+h)-f(x)}{h}$ Alternativen gibt. Hier eine Liste mit einer Kurzcharakterisierung:

Ableitungsbegriff	Charakterisierung
$\lim\limits_{h\to 0}\dfrac{f(x+h)-f(x-h)}{2h}$	Beidseitige Ableitung. Quadratische Funktionen sind ohne Grenzwertberechnung ableitbar. Numerisch sehr gute Approximationen; aber: Auch die Betragsfunktion ist differenzierbar.
$\lim\limits_{q\to 1}\dfrac{f(qx)-f(x)}{qx-x}$	Die beiden Stellen werden multiplikativ unterschieden, nicht additiv. Außer für $x=0$ äquivalent zur Standarddefinition.
$\lim\limits_{q\to 1}\dfrac{f(qx)-f(x/q)}{qx-x/q}$	Eine Kombination aus beiden obigen Ideen.

Zum Vergleich der drei Definitionen zeigt die folgende Tabelle die Ableitung der quadratische Funktion mit $f(x)=x^2$:

Rechnung	Kommentar
$\lim\limits_{h\to 0}\dfrac{(x+h)^2-(x-h)^2}{2h}=\lim\limits_{h\to 0}\dfrac{4xh}{2h}=2x$	Selbst für $h>0$ ergibt schon der Differenzquotient den exakten Wert.
$\lim\limits_{q\to 1}\dfrac{(qx)^2-(x)^2}{qx-x}=x\cdot\lim\limits_{q\to 1}\dfrac{q^2-1}{q-1}=x\cdot\lim\limits_{q\to 1}q+1=2x$	Die funktionale Abhängigkeit (Ableitung ist linear) ist sofort klar.
$\lim\limits_{q\to 1}\frac{(qx)^2-\left(\frac{x}{q}\right)^2}{qx-\frac{x}{q}}=x\cdot\lim\limits_{q\to 1}\frac{q^2-q^{-2}}{q-q^{-1}}=x\cdot\lim\limits_{q\to 1}\frac{(q-q^{-1})\cdot(q+q^{-1})}{q-q^{-1}}=2x$	Die Vorteile kombinieren sich hier nicht.

Es lohnt (für die Lehrkraft, für Studenten, für Schüler) zu erforschen (das ist authentische Forschung), bei welchen Funktionen und Regeln die Arbeit jetzt leichter wird, wo schwerer. Bei der Quadratwurzel etwa hat die multiplikative Version einen starken Aufschlag:

$$\lim_{q\to 1}\frac{\sqrt{qx}-\sqrt{x}}{qx-x}=\frac{\sqrt{x}}{x}\cdot\lim_{q\to 1}\frac{\sqrt{q}-1}{q-1}=\frac{1}{\sqrt{x}}\cdot\lim_{q\to 1}\frac{1}{\sqrt{q}+1}=\frac{1}{2\sqrt{x}}$$

Wenn man mit all diesen Definitionen die Ableitung der üblichen Funktionen sowie die Kalkülregeln beweist, stellt man fest, dass einige Rechnungen einfacher gehen, andere werden aufwändiger. Beispielsweise erleichtert die multiplikative Version die Ableitung des Logarithmus, während sie die der Exponentialfunktion erschwert. Dies erklärt sich natürlich leicht durch die Verträglichkeit von Logarithmus bzw. Exponentialfunktion mit Addition und Multiplikation – damit wird eine Reflexion der Eigenschaften der Funktionsklassen angeregt.

Neben der Ermöglichung von forschendem Lernen (Begriffe erkunden und Beziehungen zwischen ihnen herstellen) stellen diese Varianten ein Feld zum wiederholenden Üben dar: Die erste Herleitung der Ableitung von Polynomen ist zu schwer, um sie selbsttätig zu machen, aber die Verifikation mit einer zweiten Herleitung auf Basis einer der Variationen durchzuspielen, ist deutlich einfacher.

3.2: Differenzenquotient: Grenzwert später

Etwas karikiert könnte man sagen, dass der Unterricht oft schnell zum Differenzenquotient kommt, kurz über den Grenzwert spricht und dann ganz lange die Regeln des Differentialkalküls anwendet ohne die ersten beiden Konzepte noch groß zu verwenden. Man könnte überlegen, ob man den ersten Schritt nicht wesentlich verlängern könnte, um die Motivation für den Grenzwert zu steigern. Das bedeutet, dass man lange mit diskreten Strukturen arbeitet, dass Approximationen mit teilweise linearen Abschnitten an Bedeutung gewinnen.

Das zentrale Ziel der Analysis ist die Beschreibung von Änderungsprozessen im Kleinen, also die Veränderung von Funktionswerten lokal. Lokale Sichtweisen sind über $\mathbb{Z}$ nur eingeschränkt, über $\mathbb{Q}$ schon besser und über $\mathbb{R}$ gut deutbar. Eine wichtige Transformation im Laufe der schulischen Mathematik ist also der Übergang diskret → kontinuierlich. Wertetabellen geben zunächst diskrete Zuordnungen, ihre Graphen geben durch glatte Verbindung die Idee einer kontinuierlichen Zuordnung. Experimente liefern oft nur diskrete Messwerte und SuS neigen dazu, diese als Polygonzug zu verbinden. Guter Unterricht reflektiert Sinn und Unsinn des Verbindens im jeweiligen Kontext. SuS neigen auch dazu, die Graphen von Parabeln mit einer Spitze zu zeichnen, wenn sie von einer Wertetabelle ausgehen. Funktionsplotter können das Phänomen beheben – SuS sehen die Rundung und zeichnen sie – aber es ist die Frage, ob das ein oberflächlicher Erfolg ist, ob verstanden wird, warum das so ist. Liegt hier nicht eine Chance? Sollte man der diskreten Phase nicht mehr Aufmerksamkeit schenken? Evtl. sollte man den Graphen einer Funktion bewusst als Polygonzug mit verschiedenen Schrittweiten zeichnen und so eine Approximation an eine glatte Kurve erleben (in Fortsetzung der Kreisapproximation nach Archimedes). Man sieht etwa, dass die Steigung des Segmentes zu x^2, das in (0,0) beginnt, immer kleiner wird, je kleiner die Schrittweite ist (warum ist das so…. Was bedeutet es?). Damit ist auch die Idee der lokalen linearen Approximation angeregt. Einen Polygonzug zu einem Graphen kann man betrachten:

- Ganz klein: Man sieht nur $(x_0|f(x_0))$: Funktionswert. 0. Näherung
- Umgebung betrachten: Man sieht zwei Punkte, ein Segment, dem man eine Steigung zuordnen kann. 1. Näherung
- Etwas größere Umgebung betrachten: Drei Punkte, zwei Segmente. Dem kann ein Winkel dazwischen zugeordnet werden (Krümmung). 2. NäherungGeschlossene Polygonzüge sind praxisrelevant und die Gaußsche Formel kann so geschrieben werden, dass man die Struktur des Integrals sofort sieht (s. S. 231 Greefrath et al.).

Nach diesen allgemeinen Überlegungen zur Approximation mit Polygonzügen beschränken wir uns jetzt auf den Fall der ersten Näherung, also auf die Approximation einer Funktion durch eine Sekante durch zwei Stellen. Viele Anwendungen liefern Motivation für eine Definition der folgenden Art: Die Steigung eines linearen Segmentes eines Funktionsgraphen ist durch zwei Punkte $\big(x, f(x)\big), (x + \Delta x, f(x + \Delta x)$ festgelegt und man definiert:

Differenzenquotient $\Delta(f, x, \Delta x) := \frac{f(x+\Delta x)-f(x)}{\Delta x}$.

Die neue Idee dieses Abschnitts ist, jetzt nicht schnell zum Grenzwert zu kommen, sondern zunächst mit diesem Term viel anzufangen. Sind etwa Messwerte eines Bewegungsvorgangs als Funktion der Zeit x gegeben, so liefert die Messfrequenz automatisch ein passendes Δx. Generell kann man viel auf dieser Ebene machen, insbesondere auch bei Computereinsatz. Mit einem Computeralgebrasystem kann die Wirkung des Differenzenquotienten erforscht

werden. Für kleine Δx ergeben sich etwa Funktionen, die optisch der Ableitung sehr nahe kommen. Auch die algebraische Betrachtung zeigt einige interessante Gesetzmäßigkeiten.
Mit einem Computeralgebrasystem wie Maxima kann man eine kleine Zahl festlegen, etwa `dx:0.001`, und dann den Differenzenquotienten definieren als `diffquot(f,x):= (subst(x=x+dx, f)-f)/dx`. Danach zeigt der Graph von `diffquot(x^3,x)` eine Parabel (approximativ) und der Term dazu ist sehr spannend: $3x^2 + 0.003x + 10^{-6}$. Ähnlich interessant ist das mit `trigexpand` vereinfachte Ergebnis von:
`diffquot(sin(x),x)` nämlich $0.99999983\, cos(x) - 4.99999958 \cdot\ 10^{-4}\, sin(x)$.
An dieser Stelle lässt sich der Grenzwert motivieren, aber wir gehen noch etwas weiter ohne Grenzwert vor und gewinnen Kalkülregeln für den Differenzenquotienten:

$$\Delta(c \cdot f, x, \Delta x) = c \cdot \Delta(f, x, \Delta x), \Delta(f + g, x, \Delta x) = \Delta(f, x, \Delta x) + \Delta(g, x, \Delta x),$$

$$\Delta(f \circ g, x, \Delta x) = \Delta\big(f, g(x), g(x + \Delta x) - g(x)\big) \cdot \Delta(g, x, \Delta x),$$

$$\Delta(f \cdot g, x, \Delta x) = \Delta(f, x, \Delta x) \cdot g(x + \Delta\, x) + f(x) \cdot \Delta(g, x, \Delta x)$$

$$\Delta(1/f, x, \Delta x) = \frac{-\Delta(f, x, \Delta x)}{f(x) \cdot f(x + \Delta x)}$$

Außerdem $\Delta(x \mapsto x^2, x, \Delta x) = 2x + \Delta x$ und $\Delta(x \mapsto x^3, x, \Delta x) = 3x^2 + 3x\Delta x + \Delta x^2$ usw.
Der Zusammenhang mit der Monotonie kann auch festgehalten werden: Offensichtlich gilt, dass bei einer auf $[a, b]$ monoton wachsenden Funktion gilt, dass für alle $x \in [a, b]$ und alle Δx mit $x + \Delta x \in [a, b]$ gilt: $\Delta(f, x, \Delta x) \geq 0$ und diese Implikation ist sogar umkehrbar. Damit erschließt sich auch die Bestimmung von Extremstellen. Dies wird umso genauer, je kleiner Δx. Der Wunsch, Δx klein zu haben, kann auch aus physikalischen Anwendungen motiviert werden. Der Grenzwert $\Delta x \to 0$ vereinfacht einige der Kalkülregeln und schließlich löst er ein ästhetisches Problem: Die obige Produktregel ist auf der rechten Seite nicht symmetrisch zwischen den beiden Funktionen obwohl das Produkt kommutativ ist.
Nach einiger Zeit der Beschäftigung sollte sich daher eine fundamentale Einsicht durchsetzen: Der Grenzwert vereinfacht eine Reihe von Regeln! Die Motivation dafür dürfte höher sein, wenn man vorher ein paar der mühseligen Rechnungen der obigen Art gemacht hat.

4 Gesamtfazit

Die Berechnung von Ableitungen ist reicher als „Limes h gegen Null". Der vorliegende Aufsatz kann einige Ideen liefern und diese können für alternative Einstiege verwendet werden oder für die nachträgliche Konstruktionen einer parallel-Theorie, evtl. dann stärker durch Schüler entdeckend. Deswegen lohnt es sich, wenn man als Lehrkraft diese Ideen kennt, um sie bei Bedarf einsetzen zu können.

Literatur

Caratheodory, C. [1961]: Funktionentheorie. Heidelberg: Springer.
Greefrath, G., Oldenburg R., Siller, H.-St., Ulm, V., Weigand, H.-G. [2016]: Didaktik der Analysis. Heidelberg: Springer,
Oldenburg, R. [2016a]: Differentiale als Prognosen. Eine Grundvorstellung als Ausgangspunkt analytischer Begriffsbildung. Journal für Mathematik-Didaktik 37, 1, S. 55-82.
Oldenburg, R. [2016b]. Algebra at the Meta and the Object Level. Turkish Journal of Computer and Mathematics Education, Vol 6, No 3.
Oldenburg, R. [2005]: Ableitungen alternativ berechnen, *MNU* 58, 343-344.
Range, R. Michael [2016]: What is calculus? From simple algebra to deep analysis. Hackensack, NJ: World Scientific.