Differentiale als Prognosen: eine Grundvorstellung als Ausgangspunkt analytischer Begriffsbildung

Reinhard Oldenburg

Angaben zur Veröffentlichung / Publication details:

Oldenburg, Reinhard. 2016. "Differentiale als Prognosen: eine Grundvorstellung als Ausgangspunkt analytischer Begriffsbildung." *Journal für Mathematik-Didaktik* 37 (1): 55–82. https://doi.org/10.1007/s13138-016-0096-2.

Differentiale als Prognosen

Eine Grundvorstellung als Ausgangspunkt analytischer Begriffsbildung

Reinhard Oldenburg

Mail: reinhard.oldenburg@math.uni-augsburg.de, Telefon: +49 821 598 2494, Fax: +49 821 598 2278

Institut für Mathematik, Universitätsstraße 14, 86159 Augsburg

Differentiale drücken die Ableitungs-Vorstellung der lokalen Linearisierung in anschaulicher Art aus. Es wird eine intuitive Definition für Differentiale vorgeschlagen und diese wird in Beziehung zur didaktischen und fachwissenschaftlichen Literatur gesetzt. Weiter wird skizziert, wie sich die Analysis auf Basis dieser Definition entwickeln lässt und wie ein darauf aufbauender Unterrichtsgang aussehen könnte. Statt Änderungsraten werden hier Änderungen zum zentralen Beschreibungs- wie Argumentationsmittel.

Schlüsselwörter: Analysis, Differentiale, Modellbildung

Differentials as Prognoses

A basic idea as starting point of concept development in calculus

Differentials are means to express the idea of linearization in calculus. The article gives an intuitive definition of differential and discusses its relationship to the didactical and mathematical literature. It is shown how calculus can be derived from this definition and what the structure of a possible teaching sequence may be that uses differentials from the beginning and thus puts the focus on the concept of change, not on rates of change.

Key words: calculus, differentials, modelling

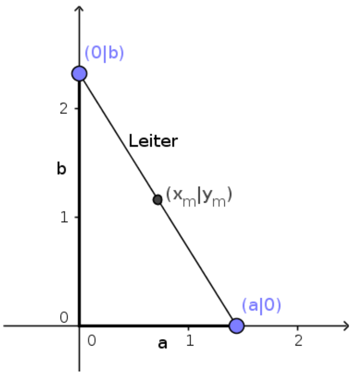
MESC: D30, D80, I40

1 Zum Stand der Analysis in der Schule

Gegenwärtig werden im Analysisunterricht der gymnasialen Oberstufe vor allem die Grundvorstellungen der lokalen Änderungsrate und der Tangentensteigung gefördert. Die Vorstellung der lokalen Linearisierung spielt kaum eine Rolle (KMK 2012, S. 19-20). Nach einer Phase der Strenge in der Schulanalysis der 1970er Jahre (Tietze 2000, S.218ff) haben sich der Formalisierungsgrad und die Anbindung an fachwissenschaftliche Strukturen in diesem Gebiet gelockert. Die Bildungsstandards für das Abitur (KMK 2012) erwähnen beispielsweise die Begriffe Folgen und Reihen nicht, es wird nur gefordert „Grenzwerte auf der Grundlage eines propädeutischen Grenzwertbegriffs insbesondere bei der Bestimmung von Ableitung und Integral [zu] nutzen“, (ebda. S. 18). Häufig fällt die Behandlung des Grenzwertkonzeptes und das der Ableitung in Eins: In den meisten Bundesländern kommen Grenzwertfragestellungen erstmals bei dem als Änderungsrate interpretierten Differenzenquotienten vor (so etwa Götz et al. 2009) und führen dann direkt zur Ableitung. Der Vorschlag von Blum und Kirsch (1979) zur Reduktion der Strenge durch die Abkehr von einer an der Hochschulmathematik angelehnten Formalisierung, wie sie im Rahmen der sogenannten Neuen Mathematik praktiziert wurde, hat eine Richtung vorgegeben, in die inzwischen sehr weit gegangen wird, so dass heute weitgehend informell gearbeitet wird. Dies geht teilweise deutlich über die Intentionen von Blum&Kirsch hinaus, die keineswegs beim rein numerischen Herantasten oder dem Konstatieren graphischer Evidenz stehen bleiben wollten, sondern z.B. für eine operative Annäherung an den Grenzwert plädiert haben (also zu einer vorgegebener maximalen Abweichung zu bestimmen, wie dicht man herangehen muss, um diese einzuhalten). Der Trend zur Entformalisierung zeigt sich deutlich beim Vergleich verschiedener Auflagen des gleichen Schulbuchs. Formalisierung ist sicher kein eigenständiges Ziel des Mathematikunterrichts – anders etwa als die Kompetenz des Argumentierens. Aber der richtige Grad an Formalisierung kann gerade das Argumentieren befördern, während eine zu geringe Formalisierung zu unnötig unpräzisen Argumenten führt und ein zu hoher Grad die Argumentation eher behindert.

Die Bildungsstandards (KMK 2012, S. 19-20) geben der Idee der Änderungsrate hohes Gewicht und diese ist in den Schulbüchern seit einigen Jahren weit verbreitet (beispielsweise Götz et al. 2009, Lergenmüller und Schmidt 2009). Es zeigt sich allerdings, dass das Konzept der Änderungsrate nicht optimal bei den Lernenden ankommt (Feudel 2015). Nur wenige Abiturienten sind demnach in der Lage, den Begriff der Ableitung im Sinne einer Änderungsrate zu erklären. Es stellt sich die Frage, ob das Konzept der Änderungsrate und damit zusammenhängend die Arbeit mit Differenzen- und Differentialquotient ein vernetzendes, kumulatives Lernen begünstigt oder teilweise durch fehlende Anknüpfungspunkte erschwert. Eine weitere Hypothese dieser Arbeit ist, dass die algebraische Struktur von Änderungsraten (nämlich Bruchterme) vielen Lernenden nicht hinreichend vertraut ist, um Änderungsraten interpretieren zu können. Deswegen wird in diesem Aufsatz vorgeschlagen, zunächst Änderungen statt Änderungsraten zu betrachten.

Die von vielen Kolleginnen und Kollegen subjektiv wahrgenommene, wenn auch schwer objektiv messbare Reduktion von Formalisierungsgrad und Strenge konnte nach ebenso subjektiver Einschätzung nur unzureichend durch einen Gewinn an tragfähigen Vorstellungen, auf deren Basis argumentiert und modelliert werden kann, kompensiert werden. In diesem Beitrag soll gezeigt werden, wie sich durch eine auf Differentialen basierende Konzeptualisierung der Analysis möglicherweise eine Verbesserung in Bezug auf die allgemeinen mathematischen Konzepte erzielen lässt. Dabei soll gezeigt werden, dass eine Formulierung der Analysis mit Differentialen eine hervorragende Sprache bereitstellt, um anschaulich zu argumentieren und authentische Anwendungsgebiete zu behandeln. Differentiale dx, dy... wie sie schon Leibniz eingeführt hat, haben in der Fachwissenschaft diverse Bedeutungen. Es gibt verschiedene Realisierungen – das wird zu diskutieren sein. Zur Motivation sei zunächst angeführt, dass Freudenthal (Mathematik als pädagogische Aufgabe, Bd. 2) die Rolle von Differentialen für das Verständnis der Analysis gewürdigt hat und eine Reihe von Anwendungen dargestellt hat („Wenn man Realdaten mittels der Analysis mathematisiert, melden sich die Differentiale von selber.“, Freudenthal 1979, S. 501). Als illustrierendes Beispiel gebe ich das folgende klassische Beispiel an (Abbildung 1): Eine Leiter der Länge l lehnt an einer senkrechten Hauswand (das kartesische Koordinatensystem ist naheliegend in den Fußpunkt der Wand gelegt). Die Entfernungen der Leiterenden seien a von der Hauswand und b vom Boden. Dann gilt $a^2+b^2=l^2, a>0, b>0$. Übergang zu Differentialen (die man an dieser Stelle verstehen sollte als linearisierte Änderungen oder wahlweise als kleine Änderungen) mittels der Rechenregeln $d(x+y)=dx+dy, d(x^2)=2xdx, d(\text{konst})=0$ liefert $a\cdot da+b\cdot db=0$. Diese Beziehung lässt sich gut interpretieren: Falls die Leiter wegrutscht ($da>0$), bewegt sich ihr oberer Endpunkt nach unten: $db=-\frac{a}{b}\cdot da<0$, und zwar besonders schnell, wenn die Leiter schon weit weg ist, also a groß und b klein ist.



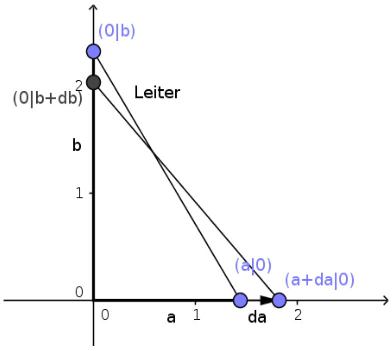


Abbildung 1: Die klassische Aufgabe der rutschenden Leiter

Die Leistung der Differentiale sowohl für das Verständnis von Konzepten der Analysis wie für die Modellbildung besteht darin, *Änderungen* beschreiben zu können. In der Schulanalysis werden diese über Änderungsraten beschrieben, aber solche Quotienten von Änderungen sind schwerer verständlich als Änderungen selbst – dies folgt schon aus den vielfältigen Feststellungen zur Schwierigkeit der Bruchrechnung. Interessanterweise lautet der Titel der wichtigen Dissertation von Hahn (2008) „Bestand und Änderung", obwohl sie von Problemen mit Änderungsraten handelt. Selbstverständlich steht die Bedeutung von Änderungsraten außer Frage. Eine Hypothese dieses Beitrags ist aber, dass es konzeptuell einfacher ist, erst Änderungen und dann Änderungsraten zu betrachten.

Ziel dieses Aufsatzes ist es, durch eine stoffdidaktische Analyse zu zeigen, dass die Verwendung von Differentialen einen Analysisunterricht ermöglichen könnte, der stärker verständnisorientiert ist. Der Begriff „Stoffdidaktik" wird von verschiedenen Autoren sehr unterschiedlich interpretiert (siehe Sträßer 2014) und diese Diskussion kann hier nicht in einer angemessenen Breite geführt werden. Für mich ist Gegenstand der Stoffdidaktik eine Strukturierung mathematischer Inhalte zum Zwecke der Lehre. Dies umfasst Elementarisierung, Explikation und Transformation. Methoden der Stoffdidaktik sind die der Mathematik, der Epistemologie sowie spezielle, lokale Methoden wie das didaktische Gedankenexperiment, bei dem eine Unterrichtseinheit durchdacht und geplant wird, ohne ausgeführt werden zu können. Dieser Methodik folgt der vorliegende Aufsatz und schließt sich daher eher an Kirsch (1977) als an Steinbring (2011) an. Kirsch beschreibt das Zugänglichmachen eines mathematischen Sachverhalts als transformatorischen Prozess. Dies liegt quer zum Bild der Position von Steinbring, nach der es einen unveränderlichen Gegenstand der Stoffdidaktik gibt. Im Gegensatz dazu gehe ich davon aus, dass mathematische Begriffe hochgradig dynamisch sind (vgl. etwa Lakatos 1979), und dass didaktische Überlegungen den genetischen Prozess mitgestalten. Dass mathematische Sätze eine über Jahrtausende hinweg eine stabile Form haben können und unendlich gültig erscheinen, wird auch in Theorien plausibel erklärt, die Mathematik als solche als revidierbar und dynamisch anerkennen, beispielsweise der Theorie der naturalisierten Erkenntnistheorie (Quine 1951). Dort sind Sätze der Mathematik prinzipiell genauso revidierbar wie solche der empirischen Wissenschaften. Dies passiert allerdings seltener, da das pragmatische Prinzip der „minimal mutilation" Veränderungen in der Mathematik i.d.R. als unökonomisch erscheinen lässt, da in der Mathematik kleine Änderungen oft weitreichende Konsequenzen im Gesamtgebäude des Denkens hätten. Dieser Aufsatz zeigt, dass eine Konzeptualisierung der Analysis mit Differentialen nur zu lokalen Änderungen führt – es gibt also innerhalb der Mathematik ausreichend Freiheit für die hier vorgeschlagene Sicht auf Differentiale.

2 Fachliche und didaktische Einordnung

Dieser Aufsatz schlägt vor, auf intuitiver Ebene Differentiale als lineare Prognosen für Änderungen zu verstehen. Genauer soll folgende intuitive Definition angestrebt werden: Das Differential dx einer Größe x ist ebenfalls eine Größe[1], nämlich die bei lokaler Betrachtung „bestmögliche" lineare Prognose der Änderung Δx. Für kleine Änderungen erwartet man, dass $\Delta x \approx dx$. Im dritten Kapitel wird dies ausgeführt. Insbesondere wird dort aufgezeigt, dass aus diesem Verständnis der Änderungsprognosen folgt, dass zur Funktionsgleichung $y = f(x)$ die Beziehung $dy = f'(x)dx$ zwischen den Differentialen gehört, wobei $f'(x)$ die übliche Ableitung ist.

Dies ist eine Definition auf einer informellen Ebene. Es wird zu zeigen sein:

- Es ist plausibel, dass man eine solche Definition mit Schülerinnen und Schülern erarbeiten kann. (Kapitel 3)
- Es ist möglich, Argumentationen auf einer solchen Definitionsbasis zu führen. (Kapitel 3)

1 Was Größen in diesem Zusammenhang sind, wird in Abschnitt 2.1 diskutiert. Insbesondere sind auch Zahlen Größen, das Differential kann also eine Zahl sein – folglich ist es nicht als infinitesimal kleines Objekt zu denken.

- Eine fachliche korrekte Exaktifizierung und Anknüpfung an fachliche Definitionen von Differentialen ist möglich (Kapitel 2)

Bevor diese Herausforderungen in Angriff genommen werden, erfolgt eine Einordnung in die mathematische und didaktische Literatur zu Differentialen. Erster Schritt dazu ist, das Verständnis des Begriffs ‚Größe' zu klären, der in der obigen Definition des Differentials verwendet wird.

2.1 Größen und Funktionen

Der Begriff der Größe spielt im vorgeschlagenen Unterrichtsgang eine zentrale Rolle und ist deswegen genauer zu explizieren. Mit dem Begriff Größe werden üblicherweise zwei Aspekte verbunden:

a) Eine Größe ist ein Wert, der eine Dimension besitzt, beispielsweise eine Länge oder eine Temperatur.
b) Eine Größe kann mit der Zeit, ggf. auch mit dem Ort, variieren.

Der erste Aspekt wird sehr häufig mit Größen assoziiert. Für die Behandlung der Differentiale ist er von untergeordneter Bedeutung. Ein angemessenes Verständnis des zweiten Aspekts dagegen ist wichtig für die Arbeit mit Differentialen.

Zum ersten Aspekt: Kirsch (1997, S. 53ff) scheidet sorgfältig Zahlen von Größen und definiert letztere als Elemente eines Größenbereichs, für den bestimmte Axiome gelten. Diese Axiome passen zwar für geometrische Längen und Flächeninhalte, schließen aber die meisten Größen der Physik wie Ort, Geschwindigkeit, Beschleunigung, Ladung, Temperatur, Impuls, Drehimpuls, Magnetisierung u.v.m. aus. Da Differentiale gerade in Hinblick auf Modellbildungsprozesse interessant sind, kann diese Theorie also nicht zugrunde gelegt werden. Insbesondere sollte für Größen in meinem Sinne anders als bei Kirsch gelten, dass die Änderung einer Größe ebenfalls eine Größe ist.
Für mich ist eine Größe eine Variable, die Werte in einem Modul oder Vektorraum annimmt (und dieses Modul oder dieser Vektorraum ist, was künftig als Größenbereich bezeichnet werden soll). Einheiten von Längen wie cm und m sind schlicht unterschiedlich lange Basisvektoren der Dimension ‚Länge'. Verhältnisse von Größen gleicher Dimension sind Zahlen. Man könnte die (reellen) Zahlen damit als eine Abstraktion von Größen sehen, deren Gemeinsamkeit durch Abstraktion von der Einheit gesehen wird. Da $1 \in \mathbb{R}$ ebenfalls als Einheit gesehen werden kann, sind auch alle reelle Zahlen Größen in diesem, hier verwendeten Sinne. Allgemeiner kann man Bijektionen zwischen Vektorräumen und Größenbereichen angeben. Wenn man bereit ist, bei allen Größen, über die man Aussagen machen will, diese Bijektionen (also im Wesentlichen die Wahl einer Einheit) anzugeben, kann man sich auf das Sprechen über Zahlen beschränken. Die in diesem Aufsatz betrachteten Differentiale sind Variablen für Größen in diesem Sinne, d.h. sie sind Zahlen oder Längen, Massen o.ä.
Zum zweiten Aspekt: Größen wurden historisch (etwa bei Newton) oft als implizite Funktionen der Zeit (und ggf. des Ortes) gesehen. Solche Größen sind also im Grunde Funktionen. Dies zeigt sich beispielsweise dann, wenn von einer stetigen Größe gesprochen wird. Das Sprechen über Größen in diesem Sinne bringt die Gefahr mit sich, nicht zwischen der Funktion und ihrem Wert zu unterscheiden. Der logisch unklare Status führte zu vielen Problemen. Epple schreibt dazu „Technische Probleme mit dem Beweis grundlegender Sätze über Grenzprozesse (…) führten immer wieder auf dieselbe Grundfrage: Wie sollte (…) der Bereich der reellen Größen technisch gefasst werden?" (Epple 1999, S, 371). Sprechweisen wie „x ist stetig", „x wächst" etc. sind sinnvoll, wenn man x als eine Funktion betrachtet, aber nicht, wenn x für eine Zahl steht. Veränderliche gibt es in der Logik nicht (Quine 1960).
Das Verständnis von Differentialen in diesem Aufsatz ist verträglich mit diesem logischen Ansatz: Größen und ihre Veränderungen werden in der Sprache der Prädikatenlogik mit den reellen Zahlen (bzw. einem Größenbereich) als Gegenstandsbereich beschrieben durch drei eigenständige

(beispielsweise reelle) Variablen $x, \Delta x, dx \in \mathbb{R}$, die sich nur durch die Konvention unterscheiden, dass sie eine Größe, eine Änderung dieser Größe und eine bestmögliche lineare Prognose[2] dieser Änderung angeben. Es ist dabei Konvention, dass man die Propositionen (hier zu lesen als Oberbegriff für Gleichungen, Ungleichungen, Mengeninklusionen etc.) zu diesen Variablen entsprechend wählt, so dass die beabsichtigte Interpretation umgesetzt wird. Es ist etwa Konvention, bei einer funktionalen Abhängigkeit zu den beiden Variablen x, y und der Proposition $y = f(x)$ noch die Variablen $dx, dy, \Delta x, \Delta y$ und die Beziehungen $dx = \Delta x$ (was ausdrückt, dass dies eine unabhängige, willkürliche Änderung ist: Es gibt keinen Grund, hier einen bestimmten Wert für Δx anzunehmen, man kann jede beliebige Änderung betrachten, daher kann man nichts anderes prognostizieren als eben diese willkürliche Änderung) sowie $\Delta y = f(x + \Delta x) - f(x), dy = f'(x)dx$ hinzuzunehmen. Was dabei $f'(x)$ ist, hängt vom Kenntnisstand ab: Wenn man über eine entwickelte Analysis verfügt, nimmt man die Ableitung, sonst sind ggf. ad hoc Überlegungen anzustellen dazu, was man hier prognostizieren kann und will.

Es ist auch erhellend, Differentiale auf den drei linguistischen Ebenen (Morris 1938) zu sehen. Auf syntaktischer Ebene erkennt man Differentiale am vorgestellten d. Einen Unterschied gibt es auf semantischer Ebene nicht: Wie andere Variablen auch sind Differentiale Referenzen auf Zahlen. Auf der pragmatischen Sprachebene unterscheiden sich Differentiale von anderen Variablen dadurch, dass sie für lineare Prognosen, resp. idealisierte Änderungen anderer Größen stehen sollen. Mit diesen Konventionen zur Sprachverwendung ist die Entwicklung der Theorie im dritten Kapitel kompatibel mit der strengen logischen Sicht.

Neben dieser strengen Sicht könnte man für den Unterricht auch eine etwas informellere Perspektive einnehmen und sagen, eine Größe sei eine Variable für eine Zahl oder ein Element eines allgemeineren Größenbereichs, die unter dem Veränderlichenaspekt im Sinne der Algebradidaktik (Malle 1993) gesehen wird und eine Interpretation in einem Kontext besitzt. Damit wird angenommen, dass die Werte einer Größe nicht willkürlich sind, sondern sich aus der Beschreibung einer Situation (kontextuelle Interpretation) ergeben. Dann kann man auch von der Änderung einer Größe sprechen Mit x ist dann via Kontext auch Δx definiert. Das x steht für eine Zahl, etwa die Temperatur an einem bestimmten Ort und zu einer bestimmten Zeit, gemessen in einer festgelegten Einheit (welcher Ort und welcher Zeitpunkt das ist, ist willkürlich). Nun ist die Änderung einer Zahl an sich nicht definiert, aber aus der kontextuellen Bedeutung (also auf pragmatischer Ebene) kann sich ein sinnvoller Wert einer Änderung ergeben. Im Beispiel betrachtet man irgendeine Änderung (des Ortes, der Zeit oder beides – dies ist völlig willkürlich; auch der Betrag ist willkürlich). Die Temperatur unter den so geänderten Umständen sei $x + \Delta x$ und ihre lineare[3] Prognose sei $x + dx$.

Typischerweise stehen mehrere Größen in bestimmten Relationen zueinander. Dann erzwingt die Linearitätsannahme im Falle zweier Differentiale, die sich auf die gleiche Änderung beziehen, ihre Proportionalität. Es ist nicht damit zu rechnen, dass die Verwendung von Größen in diesem lockeren Sinne Lernenden Probleme bereitet: Sie kennen das Sprechen über Größen aus dem Alltag und anderen Schulfächern. Im Gegenteil: Die Erfahrung (dies ist ein subjektiver Bericht) lehrt, dass auch viele Lehramtsstudenten eine Größensprechweise etwa bei der Monotonie verwenden. In der Didaktik selbst gibt es ebenfalls eine Tradition (vgl. Malle 1993, S.80 ff), reelle Variable als Veränderliche zu sehen, obwohl das, wie oben aufgeführt, im Widerspruch zur mathematischen Logik steht. Ich berühre hier also erneut die Frage der Balance von Anschaulichkeit und Strenge. Die Algebradidaktik hat sich gegen die streng logische Sicht entschieden, warum sollte die Analysisdidaktik dem nicht folgen? Unterstützt wird das auch durch den Befund (Gray et al. 2005), dass ein Verständnis von Variablen als veränderliche, in Beziehung stehenden Größen, ein positiver Prädiktor für den Erfolg in Analysis-Kursen ist. Der Unterrichtsvorschlag zu Differentialen kann unterschiedlich ausgestaltet werden, je nachdem welches Maß an formeller Strenge man anstrebt.

---

2 Falls in einer konkreten Situation ein Kriterium für „bestmöglich" oder „linear" fehlt, setzt man einfach $dx = \Delta x$ – das ist dann eben die bestmögliche lineare Prognose.

3 Die Linearitätsforderung ist dann sinnvoll, wenn man mehr als eine Größe und ihre Änderung betrachtet.

Neben den Größen selbst ist noch der Begriff der Funktion zu beleuchten, da Differentiale häufig zu funktional verbundenen Größen betrachtet werden. Oben wurde bereits erläutert, wie man qua pragmatischer Konvention Funktionen behandelt. Zusätzlich zur Funktionsgleichung $y = x^2$ beispielsweise nimmt man qua pragmatischer Konvention die Variablen dx, dy und die Beziehung $dy = 2x \cdot dx$ als Ausdruck der linearen Prognose zu den Gleichungen bzw. Propositionen, mit denen man arbeitet, hinzu.
Mit der weniger strengen Auffassung, dass Größen kontext-interpretierte Variablen unter dem Veränderlichenaspekt sind, besteht ebenfalls kein Problem mit Funktionen. Es bietet sich die Chance, an die historische, bei Euler vorhandene und heute noch in der Physik anzutreffende Sichtweise anzuknüpfen, nach der eine Funktion eine Beziehung zweier Größen (einer abhängigen und einer unabhängigen) ist. Funktionen können als spezielle, gerichtete Relationen gesehen werden und somit in den Kontext beliebiger Relationen eingebunden werden. Der Umgang mit Relationen in diesem Sinne ist didaktisch sinnvoll – dem funktionalen Denken sollte ein relationales Denken an die Seite gestellt werden. Relationales Denken ist nützlich, weil damit Funktionen, Gleichungen und Ungleichungen vereinheitlicht gesehen werden können. Beispielweise erlaubt es, Funktion und Umkehrfunktion als Ausdruck der gleichen Relation zu sehen. Der Begriff Relation wird dabei weit gefasst und umfasst sowohl die intentionale als auch die extensionale Interpretation. Letztere ist für die Formalisierung entscheidend und der Leser kann, wenn gewünscht, durchgehend in diesem Sinne für Relation n-stelliges Prädikat lesen. Funktionen sind spezielle Relationen, bei denen insbesondere die Symmetrie zwischen den in Beziehung gesetzten Größen aufgehoben wird, indem diese in abhängige und unabhängige Größen eingeteilt werden. Die Beispiele unten (insbesondere am Ende von 3.3) zeigen, dass es vorteilhaft ist, die Rollen von abhängiger und unabhängiger Größe flexibel wechseln zu können. In der Physik sind viele Größen relational verbunden. Die ideale Gasgleichung $pV = nRT$ verbindet Druck, Volumen, Stoffmenge und Temperatur eines Gases (vgl. auch Blum et al. 2015, S. 64) als Relation – erst in bestimmten Situationen mag es sinnvoll sein, eine davon als abhängige Variable auszuwählen. Ähnliche Überlegungen gelten weit über die Physik hinaus, etwa für die Beziehung zwischen Darlehens- und Guthabenzinsen. Wenn Lernende gebeten werden, Abhängigkeiten von Größen zu benennen und darzustellen, nennen sie häufig Situationen, die sich nicht mit Funktionen beschreiben lassen, sondern allgemeine Relationen darstellen (vgl. etwa Lengnink 2005, Abb. 2). Relationen sind demnach ein wesentliches Mittel zur logischen Strukturierung der wahrgenommenen Welt, aber auch der innermathematischen Konstruktionen. Die verbreitetste algebraische Ausdrucksform von Relationen sind Gleichungen. Sofern nicht nur Aussagen über einzelne Entitäten gemacht werden, sondern Dinge in Beziehung gesetzt werden, enthalten diese Gleichungen mehrere Variablen. Insofern sind multivariate Beziehungen, auch in der Form multivariater Analysis, ein sinnvoller Bildungsgegenstand für die Sekundarstufe. Dazu passt gut, dass eine Analysis mit Differentialen Relationen stark betont, indem einer Relation $g(x, y) = 0$ noch die Prognosen-Relation $\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = 0$ zugeordnet wird. Generell wird die klare Trennung zwischen univariater und multivariater Analysis aufgehoben. Schon die Beziehung $d(u + v) = du + dv$ ist am einfachsten multivariat zu verstehen (in funktionaler Sprache ist es eine Aussage zur Ableitung der Funktion $(x, y) \mapsto x + y$); eine univariate Interpretation erfordert, dass man u, v als Funktionen einer weiteren unabhängigen Variablen x sieht: $\frac{d}{dx}\big(u(x) + v(x)\big) = \frac{du(x)}{dx} + \frac{dv(x)}{dx}$. Im Vergleich zeigt sich, dass die multivariate Formulierung mit Differentialen notationell und konzeptionell die einfachere ist.

2.2 Differentiale als infinitesimale Größen

Leibniz hat Differentiale als infinitesimal kleine Größen gedacht. Was das genau bedeutet, ist schwer einzuschätzen, siehe Sonar (2011, S. 414) und Hairer und Wanner (1996, S.82). Es kann sein, dass er Differentiale als aktual unendlich klein oder nur als potentiell unendlich klein verstanden hat. Paradebeispiel ist jedenfalls sein charakteristisches Dreieck, ein infinitesimal kleines Steigungsdreieck mit Kathetenlängen dx, dy. Diese infinitesimale Sicht von Differentialen hat sich in Anwendungsfeldern bewährt, u.a., weil man eine gute Intuition dafür entwickeln kann (siehe etwa Rogers 2005). Dennoch ist es schwierig, immer korrekt zu argumentieren, und selbst große Mathematiker haben dabei Fehler gemacht (viele Beispiele findet man in Witzke 2010). Dies hat dazu geführt, dass Differentiale als unwissenschaftlich galten. Courant und Robinson (1970) schreiben: „...so wurde der Gegenstand [gemeint ist die Analysis, R. O.] mehr als ein Jahrhundert durch Formulierungen wie ‚unendlich kleine Größen', ‚Differentiale' oder, letzte Verhältnisse' usw. verschleiert". Mittlerweile aber steht mit der Nichtstandardanalysis (Robinson 1966) eine fachlich korrekte Fassung der Analysis auf Basis von infinitesimalen Zahlen zur Verfügung. Es gibt Analysiskurse auf Basis der Nichtstandardanalysis (Keisler 1976) als auch auf (gehobenem) Schulniveau (Wunderling 2007). Laugwitz und Schnitzspan (1983) haben in einem Themenheft der Zeitschrift „Der Mathematikunterricht" verschiedene Zugänge dargestellt und festgestellt, die Erfahrungen im Unterricht seien positiv, geben dafür aber nur sehr wenig Belege. Besser empirisch gestützt waren die ermutigenden Befunde von Sullivan (1976), aber trotzdem gab es grundlegende und erfolgreiche Kritik (Bishop 1977). Artigue (1991) diskutiert ebenfalls die Fähigkeit von Studiereden, mit Infinitesimalen zu argumentieren. Mein hier gemachter Vorschlag unterscheidet sich grundlegend davon, insofern darin Differentiale nicht unendlich klein sind.

2.3 Differentiale als endliche Größen

Dieser Abschnitt diskutiert fachliche Grundlagen, wie sie auch dem in Abschnitt 3 entwickelten Verständnis von Differentialen zu Grunde liegen. Der Leser könnte deswegen an dieser Stelle einen Blick auf den Anfang des dritten Abschnitts werfen.

In vielen einführenden Analysisbüchern (vor allem aus dem englischen Sprachraum – beispielsweise Thomas 2013) werden Differentiale im Anschluss an die herkömmlich definierte Ableitung als reellwertige Variablen eingeführt, für die die Beziehung $dy = f'(x)dx$ gilt. Geometrisch gedeutet sind (dx, dy) die Komponenten eines beliebigen Tangentialvektors[4] an den Funktionsgraphen, es sind also beliebig große Änderungen, aber nicht Änderungen des originalen Zusammenhangs, sondern seiner linearen Approximation. Diese Sicht ist völlig kompatibel mit dem Vorschlag dieser Arbeit, die Reihenfolge der Konzeptentwicklung ist aber bei der Definition $dy := f'(x)dx$ am logisch deduktiven Aufbau orientiert (zuerst erfolgt die fachliche Konstruktion, dann die intuitive Interpretation), während hier ein informell-intuitiver Einstieg gewählt wird.

Eine weitere Konzeption stellt das Frechetdifferential dar. Es wurde eigentlich für die Differentiation von Funktionen zwischen Banachräumen im Rahmen der Funktionalanalysis (siehe Heuser 1988, S. 331) entwickelt, kann aber als didaktisch interessante Fassung (siehe Torregrosa et al. 2006) der Differentiation reellwertiger Funktionen gesehen werden. Es entspricht der Sichtweise der Ableitung als lokaler Linearisierung. López-Gay et al. (2015) bezeichnen dies als „The differential as linear estimate" und begründen, warum dieser Ansatz für Physikstudenten besonders nützlich ist. Die Differentiale in diesem Sinne unterscheiden sich nur in der Art der Definition, nicht im Ergebnis von den oben über $dy := f'(x)dx$ eingeführten. Eingepasst in die Notation reellwertiger Funktionen, kann man die Frechet'sche Definition so formulieren: Für eine Funktion $f: \mathbb{R} \to \mathbb{R}$ wird die Änderung Δy der Funktionswerte für eine beliebige reelle Änderung $dx \in \mathbb{R}$ geschrieben als Summe $\Delta y = f(x + dx) - f(x) = L \cdot dx + \epsilon(dx) \cdot dx$ einer linearen Änderung $dy \coloneqq L \cdot dx, L \in \mathbb{R}$ und einer nichtlinearen Änderung $\epsilon: \mathbb{R} \to \mathbb{R}$. Dann nennt man f in x Frechet-differenzierbar, wenn $\lim_{dx \to 0} \epsilon(dx) = 0$ und in diesem Falle wird die Zahl L als $f'(x)$ und das Produkt $dy = L \cdot dx$ als Differential von y oder als $df(x)$ bezeichnet. Etwas kompakter geschrieben lautet diese Definition für einen Zusammenhang $y = f(x)$: $dy = f(x + dx) - f(x) - \epsilon(dx) \cdot dx$.

4 Tangentenvektoren sind nur in ihrer Richtung, nicht in ihrer Länge festgelegt. Ihre Länge kann willkürlich gewählt werden und das entspricht der willkürlichen Wahl, wie weit prognostiziert wird.

Wenn die Ableitung vorab als Grenzwert des Differenzenquotienten definiert wurde, kann man leicht zeigen, dass differenzierbare Funktionen auch Frechet-differenzierbar mit $L = f'(x)$ sind[5].

Von meinem in Abschnitt 3 vorgestellten Zugang unterscheidet sich dies nur in zweierlei Hinsicht: Die Grenzwertüberlegung, die bei mir nur intuitiv nötig ist, wird hier formal benötigt und die grundlegende Formel, mit der man etwa beim Modellieren argumentiert, ist minimal komplexer ($\Delta y = dy + \epsilon(dx)dx$ statt $\Delta y \approx dy$). Deswegen sind die Regelherleitungen auch formal etwas umfangreicher, dafür aber fachlich lückenfrei darstellbar. Man geht dann von der Form $dy = f(x + dx) - f(x) - \epsilon(x)dx$ aus. Für die quadratische Funktion mit Gleichung $y = x^2$ findet man damit: $dy = (x + dx)^2 - x^2 = 2xdx + dx^2$. Also ist hier $\epsilon(dx) = dx$. Höhere Potenzen folgen über Induktion oder mittels Binomialsatz. Auf Beweise der weiteren Ableitungsregeln in diesem Kalkül wird hier verzichtet.

Differentiale im Sinne von Frechet behandelt auch Pickert (1962). Ausgehend von der Frage der näherungsweisen Berechnung von Funktionswerten motiviert er die lineare Approximation. Weitere didaktische Überlegungen gibt er nicht, aber er legt besonderen Wert auf die Frage der Notation und schlägt z.B. vor, die unabhängige Änderungsgröße nicht als Differential dx zu notieren, sondern schlicht als h - in Anlehnung an die Rolle von h im Differenzenquotienten. Ferner schlägt er vor, die Variable, nach der differenziert wird, als Index zu kennzeichnen, so dass man unterscheiden kann $d_x(xy^2) = hy^2$ und $d_y(xy^2) = 2xyh$. Pickert verwendet also im Gegensatz zu meinem Vorschlag dx nicht als eigenständige Variable, sondern für ihn ist d eine Funktion, die auf Funktionen wirkt. Man könnte statt $d_y(xy^2)$ auch $d(y \mapsto xy^2)$ schreiben. Das hat den Nachteil, dass die Ausdehnung auf mehrere Variablen notationell schwieriger wird, während im vorliegenden Vorschlag einfach $d(xy^2) = y^2dx + 2xydy$ geschrieben wird.

Godement (2004) behandelt Differentiale als abgeleitete Größen, wobei er wie Pickert die funktionale Perspektive einnimmt. Er definiert für eine reellwertige Funktion das Differential durch $df(a, h) \coloneqq f'(a)h$. Mit x bezeichnet er die identische Funktion $x: x \mapsto x$. Für diese Funktion gilt $dx(a, h) = h$, so dass man die beiden Formeln kombinieren kann zu $df(a, h) = f'(x)dx(a, h)$. Inhaltlich unterscheidet sich das nicht von den Frechetdifferentialen. Das Bewusstmachen der Abhängigkeiten in der Notation kann eine klärende Funktion haben, weil es klar macht, dass alles von der betrachteten Stelle a abhängt und die aufeinander bezogenen Differentiale dx, dy in ihrer Größe proportional zueinander sind (h). Für das praktische Arbeiten, insbesondere beim Modellieren, ist der zusätzliche Notationsballast aber eher hinderlich. Das dy in meinem in Abschnitt 3 entwickelten Zugang entspricht dem Wert der Funktion df wie bei Godement definiert an der jeweiligen Stelle und mit $h = dx$.

Eine weitere Variante der Frechetdifferentiale $dy = f(x + dx) - f(x) - R_x(dx)$ bringt (Loomis und Sternberg 1968). Dort wird auf die Verschwindungseigenschaft des Restgliedes abgehoben, dass $\lim\limits_{dx\to 0} \frac{R_x(dx)}{dx} = 0$. Funktionen mit dieser Eigenschaft werden zu einer Funktionsklasse σ zusammengefasst. Es ist dann, in vereinfachter Notation: $dy = \Delta y + \sigma$.

5 Nach Division durch dx und Umstellen geht nämlich $L \cdot dx = f(x + dx) - f(x) - \epsilon(dx) \cdot dx$ über in $L = \frac{f(x+dx)-f(x)}{dx} - \epsilon(dx)$. Wendet man darauf den Grenzwert $dx \to 0$ an, findet man die gewöhnliche Ableitung $L = \lim\limits_{dx\to 0} \frac{f(x+dx)-f(x)}{dx} = f'(x)$. Diese Überlegungen lassen sich auch rückwärts durchlaufen, d.h. man startet von der Definition $\epsilon(dx) = \frac{f(x+dx)-f(x)}{dx} - f'(x)$ für eine im üblichen Sinne differenzierbare Funktion und sieht, dass $\lim\limits_{dx\to 0} \epsilon(dx) = 0$ gilt. Damit ist die Äquivalenz zur gewöhnlichen Formulierung der Differenzierbarkeit gezeigt.

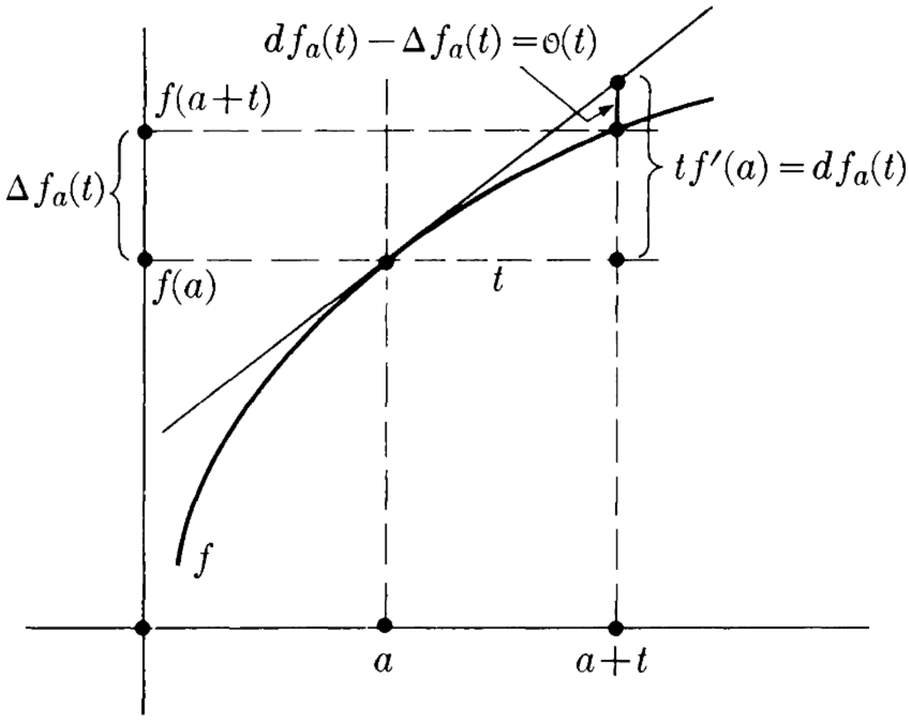


Abbildung 2: Frechetdifferentiale nach Loomis und Sternberg (1968, S. 141). Der unabhängige Zuwachs ist t (entsprechend dx bei mir und h bei Godement (2004)).

3 Ein Unterrichtsgang

Dieser Abschnitt beschreibt im Sinne eines didaktischen Gedankenexperiments einen hypothetischen Unterrichtsgang, der im Anschluss an die Sekundarstufe I direkt in die Analysis mit Differentialen einführt.

3.1 Die Begriffseinführung

Zentraler Begriff der Analysis – so wie sie hier verstanden wird – ist der der Änderung, nicht der der Änderungsrate (vgl. auch Picher (2010) und die Diskussion im zweiten Teil dieses Beitrags). Es werden daher zum Einstieg Situationen benannt, wo sich etwas ändert: Wasserpegel, Aktienkurse, Positionen, Temperaturen u.v.m. Ein Arbeitsauftrag könnte lauten: Beim Wachstum der Weltbevölkerung sind zwei Größen wichtig, die Zeit und die Zahl der jeweils lebenden Menschen. Erstellt eine möglichst lange Tabelle (mit den Spalten: Änderungsprozess, erste Größe, zweite Größe, ggf. weitere Größen) von weiteren Prozessen, bei denen sich etwas ändert. Die Lernenden können dabei viele Beispiele sowohl aus dem Mathematikunterricht (z.B. lineares und exponentielles Wachstum) als auch aus anderen Fächern und dem Alltag einbringen. Ziel ist eine Sammlung und eine Klärung der gemeinsamen Struktur, nämlich dass man Größen und ihre Änderungen betrachtet. Erarbeitet werden soll dann etwa der folgende Sachverhalt (die verwendete Mathematik ist unverändert die elementare Algebra der Sekundartstufe I):

Änderungen können räumlich oder zeitlich sein, oder auch nur einen abstrakten Vergleich betreffen. In jedem Falle gibt es mindestens eine Größe x und ihre Änderung Δx, den Änderungsprozess kann man schreiben als $x \mapsto x + \Delta x = x_{neu}$.

Es wird nun eine Reihe von Änderungsprozessen betrachtet. Dabei ist deren Verlauf bis zu einem gewissen Zeit-/Raumpunkt jeweils bekannt. Die Fragestellung (also der Arbeitsauftrag an die Lernenden) lautet nun[6]: Kann man die weitere Änderung vorhersagen und wie sieht diese aus? Wenn man beispielsweise Bahnen von Flugzeugen hat, die abstrahiert in Abb. 3 wiedergegeben

6 In dieser Lernsituation wurde bewusst auf einen Operator verzichtet.

sind, dann kann man nach Prognosen fragen, wie sich das Flugzeug nach dem letzten in der Darstellung wiedergegebenen Zeitpunkt wohl weiterbewegen wird. Dabei wird schnell klar, dass man unterschiedlich prognostizieren kann: Beispielsweise kann man in Abb. 3 vermuten, dass die erste Bahn eine Kreisbahn ist und es also kreisförmig weitergeht: Das ist die Kreisprognose. Bei anderen Bahnen liegt eine geradlinige Prognose nahe. Dabei kann also erarbeitet werden, dass Prognosen nicht eindeutig sind: Man kann verschiedene Formen wählen und die Wahl wird von verschiedenen pragmatischen Gesichtspunkten abhängen, etwa ob man eine kurz- oder langfristige Prognose anstrebt und welche Informationen man einbeziehen will. Prognosen sind also ähnlich wie Modelle nicht eindeutig gegeben. Insbesondere sollte diskutiert werden, welche Informationen man für die Prognose zu Grunde legt: Eine Prognose kann den ganzen bisherigen Verlauf in Betracht ziehen oder nur den letzten Trend. Bei der Bahn rechts oben in Abb. 3 entscheidet das über die Frage, ob eine steigende oder fallende Prognose plausibler erscheint: Wenn man den gesamten Verlauf in Betracht zieht (und annimmt, dass die Bahn von links nach rechts durchlaufen wird), dominiert eine Bewegung nach unten, die man dann in der Prognose fortsetzen wird. Bei einer Betrachtung des letzten Trends erscheint dagegen eine Fortsetzung nach oben als plausibler.

Das Erstellen einer Prognose ist also eine Modellbildung und wie in der allgemeinen Modelltheorie (Stachowiak 1973, S. 131f.) behauptet, spielen bei der Modellauswahl u.a. pragmatische Überlegungen eine Rolle. Das Modell der geradlinigen Fortsetzung erscheint besonders einfach und es gibt zudem einen physikalischen Grund: Die Kugelrollbahn (siehe Abb. 4, vgl. Oldenburg 2007) zeigt, dass auf einer Ebene rollende Kugeln, die zunächst durch eine Begrenzung auf eine gekrümmte Bahn gezwungen werden, sich nach Ende der Führung geradlinig weiterbewegen. Auch wenn dies ein Spezialfall ist, macht er deutlich, dass die lineare Prognose, die sich nur auf den letzten Trend bezieht, zumindest für einige Situationen gut passend ist. Aber es ist klar, dass es Alternativen gibt, und das sollen die Lernenden wissen und akzeptieren. Dazu ist es hilfreich wenn sie schon Erfahrungen mit offenen Modellbildungen haben, insbesondere dass sie erfahren haben, dass es sinnvoll sein kann, ein einfaches Modell zu verwenden, auch wenn es möglicherweise komplexere gibt, die näher an der vermuteten Wahrheit sind.

Ebenfalls schon hier kann deutlich werden, dass ein einziger Punkt zu keiner Prognose befähigt. Prognosen sind nur möglich, wenn man zumindest einen lokalen Blick auf die Umgebung des Prognosezeitpunktes werfen kann.

Der Gesichtspunkt der linearen Prognose wurde von Oldenburg (2012) dargestellt und durch ein HTML5-Applet (Oldenburg 2012b) unterstützt. Das Applet zeigt einen sich bewegenden Punkt, dessen Bewegung nach einiger Zeit stoppt und man soll versuchen eine lineare Prognose abzugeben, wo sich der Punkt eine Sekunde später befinden könnte. Dadurch kann das Generieren einer Prognose geübt werden.

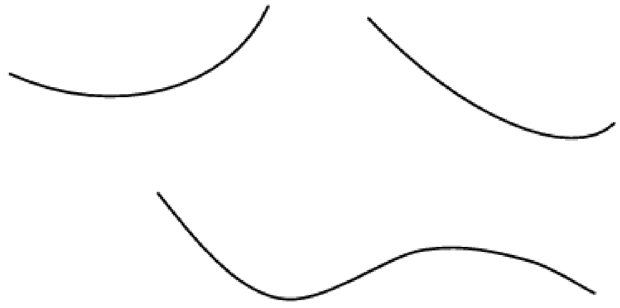


Abbildung 3: Verschiedene Bahnen – Wie geht es weiter?

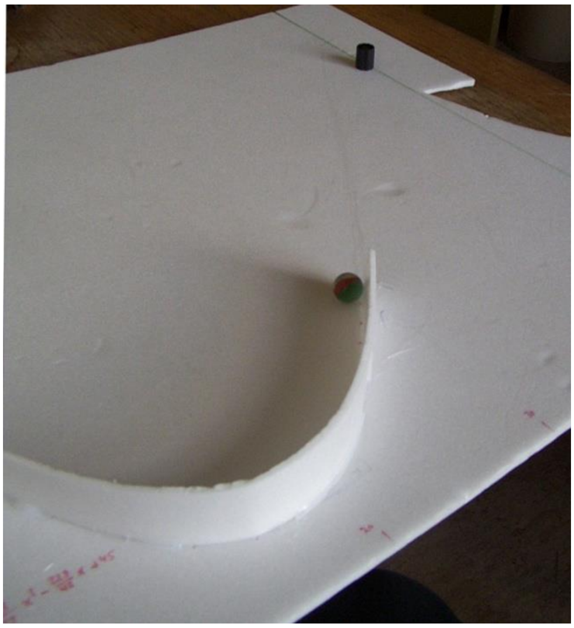


Abbildung 4: Die Kugelrollbahn realisiert eine lineare Approximation

An dieser Stelle können Differentiale 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 $\Delta x \approx dx$.

Was dabei bestmöglich genau bedeutet, wird nicht genauer angegeben – dies spiegelt wiederum die Willkürlichkeit der Modellbildung. Falls Änderungen exakt bekannt sind (oder als unabhängige Größen vorgegeben werden), braucht man nicht zu prognostizieren, man nennt dann dx einfach eine Änderung, was sprachlich viel angenehmer ist als „lineare Prognose einer Änderung“.

Diese Definition erinnert an die Definition der Wahrscheinlichkeit als bestmögliche Prognose für relative Häufigkeiten in langen Versuchsreihen. Hier wie da ist die Frage, was bestmöglich genau bedeutet. Dies sei zunächst der Intuition überlassen. Die Beziehung zwischen relativen Häufigkeiten und Wahrscheinlichkeiten findet ihre Analogie[7] in der von Δx und dx (beide Male geht es um die Beziehung einer gut erfassbaren Größe (relative Häufigkeit, Δx) und einer theoretischen Größe (Wahrscheinlichkeit, Differential dx)). Der Modellierungsaspekt macht deutlich, dass das lineare Prognosemodell nicht das einzig denkbare ist.

Eine genauere Spezifikation, was eine „bestmögliche“ lineare Prognose ist, ist erst möglich, wenn man speziellere Situationen betrachtet, insbesondere solche, bei denen nicht nur eine Größe und ihre Prognose zu betrachten ist, sondern mindestens eine zweite. Diese können funktional (Terme, Funktionen) oder relational (Gleichungen) verbunden sein. Die meisten Vorerfahrungen haben Lernende mit Funktionen. An Kurven (Graphen von Funktionen, aber auch allgemeinerer Relationen) kann man die Idee der linearen Fortsetzung geometrisch intuitiv entwickeln. Es ist zu erwarten, dass die Lernenden die Knickfreiheit intuitiv als ein Gütemaß akzeptieren. Ein Arbeitsauftrag könnte darin bestehen, dass man den Lernenden eine Reihe von Graphen wie in Abb.

7 Die Analogie zum Wahrscheinlichkeitsbegriff erstreckt sich auch darauf, dass sich der Bereich, in dem es objektiv „richtige“ Werte zu geben scheint, sukzessive ausdehnt, wenn auch nie alles umfasst. Bei der Wahrscheinlichkeit mögen zunächst nur die Werte bei Laplaceexperimenten offensichtlich, später auch bei einigen anderen, die man zu Laplaceexperimenten in Beziehung setzen kann. Bei den Differentialen ist zunächst nur bei linearen Zusammenhängen offensichtlich, wie sie linear prognostiziert werden sollten. Bei der Fortsetzung von Kurven ist aber zB auch ableitbar, dass bei Kegelschnitten die algebraisch berechenbaren Tangenten allein als bestmögliche lineare Prognose in Frage kommen.

3 gibt, die in verschiedenen Farben durch unterschiedliche Halbgeraden fortgesetzt sind. Die Lernenden sollen dann begründet eine bestmögliche Prognose auswählen.

Interessant ist in diesem Zusammenhang das auf Kirsch (1979) zurückgehende Hilfsmittel des Funktionenmikroskops. Es zeigt, dass praktisch alle schulüblichen Funktionen lokal geradlinig aussehen, d.h. lokal ist die lineare Prognose besonders gut. Da bisher völlig kalkül- und algebrafrei gearbeitet wird, gibt es keinen Grund, sich auf bestimmte einfache Funktionsklassen zu beschränken. Beispielsweise kann man den Graphen der Sinusfunktion in der Nähe von $x = 1$ betrachten und beschreiben lassen. Danach gibt man an, dass $\sin(1) \approx 0.8415, \sin(1.1) \approx 0.8912$ und fragt nach einer Schätzung von $\sin(1.2), \sin(1.05)$. An dieser Stelle kann man auch Tabellen-gestützte Prognosen betrachten, sowohl bei realweltlichen Daten als auch bei Wertetabellen von Termen. Das Phänomen der lokalen Linearität kann so graphisch und tabellarisch erforscht werden.

An dieser Stelle lässt sich aufzeigen, dass es Zusammenhänge gibt, bei denen lineare Prognosen sehr schnell sehr schlecht werden können, etwa bei der Betragsfunktion. Solche Stellen, die sich schlecht prognostizieren lassen, kann man nicht-differenzierbar nennen, muss aber eine präzise Definition des Begriffs aufschieben. Möglich ist etwa aufzuzeigen, dass man linear von $|-0.1| = 0.1, |0| = 0$ auf $|0.1| = -0.1$ schließen würde – und dass das eine ziemlich schlechte Prognose ist, während an anderen Stellen die lineare Prognose lokal sogar perfekt ist.

Bei Termen kann man gut die Linearitätsbedingung erarbeiten: Man gibt den Lernenden den Auftrag, mit Wertetabellen für verschiedene Terme das Änderungsverhalten zu untersuchen, wobei jeweils systematisch (so wie man es auch im Physikunterricht gelernt haben sollte) nur eine Variable verändert wird. Wählt man die Werte für eine Variable in einen beliebigen Ausgangswert und das 2-, 3-, 4- fache usw., und findet man dann auch bei der zweiten Größe (z.B. dem Termwert) diese Verhältnisse, dann hängen diese Größen linear zusammen.

Falls sich dies vorher noch nicht aus dem Unterrichtsgeschehen ergeben hat, empfiehlt es sich, spätestens jetzt auf Funktionen einzugehen und die zu Beginn erstellte Liste von Änderungsprozessen daraufhin anzuschauen, welche davon durch Funktionen beschrieben werden können und welche Größen dann als abhängige resp. unabhängige Variablen verstanden werden.

Linearität ist am Graphen (von Funktionen oder auch allgemeineren Relationen, das ist egal) leicht zu interpretieren. Allgemein bedeutet sie, dass jede Ver-*k*-fachung des Differentials einer unabhängigen Größe dazu führt, dass auch die Änderung der abhängigen Größe mit dem gleichen Faktor wächst. Erarbeitet werden sollte dann die algebraische Charakterisierung, dass ein Term linear in dx, dy, dz ist, wenn er die Bauart hat $Adx + Bdy + Cdz$, wobei A, B, C nicht von dx, dy, dz abhängen. Die Koeffizienten A, B, C sind die lokalen Änderungsraten. Da ich aber nicht dividiere, um diese explizit zu bestimmen, hat das Konzept der Änderungsrate bei mir zumindest zunächst keine Bedeutung. Grundsätzlich ist eine lineare Beziehung $y = ax$ natürlich äquivalent zu einer konstanten Änderungsrate $\frac{y}{x} = a$, aber der Charme der Differentiale kommt daher, dass man diese Quotientenform weitgehend vermeiden kann, wie die Beispiele zeigen werden.

Es soll die folgende Konvention zur Syntax gelten, um Uneindeutigkeiten zu vermeiden: dx^2 bedeutet $(dx)^2$ – im Gegensatz zu $d(x^2)$.

3.2 Entwicklung der Theorie: Beziehungen zwischen Differentialen

In Veränderungsprozessen sind meist mehrere Größen involviert und deswegen interessiert es, wie Änderungen zusammenhängen. Für den Unterricht geeignete Beispiele sind u.a. die rutschende Leiter, die Höhe des Wasserpegels an den Ecken eines gekippten Gefäßes (Becker und Shimada 1997), die Höhe der Enden einer Wippe. Dabei können die Größen identifiziert werden und ihr Änderungsverhalten zunächst qualitativ in Worten beschrieben werden. Angenommen, in einer Summe $x + y$ wachsen die beiden Summanden x, y um dx resp. dy auf $x + dx$ resp. $y + dy$, die neue Summe ist also $x + y + dx + dy$. Dann ist die gesamte Änderung $dx + dy$ linear in beiden Veränderungen, also ist dies die bestmögliche lineare Prognose (man muss keine weiteren

Approximationen machen, um Linearität zu erreichen). Es gilt damit die Regel $d(x+y) = dx + dy$, die durch Abb. 5 nahegelegt wird.

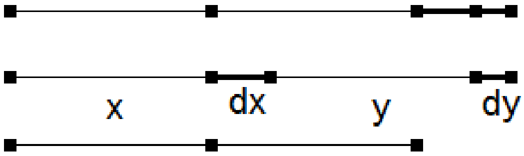


Abbildung 5: Linearität von Änderungen

Diese simple Regel hat inhaltlich leicht fassbare Folgen: Wenn eine Summe konstant ist ($x + y = c$), dann gilt $dx + dy = 0$, also $dx = -dy$, d.h. wenn die eine Größe wächst, muss die andere abnehmen.

Ein Spezialfall ist $d(2x) = d(x + x) = 2dx$ und analog gilt allgemein $d(k \cdot x) = k \cdot dx$.

Der nächste wesentliche Schritt ist die Untersuchung des Änderungsverhaltens von Produkten uv. Wiederum gibt es viele Kontexte, in denen man sich für die Änderung des Wertes eines Produktes interessieren kann. In einer ersten sinnstiftenden Kontextualisierung denkt man sich beispielsweise u als eine Gütermenge und v als Preis pro Einheit, so dass das Produkt den Gesamtpreis darstellt. Wenn dann u um du wächst, steigt der Preis um vdu, bei Wachstum von v analog um udv, zusammen genommen also um $udv + vdu$. Dieses Argument könnte aber auch von Schülern kritisiert werden: Durch Ausmultiplizieren der neuen Größen ergibt sich als Änderung $udv + vdu + dudv$. Warum diese Form nicht die angemessene ist, kann dann begründet, aber nicht bewiesen werden:

- Der letzte Term ist nicht mehr linear gemäß der obigen Erklärung der Linearität. Für eine bestmögliche lineare Prognose darf er also nicht auftreten. Ihn wegzulassen ist ein Modellierungsfehler (Modelle erheben ja nicht den Anspruch, die Wahrheit abzubilden, sondern eine pragmatische Beschreibung zu geben – dass sie sich bewährt, wird sich zeigen)
- Die Situation kann auch durch die übliche geometrische Visualisierung des Produkts als Rechteck, dessen Seitenlängen um du und dv zunehmen, erläutert werden (Tietze 2000, S. 172). Der nichtlineare Term $du \cdot dv$ entspricht einem kleinen Rechteck von relativ kleinem Flächeninhalt. Die Vernachlässigung des nichtlinearen Terms im Sinne des Modells der linearen Prognose bewahrt also die Erwartung, dass die Prognose gut ist für kleine Änderungen.

Dies spricht zusammen für die Postulierung der folgenden Produktregel:

$$d(u \cdot v) = u \cdot dv + v \cdot du.$$

Im Spezialfall des Quadrates folgt: $d(x^2) = 2 \cdot x \cdot dx$ und weiter $d(x^3) = d(x \cdot x^2) = dx \cdot x^2 + x \cdot 2xdx = 3x^2dx$ und weiter liefert Iteration mit einem informellen Induktionsschluss $d(x^n) = d(x \cdot x^{n-1}) = dx \cdot x^{n-1} + x \cdot (n-1) \cdot x^{n-2}dx = n \cdot x^{n-1}dx$.

Aus der Sekundarstufe I sind umgekehrt proportionale Zusammenhänge bekannt, bei denen also $y = \frac{c}{x}, x \cdot y = c$ gilt. Eine geometrische Visualisierung ist die Beziehung der Seitenlängen von Rechtecken, die alle den gleichen Flächeninhalt c besitzen. Man fragt dann, wie sich eine Änderung der einen Seitenlänge auf die andere auswirkt. Wenn man es breiter macht, muss es flacher werden, damit $x \cdot y = c$, oder $y = \frac{c}{x}$ gültig bleibt. Dies sind zwei relational verbundene Variablen – eine typische Situation, die zur Verwendung von Differentialen einlädt. Wenn man auf beiden Seiten die Änderungen betrachtet, findet man $ydx + xdy = 0$. Das zeigt, dass dx und dy unterschiedliche Vorzeichen haben (wenn die Breite zunimmt, nimmt die Höhe ab). Es folgt: $dy = -\frac{y}{x}dx = -\frac{c}{x^2}dx$

Erfreulicherweise ist dieses Ergebnis auch verträglich mit der oben gegebenen Regel für $d(x^n)$ und auch $d\left(x^{\frac{n}{m}}\right)$ lässt sich so finden.

Eine interessante Eigenschaft des Kalküls zeigt sich bei der Berechnung von $d((2x)^5)$. Man benötigt nämlich keine explizite Formulierung der Kettenregel, sondern diese folgt aus dem Substitutionsprinzip: Man setzt in $d(u^n) = n \cdot u^{n-1} \cdot du$ ein $n = 5, u = 2x$ und erhält $d((2x)^5) = 5 \cdot (2x)^4 \cdot d(2x) = 10 \cdot (2x)^4 dx$.

3.3. Tangenten, Grenzwerte und weitere Ableitungsregeln

Funktionsgraphen als Darstellung bestimmter Beziehungen zwischen Variablen wurden schon am Anfang in die Beschreibung integriert. Diese werden jetzt noch etwas vertieft betrachtet. Bei differenzierbaren Funktionen mit Gleichung $y = f(x)$ zeigt das Funktionsmikroskop nahezu einen geradlinigen Ausschnitt, der global eine Gerade definiert, die man Tangente nennt. Die Lernenden können beispielsweise für $y = x^2$ an der Stelle $x_0 = 1$ verschiedene Geraden (mit den Gleichungen $g_m(x) = m \cdot (x - x_0) + f(x_0)$) anlegen, die sich in der Steigung m unterscheiden. Auf rein phänomenaler Ebene (Funktionsmikroskop) kann festgestellt werden, dass die Gerade für $m = 2$ besonders gut passt, und dass man beim Heranzoomen, also in lokaler Betrachtung, fast keinen Unterscheid zwischen einem Stück dieser Geraden und dem Funktionsgraphen erkennen kann. Die bisher schon erarbeiteten Regeln für Differentiale besagen $dy = 2 \cdot dx$. Die Differentiale dx, dy können beliebig groß sein, solange sie diese Beziehung erfüllen. Graphisch gedeutet sind sie also die Kathetenlängen von Steigungsdreiecken an die Gerade mit der Steigung $\frac{dy}{dx} = 2 = m$. Dass die so gefundene Gerade als Tangente bezeichnet wird, ist eine einfache Mitteilung an die Lernenden. Die Differentiale dx und dy sind also die Komponenten eines beliebigen Vektors in die Tangentenrichtung. Die Tangentensteigung ist (Steigungsdreieck): $\frac{dy}{dx}$. Damit hat man Methoden zum Berechnen von lokalen Steigungen vieler Funktionen.

An dieser Stelle erschließen sich auch viele schulübliche Anwendungen der Änderungsrate. Als Prototyp sei die Bewegung eines Fahrzeugs genannt. Die Änderung seiner Position x ist eine zurückgelegte Streckenlänge dx und die zugehörige Änderung der Zeit dt ist die benötig Zeitdauer. Der Quotient kann dann als Geschwindigkeit interpretiert werden. Da der Quotient eine neue interessante Größe ist, ist es auch motivierbar, eine neue Bezeichnung einzuführen: $f'(x) = \frac{dy}{dx}$. Die Sprechweise kann eingeführt werden, dass man generell den Quotienten zweier Änderungen eine Änderungsrate nennt. Hier kann also die Standard-Begrifflichkeit erschlossen werden.

An dieser Stelle können auch einige Überlegungen zu Grenzwerten angestellt werden. Die oben gegebene intuitive Definition des Differentials dx einer Größe x als die bei lokaler Betrachtung bestmögliche lineare Prognose der Änderung Δx verwendet den Grenzwertbegriff nicht explizit. Er steckt aber im Begriff der „lokalen Betrachtung" und in der im zweiten Satz nachgetragenen Erwartung: „Für kleine Änderungen erwartet man, dass $\Delta x \approx dx$." Die obige Diskussion der Frechet-Differentiale hat gezeigt, wie das exaktifiziert werden kann. Unterrichtlich sollte eine solche Form aber noch nicht angestrebt werden. Möglich ist an dieser Stelle, wo einige nichttriviale Differentiationsregeln zur Verfügung (siehe 3.2) stehen, für einige exemplarische Funktionen und Stellen (etwa traditionell $y = f(x) = x^2, x_0 = 1$) systematisch die tatsächlichen Änderungen Δy und die Differentiale dy für eine Reihe von Werten von dx, etwa 0,1; 0,01; ... zu vergleichen und so die schon eingangs gestellte Forderung, dass für kleine Änderungen $\Delta x \approx dx$ sein soll, zu bestätigen. Dadurch leistet man Propädeutik für den Grenzwertbegriff. Die numerische Rechnung an dieser (im Vergleich zum traditionellen Unterrichtsgang relativ späten) Stelle ist nicht Grundlage der Ableitung, sondern wird genutzt, um eine bereits entwickelte Erwartung, dass für kleine Änderungen gilt $\Delta x \approx dx$, zu bestätigen. Gleichzeitig erschließt sich so die für die Anwendung der

Differentialrechnung nützliche Vorstellung, das dx zu denken als „a little bit of *x*" (Thompson u. Gardner 1998) oder „a little change of *x*" (Thompson 2013) – vgl. Abschnitt 2.4.

Die Ableitung der Wurzelfunktion gelingt gut mit der Sichtweise der relational verbundenen Variablen: $y = \sqrt{x} \Rightarrow y^2 = x \Rightarrow 2y \cdot dy = dx \Rightarrow dy = \frac{dx}{2y} = \frac{dx}{2\sqrt{x}}$. Diese einfache Rechnung stellt einen unvollständigen Beweis dar, weil die Frage der Differenzierbarkeit (also der Existenz der Ableitung) aufgrund der noch nicht geleisteten vollständigen Formalisierung nicht geführt werden kann – aber sie hat den Vorteil, dass man gut argumentieren kann! Die Differenzierbarkeit (im Sinne der lokalen Linearität) kann zumindest intuitiv begründet werden: Der Graph der Wurzelfunktion und der quadratischen Funktionen gehen durch Spiegelung an der Winkelhalbierenden auseinander hervor. Wenn also die eine linear prognostizierbar sein sollte, dann auch die andere – zumindest falls die Prognose nicht Richtungen ergeben sollte, die funktional nicht darstellbar sind (in (0|0) unterscheiden sich die beiden Graphen in ihrer Differenzierbarkeit!)[8] .

Nach dem gleichen Muster behandelt man die Quotientenregel ($v \neq 0$ wird angenommen):

$$y = \frac{u}{v} \Leftrightarrow y \cdot v = u \Rightarrow dy \cdot v + y \cdot dv = du \Rightarrow dy = \frac{du - y \cdot dv}{v} = \frac{v \cdot du - u \cdot dv}{v^2}$$

Dass die Kettenregel mit Differentialen besonders schön ist, ist Allgemeingut: $z = f(y) \Rightarrow dz = f'(y)dy, y = g(x) \Rightarrow dy = g'(x)dx, dz = f'\big(g(x)\big)g'(x)dx$ oder noch prägnanter: $\mathrm{d}z = \frac{dz}{dy} \cdot \mathrm{d}y$. Die Regel für die Umkehrfunktion ist ein Spezialfall.

Das übliche Argument gegen diese Herleitung trifft hier nicht, weil die Differentiale in meinem Sinne keine Infinitesimale sind und weil die Hauptschwierigkeit beim konventionellen Beweis hier ebenfalls sichtbar ist, nämlich, dass $dy = 0$ sein könnte. Es wird hier also kein falscher Beweis gelehrt. Eine auch Sonderfälle umfassende Beweisführung kann allerdings noch nicht geleistet werden.

Bei Sinus und Cosinus liegen die Dinge ähnlich. Am Einheitskreis deutet man die Koordinaten eines Kreispunktes als $\big(x = cos(t), y = sin(t)\big)$. Wenn sich t (im Bogenmaß) um ein Stückchen dt ändert, ist die lineare Prognose der Bewegung von P längs der Kreistangenten um ein Stück der Länge dt. Da die Tangente senkrecht auf dem Radiusstrahl steht, muss $\frac{dy}{dx} \cdot \frac{y}{x} = -1$ gelten (dies kann auch als Ähnlichkeitsbeziehung am Einheitskreis gedeutet werden, oder in der Form $xdx + ydy = 0$ als Skalarprodukt). Eine Lösung ist naheliegend[9]: $dy = k \cdot x, dx = -k \cdot y$. Die Länge der Tangentendifferentiale bestimmt man so, dass die prognostizierte Bewegung um $\sqrt{dx^2 + dy^2}$ gerade gleich dem Betrag der tatsächlichen Bewegung auf dem Kreis ist, also gleich dt. Damit steht fest: $dy = x \cdot dt = cos(t)dt, dx = -ydt = -sin(t)dt$.

Schwierig ist allein die Ableitung der Exponentialfunktion. Dazu später mehr.

3.4 Monotonie

Die Formalisierung von intuitiv naheliegenden Begriffen ist ein unterrichtliches Problem. Den Schwierigkeiten, die Lernende dabei haben können, setzt die Evolution der Lehrpläne im Wesentlichen eine Entformalisierung der Schulmathematik entgegen, obwohl es m.E. sinnvoller wäre, Formalisierung zu erleichtern.

8 Ich gehe allerdings davon aus, dass die Frage der Differenzierbarkeit zu Beginn von den Lernenden nicht hinterfragt wird. Die hier gegebene Argumentation zur Differenzierbarkeit (im Sinne linearer Diagnostizierbarkeit) sollte deswegen erst zu einem späteren Stadium (siehe Exaktifizierung) in Angriff genommen werden.

9 Wer diese Lösung nicht sieht, löst die Beziehung auf nach $dy = -\frac{x}{y}dx$ und setzt ein $dt^2 = dx^2 + dy^2 = dx^2 + \frac{x^2}{y^2}dx^2 = \frac{x^2+y^2}{y^2}dx^2$. Mit $x^2 + y^2 = 1$ folgt $dx = \pm ydt$. Das richtige (negative) Vorzeichen liest man aus der Bewegungsrichtung am Einheitskreis ab.

Die Definition der strengen Monotonie einer Funktion f auf einem Intervall $[a,b]$ kann formuliert werden über den Punktvergleich (d.h. $\forall x_1, x_2 \in [a,b] \wedge x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$), der eine Aussage über Änderungen ist (nämlich:$\Delta x > 0 \Rightarrow \Delta f(x) > 0$), oder über die Positivität aller Differenzenquotienten auf Teilintervallen. Letzteres führt leicht zum üblichen Kriterium des Monotoniesatzes. Die Vorteile einer Formulierung mit Änderungen soll nun am Beispiel einer Studentenschwierigkeit illustriert werden. In einer Hausarbeit stand: Die Exponentialfunktion steigt rechts vom Ursprung und fällt links davon. Diese fachlich falsche Ausdrucksweise kann man verstehen, wenn man die eigene Position in Bezug auf den Graphen und Bewegungen in der Graphenwelt ausgehend vom Ursprung betrachtet. Die Aussage ist: Wenn ich nach rechts gehe (also $\Delta x > 0$), dann nehmen die Werte zu (also $\Delta y > 0$), und wenn ich nach links gehe ($\Delta x < 0$), verkleinern sich die y-Werte ($\Delta y < 0$). Verwendet man statt Änderungen ihre linearen Prognosen kommt man zum Kriterium der Analysis. Man sieht: Differentiale bieten eine gute Form, solches Wissen zu formalisieren. Eine Maximalstelle liegt vor, wenn man weder nach rechts ($dx > 0$) noch nach links ($dx < 0$) ein Wachstum ($dy > 0$) erzielen kann. Das geht nur, wenn $dy = 0$. Dies zeigt, dass die Formalisierung mit Differentialen näher am intuitiven Begriffsverständnis sein kann und damit leichter sein dürfte, als die mit ihren Quotienten.

Auch klassische Optimierungsaufgaben lassen sich mit Differentialen schön interpretieren: Dazu betrachte ich die bekannte Aufgabe, aus einem Blatt von 20cm x 30cm durch Herausschneiden von vier Quadraten (Seitenlänge x) eine möglichst voluminöse oben offene Schachtel zu bauen. Das Volumen ist:

$V(x) = x \cdot (20 - 2x) \cdot (30 - 2x)$, seine Änderung:

$$dV = dx \cdot (20 - 2x) \cdot (30 - 2x) - 2x \cdot dx \cdot (30 - 2x) - 2x \cdot (20 - 2x) \cdot dx$$

Der Vorteil ist, dass sich dies inhaltlich verstehen lässt: Wenn man x vergrößert, wird die Schachtel höher, das Volumen vergrößert sich also um den ersten Summanden. Andererseits wird der Boden kleiner, und zwar in der Breite wie in der Länge und das liefert die beiden weiteren Summanden. Wenn man von $x = 0$ startet und x größer macht, wächst das Volumen ($dV > 0$) bis zu einem Punkt, an dem die Zunahme der Höhe nur noch so viel Volumen hinzufügt, wie man durch Breiten- und Längenabnahme verliert. Dann ist $dV = 0$ und daraus berechnet man die exakte Lage des Maximums wie gehabt.

3.5 Exaktifizierung

Es wurde schon eingangs festgestellt, dass die lineare Prognose in der Regel umso besser ist, je kleiner die Änderungen sind. Mit der Idee des Funktionenmikroskops, also mit dem Heranzoomen an Stellen eines Funktionsgraphen, kann diese lokal gute Passung der linearen Prognose schön beobachtet werden. Wie oben angeregt, sollten dabei schon zu Beginn auch Funktionen betrachtet werden, bei denen diese Erwartung enttäuscht wird. Beispielsweise hängt bei der Betragsfunktion $y = |x|$ an der Stelle $x = 0$ die Beziehung zwischen dx und dy davon ab, ob $dx > 0$ oder $dx < 0$. Informell definiert man dann, dass eine Funktion an einer Stelle differenzierbar heißt, wenn die Erwartung der guten linearen Prognostizierbarkeit für kleine Änderungen erfüllt wird. An dieser Stelle können die Lernenden exemplarisch erfahren, wie Mathematik als Prozess funktioniert (vgl. auch Lakatos 1979): Eine erste Definition war ausreichend gut, um eine Reihe von Schlüssen zu ziehen, kann sich aber ggf. bei neuen Fragen als unzulänglich erweisen, so dass eine Präzisierung erforderlich wird. Die Frage danach, wann eine gute lineare Prognose möglich ist, ist eine solche. Geeignete Beispiele für die Diskussion (jeweils mit der Stelle 0 im Fokus) könnten etwa die Funktion $f_1(x) = \begin{Bmatrix} 0, x < 0 \\ x^2, x \geq 0 \end{Bmatrix}, f_2(x) = \begin{Bmatrix} -\sqrt{-x}, x < 0 \\ \sqrt{x}, x \geq 0 \end{Bmatrix}$ sein.

Bis hierhin hat die vage Definition aber nicht geschadet. Es gehört zum authentischen Bild von Mathematik, dass Definitionen bei Bedarf ausgeschärft werden. Bei der Präzisierung des Begriffs des Differentials ist die entscheidende Herausforderung, genauer zu sagen, was es bedeuten soll, dass die Prognose für kleine dx besonders gut wird. Hierin steckt der Kern des Grenzwertbegriffs und dieser kann im Unterrichtsgang eingeführt werden: Ein Anstoß zur Weiterentwicklung kann durch die Betrachtung der Exponentialfunktion kommen. Es ist zwar anschaulich klar, dass für $y = a^x, a > 1$ und $dx > 0$ auch $dy > 0$ gelten muss, weil der Zusammenhang monoton steigend ist, und dass dy/dx umso größer wird, je größer x ist, aber eine genaue Bestimmung ist nicht möglich. Es soll dy eine lineare Prognose für $a^{x+dx} - a^x$ sein:

$$dy \approx \Delta y = f(x + dx) - f(x) = a^{x+dx} - a^x = a^x \cdot (a^{dx} - 1) = a^x \cdot \frac{a^{dx}-1}{dx} \cdot dx.$$

Dabei erwartet man, dass diese Prognose besonders gut wird, wenn dx sehr klein ist. Oben wurde diese Erfahrung auch schon numerisch am Beispiel der quadratischen Funktion gefestigt.

Man kann gleich die allgemeine Situation einer Funktionsgleichung $y = f(x)$ betrachten mit der Änderung $\Delta y = f(x + dx) - f(x)$. Für kleine dx werden beide Seiten der Gleichung sehr klein, für $dx = 0$ sind beide Seiten 0, die Gleichung also aussagelos. Um das zu vermeiden, dividiert man durch dx und erhält $\frac{\Delta y}{dx} = \frac{f(x+dx)-f(x)}{dx}$. Diesen Term erkundet man numerisch, so wie man es auch in anderen Zugängen zur Differentialrechnung macht. Wenn die lineare Prognose tatsächlich immer besser wird, je kleiner dx gewählt wird, dann sollte man diesen Verkleinerungsprozess systematisieren und das führt auf den Grenzwert, an dem sich durch mein Konzept keinerlei Änderung ergibt. Die Schüler können dann überprüfen, dass die bisher gefundenen Ableitungsregeln ebenfalls aus der Grenzwertdefinition folgen. Da die Schüler schon wissen, was sie als Ableitung von x^2 erwarten, sind die Schwierigkeiten, die sonst gehäuft auftreten, etwas auseinandergezogen. Am Ende der Untersuchung sollte stehen, dass für die Differentiale einer funktionalen Beziehung $y = f(x)$ gilt: $dy = \left(\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}\right) \cdot dx$.

An der weiteren Behandlung der Exponentialfunktion ergibt sich kein wesentlicher Unterschied zum üblichen Vorgehen.

Sobald an dieser Stelle (oder früher) die Ableitung $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ eingeführt ist, können Differentiale auch so charakterisiert werden: Zu jeder funktionalen Beziehung $y = f(x)$ definiert man weitere Variablen $dx, dy \in \mathbb{R}$ mit der Beziehung $dy = f'(x)dx$ oder allgemeiner für $g(x, y, z) = 0$ soll gelten $\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy + \frac{\partial g}{\partial z}dz = 0$. Dadurch schließt man nahtlos an den Kalkül an, wo man etwa zu $g(x, y) = x^2 + y^2 - 1 = 0$ rechnet $2xdx + 2ydy = 0$.

3.6. Didaktische Betrachtungen zur Lehre mit Differentialen

Als Zusammenfassung der bisherigen fachlichen Analyse kann festgestellt werden: Die im oben skizzierten Unterrichtsgang an den Anfang gestellte Grundvorstellung der bestmöglichen linearen Prognose muss im Laufe der weiteren Entwicklung der Analysis von den hier skizzierten Anfängen bis zum universitären Niveau nie grundlegend revidiert, sondern nur erweitert und exaktifiziert werden. Sie kann auf zwei Arten präzisiert werden, als Frechetdifferentiale oder, wie im skizzierten exemplarischen Unterrichtsgang vorgeschlagen, durch die Definition der Ableitung als Grenzwert des Differenzenquotienten.

In diesem Abschnitt folgt nun basierend auf dieser fachlichen Klärung ein Blick auf didaktische Überlegungen zu Differentialen in der Lehre.

Die Vorteile der Auffassung von Differentialen als endliche Größen für den Lernprozess sind von vielen Hochschullehrenden benannt worden. In Apostol et al. (1969) finden sich mehrere Beiträge, die – wenn auch mit leichten Variationen untereinander – Definitionsmöglichkeiten für Differentiale darstellen. Der entscheidende Unterschied zum Vorgehen in der vorliegenden Arbeit ist, dass ich hier versuche, eine Grundvorstellung von Differentialen *vor* der Formalisierung zu erreichen. Antrieb ist dabei u.a. die Hoffnung, Differentiale frühzeitig in der Modellbildung einsetzen zu können. Auch Blum (1975, S, 291) hat für die Verwendung von Differentialen in der Schule plädiert, indem er u.a. schreibt „Sie ist insbesondere bei Anwendungsproblemen äußerst zweckmäßig.“ In heutiger didaktischer Sprache bedeutet das, dass die Verwendung von Differentialen die Kompetenz des Modellierens stärken kann. Gleichzeitig warnt er: „Sprechweisen wie ‚Differentiale sind unendlich kleine Größen‘ sollten jedoch der Vergangenheit angehören.“ (ebda. S. 295). Er plädiert daher dafür, die Ableitung als Grenzwert einzuführen und im Anschluss die Vorstellung der linearen Approximation zu entwickeln. – ob dies mit Differentialen formuliert werden soll, wird bei ihm allerdings nicht klar.

Ich gebe hier einen Überblick über verschiedene Überlegungen zur Verwendung von endlichen Differenzialen in der schulischen und universitären Lehre. Schon vor fast 90 Jahren hat Walther (1929) eine Darstellung von Differentialen als endliche Größen bezogen auf die lineare Approximation gegeben. Dabei werden Differentiale im Haupttext als abgeleitete Größen betrachtet, d.h. es wird angenommen, dass die Ableitung einer Funktion schon bekannt ist und mit ihrer Hilfe werden die Differentiale definiert. In einem abschließenden Abschnitt „Unterrichtsfragen“ vertritt Walther dann den Standpunkt, dass für den Unterricht an der Schule Zugänge, die den Grenzwert an den Anfang stellen, problematisch seien, weil sie die Lernenden sehr belasten würden. Er plädiert für einen (wie er selbst schreibt) radikal geometrischen Zugang: Man zeichnet intuitiv Tangenten ein, untersucht diese und gewinnt danach die Differentiale als Katheten von Steigungsdreiecken. Walther löst sich damit vom Formalismus, aber nicht von der fachlichen Reihenfolge, erst Ableitung, dann Differentiale.

Noch stärker an den Anfang gerückt werden Differentiale in der Unterrichtskonzeption von David Tall (2012 und Referenzen darin). Er konzipierte und erprobte das Konzept des „locally straight“ Zugangs zur Analysis unter Nutzung des Funktionsmikroskops. Zentral ist der Aufbau der Vorstellung, dass glatte Kurven lokal gut durch eine Gerade, die Tangente, approximiert werden, und er definiert dx, dy als Komponenten eines (beliebig langen) Tangentenvektors. Oben wurde schon gezeigt, dass diese Sichtweise von meinem Zugang aus zugänglich ist; es ist gewissermaßen der geometrische Teil meines Zugangs. Gemeinsam ist beiden Zugängen der Verzicht auf eine präzise Definition zu Beginn und stattdessen das Aufsetzen auf intuitiven Vorerfahrungen. Der größte Unterschied ist die Integration der Modellierungsperspektive im hier vorgeschlagenen Zugang. Die Konzeptualisierung als Prognosen erschließt einige außermathematische Änderungsprozesse leichter als der eher innermathematisch-geometrische Weg von Tall, der eine Darstellung des betrachteten Zusammenhangs als Funktionsgraph erfordert. Zur Illustration sei das Leiterbeispiel aus dem ersten Abschnitt erneut aufgegriffen. Will man eine Prognose abgeben, wie sich der Mittelpunkt (x_m, y_m) der Leiter bewegen wird, so legt eine Ähnlichkeitsbetrachtung der Situation in Abb. 1 nahe: $dx_m = \frac{da}{2}, dy_m = \frac{db}{2}$. (Alternativ wendet man den ersten und den zweiten Strahlensatz mit Zentrum $(a, 0)$ an und erhält $\frac{a}{x_m} = \frac{b}{y_m} = \frac{a}{\frac{a}{2}}$=2, also $a = 2x_m, b = 2y_m$ und Differentiation liefert die genannten Beziehungen.) Möchte man dagegen den Weg über Tangenten an Funktionsgraphen gehen, muss zunächst ein passender Graph gefunden werden – je nach Wahl der Größe, die auf der Rechtsachse aufgetragen wird, kann das einige Arbeit verursachen.

Auf (sehr) elementarem Hochschulniveau gibt es ferner das Buch von Thompson u. Gardner (1998), in dem Lernenden die sprachliche Metapher an die Hand gegeben wird, d*x* bedeute „a little bit of *x*“. Die Linearität wird dort aus der Kleinheit der Änderung gefolgert (nämlich aus der Überlegung zur Größenordnung, dass für sehr kleine dx das Quadrat dx^2 noch viel kleiner ist und deswegen vernachlässigt wird, so dass z.B. $(x + dx)^2 - x^2 = 2xdx(+dx^2)$ linear, also ohne den eingeklammerten Term genähert wird). Nachteil ist, dass die Frage, was man vernachlässigen kann, nicht so klar ist. Sicher aber ist die „eine kleine Menge von“ eine für vieles brauchbare Grundvorstellung zum Differentialoperator. Diese Grundvorstellung kann in meinem Ansatz wie folgt erschlossen werden: Man betrachtet eine Größe, z.B. die Masse m der Gewichte auf einer Seite einer Waage. Ändert sich diese Masse um einen kleinen Betrag dm von m auf $m + dm$, dann ist dm ein klein bisschen vom Endzustand. Weil der sich aber nicht viel vom Anfangszustand unterscheidet, kann man auch sagen, es sei ein bisschen vom Ausgangszustand m. Besondere

Bedeutung hat diese Sichtweise bei Dichten, etwa bei Wahrscheinlichkeitsdichten: $dp = \rho(x)dx$ ist der kleine Teil der Gesamtwahrscheinlichkeit, der auf Werte der Zufallsvariablen entfällt, die zwischen x und $x + dx$ liegen. Wenn man die gesamten kleinen Stücke akkumuliert erhält man das Ganze: Das Integral konstruiert also den Bestand aus den Änderungen.

Eher exotisch dürfte der Zugang von Edwards (1994) sein. Edwards setzt eine auf die physikalische Anschauung gestützte intuitive Formulierung von 1-Differentialformen als Duale zu Vektoren an den Anfang („a rule that assigns numbers to displacements“, S. 1, konkret also eine Zuordnung, die einer Verschiebung die nötige Arbeit zuordnet) und bleibt dann lange in der rein algebraischen Sphäre. Die Ableitung bekommt zunächst einen unscheinbaren Auftritt als Wert des Korrekturfaktors beim Pullback einer Form von einem Raum in einen anderen. Für lineare Abbildungen ist dies elementaralgebraisch bestimmbar. Approximative Überlegungen und Grenzwerte kommen erst bei der Integration nicht-konstanter Formen vor. Dieser Vorschlag ist hochgradig originell. Seine algebraische Komplexität zahlt sich allerdings nur aus, wenn man tief in die Theorie einsteigt. Aus diesem Grund erscheint er für das Gymnasium wenig geeignet.

3.7. Überlegungen zu einer möglichen unterrichtlichen Umsetzung

Die obigen Ausführungen bilden schon den Kern einer möglichen Unterrichtssequenz. Hier werden noch einige ergänzende spekulative Überlegungen zur Implementation angestellt.

Es kann angenommen werden, dass die Lernenden die Bildungsstandards für den mittleren Schulabschluss erfüllen. Insbesondere sollten die Lernenden die Erfahrung gemacht haben, dass es beim Modellieren (etwa bei der Volumenermittlung von komplexen Körpern) zweckmäßig sein kann, ein einfaches Modell zu verwenden, auch wenn einem bewusst ist, dass man dadurch einen Modellierungsfehler macht. Das erklärte Ziel der Einheit ist dann, Änderungsprozesse möglichst einfach zu prognostizieren. Die skizzierten Einstiegsaktivitäten sollen die Beschäftigung mit Größen und ihren Veränderungen initiieren. Damit man sinnvolle Prognosen abgeben kann, benötigt man ein Verständnis der Prozesse und/oder der Beziehungen zwischen Relationen. Diese relationale Sicht, insbesondere auch die Sicht, dass Funktionen spezielle Relationen sind, muss zu Beginn hinreichend klar werden. Funktionales Denken spielt ebenfalls eine wichtige Rolle, aber keine exklusive. Betrachtet man etwa als (relativ simples) Beispiel den Zusammenhang von Seitenlänge a und Flächeninhalt A eines Quadrats, so kann man Prognosen für A erstellen, wenn sich a ändert, oder umgekehrt für a wenn sich A ändert. Beides ergibt sich aus der Beziehung $A = a^2$. Die Bedeutung von Relationen sowohl als Modellierungsmittel als auch als theoretisches Konzept ist in einem Unterricht mit Differentialen hoch. Ein Mathematisierungsmuster, das dies unterstützt, ist es, Wissen durch Gleichungen auszudrücken. Wenn man sich etwa für die Ableitung von Quotienten interessiert, ist es sinnvoll, diese Struktur als Gleichung hinzuschreiben $f = \frac{u}{v}$ und äquivalent, aber strukturell vereinfachend, umzuformen zu $f \cdot v = u$. Diese Strategie könnte helfen (zusammen mit den verhältnismäßig einfachen Herleitungen der Regeln des Differentialkalküls), dass Lernende einen größeren Anteil der Mathematik selbst erarbeiten können, als dies sonst häufig der Fall ist.

3.8 Didaktische Verortung

Die Planung von Unterricht mit Differentialen muss auch bedenken, wie sich den so unterrichteten Lernenden Konzepte erschließen, die in anderen Zugängen zentral sind, so dass sie in die Lage versetzt werden, beispielsweise Vorlesungen, die eher an traditionellen Ansätzen orientiert sind, zu folgen. Weiter ist zu prüfen, ob mit ihm die Vorgaben der Bildungsstandards der KMK (2012) erfüllt werden können. Diese fordern, dass primär das Verständnis der Änderungsraten am Ende bei den Schülern als Output vorhanden ist. Da nicht eindeutig festgelegt wird, wie die Aneignung der Kompetenzen vonstattengeht, ist zu prüfen, ob man auf dem alternativen Weg über Differentiale auch zu einem Verständnis von Änderungsraten, wie es die Bildungsstandards fordern, kommen kann. Eingangs wurde dargelegt, dass ein Vorteil von Differentialen darin besteht, Änderungen zu betrachten, nicht Quotienten von Änderungen. Sobald eine gewisse Vertrautheit im Umgang mit Änderungen erreicht ist, spricht vieles dafür, dass man dividiert und damit auch Änderungsratenaufgaben bearbeitet. Aus der Beziehung $\Delta y \approx dy = m \cdot dx$, die ausdrückt, dass eine lineare Prognose gemacht wird, die lokal gut sein soll, folgt $\frac{\Delta y}{dx} \approx \frac{dy}{dx} = m$. Der Quotient der linearen Änderungsprognosen zweier durch einen funktionalen Zusammenhang verbundener Größen ist also eine Zahl, je nach Interpretation die Proportionalitätskonstante der zueinander proportionalen Änderungsprognosen dx, dy, der Verstärkungsfaktor für linearisierte Änderungen, die Tangentensteigung oder die lokale Änderungsrate, die für kleine Änderungen nahe bei der mittleren Änderungsrate (auf der linken Seite des Ungefährzeichens) liegt. Die Deutung am Graphen vernetzt dieses Wissen: Der Graph ist lokal linear. Diese Sichtweise der Differenzierbarkeit als lokale Linearisierbarkeit (und der Ableitung als Steigung der bestpassenden Geraden) wird zwar in der didaktischen Literatur gleichrangig mit der Idee der Änderungsrate genannt (z.B. Danckwerts und Vogel 2006), spielt in der Praxis aber kaum eine Rolle. Kern ist die Erkenntnis, dass bestimmte Funktionen, nämlich die differenzierbaren, im Funktionsmikroskop um jeden Punkt $(x_0, f(x_0))$ des Graphen hinreichend vergrößert geradlinig erscheinen, so dass sie lokal gut durch eine Gerade approximiert werden können: $y = f(x) \approx f(x_0) + m \cdot dx, \Delta y \approx dy = m \cdot dx$. Die Steigung, die das optimal leistet, ist die Ableitung $m = f'(x_0)$.

Lokale Linearisierung wird in Tall (2012) basierend sowohl auf empirischen wie auch auf konzeptuellen Überlegungen als hervorragendes Leitprinzip für den Unterricht herausgehoben. U.a. gehört zu den Vorteilen, dass die Schüler die Weiterentwicklung von der in früheren Schuljahren erworbenen Vorstellung, dass eine Tangente in einem Punkt berührt, besser leisten. Das zeigt sich laut Tall beispielsweise darin, dass die Tangente in (0,0) an den Graphen von $f(x) = x, x < 0; f(x) = x + x^2, x \geq 0$ bei Schülern, die nach einem lokalen Linearitätskonzept unterrichtet wurden, häufiger korrekt eingezeichnet wird als bei Vergleichsgruppen. Tall führt als Erklärung für diesen Zusammenhang an, dass die Vorstellung der lokalen Linearität gut zusammenpasst mit der sensorischen Wahrnehmung von lokaler Glattheit – sie schließt damit unmittelbar an eine grundlegende körperliche Erfahrung an (dass Differentiale damit Legitimation auch aus den Erkenntnissen der embodied mathematics (Lakoff und Nunez 2000) beziehen, ist – trotz der Kritik an Lakoff und Nunez – eine wichtige Erkenntnis!). Differentiale sind somit ein gutes Mittel, die Idee der lokalen Linearisierung zum Ausdruck zu bringen. Sobald diese Idee entwickelt ist, kann man Quotienten betrachten und mit $\Delta y / \Delta x$ und dy/dx die mittlere und lokale Änderungsrate einführen. Einige der üblichen Aufgaben zur Anwendung der lokalen Änderungsrate erschließen vom Standpunkt der lokalen Linearisierung sich noch besser. Ein Beispiel ist die Frage, wie das Volumen einer Kugel wächst, wenn man eine dünne Lackschicht der Dicke Δh aufträgt. Im Sinne der lokalen Linearisierung ist sofort klar, dass der Volumenzuwachs gleich dem Produkt aus der Kugeloberfläche und Δh ist.

In Tietze et al. (2000, S. 258f) wird das Konzept der lokalen Linearisierung sehr kritisch gesehen. Die dortigen Einwände sind allerdings entweder in sich fraglich oder gelten nicht bei der Umsetzung mit Differentialen: Das erste Argument, die Bedeutung der Linearität im Mathematikunterricht solle nicht überschätzt werden, hängt von subjektiven Gewichtungen ab; das zweite, wonach lokale Linearisierung vor allem dann sinnvoll ist, wenn multivariate Fragestellungen betrachtet werden, habe ich durch ein entsprechendes Beispiel (rutschende Leiter) schon entkräftet. Das dritte Argument behauptet, die Beweise der Kalkülregeln seien nicht einfacher als traditionelle Versionen, und wird m.E. aus konzeptueller Ebene durch die obigen Betrachtungen entkräftet (neben der konzeptuellen Schwierigkeit wäre noch die empirische Schwierigkeit zu klären, das leistet diese Arbeit nicht). Der vierte Punkt schließlich behauptet einen geringen Gewinn durch die Strukturähnlichkeit von Stetigkeit und Differenzierbarkeit – auf diesen Gewinn hofft das vorgestellte Konzept erst gar nicht.

4 Fazit

Der Beitrag hat durch stoffdidaktische Analyse gezeigt, wie Differentiale so intuitiv definiert werden können, dass auf dieser Basis Theorieentwicklung und Argumentation möglich ist, und andererseits ein Exaktifizieren und Anschließen an die Standardbegrifflichkeit geleistet werden kann. Besonderes Potential haben Differentiale beim inner- und außermathematischen Modellieren, weil mit ihnen Beziehungen nicht zwischen Raten (also Bruchtermen) sondern direkt zwischen Änderungen beschrieben werden können. Beispiele sind in Oldenburg (2015) zusammengestellt. Dieser Beispielkatalog zeigt insbesondere die gute Eignung der Differentiale zum Modellieren. Dabei zeigt sich auch, dass Computeralgebrasysteme gute Hilfsmittel sind. Dies liegt daran, dass Differentiale eine anwendungsnahe Möglichkeit sind, Zusammenhänge zu algebraisieren.

Gerade an dieser Leistung mag aber auch eine kritische Würdigung ansetzen: Durch die starke Verlagerung der Argumentation ins Algebraische wird das Grenzwertkonzept, obwohl es eine zentrale Idee der Analysis ist, eher gering gewichtet. Der Beschäftigung mit der Idee des Grenzwertes kommt ein hoher Bildungswert zu. Dennoch scheint mir dieser Punkt kein großer Mangel meines Konzepts zu sein. Im Gegenteil: Dass die Notwendigkeit entfällt, die Konzepte des Grenzwertes und der Ableitung gleichzeitig einzuführen, ist für die Gestaltung der Lernprozesse eher vorteilhaft. Entweder man behandelt vorab Grenzwerte von Folgen (etwa als langfristige Prognosen bei Wachstumsprozessen), oder im Nachhinein wie oben skizziert bei der Exaktifizierung und Ausdehnung der Differentialrechnung auf die Exponentialfunktion. Der nächste Schritt nach dieser stoffdidaktischen Durcharbeitung sollte eine Erprobung in der Praxis sein. Bisher gibt es dazu noch keine Erkenntnisse, da die Erprobung durch die aktuellen Lehrpläne erschwert wird, auch wenn die grundlegenden Ziele verträglich sind.

Danksagung

Mein Dank gilt Christian Groß, Andreas Merkel, Renate Motzer, Samuel Pfeifer, Sabrina Scheffler, Wolfgang Schneider, Benedikt Weygandt für Diskussionen und das Lesen bestimmter Fassungen des Manuskripts und den Gutachtern für umfangreiche Kommentare.

Literatur

Apostol, T. M. et al. (1969). *Selected Papers on Calculus*. Belmont: AMS.

Artigue, M. (1991). Analysis. In D. Tall (Hrsg.) *Advanced Mathematical Thinking* (S. 167-198). Dordrecht: Kluwer.

Becker, J. P., Shimada, S. (1997). *The Open-Ended Approach: A New Proposal for Teaching Mathematics.* Reston: The National Council of Teachers of Mathematics, Inc.

Bishop, E. (1977). Review: H. Jerome Keisler, Elementary calculus. *Bull. Amer. Math. Soc*. 83, 205–208.

Bossel, H. (2004). *Systemzoo.* Bände 1-3, Norderstedt: BoD.

Blum, W. (1975). Ein Grundkurs in Analysis für die berufliche Oberstufe. In H. Postel et al. (Hrsg.). *Mathematik lehren und lernen* (S. 290-301), Hannover: Schroedel.

Blum, W., Kirsch, A. (Hrsg.) (1979). Anschaulichkeit und Strenge in der Analysis IV. *Der Mathematikunterricht* 25 (3).

Blum, W., Kirsch, A. (1979). Zur Konzeption des Analysisunterrichts in Grundkursen. *Der Mathematikunterricht* 25, (1), S. 6-24.

Blum, W., Vogel, S., Drüke-Noe, Ch., Roppelt, A. (Hrsg.). (2015). *Bildungsstandards aktuell: Mathematik in der Sekundarstufe II*. Hannover: Schroedel.

Courant, R., Robbins, H. (1973). *Was ist Mathematik*? Berlin: Springer.

Danckwerts, R., Vogel, D. (2006). *Analysis verständlich unterrichten*. Heidelberg: Springer.

Dray, T., Manogue, C. A. (2010). Putting Differentials Back into Calculus. *The College Mathematics Journal* 41(2) , S. 90-100.

Edwards, H. (1994). *Advanced Calculus*. Boston: Birkhäuser.

Epple, M. (1999). Das Ende der Größenlehre: Grundlagen der Analysis 1860-1910. In H. N. Jahnke (Hrsg.). *Geschichte der Analysis* (S. 371-410). Heidelberg: Spektrum.

Feudel, F. (2015). Die Ableitung als Absolute Änderung? In F. Caluori et al. (Hrsg.). *Beiträge zum Mathematikunterricht 2015* (S. 1049-1052), Münster: WTM-Verlag..

Freudenthal, H. (1979). *Mathematik als pädagogische Aufgabe*. Stuttgart: Klett.

Godement, R. (2004). *Analysis I*. Berlin: Springer.

Götz, H. et al. (Hrsg.) (2009). *Lambacher Schweizer 11 – Bayern*. Stuttgart: Klett.

Gray, S. S., Loud, B. J., & Sokolowski, C. P. (2005). Undergraduates' errors in using and interpreting algebraic variables: A comparative study. In G. M. Lloyd, M. R. Wilson, J. L. Wilkins, & S. L. Behm (Eds.), *Proceedings of the 27th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* [CD-ROM]. Eugene, OR: All Academic.

Hahn, S. (2008). *Bestand und Änderung*. Oldenburg: Didaktisches Zentrum.

Hairer, E., Wanner, G. (1996). *Analysis by its history*. Berlin: Springer.

Heuser, H. (1988). *Lehrbuch der Analysis.* Band 2. Stuttgart: Teubner.

Kirsch, A. (1979). Ein Vorschlag zur visuellen Vermittlung einer Grundvorstellung vom Ableitungsbegriff. *Der Mathematikunterricht* 25(3), 25-41.

Kirsch, A. (1977). Aspects of Simplification in Mathematics Teaching. In H. Athen, H. Kunle (Hrsg.). *Proceedings of the Third International Congress on Mathematical Education* (S. 98-120), Karlsruhe.

Kirsch, A. (1997). *Mathematik wirklich verstehen*. Köln: Aulis.

KMK (2012). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife* (Beschluss der Kultusministerkonferenz vom 18.10.2012). http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf. Gesehen am 5.4.2016.

Laugwitz, D., Schnitzspan, W. (1983): Nichtstandard-Analysis. *Der Mathematikunterricht*, 29(4), S. 37-59.

Lergenmüller, A., Schmidt, G. (Hrsg.) (2009). *Neue Wege 10 Niedersachsen*. Hannover: Schroeder.

Loomis, L. H., Sternberg, S. (1968). *Advanced Calculus*. Reading: Addison-Wesley.

Keisler, H. J. (1976). *Elementary Calculus: An Approach Using Infinitesimals*. Prindle, Weber & Schmidt.

Lakatos, I. (1979). *Beweise und Widerlegungen*. Wiesbaden: Vieweg.

Lakoff, G., Núñez, R. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.

Lengnink, K. (2005). Abhängigkeiten von Größen - zwischen Mathematikunterricht und Lebenswelt. *Praxis der Mathematik in der Schule* 47(2), 13-19.

López-Gay, R., Sáez, J. M., & Torregrosa, J. M. (2015). Obstacles to mathematization in physics: The case of the differential. *Science and Education* 24, 591-613.

Malle, G. (1993). *Didaktische Probleme der elementaren Algebra*. Wiesbaden: Vieweg.

Morris, C. W. (1938). *Foundations of the Theory of Signs*. Chicago: University of Chicago Press.

Oldenburg, R. (2012). Bewegungsvorgänge lokal betrachtet. *Praxis der Mathematik* 54(44), 25-28.

Oldenburg, R. (2012b). Lineare Prognosen von Bewegungen. http://myweb.rz.uni-augsburg.de/~oldenbre/dyna.html. Gesehen am 5.4.2016.

Oldenburg, R. (2007). Experimentell zum Ableitungsbegriff. *mathematik lehren*, Heft 141, 52-56.

Oldenburg, R. (2015). Analysis mit Differentialen: Anwendungsbeispiele. http://myweb.rz.uni-augsburg.de/~oldenbre/DifferentialeOnlineMaterialien.pdf. Gesehen am 5.4.2016.

Picher, F. (2010). Änderungen besser verstehen - Mathematik besser verstehen. In M. Helmerich et al. (Hrsg.). *Mathematik verstehen - Philosophische und didaktische Perspektiven*. Braunschweig: Vieweg.

Pickert, G. (1962). Differentiale. *Der Mathematische und Naturwissenschaftliche Unterricht* 15 (4), 145-150.

Quine, W. v.O. (1951). Two dogmas of empiricism. *Philosophical Review* 60, 20. Abgedruckt in W.v.O. Quine. *From a Logical Point of View*. Harvard University Press, Cambridge, Mass, 1953, S. 20-46.

Quine, W.v.O. (1960). Variables explained away. *Proc. Of the American Philosophical Society*, 104 (3) , 343-347.

Robinson, A (1966). *Nonstandard Analysis*.. Amsterdam: North-Holland.

Rogers, R. (2005) Putting the Differential Back into Differential Calculus. In A. Shell-Gellasch et al. *From Calculus to Computers* (S. 9-20). Washington: MAA.

Sonar, Th. (2011). *3000 Jahre Analysis*. Berlin: Springer.

Stachowiak, H. (1973). *Allgemeine Modellthorie*. Wien: Springer.

Steinbring, H. (2011). Changed Views on Mathematical Knowledge in the Course of Didactical Theory Development: Independent Corpus of Scientific Knowledge or Result of Social Constructions? In T. Rowland und K. Ruthven (Hrsg.). *Mathematical Knowledge in Teaching* (S. 43-64). Heidelberg - London - New York, Springer.

Sträßer, R. (2014). Stoffdidaktik in Mathematics Education. In St. Lermann: *Encyclopedia of Mathematics Education*. (S. 566-570). New York: Springer.

Sullivan, K. (1976). The Teaching of Elementary Calculus Using the Nonstandard Analysis Approach. *The American Mathematical Monthly* 83, (5), 370-375.

Tall, D. (2012). *How Humans learn to think mathematically*. Cambridge: Cambridge.

Thomas, B. Th., Weir, M. D., Hass, J. (2013). *Basisbuch Analysis*. München: Pearson.

Thompson, S., Gardner, M. (1998). *Calculus made easy*. New York: CreateSpace.

Thompson, P. W., Byerley, C., & Hatfield, N. (2013). A conceptual approach to calculus made possible by technology. *Computers in Schools*, 30, 124-147.

Tietze, U.-P. et al. (2000). *Mathematikunterricht in der Sekundarstufe II*. Band 1. Wiesbaden: Vieweg.

Torregrosa, J. M., Lopez-Gay, R. (2006). Mathematics in Physics Education: Scanning Historical Evolution of the Differential to Find a More Appropriate Model for Teaching Differential Calculus in Physics. *Science & Education* 15, 447–462.

Walther, A. (1929). *Begriff und Anwendungen des Differentials*. Leipzig: Teubner.

Weber, Ch. (2007). *Mathematische Vorstellungen bilden*. Bern: h.e.p. Verlag.

Witzke, I. (2010). *Die Entwicklung des Leibnizschen Calculus*. Hildesheim: Franzbecker.

Wunderling, H. (Hrsg.). (2007). A*nalysis als Infitesimalrechnung*. Berlin: Duden Paetec.