Mathematik für Wirtschaftswissenschaftler

Vorlesungsprogramm für den 26. 04. und 03. 05. 2007

(K. Steffen, Heinrich-Heine-Universität Düsseldorf, SS 2007)

4.3 Die Ableitung, Definition und Rechenregeln

Die Differentialrechnung ist ein sehr leistungsfähiger Kalkül zur Berechnung der Änderungsrate (genannt "Ableitung" $f'(x)$) von elementaren Funktionen $f(x)$. Die wichtigsten Anwendungen sind:

- *Extremstellenbestimmung bei $f(x)$ mittels Nullstellenbestimmung bei $f'(x)$;*
- *Bestimmung des Monotonieverhaltens bei $f(x)$ mit Vorzeichenuntersuchung bei $f'(x)$;*
- *Bestimmung des Konvexitätsverhaltens (Progressivität oder Degressivität der Monotonie) von $f(x)$ mittels Monotonieuntersuchung bei $f'(x)$ bzw. mittels Vorzeichenuntersuchung bei der zweiten Ableitung $f''(x)$.*

Dies ist von offensichtlicher Relevanz für die Ökonomie: Die Wichtigkeit der Extremstellenbestimmung (z.B. Gewinnmaximierung) liegt auf der Hand, und natürlich ist es auch wichtig, von ökonomische Variablen zu wissen, ob sie wachsen oder fallen und ob das in progressiver oder degressiver Weise geschieht. Wenn man ökonomische Variablen mit empirisch oder heuristisch bekanntem Monotonie- und Konvexitätsverhalten mathematisch modellieren will, so muss man dies natürlich durch elementare Funktionen mit demselben Verhalten tun. Also braucht man auch wirkungsvolle mathematische Hilfsmittel, um solche Eigenschaften nachzuweisen. Diese Hilfsmittel liefert die Differentialrechnung, indem sie die genannten Aufgaben auf viel einfachere Aufgaben wie die Nullstellenbestimmung und die Vorzeichenuntersuchung bei der Ableitung zurückführt. Damit man sich dieser Hilfsmittel bedienen kann, ist es natürlich erforderlich, dass man die Ableitung zu einer gegebenen elementaren Funktion ausrechnen kann. Davon handelt dieser Abschnitt. Die Anwendungen der Differentialrechnung folgen dann in 4.5 und 4.6.

Die Grundidee beim Konzept der Ableitung einer reellen Funktion $f(x)$ von einer Variablen ist, dass man nicht über die Größe ihrer Werte, sondern über die *Änderung* der Werte eine quantitative Aussage machen möchte. Dazu fixieren wir eine Stelle x_0 im Definitionsbereich und betrachten eine positive oder negative *Argumentänderung* $h \neq 0$, die uns von x_0 zu einer anderen Stelle $x = x_0 + h$ im Definitionsbereich bringt. Die entsprechende *Änderung der abhängigen Variablen* ist dann $f(x_0 + h) - f(x_0) = f(x) - f(x_0)$. Nun bewirken unterschiedlich große Änderungen des Arguments im Allgemeinen auch unterschiedliche Änderungen der Funktionswerte, also ist die Angabe der Änderung der Funktionswerte für sich keine aufschlussreiche Information über die Funktion, wenn man sie nicht ins Verhältnis zur Änderung der Argumente setzt. Man betrachtet daher die **Änderung der Funktionswerte relativ zur Änderung des Arguments**

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{f(x) - f(x_0)}{x - x_0} .$$

Dies wird auch der **Differenzenquotient** der Funktion f an der Stelle x_0 zur Schrittweite h genannt und oft $\frac{\Delta f}{\Delta x}$ notiert, wobei Δx ("Delta x") für die Differenz $x - x_0 = h$ der unabhängigen Variablen steht und Δf für die Differenz der Funktionswerte, also für $f(x_0+h) - f(x_0) = f(x) - f(x_0)$.

Um nun die Änderungsrate der Funktion an der Stelle x_0 zu beschreiben, muss man sich auf Änderungen $h = x - x_0$ von sehr kleinem Betrag beschränken — schließlich wird sich die Funktion an anderen Stellen mit anderer Geschwindigkeit ändern (z.B. könnte sie ja progressiv wachsen). Mathematisch bedeutet das, dass man zum Limes der Differenzenquotienten beim Grenzübergang $h \to 0$ übergehen muss. Wir setzen aber die Beherrschung des Grenzwertbegriffs nicht voraus (das ist für das Erlernen der Differentialrechnung nicht nötig, und historisch hat sich der exakte Grenzwertbegriff auch viel später herausgebildet als die Differentialrechnung), sondern beschreiben in der folgenden Definition explizit, was die Konvergenz der Differenzenquotienten bedeutet:

DEFINITION: Ist $f : I \to \mathbb{R}$ eine reelle Funktion auf einem Intervall (positiver Länge) und $x_0 \in I$, so heißt f **differenzierbar an der Stelle x_0**, wenn die Differenzenquotienten $\frac{1}{h}[f(x_0 + h) - f(x_0)]$ einen Limes $c \in \mathbb{R}$ haben bei $h \to 0$, d.h. wenn für jede gegebene Fehlerschranke $\varepsilon > 0$

$$\left| \frac{f(x_0 + h) - f(x_0)}{h} - c \right| < \varepsilon \quad \text{ist für hinreichend kleinen Betrag} \quad |h| > 0 \, ,$$

wobei natürlich nur Werte h mit $x_0 + h \in I$ zugelassen sind. Die Zahl c heißt dann die **Ableitung** oder **Differentialquotient** (auch *Änderungsrate, Änderungsgeschwindigkeit* oder *Steigung*) der Funktion f **an der Stelle x_0** und wird notiert

$$f'(x_0) \, , \quad \frac{df}{dx}(x_0) \, , \quad \left. \frac{d}{dx} \right|_{x=x_0} f(x) \, , \quad \text{etc.}$$

Ist f an jeder Stelle $x \in I$ differenzierbar, so heißt $f'(x)$ die **Ableitungsfunktion** (oder einfach die **Ableitung**) zu f **auf I**. Die Berechnung der Ableitung nennt man auch **Differenzieren** der Funktion. Ist f' selbst wieder differenzierbar auf I, so heißt die Ableitung von f' die **zweite Ableitung** oder Ableitung zweiter Ordnung von f und wird notiert $f'' = \frac{d^2}{dx^2} f(x)$. Analog erklärt man die dritte Ableitung f''' und allgemein für $n \in \mathbb{N}$ die **Ableitung n–ter Ordnung** $f^{(n)}(x) = \frac{d^n}{dx^n} f(x)$ auf I. ■

DISKUSSION: **1)** *Interpretation in den Wirtschafts– (und anderen) Wissenschaften*: In der Praxis kann man die Ableitung $f'(x_0)$ an einer Stelle x_0 und die dagegen konvergierenden Differenzenquotienten $\frac{1}{h}[f(x_0 + h) - f(x_0)]$ als gleich ansehen, wenn $|h| = |x - x_0|$ klein ist. In diesem Sinne können wir also sagen:

- *Die Ableitung $f'(x_0)$ gibt für kleine Änderungen von x_0 die Änderung des Funktionswertes relativ zur Änderung des Arguments an;*
- *das Vorzeichen der Ableitung $f'(x_0)$ beschreibt die Änderungsrichtung des Wertes von f in x_0 bei kleiner Vergrößerung von x_0.*

Die Änderung der Funktionswerte relativ zur Änderung des Arguments ist $f(x) - f(x_0)$ dividiert durch $x - x_0$, also der Differenzenquotient. Wenn $f'(x_0) > 0$ ist, so sind auch die Differenzenquotienten zu x nahe x_0 positiv, d.h. es gilt dann $f(x) > f(x_0)$ für $x > x_0$ und $f(x) < f(x_0)$ für $x < x_0$ (und x nahe bei x_0). Ist $f'(x_0) < 0$, so verhält es sich umgekehrt. In diesem Sinne bestimmt das Vorzeichen von $f'(x_0)$ die Änderungsrichtung, d.h. die Funktionswerte werden bei geringfügiger Vergrößerung von x_0 größer, wenn $f'(x_0)$ positiv ist, und sie werden kleiner, wenn $f'(x_0)$ negativ ist. (Im Fall $f'(x_0) = 0$ ändern sich die Funktionswerte bei kleiner Variation von x_0 “extrem wenig”; man sagt, dass die Funktion dann “in erster Näherung konstant” gleich $f(x_0)$ ist nahe x_0. Über die Richtung der möglichen sehr kleinen Änderungen der Funktionswerte läßt sich aber keine allgemein gültige Aussage machen.)

Im ökonomischen Kontext ist der Wert x_0 der unabhängigen Variablen häufig eine große Zahl (von Einheiten), so dass eine Erhöhung um 1 (um eine Einheit) als kleine Änderung angesehen werden kann. Unter dieser Prämisse kann man sagen:

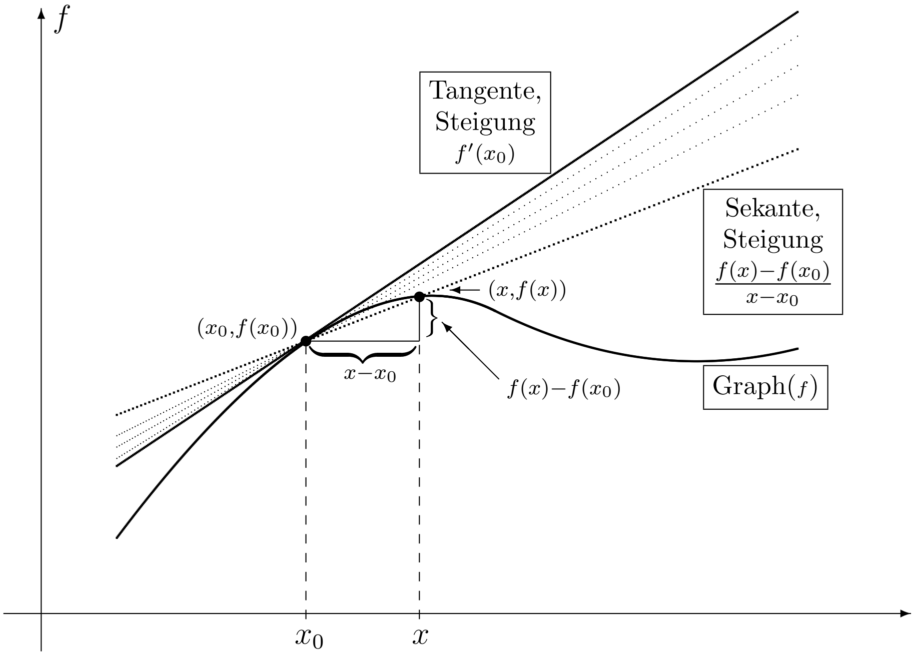
- *Der Ableitungswert $f'(x_0)$ gibt (ungefähr) die Änderung $f(x_0{+}1)-f(x_0)$ der Funktionswerte an, die durch eine Erhöhung der unabhängigen Variablen um eine (kleine) Einheit bewirkt wird.*

Genau stimmt das freilich nur bei affin linearen Funktionen $f(x) = ax + b$. Und damit es "ungefähr" stimmt, ist Voraussetzung, dass die Erhöhung um eine Einheit wirklich nur eine *kleine* Änderung der unabhängigen Variablen bedeutet. (Genauer gesagt darf sich die Änderungsrate von f innerhalb des Intervalls $[x_0, x_0 + 1]$ nur wenig ändern; schließlich kann man $f(x_0{+}1) - f(x_0)$ mit gleichem Recht als Näherungswert für $f'(x_0{+}1)$ ansehen, und wenn $f'(x_0{+}1)$ zu sehr von $f'(x_0)$ abweicht, so ist die Interpretation der Differenz $f(x_0{+}1) - f(x_0)$ als Näherung für die Ableitung $f'(x_0)$ nicht mehr haltbar.)

Übrigens ist die *Maßeinheit für die Ableitung* $f'(x_0)$ dieselbe wie für die Differenzenquotienten, also das Verhältnis der Maßeinheiten für die abhängige und die unabhängige Variable. Bei einem Kapitalstand $K(t)$ € in Abhängigkeit von t Zeiteinheiten (z.B. Jahren) wird also die Ableitung $K'(t)$ in € pro Zeiteinheit angegeben und bei einer Angebots–Funktion, welche den auf dem Markt angebotenen Output eines Gutes in Abhängigkeit vom Marktpreis p (pro Mengeneinheit) beschreibt, wird die Änderungsrate $x'(p)$ in Mengeneinheiten pro Geldeinheit ausgewiesen. Ist die Geldeinheit 1 € und der Preis hoch, sagen wir über 1000 €, so ist zwar $x'(p) \approx x(p{+}1) - x(p)$, aber hinsichtlich der Einheiten stimmt diese Gleichung nur, wenn man die Änderung der Funktionswerte noch durch die Argumentänderung 1 € dividiert; $x'(p)$ beschreibt eben nicht die Angebotsänderung bei Preissteigerung um 1 €, sondern die Angebotsänderung *relativ zur Preisänderung*.

2) *Ökomomische Terminologie*: Weil $f'(x_0)$ näherungsweise gleich $\frac{1}{h}[f(x_0{+}h) - f(x_0)]$ ist bei "marginalen" (d.h. vernachlässigbar kleinen) Zuwächsen h, wobei "in der Grenze" $h \to 0$ sogar Gleichheit gilt, heißt in den Wirtschaftswissenschaften die Ableitung f' einer Funktion f die **Marginalfunktion** oder **Grenzfunktion** zu f. Besonders glücklich scheint das nicht. Auch besteht die Gefahr der Verwechselung des Grenzfunktionswertes $f'(x_0)$ mit dem Grenzwert der Funktion $f(x)$ bei $x \to x_0$, was ein ganz anderes Konzept ist (und den Funktionswert $f(x_0)$ liefert, wenn f an der Stelle x_0 stetig ist; besser wäre schon "Grenz–Funktionsänderung" statt "Grenzfunktion", um auszudrücken, was wirklich gemeint ist.) Aber in der wirtschaftswissenschaftlichen Literatur ist diese Terminologie verbreitet. So findet man die Namen *Grenzkosten* oder *Grenzkostenfunktion* für die Ableitung $K'(x)$ einer Kostenfunktion $K(x)$ (in Abhängigkeit vom Produktions–Output x), *Grenzstückkosten* für die Änderungsrate $k'(x)$ der zugehörigen Stückkostenfunktion $k(x) = K(x)/x$, *marginale Sparquote* für die Ableitung $\frac{d}{dY}S(Y)$ einer Sparfunktion $S(Y)$ und *marginale Konsumquote* oder *Grenzhang zum Konsum* für die Ableitung $\frac{d}{dY}C(Y)$ der entsprechenden Konsumfunktion (jeweils in Abhängigkeit vom Einkommen). Auch der *Grenzsteuersatz* ist eine Ableitung, nämlich die der Einkommensteuerfunktion $S(E)$, welche die auf ein zu versteuerndes Einkommen von E € zu entrichtende Einkommensteuer angibt; hier ist die Interpretation der Ableitung als Zuwachs $S'(E) \approx S(E + 1) - S(E)$ sinnvoll, d.h. der Grenzsteuersatz wird als der Teil eines zusätzlich erzielten Einkommens von 1 € interpretiert, der als Steuer abgeführt werden muss (und normalerweise als Prozentsatz $100 \cdot S'(E)\,\%$ angegeben). Die Differentialrechnung, das ist die Theorie von Berechnung und Anwendung der Ableitung, heißt in der Wirtschaftswissenschaft **Marginalanalyse** (Analyse des Funktionsverhaltens bei marginalen Argumentänderungen).

3) *Geometrische Interpretation*: Dies ist aus der Schule am besten bekannt, aber weder für die ökonomischen Anwendungen noch für das Verständnis der Differentialrechnung besonders wichtig. Der Differenzenquotient $[f(x) - f(x_0)]/(x - x_0)$ ist die *Steigung der Sekante* des Graphen von f, die durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$ geht. Wenn f an der Stelle x_0 differenzierbar ist (und nur dann), so nähern sich diese Sekanten bei $x \to x_0$ einer Grenzgeraden durch den Punkt $(x_0, f(x_0))$, die man die *Tangente an den Graphen* in diesem Punkte nennt.



Die Steigung der Tangente ist dann der Grenzwert der Sekantensteigungen bei $x \to x_0$, also der Ableitungswert $f'(x_0)$.

- *$f'(x_0)$ ist die Steigung der Tangente an Graph(f) im Punkt $(x_0, f(x_0))$,*

und deshalb wird die Ableitung auch als "Steigung" der Funktion bezeichnet. Da die Tangente eine Gerade durch den Punkt $(x_0, f(x_0))$ ist und die Steigung $f'(x_0)$ hat, ergibt sich aus der Punkt-Steigungs-Form der Geradengleichung sofort die *Gleichung der Tangente*:

$$y = f'(x_0)(x - x_0) + f(x_0)\,.$$

4) Die Differenzierbarkeit von f an der Stelle x_0 läßt sich ausdrücken durch folgende **Zuwachsformel**, die äquivalent ist zur Existenz der Ableitung $f'(x_0) = c$:

$$\begin{aligned} f(x_0 + h) - f(x_0) &= c \cdot h + h \cdot R(h) \\ & \qquad \text{mit } c = f'(x_0) \text{ und } \lim_{h \to 0} R(h) = 0\,. \\ f(x) - f(x_0) &= c \cdot (x - x_0) + (x - x_0) \cdot R(x - x_0) \end{aligned}$$

Dividiert man nämlich die erste Gleichung durch $h \neq 0$, so besagt sie, dass die Abweichung der Differenzenquotienten vom Wert c gleich $R(h)$ ist, und $\lim_{h \to 0} R(h) = 0$ bedeutet eben, dass diese Abweichung kleiner ist als jede gegebene Fehlerschranke, wenn nur $|h|$ klein genug ist. Die Zuwachsformel kann man so lesen:

- *Die Änderung der Funktionswerte bei x_0 ist in erster Näherung proportional zur Änderung des Arguments; der Proportionalitätsfaktor ist der Ableitungswert $f'(x_0)$.*

Wenn also $f'(x_0) = 2, 1$ bzw. $\frac{1}{2}$ ist, so bewirkt eine kleine Änderung von x_0 ungefähr eine doppelt so große, gleich große bzw. halb so große Änderung des Funktionswertes $f(x_0)$ in derselben Richtung wie die Änderung des Arguments, im Fall $f'(x_0) = -2, -1$ bzw. $-\frac{1}{2}$ eine entsprechende Änderung des Funktionswertes in der entgegengesetzten Richtung, im Fall $f'(x_0) = 0$ aber praktisch keine wahrnehmbare Änderung des Funktionswertes. "In erster Näherung" bedeutet genau, dass der Fehler $hR(h)$, den man bei Ersetzen des Funktionszuwachses $f(x+h) - f(x_0)$ durch die zu h proportionale Größe $f'(x_0)h$ macht, beliebig klein ist im Verhältnis zur Größe von h, also kleiner als $|h|\varepsilon$ für ein gegebenes $\varepsilon > 0$, wenn nur $|h|$ klein genug ist.

Eine Anwendung finden diese Überlegungen in der *Fehlerrechnung*: Hat man eine (z.B. ökonomische) Größe durch eine Zahl $x_0 \in \mathbb{R}$ beschrieben, die nicht genau bekannt oder fehlerbehaftet ist, so will man abschätzen, wie groß der Fehler bei einer gemäß bekanntem Funktionsgesetz von x_0 abhängigen Größe $y_0 = f(x_0)$ sein kann. Weicht der wahre Wert $x = x_0 + \Delta x_0$ der Größe von der Zahl x_0 nur um einen kleinen Fehler Δx_0 ab, so kann man den Fehler $\Delta y_0 = y - y_0 = f(x) - f(x_0)$ der abhängigen Variablen mit ausreichender Genauigkeit durch $f'(x_0)\Delta x_0$ darstellen.

- *Der Fehler bei der abhängigen Größe* $y = f(x)$ *ist* $\Delta y_0 \approx f'(x_0)\Delta x_0$.

Wenn also $f'(x_0)$ beispielsweise den Wert 1000 hat, so führt ein kleiner Fehler bei der Bestimmung von x_0 unter Umständen zu einem etwa 1000 Mal größeren Fehler bei der Angabe von y_0. Wenn andererseits $|f'(x_0)| \leq 1$ ist, so bewirken (kleine!) Fehler bei der Angabe von x_0 keine größeren Fehler bei der Angabe von y_0.

5) Aus der Zuwachsformel liest man unmittelbar ab, dass sich der Funktionswert beliebig wenig ändert, wenn man das Argument hinreichend wenig variiert, d.h.

- *differenzierbare Funktionen* f *sind überall stetig auf ihrem Definitionsintervall;*

denn ist $\varepsilon > 0$ gegeben, so gilt für x_0 und x im Definitionsintervall I mit $|x-x_0|$ hinreichend klein ja $|f'(x_0)(x - x_0)| < \frac{1}{2}\varepsilon$ und $|(x - x_0)R(x - x_0)| < \frac{1}{2}\varepsilon$, da $|R(x - x_0)| < 1$ ist (sogar beliebig klein) für alle x nahe genug bei x_0, also liefert die Zuwachsformel $|f(x) - f(x_0)| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$ für alle $x \in I$ nahe genug bei x_0, und das ist die Stetigkeit von f an der Stelle x_0. Die Argumentation zeigt aber mehr, nämlich dass $|f(x) - f(x_0)| < L \cdot |x - x_0|$ ist für x nahe bei x_0 mit einer Dehnungsschranke $L < \infty$ (man kann hier irgendeine Zahl $L > |f'(x_0)|$ nehmen), d.h. die Differenz der Funktionswerte ist nicht nur klein für x nahe x_0, sondern höchstens von derselben Größenordnung wie $|x - x_0|$. So etwas gilt für allgemeine stetige Funktionen keineswegs (z.B. nicht für $f(x) := \sqrt{|x|}$ bei $x_0 = 0$), und deshalb *ist Differenzierbarkeit eine viel stärkere Eigenschaft als Stetigkeit.*

6) Man kann die Zuwachsformeln auch so interpretieren, dass die Funktion f nahe bei der Stelle x_0 besonders gut approximiert wird durch die lineare Funktion

$$\ell(x) := f'(x_0)(x - x_0) + f(x_0)\,,$$

genannt die **Linearisierung** von f zur Stelle x_0. Die Linearisierung von f ist diejenige lineare Funktion, die $f'(x_0)$ als Steigung hat und an der Stelle x_0 denselben Funktionswert wie f. Sie approximiert die Funktion f gut nahe bei x_0 in dem Sinne, dass gilt

$$f(x) - \ell(x) \;=\; (x - x_0) \cdot R(x - x_0)$$

wobei die rechte Seite beliebig kleinen Betrag hat im Verhältnis zu $|x-x_0|$, also wesentlich kleiner ist als der Abstand von x zu x_0, wenn x nahe genug bei x_0 ist.

Man kann sich überlegen, dass lineare Approximierbarkeit von f bei x_0 in diesem Sinne äquivalent zur Differenzierbarkeit von f in x_0 ist und dass es keine andere lineare Funktion gibt, die in demselben Sinne f bei x_0 approximiert. Der Graph der Linearisierung ℓ von f bei x_0 ist die Graphentangente im Punkt $(x_0, f(x_0))$, und die gute Approximation von $f(x)$ durch $\ell(x)$ für x nahe bei x_0 entspricht der geometrischen Tatsache, dass die Tangente den Graphen berührt und ihn nicht schneidet, wie eine Sekante. In der Mathematik *ist Linearisierung die Grundidee der Differentialrechnung*: Man approximiert eine komplizierte Funktion $f(x)$ nahe einer Stelle x_0 gut durch eine lineare Funktion und versucht, aus den (einfach zu bestimmenden) Eigenschaften der linearen Approximation Rückschlüsse auf Eigenschaften von $f(x)$ zu ziehen. Diese Idee lässt sich auch auf Funktionen von mehreren Variablen ausdehnen, was auch ökonomisch relevante Anwendungen hat (s. Kap. 5).

Der homogene lineare Anteil $\ell_0(h) = f'(x_0) \cdot h$ der Linearisierung, der in erster Näherung die Werteänderung $f(x_0{+}h) - f(x_0)$ angibt für kleine $|h|$, heißt übrigens **Differential** der Funktion f an der Stelle x_0 und wird bezeichnet $df(x_0)h$, $d_{x_0}f(h)$ oder ähnlich. Das Differential ist also die homogene lineare Funktion $\ell_0(h)$, welche die Funktionsänderungen $f(x_0{+}h) - f(x_0)$ bei $h \to 0$ am besten approximiert. Noch immer findet man, selbst in neueren Büchern, eine Erklärung der Differentiale als "unendlich kleine Funktionszuwächse df", die einem "infinitesimal kleinen Argumentzuwachs dx" entsprechen, oder ähnlich Unverständliches. Es gibt aber keine "unendlich kleinen Größen" außer Null — es gibt nur endliche Größen, die man im Grenzübergang gegen Null gehen lassen kann. Wenn in ökonomischen Kontext von Differentialen dx, dy etc. gesprochen wird, so ist eigentlich gemeint, dass man kleine Variablendifferenzen Δx, Δy betrachtet und mit diesen einen Grenzübergang gegen Null durchführt. Auch bei der Ableitung $\frac{df}{dx}$ handelt es sich — obwohl die Ableitung auch "Differentialquotient" genannt wird — nicht um einen Quotienten zweier Differentiale, d.h. einen "Quotienten von unendlich kleiner Größen", den niemand erklären kann, sondern um eine symbolische Schreibweise, die an diesen Grenzübergang mit den Differenzenquotienten $\frac{\Delta f}{\Delta x}$ erinnert. (Wenn man allerdings das oben erklärte Differential $df(x_0) = \ell_0$ durch das Differential $dx(x_0)$ der Funktion x teilt, d.h. durch die homogene lineare Funktion $dx(x_0)h = h$, so ist der Quotient $df(x_0)/dx(x_0)$ der Differentiale gleich der Konstanten $f'(x_0)$; in diesem höheren Sinne ist es dann doch richtig, dass die Ableitung $f' = \frac{df}{dx}$ der Quotient der Differentiale df und dx ist.)

7) Zur *Notation*: Das Symbol "$\frac{df}{dx}$" soll daran erinnern, dass ein Grenzwert von Differenzenquotienten $\frac{\Delta f}{\Delta x}$ beim Grenzübergang $\Delta x \to 0$ gebildet wird. Es handelt sich hier also nicht um einen Quotienten von Zahlen, sondern nur um ein zusammengesetztes Symbol für die Ableitung von f. (Ganz abwegig wäre es — was schon vorgekommen ist —, etwa das "d" in "$\frac{d}{dx}$" zu kürzen und "$\frac{df}{dx} = \frac{f}{x}$" zu schreiben.) Die "Strich–Notation" f' für die Ableitung ist sinnvoll, wenn ein Buchstabensymbol f als Funktionsname vergeben ist. Die "$\frac{d}{dx}$"–Notation ist dagegen praktisch, wenn die Funktion durch einen Term in der Variablen x gegeben ist, ohne dass sie einen Funktionsnamen erhalten hat; man schreibt den Term dann einfach direkt hinter das $\frac{d}{dx}$–Symbol, um die Ableitung der Funktion zu bezeichnen, die durch den Term definiert wird. So bezeichnet $\frac{d}{dx}(x^2 + 3\mathrm{e}^{-x})$ die Ableitung der Funktion $f(x) = x^2 + 3\mathrm{e}^{-x}$ an einer allgemeinen Stelle x, also die Ableitungsfunktion f'. Will man die Ableitung an einer bestimmten Stelle x_0 bilden, also x_0 in f' einsetzen, so kann man dies notieren, indem man "$x = x_0$" als Subskript hinzufügt, das durch einen vertikalen Strich abgetrennt ist. Für den Ableitungswert $f'(x_0)$ der gerade betrachteten Funktion f schreibt man also $\frac{d}{dx}\big|_{x=x_0}(x^2 + 3\mathrm{e}^{-x})$ oder $\frac{d}{dx}(x^2 + 3\mathrm{e}^{-x})\big|_{x=x_0}$ (und nicht etwa $(x^2 + 3\mathrm{e}^{-x})'$ — die "Strich-Schreibweise" für die Ableitung von Termen ist unüblich.)

Natürlich müssen die unabhängigen Variablen nicht immer “x” und die abhängigen nicht immer “$y = f(x)$” heißen. Also schreibt man $B'(t)$ oder $\frac{d}{dt}B$ für die Ableitung eines zeitabhängigen kontinuierlich veränderlichen Bestandes $B(t)$, und $x'(p_0)$ oder $\frac{d}{dp}\Big|_{p=p_0} x(p)$ oder $\frac{dx}{dp}(p_0)$ für die Ableitung einer Angebotsfunktion $x(p)$ bei einem bestimmten Preis $p = p_0$. Oft hängen Funktionen von mehreren Variablen oder Parametern ab, etwa die auf Taschenrechnern vorhandene Funktion x^y von $x \in \mathbb{R}_{>0}$ und von $y \in \mathbb{R}$. Ob man von “Variablen” oder von “Parametern” spricht, ist nur eine Frage des Standpunkts: Parameter betrachtet man innerhalb des Kontextes einer Diskussion als konstant. Wird z.B. y als Parameter aufgefasst, so hat man Potenzfunktionen $f(x) = x^y$ zu betrachten (für jeden Exponenten y eine andere). Wird dagegen x als Parameter aufgefasst, so hat man Exponentialfunktionen $g(y) = x^y$ zu untersuchen (für jede Basis $x > 0$ eine andere). Man kann x^y auch als Funktion von zwei Variablen (x, y) auffassen, also bei simultaner Variation von x und y studieren. Differenziert wird aber (jedenfalls in diesem Kapitel) immer nur nach *einer* reellen Variablen, alle anderen müssen während des Ableitungsprozesses konstant gehalten werden. In der Wirtschaftswissenschaft heißt dies, wie schon gesagt, **c.p.–Bedingung** (von “ceteris paribus”, d.h. die übrigen [Variablen bleiben auf] gleichen [Werten]). Bei einer Funktion von mehreren Variablen nennt man die Ableitung nach einer einzelnen Variablen bei c.p.-Bedingung für die übrigen Variablen eine **partielle Ableitung** der Funktion. Dafür ist als Notation statt des romanischen “d”ein kyrillisches “∂” üblich, also “$\frac{\partial}{\partial x}$ statt “$\frac{d}{dx}$”. (Aber “$\frac{d}{dx}$” ist nicht verkehrt; am “∂” erkennt man nur, dass die differenzierte Funktion als Funktion von mehreren gleichberechtigten Variablen aufgefasst wird.)

Für $h(x, y) = x^y$ bedeutet also

$$\frac{\partial h}{\partial x}(x_0, y_0) = \frac{\partial}{\partial x}\Big|_{x=x_0,\, y=y_0} h(x, y) := \frac{d}{dx}\Big|_{x=x_0} h(x, y_0)$$

die Ableitung $f'(x_0)$ der Funktion $f(x) = x^{y_0}$ nach x an der Stelle x_0 bei festem y_0 und

$$\frac{\partial h}{\partial y}(x_0, y_0) = \frac{\partial}{\partial y}\Big|_{x=x_0,\, y=y_0} h(x, y) := \frac{d}{dy}\Big|_{y=y_0} h(x_0, y)$$

die Ableitung $g'(y_0)$ der Funktion $g(y) = {x_0}^y$ nach y an der Stelle y_0 bei festem $x_0 > 0$. ■

BEISPIELE (***Ableitung der Grundfunktionen***)**:**

1) Konstante Funktionen $f \equiv c$ haben natürlich alle Differenzenquotienten Null, $\frac{1}{h}[f(x + h) - f(x)] = \frac{1}{h}[c - c] = 0$, und daher ist auch die Ableitung, der Grenzwert der Differenzenquotienten bei $h \to 0$, an jeder Stelle x gleich Null.

- *Konstante Funktionen haben Ableitungswert Null an jeder Stelle.*

Das ist natürlich keine Überraschung: Die Ableitung beschreibt ja die Änderungsrate einer Funktion, und konstante Funktionen ändern ihren Wert überhaupt nicht, also muss die Änderungsrate an jeder Stelle gleich Null sein. Später werden wir sehen, dass auch umgekehrt jede Funktion mit Ableitung Null überall auf einem Intervall konstant ist.

2) Lineare Funktionen $\ell(x) = ax + b$ haben Differenzenquotienten $\frac{1}{h}[a(x+h)+b - (ax+b)] = a$ unabhängig vom Argumentzuwachs h und von der Stelle x. Daher ist natürlich auch der Grenzwert der Differenzenquotienten bei $h \to 0$ gleich a an jeder Stelle x.

- *Die Ableitung einer linearen Funktion ist konstant mit dem führenden Koeffizienten der linearen Funktion als Wert.*

Der führende Koeffizient a von $ax+b$ wird deshalb ja auch die *Steigung* der linearen Funktion genannt. Geometrisch ist das Ergebnis klar, weil der Graph einer linearen Funktion eine Gerade ist, also an jeder Stelle mit seiner Tangente übereinstimmt; deshalb muss auch die Tangentensteigung, also die Ableitung der Funktion, an jeder Stelle gleich der Steigung dieser Geraden sein. Später werden wir sehen, dass auch umgekehrt jede differenzierbare Funktion mit konstanter Ableitung linear ist. (Das ist auch plausibel, aber doch beweisbedürftig: Wenn alle Tangenten an den Graphen einer Funktion dieselbe Steigung haben, warum muss dann der Graph selbst schon eine Gerade sein?) *Für lineare Funktionen ℓ gilt die Zuwachsformel exakt,*

$$\ell(x+h) = \ell(x) + \ell'(x)h \ ,$$

d.h. der Fehler $hR(h)$ verschwindet hier; und nur bei linearen Funktionen ist das so (weil diese Gleichung für $x := 0$ ja $\ell(h) = ah + b$ liefert mit $a := \ell'(0)$ und $b := \ell(0)$). Als Spezialfall halten wir für die Ableitung der Grundfunktion $\ell(x) = x$ noch fest:

$$\frac{d}{dx}x = 1 \qquad \text{an allen Stellen} \quad x \in \mathbb{R} \ .$$

3) Quadratische Funktionen $q(x) = ax^2 + bx + c$ haben Zuwächse $q(x+h) - q(x) = 2axh + ah^2 + bh = (2ax + b)h + ah^2$. Der Fehlerterm ist hier also $hR(h) = ah^2$, d.h. $R(h) = ah$. Da $R(h) = ah$ Grenzwert Null hat bei $h \to 0$, ist q differenzierbar an der Stelle x mit Ableitung $2ax+b$. (Die Zuwachsformel $f(x+h)-f(x) = Ch+hR(h)$ mit $R(h) \to 0$ bei $h \to 0$ ist ja äquivalent zur Differenzierbarkeit von f an der Stelle x mit $f'(x) = C$; siehe Punkt 4) in der vorangegangenen Diskussion.)

- *Die Ableitung der quadratischen Funktion ax^2+bx+c ist die lineare Funktion $2ax+b$.*

Dieses Ergebnis ist auch geometrisch plausibel, zumindest in qualitativer Hinsicht. Im Fall $a > 0$ ist z.B. der Graph der quadratischen Funktion q eine nach oben geöffnete Parabel mit Scheitelrechtswert $-\frac{b}{2a}$. Die Steigung dieser Parabel ist daher Null an der Stelle $-\frac{b}{2a}$ und rechts davon positiv und wachsend, links davon negativ und fallend. Genau dieses Verhalten hat die lineare Funktion $ax + b$. Allerdings zeigt diese qualitative geometrische Überlegung nicht, dass die Ableitungsfunktion linear ist; wenn man das geometrisch einsehen will, so muss man einiges mehr über die Tangentenkonstruktion zu Parabeln wissen. Die Differentialrechnung macht aber gerade solche anspruchsvolleren geometrischen Überlegungen entbehrlich; denn die einfache Ableitungsberechnung $q'(x_0) = 2ax_0+b$ oben liefert uns ja sofort die *Gleichung der Parabeltangente* im Punkt $(x_0, q(x_0))$:

$$y = (2ax_0 + b)(x - x_0) + ax_0^2 + bx_0 + c = (2ax_0 + b)x + c - ax_0^2 \ .$$

4) Für **Potenzfunktionen** x^n mit natürlichem Exponenten können wir $(x+h)^n$ mit der binomischen Formel schreiben $x^n + \binom{n}{1}x^{n-1}h + h^2[\ldots]$, wobei $[\ldots]$ eine Linearkombination von Produkten von Potenzen von x und h ist, die nicht weiter interessiert. Der Zuwachs ist also $(x + h)^n - x^n = \binom{n}{1}x^{n-1}h + hR(h)$, wobei $R(h) = h[\ldots]$ gegen Null geht mit $h \to 0$. Dies bedeutet, dass die Funktion x^n an jeder Stelle differenzierbar ist mit Ableitung $\binom{n}{1}x^{n-1}$, und da der Binomialkoeffizient $\binom{n}{1} = n$ ist, erhalten wir die fundamentale **Ableitungsformel für Potenzfunktionen**:

$$\boxed{\frac{d}{dx}x^n \ = \ nx^{n-1}} \qquad \text{für} \ \ x \in \mathbb{R} \ \ \text{und} \ \ n \in \mathbb{N} \ ,$$

also

$$\frac{d}{dx}x = 1 \ , \quad \frac{d}{dx}x^2 = 2x \ , \quad \frac{d}{dx}x^3 = 3x^2 \ , \quad \frac{d}{dx}x^4 = 4x^3 \ , \ \ldots \ .$$

Das analoge Ergebnis gilt auf $\mathbb{R}_{\neq 0}$, wenn der Exponent eine negative ganze Zahl ist. Dazu schreibt man für $n \in \mathbb{N}$ die Differenzenquotienten der Funktion x^{-n} in der Form $\frac{1}{h}[(x+h)^{-n} - x^{-n}] = \frac{1}{h}[x^n - (x+h)^n]/[x^n(x+h)^n]$, wobei der Zähler Limes $-nx^{n-1}$ hat bei $h \to 0$, wie wir gerade herausgefunden haben, während der Nenner gegen x^{2n} strebt; der Limes der Differenzenquotienten ist daher $-nx^{-n-1}$, d.h.

$$\frac{d}{dx}\,x^{-n} = -nx^{-n-1} \qquad \text{für} \;\; x \in \mathbb{R}_{\neq 0} \;\; \text{und} \;\; n \in \mathbb{N}\,,$$

also

$$\frac{d}{dx}\,\frac{1}{x} = \frac{-1}{x^2}\,, \quad \frac{d}{dx}\,\frac{1}{x^2} = \frac{-2}{x^3}\,, \quad \frac{d}{dx}\,\frac{1}{x^n} = \frac{-n}{x^{n+1}} \quad \text{für} \;\; x \neq 0\,.$$

Auch die entsprechende **Ableitungsformel für allgemeine Potenzfunktionen** x^s mit $s \in \mathbb{R}$ gilt auf $\mathbb{R}_{>0}$:

$$\boxed{\frac{d}{dx}\,x^s \;=\; sx^{s-1}} \qquad \text{für} \;\; x \in \mathbb{R}_{>0} \;\; \text{und} \;\; s \in \mathbb{R}\,.$$

- *Man differenziert eine Potenzfunktion, indem man sie mit ihrem Exponenten multipliziert und danach den Exponenten um 1 erniedrigt.*

Der Beweis ist am einfachsten mit Hilfe der Darstellung $x^s = \exp(s \ln x)$ und den späteren Rechenregeln für das Differenzieren, kann aber z.B. auch direkt mit der Bernoulli-Ungleichung $(1+t)^s > 1 + st$ für $-1 < t \neq 1$ aus 1.4 so geführt werden: $(x+h)^s - x^s = x^s[(1+\frac{h}{x})^s - 1] > x^s s \frac{h}{x}$ gilt für $x > 0$ und hinreichend kleine $|h| > 0$, und durch Vertauschen der Rollen von x und $x+h$ erhält man auch die Ungleichung $(x+h)^s - x^s < (x+h)^s \frac{h}{x+h}$ in umgekehrter Richtung. Die Differenzenquotienten $\frac{1}{h}[(x+h)^s - x^s]$ liegen folglich zwischen sx^{s-1} und $s(x+h)^{s-1}$, und da der letzte Ausdruck gegen sx^{s-1} strebt bei $h \to 0$ wegen der Stetigkeit der Potenzfunktionen, haben die Differenzenquotienten den Limes sx^{s-1} bei $h \to 0$. Für $0 < s < 1$ gilt die Bernoulli-Ungleichung zwar in der entgegengesetzten Richtung, aber das ändert nichts an der Argumentation. Die Fälle $s = 0$ (also x^s konstant 1 auf $\mathbb{R}_{>0}$) und $s = 1$ (also $x^s = x$ linear) haben wir schon behandelt; dafür gilt die Ableitungsformel offensichtlich auch.

Als *Spezialfälle* notieren wir noch die Ableitung von "reziproken Potenzfunktionen" $1/x^t$:

$$\frac{d}{dx}\,\frac{1}{x^t} = \frac{d}{dx}\,x^{-t} = -tx^{-t-1} = \frac{-t}{x^{t+1}} \qquad \text{für} \;\; x > 0 \;\; \text{und} \;\; t \in \mathbb{R}\,,$$

und die **Ableitungen der Wurzelfunktionen** $\sqrt[n]{x} = x^{1/n}$ für $x > 0$ und $n \in \mathbb{N}_{\geq 2}$:

$$\frac{d}{dx}\,\sqrt{x} = \tfrac{1}{2}x^{1/2-1} = \frac{1}{2\sqrt{x}}\,, \qquad \frac{d}{dx}\,\sqrt[n]{x}\,, = \tfrac{1}{n}x^{1/n-1} = \frac{\sqrt[n]{x}}{nx}\,.$$

5) Die Ableitung der **Exponentialfunktionen** ist am einfachsten zu sehen im Falle der natürlichen Basis e. Dafür hatten wir in 1.4 die fundamentale Ungleichung $\mathrm{e}^h > 1 + h$ für reelle $h \neq 0$ hergeleitet, und das gibt $\mathrm{e}^{x+h} - \mathrm{e}^x = \mathrm{e}^x(\mathrm{e}^h - 1) > \mathrm{e}^x h$. Andererseits ist auch $\mathrm{e}^{x+h} - \mathrm{e}^x = \mathrm{e}^{x+h}(1 - \mathrm{e}^{-h}) < \mathrm{e}^{x+h}$ und $\mathrm{e}^h = 1/\mathrm{e}^{-h} < 1/(1-h)$ für $-1 < h \neq 0$. Die Differenzenquotienten $\frac{1}{h}[\mathrm{e}^{x+h} - \mathrm{e}^x]$ liegen also zwischen e^x und $\mathrm{e}^x/(1-h)$ und streben daher bei $h \to 0$ gegen e^x. Das Resultat ist, dass $\exp(x) = \mathrm{e}^x$ an jeder Stelle $x \in \mathbb{R}$ differenzierbar ist, und wir erhalten für die **Ableitung der natürlichen Exponentialfunktion**:

$$\boxed{\frac{d}{dx}\mathrm{e}^x \;=\; \mathrm{e}^x} \qquad \text{für alle} \;\; x \in \mathbb{R}\,, \;\; \text{kurz} \qquad \boxed{\exp = \exp'} \qquad \text{auf} \;\; \mathbb{R}\,;$$

mit Worten:

- *Die natürliche Exponentialfunktion ist ihre eigene Ableitungsfunktion.*

Dies ist eine weitere Eigenschaft, welche die natürliche Exponentialfunktion und damit die natürliche Basis e auszeichnet: Bis auf einen konstanten Faktor ist exp die einzige Funktion, die ihre eigene Ableitungsfunktion ist. Man kann das Ergebnis auch aus der Exponentialreihe $\mathrm{e}^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \ldots$ herleiten, indem man sie Glied für Glied differenziert: Das Ergebnis ist $0 + \frac{1}{1!}1 + \frac{1}{2!}2x + \frac{1}{3!}3x^2 + \frac{1}{4!}4x^3 + \ldots = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \ldots = \mathrm{e}^x$. Das Vorgehen kann mit einigem mathematischen Aufwand gerechtfertigt werden und zeigt, dass man bei der Suche nach einer Funktion in Form einer Potenzreihe $f(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \ldots$, die $f = f'$ erfüllt, automatisch auf die Exponentialreihe geführt wird.

Für eine allgemeine Exponentialfunktion a^x mit Basis $0 < a \neq 1$ können wir die Differenzenquotienten schreiben $\frac{1}{h}[a^{h+x} - a^x] = a^x \frac{1}{h}[a^h - 1] = a^x \frac{1}{k}[\mathrm{e}^k - 1] \ln a$ mit $k := h \ln a$, und da mit h auch k gegen Null geht sowie $\frac{1}{k}[\mathrm{e}^k - 1]$ gegen die Ableitung der natürlichen Exponentialfunktion an der Stelle 0, also gegen $\mathrm{e}^0 = 1$, erhalten wir als Limes den Wert $a^x \ln a$. Damit ist die **Ableitung einer allgemeinen Exponentialfunktion** berechnet:

$$\boxed{\frac{d}{dx}\, a^x \;=\; a^x \cdot \ln a} \qquad \text{für } \; x \in \mathbb{R} \;\text{ und } a > 0\,,$$

was auch für $a = 1$, also $a^x = 1$ konstant und $\ln a = 0$, richtig ist.

- *Man differenziert eine Exponentialfunktion, indem man sie mit dem natürlichen Logarithmus der Basis multipliziert.*

Da $\ln a = 1$ nur für $a = \mathrm{e}$ gilt, sehen wir, dass die natürliche Exponentialfunktion die einzige ist, die Steigung 1 an der Stelle $x = 0$ hat; dies ist eine weitere Eigenschaft, welche die natürliche Basis besonders auszeichnet.

Für Exponentialfunktionen der Form $e^{ct} = (\mathrm{e}^c)^t$ lautet die Ableitungsformel

$$\frac{d}{dt}\, \mathrm{e}^{ct} = c \cdot \mathrm{e}^{ct}\,.$$

Die Ableitung einer solchen Funktion ist also proportional zu der abgeleiteten Funktion mit Proportionalitätsfaktor c; derselbe Zusammenhang zwischen f und f' besteht dann auch bei Vielfachen de^{ct} der Funktion e^{ct}. Ein solches Gesetz $f'(t) = cf(t)$ ist grundlegend für *ungehemmte kontinuierliche Wachstums- oder Zerfallsprozesse*, bei denen die Änderung eines "Bestandes" $f(t)$ zu jedem Zeitpunkt t proportional zur aktuellen Größe des Bestandes ist. Ein Beispiel ist die Kapitalentwicklung bei kontinuierlicher Verzinsung eines Anfangskapitals K_0, welche bei Jahreszinsfuß $p\%$ gegeben ist durch $K(t) = K_0\mathrm{e}^{pt/100}$ (siehe 1.3). Hier wächst der Bestand, weil der Proportionalitätsfaktor $c = \frac{1}{100}p$ positiv ist (und der Anfangswert K_0 auch). Zerfallsprozesse dagegen, wie z.B. radioaktiver Zerfall einer Substanz oder Ähnliches, werden beschrieben durch Funktionen $f(t) = d\mathrm{e}^{-ct}$ mit $c > 0$ und erfüllen $f' = -c \cdot f$ mit negativem Proportionaltätsfaktor $-c$.

6) Für die natürliche **Logarithmusfunktion** kann man die Differenzenquotienten zu Stellen $x_0 = \ln y_0$ und $x = \ln y$ schreiben $(\ln x - \ln x_0)/(x - x_0) = (y - y_0)/(\mathrm{e}^y - \mathrm{e}^{y_0})$. Das Reziproke des letzten Quotienten strebt gemäß 5) gegen $\mathrm{e}^{y_0} = x_0$ bei $y \to y_0$. Da die Logarithmusfunktion stetig ist, strebt mit $x \to x_0$ aber auch y gegen y_0, und damit die Differenzenquotienten der Logarithmusfunktion gegen $1/x_0$. Also ist $1/x_0$ die Ableitung von $\ln x$ an der Stelle x_0, und da x_0 hier eine beliebige Stelle in $\mathbb{R}_{>0}$ ist, haben wir für die **Ableitung der natürlichen Logarithmusfunktion** gezeigt:

$$\boxed{\frac{d}{dx}\, \ln x \;=\; \frac{1}{x}} \qquad \text{für } \; x \in \mathbb{R}_{>0}\,.$$

Das Argument, mit dem wir die Ableitung von ln aus Kenntnis der Ableitung von exp hier hergeleitet haben, funktioniert übrigens ganz allgemein bei der Berechnung der Ableitung der Umkehrfunktion einer streng monotonen Funktion mit Ableitung ohne Nullstelle; wir kommen darauf bei den Rechenregeln für das Differenzieren zurück. Die Ableitung anderer Logarithmusfunktionen $\log_a x = (\ln x)/(\ln a)$ ist nun einfach: Ihre Differenzenquotienten unterscheiden sich von denen der natürlichen Logarithmusfunktion alle nur um den Faktor $1/\ln a$, also erhalten wir für den Limes der Differenzenquotienten

$$\frac{d}{dx}\log_a x \;=\; \frac{1}{x\ln a} \qquad \text{für} \;\; x>0 \;\; \text{und} \;\; 0<a\neq 1\,.$$

Die Berechnung der Ableitungen der Logarithmusfunktionen füllt übrigens eine Lücke in der Liste der Ableitungsfunktionen, die wir bisher bestimmt haben. Für alle Exponenten $r \in \mathbb{R}$ außer $r=-1$ kennen wir nämlich eine Funktion, die x^r als Ableitung hat, wenigstens bis auf einen Faktor $\neq 0$, nämlich x^{r+1} mit Ableitung $(r{+}1)x^r$; die Funktion $\frac{1}{r+1}x^r$ hat dann genau x^r als Ableitung (wie man leicht überlegt oder mit der "ersten Faktorregel" weiter unten schließt). Für $r=-1$ hat aber $x^{r+1}=x^0\equiv 1$ Ableitung Null, und wir erhalten somit keine Potenzfunktion mit Ableitung $1/x$. Diese Lücke schließt gerade die natürliche Logarithmusfunktion.

7) Die Ableitung der *Kreisfunktionen* geben wir hier nur ohne Beweis an, weil sie in der Ökonomie nicht so wichtig sind:

$$\frac{d}{dx}\sin x \;=\; \cos x\,, \qquad \frac{d}{dx}\cos x \;=\; -\sin x \qquad \text{für} \;\; x\in\mathbb{R}\,.$$

8) Hier einige Beispiele für die Berechnung von **partiellen Ableitungen**: Für Potenzfunktionen x^s einer Variablen x haben wir als Ableitung sx^{s-1} berechnet. Die Differenzenquotienten von Vielfachen cx^s unterscheiden sich dann alle nur durch den Faktor c von den Differenzenquotienten zu x^s, daher ist auch die Ableitung von cx^s, also der Limes der Differenzenquotienten, gleich csx^{s-1}. Diese Argumentation beweist für differenzierbare Funktionen f ganz allgemein die **erste Faktorregel**:

$$(cf)' = c\cdot f' \qquad \text{für Konstanten} \;\; c\in\mathbb{R}\,.$$

Es folgt z.B.

$$\frac{\partial}{\partial x}\,x^r y^s = r x^{r-1} y^s\,, \qquad \frac{\partial}{\partial y}\,x^r y^s = s x^r y^{s-1}\,,$$

da bei der Ableitung nach x die Potenz y^s als konstanter Faktor anzusehen ist und bei der Ableitung nach y die Potenz x^r. Die Rechnung gilt für $x,y\in\mathbb{R}_{>0}$, bei natürlichen (bzw. ganzen) Zahlen als Exponenten sogar für beliebige $x,y\in\mathbb{R}$ (bzw. in $\mathbb{R}_{\neq 0}$). Funktionen der Form $x^r y^s$ kommen in der Wirtschaftswissenschaft als Produktionsfunktionen vor. Allgemeiner verwendet man als Ansatz für Produktionsfunktionen sog. **Cobb-Douglas-Funktionen**

$$x(r_1,r_2,\ldots,r_k) = c\cdot r_1^{s_1}\cdot r_2^{s_2}\cdot\ldots\cdot r_k^{s_k}$$

mit reellen Exponenten $s_i>0$, um in bestimmten Situationen einen Produktions-Output x (Produktionseinheiten) in Abhängigkeit vom Einsatz von k Produktionsfaktoren $r_1,r_2,\ldots,r_k>0$ (Rohstoffe, Energie, Arbeitsleistung, … jeweils in Faktoreinheiten) zu beschreiben. Die partielle Ableitung

$$\frac{\partial}{\partial r_i}x(r_1,r_2,\ldots,r_k) = c\cdot r_1^{s_1}\cdot\ldots\cdot s_i r_i^{s_i-1}\cdot\ldots\cdot r_k^{s_k} = \frac{s_i}{r_i}\cdot x(r_1,r_2,\ldots,r_k)$$

gibt dann die Änderung des Outputs bei Erhöhung des i-ten Faktoreinsatzes um eine (kleine) Einheit bei c.p.-Bedingung an, wenn also der Einsatz der anderen Faktoren unverändert bleibt.

Für die partiellen Ableitungen der auf dem Taschenrechner vorhandenen Funktion x^y von zwei Variablen $x \in \mathbb{R}_{>0}$ und $y \in \mathbb{R}$ ergibt sich

$$\frac{\partial}{\partial x}\, x^y = yx^{y-1} = \frac{x}{y}\, x^y\,, \qquad \frac{\partial}{\partial y}\, x^y = x^y \ln x\,,$$

weil bei der Ableitung nach x (bei konstantem y) eine Potenzfunktion, bei der Ableitung nach y (bei festem x) aber eine Exponentialfunktion zu differenzieren ist.

9) Einige **Ableitungen höherer Ordnung** können wir nun auch schon leicht berechnen:

$$\frac{d^2}{dx^2}\, x^s = \frac{d}{dx}\, sx^{s-1} = s(s-1)x^{s-2}$$

und allgemein

$$\frac{d^m}{dx^m}\, x^s = s(s-1)\cdot\ldots\cdot(s-m+1)\cdot x^{s-m}\,.$$

Dies gilt für $x > 0$, im Falle eines ganzen Exponenten aber auch für $x \neq 0$ und im Falle natürlicher Exponenten sogar für alle $x \in \mathbb{R}$. Insbesondere haben wir also

$$\frac{d^m}{dx^m}\, x^n \;=\; \begin{cases} \dfrac{n!}{(n-m)!}\, x^{n-m} & \text{für} \quad m < n \\ n!\ \text{ konstant} & \text{für} \quad m = n \\ 0 \quad \text{konstant} & \text{für} \quad m > n \end{cases} \qquad (\, n \in \mathbb{N}_0\,)\,.$$

Wegen $\frac{d}{dx} \ln x = 1/x$ folgt auch

$$\frac{d^m}{dx^m}\ln x \;=\; \frac{d^{m-1}}{dx^{m-1}}x^{-1} \;=\; (-1)(-2)\cdot\ldots\cdot(1{-}m)x^{-m} \;=\; (-1)^{m-1}(m{-}1)!\,\frac{1}{x^m}\,.$$

Für die natürliche Exponentialfunktion ist die Ableitung erster Ordnung die Funktion selbst, und das gilt dann auch für alle folgenden Ableitungen:

$$\frac{d^m}{dx^m}\, \mathrm{e}^x \;=\; \mathrm{e}^x \qquad \text{für alle } \; m \in \mathbb{N}\,.$$

Für allgemeine Exponentialfunktionen $a^x = \mathrm{e}^{cx}$ (mit $a > 0$ und $c = \ln a$) folgt mit $\frac{d}{dx}\mathrm{e}^{cx} = c \cdot \mathrm{e}^{cx}$ dann

$$\frac{d^m}{dx^m}\, \mathrm{e}^{cx} \;=\; c^m \mathrm{e}^{cx}\,, \qquad \frac{d^m}{dx^m}\, a^x \;=\; (\ln a)^m a^x\,.$$

Weil die Kreisfunktion $\sin$ als Ableitung $\cos$ hat und $\cos$ als Ableitung $-\sin$, ist $-\sin$ die zweite Ableitung von $\sin$ und ebenso $-\cos$ die zweite Ableitung von $\cos$. Diese beiden Funktionen lösen also die Differentialgleichung $f'' = -f$, und darin besteht z.B. ihre Bedeutung in der Physik. (Sie beschreiben Schwingungsvorgänge, bei denen die wirkende Gegenkraft zur Auslenkung proportional ist.) Die Ableitungen höherer Ordnung ergeben sich daraus dann auch sofort: Es ist $\sin^{(m)} = \sin, \cos, -\sin$ bzw. $-\cos$, wenn m mit Rest 0, 1, 2 bzw. 3 durch 4 teilbar ist, und dasselbe gilt, wenn man $\sin$ durch $\cos$ und $\cos$ durch $-\sin$ ersetzt.

10) Partielle Ableitungen höherer Ordnung erklärt man durch wiederholte Ableitung nach derselben oder auch nach verschiedenen Variablen. Mit einer Notation, die sich nun von selbst erklärt, gilt dann zum Beispiel:

$$\frac{\partial^3}{\partial x^2 \partial y} x^m y^n = m(m-1)n \cdot x^{m-2} y^{n-1} ,$$

$$\frac{\partial^2}{\partial x \partial y}\, x^y = \frac{\partial}{\partial x}(x^y \ln x) , \qquad \frac{\partial^2}{\partial y \partial x}\, x^y = \frac{\partial}{\partial y}(y x^{y-1}) ,$$

und die verbliebenen Ableitungen erster Ordnung kann man mit Hilfe der Produktregel für das Differenzieren (siehe unten) leicht ausrechnen. Das Ergebnis ist in beiden Fällen die Funktion $x^{y-1} + yx^{y-1} \ln x$, und das ist kein Zufall; denn bei “normalen” Funktionen, insbesondere bei Funktionen, die durch Rechenterme mit mehreren Variablen gegeben sind wie hier, wird in der Mathematik gezeigt:

- *Die partiellen Ableitungen höherer Ordnung sind unabhängig davon, in welcher Reihenfolge man die einzelnen Differentiationen bei c.p.-Bedingung ausführt.* ■

Wir haben nun die Ableitungen aller Grundfunktionen berechnet, aus denen die elementaren Funktionen aufgebaut werden (und noch einige Ableitungen mehr). Um alle elementaren Funktionen differenzieren zu können, brauchen wir noch Rechenregeln, mit denen wir die Ableitung “zusammengesetzter Funktionen“ aus den Ableitungen der Bestandteile berechnen können. In der Wirtschaftsmathematik kommen nur verhältnismäßig einfach zusammengesetzte Funktionen vor, deren Ableitung man mit Kenntnis dieser Rechenregeln — und der Ableitungen der Grundfunktionen — direkt ohne großen Rechenaufwand hinschreiben kann. Im Prinzip aber kann man mit Hilfe der Ableitungsrechenregeln *jede* elementare Funktion differenzieren und ihre Ableitung wieder als elementare Funktion darstellen, wie kompliziert der Rechenterm, der die Funktion definiert, auch sei!

RECHENREGELN für die Berechnung von Ableitungen und BEISPIELE:

1) Ist $f : I \to \mathbb{R}$ eine differenzierbare Funktion auf dem Intervall I (mit positiver Länge, also nicht einpunktig oder gar leer) in $\mathbb{R}$, so gilt für alle Konstanten $a, b, c, d \in \mathbb{R}$:

- **erste Faktorregel**: $\boxed{\frac{d}{dx}\,(cf(x)) = cf'(x)}$ für $x \in I$;

  einen konstanten Faktor kann man aus der Ableitung “herausziehen”;

- **erste Verschiebungsregel**: $\boxed{\frac{d}{dx}\,(f(x) + d) = f'(x)}$ für $x \in I$;

  Addition einer Konstanten ändert die Ableitung nicht (“**Konstantenregel**”);

- **zweite Faktorregel**: $\boxed{\frac{d}{dx}\, f(ax) = af'(ax)}$ für $ax \in I$;

  ein Faktor vor der unabhängigen Variablen tritt beim Ableiten zusätzlich außen auf.

- **zweite Verschiebungsregel**: $\boxed{\frac{d}{dx} f(x+b) = f'(x+b)}$ für $x + b \in I$;

  bei einer horizontalen Verschiebung der Funktion wird die Ableitung mitverschoben.

Diese Regeln lassen sich zusammenfassen zu der

- **Skalierungsregel:** $\boxed{\frac{d}{dx}\,\big[cf(ax{+}b) + d\big] = ac \cdot f'(ax+b)}$ für $ax + b \in I$.

Diese Rechenregeln sind durch einfache Überlegungen mit Differenzenquotienten sehr leicht zu beweisen. Die Differenzenquotienten von $f(x) + d$ sind dieselben wie die von f und die von $cf(x)$ unterscheiden sich von denen zu f nur um den Faktor c; daraus folgen sofort die beiden ersten Regeln. Die Funktion $f(x+b)$ hat an der Stelle x dieselben Differenzenquotienten $\frac{1}{h}[f(x+h+b) - f(x+b)]$ wie die Funktion $f(x)$ an der Stelle $x+b$; daraus ergibt sich die zweite Verschiebungsregel. Die zweite Faktorregel schließlich ist klar für $a = 0$ und folgt für $a \neq 0$ aus der Identität $\frac{1}{h}[f(a(x+h)) - f(ax)] = a\frac{1}{k}[f(ax+k) - f(ax)]$ für $k := ah$, weil mit h auch k gegen Null geht.

Der Übergang $x \mapsto x + b$ bzw. $y = f(x) \mapsto y + d$ entspricht einer Verschiebung des Nullpunkts der Zahlenskala, in der die Größen x bzw. y gemessen werden. Geometrisch läuft das auf eine horizontale bzw. vertikale Verschiebung des Graphen hinaus, und die beiden Verschiebungsregeln besagen einfach, dass dabei die Steigung des Graphen (in einander entsprechenden Punkten) unverändert bleibt. Der Übergang $x \mapsto ax$ bzw. $y \mapsto cy$ entspricht im Fall $a > 0$ bzw. $c > 0$ einer Maßstabänderung, also einem Wechsel der Maßeinheit, auf der Zahlenskala. Dies ist in ökonomischem Kontext häufig erforderlich aufgrund unterschiedlicher Einheiten (Währung, Längeneinheiten, Gewichtseinheiten, $\ldots$) in verschiedenen Ländern. Im Fall $a < 0$ bzw. $c < 0$ erfolgt zusätzlich noch eine Richtungsumkehr auf der Skala. Die Faktorregeln drücken aus, wie sich bei einer solchen Reskalierung der Variablen die Ableitung ändert. Geometrisch besagen sie, dass sich die Steigungen des Graphen um einen Faktor ändern, wenn man ihn in vertikaler Richtung mit diesem Faktor streckt bzw. staucht oder wenn man ihn in horizontaler Richtung mit dem reziproken Faktor staucht bzw. streckt; anschaulich ist das unmittelbar klar. Die Skalierungsregel für Transformation $x \mapsto ax + b$ der unabhängigen Variablen ist ein einfacher Spezialfall der späteren Kettenregel; man braucht sie sich daher nicht zu merken, wenn man die Kettenregel beherrscht. Wegen ihrer geometrischen und ökonomischen Bedeutung gehört diese Regel aber zu den entsprechenden Regeln für Skalierung der abhängigen Variablen, also zur ersten Faktor- und Verschiebungsregel.

2) Bei Anwendung der obigen Rechenregeln ist — wie bei den folgenden Rechenregeln — darauf zu achten,

- *dass die Funktionen und ihre Ableitung in einer Gleichung, die eine Rechenregel darstellt, auf beiden Seiten stets an denselben Stellen ausgewertet werden.*

Deshalb kann in der zweiten Faktorregel rechts nur $f'(ax)$ auftreten, nicht etwa $f'(x)$, und in der Skalierungsregel muss rechts f' an derselben Stelle $ax + b$ ausgewertet werden wie die linke Seite. Eine "Regel $\frac{d}{dx}f(ax) = af'(x)$" wäre (außer für $a = 1$) völlig unsinnig! Sie kann schon deswegen nicht allgemein gültig sein, weil f' an der Stelle x gar nicht definiert sein muss, wenn f nahe ax definiert ist und $ax \neq x$; und selbst wenn $f'(x)$ definiert sein sollte, so hat dieser Wert mit der Ableitung $\frac{d}{dx}f(ax)$, die nur von den Werten von f beliebig nahe bei der Stelle ax abhängt, im Allgemeinen nichts zu tun.

3) *Beispiele für die Anwendung der Skalierungsregeln*:

$$\frac{d}{dx}(cx^s) = c\frac{d}{dx}\,x^s = csx^{s-1}\,,$$

ergibt sich mit der zweiten Faktorregel, mit der ersten dagegen

$$\frac{d}{dx}(cx)^s = c \cdot \frac{d}{dy}\,\Big|_{y=cx}\,y^s = c \cdot s(cx)^{s-1} = c^s \cdot sx^{s-1}\,,$$

was hier natürlich mit $(cx)^s = c^s x^s$ auch (einfacher) aus der ersten Faktorregel folgt.

Ebenfalls mit erster und zweiter Faktorregel berechnet man

$$\frac{d}{dx}(c\mathrm{e}^x) = c\mathrm{e}^x\,, \qquad \frac{d}{dx}\mathrm{e}^{cx} = c\mathrm{e}^{cx}\,,$$

$$\frac{d}{dx}\,a^x = \frac{d}{dx}\,\mathrm{e}^{x\ln a} = (\ln a)\mathrm{e}^{x\ln a} = a^x\ln a\,,$$

was schon bekannt ist (und früher unter Vorwegnahme der Faktorregeln gezeigt wurde).

Noch drei Beispiele zur Anwendung der Skalierungsregel:

$$\frac{d}{dx}\underbrace{(2x+5)^3}_{\substack{=f(2x+5)\\ \text{mit } f(y):=y^3}} = 2\cdot\underbrace{3(2x+5)^2}_{=f'(y)=3y^2} = 6\cdot(2x+5)^2\;;$$

$$\frac{d}{dx}\underbrace{\frac{1}{\sqrt{2-x}}}_{\substack{=f(2-x)\\ \text{mit } f(y):=1/\sqrt{y}}} = (-1)\cdot\underbrace{\frac{-1}{2\sqrt{2-x}^{\,3}}}_{=f'(2-x)} = \frac{1}{2(2-x)\sqrt{2-x}}\;;$$

$$\frac{d}{dx}\frac{3x-1}{2x+7} = \frac{d}{dx}\left(\frac{3}{2}-\frac{23}{2}\underbrace{\frac{1}{2x+7}}_{\substack{=f(2x+7)\\ \text{mit } f(y):=1/y}}\right) = \frac{-23}{2}\cdot 2\cdot\frac{-1}{(2x+7)^2} = \frac{23}{(2x+7)^2}$$

So kann man offenbar jede gebrochen lineare Funktion ableiten.

4) Hier noch zwei Anwendungen der Faktor- und Konstantenregel, eine mathematische und eine in der Ökonomie. Gemäß 4.2 heißt eine Funktion f auf einem zu 0 symmetrischen Intervall gerade, wenn $g(x) = g(-x)$ gilt. Differenzieren wir beide Seiten dieser Gleichung, so erhalten wir mit der zweiten Faktorregel

$$g'(x) = \frac{d}{dx}\,g(-x) = (-1)\cdot g'(-x) = -g'(-x)\,.$$

Analog folgt für eine ungerade Funktion f, also $f(x) = -f(-x)$ die Gleichung $f'(x) = f'(-x)$ für die Ableitung. Differenzieren ändert also die Parität von Funktionen, d.h.

- *Die Ableitung einer geraden Funktion ist ungerade, die Ableitung einer ungeraden Funktion ist gerade.*

Die zweite Anwendung ist folgender "ökonomischer Lehrsatz":

Die Grenzkosten sind unabhängig von den Fixkosten.

Dies bedeutet, wenn die Kostenfunktion $K(x) = K_{\text{fix}} + K_{\text{var}}(x)$ aus konstanten Fixkosten K_{fix} und variablen Kosten $K_{\text{var}}(x)$ zusammengesetzt ist, dass die Ableitung $K'(x)$ (das ist ja mit der ökonomischen Terminologie "Grenzkosten" gemeint) unabhängig von den Fixkosten ist — und das ist klar, da ja nach der Konstantenregel $\frac{d}{dx}\,K(x) = \frac{d}{dx}\,K_{\text{var}}(x)$ ist. Der "ökonomische Lehrsatz" ist also einfach eine Anwendung der ersten Verschiebungsregel für die Ableitung. Auch andere "ökonomische Lehrsätze" erweisen sich später als einfache Anwendungen von Ableitungsrechenregeln.

5) Sind f und g differenzierbare Funktionen auf demselben Intervall I (positiver Länge) in $\mathbb{R}$, so gilt auf I die

- **Summenregel und Differenzregel:** $\boxed{(f \pm g)' \;=\; f' \pm g'}$;

  die Ableitung einer Summe / Differenz von Funktionen ist die Summe / Differenz ihrer Ableitungen.

Das ist leicht einzusehen: Die Differenzenquotienten $\frac{1}{h}[(f+g)(x+h)-(f+g)(x)]$ von $f+g$ sind ja gleich der Summe $\frac{1}{h}[f(x+h)-f(x)]+\frac{1}{h}[g(x+h)-g(x)]$ der Differenzenquotienten von f und von g, und wenn letztere bei $h \to 0$ gegen $f'(x)$ bzw. $g'(x)$ streben, so erstere offenbar gegen $f'(x)+g'(x)$. Dieselbe Regel gilt natürlich auch für Summen von mehr als zwei Funktionen, und wenn wir sie mit der ersten Faktorregel kombinieren, so erhalten wir für differenzierbare Funktionen $f_1, \ldots, f_n$ auf I und Konstanten $c_1, \ldots, c_n \in \mathbb{R}$ folgende **Verallgemeinerung der Summen- und Faktorregel**:

$$(c_1 f_1 + c_2 f_2 + \ldots + c_n f_n)' \;=\; c_1 f_1' + c_2 f_2' + \ldots + c_n f_n' \;.$$

Man sagt dazu:

- *Eine Linearkombination von Funktionen darf man gliedweise differenzieren.*

Außerdem gelten die Regeln offenbar auch für die zweite, dritte, ... Ableitung, wenn die Funktionen entsprechend oft differenzierbar sind.

- *Die (verallgemeinerte) Summenregel gilt auch für Ableitungen höherer Ordnung.*

6) *Beispiele für die Anwendung der Summenregel*: Die **Ableitung der Hyperbelfunktionen** $\cosh x = \frac{1}{2}(\mathrm{e}^{x}+\mathrm{e}^{-x})$ und $\sinh x = \frac{1}{2}(\mathrm{e}^{x}-\mathrm{e}^{-x})$ ist

$$\cosh' x = \frac{d}{dx}(\tfrac{1}{2}\mathrm{e}^{x} + \tfrac{1}{2}\mathrm{e}^{-x}) = \tfrac{1}{2}\mathrm{e}^{x} - \tfrac{1}{2}\mathrm{e}^{-x} = \sinh x \;,$$
$$\sinh' x = \frac{d}{dx}(\tfrac{1}{2}\mathrm{e}^{x} - \tfrac{1}{2}\mathrm{e}^{-x}) = \tfrac{1}{2}\mathrm{e}^{x} + \tfrac{1}{2}\mathrm{e}^{-x} = \cosh x \;.$$

Es folgt, dass die m–te Ableitung von $\cosh$ gleich $\cosh$ und die von $\sinh$ gleich $\sinh$ ist, wenn m eine gerade Zahl ist, während für ungerade Ableitungsordnung m die m–te Ableitung von $\cosh$ gleich $\sinh$ ist und die von $\sinh$ gleich $\cosh$.

Die **Ableitung von Polynomfunktionen** wird gliedweise berechnet:

$$\frac{d}{dx}\big(a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0\big) = a_n n x^{n-1} + \ldots + a_2 2x + a_1 \;.$$

Man braucht also nur jede Potenz x^k der unabhängigen Variablen durch ihre Ableitung kx^{k-1} zu ersetzen. Insbesondere sind alle Polynomfunktionen differenzierbar auf $\mathbb{R}$, und ihre Ableitung ist wieder eine Polynomfunktion. Genauer gilt:

- *Die Ableitung eines Polynoms vom Grad n ist ein Polynom vom Grad $n-1$;*
- *die n–te Ableitung ist daher konstant und die $(n+1)$–te ist Null.*

Der konstante Wert der n–ten Ableitung ist $n!a_n$. Später werden wir sehen, dass umgekehrt jede Funktion, deren Ableitung $(n+1)$–ter Ordnung auf einem Intervall überall Null ist, schon eine Polynomfunktion von Grad höchstens n ist.

7) Zwei "ökonomische Lehrsätze":

Der Grenzgewinn ist die Differenz von Grenzerlös und Grenzkosten.

Die marginale Konsumquote und die marginale Sparquote ergänzen sich zu 1.

Die erste Aussage ist $G' = E' - K'$ für die Gewinnfunktion G, die ja als Differenz von Erlösfunktion E und Kostenfunktion K definiert ist; es handelt sich also einfach um die Differenzregel für die Ableitung. Die zweite Aussage bezieht sich auf die vom Einkommen Y abhängige Konsumfunktion $C(Y)$ (der Anteil der konsumiert wird) und die zugehörige Sparfunktion $S(Y) = Y - C(Y)$ (was nicht konsumiert wird, wird gespart). Es gilt also $C(Y) + S(Y) = Y$, und wenn man beide Seiten nach Y differenziert, so ergibt sich mit der Summenregel $C'(Y) + S'(Y) = 1$. Das ist genau der Inhalt der zweiten Aussage, die also nur eine Anwendung der Summenregel ist.

Ensprechend gilt: *der Grenzstückgewinn ist die Differenz von Grenz(stück)preis und Grenzstückkosten*, d.h. $g'(x) = p'(x) - k'(x)$ für den Stückgewinn $g(x) = G(x)/x$, den Preis $p(x) = E(x)/x$ (durchschnittlicher Marktpreis bei Absatz des Output x) und die Stückkosten $k'(x) = K(x)/x$ in Abhängigkeit vom Produktions-Output x. Die Ableitungen beschreiben jeweils, um wieviele Geldeinheiten sich die Stückfunktionswerte verändern, wenn der Output um eine (kleine) Einheit erhöht wird. Den unter Ignorierung der Fixkosten entstehenden Gewinn $G_{\text{deck}}(x) = E(x) - K_{\text{var}}(x)$ nennt man **Deckungsbeitrag**; dies ist, wenn man die Fixkosten in jedem Fall zu tragen hat, die für die Entscheidung über die Aufnahme und den Umfang der Produktion maßgebende Größe. Da sich $G(x)$ und $G_{\text{deck}}(x)$ nur um die Konstante K_{fix} unterscheiden, haben sie dieselbe Ableitung: *Der Grenzgewinn ist gleich dem Grenzdeckungsbeitrag.* Aber der Grenzstückgewinn $g'(x)$ ist bei positiven Fixkosten nicht gleich dem Grenzstückdeckungsbeitrag $g'_{\text{deck}}(x)$; denn es ist $g_{\text{deck}}(x) = G_{\text{deck}}(x)/x = (G(x) + K_{\text{fix}})/x = g(x) + K_{\text{fix}}/x$ und daher $g'_{\text{deck}}(x) = g'(x) - K_{\text{fix}}/x^2$. (Für großen Produktions–Output x ist aber der Unterschied K_{fix}/x^2 vernachlässigbar klein.)

8) Sind f und g differenzierbare Funktionen auf demselben Intervall I (mit positiver Länge) in $\mathbb{R}$ so gelten:

- **Produktregel:** $\boxed{(f \cdot g)' = f' \cdot g + f \cdot g'}$;

  die Ableitung eines Produkts von zwei Funktionen bildet man, indem man jede mit der Ableitung der anderen multipliziert und die beiden Produkte addiert;

- **Reziprokenregel:** $\boxed{\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}}$, *wo* $g \neq 0$ *ist;*

  die Ableitung des Reziproken einer Funktion ist der Quotient aus der Ableitung und dem negativen Quadrat dieser Funktion;

- **Quotientenregel:** $\boxed{\left(\frac{f}{g}\right)' = \frac{gf' - fg'}{g^2}}$, *wo* $g \neq 0$ *ist;*

  die Ableitung eines Quotienten von zwei Funktionen erhält man, indem man den Nenner quadriert und im Zähler die Differenz aus dem Produkt der Nennerfunktion mit der Ableitung der Zählerfunktion und dem Produkt der Zählerfunktion mit der Ableitung der Nennerfunktion bildet.

Zum Beweis der Produktregel formt man so um: $\frac{1}{h}[f(x+h)g(x+h) - f(x)g(x)] = \frac{1}{h}[f(x+h) - f(x)]g(x+h) + f(x)\frac{1}{h}[g(x+h) - g(x)]$. Da $g(x+h)$ gegen $g(x)$ strebt mit $h \to 0$ (die Funktion ist ja stetig an der Stelle x, wenn sie dort sogar differenzierbar ist), ergibt sich mit der Konvergenz der Differenzenquotienten von f und g gegen die Ableitungen daraus sofort die Produktregel. Die Differenzenquotienten von $1/g$ formen wir ähnlich um zu $\frac{1}{h}[1/g(x+h) - 1/g(x)] = -\frac{1}{h}[g(x+h) - g(x)]/[g(x+h)g(x)]$ und sehen daraus die Konvergenz gegen $-g'(x)/g(x)^2$ bei $h \to 0$. Das ist die Reziprokenregel, und die Quotientenregel für $f/g = f \cdot (1/g)$ folgt daraus sofort mit der Produktregel. Der Beweis zeigt auch, dass mit f und g auch fg differenzierbar ist und ebenso f/g an Stellen x mit $g(x) \neq 0$. Für Produkte von drei differenzierbaren Funktionen auf I gilt dann $(fgh)' = (fg)'h + (fg)h' = f'gh + fg'h + fgh'$, und für Produkte von beliebig vielen differenzierbaren Funktionen auf demselben Intervall findet man

$$\begin{aligned}
(f_1 f_2 \cdot \ldots \cdot f_n)' &= \\
&= f_1' f_2 \cdot \ldots \cdot f_n + f_1 f_2' f_3 \cdot \ldots \cdot f_n + \ldots + f_1 \cdot \ldots \cdot f_{n-1} f_n' \\
&= f_1 f_2 \cdot \ldots \cdot f_n \left(\frac{f_1'}{f_1} + \frac{f_2'}{f_2} + \ldots + \frac{f_n'}{f_n} \right) ,
\end{aligned}$$

letzteres natürlich nur an Stellen, wo keine der Funktionen f_i den Wert 0 hat. Für die Ableitung eines Produkts von mehreren reziproken Funktionen kann man ähnliche Formeln herleiten, die weniger praktische Bedeutung haben. Für den Spezialfall eines Produkts von n gleichen Faktoren f oder $1/g$ notieren wir noch folgende Formeln für die **Ableitung einer Potenz einer Funktion** mit Exponent $n \in \mathbb{N}$:

$$(f^n)' = n f^{n-1} f' , \qquad \left(\frac{1}{g^n}\right)' = \frac{-ng'}{g^{n+1}} , \quad \text{wo } \ g \neq 0 \ \text{ ist.}$$

9) *Beispiele für die Anwendung der Produktregel*:

$$\frac{d}{dx}(x\mathrm{e}^x) = \left(\frac{d}{dx}\,x\right)\mathrm{e}^x + x\frac{d}{dx}\,\mathrm{e}^x = 1 \cdot \mathrm{e}^x + x \cdot \mathrm{e}^x = (1+x)\mathrm{e}^x ,$$

$$\frac{d}{dx}(x^s \mathrm{e}^{cx}) = (sx^{s-1})\mathrm{e}^{cx} + x^s(c\mathrm{e}^{cx}) = (s+cx)x^{s-1}\mathrm{e}^{cx} ,$$

für $c \in \mathbb{R}$, $x > 0$ und beliebige $s \in \mathbb{R}$, oder $x \neq 0$ im Fall $s \in \mathbb{Z}$ oder alle $x \in \mathbb{R}$ im Fall $s \in \mathbb{N}$. Noch ein Beispiel mit drei Faktoren:

$$\frac{d}{dt}(t^5 \mathrm{e}^{-2t} \ln t) = 5t^4 \mathrm{e}^{-2t} \ln t + t^5(-2\mathrm{e}^{-2t}) \ln t + t^5 \mathrm{e}^{-2t}\frac{1}{t} \quad \text{für } t > 0 .$$

Ableitungen von Potenzen einer Funktion:

$$\frac{d}{dx}(x^2+1)^{100} = 100(x^2+1)^{99}\frac{d}{dx}(x^2+1) = 200x \cdot (x^2+1)^{99} ,$$

$$\frac{d}{dt}\cosh^4 t = 4(\cosh^3 t)\frac{d}{dt}\cosh t = 4(\cosh^3 t)\sinh t ,$$

$$\frac{d}{dy}\frac{1}{(y^3 - 2y + 1)^2} = \frac{-2(3y^2 - 2)}{(y^3 - 2y + 1)^3} \qquad (\text{ wo der Nenner} \neq 0 \text{ ist}) .$$

Die noch ausstehende gemischte partielle Ableitung zweiter Ordnung der Funktion x^y können wir nun auch berechnen (für $x > 0$ und $y \in \mathbb{R}$):

$$\begin{aligned}
\frac{\partial^2}{\partial x \partial y}\, x^y &= \frac{\partial}{\partial x}(x^y \ln x) = \Big(\frac{\partial}{\partial x} x^y\Big) \ln x + x^y \Big(\frac{d}{dx} \ln x\Big) \\
&= y x^{y-1} \ln x + x^y \frac{1}{x} = (1 + y \ln x) x^{y-1} \,, \\
\frac{\partial^2}{\partial y \partial x}\, x^y &= \frac{\partial}{\partial y}(y x^{y-1}) = \Big(\frac{d}{dy}\, y\Big) x^{y-1} + y \Big(\frac{\partial}{\partial y}\, x^{y-1}\Big) \\
&= 1 \cdot x^{y-1} + y \cdot x^{y-1} \ln x = (1 + y \ln x) x^{y-1} \,.
\end{aligned}$$

Wir haben schon gesagt, dass die Übereinstimmung der beiden Ergebnisse kein Zufall ist; bei elementaren Funktionen wie hier ist das Ergebnis der partiellen Differentiationen höherer Ordnung immer unabhängig davon, in welcher Reihenfolge die einzelnen partiellen Ableitungen ausgeführt werden. Man erkennt, dass man mit weiterer Anwendung von Produkt und Summenregel nun auch beliebige partielle Ableitungen höherer Ordnung der Funktion x^y berechnen kann (jedenfalls im Prinzip). Zum Beispiel ist

$$\begin{aligned}
\frac{\partial^3}{\partial x^2 \partial y}\, x^y &= \frac{\partial}{\partial x}[(1 + y \ln x) x^{y-1}] \\
&= y \frac{1}{x} x^{y-1} + (1 + y \ln x)(y-1) x^{y-2} = [2y - 1 + y(y-1) \ln x] x^{y-2} \,.
\end{aligned}$$

10) *Beispiele für die Anwendung der Quotientenregel*:

$$\begin{aligned}
\frac{d}{dx}\, \frac{1}{\ln x} &= \frac{-\frac{d}{dx} \ln x}{(\ln x)^2} = \frac{-1}{x(\ln x)^2} && (x > 0)\,, \\
\frac{d}{dy}\, \frac{\ln y}{y} &= \frac{y \frac{d}{dy} \ln y - (\ln y) \frac{d}{dy}\, y}{y^2} = \frac{1 - \ln y}{y^2} && (y > 0)\,, \\
\frac{d}{dx}\, \frac{x^2 - 2x + 1}{x^3 - 2x^2 - 2} &= \frac{(x^3 - 2x^2 - 2)(2x - 2) - (x^2 - 2x + 1)(3x^2 - 4x)}{(x^3 - 2x^2 - 2)^2} \,,
\end{aligned}$$

letzteres an den Stellen $x \in \mathbb{R}$, an denen der Nenner nicht verschwindet.

Ableitung der Tangensfunktionen $\tanh = \dfrac{\sinh}{\cosh}$ und $\tan = \dfrac{\sin}{\cos}$:

$$\begin{aligned}
\frac{d}{dx} \tanh x &= \frac{d}{dx}\, \frac{\sinh x}{\cosh x} = \frac{(\cosh x) \frac{d}{dx} \sinh x - (\sinh x) \frac{d}{dx} \cosh x}{(\cosh x)^2} \\
&= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} = \frac{1}{\cosh^2 x} = 1 - \tanh^2 x \,, \\
\frac{d}{dx} \tan x &= \frac{d}{dx}\, \frac{\sin x}{\cos x} = \frac{(\cos x) \frac{d}{dx} \sin x - (\sin x) \frac{d}{dx} \cos x}{(\cos x)^2} \\
&= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x \,.
\end{aligned}$$

Die **Ableitung rationaler Funktionen** $p(x)/q(x)$ mit Polynomen p und q ist

$$\frac{d}{dx}\, \frac{p(x)}{q(x)} = \frac{q(x) p'(x) - p(x) q'(x)}{q(x)^2} \,,$$

an allen Stellen, wo das Nennerpolynom $q(x) \neq 0$ ist. Ein konkretes Beispiel mit p vom Grad 2 und q vom Grad 3 haben wir oben schon behandelt.

Man erkennt, dass die Ableitung wieder eine rationale Funktion ist. Im Unterschied zur Differentiation von Polynomfunktionen, die den Grad erniedrigt und so die Funktion durch Ableiten "einfacher" macht, erhöhen sich aber beim Ableiten von rationalen Funktionen leider im Allgemeinen der Nennergrad und der Zählergrad, so dass die abgeleitete Funktion eine "kompliziertere" rationale Funktion ist als die ursprüngliche. (Das ist besonders unangenehm, wenn man mehr als einmal ableiten will.) Genauer gilt offenbar:

- *Rationale Funktionen p/q sind differenzierbar an den Stellen, wo das Nennerpolynom nicht verschwindet; hat p den Grad m und q den Grad n, so lässt sich die Ableitung $(p/q)'$ als Quotient eines Polynoms vom Grad $\leq m{+}n{-}1$ und des Polynoms q^2 vom Grad $2n$ darstellen.*

Unter Umständen lässt sich die Darstellung der Ableitung durch Kürzen noch vereinfachen (z.B. immer, wenn q Nullstellen der Ordnung ≥ 2 hat, weil diese auch Nullstellen von q' sind, wie man zeigen kann). Wenn $m = n$ ist, also p und q denselben Grad haben, so heben sich außerdem im Zähler $qp' - pq'$ die Terme mit x^{2n-1} heraus, so dass der Zählergrad echt kleiner ist als $2n - 1$. Gebrochen-lineare Funktionen ($m = 1 = n$) haben z.B. als Ableitung immer das Reziproke eines quadratischen Polynoms, wenn sie nicht linear sind und somit konstante Ableitung haben.

11) Ableitung von Stückfunktionen / Durchschnittsfunktionen: Ist $F(x)$ eine für $x{>}0$ definierte Funktion, so heißt $f(x) := F(x)/x$ die zu F gehörige *Stückfunktion* oder *Durchschnittsfunktion*. Wir kennen das schon vom Beispiel einer Kostenfunktion $K(x)$, wo $k(x) = K(x)/x$ die Stückkosten angibt, also die Durchschnittskosten, die pro Einheit bei der Produktion von insgesamt x Einheiten entstehen. Die Stückfunktion wird oft auch mit einem Querstrich bezeichnet, d.h. man schreibt $\overline{h}(x) := h(x)/x$ für die Stückfunktion zu einer gegebenen Funktion $h(x)$. Für die *Grenzstückfunktion*, also die Ableitung der Stückfunktion, gilt gemäß der Quotientenregel:

$$f'(x) = \frac{d}{dx}\,\frac{F(x)}{x} = \frac{xF'(x) - F(x)\cdot 1}{x^2} = \frac{F'(x)}{x} - \frac{F(x)}{x^2} = \frac{F'(x) - f(x)}{x}\ .$$

Man kann das Ergebnis im wirtschaftswissenschaftlichen Jargon (nicht ganz ernst gemeint) so formulieren und — wichtiger — eine Folgerung über das Monotonieverhalten der Stückfunktion ziehen:

- *Die Grenzstückfunktion ist die Durchschnittsfunktion zur Differenz der Grenzfunktion und ihrer Stückfunktion.*
- *In Bereichen, in denen die Grenzfunktion F' größer ist als die Stückfunktion f ist die Grenzstückfunktion positiv, die Stückfunktion also wachsend; in Bereichen, in denen die Grenzfunktion kleiner ist als die Stückfunktion, ist die Grenzstückfunktion negativ, die Stückfunktion also fallend.*

(Dass eine Funktion auf einem Intervall, wo ihre Ableitung positiv ist, tatsächlich streng wächst, beweisen wir in 4.6.) Interpretation: Der Wert der Stückfunktion zu $F(x)$ wird größer (bzw. kleiner) bei Erhöhung von x um eine (kleine) Einheit, wenn dabei die Änderung von $F(x)$ größer (bzw. kleiner) ist als der Durchschnittswert $F(x)/x$. Das ist auch ökonomisch betrachtet sehr plausibel. Dieser Zusammenhang zwischen einer Funktion, der zugehörigen Durchschnittsfunktion / Stückfunktion und den Ableitungen wird oft gebraucht, z.B. mit folgender Terminologie:

- $K(x)$ *Kostenfunktion* (für die Produktion von x Output-Einheiten),
  $k(x) = K(x)/x$ *Stückkostenfunktion,*
  $K'(x)$ *Grenzkosten* ($\approx$ Erhöhungung von $K(x)$ bei Produktion einer Mehreinheit),
  $k'(x)$ *Grenzstückkosten* ($\approx$ Änderung von $k(x)$ bei Produktion einer Mehreinheit);

- $E(x)$ *Erlösfunktion, Umsatzfunktion* (für x am Markt abgesetzte Einheiten),
  $p(x) = E(x)/x$ *(Stück–)Preis* (am Markt erzielter Preis bei Absatz von x Einheiten),
  $E'(x)$ *Grenzerlös* ($\approx$ Mehrerlös bei Absatz einer Mehreinheit),
  $p'(x)$ *Grenzpreis* ($\approx$ Marktpreisänderung bei Absatz einer Mehreinheit);

- $x(r)$ *Produktionsfunktion* (Output in Abhängigkeit von Produktionsfaktor r),
  $\overline{x}(r) := x(r)/r$ *Durchschnittsproduktionsertrag, durchschnittliche Produktivität,*
  $x'(r)$ *Grenzproduktivität, Grenzproduktionsertrag* ($\approx$ Erhöhung des Outputs bei Mehreinsatz einer Faktoreinheit),
  $\overline{x}'(r)$ *Grenzdurchschnittsertrag* ($\approx$ Änderung des Durchschnittsertrags bei Mehreinsatz einer Faktoreinheit);

- $r(x)$ *Faktoreinsatzfunktion, Faktorverbrauchsfunktion* (Umkehrfunktion zu $x(r)$, gibt den Faktorverbrauch bei Produktion von x Einheiten an),
  $\overline{r}(x) := r(x)/x$ *Produktionskoeffizient* (damit ist die Zahl der produzierten Einheiten zu multiplizieren, um den Faktorverbrauch zu erhalten),
  $r'(x)$ *Grenzverbrauchsfunktion* ($\approx$ Erhöhung des Faktorverbrauchs bei Produktion einer zusätzlichen Einheit),
  $\overline{r}'(x)$ *Grenzproduktionskoeffizient* ($\approx$ Änderung des Durchschnittsfaktorverbrauchs für x produzierte Einheiten, wenn eine Einheit mehr produziert wird);

- $C(Y)$ *Konsumfunktion, Konsumquote* (in Abhängigkeit vom Einkommen Y),
  $c(Y) := C(Y)/Y$ *durchschnittliche Konsumquote,*
  $C'(Y)$ *Grenzhang zum Konsum, marginale Konsumquote* ($\approx$ Erhöhung von $C(Y)$ bei Einkommenserhöhung um eine Geldeinheit),
  $c'(Y)$ *marginale durchschnittliche Konsumquote* ($\approx$ Änderung der durchschnittlichen Konsumquote bei Einkommenserhöhung um eine Geldeinheit).

Soweit die Funktionswerte noch von andern Variablen abhängen, z.B. der Output von weiteren Produktionsfaktoren, ist beim Differenzieren die c.p.-Bedingung, also Konstanz dieser Variablen, unterstellt. Mathematisch gesehen ist für die Berechnung dieser Ableitungen nicht viel erforderlich: Die Summenregel und die Quotientenregel in sehr einfachen Spezialfällen. Das eigentliche Problem ist nicht die Berechnung von Funktionen, sondern die Entschlüsselung von Wortungetümen wie "Grenzproduktionsertrag" (oben $\frac{d}{dr}x(r)$), "marginale durchschnittliche Konsumquote" (oben $\frac{d}{dY}(C(Y)/Y)$) oder gar "Grenzdurchschnittsproduktionsfaktorverbrauchsfunktion" (oben $\frac{d}{dx}(r(x)/x)$). Die ökonomische Terminologie ist aber so gewählt, dass man nach etwas Eigewöhnung an den Namen der Funktionen jeweils erkennen kann, ob F, $f = \overline{F}$, F' oder $f' = \overline{F}'$ gemeint ist. Die Reihenfolge der im Namen vorangestellten Präfixe, die Operationen mit Funktionen bezeichnen, ist dabei von rechts nach links zu lesen. Also ist die "Grenzstückfunktion" zu $F(x)$ die Grenzfunktion der Stückfunktion $f(x) = F(x)/x$, d.h. die Ableitung $f'(x) = \frac{d}{dx}(F(x)/x)$; dagegen ist die "Stückgrenzfunktion" zu $F(x)$ die Durchschnittsfunktion der Grenzfunktion $F'(x)$, d.h. gleich $F'(x)/x$ (das ist etwas Anderes!).

Wir erläutern die Bedeutung der Berechnung von der Ableitung der Durchschnittsfunktionen noch am Beispiel der Konsumquote $C(Y)$ und Sparquote $S(Y) = Y - C(Y)$ in Abhängigkeit vom Einkommen Y. Nach *Keynes* bestehen diesbezüglich folgende empirischen Gesetze für $C(Y)$, $S(Y)$ und die Durchschnittsfunktion $c(Y) = C(Y)/Y$:

(i) Der Grenzhang zum Konsum $C'(Y)$ ist positiv (d.h. bei höherem Einkommen wird mehr konsumiert);

(ii) die marginale Sparquote $S'(Y) = 1 - C'(Y)$ ist ebenfalls positiv (d.h. bei höherem Einkommen wird auch mehr gespart);

(iii) die marginale durchschnittliche Konsumquote $c'(Y)$ ist negativ (mit wachsendem Einkommen nimmt der konsumierte Prozentsatz $c(Y) \cdot 100\%$ des Einkommens ab).

Die beiden ersten Bedingungen zusammen besagen einfach $0 < C'(Y) < 1$. Wegen $\frac{d}{dY}c(Y) = \frac{1}{Y}(C'(Y) - c(Y))$ kann die letzte Bedingung auch dadurch beschrieben werden, dass *die durchschnittliche Konsumquote größer ist als die marginale.* (Mit der in 4.4 eingeführten Terminpologie sagt man zu diesem Sachverhalt, dass *die Konsumfunktion unelastisch ist.*) Die drei Bedingungen sind erfüllbar, legen aber die Konsumfunktion $C(Y)$ noch keineswegs völlig fest. Es gibt viele Ansätze, welche alle drei Bedingungen erfüllen, z.B. $a(1 - \mathrm{e}^{-bY})$ oder $cY/(Y + d)$ mit $a, b, c, d \in \mathbb{R}_{>0}$ und $ab \leq 1$ bzw. $c/d \leq 1$. (Beim ersten Ansatz ist $C'(Y) = ab\mathrm{e}^{-bY} \in]0, 1[$ wegen $bY > 0$ und $0 < ab \leq 1$, und $c(Y) > C'(Y)$ ist äquivalent mit $1 - \mathrm{e}^{-bY} > bY\mathrm{e}^{-bY}$ bzw. $\mathrm{e}^{-bY} < 1/(1 + bY)$, was aus der fundamentalen Ungleichung $\mathrm{e}^{bY} > (1 + bY)$ für die Exponentialfunktion folgt; beim zweiten Ansatz ist $C'(Y) = cd/(Y + d)^2 \in]0, 1[$ und $c(Y) = d/(Y + d) > C'(Y)$ wegen $0 < c/d \leq 1$ und $Y > 0$.) Eine weitergehende Spezifikation der Konsumfunktion in einem mathematischen Modell, das eine reale ökonomische Situation abbilden soll, erfordert zusätzliche Annahmen oder ökonomische Überlegungen oder den Abgleich mit empirischen Daten.

12) Ist $h(y)$ eine differenzierbare Funktion auf einem Intervall J und $f(x)$ eine differenzoierbare Funktion auf einem Intervall I mit Werten in J, so dass die Verkettung $h(f(x))$ auf I definiert ist, so gilt die

- **Kettenregel:** $\boxed{\frac{d}{dx}h(f(x)) = \frac{dh}{dy}(f(x)) \cdot \frac{df}{dx}(x)}$ für $x \in I$,

  kurz: $\boxed{(h \circ f)' = (h' \circ f) \cdot f'}$ auf I;

  die Verkettung von zwei differenzierbaren Funktionen ist differenzierbar, und ihre Ableitung ist das Produkt aus der Ableitung der äußeren Funktion verkettet mit der inneren Funktion und aus der Ableitung der inneren Funktion.

Zum Beweis schreibt man für $y_0 = f(x_0)$ und $y = f(x)$ die Zuwachsformeln $f(x) - f(x_0) = f'(x_0)(x - x_0) + (x - x_0)R(x - x_0)$ und $h(y) - h(y_0) = h'(y_0)(y - y_0) + (y - y_0)T(y - y_0)$ auf und folgert

$$\begin{aligned} h(f(x)) - h(f(x_0)) &= h'(y_0)f'(x_0)(x - x_0) \\ &+ (x - x_0)\Big[h'(y_0)R(x - x_0) + f'(x_0)T(y - y_0) + R(x - x_0)T(y - y_0)\Big] . \end{aligned}$$

Hier sind nun $|R(x - x_0)|$ und $y - y_0 = f(x) - f(x_0)$, also auch $T(y - y_0)$, beliebig klein für $x \in I$ nahe genug bei x_0, d.h. der Betrag des Faktors $[\dots]$ bei $(x - x_0)$ ist kleiner als jede gegebene Fehlerschranke $\varepsilon > 0$, wenn nur x nahe genug bei x_0 ist. Das bedeutet aber gerade, dass $h(f(x))$ an der Stelle x_0 differenzierbar ist mit Ableitung $h'(y_0)f'(x_0) = h'(f(x_0))f'(x_0)$. Da x_0 hierbei eine beliebige Stelle in I ist, sind Differenzierbarkeit von $h \circ f$ auf I und die Kettenregel bewiesen.

Mit $y = f(x)$ und $z = h(y)$ wird die Kettenregel oft symbolisch so geschrieben:

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} ,$$

aber das ist sehr verkürzt und gibt zu Missverständnissen Anlass. Erstens ist hier auf der linken Seite mit “z” die mittelbare Funktion $h(f(x))$ gemeint, auf der rechten Seite aber die äußere Funktion $h(y)$; zweitens hat man beim Auswerten der Formel links und in den Faktor $\frac{dy}{dx}$ rechts dieselbe Stelle $x_0 \in I$ einzusetzen, in $\frac{dz}{dy}$ aber die zugehörige Stelle $y_0 = f(x_0) \in J$. All dies kommt in der symbolischen Notation gar nicht zum Ausdruck, ist aber im Prinzip klar, weil die Gleichung sonst keinen Sinn ergeben würde. Die Ableitung $(h\circ f)'(x_0)$ hängt ja nur ab von den Werten $(h\circ f)(x)$ mit x beliebig nahe bei x_0, also (da f stetig ist) nur von den Werten von h an Stellen y beliebig nahe bei $y_0 = f(x_0)$; daher kann auf der rechten Seite der Kettenregel auch nur die Ableitung von h an dieser Stelle $y_0 = f(x_0)$ auftreten — das Einsetzen einer anderen Stelle wie etwa x_0 wäre sinnlos und unter Umständen auch unmöglich, weil h an der Stelle x_0 nicht definiert zu sein braucht. Die Kettenregel bezieht sich, anders als Summen-, Differenz-, Produkt- und Quotientenregel, auf zwei Funktionen mit im Allgemeinen *verschiedenen* Definitionsbereichen.

Wenn man die Ableitung von $h \circ f$ nochmals differenzieren will, so muss man beachten, dass die rechte Seite der Kettenregel ein *Produkt* ist, also mit der Produktregel differenziert werden muss. Der erste Faktor $h' \circ f$ dieses Produkts ist wieder eine Verkettung und daher mit der Kettenregel abzuleiten. Es folgt also, wenn f und h zweimal differenzierbar sind:

$$\begin{aligned} (h \circ f)'' &= [(h' \circ f) \cdot f']' = (h' \circ f)' \cdot f' + (h' \circ f) \cdot (f')' \\ &= ((h'' \circ f) \cdot f') \cdot f' + (h' \circ f) \cdot f'' = (h'' \circ f) \cdot (f')^2 + (h' \circ f) \cdot f'' . \end{aligned}$$

Man kann auch Kettenregeln noch höherer Ordnung aufstellen — hier aber nicht.

13) *Beispiele für die Anwendung der Kettenregel*:

Ableitung allgemeiner Potenzen einer Funktion:

$$\frac{d}{dx} f(x)^s = \left(\frac{d}{dy} y^s \right) \Big|_{y=f(x)} \cdot f'(x) = \left(s y^{s-1} \right) \big|_{y=f(x)} \cdot f'(x) = s f(x)^{s-1} \cdot f'(x) .$$

Hier haben wir als äußere Funktion die Potenzfunktion $h(y) = y^s$ genommen, als innere die Funktion $f(x)$. Die Rechnung gilt bei positiven differenzierbaren Funktionen f für alle reellen Exponenten s und bei natürlichen Exponenten $s = n \in \mathbb{N}$ für alle differenzierbaren Funktionen $f(x)$ (bei ganzzahligen Exponenten für alle f ohne Nullstelle). Die früheren Formeln für die Ableitungen von Potenzen einer Funktion, die wir im Fall von ganzen Exponenten aus Produkt- und Quotientenregel gefolgert hatten, sind in der jetzigen allgemeineren Regel enthalten, aber noch sehr viel mehr, z.B.

$$\frac{d}{dx}(x^2+1)^{5/2} = \frac{5}{2}(x^2+1)^{3/2} \frac{d}{dx}(x^2+1) = \frac{5}{2}(x^2+1)^{3/2} \cdot 2x ,$$

und die **Ableitung von Wurzeln aus einer Funktion**, etwa

$$\frac{d}{dy} \sqrt[3]{1+y^4} = \frac{d}{dy}(1+y^4)^{1/3} = \frac{1}{3}(1+y^4)^{-2/3} \cdot 4y^3 = \frac{4y^3}{3\sqrt[3]{1+y^4}^{\,2}} ,$$

$$\frac{d^2}{dt^2} \sqrt{\ln t} = \frac{d}{dt} \left(\frac{1}{2\sqrt{\ln t}} \cdot \frac{1}{t} \right) = \frac{-1}{4\sqrt{\ln t}^{\,3}} \cdot \frac{1}{t} \cdot \frac{1}{t} + \frac{1}{2\sqrt{\ln t}} \cdot \frac{-1}{t^2} = -\frac{1+\ln t}{4\sqrt{\ln t}^{\,3}} \qquad (t > 1) .$$

Die **Ableitung von Exponentialen einer Funktion** ergibt sich mit der äußeren Funktion e^y aus der Kettenregel so (wenn $f(x)$ differenzierbar ist):

$$\frac{d}{dx}\, e^{f(x)} == \left(\frac{d}{dy} e^y\right)\Big|_{y=f(x)} \cdot f'(x) = (e^y)\,\big|_{y=f(x)} \cdot f'(x) = e^{f(x)} \cdot f'(x)\ ,$$

und für allgemeine Basen $a > 0$

$$\frac{d}{dx}\, a^{f(x)} = \frac{d}{dx}\, e^{f(x)\ln a} = e^{f(x)\ln a} \cdot f'(x) \cdot \ln a = a^{f(x)} \cdot f'(x) \cdot \ln a\ .$$

Zum Beispiel ist die Ableitung der in der (Wirtschafts–)Statistik wichtigen sogenannten *Normalverteilung* $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, deren Graph als *Gaußsche Glockenkurve* bekannt ist, gegeben durch

$$\frac{d}{dx}\, \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \cdot \frac{-2x}{2} = \frac{-x}{\sqrt{2\pi}}\, e^{-x^2/2}\ ,$$

und die Funktion x^x hat für $x > 0$ die Ableitung

$$\frac{d}{dx}\, x^x = \frac{d}{dx}\, e^{x\ln x} = e^{x\ln x} \cdot \frac{d}{dx}(x\ln x) = x^x\Big(1 \cdot \ln x + x \cdot \frac{1}{x}\Big) = (1 + \ln x)x^x\ .$$

In ökonomischem Zusammenhang stößt man auf Exponentiale von Funktionen z.B. bei *Konsumfunktionen* der Form $C(Y) = a e^{-c/Y}$, wobei a, c positive Parameter sind. Ihre Ableitung, die *marginale Konsumquote* ist dann

$$\frac{d}{dY} C(Y) = \frac{d}{dY}(a^{-c/Y}) = a e^{-c/Y} \cdot (-c) \cdot \frac{-1}{Y^2} = \frac{ac}{Y^2} e^{-c/Y}\ .$$

Bei $Y \to 0$ in $\mathbb{R}_{>0}$ streben $C(Y)$ und $C'(Y)$ gegen Null, für kleine $Y > 0$ ist der exponentielle Faktor $e^{-c/Y}$ extrem klein. Bei diesem Ansatz ist also der Konsum bis zu einem gewissen Schwellenwert des Einkommens Y praktisch Null. Das erscheint nicht realistisch, man würde eher annehmen, dass sich die Sparquote $S(Y) = Y - C(Y)$ so verhält, weil kleine Einkommen vollständig konsumiert werden. Bei $Y \to \infty$ strebt andererseits $C(Y)$ hier gegen den endlichen Grenzwert a und $C'(Y)$ gegen Null. Dies ist ein in gewissen Situationen denkbares, wenn auch kein zwingendes, Verhalten für eine Konsumfunktion; es bedeutet, dass der Konsum eine bestimmte Schranke nie übersteigt. Quintessenz: Der Ansatz scheint zur Modellierung von Konsumverhalten wenig geeignet.

Für die **Ableitung des Logarithmus einer Funktion**, die natürlich positiv sein muss (und differenzierbar), liefert die Kettenregel:

$$\frac{d}{dx}\, \ln f(x) = \left(\frac{d}{dy} \ln y\right)\Big|_{y=f(x)} \cdot f'(x) = \left(\frac{1}{y}\right)\Big|_{y=f(x)} \cdot f'(x) = \frac{f'(x)}{f(x)}\ ;$$

der Quotient f'/f heißt deshalb auch die **logarithmische Ableitung** der Funktion f. Zum Beispiel ist

$$\frac{d}{dx}\, \ln(1 + x^2) = \frac{2x}{1 + x^2}\ ,$$

und

$$\frac{d}{dt}\, \ln(\ln t) = \frac{1/t}{\ln t} = \frac{1}{t\ln t} \qquad (\text{für } \ t > 1\,,\ \text{also } \ \ln t > 0\,)\,.$$

Die logarithmisch Ableitung ist nützlich für die Differentiation von Produkten von Funktionen ohne Nullstellen; denn *die logarithmische Ableitung eines Produkts ist die Summe der logarithmischen Ableitungen seiner Faktoren*:

$$\frac{(f_1 \cdot f_2 \cdot \ldots \cdot f_n)'}{f_1 \cdot f_2 \cdot \ldots \cdot f_n} = \frac{f_1'}{f_1} + \frac{f_2'}{f_2} + \ldots + \frac{f_n'}{f_n}\ .$$

Noch ein Beispiel für die Differentiation mehrfacher Verkettungen:

$$\frac{d}{dt}\,\mathrm{e}^{\sqrt{1+t^2}} = \mathrm{e}^{\sqrt{1+t^2}}\frac{d}{dt}\sqrt{1+t^2} = \mathrm{e}^{\sqrt{1+t^2}}\frac{1}{2\sqrt{1+t^2}}\,\frac{d}{dt}(1+t^2) = \frac{t}{\sqrt{1+t^2}}\,\mathrm{e}^{\sqrt{1+t^2}}\ .$$

Hier haben wir die Verkettung $h \circ g \circ f$ von drei Funktionen $f(t) = 1 + t^2$, $g(x) = \sqrt{x}$ und $h(y) = \mathrm{e}^y$ abgeleitet, indem wir $h \circ (g \circ f)$ geklammert haben, also zuerst h als äußere Funktion und $g \circ f$ als innere Funktion aufgefasst haben, und dann die Ableitung dieser inneren Funktion $g \circ f$ wieder mit der Kettenregel berechnet haben. Man hätte auch, aber das ist meist weniger günstig für die Berechnung der Ableitung, $(h \circ g) \circ f$ klammern können; denn *die Verkettung von Funktionen ist assoziativ,* d.h. unabhängig von der Klammerung. (Der Wert $h(g(f(t)))$ hängt nicht davon ab, ob man f und g "in einem Schritt ausführt" oder g und h.) Das Ergebnis für die Ableitung der Verkettung (wenn definiert) von drei (differenzierbaren) Funktionen ist in jedem Fall

$$(h \circ g \circ f)' = (h' \circ g \circ f) \cdot (g' \circ f) \cdot f'\ ,$$

und es sollte klar sein wie die entsprechenden Regeln für die Ableitung einer Verkettung von mehr als drei Funktionen aussehen werden.

Das letzte Beispiel ist die Ableitung von Funktionen des Typs $g(x)^{f(x)}$:

$$\begin{aligned} &\frac{d}{dx}g(x)^{f(x)} = \frac{d}{dx}\,\mathrm{e}^{f(x)\ln g(x)} = \mathrm{e}^{f(x)\ln g(x)}\frac{d}{dx}[f(x)\ln g(x)] \\ = \;& g(x)^{f(x)}\Big[f'(x)\ln g(x) + f(x)\frac{d}{dx}\ln g(x)\Big] = g(x)^{f(x)}\Big[f'(x)\ln g(x) + f(x)\frac{g'(x)}{g(x)}\Big]\ . \end{aligned}$$

Hier haben wir der Reihe nach die Kettenregel für Exponentiale von Funktionen, die Produktregel und die Kettenregel für den Logarithmus einer Funktion angewendet. Voraussetzung ist natürlich, dass f und g differenzierbar sind auf dem betrachteten Intervall in $\mathbb{R}$ und dass g positiv ist. Unter den hier differenzierten Funktionentyp fallen zum Beispiel x^x, $x^{1/x}$, $(1+\frac{1}{x})^x$, usw.

14) Die letzte zu besprechende Differentiationsregel bezieht sich auf die Berechnung der **Ableitung einer Umkehrfunktion** zu einer gegebenen Funktion mit bekannter Ableitung. Dazu betrachten wir eine stetige streng monotone Funktion f auf einem Intervall I (positiver Länge) in $\mathbb{R}$. Aus 4.2 wissen wir schon, dass die Wertemenge von f dann ein Intervall J ist und dass f eine stetige und gleichsinnig streng monotone Umkehrfunktion $g\colon J \to I \subset \mathbb{R}$ besitzt, also $g(f(x)) = x$ für $x \in I$ (und $f(g(y)) = y$ für $y \in J$ wie man durch Einsetzen von $f(x)$ für y sieht). Ist f differenzierbar, so könnte man denken, dass auch g differenzierbar sein sollte — aber das ist nicht so, im Allgemeinen. Wenn nämlich auch g differenzierbar ist, so liefert Differenzieren der Gleichung $g(f(x)) = x$ mit der Kettenregel $g'(f(x)) \cdot f'(x) = 1$. Daraus sehen wir, dass nur eine Chance für Differenzierbarkeit von g an der Stelle $y = f(x)$ besteht, wenn $f'(x) \neq 0$ ist, und dann muss $g'(y) = 1/f'(x)$ sein. Die streng wachsende Funktion $f(x) = x^3$ auf $\mathbb{R}$ mit $f'(0) = 0$ hat also zwar eine stetige Umkehrfunktion (nämlich $g(x) = (\operatorname{sign} x)\sqrt[3]{|x|}$), aber diese kann an der Stelle $0 = f(0)$ nicht differenzierbar sein (sie hat an dieser Stelle unendliche Steigung in dem Sinne, dass die Differenzenquotienten gegen ∞ streben; das ist auch geometrisch klar: Ein horizontale Tangente zum Graphen von f entspricht einer vertikalen Tangente zu dem an der Diagonale gespiegelten Graphen, also zum Graphen der Umkehrfunktion). An den Stellen x aber, für die $f'(x) \neq 0$ ist, gilt für die Ableitung die

- **Umkehrfunktionsregel:** $\boxed{g'(y) = \dfrac{1}{f'(x)} \quad \text{für } y = f(x)\,,}$ (wenn $f'(x) \neq 0$);

  die Ableitung der Umkehrfunktion zu einer streng monotonen differenzierbaren Funktion f an einer Stelle im Wertebereich von f existiert und ist das Reziproke der Ableitung von f an der entsprechenden Stelle im Definitionsbereich, wenn dort nicht die Ableitung von f verschwindet.

Zum Beweis schreiben wir für $y = f(x)$ und $y_0 = f(x_0)$, also auch $x = g(y)$ und $x_0 = g(y_0)$, die Differenzenquotienten der Umkehrfunktion auf:

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \left[\frac{f(x) - f(x_0)}{x - x_0}\right]^{-1}.$$

Da mit der streng monotonen Funktion f auch die Umkehrfunktion g stetig ist (siehe 4.2), liegt x beliebig nahe bei x_0, wenn y genügend nahe bei y_0 ist, und daher liegen dann auch die Differenzequotienten von f beliebig nahe bei der Ableitung $f'(x_0)$. Wenn nun $f'(x_0) \neq 0$ ist, so liegen die Reziproken dieser Differenzenquotienten entsprechend beliebig nahe bei $1/f'(x_0)$ (das garantiert die Stetigkeit der Reziprokenbildung $z \mapsto 1/z$ auf $\mathbb{R}_{\neq 0}$), und mit der obigen Gleichung folgt, dass die Differenzenquotienten von g bei $y \to y_0$ den Limes $1/f'(x_0)$ haben, dass also $g'(y_0) = 1/f'(x_0)$ ist.

Die Differentiationsregel für Umkehrfunktionen wird mit $y = f(x)$ und $x = g(y)$ oft symbolisch

$$\frac{dx}{dy} = \frac{1}{dy/dx}$$

geschrieben, aber diese Gleichung ist nur korrekt, wenn man sie so versteht, dass beim Auswerten jeweils links eine Stelle $y = f(x) \in J$ und rechts die zugehörige Stelle $x = g(y)$ in I zu nehmen ist. (Alles andere wäre sinnlos; z.B. kann man $x \in I$ links im Allgemeinen schon deswegen nicht einsetzen, weil x gar nicht im Definitionsbereich J der Umkehrfunktion liegt.) Die rechte Seite der Gleichung ist also eigentlich als Verkettung der Ableitung df/dx mit der Umkehrfunktion $g(y)$ zu lesen oder die linke als Verkettung der Ableitung dg/dy mit $f(x)$, d.h. wenn f' keine Nullstelle auf I hat, so gilt:

$$\boxed{g' = \frac{1}{f'} \circ g} \quad \text{bzw.} \quad \boxed{g' \circ f = \frac{1}{f'}}\ ;$$

wenn die Ableitung einer streng monotonen Funktion keine Nullstelle hat, so ist die Ableitung der Umkehrfunktion gegeben durch Verkettung des Reziproken der Ableitung der Ausgangsfunktion mit der Umkehrfunktion.

Da die Ableitung einer wachsenden Funktion nichtnegativ ist (die Differenzenquotienten sind ja alle nichtnegativ) und die einer fallenden Funktion nichtpositiv, bedeutet die Voraussetzung, dass die Ableitung überall auf dem Definitionsintervall von Null verschieden ist, im Falle einer monotonen Funktion natürlich, dass sie überall positiv oder überall negativ sein muss. (Das gilt übrigens sogar ohne Voraussetzung der Monotonie: Man kann zeigen, dass eine differenzierbare Funktion ohne Ableitungsnullstelle monoton sein muss.)

Im Allgemeinen ist die Umkehrfunktion g zu einer umkehrbaren elementaren Funktion f nicht wieder elementar, d.h. man hat keine “Formel” für die Auswertung von g. Dann ist auch die Gleichung $g' = (1/f') \circ g$ keine, die eine Formel-Darstellung von $g'(y)$ als elementare Funktion von y liefert. Wenn man jedoch die Ableitung $f'(x)$ durch $f(x)$ ausdrücken kann, d.h. wenn $f'(x) = h(f(x))$ ist mit einer elementaren Funktion h ohne Nullstelle, so erhält man eine **elementare Darstellung der Ableitung der Umkehrfunktion**

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{h(f(x))} = \frac{1}{h(y)} \qquad (\, y = f(x)\,, \quad f'(x) = h(f(x)) \neq 0\,)$$

durch das Reziproke der elementaren Funktion h. Es ist durchaus möglich, dass dann zwar die Ableitung der Umkehrfunktion elementar ist, also durch einen Rechenterm definierbar, die Umkehrfunktion selbst aber nicht.

In ökonomischem Kontext werden häufig Funktion und Umkehrfunktion nicht festgelegt, wenn man zwei Variablen x und y hat, die sich gegenseitig bestimmen. Man kann dann $y = f(x)$ als Funktion von x auffassen oder $x = g(y)$ als Funktion von y. Ein Beispiel für diese Situation ist die *Preis–Nachfrage–Funktion* $x(p)$ und die *Nachfrage–Preis–Funktion* $p(x)$, siehe 4.2. In einer derartigen Situation mit "gleichberechtigten Variablen" wird die Ableitungsregel für die Umkehrfunktion $\frac{dx}{dy}(y) = 1/\frac{dy}{dx}(x)$ bzw. $\frac{dy}{dx}(x) = 1/\frac{dx}{dy}(y)$ für $y = f(x)$ und $x = g(y)$ oft verwendet, wenn die wechselseitige Abhängigkeit von x und y aus ökonomischen Gründen monoton ist, wobei man sich über die Voraussetzung, dass die Ableitungen im Nenner nicht Null sein dürfen, keine Gedanken macht. Das ist gerechtfertigt, weil man die Funktionen f, g ohnehin nicht genau kennt, also kann man auch gleich annehmen, dass sie streng monoton sind und überall positive oder überall negative Ableitung haben. Ist $f(x)$ nur schwach monoton, so lässt sich dies nämlich erreichen, indem man zu $f(x)$ eine beliebig kleine "Störung" $\pm\varepsilon x$ addiert mit $0 < \varepsilon \ll 1$.

15) *Beispiele für Berechnung der Ableitung der Umkehrfunktion*:

Die Exponentialfunktion $f(x) := \mathrm{e}^x$ ist wachsend auf $\mathbb{R}$ mit Ableitung $f'(x) = \mathrm{e}^x > 0$, also ist die Umkehrfunktion ln auf ihrem Definitionsbereich $\mathbb{R}_{>0}$ überall differenzierbar. Hier kann man die Ableitung durch die abgeleitete Funktion ausdrücken, nämlich $f'(x) = h(f(x))$ mit $h(y) = y$. Also gilt

$$\frac{d}{dy} \ln y = \frac{1}{\mathrm{e}^x} = \frac{1}{y} \qquad (\, y = \mathrm{e}^x\,)\,.$$

Diese Rechnung haben wir schon vorgeführt; denn bei der früheren Berechnung der Ableitung der Grundfunktion $\ln y$ haben wir genau so argumentiert wie bei der Berechnung von Umkehrfunktionsableitungen allgemein. Ausgehend von $f(x) = \ln x$ mit $f'(x) = 1/x = \mathrm{e}^{-f(x)}$ würde man übrigens für die Umkehrfunktion $g(y) = \mathrm{e}^y$ ganz analog finden $g'(y) = 1/\mathrm{e}^{-y} = \mathrm{e}^y$.

Die Funktion $f(x) = \sinh x = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x})$ ist streng wachsend auf $\mathbb{R}$ mit Ableitung $\sinh' x = \cosh x = \sqrt{1 + \sinh^2 x} \geq 1$ überall auf $\mathbb{R}$ und mit $\mathbb{R}$ als Wertebereich. Die Umkehrfunktion $g : \mathbb{R} \to \mathbb{R}$ heißt *Areasinus (hyperbolicus)*, weil sie mit dem Flächeninhalt von Hyperbelsektoren zusammenhängt, und wird Arsinh oder $\sinh^{-1}$ notiert (was nicht mit $1/\sinh$ verwechselt werden darf). Die Umkehrfunktionsregel gibt hier:

$$\frac{d}{dy} \operatorname{Arsinh} y = \frac{1}{\cosh x} = \frac{1}{\sqrt{1 + y^2}} \qquad (\, y = \sinh x\,)\,.$$

Auch die hyperbolische Tangensfunktion $\tanh x = (\mathrm{e}^x - \mathrm{e}^{-x})/(\mathrm{e}^x + \mathrm{e}^{-x})$ (eine typische logistische Funktion) ist streng wachsend auf $\mathbb{R}$, und die Ableitung $\tanh' x = 1/(\cosh x)^2 = 1 - (\tanh x)^2 > 0$ haben wir schon bei den Beispielen zur Quotientenregel berechnet. Der Wertebereich von tanh ist das Intervall $]-1, 1[$ und die darauf definierte Umkehrfunktion, der sog. *Areatangens (hyperbolicus)* $\operatorname{Artanh} = \tanh^{-1}$, ist differenzierbar mit Ableitung

$$\frac{d}{dy} \operatorname{Artanh} y = \frac{1}{(\cosh x)^2} = \frac{1}{1 - y^2} \qquad (\, y = \tanh x\,)\,.$$

Bei den Areafunktionen sind aber nicht nur die Ableitungen elementare Funktionen, sondern die Funktionen selbst auch. Die Gleichung $y = \sinh x = \frac{1}{2}\mathrm{e}^x - \frac{1}{2}\mathrm{e}^{-x}$ ist nämlich eine quadratische Gleichung $(\mathrm{e}^x)^2 - 2y\mathrm{e}^x - 1 = 0$ für e^x mit der einzigen Lösung $\mathrm{e}^x = y + \sqrt{1+y^2}$ (die andere Lösung der quadratischen Gleichung ist negativ und daher nicht von der Form e^x). Ebenso ist die Gleichung $y = \tanh x = (\mathrm{e}^x - \mathrm{e}^{-x})/(\mathrm{e}^x + \mathrm{e}^{-x})$ eine quadratische Gleichung $(1-y)(\mathrm{e}^x)^2 = 1+y$ für e^x mit der einzigen Lösung $\mathrm{e}^x = \sqrt{(1+y)/(1-y)}$. Logarithmieren der gefundenen Ausdrücke für e^x gibt daher die Formeln

$$\operatorname{Arsinh} y = \ln\left(y + \sqrt{1+y^2}\right), \quad \operatorname{Artanh} y = \ln\sqrt{\frac{1+y}{1-y}} = \frac{1}{2}\ln\frac{1+y}{1-y} \quad (-1 < y < 1),$$

mit denen man diese beiden Areafunktionen durch die elementaren Grundfunktionen dargestellt hat. Durch Differentiation dieser Funktionen mit der Ketten- und Quotientenregel erhält man die obigen Formeln für ihre Ableitungen natürlich auch, aber viel mühsamer.

Ganz analog, nur mit einem Vorzeichenunterschied, berechnet man die Ableitung der Umkehrfunktion zur Sinusfunktion auf dem Intervall $]-\frac{1}{2}\pi, \frac{1}{2}\pi[$, auf dem der Sinus streng wachsend ist mit Ableitung $\sin' x = \cos x = \sqrt{1-\sin^2 x}$. Die Umkehrfunktion ist auf $]-1,1[$ definiert und heißt *Arcussinus*, weil sie mit der Länge von Bögen auf Kreislinien zusammenhängt; sie wird notiert $\arcsin$ oder $\sin^{-1}$. Dafür gilt dann:

$$\frac{d}{dy}\arcsin y = \frac{1}{\cos x} = \frac{1}{\sqrt{1-y^2}} \qquad (-1 < y = \sin x < 1, -\frac{1}{2}\pi < x < \frac{1}{2}\pi).$$

Und die Tangensfunktion $\tan x = \frac{\sin x}{\cos x}$ ist streng wachsend auf $]-\frac{1}{2}\pi, \frac{1}{2}\pi[$ mit Ableitung $\tan' x = 1/(\cos x)^2 = 1 + (\tan x)^2$ und mit Wertebereich $\mathbb{R}$, so dass sich für ihre Umkehrfunktion, den *Arcustangens* $\arctan = \tan^{-1}$, die Ableitung

$$\frac{d}{dy}\arctan y = \frac{1}{(\cos x)^2} = \frac{1}{1+y^2} \qquad (y = \tan x \in \mathbb{R}, \quad -\frac{1}{2}\pi < x < \frac{1}{2}\pi)$$

ergibt. Anders als bei den Areafunktionen gibt es hier aber keine Darstellung der Umkehrfunktion $\arctan$, $\arcsin$ selbst durch andere elementare Grundfunktionen (obwohl die Ableitung des Arcustangens eine ganz einfache Funktion ist). Deshalb muss man in der Mathematik diese Funktionen zu den elementaren Grundfunktionen hinzunehmen, weil man sie insbesondere in der Integralrechnung braucht (damit eine so einfache Funktion wie $1/(1+y^2)$ auch eine elementare Stammfunktion hat, d.h. eine Funktion, deren Ableitung sie ist). Eine der Arcusfunktionen genügt dabei, da man die andere durch sie ausdrücken kann; z.B. ist $\arcsin x = \arctan(x/\sqrt{1-x^2})$ für $|x| < 1$. ■

Wir haben nun gesehen, dass die elementaren Grundfunktionen, also Konstanten und die Funktionen x, e^x, $\ln x$ (in der Mathematik zusätzlich noch $\sin x$ und $\arcsin x$), differenzierbar sind auf ihren maximalen offenen Definitionsintervallen und dass sie elementare Funktionen als Ableitungen haben. Aus den Rechenregeln für die Ableitung sehen wir weiter, dass auch jede Funktion, die aus zwei Grundfunktionen zusammengesetzt ist mittels der Operationen Addition, Subtraktion, Multiplikation, Division (auf Intervallen ohne Nennernullstellen natürlich) und Verkettung, wieder eine differenzierbare Funktion ist, deren Ableitung ebenfalls elementar ist, also mit einer Formel, welche nur diese Operationen enthält, durch die Grundfunktionen ausgedrückt werden kann. Und wenn wir nun aus solchen zusammengesetzten Funktionen der “zweiten Generation” mit den algebraischen

Operationen und mit Verkettung kompliziertere elementare Funktionen der “dritten Generation” zusammensetzen, so sind diese aufgrund der Ableitungsrechenregeln offenbar wieder differenzierbar mit elementaren Ableitungen. So fortfahrend sehen wir, dass überhaupt jede elementare Funktion, also jede durch einen Rechenterm definierbare Funktion, der aus endlich vielen Rechenoperationen und Verkettungen mit den Grundfunktionen zusammengesetzt ist, differenziert werden kann und dass das Ergebnis des Ableitens wieder eine elementare Funktion ist. Diese Ableitung kann man deshalb erneut differenzieren und erhält wieder eine elementare Funktion u.s.w. Es gilt daher der

SATZ: *Jede auf einem offenen Intervall in $\mathbb{R}$ definierte elementare Funktion f ist dort beliebig oft differenzierbar, und ihre Ableitungen $f^{(n)}$ jeder Ordnung sind wieder elementare Funktionen, die explizit berechnet werden können. Das heißt, wenn $f(x)$ durch einen Term gegeben ist, der aus den Grundfunktionen mit endlich vielen Rechenoperationen und Verkettungen aufgebaut ist, so erhält man durch Anwendung der Ableitungsrechenregeln und der Formeln für die Ableitung der Grundfunktionen auch einen entsprechenden Rechenterm für die Ableitung f' und für alle Ableitungen höherer Ordnung $f^{(n)}$.* ■

Das bedeutet also, dass man für jede Funktion, die durch einen Rechenterm gegeben ist, auch einen Rechenterm für ihre Ableitung angeben kann, im Prinzip jedenfalls; denn die Terme für die Ableitung können, namentlich bei mehrfacher Anwendung der Quotientenregel oder der Kettenregel, erheblich größer und komplizierter werden als der Term für die Ausgangsfunktion. Das größte Problem beim Ableiten von komplexen Termen ist, ihren Aufbau genau zu verstehen, so dass man Summenregel, Produktregel, Quotientenregel und Kettenregel in der richtigen Reihenfolge anwendet. Da formales Differenzieren aber eine ganz schematische Angelegenheit ist, gibt es dafür auch Rechenprogramme, die symbolisch differenzieren können, d.h. zur Eingabe eines Rechenterms für eine Funktion als Ausgabe einen Rechenterm für die Ableitung dieser Funktion liefern. (Allerdings differenzieren diese Programme auch Funktionen, die es gar nicht gibt. Zum Beispiel liefern sie für $\ln(1-\mathrm{e}^{x^2})$ die Ableitung $-2x\mathrm{e}^{x^2}/(1-\mathrm{e}^{x^2})$, die für $x \neq 0$ definiert ist, obwohl der Ausgangsterm wegen $1-\mathrm{e}^{x^2} \leq 0$ für kein $x \in \mathbb{R}$ einen Funktionswert liefert.)

Da die in der Wirtschaftsmathematik auftretenden konkreten Funktionen allesamt elementar sind oder wenigstens stückweise aus elementaren Funktionen zusammengesetzt sind, die sogar durch recht einfache Terme definiert werden, können wir abschließend sagen:

- *Jede konkrete Funktion in mathematischen Modellen ökonomischer Vorgänge kann beliebig oft differenziert werden — von einzelnen isolierten Ausnahmestellen abgesehen (Sprungstellen, Unendlichkeitsstellen, Knickstellen, ...) — und ihre Ableitungen können auch konkret ausgerechnet werden.*

4.4 Die Elastizität von Funktionen

Wir besprechen hier eine Variante des Ableitungsbegriffs, die in der Wirtschaftsmathematik besondere Bedeutung hat. Ausgangspunkt ist die Überlegung, dass der reine Zahlenwert $f'(x) \in \mathbb{R}$ der Ableitung einer Funktion $y = f(x)$, welche die Abhängigkeit zwischen zwei ökonomischen Größen beschreibt, keine ökonomische Aussagekraft besitzt, solange nicht die Einheiten spezifiziert sind, mit denen die unabhängige Variable x und die abhängige Variable y gemessen werden. Auch die Zahlenwerte der Variablen selbst haben ohne Angabe von Einheiten ja keinen Sinn: $y = 1000$ bedeutet, wenn die Einheit ein Euro ist, etwas ganz anderes, als wenn sie tausend Euro ist oder ein Dollar oder ein Yen! So verhält es sich auch bei den Ableitungen. Ist z.B. $E(x)$ eine Erlösfunktion, die den Erlös in € bei Absatz von x Mengeneinheiten eines Guts angibt, so sagt $E'(x_0) = 100$, dass der Erlös bei Absatz einer weiteren (kleinen) Mengeneinheit zusätzlich zu x_0 um 100 € zunimmt. Bei Einheiten von 1000 € wird derselbe Sachverhalt aber durch die Funktion $\widetilde{E}(x) = \frac{1}{1000}E(x)$ beschrieben, und der Grenzerlös ist dann $\widetilde{E}'(x_0) = \frac{1}{10}$. Und wenn man zu einer 1000–fach größeren Mengeneinheit übergeht, etwa Tonnen statt Kilogramm, so wird die abgesetzte Menge durch die Variable $\widetilde{x} = \frac{1}{1000}x$ beschrieben, und die Erlösfunktion als Funktion von $\widetilde{x}$ ist nun $E(1000 \cdot \widetilde{x})$ mit der Ableitung $1000 \cdot E'(x_0) = 100\,000$ an der Stelle $\widetilde{x}_0 = \frac{1}{1000}x_0$.

Für die Interpretation und Vergleichbarkeit von Zahlenangaben, die ökonomische Sachverhalte beschreiben, ist es wünschenswert, die Zahlenangaben nach Möglichkeit absolut, d.h. von der Wahl von Einheiten unabhängig, zu machen. Für die Änderungsrate einer (ökonomischen) Funktion kann das erreicht werden, indem man nicht die Änderungen $x - x_0 = h$ und $f(x) - f(x_0) = f(x_0 + h) - f(x_0)$ von Argument und Funktionswert selbst betrachtet, sondern ihre prozentualen Änderungen. Dazu nehmen wir von jetzt ab (bis zum Ende dieses Abschnitts 4.4) an, dass die Funktion f auf $\mathbb{R}_{>0}$ oder einem Teilintervall von $\mathbb{R}_{>0}$ definiert ist und Werte in $\mathbb{R}_{>0}$ hat. Beim Übergang von x_0 und $f(x_0)$ zu $x = x_0 + h$ und $f(x) = f(x_0 + h)$ ist dann die **relative Änderung des Arguments** definiert als das Verhältnis der Änderung von x_0 zur Größe von x_0,

$$\frac{x - x_0}{x_0} = \frac{h}{x_0}$$

und die **relative Änderung des Funktionswertes**

$$\frac{f(x) - f(x_0)}{f(x_0)} = \frac{f(x_0 + h) - f(x_0)}{f(x_0)}$$

entsprechend als das Verhältnis der Änderung des Funktionswertes zu seiner Größe. Die relativen Änderungen sind unabhängig von der Wahl der Einheiten, mit denen die Variablen gemessen werden, weil sich ein Wechsel zu einer anderen Einheit im Zähler und im Nenner herauskürzt. Man sagt auch **prozentuale Änderung** statt relative Änderung; genauer ändert sich x_0 um $100 \cdot \frac{x - x_0}{x_0}$ Prozent, wenn man von x_0 zu x übergeht, und $f(x_0)$ um $100 \cdot \frac{f(x) - f(x_0)}{f(x_0)}$ Prozent.

Statt der Differenzenquotienten, also dem Verhältnis der Änderungen von abhängiger und unabhängiger Variablen, bildet man nun das Verhältnis der relativen Änderungen, den sog. **relativen Differenzenquotienten**

$$\frac{\frac{f(x) - f(x_0)}{f(x_0)}}{\frac{x - x_0}{x_0}} = \frac{x_0}{f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0} = \frac{x_0}{f(x_0)} \cdot \frac{f(x_0 + h) - f(x_0)}{h}.$$

In der Ökonomie heißt dieser Quotient auch **Intervall-Elastizität** ϵ ("Epsilon") der abhängigen Variablen y auf dem Intervall zwischen x_0 und x. Die Bedeutung dieser Zahl ist, dass sich der Funktionswert um $\epsilon{\cdot}p$ Prozent erhöht ($\epsilon{\cdot}p > 0$) bzw. erniedrigt ($\epsilon{\cdot}p < 0$), wenn die abhängige Variable von x_0 nach x läuft und ihren Wert dabei um $p\,\%$ erhöht ($p > 0$) bzw. erniedrigt ($p < 0$).

Um das Änderungsverhalten der Funktion $y = f(x)$ an der Stelle x_0 zu beschreiben, muss man nun noch den Grenzübergang $h = x - x_0 \to 0$ vornehmen. Im ökonomischen Jargon geht man von der Intervall-Elastizität über zur **Punkt–Elastizität** an der Stelle x_0. Da der relative Differenzenquotient oben einfach das Produkt des konstanten Faktors $x_0/f(x_0) > 0$ mit dem (gewöhnlichen) Differenzenquotienten ist, hat der relative Differenzenquotient einen Limes, genau wenn das für den gewöhnlichen Differenzenquotienten gilt, wenn also die Ableitung $f'(x_0)$ existiert, und der Limes des relativen Differenzenquotienten ist dann das Produkt des Faktors $x_0/f(x_0)$ mit $f'(x_0)$. Damit haben wir die gewünschte Größe gefunden, die das Änderungsverhalten von $y = f(x)$ bei einer Stelle x_0 unabhängig von der Wahl von Einheiten beschreibt:

DEFINITION: Ist die Funktion $f : \mathbb{R}_{>0} \supset I \to \mathbb{R}_{>0}$ differenzierbar an der Stelle $x_0 \in I$, so heißt (lies "Epsilon f von x_0")

$$\varepsilon_f(x_0) := \frac{x_0 f'(x_0)}{f(x_0)} \qquad \text{(auch notiert } \frac{Ef}{Ex}(x_0)\,)$$

die **Elastizität** von f **an der Stelle** $\boldsymbol{x_0}$ (oder auch die **Punkt–Elastizität** im Punkt x_0). Ist f überall auf I differenzierbar, so heißt die Funktion

$$\varepsilon_f(x) := \frac{x f'(x)}{f(x)} \qquad (\,x \in I\,)$$

die **Elastizitätsfunktion** zu f auf I oder einfach auch die **Elastizität von** $\boldsymbol{f}$ **auf** $\boldsymbol{I}$. ■

DISKUSSION: **1)** Wir betonen nochmals:

- *Der Begriff der Elastizität ist nur sinnvoll für* positive *Funktionen* $f(x) > 0$ *von einer* positiven *Variablen* $x > 0$.

2) *Die Elastizität ist unabhängig von der Wahl der Einheiten*, mit denen die unabhängige oder die abhängige Variable gemessen wird. Eine Maßstabsänderung bei der abhängigen Variablen $y = f(x)$ bedeutet den Übergang von $f(x)$ zu $cf(x)$ mit einem konstanten Faktor $c > 0$, und dieser Faktor kürzt sich in der Elastizität $xf'(x)/f(x)$ gemäß erster Faktorregel für die Ableitung heraus. Eine Maßstabsänderung bei der unabhängigen Variablen bedeutet den Übergang von $f(x)$ zur Funktion $\widetilde{f}(\widetilde{x}) = f(x)$ mit der neuen Variablen $\widetilde{x} = \frac{1}{a}x$ und einem konstanten Faktor $a > 0$, und dafür ist gemäß zweiter Faktorregel die Elastizität $\varepsilon_{\widetilde{f}}(\widetilde{x})$ gleich der Elastizität $\varepsilon_f(x)$ von f an der Stelle $x = a\widetilde{x}$, die im ursprünglichen Maßstab derselben realen Größe entspricht wie der Wert $\widetilde{x}$ nach der Maßstabänderung.

- *Elastizitäten beschreiben die Änderung ökonomischer Variablen durch "absolute" reelle Zahlen, die nicht von der Festlegung irgendwelcher Einheiten abhängen und daher unmittelbar vergleichbar und ökonomisch interpretierbar sind.*

Das ist kein Wunder, denn wir haben die Elastizität ja gerade definiert, um diese Ziel zu erreichen. Praktisch bedeutet es, dass z.B. ein Unternehmen, das international auf verschiedenen Märkten tätig ist, die Preis–Nachfrage–Funktionen für ein Produkt mit den im jeweiligen Land verwendeten Währungen und Maßeinheiten für den Absatz definieren und die Elastizitäten der Nachfrage-Funktionen doch unmittelbar vergleichen kann: Größere Elastizität bedeutet größere Empfindlichkeit der Nachfrage, also größeren prozentualen Rückgang der Nachfrage bei 1 % Preiserhöhung.

3) *Ökonomische Interpretation der Elastizität*: Bei einer hinreichend kleinen Änderung der unabhängigen Variablen, sagen wir um 1 %, ist der relative Differenzenquotient praktisch schon gleich dem Limes. Der Betrag der Elastizität $|\varepsilon_f(x)|$ gibt daher an, um wieviel Prozent sich der Wert $f(x)$ der abhängigen Variablen ändert, wenn die unabhängige Variable x um 1 % erhöht wird. Dabei bestimmt das Vorzeichen der Elastizität $\varepsilon_f(x)$, ob der Funktionswert bei dieser Änderung zunimmt (nämlich bei $\varepsilon_f(x) > 0$) oder abnimmt (im Fall $\varepsilon_f(x) < 0$); denn wegen $x/f(x) > 0$ ist das Vorzeichen der Elastizität $\varepsilon_f(x)$ ja dasselbe wie das der Ableitung $f'(x)$, und kleine Zuwächse bei x bewirken eine Vergrößerung von $f(x)$, wenn $f'(x) > 0$ ist, bzw. eine Verkleinerung, wenn $f'(x) < 0$. Wenn $\varepsilon_f(x) = 0$ ist, also auch $f'(x) = 0$, so kann man über das Vorzeichen der Änderung des Wertes von $f(x)$ bei Erhöhung von x um 1 % nichts sagen, aber die prozentuale Änderung des Funktionswertes ist in diesem Falle sehr klein, viel kleiner als 1 % (bei "normalen" Funktionen, jedenfalls).

- *Der* **Betrag der Elastizität** *gibt an, um wieviel Prozent sich die abhängige Variable $y = f(x)$ (ungefähr) ändert, wenn man die unabhängige Variable x um 1 % erhöht, d.h. $f(x)$ ändert sich (ungefähr) um $|\varepsilon_f(x)|$ Prozent;*
- *das* **Vorzeichen der Elastizität** *bestimmt dabei die Richtung der Änderung der abhängigen Variablen, d.h. $f(x)$ wird größer, wenn $\varepsilon_f(x) > 0$ ist, und $f(x)$ wird kleiner, wenn $\varepsilon_f(x) < 0$ ist (jeweils bei kleiner Vergrößerung von x).*

4) Der Name "Elastizität" kommt daher, dass man in der Ökonomie eine abhängige Größe "elastisch" nennt, wenn sie auf Variationen der unabhängigen Variablen mit deutlichen Änderungen reagiert, dagegen "starr", wenn sie durch Änderungen der unabhängigen Variablen wenig beeinflusst wird. ("Elastizität" ist vielleicht keine besonders glückliche Wortwahl für das Gemeinte; "Empfindlichkeit" oder "Sensitivität" wäre zum Beispiel besser, ist aber nicht gebräuchlich.) Genauer heißt die abhängige Variable $y = f(x)$

- **elastisch** in dem Bereich, wo $|\varepsilon_f(x)| > 1$ ist (kleine relative Änderungen von x bewirken dort größere relative Änderungen von $f(x)$);
- **unelastisch** in dem Bereich, wo $|\varepsilon_f(x)| < 1$ ist (kleine relative Änderungen von x bewirken dort kleinere relative Änderungen von $f(x)$);
- **ausgeglichen elastisch** (oder **proportional elastisch**) an den Stellen x mit $|\varepsilon_f(x)| = 1$ (kleine relative Änderungen von x bewirken dort gleich große relative Änderungen von $f(x)$);
- **vollkommen unelastisch** (oder **starr**) an den Stellen x mit $\varepsilon_f(x) = 0$ (kleine Änderungen von x bewirken überhaupt keine wahrnehmbaren Änderungen des Funktionswertes $f(x)$);
- **vollkommen elastisch** an Unendlichkeitsstellen der Elastizitätsfunktion (geringste Änderungen von x bewirken dort große relative Änderungen von $f(x)$).

Zum Beispiel wird eine Nachfragefunktion $x(p)$ elastisch bezüglich des Preises sein ($|\varepsilon_x(p)| > 1$ oder genauer, da die Nachfrage mit zunehmendem Preis abnimmt, $\varepsilon_x(p) < -1$), wenn es sich um ein leicht entbehrliches Gut handelt; kleinere Preisänderungen haben dann deutliche Nachfrageänderungen zur Folge. Bei einem schwer zu entbehrenden Gut aber, das nicht ohne Weiteres zu substituieren ist, wird die Nachfragefunktion unelastisch sein (genauer $-1 < \varepsilon_x(p) < 0$). Im allgemeinen hat eine ökonomische Funktion verschiedenes Elastizitätsverhalten auf verschiedenen Teilintervallen ihres Definitionsbereiches; die Stellen des Übergangs von elastischem zu unelastischem Verhalten sind dann Punkte mit ausgeglichener Elastizität $|\varepsilon_f(x)| = 1$.

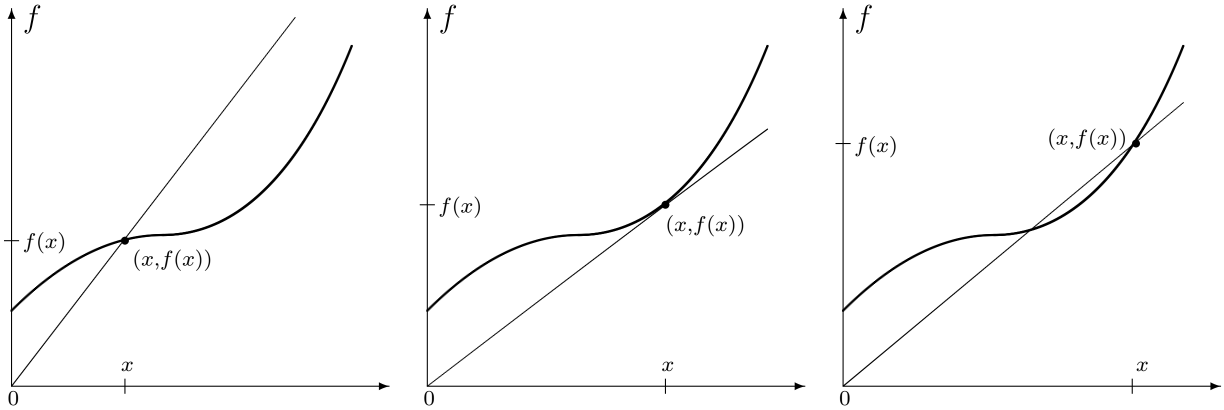
5) Man kann die Elastizität von $f(x)$ schreiben als **Quotient der Grenzfunktion und der Durchschnittsfunktion** (Stückfunktion) $\overline{f}(x) = f(x)/x$:

$$\varepsilon_f(x) = \frac{xf'(x)}{f(x)} = \frac{f'(x)}{\overline{f}(x)} .$$

Das gibt folgende Interpretation der Begriffsbildungen aus 3):

- *f ist elastisch bzw. ausgeglichen elastisch bzw. unelastisch bei x, genau wenn der Grenzfunktionswert $f'(x)$ einen größeren, denselben bzw. einen kleineren Betrag hat als der Durchschnittswert $\overline{f}(x)$.*

Geometrische Interpretation für wachsende Funktionen: Der Durchschnittswert $f(x)/x$ ist die Steigung des Strahls vom Ursprung zum Graphenpunkt $(x, f(x))$ und $f'(x)$ die Steigung des Graphen in diesem Punkt. Wenn $f'(x) < f(x)/x$ ist, so schneidet die Graphenkurve den Strahl von oben nach unten (wenn man ihr von links nach rechts fortschreitend folgt), wenn $f'(x) = f(x)/x$ ist, so berühren sich der Strahl und die Graphenkurve, im Fall $0 < f'(x) < f(x)/x$ schneidet die Graphenkurve von unten durch den Strahl.



f unelastisch in x *f ausgeglichen elastisch in x* *f elastisch in x*

Eine andere Interpretation ergibt sich aus der schon vorgeführten Berechnung der *Ableitung der Durchschnittsfunktion*:

$$\frac{d}{dx}\frac{f(x)}{x} = \frac{xf'(x) - f(x)}{x^2} = \frac{f(x)}{x^2}\Big(\frac{xf'(x)}{f(x)} - 1\Big) , \quad \text{also} \quad \boxed{\overline{f}\,'(x) = \frac{f(x)}{x^2}(\varepsilon_f(x) - 1)} .$$

Daraus liest man ab, dass $\overline{f}\,'(x) > 0$ bzw. $= 0$ bzw. < 0 genau dann gilt, wenn $\varepsilon_f(x) > 1$ bzw. $= 1$ bzw. < 1 ist. (Wir erinnern an unsere Generalvoraussetzung $x > 0$ und $f(x) > 0$ in diesem Abschnitt.) Für wachsende Funktionen f bedeutet das, weil dann $\varepsilon_f(x) \geq 0$ ist:

- *Eine wachsende Funktion f ist elastisch, ausgeglichen elastisch bzw. unelastisch, genau wo ihre Grenzdurchschnittsfunktion positiv, Null bzw. negativ ist.*

Zum Beispiel kann das früher (bei der Quotientenregel) angegebene dritte empirische Gesetz von Keynes, wonach die marginale Konsumquote $c'(Y) = \frac{d}{dY}(C(Y)/Y)$ negativ ist, auch als *Unelastizität der Konsumfunktion* $C(Y)$ ausgesprochen werden.

Für fallende Funktionen, also solche mit $\varepsilon_f < 0$, sind keine mit den obigen vergleichbaren Aussagen möglich; die Durchschnittsfunktion hat dann immer eine negative Ableitung. Auch eine schöne geometrische Interpretation wie oben gibt es dann nicht, da man die absolut genommene (negative) Steigung des Graphen im Punkt $(x, f(x))$ mit der Steigung des Strahls aus dem Ursprung durch diesen Punkt vergleichen müsste. (Man kann, aber das erscheint gezwungen, die Steigung des Graphen mit der des an der horizontalen Geraden mit Hochwert $f(x)$ gespiegelten Ursprungsstrahls vergleichen; dann ist f elastisch in x, wenn der Graph diesen gespiegelten Strahl von oben nach unten schneidet.)

6) Eine *mathematische Interpretation der Elastizität*: Die relativen Differenzenquotienten $(\frac{f(x)}{f(x_0)} - 1)/(\frac{x}{x_0} - 1)$ sind für x nahe bei x_0 ungefähr gleich $[\ln \frac{f(x)}{f(x_0)}]/\ln \frac{x}{x_0}$, weil $\ln t \approx t - 1$ ist für t nahe bei 1 (wegen $\ln t = \ln t - \ln 1 = (t-1)\frac{\ln t - \ln 1}{t-1}$ und $\ln'(1) = 1$). Daher gilt

$$\varepsilon_f(x_0) = \lim_{x \to x_0} \frac{\ln f(x) - \ln f(x_0)}{\ln x - \ln x_0} = \frac{d \ln f(x)}{d \ln x}\Big|_{x=x_0},$$

wobei der letzte Ausdruck als die Ableitung der Funktion $\ln f(\mathrm{e}^t)$ nach $t = \ln x$ an der Stelle $t_0 = \ln x_0$ zu verstehen ist. In der Tat ist nach der Kettenregel diese Ableitung gleich $\mathrm{e}^t f'(\mathrm{e}^t)/f(\mathrm{e}^t) = xf'(x)/f(x) = \varepsilon_f(x)$. Man kann dieses Ergebnis geometrisch so auffassen:

- *Die Elastizität von f ist die Steigung des Graphen von f in logarithmischer Skala, d.h. bei Auftragung des Hochwerts $\ln f(x)$ über dem Rechtswert $\ln x$ für $x > 0$.*

Da die Logarithmusfunktion Multiplikation in Addition verwandelt wegen des Logarithmusgesetzes $\ln(s \cdot t) = \ln s + \ln t$, kann man die Sache auch so ansehen: Wenn man das multiplikative Analogon der Ableitung sucht, so hat man die unabhängige Variable $x > 0$ und die abhängige Variable $y = f(x) > 0$ zunächst mit der natürlichen Logarithmusfunktion zu transformieren, dann die Ableitung von $\ln f(x)$ nach $\ln x$ zu bilden und anschließend mit der Umkehrfunktion $\exp$ zurück zu transformieren. Das Ergebnis ist $\exp(\varepsilon_f(x))$, und in diesem Sinne kann man sagen:

- *Das multiplikative Analogon der Ableitung ist für positive Funktionen $f(x)$ einer positiven Variablen x die Bildung des Exponentials der Elastizität $\exp(\varepsilon_f(x))$.*

7) *Terminologie und Notation*: Im Zusammenhang mit ökonomischen Variablen x, y spricht man von der **Elastizität von *y* bzgl. *x*** oder der ***x*–Elastizität von *y*** oder der ***x*–*y*–Elastizität**, wenn $y = f(x)$ als Funktion von x aufgefasst wird und die Elastizität dieser Funktion gemeint ist. Man schreibt dafür dann

$$\varepsilon_f(x) \qquad \text{oder} \quad \varepsilon_{y,x}\,.$$

Wird dagegen vermöge der Umkehrfunktion $x = g(y)$ die Variable x als Funktion von y aufgefasst (wenn f differenzierbar umkehrbar ist), so steht

$$\varepsilon_g(y) \qquad \text{oder} \quad \varepsilon_{x,y}$$

für die y–Elastizität von x.

- *Die Elastizität wird stets von der abhängigen Variablen gebildet; nur auf diese Variable können sich Attribute wie "elastisch" oder "unelastisch" beziehen. In der Notation $\varepsilon_y(x)$ oder $\varepsilon_{y,x}$ steht immer erst die abhängige Variable y, dann die unabhängige Variable x, in der Terminologie "x–y–Elastizität" ist es umgekehrt.*

Ist z.B. eine Nachfragefunktion $x(p)$ in Abhängigkeit vom Preis p gegeben, so heißt $\varepsilon_x(p) = \varepsilon_{x,p}$ die *Preiselastizität der Nachfrage* oder *Preis–Nachfrage–Elastizität.* Dagegen ist $\varepsilon_p(x) = \varepsilon_{p,x}$ für die Umkehrfunktion $p(x)$ zu $x(p)$ die *Nachfrageelastizität des Preises.* Entsprechend zu verstehen sind z.B. *Elastizität der Kosten bzgl. des Produktionsoutputs* $\varepsilon_K(x)$, wo $K(x)$ die Kosten für die Produktion von x Einheiten angibt, oder die *Einkommenselastizität des Konsums* $\varepsilon_{C,Y}$ für eine Konsumfunktion $C(Y)$ in Abhängigkeit vom Einkommen Y. Die Terminologie legt, wenn man sie dekodieren kann, zweifelsfrei fest, was abhängige und unabhängige Variable sind und von welchen Funktionen die Elastizität zu bilden ist.

Wie die partiellen Ableitungen werden die **partiellen Elastizitäten** einer Funktion $f(x_1, \ldots, x_n) > 0$ von mehreren Variablen erklärt als die Elastizität der Funktionen $0 < x_i \mapsto f(x_1, \ldots, x_i, \ldots, x_n)$ bei c.p.-Bedingung, also bei festgehaltenen Werten der x_j mit $j \neq i$:

$$\varepsilon_{f,x_i}(x_1, \ldots x_n) := \frac{x_i \dfrac{\partial f}{\partial x_i}(x_1, \ldots, x_n)}{f(x_1, \ldots, x_n)} .$$

Die Werte der partiellen Elastizität hängen nicht nur von x_i ab, sondern natürlich auch von den Werten der anderen (festgehaltenen) Variablen; daher ist ε_{f,x_i} eine Funktion von allen Variablen $x_1, \ldots, x_n$.

Hängt zum Beispiel die Nachfrage x_1 nach einem Gut nicht nur von dessen Preis p_1 ab, sondern auch von den Preisen $p_2, \ldots, p_k$ für *konkurrierende Güter*, so ist die Nachfrage nach dem ersten Gut eine Funktion $x_1(p_1, \ldots, p_k)$ von k Variablen und entsprechend auch die Nachfragefunktionen $x_2(p_1, \ldots, p_k), \ldots, x_k(p_1, \ldots, p_k)$ nach den $k-1$ konkurrierenden Gütern. Dann heißt ε_{x_i,p_i} die *Preiselastizität der Nachfrage für das i–te Produkt* und gibt an, um wieviel Prozent sich die Nachfrage nach dem i–ten Produkt (ungefähr) ändert, wenn sein Preis um 1 % erhöht wird. Da bei Preiserhöhung die Nachfrage abnimmt, ist also $\varepsilon_{x_i,p_i} < 0$. Dagegen gibt die sog. *Kreuz–Preiselastizität der Nachfrage* ε_{x_i,p_j} für $j \neq i$ an, um wieviel Prozent sich die Nachfrage nach dem i–ten Produkt (ungefähr) verändert, wenn der Preis für das konkurrierende Produkt Nr. j um 1 % angehoben wird. Da sich die Anhebung des Preises für ein Konkurrenzprodukt steigernd auf die Nachfrage nach dem eigenen Produkt auswirkt, sind diese Kreuz–Preiselastizitäten der Nachfrage positiv, $\varepsilon_{x_i,p_j} > 0$ für $j \neq i$. (Das Präfix "Kreuz" deutet an, dass hier die Nachfrage x_i nicht in Abhängigkeit vom Preis p_i, sondern "über Kreuz" vom Preis p_j für ein anderes Produkt, $j \neq i$, betrachtet wird.)

Auch bei Produktionsfunktionen $x(r_1, \ldots, r_k)$ die vom Input $r_1, \ldots, r_k$ von k Produktionsfaktoren abhängen, werden partielle Elastizitäten $\varepsilon_{x,r_i}(r_1, \ldots, r_k)$ betrachtet; hier heißen sie *Elastizität des Produktionsoutputs bezüglich des i-ten Produktionsfaktors.* ■

BEISPIELE (***zur Berechnung von Elastizitäten***):

Für die Berechnung von Elastizitäten muss man nur Ableitungen berechnen können; denn die Elastizität $\varepsilon_f(x)$ ist ja nichts weiter als das Produkt von $x/f(x)$ mit der Ableitung $f'(x)$. Wenn man also Ableitungen ausrechnen kann, so auch Elastizitäten.

1) Konstante Funktionen $f \equiv c$ haben Ableitung Null und daher natürlich auch Elastizität Null (auf Intervallen in $\mathbb{R}_{>0}$ und für positive Werte der Konstanten c; denn Elastizitäten sind ja nur für positive Funktionen von positiven Variablen definiert).

2) Potenzfunktionen $f(x) = x^s$ haben konstante Elastizität, und zwar ist

$$\varepsilon_{x^s,x} = \frac{x\frac{d}{dx}x^s}{x^s} = \frac{xsx^{s-1}}{x^s} = s$$

gleich dem Exponenten $s \in \mathbb{R}$ für alle $x > 0$. Dasselbe gilt natürlich für Vielfache cx^s von Potenzfunktionen mit $c > 0$ (weil sich der Faktor c bei der Rechnung herauskürzt, bzw. weil sowieso cf dieselbe Elastizitätsfunktion hat wie f wegen der Unabhängigkeit der Elastizität von der Wahl von Einheiten). Man nennt derartige Funktionen cx^s auch **isoelastische Funktionen**, weil sie an jeder Stelle $x > 0$ dieselbe Elastizität haben. Es lässt sich zeigen, dass umgekehrt jede isoelastische Funktion Vielfaches einer Potenzfunktion ist. Für Exponenten s mit $|s| > 1$ ist cx^s elastisch auf $\mathbb{R}_{>0}$, für $|s| < 1$ unelastisch, für $s = \pm 1$ überall auf $\mathbb{R}_{>0}$ ausgeglichen elastisch (das sind also die Funktionen cx und c/x, die proportionale bzw. umgekehrt proportionale Abhängigkeit beschreiben), und für $s = 0$ starr auf $\mathbb{R}_{>0}$ (konstante Funktionen).

3) Exponentialfunktionen *haben eine homogen lineare Elastizitätsfunktion*, und zwar ist für $a^x = \mathrm{e}^{cx}$ mit $c = \ln a$

$$\varepsilon_{a^x,x} = \frac{x\frac{d}{dx}\mathrm{e}^{cx}}{\mathrm{e}^{cx}} = \frac{x\mathrm{e}^{cx}c}{\mathrm{e}^{cx}} = cx = (\ln a)x$$

für alle $x > 0$. Im Bereich $x > 1/|\ln a|$ ist a^x elastisch, im Bereich $0 < x < 1/|\ln x|$ unelastisch, an der Übergangsstelle $x = 1/|\ln a|$ ausgeglichen elastisch.

4) Für "*überexponentielle Funktionen*" $f(x) = \mathrm{e}^{x^s}$ ist

$$\varepsilon_f(x) = \frac{x\mathrm{e}^{x^s}sx^{s-1}}{\mathrm{e}^{x^s}} = sx^s$$

Vielfaches einer Potenzfunktion. (Von "überexponentiell", d.h. schneller als exponentiell wachsend bei $x \to \infty$, kann man nur im Fall $s > 1$ sprechen; die Rechnung gilt aber für beliebige $s \in \mathbb{R}$.) Allgemeiner ist f

$$\varepsilon_f(x) = xg'(x) \qquad \text{für} \quad f(x) = \mathrm{e}^{g(x)} .$$

Funktionen mit polynomialer Elastizitätsfunktion $p(x) = a_n x^n + \ldots + a_1 x + a_0$ erhält man daher durch einen Ansatz $f(x) = \exp(b_n x^n + \ldots + b_1 x + b_0 \ln x)$. Die Berechnung der Elastizität bei diesem Ansatz gibt $x(b_n n x^{n-1} + \ldots + b_1 + b_0/x)$, so dass man mit der Wahl $b_n := \frac{1}{n}a_n$, $b_{n-1} := \frac{1}{n-1}a_{n-1}$, $\ldots$, $b_1 := a_1$ und $b_0 := a_0$ die gegebene Polynomfunktion $\varepsilon_f(x) = p(x)$ als Elastizitätsfunktion erhält.

5) Für die **logistische Funktion** $f(x) = b/(1 + ae^{-cx})$ mit Grenzwerten 0 bei $x \to -\infty$, b bei $x \to \infty$ und Wert $\frac{b}{1+a}$ an der Stelle $x = 0$ (bei Parametern $a, b, c > 0$) ist die Elastizitätsfunktion

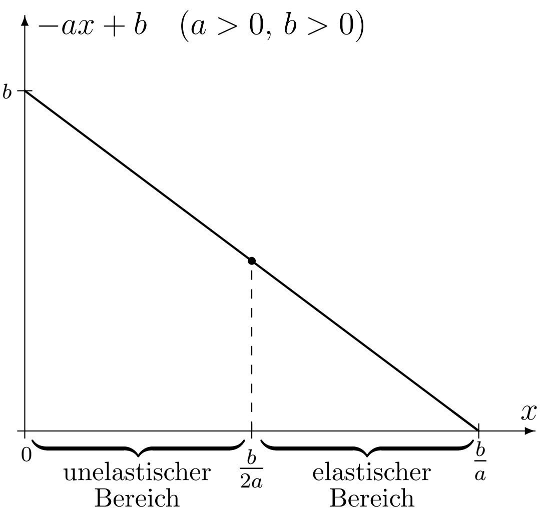
$$\varepsilon_f(x) = x \cdot \frac{1 + ae^{-cx}}{b} \cdot \frac{-ba(-c)e^{-cx}}{(1 + ae^{-cx})^2} = x\frac{ace^{-cx}}{1 + ae^{-cx}} = \frac{acx}{e^{cx} + a} .$$

Genau an den Stellen $x \in \mathbb{R}_{>0}$ mit $e^{cx} + a > acx$ ist f unelastisch. Da die Elastizitätsfunktion den Grenzwert 0 bei $x \searrow 0$ und bei $x \to \infty$ hat, ist f also unelastisch sowohl für hinreichend kleine Werte $x > 0$ als auch für hinreichend große Werte von x. Ob es dazwischen einen elastischen Bereich gibt, ob also $e^{cx} + a < acx$ ist für gewisse $x > 0$, hängt vom Parameter a ab. Mit Differentialrechnung kann man zeigen, dass $e^{cx} + a - acx$ minimal wird, wenn $e^{cx} = a$ ist, und dann ist $e^{cx} + a - acx < 0$, genau wenn $a > e^2$ gilt. In diesem Fall gibt es also einen elastischen Bereich von f um $\frac{1}{c}\ln a$, im Fall $0 < a < e^2$ ist f überall unelasisch, im Grenzfall $a = e^2$ ist f an der Stelle $x = \frac{1}{c}\ln a = \frac{2}{c}$ ausgeglichen elastisch und sonst überall unelastisch auf $\mathbb{R}_{>0}$.

6) Lineare Funktionen $\ell(x) = ax + b$ kommen in der Wirtschaftsmathematik sehr oft vor; daher ist es wichtig, ihre Elastizität zu berechnen. Auf dem maximalen Teilintervall von $\mathbb{R}_{>0}$, auf dem $\ell(x) > 0$ ist, gilt

$$\varepsilon_{ax+b,x} = \frac{x\frac{d}{dx}(ax + b)}{ax + b} = \frac{ax}{ax + b} .$$

Die Elastizität einer linearen Funktion ist also eine gebrochen lineare Funktion. Im Fall $b = 0$, $a > 0$ ist $\varepsilon_{ax,x} = 1$, die Funktion also ausgeglichen elastisch. Im Fall $b < 0$, $a > 0$ ist sie elastisch auf $]-\frac{b}{a}, \infty[$, im Fall $b > 0$, $a > 0$ ist sie überall unelastisch auf $\mathbb{R}_{>0}$. Im interessantesten Fall $\ell(x) = -ax + b$ mit $b > 0$, $a > 0$ schließlich ist die Elastizität negativ und die lineare Funktion unelastisch auf der linken Hälfte $]0, \frac{b}{2a}[$ des Definitionsintervalls, ausgeglichen elastisch an der Stelle $\frac{b}{2a}$ und elastisch auf der rechten Hälfte $]\frac{b}{2a}, \frac{b}{a}[$.



7) *Lineare Funktionen in der Ökonomie* treten z.B. als Kostenfunktionen $K(x) = K_{\text{fix}} + K_{\text{var}}(x)$ auf mit positiven Fixkosten $K_{\text{fix}} > 0$ und homogen linearen variablen Kosten $K_{\text{var}}(x) = \kappa \cdot x$ proportional zum Output x (mit Proportionalitätsfaktor $\kappa > 0$). Solche Funktionen sind immer unelastisch, d.h. eine Erhöhung des Outputs um 1 % bewirkt eine Kostensteigerung von weniger als 1 %. (Das liegt daran, dass die Fixkosten nicht mitsteigen).

Interessanter ist der Fall einer linearen Nachfragefunktion $x(p) = b - ap$ auf dem Intervall $[0, \frac{b}{a}]$, wobei $a > 0$ ist und das Minuszeichen steht, weil die Nachfrage ja bei Preiserhöhung abnimmt. Hier ist $\varepsilon_{x,p} = \frac{-ap}{b-ap}$ negativ für $0 < x < \frac{b}{a}$ und es gibt elastische und unelastische Bereiche der Nachfrage. Genauer ist $\varepsilon_{x,p} < -1$ für $\frac{b}{2a} < x < \frac{b}{a}$ (elastischer Bereich) und $-1 < \varepsilon_{x,p} < 0$ für $0 < p < \frac{b}{2a}$ (unelastischer Bereich); genau bei $p = \frac{b}{2a}$ ist die Nachfragefunktion ausgeglichen elastisch (vgl. die Berechnung in Beispiel 6) oben).

Bei zwei konkurrierenden Gütern kann man in einem einfachen Modell die Nachfragefunktionen linear $x_1(p_1, p_2) = c - a_1p_1 + a_2p_2$ und $x_2(p_1, p_2) = d + b_1p_1 - b_2p_2$ ansetzen mit positiven Parametern a_1, a_2, b_1, b_2, c, d und einer Vorzeichenwahl, die der Tatsache entspricht, dass die Nachfrage für ein Gut eine abnehmende Funktion seines Preises ist, aber zunimmt, wenn der Preis für ein konkurrierendes Produkt erhöht wird. Der ökonomisch sinnvolle Definitionsbereich der Nachfragefunktionen ist die Menge der Preis–Paare $p_1 > 0$ und $p_2 > 0$, für die x_1 und x_2 positiv ausfallen. Die *Elastizitäten der Nachfrage* sind hier

$$\varepsilon_{x_1,p_1} = \frac{-a_1p_1}{x_1(p_1,p_2)} < 0 \qquad \text{und} \qquad \varepsilon_{x_2,p_2} = \frac{-b_2p_2}{x_2(p_1,p_2)} < 0 ,$$

und die *Kreuz–Preis–Eastizitäten der Nachfrage* sind

$$\varepsilon_{x_1,p_2} = \frac{a_2p_2}{x_1(p_1,p_2)} > 0 \qquad \text{und} \qquad \varepsilon_{x_2,p_1} = \frac{b_1p_1}{x_2(p_1,p_2)} > 0 .$$

Erhöhung von Preis p_1 um $1\,\%$ (bei festem Preis p_2) verkleinert die Nachfrage x_1 also etwa um $(a_1p_1/x_1)\,\%$ und vergrößert die Nachfrage x_2 nach dem konkurrierenden Gut um etwa $(b_1p_1/x_2)\,\%$. Die Diskussion für mehr als zwei konkurrierende Güter verläuft völlig analog und ist, abgesehen von Nummerierungsproblemen, nicht schwieriger. ■

Da die Elastizität $\varepsilon_f(x) = \frac{x}{f(x)} f'(x)$ durch die Ableitung $f'(x)$ einfach ausgedrückt ist, ergeben sich aus den Rechenregeln für die Ableitung sofort entsprechende Rechenregeln für die Elastizität. Weil die Elastizität (bzw. genauer ihr Exponential) das multiplikative Analogon der Ableitung ist, werden diejenigen Rechenregeln für die Elastizität besonders einfach, die sich auf die Multiplikation und die Division von Funktionen beziehen.

RECHENREGELN für die Elastizität:

1) Für differenzierbare $f : \mathbb{R}_{>0} \supset I \to \mathbb{R}_{>0}$ und Konstanten $c > 0$, $a > 0$ gelten auf I die

- **Faktorregeln:** $\varepsilon_{cf}(x) = \varepsilon_f(x)$, $\varepsilon_{f(ax),x} = \varepsilon_f(ax)$.

Dies ist nichts weiter als die schon festgestellte Unabhängigkeit der Elastizität von der Einheitenwahl bei der abhängigen und der unabhängigen Variablen.

2) Für differenzierbare Funktionen $f, g : \mathbb{R}_{>0} \supset I \to \mathbb{R}_{>0}$ gelten auf I :

- **Produktregel:** $\varepsilon_{f\cdot g}(x) = \varepsilon_f(x) + \varepsilon_g(x)$;

  die Elastizität eines Produkts von Funktionen ist die Summe ihrer Elastizitäten;

- **Quotientenregel:** $\varepsilon_{f/g}(x) = \varepsilon_f(x) - \varepsilon_g(x)$;

  die Elastizität des Quotienten von zwei Funktionen ist die Differenz ihrer Elastizitäten.

Der Beweis ergibt sich aus der Produkt- und Quotientenregel für die Ableitung sofort:

$$\varepsilon_{f+g}(x) = \frac{x(fg)'(x)}{(fg)(x)} = \frac{x[f'(x)g(x) + f(x)g'(x)]}{f(x)g(x)} = \frac{xf'(x)}{f(x)} + \frac{xg'(x)}{g(x)} = \varepsilon_f(x) + \varepsilon_g(x) ,$$

$$\varepsilon_{f/g}(x) = \frac{x(f/g)'(x)}{(f/g)(x)} = \frac{x[f'(x)g(x) - f(x)g'(x)]}{f(x)g(x)} = \frac{xf'(x)}{f(x)} - \frac{xg'(x)}{g(x)} = \varepsilon_f(x) - \varepsilon_g(x) .$$

3) Für differenzierbare Funktionen $f : \mathbb{R}_{>0} \supset I \to J \subset \mathbb{R}_{>0}$ und $h : J \to \mathbb{R}_{>0}$ gilt die

- **Kettenregel:** $\varepsilon_{h\circ f}(x) = \varepsilon_h(f(x)) \cdot \varepsilon_f(x)$;

  die Elastizität einer Verkettung von zwei Funktionen ist das Produkt der äußeren Elastizität am Wert der inneren Funktion mit der inneren Elastizität.

Diese Rechenregel ist vollkommen analog zur Kettenregel für die Berechnung von Ableitungen und folgt daraus auch sofort:

$$\varepsilon_{(h\circ f)(x)} = \frac{x(h\circ f)'(x)}{(h\circ f)(x)} = \frac{x(h'(f(x))f'(x)}{h(f(x))} = \frac{f(x)h'(f(x))}{h(f(x))} \cdot \frac{xf'(x)}{f(x)} = \varepsilon_h(f(x))\varepsilon_f(x) \; .$$

Mit $y = f(x)$ und $z = h(y)$ wird diese Regel symbolisch so geschrieben:

$$\varepsilon_{z,x} = \varepsilon_{z,y} \cdot \varepsilon_{y,x}$$

4) Für streng monotone und differenzierbare Funktionen $f : \mathbb{R}_{>0} \supset I \to J \subset \mathbb{R}_{>0}$ mit Werteintervall J und die Umkehrfunktion $g : J \to I \subset \mathbb{R}_{>0}$ gilt an den Stellen $x \in I$ mit $\varepsilon_f(x) \neq 0$ die

- **Umkehrregel:** $\varepsilon_g(y) = \dfrac{1}{\varepsilon_f(x)}$ $\quad (y = f(x), \;\; x = g(y), \;\; \varepsilon_f(x) \neq 0)$;

  die Elastizität einer Funktion und die Elastizität ihrer Umkehrfunktion an entsprechenden Stellen sind reziprok zueinander, wenn $\neq 0$.

Auch diese Regel ist völlig analog zur Regel für die Ableitung der Umkehrfunktion und folgt daraus, da mit $\varepsilon_f(x) \neq 0$ auch $f'(x) \neq 0$ ist:

$$\varepsilon_g(y) = \frac{yg'(y)}{g(y)} = \frac{f(x)/f'(x)}{x} = \frac{1}{xf'(x)/f(x)} = \frac{1}{\varepsilon_f(x)} \; .$$

Symbolisch kann man die Umkehrregel einfach so schreiben, und so wird sie in der Wirtschaftsmathematik routinemäßig verwendet:

$$\varepsilon_{y,x} = \frac{1}{\varepsilon_{x,y}} \qquad \text{oder} \qquad \varepsilon_{x,y} \cdot \varepsilon_{y,x} = \varepsilon_{x,x} = 1 \; .$$

5) Für die Elastizität der Summe von Funktionen gibt es leider keine einfache Regel; denn

$$\varepsilon_{f+g}(x) = \frac{x(f+g)'(x)}{(f+g)(x)} = \frac{xf'(x) + xg'(x)}{f(x) + g(x)}$$

lässt sich nicht durch $\varepsilon_f(x)$ und $\varepsilon_g(x)$ ausdrücken. Auch bei Verschiebung der abhängigen oder unabhängigen Variablen ist die Elastizität der verschobenen Funktion

$$\varepsilon_{f(x)+c,x} = \frac{xf'(x)}{f(x)+c} \qquad \text{und} \qquad \varepsilon_{f(x+b),x} = \frac{xf'(x+b)}{f(x+b)}$$

im Allgemeinen nicht durch $\varepsilon_f(x)$ bzw. $\varepsilon_f(x+b)$ auszudrücken. ■

BEISPIELE (***zu den Rechenregeln für die Elastizität und ökonomischen Anwendungen***):

1) Aus der Kettenregel folgt für die *Elastizität der Potenz* $f(x)^s$ *einer Funktion* $y = f(x)$:

$$\varepsilon_{f^s}(x) = \varepsilon_{y^s,y} \cdot \varepsilon_f(x) = s \cdot \varepsilon_f(x) \ ;$$

also ist die Elastizität von $f(x)^s$ einfach das s–fache der Elastizität von $f(x)$. Für das Exponential $\mathrm{e}^{f(x)}$ ergibt sich:

$$\varepsilon_{\mathrm{e}^f}(x) = \varepsilon_{\mathrm{e}^y,y} \cdot \varepsilon_{f(x)} = y \cdot \varepsilon_{f(x)} = f(x)\varepsilon_{f(x)} = xf'(x) \ ,$$

also das Produkt von $f(x)$ mit seiner Elastizität $\varepsilon_f(x)$.

2) Die *Elastizität einer Durchschnittsfunktion* $\overline{f}(x) = f(x)/x$ ergibt sich sofort aus der schon mehrfach berechneten Ableitung der Durchschnittsfunktion. Hier folgern wir mit $\varepsilon_{x,x} = 1$ das Ergebnis aus der Quotientenregel:

$$\varepsilon_{\overline{f}}(x) = \varepsilon_f(x) - \varepsilon_{x,x} = \varepsilon_f(x) - 1$$

bzw.

$$\varepsilon_f(x) = 1 + \varepsilon_{\overline{f}}(x) \ .$$

Wir sehen erneut, dass f elastisch ist, genauer $\varepsilon_f(x) > 1$, an Stellen x positiver Ableitung (und damit auch positiver Elastizität) der Durchschnittsfunktion, während eine nichtfallende Funktion f (also $\varepsilon_f(x) \geq 0$) unelastisch ist an Stellen x negativer Ableitung der Durchschnittsfunktion, genauer $0 \leq \varepsilon_f(x) < 1$.

3) Wir betrachten eine *Produktionsfunktion* $x(r_1, r_2)$, die den Produktionsoutput x in Abhängigkeit von den eingesetzten Einheiten r_1 und r_2 zweier Produktionsfaktoren angibt und die Form

$$x(r_1, r_2) = b(a_1 r_1^s + a_2 r_2^s)^{1/s}$$

mit positiven Parametern a_1, a_2, b besitze. Die partiellen Elastizitäten, die angeben, um wieviel Prozent sich der Output erhöht, wenn der Einsatz eines Faktors um 1 % gesteigert wird (und der des anderen Faktors unverändert bleibt), sind

$$\varepsilon_{x,r_i}(r_1, r_2) = \frac{r_i \frac{1}{s}(a_1 r_1^s + a_2 r_2^s)^{1/s-1} a_i s r_i^{s-1}}{x(r_1, r_2)} = \frac{a_i r_i^s}{x(r_1, r_2)} \ .$$

Von ökonomischem Interesse ist hier die *Substitution* eines Faktors durch den anderen, wenn man den Output auf einem fixierten Niveau $x(r_1, r_2) = c$ halten will und dazu den Mindereinsatz eines Faktors durch entsprechend höheren Einsatz des anderen Faktors kompensieren muss. Die dadurch bewirkte Abhängigkeit der Faktoren voneinander beschreibt die *Substitutionsfunktion* $r_2(r_1)$, die man durch Auflösen der Gleichung $x(r_1, r_2) = c$ nach r_2 erhält, hier also

$$r_2(r_1) = \left[\frac{1}{a_2}\left(\frac{c}{b}\right)^s - \frac{a_1}{a_2} r_1^s\right]^{1/s} \ .$$

Die (negative) Ableitung dieser Funktion, genannt *Grenzrate der Substitution*, drückt aus, um wieviel Einheiten der Einsatz des zweiten Faktors reduziert werden kann, wenn der des ersten um eine kleine Einheit erhöht wird und das Output-Niveau c gehalten wird.

Diese Grenzrate berechnet sich zu

$$\frac{dr_2}{dr_1}(r_1) = \frac{1}{s}\left[\frac{1}{a_2}\left(\frac{c}{b}\right)^s - \frac{a_1}{a_2}r_1^s\right]^{1/s-1}\left(\frac{-a_1}{a_2}\right)sr_1^{s-1} = -\frac{a_1}{a_2}\left(\frac{r_1}{r_2(r_1)}\right)^{s-1} .$$

Die *Elastizität der Substitutionsfunktion*

$$\varepsilon_{r_2,r_1} = \frac{r_1}{r_2(r_1)}\frac{dr_2}{dr_1}(r_1) = -\frac{a_1}{a_2}\left(\frac{r_1}{r_2(r_1)}\right)^s$$

gibt dann an, um wieviel Prozent der Einsatz des zweiten Faktors reduziert werden kann, wenn der des ersten um 1 Prozent gesteigert wird und das Output-Niveau c unverändert bleibt.

In der Ökonomie wird noch eine andere Größe betrachtet, die sog. *Substitutionselastizität*, die (recht kompliziert) definiert wird als Quotient der relativen Änderungen des Faktorverhältnisses $r_2(r_1)/r_1$ und der Substitutionsgrenzrate $\frac{d}{dr_1}r_2(r_1)$ oder, was dasselbe ist, als Quotient der Elastizitäten dieser beiden Funktionen. Da im vorliegenden Fall die Grenzrate der Substitution $\frac{d}{dr_1}r_2(r_1)$ ein Vielfaches der $(s-1)$-ten Potenz des Faktorverhältnisses $r_2(r_1)/r_1$ ist, ergibt sich mit 1), dass die Elastizität dieser Grenzrate das $(s-1)$-fache der Elastizität des Faktorverhältnisses ist. Die Substitutionselastizität ist also hier ein konstanter Wert $\frac{1}{s-1}$. Deshalb wird eine Produktionsfunktion der obigen Form auch **CES–Funktion** genannt (von **C**onstant **E**lasticity of **S**ubstitution).

4) Eine ökonomische Anwendung von 2) ergibt sich, wenn man ein *Erlösfunktion* $E(x)$ (*Umsatzfunktion*) betrachtet. Die zugehörige Durchschnittsfunktion ist dann der *Preis* pro Mengeneinheit $p(x) = E(x)/x$, der am Markt bei Angebot und Absatz von x Mengeneinheiten des Gutes erzielt wird. Die Preisfunktion $x(p)$ nimmt man sinnvollerweise streng fallend mit Ableitung < 0 an (das ist realistisch außer bei speziellen Luxusgütern, die bei niedrigeren Preisen weniger gekauft werden). Die Umkehrfunktion $x(p)$ ist dann die *Nachfragefunktion*, die angibt, welche Menge des produzierten Gutes am Markt abgesetzt werden kann und daher auch produziert wird, wenn der Preis p ist. Die Nachfragefunktion ist natürlich auch fallend. Aus 2) oder durch direkte Rechnung aus $E(x) = x \cdot p(x)$ bzw. $E(p) = x(p) \cdot p$ bekommt man:

$$\varepsilon_E(x) = 1 + \varepsilon_p(x) \qquad \text{bzw.} \qquad \varepsilon_E(p) = 1 + \varepsilon_x(p) ,$$

was wegen $\varepsilon_E(x) = xE'(x)/E(x) = E'(x)/p(x)$ bzw. $\varepsilon_E(p) = pE'(p)/E(p) = E'(p)/x(p)$ auch so geschrieben werden kann:

$$E'(x) = p(x)(1 + \varepsilon_p(x)) \qquad \text{bzw.} \qquad E'(p) = x(p)(1 + \varepsilon_x(p)) .$$

Dies ist in der Ökonomie die sog. **Amoroso–Robinson–Beziehung** zwischen *Grenzerlös* (*Grenzumsatz*) $E'(p)$, *Preis* p und *Preiselastizität der Nachfrage* $\varepsilon_x(p)$. Mathematisch gesehen handelt es sich dabei nur um eine sehr einfache Anwendung der Produktregel, nämlich $E'(p) = \frac{d}{dp}(px(p)) = x(p) + px'(p))$, und die Definition der Elastizität $\varepsilon_x(p) = px'(p)/x(p)$. Die ökonomische Bedeutung der Beziehung besteht darin, dass sie eine wichtige Größe, den Grenzerlös, durch die von Einheitenwahlen unabhängige Preiselastizität der Nachfrage $\varepsilon_{x,p}$ bzw. die reziproke Elastizität $\varepsilon_{p,x}$ ausdrückt. Am Grenzerlös $E'(x)$, also am Mehrerlös für eine mehr produzierte und auf den Markt gebrachte Output-Einheit, ist ein Unternehmer natürlich interessiert; insbesondere wird er die Mehrproduktion unterlassen, wenn der Grenzerlös negativ ist, was bedeutet, dass durch das Mehrangebot ein solcher Preisverfall entsteht, dass er trotz größerem Umsatz weniger erlösen würde.

Um Missverständnisse zu vermeiden, weisen wir darauf hin, dass $E'(x)$ und $E'(p)$ oben für $x = x(p)$ nicht gleich sind; denn mit $E(p)$ ist eigentlich die mittelbare Funktion $E(x(p))$ gemeint (für die man einfach $E(p)$ geschrieben hat, aber besser ein anderes Buchstabensymbol gewählt hätte, weil es ja nicht dieselbe Funktion wie $E(x)$ ist), und die Ableitung berechnet sich dementsprechend gemäß der Kettenregel zu $\frac{d}{dp}E(x(p)) = E'(x(p))x'(p)$, d.h. mit $E'(p)$ ist oben eigentlich $E'(x)x'(p)$ für $x = x(p)$ gemeint. Insbesondere haben $E'(x)$ und der richtig verstandene Wert $E'(p)$ entgegengesetztes Vorzeichen, weil ja $x'(p) < 0$ ist. Wenn der Erlös bei Absatz einer Mehreinheit noch wächst, so nimmt er eben bei *abnehmendem* Preis zu (denn der Preis fällt ja bei Absatzerhöhung), d.h. als Funktion des Preises ist der Erlös fallend!

(5) Da die Nachfragefunktion $x(p)$ negative Ableitung hat, also auch negative Elastizität $\varepsilon_x(p)$, und da auch $\varepsilon_p(x) = 1/\varepsilon_x(p)$ negativ ist, kann man aus der Amoroso–Robinson–Beziehung folgendes ablesen:

- *Im unelastischen Bereich der Nachfrage ist der Grenzerlös $E'(p)$ positiv und $E'(x)$ negativ;* denn dann ist $-1 < \varepsilon_x(p) < 0$, also $1 + \varepsilon_x(p) > 0$; in diesem Bereich wächst daher der Erlös, wenn der Marktpreis (um eine kleine Spanne) steigt und er nimmt ab, wenn man die am Markt angebotene Menge des Produkts vergrößert, weil der dadurch verursachte Preisverfall den Mehrabsatz überkompensiert.
- *Im elastischen Bereich der Nachfrage ist der Grenzerlös $E'(p)$ negativ und $E'(x)$ positiv;* denn dann ist $\varepsilon_x(p) < -1$, also $1 + \varepsilon_x(p) < 0$; in diesem Bereich sinkt daher der Erlös bei geringfügiger Preiserhöhung, d.h. die dadurch verursachten Absatzeinbußen überkompensieren den erzielten höheren Preis, während der Erlös bei Erhöhung des Output x noch zunimmt trotz des dadurch bewirkten Drucks auf die Preise.

Allerdings wird der Grenzerlös den Anbieter meist weniger kümmern als sein Gewinn. Vielmehr wird er versuchen, den Absatz x so einzurichten, bzw. den Preis p so zu beeinflussen, dass sein *Gewinn* $G(x) = E(x) - K(x)$ maximal ist. Im Gewinnmaximum ist $G'(x) = 0$; denn wenn z.B. $G'(x) > 0$ ist, so sind auch die Differenzenquotienten zu kleinen Änderungen von x positiv, d.h. G nimmt rechts von x größere Werte an als an der Stelle x (und im Fall $G'(x) < 0$ analog links von G; dieses Argument ist die Grundlage der Extremstellenbestimmung, siehe 4.5). Für Kostenfunktionen $K(x)$ gilt allgemein $K'(x) > 0$, d.h. die Kosten steigen bei wachsendem Produktionsoutput (nur die Stückkosten $k(x) = K(x)/x$ sinken dabei vielleicht). Also folgt $0 < K'(x) = E'(x) - G'(x) = E'(x) = p(x)(1 + \varepsilon_p(x))$, und wegen $p(x) > 0$ weiter $\varepsilon_p(x) > -1$, und mit $\varepsilon_p(x) < 0$ auch $\varepsilon_x(p) = 1/\varepsilon_p(x) < -1$. Das Ergebnis dieser Überlegungen ist also:

- *Im Gewinnmaximum ist die Nachfrage elastisch bzgl. des Preises;* kleine prozentuale Preisänderungen bewirken dann größere prozentuale Nachfrageänderungen.

Für die Bestimmung des Output x^* mit maximalem Gewinn $G(x^*)$ bedeutet dies, dass man nach der Stelle x^* von vorneherein nur in dem Intervall (oder Bereich) suchen muss, in dem die Nachfrage elastisch bzgl. des Preises ist. ■