

1 Zur Einstimmung



© Springer-Verlag Berlin Heidelberg 2016

K.-J. Förster (Hrsg.), *Die Geschichte des Prioritätsstreits zwischen Leibniz und Newton,*
Vom Zählstein zum Computer, DOI 10.1007/978-3-662-48862-1_1

1.1 Worum geht es eigentlich?

Bereits im Altertum gelang es Archimedes (um 287–212 v. Chr.), zum einen Tangenten an gegebene Kurven, und zum anderen den Flächeninhalt unter einer gegebenen Kurve zu berechnen. Er konnte zu seiner Zeit nicht erkennen, dass diese beiden Operationen – Tangentenberechnung und Flächenberechnung – zueinander inverse Operationen sind, was sich im Hauptsatz der Differenzial- und Integralrechnung ausdrückt, der erst von Newton und Leibniz erkannt und von Leibniz aus der Geometrie eines Isaac Barrow (1630–1677) heraus in die neue Sprache der symbolischen Algebra gebracht wurde, in der er seine Kraft voll entfalten konnte.

Aber *warum*, um Gottes Willen, will man Tangenten und Flächen ausrechnen? So haben sich sicherlich Generationen von Schülerinnen und Schülern gefragt und vielleicht fragen Sie es sich in diesem Moment auch. Ich will eine Erklärung geben und zuvor versuchen klarzumachen, wie man solche Rechnungen durchführt. Dabei greife ich auf ein großes Vorbild zurück: Das Buch *Calculus made easy* des englischen Physikers Silvanus Phillips Thompson (1851–1916).

Calculus made easy erschien 1910 und ist bis heute (!) im Druck. Im Jahr 1998 gab der bekannte Autor und Wissenschaftsjournalist Martin Gardner (1914–2010) das Buch [Thompson 1998a] mit einer neuen Einführung und leicht modernisiert heraus, und im selben Jahr erschien bereits der Nachdruck der 12. Auflage der deutschen Übersetzung *Analysis leicht gemacht* [Thompson 1998b]. Die von Thompson verwendeten Erläuterungen halten moderner Strenge nicht immer stand, aber genügen für unsere Zwecke! Wenn Sie eine gute Ausbildung in Differenzial- und Integralrechnung genossen haben, können Sie die folgenden Abschnitte natürlich überspringen.

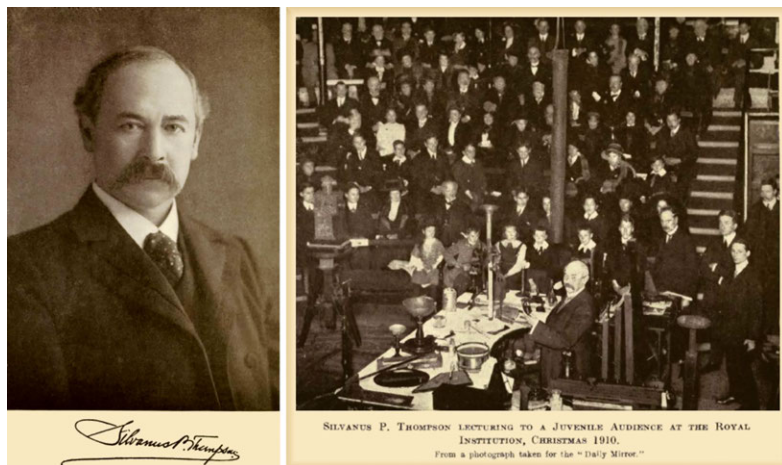


Abb. 1.1.1. Silvanus P. Thompson [Foto: Fotograf unbekannt, vor 1900 aus: Thompson and Thompson, 1920] und beim Vortrag vor einem jugendlichen Auditorium [Foto: Fotograf der „Daily Mirror“, 1910, Thompson and Thompson, 1920]

1.2 Steigung, Ableitung und Differenzialquotient

Wenn auf einem Verkehrsschild vor einer Steigung von 10 % gewarnt wird, was ist damit dann eigentlich gemeint? Muss ich bei 100 % Steigung etwa senkrecht die Wand hoch?



Abb. 1.2.1. Achtung: 10 % Steigung

Die Steigung macht eine Aussage über das Verhältnis der zu überwindenden Höhe zur gefahrenen Strecke. Müssen also auf den nächsten 5 km 250 Höhenmeter überwunden werden, dann ist die Steigung

$$s = \frac{250 \text{ m}}{5000 \text{ m}} = 0.05$$

oder 5 %. Muss man auf den nächsten 5 km 500 Höhenmeter überwinden, wäre das eine Steigung von 10 %, und wenn auf eine horizontale Strecke von 5 km auch eine Höhendifferenz von 5 km käme, dann wäre die Steigung 100 %. Abbildung 1.2.2 zeigt die Situation, wobei Δx die zu fahrende horizontale Strecke, und Δy die Höhendifferenz bezeichnet.

Mit der Steigung können wir wie in Abbildung 1.2.2 über die Tangensfunktion einen Winkel α assoziieren,

$$\tan \alpha = \frac{\Delta y}{\Delta x},$$

denn das **Steigungsdreieck** ist offenbar ein rechtwinkliges Dreieck. Der Winkel α bei unserem Beispiel von 5 % Steigung wäre

$$\alpha = \arctan 0.05 \approx 2.86^\circ,$$

10 % Steigung korrespondiert zu einem Winkel von 5.7° und 100 % Steigung führt auf $\alpha = \arctan 1 = 45^\circ$.

Nun ist eine Straße ganz sicher keine aus ebenen Stücken zusammengesetzte Fahrbahn, sondern in ihrem Profil eher eine Funktion

$$y = f(x)$$

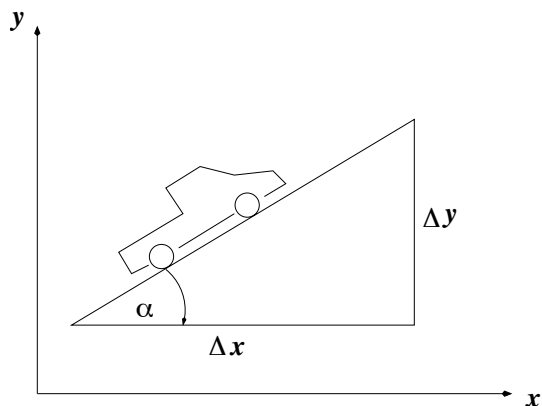


Abb. 1.2.2. Steigung und Winkel

wie in Abbildung 1.2.3 gezeigt. Wie man sieht, entspricht die ausgerechnete Steigung $\Delta y/\Delta x$ nun nicht mehr genau der wahren Steigung am Punkt x_0 und je größer wir Δx machen, umso weniger hat unsere Steigung mit der wahren Steigung zu tun. Hier setzt man nun den neuen Kalkül von Leibniz und Newton ein!

Wir nehmen an, wir machen in x -Richtung nur einen winzigen Schritt. Wirklich sehr, sehr winzig. Also eigentlich ein unendlich kleiner Schritt (oder, in der Übersetzung von Thompsons Buch, „unbegrenzt klein“ [Thompson 1998b, S. 100]), und um diesen Schritt von Δx zu unterscheiden, nennen wir ihn in Leibniz'scher Notation dx . Das „ d “ kann dabei gelesen werden als „ein winziges Stück von“.

Damit ist dx „ein winziges Stückchen von x “.

Gehen wir von x_0 aus den winzigen Schritt dx nach rechts, dann gibt die unbegrenzt kleine Größe dy die Höhendifferenz an. Wenn wir von x_0 in den Punkt $x_0 + dx$ gehen, ist aus dem Funktionswert $y_0 = f(x_0)$ gerade der Funktionswert $y_0 + dy$ geworden. Da wir nur einen unbegrenzt kleinen Schritt weg von x_0 gemacht haben, muss der **Differenzialquotient**

$$s(x_0) := \frac{dy}{dx} \quad (1.1)$$

die wahre Steigung der Fahrbahn am Punkt x_0 angeben. Diese wahre Steigung beschreibt die **Änderung** der Fahrbahn am Punkt x_0 . Aber wie berechnet man nun diesen Quotienten aus zwei unbegrenzt kleinen Größen?

Ein gutes Beispiel liefert die Parabel $y = f(x) = x^2$. Wollen wir die Steigung am Punkt $x_0 = 2$ berechnen, dann müssen wir die Änderung von $y_0 = f(x_0) = x_0^2$ zu $y_0 + dy = f(x_0 + dx)$ betrachten:

$$y_0 + dy = f(x_0 + dx) = (x_0 + dx)^2 = x_0^2 + 2x_0 dx + (dx)^2.$$

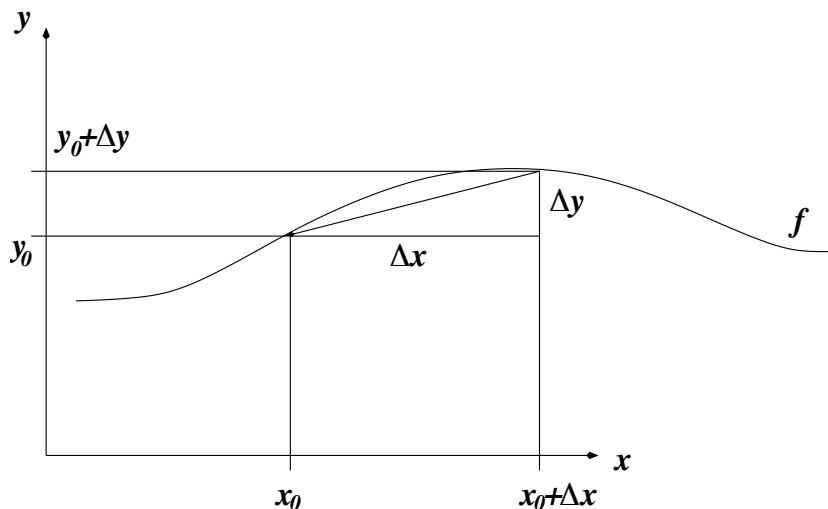


Abb. 1.2.3. Realistischer Straßenverlauf

Nun ist $y_0 = x_0^2$, so dass wir

$$dy = 2x_0 dx + (dx)^2 \quad (1.2)$$

erhalten. Die Größe dx haben wir als unbegrenzt klein angenommen. Wie groß ist dann wohl das Quadrat einer solchen Größe? Wir ahnen schon, dass wir ohne Probleme $(dx)^2 = 0$ setzen dürfen, aber warum dürfen wir *nicht* $dx = 0$ setzen? Hilfreich ist die Vorstellung der Fläche eines Quadrats mit Kantenlänge $x + dx$ wie in Abbildung 1.2.4. Der Flächeninhalt ist offenbar $(x + dx)^2 = x^2 + 2x dx + (dx)^2$ und wir machen folgende Beobachtung: Die beiden Streifen mit Flächeninhalt $x \cdot dx$ tragen offenbar noch merklich zur Fläche bei, aber das kleine Quadrat mit Flächeninhalt $(dx)^2$ ist sicher zu vernachlässigen.

Man spricht auch von **Kleinheit höherer Ordnung**, bezeichnet Größen wie $x \cdot dx$ als **Größen erster Ordnung**, Größen wie $dx \cdot dx$ als **Größen zweiter Ordnung**, usw., und sagt, dass Größen höherer Ordnung im Vergleich zu Größen erster Ordnung vernachlässigbar sind.

Damit wird nun aus unserer Gleichung (1.2) die Gleichung

$$dy = 2x_0 dx,$$

und die Division durch dx liefert die Steigung am Punkt x_0 ,

$$\frac{dy}{dx} = 2x_0.$$

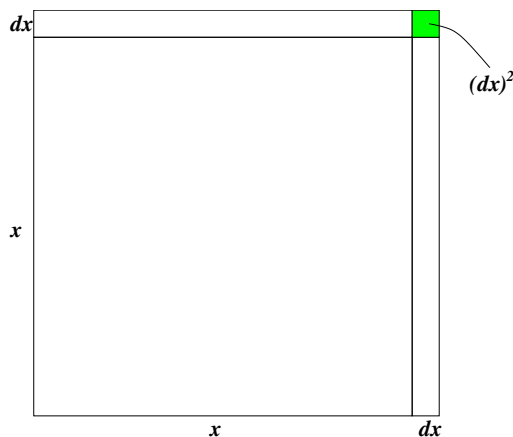


Abb. 1.2.4. Flächenbetrachtung am Quadrat

Am Punkt $x_0 = 2$ ist die Steigung der Funktion $y = x^2$ also $dy/dx(2) = 4$.

Auf diese Weise – und mit dem **Binomialtheorem** – können wir auch zeigen, dass die Steigung der Funktionen $y = x^n$, n irgendeine natürliche Zahl, immer $dy/dx = nx^{n-1}$ ist. Denn wir setzen an

$$y + dy = (x + dx)^n$$

und können nun auf den Klammerterm die Binomialentwicklung anwenden:

$$\begin{aligned} (x + dx)^n &= x^n + \frac{n}{1} x^{n-1} dx + \frac{n(n-1)}{1 \cdot 2} x^{n-2} (dx)^2 + \dots \\ &+ \dots \frac{n(n-1)(n-2) \cdots 2}{1 \cdot 2 \cdots (n-1)} x (dx)^{n-1} + \frac{n(n-1)(n-2) \cdots 2 \cdot 1}{1 \cdot 2 \cdots (n-1) \cdot n} (dx)^n. \end{aligned}$$

Diese Entwicklung kennen Sie sicher für $n = 2$, denn dann ist es der Binomische Lehrsatz $(a + b)^2 = a^2 + 2ab + b^2$. Da wir schon $(dx)^2 = 0$ gesetzt haben, können wir alle Terme mit höherer Potenz von dx wegstreichen, erhalten also

$$y + dy = x^n + nx^{n-1} dx$$

und wegen $y = x^n$ folgt der Differenzialquotient

$$\frac{dy}{dx} = nx^{n-1}.$$

Das Binomialtheorem ist nicht nur für natürliche Zahlen n als Exponenten gültig, das ist eine der wesentlichen Entdeckungen Newtons. Dann bricht die Summe zwar nicht mehr ab, sondern hat unendlich viele Summanden, aber das braucht uns nicht zu kümmern, da wir die hohen Potenzen von dx in jedem

Fall zu Null setzen. Wir können also auch Steigungen von Wurzelfunktionen $y = \sqrt[k]{x}$ berechnen, denn $\sqrt[k]{x} = x^{1/k}$:

$$y + dy = \sqrt[k]{x + dx} = (x + dx)^{1/k} = x^{1/k} + \frac{1}{k} x^{1/k-1} dx + \dots,$$

(in den Pünktchen sind die höheren Potenzen von dx versteckt), und damit folgt

$$y = x^{1/k} \quad \Rightarrow \quad \frac{dy}{dx} = \frac{1}{k} x^{1/k-1}.$$

Mit etwas mehr Einsatz gewinnt man die Differenzialquotienten weiterer Funktionen, die wir hier in einer Tabelle zusammenfassen:

$y = f(x)$	$\frac{dy}{dx}$
x^k	kx^{k-1}
e^x	e^x
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\ln x$	$\frac{1}{x}$

(1.3)

Mit den hier beschriebenen Methoden können wir leicht Regeln für Differenzialquotienten, wie die Produkt- und die Kettenregel, ableiten, und wir bedienen uns dazu wieder der heuristisch motivierten Methode in [Thompson 1998a] bzw. [Thompson 1998b].

1.2.1 Die Produktregel

Ist ein Produkt in der Form $y = f(x) \cdot g(x)$ gegeben, dann gilt

$$y + dy = (f + df) \cdot (g + dg) = fg + f dg + g df + df dg. \quad (1.4)$$

Nanu, warum schreiben wir denn jetzt nicht

$$y + dy = f(x + dx) \cdot g(x + dx)$$

wie wir es oben gemacht haben? Na ja, weil wir so einfach nicht weiterkommen¹! Wir haben hier ja keine konkreten Beispielfunktionen, sondern allgemeine Platzhalter $f(x)$ und $g(x)$. Aber $f(x + dx)$ schreiben wir einfach $f + df$, in der Vorstellung, dass der Funktionswert von f an der Stelle $x + dx$ derjenige von f an der Stelle x mit einer kleinen Korrektur df ist. Insofern ist es tatsächlich ganz gleich, ob wir $f(x + dx)$ oder $f + df$ schreiben, denn $f(x + dx) = f + df$.

¹Wenn Sie die Taylor'sche Formel kennen, dann wissen Sie: $f(x + dx) = f(x) + dx \cdot df(x)/dx + \dots$, wobei in den Pünktchen wieder die höheren Potenzen von dx stecken. Setzen wir die wieder zu Null, dann steht da tatsächlich $f(x + dx) = f(x) + df(x)$.

Blicken wir nun zurück auf (1.4). Offenbar ist $df \cdot dg$ eine Größe zweiter Ordnung und ist damit Null. Subtrahieren wir noch $y = fg$, dann folgt

$$dy = f dg + g df,$$

und wenn wir noch durch dx dividieren ergibt sich die **Produktregel**

$$y = f(x)g(x) \Rightarrow \frac{dy}{dx} = f \frac{dg}{dx} + g \frac{df}{dx}. \quad (1.5)$$

Wir können damit ein Produkt wie $y = \underbrace{e^x}_{=f(x)} \cdot \underbrace{(1+x^4)}_{=g(x)}$ einfach mit $df/dx = e^x$

und $dg/dx = 4x^3$ zu

$$\frac{dy}{dx} = f \frac{dg}{dx} + g \frac{df}{dx} = e^x \cdot 4x^3 + (1+x^4) \cdot e^x$$

berechnen.

1.2.2 Die Quotientenregel

Ist an Stelle eines Produktes ein Quotient

$$y = \frac{f(x)}{g(x)}$$

zweier Funktionen gegeben, dann führt

$$y + dy = \frac{f + df}{g + dg}$$

auf die folgende **Polynomdivision**²:

$$\begin{aligned} f + df : g + dg &= \frac{f}{g} + \frac{df}{g} - \frac{f}{g^2} dg \\ &\quad \underline{-(f + \frac{f}{g} dg)} \\ &\quad df - \frac{f}{g} dg \\ &\quad \underline{-(df + \frac{df dg}{g})} \\ &\quad -\frac{f}{g} dg - \frac{df dg}{g} \\ &\quad \underline{-\left(-\frac{f}{g} dg - \frac{f}{g^2} (dg)^2\right)} \\ &\quad -\frac{df dg}{g} + \frac{f}{g^2} (dg)^2. \end{aligned}$$

²Es macht nichts, wenn sie diese Rechnung noch nie gesehen haben. Merken Sie sich einfach, dass man Polynome ganz genau so dividieren kann wie zwei Zahlen.

Der jetzt entstandene Rest $-df dg/g + f(dg)^2/g^2$ ist aber Null, da es sich nur noch um Größen zweiter Ordnung handelt. Unsere Polynomdivision liefert also

$$y + dy = \frac{f}{g} + \frac{g df - f dg}{g^2}$$

und wenn wir jetzt $y = f/g$ subtrahieren und durch dx dividieren, dann haben wir die **Quotientenregel** hergeleitet:

$$y = \frac{f(x)}{g(x)} \Rightarrow \frac{dy}{dx} = \frac{g \frac{df}{dx} - f \frac{dg}{dx}}{g^2}. \quad (1.6)$$

Statt „Differenzialquotient“ verwendet man heute lieber die Bezeichnung **Ableitung** und schreibt

$$f'(x_0) = \frac{dy}{dx}(x_0)$$

für die Steigung der Funktion f am Punkt x_0 . Offenbar ist die Steigung von Funktionen an Maxima und Minima Null, so dass man durch Lösung der Gleichung

$$\frac{dy}{dx} = 0$$

die Punkte x finden kann, an denen Extremwerte zu vermuten sind.

Auch **höhere Ableitungen** spielen eine wichtige Rolle. In der Leibniz'schen Bezeichnungsweise ist

$$\frac{d^2y}{dx^2} := \frac{d}{dx} \left(\frac{dy}{dx} \right)$$

die Ableitung der Ableitung, usw. Zweite Ableitungen sind unter anderem nützlich, um an Extremstellen feststellen zu können, ob es sich um ein Maximum oder ein Minimum handelt. So lautet die erste Ableitung von $y = f(x) = x^2$:

$$f'(x) = 2x,$$

und die ist genau bei $x = 0$ Null. Also vermuten wir bei $x = 0$ eine Extremstelle. Die zweite Ableitung

$$f''(x) = 2$$

ist positiv, und das bedeutet, dass bei $x = 0$ ein Minimum vorliegt.

1.2.3 Die Kettenregel

Die wahre Überlegenheit der Leibniz'schen Bezeichnungsweise zeigt sich in der Kettenregel. Diese Ableitungsregel wird immer dann benötigt, wenn man verkettete Funktion wie

$$y = f(g(x))$$

zu differenzieren hat. Nennen wir den inneren Teil $u(x) := g(x)$ und den äußeren Teil $v(u) := f(u)$ und differenzieren wir beide Teile,

$$\frac{du}{dx} = g'(x), \quad \frac{dv}{du} = f'(u) = f'(g(x)),$$

dann folgt die gesamte Ableitung doch aus dem Produkt

$$\frac{dy}{dx} = \frac{dv}{du} = \frac{du}{dx} \cdot \frac{dv}{du},$$

denn das du lässt sich formal kürzen:

$$\frac{\cancel{du}}{dx} \cdot \frac{dv}{\cancel{du}} = \frac{dv}{dx},$$

also gilt

$$y = f(g(x)) \Rightarrow \frac{dy}{dx} = g'(x) \cdot f'(g(x)). \quad (1.7)$$

Diese Regel ist nicht auf zwei verkettete Formeln beschränkt, denn verwenden wir in der vierfachen Verkettung

$$y = f(g(h(k(x))))$$

die Abkürzungen

$$u(x) := k(x), \quad v(u) := h(u), \quad w(v) = g(v), \quad z(w) = f(w),$$

dann ist $y = z(x)$ und aus den einzelnen Differenzialquotienten

$$\frac{du}{dx} = k'(x), \quad \frac{dv}{du} = h'(u), \quad \frac{dw}{dv} = g'(v), \quad \frac{dz}{dw} = f'(w)$$

folgt durch Multiplikation und Kürzen

$$\frac{\cancel{du}}{dx} \cdot \frac{\cancel{dv}}{\cancel{du}} \cdot \frac{\cancel{dw}}{\cancel{dv}} \cdot \frac{dz}{\cancel{dw}} = \frac{dz}{dx},$$

also

$$\frac{dy}{dx} = \frac{dz}{dx} = k'(x) \cdot h'(k(x)) \cdot g'(h(k(x))) \cdot f'(g(h(k(x)))).$$

Betrachten wir als Beispiel die Funktion $y = \sin(e^{4x^2+x})$. Wir führen die Abkürzungen

$$u(x) = 4x^2 + x, \quad v(u) = e^u, \quad w(v) = \sin v$$

ein und differenzieren einzeln:

$$\frac{du}{dx} = 8x + 1, \quad \frac{dv}{du} = e^u, \quad \frac{dw}{dv} = \cos v.$$

Damit folgt

$$\begin{aligned}\frac{dy}{dx} &= \frac{dw}{dx} = \frac{du}{dx} \cdot \frac{dv}{du} \cdot \frac{dw}{dv} = (8x+1) \cdot e^u \cdot \cos v \\ &= (8x+1) \cdot e^{4x^2+x} \cdot \cos(e^{4x^2+x}).\end{aligned}$$

1.2.4 Die Umkehrregel

Ist $y = f(x)$ und existiert die Umkehrfunktion $x = f^{-1}(y)$, dann lautet die **Regel zur Ableitung der Umkehrfunktion**

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}. \quad (1.8)$$

Zur Illustration betrachten wir die Logarithmusfunktion

$$y = \ln x,$$

von der wir wissen, dass sie die Umkehrfunktion von $y = e^x$ ist. Durch Anwendung der e -Funktion finden wir aus $y = \ln x$ die Gleichung

$$x = e^y,$$

und daher

$$\frac{dx}{dy} = e^y = e^{\ln x} = x.$$

Aus der Regel zur Ableitung der Umkehrfunktion folgt dann

$$\frac{dy}{dx} = \frac{d \ln x}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{x}.$$

Wir haben damit gezeigt, dass $(\ln x)' = 1/x$ gilt.

1.3 Flächeninhalt, Integral und Antidifferenziation

Wir stellen uns nun vor, dass eine Strecke $[a, b]$ der Länge $x = b - a$ in unbegrenzt kleine Größen der Länge dx zerlegt wird. Wenn wir alle diese kleinen Stückchen summieren, muss natürlich die Länge x herauskommen. Wegen der unbegrenzten Kleinheit der dx ist deren Anzahl natürlich riesengroß, so dass es sinnvoll ist, für die Summe ein besonderes Symbol einzuführen. Mit Leibniz schreiben wir für diese Summe das **Integral**

$$x = \int_a^b dx = b - a,$$

wobei man bei diesem Zeichen an ein langgezogenes „S“ (für Summe) denken sollte. Die Grenzen a und b am Integral machen daraus ein **bestimmtes Integral**. Nehmen wir nun von der Länge x ein winziges kleines Stückchen, dann ergibt sich bei Anwendung von „ d “

$$dx = d \int dx = dx,$$

mit anderen Worten: **Ableitung und Integral sind zueinander inverse Operationen – Das Integral macht aus dx wieder x (d. h. neutralisiert das d) und das d bringt das Integral zum Verschwinden.**

Damit können wir Integrale als Umkehrung der Differenziation berechnen. Ist für $y = x^n$ die Ableitung nx^{n-1} , dann ist das Integral von $y = nx^{n-1}$ gerade

$$x^n = \int nx^{n-1} dx.$$

Nun ja, das ist nicht *ganz* richtig, denn nicht nur x^n ist das **unbestimmte Integral** (man sagt auch: **Stammfunktion**), sondern $x^n + \text{const.}$ Ist $f(x) = c$ eine beliebige Konstante (also f gar nicht abhängig von x), dann ist offenbar

$$y_0 + dy = f(x_0 + dx) = c$$

und wegen $y_0 = f(x_0) = c$ muss dc verschwinden. Damit gilt

$$d(x^n + c) = dx^n = d \int nx^{n-1} dx = nx^{n-1}.$$

Eine Stammfunktion ist also nicht eindeutig bestimmt sondern nur bis auf irgendwelche Konstanten, die man addieren kann.

Dieser simple Gesichtspunkt, dass wir Integrale als Umkehrung der Differenziation verstehen, die Integration also als **Antidifferenziation** bezeichnen, reicht uns schon! Eine eigenständige Bedeutung der Integration wird erst nach der Zeit von Leibniz und Newton Einzug in die Mathematik halten.

Nun aber zur geometrischen Interpretation. Was bedeutet $\int f(x) dx$ geometrisch?

In Abbildung 1.3.1 haben wir das Intervall $[a, b]$ auf der Abszisse in Stücke der Länge dx unterteilt. Im Bild sind nur vier Stücke zu sehen, aber wir stellen uns vor, es seien unendlich viele, weil dx unbegrenzt klein ist. Nehmen wir nun das Produkt $f(x) \cdot dx$, wobei auf jedem der Teilstücke x am linken Rand genommen wird (das spielt aber keine Rolle), so ist dieses Produkt gerade der grün dargestellte Flächeninhalt eines der Rechtecke. Summieren wir alle diese Rechteckflächen, dann erhalten wir den Flächeninhalt unter der Funktion f auf dem gesamten Intervall,

$$\int_a^b f(x) dx.$$

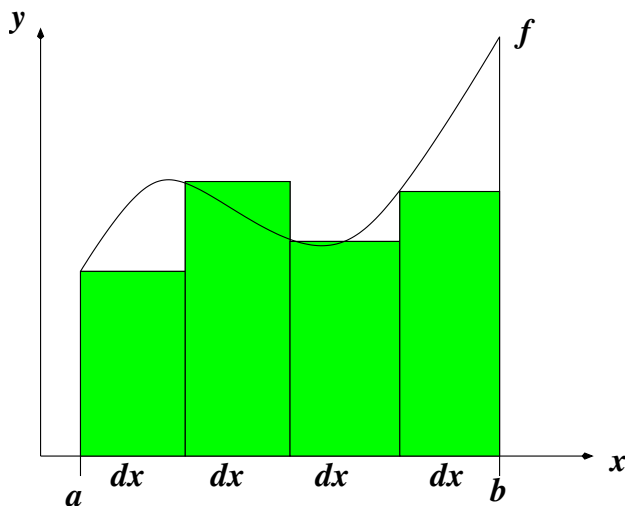


Abb. 1.3.1. Berechnung der Fläche unter einer Kurve

Ein Integral einer Funktion f über einem Intervall ist also nichts anderes als der Flächeninhalt unter dem Graphen dieser Funktion.

Wir haben schon auf den inversen Charakter von Integration und Differenziation hingewiesen. Diese wichtige Erkenntnis ist der Inhalt des **Hauptsatzes der Differenzial- und Integralrechnung**: Ist $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion, dann ist für alle $x_0 \in [a, b]$ die Integralfunktion

$$F(x) := \int_{x_0}^x f(t) dt \quad (1.9)$$

differenzierbar und es gilt $F'(x) = f(x)$ an jeder Stelle $x \in [a, b]$. Für den Wert des Integrals gilt

$$\int_a^b f(x) dx = F(b) - F(a). \quad (1.10)$$

1.4 Indivisible und Infinitesimale

Die Integration, also die Bestimmung des Flächeninhalts unter Kurven, ist ein guter Ausgangspunkt zur Klärung zweier Begriffe, die in der Geschichte der Mathematik eine bedeutende Rolle gespielt haben: Indivisible und Infinitesimale.

Die Streifen der Breite dx in Abbildung 1.3.1 nennt man **Infinitesimale**, hier ist es eine infinitesimale Fläche. Die Breite des Flächenstreifens ist dx ,

also eine unbegrenzt kleine, aber von Null verschiedene Größe. Zerlegt man das Intervall von Null bis Eins in sehr viele Teilintervalle der Länge dx , dann erhält man infinitesimale Strecken. Zerlegt man das Intervall in 10 Millionen Teile, dann ist $dx = 0.000\,000\,1$, wählt man 100 Millionen Teile, dann folgt $dx = 0.000\,000\,01$, und so weiter. Und hier sehen wir den entscheidenden Punkt: Je mehr Teilintervalle wir fordern, desto kleiner wird dx und schließlich – wenn wir immer mehr Teilintervalle betrachten – wird dx unbegrenzt klein, aber nie Null!

Die Infinitesimalen entsprechen der Auffassung des antiken Philosophen Aristoteles (384–322 v. Chr.), der in seinem Buch *Physik* [Aristoteles 1995] darlegte, dass eine Strecke nicht aus Punkten, sondern aus Infinitesimalen aufgebaut ist. Aristoteles bezeichnet die Strecke als **Kontinuum**, und propagiert die unbegrenzte Teilbarkeit. Ganz gleich, wie oft man ein Kontinuum teilt, immer werden die auch noch so klein werdenden Teile Kontinua bleiben.

Aber warum kann eine Strecke nicht aus Punkten aufgebaut sein? Das ist auch möglich und wurde von dem griechischen Philosophen Demokrit (460 oder 456 v. Chr. – um 380 v. Chr.) vertreten, der sich die Strecke aus unendlich vielen Atomen (atomos = unteilbar) zusammengesetzt dachte. Ein Punkt hat aber nach den Axiomen des Euklid (3. Jh. v. Chr.) keine Ausdehnung und ist damit eine **Indivisible** („Unteilbare“). Wollten wir die Fläche unter der Kurve in Abbildung 1.3.1 durch Indivisible beschreiben, müssten wir uns unendlich viele parallele Linien der Breite Null vorstellen, die die Fläche irgendwie „überdecken“, aber das ist ja unmöglich, da noch so viele breitenlose Linien nichts überdecken können. Die „Summe“ aller dieser unendlich vielen Indivisiblen ist dann die Fläche. Das ist schwer zu verdauen, aber Indivisiblenmathematiker wie Bonaventura Francesco Cavalieri (1598–1647) und Evangelista



Abb. 1.4.1. Demokrit (von griechischer Banknote), Marmor-Skulptur von Aristoteles (Kopie aus dem 1./2. Jh. n. Chr. von einer verlorenen Bronze-Skulptur, Musée de Louvre), Archimedes (von Briefmarke)

Torricelli (1608–1647) hatten mit dieser Idee schon Volumina von komplizierten Körpern berechnet und damit die Brauchbarkeit der Indivisiblenrechnung bewiesen, vergl. [Sonar 2011]. Das Wort „Summe“ ist ganz bewusst in Anführungszeichen gesetzt und seine Verwendung ist gefährlich! Cavalieri sprach nie von Summen, sondern von „omnes lineae“, also der Gesamtheit aller Linien, vergl. [Folkerts/Knobloch/Reich 2001, S. 299].

Am Beginn aller Analysis stand historisch die Indivisiblenmathematik und auch Leibniz ist als Indivisiblenmathematiker gestartet. Mit der Idee der Einführung eines dx , das so klein gedacht werden konnte wie nötig, ist er aber einer der Begründer der Infinitesimalmathematik geworden.

1.5 ... und wozu braucht man das?

Wir haben (hoffentlich!) klarmachen können, *wie* man Steigungen und Flächen von Funktionen berechnet, aber *warum* ist es so wichtig, so etwas zu können?

In den Zeiten von Leibniz und Newton war die Mechanik eine noch junge Wissenschaft. Eine Generation vorher hatte Galileo Galilei (1564–1642) die Fallgesetze hergeleitet. Unter Einwirkung der Erdbeschleunigung $g \approx 9.81 \text{ N/m}^2$ fällt ein Körper in der Zeit t die Strecke

$$y(t) = \frac{1}{2}gt^2.$$

Er erreicht dabei eine Geschwindigkeit (die Reibung in der Luft ist vernachlässigt) von

$$v(t) = gt$$

und die Beschleunigung im Schwerfeld der Erde ist natürlich $a(t) = g$, hängt also gar nicht von der Fallzeit ab. Mit der neuen Differenzial- und Integralrechnung konnte nun die Mechanik vollständig durchdrungen werden, ein Programm, das von Leonhard Euler (1707–1783) durchgeführt wurde [Szabó 1996]. Die Geschwindigkeit ist die Änderung der durchlaufenen Strecke, also

$$v(t) = y'(t) = \frac{dy}{dt}(t) = gt,$$

und die Beschleunigung ist die Änderung der Geschwindigkeit,

$$a(t) = v'(t) = \frac{dv}{dt}(t) = \frac{d^2y}{dt^2}(t) = g.$$



Abb. 1.5.1. Galileo Galilei ([Gemälde von Justus Sustermans, 1636] National Maritime Museum, Greenwich, London)

Anders ausgedrückt wissen wir damit, dass die Fläche unterhalb der Kurve v im Geschwindigkeits-Zeit-Diagramm gerade das Integral von v darstellt:

$$\int_0^t v(s) ds,$$

und diese Größe muss der zurückgelegte Weg sein, denn schließlich ist die Integration die Umkehrung der Differenziation! Es gilt also

$$y(t) = \int_0^t v(s) ds.$$

In seinem Buch *Discorsi e dimonstrazioni matematiche, intorno à due nuoue scienze attinenti alla meccanica e ai moti locali* (deutsch: „Unterredungen und mathematische Demonstrationen über zwei neue Wissenszweige, die Mechanik und die Fallgesetze betreffend“), veröffentlicht 1638 in Leiden, als Galileo schon in Italien im Griff der Inquisition war, finden wir folgendes Theorem [Galilei 1973, S. 158]:

„Die Zeit, in welcher irgend eine Strecke von einem Körper von der Ruhelage aus mittels einer gleichförmig beschleunigten Bewegung zurückgelegt wird, ist gleich der Zeit, in welcher dieselbe Strecke von demselben Körper zurückgelegt würde mittels einer gleichförmigen Bewegung, deren Geschwindigkeit gleich wäre dem halben Betrage des höchsten und letzten Geschwindigkeitswerthes bei jener ersten gleichförmig beschleunigten Bewegung.“

Abb. 1.5.2. Titelblatt der *Discorsi* 1638

Dieses Theorem, das Galileo durch geometrische Argumentationen beweist, ist nichts anderes als die Gleichheit der Integrale

$$y(t_E) = \int_0^{t_E} gt \, dt = \frac{v(t_E)}{2} \int_0^{t_E} dt,$$

wobei t_E eine Endzeit bezeichnet. Die Gleichheit der beiden Integrale drückt gerade die Flächengleichheit im Geschwindigkeits-Zeit-Diagramm aus, wie in Abbildung 1.5.3 gezeigt. Die gelbe und schraffierte Rechtecksfläche ist genau die gleiche Fläche wie die unter der Geraden $v(t) = gt$ von $t = 0$ bis $t = t_E$.

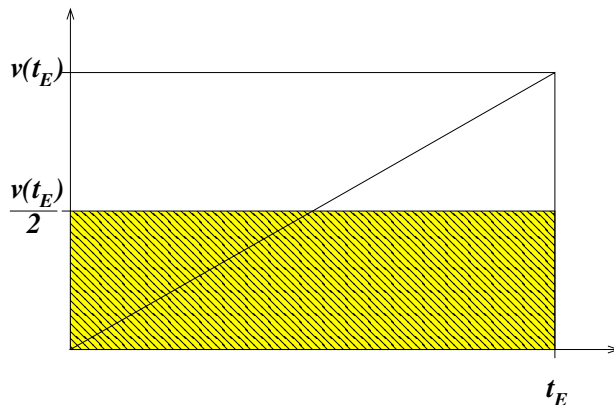


Abb. 1.5.3. Flächengleichheit im Geschwindigkeits-Zeit-Diagramm

Schon dieses einfache Beispiel lässt uns ahnen, dass man mit Hilfe der Differenzial- und Integralrechnung eine Wissenschaft wie die Mechanik vollständig beschreiben kann. Wichtige Prozesse in der Natur, in den Ingenieurwissenschaften und in der Finanz- und Wirtschaftsmathematik führen auf Differenzialgleichungen, d. h. Gleichungen, in denen die Ableitung einer Funktion auftaucht. Solche Gleichungen sind nur mit Hilfe der Differenzial- und Integralrechnung zu bearbeiten. So wundert es vielleicht auch nicht, dass Newtons und Leibnizens Differenzial- und Integralrechnung die Grundlage der technischen Revolution bildete, die im 17. Jahrhundert begann. Brücken, Schiffe, Motoren, Flugzeuge und selbst Smartphones und MP3-Player sind ohne diese Mathematik nicht denkbar. Die neue Mathematik von Newton und Leibniz führte innerhalb kürzester Zeit zu einer Revolution in nahezu allen Wissenschaften.

Mit anderen Worten: Es ging also wirklich um etwas! Hätte einer der Streithähne im Kampf um die Priorität der Entdeckung dieser Mathematik klein beigegeben, dann wäre sein Name bis heute befleckt vom Verdacht des Plagiats.

Die Geschichte des Prioritätsstreits zwischen Leibniz
and Newton

Geschichte – Kulturen – Menschen – Mit einem Nachwort
von Eberhard Knobloch

Sonar, Th.

2016, XXVII, 596 S. 250 Abb., Hardcover

ISBN: 978-3-662-48861-4