Anfang der Differentialrechnung. (Version 29.6.2025)

In der ersten Hälfte des 17. Jahrhunderts war von Fermat und später von Descartes die analytische Geometrie entwickelt worden. Damit konnten Kurven durch Gleichungen definiert werden. Beide Autoren haben für einige gleichungsdefinierte Kurven schon Tangenten bestimmen können. Aber erst die Differentialrechnung, sowohl in der von Leibniz wie in der von Newton entwickelten Form, verallgemeinerte das alte Konzept “Steigung über einem Intervall” oder “Sehnensteigung” in das neue Konzept “an die Kurve angepasste Steigung über einem Punkt” oder “Tangentensteigung”. Das war in der Geometrie ein sensationeller Erfolg noch ehe Newton die Differentialrechnung auf die Physik anwandte. Aber es war nur der Anfang der Differentialrechnung. Besonders wichtig wurde, dass man von Voraussetzungen über diese limes-definierten Ableitungen zurückschließen konnte auf Eigenschaften der differenzierten Funktion. Der Monotoniesatz besagt: *Eine Funktion mit positiver Ableitung ist streng wachsend.* Das heißt, die Ableitungen verhalten sich tatsächlich wie “Steigungen”.

Dasselbe noch einmal in einem physikalischen Kontext: Zurückgelegte Wege werden mit einer von der Zeit abhängigen Funktion $s(t)$ modelliert. Die experimentell zugänglichen Differenzenquotienten, also $(s(t_2) - s(t_1))/(t_2 - t_1)$, heißen in diesem Zusammenhang *Durchschnittsgeschwindigkeiten.* Mit den Grenzwerten dieser Durchschnittsgeschwindigkeiten werden Ableitungen $s'(t)$ der Beschreibung hinzugefügt. Wegen des Monotoniesatzes gilt wieder: Wenn diese Ableitungen positiv sind, wachsen die Wegfunktionen $s(t)$. Das bedeutet eben: die Ableitungen $s'(t)$ verhalten sich wie *momentane* Geschwindigkeiten $v(t)$. Und ebenso heißen deren Ableitungen (momentane) Beschleunigungen.

Warum ich finde, dass die heutigen Schulbücher der Analysis Unrecht tun:
Es wird wohl niemand bezweifeln, dass die Menschen schon vor 1600 Lineale an Kurven legen und ungefähre Tangenten zeichnen konnten. Das modern gewordene sogenannte “graphische Differenzieren” macht daher keinen Versuch, den Schülerinnen und Schülern zu erklären, welch enormen begrifflichen Fortschritt Leibniz und Newton erreicht haben.

Mein Taschenrechner berechnet den Differenzenquotienten $(\sin(x + h) - \sin(x))/h$ mit $h = 10^{-12}$ als 0. Da alle Taschenrechner notwendig mit endlicher Stellenzahl rechnen, bekommt man dies Ergebnis immer, wenn man ausreichend kleine Differenzen benutzt. Solche Rechnungen liefern also nur dann plausible Zahlen, wenn man sehr deutlich bei endlichen Differenzen bleibt. Sie erklären also auch nicht, inwiefern Leibniz und Newton etwas radikal Neues geschaffen haben.

Die größte Wirkung der neuen Differentialrechnung war, dass sie eine Formulierung der Gesetze der Mechanik ermöglichte. Newton’s ”Principia” waren der Beginn der Theoretischen Physik und im Gefolge davon, der auf die Gesetze der Mechanik gegründeten Industrialisierung. Die Physik steht heute zur Motivierung der Analysis auf der Schule nicht mehr zur Verfügung. Damit ist es schwierig, den Schülerinnen und Schülern verständlich zu machen, dass sie mit der Analysis eine geistige Revolution kennen lernen.

Vor Erfindung der analytische Geometrie, waren die meisten Kurven durch geometrische Konstruktionen definiert, die die Kurventangenten mitlieferten (Ellipsen, Parabeln, Hyperbeln, Rollkurven wie die Cardioide, durch Gelenkmechanismen definierte Kurven wie die Lemniskate, . . .). In der analytischen Geometrie konnten Kurven durch Formeln definiert werden, deren Tangenten die Differentialrechnung berechnete. Im Vergleich mit der theoretischen Mechanik und ihren technischen Konsequenzen ist die Berechenbarkeit von Tangenten – ohnehin nur der Anfang der Analysis – eine deutlich schwächere Motivation. Auf der Schule kommen parametrisierte Kurven $c(t) = (x(t), y(t))$ nicht vor, Kurven sind nur Graphen von Funktionen. Das schränkt die Möglichkeiten weiter ein. Also, kann man die Analysis überhaupt als Fortschritt vorführen?

Das Einfachste, was ich dazu kenne, ist die Brennspiegeleigenschaft der Parabel und ihre geometrische Konstruktion aus Brennpunkt und Leitlinie. Beides wird möglich, wenn man die Ableitung der quadratischen Parabel zur Verfügung hat. Das zeigt dann: Ableitungen ermöglichen Argumente und Einsichten, die man ohne diesen Begriff nicht haben kann. Und das muss im Analysisunterricht immer wieder vorkommen: *Ableitungen ermöglichen Argumente und Einsichten, die man ohne diesen Begriff nicht haben könnte.*

Die quadratische Parabel und der Nutzen ihrer Tangenten.

Dieser Abschnitt ist lang, weil es um die **erste** Tangente nach der Kreistangente geht.

1. Definition der Tangente durch Vergleich mit Kreisen.

Der Kreis ist die erste Kurve, bei der "Tangenten" (aus dem Lateinischen: Berührende) vorkommen. Eine Tangente ist definiert als die Gerade durch einen Kreispunkt, die dort senkrecht auf dem Radius ist. Wegen des Satzes des Pythagoras sind alle anderen Punkte der Tangente weiter vom Kreismittelpunkt entfernt als der Berührpunkt – der Kreis liegt auf einer Seite jeder seiner Tangenten. Ich formuliere eine weitere offensichtliche Eigenschaft, die sich sofort auf die Parabel übertragen läßt: Legt man einen Kreis so, dass er die x-Achse von oben berührt, so sind die rechtsseitigen Sehnensteigungen positiv, also größer als die Steigung 0 der Tangente und die linksseitigen Sehnensteigungen sind negativ.
Bei der **Normalparabel** $P(x) = x^2$ sind an der Stelle $x = a$

die rechtsseitigen (also $a < x$) Sehnensteigungen $(x^2 - a^2)/(x - a) = x + a > 2a$,
die linksseitigen (also $x < a$) Sehnensteigungen $(a^2 - x^2)/(a - x) = x + a < 2a$.

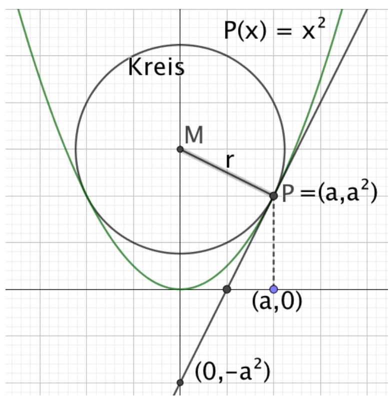
Wir erwarten wegen der Erfahrung am Kreis, dass die Tangentensteigung zwischen den rechtsseitigen und linksseitigen Sehnensteigungen liegen sollte, also $= 2a$ ist. Die Tangente ist dann die Gerade durch den Parabelpunkt (a, a^2) mit der Steigung $2a$, also

$$T_a(x) = 2a \cdot (x - a) + a^2 = 2a \cdot x - a^2.$$

Was kann man anführen, um zu rechtfertigen, diese Gerade "Tangente" zu nennen?
Die einfachste Eigenschaft ist: Die Parabel liegt oberhalb dieser Geraden, denn

$$P(x) - T_a(x) = x^2 - 2a \cdot x + a^2 = (x - a)^2 \geq 0.$$

Da das Wort Tangente "Berührende" bedeutet, wollen wir sehen, dass der Graph der Parabel an jeder Stelle zwischen einem Kreis und der gemeinsamen Tangente hindurchläuft, also mindestens so gut berührt wird wie der Kreis. Wir stellen uns vor, dass wir einen nicht zu kleinen Kreis von oben in die Parabel rutschen lassen, bis er stecken bleibt, weil er auf beiden Seiten die Parabel berührt. Sein Mittelpunkt liegt dann auf der y-Achse und der Radius zu einem Berührpunkt steht senkrecht auf der **Tangente**. Das erlaubt uns, diese Kreise zu finden: Die Gerade durch $P = (a, a^2)$ senkrecht zur Tangente heißt **Normale**.



Sie hat die Gleichung $$N_a(x) = \frac{-1}{2a} \cdot (x - a) + a^2 = \frac{-1}{2a} \cdot x + \frac{1}{2} + a^2.$$

Dieser Kreis hat also den Mittelpunkt $M = (0, N_a(0)) = (0, \frac{1}{2} + a^2)$ und der Abstand r zum Berührpunkt $P = (a, a^2)$ ist $r = \sqrt{1/4 + a^2}$. Die Kreisgleichung ist:

$$x^2 + (y - (1/2 + a^2))^2 = r^2 = |P - M|^2 = 1/4 + a^2.$$

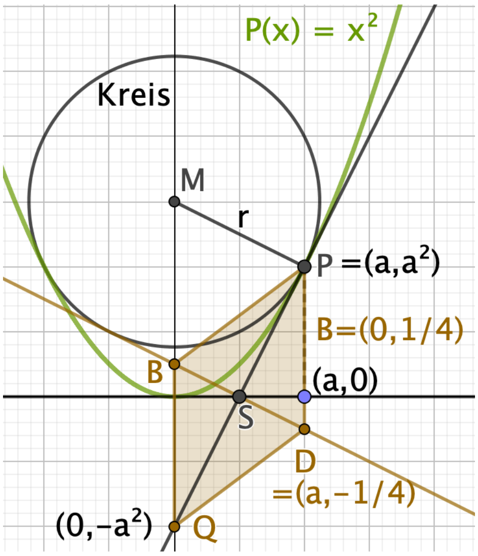
Um sicher zu sein, dass die Parabel wirklich zwischen Kreis und Kreistangente hindurch läuft, müssen wir noch einsehen, dass alle Punkte der Parabel, also alle (x, x^2), außerhalb dieses Kreises liegen. Wir müssen also zeigen

$$x^2 + \left(x^2 - (1/2 + a^2)\right)^2 \geq 1/4 + a^2$$
$$\text{oder: } x^2 + x^4 - x^2 \cdot (1 + 2a^2) + 1/4 + a^2 + a^4 \geq 1/4 + a^2$$
$$\text{oder offensichtlich: } x^4 - 2a^2 \cdot x^2 + a^4 = (x^2 - a^2)^2 \geq 0.$$

Solche einklemmenden Kreise sind *nicht eindeutig*: jeder kleinere Kreis mit derselben Tangente hat die Parabel zwischen sich und seiner Tangente. – *Verglichen mit dem, was wir vom Kreis her kennen, ist jetzt der Name* **Tangente der Parabel** *gerechtfertigt.*

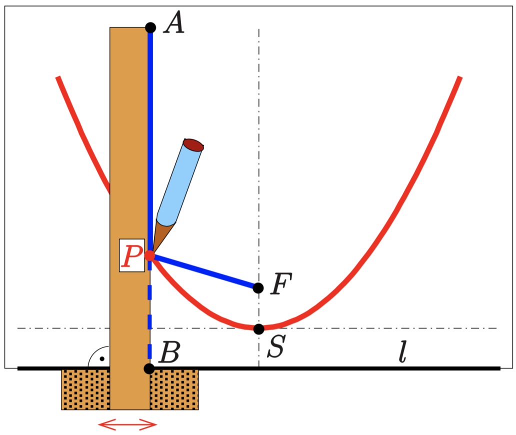
2. Folgerungen aus der Kenntnis der Parabeltangente.

2.1. Geometrische Definition der Parabel



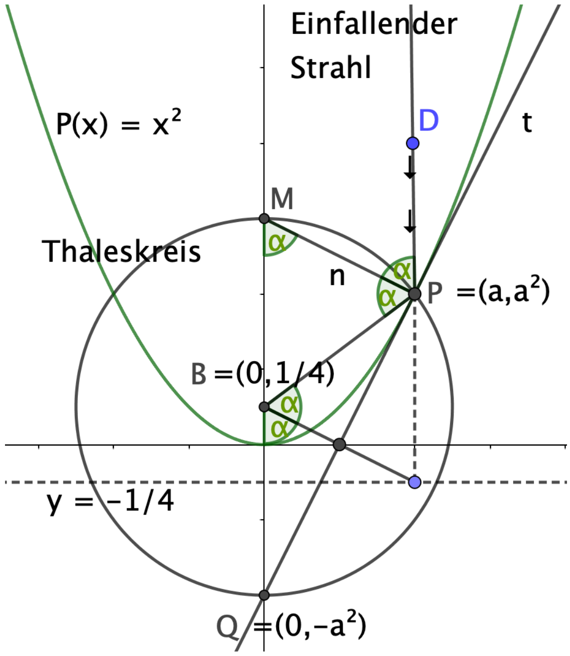
Die Tangente in $P = (a, a^2)$ schneidet die x-Achse in $S = (a/2, 0)$ und die y-Achse in $Q = (0, -a^2)$. Die vorhergehende Zeichnung wird durch die *Senkrechte zur Tangente* in S ergänzt. Da sie die Steigung $-1/2a$ hat, schneidet sie die y-Achse in $B = (0, 1/4)$ und die Vertikale $x = a$ in $D = (a, -1/4)$. Das Viereck $PBQD$ ist ein **Rhombus**, weil sich die Diagonalen halbieren und senkrecht auf einander stehen. Dieser aus der Tangente konstruierte Rhombus macht die geometrische Definition der Parabel kristallklar: es gilt $|PB| = |PD|$ oder *Die Punkte der Parabel sind gleich weit vom "Brennpunkt"* B *und der "Leitlinie"* $y = -1/4$ *entfernt.*
Die Kenntnis der Tangente der Parabel (= Graph von $P(x) = x^2$) liefert also deren geometrische Definition.

2.2. Fadenkonstruktion der Parabel



Fadenkonstruktionen von Ellipse und Hyperbel hatten schon die Griechen. Nach *Wikipedia* ist die abgebildete sehr schöne Fadenkonstruktion der Parabel von Jakob Steiner (1796-1863) gefunden worden. Die Bezeichnungen des Wikipediabildes: F ist der Brennpunkt (=Fokalpunkt) der Parabel, S ist der Parabelscheitel, die Strecke AB ist senkrecht auf der Leitlinie ℓ. Das braune Gerät wird mit der horizontalen Kante längs der Leitlinie verschoben und der in A und F befestigte Faden der Länge $|AB|$ wird mit dem Bleistift gespannt. Dadurch wird $|PB| = |PF|$ und P liegt auf der durch F und ℓ definierten Parabel.
Punkte T auf der *Winkelhalbierenden* des Winkels $\angle(FPB)$ sind gleich weit von B und F entfernt, aber für $T \neq P$ ist TB *nicht senkrecht* auf ℓ und daher T näher an ℓ als an F. Die Winkelhalbierende liegt also außerhalb der Parabel, sie ist deren Tangente.

2.3. Reflektion an der Parabel



Die "Satellitenschüsseln" genannten Empfangsantennen haben die Form von Flächen, die durch Rotation von Parabeln um ihre Achse entstehen. An blanken metallischen Flächen werden Lichtstrahlen und Radiowellen so reflektiert, dass an jeder Stelle der Einfallswinkel zur Normale gleich dem Ausfallswinkel ist. Setzt man diese physikalische Eigenschaft voraus, so folgt aus der Kenntnis der Parabeltangente von Neuem, was eben für die Parabel auf geometrische Weise bewiesen wurde:

Die Gerade MP ist senkrecht zur Tangente t, also $M = (0, \frac{1}{2} + a^2)$. Deshalb werden parallel zur Parabelachse ankommende Strahlen so reflektiert:

$\angle DPM = \alpha = \angle MPB$ (Reflektionsgesetz).

Alle übrigen mit α bezeichneten Winkel sind als Wechsel- oder Stufenwinkel gleich einem dieser beiden. Daher sind die Dreiecke PBM und PBQ gleichschenklig. Es folgt $|BM| = |BP| = |BQ|$.

Daher ist B Mittelpunkt des Kreises durch M, P und Q, also $B = (0, \frac{1}{4})$ für alle ankommenden Strahlen. Wegen des rechten Winkels bei P ist der Kreis ein Thaleskreis.

Der in der geometrischen Definition schon *Brennpunkt* genannte Punkte B hat also seinen Namen daher, dass parallel einfallende Lichtstrahlen zu diesem Punkt hin reflektiert werden. Die Kenntnis der Parabeltangente erklärt also die Form von Empfangsantennen mit dem Empfänger im Brennpunkt und ebenso die Form der Spiegel astronomischer Spiegelteleskope.

2.4 Die quadratische Parabel und der Anfang der theoretischen Physik.

Newton hat sich auch für die Geometrie von Kurven interessiert, aber sein Hauptinteresse galt der Physik. Und der berühmte Apfel fiel ebenfalls mit einem quadratischen Gesetz $s(t) = -\frac{g}{2} \cdot t^2$. Das Minuszeichen kommt daher, dass wir uns vertikale Koordinaten nach oben positiv vorstellen, der Apfel aber nach unten fällt. Eine physikalische Konstante ist notwendig, weil die Argumente Zeiten, die Werte aber Längen sind. Auf der Erdoberfläche ist diese Konstante $g = 9{,}81\,\mathrm{m/s}^2$. Dieser Zahlenwert wird im Folgenden ignoriert. Die rechtsseitigen ($t > t_0$) und linksseitigen ($t < t_0$) Differenzenquotienten $(s(t) - s(t_0))/(t - t_0)$ heißen jetzt *Durchschnittsgeschwindigkeiten*.

Die rechtsseitigen sind $\quad \dfrac{-g}{2} \cdot \dfrac{t^2 - t_0^2}{t - t_0} = -g \cdot \dfrac{t + t_0}{2} < -g \cdot t_0,$

die linksseitigen sind $\quad \dfrac{-g}{2} \cdot \dfrac{t_0^2 - t^2}{t_0 - t} = -g \cdot \dfrac{t + t_0}{2} > -g \cdot t_0.$

Es gibt nur *eine* Zahl, die zwischen die rechts- und linksseitigen Durchschnittsgeschwindigkeiten passt, sie heißt seit Newton *Momentangeschwindigkeit* $v(t_0) = -g \cdot t_0$. (Da das für jedes t_0 gilt, schreiben wir $v(t) = -g \cdot t$.)

Das sieht zunächst harmlos aus, aber die Naturgesetze sind so formuliert, dass die wirkenden Kräfte explizit die Momentangeschwindigkeiten verändern und erst über den damit bestimmten gesamten Bewegungsverlauf die Durchschnittsgeschwindigkeiten. Das erlaubte Newton die Formulierung seiner heute *drittes Newtonsches Gesetz* genannten grundlegenden Gleichung der Mechanik:

(wirkende) Kraft = Masse · (bewirkte) Beschleunigung.

Mit der in diesem Gesetz auftretenden "Beschleunigung" ist die *momentane* Änderung der *Momentan*geschwindigkeit gemeint. Ich kann mir nicht vorstellen, wie man dieses Gesetz mit den Durchschnittsgrößen formulieren könnte. Ich betrachte daher die Idee, den experimentell messbaren und begrifflich uralten Durchschnittsgeschwindigkeiten $(s(t)-s(t_0))/(t-t_0), t \neq t_0$ die **nur** im beschreibenden Modell definierbaren Momentangeschwindigkeiten $v(t) := s'(t)$ hinzuzufügen, als **phantastischen Fortschritt.**
Schon Galilei hatte an der schiefen Ebene ein quadratisches Bewegungsgesetz beobachtet. Die neue Differentialrechnung lieferte dazu die theoretischen Momentangeschwindigkeiten $s'(t) = \text{const}{\cdot}t + v_0$. Im Beispiel des fallenden Apfels sind alle Durchschnittsänderungen $(v(t) - v(t_0))/(t - t_0)$ der ebenfalls linearen Momentangeschwindigkeit $v(t) = \text{const} \cdot t + v_0$ konstant mit const $= -g$. Deshalb hat man in diesem Beispiel keine Probleme, auch die momentanen Beschleunigungen als konstant und gleich $-g$ anzusehen.

Soweit haben wir nur im physikalischen Kontext wiederholt, was im geometrischen Kontext erfolgreich war. Für die Physik ist das jedoch nicht genug, denn man möchte, im Beispiel, aus der konstanten Anziehungskraft auf das quadratische Bewegungsgesetz **schließen**. Dazu genügt zur Definition von Tangenten **nicht**, was ich bisher gesagt habe: Tangenten von Kurven sollen sich "so ähnlich" verhalten wie Kreistangenten. Und es genügt auch nicht, reelle Zahlen gegen einen Grenzwert "gehen" zu lassen. (In meiner Vorstellung "gehen" reelle Zahlen nicht, sondern jede hat ihren festen Platz in $\mathbb{R}$. Um einem Grenzwert näher zu kommen, muss man immer neue reelle Zahlen, evtl. mit großer Mühe, *finden*, die näher an dem diskutierten Grenzwert liegen.) Für das angestrebte Zurückschließen benötigt man eine Definition, die erlaubt, aus der Existenz eines Grenzwertes auf Eigenschaften zu schließen, die in seiner Definition benutzt wurden. Mit einer solchen Definition läßt sich dann der Monotoniesatz beweisen, also etwa diese Aussage: Wenn alle Tangenten einer Funktion nicht negative Steigung haben, dann ist die Funktion schwach wachsend: $x < y \;\Rightarrow\; f(x) \leq f(y)$. Daraus folgt: Wenn alle Tangenten die Steigung 0 haben, dann ist die Funktion weder wachsend noch fallend, also konstant. Stammfunktionen F einer gegebenen Funktion f sind definiert durch $F' = f$. Für zwei Stammfunktionen F_1, F_2 folgt $(F_2 - F_1)' = 0$, also (Monotoniesatz!) $(F_2 - F_1) = \text{const}$. Das liefert keine Berechnung von Stammfunktionen, aber sagt immerhin, dass diese durch den Wert $F(a)$ an einer Stelle a *eindeutig* bestimmt sind. – Da Geschwindigkeitsfunktionen als Ableitungen der Wegfunktionen **definiert** sind und ebenso Beschleunigungsfunktionen als Ableitungen der Geschwindigkeitsfunktionen, hat man: *Geschwindigkeitsfunktionen sind Stammfunktionen der Beschleunigungsfunktionen und Wegfunktionen sind Stammfunktionen der Geschwindigkeiten.* Beim Fallgesetz ist die Beschleunigung die konstante Funktion $a(t) = -g$, also gilt für den aus der Ruhe fallenden Apfel $v(t) = 0 - g \cdot t$. Der Fall des Apfels aus der Höhe H muss durch die Stammfunktion $s(t) = H - \frac{g}{2}t^2$ beschrieben werden ($s'(t) = v(t)$, $s(0) = H$). – Für diese Argumentation wird keine Integralrechnung benötigt. Für derartige physikalische Anwendungen sind Integrale eine Methode, Stammfunktionen, die man *nicht* kennt, zu konstruieren.

Probleme, die beim Aufbau der Differentialrechnung auftreten.

Eine Definition, mit der man aus Ableitungsvoraussetzungen Funktionseigenschaften folgern kann, muss mit Ungleichungen formuliert werden. Leider sind Ungleichungen sehr weitgehend aus dem Mathematikunterricht verschwunden, obwohl wir im täglichen Leben wirklich täglich die verschiedenartigsten konkreten oder abstrakten Dinge *der Größe nach* vergleichen. Um 1800, also zur Zeit von Gauß, war man noch weit von den heutigen Definitionen der Begriffe *konvergente Folge, Grenzwert, Tangente, Stetigkeit* entfernt. Während des 19. Jahrhunderts wurde intensiv daran gearbeitet. Bis 1900 hatten sich Definitionen dieser Begriffe durchgesetzt, die extrem verschieden von Formulierungen sind, die in der Umgangangssprache vorkommen. Das hat zu einem sehr festgelegten Aufbau der Analysis geführt. Ich glaube, fast alle in Deutschland an Universitäten in Mathematik ausgebildeten Personen haben die Analysis so kennen gelernt: *Zuerst werden konvergente Folgen und ihre Grenzwerte erklärt, dann folgt der Begriff stetig (oft zuerst als folgenstetig) und am Schluss wird die Definition der Differenzierbarkeit mit Hilfe des Grenzwertbegriffs oder der Stetigkeit formuliert.* Das ist so fest etabliert, dass ich noch niemanden getroffen habe, der es für ernsthaft erwähnenswert hält, dass die Differentialrechnung bereits 200 gloreiche Jahre hinter sich hatte, als schließlich Grenzwerte und Stetigkeit ihre heutige Formulierung erhielten.

Deshalb möchte ich hervorheben, dass in der oben besprochenen Ausdehnung des Begriffs Tangente von Kreistangenten auf Parabeltangenten Grenzwerte oder Stetigkeit nicht vorkommen. Dafür gibt es zwei Gründe. Erstens, die Potenzfunktionen sind so einfach, dass sich bei dem Vergleich der rechtsseitigen und linksseitigen Sehnensteigungen der Faktor $(x-a)$ im Nenner aus dem Zähler $(f(x)-f(a)) = (x^k - a^k) = (x-a)\cdot(x^{k-1}+\ldots+a^{k-1}))$ kürzen läßt, so dass man z.B. für $k=2$ sehen kann, dass nur die Zahl $2a$ als Steigung der Tangente in Frage kommt (weil $2a$ die einzige Zahl ist, die zwischen den linksseitigen und rechtsseitigen Sehnensteigungen liegt). Zweitens, bei dem Nachweis, dass die vermutete Gerade mit guten Gründen Tangente genannt werden kann, kommt *keine Division* durch $(x-a)$ vor, weil das Argument sich darauf konzentriert, wie wenig der Graph der Parabel von der vermuteten Tangente abweicht – nämlich weniger als ein passender Kreis von seiner Tangente abweicht.

Diese Argumentation läßt sich von $k=2$ auf alle Potenzfunktionen ausdehnen (und danach mit Differentiationsregeln auf alle rationalen Funktionen), aber ich kenne nur für die Parabel so schöne geometrische Eigenschaften, die unmittelbar zeigen, dass die definierten Tangenten nützlich sind. Ich finde, dass es zu wenig ist, das bloße Ausrechnen von Ableitungen als Hauptsache zu betrachten. Die Differentialrechnung hätte nicht ihr *enormes Prestige* erreicht, wenn sie nur Tangenten bestimmen und mit deren Hilfe Extremstellen finden und Wendetangenten berechnen könnte – ein Eindruck, den heutige Schulbücher erwecken.

Ehe ich zu Vorschlägen komme, muß ich auch die Schwierigkeiten ansprechen, die die endgültigen Definitionen vom Ende des 19. Jahrhunderts für die Schule verursachen. Für mich steht an erster Stelle die Tatsache, dass alle auf der Schule vorkommenden Funktionen schon Newton bekannt waren. Daher kann man keine Beispiele vorführen, für die die Definitionen so gemacht werden mußten, wie es im 19. Jahrhundert erarbeitet wurde. Ohne solche Beispiele schießt man aber mit Kanonen auf Spatzen. Das zweite Problem sind die vielen Misserfolge, die sich im Laufe der Zeit angesammelt haben. Von den Leuten, die Abitur gemacht haben, aber später nichts mehr mit Mathematik zu tun hatten, habe ich nur einen einzigen kennen gelernt, der bereit war, ein an den Analysisunterricht anknüpfendes Gespräch zu führen. Und niemand wird leugnen, dass auch noch viele Studienanfänger im Fach Mathematik erhebliche Schwierigkeiten mit der Stetigkeit haben.

Wie kann es nach der quadratischen Parabel weitergehen?

1. Übersicht über meinen Vorschlag.

Mein wichtigstes *Anliegen* ist, dass Schülerinnen und Schüler im Mathematikunterricht logische Argumentation kennen und verwenden lernen, Argumente also, die **nichts** benutzen dürfen, was nicht in Definitionen und Voraussetzungen zur Verfügung gestellt wird, und deren Folgerungen nur logische Schlussweisen benutzen.
Meine wichtigste *Vereinfachung* ist, statt der Grenzwerte von Sehnensteigungen die Approximation durch Tangenten ins Zentrum der Diskussion zu rücken – wie bei der quadratischen Parabel. Dadurch, dass man in der Definition auf die Division durch $(x-a)$ verzichtet, wird die Differentialrechnung der rationalen Funktionen unabhängig von Grenzwertdiskussionen. (Damit keine Missverständnisse entstehen: Grenzwerte sind unvermeidlich und von *unmittelbarem* Nutzen, überall wo irrationale Zahlen auftreten.) Das wird möglich durch die anschauliche Verabredung:

> (*Nicht umgangssprachliche Formulierung mit Ungleichungen weiter unten, Bild S.8)*
> Eine Funktion f hat an der Stelle $x=a$ die Tangente $T_a(x)=m\cdot(x-a)+f(a)$, wenn es zwei Kreise mit dieser gemeinsamen Tangente gibt, sodass sich der Graph von f in einem kleinen Intervall um a zwischen den beiden Kreisen hindurchzwängt.

Da ich den Nutzen von Tangenten heutzutage nicht mit Anwendungen aus der Physik vorführen kann (bzw. darf), muss ich mich beschränken auf die Fehlerkontrolle numerischer Verfahren zur Differentiation, Integration und zur Berechnung nicht-rationaler Funktionen wie

$$\sqrt{x},\ \exp,\ \sin.$$

Ein Bonus der Konzentration auf die Tangenten statt auf die Grenzwerte der Differenzenquotienten besteht darin, dass man folgende übersichtliche Formulierung für alle Differentiationsregeln hat:

> Setzt man eine kompliziertere Funktion h aus einfacheren Funktionen f,g zusammen (z.B. $h=\alpha\cdot f+\beta\cdot g,\ h=f\cdot g,\ h=f\circ g$), so erhält man die Tangenten von h, indem man die Tangenten von f und g ebenso zusammensetzt wie diese Funktionen. (Details weiter unten.)

Aus meiner Sicht sind die Differentiationsregeln nur Bequemlichkeitsregeln, weil alles Schwierige schon in den Voraussetzungen steht. Ich betrachte den Monotoniesatz und seine Folgerungen als die einzige wesentliche Satzgruppe aus den Anfängen der Differentialrechnung. Deshalb führe ich einen Beweis vor, der auf der Schule behandelt werden könnte.

2. Tangentendefinition wie bei der Parabel.

Es ist einfacher, Ungleichungen mit Parabeln zu handhaben als Ungleichungen mit Kreisen. Da wir für die Parabel schon gezeigt haben, dass sich zu jedem Punkt ein Kreis finden läßt, so dass der Graph der Parabel zwischen Kreis und Kreistangente verläuft, können wir in der folgenden Formulierung Parabeln statt Kreisen verwenden.

Tangentendefinition:

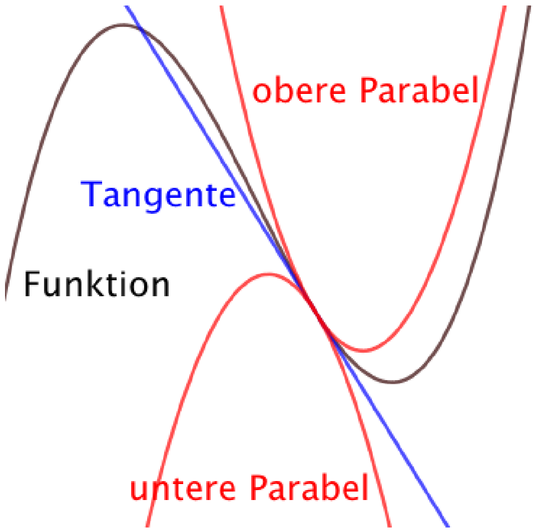
Die Funktion f hat an der Stelle a die **Tangente** $T_a(x)=m\cdot(x-a)+f(a)$, wenn es ein Intervall $[a-\delta,a+\delta]$ und eine Konstante B gibt, so dass folgende *Approximationsungleichung* gilt:

(TD)

$x\in[a-\delta,a+\delta]\Rightarrow |f(x)-T_a(x)|\le B\cdot|x-a|^2.$
ohne Beträge: $\Rightarrow -B\cdot(x-a)^2\le f(x)-T_a(x)\le B\cdot(x-a)^2.$
Bild nächste Seite.

Der Graph von f verläuft also im Intervall $[a-\delta, a+\delta]$ zwischen einer oberen und einer unteren Parabel mit der gemeinsamen Tangente T_a. Wir zeigen im nächsten Abschnitt, dass die Tangente von f bei a durch diese Definition eindeutig bestimmt ist. Das wird durch eine entsprechende Bezeichnung betont:

$$m = f'(a).$$



Funktionen, die mit der im 19. Jahrhundert akzeptierten Definition differenzierbar sind, brauchen nicht so einfach zu sein, dass ihr Graph von einer Tangente um weniger als eine quadratische Parabel abweicht. Beispiel: Die Funktion $f(x) = |x|^{\frac{3}{2}}$ ist bei $x = 0$ differenzierbar, aber es gibt kein $c > 0$, sodass $f(x) \leq c \cdot (x-0)^2$ gilt. – Funktionen, die nach der endgültigen Definition *zweimal* differenzierbar sind und für die im Intervall $[a-\delta, a+\delta]$ gilt $|f''| \leq 2B$, werden – als Folgerung aus dem Monotoniesatz – von ihrer Tangente bei a so gut approximiert, wie das in unserer schärferen *Ableitungsdefinition* verlangt wird.

3. Erste Konsequenz der Definition: Eindeutigkeit der Tangente.

In der Tangentendefinition kommt nicht explizit vor, dass die Tangente bei a *eindeutig* bestimmt ist. Wegen der Bilder, die die Schülerinnen und Schüler gesehen haben, werden sie wohl kaum Zweifel an der Eindeutigkeit haben. Obwohl man daher über die Wichtigkeit eines Beweises geteilter Meinung sein kann, füge ich einen *Beweis der Eindeutigkeit* an, denn Vertrautheit mit der Definition kann nur mit Beweisen erreicht werden, da die Anwendungen sich auf das Einüben der Differentiationsregeln konzentrieren.

Weil der Umgang mit Ungleichungen auf der Schule kaum vorkommt, schicke ich einige Bemerkungen voraus. (Ausführlicher in: Hilfestellung zum Umgang mit Ungleichungen.)

Es seien $a, b \in \mathbb{R}$ und $p, q \in \mathbb{R}_+$. Dann bedeuten folgende Ungleichungen dasselbe:

$$\begin{array}{ccccc} |a| \leq p & \Leftrightarrow & -p \leq a \leq p & \Leftrightarrow & -p \leq -a \leq p \\ |b| \leq q & \Leftrightarrow & -q \leq b \leq q & \Leftrightarrow & -q \leq -b \leq q \end{array}$$

Ungeübte können die Ungleichungsketten leichter addieren als mit den Beträgen argumentieren (zum Lesen sind die Beträge übersichtlicher), also

$$|a| \leq p \ \wedge\ |b| \leq q \quad \Rightarrow \quad -p-q \leq a \pm b \leq +p+q \quad \Leftrightarrow \quad |a \pm b| \leq p+q.$$

Beweis: Daraus wird nun gefolgert, dass *zwei* Tangenten an den Graphen einer Funktion f an der Stelle $(a, f(a))$ *übereinstimmen*. Wir setzen also für alle $x \in [a-\delta, a+\delta]$ die zwei Approximationsungleichungen für (möglicher Weise) zwei Tangenten voraus:

$$\begin{array}{c} |(f(x) - f(a)) - m_1 \cdot (x-a)| \leq B_1 \cdot (x-a)^2 \\ |(f(x) - f(a)) - m_2 \cdot (x-a)| \leq B_2 \cdot (x-a)^2 \end{array}$$

und wir schließen daraus für die Differenz der Terme in den Beträgen

$$|m_1 \cdot (x-a) - m_2 \cdot (x-a)| \leq (B_1 + B_2) \cdot (x-a)^2.$$

Diese Ungleichung kann für alle $x \neq a,\ x \in [a-\delta, a+\delta]$ durch $|x-a|$ dividiert werden:

$$0 < |x-a| \leq \delta \ \Rightarrow\ |m_1 - m_2| \leq (B_1 + B_2) \cdot |x-a|.$$

Wichtig ist, dass man hier nicht $x = a$ setzen darf, um $m_1 = m_2$ zu schließen, weil für diese Ungleichung ja $|x-a| > 0$ vorausgesetzt wird. Man muss damit durchkommen, dass man

$|x-a|$ durch Wahl von x so klein machen kann, wie man will. Dafür hat man das 2000 Jahre alte Eudoxos Axiom, das für reelle Zahlen $r \in \mathbb{R}$ z.B. so formuliert wird:

$$\forall_{n\in\mathbb{N}}\ 0 \le r \le 1/n \ \Rightarrow\ r = 0.$$

Wähle daher $\quad x := a + \dfrac{1}{n\cdot(B_1+B_2)}$, dann folgt $\ 0 \le |m_1 - m_2| \le \dfrac{1}{n}\ $ und $\ m_1 = m_2$.

Dies ist ein wichtiges Argument der Analysis, eine Vorstufe von Grenzwertargumenten, das schon Archimedes bei der Bestimmung des Flächeninhalts unter Parabeln benutzt hat.

4. Bausteine für Differentiationsregeln:
Direkt aus der Definition berechnete Beispiele.

Zweifellos werden viel mehr Ableitungen mit den Differentiationsregeln berechnet als direkt aus der Definition bestimmt. Damit dieser Prozess beginnen kann, müssen einige erste Ableitungen direkt aus der Definition hergeleitet werden.
Die höheren Potenzen werden mit der Produktregel differenziert, aber es hilft dem Umgang mit der Definition, nach der quadratischen Parabel zwei weitere Beispiele direkt zu differenzieren. Dazu verallgemeinern wir die dritte binomische Formel ein wenig:

$$\begin{aligned}
&(x-a)\cdot(x+a) = x^2 - a^2,\\
&(x-a)\cdot(x^2 + x\cdot a + a^2) = x^3 - a^3,\\
&(x-a)\cdot(x^3 + x^2\cdot a + x\cdot a^2 + a^2) = x^4 - a^4.
\end{aligned}$$

Im ersten Schritt geht es nur darum, durch Betrachten der rechtsseitigen und linksseitigen Sehnensteigungen eine *gute Vermutung* zu bekommen, wie groß die Tangentensteigung sein könnte. Daher kann ich, zur Vermeidung von Fallunterscheidungen, die Voraussetzung

$$0 < x, a \ \text{ machen.}$$

$$f(x) := x^3,\quad \frac{f(x)-f(a)}{x-a} = x^2 + x\cdot a + a^2 \begin{cases} \ge 3a^2 & \text{für } x > a\\ \le 3a^2 & \text{für } x < a\end{cases} \qquad f'(a) = 3a^2$$

$$f(x) := x^4,\quad \frac{f(x)-f(a)}{x-a} = x^3 + x^2\cdot a + x\cdot a^2 + a^3 \begin{cases} \ge 4a^3 & \text{für } x > a\\ \le 4a^3 & \text{für } x < a\end{cases} \quad f'(a) = 4a^3$$

Im zweiten Schritt muss die *Approximationsungleichung* bewiesen werden, ohne Einschränkung an die Vorzeichen von x und a. Die Potenzfunktionen sind so einfach, dass man $\delta = 1$ nehmen könnte. Ich benutze weiter δ, damit man den Einfluß der Länge des Intervalls sieht. Mit der vermuteten Steigung wird $T_a(x) := f'(a)\cdot(x-a) + f(a)$ definiert und die Differenz $f(x) - T_a(x)$ abgeschätzt. So wird die Vermutung bestätigt. Beachte dabei die oben aufgeschriebenen Faktorisierungen: $(x^k - a^k) = (x-a)\cdot(x^{k-1} + \ldots + a^{k-1})$, um Faktoren $(x-a)$ mehrmals auszuklammern.

$$\begin{aligned}
|x^3 - (3a^2\cdot(x-a) + a^3)| &= |(x^2 + x\cdot a + a^2) - 3a^2|\cdot|x-a| = |x+2a|\cdot|x-a|^2\\
&\le (3|a| + \delta)\cdot(x-a)^2 \qquad\qquad \text{also } B_{x^3} := 3|a| + \delta\\
|x^4 - (4a^3\cdot(x-a) + a^4)| &= |(x^3 + x^2\cdot a + x\cdot a^2 + a^3) - 4a^3|\cdot|x-a|\\
&= |(x^2 + x\cdot a + a^2) + (x+a)\cdot a + a^2|\cdot|x-a|^2\\
&\le (6a^2 + 4|a|\cdot\delta + \delta^2)\cdot(x-a)^2 \ \ \text{also } B_{x^4} := 6a^2 + 4|a|\cdot\delta + \delta^2
\end{aligned}$$

Wie gesagt, wegen der Produktregel braucht man diese Rechnungen nicht zu machen. Der direkte Zugriff auf die Tangentendefinition ist nur notwendig, wenn man etwas Neues macht, sonst genügen die Differentiationsregeln. Mit jeder Definition sollte aber geübt werden.

Inverse Funktion, Wurzelfunktion

Zwei weitere Funktionen, die ich direkt mit der Definition, also ohne Anwendung der Regeln, differenzieren würde, sind $f(x) = 1/x$ und $f(x) = \sqrt{x}$. Damit das Intervall $[a-\delta, a+\delta]$ nicht aus dem Definitionsbereich herausragt, wird $\delta \le |a|/2$ gewählt. Aufgeschrieben ist nur der Fall $a > 0$ und mit der Voraussetzung $x \in [a-\delta, a+\delta]$. Die *vermuteten* Tangentensteigungen folgen wieder aus der Betrachtung der rechts- und linksseitigen Sehnensteigungen. Für die Wurzelfunktion beachte: $(\sqrt{x} - \sqrt{a}) \cdot (\sqrt{x} + \sqrt{a}) = x - a$.

$$f(x) := 1/x, \quad \frac{f(x) - f(a)}{x - a} = \frac{-1}{a \cdot x} \quad \begin{cases} \ge -1/a^2 & \text{für } x > a \\ \le -1/a^2 & \text{für } x < a \end{cases} \qquad f'(a) = -1/a^2$$

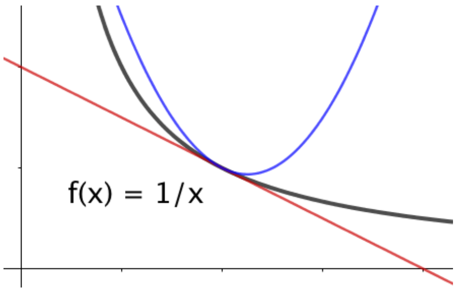
$$f(x) := \sqrt{x}, \quad \frac{f(x) - f(a)}{x - a} = \frac{1}{\sqrt{x} + \sqrt{a}} \quad \begin{cases} \le \frac{1}{2\sqrt{a}} & \text{für } x > a \\ \ge \frac{1}{2\sqrt{a}} & \text{für } x < a \end{cases} \qquad f'(a) = \frac{1}{2\sqrt{a}}$$

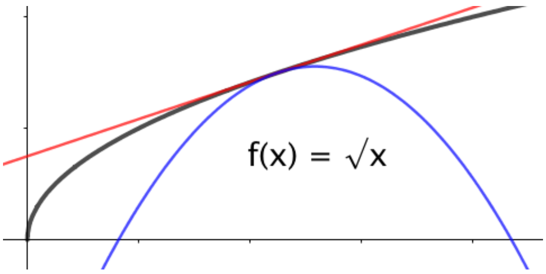
Zu *beweisen* ist wieder die *Approximationsungleichung* für die Differenz $f(x) - T_a(x)$ unter der Voraussetzung $x \in [a-\delta, a+\delta]$. Das Vorzeichen dieser Differenz zeigt der Term nach dem Gleichheitszeichen, daher bleibt der Graph in $[a-\delta, a+\delta]$ auf einer Seite der Tangente.

$$0 \le \frac{1}{x} - \left(\frac{-1}{a^2} \cdot (x-a) + \frac{1}{a}\right) = \frac{(x-a)^2}{a^2 x} \le \frac{1}{a^2(a-\delta)} \cdot (x-a)^2$$

$$0 \ge \sqrt{x} - \left(\frac{x-a}{2\sqrt{a}} + \sqrt{a}\right) = \frac{(\sqrt{a} - \sqrt{x}) \cdot (x-a)}{2\sqrt{a} \cdot (\sqrt{x} + \sqrt{a})} \ge \frac{-1}{8\sqrt{a}(a-\delta)} \cdot (x-a)^2$$

Wir lesen ab $B_{1/x} := 1/(a^2(a-\delta))$ und $B_{\sqrt{x}} := 1/(8\sqrt{a}(a-\delta))$. Meiner Meinung nach genügt das, um zu verstehen, wie die Definition auf Beispiele angewendet wird. Natürlich unterstützen Bilder das Verständnis:





Bemerkung 1. Die berechneten Tangenten von $f(x) = x^3$ schneiden den Graphen ein weiteres Mal bei $x = -2a$. Die vom Kreisbeispiel suggerierte Vorstellung, Tangenten ließen ihre Kurven auf einer Seite, ist also falsch. Auch eine Beschränkung auf die in der Definition vorkommenden kleinen Intervalle $[a-\delta, a+\delta]$ hilft nicht, denn die x-Achse sieht nicht nur wie die Tangente $T_0(x) = 0 \cdot (x-a) + 0^2$ des Graphen von $f(x) = x^3$ aus, sie ist sogar nach unserer schärferen Definition Tangente, denn $|x^3 - T_0(x)| = |(x-0)^3| \le \delta \cdot (x-0)^2$. Diese Tangente schneidet ihre Kurve im *Berührpunkt.* Man sieht daran, dass die allgemeine Charakterisierung von Tangenten durch *"lassen die Kurve auf einer Seite"* ebenso falsch ist wie die allgemeine Behauptung, eine Funktion wechsele an einer Nullstelle ihr Vorzeichen, auch wenn die Gegenbeispiele dafür (wie $f(x) = x^2$) klar in der Minderheit sind.

Bemerkung 2. Falls der Graph von g *zwischen* dem Graphen von f und dessen Tangente bei $x = a$ verläuft, so gilt die quadratische Abschätzung aus (TD) zwischen dem Graphen von f und dessen Tangente erst recht für g. Daher ist g bei $x = a$ differenzierbar und $g'(a) = f'(a)$.

Ausblick auf Transzendente Funktionen

Die transzendenten Funktionen werden meistens an einer späteren Stelle in einem Analysislehrgang behandelt. Für Leser, die die Analysis schon kennen, möchte ich als *Vorschau* erwähnen, dass auch die Ableitungen dieser Funktionen mit der verschärften Tanngentendefinition (TD) erhalten werden können. Dazu genügt nach Bemerkung 2 (Seite 10), dass die Graphen dieser Funktionen *zwischen* dem Graphen einer rationalen Funktion und dessen Tangente verlaufen.

Selbstverständlich benötigt man Grenzwerte, um die transzendenten Funktionen zu definieren. Ich finde auch nicht, dass Grenzwerte so schwierig zu erklären sind, dass deren Eliminierung aus der Schulmathematik gerechtfertigt ist. Aber, da die überall benutzten Taschenrechner meistens nur Approximationen ausgeben und auch viele Eingaben nur Approximationen sind $(1/3, \sqrt{2}, \pi, \exp(1), \ldots)$, ist es vordringlich, **Approximationen** zu erklären. Deshalb dehne ich die begonnene Erklärung, Ableitungen mit Approximationen statt mit Grenzwerten zu bestimmen, hier auf die transzendenten Funktionen aus.

Für die **Sinusfunktion** scheinen alle Autoren von der Ungleichung

$$0 \le x < \pi/2 \;\Rightarrow\; \sin x \le x \le \tan x = \sin x / \cos x \quad \text{umgestellt:} \quad x \cdot \cos x \le \sin x \le x$$

auszugehen, auch mein Text A6, S. 5. Zusammen mit $0 < a < 1 \;\Rightarrow\; a < \sqrt{a} < 1$ verwende ich das hier so:

$$0 \le x < \pi/2 \;\Rightarrow\; x \ge \sin x \ge x \cdot \cos x = x \cdot \sqrt{1 - \sin^2(x)} \ge x \cdot (1 - \sin^2(x)) \ge$$
$$\ge x - x^3, \quad \text{denn } \sin^2(x) \le x^2.$$

Hier ist der Graph von $\mathrm{tg}(x) := x$ die *schon bekannte* Tangente bei $x = 0$ der Unterfunktion $f_u(x) = x - x^3$ und der Graph von Sinus verläuft *zwischen* diesem Unterpolynom f_u und dessen Tangente (wegen $\sin(-x) = -\sin x$ im Intervall $-\pi/2 < x < \pi/2$).

Wegen Bemerkung 2, vorherige Seite, ist $\mathrm{tg}(x)$ auch die Tangente von $\sin(x)$ bei $x = 0$.

Die Ableitung an anderen Stellen folgt aus $\sin'(x + a)|_{x=0} = \sin'(0)\cos(a) + \cos'(0)\sin(a)$. Die Additionstheoreme und die Anfangsungleichung $\sin x \le x \le \tan x$ folgen aus der Definition am Einheitskreis, ausgeführt in meinem Text A6.

Die **natürliche Exponentialfunktion** wird meistens im Anschluss an die Ableitung der Exponentialfunktionen $f_a(x) := a^x$ definiert. Sie kann auch als Grenzwert von unteren und oberen Approximationen definiert werden (ausgeführt in meinem Text A2 Seite 3 - 4 mit Hilfe des multiplikativen Monotoniesatzes $f'/f \le g'g \Rightarrow g/f$ ist wachsend):

$$-n < x < n \;\Rightarrow\; \left(1 + \frac{x}{n}\right)^n \le \exp(x) \le \left(1 - \frac{x}{n}\right)^{-n},$$

insbesondere: $$-1 < x < 1 \;\Rightarrow\; (1 + x) \le \exp(x) \le (1 - x)^{-1}.$$

Auch hier ist $\mathrm{tg}(x) := 1+x$ Tangente der Oberfunktion $f_o(x) = (1-x)^{-1}$. Und der Graph von exp verläuft *zwischen* dieser Oberfunktion und ihrer Tangente bei $x = 0$. Man braucht zwar Grenzwerte um $\exp(x)$ zu definieren, aber man braucht keine Grenzwerte für die Ableitung der Grenzfunktion $\exp(x)$ bei $x = 0$, weil, wie in allen anderen Beispielen, der Graph dieser Grenzfunktion zwischen dem Graph einer *rationalen* Funktion und dessen Tangente verläuft.

Wegen Bemerkung 2, vorherige Seite, ist $\mathrm{tg}(x)$ auch die Tangente von $\exp(x)$ bei $x = 0$.

Auch für exp folgt die Differenzierbarkeit an Stellen $x \ne 0$ aus dem Additionstheorem, aus $e^{x+a} = e^x \cdot e^a$, welches direkt aus Eigenschaften der Approximationen gefolgert werden kann.

5. Differentiationsregeln

Differentiationsregeln haben den Sinn, solche Funktionen ohne Rückgriff auf die Tangentendefinition zu differenzieren, die aus "einfacheren" Funktionen – *deren Ableitungen bekannt sind* – zusammengesetzt sind. In Differentiationsregeln kommen also mehrere Funktionen vor und wir müssen unterscheiden können, welche Tangentengleichung zu welcher Funktion gehört. Deshalb verabreden wir eine etwas kompliziertere Bezeichnung:

Die Tangente des Graphen von f an der Stelle a ist T_a^f,
oder mit Argumenten: $T_a^f(x) := f'(a)\cdot(x-a)+f(a)$
und (TD) Seite 7: $[a-\delta^f, a+\delta^f] \;\Rightarrow\; |f(x)-T_a^f(x)| \le B^f\cdot(x-a)^2$.

Schließlich wird die Formulierung der Beweise übersichtlicher, wenn wir dem Unterschied zwischen Funktion und Tangente ebenfalls einen (etwas umfangreichen) Namen geben:

$U_a^f(x-a) := f(x)-T_a^f(x),\ |U_a^f(x-a)| \le B^f\cdot(x-a)^2$. Anders bezeichnet:
$h := (x-a),\ \ U_a^f(h) := f(a+h)-(f'(a)\cdot h+f(a)),\ |U_a^f(h)| \le B^f\cdot h^2$.

In den Voraussetzungen der Differentiationsregeln steht, dass die Approximationsungleichungen für die zum Zusammensetzen benutzten Funktionen erfüllt sind. Die Beweise bestehen nur aus der Kontrolle der Approximationsfehler mit der Dreiecksungleichung. Das gilt ebenso, wenn man später lernt, mit der endgültigen Differenzierbarkeitsdefinition zu arbeiten. Die folgenden Beweise zitieren also **nicht** frühere Ergebnisse sondern enthalten **alle** wesentlichen Details. Für alle Regeln gilt die übersichtliche Zusammenfassung: *Um die Tangenten der zusammengesetzten Funktion h zu erhalten, muss man mit den Tangenten der Bausteinfunktionen f, g dasselbe machen wie mit f und g.*

Vorausetzungen: Die Funktionen f, g seien an der Stelle a differenzierbar. Das heißt, es gibt $\delta^f, B^f, \delta^g, B^g$ so dass

$$x \in [a-\delta^f, a+\delta^f] \Rightarrow |f(x)-T_a^f(x)| = |U_a^f(x-a)| \le B^f\cdot(x-a)^2,$$
$$x \in [a-\delta^g, a+\delta^g] \Rightarrow |g(x)-T_a^g(x)| = |U_a^g(x-a)| \le B^g\cdot(x-a)^2.$$

Beide Ungleichungen gelten mit $\delta := \min(\delta^f,\ \delta^g) > 0$ im Intervall $[a-\delta, a+\delta]$.

Behauptung 1: Differenzieren von Linearkombinationen

Mit reellen Konstanten $\alpha, \beta \in \mathbb{R}$ definieren wir die **Linearkombination** $h := \alpha\cdot f+\beta\cdot g$. Dann ist h differenzierbar bei $x = a$ mit

$$h'(a) = \alpha\cdot f'(a)+\beta\cdot g'(a).$$

Beweis 1. Die beiden Gleichungen,

$$f(x) = T_a^f(x)+U_a^f(x-a) \quad\text{und}\quad g(x) = T_a^g(x)+U_a^g(x-a),$$

die in $[a-\delta, a+\delta]$ gelten, werden mit mit α bzw. β multipliziert und addiert. Das liefert h als die *Linearkombination der Tangenten* plus die Linearkombination der Fehler:

$$h(x) = \alpha\cdot T_a^f(x)+\beta\cdot T_a^g(x) \;\;+\;\; \alpha\cdot U_a^f(x-a)+\beta\cdot U_a^g(x-a).$$

Für die Linearkombination der Fehler liefert die Dreiecksungleichung ($|u+v| \le |u|+|v|$):

$$\left|\alpha\cdot U_a^f(x-a)+\beta\cdot U_a^g(x-a)\right| \le |\alpha\cdot B^f\cdot(x-a)^2|+|\beta\cdot B^g\cdot(x-a)^2| \le$$
$$\le \big(|\alpha|\cdot B^f+|\beta|\cdot B^g\big)\cdot(x-a)^2.$$

Mit der Definition $B^h := \big(|\alpha|\cdot B^f+|\beta|\cdot B^g\big)$ gilt dann in $[a-\delta, a+\delta]$

$$|h(x)-((\alpha\cdot f'(a)+\beta\cdot g'(a))\cdot(x-a)+h(a))| \le B^h\cdot(x-a)^2.$$

Das beweist die Differenzierbarkeit von h und $h'(a) = \alpha\cdot f'(a)+\beta\cdot g'(a)$.

Die Differentiationsregel für Linearkombinationen gilt also wegen der Dreiecksungleichung – angewandt auf die Fehler.

Behauptung 2: **Produktregel**
Das Produkt $h := f \cdot g$ ist bei $x = a$ differenzierbar und

$$h'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a).$$

Beweis 2. Die beiden Tangenten-Approximationen

$$f(x) = T_a^f(x) + U_a^f(x-a) \quad \text{und} \quad g(x) = T_a^g(x) + U_a^g(x-a)$$

die in $[a-\delta, a+\delta]$ gelten, werden multipliziert. Das liefert $h(x)$ als *Produkt der Tangenten* (als Hauptterm) plus drei Fehlerterme.
Der Beweis besteht – ohne Zitate anderer Ergebnisse – allein im Zusammenfügen aller schon in den Voraussetzungen stehenden Fehlerterme mit Hilfe der Dreiecksungleichung!
Multipliziert man zwei Zerlegungen in Haupteil plus "unwesentlichen" Fehlerteil, also

$$F_1 = H_1 + U_1, \; F_2 = H_2 + U_2,$$

so ergibt sich mit der Dreiecksungleichung $|u + v + w| \le |u| + |v| + |w|$:

$$\begin{aligned} |F_1 \cdot F_2 - H_1 \cdot H_2| &= |H_1 \cdot U_2 + H_2 \cdot U_1 + U_1 \cdot U_2| \\ &\le |H_1 \cdot U_2| + |H_2 \cdot U_1| + |U_1 \cdot U_2|. \end{aligned}$$

Die Fehlerterme sind *schon in der Voraussetzung* so abgeschätzt: $\le B^* \cdot (x-a)^2$.
Wenn man daher Schranken für $H_1,\ H_2$ finden kann, dann lassen sich die unwesentlichen Terme zusammenfassen mit einer Abschätzung $\le B_U \cdot (x-a)^2$.
$H_1,\ H_2$ sind die linearen Tangentenfunktionen, für sie finden wir mit der Dreiecksungleichung und $|x-a| \le \delta$ Schranken:

$$|H_1| = |T_a^f(x)| \le |f'(a)| \cdot \delta + |f(a)| =: M^f, \qquad |H_2| = |T_a^g(x)| \le |g'(a)| \cdot \delta + |g(a)| =: M^g,$$

mit denen wir die Fehlerterme so zusammenfassen können:

$$\begin{aligned} |H_1 \cdot U_2| + |H_2 \cdot U_1| + |U_1 \cdot U_2| &\le \big(B^f \cdot M^g + B^g \cdot M^f + B^f \cdot B^g \cdot \delta^2\big) \cdot (x-a)^2 \\ &=: B_U \cdot (x-a)^2 \end{aligned}$$

$H_1 \cdot H_2$ ist das Produkt der beiden linearen Tangentenfunktionen $T_a^f(x) \cdot T_a^g(x)$, damit

$$\begin{aligned} H_1 \cdot H_2 &= \big(f'(a) \cdot (x-a) + f(a)\big) \;\cdot\; \big(g'(a) \cdot (x-a) + g(a)\big) \\ &= f'(a) \cdot g'(a)(x-a)^2 + f'(a)(x-a) \cdot g(a) + f(a) \cdot g'(a)(x-a) + f(a) \cdot g(a) \\ &= \underline{(f'(a)g(a) + g'(a)f(a)) \cdot (x-a) + h(a)} \quad + \quad f'(a)g'(a)(x-a)^2. \end{aligned}$$

Der nicht unterstrichene quadratische Term gehört noch zu den Fehlern, er wird diesen hinzugefügt:

$$\Big|h(x) - \big((f'(a)g(a) + g'(a)f(a)) \cdot (x-a) + h(a)\big)\Big| \le (B_U + |f'(a)g'(a)|) \cdot (x-a)^2.$$

Diese quadratische Kontrolle der Differenz zwischen h und einer linearen Funktion zeigt nach Definition, dass diese lineare Funktion die Tangente von $h = f \cdot g$ bei a ist:

$$\boxed{T_a^h(x) = \; h'(a) \cdot (x-a) + h(a), \quad \text{mit} \quad h'(a) = f'(a)g(a) + g'(a)f(a).}$$

Das waren die beiden einfacheren Differentiationsregeln. Mit der **Kettenregel** ist es komplizierter, da ich in Schulbüchern der letzten 50 Jahre nur Beweise gefunden habe, die entweder den Zusatz enthalten, der Beweis gelte nur für streng monotone Funktionen, oder in denen ohne Zögern mit 0/0 erweitert wird, nämlich in der Form

$$\frac{u(v(x)) - u(v(a))}{v(x) - v(a)} \cdot \frac{v(x) - v(a)}{x - a}.$$

Deshalb möchte ich zuerst erklären, wie leicht man sich zweimal differenzierbare Funktionen verschaffen kann, deren Graphen viele horizontale Stücke enthalten, die also als innere Funktionen bei der zitierten Erweiterung *nicht vorkommen dürfen.*

> Die Funktion $f(x) := \sin(x) + x$ ist streng monoton wachsend, hat aber an allen ungeraden Vielfachen von π *horizontale* Wendetangenten. Man kann an jeder dieser Wendestellen den Graphen durchtrennen, etwas auseinander schieben und mit einer horizontalen Strecke wieder verbinden. Diese geänderte Kurve ist Graph einer zweimal differenzierbaren Funktion mit vielen Konstanzintervallen.

Einfache Kettenregeln. 1. Bei der Komposition linearer Funktionen $\ell_1(x) := m_1 \cdot x + b_1$ und $\ell_2(x) := m_2 \cdot x + b_2$ multiplizieren sich die Steigungen (= Ableitungen):

$$\ell_2(\ell_1(x)) = m_2 \cdot \ell_1(x) + b_2 = (m_2 \cdot m_1) \cdot x + (m_2 \cdot b_1 + b_2).$$

2. Wenn eine der beiden Funktionen linear ist, folgt die Kettenregel direkt durch Einsetzen aus der Approximationsungleichung (TD), Seite 7:

$$|x - a| \le \delta \;\Rightarrow\; \pm\big(\ell_2(f(x)) - \ell_2(T_a^f(x))\big) = \pm m_2 \cdot (f(x) - T_a^f(x)) \le (|m_2| \cdot B^f) \cdot |x - a|^2,$$

$$m_1 \cdot |x - a| \le \delta \;\Rightarrow\; \Big|f(\ell_1(x)) - T^f_{\ell_1(a)}(\ell_1(x))\Big| \le B^f \cdot |\ell_1(x) - \ell_1(a)|^2 \le (B^f \cdot m_1^2) \cdot |x - a|^2.$$

Die Kompositionen von ℓ_j und T^f sind die Tangenten der Kompositionen $\ell_2 \circ f,\; f \circ \ell_1$!!

Im Beweis der **allgemeinen Kettenregel** wird eine Eigenschaft der inneren Funktion benutzt, die ich als Hilfssatz vorab formulieren möchte.

Hilfssatz: Die Ableitung ist beinahe eine (lokale) Dehnungsschranke.

Genauer, aus der Approximationsungleichung der Tangente

$$x \in [a - \delta, a + \delta] \Rightarrow |f(x) - T_a^f(x)| \le B \cdot |x - a|^2$$

folgt mit der Dreiecksungleichung, dass f Abstände von der Stelle a höchstens um den Faktor $L := |f'(a)| + B \cdot \delta$ vergrößert, also

Behauptung: $\quad x \in [a - \delta, a + \delta] \;\Rightarrow\; |f(x) - f(a)| \le L \cdot |x - a|.$

Beweis. Wir benutzen folgende Variante der Dreiecksungleichung:

$$|u| = |(u - v) + v| \le |u - v| + |v|, \text{ also } |u - v| \ge |u| - |v|.$$

Dann ergibt sich aus der Approximationsungleichung der Tangente für $|x - a| \le \delta$

$$\begin{aligned} B \cdot |x - a|^2 \ge |f(x) - T_a^f(x)| &= |(f(x) - f(a)) - f'(a) \cdot (x - a)| \ge \\ &\ge |f(x) - f(a)| - |f'(a)| \cdot |x - a|, \end{aligned}$$

und damit sowie $|x - a|^2 \le \delta \cdot |x - a|$ schließlich die behauptete Dehnungsungleichung:

$$\boxed{|f(x) - f(a)| \le \big(|f'(a)| + B \cdot \delta\big) \cdot |x - a| = L \cdot |x - a|.}$$

Bei der Kettenregel haben innere und äußere Funktion unterschiedliche Definitionsbereiche. Um Verwechslungen möglichst auszuschließen, bezeichne ich die äußere Funktion und ihre Argumente mit großen Buchstaben.
Es geht jetzt um die Funktion $h := F \circ f$ oder $h(x) = F(f(x))$.
Voraussetzungen der Kettenregel: f ist differenzierbar bei a und F ist differenzierbar bei $A = f(a)$, also

$$x \in [a - \delta^f, a + \delta^g] \Rightarrow |f(x) - T_a^f(x)| \leq B^f \cdot (x - a)^2,$$
$$X \in [A - \delta^F, A + \delta^F] \Rightarrow |F(X) - T_A^F(X)| \leq B^F \cdot (X - A)^2.$$

Damit $f(x)$ in das Intervall $[A - \delta^F, A + \delta^F]$ fällt, muss mit Hilfe des Hilfssatzes weiter vorausgesetzt werden, dass δ^f klein genug ist, also notfalls verkleinert werden muss:

$$\delta^f \leq \delta^F / L \text{ und } x \in [a - \delta^f, a + \delta^f] \Rightarrow f(x) \in [A - \delta^F, A + \delta^F].$$

Diese Kontrolle der inneren Funktion ist ein wichtiger Teil des Beweises.

Behauptung: Die Funktion $h = F \circ f$ ist bei a differenzierbar mit

$$\boxed{h'(a) = F'(f(a)) \cdot f'(a).}$$

Für die Tangenten gilt $T_a^h(x) = T_A^F \circ T_a^f(x)$. Mit anderen Worten:
Die Tangente der Komposition $F \circ f$ ist die Komposition der Tangenten.

Beweis: Wie immer geht es darum, die Differenz zwischen $F \circ f$ und der erwarteten Tangente $T_a^h(x) = T_A^F \circ T_a^f(x)$ zu beherrschen.
Das legt folgenden Anfang nahe:

$$|F \circ f(x) - T_A^F \circ T_a^f(x)| = |F \circ f(x) - T_A^F \circ f(x) \ + \ T_A^F \circ f(x) - T_A^F \circ T_a^f(x)| \leq$$
$$\leq |F \circ f(x) - T_A^F \circ f(x)| \ + \ |T_A^F \circ f(x) - T_A^F \circ T_a^f(x)|.$$

und diese beiden Summanden müssen weiter behandelt werden.

Nach Voraussetzung über F und Einsatz des Hilfssatzes haben wir für den ersten Summanden (beachte $A = f(a)$):

$$|F \circ f(x) - T_A^F \circ f(x)| \leq B^F \cdot (f(x) - f(a))^2 \leq B^F \cdot L \cdot (x - a)^2.$$

Nach Voraussetzung über f haben wir für den zweiten Summanden (beachte, dass T_A^F die Steigung $F'(A)$ hat, also $T_A^F(Y) - T_A^F(Z) = F'(A) \cdot (Y - Z)$):

$$|T_A^F(f(x)) - T_A^F(T_a^f(x))| = \Big|F'(A) \cdot \big(f(x) - (f'(a)(x - a) + f(a))\big)\Big|$$
$$\leq |F'(A)| \cdot B^f \cdot (x - a)^2.$$

Beides zusammen beweist die Behauptung:

$$\boxed{\big|F \circ f(x) - T_A^F \circ T_a^f(x)\big| \leq \big(B^F \cdot L + |F'(A)| \cdot B^f\big) \cdot (x - a)^2.}$$

Ich wiederhole: Die Tangente der Komposition $F \circ f$ bei a ist die Komposition der Tangente von F bei $A = f(a)$ und der Tangente von f bei a. Das ist ein sehr schönes Ergebnis.

Außerdem: *Der Beweis beruht wirklich nur auf der Dreiecksungleichung.*

Damit ist gesagt, was zunächst zu Definition und Berechnung von Ableitungen gesagt werden muss. Später, wenn man (mit der Vollständigkeit!) Umkehrfunktionen, exp und sin behandelt, kommen weitere Differentiationsregeln hinzu. Aber wir müssen noch auf das für Physik und numerische Mathematik wichtige Problem zurückkommen:

Wie schließt man zurück von Informationen über die Ableitung auf Eigenschaften der Funktion?

In der Hochschulmathematik gehen solche Resultate von dem *Mittelwertsatz der Differentialrechnung* aus, der aber einen so aufwendigen Beweis hat, dass er nicht zu meiner Vorstellung von "Einführung in die Differentialrechnung" passt. (Ich erinnere daran, dass der Begriff "stetig" 200 Jahre jünger als die Differentialrechnung ist. Euler und Gauß kannten den Begriff "stetig" nicht.) Eine unmittelbare Folge des Mittelwertsatzes ist der Monotoniesatz:

Eine Funktion mit positiver Ableitung ist wachsend.

Ich habe in dem Text A5: *Loblied auf den Monotoniesatz* auf Seite 7 einen Beweis aufgeschrieben, der direkt von der ϵ-δ-Definition der Differenzierbarkeit ausgeht, natürlich die Vollständigkeit benötigt und wie in allen ähnlichen Fällen indirekt ist, aber *keine weiteren* Begriffe benutzt. Damit ist die Einführung in die Differentialrechnung beendet, die weitere Entwicklung wird von Anwendungen des Monotoniesatzes dominiert, die Arbeit mit Analysis beginnt.

Für Interessenten füge ich noch einen erheblich elementareren Beweis des Monotoniesatzes an, der aber das Reizwort "gleichmäßig" benutzt.
Da wir bis hier nur mit Funktionen zu tun haben, die schon mit der vorgeschlagenen schärferen Differenzierbarkeitsdefinition als differenzierbar erkannt werden können, kann man den Beweis aus Text A5 weiter vereinfachen. Ich gehe noch einen Schritt weiter. In vielen weit über die Schule hinausreichenden Situationen kann man die in der Definition auftretenden Konstanten δ^f, B^f in dem gerade betrachteten Intervall *unabhängig* von der Stelle a wählen (man sagt auch: *gleichmäßig wählen*). Dann kann man im Beweis sogar auf die Vollständigkeit verzichten und ihn als direkten Beweis organisieren.
Dass man einen Monotoniesatz ohne Vollständigkeit, also z.B. nur mit Kenntnis der rationalen Zahlen, beweisen kann, liegt nicht daran, dass wir nur eine Verschärfung der Differenzierbarkeitsdefinition betrachten, es geht mit der ϵ-δ-Definition ebenso, wenn man die entsprechenden Gleichmäßigkeitsvoraussetzungen macht. Deshalb denke ich, dass diese Version des Monotoniesatzes eine interessante Variante ist. Bei den Differenzierbarkeits**definitionen** spielt es keine entscheidende Rolle, ob man an die reellen oder nur an die rationalen Zahlen denkt, denn diese Definitionen sagen nur, wie man erkennt, dass eine *gegebene* Zahl m Ableitung $f'(a)$ von f bei a ist. Nur wenn zur Definition von f oder f' irrationale Zahlen nötig sind, kommt man nicht mit den rationalen Zahlen aus.

Das Wort "gleichmäßig" mag eine gewisse Abneigung auslösen, weil es zuerst in der Kombination "gleichmäßig stetig" auftritt und dort mit dem eher schwierigen Beweis verbunden ist, dass eine auf $[a, b]$ stetige Funktion dort sogar(!) gleichmäßig stetig ist. Meiner Meinung nach macht das Adjektiv "gleichmäßig" Begriffe der Analysis einfacher, nicht schwieriger. Und das in so großem Maße, dass die (oft geringe) Mehrarbeit zum Nachweis der Gleichmäßigkeit einer Voraussetzung mehr als gerechtfertigt ist. Zum Beispiel erlaubt eine so verbesserte Voraussetzung oft, einen indirekten Beweis in einen direkten zu verwandeln, wie im folgenden Beispiel.

Der rationale Monotoniesatz

Voraussetzung: Die Funktion f sei in dem Intervall $[\alpha, \omega]$ **gleichmäßig** differenzierbar. Das soll heißen: Es gibt Konstanten δ^f, B^f so dass *für alle* $a \in [\alpha, \omega]$ ein $m \in \mathbb{R}$ so existiert:

$$x \in [a - \delta^f, a + \delta^f] \cap [\alpha, \omega] \;\Rightarrow$$
$$-B^f \cdot (x-a)^2 \le f(x) - (m \cdot (x-a) + f(a)) \le B^f \cdot (x-a)^2$$

Hauptvoraussetzung: Für die Ableitung gilt: $f' \geq 0$.

Behauptung: $\boxed{x < y \in [\alpha, \omega] \;\Rightarrow\; f(x) \le f(y).}$

Beweis: Wir werden mehrfach benutzen, dass aus $f' \ge 0$ folgt

$$\xi < \eta \in [x, y] \text{ und } \eta - \xi < \delta^f \Rightarrow$$
$$\Rightarrow -B \cdot (\eta - \xi)^2 \le f(\eta) - f(\xi) - f'(\xi) \cdot (\eta - \xi) \le f(\eta) - f(\xi).$$

Wähle $n > (y-x)/\delta^f$ und teile das Intervall $[x, y]$ in n gleiche Teile, deren Länge $< \delta^f$ ist. Wende auf jedes Teilintervall die eben hergeleitete Ungleichung $-B \cdot (\eta - \xi)^2 \le f(\eta) - f(\xi)$ an. Diese n Ungleichungen werden summiert, wobei sich an jeder inneren Intervallgrenze die Funktionswerte wegheben, $f(linkes\ Intervallende) - f(rechter\ Intervallanfang) = 0$. Es bleibt:

$$n > (y-x)/\delta^f \;\Rightarrow\; n \cdot \big(-B \cdot (y-x)^2/n^2\big) = \tfrac{-B(y-x)^2}{n} \le f(y) - f(x).$$

Nun argumentieren wir wie Archimedes, dass diese Ungleichungen zur Folge haben

$$0 \le f(y) - f(x),$$

also die Behauptung des Monotoniesatzes.

Ich wiederhole: Die Voraussetzung der *gleichmäßigen* Differenzierbarkeit macht diesen Beweis möglich. Wenn man mit der ϵ-δ-Definition der Differenzierbarkeit arbeitet, muss man verlangen, dass zu jedem $\epsilon > 0$ das $\delta > 0$ *gleichmäßig* gewählt werden kann. Dann kann man diesen Beweis anpassen. Natürlich muss man sich mehr anstrengen, um die notwendigen Voraussetzungen als *gleichmäßig* nachzuweisen. Ob man das will oder nicht, ist Geschmacksache. Ich finde auf jeden Fall erwähnenswert, dass man bei gleichmäßigen Voraussetzungen bis hier ohne die Vollständigkeit auskommen kann. Dagegen ist die Vollständigkeit gänzlich unverzichtbar, um die transzendenten Funktionen exp, log, sin, cos, tan zu definieren. Selbst bei den Quadratwurzeln ist die Vollständigkeit das bequemere Hilfsmittel, obwohl man sich wegen $\sqrt{a} \cdot \sqrt{a} = a$ und $x^2 < a \Leftrightarrow |x| < \sqrt{a}$ ein gutes Stück weit mit symbolischem Rechnen bewegen kann.

Zusatz: Was ist mit den strikten Ungleichungen, gilt $f' > 0 \Rightarrow f(x) < f(y)$?

Im Archimedes Argument hilft die strenge Ungleichung in der Voraussetzung nicht. Aus $\forall_{n \in \mathbb{N}} \; -1/n < a$ folgt eben nur $0 \le a$. Daher kann man den obigen Beweis nicht auf strikte Ungleichungen ausdehnen. Man kann jedoch eine einfache Bemerkung über monotone Funktionen f benutzen:

Hat man $x < y$ mit $f(x) = f(y)$ so folgt aus der Monotonie

$$\xi \in (x, y) \Rightarrow f(x) \le f(\xi) \le f(y) = f(x).$$

Daher ist f konstant in $[x, y]$ und man hat für $\xi \in (x, y)$: $f'(\xi) = 0$.

Wenn also im Monotoniesatz statt $f' \ge 0$ vorausgesetzt wird $f' > 0$, so schließt man zuerst, dass f schwach wachsend ist und zeigt mit der letzten Bemerkung (indirekt!), dass $f(x) = f(y)$ unmöglich ist, also dass f streng wachsend sein muss.

Kommentar zur Vollständigkeit

Selbstverständlich kommen auf der Schule Quadratwurzeln früher vor als die Differentialrechnung. Deshalb haben Schülerinnen und Schüler gesehen, dass mit mehr Zahlen als den rationalen Zahlen gerechnet werden kann. In diesem Sinne sind sie der Vollständigkeit der reellen Zahlen schon begegnet, ehe sie mit Differentialrechnung beginnen. In dem Text Z2 habe ich auf den Seiten 7 bis 9 beschrieben, wie man mit der dritten binomischen Formel, also weit vor der Differentialrechnung, Quadratwurzeln berechnen kann (und mit der Ausdehnung dieser Formel auf höhere Potenzen auch n-te Wurzeln). Zum ersten Kontakt mit der Vollständigkeit habe ich in dem Text A3 *"Die Entwicklung der Grenzwertdefinition beginnt bei Archimedes"* geschrieben. Für die Vollständigkeit gilt: *Wenn man mit irrationalen Zahlen zu tun hat, braucht man sie – wenn man mit den rationalen Zahlen auskommt, braucht man sie nicht.* Da auch das Vollständigkeitsaxiom der reellen Zahlen aus der zweiten Hälfte des 19. Jahrhunderts stammt, habe ich auf den vorhergehenden Seiten beim Aufbau der Anfänge der Differentialrechnung betont, dass man diesen Anfang mit den rationalen Zahlen bewältigen kann. Klar, dass diese Betonung Geschmacksache ist, denn die Differentialrechnung wurde nicht für diese Anfänge geschaffen. Schon bei den nächsten Schritten – Umkehrfunktionen, Exponentialfunktionen, trigonometrische Funktionen – hängen die *Definitionen* unvermeidbar von der Vollständigkeit ab. Das steht nicht im Widerspruch dazu, dass frühe Tafeln trigonometrischer Funktionen viel älter und die Neperschen Logarithmentafeln immer noch 250 Jahre älter als die Vollständigkeit sind, denn diese Tafelwerke wurden auch zu ihrer Entstehungszeit als *Approximationen* aufgefasst. Und Approximationen dieser Funktionswerte sind die Ausgaben der Taschenrechner auch heute noch – anders geht es ja gar nicht.

Kommentar zu n-ten Wurzeln

Es ist nicht üblich, wie im Text Z2, Wurzeln mit der dritten binomischen Formel zu berechnen, sondern – als Anwendung bzw. zur Illustrierung der Differentialrechnung – mit dem Newton Verfahren. Das habe ich in dem Text A3 ab Seite 6 beschrieben. Man benötigt die n-ten Wurzeln zur *Definition* der Potenzen mit rationalen Exponenten:

$$a^{\frac{m}{n}} := \sqrt[n]{a^m}.$$

Ich finde daher, dass Berechnungsverfahren dieser (meistens) irrationalen Zahlen angesprochen werden sollten. Da die erwähnten Verfahren Approximationen von oben und von unten liefern und da man durch Potenzieren *überprüfen* kann, wie gut die erhaltenen Werte sind, kann man den Begriff *Approximation* betonen und diese Berechnungen behandeln, bevor Grenzwerte auftreten. Man muss dabei kein schlechtes (mathematisches) Gewissen haben, denn schon die Babylonier haben mit Approximationen gerechnet. Auch die Seefahrer, die mit den Tafeln der Logarithmen trigonometrischer Funktionen ihre Position auf hoher See bestimmten, haben sich auf Approximationen verlassen. Und für heutige Benutzer von Taschenrechnern genügt das Verständnis von Approximationen zur Erklärung der Ausgaben dieser Werkzeuge.

Kommentar zu Exponentialfunktionen

Die Exponentialfunktionen $f_a(x) := a^x$ gehören zu den am unvollständigsten behandelten Themen der Schulmathematik. Da Grenzwerte nicht mehr eingesetzt werden, fehlt schon eine Definition für irrationale x. Und selbst für rationale x wird das Potenzgesetz $a^{x+y} = a^x \cdot a^y$ zwar erwähnt, aber kaum begründet. Der Text A2 enthält zwei vollständige Behandlungen.

1. Auf den Seiten 8 - 12 wird der in Schulbüchern begonnene Weg ausgeführt: Beginn und Potenzgesetze mit *ganzen* Exponenten. *Rationale* Exponenten mit n-ten Wurzeln. *Irrationale* Exponenten mit Hilfe der Ungleichung

$$a \geq 1,\ h \in \mathbb{Q},\ 0 \leq h \leq 1 \quad \Rightarrow \quad a^h - a^0 \leq (a-1) \cdot h.$$

Aus dieser folgt: Wenn $\{x_n\}$ eine konvergente Folge *rationaler* Zahlen ist, dann ist auch $\{a^{x_n}\}$ eine konvergente Folge und man kann definieren: $a^{\lim x_n} := \lim a^{x_n}$ (Vollständigkeit von $\mathbb{R}$).
Die Potenzgesetze werden auf alle reellen Exponenten ausgedehnt und mit ihrer Hilfe die Ableitung der Exponentialfunktionen bestimmt. Die Funktion

$$\boxed{\begin{array}{l} g(x) := f_a(x/f_a'(0)) \ \text{ hat die Ableitung } \ g'(x) = g(x). \\ \text{Sie ist also die natürliche Exponentialfunktion } \ g(x) = \exp(x). \\ \text{Weiter folgt } \ \exp(f_a'(0)) = a \ \text{ also } \ln(a) = f_a'(0),\ f_a'(x) = \ln(a) \cdot f_a(x). \end{array}}$$

2. Auf den Seiten 3 - 7 wird ein Weg beschrieben, der die *Differentialrechnung für rationale Funktionen* einschließlich des für diese Funktionen behandelten Monotoniesatzes voraussetzt. (Vergleiche die ersten 17 Seiten dieses Textes.) Die rationalen(!) Funktionen

$$h_n(x) := \left(1 + \frac{x}{n}\right)^n \quad \text{und} \quad H_n(x) := 1/h_n(-x) \quad \text{mit} \quad x < n$$

bilden für jedes $x \in \mathbb{R}$ eine Intervallschachtelung $[h_n(x), H_n(x)]$, durch die mit der Vollständigkeit eine Grenzfunktion $h_\infty(x)$ definiert wird, die bei $x = 0$ die Ableitung $h_\infty'(0) = 1$ hat. Die Approximationen h_n, H_n sind gut genug, sodass die Potenzgesetze für h_∞ bewiesen werden können. Daraus folgt wie üblich $h_\infty' = h_\infty$. Man hat also von Neuem die natürliche Exponentialfunktion $\exp(x) = h_\infty(x)$ zusammen mit den numerischen Approximationen

$$\boxed{\left(1 + \frac{x}{n}\right)^n \leq \exp(x) \leq \left(1 - \frac{x}{n}\right)^{-n}.}$$

Und bessere Approximationen in den Experimenten 1 - 5 auf den Seiten 5 - 6 des Textes A2.

Kommentar zu den trigonometrischen Funktionen

In dem Text A6 erkläre ich, was man über sin und cos weiß bzw. nicht weiß, wenn man die Definition am Einheitskreis kennt. Mit Hilfe der Additionstheoreme und drei spezieller Dreiecke kann man sin und cos für alle ganzzahligen Vielfachen von 3° bestimmen. Weiter kann man Winkel halbieren und mit Quadratwurzeln deren sin- und cos-Werte ausrechnen. Damit kann man für alle Winkel $k \cdot 3°/2^n$ deren sin- und cos-Werte bestimmen. Weiter kommt man ohne Grenzwerte nicht.
Außerdem erlaubt die Definition am Einheitskreis zusammen mit den Additionstheoremen sin und cos zu differenzieren: $\sin' = \cos$, $\cos' = -\sin$. Danach genügt der Monotoniesatz, um wesentlich einfachere Berechnungsverfahren herzuleiten (Abbildung Seite 10 in A6).

Wenn man bis hierhin gekommen ist, hat man nach meiner Erfahrung genügend mit Ungleichungen geübt, um die endgültigen Definitionen von Grenzwert und Differenzierbarkeit gut benutzen zu können. Um zu zeigen, dass das Umgehen mit quadratischen Fehlern eine gute Vorbereitung auf die endgültigen Definitionen ist, führe ich den Beweis für die Differenzierbarkeit von Umkehrfunktionen zuerst noch einmal mit quadratischen Fehlern, danach die endgültige Version mit ϵ und δ.

Umkehrfunktionen

Umkehrfunktionen werden nur für streng monotone Funktionen definiert, wenn nötig nur auf Teilintervallen ihres Definitionsbereichs. Zu einer solchen Funktion $x \to f(x) =: y$ ist die Umkehrfunktion u, $y \to u(y)$, definiert durch

$$u(f(x)) = x \text{ oder gleichwertig } f(u(y)) = y.$$

Zeichnet man den Graphen von f auf eine durchsichtige Folie und sieht sie so von der Rückseite an, dass die y-Achse horizontal nach rechts zeigt, so sieht man den Graphen von u. Zeichnet man die Graphen von f und u in dasselbe Koordinatensystem, so liegen die beiden spiegelsymmetrisch zur Winkelhalbierenden des ersten Quadranten. Man erwartet daher, dass die Umkehrfunktion differenzierbar ist.
Die Umkehrfunktion der streng monotonen Funktion $f(x) := x^3$ ist nicht differenzierbar bei $y = 0$. Da wir an der Ableitung der Umkehrfunktion interessiert sind, verschärfen wir die Voraussetzungen noch etwas: Wir betrachten nur Funktionen f, deren *Ableitung nicht verschwindet* - nötigenfalls wieder durch Einschränkung auf ein Teilintervall des Definitionsbereichs. Außerdem kann $f' > 0$ angenommen werden, andernfalls betrachte $-f(x)$.

Differentiationsregel für Umkehrfunktionen

Bezeichnungen: Tangente $T_a^f(x) = f(a) + f'(a)(x-a)$, $y = f(x)$, $b = f(a)$, $u(y) = x$.
Voraussetzung 1: $f' > 0$ und $|x-a| \le \delta^f \Rightarrow |f(x) - T_a^f(x)| \le B(x-a)^2$.
Behauptung: $T_b^u(y) = u(b) + (1/f'(a))(y-b)$, also $u'(b) = 1/f'(a) = 1/f'(u(b))$.
Beweis:
Wir schreiben die Voraussetzung ohne Beträge und passen die Bezeichnungen schrittweise an die Umkehrfunktion an ($|x-a| \le \delta^f$ ist immer vorausgesetzt):

$$(1) \qquad -B(x-a)^2 \le f(x) - f(a) - f'(a)(x-a) \le B(x-a)^2 \quad \text{oder}$$
$$(2) \qquad -B(x-a)^2 \le y - b - f'(a)(u(y) - u(b)) \le B(x-a)^2 \quad \text{oder}$$
$$(3) \qquad -\frac{B}{f'(a)} \cdot (x-a)^2 \le \frac{1}{f'(a)}(y-b) - u(y) + u(b) \le \frac{B}{f'(a)} \cdot (x-a)^2.$$

Der mittlere Term in (3) ist schon die Differenz zwischen $u(y)$ und der behaupteten Tangente

$$T_b^u(y) = u(b) + (1/f'(a))(y-b).$$

Nun müssen zunächst die quadratischen Fehlerterme $(x-a)^2$ durch $(y-b)^2$ ersetzt werden, dann muss gezeigt werden, dass es genügt, statt $|x-a| \le \delta^f$ vorauszusetzen $|y-b| \le \delta^u$. Dazu machen wir nur die Hälfte der Umbenennungen von Zeile 1 zu Zeile 2 und ordnen die Terme etwas anders:

$$\big(f'(a) - B(x-a)\big) \cdot (x-a) \le y - b \le \big(f'(a) + B(x-a)\big) \cdot (x-a).$$

Die linke dieser Ungleichungen verspricht Erfolg. Wir müssen nur durch Verkleinern von δ^f

dafür sorgen, dass $B|x-a|$ nicht größer als $f'(a)/2$ wird. Wir verschärfen also die Voraussetzung durch

$$\delta^f := \min\left(\delta^f, \frac{f'(a)}{2B}\right).$$

Dann folgt $\frac{f'(a)}{2} \leq f'(a) - B(x-a)$, damit $|x-a| \leq \frac{2}{f'(a)} \cdot |u-b|$ und wir sind am Ziel:

Definiere $$\delta^u := \frac{f'(a)}{2} \cdot \delta^f$$

und erhalte die für die Differenzierbarkeit von u gewünschte quadratische Abschätzung:

$$(4)\quad |y-b| \leq \delta^u \;\Rightarrow\; |x-a| \leq \delta^f \;\Rightarrow\; \left|u(y) - \left(u(b) + \frac{1}{f'(a)}(y-b)\right)\right| \leq \frac{2B}{f'(a)^2}(y-b)^2.$$

Und nun die endgültige, die klassische Version:

Bezeichnungen: Tangente $T_a^f(x) = f(a) + f'(a)(x-a)$, $y = f(x)$, $b = f(a)$, $u(y) = x$.
Voraussetzung 2: $f' > 0$ und $\forall_{\epsilon^f>0}\exists_{\delta^f>0}$ sodass $|f(x) - T_a^f(x)| \leq \epsilon^f|x-a|$.
Behauptung: $T_b^u(y) = u(b) + (1/f'(a))(y-b)$, also $u'(b) = 1/f'(a) = 1/f'(u(b))$.
Beweis:
Wir schreiben die Voraussetzung ohne Beträge und passen die Bezeichnungen schrittweise an die Umkehrfunktion an ($|x-a| \leq \delta^f$ ist immer vorausgesetzt):

(1) $\quad -\epsilon^f \cdot |x-a| \leq f(x) - f(a) - f'(a)(x-a) \leq \epsilon^f \cdot |x-a|$ oder

(2) $\quad -\epsilon^f \cdot |x-a| \leq y - b - f'(a)(u(y) - u(b)) \leq \epsilon^f \cdot |x-a|$ oder

(3) $\quad -\dfrac{\epsilon^f}{f'(a)} \cdot |x-a| \leq \dfrac{1}{f'(a)}(y-b) - u(y) + u(b) \leq \dfrac{\epsilon^f}{f'(a)} \cdot |x-a|.$

Der mittlere Term in (3) ist schon die Differenz zwischen $u(y)$ und der behaupteten Tangente

$$T_b^u(y) = u(b) + (1/f'(a))(y-b).$$

Nun müssen erstens die Fehlerterme $\epsilon^f \cdot |x-a|$ durch $\epsilon^u \cdot |y-b|$ ersetzt werden und zweitens muss gezeigt werden, dass es genügt, statt $|x-a| \leq \delta^f$ vorauszusetzen $|y-b| \leq \delta^u$ – und natürlich zu *jedem* ϵ^u.
Dazu machen wir nur die Hälfte der Umbenennungen von Zeile 1 zu Zeile 2 und ordnen die Terme etwas anders (beachte $x > a \;\Leftrightarrow\; u > b$):

(2a) $\quad (f'(a) - \epsilon^f) \cdot |x-a| \leq |y-b| \leq (f'(a) + \epsilon^f) \cdot |x-a|.$

Die linke dieser Ungleichungen verspricht Erfolg. Wir müssen nur von vornherein $\epsilon^f < \frac{f'(a)}{2}$ wählen und nur zu solchen ϵ^f ein δ^f bestimmen. Dann haben wir immer $|x-a| \leq \frac{2}{f'(a)} \cdot |y-b|$ und können damit in (3) $|x-a|$ durch $\frac{2}{f'(a)} \cdot |y-b|$ ersetzen.

Definiere schließlich $\quad \delta^u := \dfrac{f'(a)}{2} \cdot \delta^f$, dann gilt $|y-b| \leq \delta^u \;\Rightarrow\; |x-a| \leq \delta^f$.

Damit folgt aus (3) zu gegebenem $\epsilon^u > 0$ mit der Wahl $\epsilon^f := \frac{f'(a)^2}{2}\epsilon^u$ und dazu δ^f, dann δ^u:

$$|y-b| \leq \delta^u \;\Rightarrow -\epsilon^u \cdot |y-b| \leq \frac{1}{f'(a)}(y-b) - u(y) + u(b) \leq \epsilon^u \cdot |y-b|.$$

qed