Die Erfindung der Ableitung war ein sensationeller Erfolg

Zusammenfassung

Die Erklärung der Experten können Schülerinnen und Schüler, die gerade die ersten Schritte in Analysis hinter sich haben, noch nicht verstehen. Ich möchte eine Antwort auf Schulniveau versuchen.

Dazu stelle ich auf den ersten vier Seiten zusammen, welche auf der Schule vorkommenden und mit Analysis zusammenhängenden Themen die Mathematiker schon vor Erfindung der Differentialrechnung beherrschten:
Für die Monotoniediskussion der Potenzfunktionen braucht man nicht deren Ableitung. Steigung als quantitatives Maß für Wachstum ist zunächst nur für Geraden definiert. Ich weiß von einem Schülervorschlag, die Steigung eines nichtlinearen Graphen als Steigung der Tangente zu definieren. Aber die Schülerinnen und Schüler kennen vor der Differentialrechnung nur die Definition der Kreistangente. Deshalb zeige ich (Seiten 3-5), wie man in einfachen Fällen Geraden findet, die den Funktionsgraphen erstens treffen und zweitens ihn auf einer Seite lassen. In diesen Fällen kann man also die Schülerantwort benutzen. Aber, diese punktweisen Steigungen kann man nicht so zusammensetzen, wie über Intervallen definierte Steigungen (Seite 6). Insbesondere ist sehr einfach ein Monotoniesatz zu beweisen, der für über Intervallen gegebene Steigungen sagt, was später der Monotoniesatz der Differentialrechnung (erwähnt und formuliert auf Seite 8) für die punktweisen Steigungen behauptet. Ich erkläre, warum die beschriebene *einseitige Berührung* keine Tangentendefinition ist und was man genauer über Tangenten wissen muss, um den Monotoniesatz beweisen zu können.

Was kann man vor dem Differenzieren schon definieren oder beantworten?

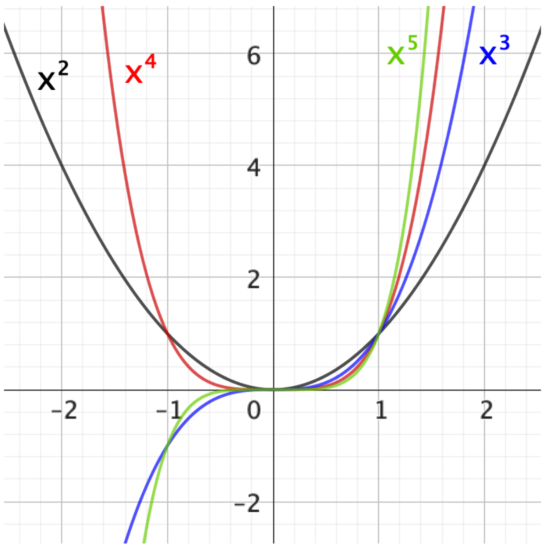
Monotonie von Funktionen in einem Intervall

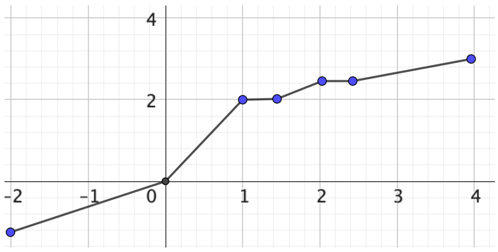
Eine Funktion f heißt in dem Intervall $[a, b]$ wachsend oder steigend oder monoton steigend, wenn gilt: $$x, y \in [a, b],\ x < y \ \Rightarrow\ f(x) \le f(y).$$

Fügt man das Adjektiv "streng" hinzu, also streng wachsend oder streng monoton steigend, so bedeutet das:
$$x, y \in [a, b],\ x < y \ \Rightarrow\ f(x) < f(y).$$

Mit fallend oder monoton fallend sind die umgekehrten Ungleichungen zwischen $f(x)$ und $f(y)$ gemeint:
$$x, y \in [a, b],\ x < y \ \Rightarrow\ f(x) \ge f(y), \text{ bzw. mit "streng": } f(x) > f(y).$$

Bemerkung: In der Mathematik wird "monoton" etwas anders verwendet:
Eine monotone Funktion ist in dem ganzen Intervall *wachsend* **oder** in dem ganzen Intervall *fallend*. *Monoton fallend* und *fallend* sagt einfach dasselbe. Allerdings hat man eine Kollision mit der Umgangssprache: Die konstanten Funktionen sind beides: *fallend* und *wachsend*. Deshalb bevorzuge ich analog zu "streng" den Zusatz "schwach" bei den schwachen Ungleichungen $\le$, $\ge$.





Graph einer in $[-2, 4]$ (schwach) wachsenden Funktion.
Funktionen, die monoton aber nicht streng monoton sind, sind auf gewissen Teilintervallen konstant.

Die Potenzen $f(x) := x^n,\ n = 1, 2, 3, \ldots$ sind in $[0, \infty)$ streng wachsend, denn aus $0 \le x < y$ folgt $x^n < y^n$.
Die geraden Potenzen sind symmetrisch zur y-Achse, $f(-x) = +f(x)$, also sind sie streng fallend in $(-\infty, 0]$.
Sie haben bei 0 das *Minimum* $f(0) = 0$, denn $x \ne 0 \ \Rightarrow\ x^{2n} > 0$.
Die ungeraden Potenzen sind punktsymmetrisch zu $(0|0)$, $f(-x) = -f(x)$, also sind sie auch in $(-\infty, 0]$ streng wachsend und damit in ganz $\mathbb{R}$ streng wachsend.
Diese Beispiele kann man auch noch in Richtung der x-Achse verschieben, betrachte $f(x) := (x - a)^n$ und diskutiere ohne Kenntnis von Ableitungen wie eben.

Steigungen, also quantitatives Wachstum

Für lineare Funktionen $f(x) = m \cdot x + b$ kann man Steigungen definieren. Zu je zwei reellen Zahlen $x_1 < x_2 \in \mathbb{R}$ definiert man: *Steigung* $:= (f(x_2) - f(x_1))/(x_2 - x_1)$. In diesem Fall findet man immer denselben Wert: *Steigung* $= m$. Diese Steigungen sind ein *quantitatives Maß* für Wachstum: Je größer m, um so stärker wächst f. Wenn man versucht, denselben Quotienten bei anderen Funktionen, z.B. Potenzen mit $n \geq 2$, zu benutzen, so berechnet man wegen $x_1 < x_2$ offensichtlich die Steigungen von Sekanten. Man bekommt nicht eine Zahl, die man die Steigung des Graphen von f an einer bestimmten Stelle nennen könnte. Ein Schülerkommentar schlug vor, die Steigung der "Tangente" zu nehmen!! Aber sie kennen nur für den Kreis eine Definition: Die Tangente in einem Kreispunkt steht senkrecht auf dem Radius zu diesem Kreispunkt. Diese Definition läßt sich nicht verallgemeinern.

Die Kreistangente hat noch eine weitere Eigenschaft:

Sie läßt den Kreis auf einer Seite.

Das ist zwar keine Eigenschaft, mit deren Hilfe man viel beweisen kann, aber sie erlaubt, bei einfachen Funktionen Tangenten zu finden als *solche* Geraden, die den Funktionsgraph zwar treffen, aber ihn auf einer Seite lassen. Dazu formuliere ich diese Eigenschaft der Kreistangente noch etwas um.
(Sprachregelung: Eine "Sekante" einer Kurve ist eine Gerade, die die Kurve schneidet. Eine "Sehne" einer Kurve ist eine Strecke, die zwei Kurvenpunkte verbindet. Die Verlängerung einer Sehne zu einer Geraden ist also eine Sekante. Sehnensteigungen und Sekantensteigungen sind dasselbe: Steigungen von Geraden.)

Nun die Umformulierung: Legt man einen Kreis so, dass er die x-Achse im Berührpunkt P von oben berührt, dann sind die von P ausgehenden Sehnensteigungen *rechts von P positiv* und *links von P negativ.* Und noch anders gesagt:

Die rechtsseitigen Sehnensteigungen sind <u>größer</u> als die Tangentensteigung,
die linksseitigen Sehnensteigungen sind <u>kleiner</u> als die Tangentensteigung.

Das läßt sich bei der quadratischen Parabel, dem Graph von $f(x) := x^2$, wiederholen. Wähle auf der Parabel die folgenden drei Punkte L, P, R mit $a \in \mathbb{R}$ beliebig und einer kleinen Schrittweite $h > 0$. In dem folgenden mit GeoGebra hergestellten Bild kann a mit einem Regler verändert werden. (R ist *r*echts von P und L *l*inks.)

$$P := (a \mid a^2),\ R := (a+h \mid (a+h)^2),\ L := (a-h \mid (a-h)^2).$$

Wir beobachten, dass die *rechtsseitigen* Sehnensteigungen $> 2a$ sind:

Steigung zwischen P und R: $\quad m_{pr} := \dfrac{(a+h)^2 - a^2}{h} = 2a + h \ > \ 2a,$

und wir beobachten, dass die *linksseitigen* Steigungen $< 2a$ sind:

Steigung zwischen L und P: $\quad m_{pl} := \dfrac{a^2 - (a-h)^2}{h} = 2a - h \ < \ 2a.$

Der Unterschied dieser Sehnensteigungen zu $2a$ ist so klein wie die Schrittweite h.

Ihr Mittelwert ist die Sehnensteigung zwischen L und R und sogar unabhängig von der Große von h:

$$m_m := \frac{m_{pr} + m_{pl}}{2} = \frac{(a+h)^2 - (a-h)^2}{(a+h) - (a-h)} = 2a.$$

Die Gerade mit der Steigung $m = 2a$ durch den Punkt $(a \mid a^2)$, also

$$t_a(x) := 2a \cdot (x - a) + a^2 = 2a \cdot x - a^2,$$

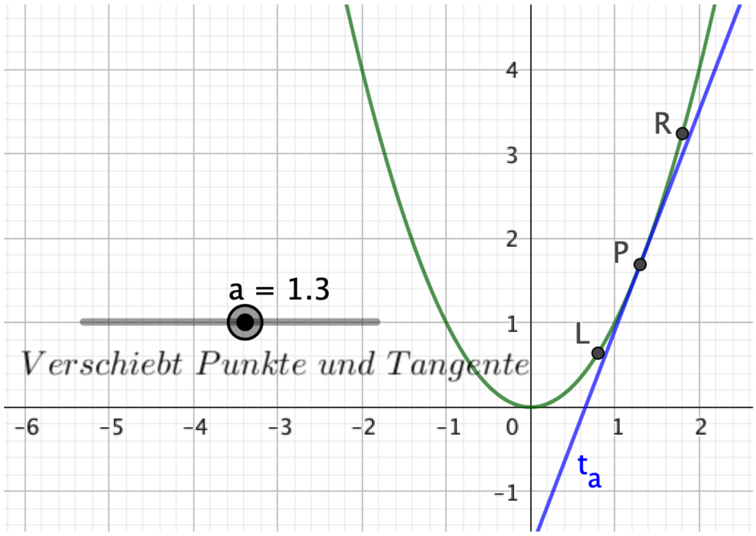
ist daher etwas besonderes.

Behauptung: Sie läßt die Parabel auf einer Seite.

Für das Auge zeigt das folgende Bild die Behauptung:

GeoGebra Beschreibung:

$$P = (a \mid a^2)$$
$$R = (a + h \mid (a+h)^2$$
$$L = (a - h \mid (a-h)^2$$
$$m_{pr} = \frac{R_y - P_y}{R_x - P_x} > 2a$$
$$m_{pl} = \frac{P_y - L_y}{P_x - L_x} < 2a$$
$$m_m = \frac{m_{pr} + m_{pl}}{2} = 2a$$



Mit der dritten binomischen Formel kann man es in einer Zeile nachrechnen:
Rechnerischer Beweis der Behauptung (mit $x^2 - a^2 = (x-a)(x+a)$):

$$f(x) - t_a(x) = x^2 - (2a \cdot (x-a) + a^2) = (x-a) \cdot (x + a - 2a) = (x-a)^2 \geq 0.$$

Ich glaube, dass auch in der griechischen Antike Parabeltangenten dadurch definiert waren, dass sie die Parabel auf einer Seite lassen, denn diese Eigenschaft konnte ohne die Formeln der analytischen Geometrie ausgedrückt werden. Für die quadratische Parabel haben wir daher eine zufriedenstellende Tangentendefinition. Und:
Die Steigung der Parabel im Berührpunkt wird definiert(!) als die Steigung dieser besonderen Gerade, eben als die Steigung der Tangente.
Natürlich muss dazu *vorher* geklärt sein, welche Geraden man als Tangenten bezeichnet. Man folgt also dem am Anfang erwähnten Schülervorschlag!

Diesen Erfolg kann man auf die kubische Parabel $f(x) = x^3$ ausdehnen, wenn man sie nur über der *positiven* reellen Achse betrachtet, also $x \geq 0$ voraussetzt.
Ähnlich wie eben werden drei mit einem Schieber (a) verschiebbare Punkte L, P, R auf dem Graphen dieser Funktion gewählt (mit $a - h \geq 0$) :

$$P := (a \mid a^3),\ \ R := (a+h \mid (a+h)^3),\ \ L := (a-h \mid (a-h)^3).$$

Beobachte, dass für die *rechtsseitigen* Sehnensteigungen gilt:

$$m_{pr} := \frac{(a+h)^3 - a^3}{h} = 3a^2 + 3ah + h^2 \ > \ 3a^2$$

und für die *linksseitigen* gilt unter der zusätzlichen Voraussetzung $0 \leq a - h$:

$$m_{pl} := \frac{a^3 - (a-h)^3}{h} = 3a^2 - 3ah + h^2 \ < \ 3a^2$$

Wir lesen die Erwartung ab, dass die Tangente die Steigung $3a^2$ haben sollte, also folgende Gerade sein sollte:

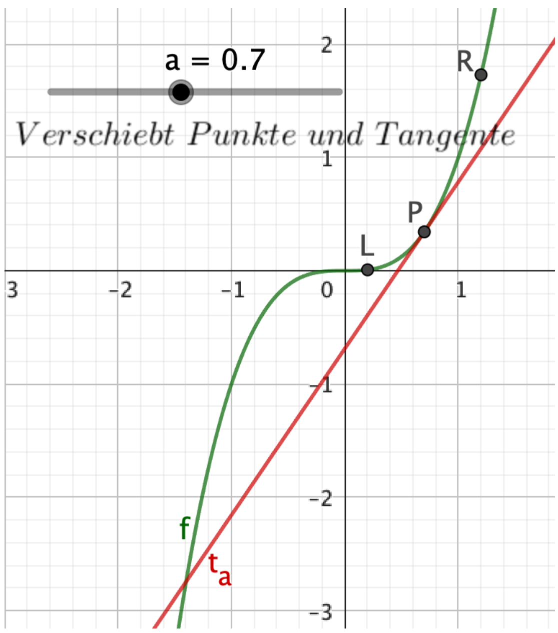
$$t_a(x) := 3a^2 \cdot (x - a) + a^3 = 3a^2 \cdot x - 2a^3.$$

Wir rechnen mit $x^3 - a^3 = (x-a) \cdot (x^2 + ax + a^2)$ nach, dass für $x \geq 0$ der Graph *oberhalb* dieser Geraden verläuft:

$$\begin{aligned} f(x) - t_a(x) &= x^3 - a^3 - 3a^2 \cdot (x-a) = (x-a) \cdot (x^2 + xa + a^2 - 3a^2) \\ &= (x-a)^2 \cdot (x + a + a) \geq 0. \end{aligned}$$

Aber wir sehen auch, dass die so gefundene Tangente den Graphen noch einmal in $(-2a \mid (-2a)^3)$ schneidet, allerdings erst links von 0, was uns nicht stört, da wir den Graphen nur über der positiven x-Achse betrachten.

$$P = (a \mid a^3)$$
$$R = (a+h \mid (a+h)^3$$
$$L = (a-h \mid (a-h)^3$$
$$m_{pr} = \frac{R_y - P_y}{R_x - P_x} > 3a^2$$
$$m_{pl} = \frac{P_y - L_y}{P_x - L_x} < 3a^2$$
$$t_a(x) = 3a^2 \cdot (x-a) + a^3$$



Wir können so zwar Tangenten für einfache Funktionen finden, aber die Methode erlaubt nicht, den die Analysis (von Funktionen einer Variablen) dominierenden Monotoniesatz zu beweisen. Wir benutzen ja nur, was vor Newton bekannt war.

Über Intervallen gegebene Steigungen sind einfach zu benutzen!

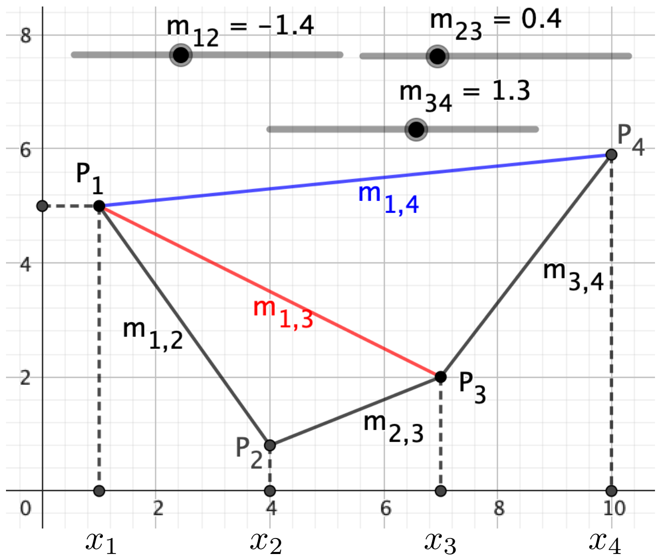
Wir betrachten eine Reihe aneinander gelegter Intervalle, auf denen Steigungen gegeben sind. Genauer, gegeben seien die Zahlen $x_1 < x_2 < \ldots < x_n$.
Diese definieren $n-1$ aneinander liegende Intervalle $[x_1, x_2], [x_2, x_3], \ldots, [x_{n-1}, x_n]$.
Für diese Intervalle $[x_{j-1}, x_j]$ seien die wählbaren Steigungen $m_{j-1,j} \in \mathbb{R}$ definiert.
Dann kann ein Streckenzug $P_1 = (x_1|y_1), P_2 = (x_2|y_2), \ldots, P_n = (x_n|y_n)$ schrittweise so berechnet werden, dass er im Intervall $[x_{j-1}, x_j]$ die gegebene Steigung $m_{j-1,j}$ hat, nämlich $y_j := y_{j-1} + m_{j-1,j} \cdot (x_j - x_{j-1}),\ j = 2$, dann $j = 3$ usw.
Mit diesem Streckenzug haben wir den Graphen einer stückweise linearen Funktion mit den *gegebenen* Steigungen gefunden! (Vergleiche: Stammfunktionen finden)

Beispiel mit $n = 4$:

Gegeben:
$x_1 < x_2 < x_3 < x_4$,
$y_1, P_1 = (x_1|\,y_1)$
und wählbare Steigungen
$m_{12},\ m_{23},\ m_{34} \in \mathbb{R}$.

Gesucht: $y_2,\ y_3,\ y_4$ und damit $P_2 = (x_2\,|\,y_2), P_3, P_4$

Berechne:
$y_2 = y_1 + m_{12} \cdot (x_2 - x_1)$,
$P_2 = (x_2\,|\,y_2)$, usw.



Man kann auch Aussagen über Steigungen zwischen *anderen* als den Intervall-Endpunkten machen: *Sie liegen immer* **zwischen** *der kleinsten und der größten gegebenen Steigung, die zwischen den betrachteten Punkten vorkommt.* Ich glaube, dass das in der Quotientenform unübersichtlicher ist, als wenn man die Nenner hoch multipliziert:

Statt $\quad m_{jk} = \dfrac{y_k - y_j}{x_k - x_j} \quad$ benutze $\quad (y_k - y_j) = m_{jk} \cdot (x_k - x_j)$.

Zum Beispiel können die drei Summanden, die $y_4 - y_1$ liefern, vergrößert werden mit Hilfe von $a \le b, 0 < c \ \Rightarrow\ a \cdot c \le b \cdot c$ oder verkleinert mit $a \ge b, 0 < c \ \Rightarrow\ a \cdot c \ge b \cdot c$.
Abkürzung: $\max(m_{jk}) := \max(m_{12}, m_{23}, m_{34})$. Nämlich mit $x_{j+1} - x_j > 0$ so:

$$y_4 - y_1 = m_{12} \cdot (x_2 - x_1) + m_{23} \cdot (x_3 - x_2) + m_{34} \cdot (x_4 - x_3)$$

$$\begin{cases} \le\ \max(m_{jk}) \cdot (x_2 - x_1\ +\ x_3 - x_2\ +\ x_4 - x_3) = \max(m_{jk}) \cdot (x_4 - x_1) \\ \ge\ \min(m_{jk}) \cdot (x_2 - x_1\ +\ x_3 - x_2\ +\ x_4 - x_3) = \min(m_{jk}) \cdot (x_4 - x_1). \end{cases}$$

Das ist der Beweis des "Monotoniesatzes" für Steigungen, die für Intervalle gegeben sind. Was man mit Steigungen anfangen kann, die nicht zu ganzen Intervallen, sondern nur zu Punkten gegeben sind, war 2000 Jahre lang ein Rätsel – ausgedrückt z. B. durch die Formulierung von Demokrit: *Die Materie muss aus Atomen bestehen, da man unendlich kleine Teile nicht zu etwas Endlichem zusammensetzen kann.*

Was die Erfindung der Differentialrechnung so erfolgreich machte

Die endgültige Antwort ist für den Anfang zu allgemein

Wer mit Kenntnissen in Mathematik, Physik und Technik auf die Mathematikgeschichte zurückblickt, wird wohl als den größten Erfolg der Analysis ansehen, dass die neuen Begriffe erlaubten, *die Naturgesetze präzise mit Gleichungen zu formulieren.* Diese Antwort so zu verstehen, dass sie nicht nur Propaganda sondern eine konkrete Erklärung ist, setzt voraus, die Analysis schon zu kennen. Wer gerade beginnt, Analysis zu lernen, muss meiner Meinung nach mit einer weniger globalen Erklärung angesprochen werden. Das möchte ich versuchen.

Das eben benutzte Kriterium ist keine Tangentendefinition

Erstens, es gibt Ausnahmetangenten: Die x-Achse ist bei $x = 0$ Tangente des Graphen von $f(x) = x^3$, aber sie wechselt im Berührpunkt von einer Seite des Graphen auf die andere Seite. Das ist kein Einzelfall, denn Tangenten in Wendepunkten wechseln immer von einer Seite auf die andere Seite.

Zweitens, die x-Achse berührt den Graphen von $g(x) = |x|$ in $(0\,|\,0)$ von unten, aber sie ist *nicht Tangente* im Sinne der Analysis. Es gibt sehr viele Beispiele solcher einseitigen Berührungen, die *keine Tangenten* sind. Es ist zwar richtig, dass Tangenten eines Funktionsgraphen diesen oft wenigstens ein Stück weit auf einer Seite lassen, aber diese Eigenschaft erfasst nicht, was mit Tangente gemeint ist. Die Formulierung "Berührung von einer Seite" ist zu schwach, um die für die Analysis wesentlichen Konsequenzen ziehen zu können.

Wir haben dasselbe Problem wie die antiken Mathematiker

In der Antike gab es keine Koordinatensysteme und keine Beschreibung von Funktionen mit Formeln. Diese Mittel stehen uns erst seit Descartes (1596 - 1650) und Fermat (1607 - 1665) zur Verfügung, also nur wenige Jahrzehnte länger als Newtons Entdeckung. Ich möchte hier nicht die Arbeitsweise griechischer Mathematiker genau schildern, ich lasse es bei Demokrits Formulierung: *Unendlich kleine Teile kann man nicht zu etwas Endlichem zusammensetzen* und übersetze dies schon damals diskutierte Problem in den Kontext von Funktionen, für die wir wie auf den Seiten 3 - 5 Tangenten finden konnten. Wir denken also an Funktionen, deren Tangenten den Graphen (der Funktion f im betrachteten Definitionsbereich D_f) auf einer Seite lassen. Es ist kein Problem zu verabreden, dass für alle $x \in D_f$ die bekannte Steigung $m(x)$ der Tangente an den Graphen im Punkte $(x\,|\,f(x))$ auch als *Steigung des Graphen* in diesem <u>*Punkt*</u> bezeichnet wird.

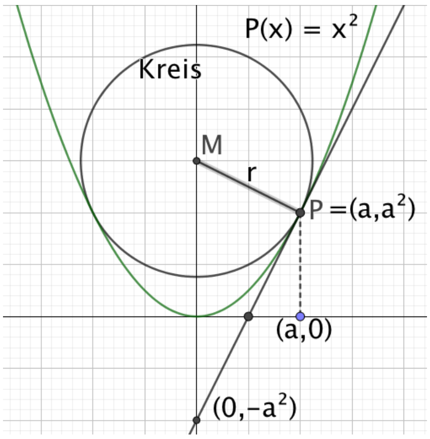
Nun ist offensichtlich, dass man diese "punktweisen" Steigungen *nicht so* zusammensetzen kann, wie ich das auf Seite 4 für über Intervallen definierte Steigungen vorgeführt habe. Selbst die moderne Analysis hat kein Konzept, wie punktweise Steigungen zusammengesetzt werden könnten, wenn man wirklich *keine weiteren Voraussetzungen* hat, als dass für jedes $x \in [a, b]$ eine *irgendwie* definierte Steigung $m(x)$ gegeben ist. So interpretiert hat Demokrit also noch heute recht.

Zusätzliche Voraussetzungen, die Folgerungen möglich machen

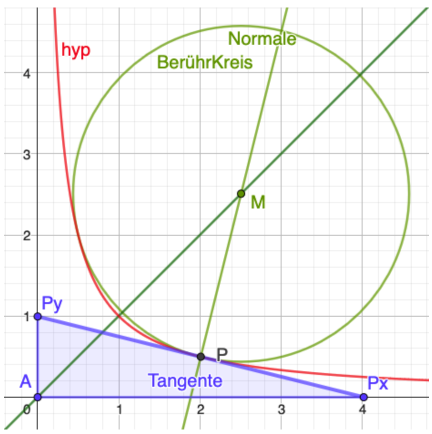
Der auf der Schule gern als selbstverständlich zitierte *Monotoniesatz* enthält solche zusätzlichen Voraussetzungen. Gegeben sind mit $f'(x)$ bezeichnete punktweise Steigungen, die *als Tangentensteigungen einer Funktion f erhalten wurden*!! Das Problem, wie sich die punktweisen Steigungen als Steigungen eines Funktionsgraphen aneinander fügen lassen, ist durch diese Voraussetzung gelöst. Dann läßt sich folgender Monotoniesatz für die punktweisen Steigungen beweisen:

Monotoniesatz:
$$\left.\begin{array}{r} f'(x) \geq 0 \text{ für } x \in (a,b) \\ x < y \in [a,b] \end{array}\right\} \Rightarrow f(x) \leq f(y).$$

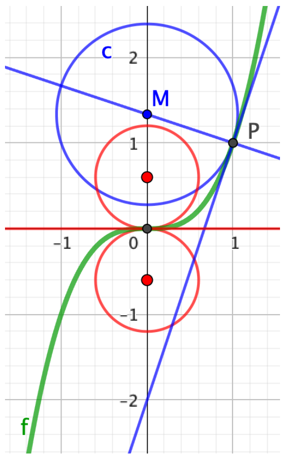
In Worten: Falls $f' \geq 0$ ist, gibt es keine negativen Sekantensteigungen. Durch diesen Satz bekommen die eigentlich unanschaulichen punktweisen Steigungen den anschaulichen Sinn, den auch ihr Name (Ableitung = Steigung) ausdrückt.
Allerdings läßt sich der Satz noch nicht mit den bisher formulierten Eigenschaften von Tangenten beweisen. Ich versuche, an Bildern einfacher Funktionen zu erklären, was noch fehlt.



Parabel: $\mathrm{P}(x) = x^2$



Hyperbel: $\mathrm{hyp}(x) = 1/x$



Potenz: $f(x) = x^3$

In allen drei Fällen wird auf der Normalen im Punkt P des Funktionsgraphen ein Punkt M gewählt, so dass das Innere des Kreises um M durch P den Funktionsgraphen *nicht* trifft. In den ersten beiden Fällen ist M als Schnitt zwischen Normale und Symmetrieachse gewählt. Man kann mit den binomischen Formeln nachrechnen, dass der Funktionsgraph zwischen diesem Kreis und der gemeinsamen Tangente von Kreis und Funktionsgraph hindurch läuft. Mit anderen Worten:

Der Funktionsgraph berührt die Tangente besser als der gewählte Kreis!

Im Wendepunkt der kubischen Potenz kann man das nicht erreichen, aber man kann auf jeder Seite des Graphen einen Kreis wählen (rot), sodass die beiden Kreise eine gemeinsame Tangente haben und der Funktionsgraph zwischen diesen Kreisen verläuft. Auf diese Weise haben wir die (bisherige) einseitige Berührung der Tangente zu einer beidseitigen Kontrolle verbessert. – Eine solche *beidseitige Kontrolle* der Berührung von Funktionsgraph und Tangente muss in der Tangentendefinition enthalten sein, um den Monotoniesatz beweisbar zu machen.

Wieso liefert $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a)$ solch eine Kontrolle?

Das ist nicht mehr leicht zu erklären, seit Grenzwerte nicht mehr mit Ungleichungen sondern nur noch undeutlich verbal beschrieben werden, etwa "h geht gegen 0". In der vollständigen Grenzwertdefinition kommt üblicherweise ein Symbol ϵ vor, das von einer beinahe magischen Aura umgeben ist. Aber eigentlich haben es sich die Mathematiker nur zur Gewohnheit gemacht, eine *beliebig wählbare(!) Fehlerschranke* mit dem griechischen Buchstaben ϵ zu bezeichnen. Eine "Fehlerschranke" ist eine Zahl, die größer ($\geq$) als der Unterschied zwischen einer theoretisch definierten Zahl und einem stattdessen verwendeten numerischen Wert ist. Z.B. ist $0,1$ eine Fehlerschranke für den Unterschied zwischen $\sqrt{2}$ und $1,4$. Eine "wählbare" Fehlerschranke ϵ enthält zusätzlich die Aufforderung, in einer bevorstehenden Rechnung genauer als ϵ zu sein. (Das Benutzen von ϵ ist neben der Gewohnheit auch bequem, denn das Auftreten von ϵ sagt dem Leser ohne weitere Worte, dass von einer wählbaren Fehlerschranke die Rede ist.) In der Definition von $\lim_{h\to 0}$ in

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a)$$

kommt es dann darauf an, für *jede* gewählte Fehlerschranke $\epsilon > 0$ aus Einschränkungen an $h \neq 0$ zu folgern, dass folgende Ungleichung gilt:

$$(*) \qquad \left| \frac{f(a+h)-f(a)}{h} - f'(a) \right| \leq \epsilon.$$

Die Einschänkungen an h werden gewohnheitsmäßig mit einem weiteren griechischen Buchstaben ausgedrückt: Zu jeder gewählten Fehlerschranke $\epsilon > 0$ muss eine Zahl $\delta > 0$ gefunden werden, sodass man für die gegebene Funktion f aus der Voraussetzung $0 < |h| \leq \delta$ folgern kann, dass die Ungleichung $(*)$ gilt. Solche Satzkonstruktionen kommen in der Umgangssprache nicht vor und deshalb ist die Grenzwertdefinition unbeliebt.
Wenn man $(*)$ mit $|h|$ multipliziert und $f'(a) \cdot h + f(a)$ als Tangentengleichung in $x = a$ erkennt, dann sieht man, dass eine beidseitige Kontrolle für den Unterschied zwischen Funktionsgraph und Tangente unmittelbar äquivalent zu $(*)$ ist:

$$|h| \leq \delta \;\Rightarrow\; |f(a+h) - \big(f'(a)\cdot h + f(a)\big)| \leq \epsilon \cdot |h|.$$

Diese Kontrolle der Berührung von Funktionsgraph und Tangente ist etwas schwächer als die vorher mit Kreisen formulierte Kontrolle. Sie ist auch wesentlich unanschaulicher, weil sie nicht als *eine* Ungleichung verstanden werden soll, sondern als *unendlich viele* Ungleichungen – eben zu *jeder* Wahl der Fehlerschranke ϵ.
Mit diesem Verständnis des Grenzwertbegriffs läßt sich der Monotoniesatz beweisen. Die anschaulichere beidseitige Kontrolle durch Kreise vereinfacht den Beweis etwas – die Ungleichungen verändern sich kaum, aber die zugehörige Sprache ist weniger anspruchsvoll.

Der Monotoniesatz ist nur der Anfang des Erfolgs der Analysis

Ich möchte nicht, dass meine Formulierungen den Eindruck hervorrufen, der Monotoniesatz sei der eigentliche Erfolg der Analysis. Er ist nur ein wichtiger Anfangserfolg, der schon in der Schule angesprochen wird. Da die Ableitung wegen des Monotoniesatzes auch den Namen *Steigung des Graphen* erhalten hat, wirkt der Monotoniesatz leicht wie eine Selbstverständlichkeit. Schon eine zweimal wiederholte Anwendung dieses Satzes liefert folgende wesentlich weniger selbstverständliche Eigenschaft: Hat man für eine Funktion f in einem Intervall $[a, b]$ die Voraussetzung $|f''| \leq 2B$, dann ist der Unterschied zwischen dem Graphen von f und seiner Sehne zwischen a und b höchstens $B \cdot (x - a)(b - x)$:

$$\left. \begin{array}{r} |f''| \leq 2B \ \text{ in } [a, b] \\ x \in [a, b] \end{array} \right\} \Rightarrow \left| f(x) - \frac{f(a) \cdot (b - x) + f(b) \cdot (x - a)}{b - a} \right| \leq B(x - a)(b - x).$$

Dies ist eine Fehlerschranke beim Interpolieren. Alle Fehlerschranken der numerischen Mathematik für eindimensionale Funktionen entstehen auf ähnliche Weise.

Für die theoretische Physik besonders wichtig ist, dass neue Funktionen mit Hilfe von Ableitungseigenschaften definiert werden können. Ein einfaches Beispiel ist die Exponentialfunktion. Sie kann definiert und numerisch berechnet werden mit Hilfe der Eigenschaft (genannt "Differentialgleichung") $f' = f,\ f(0) = 1$. Allerdings sind eindimensionale Situationen die Ausnahme, sodass die Details erheblich komplizierter werden. Ich begnüge mich mit der Zusammenfassung: Da die meisten Naturgesetze aus den Daten eines Zustands seine Ableitung berechnen, kann man ohne das Konzept der Differentialrechnung die Naturgesetze gar nicht formulieren.

Was ist mit dem Mittelwertsatz der Differentialrechnung?

Es ist in der Mathematik üblich, den folgenden Mittelwertsatz als die wesentliche Aussage zu betrachten und den Monotoniesatz als eine offensichtliche Folgerung ("Wäre der Monotoniesatz falsch, so hätte man sofort einen Widerspruch zum Mittelwertsatz"). Ich formuliere diesen Satz hier nur, ohne ihn zu erklären:

Mittelwertsatz: $\left\{ \begin{array}{l} f \text{ sei stetig in } [a, b] \text{ und differenzierbar in } (a, b), \text{ dann gilt:} \\ \quad \text{Es gibt eine reelle Zahl } \xi \in (a, b) \text{ mit } f'(\xi) = \frac{f(b) - f(a)}{b - a}. \end{array} \right.$

In Worten: *Zu jeder Sehne gibt es eine parallele Tangente.*

Im Kontext des Anfangsunterrichts in Analysis gebe ich zu bedenken: Erstens ist der begriffliche Aufwand für einen Beweis des Mittelwertsatzes erheblich höher als für einen nur die Definition der Ableitung benutzenden Beweis des Monotoniesatzes. Und zweitens kenne ich nur einen einzigen mit Hilfe des Mittelwertsatzes geführten Beweis, den ich nicht schon mit dem Monotoniesatz führen kann, nämlich den Beweis der Symmetrie der zweiten partiellen Ableitungen – also einer Anwendung weit jenseits der Schule.