Der Monotoniesatz ist ein phantastisches Werkzeug.

Zuächst gab es nur (Sehnen-)Steigungen einer Funktion f über einem Intervall $[a,b]$, nämlich $m = (f(b) - f(a))/(b-a)$. Steigungen in einem Punkt machten keinen Sinn. Seit Newton und Leibniz kann man Steigungen in einem Punkt definieren, $m = f'(a)$, nicht experimentell, nur in der theoretischen Beschreibung. Erstaunlicher Weise kann man aus Voraussetzungen über diese "punktweisen Steigungen" Folgerungen ziehen, die weitreichende Konsequenzen haben. Der folgende Satz wird in allen Schulbüchern als anschaulich klar vorgestellt:

Der Monotoniesatz

Gegeben sei eine differenzierbare Funktion $f : [a,b] \to \mathbb{R}$
Voraussetzung: $x \in [a,b] \Rightarrow f'(x) \geq 0$.
Behauptung: $x < y \in [a,b] \Rightarrow f(x) \leq f(y)$.

Egal, als wie anschaulich dieser Satz hingestellt wird: Fast jeder, der den Beweis nicht gesehen hat, wird bei einem eigenen Beweisversuch scheitern. Und alle Ungleichungen der Analysis, die auftreten, ehe man das sogenannte Maximumprinzip kennen gelernt hat, werden mit dem Monotoniesatz (bzw. seiner n-dimensionalen Version, dem Schrankensatz) bewiesen. Der Monotoniesatz muss auf der Schule nicht bewiesen werden. Aber ohne seine Würdigung ist der Unterricht kein Analysisunterricht. Ohne ihn handelt es sich nur um die Einbettung von Fachausdrücken in die Umgangssprache. – Wir kommen zu Folgerungen.

Die Anwendungsmöglichkeiten nehmen sehr zu, wenn man den Monotoniesatz für zwei Funktionen formuliert:

Funktionen mit größerer Ableitung wachsen schneller.

Gegeben seien differenzierbare Funktionen $f, g : [a,b] \to \mathbb{R}$

Voraussetzung: $x \in [a,b] \Rightarrow f'(x) \leq g'(x)$.

Behauptung: $x < y \in [a,b] \Rightarrow g(x) - f(x) \leq g(y) - f(y)$ oder $f(y) - f(x) \leq g(y) - g(x)$.

Beispiel Lipschitz-Schranke: $-L \leq f'(x) \leq L \Rightarrow -L\cdot(y-x) \leq f(y) - f(x) \leq L\cdot(y-x)$,
In Worten: *Ableitungsschranken sind "Lipschitzschranken".*

Beweis: Die Funktion $h := g - f$ erfüllt die Voraussetzungen des Monotoniesatzes.

Statt des vorhergehenden Vergleichs durch Differenzbildung kann man auch Quotienten positiver Funktionen vergleichen:

Der multiplikative Monotoniesatz.

Gegeben seien positive differenzierbare Funktionen $f, g : [a,b] \to \mathbb{R}_+$

Voraussetzung: $x \in [a,b] \Rightarrow f'(x)/f(x) \leq g'(x)/g(x)$.

Behauptung: $x < y \in [a,b] \Rightarrow g(x)/f(x) \leq g(y)/f(y)$ oder $f(y)/f(x) \leq g(y)/g(x)$.

Beispiel: $h'/h \leq 1 \leq H'/H \Rightarrow h(x)/h(0) \leq \exp(x) \leq H(x)/H(0)$.

Beweis: Die Funktion g/f erfüllt die Voraussetzungen des Monotoniesatzes wegen

$$(g/f)' = (g \cdot \frac{1}{f})' = g'/f + g\cdot(-f'/f^2) = g/f \cdot (g'/g - f'/f) \geq 0.$$

1. Die Lipschitz-Schranke (erstes Beispiel oben) besitzt eine überraschende Anwendung auf konvergente Funktionenfolgen f_n mit **derselben** Ableitungsschranke: $-L \le f_n' \le +L$ für alle n. Obwohl man über die Ableitung der Grenzfunktion f_∞ noch nichts weiß, folgt deren Stetigkeitsabschätzung:
Zu jeder Fehlerschranke $1/k$ gibt es Indices n_x, n_y, so dass für alle $n \ge n_x, n_y$ gilt
$\big|f_\infty(x) - f_n(x)\big| \le 1/k,\ \big|f_\infty(y) - f_n(y)\big| \le 1/k,$ also $\big|f_\infty(y) - f_\infty(x)\big| \le \big|f_n(y) - f_n(x)\big| + 2/k$
Nun liefert Archimedes' Argument $\big((\forall_k\ a \le b + 1/k) \Rightarrow a \le b\big)$:

$$\forall_{n\in\mathbb{N}}\ \big|f_n(y) - f_n(x)\big| \le L \cdot |y - x| \ \Rightarrow\ \big|f_\infty(y) - f_\infty(x)\big| \le L \cdot |y - x|.$$

2. Der multiplikative Monotoniesatz erlaubt zum Beispiel eine unerwartet einfache Behandlung der Exponentialfunktion. (Vergleiche Text A2 und Beispiel 6 in Text A3.)

Für die Funktion $h_n(x) := (1 + x/n)^n$ gilt: $(h_n'/h_n)(x) = 1/(1 + x/n) \le 1$.

Für die Funktion $H_n(x) := 1/h_n(-x) = 1/(1-x/n)^n$ gilt: $(H_n'/H_n)(x) = 1/(1-x/n) \ge 1$.

Daher folgen aus dem multiplikativen Monotoniesatz mühelos obere und untere Approximationen der Exponentialfunktion (natürlich mit $x < n$ in der rechten Ungleichung)

$$0 \le x \quad \Rightarrow \quad (1 + x/n)^n \le \exp(x) \le (1 - x/n)^{-n}$$

Genauere Aussagen folgen aus **zweifacher** *Anwendung des Monotoniesatzes.*

Wir diskutieren zuerst, wie groß die Abweichung von der Tangente ist. Die Ungleichungen sind in beiden Richtungen zu lesen. Einerseits: Wie viel sagt die Ableitung über Differenzen von Funktionswerten aus? Andererseits: Wie genau kann man Ableitungen aus Differenzenquotienten bestimmen?

Abweichung von der Tangente.

Gegeben sei eine zweimal differenzierbare Funktion $f : [a, b] \to \mathbb{R}$
Voraussetzung: $x \in [a, b] \Rightarrow -B \le f''(x) \le +B$.
Behauptung: $x, c \in [a, b] \Rightarrow |f(x) - f(c) - f'(c) \cdot (x - c)| \le 0.5B \cdot |x - c|^2$
Beweis:
Der Wachstumsvergleich von zwei Funktionen (Monotoniesatz) liefert für die Ableitung f'

$$x \ge c \Rightarrow \pm(f'(x) - f'(c)) \le +B \cdot (x - c) \text{ und}$$
$$x \le c \Rightarrow \pm(f'(x) - f'(c)) \le -B \cdot (x - c).$$

Die Differenz zwischen Funktion und Tangente ist $h(x) := f(x) - f(c) - f'(c) \cdot (x - c)$.
Wir haben also:

$$x \ge c \Rightarrow \pm h'(x) \le +0.5B\big((x - c)^2\big)' \text{ und}$$
$$x \le c \Rightarrow \pm h'(x) \le -0.5B\big((x - c)^2\big)'.$$

Der Monotoniesatzes mit $h(c) = 0$ gibt

$$\pm h(x) \le 0.5B\cdot(x - c)^2\Big|_c^x, \text{ bzw. } \le -0.5B \cdot (x - c)^2\Big|_x^c.$$
$$\text{also } |h(x)| \le 0.5B(x - c)^2, \text{wie behauptet.}$$

Bemerkung (Gleicher Beweis):
Die Voraussetzung $0 \le f'' \le B$ liefert, dass der Graph von f an jeder Stelle c oberhalb seiner Tangente und unterhalb einer quadratischen Parabel verläuft:

$$0 \le f(x) - f(c) - f'(c) \cdot (x - c) \le 0.5B \cdot (x - c)^2.$$

Ebenso interessant ist es, Abweichungen von Sehnen zu kontrollieren. Wir zeigen zuerst:

Funktionen mit $f'' \ge 0$ liegen unterhalb jeder Sehne.

Gegeben sei eine zweimal differenzierbare Funktion $f : [\alpha, \omega] \to \mathbb{R}$
Voraussetzung: $f'' \ge 0$.
Behauptung: Der Graph von f liegt unterhalb jeder Sehne.
Beweis:
Die Sehne über $[a, b] \subset [\alpha, \omega]$ ist: $S_{ab}(x) := \big(f(b) \cdot (x - a) + f(a) \cdot (b - x)\big)/(b - a)$.
Für die Hilfsfunktion $h(x) := f(x) - S_{ab}(x)$ gilt $h'' = f'' \ge 0$. Daher ist h' wachsend.
Ziel ist, an jeder Stelle $c \in [a, b]$ zu zeigen: $h(c) \le 0$. Schon bekannt ist $h(a) = h(b) = 0$.
Da wir über das Vorzeichen von h' nichts wissen, brauchen wir eine Fallunterscheidung:
Entweder $h'(c) \ge 0$, dann folgt für $x \in [c, b]$, weil h' wächst: $0 \le h'(c) \le h'(x)$. Nach dem Monotoniesatz wächst h in $[c, b]$, so dass aus $h(b) = 0$ folgt $h(c) \le 0$.
Oder $h'(c) < 0$, dann folgt für $x \in [a, c]$, weil h' wächst: $h'(x) \le h'(c) < 0$. Nach dem Monotoniesatz fällt h in $[a, c]$, so dass $0 = h(a) \ge h(c)$ folgt.
In beiden Fällen haben wir also die Behauptung $h(c) \le 0$.

Bemerkung:
Die Voraussetzung $0 \le f'' \le B$ liefert: Der Graph von f verläuft unterhalb der Sehne und oberhalb einer quadratischen Parabel - zwischen den Endpunkten der Sehne, $a \le x \le b$:

$$\boxed{0.5B \cdot (x - b) \cdot (x - a) \;\le\; f(x) - S_{ab}(x) \;\le\; 0,}$$

denn für die Hilfsfunktion $h(x) := 0.5B \cdot (x - b) \cdot (x - a) + S_{ab}(x) - f(x)$ gilt:
$h'' = B - f'' \ge 0$ und $h(a) = h(b) = 0$, so dass für $x \in [a, b]$ wie vorher folgt: $h(x) \le 0$.

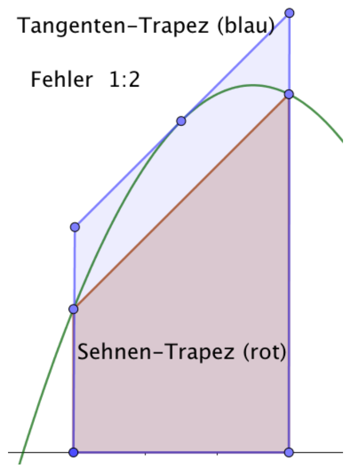
Dies sind gute Abschätzungen, aus denen wir praktische Folgerungen ziehen.
Erstens:

Die Fehlerabschätzung bei **linearer Interpolation** ist gleich dem Unterschied zwischen Funktion und Sehne.

Zweitens:

Fehlerabschätzung bei numerischer Integration:

Die beiden einfachsten Verfahren zur **numerischen Integration** einer Funktion f in einem (kleinen) Intervall $[a, b]$ (die auch in Schulbüchern vorkommen) sind:



Sehnen-Trapez Regel: $\displaystyle\int_a^b f(x)dx \approx \frac{f(a)+f(b)}{2} \cdot (b-a),$

Tangenten-Trapez Regel: $\displaystyle\int_a^b f(x)dx \approx f(\frac{a+b}{2}) \cdot (b-a).$

Setzen wir zusätzlich $0 \leq f'' \leq B$ voraus, so können wir die vorhergehenden Abschätzungen zur Fehlerkontrolle einsetzen; z.B. folgen die linken Abschätzungen $(0 \leq \ldots)$ daraus, dass f unterhalb seiner Sehnen und oberhalb seiner Tangenten verläuft:

$$0 \leq \frac{f(a)+f(b)}{2} \cdot (b-a) - \int_a^b f(x)dx \leq \int_a^b 0.5B(b-x)(x-a)dx = \frac{B}{12}(b-a)^3,$$

$$0 \leq \int_a^b f(x)dx - f(\frac{a+b}{2}) \cdot (b-a) \leq \int_a^b 0.5B(x-\frac{a+b}{2})^2 = \frac{B}{24}(b-a)^3.$$

Die Tangenten-Trapez Regel hat also einen etwa halb so großen Fehler wie die Sehnen-Trapez Regel. Außerdem ist der eine Wert zu klein, der andere zu groß. Wenn man die beiden Verfahren also im Verhältnis 2:1 mittelt, sollte man eine bessere numerische Formel bekommen. In der Tat:

$$\int_a^b f(x)dx \approx \frac{1}{3}\frac{f(a)+f(b)}{2} \cdot (b-a) + \frac{2}{3} f(\frac{a+b}{2}) \cdot (b-a) = \big(f(a) + 4f(\frac{a+b}{2}) + f(b)\big) \cdot \frac{b-a}{6}$$

wir bekommen die viel benutzte **Simpson Regel**, die Polynome bis zum Grad 3 exakt integriert! Aus Voraussetzungen an die vierten Ableitungen gewinnt man mit entsprechend oft wiederholter Anwendung des Monotoniesatzes eine Fehlerabschätzung der Simpson Regel: $\leq \max|f^{(4)}| \cdot (b-a)^5/2880$. (Beweis Seite 10)

Nächstes Ziel:

Im Zusammenhang mit der Integration werden in den Schulbüchern Produktsummen betrachtet. Das sind Riemann Summen zu äquidistanten Einteilungen. Um der Entwicklung dort zu folgen, muss man verstehen, wie wenig diese Produktsummen von der Differenz $F(b) - F(a)$ der Werte einer Stammfunktion F des Integranden f verschieden sind. Für unsere Abschätzung ist es nicht nötig, dass die Produktsummen zu äquidistanten Einteilungen gebildet werden, denn wir werden die Fehlerabschätzung für jedes Teilintervall einzeln machen und dann diese Fehler addieren. (Die wichtige Intervalladditivität der Integrale ist mit äquidistanten Einteilungen kaum zu beweisen.) Der folgende Vergleich zwischen Riemann Summen und Stammfunktionen ist **keine** Konstruktion des Integrals; er erklärt nur, warum man hoffen kann, mit Integralen Stammfunktionen zu konstruieren und auch, warum man keine Integrale braucht, wenn man Stammfunktionen hat.

Riemann Summen und Stammfunktionen.

Gegeben: Eine differenzierbare Funktion $f : [a, b] \to \mathbb{R}$, eine Stammfunktion F $(F' = f)$.
Zusatzvoraussetzung: Eine Schranke $|F''| = |f'| \le L$, um ein möglichst übersichtliches Ergebnis zu erhalten.
Gewählt: $x \le \xi \le y \in [a, b]$.

Behauptung: $|F(y) - F(x) - f(\xi) \cdot (y - x)| \le 0.5L \cdot (y - x)^2$

Beweis:
Erst mit der Dreiecksungleichung, dann mit der Abweichung von der Tangente und schließlich wegen
$(y - x)^2 = (y - \xi + \xi - x)^2 = (y - \xi)^2 + (\xi - x)^2 + 2(y - \xi)(\xi - x) \ge (y - \xi)^2 + (\xi - x)^2$

gilt:
$$\begin{aligned} & |F(y) - F(\xi) + F(\xi) - F(x) - f(\xi) \cdot (y - x)| \\ \le\ & |F(y) - F(\xi) - f(\xi) \cdot (y - \xi)| + |F(\xi) - F(x) - f(\xi) \cdot (\xi - x)| \\ \le\ & 0.5L \cdot (y - \xi)^2 + 0.5L \cdot (\xi - x)^2 \\ \le\ & 0.5L \cdot (y - x)^2 \end{aligned}$$

Dies Ergebnis wird jetzt angewandt auf die n Teilintervalle $[x_{i-1}, x_i]$ einer Einteilung $a = x_0 < x_1 < \dots x_n = b$ des Intervalls $[a, b]$ mit den Zwischenstellen $\xi_i \in [x_{i-1}, x_i]$.

$$\begin{aligned} & \Big|F(b) - F(a) - \sum_{i=1}^{i=n} f(\xi_i) \cdot (x_i - x_{i-1})\Big| = \\ & \Big|\sum_{i=1}^{i=n} F(x_i) - F(x_{i-1}) - f(\xi_i) \cdot (x_i - x_{i-1})\Big| \le \\ & \sum_{i=1}^{i=n} \Big|F(x_i) - F(x_{i-1}) - f(\xi_i) \cdot (x_i - x_{i-1})\Big| \le \\ & \sum_{i=1}^{i=n} 0.5L \cdot (x_i - x_{i-1})^2 \le 0.5L \cdot (b - a) \cdot \max_{i=1\dots n} |x_i - x_{i-1}|. \end{aligned}$$

Hier ist $L_{ab} := 0.5L \cdot (b - a)$ eine Konstante, die von der Einteilung nicht abhängt. Der Unterschied zwischen der Differenz $F(b) - F(a)$ und der Summe $\sum_{i=1}^{i=n} f(\xi_i) \cdot (x_i - x_{i-1})$ wird also durch die Länge des größten Teilintervalls kontrolliert:

$$\boxed{\Big|F(b) - F(a) - \sum_{i=1}^{i=n} f(\xi_i) \cdot (x_i - x_{i-1})\Big| \le L_{ab} \cdot \max_{i=1\dots n} |x_i - x_{i-1}|}$$

Also, wenn nur die Länge des längsten Teilintervalls gegen 0 konvergiert, dann haben all diese Riemannsummen denselben Grenzwert $F(b) - F(a)$.

Die hier auftretenden Formeln sind nicht komplizierter als die, die ich in den Schulbüchern gesehen habe. Der wesentliche Unterschied besteht darin, dass die behaupteten Zusammenhänge durch Argumentieren mit dem Monotoniesatz logisch begründet sind. Und dies logische Begründen ist, was die Mathematik von einer Lehre von Rezepten unterscheidet.

Approximation von Ableitungen durch Differenzenquotienten

Der Monotoniesatz folgert nicht nur aus Ableitungsvoraussetzungen Eigenschaften der Funktion, er kann auch in umgekehrter Richtung benutzt werden. Die am Anfang hergeleitete Abweichung von der Tangente

$$|f(x) - f(c) - f'(c) \cdot (x-c)| \le \tfrac{1}{2} B \cdot |x-c|^2$$

kann man auch als Fehlerabschätzung für die **numerische Differentiation** lesen, wobei mit numerischer Differentiation gemeint ist, Ableitungen durch Differenzenquotienten zu approximieren:

$$\left| \frac{f(x) - f(c)}{x-c} - f'(c) \right| \le \frac{1}{2} B \cdot |x-c|$$

Eine zweimalige Anwendung des Monotoniesatzes verbessert diese numerische Differentiation erheblich, da man mit größeren Werten von h weniger Stellenverlust hat und daher wegen des kleineren Formelfehlers größere Genauigkeit erreicht:

Voraussetzung: $x \in [a,b] \Rightarrow |f'''(x)| \le C$ und $(c \pm h) \in [a,b]$

Behauptung: $\left| \dfrac{f(c+h) - f(c-h)}{2h} - f'(c) \right| \le \dfrac{C}{6} h^2$

Bemerkung: Die Formel ist exakt für quadratische Polynome ($C = 0$).

Beweis:
Die aus Schranken für die zweite Ableitung hergeleitete Abweichung von der Tangente kann direkt für f' benutzt werden, denn $f''' = (f')''$:

$x, c \in [a,b] \Rightarrow$

$$-\frac{C}{2} \cdot (x-c)^2 \le f'(x) - f'(c) - f''(c) \cdot (x-c) \le \frac{C}{2} \cdot (x-c)^2 = \left(\frac{C}{6} \cdot (x-c)^3 \right)'$$

Alle Terme dieser Ungleichungskette sind Ableitungen, daher ergibt der Monotoniesatz:

$$-\frac{C}{6} \cdot |x-c|^3 \;\le\; f(x) - f(c) - f'(c)(x-c) - \frac{1}{2} f''(c)(x-c)^2 \;\le\; \frac{C}{6} \cdot |x-c|^3$$

Wähle für x nach einander $x = c+h,\ x = c-h,$ subtrahiere und dividiere durch $2h$

$$-\frac{C}{6} \cdot |h|^2 \;\le\; \frac{f(c+h) - f(c-h)}{2h} - f'(c) \;\le\; \frac{C}{6} \cdot |h|^2 \qquad q.e.d.$$

Die Verbesserung kommt dadurch zustande, dass die quadratischen Terme beim Subtrahieren wegfallen. Diese numerische Differentiation lohnt sich, falls die Berechnung von f' erheblich aufwendiger ist als zwei Berechnungen von f, oder falls f' gar nicht zur Verfügung steht. – Da beim Subtrahieren nahe benachbarter Zahlen Stellen verloren gehen, kann es sich bei fünfmal differenzierbaren Funktionen lohnen, durch Mittelung von zwei Rechnungen mit unterschiedlicher Schrittweite die Fehlerordnung h^4 zu erreichen, so dass die Differenzen mit $h = 2^{-7}$ ausgeführt werden können. Durch den großen Wert von h ist der Stellenverlust klein und h als Zweierpotenz verhindert nunötige Rundungsfehler im Rechner. Beispiel (beachte $4f'''(c)h^3/2h - f'''(c)(2h)^3/4h = 0$):

$$\left| \left(\frac{4}{3} \frac{f(c+h) - f(c-h)}{2h} - \frac{1}{3} \frac{f(c+2h) - f(c-2h)}{4h} \right) - f'(c) \right| \le const(f^{(5)}) \cdot h^4$$

Diese Formel berechnet Ableitungen von Polynomen bis zum Grad 4 exakt ($f^{(5)} = 0$). Wählt man $h = 1/128$, so erhält man die Ableitung von Sinus auf 10 Stellen.

Die folgende Diskussion ist nicht anspruchsvoller als das Verständnis der Definition der Ableitung als Grenzwert.

Wie kann man dem Monotoniesatz näher kommen?

Da Ableitungen als Grenzwerte definiert sind, kann man nicht mit endlich vielen Rechnerexperimenten Aussagen über Ableitungen gewinnen. Man wird erwarten, dass Beweise für Aussagen über Ableitungen Grenzwerte benutzen. Die ersten Beweise sollten nicht schwieriger sein als das Verständnis der Definition. Vielleicht wird das erreicht mit folgendem ersten Ergebnis, dem **Satz**:

Voraussetzung: f sei auf $[a, b]$ differenzierbar und $(f(b) - f(a))/(b - a) = m$.
Behauptung: Es gibt ein $\xi \in [a, b]$ mit $f'(\xi) \geq m$.
Bemerkung: Der *Mittelwertsatz der Differentialrechnung* besagt sogar:

Es gibt ein $\xi \in (a, b)$ mit $f'(\xi) = m$.

Aber dessen Beweis ist ein gutes Stück aufwendiger als der folgende

Beweis:

Wir werden eine Intervallschachtelung definieren, deren Grenzwert ξ die gewünschte Eigenschaft $f'(\xi) \geq m$ hat.

Da man die Sehne $S_{ab}(x) = \big(f(a) \cdot (b - x) + f(b) \cdot (x - a)\big)/(b - a)$ der Steigung m vor dem Beweis abziehen und nach dem Beweis wieder addieren kann, genügt es, den Beweis für $m = 0$ zu führen.

Setze $a_0 = a,\ b_0 = b,\ c = (a + b)/2$. Fallunterscheidung: Die Sehnensteigung von f ist über $[a, c]$ oder über $[c, b]$ mindestens 0. Denn aus $f(c) - f(a) < 0$ und $f(b) - f(c) < 0$ folgt durch Addition $f(b) - f(a) < 0$ im Widerspruch zur Voraussetzung. Im ersten Fall setze $a_1 = a_0,\ b_1 = c$, im zweiten Fall $a_1 = c,\ b_1 = b_0$. Dieser Schritt kann beliebig oft wiederholt werden und liefert Intervalle a_n, b_n der Länge $b_n - a_n = (b_0 - a_0)2^{-n}$ und $f(b_n) - f(a_n) \geq 0$.

Den Grenzwert der Intervallschachtelung nennen wir ξ. Wir halten fest $\forall_n\ \xi \in [a_n, b_n]$.

In jedem Intervall haben wir die Fallunterscheidung $f(\xi) - f(a_n) \geq 0$ oder $f(b_n) - f(\xi) \geq 0$. Im ersten Fall (oder falls $\xi = b_n$) setzen wir $c_n := a_n$, im zweiten Fall (oder falls $\xi = a_n$) setzen wir $c_n := b_n$. Dann ist immer $c_n - \xi \neq 0$ und $(f(c_n) - f(\xi))/(c_n - \xi) \geq 0$. Die Folge $\{c_n\}$ konvergiert gegen ξ und liefert $f'(\xi) \geq 0$.

Dieser Satz leitet aus Voraussetzungen über f eine Eigenschaft von f' her. Sehr oft ist man an Folgerungen in der umgekehrten Richtung interessiert. Das kann man aus diesem Satz, benutzt für $m = 0$, herausholen:

Voraussetzung: $f' < 0$ in $[a, b]$.
Behauptung: Es kann kein Teilintervall $[x, y] \subset [a, b]$ geben mit $f(y) - f(x) \geq 0$.
Indirekter Beweis: Ein solches Teilintervall würde die Voraussetzungen des Satzes erfüllen, so dass der Satz ein ξ liefert mit $f'(\xi) \geq 0$ - im Widerspruch zur Voraussetzung $f' < 0$ in $[a, b]$. Es gibt also kein solches Teilintervall, d.h. $\forall_{x<y\in[a,b]}\ f(y) < f(x)$.
Zusammenfassung: Eine Funktion f mit $f' < 0$ in $[a, b]$ ist streng fallend. Und das ist schon die einfachste Form des Monotoniesatzes.

Beinahe ein Gegenbeispiel zum Monotoniesatz: Die Cantortreppe

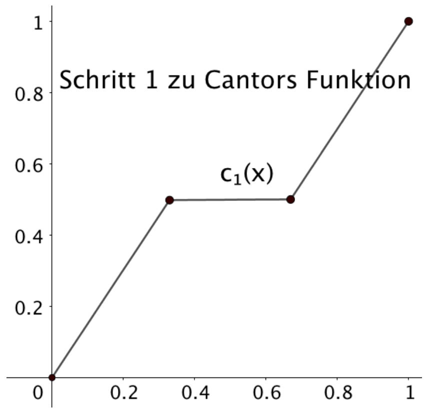
Cantor hat eine stetige, schwach monotone Funktion cantor : $[0,1] \mapsto [0,1]$ definiert mit folgenden Eigenschaften:

$$\mathrm{cantor}(0) = 0,\ \mathrm{cantor}(1) = 1,\ \mathrm{cantor}'(\mathrm{x}) = 0 \text{ “fast überall”}.$$

Hier bedeutet *fast überall*, dass die Teilmenge von [0, 1], auf der die Ableitung **nicht** definiert ist, eine sogenannte Nullmenge ist. Solche Ausnahmemengen sind so klein, dass sie die Integrierbarkeit und das Integral einer Funktion nicht beeinträchtigen. Daher ist $\int_0^1 \mathrm{cantor}'(\mathrm{x})\mathrm{dx} = 0$. Mit anderen Worten: *Die Funktion* cantor *kann nicht als Integral ihrer integrierbaren Ableitung rekonstruiert werden.* Die (spärlichen) Ausnahmestellen, an denen die Ableitung nicht definiert ist, verhindern die Gültigkeit des Monotoniesatzes.

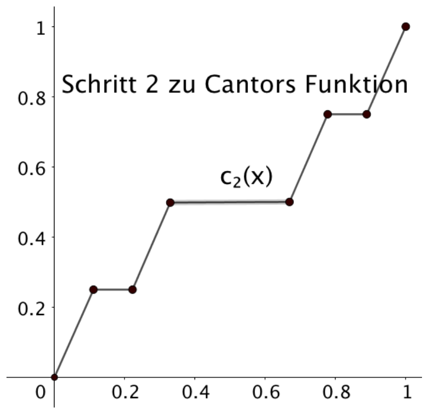
Cantors Funktion wird als Grenzwert einer Folge c_n stetiger, schwach monotoner Funktionen definiert. Der Anfang ist:

$$c_1(x) := \begin{cases} 1.5 \cdot x & \text{für } 0 \le x \le 1/3 \\ 0.5 & \text{für } 1/3 \le x \le 2/3 \\ 0.5 + 1.5 \cdot (x - 2/3) & \text{für } 2/3 \le x \le 1 \end{cases}$$



Danach werden die Funktionen auf immer mehr (aber immer kürzeren) Teilintervallen konstant. Auf den Konstanzintervallen ändern sich bei späteren Schritten die Werte nicht mehr:

$$c_{n+1}(x) := \begin{cases} 0.5 \cdot c_n(3x) & \text{für } 0 \le x \le 1/3 \\ 0.5 & \text{für } 1/3 \le x \le 2/3 \\ 0.5 + 0.5 \cdot c_n(3x - 2) & \text{für } 2/3 \le x \le 1 \end{cases}$$



Die Funktion c_1 ist auf Intervallen der Gesamtlänge 2/3 nicht konstant. Die Funktion c_n ist auf (endlich vielen) Intervallen der Gesamtlänge $(2/3)^n$ nicht konstant. Die Funktion cantor ist auf unendlich vielen Intervallen der Gesamtlänge 1 (mit unterschiedlichen Werten je Intervall) konstant.

Ab hier eher kein Thema für die Schule, aber Illustration der Kraft des Monotoniesatzes.

Es gibt Situationen, in denen Voraussetzungen an die vierte Ableitung natürlich sind. Die dann notwendige mehrmalige Anwendung des Monotoniesatzes muss mit Argumenten kombiniert werden, die an Kurvendiskussionen erinnern. Erstes Beispiel:

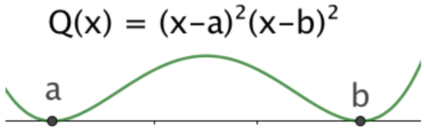
Kubische Interpolation

Voraussetzung: f sei auf $[a, b]$ viermal differenzierbar und $x \in [a, b] \Rightarrow |f^{(4)}(x)| \leq D$.
P sei das kubische Polynom mit
$f(a) = P(a),\ f'(a) = P'(a),\ f(b) = P(b),\ f'(b) = P'(b)$.

Behauptung: $x \in [a, b] \Rightarrow |f(x) - P(x)| \leq \dfrac{D}{384} \cdot (b - a)^4$

Der Hauptteil des Beweises besteht in folgendem **Hilfssatz**:

Voraussetzung: h sei auf $[a, b]$ viermal differenzierbar und $h^{(4)} \geq 0$
Ferner sei $h(a) = h'(a) = h(b) = h'(b) = 0$.



Behauptung: $x \in [a, b] \Rightarrow h(x) \geq 0$.

Bemerkung: Hat man z.B. zehnstellige Sinus- und Kosinustabellen in 1° Abstand, dann liefert diese Interpolation die Zwischenwerte mit einem Fehler $< 2.5 \cdot 10^{-10}$, also TR-konkurrenzfähig.

Beweis.
Schritt 1: Es genügt, die Behauptung unter der Voraussetzung $h^{(4)} > 0$ zu beweisen. Denn, falls h in $[a, b]$ auch negative Werte hat, so kann man zu h die Funktion $p(x) := \epsilon \cdot (x - a)^2 (x - b)^2$ mit so kleinem ϵ addieren, dass die Summe $h + p$ weiterhin auch negative Werte hat, aber $(h + p)^{(4)} > 0$ ist.
Schritt2: Zunächst folgt aus $h^{(4)} > 0$, dass $h^{(3)}$ streng wachsend ist und h'' streng konvex. Daher ist h nicht konstant und wegen $h(a) = h(b) = 0$ muss h' in $[a, b]$ positive und negative Werte haben, also mindestens ein positives Maximum und ein negatives Minimum. An diesen Extremstellen muss h'' null sein. Da eine streng konvexe Funktion höchstens zwei Nullstellen haben kann, folgt jetzt: h'' hat genau zwei Nullstellen $n_1, n_2 \in (a, b)$. Zwischen diesen Nullstellen ist h'' negativ und außerhalb positiv.
Schritt 3: Daher ist die Funktion h konvex in $[a, n_1]$, $[n_2, b]$ und h ist konkav in $[n_1, n_2]$. Konvexe Funktionen sind oberhalb ihrer Tangenten und die x-Achse ist Tangente in a und b. Deshalb ist $h \geq 0$ in $[a, n_1]$ und $[n_2, b]$. Konkave Funktionen sind oberhalb ihrer Sehnen und die Sehne über $[n_1, n_2]$ hat Endwerte ≥ 0. Deshalb gilt $h \geq 0$ auch in $[n_1, n_2]$.

Wir verbessern den Hilfssatz zu der behaupteten expliziten Abschätzung für $f - P$. Definiere das Polynom $Q(x) = \frac{D}{24}(x - a)^2 (x - b)^2$. Dann erfüllen die beiden Funktionen $h := Q \pm (f - P)$ die Voraussetzungen des Hilfssatzes, wir haben also $|f - P| \leq Q$. Außerdem ist $Q((a + b)/2) = \frac{D}{24} \cdot \frac{(b-a)^4}{16}$ das Maximum von Q in $[a, b]$. Das beweist die am Anfang behauptete Abschätzung $|f - P| \leq \frac{D}{384}(b - a)^4$.

Auf sehr ähnliche Weise bekommt man als zweites Beispiel die auf Seite 4 zitierte

Fehlerabschätzung der Simpson Regel.

Diese numerische Integrationsformel integriert eine Funktion f auf dem Intervall $[a, b]$ so:

$$\int_a^b f(x)\,dx \approx \big(f(a) + 4f(\frac{a+b}{2}) + f(b)\big) \cdot (b-a)/6.$$

Die rechte Seite ist das Integral desjenigen kubischen Polynoms P, das definiert ist durch

$$P(a) = f(a),\ P(\frac{a+b}{2}) = f(\frac{a+b}{2}),\ P(b) = f(b),\ P'(\frac{a+b}{2}) = f'(\frac{a+b}{2}).$$

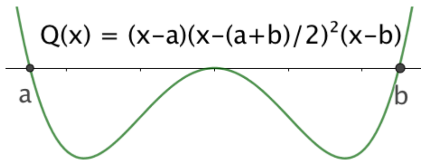
Behauptung: Eine Schranke $|f^{(4)}| \leq D$ liefert die Fehlerabschätzung

$$\left|\int_a^b (f-P)(x)\,dx\right| \leq \frac{D}{2880}(b-a)^5.$$

Falls f nicht viermal differenzierbar ist, muss diese Simpson Integration **nicht** besser sein als die einfacheren Trapez-Regeln.

Zum Beweis verwenden wir wieder ein Polynom Q mit denselben Nullstellen wie $f - P$ und $Q^{(4)} = D$ nämlich

$$Q(x) := \frac{D}{24}(x-a)(x-\frac{a+b}{2})^2(x-b).$$



Für die Funktion $h := Q \pm (f - P)$ gilt ein ähnlicher **Hilfssatz** wie im ersten Beispiel:

Voraussetzung: $h^{(4)} \geq 0,\ h(a) = h(\frac{a+b}{2}) = h(b) = h'(\frac{a+b}{2}) = 0.$

Behauptung : $x \in [a, b] \ \Rightarrow\ h(x) \leq 0,$

und daraus: $$\left|\int_a^b (f-P)(x)\,dx\right| \leq -\int_a^b Q(x)\,dx = \frac{D}{2880}(b-a)^5.$$

Beweis:

Schritt 1. Falls die Funktion h keine positiven Werte hat, ist nichts zu beweisen. Andernfalls können wir $\epsilon \cdot Q < 0$ addieren, so dass weiter positive Werte $h(x) > 0$ vorkommen und die Ungleichung $h^{(4)} > 0$ strikt ist.

Schritt 2. Wegen $h^{(4)} > 0$ ist h'' streng konvex, hat also höchstens zwei Nullstellen. Außerdem sind weder h noch h' stückweise konstant. Wegen $h(a) = h(\frac{a+b}{2}) = h(b) = 0$ muss h' sowohl bei $n_1 \in (a, \frac{a+b}{2})$ als auch bei $n_2 \in (\frac{a+b}{2}, b)$ einen Vorzeichenwechsel haben. Wegen $h'(\frac{a+b}{2}) = 0$ haben wir (mindestens) drei Nullstellen von h' und damit genau zwei Nullstellen von h'' bei $n_3 \in (n_1, \frac{a+b}{2})$ und bei $n_4 \in (\frac{a+b}{2}, n_2)$. Die konvexe Funktion h'' ist also positiv in (a, n_3), (n_4, b) und negativ in (n_3, n_4).

Schritt 3. Die Funktion h ist also konvex in $[a, n_3]$, $[n_4, b]$ und konkav in $[n_3, n_4]$. Konkave Funktionen sind unterhalb ihrer Tangenten und die x-Achse ist Tangente in $\frac{a+b}{2}$, also $h \leq 0$ in $[n_3, n_4]$. Eine konvexe Funktion ist unterhalb ihrer Sehnen und die Sehnen über $[a, n_3]$ und $[n_4, b]$ haben Endwerte ≤ 0. Daher folgt $h \leq 0$ auch in diesen Intervallen.

Differentialgleichungen kommen zwar auf der Schule nicht vor, aber sie sind seit Newton das Werkzeug zur Formulierung von Naturgesetzen. Diese Rolle könnten sie ohne einen Eindeutigkeitssatz nicht spielen.

Der wichtigste

Eindeutigkeitssatz für gewöhnliche Differentialgleichungen

wird ohne zusätzliche Begriffe unmittelbar aus dem Monotoniesatz hergeleitet werden.

Wie kann man sich eine Differentialgleichung (DGL) vorstellen?

Die Definition der Differenzierbarkeit einer Funktion f an der Stelle $x = a$ verlangt, dass für **alle** x aus einer (eventuell sehr kleinen) Umgebung der Stelle a die Funktionswerte $f(x)$ bekannt sind. Das fällt nicht immer auf, weil die am besten bekannten Funktionen durch Formeln für ihre Funktionswerte gegeben sind, so dass man automatisch mit "allen" Werten rechnet. Im allgemeinen kann man nicht allein aus dem Funktionswert $f(a)$ die Ableitung $f'(a)$ berechnen.

Bei der Formulierung der Naturgesetze treten nun Funktionen auf, für die es doch möglich ist, aus a und $f(a)$ die Ableitung $f'(a)$ zu berechnen. Zu der Funktion f gibt es dann eine Abbildung $F : (x, y) \to F(x, y)$, ein "Naturgesetz", das diese Berechnung leistet: $f'(x) = F(x, f(x))$. Auf diese Weise bestimmt die Funktion f (auch "Lösung" der DGL genannt) an jeder Stelle mit Hilfe des Naturgesetzes, wie sie sich ändert. Anschaulich versteht man das am besten, wenn man sich die unabhängige Variable x als "Zeit" vorstellt:

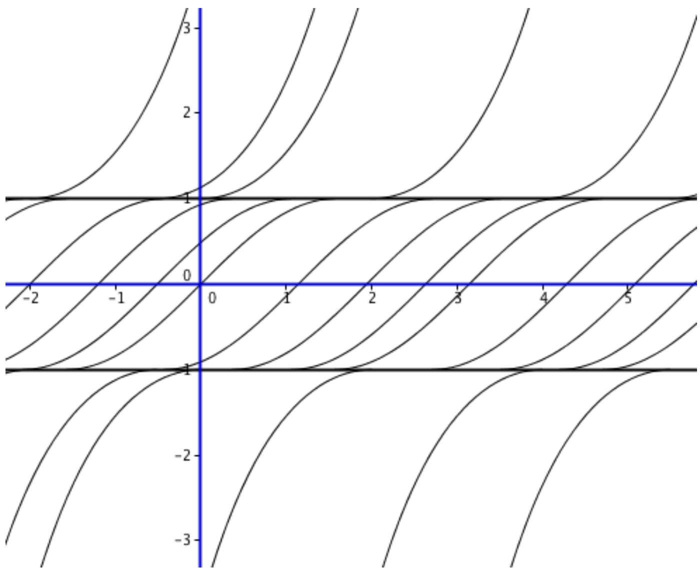
Die zeitliche Entwicklung der Funktion f ist durch das Naturgesetz F festgelegt.

Daher können mit solchen Funktionen f Vorhersagen gemacht werden. In der Mathematik betrachtet man nun solche Differentialgleichungen $f'(x) = F(x, f(x))$ auch ohne dass F durch ein Naturgesetz gegeben ist. Das folgende Beispiel zeigt, dass der Verlauf von f **nicht** durch einen Startwert $f(a)$ festgelegt sein muss. Wegen der naturwissenschaftlichen Anwendungen ebenso wie aus mathematischer Neugier ist man an Eigenschaften von F interessiert, aus denen folgt, dass Lösungen der DGL $f'(x) = F(x, fa(x))$ durch einen Startwert *eindeutig bestimmt* sind. Der folgende Satz ist ein solcher Eindeutigkeitssatz.

Hier ist ein **Beispiel** von Lösungskurven einer DGL, die **nicht** durch Anfangswerte eindeutig bestimmt sind. Die Geraden $y = 1$ und $y = -1$ sind ebenfalls Lösungen, sie heißen "singuläre Lösungen". Sie erlauben, die Lösungskurvenstücke innerhalb des Streifens $-1 \le y \le 1$ mit den Lösungskurven außerhalb dieses Streifens durch mehr oder weniger lange Stücke dieser singulären Lösungen zu verbinden.

Die zu dem Bild gehörende Differentialgleichung lautet:

$$f'(x) = +\big((f(x)^2 - 1)^2\big)^{1/4}.$$



Der bekannteste Eindeutigkeitssatz verlangt eine "Lipschitzbedingung"

Voraussetzung des Eindeutigkeitssatzes: Für die die DGL definierende Funktion F gibt es eine ("Lipschitz")-Konstante L, so dass für alle Argumente gilt

$$|F(x, y_2) - F(x, y_1)| \le L \cdot |y_2 - y_1|.$$

Vor dem Beweis fügen wir ein **Illustrationsbeispiel** ein. Gilt für eine Funktion f die Abschätzung $|f'| \le L \cdot |f|$, so ist die Hilfsfunktion $h(x) := f^2(x) \cdot \exp(-2L \cdot x)$ allein wegen des **Monotoniesatzes** schwach fallend (also $0 \le h(x) \le h(0)e^{2Lx}$), denn $f \cdot f' \le L \cdot f^2$ und daher

$$h'(x) = 2f(x) \cdot f'(x) \cdot e^{-2L \cdot x} - 2L \cdot f^2(x) \cdot e^{-2L \cdot x} \quad \le \quad (2L - 2L) \cdot h(x) = 0.$$

Beweis des Eindeutigkeitssatzes: Gegeben seien zwei Lösungen f_1, f_2 der DGL, also $f_1'(x) = F(x, f_1(x)),\ f_2'(x) = F(x, f_2(x))$.
Dann gilt für die Differenzfunktion $g(x) := f_2(x) - f_1(x)$:

$$|g'(x)| = |F(x, f_2(x)) - F(x, f_1(x))| \quad \le \quad L \cdot |f_2(x) - f_1(x)| = L \cdot |g(x)|.$$

Das Illustrationsbeispiel liefert also $|f_2(x) - f_1(x)| \le |f_2(0) - f_1(0)| \cdot e^{Lx}$.
Diese Ungleichung kontrolliert das Anwachsen anfänglicher Unterschiede und liefert die eindeutige Bestimmtheit durch die Startwerte, denn aus $f_2(0) = f_1(0)$ folgt $f_2(x) = f_1(x)$.

Bemerkung. Es gibt eine ganze Reihe von Fällen, in denen es leichter ist eine Lösung einer DGL zu raten als solch eine Lösung mit einem Lösungsverfahren zu bestimmen. Ein Eindeutigkeitssatz gibt dann die bequeme Versicherung: *Neben der geratenen Lösung $f(x)$ gibt es* **keine** *weitere Lösung $g(x)$, so dass an irgendeiner Stelle a gilt $f(a) = g(a)$!* Deshalb ist ein Eindeutigkeitssatz auch schon nützlich, bevor man einen Existenzsatz für Lösungen hat.