Mathematik für Wirtschaftswissenschaftler

Vorlesungsprogramm für den 10. 05. 2007

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

4.5 Extremstellenbestimmung bei Funktionen einer Variablen

Der Zweck der Differentialrechnung ist nicht, für alle möglichen mehr oder weniger komplizierten Funktionen f die Ableitung "zum Spaß" auszurechnen — machbar ist das ja für alle elementaren Funktionen —, sondern aus der Untersuchung der Ableitung f' und evtl. auch noch höherer Ableitungen f'', f''', ... Rückschlüsse auf die Funktion f selbst zu ziehen. Die vielleicht überzeugendste Anwendung der Differentialrechnung in diesem Sinne ist das Verfahren zur Extremstellenberechnung für eine reelle Funktion mitttels Nullstellenbestimmung bei der Ableitung f'. Um dieses Verfahren anwenden zu können, muss man erstens die Ableitung f' ausrechnen können; das geht mit den Ableitungen der Grundfunktionen und den Rechenregeln für die Ableitung aus 4.3 für die in der Wirtschaftsmathematik vorkommenden Funktionen immer. Zweitens muss man die Nullstellen von f' bestimmen können; das ist explizit nur bei sehr einfach aufgebauten Funktionen möglich. Näherungsweise, mit der Intervallhalbierungsmethode oder mit fortgeschritteneren numerischen Verfahren, ist die Nullstellenbestimmung aber immer mit der erforderlichen Genauigkeit durchführbar, während es zur direkten Extremstellenbestimmung bei einer gegebenen Funktion kein gutes Näherungsverfahren gibt (weil man ohne nähere Informationen über die Funktion aus der Berechnung ihrer Werte an endlich vielen Stellen in einem Intervall keinerlei Rückschlüsse auf die Lage der Extremstellen in diesem Intervall ziehen kann). Zunächst zur Terminologie:

DEFINITIONEN und DISKUSSION (*Extremstellen und -werte*):

1) Für eine gegebene reelle Funktion $f : \mathbb{R} \supset D \to \mathbb{R}$ heißt eine Stelle $x_* \in D$ **(absolute) Minimumstelle** von f auf D, wenn f an keiner anderen Stelle in D einen kleineren Funktionswert hat, wenn also gilt $f(x_*) \le f(x)$ für alle $x \in D$. Der Funktionswert an einer solchen Stelle heißt dann das **(absolute) Minimum** der Funktion f auf dem Definitionsbereich D, und man sagt, dass f auf D an der Stelle x_* **ein Minimum annimmt**. Wird das Minimum allein an der Stelle x_* angenommen, gilt also $f(x) > f(x_*)$ für alle $x \in D$ mit $x \ne x_*$, so heißt x_* die **eindeutige Minimumstelle** zu f auf D, und wir sagen, dass f ein **einziges Minimum** auf D an der Stelle x_* hat. Entsprechend heißt $x^* \in D$ eine **(absolute) Maximumstelle** von f auf I, und f **nimmt ein Maximum an** im Punkt x^* auf D, wenn $f(x^*) \ge f(x)$ ist für alle $x \in D$. Der Wert $f(x^*)$ heißt dann das **(absolute) Maximum** der Funktion f auf D, und f hat auf D ein **einziges Maximum** im Punkt x^*, wenn sogar $f(x) < f(x^*)$ gilt für alle $x \in D$ mit $x \ne x^*$, in welchem Falle x^* die **eindeutige Maximumstelle** von f auf D genannt wird. Der Sammelbegriff für Maximum- und Minimumstellen ist **(absolute) Extremstellen**, und das Maximum wie das Minimum nennt man einen **(absoluten) Extremwert** oder **Extremum** oder **Optimum** von f auf D. Als Notation für das Minimum und Maximum von f auf D verwendet man, wenn sie existieren:

$$\max_{x\in D} f(x)\,, \quad \min_{x\in D} f(x)\,, \qquad \text{oder kurz} \qquad \max_D f\,, \quad \min_D f\,.$$

2) Man sollte

- *die Extremstellen von den Extremwerten (Extrema) unterscheiden!*

Normalerweise interessieren bei einer Extremwertaufgabe nicht nur die Extremwerte der untersuchten Funktion f, sondern auch und vor allem die Stellen im Definitionsbereich, an denen diese Werte angenommen werden. Man möchte eben z.B. wissen, wie man den Produktions–Output x wählen muss, damit der Gewinn $G(x)$ maximal ausfällt, und nicht nur, wie groß der maximale Gewinn ist. Kennt man die Extremstellen, so kann man die Extremwerte einfach durch Auswerten der Funktion in diesen Punkten (Einsetzen in die Funktion) ausrechnen.

3) Wenn von Maximum und Minimum die Rede ist, so sollte man sich immer zunächst vor Augen halten:

- *Extremstellen brauchen nicht zu existieren; und wenn sie existieren, so brauchen sie nicht eindeutig zu sein!*

Das Letzte ist klar; eine Funktion kann natürlich ihr Maximum und / oder Minimum an mehreren verschiedenen Stellen annehmen (Extremes Beispiel: Konstante Funktion). Bzgl. der Existenz haben wir die Lage schon in 4.2 analysiert: Was *immer* existiert, wenn wir dafür auch nötigenfalls die Werte ∞ oder $-\infty$ zulassen, ist das *Infimum* $\inf_D f$ und das *Supremum* $\sup_D f$, also die größte untere Schranke und die kleinste obere Schranke für die Funktionswerte von f auf D. Wenn $\sup_D f = \infty$ ist, so nimmt die Funktion beliebig große Werte auf der Menge D an; dann kann es natürlich kein Maximum, keinen größten Funktionswert, geben; denn ∞ ist kein zugelassener Funktionswert. *Maximum und Minimum sind immer endliche Werte, soweit sie existieren.* Ebenso kann kein Minimum von f auf D existieren, wenn $\inf_D f = -\infty$ ist, wenn also die Funktion negative Werte beliebig großen Betrags annimmt. Aber auch wenn Supremum und Infimum endlich sind, gibt es im allgemeinen keine Extrema, weil Supremum bzw. Infimum nicht als Funktionswerte auf D angenommen werden müssen. (Genau wenn es ein Funktionswert auf D ist, ist das Supremum gleich dem Maximum bzw. das Infimum gleich dem Supremum.) Es kann z.B. sein, dass die Funktionswerte nur bei Annäherung an nicht zu D gehörende Randpunkte gegen Supremum oder Infimum streben, aber diese Werte auf D nie erreichen. Oder die Funktion ist vielleicht an den Stellen in D, "wo eigentlich die Extremstellen sein müssten", unstetig und hat dort nicht die "richtigen" Werte. Simple Beispiele für diese Situation haben wir in 4.2 schon gesehen. Aus 4.2 rekapitulieren wir auch den wichtigsten Satz über die Existenz von Extremstellen:

- **Satz vom Maximum und Minimum – Extremstellensatz:** *Auf einem kompakten Intervall $I = [a, b]$ in $\mathbb{R}$ hat jede stetige reellwertige Funktion mindestens eine (absolute) Minimumstelle und mindestens eine (absolute) Maximumstelle.*

4) Bei der Anwendung des Extremstellensatzes ist die Stetigkeit der Funktion normalerweise kein Problem, weil ja alle elementaren Funktionen, also alle durch eine konkrete Formel gegebenen Funktionen, stetig sind (sogar beliebig oft differenzierbar). Auch ist der Bereich D, zu dem man die Extrema der Funktion bestimmen will, meistens ein Intervall, allerdings nicht immer ein kompaktes. Wir erinnern an diese Begriffsbildung aus 1.4: Ein kompaktes Intervall ist ein beschränktes Intervall, das seine Endpunkte a, b enthält, wobei wir hier immer $a < b$ annehmen. (Das leere Intervall schließen wir als Definitionsbereich grundsätzlich aus, und einpunktige Intervalle $[a, a]$ sind auch nicht besonders interessant.)

Es kommt aber vor, dass man Extremstellen zu einer Funktion auf einem nicht–kompakten Intervall bestimmen muss, z.B. auf einem unbeschränkten Intervall mit ∞ oder $-\infty$ als "uneigentlichem Randpunkt" wie $\mathbb{R} =]-\infty, \infty[$ oder $[0, \infty[$, oder auf einem beschränkten Intervall, das nicht seine beiden Randpunkte enthält wie $]0,1[$ oder $[0,1[$. Hier gilt der

- **Zusatz zum Extremstellensatz:** *Ist das Definitionsintervall I der stetigen Funktion f nicht kompakt, und hat die Funktion in der Nähe der nicht zu I gehörenden (evtl. unendlichen) Randpunkte nur Werte, die größer [bzw. kleiner] sind als ein Funktionswert auf I, so nimmt f auf I ein absolutes Minimum [Maximum] an.*
- *Das gilt insbesondere dann, wenn die Funktion den Grenzwert $+\infty$ [bzw. $-\infty$] hat bei Annäherung aus I an die nicht zu I gehörenden (evtl. unendlichen) Randpunkte.*

Die Voraussetzung hier bedeutet nämlich, dass es ein kompaktes Intervall $[a,b] \subset I$ und eine Stelle $x_0 \in [a,b]$ gibt, derart dass $f(x) > f(x_0)$ ist für alle Punkte x aus I, die außerhalb von $[a,b]$ liegen. Dann ist aber das Minimum von f auf dem kompakten Intervall $[a,b]$ offenbar auch schon das Minimum von f auf dem ganzen Intervall I, und alle Minimumstellen von f auf I liegen schon in $[a,b]$. Die Voraussetzung ist insbesondere erfüllt, wenn die Funktion einen unendlichen Grenzwert bei Annäherung an die nicht zu I gehörenden Randpunkte hat. Ist also z.B. $I = \mathbb{R}_{>0}$ und gilt $\lim_{x\to\infty} f(x) = \infty$ sowie $\lim_{x\searrow 0} f(x) = \infty$, so hat f mindestens eine Minimumstelle in I (sofern f stetig ist).

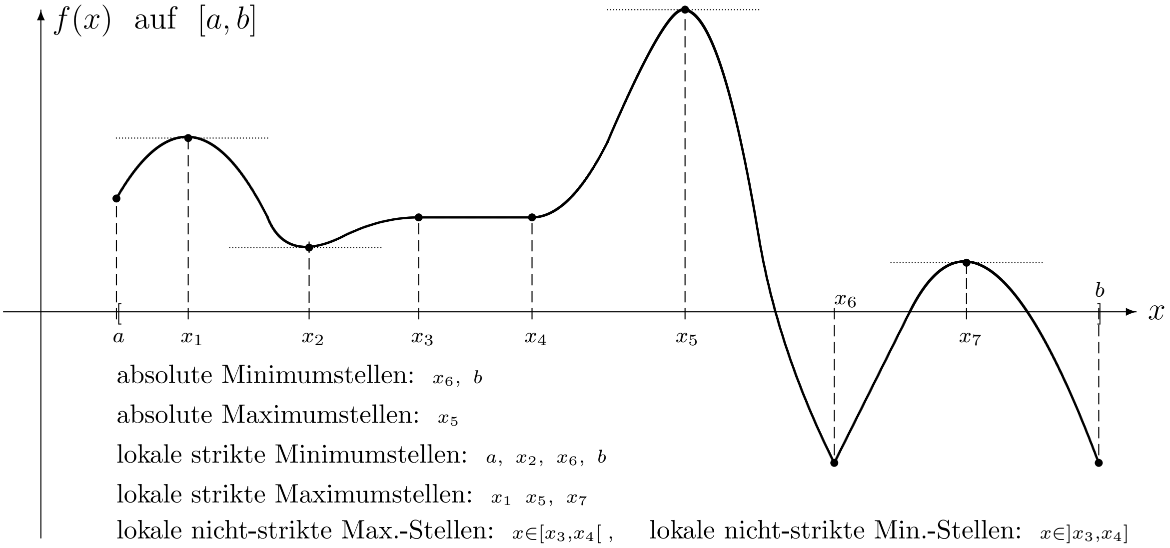
5) In der reellen Analysis ist man oft nicht sehr präzise bei der Angabe von Definitionsbereichen; bei Funktionen, die durch Rechenterme gegeben sind, denkt man sich z.B. immer den maximalen Definitionsbereich des Terms in $\mathbb{R}$. Bei Extremwertaufgaben muss man aber genauer sein; denn

- *absolute Extremstellen und Extremwerte einer Funktion hängen davon ab, welcher Definitionsbereich D, betrachtet wird; daher muss man den Definitionsbereich genau spezifizieren, wenn von Extremstellen oder Extremwerten die Rede ist.*

Eine Maximumstelle zu f auf D ist im Allgemeinen keine mehr zu f auf einem größeren Definitionsbereich $\widetilde{D} \supset D$; denn bei Vergrößerung von D können ja neue Funktionswerte hinzukommen, die größer sein können als das Maximum von f auf D. Aus demselben Grund ist z.B. auch das absolute Maximum zu f auf D im Allgemeinen größer als das Maximum von f auf einem kleineren Definitionsbereich $D' \subset D$, vorausgesetzt die beiden Maxima existieren überhaupt. Nur wenn D' schon eine der Maximumstellen zu f auf D enthält, ist das Maximum von f auf D' dasselbe wie auf D (und dann sind die Maximumstellen zu f auf D' genau diejenigen Maximumstellen von f auf D, die schon in D' liegen).

6) Eine Stelle $x_* \in D$ heißt **lokale Minimumstelle** oder **relative Minimumstelle** zu der reellen Funktion f auf D, wenn $f(x) \geq f(x_*)$ gilt für alle hinreichend nahe bei x_* gelegenen $x \in D$, d.h. genauer, wenn es $\delta > 0$ gibt, so dass x_* absolute Minimumstelle zu f auf der "Umgebung" $D \cap]x_* - \delta, x_* + \delta[$ von x_* in D ist. Der Funktionswert $f(x_*)$ heißt dann ein **relatives Minimum** oder **lokales Minimum** zu f auf D. Kann dabei $\delta > 0$ so gewählt werden, dass x_* die einzige Minimumstelle zu f in der Umgebung ist, dass also gilt $f(x) > f(x_*)$ für alle von x_* verschiedenen Stellen $x \in D$ nahe genug bei x_*, so heißt x_* **lokale strikte Minmimumstelle** oder **lokale isolierte Minimumstelle** von f auf D, und man sagt von der Funktion, dass sie an der Stelle x_* ein **striktes relatives Minimum** oder ein **striktes lokales Minimum** bzgl. D hat. Entsprechend erklärt man **lokale (strikte) Maximumstellen** x^* zu f und **lokale (strikte) Maxima** $f(x^*)$, wobei auch wieder "relativ" statt "lokal" und "isoliert" statt "strikt" gesagt werden kann. Der Sammelbegriff ist **lokale Extremstellen** und **lokale Extremwerte**.

Die folgende Abbildung veranschaulicht die Begriffsbildungen:



Minimum- bzw. Maximumstellen von f auf D sind natürlich auch relative Minimum- bzw. Maximumstellen; das Umgekehrte ist jedoch im Allgemeinen nicht richtig. Um die Extremstellen auf D von den (nur) relativen Extremstellen abzuheben, verwenden wir das Attribut “*absolut*”. Man spricht auch von **globalen Extremstellen** zu f auf D, statt von absoluten Extremstellen, weil “lokal – global” ein Gegensatzpaar ist. Wenn von Extremstellen oder -werten ohne Attribut die Rede ist, so meinen wir immer die absoluten.

■

Es ist anschaulich klar, dass die Tangente an den Graphen in lokalen Extremstellen horizontal sein muss, wie es in der Abbildung auch angedeutet ist. Mit anderen Worten: An solchen Stellen muss die Ableitung der Funktion verschwinden; denn das ist ja die Steigung der Tangente. Diese einfache Überlegung ist die geometrische Grundlage des notwendigen Kriteriums erster Ordnung für Extremstellen, das den ersten Teil des folgenden Satzes bildet. Der zweite Teil ist ein hinreichendes Kriterium, mit dem man Stellen mit Steigung Null als lokale, unter Umständen sogar als globale Extremstellen identifizieren kann.

SATZ (Ableitungskriterien erster Ordnung für Extremstellen):

(i) (notwendiges Kriterium) *Hat die Funktion $f : I \to \mathbb{R}$ eine lokale oder gar absolute Extremstelle im inneren Punkt x_0 des Intervalls I (kein Randpunkt) und ist f dort differenzierbar, so gilt notwendigerweise*

$$\boxed{f'(x_0) = 0} \, .$$

(ii) (hinreichendes Kriterium) *Die Funktion $f : I \to \mathbb{R}$ sei differenzierbar mit einer Ableitungsnullstelle $f'(x_0) = 0$ in einem inneren Punkt x_0 des Intervalls I. Hat die Ableitung f' bei x_0 einen Vorzeichenwechsel vom Typ $-/+$, so liegt in x_0 eine lokale strikte Minimumstelle vor, hat sie dagegen einen Vorzeichenwechsel vom Typ $+/-$, so liegt dort eine lokale strike Maximumstelle vor.*

(iii) (hinreichendes Sattelpunktkriterium) *Hat in (ii) die Ableitung an der Stelle x_0 eine isolierte Nullstelle ohne Vorzeichenwechsel, so liegt dort ein strikter Sattelpunkt von f vor und jedenfalls keine lokale oder gar absolute Extremstelle; f ist sogar streng monoton auf einem Intervall um x_0.*

Da in (ii) nichts über die Größe des (evtl. sehr kleinen) Intervalls um x_0 gesagt wird, auf dem x_0 absolute Minimum- bzw. Maximumstelle ist, ist folgender Zusatz noch wichtig:

ZUSATZ: (iv) (Monotonie zwischen den lokalen Extremstellen) *Auf Intervallen ohne innere lokale Extremstellen ist eine stetige Funktion f streng monoton.*

(v) (Bereich absoluter Extremalität) *Daher ist eine innere lokale isolierte Minimumstelle [Maximumstelle] x_0 der stetigen Funktion $f : I \to \mathbb{R}$ eindeutige absolute Minimumstelle [Maximumstelle] von f mindestens auf dem größten Intervall J in I mit $x_0 \in J$, das keine andere lokale Extremstelle ($\neq x_0$) im Inneren hat.*

Das bedeutet also, dass eine isolierte lokale Minimumstelle x_0 von f auf I eine absolute Minimumstelle auf dem Intervall $J \subset I$ ist, das sich von x_0 aus nach links und rechts bis zur nächsten lokalen Extremstelle erstreckt oder, wenn keine solche Stelle kommt, bis zum nächsten Randpunkt von I oder, wenn kein Randpunkt kommt, sogar bis $-\infty$ bzw. bis ∞. Insbesondere ist x_0 absolute Minimumstelle von f auf ganz I, wenn f außer x_0 in I keine inneren lokalen Extremstellen mehr hat. Bei komplizierten (nichtelementaren) stetigen Funktionen ist es möglich, dass sich lokale Extremstellen von rechts und links gegen x_0 häufen; dann ist $J = \{x_0\}$ und die Aussage (ii) leer. Bei elementaren Funktionen jedoch liegen die lokalen Extremstellen isoliert, und x_0 hat positiven Abstand zur nächsten lokalen Extremstelle nach links und nach rechts (wenn eine kommt, bevor I endet).

Wir geben einen *Beweis* des Satzes und des Zusatzes und erklären dabei auch (obwohl es sich fast von selbst versteht), was unter einem "Vorzeichenwechsel" zu verstehen ist. Gilt $f'(x_0) > 0$ an einer inneren Stelle x_0, so sind auch alle Differenzenquotienten $\frac{f(x)-f(x_0)}{x-x_0}$ positiv für x hinreichend nahe bei x_0; denn diese Quotienten liegen ja beliebig nahe bei $f'(x_0)$ für x nahe genug bei x_0. Dies bedeutet aber, dass insbesondere $f(x) > f(x_0)$ ist für $x > x_0$ nahe genug bei x_0 und $f(x) < f(x_0)$ für $x < x_0$ nahe bei x_0; also hat dann f gewiss keine lokale Maximum- oder Minimumstellestelle in x_0. Dasselbe gilt natürlich im Fall $f'(x_0) < 0$ (argumentiere analog oder gehe einfach zu $-f$ über). Damit ist die Aussage (i) schon bewiesen: Ist x_0 eine innere lokale Extremstelle und f dort differenzierbar, so kann $f'(x_0)$ weder positiv noch negativ sein, also gilt $f'(x_0) = 0$.

Für (ii) betrachten wir die Situation, dass f' bei x_0 einen **Vorzeichenwechsel vom Typ $-/+$** hat, d.h. natürlich, dass $f'(x) > 0$ ist auf einem kleinen Intervall $]x_0, x_0 + \delta[$ in I rechts von x_0 (mit $\delta > 0$) und $f'(x) < 0$ auf einem kleinen Intervall $]x_0 - \delta, x_0[$ in I links von x_0. Für $x_0 < x_1 < x_0 + \delta$ ist dann $f'(x_1) > 0$ und daher, wie wir ja gerade überlegt hatten, $f(x) > f(x_1)$ für $x > x_1$ mit x nahe genug bei x_1. Es muss dann aber sogar $f(x_2) > f(x_1)$ sein für alle $x_1 < x_2 < x_0 + \delta$; denn andernfalls hätte f auf dem kompakten Intervall $[x_1, x_2]$ eine Maximumstelle, die kein Randpunkt sein kann, und dort wäre eine Nullstelle von f' im Widerspruch zu $f' > 0$ auf $]x_0, x_0 + \delta[$. Damit ist gezeigt, dass f auf $]x_0, x_0+\delta[$ streng wachsend ist, und wegen Stetigkeit von f an der Stelle x_0 folgt dasselbe auf $[x_0, x_0+\delta[$. Insbesondere ist $f(x) > f(x_0)$ für $x_0 < x < x_0+\delta$. Genau so sieht man (oder betrachte $f(-x)$ statt $f(x)$), dass auch $f(x) > f(x_0)$ ist für $x_0 - \delta < x < x_0$, und daher ist x_0 einzige Minimumstelle von f auf dem Intervall $]x_0 - \delta, x_0 + \delta[\subset I$. Bei einem **Vorzeichenwechsel vom Typ $+/-$** von f' bei x_0, d.h. f' ist positiv links von x_0 und negativ rechts von x_0 in einer genügend kleinen Umgebung, schließt man ganz analog (oder indem man zu $-f$ übergeht), dass x_0 lokale isolierte Maximumstelle ist.

Bei einer **isolierten Nullstelle ohne Vorzeichenwechsel** x_0 von f' unterscheidet man den **Typ +/+** und den **Typ −/−**; der erste Typ bedeutet natürlich, dass f' positiv ist beiderseits von x_0, der zweite, dass f' negativ ist beiderseits von x_0, jeweils in hinreichender Nähe zu x_0 jedenfalls. Dafür zeigt die obige Argumentation, dass f auf einem ganzen Intervall $]x_0-\delta, x_0+\delta[$ um x_0 in I streng wachsend ist beim Typ +/+ bzw. streng fallend beim Typ −/−. Insbesondere ist x_0 ein **strikter Sattelpunkt**, d.h. $f'(x_0)=0$ und $f(x)$ ist für x nahe x_0 auf einer Seite von x_0 größer und auf der anderen Seite kleiner als $f(x_0)$. Dort kann keine innere lokale Extremstelle vorliegen, und damit ist auch (iii) und der ganze Satz bewiesen.

Für Teil (iv) des Zusatzes überlegt man sich, wenn f nicht streng monoton ist auf einem Intervall, dass es dann darin $x_1 < x_2 < x_3$ gibt mit $f(x_2) \geq f(x_1)$ und $f(x_2) \geq f(x_3)$ oder mit $f(x_2) \leq f(x_1)$ und $f(x_2) \leq f(x_3)$. Im ersten Fall hat die stetige Funktion f eine innere Maximumstelle in $[x_1, x_3]$, im zweiten Fall eine innere Minimumstelle, und das beweist (iv). Daraus folgt dann (v), weil f auf einem Intervall $[x_0, x_2]$ ohne innere Extremstelle natürlich nur streng wachsend (und nicht fallend) sein kann, wenn x_0 lokale strikte Minimumstelle ist, und auf einem Intervall $[x_1, x_0]$ entsprechend nur streng fallend (und nicht wachsend). Damit ist auch der Zusatz bewiesen. ■

Da die Nullstellen der Ableitung bei der Extremstellenbestimmung eine so große Rolle spielen, hat man dafür besondere Namen eingeführt. “Stationär” bedeutet dabei, dass sich die Funktionswerte bei kleinen Änderungen einer solchen Stelle praktisch nicht ändern.

DEFINITION: Die Nullstellen der Ableitung einer differenzierbaren Funktion f einer Variablen heißen die **kritischen Punkte** oder die **stationären Stellen** von f. Wenn f' in einem kritischen Punkt einen Vorzeichenwechsel hat bzw. eine isolierte Nullstelle ohne Vorzeichenwechsel, so nennen wir den Typ −/+ oder +/− bzw. +/+ oder −/− dieser Nullstelle von f' auch den **Typ des kritischen Punktes**. ■

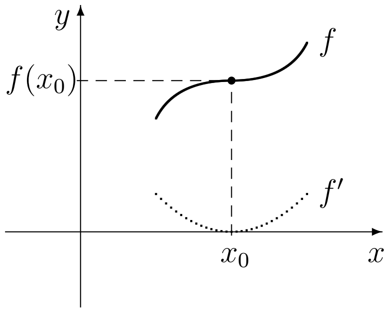
DISKUSSION: **1)** Ist die Funktion f an allen inneren Stellen des Intervalls I differenzierbar, so sagt das notwendige Kriterium:

- *Alle inneren lokalen Extremstellen von f auf I sind unter den kritischen Punkten von f zu finden.*

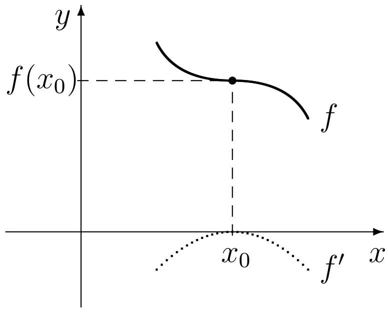
2) *Warnung*: Das notwendige Kriterium für lokale Extremstellen ist aber nicht hinreichend, d.h.

- *innere kritische Punkte müssen aber nicht unbedingt lokale Extremstellen sein!*

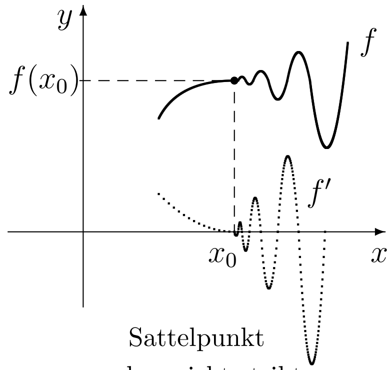
Es kann eben auch **Sattelpunkte** geben, d.h. Ableitungsnullstellen x_0, die keine lokalen Extremstellen sind. Die Funktion nimmt dann an Stellen beliebig nahe bei x_0 sowohl größere Werte als $f(x_0)$ an als auch kleinere Werte. Wenn dabei die Funktion in einer Umgebung $]x_0-\delta, x_0+\delta[$ auf der einen Seite von x_0 echt kleiner als $f(x_0)$ ist und auf der anderen Seite echt größer als $f(x_0)$, so sprechen wir wie oben schon von einem **strikten Sattelpunkt** oder einem **speziellen Sattelpunkt**. Wenn x_0 ein kritischer Punkt vom Typ +/+ oder −/− ist, so liegt z.B. ein spezieller Sattelpunkt vor, das besagt gerade das Sattelpunktkriterium (iii). Im Beweis haben wir sogar gezeigt, dass dann f streng monoton ist auf einer Umgebung; das ist eine noch stärkere Sattelpunkteigenschaft. Dort sieht die Funktion in der Nähe tatsächlich so aus, wie man es bei einem Sattelpunkt erwartet, und andere Typen von Sattelpunkten können bei elementaren Funktionen nicht auftreten. Die folgende Abbildung illustriert die verschiedenen Typen von Sattelpunkten:



strikter Sattelpunkt
Nullstelle bei f' vom Typ $+/+$

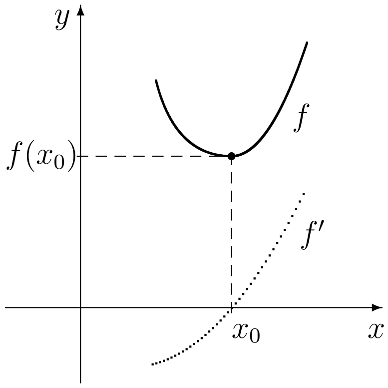


strikter Sattelpunkt
Nullstelle bei f' vom Typ $-/-$

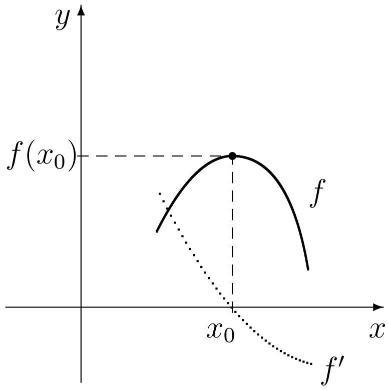


Sattelpunkt
aber nicht strikt

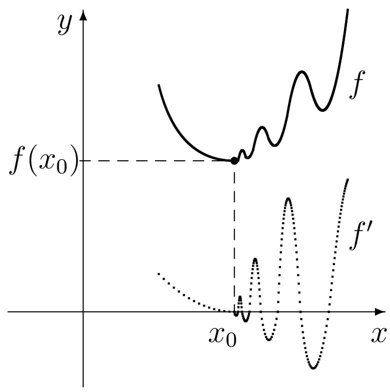
3) Genau so wenig ist das hinreichende Kriterium notwendig: Es gibt lokale strikte Extremstellen x_0, bei denen die Ableitung keinen Vorzeichenwechsel vom Typ $-/+$ oder $+/-$ hat, sondern auf auf einer oder auf beiden Seiten von x_0 beliebig nahe bei x_0 positive und negative Werte annimmt. Dann ist die Funktion nicht, wie es sich beim Beweis von (ii) im Falle von Vorzeichenwechseln ergab, auf beiden Seiten von x_0 streng gegensinnig monoton nahe bei x_0. Allerdings kommt das nur bei nichtelementaren Funktionen vor.



lokale isolierte Min.-Stelle
Nullstelle bei f' vom Typ $-/+$

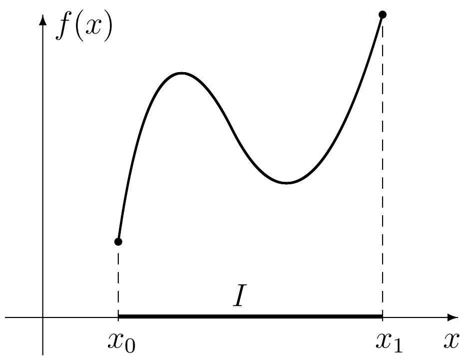


lokale isolierte Max.-Stelle
Nullstelle bei f' vom Typ $+/-$



lokale isolierte Min–Stelle
Nullstelle bei f' von anderem Typ

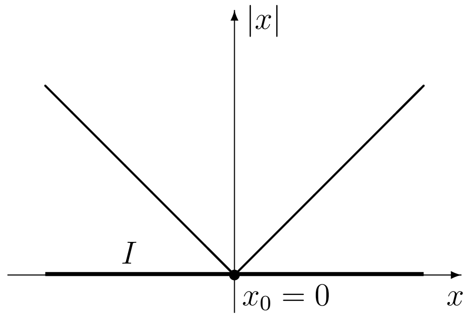
4) In (lokalen oder absoluten) **Randextremstellen** x_0 von f auf I, d.h. x_0 ist Randpunkt des Intervalls I und z.B. $f(x) \geq f(x_0)$ für alle $x \in I$ hinreichend nahe bei x_0 im Fall einer lokalen Randminimumstelle, *gilt das Ableitungskriterium* $f'(x_0) = 0$ *im Allgemeinen nicht*. Man kann hier aber noch etwas über das **Vorzeichen der Ableitung** $f'(x_0)$ sagen, wenn sie existiert. Ist x_0 linker Randpunkt und x_0 lokale Minimumstelle zu f auf I, so kann offenbar nicht $f'(x_0) < 0$ sein (sonst wären die Funktionswerte rechts von x_0 nahe bei x_0 ja kleiner als $f(x_0)$). Entsprechend folgt $f'(x_0) \geq 0$ bei einer rechten lokalen Randmaximumstelle und $f'(x_0) \leq 0$ im Fall einer linken lokalen Randmaximumstelle oder einer rechten lokalen Randminimumstelle. Durch Vorzeichenbetrachtung bei den Ableitungswerten in den zu I gehörenden Randpunkten kann man also diese evtl. aus dem Kandidatenkreis für Maximumstellen oder für Minimumstellen von f auf I eliminieren. Aber das lohnt meistens die Mühe nicht. Grundsätzlich darf man jedoch nicht vergessen:



Randextremstellen x_0 und x_1

- *(Lokale) Extremwerte können auch in Randpunkten angenommen werden, unter Umständen sogar nur in Randpunkten!*

5) Lokale oder absolute Extremstellen können auch in **Nichtdifferenzierbarkeitsstellen** der Funktion liegen, wenn sie nicht überall auf dem betrachteten Intervall differenzierbar ist. Ein einfaches Beispiel ist die Betragsfunktion $f(x) = |x|$ auf einem Intervall I mit innerem Punkt $x_0 = 0$, die offenbar an der "*Knickstelle*" $x_0 = 0$ ihr absolutes Minimum annimmt und nur dort. Das notwendige Kriterium $f'(x_0) = 0$ gilt hier nicht, weil f an dieser Stelle gar nicht differenzierbar ist.



Minimumstelle von $f(x) = |x|$

- *(Lokale) Extremwerte können auch in Nichtdifferenzierbarkeitsstellen angenommen werden, unter Umständen sogar nur dort (wenn die Funktion nicht überall im Grundintevall differenzierbar ist).*

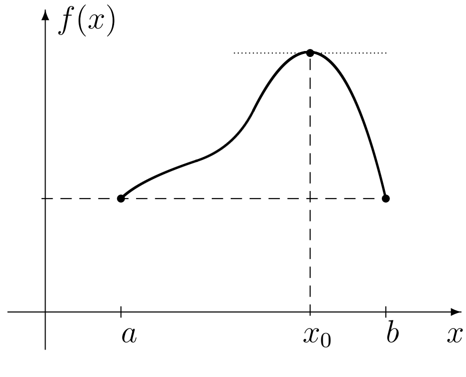
Auch Unstetigkeitsstellen von f in I gehören zu den Nichtdifferenzierbarkeitsstellen. Für die Bestimmung von Maximum und Minimum von f auf I muss dann natürlich klar sein, wie der Funktionswert an solch einer Unstetigkeitsstelle genau definiert ist.

6) Alles Gesagte gilt sinngemäß auch für beliebige Definitionsbereiche $D \subset \mathbb{R}$, nicht nur für Intervalle. Innere Punkte x_0 von D sind dann solche, für die ein ganzes Intervall positiver Länge mit Mittelpunkt x_0 in D enthalten ist. (Man kann sich also von x_0 aus etwas innerhalb von D nach rechts und links bewegen.) Die anderen Punkte von D sind die zu D gehörenden Randpunkte. Es kommt vor, dass ein Definitionsbereich D aus mehreren offenen Intervallen zusammengesetzt ist, die sich nicht überlappen. Bei einer rationalen Funktion mit **Polstellen** z.B. sind das die offenen Intervalle zwischen zwei benachbarten Polstellen sowie zwischen $-\infty$ und der kleinsten Polstelle und zwischen der größten Polstelle und $+\infty$. Das Minimierungssproblem etwa kann man dann auf jedem dieser Intervalle separat betrachten, und wenn es in jedem eine oder mehrere Minimumstellen gibt, so sind die mit dem insgesamt kleinsten Funktionswert die absoluten Minimumstellen der rationalen Funktion auf D. (Dies ist so, genau wenn die rationale Funktion an jeder Polstelle den Grenzwert $+\infty$ hat und bei $x \to \infty$ sowie $x \to -\infty$ Grenzwerte, die nicht kleiner sind als alle Funktionswerte auf dem jeweiligen Intervall mit Endpunkt $+\infty$ bzw. $-\infty$.)

7) Die hauptsächliche Anwendung der Ableitungskriterien für Extremstellen besteht darin, dass man die Extremstellen mit Hilfe der Berechnung der Nullstellen der Ableitung, bestimmen kann. Mit diesem Verfahren werden wir uns gleich intensiv beschäftigen. Hier erwähnen wir noch eine Anwendung in umgekehrter Richtung, nämlich die Existenz von Nullstellen der Ableitung, wenn man innere Extremstellen hat. Paradebeispiel ist der

SATZ von Rolle: *Ist die Funktion f stetig auf dem Intervall $[a, b]$ und differenzierbar im Inneren, und hat sie an den Randstellen a und b denselben Funktionswert, so besitzt die Ableitung f' mindestens eine Nullstelle x_0 echt zwischen a und b.*

Der *Beweis* liegt aufgrund der Abbildung nahe: f hat in dem kompakten Intervall $[a, b]$ Maximumstellen und Minimumstellen, und mindestens eine davon liegt im Innern $]a, b[$, weil wegen $f(a) = f(b)$ andernfalls f konstant wäre. An der inneren Extremstelle verschwindet dann die Ableitung. ■



Beweis des Satzes von Rolle

Die hauptsächliche Anwendung der Ableitungskriterien für Extremstellen ist, wie gesagt, ein Verfahren zur Bestimmung aller lokalen und absoluten Extremstellen einer differenzierbaren Funktion. Wir beschreiben dieses Verfahren jetzt im Detail für den Fall, dass man die absoluten Extremstellen und Extrema einer gegebenen Funktion auf einem gegebenen Intervall bestimmen möchte. Das ist in der Praxis der wichtigste Fall:

VERFAHREN zur Bestimmung der Extremstellen einer gegebenen reellen Funktion f auf einem gegebenen Intervall I in $\mathbb{R}$:

- **Kritische Stellen berechnen**, also die Nullstellen der Ableitung f' (im Inneren von I);
- **vollständige Kandidatenliste erstellen**, in der die kritischen Punkte von f im Inneren von I, die zu I gehörenden Randpunkte von I und gegebenenfalls auch die inneren Nichtdifferenzierbarkeitsstellen von f in I aufgelistet sind; alle lokalen und erst recht alle absoluten Extremstellen von f auf I kommen dann in dieser Liste vor – das ist mit "Vollständigkeit" gemeint;
- **Existenz klären**, d.h. die Frage, ob f auf I überhaupt ein Maximum bzw. Minimum haben muss; wenn f auf I stetig ist, so geht das mit dem Satz vom Maximum und Minimum bei kompaktem Intervall I und evtl. mit zusätzlicher Betrachtung des Grenzverhaltens von f bei den nicht zu I gehörenden oder unendlich fernen Randpunkten, falls I nicht kompakt ist (siehe den am Anfang dieses Abschnitts angegebenen Zusatz zum Extremstellensatz);
- **Wertevergleich durchführen**, d.h. die Funktionswerte von f an den Stellen, die Kandidaten sind, der Größe nach vergleichen:
- **Schlussfolgerung**: *Wenn die Kandidatenliste vollständig ist und wenn das Maximum / Minimum von f auf I existiert, so sind die Kandidaten mit dem größten bzw. kleinsten Funktionswert die absoluten Maximum- bzw. Minimumstellen von f auf I.*

DISKUSSION: **1)** *Warnung*: Ohne die Existenz geklärt zu haben, kann man aus dem Vergleich der Funktionswerte an den Stellen der Kandidatenliste keine Rücksclüsse über die absoluten Extremstellen und Extremwerte von f auf I ziehen! Es gibt nämlich in diesem Fall möglicherweise überhaupt keine absoluten Maximum- oder Minimumstellen, und wenn es keine Extremstellen gibt, so können auch die Kandidaten mit dem größten bzw. kleinsten Funktionswert keine sein! Das einfachste Beispiel ist $f(x) = x^3$ auf $I = \mathbb{R}$; hier hat $f'(x) = 3x^2$ nur die Nullstelle $x_0 = 0$ und die vollständige Kandidatenliste enthält nur diese eine Stelle, aber dennoch nimmt f an dieser Stelle kein Extremum an, nicht einmal lokal; denn die Funktion ist ja streng monoton. (Es handelt sich bei diesem kritischen Punkt um das Paradebeispiel eines strikten Sattelpunktes.) Ein Blick auf das Grenzverhalten von x^3 bei $x \to \pm\infty$ hätte sofort geklärt, dass die Funktion beliebig große und beliebig negative Werte hat und deshalb auf $\mathbb{R}$ keine absolute Extremstelle haben kann. *Der häufigste Fehler* bei der Extremwertberechnung ist, diese Möglichkeit zu übersehen und z.B. eine Stelle als Maximumstelle anzugeben, weil sie von allen Kandidaten den größten Funktionswert hat, obwohl die Funktion noch größere (vielleicht beliebig große) Werte auf I annimmt.

2) Auf der anderen Seite ist klar: *Wenn die vollständige Kandidatenliste leer ist*, wenn es also keine Kandidaten gibt, *so gibt es auch keine Extremstellen*, nicht einmal lokale. Das gilt z.B. für Funktionen mit überall positiver Ableitung auf $I = \mathbb{R}$ wie x, $x + x^3$, e^x, $\tanh x$, $\ldots$. Auf kompakten Teilintervallen $[a, b]$ von $\mathbb{R}$ haben solche differenzierbaren Funktionen dagegen sehr wohl ein absolutes Maximum und Minimum, aber das kann nur in den Randpunkten angenommen werden, weil es ja keine inneren Kandidaten gibt. *Enthält die vollständige Kandidatenliste nur einen Punkt*, so ist ferner klar, dass es höchstens eine Extremstelle gibt, d.h. also dass Maximum oder Minimum oder beide (!) von f auf I nicht angenommen werden. (Die eine Stelle kann nicht Minimum- und Maximumstelle zugleich sein, weil sonst f konstant wäre und daher $f' = 0$ überall auf I; wir nehmen hier natürlich positive Länge von I an.)

3) Die hinreichenden Kriterien für lokale Extremstellen kann man bei einer differenzierbaren Funktion $f : I \to \mathbb{R}$ anwenden zur

> **Bestimmung aller inneren lokalen Extremstellen** durch Untersuchung des Vorzeichens von f' bei den inneren kritischen Punkten und zur
>
> **Verkleinerung der Kandidatenliste für absolute Extremstellen**.

Die Nullstellen von nichtkonstanten elementaren Funktionen f, und das gilt auch für ihre Ableitungen f', sind immer isoliert und von einem der vier möglichen Typen $-/-$, $-/+$, $+/-$, $+/+$. Mit der Berechnung der Nullstellen von f' kann man in den meisten Fällen ohne besonderen Mehraufwand auch gleich den Typ dieser Nullstellen mitbestimmen und erhält damit die inneren lokalen Extremstellen (Typen $-/+$ oder $+/-$) und Sattelpunkte (Typen $-/-$ und $+/+$). Auch bei zu I gehörenden Randpunkten lassen sich hinreichende Ableitungskriterien für lokale isolierte Extremstellen angeben. Ist z.B. $x_0 \in I$ linker Randpunkt und $f'(x_0) > 0$ oder f stetig in x_0 mit $f'(x) > 0$ für $x > x_0$ nahe bei x_0, so liegt in x_0 eine lokale isolierte Randminimumstelle vor, d.h. es ist $f(x) > f(x_0)$ für $x > x_0$ nahe bei x_0.

Die Sattelpunkte kann man aus der Kandidatenliste für die absoluten Extremstellen von f auf I natürlich streichen, ebenso die lokalen strikten Maximumstellen (Typ $+/-$) aus der Kandidatenliste für absolute Minimumstellen und die lokalen strikten Minimumstellen (Typ $-/+$) aus der Kandidatenliste für die absoluten Maximumstellen von f auf I. Wenn es aber nur um die Bestimmung der absoluten Extremstellen geht, so lohnt sich der Aufwand meistens nicht, den man hat, um die Kandidatenliste auf diese Weise etwas verkleinern zu können. Es ist dann einfacher, diese mögliche Verkleinerung der Liste zu unterlassen und beim Wertevergleich einige wenige Funktionswerte mehr zu berechnen.

4) Gibt es nur endlich viele innere kritische Punkte $x_1 < x_2 < \ldots < x_n$ der untersuchten Funktion $f : I \to \mathbb{R}$ — und das ist in der Praxis immer der Fall —, so ist die

- *Bestimmung der absoluten Extremstellen durch* ***Vergleich der Randwerte mit dem kleinsten und größten inneren kritischen Wert***

möglich. Wir nehmen dazu an, dass f im Inneren von I differenzierbar ist und dass $a < b$ die evtl. unendlichen Randpunkte von I sind, sowie $f(a+) = \lim_{x \downarrow a} f(x)$ sowie $f(b-) = \lim_{x \uparrow b} f(x)$ die Grenzwerte der Funktion am Rand, genannt *Randwerte* zu f auf I. Diese Grenzwerte existieren, wenn man dafür auch die unendlichen Werte $\pm\infty$ zuläßt, weil f auf dem Intervall $]a, x_1]$ links vom kleinsten kritischen Punkt und auf dem Intervall

$[x_n, b[$ rechts vom größten kritischen Punkt streng monoton ist (gemäß der Aussage (iv) des Zusatzes zu den Ableitungskriterien für Extremstellen). Falls a oder b zu I gehören, so nehmen wir ferner an, dass f dort noch stetig ist, d.h. es ist dann einfach $f(a+) = f(a)$ bzw. $f(b-) = f(b)$ der Funktionswert im Randpunkt (und damit insbesondere endlich). Wenn nun die Randwerte z.B. über dem kleinsten der kritischen Werte $f(x_i)$ liegen, so liegen wegen der strengen Monotonie von f auf den Intervallen zwischen den kritischen Stellen und auf $]a, x_1]$ sowie $[x_n, b[$ überhaupt alle Funktionswerte über dem minimalen kritischen Wert. Also ist in diesem Fall der minimale kritische Wert das Minimum von f auf I und wird genau an den kritischen Stellen mit diesem Wert angenommen. Ist einer der Randwerte gleich dem kleinsten inneren kritischen Wert, der andere nicht kleiner, so gilt dasselbe mit der Modifikation, dass das Minimum jetzt auch noch zusätzlich in den zu I gehörenden Randpunkten erreicht wird, in denen der Funktionswert gleich dem minimalem inneren kritischen Wert ist. Liegt aber ein Randwert unterhalb, so kann das Infimum von f auf I höchstens in Randpunkten angenommen werden, die zu I gehören. Es gilt also unter diesen Annahmen:

- *Sind die Randwerte nicht kleiner als der kleinste kritische Wert von f im Innern, so ist der kleinste innere kritische Wert das absolute Minimum von f auf I; die absoluten Minimumstellen von f auf I sind in diesem Fall genau die inneren kritischen Punkte mit kleinstem Funktionswert und die in I liegenden Randpunkte, an denen dieser Wert evtl. auch noch angenommen wird.*
- *Liegt ein Randwert aber unter dem kleinsten inneren kritischen Wert, so wird das Minimum von f nur in den zu I gehörenden Randpunkten angenommen, wo der kleinste Randwert erreicht wird (also gar nicht, wenn die Randpunkte nicht zu I gehören oder wenn der kleinste Randwert $-\infty$ ist).*

Ganz entsprechende Aussagen gelten natürlich für den Vergleich der Randwerte mit dem größten inneren kritischen Wert zur Bestimmung der Maximumstellen von f auf I; wir brauchen diese Aussagen nicht mehr explizit zu formulieren. Auch kann man endlich viele Nichtdifferenzierbarkeitsstellen im Inneren zulassen, an denen f noch stetig ist. Man behandelt diese Stellen dann einfach so, als wären es kritische Punkte, und die obigen Aussagen gelten dann genauso.

5) Wir halten noch gesondert fest, was die Überlegungen in 4) ergeben, wenn f nur einen oder gar keinen inneren kritischen Punkt hat. Wie in 4) nehmen wir dabei an, dass f im Innern von I differenzierbar ist und in den Randpunkten von I, soweit sie in I liegen, noch stetig.

- *Hat f nur einen inneren kritischen Punkt x_0 und ist dieser eine Nullstelle vom Typ $-/+$ von f' [bzw. vom Typ $+/-$], so ist f beiderseits von x_0 gegensinnig streng monoton und x_0 ist die eindeutige absolute Minimumstelle [Maximumstelle] von f auf I. Ein Maximum [Minimum] von f auf I wird dann nur in den in I liegenden Randpunkten erreicht, in denen f den gößten [kleinsten] Randwert annimmt (also gar nicht, wenn dieser Randwert nicht angenommen wird).*
- *Hat f keine inneren kritischen Punkte oder nur Sattelpunkte, so ist f streng monoton auf I und nimmt ein Minimum bzw. Maximum genau in den Randpunkten mit dem kleineren bzw. größeren Randwert an, sofern diese in I liegen (sonst gar nicht).*

■

BEISPIELE (***zur Extremstellenbestimmung***):

(1) Lineare Funktionen $\ell(x) = ax + b$ mit Steigung $a \neq 0$ treten in mathematischen Modellen der Ökonomie häufig auf, weil es die einfachsten nichtkonstanten Funktionen sind. Sie haben keine kritischen Punlkte, da f' konstant gleich $a \neq 0$ ist. Folglich gilt:

- *Nichtkonstante lineare Funktionen haben keine kritischen Punkte; sie nehmen also ein Maximum oder Minimum höchstens in Randpunkten an.*

Genauer ist es so, wenn z.B. die Steigung a positiv ist, also die lineare Funktion ℓ streng wachsend ist: Ein Minimum wird genau dann angenommen, wenn das Definitionsintervall I von unten beschränkt ist und wenn sein linker Endpunkt $x_{\min}$ zu I gehört; dann ist $x_{\min}$ die einzige Minimumstelle in I und $\ell(x_{\min}) = ax_{\min} + b$ das Minimum. Analog wird ein Maximum von ℓ auf I genau dann angenommen, wenn I von oben beschränkt ist und wenn der rechte Endpunkt $x_{\max}$ in I liegt; dann ist $x_{\max}$ die einzige Maximumstelle und $\ell(x_{\max}) = ax_{\max} + b$ das Maximum von ℓ auf I. Maximum und Minimum werden also dann und nur dann von ℓ angenommen, wenn $I = [x_{\min}, x_{\max}]$ ein kompaktes Intervall ist. Offenbar gilt die Argumentation nicht nur für lineare (nichtkonstante) Funktionen, sondern überhaupt für jede streng monotone Funktion:

- *Streng monotone Funktionen nehmen Maximum und Minimum höchstens am Rande des Definitionsintervalls an.*

In der Ökonomie ist das sinnvolle Definitionsintervall oft durch Kapazitätsgrenzen bestimmt; z.B. ist es für eine Kostenfunktion $K(x)$ das Intervall $[0, x_{\max}]$ zwischen dem kleinstmöglichen Produktionsoutput $x = 0$ und dem durch die Produktionskapazität bestimmten größtmöglichen Output $x_{\max}$. Die Tatsache, dass eine lineare Funktion auf einem kompakten Intervall die Extremwerte in den Randpunkten annimmt (und nur dort, wenn sie Steigung $\neq 0$ hat), wird dann so formuliert:

- *Lineare ökonomische Funktionen nehmen ihre Extremwerte nur bei Erreichen der Kapazitätsgrenzen an.*

(2) Für stetige **stückweise lineare Funktionen** ist die Situation auch noch übersichtlich. Solche Funktionen sind auf den Teilintervallen einer Zerlegung des Grundintervalls durch lineare Funktionen definiert, wobei die linearen Funktionen für zwei benachbarte Intervalle jeweils zusammenpassen, d.h. in ihrem gemeinsamen Randpunkt denselben Wert haben. Dabei brauchen die Steigungen der beiden linearen Funktionen nicht übereinzustimmen, man spricht dann von einer *Knickstelle* der stückweise linearen Funktion. Einfachstes Beispiel ist die auf $\mathbb{R}_{\geq 0}$ und auf $\mathbb{R}_{\leq 0}$ lineare Betragsfunktion mit den Steigungen -1 links von Null und $+1$ rechts von Null; bei $x = 0$ liegt das Paradebeispiel einer Knickstelle vor. geometrisch kann man stückweise lineare Funktionen durch die Eigenschaft beschreiben, dass ihr Graph ein Streckenzug ist (wobei ganz links oder ganz rechts auch Halbgeraden statt Strecken endlicher Länge auftreten können).

Kandidaten für Extremstellen einer stückweise linearen Funktion f auf einem Intervall I sind dann nur die Randpunkte von I (soweit sie in I liegen), die inneren Knickstellen und die Punkte in Intervallen, auf denen f konstant ist, also Steigung Null hat; denn in allen anderen Punkten ist f differenzierbar, sogar lokal linear, mit Ableitung $\neq 0$. Bei den Knickstellen kann man noch diejenigen aussortieren, bei denen f streng monoton ist, d.h. positive Steigung auf beiden Seiten der Stelle hat oder negative Steigung auf beiden Seiten. Einfache Beispiele zeigen, das die Extremstellen dann in den verbliebenen Kandidaten jeder Art — Randpunkte, Knickstellen mit Steigungen unterschiedlichen Vorzeichens und Stellen in Konstanzintervallen — liegen können.

(3) Quadratische Funktionen $q(x) = ax^2 + bc + c$ (mit $a \neq 0$) haben $q'(x) = 2ax + b$ als Ableitung und als einzige Nullstelle von q' auf $\mathbb{R}$ den Scheitelrechtswert $x_0 = -\frac{b}{2a}$. Mit quadratischer Ergänzung $f(x) = a(x + \frac{b}{2a})^2 + c - \frac{b^2}{4a}$ sieht man sofort, dass im Fall $a > 0$ eine absolute Minimumstelle von q auf $\mathbb{R}$ liegt, im Fall $a < 0$ eine absolute Maximumstelle, wobei q auf beiden Seiten dieser Stelle streng monoton ist.

- *Quadratische Funktionen q haben einen einzigen kritischen Punkt x_0 in $\mathbb{R}$ und dieser ist absolute Minimum- oder Maximumstelle von q auf $\mathbb{R}$; auf den Halbgeraden links und rechts von x_0 ist q gegensinnig streng monoton.*
- *Für Extremstellen von q auf einem Intervall $I \subset \mathbb{R}$ gibt es maximal drei Kandidaten: Die Randpunkte von I, die zu I gehören, und die Stelle x_0, wenn sie in I liegt.*

Natürlich ist x_0, wenn x_0 in I liegt und $a > 0$ ist, auch eindeutige Minimumstelle von q auf I. Ein Maximum nimmt q auf I dann nur an, wenn beide Randpunkte endlich sind und wenn der mit dem größeren Funktionswert, bzw. im Fall gleicher Randwerte wenigstens einer der beiden Randpunkte, zu I gehört.

Angenommen, wir würden die quadratische Ergänzung nicht kennen, dann könnten wir im Fall $a > 0$ so argumentieren: Wegen $q(x) \to \infty$ bei $x \to \infty$ und bei $x \to -\infty$ (und wegen Stetigkeit von q) nimmt q ein absolutes Minimum auf $\mathbb{R}$ an; einziger Kandidat für eine Extremstelle ist aber x_0; also muss in x_0 eine eindeutige absolute Minimumstelle vorliegen, und auf $]-\infty, x_0]$ ist q streng fallend, auf $[x_0, \infty[$ streng wachsend, weil ja stetige Funktionen auf Intervallen ohne innere kritische Punkte streng monoton sind (und hier z.B. wegen der Grenzwerte im Unendlichen nur die angegebene Richtung der Monotonie in Frage kommt). Der Vorteil dieser Argumentation ist, dass sie gar nicht auf quadratische Funktionen beschränkt ist, sondern sehr viel allgemeiner das Folgende zeigt:

- *Ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar mit genau einem kritischen Punkt x_0 in $\mathbb{R}$ und mit Grenzwerten $> f(x_0)$ bei $x \to \pm\infty$, so ist x_0 eindeutige absolute Minimumstelle von f auf $\mathbb{R}$, und f ist auf $]-\infty, x_0]$ streng fallend, auf $[x_0, \infty[$ streng wachsend.*

Für die Extremstellen dieser Funktion f auf einem Intervall $I \subset \mathbb{R}$ gibt es dann ebenfalls maximal drei Kandidaten: Die Randpunkte von I, soweit sie zu I gehören, und die Stelle x_0, die, wenn sie in I liegt, natürlich auch eindeutige absolute Minimumstelle auf I ist.

(4) In der Ökonomie treten quadratische Funktionen insbesondere als **quadratische Gewinnfunktionen** $G(x) = E(x) - K(x) = x \cdot p(x) - K_{\text{var}}(x) - K_{\text{fix}}$ auf, nämlich wenn die Nachfragefunktione $x(p)$ und damit auch die Preisfunktion $p(x)$ (Umkehrfunktion) linear ist, $p(x) = b - ax$, und wenn die variablen Kosten $K_{\text{var}} = cx$ homogen linear sind, so dass also

$$G(x) = x(b - ax) - cx - K_{\text{fix}} = -ax^2 + (b - c)x - K_{\text{fix}}$$

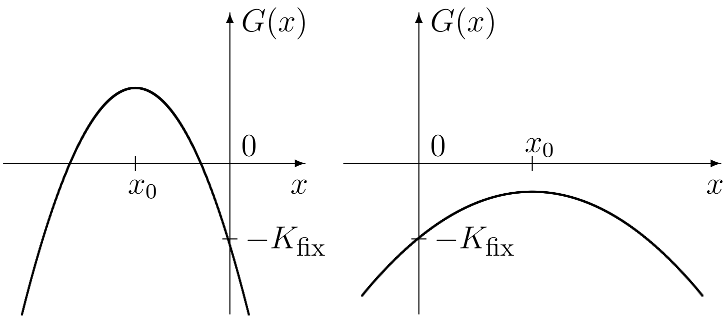
ist. (Auch einen quadratischen Anteil der variablen Kosten könnte man zulassen, dann wäre $G(x)$ immer noch quadratisch, wenn der führende Koeffizient nicht Null wird.) Hier ist natürlich $a > 0$; denn bei zunehmendem und auf den Markt gebrachten Output x wird ein kleinerer Preis pro abgesetzte Einheit erzielt. Also hat $G(x)$ ein eindeutiges Maximum auf $\mathbb{R}$ bei $x_0 = \frac{b-c}{2a}$. Hier geht es aber nicht um das Maximum auf $\mathbb{R}$, sondern um das auf $\mathbb{R}_{\geq 0}$ oder auf dem Intervall $[0, x_{\max}]$ von 0 bis zum maximal möglichen Output $x_{\max}$. Außerdem sind nur Werte $x < \frac{b}{a}$ sinnvoll, da sonst der Preis $p(x)$ nicht positiv wäre. Für $x \geq \frac{b}{a}$ ist aber $G(x) < 0$, so dass solche Werte des Outputs x ohnehin außer Betracht bleiben, wenn man Produktion mit positivem Gewinn analysiert.

Wir müssen nun drei Fälle unterscheiden: Im Fall $b \leq c$ ist $x_0 \leq 0$ und die Maximumstelle ist $x = 0$, weil $G(x)$ ja rechts von x_0 fällt; wegen $G(0) = -K_{\text{fix}} \leq 0$ ist dann keine Produktion mit Gewinn möglich. Im Fall $b > c$ ist $x_0 > 0$ eindeutige Maximumstelle von $G(x)$ auf $\mathbb{R}$ und auf $\mathbb{R}_{\geq 0}$. Der maximale Gewinn auf $\mathbb{R}_{\geq 0}$ ist dann $G(x_0) = \frac{(b-c)^2}{4a} - K_{\text{fix}}$, und nur wenn diese Zahl positiv ist, ist Produktion mit Gewinn möglich. Die *Gewinnzone*, also das Positivitätsintervall von $G(x)$ liegt dann zwischen den Stellen $\frac{1}{2a}(b-c\pm\sqrt{(b-c)^2 - 4aK_{\text{fix}}})$. Wenn x_0 in $[0, x_{\max}]$ liegt, so wird der größte Gewinn auch durch den Output x_0 innerhalb der Kapazitätsgrenzen erzielt. Wenn allerdings $x_0 > x_{\max}$ ist, d.h. $b - c > 2ax_{\max}$, so ist die Kapazitätsgrenze $x_{\max}$ die eindeutige Maximumstelle im Intervall $[0, x_{\max}]$. Produktion mit Gewinn ist in diesem Fall nur möglich, wenn diese Stelle noch in der Gewinnzone liegt, wenn also $\frac{1}{2a}(b-c-\sqrt{(b-c)^2 - 4aK_{\text{fix}}}) < x_{\max}$ ist. Wir fassen das Ergebnis der Disussion zusammen:

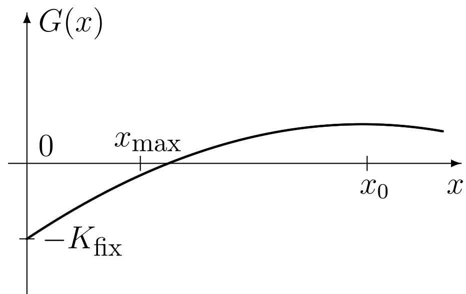
- *Ohne Kapazitätsgrenze* ist Produktion mit Gewinn möglich, genau wenn
$$b > c \quad \text{und} \quad (b-c)^2 - 4aK_{\text{fix}} > 0\,;$$
der maximale Gewinn wird dann erreicht genau bei $x_0 = (b-c)/2a$.

- *Mit Kapazitätsgrenze* $x_{\max}$ ist Produktion mit Gewinn möglich, genau wenn
$$b > c \qquad (b-c)^2 - 4aK_{\text{fix}} > 0 \quad \text{und} \quad \sqrt{(b-c)^2 - 4aK_{\text{fix}}} > \tfrac{1}{2a}(x_0 - x_{\max})\,;$$
der maximale Gewinn wird dann erreicht genau bei $x^* = \min\{x_0, x_{\max}\}$.

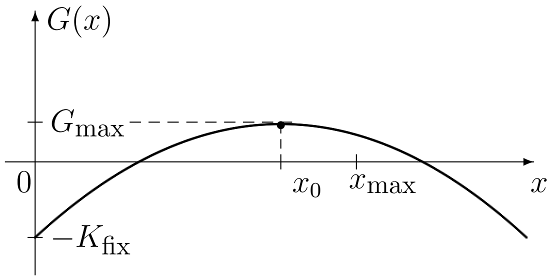
Das leuchtet ein: Wenn der maximale Gewinn ohne Kapazitätsbegrenzung bei einem Output erreicht würde, der oberhalb der realen Kapazitätsgrenze liegt, so bringt Produktion mit maximal Output den größten möglichen Gewinn. Die folgenden Abbildungen illustrieren die verschiedenen möglichen Situationen:



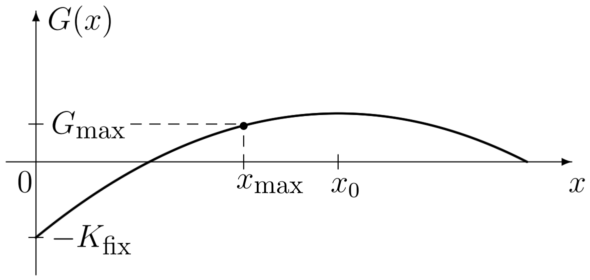
überhaupt kein Gewinn möglich



kein Gewinn innerhalb der Kapazitätsgrenze möglich



maximaler Gewinn unterhalb der Kapazitätsgrenze

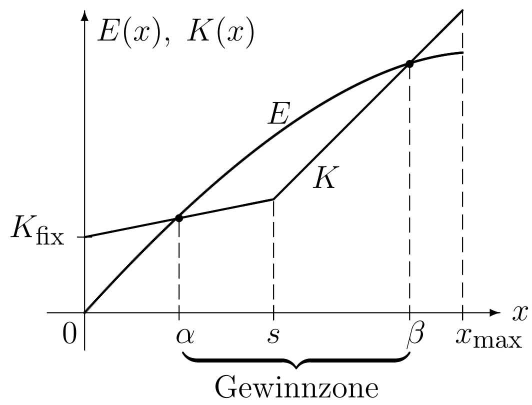


maximaler Gewinn genau an der Kapazitätsgrenze

(5) Stückweise quadratische Gewinnfunktionen entstehen, wenn die Erlösfunktion wie in 4) ist, aber die Kostenfunktion stückweise linear mit unterschiedlichen Steigungen. Es ist ökonomisch sinnvoll, die Erlösfunktion $E(x) = x \cdot p(x)$ als differenzierbar anzunehmen, d.h. man rechnet nicht mit sprunghaften Änderungen des Preises oder des Grenzpreises. Dagegen sind Nichtdifferenzierbarkeitsstellen bei Kostenfunktionen durchaus realistisch. Zum Beispiel ist denkbar, dass bei Überschreitung bestimmter Schwellenwerte durch den Output x die übersteigende Menge nur mit höheren Stück–variablen Kosten produziert werden kann, etwa wegen Lohnzuschlägen für Nacht- oder Wochenendarbeit oder wegen des Einsatzes weniger rentabler Produktionsanlagen. Dann ist die Kostenfunktion zwar noch stetig am Schwellenwert, aber nicht mehr ihre Ableitung. Entstehen bei der Inbetriebnahme weiterer Produktionsaggregate zusätzliche Fixkosten, so hat man sogar einen Sprungstelle der Kostenfunktion wegen des Fixkostensprungs.

Als einfaches Modell betrachten wir eine lineare Nachfragefunktion mit zugehöriger quadratischer Erlösfunktion $E(x) = x \cdot p(x) = -ax^2 + bx$ wie in 4) auf dem Intervall von $x = 0$ bis $x_{\max}$ (Kapazitätsgrenze) und eine stückweise lineare Kostenfunktion $K(x) = c_1 x + K_{\text{fix}}$ für $0 \le x \le s$ mit einem Schwellenwert $0 < s < x_{\max}$ und $K(x) = c_2(x - s) + K(s)$ für $s < x \le x_{\max}$, wobei $0 < c_1 < c_2$ sei. Man kann zunächst durch Auftragen der Graphen von $E(x)$ und von $K(x)$ (oder durch Auflösen der Ungleichung $E(x) > K(x)$) überprüfen, ob es überhaupt Output-Werte $x \in [0, x_{\max}]$ gibt, für die mit Gewinn produziert werden kann, für die also $G(x) = E(x) - K(x) > 0$ ist. Die Menge dieser x, also die *Gewinnzone*, ist dann ein Teilintervall $]\alpha, \beta[$ von $[0, x_{\max}]$ mit $0 < \alpha < \beta \le x_{\max}$ und $G(\alpha) = G(\beta) = 0$ oder ein Teilintervall $]\alpha, x_{\max}]$ mit $G(\alpha) = 0 < G(x_{\max})$, wie man sich überlegen kann (weil E streng konkav und K konvex, also $G = E - K$ streng konkav, ist und $G(0) = -K_{\text{fix}} < 0$). Die folgende Abbildung zeigt die Situation, wenn $\beta < x_{\max}$ ist. Der Bereich zwischen dem Graphen der Kostenfunktion und der Erlösfunktion heißt "*Gewinnlinse*"; die Projektion der Gewinnlinse auf die x-Achse ist die Gewinnzone.

Ist die Gewinnzone nicht leer, so nimmt die Gewinnfunktion darauf ein positives Maximum an; denn ihr Maximum auf dem kompakten Intervall $[\alpha, \beta]$ kann nicht in α oder β angenommen werden und ihr Maximum auf $[\alpha, x_{\max}]$ nicht in α, weil $G(\alpha) = 0$ und $G(\beta) = 0$ ist und G positive Werte im Innern des Intervalls hat. Die Aufgabe ist nun, die Stelle zu berechnen, an der das Maximum angenommen wird, also den Produktions-Output x^* mit maximalem Gewinn.

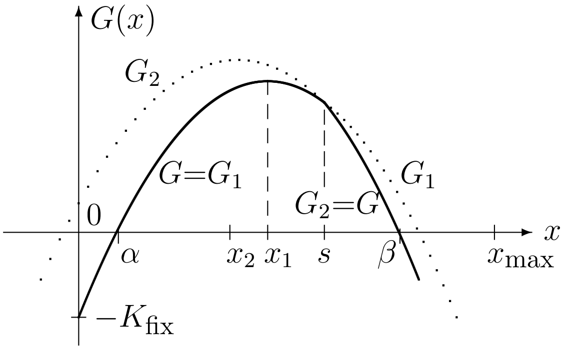


In unserer konkreten Situation ist die Gewinnfunktion aus zwei quadratischen Funktionen zusammengesetzt:

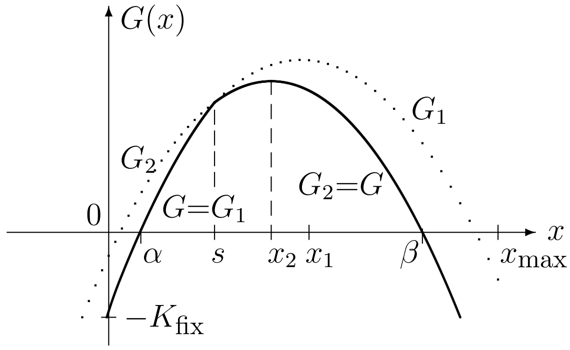
$$G(x) = \begin{cases} G_1(x) = -ax^2 + (b - c_1)x - K_{\text{fix}} & \text{für } 0 \le x \le s\,, \\ G_2(x) = -ax^2 + (b - c_2)x + c_2 s - K(s) & \text{für } s \le x \le x_{\max}\,, \end{cases}$$

wobei $G_1'(s) = G_2'(s) + (c_2 - c_1)$ ist, also G in s eine Knickstelle hat, in der die Steigung sprunghaft um $c_2 - c_1$ abnimmt.

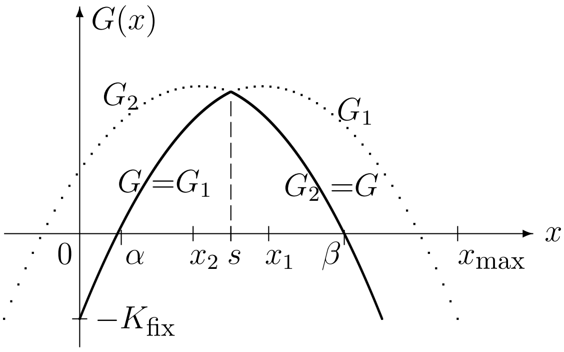
Die Kandidatenliste für die absoluten Maximumstellen von G in der Gewinnzone besteht aus der Nichtdifferenzierbarkeitsstelle s, dem Randpunkt $x_{\max}$, wenn er in der Gewinnzone liegt, und den kritischen Punkten $x_1 = (b - c_1)/2a$ von G_1, sofern x_1 in $]\alpha, s[$ liegt, sowie $x_2 = (b - c_2)/2a$, sofern x_2 in $]s, x_{\max}[$ liegt. Je nach Situation gibt es also einen bis vier Kandidaten in der Gewinnzone, und in genau einem davon wird das Maximum angenommen, weil wir ja schon überlegt haben, dass es ein Maximum in der Gewinnzone gibt, und weil wir auch schon bemerkt haben, dass G streng konkav ist. (Konkave Funktionen mit zwei Maximumstellen müssen dazwischen konstant gleich dem Maximum sein, können also nicht streng konkav sein.) Beispiele zeigen, das das Gewinnmaximum in jedem der vier Punkte $x_1, x_2, s, x_{\max}$ liegen kann.



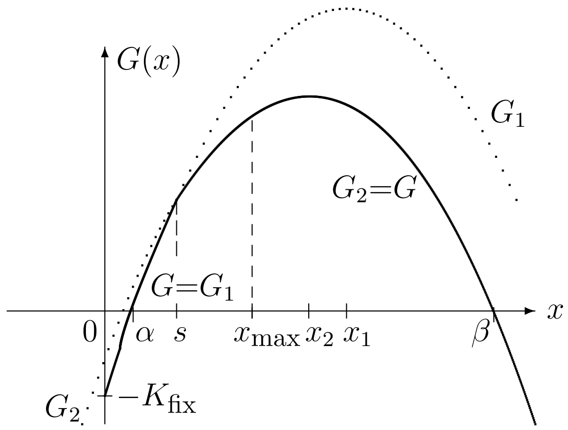
Gewinnmaximum in x_1



Gewinnmaximum in x_2



Gewinnmaximum in s

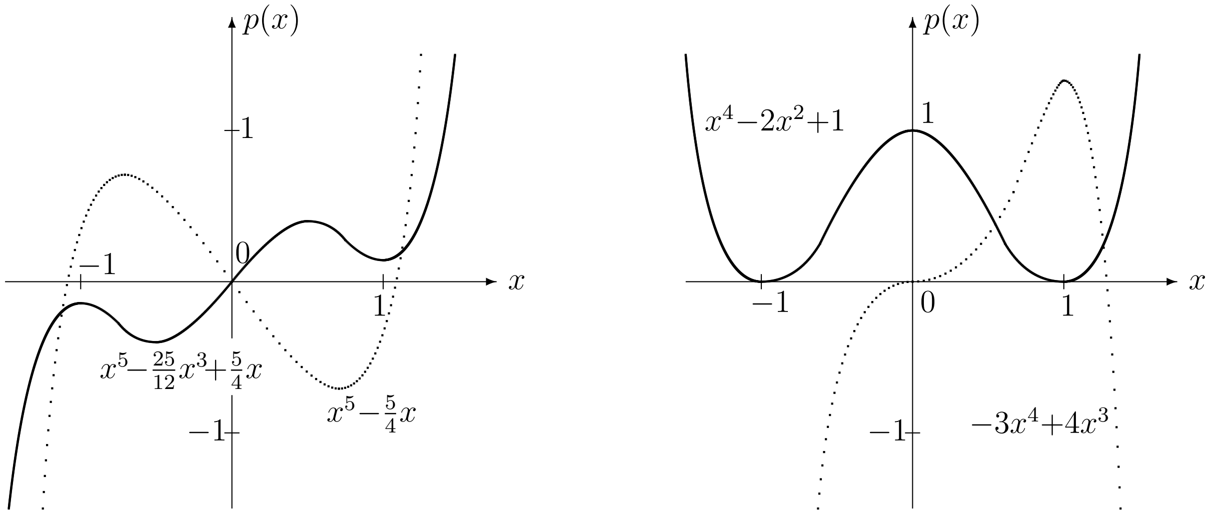


Gewinnmaximum in x_{max}

(6) Polynomfunktionen $p(x) = a_n x^n + \ldots + a_1 x + a_0$ vom Grad n haben eine Polynomfunktion $p'(x)$ vom Grad $n-1$ als Ableitung, und diese hat höchstens $n-1$ Nullstellen (mit Vielfachheit gerechnet). Also gibt es *höchstens* $n-1$ *kritische Punkte* in $\mathbb{R}$. Da die Nullstellen von Polynomfunktionen entweder ungerade Vielfachheit haben und dann Nullstellen mit Vorzeichenwechsel sind, oder gerade Vielfachheit besitzen und dann Nullstellen des Typs $+/+$ oder $-/-$ sind, *sind die kritischen Punkte von* $p(x)$ *entweder lokale isolierte Extremstellen oder strikte Sattelpunkte.* Da zwischen zwei lokalen isolierten Minimumstelle eine lokale Maximumstelle liegen muss und zwischen zwei lokalen isolierten Maximumstellen eine lokale Minimumstelle, *folgen die lokalen Minimumstellen und Maximumstellen im Wechsel aufeinander*, wobei dazwischen nur Sattelpunkte liegen können.

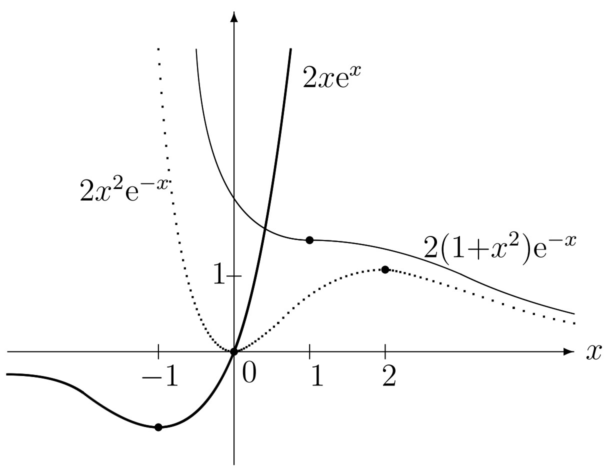
Bei ungeradem Grad $n = 2m+1 \geq 3$ von $p(x)$ braucht es überhaupt keinen kritischen Punkt zu geben, wie die streng monotonen Funktionen $x^{2m+1}+x$ mit Ableitung ≥ 1 überall auf $\mathbb{R}$ zeigen. Es ist auch möglich, dass es nur Sattelpunkte gibt, aber keine lokalen Extremstellen, wie man an x^{2m+1} sieht mit einem einzigen kritischen Punkt $x = 0$, der Sattelpunkt ist. Wenn es aber lokale Extremstellen gibt, so mindestens zwei, bzw. eine gerade positive Anzahl, genauer gesagt. Das gilt, weil $p'(x)$ geraden Grad hat und daher bei Zählung mit Vielfachheit eine gerade Anzahl reeller Nullstellen (nach Ausfaktorisieren aller Linearfaktoren zu reellen Nullstellen von $p'(x)$ bleibt ein Restpolynom ohne reelle Nullstelle, also ein Restpolynom von geradem Grad) und weil die Sattelpunkte Nullstellen gerader Vielfachheit von $p'(x)$ sind, die lokalen Extremstellen aber Nullstellen von ungerader Vielfachheit. Da die lokalen Minimumstellen und Maximumstellen im Wechsel aufeinander folgen, *gibt es also gleich viele lokale Minimumstellen und Maximumstellen.* Absolute Extremstellen von $p(x)$ auf $\mathbb{R}$ existieren natürlich nicht; denn $p(x)$ nimmt ja beliebig große Werte und auch negative Werte von beliebig großem Betrag an.

Bei geradem Grad $n = 2m \geq 2$ von $p(x)$ hat $p'(x)$ mindestens eine Nullstelle in $\mathbb{R}$, also gibt es mindestens einen kritischen Punkt, und evtl. nicht mehr, wie x^{2m} mit einem einzigen kritischen Punkt in $x = 0$ zeigt, der absolute Minimumstelle ist. Allgemein kann man Folgendes sagen: Weil $p(x)$ bei geradem Grad denselben unendlichen Grenzwert ∞ bzw. $-\infty$ bei $x \to \pm\infty$ hat, nimmt $p(x)$ sogar ein absolutes Minimum bzw. Maximum auf $\mathbb{R}$ an (aber natürlich nur eins von beiden). Im Fall $a_n > 0$ hat $p(x)$ den Limes ∞ bei $x \to \pm\infty$, also besitzt die Polynomfunktion dann *mindestens eine absolute Minimumstelle und insgesamt eine lokale Minimumstelle mehr als lokale Maximumstellen*, weil diese Stellen ja im Wechsel aufeinander folgen und weil zwischen einer lokalen isolierten Maximumstelle und $\pm\infty$ noch eine lokale Minimumstelle kommen muss. Im Fall $a_n < 0$ hat $p(x)$ natürlich den Limes $-\infty$ bei $x \to \pm\infty$, besitzt also *mindestens eine absolute Maximumstelle und insgesamt eine lokale Maximumstelle mehr als lokale Minimumstellen.*

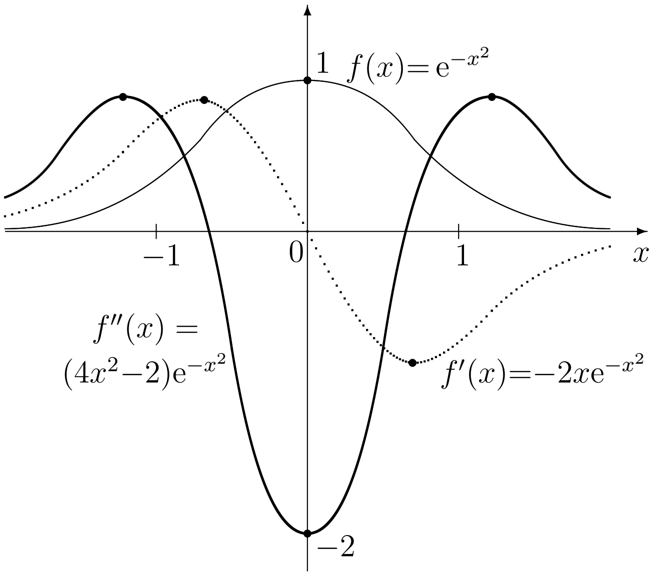


(7) Bei Funktionen der Form $f(x) = p(x)\mathrm{e}^{q(x)}$ mit Polynomfunktionen p, q ist die Bestimmung der kritischen Punkte und ihres Typs ebenfalls mit Berechnung der Nullstellen und der Nullstellenvielfachheiten einer Polynomfunktion zu erledigen. Gemäß Produkt- und Kettenregel ist nämlich $f'(x) = [p'(x) + p(x)q'(x)]\mathrm{e}^{q(x)}$, und weil $\mathrm{e}^{q(x)} > 0$ ist für alle $x \in \mathbb{R}$, sind die Nullstellen von $f'(x)$ die des Polynoms $p'(x) + p(x)q'(x)$ und der Nullstellentyp $-/-$, $-/+$, $+/-$ bzw. $+/+$ ist bei $f'(x)$ ebenfalls derselbe wie bei dieser Polynomfunktion. Das Grenzverhalten von $f(x)$ im Unendlichen ist bestimmt durch die führenden Terme $a_m x^m$ von $p(x)$ und $b_n x^n$ von $q(x)$; der Grenzwert ist $\pm\infty$ mit dem Vorzeichen von a_n bei $x \to \infty$ und mit dem Vorzeichen von $(-1)^m a_m$ bei $x \to -\infty$, wenn bei dem Grenzübergang $q(x) \to \infty$ strebt, bzw. gleich 0, wenn dabei $q(x) \to -\infty$ geht (der exponentielle Faktor $\mathrm{e}^{q(x)}$ strebt dann viel schneller gegen Null, als $|p(x)|$ gegen Unendlich gehen kann). Wenn $f(x)$ bei $x \to \infty$ und bei $x \to -\infty$ denselben Grenzwert 0, ∞ oder $-\infty$ hat, so besitzt f mindestens eine absolute Extremstelle auf $\mathbb{R}$, bei Grenzwert 0 sogar eine absolute Minimumstelle und eine absolute Maximumstelle, falls f auf $\mathbb{R}$ positive und negative Werte hat. Im Falle entgegengesetzt unendlicher Grenzwerte von $f(x)$ bei $x \to \infty$ und bei $x \to -\infty$ gibt es natürlich keine absoluten Extremstellen auf $\mathbb{R}$. Ist aber ein Grenzwert von $f(x)$ unendlich, der andere Null, so kommt es darauf an, ob f Nullstellen auf $\mathbb{R}$ hat oder nicht; falls ja, so besitzt f mindestens eine absolute Extremstelle auf $\mathbb{R}$, andernfalls aber keine (weil im letzten Fall der Wertebereich von f auf $\mathbb{R}$ gleich $]0,\infty[$ oder gleich $]-\infty,0[$ ist und keinen größten oder kleinsten Wert enthält, während im ersten Fall entweder schon die Nullstelle absolute Extremstelle ist oder f einen Wert annimmt, der kleiner ist als seine beiden Grenzwerte im Unendlichen bzw. größer als diese beiden Grenzwerte, so dass ein absolutes Minimum auf $\mathbb{R}$ existiert bzw. ein absolutes Maximum).

Konkret hat z.B. $f(x) := x\mathrm{e}^x$ mit Ableitung $f'(x) = (1+x)\mathrm{e}^x$ genau einen kritischen Punkt $x = -1$, und da f' dort einen Vorzeichenwechsesl $-/+$ hat, ist $x = -1$ lokale strikte Minimumstelle von f, sogar eindeutige absolute Minimumstelle von f auf $\mathbb{R}$, weil es keine anderen kritischen Punkte gibt, so dass $x = 1$ absolute Minimumstelle bis zu den unendliche fernen Randpunkten ∞ und $-\infty$ ist. (Da $f(x)$ die Grenzwerte ∞ bei $x \to \infty$ hat und 0 bei $x \to -\infty$ und da f negative Werte annimmt, ist auch aus anderen Gründen klar, dass f eine absolute Minimumstelle auf $\mathbb{R}$ hat; und diese kann nur im einzigen kritischen Punkt $x = 1$ liegen.) Bei $g(x) := x^2\mathrm{e}^{-x}$ mit $g'(x) = (2x - x^2)\mathrm{e}^{-x}$ haben wir zwei kritische Punkte, $x = 0$ mit Vorzeichenwechsel $-/+$ bei g' und $x = 2$ mit Vorzeichenwechsel $+/-$; also ist $x = 0$ lokale strikte Minimumstelle und $x = 2$ eine lokale strikte Maximumstelle. Da $g(x)$ beliebig große Werte annimmt (Grenzwert ∞ bei $x \to -\infty$), kann diese Funktion natürlich keine absolute Maximumstelle auf $\mathbb{R}$ haben. An der Stelle $x = 0$ ist aber $g(x) = 0$, und weil g sonst nur positive Werte hat, liegt dort die eindeutige absolute Minimumstelle auf $\mathbb{R}$. Die Funktion $h(x) := (1+x^2)\mathrm{e}^{-x}$ mit $h'(x) = (2x - 1 - x^2)\mathrm{e}^{-x} = -(1-x)^2\mathrm{e}^{-x}$ schließlich hat wieder nur einen kritischen Punkt $x = 1$, aber der ist ohne Vorzeichenwechsel vom Typ $-/-$, also hat h dort einen strikten Sattelpunkt und überhaupt keine lokalen Extremstellen auf $\mathbb{R}$. Vielmehr ist h streng fallend auf $\mathbb{R}$ mit Wertebereich $]0,\infty[$. Die Abbildung zeigt die vertikal mit dem Faktor 2 gestreckten Graphen.



Beispiele mit Grenzwert 0 bei $x \to \infty$ und bei $x \to -\infty$ erhält man in der Form $p(x)\mathrm{e}^{-x^2}$, also z.B. die Funktion $f(x) := \mathrm{e}^{-x^2}$, die bis auf horizontale und vertikale Skalierung die *Normalverteilung* $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$ ist und als Graph die bekannte Gaußsche “Glockenkurve” hat, sowie die Ableitung $f'(x) = -2x\mathrm{e}^{-x^2}$ sowie die zweite Ableitung $f''(x) = (4x^2 - 2)\mathrm{e}^{-x^2}$.



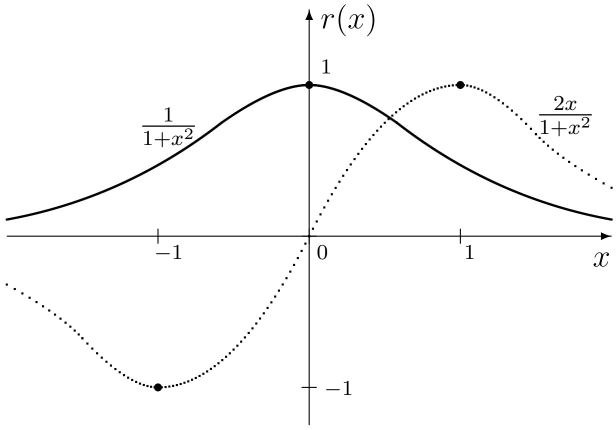
Einziger kritischer Punkt von f ist $x = 0$, und dort hat f' einen Vorzeichenwechsel vom Typ $+/-$, also liegt eine lokale isolierte Maximumstelle vor, die sogar einzige absolute Maximumstelle von f auf $\mathbb{R}$ ist, weil es keine weiteren kritischen Punkte gibt. (Alternativ überlegt man, dass f ein absolutes Maximum auf $\mathbb{R}$ annehmen muss, weil es Grenzwert 0 im Unendlichen hat und positive Werte annimmt; die Maximumstelle kann dann nur der einzige kritische Punkt $x = 0$ sein.) Die kritischen Punkte von f' sind die Nullstellen $\pm 1/\sqrt{2}$ von f'', dort hat f'' Vorzeichenwechsel vom Typ $+/-$ und vom Typ $-/+$, also liegt eine lokale strikte Maximumstelle und eine lokale strikte Minimumstelle von f' vor. (Diese Stellen sind die Wendepunkte von f, d.h. die Punkte, bei denen ein Übergang von konvexem zu konkavem Funktionsverhalten oder umgekehrt stattfindet. Wir kommen auf Wendepunkte in 4.6 zu sprechen.) Da $f'(x)$ Grenzwerte Null bei $x \to \pm\infty$ und positive wie negative Werte hat, nimmt f' ein positives absolutes Maximum auf $\mathbb{R}$ an, das nur in $-1/\sqrt{2}$ liegen kann, und ein negatives absolutes Minimum, das im zweiten kritischen Punkt $1/\sqrt{2}$ liegen muss. Die kritischen Stellen von f'' sind die Nullstellen von $f'''(x) = (12x - 8x^3)\mathrm{e}^{-x^2}$, also die drei Stellen $x = 0$ und $x = \pm\sqrt{3/2}$; bei $x = 0$ liegt ein Vorzeichenwechsel $-/+$ von f''' vor, also eine lokale isolierte Minimumstelle von f'', bei $\pm\sqrt{3/2}$ haben wir einen Vorzeichenwechsel $+/-$ von f''', also eine lokale strikte Maximumstelle von f''. Weil $f''(x)$ Grenzwert 0 bei $x \to \pm\infty$ und positive wie auch negative Werte hat, nimmt f'' ein absolutes Minimum auf $\mathbb{R}$ an, was nur an der Stelle $x = 0$ liegen kann, und auch ein absolutes Maximum, was nur bei $x = \sqrt{3/2}$ oder $x = -\sqrt{3/2}$ passieren kann und tatsächlich an diesen beiden Stellen angenommen wird, weil f'' dort denselben Funktionswert hat (bzw. weil f'' eine gerade Funktion ist.)

(8) Rationale Funktionen $r(x) = p(x)/q(x)$ mit Polynomen $p(x) = a_m x^m + \ldots + a_1 x + a_0$ und $q(x) = b_n x^n + \ldots + b_1 x + b_0$ haben eine rationale Funktion $r'(x) = [p'(x)q(x) - p(x)q'(x)]/q(x)^2$ als Ableitung auf ihrem maximalen Definitionsbereich $D \subset \mathbb{R}$. (Der maximale Definitionsbereich ist $\mathbb{R}$ ohne die Polstellen; wir nehmen an, dass hebbare Definitionslücken behoben sind, also $p(x)$ und $q(x)$ keine gemeinsamen Nullstellen haben.) Die Nullstellen von $r'(x)$ sind daher dieselben wie die des Polynoms $p'(x)q(x) - p(x)q'(x)$, und der Nullstellentyp $-/-$, $-/+$, $+/-$ bzw. $+/+$ ist auch derselbe. Da $p'(x)q(x) - p(x)q'(x)$ Grad $m+n-1$ hat im Fall $m \neq n$ und Grad $\leq 2n-2$ im Fall $m = n$, *gibt es höchstens* $m+n-1$ *kritische Punkte, im Fall* $m = n$ *sogar höchstens* $2n-2$, *und alle kritischen Punkte sind lokale strikte Extremstellen oder strikte Sattelpunkte.* Außerdem folgen die Vorzeichenwechsel $-/+$ und $+/-$ im Wechsel aufeinander, wobei höchstens Nullstellen ohne Vorzeichenwechsel dazwischen liegen. Dabei muss man allerdings berücksichtigen, dass Polstellen der Ordnung $\ell \geq 2$ auch Nullstellen der Vielfachheit $\ell - 1$ von $p'(x)q(x) - p(x)q'(x)$ sind. Wenn wir Polstellen gerader Ordnung als “uneigentliche lokale Extremstellen” von $r(x)$ ansehen und genauer solche mit beidseitigem Grenzwert $+\infty$ von $r(x)$ als “uneigentliche lokale Maximumstellen” und solche mit beidseitigem Limes $-\infty$

als “uneigentliche lokale Minimumstellen”, so *folgen die lokalen (evtl. uneigentlichen) Minimumstellen und Maximumstellen im Wechsel aufeinander, wobei nur Sattelpunkte oder Polstellen ungerader Ordnung dazwischen liegen können.* (Polstellen ungerader Ordnungkann man als “uneigentliche Sattelpunkte” ansehen, weil $r(x)$ auf den beiden Seiten nahe einer solchen Stelle gegensinnig streng monoton ist.) Bei Mitzählung der uneigentlichen Extremstellen ist daher die Zahl der lokalen Minimumstellen gleich der Zahl der lokalen Maximumstellen, oder beide Zahlen unterscheiden sich nur um 1; der erste Fall liegt vor, wenn $m+n$ ungerade ist oder wenn $m = n$ ist und $r(x) - a_n/b_n = \widetilde{p}(x)/q(x)$ mit $\widetilde{p}$ vom Grad $\widetilde{m} < n$ und $\widetilde{m} + n$ ungerade, ansonsten hat man den zweiten Fall.

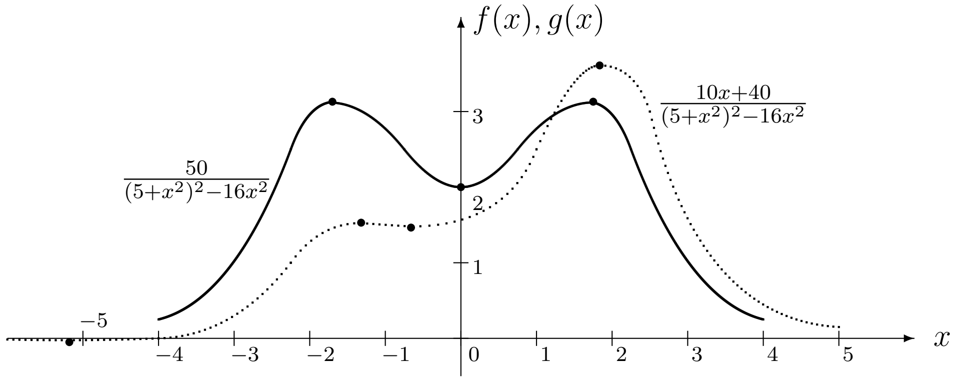
Wir diskutieren noch die Existenz absoluter Extremstellen: Auf einem offenen Intervall in $\mathbb{R}$, das $-\infty$ oder Polstellen oder ∞ als Randpunkte hat, nimmt die rationale Funktion $r(x)$ ein absolutes Minimum an, wenn die Grenzwerte an beiden Intervallenden ∞ sind, und ein absolutes Maximum, wenn die Grenzwerte beide $-\infty$ sind. Sind die Grenzwerte an den Intervallenden aber entgegengesetzt unendlich, so nimmt $r(x)$ auf einem solchen Intervall alle reellen Werte an und hat somit darauf keine absoluten Extremstellen. An Polstellen sind die einseitigen Grenzwerte immer $\pm\infty$, im Fall $m > n$ auch im Unendlichen. Im Fall $m \leq n$ aber ist der Grenzwert von $r(x)$ gleich a_n/b_n bei $x \to \pm\infty$ (mit $a_n := 0$ für $m < n$). Gibt es keine Polstellen, so hat dann $r(x)$ mindestens ein absolute Extremstelle auf $\mathbb{R}$; wenn sich $r(x)$ seinem Grenzwert bei $x \to \infty$ und bei $x \to -\infty$ von unterschiedlichen Seiten nähert, so gibt es sogar eine absolute Minimumstelle und eine absolute Maximumstelle auf $\mathbb{R}$. Gibt es aber Polstellen und ist x_1 die kleinste, x_k die größte, so muss man die Existenz von Extremstellen auf $]-\infty, x_1[$ und auf $]x_k, \infty[$ im Fall $m \leq n$ noch gesondert diskutieren. Hat $r(x)$ auf einem solchen Intervall nur Werte echt zwischen dem Grenzwert a_n/b_n im Unendlichen und dem Grenzwert $\pm\infty$ bei Annäherung an die Randpolstelle, so gibt es keine absoluten Extremstellen auf diesem Intervall, andernfalls mindestens eine. Ein absolutes Minimum auf dem maximalen Definitionsbereich D nimmt die rationale Funktion $r(x)$ genau dann an, wenn sie an allen Polstellen den Grenzwert $+\infty$ hat (insbesondere müssen dann alle Polstellen gerade Ordnung haben) und wenn sie im Fall $m > n$ auch bei $x \to \pm\infty$ den Grenzwert $+\infty$ hat, bzw. im Fall $m \leq n$ einen Wert annimmt, der nicht größer ist als der Grenzwert im Unendlichen. Die absoluten Minimumstellen auf D sind dann genau die Minimumstellen in den einzelnen Intervallen zwischen den Polstellen und $\pm\infty$, an denen $r(x)$ den kleinsten Wert hat. Analoges gilt für die Existenz und Bestimmung von absoluten Maximumstellen zu $r(x)$ auf D.

(9) Die einfachsten rationalen Funktionen $r(x)$ ohne Polstelle sind solche mit einem quadratischen Nenner $q(x)$ ohne Nullstelle auf $\mathbb{R}$, wie z.B. $1+x^2$. Betrachten wir Funktionen mit Grenzwert Null im Unendlichen, so kann das Zählerpolynom höchstens Grad 1 haben. Bei konstantem Zähler $p \neq 0$ ist die Gleichung $r'(x) = 0$ für die kritischen Punkte äquivalent mit der linearen Gleichung $q'(x) = 0$, also gibt es dann genau einen kritischen Punkt, und weil dort $r'(x)$ einen Vorzeichenwechsel hat, liegt eine lokale strikte Extremstelle vor. Da in diesem Fall $r(x) = p/q(x)$ nur positive bzw. nur negative Werte hat, handelt es sich sogar um eine eindeutige absolute Extremstelle auf $\mathbb{R}$. (Die Existenz einer absoluten Extremstelle ist auch klar, weil $r(x)$ den Limes 0 im Unendlichen hat. Ist das Zählerpolynom $p(x) = ax+b$ linear mit $a \neq 0$, so ist die Gleichung für die kritischen Punkte äquivalent mit der quadratischen Gleichung $aq(x) - (ax + b)q'(x) = 0$. Diese hat genau zwei Lösungen, und diese beiden kritischen Stellen sind die eindeutige absolute Minimum- und Maximumstelle von $r(x)$ auf $\mathbb{R}$; denn weil hier $r(x) = (ax+b)/q(x)$ positive und negative Werte besitzt und im Unendlichen den Grenzwert Null hat, gibt es ja beide Extremstellen auf $\mathbb{R}$.

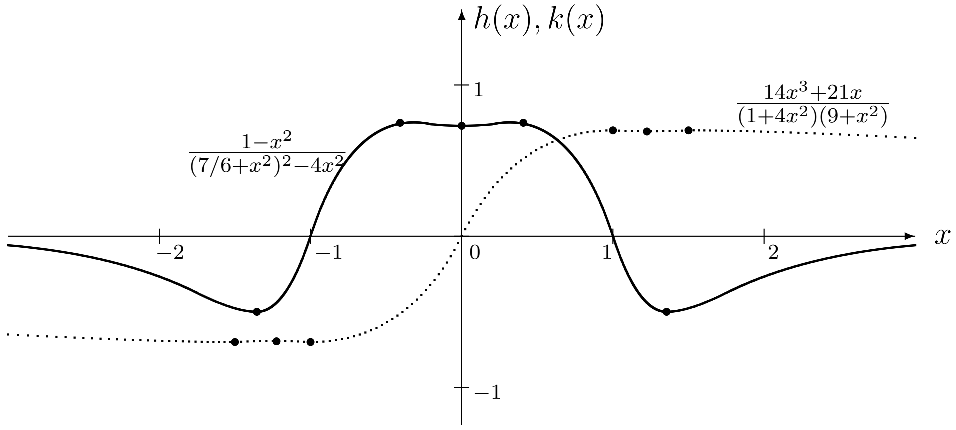


Die Abbildung zeigt die Situation für die konkreten Funktionen $1/(1+x^2)$ und $2x/(1+x^2)$. Bei der ersten ist $x = 0$ der einzige kritische Punkt, und es ist klar, dass dort eine absolute Maximumstelle auf $\mathbb{R}$ liegt. Bei der zweiten ist $2(1+x^2)-(2x)^2 = 0$ bzw. äquivalent $x^2 = 1$ die Gleichung für die kritischen Punkte, also sind 1 und -1 die kritischen Stellen, und da die Funktion positiv auf $\mathbb{R}_{>0}$ ist und negativ auf $\mathbb{R}_{<0}$, liegt in $x = 1$ das absolute Maximum, in $x = -1$ das absolute Minimum.

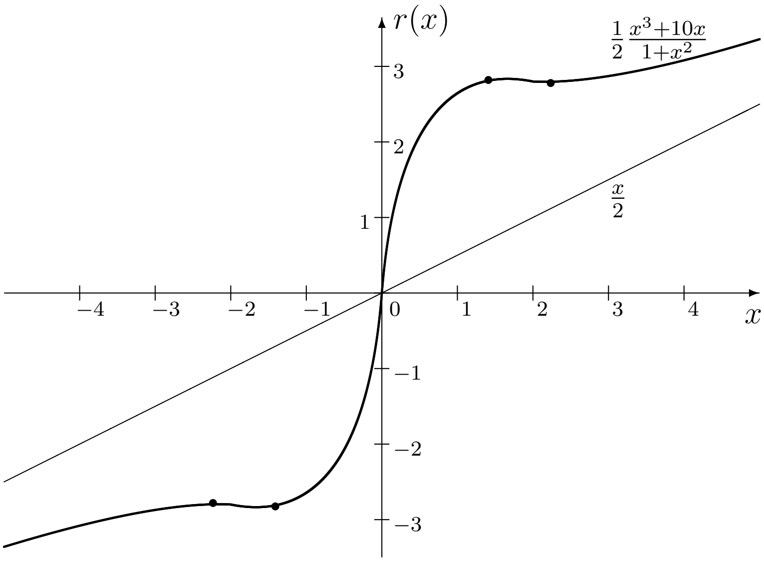
Mehr als zwei kritische Punkte können auftreten, wenn wir ein Nennerpolynom $q(x)$ vom Grad 4 ohne Nullstelle auf $\mathbb{R}$ und ein Zählerpolynom $p(x)$ vom Grad $m \leq 3$ betrachten; die Gleichung für die kritischen Punkte $p'(x)q(x) - p(x)q'(x) = 0$ hat dann den Grad $4+m-1$, also gibt es bis zu $4+m-1 \leq 6$ kritische Stellen. Darunter ist mindestens eine absolute Extremstelle von $r(x) = p(x)/q(x)$ und, wenn $p(x)$ positive und negative Werte annimmt, sogar mindestens eine absolute Minimumstelle und mindestens eine absolute Maximumstelle, weil $r(x)$ im Unendlichen den Grenzwert Null hat. Konkret betrachten wir $f(x) := 2\cdot\frac{5}{1+(x-2)^2}\cdot\frac{5}{1+(x+2)^2} = \frac{50}{(5+x^2)^2-16x^2}$ und $g(x) := \frac{10x+40}{(5+x^2)^2-16x^2}$. Durch Ableiten von $f(x)$ mit der Quotientenregel erhält man für die kritischen Punkte die Gleichung $32x - 4x(5+x^2) = 0$ oder $x^3 = 3x$ mit den Lösungen $0, \pm\sqrt{3}$ und den Funktionswerten $f(0) = 2$, $f(\pm\sqrt{3}) = \frac{50}{16} = 3.125$. Da $f(x)$ den Limes 0 im Unendlichen hat, sind folglich $x = \pm\sqrt{3}$ absolute Maximumstellen auf $\mathbb{R}$ und $x = 0$ eine lokale strikte Minimumstelle. Bei $g(x)$ findet man durch Ableiten die Gleichung $10(-3x^4 - 16x^3 + 6x^2 + 48x + 25) = 0$ für die kritischen Stellen. Die Lösungen können wir zwar nicht explizit bestimmen, aber durch Berechnung einiger Werte der linken Seite dieser Gleichung kann man sich davon überzeugen, dass es vier Vorzeichenwechsel gibt, und mit einem Näherungsverfahren (Intervallhalbierung) findet man $x_1 \approx -5.18352$ und $x_3 \approx -0.65842$ als kritische Punkte vom Typ $-/+$ sowie $x_2 \approx -1.32884$ und $x_4 = 1.83745$ als kritische Stellen vom Typ $+/-$. Vergleich der Werte $f(x_1) \approx -0.02021$, $f(x_2) \approx 1.52436$, $f(x_3) \approx 1.47944$ und $f(x_4) \approx 3.61641$ zeigt dann, dass in x_4 das absolute Maximum und in x_1 das absolute Minimum von $g(x)$ auf $\mathbb{R}$ liegt (letzteres ist sowieso klar, da $g(x)$ positive und negative Werte und wegen $\lim_{x\to\pm\infty} g(x) = 0$ ein positives Maximum und ein negatives Minimum auf $\mathbb{R}$ hat, und da $g(x) > 0$ ist für $x > -4$). Bei x_2 bzw. x_3 liegen eine lokale strike Maximumstelle bzw. Minimumstelle vor; diese lokalen Extrema sind sehr schwach ausgeprägt, so dass sie in der Abbildung zu einem Sattelpunkt zusammenzufallen scheinen.



Beispiele mit 5 bzw. 6 kritischen Punkten erhält man durch geeignete Wahl des Zählers als quadratisches bzw. kubisches Poylnom und des Nenners als Polynom vom Grad 4 ohne Nullstelle auf $\mathbb{R}$. Hier betrachten wir $h(x) := \frac{1-x^2}{(7/6+x^2)^2-4x^2}$ und $k(x) := \frac{2x}{1+4x^2} + \frac{x/3}{1+x^2/9} = \frac{14x^3+21x}{(1+4x^2)(9+x^2)}$. Für den Zähler der Ableitung von $h(x)$ berechnet man mit der Quotientenregel den Ausdruck $-2x[(7/6+x^2)^2-4x^2]+(x^2-1)2x[2(7/6+x^2)-4] = 2x(x^4-2x^2+11/36)$ und findet durch Lösen der biquadratischen Gleichung die kritischen Punkte $x=0,\ \pm\sqrt{1/6},\ \pm\sqrt{11/6}$. Bei $x=0$ hat $h'(x)$ einen Vorzeichenwechsel $-/+$, also liegt dort eine lokale strikte Minimumstelle vor. Da $h(x)$ positive und negative Werte annimmt und den Grenzwert Null im Unendlichen hat, nimmt $h(x)$ auf $\mathbb{R}$ ein Maximum und ein Minimum an. Beachtet man noch, dass $h(x)$ eine gerade Funktion ist, so ergeben sich $\pm\sqrt{11/6}$ als absolute Minimumstellen und $\pm\sqrt{1/6}$ als absolute Maximumstellen. Für $k(x)$ erhält man durch Nullsetzen der Ableitung die Gleichung $(1-x^2)(8x^4-30x^2+27)=0$ vom Grad 6 für die kritischen Punkte und man findet die Lösungen $\pm 1,\ \pm\sqrt{3/2},\ \pm\frac{3}{2}$. Betrachtung der Vorzeichenwechsel bei $k'(x)$ zeigt, dass 1 und $\frac{3}{2}$ lokale strikte Maximumstellen sind, $\sqrt{3/2}$ dagegen eine lokale isolierte Minimumstelle. Da $k(x)$ eine ungerade Funktion ist und wegen des Grenzwertes Null im Unendlichen Maximum und Minimum auf $\mathbb{R}$ annimmt, ist damit schon klar, dass nur 1 und $\frac{3}{2}$ als absolute Maximumstellen in Frage kommen und nur -1 oder $-\frac{3}{2}$ als absolute Minimumstellen. Werteberechnung gibt $k(1) = \frac{7}{10} = k(\frac{3}{2})$, also sind 1 und $\frac{3}{2}$ die absoluten Maximumstellen und -1 und $-\frac{3}{2}$ die absoluten Minimumstellen von $k(x)$ auf $\mathbb{R}$. Das lokale Minimum $k(\sqrt{3/2}) \approx 0.69985$ ist so schwach ausgeprägt, dass es in der Abbildung nicht zu erkennen ist.



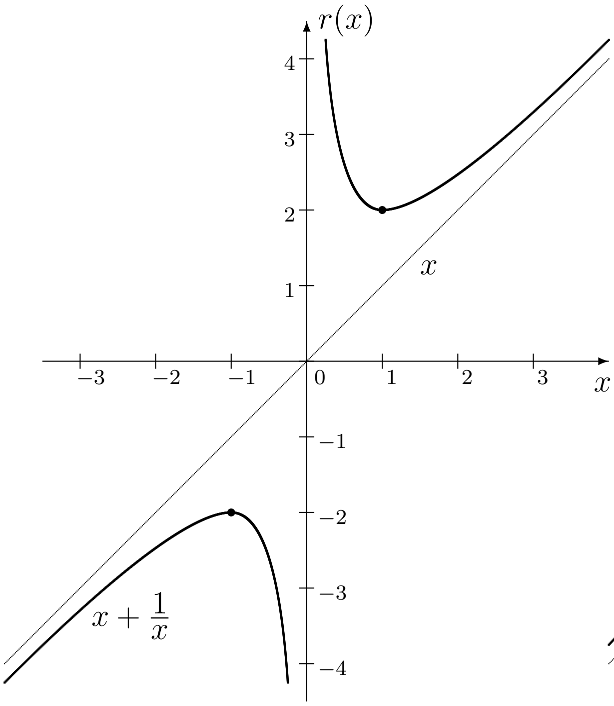
(10) Wir betrachten nun rationale Funktionen $r(x) = p(x)/q(x)$ mit unendlichem Grenzwert im Unendlichen, aber immer noch ohne Polstellen. (Den Fall eines endlichen Grenzwerts $\neq 0$ kann man durch Subtraktion einer Konstanten, nämlich des Quotienten der führenden Koeffizienten von p und q, auf den oben behandelten Fall des Grenzwerts Null im Unendlichen zurückführen.) Der einfachste Fall liegt vor, wenn q quadratisch ist, etwa $q(x) = 1+x^2$, und $p(x)$ ein kubisches Polynom. Dann hat $r(x)$ bei $x\to\infty$ und $x\to-\infty$ unendliche Grenzwerte mit entgegengesetztem Vorzeichen, also nimmt $r(x)$ nach dem Zwischenwertsatz alle reellen Zahlen als Werte an und hat natürlich keine absoluten Extremstellen auf $\mathbb{R}$. Die Gleichung $p'(x)q(x)-p(x)q'(x)=0$ ist jetzt vom Grad 4, also gibt es höchstens 4 kritische Punkte, unter Umständen aber gar keinen, wie das Beispiel $p(x) := x^3+x$, $q(x) := x^2+1$, $r(x) = x$ zeigt. Ein Beispiel mit der maximal möglichen Zahl von vier kritischen Punkten, von denen dann zwei lokale isolierte Minimumstellen und zwei lokale isolierte Maximumstellen sein müssen, erhält man in der Form $r(x) := ax+b\frac{x}{1+x^2}$ mit geeigneten Parametern $a>0$ und $b>0$. Ist a klein genug, so erzeugen die Minimumstelle von $\frac{x}{1+x^2}$ in $x = -1$ und die Maximumstelle in $x = 1$ entsprechende lokale Extremstellen von $r(x)$ nahe -1 und 1; wegen des Grenzverhaltens von $r(x)$ bei $x \to -\infty$ und $x \to \infty$ gibt es dann außerdem noch eine weitere lokale Maximumstelle links von $x = -1$ und eine weitere lokale Minimumstelle rechts von 1. Konkret hat $r(x) := \frac{1}{2}x + \frac{9}{2}\frac{x}{1+x^2} = \frac{1}{2}\frac{x^3+10x}{1+x^2}$ die biquadratische Gleichung $2x^4 - 14x^2 + 20 = 0$ für die kritischen Punkte, welche die Lösungen $x = \pm\sqrt{2}, \pm\sqrt{5}$ hat Davon sind $-\sqrt{5}$ und $\sqrt{2}$ lokale strikte Maximumstellen und $-\sqrt{2}$ und $\sqrt{5}$ lokale strikte Minimumstellen. Bei $x \to \pm\infty$ ist $r(x)$ asymptotisch zu $\frac{x}{2}$.

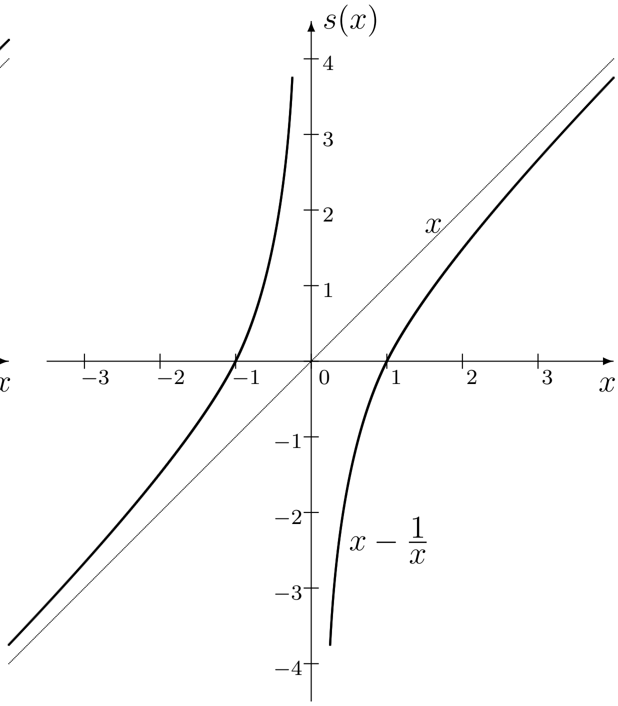


Ist $r(x) = p(x)/q(x)$ mit p vom Grad 4 und q vom Grad 2 ohne Nullstelle, so hat $r(x)$ bei $x \to \infty$ und $x \to -\infty$ denselben unendlichen Grenzwert, also in jedem Fall mindestens eine absolute Extremstelle auf $\mathbb{R}$ (eine absolute Minimumstelle, wenn der Grenzwert ∞ ist, eine absolute Maximumstelle, wenn er $-\infty$ ist). Die Gleichung für die kritischen Punkte hat jetzt den Grad 5, also gibt es maximal 5 kritische Stellen, und man kann auch konkrete Beispiele mit 5 kritischen Punkten finden; jedoch ist das gar nicht so einfach, und wir gehen daher weiter zu

(11) wichtigeren Beispielen der *Extremstellenbestimmung bei rationalen Funktionen mit Polstellen*: Gebrochen lineare Funktionen $r(x) = \frac{ax+b}{cx+d}$ mit $c \neq 0$ und $ad - bc \neq 0$ (so dass r nicht konstant ist) haben keine kritischen Punkte, da $r'(x) = \frac{ad-bc}{(cx+d)^2}$ keine Nullstellen hat. Die Graphen der gebrochen linearen Funktionen unterscheiden sich vom Graphen der Funktion $\frac{1}{x}$, also von der Normalhyperbel, alle nur durch Verschiebung und Achsenstreckungen oder -stauchungen mit demselben Faktor auf beiden Achsen. Insbesondere besitzen solche Funktionen keine lokalen Extremstellen und sind gleichsinnig streng monoton auf den beiden Intervallen, in die $\mathbb{R}$ durch die Polstelle $-\frac{d}{c}$ zerlegt wird.

Wenn wir also rationale Funktionen mit genau einer Polstelle (von der Ordnung 1) und mit lokalen Extremstellen suchen, so müssen wir mindestens quadratische Zähler ins Auge fassen, also etwa $r(x) = \frac{1}{x}(ax^2 + bx + c) = ax + b + c/x$ betrachten mit $a \neq 0 \neq c$. Funktionen dieser Form (mit $a < 0$, $c < 0$) treten z.B. als Gewinnfunktionen auf $\mathbb{R}_{>0}$ in bestimmten ökonomischen Modellen auf. Die Gleichung für die kritischen Punkte ist dann $a - c/x^2 = 0$ oder $x^2 = \frac{c}{a}$, also gibt es genau zwei kritische Punkte $\pm\sqrt{\frac{c}{a}}$, wenn a und c dasselbe Vorzeichen haben, keine dagegen bei verschiedenen Vorzeichen. Im ersten Fall hat $r(x)$ denselben unendlichen Grenzwert an den beiden Endpunkten des Intervalls $]0, \infty[$ und den entgegengesetzt unendlichen Grenzwert an den beiden Endpunkten des Intervalls $]-\infty, 0[$; also hat $r(x)$ auf einem dieser Intervalle eine absolute Minimumstelle und auf dem anderen eine absolute Maximumstelle, und diese Stellen müssen gerade die beiden kritischen Punkte sein. (Auf dem maximalen Definitionsbereich $\mathbb{R}_{\neq 0}$ hat r wie überhaupt jede rationale Funktion mit Polen ungerader Ordnung natürlich keine absoluten Extremstellen.) Im zweiten Fall ist $r(x)$ auf beiden Seiten von 0 gleichsinnig streng monoton. Die Abbildungen zeigen die Situation für $r(x) := x + \frac{1}{x}$ und $s(x) := x - \frac{1}{x}$.



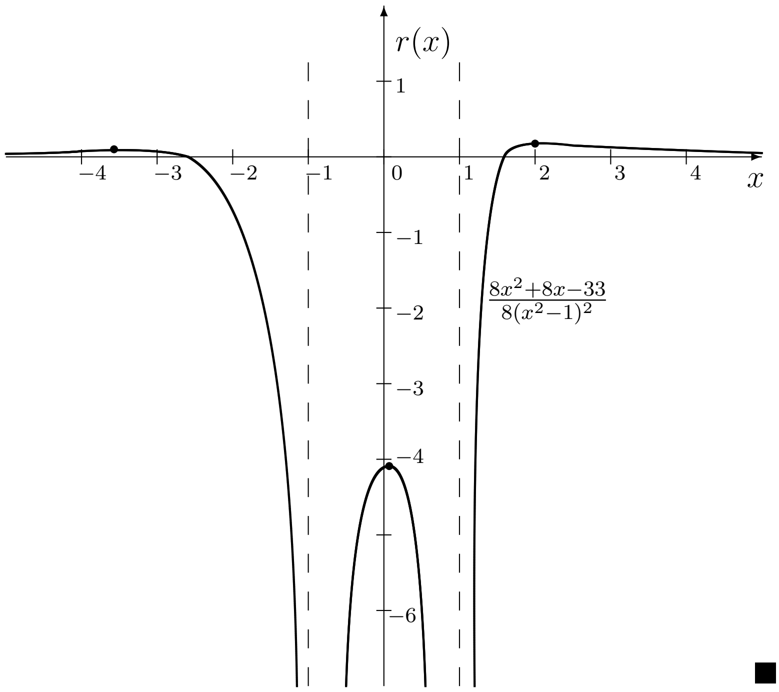


Bis zu drei kritische Punkte können auftreten bei $r(x) = \frac{ax^3+bx^2+c}{x} = ax^2 + bx + c\frac{1}{x}$, wobei wir $a > 0$ und $c > 0$ annehmen. (Andere Fälle kann man durch Übergang von x zu $-x$ oder von $r(x)$ zu $-r(x)$ darauf zurückführen.) Dann hat $r(x)$ bei $x \downarrow 0$ und bei $x \to \infty$ denselben Grenzwert ∞, also nimmt $r(x)$ auf dem Intervall $]0, \infty[$ ein absolutes Minimum an. Auf dem anderen Intervall $]-\infty, 0[$ hat $r(x)$ am Rand entgegengesetzt unendliche Grenzwerte, also nimmt die Funktion $r(x)$ dort alle reellen Zahlen als Werte an. Die Gleichung für die kritischen Punkte lautet hier $2ax^3 + bx^2 - c = 0$ und hat maximal 3 Lösungen, von denen eine die Minimumstelle x_* auf $]0, \infty[$ sein muss. Andere Lösungen müssen in $]-\infty, 0[$ liegen; denn hätte $r(x)$ in $]0, \infty[$ noch eine lokale Maximumstelle oder einen Sattelpunkt, so würde $r(x)$ den entsprechenden Funktionswert mit Vielfachheit mindestens 4 annehmen, was aber nicht möglich ist, da Zähler- und Nennergrad kleiner als 4 sind. Gibt es außer x_* keine kritischen Punkte oder nur einen Sattelpunkt, so ist $r(x)$ auf $]-\infty, 0[$ streng fallend; gibt es noch zwei weitere kritische Punkte, so sind diese eine lokale Minimumstelle und eine lokale Maximumstelle von $r(x)$ im Intervall $]-\infty, 0[$.

Bei einem Pol zweiter Ordnung gibt es keine kritischen Punkte, wenn das Zählerpolynom konstant ist wie etwa bei $r(x) = \frac{1}{x^2}$. Ist das Zählerpolynom linear wie bei $r(x) = \frac{x+c}{x^2}$ mit $c \neq 0$, so besitzt die rationale Funktion eine absolute Extremstelle, weil sie z.B. im Fall $c > 0$ auf $]0, \infty[$ positiv ist und auf $]-\infty, 0[$ ein negatives Minimum annimmt wegen der Grenzwerte 0 und ∞ am Rande und wegen $r(x) < 0$ für $x < -c$. Da die Gleichung für die kritischen Punkte $0 = r'(x) = -x^{-3}(x + 2c)$ nur die eine Lösung $x = -2c$ hat, liegt an dieser Stelle das absolute Extremum, und weitere lokale Extremstellen oder Sattelpunkte gibt es nicht. Der Fall eines quadratischen Zählers $r(x) = \frac{ax^2+bx+c}{x^2}$ unterscheidet sich nur um die additive Konstante a von den vorherigen Fällen. Im Falle eines kubischen Zählers $r(x) = \frac{x^3+bx+c}{x^2} = x + b\frac{1}{x} + c\frac{1}{x^2}$, etwa mit $c > 0$, haben wir an beiden Randpunkten von $]0, \infty[$ den Grenzwert ∞ und daher eine absolute Minimumstelle x_* in diesem Intervall. Am Rande von $]-\infty, 0[$ hat $r(x)$ dagegen entgegengesetzt unendliche Grenzwerte, nimmt also auf diesem Intervall alle reellen Zahlen als Werte an. Die Gleichung für die kritischen Stellen ist hier $x^3 - bx - 2c = 0$ und hat maximal drei Lösungen, von denen x_* eine ist. Die Situation ist also ganz ähnlich wie oben $r(x) = \frac{1}{x}p(x)$ mit p vom Grad 3 (und lässt sich durch Übergang von x zu $\frac{1}{x}$ auch darauf zurückführen).

Bei $r(x) = \frac{x^4+ax^3+bx+c}{x^2}$ mit $c > 0$ haben wir auf den Intervallen $]0, \infty[$ und $]-\infty, 0[$ jeweils den Limes ∞ an beiden Randpunkten; daher gibt es in jedem Intervall eine absolute Minimumstelle und die mit dem kleineren Funktionswert ist dann absolute Minimumstelle von $r(x)$ auf $\mathbb{R}_{\neq 0}$ (bzw. beide sind absolute Minimumstellen auf $\mathbb{R}_{\neq 0}$, wenn die Funktionswerte an diesen beiden Stellen gleich sind). Die Gleichung für die kritischen Punkte lautet hier $2x^4 + ax^3 - bx - 2c = 0$ und hat maximal vier Lösungen, zu denen natürlich die beiden Minimumstellen in den Intervallen $]0, \infty[$ und $]-\infty, 0[$ gehören. Eventuell vorhandene weitere kritische Punkte müssen dann eine zusätzliche lokale Minimum- und eine zusätzliche lokale Maximumstelle in einem dieser Intervalle sein oder ein Sattelpunkt. Das ergibt sich aus dem Grenzverhalten am Rande der Intervalle und der Tatsache, dass lokale Minimum- und Maximumstellen im Wechsel aufeinander folgen. Ist $r(x) = \frac{x^4+c}{x^2}$ gerade, so kann es folglich keine weiteren kritischen Punkte geben.

Wir betrachten schließlich noch ein Beispiel mit zwei Polstellen zweiter Ordnung $r(x) := \frac{8x^2+8x-33}{8(x^2-1)^2}$. Bei ± 1 hat diese Funktion Pole zweiter Ordnung mit dem Grenzwert $-\infty$, weil das Zählerpolynom dort einen negativen Wert hat. Da der Grenzwert von $r(x)$ bei $x \to \pm\infty$ Null ist und da $r(x)$ positiv ist für große Werte des Betrags von x, nimmt folglich $r(x)$ auf jedem der drei Intervalle $]-\infty, -1[$, $]-1, 0[$, $]1, \infty[$ ein Maximum an. Die Gleichung für die kritischen Punkte lautet hier $(x^2-1)(16x+8) - 4x(8x^2+8x-33) = 0$ oder äquivalent $2x^3 + 3x^2 - \frac{29}{2}x + 1 = 0$. Als eine Lösung kann man $x_1 = 2$ erraten. Durch Lösung einer quadratischen Gleichung findet man dazu noch die weiteren Lösungen $x_{2,3} = -\frac{7}{4} \pm \sqrt{53}$, d.h. $x_2 \approx 0.07$ und $x_3 \approx -3.57$. Der Wertevergleich $r(x_1) = \frac{5}{24} \approx 0.21$, $r(x_2) \approx -4.09$, $r(x_3) \approx 0.04$ erweist alsdann $x_1 = 2$ als die absolute Maximumstelle von r auf dem maximalen Definitionsbereich $\mathbb{R}$ ohne ± 1. Beachte, dass die $-\infty$–Stellen ± 1 hier als "uneigentliche" lokale Minimumstellen gelten, so dass also die lokalen Maximumstellen und Minimumstellen der Funktion im Wechsel aufeinander folgen, wenn man den Bereich $\mathbb{R}$ durchläuft.



■

Diese Beispiele demonstrieren den praktischen Nutzen der Extremstellenbestimmung mittels Differentialrechnung: Man kann damit (wenn man die Nullstellen der Ableitung berechnen kann) die Maxima und Minima von schon recht komplizierten elementaren Funktionen bestimmen, bei denen man mit direkten Überlegungen überhaupt keine Chance hätte. Die notwendige Bedingung für innere Extremstellen einer reellen Funktion von einer Variablen, also das Verschwinden der Ableitung $f'(x_0)$, hat aber auch theoretische Anwendungen in der Ökonomie. Sie ist nämlich der simple mathematische Hintergrund von vielen "ökonomischen Gesetzen". Wir geben dazu nun einige

BEISPIELE (***ökonomische Anwendungen des notwendigen Kriteriums für Extremstellen***):

(1) *Im Gewinnmaximum verschwindet der Grenzgewinn, wenn nicht an der Grenze der Produktionskapazität produziert wird;*

und dann sind Grenzerlös und Grenzkosten gleich.

Das erste ökonomische Gesetz ist nichts anderes als die Aussage, dass für den Grenzgewinn $G'(x^*) = 0$ gilt in inneren Maximumstellen x^* der Gewinnfunktion $G(x)$ als Funktion des Produktions–Output $[0, x_{\max}]$. Dass nicht an der Grenze der Produktionskapazität produziert wird, bedeutet $x^* < x_{\max}$, und natürlich ist unterstellt, dass überhaupt etwas produziert wird, also $x^* > 0$, so dass x^* im Inneren $]0, x_{\max}[$ des Definitionsintervalls liegt. Da der Gewinn $G(x)$ definiert ist als Differenz $E(x) - K(x)$ von Erlös und Kosten, folgt aus der Differenzregel für die Ableitung $E'(x) - K'(x) = 0$, also das zweite "ökonomische Gesetz". Ökonomisch interpretiert bedeutet dies, dass der Mehrerlös für eine zusätzlich produzierte (kleine) Einheit (fast) exakt kompensiert wird durch die Mehrkosten dieser Produktionsausweitung, und das ist auch aus rein ökonomischer Sicht ein plausibles Kriterium für Produktion mit maximalem Gewinn: Würde der Mehrerlös die Mehrkosten übersteigen, so könnte man den Gewinn durch Mehrproduktion steigern (solange die Kapazitätsgrenze nicht erreicht ist); wäre der Erlös für die zusätzlich produzierte Einheit aber kleiner als die Mehrkosten, so gälte dies auch für die letzte produzierte Einheit (weil kleine Einheiten angenommen sind), also würde eine Verringerung der Produktion um eine Einheit den Gewinn vergrößern.

(2) Natürlich gilt eine zum ersten "Gesetz" analoge Aussage für jede (differenzierbare) ökonomische Funktion $f(x)$ einer Variablen: An einer inneren Extremstelle x_0 von f verschwindet der Wert der zugehörigen Grenz– bzw. Marginalfunktion $f'(x_0)$. Sind abhängige und unabhängige Variable positiv, so ist äquivalent, dass f dort vollkommen elastisch ist, dass also die Elastizität $\varepsilon_f(x_0) = x_0 f'(x_0)/f(x_0)$ verschwindet In dieser Terminologie aus 4.4 lautet also das notwendige Kriterium für innere Extremstellen:

In inneren Extremstellen ist eine ökonomische Funktion vollkommen elastisch.

(3) Wird das Gewinnmaximum an der Kapazitätsgrenze $x^* = x_{\max}$ erreicht, so gilt dort das notwendige Kriterium für Randextremstellen $0 \leq G'(x_{\max})$, also $E'(x_{\max}) \geq K'(x_{\max})$; denn im anderen Fall wäre ja $G(x) > G(x_{\max})$ für x nahe $x_{\max}$ (und $x < x_{\max}$).

Wird der maximale Gewinn an der Kapazitätsgrenze erreicht, so ist der Grenzgewinn nichtnegativ und der Grenzerlös nicht kleiner als die Grenzkosten.

Schreiben wir den Erlös $E(x) = x \cdot p(x)$ mit Nachfragefunktion $x(p)$ und umkehrender Preisfunktion $p(x)$, so haben wir gemäß Produktregel $E'(x) = p(x) + xp'(x)$. Für einen monopolistischen Anbieter ist $x'(p) < 0$ (eine Preiserhöhung reduziert die Nachfrage), also auch $p'(x) = 1/x'(p) < 0$. Außerdem gilt natürlich $K'(x) > 0$ (Mehrproduktion erhöht die Kosten). In einer Maximumstelle x^* für den Gewinn, sei sie im Inneren oder an der Kapazitätsgrenze, gilt also $0 < K'(x^*) \leq E'(x^*) = p(x^*) + x^*p'(x^*)$, und es folgt $0 > \varepsilon_p(x^*) = x^*p'(x^*)/p(x^*) > -1$, also $\varepsilon_x(p^*) = 1/\varepsilon_p(x^*) < -1$ für den zu x^* gehörenden Preis $p^* = p(x^*)$. Das haben wir für eine innere Gewinnmaximumstelle schon in 4.4 gesehen (als Folgerung aus der Amoroso–Robinson–Beziehung). Wenn wir uns daran erinnern, dass $x(p)$ "elastisch" heißt in Bereichen, wo $|\varepsilon_x(p)| > 1$ ist (hier also $\varepsilon_x(p) < -1$), so können wir also folgendes ökonomische Gesetz aussprechen:

Bei Produktion mit maximalem Gewinn ist die Nachfrage elastisch bzgl. des Preises.

(4) Die Stückgewinnfunktion $g(x) = G(x)/x$ strebt normalerweise gegen $-\infty$ bei $x \searrow 0$ (wegen der Fixkosten), nimmt also eine Maximum an (mindestens) einer Stelle ξ^* auf $]0, x_{\max}]$ an (und wir nehmen an, dass dieses Maximum positiv ist). Gemäß Quotientenregel gilt $g'(\xi^*) = [\xi^* G'(\xi^*) - G(\xi^*)]/(\xi^*)^2 = [G'(\xi^*) - g(\xi^*)]/\xi^*$ und das notwendige Kriterium gibt $g'(\xi^*) = 0$, wenn $\xi^* < x_{\max}$ ist, bzw. $g'(\xi^*) \geq 0$, wenn $\xi^* = x_{\max}$ ist. Es folgt $G'(\xi^*) \geq g(\xi^*)$ mit Gleichheit, wenn ξ^* innere Maximumstelle von $g(x)$ ist. Für $0 < x < \xi^*$ mit $G(x) > 0$ gilt dann $G(\xi^*) = \xi^* g(\xi^*) \geq \xi^* g(x) = \frac{\xi^*}{x} \cdot G(x) > G(x)$, so dass $G(x)$ kein Maximum an einer Stelle $x < \xi^*$ haben kann. Im Fall $\xi^* < x_{\max}$ ist $G'(\xi^*) = g(\xi^*) > 0$, also kann auch ξ^* selbst keine Maximumstelle von $G(x)$ sein und es folgt $x^* > \xi^*$.

> *Der Produktions–Output ξ^* mit maximalem Stückgewinn ist nicht kleiner als der Output x^* mit maximalem Gewinn.*
>
> *Bei Produktion mit maximalem Stückgewinn ist der Stückgewinn nicht kleiner als der Grenzgewinn und beide sind gleich, wenn die Kapazität nicht ausgelastet ist.*

Das ist ökonomisch sehr plausibel: Wenn der Grenzgewinn, also der Gewinn für die Produktion der letzten (kleinen) Einheit, kleiner ist als der Durchschnittsgewinn, so wird der Durchschnittsgewinn durch Minderproduktion dieser Einheit gesteigert. Und solange der Stückgewinn durch Produktionsausweitung wächst, nimmt erst recht der Gewinn zu, also wird das Gewinnmaximum bei größerem Output erreicht als das Stückgewinnmaximum (oder beide an der Kapazitätsgrenze).

$\overline{(4)}$ Dieselbe Rechnung wie in (4) lässt sich für die Ableitung der Durchschnittsfunktion $\overline{f}(x) = f(x)/x$ zu einer beliebigen positiven Funktion $f(x)$ einer positiven Variablen x durchführen und gibt $\overline{f}\,'(x) = [xf'(x) - f(x)]/x^2 = [f'(x) - \overline{f}(x)]/x = f(x)[\varepsilon_f(x) - 1]/x^2$. Also gilt folgende *Charakterisierung von kritischen Punkten der Durchschnittsfunktion*:

$$\overline{f}\,'(x_0) = 0 \qquad \Longleftrightarrow \qquad f'(x_0) = \overline{f}(x_0) \qquad \Longleftrightarrow \qquad \varepsilon_f(x_0) = 1.$$

Diese einfache Beobachtung ist der inhaltliche Kern vieler ökonomischer Gesetze, z.B.:

> *Im Betriebsoptimum sind Grenzkosten und Stückkosten gleich und die Kosten ausgeglichen elastisch bzgl. des Produktions–Output.*
>
> *Im Betriebsminimum sind Grenzkosten und durchschnittliche variable Kosten gleich und die variablen Kosten ausgeglichen elastisch bzgl. des Output.*
>
> *Bei maximalem durchschnittlichen Produktionsertrag stimmen für jeden Produktionsfaktor Grenzproduktivität und durchschnittlicher Produktionsertrag überein und der Produktionsertrag ist ausgeglichen elastisch bzgl. des Faktoreinsatzes.*

Hierzu muss man nur wissen, dass unter **Betriebsoptimum** die Produktion bei minimalen Stückkosten und unter **Betriebsminimum** die Produktion bei minimalen Stück–variablen Kosten verstanden wird. Im Betriebsoptimum liegt also eine Minimumstelle x_0 der Stückkostenfunktion $k(x) = K(x)/x$ vor, und es folgt $K'(x_0) = k(x_0)$ bzw. äquivalent $\varepsilon_K(x_0) = 1$. Im Betriebsminimum hat man andererseits eine Minimumstelle x_1 der Funktion $\overline{K_{\text{var}}}(x) = K_{\text{var}}(x)/x$, also gilt dort $K'_{\text{var}}(x_1) = \overline{K_{\text{var}}}(x)$ und $\varepsilon_{K_{\text{var}}}(x_1) = 1$, wobei wegen Konstanz der Fixkosten K_{fix} auch $K'(x_1)$ statt $K'_{\text{var}}(x_1)$ geschrieben werden kann. Im dritten Gesetz ist der **Produktionsertrag** $x(r)$ in Output–Einheiten in Abhängigkeit von einem Produktionsfaktor r in Faktoreinheiten (bei c.p.–Bedingung für die anderen Produktionsfaktoren) gemeint. In diesem Kontext heißt $\overline{x}(r) = x(r)/r$ **durchschnittlicher Produktionsertrag** (gibt den Produktions–Output pro eingesetzter Faktoreinheit an) und $x'(r)$ **Grenzproduktivität** (gibt die Output–Erhöhung an, die durch den zusätzlichen Einsatz einer – kleinen – Faktoreinheit erreicht wird). In einer Maximumstelle r_0 von $\overline{x}(r)$ gilt also $x'(r_0) = \overline{x}(r_0)$ und $\varepsilon_x(r_0) = 1$, das ist das dritte angegebene Gesetz.

Allerdings gelten die drei angegebenen Gesetze nur unter der Prämisse, dass das jeweilige Minimum bzw. Maximum der Durchschnittsfunktion im Inneren des Definitionsintervalls angenommen wird, also nicht etwa in einer Kapazitätsgrenze $x_{\max}$ bzw. $r_{\max}$; denn sonst kann man ja nicht schließen, dass die Ableitung der Durchschnittsfunktion an dieser Extremstelle verschwindet. (Der andere Randpunkt 0 des Definitionsintervalls ist in ökonomisch sinnvollen Situationen als Minimum– bzw. Maximumstelle ausgeschlossen.) Ist $x_0 = x_{\max}$ bzw. $x_1 = x_{\max}$, so kann man an dieser Randminimumstelle nur $k'(x_0) \leq 0$ bzw. $\overline{K_{\text{var}}}'(x_1) \leq 0$ schließen und erhält die Aussage, dass die an der jeweiligen Minimumstelle die Grenzkosten höchstens so groß sind wie die Stückkosten bzw. die stückvariablen Kosten und dass die Kosten bzw. die variablen Kosten jedenfalls nicht elastisch sind. Ist $r_0 = r_{\max}$, so gilt an der Randmaximumstelle $x'(r_0) \geq 0$, also kann man noch aussagen, dass der durchschnittliche Produktionsertrag dort jedenfalls nicht unelastisch ist und höchstens so groß wie die Grenzproduktivität.

(5) *Optimale Produktion*: Wir betrachten eine *Produktions(ertrags)funktion* $x(r_1, \ldots, r_k)$, die den Produktions–Output in Abhängigkeit von mehreren eingesetzten Produktionsfaktoren $r_1, \ldots, r_k$ angibt. Sind $p_i(r_i)$ die *Faktorpreise* (Preis für eine Einheit des i–ten Faktors bei Einkauf von r_i Einheiten), so hat man Produktionskosten $K(r_1, \ldots, r_k) = r_1 p_1(r_1) + \ldots + r_k p_k(r_k) + K_{\text{fix}}$ und den Gewinn $G(r_1, \ldots, r_k) = E(x(r_1, \ldots, r_k)) - K(r_1, \ldots, r_k)$ in Abhängigkeit von den Produktionsfaktoren. Dabei ist $E(x)$ der Erlös für x produzierte Einheiten (die am Markt auch abgesetzt werden können, wie wir annehmen wollen). Werden mit denselben Produktionsfaktoren mehrere nicht konkurrierende Produkte $x_1, \ldots, x_l$ hergestellt, so kann man ihre Gewinnfunktionen $G_j(r_1, \ldots, r_k)$, $j = 1 \ldots l$, einfach addieren, um die Gesamtgewinnfunktion zu erhalten. Bei konkurrierenden Produkten hängt aber die Gesamtgewinnfunktion in komplizierterer Weise von den Outputs $x_1, \ldots, x_l$ und damit mittelbar von den Faktoren $r_1, \ldots, r_k$ ab; darauf gehen wir hier nicht weiter ein.

Betrachten wir nun alle Faktoreinsatzmengen als konstant bis auf eine, etwa $r = r_i$, so hat die Gewinnfunktion bei dieser c.p.–Bedingung die Gestalt

$$G(r) = E(x(r)) - rq(r) - K_{\text{fix}}$$

mit dem Faktorpreis $q(r) = p_i(r)$ und einer Fixkosten–Konstanten K_{fix}. Liegt beim Einsatz von r_0 Faktoreinheiten ein Gewinnmaximum vor (und ist r_0 nicht der maximal mögliche Faktoreinsatz), so gilt gemäß Ketten und Produktregel für den Outputwert $x_0 = x(r_0)$:

$$0 = G'(r_0) = E'(x_0) \cdot x'(r_0) - q(r_0) - r_0 q'(r_0)$$

oder äquivalent

$$\underbrace{E'(x_0)}_{\text{Grenzerlös}} \cdot \underbrace{x'(r_0)}_{\text{Grenzproduktivität}} = \underbrace{q(r_0)}_{\text{Faktorpreis}} \cdot [1 + \underbrace{\varepsilon_q(r_0)}_{\text{Elastizität des Faktorpreises}}] \,.$$

Wenn die Anzahl der eingekauften Faktoreinheiten den Faktorpreis nicht beeinflusst (“vollkommene Konkurrenz auf dem Faktormarkt”), so ist der Faktorpreis q konstant und die Bedingung vereinfacht sich zu

$$E'(x_0) \cdot x'(r_0) = q \,,$$

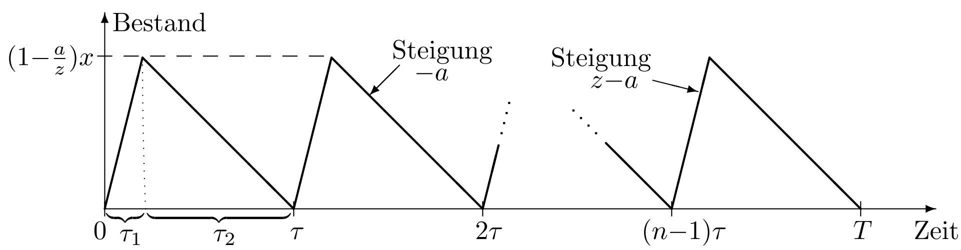
in Worten:

> *Das Produkt aus Grenzerlös und Grenzproduktivität ist gleich dem konstanten Faktorpreis, wenn im Gewinnmaximum produziert wird.*

Beim Produktionsfaktor Arbeit z.B. bedeutet dies, dass ein Unternehmer zusätzliche Arbeitskräfte zur Produktionssteigerung einsetzt, solange der mit einer Arbeitseinheit zusätzlich produzierbare Output $x(r+1) - x(r) \approx x'(r)$, bewertet mit dem Grenzerlös $E'(x)$, noch größer ist als die Kosten der Arbeitseinheit. Bei konstantem Marktpreis p für das Produkt ist $E(x) = p \cdot x$ und $E'(x) = p$ konstant. Es werden dann zusätzliche Arbeitskräfte eingestellt, solange die Kosten für eine Arbeitseinheit kleiner als der Marktwert $p \cdot x'(r) \approx p \cdot [x(r+1) - x(r)]$ des damit zusätzlich herstellbaren Outputs sind. (Das würde ein Produzent natürlich auch tun, wenn er ganz "naiv" agiert, ohne mit Hilfe der Differentialrechnung den Gewinn–maximierenden Faktoreinsatz auszurechnen.)

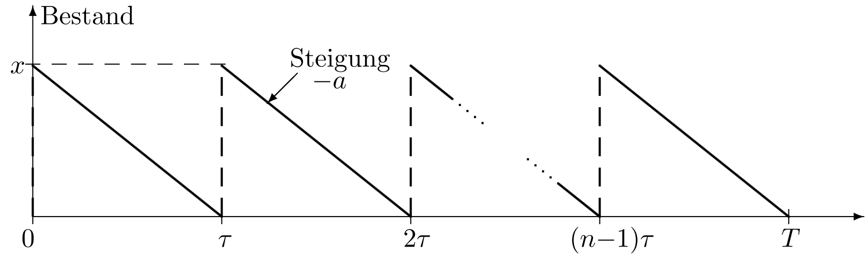
(6) *Optimale Lagerhaltung*: Lagerhaltungsprobleme können – auch mathematisch – sehr kompliziert sein. Hier betrachten wir nur einen einfachen Fall und gehen davon aus, dass aus einem Lager ein Gut mit bekannter konstanter Rate abgeht und in gleichen Zeitabständen wieder aufgefüllt wird. Man kann sich dabei ein *Einlieferungslager* vorstellen, in dem ein Rohstoff oder Vorprodukt für eine Produktion bereitgehalten wird, oder ein *Auslieferungslager*, in dem ein produziertes Gut zum Absatz am Markt aufbewahrt wird. In beiden Fällen entstehen einerseits *Lagerkosten*, die es angeraten sein lassen, den Lagerbestand nicht zu hoch werden zu lassen (von Kapazitätsgrenzen sehen wir hier ganz ab), und andererseits mit jedem Auffüllvorgang (Nachbestellung beim Eingangslager, Einlieferung einer neuen Produktionscharge beim Auslieferungslager) verbundene sog. *Rüstkosten*, die es nahe legen, die Auffüllmengen jeweils groß zu wählen. Dies sind ersichtlich gegenläufige Kostenarten, und es entsteht das Problem, die *optimale Losgröße* (optimale Bestellmenge, optimale Auflagengröße) festzulegen, so dass langfristig die Summe der Lagerkosten und der Rüstkosten minimiert wird.

Wir nehmen an, dass in einem (großen) Zeitraum T (Zeiteinheiten) die *Gesamtmenge* X (Mengeneinheiten) des gelagerten Gutes mit bekannter konstanter *Abgangsrate* a (Mengeneinheiten pro Zeiteinheit, $X = a \cdot T$) abgeht und dass bei jedem Auffüllvorgang die *Losgröße* x (Mengeneinheiten) mit konstanter *Zugangsrate* z (Mengeneinheiten pro Zeiteinheit) zugeht, wobei natürlich $z > a$ sein muss (sonst würde der Lagerbestand nie zunehmen). Der Extremfall $z = \infty$ ist dabei auch zugelassen und bedeutet, dass der Zugang "schlagartig" erfolgt. Der zeitliche Abstand zwischen zwei Auffüllvorgängen sei $\tau = \frac{T}{n}$ (Zeiteinheiten, $n \in \mathbb{N}$). Es wäre unsinnig, langfristig mehr zu lagern als abgeht, also wird man die Losgröße x und die Zeitspanne τ so einrichten, dass $x = \frac{X}{n} = \frac{X \cdot \tau}{T} = a \cdot \tau$ ist, d.h. langfristig wird nur so viel aufgefüllt, wie auch abgeht. Der Lagerbestand entwickelt sich nun so, dass er ab den Zeitpunkten $0, \tau, 2\tau, \ldots, (n-1)\tau$ jeweils während eines Zeitraums $\frac{x}{z}$ mit Rate z aufgefüllt wird (d.h. nach der Zeit $\frac{x}{z}$ sind gerade x Mengeneinheiten nachgefüllt), während gleichzeitig in den Auffüllzeiträumen und außerhalb dieser Zeiten ein Abgang mit Rate a erfolgt. Für den Zeitraum $\tau_1 = \frac{x}{z}$ hat man per Saldo dann einen Zugang mit Rate $z-a$, wobei der Bestand vom angenommenen Anfangswert 0 bis auf $(z-a)\tau_1 = (1-\frac{a}{z})x$ anwächst, im anschließenden Zeitraum $\tau_2 = \tau-\tau_1$ fällt er dann von $(1-\frac{a}{z})x$ auf $(1-\frac{a}{z})x - a\tau_2 = x - a(\tau_1+\tau_2) = x - a\tau = 0$.



Lagerbestand im Fall $a < z < \infty$

Im Fall $z = \infty$ springt der Lagerbestand zu den Zeitpunkten 0, τ, $2\tau, \ldots$ jeweils von 0 auf x und wird danach während des Zeitraums τ mit Rate a wieder auf 0 abgebaut.



Lagerbestand im Fall $a > 0,\ z = \infty$

Der *durchschnittliche Lagerbestand* ist wegen der linearen zeitlichen Entwicklung in den Zeitintervallen $[0, \tau_1]$ und $[\tau_1, \tau]$ im Fall $z < \infty$ bzw. im Intervall $[0, \tau]$ im Fall $z = \infty$ jeweils gerade die Hälfte des Maximalwertes, also gleich $\frac{1}{2}(1-\frac{a}{z})x$ bzw. $\frac{1}{2}x$. Für die Lagerkosten machen wir die einfachste mögliche Annahme, dass sie proportional zum durchschnittlichen Lagerbestand sind mit einem Proportionalitätsfaktor k_{L} (in Geldeinheiten pro Zeit- und Mengeneinheit). Für den Gesamtzeitraum T ergeben sich damit als *Lagerkosten*:

$$K_{\mathrm{L}} = \frac{1}{2}\Big(1 - \frac{a}{z}\Big)x \cdot T \cdot k_{\mathrm{L}} \qquad \text{(mit } \tfrac{a}{z} := 0, \text{ wenn } z = \infty).$$

Für die *Rüstkosten* nehmen wir an, dass sie bei jedem Auffüllvorgang in derselben Höhe k_{R} (Geldeinheiten) entstehen, so dass im Zeitraum T insgesamt

$$K_{\mathrm{R}} = n \cdot k_{\mathrm{R}} = \frac{X}{x} \cdot k_{\mathrm{R}}$$

anfallen. Damit haben wir die *Gesamtkosten* der Lagerhaltung für den Zeitraum T in Abhängigkeit von der Losgröße x ermittelt (wenn wir noch $X = a \cdot T$ beachten):

$$K(x) = \frac{a}{2}\Big(1 - \frac{a}{z}\Big) \cdot k_{\mathrm{L}} \cdot X \cdot x + k_{\mathrm{R}} \cdot X \cdot \frac{1}{x} \qquad \text{(mit } \tfrac{a}{z} := 0, \text{ wenn } z = \infty).$$

Die Aufgabe ist jetzt, diese Funktion auf $\mathbb{R}_{>0}$ zu minimieren. Dies ist im Vergleich zur Aufstellung der Kostenfunktion einfach: Zunächst sieht man (wegen $1-\frac{a}{z} > 0$), dass $K(x) \to \infty$ strebt bei $x \to \infty$ und bei $x \searrow 0$, also gibt es (mindestens) eine absoluten Minimumstelle x_* und die lässt sich aus der Ableitungsbedingung $K'(x_*) = 0$ bestimmen. Es ist

$$K'(x) = \frac{a}{2}\Big(1 - \frac{a}{z}\Big) \cdot k_{\mathrm{L}} \cdot X - k_{\mathrm{R}} \cdot X \cdot \frac{1}{x^2} \qquad \text{(mit } \tfrac{a}{z} := 0, \text{ wenn } z = \infty),$$

und die Gleichung $K'(x) = 0$ hat eine einzige positive Lösung x_*, die dann auch die gesuchte Minimumstelle ist. Das Ergebnis ist die **Andlersche Losgrößenformel** für die Kosten–optimale Losgröße:

$$x_* = \sqrt{\frac{k_{\mathrm{R}}}{k_{\mathrm{L}}} \cdot \frac{2z}{a(z-a)}} \quad \text{im Fall } z < \infty \text{ bzw.} \qquad x_* = \sqrt{\frac{k_{\mathrm{R}}}{k_{\mathrm{L}}} \cdot \frac{2}{a}} \quad \text{im Fall } z = \infty.$$

Das Ergebnis x_* hängt nicht ab vom Gesamtzeitraum T und der in dieser Zeit aus dem Lager entnommenen Menge X, sondern nur von den Abgangs– bzw. Zugangsraten a und z und von den Konstanten k_{L} und k_{R}, welche die Lager– bzw. Rüstkosten festlegen. Im Allgemeinen wird X kein ganzzahliges Vielfaches von x_* sein und damit T kein ganzzahliges Vielfaches von dem zur optimalen Losgröße gehörenden Zeitabstand $\tau_* = \frac{x_*}{a}$ zwischen zwei Auffüllvorgängen. In der Praxis wählt man die tatsächlichen Losgrößen x und Perioden τ natürlich als ganzzahlige Vielfache gewisser Einheiten, und zwar möglichst nahe bei den berechneten optimalen Werten x_* und τ_*. ■

Das letzte Beispiel ist typisch für Anwendungen der Mathematik in der Wirtschaftswissenschaft: Die eigentliche Schwierigkeit ist die Aufstellung einer elementaren Formel für die zu optimierende ökonomische Funktion, die zunächst nur indirekt durch ökonomische Sachverhalte beschrieben ist. Die Bestimmung der Extremstellen mittels Differentialrechnung ist dann meist eine ganz einfache Sache. Dabei ist die Existenz der gesuchten Minimum– oder Maximumstelle aufgrund der ökonomischen Interpretation der Funktion meist von vorneherein klar, so dass man die Existenzfrage beiseite lassen und sich direkt der Berechnung der Ableitungsnullstellen zuwenden kann. Zu beachten ist aber stets, dass das Optimum evtl. in Randpunkten des ökonomisch sinnvollen Definitionsbereichs erreicht wird oder auch in Nichtdifferenzierbarkeitsstellen der Funktion – das Optimum muss also nicht immer in einer Ableitungsnullstelle liegen! Im Übrigen sollte man die Bedeutung der mathematischen Optimierung nicht überschätzen: Ökonomische Funktionen sind oft nicht exakt bekannt und Änderungen der geschätzten Parametern können zu ganz anderen Extremstellen führen. Die Rechenergebnisse müssen immer kritisch hinterfragt werden, ob sie auch ökonomisch sinnvoll sind!

Wir schließen diesen Abschnitt mit Kriterien höherer Ableitungsordnung für lokale Extrema, die manchmal nützlich sein können, um den Typ eines kritischen Punkts zu bestimmen (lokale Minimumstelle oder lokale Maximumstelle oder Sattelpunkt?), deren Bedeutung aber oft – vor allem in der Schulmathematik – überschätzt wird.

SATZ (**Kriterien höherer Ordnung für lokale Extrema**):
Ist $f'(x_0) = 0, \ldots, f^{(k-1)}(x_0) = 0$ *und* $f^{(k)}(x_0) \neq 0$ *die nichtverschwindende Ableitung kleinster Ordnung* $k \in \mathbb{N}_{\geq 2}$ *an der inneren Stelle* x_0 *des Definitionsintervalls der reellen Funktion* f*, so gilt:*

(i) (notwendiges Kriterium) *Ist* x_0 *lokale Minimumstelle von* f*, so ist* k *gerade und* $f^{(k)}(x_0) \geq 0$*; ist* x_0 *lokale Maximumstelle von* f*, so ist* k *gerade und* $f^{(k)}(x_0) \leq 0$.

(ii) (hinreichendes Kriterium) *Ist* k *gerade und* $f^{(k)}(x_0) > 0$*, so ist* x_0 *strikte lokale Minimumstelle von* f*; ist* k *gerade und* $f^{(k)}(x_0) < 0$*, so ist* x_0 *strikte lokale Maximumstelle von* f.

(iii) (Sattelpunktkriterium) *k ist genau dann ungerade, wenn* x_0 *ein Sattelpunkt von* f *ist, und zwar liegt dann sogar ein strikter Sattelpunkt vor.*

Der *Beweis* ist sehr einfach mit Hilfe des folgenden **Ableitungskriteriums für Vorzeichenwechsel** einer differenzierbaren reellen Funktion g an einer inneren Nullstelle x_0 ihres Definitonsintervalls:

- *Ist* $g(x_0) = 0$ *und* $g'(x_0) > 0$*, so hat* g *bei* x_0 *einen Vorzeichenwechsel vom Typ* $-/+$ *; dasselbe gilt, wenn* g' *in* x_0 *eine Nullstelle vom Typ* $+/+$ *hat.*
- Ist $g(x_0) = 0$ und $g'(x_0) < 0$, so hat g bei x_0 einen Vorzeichenwechsel vom Typ $+/-$; dasselbe gilt, wenn g' in x_0 eine Nullstelle vom Typ $-/-$ hat.

Das ist klar; denn aus $g'(x_0) > 0$ folgt für die Differenzenquotienten $\frac{g(x)-g(x_0)}{x-x_0} > 0$ für alle $x \neq x_0$ hinreichend nahe bei x_0, und damit $g(x) < g(x_0) = 0$ für $x < x_0$ und $g(x) > g(x_0) = 0$ für $x > x_0$ (und x nahe bei x_0). Ist $g'(x_0) = 0$, so kann man natürlich nicht auf einen Vorzeichenwechsel von g bei x_0 schließen; es könnte ja auch eine lokale Extremstelle vorliegen. Ist aber g' positiv beiderseits der Nullstelle x_0, so ist g streng monoton auf einem Intervall um x_0 (nach dem Sattelpunktkriterium (iii) vom ersten Satz dieses Abschnitts), und zwar natürlich streng wachsend (sonst wäre $g'(x) \leq 0$ für x nahe x_0); also hat auch in diesem Fall g bei x_0 einen Vorzeichenwechsel $-/+$.

Ist nun z.B. $f^{(k)}(x_0) > 0$ wie in Aussage (ii) des Satzes, so hat $f^{(k-1)}$ bei x_0 einen Vorzeichenwechsel $-/+$, also besitzt $f^{(k-2)}$ ein striktes lokales Minimum in x_0 (nach dem hinreichenden Ableitungskriterium erster Ordnung). Ist auch $f^{(k-2)}(x_0) = 0$, so ist x_0 eine Nullstelle vom Typ $+/+$ für $f^{(k-2)}$, also hat $f^{(k-3)}$ einen strikten Sattelpunkt in x_0 und, wenn auch $f^{(k-3)}(x_0) = 0$ ist, einen Vorzeichenwechsel vom Typ $-/+$. So fortfahrend gelangt man für gerade k zu dem Ergebnis, dass f in x_0 eine strikte lokale Minimumstelle hat, und für ungerade k zu dem Resultat, dass in x_0 ein strikter Sattelpunkt von f vorliegt (wobei f auf einem Intervall um x_0 streng wächst). Die Aussage (ii) ist damit bewiesen und auch, dass x_0 Sattelpunkt sein muss, wenn k ungerade ist. Die restlichen Aussagen ergeben sich daraus rein logisch, weil z.B. strikte lokale Maximumstellen und Sattelpunkte keine lokalen Minimumstellen sein können.

DISKUSSION: **(1)** Die Kriterien höherer Ordnung werden meist mit Ableitungsordnung $k = 2$ angewendet. Zur Bestimmung des Typs der kritischen Punkte wird dazu nach der Nullstellenbestimmung bei f' sofort die zweite Ableitung f'' berechnet, um damit strikte lokale Extremstellen zu identifizieren, oder gar noch f''', um Sattelpunkte zu bestimmen. Der Nutzen dieser schematischen Vorgehensweise ist aber gering:

- *Der Aufwand für die Berechnung von f'', f''' oder gar noch höheren Ableitungen zur Bestimmung des Typs der kritischen Punkte einer Funktion f lohnt sich oft nicht; die Anwendung des hinreichenden Kriteriums erster Ordnung, das auf der Bestimmung der Vorzeichen von f' nahe den kritischen Punkten beruht, oder eine direkte Inspektion der Funktion geben die gewünschte Information meistens einfacher.*

Ganz besonders gilt das, wenn – wie meistens – nur *absolute* Extremstellen gesucht sind. Mit den Kriterien höherer Ordnung kann man dann die Kandidatenliste im Allgemeinen nur wenig verkleinern, indem man z.B. die Sattelpunkte streicht und, wenn absolute Minimumstellen (bzw. Maximumstellen) gefragt sind, die strikten lokalen Maximumstellen (bzw. die strikten lokalen Minimumstellen). Da ist es meistens weniger aufwendig, durch Vergleich der Funktionswerte an allen Stellen der ursprünglichen Kandidatenliste direkt die Extremstellen herauszufinden (wenn deren Existenz sicher ist). Wir empfehlen daher nicht, bei der Extremstellenbestimmung immer gleich schematisch die Werte der zweiten Ableitung in den kritischen Punkten zu berechnen.

(2) Ist $f'(x_0) = 0$ und auch $f''(x_0) = 0$, so kann man über den Typ des kritischen Punktes x_0 keine Aussage machen: Es kann eine lokale Maximumstelle vorliegen oder eine lokale Minimumstelle oder ein Sattelpunkt.

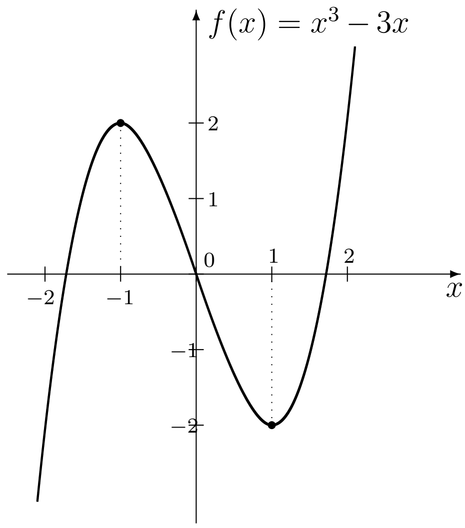
- *Auch in strikten lokalen Extremstellen x_0 kann durchaus $f''(x_0) = 0$ sein.*

Einfachstes Beispiel ist $f(x) = x^4$ mit isolierter Minimumstelle $x_0 = 0$, wo auch $f''(x) = 12x^2$ verschwindet (während bei $f(x) = x^3$, wo ebenfalls $f'(0) = f''(0) = 0$ ist, ein Sattelpunkt vorliegt). Insofern besteht zwischen den notwendigen und den hinreichenden Kriterien zweiter Ordnung eine Lücke: Die notwendigen Kriterien sind nicht hinreichend und die hinreichend nicht notwendig für isolierte lokale Extremstellen. Es gibt übrigens beliebig oft differenzierbare Funktionen f, die an einer Stelle x_0 eine strikte Extremstelle haben oder auch einen Sattelpunkt und dabei $f^{(k)}(x_0) = 0$ für alle Ableitungsordnungen $k \in \mathbb{N}$. Solche Funktionen können aber nicht elementar sein und kommen deshalb in der Mathematik für die Wirtschaftswissenschaft nicht vor. ■

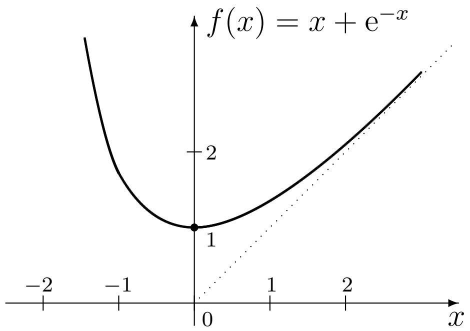
Wir geben noch einige abschließende Beispiele zur Bestimmung des Typs von kritischen Punkten (lokale Minimumstelle, lokale Maximumstelle, Sattelpunkt) und der Schlussfolgerungen, die man daraus bzgl. der absoluten Extremstellen ziehen kann. Man sollte sich dabei vor Augen halten, dass man die abgebildeten Graphen der Funktionen zunächst nicht genau kennt und die Extremstellendiskussion vielmehr benutzt, um Informationen über den Funktionsverlauf zu gewinnen.

BEISPIELE (**zur Bestimmung des Typs von kritischen Punkten**):

(1) $f(x) = x^3 - 3x$ hat auf $\mathbb{R}$ die Ableitung $f'(x) = 3x^2 - 3$ und die einzigen kritischen Punkte -1 und 1. Da f von unten und von oben unbeschränkt ist auf $\mathbb{R}$, gibt es keine absoluten Extremstellen. Aus $f''(x) = 6x$ entimmt man $f''(-1) < 0$ und $f''(1) > 0$, also liegt in -1 ein striktes lokales Maximum vor und in 1 ein striktes lokales Minimum. Das kann man auch ohne Berechnung der zweiten Ableitung daraus sehen, dass $f'(x) = 3(x-1)(x+1)$ bei -1 einen Vorzeichenwechsel vom Typ $+/-$ und bei 1 einen Vorzeichenwechsel vom Typ $-/+$ hat. Es ergibt sich sogar direkt aus den Werten $f(0) = 0$, $f(-1) = 2$, $f(1) = -2$ und den Grenzwerten $\lim_{x\to\infty} f(x) = \infty$, $\lim_{x\to-\infty} f(x) = -\infty$. Nach dem Zusatz zu den hinreichenden Kriterien erster Ordnung ist -1 eindeutige absolute Maximumstelle für f auf dem größten Intervall $]-\infty, 1]$ bis zum nächsten kritischen Punkt (in Wahrheit sogar noch auf dem größeren Intervall $]-\infty, 2[$), und 1 ist eindeutige absolute Minimumstelle zu f auf dem Intervall $[-1, \infty[$ (und sogar auf $]-2, \infty[$).



(2) $f(x) = x + \mathrm{e}^{-x}$ auf $\mathbb{R}$ hat die Ableitung $f'(x) = 1 - \mathrm{e}^{-x}$ und einen einzigen kritischen Punkt $x_0 = -\ln 1 = 0$. Wegen $f''(0) = \mathrm{e}^{-x}\big|_{x=0} = 1 > 0$ liegt dort ein striktes lokales Minimum. Nach dem Zusatz (v) zu den Kriterien erster Ordnung ist $x = 0$ eindeutige Minimumstelle auf dem größten Intervall um 0 ohne weitere kritische Punkte, also sogar eindeutige Minimumstelle von f auf $\mathbb{R}$. Das sieht man auch direkt an den Grenzwerten $\lim_{x\to\infty} f(x) = \infty$ und $\lim_{x\to-\infty} f(x) = \infty$ (genauer ist $f(x)$ asymptotisch gleich mit x bei $x \to \infty$ und mit e^{-x} bei $x \to -\infty$)



(3) $f(x) = x^n$ mit natürlichem Exponenten $n \in \mathbb{N}_{\geq 2}$ hat im einzigen kritischen Punkt $x = 0$ Ableitungswerte $f'(0) = f''(0) = \ldots = f^{(n-1)}(0) = 0$ und $f^{(n)}(0) = n! > 0$. Ist n ungerade, so liegt in $x = 0$ folglich ein Sattelpunkt. Ist n gerade, so liegt in $x = 0$ eine strikte lokale Minimumstelle, die sogar eindeutige absolute Minimumstelle von f auf ganz $\mathbb{R}$ ist, da es außer $x = 0$ keine weiteren kritischen Punkte gibt. (All dies kann man der Funktion natürlich direkt ansehen. Es geht hier aber darum, die Anwendung der Ableitungskriterien höherer Ordnung zu demonstrieren.)

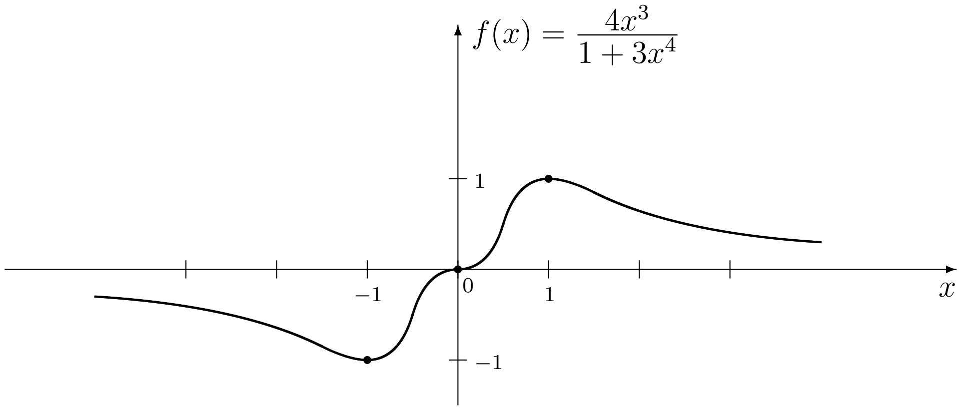
(4) Die rationale Funktion

$$f(x) = \frac{4x^3}{1+3x^4}$$

hat keine Polstellen auf $\mathbb{R}$ und ihre Ableitung ist

$$f'(x) = 4\frac{3x^2(1+3x^4) - 12x^3 \cdot x^3}{(1+3x^4)^2} = 4\frac{3x^2 - 3x^6}{(1+3x^4)^2}\,.$$

Wegen $x^2 - x^6 = -x^2(x^4-1) = -x^2(x^2-1)(x^2+1) = -x^2(x-1)(x+1)(x^2+1)$ sind $0, -1, 1$ die kritischen Punkte in $\mathbb{R}$. Die Berechnung der zweiten Ableitung f'' ist hier bereits so kompliziert, dass wir lieber darauf verzichten und zusehen, ob wir den Typ der kritischen Punkte nicht schon mit den Kriterien erster Ordnung bzw. mit direkter Inspektion der Funktion ermitteln können. (Das ist typisch für die Diskussion rationaler Funktionen, weil wiederholtes Differenzieren mit der Quotientenregel den Zählergrad und den Nennergrad im Allgemeinen erhöhen.) Bei $x = -1$ hat f' einen Vorzeichenwechsel vom Typ $-/+$, bei $x = 1$ einen Vorzeichenwechsel vom Typ $+/-$ und nahe $x = 0$ ist f' auf beiden Seiten positiv. Also haben wir eine strikte lokale Minimumstelle bei $x = -1$, eine strikte lokale Maximumstelle bei $x = 1$ und einen strikten Sattelpunkt bei $x = 0$. Nach dem Zusatz zu den Kriterien erster Ordnung über den Bereich absoluter Extremalität ist $x = -1$ absolute Minimumstelle von f auf $]-\infty, 1]$ und $x = 1$ ist absolute Maximumstelle von f auf $[-1, \infty[$. Ein Blick auf die Grenzwerte $\lim_{x\to\pm\infty} f(x) = 0$ und auf die Wertemenge von f, die sowohl positive als auch negative Werte enthält, zeigt sogar, dass f auf $\mathbb{R}$ ein positives Maximum hat und ein negatives Minimum, und dafür kommen wegen $f(0) = 0$ nur die kritischen Punkte $x = 1$ und $x = -1$ in Frage. Somit ist $x = 1$ sogar eindeutige absolute Maximumstelle von f auf ganz $\mathbb{R}$ und -1 ist eindeutige absolute Minimumstelle.



■

Man sieht im letzten Beispiel: *Alle gewünschten Informationen haben wir ohne Berechnung der zweiten Ableitung erhalten.* Außerdem würde auch die Berechnung der zweiten Ableitung wegen $f''(0) = 0$ über den Typ des kritischen Punktes $x = 0$ immer noch keinen Aufschluss geben, sondern erst die (in diesem Beispiel wirklich unangenehme) Berechnung der dritten Ableitung $f'''(0) > 0$. Das demonstriert, was wir oben zu den Kriterien höherer Ordnung gesagt haben, nämlich dass sie weniger nützlich sind als oft geglaubt wird und dass man den Typ der kritischen Punkte in konkreten Fällen meist schon durch Vorzeichendiskussion bei der Ableitung oder durch direkte Inspektion der Funktion bestimmen kann.