

Mathematik für Wirtschaftswissenschaftler

Vorlesungsprogramm für den 05. 04. und 12. 04. 2007

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

Kapitel 4: Analysis der Funktionen einer Variablen

4.1 Der Funktionsbegriff, elementare Funktionen

Der mathematische Funktionsbegriff ist grundlegend für viele Anwendungen der Mathematik auf die Modellierung ökonomischer Vorgänge und für den Einsatz leistungsfähiger mathematischer Theorien, wie z.B. der Differentialrechnung, zur Lösung ökonomisch motivierter Probleme. Im Grunde ist das Konzept einer Funktion, oder etwas allgemeiner einer Abbildung zwischen zwei Mengen, ein undefinierter Grundbegriff, den man wie auch das Konzept einer Menge nicht mathematisch streng durch bereits eingeführte Begriffe definiert, sondern nur so genau beschreibt, dass man damit arbeiten und darüber sprechen kann, ohne Missverständnisse befürchten zu müssen. Daher ist die folgende "Definition" auch keine Definition im mathematischen Sinne, sondern eben nur eine hinreichend genaue Beschreibung des Funktionsbegriffs.

DEFINITION: Eine **reelle Funktion einer Veränderlichen** ist eine Abbildung f , die jeder Zahl x aus einer Teilmenge D von \mathbb{R} (dem **Definitionsbereich** oder der **Definitionsmenge**; nichtleer vorausgesetzt) genau eine reelle Zahl y zuordnet. Man schreibt dann $y = f(x)$ für die x zugeordnete Zahl und nennt diese den **Funktionswert** an der Stelle x . Die Menge aller Werte, welche die Funktion an den Stellen des Definitionsbereichs D annimmt, heißt die **Wertemenge** oder das **Bild** von f auf D . ■

DISKUSSION, TERMINOLOGIE und NOTATION zum Funktionsbegriff:

1) "Abbildung" ist wie "Menge" ein undefinierter Grundbegriff. Das Wesentliche an einer **Abbildung** f von einer Menge D ("*Definitionsmenge*"; nichtleer vorausgesetzt) in eine Menge Z ("*Zielmenge*" oder "*Wertevorrat*") ist, dass f eine Vorschrift ist, die *jedem* Element x aus D *genau einen* Wert $y = f(x)$ aus der Menge Z zuordnet. Man sagt dann auch dass $x \in D$ durch die Abbildungsvorschrift auf $y = f(x)$ *abgebildet* wird und nennt $y = f(x)$ das *Bild von x* unter der Abbildung f . Die Zuordnungsvorschrift kann durch eine Formel definiert sein, gewissermaßen ein Programm, das bei Eingabe von $x \in D$ die Ausgabe $f(x) \in Z$ ergibt, oder durch eine verbale Festlegung oder sonst irgendwie — es muss eben nur für jedes Element x der Definitionsmenge zweifelsfrei festgelegt sein, welches Element y aus der Zielmenge ihm zugeordnet ist! Dagegen wird weder gefordert, dass jedes Element der Zielmenge überhaupt irgendeinem Element der Definitionsmenge zugeordnet ist, der Wertevorrat Z kann also Elemente y enthalten, die keine Werte der Abbildung f sind, noch wird verlangt, dass jedes $y \in Z$ höchstens einem $x \in D$ zugeordnet ist, es kann also verschiedene Elemente in der Definitionsmenge geben, die auf dasselbe Element der Zielmenge abgebildet werden.

Statt “Abbildung” sagt man auch **Funktion**, wenn die Zielmenge ein Zahlbereich ist, die Werte der Abbildung also Zahlen sind. Von einer **reellen Funktion** (oder *reellwertigen Funktion*) spricht man, wenn $Z = \mathbb{R}$ die Zielmenge ist, wenn die Funktionswerte also reelle Zahlen sind. Und die Abbildung heißt **Funktion von einer Veränderlichen / Variablen** (oder genauer Funktion von einer *reellen Variablen*), wenn ihre Definitionsmenge ebenfalls aus reellen Zahlen besteht, wenn also D Teilmenge von \mathbb{R} ist, z.B. ein (nichtleeres und auch nicht zu einem Punkt degeneriertes) Intervall. Sind Definitionsbereich und Wertemenge der Funktion Teilmengen der Zahlengeraden \mathbb{R} , so haben wir die in der obigen Definition betrachtete Situation von einer *reellen Funktion einer Veränderlichen*. Solche Funktionen bilden den Gegenstand dieses Kapitels.

Eine andere Situation liegt vor, wenn die Elemente des Definitionsbereichs Paare ($n = 2$) oder n -gliedrige Folgen (x_1, \dots, x_n) von reellen Zahlen ($n \geq 2$) sind. Dann hat man eine *Funktion* $f(x) = f(x_1, \dots, x_n)$ von mehreren reellen Veränderlichen, genauer eine von n (reellen) Variablen. Mit der Analysis solcher Funktionen befassen wir uns im nächsten Kapitel. Sind dagegen die Funktionswerte m -gliedrige Folgen $f(x) = (f_1(x), \dots, f_m(x))$ von reellen Zahlen ($m \geq 2$), so spricht man von einer *vektorwertigen Funktion*, einer \mathbb{R}^m -wertigen, genauer gesagt. Dies ist im Prinzip nichts anderes als ein System von m reellen Funktionen f_1, \dots, f_m , den sog. *Komponentenfunktionen von f* , und man kann die Ergebnisse der Analysis von reellen Funktionen auf jede Komponentenfunktion f_i anwenden.

Ebenfalls außerhalb der Betrachtungen dieses Kapitels sind Situationen, in denen der Definitionsbereich eine “diskrete” Menge ist, wie z.B. die Menge der natürlichen Zahlen \mathbb{N} . Eine Abbildung, die jedem $k \in \mathbb{N}$ ein Element y_k der Zielmenge Z zuordnet, ist nichts anderes als eine unendliche Folge (y_1, y_2, \dots) in Z . Auf solche “diskrete” Abbildungen kann man die Analysis der Funktionen einer Veränderlichen nicht anwenden; denn dafür ist wesentlich, dass die Variable “kontinuierlich” im Definitionsbereich variieren kann, also z.B. alle reellen Zahlenwerte in einem Intervall (positiver Länge) annimmt. Zwar sind ökonomische Größen eigentlich immer “diskrete Variablen”, nämlich Vielfache kleinster Einheiten. Man denke etwa an Stückzahlen einer Produktion oder an Geldbeträge. Für die mathematische Modellierung ökonomischer Vorgänge in der Analysis ist es aber notwendig, mit kontinuierlichen Variablen zu arbeiten. Dass die mathematische Theorie mit idealisierten kontinuierlichen Variablen überhaupt Ergebnisse von praktischer ökonomischer Relevanz liefert, liegt daran, dass die diskreten Änderungen der realen ökonomischen Größen in vielen Situationen relativ klein sind (etwa 1 Stück bei großer Stückzahl oder 1 Eurocent bei einem großen Eurobetrag).

2) Im Zusammenhang mit einer gegebenen Funktion f gibt es eine statische und eine dynamische Vorstellung: Wenn man ein bestimmtes Element $x \in D$ fixiert, so nennt man dies eine **Stelle** bzw. einen **Punkt** im Definitionsbereich D oder ein **Argument** der Funktion; das Argument ist gewissermaßen die Eingabe, die man dem Funktionsprogramm geben muss, damit es den zugehörigen **Funktionswert** $f(x)$ berechnen kann. Die Bestimmung eines Funktionswertes $f(x)$ bezeichnet man daher auch als **Einsetzen des Arguments x in die Funktion** oder **Auswertung der Funktion an der Stelle x** . Stellt man sich dagegen vor, dass x alle Elemente des Definitionsbereichs durchläuft und $y = f(x)$ die zugeordneten Elemente des Wertebereichs, so bezeichnet man x als **unabhängige Variable** und y als **abhängige Variable** oder Veränderliche. Man denkt sich hier, dass man die Stellen x innerhalb von D willkürlich variiert und dass die durch die Funktionsvorschrift bestimmten Werte von y sich in Abhängigkeit von x mitverändern. Nehmen x, y dabei wie oben nur reelle Zahlen als Werte an, so spricht man von *reellen Variablen*.

Bei der mathematischen Modellierung der Abhängigkeiten zwischen verschiedenen ökonomischen Größen durch Funktionen von einer reellen Variablen muss man zunächst entscheiden, welche Größe man als unabhängige Variable und somit als willkürlich veränderbar (innerhalb eines gewissen Bereichs) ansehen will und welche als abhängige Variable. In der realen ökonomischen Situation sind meistens noch weitere Variablen präsent, die wir aber als fixierte Parameter betrachten müssen, wenn wir die Situation durch *eine* reelle Funktion von *einer* reellen Variablen modellieren wollen. Diese Bedingung, dass die außer den gewählten unabhängigen und abhängigen Variablen noch zusätzlich vorkommenden ökonomischen Größen (hypothetisch) auf konstanten Werten gehalten werden, heißt in der Wirtschaftswissenschaft **ceteris-paribus-Bedingung** oder kurz **c.p.-Bedingung**, was bedeutet “die übrigen (Variablen) werden auf gleichbleibenden Werten gehalten”.

Nun hängen in der Ökonomie fast alle Größen voneinander ab, und es ist durchaus nicht immer angemessen, eine der ökonomischen Variablen als “unabhängig” und eine andere als “abhängig” auszuzeichnen. Manchmal ist die Sache klar, etwa wenn ein Monopolist den Preis für sein Produkt beliebig festsetzen kann und sich die Nachfrage nach diesem Produkt als Funktion des Preises ergibt (hoher Preis – kleine Nachfrage, niedriger Preis – große Nachfrage). In einer Konkurrenzsituation (Polypole) aber, in der dasselbe Produkt von mehreren Produzenten hergestellt wird, stellt sich ein Marktpreis ein, der nicht (ohne Absprache in einem Kartell) willkürlich verändert werden kann. Trotzdem ist auch hier die Nachfrage als eine Funktion des Preises anzusehen, und wenn man diesen Zusammenhang untersuchen will, so muss man den Preis als unabhängige Variable denken, auch wenn er in der Realität gar nicht willkürlich variiert werden kann. Die Auszeichnung einer unabhängigen und einer abhängigen Variablen ist immer dann notwendig, wenn man die ökonomische Situation durch *eine* Funktion von *einer* Variablen beschreiben will, und sie ist Voraussetzung für die Anwendung der wirkungsvollen Ergebnisse und Techniken, welche die Analysis der Funktionen von einer Veränderlichen zur Verfügung stellt. Wenn eine solche Vorgehensweise sachlich gänzlich unangebracht ist, insbesondere wenn zwischen den betrachteten Größen gar kein funktionaler Zusammenhang in dem Sinne besteht, dass die Werte der einen die jeweiligen Werte der anderen eindeutig festlegen, so wird man die angemessene mathematische Modellierung der Situation nicht mit einer reellen Funktion einer Veränderlichen vornehmen können.

3) Zur Definition einer Funktion gehört grundsätzlich auch die Angabe des Definitionsbereichs. Aber die Zuordnungsvorschrift ist das Primäre, und daher nimmt man es oft mit der Angabe des Definitionsbereichs nicht so genau; er ergibt sich meist aus dem Kontext. Ist z.B. die Funktionsvorschrift durch einen Rechenterm gegeben, in dem die Variable x und sonst nur Konstanten (konkrete Zahlen oder als fixiert gedachte Parameter) vorkommen, so ist als Definitionsmenge meistens der sog. **maximale Definitionsbereich** des Terms gemeint, d.h. die Menge aller reellen Zahlen x für die man die durch den Term beschriebene Rechenvorschrift ausführen kann. Im Extremfall, wie etwa bei $\sqrt{-1-x^2}$ oder $\ln(-x^2)$, ist das die leere Menge, so dass der Term gar keine Funktion definiert. (Die leere Menge lassen wir als Definitionsbereich nicht zu.) Im “Normalfall” wird aber der maximale Definitionsbereich ein Intervall sein, etwa ganz \mathbb{R} oder $\mathbb{R}_{\geq 0}$ oder $\mathbb{R}_{> 0}$, aus dem evtl. noch einige Stellen zu entfernen sind, wenn für diese Ausnahmewerte von x der Wert des Terms nicht definiert ist (etwa weil dafür ein auftretender Nenner Null wird). Die entfernten Stellen nennt man auch **Definitionslücken** der durch den Term definierten Funktion; z.B. hat $f(x) = 1/(1-x^2)$ auf \mathbb{R} die Definitionslücken -1 und $+1$, also den maximalen Definitionsbereich $\mathbb{R} \setminus \{-1, 1\}$ (lies: “ \mathbb{R} ohne -1 und 1 ”).

Natürlich kann man den Definitionsbereich D einer Funktion f immer zu einer kleineren (nichtleeren) Menge $E \subset D$ verkleinern; man betrachtet dann dieselbe Funktionsvorschrift, setzt aber nur noch Stellen aus E ein. Dadurch erhält man eine Funktion auf E , die man **Einschränkung** der zuvor auf D definierten Funktion auf die kleinere Menge E nennt. Zum Beispiel kann man die auf ganz \mathbb{R} definierte Quadrat-Funktion $f(x) = x^2$ auf den Bereich der positiven Zahlen $\mathbb{R}_{>0}$ einschränken; man betrachtet die Funktion eben nur “auf $\mathbb{R}_{>0}$ ”, wie man sagt, um anzuzeigen, dass als Definitionsbereich $\mathbb{R}_{>0}$ gemeint ist und nicht \mathbb{R} . Gerade für ökonomische Variablen sind oft nur positive oder nichtnegative Werte sinnvoll; also wird man für Funktionen von solchen Veränderlichen den Definitionsbereich $\mathbb{R}_{>0}$ oder $\mathbb{R}_{\geq 0}$ nehmen, auch wenn der maximale Definitionsbereich größer ist.

Allerdings muss man sich vor Augen halten, dass *wichtige Eigenschaften einer Funktion nicht nur von der Funktionsvorschrift, sondern auch vom Definitionsbereich abhängen!* Beispielsweise hat die Funktion $f(x) = x^2$ auf $\mathbb{R}_{>0}$ keine Nullstelle, auf \mathbb{R} aber schon; bei Verkleinerung des Definitionsbereichs können eben Nullstellen “verloren gehen”. Und auf $\mathbb{R}_{>0}$ ist $f(x) = x^2$ eine positive Funktion (d.h. sie hat nur positive Werte), auf \mathbb{R} aber nicht (weil sie den Wert 0 annimmt in $x = 0$). Auf $\mathbb{R}_{>0}$ ist die Funktion außerdem **injektiv** (oder: *eineindeutig, eins-zu-eins*), d.h. sie hat an verschiedenen Stellen in $\mathbb{R}_{>0}$ auch verschiedene Werte. Auf dem Definitionsbereich \mathbb{R} ist sie aber nicht injektiv (weil sie z.B. an den Stellen 1 und -1 denselben Wert hat). Positivität und Injektivität einer Funktion können also bei Vergrößerung des Definitionsbereichs verloren gehen. Wenn die diskutierten Eigenschaften einer Funktion nicht nur von der Funktionsvorschrift, sondern auch vom Definitionsbereich abhängen, ist die Angabe des genauen Definitionsbereichs natürlich unerlässlich!

Zur Definition einer Abbildung gehört neben der Abbildungsvorschrift und dem Definitionsbereich genau genommen auch noch die Angabe der Zielmenge Z . Zum Beispiel kann man die Quadrierfunktion $f(x) = x^2$ mit Definitionsbereich \mathbb{R} als eine Abbildung in die Menge \mathbb{R} der reellen Zahlen auffassen, aber auch als eine Abbildung in die kleinere Zielmenge $\mathbb{R}_{\geq 0}$, weil die Werte ja nichtnegative Zahlen sind. Streng genommen definieren verschiedene mögliche Wahlen der Zielmenge verschiedene Abbildungen. In der Analysis nimmt man es aber nicht so genau und vereinbart, dass *zwei Funktionen gleich sind, wenn sie denselben Definitionsbereich haben und für alle Elemente des Definitionsbereichs dieselben Werte*. Die Funktionsvorschrift kann dabei durchaus unterschiedlich aussehen; z.B. definieren $f(x) := |x|$ und $g(x) := \sqrt{x^2}$ dieselbe Funktion auf \mathbb{R} , also gilt $f = g$. Für reelle Funktionen wie hier ist dabei die Zahlengerade \mathbb{R} als Zielmenge zu denken, für vektorwertige Funktionen mit m Komponenten der m -dimensionale Zahlenraum \mathbb{R}^m .

Eine Eigenschaft von Abbildungen/Funktionen, die von der Zielmenge wesentlich abhängt, ist die Surjektivität. Eine Abbildung f mit Zielmenge Z heißt **surjektiv**, wenn jedes Element $y \in Z$ das Bild $y = f(x)$ von mindestens einem Punkt x aus dem Definitionsbereich ist. Man sagt dann auch, dass f eine **Abbildung auf Z** ist (statt “in Z ”). Diese Eigenschaft kann man immer erreichen, indem man die Zielmenge nötigenfalls zur Wertemenge der Abbildung verkleinert. Surjektivität geht andererseits verloren, sobald man die Zielmenge einer surjektiven Abbildung echt vergrößert (ohne Definitionsbereich und Abbildungsvorschrift zu ändern). Wenn man von der Surjektivität einer reellen Funktion f spricht, so muss man die ins Auge gefasste Zielmenge $Z \subset \mathbb{R}$ genau angeben. Zum Beispiel ist die Quadrat-Funktion $f(x) = x^2$ mit Definitionsbereich \mathbb{R} nicht surjektiv, wenn man \mathbb{R} als Zielmenge meint (weil -1 kein Wert ist), wohl aber surjektiv, wenn $\mathbb{R}_{\geq 0}$ als Zielmenge gewählt wird. Es handelt sich also um eine Abbildung “in \mathbb{R} ” und “auf $\mathbb{R}_{\geq 0}$ ”.

4) *Notation:* Um anzugeben, dass f eine Abbildung von der Menge D in die Menge Z ist, schreibt man allgemein $f: D \rightarrow Z$ (lies: “ f von D nach Z ”). Sind für Definitionsbereich und Zielmenge der Abbildung f keine Buchstabensymbole vergeben, so ist $\text{Def}(f)$ eine übliche Bezeichnung für ihren Definitionsbereich und $\text{Ziel}(f)$ für ihre Zielmenge. Die Wertemenge von f bezeichnen wir suggestiv mit $f(D) = \{f(x) : x \in D\}$; eine andere übliche Notation für diese auch als Bild der Abbildung bezeichnete Teilmenge von Z ist $\text{Bild}(f)$. Dass f eine reelle Funktion von einer reellen Veränderlichen ist, drückt man also aus mit den Schreibweisen

$$f: \mathbb{R} \supset D \rightarrow \mathbb{R} \quad \text{oder} \quad \text{Def}(f) \subset \mathbb{R}, \text{Bild}(f) \subset \mathbb{R}$$

wobei die Angabe der “Grundmenge” \mathbb{R} in der ersten Notation auch unterbleiben kann, wenn klar ist, dass der Definitionsbereich D auf der Zahlengeraden liegt. Weiter zeigt

$$f: x \mapsto y \quad \text{oder} \quad x \xrightarrow{f} y \quad \text{oder} \quad y = f(x),$$

an, dass $y \in \mathbb{R}$ der Funktionswert ist, der dem Element x aus dem Definitionsbereich durch die Funktion zugeordnet wird (lies: “ x geht über in y unter f ” oder “ y ist Bild / Wert von x unter f ”).

Die Funktionsvorschrift selbst wird meistens durch einen Rechterm $T(x)$ in der unabhängigen Variablen angegeben. Man schreibt also $f(x) := x^2$ für die Vorschrift “quadrieren” oder $f(x) := 1/(1-x^2)$ für die Vorschrift “die Differenz von 1 und dem Quadrat nehmen und dann den Kehrwert bilden”. (Der Doppelpunkt vor dem Gleichheitszeichen zeigt an, dass es sich hier um eine Definition der Werte $f(x)$ handelt, nicht etwa um eine Gleichung zwischen zwei definierten Termen, die je nach Wahl von x richtig oder falsch sein kann; er wird aber oft weggelassen.) Es ist auch üblich, die gemeinte Funktion überhaupt nicht mit einem Buchstabensymbol zu benennen, wenn das entbehrlich ist, sondern die Funktionsvorschrift allein durch den Rechterm anzugeben; man spricht also einfach von der Funktion x^2 bzw. $1/(1-x^2)$ oder genauer, wenn der Definitionsbereich wichtig ist, z.B. von der Funktion x^2 auf $\mathbb{R}_{>0}$ bzw. von $1/(1-x^2)$ auf $]1, \infty[$.

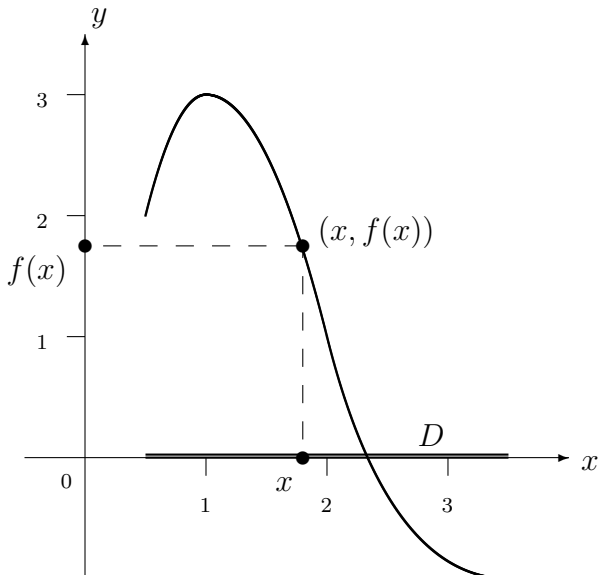
Natürlich kann für die unabhängige Variable auch ein anderes Buchstabensymbol stehen als “ x ” und für die Funktion ein anderes als “ f ”. In den Wirtschaftswissenschaften sind für viele ökonomischen Variablen ganz bestimmte Buchstabensymbole üblich, etwa “ p ” für einen Preis, “ K ” für Kosten, “ E ” für Erlös (Umsatz), “ G ” für Gewinn, “ x ” oder “ q ” für die produzierte, am Markt abgesetzte oder nachgefragte Menge eines Produkts. Dann hat man es also mit Funktionen wie $p(x)$, $K(x)$, $E(x)$ oder $G(x)$ zu tun, aber auch mit $E(p)$, $G(p)$ und sogar $x(p)$, wo nun “ p ” die unabhängige Variable ist und “ x ” die abhängige.

5) Die **graphische Darstellung** einer reellen Funktion f von einer Veränderlichen erfolgt durch Zeichnen des sog. **Funktionsgraphen**, das ist die Menge alle Zahlenpaare $(x, f(x))$ mit x aus dem Definitionsbereich $D \subset \mathbb{R}$,

$$\text{Graph}(f) := \{(x, y) : x \in D, y = f(x)\}.$$

Da hier die unabhängige Variable x und die abhängige Variable $y = f(x)$ reelle Werte annehmen, ist dies eine Teilmenge der Ebene \mathbb{R}^2 . Ein Vorteil unserer Beschränkung auf reelle Funktionen von einer Veränderlichen in diesem Kapitel ist die Möglichkeit der Veranschaulichung von qualitativen Funktionsverläufen durch Zeichnen entsprechender “Graphenkurven” in der “Zeichenebene”. Der Graph einer reellen Funktion von zwei Veränderlichen ist dagegen eine “Fläche” im Raum \mathbb{R}^3 und weniger einfach zu zeichnen bzw. zu veranschaulichen. Bei Funktionen von mehr als zwei Veränderlichen gibt es überhaupt keine adäquate graphische Darstellung.

Beim Zeichnen einer Graphenkurve wird im Prinzip für jede Stelle (in der Praxis natürlich nur für einige wenige ausgewählte Stellen) $x \in D$ der Punkt mit Rechtswert x und mit Hochwert $y = f(x)$ in der Koordinatenebene aufgetragen. Das Charakteristische an einem Funktionsgraphen in \mathbb{R}^2 ist, dass er von jeder vertikalen Geraden (parallel zur y -Achse) in höchstens einem Punkt geschnitten wird, nämlich von den vertikalen Geraden, die durch ihren Definitionsbereich auf der x -Achse verlaufen, in genau einem Punkt, von den andern vertikalen Geraden überhaupt nicht. Dies bedeutet eben gerade, dass jedem Punkt aus D genau ein Wert durch die Funktion zugeordnet ist, und den Punkten aus $\mathbb{R} \setminus D$ keiner.



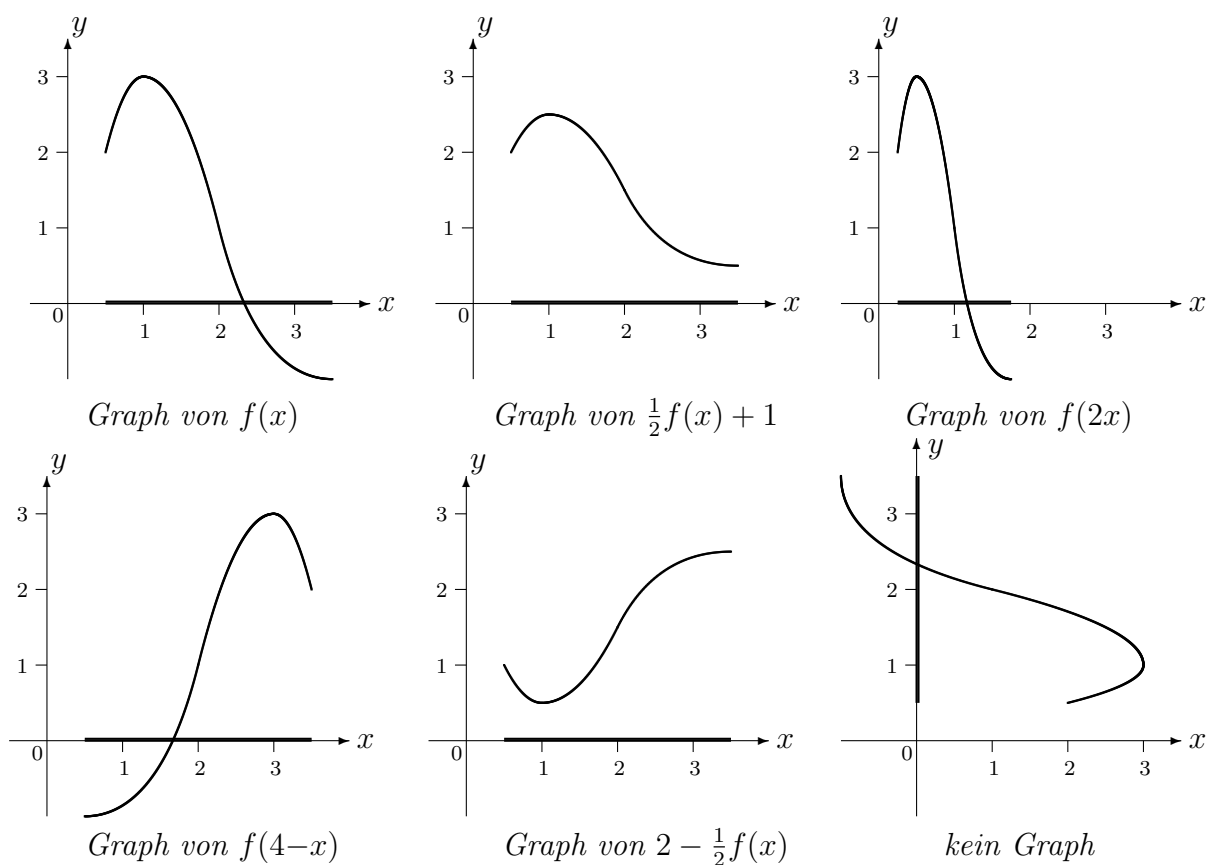
Graph der Funktion f auf $D = [0.5, 3.5]$

Wenn man den Graphen einer Funktion zeichnet, so hängt die geometrische Form der Graphenkurve von den Nullpunkten und Längeneinheiten ab, die man auf den Achsen wählt. Eine Veränderung der Einheiten bewirkt eine horizontale und/oder vertikale Streckung oder Stauchung des Graphen. Bei nichtlinearen Verzerrungen, etwa dem Auftragen von $\ln y = \ln f(x)$ über $\ln x$ (wie beim Auftragen in Logarithmuspapier), kann sich die Form sogar noch deutlicher verändern; z.B. wird so aus einer gekrümmten Graphenkurve unter Umständen eine gerade Strecke. Die geometrisch verschiedenen Kurven, die man so erhält, stellen aber alle dieselbe Funktion f dar.

6) Die Verschiebungen bzw. Streckungen oder -stauchungen, denen Funktionsgraphen bei Verschiebung bzw. Maßstabsänderung der Skalen auf den Achsen unterliegen, entsprechen einfachen algebraischen Operationen mit Funktionen. Hierzu geben wir folgende Tabelle an, in der die geometrischen Operationen mit den Graphen den entsprechenden algebraischen Operationen mit den Funktionen gegenübergestellt sind:

Graph(g) entsteht aus Graph(f) durch	wenn $g(x) =$
horizontale Verschiebung um c	$f(x - c)$
vertikale Verschiebung um d	$f(x) + d$
Streckung / Stauchung der Hochwerte mit Faktor $s > 0$	$sf(x)$
Streckung / Stauchung der Rechtswerte mit Faktor $r > 0$	$f(\frac{1}{r} x)$
Spiegelung an der horizontalen Achse	$-f(x)$
Spiegelung an der vertikalen Achse	$f(-x)$
Punktspiegelung am Ursprung	$-f(-x)$
Spiegelung an der Diagonalen $\{(x, y) : x = y \in \mathbb{R}\}$ also Komponentenvertauschung $(x, y) \mapsto (y, x)$	$f^{-1}(x)$

Dabei ist die horizontale *Verschiebung* um c nach rechts auszuführen, wenn $c > 0$ ist, und um $|c|$ nach links, wenn $c < 0$ ist. Analog erfolgt die vertikale Verschiebung um d nach oben, wenn $d > 0$ ist, und um $|d|$ nach unten, wenn $d < 0$. Von einer *Streckung* der Hoch- bzw. Rechtswerte sprechen wir, wenn sie alle mit demselben Faktor $s > 1$ multipliziert (z.B. alle verdoppelt) werden; bei Multiplikation mit einem Faktor $0 < s < 1$ (z.B. bei Halbierung der Werte) spricht man dagegen von einer *Stauchung*. Natürlich kann man auch mehrere der genannten Operationen hintereinander ausführen; z.B. entsteht der Graph von $-f(-x)$ aus dem von $f(x)$ auch durch Spiegelung an der horizontalen und an der vertikalen Achse nacheinander, und der Graph von $g(x) = 2 - \frac{1}{2}f(x)$ entsteht aus dem von f durch vertikale Stauchung mit dem Faktor $\frac{1}{2}$, anschließende Spiegelung an der horizontalen Achse und schließlich Verschiebung um 2 nach oben. Hier einige Abbildungen zur Veranschaulichung des Sachverhalts:



7) Die letzte Abbildung illustriert das Problem, das bei Spiegelung an der Diagonalen zu den beiden Koordinatenachsen auftaucht: Bei dieser Spiegelung braucht das Spiegelbild eines Graphen kein Funktionsgraph zu sein, d.h. das Spiegelbild kann vertikale Geraden in mehr als einem Punkt treffen. Die Spiegelung an der Diagonalen ist nämlich einfach die Komponentenvertauschung $(x, y) \mapsto (y, x)$ in der Ebene, und dabei gehen die horizontalen in die vertikalen Geraden über. Ein Graph kann aber mit einer horizontalen Geraden durchaus mehrere Punkte, sogar ganze Intervalle, gemeinsam haben, d.h. die Funktion kann denselben Wert an verschiedenen Stellen annehmen, wie z.B. f in der Abbildung alle Werte $y \in [2, 3[$. Das an der Diagonalen gespiegelte Bild eines Graphen ist infolgedessen dann und nur dann wieder ein Funktionsgraph, wenn der ursprüngliche Graph jede horizontale Gerade in höchstens einem Punkt schneidet, wenn also die Funktion jeden Wert an höchstens einer Stelle annimmt. Mit anderen Worten: Genau wenn die ursprüngliche Funktion injektiv ist, also an verschiedenen Stellen verschiedene Werte hat, ist das Spiegelbild ihres Graphen bzgl. der Diagonalen wieder Graph einer Funktion.

Die Funktion, deren Graph dann das Spiegelbild von $\text{Graph}(f)$ bzgl. der Diagonalen ist, heißt die **Umkehrfunktion** zu f und wird f^{-1} notiert. *Warnung:* Der Exponent -1 in dieser Notation darf nicht mit Bildung der reziproken Funktion $\frac{1}{f}$ verwechselt werden; manche Autoren schreiben daher f^{-1} für die Umkehrfunktion. Die Umkehrfunktion zu $f(x) := x^2$ auf dem Definitionsbereich $\mathbb{R}_{>0}$ ist z.B. die Wurzelfunktion $0 < y \mapsto \sqrt{y}$; die reziproke Funktion zu f ist dagegen $0 < x \mapsto \frac{1}{x^2}$. Wenn nicht ausdrücklich anderes gesagt wird, so meinen wir mit f^{-1} stets die Umkehrfunktion zu f und nicht die reziproke Funktion.

Die Umkehrfunktion zu f ist genau dann definiert, wenn f injektiv ist; eine injektive Abbildung nennt man daher auch **umkehrbare Funktion**. Die Umkehrbarkeit einer Funktion hängt nicht nur von der Funktionsvorschrift ab, sondern ganz wesentlich auch vom Definitionsbereich. Zum Beispiel ist $f(x) := x^2$ auf $\mathbb{R}_{>0}$ umkehrbar mit Umkehrfunktion $f^{-1}(y) = \sqrt{y}$, aber auf \mathbb{R} nicht (da z.B. $f(-1) = f(1)$ ist).

Ist f umkehrbar, so gibt es zu jedem Wert y von f genau eine Stelle x in $\text{Def}(f)$ mit Funktionswert $y = f(x)$, und die Umkehrfunktion ordnet dem Element y des Wertebereichs von f dann dieses eindeutige zu y gehörende "Urbild" x unter der Abbildung f zu:

$$f^{-1}(y) = x \iff x \in \text{Def}(f) \quad \text{und} \quad f(x) = y.$$

Konkret berechnet man einen Umkehrfunktionswert $f^{-1}(y)$, indem man die Gleichung $f(x) = y$ nach x auflöst; die Umkehrbarkeit von f garantiert die Existenz einer eindeutigen Lösung $x \in \text{Def}(f)$, und diese Lösung ist der Umkehrfunktionswert $f^{-1}(y)$. Der Definitionsbereich der Umkehrfunktion ist (wenn sie existiert) die Wertemenge $\text{Bild}(f)$ der Ausgangsfunktion, und die Zielmenge der Umkehrfunktion ist der Definitionsbereich $\text{Def}(f)$ der Ausgangsfunktion (oder die Grundmenge, in der dieser Definitionsbereich liegt, etwa \mathbb{R} im Falle einer Funktion f von einer reellen Veränderlichen). Der Graph der Umkehrfunktion einer injektiven Funktion f ist also

$$\text{Graph}(f^{-1}) = \{(y, x) : y \in \text{Bild}(f), x \in \text{Def}(f), y = f(x)\},$$

und man erkennt, dass er aus

$$\text{Graph}(f) = \{(x, y) : x \in \text{Def}(f), y \in \text{Bild}(f), y = f(x)\}$$

durch Komponentenvertauschung $(x, y) \mapsto (y, x)$ entsteht, also durch Spiegelung an der Diagonalen zu den beiden Koordinatenachsen, wie oben behauptet.

Die Wirkung einer injektiven Abbildung f wird durch die Umkehrabbildung gewissermaßen "umgekehrt", d.h. Anwendung von f und danach f^{-1} auf ein Argument $x \in \text{Def}(f)$ gibt das Argument x zurück, $f^{-1}(f(x)) = x$. In gleichem Sinne kehrt f auch die Wirkung von f^{-1} um, wenn f^{-1} existiert; denn für $y \in \text{Bild}(f)$ ist $f^{-1}(y)$ das eindeutige Element $x \in \text{Def}(f)$ mit $f(x) = y$, also gilt $f(f^{-1}(y)) = y$. Mit f ist daher auch f^{-1} umkehrbar, und f ist die Umkehrabbildung zu f^{-1} , in Formeln: $(f^{-1})^{-1} = f$. ■

Nun führen wir die Funktionen ein, mit denen die meisten in der Ökonomie relevanten funktionalen Abhängigkeiten von zwei Größen mathematisch modelliert werden können. Diese Funktionen sind aus einfachen Grundfunktionen aufgebaut mit Hilfe der Grundrechenarten und der Verkettung (Einsetzen einer Funktion in eine andere); sie sind also durch einen "Rechenterm" gegeben, der mit einem wissenschaftlichen Taschenrechner oder einem Computer-Rechenprogramm an jeder Stelle ausgewertet werden kann (soweit sich nicht so große Zahlen ergeben, dass sie im Rechner nicht mehr darstellbar sind).

DEFINITION: Die **elementaren Grundfunktionen** sind die identische Funktion x , die natürliche Exponentialfunktion e^x und die natürliche Logarithmusfunktion $\ln x$. Eine (ökonomische) **elementare Funktion** ist eine reelle Funktion von einer Veränderlichen, die durch einen “Rechenterm” (eine “Formel”) definiert ist, der aufgebaut ist aus Konstanten (Zahlen aus \mathbb{R}) und den elementaren Grundfunktionen mittels der vier Grundrechenarten $+$, $-$, \cdot , $:$ und der Verkettung. ■

DISKUSSION: 1) Wir haben in dieser Definition “Rechenterm” bzw. “Formel” in Anführungszeichen gesetzt, weil es keine im mathematischen Sinne streng definierten Konzepte sind. Es ist auch nicht nötig, genauer darauf einzugehen, was man eigentlich unter einer Formel zu verstehen hat (das exakt zu beschreiben, ist mathematisch recht anspruchsvoll, aber möglich). Vielmehr stellen wir uns unter einem “Rechenterm” einfach eine Rechenvorschrift vor, die man mit Hilfe eines wissenschaftlichen Taschenrechners oder eines entsprechenden Programms in einer mathematischen Programmiersprache ausführen oder sich als beliebig genau ausführbar denken kann, derart dass man für die Eingabe einer reellen Zahl x aus dem Definitionsbereich des Terms ein eindeutiges Rechenergebnis y als Ausgabe erhält. Man unterstellt dabei für die Theorie, dass der Rechner beliebig genau rechnen kann und Zwischenergebnisse in beliebiger Anzahl speichern kann. In der Praxis rechnet man natürlich mit begrenzter Genauigkeit, und die Anzahl der Rechenschritte, die zur Funktionsauswertung erforderlich sind, ist relativ klein.

- *Elementare Funktionen sind durch einen Rechenterm gegeben, der zu jeder Eingabe einer reellen Zahl aus seinem Definitionsbereich ein Ergebnis liefert, das im Prinzip mit einem Rechner berechnet werden kann, der über die Grundfunktionen e^x und $\ln x$ verfügt und unbegrenzte Genauigkeit und Speicherkapazität hat.*

2) Die drei Grundfunktionen x , e^x , $\ln x$ die wir in der obigen Definition zu Grunde gelegt haben, reichen aus, um die für die mathematische Modellierung ökonomischer Vorgänge benötigten Funktionen aufzubauen. In der Mathematik und Physik kommen als weitere Grundfunktionen noch Kreisfunktionen und ihre Umkehrfunktionen dazu, z.B. die periodische Sinusfunktion $\sin x$ und die Umkehrfunktion $\sin^{-1} = \arcsin$ zur Einschränkung des Sinus auf das Intervall $]-\frac{1}{2}\pi, \frac{1}{2}\pi[$, die auch auf jedem wissenschaftlichen Taschenrechner vorhanden sind. In der Ökonomie gibt es zwar Zyklen, wobei sich ökonomische Vorgänge in ähnlicher Form erneut abspielen, aber es gibt keine exakte Periodizität, bei der in gleichen Zeiträumen genau dieselben Entwicklungen stattfinden. Daher treten die exakt periodischen Kreisfunktionen und ihre Umkehrfunktionen in der Wirtschaftsmathematik nicht auf (mit Ausnahme einiger volkswirtschaftlicher Zyklenmodelle) und sind in obiger Definition nicht als Grundfunktionen aufgeführt. Die gegebene Definition erfasst infolgedessen nicht alle elementaren Funktionen der Mathematik, sondern nur spezielle elementare Funktionen, die man “ökonomische” elementare Funktionen nennen kann.

Die Grundfunktion $f(x) = x$ wird übrigens “identische Funktion” genannt, weil das Bild eines jeden Arguments x identisch mit diesem Element x ist. Allgemein nennt man die Abbildung mit Definitionsmenge D , die jedem Element $x \in D$ das Element x selbst zuordnet, die *identische Abbildung* oder *Identität* auf der Menge D .

3) Die Rechenoperationen bei der Berechnung eines Funktionswertes einer elementaren Funktion führt man immer mit reellen Zahlen aus, nämlich mit den entsprechenden Werten der Grundfunktionen, aus denen der elementare Funktionsterm aufgebaut ist. Dementsprechend sind für zwei reelle Funktionen f und g mit demselben Definitionsbereich

bereich D in \mathbb{R} die Summe $f + g$, die Differenz $f - g$, das Produkt $f \cdot g = fg$ und der Quotient f/g *wertweise* erklärt, d.h. man definiert die **Rechenoperationen mit Funktionen** durch

$$(f * g)(x) := f(x) * g(x) \quad (* \in \{+, -, \cdot, :\}),$$

wobei “ $*$ ” für eine der vier Grundrechenarten steht und der Definitionsbereich von $f \pm g$ und fg wieder D ist, der von $f : g = f/g$ aber natürlich nur die Menge der Stellen $x \in D$, an denen die Nennerfunktion g keine Nullstelle hat. Rechenoperationen mit elementaren Funktionen haben als Ergebnis stets wieder eine elementare Funktion.

4) Eine weitere wichtige Operation, die zum Aufbau der Rechenterme, die eine elementare Funktion definieren, verwendet werden kann, ist die **Verkettung** (auch **Komposition** oder **Hintereinanderausführung** genannt) $h \circ f$ von zwei Funktionen h und f . Hierbei handelt es sich um das *Einsetzen einer Funktion/eines Terms $f(x)$ für die unabhängige Variable in einer Funktion $h(y)$* , d.h. der Definitionsbereich der Verkettung $h \circ f$ besteht aus den Stellen x im Definitionsbereich von f , deren Wert $f(x)$ im Definitionsbereich von h liegt, und der Funktionswert von $h \circ f$ an diesen Stellen x ist dann definiert durch

$$(h \circ f)(x) := h(f(x)) \quad (x \in \text{Def}(f) \text{ mit } f(x) \in \text{Def}(h)).$$

Man kann sich vorstellen, dass die Ausgabe $f(x)$, welche die Funktionsvorschrift f zur Eingabe x ergibt, wieder als Eingabe für die weitere Funktionsvorschrift h verwendet wird, wodurch das Ergebnis $(h \circ f)(x) = h(f(x))$ entsteht. Dementsprechend heißt hier f die **innere Funktion** (oder die *vorangehende Funktion*) und h die **äußere Funktion** (oder die *nachfolgende Funktion*), und man liest $h \circ f$ als “ h nach f ” (auch “ h Kreis f ” wegen des für die Verkettung verwendeten Operationssymbols “ \circ ”). Für die Funktion $f(x) := x^2$ (Quadrieren) und $h(x) := e^x$ (Exponentieren) ergibt sich z.B.

$$(h \circ f)(x) = e^{x^2}, \quad (f \circ h)(x) = (e^x)^2 = e^{2x}, \quad (f \circ f)(x) = (x^2)^2 = x^4, \quad (h \circ h)(x) = e^{e^x}.$$

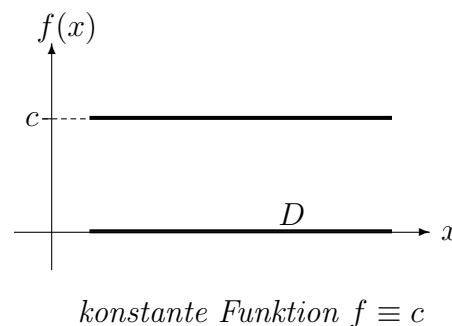
Die Verkettung von zwei elementaren Funktionen ist wieder eine elementare Funktion (sofern ihr Definitionsbereich nichtleer ist). Für die Verkettung einer umkehrbaren Funktion f mit ihrer Umkehrfunktion gilt (s.o.) $f^{-1}(f(x)) = x$ und $f(f^{-1}(y)) = y$ für alle $x \in \text{Def}(f)$ bzw. alle $y \in \text{Bild}(f)$, d.h. die Komposition $f^{-1} \circ f$ ist die Identität auf dem Definitionsbereich von f und $f \circ f^{-1}$ ist die Identität auf dem Definitionsbereich von f^{-1} .

5) Mit den algebraischen Operationen und mit der Verkettung kann man aus den elementaren Grundfunktionen alle Rechenterme aufbauen, die jemals in mathematischen Modellen der Ökonomie benötigt werden, sogar noch viel kompliziertere Rechenterme. Dabei ist der Aufbau der Terme, also die Art und Reihenfolge der auszuführenden Rechenoperationen, keineswegs durch die Funktion festgelegt, die der Term definiert. Mit andern Worten: Verschiedene Terme können durchaus dieselbe elementare Funktion definieren (und bei komplizierten Termen kann es eine schwierige Aufgabe sein zu entscheiden, ob das der Fall ist oder nicht). Zum Beispiel kann die Funktion e^{2x} , die oben als Ausführung $f \circ h$ der Quadrierfunktion nach der Exponentialfunktion aufgefasst wurde, auch als Produkt $(h \cdot h)(x) = e^x \cdot e^x$ oder als Verkettung $(h \circ g)(x) = \exp(2x)$ der Exponentialfunktion mit der inneren Funktion $g(x) := 2x$ angesehen werden. Insofern ist eine elementare Funktion nicht dasselbe wie ein Rechenterm, sondern Rechenterme enthalten mehr Information als nur die Ausgaben zu allen möglichen Eingaben, nämlich eine ganz bestimmte Abfolge von Rechenoperationen und Verkettungsoperationen, gewissermaßen ein Programm zur Funktionswertberechnung. Es kann auch sein, dass ein formaler Rechenterm gar keine Funktion definiert, nämlich wenn sein Definitionsbereich leer ist (was wir bei Funktionen nicht zugelassen haben). Schon genannte Beispiele sind $\ln(-x^2)$, $\sqrt{-1-x^2}$, $\sqrt{-e^x}$, ... ■

DISKUSSION (die einfachsten elementaren Funktionen: **Polynomfunktionen**):

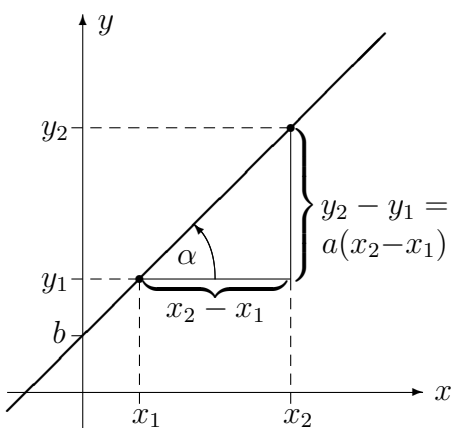
1) **Konstante Funktionen** $f(x) = c$ für alle $x \in D$ nehmen an allen Stellen ihres Definitionsbereichs denselben Wert an und sind natürlich nicht besonders interessant; denn sie beschreiben ja gar keine echte Abhängigkeit der Variablen $y = f(x)$ von x . Weder variiert y hier, noch hängt y von x ab (y hat ja immer denselben Wert c , ganz gleich wie x in D gewählt ist.) Dennoch wäre es nicht sinnvoll, konstante Funktionen aus den Betrachtungen ganz auszuschließen, und auch ökonomische abhängige Variablen können gelegentlich konstant sein, zumindest über gewissen Teilintervallen des Laufbereichs der unabhängigen Variablen. Man schreibt manchmal $f \equiv c$, um anzuzeigen, dass f eine konstante Funktion mit Wert c ist. (Statt diese Notation zu verwenden, kann man auch einfach sagen, dass “ $f = c$ konstant” ist.) Die konstante Funktion $f \equiv 0$ mit Wert Null heißt die **Nullfunktion**.

Der Graph der konstanten Funktion $f \equiv c$ auf \mathbb{R} ist die Gerade parallel zur horizontalen Koordinatenachse mit dem Achsenabschnitt c auf der vertikalen Achse; der Graph der Nullfunktion auf \mathbb{R} ist die horizontale Koordinatenachse. Ist der Definitionsbereich eine echte Teilmenge D von \mathbb{R} , so ist der Graph der konstanten Funktion $f \equiv c$ das über D gelegene Teilstück der horizontalen Geraden mit Achsenabschnitt c .



2) **Lineare Funktionen** haben die Form $\ell(x) = ax + b$; sie sind aus der identischen Funktion x und zwei Konstanten $a, b \in \mathbb{R}$ durch Rechenoperationen aufgebaut, also elementar. Im Fall $a = 0$ ist die Funktion konstant mit dem Wert b , also betrachten wir hier den Fall $a \neq 0$. Im Fall $b = 0$ heißt die Funktion auch **homogen linear** oder eine **Proportionalität**; sie beschreibt dann **proportionale Abhängigkeit** der abhängigen und unabhängigen Variablen (entspricht dem “Dreisatz”), d.h. $y = ax$ ist proportional zu x . Will man betonen, dass $b \neq 0$ sein darf, so nennt man $\ell(x) = ax + b$ auch **affin linear**.

Geometrisch sind lineare Funktion $\ell(x) = ax + b$ auf \mathbb{R} dadurch charakterisiert, dass *der Graph in der Zeichenebene eine Gerade ist*, die natürlich (als Graph) nicht vertikal sein kann. Dabei ist b der sog. **Achsenabschnitt** auf der vertikalen Achse, d.h. der Endpunkt der von 0 aus auf der vertikalen Achse abgetragenen Strecke, durch den die Gerade geht, und a ist die **Steigung** der Geraden, also der Quotient $(y_2 - y_1)/(x_2 - x_1)$ der Hochwertdifferenz und Rechtswertdifferenz von zwei beliebigen Punkten (x_1, y_1) und (x_2, y_2) auf der Geraden. (Das sieht man aus $y_2 - y_1 = (ax_2 + b) - (ax_1 + b) = a(x_2 - x_1)$. Die Steigung wird oft auch mit dem Buchstaben “ m ” bezeichnet.) Der im Bogenmaß gemessene Winkel $\alpha \in]-\frac{1}{2}\pi, \frac{1}{2}\pi[$ zwischen der horizontalen Achse und der Geraden, positiv gerechnet, wenn die Steigung positiv ist, und negativ, wenn die Steigung negativ ist (sowie $\alpha = 0$ bei Steigung $a = 0$), heißt auch *Steigungswinkel* der Geraden; sein Tangens ist dann die Steigung $a = \tan \alpha$. (Das kann man als Definition der Tangensfunktion auffassen, wenn man sich in Trigonometrie nicht auskennt.)



lineare Funktion $\ell(x) = ax + b$

Eine lineare Funktion $\ell(x) = ax + b$ ist festgelegt durch Angabe ihrer Steigung a und ihres Achsenabschnitts b . Statt b kann man auch einen beliebigen Punkt (x_1, y_1) auf dem Graphen angeben; denn aus $y_1 = ax_1 + b$ folgt $b = y_1 - ax_1$. Das gibt die sogenannte

$$\text{Punkt-Steigungs-Form} \quad y = a(x - x_1) + y_1$$

der Gleichung für die Gerade mit Steigung a durch (x_1, y_1) . Da man die Steigung mit Hilfe eines zweiten Punktes (x_2, y_2) auf der Geraden durch $a = (y_2 - y_1)/(x_2 - x_1)$ ausdrücken kann, erhält man mit leichter Umrechnung die

$$\text{Zwei-Punkte-Form} \quad y = \frac{x_2 - x}{x_2 - x_1} y_1 + \frac{x - x_1}{x_2 - x_1} y_2$$

der Gleichung für die Gerade durch (x_1, y_1) und (x_2, y_2) mit $x_1 \neq x_2$ bzw. für die eindeutige lineare Funktion $y = \ell(x)$, welche an den beiden gegebenen Stellen x_1 und x_2 die gegebenen Werte y_1 und y_2 hat. Ist die lineare Funktion ℓ homogen linear, ihr Graph also eine Gerade durch den Ursprung, so ist sie schon völlig festgelegt allein durch Angabe der Steigung a bzw. durch die Angabe eines einzigen Funktionswertes $y_1 = \ell(x_1)$ an einer Stelle $x_1 \neq 0$, nämlich $\ell(x) = ax = (y_1/x_1)x$.

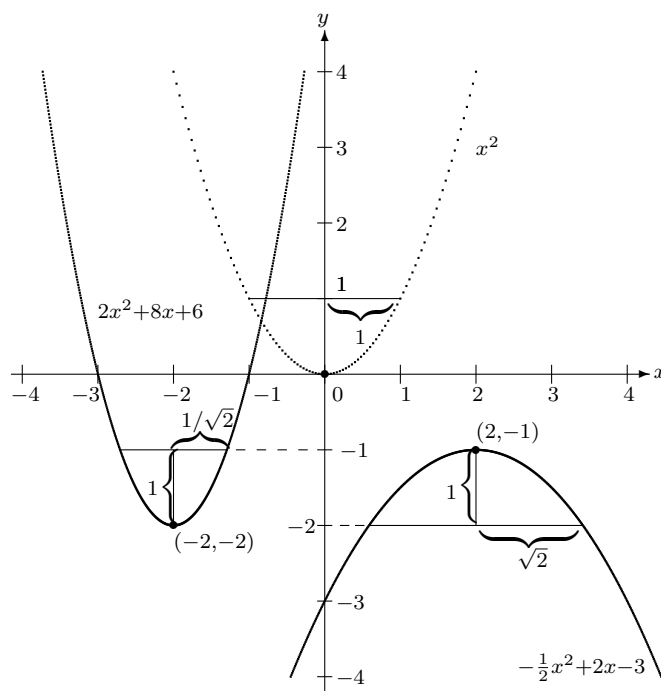
Die Umkehrfunktion einer nichtkonstanten linearen Funktion $\ell(x) = ax + b$ existiert auf \mathbb{R} und ist wieder linear. Durch Auflösung der linearen Gleichung $y = ax + b$, die genau im Fall $a \neq 0$ eindeutig möglich ist, erhält man den Umkehrfunktionswert $\ell^{-1}(y) = x = \frac{1}{a}y - \frac{b}{a}$; die Umkehrfunktion hat also die zur Steigung a von ℓ reziproke Steigung $\frac{1}{a}$ und die Nullstelle $-\frac{b}{a}$ von ℓ als Achsenabschnitt. Das ist auch geometrisch klar, weil ja der Graph von ℓ^{-1} durch Spiegelung der Geraden $\text{Graph}(\ell)$ an der Diagonalen zu den beiden Koordinatenachsen entsteht.

Zur Beschreibung von Abhängigkeiten zwischen ökonomischen Variablen x, y werden lineare Funktionen sehr oft verwendet, allein schon deswegen, weil der Zusammenhang zwischen x und y nicht genau bekannt und eine lineare Abhängigkeit $y = \ell(x)$ die einfachste Annahme ist. Besonders oft kommen homogene lineare Funktionen vor, weil sie Proportionalität von x und y beschreiben. Nimmt man einen linearen Zusammenhang $y = ax + b$ an, kennt aber die Koeffizienten a und b nicht, so darf man nicht einfach die Zwei-Punkte-Form der Geradengleichung für zwei "Datenerhebungen" (x_1, y_1) (x_2, y_2) nehmen, sondern man muss viele Daten (x_i, y_i) , $i = 1 \dots k$, erheben und dann die lineare Funktion bestimmen, die dazu "am besten passt". Dies ist eine Aufgabe der sog. *Regressionsanalyse* in der Statistik. Geometrisch handelt es sich darum, eine Gerade in der Zeichenebene zu bestimmen, welche sich den gegebenen Punkten (x_i, y_i) möglichst gut annähert. Die Antwort hängt davon ab, wie man die Abweichung einer Geraden $y = ax + b$ von den Punkten misst. Bei der *Methode der kleinsten Quadrate* nimmt man dazu die Summe der Quadrate $(y_i - ax_i - b)^2$ der vertikalen Abweichungen und bestimmt dann a und b so, dass diese Summe möglichst klein wird. Wir haben das schon in 1.5 im Zusammenhang mit dem quadratischen Mittel diskutiert und gezeigt, wie man die Lösung dieses Problems, die sog. *lineare Regressionsfunktion* oder *Regressionsgerade* berechnet. (Warnung: Die lineare Regressionsfunktion kann die Abhängigkeit zwischen x und y natürlich nicht gut darstellen, wenn diese in Wahrheit nichtlinearer Natur ist. Man sollte also bei der Berechnung linearer Regressionsfunktionen gute Gründe für die Annahme einer linearen Abhängigkeit zwischen x und y haben oder zumindest diese Annahme im Nachhinein durch entsprechende statistische Analyse des Datenmaterials rechtfertigen.)

3) Quadratische Funktionen $q(x) = ax^2 + bx + c$ (mit $a, b, c \in \mathbb{R}$, $a \neq 0$) sind gewissermassen die einfachsten nichtlinearen Funktionen und kommen deshalb auch in ökonomischen Modellen sehr oft vor, wenn kein linearer Zusammenhang zwischen zwei Variablen x und y vorliegt. Da die Funktion x^2 das Produkt der identischen Funktion mit sich selbst ist, sind quadratische Funktionen elementar. Man kann die Diskussion der quadratischen Funktionen auf die Untersuchung der Potenzfunktion x^2 , zurückführen durch Umformung mit **quadratischer Ergänzung**

$$q(x) = ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2}\right) - \frac{b^2}{4a} + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{d}{4a},$$

wo $d := b^2 - 4ac$ die sog. *Diskriminante* der quadratischen Funktion ist. Der Graph der Potenzfunktion x^2 ist die bekannte *Normalparabel*, und aus der quadratischen Ergänzung liest man nun ab, dass der Graph von q daraus im Fall $a > 0$ entsteht durch vertikale Streckung / Stauchung mit dem Faktor a (oder äquivalent, durch horizontale und vertikale Stauchung / Streckung mit dem Faktor $1/a$) und anschließende horizontale Verschiebung um $-b/2a$ und vertikale Verschiebung um $-d/4a$. Der Graph der quadratischen Funktion ist daher eine nach oben geöffnete Parabel mit *Scheitelpunkt* $(-\frac{b}{2a}, -\frac{d}{4a})$, und a bestimmt die Breite dieser Parabel, z.B. ist die Länge der horizontalen Sehne im Abstand 1 zum Scheitel gleich $2/\sqrt{a}$. Wenn $a < 0$



ist, so kommt zu der Streckung / Stauchung mit Faktor $|a|$ noch eine Spiegelung an der horizontalen Achse hinzu, der Graph ist dann eine nach unten geöffnete Parabel. Folglich hat die quadratische Funktion q **genau eine Extremstelle** $\frac{-b}{2a}$ und nimmt dort ihr Minimum auf \mathbb{R} an, wenn $a > 0$, bzw. ihr Maximum auf \mathbb{R} , wenn $a < 0$, jeweils mit Wert $\frac{-d}{4a}$.

Auf ganz \mathbb{R} ist eine quadratische Funktion $q(x) = ax^2 + bx + c$ nie umkehrbar, weil sie an Stellen, die zum Scheitelrechtswert $\frac{-b}{2a}$ symmetrisch liegen, denselben Funktionswert hat. Auf dem Intervall $[\frac{-b}{2a}, \infty[$ rechts vom Scheitelrechtswert ist q aber injektiv, und man erhält den Umkehrfunktionswert $q^{-1}(y)$ als eindeutige Lösung $\frac{-b}{2a} + \frac{1}{2|a|}\sqrt{d + 4ay}$ der quadratischen Gleichung $y = ax^2 + bx + c$ in diesem Intervall. Dabei liegt y im Definitionsbereich von q^{-1} , ist also eine Wert von q , genau wenn $d + 4ay \geq 0$ ist. Die ebenfalls existierende Umkehrfunktion zu q auf $]-\infty, \frac{-b}{2a}]$ hat denselben Definitionsbereich und ist gegeben durch $y \mapsto \frac{-b}{2a} - \frac{1}{2|a|}\sqrt{d + 4ay}$.

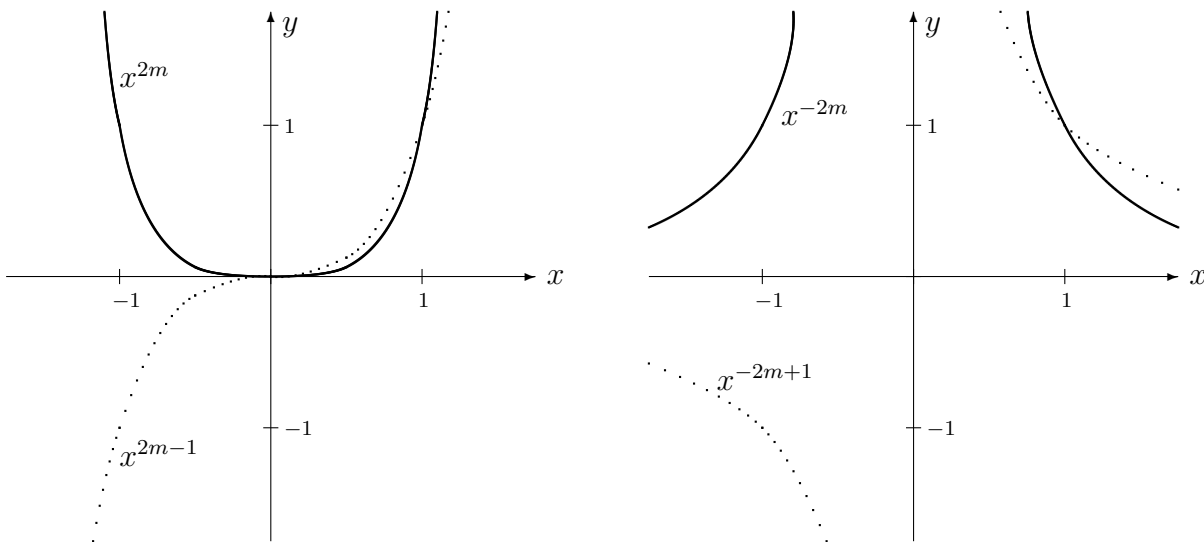
Eine quadratische Funktion $q(x) = ax^2 + bx + c$ ist vollständig bestimmt durch ihre Werte y_i an drei verschiedenen Stellen x_i , $i = 1 \dots 3$. Durch Einsetzen erhält man drei Gleichungen $ax_i^2 + bx_i + c = y_i$ für die unbekannt Koeffizienten a, b, c , die sich eindeutig lösen lassen. Das gibt für die gesuchte quadratische Funktion folgende **Drei-Punkte-Formel**

$$q(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} y_1 + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} y_3,$$

die auch *quadratische Lagrange-Interpolationsformel* heißt, weil sie die Funktionswerte quadratisch zwischen den drei Stellen x_i mit gegebenen Funktionswerten y_i interpoliert.

Allerdings ist diese Funktion nur dann quadratisch, wenn die drei Punkte (x_i, y_i) nicht auf einer Geraden liegen, sonst heben sich die quadratischen Terme heraus und q ist eine lineare Funktion. Geometrisch bedeutet die Existenz und Eindeutigkeit der quadratischen Interpolationsfunktion, dass es zu drei nicht auf einer Geraden liegenden Punkten in der Zeichenebene mit verschiedenen Rechtswerten genau eine Parabel mit vertikaler Symmetrieachse gibt, welche durch diese Punkte geht.

4) Potenzfunktionen mit ganzen Exponenten x^n auf \mathbb{R} (wenn $n \in \mathbb{Z}_{>0}$) bzw. auf $\mathbb{R}_{\neq 0}$ (wenn $n \in \mathbb{Z}_{<0}$) sind ebenfalls elementar, weil sie durch algebraische Rechenoperationen aus der identischen Funktion entstehen (Produkt von n Kopien der Funktion x bzw. Division der Konstanten 1 durch dieses Produkt). Für $n = 0$ hat man die konstante Funktion $\equiv 1$ mit Ausnahme der Definitionslücke bei $x = 0$, die man hier durch die Festsetzung $0^0 := 1$ beseitigen kann. Für $n = 1$ und $n = 2$ haben wir die Funktionen x und x^2 in 2) und 3) schon betrachtet. Für $n \in \mathbb{N}_{\geq 3}$ sieht der Graph von x^n über $[0, \infty]$ aus wie eine nahe 0 “abgeflachte” und rechts von 1 “aufgestellte” Parabel; über $]-\infty, 0]$ ist der Graph das Spiegelbild an der vertikalen Achse, wenn n gerade ist, bzw. das Spiegelbild am Ursprung, wenn n ungerade ist. Es besteht hier also ein Unterschied im Vorzeichen der Werte von x^n für $x < 0$ bei geraden Exponenten ($x^n > 0$) und bei ungeraden Exponenten ($x^n < 0$). Dieser Vorzeichenunterschied besteht auch bei den reziproken Funktionen x^{-n} . Für $n = 1$ ist der Graph von $x^{-1} = \frac{1}{x}$ über $\mathbb{R}_{\neq 0}$ die bekannte Normalhyperbel. Für größere $n \in \mathbb{N}_{\geq 2}$ verlaufen die Graphen über $\mathbb{R}_{>0}$ links von 1 steiler und rechts von 1 näher bei der horizontalen Achse. Die Graphen über $\mathbb{R}_{<0}$ ergeben sich dann wieder durch Spiegelung an der vertikalen Achse, wenn n gerade ist, bzw. am Ursprung, wenn n ungerade ist. Die Abbildungen illustrieren die Potenzfunktionen x^n für gerade $n = 2m$ bzw. ungerade $n = 2m - 1$ in $\mathbb{N}_{\geq 2}$ sowie die dazu reziproken Potenzfunktionen auf $\mathbb{R}_{\neq 0}$ für $n \in \mathbb{N}$.



5) Polynomfunktionen, auch **ganzrationale Funktionen** oder kurz **Polynome** genannt, sind Linearkombinationen der Potenzfunktionen x^k mit nichtnegativen ganzen Exponenten, also elementare Funktionen von der Form

$$p(x) = \sum_{k=0}^n a_k x^k = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

mit Zahlen $a_0, a_1, \dots, a_n \in \mathbb{R}$, genannt die **Koeffizienten** der Polynomfunktion. (In der Summe ist der Faktor x^0 als 1 zu interpretieren, auch wenn $x = 0$ ist.) Normalerweise

ist der *führende Koeffizient* $a_n \neq 0$ (sonst würde man diesen Summanden ja gar nicht aufschreiben), und dann heißt die Zahl $n \in \mathbb{N}_0$ der **Grad** des Polynoms. (Der Nullfunktion, die man erhält, wenn alle Koeffizienten Null sind, ordnet man keinen Grad zu.) Ist $a_n = 1$, so nennt man das Polynom *normiert*. Der Koeffizient a_0 heißt *absolutes Glied* oder *Achsenabschnitt*, weil es den Hochwert des Schnittpunkts $(0, p(0)) = (0, a_0)$ vom Graphen mit der vertikalen Achse angibt.

Die Polynomfunktionen vom Grad 0 sind also die konstanten Funktionen ($\neq 0$), die nicht-konstanten Polynome vom Grad 1 sind die linearen Funktionen, die nichtlinearen vom Grad 2 sind die quadratischen Funktionen, und die Potenzfunktion x^n ist ein normiertes Polynom vom Grad n für jedes $n \in \mathbb{N}_0$. Summen $p(x) + q(x)$ und Produkte $p(x) \cdot q(x)$ von Polynomfunktionen sind offenbar wieder welche, ebenso die Verkettung $q(p(x))$; denn man kann ja die Potenzen $p(x)^l$ durch Ausmultiplizieren wieder als Linearkombination von Potenzen x^k darstellen. Dabei ist der Grad von $p(x) \cdot q(x)$ die Summe und der von $q(p(x))$ das Produkt der Grade von p und q . Der Grad von $p(x) + q(x)$ ist höchstens so groß wie das Maximum der Grade von p und q , kann aber wegen Kürzungen kleiner sein.

Wenn man eine Nullstelle x_1 der Polynomfunktion $p(x)$ oben kennt, so kann man (wie bereits in 1.2 besprochen) stets den zugehörigen **Linearfaktor abspalten**, d.h.

$$p(x) = (x - x_1) \cdot \tilde{p}(x)$$

zerlegen, wobei \tilde{p} eine eindeutig bestimmte Polynomfunktion mit einem um 1 kleineren Grad als p ist. Dazu bestimmt man die Koeffizienten $\tilde{a}_{n-1} = a_n, \dots, \tilde{a}_1, \tilde{a}_0$ von \tilde{p} in dieser Reihenfolge so, dass beim Ausmultiplizieren der rechten Seite vor jeder Potenz x^k der Koeffizient a_k der linken Seite entsteht. Wenn \tilde{p} wieder die Nullstelle x_1 hat, so kann man den Linearfaktor $x - x_1$ erneut abspalten usw. bis zu einer Zerlegung

$$p(x) = (x - x_1)^{n_1} \cdot \hat{p}(x) \quad (\hat{p}(x_1) \neq 0),$$

bei der das "Restpolynom" \hat{p} an der Stelle x_1 nicht mehr den Wert Null annimmt. Man nennt die natürliche Zahl n_1 , also die Anzahl der abspaltbaren Linearfaktoren $x - x_1$, die **Vielfachheit der Nullstelle** x_1 (auch *Ordnung* der Nullstelle). Wenn nun \hat{p} eine andere Nullstelle x_2 hat, so kann man das Verfahren fortsetzen und endet mit einer Darstellung

$$p(x) = (x - x_1)^{n_1} (x - x_2)^{n_2} \cdot \dots \cdot (x - x_\ell)^{n_\ell} \cdot q(x),$$

wobei q eine Polynomfunktion ohne Nullstelle auf \mathbb{R} ist (wie z.B. $q(x) = x^2 + 1$). Es folgt:

- Eine Polynomfunktion vom Grad n hat höchstens n reelle Nullstellen, selbst wenn man jede mit ihrer Vielfachheit zählt.
- Der Wert $p(x)$ ändert das Vorzeichen, wenn x eine Nullstelle ungerader Ordnung überschreitet, nicht aber, wenn die Nullstelle gerade Ordnung hat.
- Zwei Polynomfunktionen vom Grad $\leq n$, die an $n + 1$ Stellen in \mathbb{R} denselben Wert haben, stimmen an allen Stellen in \mathbb{R} überein und haben dieselben Koeffizienten.

Die erste Aussage ist klar, da sich beim Abspalten eines Linearfaktors der Grad des Restpolynoms um 1 verkleinert. Die zweite liest man aus der obigen Faktorzerlegung ab, weil nur die Faktoren $x - x_i$ ihr Vorzeichen ändern, wenn x durch x_i läuft, $q(x)$ aber nicht (siehe 4.2). Die dritte Aussage ergibt sich, indem man die Differenz $p(x) - q(x)$ der beiden Polynomfunktionen betrachtet; dies ist ein Polynom $\sum_{k=0}^n c_k x^k$ vom Grad $\leq n$ oder die Nullfunktion, und da es mehr als n Nullstellen hat, kommt nur die zweite Alternative in Frage. Dann müssen p und q auch dieselben Koeffizienten haben; sonst wäre $p(x) - q(x) \neq 0$ für große x , da der führende Summand $c_k x^k \neq 0$ die anderen dominiert für große x .

Wenn man auch komplexe Nullstellen zuläßt, so kann man das Verfahren sogar fortsetzen bis zu einer *vollständigen Zerlegung in Linearfaktoren*, wobei oben dann $q(x)$ durch den führenden Koeffizienten $a_n \neq 0$ von p zu ersetzen ist. Z.B. kann $x^2 + 1$ mit der komplexen Nullstelle $\mathbf{i} = \sqrt{-1}$ zerlegt werden $x^2 + 1 = (x - \mathbf{i})(x + \mathbf{i})$. Außerdem kann bewiesen werden, dass jede Polynomfunktion vom Grad $n \geq 1$ eine Nullstelle im Bereich der komplexen Zahlen \mathbb{C} besitzt. Es gilt dann also der *Fundamentalsatz der Algebra*: Eine Polynomfunktion vom Grad n hat, bei Zählung mit Vielfachheiten, genau n Nullstellen im Bereich der komplexen Zahlen \mathbb{C} . Dies ist ein großer Vorteil dieses Zahlbereichs, der \mathbb{R} erweitert. In ökonomischem Zusammenhang interessieren wir uns aber nur für reelle Nullstellen.

Im Reellen ist der Fundamentalsatz der Algebra nicht richtig, es gibt eben Polynomfunktionen ohne reelle Nullstelle wie $x^2 + 1$ oder Potenzen davon. Durch Rechnen im Bereich \mathbb{C} der komplexen Zahlen kann man sogar zeigen, dass jedes reelle Polynom ohne Nullstelle in \mathbb{R} ein Produkt von quadratischen Funktionen $(x - \xi_j)^2 + \eta_j^2$ ohne Nullstelle ist ($\eta_j \neq 0$). Ist dagegen der Grad n von p ungerade, so hat der führende Summand $a_n x^n$ unterschiedliches Vorzeichen für positive und für negative Werte von x . Da außerdem für große $|x|$ der führende Term viel größeren Betrag hat als die Summe der anderen Summanden, nimmt folglich $p(x)$ positive und negative Werte an. Daher folgt aus dem Zwischenwertsatz (siehe 4.2): *Polynomfunktionen von ungeradem Grad haben mindestens eine reelle Nullstelle*. Es ist also kein Zufall, dass die obigen Beispiele von Polynomen ohne reelle Nullstelle von geradem Grad sind.

Will man sich einen Überblick über den Verlauf einer vorgelegten Polynomfunktion $p(x) = a_n x^n + \dots + a_1 x + a_0$ verschaffen, so schaut man sich zunächst den führenden Term $a_n x^n$ an; denn **der führende Summand bestimmt das Verhalten im Unendlichen**. Um das einzusehen, klammert man x^n aus:

$$p(x) = a_n x^n + x^n \left[a_{n-1} \frac{1}{x} + \dots + a_1 \frac{1}{x^{n-1}} + a_0 \frac{1}{x^n} \right].$$

Für große Werte von $|x|$ ist der Faktor [...] so klein, wie man will. Dies bedeutet, dass der relative Fehler, den man bei Ersetzung von $p(x)$ durch $a_n x^n$ begeht, also der Fehler $|p(x) - a_n x^n|$ dividiert durch den Betrag von $a_n x^n$, kleiner als jede gegebene Fehlerschranke ist, wenn man $|x|$ groß genug wählt. Man sagt für diesen Sachverhalt:

$$p(x) \text{ ist asymptotisch gleich mit } a_n x^n \text{ bei } x \rightarrow \pm\infty,$$

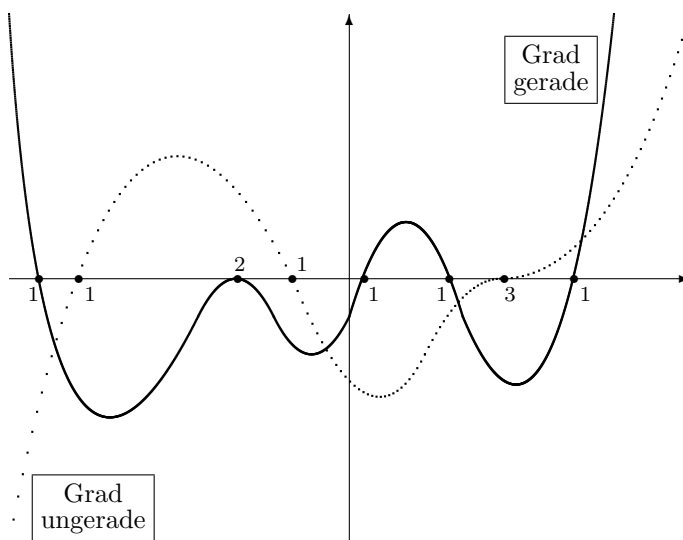
und notiert dafür

$$p(x) \sim a_n x^n \quad (|x| \rightarrow \pm\infty), \quad \lim_{x \rightarrow \pm\infty} \left| \frac{p(x) - a_n x^n}{x^n} \right| = 0.$$

Insbesondere folgt, dass $p(x)$ dasselbe Grenzverhalten im Unendlichen hat wie $a_n x^n$, d.h. für hinreichend große Werte von x und Grad $n \geq 1$ ist $p(x)$ größer als jede gegebene positive Schranke, wenn $a_n > 0$ ist, bzw. kleiner als jede gegebene negative Schranke, wenn $a_n < 0$ ist. Für negative x von hinreichend großem Betrag haben wir dann dieselben Aussage für $p(x)$, wenn n gerade ist, bzw. für $-p(x)$, wenn n ungerade ist. Wir halten für Polynomfunktionen vom Grad $n \geq 1$ dieses **Verhalten im Unendlichen** fest in folgender symbolischer Beschreibung:

$$\lim_{x \rightarrow \infty} p(x) = \left\{ \begin{array}{ll} \infty, & \text{wenn } a_n > 0 \\ -\infty, & \text{wenn } a_n < 0 \end{array} \right\} = (-1)^n \lim_{x \rightarrow -\infty} p(x) \quad (n \geq 1).$$

Das **Vorzeichen der Werte** einer Polynomfunktion $p(x)$ kann man so bestimmen: Für große x ist es das Vorzeichen des führenden Koeffizienten a_n . Bei Verkleinerung von x ändert es sich, wenn man eine Nullstelle ungerader Ordnung überschreitet, nicht aber bei Überquerung einer Nullstelle gerader Ordnung; das haben wir ja oben gesehen. So bestimmt man das Vorzeichen der Werte auf ganz \mathbb{R} . Die Abbildung zeigt die verschiedenen Möglichkeiten; die Zahlen geben die Vielfachheiten der jeweiligen Nullstellen an. Natürlich kann man bei der Bestimmung der Vorzeichenwerte auch an irgendeiner Stelle x mit $p(x) \neq 0$ beginnen und von dort aus nach links und rechts verfolgen, wie sich das Vorzeichen von $p(x)$ beim Überschreiten von Nullstellen ändert (oder nicht). ■



Polynomfunktionen sind sehr “flexibel”, man kann damit beliebige Abhängigkeiten zwischen zwei Variablen sehr genau darstellen, jedenfalls über einem kompakten Intervall als Definitionsbereich. Wir geben zwei Sätze an, welche diesen Punkt illustrieren:

Interpolationssatz: *Zu $n+1$ Stellen in \mathbb{R} und $n+1$ reellen Werten, gibt es stets genau eine Polynomfunktion vom Grad höchstens n , die an den gegebenen Stellen die gegebenen Werte hat.*

Approximationssatz: *Zu jeder stetigen Funktion f auf einem kompakten Intervall $[a, b]$ und zu jeder (noch so klein) gegebenen Fehlerschranke $\varepsilon > 0$ gibt es eine Polynomfunktion p , die f bis auf einen Fehler höchstens ε approximiert in dem Sinne, dass $|f(x) - p(x)| < \varepsilon$ ist für alle Stellen $x \in [a, b]$.*

Die erste Aussage ist nicht schwer zu beweisen: Sind $x_0, x_1, \dots, x_n \in \mathbb{R}$ die gegebenen Stellen (sog. *Stützstellen*) und $y_0, y_1, \dots, y_n \in \mathbb{R}$ die gegebenen Werte, so genügt es, für $j = 0 \dots n$ ein Polynom $L_j(x)$ vom Grad n anzugeben, das an der Stelle x_j den Wert 1, an den Stellen x_i mit $i \neq j$ den Wert 0 hat; denn dann ist offenbar $p(x) := \sum_{j=0}^n y_j L_j(x)$ eine Polynomfunktion vom Grad $\leq n$ mit $p(x_j) = y_j$ für alle $j = 1 \dots n$, und es ist auch die einzige, weil ja zwei Polynomfunktionen vom Grad $\leq n$ schon übereinstimmen, wenn sie an $n+1$ Stellen dieselben Werte haben. Die sog. *Lagrange-Polynome* $L_j(x)$ kann man leicht finden: Man multipliziert einfach alle Linearfaktoren $x - x_i$ mit $i \neq j$ aneinander und dividiert durch den Wert dieses Produkts an der Stelle $x = x_j$. Resultat ist die *Lagrange-Interpolationsformel*, welche die Zwei-Punkte-Formel ($n+1 = 2$) für lineare und die Drei-Punkte-Formel ($n+1 = 3$) für quadratische Funktionen verallgemeinert:

$$p(x) = y_0 L_0(x) + y_1 L_1(x) + \dots + y_n L_n(x) \quad \text{mit}$$

$$L_j(x) = \frac{(x - x_0) \cdot \dots \cdot (x - x_{j-1}) \cdot (x - x_{j+1}) \cdot \dots \cdot (x - x_n)}{(x_j - x_0) \cdot \dots \cdot (x_j - x_{j-1}) \cdot (x_j - x_{j+1}) \cdot \dots \cdot (x_j - x_n)}$$

ist das eindeutige Polynom $p(x)$ mit Werte y_j an den Stellen x_j für $j = 0 \dots n$. Es heißt das *Interpolationspolynom* zu den gegebenen $n+1$ Stützstellen und Werten, weil es Funktionswerte zwischen den Stützstellen “interpoliert” (außerhalb des Intervalls zwischen der kleinsten und der größten Stützstelle spricht man von “Extrapolation” durch $p(x)$).

Ist eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ gegeben, so kann man im Definitionsintervall $n+1$ Punkte x_j wählen, z.B. äquidistant $x_j = a + \frac{j}{n}(b-a)$ ($j = 0 \dots n$), und das Interpolationspolynom p vom Grad $\leq n$ bilden, das an den Stellen x_j dieselben Werte hat wie die Funktion f . Das heißt aber keineswegs, dass p auch zwischen den Stützstellen Werte hat, die nahe bei den Werten von f liegen, selbst wenn f ein ganz einfache Funktion ist wie etwa $1/(x^2+1)$ und wenn man n sehr groß macht, also die Stützstellen, an denen p mit f übereinstimmt, sehr dicht im Definitionsintervall wählt. Die Bedeutung des sog. *Approximationssatzes von Weierstraß* oben besteht darin, dass man die Polynomfunktion diesem Satz zufolge so wählen kann — auf andere Weise als beim Interpolationspolynom —, dass die Abweichung zwischen den Werten der stetigen Funktion f und den Werten von p *überall* im Definitionsintervall $[a, b]$ klein ist, so klein wie man will. (Zum Begriff der Stetigkeit siehe 4.2; alle elementaren Funktionen sind stetig.) Der Satz ist nicht so leicht zu beweisen, dass wir hier weiter darauf eingehen können, und man zahlt einen Preis, wenn man die Approximation sehr gut machen will: Der Grad von p wird dann unter Umständen sehr groß, und das macht das eigentliche Ziel der Approximation zunichte, nämlich eine “komplizierte” Funktion f durch eine “einfache” Funktion p ohne großen Fehler zu ersetzen. Zwar gibt es ein Rechenverfahren, mit dem man die Werte einer Polynomfunktion $p(x) = a_n x^n + \dots + a_1 x + a_0$ mit relativ wenig Aufwand ausrechnen kann, nämlich das sog. *Horner-Schema*

$$p(x) = (\dots((a_n \cdot x + a_{n-1}) \cdot x + a_{n-2}) \cdot x + \dots + a_1) \cdot x + a_0,$$

mit dem man den Wert $p(x)$ durch nur n Multiplikationen und n Additionen (“von innen nach außen” fortschreitend) erhält; aber wenn der Grad des Polynoms z.B. 1000 ist, so ist das schon sehr aufwendig. Deshalb kommt man in der Praxis nicht mit Polynomfunktionen alleine aus, sondern braucht noch weitere Funktionen.

Es gibt auch noch einen weiteren Grund, warum man bei der mathematischen Modellbildung in den Naturwissenschaften und auch in der Wirtschaftswissenschaft nicht mit Polynomfunktionen auskommt: Polynomfunktionen haben “globale” Eigenschaften, die sie als Modellfunktionen zur Beschreibung gewisser Abhängigkeiten zwischen Variablen grundsätzlich ungeeignet machen. So nehmen nichtkonstante Polynomfunktionen auf einem unbeschränkten Intervall wie $]0, \infty[$ stets Werte von beliebig großem Betrag an und sie können andererseits auf beschränkten Intervallen nicht beliebig große Werte annehmen. Es gibt aber ökonomische Situationen, in denen man von der Modellfunktion gerade die gegenteiligen Eigenschaften fordert, weil die betrachtete abhängige ökonomische Variable eben nicht beliebig groß werden kann bzw. weil sie bei Annäherung der unabhängigen Variablen an einen endlichen Wert schon beliebig groß wird.

Betrachten wir zum Beispiel eine einfache lineare Kostenfunktion $K(x) = K_{\text{fix}} + \kappa \cdot x$, bei der sich die Kosten für die Produktion von x Einheiten eines Gutes zusammensetzen aus einem zu x proportionalen Anteil $\kappa \cdot x$ und einem konstanten Fixkostenanteil K_{fix} der für das Betreiben der Produktionsanlage unabhängig von der produzierte Menge anfällt. Dann hat die zugehörige Stückkostenfunktion $k(x) = K_{\text{fix}}/x + \kappa$, welche die Kosten pro Einheit bei Produktions-Output x angibt, eine rechtsseitige Unendlichkeitsstelle bei $x = 0$ und bei $x \rightarrow \infty$ strebt sie gegen den endlichen Grenzwert κ . Dieses Verhalten ist auch bei ökonomischer Betrachtungsweise durchaus einsichtig: Bei sehr kleinen Produktionsmengen wirken sich die Fixkosten stark aus, so dass extrem hohe Stückkosten vorliegen, bei sehr großen Produktionsmengen fallen die Fixkosten dagegen kaum noch ins Gewicht und die Stückkosten sind praktisch gleich dem Proportionalitätsfaktor κ .

Nun gibt es zwar in der Realität weder beliebig kleine Produktionsmengen, noch beliebig große, aber es ist gewiss nicht sinnvoll, eine Funktion mit einem solchen Verhalten wie diese Stückkostenfunktion $k(x)$ durch Polynomfunktionen $p(x)$ zu approximieren, die bei $x \rightarrow \infty$ und bei $x \downarrow 0$ qualitativ ein völlig anderes Verhalten zeigen als die zu modellierende ökonomische Funktion! Also müssen wir die Klasse der Funktionen, die wir für die Modellierung ökonomischer Abläufe und Abhängigkeiten einsetzen wollen, erweitern und dafür sorgen, dass z.B. eine Funktion wie die Stückkostenfunktion $k(x) = (K_{\text{fix}} + \kappa \cdot x)/x$ oben mit zu dieser Klasse gehört. Diese Funktion ist kein Polynom, lässt sich aber als Quotient von zwei (linearen) Polynomen darstellen, und das bringt uns zu folgender Definition:

DEFINITION und DISKUSSION (*rationale Funktionen*):

1) Unter einer **rationalen Funktion** versteht man eine Funktion, die als Quotient $r = p/q$ von zwei Polynomfunktionen p und q darstellbar ist, also die Form

$$r(x) = \frac{p(x)}{q(x)} = \frac{a_m x^m + \dots + a_1 x + a_0}{b_n x^n + \dots + b_1 x + b_0}$$

hat. Der maximale Definitionsbereich des Terms $p(x)/q(x)$ ist \mathbb{R} ohne die endlich vielen Nullstellen des Nennerpolynoms, die auch **Definitionslücken** heißen (aber evtl. noch beseitigt werden können; natürlich ist die Nullfunktion hier nicht als Nenner zugelassen). Der Name kommt von der Analogie zu den rationalen Zahlen, also den als Bruch p/q von zwei ganzen Zahlen darstellbaren Zahlen, und weil die Brüche mit Nenner $q = 1$ dabei die ganzen Zahlen geben, nennt man rationale Funktionen mit konstantem Nennerpolynom $q \equiv 1$ (also Polynomfunktionen), eben auch *ganzrationale Funktionen* und solche, die so nicht darstellbar sind, *gebrochen-rationale Funktionen*. Summen, Differenzen, Produkte und Quotienten von rationalen Funktionen sind wieder welche (während Quotienten von Polynomen im Allgemeinen nicht wieder Polynome sind), ebenso auch Verkettungen (Einsetzen einer rationalen Funktion in eine andere). Alle rationalen Funktionen gehen durch die Rechenoperationen aus der identischen Grundfunktion x hervor, sind also elementar.

Es sind zwei ganz verschiedene Arten von Definitionslücken zu unterscheiden: Ist x_0 eine Nullstelle des Nennerpolynoms q , so kommt es darauf an, ob x_0 zugleich Nullstelle des Zählerpolynoms p ist und gegebenenfalls, ob die Nullstellenvielfachheit m_0 bei p mindestens so groß ist wie die Nullstellenvielfachheit n_0 bei q oder nicht. Klammern wir die maximale Anzahl von Linearfaktoren aus,

$$p(x) = (x - x_0)^{m_0} \tilde{p}(x), \quad q(x) = (x - x_0)^{n_0} \tilde{q}(x) \quad \text{mit} \quad \tilde{p}(x_0) \neq 0 \neq \tilde{q}(x_0),$$

so können wir im ersten Fall $m_0 \geq n_0$ alle Linearfaktoren im Nenner wegkürzen und erhalten eine Darstellung derselben rationalen Funktion $r = p/q$ durch einen Term

$$r(x) = \frac{\tilde{p}(x)}{\tilde{q}(x)} (x - x_0)^{m_0 - n_0} \quad (m_0 \geq n_0),$$

bei dem x_0 keine Definitionslücke mehr ist. Durch die Kürzung der gemeinsamen Linearfaktoren ist also die Definitionslücke behoben, und deshalb nennt man Definitionslücken dieser Art auch **hebbare Definitionslücken**. An solchen Stellen hat die rationale Funktion gar nichts Besonderes, im Graphen fehlte sozusagen lediglich ein Punkt, durch den der Graph "glatt" durchgeht. Bei einer hebbaren Definitionslücke hat man nur eine ungünstig erweiterte Bruchdarstellung der rationalen Funktion vorliegen, welche Nullstellen im Nenner produziert, die weggekürzt werden können. Den **maximalen Definitionsbereich** der rationalen Funktion erhält man, wenn man diese Kürzung für alle Nennernullstellen vorgenommen hat, d.h. wenn man die rationale Funktion so darstellt, dass Zähler- und Nennerpolynom keine gemeinsamen Nullstellen mehr haben.

Die so **gekürzte Darstellung** der rationalen Funktion, bei der Zählerpolynom und Nennerpolynom keine gemeinsamen reellen Nullstellen haben, kann dann unter Umständen noch weiter gekürzt werden, indem man gemeinsame Polynomfaktoren ohne reelle Nullstelle (wie etwa $x^2 + 1$) bei Zähler und Nenner abspaltet, jedoch sind solche gemeinsamen Faktoren schwerer zu finden als gemeinsame reelle Nullstellen (es geht z.B., indem man gemeinsame nichtreelle Nullstellen in \mathbb{C} bestimmt), und die *vollständig gekürzte Darstellung* der rationalen Funktion, bei der Zähler- und Nennerpolynom kleinstmöglichen Grad haben, braucht man nicht unbedingt herzustellen.

2) Ganz anders ist die Situation bei einer Nennernullstelle x_0 , die keine Zählernullstelle ist ($m_0 = 0$) oder eine Nullstelle des Zählerpolynoms von kleinerer Vielfachheit m_0 als die Vielfachheit n_0 derselben Nullstelle bzgl. des Nennerpolynoms. Solch eine Nullstelle des Nennerpolynoms nennt man eine **Polstelle** der rationalen Funktion $r(x) = p(x)/q(x)$. Nach Kürzen aller dem Zähler und Nenner gemeinsamen Linearfaktoren $x - x_0$ zu der Polstelle x_0 erhält man eine Darstellung

$$r(x) = \frac{\tilde{p}(x)}{\tilde{q}(x)} \frac{1}{(x - x_0)^{n_0 - m_0}} \quad (m_0 < n_0, \tilde{p}(x_0) \neq 0 \neq \tilde{q}(x_0)).$$

Wenn sich hier nun x der Stelle x_0 nähert, so wird der Nenner beliebig klein, während der Zähler gegen den Wert $\tilde{p}(x_0) \neq 0$ strebt, also wird $|r(x)|$ beliebig groß bei $x \rightarrow x_0$; in Worten:

- *Polstellen einer rationalen Funktion sind Unendlichkeitsstellen.*

Das bedeutet insbesondere, dass solche Definitionslücken durch keine Festsetzung des Wertes $r(x_0)$ zu beseitigen sind. Der maximale Definitionsbereich der rationalen Funktion ist also \mathbb{R} ohne ihre Polstellen.

Eine genauere Aussage über die Art der Unendlichkeitsstelle in einem Pol x_0 können wir machen, wenn wir die **Ordnung des Pols** bestimmen, das ist die Vielfachheit der Nennernullstelle x_0 , wenn das Zählerpolynom dort keine Nullstelle hat ($m_0 = 0$), bzw. die Differenz $n_0 - m_0$ der Nullstellenvielfachheit von x_0 in Nenner und Zähler andernfalls. (Am besten arbeitet man immer mit der gekürzten Darstellung der rationalen Funktion; dann gibt es keine gemeinsamen Nullstellen von Zähler und Nenner.) Da nun $(x - x_0)^{n_0 - m_0}$ dasselbe Vorzeichen für $x < x_0$ wie für $x > x_0$ hat, wenn $n_0 - m_0$ gerade ist, und entgegengesetztes Vorzeichen andernfalls, sehen wir aus der obigen Darstellung von $r(x)$ (in der $\tilde{p}(x)/\tilde{q}(x)$ nahe bei der Zahl $\tilde{p}(x_0)/\tilde{q}(x_0) \neq 0$ liegt und das Vorzeichen nicht wechselt, wenn x nahe bei x_0 ist):

- *In einem Pol gerader Ordnung hat die rationale Funktion eine Unendlichkeitsstelle ohne Vorzeichenwechsel, in einem Pol ungerader Ordnung eine Unendlichkeitsstelle mit Vorzeichenwechsel.*

Das heißt also: Der Grenzwert von $r(x)$ ist bei $x \uparrow x_0$ (wenn sich x von unten/links an x_0 annähert) und bei $x \downarrow x_0$ (wenn sich x von oben/rechts an x_0 annähert) derselbe unendliche Wert $+\infty$ bzw. $-\infty$, wenn die Ordnung der Polstelle x_0 gerade ist, die Grenzwerte von den beiden verschiedenen Seiten sind aber entgegengesetzt unendlich, wenn die Polordnung ungerade ist. Der *Prototyp* einer Polstelle k -ter Ordnung ist der Nullpunkt bei der Funktion $1/x^k$, und das Grenzverhalten einer rationalen Funktion bei einer Polstelle x_0 der Ordnung k ist so wie das Grenzverhalten von $\pm 1/x^k$ bei $x = 0$ (wobei das Vorzeichen das von $\tilde{p}(x_0)/\tilde{q}(x_0)$ oben ist).

3) Die Bestimmung der Vorzeichen der Werte einer rationalen Funktion kann ähnlich wie bei Polynomen dadurch erfolgen, dass man das Verhalten im Unendlichen und die Vorzeichenwechsel beim Überschreiten von Nullstellen und von Polstellen untersucht. Die **Nullstellen einer rationalen Funktion** $r(x) = p(x)/q(x)$ sind natürlich die Nullstellen des Zählerpolynoms $p(x)$, soweit es sich nicht um Polstellen oder um Definitionslücken handelt, bei der in der gekürzten Darstellung keine Nullstelle vorliegt. Mit andern Worten: Nullstellen von $r(x)$ sind die Nullstellen x_0 von $p(x)$, die keine Nullstelle von $q(x)$ sind oder nur Nullstellen kleinerer Vielfachheit n_0 von q als die Nullstellenvielfachheit m_0 bezüglich p . Die **Ordnung** oder **Vielfachheit der Nullstelle** x_0 von r wird entsprechend definiert als die Vielfachheit m_0 dieser Nullstelle bzgl. des Polynoms p , wenn $q(x_0) \neq 0$ ist, bzw. als die Differenz $m_0 - n_0$ der Vielfachheiten bzgl. Zähler- und Nennerpolynom, wenn x_0 auch Nennernullstelle ist. Man hat dann eine Darstellung $r(x) = (x - x_0)^{m_0 - n_0} \cdot \tilde{p}(x) / \tilde{q}(x)$ mit $\tilde{p}(x_0) \neq 0$ und $\tilde{q}(x_0) \neq 0$ und sieht wie oben:

- Bei einer Nullstelle ungerader Ordnung hat die rationale Funktion einen Vorzeichenwechsel, bei einer Nullstelle gerader Ordnung keinen.

Das bedeutet also nahe bei x_0 , dass $r(x) < 0$ ist auf einer Seite nahe x_0 und $r(x) > 0$ auf der anderen Seite, wenn x_0 Nullstelle ungerader Ordnung von r ist, während $r(x) > 0$ ist auf beiden Seiten oder $r(x) < 0$ auf beiden Seiten im Falle einer Nullstelle gerader Ordnung. Die Situation ist hier genau so wie bei Polynomen. Man muss nur aufpassen, dass man keine Nullstellen übersieht, die Definitionslücken sind, also Nennernullstellen, die auch Zählernullstellen sind, und zwar mit größerer Vielfachheit als beim Nenner.

Was die Anzahl der Null- und Polstellen von rationalen Funktionen angeht, so ist folgendes unmittelbar klar, weil Polynomfunktionen höchstens so viele Nullstellen haben (mit Vielfachheiten gerechnet) wie ihr Grad angibt:

- Eine nichtkonstante rationale Funktion $r(x) = p(x)/q(x)$ hat höchstens so viele Nullstellen (mit Vielfachheit gerechnet) wie der Grad m des Zählerpolynoms p ,
- höchstens so viele Polstellen (mit Ordnung gezählt) wie der Grad n des Nennerpolynoms q und
- für jedes $c \in \mathbb{R}$ höchstens $\max(m, n)$ Stellen x mit Funktionswert $r(x) = c$.

Die letzte Aussage gilt, weil $r(x) = c$ äquivalent ist mit $p(x) - cq(x) = 0$ und $q(x) \neq 0$, wobei das Polynom $p(x) - cq(x)$ höchstens den Grad $\max(m, n)$ hat. Dabei ist nicht gesagt, dass jede gegebene Zahl $c \in \mathbb{R}$ als Wert angenommen wird! Die Aussagen werden natürlich am schärfsten, wenn man die Bruchdarstellung $r(x) = p(x)/q(x)$ so weit wie möglich gekürzt hat, so dass die Grade von p und q kleinstmöglich sind.

4) Will man sich einen Überblick über den Verlauf einer rationalen Funktion $r(x)$ verschaffen, so ist neben den Polstellen und Nullstellen das “Verhalten im Unendlichen” wichtig, d.h. das Verhalten für großer Werte von $|x|$. Wie bei Polynomfunktionen gilt auch hier: Die **führenden Summanden bestimmen das Verhalten im Unendlichen**. Damit ist gemeint, wenn etwa

$$r(x) = \frac{a_m x^m + \dots + a_1 x + a_0}{b_n x^n + \dots + b_1 x + b_0} \quad \text{mit } a_m \neq 0 \neq b_n$$

ist, dass man im Zähler und Nenner alle Summanden bis auf den ersten weglässt und die “vereinfachte” Funktion $a_m x^m / b_n x^n = (a_m / b_n) x^{m-n}$ betrachtet, deren Verhalten bei $x \rightarrow \infty$ und $x \rightarrow -\infty$ unmittelbar klar ist.

Die Differenz von $r(x)$ und dieser vereinfachten Funktion kann man schreiben

$$r(x) - \frac{a_m x^m}{b_n x^n} = x^{m-n} \cdot \left[\frac{a_m + a_{m-1}x^{-1} + \dots + a_1 x^{1-m} + a_0 x^{-m}}{b_n + b_{n-1}x^{-1} + \dots + b_1 x^{1-n} + b_0 x^{-n}} - \frac{a_m}{b_n} \right],$$

wobei der Faktor [...] so klein ist wie man will, wenn $|x|$ groß genug ist. Dies bedeutet, dass der relative Fehler bei Ersetzung von $r(x)$ durch $(a_m x^m / b_n x^n)$, das ist der Fehler $|r(x) - (a_m x^m / b_n x^n)|$ dividiert durch $|(a_m x^m / b_n x^n)|$, kleiner als jede gegebene positive Fehlerschranke ist für hinreichend große Werte von $|x|$. Man sagt zu diesem Sachverhalt

$$r(x) \text{ ist asymptotisch gleich mit } \frac{a_m x^m}{b_n x^n} \text{ bei } x \rightarrow \pm\infty,$$

und notiert dafür

$$r(x) \sim \frac{a_m x^m}{b_n x^n} \quad (|x| \rightarrow \pm\infty), \quad \lim_{x \rightarrow \pm\infty} x^{n-m} \left| r(x) - \frac{a_m x^m}{b_n x^n} \right| = 0.$$

Um das Verhalten einer rationalen Funktion im Unendlichen zu beschreiben, müssen wir drei Fälle unterscheiden: Wenn der Zählergrad m größer als der Nennergrad n ist, so strebt x^{m-n} gegen ∞ bei $x \rightarrow \infty$ und entsprechend $(a_m x^m) / (b_n x^n)$ gegen $\sigma\infty$, wo $\sigma = \pm 1$ das Vorzeichen von a_m / b_n ist; bei $x \rightarrow -\infty$ strebt dann $(a_m x^m) / (b_n x^n)$ gegen $(-1)^{m-n} \sigma\infty$. Die rationale Funktion $r(x)$ hat dieselben unendlichen Grenzwerte bei $x \rightarrow \pm\infty$ wie $(a_m x^m) / (b_n x^n)$ und sie wächst dem Betrag nach dabei auch genau so schnell. Im zweiten Fall $m = n$ hat $r(x)$ den endlichen Grenzwert $a_n / b_n \neq 0$ bei $x \rightarrow \pm\infty$, d.h. $|r(x) - a_n / b_n|$ ist kleiner als jede gegebene Fehlerschranke $\varepsilon > 0$, wenn $|x|$ groß genug ist. Für große Werte von $|x|$ verläuft dann also die rationale Funktion $r(x)$ beliebig nahe bei der konstanten Funktion mit Wert $a_n / b_n \neq 0$. Im dritten Fall $m < n$ schließlich, wo der Zählergrad kleiner als der Nennergrad ist, hat $r(x)$ den Grenzwert Null bei $x \rightarrow \pm\infty$ und genauer wird $|r(x)|$ genau so schnell klein mit wachsendem $|x|$ wie $|x|^{m-n}$. Wir fassen das zusammen zu folgenden Aussagen über **das Verhalten einer rationalen Funktion im Unendlichen**:

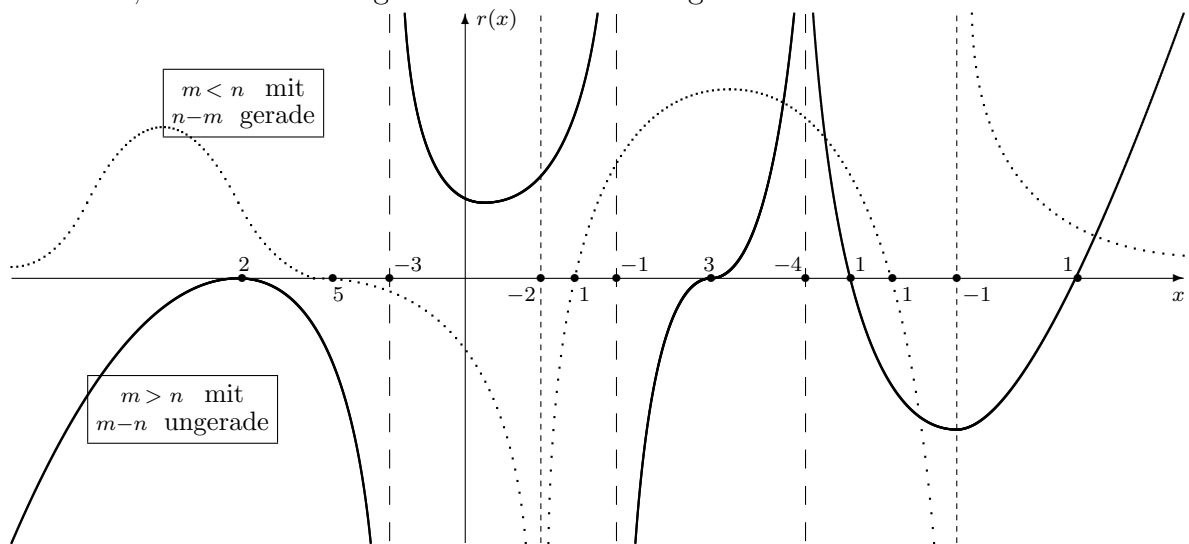
$$\underline{\text{Fall } m > n:} \quad \lim_{x \rightarrow \infty} r(x) = \sigma \cdot \infty = (-1)^{m-n} \lim_{x \rightarrow -\infty} r(x) \quad (\sigma \text{ das Vorzeichen von } \frac{a_m}{b_n});$$

$$\underline{\text{Fall } m = n:} \quad \lim_{x \rightarrow \infty} r(x) = \frac{a_n}{b_n} = \lim_{x \rightarrow -\infty} r(x);$$

$$\underline{\text{Fall } m < n:} \quad \lim_{x \rightarrow \infty} r(x) = 0 = (-1)^{m-n} \lim_{x \rightarrow -\infty} r(x).$$

Man benötigt hier für das Verständnis keine mathematische Grenzwerttheorie; es genügt eine intuitive Vorstellung, so wie es oben beschrieben ist. Man braucht sich nur zu merken, dass $r(x)$ grundsätzlich dasselbe Grenzverhalten bei $x \rightarrow \infty$ und bei $x \rightarrow -\infty$ hat wie die einfache Funktion $(a_m x^m) / (b_n x^n)$, die man erhält, wenn man im Zähler und Nenner der Bruchdarstellung der rationalen Funktion nur die führenden Summanden berücksichtigt. Das Grenzverhalten kann man also sofort durch Inspektion der Grade von Zähler- und Nennerpolynom und der führenden Koeffizienten erkennen, ohne noch irgendeine Rechnung ausführen zu müssen! Wir bemerken noch, dass eine rationale Funktion bei $x \rightarrow \infty$ und bei $x \rightarrow -\infty$ stets dieselben Grenzwerte hat mit einer Ausnahme, nämlich wenn der Zählergrad um eine ungerade Zahl größer ist als der Nennergrad, in welchem Falle die beiden Grenzwerte entgegengesetzt unendlich sind.

5) Das **Vorzeichen der Werte einer rationalen Funktion** r kann man nun so bestimmen: Für große Werte von x hat $r(x)$ dasselbe Vorzeichen wie der Quotient $\frac{a_m}{b_n}$ der führenden Koeffizienten in Zähler- und Nennerpolynom. Das sieht man aus der vorangegangenen Diskussion des Grenzverhaltens; es gilt unabhängig davon, ob der Grenzwert bei $x \rightarrow \infty$ unendlich oder endlich $\neq 0$ oder Null ist. Wenn man dann x verkleinert, so kann sich das Vorzeichen von $r(x)$ höchstens bei Überschreitung einer Null- oder Polstelle ändern (weil die Funktion zwischen den Polstellen stetig ist und daher bei Vorzeichenänderung in einem Intervall gemäß Zwischenwertsatz auch eine Nullstelle in diesem Intervall haben muss; siehe 4.2). Ist die Vielfachheit der überschrittenen Nullstelle bzw. die Ordnung der überschrittenen Polstelle ungerade, so ändert sich das Vorzeichen von $r(x)$ dort auch wirklich, d.h. die Funktionswerte haben auf den beiden Seiten der Stelle nahebei verschiedene Vorzeichen, bei Nullstellen gerader Vielfachheit oder Polstellen gerader Ordnung ist das Vorzeichen dagegen auf beiden Seiten der Stelle nahebei gleich. (Man darf dabei nicht vergessen, die hebbaren Definitionslücken, in denen das Zählerpolynom von höherer Vielfachheit verschwindet als das Nennerpolynom, auch als Nullstellen zu berücksichtigen; sonst übersieht man unter Umständen einen Vorzeichenwechsel.) Die folgende Abbildung zeigt verschiedene Möglichkeiten; die positiven Zahlen geben die Vielfachheit der Nullstellen an, die negativen die negative Ordnung (= "negative Vielfachheit") der Polstellen; m ist der Zählergrad und n der Nennergrad.



Aus dem Zwischenwertsatz in 4.2 folgt übrigens, dass $r(x)$ auf einem Intervall mit Polstellen oder $\pm\infty$ als Randpunkten alle reellen Zahlen als Werte annimmt, wenn $r(x)$ entgegengesetzt unendliche Grenzwerte an den beiden Randpunkten dieses Intervalls hat und keine weiteren Polstellen im Intervall. Hat $r(x)$ dagegen gleiche Grenzwerte ∞ (bzw. $-\infty$) an den beiden Endpunkten eines Intervalls ohne innere Polstellen, so nimmt $r(x)$ auf diesem Intervall ein endliches Minimum an (bzw. ein endliches Maximum) und jeden größeren Wert (bzw. jeden kleineren) an mindestens zwei Stellen in diesem Intervall. Ein Blick auf die Abbildung macht diesen Sachverhalt klar.

6) Das Ergebnis in (5), dass sich eine rationale Funktion $r(x)$ bei $x \rightarrow \pm\infty$ asymptotisch wie der Quotient $a_m x^m / b_n x^n$ der führenden Summanden im Zähler und Nenner verhält, läßt sich noch verbessern, wenn der Zählergrad m größer als der Nennergrad n ist. Dieses Ergebnis garantiert nämlich nur, dass der "Fehler" $|r(x) - a_m x^m / b_n x^n|$ für große Werte $|x|$ klein ist im Verhältnis zu $|x|^{m-n}$, aber im Fall $m > n$ bedeutet das nicht, dass der Fehler klein wird bei $x \rightarrow \pm\infty$, weil ja $|x|^{m-n}$ gegen unendlich geht.

Eine einfache Asymptotenfunktion $s(x)$ zu $r(x)$, für die der Fehler $|r(x) - s(x)|$ tatsächlich beliebig klein wird für große Werte von $|x|$, erhält man durch Polynomdivision mit Rest. Dies ist ein Algorithmus, mit dem man den Quotienten $p(x)/q(x)$ von zwei Polynomen in der Form $s(x) + \tilde{p}(x)/q(x)$ darstellen kann mit einem Polynom s und einem Polynom \tilde{p} von kleinerem Grad als q . Wenn der Grad m von p kleiner ist als der von q , so setzt man $s := 0$ und ist schon fertig. Im Fall $m \geq n$ spaltet man in einem ersten Schritt im Zähler den Summanden $(a_m/b_n)x^{m-n}q(x)$ ab, der genau denselben führenden Term a_mx^m wie $p(x)$ hat, und schreibt

$$\text{mit } \frac{p(x)}{q(x)} = \frac{\frac{a_m}{b_n}x^{m-n}q(x) + p_1(x)}{q(x)} = c_{m-n}x^{n-m} + \frac{p_1(x)}{q(x)},$$

$$c_{m-n} := \frac{a_m}{b_n} \quad \text{und} \quad p_1(x) := p(x) - \frac{a_m}{b_n}x^{m-n}q(x),$$

wobei der Grad m_1 von p_1 kleiner als m ist, weil beide Polynome in dieser Differenz denselben führenden Term a_mx^m besitzen. Wenn nun p_1 kleineren Grad als q hat, so setzt man $s(x) := (a_m/b_n)x^{m-n}$ und $\tilde{p} := p_1$ und ist fertig. Wenn nicht, so wiederholt man das Verfahren mit $p_1(x)/q(x)$ statt $p(x)/q(x)$ und erhält $p(x)/q(x) = c_{m-n}x^{m-n} + c_{m_1-n}x^{m_1-n} + p_2(x)/q(x)$ mit $\text{Grad}(p_2) < m_1$. Nun setzt man das Verfahren nötigenfalls mit $p_2(x)/q(x)$ fort usw., bis schließlich der Zählergrad kleiner als der Grad des Nenners q ist. Resultat ist die **Polynomdivision mit Rest**:

$$\frac{p(x)}{q(x)} = s(x) + \frac{\tilde{p}(x)}{q(x)} \quad \text{mit} \quad \text{Grad}(\tilde{p}) < \text{Grad}(q) \quad \text{und}$$

$$s(x) = c_{m-n}x^{m-n} + \dots + c_1x + c_0 \quad \text{vom Grad } m - n.$$

Weil im “Rest” $\tilde{p}(x)/q(x)$ das Zählerpolynom kleineren Grad als das Nennerpolynom hat, geht dieser Term gegen Null bei $x \rightarrow \pm \infty$, also ist $s(x)$ tatsächlich eine **polynomiale Asymptotenfunktion** für die rationale Funktion $r(x) = p(x)/q(x)$ bei $x \rightarrow \infty$ und bei $x \rightarrow -\infty$ in dem Sinne, dass der “Fehler” $|r(x) - s(x)|$ kleiner ist als jede gegebene Fehler-schranke $\varepsilon > 0$, wenn nur $|x|$ groß genug ist. Am Vorzeichen des Restes für große $x > 0$ bzw. betragsgroße $x < 0$, das sich an den Graden und den führenden Koeffizienten von \tilde{p} und q unmittelbar ablesen läßt, kann man außerdem noch erkennen, ob sich $r(x)$ für große $x > 0$ bzw. betragsgroße $x < 0$ von oben oder von unten der Asymptotenfunktion $s(x)$ nähert. Übrigens ist eine polynomiale Asymptotenfunktion im obigen Sinne eindeutig bestimmt; denn wenn die Differenz von zwei Polynomen Grenzwert Null im Unendlichen hat, so müssen beide Polynome übereinstimmen. ■

BEISPIELE (von rationalen Funktionen):

(1) Die einfachste nichtpolynomiale rationale Funktion ist $r(x) = \frac{1}{x}$.

Diese Funktion hat eine Polstelle in $x = 0$, die gewissermaßen der Prototyp einer Polstelle erster Ordnung ist. Bei $x \rightarrow \pm \infty$ strebt $1/x$ gegen Null, und diese Funktion ist auch das einfachste Beispiel einer Funktion mit Grenzwert Null im Unendlichen. Der Graph der Funktion $1/x$ ist die bekannte *Normalhyperbel*, und dass die Funktion bei Null eine Polstelle und im Unendlichen den Limes Null hat, bedeutet geometrisch, dass die Hyperbel die vertikale und die horizontale Achse als Asymptotengeraden besitzt. Funktionen der Form $y = b/x$ mit $b \neq 0$ kommen auch in ökonomischem Zusammenhang vor; denn sie beschreiben **umgekehrt proportionale Abhängigkeit** der Variablen y von x , also die Abhängigkeit, die dem *umgekehrten Dreisatz* entspricht. Übrigens ist dann auch $x = b/y$ umgekehrt proportional zu y mit demselben Proportionalitätsfaktor; die Funktion ist also gleich ihrer Umkehrfunktion. Für den Graphen bedeutet dies, dass er zur Diagonalen $\{(x, y) : x = y\}$ spiegelsymmetrisch ist.

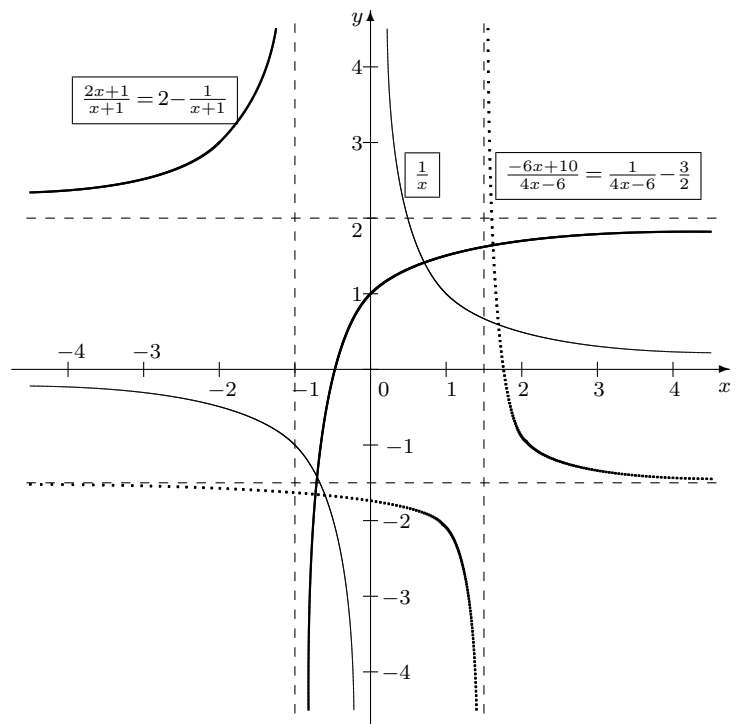
(2) **Gebrochen–lineare Funktionen** sind rationale Funktionen der Form

$$r(x) = \frac{ax + b}{cx + d} \quad (ad - bc \neq 0),$$

also Quotienten von zwei linearen Funktionen. Die Voraussetzung $ad - bc$ bedeutet, dass die lineare Zählerfunktion kein konstantes Vielfaches der Nennerfunktion ist. Im Fall $c = 0$ ist $r(x)$ eine lineare Funktion. Im Fall $c \neq 0$, der hier allein interessiert, hat die Funktion eine Polstelle erster Ordnung in $x = -d/c$ (beachte, dass der Zähler wegen der Voraussetzung $ad - bc \neq 0$ an dieser Stelle keine Nullstelle haben kann). Bei $x \rightarrow \pm \infty$ hat $r(x)$ denselben Grenzwert wie ax/cx also den endlichen Grenzwert a/c . Mit Polynomdivision kann man noch etwas mehr sehen:

$$r(x) = \frac{ax + b}{cx + d} = \frac{\frac{a}{c}(cx + d) + b - \frac{a}{c}d}{cx + d} = \frac{a}{c} + \frac{bc - ad}{c^2} \frac{1}{x + \frac{d}{c}}$$

zeigt, dass die Funktion bei $x \rightarrow \infty$ von unten gegen den Grenzwert a/c strebt, wenn $ad - bc > 0$ ist, bzw. von oben, wenn $ad - bc < 0$. Bei $x \rightarrow -\infty$ verhält es sich dagegen umgekehrt, die Funktion ist also für große $|x|$ auf einer Seite von Null größer und auf der anderen Seite kleiner als der Grenzwert a/c , dem sie sich im Unendlichen nähert. Außerdem liest man ab, dass der Graph der gebrochen–linearen Funktion r aus dem Graphen der Funktion $1/x$ entsteht durch horizontale Verschiebung um $-\frac{d}{c}$, durch anschließende vertikale Streckung oder Stauchung mit dem Faktor $(bc - ad)/c^2$ (was auch eine Spiegelung an der horizontalen Achse beinhaltet, wenn der Faktor negativ ist) und schließlich durch vertikale Verschiebung um $\frac{a}{c}$. Im Wesentlichen sind also die (echt, d.h. $c \neq 0$) gebrochen–linearen Funktionen nur leichte Modifikationen der Funktion $1/x$; die Graphen sind stets zwei Hyperbeläste mit einer horizontalen und einer vertikalen Asymptotengeraden.

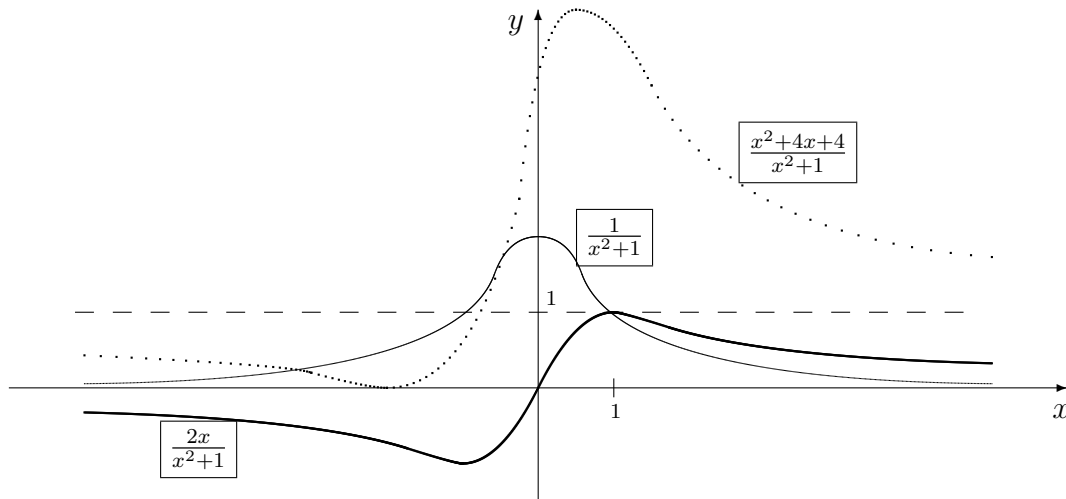


Die Umkehrfunktion einer gebrochen–linearen Funktion $r(x) = \frac{ax+b}{cx+d}$ existiert immer und ist wieder eine gebrochen–lineare Funktion mit Polstelle in a/c ; durch Auflösen von $y = (ax + b)/(cx + d)$ erhält man $r^{-1}(y) = x = (dy - b)/(-cy + a)$. In ökonomischem Zusammenhang sind wir auf eine gebrochen–lineare Funktion gestoßen, als wir oben die Stückkostenfunktion $k(x) = (K_{\text{fix}} + \kappa \cdot x)/x$ zu einer linearen Kostenfunktion betrachtet haben. Diese Funktion $k(x)$ hat eine Polstelle in 0 und den Grenzwert κ bei $x \rightarrow \infty$, und das ist auch ökonomisch sinnvoll.

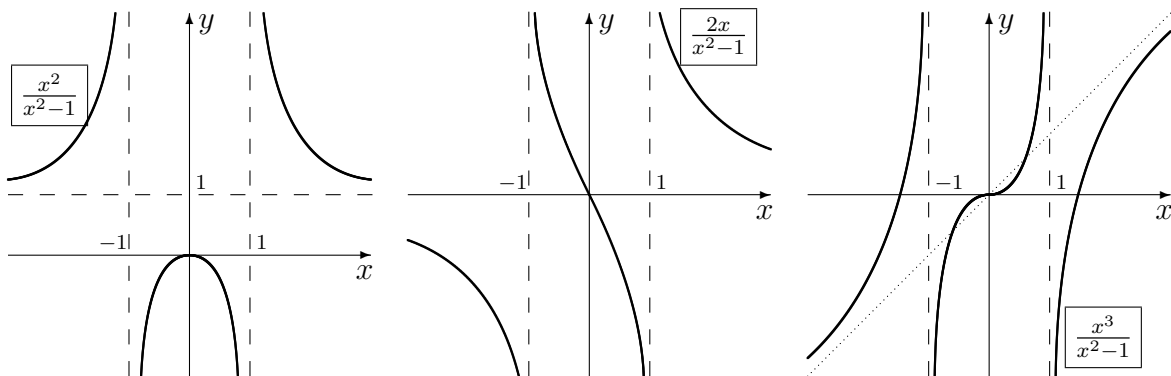
(3) Bei quadratischen Nennern

$$r(x) = \frac{p(x)}{ax^2 + bx + c} \quad (a \neq 0)$$

gibt es bzgl. der Polstellen folgende Möglichkeiten: Keine Polstelle vorhanden, wenn der Nenner keine Nullstellen hat, zwei Pole erster Ordnung, wenn der Nenner zwei Nullstellen hat, die nicht auch Zählernullstellen sind (sonst kann man einen Linearfaktor kürzen), eine Polstelle der Ordnung 2, wenn der Nenner genau eine Nullstelle hat (die dann Vielfachheit 2 haben muss) und der Zähler an dieser Stelle nicht Null ist. Für das Verhalten im Unendlichen gilt: Der Grenzwert von $r(x)$ bei $x \rightarrow \pm \infty$ ist 0, wenn der Zählergrad ≤ 1 ist, der Limes ist ein endlicher Wert $\neq 0$, wenn auch der Zählergrad 2 ist. In diesen Fällen hat man eine auf ganz \mathbb{R} beschränkte Funktion vorliegen (d.h. alle Werte liegen zwischen zwei festen reellen Zahlen), wenn es keine Polstellen gibt. Der Limes von $r(x)$ bei $x \rightarrow \pm \infty$ ist dagegen unendlich, wenn der Zählergrad > 2 ist (wobei man für $x \rightarrow \infty$ und $x \rightarrow -\infty$ denselben unendlichen Grenzwert erhält, wenn der Zählergrad gerade ist, und entgegengesetzte unendliche Werte andernfalls).



beschränkte rationale Funktionen auf \mathbb{R} (ohne Pole, endlicher Limes bei $x \rightarrow \pm \infty$)

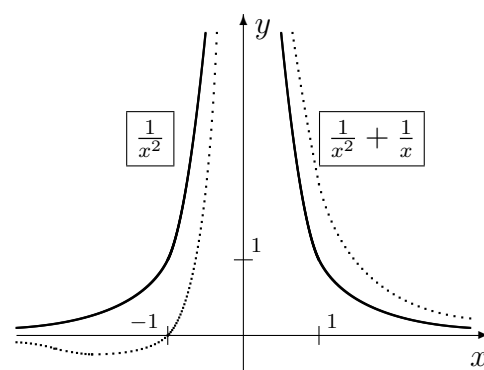


drei rationale Funktionen mit Polen erster Ordnung in $+1$ und -1

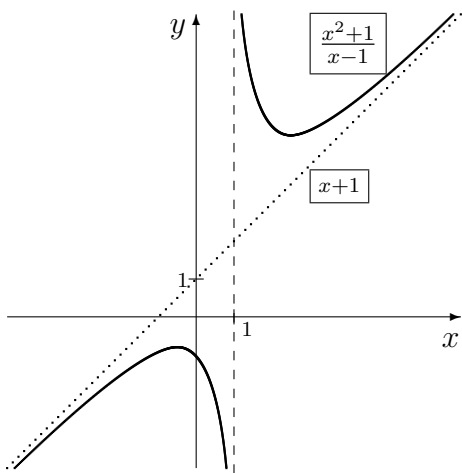
Die erste der drei Funktionen in der letzten Abbildung ist übrigens ein Beispiel dafür, dass eine rationale Funktion ein ganzes Intervall (hier $]0, 1[$) von Werten auslassen kann, obwohl sie beliebig große und beliebig kleine Werte auf \mathbb{R} annimmt. Der Zwischenwertsatz (siehe 4.2) gilt eben für rationale Funktionen mit Polstellen auf \mathbb{R} im Allgemeinen nicht, weil Polstellen Unstetigkeitsstellen sind.

Auch bei der rationalen Funktion $(2x+2)/(x^2-1)$ hat der Nenner die Nullstellen $+1$ und -1 , aber die rationale Funktion hat dennoch nur eine Polstelle; denn $x = -1$ ist auch Nullstelle des Zählers, und Kürzen des Linearfaktors $x + 1$ gibt die Formel $2/(x - 1)$ für dieselbe rationale Funktion mit nun behobener Definitionslücke bei $x = -1$. *Das Vorhandensein einer Polstelle läßt sich nicht allein am Verschwinden des Nenners erkennen* — es muss auch untersucht werden, ob der Zähler nicht an derselben Stelle eine Nullstelle hat, und falls ja, ob die Vielfachheit der Nullstelle des Zählers kleiner ist als die Vielfachheit derselben Nullstelle des Nenners!

Der Prototyp einer Polstelle zweiter Ordnung ist natürlich $x = 0$ bei der Funktion $1/x^2$. Ebenfalls eine Polstelle zweiter Ordnung in $x = 0$ hat $1/x + 1/x^2 = (x + 1)/x^2$; diese Funktion nähert sich ihrem Grenzwert 0 von verschiedenen Seiten bei $x \rightarrow \infty$ bzw. $x \rightarrow -\infty$, während $1/x^2$ natürlich stets positiv ist und sich seinem Grenzwert 0 bei $x \rightarrow \pm \infty$ von oben nähert. Bis auf eine horizontale und eine vertikale Verschiebung und Streckung oder Stauchung sowie Spiegelung an einer Achse sieht der Graph jeder rationalen Funktion mit einem Nenner $(x - x_0)^2$ und einem Zähler vom Grad ≤ 2 so aus wie der von $1/x^2$ oder der von $1/x + 1/x^2$. Dazu überlegt man sich, dass solch eine Funktion bis auf eine additive Konstante als Linearkombination $a/(x - x_0)^2 + b/(x - x_0)$ der beiden sog. *Partialbrüche* $1/(x - x_0)^2$ und $1/(x - x_0)$ geschrieben werden kann, und formt dies im Fall $b \neq 0$ noch um zu $(b^2/a)[(1/(\frac{b}{a}x - \frac{b}{a}x_0)^2 + 1/(\frac{b}{a}x - \frac{b}{a}x_0))]$.



(4) Polynomiale Asymptotenfunktionen sind oben schon vorgekommen. Zum Beispiel haben die rationalen Funktionen $r(x)$, die einen endlichen Grenzwert im Unendlichen haben (das ist dann bei $x \rightarrow \infty$ und bei $x \rightarrow -\infty$ automatisch derselbe Grenzwert), eine konstante Asymptotenfunktion, deren Wert eben dieser Limes im Unendlichen ist. Bei der Abbildung des Graphen der Funktion $x^3/(x^2 - 1) = x + x/(x^2 - 1)$ in 3) oben haben wir den Graphen der linearen Asymptotenfunktion $s(x) = x$ auch schon eingezeichnet und berücksichtigt, dass die Funktion für große x größere Werte als diese Asymptotenfunktion hat, für betragsgroße negative x aber kleinere Werte. Hier ein weiteres Beispiel mit linearer



Asymptotenfunktion im Unendlichen (wir führen die zwei Schritte der Polynomdivision gleich auf einmal durch, indem wir im Zähler passende Vielfache des Nenners als Summanden abspalten):

$$\frac{x^2+1}{x-1} = \frac{x(x-1) + (x-1) + 2}{x-1} = x + 1 + \frac{2}{x-1}.$$

Diese rationale Funktion hier hat also die lineare Asymptotenfunktion $x + 1$ im Unendlichen, und sie nähert sich ihr von oben bei $x \rightarrow \infty$ und von unten bei $x \rightarrow -\infty$. Bei $x = 1$ hat sie eine Polstelle erster Ordnung, wobei sie bei Annäherung von rechts an diese Stelle positive Werte, also den Limes ∞ hat, bei Annäherung von links den Limes $-\infty$ (bei einer Polstelle erster Ordnung gibt es ja einen Vorzeichenwechsel).

Auch diese Funktion läßt ein Intervall in \mathbb{R} als Werte aus (nämlich alle Zahlen zwischen $2 - 2\sqrt{2}$ und $2 + 2\sqrt{2}$).

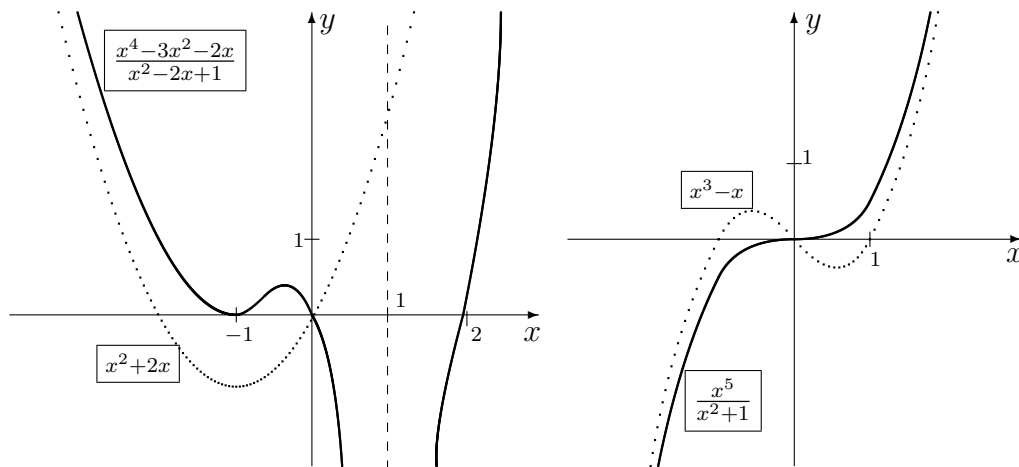
Hier noch ein Beispiel mit Asymptotenpolynom vom Grad 2:

$$\frac{x^4 - 3x^2 - 2x}{x^2 - 2x + 1} = \frac{x^2(x^2 - 2x + 1) + 2x(x^2 - 2x + 1) - 4x}{x^2 - 2x + 1} = x^2 + 2x - \frac{4x}{x^2 - 2x + 1},$$

und eines mit Asymptotenfunktion vom Grad 3:

$$\frac{x^5}{x^2 + 1} = \frac{x^3(x^2 + 1) - x(x^2 + 1) + x}{x^2 + 1} = x^3 - x + \frac{x}{x^2 + 1}.$$

Die Polynomdivision führt man bei solchen noch relativ übersichtlichen Termen am einfachsten durch, indem man mit x -Potenzen multiplizierte Vielfache des Nenners im Zähler so ergänzt, dass ein Polynom von kleinerem Grad als der Nenner übrig bleibt. Ein Blick auf die Vorzeichen der Reste zeigt jeweils, wo die rationale Funktion über und wo sie unter ihrer Asymptotenfunktion verläuft. Daraus entnimmt man dann folgenden qualitativen Funktionsverlauf:



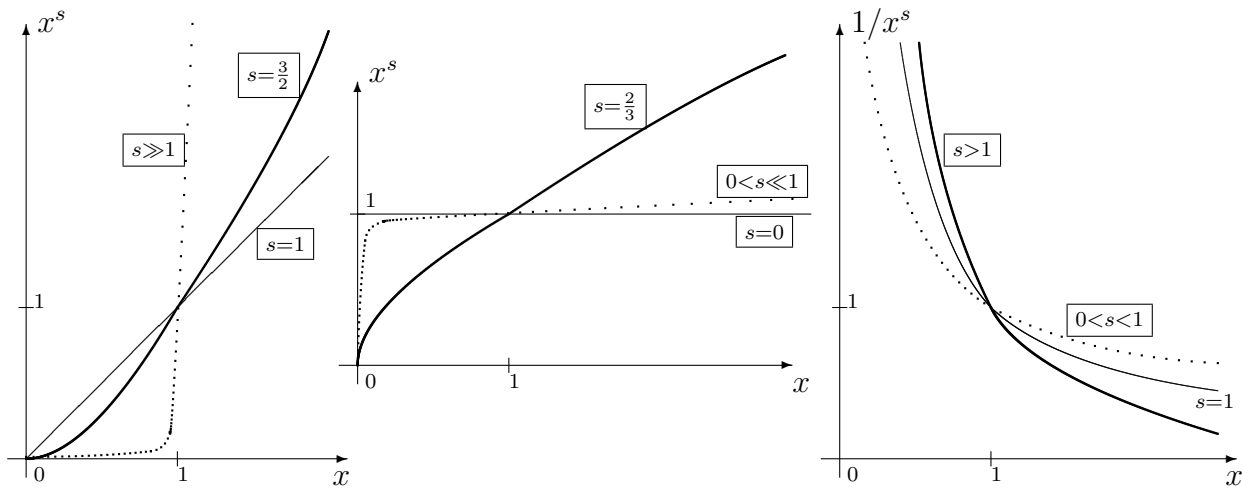
Rationalen Funktionen können, wie die vorangehenden Beispiele belegen, sehr unterschiedliches Funktionsverhalten modellieren. Trotzdem läßt sich mit ihnen noch nicht jede in der Natur und auch in der Ökonomie auftretende Abhängigkeit zwischen zwei Variablen gut beschreiben. Zum Beispiel haben rationale Funktionen $r(x)$ bei $x \rightarrow \infty$ stets ein asymptotisches Verhalten wie eine Potenz x^k mit einem ganzen Exponenten k ; denn $r(x)$ verhält sich ja wie der Quotient $a_m x^m / b_n x^n$ der führenden Summanden des Zähler- und des Nennerpolynoms in einer Bruchdarstellung der Funktion als Quotient von zwei Polynomen. Ein “gebrochenes” Wachstum wie etwa $\sqrt{x} = x^{1/2}$ oder $\sqrt[3]{x^4} = x^{4/3}$ bei $x \rightarrow \infty$ läßt sich mit rationalen Funktionen also nicht beschreiben, ebensowenig exponentielles Wachstum a^x für $x \rightarrow \infty$ bei Basis $a > 1$, also ein Wachstum, das schneller ist als das jeder Potenz x^k , und auch nicht logarithmisches Wachstum $\log_a x$ bei $x \rightarrow \infty$, also ein Wachstum gegen unendlich, das langsamer ist als bei jeder Potenzfunktion x^k mit $k \in \mathbb{N}$ und auch langsamer als bei jeder Wurzelfunktion $\sqrt[n]{x} = x^{1/n}$. Diese Typen des Wachstums kommen aber in theoretischen Modellen ökonomischer Vorgänge durchaus vor, z.B. Wurzelfunktionen als Modell von Größen, die degressiv wachsend gegen unendlich gehen, wie etwa die Kosten $K(x)$ einer mit wachsendem Output x immer günstigeren Produktion (d.h. abnehmende Stückkosten), oder Exponentialfunktionen als Modell für einen kontinuierlichen Vorgang ungehemmten Wachstums, bei dem sich eine Größe in gleichen Zeitabständen stets um denselben Faktor > 1 vergrößert, wie z.B. bei Aufzinsung mit kontinuierlicher Verzinsung. Zwar gibt es in der Realität kein ungehemmtes Wachstum über beliebig lange Zeiträume, aber es ist sicher wünschenswert, solche “theoretischen” ökonomischen Vorgänge mathematisch simulieren zu können.

Ein anderes Phänomen, das nicht mit rationalen Funktionen modelliert werden kann, ist das Vorhandensein von zwei unterschiedlichen endlichen Grenzwerten im Unendlichen bei einer Funktion f , also $\lim_{x \rightarrow \infty} f(x) = b$ und $\lim_{x \rightarrow -\infty} f(x) = a$ mit $a \neq b$. Solche Funktionen werden auch in der Wirtschaftsmathematik verwendet, um z.B. wachsende zeitabhängige Größen zu modellieren, die in endlicher Zeit einen gewissen maximalen Sättigungswert praktisch (im Rahmen der Rechengenauigkeit) erreichen. Eine rationale Funktion hat aber, wie wir gesehen haben, bei ∞ und bei $-\infty$ immer denselben Grenzwert, wenn diese Grenzwerte endlich sind. Solch eine Funktion kann daher nicht monoton wachsend auf \mathbb{R} sein und dabei endliche Grenzwerte im Unendlichen besitzen. Also kann man mit rationalen Funktionen einen monotonen Wachstumsvorgang mit endlicher oberer Sättigungsgrenze nicht modellieren.

Die Konsequenz aus alledem ist, dass man bei der mathematischen Modellierung ökonomischer Abläufe und Variablenabhängigkeiten nicht mit rationalen Funktionen auskommt, sondern die Menge der betrachteten mathematischen Funktionen noch etwas erweitern muss. Da die algebraischen Rechenoperationen und Verkettungen nicht aus dem Bereich der rationalen Funktionen herausführen, kann diese Erweiterung nur dadurch erfolgen, dass man zur identischen Funktion x noch weitere elementare Grundfunktionen hinzufügt, und deshalb haben wir die Exponentialfunktion e^x und die Logarithmusfunktion $\ln x$ mit zu den elementaren Grundfunktionen genommen. Weil wir damit dann "Rechterme" bilden können, in denen außer Potenzen x der unabhängigen Variablen auch noch Exponentiale $\exp(x) = e^x$ und Logarithmen $\ln x$ auftreten können, haben wir dann auch Funktionen zur Verfügung mit einem Verhalten wie oben, das sich durch rationale Funktionen nicht modellieren lässt. Mit dieser Erweiterung der Klasse der rationalen Funktionen zur Klasse der elementaren Funktionen sind in der Tat alle mathematischen Funktionen erfasst, die in Modellen für ökonomische Vorgänge und Gesetzmäßigkeiten Verwendung finden. (Etwas allgemeiner sind noch "stückweise elementare Funktionen", die durch Aneinandersetzen von verschiedenen elementaren Funktionen auf aneinanderstoßenden Intervallen definiert werden; siehe das Ende dieses Abschnitts 4.1. Die Behandlung solcher zusammengesetzten Funktionen läuft aber letztlich auch auf die Diskussion der einzelnen Bausteine hinaus, die elementarer Funktionen sind. Wir haben schon erwähnt, dass man in Mathematik und Naturwissenschaften noch weitere Grundfunktionen benötigt, nämlich eine Kreisfunktion und die Umkehrfunktion einer Kreisfunktion, die aber für mathematische Modelle ökonomischer Vorgänge nicht wichtig sind.)

BEISPIELE (*weitere elementare Funktionen*):

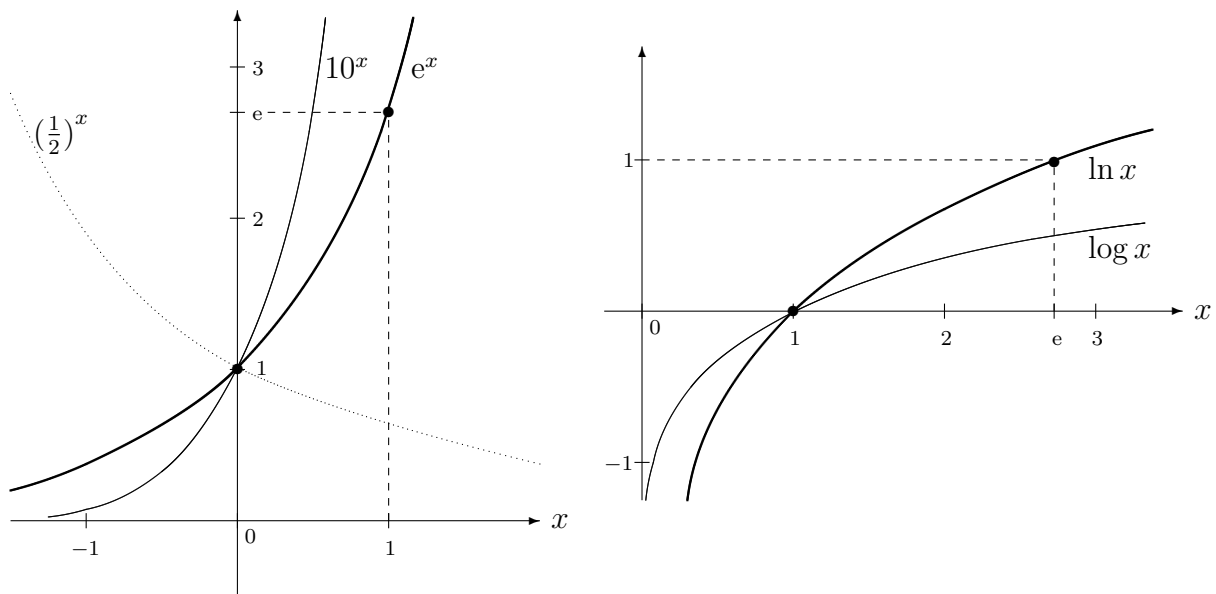
(1) Die **allgemeinen Potenzfunktionen** x^s mit reellem Exponenten s sind elementar auf $\mathbb{R}_{>0}$, wie man aus der Darstellung $x^s = \exp(s \ln x)$ als Verkettung der Exponentialfunktion mit dem s -fachen der Logarithmusfunktion erkennt. Für $s > 0$ kann man die Funktion an der Stelle $x = 0$ noch durch die Festsetzung $0^s := 0$ definieren, was der Grenzwert von x^s bei $x \downarrow 0$ ist. (Die so auf $\mathbb{R}_{\geq 0}$ definierte Funktion x^s ist aber – streng genommen – für nichtganze $r > 0$ nicht mehr elementar. Wenn z.B. $x^{1/2} = \sqrt{x}$ auf $\mathbb{R}_{\geq 0}$ elementar wäre, so auch die Verkettung der Quadratwurzelfunktion mit der Quadratfunktion $|x| = \sqrt{x^2}$ auf ganz \mathbb{R} . Aber die Betragsfunktion ist nicht elementar bei $x = 0$, weil ihr Graph dort eine Knickstelle hat, was bei elementaren Funktionen ausgeschlossen ist, wie mit Differentialrechnung in 4.3 gezeigt wird.) Die folgenden Abbildungen zeigen den qualitativen Verlauf der Graphen von allgemeinen Potenzfunktionen für verschiedene Exponenten s :



die Graphen der Potenzfunktionen x^s und $1/x^s$ auf $\mathbb{R}_{>0}$

Für sehr große Exponenten $s \gg 1$ sehen die Graphen von x^s und $x^{1/s}$ aus wie “rechte Winkel mit einem Schenkel der Länge 1 und einem unendlich langen Schenkel” und die Graphen von x^{-s} und $x^{-1/s}$ wie “rechte Winkel mit zwei unendlich langen Schenkeln”.

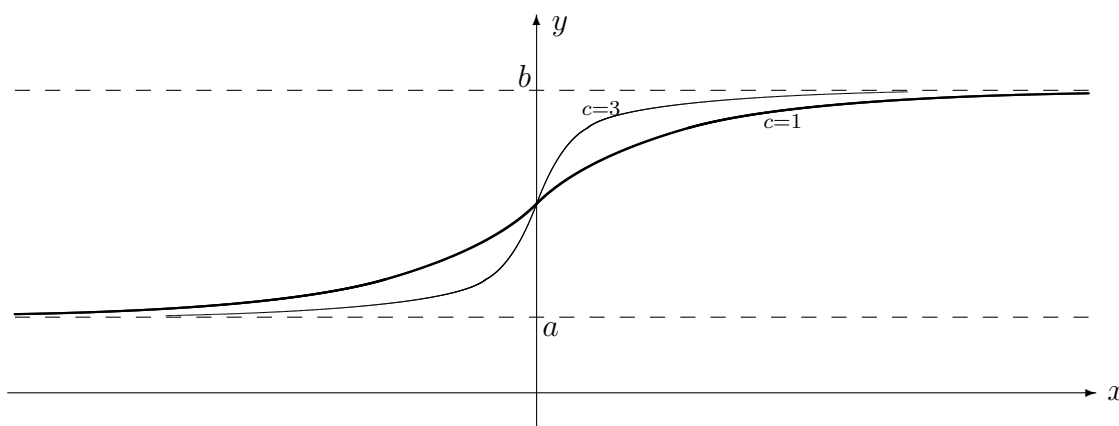
(2) Allgemeine Exponentialfunktionen $a^x = \exp(x \ln a)$ sowie **allgemeine Logarithmusfunktionen** $\log_a x = (\ln x) / \ln a$ zur Basis $a > 0$ (und $a \neq 1$ im Fall der Logarithmusfunktion) sind elementare Funktionen auf \mathbb{R} bzw. auf $\mathbb{R}_{>0}$; denn a^x ist die Verkettung der äußeren Funktion $\exp(y) = e^y$ mit der linearen inneren Funktion $x \mapsto x \ln a$ und \log_a ist ein Vielfaches der natürlichen Logarithmusfunktion \ln . Der Graph von a^x entsteht aus dem der natürlichen Exponentialfunktion e^x durch Streckung / Stauchung der Rechtswerte mit dem Faktor $\frac{1}{|\ln a|}$, wobei im Fall $0 < a < 1$ noch eine Spiegelung an der vertikalen Achse hinzukommt. Die Logarithmusfunktion \log_a ist die Umkehrfunktion zur Exponentialfunktion a^x auf \mathbb{R} ; ihr Graph entsteht also aus dem Graphen dieser Exponentialfunktion durch Spiegelung an der Diagonalen zu den beiden Koordinatenachsen und aus dem Graphen der natürlichen Logarithmusfunktion durch Streckung / Stauchung der Hochwerte mit dem Faktor $\frac{1}{\ln a}$ (wenn $a > 1$). Für sehr große Werte der Basis a bzw. (was auf dasselbe hinausläuft) “aus großer Entfernung betrachtet” sehen die Graphen der Exponentialfunktionen und Logarithmusfunktionen fast so aus wie die Vereinigung einer vertikalen und einer horizontalen Halbgeraden, die sich im Ursprung treffen.



(3) Eine elementare Funktion auf \mathbb{R} , die sich bei $x \rightarrow \infty$ bzw. bei $x \rightarrow -\infty$ zwei verschiedenen endlichen Grenzwerten b bzw. a (“Sättigungswerte”) annähert und infolgedessen nicht rational sein kann, ist die (im Fall $b > a$ streng wachsende) **logistische Funktion**

$$f(x) = \frac{a + be^{cx}}{1 + e^{cx}} = b - \frac{b-a}{e^{cx} + 1} = a + \frac{b-a}{1 + e^{-cx}},$$

wobei der Parameter $c > 0$ reguliert, wie schnell sich die Funktionswerte $f(x)$ dem Wert b bei $x \rightarrow \infty$ bzw. dem Wert a bei $x \rightarrow -\infty$ annähern (in jedem Fall mit exponentieller Geschwindigkeit, je größer c desto schneller). Dieses Grenzverhalten liest man aus den angegebenen Darstellungen der Funktion unmittelbar ab, da $e^{cx} = (e^c)^x$ beliebig groß wird für große Werte der Variablen x und e^{-cx} beliebig groß für betragsgroße negative Werte von x .



Die Abbildung zeigt die Graphen der logistischen Funktion mit $b - a = 1$ und $c = 1$ sowie den durch Stauchung der Rechtswerte mit Faktor $\frac{1}{3}$ daraus entstehenden Graphen der logistischen Funktion zum Parameter $c = 3$. Die Graphen nähern sich ihren horizontalen Asymptotengeraden “exponentiell schnell”, so dass sie schon für nicht sehr große Werte von $|x|$ im Rahmen der Rechengenauigkeit nicht mehr von den Asymptotengeraden zu unterscheiden sind. Die Annäherung an die Asymptoten erfolgt hier wesentlich schneller als z.B. bei einer Hyperbel wie der Graph von $\frac{1}{x}$. Wir haben hier die logistische Funktion so definiert, dass sie als Wert $f(0) = \frac{1}{2}(a+b)$ an der Stelle $x = 0$ das arithmetische Mittel ihrer Sättigungswerte hat; der Graph ist dann punktsymmetrisch zum Punkt $(0, \frac{a+b}{2})$. Durch horizontale Verschiebung kann man an der Stelle $x = 0$ beliebige andere Werte zwischen a und b erreichen. Auch die verschobenen Funktionen $x \mapsto f(x-x_0)$, die den Wert $\frac{1}{2}(a+b)$ an der Stelle $x = x_0$ annehmen, heißen “logistische Funktionen”.

Eine andere Darstellung der logistischen Funktionen beruht auf den *Hyperbelfunktionen* $\cosh x = \frac{1}{2}(e^x + e^{-x})$ und $\sinh x = \frac{1}{2}(e^x - e^{-x})$, die ihren Namen tragen, weil sie die Hyperbelgleichung $(\cosh x)^2 - (\sinh x)^2 = 1$ erfüllen. Der Quotient des hyperbolischen Sinus und Kosinus heißt *hyperbolischer Tangens* $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ und stimmt überein mit der logistischen Funktion zu den speziellen Parametern $a = -1$, $b = 1$, $c = 2$. Die oben definierte logistische Funktion f zu allgemeinen Parametern a, b, c lässt sich dann durch den hyperbolischen Tangens ausdrücken in der Form $f(x) = \frac{a+b}{2} + \frac{b-a}{2} \tanh(\frac{1}{2}cx)$. Die Hyperbelfunktionen sind durch Rechenterme gegeben, welche nur die Grundfunktion e^x enthalten (dazu schreibe e^{-x} als $1/e^x$). Also handelt es sich bei den Hyperbelfunktionen um elementare Funktionen. ■

Durch algebraische Rechenoperationen und Verkettung kann man eine unübersehbare Vielfalt von elementaren Funktionen erzeugen. Zum Beispiel sind Linearkombinationen von allgemeinen Potenzfunktionen $c_1x^{s_1} + c_2x^{s_2} + \dots + c_nx^{s_n}$ bzw. von Exponentialfunktionen $c_1a_1^x + c_2a_2^x + \dots + c_na_n^x$ elementare Funktionen auf $\mathbb{R}_{>0}$ bzw. auf \mathbb{R} , und Quotienten von solchen Funktionen sind elementar auf Definitionsbereichen ohne Nennernullstellen. Und man kann Wurzeln oder Logarithmen von solchen Funktionen bilden, sogar verschachtelte Wurzelausdrücke, davon dann wieder Linearkombinationen oder Produkte oder Quotienten \dots . Es lassen sich elementare Rechenterme von solcher Komplexität ausdenken, dass sie nicht als Formel auf eine Seite gedruckt werden können. Da man aus elementaren Funktionen durch algebraische Rechenoperationen und Verkettung immer neue elementare Funktionen zusammensetzen kann, gibt es im Prinzip keine Grenze für die Komplexität. Eine Schwierigkeit bei sehr komplexen Rechentermen ist allerdings die Bestimmung eines sinnvollen bzw. des maximalen Definitionsbereichs — es kann ja durchaus sein, dass der formale Rechenterm einen leeren Definitionsbereich hat. Die konkreten elementaren Funktionen, die in der Wirtschaftsmathematik auftreten, sind aber meist relativ einfach aufgebaut, wie z.B. Polynome von kleinem Grad oder Quotienten von Polynomen kleinen Grades oder wie die logistischen Funktionen in Beispiel (3) oben.

Aber selbst in der Wirtschaftsmathematik kommt man nicht ganz mit elementaren Funktionen aus. Zum Beispiel ist in einer Situation, wo eine ökonomische Variable $y = f(x)$ als elementare Funktion einer anderen Variablen x dargestellt ist und x auch umgekehrt als Funktion von y auffassbar ist, also die Umkehrfunktion f^{-1} existiert, die Variable x nicht unbedingt durch einen Rechenterm in der Variablen y ausdrückbar. *Die Umkehrfunktion einer elementaren Funktion ist nämlich im Allgemeinen nichtelementar* (sofern sie überhaupt existiert). Mit anderen Worten: Wenn die linke Seite der Gleichung $f(x) = y$ ein elementarer Rechenterm ist und die Gleichung für x, y aus gewissen Intervallen eindeutig lösbar ist, so kann man doch die Lösung x im Allgemeinen nicht durch eine elementare Formel aus y berechnen. Das ist schon bei Polynomfunktionen so, z.B. bei der (auf \mathbb{R} umkehrbaren) Funktion $p(x) = x^5 + x$. Eine Ausnahme bilden nur Polynome von kleinem Grad ≤ 4 , weil es dafür Lösungsformeln gibt. Zum Beispiel ist eine quadratische Funktion $q(x) = ax^2 + bx + c$ auf dem Intervall $]-\frac{b}{2a}, \infty[$ umkehrbar, und die Lösungsformel für quadratische Gleichungen gibt den elementaren Ausdruck $q^{-1}(y) = \frac{-b}{2a} + \frac{1}{2|a|} \sqrt{d + 4ay}$ für die Umkehrfunktion ($d = b^2 - 4ac$ die Diskriminante und $y \in \mathbb{R}$ mit $d + 4ay > 0$). Bei anderer Vorzeichenwahl der Wurzel erhält man die Umkehrfunktion zu q auf $]-\infty, \frac{-b}{2a}[$.

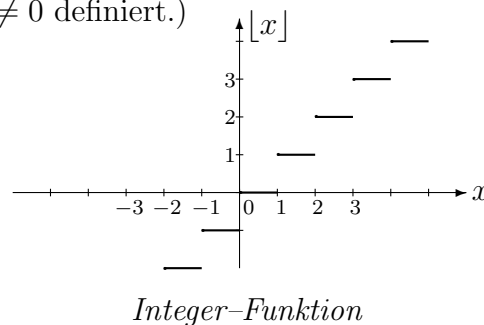
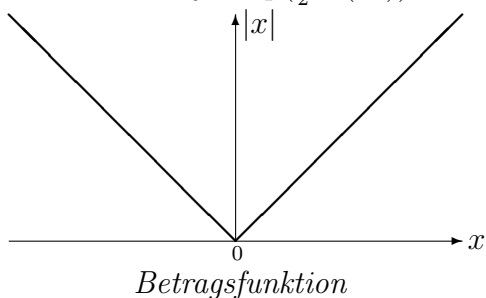
Oft hat man es in der Wirtschaftsmathematik auch mit Funktionen zu tun, die zwar nicht elementar im Sinne der Definition sind, aber aus verschiedenen elementaren Funktionen zusammengesetzt sind. Allgemein führt das **Aneinandersetzen von Funktionen** $f: \mathbb{R} \supset I \rightarrow \mathbb{R}$ und $g: \mathbb{R} \supset J \rightarrow \mathbb{R}$ mit disjunkten Definitionsbereichen I und J zu einer "Vereinigungsfunktion" $f \cup g$, welche die Vereinigung $I \cup J$ als Definitionsbereich hat und an Stellen $x \in I$ dieselben Werte wie f , an Stellen $x \in J$ aber dieselben Werte wie g . Der Graph von $f \cup g$ entsteht dann also, indem man den Graphen von f über I mit dem Graphen von g über J zusammensetzt. Wir denken dabei z.B. an nichtüberlappende Intervalle $I = [x_0, x_1[$, $J = [x_1, x_2]$, die einen Randpunkt x_1 gemeinsam haben; dann ist $I \cup J$ das Intervall $[x_0, x_2]$. Enthalten beide Intervalle den gemeinsamen Randpunkt x_1 , in dem sie aneinandergesetzt werden, so hat man bei der Definition des Funktionswertes von $f \cup g$ an der Stelle x_1 die Wahl zwischen $f(x_1)$ und $g(x_1)$. Ob man sich an dieser Stelle für $f(x_1)$ oder $g(x_1)$ oder eine andere Zahl als Wert von $f \cup g$ entscheidet, ist in der Praxis aber irrelevant. Man muss jedoch im Fall $f(x_1) \neq g(x_1)$ beachten, dass die zusammengesetzte Funktion dann ihre Werte bei Überschreiten von x_1 sprunghaft ändert.

Wenn nun f und g elementare Funktionen sind, so nennen wir die daraus zusammengesetzte Funktion $f \cup g$ eine **stückweise elementare Funktion**. Dies ist dann im Allgemeinen keine elementare Funktion auf $I \cup J$, weil es evtl. keinen elementaren Rechterm gibt, der die Funktionswerte für *alle* $x \in I \cup J$ berechnet, sondern eben nur zwei *verschiedene* Rechenterme, von denen einer die Funktionswerte über dem Teilbereich I des Definitionsbereichs berechnet, der andere die Funktionswerte über dem Teilbereich J . Die zusammengesetzte Funktion $f \cup g$ kann auch in der Tat an einer Stelle, an der die Intervalle I und J aneinanderstoßen, ein Verhalten haben, das bei elementaren Funktionen ausgeschlossen ist, z.B. eine ‘‘Sprungstelle’’ oder eine ‘‘Knickstelle’’ im Graphen.

Die Konstruktion der ‘‘Vereinigungsfunktion’’ geht analog, wenn mehr als zwei (endlich oder unendlich viele) Funktionen f_1, f_2, f_3, \dots auf Intervallen I_1, I_2, I_3, \dots gegeben sind, die paarweise disjunkt sind (d.h. keine zwei davon haben ein Element gemeinsam). Die Vereinigungsfunktion f ist dann definiert auf der Vereinigung $I = I_1 \cup I_2 \cup I_3 \cup \dots$ dieser Intervalle und hat als Werte $f(x) := f_k(x)$, wenn das Argument x im Teilintervall I_k liegt. Sind alle f_k elementare Funktionen, so heißt f wieder eine stückweise elementare Funktion. Die Analyse einer stückweise elementaren Funktion läuft im Wesentlichen auf die Behandlung der verschiedenen elementaren Funktionen $f_k : I_k \rightarrow \mathbb{R}$ hinaus, aus denen sie zusammengesetzt ist, wobei an den Stellen, an denen zwei Teilintervalle I_k zusammenstoßen, unter Umständen noch gesonderte Betrachtungen erforderlich sind.

BEISPIELE (stückweise elementare Funktionen):

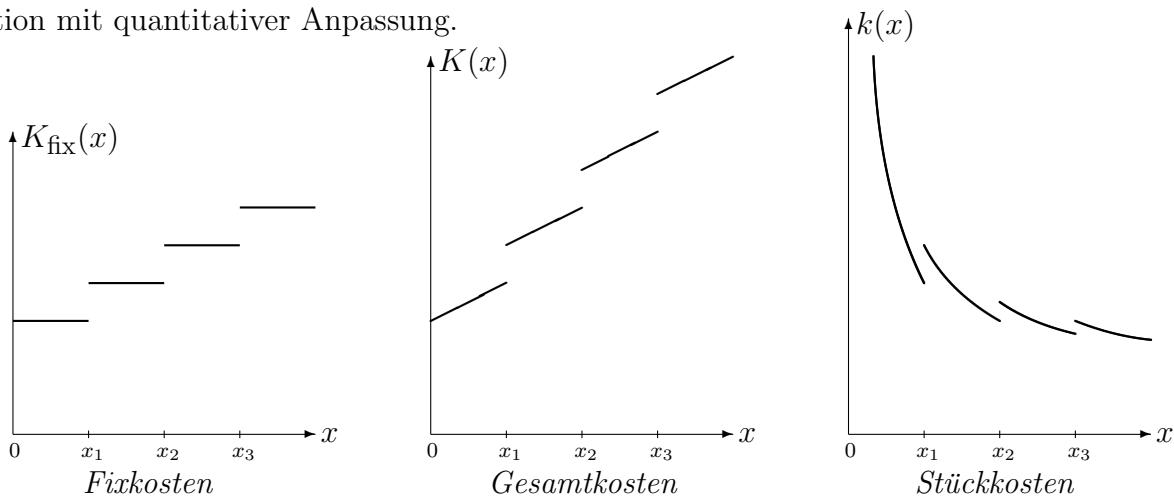
(1) Die **Betragsfunktion** $f(x) = |x|$ ist eine **stückweise lineare Funktion** auf \mathbb{R} , d.h. aus linearen Funktionen zusammengesetzt, nämlich aus $g(x) = x$ auf $[0, \infty[$ und aus $f(x) = -x$ auf $] -\infty, 0[$. Die Betragsfunktion ist aber keine elementare Funktion auf \mathbb{R} , es gibt also keine elementaren Rechterm, der für alle $x \in \mathbb{R}$ den Wert $|x|$ hat. (Das kann man an der Knickstelle des Graphen erkennen: In der Differentialrechnung 4.3 wird gezeigt, dass die Graphen elementarer Funktionen überall Tangenten und nirgends Knickstellen haben. Der Term $\sqrt{x^2}$ definiert die Betragsfunktion nur auf $\mathbb{R}_{\neq 0}$ als elementare Funktion, weil $\sqrt{}$ nur auf $\mathbb{R}_{>0}$ elementar ist. Durch die Grundfunktionen ausgedrückt lautet der Term ja $\exp(\frac{1}{2} \ln(x^2))$ und ist nur für $x \neq 0$ definiert.)



(2) Die **Integer-Funktion** (*Größtes-Ganzes-Funktion*, ‘‘Gauß-Klammer’’) $f(x) = [x]$, die jede reelle Zahl zu einer ganzen Zahl abrundet, ist eine **Treppenfunktion**, d.h. aus konstanten Funktionen zusammengesetzt. Hier ist \mathbb{R} zerlegt in die Intervalle $[k, k+1[$ mit $k \in \mathbb{Z}$, und für $x \in [k, k+1[$ ist die Funktionsvorschrift einfach $[x] := k$. Die analoge Aufrundungsfunktion ist definiert durch $[x] := k+1$ für $x \in]k, k+1]$ mit $k \in \mathbb{Z}$. Ihr Graph entsteht aus dem der Integer-Funktion durch vertikale Verschiebung um 1 nach oben mit Ausnahme der über ganzzahligen Argumenten k gelegenen Punkte, in denen $[k] = k = [k]$ gilt. Beide Funktion haben Sprungstellen mit Sprunghöhe 1 in den ganzen Zahlen. Mit diesen Funktionen kann man z.B. eine Kapitalentwicklung bei gleichbleibenden Ratenzahlungen ohne Verzinsung modellieren. Für allgemeinere Kontenentwicklungen hat man Treppenfunktionen mit Sprüngen in evtl. variierenden Höhen und Abständen.

(3) Ein ökonomisches Beispiel, in dem stückweise elementare Funktionen mit Sprungstellen auftreten, ist eine **Produktion mit quantitativer Anpassung** der Produktionsanlagen. Hier wird, wenn der Produktions-Output x (Mengeneinheiten) bestimmte Schwellenwerte $x_1 < x_2 < \dots$ übersteigt, jeweils ein neues Produktionsaggregat in Betrieb genommen. Das führt zu einem sprunghaften Anstieg der *Fixkosten* $K_{\text{fix}}(x)$, die sich additiv zusammensetzen aus den konstanten *Fixgrundkosten* für die Unterhaltung der Produktionsstätte und den Fixkosten für alle Produktionsaggregate, die zur Erzeugung des Outputs x betrieben werden müssen. Die Fixkostenfunktion hat also Spünge bei den Schwellenwerten x_i , bei denen jeweils ein Aggregat zugeschaltet werden muss.

Zu den Fixkosten kommen noch die *variablen Kosten* $K_{\text{var}}(x)$, welche die zur Erzeugung des Outputs x benötigten Rohstoffkosten und Energiekosten beinhalten und in einfachen Fällen proportional zu x sind, also $K_{\text{var}}(x) = \kappa x$ mit einem Proportionalitätsfaktor $\kappa \in \mathbb{R}_{>0}$ (den sog. *stückvariablen Kosten*). Die *Gesamtkosten* $K(x) = K_{\text{fix}}(x) + K_{\text{var}}(x)$ sind dann eine stückweise lineare Funktion, die in den Intervallen $]x_i, x_{i+1}[$ zwischen den Schwellenwerten jeweils affin linear mit Steigung κ ist (aber mit von Intervall zu Intervall verschiedenen Achsenabschnitten) und in den Schwellenwerten Sprungstellen mit derselben Sprunghöhe wie die Fixkostenfunktion hat. Als *Stückkostenfunktion* $k(x) = K(x)/x = \frac{1}{x}K_{\text{fix}}(x) + \kappa$ ergibt sich hier eine stückweise rationale Funktion mit Sprungstellen abnehmender Höhe (weil die Fixkostensprünge bei größeren Outputs die Stückkosten weniger erhöhen) und monoton abnehmendem Funktionsverlauf dazwischen (weil die stückvariablen Kosten konstant sind und die Fixkosten sich bei größeren Outputs weniger auf die Stückkosten auswirken). Die Abbildungen zeigen typische Funktionsverläufe bei Produktion mit quantitativer Anpassung.



In realistischen Situationen sind die stückvariablen Kosten für ein Produktionsaggregat oft nicht konstant, sondern sinken mit wachsender Auslastung. Die Gesamtkosten zeigen dann zwischen den Schwellenwerten keinen linearen, sondern einen degressiv wachsenden Verlauf (die Steigung nimmt mit wachsendem x ab). Es ist intuitiv klar, dass man dann kein neues Produktionsaggregat in Betrieb nehmen wird, so lange nicht alle bereits betriebenen Aggregate voll ausgelastet sind. Anders ist die Situation, wenn die stückvariablen Kosten eines Aggregats mit wachsender Auslastung steigen, weil etwa die Anlage bei starker Auslastung ungünstiger arbeitet. Dann wird es für einen gewünschten Output unter Umständen trotz höherer Fixkosten günstiger sein, zusätzliche Aggregate in Betrieb zu nehmen und alle betriebenen Aggregate in einem günstigen Auslastungsbereich zu halten, statt den Output mit weniger Aggregaten in einem ungünstigen Auslastungsbereich zu produzieren. Die kostengünstigste Produktionsweise in einem solchen Fall zu bestimmen, kann eine recht anspruchsvolle wirtschaftsmathematische Optimierungsaufgabe sein. ■