Mikro 2, Sommersemester 2003

Karl Dietrich
dietrich@vwl.uni-hannover.de

Version 1.6 vom 23. Juni 2003

Gliederung

1. Marktnachfrage und Marktgleichgewicht
2. Produktion
3. Gewinnmaximierung
4. Kostenminimierung
5. Kostenkurven
6. Angebot einer Firma
7. Marktangebot einer Industrie
8. Monopol
9. Monopolistisches Verhalten
10. Faktormärkte
11. Oligopol
12. Spieltheorie

Marktnachfrage

Wir behandeln heute

- die ersten Abschnitte aus Kapitel 15 — Marktnachfrage — bis zu den Elastizitäten (15.1–15.4)
- Teile aus Kapitel 16 — Gleichgewicht auf einem Markt.

Der erste Punkt ist im wesentlichen als Anknüpfung an Mikro 1 vom letzten Semester gedacht.

Einflüsse auf die individuelle Nachfrage

Die individuelle Nachfragekurve eines Verbrauchers nach einem Gut der Sorte 1 beschreibt seine optimale Wahl der Gütermenge x_1

- bei alternativen Preises des Gutes p_1,
- den alternativen Preisen anderer Güter — wir unterstellen jedoch nur ein weiteres Gut — die Sorte 2 — p_2,
- alternativen Konsumsummen (oder Einkommensniveaus) m.

Die Präferenzen des Verbrauchers finden ihren Ausdruck in der Lage und Gestalt der Indifferenzkurven und üben auch einen wesentlichen Einfluss aus. Dieser Einfluss wird jedoch bei der Betrachtung der individuellen Nachfrage konstant gehalten.

Allgemeine und direkte Nachfragekurve

Die allgemeine Nachfragekurve eines Verbrauchers i für Gut 1 ist eine Funktion (mit mehreren Variablen), die jeder Preise-Einkommens-Konstellation seine optimale Verbrauchsmenge x_{i1} zuordnet:

$$x_{i1} = x_{i1}(p_1, p_2, m_i)$$

Seine direkte Nachfragekurve ergibt sich, wenn seine optimale Verbrauchsmenge bei konstantem Preis des zweiten Gutes $p_2 = \bar{p}_2$ und konstantem Einkommen $m_i = \bar{m}_i$ betrachtet wird:

$$d_{i1} = d_{i1}(p_1) = x_{i1}(p_1, \bar{p}_2, \bar{m}_i)$$

Allgemeine Marktnachfrage

Fasst man die nachgefragten Mengen x_{i1} der Verbraucher $i = 1, 2, \ldots n$ zusammen, entsteht die allgemeine Marktnachfragekurve X_1.

$$X_1 = X_1(p_1, p_2, m_1, \ldots m_n) = \sum_{i=1}^{n} x_{i1}(p_1, p_2, m_i)$$

Die Eigenschaften der individuellen (allgemeinen) Nachfragekurven vererben sich an das Aggregat X_1.

Dazu kommt, dass X_1 durch die Anzahl der Nachfrager beeinflusst wird.

Marktnachfrage

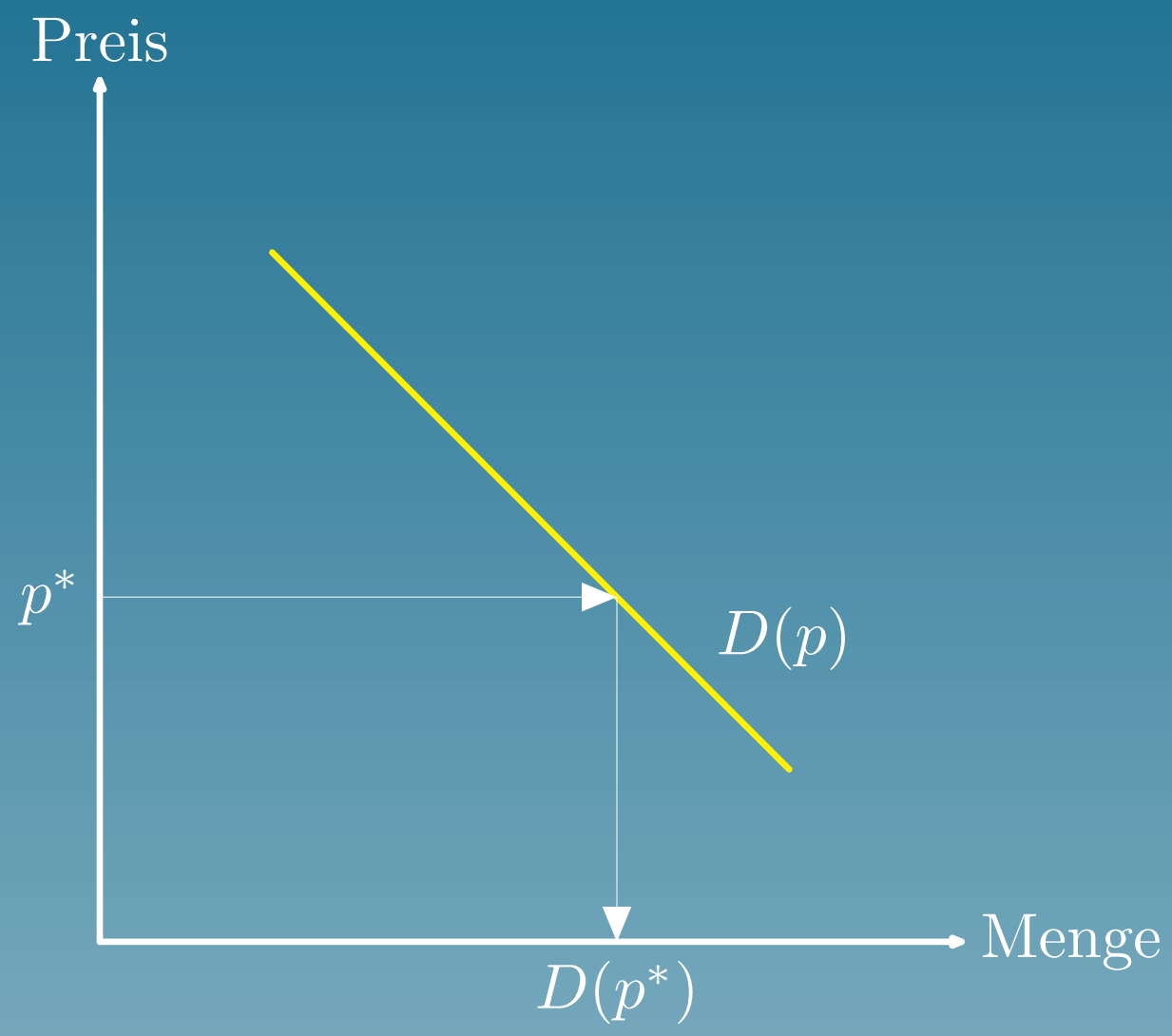
Werden die direkten Nachfragekurven d_{i1} für Gut 1 aller Verbraucher aggregiert, ergibt sich die Marktnachfrage

$$D_1(p_1) = \sum_{i=1}^{n} d_{i1}(p_1)$$

Meist wird nur ein einziges Gut betrachtet, dann lässt man den Index 1 weg und bezeichnet mit $q = D(p)$ die Marktnachfrage nach dem Gut in Abhängigkeit von seinem Preis.

Aber nach wie vor sind alle zuvor genannten Einflüsse auf die Marktnachfrage vorhanden — sie werden lediglich zum Zwecke der Vereinfachung vernachlässigt.

Marktnachfragekurve



Intensive und extensive Nachfrageänderungen

Wenn sich der Preis ändert, werden die Nachfrager darauf reagieren. Sinkt der Preis, dann werden die bereits am Markt befindlichen Verbraucher ihre Käufe ausdehnen und zugleich werden auch neue Käufer, denen das Gut bisher zu teuer war, hinzukommen. Daher wird zwischen in- und extensiven Nachfrageänderungen unterschieden.

Die durch Vermehrung oder Verminderung der *Anzahl* der Nachfrager hervorgerufene Nachfrageänderung heißt extensiv.

Die durch Vergrößerung oder Verkleinerung der individuellen Nachfrage bei gleicher Anzahl der Verbraucher hervorgerufene Nachfrageänderung heißt intensiv.

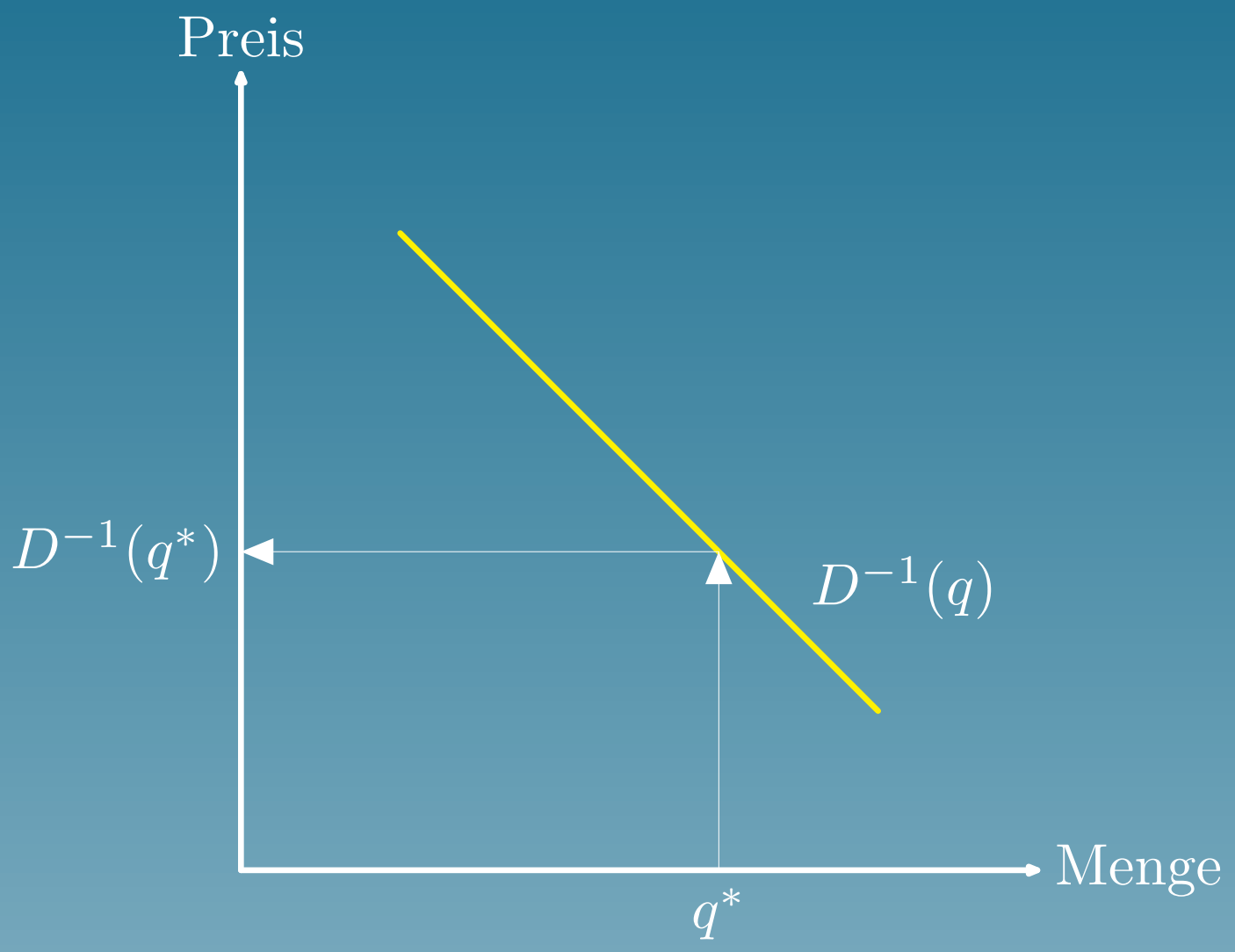
Inverse Marktnachfragekurve

Die Marktnachfragekurve ordnet jedem Preis p eine Gütermenge $q = D(p)$ zu. Diese Gütermenge wird auf der Rechts-Achse dargestellt.

Die inverse Marktnachfragekurve dreht dagegen die Kausalrichtung um: jeder Gütermenge q wird jener Reservationspreis $D^{-1}(q)$ zugeordnet, der alle Verbraucher veranlasst gerade diese Menge als Summe ihrer individuellen Haushaltsoptima zu planen. $D^{-1}(q)$ ist jener *Preis*, der bei den Verbrauchern eine Gesamtnachfrage in Höhe von x hervorruft. Dieser Preis ist ein Maß für die Zahlungsbereitschaft der Verbraucher bei der Gesamtmenge q.

Im mathematischen Sprachgebrauch ist $D^{-1}(q)$ die Umkehrfunktion von D, wird aber auch dort als inverse Funktion bezeichnet.

Graph der inversen Marktnachfragekurve

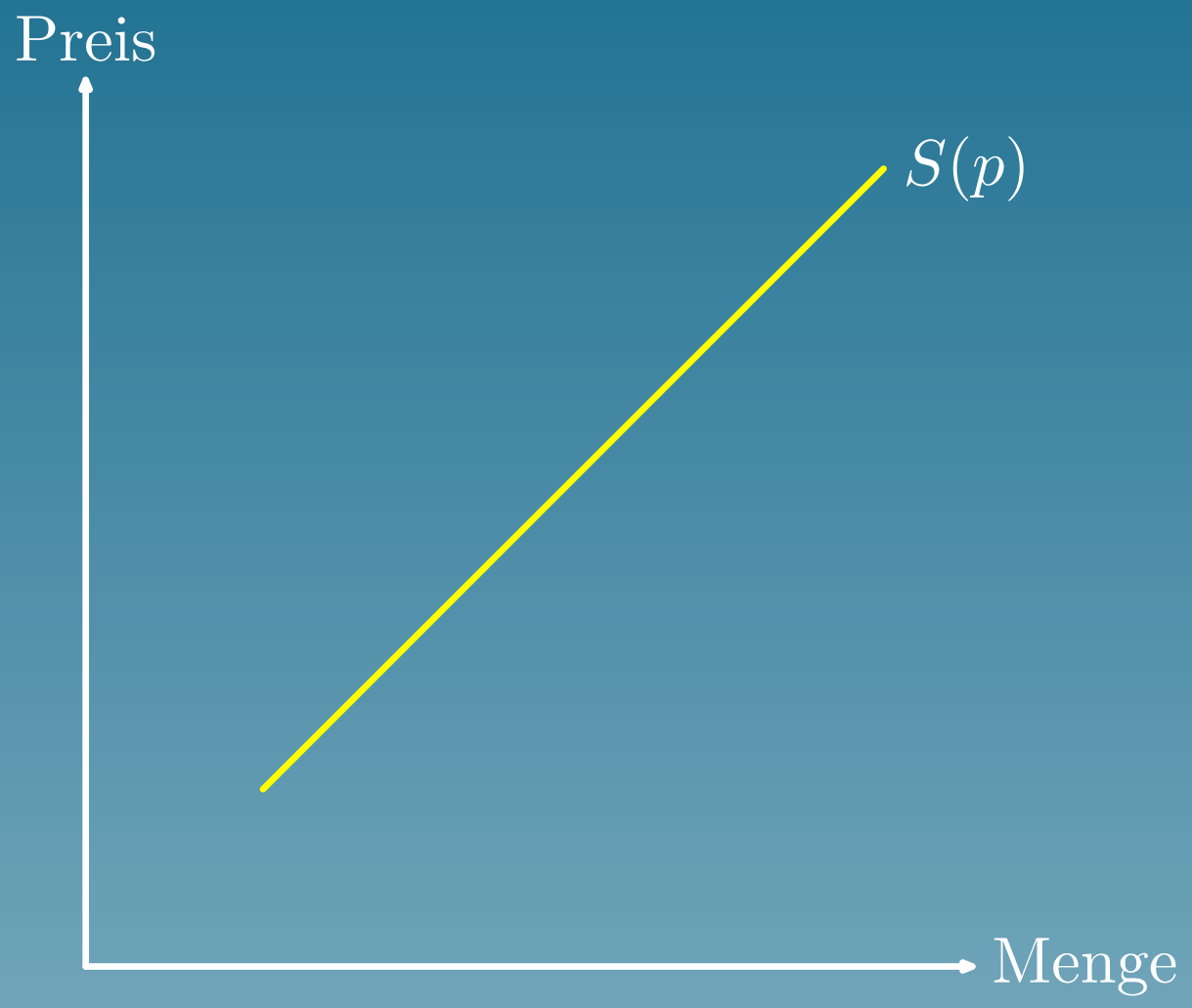


Marktangebot

Der Marktnachfrage wird nun das Angebot gegenüber gestellt. In den nächsten Wochen wird das Verhalten der Anbieter genauer betrachtet. Für heute muss es genügen, eine Marktangebotskurve zu kennen. Es soll das Zusammentreffen von Anbietern und Nachfragern auf einem isolierten Markt betrachtet werden.

Die Angebotskurve $S(p)$ ordnet jedem Preis p die Menge q des Gutes zu, welche alle Anbieter gemeinsam bereit sind zu diesem Preis zu verkaufen. Diese Kurve ist im Preis-Mengen-Diagramm ansteigend, denn je höher der Preis, desto größer wird auch die Verkaufsbereitschaft sein, vorausgesetzt, die potenziellen Anbieter sind im Besitz ein hinreichend großen Menge des Gutes.

Marktangebotskurve



Konkurrenzmarkt

Ein Konkurrenzmarkt ist dadurch gekennzeichnet, dass alle Marktteilnehmer (also aller Anbieter und Nachfrager) den Marktpreis als eine Größe ansehen, die sie *nicht* beeinflussen können.

Man spricht auch vom Verhalten der Mengenanpassung aller Marktteilnehmer, weil sich ihre Planungen bezüglich der angebotenen und nachgefragten Menge des Gutes nach dem Preis richten.

Auch wenn der Marktpreis durch einen einzelnen Marktteilnehmer nicht spürbar beeinflusst wird, so kommt durch das Zusammenwirken aller ein bestimmter Marktpreis oder Konkurrenzpreis zustande.

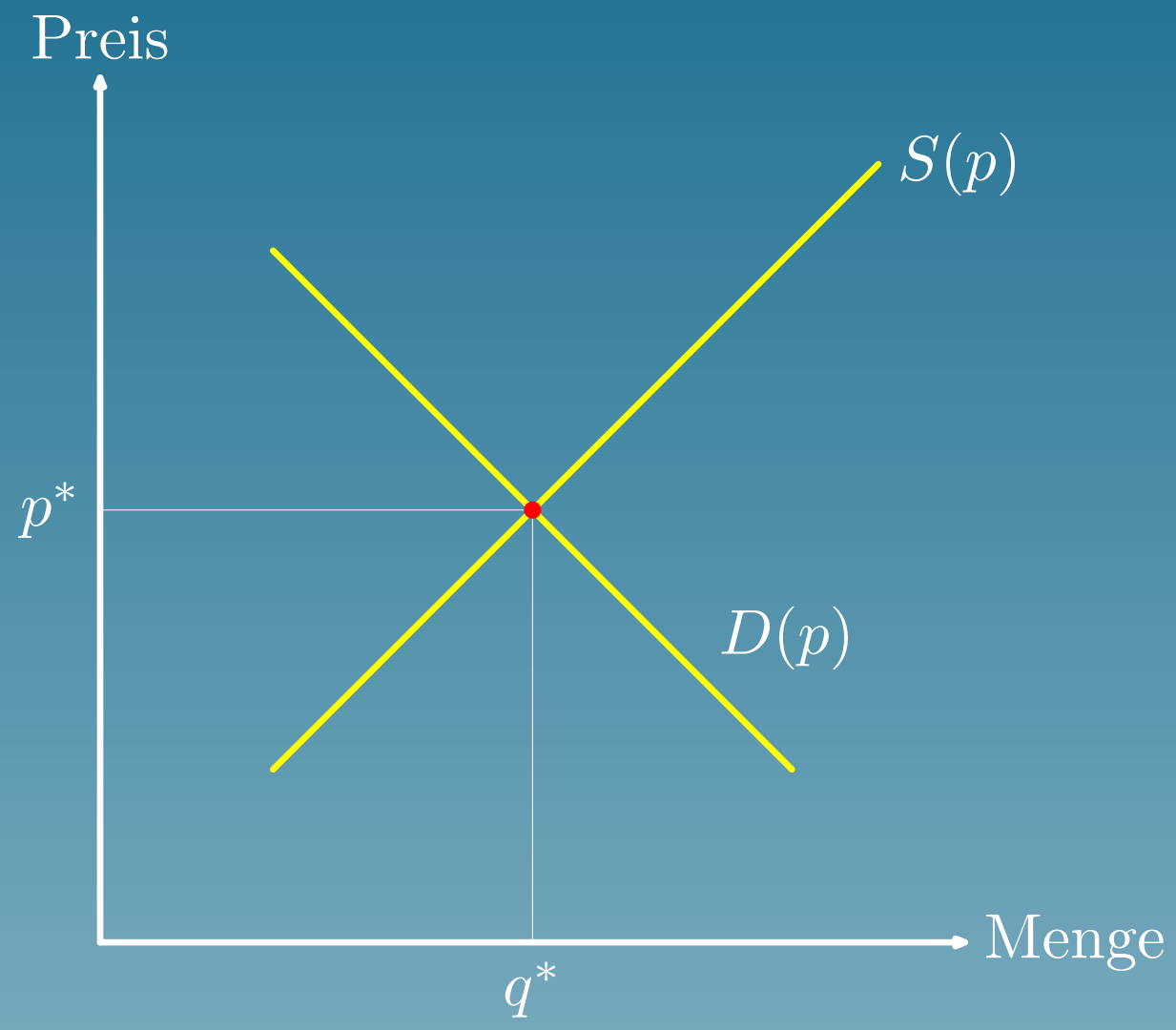
Marktgleichgewicht

Der Gleichgewichtspreis auf einem (anonymen) Konkurrenzmarkt kommt durch einen Wettbewerbsmechanismus zustande, der in seinen Details nicht zu beschreiben ist. Wohl aber das Resultat:

Im Gleichgewicht müssen angebotene und nachgefragte Gütermenge gleich sein. Andernfalls hätten noch manche Marktteilnehmer einen Anreiz neu zu verhandeln. Beim Gleichgewichtspreis geben alle Marktteilnehmer „ihre besten Antworten“ (so wie bei jedem anderen Preis auch), und zugleich sind diese Antworten auch durchführbar: kein Nachfrager muss verzichten und kein Anbieter behält einen Teil seines Besitzes, den er eigentlich veräußern möchte.

Die mathematische Gleichgewichtsbedingung lautet $D(p^*) = S(p^*)$.

Gleichgewicht = Schnittpunkt von Angebots- und Nachfragekurve



Der Rest von Varians Kapitel 16 wurde nicht besprochen, ist aber gleichwohl relevant.

Produktion

Wir behandeln das mit *Technologie* überschriebene 17. Kapitel von Varians Lehrbuch (in der Fünften Auflage ist dies Kapitel 18). Ganz bewusst ist hier ein anderer Titel gewählt.

Unter Produktion wird die Herstellung von Gütern (und Diensten) zum Zwecke der Bedarfsdeckung verstanden. Sie ist der „allgemeine Stoffwechsel zwischen Mensch und Natur“.

Die Produktion findet jedoch immer im Rahmen bestimmter gesellschaftlicher Gegebenheiten statt.

Produktion in industrialisierten Marktwirtschaften

Die Produktion in den *industrialisierten Marktwirtschaften* ist gekennzeichnet durch

- hochgradige *Arbeitsteilung*
- innerhalb und zwischen *Firmen*
- die sich fast ausnahmlos in *Privatbesitz* befinden.

Die Institution Firma ist eine Erfindung des Kapitalismus. Ihr Innenleben ist der hauptsächliche Gegenstand der BWL, aber auch die (moderne) Mikroökonomik hat Einiges an Erklärung dazu einzubringen.

Firma als Black Box

Wir betrachten die Firma — den ökonomischen Ort der Produktion — als eine Black Box, abstrahieren also von ihrem Innenleben. Wir registrieren, was hineingeht und (wie ein früherer Bundeskanzler zu sagen pflegte) „was hinten rauskommt“.

Alles, was die Firma zum Produzieren verwendet, heißt Input oder Produktionsfaktor, und das Resultat der Produktion heißt Output oder Produktionsergebnis.

Unterteilung des Inputs

Der Input lässt sich weiter unterteilen in die Gruppen

- Arbeitsleistung,
- Land,
- (physisches) Kapital und
- Rohstoffe.

Eine andere Unterscheidung — die vor allem bei dynamischen Betrachtungen wesentlich ist — betrifft die Dauerhaftigkeit der Inputs. Ein Kapitalgut, z.B. ein Schweißapparat, kann über einen langen Zeitraum benutzt werden, während die Blechteile, die damit bearbeitet werden, sofort Teil des Outputs sind und nicht mehr anderweitig eingesetzt werden können.

Input als Stromgröße

Wir wollen die Inputs generell als Stromgrößen ansehen. Nicht die Arbeitskräfte mit „Haut und Haaren“, sondern nur ihre Arbeitsleistung pro Tag (oder Woche etc) zählt zum Input. Genauso zählt nicht eine Maschine als solche, sondern ihre Leistungsabgabe während einer bestimmten Periode als Input (beispielsweise 60 Stunden Maschinenlaufzeit pro Woche).

Output

Der Output, das Produktionsergebnis, ist teilweise erwünscht und teilweise unerwünscht. Wir betrachten jedoch nur das Erwünschte, während die nicht erwünschten Kuppelprodukte (Schadstoffe) vernachlässigt werden. Weiter wird unterstellt, dass der Output nur aus einer einzigen Gütersorte besteht, die verkauft werden soll.

Falls die Firma Teile ihres Produktionsergebnisses selbst verwendet (der eigene Stromverbrauch der Stadtwerke), dann kann zwischen Brutto- und Nettoproduktion unterschieden werden. Diese Unterscheidung ist für gesamtwirtschaftliche Betrachtungen von großer Bedeutung. Wir betrachten ausschließlich das Nettoprodukt.

Input-Output Schema

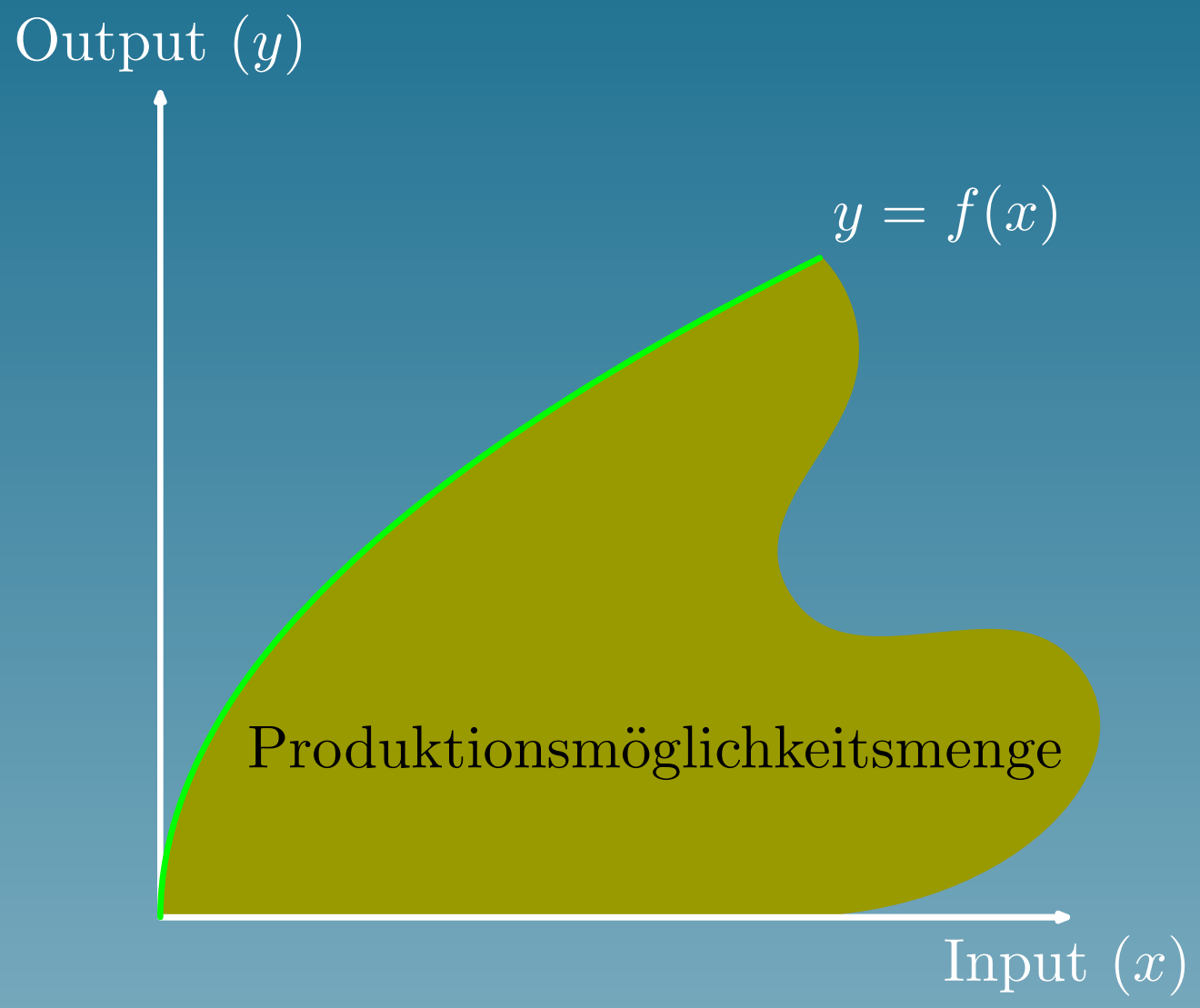
Wird ausschließlich die Beziehung zwischen dem Input(strom) und dem Output(strom) betrachtet, erhält man eine Beschreibung der Technologie.

Produktionsmöglichkeitsmenge

Angenommen, es gibt nur einen Input. Dann kann mit einem bestimmten Quantum davon ganz unterschiedlich verfahren werden. Folglich gibt es einen ganzen Bereich von möglichen Produktionsresultaten, von Nichts (bei völliger Verschwendung) über Wenig (bei ineffizienter Organisation in der Firma oder technisch veralteten Verfahren) bis hin zum Maximum.

Der obere Rand dieser Produktionsmöglichkeitsmenge, der maximale Output y bei einem bestimmten Input x, heißt Produktionsfunktion $y = f(x)$.

Graphik der Produktionsmöglichkeitsmenge



Monotonie der Technologie

Wenn der Output y_1 durch Einsatz von x_1 Einheiten des Inputs hergestellt werden kann, dann soll es auch möglich sein, denselben Output mit mehr Einheiten des Inputs herzustellen. Diese Eigenschaft wird als Monotonie der Technologie bezeichnet.

Die voranstehende Produktionsmöglichkeitsmenge erfüllt offensichtlich diese Eigenschaft: Wenn ein Punkt (x, y) innerhalb der Menge liegt, dann liegt auch jeder Punkt waagerecht rechts davon in der Menge.

Zwei Inputs

So wie in der Haushaltstheorie unterstellt wird, der Verbraucher könne zwei Güter kaufen, wird jetzt unterstellt, es können zwei verschiedene Inputs in der Firma eingesetzt werden. Dadurch eröffnet sich die Möglichkeit unterschiedlich zusammengesetzte Inputbündel zu betrachten und zugleich lassen sich die Produktionsfunktionen noch dreidimensional darstellen.

Wir betrachten also Produktionsfunktionen $y = f(x_1, x_2)$. Dabei ist y der maximale Output, der sich mittels des Inputbündels (x_1, x_2) herstellen lässt.

Spezielle Produktionsfunktionen

Wir haben bereits besondere Funktionstypen als Nutzenfunktionen kennen gelernt:

- Leontief-Funktion,
- Cobb-Douglas-Funktion
- CES-Funktion

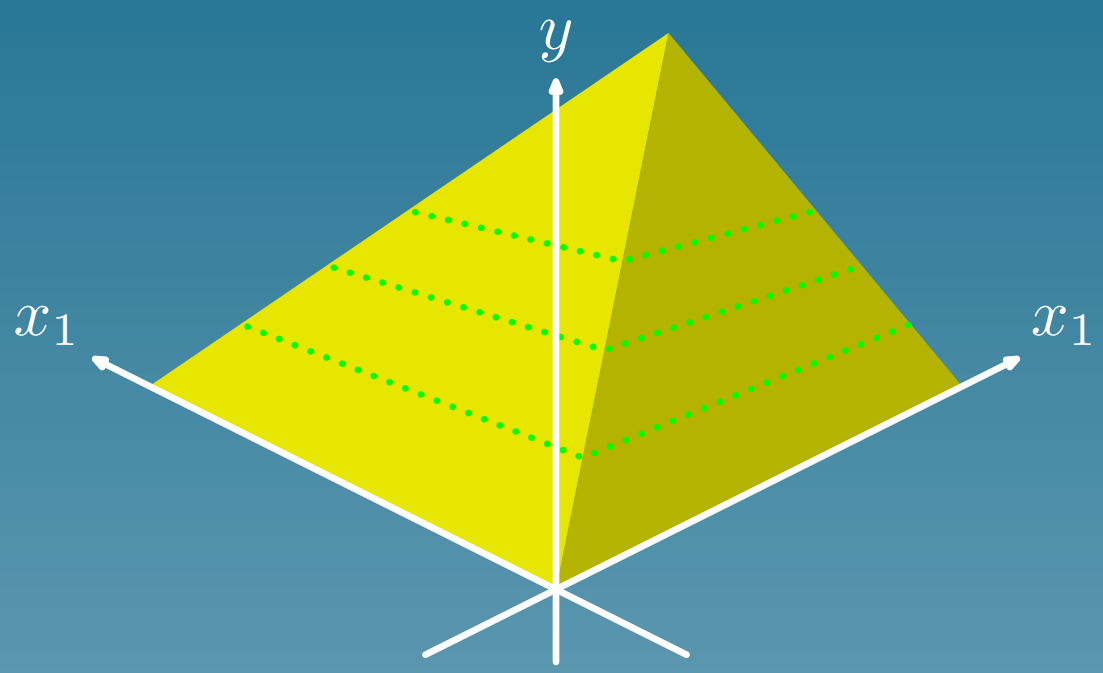
Dieselben Funktionstypen werden auch als Produktionsfunktionen benutzt. Im Unterschied zum nicht beobachtbaren (und nur ordinal messbaren) Nutzen ist jedoch der Output eines Produktionsprozesses sehr wohl beobachtbar und er lässt sich auch kardinal messen (in Stück, kg etc).

Limitationale und substitutionale Produktionsfunktionen

Die Leontief-Funktion $y = \min\{ax_1, bx_2\}$ ist eine sogenannte limitationale Produktionsfunktion. Der Produktionsfaktor 1 kann nicht verringert und durch den zweiten Input substitutiert werden, um das Produktionsniveau aufrecht zu erhalten. Jeder Input limitiert den Output, daher ergibt sich keine Wahlmöglichkeit hinsichtlich der Zusammensetzung des Inputbündels.

Bei den substitutionalen Produktionsfunktionen ist dies jedoch mit mehr oder weniger Aufwand möglich. Die Cobb Douglas Funktion $y = Ax_1^a x_2^b$ ist ein Beispiel für eine solche Produktionsfunktion. Es gibt sehr viele unterschiedlich zusammengesetzte Inputbündel, die alle denselben Output hervorbringen.

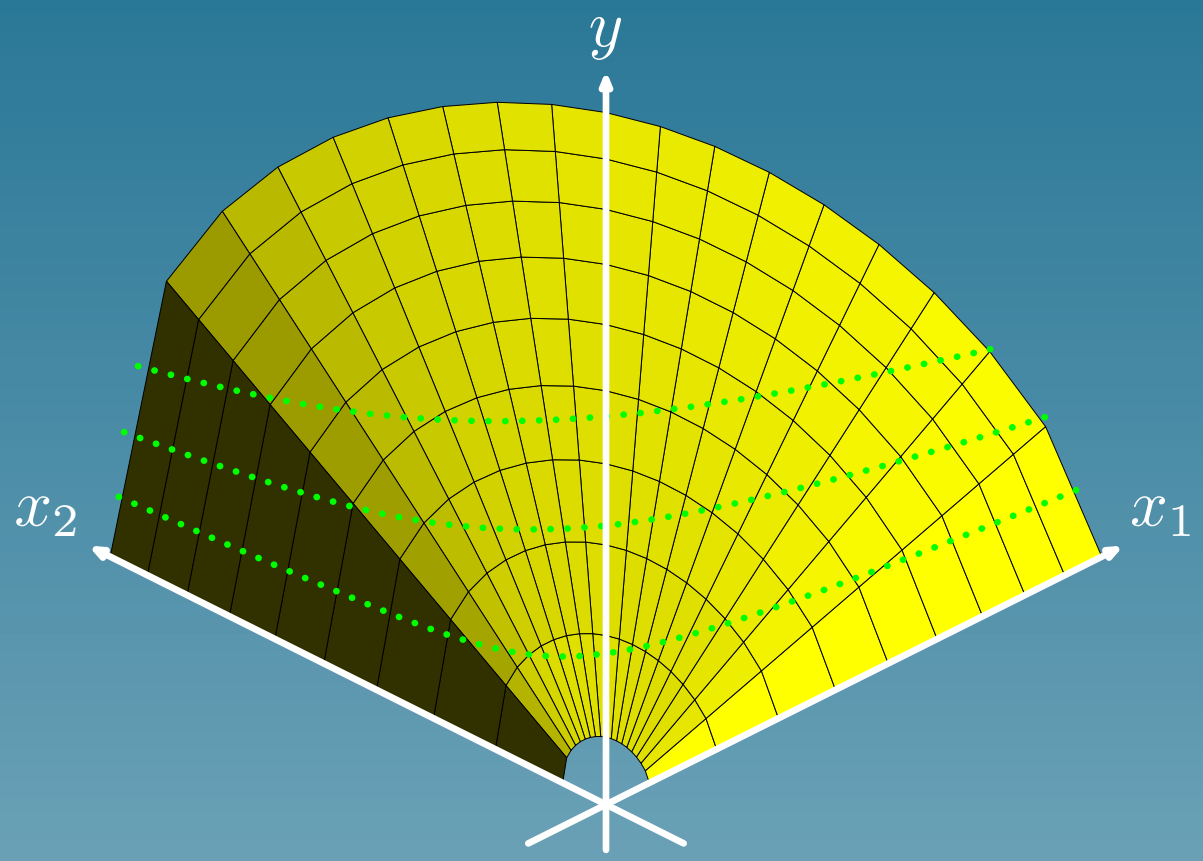
Ertragsgebirge der Leontief Produktionsfunktion



Die äußere Gestalt erinnert an eine schiefe Pyramide.

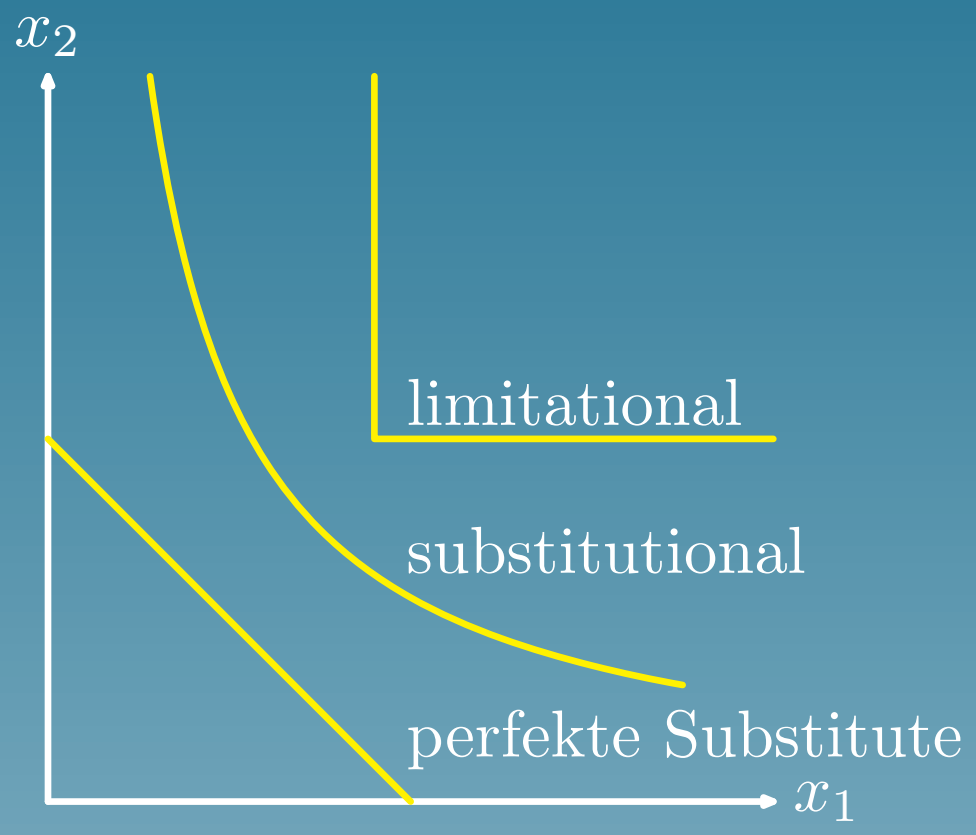
Ihre Isoquanten — Inputkonstellationen, die zum selben Output führen — verlaufen rechtwinklig.

Ertragsgebirge der Cobb-Douglas Produktionsfunktion



Die Gestalt erinnert an eine schiefe Schultüte.

Isoquanten substitutionaler und limitationaler Produktionsfunktionen



Konvexität der Isoquanten

Isoquanten sind Kurven gleichen Outputs. Mathematisch stellen sie Niveau- oder Höhenlinien einer Produktionsfunktion dar. Die Isoquanten sollen eine konvexe Krümmung aufweisen.

Auch die Isoquanten einer Produktionsfunktion mit perfekter Substitution $y = ax_1 + bx_2$, die parallele Strecken darstellen, sind konvex — aber nicht streng konvex.

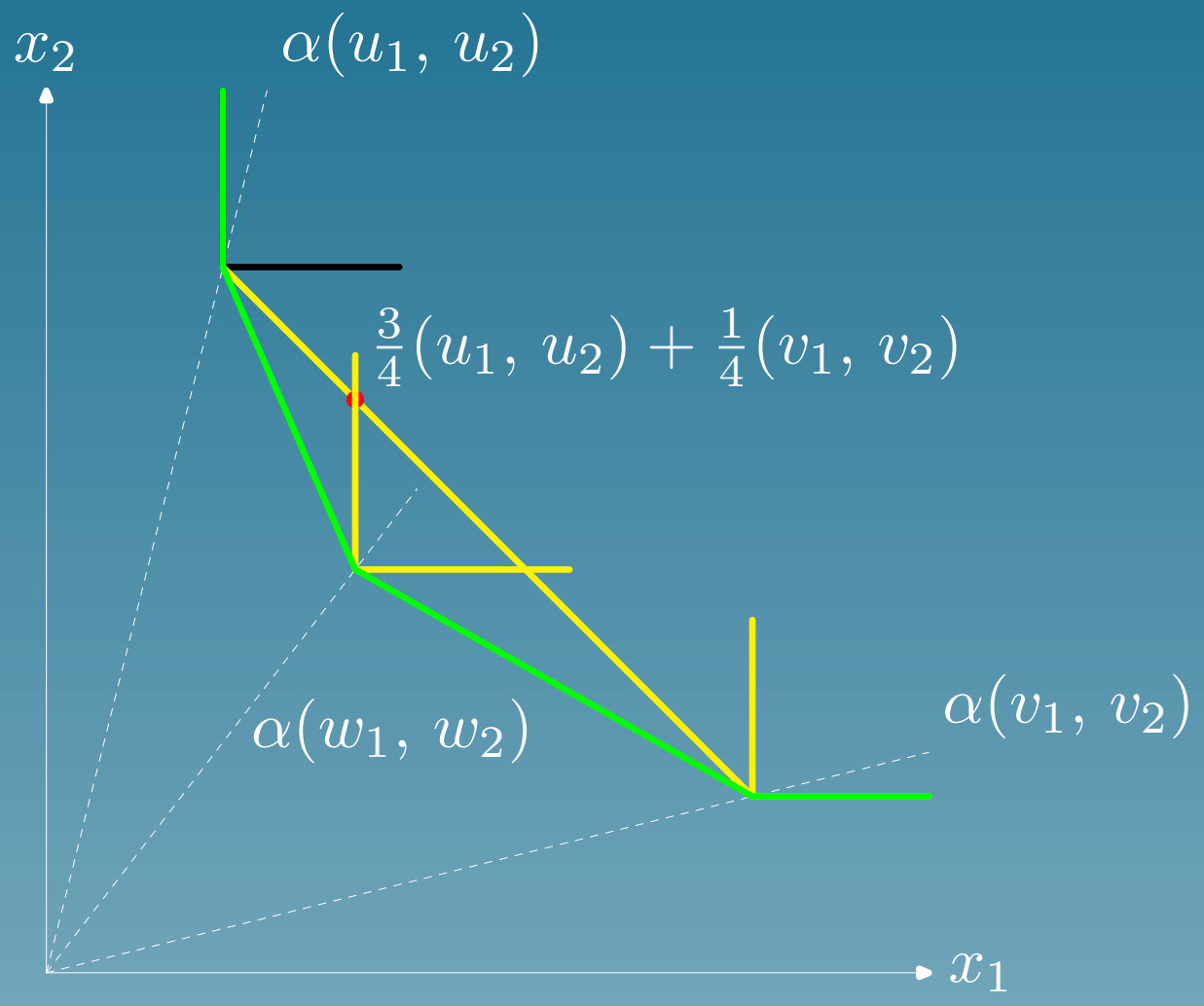
Konvexkombination von Aktivitäten

Limitationale oder Leontiefsche Produktionsfunktionen werden häufig zur Beschreibung einzelner Aktivitäten des Produktionsprozesses benutzt. Aus dieser Aktivitätsanalyse (Operations Research) könne dann auf die Produktionsfunktion geschlossen werden.

Angenommen, man kann einen bestimmten Output (z.B. Aushub einer Baugrube mit dem Volumen $30m^3$ an einem Tag) durch zwei unterschiedliche Aktivitäten (u_1, u_2) oder (v_1, v_2) herstellen. Dann ist derselbe Output auch dadurch zu erzielen, indem man eine Konvexkombination der beiden Aktivitäten wählt.

$$t\big(u_1, u_2\big) + (1-t)\big(v_1, v_2\big) \quad 0 \leq t \leq 1$$

Konvexkombination und konvexe Hülle von Aktivitäten

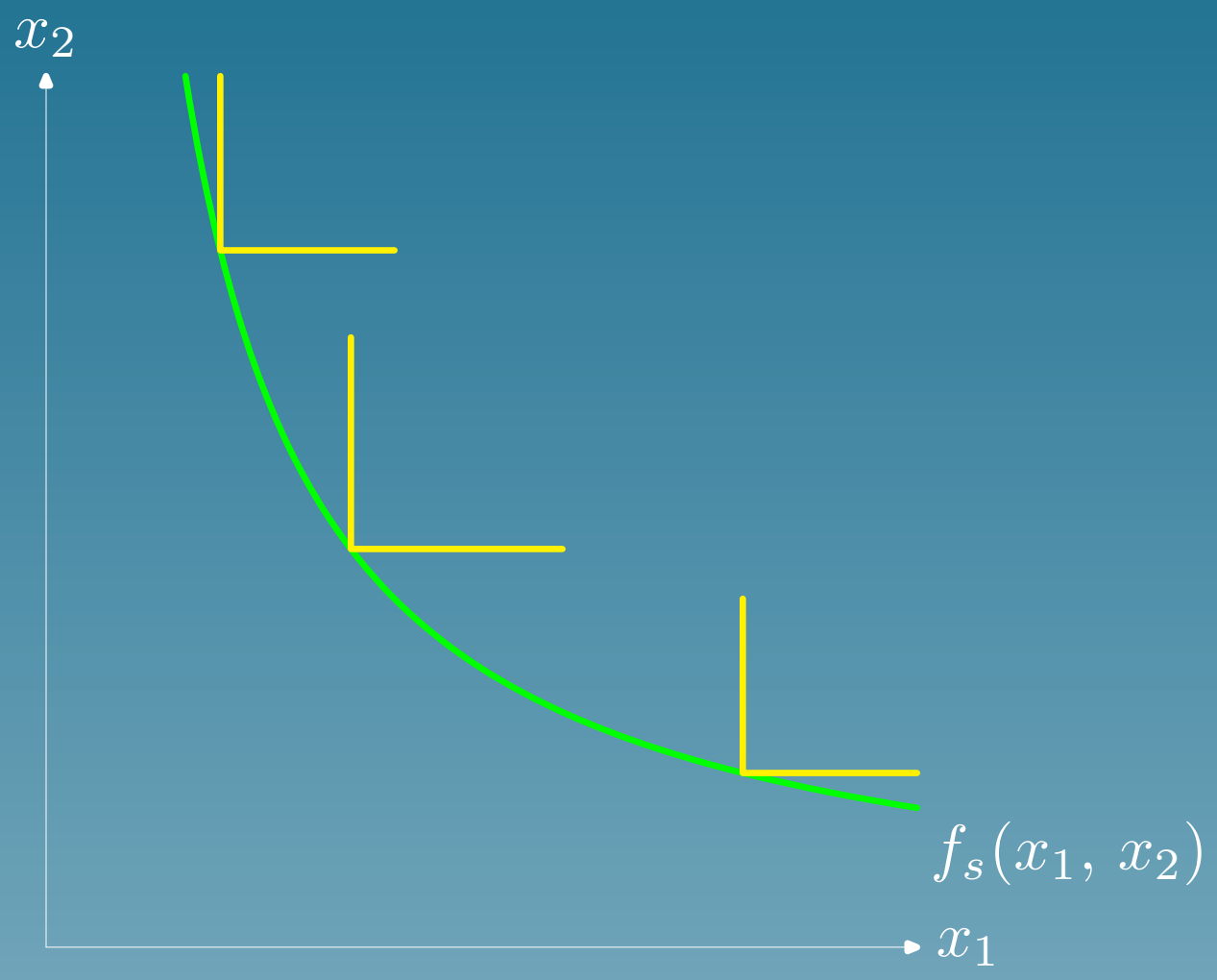


Substitution als Wechsel der Aktivitäten

Es liegt nun zwar nahe sehr viele unterschiedliche Aktivitäten zu unterstellen, die alle zur Produktion derselben Gütersorte benutzt werden können. Die konvexe Hülle dieser Aktivitäten nähert sich dann mehr und mehr der Gestalt einer substitutionalen Isoquante an. Das Ersetzen des einen Inputs durch den anderen könnte dann einfach als Übergang von einer Aktivität zu einer anderen gedeutet werden.

Allerdings ist diese Interpretation falsch. Die Darlegung der Gründe würde zu weit führen. Man bezeichnet die obige falsche Interpretation als *surrogate Produktionsfunktion*. Natürlich lässt sich eine surrogate Produktionsfunktion zeichnen, aber die ökonomische Interpretation dieser als substitutionale Isoquante ist falsch.

Surrogate Produktionsfunktion



Grenzproduktivität und -produkt eines Inputs

Angenommen, die Produktionsfunktion $y = f(x_1,\, x_2)$ ist substitutional, und es ist das Inputbündel (u_1, u_2) in der Firma eingesetzt. Wie ändert sich der Output, wenn der erste Input geringfügig um $\Delta x_1 > 0$ vermehrt wird?

$$\Delta y = f\big(u_1 + \Delta x_1,\, u_2\big) - f\big(u_1,\, u_2\big)$$

Diese Veränderung heißt *Durchschnittsprodukt* des Inputs 1 zwischen u_1 und $u_1 + \Delta x_1$ an der Stelle (u_1, u_2).

Wir können auch den Newton-Quotienten bilden:

$$\frac{\Delta y}{\Delta x_1} = \frac{f\left(u_1 + \Delta x_1,\, u_2\right) - f\left(u_1,\, u_2\right)}{\Delta x_1}$$

Der Newton-Quotient heißt im Zusammenhang mit einer substitutionalen Produktionsfunktion *Durchschnittsproduktivität* des Inputs 1 zwischen u_1 und $u_1 + \Delta x_1$ an der Stelle (u_1, u_2).

Er misst die durchschnittliche Rate, mit der sich der Output im Verhältnis zum ersten Input verändert.

Wir betrachten nun den Grenzwert des Newton-Quotienten für $\Delta x_1 \to 0$. Man beachte, es wird *nicht* durch Null dividiert, sondern überlegt, wie sich die durchschnittliche Rate — der Bruch — verhält, wenn der Nenner gegen Null strebt (und nicht Null ist).

$$\lim_{\Delta x_1 \to 0} \frac{\Delta y}{\Delta x_1} = \lim_{\Delta x_1 \to 0} \frac{f\left(u_1 + \Delta x_1,\, u_2\right) - f\left(u_1,\, u_2\right)}{\Delta x_1} = f_1'(u_1,\, u_2)$$

Das Resultat dieses Grenzübergangs ist (sofern der Grenzwert existiert) die erste partielle Ableitung der Produktionsfunktion nach ihrem ersten Argument an der Stelle (u_1, u_2).

Die erste partielle Ableitung einer substitutionalen Produktionsfunktion nach dem ersten Input an einer Stelle heißt *Grenzproduktivität* des ersten Inputs an dieser Stelle $GP_1 = f_1'$.

Wir betrachten nun das erste partielle Differential von f an der Stelle (u_1, u_2):

$$\mathrm{d}\,y = f_1'\big(u_1, u_2\big)\,\mathrm{d}\,x_1$$

Hierbei ist f_1' die fest vorgegebene partielle Ableitung und $\mathrm{d}\,x_1$ eine *beliebig wählbare* Zahl. Setzen wir diese Zahl gleich der geringfügigen Vermehrung des ersten Inputs, dann erhalten wir eine lineare Näherung (Approximation) für Δy:

$$\Delta y \approx f_1'\big(u_1, u_2\big)\Delta x_1$$

Der rechts stehende Ausdruck heißt *Grenzprodukt* des ersten Inputs an der Stelle (u_1, u_2). Es handelt sich dabei um eine lineare Approximation des eingangs definierten Durchschnittsproduktes des ersten Inputs.

Wir unterstellen für den Rest des Mikro-2-Kurses, dass die substitutionale Produktionsfunktion $y = f(x_1, x_2)$ überall (in ihrem Definitionsbereich) zweimal stetig differenzierbar ist.

Dann können wir die zuvor genannten Maßzahlen überall und auch für den zweiten Input berechnen.

Im Lehrbuch von Varian gehen die Bezeichnungen durcheinander, sodass eine Klarstellung angebracht ist. In Mathe 2 werden die analytischen Werkzeuge sehr ausführlich dargestellt. Hier wurden sie lediglich angewandt und sind sicher nicht auf Anhieb zu begreifen.

Beispiel 1

Angenommen, die Technik der Firma wird durch die substitutionale Cobb-Douglas Produktionsfunktion $y = f(x_1, x_2) = 15x_1^{1/3}x_2^{2/3}$ beschrieben. Es wird das Inputbündel $(1000, 125)$ eingesetzt. Der Output beläuft sich dann auf $f(1000, 125) = 15(10)(25) = 3750$. Nun soll der erste Input geringfügig um eine Einheit auf 1001 erhöht werden, also $\Delta x_1 = 1$. Dadurch erhöht sich der Output auf $f(1001, 125) = 3751.2496$ und damit ist das Durchschnittsprodukt des ersten Inputs an dieser Stelle $\Delta y = 1.2496$.

Die Durchschnittsproduktivität des ersten Inputs $\Delta y/\Delta x_1$ beträgt ebenfalls 1.2496, da $\Delta x_1 = 1$. Die beiden Maßzahlen sind numerisch gleich, aber sie unterscheiden sich in ihren Dimensionen.

Die Grenzproduktivität GP_1 des ersten Inputs erhalten wir, indem zunächst allgemein die erste partielle Ableitung der Produktionsfunktion gebildet wird. Danach wird dieser Ausdruck an der Stelle $(1000,\, 125)$ ausgewertet:

$$f_1' = 15\frac{1}{3}x_1^{-2/3}x_2^{2/3} \quad \rightsquigarrow \quad f_1'(1000,\, 125) = 5\left(\frac{125}{1000}\right)^{2/3} = \frac{5}{4}$$

Also ist $GP_1 = 1.25 \neq 1.2496 = \Delta y/\Delta x_1$.

Nun zum Grenzprodukt des ersten Inputs. Wir kennen die partielle Ableitung und wir kennen die Veränderung des Inputs, also brauchen wir beides nur zu multiplizieren und erhalten:

$$f_1'(u_1, u_2)\Delta x_1 = \frac{5}{4} \cdot 1 = 1.25$$

Dies ist eine lineare Annäherung an den eingangs berechneten exakten Wert von 1.2496.

Grenzrate der technischen Substitution

Die Grenzrate der Substitution gibt die Steigung einer Indifferenzkurve an einer Stelle an.

Die Grenzrate der technischen Substitution (GRTS) gibt die Steigung einer Isoquante an einer Stelle an.

Eine Isoquante der Produktionsfunktion $y = f(x_1, x_2)$ beschreibt alle Inputbündel, die denselben Output $y = c$ hervorbringen. Das Outputniveau c ist dabei fest vorgegeben. Wenn der Typ der Produktionsfunktion bekannt ist, kann man die implizite Funktion $f(x_1, x_2) = c$ zwischen den beiden Inputs nach einem Input auflösen und davon die normale Ableitung bilden. Das direkte Verfahren beruht auf dem Satz über implizite Funktionen.

Angenommen, die Technik der Firma wird durch eine substitutionale Produktionsfunktion $y = f(x_1, x_2)$ beschrieben und derzeit wird das Inputbündel (u_1, u_2) eingesetzt. Der Output beläuft sich auf $c = f(u_1, u_2)$. Nun will die Firma auf eine geringfügige Menge Δx_2 des zweiten Inputs verzichten (weil dieser sich verteuert hat), aber das Produktionsniveau c soll (aufgrund langfristiger Lieferverpflichtungen) beibehalten werden. Wieviel mehr von Input 1 muss sie einsetzen, um die Verminderung von Δx_2 zu kompensieren?

Der Satz über implizite Funktionen besagt, wenn $f_1'(u_1,\, u_2) \neq 0$, dann gilt

$$\Delta x_1 \approx -\frac{f_2'(u_1,\, u_2)}{f_1'(u_1,\, u_2)}\,\Delta x_2$$

Die notwendige Kompensation Δx_1 durch den ersten Input ist annähernd proportional der geringfügigen Veränderung des zweiten Inputs, wobei der Proportionalitätsfaktor gleich dem negativen Verhältnis der beiden Grenzproduktivitäten ist.

Beispiel 2

Wir betrachten wieder die Cobb-Douglas Produktionsfunktion $y = 15x_1^{1/3}x_2^{2/3}$ und ein aktuelles Inputbündel $(1000,\, 125)$. Der Output beträgt $c = 3750$ und soll aufrecht erhalten werden. Nun wird der zweite Input geringfügig von 125 auf 124 Einheiten vermindert. Dadurch sinkt der Output auf $f(1000,\, 124) = 3729.9732$ und soll durch vermehrten Einsatz des ersten Input ausgeglichen werden.

Das Verhältnis der beiden Grenzproduktivitäten ist

$$\frac{f_2'(x_1,\, x_2)}{f_1'(x_1,\, x_2)} = \frac{10\left(\frac{x_1}{x_2}\right)^{1/3}}{5\left(\frac{x_2}{x_1}\right)^{2/3}} = 2\frac{x_1}{x_2}$$

Auswerten an der Stelle $(1000, 125)$ ergibt den Faktor 16. Dieser ist negativ zu nehmen und mit $\Delta x_2 = -1$ zu multiplizieren, wodurch sich $\Delta x_1 = 16$ als die ungefähre Aufstockung beim Input 1 ergibt.

Wir machen die Probe mit dem neuen Inputbündel $(1016, 124)$ und erhalten als Output (Funktionswert) $y = 3749.7613$, also fast wieder das gewünschte Niveau. Der exakte Wert für die kompensierende Erhöhung von Input 1 beläuft sich auf 16.1942 Einheiten.

Gesetz der sinkenden Grenzproduktivität

Für substitutionale Produktionsfunktionen wird unterstellt, dass die Grenzproduktivitäten der Inputs positiv sind — mit zunehmendem Input einer Sorte wird der Output größer. Das Gesetz von der sinkenden Grenzproduktivität verlangt nun weiter, dass die Grenzproduktivität jedes Inputs geringer wird, je mehr von dem Input bereits eingesetzt ist.

Mathematisch ausgedrückt: die ersten partiellen Ableitungen einer substitutionalen Produktionsfunktion sind alle positiv, während die zweiten partiellen Ableitungen (ohne die Kreuzableitungen) alle negativ sind

$$f_1'(x_1,\, x_2) > 0 \qquad f_{11}''(x_1,\, x_2) < 0$$

$$f_2'(x_1,\, x_2) > 0 \qquad f_{22}''(x_1,\, x_2) < 0$$

Beispiel 3

Die Cobb-Douglas Produktionsfunktion $y = Ax_1^a x_2^b$ hat positive Grenzproduktivitäten, falls $a > 0$ und $b > 0$ gilt, denn

$$\frac{\partial y}{\partial x_1} = aAx_1^{a-1}x_2^b \quad \text{und} \quad \frac{\partial y}{\partial x_2} = bAx_1^a x_2^{b-1}$$

Sie erfüllt das Gesetz der sinkenden Grenzproduktivität, falls $a < 1$ und $b < 1$, denn

$$\frac{\partial^2 y}{\partial x_1^2} = (a-1)aAx_1^{a-2}x_2^b \quad \text{und} \quad \frac{\partial^2 y}{\partial x_2^2} = (b-1)bAx_1^a x_2^{b-2}$$

Zusammenfassend muss also $0 < a < 1$ und $0 < b < 1$ für die Parameter der Cobb-Douglas Produktionsfunktion gelten, wenn sie diese sogenannten neoklassischen Eigenschaften aufweisen soll.

Gesetz der sinkenden GRTS

Dieses Gesetz ist Ausdruck der konvexen Krümmung der Isoquanten einer substitutionalen Produktionsfunktion. Es wurde sehr ausführlich im zweiten Kapitel meines Mikro-1-Skriptes besprochen (und kritisch von mehreren Seiten beleuchtet). Dem ist nichts hinzuzufügen.

Kurz- und langfristige Betrachtungen

Wir unterscheiden zwischen kurz- und langfristiger Betrachtung, je nachdem ob einige Inputs unveränderlich sind oder ob alle Inputs variabel sind. Die Unterscheidung basiert somit *nicht* auf der normalen Kalenderzeit, sondern auf der firmenspezifischen Technik.

Wird die grobe Unterteilung der Inputs in Kapital und Arbeit vorgenommen, dann wird meist unterstellt, dass der Kapitaleinsatz kurzfristig nicht verändert werden kann. Man spricht in diesem Zusammenhang auch von einer konstanten Betriebsgröße.

Manchmal wird auch noch eine sogenannte Marktperiode betrachtet. Während der Marktperiode kann kein Input verändert werden, daher ist auch der Output konstant.

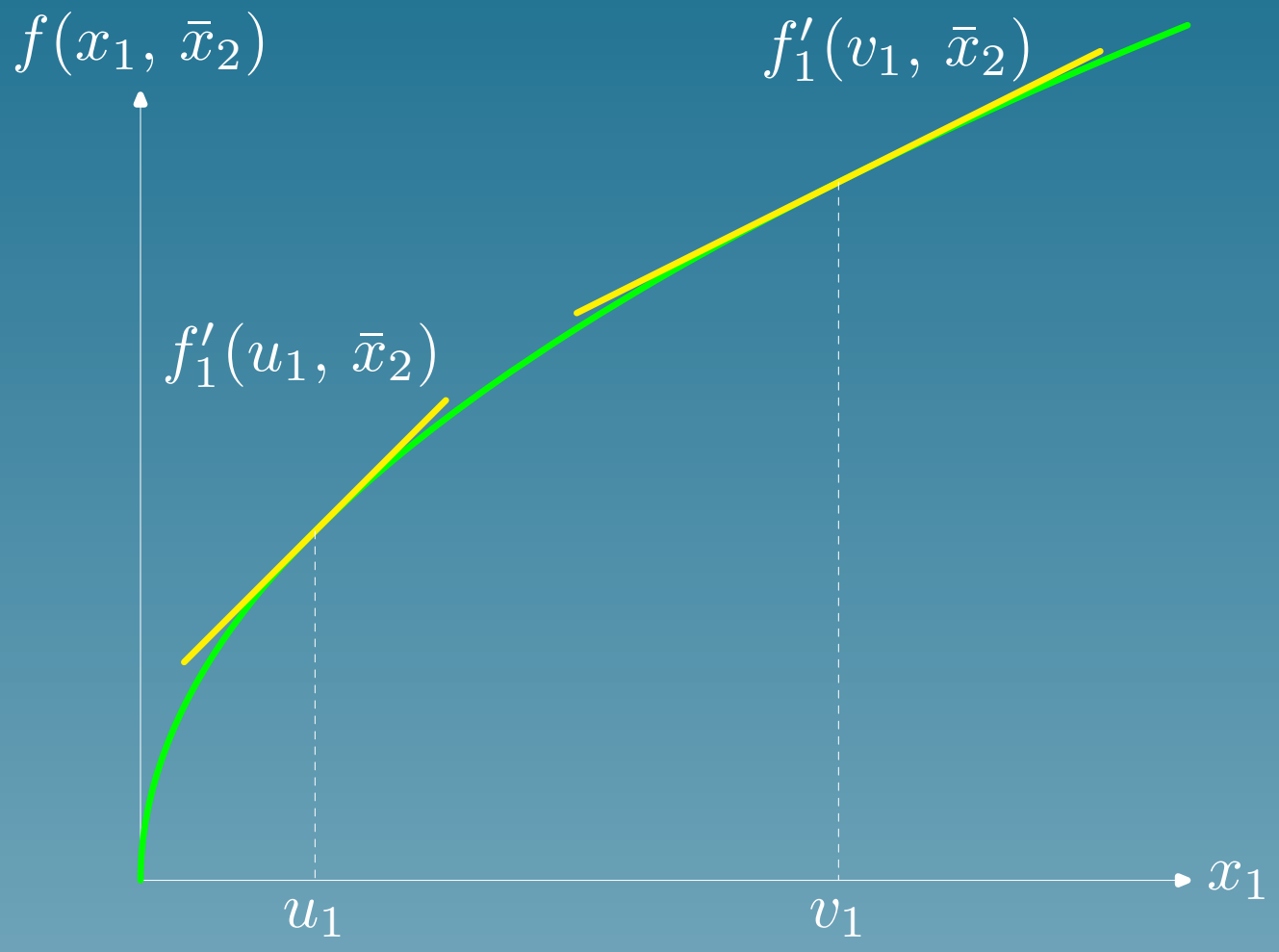
Ertragskurve und kurzfristige Produktionsfunktion

Angenommen der Input 2 ist kurzfristig in Höhe von $\bar{x}_2$ fixiert. Die kurzfristige Produktionsfunktion $y = f(x_1, \bar{x}_2)$, bei der nur der erste Input variabel ist, wird auch als Ertragskurve des ersten Inputs bezeichnet.

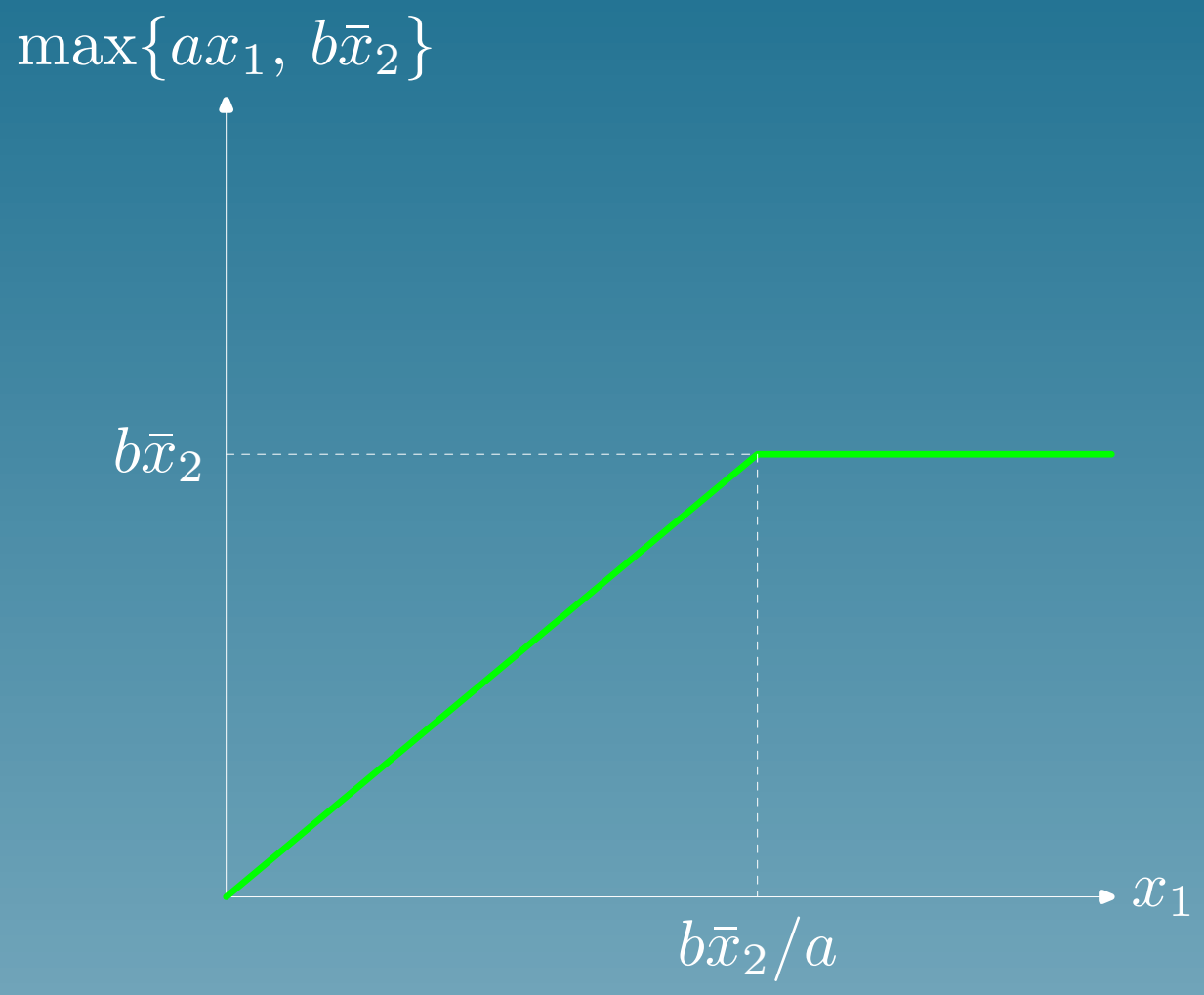
Im mathematischen Sprachgebrauch handelt es sich dabei um die erste partielle Funktion der Produktionsfunktion $f(x_1, x_2)$.

Bei der (kurzfristigen) Ertragskurve unterscheidet sich der typische Verlauf einer neoklassischen Produktionsfunktion (substitutional mit positiver und fallender Grenzproduktivität des variablen Inputs) von der limitationalen Produktionsfunktion.

Neoklassisches Ertragsgesetz



Limitationales Ertragsgesetz



Langfristige Betrachtung der Produktion

Bei langfristigen Betrachtungen wird meist unterstellt, dass alle Inputs proportional gleich verändert werden. Der Input-Mix (die Zusammensetzung des Inputbündels) ist konstant, und es werden Vielfache davon in der Firma eingesetzt.

Die Konstanz des Input-Mixes ist bei limitationalen Produktionsfunktionen unmittelbar einsichtig: nur die technisch effizienten Eckpunkte der Isoquanten kommen als Inputbündel in Frage, wenn alle Inputs für die Firma variabel sind.

Bei substitutionalen Produktionsfunktionen wird die Konstanz ökonomisch begründet durch konstante Preise der Inputs. Diese Begründung wird im Rahmen der Kostenminimierung erörtert.

Konstante Skalenerträge

Man stelle sich eine Firma vor, die mit einem bestimmten Input-Mix produziert. Nun wird die Firma identisch repliziert (geklont), dann wird der Output beider Firmen gerade doppelt so groß wie vorher sein. M.a.W. die Verdoppelung *aller* Inputs verdoppelt auch den Output.

Wenn sich der Output infolge einer proportionalen Änderung des Einsatzes *aller* Inputs in derselben Proportion verändert, spricht man von konstanten Skalenerträgen.

Im mathematischen Sprachgebrauch ist in diesem Fall die Produktionsfunktion homogen vom Grad 1 oder linearhomogen.

$$f(tx_1, tx_2) = tf(x_1, x_2) \quad \text{für} \quad t > 0$$

Beispiel 4

Die Leontief Produktionsfunktion $y = \max\{ax_1, bx_2\}$ hat konstante Skalenerträge (ist linearhomogen), denn für $t > 0$ gilt:

$$f(tx_1, tx_2) = \max\{a(tx_1), b(tx_2)\} = t\max\{ax_1, bx_2\} = tf(x_1, x_2)$$

Die Cobb-Douglas Produktionsfunktion hat konstante Skalenerträge, falls die Summe der beiden Parameter a und b gleich Eins ist, denn

$$f(tx_1, tx_2) = A(tx_1)^a(tx_2)^b = t^{a+b}Ax_1^a x_2^b = t^{a+b}f(x_1, x_2)$$

Am Beispiel der Cobb-Douglas Produktionsfunktion sieht man, dass das Gesetz der sinkenden Grenzproduktivität eines Inputs mit konstanten Skalenerträgen durchaus vereinbar ist.

Gesetz von Wicksell

Ist $f(x_1, x_2)$ eine Produktionsfunktion mit konstanten Skalenerträgen (homogen vom Grad Eins), welche positive aber abnehmende Grenzproduktivitäten der beiden Inputs aufweist ($f_1' > 0$, $f_{11}'' < 0$ und $f_2' > 0$, $f_{22}'' < 0$), dann sind die partiellen Kreuzableitungen positiv ($f_{12}'' > 0$), d.h. die Grenzproduktivität eines Inputs steigt, wenn vom anderen mehr eingesetzt wird.

Der Beweis des Gesetzes basiert auf dem Satz von Euler über homogene Funktionen. Sowohl der Satz von Euler als auch der Beweis des Gesetzes von Wicksell wird in Mathe 2 behandelt.

Steigende Skalenerträge

Falls sich der Output bei einer proportionalen *Ausdehnung* aller Inputs überproportional vergrößert, spricht man von steigenden Skalenerträgen.

$f(tx_1, tx_2) > tf(x_1, x_2)$ für $t > 1$

Steigende Skalenerträge können beispielsweise bei homogenen Produktionsfunktionen vorkommen, deren Homogenitätsgrad größer als Eins ist. Weiter können sie auch bei homothetischen Produktionsfunktionen vorkommen, wenn die zugrunde liegende homogene Funktion (die mit einer monoton wachsenden Funktion verknüpft ist) einen Homogenitätsgrad $k > 1$ aufweist.

Fallende Skalenerträge

Falls sich der Output bei einer proportionalen Ausdehnung aller Inputs unterproportional vergrößert, weist die Produktionsfunktion fallende Skalenerträge auf.

$$f(tx_1, tx_2) < tf(x_1, x_2) \quad \text{für} \quad t > 1$$

Die Argumentation im Lehrbuch von Varian, dass in diesem Fall wohl ein Denkfehler vorliegt, ist stichhaltig.

Ein ähnliches Argument (vgl. den Aufsatz von Sraffa) gilt jedoch auch für den Fall der steigenden Skalenerträge. Allerdings wird dies heutzutage von vielen nicht mehr wahrgenommen.

Gewinnmaximierung

Das unternehmerische Ziel der (privaten) Eigentümer der Firma soll ein maximaler Gewinn sein.

Die Firma soll sich auf den Beschaffungsmärkten für Inputs und dem Absatzmarkt für ihr Produkt vollständiger Konkurrenz gegenüber sehen. Daher plant die Firma ihre Produktion so, als ob sie keinen Einfluss auf die Preise der Inputs und den Output hat.

Der volkswirtschaftliche Begriff des Gewinns

Der Gewinn in der Volkswirtschaftslehre ist immer definiert als Erlös minus Kosten.

Gewinn = Erlös − Kosten

Der Erlös ist gleich dem zu Marktpreisen bewerteten Output.

Die Kosten sind — bei inputorientierter Betrachtung — gleich der Summe der mit Marktpreisen bewerteten Inputmengen, und zwar *aller* zur Produktion benötigter Inputs unabhängig vom jeweiligen Besitzer.

Das Wort Gewinn hat im alltäglichen Sprachgebrauch eine davon abweichende Bedeutung.

Kurzfristiger Gewinn

Wir betrachten eine Firma, deren Produktion durch eine (substitutionale) Produktionsfunktion mit zwei Inputs und einem Output gekennzeichnet ist. Der zweite Input ist fixiert auf $\bar{x}_2$. Die kurzfristige Produktionsfunktion lautet dann $y = f\left(x_1, \bar{x}_2\right)$.

Wird der Preis des Outputs mit p bezeichnet und die beiden Preise der Inputs mit w_1 und w_2, dann lautet der kurzfristige Gewinn bei inputorientierter Betrachtung

$$\pi = py - w_1 x_1 - w_2 \bar{x}_2 = pf\left(x_1, \bar{x}_2\right) - w_1 x_1 - w_2 \bar{x}_2$$

Kurzfristige Gewinnmaximierung

Die Firma soll ihre Produktion kurzfristig so planen, dass sich ein maximaler Gewinn ergibt. Dabei steht ihr nur die Wahl der Einsatzmenge des ersten Inputs frei, denn kurzfristig ist der zweite Input konstant. Die Inputspreise w_1 und w_2 sind exogene Marktpreise, dasselbe gilt für den Preis p ihres Outputs.

Daher lautet ihre Optimierungsaufgabe: Maximiere durch Wahl von x_1 den Gewinn π.

$$\max_{x_1} pf\left(x_1, \bar{x}_2\right) - w_1 x_1 - w_2 \bar{x}_2$$

Gewinnmaximale Input-Regel

Der Gewinn wird maximal, wenn der Grenzgewinn gleich Null und die zweite Ableitung des Gewinns negativ ist. Ist x_1^* die gewinnmaximale Menge des ersten Inputs, dann muss gelten:

$$pf_1'(x_1^*, \bar{x}_2) - w_1 = 0 \quad \text{und} \quad pf_{11}''(x_1^*, \bar{x}_2) < 0$$

Die Bedingung erster Ordnung $pf_1' = w_1$ kann für positive Preise nur dann erfüllt werden, wenn die Grenzproduktivität des ersten Inputs an der Stelle x_1^* positiv ist. Die Bedingung zweiter Ordnung kann für positive Preise nur dann erfüllt werden, wenn die Veränderung der Grenzproduktivität des ersten Inputs an der Stelle x_1^* negativ ist.

Wertgrenzproduktivität = Inputpreis

Die Bedingung erster Ordnung für ein (inputorientiertes) Gewinnmaximum kann auch umgestellt werden zu

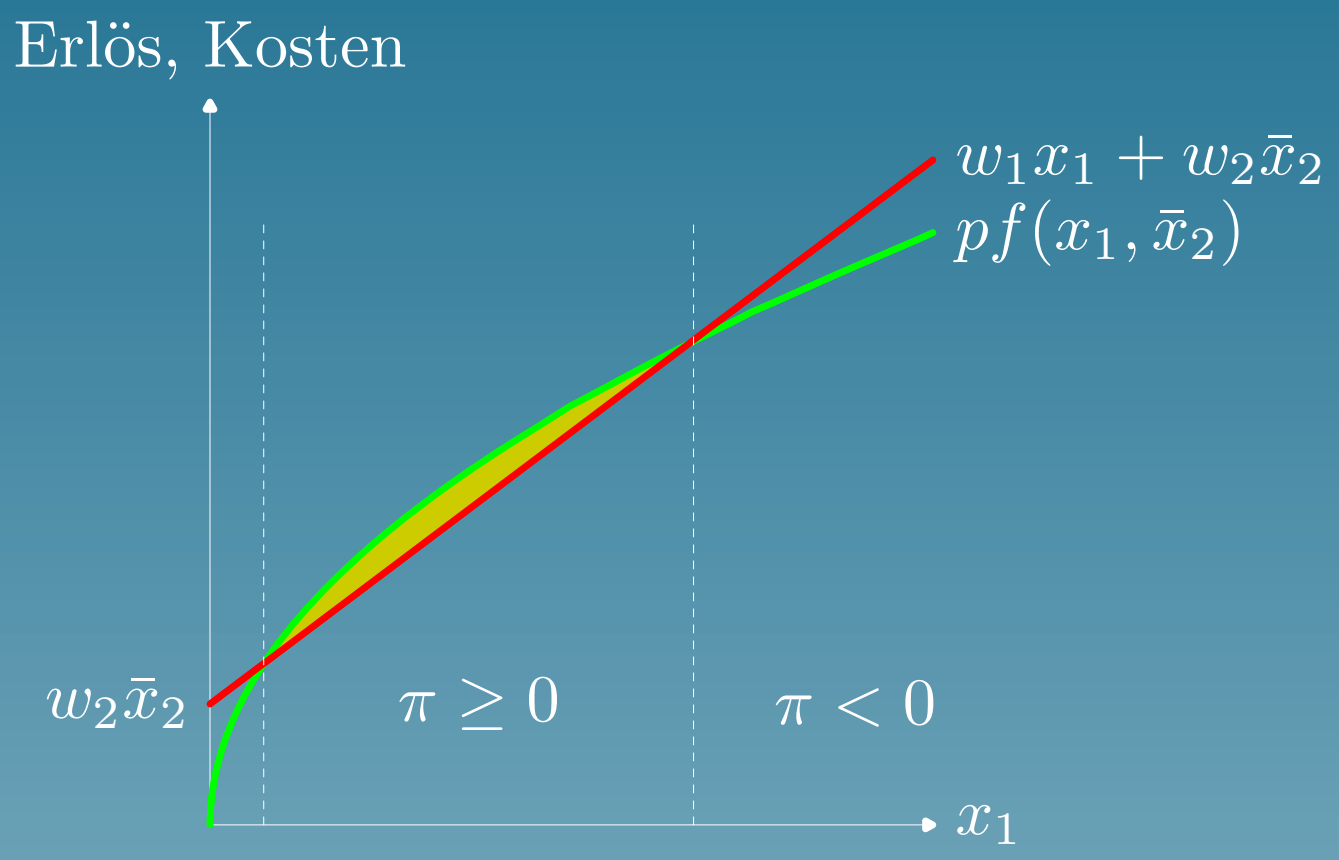
$$pf_1'\big(x_1^*,\, \bar{x}_2\big) = w_1$$

Die linke Seite bezeichnet dabei die *Wert*grenzproduktivität des ersten Inputs, die rechte Seite ist der Preis für den ersten Input.

Falls $pf_1' > w_1$ und die Firma eine kleine Einheit Δx_1 des ersten Inputs zusätzlich einsetzt, erhöht sich der Output um $f_1'\Delta x_1$ und damit bei konstantem Preis p der Erlös um $pf_1'\Delta x_1$. Zugleich erhöhen sich aber auch die Kosten um $w_1\Delta x_1$. Solange aber der zusätzliche Erlös größer als die zusätzlichen Kosten sind, steigt der Gewinn mit zunehmender Beschäftigung des ersten Inputs.

Wegen seiner sinkenden Grenzproduktivität wird jedoch der zusätzliche Erlös bei schrittweiser Erhöhung des ersten Inputs immer geringer, während die zusätzlichen Kosten $w_1\Delta x_1$ konstant sind. Daher wird ein Punkt erreicht, ab dem der Gewinn bei der schrittweisen Erhöhung des ersten Inputs nicht mehr weiter steigt, sondern sinkt.

Geometrische Betrachtung des kurzfristigen Gewinns



Kurzfristige Kosten- und Erlöskurven

Die Gerade $k = w_1 x_1 + w_2 \bar{x}_2$ stellt die kurzfristigen Kosten dar, die sich aus variablen Kosten $(w_1 x_1)$ und fixen Kosten $(w_2 \bar{x}_2)$ zusammensetzen.

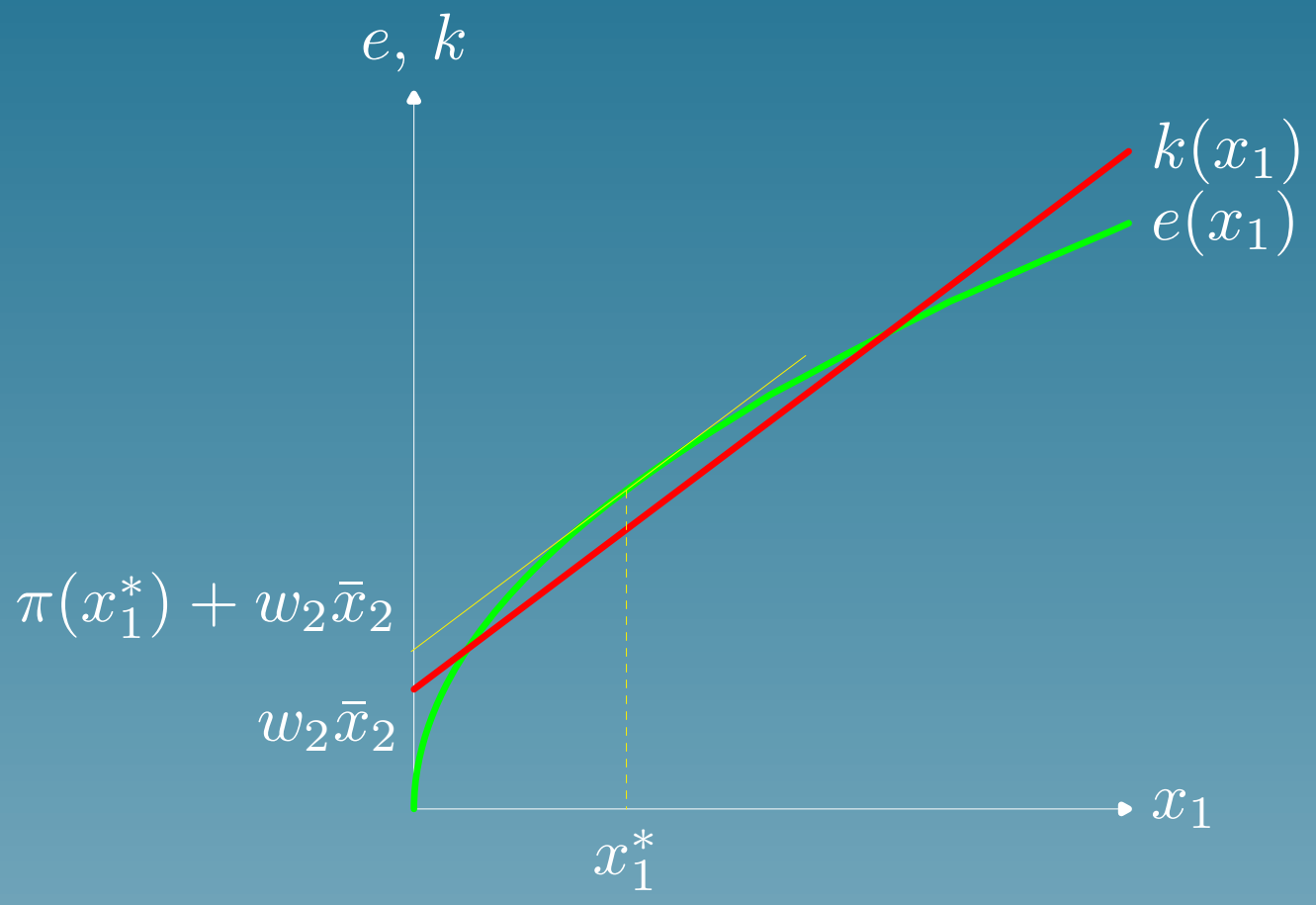
Die konkave Kurve $e = pf(x_1, \bar{x}_2)$ ist die kurzfristige Erlöskurve.

Die vertikale Differenz $e - k$ zwischen diesen beiden Kurven ist der kurzfristige Gewinn.

Da die Kostenkurve für sehr kleine und sehr große Inputmengen x_1 oberhalb der Erlöskurve verläuft, ist der Gewinn in diesen Bereichen negativ. An den (beiden) Schnittpunkten der Kosten- und Erlöskurve ist der kurzfristige Erlös gleich den kurzfristigen Kosten und somit der Gewinn gleich Null. Zwischen den beiden Schnittpunkten ist der kurzfristige Gewinn positiv.

Beachte: Bei einer Wanderung entlang der kurzfristigen Kostenkurve ist der kurzfristige Gewinn in jedem Punkt ein anderer.

Isogewinnpunkte



Geometrische Konstruktion des kurzfristigen Gewinnmaximums

Das kurzfristige Gewinnmaximum ist gekennzeichnet durch die (inputorientierte) Regel $pf_1' = w_1$. Es ist die Stelle, an der die kurzfristige Erlöskurve dieselbe Steigung wie die kurzfristige (inputorientierte) Kostenkurve hat.

Geometrisch kann man diesen Punkt gleicher Steigung konstruieren, indem man die kurzfristige Kostenkurve solange parallel verschiebt, bis sie die Erlöskurve tangiert. Der Ordinatenabschnitt dieser Parallenen zur Kostenkurve ist gleich dem maximalen Gewinn plus den Fixkosten.

Komparativ statische Analyse des kurzfristigen Gewinnmaximums

Wie reagiert die gewinnmaximierende Firma auf Veränderungen der drei Preise p, w_1 und w_2?

Sie wird ihren optimalen Produktionsplan anpassen, indem sie die Einsatzmenge des ersten Inputs in geeigneter Weise ändert.

Dabei ist zu vermuten, dass

- bei einer Steigerung des Preises ihres Outputs $\Delta p > 0$ der Einsatz des ersten Inputs erhöht wird $\Delta x_1 > 0$;
- bei einer Steigerung des Preises ihres variablen ersten Inputs $\Delta w_1 > 0$ sein Einsatz verringert wird $\Delta x_1 < 0$;
- bei geringfügigen Steigerungen des Preises ihres fixen zweiten Inputs der Einsatz des ersten Inputs gleich bleibt $\Delta x_1 = 0$;
- bei drastischen Steigerungen des Preises ihres fixen Inputs die Produktion eingestellt wird, d.h. $x_1 = 0$ gewählt wird.

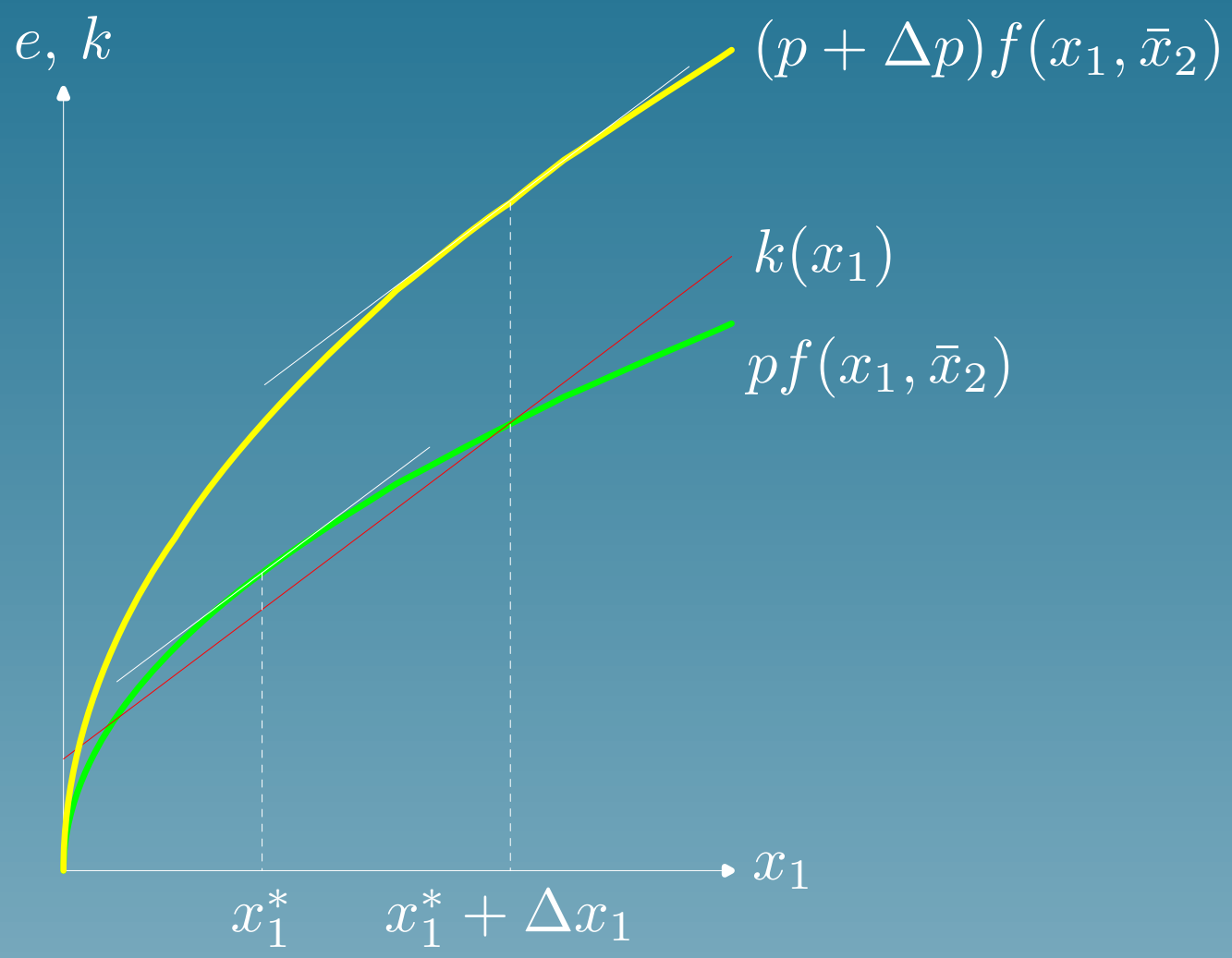
Erhöhung des Preises für den Output

Angenommen, das alte kurzfristige Gewinnmaximum wurde durch die Inputmenge x_1^* erreicht. Für dieses Quantum war die inputorientierte Regel $pf_1'(x_1^*, \bar{x}_2) = w_1$ erfüllt. Nun steigt der Preis und beim neuen Preis $p + \Delta p$ gilt dann $(p + \Delta p)f_1'(x_1^*, \bar{x}_2) > w_1$.

Daher lohnt es sich für die Firma, den Einsatz des ersten Inputs zu erhöhen, solange, bis der Ertragszuwachs wieder gleich dem konstanten Preis w_1 ist. Im neuen Gewinnmaximum muss gelten:

$$(p + \Delta p)f_1'(x_1^* + \Delta x_1, \bar{x}_2) = w_1$$

Reaktion auf $\Delta p > 0$



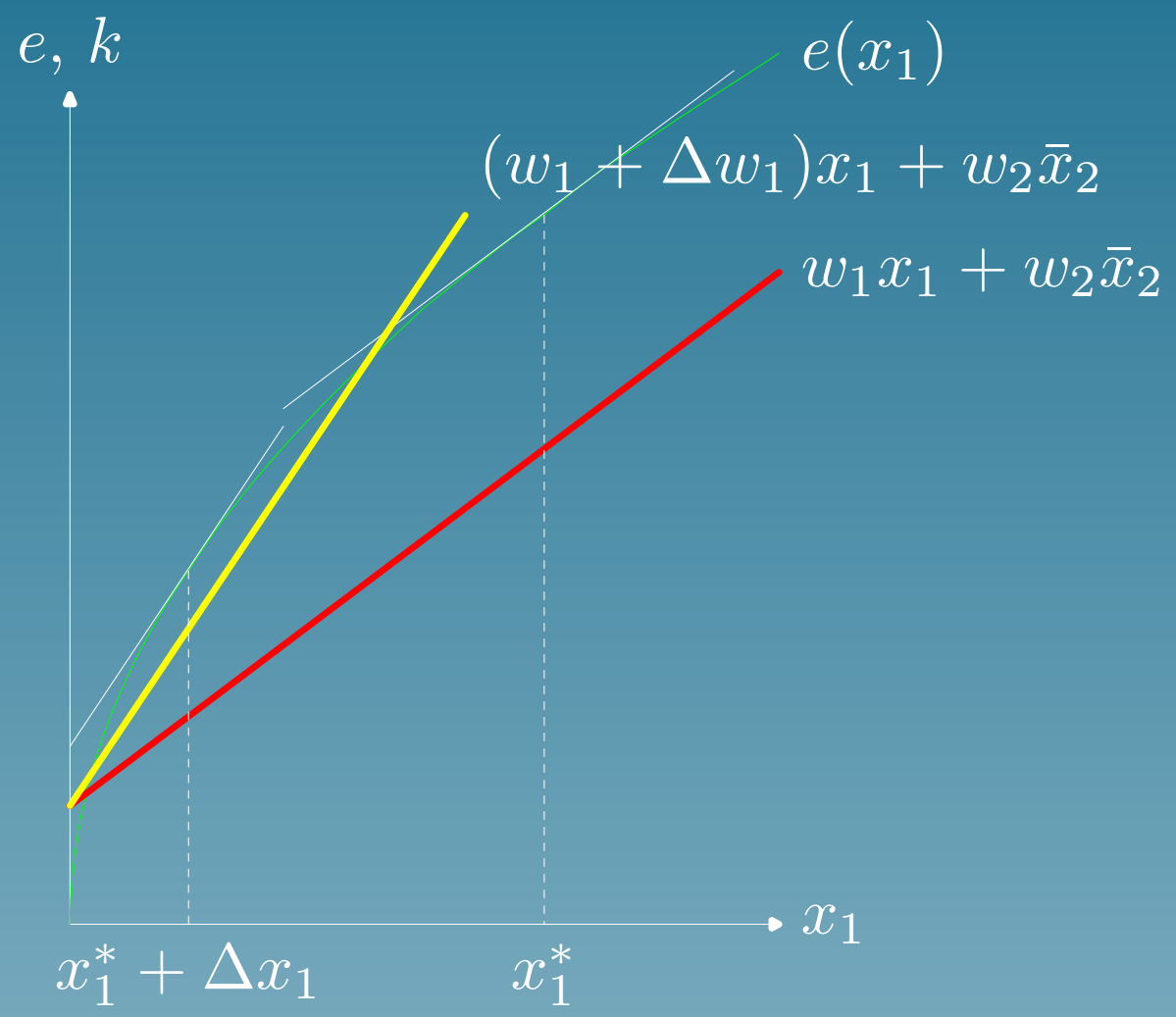
Erhöhung des Preises für den variablen Input

Angenommen, der Preis des variablen Inputs erhöht sich um Δw_1. Nun gilt beim alten gewinnmaximalen Input x_1^* die Ungleichung $pf_1'(x_1^*, \bar{x}_2) < w_1 + \Delta w_1$. Um die linke Seite zu vergrößern, muss die Grenzproduktivität des ersten Inputs steigen. Dies gelingt der Firma, indem sie die Einsatzmenge des ersten Inputs verringert. Im neuen Gewinnmaximum muss dann wieder

$$pf_1'(x_1^* + \Delta x_1, \bar{x}_2) = w_1 + \Delta w_1$$

erfüllt sein.

Reaktion auf $\Delta w_1 > 0$

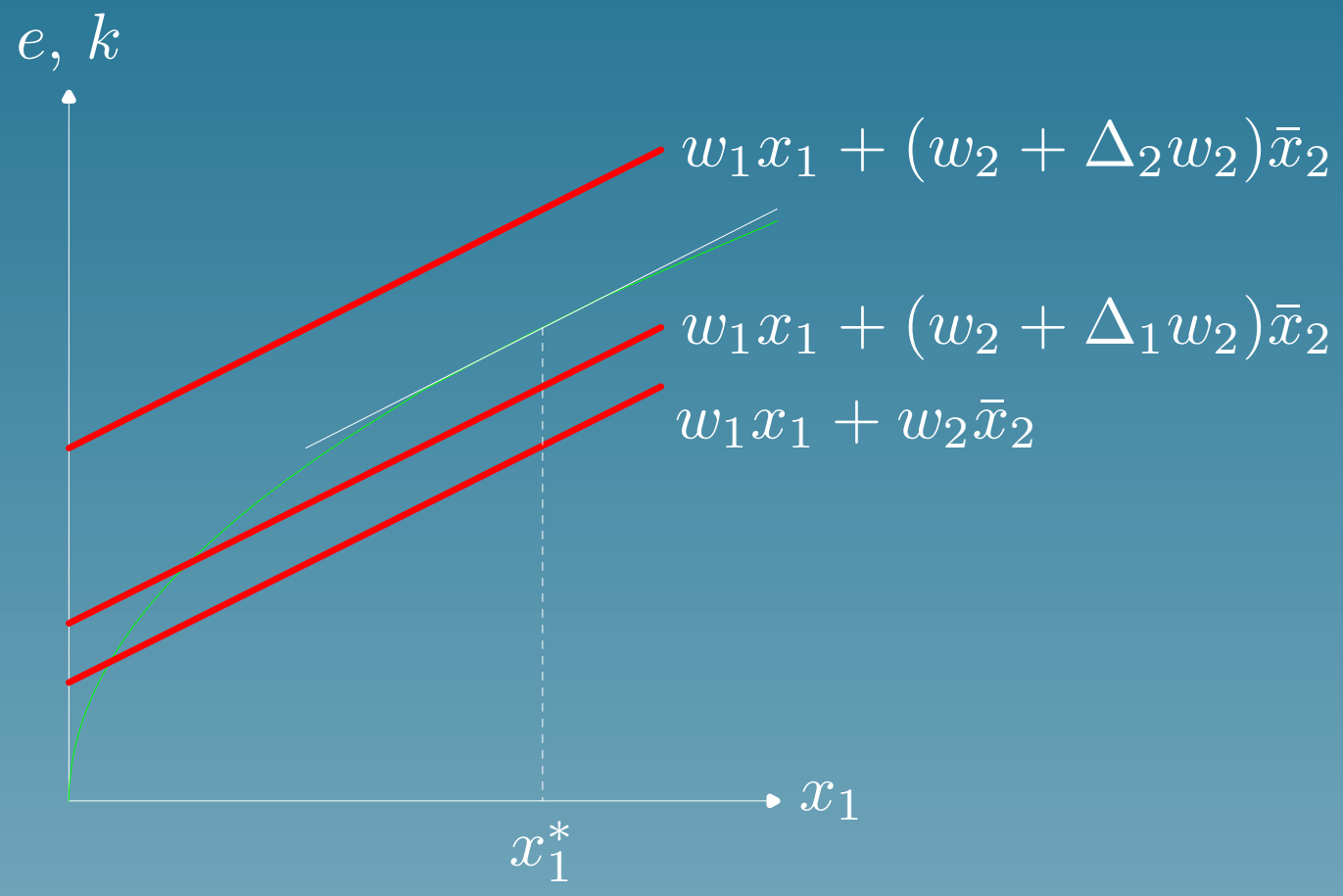


Erhöhung des Preises für den fixen Input

Bei einer geringfügigen Steigerung $\Delta_1 w_2$ des Preises für den kurzfristig konstanten Input erhöhen sich die fixen Kosten. In gleichem Maße verringert sich der kurzfristige Gewinn. Solange der Gewinn noch positiv bleibt, hat dies jedoch keine Anpassung beim variablen Input zur Folge.

Bei einer drastischen Steigerung $\Delta_2 w_2$ des Preises für den kurzfristig konstanten Input wird jedoch der Gewinn negativ auch bei optimaler Wahl des variablen Inputs. Die Firma wird daher die Produktion völlig einstellen. (Eigentlich ist die Firma indifferent zwischen Produzieren und nicht Produzieren. Gibt es jedoch quasi-fixe Kosten, wird sie die Produktion definitiv einstellen.)

Reaktionen auf $\Delta w_2 > 0$

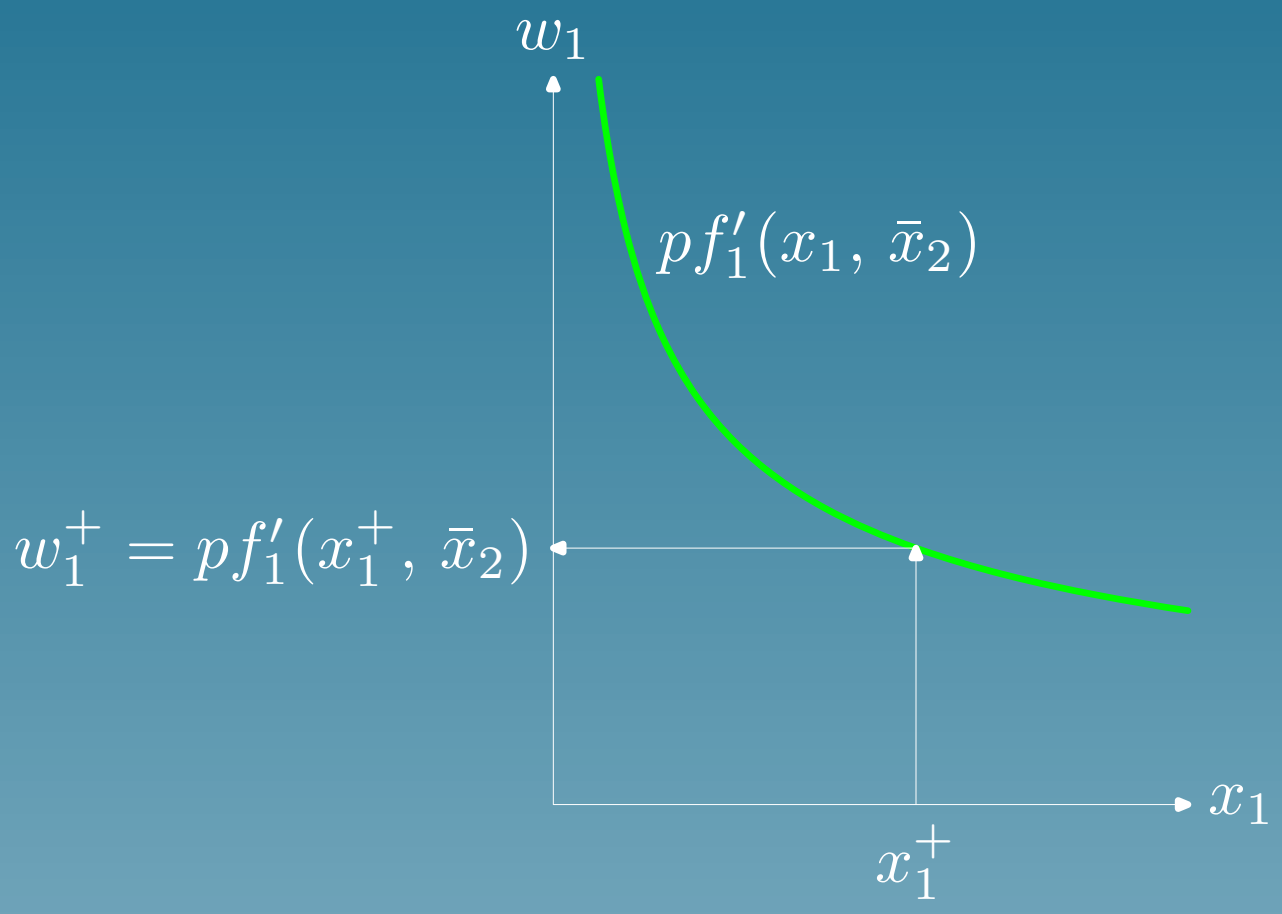


Inverse Inputnachfragekurve

Die kurzfristige gewinnmaximale Inputregel $pf_1'(x_1, \bar{x}_2) = w_1$ kann auch als inverse Nachfragekurve für den variablen Input interpretiert werden. Dazu werden die beiden Preise p und w_2 fixiert.

Bei einer bestimmten Inputmenge x_1^+ des ersten Inputs wird nun gefragt, welchen Preis die Firma pro Mengeneinheit zu zahlen bereit ist. Die Antwort lautet: beim Einsatz von x_1^+ Einheiten beläuft sich das Wertgrenzprodukt des ersten Inputs auf $pf_1'(x_1^+, \bar{x}_2)$; dies ist der Preis den die Firma höchstens zu zahlen bereit ist. Diese Überlegung gilt für beliebige Inputmengen x_1, daher stellt die Kurve der Wertgrenzproduktivität des ersten Inputs die Reservationspreise der gewinnmaximierenden Firma für ihn dar.

Wertgrenzproduktkurve = inverse Inputnachfrage



Langfristiges Gewinnmaximum

In der langen Frist sind alle Inputs variabel. Die Firma hat nun durch Wahl von x_1 und x_2 ihren Gewinn zu maximieren.

$$\max_{x_1,\, x_2} pf\left(x_1,\, x_2\right) - w_1 x_1 - w_2 x_2$$

Es gibt folglich zwei inputorientierte Gewinnmaximierungsregeln. Die beiden Inputs sind so zu wählen, dass gleichzeitig gilt:

$$pf_1'\left(x_1^*,\, x_2^*\right) = w_1$$
$$pf_2'\left(x_1^*,\, x_2^*\right) = w_2$$

Dies ist ein simultanes nichtlineares Gleichungssystem, denn f_1' wird nun auch durch x_2 beeinflusst und f_2' durch x_1.

Allgemeine Inputnachfragekurven

Falls die Hesse-Matrix der Produktionsfunktion den Rang 2 hat, kann das Gleichungssystem in der Nähe des gewinnmaximalen Inputbündels (x_1^*, x_2^*) nach den beiden Inputs in Abhängigkeit der drei Preise aufgelöst werden:

$$x_1 = g_1(w_1, w_2, p) \quad \text{und} \quad x_2 = g_2(w_1, w_2, p)$$

Dies sind die allgemeinen Inputnachfragekurven.

Langfristiges Gewinnmaximum und Skalenerträge

Angenommen, die Produktionsfunktion weist konstante Skalenerträge auf. Weiter sei das langfristige Gewinnmaximum beim Inputbündel (x_1^*, x_2^*) erreicht. Der Gewinn beläuft sich dann auf

$$\pi^* = pf(x_1^*, x_2^*) - w_1 x_1^* - w_2 x_2^*$$

Wie ändert sich der Gewinn π^*, wenn beide Inputs verdoppelt werden?

Der Output verdoppelt sich, und bei konstanten Preisen p, w_1 und w_2 verdoppeln sich Erlös und Kosten, sodass sich auch π^* verdoppelt.

Dies ist jedoch nur bei $\pi^* = 0$ eine sinnvolle Aussage.

Offenbarte Profitabilität

Die gewinnmaximierende Firma offenbart mit ihren Entscheidungen zweierlei:

- die gewählten Inputmengen und der zugehörige Output sind ein zulässiger Produktionsplan — Element der Produktionsmöglichkeitsmenge
- diese Wahl erbringt einen höheren Gewinn als andere.

Beobachtete Daten

Angenommen, wir beobachten zwei Entscheidungen der Firma.

Im Zeitpunkt t gelten die Preise (p^t, w_1^t, w_2^t) und die Firma plant (y^t, x_1^t, x_2^t). Zu einem anderen Zeitpunkt s gelten die Preise (p^s, w_1^s, w_2^s) und die Firma plant (y^s, x_1^s, x_2^s).

Wenn zwischen den beiden Zeitpunkten keine technischen Änderungen eingetreten sind, muss die Entscheidung im Zeitpunkt t profitabler sein als mit dem Plan von s. Genauso muss der Produktionsplan im Zeitpunkt s, der angesichts der s-Preise gewählt wurde, profitabler sein als der von t.

Schwaches Axiom der offenbarten Profitabilität

Folglich müssen bei einer gewinnmaximierenden Firma die beiden Ungleichungen erfüllt sein.

$$p^t y^t - w_1^t x_1^t - w_2^t x_2^t \geq p^t y^s - w_1^t x_1^s - w_2^t x_2^s$$
$$p^s y^s - w_1^s x_1^s - w_2^s x_2^s \geq p^s y^t - w_1^s x_1^t - w_2^s x_2^t$$

Werden Entscheidungen beobachtet, die diese Ungleichungen verletzen, dann hat die Firma in wenigstens einer der Perioden ihren Gewinn nicht maximiert. Daher bezeichnen wir diese empirisch nachprüfbaren Ungleichungen als *Schwaches Axiom der offenbarten Profitabilität* (WAPM).

Empirscher Test der komparativ statischen Analyse einer gewinnmaximierenden Firma

Falls die Entscheidungen der Firma das WAPM erfüllt, können die beiden Ungleichungen umgeformt und zusammengefasst werden zu

$$(p^t - p^s)(y^t - y^s) - (w_1^t - w_1^s)(x_1^t - x_1^s) - (w_2^t - w_2^s)(x_2^t - x_2^s) \geq 0$$

Die Formel des WAPM wird übersichtlicher, wenn die Veränderungen zwischen t und s mit Δ bezeichnet werden.

$$\Delta p \Delta y - \Delta w_1 \Delta x_1 - \Delta w_2 \Delta x_2 \geq 0$$

Damit kann man nun alle Reaktionen der Firma auf Preisänderungen, die im Rahmen der komparativ statischen Analyse gewonnen wurden, empirisch überprüfen.

Angenommen, nur der Preis für den Output hat sich zwischen t und s geändert, dann ist $\Delta p \neq 0$ und $\Delta w_1 = \Delta w_2 = 0$ und obige Ungleichung besagt: $\Delta p \Delta y \geq 0$, d.h. Outputpreis und Outputmenge verändern sich in der gleichen Richtung.

Ist dagegen nur $\Delta w_1 \neq 0$, erhalten wir $-\Delta w_1 \Delta x_1 \geq 0$, was bedeutet, dass sich Preis und Menge des ersten Inputs in gegensätzlicher Richtung verändern.

Aus den empirischen Beobachtungen lässt sich also bei Gültigkeit des WAPM zeigen, dass die Firma sich so verhält, als ob sie ihre Gewinne maximiert.

Rekonstruktion der Produktionsmöglichkeitsmenge

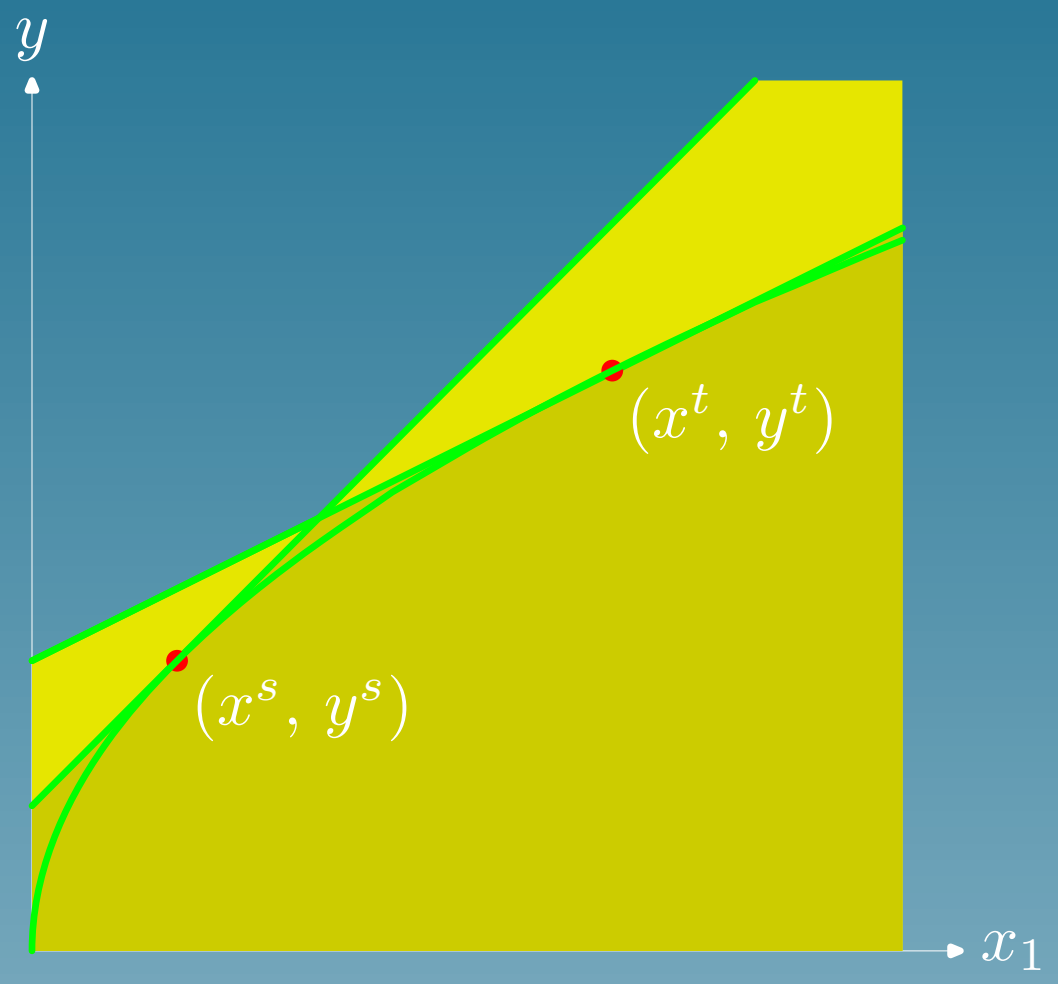
Im Rahmen der Haushaltstheorie konnte mit dem schwachen Axiom der offenbarten Präferenzen aus den beobachtbaren Kaufentscheidungen der Verlauf einer Indifferenzkurve angenähert werden. Analog kann aus den beobachteten Produktionsplänen einer Firma auf die ungefähre Gestalt ihrer Produktionsmöglichkeitsmenge geschlossen werden.

Dazu betrachten wir eine Firma, die nur einen Input in der Produktion einsetzt. In einer Periode t beobachten wir, dass sie angesichts der Preise p^t und w^t den Output y^t durch Einsatz von x^t Einheiten des Inputs produziert.

Daraus lässt sich der Gewinn π^t berechnen als $\pi^t = p^t y^t - w^t x^t$.

Falls sich die Firma gewinnmaximierend verhält, muss die Gerade $y = (\pi^t - w^t x)/p^t$ an der Stelle (x^t, y^t) die Produktionsmöglichkeitsmenge berühren. Weiter muss die Produktionsmöglichkeitsmenge unterhalb dieser Geraden liegen.

WAPM und Produktionsmöglichkeitsmenge



Gewinnmaximierung und Kostenminimierung

Eine Firma, die ihren Gewinn maximiert und dabei den Output y produziert, wird zugleich die Produktionskosten zur Herstellung dieses Outputs minimieren. Wären die Kosten *nicht* minimal, dann gäbe es einen preiswerteren Weg, den Output y zu produzieren und die Firma hätte noch nicht den Gewinn maximiert.

Daher kann die optimale (gewinnmaximale) Produktionsplanung einer Firma in zwei (analytischen) Schritten erfolgen.

1. bestimme die minimalen Kosten zur Herstellung einer beliebigen Menge des Outputs — das ist die Frage nach den outputorientierten Kosten;
2. wähle das Outputniveau, welches den maximalen Gewinn einbringt.

Kostenminimierung

Die Planung einer gewinnmaximierenden Firma kann in zwei Teilaufgaben zerlegt werden.

1. Welches sind die minimalen Kosten, zu dem sich ein bestimmter Output produzieren lässt?
2. Welches ist der gewinnmaximale Output?

In dieser und der nächsten Vorlesung diskutieren wir die erste Frage. Danach wird die zweite Frage betrachtet, sie führt uns zur Marktangebotskurve.

Aufgabenstellung der Kostenminimierung

Minimiere durch Wahl der Inputmengen x_1 und x_2 die Kosten $c = w_1 x_1 + w_2 x_2$, die bei der Herstellung des (vorgegebenen) Outputs $y = f(x_1, x_2)$ anfallen.

$$\min_{x_1,\, x_2} w_1 x_1 + w_2 x_2 \quad \text{u d N} \quad f(x_1, x_2) = y$$

Die Preise w_1 und w_2 der Inputs und der Outputlevel y sind exogene Größen.

Eine ähnliche Aufgabenstellung wurde bereits in der Haushaltstheorie untersucht. Dort war es das Ausgabenminimum eines Verbrauchers, das zu den kompensierten Nachfragekurven führte.

Analytische Lösung

Man ordnet der Nebenbedingung einen Lagrange-Multiplikator λ zu, bildet die Lagrange-Funktion

$$\mathcal{L}(x_1,\, x_2) = w_1 x_1 + w_2 x_2 - \lambda\big(f(x_1,\, x_2) - y\big)$$

und stellt die Bedingungen erster Ordnung auf

$$\mathcal{L}_1' = w_1 - \lambda f_1'(x_1,\, x_2) = 0 \qquad \text{(i)}$$

$$\mathcal{L}_2' = w_2 - \lambda f_2'(x_1,\, x_2) = 0 \qquad \text{(ii)}$$

$$f(x_1,\, x_2) = y \qquad \text{(iii)}$$

Jede Lösung $(x_1^*,\, x_2^*,\, \lambda)$ des nichtlinearen Gleichungssystems (i)–(iii) ist eine potenzielle Lösung der Optimierungsaufgabe und wird als Minimalkostenkombination (MKK) bezeichnet.

Eigenschaften der Lösungen

Da $w_i > 0$ und $f_i' > 0$, muss in (i) und (ii) der Lagrange-Multiplikator positiv sein $\lambda > 0$.

Es gilt

$$\lambda = \frac{w_1}{f_1'} = \frac{w_2}{f_2'} \quad \rightsquigarrow \quad \frac{w_1}{w_2} = \frac{f_1'}{f_2'} = -\text{GRTS}$$

Vergleicht man (i) und (ii) mit den gewinnmaximalen Inputregeln, dann tritt hier der Lagrange-Multiplikator an die Stelle des Preises p für den Output. Wohlgemerkt: hier werden die Kosten minimiert, dort wurde der Gewinn maximiert. Beide Ansätze sind dual.

Der Lagrange-Multiplikator gibt an, um wieviel sich der Wert der Zielfunktion bei einer MKK verändert, wenn der Output y variiert. Dies sind die sogenannten Grenzkosten.

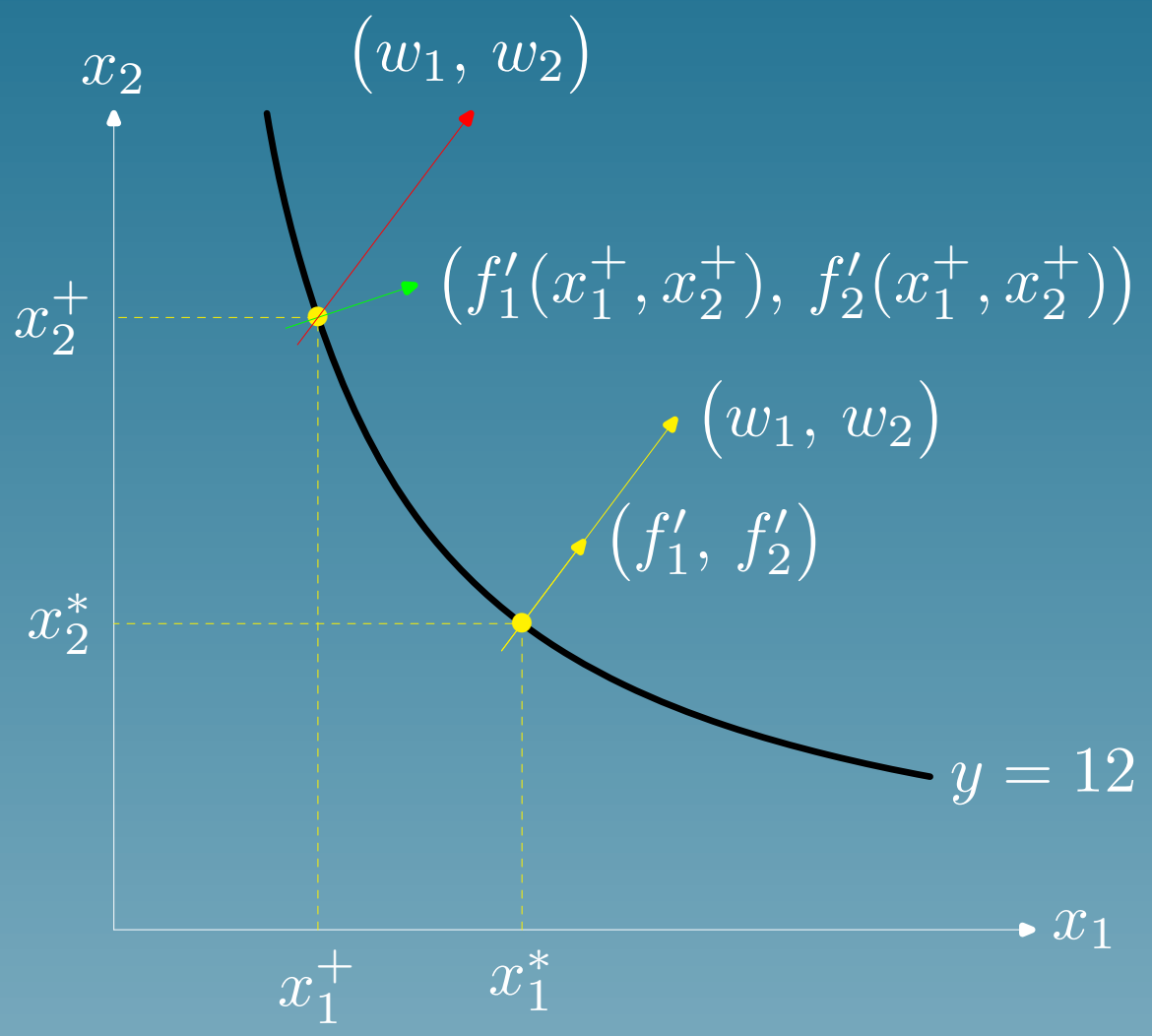
Da die Zielfunktion $c = w_1x_1 + w_2x_2$ linear und der Lagrange-Multiplikator positiv ist, kann von der Krümmung der Produktionsfunktion auf die Krümmung der Lagrange-Funktion geschlossen werden. Die Lagrange-Funktion ist konvex (nicht streng konvex), falls die Produktionsfunktion konkav ist. In diesem Fall sind die Bedingungen erster Ordnung (i)–(iii) auch hinreichend dafür, dass ein Minimum und kein Maximum der Zielfunktion vorliegt.

Kostenfunktion

Der Wert der Zielfunktion c ist gleich den minimalen Kosten, die bei der Herstellung des Outputs y und bei gegebenen Preisen der Inputs entstehen. Diese exogenen Größen beeinflussen die minimalen Kosten, daher gilt:

$$c = c(w_1, w_2, y)$$

Wanderung entlang einer Isoquante

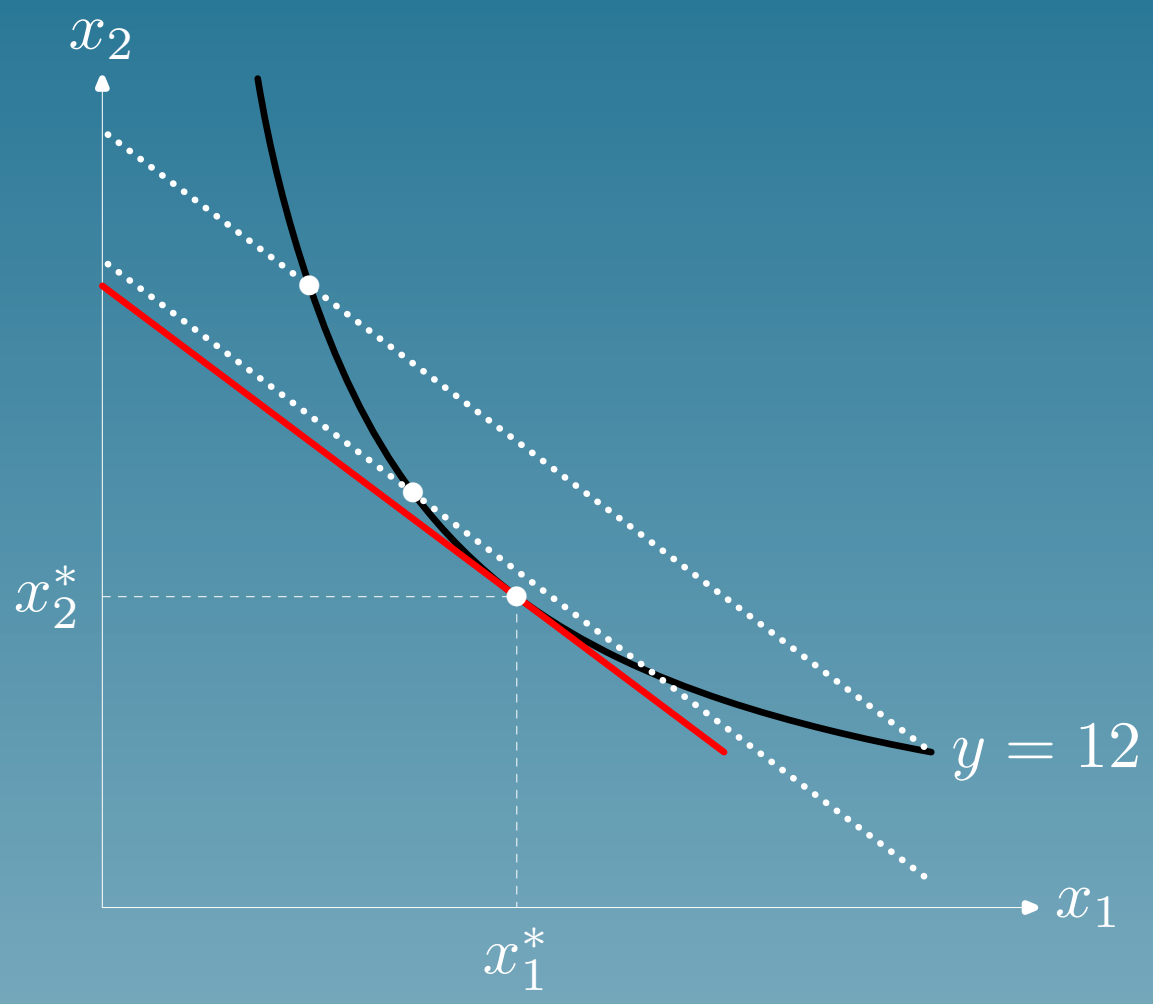


Isokostenlinie

Eine Höhenlinie $w_1x_1 + w_2x_2 = C$ mit konstantem C heißt Isokostenlinie. Es ist eine Gerade — ähnlich der Isoausgabenlinie in der Haushaltstheorie — deren Steigung durch $-w_1/w_2$ gegeben ist.

Statt entlang einer Isoquante zu wandern, bis der Gradient der Produktionsfunktion in Richtung des Inputpreisvektors zeigt, kann man auch eine Isokostenlinie solange in Richtung Ursprung verschieben, bis sie gerade noch die vorgegebene Isoquante tangiert.

Verschiebung einer Isokostenlinie



MKK und Inputpreise

Falls der Preis w_1 des ersten Inputs steigt, gibt es bei einer substitutionalen Isoquante eine neue MKK, die links oberhalb der alten liegt.

In der neuen MKK wird weniger vom teurer gewordenen Input eingesetzt. Um das Produktionsniveau zu halten, muss nun der andere Input vermehrt eingesetzt werden.

Bei einer limitationalen Isoquante ändert sich die MKK nicht, wenn sich die Preise der Inputs ändern.

Tritt bei beiden Inputpreisen eine proportional gleiche Preisänderung ein, ändert sich die MKK nicht.

Abgeleitete Inputnachfragekurven

Die MKK (x_1^*, x_2^*) ist abhängig vom gewählten Produktionsniveau y und von den beiden Inputpreisen w_1 und w_2.

Die Beziehung zwischen y, w_1, w_2 und dem kostenminimalen Inputbündel (der MKK) ergibt zwei sogenannte *abgeleitete Inputnachfragekurven*

$$x_1 = x_1(w_1, w_2, y) \quad \text{und} \quad x_2 = x_2(w_1, w_2, y)$$

Offenbarte Kostenminimierung

Die komparativ statische Untersuchung zur MKK lässt sich empirisch überprüfen.

Angenommen, wir beobachten bei den Preisen (w_1^t, w_2^t), dass die Firma ein Inputbündel (x_1^t, x_2^t) einsetzt und bei den Preisen (w_1^s, w_2^s) ein Inputbündel (x_1^s, x_2^s). Wenn wir unterstellen, dass der Output in beiden Fällen gleich y ist, dann muss bei kostenminimierendem Verhalten gelten

$$w_1^t x_1^t + w_2^t x_2^t \leq w_1^t x_1^s + w_2^t x_2^s$$
$$w_1^s x_1^s + w_2^s x_2^s \leq w_1^s x_1^t + w_2^s x_2^t$$

Die beiden Ungleichungen lassen sich zusammenfassen zu

$$(w_1^t - w_1^s)(x_1^t - x_1^s) + (w_2^t - w_2^s)(x_2^t - x_2^s) \leq 0$$

Schreibt man die Veränderungen der Preise und Mengen als Δ, wird der Ausdruck übersichtlicher

$$\Delta w_1 \Delta x_1 + \Delta w_2 \Delta x_2 \leq 0$$

Hat sich nur der Preis des ersten Inputs verändert (also $\Delta w_2 = 0$), dann gilt

$$\Delta w_1 \Delta x_1 \leq 0$$

Faktorpreis und Faktormenge bewegen sich in unterschiedliche Richtungen.

Homogenität der Kostenfunktion

Der Wert der Zielfunktion im Optimum ist

$$c = c(w_1, w_2, y) = w_1 x_1(w_1, w_2, y) + w_2 x_2(w_1, w_2, y)$$

Da sich bei einer proportionalen Änderung der Inputpreise die MKK nicht ändert, ist die Kostenfunktion linearhomogen in den Inputpreisen.

$$c(tw_1, tw_2, y) = tc(w_1, w_2, y) \quad \text{für } t > 0$$

Skalenerträge und Kostenfunktion

Wir betrachten die Beziehung zwischen den minimalen Kosten und dem Output bei gegebenen Preisen der beiden Inputs.

Wie hängt die Kostenfunktion $c(w_1, w_2, y)$ von y ab?

Angenommen, die Produktionsfunktion ist linearhomogen, d.h. $f(tx_1, tx_2) = tf(x_1, x_2)$ für $t > 0$. Der Output ändert sich bei einer proportionalen Vervielfachung beider Inputs in der gleichen Proportion. Das gilt auch für die besonderen MKK. Werden diese proportional verändert (was bei Konstanz der Inputpreise zutrifft), dann ändern sich der Output und die Kosten in derselben Proportion.

$$c(w_1, w_2, y) = yc(w_1, w_2, 1)$$

Steigende bzw. fallende Skalenerträge und Kostenfunktion

Bei steigenden Skalenerträgen steigt der Output bei einer Verdoppelung aller Inputs auf mehr als das Doppelte. Soll der Output verdoppelt werden, sind also beide Inputs zwar vermehrt einzusetzen, aber weniger als das Doppelte. Daher werden sich die minimalen Kosten unterproportional zu den Inputs entwickeln.

Umgekehrt verhält es sich bei fallenden Skalenerträgen. In diesem Fall steigen die Kosten überproportional zum Output.

Beides lässt sich anhand der sogenannten Durchschnittskosten einfacher ausdrücken.

Durchschnittskosten

Die Durchschnittskosten $AC(y)$ sind definiert als die minimalen Kosten zur Produktion von y (bei gegebenen Inputpreisen), dividiert durch den Output.

$$AC(y) = \frac{c(w_1,\, w_2,\, y)}{y}$$

Bei konstanten Skalenerträgen sind die minimalen Kosten direkt proportional den sogenannten Einheitskosten $c(w_1,\, w_2,\, 1)$ und die Durchschnittskosten, oder Stückkosten $AC(y)$ sind konstant.

$$AC(y) = \frac{c(w_1,\, w_2,\, y)}{y} = \frac{yc(w_1,\, w_2,\, 1)}{y} = c(w_1,\, w_2,\, 1)$$

Durchschnittskosten und Skalenerträge

Angenommen, die Produktionsfunktion ist homogen vom Grad k.

$$f(tx_1,\, tx_2) = t^k f(x_1,\, x_2) \quad \text{für } t > 0$$

Die Durchschnittskosten belaufen sich auf

$$AC(y) = \frac{c(w_1,\, w_2,\, 1)}{t^k} = t^{-k} c(w_1,\, w_2,\, 1)$$

Ist $k = 1$, sind die Durchschnittskosten konstant.

Ist $k > 1$ (steigende Skalenerträge), dann ist $AC(y)$ eine fallende Funktion des Outputs.

Ist $k < 1$ (sinkende Skalenerträge), dann steigt $AC(y)$ mit dem Output.

Lang- und kurzfristige Kosten

Langfristig sind alle Inputs variabel, kurzfristig sind einige Inputs fixiert und der Rest ist variabel. Der kurzfristige Entscheidungsspielraum der Firma ist enger als der langfristige.

Angenommen, der zweite Input ist kurzfristig auf ein bestimmtes Quantum $x_2 = \bar{x}_2$ fixiert. Dann lautet die kurzfristige Optimierungsaufgabe

$$\min_{x_1} w_1 x_1 + w_2 \bar{x}_2 \quad \text{u d N} \quad f(x_1, \bar{x}_2) = y$$

Als Lösung ergibt sich die kurzfristige MKK

$$x_1 = x_1^s(w_1, w_2, y)$$

$$x_2 = \bar{x}_2$$

Kurzfristige Kostenfunktion

Der Wert der Zielfunktion bei der kurzfristigen MKK bei alternativen Outputniveaus ist die kurzfristige Kostenfunktion

$$c_s(y, \bar{x}_2) = w_1 x_1^s(w_1,\, w_2,\, y) + w_2 \bar{x}_2$$

Kostenkurven

Die (outputorientierte) Kostenkurve $c(y)$ wird untersucht.

Sie stellt die minimalen Kosten dar, zu welchen sich der Output y produzieren lässt, gegeben die Preise der Inputs w_1 und w_2.

Zunächst werden kurzfristige, dann langfristige Kostenkurven betrachtet.

Kurzfristig gibt es Fixkosten F, die unabhängig vom Output anfallen, und variable Kosten c_v, die von der Höhe des Outputs abhängen.

Langfristig gibt es keine Fixkosten.

Kurzfristige Durchschnittskosten

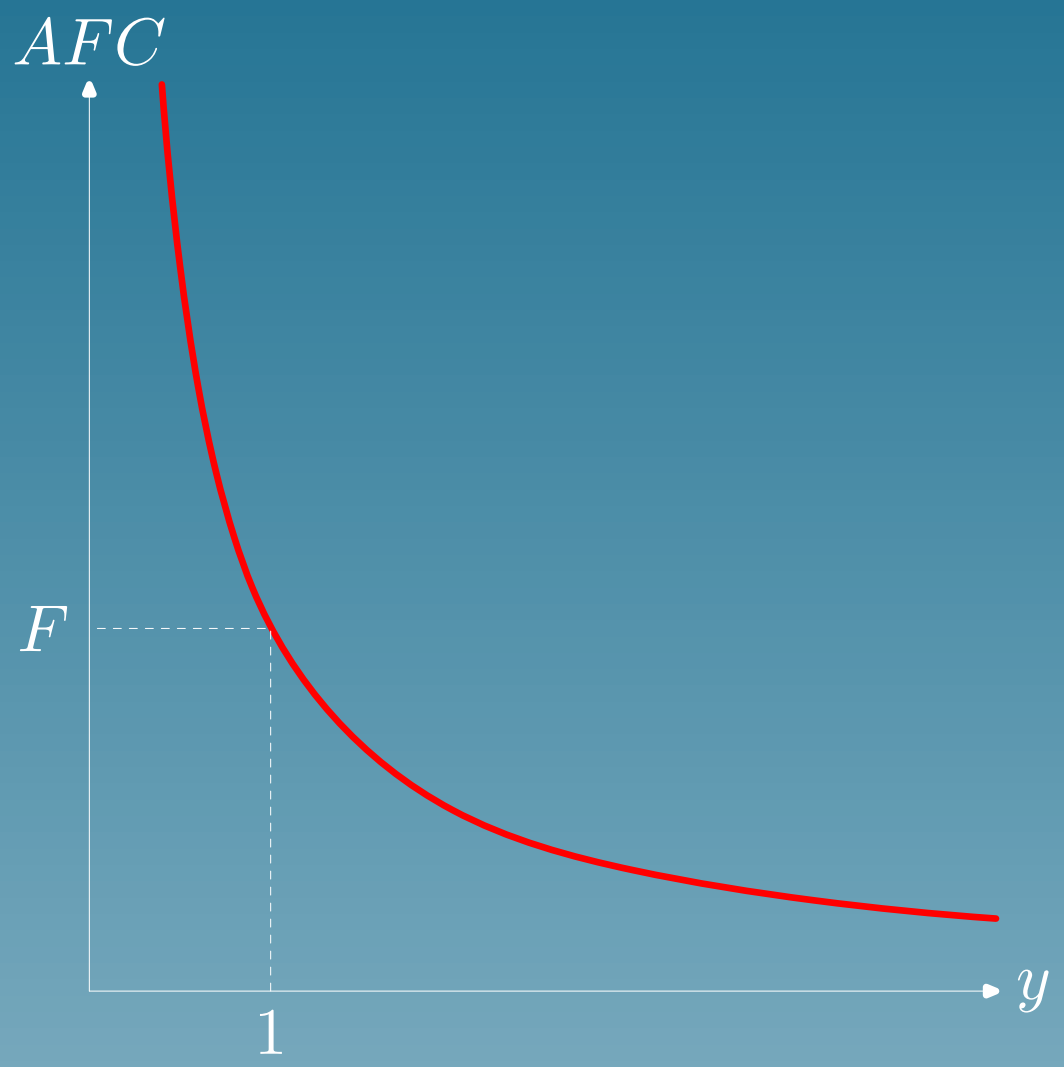
Die kurzfristigen Durchschnittskosten messen die Kosten pro Outputeinheit. Sie werden auch als kurzfristige Stückkosten bezeichnet.

Ausgehend von den kurzfristigen Kosten $c(y) = c_v(y) + F$ erhält man die Durchschnittskosten, indem durch den Output dividiert wird:

$$AC(y) = \frac{c(y)}{y} = \frac{c_v(y)}{y} + \frac{F}{y} = AVC(y) + AFC(y)$$

Die kurzfristigen Durchschnittskosten AC sind die Summe aus den durchschnittlichen variablen Kosten AVC und den durchschnittlichen Fixkosten AFC.

Graphik einer durchschnittlichen Fixkostenkurve



Durchschnittliche Fixkosten

Die Kurve der durchschnittlichen Fixkosten

$$AFC(y) = \frac{F}{y}$$

ist eine Hyperbel.

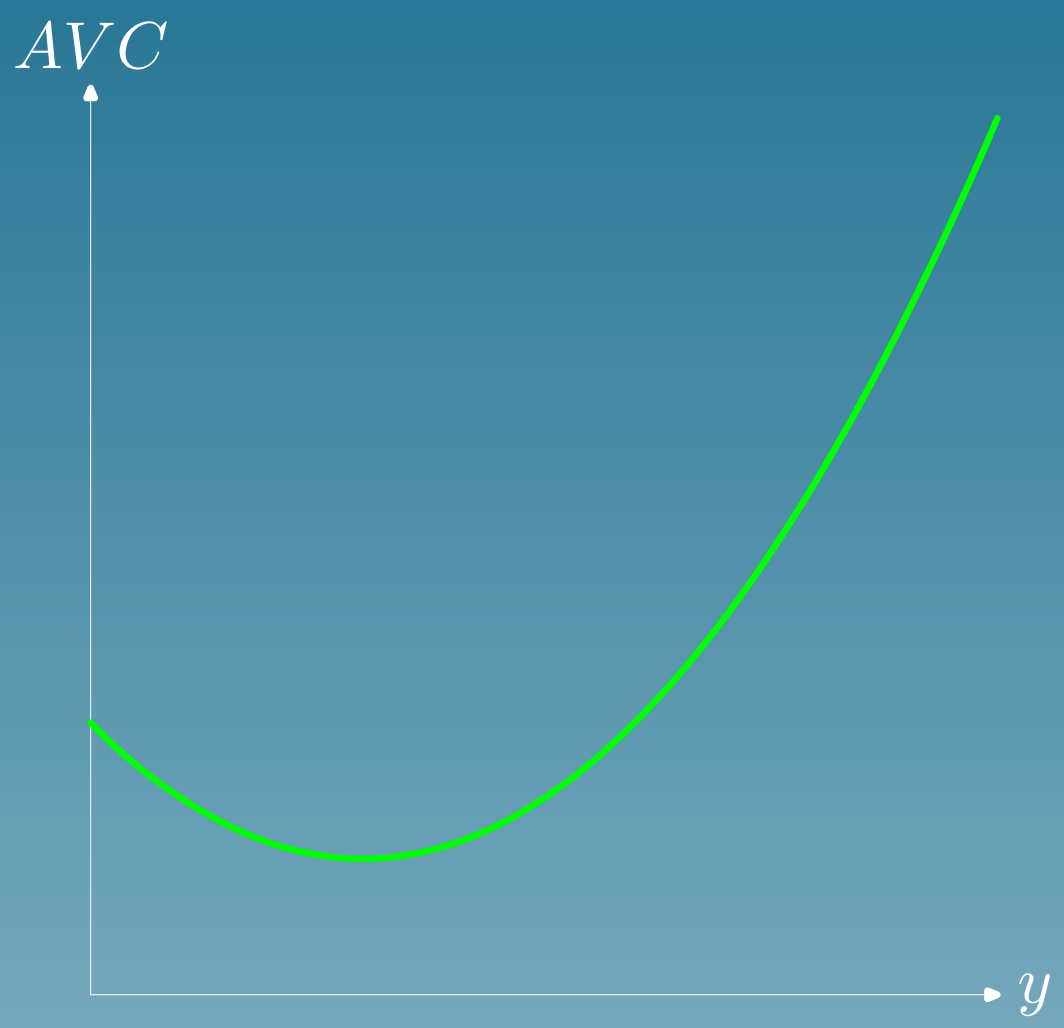
Ihre Lage ist eindeutig durch die Fixkosten F festgelegt.

Kurzfristige variable Durchschnittskosten

Die Gestalt der kurzfristigen variablen Durchschnittskostenkurve wird maßgeblich durch die kurzfristige Produktionsfunktion festgelegt.

Man unterstellt in der Regel einen U-förmigen Verlauf.

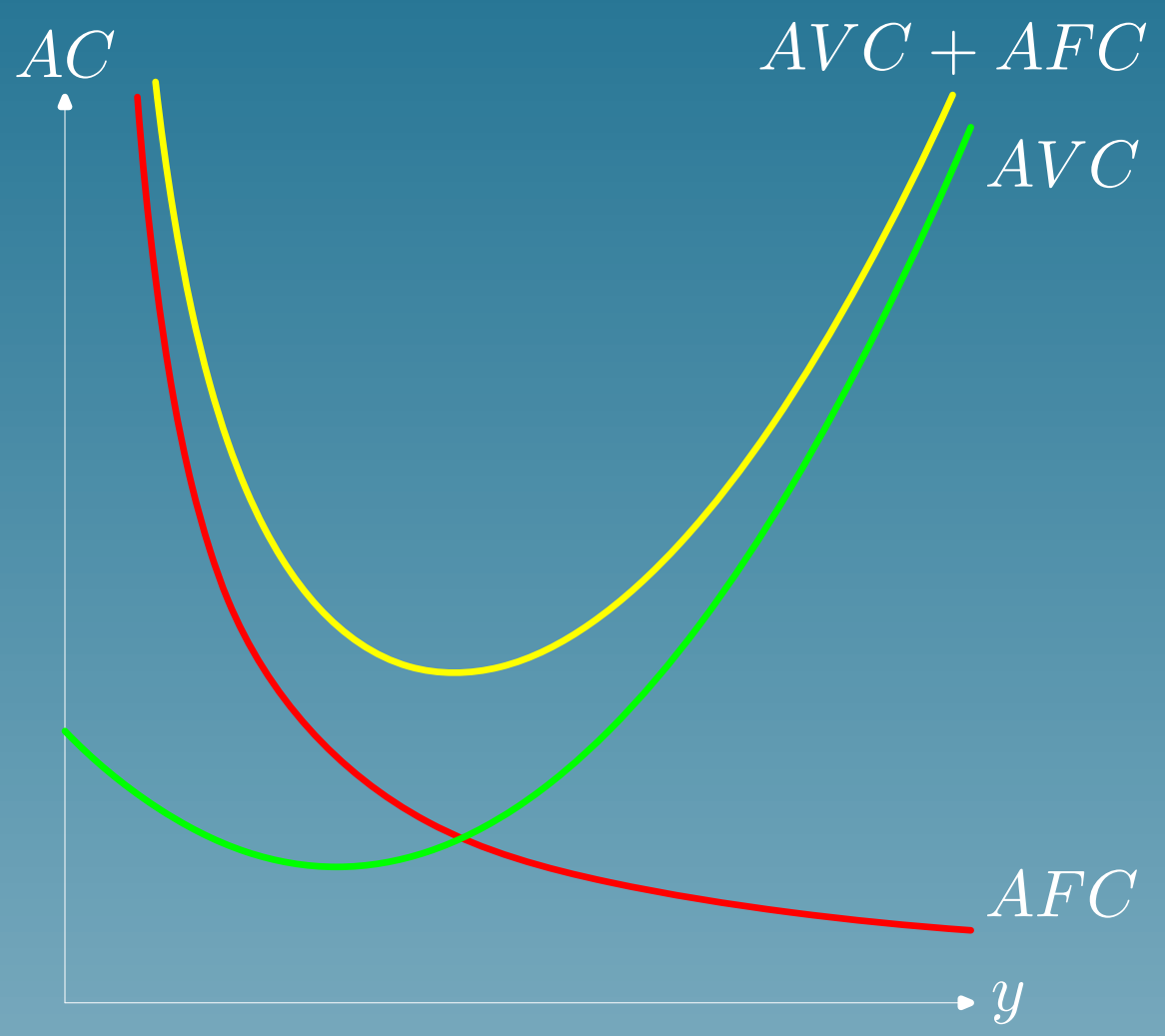
U-förmige kurzfristige variable Durchschnittskosten



Kurzfristige Durchschnittskosten

Durch vertikale Addition der durchschnittlichen variablen und fixen Kosten ergibt sich die Kurve der kurzfristigen Durchschnittskosten. Da für sehr kleine Outputniveaus die durchschnittlichen Fixkosten dominieren und bei sehr großen Outputniveaus die durchschnittlichen variablen Kosten, hat die Kurve der kurzfristigen Durchschnittskosten eine U-förmige Gestalt.

Typische Verläufe der kurzfristigen Durchschnittskosten



Grenzkosten

Im Zusammenhang mit der Kostenminimierungsaufgabe wurde bereits erwähnt, dass der Lagrange-Multiplikator gleich den Grenzkosten der Produktion ist.

Die Grenzkosten MC geben an, mit welcher Rate sich die minimalen Kosten bei einer kleinen Variation des Outputs ändern.

$$MC(y) = \lim_{h \to 0} \frac{c(y+h) - c(y)}{h} = c'(y)$$

Grenzkosten und variable Kosten

Da sich die (outputorientierten) Kosten aus den variablen und den fixen Kosten zusammensetzen, letztere aber bei jedem Outputniveau konstant sind, können die Grenzkosten auch als die Veränderungsrate der kurzfristigen variablen Kosten im Bezug auf eine kleine Outputveränderung angesehen werden.

$$MC(y) = c_v'(y)$$

Die Grenzkosten sind der Grenzwert eines Newton-Quotienten und nicht der Newton-Quotient wie bei Varian angegeben.

Grenzkosten und Durchschnittskosten

Die Ableitung der Durchschnittskosten ergibt

$$\frac{\partial \frac{c(y)}{y}}{\partial y} = \frac{yc'(y) - c(y)}{y^2} = \frac{1}{y}\left(c'(y) - \frac{c(y)}{y}\right) = \frac{1}{y}\big(MC(y) - AC(y)\big)$$

Der Ausdruck in der Klammer ist gleich Null — und damit die Ableitung der Durchschnittskosten — wenn die Grenzkosten gleich den Durchschnittskosten sind. MaW schneidet die Grenzkostenkurve die Durchschnittskostenkurve in deren Minimum.

Im Bereich fallender Durchschnittskosten ist ihre Ableitung negativ, dort muss also $MC < AC$ gelten.

Im Bereich steigender Durchschnittskosten ist ihre Ableitung positiv, also muss dort $MC > AC$ gelten.

Grenzkosten und variable Durchschnittskosten

Für $y \to 0$ werden die variablen Durchschnittskosten $AVC(y) = c_v(y)/y$ zu einem unbestimmten Ausdruck $0/0$, denn $c_v(0) = 0$. Mittels der Regel von l'Hospital kann jedoch der (linksseitige) Grenzwert bestimmt werden.

$$\lim_{y \to 0} AVC(y) = \lim_{y \to 0} \frac{c_v'(y)}{y'} = c_v'(0) = MC(0)$$

Daher ist auch die Beziehung $MC(1) = AVC(1)$ im Varian nicht korrekt. Vielmehr gilt $MC(0) = AVC(0)$, d.h. die Grenzkosten sind gleich den variablen Durchschnittskosten beim Outputniveau Null.

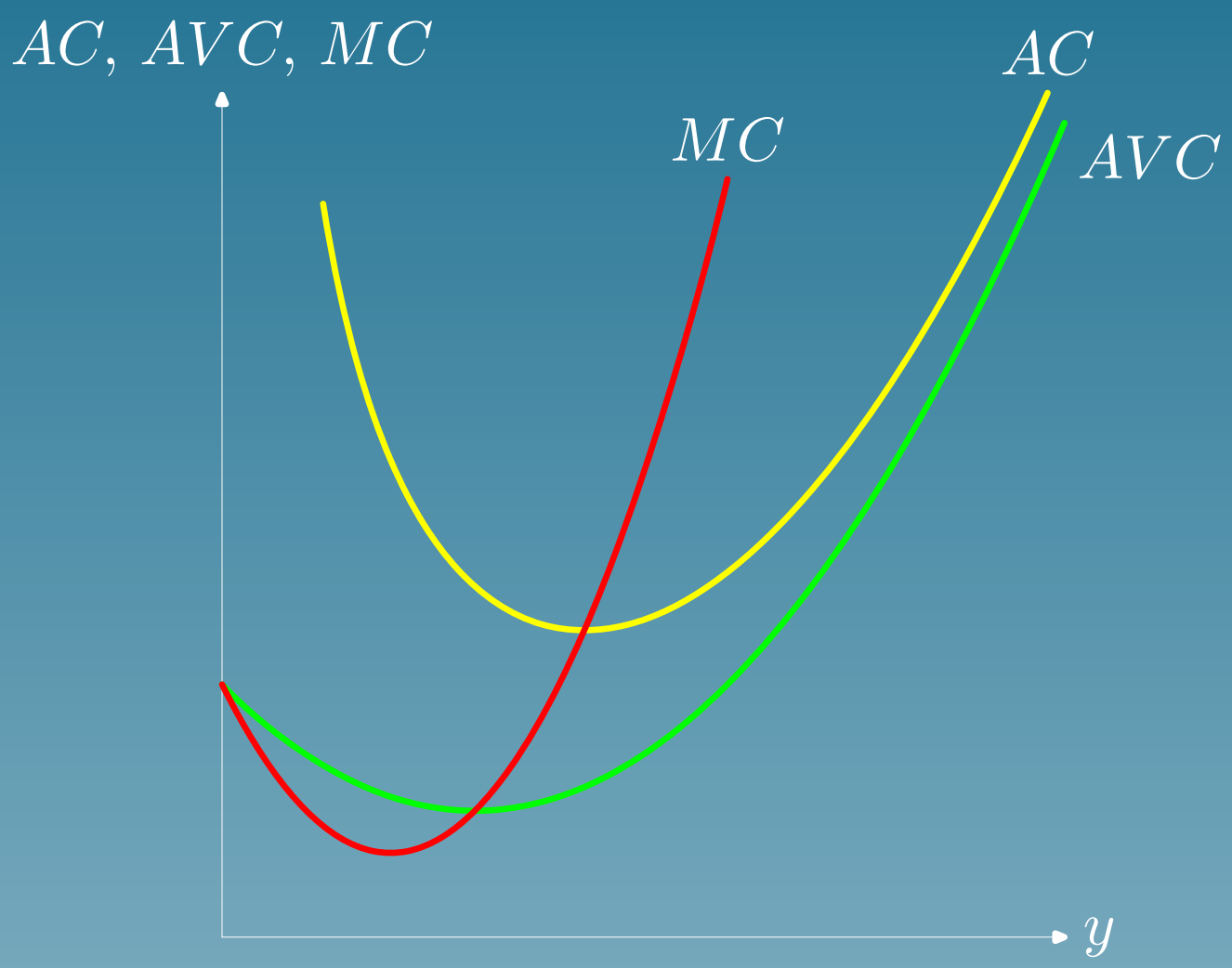
Die Ableitung der variablen Durchschnittskosten ergibt

$$\frac{\partial AVC(y)}{\partial y} = \frac{\partial \frac{c_v(y)}{y}}{\partial y} = \frac{yc_v'(y) - c_v(y)}{y^2} = \frac{1}{y}\big(MC(y) - AVC(y)\big)$$

Damit gilt mutatis mutandis dasselbe wie für die Beziehung zwischen der Durchschnittskostenkurve und der Grenzkostenkurve:

- die Grenzkostenkurve schneidet die AVC in deren Minimum,
- im Bereich fallender variabler Durchschnittskosten verlaufen die Grenzkosten unterhalb,
- im Bereich steigender variabler Durchschnittskosten verlaufen die Grenzkosten oberhalb.

Typische Konstellation von MC, AC und AVC



Grenzkosten und variable Kosten

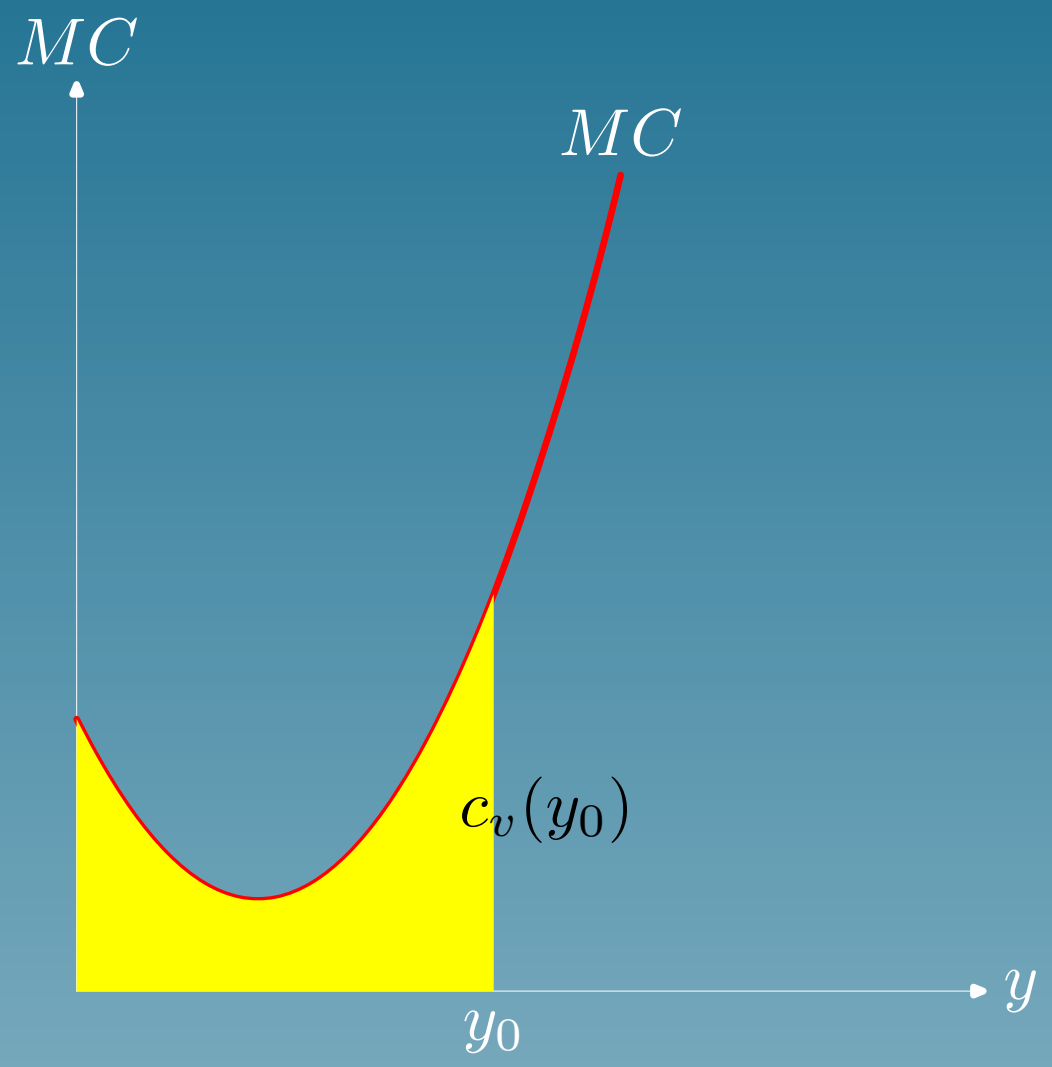
Die Grenzkosten ergeben sich als Ableitung der variablen Kosten $MC(y) = c_v'(y)$. Der Fundamentalsatz der Integralrechnung besagt nun, dass umgekehrt die variablen Kosten die Aufleitung (eine Stammfunktion) der Grenzkosten sind.

$$c_v(y) = \int MC(y)\,\mathrm{d}\,y$$

Das bestimmte Integral zwischen 0 und y ist die Fläche unter der Grenzkostenkurve. Es gilt wegen $c_v(0) = 0$

$$\int_0^y MC(y)\,\mathrm{d}\,y = c_v(y) - c_v(0) = c_v(y)$$

Fläche unter der Grenzkostenkurve = variable Kosten



Lang- und kurzfristige Kosten

Bisher wurde der konstante Input (die Betriebsgröße) bei den kurzfristigen Durchschnittskosten nicht ausdrücklich beachtet. Bei der Unterscheidung zwischen lang- und kurzfristigen Kostenkurven ist dies jedoch wesentlich.

Die kurzfristige Kostenfunktion $c_s(y, k^*)$ kann bei der Betriebsgröße k^* nicht kleiner als die langfristige Kostenfunktion $c(y)$ sein.

$$c(y) \leq c_s(y, k^*)$$

Es gibt jedoch ein Outputniveau y^*, bei welchem die Betriebsgröße k^* optimal ist und dort sind die kurz- und langfristigen Kosten gleich

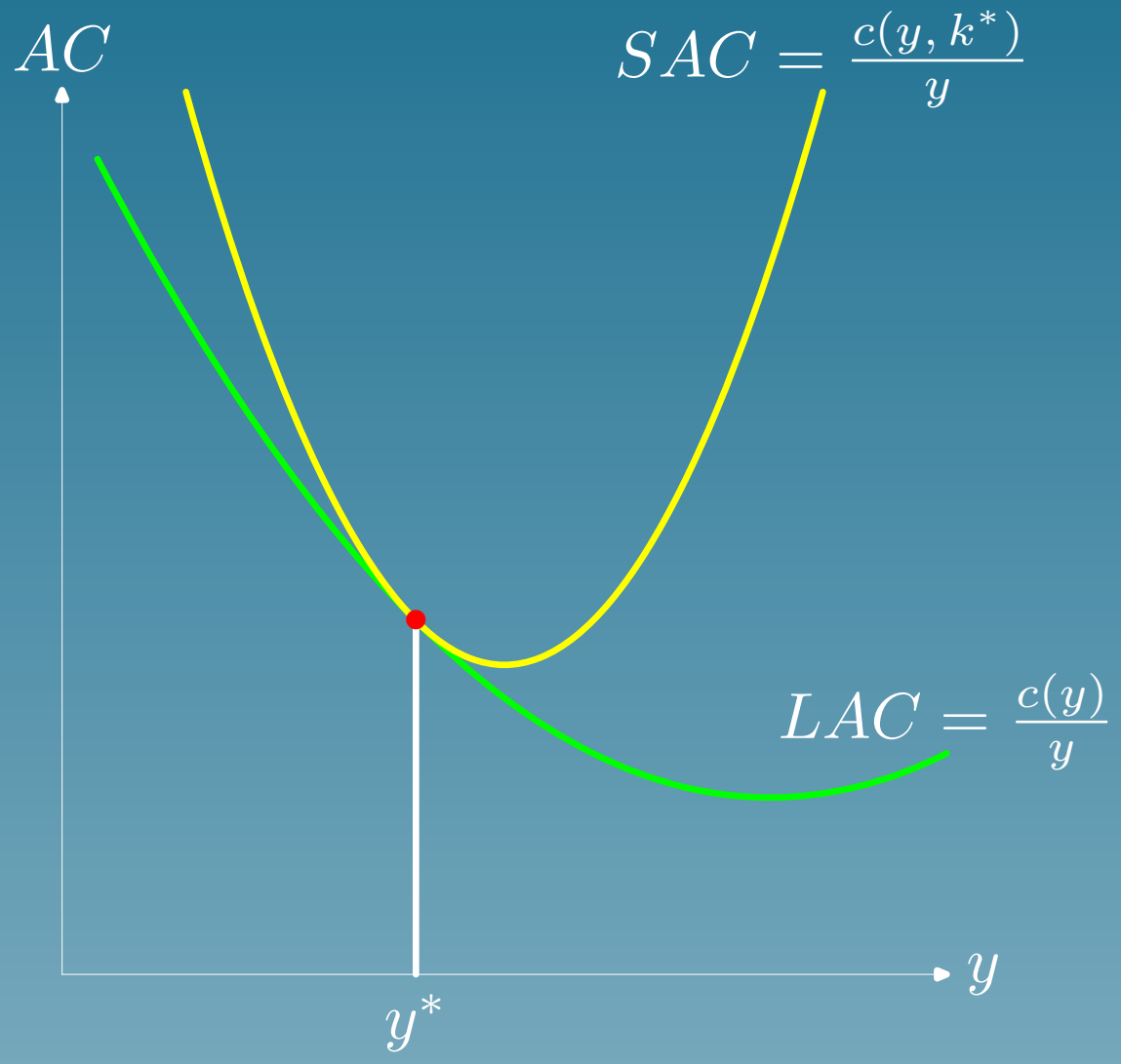
$$c(y^*) = c_s(y^*, k^*)$$

Lang- und kurzfristige Durchschnittskosten

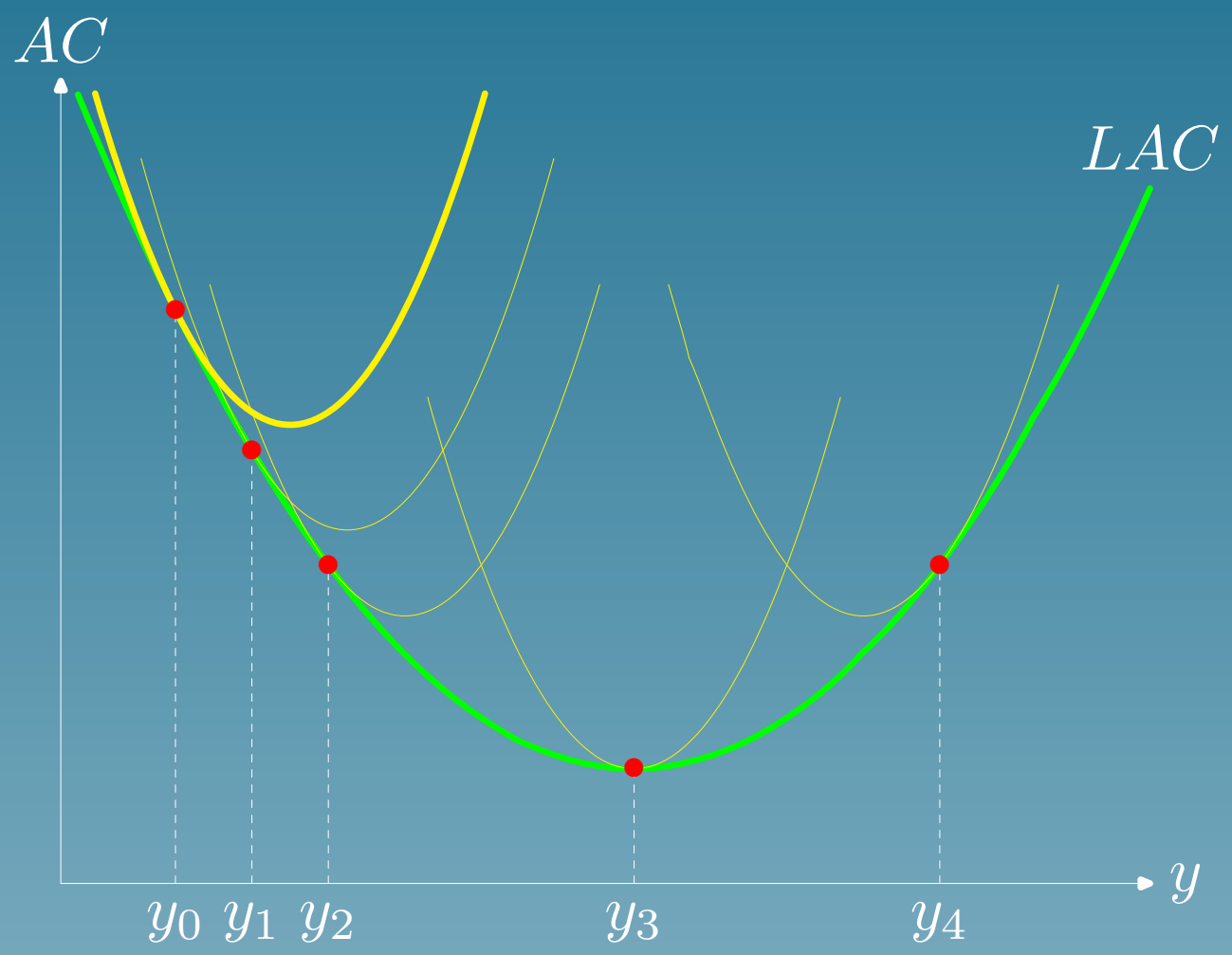
Die Beziehungen zwischen lang- und kurzfristigen Kostenkurven gelten auch für die lang- und kurzfristigen Durchschnittskostenkurven LAC und SAC.

$$LAC(y) \leq SAC(y, k^*) \quad \text{und} \quad LAC(y^*) = SAC(y^*, k^*)$$

Die SAC berührt die LAC von oben



LAC als untere Umhüllende aller SAC's zu alternativen Betriebsgrößen



Kurz- und langfristige Grenzkosten

Die langfristigen Grenzkosten eines bestimmten Outputs y sind gleich den kurzfristigen Grenzkosten bei optimaler Betriebsgröße zur Herstellung von y.

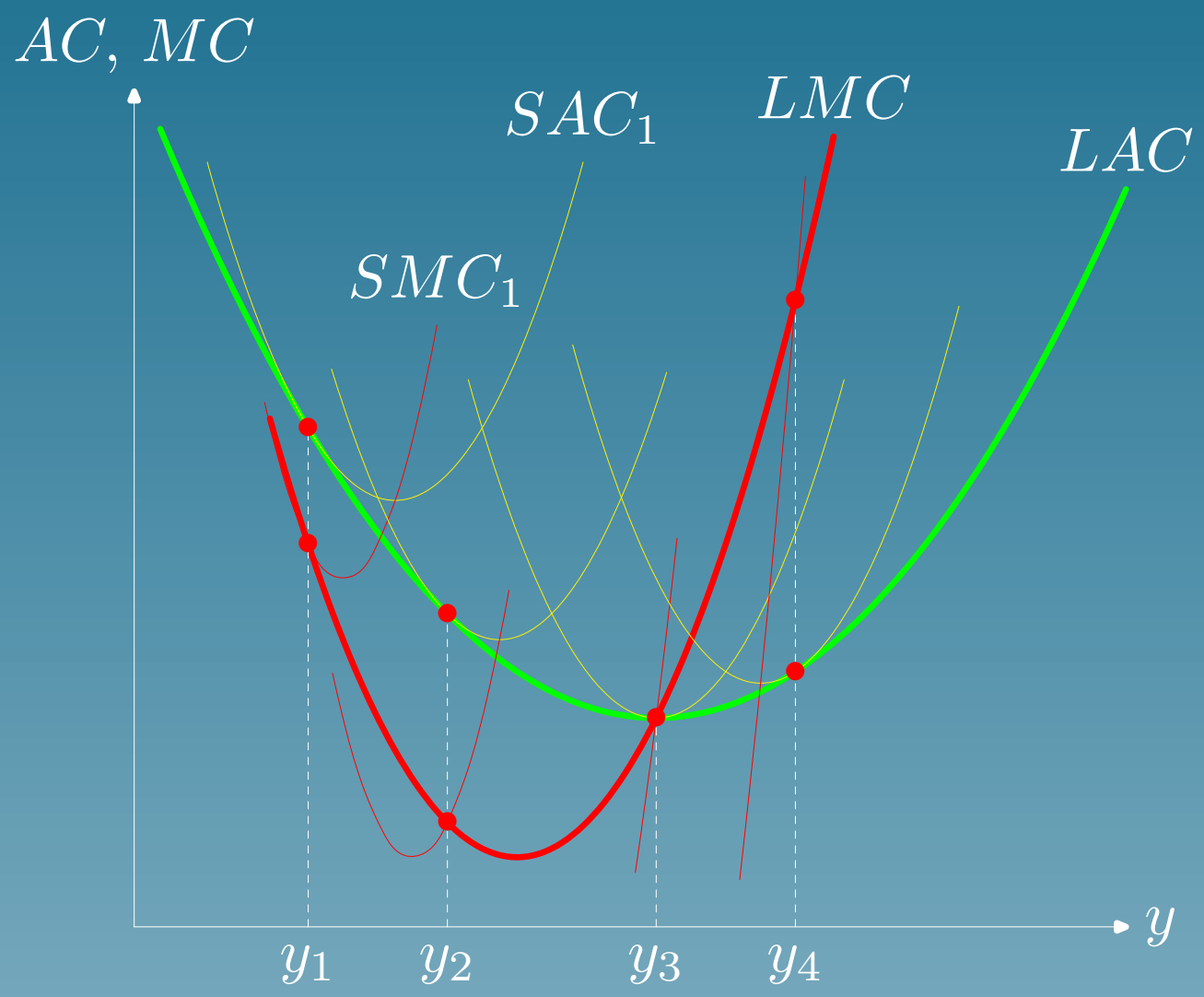
Ist k_0 die optimale Betriebsgröße zur Herstellung von y_0, dann berührt die kurzfristige Durchschnittskostenkurve $SAC(y, k_0)$ an der Stelle y_0 die langfristige Durchschnittskostenkurve:

$$SAC(y_0,\, k_0) = LAC(y_0)$$

und an dieser Stelle stimmen auch kurz- und langfristige Grenzkosten überein:

$$SMC(y_0,\, k_0) = c_s'(y_0,\, k_0) = c'\big(y_0, k(y_0)\big) = LMC(y_0)$$

Kurz- und langfristige Grenz- und Durchschnittskosten



Angebot einer Firma

Die Firma unterliegt bei ihren Entscheidungen mit dem Ziel der Gewinnmaximierung mehreren Beschränkungen:

- sie hat ihre Produktionsmöglichkeitsmenge zu beachten,
- sie konkurriert auf den Märkten für ihre Inputs mit anderen Firmen, welche diese Ressourcen anderweitig einsetzen möchten,
- sie konkurriert auf dem Markt für ihren Output mit Anbietern desselben oder sehr ähnlicher Güter.

Die Einflüsse der Technik und der Preise ihrer Inputs auf die Kosten der Firma wurden bereits besprochen. Jetzt geht es um die Einflüsse, welchen sie vom Absatzmarkt her ausgesetzt ist.

Vollständige Konkurrenz auf dem Absatzmarkt

Der Absatzmarkt der Firma soll durch vollständige Konkurrenz gekennzeichnet sein:

- es gibt sehr viele Nachfrager mit verschwindend kleinem Kaufanteil,
- es gibt neben der Firma sehr viele andere Anbieter mit jeweils verschwindend kleinem Verkaufsanteil,
- alle Marktteilnehmer sind vollständig informiert,
- seitens der Nachfrager gibt es weder sachliche noch räumliche noch zeitliche noch persönliche Präferenzen für die Angebote bzw. Anbieter.

Der Marktpreis ist ein Datum

Auf einem (vollkommenen) Markt mit vollständiger Konkurrenz kann es nur einen einheitlichen Preis geben, den bereits früher dargestellten Gleichgewichtspreis.

Alle Marktteilnehmer (auch die Firma) betrachten diesen Preis als ein Datum, auf welches sie durch ihre Käufe und Verkäufe keinen Einfluss nehmen können.

Unvollkommenheiten des Marktes geben Anlass zu Preisdiskriminierung. Dieses Thema wird später behandelt.

Die Preisabsatzfunktion der Firma

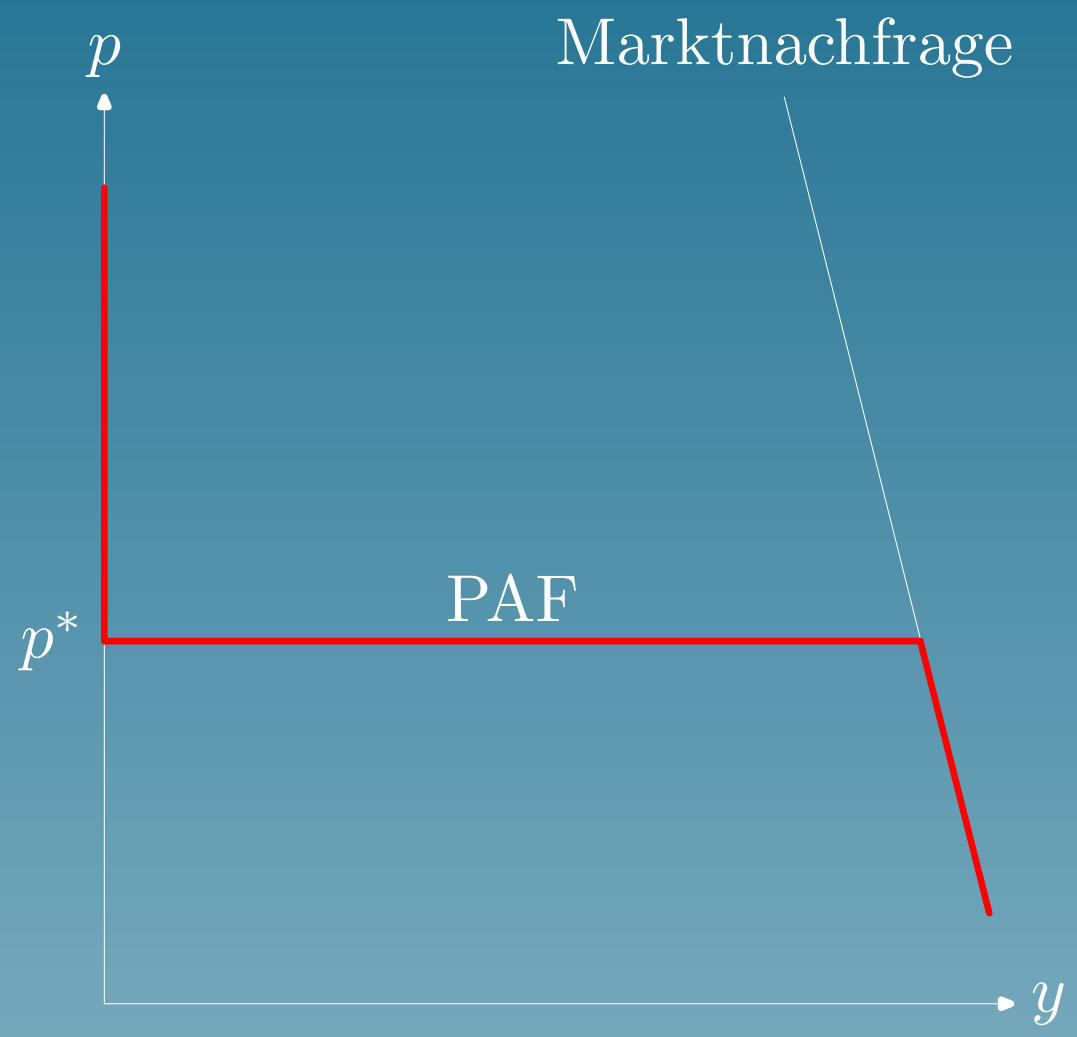
Angenommen, der Gleichgewichtspreis auf dem Absatzmarkt der Firma beträgt p^*.

Falls die Firma für ihr Produkt einen höheren Preis als p^* fordert, findet sie keinen Abnehmer.

Falls die Firma für ihr Produkt einen geringeren Preis als p^* fordert, sieht sie sich der gesamten Marktnachfrage gegenüber. Dies ist jedoch ein Güterquantum, das sehr weit oberhalb ihrer kurzfristigen Kapazitätsgrenze liegt.

Falls die Firma für ihr Produkt den Gleichgewichtspreis p^* fordert, dann kann sie — angesichts ihrer kurzfristigen Kapazitätsgrenze — beliebig viel absetzen.

Preisabsatzfunktion der Firma und Marktnachfragekurve



Die Angebotsentscheidung einer Firma

Es soll der Gewinn der Firma maximiert werden. Dieser ist nach wie vor die Differenz zwischen Erlös und Kosten. Diesmal werden aber Erlös und Kosten outputorientiert betrachtet.

$\pi(y) = py - c(y)$ (outputorientierter Gewinn)

Zur Erinnerung: der inputorientierte Gewinn wurde mittels der Produktionsfunktion $y = f(x_1, x_2)$ notiert als

$$\pi(x_1, x_2) = pf(x_1, x_2) - w_1x_1 - w_2x_2$$

Gewinnmaximierende Outputregel

Die Firma soll — bei gegebenem Marktpreis (und gegebenen Inputpreisen) — jenen Output y herstellen, bei welchem der (outputorientierte) Gewinn maximal wird:

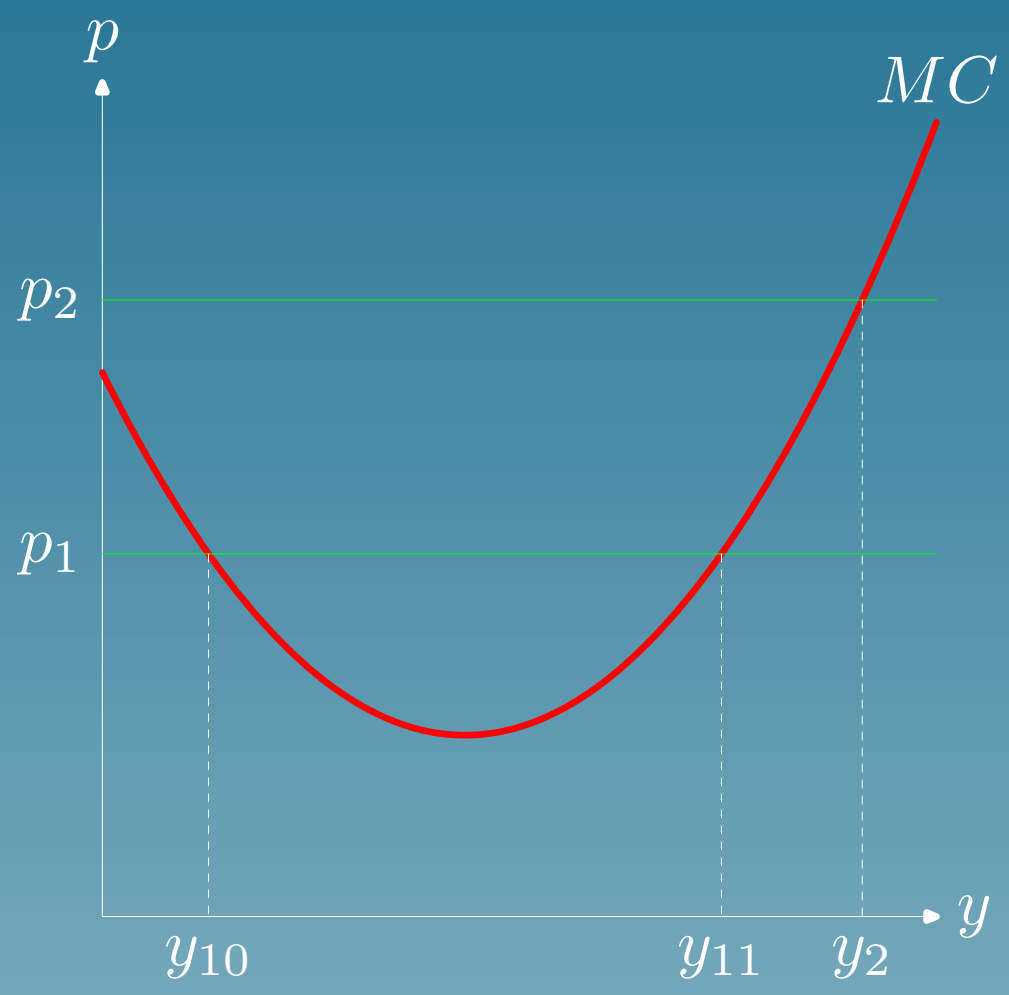
$$\max_y \pi(y) = py - c(y)$$

Die Bedingung erster Ordnung für ein Maximum verlangt, dass die Ableitung der Gewinnfunktion Null ist.

$$\pi'(y) = p - c'(y) = 0 \quad \rightsquigarrow \quad p = MC(y)$$

Das sind jene Outputniveaus, bei denen die Grenzkosten zu ihrer Herstellung gleich dem (vorgegebenen) Marktpreis sind.

Preis = Grenzkosten



Bedingung zweiter Ordnung

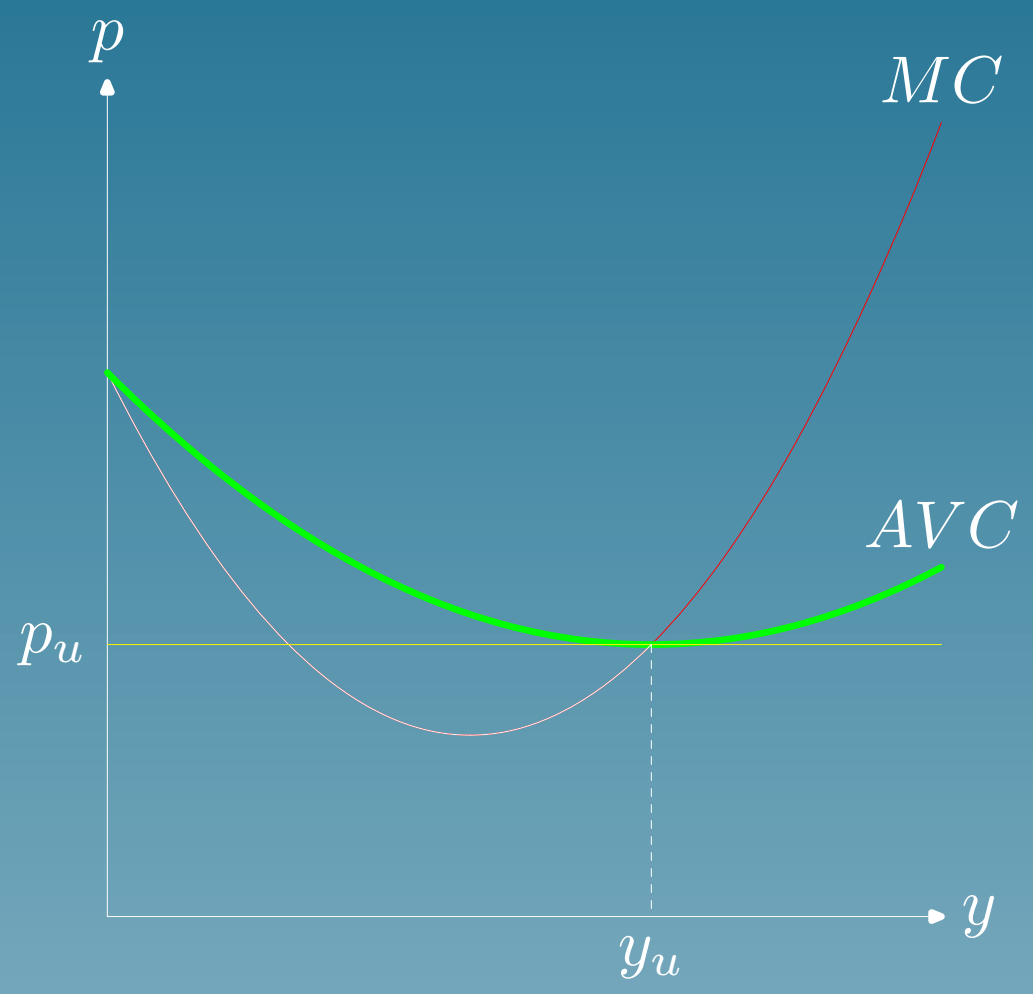
Die Bedingung zweiter Ordnung verlangt, dass an den kritischen Stellen die zweite Ableitung des Gewinnes negativ ist.

$$\pi''(y) = -MC'(y) < 0 \quad \rightsquigarrow \quad MC'(y) > 0$$

Daher sind bei gegebenem Marktpreis nur die Outputs y auf dem ansteigenden Teil der Grenzkostenkurve gewinnmaximal.

In der Abbildung ist also beim Marktpreis p_1 der Output y_{11} gewinnmaximal, wohingegen der Output y_{10} der Firma einen maximalen *Verlust* bescheren würde.

Die kurzfristige Angebotsschwelle



Falls der Marktpreis kleiner als das Minimum der kurzfristigen Durchschnittskosten ist, wird die Firma nichts produzieren.

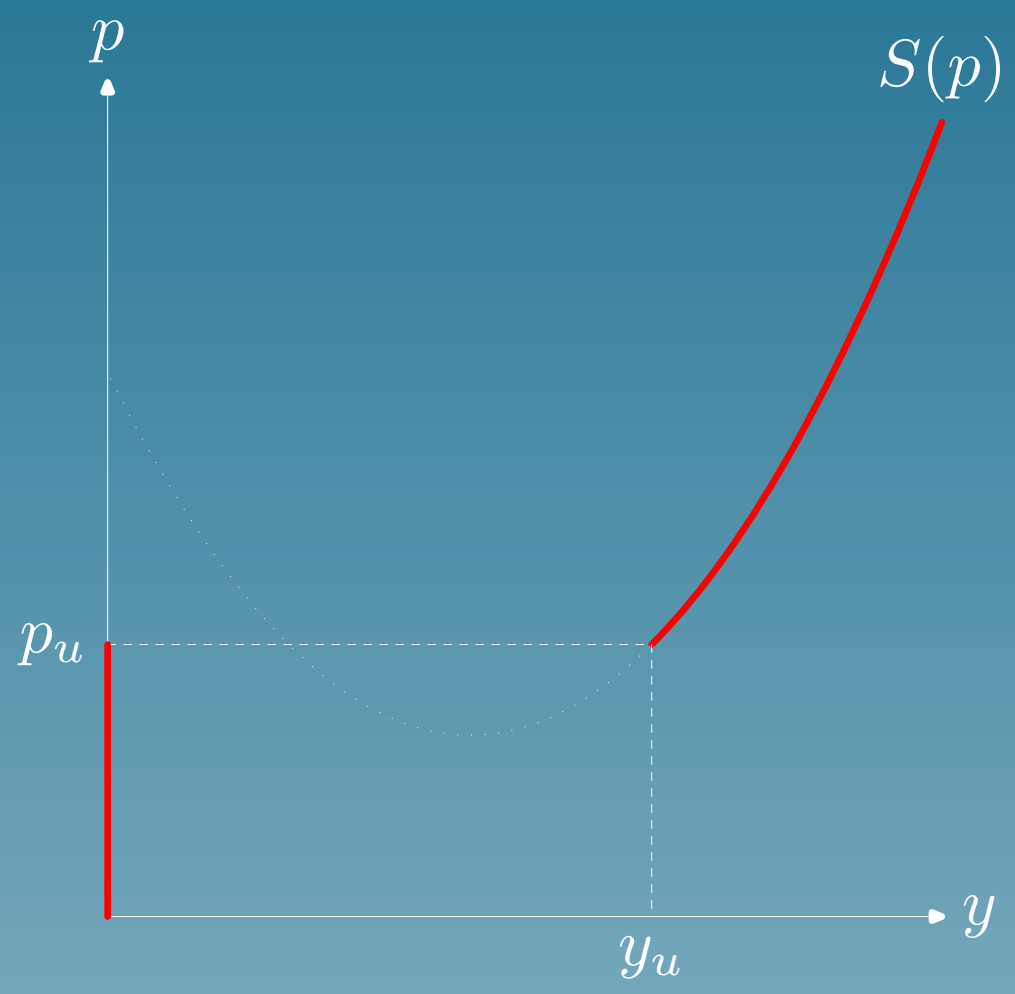
In diesem Fall deckt der Preis *nicht* die variablen Stückkosten, sodass bei einer Produktion $y > 0$ der Verlust größer als die Fixkosten wäre.

$$p < AVC_{\min} \implies py < c_v(y) \implies \pi(y) < -F = \pi(0)$$

Darum wird die gewinnmaximierende Firma die Produktion einstellen, wenn der Preis unter dem Minimum der AVC liegt.

Umgekehrt lautet die Bedingung für die Beteiligung am Marktangebot für eine gewinnmaximierende Firma, dass der Marktpreis auf oder oberhalb ihrer sogenannten Angebotsschwelle $AVC_{\min}$ verläuft.

Kurzfristige Angebotskurve



Die kurzfristige Angebotskurve $y = S(p)$ ordnet einem (beliebigen) Marktpreis p den gewinnmaximalen Output der Firma zu.

Für Preise unterhalb der Angebotsschwelle p_u ist der gewinnmaximale Output gleich Null und für Preise auf oder oberhalb der Gewinnschwelle ist der Output jener, bei dessen Herstellung gerade Grenzkosten in Höhe des Preises anfallen.

$$y = S(p) = \begin{cases} 0 & \text{für } \; p < p_u \\ MC^{-1}(p) & \text{für } \; p \geq p_u \end{cases}$$

Beispiel 5

Angenommen, die kurzfristige Kostenfunktion der Firma lautet

$$c(y,\, k) = \frac{1}{3}y^3 - y^2 + \frac{3}{2}y + 2$$

Dann ist die kurzfristige Grenzkostenkurve gegeben durch

$$MC(y) = c'(y,\, k) = y^2 - 2y + \frac{3}{2}$$

Die kurzfristige Angebotskurve ist die Umkehrfunktion dieser Grenzkosten, die oberhalb der Angebotsschwelle verlaufen. Bevor wir also die Angebotskurve berechnen können, sind p_u und y_u der Firma zu bestimmen.

Die kurzfristigen variablen Durchschnittskosten belaufen sich auf

$$AVC(y) = \frac{1}{3}y^2 - y + \frac{3}{2}$$

und sind minimal, für

$$AVC'(y) = 0 = \frac{2}{3}y - 1 \quad \rightsquigarrow \quad y_u = \frac{3}{2}$$

Die kurzfristige Angebotsschwelle der Firma ist somit

$$p_u = AVC(y_u) = \frac{1}{3}\frac{9}{4} - \frac{3}{2} + \frac{3}{2} = \frac{3}{4}$$

Nun ist die outputorientierte Bedingung für den maximalen Gewinn $p = MC(y)$ für Preise oberhalb der Gewinnschwelle nach y aufzulösen.

MaW ist die in y quadratische Gleichung $y^2 - 2y + \frac{3}{2} - p = 0$ für $p \geq p_u$ zu lösen. Es ergibt sich

$$y = 1 + \sqrt{p - \frac{1}{2}}$$

Damit lautet die kurzfristige Angebotskurve der Firma

$$y = S(p) = \begin{cases} 0 & \text{für } \; p < \frac{3}{4} \\ 1 + \sqrt{p - \frac{1}{2}} & \text{für } \; p \geq \frac{3}{4} \end{cases}$$

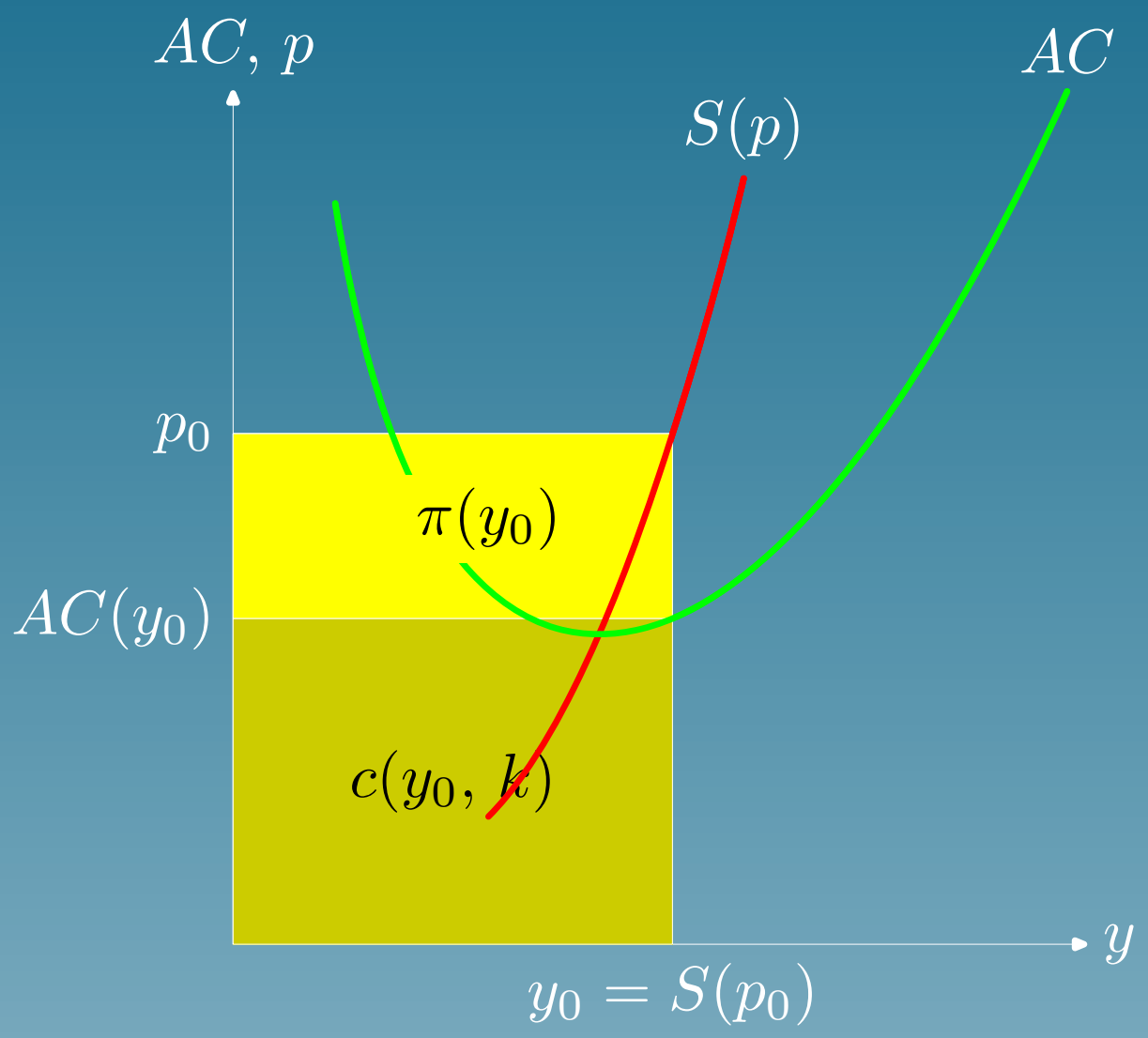
Flächenbetrachtungen

Die kurzfristigen Durchschnitts- und Grenzkostenkurven gestatten es zu einem gegebenen Marktpreis

- den Gewinn und
- die Produzentenrente

als Flächen darzustellen.

Gewinn als Flächenstück



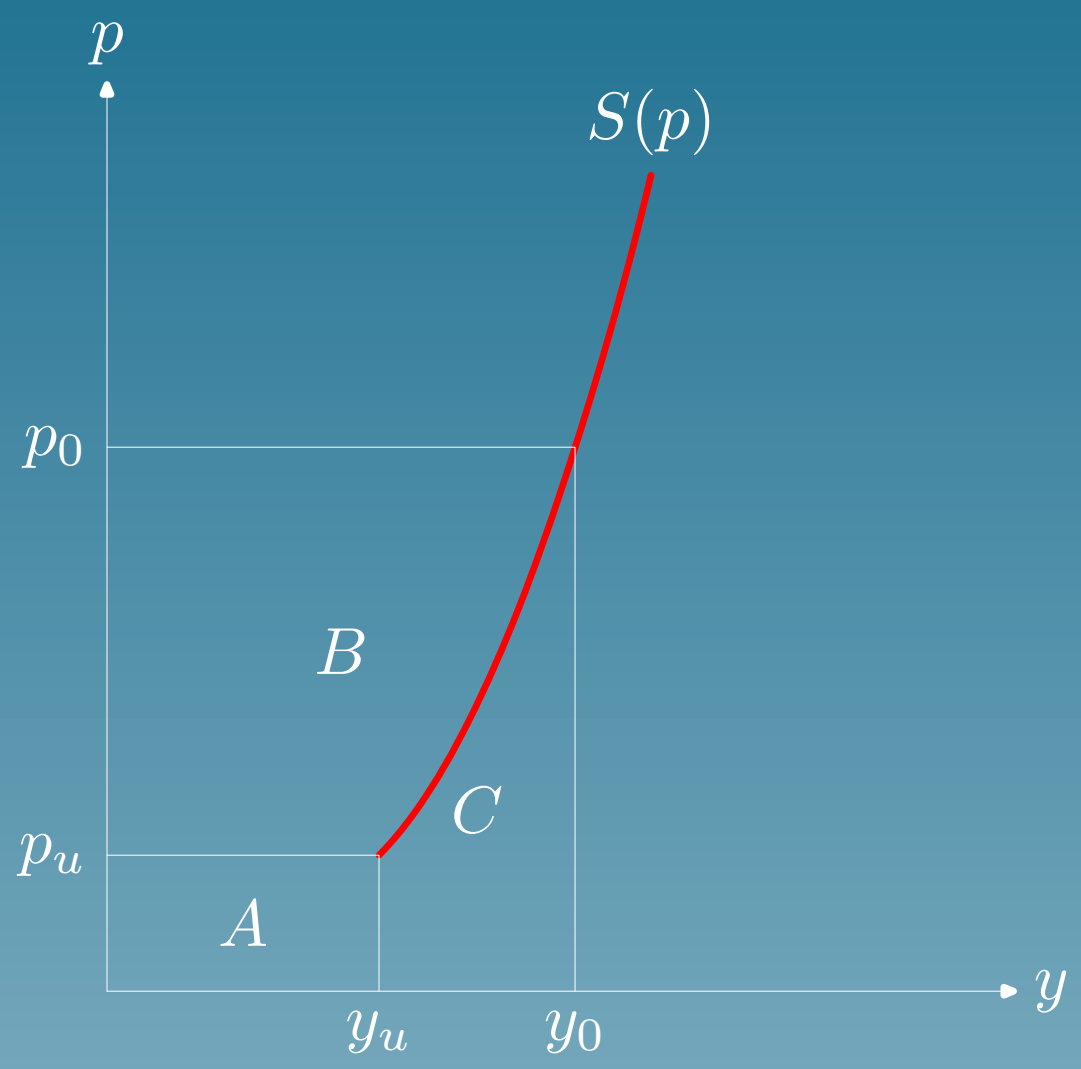
Produzentenrente

Die kurzfristige Produzentenrente der Firma ist, solange sie als Produzent oder Anbieter aktiv ist — analog zur Konsumentenrente — die Fläche zwischen der Preis-Achse und ihrer kurzfristigen Angebotskurve.

Falls der Marktpreis unter ihrer Angebotsschwelle liegt, produziert sie nicht, folglich gibt es auch keine sinnvolle Produzentenrente.

Anders als bei der Konsumentenrente, die aufgrund der ordinalen Nutzenfunktion nur sehr allgemein charakterisiert werden kann, ist bei der Produzentenrente eine direkte Beziehung zum Gewinn der Firma gegeben.

Zerlegung des Erlöses in Teilflächen



Der Erlös der gewinnmaximierenden Firma beim Preis $p_0 > p_u$ ist gleich der Rechtecksfläche $p_0 y_0$. Diese kann in die drei Teilflächen A, B und C zerlegt werden, wobei die Teilfläche B die Produzentenrente ist.

$$p_0 y_0 = A + B + C$$

Wir berechnen die Teilflächen A und C, um einen expliziten Ausdruck für die Produzentenrente B zu erhalten.

$$A = p_u y_u = AVC(y_u)\, y_u = c_v(y_u)$$

Die Teilfläche A ist gleich den minimalen variablen Kosten.

$$C = \int_{y_u}^{y_0} MC(y)\,\mathrm{d}\,y = c_v(y_0) - c_v(y_u)$$

Die Teilfläche C ist die Fläche unter der Grenzkostenkurve, und als bestimmtes Integral gleich der Differenz der variablen Kosten an Ober- und Untergrenze des Integrals.

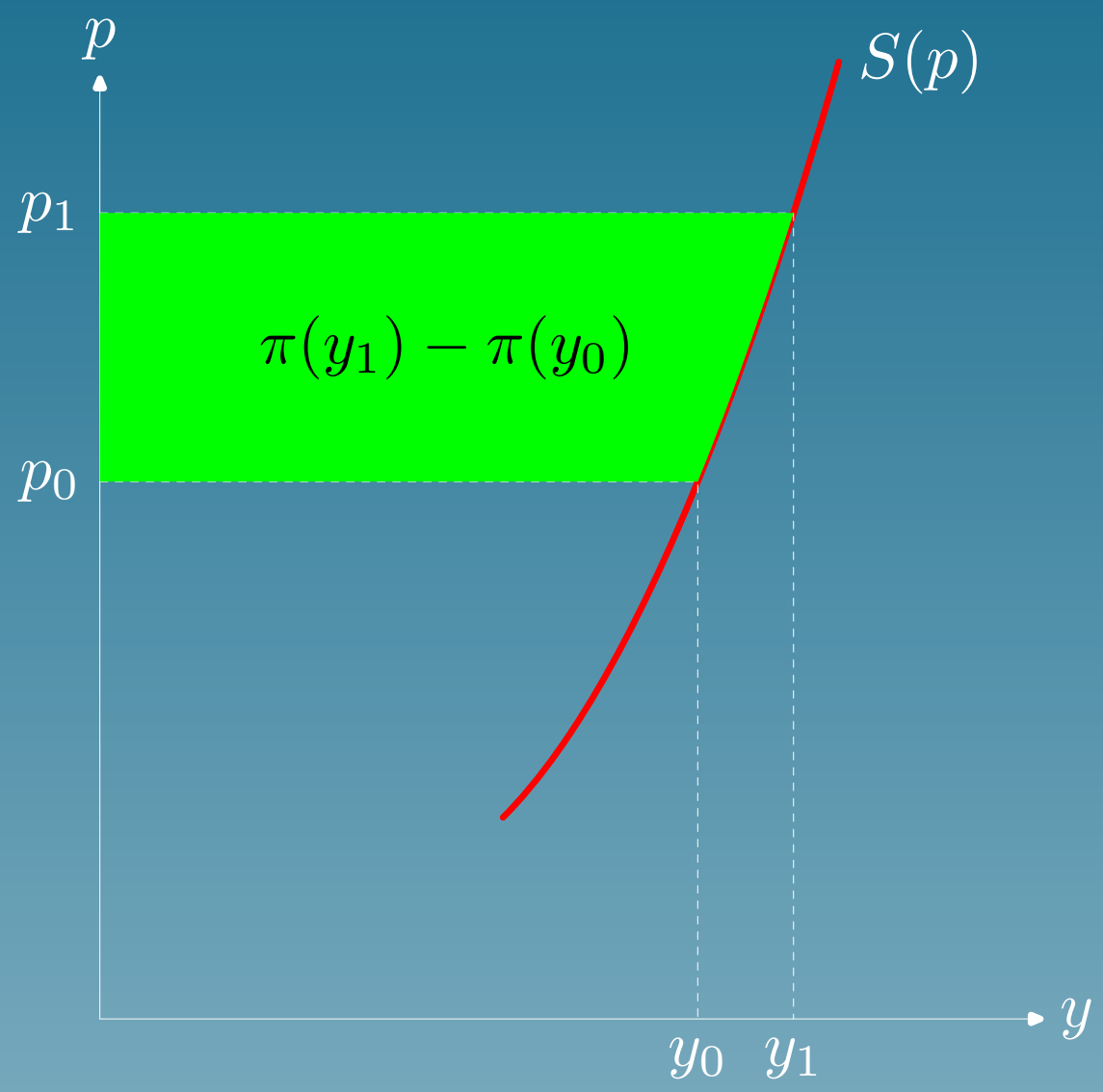
Die Produzentenrente ergibt sich somit als

$$
\begin{aligned}
B &= p_0 y_0 - A - C \\
&= p_0 y_0 - c_v(y_u) - c_v(y_0) + c_v(y_u) \\
&= p_0 y_0 - c_v(y_0) = p_0 y_0 - c(y_0) + F \\
&= \pi(y_0) + F
\end{aligned}
$$

Summe aus dem Gewinn $\pi(y_0)$ und den Fixkosten.

Die Veränderung der Produzentenrente, welche durch eine Preisänderung ausgelöst wird, ist daher gleich der Gewinnänderung.

Veränderung der Produzentenrente = Gewinnänderung



Langfristiges Angebot einer Firma

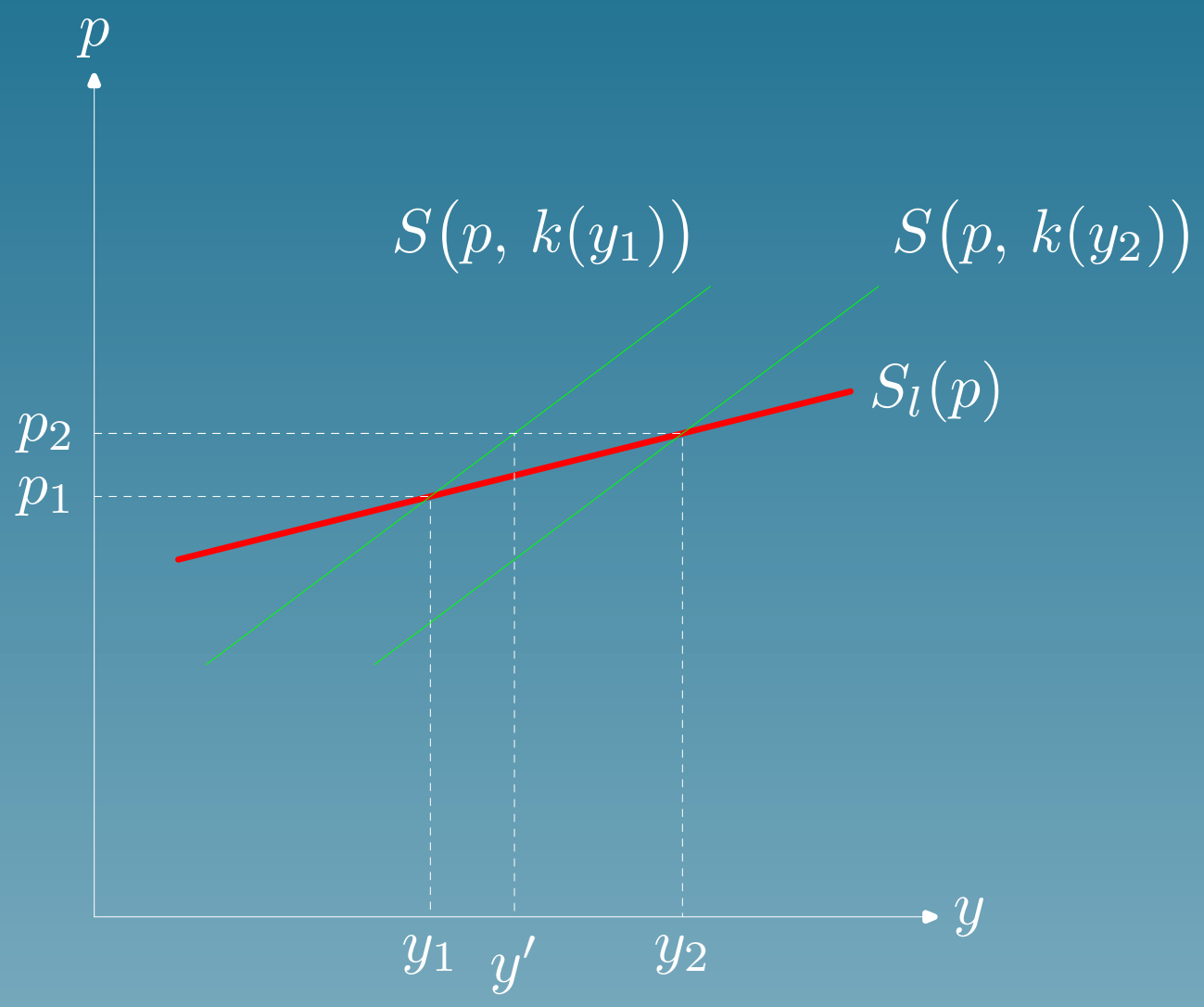
Langfristig gibt es keine Fixkosten, da alle Inputs veränderlich sind.

Die Firma wird bei gegebenem Marktpreis langfristig jenen Output produzieren, bei dem die langfristigen Grenzkosten gleich diesem Preis sind.

$$p = LMC(y) = MC\big(y, k(y)\big)$$

MaW auch die Betriebsgröße wird langfristig optimal angepasst.

Kurz- und langfristige Angebotskurven



Langfristige Angebotsschwelle

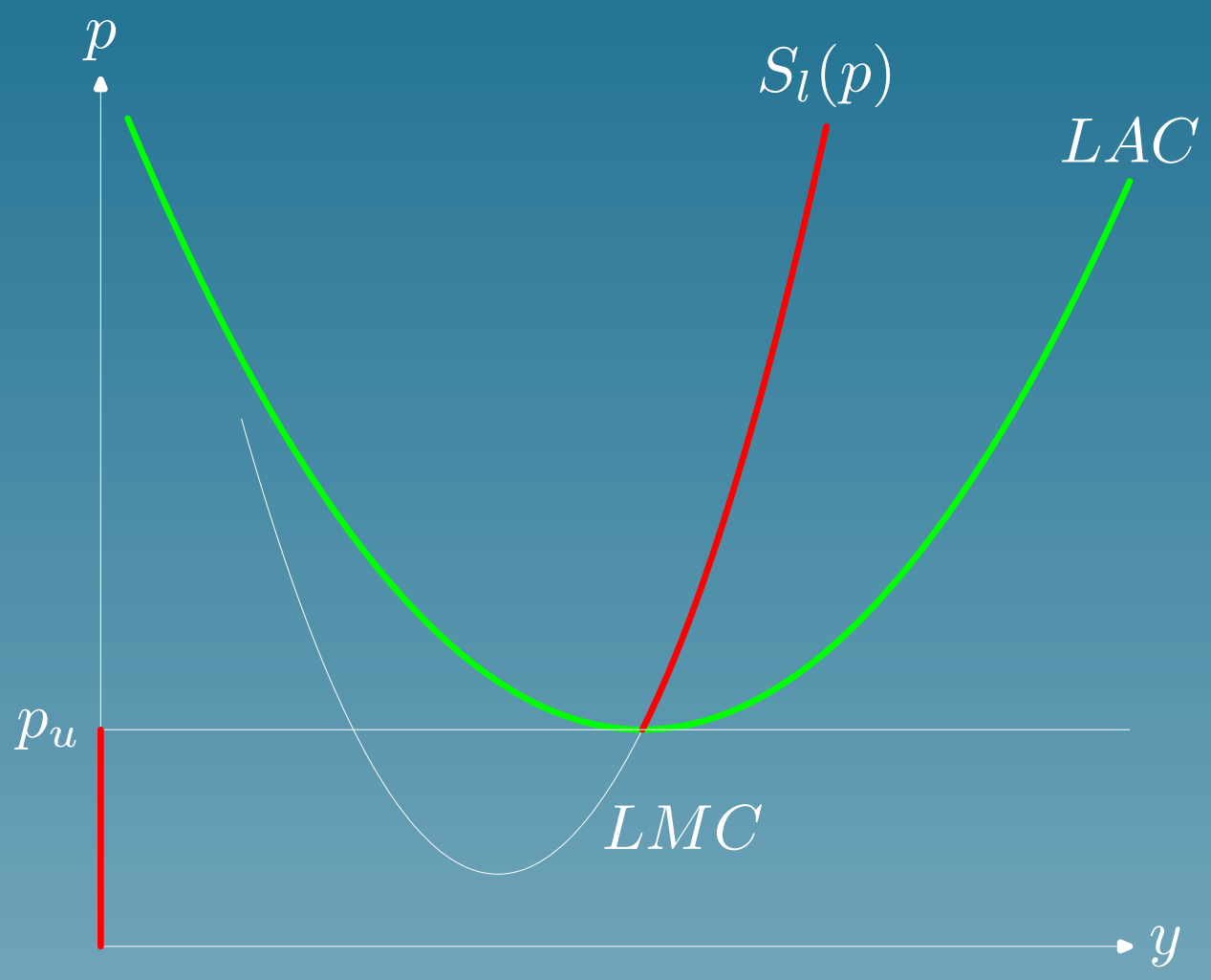
Die Firma wird sich langfristig nur dann als Produzent und Anbieter betätigen, wenn sie einen nicht-negativen Gewinn erwirtschaftet.

$$\pi = py - c(y) \geq 0 \quad \rightsquigarrow \quad p \geq \frac{c(y)}{y} = LAC(y)$$

Die langfristige Angebotskurve ist daher der Teil der langfristigen Grenzkostenkurve, der über der langfristigen Durchschnittskostenkurve verläuft.

Beachte: auch bei einem langfristigen Gewinn in Höhe von Null werden alle an der Produktion beteiligten Inputs mit ihren jeweiligen Inputpreisen entlohnt. Ein positiver Gewinn zeigt daher an, dass manche Inputs (die Eigentümer der Firma bei Privateigentum) mehr als ihre „normale“ Entlohnung erhalten.

Langfristige Angebotskurve



Marktangebot einer Industrie

Unter Industrie sind alle Firmen zu verstehen, die Produkte herstellen, welche sich in den Augen der Verbraucher nicht unterscheiden.

Die Produktion in einer Industrie ist daher das Marktangebot für ein bestimmtes Gut.

Die Marktangebotskurve, welche zu Beginn von Mikro 2 einfach unterstellt wurde, kann jetzt — analog zur Marktnachfragekurve — aus den Angebotskurven aller Firmen der Industrie durch horizontale Aggregation bestimmt werden.

Wie üblich wird dabei zwischen dem kurz- und dem langfristigen Marktangebot einer Branche unterschieden.

Aggregation der Angebotskurven einzelner Firmen

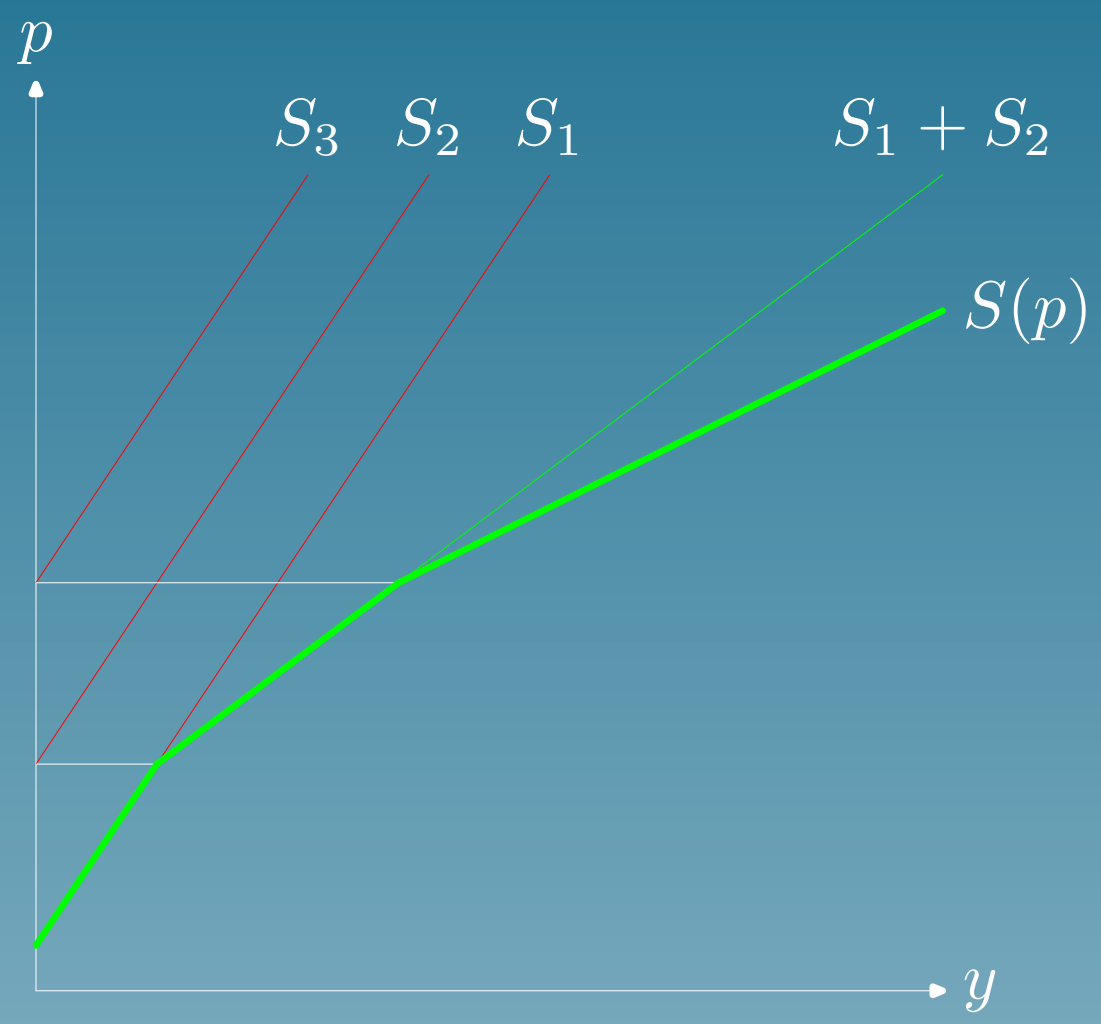
Angenommen, es gibt kurzfristig n Firmen in der Industrie. Die Angebotsfunktion der i-ten Firma wird mit $S_i(p)$ bezeichnet.

Die Summe über alle Angebotsmengen — jeweils für einen bestimmten Preis p — ergibt die Marktangebotskurve $S(p)$.

$$S(p) = \sum_{i=1}^{n} S_i(p)$$

Geometrisch ergibt sich die Angebotskurve der Branche durch eine horizontale Addition aller individueller Angebotskurven.

Horizontale Aggregation individueller Angebotskurven



Kurzfristiges Marktgleichgewicht

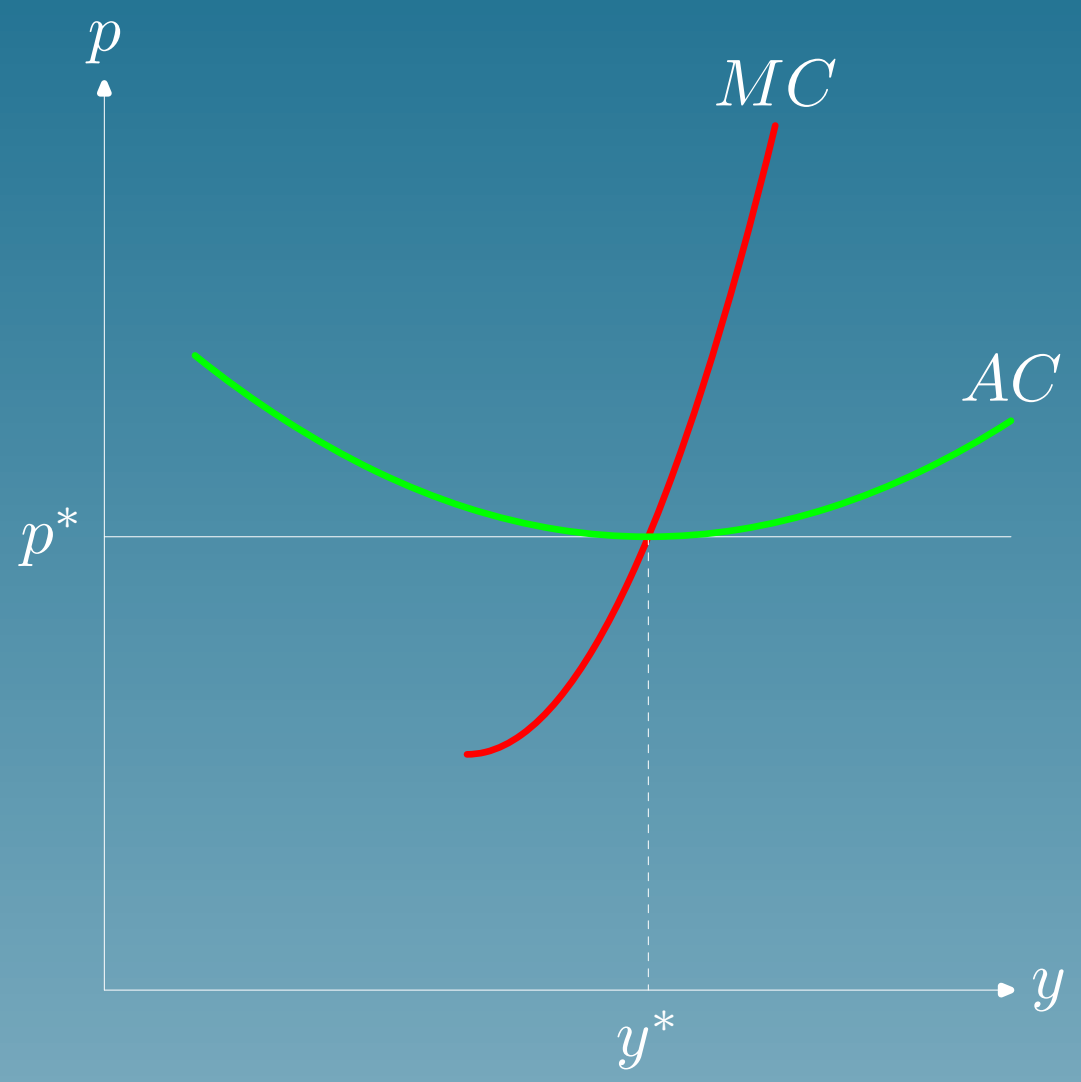
Das kurzfristige Marktgleichgewicht ist der Schnittpunkt zwischen der Marktnachfragekurve und der kurzfristigen Angebotskurve der Industrie.

Beim Gleichgewichtspreis p^* kann auf die individuellen Angebote der Firmen zurückgeschlossen werden.

Die kurzfristigen Gewinne einzelner Firmen können beim Gleichgewichtspreis p^*

- Null sein, diese werden *Marginalanbieter* genannt;
- positiv sein, diese werden *Intramarginalanbieter* genannt;
- negativ sein, diese Firmen werden langfristig nicht mehr als Anbieter auftreten.

Situation eines Marginalanbieters

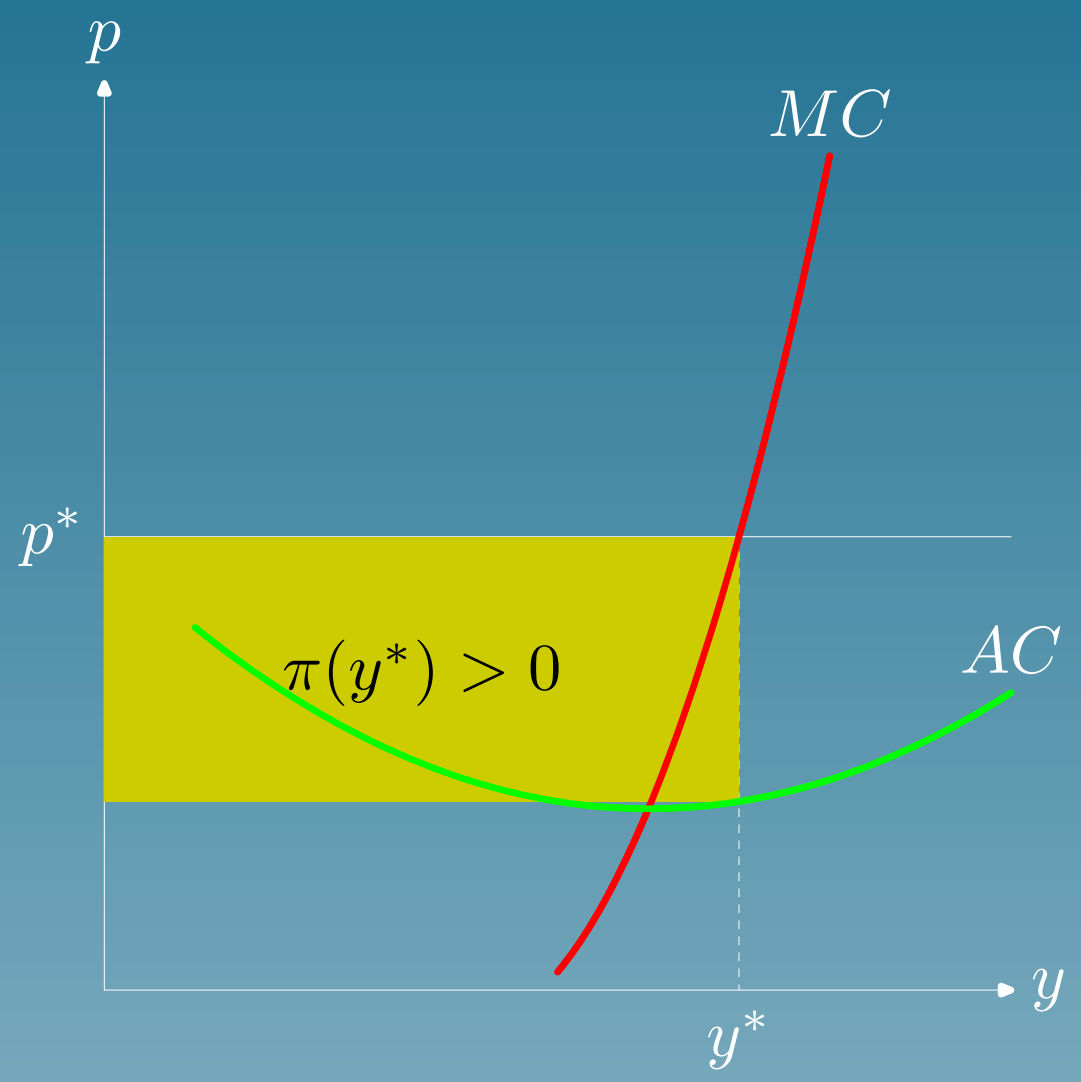


Der Marktpreis p^* ist gleich dem Minimum der Durchschnittskosten des Marginalanbieters. Dieser Preis im Schnittpunkt von kurzfristiger Grenzkosten- und Durchschnittskostenkurve wird *Gewinnschwelle* genannt.

Da beim Marginalanbieter $p^* = c(y^*, k)/y^*$ gilt, ist sein Gewinn gleich Null

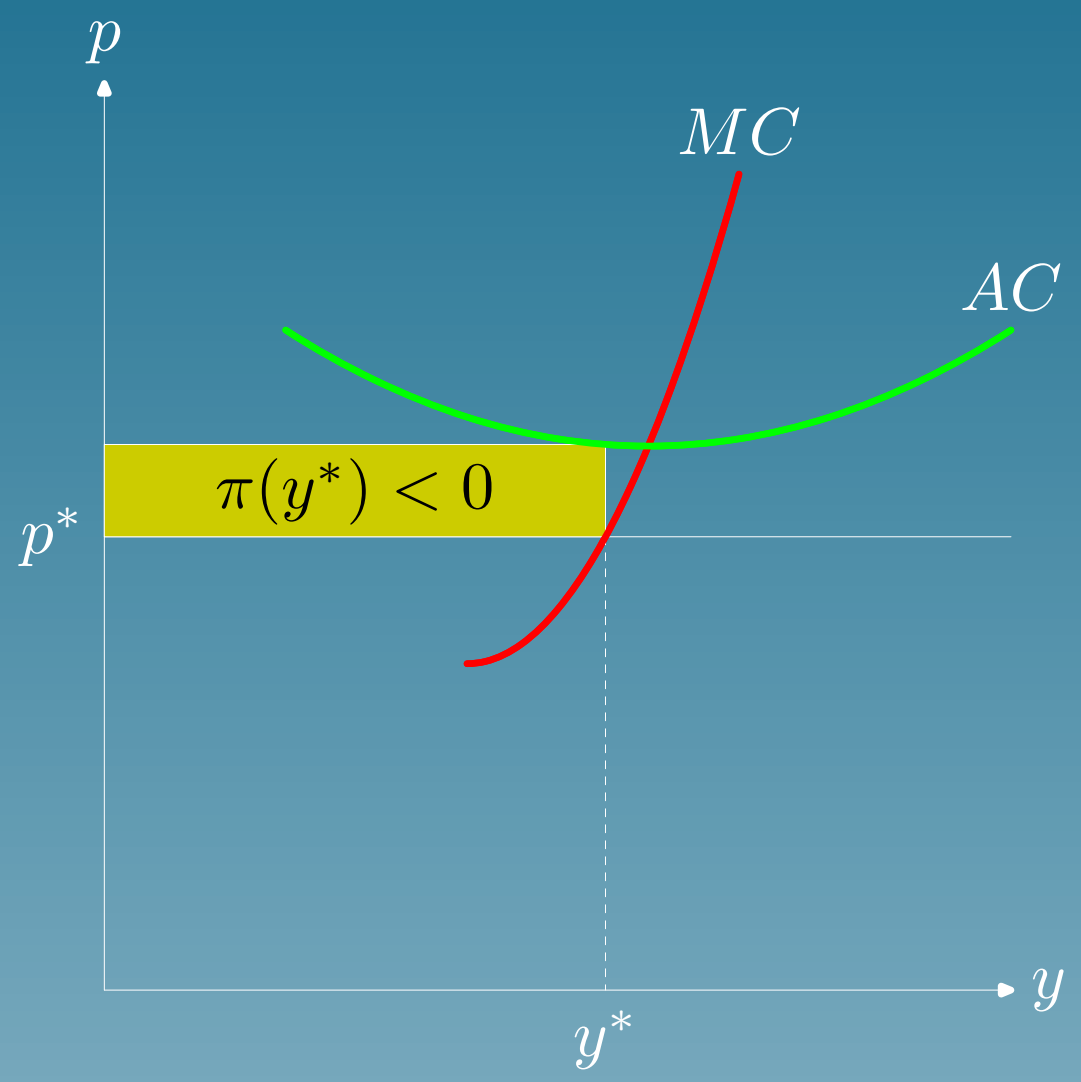
$$\pi(y^*) = p^* y^* - c(y^*, k) = 0$$

Intramarginalanbieter



Der Marktpreis p^* ist größer als das Minimum der Durchschnittskosten bei einem Intramarginalanbieter. Sein volkswirtschaftlicher Gewinn ist positiv.

Kurzfristiger Verlust



Liegt das Durchschnittskostenminimum einer Firma unter dem Marktpreis p^*, so macht sie einen Verlust.

Falls es einem solchen Anbieter nicht gelingt, langfristig durch Anpassung der fixen Inputs den Verlust zu beseitigen, wird er aus dem Markt ausscheiden.

Langfristige Marktzu- und austritte

Im langfristigen Marktgleichgewicht haben die Firmen ihre Betriebsgröße an den Marktpreis angepasst.

Weiter kann es jedoch langfristig zu Marktaustritten und Marktzutritten von Firmen kommen.

In den meisten Industrien gibt es einen freien Marktzugang. In manchen gibt es aber auch Zutrittsbarrieren.

Märkte, bei denen unbeschränkter Zutritt möglich ist, werden auch *umkämpft* genannt.

Repräsentative Firma in der langen Frist

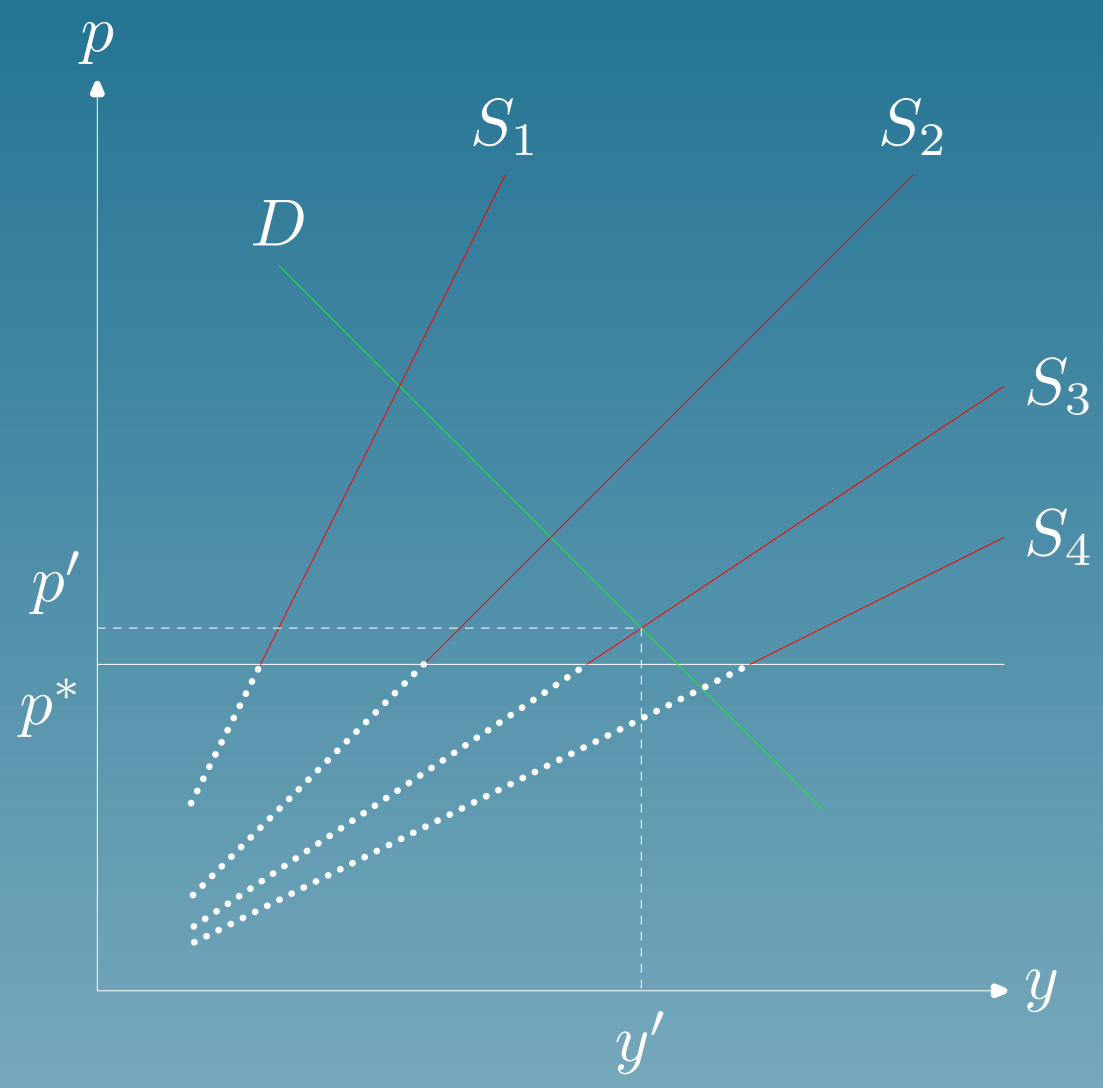
Es wird unterstellt, dass alle potenziellen Anbieter dieselbe Produktionstechnik benutzen, und daher identische langfristige Kostenkurven $c(y)$ haben.

Das Minimum der langfristigen Durchschnittskosten wird beim Output y^* einer repräsentativen Firma erreicht. Die langfristige Preisuntergrenze — die langfristige Angebotsschwelle der Firma — ist somit bestimmt durch

$$p^* = \frac{c(y^*)}{y^*}$$

Zur weiteren Vereinfachung wird unterstellt, dass die langfristigen Grenzkosten linear sind.

Langfristige Anzahl der Anbieter

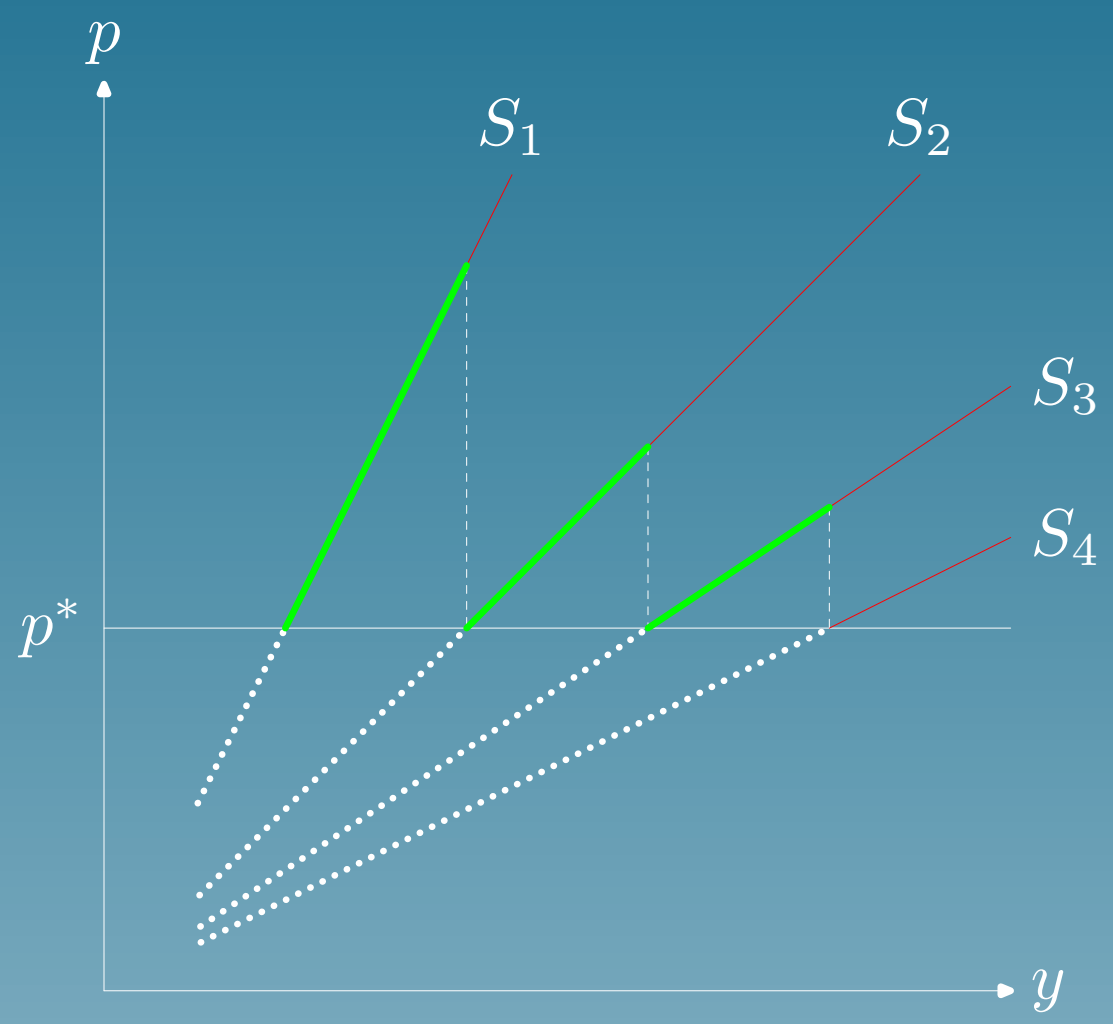


Im dargestellten Marktdiagramm können langfristig nur 3 Anbieter existieren. Die (horizontal) aggregierte langfristige Angebotskurve ist S_3.

Sollte ein vierter Anbieter hinzukommen, würde der langfristige Marktpreis unter die Angebotsschwelle sinken und damit für jede Firma der Industrie zu Verlusten führen, was jedoch mit einem langfristigen Gleichgewicht nicht vereinbar ist.

Der langfristige Gleichgewichtspreis p' liegt jedoch in der Nähe der Angebotsschwelle p^*.

Langfristige Angebotskurve bei identischen Anbietern



Langfristiges Marktgleichgewicht

Die langfristige Marktangebotskurve ist (annähernd) eine horizontale Linie, ihr Abstand von der Mengen-Achse ist das Minimum der Gewinnschwellen aller Firmen einer Industrie (also das Minimum aller Durchschnittskostenminima).

Falls es kurzfristig Anbieter mit einer höheren Gewinnschwelle gibt, so werden sie auf Dauer ausscheiden, oder auf die Technik mit den geringeren Kosten umsteigen.

Liegt dagegen der kurzfristige Gleichgewichtspreis noch über der Gewinnschwelle einiger Anbieter, so wird der überdurchschnittliche Kapitalertrag neue Anbieter auf den Markt locken, mit der Folge, dass sich die kurzfristige Angebotskurve nach rechts verlagert.

Langfristiger Gleichgewichtspreis

Der sich langfristig einstellende Marktpreis bei vollkommener Konkurrenz ist bestimmt durch das gesellschaftlich niedrigste Durchschnittskostenminimum. Daher werden die Nachfrager zu den gesellschaftlich geringsten Kosten mit dem Gut versorgt. Die Nachfrage hat zwar Einfluss auf die am Markt abgesetzte Menge, aber nicht mehr auf den Preis.

Offensichtlich ist dieser langfristig sich durchsetzende Wettbewerb um die kostengünstigste Herstellung des Gutes eine Eigenschaft der Konkurrenz, die auch wirtschaftspolitisch erwünscht ist. Dies kommt auch im Leitbild des funktionsfähigen Wettbewerbs zum Ausdruck.

Beschränkter Marktzutritt

Es gibt verschiedene Ursachen für einen beschränkten Marktzugang.

Alle hängen aber mehr oder weniger direkt damit zusammen, dass die kurzfristig nicht variierbaren Inputs auch langfristig fixiert sind und damit verhindern, dass neue Firmen hinzukommen.

Die Beschränkung kann durch die „Natur“ vorgegeben sein, wie beispielsweise bei den nicht-regenerierbaren natürlichen Rohstoffen (Metalle, fossile Energieträger etc), oder staatlich verordnet sein, wie beispielsweise die Einschränkung der Niederlassungsfreiheit bei Apotheken und Ärzten.

Langfristiges Gleichgewicht bei beschränktem Marktzutritt

Wenn der Marktzutritt frei ist, wird der Wettbewerb langfristig dazu führen, dass sich der Preis auf das niedrigste Niveau aller Durchschnittskostenminima der beteiligten Firmen einpendelt und alle Gewinne verschwinden.

Ist der Marktzutritt dagegen beschränkt, dann kommt der Wettbewerb um die kostengünstigste Technik nicht zustande und einige Firmen erzielen auf den ersten Blick einen positiven Gewinn.

Dieser erste Eindruck ist jedoch falsch. Diese Firmen setzen Inputs ein, die anderen nicht zugänglich sind und diese Inputs erzielen sogenannte Renten.

Theorie des Marktes

Die Haushaltstheorie lieferte eine Marktnachfragekurve. Die Unternehmenstheorie lieferte eine individuelle Angebotskurve. Die Markttheorie soll nun die *Koordination* dieser unterschiedlichen Pläne erklären.

Es wird das Zusammenspiel von Angebot und Nachfrage auf einem einzelnen Markt betrachtet. Die Interdependenz der Märkte wird ausgeblendet.

Es gibt ein allgemeines Einteilungsschema, das sich orientiert an

- der Marktstruktur,
- dem Marktverhalten,
- dem Marktergebnis.

Quantitatives Marktformenschema

Je nach quantitativer Besetzung der beiden Marktseiten in einer, wenige und viele erhält man folgendes Schema

	Nachfrager		
Anbieter	viele	wenige	einer
viele	Polypol	Oligopson	Monopson
wenige	Oligopol		
einer	Monopol		bilaterales Monopol

Es gibt folgende lose Zusammenhänge zwischen quantitativer Marktform und Verhaltensweisen

- Polypol $\Longrightarrow$ Mengenanpassung bei Anbietern und Nachfragern
- Monopol $\Longrightarrow$ Mengen- oder Preisfixierung beim Anbieter
- Oligopol $\Longrightarrow$ Reaktionsverbundenheit der Anbieter, strategisches Verhalten (wird in der Spieltheorie näher untersucht).

Vollkommene und unvollkommene Märkte

Wenn in den Augen der Nachfrager die gehandelten Güter völlig gleichartig (homogen) sind, spricht man von einem *vollkommenen* Markt. Er ist gekennzeichnet dadurch dass die Nachfrager weder

- räumliche
- sachliche
- zeitliche
- persönliche

Präferenzen besitzen.

Ist hingegen das Gut inhomogen, d.h. besitzen die Nachfrager eine der oben genannten Vorlieben, dann handelt es sich um einen unvollkommenen Markt.

Monopol

- Es gibt nur *einen* Anbieter.
- Der Marktzutritt für weitere Anbieter ist *nicht* möglich.
- Der Monopolist produziert ein homogenes Gut.
- Der Monopolist ist Gewinnmaximierer und verhält sich auf seinem Absatzmarkt und seinen Beschaffungsmärkten als Mengenanpasser.
- Es gibt sehr viele, kleine Nachfrager.

Unterschiede zwischen Monopol und Konkurrenz

Der wesentliche Unterschied zur vollständigen Konkurrenz ist der einzelne Anbieter, der sich der gesamten Marktnachfrage gegenübersieht. Seine PAF ist gleich der Marktnachfragekurve. Wenn der Monopolist seine Angebotsmenge variiert, dann verändert sich der Marktpreis.

Dieser Unterschied kommt in der Erlösfunktion und damit auch der Gewinnfunktion des Monopolisten zum Ausdruck. Für ihn ist der Marktpreis *kein Datum*.

Gewinnmaximum im Monopol

Die inverse Marktnachfragekurve (die PAF des Monopolisten) wird mit $p(y)$ bezeichnet und die Kostenfunktion mit $c(y)$. Die Erlösfunktion lautet dann

$$r(y) = p(y)y$$

Die Gewinnmaximierungsaufgabe für ein Monopol ist praktisch dieselbe wie die einer Firma unter Konkurrenz:

$$\max_y \pi_m(y) = r(y) - c(y)$$

Der Unterschied ist jedoch, dass nun der Einfluss der angebotenen Menge auf den Marktpreis zu berücksichtigen ist.

Der Gewinn wird maximal für $\pi'_m = 0$ und $\pi''_m < 0$

$$MR = r'(y) = c'(y) = MC$$

also weiterhin bei dem Output, bei welchem der Grenzerlös gleich den Grenzkosten ist.

Im Falle der Konkurrenz ist der Marktpreis ein Datum und der Grenzerlös folglich gleich diesem Preis.

Grenzerlös im Monopol

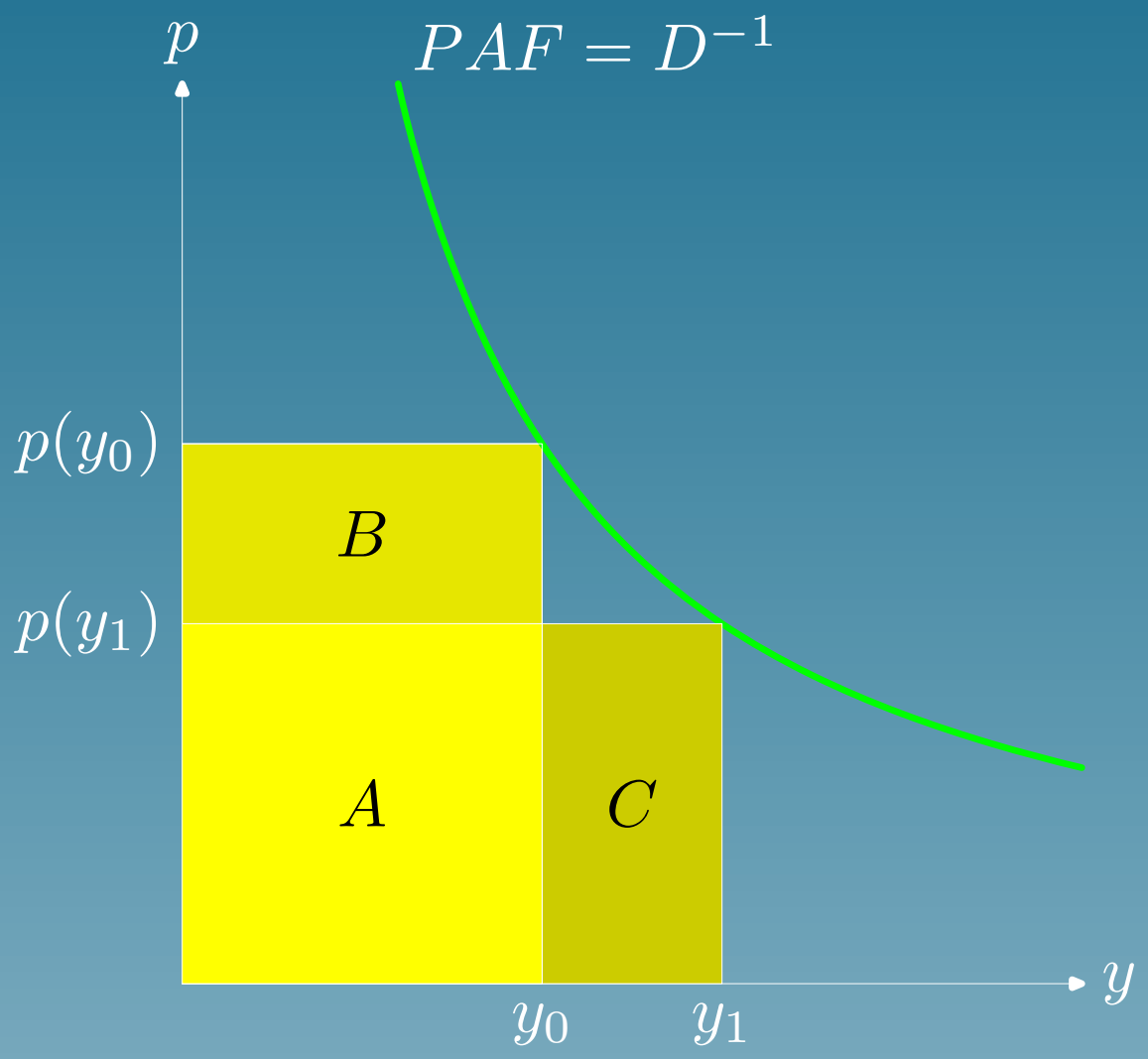
Im Monopol besteht der Grenzerlös jedoch aus zwei Termen:

$$r'(y) = \big(p(y)\,y\big)' = p'(y)\,y + p(y)$$

Der erste Term $p'(y)\,y$ ist bei einer normalen (fallenden) Marktnachfrage negativ. Wenn der Monopolist sein Angebot marginal ändert, so hat dies einen Einfluss auf den Preis, zu welchem er seine gesamte Menge absetzen kann.

Der zweite Term ist der bereits bekannte Effekt, wenn der Absatz — bei konstantem Preis — variiert, dann ändert sich der Erlös direkt proportional mit dem Preis.

Erlösänderung im Monopol



Bei einem Angebot in Höhe von y_0 beträgt der Erlös des Monopolisten $r(y_0) = p(y_0)y_0$. Dies ist gleich den beiden Rechtecksflächen $A + B$.

Erhöht der Monopolist sein Angebot auf y_1, ergibt sich als neuer Erlös $r(y_1) = p(y_1)y_1$, was $A + C$ entspricht.

Seine Erlösänderung setzt sich aus zwei Effekten zusammen, sein neuer Erlös bei y_1 ist um die Fläche B kleiner und zugleich um die Fläche C größer als beim alten Angebot y_0.

Für die Erlösänderung gilt:

$$\begin{aligned} r(y_1) - r(y_0) = \Delta r &= p(y_1)y_1 - p(y_0)y_0 \\ &= p(y_0 + \Delta y)\big(y_0 + \Delta y\big) - p(y_0)y_0 \\ &= \big(p(y_0 + \Delta y) - p(y_0)\big)y_0 + p(y_0 + \Delta y)\Delta y \end{aligned}$$

Also ergibt sich als Änderungsrate des Erlöses der Newton-Quotient

$$\frac{\Delta r}{\Delta y} = \frac{p(y_0 + \Delta y) - p(y_0)}{\Delta y} y_0 + p(y_0 + \Delta y)$$

und im Grenzübergang für $\Delta y \to 0$ erhalten wir das frühere Resultat

$$\lim_{\Delta y \to 0} \frac{\Delta r}{\Delta y} = r'(y_0) = p'(y_0)y_0 + p(y_0)$$

Grenzerlös und Preiselastizität der Nachfrage

Wird im Ausdruck für den Grenzerlös $r'(y) = p + p'(y)\,y$ der Preis ausgeklammert, ergibt sich

$$r'(y) = p(y)\left(1 + \frac{\partial p}{\partial y}\frac{y}{p(y)}\right) = p(y)\left(1 + \frac{1}{\epsilon(y)}\right)$$

wobei ϵ die direkte Preiselastizität der Nachfrage bezeichnet.

Diese Beziehung zwischen Grenzerlös und Preiselastizität der Nachfrage ist auch unter dem Namen Amoroso-Robinson-Relation bekannt.

Im Falle der vollständigen Konkurrenz ist die PAF der Firma eine horizontale Linie in Höhe des Marktpreises und damit $\epsilon(y) = -\infty$, daher gilt bei ihr $r'(y) = p$.

Lineare Marktnachfrage und PAF im Monopol

Angenommen, die Marktnachfragekurve $y = D(p)$ ist linear

$$y = \alpha - \beta p \quad \alpha, \beta > 0$$

Die Preis-Absatz-Funktion des Monopolisten ist die inverse Nachfragekurve

$$p(y) = a - by \quad \text{wobei} \quad a = \frac{\alpha}{\beta}, \; b = \frac{1}{\beta}$$

Die Größe a ist dabei der sogenannte Prohibitivpreis und a/b ist die sogenannte Sättigungsmenge.

Lineare Marktnachfrage und Grenzerlös im Monopol

Bei einer linearen Marktnachfrage und damit einer ebenfalls linearen PAF ist der Erlös eines Monopols eine nach unten geöffnete Parabel

$$r(y) = p(y)\,y = ay - by^2$$

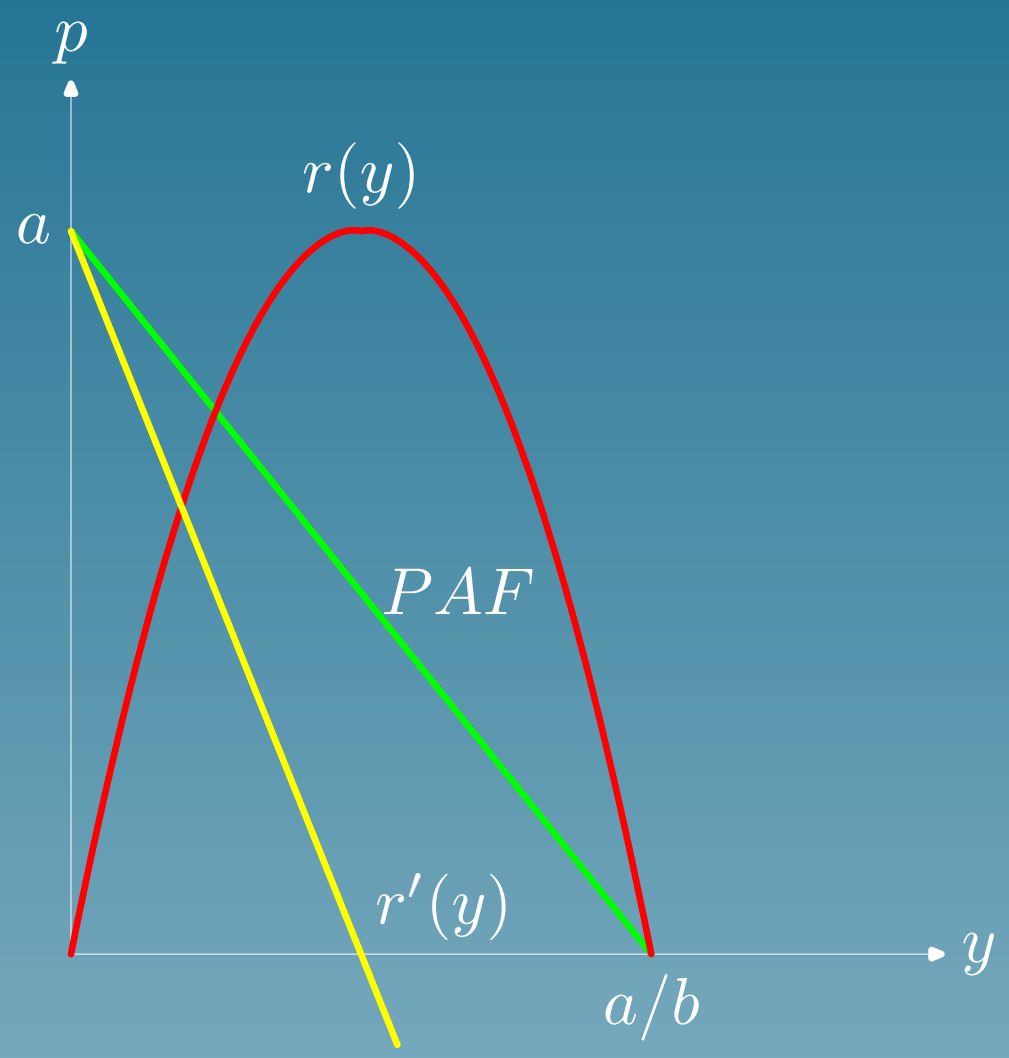
Der Grenzerlös

$$r'(y) = a - 2by$$

ist eine Gerade, die doppelt so steil ist, wie die PAF.

Sie besitzt denselben Ordinatenabschnitt a (Prohibitivpreis) wie die PAF und ihr Abszissenabschnitt $a/2b$ ist die Hälfte der Sättigungsmenge a/b.

Lineare PAF, Erlös- und Grenzerlöskurve im Monopol



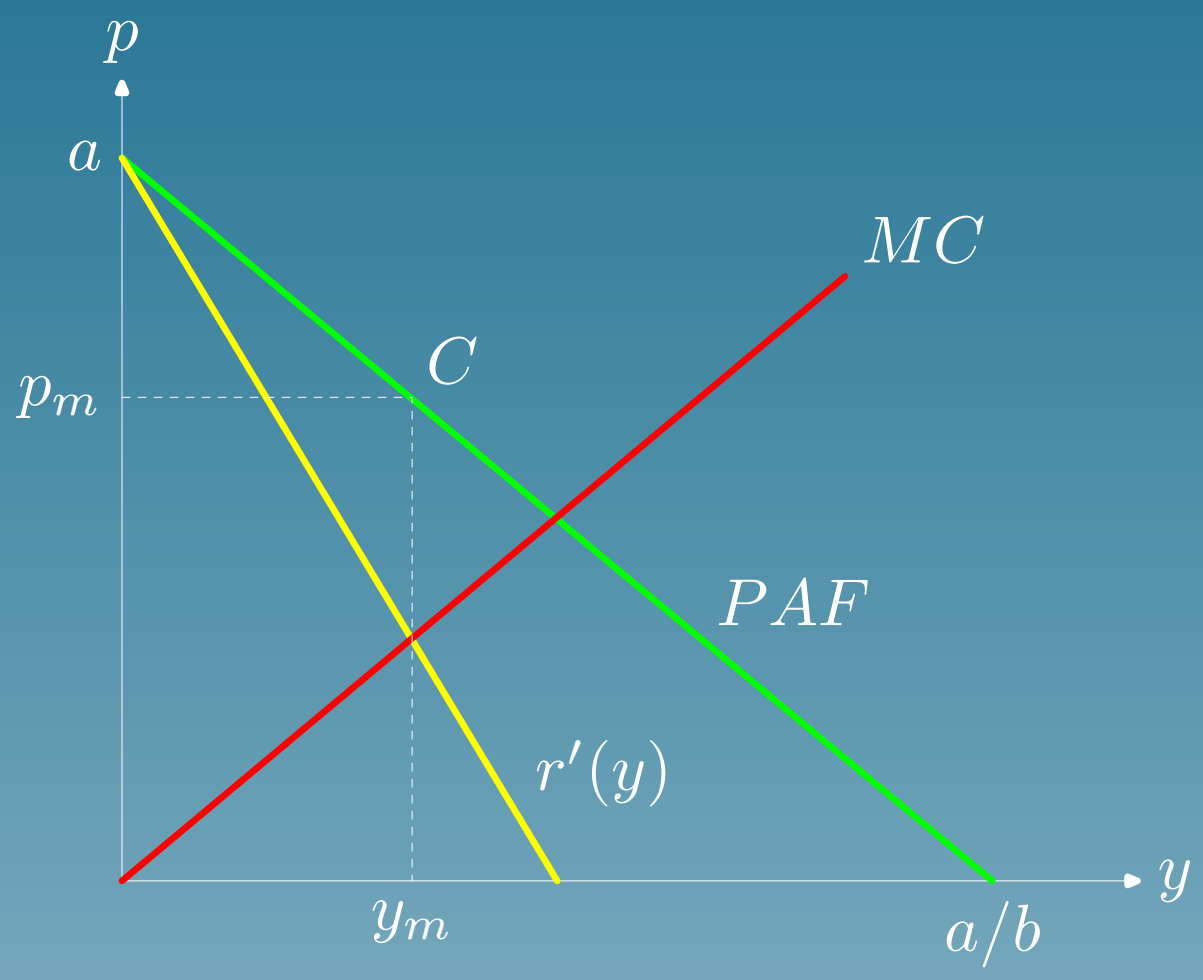
Gewinnmaximale Preis-Mengen-Kombination im Monopol

Das Monopol plant seine gewinnmaximale Produktionsmenge *nicht* zu alternativen potenziellen Gleichgewichtspreisen, sondern mit Berücksichtigung seines Einflusses auf den Marktpreis. Daher gibt es auch keine Angebotskurve, sondern nur eine einzige Preis-Mengen-Kombination, welche zum maximalen Monopolgewinn führt.

Dieser Punkt im Preis-Mengen-Diagramm heißt Cournot'scher Punkt.

Es handelt sich um den Punkt auf der PAF des Monopols, unter dem sich Grenzkosten- und Grenzerlöskurve schneiden.

Cournot'scher Punkt



Monopolgrad

Der Preis p_m im Monopol ist größer als die Grenzkosten.

Die Rate, mit welcher der Preis p_m von den Grenzkosten abweicht, ist wegen der gewinnmaximalen Outputregel $r'(y_m) = MC(y_m)$ und der Amoroso-Robinson-Beziehung einfach der Kehrwert der Preiselastizität der Nachfrage.

$$MC(y_m) = r'(y_m) = p_m \left(1 + \frac{1}{\epsilon(y_m)}\right) \quad \rightsquigarrow \quad \frac{MC - p_m}{p_m} = \frac{1}{\epsilon(y_m)}$$

Da unter Konkurrenz der Preis gleich den Grenzkosten ist, kann der Kehrwert der Preiselastizität als Maß für die Marktmacht eines Monopols angesehen werden.

Diese Überlegungen lassen sich auch auf eine Industrie mit n Firmen übertragen. Ist y_i das gewinnmaximale Angebot von Firma $i = 1, \ldots n$, und y das aggregierte Angebot, dann gilt für die i-te Firma:

$$\Delta r(y_i) = p(y) + \frac{\Delta p}{\Delta y} y_i$$

Ist $s_i = y_i/y$ der Marktanteil von Firma i, dann gilt

$$\begin{aligned}\Delta r'(y_i) &= p(y) + \frac{\Delta p}{\Delta y}\frac{s_i}{y} \\ &= p(y)\left(1 + \frac{s_i}{\epsilon(y)}\right)\end{aligned}$$

Der Monopolgrad der i-ten Firma beträgt damit

$$\frac{MC_i - p(y)}{p(y)} = \frac{s_i}{\epsilon(y)}$$

Er ist um so geringer,

- je geringer der Marktanteil s_i der Firma und
- je kleiner die Preiselastizität der Nachfrage $\epsilon(y)$ ist.

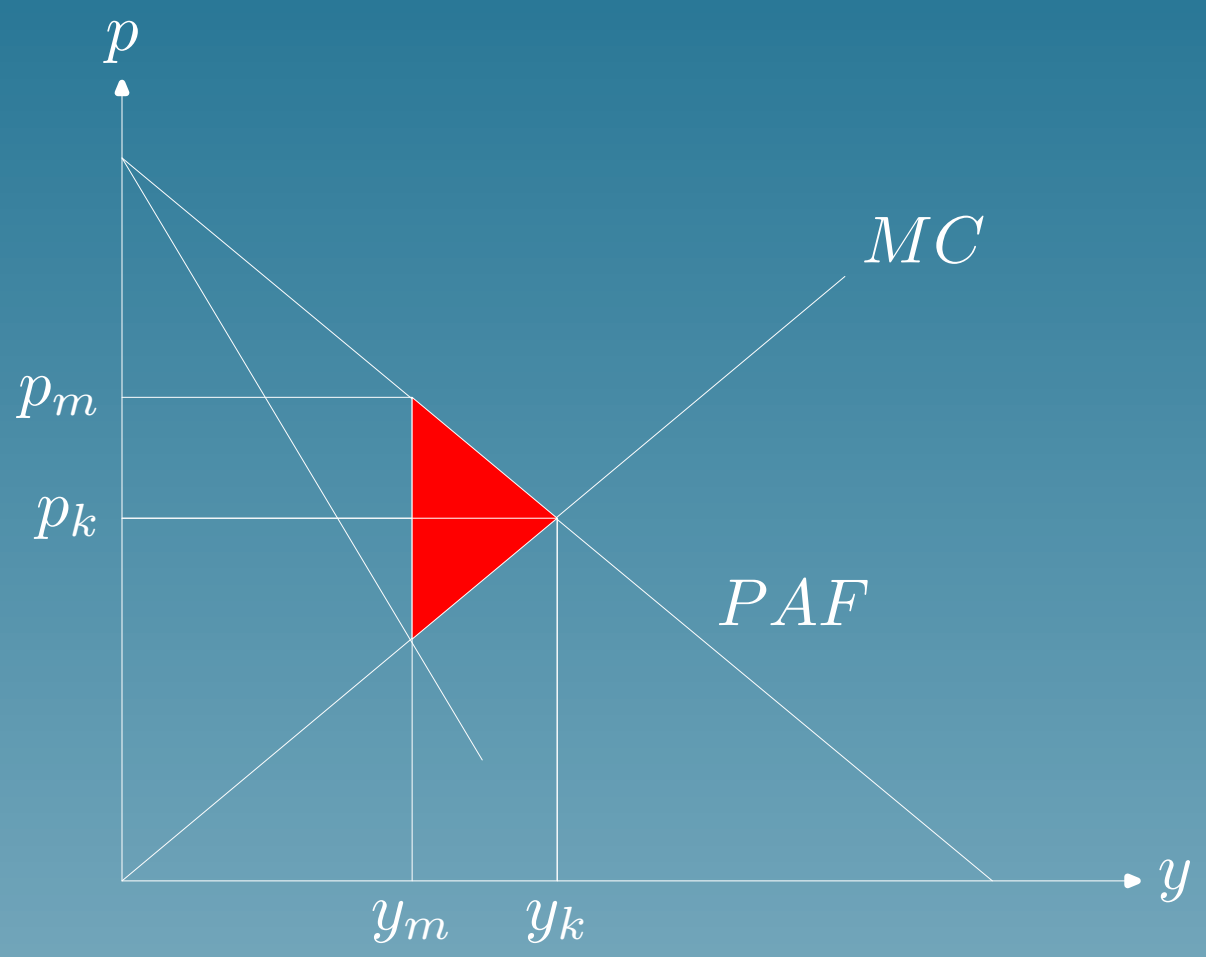
Beachte: die Preiselastizität der Nachfrage ist eine negative (und dimensionslose) Zahl.

Ineffizienz des Monopols

Der Cournot'sche Punkt liegt oberhalb der Grenzkostenkurve auf der PAF des Monopols. Wenn es gelänge, den Monopolisten zu veranlassen, sich wie eine Firma unter vollständiger Konkurrenz zu verhalten, würde der Output am Schnittpunkt der Grenzkostenkurve mit der PAF gewählt, einem Punkt, der rechts unterhalb des Cournot'schen Punktes liegt.

Die Verbraucher (Gesellschaft) hätten dann einen geringeren Preis zu entrichten und würden auch mehr des Gutes erhalten. Die Konsumentenrente wäre größer und die Produzentenrente kleiner als im Cournot'schen Punkt. Neben dieser (eventuell unerwünschten) Verteilung der Rente auf Konsumenten und Produzenten verursacht ein Monopol auch einen gesellschaftlichen Wohlfahrtsverlust.

Harberger-Dreieck eines Monopols



Monopolkontrolle

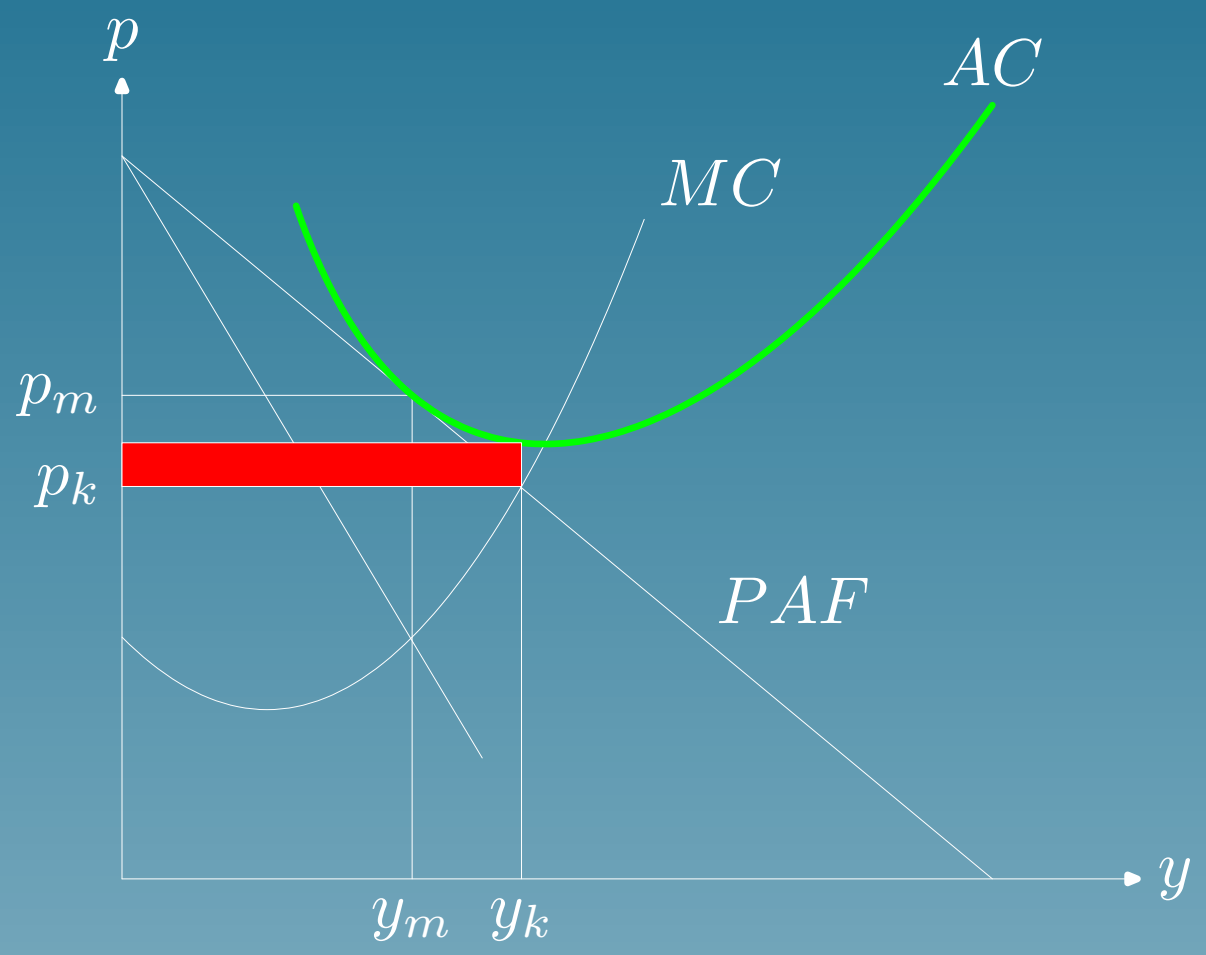
Es liegt nahe den gesellschaftlichen Wohlfahrtsverlust zu beseitigen, indem der Staat dem Monopol vorschreibt, seinen Preis auf p_k zu senken oder seinen Output auf y_k zu erhöhen.

Dabei bleibt jedoch der Gewinn des Monopols unberücksichtigt. Erst wenn die Durchschnittskosten bekannt sind, kann dazu eine Aussage gemacht werden.

Falls die Durchschnittskostenkurve des Monopols durch seinen Cournot'schen Punkt läuft, dann produziert es dort gerade kostendeckend. Eine staatlich verordnete Preissenkung würde zu Verlusten führen.

Ein Monopol, dessen Durchschnittskostenkurve durch seinen Cournot'schen Punkt läuft, wird natürliches Monopol genannt.

Natürliches Monopol



Monopolistisches Verhalten

Wenn ein Alleinanbieter (Monopolist) seinen Preis anhebt, wird er einen Teil seines Absatzes verlieren. Wenn eine Firma unter vollständiger Konkurrenz ihren Preis anhebt, verliert sie ihre gesamte Kundschaft.

In der Realität wird sich eine Firma innerhalb einer Industrie irgendwo zwischen diesen beiden Extremen finden. Sie hat einen gewissen Spielraum für die Preisgestaltung ihres Produktes, kann sich aber nicht ganz so wie ein Monopol verhalten.

Drei Formen der Preisdiskriminierung

Wenn es dem Monopolisten gelingt, sein Produkt zu unterschiedlichen Preisen zu verkaufen, spricht man von Preisdiskriminierung.

Voraussetzung für Preisdiskriminierung ist, dass die Nachfrager das Produkt nicht untereinander weiterverkaufen können. Andernfalls könnten Zwischenhändler auftreten, welche durch Arbitragegeschäfte Gewinne machen könnten.

Preisdiskriminierung ist vor allem bei Dienstleistungen anzutreffen, da hier der Verbraucher beim Produzieren mitwirken muss (z.B. Personenbeförderung).

Es werden davon drei Arten der Preisdiskriminierung unterschieden.

Drei Formen der Preisdiskriminierung

1. Preisdiskriminierung ersten Grades liegt vor, wenn der Monopolist für unterschiedliche Mengeneinheiten verschiedene Preise setzt und sich diese Preise zusätzlich von Käufer zu Käufer unterscheiden können.
2. Preisdiskriminierung zweiten Grades liegt vor, wenn das Monopol unterschiedliche Gütereinheiten zu unterschiedlichen Preisen verkauft, aber unterschiedliche Käufer, welche dasselbe Quantum kaufen, denselben Preis zahlen (z.B. beim Mengenrabatt).
3. Preisdiskriminierung dritten Grades liegt vor, wenn der Monopolist sein Gut zu unterschiedlichen Preisen an bestimmte Käufergruppen verkauft. Ein einzelner Kunde hat dabei unabhängig von seiner Kaufmenge einen bestimmten Preis zu entrichten.

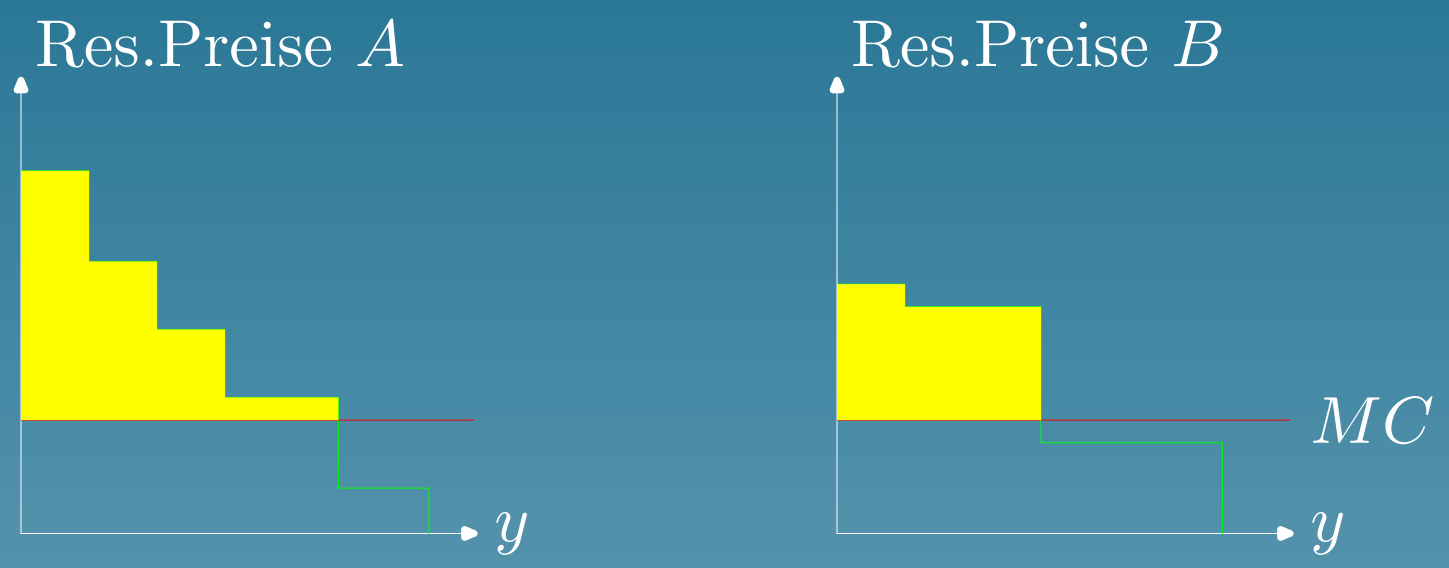
Preisdiskriminierung ersten Grades

Diese Form wird auch als vollständige Preisdiskriminierung bezeichnet.

Jede einzelne Gütereinheit wird dabei an die Person mit der jeweils höchsten Zahlungsbereitschaft verkauft. Das Monopol vereinnahmt dadurch die gesamte Konsumentenrente.

Da der Monopolist jede Gütereinheit zu unterschiedlichen Preisen verkaufen kann, wird seine letzte Gütereinheit zu einem Preis verkauft, bei deren Produktion Grenzkosten in Höhe der Grenzkosten dieses Preises anfallen.

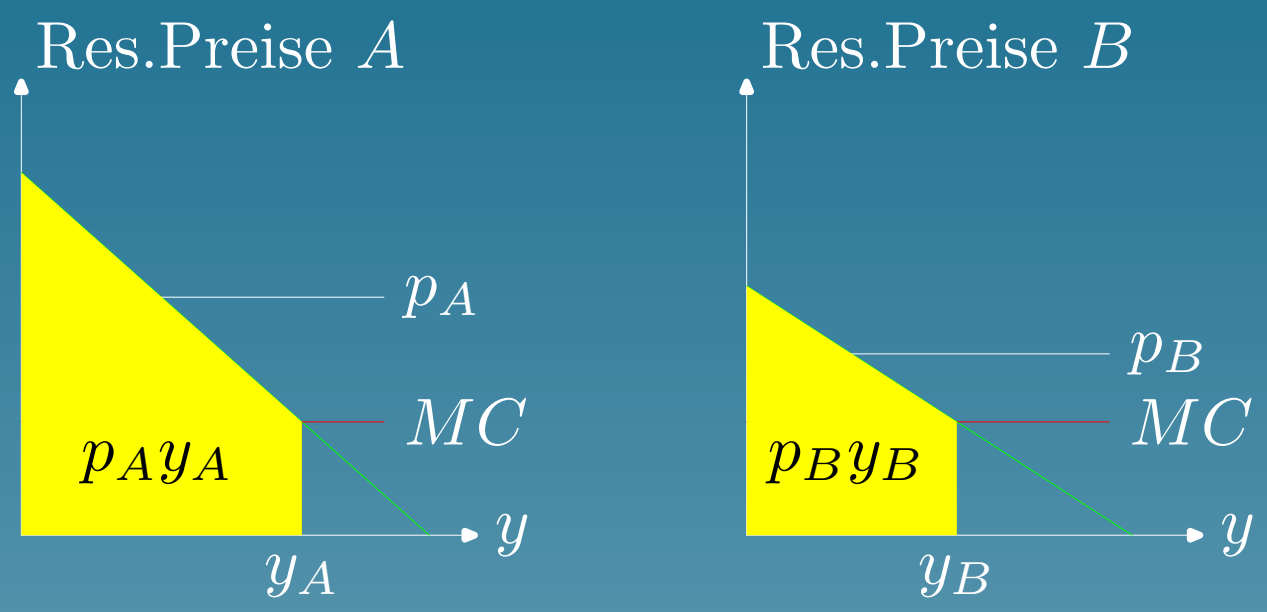
Monopolist vereinnahmt die Konsumentenrente



Der Monopolist verkauft an den Kunden A eine erste Einheit zum Preis von dessen höchster Zahlungsbereitschaft (Reservationspreis), danach eine zweite Einheit usw.

Das gleiche Verfahren wird beim Kunden B angewandt usw.

Take it or leave it



Der Monopolist unterbreitet dem Kunden A die Option, die Menge y_A zum Preis von p_A zu kaufen. Der Preis p_A ist der Mittelwert aus dem Reservationspreis von A für seine erste Gütereinheit und den Grenzkosten. Beim Kunden B wird genauso verfahren.

Effizienz der Preisdiskriminierung ersten Grades

Bei Preisdiskriminierung ersten Grades entsteht kein gesellschaftlicher Wohlfahrtsverlust.

Lediglich die Verteilung der Renten ist anders als bei vollständiger Konkurrenz.

Sowohl die vollständige Konkurrenz also auch das Monopol mit Preisdiskriminierung ersten Grades sind daher Pareto-Effizient.

Preisdiskriminierung zweiten Grades

Bei dieser Form werden unterschiedliche Einheiten desselben Gutes zu unterschiedlichen Preisen verkauft. Zugleich wird nicht zwischen Kunden differenziert.

Typische Beispiele sind

- der „Sechserpack“ ist billiger als sechs Einzelteile
- Elektrizitätstarife für Großverbraucher sind niedriger als für Kleine.

Preisdiskriminierung dritten Grades

Unterschiedliche Käufer zahlen dabei unterschiedliche Preise. Es gibt jedoch keinen Mengenrabatt.

Typische Beispiele dafür sind Eintrittsermäßigungen für Schüler und Studenten oder Senioren.

Aufspaltung des Marktes

Preisdiskriminierung ist nur möglich, wenn sich die Käufer in mindestens zwei Gruppen aufspalten lassen und ein Handel der Käufer untereinander nicht möglich ist.

Angenommen, es gibt zwei Teilmärkte mit den inversen Nachfragekurven (PAF's des Monopolisten) $p_1(y_1)$ und $p_2(y_2)$. Der preisdiskriminierende Monopolist wird seine Produktion $y = y_1 + y_2$ so wählen und auf die beiden Teilmärkte aufteilen, dass sein Gewinn maximal ist.

$$\max_{y_1,\, y_2} p_1(y_1)\, y_1 + p_2(y_2)\, y_2 - c(y_1 + y_2)$$

Gewinnmaximum des preisdiskriminierenden Monopols

Die beiden Bedingungen für ein Gewinnmaximum lauten

$$p_1'(y_1)\,y_1 + p_1(y_1) = c'(y_1 + y_2)$$
$$p_2'(y_2)\,y_2 + p_2(y_2) = c'(y_1 + y_2)$$

oder mittels der Amoroso-Robinson-Beziehung

$$p_1\left[1 - \frac{1}{|\epsilon(y_1)|}\right] = p_2\left[1 - \frac{1}{|\epsilon(y_2)|}\right] = c'(y)$$

Man erkennt, dass die Preise nur dann verschieden sind, wenn sich die Preiselastizitäten der Teilmärkte unterscheiden.

Beispiel 6

Wenn es zwei Teilmärkte mit linearen Nachfragekurven $y_1 = a - bp_1$ und $y_2 = c - dp_2$ gibt und die Grenzkosten der Produktion Null sind, dann ergibt sich folgende Kalkulation.

Die beiden PAF's (inversen Nachfragekurven beim Monopol) lauten

$$p_1 = \frac{a}{b} - \frac{y_1}{b} \quad \text{und} \quad p_1 = \frac{c}{d} - \frac{y_2}{d}$$

Die Mengen y_1 und y_2 sind so zu wählen, dass die Grenzerlöse auf den beiden Teilmärkten gleich Null (den Grenzkosten) sind.

$$(p_1 y_1)' = \frac{a}{b} - \frac{2y_1}{b} = 0 \quad \rightsquigarrow \quad y_1 = a/2$$

Auf jedem Teilmarkt wird die Hälfte der Sättigungsmenge angeboten

$$y_1^* = a/2 \quad \text{und} \quad y_2^* = c/2$$

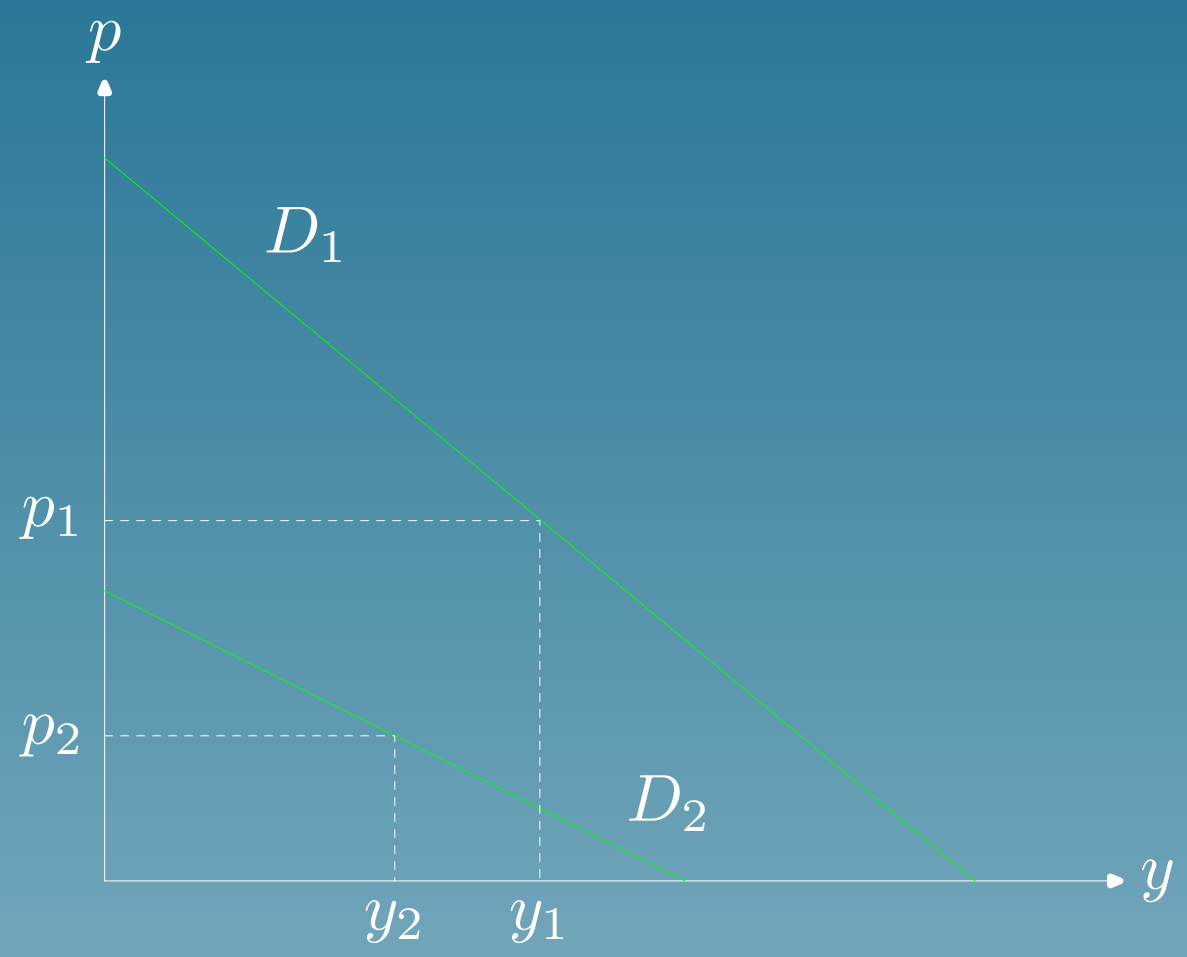
Die Preise auf den beiden Teilmärkten sind dann

$$p_1^* = \frac{a}{b} - \frac{y_1^*}{b} = \frac{a}{b} - \frac{a}{2b} = \frac{a}{2b} \quad \text{und} \quad p_2^* = \frac{c}{2d}$$

Der Gewinn des Monopols beläuft sich

$$\pi(y_1^*, y_2^*) = p_1^* y_1^* + p_2^* y_2^* = \frac{a^2}{4b} + \frac{c^2}{4d}$$

Wohlfahrtszuwachs durch Preisdiskriminierung



Unterstellt sind Grenzkosten in Höhe von Null. Wenn das Monopol nur einen Preis setzen darf, wird es nur den Teilmarkt 1 versorgen. Der Preis p_1 ist höher als der Prohibitivpreis auf dem Teilmarkt 2.

Falls aber Preisdiskriminierung erlaubt wird, dann versorgt das Monopol auch den zweiten Teilmarkt mit der Menge y_2 zum Preis von p_2. Dies ist eindeutig ein Wohlfahrtszuwachs.

Andere Arten der Preisgestaltung

Es gibt noch eine Vielzahl weiterer Arten der diskriminierenden Preisgestaltung.

Darunter sind vor allem

- die Bündelung der Produkte und
- die Aufspaltung in Grundgebühr und verbrauchsabhängigen Tarif

von praktischer Relevanz.

Monopolistische Konkurrenz

Märkte mit monopolistischer Konkurrenz sind durch die folgenden Merkmale gekennzeichnet:

1. Wie bei vollständiger Konkurrenz gibt es sehr viele Anbieter und Nachfrager mit jeweils verschwindend geringem Marktanteil (atomistische Angebots- und Nachfragestruktur).
2. Es gibt Präferenzen (räumlicher, zeitlicher, sachlicher oder persönlicher Art) seitens der Nachfrager; der betrachtete Markt ist unvollkommen.
3. Ein Anbieter kennt zumindest die Nachfrage des von ihm angebotenen Gutes, es kann auch vollständige Preisinformation bei allen Marktteilnehmern vorliegen.

Unterschiedliche Preise bei monopolistischer Konkurrenz

Aufgrund der Unvollkommenheit des Marktes (Eigenschaft 2) bzw. wegen der möglichen unvollkommenen Information (Eigenschaft 3) kann es zu unterschiedlichen Preisen auf dem Markt kommen.

Diese Preise werden von den Nachfragern als Daten hingenommen, an die sie ihre Mengenplanungen anpassen.

Da die Marktanteile der Marktteilnehmer gemäß Eigenschaft 1 verschwindend klein sind, wirkt sich die Preissetzung oder Preisänderung eines einzelnen Anbieters nicht merklich auf die Angebotssituation seiner Mitkonkurrenten aus; sie haben keine Veranlassung auf preispolitische Maßnahmen zu reagieren.

Chamberlin's Ansatz

Anbieter i ist repräsentativ für die große Zahl n aller Firmen der Industrie. Sein Absatz y_i wird negativ von seinem eigenen Preis p_i und positiv von den Preisen der übrigen Anbieter beeinflusst. Seine Preisabsatzfunktion ist linear:

$$y_i = -a_i\, p_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} b_{ij} p_j + c \qquad a,\, b,\, c > 0$$

Die einzelnen b_{ij} sind zwar vernachlässigbar klein, die Summe selbst jedoch nicht. Der Einfachheit halber wird $b_{ij} = b$ gesetzt.

Chamberlin unterscheidet zwei Typen von PAF's:

1. Die PAF von Anbieter i für den Fall, dass alle anderen Anbieter stets die gleichen Preise setzen, $p_i = p_j$, $j = 1, 2 \ldots n$.
2. Die PAF bei isolierter Preisvariation bei Anbieter i: $p_j = \bar{p}_j$.

Im Fall 1 wird die PAF zu

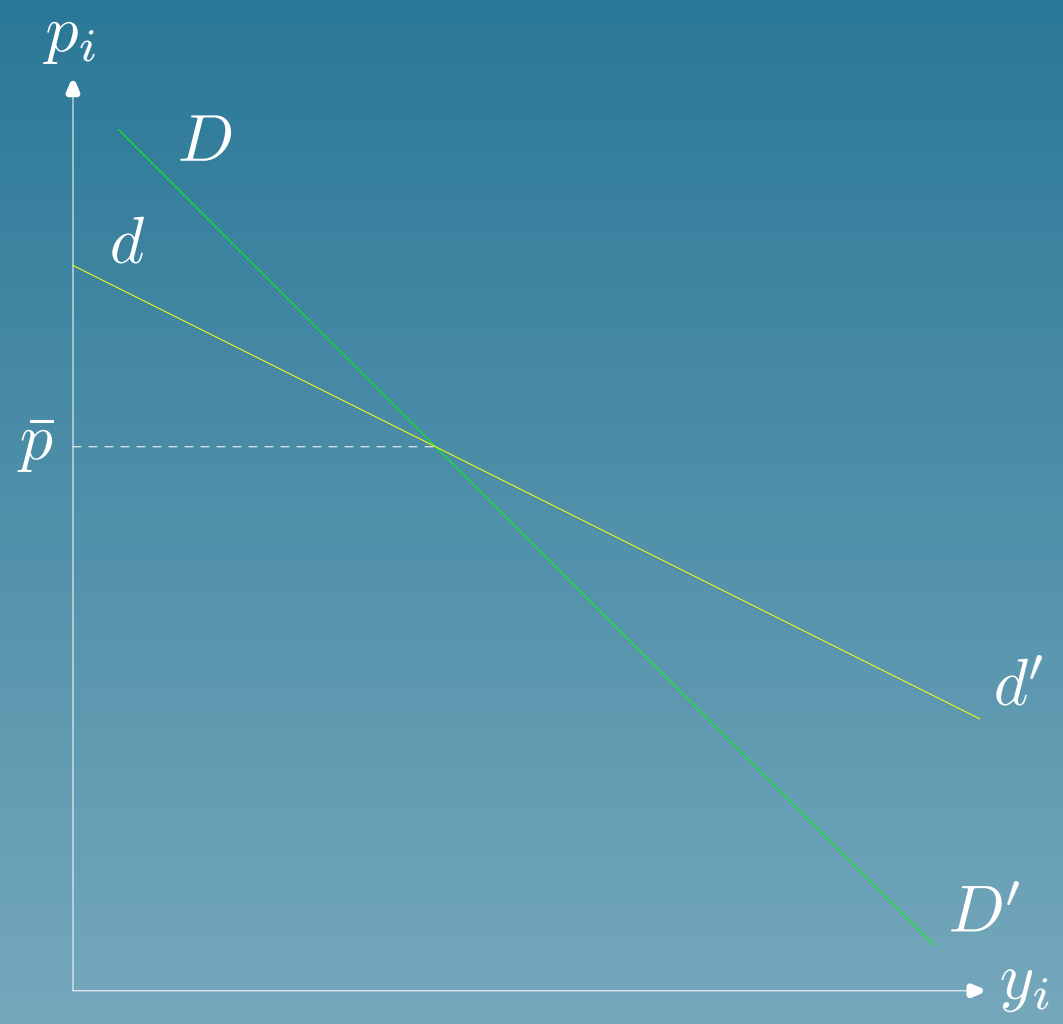
$$y_i = (-a + b(n-1))\, p_i + c \;\rightarrow\; p_i = -\frac{y_i}{a - b(n-1)} + \frac{c}{a - b(n-1)}$$

und im zweiten Fall:

$$y_i = -ap_i + k \quad \rightarrow \quad p_i = -\frac{y_i}{a} + \frac{k}{a} \quad \text{wobei} \quad k = b\sum \bar{p}_j + c$$

Die PAF vom Typ 2 verläuft flacher als die von Typ 1, da $-1/a > -1/(a - b(n-1))$.

DD' und dd' Kurven



Die DD'-Kurve stellt die PAF vom Typ 1 dar, die dd'-Kurve die vom Typ 2.

Konstruktionsbedingt schneidet eine dd'-Kurve die DD'-Kurve an der Stelle, wo alle Preise gleich sind (hier als $\bar{p}$ angenommen).

Gewinnmaximum bei Anbieter i

Der Anbieter legt seiner gewinnmaximalen Produktionsplanung die dd'-Kurve zugrunde, denn er kann bei eigenen preispolitischen Aktionen mit konstanten Preisen der Mitanbieter rechnen, da sich seine Aktionen nur unmerklich auf jeden einzelnen Konkurrenten auswirken.

Er verhält sich zu „seinem“ Markt wie ein Monopolist, also gemäß der Regel $MR = MC$.

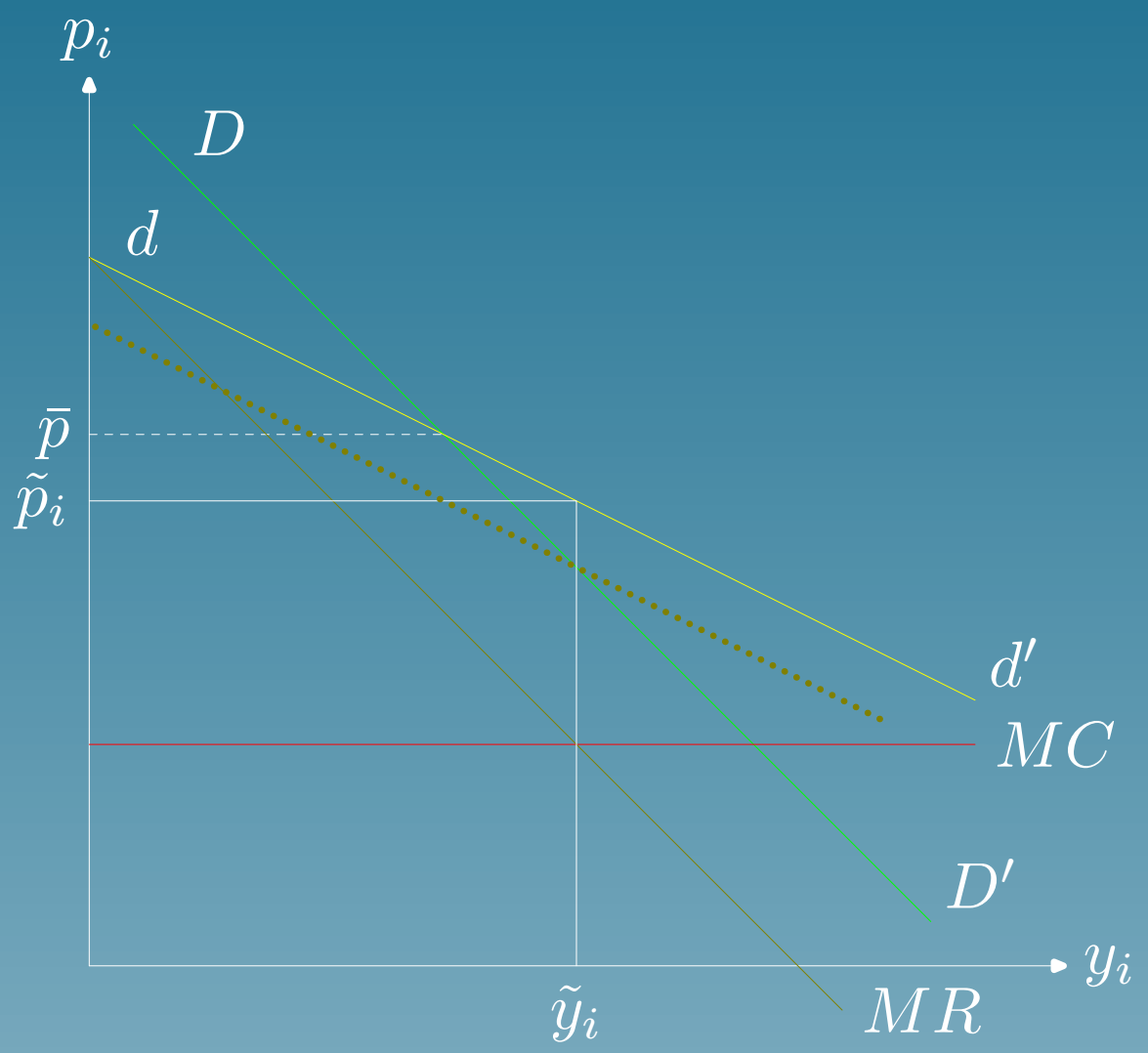
Sein Grenzerlös (mittels der dd'-Kurve) beläuft sich auf

$$MR = -\frac{2\,y_i}{a} + \frac{k}{a}$$

Da nicht nur die Absatz-, sondern auch die Kostensituation bei allen Anbietern gleich ist (Symmetrie-Annahme), werden sich alle gleich verhalten, sodass ein beliebiger Anbieter repräsentativ für alle ist. Die Lösung kann daher nur auf der DD'-Kurve liegen. Wie sie zustande kommt, lässt sich geometrisch veranschaulichen.

Der Einfachheit halber werden konstante Grenzkosten unterstellt.

Geometrische Darstellung des Gewinnmaximums

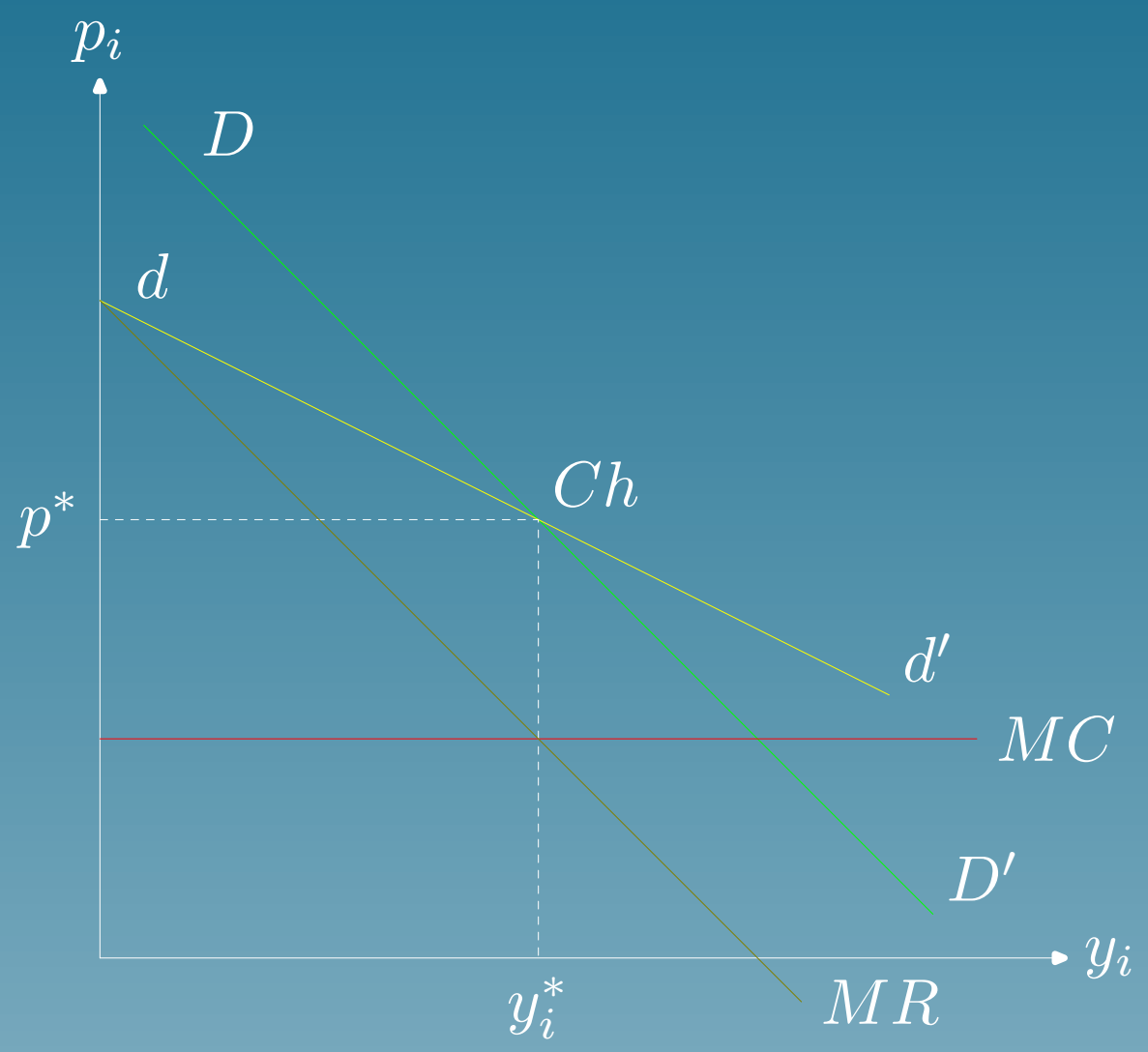


Anbieter i verhält sich wie ein Monopolist mit einer PAF wie dd'. Sein gewinnmaximierender Preis beträgt $\tilde{p} < \bar{p}$.

Da aber alle Anbieter sich wie er verhalten, ist seine Annahme, die übrigen würden den Preis $\bar{p}$ fordern, falsch, alle fordern den Preis $\tilde{p}$, was dann zu einer ganz neuen dd'-Kurve führt, die oben strichliert ist.

Der Anpassungsprozess endet, wenn das individuelle Gewinnmaximum auf der DD'-Kurve zu liegen kommt.

Chamberlin-Punkt



Im Chamberlinpunkt Ch sind die dezentralen Planungen der Anbieter konsistent.

Er ist ein sogenanntes *Nash-Gleichgewicht*: Anbieter i erwartet, dass die übrigen Anbieter den Preis p^* setzen, dieser Preis führt zugleich bei ihm zum Gewinnmaximum.

Langfristiges Gleichgewicht bei monopolistischer Konkurrenz

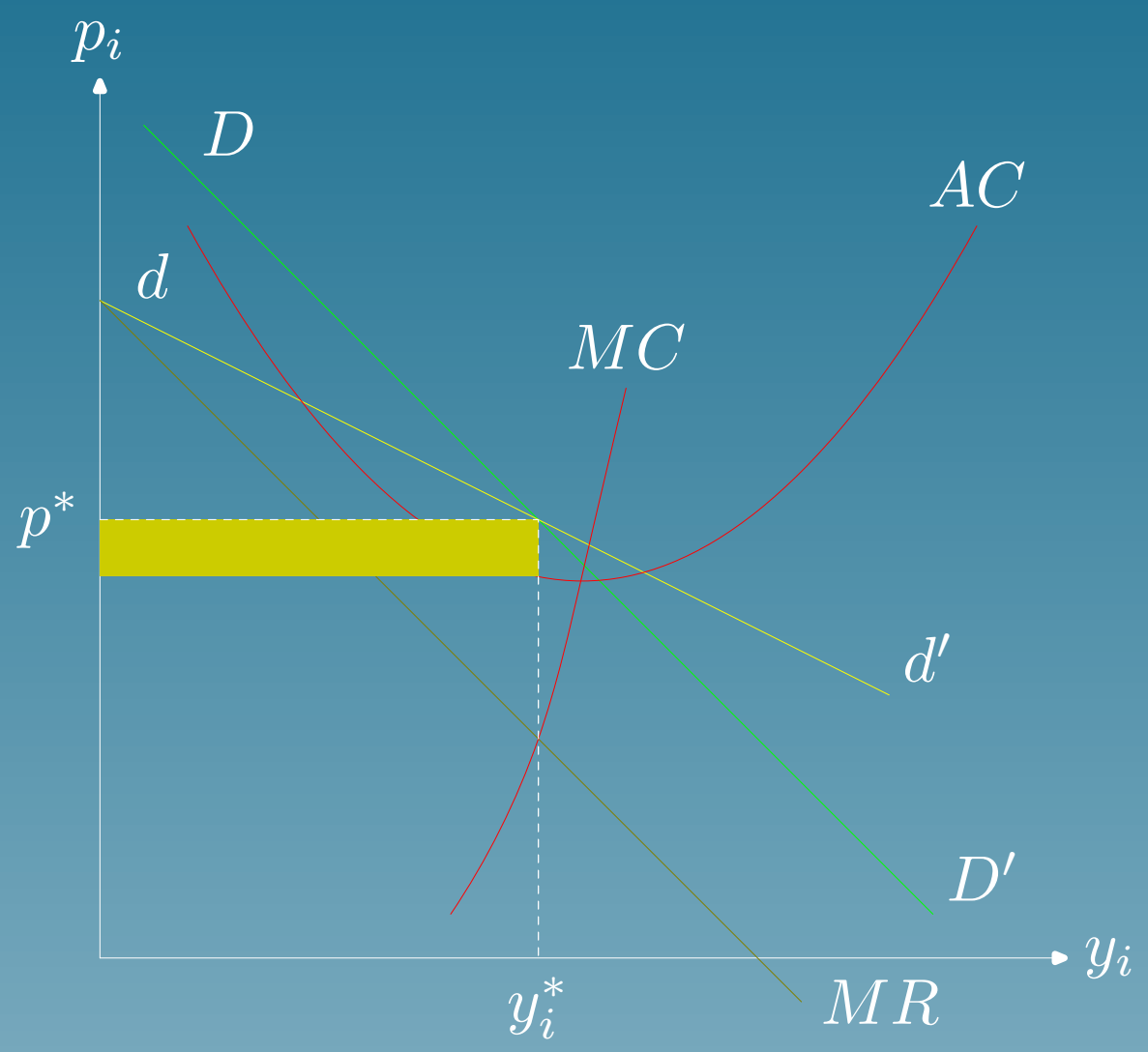
Langfristige Betrachtungen gehen von zwei zusätzlichen Voraussetzungen aus:

- die Betriebsgröße der Firmen ist veränderlich
- der Marktzutritt ist frei, die Anzahl der Anbieter ist veränderlich.

Neue Anbieter werden auf dem Markt tätig, falls sie dort einen positiven Gewinn erzielen können. Während der einzelne neue Anbieter die PAF der bereits am Markt befindlichen Anbieter nur unmerklich beeinflusst, hat der Zustrom einer größeren Anzahl eine Linksverschiebung und Abflachung der PAF zur Folge.

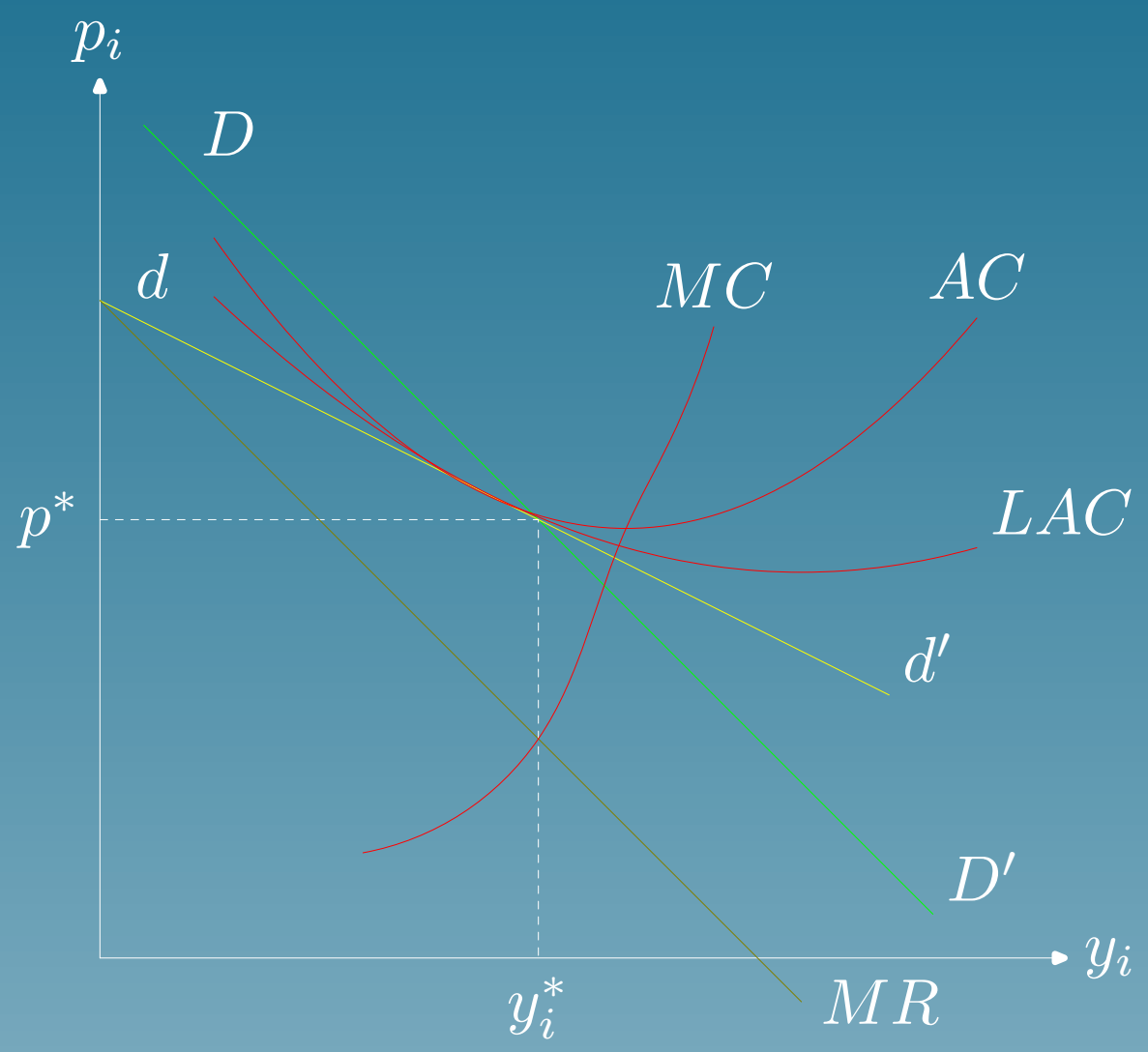
Gesucht ist die Preis-Mengen-Kombination des repräsentativen Anbieters, bei der der Zustrom neuer Anbieter aufhört.

Intramarginal-Anbieter unter monopolistischer Konkurrenz



Die repräsentative Firma erzielt einen positiven Gewinn in Höhe der schraffierten Fläche. Dies kann also kein langfristiges Gleichgewicht sein.

Chamberlin'sche Tangentenlösung



Durch den Zustrom neuer Anbieter verändern sich die DD'- und die dd'-Kurve. Das Gruppengleichgewicht ist bei einem Nullgewinn erreicht, wo die dd'-Kurve die LDK berührt (daher der Name Tangentenlösung).

Beurteilung des langfristigen Gleichgewichts der monopolistischen Konkurrenz

Im Unterschied zum langfristigen Gleichgewicht bei vollständiger Konkurrenz (waagerechte PAF) wird hier das Gruppengleichgewicht immer links vom Betriebsoptimum und links vom langfristigen Durchschnittskostenminimum realisiert, woraus auch langfristige Überkapazitäten entstehen.

Die Tangentenlösung ist daher zwar suboptimal, zumal auch die Verbraucher mit geringerer Menge und höherem Preis als bei vollständiger Konkurrenz versorgt werden, dem steht aber die Produktdiversifikation als Vorteil gegenüber.

Faktormärkte

Die allgemeine outputorientierte Gewinnmaximierungsregel für eine Firma lautet: produziere die Menge, bei der die Grenzkosten so groß wie der Grenzerlös ist.

In Abhängigkeit der Marktform und der kurz- oder langfristigen Betrachtung nimmt diese allgemeine Regel besondere Gestalt an.

Nun soll auch die inputorientierte Gewinnmaximierungsregel bei unterschiedlichen Marktformen betrachtet werden. Ursprünglich wurde dabei unterstellt, dass die Firma auf den Beschaffungsmärkten und dem Absatzmarkt eine von vielen kleinen Firmen ist. Nun wird die Faktornachfrage eines Monopols betrachtet.

Inputorientierte Amoroso-Robinson-Beziehung

Angenommen, die Firma besitzt ein Monopol für den Absatzmarkt und erzeugt ihren Output mittels eines Inputs gemäß der Produktionsfunktion $y = f(x)$. Ihre PAF ist die inverse Marktnachfragekurve $p(y)$.

Der inputorientierte Erlös ist dann

$$py = p(y)y = p\big(f(x)\big)f(x) = R(x)$$

Für den Grenzerlös erhalten wir mittels der Kettenregel

$$R'(x) = p'(y)f'(x)f(x) + p(y)f'(x) = p\big(f(x)\big)\left[1 - \frac{1}{|\epsilon|}\right]f'(x)$$

Inputorientierter Gewinn eines Monopols

Die inputorientierten Kosten sind bei einem Input gegeben durch wx, wobei w den Preis des Inputs bezeichnet. Unterstellt wird, dass die Firma den Inputpreis w als Konstante betrachtet.

Der inputorientierte Gewinn des Monopols

$$\pi(x) = R(x) - wx$$

wird maximal, wenn der Grenzgewinn Null ist, also für

$$R'(x) = w \quad \rightsquigarrow \quad p\big(f(x)\big)\left[1 - \frac{1}{|\epsilon|}\right] f'(x) = w$$

Vergleich mit der Inputregel bei vollständiger Konkurrenz

Beim erstmaligen Betrachten des Gewinnmaximums — unter vollständiger Konkurrenz — stand auf der linken Seite nur der Ausdruck pf', die sogenannte Wertgrenzproduktivität des Inputs.

Beim Monopol steht dort ein Ausdruck, der geringer als pf' ist.

Die gewinnmaximierende Inputregel $pf'(x) = w$ ist bekanntlich die inverse Nachfragekurve nach dem Input der Firma bei vollständiger Konkurrenz. Analog ist $R'(x) = w$ die inverse Nachfragekurve nach dem Input bei einem Monopol.

Beispiel 7

Angenommen, die inverse Marktnachfragekurve (PAF des Monopols) ist linear $p(y) = a - by$ und die Produktionsfunktion ist ebenfalls linear $y = f(x) = \alpha x$.

Um eine Einheit des Gutes zu produzieren werden $1 = \alpha x$ also $1/\alpha$ Einheiten des Inputs benötigt. Diese kosten die Firma w/α. Die outputorientierten Kosten belaufen sich somit auf

$$c(y) = y\frac{w}{\alpha}$$

Die Grenzkosten sind gleich den Durchschnittskosten und gleich den Einheitskosten

$$c'(y) = \frac{c(y)}{y} = c(1) = \frac{w}{\alpha}$$

Der outputorientierte Gewinn für das Monopol ist

$$\pi_m(y) = p(y)y - c(y) = ay - by^2 - y\frac{w}{\alpha}$$

Der inputorientierte Gewinn ergibt sich, indem y durch αx ersetzt wird.

$$\pi_m(x) = p\big(f(x)\big)f(x) - c\big(f(x)\big) = a\alpha x - b\alpha^2 x^2 - wx$$

Die Inputnachfrage ergibt sich aus der inputorientierten Gewinnmaximierungsregel $\pi'_m(x) = 0$

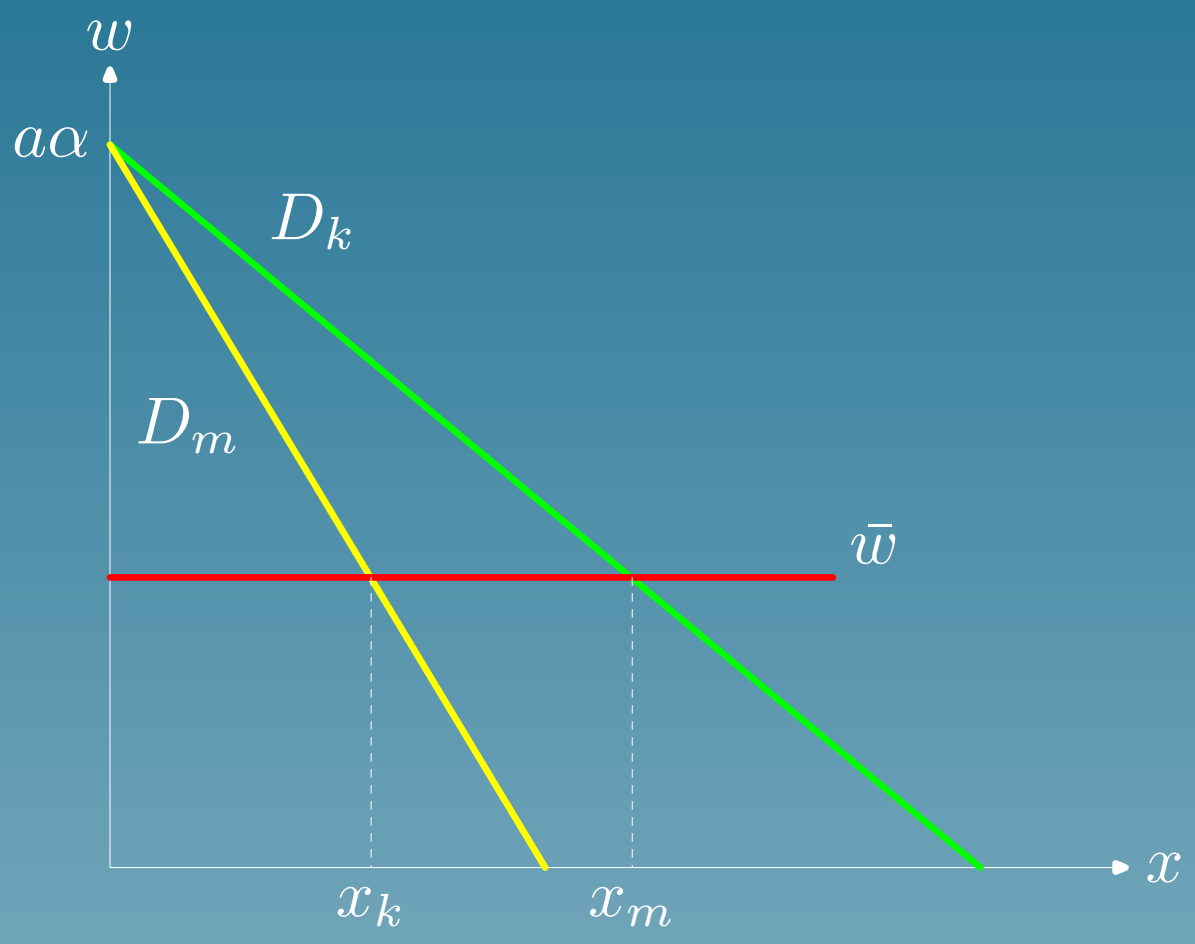
$$a\alpha - 2b\alpha^2 x = w \quad \rightsquigarrow \quad x = \frac{a\alpha - w}{2b\alpha^2} = D_m(w)$$

Die Inputnachfrage einer Firma, die auf ihrem Absatzmarkt der Konkurrenz ausgesetzt ist, lautet hingegen

$$p = c'\big(f(x)\big) \quad \rightsquigarrow \quad a - b\alpha x = \frac{w}{\alpha} \quad \rightsquigarrow \quad x = \frac{a\alpha - w}{b\alpha^2} = D_k(w)$$

Beide Inputnachfragekurven sind in diesem Fall linear, die des Monopols verläuft jedoch links unterhalb von der der Firma unter Konkurrenz.

Faktornachfrage bei Konkurrenz und Monopol



Monopson

Gibt es auf einem Markt nur einen Nachfrager und viele kleine Anbieter, dann handelt es sich um die Marktform des Monopsons. In der Regel ist dieser Nachfrager eine Firma oder eine öffentliche Hand, aber kein privater Haushalt.

Der alleinige Nachfrager ist sich bewusst, dass er durch seine Nachfrage den Marktpreis beeinflusst, analog zum Monopol. Er plant seine Nachfrage anhand der Preisbeschaffungsfunktion PBF, welche gleich der inversen Angebotskurve ist — analog zur PAF des Monopols.

Gewinn im Monopson

Da der Preis w des Inputs x durch die Nachfrage des Monopsons beeinflusst wird, betrachtet man den inputorientierten Gewinn.

$$\pi(x) = pf(x) - w(x)x$$

dabei ist $f(x)$ die Produktionsfunktion und $w(x)$ die PBF des Monopsons.

Die Grenzkosten belaufen sich nun auf

$$c' = w'(x)x + w(x) = w\left[1 + \frac{1}{\eta}\right]$$

wobei $\eta > 0$ die Angebotselastizität für den Input bezeichnet.

Gewinnmaximum im Monopson

Der (inputorientierte) Gewinn ist zu maximieren

$$\max_x \pi(x) = pf(x) - w(x)x \quad \rightsquigarrow \quad pf'(x) = w\left[1 + \frac{1}{\eta}\right]$$

Links steht der (inputorientierte) Erlös und rechts die (inputorientierten) Grenzkosten.

Auch im Monopson wird der Gewinn nach der allgemeinen Regel für die Inputmenge maximal, wo Grenzerlös gleich Grenzkosten ist.

Kosten bei linearer Preisbeschaffungsfunktion PBF

Angenommen, die inverse Angebotskurve (PBF des Monopsons) ist linear

$$w(x) = a + bx$$

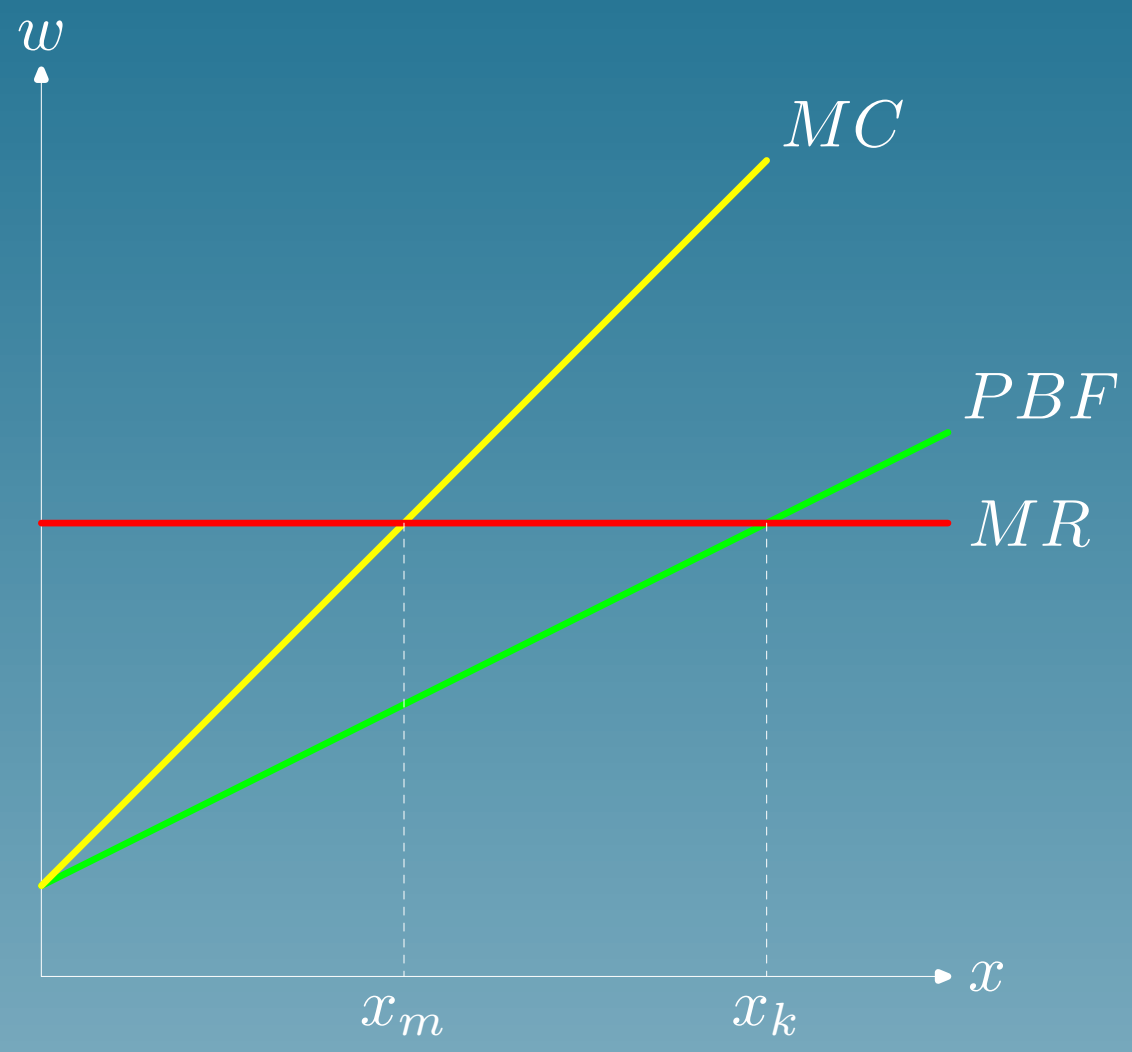
Die inputorientierten Kosten beim Monopson sind dann quadratisch

$$c(x) = w(x)x = ax + bx^2$$

und die inputorientierten Grenzkosten sind wieder linear, steigen aber doppelt so stark wie die PBF

$$c'(x) = a + 2bx$$

Gewinnmaximum bei linearer PBF und Produktionsfunktion



Oligopol

Gibt es nur wenige Anbieter, dann liegt die Marktform des Oligopols vor. Die einzelnen Anbieter haben einen merklichen Einfluss auf das Marktgeschehen, und ihre Entscheidungen sind durch eine strategische Interdependenz gekennzeichnet.

Betrachtet werden nur sogenannte Duopole, das sind Oligopole mit zwei Anbietern.

Auch wenn die Spieltheorie erst im Anschluss behandelt wird, kommen schon jetzt Begriffe und Betrachtungsweise der (mathematischen) Theorie der strategischen Spiele zur Sprache.

Strategien im Duopol

Bei nur zwei Anbietern gibt es insgesamt vier Größen, welche sie wählen können: die beiden Preise und die beiden Mengen.

Kann eine Firma den Preis oder ihre Menge vor der anderen festlegen, liegt ein sogenanntes sequentielles Spiel vor. Die beginnende Firma hat dabei den „Vorteil des ersten Zuges“, während die andere lediglich darauf reagiert. Im Rahmen der Oligopoltheorie nennt man diese Verhaltensweisen Preis- oder Mengenführerschaft und Preis- oder Mengengefolgschaft.

Beide Firmen können ihre Preise oder Mengen auch simultan planen, wobei Prognosen über die Aktivität des Mitanbieters gebildet werden. Diese Situation wird simultanes Spiel genannt. Auch hierbei gibt es Mengen- oder Preisstrategien.

Schließlich können sich beide Firmen auch auf ein gemeinsames Verhalten verabreden. Dies ist dann ein kooperatives Spiel und wird im Rahmen der Oligopoltheorie als ein Kartell bezeichnet.

Mengenführerschaft und -gefolgschaft (Stackelberg)

Firma 1 ist der Mengenführer und Firma 2 passt sich mit ihrer Planung an die von Firma 1 an. Beide wollen ihren Gewinn maximieren. Der Mengenführer (Firma 1) berücksichtigt bei der Planung des Outputs die Reaktion von Firma 2. Umgekehrt nimmt die Firma 2 die angebotene Menge y_1 von Firma 1 als gegeben hin.

Da Firma 1 die Reaktion von Firma 2 einplant, ist diese zuerst zu betrachten.

Die gemeinsame PAF ist die inverse Marktnachfragekurve $p(y_1 + y_2)$.

Planung von Firma 2

Die Firma 2 wählt — bei gegebenem Output y_1 des Mengenführers — jenen Output y_2, welcher ihren Gewinn maximiert.

$$\max_{y_2} p(y_1 + y_2)y_2 - c_2(y_2)$$

Nach wie vor gilt die allgemeine Regel $MR = MC$, nur wird diesmal der Grenzerlös des Mengenfolgers durch y_1 beeinflusst.

$$MR_2 = p'y_2 + p = c_2' = MC_2$$

Reaktionsfunktion bei linearer PAF

Angenommen, die gemeinsame PAF ist linear $p(y_1 + y_2) = a - b(y_1 + y_2)$ und die Grenzkosten der Produktion sind bei beiden Firmen gleich Null. Dann ergibt sich als Gewinn der Firma 2

$$\pi_2(y_1, y_2) = p(y_1 + y_2)y_2 = ay_2 - by_1y_2 - by_2^2$$

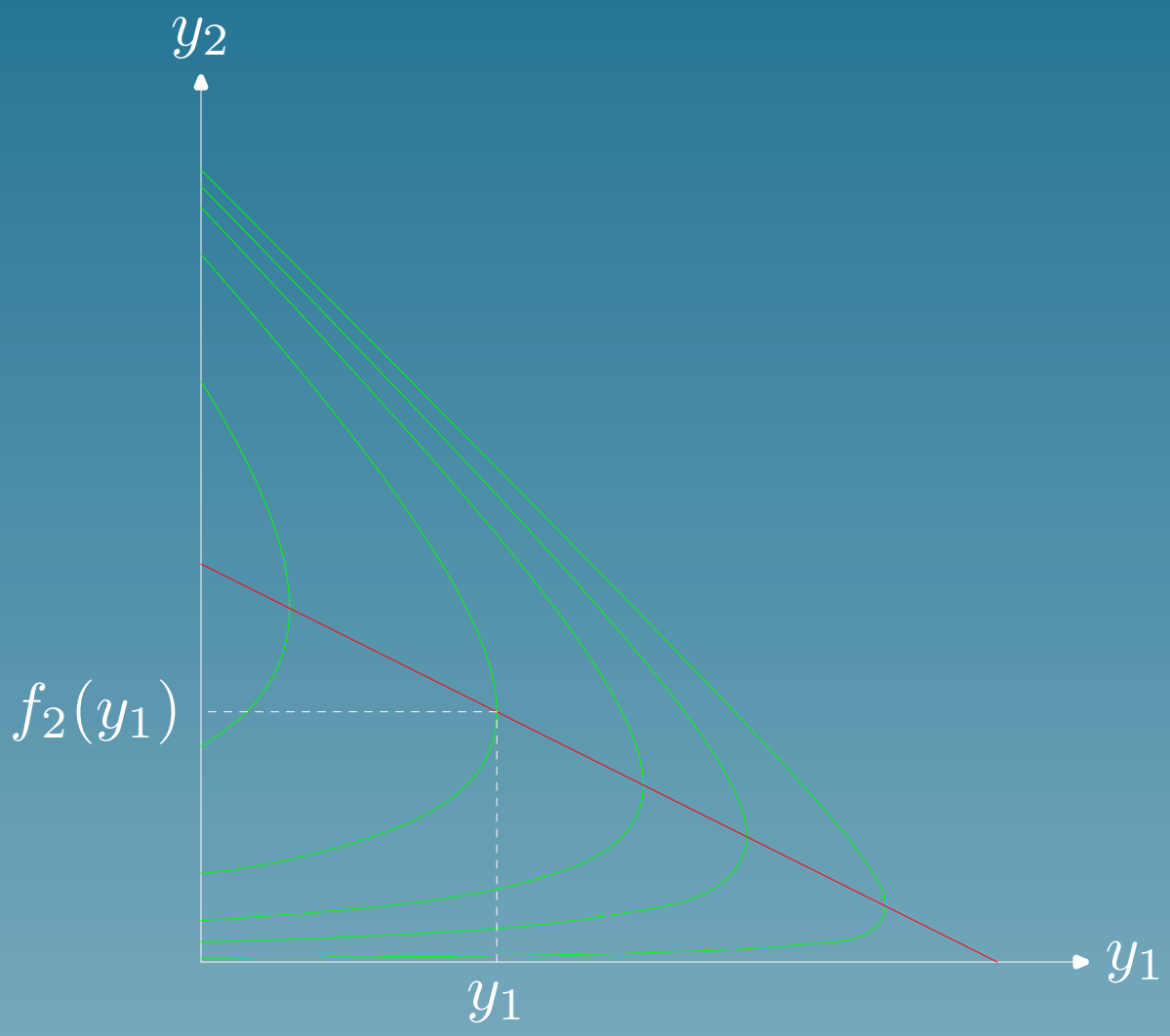
der vom Output y_1 des Mengenführers beeinflusst wird.

Die erste Ableitung gleich Null gesetzt ergibt

$$\frac{\partial \pi_2(y_1, y_2)}{\partial y_2} = 0 \rightsquigarrow a - by_1 - 2by_2 = 0 \rightsquigarrow y_2 = \frac{a - by_1}{2b} = f_2(y_1)$$

Der gewinnmaximale Output von Firma 2 in Abhängigkeit des Outputs von Firma 1 wird als Reaktionsfunktion des Mengenfolgers bezeichnet. Sie ist die beste Antwort auf die Entscheidung des Mengenführers.

Reaktionsfunktion und Isogewinnlinien von Firma 2



Planung des Mengenführers (Firma 1)

Firma 1 maximiert ihren Gewinn unter Beachtung der Reaktion von Firma 2

$$\max_{y_1} p(y_1 + y_2)y_1 - c_1(y_1) \quad \text{wobei} \quad y_2 = f_2(y_1)$$

Wird die Reaktionsfunktion von Firma 2 eingesetzt, dann ist der zu maximierende Gewinn von Firma 1 nur noch von ihrem Output abhängig.

$$\pi_1(y_1) = p\big(y_1 + f_2(y_1)\big)y_1 - c_1(y_1)$$

Gewinnmaximum des Mengenführers

Der Gewinn des Mengenführers wird maximal für

$$MR_1 = p'[1 + f_2']y_1 + p = c_1' = MC_1$$

wobei der Term $p' f_2' y_1$ im Grenzertrag neu ist. Er gibt die Preissenkung an, welche durch die Reaktion der Firma 2 ausgelöst wird.

Im Falle einer linearen Preisabsatzfunktion $p = a - b(y_1 + y_2)$ und Grenzkosten in Höhe von Null erhalten wir die spezielle Gewinnfunktion

$$\pi_1 = ay_1 - by_1^2 - by_1 y_2 \quad \text{wobei} \quad y_2 = \frac{a - by_1}{2b}$$

Einsetzen der Reaktionsfunktion von Firma 2 ergibt

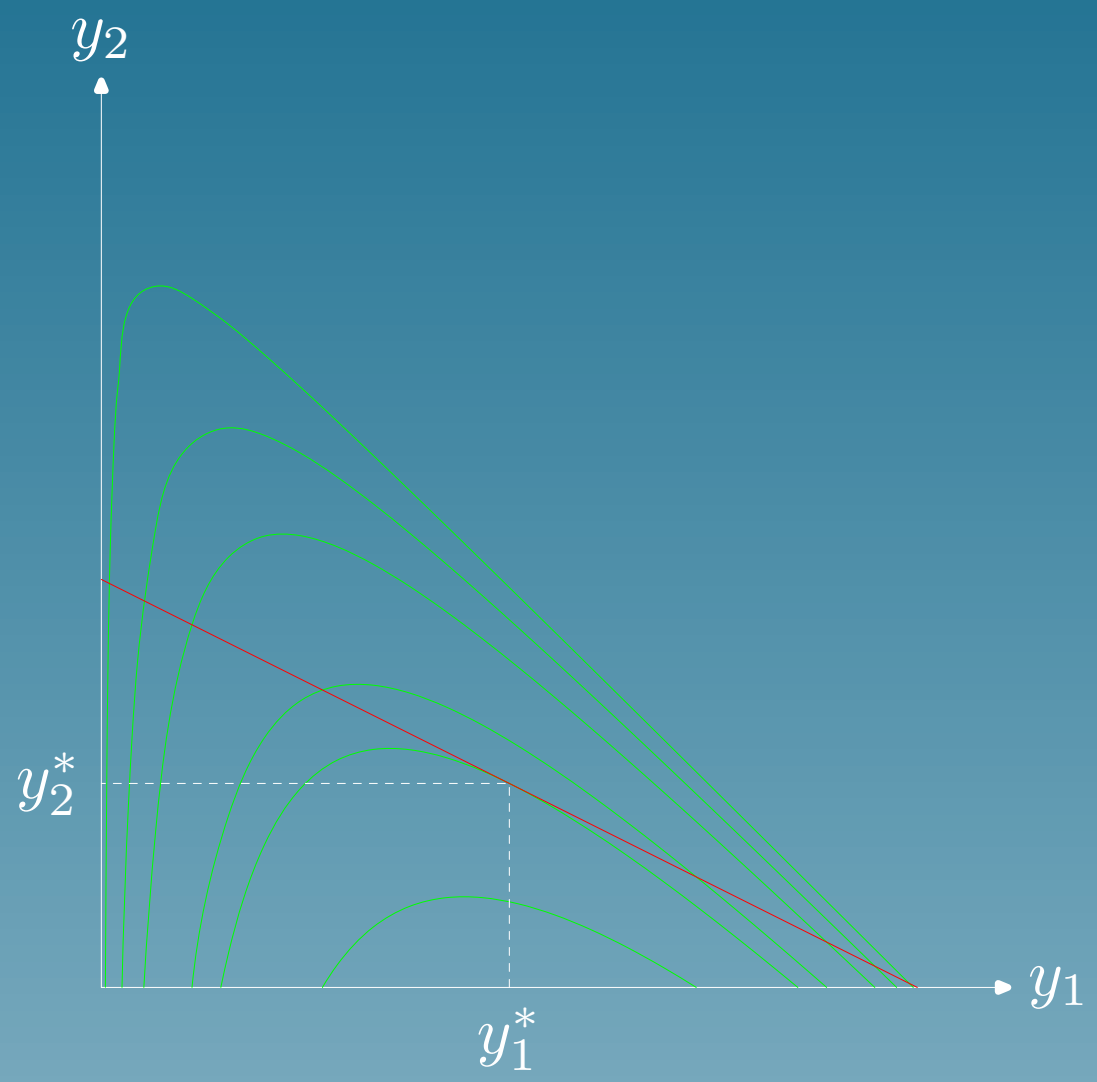
$$\pi_1 = \frac{a}{2}y_1 - \frac{b}{2}y_1^2$$

und der gewinnmaximale Output des Mengenführers beläuft sich auf

$$\pi_1' = 0 = \frac{a}{2} - by_1 \quad \rightsquigarrow \quad y_1 = \frac{a}{2b}$$

Der Mengenführer wählt die Hälfte der Sättigungsmenge, also denselben Output, den auch ein Monopolist bei dieser PAF und Grenzkosten von Null wählen würde.

Isogewinnlinien von 1 und Reaktionsfunktion von 2



Preisführerschaft

Anstelle der Menge setzt der Marktführer (Firma 1) den Preis und der Nachahmer (Firma 2) akzeptiert diesen Preis.

Ähnlich wie beim Mengenführer plant der Preisführer den Preis unter Berücksichtigung der dabei angebotenen Menge des Nachahmers. Dieser verhält sich wie eine Firma bei vollständiger Konkurrenz, d.h. Firma 2 plant bei einem gegebenen Preis p jenen Output y_2, bei welchem die Grenzkosten gleich diesem Preis sind.

Die Angebotskurve $S_2(p)$ von Firma 2 ist jener Teil ihrer Grenzkostenkurve, der oberhalb der Angebotsschwelle verläuft.

Restnachfrage des Preisführers

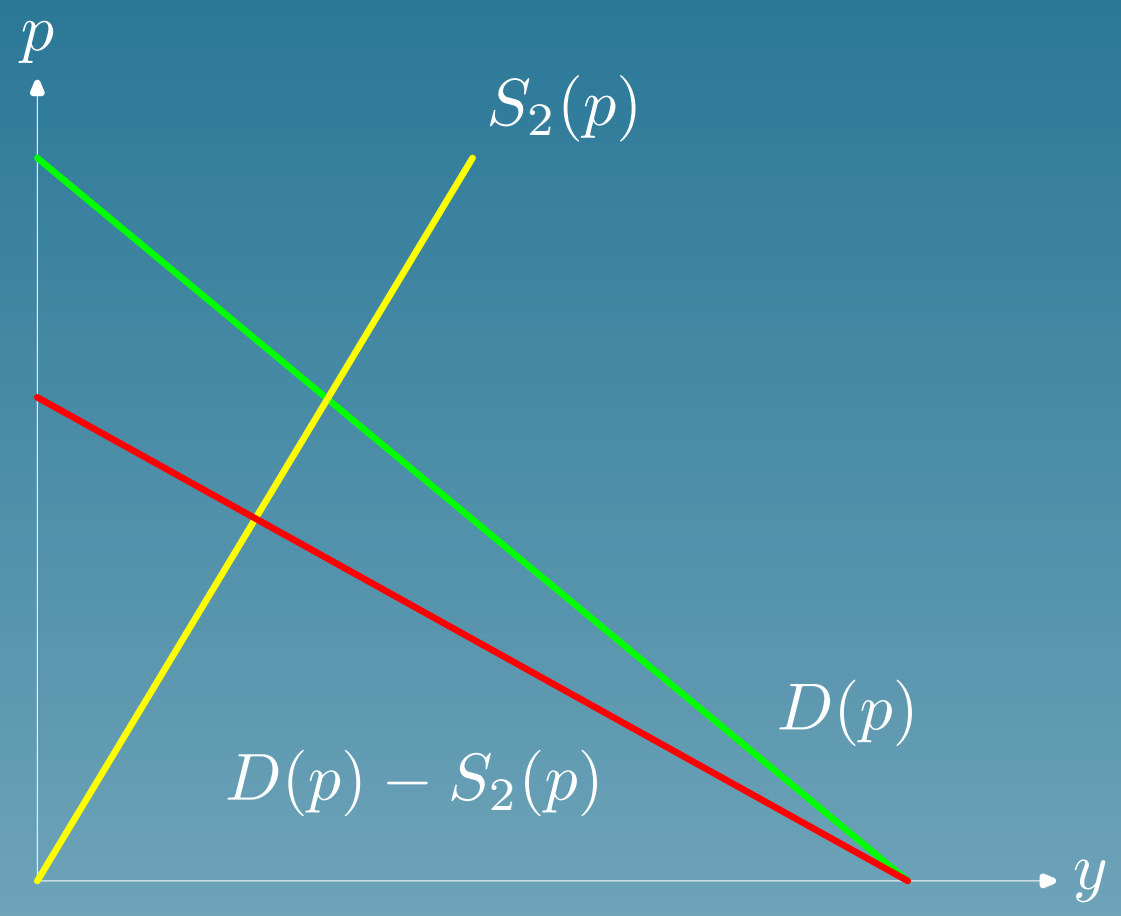
Wenn der Preisführer einen Preis p setzt, wird Firma 2 die Menge $S_2(p)$ anbieten und damit verbleibt für ihn die Restnachfrage $y_1 = D(p) - S_2(p)$.

Diese Restnachfrage ist seine inverse PAF und wird der Planung des maximalen Gewinns zugrunde gelegt.

Falls der Preisführer mit konstanten Grenzkosten (und Durchschnittskosten) produziert $c_1(y_1) = \alpha y_1$, kann sein Gewinn auch direkt in Abhängigkeit des Preises formuliert werden:

$$py_1 - \alpha y_1 = (p - \alpha)y_1 = (p - \alpha)\big(D(p) - S_2(p)\big) = \pi_1(p)$$

Restnachfrage als horizontale Subtraktion



Gewinnmaximum des Preisführers

Der Preisführer wählt den Preis, bei dem sein Gewinn $\pi_1(p)$ maximal wird

$$\frac{\partial \pi_1(p)}{\partial p} = \big(D(p) - S_2(p)\big) + (p - \alpha)\big(D'(p) - S_2'(p)\big) = 0$$

$$\rightsquigarrow \quad MR_1 = p + \frac{D(p) - S_2(p)}{D'(p) - S_2'(p)} = \alpha = MC_1$$

Beispiel 8

Angenommen, die Marktnachfragekurve ist linear $D(p) = a - bp$ und die Angebotskurve des Preisnachahmers (Firma 2) ist ebenfalls linear $S_2(p) = cp$.

Die Restnachfrage, der sich der Preisführer (Firma 1) gegenüber sieht, lautet dann

$$y_1 = D(p) - S_2(p) = a - (b + c)p$$

Die (outputorientierten) Kosten des Preisführers sollen linear sein

$$c_1(y_1) = \alpha y_1$$

Damit ist der Gewinn in Abhängigkeit des Preises gegeben durch

$$py_1 - c_1(y_1) = (p - \alpha)y_1 = (p - \alpha)\big(a - (b + c)p\big) = \pi_1(p)$$

Der gewinnmaximale Preis ergibt sich, indem man die Gewinnfunktion des Preisführers nach dem Preis abgeleitet, den Ausdruck Null setzt und nach p auflöst.

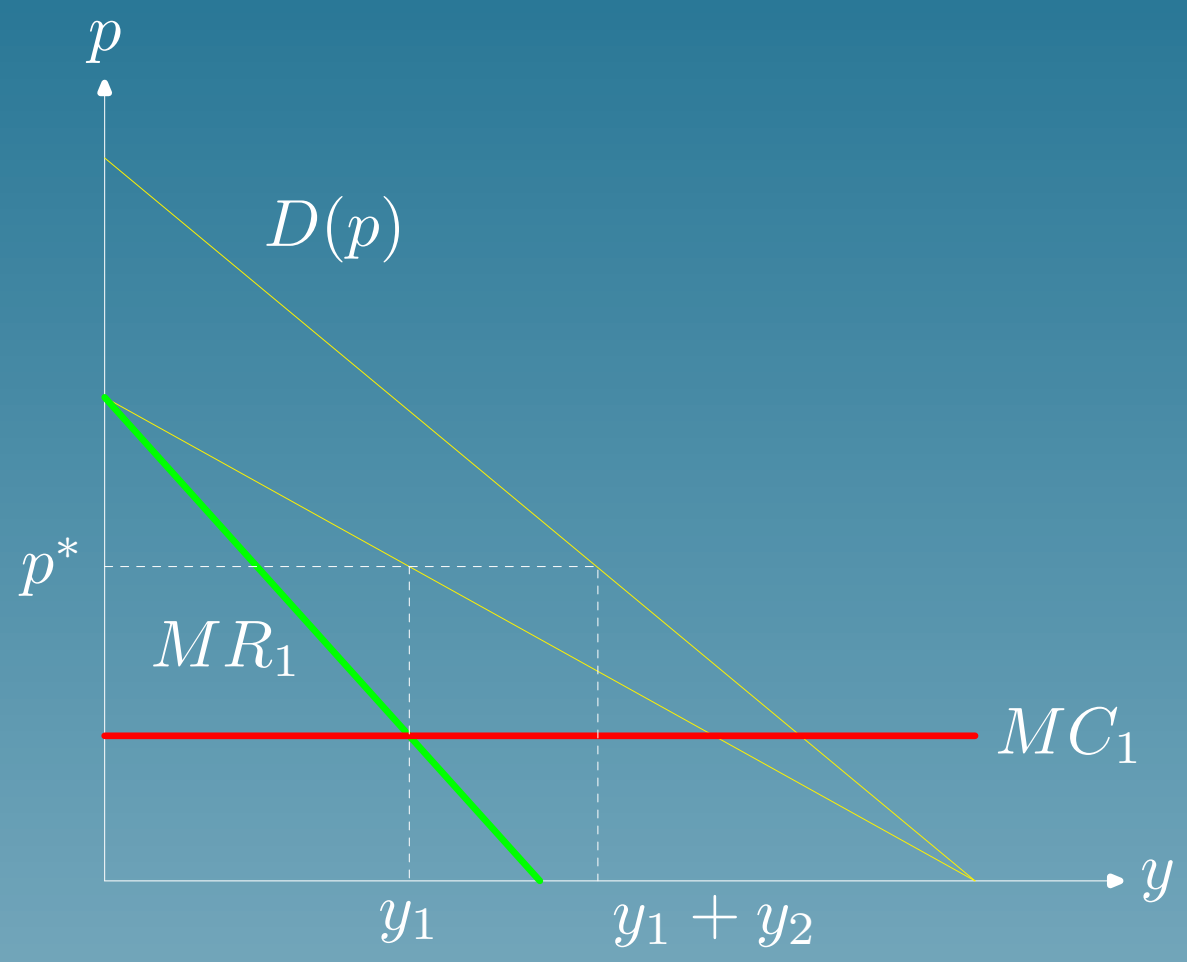
$$\frac{\partial \pi_1(p)}{\partial p} = a + \alpha(b+c) - 2(b+c)p = 0 \quad \rightsquigarrow \quad p^* = \frac{a}{2(b+c)} + \frac{\alpha}{2}$$

Das gewinnmaximale Angebot des Preisführers ist damit

$$y_1^* = D(p^*) - S_2(p^*) = a - (b+c)p^* = \frac{a}{2} - \frac{\alpha(b+c)}{2}$$

Der Punkt (y_1^*, p^*) ist der Cournot'sche Punkt des Preisführers auf seiner Restnachfragekurve.

Aufteilung des Angebotes bei Preisführerschaft



Simultane Mengenplanung (Cournot)

Beide Firmen planen ihren gewinnmaximalen Output auf der Grundlage einer Prognose des Angebots der jeweils anderen Firma.

Bezeichnet y_2^e die Prognose von Firma 1 über den Output von Firma 2, dann lautet ihre Optimierungsaufgabe:

$$\max_{y_1} p(y_1 + y_2^e)y_1 - c_1(y_1)$$

Analog maximiert Firma 2 ihren Gewinn, gegeben ihre Erwartung y_1^e über den Output von Firma 1

$$\max_{y_2} p(y_1^e + y_2)y_2 - c_2(y_2)$$

Reaktionsfunktionen im Cournot-Modell

Der gewinnmaximale Output von Firma 1 wird beeinflusst durch y_2^e. Dies führt zu ihrer Reaktionsfunktion (beste Antwort)

$$y_1 = f_1(y_2^e)$$

Dasselbe gilt auch bei Firma 2, ihre jeweils beste (gewinnmaximierende) Antwort auf einen prognostizierten Output von Firma 1 ist die Reaktionsfunktion

$$y_2 = f_2(y_1^e)$$

Cournot-Gleichgewicht

Das Cournot-Gleichgewicht ist dadurch gekennzeichnet, dass die Produktionsentscheidungen und -prognosen der beiden Firmen wechselseitig miteinander vereinbar sind:

$$y_1^* = f_1(y_2^*) \quad \text{und} \quad y_2^* = f_2(y_1^*)$$

Dies ist gerade in den Schnittpunkten der beiden Reaktionsfunktionen erfüllt (falls sie sich bei positiven Werten schneiden).

Beispiel 9

Angenommen, die inverse Marktnachfragekurve (die gemeinsame PAF) ist linear $p(y_1+y_2) = a-b(y_1+y_2)$ und die Grenzkosten der Produktion sind in beiden Firmen Null.

Dann lautet die Gewinnmaximierungsaufgabe von Firma 1

$$\max_{y_1}\big(a - b(y_1 + y_2^e)\big)y_1$$

Die Ableitung des Gewinns Null gesetzt und nach y_1 aufgelöst ergibt die Reaktionsfunktion von Firma 1

$$\big(ay_1 - by_1^2 - by_2^e y_1\big)' = a - 2by_1 - by_2^e = 0 \quad \rightsquigarrow \quad y_1 = \frac{a - by_2^e}{2b}$$

Dasselbe Verfahren ergibt eine Reaktionsfunktion für die Firma 2

$$y_2 = \frac{a - by_1^e}{2b}$$

Falls nun Firma 1 die Produktion von Firma 2 korrekt prognostiziert und umgekehrt, ergibt sich das Gleichungssystem

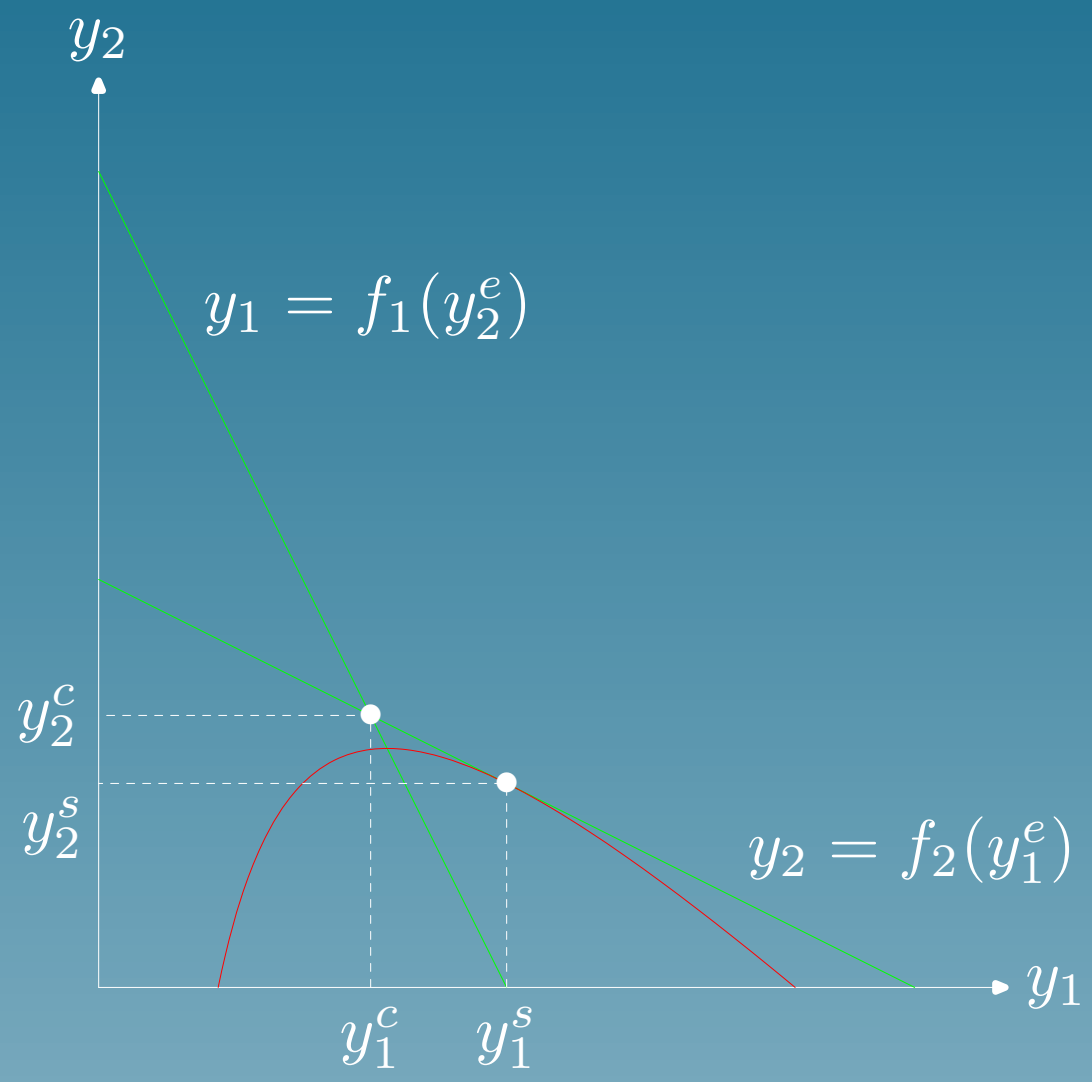
$$y_1^* = \frac{a - by_2^*}{2b} \quad \text{und} \quad y_2^* = \frac{a - by_1^*}{2b}$$

mit der Lösung

$$y_1^* = y_2^* = \frac{a}{3b}$$

d.h. jede Firma bietet ein Drittel der Sättigungsmenge an.

Schnittpunkt der Reaktionsfunktionen



Das Cournot-Gleichgewicht befindet sich im Schnittpunkt der beiden Reaktionsfunktionen.

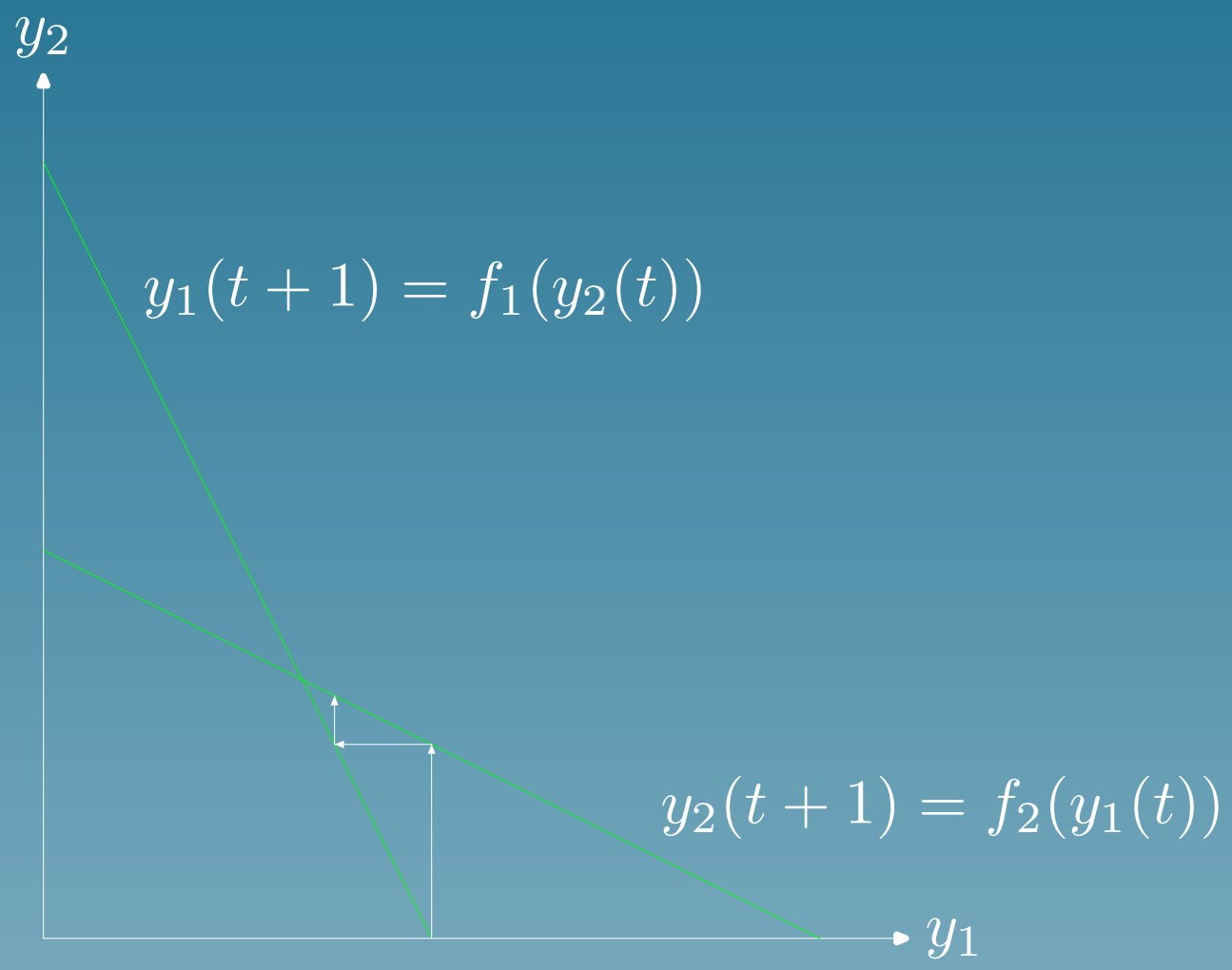
Das Stackelberg-Gleichgewicht befindet sich — bei gleichen Rahmenbedingungen — im Tangentialpunkt einer Isogewinnlinie des Marktführers mit der Reaktionsfunktion des Nachahmers.

Sequentielle Mengenplanung

Man kann das Cournot-Modell auch dynamisch interpretieren. Beide Duopolisten planen dabei ihr Angebot unter Berücksichtigung des zuletzt beobachteten Angebots des Mitanbieters.

$$y_1(t+1) = f_1\big(y_2(t)\big) \quad \text{und} \quad y_2(t+1) = f_2\big(y_1(t)\big)$$

Anpassung an das Cournot-Gleichgewicht



Simultane Preisplanung (Bertrand)

Angenommen, die beiden Firmen bieten ein identisches Gut an und ihre Grenzkosten sind konstant. Wenn beide denselben Preis setzen, teilen sie sich die Marktnachfrage und wenn sie unterschiedliche Preise setzen, verliert die Firma mit dem höheren Preis ihren Absatz.

Die Nachfrage der sich Firma $i(=1,\,2)$ gegenüber sieht lautet

$$D_i(p_i,\,p_j) = \begin{cases} D(p_i) & \text{falls } \; p_i < p_j \\ \frac{1}{2}D(p_i) & \text{falls } \; p_i = p_j \\ 0 & \text{falls } \; p_i > p_j \end{cases}$$

Der Gewinn von Firma i beträgt $\pi_i(p_i,\,p_j) = (p_i - c)D_i(p_i,\,p_j)$.

Bertrand-Gleichgewicht

Der Gleichgewichtspreis (das Bertrand-Gleichgewicht) ist gleich den Grenzkosten

$$p* = p_i^* = p_j^* = c$$

Dieses überraschende Resultat lässt sich folgendermaßen beweisen:

Angenommen, $p_i > p_j > c$, dann wird Firma i keinen Gewinn machen, wohl aber Firma j. In dieser Situation besteht die beste Antwort von Firma i darin ihren Preis etwas unterhalb den von Firma j zu setzen, also $p_i = p_j - \epsilon > c$ (dabei bezeichnet ϵ eine kleine positive Zahl). Dadurch erhält Firma i die gesamte Nachfrage und erzielt einen positiven Gewinn.

Die beste Antwort von Firma j besteht nun darin ihrerseits den Preis knapp unterhalb von p_i zu setzen, um damit die Nachfrage zurückzugewinnen. Damit ist gezeigt, dass unterschiedliche Preise oberhalb der Grenzkosten keinen Bestand haben und somit kein Gleichgewicht sind.

Angenommen, $p_i = p_j > c$, dann teilen sich beide Firmen den Markt, aber jede hat einen Anreiz den Preis zu senken, da dies ihren Absatz verdoppelt. Folglich können auch identische Preise, die größer als die Grenzkosten sind, keinen Bestand haben.

Falls nun $p_i > p_j = c$ gilt, könnte Firma j einen positiven Gewinn erzielen, indem sie ihren Preis über die Grenzkosten anhebt, aber noch unter p_i bleibt. Dann liegt aber wieder die erste Situation vor, von der wir wissen, dass sie kein Gleichgewicht sein kann.

Gemeinsame Gewinnmaximierung

Abgestimmte Verhaltensweisen von Firmen zur gemeinsamen Gewinnmaximierung werden auch Kollusion genannt, und die Firmengruppe wird auch als Kartell bezeichnet. Im Rahmen der Spieltheorie spricht man in diesem Fall von kooperativen Strategien.

Im Falle von zwei Firmen (eines Duopols) sind jene Outputs y_1 und y_2 zu wählen, welche den Gewinn der Industrie (und nicht einer Firma) maximieren:

$$\max_{y_1,\, y_2} \; p(y_1 + y_2)\big(y_1 + y_2\big) - c_1(y_1) - c_2(y_2)$$

Die Bedingungen erster Ordnung verlangen, dass die Ableitungen des gemeinsamen Gewinns nach den beiden Outputs Null werden:

$$\frac{\partial p}{\partial(y_1+y_2)}\frac{\partial y_1+y_2}{\partial y_1}(y_1+y_2)+p\frac{\partial y_1+y_2}{\partial y_1}-c_1'=0$$
$$\frac{\partial p}{\partial(y_1+y_2)}\frac{\partial y_1+y_2}{\partial y_2}(y_1+y_2)+p\frac{\partial y_1+y_2}{\partial y_2}-c_2'=0$$

Da $\partial(y_1+y_2)/\partial y_1=1$ und $\partial(y_1+y_2)/\partial y_2=1$, ergibt sich

$$p'(y_1+y_2)+p=c_1'(y_1)=MC_1$$
$$p'(y_1+y_2)+p=c_2'(y_2)=MC_2$$

Da beide linken Seiten gleich sind müssen die Grenzkosten ebenfalls gleich sein $c_1'(y_1)=c_2'(y_2)$.

Aufteilung der Produktion und des Gewinns

Bei gemeinsamer Gewinnmaximierung handeln beide Firmen so wie ein Monopol mit zwei Betriebsstätten. Die Aufteilung der Produktion auf die beiden Firmen erfolgt nach dem Kriterium gleicher Grenzkosten. Sind die Grenzkostenkurven beider Firmen beispielsweise identisch, dann produzieren beide die Hälfte. Bei unterschiedlichen Grenzkostenkurven produziert die Firma mit den niedrigeren Grenzkosten mehr als die andere.

Sind die Grenzkosten konstant, aber verschieden, dann kann die Bedingung $c_1'(y_1) = c_2'(y_2)$ bei keiner Aufteilung der Produktion erfüllt werden.

Anreiz zum Bruch der Vereinbarung

Angenommen, die Bedingungen zur gemeinsamen Gewinnmaximierung sind für die Angebotsmengen y_1^* und y_2^* erfüllt. Dann hat jede Firma einen Anreiz entgegen der Vereinbarung mehr zu produzieren, da dadurch ihr Gewinn steigt.

Wenn beispielsweise Firma 1 ihren Output um Δy_1 vergrößert, dann sinkt der Preis und ihr Absatz steigt und ihre Grenzkosten werden zumindest nicht fallen. Die Veränderungsrate ihres Gewinns beträgt

$$\frac{\Delta \pi_1(y_1^*, y_2^*)}{\Delta y_1} = \frac{p(y_1^* + \Delta y_1 + y_2^*) - p(y_1^* + y_2^*)}{\Delta y_1} y_1^* + p(y_1^* + \Delta y_1 + y_2^*) - \frac{c_1(y_1^* + \Delta y_1) - c_1(y_1^*)}{\Delta y_1}$$

Im Grenzübergang $\Delta y_1 \to 0$ ergibt sich für den Newton-Quotienten

$$\lim_{\Delta y_1 \to 0} \frac{\Delta \pi_1(y_1^*, y_2^*)}{\Delta y_1} = \frac{\partial \pi_1}{\partial y_1} = p'(y_1^* + y_2^*)y_1^* + p(y_1^* + y_2^*) - c_1'(y_1^*)$$

Die Bedingung erster Ordnung für das gemeinsame Gewinnmaximum lautet

$$p'(y_1^* + y_2^*)y_1^* + p'(y_1^* + y_2^*)y_2^* + p(y_1^* + y_2^*) - c_1'(y_1^*) = 0$$

Daher führt wegen

$$\frac{\partial \pi_1(y_1^*, y_2^*)}{\partial y_1} = -p'(y_1^* + y_2^*)y_2^* > 0$$

eine Produktionssteigerung über y_1^* hinaus zu einem größeren Gewinn bei Firma 1.

Strafmaßnahmen im Kartell

Die an einem Kartell beteiligten Firmen haben einen Anreiz von der Vereinbarung abzuweichen. Eine Möglichkeit, um die Mitglieder „bei der Stange zu halten“, besteht darin, Strafen bei Fehlverhalten anzudrohen.

Im Falle eines Duopols mit zwei identischen Firmen besteht eine solche Strafandrohung beispielsweise in folgendem Kommunique:

- Solange Du (die andere Firma) die Hälfte der Monopolmenge produzierst, mache ich das gleiche.
- Sollte ich Dich jedoch beim Versuch erwischen, mehr als diese Gütermenge zu produzieren, werde ich zukünftig nur noch die Cournot-Menge produzieren.

Glaubwürdigkeit der Grim-Strategie

Im Rahmen der Spieltheorie wird das voranstehende Kommunique als die sogenannte Grim-Strategie beim wiederholten Spiel bezeichnet.

Falls beide Firmen die Hälfte der Monopolmenge produzieren, erzielt jede einen Gewinn in Höhe von π_m pro Periode. Der Gegenwartswert (Firmenwert) im Falle des dauerhaften Kartells ist somit $\pi_m + \pi_m/r$, wobei r der Zinssatz ist.

Falls jedoch eine Firma mehr produziert, dann kann sie dies nur eine Periode lang absetzen, danach gibt es nur noch die Cournot-Mengen. Der Gegenwartswert dieser Gewinne beläuft sich auf $\pi_d + \pi_c/r$.

Die Grim-Strategie stellt eine echte Strafe dar, wenn sich das Abweichen von der Kartell-Lösung für eine Firma nicht lohnt, also gilt

$$\pi_m + \frac{\pi_m}{r} > \pi_d + \frac{\pi_c}{r} \quad \rightsquigarrow \quad r < \frac{\pi_m - \pi_c}{\pi_d - \pi_m}$$

Spieltheorie

Die Verhaltensweisen der Firmen im Duopol sind ein Beispiel für strategische Interaktion von Wirtschaftseinheiten. Diese können auch mittels der Spieltheorie untersucht werden.

Mittels der Spieltheorie können eine Vielzahl anderer sozialer Interaktionen untersucht werden (beispielsweise Wahlkämpfe, internationales Konfliktmanagement, Wettrüsten etc).

Strategische Spiele

Ein strategisches Spiel besteht aus einem Satz abstrakter Regeln, welche das Verhalten der Spieler und das Spielergebnis auf der Grundlage der durchgeführten Spielzüge beschreiben.

Um ein wohldefiniertes strategisches Spiel zu charakterisieren, müssen die Regeln beschreiben

- wer die Spieler sind,
- ob der Zufall eine Rolle spielen soll oder nicht,
- in welcher Reihenfolge die Spieler ihre Züge machen und welche Informationen sie bei jedem Zug haben,
- wie am Ende das Spielergebnis zustande kommt.

Beispiel 10

Spieler A kann „oben“ oder „unten“ auf seinen Zettel schreiben und Spieler B kann „links“ oder „rechts“ auf seinen Zettel schreiben. Das sind die Strategien, welche die Spieler wählen können.

Nachdem die Spieler ihre Zettel ausgefüllt haben, werden sie eingesammelt und das Ergebnis ermittelt. Zu jeder Kombination von Strategien gibt es unterschiedliche Punkte für die Spieler. Diese sind in einer Auszahlungsmatrix zusammengefasst.

Normalform eines Spieles

Die sogenannte Normalform eines Spiels fasst die Strategien und die Auszahlungsmatrix in einer Tabelle zusammen.

		Spieler B	
		links	rechts
Spieler A	oben	$1,\ 2$	$0,\ 1$
	unten	$2,\ 1$	$1,\ 0$

Die linke Zahl in einem Kästchen stellt die Auszahlung an Spieler A dar und die rechte die an Spieler B.

Diese Darstellung ist insbesondere für Spiele geeignet, bei denen die Spieler gleichzeitig ihre Entscheidungen treffen.

Dominante Strategie

Bei der vorliegenden Auszahlungsmatrix ist es für Spieler A immer besser „unten“ zu schreiben. Falls Spieler B „links“ spielt, erhält er 2 (hätte Spieler A „oben“ gewählt, dann hätte er nur 1 als Auszahlung). Falls Spieler B andererseits „rechts“ wählt, bekommt er 1 (anstelle von 0, wenn er „oben“ gewählt hätte).

Für Spieler B ist es ebenfalls immer besser „links“ zu wählen.

Man bezeichnet eine Strategie als dominant, wenn sie eine höhere Auszahlung bewirkt als die anderen Strategien des Spielers und zwar bei jeder Wahl der Strategien der übrigen Spieler.

Gleichgewicht in dominanten Strategien

Falls es für einen Spieler eine dominante Strategie gibt, dann wird er diese allen übrigen vorziehen.

Falls jeder Spieler eine dominante Strategie besitzt, ist der Spielverlauf und das Ergebnis des Spiels vorhersagbar: Alle Spieler wählen ihre dominante Strategie. Man nennt diesen Zustand ein Gleichgewicht in dominanten Strategien.

Das voranstehende Spiel besitzt ein Gleichgewicht in dominanten Strategien.

Beispiel 11

Ein reales Beispiel eines Spiels mit dominanter Strategie ist die sogenannte Zweit-Preis-Auktion. Bei dieser Form der Versteigerung notiert jeder Teilnehmer sein Gebot auf einem Zettel. Der Auktionator erteilt dem höchsten Gebot den Zuschlag, der Gewinner muss jedoch nur den Preis des zweithöchsten Gebotes entrichten.

Die dominante Strategie jedes Teilnehmers besteht darin sein individuelles Höchstgebot auf dem Zettel zu notieren. Daher erzwingt die Zweit-Preis-Auktion die Ehrlichkeit der Teilnehmer.

Nash-Gleichgewicht

Das Nash-Gleichgewicht ist das grundlegende Gleichgewichtskonzept der Spieltheorie. Es beschreibt einen Zustand, bei dem jeder Spieler nicht von seiner gewählten Strategie abweichen will unter der Voraussetzung, dass die Strategien der Mitspieler gegeben sind.

Bezeichnen S_i die Strategiemengen der Spieler $i = 1, 2, \ldots, n$, s_i eine Strategie von Spieler i und $\pi_i(s_1, s_2, \ldots, s_n)$ die Auszahlung, dann ist $(s_1^*, s_2^*, \ldots, s_n^*)$ ein Nash-Gleichgewicht, wenn gilt:

$$\pi_i(s_1^*, s_2^*, \ldots, s_i^*, \ldots, s_n^*) \geq \pi_i(s_1^*, s_2^*, \ldots, s_i^+, \ldots, s_n^*)$$

für alle i und für alle Strategien $s_i^+ \in S_i$.

Um die Definition des Nash-Gleichgewichts zu verstehen, muss man unterstellen, dass die Spieler einen Zustand $(s_1^*, s_2^*, \ldots, s_n^*)$ erreicht haben, ohne zu fragen wie er zustande kam.

Davon ausgehend wird nun geprüft, ob ein Spieler i von seiner Strategie abweichen möchte, während die übrigen $n-1$ Spieler ihre Wahl beibehalten. Falls nun der Spieler i keinen Anreiz hat von s_i^* abzuweichen und dies reihum für alle Spieler gilt, liegt ein Nash-Gleichgewicht vor.

Die Strategie s_i^* ist die beste Antwort von Spieler i auf die Wahl von s_j^* der übrigen Spieler.

Beispiel 12

Es gibt im folgenden Spiel zwei Nash-Gleichgewichte in reinen Strategien.

		Spieler B	
		links	rechts
Spieler A	oben	$2,\ 1$	$0,\ 0$
	unten	$0,\ 0$	$1,\ 2$

Beachte, es gibt in diesem Spiel keine dominanten Strategien.

Gemischte Strategien

Es gibt auch Spiele (beispielsweise Papier-Stein-Schere), die zunächst kein Nash-Gleichgewicht haben. Werden die Strategien jedoch randomisiert, d.h. mit bestimmten Wahrscheinlichkeiten gewählt, dann gibt es immer (mindestens) ein Nash-Gleichgewicht. Randomisierte Strategien werden auch gemischte Strategien genannt.

Die Wahrscheinlichkeiten, mit denen ein Spieler seine reinen Strategien wählt, sind so bestimmt, dass die erwarteten Auszahlungen seines Gegenspielers bei allen Handlungsalternativen gleich sind.

Beispiel 13

Es gibt im folgenden Spiel kein Nash-Gleichgewicht in reinen Strategien.

		Spieler B	
		links	rechts
Spieler A	oben	$0,\ 0$	$0,\ -1$
	unten	$1,\ 0$	$-1,\ 3$

Wenn jedoch Spieler A in $3/4$ aller Fälle „oben“ wählt und mit der Wahrscheinlichkeit $1/4$ „unten“, während Spieler B seine Strategien je mit Wahrscheinlichkeit $1/2$ spielt, dann sind diese gemischten Strategien ein Nash-Gleichgewicht.

Die Berechnung der Wahrscheinlichkeiten für die gemischte Strategie von Spieler A ergibt sich wie folgt:

Mit Wahrscheinlichkeit $0 < q < 1$ wird „oben“ gewählt und mit der Gegenwahrscheinlichkeit $1-q$ „unten“. Die erwartete Auszahlung an Spieler B ist dann im Falle, dass er „links“ spielt,

$$E\,\pi_B(\text{links}) = q(0) + (1-q)(0)$$

und falls er „rechts“ spielt

$$E\,\pi_B(\text{rechts}) = q(-1) + (1-q)(3)$$

Die Wahrscheinlichkeit q ist nun so zu wählen, dass beide erwarteten Auszahlungen gleich sind.

$$q(0) + (1-q)(0) = 0 = q(-1) + (1-q)(3) \quad \rightsquigarrow \quad q = \frac{3}{4}$$

Gefangenen-Dilemma

Das Gefangenen-Dilemma ist ein Spiel, dessen Gleichgewicht in dominanten Strategien nicht pareto-optimal ist.

		Spieler B	
		links	rechts
Spieler A	oben	$-3,\ -3$	$0,\ -6$
	unten	$-6,\ 0$	$-1,\ -1$

Diese Situation ist typisch für Kartelle.

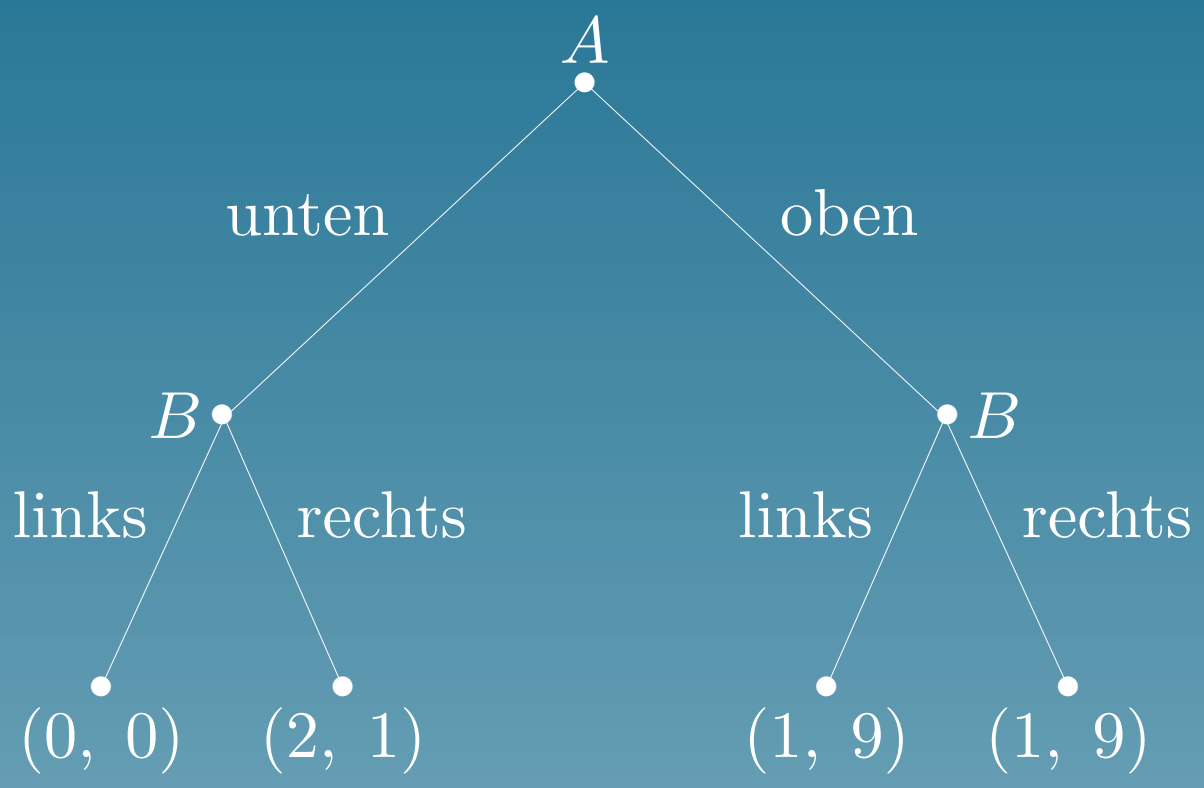
Sequentielle Spiele

Sehen die Spielregeln vor, dass einer der Spieler den ersten Zug macht und der andere darauf reagiert usw. handelt es sich um ein sequentielles Spiel.

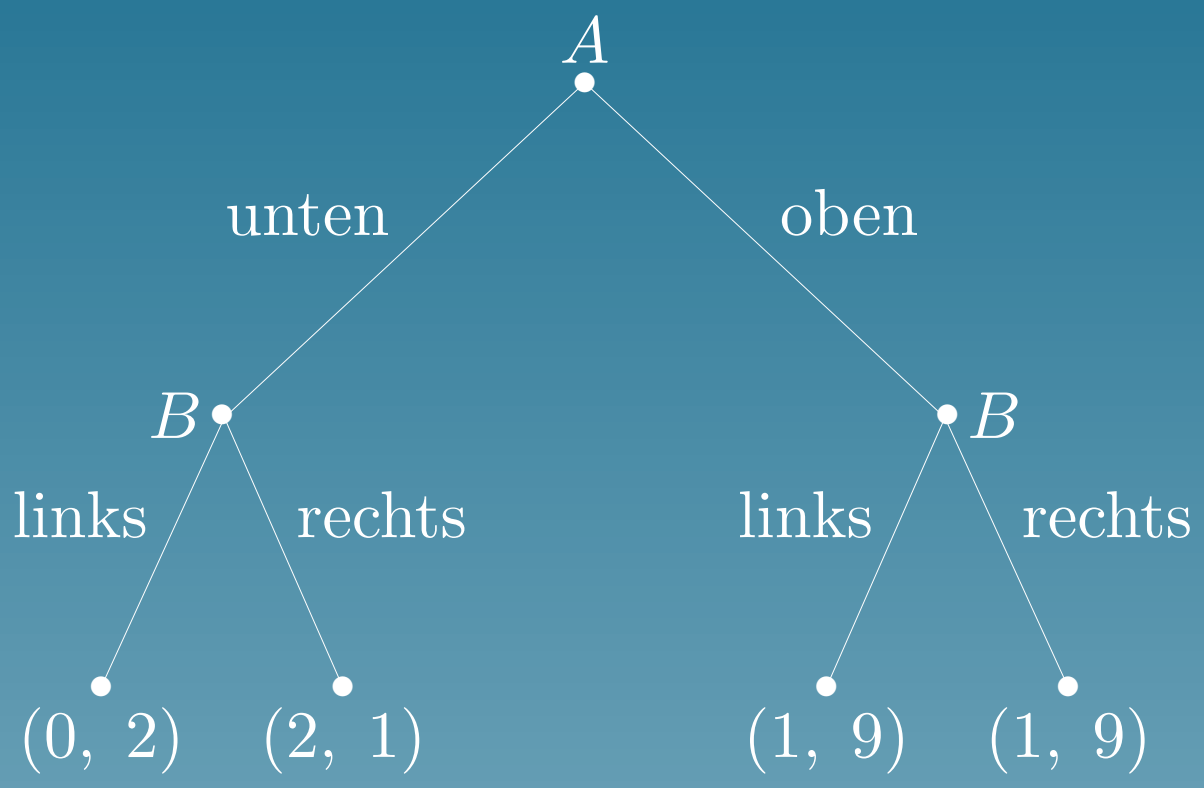
Das Stackelbergmodell ist ein Beispiel für ein sequentielles Spiel, bei dem der Vorteil des ersten Zuges deutlich zutage tritt (in manchen Spielen ist dies auch ein Nachteil).

Sequentielle Spiele lassen sich durch die sogenannte extensive Form eines Spiels besonders gut darstellen. Dazu werden die Spielzüge in einem Spielbaum angeordnet.

Extensive Form eines Spieles



Sequentielle Spiele werden mittels einer Rückwärtsinduktion untersucht. Dazu geht man von einem vorletzten Knoten (hier für Spieler B) aus und bestimmt seine beste Antwort. Danach geht man einen Knoten zurück (hier zum ersten Zug von A).

Das Spiel enthält die unglaubwürdige Drohung, dass Spieler B „links“ wählt, wenn zuvor Spieler A den Zug „unten“ gemacht hat.

Sequentielles Spiel mit glaubwürdiger Drohung