Fixpunkte (und einige weitere grundlegende Existenzsätze) **und ihre Anwendungen in Spieltheorie und Optimierung**

(Themen und Summary, Proseminar WiSem 2005/06 B. Kummer)

1 Projektionssatz im Hilbert Raum
2 Trennungssätze
3 Konvexe Hülle und Satz von Caratheodory
4 Brouwer's Fixpunktsatz
5 Kakutani's Fixpunktsatz
6 Ekeland's variational principle
7 MinimaxSatz und Nash-Gleichgewicht im R^n (über Kakutani)
8 MinimaxSatz (über Induktionsbeweis)
9 Dualität in konvexer Optimierung (als Anwendung des Minimax Satzes)
10 Variationsungleichung (Existenz über Zerlegung der Einheit)
11 Lineare Optimierung und Matrixspiele (gemischte Strategien, Interpretation)
12 Der Julia Robinson Algorithmus (fiktives Spielen) für Matrixspiele
13 Schnelle Modifikation dieses Algorithmus und numerische Beispiele
14 Lineare Optimierung als Matrixspiel
15 Nash-Verhandlungslösungen
16 Drohungen im Bimatrixspiel (nach Owen)
17 Michael's selection theorem
18 NM-Lösung
19 Core und balancierte Spiele
20 Shapley- Vektor (Existenz/Eindeutigkeit)
21 Auman's Drohstabilität
22 Edgeworth Markt
23 Walras Gleichgewicht
24 Nutzensfunktionen (Existenz)

1 Projektionen im (reellen) Hilbert Raum

das ist ein Banach Raum über R mit Skalarprodukt $\langle x , y\rangle$, so dass $\| x \|^2 = \langle x , x\rangle$.

1.1 Projektionssatz

Vor: X reeller Hilbert-Raum, $C \subset X$, konvex, abgeschlossen, nichtleer, $x \in X \setminus C$.

Beh. Es gibt ein ein $z \in C$ mit $\| x - z \| = \text{dist}(x, C) := \inf \{ \| x - \xi \| / \ \xi \in C \}$.

(im komplexen Hilbert Raum mit C = Teilraum von X: derselbe Beweis)

Beweis

Sei $\text{dist}(x, C) = d$. Man nutzt die **Parallellogrammgleichung**

(1.1) $\| a + b \|^2 + \| a - b \|^2 = 2 (\| a \|^2 + \| b \|^2)$,

die durch direktes Ausrechnen folgt. Sei $\{ z_n \}$ eine Folge in C mit $\| x - z_n \|^2 < d^2 + n^{-2}$.

Man zeigt Konvergenz der z_n durch Abschätzen von $\| z_m – z_n \|^2$.

Dazu wird (1.1) auf $\mathbf{a = x – z_n}$ und $\mathbf{b = x - z_m}$ angewandt.

$$\| z_m - z_n \|^2 + \| 2x - (z_m + z_n) \|^2 = 2 (\| x - z_n \|^2 + \| x - z_m \|^2).$$

Mit $(z_m + z_n) / 2 \in C$ und

$$\| 2x - (z_m + z_n) \|^2 = 4 \| x - (z_m + z_n) / 2 \|^2 \geq 4 d^2$$

ergibt das

$$\| z_m - z_n \|^2 \leq 2 (\| x - z_n \|^2 + \| x – z_m \|^2) - 4 d^2 \leq 2 (d^2 + n^{-2} + d^2 + m^{-2}) - 4 d^2 = 2 (n^{-2} + m^{-2}) .$$

Die Folge z_n ist damit eine Cauchy-Folge (= Fundamentalfolge), $z = \lim z_n$ ist der gesuchte Punkt. ♦

Einge Bemerkungen und Folgerungen

Ein solches z heißt ***Projektion*** von x auf C. Es ist *eindeutig*, was man über die Norm wie im R^n sofort ausrechnet (Bei $z' \neq z''$ wird der Abstand zu $½ (z' + z'')$ kürzer).

Man kann auch so argumentieren: Gibt es zwei Punkte $z' \neq z''$ mit kleinstem Abstand zu x, kann man obige z_n so auswählen: $z_n = z'$ für gerade n und $z_n = z''$ für ungerade n. Damit hat man einen Widerspruch zur bewiesenen Konvergenz der *ganzen* Folge.

Normaleneigenschaft
Es ist $\| y - p(y) \| \leq \| y - x \| + \| x - p(x) \|$. Sei $z = p(x)$ die Projektion von x auf C.
Dann gilt mit $0 < t < 1$ und $c \in C$ für den Punkt $a = c - p(x)$ (wegen Konvexität von C):

$$p(x) + t\,a \in C$$

und $h(t) := \| x - (p(x) + t\,a) \|^2 = \| x - p(x) \|^2 + t^2 \| a \|^2 - 2\,t \langle x - p(x), a\rangle$.
Wäre $\langle x - p(x), a\rangle > 0$, würde $h(t) < \|x - p(x)\|^2$ für kleine t folgen, was der Projektionseigenschaft widerspricht. Also gilt

(1.2) $\langle x - p(x), c - p(x)\rangle \leq 0 \quad \forall\ c \in C$, d.h. $x - p(x) \in N_C(p(x))$,

Def. $N_C(y) = \{u \in X \mid \langle u, c - y\rangle \leq 0\ \forall\ c \in C \}$ ist für $y \in C$ der *Normalenkegel an C in y* .
Andererseits charakterisiert (1.2) die Projektion p(x) vollständig, d.h. (1.2) und $p(x) \in C$ liefert gerade, dass p(x) die Projektion von x auf C ist. ♦

Orthogonalraum:
Sei U ein (abgeschl.) Teilraum von X. Dann gibt es zu jedem $x \in X$ *eindeutige* $u \in U$ und $v \in V$ mit

(1.3) $x = u + v$,

wenn nur V der durch $\{ v \mid \langle v , u \rangle = 0 \ \ \forall\, u \in U \} = N_U(0)$ definierte Orthogonalraum zu U ist.
Beweis: Man setze $u = p_U(x)$ (Proj. auf U), $v = x - u$. Damit gilt **(1.3)** sowie $u \in U$ und $v \in V$.
Eindeutigkeit: Wenn auch $x = u' + v'$, so folgt $u - u' = v' - v$, $u - u' \in U$ und $v' - v \in V$.
Weil nun $u - u' \in V \cap U$, muß gelten $\langle u- u' , u- u' \rangle = 0$, d.h. $u = u'$. ♦

Schwarzsche** (Bonjakowskische)* ***Ungleichung
Die Ungleichung $| \langle a , b\rangle |^2 \leq \langle a, a\rangle \langle b, b\rangle$, die letzlich (auf den Axiomen des Skalarproduktes basierend) die Beziehung zwischen Skalarprodukt und Norm im Hilbert-Raum herstellt, heißt Schwarzsche Ungleichung.

Beweis:
Wir schreiben den Beweis gleich für einen komplexen Hilbert Raum auf. Man betrachtet, für $b \neq 0$ und komplexe t, zunächst

$$0 \leq \langle a+ t\, b, a+ t\, b \rangle = \langle a,a\rangle + \langle a, t\, b\rangle + \langle t\, b, a\rangle + | t |^2 \langle b, b\rangle = \langle a,a\rangle + \mathrm{conj}(t) \langle a, b\rangle + t \langle b, a\rangle + | t |^2 \langle b, b\rangle$$

und setze $t = - \langle a, b) / \langle b,b\rangle$. Dann ist

$$0 \leq \langle a, a\rangle - [\langle b, a\rangle / \langle b,b\rangle] \langle a, b\rangle - [\langle a, b\rangle / \langle b,b\rangle] \langle b, a\rangle + | \langle a, b\rangle |^2 / \langle b, b\rangle ,$$

$$0 \leq \langle a, a\rangle \langle b, b\rangle - \langle b, a\rangle \langle a, b\rangle - \langle a, b\rangle \langle b, a\rangle + | \langle a, b\rangle |^2$$

Wegen $c\,\mathrm{conj}(c) = | c |^2$ erhält man so $\langle a,b\rangle \langle b,a\rangle = | \langle a, b\rangle |^2$ und $0 \leq \langle a,a\rangle \langle b, b\rangle - | \langle a, b\rangle |^2$. ♦

Lipschitzeigenschaft der Projektion
Mittels (1.2) zeigt man leicht, dass p(.) nichtexpansiv ist, d.h., $\| p(y) - p(x) \| \leq \| y - x \|$.
Wir erhalten nämlich $\langle x - p(x), p(y) - p(x) \rangle \leq 0$ mit $c = p(y)$

$\langle y - p(y), p(x) - p(y) \rangle \leq 0$ mit $c = p(x)$

und so $\langle x - p(x) - (y - p(y)), p(y) - p(x) \rangle \leq 0$.
Ausmultiplizieren ergibt $\langle x - y, p(y) - p(x)\rangle + \|p(y) - p(x)\|^2 \leq 0$,
und mittels Schwarzscher Ungleichung

$$\|p(y) - p(x)\|^2 \leq \langle x - y, p(x) - p(y)\rangle \leq \|x - y\| \|p(y) - p(x)\|.$$

die Behauptung. ♦

2 Trennungssätze

Satz 1.2 Sei $M \subset R^n$ nichtleer, abgeschl. konvex; $0 \notin M$. Dann exist. ein $u \in R^n$ mit
$0 < \inf_{x \in M} \langle u , x \rangle$.

Beweis:
Man nehme $u \in M$ mit kleinster *Euklidischer* Norm (u = Proj. von 0 auf M) und rechne nach.
Siehe auch Normaleneigenschaft der Projektion im Hibert Raum (wo Satz 1.2 ebenfalls gilt). ♦

Allgemeiner;
Wenn $0 \in \mathrm{bd}\, M$ (der Rand von M), so Grenzübergang: Man nehme Punkte $x^k \in R^n \setminus M$, $x^k \to 0$, setze $M^k = -x^k + M$ (Verschiebung von M) und wähle u^k als normierten Trennungsvektor von 0 und M^k mit Häuf.-Punkt u. Dann gilt noch

$$0 \leq \inf_{x \in M} \langle u , x \rangle, \qquad u \neq 0.$$

(hierzu braucht man R^n; im Hilbert Raum unendl. Dimension gibt es *beschränkte Folgen ohne konvergente Teilfolge,* etwa die Folge der Einheitsvektoren im l^2).

Trennung konvexer Mengen

Seien M, N konvexe, nichtleere Mengen im R^n, $\operatorname{int} M \cap N = \emptyset$, $\operatorname{int} M \neq \emptyset$.
Dann exist. ein $u \neq 0$ mit $\sup_{x \in M} \langle u, x \rangle \leq \inf_{y \in N} \langle u, y \rangle$.

Man benutze, dass die konvexe Menge X = N - M (elementweise Operation) innere Punkte besitzt (da int M nicht leer ist), aber $0 \notin \operatorname{int} X$ ist. Trennung von 0 und X liefert das Ergebnis. ♦

Trennungssatz (allgemein im normierten Raum)

Satz 1.3

Vor. Sei X ein (linearer) normierter Raum über R, $M \subset X$ sei offen, konvex, nichtleer, $0 \notin M$.
Beh. Es exist. $x^* \in X^* \setminus \{0\}$ mit $0 \leq x^*(x)$ $\forall x \in M$.

Bemerkung X^* ist der lineare Raum aller stetigen, additiven und homogenen Abbildungen $x^*: X \to R$, kurz der stetigen linearen Funktionale. Zumeist schreibt man $x^*(x)$ als (x^*, x).

Beweis:

Wir wählen ein festes $x^o \in M$ und betrachten den eindimensionalen Unterraum $U = \{ t x^o / t \in R \}$
Setzt man $g^o(t x^o) = t$, gilt offenbar

$$0 \leq g^o(x) \quad \forall x \in M \cap U.$$

Also erfüllt g^o die Forderung an x^* für $M \cap U$. Man erweitert nun g^o zu einer auf ganz X definierten linearen Abbildung. Sei dazu $G \supset U$ ein Unterraum von X (abgeschlossen oder nicht) und $g: G \to R$ additiv und homogen mit

$$0 \leq g(x)\ \forall x \in M \cap G, \quad \text{sowie } g = g^o \quad \text{auf } U.$$

Unter diesen Paaren (g , G) definieren wir eine Halbordnung:

$$(g, G) < (g', G') \text{ falls } G \subset G' \text{ und } g' = g \text{ für } x \in G.$$

Die Menge dieser (zulässigen) Paare ist nicht leer, z.B. nehme man G = U und setze $g = g^o$.
Für eine beliebige Kette (g^α, G^α) mit

$$(g^\alpha, G^\alpha) < (g^\beta, G^\beta) \quad \text{falls } \alpha < \beta$$

findet man dann eine obere Schranke (g', G') in $G' = \cup G^\alpha$ und $g'(x) = g^\alpha(x)$ falls $x \in G^\alpha$.
Damit gibt es *(Zornsches Lemma)* ein *maximales* Element bzgl. der Halbordnung, wir bezeichnen es erneut mit (g , G).

Wir zeigen indirekt, dass G = X gelten muß. Angenommen, es gibt ein $p \in X \setminus G$.
Dann gilt $G' := \operatorname{lin}(G \cup \{p\}) \supset G$, und $x' \in G'$ hat eine eindeutige Darstellung

$$x' = x + t p \quad (x \in G, \ t \in R).$$

Eindeutig:

Aus $x' = y + s p$ ($y \in G$, $s \in R$) folgt $0 = y - x + (s - t) p$. Da 0 und $y - x$ aus G sind, liegt auch $(s - t) p$ in G. Letzteres liefert $s = t$ wegen $p \notin G$. Somit ist auch $y = x$.

Wir definieren nun

$$\mathbf{g'(x') = g(x) + a\, t} \qquad \text{mit rellem a.}$$

Damit (g', G') zulässig wird, brauchen wir ein *spezielles a,* so dass

(1.4) $0 \leq g(x) + a t$ für alle $(x, t) \in (G, R)$ mit $x + t p \in M$.

Um a zu finden, unterscheiden wir positive und negative t , d.h. wir betrachten

$$(x, t) \text{ mit } t > 0 \quad \text{und} \quad (y, s) \text{ mit } s < 0.$$

Der Fall t = 0 liefert in (1.4) $0 \leq g(x)\ \forall\ x \in M \cap G$ nach Wahl von (g, G), was a nicht einschränkt. Also nimmt Forderung (1.4) die Form an:

(1.5) $0 \leq g(x) + a t \quad \forall (x, t) \in (G, R)$, $t > 0$, so dass $x + t p \in M$.

(1.6) $0 \leq g(y) + a s \quad \forall (y, s) \in (G, R)$, $s < 0$, so dass $y + s p \in M$.

Seien (x, t) und (y, s) wie oben beliebig fixiert. Wir betrachten (z, 0) als Konvexkombination mit

$$0 = \lambda t + (1-\lambda) s, \qquad z = \lambda x + (1-\lambda) y.$$

Dann ist $\lambda = -s / (t - s)$, $1 - \lambda = t / (t - s)$ und $z = (t - s)^{-1} (t y - s x)$

Dies bleibt auch richtig, wenn ein Skalar, s oder t, Null ist. Der Punkt z gehört offenbar zu G. Als Konvexkombination zweier Punkte aus M, ist $z = z + 0 p \in M$. Folglich gilt

$$0 \leq g(z) = (t - s)^{-1} (- s\, g(x) + t\, g(y))$$

und $s\, g(x) \leq t\, g(y)$ bzw. nach Division durch $-s t > 0$,

$$-g(x) / t \leq -g(y) / s .$$

Damit können wir a so wählen, dass die entscheidende Ungleichung

$\sup_{(x,t)}\ -g(x)/t \ \le\ a\ \le\ \inf_{(y,s)}\ -g(y)/s$

gilt. Sie liefert $-g(x) \le a\,t$ und $a\,s \ge -g(y)$, wonach (1.5) und (1.6) richtig sind. Wir erhalten daher einen Widerspruch zur Maximalität von (g, G), sofern $G \neq X$ ist. Also ist tatsächlich **G = X.**

Nun ist g additiv und homogen und erfüllt $0 \le g(x)\ \forall\ x \in M$. Wegen $g = g^o$ auf U ist g nicht identisch Null. Sei schließlich $y \in \operatorname{int} M$, und die Kugel $B(y, \varepsilon) = y + \varepsilon B$ enthalten in M. Dann gilt

$$0 \le g(y) + \varepsilon\, g(b) \quad \forall\ b \in B,$$

also ist g auf B durch $g(b) \ge -\varepsilon^{-1} g(y) =: C$ nach unten beschränkt. Als homogenes Funktional ist g dann auch nach oben durch |C| beschränkt und somit stetig in 0. Folglich (Additivität) ist g überall stetig, und $x^* = g$ liefert die Behauptung. ♦

3 Konvexe Hülle und Satz von Caratheodory

Für $X \subset R^n$ sei

$\operatorname{conv} X = \{ y \in R^n \ /\ \exists N\ \exists \lambda_i > 0\ (i = 0, 1, ,,,, N)\ \exists x^i \in X$ mit $\Sigma\ \lambda_i = 1$ und $y = \Sigma\ \lambda_i x^i \}$

die **konvexe Hülle** von X. Analog im linearen Raum über R.

Satz 1.4 (von Caratheodory):

Sei $X \subset R^n$, $X \neq \emptyset$. Dann gilt $h \in \operatorname{conv} X \Leftrightarrow \exists x^i \in M$, $i = 0, 1, ..., $ **$N \le n$** mit $h \in \operatorname{conv}\{x^o, ... x^N\}$.

Beweis:

Man nimmt *minimales* N, so dass $h \in \operatorname{conv}\{x^o, ... x^N\}$, $x^i \in M$. Dann ist

$$h = \Sigma\ \lambda_i x^i, \quad \Sigma\ \lambda_i = 1, \quad \lambda_i > 0.$$

Wenn $N > n$, so sind die Vektoren $x^i - x^o \in R^n$ (i = 1, ... ,N) linear abhängig, also gibt es (nichttriviale) r_i mit $0 = \Sigma\, r_i (x^i - x^o)$, d.h., $0 = \Sigma_{i>0}\ r_i x^i - (\Sigma_{i>0}\ r_i)\, x^o$.

Damit gibt es auch Zahlen α_i $(0 \le i \le N)$, so dass $\Sigma_i\ \alpha_i = 0$, $\alpha \neq 0$ und $\Sigma_i\ \alpha_i x^i = 0$.

Also gilt $h = \Sigma\ (\lambda_i + t\,\alpha_i)\, x^i$, $\Sigma\ (\lambda_i + t\,\alpha_i) = 1$ $(\alpha \neq 0)$.

Mit geeignetem t erhält man so eine Darstellung für h, die weniger positive $\lambda'_i = \lambda_i + t\,\alpha_i$ hat. Wid. ♦

Folgerung: Im R^n ist die konvexe Hülle einer kompakten Menge abgeschlossen.

Aber: Das stimmt nicht notw. für die konvexe Hülle einer kompakten Menge im Hilbert - Raum !
M sei der Nullpunkt vereinigt mit allen Elementen der Form
$x(k) = (0, 0, .., 1/k, 0, ...)$ in l^2, $k = 1, 2, ...,$ 1/k an Stelle k.
conv M enthält nicht das Element $y = \Sigma\ 2^{-k} x(k)$ aus der Abschließung von conv M .

4 Brouwer's Fixpunktsatz

Satz 1.5 Brouwer's Fixpunktsatz

Vor. $S \subset R^n$ konvex, kompakt, nichtleer , $f\colon S \to S$ stetig,

Beh. Es gibt ein $x^* \in S$, so dass $f(x^*) = x^*$ (Fixpunkt von f).

Wir beweisen den Satz zuerst für das Einheitssimplex $S = \{ x \in R^n \ /\ \Sigma\, x_i = 1, x_i \ge 0 \}$.
Die Erweiterung auf bel. konvexe, kompakte Mengen $S \neq \emptyset$ wird dann (fast) eine ÜA.
Zunächst beschäftigen wir uns mit ***Sperners Lemma.***
Man definiere eine Simplexunterteilung induktiv: *(Für n = 2, ein Intervall, kanonisch)*
1. einfache Unterteilung: über konvexe Hülle von einfach unterteilter Seite und Schwerpunkt.
2. mehrfache Unterteilung: Wiederholung der Prozedur auf Teilsimplizes.

Damit ist klar, was eine N-fache Unterteilung sein soll.
Knoten der Unterteilung seien die Ecken der Unterteilungssimplizes. Der *Durchmesser* (letzterer) geht mit der Ordnung der Unterteilung gegen Null (Induktion). Der Durchmesser einer Menge M (auch diam M) ist das Supremum der Abstände zweier ihrer Elemente.

Indexfunktion:

Eine Funktion $j = j(x) \in \{1,...,n\}$ für $x \in S$ heißt zulässige Indexfunktion, sofern gilt:

$$j(x) \in \{\pi 1, ...\ \pi k\} \text{ falls } x \in \operatorname{conv}\{e_{\pi 1}, ...\ e_{\pi k}\}.$$

Sperner's Lemma: *Sei $j(.)$ auf den Knoten (Ecken) einer N-fachen Unterteilung eine zulässige Indexfunktion. Dann gibt es ein Teilsimplex T der Unterteilung, so dass den Ecken p_1 , ,..., p_n von T alle Nummern 1, ...,n zugeordnet sind.*

Beweis:

Betrachte Teilsimplizes T(k) und definiere:

Ausgezeichnete Seite: Seite eines Teilsimplex' T(k) mit zugeordneten j-Werten 1,...,n-1.

"Normales" T(k): zugeordnete j-Werte sind 1,...,n.

Man zeigt per Induktion über die Dimension des Ausgangssimplex:

Es gibt stets eine ungerade Anzahl normaler Teilsimplizes.

Für dim = 0 (nur ein Punkt) trivial.

Idee: Sehen und gesehen werden !

1) sehen

Wir laufen durch alle T(k) und zählen die zu sehenden *ausgezeichneten Seiten* σ ; Anzahl sei t(k):

Wenn T(k) normal, so ist t(k) = 1. Sonst: t(k) = 0 oder t(k) = 2.

Also ist $h := \Sigma\, t(k)$ gerade $\Leftrightarrow$ Anzahl normaler T(k) ist gerade.

2) gesehen werden

Wir fragen nun, wie oft eine feste ausgezeichnete Seite σ gesehen wird.

Fall 1: σ in Original- Simplexseite der Knoten $e_1, \dots, e_{\pi\text{-}1}$: Genau einmal.

Fall 2: σ nicht in Original- Simplexseite der Knoten $e_1, \dots, e_{\pi\text{-}1}$:

Dann liegt sie wegen der Zulässigk.Bedingung auch in keiner anderen Seite des Originalsimplex S.

Sie wird deshalb von genau zwei Teilsimplizes T(k), T(k') aus gesehen (die sich an σ “spiegeln”).

Bezeichnet Σ_1 und Σ_2 die *Anzahl der ausgezeichneten Seiten* zu beiden Fällen (die in Original- Simplexseite der Knoten $e_1, \dots, e_{\pi\text{-}1}$ bzw. die restlichen) , so folgt $h = \Sigma_1 + 2\,\Sigma_2$.

Damit ist h gerade $\Leftrightarrow$ Σ_1 gerade.

Σ_1 ist aber die Anzahl normaler Teilsimplizes in einem Simplex kleinerer Dimension !

Daher ist Σ_1 ungerade, folglich auch h ungerade, was zu zeigen war. ♦

Die spezielle zulässige Indexfunktion zum Beweis des Brouwerschen Fixpunktsatzes:

Sei $S = \{ x \in R^n \,/\, \Sigma\, x_i = 1,\ x_i \geq 0 \}$ und $f\colon S \to S$ stetig.

Zu gegebenem x sei $I(x) = \{ i \mid x_i \geq f_i(x)\}$. Außerdem betrachten wir die (abgeschlossenen) Mengen

$X_i := \{ x \mid x_i \geq f_i(x) \}$.

Dann ist $e_i \in X_i$ (i-ter Einheitsvektor, eine Ecke von S). Weiter sieht man leicht wegen der Summenbedingung: x Fixpunkt von f $\Leftrightarrow$ $x \in \cap_i X_i$.

Außerdem gilt:

$$x \in \operatorname{conv}\{ e_1, \dots, e_k \} \Rightarrow \Sigma_{1\leq i \leq k}\, x_i = 1 \Rightarrow \Sigma_{1\leq i \leq k}\, f_i(x) \leq 1 \Rightarrow x \in \cup_{1\leq i \leq k} X_i .$$

Sonst ist die Summenbedingung für f(x) verletzt.

Also kann man jedem $x \in \operatorname{conv}\{ e_1, \dots, e_k \}$ ein $i \in \{ 1 \dots k\}$ zuordnen mit $x \in X_i$.

Das gilt analog für jede Auswahl $\{ e_{\pi 1}, \dots, e_{\pi k} \}$ von k Ecken e_i . Jedem x ordnen wir nun die *kleinste* Seite S(x) von S zu, in der x liegt. Das ist eine konvexe Hülle von Ecken $\operatorname{conv}\{ e_{\pi 1}, \dots, e_{\pi k} \}$ wobei k und π von x abhängen (mit minimalem k). Weiter wählen wir ein $i(x) \in \{\pi 1, \dots \pi k\}$ mit $x \in X_{\pi i(x)}$.

Dann ist diese Zuordnung i = i(x) eine spezielle zulässige Indexfunktion j.

Nach dem Lemma von Sperner gibt es zu jeder N-fachen Unterteilung von S ein normales Teilsimplex T(k). Für $x(N) \in T(k)$ folgt dann: $\operatorname{dist}(x(N), X_i) \leq \operatorname{diam} T(k) \to 0$ (für $N \to \infty$).

Nimmt man einen Häuf.-Punkt x* der x(N), erfüllt er damit $\operatorname{dist}(x^*, X_i) = 0$ und, da alle X_i abgeschlossen sind, $x^* \in \cap_i X_i$. Also ist der Brouwersche Fixpunktsatz für das Einheitssimplex S richtig. ♦

Behauptung für beliebige konvexe kompakte Mengen:

1. Zunächst erweitere man den Satz auf ein beliebiges Simplex $S = \operatorname{conv}\{a^1, \dots, a^n\}$ mittels

$$x \in S \Leftrightarrow x = \Sigma_k \lambda_k a^k , \quad \lambda \in \Lambda := \{ \lambda \in R^n \,/\, \Sigma_k \lambda_k = 1,\ \lambda_k \geq 0 \},$$

unter Berücksichtigung, dass die Zuordnung $\lambda = u(x)$ zwischen x und λ in beiden Richtungen eindeutig und stetig (linear) ist. Das folgt daraus, dass die Vektoren $a^2 - a^1, \dots, a^n - a^1$ (nach Definition eines Simplex) linear unabhängig sind.

Damit gilt:

x ist Fixpunkt von $f: S \to S \quad \Leftrightarrow \quad \lambda = u(x)$ ist Fixpunkt von $g = u \circ f \circ u^{-1}: \Lambda \to \Lambda$.
Die Anwendung des schon bewiesenen Teils auf das Einheitssimplex Λ liefert die Behauptung für S.

2. Für konvexe kompakte Mengen $S \neq \emptyset$ wähle man ein Simplex Σ, das S enthält.
Sei $\pi(x) = \mathrm{proj}_S(x)$ die Projektion von x auf S. Sie ist stetig. Durch $g(x) = f(\pi(x))$ für $x \in \Sigma$ läßt sich dann f stetig auf Σ erweitern; $g(x) = f(x)$ für $x \in S$, $g: \Sigma \to S \subset \Sigma$.
Da g und f dieselben Fixpunkte besitzen, folgt so die Behauptung für S. ♦

Brouwer's Fixpunktsatz und "Arbeitslemma" :

Arbeitslemma:
Sei $S = \{x \in R^n \mid \Sigma x_i = 1, \; x_i \geq 0 \; \forall i\}$ und $F_i = \{x \in S \mid x_i = 0\}$ (die i-te Seite von S)
Seien $X_i \subset S$ abgeschlossene Mengen mit $F_i \subset X_i \; \forall i$ und $\cup_i X_i = S$. Beh. $\cap_i X_i \neq \emptyset$.

Wieso dieser Name ? Man deute x_i als Arbeitspensum, das Person i zu erledigen hat; x ist eine Aufteilung der gesamten zu erledigenden Arbeit. X_i sei die Menge der Aufteilungen x, mit welchen i einverstanden ist. Die Bedingung $\cup_i X_i = S$ sagt, dass stets wenigstens *einer* einverstanden sein soll. Die Behauptung liefert, dass bei geschickter Aufteilung dann sogar *alle* einverstanden sind.

Beweis (mittels Brouwer's FPSatz):
Indirekt. Angenommen $\cap_i X_i = \emptyset$. Sei $d_i(x) = \mathrm{dist}(x, X_i)$; $x \in S$.
Dann ist d_i stetig und $s(x) := \Sigma_i d_i(x) > 0$ weil stets wenigstens ein $d_i(x) > 0$ positiv ist. Die Funktion f(x) mit Komponenten

$$f_i(x) = d_i(x) / s(x)$$

bildet S in sich ab. Weil $\cup_i X_i = S$ gilt, liegt x in wenigstens einem X_i, $(i = i(x))$, also ist stets ein $f_i(x)$ Null. Daher bildet f sogar auf den Rand $\mathrm{bd}\, S = \cup F_i$ von S ab.
Auch folgt $f(x) \in F_i$, falls $x \in F_i \subset X_i$.
Wir betrachten nun zu f(x) den am Zentrum $z = (1/n, \ldots, 1/n)$ des Simplex S gespiegelten Randpunkt g(x). Er ist beschrieben durch

$$z = \lambda\, g(x) + (1-\lambda)\, f(x), \quad 0 < \lambda < 1, \quad g(x) \in \mathrm{bd}\, S .$$

Wegen $f(x) \in \mathrm{bd}\, S$ sind dabei g(x) und λ eindeutig und stetig definiert.
Also gilt $g: S \to \mathrm{bd}\, S$ und wegen der Spiegelung für jedes $x \in F_i$ auch $f(x) \in F_i$ und $g(x) \notin F_i$. Für $x \in \mathrm{bd}\, S$ ist deshalb $g(x) \neq x$. Die Punkte aus int S können ebenfalls keine Fixpunkte von g sein, weil sie auf den Rand abgebildet werden. Also wäre g eine Gegenbeispiel zu Brouwer's FPSatz. ♦

Umkehrung:
Aus dem Arbeitslemma folgt Brouwer's FPSatz (zunächst für das angegebene Simplex)
Angenommen $f: S \to S$ ist stetig. Wir definieren die abgeschl. Mengen $X_i = \{x \in S \mid f_i(x) \geq x_i\}$.
Die Ungleichung ist hier im Vergleich zum Beweis mit Sperners Lemma gerade umgekehrt.
Hiermit folgt jetzt $F_i \subset X_i$. Weiter sieht man: Gäbe es ein $x \in S \setminus \cup_i X_i$, würde folgen $f_i(x) < x_i \; \forall i$, also auch $\Sigma f_i(x) < 1$ im Gegensatz zu $f(x) \in S$. Also gilt $\cup_i X_i = S$. Damit sind die Voraussetzungen des Arbeitslemmas erfüllt. Mit $x \in \cap_i X_i$ erhalten wir nun $f_i(x) \geq x_i \; \forall i$ und wegen der Summenbedingung auch die Behauptung $f_i(x) = x_i \; \forall i$. ♦

Brouwer's Fixpunktsatz und Retrakte:
Eine Teilmenge $C \subset X$ (im metr. Raum) heisst Retrakt von X, wenn es eine stetige Abbildung $f: X \to C$ gibt mit $f(c) = c \;\; \forall c \in C$. Sei X = B die abgeschl. Einheitskugel im R^n und C = bd B ihr Rand.
Behauptung (unser Retraktsatz): Der Rand bd B ist kein Retrakt von B.
Angenommen doch: Dann kann man $g(x) = -f(x)$ setzen, wonach $g: B \to B$ keinen Fixpunkt hätte. Also: Brouwer's Satz beweist die Behauptung.
Umgekehrt: Die Retrakt-Behauptung sei schon bewiesen. Um Brouwer's Satz (für die Einheitskugel) zu beweisen, nimmt man an, dass ein stetiges $g: B \to B$ ohne Fixpunkt existiert. Dann gibt es ein $c > 0$ so dass $d(x, g(x)) \geq c \;\forall\; x \in B$. Zu festem x betrachte man den Strahl, der in $g(x) \in B$ beginnt und durch x verläuft. Er schneidet (in Vorwärtsrichtung) den Rand von B in genau einem Punkt f(x). Wegen $d(x, g(x)) \geq c$ ist $f = f(x)$ stetig auf B. Wenn $x \in \mathrm{bd}\, B$, so folgt nach Konstruktion offenbar $f(x) = x$. Also würde f zeigen, dass C = bd B Retrakt von B ist. Damit muss Brouwer's Satz für die

Kugel richtig sein. Hiermit erweitert man ihn wie im Falle von Simplizes auf beliebige konvexe, kompakte Mengen. ♦

5 Kakutani's Fixpunktsatz

Satz 1.6

Vor. Sei S konvex, kompakt, nichtleer im R^n, $F: S \rightrightarrows S$ eine (mehrwertige) Abbildung mit
$\emptyset \neq F(x) \subset S$, $F(x)$ konvex, kompakt, gph $F := \{ (x, y) / x \in S,\ y \in F(x) \}$ abgeschlossen.

Beh. Es gibt ein $x^* \in S$, so dass $x^* \in F(x^*)$ (Fixpunkt von F).

Bemerkungen

Die Symbolik $F: S \rightrightarrows S$ ist üblich um zu zeigen, dass die Bilder Mengen sind; man könnte auch schreiben $F: S \to P(S)$ (Potenzmenge).
Die Menge gph F heißt Graph von F. Man nennt eine mehrdeutige Abbildung F abgeschlossen, wenn gph F abgeschlossen im Produktraum ist.
Sind die Bildmengen alle einelementig, kann man F als Funktion verstehen. Die Abgeschlossenheit von F bedeutet dann (weil in eine kompakte Menge abgebildet wird) gerade Stetigkeit.

Beweis:

Diesmal nehmen wir an, $S = B$ sei die Euklidische Einheitskugel des R^n, und es gelte

$x \notin F(x)\ \forall\, x \in S$.

Wir betrachten nur Punkte x aus S. Dann gibt es (Trennungssatz), zu jedem x ein $z \in R^n$ als Trennungsvektor für x und F(x); also ein z mit

$t(z, x) := \inf \{ \langle z,\ y - x \rangle /\ y \in F(x) \} > 0.$

Also ist $\Omega(x) := \{ z /\ t(z, x) > 0 \}$ nie leer, wonach

(1.7) $S \subset \cup_z\ \Omega^{-1}(z)$ mit $z \in R^n$ und $\Omega^{-1}(z) = \{ x /\ t(z, x) > 0 \}$

gilt.

Wir zeigen

(1.8) $\Omega^{-1}(z)$ ist (relativ) offen in S.

Andernfalls existieren z und x mit $z \in \Omega(x)$ sowie eine Folge $x^k \to x$, so dass $x^k \in S \setminus \Omega^{-1}(z)$, d.h.,

$\inf \{ \langle z,\ y - x^k \rangle /\ y \in F(x^k) \} \le 0.$

Da $F(x^k)$ kompakt ist, gibt es weiter $y^k \in F(x^k)$ mit $\langle z,\ y^k - x^k \rangle \le 0$. Nach Auswahl einer geeigneten Teilfolge darf man annehmen, dass die Folge der y^k (enthalten in der kompakten Menge S) konvergiert, $y^k \to y$. Wegen $(x^k, y^k) \in$ gph F und der Abgeschlossenheit von gph F, ist auch $(x, y) \in$ gph F. Wir erhalten so den Widerspruch $\langle z,\ y - x \rangle \le 0$ zu $z \in \Omega(x)$, also gilt (1.8).

Folglich gibt es endlich viele (Überdeckungssatz) Elemente z^p ($p = 1, \dots , N$) mit

(1.9) $S \subset \cup_p\ \Omega^{-1}(z^p)$.

Sei $A_p = S \setminus \Omega^{-1}(z^p)$ und $d_p(x) = \operatorname{dist}(x, A_p)$. Wegen (1.9) und der Abgeschlossenheit aller A_p gilt nun $d(x) := \Sigma_p\, d_p(x) > 0$ für alle x. Damit können wir stetige Funktionen

$\lambda_p(x) = d_p(x) / d(x)$ und $z(x) = \Sigma_p\, \lambda_p(x)\, z^p$ definieren.

Bemerkung:

Stetige Funktionen λ_p , die in dieser (und allgemeinerer) Weise einer offenen Überdeckung von S (hier (1.9)) *zugeordnet sind, heißen Zerlegung der 1 (Zerlegung der Einheit), wenn* $\Sigma_p\, \lambda_p(x) = 1$ und $\lambda_p(x) > 0 \Rightarrow x \in \Omega^{-1}(z^p)$.

Um zu sehen, dass $z(x) \in \Omega(x)$, betrachten wir irgendein $y \in F(x)$, $x \in S$. Dann gibt es wenigstens ein positives $\lambda_p(x)$, und aus $\lambda_p(x) > 0$ folgt $x \in \Omega^{-1}(z^p)$ und $\lambda_p(x) \langle z^p,\ y - x \rangle > 0$.
Also ist (wegen $\Sigma_p\, \lambda_p(x) = 1$) auch $\langle z(x),\ y - x \rangle = \Sigma_p\, \lambda_p(x) \langle z^p,\ y - x \rangle > 0$.
Da $y \in F(x)$ beliebig gewählt war, folgt $z(x) \in \Omega(x)$. Insbesondere ist dann $z(x) \neq 0$.
Die Funktion

$f(x) = z(x) / \| z(x)\|$

bildet daher die Kugel S stetig auf ihren Rand ab und erfüllt offenbar ebenfalls $f(x) \in \Omega(x)$. Nach dem Satz von Brouwer besitzt f einen Fixpunkt: $x^* = f(x^*) \in \Omega(x^*)$.
Er erfüllt $\| x^* \| = \langle x^*,\ x^* \rangle = 1$ und $0 < \inf \{ \langle x^*,\ y - x^* \rangle /\ y \in F(x^*) \}$. Letzteres bedeutet

$\langle x^*,\ y \rangle > 1 \quad \forall\ \ y \in F(x^*)$

und liefert $F(x^*) \cap B = \emptyset$, im Gegensatz zu $\emptyset \neq F(x^*) \subset S$. Dies beweist den Satz für die Einheitskugel S.
Wie im Falle des Brouwerschen Fixpunktsatzes stellen wir nun fest, dass der Satz auch für jede Kugel $S = r\,B$ ($r > 0$) gilt. Ist schließlich S konvex und kompakt, wählen wir r so groß, dass $S \subset r\,B$ und beweisen den Satz durch Betrachtung von $G(x) = F(\pi(x))$ auf $r\,B$ mit $\pi(x)$ = Proj. von x auf S. ♦

6 Ekeland's variational principle (1974)

Satz 1.7
Let X be a complete metric space and $\phi: X \to R \cup \{\infty\}$ be a l.s.c. (= unterhalb stetig) functional with a finite infimum. Let ε and α be positive, and let x satisfy $\phi(x) \leq \varepsilon + \inf_X \phi$.
Then, there is some $z \in X$ such that

$$d(z, x) \leq \alpha, \quad \phi(z) \leq \phi(x) \quad \text{and} \quad \phi(\xi) + (\varepsilon / \alpha)\, d(\xi, z) \geq \phi(z) \quad \forall \xi \in X .$$

Bemerkungen
1 Man erhält zwar keinen Minimialpunkt von ϕ, aber einen globalen Minimalpunkt (nämlich z) der veränderten Funktion $f_{\alpha\,\varepsilon\,z}(\xi) = \phi(\xi) + (\varepsilon / \alpha)\, d(\xi, z)$, die man so spezialisieren kann, dass sie sich „wenig“ von ϕ unterscheidet.
2 ϕ heißt *unterhalb stetig*, wenn alle unteren Niveaumengen $\{ \xi \in X \mid \phi(\xi) \leq c \}$ $(c \in R)$ abgeschlossen sind; äquivalent dazu: $\liminf_{\xi' \to \xi} \phi(\xi') \geq \phi(\xi)\ \forall \xi$.
3 Die Punkte mit $\phi(\xi) = \infty$ spielen im Beweis keine Rolle. Durch die Voraussetzung läßt man nur Funktionen zu wie $\phi(\xi) = |\xi|^{-1}$ für $\xi \neq 0$ und $\phi(0) = \infty$ auf dem vollst. metr. Raum $X = R$.
4 Man kann jedoch auch einfach, bei gegebenem x, zum ebenfalls vollständigen Raum

$$X' = \{ \xi \in X \mid \phi(\xi) \leq \phi(x) \} \qquad \text{übergehen.}$$

Proof:
Put $h(z) = \inf_{\xi \in X} \phi(\xi) + (\varepsilon / \alpha)\, d(\xi, z)$.
For arbitrary ξ, z and z' in X, we observe

$$\phi(\xi) + (\varepsilon / \alpha)\, d(\xi, z) \leq \phi(\xi) + (\varepsilon /\alpha) [\, d(\xi, z') + d(z', z) \,].$$

Taking the infimum over all $\xi \in X$ on both sides, we obtain $h(z) \leq h(z') + (\varepsilon /\alpha)\, d(z', z)$.
Therefore, h is a Lipschitz function (since z and z' may change the place).
To construct a sequence z^k we set $z^o = x$.
If $h(z^o) \geq \phi(z^o)$ then $z = z^o$ satisfies the requirements of the theorem.
Thus, beginning with k = 0, assume that $h(z^k) < \phi(z^k)$. Then one finds some z^{k+1} such that

$$\textbf{(1.10)} \qquad \phi(z^{k+1}) + (\varepsilon / \alpha)\, d(z^{k+1}, z^k) < \phi(z^k)$$

and, in addition,

$$\textbf{(1.11)} \qquad \phi(z^{k+1}) + (\varepsilon / \alpha)\, d(z^{k+1}, z^k) < h(z^k) + 2^{-k}.$$

From (1.10) and $\phi(z^o) \leq \inf_X \phi + \varepsilon$, we obtain for each k,

$$(\varepsilon /\alpha)\ \Sigma_{s \leq k}\, d(z^{s+1}, z^s) < \Sigma_{s \leq k} (\phi(z^s) - \phi(z^{s+1})) = \phi(z^o) - \phi(z^{k+1}) \leq \varepsilon.$$

This yields particularly

$$\textbf{(1.12)} \qquad d(z^{k+1}, z^o) \leq \Sigma_{s \leq k}\, d(z^{s+1}, z^s) \leq \alpha .$$

By (1.10) and (1.12), $z = z^k$ is the point in question whenever $\phi(z^k) \leq h(z^k)$.
Otherwise (1.12) shows that the Cauchy sequence z^k has a limit z^* in the complete space X.
Clearly, $d(z^*, z^o) \leq \alpha$. Since ϕ is l.s.c., it follows by (1.10),

$$\phi(z^*) \leq \liminf \phi(z^k) \leq \phi(z^o).$$

Finally, recalling that h is continuous, we infer due to (1.11),

$$\phi(z^*) \leq \liminf [\, \phi(z^{k+1}) + (\varepsilon /\alpha)\, d(z^{k+1}, z^k) \,] \leq \limsup [\, h(z^k) + 2^{-k} \,] \leq h(z^*). \qquad ♦$$

Die häufigste Anwendung erfolgt mit $\alpha = \sqrt{\varepsilon}$.
Für eine differenzierbare Funktion $\phi : X = R^n \to R$ folgt so z.B., dass in z nicht nur $\phi(z) - \inf_X \phi$ klein ist, sondern auch die Norm des Gradienten $\| D\phi(z) \| \leq \sqrt{\varepsilon}$ klein sein muss.

Denn wir haben mit $\xi = z + u$ und $\phi(z+u) = \phi(z) + D\phi(z)\,u + o(u)$,

$$\phi(z) + D\phi(z)\,u + o(u) + \sqrt{\varepsilon}\,\|u\| \geq \phi(z) \qquad \forall\, u \in R^n,$$

also
$$D\phi(z)\,u + o(u) + \sqrt{\varepsilon}\,\|u\| \geq 0 \qquad \forall\, u \in R^n.$$

Nimmt man $u = -t\, D\phi(z) / \|D\phi(z)\|$ (im Falle $D\phi(z) \neq 0$) und kleine $t > 0$, folgt so

$$-t\,\|D\phi(z)\| + o(u) + \sqrt{\varepsilon}\; t \geq 0,$$

was nach Division durch t mit $t \downarrow 0$ und $o(u) / t \to 0$ die Behauptung $\|D\phi(z)\| \leq \sqrt{\varepsilon}$ liefert.

Wie folgt **Banach's** Fixpunktsatz aus Ekeland's Principle ?
Für $f\colon X \to X$ (kontraktiv mit Konstante $k < 1$) definiere man $\phi(x) = d(x, f(x))$. Anwendung von Ekeland mit $\alpha = \sqrt{\varepsilon}$ liefert (mit beliebig kleinem $\varepsilon > 0$) die Existenz von $z = z(\varepsilon)$, so dass

$$\phi(\xi) + \sqrt{\varepsilon}\, d(\xi, z) \geq \phi(z) \quad \forall\, \xi \in X.$$

Nun setze man $\xi = f(z)$. Dann folgt wegen $k\, d(z, f(z)) \geq d(f(z), f(f(z)))$ (aus Kontraktivität)

$$k\, d(z, f(z)) + \sqrt{\varepsilon}\, d(f(z), z) \geq d(f(z), f(f(z))) + \sqrt{\varepsilon}\, d(f(z), z)$$
$$= \phi(f(z)) + \sqrt{\varepsilon}\, d(f(z), z) \geq \phi(z) = d(z, f(z)).$$

Also gilt

$$\sqrt{\varepsilon}\, d(f(z), z) \geq (1-k)\, d(z, f(z)).$$

Da $1 - k > 0$, ist dies mit $\sqrt{\varepsilon} < 1 - k$ nur dann richtig, wenn $d(z, f(z)) = 0$.
Die Einzigkeit des Fixpunktes ist trivial: $d(z, z') = d(f(z), f(z')) \leq k\, d(z, z') \Rightarrow z = z'$.

7 Minimaxsatz im R^n und Nash-Gleichgewicht (über Kakutani)

Eine Funktion $g\colon M \to R$, definiert auf einer konvexen Menge M, heisst konkav, wenn $-g$ auf M konvex ist. g heisst konvex, falls $g(\lambda u + (1-\lambda)v) \leq \lambda\, g(u) + (1-\lambda) g(v)$ $\forall\, u, v \in M$, $\lambda \in (0, 1)$.
strikt konvex = konvex und für $u \neq v$, $\lambda \in (0, 1)$ gilt nie die Gleichheit.
Wichtigste Folgerungen aus Konvexität:
Jeder lokale Minimalpunkt von f bzgl. M ist auch ein globaler.
Die Menge der Minimalpunkte ist konvex, für strikt konvexe Funktionen höchstens einelementig.

Minimax Satz
Seien X und Y nichtleere, konvexe, kompakte Teilmengen des R^n und $f\colon (X,Y) \to R$ stetig.
Sei weiter $f(\,.\,, y)\colon X \to R$ konvex $\forall\, y \in Y$ und $f(x, \,.\,)\colon Y \to R$ konkav $\forall\, x \in X$.
Dann existiert ein Punkt $(u, v) \in (X, Y)$ mit

(1) $\quad f(u, y) \leq f(u, v) \leq f(x, v) \quad \forall\, (x, y) \in (X, Y)$.

Beweis 1 (über Kakutani's Fixpunktsatz)
Seien stets $(x, y) \in (X, Y)$, $(u, v) \in (X, Y)$. Für beliebige (x, y) sei

$$U(y) = \{\, u \,/\, f(u, y) \leq f(\xi, y) \quad \forall\, \xi \in X \,\}.$$

Das ist die Menge der Minimalpunkte von $f(\,.\,, y)$ bezüglich X; wir nennen sie auch $\arg\min_X f(\,.\,, y)$.
Analog bilden wir

$$V(x) = \{\, v \,/\, f(x, v) \geq f(x, \upsilon) \quad \forall\, \upsilon \in Y \,\},$$

die Menge der Maximalpunkte von $f(x, \,.\,)$ bezüglich Y; $\arg\max_Y f(x, \,.\,)$. Forderung (1) bedeutet nichts anderes als

$$u \in U(v) \text{ und } v \in V(u).$$

Definiert man eine Abbildung $F(x, y) = (U(y), V(x)) := \{\, (u, v) \,/\, u \in U(y),\ v \in V(x) \,\}$, heisst das auch $(u, v) \in F(u, v)$.
Es ist also zu zeigen, dass F als mehrdeutige Abbildung von (X, Y) in sich einen Fixpunkt besitzt. Wir prüfen die Voraussetzungen des Kakutani Satzes.
U(y) als Menge der Minimialpunkte von $f(\,.\,, y)$ ist konvex (kann man direkt aus der Definition von konvex zeigen) und nichtleer (da X kompakt und f stetig). Dasselbe folgt analog für V(x).
Damit ist auch F(x, y) nichtleer und konvex im (R^n, R^n). Wir brauchen noch, dass graph F abgschlossen ist, d.h.: aus $(x^k, y^k) \to (x, y)$, $(u^k, v^k) \to (u, v)$ und $(u^k, v^k) \in F(x^k, y^k)$ muss folgen dass $(u, v) \in F(x, y)$.
Aber $(u^k, v^k) \in F(x^k, y^k)$ bedeutet

$$f(u^k, y^k) \leq f(\xi, y^k) \quad \forall\, \xi \in X, \qquad f(x^k, v^k) \geq f(x^k, \upsilon) \quad \forall\, \upsilon \in Y,$$

und zu zeigen ist

$f(u , y) \leq f(\xi , y) \qquad \forall \xi \in X , \qquad f(x , v) \geq f(x, \upsilon) \qquad \forall \upsilon \in Y .$

Wäre im Gegenteil etwa

$f(u , y) > f(\xi^o , y)$ für ein $\xi^o \in X$, so würden wir aus Stetigkeitsgründen

auch $f(u^k , y^k) > f(\xi^o , y^k)$ erhalten, im Widerspruch zu $(u^k, v^k) \in F(x^k, y^k)$. Ganz analog muss auch

$f(x , v) \geq f(x, \upsilon) \qquad \forall \upsilon \in Y$

richtig sein. Also ist graph F abgschlossen. Damit liefert Kakutanis Satz die Behauptung. ♦

Nash-Gleichgewicht

Analog zu oben kann man auch jetzt vorgehen: In einem *n-Personenspiel* mögen die Spieler konvexe, kompakte *Strategiemengen* $X_i \subset R^m$ (i = 1,..,n) besitzen (nicht leer) und bei Wahl der konkreten Strategien $x_i \in X_i$ jeweils den Gewinn $f_i(x_1, \dots , x_n)$ erhalten. Kooperation bei der Strategiewahl sei verboten.

Dann muss Spieler i mit der Situation $x = (x_1, \dots , x_n) \in (X_1 , \dots , X_n)$ zufrieden sein (er kann jedenfalls als einziger nicht vorteilhaft abweichen), wenn gilt

$(*)_i \qquad f_i (x_1, .., x_{i-1} , \xi_i , x_{i+1} , \dots , x_n) \leq f_i (x_1, \dots , x_n) \qquad \forall \xi_i \in X_i$ (analog für i = 1, i = n).

Für die linke Seite schreibt man auch $f_i (x // \xi_i)$. Erfüllt ein $x \in (X_1 , \dots , X_n)$ die Bedingung $(*)_i$ für alle i, so heisst x Gleichgewichtssituation (auch Nash-Gleichgewicht (GGW) oder Gleichgewichtspunkt.

Somit ist ein Sattelpunkt ein Nash-Gleichgewicht, wenn n = 2 und $f_1 = - f$, $f_2 = f$ gilt. Die Existenz eines Nash-GGW kann unter analogen Bedingungen mittels Beweis 1 gezeigt werden.

Satz

Seien alle f_i stetig und $f_i (x // \xi_i)$ in ξ_i konkav. Dann gibt es wenigstens ein Nash-GGW.

Beweis:

Man ordne jedem $x \in (X_1 , \dots , X_n)$ die Mengen der Maximalpunkte

$F_i (x) = \text{argmax} \{ f_i (x // \xi_i) \mid \xi_i \in X_i \} \subset X_i$

zu. Auf die Abbildung $F(x) = (F_1 (x) , \dots , F_n (x))$ mit konvexen kompakten Bildmengen im Produktraum kann man (wieder mit denselben Argumenten) den Kakutani-Satz anwenden. Jeder Fixpunkt x ist dann ein Nash-GGW. ♦

8 Minimaxsatz (über vollst. Induktion)

Der Minimaxsatz muss nicht über Kakutanis Satz bewiesen werden, man kann auch vollständige Induktion über die Dimension des Raumes nutzen. Man zeigt leicht

Sind (u’, v’) und (u’’, v’’) Sattelpunkte , so auch (u’, v’’) und (u’’, v’).

Wir bilden nun $f_\varepsilon (x, y) = f(x, y) + \varepsilon \|x\|^2 - \varepsilon \|y\|^2$ mit Euklidischer Norm und halten zunächst $\varepsilon > 0$ fest. Dann wird f_ε strikt konvex-konkav (d.h. die Konvexitäts/ Konkavitäts-ungleichung gilt streng für $\lambda \in (0, 1)$ und $x_1 \neq x_2$ bzw. $y_1 \neq y_2$) und hat deshalb eindeutige Minima (Maxima) bzgl x (bzw. y).

Auch Sattelpunkte (u, v) , sofern existent, sind dann eindeutig, was wir in der Folge ausnutzen. Wir beweisen den Satz zuerst für f_ε .

Der Satz ist trivial für dim X + dim Y = 0 (dann sind X und Y einelementig).

Er sei schon richtig für dim X + dim Y = k, wir betrachten den Fall dim X + dim Y = k+1.

Dann enthält eine beider Mengen, o.B.d.A. X , mindestens 2 Punkte a, b (damit auch eine Strecke).

Sei $c = b – a$ und H(t) die affine Hyperebene $H(t) = \{ x / c^T x = t \}$; ($c^T x$ = Skalarprodukt)

Sei $X(t) = X \cap H(t)$. Man sieht leicht, dass $\dim X(t) < \dim X$ und dass

$X(t) \neq \emptyset$ genau dann gilt, wenn t in einem gewissen kompakten Interval $[t_1 , t_2]$ liegt.

Auf die Funktion f_ε und die Mengen X(t), Y (für $t \in [t_1 , t_2]$) können wir den Satz anwenden.

Damit gibt es einen eindeutigen Sattelpunkt (x(t), y(t)) für f_ε und X(t), Y.

Wenn ein (x(t), y(t)) Sattelpunkt für f_ε und X, Y ist, sind wir fertig.

Andenfalls muss es zu jedem $t \in [t_1 , t_2]$ ein $\xi(t) \in X$ geben mit $f_\varepsilon (\xi(t), y(t)) < f(x(t), y(t))$. Inbesondere können wir dasjenige eindeutige $\xi(t) \in X$ wählen, das die linke Seite bzgl. ξ minimiert.

Da (x(t), y(t)) schon Sattelpunkt für das reduzierte Problem war, gilt $\xi(t) \in X \setminus H(t)$.

Man überlegt sich als nächstes (am einfachsten indirekt),
dass sowohl die eindeutigen Sattelpunkte (x(t), y(t)) als auch $\xi(t)$ stetig von t abhängen müssen. Ferner gilt nach Konstruktion von H(t),

$c^T \xi(t) > c^T x(t) = t$ für $t = t_1$.

$c^T \xi(t) < c^T x(t) = t$ für $t = t_2$.

Also hat die stetige Funktion $h(t) = c^T \xi(t) - t$ eine Nullstelle t_o im Intervall $[t_1 , t_2]$. Dort gilt aber $\xi(t_o) \in H(t_o)$, was schon ausgeschlossen war. Damit besitzt f_ε wie behauptet einen Sattelpunkt in Bezug auf X und Y. Sei nun (x^k , y^k) ein Sattelpunkt zu $\varepsilon = 1/k$. Wegen Kompaktheit von X und Y besitzt diese Folge einen Häufungspunkt (u, v) , von dem man leicht (mittels Stetigkeit von f) zeigt, dass er ein Sattelpunkt für f ist. ♦

9 Dualität in konvexer Optimierung (als Anwendung des Minimax Satzes)

Wir betrachten hier die Aufgabe

(P) min $f(x)$, $x \in M$; $M = \{ x \in R^n \mid g_i (x) \leq 0 \;\; \forall\, i = 1,\dots ,m \}$

mit *konvexen* Funktionen $f, g_i : R^n \to R$ (eine klassische Aufgabe der konvexen Optimierung).
Die Funktion

$L(x, y) = f(x) + \Sigma\, y_i\, g_i(x)$

heisst ***Lagrange Funktion*** dieser Aufgabe. Ihre Sattelpunkte (x^*, y^*) mit $y^* \geq 0$ und

(D0) $L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall\, x \in R^n ,\; y \in R^n , y \geq 0$

spielen eine entscheidende Rolle für Aufgabe (P).

Bemerkung:
Sind f, g_i differenzierbar, so folgt aus **(D0)**, da x^* die Funktion $L(. , y^*)$ minimiert, die Bedingung

(*) $0 = D_x L(x^*, y^*) = Df(x^*) + \Sigma\, y^*_i\, Dg_i(x^*)$,

was eine direkte Verallgemeinerung der notwendigen Bedingung $Df(x^*) = 0$ für einen (lokalen) Minimalpunkt x^* von f darstellt. Wir verzichten hier auf die Differenzierbarkeit, notieren aber
ohne Beweis: Konvexe Funktionen auf dem R^n sind überall stetig.
In linearen Räumen unendlicher Dimension müsste man Stetigkeit extra voraussetzen.
Zusatzvoraussetzungen: Gelte

(i) $\lim f(x) = \infty$ für $\|x\| \to \infty$ und

(ii) $g_i(x^o) < 0 \;\; \forall i$ mit einem gewissen $x^o \in R^n$ (x^o ist ein sogenannter *Slater*punkt).

(Dualitäts-) Satz
Unter diesen Voraussetzungen hat (P) eine Lösung x^*, und für jede Lösung x^* existiert ein $y^* \geq 0$, so dass (x^*, y^*) die Sattelpunktbedingung (D0) erfüllt.

Bemerkung:
Der Punkt y^* maximiert die ("duale") Funktion $h(y) = \inf_x \; L(x, y)$ bzgl. $y \geq 0$,
während x^* die Funktion $s(x) = \sup_{y \geq 0} L(x, y)$
über R^n minimiert mit $s(x^*) = f(x^*) = h(y^*)$.
Man sieht sofort: $s(x) = \infty \Leftrightarrow \exists\, i$ mit $g_i(x) > 0$ (d.h. x unzulässig für (P)). Andernfalls: $s(x) = f(x)$.
Beweis des Satzes
Existenz von x*:
Die Menge $M_o := \{ x \,/\, f(x) \leq f(x^o) \} \cap M$ ist abgeschlossen und nicht leer. Sie ist ausserdem beschränkt wegen Voraussetzung (i). Damit existiert ein Minimalpunkt x^* von f bzgl. M_o . Er löst offenbar (P).
Existenz von y*:
Sei $r > 0$ hinreichend gross, so dass $\|x^*\| < r$. Wir betrachten das Sattelpunktproblem (D0) für L in Bezug auf die Mengen $X_r = \{ x \,/\, \|x\| \leq r \}$ und $Y_r = \{ y \geq 0 \,/\, \|y\| \leq r \}$ mit Maximum-Norm. Nach dem Minimax-Satz gibt es einen Punkt $(x'_r, y'_r) \in (X_r ,\; Y_r)$ mit

$L(x'_r, y) \leq L(x'_r, y'_r) \leq L(x, y'_r) \quad \forall\, x \in X_r ,\; y \in Y_r$,

und explizit

(D1) $f(x'_r) + \langle y, g(x'_r) \rangle \le f(x'_r) + \langle y'_r, g(x'_r) \rangle \le f(x) + \langle y'_r, g(x) \rangle \quad \forall x \in X_r, \quad y \in Y_r$.

Rechts können wir **$x = x^*$** setzen um zu sehen, dass wegen $g(x^*) \le 0$ und $y'_r \ge 0$ gilt:
(D2) $f(x'_r) + \langle y'_r, g(x'_r) \rangle \le f(x^*) + \langle y'_r, g(x^*) \rangle \le f(x^*)$.
Mit **$x = x^o$ und $y = 0$** folgt ausserdem, wenn wir o.B.d.A. $g_i(x^o) < -\varepsilon \ \forall i$ annehemen:
(D3) $f(x'_r) \le f(x'_r) + \langle y'_r, g(x'_r) \rangle \le f(x^o) + \langle y'_r, g(x^o) \rangle \le f(x^o) - \varepsilon \ \Sigma_i \ y'_{r\,i}$.

Links setzen wir **$y = 0$** um zu sehen, dass

$$f(x'_r) \le f(x'_r) + \langle y'_r, g(x'_r) \rangle \le f(x^*).$$

Wegen (i) liegen daher alle x'_r in der *kompakten* Menge $\{ x / f(x) \le f(x^*) \}$, bleiben also beschränkt für grosse r und besitzen einen Häufungspunkt für $r \to \infty$.
Damit kann wegen (D3) auch die Norm von $y'_r \ge 0$ nicht divergieren. Also $\|x'_r\| + \|y'_r\| < C < \infty \ \forall r$.
Wir fixieren nun $r > C$.
Dann muss $(x', y') = (x'_r, y'_r)$ die Sattelpunktbedingung (D0) erfüllen.
Wäre nämlich
$L(\xi, y') < L(x', y')$ für irgendein ξ, so folgte aus der Konvexität von L in x für alle $\lambda \in (0, 1)$

$$L(\lambda \xi + (1-\lambda) x', \ y') \le \lambda L(\xi, y') + (1-\lambda) L(x', y') < L(x', y').$$

Der Punkt $x'' = \lambda \xi + (1-\lambda) x'$ liegt dann wegen $x' \in X_r$ für kleine $\lambda > 0$ ebenfalls in X_r und verletzt die Bedingung (D2). Analog zeigt man, da L in y linear ist, dass die rechte Ungleichung in (D0) gilt.
Wir zeigen jetzt, dass x' ebenfalls Aufgabe (P) löst: Die linke Ungleichung von (D0),

$$\langle y, g(x') \rangle \le \langle y', g(x') \rangle \quad \forall y \ge 0,$$

sagt uns, dass nur $g(x') \le 0$ und $\langle y', g(x') \rangle = 0$ gelten kann. Also ist x' zulässig und erfüllt wegen (D2) auch $f(x') \le f(x^*)$. Der Punkt x' ist also ebenfalls optimal für (P).
Der Satz ist daher bewiesen, *sofern* (P) nur eine Lösung x^* ($= x'$) besitzt.
Andernfalls
ersetzte man die Zielfunktion f durch $f_\delta(x) = f(x) + \delta \| x - x^* \|^2$ (Euklidische Norm, $\delta > 0$). Dann ist x^* die *eindeutige* Lösung, man erhält einen entsprechenden Sattelpunkt $(x^*, y^*(\delta))$,und findet das gesuchte y^* als Häufungspunkt von $y^*(\delta)$ für $\delta \downarrow 0$. Die nötige Beschränkheit aller $y^*(\delta)$ folgt wieder wie oben mittels Voraussetzung (ii) und Ungleichung (D3). ♦

Bemerkung:
Durch Übergang zu $f_\delta(x) = f(x) + \delta \| x - x^* \|^2$ und $\delta \downarrow 0$ kann man auch auf die Voraussetzung (i) verzichten (sie wird dann durch Lösbarkeit von (P) ersetzt). Hierzu braucht man aber mehr Wissen über konvexe Funktionen; zumindest dass nun wieder (i) für f_δ gilt.

Das meist zitierte Buch zur konvexen Optimierung (in endlicher Dimension) ist wahrscheinlich [Roc70], den aktuellen Stand kann man in etwa erahnen aus [Cla83], [HirLem93], [OutKocZow98] (in letzterem sind auch Spielmodelle direkt erfasst) oder [KK02].

10 Variationsungleichungen (Existenz von Lösungen über Zerlegung der Einheit)

Es sei $\emptyset \ne S \subset R^n$ konvex und kompakt und $f: S \to R^n$ stetig. Es sei ein $x \in S$ gesucht, so dass gilt
(1) $\langle f(x), y - x \rangle \le 0$ für alle $y \in S$.
Diese Bedingung nennt man auch *Variationsungleichung.*
Setzt man $u(x) = x + f(x)$, bedeutet (1) gerade, dass x die Projektion von u(x) auf S sein soll bzw. f(x) im Normalenkegel $N_S(x)$ von S im Punkt x zu liegen hat.

Satz: Unter obigen Voraussetzungen gibt es stets eine Lösung.
Beweis (indirekt)
Angenommen für jedes $x \in S$ ist die Menge $F(x) = \{ y \in S \mid \langle f(x), y - x \rangle > 0 \}$ nicht leer.
Die Mengen $F(x) \subset S$ sind konvex, und die Mengen

$$F^{-1}(y) = \{ x \in S \mid \langle f(x), y - x \rangle > 0 \} \qquad (y \in S)$$

sind (relativ) offen weil f stetig ist.

Weiter gilt $S \subset \cup_{y \in S} F^{-1}(y)$, denn zu jedem $x \in S$ gibt es nach unserer Annahme ein $y \in F(x)$; also ist dann $x \in F^{-1}(y)$. Der Überdeckungssatz liefert die Existenz endlich vieler $y^1, \dots, y^p \in S$ mit

$$S \subset \cup_{k=1,\dots,p} F^{-1}(y^k).$$

Zerlegung der Einheit:
Mit den abgeschlossenen Mengen $A_k = S \setminus F^{-1}(y)$ und den stetigen Abstandsfunktionen

$$d_k(x) = \mathrm{dist}(x, A_k)$$

folgt nach früherem Muster $s(x) := \Sigma\, d_k(x) > 0$. Wir bilden wieder

$$\lambda_k(x) = d_k(x)/s(x) \quad \text{und} \quad g(x) = \Sigma\, \lambda_k(x)\, y^k.$$

Aus $\lambda_k(x) > 0$ folgt $d_k(x) > 0$ und so $x \in F^{-1}(y^k)$, $y^k \in F(x)$ und $\langle f(x), y^k - x \rangle > 0$.
Das liefert nach Multiplikation mit λ_k und Addition wegen $\Sigma\, \lambda_k(x) = 1$,

$$\langle f(x), g(x) - x \rangle > 0 \quad \forall x \in S.$$

Insbesondere ist stets $g(x) \neq x$. Andererseits gilt für $\lambda_k(x) > 0$ auch $y^k \in S$ und – weil S konvex ist – $g(x) \in S$. Da g stetig ist, ergibt sich ein Widerspruch zu Brouwer's Fixpunktsatz. ♦

Spezialisierungen:

1. Ist $S = [a, b]$ ein Intervall in R, soll gelten $x \in [a, b]$ und

$$f(x) = 0 \qquad \text{oder} \quad x = a \text{ und } f(x) \le 0 \qquad \text{oder } x = b \text{ und } f(x) \ge 0.$$

2. Sei $f(x) = Dh(x)$ der Gradient einer C^1 Funktion $h: R^n \to R$, die auf S maximiert werden soll. Bedingung

(1) $\quad \langle f(x), y - x \rangle \le 0 \quad$ für alle $y \in S$

verlangt, dass keine Richtungsableitung von h, genommen in x in Richtung $y - x$, positiv sein darf. *Das ist eine notwendige (und für konkave h auch hinreichende) Bedingung dafür, dass x ein lokaler Maximalpunkt für h auf S ist. Wäre nämlich $\langle f(x), y - x \rangle > 0$ und $y \in S$, würde für Punkte $z(t) = x + t\,(y\text{-}x)$ mit kleinem positiven t sowohl $h(z(t)) > h(x)$ als auch (wegen* Konvexität *von S) $z(t) \in S$ folgen, wonach x kein lokaler MaximalPunkt für h auf S sein kann.*
In diesem Spezialfall kann man also x finden, indem man h auf der kompakten Menge S maximiert. Ist f (wie im Satz) kein Gradient, braucht man allerdings einen „besseren" Existenzbeweis.

3. Mit $x = (\xi, \lambda)$, $S = (X, \Lambda)$, $h: S \to R$ stetig differenzierbar (um das zu sagen, sei h noch auf einer Umgebung von S definiert) und $f = (-D_\xi h(\xi, \lambda),\ D_\lambda h(\xi, \lambda))$, bedeutet (1)

$$\langle -D_\xi h(\xi, \lambda), \xi' - \xi \rangle + \langle D_\lambda h(\xi, \lambda), \lambda' - \lambda \rangle \le 0 \quad \text{für alle } (\xi', \lambda') \in S.$$

Hier können ξ', λ' unabhängig voneinander gewählt werden. Daher wird die Ungleichung zu

$$\langle D_\xi h(\xi, \lambda), \xi' - \xi \rangle \ge 0 \qquad \text{für alle } \xi' \in X,$$
$$\langle D_\lambda h(\xi, \lambda), \lambda' - \lambda \rangle \le 0 \qquad \text{für alle } \lambda' \in \Lambda.$$

Das ist nun analog zu Spez. 2 eine notwendige Bedingung dafür, dass h in (ξ, λ) minimal bzgl. $\xi' \in X$ und maximal bzgl. $\lambda' \in \Lambda$ ist; also eine notwendige Bedingung für einen Sattelpunkt von h. Ist h konvex in ξ und konkav in λ, ist die Bedingung auch hinreichend; wir erhalten so erneut die Existenz eines Sattelpunktes (allerdings nur unter C^1 – Voraussetzung für h), siehe Minimaxsatz.

Der angegeben Beweis des Existenzsatzes läßt sich (mit derselben Grundidee) auf allgemeine Funktionenräume (etwa Banach Räume) verallgemeinern. Aus dem Sklararprodukt wird dann $f(x)(y\text{-}x)$ wobei $f(x)$ ein lineares (beschränktes) Funktional auf X ist (diese Funktionen bilden den Dualraum X^*), das sich mit x stetig ändern darf; $f(x) \in X^*$, $x \to f(x)$ stetig als nichtlineare Funktion von X in X^* (die Norm $\|x^*\|$ ist das Supremum der Werte auf B: $\|x^*\| = \sup \{x^*(x) \mid \|x\| = 1\}$).

In allgemeinerem Kontext kann man $f(x)$ aus (1) als ein ebenfalls gesuchtes Element einer Menge $H(x) \subset X^*$ betrachten, wobei nun H geeignete Voraussetzungen erfüllen muß. Diese haben jedoch stark genug zu sein, um analoge Eigenschaften von F wie im Beweis zu sichern.

11 Lineare Optimierung und Matrixspiele (gemischte Strategien, Interpretation)

In einem (m, n)-Matrixspiel A haben 2 Spieler original m bzw. n Strategien, und Spieler 1 erhält von Spieler 2 bei Anwendung der i-ten und k-ten Strategie der entsprechenden Spieler den Gewinn $a_{i\,k}$.

Die Auswahl der beiden Strategien soll unabhängig voneinander erfolgen, d.h. Sp. 1 sieht erst nach dem Spiel (oder besser nach einer einzelnen *Partie* des Spiels), was Sp. 2 gewählt hat und umgekehrt.
So kann jede Matrix als ein Spiel interpretiert werden. Beide Spieler habe ihre Strategie zu wählen, ohne die des Gegeners zu kennen.
Ein *Gleichgewichtspunkt* in diesem Spiel ist ein Stategienpaar (i0, k0), so dass

$$\mathbf{a_{i0\,k} \geq a_{i0\,k0} \;\; \forall\, k \quad und \quad a_{i\,k0} \leq a_{i0\,k0} \;\; \forall\, i ,}$$

d.h., wenn ein Spieler allein seine Strategie ändert, kann sich sein Gewinn nur verringern (Gewinn kann auch negativ sein).
Ein Gleichgewichtspunkt muss nicht existieren, Beispiel für A (Zahl-Wappen Spiel):

$$\begin{matrix} 1 & -1 \\ -1 & 1 \end{matrix} .$$

Wir nehmen deshalb an, das Spiel werde hinreichend oft gespielt. Dann kann es sinnvoll sein, dass Spieler 1 seine Strategien mit gewissen Wahrscheinlichkeiten wählt. Sei daher x_i die Wahrscheinlichkeit, mit der er die i-te Originalstrategie wählt.
Dann wird $x = (x_1 , \dots , x_m)$ ein Vektor im Simplex $X = \{ x \in R^m / \; x_i \geq 0 \; \forall\, i ; \;\; \Sigma x_i = 1 \}$.
Sein Gewinn gegen Strategie k von Spieler 2 wird eine Zufallsgrösse mit dem *Erwartungswert* $E(x, k) = \Sigma x_i \; a_{i\,k}$. In diesem Zusammenhang nennt man x eine *gemischte* Strategie (für Spieler 1). Dasselbe könnte Spieler 2 tun. Seine gemischten Strategien bilden das Simplex $Y = \{ y \in R^n / \; y_i \geq 0 \; \forall\, i ; \;\; \Sigma y_i = 1 \}$.
Der Erwartungswert des Gewinns bei (unabhängiger) Anwendung von x und y wird so

$$E(x, y) = \Sigma_k \, y_k \, E(x, k) = \Sigma_k \, y_k \; (\Sigma_i \, x_i \; a_{i\,k}) = \Sigma_i \; \Sigma_k \; x_i \; a_{i\,k} \, y_k = x^T A \, y .$$

Das neue Spiel mit den Strategiemengen X und Y und der Gewinnfunktion $f(x, y) = x^T A \, y$ heisst ***gemischte Erweiterung des Matrixspiels*** A.
Ein *Gleichgewichtspunkt (u, v) in Bezug auf gemischte Strategien* ist definiert durch

$$\mathbf{x^T A\, v \;\leq\; u^T A\, v \;\leq\; u^T A\, y \qquad \forall\, x \in X \quad \forall\, y \in Y.}$$

Hauptsatz für Matrixspiele (John v. Neumann, ca 1925).
In der *gemischten Erweiterung* des Spiels A existiert stets ein Gleichgewichtspunkt (GGP), und der Wert $\mathbf{u^T A\, v}$ ist für alle Gleichgewichtspunkte (u, v) derselbe (er heisst *Wert des Spiels*).

Beweis
Tatsächlich ist die Funktion $f(x, y) = x^T A \, y$ bilinear, also konvex-konkav im Sinne des Minimax-satzes; und X und Y sind konvex und kompakt. Also ist die Existenz gesichert. Hat man einen weiteren Gleichgewichtspunkt (u', v'), so folgt

$u'^T A\, v \leq u^T A\, v \leq u^T A\, v'$ weil (u, v) Gleichgewichtspunkt ist und
$u^T A\, v' \leq u'^T A\, v' \leq u'^T A\, v$ weil (u', v') Gleichgewichtspunkt ist.

Das liefert die Behauptung 2 wegen

$$u'^T A\, v' \leq u'^T A\, v \leq u^T A\, v \leq u^T A\, v' \leq u'^T A\, v' . \qquad \blacklozenge$$

Bemerkung:
Gemischte Strategien sind auch in allgemeineren Fällen Wahrscheinlichkeitsverteilungen über der Menge der Originalstrategien, und die neue Gewinnfunktion ist dann der Erwartungswert des nun zufälligen Originalgewinns. In dieser Weise lassen sich auch kompliziertere Spiele behandeln, wenn in den Originalstrategien keine Lösungen existieren.
Letzteres liegt i.a. daran, dass keine sogenannte "vollständige Information" vorliegt, was hier bedeutet, dass Spieler 1 bei seiner Entscheidung für eine Strategie nichts über die Wahl der Strategie von Sp. 2 weiss. Im Gegensatz dazu ist etwa Schach ein Spiel mit vollständiger Information (Strategien sind dort Funktionen, die jeder möglichen Position einen Zug zuordnen). In solchen Spielen existieren i.a. Lösungen in den Originalstrategien. Zur exakten Fassung des hier intuitiven und reichlich vagen Begriffs der Information und zu Spielen mit vollständiger Information, siehe auch [Ber57], [Owe71] , [Kum79], [RaScZa79].

Matrixspiel als lineare Optimierungsaufgabe
Um ein Matrixspiel mittels linearer Optimierung (und ihrer Standard Methoden) lösen zu können, formulieren wir jetzt die Aufgabe, den Wert und einen Gleichgewichtspunkt (GGP) im (gemischten) Matrixspiel zu finden, einfach um. Hierfür definieren wir

$$\boldsymbol{\varphi(v) = \max_{x \in X} x^T A v, \quad \psi(u) = \min_{y \in Y} u^T A y}$$

und bemerken, dass aufgrund der Definition gilt

(u, v) ist ein GGP

$\Leftrightarrow \quad x^T A v \le u^T A v \le u^T A y \qquad \forall x \in X \quad \forall y \in Y.$

$\Leftrightarrow \quad \boldsymbol{\varphi(v) \le u^T A v \le \psi(u).}$

Andererseits gilt trivial wegen $u \in X$ und $v \in Y$

$$\boldsymbol{\varphi(v) \ge u^T A v \ge \psi(u).}$$

Damit ist (u, v) ein GGP genau dann, wenn v die Funktion φ auf Y minimiert und u die Funktion ψ auf X maximiert. Wegen der Existenz eines GGP wissen wir, dass beide Extremalwerte gleich (dem Wert des Spiels) sind. Also haben wir nur

$$\boldsymbol{\min \{ \varphi(y) \ / \ y \in Y \} \quad \text{und} \quad \max \{ \psi(x) \ / \ x \in X \},}$$

zu berechnen, und jedes Lösungspaar (u, v) bildet einen GGP. (φ ist konvex, ψ ist konkav)
Als nächstes vereinfachen wir beide Funktionen zu

$$\varphi(y) = \max_i A_{i.} y, \quad \psi(x) = \min_j x^T A_{.j} \ .$$

Für festes y wird das Maximum $\varphi(y) = \max_{x \in X} x^T (A y)$ angenommen, etwa in $x^o \in X$.
Dann ist in Komponentenschreibweise und mit $e(i) \in X$ als Einheitsvektor $A_{i.}$ = Zeile i von A:

(*) $\quad \varphi(y) = \Sigma x^o_i A_{i.} y = \Sigma x^o_i e(i)^T A y$.

Gilt $e(i)^T A y < \varphi(y) \ \forall i$, so folgt für die Konvexkombination (summiert über i mit $x^o_i > 0$):

$$\Sigma x^o_i e(i)^T A y < \Sigma x^o_i \varphi(y) = \varphi(y)$$

im Widerspruch zu (*). Also existiert ein i mit $\varphi(y) \le e(i)^T A y = A_{i.} y$. Andererseits ist e(i) in X, also auch stets $\varphi(y) \ge e(i)^T A y$. Zusammengefasst, heisst beides

$$\varphi(y) = \max_{x \in X} x^T (A y) = \max_i e(i)^T A y = \max_i A_{i.} y \ .$$

Unsere Aufgabe $\boldsymbol{\min \{ \varphi(y) \ / \ y \in Y \}}$ kann deshalb geschrieben werden als

$\min \{ \max_i A_{i.} y \ / \ y \in Y \}$

$= \min \{ t \ / \ t \ge \max_i A_{i.} y, \quad y \in Y \}$ (mit Variablen $t \in R$ und y)

$\boldsymbol{= \min \{ t + 0^T y \ / \ t \ge A_{i.} y \quad \forall i, \ y \in Y \}.}$

Das ist eine lineare Optimierungsaufgabe. Analog kann man die Aufgabe $\max \{ \psi(x) \ / \ x \in X \}$ studieren und erhält sie in äquivalenter Form als

$\boldsymbol{\max \{ s + 0^T x \ / \ s \le x^T A_{.k} \quad \forall k, \ x \in X \}.}$ ($A_{.k}$ = Spalte k von A)

Spezialfall:
In einem *schiefsymmetrischen* Matrixspiel, d.h., m = n und $A^T = -A$, ist der Wert stets 0, und beide Spieler besitzen dieselben optimalen Stategien; denn $a_{ii} = 0$ und

$\varphi(v) = \max_{x \in X} x^T A v \ge v^T A v = \Sigma_{i > k} (v_i a_{ik} v_k + v_i -a_{ik} v_k) = 0$

und $\quad \psi(u) = \min_{y \in Y} u^T A y \le u^T A u = 0.$

12 Der Julia Robinson Algorithmus (fiktives Spielen) für Matrixspiele

Um ein Matrixspiel A (in gemischten Strategien und nach J. Robinson) *direkt zu lösen*, nehme man folgendes an:
Beide Spieler mögen das Spiel schon N Mal in Originalstrategien ausgetragen haben, wobei Spieler 1 seine i-te Strategie genau P_i mal angewendet habe, Sp. 2 seine j-te Strategie genau Q_j mal.

Für das nächste Mal denke sich Sp. 2, dass Sp. 1 die gemischte Strategie x = P / N anwendet und benutze dagegen eine beste Antwort in Originalstrategien: Spieler 2 wählt also eine Strategie k = k0, die die Summe

$\Sigma_i x_i a_{ik}$ bzgl. k *minimiert* (und ändert damit Q_{k0} um 1)

Analog möge Spieler 1 handeln: Er bildet y = Q/N und wählt eine Strategie i0, die die Summe

$\Sigma_k\, a_{ik}\, y_k$ bzgl. i *maximiert* (und ändert damit P_{i0} um 1).
Im Ergebnis entstehen so Folgen absoluter und relativer Häufigkeiten P(N), Q(N), x(N), y(N), wobei

$$Q(N+1) = Q(N) + e^{k0} \ \ (\text{in Rn}), \qquad P(N+1) = P(N) + e^{i0} \ \ (\text{in Rm}).$$

Die Strategiewahl für den ersten Schritt sei irgendwie willkürlich erfolgt, z.B. P und Q als Einheits-Vektoren, N = 1.

Satz [Rob51]
Jeder Häufungspunkt (x*, y*) der Folge (x(N), y(N)) ist ein GGP der gemischte Erweiterung des Matrixspiels A.

Da ein Häufungspunkt in der kompakten Menge (X, Y) stets existiert, ist dies zugleich ein Existenzsatz.
Der Beweis in [Rob51] ist deshalb originell, weil man induktiv wieder eine stärkere Aussage beweist als formal nötig (s. auch Sperner's Lemma). Der Beweis in [RaScZa79] ist dagegen falsch, weil auf diesen Trick gerade verzichtet wird.

Bemerkungen:
Als Fehlergrösse kann man $f(x, y) = \max_i \Sigma_k\, y_k\, a_{ik} - \min_k \Sigma_i\, x_i\, a_{ik}$ (= 0 für GGP) benutzen. Der Fehler ist nicht monoton in N und konvergiert i.a. sehr langsam.
Beispiel: Für die Matrix

0	1	-2
-1	0	3
2	-3	0

und Anfangsstrategien (e^1, e^1) braucht man N >= 1998000 Iterationen, um f(x, y) < 1E-3 zu erreichen.

Unter den Matrixspielen sind die *schiefsymmetrischen* (m = n und $A^T = -A$) besonders einfach, weil ihr Wert Null ist und optimale Strategien beider Spieler identisch sind. Zum anderen sind sie allgemein genug, weil sich jede lineare Optimierungsaufgabe und jedes Matixspiel auf ein schiefsymmetrisches zurückführen lassen (siehe Lineare Optimierung als Matrixspiel; im Moment *glauben* wir das).

Wir betrachten deshalb den Algorithmus und Robinson's Satz für diesen **Spezialfall** (Der Beweis wird etwas einfacher weil weniger technisch).
Beweis
Put $\max z = \max_k z_k$, $\min z = \min_k z_k$ for $z \in R^n$ and $X(n) = \{ x \in R^n \mid \Sigma x_i = 1,\ x_i \geq 0\ \forall i \}$.
Rows i and columns k of a matrix A shall be denoted by $A_{i\,.}$ and $A_{.\,k}$, respectively. Using a simplification, possible due to $A^T = -A$, J. Robinson's method can be reduced to the following procedure with *arbitrary initial vectors* $z(0) \in R^n$:

Given any $z(0) \in R^n$ and $y(0) = 0 \in R^n$, step $s \geq 0$ requires to find $i = i(s)$ with $z(s)_i = \max z(s)$ and to put $y(s+1) = y(s) + e^i$, $z(s+1) = z(s) + A_{.\,i}$.

Then, it holds $z(s) = z(0) + A\,y(s)$
and $\min z(0) + \max A\,y(s) \leq \max z(s) \leq \max z(0) + \max A\,y(s) \qquad \forall s.$
After division by $s = \Sigma_{1 \leq k \leq n}\, y(s)_k$ one thus obtains $x(s) = y(s) / s \in X(n)$ and the error estimate $(\max z(s) - \max z(0)) / s \leq \max A\,x(s) \leq (\max z(s) - \min z(0)) / s \qquad \forall s > 0.$

We recall the original formulation of

J. Robinson's theorem: Given any $\varepsilon > 0$, it holds
$\mathbf{\max z(s) - \max z(0) \leq \varepsilon\, s \quad if \quad s > s(\varepsilon)}$ **,**
where $s(\varepsilon)$ depends on the (max element $a = \max a_{ik}$ of) matrix A only (and not on z(0)).

Proof:
The statement holds for n = 1 since A =(0). Let it hold for n-1. Then there is some σ such that
(*) $\mathbf{\max z'(s) - \max z'(0) \leq \tfrac{1}{2}\, \varepsilon\, s}$ **for all** $\mathbf{s \geq \sigma}$

if the algorithm is applied to any of the n sub-matrices A' , obtained from A by deleting some row and the related column. We consider

$$N > \max \{ 1, \; 4a/\varepsilon \}$$

intervals of iterations of length σ :

$$z(0), \mathbf{z(1),\dots, z(\sigma),\;\; z(\sigma+1),\dots, z(2\sigma),\;\; z(2\sigma+1),\dots, z(3\sigma)}, \;\dots\;, \mathbf{z((N-1)\sigma+1),\dots, z(N\sigma)}.$$

Since $A = -A^T$, we have $\langle y(s), A\, y(s) \rangle = 0$ and $\langle y(s), z(s) \rangle = \langle y(s), z(0) + A\, y(s) \rangle = \langle y(s), z(0) \rangle$,

hence $\langle y(s), z(s) - z(0) \rangle = 0.$

So $z(s)_i > z_i(0)\ \forall i$ cannot hold for $s > 0$. In consequence, there is some $k = k(s)$ such that

(**) $\quad \mathbf{z(s)_k \le z_k(0) \le \max z(0).}$

Type 1 intervals:

If, for some step s in the m-th interval $I_m = [\,(m-1)\sigma,\dots, m\sigma\,]$, there is some k such that $z(s)_k \le \max z(0)$ *and* component z_k is maximal at some other step $s' \in I_m$, then we obtain

$$\max z(m\sigma) \le z_k(s) + \mathbf{a\,\sigma} \le \max z(0) + \mathbf{a\,\sigma}$$

since component z_k as well as $\max z$ change at most by the value a at each of the σ iterations.

Type 2 intervals:

If type 1 does not happen, then all components z_k satisfying $z_k(s) \le \max z(0)$ with some $s \in I_m$ (where the set of such k is not empty) are *never maximal* in interval I_m. We fix some of them.

By definition, then the algorithm generates during this time the same steps as for the sub-matrix $A^{(-k)}$ without row and column k and with the reduced vectors $z' \in R^{n-1}$ being z without component z_k.

Thus , identifying $z'((m-1)\sigma)$ with the initial vector $z'(0)$ of sub-game $A^{(-k)}$, we have from (*)

$$\max z(m\sigma) \le \max z((m-1)\sigma) + \tfrac{1}{2}\, \boldsymbol{\varepsilon\,\sigma}.$$

After the last interval of type 1, there are at most N intervals of type 2; the same is true if type 1 does not occur at all. Hence we obtain

$$\mathbf{\max z(N\sigma) - \max z(0) \le a\,\sigma + \tfrac{1}{2} N\,\varepsilon\,\sigma .}$$

Setting finally $q = N\sigma + p$ where $0 \le p < \sigma$, we have $\max z(q) \le \max z((N+1)\sigma)$.

Since $\sigma \le q/N$ this yields $\max z(q) - \max z(0) \le a\,q / N + \tfrac{1}{2}\varepsilon\, q\,(N+1)/N$ and by the choice of N,

(***) $\quad \mathbf{\max z(q) - \max z(0) < \varepsilon\, q}$. ♦

Notes:

1 Formula (***) holds if $q = N\sigma + p < (N+1)\sigma$, $N > \max\{1, 4a/\varepsilon\}$ (σ, $\tfrac{1}{2}\varepsilon$ from n-1 games).

2 max z increases only if the active index i(s) changes.

3 Really, it would be enough to show the theorem for z(0) = 0 only. However, then the induction-proof does not work since $z'((m-1)\sigma)$ cannot be identified with an initial vector !

13 Schnelle Modifikation dieses Algorithmus und numerische Beispiele

Man kann den Algorithmus wesentlich verbessern (B. Kummer, Spieltheorie Vorlesungen und verschiedene Vorträge ab 1990): Dazu muss man nur zusätzlich ausrechnen, wir oft hintereinander das Strategienpaar (k0, i0) aus Schritt N auch in den folgenden Schritten "beste Antwort" ist:

Setzt man in Schritt N:

$$V = P^T A, \quad U = A\,Q \quad \text{und ist } U_{i0} \text{ maximal}, \quad V_{k0} \text{ minimal},$$

so bleiben dieselben Komponenten von (U, V) - nach Konstruktion - auch minimal bzw. maximal in den Schritten N, … , N+h, solange gilt

$$A_{i0\,.}\,(Q + h\,e^{k0}) \ge A_{i\,.}\,(Q + h\,e^{k0}) \quad \text{bzw.} \quad U_{i0} + h\,a_{i0,k0} \ge U_i + h\,a_{i,k0}$$

und $\quad (P + h\,e^{i0})\,A_{.\,k0} \le (P + h\,e^{i0})\,A_{.\,j} \quad$ bzw. $\quad V_{k0} + h\,a_{i0,k0} \le V_j + h\,a_{i0,j}$.

Zusammengefasst liefert dies die einfachen Bedingungen

(1) $\quad U_{i0} - U_i \ge h\,(a_{i,k0} - a_{i0,k0}) \quad \forall\ i = 1, \dots, m \quad$ mit $\quad a_{i,k0} - a_{i0,k0} > 0 \quad$ und

(2) $\quad V_{k0} - V_j \le h\,(a_{i0,j} - a_{i0,k0}) \quad \forall\ j = 1, \dots, n \quad$ mit $\quad a_{i0,j} - a_{i0,k0} < 0.$

Bildet nicht schon (i0, k0) eine GGS in reinen Strategien, so gibt es ein maximales, natürliches h ≥ 0 mit (1) und (2),

das sich durch weniger als n+m Divisionen ermitteln läßt.

Man weiss damit, dass erst im J-Robinson-Schritt N+h+1 ein Indexwechsel erfolgen *muss* (in Schritt N+h realisieren (i0, k0) noch die Extrema.

Also kann man alle diese h+1 Schritte gleichzeitig machen.

Man addiert auf P_{k0} , Q_{i0} und N einfach h+1 und erhält so die nächsten P' und Q'.
Anschliessend bestimmt man für N' = N+h+1 erneut das entsprechende (maximale) h',
und aggregiert weitere Originalschritte in derselben Weise.

Die neuen Schritte sind jetzt wegen der Bestimmung von h etwa 4-8 mal "teurer" (in Aufwand oder Rechenzeit bei integer Matrizen) als die alten. Sie modellieren allerdings viele Originalschritte, obgleich in manchen Iterationen auch nur h = 0 sein kann. Es läßt sich jedoch leicht zeigen:

Satz

Der Fehler beider Algorithmen nach N (originalen bzw. neuen) Schritten erfüllt

$$F_{new}(N) / F_{old}(N) \to 0 \quad \text{für} \quad N \to \infty$$

sofern alle a_{ik} rational sind.

Beweis: Siehe Abschätzung (Vergleich der Fehler) für den schiefsymmetrischen Fall unten. ♦

Mehr noch:

Bei allen (von sehr vielen, $> 10^7$) praktischen Tests mit Matrix-Dimension von (5, 5) bis $(10^6, 10^6)$ galt ausserdem mit $m(A) = \max_{i,k} |a_{i,k}|$

(3) $\mathbf{F_{new}(N) \leq 2\ m(A)\ (n+m)\ /\ N}$ **(noch nicht allgemein bewiesen).**

Beispiel *für die schiefsymmetrische (50, 50) Matrix mit* $a_{ii} = 0$

$a_{ik} = 1 + i - [k/2]$ *für* $k > i$, $[k/2]$ = *grösste ganze Zahl g mit* $g \leq k/2$;
$a_{ik} = - a_{ki}$ *sonst* .

Man benötigt $N_{new} = 7423$ *bzw.* $N_{old} = 5808400$ *Iterationen , um den Fehler kleiner als 0.02 zu machen mit Anfangsstrategien* (e^1, e^1).

Ist Formel (3) generell richtig, so folgt für geforderte absolute Genauigkeit ε:

$$F_{new}(N) < \varepsilon \quad \text{falls} \quad 2\, m(A)\, (n+m) \,/\, N < \varepsilon.$$

Also reichen $N > 2\,(n+m)\ m(A) / \varepsilon$ (neue) Iteration. Jede dieser Iteration besteht aus (grob gesagt) etwa 5 (n+m) Elemetaroperationen. Damit würden

(4) $\mathbf{N_E > 10\ (n+m)^2 / \alpha}$

Elementaroperationen reichen, um den relativen Fehler kleiner als $\alpha = \varepsilon / m(A)$ zu machen. So schnell ist (bei grossem n, m und festem α) kein anderes, z. Zt. bekanntes Verfahren, da deren Aufwand mindestens von der Ordung $(n+m)^3$ ist.

Vereinfachung für Schiefsymmetrie:

Ist m = n und $A^T = -A$, heißt A schiefsymmetrisch. Dann gilt wegen $x^T A x \equiv 0$:

(x, y) ist GGP ⇔ x und y (aus dem Einheitssimplex S) lösen $A x \leq 0$.

Damit läßt sich auch der modifizierte Algorithmus vereinfachen:

Beginnend mit $Z^0 = 0$, $Y^0 = 0$ (in R^n) bestimmt man in Schritt $s \geq 0$:
i(s) so, dass Komponente i(s) von Z^s maximal ist
und r so, dass $r = \min_k \{ (Z^s_{i(s)} - Z^s_k) / a_{k\, i(s)} \mid a_{k\, i(s)} > 0 \}$
(wenn $a_{k\, i(s)} \leq 0\ \forall k$, löst e^{i0} das Spiel).
Nun setzt man h = [r] +1 (ganzer Teil von r +1) und bildet einfach
$Y^{s+1} = Y^s + h\, e^{i(s)}$, $Z^{s+1} = Z^s + h\, A_{.\, i(s)}$.

Dann ist für s > 0: $Z^s = A\, Y^s$, $x(s) = Y^s / s \in S$ und $A\, x(s) = Z^s / s$.

Der Original-Algorithmus würde dieselben Schritte mit konstantem $h \equiv 1$ verlangen.

Vergleich der Fehler $\mathbf{f(s) = \max_i A_{i.}\, x(s) = \max_i Z^s_i / s}$

für beide Algorithmen:

Seien alle $a_{i,k}$ ganzzahlig. Sei N(s) die Zahl der J. Robinson Schritte, die durch s Schritte des modifizierten Algorithmus modelliert wurden.

Wenn kein Indexwechsel beim R-Algorithmus erfolgen muss, bleibt die maximale Komponente von Z konstant (da $a_{ii} = 0$); ansonsten wächst sie um mindestens 1 und höchstens m(A). Daher gilt

$$s / N(s) \leq f_{ROB}(N(s)) \leq s\ m(A) / N(s)$$

Andererseits ist auch $s / N(s) \leq \max_i Z^s_i / s = f(s) = f_{ROB}(N(s))$ richtig.

Hiermit folgt sofort

(5) $\quad \mathbf{s \;\leq\; N(s)\; f_{ROB}(\,N(s)\,)},$

d.h. insbesondere $s \ll N(s)$, weil (nach J.Rob.) $f_{ROB}(N(s)) \to 0$ für $s \to \infty$. ♦

Die kritische Formel (3) reduziert sich im aktuellen Fall auf die Form

(3)' $\quad \mathbf{f_{new}(s) \;\leq\; m(A)\; n / s}$.

Wir zeigen nun: Jedes lineare Optimierungsproblem (und jedes Matrixspiel) läßt sich in ein schiefsymmetrisches Spiel umwandeln. Daher reicht es völlig, allein diesen simplen Fall zu studieren.

14 Lineare Optimierung als Matrixspiel

Zunächst leiten wir einige Fakten über lineare Optimierungsaufgaben her. Unser Ausgangspunkt ist die lineare Optimierungsaufgabe

(LP) $\quad \max \{ c^T x \;/\; A x \leq b,\; x \geq 0 \}$

(A ist eine (m, n)-Matrix, $c, x \in R^n$, $b \in R^m$; Ungleichungen sind komponentenweise zu verstehen).
Als ihre Dualaufgabe bezeichnet man

(LD) $\quad \min \{ b^T y \;/\; A^T y \geq c,\; y \geq 0 \}$.

Wegen Nichtnegativität der Variablen folgt für beliebige zulässige Punkte x, y der beiden Probleme

$$y^T A x = \Sigma_k\, y_k\, A_{k.} x \;\leq\; \Sigma_k\, y_k\, b_k = y^T b \quad \text{und analog} \quad x^T A^T y \geq x^T c;$$

also auch

(1.1) $\quad c^T x \;\leq\; x^T A^T y = y^T A x \;\leq\; y^T b$.

Damit gilt im Falle der Lösbarkeit beider Aufgaben für ihre Extremalwerte v_P und v_D stets $v_P \leq v_D$.

Wir beweisen, dass sogar $v_P = v_D$ gilt (das ist der *Dualitätssatz der linearen Optimierung*).

Sei x optimal für (LP). Wir betrachten ein beliebiges $u \in R^m$ mit

$u_i \geq 0$ falls $x_i = 0$,

$A_{k.}\, u \leq 0$ falls $A_{k.}\, x = b_k$ (nur diese Indizes i, k werden in der Folge betrachtet).

Dann ist x + t u zuläesig für (LP) bei kleinen t > 0. Also folgt $c^T u \leq 0$ wegen Optimalität von x.
Wir zeigen damit, dass sich c^T schreiben lässt als nichtnegative Linearkombination der entsprechenden Vektoren $-e^i$ und $A_{k.}$ (das ist die nichttriviale Richtung des *Lemmas von Farkas*).
Andernfalls wäre c kein Element des abgeschlossenen konvexen Kegels $H \subset R^n$ aller Vektoren

$$h = \Sigma\, \mu_k\, A_{k.} - \Sigma\, \lambda_i\, e^i \quad \text{darstellbar mit} \quad \mu_k \geq 0,\; \lambda_i \geq 0 .$$

Also könnte man c von H trennen, d.h., es gibt ein $w \in R^n$, so dass

$$\langle h, w \rangle < \langle c, w \rangle \quad \forall\, h \in H ,$$

folglich $\quad \Sigma\, \mu_k\, A_{k.}\, w - \Sigma\, \lambda_i\, e^i\, w < \langle c, w \rangle \quad \forall\, \mu_k \geq 0,\; \lambda_i \geq 0$.

Hieraus erhalten wir $0 < \langle c, w \rangle$ und $A_{k.}\, w \leq 0$, $e^i w \geq 0$ im Widerspruch zu den Eigenschaften von u.
Somit ist

$$\mathbf{c^T = \Sigma\, \mu_k\, A_{k.} - \Sigma\, \lambda_i\, e^i \quad \text{mit gewissen} \quad \mu_k \geq 0,\; \lambda_i \geq 0 \quad \text{richtig.}}$$

Setzen wir die restlichen Komponenten μ_r (zu $A_{r.}\, x < b_r$) Null, wird $\mu \in R^m$ zulässig für (LD), und

$$b^T \mu = \Sigma\, \mu_k\, b_k = \Sigma\, \mu_k\, A_{k.}\, x = c^T x + \Sigma\, \lambda_i\, e^i\, x = c^T x.$$

Damit folgt $\inf \{ b^T y \mid A^T y \geq c,\; y \geq 0 \} \leq b^T \mu \leq c^T x = v_P$, also auch $v_D \leq v_P$ und $v_D = v_P$. ♦

Im Falle von $c^T x = b^T y$ erhält man also Optimalität zulässiger Punkte x, y in (LP) bzw. (LD). In analoger Weise kann man von einer Lösung der Dualaufgabe ausgehen und zeigen, dass auch die primale lösbar ist. Man erhält so den **Dualitätssatz (d. lin. Optimierung):** *Die Aufgaben (LP) und (LD) sind stets gleichzeitig lösbar und haben im Falle der Lösbarkeit denselben Extremalwert.* ♦

Außerdem: Aus $-\infty < \inf \{ b^T y \mid A^T y \geq c,\; y \geq 0 \}$ und der Existenz eines zulässigen y folgt über Induktion bzgl. n+m (wir beweisen es hier nicht), dass das Infimum (in einem zulässigen Punkt) auch angenommen wird (analog für Primalaufgabe). Dies ist der ***Existenzsatz der linearen Optimierung.***

Insbesonder sind beide Zielfunktionen über den entsprechenden Restriktionsmengen beschränkt - nach oben bzw. unten siehe (1.1) - wenn es zulässige Punkte für beide Aufgaben gibt, Dies sichert also schon die Lösbarkeit.

Wir folgen nun [Vor74] und betrachten das nachstehende schiefsymmetrische Matrix-Spiel G:

$$G = G(A,b,c): \quad \begin{matrix} 0 & -A^T & c \\ A & 0 & -b \\ -c^T & b^T & 0 \end{matrix}.$$

Wegen $G^T = -G$ ist sein Wert 0, und die optimalen Strategien $s = (\xi , \eta , \tau)^T$ im Simplex X(n+m+1) von Spieler 2 sind durch $Gs \le 0$ charakterisiert, weil dies gerade $\boldsymbol{\sigma}^T \mathbf{G}\, \mathbf{s} \le 0$ für alle Strategien $\boldsymbol{\sigma}$ von Sp. 1 liefert. Man beachte: Optimal für Sp. 2 ist hier dasselbe wie optimal für Sp.1; wir reden also nur von optimalen Strategien.

Zusammenhang zwischen G, (LP) und (LD):

(i) Sei $s = (\xi , \eta , \tau)^T$ eine optimale Strategie für G mit $\tau > 0$. Dann bilden

(5.1) $x = \xi / \tau$ und $y = \eta / \tau$

ein Paar von primal-dual Lösungen zu (LP), (LD).

Denn die Ungleichungen zu $Gs \le 0$ liefern Zulässigkeit

$0\,\xi \;\; - A^T \eta \;\; + c\,\tau \le 0 \;\; \Rightarrow \;\; - A^T y \; + c \le 0,$

$A\,\xi \;\; -0 \;\; - b\,\tau \le 0 \;\; \Rightarrow \;\; A\, x \; - b \le 0,$

und Optimalität:

$-c^T \xi \; + \; b^T \eta \; + 0 \le 0 \;\; \Rightarrow \;\; b^T y \le c^T x$ (Mit (1.1) muss $b^T y$ minimal, $c^T x$ maximal sein).

(ii) Umgekehrt, ist (x, y) ein Paar von Lösungen zu (LP), (LD), so bilde man

(5.2) $\tau = (1 + \Sigma x_i + \Sigma y_k)^{-1}$, $\xi = \tau x$ and $\eta = \tau y$.

Dann ist $s = (\xi , \eta , \tau)^T$ optimal im Spiel und $\tau > 0$.

Denn es folgt: Die Komponenten von s sind nicht negativ, ihre Summe ist 1 und

$0\,\xi \;\; - A^T \eta \;\; + c\,\tau \;\le\; 0$

$A\,\xi \;\; -0 \;\; - b\,\tau \;\le\; 0$.

Nutzt man auch $c^T x = b^T y$, so folgt $-c^T \xi \; + \; b^T \eta \; + 0\,\tau \; = 0.$

Also löst dann $s = (\xi , \eta , \tau)$ das Spiel.

Die Spiel-Lösungen mit $\tau > 0$ und primal-dual Lösungen der Optimierungsaufgaben sind also durch (5.1) und (5.2) miteinander verbunden.

- Hat das Spiel nur Lösungen mit $\tau = 0$, so sind beide Aufgaben (LP), (LD) unlösbar.
  Denn aus der Lösbarkeit der einen folgt Lösbarkeit der anderen, und (ii) sichert $\tau > 0$ für eine optimale Strategie.
- Sind beide Aufgaben (LP), (LD) unlösbar, so hat das Spiel nur Lösungen mit $\tau = 0$.
  (siehe (i)).
- Hat das Spiel Lösungen mit $\tau > 0$ und $\tau = 0$, so folgt aus der Konvexität der Lösungsmenge, dass es optimale Strategien s^v ($v = 1, 2, ...$) gibt mit dazugehörigen $\tau_v \to +0$.

Aus (5.1) und $x^v = \xi^v / \tau_v$ und $y^v = \eta^v / \tau_v$ folgt dann, dass wenigstens eine der Lösungsmengen von (LP) und (LD) unbeschränkt ist, weil nicht gleichzeitig $\|\xi^v\|$ und $\|\eta^v\|$ verschwinden können (Summenbedingung des Simplex).

Kennt man alle Lösungen des Spieles G, hat man also volle Information über die Lösungen von (LP) und (LD). Die minimale Komponente τ^* bzgl. aller Lösungen $s = (\xi , \eta , \tau)$ zu G spielt die Rolle einer condition number für (LP) und (LD), wobei $\tau^* = 0$ für Degeneriertheit steht (unbeschränkte oder leere Lösungsmengen) und grosses τ^* sichert, dass (LP) und (LD) Lösungen mit kleiner Norm besitzen.

15 Nash-Verhandlungslösungen

Zwei Spieler mögen die Gelegenheit haben, sich einen Gesamtgewinn g zu teilen; der erste (er besitze schon x0) bekommt x, der zweite (er besitze y0) erhält y; $x+y \le g$.

Sei weiter U(h) und V(h) der vermeintliche Nutzen, den beide jeweils einer Geldmenge $h \ge 0$ zuschreiben. Dann erhöht sich dieser um

$\delta U(x) = U(x0+x) - U(x0)$ bzw. $\delta V(y) = V(y0+y) - V(y0)$,

und gewöhnlich sind U und V monotone, konkave Funktionen, die unterproportial wachsen; etwa

$U = c_u \ln (h + 1)$, $V = c_v \ln (h + 1)$.

Problem:

Welche Gewinn- (bzw. Nutzensaufteilung) ist „gerecht“ ? Dazu betrachte man alle $(u, v) \in R^2$ mit

$u = U(x0+x)$, $v = V(y0+y)$, $x+y \le g$, $x, y \ge 0$.

Sie mögen eine ***konvexe abgeschlossene Menge S*** bilden, von der uns der Nord-Ost-Teil

$S^+ = \{ (u, v) \in S \mid u \geq u^o, v \geq v^o \}$

mit $u^o := U(x0)$, $v^o := V(y0)$, $(u^o, v^o) \in S$ interessiert.

Die „gerechte Lösung" (u^*, v^*) sei nun durch eine Funktion Φ definiert, die jeder konvexen abgeschlossenen Menge $S \subset R^2$ mit $(u^o, v^o) \in S$ und beschränktem S^+, ein

$(u^*, v^*) = \Phi(S, u^o, v^o)$ aus S^+ zuordnet.

Was gerecht bzw. vernünftig ist, forderte Nash mittels folgender **Axiome.**

(1) Pareto-Optimalität:
(u^*, v^*) erfülle, dass kein $(u, v) \in S$ exist. mit $(u, v) \geq (u^*, v^*)$ und $(u, v) \neq (u^*, v^*)$.

(2) Symmetrie: Wenn S^+ symmetrisch ist, d.h. $(u, v) \in S^+ \Leftrightarrow (u, v) \in S^+$, so sei $u^* = v^*$.

(3) Invarianz gegenüber verbotenen Alternativen: Wenn $(u^*, v^*) \in T^+ \subset S^+$ und $(u^o, v^o) \in T$, so sei $(u^*, v^*) = \Phi(T, u^o, v^o)$.

(4) Invarianz gegenüber linearen Transformationen: Wenn L eine positive lineare Transformation des R^2 ist; $u' = \alpha_1 u + \alpha_o$, $v' = \beta_1 v + \beta_o$ mit $\alpha_1 > 0$, $\beta_1 > 0$, so sei Φ invariant gegenüber L, d.h. $\Phi(LS, Lu^o, Lv^o) = L(u^*, v^*)$.

Satz (Nash) Es gibt genau eine Funktion Φ mit den Eigenschaften (1),…,(4). Sie ordnet (S, u^o, v^o), jeweils die eindeutige Lösung (u^*, v^*) der Aufgabe

(A1) $\max \{ f(u, v) \mid (u, v) \in S^+ \}$, *sofern* $(u, v) \in S^+$ *mit* $f(u, v) := (u - u^o)(v - v^o) > 0$ *existiert,* bzw.

(A2) $\max \{ u + v \mid (u, v) \in S^+ \}$ *andernfalls*

zu.

Beweis

In Situation (A2) ist S^+ einfach eine Strecke, und die zu beweisenden Aussagen folgen sofort aus Axiom (1). Wir nehmen daher an, dass Situation (A1) vorliegt.

Existenz von Φ:

Lösungen von (A1) existieren wegen Kompaktheit von S^+ und Stetigkeit der Zielfunktion. Die Lösungen sind (wegen Konvexität von S^+ und der speziellen Form von f) auch eindeutig.

Löst (u^*, v^*) (A1), so folgt als notwendige Optimalitätsbedingung

(O1) $\langle Df(u^*, v^*), (u - u^*, v - v^*) \rangle \leq 0 \quad \forall (u, v) \in S^+$.

Zusammen mit $(u^*, v^*) \in S^+$ ist (O1) auch hinreichend für Optimalität von (u^*, v^*) (was man direkt nachrechnen kann). Die linke Seite in (O1) ist aber gerade

$(v^* - v^o)(u - u^*) + (u^* - u^o)(v - v^*)$.

Nach einer linearen Tansformation wie angegeben, werden deshalb die Ungleichungen in (O1) zu

(O2) $\alpha_1 \beta_1 (v^* - v^o)(u - u^*) + \alpha_1 \beta_1 (u^* - u^o)(v - v^*) \leq 0 \quad \forall (u, v) \in S^+$

und zeigen, dass die Invarianz (4) für die Lösungen von (A1) erfüllt ist.

Die Gültigkeit von (1), (2), (3) für die Lösungen von (A1) weist man sofort direkt nach.

Eindeutigkeit

Sei $(u^*, v^*) = \Phi(S, u^o, v^o)$ und (u^{**}, v^{**}) Lösung von (A1). Da $(u^*, v^*) = (u^{**}, v^{**})$ zu zeigen ist, nehmen wir das Gegenteil an und bilden $T = \operatorname{conv} \{ (u^o, v^o), (u^*, v^*), (u^{**}, v^{**}) \}$.

Nach (3) ist auch $(u^*, v^*) = \Phi(T, u^o, v^o)$. Außerdem bleibt (u^{**}, v^{**}) offenbar auch Lösung von (A1) für dieses Teilsystem (T, u^o, v^o).

Seien nun positive α_1 und β_1 sowie reelle α_o und β_o so gewählt, dass (u^o, v^o) in den Ursprung und T in ein symmetrisches (zur Achse u = v) Dreieck $T' = L(T)$ transformiert wird (solche Konstanten kann man leicht ausrechnen). Für $(T', 0, 0)$ ist offenbar der NordOstPunkt mit $u^1 = v^1$ die Lösung von (A1). Wegen (2) gilt auch $(u^1, v^1) = \Phi(T', 0, 0)$.

Weil auch die Lösungen von (A1) invariant gegenüber positiven Transformationen sind (wie oben bewiesen), folgt nun mit der inversen Transformation L^{-1} wegen $(T, u^o, v^o) = L^{-1}(T', 0, 0)$,

$L^{-1}(u^1, v^1) = \Phi(T, u^o, v^o)$ und $L^{-1}(u^1, v^1)$ löst (A1) für (T, u^o, v^o).

Das zeigt aber $(u^*, v^*) = (u^{**}, v^{**})$. ♦

16 Drohungen im Bimatrixspiel (nach Owen)

Wir betrachten ein Bimatrixspiel. Es unterscheidet sich von einem Matrixspiel dadurch, dass es nicht antagonistisch ist, also erhält Sp. 1 bei Anwendung gemischter Strategien x, y aus den Simplizes X bzw. Y den Gewinn $x^T A y$, Spieler 2 den Gewinn $x^T B y$ mit entsprechenden (m, n) Matrizen.
Jetzt sei Kooperation erlaubt; beide Spieler können also gemeinsam die Summe

$$x^T A y + x^T B y$$

maximieren, der maximale Wert sei $g > 0$, und nun darüber streiten, wie sie g aufteilen.
Falls sie sich nicht einigen können, bleibt nur, dass sie Strategien x und y wählen und den entsprechenden Gewinn $u_o(x, y) = x^T A y$ bzw. $v_o(x, y) = x^T B y$ erhalten.

Setzt man der Einfachkeit halber Nutzen = Geld, so spielen nun $u_o(x, y)$ und $v_o(x, y)$ die Rolle von u_o und v_o im Nash-Verhandlungsproblem mit
$S = \{ (u, v) \mid u + v \leq g \}$. Die Nash-Verhandlungslösung ist offenbar

$$u^* = u_o + h / 2 ,$$
$$v^* = v_o + h / 2,$$

wobei h der Überschuß $h = g – (u_o + v_o)$ ist. Das Ergebnis hängt von (x, y) ab.

Eine Möglichkeit der Wahl von (x, y) wäre, sich den jeweils garantierten Mindestgewinn zu sichern, d.h. die Spieler wählen (x', y'), so dass

$$\min_y x'^T A y = \max_x \min_y x^T A y \qquad (= U_1)$$
$$\min_x x^T A y' = \max_y \min_x x^T B y \qquad (= V_1).$$

Im Resultat wäre

$$u_o = x'^T A y' \geq U_1 \quad \text{bzw.} \quad v_o = x'^T B y' \geq V_1 .$$

Wir können aber auch „*Drohungen*“ berücksichtigen. Das soll heißen:
Spieler 1 nimmt bei der Wahl von x in Kauf, dass sein Gewinn $x^T A y$ recht klein wird (bei ungünstigem y), wenn der Gewinn des Gegners noch viel kleiner wird (z.B. wenn x „Streik“ bedeutet, der mit hohem Verlust von Sp. 2 verbunden ist). Analog kann Sp. 2 denken.

Beispiel

$$A = \begin{matrix} -3 & -2 \\ 4 & 2 \end{matrix} \qquad B = \begin{matrix} -33 & -20 \\ 10 & 8 \end{matrix}$$

$g = 4 +10$.
Strategie $x' = e^2$ garantiert Sp.1 mindestens den Gewinn 2.
Strategie $y' = e^2$ garantiert Sp.2 mindestens den Gewinn -20.
Wählen beide diese Strategien, wird $(u_o , v_o) = (2, 8)$ und

$$(u^*, v^*) = (4 , 10) \text{ wegen } h = 4.$$

Nun wähle Sp. 1 $x = e^1$, womit er (gegen $y = e^1$) den „Gewinn“ –3 in Kauf nimmt.
Welche gemischte Strategie y Spieler 2 auch wählt, immer gilt

$$u_o = x^T A y \in [-3, -2] \quad \text{und} \quad v_o = x^T B y \in [-33, -20].$$

Mit minimalem u_o und maximalem v_o kann man u^* im Verhandlungsproblem abschätzen.
Damit wird $u^* \geq -3 + h/2$, wobei $h \geq 14 - (-2 - 20) = 36$ ist, also

$$u^* \geq -3 + 18 = 15 \quad \text{und} \quad \text{demzufolge } v^* \leq -1.$$

Es ist also durchaus von Nutzen, wenn Sp. 1 droht, bei Nicht-Einigung $x = e^1$ zu wählen.

Was sind nun „optimale Drohungen“ ?
Im resultierenden Verhandlungsproblem nimmt $x^T A y$ den Platz von u_o ein, analog ist $v_o = x^T B y$. Damit hängt die Lösung (u^*, v^*) von den „Drohstrategien“ x und y wie folgt ab:

$$u^* = u^*(x, y) = x^T A y + (g – (x^T A y + x^T B y)) / 2 ,$$
$$v^* = v^*(x, y) = x^T B y + (g – (x^T A y + x^T B y)) / 2 .$$

Man erhält so eine neues Spiel zur Wahl von x und y mit ebendiesen Gewinnen.
Bis auf die Konstante g wird so

$$u^*(x, y) = (x^T A y - x^T B y) / 2 = ½ \ x^T (A - B) y$$
$$v^*(x, y) = (x^T B y - x^T A y) / 2 = ½ \ x^T (B - A) y.$$

Das neue Spiel ist also gerade das (antagonistische) Matrixspiel mit der Matrix A – B (der Faktor ändert nichts). Also gilt der

Satz Das Matrixspiel A–B liefert optimale Drohungen für beide Spieler im beschriebenen Bimatrixspiel (A, B) mit Drohungen.

17 E. Michael's selection Theorem (1956)

Satz (vereinfachte R^n -Version)
Es sei $F: X \to Y$ eine mehrwertige, uhs. Abbildung mit nichtleeren, konvexen Bildern $F(x) \subset Y$, $X \subset R^n$, $Y = R^m$, X nichtleer und kompakt.
Dann existiert eine stetige Funktion $f: X \to Y$ mit $f(x) \in F(x)\ \forall x \in X$.

Dabei steht uhs. für unterhalb stetig und bedeutet *für mehrwertige Abbildungen F*, dass für jede offene Menge $\Omega \subset Y$ die Urbildmenge $F^-(\Omega) := \{ x \mid F(x) \cap \Omega \neq \emptyset \}$ ebenfalls (relativ) offen in X ist. Äquivalent zu uhs sind:
(i) $\lim_{x' \to x} \text{dist}(y, F(x')) = 0\ \forall (x, y)$ mit $y \in F(x)$ [diese Paare bilden gph F] bzw.
(ii) $\forall (x, y) \in$ gph F gibt es eine *in x stetige* Auswahlfunktion $g_{x,y}$ für F mit $g_{x,y}(x) = y$.
Beweis.
Teil 1
Konstruktion von stetigem f, so dass $f(x) \in$ cl F(x) (Abschließung von F(x)) $\forall x \in X$.
Zu gegebenem $\varepsilon > 0$ sei $F_\varepsilon(x) = \{ y' / \text{dist}(y', F(x)) < \varepsilon \}$. Dann ist $F(x) \subset F_\varepsilon(x)$, und die Bilder der ("um ε aufgeblasenen") mehrd. Abbildund F_ε bleiben konvex. Wir zeigen zuerst die Existenz einer stetigen Auswahl für $F_\varepsilon(x)$.
Ist $y' \in F_\varepsilon(x)$, so gibt es ein $y \in F(x)$ mit $d(y', y) = \delta < \varepsilon$. Weil die offene Kugel $K(y, \varepsilon-\delta)$ stets Punkte aus $F(x')$ enthält, sofern $x' \in X$ und $d(x', x)$ hinreichend klein ist (uhs), folgt für diese x' auch $y' \in F_\varepsilon(x')$. *Damit sind die Urbildmengen $F_\varepsilon^-(y') = \{ x \in X \mid y' \in F_\varepsilon(x) \} \subset X$ alle offen.*
Zu jedem $x \in X$ gibt es weiter ein $y \in F(x)$; also ist dann $x \in F^-(y) \subset F_\varepsilon^-(y)$.
Die Mengen $F_\varepsilon^-(y')$ mit $y' \in Y$ bilden daher eine offene Überdeckung der kompakten Menge X. Nach Heine-Borel gibt es somit *endlich viele y_i*, mit denen schon $X \subset \cup F_\varepsilon^-(y_i)$ gilt. Nach bekanntem Vorbild konstruiert man wieder eine Zerlegung der Einheit für diese Überdeckung, wir erhalten also stetige Funktionen λ_i mit den Eigenschaften:

$$\lambda_i(x) \geq 0,\quad \Sigma_i \lambda_i(x) = 1 \quad \text{und} \quad \lambda_i(x) > 0 \Rightarrow x \in F_\varepsilon^-(y_i)\,.$$

Damit definieren wir

$$f(x) = \Sigma_i \lambda_i(x)\, y_i\,.$$

Dann ist f stetig und f(x) jeweils eine Konvex-Kombination von Punkten y_i mit $x \in F_\varepsilon^-(y_i)$ bzw. $y_i \in F_\varepsilon(x)$. Wegen Konvexität von $F_\varepsilon(x)$ gilt also $f(x) \in F_\varepsilon(x)$; f ist eine stetige Auswahlfunktion für $F_\varepsilon(x)$. Wir nennen eine solche Funktion f nun f_ε.
Konstruktion einer stetigen Auswahlfunktion für cl F: Hat man f_ε, kann man die Abbildung

$$G_\varepsilon(x) = F(x) \cap \{ y \mid d(y, f_\varepsilon(x)) < \varepsilon \} \subset F(x)$$

definieren, die wegen $\text{dist}(f_\varepsilon(x), F(x)) < \varepsilon$ erneut nichtleere, konvexe Bider besitzt und (wie man leicht direkt zeigt) wieder uhs ist.
Beginnend mit $F_1 = F$, ist dann mit $\varepsilon_k = 4^{-k}$ folgende Konstruktion möglich:
Zu F_k wähle man f_k als stetige ε_k- Auswahl und bilde

$$F_{k+1}(x) = F_k(x) \cap \{ y \mid \text{dist}(y, f_k(x)) < \varepsilon_k \}.$$

Da $F_{k+1}(x)$ wieder nichtleer und konvex ist und F_{k+1} auch uhs bleibt, lässt sich die Konstruktion tatsächlich wiederholen. Man erhält so eine Folge von Abbildungen $F_{k+1} \subset F_k$ und Funktionen, die in allen Argumenten x erfüllen:

$$d(f_k, F_k) < \varepsilon_k, \quad \text{also erst recht} \quad d(f_k, F) < \varepsilon_k,$$

und $d(f_k, f_{k+m}) \leq d(f_k, f_{k+1}) + d(f_{k+1}, f_{k+2}) + ... + d(f_{k+m-1}, f_{k+m})$.
Mit $y \in F_{k+1}(x)$, so dass $d(f_{k+1}, y) < \varepsilon_{k+1}$, folgt so per Definition

$$d(f_k, f_{k+1}) \leq d(f_k, y) + d(y, f_{k+1}) < \varepsilon_k + \varepsilon_{k+1} < 2^{-k}.$$

Die stetigen Funktionen f_k bilden somit eine Fundamental (Cauchy-) Folge im vollständigen Raum C(X). Also existiert $f = \lim f_k$ in C(X) und erfüllt wegen $d(f_k(x), F(x)) < \varepsilon_k$ auch $f(x) \in \mathrm{cl}\, F(x)\ \forall x$. Der Satz gilt also zumindest dann, wenn die Bilder abgeschlossen sind.

Teil 2

Konstruktion von stetigem g, so dass $g(x) \in \mathrm{relint}\, F(x) \subset F(x)\ \forall x \in X$.

Def. relint M (relatives Innere) für konvexe Mengen M: Sei U_M der kleinste affine Teilraum, der M enthält, also $U_M = \{ u \mid \exists N, \exists r_k \in R, \exists y_k \in M: u = \Sigma r_k y_k, \Sigma r_k = 1, k = 1,\ldots,N \}$. Für $M \subset R^m$ ist U_M abgeschlossen, und es reicht $N \le m+1$ zu wählen; vergleiche mit conv M (Caratheodory).

Man definiert: $y \in \mathrm{relint}\, M$, falls $B(y, \varepsilon) \cap U_M \subset M$ für ein $\varepsilon > 0$ gilt. In der Folge benutzen wir

1. falls $y \in \mathrm{relint}\, M$, $y' \in \mathrm{cl}\, M$ und $\lambda \in (0, 1)$, so ist $z(\lambda) = \lambda y + (1-\lambda) y' \in \mathrm{relint}\, M$ (man wähle für $z(\lambda)$ jeweils $0 < \varepsilon(\lambda) < \lambda \varepsilon$)
2. relint M = relint (cl M).

Erfülle eine stetige Funktion f zunächst $f(x) \in \mathrm{cl}\, F(x)\ \forall x$.

Mit

$$G(x) = F(x) \cap B(f(x), 1) \qquad \text{(die } \textit{offene} \text{ Kugel um f(x))}$$

ist G wieder uhs, und alle Bilder sind beschränkt. Sei $Y_0 = \{y_1, y_2, \ldots\}$ abzählbar und dicht in $Y = R^m$.

Wir bilden mit *offenen* Kugeln $B(y_p, 1/k)$

(*) $$X_{p,k} = \{ x \in X \mid \mathrm{cl}\, G(x) \cap B(y_p, 1/k) \neq \varnothing \}.$$

Da G uhs ist (somit auch cl G) , sind dann alle $X_{p,k}$ offen.

Damit kann jedes $X_{p,k}$ auch als *abzählbare Vereinigung abgeschlossener Mengen* geschrieben werden: $X_{p,k} = \cup_q A_{p,k,q}$ (q = 1,2,3, ….);

z.B. so: Man nimmt eine abzb. dichte Teilmenge W von $X_{p,k}$, betrachtet alle abgeschl. Kugeln mit rationalem Radius um $w \in W$, die ganz in $X_{p,k}$ liegen, und vereinigt sie alle.

Nun definiert man

$$G_{p,k,q}(x) = \mathrm{cl}\, G(x) \qquad \text{falls } x \in X \setminus A_{p,k,q}$$

und

$$G_{p,k,q}(x) = \mathrm{cl}\, G(x) \cap B(y_p, 1/k) \qquad \text{falls } x \in A_{p,k,q}.$$

Weil alle Mengen $A_{p,k,q}$ abgeschlossen sind und $G_{p,k,q}$ auf $A_{p,k,q}$ eine uhs Teilabbildung von cl G ist, wird $G_{p,k,q}$ insgesamt uhs auf X.

Damit gibt es (wie schon gezeigt) stetige $f_{p,k,q}$ mit

$$f_{p,k,q}(x) \in \mathrm{cl}\, G_{p,k,q}(x)\ \forall x \in X.$$

Dichtheit der Bilder. Wir zeigen: Die abzählbare Menge a(x) aller Punkte $f_{p,k,q}(x)$ ist dicht in cl G(x).

Sei $y \in \mathrm{cl}\, G(x)$ und $\varepsilon > 0$.

Dann gibt es ein k mit $1/k < \varepsilon$ und ein p mit $y \in B(y_p, 1/k)$. Damit folgt nach Definition $x \in X_{p,k}$, womit $x \in A_{p,k,q}$ für ein q gilt.

Mit diesem q folgt weiter, dass auch

$$f_{p,k,q}(x) \in \mathrm{cl}\, G_{p,k,q}(x) = \mathrm{cl}(\mathrm{cl}\, G(x) \cap B(y_p, 1/k))$$

richtig ist. Also muss gelten

$$\| f_{p,k,q}(x) - y \| \le \| f_{p,k,q}(x) - y_p \| + \| y_p - y \| \le 2/k < 2\varepsilon.$$

Das bedeutet : *Die abzählbare Menge a(x) aller Punkte $f_{p,k,q}(x)$ ist dicht in cl G(x).*

Schließlich lassen sich alle $f_{p,k,q}(x)$ umnummerieren in $g_1(x), \ldots g_\nu(x), \ldots$. Bildet man nun mit den *beschränkten und stetigen* Funktionen g_ν

$$g(x) = \Sigma_\nu\, 2^{-\nu} g_\nu(x),$$

so wird auch g stetig.

Weil (für jedes x wegen der gezeigten Dichtheit) wenigstens ein $g_\nu(x)$ aus relint G(x) ist, besitzt auch g(x) diese Eigenschaft, denn es gilt

$$g(x) = 2^{-\nu} g_\nu(x) + \Sigma_{\mu \neq \nu} 2^{-\mu} g_\mu(x) = 2^{-\nu} g_\nu(x) + (1-2^{-\nu}) \Sigma_{\mu \neq \nu} (1-2^{-\nu})^{-1} 2^{-\mu} g_\mu(x)$$

wobei $\Sigma_{\mu \neq \nu} (1-2^{-\nu})^{-1} 2^{-\mu} g_\mu(x) \in \mathrm{cl}\, G(x)$ ist.

Man erhält so die gesuchte Funktion: $g(x) \in \mathrm{relint\, cl}\, G(x) = \mathrm{relint}\, G(x) \subset \mathrm{relint}\, F(x)$. ♦

Folgerung:

Sei F eine mehrwertige, uhs Abbildung, definiert auf einer konvexen, kompakten Menge $\emptyset \neq S \subset R^n$ mit $\emptyset \neq F(x)$ konvex und $F(x) \subset S \ \forall x \in S$. Dann besitzt F einen Fixpunkt.
Beweis: Man wende Brouwer's FPSatz auf eine stetige selection von F an. ♦

Besonders interessant ist dies, wenn x stets ein Randpunkt von F(x) ist
(wie etwa unten für die „Dominanzmengen" und ihren konvexen Hüllen in einem Markt).
Mit einer selection der Form $f(x) \in$ relint F(x) und jedem Fixpunkt x* von f folgt dann
$x^* = f(x^*) \in$ relint $F(x^*)$. Das impliziert zusätzlich $F(x^*) = \{x^*\}$
(und bedeutet dann, dass x* nicht durch andere Elemente dominiert werden kann).

Kooperative Spiele und ökonomisches Gleichgewicht

Spiele mit Kooperation
Sei v die charakteristische Funktion eines n-Personen Spiels (I, v). Hierbei ist I = {1...n} eine endliche Menge (von Spielern) und v(S) ist definiert für alle Teilmengen S von I: alle v(S) sind reell, $v(\emptyset) = 0$, man deutet v(S) als garantierten Mindestgewinn für die *Koalition* S, wenn die Spieler aus S vernunftig kooperieren.
Mit dieser Interpretation folgt, dass v superadditiv ist:

$$v(S \cup T) \geq v(s) + v(T) \ \text{ falls } \ S \cap T = \emptyset.$$

Es bleibt das Problem (und darum entbrennt der Streit) den bei Kooperation aller Spieler möglichen Gewinn v(I) „geeignet" aufzuteilen. Die möglichen Aufteilungen, die für jeden Spieler interessant sind, werden durch die Menge

$$X = \{ x \in R^n \mid \Sigma_{i \in I} x_i = v(I), \quad x_i \geq v(\{i\}) \ \forall i \}$$

gegeben. Statt v({ i}) wird oft einfacher v(i) geschrieben.
Wir setzen (o.B.d.A.) zur Vereinfachung voraus, dass $v(i) = 0 \ \forall i$, wonach auch $v(S) \geq 0 \ \forall S$ folgt.
Ein Spiel heisst wesentlich (nichttrivial), wenn v(I) > 0.

Dominanz:
Sind x und y aus X, so sagt man y dominiert x bzgl. S, kurz $x <_S y$ wenn gilt:

$$y_i > x_i \ \forall i \in S \quad \text{und} \quad \Sigma_{i \in S} \ y_i \leq v(S).$$

Die zweite Bedingung sagt, dass die für Spieler aus S bessere Aufteilung nicht "utopisch ist", denn sie könnten die Gewinsumme tatsächlich durch Kooperation erreichen.
Man sagt y *dominiert* x, wenn $x <_S y$ für wenigsten eine Koalition S gilt.
Interpretation:
Wird der Gewinn der einzelnen Spieler durch die Wahl von Strategien und Gewinnfunktionen $f_k(s_1, \dots s_n)$ geregelt, wird v(K) der größte Gesamtgewinn, den sich die Spieler aus K durch Kooperation sichern können (auch wenn sich die restlichen Spieler der Komplementärkoalition S = I \ K gegen die Koalition verbünden); damit ist v(K) ein max–min-Ausdruck.

$$v(K) = \max_{a \in \sigma(K)} \min_{b \in \sigma(S)} \Sigma_{k \in K} f_k(a, b),$$

wenn a, b die Strategien der entsprechenden Koaltionen sind; σ(K) = Menge "der Strategien" der Koalition K. Wir benutzen diese Darstellung in der Folge nicht, sondern nur die Eigenschaft der Superadditivität.
Die folgenden Modelle haben ein Sinn, wenn mögliche Gewinne der Spieler aufgeteilt und übertragen werden können (Spiele mit Seitenzahlungen)

18 von Neumann-Morgenstern Lösung im klassischen Kooperativspiel

Def. Eine Teilmenge $M \subset X$ heisst NM- Lösung, wenn gilt
(1) die Elemente aus M dominieren sich nicht gegenseitig und
(2) zu jedem $x \in X \setminus M$ gibt es ein $y \in M$, das x dominiert.

Dieser Lösungsbegriff hat seine Eigenarten.

Wir vergleichen zunächst mit den Lösungen einer Extremalaufgabe $\min \{ f(x) \mid x \in X \}$. Man definiere in X „Dominanz“ durch „y dominiert x , wenn $f(y) < f(x)$“.
Eine NM-Lösung M beschreibt dann offenbar gerade die Menge der globalen Minimalpunkte von f und ist daher eindeutig bestimmt.

Wir betrachten nun das folgende 3-Personenspiel
Beispiel 1: $v(S) = 0$, wenn S aus nur einem Element besteht (oder $S = \emptyset$ ist), ansonsten sei $v(S) = 1$.

Die „gerechte“ Verteilung

$x = (1/3, 1/3, 1/3)$ wird bzgl. $S = \{1,2\}$

offenbar dominiert durch

$y = (½ , ½, 0)$,

aber auch durch

$y = (p, 1-p, 0)$, $1/3 < p < 2/3$.

Umgekehrt werden solche y nicht von x dominiert, welche Koalition man auch betrachtet. Deshalb ist $M = \{ (1/3, 1/3, 1/3) \}$ sicher keine NM-Lösung.

Die Mengen $\{y \mid y \text{ dominiert } x\}$ und $\{ y \mid y \text{ wird von } x \text{ dominiert} \}$ kann man und sollte man sich gut am Einheitssimplex verdeutlichen !

Dann sieht man am einfachsten, dass die aus 3 Elementen bestehende Menge

$M = \{ (½ , ½, 0), (½ , 0, ½), (0, ½ , ½) \}$

eine NM-Lösung ist.
Aber sie ist nicht die einzige. Man betrachte etwa

$M_o = \{ (p, 1-p, 0) \mid 0 \le p \le 1 \}$ (und analog die Mengen mit 0 an anderer Stelle)

oder

$M_\varepsilon = \{ (p, 1-p, \varepsilon) \mid 0 \le p \le 1-\varepsilon \}$ (und analog die Mengen mit ε an anderer Stelle)

Wie gross darf ε im letzten Fall werden damit M_ε eine NM-Lösung bleibt ?

Insgesamt erhält man aus der Diskussion dieser Fälle:

NM-Lösungen sind auch in einfachen Spieler nicht eindeutig bestimmt und allgemein nicht leicht zu finden.
Lange hat man wenigstens die Existenz nachweisen wollen. Dann wurde ein Spiel ($n = 10$) gefunden, das tatsächlich keine NM-Lösung besitzt.

Zum Problem, NM-Lösungen zu finden, haben v. Neumann und Morgenstern allerdings für kleine n (< 5) schon erhebliche Arbeit in ihrer Monographie geleistet.

Das härteste Problem scheint allerdings in einer (für Anwendungen) akzeptablen Interpretation dieser Lösungen zu liegen. Deshalb gibt es auch zahlreiche andere Arten von Lösungsbegriffen.

19 Core und balancierte Spiele

Definition des „cores“ C : Das ist die Menge derjenigen x aus X, die durch kein y (bezüglich irgendeiner Koalition S) dominiert werden.

Damit erhält man $x \in C \iff$

(C0) $\Sigma_{i \in S}\, x_i \ \ge\ v(S)$ für alle S und $x \in X$.

Hierbei durchläuft S alle Koalitionen (Teilmengen von I) , die leere Menge und I ausgenommen.
Allerdings: Die Menge C kann leer sein.
Man nehme obiges Beispiel mit $n =3$, $v(S) = 0$ wenn $\text{card } S \le 1$, $v(S) = 1$ wenn $\text{card } S > 1$.

Also braucht man zusätzliche Bedingungen.
Die folgenden Überlegungen basieren im Wesentlichen auf dem Dualitätssatz der linearen Optimierung.

Def. Ein System σ von Teilmengen S von I heißt **balanciert**, wenn es positive y(S) ($S \in \sigma$) derart gibt dass $\Sigma_{S:\, i \in S}\ y(S) = 1$ für alle $i \in I$.

Def. Gilt die Ungleichung

(*) $\quad \boldsymbol{\Sigma_{S \in \sigma}\ v(S)\, y(S) \ \leq\ v(I)}$

für *jedes* balancierte Mengensystem σ, wobei über S aus σ summiert wird, so heißt das Spiel balanciert. Damit gilt der

Satz 1: Der core C ist nicht leer gdw. das Spiel balanciert ist.

Beweis:
Wir denken uns zum linearen System (C0) die triviale Zielfunktion
min $\Sigma\, 0\, x_i$ ($x \in R^n$) und fassen alles als LO-Aufgabe (P) auf. Sie läßt sich dualisieren.
Die Dualaufgabe hat die Variablen y(S) und y(I) und ergibt sich als
(D) max $v(I)\, y(I) + \Sigma\ v(S)\, y(S)$
wobei Restriktionen durch jeden Index $i \in I$ definiert werden:
$y(I) + \Sigma_{S:\, i \in S}\ y(S) = 0$ für alle i , $y(S) \geq 0$, y(I) beliebig.
Der *Dualitätssatz der linearen Optimierung* besagt insbesondere, dass Primal- und Dualaufgabe stets gleichzeitig lösbar (bzw. unlösbar) sind.
Damit folgt: C ist nicht leer gdw. (D) lösbar.
Nun sieht man:
(D) hat einen (trivialen) zulässigen Punkt (y = 0).
Der Restriktionsbereich ist ein Kegel. Damit ist (D) lösbar gdw. die duale ZF stets ≤ 0 bleibt, für alle zulässigen y.
D.h. $-\ y(I)\, v(I) \ \geq\ \Sigma\ v(S)\, y(S)$ (offenbar ist dann $t := -y(I) \geq 0$ s. Restr.)
Sei $B(i, y) = \Sigma_{S:\, i \in S}\ y(S)$.
Dann erhalten wir:
(D) ist lösbar gdw.
Für alle y(.) mit $y(S) \geq 0$ und $B(i , y) = t\ \forall i$ gilt $\Sigma\ v(S)\, y(S) \leq t\, v(I)$.
Für t = 0 ist dies trivial erfüllt. Also ist nur der Fall t > 0 interessant.
In diesem Fall ist die Bedingung genau dann erfüllt, wenn sie für t = 1 gilt.
Letzteres bedeutet (nach Definition) aber gerade, dass das Spiel balanciert ist. ♦

Interpretation:
Wir interpretieren y(S) als den *Beitrag*, den ein Mitglied der Koalition *S für seine Mitgliedschaft* zu entrichten hat. Dann ist B(i , y) die von i zu zahlende totale Beitragssumme.
Die Restriktionen (mit t = 1) verlangen, dass die von jedem Spieler zu zahlende Beitragssumme konstant 1 ist. In diesem Fall nennen wir den Beitragsvektor y zulässig.
Die Bedingung dafür, dass C nicht leer ist, besagt somit: Für jedes System y von zulässigen Beiträgen ist (*) erfüllt.
Bedingung (*) verlangt, dass der Gewinn, der maximal verteilt werden kann – nämlich v(I) - „hinreichend groß" ist.

20 Shapley- Vektor (Existenz / Eindeutigkeit / konkrete Formel)

Sei (I, v) ein superadditives Kooperativspiel . D.h., wie schon erklärt, I ist die Spielermenge 1...n; v ist auf der Potenzmenge 2^I von I reellwertig definiert und erfüllt
$v(K \cup L) \geq v(K) + v(L)$ falls $K, L \subset I$ und $K \cap L = \emptyset$, $v(\emptyset) = 0$ per Definition.
Sind u und v charakteristische Funktionen und $\alpha, \beta \geq 0$, so ist auch $\alpha\, u + \beta\, v$ eine charakteristische Funktion.
Ein Spieler i heißt *Strohmann*, wenn stets $v(S \cup \{ i \}) = v(S) + v(i)$ gilt, sofern $i \notin S$.
Hierbei steht v(i) abkürzend für v({i}).

Problem: Wie soll $v(I)$ auf die n einzelnen Spieler „gerecht“ aufgeteilt werden ?
Die folgenden Bedingungen sind Shapleys "Axiome der Gerechtigkeit".

Gesucht wird eine Funktion Φ, so dass $\Phi = \Phi(v) \in \mathbb{R}^n$ (das soll die „gerechte“ Aufteilung des Gewinns sein) und

(1) $\Phi_i(v) \geq v(i) \ \forall i$ und $\Phi_i(v) = v(i)$ für jeden Strohmann i.

(2) $\Sigma_i \Phi_i(v) = v(I)$

(3) $\Phi(u+v) = \Phi(u) + \Phi(v)$

(4) Wenn $\pi : I \leftrightarrow I$ (Permutation) und $v(\pi K) = v(K) \ \forall K$, so $\Phi_i(v) = \Phi_{\pi i}(v)$.

Hierbei steht πK für die Menge $\{\pi i \,/\, i \in K\}$.

Satz (L.S. Shapley A value for n-person games. Annals of Mathematics Studies 28 (1953), 307-317)
Es gibt genau eine Funktion Φ mit den obigen Eigenschaften. Sie ordnet jedem v den Vektor $\Phi(v)$ mit den Komponenten

(5) $\Phi_i(v) = \Sigma_{S \ni i}\ r(s)\ [\, v(S) - v(S \setminus \{i\})\,]$

zu, wobei $r(s) = (n-s)!\ (s-1)!\ / n!$ und $s = \text{card } S$. Summiert wird über alle $S \subset I$ mit $i \in S$.

Beweis:

1. Vorbetrachtungen

Es reicht, Spiele mit $v(i) = 0$ zu betrachten, da man ansonsten $u(K) = -\Sigma_{i \in K} v(i)$ definieren und zum Spiel $v' = v + u$ übergehen könnte. Wir machen aber davon nicht Gebrauch.
Angenommen Φ existiert entsprechend den Axiomen.
Seien u', u'' und v charakteristische Funktionen mit $v = u'' - u'$.
Dann gilt $u'' = v + u'$ und deshalb wegen (3), $\Phi(u'') = \Phi(v) + \Phi(u')$, also auch

$$\Phi(v) = \Phi(u'') - \Phi(u').$$

Sei weiter $v = \Sigma_k c_k u_k$ ein Spiel, das sich als Linearkombination von Spielen u_k schreiben läßt.

Spaltet man auf $K^+ = \{ k \,/\, c_k > 0 \}$, $K^- = \{ k \,/\, c_k < 0 \}$, so gilt

$$v = \Sigma_{k\in K^+}\ c_k u_k - \Sigma_{k\in K^-}\ (-c_k u_k) = u'' - u'.$$

Jede Summe u'', u' ist eine charakteristische Funktion.
Aus (3) folgt $\Phi(u'') = \Sigma_{k\in K^+}\ \Phi(c_k u_k)$, $\Phi(u') = \Sigma_{k\in K^-}\ \Phi(-c_k u_k)$ und somit auch

$$\Phi(v) = \Sigma_{k\in K^+}\ \Phi(c_k u_k) - \Sigma_{k\in K^-}\ \Phi(-c_k u_k).$$

Haben weiter alle u_k die Eigenschaft, dass

$$\Phi(\lambda u_k) = \lambda \Phi(u_k) \quad \text{für alle } \lambda \geq 0 \quad \text{(positive Homogenität von } \Phi \text{ für } u_k)$$

gilt, erhalten wir so

$$\Phi(v) = \Sigma_{k\in K^+}\ c_k \Phi(u_k) + \Sigma_{k\in K^-}\ c_k \Phi(u_k).$$

2. Darstellung von v und Eindeutigkeit von Φ

Wir kommen nun zu einer speziellen Darstellung von v. Für festes $S \subset I$ sei $u = u_S$ die charakteristische Funktion

$$u_S(K) = 1 \text{ falls } S \subset K \ ; \quad u_S(K) = 0 \text{ sonst.}$$

u_S zeigt an, dass eine Koalition K genau dann gewinnt (nämlich die Summe 1), wenn alle Spieler aus S beteiligt sind.
Nach (1), (2), (4) ist $\Phi(u_S)$ eindeutig bestimmt durch

$$\Phi_i(u_S) = 1/s \text{ falls } i \in S, \quad \text{wobei } s = \text{card } S$$

$$\Phi_i(u_S) = 0 \quad \text{sonst.}$$

Dieselben Axiome liefern analog:

$$\Phi(\lambda u_S) = \lambda \Phi(u_S) \quad \text{für alle } \lambda \geq 0. \quad \text{(positive Homogenität von } \Phi \text{ für } u_S)$$

Wir wollen nun v als Summe der Form

(*) $v = \Sigma_S\ \lambda(S)\, u_S$ mit reellen Koeffizienten $\lambda(S)$, $\emptyset \neq S \subset I$

schreiben. Nach den Vorbetrachtungen folgt dann

$$\Phi(v) = \Sigma_S\ \lambda(S)\, \Phi(u_S) \qquad \text{(auch wenn gewisse } \lambda(S) \text{ negativ sind).}$$

Wenn wir zeigen, dass Koeffizienten $\lambda(S)$ in (*) stets eindeutig existieren (abhängig von gegeb. v), muß demnach auch $\Phi(v)$ eindeutig bestimmt sein.

Eindeutigkeit:

Bedingung (*) liefert eine Gleichung für jede Koalition K, $\emptyset \neq K \subset I$, nämlich

$(*)_K$ $\quad v(K) = \Sigma_S \; \lambda(S)\, u_S(K)$.

Hier sind alle S mit $u_S(K) \neq 0$ interessant, also alle S mit $S \subset K$ und damit $u_S(K) = 1$.

$(*)_K$ bedeutet somit

(6) $\quad v(K) = \Sigma_S \; \lambda(S) + \lambda(K)$ $\quad$ wobei $S \subset K$ und $S \neq K$.

Liegen schon alle $\lambda(S)$ mit $\emptyset \neq S \subset K$ und $S \neq K$ fest, so ist $\lambda(K)$ eindeutig bestimmt als

(7) $\quad \lambda(K) = v(K) - \Sigma_S \; \lambda(S)$ $\quad$ wobei $S \subset K$ und $S \neq K$.

Für Koalitionen K mit card K = 1 ist Eindeutigkeit klar: $\lambda(K) = v(K)$.

Aus der Eindeutigkeit aller $\lambda(S)$ für card $S < p$ folgt mit (7) aber die Eindeutigkeit für card $S = p$.

Existenz:

Es ist schon klar, dass sich (6) für card K = 1 erfüllen läßt, dazu muß nur $\lambda(i) = v(i)$ sein.

Angenommen, alle $\lambda(S)$ sind für card $S = s < p$ bereits so definiert, dass alle Gleichungen (6) mit card $K < p$ erfüllt werden.

Wir betrachten (6) für beliebiges festes K mit card K = p. Das heißt mit $S \subset K$ und $S \neq K$,

(7) $\quad \lambda(K) = v(K) - \Sigma_{S \subset K,\, S \neq K} \; \lambda(S)$.

Dies läßt sich zur korrekten Definition von $\lambda(K)$ verwenden, weil die rechte Seite nur von den Werten $\lambda(S)$, card $S < p$ abhängt, und daher kein Widerspruch zur Definition von $\lambda(K')$ für die restlichen K' mit card K' = p entsteht.

Damit sind Existenz und Eindeutigkeit der Darstellung (*), und somit die **Eindeutigkeit von Φ** klar.

3. Existenz von Φ .

Es bleibt zu zeigen, dass die Funktion Φ aus (5) die geforderten Eigenschaften besitzt.

Dazu machen wir uns klar, was $\Phi_i(v) = \Sigma_{S \ni i} \; r(s) \; [\, v(S) - v(S \setminus \{i\}) \,]$ bedeutet.

Mit einer festen Reihenfolge, d.h. einer Pemutation π aller Spieler, denken wir uns, dass die Spieler ein Zimmer betreten: Zuerst Spieler $\pi 1$, dann $\pi 2$, usw. Nach Schritt s sind die Spieler der Menge S $= \{\pi 1, ..., \pi s\}$ versammelt, und Spieler $i = \pi s$ möge den Betrag

$$d_i(\pi) = v(S) - v(S \setminus \{i\})$$

erhalten, also den Gewinnzuwachs der "Zimmerkoalition" durch Beitritt von $i = \pi s$.

Damit wird ein Vektor $d(\pi) \in R^n$ definiert. Er erfüllt offenbar

(8) $\quad d_i(\pi) \geq v(i)$ $\qquad$ wegen Superadditivität, wobei $d_i(\pi) = v(i)$ für jeden Strohmann,

(9) $\quad \Sigma_i \; d_i(\pi) = \Sigma_s [\, v(\{\pi 1, ..., \pi s\}) - v(\{\pi 1, ..., \pi s\} \setminus \{\pi s\}) \,] = v(I) - v(\emptyset) = v(I)$.

Betrachtet man $d(\pi)$ auch als Funktion von v, $d(\pi) = d(\pi, v)$, so ist $d(\pi, .)$ additiv:

(10) $\quad d(\pi, v+u) = d(\pi, v) + d(\pi, u)$.

Sei nun $\phi = \phi(v)$ das arithmetische Mittel aller n! Vektoren $d(\pi)$, $\phi = (1/n!)\, \Sigma_\pi \, d(\pi)$.

Die Eigenschaften (8), (9), (10) übertragen sich elementar auf ϕ. Damit erfüllt $\phi = \phi(v)$ die Axiome (1), (2) und (3).

Man erhält weiter

$$n!\, \phi_i(v) = \Sigma_\pi \; d_i(\pi) = \Sigma_{\pi:\ \pi(s) = i} \; [\, v(\{\pi 1, ..., \pi s\}) - v(\{\pi 1, ..., \pi s\} \setminus \{\pi s\}) \,] .$$

Dieselbe Differenz $[\, v(S) - v(S \setminus \{i\}) \,]$ tritt hier so oft auf wie es Permutationen π gibt mit $\{\pi 1, ..., \pi s\} = S$ und $\pi(s) = i$.

Permutiert man die s-1 Elemente von $S \setminus \{i\}$ sowie die letzten n-s Elemente von $I \setminus S$ unabhängig voneinander, sind das gerade $(s-1)!\ (n-s)!$ Möglichkeiten.

Das zeigt $\phi(v) = \Phi(v)$. Schließlich zeigt die Darstellung (5), dass auch (4) erfüllt ist. Damit hat $\Phi(v)$ alle geforderten Eigenschaften. ♦

Man beachte, dass wir es uns erspart haben, die häßlichen Lösungen $\lambda(S) = \Sigma_{T \subset S} \; (-1)^{s-t}\, v(T)$ von (*) explizit betrachten zu müssen! Man vergleiche etwa mit [RaScZa79]. Zur Theorie kooperativer Spiele; siehe auch [Ros71].

21 Drohungen und Gegendrohungen (Auman / Maschler); siehe z.B. Owen (1971)

Wir kehren zu einem Kooperativspiel (I, v) mit superadditiver charakteristischer Funktion v und übertragbarem Gewinn zurück, und betrachten eine Koalitionsstruktur $\kappa = \{K_r\}$ von n Spielern, d. h. $K_1, \dots K_{m1}$ bilden eine Zerlegung der Spielermenge I. Dazu betrachten wir einen Punkt $x \in R^n$ (Gewinnverteilung), so dass für alle K_r gilt

$$x(K_r) := \Sigma_{i \in Kr}\, x_i = v(K_r).$$

Die Menge dieser x sei $X(\kappa)$, und ein Paar $(x, \{K_r\})$ heiße ***Konfiguration*.**

In der Folge werden wir noch voraussetzen, dass $x_i \geq v(i)$ für alle i gilt. Derartige Konfigurationen heißen auch individuell rational (i. r.). Ist dagegen sogar $x(S) \geq v(S)$ für jede in einer Koalition K_r enthaltenen Spielermenge S, so spricht man von koalitionsrationalen Konfigurationen (k.r.) . Letztere müssen nicht notwendig existieren, wie im Falle von nur einer Koalition $K_1 = I$ zu sehen ist (core kann leer sein, Satz 1).

Wir definieren jetzt Drohungen und Gegendrohungen gegen diese Drohungen .
Beides hat sowohl für ind. als auch koalit. - rationale Konfigurationen Sinn.
Allerdings wird der entsprechende Existenzsatz unten nur für einen Spezialfall bewiesen, *(wenn S und T aus genau einem Element bestehen; siehe (1)' usw),* so dass die allgemeineren Definitionen hier nicht weiter gebraucht werden.

Zunächst brauchen wir allerdings den Begriff des *Partners.*
Sei S eine Teilmenge von I. Wir nennen dann die Menge $\cup K_r$, so dass K_r die Menge S schneidet, Menge der Partner von S in der Koalitionsstruktur κ. Formal $P(S, \kappa)$.
Besteht S nur aus Spieler s, so ist das einfach diejenige Menge K_r, welche s enthält.
Seien nun S und T zwei disjunkte, nichtleere Teilmengen derselben Koalition K_{r0}. Wir betrachten eine neue Konfiguration $(y, \sigma) = (y, \{S_r\})$ (mit nicht notwendig ebenso vielen Koalitionen) .
Diese heißt
Drohung von S gegen T, sofern

(1) kein $t \in T$ zur Partnermenge $P(S, \sigma)$ gehört ,
(2) $y_i \geq x_i$ für alle $i \in P(S, \sigma)$
(3) $y_i > x_i$ für alle $i \in S$ gilt.

Eine dritte Konfiguration $(z, \tau) = (z, \{T_r\})$ heißt **Gegendrohung** von T gegen die obige Drohung , sofern

(4) S keine Teilmenge von $P(T, \tau)$ ist ,
(5) $z_i \geq x_i$ für alle $i \in P(T, \tau)$
(6) $z_i \geq y_i$ für alle $i \in P(T, \tau) \cap P(S, \sigma)$ gilt.

Aus der Definition folgen einige simple , aber wichtige Bemerkungen .
1. Eine Drohung von S gegen T ist eine Drohung gegen jeden Spieler t aus T.
2. Spieler $t \in T$ findet u. U. leichter eine Gegendrohung als die ganze Koalition T .
3. Hat T gegen eine Drohung von S keine Gegendrohung , so bleibt das auch nach hinreichend kleinen Änderungen von $x \in X(\kappa)$ mit derselben Koalitionsstruktur κ richtig . (Der nötige Spielraum für die entsprechende Änderung von y ergibt sich aus Ungleichung (3)) .
4. Angenommen S hat mit (y, σ) eine nicht parierbare Drohung gegen einen Spieler t . Dann hat jeder Spieler $s \in S$ eine nicht parierbare Drohung gegen t.

Die 4. Aussage sieht man so: s faßt alle Partner $P(S, \sigma)$ zu einer Koalition S´ zusammen und läßt die übrigen Koalitionen von σ unverändert . Damit entsteht eine Koalitionsstruktur σ'. Allen Spielern gibt er zunächst den Gewinn y_i. Da v superadditiv ist, kann die Summe der y_i über i

$\in S'$ höchstens kleiner als $v(S')$ sein. In diesem Fall wird der Überschuß irgendwie auf alle Partner aufgeteilt.
Da $P(S, \sigma) = S' = P(S', \sigma')$ ist, kann t auch gegen diese Drohung von s nichts ausrichten .

Eine Konfiguration heißt **stabil**, wenn es zu jeder Drohung eine Gegendrohung gibt. Sei M die Menge dieser Konfigurationen. In der obigen Allgemeinheit läßt sich noch nichts über die Existenz stabiler Konfigurationen sagen. Selbst wenn man annmmt, dass es k.r. Konfigurationen gibt , und nur solche betrachtet, ist dies unklar. Die Frage bleibt auch noch offen, wenn man nur indiv. rat. Konfigurationen zuläßt. Eine positive Aussage kann für den Fall indiv. rat. Konf. gemacht werden, unter der Einschränkung , dass im Falle einer Drohung von S gegen einen einzigen Spieler t dieser Spieler stets eine Gegendrohung besitzen soll ; also nur ein-elementige T betrachtet werden. Die entsprechende stabile Menge M wird mit $M^i{}_1$ bezeichnet.

Satz: Zu jeder Koalitionsstruktur κ gibt es ein $x \in X(\kappa)$, so dass (x, κ) eine individuell rationale stabile Konfiguration in bezug auf Gegendrohungen *einzelner* Spieler ist, d.h. zu $M^i{}_1$ gehört .
Beweis: Wir dürfen voraussetzen, dass $v(i) = 0$ für alle i .
Nach Bemerkung 4 müssen wir nur den Fall betrachten , dass die Drohkoalitionen ebenfalls nur aus einem Spieler s bestehen. (Um zu vereinfachen, kann man gleich den Fall untersuchen, dass jeweils nur ein Spieler droht bzw. „gegendroht“)
Wir halten (x, κ) fest und schreiben $s >> t$ falls s eine nicht parierbare Drohung gegen t besitzt. Das heißt
(1)´ t ist nicht in $P(s, \sigma)$,
(2)´ $y_i \geq x_i$ für alle $i \in P(s, \sigma)$ und
(3)´ $y_s > x_s$,
und es gibt keine dritte Konfiguration $(z, \tau) = (z, \{T_r\})$ mit den Eigenschaften
(4)´ s nicht aus $P(t, \tau)$,
(5)´ $z_i \geq x_i$ für alle $i \in P(t, \tau)$
(6)´ $z_i \geq y_i$ für alle $i \in P(t, \tau) \cap P(s, \sigma)$.
Diese Bedingungen betreffen nur die Partner von s in σ, also die Koalition S , in der sich dann s, aber nicht t befindet , und die Koalition T von τ , die t aber nicht s enthält.
Also ist

$$s >> t$$

genau dann, wenn es eine Koalition S und dazu ein y_S (mit Komponenten y_i, $i \in S$) gibt , dass
(1)´´ $\Sigma\, y_S = v(S)$
(2)´ $y_i \geq x_i$ für alle $i \in S$ und (3)´ $y_s > x_s$ gilt
und kein Paar (z_T, T) mit
(4)´´ $\Sigma\, z_T = v(T)$
(5)´ $z_i \geq x_i$ für alle $i \in T$ und (6)´ $z_i \geq y_i$ für alle $i \in T \cap S$
existiert.

t ist stark, bzgl. x wenn fuer alle S ohne t und y
(1)´´ $\Sigma\, y_S = v(S)$
(2)´ $y_i \geq x_i$ für alle $i \in S$ und $y_s \geq x_s$ (Quasidrohung) gilt
dass $(z_T, T=T(S, x))$ mit s nicht in T exist, so dass
(4)´´ $\Sigma\, z_T = v(T)$
(5)´ $z_i \geq x_i$ für alle $i \in T$ und $z_i \geq y_i$ für alle $i \in T \cap S$ (t ist nicht dabei) gilt.
Trivialfall x_t = 0.
Sonst hat man ein positives x_t , das ist später wichtig.

Y_t (x) ist abgeschlossen: Wenn t bedroht werden kann, sind alle 2. T-Ungl. Systems in z unlösbar. Die Menge dieser x (und dazu passender Drohungen y) ist offen. Anders herum: die Menge aller x, so dass t schwach ist, ist offen, die Kompl.Menge (stark) Y_t (x) ist abgeschlossen (Vereinigung von Polyedern).

Die erste Gruppe von (Droh-) Bedingungen verlangt offenbar, dass die Zahl $e(S, x) = v(S) - \Sigma x_S$ positiv ist. Sie heißt auch Exzess von S bzgl. x .

Wir zeigen als erstes , dass die **Relation >> azyklisch** ist .

Dazu nehmen wir s(1) >> ... >> s(m) >> s(1) an. Alle s(k) liegen dann in derselben Koalition K_r und benutzen entsprechende Drohkoalitionen S(k) gegen den nächsten s(k+1) [m +1 := 1] . Möge e(S(k) , x) für k = k0 **maximal** sein.
Der durch seinen Vorgänger s = (k0-1) bedrohte Spieler t könnte dann mittels seiner eigenen Drohkoalition **T = S(k0)** eine Gegendrohung finden, **sofern** s = (k0-1) nicht in S(k0) enthalten ist. Die Existenz einer Gegendrohung war aber ausgeschlossen, also (k0-1) $\in$ T. Damit könnte aber auch (k0 -1) mittels **T = S(k0)** eine Gegendrohung finden, wenn nicht (k0 - 2) $\in$ **T** ist , usw. Schließlich erhält man so, dass alle diese Spieler in T sind, wonach aber S(k0) keine Koalition sein kann, mit der k0 gegen k0+1 droht. Also ist die Relation azyklisch.

Damit gibt es stets einen Spieler i (x, r) $\in K_r$, der nicht schwächer als die übrigen ist , der also immer eine Gegendrohung findet . Wir fixieren nun ein $x \in X(\boldsymbol{\kappa})$ und eine Koalition K_r , und zerlegen x in der Form $x = (x(r), x_N(r))$, wobei die Komponenten von x(r) zu K_r gehören, die von $x_N(r)$ zum "Restvektor" mit Indizes außerhalb K_r .

Die Mengen Y_i (x) und Überdeckung von X(r) : Den Vektor x(r) variieren wir im Simplex X(r) der erlaubten Gewinnverteilungen der Koalition K_r , betrachten also y(r) ≥ 0, so dass wieder

$\Sigma y_i(r) = v(K_r)$ ist .

Für jedes i definieren wir nun

Y_i (x) als die Menge derjenigen y(r) , so dass in $(y(r), x_N(r))$ Spieler i nicht schwach ist
(jede Drohung parieren kann, abgeschl. Menge; neues y, nicht das aus der Drohung)
Kompl. offen, also Y_i (.) abgeschl.

Hierbei ist $r = r(i)$ so zu wählen , dass $i \in K_r$.
Da die Relation >> azyklisch ist, ist bei jeder Wahl von y(r) wenigstens ein Spieler $i \in K_r$ nicht schwach, also

$\cup_{i \in Kr} Y_i(x) = X(r)$.

Zu Y_i (x) gehört offenbar die durch $y_i = 0$ beschriebene Seite des Simplex X(r) (Wer nichts zu verlieren hat, kann jede Drohung - durch die Einzelkoalition - parieren). Die Mengen Y_i (x) und ebenso die Abbildung, die x gerade Y_i (x) zuordnet, sind abgeschlossen (d.h. ihr graph ist eine abgeschl. Menge im entsprechenden Produktraum; vergl. Bemerkung 3, Komplementärmenge is offen) .

Wir brauchen dass die Abbildung Y_i unterhalb stetig ist.

Angenommen Spieler t ist stark in (x, kappa) und x_t > 0. Wir bilden x(r)_eps, indem sich i den Gewinn x_t – eps nimmt und den Rest auf die uebrigen Komponenten von x(r) verteilt. Der Rest $x_N(r)$ bleibe zunächst fest. Jede Drohung (y, sigma) eines Spielers s gegen t in der neuen Konfiguration ist auch eine Drohung gegen t in der alten Konfiguration. Sie konnte mittels der Gegendrohung (z, tau) beantwortet werden. Einziger Unterschied gegenüber vorher ist jetzt die Bedingung

(5)´ $z_i \geq x_i$ (neu) für alle $i \in T \setminus S$ falls i in K_r liegt und nun größer ist.

Für i = t ist x_i (neu) = x_i – eps. Fuer i <> t ist x_i (neu) = x_i + eps / (| K_r | - 1) .

Da s nicht in T liegt, müssen höchstens (| K_r | - 2) Spieler aus T mehr erhalten (nicht der Spieler s).

Das geht mit verändertem z(neu), wenn man z(neu)_t = x_t - eps = x_t (neu) setzt und die übrigen Komponenten auf x_i (neu) erhöht, $i \in T \setminus S$, i in K_r .

Dann bleibt ein Überschuss, nämlich mindestens delta = eps / (| K_r | - 1). Diesen kann t auf alle Spieler aus T gleichmäßig verteilen und damit auch wieder die Summenbedingung erfüllen.
Damit hat man noch immer eine Gegendrohung, wenn die Komponenten aus x_rest wenig verändert sind: z.B. keine mehr als um delta / n erhöht wird.
Folgerung: Die Abb. Y_i(x) ist unterhalb stetig.

Sei $D_r(x)$ der (abgeschl.) Durchschnitt aller $Y_i(x)$ bzgl. $i \in K_r$.
Wir müssen zeigen, dass ein $x \in X(\kappa)$ existiert mit
(7) $x(r) \in D_r(x)$ ***für alle r .***
Nach dem "Arbeitslemma" ist $D_r(x)$ nicht leer. Hätten wir nur ein r, wäre das schon die Behauptung.

Die folgenden Überlegungen beweisen (nebenbei) auch das "Arbeitslemma" .
Angenommen, (7) ist falsch für alle x. Dann betrachten wir zu gegebenem x ein festes r = r(x) mit
$$x(r) \notin D_r(x) = \cap_{i \in K_r} Y_i(x)$$
und bilden für $i \in K_r$ den Abstand $d_{ir}(x)$ zwischen x (r) und $Y_i(x)$.
Ist er positiv, so gilt $x_i(r) > 0$, denn wer nichts hat, ist stark. Nach Wahl von r ist auch
$$c_r(x) := \Sigma_{i \in K_r} d_i(x) \quad \text{positiv.}$$
Die Funktion h^r mit
$$h^r{}_i(x) = v(K_r)\ d_i(x) \ / \ c_r(x) \quad , \quad i \in K_r$$
ist deshalb wohldefiniert und bildet x in den Rand bd X(r) ab, weil wenigstens ein $d_i(x)$ Null ist.
Wir bilden nun $f^r(x)$ wieder durch Spiegelung von $h^r(x)$ am Zentrum z^r des Simplex´ X(r), d. h.
$$f^r(x) = z^r + t \ (z^r - h^r(x)) \quad \text{mit maximalem } t\text{, so dass noch}$$
$$z^r + t \ (z^r - h^r(x)) \in X(r) .$$
Die resultierende Abbildung ist wohldefiniert und stetig in x solange nur
$x(r) \notin D_r(x)$ richtig bleibt (also auf einer offenen Menge).
In diesem Falle ist auch $f^r(x) \neq x(r)$ denn:
Für $x(r) \in$ int X(r) ist das offensichtlich weil $f^r(x)$ im Rand liegt .
Für $x(r) \notin$ int X(r), ist eine Komponente $x_i(r)$ Null , also auch die Komponente $h^r{}_i(x)$ und somit *nicht* die gespiegelte Komponente $f^r{}_i(x)$.

Wir betrachten schließlich die Abbildung F(x), die jeweils die konvexe Hülle von x(r) und $f^r(x)$ zuordnet, wenn $x(r) \notin D_r(x)$, und ansonsten durch den Punkt x(r) allein definiert ist. Sie ist uhs und bildet die Menge der zulässigen x in sich ab.
Mit einer stetigen Auswahlfunktion g = g(x) aus relint F(x) erhalten wir einen Fixpunkt x=g(x) aus relint F(x). Dieser muss aber x(r) in D(r) für alle r erfüllen,

was den Satz beweist. ♦

Ergänzung: Es geht auch anders, aber mit mehr Aufwand: Die Menge $D_r(x)$ hängt nur von $x_N(r)$ ab. Ist $t \in (0, 1)$, bleibt auch $t \ f^r(x) + (1 - t) \ x(r) \neq x(r)$, wenn dist (x(r) , $D_r(x)$) > 0. Wir setzen t speziell als stetig in x an: t = t (x , r) = dist (x(r) , $D_r(x)$) an . Dazu muss man aber noch extra beweisen, dass auch dieser Abstand stetig ist, worauf wir hier verzichten. Der Abstand sei dabei so festgelegt, dass er Werte < 1 liefert, damit die Punkte zulässig bleiben. [Man nehme etwa den Euklidischen Abstand d und setze dist = d / (1 + d)]. Wir definieren schließlich $f^r(x) = 0$ falls $x(r) \in D_r(x)$ ist, und setzen g (x) aus den Komponenten der Funktionen

$$g_r(x) = t(x,r)\, f^r(x) + (1 - t(x,r))\, x(r)$$

zusammen. Sie ist stetig, weil $x(r) \in D_r(x)$ stets $t = 0$ impliziert, die Unstetigkeit von f^r auf dem Rand von $D_r(x)$ also verschwindet. So bildet g die Menge $X(\kappa)$ stetig in sich ab und hat keinen Fixpunkt, sofern (7) falsch ist. Also muß (7) für ein x richtig sein, was zu beweisen war .

22 Edgeworth Markt

auch (Tausch-) MARKT und CORE (exchange-economy), Francis Ysidro Edgeworth (1845-1926)

Ein Markt M (auch economy, Tauschmarkt, Ökonomie) ist durch folgende Bestandteile gegeben:
1. Eine Menge I von n Elementen (Spieler, Agenten, Konsumenten)
2. n Vektoren $x0(i) \in R^m$, $i \in I$, jeder besitze nichtnegative Komponenten; kurz $x0(i) \in R+^m$.
3. n Relationen <=i , die für Paare der Menge $R+^m$
   erklärt sein sollen; und zwar transitiv, reflexiv, vollständig und stetig.

Wir lesen

$$\alpha \;\; >=i \;\; \beta$$

als " α ist für i mindestens so gut wie β" , schreiben

$$\alpha \;\; >i \;\; \beta \quad \text{(lies } \alpha \text{ besser als } \beta \text{ für i),}$$

wenn $\alpha >=i \beta$, nicht aber $\beta >=i \alpha$ gilt.
"Stetigkeit" der Relation bedeutet per Definition, dass die Mengen

$$\Gamma_i(\beta) := \{\alpha \;/\; \alpha >=i \beta\} \qquad \text{und} \qquad \Gamma_i^-(\alpha) := \{\beta \;/\; \alpha >=i \beta\}$$

stets abgeschlossen sind und wird hier generell vorausgesetzt.

Konvexität des Marktes bedeutet per Definition, dass die Mengen $\Gamma_i(\beta)$ konvex sind.
Monotonie bedeutet dass aus $\alpha \in \Gamma_i(\beta)$ und $\alpha'_k \geq \alpha_k$ für alle Komponenten zu Spieler i folgt, dass auch $\alpha' \in \Gamma_i(\beta)$ gilt (mehr zu haben ist nicht schädlich).

Bemerkungen:

1. Gewöhlich unterstellt man, dass die Relationen durch stetige Funktionen ui(.) von $R+^m$ in R mittels $\alpha >i \beta$ gdw. $ui(\alpha) > ui(\beta)$ repräsentiert werden können. Die geforderten Eigenschaften sind dann trivialerweise erfüllt.
Sind *konvexe und monotone Präferenzrelationen* gegeben, kann man auf einfache Weise (theoretisch) Nutzensfunktionen definieren , die diese vollständig beschreiben in der Form

$$\alpha \;\; >i \;\; \beta \qquad \text{genau dann , wenn} \qquad u_i(\alpha) > u_i(\beta) \; .$$

Man definiert $u_i(\alpha)$ als das kleinste aller λ , so dass der Punkt $\lambda(1, \ldots, 1)$ in der Menge $\Gamma_i(\alpha)$ liegt . Die Funktion ist wegen der Monotonie der Vorzugsrelation wohldefiniert, wegen der Stetigkeit der Abbildung Γ_i auch stetig und wegen der Konvexität der Mengen $\Gamma_i(\alpha)$ quasikonkav, d. h.

$$u_i(t\alpha + (1-t)\beta) \geq \min\{u_i(\alpha), u_i(\beta)\} \quad \text{für alle } t \in (0, 1)$$

Nachrechnen ! (Hinten ausführlicher)
Ist der Graph von Γ_i sogar konvex , graph $\Gamma_i = \{ (\beta, \alpha) \mid \beta \in \Gamma_i(\alpha) \}$, so sind die Funktionen $u_i(\,.\,)$ sämlich konkav. Die Bestimmung möglichst adäquater Nutzensfunktionen, die sich andererseits numerisch in obiger Aufgabe verwerten lassen , ist eines Hauptprobleme für die effektive Anwendung all dieser Marktmodelle.

2. Vektoren α, β aus $R+^m$ werden als "Güterbündel" für den jeweiligen Spieler aufgefaßt. x0(i) ist das Bündel, das den Anfangsbesitz von i repräsentiert.
3. Eine Umverteilung (allocation, redistribution) ist ein Element x der Menge

$$\mathbf{X = \{ x = (x(1),...,x(n)) \;/\; x(i) \in R+^m ,\;\; \Sigma\, x(i) = \Sigma\, x0(i) \}.} \qquad \textbf{x ist also eine Matrix.}$$

Der Core C(M).
Der core des Marktes M, bezeichnet mit C(M), ist jetzt die Menge all der Umverteilungen x, die für jede Koalition S von Spielern folgendes erfüllen:

$(1)_S$ Es *gibt kein* $y \in X$ mit $y(s) >_s x(s)$ für alle $s \in S$
[später auch beschrieben durch $y(s) \in \text{int}\,\Gamma_s(x(s)) = \Omega_s(x(s))$]
und
$\Sigma_{s \in S}\, y(s) = \Sigma_{s \in S}\, x0(s)$; keine "Dominanz" bzgl. S.

Die Menge **DS(x)** dieser dominierenden y ist stets eine konvexe Teilmenge von X. Ihre konvexe Hülle mit dem Element x ebenfalls.
D(x) sei die konvexe Hülle aller DS(x) und x. Die Abb. ist uhs nach Vorauss. für Γ_s .
Mit reichlich Aufwand zeigt man, dass x aus **C(M)** gleichbedeutend mit **D(x) ={x} ist.**

Vorgehensweise:
Man kann zuerst zeigen, dass C(M) nicht leer ist.
Danach:
(x,p) =(x(p),p) ist Walras GGW gdw. wenn x^q in jedem Zwillingsspiel ein Edgeworth GGW ist.
Dazu braucht man:

1 Aus Walras GGW folgt Edgeworth GGW.
Hat man schon die Existenz eines Walras GGW, folgt also auch **C(M) $\neq \emptyset$.**

Ferner (Zusammenhang) :

2 Ist x in X zu keinem Preis ein Walras GGW,
so ist x auch kein Edgeworth GGW in einem gewissen Zwillingsmarkt M^q.
3 Ist y aus C(M^q), so kann man umverteilen, so dass auch ein Element x^q in C(M^q) liegt.
4 (trivial) Ist x^{q+1}) in C(M^{q+1}) so ist auch x^q in C(M^q) .

Damit wird die Menge X* aller x, die zu einem Walras GGW gehören, gerade der Durchschnitt aller symmetrischen Lösungen X[q], q nat. Zahl.

Wir zeigen den
Existenzsatz: Für konvexe monotone Märkte M gilt C(M) $\neq \emptyset$.
Beweis: Unser Problem lautet jetzt:
Wir suchen x mit D(x) = { x } vergl. AbschlussBemerkung zu E. Michael.
Nach Michaels selection theorem [Mich56] gibt es (wegen der Konvexitäts- und Stetigkeits-annahmen; speziell weil D eine sogenannte unterhalb stetige mehrd. Abbildung wird) eine
stetige Funktion f mit $f(x) \in$ relint D(x) für alle x.
Sie hat (Brouwer) einen Fixpunkt x*. Damit folgt $x^* = f(x^*) \in \text{rel int } D(x^*)$ und aus der Definition von D: $D(x^*) = \{ x^* \}$. Das ist gerade unsere Behauptung. ♦

Man braucht die folgenden spezielleren Sachen nicht. Die restlichen Vektoren $y(i), i \in I \setminus S$ würden sich entsprechend definieren lassen, so dass die Gesamtsumme gerade $\Sigma_I\, x0(i)$ wird, nämlich yi = x0(i). Man kann stets y(i) = x0(i) für die restlichen i setzen. (Es ist nun $x \in \text{bd}\, DS(x)$!)
Sei
$DS(x) = \{ y \in X \mid y$ dominiert x bzgl. S und $y(i) = x0(i)$ für $s \notin S \} \cup \{ x \}$
Kann man o.B.d.A. sichern, dass $y(i) = x(i)$ für $i \notin S$?
Ja, denn Dominanz bzgl. S stellt dazu keine Bedingung. Wir setzen die y(i) einfach so an. Wenn Dominanz überhaupt vorliegt, dann auch eine solche ! *Summenbed. erfüllt, alle Spieler sind einverstanden, die Spieler $\notin S$ auch, da sie nichts verlieren.*
$y \in DS(x) \Rightarrow y(k) = x(k)\ k \notin S\ ,\ \Sigma_S\, y(k) = \Sigma_S\, x(k) \qquad \Rightarrow y \in X.$
$y \in \text{conv}(DS(x)\ ,\ DT(x))$ d.h. $y = \lambda_1 y^1 + \lambda_2 y^2$ u.s.w. … $\qquad \Rightarrow y \in X.$
Sei nun D(x) die konvexe Hülle aller DS(x) . D(x) = conv $\cup$ S DS(x). Offenbar ist $x \in D(x)$.

q- linge:
Im folgenden Abschnitt benötigen wir das ***q-fache Produkt M^q des Ausgangsmarktes M .***
Dazu sei q eine natürliche Zahl. Ausgehend von Markt M kann man einen neuen bilden mit genau qn Spielern und jeweils q Spielern des Types i. Sie mögen jeweils dasselbe Anfangsbündel x0(i) und dieselben Präferenzen wie Spieler i in M besitzen. also quasi **"q- linge"** sein. (Die q- ling Idee stammt –wahrscheinlich- von W. Hildenbrand)

Der Existenzsatz gilt offenbar auch für M^q .

Für jedes Element $x \in X$ ist dann auch das Element $(x)^q$, welches jedem Spieler des Typs i gerade x(i) zuordnet, eine Umverteilung in M^q . Es gibt natürlich auch andere Umverteilungen, z.B könnte sich ein q-ling alles nehmen und seinen Brüdern nichts abgeben. Das wird uns aber nicht interessieren. Zunächst betrachten wir weiter den Originalmarkt M.

23 Walras Gleichgewicht *Léon Walras (1834-1910)*

Wir betrachten den Markt M unter der weiteren Annahme, dass den m Gütern positive Preise $p_1,...,p_m$ zugeordnet werden können. Ist ein Preisvektor p (o.B.d.A. $\Sigma_\kappa p_k = 1$, $p_k > 0$) fixiert, so definiert er über

$$B(i, p) := \{ \alpha \in R^{+m} / \langle \alpha , p \rangle \leq \langle x0(i) , p \rangle \}$$

die sogenannte *Budget-Menge* für Spieler i. Das sind die Güterbündel, die i entsprechend seinem Anfangsbesitz und Preis kaufen kann. Offenbar ist i an solchen $x(i) \in B(i, p)$ interessiert, für die kein besseres

(2) $\alpha >i \; x(i)$, $\alpha \in B(i,p)$ im Sinne seiner Nutzensrelation existiert („maximale" x(i)).

Die resultierende Matrix x muss dann allerdings nicht notwendig zu X gehören, wenn die Preise schlecht gewählt wurden. Ist

(3) $x \in X$,

so heißt das Paar (x, p) ***Walras-Gleichgewicht*** (competitive- equilibrium, Preis-Gleichgew., Konkurrenz-Gleichgew.).

Die Menge dieser Punkte bezeichnen wir hier mit **W(M),** die zugehörigen x mögen die Menge **X*** bilden.

Angenommen <=i wird über Nutzensfunktionen u_i realisiert. Bedingung (2) heißt dann einfach

(2)' $x(i) \in \arg\max \{ ui(\alpha) / \alpha \in B(i, p) \}$ (d.h. x(i) soll ui maximieren bzgl. B(i,p)).

Nehmen wir weiter an, es gebe für jedes p eindeutige Lösungen x(i, p). Dann lässt sich ihnen der "Defekt" $d(x, p) = | \langle \Sigma x(i, p) - \Sigma x0(i) , p \rangle |$

zuordnen, der im Gleichgewichtsfall gerade Null sein muß.

1. Zusammenhang zum Edgeworth-Markt

In jedem Falle gilt

Lemma 1: Ist (x, p) aus W(M), so ist x aus C(M), also $X^* \subset C(M)$.

Beweis: **Sei (x, p) aus W(M).**

Wenn ***x nicht aus C(M),*** gäbe es eine Koalition S und y mit

$y(s) >s \; x(s)$ $(\forall s \in S)$ und $\Sigma y(s) = \Sigma x0(s)$.

Wegen der ersten Bedingung und $(x, p) \in W(M)$ muss aber für den Wert der Warenbündel gelten:

$\langle y(s), p \rangle > \langle x0(s), p \rangle \; \forall s$ (y(s) ist zu teuer für alle Spieler aus S) ,

denn sonst „maximiert" x nicht den Nutzen von s über der Budgetmenge, d.h., es wäre x nicht maximal bzgl. (2) (Mit Relation oder Nutzensfunktion).

Also gilt auch

$\Sigma\_s \langle y(s), p \rangle > \Sigma\_s \langle x0(s), p \rangle$, $\langle \Sigma \; y(s) – x0(s) , p \rangle > 0$.

Andererseits besteht jedoch Gleichheit wegen der Dominanzgleichung $\Sigma y(s) = \Sigma x0(s)$. ♦

Bemerkung:

Analog folgt aus $(x, p) \in W(M)$ die core-Eigenschaft für $(x)^q$ im q-fachen Produkt M^q .

Preise als Stützfunktionale

Wir fragen nun umgekehrt, wann ein Element x des cores C(M) mit einem geeigneten p zu einem Walras - GGW wird.

Seien dazu die Relationen >=i sämtlich *monoton* in dem Sinne, dass aus $\beta \geq 0$, $\beta \neq 0$ jeweils $\alpha + \beta >i \; \alpha$ folgt
(irgendetwas mehr und von allem nicht weniger ist immer besser).

Wir setzen außerdem voraus , dass die Mengen $\Omega i (x(i)) := \{ \beta / \beta >i \; x(i) \}$

(4) cl Ωi (x(i)) = Γ_i (x(i))
erfüllen.

Kommentar: Es ist x $\in$ C(M), wenn x $\in$ X und wenn folgendes nicht möglich ist: y(i) $\in$ Ωi (x(i)) $\forall$ i $\in$ S *und* Σ_ i $\in$ S y(i) <= Σ_ i $\in$ S x0(i).

Also ist x $\in$ C(M), wenn aus y(i) $\in$ Ωi (x(i)) $\forall$i $\in$ S *folgt dass* Σ_S y(i)_j0 > Σ_S x0(i)_j0 für wenigstens ein Gut j0 gilt.
Das ist mit der Budget-Bedingung bei „schlechten Preisen" möglich.
Die Budget-Bedingung verlangt: Σ_j p_j y(i)_j <= Σ_j p_j x0(i)_j . Sie kann auch dann erfüllt sein, wenn alle i $\in$ S mit y(i) insgesamt zu viel von Gut j0 verlangen, nämlich wenn p_j0 zu klein ist. Also ist x $\in$ C(M) sicher nicht hinreichend dafür, dass x mit jedem Preis eine Walras Lösung wird.
Die „zu kleine Preise" p_j0 können sich ändern, wenn man andere Koalitionen S betrachtet. Deshalb ist die Frage, ob es überhaupt einen passenden Walras Preis zu x $\in$ C(M) gibt, nicht trivial. Alle Preise zu erhöhen, geht nicht wegen Σ p_j = 1.

Satz 2: Die Relationen >i seien monoton und mögen (4) erfüllen.
Ein Element x $\in$ X (insbesondere also ein core-Element) ist mit einem gewissen positiven Preisvektor p ein Walras GGW **dann und nur dann,** wenn 0 nicht innerer Punkt der konvexen Hülle conv ($\cup_i$ M(i, x)) der (Besser-) Mengen M(i , x) = Γ_i (x(i)) - x0(i) ist.

Beweis:
Wir bemerken zunächst, dass x mit einem beliebigen Vektor p $\neq$ 0 nur dann die Bedingungen für ein Walras Gleichgewicht erfüllen kann, wenn jede Komponente von p positiv ist. Dies ist eine Folge der Monotonievoraussetzung an die Vorzugsrelationen.
Wegen *x $\in$ X* brauchen wir nur noch ein p $\neq$ 0, so dass mit jedem i gilt:
Alle β aus der (offenen) Besser-Menge Ωi (x(i)) = { β / β >i x(i) } müssen im Halbraum
$\langle$ β, p $\rangle$ > $\langle$ x0(i), p $\rangle$ liegen.
Nach (4) bedeutet das mit der Budget-Gleichung: p ist Walras Preis genau dann, wenn gilt:
(*) β $\in$ Γi (x(i)) = cl Ωi (x(i)) liegt stets im Halbraum H(p, i) = { β / $\langle$ β , p $\rangle$ $\geq$ $\langle$ x0(i) , p $\rangle$ }.

(*) ist also eine *notwendige und hinreichende Bedingung* dafür, dass ein Element x $\in$ X zusammen mit einem Vektor p (der dann notwendigerweise positiv ist) ein Walras GGW wird. Transformiert man mittels der um x0 verschobenen Besser-Mengen
M(i, x) = Γ_i (x(i)) - x0(i) (sie liegen im R^m),
ist (*) äquivalent zu
0 $\leq$ min (M(i, x), p) für alle i $\in$ I.
Aufgrund des Trennungssatzes existiert ein solches p $\neq$ 0 dann und nur dann, wenn 0 nicht aus dem Innern der konvexen Hülle der Vereinigung aller Mengen M(i, x) ist. Das war zu zeigen. ♦

Also muß neben der Existenz von core-Elementen untersucht werden, was
0 $\in$ int conv ($\cup_i$ M(i, x)) .
bedeutet.
Dazu nehmen wir an: Die Mengen Γ_i (α) seinen konvex , die Relationen monoton.
Solche Märkte werden auch *konvex und monoton* genannt.

Wegen der vorausgesetzten Monotonie ist dann int M(i, x) = Ωi (x(i)) - x(i) nicht leer, was sofort auch (4) liefert.

Satz 3 (Walras GGW und q-linge): *Für konvexe, monotone Märkte M gilt:*
Eine Umverteilung x bildet mit einem gewissen positiven Preisvektor p ein Walras-Gleichgewicht genau dann, wenn mit jedem natürlichen q ihr q-faches Produkt $(x)^q$ zum core $C(M^q)$ gehört.
Beweis:
Möge x mit keinem p zu W(M) gehören. Nach Satz 2 gibt es dann endlich viele Elemente in der Menge

$\cup_i$ (int M(i , x)), sagen wir $r^1,.., r^N$, so dass schon

$0 \in \mathrm{conv}\{ r^\mu \mid \mu = 1, \dots , N\}$.

Es gibt also Zahlen w_μ mit $w_\mu > 0$, $\Sigma\, w_\mu = 1$, $0 = \Sigma\, w_\mu r^\mu$.

Sei

$M(j ,x) = \Gamma_j (x(j)) - x0(j)$,

eine solche Menge, deren Inneres gewisse r^μ enthält; die Indexmenge dieser j nennen wir J. Wir bilden „Mittelwerte" der Güterbündel in M(j ,x) mit Hilfe der Koeffizienten w_μ :

$y(j) = \Sigma_\mu\, w_\mu r^\mu / s_j$ mit $s_j = \Sigma\, w_\mu$, summiert über μ mit $r^\mu \in$ int M(j , x).

Dann ist auch y(j) aus int M(j , x) und

(**) $0 = \Sigma_J\, s_j y(j)$, $\Sigma s_j = 1$, $s_j > 0$, $y(j) \in \Omega j\, (x(j)) - x0(j)$

bzw. $x0(j) + y(j) \in \Omega i\, (x(j))$.

Wir bemerken nun, dass sich durch beliebig kleine Änderungen der Punkte y(j) in Gleichung (**) erreichen läßt, dass (**) mit *rationalen* s_j gilt.
Da die Mengen Ωi offen sind, ändern kleine solche Änderungen nichts an den Präferenzbeziehungen.
Also dürfen wir unterstellen, dass alle s_j *in (**) rational* sind. Dann lassen sie sich als

$s_j = p_j / q$ (index j in J)

mit einheitlichem q und natürlichem p_j schreiben. (Jetzt ist p kein Preisvektor !)

An dieser Stelle betrachten wir das q-fache Produkt des Ausgangsmarktes M und bilden dort die

Koalition S, die aus jeweils p_j Spielern des Typs $j \in J$ besteht.

Ihnen ordnen wir das neue, für jeden Spieler bessere Bündel x0(j) + y(j) zu.
Die Summe aller neuen Bündel von S wird so

$$\boldsymbol{\Sigma\, p_j\, (y(j) + x0(j)) = q\, \Sigma\, s_j (y(j) + x0(j)) = 0 + q\, \Sigma\, s_j\, x0(j) = \Sigma\, p_j\, x0(j) .}$$

Die Summe ist also gleich der Summe der Ausgangsbündel dieser Koalition.
Damit ist das speziell konstruierte q-fache Produkt der Umverteilung x (die mit keinem p zu W(M) gehört) ***kein*** core- Element im speziellen q-fachen Produkt des Ausgangsmarktes.

Umgekehrt muß für $(x, p) \in W(M)$ das q-fache Produkt von x auch zu $C(M^q)$ gehören (s.Beweis von Lemma 1). Wir haben daher alles gezeigt. ♦

Um die Menge X* aller $x \in X$, die mit geeignetem p aus W(M) sind, zu beschreiben, definieren wir

$\mathbf{X(q) = \{ x \in X \mid (x)^q \in C(M^q) \}}$.

Satz 3 besagt offenbar gerade dass $X^* = \cap_q X(q)$ ist.

Da sich mit wachsendem q die Koalitionsmöglichkeiten der Spieler in M^q nur vergrößern, ist

$X(q+1) \subset X(q)$.

Die Mengen X(q) sind nach Def. des cores wegen der Stetigkeitsbedingung für die Präferenzen abgeschlossen und beschränkt. Sie bilden daher ein *zentriertes Mengensystem,* wonach ihr Durchschnitt nicht leer ist, sofern jede Menge X(q) nicht leer ist.
Also folgt:

$X^* \neq \varnothing$, wenn $X(q) \neq \varnothing$ für all q.

Man beachte, dass X(q) eine echte Teilmenge des cores $C(M^q)$ sein kann, da in $(x)^q$ alle q-linge des Typs i dasselbe bekommen.
Mit dem folgenden Lemma und dem Exisenzsatz für Edgeworth GGW, *$C(M) \neq \varnothing$* ist der Existenzbeweis **$X^* \neq \varnothing$** abgeschlossen.

Lemma 2: *Sei M ein konvexer, monotoner Markt . Dann ist $X(q) \neq \varnothing$, falls $C(M^q) \neq \varnothing$.*
Beweis:

Sei $x \in C(M^q)$. Wir halten irgendein i fest. Den Spielern i1,...,iq des Typs i seien die Güterbündel

(5) $x(i1) \;<=_i\; x(i2) \;<=_i\; \ldots \;<=_i\; x(iq)$

zugeordnet. Wir bilden nun (den Mittelwert)
$\alpha i = (x(i1) + \ldots + x(iq)) / q$, setzen $z(i1) = \ldots = z(iq) = \alpha i$ sowie $z(.) = x(.)$ für alle übrigen Spieler.
Da x zum core $C(M^q)$ gehört ist , können wir zunächst zeigen , dass in (5) jeweils $=_i$ gelten muß. Dazu betrachten wir die n Spieler mit kleinstem x(i1).
Gilt fuer ein gewisses fixiertes i, etwa i = k nicht die Gleichheit, hätten wir wegen Monotonie und Konvexität der Präferenzen

$$x(k1) <_k \alpha k$$

und für die restlichen Spieler sicher noch

$$x(i1) <=_i \alpha i .$$

Im Bündel αk besitzt Spieler k1 von wenigstens einem Gut j einen positiven Anteil. Indem er davon einen kleinen Teil gleichmäßig auf die übrigen Mitglieder der

Koalition $S = \{ i1 \;/\; i = 1 ,\ldots , n \}$

verteilt, entsteht eine Umverteilung z', die x bezüglich S dominiert. Das widerspricht der Voraussetzung $x \in C(M^q)$. Also gilt in (5) tatsächlich die Gleichheit.
Damit werden aber x und z stets gleichzeitig dominiert.
Folglich hat auch z mit symmetrischer Güterverteilung der Form (x)^q die core- Eigenschaft, was zu beweisen war. ♦

2. Existenz von Walras Gleichgewichtspunkten (mittels Brouwer's Fixpunktsatz).
Der folgende Beweis für die Existenz eines Walras GGW ist kürzer als oben, vereinfacht jedoch ein wenig und stellt nicht den obigen Zusammenhang zum Edgeworth Gleichgewicht her.
Wir betrachten einen *konvexen* , *monotonen* Markt. Die Präferenzrelationen der n Teilnehmer seien durch "Nutzensfunktionen" u_i für die einzelnen Spieler gegeben , so dass

$$u_i (\alpha) < u_i (\beta)$$

gerade bedeutet, dass i das Güterbündel β dem Bündel α vorzieht. Ist p ein beliebiger Preisvektor, so sind offenbar die n Aufgaben

A(i , p) $\max_\alpha \{ \; u_i(\alpha) \;|\; \langle p , \; \alpha - x^0(i) \rangle \le 0 \}$

zu untersuchen.
Wir suchen ein solches p, dass gewisse Lösungen $x^i(p)$ [die die Ungl. sicher als Gleichung erfüllen wegen Monotonie der u_i] mit dem Gesamtangebot des Marktes kompatibel sind, d.h.,

$\Sigma_i \, x^i(p) \le c$ mit

$c = \Sigma_i \, x^0(i) \in R^m$

erfüllen.
Angenommen, die Aufgaben sind eindeutig lösbar und ihre Lösungen sind stetig in p
Hierzu gibt es zahlreiche Bedingungen aus der parametrischen Optimierung, die aber jetzt nicht weiter verfolgt werden. Es reicht z.B. vorauszusetzen, dass (1) alle u_i streng konkav sind und (2) a priori für die Menge aller „sinnvollen Umverteilungen" vorauszusetzen, dass sie in einer hinreichend grossen Kugel liegen.
Beweis der Existenz eines passenden Preises:
Indirekt:
Angenommen für alle p mit $\Sigma_j \, p_j = 1$, $p_j \ge 0$ (kurz $p \in S$) erfüllen die Lösungen $x^i(p)$ von A(i,p)

$$w_j (p) := \max\{ 0 \;, \; \Sigma_i \, x^i(p)_j \; - \; c_j \; \} > 0 \text{ für wenigstens ein (Gut) } j \in \{1, 2,\ldots, m\}.$$

Wir setzen $s(p) = \Sigma \, w_j(p) > 0$ und bilden zu p einen neuen Preisvektor p' durch

$$p'_j = (p_j + w_j(p)) / (1 + s(p)).$$

Er liegt wieder in S und die Abbildung $p \to p'$ ist stetig, hat also einen Fixpunkt p.
Der Fixpunkt erfüllt

$$p'_j = p_j = (p_j + w_j(p)) / (1 + s(p)) \quad \forall j ,$$

also

$$p_j \; (1 + s(p)) = p_j + w_j (p) \quad \forall j$$

bzw. $\qquad p_j\, s(p) = w_j(p) \qquad \forall\, j$

Wenn $p_j = 0$, so wird Gut j sicher nicht ausreichen, also $w_j > 0$ sein. Mit $s(p) > 0$ ist dies nicht möglich. Wir erhalten daher sowohl $p_j > 0$ als auch

(*) $\quad \Sigma_i\, x^i(p)_j - c_j = w_j(p) = p_j\, s(p) > 0 \quad \forall\, j \qquad$ (Kein Gut reicht aus), $\quad c = \Sigma_i\, x^0(i) \in R^m$

Andererseits sichern die Nebenbedingungen der Aufgaben A(i, p):

$$\Sigma_j\ p_j\ (x^i(p)_j - x^0(i)_j) \le 0 \quad \forall\, i$$ Summe über die Güter j , schlecht lesbar

und folglich, nach Summe über i,

(**) $\quad \Sigma_i\ (\ \Sigma_j\ p_j\ (x^i(p)_j - x^0(i)_j)\ ) \le 0.$

Die linke Seite lässt sich umschreiben als

$$\Sigma_i\ \Sigma_j\ p_j\,(x^i(p)_j - x^0(i)_j) = \Sigma_j\ \Sigma_i\ p_j\,(x^i(p)_j - x^0(i)_j)$$
$$= \Sigma_j\ p_j\ \Sigma_i\ (x^i(p)_j - x^0(i)_j) = \Sigma_j\ p_j\ (\ \Sigma_i\, x^i(p)_j - c_j\ ).$$

Nach (*) ist der letzte Ausdruck positiv im Widerspruch zu (**). Folglich muss ein Walras-Preisvektor existieren. ♦

Bemerkungen zur Berechnung

Sei y_i ein zu A(i , p) gehörende Lagrange - Multiplikator (mit Dimension 1). Er exist. zu jeder Lösung, weil die Nebenbedingungen in α affin-linear sind.

Dann muss gelten

$$Du_i(x^i(p)) + y_i\ p = 0 \in R^m,\ \langle p,\ x^i(p) - x^0(i) \rangle = 0 \ \text{ und } \ \Sigma_i\ x^i(p) \le c^0 \in R^m.$$

Die beiden ersten Bedingungen besagen gerade (für konkave, differenzierbare u_i sogar äquivalent), dass $x^i(p)$ die Aufgabe A(i, p) löst.

Mit neuen Variablen $z^i = x^i(p) - x^0(i)$ und $v_i(z^i) = u_i(\ x^0(i) + z^i\ )$ wird dies für gesuchtes (z, y, p) zu

$$\mathbf{Dv_i(z^i) + y_i\ p = 0 \in R^m,\quad \langle p,\ z^i \rangle = 0,\quad \Sigma_i\, z^i \le 0 \in R^m \qquad (i = 1, \dots, n).}$$

Das ist ein (i.a. nichtlineares)

Komplementaritäts-Problem wegen der Komplementaritäts-Bedingung $\langle p , z^i \rangle = 0$.

Bei m Gütern ist dies ein System mit $n\,m + n + m$ reellen Variablen in (x, y, p) und ebenso vielen Gleichungen. Sind die u_i quadratisch (einfachster Typ praktisch relevanter Nutzensfunktionen), also Du_i linear, so hat das System seine einfachste Gestalt.

24 Vorzugsrelation und Nutzensfunktion

Auf $X = R^n$ sei eine mehrdeutige Abbildung $F: X \to X$ gegeben.

Die Bilder F(x) seien *konvexe und abgeschlossene* Teilmengen von X.

Interpretation: $y \in F(x) \iff y$ ist besser als x oder genauso gut. Die Relation sei

reflexiv, $\quad x \in F(x)$,

transitiv, $\quad y \in F(x) \Rightarrow F(y) \subset F(x)$ $\qquad$ und

antisymmetrisch in dem Sinne, dass $y \in \operatorname{int} F(x) \Rightarrow x \notin F(y)$.

Außerdem sei sie *monoton*: $y \in F(x) \Rightarrow y + h \in F(x)$, falls $h \ge 0$. Die Klasse der Abbildungen F mit obigen Eigenschaften heiße Φ.

Beispiel: Vektor-Halbordnung, $y \in F(x) \iff y \ge x$.

Mit jedem $h > 0$ folgt so $y \in F(x) \Rightarrow y + h \in \operatorname{int} F(x)$ denn $h' > 0$ für kleine $\| h' - h \|$.

Gesucht wird: Eine Funktion $u: X \to R$ mit

(1) $\qquad y \in F(x) \iff u(y) \ge u(x).$

u heißt Nutzensfunktion zu F.

Notwendige Bedingung:

Es existiere eine Nutzensfunktion u zu F. Für beliebige Elemente x, y gilt dann $y \in F(x)$ oder $x \in F(y)$, d.h. die Relation F muß ***vollständig*** sein (alles ist uneigentlich vergleichbar). Damit scheidet die Vektor-Halbordnung F aus !

Wegen Antisymmetrie hat $y \in \operatorname{int} F(x)$ zur Folge, dass $x \notin F(y)$, also gilt nach (1):

$y \in \text{int } F(x) \;\; \Rightarrow \;\; u(y) > u(x)$.
Außerdem impliziert (1) dass $u(x) = u(y)$ gerade $y \in F(x)$ und $x \in F(y)$ bedeutet. In diesem Falle muß also $y \in \text{bd } F(x)$ und $x \in \text{bd } F(y)$ gelten, d.h. x und y sind äquivalent im Sinne der "besser" - Relation F.
Wünschenswert wäre, dass auch die Umkehrung
(1)' $\quad y \in \text{bd } F(x) \Rightarrow u(x) = u(y)$
gilt. Diese Forderung ist nicht hart.
Jede stetige Nutzensfunktion u für F erfüllt (1)' :
$y \in \text{bd } F(x) \Rightarrow$ Es gibt eine konvergent Folge $y' \to y$ mit $y' \notin F(x) \Rightarrow u(y') < u(x)$.
Weil auch $u(y) \geq u(x)$ und u stetig in y ist, kann nur $u(y) = u(x)$ gelten. ♦

Gibt es eine Nutzensfunktion u für F, die (1)' erfüllt (z.B. eine stetige Nutzensfunktion), so ist die Implikation
(2) $\quad y \in \text{bd } F(x) \Rightarrow x \in F(y)$
richtig. Denn $y \in \text{bd } F(x) \Rightarrow u(x) = u(y) \Rightarrow x \in F(y)$.
Bed. (2) ist wegen Antisymm. nur mit $x \in \text{bd } F(y)$ möglich und liefert wegen Transitivität $F(x) = F(y)$.
Def.
Schließlich heiße F *stetig,* wenn aus $y' \to y$ und $x \in F(y')$ folgt, dass $x \in F(y)$.
Mit anderen Worten, F heißt stetig, wenn alle Mengen $F^{-1}(x)$ abgeschlossen sind.
Konstruktion von u: Sei $F \in \Phi$.
Für $r \in R$ betrachten wir den Vektor $r\,1 \in R^n$ in (1,...,1) - Richtung.
Wegen Monotonie und Reflexivität gibt es zu jedem x ein $r = r(x)$ mit $r\,1 \in F(x)$, da die Gerade aller $r\,1$ die Menge $x + R^n_+ \subset F(x)$ schneidet.
Def. u(x) sei das kleinste r mit $r\,1 \in F(x)$.
Ein solches r existiert, denn wenn r hinreichend klein (negativ) ist, gilt $r\,1 < x$ und somit $x \in \text{int } F(r\,1)$ wegen Monotonie. Also ist $r\,1 \notin \text{int } F(x)$ wegen Antisymmetrie.
Damit ist
(3) $\quad u(x) = \inf \{ r \,/\, r\,1 \in F(x) \}$
endlich und wird angenommen, weil F(x) abgeschlossen ist. Nun hat $y \in F(x)$ zur Folge dass $F(y) \subset F(x)$, also auch $u(y) \geq u(x)$.
D.h. *u ist monoton bzgl. der Relation F.*

Wir zeigen weiter
(4) $\quad$ **$u(r\,1) = r$.**
Klar ist $u(r\,1) =: s \leq r$ da $r\,1 \in F(r\,1)$. Wäre $s < r$, würde $s\,1 \in F(r\,1)$ nach (3) gelten.
Aber wegen $r\,1 > s1$ und Monotonie ist auch $r\,1 \in \text{int } F(s\,1)$, was der Antisymmetrie widerspricht.
Nach (4) ist u stetig auf der Diagonalen $r\,1$.
Nach (3) gilt offenbar $u(x)\,1 \in \text{bd } F(x)$.
Sei (2) richtig.
Dann ist $F(u(x)\,1) = F(x)$ und $u(x) = u(u(x)\,1)$. Die Bilder von F werden dann schon von allen $F(r\,1)$ erzeugt. Es ist nun $u(x) = r$ gdw. $F(r\,1) = F(x)$.

Lemma
Erfüllt $F \in \Phi$ Bedingung (2), so genügt u aus (3) der Forderung (1), d.h. u ist Nutzensfunktion für F.

($\Rightarrow$) Sei $y \in F(x)$, z.z. $u(y) \geq u(x)$. Das ist klar wegen Transitivität: $F(y) \subset F(x) \Rightarrow u(y) \geq u(x)$.
($\Leftarrow$) Sei $u(y) \geq u(x)$. z.z. $y \in F(x)$. Da $u(y)\,1 \geq u(x)\,1 \in F(x)$, folgt $u(y)\,1 \in F(x)$.
Wenn $y \notin F(x)$, verbinde man $u(y)\,1$ und y durch eine Strecke S. Sie schneidet den Rand von F(x) in einem Punkte $z \in \text{bd } F(x)$. Dann folgt nach (2) $\quad x \in F(z)$.
Es ist wegen Konvexität von F(y): $S \subset F(y)$, also auch $z \in F(y)$.
Daher gilt wegen Transitivität und (2):
$F(z) \subset F(y)$, $F(x) = F(z)$.
Also folgt nach Defin. (3) auch $\quad u(x) = u(z) \geq u(y)$, d.h. $u(x) = u(y) = u(z)$.

Wir erhalten so $F(x) = F(u(x)) = F(u(y)) = F(y)$.
Damit liegt S ganz in $F(y) = F(x)$; insbesondere ist also $y \in F(x)$. ♦

Folgerung:
(i) Ist $F \in \Phi$ vollständig und stetig, so gibt es eine Nutzensfunktion für F.
(ii) Hat umgekehrt $F \in \Phi$ eine stetige Nutzensfunktion, so ist F vollständig und stetig.

Beweis:
(i) (2) ist stets erfüllt, wenn F vollständig und stetig ist:
$y \in \mathrm{bd}\, F(x) \Rightarrow$ Es gibt $y' \to y$ mit $y' \notin F(x) \Rightarrow x \in F(y') \Rightarrow x \in F(y)$.
Also existiert eine Nutzensfunktion nach dem Lemma.
(ii) Denn erstens ist F vollständig, weil überhaupt eine Nutzensfunktion existiert,
und $y \in F^{-1}(x) \Leftrightarrow x \in F(y) \Leftrightarrow u(x) \geq u(y)$ zeigt, dass F stetig ist.

Bemerkungen
Nutzensfunktionen sind nicht eindeutig bestimmt. Aus (1) folgt noch nicht die Stetigkeit von u.
In R könnte gelten: $y \in F(x) \Leftrightarrow y \geq x$. Die spezielle Funktion u nach (3) wird dann gerade $u(x) = x$.
Allerdings wird (1) auch erfüllt von der Funktion $u(x) = x$ für $x < 0$ und $u(x) = x + 1$ für $x \geq 0$.

Satz.
$F \in \Phi$ sei vollständig und stetig. Dann definiert (3) eine stetige Nutzensfunktion zu F.
Hat umgekehrt $F \in \Phi$ eine stetige Nutzensfunktion, so ist F vollständig und stetig.

Beweis.
Die zweite Aussage ist schon bewiesen. F erfüllt (2), also definiert (3) eine Nutzensfunktion u. Es bleibt die Stetigkeit zu zeigen. Gelte $x' \to x$.
Angenommen $u(x') > u(x) + \varepsilon =: p$. Dann ist $p\,1 \notin F(x')$ also $x' \in F(p\,1)$.
Da $F(p\,1)$ abgeschlossen ist, folgt $x \in F(p1)$ und so $F(x) \subset F(p1)$.
Das liefert $u(x) \geq u(p\,1) = p = u(x) + \varepsilon$. Wid.
Angenommen $p' = u(x') < u(x) - \varepsilon = p$. Dann ist $p'\,1 \in F(x')$, also auch $p1 \in F(x')$, d.h. $x' \in F^{-1}(p1)$.
Da $F^{-1}(p1)$ abgeschlossen ist, folgt $x \in F^{-1}(p1)$, also $p1 \in F(x)$ und deshalb $u(x) \leq p = u(x) - \varepsilon$. Wid.
♦

Sei $F \in \Phi$. Wann ist u aus (3) eine konkave Nutzensfunktion ?
(5) $u(\lambda x + (1-\lambda) x') \geq \lambda u(x) + (1-\lambda) u(x')$. $\lambda \in (0, 1)$.
u konkave Nutzensfunktion $\Rightarrow$ u stetig und F vollständig.
(5) $\Leftrightarrow$ Die Menge $\mathrm{hypo}\, u := \{ (r, x) / r \leq u(x) \}$ ist konvex.
Weiter ist $r \leq u(x) \Leftrightarrow u(r\,1) \leq u(x) \Leftrightarrow x \in F(r\,1)$.

Also gilt (5) genau dann wenn die Menge M aller Paare (r, x) mit $x \in F(r\,1)$ konvex ist.

$M = \mathrm{graph}\, F_U$, wenn F_U die Einschränkung von F auf die Diagonale $U = \{ r\,1 / r \in R\}$ ist.

Cournot Gleichgewicht
n Produzenten eines Gutes produzieren davon jeweils x_i Einheiten.
Der mögliche Verkaufspreis p hängt von der Summe s = sum x_k ab; einfachster Fall
p(x) = a - b s .
Es entstehen Produktionskosten c_i(x_i) fuer die Produzenten. Ihr Gewinn sei
f_i(x) = p(x) x_i – c_i(x_i).
Maximieren von f_i bzgl. x_i (falls max existiert) .
(Dp/dx_i) x_i + p(x) - c'_i = 0; Dp/dx_i) = - b; - b x_i + a - b s - c'_i = 0,
x_i + s = (a - c'_i) / b , aufsummieren ergibt s + n s = (n a - sum c'_k) / b
also (n+1) s = (n a - sum c'_k) / b und
x_i = [a - c'_i - (n a - sum c'_k) / (n+1)] / b .
Sind die Kosten linear, ist c'_i konstant und x explicit auszurechnen.
2. Ableitung von f_i nach x_i wird - 2b < 0 , also hat man jeweils Maxima bzgl. x_i.
Man braucht dazu als Voraussetzung: 0 < a - c'_i - (n a - sum c'_i) / (n+1) ,
c'_i - sum c'_k) / (n+1) < a (1 - n/(n+1))
(n+1) c'_i - sum c'_k < n a
Ausserdem Preis positiv:
0 < a - b s = a - (n a - sum c'_k) / (n+1)

$0 < (n+1)\, a - n\, a - \text{sum } c'_k$
$a > \text{sum } c'_k$.

Diagonalspiele: Matrixspiel mit $A = \text{diag}(p_1, \dots, p_n)$, p_i positiv.
(x, y) GGS impliziert dass alle x_i, y_j pos. sind.

Wert der Spiels $v > 0$: Man wähle $x = (1/n,...1/n)$, um das zu sehen.

$x_j\, p_j = x A e_j >= v = x A y$ forall j; also $x_j > 0$, analog $y_j > 0$.

$v = x A y = \text{sum}_j\, x_j\, A_j\, y = \text{sum}_j\, x_j\, p_j\, y_j$ liefert $x_j p_j = v = y_j\, p_j$.

Also bis auf einen pos. Faktor $x_j = y_j = 1 / p_j$; Faktor so, dass Summe jeweils 1 wird,
also $(\text{sum}_j\, 1/p_j)^{-1}$;

Literatur

[Ber57} C. Berge Théorie générale des jeux à n personnes. Gauthier-Villars Paris 1957
[Cla83] F.H. Clarke Optimization and Nonsmooth Analysis. Wiley, NewYork, 1983
[Eke74] I. Ekeland On the variational principle. J. Math. Analysis & Appl. 47, No.2 (1974), 324-353
[Hil74] W.Hildenbrand Core and Equilibria of a Large Economy Princeton 1974
[Sel75] R. Selten Reexamination of the perfectness concept for equilibrium points in extensive games Intern. Journ. of Game Theory 4 (1975) 25-55
[HirLem93] J.-B. Hiriart-Urruty & C. Lemarechal Convex Analysis and Minimization Algorithms I, II. Springer, Berlin, Heidelberg, New York, 1993
[Ka41] S. Kakutani A generalization of Brouwer's fixed point theorem Duke Mathematical Journal 8 (1941), 457-459
[KaAk64] L.W. Kantorovich, G.P. Akilov Funktionalanalysis in normierten Räumen. Akad.V. Berlin 1964
[KK02] D. Klatte, B. Kummer. Nonsmooth equations in optimization: regularity, calculus, methods and applications. Series Nonconvex Optimization and Its Applications, Kluwer, Dordrecht-Boston-London, 2002.
[Kum79] B. Kummer. Spiele auf Graphen. Deutsch.V. d. Wiss. Berlin 1979, Birkhäuser 1980, MIR (russ.) 1982.
[Lue73] D.E. Luenberger, Introduction to Linear and Nonlinear Programming. Addison--Wesley, Reading, Massachusetts, 1973
[Mich56] E. Michael. Continuous selections I., Annals of Math., Vol. 63, (1956), pp 361—382.
[NM61] J.v. Neumann, O. Morgenstern Theory of Games and Economic Bahavior. Princeton 1944 deutsch: Spieltheorie und wirtschafliches Verhalten. Würzburg 1961 (und spätere Auflagen)
[OutKocZow98] J. Outrata, M. Kocvara, J. Zowe Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer 1998
[Owe71] G. Owen Spieltheorie. Springer, Berlin 1971 (Übersetzung)
[Rob51] J. Robinson An iterative method of solving a game. Annals of Mathematics 54, No 2 (1951) 296 - 301
[RaScZa79] B. Rauhut, N. Schmitz, E.-W. Zachow Spieltheorie Teubner Studienbücher Mathematik, Stuttgart 1979
[Roc70] R.T. Rockafellar Convex Analysis. Princeton University Press, Princeton 1970
[RoWe] R.T. Rockafellar and R. J._B. Wets Variational Analysis. Springer, Berlin 1998
[Ros71] J. Rosenmüller Kooperative Spiele und Märkte Berlin 1971
[Sea] Sea Optimization

[Sha53] L.S. Shapley A value for n-person games. Annals of Mathematics Studies 28 (1953), 307-317
[Sto&Wit70], J. Stör and C. Witzgall Convexity and Optimization in Finite Dimensions I. Springer, Berlin, 1970
[Vor74] N.N. Vorobiev Theory of games: Lessons for Economics and Cybernetics. (russ.), University press, Univ. of Leningrad 1974
[Vor75] N.N. Vorobiev Entwicklung der Spieltheorie. Deutsch.V. d. Wiss., Berlin 1975
[Vor84] N.N. Vorobiev Foundations of Game Theory- Noncooperative Games (russ.). Nauka, Moscow 1984
[Zei76], E. Zeidler Vorlesungen über nichtlineare Funktionalanalysis I – Fixpunktsätze. Teubner V., Leipzig 1976
Weitere interessante Literatur: Bolza, Fichtenholz, Funk, Ioffe/Tichomirov, Polyak, Smirnov, Troutman.
Für aktuellere Literatur zu Spielen auch aus ökonomischer Sicht,
sehe man z.B. auch unter “Mikroökonomie, Marktmodelle, R. Selten und W.Hildenbrand” nach.