

CLEMENS PUPPE

Skript zur Vorlesung

# Ökonomische Theorien der Gerechtigkeit

Sommersemester 2004

Institut für Wirtschaftstheorie und Operations Research  
Universität Karlsruhe  
Kollegengeb. IV am Schloss  
Kaiserstr. 12  
D – 76128 Karlsruhe  
Germany  
Telefon (0721) 608 3077  
Telefax (0721) 608 3082

## Vorbemerkung

Dieses Skriptum ersetzt weder eine Vorlesung noch ein Lehrbuch. Es besteht aus einer Zusammenstellung von Definitionen, Sätzen und einigen Beispielen, die nur teilweise motiviert und veranschaulicht werden. Die Bedeutung der Resultate wird kaum diskutiert. Diese Aufgabe muß die zugehörnde Vorlesung erfüllen. Herrn Lutz Limburg danke ich für die gewissenhafte Umsetzung der Graphiken in L<sup>A</sup>T<sub>E</sub>X-Format

Karlsruhe, im SS 2004

C. P.

# 1 Faire Allokationen in Ökonomien ohne Produktion

## 1.1 Grundlegende Definitionen und Notation

Es bezeichne  $N = \{1, \dots, n\}$  die Menge der Agenten. Jeder Agent  $i \in N$  ist durch eine Präferenzrelation  $\succeq_i$  auf dem Güterraum  $\mathbf{R}_+^L$  charakterisiert, wobei  $\{1, \dots, L\}$  die Menge der Güter ist. Die Relation  $\succeq_i$  heißt eine *Präferenzordnung*, falls  $\succeq_i$  reflexiv, transitiv und vollständig ist. Wir werden im folgenden immer annehmen, daß die individuellen Präferenzen diese Bedingungen erfüllen. Zusätzlich nehmen wir an, daß die Präferenzordnungen der Agenten *stetig* sind.<sup>1</sup> Dies garantiert, daß die Präferenzordnungen durch eine *ordinale Nutzenfunktion* repräsentiert werden können, d.h. es gibt eine Funktion  $u_i : \mathbf{R}_+^L \rightarrow \mathbf{R}$  derart, daß für alle Güterbündel  $x_i, y_i$  gilt:

$$x_i \succeq_i y_i \Leftrightarrow u_i(x_i) \geq u_i(y_i).$$

Die Präferenzordnung  $\succeq_i$  heißt *monoton*, falls gilt

$$x_i^l > y_i^l \text{ für alle } l \in \{1, \dots, L\} \Rightarrow x_i \succ_i y_i,$$

d.h. ein Mehrkonsum von jedem Gut wird strikt bevorzugt. Wir werden im folgenden immer annehmen, daß die individuellen Präferenzen *monoton* sind. Manchmal werden wir auch die folgende stärkere Bedingung fordern. Es bezeichne  $x_i \geq y_i \Leftrightarrow$  für alle  $l, x_i^l \geq y_i^l$ . Die Präferenzordnung  $\succeq_i$  heißt *streng monoton* falls gilt:

$$[x_i \geq y_i \text{ und } x_i \neq y_i] \Rightarrow x_i \succ_i y_i,$$

d.h. das Güterbündel  $x_i$  wird strikt gegenüber  $y_i$  bevorzugt, falls  $x_i$  von jedem Gut mindestens soviel enthält wie  $y_i$  und von mindestens einem Gut echt mehr.

Die Präferenzordnung  $\succeq_i$  heißt *konvex*, falls für alle  $x_i, y_i, z_i$  und alle  $t \in [0, 1]$  gilt:

$$[x_i \succeq_i z_i \text{ und } y_i \succeq_i z_i] \Rightarrow t \cdot x_i + (1 - t) \cdot y_i \succeq_i z_i.$$

---

<sup>1</sup>Einige Standardbegriffe – so wie hier die Stetigkeit – werden im folgenden nicht präzise definiert. Die genauen Definitionen finden sich in den meisten Lehrbüchern der Mikroökonomie, z.B. Mas-Colell, Whinston und Green: “Microeconomic Theory”, Oxford University Press, 1995.

Die Präferenzordnung  $\succeq_i$  heißt *streng konvex*, falls für alle  $x_i, y_i, z_i$  und alle  $t \in (0, 1)$  gilt:

$$[x_i \succeq_i z_i \text{ und } y_i \succeq_i z_i] \Rightarrow t \cdot x_i + (1 - t) \cdot y_i \succ_i z_i.$$

Die verfügbaren Ressourcen der Ökonomie seien mit  $\Omega = (\omega^1, \dots, \omega^L) \in \mathbf{R}_+^L$  bezeichnet. Die Zahl  $\omega^l$  gibt also die verfügbare Menge von Gut  $l$  an. Eine *Allokation* ist eine Liste  $z = (z_1, \dots, z_n) = ((z_1^1, \dots, z_1^L), \dots, (z_n^1, \dots, z_n^L))$ , wobei  $z_i$  das Güterbündel von Agent  $i$  ist. Eine Allokation ist also ein Element von  $\mathbf{R}_+^{L \cdot n}$ . Eine Allokation  $z$  heißt *erreichbar*, falls  $\sum_{i=1}^n z_i = \Omega$ , also falls

$$\left( \sum_{i=1}^n z_i^1, \dots, \sum_{i=1}^n z_i^L \right) = (\omega^1, \dots, \omega^L).$$

Die Menge aller erreichbaren Allokationen sei mit  $Z$  bezeichnet.

Für die folgende Analyse besteht eine *Ökonomie* aus der Angabe der individuellen Präferenzordnungen  $\succeq_1, \dots, \succeq_n$  und der Menge  $\Omega$  an verfügbaren Ressourcen.

Oben haben wir die individuellen Präferenzen auf dem Raum der Güterbündel definiert. Wir können die Präferenzen aber auch als Präferenzen über Allokationen auffassen, indem wir für alle  $x, y \in \mathbf{R}_+^{L \cdot n}$

$$x \succeq_i y \Leftrightarrow x_i \succeq_i y_i \tag{1.1}$$

setzen. Die Äquivalenz (1.1) drückt aus, daß die Präferenz über Allokationen für jeden Agenten nur vom *eigenen* Konsum (Güterbündel) abhängt.

Eine erreichbare Allokation  $z \in Z$  heißt *effizient*, falls es keine andere erreichbare Allokation  $z' \in Z$  gibt, derart daß für alle  $i \in N$  gilt  $z' \succeq_i z$  und für ein  $j \in N$  sogar  $z' \succ_j z$ . Die Menge der effizienten Allokationen wird auch die *Kontraktkurve* genannt.

Wir wollen im folgenden einige Kriterien formulieren, mit dem Ziel, dadurch "gerechte" von "nicht gerechten" Allokationen zu unterscheiden.

## 1.2 Gerechte Allokationen im Fall eines teilbaren Gutes

Betrachten wir zuerst den Fall  $L = 1$ . Wegen der Monotonie der Präferenzen haben in diesem Fall alle Agenten *diesselbe* (ordinale) Präferenz (nämlich: "mehr ist besser"). Das heißt aber insbesondere, daß alle Agenten identisch sind. Daraus folgt, daß die einzig gerechte Aufteilung der Ressourcen durch

die Gleichverteilung  $(\frac{\Omega}{n}, \dots, \frac{\Omega}{n})$  gegeben ist. Im gegebenen Rahmen ließe sich nämlich eine Ungleichbehandlung identischer Agenten nicht rechtfertigen. Man beachte, daß die Gleichverteilung auch effizient ist. Dies gilt im Fall  $L > 1$  im allgemeinen nicht mehr.

### 1.3 No-Domination und Equal-Division-Lower-Bound

Wir kommen nun zum interessanteren (und schwierigeren) Fall  $L > 2$ .

**Definition** Eine Allokation  $z \in Z$  erfüllt die Begingung *no-domination*, falls für kein Paar  $i, j$  von Agenten  $z_i^l \geq z_j^l$  für alle Güter  $l$  und  $z_i^{l_0} > z_j^{l_0}$  für ein Gut  $l_0$  gilt. Die Menge aller Allokationen, die no-domination erfüllen wird mit  $ND$  bezeichnet.

Das folgende Edgeworth-Diagramm illustriert die Menge  $ND$  im Fall  $n = 2$  und  $L = 2$ .

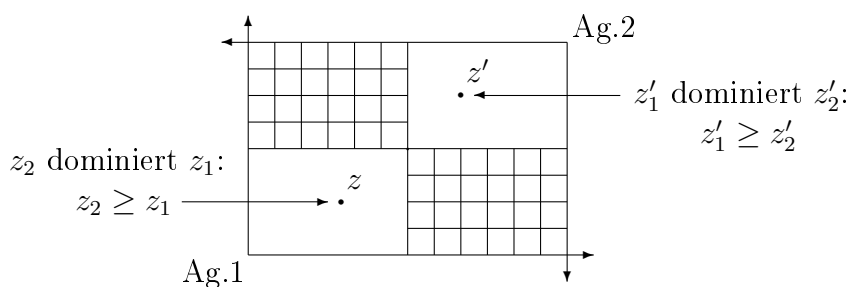


Abb.1: Die Menge  $ND$  besteht aus den beiden karierten Rechtecken

Bei monotonen Präferenzen liefert no-domination ein vernünftiges Minimalkriterium für eine gerechte Allokation: Kein Agent soll von *jedem* Gut mehr als eine andere Agentin konsumieren. Beachte, daß dieses Kriterium ohne Kenntnis der individuellen Präferenzen verifiziert werden kann.

**Definition** Eine Allokation  $z \in Z$  erfüllt die *equal-division-lower-bound*, falls für alle  $i$  gilt  $z \succeq_i \frac{\Omega}{n}$ . Die Menge aller Allokationen, die die equal-division-lower-bound erfüllen, wird mit  $EDLB$  bezeichnet.

Das folgende Diagramm illustriert die Menge  $EDLB$  im Fall  $n = 2$  und  $L = 2$ .

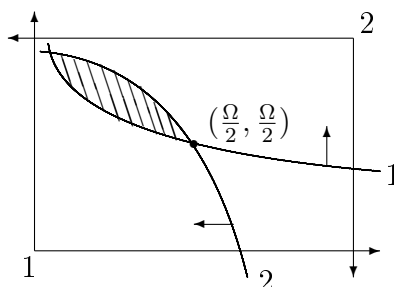


Abb.2: Die Menge EDLB besteht aus der schraffierten Linse

### 1.4 Neidfreie Allokationen

Wir kommen nun zum folgenden zentralen Kriterium:

**Definition** Eine Allokation  $z \in Z$  heißt *neidfrei (envy-free)*, falls für alle Paare  $i, j$  gilt  $z_i \succeq_i z_j$ . Die Menge aller neidfreien Allokationen wird mit  $EF$  bezeichnet.

Graphisch lassen sich neidfreie von nicht-neidfreien Allokationen im Fall  $n = 2$  und  $L = 2$  wie folgt unterscheiden. Gegeben  $z = (z_1, z_2) \in Z$ , bezeichne  $\bar{z}$  die Allokation, die dadurch entsteht, daß Agent 1 und 2 ihre Güterbündel austauschen, also

$$\bar{z} := (z_2, z_1).$$

Geometrisch erhält man die Allokation  $\bar{z}$ , indem man die Allokation  $z$  am Mittelpunkt  $(\frac{\Omega}{2}, \frac{\Omega}{2})$  der Edgeworth-Box spiegelt. Agent 1 ist neidfrei bei der Allokation  $z$  genau dann, wenn  $\bar{z}$  unterhalb (oder auf) der Indifferenzkurve durch  $z$  liegt. D.h. Agent 1 würde ihr Güterbündel nicht mit dem von Agent 2 tauschen wollen. Analog dazu ist Agent 2 bei  $z$  genau dann neidfrei, wenn  $\bar{z}$  oberhalb (oder auf) der Indifferenzkurve von Agent 2 durch  $z$  liegt. Die folgenden Diagramme illustrieren die verschiedenen Möglichkeiten.

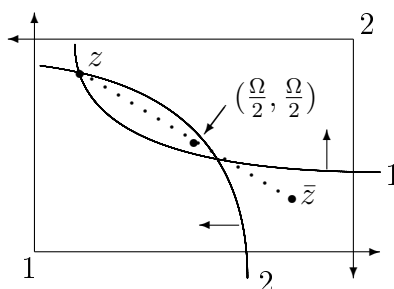


Abb.3:  $z$  ist neidfrei

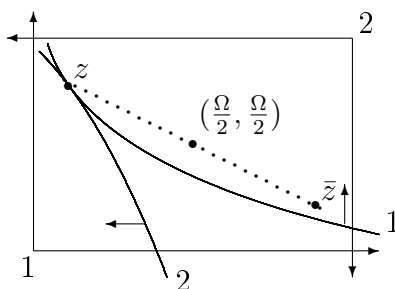


Abb.4:  $z$  ist nicht neidfrei (Agent 1 ist neidisch)

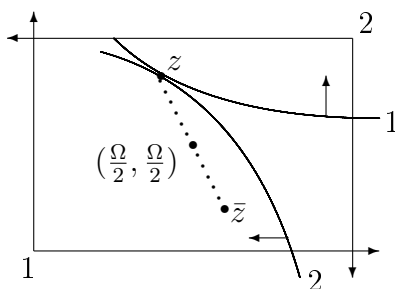


Abb.5:  $z$  ist nicht neidfrei (Agent 2 ist neidisch)

Beachte, daß beide obigen Allokationen effizient sind. *Gegenseitiger Neid* bei der Allokation  $z$  (also  $z_2 \succ_1 z_1$  und  $z_1 \succ_2 z_2$ ) bedeutet, daß  $z$  nicht effizient ist: Beide Agenten würden sich durch Tausch ihrer Güterbündel besser stellen:

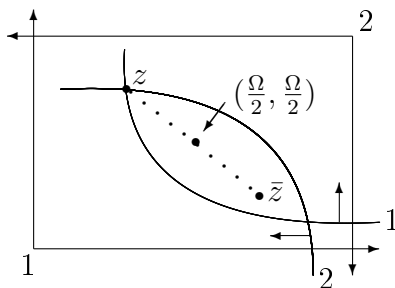


Abb.6:  $z$  ist nicht effizient, da beide Agenten neidisch sind

Andererseits impliziert Neidfreiheit nicht notwendigerweise Effizienz. Beispielsweise ist die Gleichverteilung  $(\frac{\Omega}{n}, \dots, \frac{\Omega}{n})$  immer neidfrei, aber im allgemeinen nicht effizient (siehe z.B. die Gleichverteilung in Abbildung 2 oben). Ein weiteres Beispiel einer neidfreien aber nicht effizienten Allokation ist die Allokation  $z$  in Abbildung 3.

**Notation** Im folgenden bezeichnen wir mit  $ND_P$ ,  $EDLB_P$  bzw.  $EF_P$  die Menge aller *effizienten* Allokationen in  $ND$ ,  $EDLB$  bzw.  $EF$ .

Wir wollen nun untersuchen, unter welchen Umständen  $EF_P$  nicht-leer ist, d.h. unter welchen Umständen neidfreie und effiziente Allokationen existieren. Die Antwort lautet: Sofern es ein Walrasianisches Gleichgewicht ausgehend von der Gleichverteilung der Ressourcen gibt, gibt es auch mindestens eine neidfreie und effiziente Allokation.

**Definition** Eine Allokation  $x^* \in Z$  zusammen mit einem Preisvektor  $p^* = ((p^1)^*, \dots, (p^L)^*)$  heißt ein *Walrasianisches Gleichgewicht ausgehend von der Gleichverteilung (EDWEq)*, falls für alle Agenten  $i$  die folgende Bedingung erfüllt ist. Für alle  $z_i$  in  $i$ 's Budgetmenge

$$B_i(p^*) := \{z_i \in \mathbf{R}_+^L : p^* \cdot z_i \leq p^* \cdot \frac{\Omega}{n}\}$$

gilt  $x_i^* \succeq_i z_i$ . Mit anderen Worten,  $x_i^*$  ist nutzenmaximal in  $i$ 's Budgetmenge (die wiederum durch die Gleichverteilung der Ressourcen und den Preisvektor  $p^*$  bestimmt ist).

**Satz 1.1** *Es sei  $[x^*, p^*]$  ein EDWEq. Dann ist die Allokation  $x^*$  effizient und neidfrei.*

**Beweis** Zunächst bemerken wir, daß wegen der Monotonie der individuellen Präferenzen jedes Walrasianische Gleichgewicht effizient ist (das ist das 1. Wohlfahrtstheorem). Ein genauer Beweis dieser Aussage findet sich z.B. in Mas-Colell, Whinston und Green (1995, S.549). Die Intuition ist aber sehr einfach: Angenommen wir wollen Agent  $i$  ausgehend von  $x^*$  verbessern. Da  $x^*$   $i$ 's Nutzen in ihrem Budget maximiert, ist das nur möglich, wenn  $i$  ein Güterbündel ausserhalb ihres Budgets erhält. Wegen der Erreichbarkeit muß dann aber ein anderer Agent  $j$  ein Güterbündel strikt innerhalb seines Budgets  $B_j(p^*)$  konsumieren. Wegen der Monotonie von  $\succeq_j$  würde Agent  $j$  dadurch schlechter gestellt.

Wir müssen also nur noch zeigen, daß  $x^*$  neidfrei ist. Das folgt aber sofort aus der Tatsache, daß alle Agenten aus derselben Budgetmenge wählen. Es gilt nämlich  $B_i(p^*) = B_j(p^*)$  für alle  $i, j$ , da alle Agenten dieselbe Anfangsausstattung besitzen.

Die folgende Abbildung illustriert die Lage eines EDWEq.



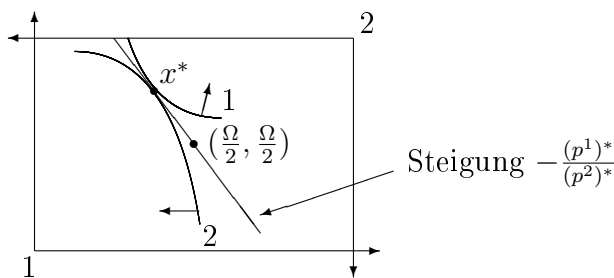


Abb.7:  $x^*$  ist ein EDWEq

**Bemerkung** Im Beweis von Satz 1.1 wurde Konvexität der individuellen Präferenzen nicht explizit benutzt. Man beachte aber, daß in Satz 1.1 die *Existenz* eines Walrasianischen Gleichgewichts vorausgesetzt ist. Zu den Standardbedingungen, die die Existenz eines solchen Gleichgewichts garantieren gehört aber auch die Konvexität der Präferenzen.

Wir beschliessen diesen Abschnitt mit einer graphischen Bestimmung aller neidfreien Allokationen im Fall  $n = 2$  und  $L = 2$ . Eine Allokation  $z \in Z$  liegt auf der *envy-boundary* von Agent 1, falls gilt  $z \sim_1 \bar{z}$ . Die envy-boundary von Agent 1 hat typischerweise die folgende Gestalt:

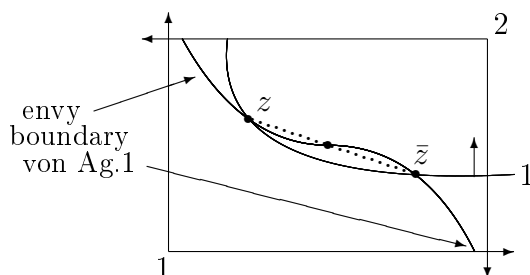


Abb.8: Die envy-boundary von Agent 1

Beachte, daß es Allokationen auf der envy-boundary von Agent 1 geben kann, die nicht die equal-division-lower-bound erfüllen. Die Allokation  $z$  in Abbildung 8 ist ein solcher Punkt.

Trägt man analog auch die envy-boundary von Agent 2 in das Diagramm ein, so ergibt sich folgendes Bild:

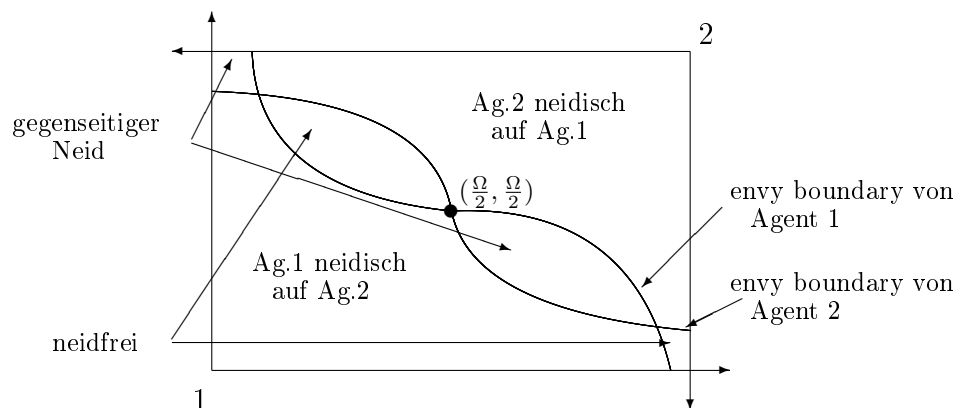


Abb.9: Die Menge aller neidfreien Allokationen im Fall  $n = L = 2$

## 1.5 Gleichbehandlung und Pareto-Indifferenz

**Definition** Eine Lösung (des Gerechtigkeitsproblems) ist einfach eine Teilmenge  $L(Z) \subseteq Z$  der Menge aller erreichbaren Allokationen.

Besteht man auf dem Kriterium der Effizienz, so sind die bisher vorgestellten Kandidaten für eine Lösung also  $ND_P$ ,  $EDLB_P$  und  $EF_P$ .

Die folgenden beiden Eigenschaften einer Lösung scheinen plausibel zu sein.

**Bedingung (ETE)** (Equal-treatment-of-equals) Falls  $\succeq_i = \succeq_j$  für alle  $i, j$ , so soll für jedes Element  $z \in L(Z)$  gelten  $z_i \sim_i z_j$  für alle  $i, j$ .

**Bedingung (PI)** (Pareto-Indifferenz) Falls  $z \in L(Z)$  und für alle  $i, z' \sim_i z$ , so auch  $z' \in L(Z)$ .

Die Bedingung (ETE) besagt, daß alle Agenten in jeder Lösung indifferent zwischen ihrem Güterbündel und dem jedes anderen Agenten sein sollen, falls alle Agenten dieselbe Präferenzordnung besitzen. Die Bedingung (PI) besagt, daß mit  $z$  auch  $z'$  eine Lösung sein soll, falls alle Agenten indifferent zwischen  $z'$  und  $z$  sind.

Es ist klar, daß die Lösungen  $EF$  bzw.  $EF_P$  die Bedingung (ETE) erfüllen: Ist nämlich  $z$  neidfrei, so gilt  $z_i \succeq_i z_j$  und  $z_j \succeq_j z_i$ ; wegen  $\succeq_i = \succeq_j$  gilt daher  $z_i \sim_i z_j$ . Andererseits erfüllt die Lösung  $ND_P$  die Bedingung (ETE) nicht (Gegenbeispiel: Übungsaufgabe). Die Lösung  $EDLB_P$  erfüllt (ETE) nur, wenn die individuellen Präferenzen konvex sind (Beweis: Übungsaufgabe).

Im Gegensatz zur Lösung  $EDLB_P$ , erfüllen die Lösungen  $ND_P$  und  $EF_P$  die Bedingung (PI) nicht. Die folgende Abbildung zeigt, daß sowohl  $ND_P$  als auch  $EF_P$  die Bedingung (PI) verletzen können: Die Allokationen  $z$  und  $z'$  sind Pareto-indifferent,  $z$  ist neidfrei und erfüllt no-domination, aber  $z'$  ist weder neidfrei noch erfüllt  $z'$  no-domination:

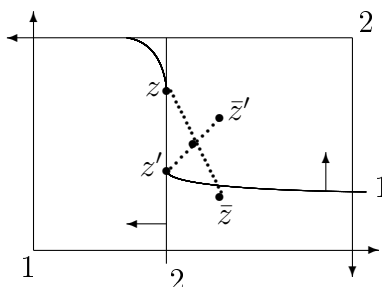


Abb.10: Die Lösungen  $ND_P$  und  $EF_P$  erfüllen nicht (PI)

Ein weiteres Beispiel dafür, daß die Bedingung (PI) durch die Lösung  $EF_P$  verletzt sein kann, ist das folgende:

**Beispiel** Betrachte  $n = 3$ ,  $L = 2$ ,  $\Omega = (1, 1)$  und die Nutzenfunktionen  $u_1(x^1, x^2) = \min\{x^1, x^2\}$ , sowie  $u_2(x^1, x^2) = u_3(x^1, x^2) = x^1 + x^2$ . Die Allokationen  $z$  und  $z'$  seien wie folgt gegeben:  $z = ((0, 0), (1, 0), (0, 1))$  und  $z' = ((0, 0), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$ . Es gilt  $z \in EF_P$  und  $z \sim_i z'$  für  $i = 1, 2, 3$ , aber  $z' \notin EF$ .

### 1.6 Egalitäre Äquivalenz

Wir kommen nun zu einem anderen wichtigen in der Literatur vorgeschlagenen Kriterium:

**Definition** Eine Allokation  $z \in Z$  heißt *egalitär äquivalent* (*egalitarian equivalent*), falls es ein Güterbündel  $x_0 \in \mathbf{R}_+^L$  gibt, derart daß für alle Agenten  $i$  gilt  $z_i \sim_i x_0$ . Es ist dabei *nicht* gefordert, daß die Allokation  $(x_0, \dots, x_0)$  erreichbar ist. Die Menge aller egalitär äquivalenten Allokationen wird mit  $EE$  bezeichnet, die Menge der zusätzlich effizienten Allokationen mit  $EE_P$ .

Die folgende Abbildung illustriert das Kriterium der egalitären Äquivalenz für den Fall  $n = 2$  und  $L = 2$ . Im Gegensatz zu den obigen Abbildungen handelt es sich nicht um ein Edgeworth Diagramm, da der Konsum beider Agenten vom üblichen Nullpunkt her abgetragen ist. Graphisch bedeutet daher Erreichbarkeit von  $z$ , daß  $z_1$  und  $z_2$  durch Spiegelung am Punkt  $\frac{\Omega}{2}$  auseinander

hervorgehen. Beachte auch, daß die Allokation  $(x_0, x_0)$  in der abgebildeten Situation nicht erreichbar ist (denn das "Referenzbündel"  $x_0$  enthält von beiden Gütern mehr als die Hälfte der jeweils verfügbaren Menge).

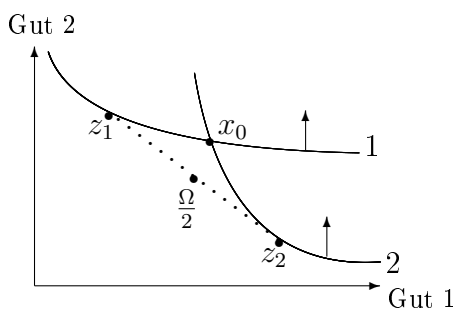


Abb.11: Eine egalitär äquivalente Allokation  $z$

**Satz 1.2** *Es existieren immer effiziente und egalitär äquivalente Allokationen, d.h.  $EE_P$  ist nicht-leer.*

**Beweis** Es sei  $r \in \mathbf{R}_+^L$  ein beliebiges Güterbündel mit  $r^l > 0$  für alle  $l$ . Weiterhin sei  $u_i$  eine Nutzenfunktion für Agent  $i$  mit  $u_i(0, \dots, 0) = 0$ . Betrachte für  $t \geq 0$  den Nutzenvektor  $(u_1(t \cdot r), \dots, u_n(t \cdot r))$ . Es sei  $t^*$  das maximale  $t$  derart, daß es eine erreichbare Allokation  $z \in Z$  mit

$$(u_1(z), \dots, u_n(z)) = (u_1(t^* \cdot r), \dots, u_n(t^* \cdot r))$$

gibt. Offensichtlich ist dieses  $z$  effizient und es gilt  $z \sim_i x_0$ , falls  $x_0 := t^* \cdot r$ .

**Bemerkung 1** Offensichtlich erfüllt die Lösung  $EE_P$  sowohl die Bedingung (ETE) als auch (PI).

**Bemerkung 2** Falls  $n = 2$ , gilt  $EF \subseteq EE$ . Dies läßt sich anhand der folgenden Abbildung einsehen: Ist  $z$  neidfrei, so muss  $z_2$  unterhalb der Indifferenzkurve von Agent 1 durch  $z_1$  liegen; ebenso muss  $z_1$  unterhalb der Indifferenzkurve von Agent 2 durch  $z_2$  liegen. Aber dann müssen sich die Indifferenzkurven schneiden. Der Schnittpunkt ist dann das Referenzbündel  $x_0$  aus der Definition von  $EE$ .

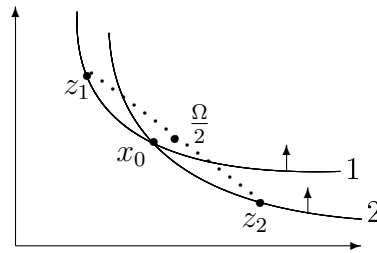


Abb.12: Im Fall  $n = 2$  impliziert Neidfreiheit egalitäre Äquivalenz

Für  $n > 2$  gilt diese Aussage nicht mehr. Die folgende Abbildung zeigt für  $n = 3$  das EDWEq  $x^*$ . Nach Satz 1.1 ist  $x^*$  ein Element von  $EE_P$ . Andererseits ist  $x^*$  nicht egalitär äquivalent, da es keinen Punkt gibt, in dem sich alle drei Indifferenzkurven schneiden.

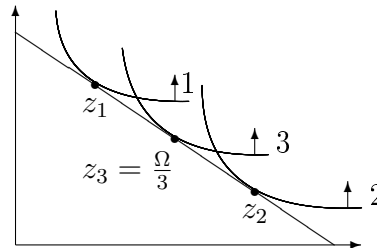


Abb.13: Das EDWEq  $x^*$  ist nicht egalitär äquivalent

**Bemerkung 3** Das folgende Beispiel zeigt, daß umgekehrt aus der egalitären Äquivalenz nicht die Neidfreiheit folgt. Es sei  $n = 2$ ,  $L = 2$ ,  $\Omega = (1, 1)$ ,  $u_1(x^1, x^2) = \min\{x^1, x^2\}$  und  $u_2(x^1, x^2) = x^1 + x^2$ . Dann ist die Allokation  $z = ((0, 0), (1, 1))$  effizient und egalitär äquivalent mit dem Referenzbündel  $x_0 = (2, 0)$ . Aber natürlich ist  $z$  nicht neidfrei (Agent 1 beneidet Agent 2).

Das Beispiel wirft ein ernstes Problem für das Konzept der egalitären Äquivalenz auf: Im Beispiel gilt  $z \in EE_P$ , obwohl  $z$  no-domination verletzt. Man kann sogar Ökonomien konstruieren, in denen *jede* effiziente und egalitär äquivalente Allokation no-domination verletzt! Diesem (gewichtigen) Nachteil es Kriteriums der egalitären Äquivalenz steht der Vorteil gegenüber, daß  $EE_P$  unter sehr allgemeinen Umständen nicht-leer ist. Diese Tatsache ist insbesondere in Ökonomien mit Produktion, denen wir uns nun zuwenden wollen, von Bedeutung.

## 2 Faire Allokationen in Ökonomien mit Produktion

Wir betrachten neben den bisherigen Konsumgütern nun noch ein weiteres Gut, nämlich "Freizeit". Es bezeichne  $l_i$  die von Agent  $i$  konsumierte Freizeit. Der Einfachheit halber normieren wir die gesamte zur Verfügung stehende Zeit pro Periode auf 1, d.h.  $l_i \in [0, 1]$ . Wir können also  $1 - l_i$  mit der Arbeitszeit von  $i$  identifizieren. Zur Vereinfachung betrachten wir ausser Freizeit nur noch ein anderes Gut, nämlich  $x$  ("Konsum"), das durch Arbeit hergestellt werden muß. Eine Allokation  $z$  ist eine Liste  $z = ((l_1, x_1), \dots, (l_n, x_n))$ . Eine solche Allokation ist neidfrei, falls für alle  $i, j$  gilt

$$(l_i, x_i) \succeq_i (l_j, x_j).$$

Die Existenz von effizienten und egalitär äquivalenten Allokationen ist auch in diesem erweiterten Rahmen unter sehr allgemeinen Bedingungen gesichert (mit einem ähnlichen Argument wie in Satz 1.2). Andererseits brauchen effiziente und neidfreie Allokationen nicht mehr zu existieren, wie das folgende Beispiel zeigt.

**Beispiel** Es gebe zwei Agenten, ein produziertes Gut  $x$  und Freizeit. Die Nutzenfunktionen seien wie folgt gegeben:

$$\begin{aligned} u_1(l_1, x_1) &= l_1 + \frac{11}{10}x_1, \\ u_2(l_2, x_2) &= l_2 + 2x_2. \end{aligned}$$

Beachte, daß Freizeit (relativ zu Konsum) für Agent 1 wichtiger ist als für Agent 2.

Die Produktionsmöglichkeiten sind durch folgende Bedingung beschrieben. Eine Allokation  $((l_1, x_1), (l_2, x_2))$  ist erreichbar, falls

$$x_1 + x_2 \leq 1 - l_1 + \frac{1}{10}(1 - l_2).$$

Beachte, daß Agent 1 produktiver als Agent 2 ist: mit dem Einsatz einer Arbeitseinheit produziert Agent 1 zehnmal mehr Einheiten des Konsumgutes.

Wir wollen nun eine notwendige Bedingung für die Effizienz einer erreichbaren Allokation  $z$  ableiten.

**Behauptung** Falls  $z$  effizient ist, dann gilt (i)  $l_1 = 0$  (Agent 1 hat keine Freizeit) und (ii) [ $l_2 < 1$  und  $x_2 = 0$ ] oder  $l_2 = 1$  (Agent 2 konsumiert nichts oder arbeitet nicht).

**Beweis (i)** Angenommen  $z$  ist erreichbar und  $l_1 > 0$ . Betrachte die erreichbare Allokation  $(\tilde{z}_1, z_2)$ , wobei  $\tilde{z}_1 = (l_1 - \epsilon, x_1 + \epsilon)$  und  $\epsilon > 0$  genügend klein. Wegen der Form von  $u_1$  gilt, daß Agent 1 die Allokation  $(\tilde{z}_1, z_2)$  strikt bevorzugen würde. Da Agent 2 unberührt bleibt, ist also  $z$  nicht effizient. Das heißt aber, daß in jeder effizienten Allokation  $l_1 = 0$  gelten muß.

**(ii)** Angenommen  $z$  ist erreichbar mit  $l_2 < 1$  und  $x_2 > 0$ . Betrachte die erreichbare Allokation  $(z_1, \tilde{z}_2)$ , wobei  $\tilde{z}_2 = (l_2 + \epsilon, x_2 - \frac{\epsilon}{10})$  und  $\epsilon > 0$  genügend klein. Man prüft nun leicht nach, daß  $u_2(\tilde{z}_2) > u_2(z_2)$ , d.h. Agent 2 würde  $\tilde{z}_2$  strikt gegenüber  $z_2$  bevorzugen. Das bedeutet aber, daß  $z$  nicht effizient sein kann. Daraus folgt, daß in jeder effizienten Allokation  $l_2 = 1$  oder [ $l_2 < 1$  und  $x_2 = 0$ ] gelten muß.

Wir haben also bisher gezeigt, daß jede effiziente Allokation entweder von der Form

a)  $z = ((0, 1 + \frac{1-l_2}{10}), (l_2, 0))$  für ein  $l_2 \in [0, 1]$ , oder

b)  $z = ((0, 1 - x_2), (1, x_2))$  für ein  $x_2 \in [0, 1]$

sein muß. Wir wollen nun zeigen, daß keine dieser Allokationen neidfrei sein kann.

Betrachte zunächst die Allokationen der Form a). Man sieht leicht, daß für alle solche Allokationen  $u_2(0, 1 + \frac{1-l_2}{10}) > u_2(l_2, 0)$  gilt, d.h. Agent 2 beneidet Agent 1. Betrachte nun die Allokationen der Form b). Agent 1 ist bei diesen Allokationen genau dann neidfrei, wenn gilt  $u_1(0, 1 - x_2) \geq u_1(1, x_2)$ . Man rechnet leicht nach, daß dies genau dann der Fall ist, wenn  $x_2 \leq \frac{1}{22}$ . Agent 2 ist bei den Allokationen der Form b) genau dann neidfrei, wenn gilt  $u_2(1, x_2) \geq u_2(0, 1 - x_2)$ . Man rechnet leicht nach, daß dies genau dann gilt, wenn  $x_2 \geq \frac{1}{4}$ . Da  $x_2$  nicht gleichzeitig kleiner gleich  $\frac{1}{22}$  und grösser gleich  $\frac{1}{4}$  sein kann, folgt also auch für die Allokationen der Form b), daß keine von ihnen neidfrei ist. Insgesamt ist damit gezeigt, daß es in der gegebenen Ökonomie keine effizienten und neidfreien Allokationen geben kann.

**Bemerkung 1** Das Problem im obigen Beispiel liegt darin, daß der produktivere Agent 1 gleichzeitig mehr Wert auf Freizeit legt. Im Fall  $L = 2$  kann man zeigen, daß  $EF_P$  nicht-leer ist, falls die produktiveren Agenten auch immer mehr Wert auf Konsum legen ("single-crossing property"). Allgemein ist  $EF_P$  immer nicht leer, falls entweder alle Agenten gleich produktiv sind

oder falls sie dieselbe relative Wertschätzung für Konsum haben. Das Problem mit diesen positiven Resultaten ist, daß die Voraussetzungen extrem restriktiv sind (Warum soll es nicht produktivere Agenten mit einer höheren Wertschätzung für Freizeit geben?).

**Bemerkung 2** In der Literatur wurden einige Vorschläge gemacht, um dem Problem der Nicht-Existenz von effizienten und neidfreien Allokationen zu begegnen. Einer davon lautet folgendermassen.

**Definition** Eine Allokation  $z$  heißt *output-adjusted envy-free*, falls für alle  $i$  gilt  $(l_i, x_i) \succeq_i (l_i(\tilde{x}_j), x_j)$ , wobei  $\tilde{x}_j$  das von  $j$  tatsächlich produzierte Güterbündel bezeichnet und  $1 - l_i(\tilde{x}_j)$  die Arbeitszeit, die *Agent*  $i$  benötigen würde, um  $\tilde{x}_j$  herzustellen.

Man kann zeigen, daß es immer effiziente und output-adjusted envy-free Allokationen gibt. Der Vorschlag birgt aber beträchtliche Probleme. Erstens ist die Definition nur bei konstanten Skalenerträgen anwendbar (sonst macht die Phrase “das von  $j$  tatsächlich produzierte Güterbündel” keinen Sinn). Zweitens ist das Kriterium unter Gerechtigkeitsaspekten durchaus angreifbar. Die vorgeschlagene Definition begünstigt nämlich die produktiveren Agenten in einer ungerechtfertigten Weise. Zum Beispiel wird kein Unterschied zwischen angeborenen und (kostspielig) erworbenen Fähigkeiten (Produktivität) gemacht.

### 3 Faire Allokation unteilbarer Güter

Bisher haben wir angenommen, daß alle Güter beliebig teilbar sind. Viele Güter (z.B. Immobilien, Möbelstücke, zu verrichtende Arbeiten, etc.) sind aber nicht beliebig teilbar. Dieser Tatsache wollen wir nun Rechnung tragen und in diesem Abschnitt auch unteilbare Güter betrachten.

Die verfügbaren Ressourcen der Ökonomie bestehen aus einer Menge von unteilbaren Gütern  $A = \{\alpha, \beta, \gamma, \dots\}$  und einem Geldbetrag  $M \geq 0$ . Der Einfachheit halber werden wir im folgenden annehmen, daß es genauso viele Güter wie Agenten gibt, also daß die Anzahl der Elemente von  $A$  gleich  $n$  ist.<sup>2</sup> Zusätzlich werden wir annehmen, daß jeder Agent genau ein unteilbares

---

<sup>2</sup>Diese Annahme ist weniger restriktiv als es den Anschein hat. Gibt es beispielsweise weniger unteilbare Güter als Agenten, so könnte man “virtuelle” Güter einführen, aus denen kein Agent einen Nutzen zieht. Gibt es mehr unteilbare Güter als Agenten, so könnte man “virtuelle” Agenten einführen, die nur aus Geld Nutzen ziehen.



Gut erhalten muß. Eine erreichbare Allokation  $z \in Z$  ist daher eine Liste  $(m, \sigma)$ , wobei  $m = (m_1, \dots, m_n) \in \mathbf{R}^n$  der Bedingung  $\sum_{i=1}^n m_i = M$  genügt und  $\sigma : N \rightarrow A$  eine eins-zu-eins Abbildung ist. Für jedes  $i$ , gibt  $m_i$  den Geldbetrag an, den Agent  $i$  erhält (beachte, daß  $m_i$  negativ sein kann). Die Abbildung  $\sigma$  gibt an, welcher Agent welches Gut erhält, wobei  $\sigma(i)$  das Gut von Agent  $i$ . Also falls  $z = (m, \sigma)$ , so ist  $z_i = (m_i, \sigma(i))$  das Bündel von Agent  $i$ .

Wir werden annehmen, daß die Präferenzen  $\succeq_i$  der Agenten zusätzlich zu den üblichen Annahmen (Reflexivität, Transitivität und Vollständigkeit) folgende beiden Bedingungen erfüllen. Für alle  $\alpha \in A$  und  $m_i > m'_i$  gilt  $(m_i, \alpha) \succ_i (m'_i, \alpha)$ , d.h. die Präferenzen sind strikt monoton im Geldbetrag. Ausserdem gelte für alle  $i$  das folgende "Kompensationsprinzip". Zu jedem Bündel  $(m_i, \alpha)$  und jedem Gut  $\beta \in A$  gibt es einen Geldbetrag  $m'_i \in \mathbf{R}$  derart, daß  $(m_i, \alpha) \sim_i (m'_i, \beta)$ . Diese Bedingung bedeutet, daß ausgehend von jedem Bündel  $(m_i, \alpha)$  die Agentin  $i$  durch einen Geldbetrag  $m'_i$  dafür kompensiert werden kann, daß sie statt dem Gut  $\alpha$  das Gut  $\beta$  erhält. Graphisch können die Präferenzen wie folgt veranschaulicht werden. Für jedes Gut in  $A$  wird eine "Geldachse" eingezeichnet. Je weiter rechts ein Bündel auf einer dieser Achsen liegt, desto mehr Geld enthält es. Eine Indifferenzkurve verbindet indifferente Bündel auf den verschiedenen Achsen (wegen der strikten Monotonie in Geld kann ein Bündel nie indifferent zu einem Punkt auf derselben Achse sein).

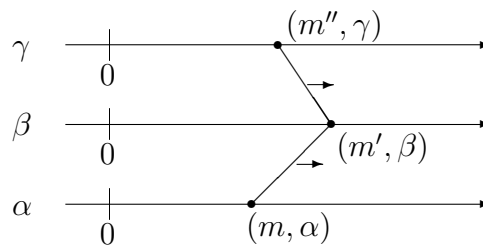


Abb.14: Indifferenzkurven mit unteilbaren Gütern und Geld

Eine Allokation  $(m, \sigma)$  ist neidfrei, falls für alle Paare von Agenten  $i, j$  gilt  $(m_i, \sigma(i)) \succeq_i (m_j, \sigma(j))$ . Die folgenden Abbildungen zeigen jeweils eine nicht-neidfreie und eine neidfreie Allokation.

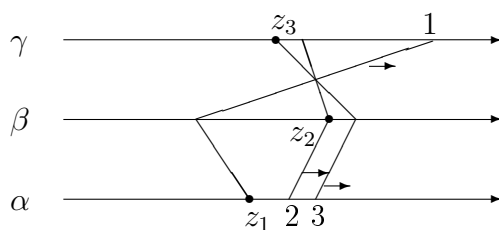


Abb.15: Agent 1 beneidet bei  $z$  Agent 2

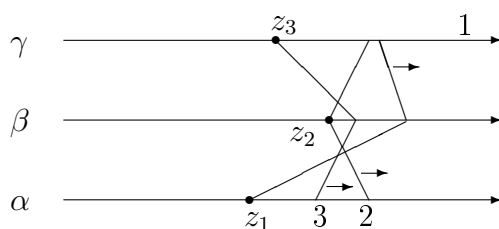


Abb.16: Eine neidfreie Allokation  $z$

Verlangt man  $m_i \geq 0$ , dann kann es sein, daß es keine neidfreien Allokationen gibt. Ist zum Beispiel  $n = 2$ ,  $M = 0$  und bevorzugen beide Agenten dasselbe Gut, so kann die Agentin, die das schlechtere Gut erhält dafür nicht kompensiert werden und ist daher auf den anderen Agenten neidisch. Andererseits gilt, falls  $m_i$  auch negative Werte annehmen darf, das folgende Resultat.

**Satz 1.3** Für genügend grosses  $M$  existieren immer neidfreie Allokationen

**Beweis** Wir betrachten nur den Fall  $n = 2$  (der allgemeine Fall ist etwas schwieriger). Es seien  $\alpha$  und  $\beta$  die beiden unteilbaren Güter. Betrachte das Bündel  $(0, \alpha)$  und wähle (gemäß des Kompensationsprinzips) einen Geldbetrag  $m$  so, daß gilt  $(0, \alpha) \sim_1 (m, \beta)$ . Wenn  $M$  genügend gross ist, existiert immer ein solcher Betrag  $m \leq M$ . Beachte, daß  $m$  aber auch negativ sein kann (nämlich dann, wenn Agent 1  $(0, \beta)$  strikt gegenüber  $(0, \alpha)$  bevorzugt). Wähle nun eine Allokation derart, daß Agent 2 das bevorzugte der beiden Bündel  $(0, \alpha)$  und  $(m, \beta)$  erhält. Das andere dieser beiden Bündel erhält Agent 1. Offensichtlich ist diese Allokation neidfrei (denn Agent 2 erhält sein bevorzugtes Bündel und Agent 1 ist indifferent). Verteile nun den eventuell noch zur Verfügung stehenden Geldbetrag auf die beiden Bündel so, daß

Agent 1 immer noch indifferent bleibt.<sup>3</sup> Nun gebe wiederum Agent 2 das bevorzugte der beiden Bündel. Die resultierende Allokation ist neidfrei und erreichbar.

Im Gegensatz zum Fall beliebig teilbarer Güter gilt im hier betrachteten Rahmen das folgende wichtige Resultat.

**Satz 1.4** *Jede neidfreie erreichbare Allokation ist effizient.*

**Beweis** Es sei  $z = (m, \sigma)$  eine neidfreie Allokation. Für jedes  $i \in N$  und jedes  $\alpha \in A$  sei  $m_\alpha^i \in \mathbf{R}$  der Betrag, für den  $z_i \sim_i (m_\alpha^i, \alpha)$  gilt. Wegen  $z_i \succeq_i z_j$  für alle  $i, j$ , gilt  $(m_{\sigma(j)}^i, \sigma(j)) \succeq_i (m_j, \sigma(j))$  für alle  $i, j$ . Insbesondere gilt also wegen der Monotonie der Präferenzen im Geldbetrag

$$m_{\sigma(j)}^i \geq m_j. \quad (3.1)$$

Angenommen nun es gibt eine Allokation  $z' = (m', \sigma')$  mit  $z' \succeq_i z$  für alle  $i$ , wobei mindestens eine Präferenz strikt ist. Dann gilt für alle  $i$

$$(m'_i, \sigma'(i)) \succeq_i (m_i, \sigma(i)) \sim_i (m_{\sigma'(i)}^i, \sigma'(i)),$$

wobei mindestens eine der Präferenzen auf der linken Seite strikt ist. Wegen der Monotonie im Geldbetrag gilt daher  $m'_i \geq m_{\sigma'(i)}^i$  mit mindestens einer strikten Ungleichung. Wegen (3.1) gilt daher für alle  $i$

$$m'_i \geq m_{\sigma'(i)}^i \geq m_{\sigma^{-1}(\sigma'(i))} \quad (3.2)$$

mit mindestens einer strikten Ungleichung. Da  $\{\sigma^{-1}(\sigma'(i)) : i \in N\} = A$ , gilt also wegen (3.2),

$$\sum_{i=1}^n m'_i > \sum_{i=1}^n m_i = M,$$

also ist  $z'$  nicht erreichbar. Aber das heißt, daß  $z$  effizient ist.

Egalitär äquivalente Allokationen existieren auch im Fall unteilbarer Güter unter den getroffenen Annahmen. Dies kann man sich wie folgt klarmachen. Betrachte für ein beliebiges Gut  $\alpha$  das Referenzbündel  $x_0 = (m_0, \alpha)$ . Wähle nun eine Zuordnung  $\sigma$  der unteilbaren Güter, und wähle  $m_i$  so daß für alle  $i$  gilt  $(m_i, \sigma(i)) \sim_i x_0$ . Beachte, daß natürlich  $m_j = m_0$  sein muß für den Agenten  $j$ , der Gut  $\alpha$  erhält. Ist  $m_0$  sehr klein (eventuell negativ) wird auch

---

<sup>3</sup>Damit dies immer möglich ist, benötigt man strenggenommen noch eine geeignete Stetigkeitsbedingung.

$\sum_i m_i$  klein. Ist  $m_0$  hingegen groß, so muß auch  $\sum_i m_i$  groß sein. Für einen Zwischenwert von  $m_0$  ist daher die Allokation  $z = (m, \sigma)$  erreichbar, und per Konstruktion ist diese Allokation egalitär äquivalent. Beachte, daß  $\alpha$  beliebig gewählt werden kann, insbesondere gibt es also mindestens so viele egalitär äquivalente Allokationen wie Güter. Die folgende Abbildung illustriert eine egalitär äquivalente Allokation für  $n = 3$

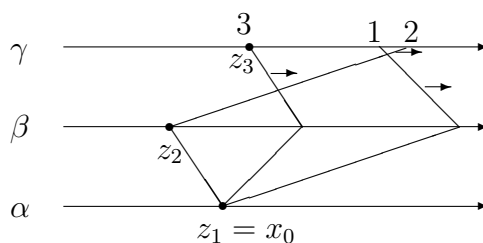


Abb.17: Eine egalitär äquivalente Allokation  $z$

### Literaturhinweise

Die Darstellung in Kapiteln 1 – 3 stützt sich auf ein unveröffentlichtes Buchmanuskript von W. Thomson mit dem Titel “Fair Allocation Rules”. Der Begriff der Neidfreiheit wurde von D. Foley: “Resource Allocation and the Public Sector”, *Yale Economic Essays* 7 (1967), S.45-98, in die Literatur eingeführt. Eine Version von Satz 1.1 findet sich in H. Varian: “Equity, Envy, and Efficiency”, *Journal of Economic Theory* 9 (1974), S.63-91. Der Begriff der egalitären Äquivalenz stammt von E. Pazner und D. Schmeidler: “Egalitarian Equivalent Allocations: A New Concept of Economic Equity”, *Quarterly Journal of Economics* 92 (1978), S.671-687; in dieser Arbeit wird auch Satz 1.2 bewiesen. Das Beispiel in Kapitel 2 zur Nicht-Existenz von effizienten und neidfreien Allokationen ist eine leichte Abwandlung des Beispiel von E. Pazner und D. Schmeidler: “A Difficulty in the Concept of Fairness”, *Review of Economic Studies* 41 (1974), S.441-443. Satz 1.3 über die Existenz von neidfreien Allokationen mit unteilbaren Gütern stammt von E. Maskin: “On the Fair Allocations of Indivisible Goods” in *Arrow and the Foundations of the Theory of Economic Policy*, G. Feiwel (Hrsg.), MacMillan 1987. Satz 1.4 findet sich in L.-G. Svensson: “Large Indivisibilities: An Analysis with Respect to Price Equilibrium and Fairness”, *Econometrica* 51 (1983), S.939-954.

## 4 Arrow's Unmöglichkeitstheorem

In diesem Kapitel wollen wir die Möglichkeit untersuchen, eine gerechte Allokation, oder allgemeiner: einen gerechten sozialen Zustand, mit Hilfe eines geeigneten *Wahlverfahrens* herbeizuführen. Betrachten wir folgendes Beispiel. Fünf Gemeinden  $A, B, C, D, E$  mit gleicher Einwohnerzahl müssen entscheiden, wo eine gemeinsam zu nutzende Einrichtung (z.B. ein Kinocenter, oder ein Krankenhaus, etc.) errichtet werden soll. Die Gemeinden sind durch ein Straßennetz miteinander verbunden, und die öffentliche Einrichtung kann an jeder Stelle entlang des Straßennetzes errichtet werden. Die folgende Abbildung zeigt das Straßennetz mit der Lage der Gemeinden und den jeweiligen Entfernungen in Kilometern (der Punkt  $w_0$  bezeichnet nur die Weggabelung):

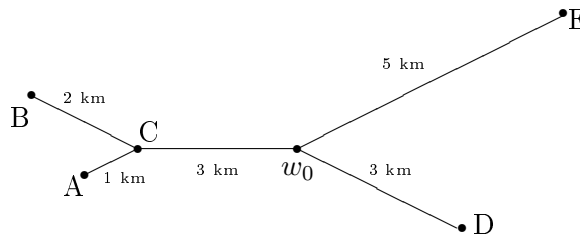


Abb.18: Wahl des Standorts einer öffentlichen Einrichtung

Betrachten wir zuerst die *paarweise Mehrheitsabstimmung*. Man prüft leicht nach, daß der Standort in der Gemeinde  $C$  gegen jeden anderen Standort  $x$  entlang des Straßennetzes bei einer Mehrheitswahl gewinnt, falls alle Gemeinden einen näheren Standort bevorzugen. Liegt nämlich  $x$  rechts von  $C$ , so wird die Mehrheit bestehend aus  $A, B$  und  $C$  für  $C$  und gegen  $x$  stimmen. Liegt andererseits  $x$  auf einem der Wege links von  $C$ , so werden

immer mindestens  $C, D$  und  $E$  für  $C$  stimmen. Eine Alternative, die gegen jede andere Alternative bei paarweiser Mehrheitsabstimmung gewinnt, heißt *Condorcet-Gewinner*. Mit anderen Worten, im obigen Beispiel ist der Standort  $C$  ein Condorcet-Gewinner. Das Problem der paarweisen Mehrheitsabstimmung ist aber, daß ein Condorcet-Gewinner im allgemeinen nicht existieren braucht.

**Beispiel: Das “Wahlparadox”** Betrachte drei Agenten 1, 2, 3, die über die drei Alternativen  $x, y, z$  zu entscheiden haben. Die Präferenzen der Agenten seien wie folgt gegeben (weiter oben stehende Alternativen werden strikt bevorzugt):

1	2	3
$x$	$y$	$z$
$y$	$z$	$x$
$z$	$x$	$y$

In dieser Situation gewinnt die Alternative  $x$  gegen  $y$  mit einer Mehrheit von 2 : 1 Stimmen; außerdem gewinnt  $y$  gegen  $z$  mit 2 : 1 Stimmen. Insbesondere kommen also weder  $y$  noch  $z$  als Gewinner der Wahl in Frage. Das Problem ist, daß auch  $x$  verliert, und zwar gegen  $z$ : wiederum eine Mehrheit von 2 : 1 würde beim paarweisen Vergleich zwischen  $x$  und  $z$  für  $z$  stimmen. Mit anderen Worten: *keine* Alternative kommt in dieser Situation als Ergebnis der paarweisen Mehrheitsabstimmung in Frage.

In seinem berühmten Buch “Social Choice and Individual Values” (1951, 2. Auflage 1963) hat Kenneth Arrow folgende Frage gestellt: Gibt es überhaupt ein geeignetes Wahlverfahren, das in jeder möglichen Situation eine konsistente Entscheidung herbeiführt? Sein berühmtes “Unmöglichkeitstheorem” beantwortet diese Frage negativ.

Es bezeichne  $X$  eine endliche Menge von *sozialen Zuständen* (z.B. Güterallokationen). Die Menge der Agenten sei  $N = \{1, \dots, i, \dots, n\}$ . Jeder Agent  $i$  besitzt eine reflexive, transitive und vollständige Präferenzordnung  $\succeq_i$  über den Elementen von  $X$ . Es bezeichne  $\mathcal{R}$  die Menge aller solchen Präferenzordnungen.

**Definition** Eine *Arrowsche Soziale Wohlfahrtsfunktion* (SWF) ist eine Abbildung  $F : \mathcal{R}^n \rightarrow \mathcal{R}$ , die jedem Profil  $(\succeq_1, \dots, \succeq_n)$  von individuellen Präferenzordnungen eine *soziale Präferenz*  $\succeq = F(\succeq_1, \dots, \succeq_n)$  zuordnet.

Beachte, daß wir in dieser Definition angenommen haben, daß eine SWF für

alle logisch möglichen Kombinationen von individuellen Präferenzordnungen eine soziale Präferenz liefern muß, d.h. eine SWF erfüllt folgende Bedingung.

**Bedingung (U)** (Unrestricted Domain)  $F$  ist auf der Menge  $\mathcal{R}^n$  aller logisch möglichen Präferenzprofile definiert.

Beachte außerdem, daß die paarweise Mehrheitswahlregel keine SWF im obigen Sinne ist, da die erzeugte soziale Präferenz nicht immer transitiv ist. Z.B. gilt im obigen Wahlparadox für die soziale Präferenz  $x \succ y$ ,  $y \succ z$  und  $z \succ x$ . Wir wollen nun anhand einiger Beispiele die zusätzlichen Bedingungen, die Arrow an eine SFW stellt, motivieren.

**Beispiel 1** Wähle eine feste Präferenzordnung  $\succeq^0$  auf  $X$ , und definiere  $F^0(\succeq_1, \dots, \succeq_n) := \succeq^0$  für alle Profile  $(\succeq_1, \dots, \succeq_n)$ . Mit anderen Worten,  $F^0$  schreibt die soziale Präferenz  $\succeq_0$  vor, egal welche Präferenzen die Individuen haben. Das Problem mit dieser SWF ist offensichtlich, daß die individuellen Präferenzen für die Entscheidung überhaupt keine Rolle spielen. Es sei zum Beispiel  $y$  die beste soziale Alternative gemäß der sozialen Präferenz  $\succeq^0$ . Dann würde immer  $y$  gewählt werden, und zwar auch dann, wenn es eine Alternative  $x$  gibt, die *jeder* Agent strikt gegenüber  $y$  bevorzugt. Um eine solche Situation zu vermeiden, lautet Arrow's erste Forderung.

**Bedingung (WP)** (Weak Pareto) Falls für alle  $i$  gilt  $x \succ_i y$ , dann soll gelten  $x \succ y$ , wobei  $\succeq$  die zum Profil  $(\succeq_1, \dots, \succeq_n)$  gehörende soziale Präferenz ist (also  $\succeq = F(\succeq_1, \dots, \succeq_n)$ ).

**Beispiel 2** Wähle einen festen Agenten  $k$  und setze für alle Präferenzprofile  $(\succeq_1, \dots, \succeq_n)$ ,  $F^k(\succeq_1, \dots, \succeq_n) := \succeq_k$ . Mit anderen Worten, egal wie die Präferenzen der Agenten im einzelnen aussehen, die soziale Präferenz stimmt immer mit der Präferenz von Agent  $k$  überein (was immer  $k$  für Präferenzen haben mag). Eine solche SWF heißt "diktatorisch" (mit Diktator  $k$ ). Allgemein:

**Definition** Eine SWF  $F$  heißt *diktatorisch*, falls es einen Agenten  $k$  gibt, derart daß für alle  $(\succeq_1, \dots, \succeq_n)$  gilt:  $x \succ_k y \Rightarrow x \succ y$ , wobei  $\succeq = F(\succeq_1, \dots, \succeq_n)$ .

**Beispiel 3 (Die Borda-Regel)** Das folgende Wahlverfahren definiert eine SWF, die (WP) erfüllt und nicht-diktatorisch ist. Die Menge  $X$  habe  $m$  Elemente. Jeder Agent vergibt an jede Alternative Punkte, und zwar nach folgendem Schema: die beste Alternative bekommt  $m$  Punkte, die zweitbeste  $m - 1$  Punkte, usw.<sup>4</sup> Sodann werden die individuellen Punkte für jede Alter-

<sup>4</sup>Der Einfachheit halber, nehmen wir hier an, daß in den individuellen Präferenzen keine Indifferenzen vorkommen.

native aufsummiert. Die soziale Präferenz ergibt sich dann aus der Punktesumme der Alternativen. Betrachte das folgende konkrete Beispiel mit drei Agenten und vier Alternativen:

1	2	3
$x$	$x$	$y$
$y$	$y$	$w$
$z$	$w$	$z$
$w$	$z$	$x$

Die Punktesumme für  $x$  ist 9, für  $y$  ist sie 10, für  $z$  ist sie 5 und für  $w$  lautet sie 6; d.h. die soziale Präferenz gemäß der Bordaregel lautet:  $y \succ_{Borda} x \succ_{Borda} w \succ_{Borda} z$ . Angenommen nun, die Agenten erhalten die zusätzliche Information, daß  $z$  und  $w$  so entscheidende Nachteile mit sich bringen, daß sie nicht ernsthaft als Alternativen zu  $x$  und  $y$  in Betracht kommen. Mit anderen Worten, nehmen wir an, die individuellen Rangordnungen verändern sich auf folgende Weise:

1	2	3
$x$	$x$	$y$
$y$	$y$	$x$
$z$	$w$	$w$
$w$	$z$	$z$

Der einzige Unterschied zu der vorhergehenden Situation ist, daß nun die Alternativen  $z$  und  $w$  immer unterhalb der Alternativen  $x$  und  $y$  stehen; die relative Anordnung von  $x$  und  $y$  sowie die relative Anordnung von  $z$  und  $w$  sind unverändert geblieben (tatsächlich hat sich nur die Präferenz von Agent 3 geändert). Berechnen wir aber die Punktesumme in der zweiten Situation, so ergibt sich, daß  $x$  nun 11 Punkte bekommt,  $y$  aber immer noch 10. Mit anderen Worten, obwohl sich die relative Position zwischen  $x$  und  $y$  (den beiden aussichtsreichen Kandidaten) nicht verändert hat, gewinnt nun  $x$  statt  $y$ . Die folgende Bedingung schließt ein solches Verhalten einer Wahlregel aus.

**Bedingung (I<sup>2</sup>)** (Binary Independence) Betrachte zwei Profile von Präferenzen  $(\succeq_1, \dots, \succeq_n)$  und  $(\succeq'_1, \dots, \succeq'_n)$  mit der jeweils zugehörigen sozialen Präferenz  $\succeq = F(\succeq_1, \dots, \succeq_n)$  und  $\succeq' = F(\succeq'_1, \dots, \succeq'_n)$ . Falls für zwei Alternativen  $x$  und  $y$  und alle  $i$  gilt:  $x \succeq_i y \Leftrightarrow x \succeq'_i y$  (falls sich also die relative Position von  $x$  und  $y$  für keinen Agenten beim Übergang von  $\succeq$  zu  $\succeq'$  ändert), so soll



gelten  $x \succeq y \Leftrightarrow x \succeq' y$  (d.h. die relative Position von  $x$  und  $y$  soll auch in den jeweiligen sozialen Rangordnungen dieselbe bleiben).

Die Bedingung  $(I^2)$  besagt also: Die einzige Information, die benötigt wird, um die relative Position von  $x$  und  $y$  in der sozialen Präferenz zu bestimmen, ist die relative Position *dieser beiden* Alternativen in allen individuellen Präferenzordnungen.

**Satz 4.1 (Arrow's Unmöglichkeitstheorem)** *Falls die Menge  $X$  mindestens drei Elemente hat, so ist jede SWF, die den Bedingungen  $(U)$ ,  $(WP)$  und  $(I^2)$  genügt, diktatorisch.*

Der Beweis von Satz 4.1 stützt sich auf zwei Hilfssätze (Lemmas), für die wir noch folgende Definitionen benötigen. Wir betrachten im folgenden immer eine fest vorgegebene SWF  $F$ , die den Bedingungen  $(U)$ ,  $(WP)$  und  $(I^2)$  genügt. Die zu einem Profil  $(\succeq_1, \dots, \succeq_n)$  gehörige soziale Präferenz unter  $F$  wird immer mit  $\succeq$  bezeichnet.

**Definition** Eine Gruppe  $L \subseteq N$  heißt *entscheidungskräftig für  $x$  gegen  $y$* , falls gilt: für alle  $(\succeq_1, \dots, \succeq_n)$ , [ $x \succ_i y$  für alle  $i \in L$  impliziert  $x \succ y$ ]. D.h. die Gruppe  $L$  ist entscheidungskräftig für  $x$  gegen  $y$ , falls eine einstimmige strikte Präferenz für  $x$  gegenüber  $y$  innerhalb der Gruppe  $L$  immer zu einer strikten sozialen Präferenz für  $x$  gegenüber  $y$  führt.

Eine Gruppe  $L \subseteq N$  heißt *schwach entscheidungskräftig für  $x$  gegen  $y$* , falls gilt: für alle  $(\succeq_1, \dots, \succeq_n)$ , [ $(x \succ_i y$  für alle  $i \in L$  und  $y \succ_i x$  für alle  $i \notin L$ ) impliziert  $x \succ y$ ]. D.h. die Gruppe  $L$  ist schwach entscheidungskräftig für  $x$  gegen  $y$ , falls eine einstimmige strikte Präferenz für  $x$  gegenüber  $y$  innerhalb der Gruppe  $L$  immer zu einer strikten sozialen Präferenz für  $x$  gegenüber  $y$  führt, falls alle anderen Agenten die entgegengesetzte Präferenz haben. Beachte, daß jede entscheidungskräftige Gruppe auch schwach entscheidungskräftig ist (denn sie kann ihre Präferenz durchsetzen, egal welche Präferenz alle anderen Agenten haben).

Im folgenden wollen wir folgende Abkürzungen benutzen:  $e(L, x, y)$  bedeute "Gruppe  $L$  ist entscheidungskräftig für  $x$  gegen  $y$ ";  $se(L, x, y)$  bedeute "Gruppe  $L$  ist schwach entscheidungskräftig für  $x$  gegen  $y$ "; schließlich bedeute  $e(L)$ : "Gruppe  $L$  ist entscheidungskräftig für jede Alternative gegen jede andere Alternative".

**Lemma 4.1** *Angenommen es gibt ein bestimmtes Paar  $a, b \in X$  derart, daß  $se(L, a, b)$ . Dann gilt  $e(L)$  (d.h.  $e(L, x, y)$  für alle  $x, y \in X$ .)*

**Beweis** Wir betrachten hier nur den Fall, daß  $x \neq a, b$  und  $y \neq a, b$ . Die Fälle  $x = a, x = b$  bzw.  $y = a, y = b$  lassen sich mit völlig analogen Argumenten behandeln.<sup>5</sup>

Aufgrund der Bedingung (U) können wir folgendes Profil von Präferenzen betrachten:

$i \in L$	$i \in N \setminus L$
$x$	$x \succ_i a$
$a$	$b \succ_i y$
$b$	$b \succ_i a$
$y$	$\vdots$
$\vdots$	$\vdots$

Beachte, daß die Anordnung der Alternativen  $x$  und  $y$  für die Agenten in  $N \setminus L$  offengelassen wird (insbesondere ist für  $i \notin L$  sowohl  $x \succ_i y$  als auch  $y \succ_i x$  verträglich mit den Angaben im obigen Präferenzprofil). Bei jedem Präferenzprofil mit den angegebenen Eigenschaften gilt  $x \succ_i a$  für alle  $i \in N$  und  $b \succ_i y$  für alle  $i \in N$ . Wegen (WP) gilt also  $x \succ a$  und  $b \succ y$ . Annahmegemäß gilt  $se(L, a, b)$ ; nun sind aber die Agenten in  $L$  die einzigen, die  $a$  gegenüber  $b$  strikt bevorzugen (alle anderen Agenten haben die entgegengesetzte strikte Präferenz). Also gilt wegen  $se(L, a, b)$ ,  $a \succ b$ . Wegen der Transitivität der sozialen Präferenz gilt daher insgesamt:  $x \succ y$ . Mit anderen Worten, in der obigen Situation kann die Gruppe  $L$  ihre Präferenz für  $x$  gegen  $y$  durchsetzen, und zwar unabhängig davon, wie die relative Position von  $x$  und  $y$  in den Präferenzen aller anderen Agenten ist. Wegen (I<sup>2</sup>) muß diese Aussage aber *immer* dann gelten, wenn die relative Position von  $x$  und  $y$  so wie angegeben ist (unabhängig davon, wie  $a$  und  $b$  angeordnet sind). Das heißt aber  $e(L, x, y)$  für alle  $x, y \in X$ .

**Lemma 4.2** *Angenommen  $e(L)$  für eine Gruppe  $L$  mit mindestens zwei Mitgliedern. Dann gibt es eine echte Untergruppe  $K \subset L$  mit  $e(K)$ .*

**Beweis** Falls  $L$  mindestens zwei Mitglieder hat, kann man  $L$  in zwei nicht-leere Untergruppen  $L'$  und  $L''$  aufspalten mit  $L' \cap L'' = \emptyset$  und  $L' \cup L'' = L$ . Wegen (U) können wir folgendes Präferenzprofil betrachten.

---

<sup>5</sup>Ein vollständiger Beweis findet sich in W.Bossert und F.Stehling: "Theorie kollektiver Entscheidungen", Springer Verlag (1990).

$i \in L'$	$i \in L''$	$i \in N \setminus L$
$a$	$b$	$c$
$b$	$c$	$a$
$c$	$a$	$b$
$\vdots$	$\vdots$	$\vdots$

Wegen  $e(L)$  gilt in dieser Situation  $b \succ c$ . Wegen der Vollständigkeit der sozialen Präferenz muß entweder (i)  $b \succ a$ , oder (ii)  $a \succeq b$  gelten. Betrachten wir zunächst den ersten Fall:

*Fall (i):* Die Agenten in  $L''$  sind die einzigen, die  $b$  gegenüber  $a$  bevorzugen. Falls also  $b \succ a$  gilt, so kann also die Gruppe  $L''$  in der obigen Situation ihre Präferenz durchsetzen, obwohl alle anderen Agenten die entgegengesetzte Präferenz haben. Wegen  $(I^2)$  gilt dies aber in jeder Situation, in der die relative Position von  $a$  und  $b$  so ist wie im angegebenen Präferenzprofil. Aber das bedeutet  $se(L'', b, a)$ , und daher wegen Lemma 4.1,  $e(L'')$ . In Fall (i) können wir also  $K = L''$  setzen.

*Fall (ii):* Falls  $a \succeq b$ , so gilt wegen  $b \succ c$  und der Transitivität  $a \succ c$ . Nun sind aber die Agenten in  $L'$  die einzigen, die  $a$  gegenüber  $c$  bevorzugen. Analog zu Fall (i) muß wegen  $(I^2)$  also  $se(L', a, c)$  gelten. Wiederum wegen Lemma 4.1 gilt dann  $e(L')$ , und wir können daher in Fall (ii)  $K = L'$  setzen.

**Beweis von Satz 4.1** Wegen (WP) gilt  $e(N)$ , d.h. die Gruppe aller Agenten ist entscheidungskräftig für jedes Paar von Alternativen. Wegen Lemma 4.2 können wir eine echte Untergruppe bilden, die immer noch entscheidungskräftig ist. Solange diese Untergruppe mindestens zwei Agenten enthält, können wir nach Lemma 4.2 wieder eine echte Untergruppe finden, die entscheidungskräftig ist. Dieser Prozess endet genau dann, wenn die letzte entscheidungskräftige Untergruppe nur noch einen Agenten enthält. Ein einzelner entscheidungskräftiger Agent ist aber definitionsgemäß ein Diktator.

## 5 Soziale Wohlfahrtsfunktionale und interpersoneller Nutzenvergleich

Betrachten wir nochmals das Beispiel der Wahl des Standorts einer öffentlichen Einrichtung aus dem letzten Kapitel (siehe Abb. 18). Wir hatten gesehen, daß der Standort  $C$  unter allen Standorten entlang des Straßennetzes der Condorcetgewinner ist (d.h.  $C$  gewinnt bei paarweiser Mehrheitsabstimmung gegen jeden anderen Standort). Der Standort  $C$  hat noch eine andere interessante Eigenschaft: er minimiert unter allen Standorten entlang des Straßennetzes die *Summe* der Distanzen zu den fünf Gemeinden. Identifiziert man den Nutzen  $u_i(x)$  eines Standorts  $x$  für eine Gemeinde  $i$  mit dem negativen Wert der Distanz, d.h. setzt man  $u_i(x) = -d(x, i)$ , so ist also der Standort  $C$  die *utilitaristische Lösung*: für jeden anderen Standort  $x$  gilt:

$$\sum_{i=A}^E u_i(C) > \sum_{i=A}^E u_i(x).$$

Können wir sagen, daß daher  $C$  der “gerechteste” Standort ist? Angenommen die öffentliche Einrichtung, um die es geht, sei ein Krankenhaus. Das gemeinsam zu nutzende Krankenhaus in Gemeinde  $C$  zu errichten, würde zwar die gesamten Fahrtkosten minimieren, bevorzugt aber die Einwohner der Gemeinde  $C$  in einer starken Weise. Im Gegensatz zu den Einwohnern von  $C$ , die keine Fahrtkosten haben, müßten beispielweise die Einwohner der Gemeinde  $E$  ganze 8 Kilometer zum Krankenhaus zurücklegen. Man könnte argumentieren, daß Gerechtigkeit verlangt, die Unterschiede zwischen den Entfernungen so klein wie möglich zu halten. Gemäß dieses Prinzips wäre der geeignete Standort die Weggabelung  $w_0$ . Denn in diesem Punkt ist die maximale Entfernung einer Gemeinde am geringsten (die Gemeinden  $B$  und  $E$  müssen jeweils 5 Kilometer zurücklegen). Mit anderen Worten, der Standort  $w_0$  ist die sogenannte *maximin Lösung*: für jeden anderen Standort  $x$  gilt:

$$\min_{i=A, \dots, E} u_i(w_0) > \min_{i=A, \dots, E} u_i(x).$$

Wir haben also nun schon zwei mögliche Kandidaten für die “gerechte” Wahl eines Standorts: die utilitaristische Lösung und die maximin (oder Rawlsche, oder egalitäre) Lösung. Sowohl die utilitaristische als auch die maximin Bewertung von sozialen Zuständen sind Beispiele von sogenannten “Sozialen Wohlfahrtsfunktionalen”. Diese wollen wir nun allgemein definieren und untersuchen.

## 5.1 Soziale Wohlfahrtsfunktionale

Es sei  $X$  wieder eine Menge von sozialen Zuständen oder sozialen Alternativen, wobei  $X$  im Gegensatz zum vorhergehenden Kapitel auch unendlich sein darf. Es bezeichne  $\mathcal{U}$  die Menge aller möglichen Nutzenfunktionen auf  $X$ .

**Definition** Ein *soziales Wohlfahrtsfunktional* (SWF1) ist eine Abbildung  $F : \mathcal{U}^n \rightarrow \mathcal{R}$ , die jedem Profil  $u = (u_1, \dots, u_n)$  von Nutzenfunktionen eine soziale Präferenzordnung  $\succeq_u = F(u)$  zuordnet.

Wie schon bei einer Arrowschen SWF haben wir in der Definition bereits folgende Bedingung vorausgesetzt:

**Bedingung (U)** (Unrestricted Domain) Ein SWF1 ist auf der Menge aller logisch möglichen Nutzenfunktionsprofile definiert.

Ähnlich zu den Bedingungen aus Kapitel 4 werden wir folgende weitere Bedingungen fordern:

**Bedingung (WP)** (Weak Pareto) Falls beim Profil  $u$  für alle  $i$  gilt  $u_i(x) > u_i(y)$ , so soll gelten  $x \succ_u y$ , wobei  $\succeq_u = F(u)$ .

**Bedingung ( $I_u^2$ )** (Binary Independence) Betrachte zwei Nutzenfunktionsprofile  $u$  und  $u'$  mit den jeweiligen sozialen Präferenzordnungen  $\succeq_u = F(u)$  und  $\succeq_{u'} = F(u')$ . Falls für die beiden Alternativen  $x$  und  $y$  und alle  $i$  gilt:  $u_i(x) = u'_i(x)$  und  $u_i(y) = u'_i(y)$ , so soll gelten  $x \succeq_u y \Leftrightarrow x \succeq_{u'} y$ .<sup>6</sup>

**Bedingung (PI)** (Pareto Indifferenz) Falls für zwei Alternativen  $x$  und  $y$  gilt:  $u_i(x) = u_i(y)$  für alle  $i$ , so soll  $x \sim_u y$  gelten, wobei  $\sim_u$  die zu  $\succeq_u = F(u)$  gehörende Indifferenzrelation ist.

Die beiden Bedingungen ( $I_u^2$ ) und (PI) lassen sich zu einer einzigen Bedin-

---

<sup>6</sup>Die Bedingung ( $I_u^2$ ) ist etwas schwächer als die Bedingung ( $I^2$ ), da nicht nur vorausgesetzt wird, daß die relative Position von  $x$  und  $y$  in jeder individuellen Präferenz unverändert bleibt, sondern sogar die zu  $x$  und  $y$  gehörenden individuellen Nutzenwerte.

gung zusammenfassen:

**Bedingung (INW)** (Independence of Non-Welfare Characteristics) Für alle  $u, u', x, y, x', y'$ : falls  $u(x) = u'(x')$  und  $u(y) = u'(y')$ , so soll gelten  $x \succeq_u y \Leftrightarrow x' \succeq_{u'} y'$ .

Die letzte Bedingung, die auch unter dem Namen "welfarism" bekannt ist, besagt also, daß die relative Anordnung von zwei Alternativen  $x$  und  $y$  in der sozialen Präferenz nur von den *von diesen beiden* Alternativen erzeugten Nutzenvektoren  $(u_1(x), \dots, u_n(x))$  bzw.  $(u_1(y), \dots, u_n(y))$  abhängen soll.

**Lemma 5.1** *Ein SWFl  $F$ , das (INW) erfüllt, genügt auch den Bedingungen  $(I_u^2)$  und (PI). Umgekehrt erfüllt jedes SWFl, das den Bedingungen (U),  $(I_u^2)$  und (PI) genügt, auch die Bedingung (INW).*

(Beweis: Übungsaufgabe)

Die Bedeutung der Bedingung (INW) ist, daß sie die Bestimmung der sozialen Präferenz für ein gegebenes Profil von Nutzenfunktionen dramatisch vereinfacht. Im allgemeinen wird die soziale Rangfolge von  $x$  und  $y$  von dem Profil der *gesamten* Nutzenfunktionen abhängen; unter (INW) aber nur von den zu  $x$  und  $y$  gehörenden Nutzenzahlen. Dies läßt sich wie folgt präzisieren.

**Definition** Eine Abbildung  $W : \mathbf{R}^n \rightarrow \mathbf{R}$ , die jedem Vektor  $(\nu_1, \dots, \nu_n)$  von Nutzenzahlen eine reelle Zahl  $W(\nu_1, \dots, \nu_n)$  zuordnet, heißt eine *Bergson-Samuelson Wohlfahrtsfunktion* (BS-swf).

**Definition** Eine BS-swf  $W$  repräsentiert das SWFl  $F$ , falls für alle  $x, y \in X$  und alle  $u$  gilt:

$$x \succeq_u y \Leftrightarrow W(u(x)) \geq W(u(y)),$$

wobei  $\succeq_u = F(u)$ .

**Lemma 5.2** *Es sei  $F$  ein SWFl, das die Bedingung (U) und eine geeignete Stetigkeitsbedingung erfüllt. Dann erfüllt  $F$  die Bedingung (INW) genau dann, wenn  $F$  durch eine BS-swf  $W$  repräsentiert werden kann.*

**Beweis** Es ist klar, daß die Möglichkeit einer Repräsentation durch eine BS-swf die Bedingung (INW) voraussetzt. Wir wollen nun umgekehrt zeigen, daß die Bedingung (INW) auch hinreichend für die Existenz einer solchen Repräsentation ist. Gegeben  $F$  definiere  $W$  wie folgt. Für zwei beliebig vorgegebene Vektoren  $\nu = (\nu_1, \dots, \nu_n)$  und  $\mu = (\mu_1, \dots, \mu_n)$  wähle zwei Alternativen  $x$  und  $y$ , sowie ein Profil  $u$  von Nutzenfunktionen derart, daß  $u(x) = \nu$  und  $u(y) = \mu$  (dies ist möglich wegen Bedingung (U)). Definiere nun eine Präferenzordnung

$\succeq_F$  auf  $\mathbf{R}^n$  durch:

$$\nu \succeq_F \mu \Leftrightarrow x \succeq_u y,$$

wobei  $\succeq_u = F(u)$ . Wegen der Bedingung (INW) hängt diese Definition nicht von der Wahl von  $x, y$  und  $u$  ab. Man kann leicht nachweisen, daß die so definierte Relation  $\succeq_F$  vollständig und transitiv ist. Die "geeignete Stetigkeitsbedingung" in der Aussage von Lemma 5.2 ist nun, daß  $\succeq_F$  als Relation auf  $\mathbf{R}^n$  stetig sein soll. Dann existiert eine "Nutzenrepräsentation"  $W : \mathbf{R}^n \rightarrow \mathbf{R}$  für  $\succeq_F$ , und man sieht leicht, daß  $W$  eine repräsentierende BS-swf für  $F$  ist.

Bevor wir uns dem Problem des interpersonellen Nutzenvergleichs zuwenden, wollen wir einige Beispiele betrachten:

**Beispiel 1: Das utilitaristische SWF1** Das utilitaristische SWF1  $F^{util}$  ist wie folgt definiert:  $F^{util}(u_1, \dots, u_n) := \succeq_u^{util}$ , wobei

$$x \succeq_u^{util} y \Leftrightarrow \sum_{i=1}^n u_i(x) \geq \sum_{i=1}^n u_i(y).$$

Die repräsentierende BS-swf  $W^{util}$  lautet also:  $W^{util}(\nu_1, \dots, \nu_n) := \sum_{i=1}^n \nu_i$ . Die Indifferenzkurven von  $W^{util}$  sehen im Fall  $n = 2$  wie folgt aus:

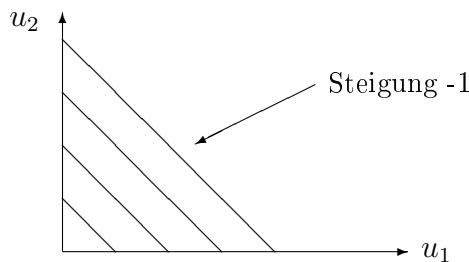


Abb.19: Indifferenzkurven von  $W^{util}$

**Beispiel 2: Das maximin SWF1** Das maximin SWF1  $F^{mm}$  ist wie folgt definiert:  $F^{mm}(u_1, \dots, u_n) := \succeq_u^{mm}$ , wobei

$$x \succeq_u^{mm} y \Leftrightarrow \min_{i=1, \dots, n} u_i(x) \geq \min_{i=1, \dots, n} u_i(y).$$

Die repräsentierende BS-swf  $W^{mm}$  lautet also:  $W^{mm}(\nu_1, \dots, \nu_n) := \min_i \nu_i$ . Die Indifferenzkurven von  $W^{mm}$  sehen im Fall  $n = 2$  wie folgt aus:

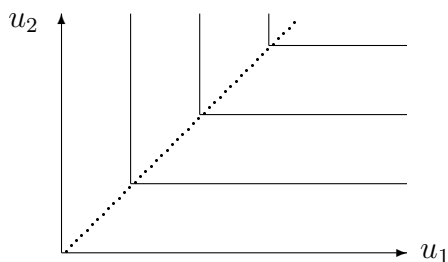


Abb.20: Indifferenzkurven von  $W^{mm}$

**Beispiel 3: Diktatorische SWFl** Ein SWFl heißt *diktatorisch* mit Diktator  $k$ , falls für alle  $u$  gilt  $x \succeq_u y \Leftrightarrow u_k(x) \geq u_k(y)$ . Eine repräsentierende BS-swf lautet daher  $W(\nu_1, \dots, \nu_n) = \nu_k$ . Die Indifferenzkurven sehen im Fall  $n = 2$  und  $k = 1$  wie folgt aus:

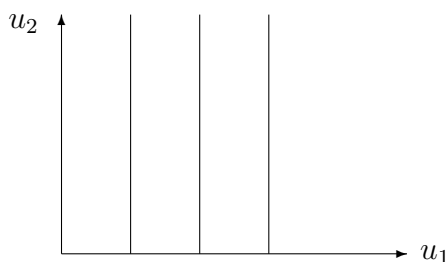


Abb.21: Indifferenzkurven einer BS-swf mit Agent 1 als Diktator

**Beispiel 4** Die folgende SWFl erfüllt nicht die Bedingung (INW) und kann folglich auch nicht mit Hilfe einer BS-swf repräsentiert werden. Definiere  $\hat{F}$  wie folgt: falls alle Nutzenfunktionen im Profil  $u$  identisch sind, setze  $x \succeq_u y \Leftrightarrow u_1(x) \geq u_1(y)$  (beachte, alle Agenten haben dieselbe Nutzenfunktion wie Agent 1 in diesem Fall); falls die Nutzenfunktionen im Profil  $u$  nicht identisch sind, setze  $x \sim_u y$  für alle  $x, y \in X$  (falls also verschiedene Agenten verschiedene Nutzenfunktionen haben, erklärt die zu diesem Profil gehörende soziale Präferenzordnung *alle* Alternativen als indifferent). Es ist klar, daß die soziale Anordnung zweier Alternativen  $x$  und  $y$  in diesem Fall nicht alleine von den Nutzenvektoren  $u(x)$  und  $u(y)$  abhängt, sondern von den vollständigen Nutzenfunktionen. Daher kann es für das so definierte SWFl  $\hat{F}$  auch keine repräsentierende BS-swf geben.



## 5.2 Interpersoneller Nutzenvergleich

Betrachten wir nochmals die utilitaristische bzw. die maximin Lösung bei der Wahl eines geeigneten Standorts einer öffentlichen Einrichtung im obigen Beispiel der fünf Gemeinden. Der für das folgende wichtige Punkt ist, daß beide Lösungen den Nutzen (im Beispiel: die negative Entfernung) verschiedener Gemeinden miteinander in Beziehung setzen, d.h. einen *interpersonellen Nutzenvergleich* anstellen. Die utilitaristische Lösung verlangt die Summierung der Nutzen, die maximin Lösung verlangt eine Entscheidung darüber, welcher beteiligte Agent den geringsten Nutzen besitzt. Ein wichtiger Unterschied der beiden Lösungen liegt darin, daß die utilitaristische Lösung "kardinale" individuelle Nutzenfunktionen voraussetzt, während die maximin Lösungen nur "ordinale" individuelle Nutzenfunktionen annehmen muß.

**Definition** Eine Nutzenfunktion  $u_i$  heißt *kardinal*, wenn sie bis auf positive affine Transformation eindeutig bestimmt ist. Dabei heißt eine Transformation  $f(u_i)$  *positiv affin*, falls  $f$  von der Form  $f(u_i) = a \cdot u_i + b$ , wobei  $a > 0$  und  $b \in \mathbf{R}$ . Kardinalität einer Nutzenfunktion bedeutet, daß *Nutzendifferenzen* sinnvoll interpretiert werden können: falls für eine kardinale Nutzenfunktion  $u_i$  gilt  $u_i(x) - u_i(z) > u_i(w) - u_i(y)$ , so läßt sich dies interpretieren als "der Nutzenzuwachs für Agent  $i$  ist beim Übergang von  $z$  zu  $x$  größer als beim Übergang von  $y$  zu  $w$ ". Der Grund ist, daß die Ungleichung  $u_i(x) - u_i(z) > u_i(w) - u_i(y)$  unverändert bleibt, falls man  $u_i$  mit einer positiven Zahl  $a$  multipliziert und eine beliebige Konstante  $b$  addiert.<sup>7</sup>

Im Gegensatz dazu heißt eine Nutzenfunktion  $u_i$  *ordinal*, wenn sie bis auf eine beliebige positive monotone Transformation  $f(u_i)$  eindeutig bestimmt ist. In diesem Fall haben weder absolute Nutzenzahlen noch Nutzendifferenzen eine inhaltliche Bedeutung. Eine ordinale Nutzenfunktion beinhaltet nicht mehr Information als eine Präferenzordnung: alles was sich aussagen läßt ist, daß  $u_i(x) > u_i(y)$  genau dann gilt, wenn  $i$  die Alternative  $x$  der Alternative  $y$  strikt bevorzugt.

<sup>7</sup>Wichtige Beispiele kardinaler Nutzenfunktionen sind die von-Neumann-Morgenstern Nutzenfunktionen der Erwartungsnutzentheorie. Ist  $u_i$  eine von-Neumann-Morgenstern Nutzenfunktion, so gilt  $u_i(x) - u_i(z) > u_i(w) - u_i(y)$  genau dann, wenn  $i$  die Lotterie  $(x; \frac{1}{2}, y; \frac{1}{2})$  ( $x$  mit Wahrscheinlichkeit  $\frac{1}{2}$  und  $y$  mit Wahrscheinlichkeit  $\frac{1}{2}$ ) gegenüber der Lotterie  $(w; \frac{1}{2}, z; \frac{1}{2})$  strikt bevorzugen würde.

**Bedingung (ONC)** Ein soziales Wohlfahrtsfunktional  $F$  heißt *vom Typ ONC* (ordinal **n**-comparable, benutzt nur ordinale Nutzeninformation und keinen interpersonellen Vergleich), falls  $F(u) = F(\tilde{u})$  für alle Nutzenfunktionsprofile  $u, \tilde{u}$  mit

$$(\tilde{u}_1, \dots, \tilde{u}_n) = (f_1(u_1), \dots, f_n(u_n)),$$

wobei  $f_1, \dots, f_n$  beliebige positiv monotone Transformationen.

**Bedingung (OFC)** Ein soziales Wohlfahrtsfunktional  $F$  heißt *vom Typ OFC* (ordinal **f**ully comparable, benutzt nur ordinale Nutzeninformation aber vollen interpersonellen Vergleich), falls  $F(u) = F(\tilde{u})$  für alle Nutzenfunktionsprofile  $u, \tilde{u}$  mit

$$(\tilde{u}_1, \dots, \tilde{u}_n) = (f(u_1), \dots, f(u_n)),$$

wobei  $f$  eine (gemeinsame) positive monotone Transformation.

Der wesentliche Unterschied zwischen den Bedingungen (ONC) und (OFC) ist also, daß bei (ONC) das SWFl invariant sein soll gegenüber beliebigen *unabhängigen* Transformationen der individuellen Nutzenfunktionen, während bei (OFC) die Invarianz nur dann gefordert wird, falls die individuellen Nutzenfunktionen *in gleicher Weise* transformiert werden. Die entsprechenden Bedingungen für kardinale Nutzenfunktionen lauten wie folgt.

**Bedingung (CNC)** Ein soziales Wohlfahrtsfunktional  $F$  heißt *vom Typ CNC* (cardinal **n**-comparable, benutzt kardinale Nutzeninformation und keinen interpersonellen Vergleich), falls  $F(u) = F(\tilde{u})$  für alle Nutzenfunktionsprofile  $u, \tilde{u}$  mit

$$(\tilde{u}_1, \dots, \tilde{u}_n) = (a_1 \cdot u_1 + b_1, \dots, a_n \cdot u_n + b_n),$$

wobei  $a_1, \dots, a_n > 0$  und  $b_1, \dots, b_n \in \mathbf{R}$ .

**Bedingung (CFC)** Ein soziales Wohlfahrtsfunktional  $F$  heißt *vom Typ CFC* (cardinal **f**ully comparable, benutzt kardinale Nutzeninformation und vollen interpersonellen Vergleich), falls  $F(u) = F(\tilde{u})$  für alle Nutzenfunktionsprofile  $u, \tilde{u}$  mit

$$(\tilde{u}_1, \dots, \tilde{u}_n) = (a \cdot u_1 + b, \dots, a \cdot u_n + b),$$

wobei  $a > 0$  und  $b \in \mathbf{R}$ .

**Bedingung (CUC)** Ein soziales Wohlfahrtsfunktional  $F$  heißt *vom Typ*

*CUC* (cardinal **u**nit-comparable, benutzt kardinale Nutzeninformation und interpersonellen Vergleich der Nutzeinheiten), falls  $F(u) = F(\tilde{u})$  für alle Nutzenfunktionsprofile  $u, \tilde{u}$  mit

$$(\tilde{u}_1, \dots, \tilde{u}_n) = (a \cdot u_1 + b_1, \dots, a \cdot u_n + b_n),$$

wobei  $a > 0$  und  $b_1, \dots, b_n \in \mathbf{R}$ .

Es ist offensichtlich, daß folgende Implikationen gelten:  $(CNC) \Rightarrow (CUC) \Rightarrow (CFC)$ ,  $(ONC) \Rightarrow (OFC)$ , sowie  $(ONC) \Rightarrow (CNC)$  und  $(OFC) \Rightarrow (CFC)$ . Das utilitaristische SWFl  $F^{util}$  ist vom Typ  $(CUC)$  (und daher auch vom Typ  $(CFC)$ ), das maximin SWFl  $F^{mm}$  ist vom Typ  $(OFC)$  (und daher auch vom Typ  $(CFC)$ , aber nicht vom Typ  $(CUC)$ ).

Wir wollen nun das Arrowsche Unmöglichkeitstheorem unter Benutzung des Begriffs eines SWFl umformulieren und einen alternativen Beweis vorstellen. Wir werden dabei sehen, daß sich die Unmöglichkeit darauf zurückführen läßt, daß der Arrowsche Rahmen interpersonellen Nutzenvergleich ausschließt. Eine Arrowsche SWF im Sinne von Kapitel 4 ist eine Abbildung die jedem Profil von *ordinalen* Präferenzen eine soziale Rangordnung zuordnet. Faßt man eine Arrowsche SWF als ein SWFl  $F$  auf, so bedeutet dies also, daß  $F$  vom Typ  $(ONC)$  sein muß. Das folgende Theorem zeigt, daß die Unmöglichkeit sogar dann bestehen bleibt, wenn man auf individueller Ebene kardinale Information zuläßt (solange man keinen interpersonellen Vergleich vornimmt).

**Satz 5.1** *Die Menge  $X$  habe mindestens drei Elemente. Es sei  $F$  ein SWFl, das den Bedingungen  $(U)$ ,  $(I_u^2)$ ,  $(WP)$ ,  $(PI)$  und einer geeigneten Stetigkeitsbedingung genügt. Dann ist  $F$  genau dann vom Typ  $(ONC)$  oder  $(CNC)$ , wenn  $F$  diktatorisch ist.*

**Beweis** (vgl. Roemer (1996, Theorem 1.2\*, S.31)<sup>8</sup>) Man kann leicht überprüfen, daß jedes diktatorische SWFl alle genannten Bedingungen erfüllt. Wir zeigen nun die umgekehrte Aussage. Aufgrund der Lemmata 5.1 und 5.2 gibt es zu  $F$  eine repräsentierende BS-swF  $W$ . Wir wollen die Gestalt der Indifferenzkurven der Funktion  $W$  bestimmen. Dabei beschränken wir uns der Einfachheit halber auf den Fall  $n = 2$ .

*Schritt 1:* Betrachte zwei beliebige Punkte  $\nu = (\nu_1, \nu_2) \in \mathbf{R}^2$  und  $\nu' = (\nu_1 + k_1, \nu_2 + k_2) \in \mathbf{R}^2$ , die auf derselben  $W$ -Indifferenzkurve liegen (also

<sup>8</sup>Roemer (1996) bezieht sich im folgenden immer auf das Buch J. Roemer: "Theories of Distributive Justice", Harvard University Press (1996).

$W(\nu) = W(\nu')$ ). Wegen (WP) kann weder ( $k_1 > 0$  und  $k_2 > 0$ ) noch ( $k_1 < 0$  und  $k_2 < 0$ ) gelten. Angenommen es gilt  $k_1 \geq 0$  und  $k_2 \leq 0$ . Betrachte die folgende Transformation  $(f_1, f_2)$ :

$$\begin{aligned} f_1(t) &= 2t - \nu_1, \\ f_2(t) &= (t + \nu_2)/2. \end{aligned}$$

Offensichtlich ist  $(f_1, f_2)$  sowohl unter (ONC) als auch unter (CNC) eine zulässige Transformation (da jedes  $f_i$  eine positive affine Transformation). Man macht sich leicht klar, daß mit  $F$  auch  $W$  vom Typ (ONC) bzw. (CNC) sein muß, d.h. es muß wegen  $W(\nu_1, \nu_2) = W(\nu'_1, \nu'_2)$  auch  $W(f_1(\nu_1), f_2(\nu_2)) = W(f_1(\nu'_1), f_2(\nu'_2))$  gelten. Nun gilt aber

$$(f_1(\nu_1), f_2(\nu_2)) = (\nu_1, \nu_2)$$

und

$$(f_1(\nu'_1), f_2(\nu'_2)) = (f_1(\nu_1 + k_1), f_2(\nu_2 + k_2)) = (\nu_1 + 2k_1, \nu_2 + k_2/2).$$

Daraus folgt (wegen Transitivität der Indifferenz), daß  $(\nu_1 + k_1, \nu_2 + k_2)$  und  $(\nu_1 + 2k_1, \nu_2 + k_2/2)$  auf derselben  $W$ -Indifferenzkurve liegen müssen. Würde nun  $k_1 > 0$  und  $k_2 < 0$  gelten, so erhielte man  $\nu_1 + 2k_1 > \nu_1 + k_1$  und  $\nu_2 + k_2/2 > \nu_2 + k_2$ , was wegen (WP) nicht möglich ist. Also muß gelten: entweder  $k_1 = 0$  oder  $k_2 = 0$ . Analog zeigt man dieselbe Schlußfolgerung für den Fall  $k_1 \leq 0$  und  $k_2 \geq 0$ .

*Schritt 2:* Wir haben also bisher gezeigt, daß zwei beliebige  $W$ -indifferente Punkte  $\nu$  und  $\nu'$  entweder dieselbe erste Koordinate oder dieselbe zweite Koordinate haben müssen. Angenommen für  $\nu$  und  $\nu'$  stimmt die erste Koordinate überein, d.h.  $k_1 = 0$ . Dann muß die Indifferenzkurve durch den Punkt  $\nu$  eine vertikale Gerade sein. Ist nämlich  $\nu''$  ein weiterer zu  $\nu$  indifferenter Punkt, so muß wegen Schritt 1 angewandt auf  $\nu$  und  $\nu''$  gelten, daß diese beiden Punkte entweder in der ersten oder der zweiten Koordinate übereinstimmen. Würden sie aber nicht in der ersten Koordinate übereinstimmen, sondern in der zweiten, dann wären  $\nu''$  und  $\nu'$  zwei indifferente Punkte, die weder in der ersten noch in der zweiten Koordinate übereinstimmen. Das kann aber nach Schritt 1 nicht sein.

Analog zeigt man, daß im Fall  $k_2 = 0$  die Indifferenzkurve durch  $\nu$  eine horizontale Gerade sein muß.

*Schritt 3:* Bisher haben wir gezeigt, daß alle Indifferenzkurven entweder vertikale oder horizontale Geraden sein müssen. Ist aber *eine* Indifferenzkurve

vertikal, so müssen *alle* Indifferenzkurven vertikal sein (weil sich Indifferenzkurven nicht schneiden können). Analog sind alle Indifferenzkurven horizontal, sobald eine davon horizontal ist. Mit anderen Worten: entweder alle Indifferenzkurven sind vertikale Geraden, oder alle Indifferenzkurven sind horizontale Geraden. Aber dies bedeutet: entweder Agent 1 ist ein Diktator, oder Agent 2 ist ein Diktator.

Satz 5.1 zeigt den Grund für die Arrowsche Unmöglichkeit der Präferenzaggregation: das Fehlen der Möglichkeit eines interpersonellen Nutzenvergleichs.<sup>9</sup> Was Satz 5.1 über die Aussage von Theorem 4.1 hinaus noch hinzufügt, ist die Erkenntnis, daß auch *kardinale* individuelle Nutzeninformation nicht hilft, solange man diese nicht interpersonell vergleicht. Welche Möglichkeiten erhalten wir aber, wenn wir interpersonellen Vergleich zulassen?

**Bedingung (AN)** (Anonymität) Das Nutzenfunktionsprofil  $u'$  entstehe aus  $u$  durch eine Permutation (also für eine Permutation  $\sigma : N \rightarrow N$  gelte  $u'_i = u_{\sigma(i)}$ ). Dann soll gelten  $F(u) = F(u')$ .

Die Bedingung (AN) ist sehr plausibel: die soziale Präferenz soll nicht von der Identität der Agenten abhängen, sondern nur von der *Verteilung* der Nutzenfunktionen. Beachte, daß ein diktatorisches SWFl diese Bedingung verletzt. Läßt sich  $F$  durch eine BS-swF  $W$  repräsentieren, so bedeutet (AN), daß die  $W$ -Indifferenzkurven im Fall  $n = 2$  symmetrisch zur Winkelhalbierenden sein müssen.

**Definition** Ein SWFl  $F$  heißt *positionell diktatorisch vom Rang  $k$* , falls  $x \succeq_u y$  genau dann gilt, wenn der Nutzen des  $k$ -bestgestellten Agenten im Zustand  $x$  mindestens so groß ist wie der Nutzen des  $k$ -bestgestellten Agenten im Zustand  $y$ . Eine positionelle Diktatur vom Rang  $k = n$  entspricht also dem maximin SWFl. Für  $k = 1$  ergäbe sich ein "maximax" Prinzip.<sup>10</sup>

---

<sup>9</sup>Beachte, daß auch die Borda-Regel einen interpersonellen Nutzenvergleich anstellt. Man kann die Borda-Regel als utilitaristische Lösung mit speziellen Nutzenfunktionen interpretieren, nämlich mit den durch die Borda-Punkte definierten Nutzen.

<sup>10</sup>Dieses Prinzip ("Maximiere den Nutzen des am bestgestellten Agenten") ist allerdings von einem Gerechtigkeitsstandpunkt nicht sehr attraktiv.

**Satz 5.2** *Es sei  $F$  ein SWFl, das den Bedingungen  $(U)$ ,  $(I_u^2)$ ,  $(WP)$ ,  $(PI)$ ,  $(AN)$  und einer geeigneten Stetigkeitsbedingung genügt. Dann ist  $F$  vom Typ  $(OFC)$  genau dann, wenn es positionell diktatorisch ist.*

**Beweis** (vgl. Roemer (1996, Theorem 1.3, S.33) Man überprüft wieder leicht, daß jedes positionell diktatorische SWFl die genannten Bedingungen erfüllt. Um die umgekehrte Aussage zu zeigen, studieren wir wieder die Gestalt der  $W$ -Indifferenzkurven für eine repräsentierende BS-swf  $W$ .

*Schritt 1:* Betrachte einen Punkt  $\nu = (\nu_1, \nu_1)$  auf der Winkelhalbierenden und einen zu  $\nu$  indifferenten Punkt  $\nu' = (\nu_1 + k_1, \nu_1 + k_2)$ . Wir beschränken uns zunächst auf den Bereich unterhalb der Winkelhalbierenden, d.h. wir betrachten den Fall  $k_1 \geq 0$  und  $k_2 \leq 0$ . Es sei  $f$  die folgende Transformation:

$$f(t) = \begin{cases} 2t - \nu_1 & \text{falls } t \geq \nu_1 \\ (t + \nu_1)/2 & \text{falls } t \leq \nu_1 \end{cases} .$$

Da  $f$  monoton steigend, ist  $(f, f)$  eine zulässige Transformation unter  $(OFC)$  (beachte, daß  $f$  nicht affin ist). Es muß also wegen  $W(\nu) = W(\nu')$  auch  $W(f(\nu_1), f(\nu_1)) = W(f(\nu_1 + k_1), f(\nu_1 + k_2))$  gelten. Nun gilt aber

$$(f(\nu_1), f(\nu_1)) = (\nu_1, \nu_1)$$

und

$$(f(\nu_1 + k_1), f(\nu_1 + k_2)) = (\nu_1 + 2k_1, \nu_1 + k_2/2).$$

Daraus folgt (wegen Transitivität der Indifferenz), daß  $(\nu_1 + k_1, \nu_1 + k_2)$  und  $(\nu_1 + 2k_1, \nu_1 + k_2/2)$  auf derselben  $W$ -Indifferenzkurve liegen müssen. Würde nun  $k_1 > 0$  und  $k_2 < 0$  gelten, so erhielte man  $\nu_1 + 2k_1 > \nu_1 + k_1$  und  $\nu_1 + k_2/2 > \nu_1 + k_2$ , was wegen  $(WP)$  nicht möglich ist. Also muß entweder  $k_1 = 0$  oder  $k_2 = 0$  gelten.

*Schritt 2:* Völlig analog zum Beweis von Satz 5.1 folgt aus Schritt 1, daß die Indifferenzkurve durch  $\nu$  im Bereich unterhalb der Winkelhalbierenden entweder vertikal oder horizontal sein muß.

*Schritt 3:* Daraus folgt (wieder unter Benutzung von  $(WP)$ ), daß entweder *alle* Indifferenzkurven unterhalb der Winkelhalbierenden vertikal sind, oder *alle* Indifferenzkurven sind in diesem Bereich horizontal. Graphisch sehen die Indifferenzkurven unterhalb der Winkelhalbierenden also so aus:

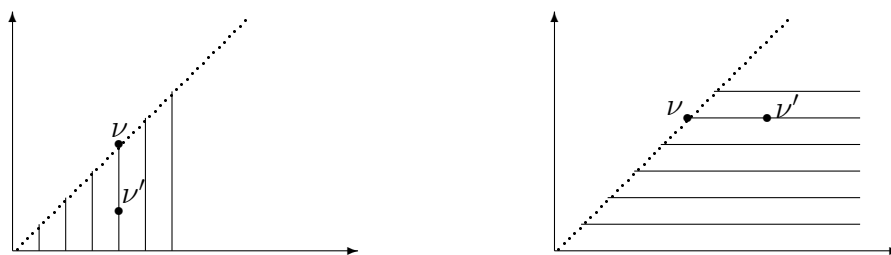


Abb.22: Indifferenzkurven unterhalb der Winkelhalbierenden  
 Links: Fall a, Rechts: Fall b

Schritt 4: Völlig analog zeigt man für den Bereich oberhalb der Winkelhalbierenden (d.h.  $k_1 \leq 0$  und  $k_2 \geq 0$ ), daß die Indifferenzkurven auch dort entweder alle vertikal oder alle horizontal sein müssen:

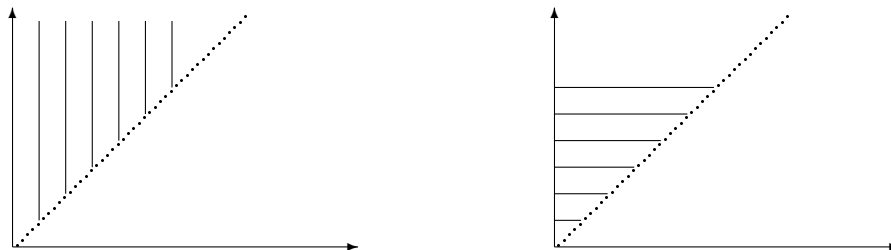


Abb.23: Indifferenzkurven oberhalb der Winkelhalbierenden  
 Links: Fall A, Rechts: Fall B

Schritt 5: Es verbleiben also zunächst noch vier Möglichkeiten, wie die  $W$ -Indifferenzkurven insgesamt aussehen können: Fall a kombiniert mit Fall A oder B und Fall b kombiniert mit Fall A oder B. Die Bedingung (AN) schließt aber die Kombinationen aA (Agent 1 ist Diktator) und bB (Agent 2 ist Diktator) aus. Also verbleiben genau die Möglichkeiten aB (“maximax”) und bA (“maximin”).

Ersetzen wir in Satz 5.2 die Bedingung (OFC) durch (CUC) so erhalten wir statt des maximin SWFL das utilitaristische SWFL.

**Satz 5.3** *Es sei  $F$  ein SWFL, das den Bedingungen (U),  $(I_u^2)$ , (WP), (PI), (AN) und einer geeigneten Stetigkeitsbedingung genügt. Dann ist  $F$  vom Typ (CUC) genau dann, wenn  $F$  das utilitaristische SWFL  $F^{util}$  ist.*

**Beweis** (vgl. Roemer (1996, Theorem 1.4, S.35)). Man prüft leicht nach, daß

$F^{util}$  alle genannten Bedingungen erfüllt. Wir zeigen nun, daß  $F^{util}$  tatsächlich das einzige SWFl, das diese Bedingungen erfüllt. Es gelte wieder der Einfachheit halber  $n = 2$ , und es sei  $W$  eine repräsentierende BS-swfl. Betrachte zwei beliebige  $W$ -indifferente Punkte  $\nu = (\nu_1, \nu_2)$  und  $(\nu_1 + k_1, \nu_2 + k_2)$ . Für jede positive reelle Zahl  $r$  seien die Transformationen  $f_1$  und  $f_2$  wie folgt definiert:

$$\begin{aligned} f_1(t) &= r \cdot t + (1 - r) \cdot \nu_1 \\ f_2(t) &= r \cdot t + (1 - r) \cdot \nu_2 \end{aligned}$$

Da  $F$  nach Voraussetzung vom Typ (CUC) ist, ist  $f = (f_1, f_2)$  offensichtlich eine zulässige Transformation. Da  $W(\nu_1, \nu_2) = W(\nu_1 + k_1, \nu_2 + k_2)$  muß also auch  $W(f_1(\nu_1), f_2(\nu_2)) = W(f_1(\nu_1 + k_1), f_2(\nu_2 + k_2))$  gelten. Es ist aber  $(f_1(\nu_1), f_2(\nu_2)) = (\nu_1, \nu_2)$  und  $(f_1(\nu_1 + k_1), f_2(\nu_2 + k_2)) = (\nu_1 + r \cdot k_1, \nu_2 + r \cdot k_2)$ . Es liegen also  $(\nu_1, \nu_2)$  und  $(\nu_1 + r \cdot k_1, \nu_2 + r \cdot k_2)$  auf derselben Indifferenzkurve. Da  $r > 0$  beliebig gewählt war, gilt dies für *alle* positiven  $r$ . Mit anderen Worten, die Indifferenzkurve durch  $\nu$  ist eine Gerade mit Steigung  $k_2/k_1$ . Wegen (WP) gilt  $k_2/k_1 \leq 0$ , und wegen (AN) muß sogar  $k_2/k_1 = -1$  gelten. Mit anderen Worten, die  $W$ -Indifferenzkurven sehen wie in Abbildung 19 aus. Das heißt aber, daß  $F$  das utilitaristische SWFl sein muß.



## 6 Axiomatic Bargaining

Satz 5.1 aus dem vorangehenden Kapitel zeigte, daß der Übergang von ordinaler zu kardinaler individueller Nutzeninformation keinen Ausweg aus der Arrowschen Unmöglichkeit liefert, solange man keinen interpersonellen Nutzenvergleich vornimmt. Heißt dies, daß man kardinale Nutzeninformation auf individueller Ebene unter keinen Umständen ohne interpersonellen Vergleich ausnutzen kann? Betrachte das folgende Beispiel eines SWFl.

**Beispiel: Das Nash SWFl** Es sei  $x_d \in X$  ein fest gewählter sozialer Zustand, z.B. der “status quo”. Definiere ein SWFl  $F^{Nash}$  folgendermaßen:  $F^{Nash}(u_1, \dots, u_n) := \sum_u^{Nash}$ , wobei

$$x \succeq_u^{Nash} y \Leftrightarrow \prod_{i=1}^n (u_i(x) - u_i(x_d)) \geq \prod_{i=1}^n (u_i(y) - u_i(x_d))$$

für alle  $x, y \in X$  derart, daß  $u_i(x) \geq u_i(x_d)$  und  $u_i(y) \geq u_i(x_d)$  für alle  $i$ .

Das Nash SWFl bewertet also soziale Zustände gemäß dem *Produkt* der individuellen Nutzenzuwächse gegenüber einem “status quo” Punkt  $x_d$ . Das Interessante an  $F^{Nash}$  ist, daß es vom Typ (CNC) (aber nicht vom Typ (ONC)) ist. Mit anderen Worten, die von  $F^{Nash}$  induzierte soziale Präferenz ist invariant gegenüber individuell unabhängigen positiv affinen Transformationen. Dies läßt sich leicht wie folgt nachweisen. Betrachte die positiv affinen Transformationen  $\tilde{u}_i = a_i \cdot u_i + b_i$ ; dann gilt für alle  $x, y$ :

$$\begin{aligned} \prod_{i=1}^n (\tilde{u}_i(x) - \tilde{u}_i(x_d)) &\geq \prod_{i=1}^n (\tilde{u}_i(y) - \tilde{u}_i(x_d)) \\ \Leftrightarrow \prod_{i=1}^n (a_i u_i(x) + b_i - (a_i u_i(x_d) + b_i)) &\geq \prod_{i=1}^n (a_i u_i(y) + b_i - (a_i u_i(x_d) + b_i)) \\ \Leftrightarrow \left( \prod_{i=1}^n a_i \right) \cdot \prod_{i=1}^n (u_i(x) - u_i(x_d)) &\geq \left( \prod_{i=1}^n a_i \right) \cdot \prod_{i=1}^n (u_i(y) - u_i(x_d)) \\ \Leftrightarrow \prod_{i=1}^n (u_i(x) - u_i(x_d)) &\geq \prod_{i=1}^n (u_i(y) - u_i(x_d)) \end{aligned}$$

Wie verträgt sich die Tatsache, daß  $F^{Nash}$  vom Typ (CNC) ist, also keinen interpersonellen Vergleich anstellt, mit der Aussage von Satz 5.1, daß sich ohne interpersonellen Nutzenvergleich kein geeignetes SWFl konstruieren läßt? Die Antwort lautet: Das Nash SWFl erfüllt zwar alle anderen Bedingungen von Satz 5.1, aber nicht die Binary Independence ( $I_u^2$ ). Es ist nämlich klar, daß die soziale Anordnung zweier Alternativen  $x$  und  $y$  nicht nur von den Nutzenvektoren  $(u_1(x), \dots, u_n(x))$  und  $(u_1(y), \dots, u_n(y))$  abhängt, sondern auch von den individuellen Nutzenwerten des status quo  $x_d$ .

Wie läßt sich die Wahl derjenigen Alternative mit dem höchsten Nutzenprodukt begründen?

## 6.1 Die Nash Bargaining Lösung

Stellen wir uns zwei Agenten vor, die eine gewisse Menge von Gütern besitzen und über eine Verteilung dieser Güter verhandeln.<sup>11</sup> Um ein konkretes Beispiel vor Augen zu haben, nehmen wir an, Agent 1 besitzt die beiden unteilbaren Güter  $\alpha$  und  $\beta$ , Agent 2 besitzt  $\gamma$  und  $\delta$ . Bezeichnen wir diese Anfangsallokation mit  $x_d$  und den zugehörigen Nutzenvektor mit  $d = (d_1, d_2)$ . Die Idee ist, daß jeder seine Anfangsausstattung behält, falls bei der Verhandlung keine Einigung zustande kommt. Daher nennen wir  $d$  auch den “disagreement” Nutzen, oder “Drohpunkt”.

Es bezeichne  $X = \{x_1, \dots, x_m\}$  die Menge der möglichen Allokationen. Zusätzlich zu diesen *deterministischen* Allokationen, können wir auch *Lotterien* über diesen Allokationen betrachten. Es bezeichne  $\mathcal{L}(X)$  die Menge aller Lotterien über  $X$ . Ein typisches Element  $p \in \mathcal{L}(X)$  ist also eine Liste  $p = (x_1; p_1, \dots, x_m; p_m)$  mit  $\sum_{j=1}^m p_j = 1$ , wobei  $p_j$  die Wahrscheinlichkeit dafür bezeichnet, daß die Allokation  $x_j$  realisiert wird. Wir nehmen an, daß die Präferenzen der Agenten die Erwartungsnutzenform haben. D.h. für jeden Agenten  $i$  existiert eine Funktion  $u_i : X \rightarrow \mathbf{R}$  derart, daß

$$p \succeq_i q \Leftrightarrow \sum_{j=1}^m p_j \cdot u_i(x_j) \geq \sum_{j=1}^m q_j \cdot u_i(x_j). \quad (6.1)$$

Die Nutzenfunktionen  $u_i$  in Formel (6.1) heißen von-Neumann-Morgenstern Nutzenfunktionen und sind bis auf positive affine Transformationen eindeutig bestimmt. Mit anderen Worten, die  $u_i$  sind kardinale Nutzenfunktionen.

<sup>11</sup>In diesem Kapitel betrachten wir der Einfachheit halber nur zwei Agenten. Alle Resultate lassen sich aber ohne Schwierigkeiten für eine beliebige Anzahl von Agenten verallgemeinern.

Schließt man die Möglichkeit von Lotterien über Allokationen mit ein, so sieht die Nutzenmöglichkeitenmenge eines Verhandlungsproblems typischerweise folgendermaßen aus:

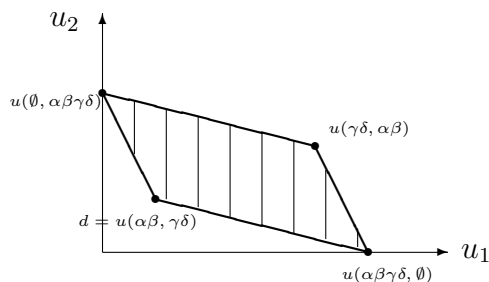


Abb.24: Nutzenmöglichkeitenmenge eines Verhandlungsproblems

Man beachte, daß sich beide Agenten in Abb. 24 durch Tausch ihrer Anfangsausstattung verbessern würden. Die Erwartungsnutzenannahme garantiert, daß jeder Nutzenvektor in der *konvexen Hülle* der deterministischen Allokationen durch eine geeignete Lotterie erreichbar ist. Zusätzlich nehmen wir noch “free disposal” an, d.h. mit jedem erreichbaren Nutzenvektor ist auch jeder komponentenweise geringerer Nutzenvektor erreichbar. Dies motiviert die folgende Definition.

**Definition** Ein *Verhandlungsproblem* ist ein Paar  $(S, d)$ , wobei  $S \subseteq \mathbf{R}^2$  konvex und komprehensiv<sup>12</sup> ist, und  $d \in S$ . Außerdem nehmen wir an, daß  $S$  nach oben beschränkt ist und mindestens einen Punkt  $s = (s_1, s_2)$  enthält derart, daß  $s_1 > d_1$  und  $s_2 > d_2$ .

Unter einer “Lösung” eines Verhandlungsproblems verstehen wir einfach die Zuordnung eines Nutzenvektors in  $S$  für jedes Paar  $(S, d)$ :

**Definition** Eine *Verhandlungslösung* ist eine Abbildung  $F$ , die jedem Verhandlungsproblem  $(S, d)$  ein Element  $F(S, d) \in S$  zuordnet.

Beachte, daß eine Verhandlungslösungen in diesem Sinne keine soziale Präferenzordnung definiert (so wie ein SWF1), sondern nur *eine* “beste” Alternative auswählt.

Welche Eigenschaften einer Lösung sollte man sinnvollerweise fordern?

<sup>12</sup>D.h. falls  $s = (s_1, s_2) \in S$ , so ist auch jeder Punkt  $t = (t_1, t_2)$  mit  $t \leq s$  ein Element von  $S$ . Wir benutzen hier, wie auch im Text, die Notation  $t \leq s$  für  $(t_1 \leq s_1$  und  $t_2 \leq s_2)$ .

**Bedingung (P)** (Pareto) Für jedes  $(S, d)$  soll  $F(S, d)$  effizient in  $S$  sein: es darf keinen Punkt  $s \in S$  geben mit  $s \geq F(S, d)$  und  $s \neq F(S, d)$ .

**Bedingung (S)** (Symmetrie) Falls  $d = 0$  und  $S$  achsensymmetrisch ist, so soll  $F_1(S, 0) = F_2(S, 0)$  gelten.

Die Bedingung der Symmetrie besagt, daß beide Agenten denselben Nutzen erhalten sollen, falls das Problem symmetrisch ist. Von einem normativen Standpunkt ist dies eine sehr plausible Forderung; von einem positiven Standpunkt reflektiert die Bedingung die Annahme, daß beide Agenten dieselbe "Verhandlungsmacht" haben.

Die folgende Bedingung besagt, daß die Lösung nicht von der Wahl der von-Neumann-Morgenstern Nutzenfunktionen abhängen soll. Die Bedingung ist analog zur Bedingung (CNC) aus Kapitel 5.

**Bedingung (S.INV)** (Skaleninvarianz) Betrachte für  $a_1, a_2 > 0$  und beliebige  $b_1, b_2 \in \mathbf{R}$  das Problem  $(aS + b, ad + b)$ , wobei  $aS + b := \{(a_1s_1 + b_1, a_2s_2 + b_2) : (s_1, s_2) \in S\}$  und  $ad + b = (a_1d_1 + b_1, a_2d_2 + b_2)$ . Dann soll gelten:  $F(aS + b, ad + b) = aF(S, d) + b$ .

Als letzte Bedingung fordern wir, daß sich bei einer Verkleinerung der Nutzenmöglichkeitenmenge die Lösung nicht verändern soll, falls die ursprüngliche Lösung immer noch erreichbar ist.

**Bedingung (CC)** (Contraction Consistency) Angenommen  $T \subseteq S$  und  $F(S, d) \in T$ . Dann soll gelten:  $F(T, d) = F(S, d)$ .

**Definition** Die *Nash-Lösung*  $F^N(S, d)$  ordnet jedem Verhandlungsproblem den Punkt zu, für den das Produkt der Nutzen unter allen individuell rationalen am größten ist (dabei heißt  $s$  *individuell rational*, falls  $s_i \geq d_i$ ):

$$F^N(S, d) := \arg \max_{s \in S, s \geq d} (s_1 - d_1)(s_2 - d_2).$$

Man prüft leicht nach, daß  $F^N$  alle obigen Bedingungen erfüllt. Umgekehrt gilt der folgende von Nash 1950 bewiesene Satz.

**Satz 6.1** *Es gibt genau eine Verhandlungslösung, die die Bedingungen (P), (S), (S.INV) und (CC) erfüllt, nämlich  $F^N$ .*

**Beweis** Betrachte ein beliebiges Verhandlungsproblem  $(S, d)$  und die zugehörige Nash-Lösung  $F^N(S, d)$ . Betrachte eine affine Transformation  $(\tilde{S}, \tilde{d})$ , die den Drohpunkt  $d$  in den Nullpunkt  $\tilde{d} = (0, 0)$  überführt und den Punkt  $F^N(S, d)$  in den Punkt  $(1, 1)$ . Da  $F^N$  die Bedingung (S.INV) erfüllt, muß

gelten  $F^N(\tilde{S}, 0) = (1, 1)$ . Da  $F^N(\tilde{S}, 0)$  das Produkt  $s_1 \cdot s_2$  in  $\tilde{S}$  maximiert, muß die Steigung der Paretogrenze von  $\tilde{S}$  im Punkt  $(1, 1)$  gleich  $-1$  sein. Betrachte das Dreieck  $R$ , wie in Abbildung 25 gezeigt:

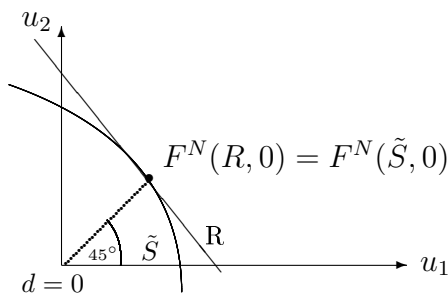


Abb.25: Die Verhandlungsprobleme  $\tilde{S}$  und  $R$

Jede Verhandlungslösung  $F$ , die die Bedingungen (P) und (S) erfüllt, muß für  $(R, 0)$  die Lösung  $F(R, 0) = (1, 1)$  vorschreiben, da  $R$  symmetrisch ist. Insbesondere muß also die Lösung  $F$  mit  $F^N$  auf  $R$  übereinstimmen. Wegen Bedingung (CC) muß aber auch  $F(\tilde{S}, 0) = (1, 1) = F^N(\tilde{S}, 0)$  gelten, und schließlich wegen (S.INV) auch  $F(S, d) = F^N(S, d)$ . Mit anderen Worten, jede Lösung, die die vier Bedingungen erfüllt, muß auf dem beliebig vorgegebenen Problem  $(S, d)$  mit der Nash-Lösung übereinstimmen. Aber das heißt:  $F$  ist die Nash-Lösung.

## 6.2 Andere Verhandlungslösungen

Ist die Nash-Lösung ein ernstzunehmender Vorschlag für eine "gerechte" Lösung eines Verhandlungsproblems? Die problematischste Bedingung scheint die Bedingung (CC) zu sein. Tatsächlich ist bei (CC) (im Gegensatz z.B. zur Bedingung (S)) nicht klar, welches "ethische" Prinzip diese Bedingung rechtfertigen könnte. Betrachten wir das folgende Beispiel. Gegeben sei das symmetrische Verhandlungsproblem  $(S, 0)$  wie in Abbildung 26 gezeigt. Man kann sich z.B. vorstellen, daß bestimmt werden soll, welcher der beiden Agenten ein unteilbares Gut  $\alpha$  mit  $u_1(\alpha) = u_2(\alpha) = 1$  erhalten soll. Jede Verhandlungslösung  $F$ , die die Bedingungen (P) und (S) erfüllt, schreibt in dieser Situation die eingezeichnete Lösung  $F(S, 0) = (\frac{1}{2}, \frac{1}{2})$  vor. Dies entspricht dem (ex-ante) Erwartungsnutzen bei einem Wurf einer fairen Münze. Betrachten wir nun das kleinere Problem  $(R, 0)$  in Abb. 26. Die Bedingung (CC) verlangt, daß die Lösung unverändert bleibt, d.h. auch beim Problem  $(R, 0)$  wäre die Lösung

der Punkt auf der Winkelhalbierenden. Man könnte aber argumentieren, daß dies nicht fair ist, denn Agent 1 erhält nun den für ihn maximal möglichen Nutzen, während Agent 2 dieselbe Konzession wie vorher machen muß.

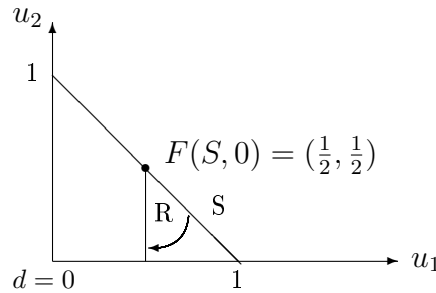


Abb.26: Die Bedingung (CC)

**Definition** Die *Kalai-Smorodinsky-Lösung*  $F^{KS}$  ordnet jedem Verhandlungsproblem den Punkt auf der Pareto­grenze zu, bei dem das Verhältnis der Nutzen genauso groß ist wie im "Idealpunkt"  $I(S, d)$ , wobei  $I_i(S, d)$  der maximale Nutzen für Agent  $i$  unter allen individuell rationalen Punkten von  $(S, d)$  ist (vgl. Abbildung 27).

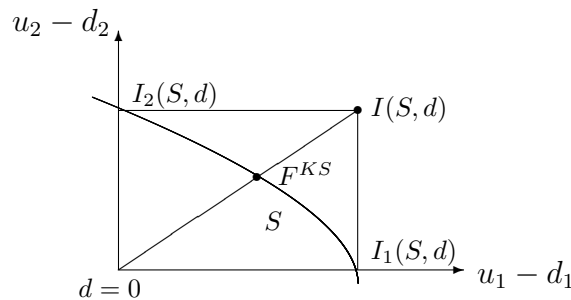


Abb.27: Die Kalai-Smorodinsky-Lösung

**Bedingung (I.MON)** (Individuelle Monotonie) Es seien  $(S, d)$  und  $(T, d)$  zwei Verhandlungsprobleme mit  $T \subseteq S$ . Falls  $I_1(S, d) = I_1(T, d)$ , so soll gelten:  $F_2(S, d) \geq F_2(T, d)$ ; analog, falls  $I_2(S, d) = I_2(T, d)$ , soll gelten:  $F_1(S, d) \geq F_1(T, d)$ .

Die Bedingung (I.MON) besagt also, daß sich bei einer Vergrößerung der Nutzenmöglichkeiten der Nutzen eines Agenten nicht verringern darf, falls sich der maximale Nutzen des anderen Agenten nicht vergrößert.

**Satz 6.2** *Es gibt genau eine Verhandlungslösung, die die Bedingungen (P), (S), (S.INV) und (I.MON) erfüllt, nämlich  $F^{KS}$ .*

**Beweis** Man prüft leicht nach, daß  $F^{KS}$  alle genannten Bedingungen erfüllt. Betrachte ein beliebiges Verhandlungsproblem  $(S, d)$ . Transformiere  $(S, d)$  affin so, daß für das entstehende Problem  $(\tilde{S}, \tilde{d})$  gilt:  $\tilde{d} = (0, 0)$  und  $I(\tilde{S}, \tilde{d}) = (1, 1)$ . Bezeichne die Kalai-Smorodinsky-Lösung von  $(\tilde{S}, 0)$  mit  $(t, t)$ . Nun betrachte das kleinere Problem  $(R, 0)$  wie in Abbildung 28 gezeigt. Für jede Lösung  $F$ , die (P) und (S) erfüllt, muß gelten  $F(R, 0) = (t, t)$ . Wegen (I.MON) muß aber auch  $F(\tilde{S}, 0) = (t, t)$  gelten (weil sich der Idealpunkt beim Übergang von  $R$  zu  $\tilde{S}$  nicht verändert). Insbesondere stimmt also jede Lösung, die (P), (S) und (I.MON) erfüllt, auf  $(\tilde{S}, 0)$  mit  $F^{KS}$  überein. Wegen (S.INV) muß dies aber auch für das ursprüngliche Problem  $(S, d)$  gelten.

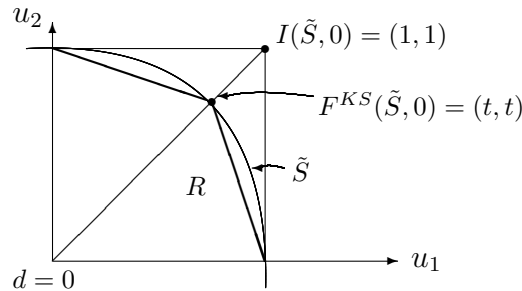


Abb.28: Die Verhandlungsprobleme  $\tilde{S}$  und  $R$

Die Intuition hinter der Bedingung (I.MON) ist, daß eine Vergrößerung der Nutzenmöglichkeitenmenge (unter bestimmten Umständen) für einen Agenten nicht schädlich sein darf. Formuliert man diese Intuition ohne die Einschränkung, daß der maximale Nutzen eines Agenten bei der Vergrößerung unverändert bleibt, so ergibt sich die folgende Bedingung:

**Bedingung (MON)** (Monotonie) Falls  $T \subseteq S$ , so soll gelten:  $F(S, d) \geq F(T, d)$ .

Man kann sich leicht durch entsprechende Beispiele klarmachen, daß weder die Nash-Lösung noch die Kalai-Smorodinsky-Lösung die Bedingung (MON) erfüllen. Gibt es überhaupt eine Verhandlungslösung, die (MON) erfüllt?

**Definition** Betrachte die Menge aller Verhandlungsprobleme  $(S, d)$  mit  $d = (0, 0)$ . Die *egalitäre Lösung*  $F^{egal}$  ordnet jedem solchen Problem den Schnittpunkt der Pareto-Grenze mit der Winkelhalbierenden zu.

**Satz 6.3** Betrachte die Menge aller Verhandlungsprobleme  $(S, 0)$ , wobei  $S$  zusätzlich zu den bisherigen Bedingungen noch folgende Bedingung erfülle: Der Rand von  $S$  enthalte im positiven Orthanten keine Geradenstücke parallel zu den Achsen. Es gibt genau eine Abbildung  $F$ , die jedem solchen Verhandlungsproblem  $S$  einen Punkt  $F(S) \in S$  zuordnet und die Bedingungen (P), (S) und (MON) erfüllt, nämlich  $F^{egal}$ .

**Beweis** Betrachte ein beliebiges Problem  $S$ . Es sei  $F^{egal}(S) = (t, t)$ ; betrachte das Problem  $R$ , das den Punkt  $(t, t)$  mit den Achsabschnitten  $(0, t + \epsilon)$  bzw.  $(t + \epsilon, 0)$  verbindet, wobei  $\epsilon > 0$  genügend klein ist, sodaß  $R \subseteq S$ . Solch ein  $\epsilon$  existiert immer, da der Rand von  $S$  keine zu den Achsen parallelen Geradenstücke enthält (siehe Abbildung 29).

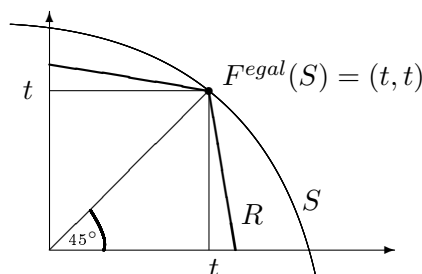


Abb.29: Die Verhandlungsprobleme  $S$  und  $R$

Wegen den Bedingungen (P) und (S) muß die Lösung  $F$  von  $R$  der Punkt  $(t, t)$  sein. Wegen (MON) darf der Nutzen keines Agenten beim Übergang von  $R$  zu  $S$  sinken. Dies ist aber nur möglich, falls die Lösung  $F(S)$  in  $S$  ebenfalls der Punkt  $(t, t)$  ist. Aber das heißt, die Lösung  $F$  muß mit der egalitären Lösung  $F^{egal}$  übereinstimmen.

**Bemerkung 1** Man kann leicht sehen, daß die egalitäre Lösung nicht die Bedingung (S.INV) erfüllt. Da aber die einzige Lösung, die (P), (S) und (MON) erfüllt, die egalitäre Lösung ist, folgt, daß es *keine* Verhandlungslösung gibt, die sowohl (P), (S), (S.INV) als auch (MON) erfüllt.

**Bemerkung 2** Die egalitäre Lösung entspricht natürlich dem maximin Prinzip angewandt auf Verhandlungsprobleme (mit  $d = (0, 0)$ ). Die Eigenschaft, daß bei einer Erweiterung der Nutzenmöglichkeiten (z.B. durch eine Erhöhung der Anfangsausstattung) kein Agent schlechter gestellt wird, ist auch in anderen Kontexten eine charakteristische Eigenschaft des maximin Prinzips, die dieses von anderen Methoden der Bewertung (z.B. einer utilitaristischen)



unterscheidet (siehe z.B. Roemer (1996, Th. 1.9, S.45 und Th. 3.2, S.102)).

**Bemerkung 3** Natürlich läßt sich auch eine “utilitaristische” Verhandlungslösung definieren:  $F^{util}(S, 0) := \arg \max_{s \in S, s \geq 0} (s_1 + s_2)$  (im Fall  $d = (0, 0)$ ). Wie die egalitäre Lösung erfüllt diese aber nicht die Bedingung (S.INV) und kann daher nicht als “echte” Verhandlungslösung (im Sinne von Nash) aufgefaßt werden.