

# KOOPERATION IM GEFANGENEN-DILEMMA DURCH ENDOGENES LERNEN

Carsten Vogt

*Inauguraldissertation zur  
Erlangung des Grades eines Doktors der  
Wirtschaftswissenschaft durch die  
Fakultät für Wirtschaftswissenschaft  
an der Otto-von-Guericke-Universität*

# INHALT

<b>1.</b>	<b>Einleitung</b>	<b>1</b>
<b>2.</b>	<b>Kooperation in Dilemmata: Experimentelle Evidenz</b>	<b>7</b>
2.1	Andreoni/Miller (1991)	8
2.1.1	Darstellung der Studie	8
2.1.2	Kritik	14
2.2	Cooper et al (1996)	16
2.3	Evidenz in Public Good Experimenten	19
2.4	Offerman et al (1996)	23
2.5	Zusammenfassung	25
<b>3.</b>	<b>Robert Frank : Die strategische Funktion von Emotionen</b>	<b>28</b>
3.1	Eine nicht-technische Diskussion	28
3.2	Frank II: Eine modelltheoretische Diskussion	34
3.3	Kritik des Frank-Modells	40
3.3.1	Die Harrington Kritik	41
3.3.2	Die experimentelle Perspektive: Sind Signale wichtig?	46
3.3.3.1	Zuckerman et al (1981)	46
3.3.3.2	Ökonomische Untersuchungen	49
3.4	Robson (1990)	56
3.5	Ausblick	57
<b>4.</b>	<b>Ein Modell des Erfahrungslernens für random matching games</b>	<b>58</b>
<b>5.</b>	<b>Das Gefangenen Dilemma mit einer lernfähigen Mutante</b>	<b>66</b>
5.1	Payoffs	66
5.2	Konzeptionelle Vorüberlegungen	69
5.3	Basis-Modell: Subpopulationen finiter Größe	71
5.4	Diskussion der Resultate	79
5.5	Eine alternative Variante des Basismodells	100
5.6	Interpretation und Diskussion der Ergebnisse	106
<b>6.</b>	<b>Robustheit der Resultate</b>	<b>107</b>
6.2	Diskussion	118
6.3	Appendix	125
<b>7.</b>	<b>Überleben von Kooperation in einer stochastischen Welt</b>	<b>133</b>
7.1	Motivation	133
7.2	Ein stochastisches Modell der Kooperation	135
7.3	Diskussion	156
7.4	Ein Beispiel: Die Best-Reply Dynamik	161
7.5	Diskussion	170
7.6	Appendix	175
<b>8.</b>	<b>Zusammenfassung</b>	<b>180</b>
<b>9.</b>	<b>Ausblick</b>	<b>186</b>



## 1. Einleitung

Kooperation im Gefangenen Dilemma ist ein ausgesprochen persistentes empirisches Phänomen. Zahllose experimentelle Studien existieren mittlerweile, die diesen Befund reproduzieren konnten: Menschen verhalten sich nicht, wie vom ökonomischen Standardmodell (basierend auf den beiden Annahmen: Rationalität und Eigennutzstreben) prognostiziert. Sowohl in einmaligen als auch wiederholten Dilemmaspielen zeigt sich eine signifikant positive Kooperationsbereitschaft, obwohl Defektion die dominante bzw. teilspielperfekte Strategie wäre.

Schon frühzeitig etablierte sich die Idee, daß Kooperation durch reziprokes, also: auf Gegenseitigkeit beruhendes Verhalten entstehen kann. Dieser Ansatz geht mindestens bis auf Trivers berühmten Aufsatz von 1973 über „reciprocal altruism“ zurück. Hintergrund dieser biologischen Arbeit sind Verhaltensweisen von Tieren wie wechselseitige Fellpflege, Symbiosen mit „Cleaner“-Fischen, Warnrufe von Vögeln, die scheinbar nicht dem eigenen Interesse entsprechen und die nicht mit dem Modell der kin-selection (Hamilton 1964) erklärt werden können. Von Vampirfledermäusen etwa werden erstaunliche kooperative Leistungen berichtet. Es ist von dieser Spezies bekannt, daß ihre Vertreter ohne Säugetierblut nicht länger als zweieinhalb Tage überleben können. Von den nächtlichen Beutezügen kehren im Durchschnitt 10 % der Tiere erfolglos zurück. Satte Vampire füttern dann hungerleidende Artgenossen, ein Verhalten, das unter der Prämisse eigennütziger Fitnessmaximierung kaum erklärlich erscheint (vgl. Sherratt und Roberts (1998)). Die zentrale Idee von Trivers ist, daß sich derartige kooperative Verhaltensweisen möglicherweise in einem wiederholten Interaktionszusammenhang als individuell rational erklären lassen.

In der Tat hat die (ökonomische) Spieltheorie dann einen wesentlichen Beitrag zur Klärung der Frage geleistet, ob kooperative Strategien in wiederholten Spielen gleichgewichtig sein können. Das klassische Folk-Theorem konstatiert, daß sich jeder individuell rationale Auszahlungsvektor in einem unendlich wiederholten Gefangenen-Dilemma als Nash-Gleichgewicht stützen läßt. Aumann und Shapley (1976) sowie Rubinstein (1979) haben gezeigt, daß das Folk-Theorem auch für die weit schärfere Anforderung der Teilspielperfektheit in unendlich wiederholten Spielen ohne Diskontierung gilt. Friedman (1971) sowie Fudenberg und Maskin (1986) zeigen, daß das Folktheorem im wesentlichen auch bei Diskontierung gilt. Der zentrale Gedanke ist auch hier eine Form reziproken Verhaltens. Friedman beispielsweise spezifiziert die gleichgewichtigen Strategien wie folgt: Jede Abweichung von den gleichgewichtigen Strategien wird mit dem Spiel der Nash-Gleichgewichtsstrategien des Basisspiels bestraft. Auf diese Weise kann jeder Auszahlungsvektor, der den nashgleichgewichtigen Auszahlungsvektor des Basis

spiels paretodominiert als teilspielperfektes Gleichgewicht stabilisiert werden. Fudenberg und Maskin nehmen für den 2-Spieler-Fall an, daß Abweichungen vom teilspielperfekten Gleichgewicht mit Minimax-Verhalten für eine bestimmte Anzahl von Spielwiederholungen bestraft wird (Fudenberg und Maskin (1986), Theorem 1). Für den n-Spieler-Fall wird angenommen, daß Bestrafung von Abweichung belohnt werden kann. Es scheint also, daß reziprokes Verhalten (hier: spiele die gleichgewichtige Strategie solange der/die anderen auch Gleichgewicht spielen, aber bestrafe abweichendes Verhalten in geeigneter Weise) eine zentrale Rolle bei der Entstehung von Kooperation in Dilemmasituationen zukommt.

Jedoch, die Folktheoreme für unendlich wiederholte Spiele können der Weisheit letzter Schluß nicht sein. Sowohl in real-life-Situationen als auch in experimentellen Anordnungen hat man es immer mit endlichen Spielen zu tun. Dort aber greift das vertraute Argument der Rückwärtsinduktion mit der Folge, daß im Gefangenen-Dilemma die spieltheoretische Lösung eindeutig Defektion als einziges Gleichgewicht auf allen Stufen des wiederholten Spiels vorsieht. Wir beobachten jedoch regelmäßig kooperatives Verhalten in diesen Situationen. Wie ist dieses robuste Phänomen zu erklären?<sup>1</sup>

Natürlich hat sich die Spieltheorie sehr schnell der Frage angenommen, ob Folktheoreme nicht auch für endlich wiederholte Spiele bewiesen werden können. An dieser Stelle ist als zentrale Arbeit das gang-of-four Modell zu nennen. Kreps et al (1982) zeigen, daß Unsicherheit über den Typ des Gegenspielers Kooperation im GD erklären kann. Nimmt einer der beiden Spieler an, daß sein Gegenüber ein „crazy player“ (etwa ein Tit for Tat Spieler) ist, so kann es sich auch für einen strikt rationalen und eigennützigen Spieler lohnen, selbst wie ein kooperativer Typ zu erscheinen. Die Intuition ist einfach: Verhält sich der rationale Defektor zunächst kooperativ, so können die dadurch erzielten Kooperationsgewinne – bei hinreichend geringer Diskontierung – die Gewinne aus einmaliger Ausbeutung

---

<sup>1</sup> Daneben sind die Folk Theorem-Resultate noch aus einem anderen Grund unbefriedigend: Sie sind nämlich ausgesprochen wenig prediktiv. Man sieht sich in unendlich wiederholten Spielen plötzlich mit einer explosionsartig wachsenden Menge von teilspielperfekten Gleichgewichten konfrontiert, aber die Theorie wiederholter Spiele erlaubt zunächst keine Aussagen darüber, welches der unzähligen Gleichgewichte gespielt wird, bzw. werden sollte. Während also im one-shot-Kontext bezüglich der Erklärung von Kooperation gilt: „nothing goes“ (Pflüger (1997)), wandelt sich dieser Erklärungsnotstand bei unendlicher Wiederholung zu einem (nearly) „anything goes“. Fudenberg und Tirole (1991) sprechen davon, daß das Folk Theorem „zu erfolgreich“ sei. Dieser Euphemismus wird pointiert durch Güth zurecht gerückt, der für die Anwendung des Folk-Theorems in der Industrieökonomik feststellt: „Da das Folk Theorem alle überhaupt zu erwartenden Marktpreisentwicklungen vorhersagt, kann man die Bereitschaft, das Folk Theorem zu akzeptieren, als eine Suizidneigung der Industrieökonomik interpretieren.“ (Güth, 1992, S.287)

des Gegenspielers überwiegen. Es lohnt also auch für einen rationalen und eigenützigen Spieler, eine Reputation als kooperativer Spieler aufzubauen.

Als Stärke des Modells wird dabei gelegentlich aufgefaßt, daß es bereits genügt die Erwartung zu hegen, der Gegner sei „irrational“, während beide Spieler tatsächlich vollständig rational und „selfish“ sein können. Das gang-of-four Modell erlaube es damit, Kooperation durch die minimal denkbare Abweichung von den Annahmen der Standardtheorie zu erklären. Diese Auffassung freilich blieb nicht unwidersprochen. Zum einen erhebt sich die Frage, wie eine solche Inkonsistenz zwischen den beliefs der Spieler und der tatsächlichen Typzusammensetzung einer Population gerechtfertigt werden könnte. Hierauf hat Samuelson (1987) hingewiesen. Langfristig sollte man erwarten, daß Spieler die tatsächliche Typverteilung lernen. Wenn eine Population also tatsächlich aus rationalen Defektoren besteht, dann sollte dieses Faktum in absehbarer Zeit common knowledge sein, und damit träfe wieder die Prognose der Standardtheorie für endlich wiederholte Spiele zu. Andererseits kann man die Frage erheben, inwiefern das Modell die in Experimenten beobachtete Kooperation tatsächlich „erklärt“. Einige stilisierte Fakten der experimentellen Literatur sind durchaus konsistent mit den Vorhersagen des Reputationsmodells, so der im Verlauf wiederholter Spiele sinkende Grad der Kooperationsbereitschaft. Andere Fakten hingegen lassen sich mit der Theorie von Kreps et al unter keinen Umständen rationalisieren, so die signifikant positive Kooperation in Schlußrunden und die ebenfalls signifikant positiven Kooperationsraten in one-shot GD-Experimenten. In diesen Situationen besteht kein Anlaß zum Reputationserwerb. Warum aber beobachtet man jedoch auch bei erfahrenen Spielern (Andreoni und Miller 1991, Cooper et al 1996) persistent kooperatives Verhalten in diesem Kontext?

Andreoni und Miller (1991) haben eine ebenso naheliegende wie natürliche Reinterpretation des Reputationsmodells vorgeschlagen. Anstatt von der schwer zu rechtfertigenden fiktiven Rolle von der realen Existenz kooperativer Typen ausgegangen werden, das heißt von Spielern, die mit einer Präferenz (oder psychologisch gesprochen: Disposition) für Kooperation ausgestattet sind. Eine solche Sicht der Theorie könnte in der Tat sowohl wesentliche qualitative Aspekte der Beobachtungen aus wiederholten Spielen erklären, hätte aber andererseits auch keinerlei Schwierigkeiten mit Kooperation in one-shot-Situationen.

Ein Blick in die experimentelle Literatur gibt in der Tat eindeutige Hinweise darauf, daß reale Experimentalpopulationen äußerst heterogen sind. Menschen scheinen sich systematisch hinsichtlich ihrer grundsätzlichen Verhaltensorientierungen zu unterscheiden. Während lange Zeit in der experimentellen Ökonomik eine ausschließlich aggregierte Betrachtung im Vordergrund stand zeigen neuere Studien, die sich explizit mit der Analyse von Individualverhalten beschäftigen, daß es

insgesamt zwei bis drei mehr oder weniger große Kategorien von Verhaltenstypen zu geben scheint, einen (schwach oder bedingt) kooperativen Typus, einen stark defektiven Typus und möglicherweise auch einen unbedingt kooperativen Typ, der immer kooperiert, unabhängig vom Verhalten seiner Mitspieler. Es erhebt sich damit die Frage, wie eine solche Koexistenz unterschiedlicher Typen zu erklären ist.

Als Ansatz bieten sich evolutionäre Modelle an. Robert Frank (1988) hat eine bemerkenswerte Theorie vorgestellt, in deren Zentrum die strategische Funktion von Emotionen steht und die letztlich genau auf die Frage der Koexistenz unterschiedlicher Verhaltensdispositionen abzielt. Wesentlich für seinen Ansatz ist die Annahme der Existenz eines hinreichend „fälschungssicheren“ Signals, welches mit ausreichender Verlässlichkeit Auskunft über die Typzugehörigkeit eines Gegenübers gibt. Abhängig vom empfangenen Signalwert kann dann das Verhalten konditioniert werden. Frank (1987) kann im Rahmen eines formalen Modells (scheinbar) plausibel machen, daß es zu einem stabilen Gleichgewicht kommen kann, in dem ehrliche Typen (solche, die im GD ein Kooperationsversprechen halten) und unehrliche Typen koexistieren.

Unglücklicherweise ist das Modell mit fundamentalen Mängeln behaftet, wie Harrington (1989) deutlich gemacht hat. Die Resultate basieren letztlich auf einer kaum zu rechtfertigenden ad hoc Annahme, nämlich der bereichsweise perfekter Signale (vgl. 3.3.1). Weicht man diese auf, bricht das Frank-Modell zusammen. Auch verwandte Modelle wie das „secret-handshake-Modell“ von Robson (1990) können letztlich den oben skizzierten experimentellen Befund nicht erklären.

Vor diesem Hintergrund schlägt die vorliegende Arbeit einen alternativen Weg ein. Es spricht vieles dafür, daß zur Erklärung von Kooperation in Dilemmasituationen Reziprozität<sup>2</sup> unverzichtbar ist. Reziprokes Verhalten wird hier jedoch nicht ermöglicht aufgrund der Annahme der Existenz hinreichend verlässlicher Signale. Vielmehr wird dieses implementiert auf der Basis eines lerntheoretischen Ansatzes. Wir führen, anknüpfend an Amann und Yang (1994) einen lernfähigen kooperativen Typ in das Gefangenen-Dilemma ein, der die Zusammensetzung der Population mit der er interagiert, sukzessive im Verlauf des matching games erlernt. Auf der Grundlage des so generierten Erfahrungswissens konditioniert die Mutante ihr Verhalten, insbesondere wird eine Ausbeutungsaversion als wirksames Motiv unterstellt, das heißt die Mutante verweigert die Interaktion mit be

---

<sup>2</sup> Dabei soll unter Reziprozität ganz allgemein ein Verhalten verstanden werden, das auf Gegenseitigkeit beruht, das also, wie etwa die Strategie Tit for Tat, Kooperation mit Kooperation belohnt, Defektion aber mit Entzug der Kooperation bestraft. Ob ein derartiges Verhalten dabei strategisch oder intentional begründet ist, spielt für unsere Belange zunächst keine Rolle.

kannten Defektoren. Wir zeigen dann, daß für geeignete Werte der Modellparameter ein global stabiles, inneres Gleichgewicht mit positiven Populationsanteilen aller drei Strategien existiert. Die Ergebnisse des Basismodells sind im wesentlichen robust gegenüber der Einführung einer weiteren lernfähigen, jedoch grundsätzlich defektiv orientierten Mutante. Es wird außerdem gezeigt, daß ein Wechsel des Modellrahmens von einem deterministischen dynamischen System hin zu einer stochastischen Welt das Überleben von Kooperation nicht zerstört.

Im einzelnen gliedert sich die vorliegende Arbeit wie folgt. Kapitel 2 referiert den experimentellen Befund bezüglich des Verhaltens in Dilemmasituationen. Besondere Aufmerksamkeit wird dabei dem lange Zeit wenig beachteten Individualverhalten geschenkt. Zusätzlich zu 2-Personen-GD Experimenten (Kapitel 2.1 und 2.2) erbringen Public-good Experimente Belege für die Existenz stabiler und systematisch verschiedener Verhaltensdispositionen (Kapitel 2.3). Schließlich wird mit Offerman et al (1996) eine Studie referiert (Kapitel 2.4), die mit sozialpsychologischen Techniken der Typenmessung zu analogen Ergebnissen kommt. Kapitel 3 widmet sich der Darstellung der Theorie von Frank, die als der prominenteste Versuch angesehen werden kann, die Koexistenz unterschiedlicher verhaltensmäßiger Dispositionen zu rationalisieren. Kapitel 3.1 liefert eine nicht-technische Diskussion der zentralen Ideen, Kapitel 3.2 referiert eine modelltheoretische Präzisierung. Kapitel 3.3 erläutert die Harrington-Kritik am Frank-Modell und zeigt, daß Franks Ergebnisse, „while fascinating, are based upon an implausible foundation“ (Harrington 1989, S. 588). Kapitel 3.4 schließlich wirft von empirischer Warte aus Zweifel an Franks Theorie auf. Zentral für die Theorie ist die Annahme hinreichend fälschungssicherer Signale, die auch mit ausreichender Verlässlichkeit von anderen rezipiert werden können. Experimentelle Untersuchungen legen jedoch den Verdacht nahe, daß reale Menschen nur in sehr begrenztem Maße in der Lage sind, den „Typ“ eines anderen anhand bestimmter physiologischer Signale zu erkennen. Im Kapitel 4 wird das Fundament für die gesamte weitere Analyse gelegt, nämlich die Formulierung eines Modells des Erfahrungslernen, das in matching games implementierbar ist. Kapitel 5 enthält das Basismodell: Wir führen eine lernfähige, kooperative Mutante als dritte Strategie in das GD ein und wenden dann Standardinstrumente der evolutionären Spieltheorie an, um das langfristige Gleichgewicht der Population zu ermitteln. Es zeigt sich, daß für hinreichend großen Lernerfolg der mutanten Strategie ein global stabiles Dreitypen-Gleichgewicht existiert. Kapitel 6 diskutiert die Frage der Robustheit der Resultate des Basismodells gegenüber einer weiteren mutanten Strategie. Gewisse konzeptionelle Probleme, die in Kapitel 5.2 diskutiert werden, wie auch die Kritik etwa von Foster und Young (1992) an deterministischen evolutionären Modellen, lassen es ratsam erscheinen, die Frage des Überlebens von



Kooperation in einer stochastischen Welt zu untersuchen. Kapitel 7 widmet sich dieser Aufgabe. Als zentral für die Ergebnisse erweist sich die Einführung der Möglichkeit der Mutation von Spielern. In einem äußerst allgemein gehaltenen Modellrahmen, der sich an der Arbeit von Kandori et al (1993) orientiert, in dem die selektive Komponente des evolutionären Prozesses lediglich durch eine monotone Beziehung zwischen resultierendem Payoff und reproduktivem Erfolg charakterisiert ist, werden strukturell analoge Ergebnisse zu den Resultaten des deterministischen Modells abgeleitet. Kapitel 7.4 verdeutlicht schließlich die Logik des „mutation counting“ anhand des Beispiels der Best-Response-Dynamik.

## 2. Kooperation in Dilemmata: Experimentelle Evidenz

Es zählt zu den am besten gesicherten Ergebnissen psychologischer als auch ökonomischer experimenteller Forschung, daß Versuchspersonen sich in Dilemmaspielen, entgegen der Prognose der ökonomischen Standardtheorie, kooperativ verhalten. Diese Beobachtung ist in einer kaum noch überschaubaren Zahl von Experimenten, sowohl für 2-Personen-Gefangenen-Dilemma-Spiele als auch für Öffentliche-Gut-Spiele<sup>3</sup>, wiederholt gemacht worden, und sie erweist sich als ausgesprochen robust gegenüber Änderungen des experimentellen Designs. Wohl schwankt das Ausmaß der beobachteten Kooperation, das qualitative Phänomen einer signifikant positiven Kooperationsbereitschaft konnte jedoch, trotz teils intensiver Bemühungen vor allem von ökonomischer Seite (vgl. Kim und Walker (1984)), nicht zum Verschwinden gebracht werden.

Der typische Standardbefund für endlich wiederholte Dilemmaspiele läßt sich wie folgt charakterisieren. Das Ausmaß kooperativen Verhaltens ist in den ersten Spielrunden am größten, mit Kooperationsraten, die sich durchaus in der Größenordnung zwischen 40 und 60% aller Beobachtungen bewegen. Die gemessene Kooperation fällt dann mit zunehmender Anzahl von Spielwiederholungen ab und erreicht üblicherweise in der Schlußrunde ihr globales Minimum. Dennoch wird auch in der letzten Runde in statistisch signifikantem Maße kooperiert, was natürlich in besonders krasssem Widerspruch zur spieltheoretischen Prognose für endlich iterierte Dilemmata steht. (Vgl. beispielhaft für zahllose weitere Arbeiten Schneider und Pommerehne (1981), Marwell und Ames (1981), Isaac, Walker und Thomas (1984), Isaac und Walker (1988), Isaac, Walker und Williams (1994). Speziell für das 2-Personen-Gefangenen-Dilemma vgl. etwa Andreoni und Miller (1991) sowie Cooper et al (1996)).

Ungewöhnlich lange Zeit dominierte in der experimentellen Literatur eine ausschließlich aggregierte Betrachtungsweise. Fast alle der oben zitierten Arbeiten betrachten lediglich „durchschnittliches Verhalten“, also etwa den durchschnittlichen Beitrag zu einem zu produzierenden öffentlichen Gut im Zeitablauf. Die Fixierung auf aggregierte Daten<sup>4</sup> hat dabei den Blick auf ein bemerkenswertes

---

<sup>3</sup> Für einen Literaturüberblick über GD-Spiele vgl. beispielsweise Dawes (1980). Dort werden auch systematisch Faktoren des Versuchsdesigns identifiziert, die die gemessene Kooperationsbereitschaft beeinflussen, wie etwa das Kommunikationsdesign. Desweiteren Dawes und Thaler (1988). Für Überblicke über Public Good Experimente: Ledyard (1995) und Weimann (1994). Eine frühe Referenz für 2-Personen-GD-Spiele sind Rapoport und Chammah (1965).

<sup>4</sup> Diese wird verständlich, wenn man sich den Ausgangspunkt zumindest der ökonomischen experimentellen Forschung vergegenwärtigt. Aus Sicht der neoklassischen Standardtheorie des

Faktum verstellt, nämlich, daß es gewisse, systematische Unterschiede in den Verhaltensweisen von Versuchsteilnehmern gibt und offensichtlich mehrere Typen existieren, die sich hinsichtlich ihrer Kooperationsbereitschaft deutlich unterscheiden. In den folgenden Abschnitten werden einige Arbeiten vorgestellt, die sich systematisch der Untersuchung des Individualverhaltens in Dilemmaspielen, vor allem 2-Personen-GD-Spielen widmen.

## 2.1 Andreoni/Miller (1991)

### 2.1.1 Darstellung der Studie

Als eine erste wichtige Arbeit, die sich ausdrücklich auch mit der Frage der Existenz unterschiedlicher Typen beschäftigt, sei die Arbeit von Andreoni und Miller dargestellt. Hintergrund der Arbeit ist das sogenannte „gang-of-four“-Modell von Kreps et al (1982), das als eine der führenden Theorien zur Erklärung von kooperativem Verhalten im 2-Personen-GD gelten kann. Kreps et al weichen bekanntermaßen die common knowledge Annahme bezüglich vollständiger Rationalität beider Spieler auf, um so den unvermeidbaren Konsequenzen der Logik der Rückwärtsinduktion im Standardmodell zu entgehen. Nimmt nämlich einer der beiden Spieler – mit einer möglicherweise auch nur kleinen Wahrscheinlichkeit – an, daß sein Gegenspieler sich nicht vollständig rational verhält, so kann es auch für einen vollständig rationalen Spieler lohnen, zumindest in den ersten Runden des endlich wiederholten GD zu kooperieren. Die Logik des Modells ist intuitiv sofort am Beispiel eines Tit for Tat (TfT) Spielers erkennbar. Nehme Spieler 1, ein rationaler Defektor an, daß Spieler 2 ein TfT-Typ ist. Dann weiß 1, daß defektives Verhalten seinerseits die Defektion des TfT Spielers nach sich zieht. Wenn Spieler 1 nun die langfristigen Kosten aus defektivem Verhalten hinreichend hoch bewertet (nämlich höher als den kurzfristigen Ausbeutungsgewinn aus Defektion), dann lohnt es sich auch für ihn, zumindest bis zu einem bestimmten Zeitpunkt, zu kooperieren. Dieses Resultat ist unter dem Schlagwort der Reputationshypothese in die Literatur eingegangen: Es lohnt sich (unter Umständen) auch für einen strikt rationalen Defektor, wie ein Kooperationspartner zu erscheinen.

Auf den ersten Blick erscheint es als unbezweifelbare Stärke des Reputationsmodells, daß lediglich ein Spieler einen belief halten muß, sein Gegenspieler sei nicht vollständig rational. Insbesondere erlaubt das gang-of-four-Modell ein weiteres

---

homo oeconomicus besteht keinerlei Anlaß davon auszugehen, daß Menschen bezüglich ihrer Verhaltensdisposition systematisch differieren. Vielmehr sollten alle Vpn unterschiedslos vollständig rationale und eigennützige Entscheider sein.

Festhalten an der strikten Rationalitätsannahme der Standardtheorie, denn: Es ist keineswegs erforderlich, daß einer der beiden Spieler tatsächlich ein TftT Spieler ist. Es ist vielmehr ausreichend, um einmütige Kooperation in den ersten Runden des wiederholten GD zu erzeugen, daß einer der beiden Spieler einen entsprechenden belief hält, tatsächlich aber beide rationale Defektoren sind.

Was zunächst als Stärke erscheint, kann jedoch auch als die entscheidende Schwäche des Modells interpretiert werden. Die eben vorgeschlagene Interpretation des Modells verlangt von den Spielern nämlich, beliefs zu halten, die mit der tatsächlichen Typenzusammensetzung der Population nicht konsistent sind. Auf diesen Umstand hat Larry Samuelson (1987) hingewiesen, und es erhebt sich natürlich die Frage, wie eine derartige Inkonsistenz, jedenfalls über längere Zeit, gerechtfertigt werden könnte. Aus der Perspektive sowohl lern- wie auch evolutionstheoretischer Ansätze würde man regelmäßig argumentieren, daß eine solche Inkonsistenz in langer Frist beseitigt wird. Andreoni und Miller schlagen daher vor, einfach von der realen (und nicht nur fiktiven) Existenz kooperativer Typen<sup>5</sup> auszugehen. Das Ergebnis des Reputationsmodells bliebe erhalten, denn auch gegenüber einem real existierenden kooperierenden Typ lohnt es sich natürlich weiterhin für einen rationalen Defektor, eine Reputation als Kooperationspartner aufzubauen. Der Vorteil des vorgeschlagenen Weges besteht in der Beseitigung der beschriebenen und schwer zu rechtfertigenden Inkonsistenz. Ziel der Untersuchung von Andreoni und Miller ist es nun, einerseits die Reputationshypothese von Kreps et al experimentell zu testen. Andererseits wird der Frage nach der „richtigen“ Interpretation des gang-of-four-Modells nachgegangen, also der Frage, ob kooperative Spielertypen tatsächlich existieren.

Das Andreoni/Miller-Experiment besteht aus insgesamt vier Anordnungen.

i) Partner

14 Versuchspersonen spielten 20-mal hintereinander ein endlich wiederholtes 2 Personen GD über 10 Runden. Am Beginn eines jeden 10-Runden Spiels wurden die Paarungen zufällig neu bestimmt. Über die jeweiligen

---

<sup>5</sup> Die Arbeit von Andreoni und Miller leidet unter einer gewissen terminologischen Unschärfe. Statt von kooperativen Typen sprechen die Autoren regelmäßig von Altruisten. Ohne allzu tief in die Auseinandersetzung um die Definition von Altruismus einzusteigen, kann doch mindestens gesagt werden, daß darunter eine Form uneigennütigen Verhaltens verstanden werden sollte. Was Andreoni und Miller wie ja auch Kreps et al vorschwebt, ist aber die (fiktive oder reale) Existenz von bedingt kooperativen Typen. Statt Altruismus wäre also hier Reziprozität der treffendere Begriff. Reziprokes Verhalten ist aber nicht notwendigerweise uneigennützig. Wir ziehen es daher im weiteren vor, von kooperativen Typen zu sprechen.

- 10 Runden blieben die Paarungen aber konstant, und die Spieler wußten dies.
- ii) Fremde  
Die Vpn wurden nunmehr in jeder der insgesamt 200 Spielrunden zufällig neu gematcht, d.h. die Spieler spielten 200 mal hintereinander ein one-shot GD.
  - iii) Computer50  
Die Versuchsanordnung war identisch zu der Partner-Anordnung, jedoch mit folgender Änderung: Es bestand nunmehr eine 50%-Chance gegen einen Computer zu spielen, der auf die Tit for Tat Strategie programmiert war. Diese Information war den Spielern bekannt.
  - iv) Computer0  
Computer0 war identisch zu Computer50, jedoch betrug die Wahrscheinlichkeit, gegen die Maschine zu spielen, nunmehr lediglich 1/1000, also 0,1%.

Andreoni/Miller leiten nun folgende Hypothesen ab. Wenn die Reputationshypothese richtig ist, dann ist auf jeden Fall zu erwarten, daß Partner mehr kooperieren als Fremde. Nur wenn die Zusammensetzung der Paare über einen gewissen Zeitraum konstant ist, macht Reputationsbildung überhaupt Sinn. Bezüglich der Fremden-Anordnung ist die theoretische Prognose natürlich, daß keine Kooperation beobachtet werden sollte. In einem one-shot GD prognostiziert auch das Reputationsmodell von Kreps et al, nicht zu kooperieren. Zumindest würde man erwarten, daß Vpn dies im Spielverlauf schließlich lernen. Sollte es zu Beginn der 200 one-shot GD also zu kooperativem Verhalten kommen, so würde man dennoch erwarten, daß nach einer hinreichenden Anzahl von Wiederholungen die spieltheoretische Lösung gelernt wurde.

Bezüglich der „Maschinen-Anordnungen“ argumentieren Andreoni/Miller wie folgt. Vpn in Computer50 haben gegenüber Partnern die größere Gewißheit, daß sie gegen einen kooperativen Automaten spielen können. Das Ausmaß der Kooperation sollte in Computer50 daher höher ausfallen. Der Sinn der Computer0-Anordnung erschließt sich aus folgender Überlegung: Wenn tatsächlich in Computer50 mehr Kooperation als in der Partner-Anordnung beobachtet wird, dann könnte dies auch daher rühren, daß nunmehr Tit for Tat common knowledge ist. In Computer0 ist Tit for Tat ebenfalls common knowledge, jedoch ist die Wahrscheinlichkeit gegen Tft zu spielen faktisch Null. Wenn common knowledge von Tft also ausreicht, um Kooperation zu generieren, dann sollte die beobachtete Kooperationsrate in Computer0 höher ausfallen als in der Partner-Anordnung.

Eine für unsere Belange wichtige Hypothese betrifft schließlich die zeitliche Entwicklung des Zeitpunkts der ersten Defektion. Wenn die Population der Versuchspersonen tatsächlich aus rationalen Defektoren bestünde (wie das Kreps et al Modell es unterstellt), dann sollte diese Tatsache allmählich im Verlauf der 200 Iterationen gelernt werden und Spieler sollten ihre beliefs über die Populationszusammensetzung entsprechend anpassen. Man würde so erwarten, daß es im Zeitablauf schwieriger wird, Kooperation im endlich wiederholten GD zu stützen und sich der Zeitpunkt der ersten Defektion (Zeitpunkt der ersten Abweichung von einmütiger Kooperation) über die 20 Wiederholungen des 10-Perioden-GD betrachtet zeitlich nach vorn verlagert. Zumindest würde man auf jeden Fall nicht erwarten, daß dieser Zeitpunkt sich zeitlich nach hinten verschiebt. Nimmt man hingegen an, daß in der Population unter anderem auch kooperative Spieler real existieren, so wäre eine solche Beobachtung mit dieser Annahme kompatibel. Die Resultate des Experiments werden in folgender Graphik zusammengefaßt. Dargestellt sind die durchschnittlichen Kooperationsraten im 10-Perioden-GD über alle 20 Wiederholungen.

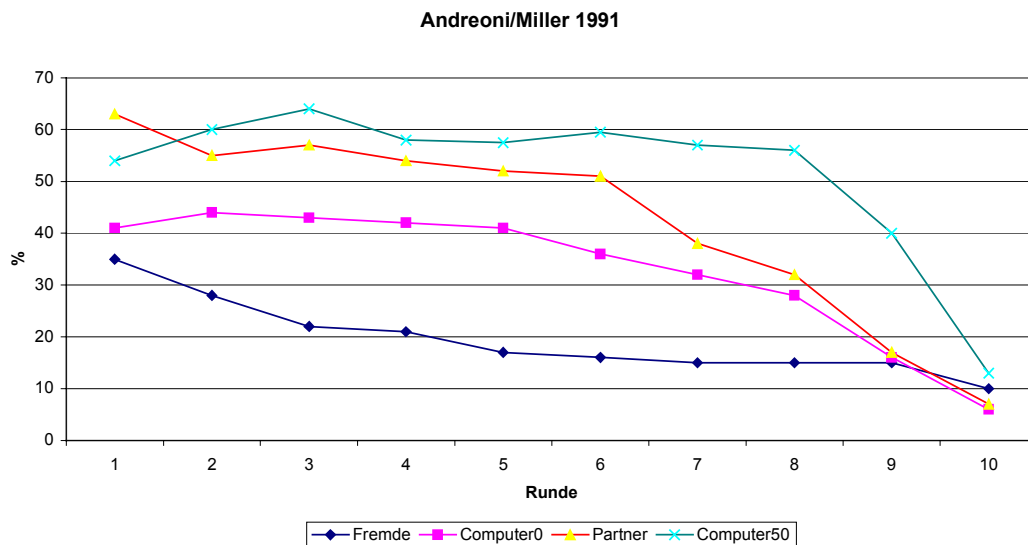


Abbildung 1: Durchschnittliche Kooperation im 10-Perioden-GD

Ein Großteil der Beobachtungen ist konsistent mit der Reputationshypothese. So sind Partner deutlich kooperativer als Fremde. Dasselbe gilt für die Computer50-Anordnung. Im übrigen ist natürlich der fallende Verlauf der Kooperationsrate im Zeitablauf in den Anordnungen Partner, Computer50 und Computer0 ebenfalls völlig konsistent mit der Reputationshypothese.

In unserem Kontext ist jedoch zunächst vor allem folgende Beobachtung von Bedeutung. Entgegen der theoretischen Erwartung zeigt sich auch in der Fremden-

Anordnung ein beachtliches Ausmaß an kooperativem Verhalten. Mehr noch, dieses erweist sich als im Zeitablauf relativ konstant. Betrachtet man das Ausmaß der Kooperation in den ersten fünf Runden und den letzten fünf Runden, so zeigt sich kein statistisch signifikanter Unterschied ( $z=1.36$ ). Dieser Befund steht natürlich im Widerspruch zur Reputationshypothese und deutet eher darauf hin, daß es tatsächlich Spieler mit einer Präferenz für Kooperation gibt<sup>6</sup>.

Eine weitere Beobachtung wirft Zweifel auf, ob die Reputationshypothese das beobachtete Verhalten tatsächlich umfassend erklären kann. Betrachtet man den durchschnittlichen Zeitpunkt der ersten Defektion, so stellt man sowohl für Partner als auch für Computer50 fest, daß sich dieser im Zeitablauf in die hintere Spielhälfte verlagert.

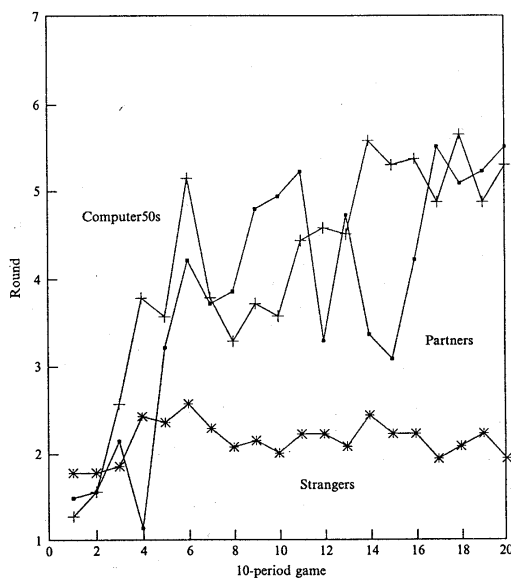


Abbildung 2: Durchschnittlicher Zeitpunkt der ersten Defektion

Partner etwa beginnen mit Defektion in den ersten 10-Perioden-GD ungefähr in der zweiten Runde. Betrachtet man nun die letzten vier Wiederholungen des 10-Perioden-Spiels, so stellt man fest, daß dort im Durchschnitt zwischen der fünften und sechsten Runde das erste mal defektiert wurde. Fast identische Beobachtungen ergeben sich für Computer50. Diese Beobachtung ist schwerlich vereinbar mit dem Reputationsmodell, wenn unterstellt wird, daß die Population tatsächlich nur aus rein rationalen Defektoren besteht. Sie ist hingegen vereinbar mit der

<sup>6</sup> Streng genommen können auch in dieser Anordnung Reputationseffekte auftreten. Zwar ist es nicht möglich, Kooperationsbereitschaft gezielt gegenüber einzelnen Spielern zu signalisieren, jedoch könnten Spieler versuchen, ein kooperatives „Gruppenklima“ zu erzeugen. Vgl. hierzu auch Kandori (1992).

Hypothese, daß kooperative Individuen existieren, daß deren Populationsanteil von den Spielern im Zeitablauf gelernt wird und beliefs über die Populationszusammensetzung von allen Spielern entsprechend angepaßt werden. Bei den Fremden hingegen bleibt der durchschnittliche Zeitpunkt der ersten Defektion weitgehend konstant.

Die bemerkenswerte Konstanz des Verhaltens der Vpn in der Fremden-Anordnung veranlaßt Andreoni/Miller zu der Hypothese, daß die Spieler ein Nash-GG mit unvollständiger Information über das Ausmaß der Kooperationsbereitschaft ihrer Mitspieler spielen. Dieses Gleichgewicht impliziert die Existenz dreier Typen, des strikten Defektors, des unbedingten Kooperators und eines dritten Typs, der zwischen Kooperation und Defektion mischt. Das Interessante ist nun, daß sich diese drei Typen tatsächlich in Andreonis und Millers subject pool identifizieren lassen. Unter einem Defektor wird dabei eine Vpn verstanden, die fast immer defektiert. Legt man ein 99%-Konfidenzintervall zu Grunde, so bedeutet dies, daß ein Defektor in mindestens 93,7% aller Interaktionen defektieren muß. Diese Klassifikation angenommen, lassen sich fünf der 14 Vpn eindeutig als Defektoren einordnen, mit beobachteten Defektionshäufigkeiten zwischen 97 und 100 %. Verfährt man in derselben Weise für die kooperative Strategie, so läßt sich genau eine Vpn als Kooperator einstufen (Kooperationshäufigkeit 94 %). Weitere sieben Vpn mischten zwischen Kooperation und Defektion mit einer durchschnittlichen Kooperationshäufigkeit von 20% (Streuung von 10 % bis 32 %). Eine Vpn mischte, aber mit einer deutlich höheren durchschnittlichen Kooperationshäufigkeit von 42 %. Interessanterweise kommt es zu einer bemerkenswerten Übereinstimmung der auf der Basis der mixed-equilibrium-Hypothese prognostizierten und empirisch beobachteten Werte der Kooperationshäufigkeiten. Andreoni und Miller resümieren:

*„This suggests that the behaviour in the stranger condition is consistent with an imperfect-information equilibrium in which individuals share a common prior on the probability of experiencing cooperation [...] of about 0.20.“* (Andreoni und Miller (1991), S. 581)

Dieses Ergebnis ist um so bemerkenswerter, als bereits zuvor Camerer und Weigelt (1988) den „homemade prior“ für erwartete Kooperation auf 17% schätzten.

Zusammenfassend kann man festhalten: Andreoni und Miller finden Evidenz dafür, daß Reputationsbildung im endlich wiederholten GD eine Rolle spielt. Allerdings kann längst nicht die Gesamtheit der Beobachtungen durch das gang-of-four-Modell rationalisiert werden. Insbesondere die konstante und signifikante Kooperationsbereitschaft im one-shot-GD läßt sich reputationstheoretisch nicht



erklären. Sowohl die Beobachtungen aus den one-shot-GD, als auch die Beobachtung, daß der durchschnittliche Defektionszeitpunkt für Partner und Computer50 sich zeitlich nach hinten verlagert, deuten eher auf die Existenz kooperativer Typen hin.

### 2.1.2 Kritik

Die Studie von Andreoni und Miller kann in mehrerlei Hinsicht kritisiert werden. Zum einen verbleibt die Ableitung der Hypothesen doch einigermaßen im Vagen. Beispielsweise setzt die Prognose über die Entwicklung des Defektionszeitpunktes die Gültigkeit der Lernrichtungstheorie von Selten und Stoecker (1986) voraus. Dies wird aber nicht explizit gesagt. Andreoni/Miller verweisen lediglich unspezifisch auf das Papier von Selten und Stoecker. Dieser und andere mögliche Kritikpunkte an der Hypothesenbildung sollen uns aber im weiteren nicht beschäftigen. Schwerwiegender im Kontext unseres Interesses wiegt die Tatsache, daß Andreoni/Miller lediglich eine geringe Anzahl unabhängiger Beobachtungen vorweisen können. Die „Altruismushypothese“ stütze sich einerseits auf die Beobachtung der zeitlichen Verlagerung des durchschnittlichen Zeitpunkts der ersten Defektion. Hierzu existieren aber lediglich zwei unabhängige Beobachtungen (das Partner-treatment und das Computer50-treatment)<sup>7</sup>, was als methodisch inakzeptabel bezeichnet werden muß.

Zum anderen waren zentral für die Schlußfolgerungen Andreoni/Millers die Beobachtungen im one-shot-GD (Fremden-Anordnung). Man kann aber mit Fug und Recht die Frage erheben, ob es sich bei der gewählten Versuchsanordnung tatsächlich um 200 one-shot Gefangenen-Dilemmata handelte. Zum einen hat Kandori (1991) gezeigt, daß gewisse Gruppenreputationseffekte auch dann strategisches Verhalten gebieten können, wenn die Historien des Spiels den Spielern nicht bekannt sind, also eigentlich Reputation im Krepsschen Sinne keine Rolle spielen kann. Zum anderen aber liegt klar auf der Hand, daß im gewählten setting die Vpn möglicherweise vom strategischen Motiv des Reputationserwerbs getrieben wurden. Man vergegenwärtige sich, daß eine ausgesprochen kleine Zahl von Personen (n=14) das one-shot-GD 200mal hintereinander spielte. Es ist sofort ersichtlich, daß eine positive Wiedertreffwahrscheinlichkeit existiert, mit der Folge, daß es durchaus lohnend erscheinen kann, durch gegenwärtiges Verhalten

---

<sup>7</sup> Die Computer0-Anordnung wäre grundsätzlich ebenfalls geeignet, die Hypothese der Existenz kooperativer Typen zu stützen, da sie sich ja faktisch von der Partner-Anordnung nicht unterscheidet. Die entsprechenden Resultate werden jedoch von Andreoni und Miller nicht referiert.

das zukünftige Verhalten anderer Spieler zu beeinflussen. Im übrigen sei hier erwähnt, daß die Wiedertreff-Wahrscheinlichkeit mit Hilfe des in Kapitel 4 dieser Arbeit entwickelten Lernmodells exakt berechnet werden kann. Vorgehend auf Resultate, die in Kapitel 4 abgeleitet werden, zeigen wir in Abbildung 3 einen Plot der Wiedertreffwahrscheinlichkeit. Die Funktion  $f(x)$  gibt dabei die Wahrscheinlichkeit an, daß ein bestimmter Spieler bekannt ist und wiedertreffend wird. Wie man sieht, konvergiert diese Wahrscheinlichkeit nach zirka 80 Iterationen gegen den Wert der empirischen relativen Häufigkeit des Spielers von  $1/14$ . Mit anderen Worten: Bereits nach etwa 80 one-shot-GD konnten die Spieler in den folgenden Spielen mit der relativen Häufigkeit des

$$t := 1, 2, \dots, 100 \quad n := 14 \quad x := \frac{1}{14}$$

$$f(x) := x \cdot \left[ 1 - \left( \frac{n-1}{n} \right)^t \right] \quad g(x) := \frac{1}{14}$$

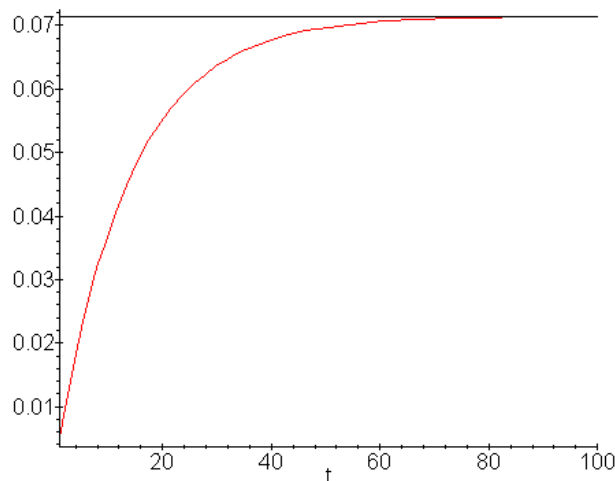


Abbildung 3: Wiedertreffwahrscheinlichkeit für  $t=1, \dots, 100$  bei random matching

Vorkommens eines Spielers in der Population von  $1/14$  als Schätzer davon ausgehen, daß sie diesen Spieler bereits einmal getroffen haben und ihn nun wiedertreffen würden (in jeder folgenden Iteration des „one-shot-GD“!). Dies ist offensichtlich keine echte one-shot-Anordnung!

Noch etwas deutlicher formuliert: Bezogen auf die Gesamtpopulation der 14 Spieler konnte jede Vpn nach zirka 80 Spielen mit Sicherheit davon ausgehen, nur noch bekannte Spieler wiederzutreffen. Unter diesen Bedingungen kann man a priori nicht ausschließen, daß die beobachtete Kooperation im Fremden-treatment auf Reputationserwägungen zurückzuführen ist.

## 2.2 Cooper et al (1996)

Glücklicherweise existiert mit Cooper et al eine Studie, die einerseits die erwähnten methodischen Mängel von Andreoni/Miller (1991) vermeidet, andererseits aber in der Lage ist, wesentliche Resultate zu replizieren. Insbesondere die one-shot games wurden mit großer Vorsicht konstruiert. 40 Vpn spielten 20 one-shot GD. Die 40 Spieler wurden in zwei gleich große Gruppen aufgeteilt, eine „rote Gruppe“ und eine „blaue Gruppe“. In jedem matching-Zeitpunkt wurden je ein roter und je ein blauer Spieler gematcht. Der verwendete matching-Mechanismus stellte sicher, daß es zu keinerlei Wiederbegegnungen kommen konnte<sup>8</sup>. Die Prozedur orientiert sich an Townsends turnpike-Modell (Townsend (1980)). Man stelle sich zwei Gruppen von Reisenden vor, die einen Schlagbaum passieren. Die eine Gruppe von Reisenden bewege sich in westlicher, die andere in östlicher Richtung. Beide Gruppen reisen mit identischer Geschwindigkeit. In jedem Zeitpunkt treffen zwei Personen aus den beiden unterschiedlichen Gruppen am Schlagbaum zusammen, spielen ihr Spiel und bewegen sich anschließend in entgegengesetzten Richtungen voneinander fort. Ein erneutes Zusammentreffen in der Zukunft ist ausgeschlossen. Dieser Mechanismus wurde den Teilnehmern des Experiments genau erklärt. Die one-shot-GD wurden dann unter dem beschriebenen matching-Verfahren durchgeführt. Spiele fanden anonym statt, d.h. keiner der Spieler kannte in irgendeinem Zeitpunkt die Identität seines Mitspielers. Ebenso kannte keine Vpn die Geschichte der Entscheidungen anderer Spieler. Es gab also in der Tat keinerlei Grund für strategisch motiviertes, kooperatives Verhalten. Das Ergebnis der one-shot Experimente ähnelt sehr genau dem bereits aus Andreoni/Miller bekannten. Man beobachtet signifikant positive Kooperationsraten, die zeitlich relativ konstant bleiben und im Durchschnitt 22 % betragen. Auch die Häufigkeitsverteilung der beobachteten Verhaltensweisen ähnelt im wesentlichen der von Andreoni/Miller:

---

<sup>8</sup> Man beachte jedoch, daß auch hier wieder indirekte Reputationsüberlegungen eine Rolle spielen können. Vgl. Fußnote 6.

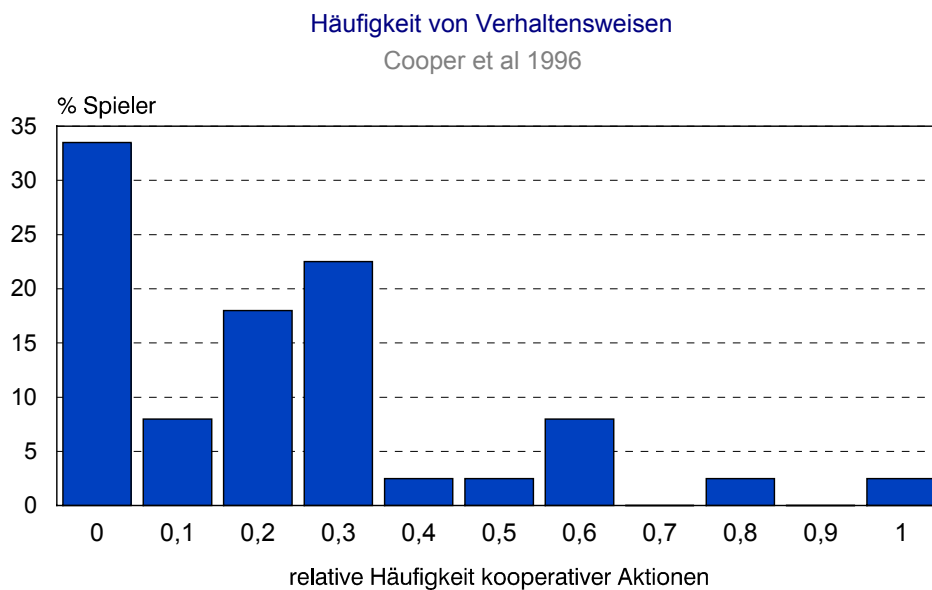


Abbildung 4: Häufigkeitsverteilung von Verhaltensweisen

Es existiert ein Anteil strikter Defektoren von gut einem Drittel (Andreoni/Miller: 35,7%) sowie wiederum ein nur sehr kleiner Anteil strikter Kooperative. Wie bei Andreoni und Miller auch, handelt es sich hierbei um genau eine Vpn. Dazwischen existieren verschiedene mischende Verhaltensweisen. Wie bei Andreoni/Miller existiert hier ein großer Anteil von schwach kooperativen Spielern mit Kooperationsraten zwischen 10 und 30 %. Ihr Anteil beläuft sich auf 48,5 % (Andreoni/Miller: 50%). Daneben existieren einige Spieler (15%) mit einer etwas ausgeprägteren Kooperationswilligkeit.

Die Arbeit von Cooper et al ist aber nicht nur interessant, weil sie geeignet erscheint, die Resultate von Andreoni/Miller zu rehabilitieren, sondern auch, weil sie in der Bestimmung von Typen methodisch neue Wege beschreitet. Die simpelste Methode zur Bestimmung des Anteils kooperativer Typen an der Gesamtpopulation ist natürlich, die empirisch beobachtete durchschnittliche Kooperationshäufigkeit am Beginn des Spiels als Schätzer zu verwenden. Allerdings wird hierbei implizit unterstellt, daß Defektoren immer defektieren und Kooperative immer kooperieren. Realistischer ist es davon auszugehen, daß Spieler gelegentlich Fehler begehen. Wir wollen hier nicht die Einzelheiten des Modells referieren. Im Grunde wird lediglich angenommen, daß die prinzipielle Verhaltensdisposition (defektieren oder kooperieren) von einer Zufallsvariable überlagert wird. Auf der Basis einer Maximum Likelihood Schätzung gelangen

Cooper et al dann zu folgender Verteilung<sup>9</sup>: Bezogen auf die ersten zehn Wiederholungen des one-shot-GD erweisen sich 15 % der Vpn als kooperative Typen, 62.5 % als Defektoren und gut 22 % lassen sich nicht eindeutig klassifizieren. Bezogen auf die zweiten zehn Wiederholungen erweisen sich 12.5 % als kooperativ, 85 % als defektiv und 2.5 % sind nicht eindeutig zuzuordnen. Auch unter Einbeziehung der Möglichkeit, daß Spieler Fehler begehen, läßt sich ein signifikant von Null verschiedener Anteil der Versuchspopulation eindeutig als kooperativer Typ klassifizieren.

Die vorgeschlagene Klassifikation erweist sich als ausgesprochen erklärungs-mächtig für die beobachteten Spieldausgänge. Man kann nämlich aufgrund der ermittelten relativen Häufigkeiten der Typen sowie der ebenfalls aus dem Maximum-Likelihood-Modell ableitbaren Fehlerhäufigkeiten die Wahrscheinlichkeiten für alle möglichen Spieldausgänge prognostizieren und diesen Wert dann mit dem empirisch beobachtbaren vergleichen. So ergibt sich auf der Grundlage des statistischen Modells von Cooper et al die Wahrscheinlichkeit für den Spieldausgang (Defektion, Defektion) als Summe aus: i) der Wahrscheinlichkeit, daß zwei defektive Typen aufeinandertreffen und keinen Fehler begehen; ii) der Wahrscheinlichkeit für ein Zusammentreffen zweier kooperativer Typen, die beide einen Fehler begehen; iii) der Wahrscheinlichkeit, daß ein Defektor, der keinen Fehler begeht, auf einen Kooperationspartner trifft, der einen Fehler macht, sowie iv) der Wahrscheinlichkeit, daß ein fehlerfreier Defektor (oder ein Kooperationspartner, der einen Fehler begeht) auf einen „Anderen“ trifft, der zufällig defektiert und schließlich v) der Wahrscheinlichkeit dafür, daß zwei „Andere“ sich treffen und beide defektieren. Verfährt man in der skizzierten Weise, dann erhält man für alle möglichen Spieldausgänge prognostizierte Häufigkeiten, die in der folgenden Tabelle mit den tatsächlichen Häufigkeiten verglichen werden.

Spieldausgang	Theoretische Prognose	emp. rel. Häufigkeit
Defektion, Defektion	0,606	0,620
Koop., Koop.	0,049	0,065
Koop., Defektion	0,172	0,155
Defektion, Koop.	0,172	0,160

Tabelle 1: Erklärungskraft des Maximum-Likelihood-Modells von Cooper et al

---

<sup>9</sup> Für die Klassifikation der Spieler wird folgendes angenommen: Ein Spieler, der eine Kooperationshäufigkeit oberhalb von 50 % aufweist, gilt als Kooperationspartner, ein Spieler der in weniger als 50% aller Aktionen kooperiert, gilt als Defektor und Individuen, bei denen die Kooperationshäufigkeit genau 50 %, betragt werden als „Andere“ bezeichnet.

Der Grad der Übereinstimmung kann ohne Übertreibung als beeindruckend bezeichnet werden und wird durch einen Signifikanztest untermauert (vgl. Cooper et al, S.204).

Weitere Evidenz für Typen findet sich in der Untersuchung von McKelvey und Palfrey (1993) zum Centipede-game. Auch dort kann ein signifikant positiver Anteil der Population als kooperativer Typus identifiziert werden.

### 2.3 Evidenz in public good Experimenten

Weitere Evidenz für die Existenz von Typen, die sich hinsichtlich ihrer Kooperationsbereitschaft unterscheiden, läßt sich in der Literatur zu public good (PG) Experimenten finden. Die Tatsache, daß GD- und PG-Experimente eine äußerst ähnliche Anreizstruktur aufweisen, sowie die Tatsache, daß die Resultate in beiden Anordnungen nahezu identisch sind, läßt es gerechtfertigt erscheinen, PG-Experimente zu unseren bisherigen Ausführungen hinzu zu ziehen.

Eine erste ökonomische Arbeit, die sich mit dem Individualverhalten beschäftigt ist die von Isaac, Walker und Thomas (1984). Bezeichne  $m_i$  den Beitrag von Spieler  $i$  zum öffentlichen Gut, angegeben in % des ihm zur Verfügung stehenden Budgets. Isaac, Walker und Thomas klassifizieren das beobachtete Individualverhalten in einem 10-Runden Standard-PG Spiel dann wie folgt:

Complete strong free riding:	$m_i = 0$
Incomplete strong free riding:	$0 < m_i \leq 33.33\%$
Weak free riding:	$33.33\% < m_i \leq 66.66\%$
Incomplete Lindahl:	$66.66\% < m_i < 100\%$
Complete Lindahl:	$m_i = 100\%$

Bezogen auf die insgesamt 560 zugrunde liegenden Beobachtungen ergibt sich dann folgende Häufigkeitsverteilung:

	Rd.											
	1	2	3	4	5	6	7	8	9	10	Total	%
complete strong free riding	12	15	13	11	12	16	20	20	22	25	166	30
incomplete strong free riding	11	8	10	10	7	8	4	9	8	4	79	14
weak free riding	15	16	16	15	18	20	13	15	14	9	151	27
Incomplete Lindahl	6	4	6	6	7	4	9	7	3	5	57	10
Complete Lindahl	12	13	11	14	12	8	10	5	9	13	107	19

Tabelle 2: Häufigkeiten beobachteter Verhaltensweisen

Quelle: Isaac, Walker und Thomas (1984), S.126

Immerhin in 19 % aller Beobachtungen verhalten sich Vpn perfekt kooperativ und investieren ihre gesamte Anfangsausstattung in das öffentliche Gut, während sich in einem wiederum beträchtlichen Prozentsatz der Beobachtungen Vpn als strikte Defektoren erweisen. Ihr Anteil liegt mit etwa einem Drittel in der Größenordnung der Beobachtungen von Andreoni/Miller und Cooper et al (1996). Insgesamt ist das obige Bild weitgehend kompatibel mit den Resultaten aus den GD-Experimenten. Jedoch gilt es zu beachten, daß es sich bei der obigen Übersicht um eine Klassifikation von **Aktionen** handelt, die keine Auskunft darüber gibt, inwiefern sich Individuen tatsächlich konstant kooperativ, schwach kooperativ oder defektiv verhalten haben. Die folgende Graphik legt allerdings die Vermutung nahe, daß es eine gewisse zeitliche Konstanz des Verhaltens auch bezogen auf Individuen gibt. Dargestellt ist die zeitliche Entwicklung der relativen Häufigkeiten von Verhaltensweisen, diesmal grob aggregiert in die drei Kategorien „strong free riding“ ( $0 \leq m_i \leq 33.33\%$ ), „weak free riding“ ( $33.33\% < m_i \leq 66.66\%$ ) und „Lindahl“ ( $66.66\% < m_i \leq 100\%$ ).

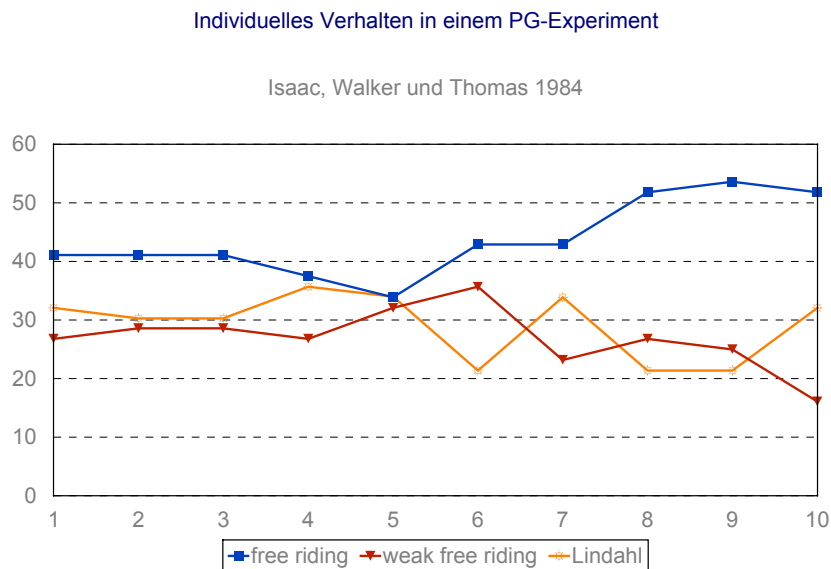


Abbildung 5: Zeitliche Entwicklung der Verhaltenskategorien bei Isaac et al

Die Abbildung enthüllt eine bemerkenswerte Konstanz der Lindahl-Spieler, deren Anteil auch in der Schlußrunde noch 32.1% beträgt, was exakt ihrer Häufigkeit in Runde 1 entspricht. Bei den übrigen beiden Typen findet eine Verlagerung von schwachem zu starkem Freifahrerverhalten statt, was ursächlich für den stets beobachtbaren Abfall der Kooperationswilligkeit gegen Spielende ist. Bis etwa

Runde 7 sind jedoch auch bei den Freifahrertypen die Häufigkeiten erstaunlich konstant. Das Ausbleiben eines Schlußrundeneffekts bei den Lindahl-Spielern gibt auch Isaac, Walker und Thomas Anlaß zu Erstaunen und zu folgender Vermutung:

*„Why did some individuals contribute positive amounts even in the 10<sup>th</sup> period, when the learning and multiperiod gaming aspects were presumably of minimal effect? We conjecture that this represents a core of people for whom utility functions are not completely selfish or who otherwise wish to believe in good guy fashion.“* (Isaac, Walker und Thomas (1984), S.141)

Wiederum deuten die experimentellen Beobachtungen also auf die Existenz unterschiedlicher Typen hin. Es scheint Menschen zu geben, die mit einer stabilen Präferenz für kooperatives Verhalten ausgestattet sind. Dieser Befund wird bestätigt durch eine Arbeit von Weimann (1994), die sich explizit mit dem Individualverhalten in einem Standard-PG Experiment beschäftigt. Gespielt wurde wiederum ein 10-Runden-PG Spiel, wobei das Spiel die Struktur eines n-Personen GD hatte (n=5). Weimann verwendet eine Isaac et al (1984) folgende Klassifikation des individuellen Verhaltens. Vpn, die im zeitlichen Durchschnitt über alle 10 Runden weniger als 1/3 ihrer Ausstattungen in das öffentliche Gut investierten, wurden als Freifahrer interpretiert, solche Individuen, die oberhalb von 2/3 ihrer Anfangsausstattungen in das öffentliche Gut einzahlten, wurden als Kooperative begriffen und Vpn, die zwischen den beiden Extremen lagen, bildeten die Kategorie der schwachen Freifahrer. Die Grundgesamtheit der Vpn über alle Experimente betrug 104. 19 Personen konnten als Freifahrer identifiziert werden, 26 Akteure waren kooperativ im Sinne der obigen Klassifikation, und eine Mehrheit von 59 Vpn fiel in die Kategorie schwache Freifahrer. Die zeitliche Entwicklung des durchschnittlichen Beitragsverhaltens, disaggregiert nach den drei Typen, zeigt die folgende Abbildung.



## Weimann 1994

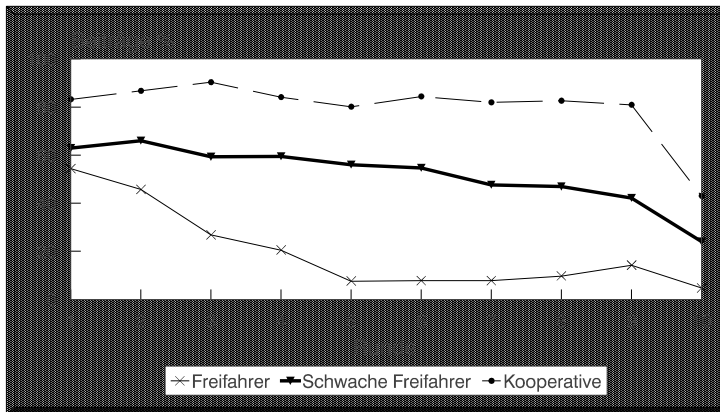


Abbildung 6: Zeitliche Entwicklung des Verhaltens der drei Typen bei Weimann

Bemerkenswert ist daran wiederum die Konstanz des Verhaltens der kooperativen Typen und auch das Niveau der Kooperationsbereitschaft. Immerhin werden über die ersten neun Runden im Durchschnitt zwischen 80 und 90% der Anfangsausstattungen in die öffentliche Anlage investiert. Die zeitliche Konstanz des Verhaltens ist vollkommen kompatibel mit der Beobachtung der Konstanz des Lindahl-Verhaltens bei Isaac, Walker und Thomas (1984). Die Vermutung liegt nahe, daß auch der gleichbleibende Anteil von Lindahl-Aktionen dort zurückzuführen ist auf einen konstanten Anteil kooperativer **Spieler** mit einer ebenfalls gleichbleibenden Kooperationswilligkeit. Im Unterschied zu Isaac, Walker und Thomas beobachtet Weimann jedoch den klassischen Schlußrundeneffekt auch für die kooperativen Spieler.

Als eine dritte Public-good-Studie, die starke Evidenz für die Existenz unterschiedlicher Typen präsentiert, sei die Arbeit von Isaac, Walker und Williams (1994) (fortan IWW) erwähnt. Auch dort werden drei charakteristische Typen von Verhaltensweisen identifiziert (vgl. IWW 1994, Fig.13, S.31). Zum einen existiert offensichtlich auch bei IWW ein stark defektiv orientierter Verhaltenstypus, der sich häufig als starker free rider verhält und nur gelegentlich positive Beiträge zum öffentlichen Gut leistet, aber auf einem sehr niedrigen Niveau. Beitragsraten unter 20 % sind für ihn typisch (vgl. ebd., subject 2). Zum zweiten wird ein Typ präsentiert (vgl. ebd., subject 10), der im Durchschnitt kooperativ ist (also gemäß der Klassifikation von Weimann oberhalb von 66% beiträgt) und häufig 100% seiner Ausstattung in das öffentliche Gut investiert. Schließlich existiert wiederum ein dritter Typ, der mit kooperativen Beiträgen beginnt (um 60%), die im Zeitverlauf aber fallen (der individuelle Beitragsverlauf ist qualitativ identisch mit dem aggregierten Verlauf in PG-Experimenten). Die Stärke der Arbeit von IWW besteht in der ausgesprochen großen Zahl von unabhängigen Beobachtungen. Die

Auswertung kann sich auf eine Datenbasis von 87 PG-Experimenten in unterschiedlichen Anordnungen stützen, an denen insgesamt 1908 Vpn teilnahmen. Diese Stärke wird aber bei der Präsentation von Individualverhalten zum Verhängnis: Es existiert keine befriedigende Methode, eine derartig große Anzahl individueller Spielverläufe zu repräsentieren. So behelfen sich IWW damit, die drei oben skizzierten Verhaltensweisen anhand dreier Vpn darzustellen und zugleich zu versichern,

*„[that] these three subjects typify three types of behavior we regularly observe.“*  
(ebd.,S.30)

Dem geneigten Leser bleibt keine andere Wahl, als dieser Versicherung zu vertrauen (oder aber sich der Mühe zu unterziehen, die Originaldaten zu reanalysieren)<sup>10</sup>.

## 2.4 Offerman et al (1996)

Offerman et al (1996) haben darauf hingewiesen, daß es grundsätzlich problematisch ist, aus den offenbarten Verhaltensweisen in wiederholten Spielen auf Verhaltensdispositionen zu schließen. Dies deshalb, weil in wiederholten Interaktionen Anlaß zu strategischem Verhalten besteht. So ist also nicht klar, ob das Spielen der kooperativen Strategie im GD oder PG-Spiel Ausdruck einer grundlegenden kooperativen Orientierung des betrachteten Spielers ist, oder aber aus Reputationsgründen motiviert ist. Dieser Vorbehalt trifft damit sämtliche in 2.3 zitierten Studien. Davon unberührt bleiben aber natürlich die Beobachtungen aus echten one-shot GD-Experimenten (Cooper et al (1996)). Dennoch soll die Arbeit von Offerman et al kurz referiert werden, weil sie zusätzliche und überzeugende Evidenz für die Existenz von Typen liefert.

Aus den genannten Gründen fordern Offerman et al eine vom strategischen Experimentkontext unabhängige Messung von Verhaltensdispositionen. Sie implementieren dies in ihrem Untersuchungskontext (einem Step-level public good Experiment) wie folgt. Vor dem eigentlichen Experiment hatten die Vpn Entscheidungen über die Aufteilung einer Summe Geld zwischen sich und einem anderen Spieler zu treffen. Dies geschah in der Art und Weise, daß den Vpn jeweils zwei komplette Allokationen zur Auswahl gestellt wurden. Bezeichne  $x$  den Betrag, den eine Allokation für den betreffenden Spieler vorsieht und  $y$  den Be

---

<sup>10</sup> Entgegen einer in derartigen Situationen gern zitierten Leninschen Devise entscheiden wir uns für Vertrauen.

trag für den anderen Spieler. Es existierten insgesamt 24 verschiedene Allokationen mit der Eigenschaft  $x^2 + y^2 = 15^2$ . Mit anderen Worten: Die Allokationen lagen also auf einem Kreis mit Radius 15.  $x+y$  war nicht konstant über die verschiedenen Allokationen. Eine Allokation kann dann graphisch einfach repräsentiert werden durch den Vektor, der auf den entsprechenden Punkt auf dem Kreissegment deutet. Die 24 Vektoren waren nun gleichmäßig über den relevanten Bereich des Kreises verteilt (was der relevante Bereich ist, wird sofort klar werden), und die Vpn hatten sich in paarweisen Vergleichen von benachbarten Vektoren für jeweils einen der beiden zu entscheiden. Durch diese Abfolge von Vergleichen hoffte man den am meisten präferierten Vektor für die jeweilige Vpn zu finden. Dieser Vektor gilt dann als Maß für die Disposition der Vpn auf der Basis der folgenden Klassifikation: Vektoren, die im Bereich von  $+112.5^\circ$  und  $67.5^\circ$  liegen, werden als Indikator für eine altruistische Orientierung gedeutet. Liegt der ermittelte Vektor im Winkelbereich von  $67.5^\circ$  bis  $22.5^\circ$ , so deutet dies auf eine kooperative Orientierung hin. „Individualisten“ sind charakterisiert durch einen Vektor zwischen  $+22.5^\circ$  und  $-22.5^\circ$ , „Kompetitive“ weisen einen Vektor zwischen  $-22.5^\circ$  und  $-67.5^\circ$  auf, während schließlich Vpn mit einem Vektor zwischen  $-67.5^\circ$  und  $-112.5^\circ$  als „Aggressive“ bezeichnet werden. Wie Offerman et al anmerken, ist diese Klassifikation – wie jede andere auch – natürlich letztlich willkürlich, jedoch naheliegend. Von einem Altruisten erwartet man, daß er nur den Nutzen des anderen im Blick hat. Sein Vektor sollte also etwa bei  $x = 0$  liegen. Im Gegensatz hierzu hat der Individualist nur sein eigenes Wohlergehen im Auge. Sein Vektor

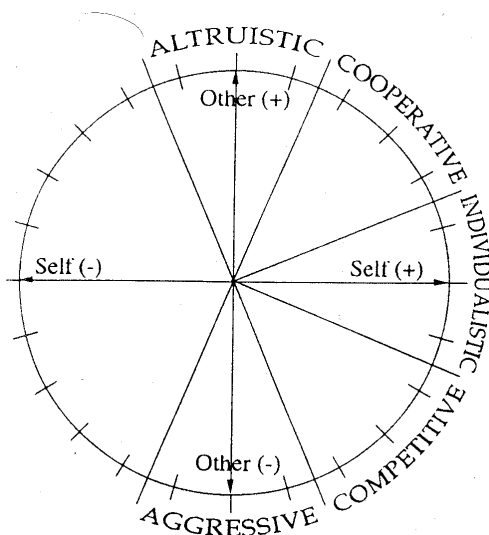


Abbildung 7: Der „value orientation circle“  
Quelle: Offerman et al (1996), S. 6

sollte also etwa bei  $y=0$  liegen. Extreme Orientierungen sind „kompetitiv“ und „aggressiv“. Solche Individuen ziehen offenbar Nutzen aus dem Disnutzen des anderen.

Die Typisierung des subject pools ergibt folgende Ergebnisse: Die überwiegende Mehrzahl der Vpn wies nach dem „value orientation circle“ eine individualistische Orientierung auf. 121 von 186 Individuen, also 65% aller Versuchspersonen, fielen in diese Kategorie. Aber immerhin 50 Spieler fielen in die Kategorie „kooperativ“, was 27% aller Vpn entspricht. Ausgesprochen rar waren extreme Orientierungen. Lediglich 8 Vpn konnten als „kompetitiv“ identifiziert werden (4%) und nur ein Spieler als „aggressiv“. Altruisten konnten nicht nachgewiesen werden. Im wesentlichen scheinen also zwei Typen von Spielern zu existieren. Dieses Ergebnis ist in voller Übereinstimmung mit der psychologischen Literatur, wo regelmäßig die zwei Typen der Individualisten und Kooperativen nachgewiesen werden können (vgl. Kelley und Stahelski (1970), Kuhlman und Wimberley (1976), Liebrand (1984)). Bezüglich der übrigen Typen ist die Evidenz unterschiedlich. Liebrand (1984) etwa findet auch 5% Altruisten.

Hervorzuheben gegenüber der psychologischen Literatur ist die Tatsache, daß Offerman et al einen grundlegenden methodischen Mangel vermeiden, indem sie ihren Vpn monetäre Anreize bieten. Die Allokationsentscheidungen in vielen psychologischen Experimenten zeitigten keinerlei reale Konsequenzen. Dies scheint im Kontext der vorliegenden Entscheidungssituation in der Tat inakzeptabel zu sein. Umso bemerkenswerter ist die Tatsache, daß sich die Ergebnisse früherer psychologischer Studien als durchaus robust erweisen.

Es scheint jedoch einen Effekt der monetären Anreize auf die Güte der Messung zu geben. Ist erst einmal der Motivationsvektor für eine Vpn ermittelt, so kann ex post beurteilt werden, ob sich die Vpn in den paarweisen Vergleichen benachbarter Allokationen konsistent verhalten hat. Konsistent ist eine Entscheidung natürlich genau dann, wenn die betreffende Person den Vektor wählt, der näher an ihrem gemessenen Motivationsvektor liegt. Offerman et al ermitteln nun eine durchschnittliche Konsistenz aller Entscheidungen von 90%, was deutlich über den Werten anderer Studien liegt. Liebrand (1984) etwa berichtet Konsistenzniveaus von 76 und 80%.

## 2.5 Zusammenfassung

Insgesamt deutet die experimentelle Literatur auf die Existenz mehrerer Spielertypen hin. Experimentalpopulationen erweisen sich als keineswegs homogen, es sind vielmehr systematische und persistente Unterschiede in den Verhaltensdispositionen der Spieler zu entdecken. Die Studien von Miller und Andreoni (1991)

sowie Cooper et al (1996) zu GD-Spielen legen die Existenz von drei Typen nahe: einem strikt kooperativen Typ, einem strikt defektiven Typ und einem dritten Typ, der zwischen der kooperativen und defektiven Strategie mischt. Ein qualitativ identischer Befund ergibt sich aus der Literatur zu public good Spielen (N-Personen-GD). Weimann (1994) findet hoch-kooperative Typen, schwache Freifahrer und Defektoren. Ähnliche Beobachtungen machen Isaac, Walker und Williams (1994). Offerman et al (1996) haben darauf aufmerksam gemacht, daß der Schluß von beobachteten Verhaltensweisen auf zugrundeliegende Dispositionen in strategischen Spielkontexten nicht zwingend ist. Nimmt man diesen methodischen Einwand ernst, so bleibt zunächst festzuhalten, daß der Befund aus echten one-shot GD-Experimenten (Cooper et al) hiervon unberührt bleibt. Zum anderen zeigt die Arbeit von Offerman et al, daß auch eine vom strategischen Kontext unabhängige Messung von Verhaltensdispositionen die Identifizierung von im wesentlichen zwei Typen erlaubt, nämlich Individualisten und Kooperative. Dieser Befund bestätigt frühere psychologische Arbeiten, die regelmäßig die genannten zwei Typen finden. In welche Verhaltensweisen in GD und NGD-Experimenten sich die von Offerman et al identifizierten zwei Typen transformieren, läßt sich dabei a priori nicht eindeutig bestimmen. Das aktuelle Verhalten hängt zusätzlich zur Disposition auch von den Erwartungen der Spieler über die Zusammensetzung des subject pools ab<sup>11</sup>. Zumindest kann aber festgehalten werden, daß die „Messung“ von Individualisten und Kooperativen den Beobachtungen aus den GD-Spielen keinesfalls widerspricht.

Wir formulieren daher abschließend folgendes

#### Stilisiertes Faktum der experimentellen Evidenz:

Es existieren mindestens zwei verschiedene Verhaltensdispositionen (Typen), nämlich eine strikt defektive Disposition und eine schwach kooperative. Ein dritter, strikt kooperativer Typus wird beobachtet, allerdings mit einem äußerst geringen Populationsanteil. Diese drei Typen scheinen in stabiler Weise zu koexistieren.

Ein zentrales Anliegen der vorliegenden Arbeit wird es sein, einen Beitrag zur Erklärung des obigen stilisierten Faktums zu leisten. Bevor jedoch die eigenen theoretischen Bemühungen skizziert werden, soll zuvor eine der prominentesten und wichtigsten Theorien, die grundsätzlich geeignet scheint, die Koexistenz verschiedener Verhaltensdispositionen erklären zu können, vorgestellt werden. Das

---

<sup>11</sup> Allerdings stellen Ockenfels und Weimann (1999) fest, daß das Verhalten ihrer Vpn gegenüber der Gruppenzusammensetzung invariant ist.

folgende Kapitel widmet sich der Darstellung der Theorie der strategischen Funktion von Emotionen von Robert Frank. Kapitel 3.1 erläutert zunächst die wesentlichen Grundzüge der Theorie auf nicht technischem Niveau. Kapitel 3.2 liefert eine modelltheoretische Präzision der Theorie. Letztere bietet auch die Basis für eine eingehende Kritik (Kapitel 3.3.1). Kapitel 3.3.3 schließlich hinterfragt Franks Theorie vor dem Hintergrund experimenteller Arbeiten, sowohl aus dem psychologischen, als auch dem ökonomischen Bereich.

### 3. Robert Frank: Die strategische Funktion von Emotionen

*I need someone to believe in  
someone to trust  
I'd trust a country man, rather than a town man  
You can judge by his eyes, take a look if you can*

*P. Gabriel*

#### 3.1 Eine nicht-technische Diskussion

Ausgangspunkt der Frankschen Argumentation ist die Sicht des Kooperationsproblems als eines commitment-Problems. Wäre es für die Spieler im Zwei-Personen Gefangenen Dilemma möglich, eine glaubwürdige Selbstbindung auf die kooperative Strategie einzugehen, dann wäre damit in der Tat das Kooperationsproblem gelöst. Die Schwierigkeit besteht allerdings darin, einen überzeugenden Grund anzugeben, wie ein derartiges Kooperationsversprechen Glaubwürdigkeit erlangen kann.

Die Sichtweise des Kooperationsproblems als eines commitment-Problems geht auf Schelling (1960) zurück. In seinem Buch „The strategy of conflict“ diskutiert er Möglichkeiten einer glaubwürdigen Selbstbindung. Letztlich laufen derartige Lösungsansätze darauf hinaus, die Anreizstruktur des Gefangenen Dilemmas zu ändern. Dies kann etwa dadurch geschehen, daß – aus näher zu bestimmenden Gründen – sich die materiellen Payoffs des Spiels ändern. Schelling diskutiert das bekannte Beispiel eines Entführers, der vor der Frage steht, ob er sein Opfer freilassen soll oder nicht. Klarerweise handelt es sich hierbei um ein sequentielles Gefangenen Dilemma, und ebenso klar ist auch die einzige teilspielperfekte Lösung. Das Versprechen, die Identität des Kidnappers nach der Freilassung nicht preiszugeben, ist nicht glaubwürdig. In Antizipation dessen wird der Kidnapper seine Geisel nicht freilassen, sondern umbringen. Das Kooperationsversprechen entpuppt sich als purer cheap talk. Anders hingegen stellt sich die Lage dar, wenn es beispielsweise eine Handlung gäbe, die dazu geeignet ist, das Opfer nach der Freilassung erpressbar zu machen. Wenn die hiermit verbundenen negativen payoffs hinreichend hoch sind, dann ändert sich die Anreizstruktur des Spiels derart, daß eine Freilassung nun in den Bereich des Möglichen rückt.

Frank knüpft nun genau an diesen Gedanken der Änderung der Anreizstruktur des GD an, wie dies im übrigen wohl mehr oder weniger alle Theorien tun, die mit der Lösung des Kooperationsproblems befaßt sind. Allerdings dreht es sich bei ihm nicht um eine Änderung der materiellen payoffs. Wie noch ausführlicher darzulegen sein wird, können durch den Bruch eines Kooperationsversprechens vielmehr auch psychische Kosten entstehen, bei deren Berücksichtigung Defektion möglicherweise nicht mehr dominante Strategie ist.

Die zentrale Idee ist, daß Emotionen eine wichtige strategische Funktion in Situationen wie dem Gefangenen Dilemma zukommen kann. Die Argumentation Franks bezieht sich jedoch nicht nur, wenn auch hauptsächlich, auf das Kooperationsproblem, sondern vielmehr allgemein auf Situationen, in denen es möglicherweise „besser“ wäre, eine scheinbar irrationale Strategie zu verfolgen. Betrachtet sei beispielsweise folgende Situation (Frank 1988, S.5): B droht gegenüber A mit Vergeltung, wenn A eine bestimmte schädigende Handlung gegenüber dem B ausführt. Es ist aber bekannt, daß die Kosten der Vergeltung prohibitiv hoch sind. Die Ankündigung von Vergeltungsmaßnahmen wäre unter diesen Umständen offensichtlich eine leere Drohung. Man betrachte nun den Fall, daß B rachsüchtig ist, d.h. B besitze eine gewissermaßen natürliche Präferenz für Vergeltung, unabhängig von den Kosten der Rache. Wenn diese Tatsache außerdem auch allseits, und insbesondere dem A bekannt ist, dann kann eine solche, letztlich biologisch begründete Präferenz als glaubwürdige Drohung funktionieren. Der entscheidende Punkt ist, daß B sich mit einer derartigen Präferenz gegenüber einer Situation ohne den „biologischen bias“ verbessert. Aus der Sicht des B wäre es also durchaus wünschenswert, mit einer solchen Prädisposition ausgestattet zu sein.

Ein weiteres Beispiel mag die strategische Funktion von Emotionen verdeutlichen. Es ist eine – hoffentlich – allgemein verbreitete Erscheinung, daß Menschen über ein Gewissen verfügen, eine innere Instanz, die über das „richtig“ oder „falsch“, das „gut“ oder „schlecht“ der eigenen Verhaltensweisen urteilt. Vermutlich die meisten von uns werden schon einmal die Erfahrung gemacht haben, daß sie im Falle eines ethisch-moralischen Fehlverhaltens sich „irgendwie schlecht fühlen“, innere Schuldgefühle aufkeimen. Betrachten wir nun folgende Situation. Ein Individuum A gibt in einem GD ein Kooperationsversprechen gegenüber dem Mitspieler B ab. A verfüge über ein Gewissen und entwickle Schuldgefühle, wann immer ein derartiges Versprechen von ihm gebrochen wird. Wenn diese psychischen Kosten hinreichend hoch sind, wird er versuchen, diese zu vermeiden und sich ehrlich verhalten, also das Kooperationsversprechen einhalten. Nehmen wir weiterhin einmal an, die Fähigkeit zur Entwicklung solcher Schuldgefühle sei signalisierbar, und zwar hinreichend fälschungssicher signalisierbar. Man kann sich dann durchaus vorstellen, daß sich in einer Population von zwei stilisierten Typen (solchen mit und solchen ohne Gewissensinstanz) die „gewissenhaften“ Typen zusammenfinden und im GD erfolgreich kooperieren. In Abhängigkeit von der Verlässlichkeit des Signals wird man dann unter anderem auch evolutionär stabile Zustände erwarten können, in denen beide Typen koexistieren. Diese Idee wird in der formalen Diskussion der Frankschen Theorie detaillierter ausgearbeitet.



Die Beispiele sollten folgendes verdeutlicht haben:

- i) Die Tatsache, daß jemand prädisponiert ist, in bestimmten Situationen in einer bestimmten Weise emotional zu reagieren, kann die Anreizstruktur dieser Situationen entscheidend ändern. Sie kann bewirken, daß nunmehr, unter Berücksichtigung aller relevanten Payoffs, Strategien rational werden, die in einem Standardtheorierahmen unter Zugrundelegung der Eigennutzhypothese und nur unter Berücksichtigung der materiellen Payoffs als irrational qualifiziert werden müßten.
- ii) Emotionale, verhaltensmäßige Prädispositionen sind also grundsätzlich dazu geeignet, Menschen auf bestimmte Strategien zu „committen“. Eine emotionale Prädisposition kann demnach technisch auch begriffen werden als eine Beschneidung des Strategienraums durch „die Natur“. Damit ein solches committment in einer strategischen Interaktion wirksam werden kann, müssen die Mitspieler dieses committment kennen.

Die Punkte (i) und (ii) werfen natürlich einige Fragen auf. Zu (i) kann sehr grundsätzlich die Frage der Existenz der behaupteten Prädispositionen aufgeworfen werden. Das klassische behavioristische Paradigma innerhalb der Psychologie behauptet schließlich, der Mensch komme als tabula rasa auf die Welt, der sich im Zuge von Sozialisation und Enkulturation Einflüsse einprägen. Die Existenz angeborener Prädispositionen wird bestritten. Hierauf ist folgendes zu entgegnen: Wenngleich Frank oft den Begriff der Prädisposition im Sinne einer angeborenen Präferenz verwendet, so ist dies jedoch nicht die einzige Verwendungsvariante. Eine Prädisposition kann grundsätzlich auch erworben werden (vgl. Frank (1988), S.6). Der gesamte Bereich habituellen Verhaltens gehört hierher. So ist also der Franksche Ansatz nicht von vornherein grundsätzlich auf einen genetisch-biologischen festgelegt, sondern prinzipiell auch für eine Prägungstheorie offen. Jedoch ist kaum zu übersehen, daß Frank den Begriff der Prädisposition häufig in einem strikt biologischen Sinn verwendet und damit implizit deren Existenz in jenem Sinne voraussetzt. Eine fundierte Begründung des Frankschen Ansatzes würde damit letztlich eine Kritik des klassischen Behaviorismus erfordern. Da es uns an dieser Stelle vornehmlich um eine Darstellung der Frankschen Theorie zu dem Zweck geht, selbige ihrerseits zu kritisieren (vgl. Kapitel 3.3 dieser Arbeit), erscheint ein derartiges Unterfangen als überdimensioniert. Gleichwohl sei darauf hingewiesen, daß das behavioristische Paradigma in den letzten 25 Jahren durchaus Risse bekommen hat. Der Behaviorismus basiert auf den Beobachtungen Skinners (1938) und bereits früher Pawlows, daß bestimmte physiologische Reaktionen, die von einem unkonditionierten Stimulus ausgelöst werden, auf einen – scheinbar beliebigen anderen – konditionierten Stimulus übergehen können. Ver

abreicht man Laborratten etwa übel schmeckendes Wasser, was bestimmte Übelkeitssymptome bei den Tieren hervorruft, und setzt sie simultan einem zweiten Stimulus (etwa einem schrillen Klingellaut) aus, so kann bereits nach wenigen Wiederholungen der unkonditionierte Stimulus weggelassen werden – die Ratten entwickeln die besagten Symptome nunmehr auch dann, wenn sie nur den Klingellaut hören. Diese Beobachtungen führten letztlich zu der Vorstellung, daß auch der Mensch durch Konditionierung lernt und prinzipiell beliebig prägbar ist. Diese Schlußfolgerung hat sich mittlerweile als falsch erwiesen, allerdings hat der entsprechende Nachweis vergleichsweise lange gedauert. Erst in den sechziger Jahren zeigten Garcia und Koelling (1966), daß Ratten nicht auf beliebige Stimuli geprägt werden können. Eine erste Gruppe von Ratten wurde einem Elektroschock ausgesetzt, während sie Wasser trank. Eine zweite Gruppe wurde während des Trinkens mit Röntgenstrahlen beschossen (was Übelkeit hervorruft). Beiden Gruppen wurde simultan ein zweiter Stimulus präsentiert (Lärm und ein Blitzlicht in der ersten Gruppe, in der zweiten Gruppe wurde dem Wasser ein bestimmter Geschmacksstoff beigefügt). Beide Gruppen erlernten schnell die typische Skinnersche Vermeidungsreaktion. Wurde ihnen nach wenigen Wiederholungen nur der konditionierte Stimulus präsentiert (das Blitzlicht bzw. der Geschmack) hörten sie auf zu trinken, sobald das Signal kam. In zwei weiteren Anordnungen wurden nunmehr die Signale vertauscht. Lärm und Blitzlicht wurde kombiniert mit Röntgenstrahlung, während der Geschmacksstoff mit dem Elektroschock kombiniert wurde. Die erstaunliche Beobachtung war, daß in beiden Anordnungen eine Vermeidungsreaktion nicht erlernt wurde. Frank (ebd., S. 148 f.) interpretiert den Befund wie folgt:

*„Rats apparently have no difficulty learning to associate a noise and light with an electric shock, or a distinct flavor with nausea. Garcia and Koelling argued that the rat's brain evolved to be especially receptive to such associations. A causal relationship between nausea and the most recent novel taste is, after all, inherently plausible. Physical pain, similarly, is frequently linked both to sounds and changes in the visual field. By contrast, a causal connection between novel taste and nonabdominal physical pain is inherently implausible. Nor are there likely to have been many circumstances in nature where a nonfood stimulus triggered feelings of nausea. Because some types of association are so implausible, the rat's brain can save capacity by not even considering them. Any rat that devoted scarce cognitive resources to investigating such relationships would probably not have survived for long.“*

Die Garcia-Experimente legen den Schluß nahe, daß Ratten (aber warum nicht auch Menschen?) mit gewissen fertigen Verschaltungen im zentralen Nervensystem auf die Welt kommen und nicht als eine tabula rasa. Die Möglichkeiten der Prägung (des Umweltlernens) sind offenkundig begrenzt. Es gibt neurale Restriktionen, die, so die naheliegende Vermutung, im Verlauf der Evolution als vorteilhaft selektiert wurden. Der Schluß, daß in gleichartiger Weise menschliches Verhalten restringiert sein könnte, ist zwar nicht zwingend aus den zitierten Beobachtungen zu ziehen, er erscheint aber im Lichte der Garcia-Befunde als plausibel. Wir wollen deshalb im folgenden die Franksche Vermutung angeborener Verhaltensdispositionen akzeptieren. Diese Frage ist von nicht ganz unerheblicher Bedeutung: Angeborene Verhaltensweisen oder Verhaltensweisen, die sich nur im Korridor einer vorgegebenen Disposition entwickeln können, befinden sich nämlich, zumindest teilweise, außerhalb der individuellen Kontrolle durch das Bewußtsein. Gerade diese Eigenschaft macht sie zu geeigneten Kandidaten für erfolgreiche „commitment-devices“. Jemand der rachsüchtig, jähzornig, ausbeutungsavers etc. **von Natur** aus ist, wird sich gemäß seiner Prädisposition verhalten, **weil er nicht anders kann!**

Selbst wenn eine Prädisposition existiert, müssen, damit sie strategisch wirksam werden kann, die Mitspieler die Prädisposition kennen (Punkt (ii)). Offenkundig reicht die rein verbale Kommunikation des Sachverhaltes einer solchen natürlichen Präferenz nicht aus. Die Behauptung „Ich bin vertrauenswürdig und halte Kooperationsversprechen. Vertrau mir also!“ im GD ist cheap talk, selbst dann, wenn diese Präferenz tatsächlich vorhanden ist. Der/die Mitspieler können schlicht die Behauptung allein aufgrund des in ihr kommunizierten Inhaltes nicht verifizieren. Damit sind wir bei einem Schlüsselement der Frankschen Theorie, der Annahme nämlich, daß verhaltensmäßige Prädispositionen aufgrund bestimmter Signale identifiziert werden können. Dabei handelt es sich um körperliche Symptome oder mehr oder weniger beobachtbare körperliche Reaktionen. Frank nennt hier eine ganze Reihe von Beispielen: So könne vor allem der Gesichtsausdruck einer Person Aufschluß über deren Gefühlszustand geben (ebd., S. 120 ff). Die Tatsache, daß für bestimmte elementare Gefühlslagen kulturübergreifende Gesichtsausdrücke existieren, spreche im übrigen für eine biologische Prägung in dieser Hinsicht. Als weitere Signalquellen werden die Augen (S. 126) genannt sowie eine Reihe weiterer, zentralnervös verursachter Reaktionen: Pupillenerweiterung (S. 127), Erröten, etwa infolge eines Schuldgefühls (S. 127), das Ausmaß des Speichelflusses (S. 129), Färbung und Höhe der Stimme (S. 129 f.), Körperhaltung und –sprache allgemein (S.130 f.).

Ein wichtiges weiteres Element der Frankschen Theorie ist die Annahme, daß Signale, zumindest in einem bestimmten Maße, fälschungssicher sein müssen, ansonsten würde sich ihr Informationswert sofort auf Null reduzieren. Man stelle sich etwa eine Population, bestehend aus defektiven und kooperativen Spielern, vor. Es sei der Einfachheit halber angenommen, es existiere ein perfektes Signal, welches den Typ eines Spielers offenbart. Sicherlich werden in diesem Fall Kooperative nur mit kooperativen Spielern interagieren, während die Defekteure unter sich bleiben. Da Kooperation den höheren Payoff sichert, werden Defekteure langfristig aussterben. Man stelle sich nun vor, perfekte Mimikry sei möglich, das heißt, das Signal, ein kooperativer Typ zu sein, könne perfekt und kostenlos imitiert werden. Defekteure werden nun natürlich vortäuschen, kooperativ zu sein, sich mit Kooperativen matchen und diese ausbeuten. Das Ergebnis ist, daß langfristig nur die Mimikry-Defekteure überleben werden. Der Selektionsvorteil, den das Signal ursprünglich begründete, wird durch die Möglichkeit perfekter Mimikry vollständig entwertet.

In diesem Zusammenhang gewinnt die Tatsache an Bedeutung, daß es sich bei den oben genannten Signalen um körperliche Symptome und Reaktionen handelt, die vollständig oder zumindest teilweise der willentlichen Kontrolle entzogen sind. Der Lügner, der bei seiner Lüge ertappt wird und dessen Pein sich in der Errötung seines Gesichts äußert, kann eben diese Reaktion nicht kontrollieren und steuern, er ist ihr vielmehr ausgeliefert. „Flight-or-fight“-Situationen provozieren autonome Reaktionen des vegetativen Nervensystems, die ebenfalls dem freien Willen entzogen sind: Die Herzschlagrate erhöht sich, ebenso die Atmungsfrequenz, die Stimmhöhe steigt, die Transpiration nimmt zu etc. Frank verweist ferner darauf, daß eine ganze Reihe von Muskelgruppen des Gesichts für große Teile der menschlichen Population willentlicher Beeinflussung entzogen sind. Kurzum: Es scheint also eine ganze Reihe von Reaktionsweisen beim Menschen zu existieren, die dieser nicht oder nur unvollständig willentlich kontrollieren kann und die in einer direkten Beziehung zu dem verursachenden emotionalen Zustand stehen. Da diese Signale in einem bestimmten Grade fälschungssicher sind, erlauben sie Außenstehenden, mit einer gewissen Verlässlichkeit auf den zugrundeliegenden emotionalen Zustand zu schließen. Da dieser positiv korreliert ist mit elementaren verhaltensmäßigen Prädispositionen, erlauben Signale letztlich den Schluß auf den Typ eines Gegenübers. Das Skelett der Frankschen Theorie läßt sich schematisch etwa wie folgt darstellen.

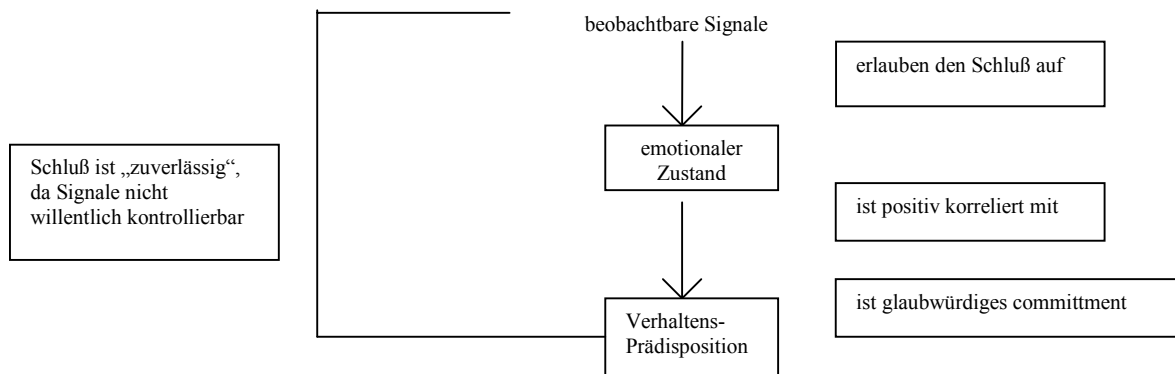


Abb. 8: Schematische Darstellung der Frankschen Theorie

Die Existenz fälschungssicherer, damit vertrauenswürdiger Signale, eröffnet offensichtlich die Möglichkeit zu selektiver Interaktion. Damit wird es beispielsweise möglich, daß nur Spieler miteinander interagieren, deren Signal einen geeigneten „Wert“ annimmt. So könnte man sich vorstellen, daß in einem GD nur jeweils hinreichend kooperative Typen miteinander interagieren. Da jeder der Spieler weiß, daß der andere mit hinreichend hoher Wahrscheinlichkeit glaubwürdig (biologisch) auf die kooperative Strategie committet ist, werden beide kooperieren. Es wird nunmehr also möglich, die Kooperationsvorteile, zumindest partiell, zu realisieren. Damit ist keineswegs mehr klar, daß eine derartige Population unter der Annahme der Selektion fitterer Strategien notwendigerweise in einmütiger Defektion enden wird. Man kann vielmehr zeigen, daß man im allgemeinen gemischte Populationszustände, also Koexistenz von Defektoren und Kooperatoren erwarten darf. Allerdings, so wird sich herausstellen, ist dieses Ergebnis nur zwingend, wenn einerseits die bisher skizzierte Franksche Theorie tatsächlich richtig ist und andererseits einige weitere, durchaus kritische Annahmen erfüllt sind. Bevor wir uns allerdings der Kritik der Frankschen Theorie zuwenden, soll zunächst die Argumentation im Rahmen eines formalen Modells präzisiert werden.

### 3.2 Frank II: Eine modelltheoretische Diskussion

Betrachtet sei eine Population, deren Mitglieder zwei Typausprägungen aufweisen können. Entweder handelt es sich um ehrliche Spieler (H für honest) oder um unehrliche (D für dishonest). H-Spieler sind mit einem Gewissen ausgestattet, d.h. wann immer sie ein (Kooperations-) Versprechen abgeben und dieses brechen entstehen ihnen Schuldgefühle (psychische Kosten, negative innere Payoffs). Diese Kosten können die Anreizstruktur in einem GD für diesen Typ von Spielern

derart ändern, daß sich Kooperation nunmehr tatsächlich für ihn lohnt. Ein Populationsmitglied ist mit folgenden Möglichkeiten konfrontiert:

- i) Es sucht sich einen Partner und geht mit diesem ein Joint Venture ein, in dem Effizienzgewinne realisiert werden können. Es wird angenommen, daß sich Spieler gemäß ihres Typs verhalten, d.h. ein H-Typ wird sich, wenn er das joint venture eingeht, immer ehrlich verhalten, ein D-Typ immer unehrlich. Treffen also in einem Joint Venture zwei H-Typen aufeinander, so kooperieren beide und die Effizienzgewinne können realisiert werden ( $x_3$ ), trifft ein H-Typ auf einen D-Typ, so wird der Ehrliche ausgebeutet ( $x_1$ ) und der D-Typ erzielt den maximal möglichen Payoff ( $x_4$ ), treffen zwei D-Typen einander, so erzielen beide nur  $x_2$ , wobei gelte  $x_4 > x_3 > x_2 > x_1$ . Die Payoffstruktur des Joint Ventures ist identisch zu der eines GD und in folgender Matrix zusammengefaßt.

	H	D
H	$x_3$	$x_1$
D	$x_4$	$x_2$

Tabelle 3: Ein Gefangenen Dilemma

- ii) Spieler können eine work-alone option wählen, die ihnen einen Payoff von  $x_2$  einbringt.

Das Spiel ist one-shot, es bestehen also keine Möglichkeiten, nicht-kooperatives Verhalten im joint venture in der Zukunft zu bestrafen. Diese Modellierungsweise schließt die Möglichkeit von Reputationseffekten aus. Dem Gewissen der H-Spieler kommt hier potentiell die Rolle des commitment device zu. Wie oben bereits ausführlich dargelegt wurde, reicht die bloße Existenz einer Verhaltensprädisposition nicht aus, um das Kooperationsproblem zu lösen. Potentielle Partner im joint venture müssen auch in der Lage sein, diese Prädisposition zu erkennen. Dabei ist es offensichtlich nicht ausreichend, diese einfach verbal zu kommunizieren („I have a conscience. Trust me.“, Frank (1987), S.594), da es sich hierbei schlicht um cheap talk handeln würde. Benötigt wird ein glaubwürdiges Signal. Wie oben ausführlich dargelegt wurde, existieren eine Reihe glaubwürdiger (weil nicht willentlich kontrollierbarer körperlicher) Reaktionen, die über den emotionalen Zustand eines Menschen Auskunft zu geben geeignet sind. Der Gefühlszustand seinerseits mag wiederum Auskunft geben über die in Frage stehende Prädisposition. Beispielhaft könnte man sich im vorliegenden Fall das unfreiwillige Erröten vorstellen, wenn jemand lügt. Es sei also im folgenden ange

nommen, daß tatsächlich ein solches beobachtbares Signal existiere, das über die Vertrauenswürdigkeit der Spieler Auskunft gibt. Für den  $i$ -ten Agenten der Population sei das Signal gegeben als

$$S_i = \mu_i + \varepsilon_i, \tag{3-1}$$

wobei  $\varepsilon_i$  i.i.d. mit Erwartungswert Null sei,  $\mu_i = \mu_H$ , falls  $i$  ein H-Typ ist, und  $\mu_i = \mu_D$ , falls  $i$  ein D-Typ ist. Das Signal sei nur bezüglich seiner  $\mu$ -Komponente vererbbar.

Der Fall eines perfekten Signals wurde oben (S.31 f.) bereits kurz diskutiert, wir gehen daher sofort zum interessanteren allgemeinen Fall eines imperfekten Signals über. Zu Zwecken größerer Anschaulichkeit arbeitet Frank mit zwei speziellen Dichtefunktionen<sup>12</sup>  $f_H(S), f_D(S)$ , die in Abbildung 9 wiedergegeben sind. Das Signal  $S$  sei also für die H-Typen auf dem Träger  $[L_H, U_H]$  und für die D-Typen auf dem Träger  $[L_D, U_D]$  verteilt. Niedrige Werte von  $S$  signalisieren geringe Vertrauenswürdigkeit. Man beachte, daß die beiden Träger nicht identisch sind. Dies hat zur Folge, daß Subjekte mit Signalen in den Intervallen  $[L_D, L_H]$

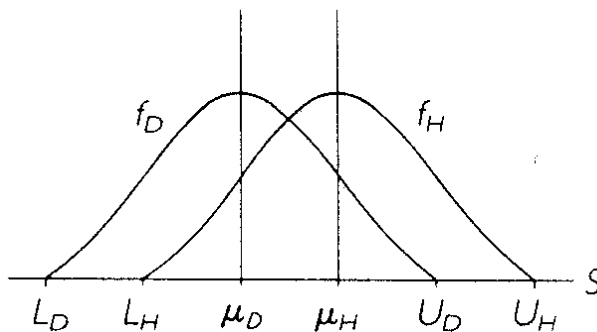


Abb. 9: Dichtefunktionen des Signals für Vertrauenswürdigkeit

Quelle: Frank (1988), S. 263

sowie  $[U_D, U_H]$  mit Sicherheit als D-, bzw. H-Typen identifiziert werden können (diese Besonderheit wird in Kapitel 3.3.1 einer ausführlichen Kritik unterzogen). Bezeichne nun  $h$  den Populationsanteil der H-Typen. Die Wahrscheinlichkeit, daß ein Individuum  $j$  mit dem emittierten Signal  $S_j$  ein H-Typ ist, berechnet sich dann nach Bayes Formel als

$$Pr\{H_j | S_j\} = \frac{hf_H(S_j)}{hf_H(S_j) + (1-h)f_D(S_j)}. \tag{3-2}$$

<sup>12</sup> Zu der Frage der Konstruktion der Dichten aus Dichtefunktionen der Standard-Normalverteilung vgl. Frank 1987, S.596, Fußnote 8.

Betrachtet sei dann das Entscheidungsproblem des Spielers  $i$ , der ein H-Typ ist und entscheiden muß, ob er mit dem Individuum  $j$ , das ein Signal  $S_j$  aussendet ein joint venture eingeht oder die work alone option wählt. Die Bedingung dafür, daß  $i$  das joint venture wählt, ist gegeben als

$$E(X_{ij}|S_j) = x_3 Pr\{H_j|S_j\} + x_1 [1 - Pr\{H_j|S_j\}] \geq x_2. \quad (3-3)$$

Bezeichne nun  $S^* = S^*(h)$  den kritischen Wert, für den (3-3) mit Gleichheit erfüllt ist, für den also  $i$  indifferent ist zwischen joint venture und work alone option. Man kann leicht zeigen, daß  $S^*(h)$  eine in  $h$  abnehmende Funktion ist. Für alle Signalwerte  $S > S^*$  wird  $i$  das joint venture wählen, bei allen Werten unterhalb  $S^*$  die work alone option. Was zur Komplettierung des Modells noch benötigt wird, ist eine Annahme über den matching-Prozess. Jeder Spieler würde natürlich gern mit demjenigen Individuum interagieren, das den höchsten Signalwert trägt. Für einen H-Typ maximiert dies die Kooperationswahrscheinlichkeit, für einen D-Typ die Ausbeutungswahrscheinlichkeit. Der „Besitzer“ dieses Signals wird aber seinerseits nicht bereit sein, mit beliebigen anderen Individuen zu interagieren, sondern er wird vielmehr bestrebt sein, ein Individuum mit gleich oder ähnlich hohem Signalwert zu attrahieren. Man sieht leicht, daß alle Individuen, deren Signalwert oberhalb  $S^*$  liegt, sich in Paaren annähernd gleicher Signalwerte zu joint ventures zusammenfinden werden. In absteigender Reihenfolge, ausgehend vom maximalen Signalwert, bilden die unmittelbar benachbarten Signale auf dem Signalkontinuum Paare.

Wir stellen uns nun folgende Frage: Betrachtet sei eine unendlich große Population, die vollständig aus D-Typen besteht und in der ein marginal kleiner Anteil von H-Mutanten auftaucht. Kann es den Mutanten gelingen, die Population zu invadieren? Wir bestimmen zur Beantwortung dieser Frage zunächst den aggregierten erwarteten Payoff der H-Mutanten für den Fall, daß  $h \rightarrow 0$ . Dieser berechnet sich als

$$\lim_{h \rightarrow 0} E(H) = x_3 \int_{U_D}^{U_H} f_H(S) dS + x_2 \int_{L_H}^{U_D} f_H(S) dS. \quad (3-4)$$

Der zweite Summand resultiert aus der Tatsache, daß in einer Population, die faktisch zu 100% aus D-Typen besteht, die work alone option die beste Antwort ist. Dies gilt lediglich dann nicht, wenn ein Gegenüber sicher als ein H-Typ identifiziert werden kann - dies ist aber bekanntermaßen nur im Intervall  $[U_D, U_H]$  der Fall (erster Summand). Da H-Typen unter sich bleiben, so gilt dies natürlich auch für D-Typen (sie finden schlicht keinen H-Typ, der bereit wäre eine Interaktion



mit ihnen einzugehen) und der erwartete Payoff der D-Typen für  $h \rightarrow 0$  beträgt gerade  $x_2$ . Damit ist aber klar, daß

$$\lim_{h \rightarrow 0} E(H) > \lim_{h \rightarrow 0} E(D) = x_2. \quad (3-5)$$

Unter jeder payoff-monotonen Selektionsdynamik, d.h. unter jeder Dynamik, die eine positive Korrelation zwischen erzieltm Payoff und Wachstumsrate einer Strategie annimmt, wird damit der Populationsanteil der H-Typen zunehmen.

Betrachtet sei nun der Grenzfall  $h \rightarrow 1$ . In einer Population, die fast ausschließlich aus H-Typen besteht, ist es für einen H natürlich beste Antwort, immer das joint venture zu wählen. Der erwartete Payoff der H für  $h \rightarrow 1$  strebt daher gegen  $x_3$ . Da H-Typen immer auf das joint venture eingehen, finden D-Typen im Intervall  $[L_H, U_D]$  immer einen Partner, den sie ausbeuten können. Bei Signalen unterhalb von  $[L_H]$  jedoch ist klar, daß es sich um einen D-Typ handeln muß. Ein H würde deshalb eine Interaktion ablehnen. Der erwartete Payoff der D-Typen ist damit gegeben als

$$\lim_{h \rightarrow 1} E(D) = x_2 \int_{L_D}^{L_H} f_D(S) dS + x_4 \int_{L_H}^{U_D} f_D(S) dS. \quad (3-6)$$

Dieser Ausdruck ist dann größer als  $x_3$ , wenn  $x_4$  hinreichend groß wird oder die Träger der Dichtefunktionen sich hinreichend stark überlappen. Im folgenden sei angenommen, daß tatsächlich  $\lim_{h \rightarrow 1} E(D) > x_3$ . Dann ist klar, daß der Populationsanteil der H-Typen abnimmt (wiederum „vernünftige“ Selektionsdynamiken vorausgesetzt).

Die bisher entwickelte Argumentation deutet bereits stark auf ein inneres Gleichgewicht hin: Ausgehend von einer monomorphen Population von D-Typen kann eine marginale Zahl von H-Mutanten die Population invadieren, unter geeigneten Annahmen an Payoffs und/oder Dichtefunktionen gilt analoges für den entgegengesetzten Grenzfall auch – D-Mutanten invadieren die monomorphe H-Population. Um jedoch genaueres, insbesondere über Eindeutigkeit und Stabilität aussagen zu können, wäre es wünschenswert, den Verlauf der erwarteten Payoffs als Funktionen von  $h$  zu kennen.

Gegeben Gleichungen (3-2) und (3-3) kann man leicht den kritischen Wert  $S^*$  als eine Funktion von  $h$  bestimmen. Gegeben  $S^*(h)$  lassen sich leicht die Populationsanteile von H- und D-Typen bestimmen, deren Signal oberhalb  $S^*$  liegt. Bezeichne  $P_H$  den Populationsanteil der H-Typen,  $P_D$  den der D-Typen mit  $S > S^*$ . Diese sind dann gegeben als

$$P_H = \int_{S^*}^{U_H} f_H(S) dS, \tag{3-7}$$

$$P_D = \int_{S^*}^{U_H} f_D(S) dS. \tag{3-8}$$

Der erwartete Payoff für einen H ermittelt sich dann wie folgt. Für alle  $S > S^*$  wird sich ein H auf das joint venture einlassen. Bezeichne

$$\lambda = \frac{hP_H}{hP_H + (1-h)P_D} \tag{3-9}$$

den Anteil aller Individuen mit  $S > S^*$ , die tatsächlich H sind. Mit Wahrscheinlichkeit  $\lambda$  wird also im joint venture der kooperative Payoff realisiert, mit Wahrscheinlichkeit  $1 - \lambda$  wird ein D-Typ getroffen, der den H ausbeutet. Mit Wahrscheinlichkeit  $1 - P_H$  wird die work alone option gewählt. Der gesamte erwartete Payoff ermittelt sich damit zu

$$E(H) = (1 - P_H)x_2 + P_H[\lambda x_3 + (1 - \lambda)x_1]. \tag{3-10}$$

Analog ermittelt man

$$E(D) = (1 - P_D)x_2 + P_D[\lambda x_4 + (1 - \lambda)x_{21}]. \tag{3-11}$$

Beide Payoffs sind wachsend in  $h$ . Für die speziellen Dichten aus Abbildung 9 kann Frank zeigen, daß sich (3-10) und (3-11) exakt einmal schneiden. Das innere Gleichgewicht ist also eindeutig. Außerdem schneiden sich (3-10) und (3-11) in der „richtigen Weise“, nämlich so, daß das gemischte Gleichgewicht stabil ist.

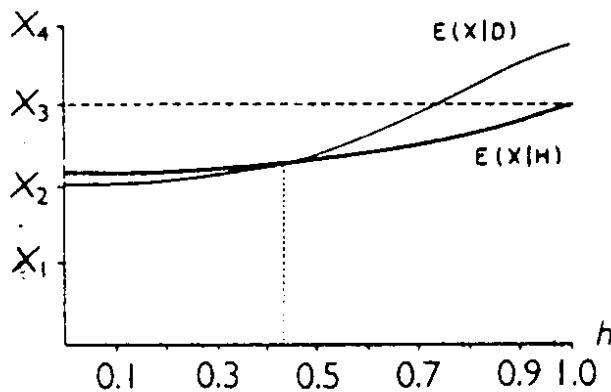


Abb. 10: Erwartete Payoffs in Abhängigkeit von  $h$ ; Quelle: Frank (1988), S.266

Für alle  $h < h^*$  gilt offensichtlich gerade  $E(H) > E(D)$ , so daß in diesem Bereich der Populationsanteil der H steigt, während für alle  $h > h^*$  gilt  $E(H) < E(D)$ , so daß also h sinkt. Natürlich hängt der Gleichgewichtswert  $h^*$  von den Parametern des Modells ab. Allgemein gilt jedoch, daß  $0 < h^* < 1$ .

Frank hat damit – auch formal – folgendes gezeigt:

- i) Gegeben eine Mutante, die in ihrem Verhalten biologisch prädisponiert ist, die also beispielsweise über ein Gewissen verfügt;
- ii) ferner gegeben, daß die „Gewissensbisse“ so hohe psychische Kosten bei der Mutante verursachen, daß diese es im GD vorzieht, die kooperative Strategie zu spielen (gegeben, sie hat sich auf das joint venture eingelassen);
- iii) außerdem gegeben, daß die Vertrauenswürdigkeit einer gewissenhaften Person hinreichend verlässlich signalisiert werden kann, so daß die Verhaltensprädisposition ein glaubwürdiges commitment für Mitspieler darstellt; dann läßt sich zeigen, daß
- iv) diese Mutante erfolgreich eine D-Population invadieren kann, und
- v) unter den speziellen Annahmen bezüglich der Verteilungen des Signals ein eindeutiges und stabiles inneres Gleichgewicht existiert, in dem ehrliche und unehrliche Typen koexistieren.

### 3.3 Kritik des Frank-Modells

Es gibt zahlreiche Möglichkeiten, das Frank-Modell zu kritisieren. Diese resultieren allein schon aus der Tatsache der komplexen, indirekt evolutionären Theorie, die ihm zugrunde liegt. Im folgenden sollen jedoch nur zwei der wichtigsten Kritikpunkte behandelt werden. In Abschnitt 3.3.1 wird das Augenmerk auf eine scheinbar harmlose Annahme des Modells gelenkt, nämlich die nicht-identischer Träger der Dichtefunktionen des Signals für H-Typen und D-Typen. Harrington (1989) hat gezeigt, daß diese Annahme für die Resultate in der Tat fundamental ist. Unsere Diskussion wird ergeben, daß die Annahme als ad hoc qualifiziert werden muß.

Die Annahme der Existenz hinreichend fälschungssicherer Signale ist offenkundig das Herzstück der Frankschen Theorie. Man kann diesen Kern unter mehreren Gesichtspunkten angreifen. Zum einen kann hinterfragt werden, ob empirisch überhaupt die Existenz von Signalen, die Aufschluß über die emotionale Verfassung eines Menschen geben, nachgewiesen werden kann. Eine solche Kritik trägt in der Tat nicht sehr weit. Nicht nur die anekdotische Evidenz, die vermutlich die meisten Menschen mit derartigen Symptomen beisteuern könnten, läßt diesen Einwand schwach erscheinen. Ein Blick in die einschlägige psychologische Lite

ratur (vgl. Zuckerman (1981) und die dort zitierte Literatur) lehrt, daß die von Frank angenommenen Signale tatsächlich existieren. Es scheint auch kaum außer Frage zu stehen, daß diese Signale in einem gewissen Grade autonomen körperlichen Reaktionen entspringen, so daß sie willentlicher Kontrolle (partiell) entzogen sind. Damit scheint auch die Annahme der Fälschungssicherheit wenig problematisch. Jedoch impliziert die Franksche Theorie auch, daß derartige existierende Signale von anderen zuverlässig wahrgenommen werden. Diese Annahme, so wird sich zeigen, erweist sich jedoch als durchaus problematisch.

### 3.3.1 Die Harrington-Kritik

Franks Theorie erscheint faszinierend, intuitiv plausibel und auch in ihrer Formalisierung sehr überzeugend. Zumindest aber die modellhafte Formulierung hält einer näheren Untersuchung nicht stand. Es zeigt sich vielmehr, „that these [Frank’s] findings, while fascinating, are based upon an implausible foundation“ (Harrington 1989, S.588). Als kritisch erweist sich die Annahme unterschiedlicher Träger der Dichtefunktionen des Signals. Gibt man diese Annahme auf, brechen Franks Resultate vollständig zusammen. Es ist das Verdienst Harringtons, dies gezeigt zu haben.

Etwas weniger technisch formuliert lautet die zentrale Annahme bei Frank, daß Bereiche für das Signal existieren, in denen der Typ des Gegenübers perfekt identifiziert werden kann. Zur Erinnerung seien die beiden Dichtefunktionen noch einmal dargestellt:

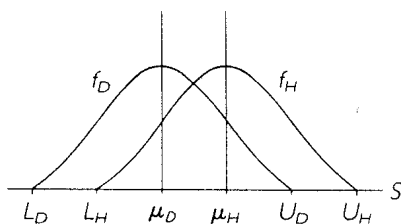


Abbildung 9: Dichten des Signals für Vertrauenswürdigkeit

In den Intervallen  $[L_D, L_H]$  sowie  $[U_D, U_H]$  kann der Typ eines Spielers mit Sicherheit aus dem empfangenen Signal bestimmt werden. Harrington stellt zu recht fest, daß diese Annahme a priori keineswegs zwingend ist. Während Harrington allerdings nach dieser Feststellung sofort dazu übergeht, die Konsequenzen der Aufgabe dieser Annahme deutlich zu machen, scheint eine tiefere Reflexion des Problems angezeigt. Was hat Frank dazu bewogen, nichtidentische Träger der Verteilungen zu unterstellen? Es war bereits darauf hingewiesen worden, daß Signale nur dann ihre Funktion erfüllen, wenn es nicht möglich ist, die Signale perfekt zu fälschen. Frank (1987) verweist nur kurz auf genau diesen Punkt.

Er behauptet, die relative Position der Dichtefunktionen spiegele genau die Annahme imperfekter Mimikry wider: „...the relative position of the two densities reflects imperfect mimicry...“ (ebd., S. 596). Aber diese Behauptung ist im allgemeinen falsch. Die Annahme nichtidentischer Träger der Dichtefunktionen ist keine notwendige Bedingung, um imperfekte Mimikry zu modellieren. Um diesen Punkt zu verstehen sollte man etwas präziser fassen, was imperfekte mimikry charakterisiert. Frank (1988) selbst erläutert an zahlreichen Stellen seines Buches, daß es lediglich darauf ankomme, daß Signale in einem statistischen Sinne zuverlässig sind:

*„The critical assumption behind the commitment model, again, is that people can make reasonable inferences about character traits in others. By „reasonable inference“ I do not mean that it is necessary to be able to predict other people’s emotional predispositions with certainty. Just as a weather forecast of a 20 percent chance of rain can be invaluable to someone who must plan outdoor activities, so can probabilistic assessments of character traits be of use to people who must choose someone to trust.“* (ebd., S.12)

Und kurz zuvor heißt es:

*„Much of the time, the practical means for accomplishing these commitments will be emotions that have observable symptoms. I will later survey persuasive evidence that at least some of these emotions are inborn. But even if they were transmitted only by cultural indoctrination, they would serve equally well. What is necessary in either case is that people who have them be observably different, **on the average** [Herv. v.A.], from those who do not.“* (ebd., S.11)

Typen müssen im Durchschnitt diskriminierbar sein. Übersetzt in technische Termini heißt das: Die Dichtefunktionen des Signals müssen sich hinsichtlich ihres Mittelwertes unterscheiden. Es ist aber in keinster Weise erkennbar, warum es zur Abbildung imperfekter Mimikry notwendig sein sollte, daß sich die Dichten auch hinsichtlich ihres Trägers unterscheiden.

Natürlich spiegelt die obige Abbildung 9 auch eine Situation imperfekter Mimikry wider, denn die beiden Dichten unterscheiden sich ja auch hinsichtlich ihres Mittelwertes. Die Tatsache ihrer unterschiedlichen Träger ist jedoch erlässlich, wird schlicht nicht benötigt, um das Geforderte abzubilden. Methodisch gesehen handelt es sich also um eine ad-hoc-Annahme.

Damit dürften Hintergrund und Berechtigung der Harrington-Kritik hinreichend erhellt sein und wir können uns der Kritik selbst zuwenden.  $f_H(S)$  und  $f_D(S)$

haben nun also denselben Träger, d.h.  $L_H = L_D = L$  sowie  $U_H = U_D = U$ . Die bedingte Wahrscheinlichkeit, daß ein Akteur mit Signal  $S_j$  ein H-Typ ist, ist dann einfach

$$Pr(H|S_j) = \frac{hf_H(S_j)}{hf_H(S_j) + (1-h)f_D(S_j)}. \quad (3-12)$$

Für  $f_H(S_j) > 0, f_D(S_j) > 0 \forall S_j \in [L, U]$  gilt natürlich  $Pr(H|S_j) \in (0,1) \forall h \in (0,1)$ . Es existiert also kein perfektes Signal. Der erwartete Payoff für einen H-Typ ist dann gegeben als

$$E(H|S_j) = Pr(H|S_j)x_3 + (1 - Pr(H|S_j))x_1. \quad (3-13)$$

Ein H-Typ wird natürlich nur dann das joint venture eingehen, falls  $E(H|S_j) \geq x_2$ . Bezeichne nun  $\Gamma(h) = \{S_j | E(H|S_j) \geq x_2\}$  die Menge aller Signale, für die das joint venture vorgezogen wird. Falls also  $S_j \notin \Gamma(h)$  wird ein H-Typ die work alone option wählen, andernfalls wird er einen Partner für ein joint venture mit annähernd gleichem Signalwert finden. Es gilt dann folgender

**SATZ: (Harrington 1989)**

$\exists \underline{h} \in (0,1)$ , so daß  $\Gamma(h) = \emptyset \forall h \in [0, \underline{h}]$ .

**Beweis:**

$$\begin{aligned} E(H|S_j) \geq x_2 &\Leftrightarrow \\ Pr(H|S_j)x_3 + (1 - Pr(H|S_j))x_1 &\geq x_2 \Leftrightarrow \\ \frac{hf_H(S_j)}{hf_H(S_j) + (1-h)f_D(S_j)}x_3 + \frac{(1-h)f_D(S_j)}{hf_H(S_j) + (1-h)f_D(S_j)}x_1 &\geq x_2 \Leftrightarrow \\ h \geq \hat{h}(S_j) &\equiv \frac{f_D(S_j)(x_2 - x_1)}{f_H(S_j)(x_3 - x_2) + f_D(S_j)(x_2 - x_1)} \end{aligned} \quad (3-14)$$

(3-14) ist also nur eine Umformung der Bedingung dafür, daß ein H-Typ sich auf ein joint venture einläßt. Klar ist: Da  $x_3 > x_2 > x_1$  und  $f_H(S_j) > 0, f_D(S_j) > 0$  folgt  $\hat{h}(S_j) \in (0,1) \forall S_j \in [L, U]$ . Unter der Annahme, daß die Funktion  $\hat{h}(S_j)$  stetig in  $S$  und das Intervall  $[L, U]$  kompakt ist, existiert ein  $\underline{h} = \min\{\hat{h}(S_j) | S_j \in [L, U]\}$ .

Klar:  $\hat{h} \in (0,1) \Rightarrow h \in (0,1) \forall S_j \in [L, U]$ . Offensichtlich gilt:

$\forall h \in [0, \underline{h}] \Rightarrow h < \hat{h}(S_j) \forall S_j \in [L, U]$ , also ist (3-14) nicht erfüllt, also gilt

$\Gamma(h) = \emptyset \forall h \in [0, \underline{h}]$ .

□

Die Intuition des Beweises ist vollkommen einfach. Gegeben wir befinden uns in einer Population, die fast ausschließlich aus D-Typen besteht, und ferner gegeben, daß kein perfektes Signal existiert. Selbst wenn das Signal sehr gut ist (annähernd perfekt), ist es bedeutungslos gegenüber der Tatsache, daß die Population zu faktisch 100% aus D-Typen besteht. Die Wahrscheinlichkeit, trotz eines solchen H-Signals einen D-Typen zu treffen ist so hoch, daß sich die Interaktion niemals lohnt. Das Signal wird durch die extreme Populations-Häufigkeitsverteilung gänzlich entwertet. Harrington verdeutlicht das Problem an einem numerischen Beispiel: Angenommen, für ein gegebenes Signal  $S'$  gelte  $f_H(S') = 100f_D(S')$ ,  $S'$  ist also ein extrem gutes Signal für Ehrlichkeit. Für  $h=0.001$  ist dennoch die a posteriori-Wahrscheinlichkeit, daß der Emittent des Signals ein D-Typ ist, 0.91!

Die Folgen sind klar: Wenn der kleine Anteil von H-Mutanten immer die work alone option wählt, dann unterscheidet sich ihr Payoff nicht von dem der D-Typen und der Populationsanteil der H-Typen wird nicht steigen. Die monomorphe D-Population kann also nicht erfolgreich unterwandert werden.

Umgekehrt kann man jedoch sehr wohl zeigen, daß die Invasion einer monomorphen H-Population durch eine kleine Zahl von D-Mutanten immer erfolgreich ist. Der Grund ist leicht einzusehen: In einer Population von fast ausschließlich ehrlichen Individuen lautet die Strategie eines H „interagiere immer“. Der H-Typ würde seine Entscheidung selbstverständlich dann konditionieren, wenn ein perfektes Signal die mutanten D-Typen entlarven würde (wie bei Frank). Dies ist hier aber nicht verfügbar. Die erwarteten Payoffs der Typen sind daher

$$\begin{aligned} E(H) &= hx_3 + (1-h)x_1, \\ E(D) &= hx_4 + (1-h)x_2. \end{aligned} \quad (3-15)$$

Da  $x_4 > x_3 > x_2 > x_1$  ist  $E(D) > E(H)$ , also wird der Populationsanteil der D-Typen zunehmen.

Was läßt sich nun für den inneren Bereich zwischen den beiden monomorphen Populationszuständen sagen? Ohne weitere Annahmen sind keine eindeutigen Ergebnisse ableitbar. Frank hatte Dichtefunktionen spezifiziert und dadurch die Eindeutigkeit und Stabilität eines inneren Gleichgewichtes gesichert. Ohne eine solche, letztlich arbiträre Maßnahme ist prinzipiell alles möglich: Die erwarteten Payoffs für H- und D-Typen sind allgemein gegeben als

$$E(H) = (1 - P_H)x_2 + P_H[\lambda x_3 + (1 - \lambda)x_1], \quad (3-16)$$

$$E(D) = (1 - P_D)x_2 + P_D[\lambda x_4 + (1 - \lambda)x_{21}], \quad (3-17)$$

wobei

$$P_H = \int_{S^*}^U f_H(S) dS, \quad (3-18)$$

$$P_D = \int_{S^*}^U f_D(S) dS, \quad (3-19)$$

$$\lambda = \frac{hP_H}{hP_H + (1-h)P_D}. \quad (3-20)$$

(3-16) und (3-17) sind beide in  $h$  monoton wachsende Funktionen. Es läßt sich ferner zeigen, daß entweder beide Funktionen simultan konvex oder konkav sein müssen. Dies läßt aber Raum für eine große Zahl möglicher Konstellationen: (3-16) und (3-17) weisen keinen Schnittpunkt auf, der Schnittpunkt ist eindeutig oder es existieren multiple innere Gleichgewichte. Sofern ein oder mehrere innere Gleichgewichte existieren können diese lokal stabil oder aber instabil sein. Die von der Frankschen Analyse suggerierte Existenz eines eindeutigen und global stabilen inneren Gleichgewichts ist also keineswegs zwingend. Sie hängt vielmehr entscheidend von zusätzlichen Annahmen ab, im Frankschen Fall betreffen diese die Dichtefunktionen. Warum selbige aber eine so höchst spezielle Form wie die von Frank unterstellte annehmen sollten, wird nicht näher erläutert. Es zeigt sich einmal mehr, daß die Frankschen Ergebnisse von einer ad hoc Annahme abhängen.

Man kann darüberhinaus zeigen, daß die Payofffunktionen keinen Schnittpunkt aufweisen, falls der Ausbeutungspayoff  $x_4$  hinreichend hoch ist. Dies ist Gegenstand des folgenden Satzes.

**SATZ: (Harrington 1989)**

$\exists$  endliches  $\bar{x}$ , so daß gilt :  $\forall x_4 > \bar{x} \Rightarrow E(D) > E(H) \forall h \in (\underline{h}, 1]$ .

**Beweis:** siehe Harrington (1989), S. 591.

Unter den Bedingungen obigen Satzes existiert also kein inneres Gleichgewicht. Ausgehend von einer monomorphen H-Population gilt vielmehr, daß D-Mutanten diese erfolgreich invadieren können und auch danach im gesamten Intervall  $(\underline{h}, 1]$  den höheren Payoff erzielen. Also sinkt der Populationsanteil der H-Typen  $h$  solange bis  $h = \underline{h}$ . Beide Typen sind dann nicht mehr unterscheidbar in ihren Payoffs. Wenngleich also eine kleine Zahl von H-Typen übrig bleiben mag, so gilt dennoch, daß in allen Populationszuständen  $h \in [0, \underline{h})$  keine Kooperation beobachtet wird. H-Typen gehen nämlich keine joint ventures ein. Es existiert also kein evolutionär stabiler Zustand, in dem kooperatives Verhalten zu beobachten wäre.



Zusammenfassend läßt sich also festhalten: Gibt man die ad hoc Annahme der Existenz von Bereichen eines perfekten Signals auf, so stürzen die Resultate des Frank-Modells entweder zusammen oder müssen doch erheblich relativiert werden. Erstens ist festzuhalten, daß eine Population von D-Typen nunmehr durch die Mutation eines marginalen Populationsanteils zu H-Typen von diesen nicht mehr unterwandert werden kann. Kooperation breitet sich nicht aus. Zweitens ist festzuhalten, daß eine Population von H-Typen lokal immer erfolgreich invadiert werden kann. Und drittens kann diese Population auch global erfolgreich invadiert werden, wenn nur der Ausbeutungspayoff hinreichend hoch ist. Insgesamt also zeichnet das Frank-Modell für die Chancen des Überlebens von Kooperation ein deutlich zu optimistisches Bild.

### 3.3.2 Weitere Kritikpunkte

Harringtons Kritik bezog sich auf die Annahme perfekter Signale. Die Darstellung und Diskussion des Frank-Modells sollte jedoch noch weitere kritikwürdige Eigenschaften der Theorie zutage gefördert haben. Die Akteure im Frank-Modell führen ein rationales Erwartungswertkalkül durch: Sie wägen die erwarteten Gewinne eines joint ventures mit den Erträgen der work alone option ab. Es muß entschieden werden, ob mit dem Träger eines bestimmten Signals eine Interaktion eingegangen wird. Die Bedingung hierfür war

$$E(X_{ij}|S_j) = x_3 Pr\{H_j|S_j\} + x_1 [1 - Pr\{H_j|S_j\}] \geq x_2,$$

$$\text{wobei } Pr\{H_j|S_j\} = \frac{hf_H(S_j)}{hf_H(S_j) + (1-h)f_D(S_j)}.$$

Was also offensichtlich erforderlich ist, um die Entscheidung treffen zu können, ist die Kenntnis einerseits der Dichtefunktionen, andererseits der Populationszusammensetzung  $h$ . Das Frank-Modell unterstellt also, daß die Spieler in zweifachem Sinne über globale Information verfügen: globale Information über die Verteilung des Signals und die globale Zusammensetzung der Population.

### 3.3.3 Die experimentelle Perspektive: Sind Signale wichtig ?

#### 3.3.3.1. Zuckerman et al. (1981)

Franks Theorie basiert maßgeblich auf der Annahme, daß Signale Information über den Verhaltenstyp von Spielern offenbaren. Neben der, wie ausführlich gezeigt, problematischen Annahme, daß solche Signale bereichsweise perfekt sein

müssen, stellt sich die Frage, ob sie – bzw. ihre Wirkung – sich empirisch nachweisen lassen. Es existiert eine breite psychologische Literatur, die sich dieser Fragestellung in zahlreichen Variationen angenommen hat. Dabei scheint relativ unstrittig zu sein, daß Signale im Frankschen Sinne existieren. Beispielsweise läßt sich sehr wohl nachweisen, daß Lügen von anderen körperlichen Reaktionen begleitet wird als das Berichten der Wahrheit (vgl. Zuckerman et al. 1981, S. 12). Der Lügner gerät in einen Zustand emotionalen „Aufruhrs“, und diese Änderung des emotionalen Zustands könnte prinzipiell als Signal für einen außenstehenden Beobachter fungieren. Die interessierende Frage im Kontext der Frankschen Theorie muß also wie folgt präzisiert werden: Sind außenstehende Beobachter in der Lage, derartige Signale wahrzunehmen und daher statistisch signifikant besser als der Zufall in der Bestimmung der verhaltensmäßigen Disposition einer Versuchsperson? Können Menschen also, beispielsweise, Lügner von Nichtlügnern mit einer hinreichenden Zuverlässigkeit unterscheiden?

Zuckerman et al. (1981) referieren die Ergebnisse von insgesamt 35 einschlägigen Untersuchungen. Die Resultate wurden von ihnen differenziert hinsichtlich der untersuchten „Signalkanäle“. Vier „channels“ werden unterschieden: (i) der facial channel („Gesichtskanal“), (ii) der body channel („Körperkanal“), (iii) der speech channel („Sprachkanal“) und (iv) der tone-of-voice channel („Klangkanal“). (i) umfaßt alle Signale, die vom Körper oberhalb des Halses ausgesandt werden, (ii) alle übrigen Körpersignale. Speech meint die Sprache in ihrer Gesamtheit von Inhalt und Klang, während (iv) den Klang der Sprache ohne Inhalt meint. Die Ergebnisse in Tabelle 4 sind in Einheiten der Standardabweichung angegeben: Ein positiver Wert indiziert, daß die Vpn in der Entdeckung von Lügen überzufällig gut waren, negative Werte bedeuten, daß die Beobachter schlechter als der Zufall waren. In Klammern ist die Anzahl der zugrundeliegenden Studien angegeben.

Visual Cues					
	Face		No Face		Means
	Body	No Body	Body	No Body	
Speech	1.00 (21)	0.99 (9)	1.49 (3)	1.09 (12)	1.14
No Speech	0.35 (6)	0.05 (7)	0.43 (4)	0.00	0.21
Means	0.68	0.52	0.96	0.54	
	0.6		0.75		
Transcript only: 0.70 (6)					
Tone only: 0.20 (4)					

Tabelle 4: “Accuracy of Detecting Deception”, Quelle: Zuckerman et al (1981), S. 26

Überraschend ist das schlechte Abschneiden des face channels. Das Gesicht allein ist offensichtlich keine sehr ergiebige Quelle von Signalen, die zuverlässig Auskunft darüber gibt, ob jemand täuscht/lügt oder nicht. Der Wert von  $d=0.05$  ist nicht signifikant von  $d=0.00$  verschieden. Mehr noch, der face channel scheint darüberhinaus sogar irreführende Information zu generieren: Im Durchschnitt waren Vpn in No-Face-conditions erfolgreicher ( $d=0.75$ ) als in Face-conditions ( $d=0.6$ ). Frank (1988) hingegen betont nachdrücklich die Rolle insbesondere des Gesichtes bei der Generierung von Informationen über die Disposition eines Menschen (ebd., S.120 ff). Das Gesicht scheint aus physiologischer Sicht in der Tat dazu prädestiniert zu sein, in Frank's Theorie eine bedeutende Rolle zu spielen. Wie Frank ausführlich darlegt (ebd., S. 121 ff), finden sich im menschlichen Gesicht tatsächlich zahlreiche Muskelgruppen, die für die meisten Menschen willentlicher Kontrolle entzogen sind. Der Gesichtsausdruck ist damit nur partiell kontrollierbar und daher potentiell eine Quelle für Signale, die den wahren emotionalen Zustand eines Individuums offenbaren. Die empirische Evidenz jedoch zeigt klar, daß, wenn solche Signale existieren, sie von beobachtenden Versuchspersonen offenbar nicht wahrgenommen werden, jedenfalls nicht so, daß daraus eine überzufällig gute Trefferquote resultiert.

Franks Theorie betont besonders die Rolle nonverbaler Signale. Genannt wurde bereits der Gesichtsausdruck. Körperhaltung, Pupillenerweiterung, Transpiration und Speichelfluß sind weitere Indikatoren, die von Frank explizit diskutiert werden. Vor diesem Hintergrund ist ein weiteres Ergebnis außerordentlich überraschend. Wir zitieren Zuckerman et al. (1981, S.27):

Tabelle 4 „also reveals that tone only and transcript only both provide cues to deception. For both channels, accuracy of deception was above chance, although **accuracy of transcript was much higher** (Herv.d.A.). Furthermore, the difference between accuracy of tone (0.20) and that of speech (1.09) can be conceptualized as the amount of information added by paralanguage cues. The surprising finding, of course, is the power (i.e. the accuracy) of the word, either written or spoken. The assumption that nonverbal channels are more important in the communication of deception than the verbal cues is simply not true [...].“

Besonders erwähnenswert ist die Tatsache, daß bereits der nackte Inhalt, also das Transskript einer Kommunikation ausreicht, um beachtliche Trefferquoten oberhalb derer anderer isolierter Kanäle zu erzielen (transcript only:  $d=0.70$ , face only:  $d=0.05$ , body only:  $d=0.43$ ).

Allgemein ist zu sagen, daß – abgesehen von der face-only-condition – alle  $d$ -Werte signifikant von Null verschieden sind. D.h., Beobachter waren in allen anderen Versuchsanordnungen in der Einschätzung von Täuschung/Lüge versus Wahrheit signifikant besser als der Zufall. Diese Resultate sind allerdings mit Vorsicht zu interpretieren, wie Zuckerman et al. (ebd., S.26) selbst zu bedenken geben:

*„It should be noted that, although all  $d$ 's in the table are positive and some are substantial, in reality even a big effect may correspond to only a few percentage points above chance accuracy. In fact, the examination of actual percentage accuracy indicates that most of the results fall in the 0.45-0.60 range with a chance level of 0.50 [...]“.*

Zusammenfassend läßt sich feststellen:

- i) Der von Frank (1988) besonders hervorgehobene face channel als Quelle fälschungssicherer Signale erweist sich überraschenderweise als nicht informativ. Das Gesicht generiert keine oder sogar irreführende Information über die Disposition des Gegenübers.
- ii) Die Betonung der Rolle nonverbaler Signale bei der Typoffenbarung läßt sich im Licht der empirischen Ergebnisse nicht aufrecht erhalten. Der bloße Inhalt von sprachlicher Kommunikation läßt bereits Aufschluß zu.
- iii) Angesichts der doch eher bescheidenen Trefferquoten zwischen 45 und 60% scheint es angebracht zu sein, von einer nur schwach ausgeprägten Fähigkeit zur Erkennung von Zuständen/Dispositionen anderer zu sprechen.

### 3.3.3.2. Ökonomische Untersuchungen

Die in der Psychologie angewandte experimentelle Methode kann unter einer ganzen Reihe von Gesichtspunkten kritisiert werden (vgl. Weimann (1995), Hey (1991)). In unserem speziellen Kontext spielen vor allem zwei Aspekte eine zentrale Rolle. Zum einen muß man sich die typische Versuchsanordnung einer psychologischen lie-detection oder deception-detection-study verdeutlichen. Üblicherweise werden Versuchspersonen dort vom Experimentator angewiesen, zu lügen oder die Wahrheit zu sagen. Beispielsweise sollten Vpn schlecht über jemanden reden, den sie nachweislich mochten und umgekehrt. Das Problem einer solchen Versuchsanordnung besteht darin, daß aus der Anweisung durch den Ex

perimentator eine moralische Entlastung der Vpn resultieren kann. Jemand der angewiesen wird zu lügen, wird dieses Verhalten anders bewerten und empfinden als jemand, der die freie Wahl hat. Die Verantwortung für die Lüge geht im ersten Fall auf den Versuchsleiter über, was im Ergebnis dazu führen kann, daß die Vpn lügen nicht mehr als verwerflich empfindet und deshalb bestimmte emotionale Zustände und damit korrelierte körperliche Symptome nicht entwickelt. Gegeben diese Überlegungen sind richtig, dann führt die skizzierte Versuchsanordnung zu einer sehr konservativen Schätzung der lie-detection-capability.

Der zweite Aspekt betrifft die Anreizstruktur psychologischer Experimente. Üblicherweise wird dort nicht mit monetären Anreizen gearbeitet, was aus ökonomischer Sicht natürlich höchst problematisch ist. Die Kritik an dieser Praxis der psychologischen Forschung ist so alt wie die Disziplin der experimentellen Ökonomik. Im Kontext der lie-detection studies wird die Berechtigung der ökonomischen Kritik besonders deutlich: Wenn lügen sich lohnt, weil die Höhe der Auszahlung einer Vpn davon abhängt, wie gut sie lügt, dann ist die Vermutung nahelegend, daß Vpn versuchen werden, besonders gut zu lügen. In welche Richtung dieser Effekt wirkt, ist a priori unklar: Sind Vpn bei diesem Bestreben erfolgreich, so würde dies die gemessene lie-detection-capability nach unten korrigieren. Andererseits: Wenn dieses Bestreben zu einem besonders kontrollierten Verhalten führt, so kann dieses selbst zu einem Signal für eine Lüge werden und den Lügner, entgegen aller Absicht, verraten.

Die Richtung des resultierenden Gesamteffekts ist ebenfalls a priori nicht sicher. Klar sein sollte nach der obigen Diskussion allerdings, daß sowohl moralische Entlastung als auch monetäre Anreize die Messung der lie-detection beeinflussen können. Die psychologischen Studien kontrollieren diese Effekte nicht. Ihre Resultate sind daher, zunächst, mit Vorsicht zu interpretieren.

Angesichts der Popularität der Frankschen Theorie ist es verwunderlich, daß es kaum experimentelle Untersuchungen von Ökonomen gibt, die sich ihrer Überprüfung annehmen. Vielleicht ist es die ausgesprochene Plausibilität der Theorie, die dazu geführt hat, daß kaum jemand ihre experimentelle Überprüfung für nötig befunden hat. Wir werden auf Gründe für diese Plausibilität später noch zu sprechen kommen.

### *Frank et al (1993)*

Frank selbst hat in einer Arbeit mit Gilovich und Regan (1993) seine Theorie einer experimentellen Überprüfung unterzogen. Bevor die Vpn ein one-shot Prisoner's Dilemma spielten, hatten sie Gelegenheit, sich kennen zu lernen und ohne jegliche Restriktionen zu kommunizieren. Im Anschluß an die Kommunikations

phase des Experiments sollten die Vpn das Verhalten ihrer Kommunikationspartner im Gefangenen Dilemma prognostizieren. Das Hauptergebnis der Studie ist, daß die Vpn überzufällig gut das Verhalten ihrer Mitspieler vorhersagten. Frank wertet dieses Resultat als Beleg für die Richtigkeit seiner Theorie. Zweifel an dieser Interpretation scheinen jedoch angebracht zu sein, leidet die Studie doch an zwei methodischen Mängeln. Zum einen wurde offensichtlich die Kommunikation nicht dokumentiert, zumindest enthält das paper keine entsprechenden Hinweise. Es ist also nicht klar, ob die Kommunikation nicht explizite oder implizite Absichtserklärungen bezüglich des Verhaltens im GD enthielt. Ockenfels und Selten (1998, S.1) stellen hierzu fest:

*„It is possible that some subjects openly or implicitly declared that they did not want to cooperate and therefore were expected not to do so. Therefore the experiment does not clarify the question what was deducible from verbal statements and what was due to the recognition of involuntary signals.“*

Der zweite methodische Mangel der Studie betrifft die Anreizstruktur. Um es für die Teilnehmer des Experiments unmöglich zu machen, aus den Auszahlungen auf das Verhalten von Mitspielern im Experiment schließen zu können, wurde zur Endauszahlung eine Zufallskomponente addiert. Das Verhalten im Experiment konnte daher die Endauszahlung nur noch schwach beeinflussen. Insbesondere bestanden aber, wie in den psychologischen Studien auch, keine monetären Anreize bezüglich des Vorhersageverhaltens.

#### *Ockenfels und Selten (1998)*

Ockenfels und Selten (1998) versuchen deshalb, die Frank-Hypothese in einem explizit ökonomischen Kontext mit wirksamen monetären Anreizen zu testen. Der ökonomische Kontext ist eine bargaining Situation, in der zwei Vpn über die Aufteilung eines Kuchens verhandeln. Dieser Kuchen hat die Größe von 30 DM. Es war sichergestellt, daß die Verhandler sich nicht kannten. Vor dem eigentlichen Experiment wurde zufällig entschieden, ob eine Vpn Kosten oder keine Kosten hat. Beide Verhandler würfelten. Im Falle einer 1, 2 oder 3 hatte die betreffende Vpn Kosten in Höhe von 12 DM, im Falle einer 4, 5 oder 6 hatte sie keine Kosten. Im Falle von Kosten wurden die 12 DM von der Endauszahlung (d.h. dem erzielten Verhandlungsergebnis, also der Aufteilung der 30 DM) abgezogen. Wurde keine Einigung über die Aufteilung des Kuchens erzielt, so gingen beide Verhandler leer aus. Die Verhandlungsdauer betrug jeweils, also pro Paar, maximal zehn Minuten. Vpn konnten die Verhandlung jedoch auch vorzeitig be-

enden. Die Verhandlungen wurden in einem Hörsaal durchgeführt, in dem sich außerdem eine Anzahl von Zuschauern befand. Das Experiment bestand insgesamt aus zwei Durchgängen. In jedem Durchgang verhandelten 24 Vpn in 12 Paaren über die Aufteilung der 30 DM. In Durchgang 1 nahmen 42 Vpn als Zuschauer der Verhandlungen teil, in Durchgang 2 46 Vpn. Die Verhandlungspartner füllten beide separat ein Formblatt aus, in dem sie die nach ihrer Ansicht erzielte Aufteilung festhielten sowie einschätzen sollten, ob ihr Partner Kosten hatte oder nicht. Die Aufgabe der Zuschauer war ebenfalls, die Kostensituation beider Verhandler einzuschätzen. Die monetären Anreize für die Zuschauer waren dabei so gestaltet, daß es sich lohnte, besonders gute Prognosen abzugeben.

Hauptfragestellung war, ob es Verhandlern bzw. Zuschauern gelingen würde, die Kostensituation überzufällig gut einzuschätzen. Dies war der Fall. Zuschauern gelang die zutreffende Einschätzung in 55.2% aller Fälle. Die Verhandler selbst schnitten mit 52.1% etwas schlechter ab. Die Resultate bewegen sich allerdings auf einem recht bescheidenen Niveau und entsprechen von der Größenordnung her den Trefferquoten, die Zuckerman et al. (1981) berichten.

Des weiteren zeigt sich in der weiteren Auswertung (Ockenfels und Selten 1998, S. 9), daß die Wahrscheinlichkeit einer zutreffenden Einschätzung der Kostensituation von Verhandler zu Verhandler sehr stark variiert:

*„Our findings suggest that the probabilities of a correct guess by an onlooker are different for different bargainers. It appears that the cost situation is easy to assess in some cases and very difficult to be correctly guessed in some other cases.“*  
(*ibd.*, S.9).

Natürlich erhebt sich unverzüglich die Frage nach der Ursache der beobachteten Varianz. A priori ist nicht auszuschließen, daß diese lediglich den Reflex einer entsprechenden Verteilung eines mehr oder weniger zuverlässigen Signals über die Verhandler darstellt. Wie sich jedoch zeigt, sind es objektive Charakteristika des Verhandlungsprozesses, die die Ergebnisse treiben. Das Hauptgewicht kommt zwei Faktoren zu: (i) dem letztendlich erzielten Verhandlungsergebnis und (ii) der Verhandlungsdauer. Ad (i): Fünf Verhandlungen endeten mit einer asymmetrischen Verteilung der 30 DM. Die Zuschauer tendierten in diesen Fällen dazu, einen höheren Bruttopyoff als Indikator für das Vorhandensein von Kosten zu interpretieren, während ein niedrigerer Bruttopyoff als Indikator für die Abwesenheit von Kosten angesehen wurde. Letzteres erweist sich in der Tat als guter Indikator: In allen fünf Verhandlungen mit asymmetrischem Ergebnis hatte der Verhandler mit dem niedrigeren Bruttopyoff tatsächlich keine Kosten. Da 93.6 % aller Zuschauer dies vermuteten, verbessert dies die durchschnittliche Treffer

quote beträchtlich. Allerdings ist die erste Vermutung, daß ein höherer Payoff Kosten signalisiert, durch die Daten nicht gerechtfertigt. In den betreffenden 5 Verhandlungen vermuten 75 % der Zuschauer diesen Zusammenhang, während tatsächlich nur eine Vpn Kosten hatte. ad (ii): In fünf weiteren Verhandlungen war die Verhandlungsdauer extrem kurz, nämlich weniger als 30 Sekunden. In allen Fällen wurde eine Gleichaufteilung vereinbart. Die sehr kurze Verhandlungszeit kann als Indikator für die Abwesenheit von Kosten interpretiert worden sein. In der Tat ist die Verhandlungszeit in diesen Fällen ein sehr guter Indikator: Neun der Zehn Verhandler hatten keine Kosten. Die durchschnittliche Trefferquote der Zuschauer betrug, bezogen auf die 5 Verhandlungen bis zu einer halben Minute, beachtliche 71.1 %. Demgegenüber war die Trefferquote für die längeren Verhandlungen erheblich niedriger, nämlich nur 51 %.

Diese Beobachtungen legen eine Bereinigung der Daten nahe. Ockenfels/Selten bestimmen eine korrigierte Trefferquote, indem zum einen die 10 Vpn aus den extrem kurzen Verhandlungen aus der Grundgesamtheit eliminiert werden, zum anderen jene 5 Vpn, die in den Verhandlungen mit asymmetrischem Ergebnis den niedrigeren Bruttopyoff realisierten. Der Effekt ist spürbar, die so gewonnene durchschnittliche Trefferquote der Zuschauer beträgt nur noch 44.5 %. Unter Berücksichtigung des 99%-Konfidenzintervalls (untere Grenze: 41.07%, obere Grenze: 47.99%) kann man schließen, daß

*„without the criteria of extremely short bargaining time and lower gross payoff in an asymmetric bargaining result, the capability of the onlookers to make correct judgements about the true cost situations is worse than random. The upper limit of the confidence interval is clearly below 50 %.“* (Ockenfels und Selten 1998, S. 12).

Man kann allerdings die Frage erheben, ob damit tatsächlich bereits die Franksche Hypothese als widerlegt gelten kann – wenigstens für bargaining Situationen. Im weitesten Sinne besagt die Frank-Hypothese, daß sich der wahre emotionale Zustand eines Menschen unwillentlich anderen mitteilt, weil bestimmte körperliche, nonverbale und zumindest partiell fälschungssichere Signale existieren, die vom Sender nur unvollkommen kontrolliert werden können. Die Kenntnis des wahren emotionalen Zustands erlaubt dann wiederum eine mehr oder weniger zuverlässige Prognose der Verhaltensdisposition. Es ist, wenigstens für den Autor nicht erkennbar, inwiefern diese Hypothese durch die Studie von Ockenfels und Selten geprüft worden sein sollte. In einer engeren Interpretation, etwa im Sinne des formalen Modells von Frank (1987) lautet die Frank-Hypothese: Man kann nicht beliebig, also willentlich und perfekt, lügen. Der Lügner entlarvt sich durch be



stimmte autonome Reaktionen wie Erröten, Nervosität etc. In diesem Sinne können die deception-detection studies, die in Zuckerman et al. (1981) referiert werden, durchaus als ein Test der Frank-Hypothese gesehen werden. Eine Lüge kann sehr wohl dazu geeignet sein, einen bestimmten, meist wohl unangenehmen Gefühlszustand zu provozieren, der dann möglicherweise an bestimmten Symptomen von Außenstehenden erkannt werden kann. Warum aber sollte eine bestimmte Kostensituation in einem Verhandlungsexperiment irgendeine emotionale Reaktion hervorrufen? Glücklicherweise testeten Ockenfels und Selten die Frank-Hypothese auch in dem o.g. engeren Sinne, wenngleich auch dieser Aspekt der Analyse eine völlig untergeordnete Rolle spielt. Erfasst wurde nämlich auch, ob ein Verhandler während der Verhandlung behauptete, Kosten zu haben. Da die wahre Kostensituation der Vpn bekannt war, war ex post bekannt, wer während der Verhandlung gelogen hatte und wer nicht. 12 Verhandler behaupteten wahrheitsgemäß Kosten zu haben, während 12 weitere Verhandler unrichtigerweise diese Behauptung erhoben. Die Ermittlung der Zuschauer-Trefferquoten zeichnet tatsächlich kein allzu günstiges Bild für die Frank-Hypothese. Die Erkennungsrate bei wahrheitsgemäßen Behauptungen lag bei lediglich 47 % (die Trefferquote bleibt unverändert, wenn eine Bereinigung der Daten nach objektiven Charakteristika der Verhandlungen vorgenommen wird). Völlig überraschenderweise vermuten die Zuschauer jedoch zu 55% tatsächlich Kosten bei denen, die gelogen haben. Mit anderen Worten, den Lügern wurde durchschnittlich mehr geglaubt als denjenigen, die wahrheitsgemäße Behauptungen erhoben. Dieses Resultat befindet sich allerdings tatsächlich nicht im Einklang mit der Frank-Hypothese.

### *Brosig (2000)*

Ein etwas optimistischeres Bild für die Frank-Hypothese zeichnet die Untersuchung von Brosig (2000). Der Versuchsaufbau ähnelt stark dem von Frank et al. (1993). Vpn konnten unrestringiert kommunizieren, bevor sie ein 2-Personen-GD spielten. Es war sichergestellt, daß die Vpn sich nicht kannten. Es war ferner gewährleistet, daß aus den Auszahlungen nicht auf die im GD getroffenen Entscheidungen geschlossen werden konnte. Die Studie unterscheidet sich von Frank et al. im wesentlichen in zwei Punkten: Zum einen wird ein Pretest zur Typenermittlung nach dem Value-orientation-circle (s.o.) durchgeführt. Zum anderen wurde die Kommunikation dokumentiert (per Videoaufzeichnung). Dies ermöglichte es, explizite Absichtserklärungen, im GD zu defektieren, aus dem Sample zu entfernen. Die wesentlichen Resultate lassen sich wie folgt zusammenfassen. Zunächst finden sich auch bei Brosig die üblichen zwei großen Typengruppen der „Individualisten“ (59.9 %) sowie der „Kooperativen“ (37.3 %). Sodann stellt man fest,

daß die Trefferquote bezüglich des Verhaltens im GD mit 67 % leicht oberhalb der erwarteten Zufallsrate von 59 %<sup>13</sup> liegt. Die Vpn waren durchschnittlich also überzufällig gut in der Prognose des Verhaltens ihres Mitspielers im GD. Allerdings ist anzumerken, daß sich das Ausmaß der Typerkennung wiederum auf jenem bescheidenen Niveau bewegt, das bereits Zuckerman et al. berichten. Interessant ist jedoch folgender Befund: Während die Fähigkeit zur Typidentifizierung für die individualistischen Vpn wenig ausgeprägt war (die tatsächliche Trefferquote betrug 63 % gegenüber erwarteten 60% bei Zufallsverhalten), so zeigte sie sich bei den kooperativen Vpn umso deutlicher: Die tatsächliche Trefferquote lag hier bei 70 % gegenüber erwarteten 59 %.

*Gilovich et al. (1998)*

Angesichts der doch eher schwachen experimentellen Evidenz verwundert die ausgesprochene Popularität der Frankschen Theorie. Eine mögliche Ursache für diese Diskrepanz haben Gilovich et al. (1998) aufgezeigt. Der Tenor ihrer Arbeit lautet, daß Menschen die Wahrscheinlichkeit, als Lügner enttarnt zu werden, systematisch und dramatisch überschätzen. Die Autoren nennen diesen Effekt „the illusion of transparency“. In einer typischen lie-detection-Anordnung wurde erhoben, mit welcher Wahrscheinlichkeit Lügner damit rechneten, von anderen Beteiligten des Experimentes als solche enttarnt zu werden. Es zeigte sich, daß die Lügner mit einer Entdeckungswahrscheinlichkeit von 48,8% rechneten. Die tatsächliche Entdeckungsrate betrug jedoch nur 25,6 %. In der gewählten Versuchsanordnung entsprach eine Entdeckungshäufigkeit von 25 % gerade der Zufallsrate. Interessanterweise zeigt sich also außerdem, daß Lügner nicht überzufällig gut entdeckt wurden. Diese Beobachtung wird von Gilovich et al. in einer ganzen Reihe weiterer Anordnungen erhärtet (vgl. ebd., S.335 ff). Es zeigt sich also einmal mehr, daß es offensichtlich nicht so einfach ist „to read one’s emotional state“, wie die Theorie von Frank suggeriert. Andererseits aber glauben wir, daß Außenstehende unsere emotionale Verfassung erkennen können müßten. Dies ist subjektiv verständlich, sind doch viele der von Frank diskutierten Signale mit erheblichem körperlichen und/oder emotionalen Aufruhr verbunden. Wir tendieren dazu zu meinen, daß dieser Aufruhr anderen nicht verborgen bleiben kann. Diese Vermutung aber erweist sich als falsch – eben als eine „Illusion of Transparency“, wie Gilovich et al. gezeigt haben.

---

<sup>13</sup> Zur Berechnung der Zufallsrate vgl. Brosig (2000), S. 6, Fußnote 13.

### 3.4 Robson (1990)

Ein weiteres, bekanntes signaltheoretisches Modell stammt von Robson. Die Arbeit beschäftigt sich allgemein mit der Frage, wie Zustände, in denen eine Population eine ineffiziente, jedoch evolutionär stabile Strategie spielt, verlassen werden können. Betrachtet sei folgendes GD:

	coop	defect
coop	3	1
defect	4	2

Tab. 5: GD

Es sei nun angenommen, daß ein mutanter Typ in einer formal unendlichen Population auftaucht. Diese Mutante zeichnet sich dadurch aus, daß sie ein (kostenloses!) Signal trägt, welches von anderen Mutanten erkannt wird, jedoch nicht von den Mitgliedern der alten Population. Es sei weiterhin angenommen, daß die Mutante gegen die alte Population „defect“ spielt, jedoch gegenüber anderen Mutanten kooperiert (daher der Name „secret handshake model“). Die Payoffs des 3x3 Spiels sind also gegeben durch

	coop	defect	mutant
coop	3	1	1
defect	4	2	2
mutant	4	2	3

Tab. 6 GD mit Mutante

Wie man sich leicht klar macht, ist eine marginale Anzahl von Mutanten der oben beschriebenen Art immer in der Lage, die alte Population erfolgreich zu invadieren. Einzige evolutionär stabile Strategie des 3x3-Spiels ist also „mutant“. Soweit scheint die Einführung der Mutante einen eleganten Ausweg aus dem Kooperationsproblem im GD zu weisen – noch dazu einen Weg, der zu einem effizienten Zustand führt! Frank (1987, 1988) hatte bereits auf die fatalen Konsequenzen perfekter Mimikry hingewiesen. Robson stellt selbst fest:

*„The difficulty is just that such an ESS is a „sitting duck“ for the introduction of still another mutant, one which would prey on the first mutant.“* (Robson (1990), S.387)

Die Payoffmatrix des entsprechenden 4x4 Spiels stellt sich dann wie folgt dar:

	coop	defect	Mutant1	Mutant2
coop	3	1	1	1
defect	4	2	2	2
Mutant1	4	2	3	1
Mutant2	4	2	4	2

Tab. 7 GD mit zwei Mutanten

Man macht sich leicht klar, daß nunmehr keine ESS mehr existieren. Um Aussagen über die langfristige Entwicklung der Population treffen zu können, muß also ein dynamisches System spezifiziert und analysiert werden. Legt man eine Entwicklung der Populationsanteile gemäß der Replikatorodynamik (vgl. Kapitel 5 dieser Arbeit) zugrunde, so stellt man fest, daß alle generischen Lösungen gegen Populationszustände konvergieren, in denen einfache Defektoren und die zweite Mutante überleben. Kooperative Strategien sterben also aus.

### 3.5 Ausblick

Die diskutierten signaltheoretischen Arbeiten weisen beide erhebliche Defizite auf. Weder das Frank-Modell noch das secret-handshake-Modell sind letztlich in der Lage, das stilisierte Faktum der stabilen Ko-Existenz verschiedener Verhaltensdispositionen zu erklären. Man kann versuchen, dieses Defizit im Rahmen eines signaltheoretischen Ansatzes zu beheben. Ein entsprechendes Modell haben Amann und Yang (1998) vorgestellt. Kapitel 3.3 hatte jedoch auch von empirischer Warte her Zweifel aufgeworfen, wie relevant Signale tatsächlich für die Entstehung kooperativen Verhaltens sind. Vor dem Hintergrund der doch nur sehr schwach ausgeprägten Evidenz für den Signalansatz, möchte die vorliegende Arbeit einen konzeptionell alternativen Weg beschreiten. Wir werden im folgenden zeigen, daß Erfahrungslernen das Gewünschte leistet. Kapitel 4 entwickelt ein Modell des Erfahrungslernens, das in random-matching games implementierbar ist. Kapitel 5 stellt das Basis-Modell der Arbeit vor, ein modifiziertes 3x3 GD. Eine lernfähige Mutante wird in das GD eingeführt. Diese lernt andere Akteure durch Interaktion mit ihnen kennen und speichert die erworbene Information in einem Gedächtnis ab. Unter der Annahme, daß Lernen kostspielig ist, kann dann gezeigt werden, daß ein global stabiler, vollständig gemischter Populationszustand existiert. Kapitel 6 und 7 diskutieren weitergehende Modellerweiterungen.

#### 4. Ein Modell des Erfahrungslernens für random matching games

Betrachtet sei eine Population der endlichen Größe  $n+1$ . Die Population bestehe anfänglich aus zwei Verhaltenstypen in beliebiger Zusammensetzung, nämlich Hawks (H) und Doves (D). Bezeichne  $x_H$  den Populationsanteil der Falken. Falken sind defektive Automaten, Tauben kooperative. Wir nehmen nun an, daß ein mutanter Verhaltenstyp in dieser Population auftaucht. Sei diese Mutante als knowledgeable type, kurz K-Typ bezeichnet. Sie zeichnet sich dadurch aus, daß sie lernfähig ist. Wir nehmen an, daß es eine endliche Anzahl von Zeitpunkten  $t=0, \dots, T$  gibt, in denen die Mitglieder der Population miteinander interagieren. Präziser gesagt wird angenommen, daß die Population in jedem der  $T+1$  Zeitpunkte zufällig und paarweise gematcht wird. Jedes so gematchte Paar von Spielern spielt das one-shot GD und wird danach in die Population „zurückgelegt“. Dann beginnt der matching-Vorgang von Neuem, insgesamt  $T+1$ -mal. Der K-Typ ist nun gerade dadurch ausgezeichnet, daß er über ein Gedächtnis verfügt. Dieses Gedächtnis wird im Zeitablauf mit Inhalten gefüllt. Wir nehmen an, daß der K-Typ nach jeder Interaktion das betreffende Individuum kennt und dazu in der Lage ist, genau dieses Individuum in späteren Zeitpunkten wieder zu erkennen, falls es zu einem Wiedertreffen kommt. Sobald also ein bestimmtes Individuum einmal getroffen wurde, ist dem K-Typ bekannt, zu welchem Verhaltenstypus das betreffende Individuum gehört.

Aus Gründen der Einfachheit betrachten wir zunächst den Fall einer singulären Mutante. Außer dem einen K-Typ existieren also nur Falken und Tauben in irgendeiner beliebigen Zusammensetzung  $\{x_H, 1 - x_H\}$ . Die grundlegende Verhaltensannahme bezüglich des K-Typ ist die, daß er sich bedingt kooperativ verhält. Genau gesagt verhält sich der K-Typ zunächst einmal grundsätzlich kooperativ. Trifft er jedoch einen Falken, der ihm bereits bekannt ist, so nehmen wir an, daß der K-Typ die Interaktion mit diesem Falken verweigert. Wie später noch erläutert wird, erzielen beide Spieler in diesem Fall den Wert einer outside-option. Die Populationsanteile der Typen seien während der  $T+1$  matching-Perioden konstant. Man kann  $T+1$  als die Lebensdauer einer Generation interpretieren. Populationsanteile ändern sich lediglich über mehrere Generationen hinweg, wie später präziser erläutert wird. Innerhalb einer Generation jedoch bleibt die Zusammensetzung der Population unverändert. Generationen überlappen nicht. Die Frage ist dann, ob es möglich ist, in einer präzisen Weise anzugeben, wie der K-Typ lernt, d.h. wie sich sein Wissen über die Falken im Zeitablauf ändert. Ist es also möglich, das Wachstum des Wissens des K-Typs zu beschreiben?

Ein Maß für das Wissen des K-Typs ist die Wahrscheinlichkeit, einen bekannten Falken zu treffen. Die Frage nach dem Wachstum des Wissens kann also dahin

gehend operationalisiert werden, ob es möglich ist anzugeben, wie sich die Wahrscheinlichkeit, einen bekannten Falken zu treffen, im Zeitablauf ändert. Das Problem kann technisch als ein kombinatorisches begriffen werden. Die Population ist eine Urne konstanter Größe und konstanter Zusammensetzung. Seien die Hawks durch schwarze Kugeln symbolisiert, die Doves durch weiße. Der matching-Prozeß kann einfach interpretiert werden als wiederholtes Ziehen aus der Urne mit Zurücklegen. Allerdings, und dies begründet die Nicht-Standard-Eigenschaft dieses Problems, werden schwarze Kugeln markiert, beispielsweise mit einem weißen Kreuz. Die obige Frage nach der Wahrscheinlichkeit, einen bekannten Falken zu treffen, entspricht dann in der technischen Formulierung der Frage nach der Wahrscheinlichkeit, in einem beliebigen Zeitpunkt t eine markierte schwarze Kugel zu ziehen. Die Schwierigkeit entspringt nun gerade der Tatsache, daß sich der Anteil der markierten Kugeln während des Ziehungsprozesses ändert. Er erhöht sich, um genau zu sein, aber wie ?

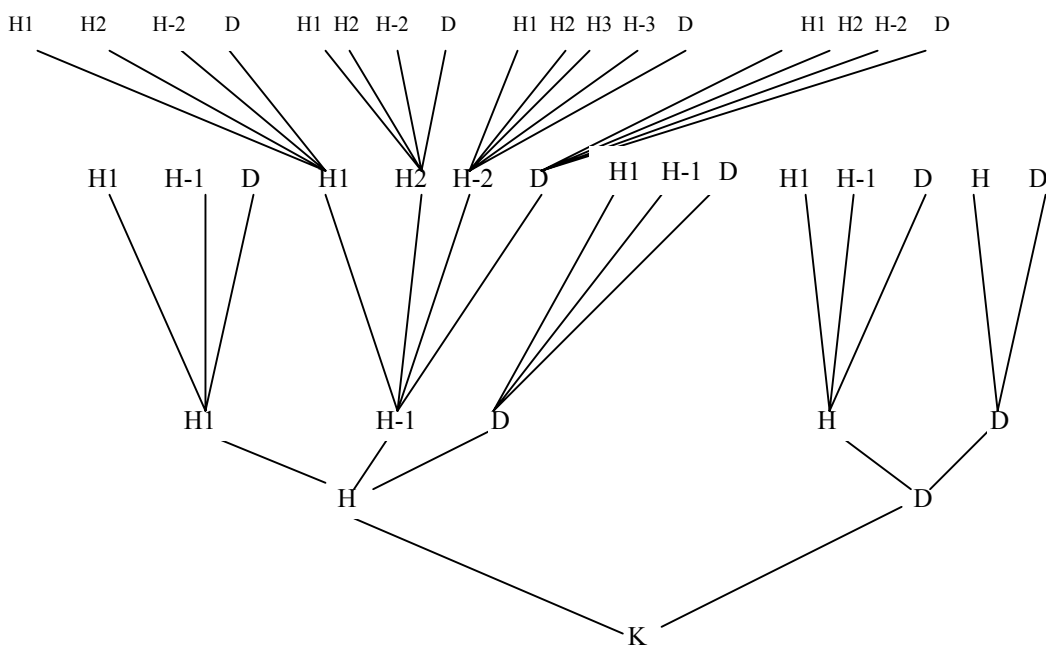


Abbildung 11: Veränderliche Zusammensetzung der Urne

Abbildung 11 verdeutlicht das Problem. Aufgrund der schnell wachsenden Komplexität des Problems zeigt die Darstellung lediglich die ersten vier Ziehungen aus der Urne und für die vierte Ziehung ist lediglich ein Zweig des „Baums“ vervollständigt. Angenommen, der K-Typ trifft im Zeitpunkt  $t=0$  einen Falken. Im Zeitpunkt  $t=1$  ist dann folgendes möglich: Er trifft eine Taube (D), er trifft den zuvor getroffenen Falken wieder (H1), oder aber er trifft einen noch unbekanntem Falken (H-1). Bereits nach einer Ziehung also besteht die Urne, aus der Perspektive der K-Mutante, nicht mehr länger nur aus zwei Elementen, sondern vielmehr aus

dreien. Der Umstand, daß der K-Typ lernt, äußert sich gerade darin, daß die Teilpopulation der Falken im Ziehungsablauf in bekannte und unbekannte Falken zerfällt, wobei sich die Anteile im Zeitablauf ändern. In Abhängigkeit von den Ereignissen in der ersten und zweiten Ziehung stellt sich die Situation in der dritten Ziehung noch komplexer dar. Gesetzt der K-Typ trifft in der zweiten Ziehung nicht den bekannten Falken wieder, sondern einen noch unbekanntem. Dann kann in der dritten Ziehung folgendes geschehen: Er trifft den bekannten Falken aus der ersten Ziehung wieder (H1), er trifft den in der zweiten Ziehung gelernten Falken wieder (H2), er trifft eine Taube (D) oder er trifft einen der auch nach zwei Ziehungen noch unbekanntem Falken (H-2) usw... Woran wir nun interessiert sind, ist genau die Wahrscheinlichkeit, in einem beliebigen Zeitpunkt ein Element aus der Menge der bekannten Falken (der markierten schwarzen Kugeln) zu ziehen. Die Lösung des Problems enthält der folgende

### SATZ 4.1

*Betrachtet sei eine Population bestehend aus strikt kooperativen (Doves) und strikt defektiven Automaten (Hawks). Bezeichne  $x_H$  den Populationsanteil der Hawks, der während des Ziehungsverlaufs konstant bleibt und sei  $n+1$  die konstante, endliche Größe der Population. Betrachtet sei ein singulärer K-Mutant. Unter pairwise random matching gilt dann: Die Wahrscheinlichkeit für die K-Mutante, in einem beliebigen Zeitpunkt  $t+1$  einen bekannten Hawk zu treffen ist gegeben als*

$$\Pr(H_K(t+1)) = x_H \left( 1 - \left( \frac{n-1}{n} \right)^t \right). \quad (4-1)$$

### Beweis:

Betrachtet sei ein beliebiger Zeitpunkt  $t$ . In  $t$  sind drei Ereignisse möglich: (i) Der K-Typ trifft einen bekannten Hawk. Sei die Wahrscheinlichkeit hierfür mit  $\Pr(H_K(t))$  bezeichnet. (ii) Er trifft einen unbekanntem Hawk. Diese Wahrscheinlichkeit ist als das Residuum  $x_H - \Pr(H_K(t))$  gegeben. (iii) Der Mutant trifft eine Taube mit Wahrscheinlichkeit  $1 - x_H$ . Ad (i): Gegeben ein bekannter Hawk wird getroffen. Klar: Die Menge der bekannten Hawks hat sich nicht erhöht. Die Wahrscheinlichkeit also, einen bekannten Hawk in  $t+1$  zu treffen, bleibt unverändert, d.h.  $\Pr(H_K(t+1) | H_K \text{ in } t) = \Pr(H_K(t))^2$ . Ad (ii): Gegeben, in  $t$  wurde ein unbekanntem Hawk getroffen, dann sind zwei Möglichkeiten zu berücksichtigen. Zum einen kann exakt dieser Hawk wiedergetroffen werden. Dieses Ereignis hat die Wahrscheinlichkeit  $1/n$ . Oder aber es wird einer der übrigen noch

unbekannten Hawks getroffen. Die Wahrscheinlichkeit, in  $t+1$  einen bekannten Hawk zu treffen erhöht sich also gerade um  $1/n$ , d.h.

$$\Pr(H_K(t)|H_U \text{ in } t) = \left( \Pr(H_K(t)) + \frac{1}{n} \right) (x_H - \Pr(H_K(t))).$$

Schließlich ad (iii): Wurde eine Dove in  $t$  getroffen, so erhöht sich die Menge der bekannten Hawks ebenfalls nicht (wie unter (i)), und die Wahrscheinlichkeit, einen bekannten Hawk zu treffen, bleibt konstant, d.h.  $\Pr(H_K(t)|D \text{ in } t) = \Pr(H_K(t))(1 - x_H)$ . Die totale Wahrscheinlichkeit, im Zeitpunkt  $t+1$  einen bekannten Falken zu treffen, ergibt sich damit als

$$\begin{aligned} \Pr(H_K(t+1)) &= \\ \Pr(H_K(t))^2 + (x_H - \Pr(H_K(t))) &\left( \Pr(H_K(t)) + \frac{1}{n} \right) + (1 - x_H) \Pr(H_K(t)). \end{aligned} \quad (4-2)$$

Dieser Ausdruck vereinfacht sich zu

$$\Pr(H_K(t+1)) = \left( 1 - \frac{1}{n} \right) \Pr(H_K(t)) + \frac{1}{n} x_H. \quad (4-3)$$

Diese Beziehung gibt nun ganz allgemein an, wie die Wahrscheinlichkeiten, einen bekannten Hawk zu treffen, für zwei benachbarte, beliebige Zeitpunkte zusammenhängen. Da die Beziehung für beliebige  $t, t+1$  gilt, so kann man offensichtlich  $\Pr(H_K(t))$  durch eine analoge Beziehung für die benachbarten Zeitpunkte  $t$  und  $t-1$  substituieren. Also: Sicherlich gilt für  $t$  und  $t-1$  die Beziehung

$$\Pr(H_K(t)) = \left( 1 - \frac{1}{n} \right) \Pr(H_K(t-1)) + \frac{1}{n} x_H. \quad (4-4)$$

Einsetzen von (4-4) in (4-3) liefert

$$\Pr(H_K(t+1)) = \left( 1 - \frac{1}{n} \right)^2 \Pr(H_K(t-1)) + \frac{1}{n} x_H \left( 1 - \frac{1}{n} \right) + \frac{1}{n} x_H. \quad (4-5)$$

Nun ist klar, daß man den unbekanntes Wahrscheinlichkeitsterm gänzlich eliminieren kann, einfach durch sukzessive Substitution. Um die evolvierende Struktur noch etwas deutlicher zu Tage treten zu lassen, substituieren wir ein weiteres Mal. Für die Zeitpunkte  $t-1, t-2$  gilt



$$Pr(H_K(t-1)) = \left(1 - \frac{1}{n}\right) Pr(H_K(t-2)) + \frac{1}{n} x_H. \quad (4-6)$$

Einsetzen von (4-6) in (4-5) und vereinfachen führt zu

$$\begin{aligned} & Pr(H_K(t+1)) \\ &= \left(1 - \frac{1}{n}\right)^3 Pr(H_K(t-2)) + \frac{1}{n} x_H \left(1 - \frac{1}{n}\right)^2 + \frac{1}{n} x_H \left(1 - \frac{1}{n}\right) + \frac{1}{n} x_H. \end{aligned} \quad (4-7)$$

Führt man diese Methode des Lösens durch Rückwärts-Iteration nun genau  $t$ -mal durch, so erhält man

$$Pr(H_K(t+1)) = Pr(H_K(\tau=0)) \left(1 - \frac{1}{n}\right)^t + \frac{1}{n} x_H \sum_{\tau=0}^{t-1} \left(1 - \frac{1}{n}\right)^\tau. \quad (4-8)$$

Klarerweise gilt  $Pr(H_K(\tau=0)) = 0$ , so daß (4-8) sich vereinfacht zu

$$Pr(H_K(t+1)) = \frac{1}{n} x_H \sum_{\tau=0}^{t-1} \left(1 - \frac{1}{n}\right)^\tau = \frac{1}{n} x_H \sum_{\tau=0}^{t-1} \left(\frac{n-1}{n}\right)^\tau. \quad (4-9)$$

Die Summe der geometrischen Reihe bestimmt sich leicht zu

$$\sum_{\tau=0}^{t-1} \left(\frac{n-1}{n}\right)^\tau = n \left(1 - \left(\frac{n-1}{n}\right)^t\right). \quad (4-10)$$

Einsetzen in (4-9) liefert dann das gewünschte Resultat:

$$Pr(H_K(t+1)) = x_H \left(1 - \left(\frac{n-1}{n}\right)^t\right).$$

□

Aus Gründen der Einfachheit wurde das obige Resultat abgeleitet für den Fall einer singulären K-Mutante. Es ändert sich jedoch nichts, wenn man zuläßt, daß die K-Typen eine Subpopulation mit beliebiger absoluter Größe oberhalb von 1 darstellen. In diesem Fall betrachtet man das Problem aus der Perspektive eines singulären, zufällig gezogenen K-Typs, der nun natürlich auch andere K-Typen treffen kann. In diesem Fall zerfällt die Urne aus Sicht eines solchen K-Mutanten in insgesamt vier verschiedene Elemente, nämlich bekannte und unbekannte

Hawks, Doves und andere K-Typen. Letztere können aber einfach zusammengefaßt werden, da sie die zeitliche Entwicklung der Wahrscheinlichkeit, einen bekannten Hawk zu treffen, gar nicht beeinflussen. Der Populationsanteil dieser zusammengefaßten Gruppe der kooperativen Typen ergibt sich wiederum residual zu dem der Falken insgesamt, so daß sich an der Ableitung obiger Formel nichts ändert.

Diese Argumentation deutet bereits an, daß das oben entwickelte Lernmodell wesentlich allgemeiner ist, als es auf den ersten Blick scheinen mag. Dies ist in der Tat der Fall. Es ist anwendbar auf Populationen mit einer beliebigen Anzahl verschiedener Elemente. Betrachtet sei also eine Population (Urne) endlicher Größe  $n+1$ , die aus  $i=1, \dots, k$  verschiedenen Typen (Elementen)  $m^i$  zusammengesetzt ist. Die Population spielt wieder ein pairwise random matching game. Wir stellen uns die Frage, mit welcher Wahrscheinlichkeit  $Pr(m_K^i(t+1))$  in einem beliebigen Zeitpunkt  $t+1$  ein bekanntes Element des Typs  $m^i$  gezogen wird. Bezeichne dabei  $x_i$  die relative Häufigkeit des  $i$ -ten Elements in der Urne (den Populationsanteil des  $i$ -ten Typs). Die Argumentation ist völlig analog zum obigen Beweis. Wird in  $t$  ein bereits bekanntes Element des Typs  $m^i$  gezogen, so ändert sich die Wahrscheinlichkeit in  $t+1$  nicht. Wird irgendein Element eines anderen, von  $i$  verschiedenen Typs gezogen, so bleibt die Wahrscheinlichkeit, in  $t+1$  einen bekannten Vertreter von  $m^i$  zu treffen, ebenfalls unverändert. Lediglich wenn in  $t$  ein bislang unbekanntes Element des Typs  $m^i$  getroffen wird, ändert sich die Wahrscheinlichkeit für den Zeitpunkt  $t+1$ . Die entsprechenden Wahrscheinlichkeiten sind gegeben als

$$Pr(m_K^i(t+1) | m_K^i \text{ in } t) = Pr(m_K^i(t))^2, \quad (4-11)$$

$$Pr(m_K^i(t) | m_U^i \text{ in } t) = \left( Pr(m_K^i(t)) + \frac{1}{n} \right) (x_i - Pr(m_K^i(t))), \quad (4-12)$$

$$Pr(m_K^i(t+1) | m^j, j \neq i \text{ in } t) = Pr(m_K^i(t)) (1 - x_i). \quad (4-13)$$

Man bestimme die totale Wahrscheinlichkeit und löse in gewohnter Manier durch Rückwärts-Iteration, so erhält man

$$Pr(m_K^i(t+1)) = x_i \left( 1 - \left( \frac{n-1}{n} \right)^t \right). \quad (4-14)$$

Einige Eigenschaften des Lernmodells faßt das folgende Lemma zusammen.

**Lemma 4.1**

i) für finites t:

$$\lim_{n \rightarrow \infty} \Pr(m_K^i(t+1)) = 0 \quad (4-15)$$

ii) für finites n:

$$\lim_{t \rightarrow \infty} \Pr(m_K^i(t+1)) = x_i \quad (4-16)$$

iii) Die Änderung der Wahrscheinlichkeit für zwei Zeitpunkte t und t+1 ist gegeben als

$$\begin{aligned} \Delta \Pr(m_K^i(t+1)) &= x_i \left( \frac{n-1}{n} \right)^t \left( \frac{1}{n-1} \right) \\ \forall n > 1, x_i > 0 \text{ gilt: } \Delta \Pr(m_K^i(t+1)) &> 0 \end{aligned} \quad (4-17)$$

**Beweis:**

(i) und (ii) sind offensichtlich. ad (iii):

$$\begin{aligned} \Delta \Pr(m_K^i(t+1)) &= x_i \left( 1 - \left( \frac{n-1}{n} \right)^t \right) - x_i \left( 1 - \left( \frac{n-1}{n} \right)^{t-1} \right) \\ &= x_i \left[ \left( \frac{n-1}{n} \right)^{t-1} - \left( \frac{n-1}{n} \right)^t \right] \\ &= x_i \left( \frac{n-1}{n} \right)^{t-1} \left[ \left( \frac{n-1}{n} \right)^{-1} - 1 \right] \\ &= x_i \left( \frac{n-1}{n} \right)^{t-1} \left[ \frac{n - (n-1)}{n-1} \right] \\ &= x_i \left( \frac{n-1}{n} \right)^{t-1} \left[ \frac{1}{n-1} \right] > 0 \forall x_i > 0, n > 1 \end{aligned}$$

□

Man kann außerdem leicht zeigen, daß die Wissenszunahme im Zeitablauf abnimmt. Um dies zu zeigen, betrachten wir die Änderungen der Wahrscheinlichkeit, ein bekanntes Element zu ziehen, für drei aufeinanderfolgende Zeitpunkte t-1, t und t+1. Sei  $\Delta_1 = \Pr(m_K^i(t+1)) - \Pr(m_K^i(t))$ ,  $\Delta_2 = \Pr(m_K^i(t)) - \Pr(m_K^i(t-1))$ .

Wir betrachten dann

$$\Delta_1 - \Delta_2 = x_i \left( \frac{n-1}{n} \right)^t \left[ \frac{1}{n-1} \right] - x_i \left( \frac{n-1}{n} \right)^{t-1} \left[ \frac{1}{n-1} \right]$$

$$\begin{aligned}
&= x_i \left( \frac{n-1}{n} \right)^t \left[ \frac{1}{n-1} \right] \left[ 1 - \left( \frac{n-1}{n} \right)^{-1} \right] \\
&= x_i \left( \frac{n-1}{n} \right)^t \left[ \frac{1}{n-1} \right] \left[ 1 - \left( \frac{n}{n-1} \right) \right] \\
&= x_i \left( \frac{n-1}{n} \right)^t \left[ \frac{1}{n-1} \right] \left[ \frac{n-1-n}{n-1} \right] \\
&= -x_i \left( \frac{n-1}{n} \right)^t \left[ \frac{1}{n-1} \right]^2,
\end{aligned} \tag{4-18}$$

was ganz offensichtlich negativ ist für alle  $n > 1$ . Der Lernprozeß ist also insgesamt charakterisiert durch positiven, aber abnehmenden Wissenszuwachs im Zeitablauf.

## 5. Das GD mit einer lernfähigen Mutante

### 5.1. Payoffs

Wir wollen in den folgenden Abschnitten der Frage nachgehen, ob und unter welchen Umständen sich eine lernfähige kooperative Mutante evolutionär durchsetzt. Dazu betrachten wir eine Population, die zunächst in beliebiger Zusammensetzung aus defektiven Typen, sogenannten Hawks, und kooperativen Typen, sogenannten Doves besteht. In dieser Population tauche nun ein Typ auf, der als knowledgeable Dove bezeichnet sei. Diese Mutante zeichnet sich wesentlich durch zwei Eigenschaften aus: i) Sie ist lernfähig im Sinne des zuvor entwickelten Lernmodells. Das heißt, die Mutante lernt durch sukzessive Interaktion andere Akteure aus der Population kennen; (ii) die von ihr verfolgte Strategie. Es sei angenommen, daß der K-Typ ausbeutungsavers ist. Dies äußere sich dergestalt, daß, wann immer der K-Typ einen bekannten Defektor trifft, er die Interaktion mit diesem Akteur verweigert. Es wird angenommen, daß in diesem Fall beide Akteure eine outside-option wahrnehmen, die ihnen einen Payoff sichert, der unterhalb des Payoffs bei einmütiger Kooperation liegt, jedoch den Payoff bei einmütiger Defektion nicht unterschreitet. Sei der Wert dieser outside option mit  $\lambda$  bezeichnet. Die Payoffs des nicht modifizierten GDs seien die in Tabelle 8 gegebenen. Dann gelte  $1 \leq \lambda < 3$ .

	D	H
D (Koop)	3	0
H (Defektion)	6	1

Tabelle 8: Gefangenen-Dilemma

Die Payoffstruktur des modifizierten GD ist damit noch nicht vollständig charakterisiert. Es sei ferner angenommen, daß die Mutante aufgrund ihrer Lernfähigkeit Kosten zu tragen hat. Diese Annahme läßt sich in unterschiedlicher Weise begründen. Die erste Begründungsvariante führt zu einer letztlich strikt biologischen Interpretation des Modells. In der biologischen Evolutionstheorie ist es Konsens, daß Innovationen, die zu einem selektiven Vorteil führen, zumeist mit bestimmten Kosten untrennbar verbunden sind. Zahlreiche Beispiele aus der Entwicklungsgeschichte von Tierarten belegen diese These. Eines der bekanntesten ist das der Giraffe. Der im Vergleich zu ihrem entwicklungsgeschichtlichen Vorgänger längere Hals führte in einer Welt knapperer Nahrungsressourcen (aufgrund vorangegangener klimatischer Veränderungen) zu einem Selektionsvorteil: Die Blätter der

in der Savanne wachsenden Bäume konnten als Nahrungsreservoir nunmehr erschlossen werden. Dieser Vorteil wurde jedoch durch einen gravierenden Nachteil erkaufte. Die Zahl der Halswirbel erhöhte sich nämlich nicht, sondern verharrte konstant bei sieben. Selten (1990, S.8) beschreibt die Konsequenzen und deutet eine Erklärung an:

*„This is very inconvenient for the giraffe. It has difficulties lying down to sleep and standing up quickly in danger. It actually sleeps very little. Why did evolution fail to increase the number of neck bones? The answer is simple: **Evolution cannot change many things at once.** A change of the number of neck bones alone would be disastrous. Muscles, nerves, and other things would have to be adjusted. This would require many simultaneous mutations.“*

Simultane Mutationen sind jedoch extrem unwahrscheinlich, so daß sie praktisch niemals beobachtet werden.

Bezüglich der Hominidenevolution ist relativ unstrittig, daß vor zirka 2.5 Millionen Jahren ein rapides Schädelwachstum mit einer entsprechenden Vergrößerung des Hirnvolumens einsetzte. Dieses Phänomen wird von einigen Anthropologen als der entscheidende Quantensprung in der intellektuellen Menschheitsentwicklung angesehen. Der Vorteil größerer intellektueller Kapazitäten wurde aber auch in diesem Fall durch einen Nachteil erkaufte. Die Vergrößerung des Schädels machte letztlich eine Verkürzung der Schwangerschaftsperiode erforderlich mit der Folge, daß Säuglinge nun früher geboren wurden und eine längere Periode intensiverer elterlicher Zuwendung benötigten. Der erhöhte Ressourceneinsatz (Zeit!) für die „Aufzucht“ von Kindern kann durchaus als Preis des intellektuellen Fortschritts der Menschheit gedeutet werden.

Was bedeuten diese Überlegungen für das hier vorliegende Modell? Man kann in der Tat der Auffassung sein, die Lernfähigkeit unseres mutanten Typs stelle im biologischen Sinne eine evolutionäre Innovation dar. Etwa könnte man argumentieren, daß Lernfähigkeit an das Vorhandensein bestimmter, hinreichend entwickelter Hirnregionen gebunden ist. Offenkundig erfordert Erfahrungslernen ein Gedächtnis. Letzteres ist aber anatomisch gebunden an die Existenz einer Großhirnrinde, welche ihrerseits erst ab einem bestimmten entwicklungsgeschichtlichen Zeitpunkt nachweisbar ist. Möglicherweise läßt sich die Entstehung von Gedächtnis und Lernfähigkeit sogar zeitlich direkt in Verbindung bringen mit dem oben erwähnten „intellektuellen Quantensprung“. Ich denke jedoch, daß eine derartige Modellinterpretation sich reichlich spekulativ ausnimmt, ganz abgesehen davon, daß sie dem Erklärungsziel der vorliegenden Arbeit wenig dienlich ist. Die Analyse des hier erst skizzenhaft entwickelten Modells wird nämlich unter ande

rem zeigen, daß eine Koexistenz aller drei Typen – Doves, Hawks und K-Typen – möglich ist. In einer streng biologischen Interpretation bedeutet dies die Koexistenz von Akteuren mit deutlich unterschiedlich entwickelten kognitiven Kapazitäten, also die Koexistenz von Hominiden auf unterschiedlichen Entwicklungsstufen. Aus ökonomischer Sicht ist ein derartiges Resultat als einigermaßen exotisch zu beurteilen. Möglicherweise mag unser Modell auch zweckdienlich sein, um bestimmte Phasen der menschlichen Phylogenese mit zeitlicher Überlappung verschiedener Hominidentypen zu erklären (diese Phasen sind ja empirisch belegt!, siehe z.B. Tattersall (2000)). Jedoch ist erklärtes Ziel dieser Arbeit, einen Hinweis zu geben zur Erklärung der Koexistenz verschiedener, persistenter Verhaltensdispositionen, die in aktuellen ökonomischen und psychologischen Experimenten mit heutigen Vertretern des homo sapiens sapiens beobachtet wurden. Mögliche paläobiologische Anwendungsmöglichkeiten unseres Modells werden daher im weiteren nicht berücksichtigt.

Eine zweite, überzeugendere Begründungsvariante für Kosten der Lernfähigkeit bietet sich jedoch an. Diese Kosten könnten bezeichnet werden als Gedächtniskosten oder Gedächtnistrainings-Kosten. Die Idee ist, daß die Aufrechterhaltung der Funktionsfähigkeit eines Gedächtnisses gewisse Ressourcen beansprucht. Etwas salopp formuliert könnte man sich vorstellen, daß ein K-Typ sich jeden Tag einem brain-jogging unterzieht, um seine Memorierfähigkeit zu erhalten. Es ist bekannt, daß die Leistungsfähigkeit des Gedächtnisses ganz maßgeblich von dessen Training abhängt. Oder aber man stelle sich vor, daß der K-Typ im Gegensatz zu den anderen Typen sich der Mühe unterzieht, eine Kartei anzulegen, in der die bisherigen Interaktionen archiviert werden, um etwa dann die gewünschte Information über einen Akteur zu liefern, wenn das Gedächtnis einmal versagen sollte. Die Idee ähnelt Vorstellungen von Conlisk (1980) oder etwa Guttman (1996), denen zufolge jede Form optimierendes Verhaltens mit Kosten verbunden ist. Nun kann die K-Mutante unseres Modells tatsächlich in einem gewissen Sinne als optimierender Typ interpretiert werden. Gegeben ihre grundsätzlich kooperative Verhaltensdisposition kann die Verweigerung der Interaktion mit einem bekannten Hawk auch schlicht als best-reply Verhalten gedeutet werden. Die Wahl der outside-option im Angesicht eines bekannten Defektors ist, gegeben der Strategieraum weist daneben lediglich die Alternativen „kooperiere“ oder „defektiere“ aus, die beste Antwort.

Motiviert durch die obigen Überlegungen sei daher angenommen, daß der K-Typ in jedem Zeitpunkt  $t$  Gedächtniskosten von  $c$ ,  $0 < c < 1$  zu tragen hat. Die Annahmen bezüglich der Payoffstruktur des modifizierten GD sind damit komplettiert. Die gesamte Payoffstruktur wird in der folgenden Tabelle zusammengefaßt.

	Dove	Hawk (kn.)	Hawk (unkn.)	K-Typ
Dove	3	0	0	3
Hawk (known)	6	1	1	$\lambda$
Hawk (unknown)	6	1	1	6
K-Typ	3-c	$\lambda$ -c	-c	3-c

Tabelle 9: Gefangenen Dilemma mit lernfähiger Mutante

### 5.2 Konzeptionelle Vorüberlegungen

Es wird im folgenden darum gehen, das obige modifizierte GD mit Methoden der evolutionären Spieltheorie zu lösen. Wie bereits erwähnt betrachten wir eine Population, deren Mitglieder zufällig und paarweise gematcht werden, und die dann das 2-Personen-one-shot GD spielen. Die Spieler erhalten ihren Payoff, werden in die Urne „zurückgelegt“ und der matching-Prozess beginnt von neuem. Insgesamt wird aus der Population T+1-mal gezogen, wobei T endlich ist.

Um das obige Spiel „zu lösen“, werden wir das Standard-Instrumentarium der evolutionären Spieltheorie verwenden. Dies besteht im wesentlichen aus der sogenannten Replikator Dynamik. Bezeichne  $\bar{\pi}_i$  den durchschnittlichen Payoff, den ein Spieler des Typs i erzielt, und bezeichne  $\bar{\pi}$  den durchschnittlichen Populationspayoff (d.h.  $\bar{\pi}$  ist das gewichtete arithmetische Mittel der Durchschnittspayoffs aller Typen i, als Gewichte werden die Populationsanteile der Typen i verwendet). Bezeichne  $x_i$  den Anteil aller Spieler des Typs i an der Gesamtpopulation.  $x_i$  ist also die relative Häufigkeit des Typs i. Dann ist die Replikator-Dynamik definiert als

$$\dot{x}_i = x_i(\bar{\pi}_i - \bar{\pi}), \forall i. \tag{5-1}$$

Die zeitliche Änderung der Populationsanteile wird also als direkt proportional zur Differenz zwischen dem Durchschnittspayoff eines Typs i und dem Populationsdurchschnitt angenommen. Solche Strategien, die eine durchschnittliche Auszahlung oberhalb des Populationsdurchschnitts erzielen, weisen also eine positive zeitliche Änderung auf. Mit anderen Worten nimmt also die relative Häufigkeit von überdurchschnittlich guten Strategien im Zeitablauf zu, während die zeitliche Änderungsrate für unterdurchschnittliche Strategien negativ ist. Strategien, die persistent schlechter als der Durchschnitt abschneiden, sterben schließlich aus. Die Replikator Dynamik (RD) stellt eine naheliegende Möglichkeit der Formali



sierung des Darwinschen Gedanken des survival of the fittest dar. Sie ist in der Tat jedoch nur eine mögliche Formalisierung unter vielen anderen, die ebenfalls denkbar sind. Allerdings kann ohne Übertreibung behauptet werden, daß die RD sicherlich die populärste Dynamik in der evolutionären Spieltheorie darstellt, deren Eigenschaften außerdem überaus gut dokumentiert sind (vgl. neben anderen Weibull 1995, Vega-Redondo 1996, Hofbauer und Sigmund 1988).

Was nun in evolutionären Standardmodellen regelmäßig unterstellt wird, ist eine unendlich große Population. Die Funktion dieser Fiktion besteht darin, sicherzustellen, daß das Gesetz der großen Zahlen gilt. Dessen Gültigkeit sichert, daß im Grenzübergang der tatsächlich realisierte durchschnittliche Payoff mit dem theoretisch erwarteten Payoff koinzidiert.

Ganz offensichtlich ergibt sich damit ein Problem bezüglich des Lernmodells: Letzteres wurde ja entwickelt für den Fall einer endlichen Population. In einer infiniten Population jedoch kann Lernen endlicher Zeitdauer keine Effekte haben. Es erhebt sich damit die keineswegs triviale Frage, wie Lernen und evolutionäre Analyse überhaupt miteinander vereinbart werden können. Gibt es Möglichkeiten, einerseits die Annahme einer infiniten Population aufrecht zu erhalten, andererseits aber nicht triviale Effekte des Erfahrungslernens in dieser Population zu generieren?

Man kann auf sehr unterschiedliche Weisen auf dieses Problem reagieren. Die radikalste besteht in der Aufgabe der Annahme einer unendlich großen Population. Dies zieht jedoch unweigerlich große konzeptionelle Änderungen nach sich. In einer strikt endlichen Population gilt das Gesetz der großen Zahl natürlich nicht mehr, und die gewünschte Koinzidenz von realisiertem durchschnittlichen Payoff und Erwartungswert stellt sich nicht ein. Der durchschnittliche Payoff ist vielmehr korrekterweise als eine Zufallsvariable mit strikt positiver Varianz aufzufassen. In letzter Konsequenz kann daher der gänzlich deterministische Analyserahmen nicht aufrecht erhalten werden. Die Entwicklung der Populationsanteile folgt vielmehr einem stochastischen Prozeß. Ein entsprechendes Modell für unsere hier vorliegende Fragestellung wird in Kapitel 7 dieser Arbeit vorgestellt.

Etwa am anderen Ende der Skala möglicher Reaktionsweisen befindet sich ein Vorschlag, dessen Lösung letztlich in der Ignorierung des Problems besteht. Er lautet: Man stelle sich eine sehr große, jedoch endliche Population vor. In einer solchen Population gilt das Gesetz der großen Zahl näherungsweise. Der Fehler, den man begeht, wenn man mit dem Erwartungswert rechnet, wird für eine hinreichend große Population entsprechend klein und damit tolerabel. Nun ist es zweifelsohne richtig, daß die Abweichung des empirischen Durchschnitts vom theoretischen Erwartungswert mit zunehmender Zahl der Wiederholungen eines Zufallsexperiments kleiner wird. Eben dies ist ja die Aussage des Gesetzes der großen

Zahl. Ebenso unstrittig richtig ist aber auch, daß der Fehler lediglich im Grenzübergang infinitesimal klein wird. Rein formal muß man also eine infinite Population verlangen, wenn man mit Erwartungswerten rechnen möchte.

Gibt es möglicherweise eine mittlere Position, die einerseits formal einwandfrei ist, andererseits aber den radikalen Wechsel des deterministischen Modellrahmens vermeidet? Man kann in der Tat die Verwendung der (deterministischen) Replikator Dynamik retten durch geeignete interpretatorische Kunstgriffe. Es bieten sich im wesentlichen zwei an. Zum ersten kann man die Vorstellung einer unendlich großen Gesamtpopulation bemühen, die in Subpopulationen endlicher Größe zerfällt<sup>13</sup>. Unter der Annahme, daß Interaktionen nur zwischen den Mitgliedern einer Subgruppe, nicht aber zwischen Mitgliedern verschiedener Subgruppen stattfinden, ist das Lernmodell anwendbar und zeitigt im allgemeinen nicht triviale Effekte. Andererseits gilt in der Aggregation der Payoffs *über die Subgruppen* hinweg wieder das Gesetz der großen Zahl, so daß Durchschnitts- und erwarteter Payoff tatsächlich zusammenfallen. Ein entsprechendes Modell wird im folgenden Abschnitt 5.3 dargestellt. Zum zweiten leistet folgende Konstruktion das Gewünschte: Man betrachte eine endliche Population und führe dann einen simultanen Grenzübergang durch, in dem Populationsgröße und Zeit gegen Unendlich streben. Während bei partiellen Grenzübergängen die Resultate des Lernmodells notwendigerweise triviale sind – entweder wird überhaupt keine Information generiert oder aber der Lernprozeß konvergiert gegen perfektes Wissen – hängt das Resultat im simultanen Fall vom Verhältnis der Konvergenzraten beider Größen ab, ist also im allgemeinen nicht trivialer Art. Dieser Ansatz wird in Kapitel 5.5 diskutiert.

### 5.3 Basis-Modell : Subpopulationen finiter Größe

Wenden wir uns aber zunächst dem intuitiv vielleicht plausibleren Ansatz zu und betrachten eine Population, die aus einer unendlichen Anzahl von Subpopulationen endlicher Größe besteht. Sei die absolute Größe einer Subpopulation  $n+1$ , wobei  $n$  eine ungerade ganze Zahl ist. Wir nehmen an, daß Interaktionen innerhalb der Subpopulationen stattfinden, nicht aber zwischen Mitgliedern verschiedener Subpopulationen. Die Mitglieder einer jeden Gruppe werden also paarweise und zufällig gematcht und spielen das simultane one-shot GD. Diese Prozedur wird genau  $T+1$ -mal wiederholt, wobei  $T$  endlich ist.

Bezeichne nun  $Q$  die Menge aller möglichen Gruppenzusammensetzungen und bezeichne  $Q_j$  ein Element aus  $Q$ , d.h.  $Q_j$  ist ein Tripel  $(H_j, D_j, K_j)$  mit der

---

<sup>13</sup> Den Anstoß zu dieser Modellierung gab Erwin Amann, dem ich dafür herzlich danke.

Eigenschaft  $H_j + D_j + K_j = n + 1$ . Bezeichne ferner  $q$  die Anzahl aller möglichen Tripel  $(H_j, D_j, K_j)$ . Da  $n+1$  endlich ist, muß auch  $q = \#\{(H_j, D_j, K_j) \mid H_j + D_j + K_j = n + 1\}$  endlich sein. Daher  $Q = \{Q_1, \dots, Q_q\}$ .

Betrachtet sei nun ein beliebiger gegebener Populationszustand  $(x_H, x_D, x_K)$ ,  $\sum_i x_i = 1, i = H, D, K$ . Klarerweise wird ein jeder solcher

Populationszustand realisiert durch eine Häufigkeitsverteilung auf der Menge  $Q$ .

Bezeichne  $f_{Q_j}$  die relative Häufigkeit von Tripel  $Q_j$ ,  $\sum_{j=1}^q f_{Q_j} = 1$ . Dann gilt offen-

sichtlich  $(x_H, x_D, x_K) = \sum_{j=1}^q \frac{Q_j}{n+1} f_{Q_j}$ . Bezeichne nun  $r_{Q_j}$  die absolute Häufigkeit des Tripels  $Q_j$ . Die absolute Größe der Gesamtpopulation ergibt sich damit aus

$\sum_{j=1}^q Q_j r_{Q_j} = \sum_{j=1}^q (n+1) r_{Q_j}$ . Da die Gesamtpopulation unendlich groß ist, andererseits  $(n+1)$  sowie  $q$  endliche Größen sind, muß für jedes  $Q_j$  mit  $f_{Q_j} > 0$   $r_{Q_j}$  unendlich betragen.

Dies ist der entscheidende Punkt des Arguments: Da jedes Tripel mit positiver relativer Häufigkeit eine infinite Anzahl von Realisationen aufweisen muß, gilt bezüglich eines solchen Tripels das Gesetz der großen Zahl. Dies impliziert, daß der durchschnittliche Payoff bezüglich aller vorkommenden Gruppen, die zum Tripel  $Q_j$  gehören, dem theoretisch erwarteten Payoff entspricht. Der Gesamtpayoff für einen Spieler des Typs  $i$  ergibt sich dann einfach als gewichtete Summe der Payoffs über alle Tripel  $Q_j$ . Formal: Bezeichne  $\bar{\pi}_{Q_j}^i$  den durchschnittlichen Payoff des Typs  $i$  der innerhalb aller Gruppen vom Typ  $Q_j$  erzielt wird. Dann gilt

$$\bar{\pi}_{Q_j}^i = E\pi_{Q_j}^i \text{ und } \bar{\pi}^i = \sum_{j=1}^q \bar{\pi}_{Q_j}^i f_{Q_j} = \sum_{j=1}^q E\pi_{Q_j}^i f_{Q_j}.$$

Damit sind alle wesentlichen Vorarbeiten für die dynamische Analyse des Spiels abgeschlossen. Lemma 5.1 gibt Auskunft über die Gleichgewichte des Spiels. Es ist wohlbekannt, daß jeder lokal stabile Populationszustand unter der Replikator Dynamik zugleich ein Nash-Gleichgewicht (NGG) des zugrunde liegenden Basis-spiels sein muß (Bomze 1986). Wir konzentrieren uns daher von vornherein auf solche Fixpunkte der Replikator Dynamik, die zugleich NGG sind.

### Lemma 5.1: (Existenz)

*Betrachtet sei eine Population bestehend aus einer unendlichen Zahl von Subpopulationen der finiten absoluten Größe  $n+1$ ,  $n \in \mathbb{N}$ ,  $n$  ungerade. Jede Subpopulation spiele das 3x3-Gefangenen-Dilemma mit einer lernfähigen kooperativen Mutante im pairwise random matching Verfahren  $T+1$ -mal. Sei die Ent*

wicklung der Populationsanteile der drei Typen Hawk, Dove und knowledgeable Dove durch die Replikator Dynamik beschrieben. Dann gilt:

i) für  $c < \frac{(n+1)\lambda(1-k)^2(\lambda-6) + \lambda(1-k)(5+3n)}{n(2+(1-k)(\lambda-6))}$  existiert ein inneres Nash-

Gleichgewicht

$$(x_H^*, x_K^*, x_D^*) = \left\{ \frac{n}{n+1} \frac{c}{\lambda(1-k)}, \frac{n(2c-3\lambda(1-k))-5\lambda(1-k)}{(n+1)\lambda(\lambda-6)(1-k)^2}, 1-x_H^*-x_K^* \right\};$$

ii) für  $\frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5]-2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} < c < \frac{n+1}{n} \lambda(1-k) - \frac{3+n}{n} - 1$

$$\text{oder } \frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5]-2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} > c > \frac{n+1}{n} \lambda(1-k) - \frac{3+n}{n} - 1$$

ist der Zwei-Typen-Fixpunkt

$$(x_H^*, x_K^*, x_D^*) = \left\{ 1-x_K^*, \frac{\lambda(1-k) - (1+c) \frac{n}{n+1} - \frac{3}{n+1}}{((1-k)(2\lambda-6)+2)}, 0 \right\} \text{ ein Nash-}$$

Gleichgewicht;

iii) für  $c > \lambda(1-k) \frac{n+1}{n} - \frac{n+3}{n}$  ist der Fixpunkt  $(x_H^*, x_K^*, x_D^*) = \{1, 0, 0\}$  ein

$$\text{Nash-Gleichgewicht, wobei } k = \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right).$$

**Beweis:** siehe Appendix

Im Kontext einer explizit dynamischen Theorie ist die Frage der Existenz von Fixpunkten nur von untergeordneter Bedeutung. Weit bedeutsamer ist natürlich die Frage ihrer Stabilität: Reichen bereits marginale Auslenkungen aus einem Fixpunkt, damit dieser verlassen wird, so wird man kaum erwarten können, einen solchen Fixpunkt jemals empirisch zu beobachten. Der folgende Satz faßt die Ergebnisse der Stabilitätsanalyse zusammen.

**Satz 5.1 (Stabilität):**

i) Für

$$c < \min \left\{ \frac{(n+1)\lambda(1-k)^2(\lambda-6) + \lambda(1-k)(5+3n)}{n(2+(1-k)(\lambda-6))}, (1-k) \left( 2\lambda \frac{(n+2)}{n} - 3 \frac{(n+3)}{n} \right) \right\}$$

ist der innere Fixpunkt global stabil.

ii) Für 
$$\frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5]-2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} < c < \frac{n+1}{n}\lambda(1-k) - \frac{3+n}{n} - 1$$

ist der Zwei-Typen-Fixpunkt mit Falken und K-Typen global stabil.

iii) Für  $c > \lambda(1-k)\frac{n+1}{n} - \frac{n+3}{n}$ , ist die monomorphe Randlösung mit Falken allein global stabil.

**Beweis:**

Bevor wir uns mit der Fragestellung globaler Stabilität befassen, ist es sinnvoll, die Frage asymptotischer lokaler Stabilität zu untersuchen.

ad i) Die Spur der Jakobimatrix, ausgewertet im inneren Nash-Gleichgewicht, ergibt sich als

$$trJ = \frac{(1-k)[n(3-2\lambda)-4\lambda+9]+cn}{(1-k)(n+1)(1-k)(6-\lambda)} 2 \frac{c}{\lambda}. \tag{5-2}$$

Die Determinante stellt sich etwas komplizierter dar:

$$DetJ = (2cn - \lambda(1-k)(3n+5)) \frac{c}{\lambda^2} \cdot \frac{\lambda^2(n+1)(1-k)^2 - \lambda(3n(1-k(3-2k))+1+k(7-6k))+c(n(4-6k-\lambda(1-k)))}{n(\lambda-6)(k-1)^3(n+1)} \tag{5-3}$$

Ein Standardresultat der Theorie dynamischer Systeme im  $\mathbb{R}^2$  besagt, daß ein Fixpunkt dann lokal stabil ist, wenn die Spur der Jakobimatrix des zugrundeliegenden dynamischen Systems negativ ist, während die Determinante ein positives Vorzeichen tragen muß. Beginnen wir mit der Spur:

$$trJ < 0 \Leftrightarrow c < (1-k) \left( 2\lambda \frac{(n+2)}{n} - 3 \frac{(n+3)}{n} \right) \tag{5-4}$$

Die Bestimmung des Vorzeichens der Determinante erfordert etwas größeren argumentativen Aufwand. Man kann zunächst leicht zeigen, daß  $0 \leq k < 1$ . Ferner galt:  $\lambda < 3$ . Dann gilt klarerweise  $n(\lambda-6)(k-1)^3(n+1) > 0$ . Wir betrachten als nächstes den Term  $(2cn - \lambda(1-k)(3n+5)) \frac{c}{\lambda^2}$ . Wegen

$$x_K^* > 0 \Leftrightarrow \frac{n(2c - 3\lambda(1-k)) - 5\lambda(1-k)}{(n+1)\lambda(\lambda-6)(1-k)^2} = \frac{2cn - \lambda(1-k)(3n+5)}{(n+1)\lambda(\lambda-6)(1-k)^2} > 0$$

muß offensichtlich gelten  $(2cn - \lambda(1-k)(3n+5)) \frac{c}{\lambda^2} < 0$ . Damit ergibt sich:

$$DetJ > 0$$

$$\Leftrightarrow \lambda^2(n+1)(1-k)^2 - \lambda(3n(1-k(3-2k)) + 1 + k(7-6k)) + c(n(4-6k - \lambda(1-k))) < 0$$

$$\Leftrightarrow c < \frac{(n+1)\lambda(1-k)^2(\lambda-6) + \lambda(1-k)(5+3n)}{n(2 + (1-k)(\lambda-6))} \quad (5-5)$$

Offensichtlich ist (5-5) mit der Existenzbedingung aus Lemma 5.1 identisch. Die Determinante ist also positiv, sobald der innere Fixpunkt existiert.

ad ii) Um den zweiten Teil von Lemma 1 zu beweisen, erweist es sich als einfacher, direkt die Eigenwerte der Jakobimatrix im Zwei-Typenfixpunkt zu berechnen. Diese sind gegeben als

$$\varepsilon_1 = \frac{(n+1)[6\lambda k^2 - \lambda^2(1-k)^2] - cn[(\lambda-6)(1-k) + 2] + \lambda[3n(1-3k) + 1 - 7k]}{n[(2\lambda-6)(1-k) + 2]}, \quad (5-6)$$

$$\varepsilon_2 = (\lambda(1-k)(n+1) - 3 - n(1+c)) \frac{n(3-c) + 1 - (n+1)(\lambda(1-k) + 6k)}{n[(2\lambda-6)(1-k) + 2]}. \quad (5-7)$$

Wir betrachten zunächst  $\varepsilon_1$ . Zwei mögliche Fälle sind zu unterscheiden.  $[(2\lambda-6)(1-k) + 2] < 0$  (was natürlich  $(\lambda-6)(1-k) + 2 < 0$  impliziert). Lokale Stabilität erfordert dann

$$\varepsilon_1 < 0 \Leftrightarrow c > \frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6) + 5] - 2\lambda n(1-k)}{n[(1-k)(\lambda-6) + 2]} \equiv c_1. \quad (5-8)$$

Für  $[(2\lambda-6)(1-k) + 2] < 0$  ist der Zwei-Typenfixpunkt ein NGG dann und nur dann, wenn  $c < c_1$  (vgl. (5-73), Appendix). Falls also der Zwei-Typenfixpunkt ein Nash-Gleichgewicht ist, so ist dieses instabil.

Für  $[(2\lambda-6)(1-k) + 2] > 0$  müssen zwei Unterfälle unterschieden werden:

a)  $(\lambda-6)(1-k) + 2 < 0$ :

$$\varepsilon_1 < 0 \Leftrightarrow c < c_1, \quad (5-9)$$

was wiederum einen Widerspruch konstituiert zur Existenz eines NGG in dieser Parameterregion (vgl. (5-72), Appendix).

b)  $(\lambda - 6)(1 - k) + 2 > 0$ :

$$\varepsilon_1 < 0 \Leftrightarrow c > c_1, \quad (5-10)$$

was klarerweise mit der Existenzbedingung aus Lemma 5.1 identisch ist. Also existiert in diesem Parameterbereich das NGG und der erste Eigenwert der Jakobimatrix ist negativ.

In vollkommen analoger Weise ermittelt man für den zweiten Eigenwert, daß dieser ebenfalls für  $[(2\lambda - 6)(1 - k) + 2] > 0$  ein negatives Vorzeichen tragen kann. Dies ist der Fall, wenn folgende Bedingung erfüllt ist:

$$c_3 \equiv \frac{n+1}{n} \lambda(1-k) - \frac{3+n}{n} > c > \frac{n+1}{n} [(1-k)(6-\lambda) + 2] - \frac{3+n}{n} \quad (5-11)$$

Diese Ungleichungen sind identisch mit (5-66) (Appendix) und (5-67) (Appendix), den Bedingungen für die Existenz des Zwei-Typenfixpunktes in diesem Parameterbereich. Der zweite Eigenwert ist also negativ, wann immer der Fixpunkt existiert, der erste Eigenwert wird negativ, wenn der Zwei-Typenfixpunkt ein Nash-Gleichgewicht ist. Beide Eigenwerte sind also negativ, falls gilt

$$c_3 > c > c_1. \quad (5-12)$$

ad iii) Die Determinante der Jakobimatrix im Hawk-Fixpunkt ergibt sich als:

$$\begin{aligned} \text{Det}J &= -\frac{3+n}{n} \left[ \frac{(n+1)\lambda(1-k)}{n} - \frac{n(1+c)+3}{n} \right] > 0 \Leftrightarrow \\ &c > \frac{n+1}{n} \lambda(1-k) - \frac{3+n}{n} \end{aligned} \quad (5-13)$$

Die Spur berechnet sich als:

$$\text{tr}J = -\frac{3+n}{n} + \lambda \frac{(n+1)(1-k)}{n} - \frac{n(1+c)+3}{n} < 0$$

$$\Leftrightarrow c > \frac{n+1}{n} \lambda(1-k) - \frac{2n+6}{n} \quad (5-14)$$

Man zeigt sofort, daß (5-13) restriktiver als (5-14) ist. Dies vervollständigt die Beweisführung bezüglich lokaler Stabilität.

Um zu den gewünschten Aussagen über globale Stabilität zu gelangen, sind einige weitere Vorarbeiten notwendig.

### Lemma 5.2

*Betrachtet sei eine beliebige 2x2 Matrix A. Falls  $\text{tr}A > 0$ , dann gilt: Mindestens ein Eigenwert von A ist positiv oder beide komplex konjugierten Eigenwerte haben positive Realteile.*

**Beweis:** Der Beweis ist direkt.

### Definition (Permanenz):

Eine Replikator Dynamik heie permanent, falls ein  $\delta > 0$  existiert, so da fr alle Trajektorien  $v(\cdot)$  mit  $v_i(0) > 0$  fr alle  $i = 1, 2, \dots, m$  gilt  $\liminf_{t \rightarrow \infty} v_i(t) > \delta$ .

### Satz 5.2 (Hofbauer und Sigmund 1987):

Sei  $p$  ein Fixpunkt im Inneren des Simplex  $\Delta_n$  und gelte  $pAx > xAx$  fr alle Fixpunkte  $x \in \text{bd}\Delta_n$ . Dann ist die Replikator Dynamik permanent.

**Beweis:** siehe Vega-Redondo (1996) S. 65 f.

### Satz 5.3 (Hofbauer and Sigmund 1987):

Sei  $v(\cdot) \subset \text{int}\Delta_n$  eine beliebige Trajektorie einer permanenten Replikator Dynamik. Dann gilt  $\lim_{T \rightarrow \infty} \frac{1}{T} \int_{t=0}^T v(t) dt = v^*$  mit  $v^*$  als dem eindeutigen inneren Fixpunkt.

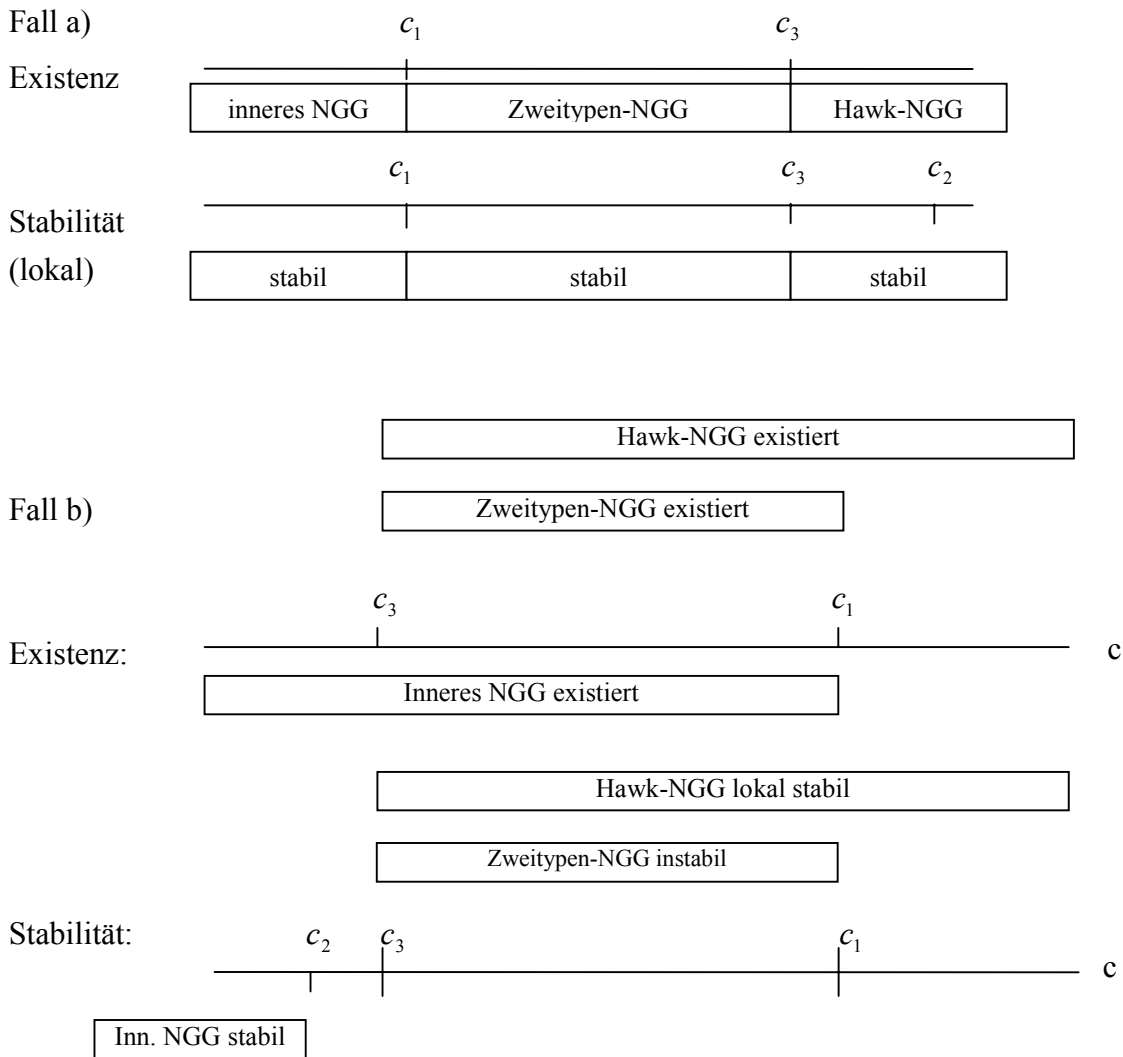
**Beweis:** siehe Vega-Redondo (1996) S.67 f.

### Beweis des Satzes 5.1:

Wir sind nun in der Lage Satz 5.1 zu beweisen, d.h. zu zeigen, da die Nash-GG aus Lemma 5.1 nicht nur lokal, sondern auch global stabil sind. Zunchst erweist es sich als sinnvoll die Frage zu klren, ob eine Koexistenz mehrerer lokal stabiler NGG mglich ist. Bezglich der Existenz von NGG waren zwei Flle zu unterscheiden,  $c_1 < c_3$  oder  $c_3 < c_1$ . Im ersten Fall zeigt man leicht, da dann auch gel



ten muß  $c_2 > c_3$ , und im zweiten Fall trifft gerade das Gegenteil zu. Die möglichen Konstellationen lassen sich in folgender Übersicht ablesen:



Beginnen wir mit Fall (a). Für  $c > c_1$  existiert kein innerer Fixpunkt. Wir wissen aus dem Poincaré-Bendixson-Theorem, daß die Grenzmenge eines dynamischen Systems in der Ebene entweder ein Fixpunkt oder aber ein Grenzzyklus sein muß. Hofbauer und Sigmund (1988) haben gezeigt, daß die Replikator Dynamik für Dimensionen  $n \leq 3$  keine isolierten inneren Grenzzyklen aufweisen kann. Weder ein innerer Fixpunkt noch ein innerer Grenzzyklus kommen also in diesem Fall als Grenzmenge in betracht. Der einzig verbleibende Kandidat für einen Grenzzyklus wäre der Rand des Simplex  $bd\Delta_3$ . Der Rand kann jedoch kein Zyklus sein, denn dies würde der Existenz lokal stabiler Fixpunkte widersprechen. Wie gezeigt können im Fall (a) keine stabilen Fixpunkte koexistieren. Entweder befinden wir uns also im Parameterbereich  $c_1 < c < c_3$ , für den somit der Zweitypenfixpunkt global stabil ist. Oder aber wir befinden uns im Bereich  $c > c_3$ , für den der Hawk-Fixpunkt global stabil ist.

Kommen wir nun zu Fall (b). Klarerweise gilt, daß für  $c > c_2$  die Spur der Jakobi-matrix ausgewertet im inneren Fixpunkt positiv wird. Aus Lemma 5.2 können wir damit sofort auf die Instabilität des inneren Fixpunktes schließen. Dann eröffnen sich zwei Möglichkeiten: i) Im Bereich  $c_2 < c < c_3$  existieren keine lokal stabilen Fixpunkte. Innere Grenzzyklen scheiden durch das Hofbauer-Sigmund-Argument wiederum als Grenzmenge aus. Es folgt, daß jede innere Trajektorie für  $t \rightarrow \infty$  schließlich  $bd\Delta_3$  erreicht, der in diesem Parameterbereich tatsächlich die eindeutige Grenzmenge darstellt. (ii) Für  $c > c_3$  ist der Hawk-Fixpunkt der einzige lokal stabile Fixpunkt, also ist er global stabil.

Betrachten wir schließlich die Fälle, in denen der innere Fixpunkt lokal stabil ist, also den Parameterbereich  $c < \min\{c_1, c_2\}$ . In beiden Fällen gilt, daß keine weiteren NGG existieren. Neben dem inneren Fixpunkt kommt also wiederum nur der Rand  $bd\Delta_3$  als Grenzmenge in betracht. Es läßt sich jedoch relativ einfach zeigen, daß die Replikator Dynamik gegenüber allen Randfixpunkten (d.h. den drei monomorphen Populationszuständen in den Ecken des Simplex sowie dem Zweitypenfixpunkt auf dem HK-Rand des Simplex) für  $c < \min\{c_1, c_2\}$  permanent ist. Also ist der innere Fixpunkt global stabil.

□

### 5.4 Diskussion der Resultate

Die Einführung einer lernfähigen, kooperativen Mutante in das Gefangenen Dilemma zeitigt bemerkenswerte Konsequenzen. Nicht nur der mutante kooperative Verhaltenstyp ist in der Lage, sich evolutionär zu behaupten. Vielmehr profitieren auch die unbedingt kooperativen Typen von der Existenz des K-Typen. Dies ist genau dann der Fall, wenn – für gegebenen Wert der outside-option  $\lambda$  und gegebene Lernbedingungen (abgebildet durch den Parameter  $k \equiv \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right)$ ) – die Kosten des Lernens  $c$  hinreichend niedrig sind.

In diesem Fall erweist sich der eindeutige innere Fixpunkt der Replikator Dynamik als global stabil. Unabhängig von den Anfangsbedingungen wird man langfristig immer in das innere Gleichgewicht gelangen. Die Ursache für dieses Modellergebnis ist natürlich darin zu sehen, daß durch Einführung des K-Typs „Immer kooperieren“ im Basisspiel nun nicht länger automatisch eine dominierte Strategie ist. Der K-Typ nützt nicht nur sich selbst durch sein Lernen. Er erhöht nicht nur seinen eigenen Payoff durch die Wahl der outside-option im Angesicht bekannter, gelernter Defektoren, sondern er reduziert gleichzeitig den Erfolg der defektiven Strategie. Dies eröffnet den naiven Kooperatoren gewissermaßen eine

Nische – im Windschatten der erfolgreichen K-Typen gedeihen auch die einfachen Tauben.

Die Existenz des stabilen inneren Gleichgewichtes hängt von den Modellparametern ab. Sind outside-option sowie Populationsgröße und Lebenszeit einer Generation fixiert, so muß der Kostenparameter  $c$  hinreichend niedrig sein, damit der innere Fixpunkt stabil ist. Anders herum gesagt muß der K-Typ, für gegebenen Wert der outside-option sowie gegebenen Parameter  $c$  hinreichend erfolgreich lernen, damit die kooperativen Typen überleben können. Dieses Resultat ist intuitiv absolut plausibel und läßt sich auch formal leicht nachvollziehen. Für die komparative Statik sei aus Gründen der einfacheren Darstellbarkeit angenommen, daß die Subpopulationen „hinreichend groß“ seien, so daß näherungsweise gilt  $\frac{1}{n} \cong 0$ , sowie  $\frac{n+1}{n} \cong 1$ . Die kritischen Werte für den Parameter  $c$  vereinfachen sich dann zu

$$c_1 \equiv \frac{\lambda(1-k)((1-k)(\lambda-6)+3)}{(2+(1-k)(\lambda-6))}, \quad c_2 \equiv (1-k)(2\lambda-3), \quad c_3 \equiv \lambda(1-k)-1.$$

Exemplarisch sei der Parameter  $c_2$  untersucht. Die partielle Ableitung nach  $k$  ergibt natürlich gerade  $3-2\lambda$ . Damit die Restriktion  $c < c_2$  erfüllt werden kann, muß selbstverständlich gelten  $c_2 > 0 \Leftrightarrow \lambda > \frac{3}{2}$ . Für diese Werte der outside-option wird die partielle Ableitung nach  $k$  offensichtlich negativ. Eine Senkung des Parameters  $k$  führt also zu einer Erhöhung des kritischen Wertes  $c_2$ . Eine Abnahme von  $k$  ist aber gerade als eine Verbesserung der Lernbedingungen für die K-Typen zu interpretieren. Der Parameter  $k$  war gegeben als  $k \equiv \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right)$ . Betrachtet sei die partielle Ableitung nach  $n$ :

$$\frac{\partial k}{\partial n} = \frac{1}{T+1} \left[ 1 - \left( \frac{n-1}{n} \right)^T \left( \frac{n+T}{n} \right) \right] \quad (5-15)$$

Wir werden zeigen, daß dieser Ausdruck für alle  $T > 0$ ,  $n > 1$  positiv sein muß.

$$\begin{aligned} \frac{1}{T+1} \left[ 1 - \left( \frac{n-1}{n} \right)^T \left( \frac{n+T}{n} \right) \right] > 0 &\Leftrightarrow \\ \frac{n}{n+T} > \left( \frac{n-1}{n} \right)^T &\Leftrightarrow \end{aligned}$$

$$T > \frac{\ln\left(\frac{n+T}{n}\right)}{\ln\left(\frac{n}{n-1}\right)} \equiv \tilde{T} \quad (5-16)$$

Bezeichne im folgenden  $f(T) = T$  und  $g(T) = \frac{\ln\left(\frac{n+T}{n}\right)}{\ln\left(\frac{n}{n-1}\right)}$ . Zunächst bemerkt

man, daß beide Funktionen für  $T \rightarrow 0$  gegen den Ursprung konvergieren. Sodann zeigen wir, daß die Funktion  $g(T)$  immer eine Steigung kleiner eins aufweist und deshalb global unterhalb der Funktion  $f(T) = T$  verlaufen muß.

$$\begin{aligned} \frac{\partial \tilde{T}}{\partial T} &= \frac{1}{\ln\left(\frac{n}{n-1}\right)} \frac{1}{(n+T)} < 1 \Leftrightarrow \\ T &> \frac{1}{\ln\left(\frac{n}{n-1}\right)} - n \end{aligned} \quad (5-17)$$

Nun ist aber klar, daß die rechte Seite der Ungleichung (5-17) negativ ist, denn:

$$\begin{aligned} \frac{1}{\ln\left(\frac{n}{n-1}\right)} - n < 0 &\Leftrightarrow \\ n \ln\left(\frac{n}{n-1}\right) &> 1 \Leftrightarrow \\ \exp\left[n \ln\left(\frac{n}{n-1}\right)\right] &> e \Leftrightarrow \\ \left[\exp\left[\ln\left(\frac{n}{n-1}\right)\right]\right]^n &> e \Leftrightarrow \\ \left(\frac{n}{n-1}\right)^n &> e \Leftrightarrow \\ \left(1 + \frac{1}{n-1}\right)^n &> e \end{aligned} \quad (5-18)$$

Diese Bedingung ist aber für alle  $n > 1$  erfüllt, denn man macht sich schnell klar, daß die Folge, aus dem Unendlichen kommend, gegen die Eulersche Zahl konvergiert:

$$\lim_{n \rightarrow 1} \left(1 + \frac{1}{n-1}\right)^n = +\infty \quad (5-19)$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n-1}\right)^n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \quad (5-20)$$

$$\frac{\partial}{\partial n} \left(1 + \frac{1}{n-1}\right)^n = n \left(1 + \frac{1}{n-1}\right)^{n-1} \left(-\frac{1}{(n-1)^2}\right) < 0 \quad (5-21)$$

Eine Abnahme der Populationsgröße bei gegebener Lebensdauer (Lerndauer) einer Generation  $T+1$  äußert sich also in sinkendem  $k$ . In diesem Sinne signalisiert ein kleineres  $k$  also bessere Lernbedingungen für den  $K$ -Typ. Und eine Abnahme von  $k$  zieht eine steigende kritische Grenze  $c_2$  nach sich. Die Bedingung für globale Stabilität des inneren Gleichgewichts wird also bei sinkendem  $k$  weniger restriktiv, d.h. der Parameterbereich, für den globale Stabilität gilt, vergrößert sich.

Völlig analog kann man zeigen, daß der Parameter  $k$  sinkt, wenn die Lerndauer  $T$  ceteris paribus zunimmt. Der Parameter  $k$  war gegeben als

$$k \equiv \frac{n}{T+1} \left(1 - \left(\frac{n-1}{n}\right)^{T+1}\right). \text{ Wir wollen die partielle Ableitung } k \text{ nach } T \text{ bestimmen.}$$

Betrachtet sei als erstes der Term  $\left(\frac{n-1}{n}\right)^{T+1}$ . Nach der Standard-

Ableitungsregel  $\frac{\partial f(x)}{\partial x} = a^x \ln(a)$  für  $f(x) = a^x$  bestimmt sich die Ableitung des

Polynoms nach  $T$  als:

$$\frac{\partial}{\partial T} \left(\frac{n-1}{n}\right)^{T+1} = \left(\frac{n-1}{n}\right)^{T+1} \ln\left(\frac{n-1}{n}\right) \quad (5-22)$$

Die gesamte partielle Ableitung des Parameters nach  $T$  erhält man als

$$\frac{\partial k}{\partial T} = -\frac{n}{(T+1)^2} + \frac{n}{(T+1)^2} \left(\frac{n-1}{n}\right)^{T+1} - \frac{n}{(T+1)} \left(\frac{n-1}{n}\right)^{T+1} \ln\left(\frac{n-1}{n}\right). \quad (5-23)$$

Wir werden nun zeigen, daß dieser Ausdruck für alle  $n, T > 1$  negativ ist.

$$\begin{aligned}
 & -\frac{n}{(T+1)^2} + \frac{n}{(T+1)^2} \left(\frac{n-1}{n}\right)^{T+1} - \frac{n}{(T+1)} \left(\frac{n-1}{n}\right)^{T+1} \ln\left(\frac{n-1}{n}\right) < 0 \Leftrightarrow \\
 & -\left(\frac{n-1}{n}\right)^{T+1} \left[ \frac{n}{(T+1)} \ln\left(\frac{n-1}{n}\right) - \frac{n}{(T+1)^2} \right] - \frac{n}{(T+1)^2} < 0 \Leftrightarrow \\
 & -\left(\frac{n-1}{n}\right)^{T+1} \left[ (T+1) \ln\left(\frac{n-1}{n}\right) - 1 \right] < 1 \Leftrightarrow \\
 & \left[ 1 - (T+1) \ln\left(\frac{n-1}{n}\right) \right] < \left(\frac{n}{n-1}\right)^{T+1} \tag{5-24}
 \end{aligned}$$

Wir müssen wiederum ein indirektes Argument bemühen, um zu zeigen, daß (5-24) erfüllt ist. Bezeichne im folgenden  $f(T) = \left[ 1 - (T+1) \ln\left(\frac{n-1}{n}\right) \right]$  sowie

$g(T) = \left(\frac{n}{n-1}\right)^{T+1}$ . Man ermittelt zunächst leicht

$$\frac{\partial f(T)}{\partial T} = \ln\left(\frac{n}{n-1}\right) > 0 \forall n > 1, \tag{5-25}$$

sowie

$$\frac{\partial g(T)}{\partial T} = \ln\left(\frac{n}{n-1}\right) \left(\frac{n}{n-1}\right)^{T+1} > 0 \forall n > 1. \tag{5-26}$$

Es gilt, daß die Funktion  $g(T)$  global die größere Steigung aufweist. Dies zeigt man leicht:

$$\begin{aligned}
 & \frac{\partial g(T)}{\partial T} > \frac{\partial f(T)}{\partial T} \Leftrightarrow \\
 & \ln\left(\frac{n}{n-1}\right) \left(\frac{n}{n-1}\right)^{T+1} > \ln\left(\frac{n}{n-1}\right) \Leftrightarrow \\
 & (T+1) \ln\left(\frac{n}{n-1}\right) > 0 \tag{5-27}
 \end{aligned}$$

was offensichtlich für alle  $n > 1$  der Fall ist. Man macht sich ferner schnell klar, daß

$$\lim_{T \rightarrow \infty} f(T) = \lim_{T \rightarrow \infty} g(T) = +\infty, \tag{5-28}$$

$$\lim_{T \rightarrow 0} f(T) = 1 + \ln\left(\frac{n}{n-1}\right), \quad \lim_{T \rightarrow 0} g(T) = \frac{n}{n-1}. \tag{5-29}$$

Wir zeigen nun:

$$\begin{aligned} \lim_{T \rightarrow 0} f(T) < \lim_{T \rightarrow 0} g(T) &\Leftrightarrow \\ 1 + \ln\left(\frac{n}{n-1}\right) < \frac{n}{n-1} \end{aligned} \quad (5-30)$$

(5-30) aber ist für alle  $n > 1$  erfüllt, wie man sich leicht klar macht:

$$\lim_{n \rightarrow 1} \left(\frac{n}{n-1}\right) = \lim_{n \rightarrow 1} 1 + \ln\left(\frac{n}{n-1}\right) = +\infty \quad (5-31)$$

$$\lim_{n \rightarrow \infty} \left(\frac{n}{n-1}\right) = \lim_{n \rightarrow \infty} 1 + \ln\left(\frac{n}{n-1}\right) = 1 \quad (5-32)$$

$$\frac{\partial}{\partial n} \left(\frac{n}{n-1}\right) = -\frac{1}{(n-1)^2} < 0 \quad (5-33)$$

$$\frac{\partial}{\partial n} \left(1 + \ln\left(\frac{n}{n-1}\right)\right) = -\frac{(n-1)}{n} \frac{1}{(n-1)^2} < 0 \quad (5-34)$$

$$\frac{\partial^2}{\partial n^2} \left(\frac{n}{n-1}\right) = -\frac{1}{(n-1)^3} < 0 \quad (5-35)$$

$$\frac{\partial^2}{\partial n^2} \left(1 + \ln\left(\frac{n}{n-1}\right)\right) = -\frac{2n+1}{(n(n-1))^2} < 0 \quad (5-36)$$

Beide Funktionen sind konkav und konvergieren aus dem Unendlichen kommend gegen den Wert eins. Man sieht jedoch leicht, daß die Funktion  $f(T)$  global die größere Steigung aufweist:

$$\begin{aligned} \frac{\partial}{\partial n} \left(1 + \ln\left(\frac{n}{n-1}\right)\right) &> \frac{\partial}{\partial n} \left(\frac{n}{n-1}\right) \Leftrightarrow \\ -\frac{(n-1)}{n} \frac{1}{(n-1)^2} &> -\frac{1}{(n-1)^2} \Leftrightarrow \\ \frac{n-1}{n} &< 1 \end{aligned} \quad (5-37)$$

(5-37) ist klarerweise für alle  $n > 1$  erfüllt. Damit ist gezeigt, daß (5-30) für alle  $n > 1$  gilt. (5-30) und (5-27) zusammen genommen ergeben damit, daß die partielle Ableitung des Parameters  $k$  für alle  $n > 1, T > 0$  negativ ist. Eine Erhöhung der Lebensdauer einer Generation führt bei gegebener Subpopulationsgröße  $n$  somit zu einer Abnahme von  $k$ . Sinkendes  $k$  signalisiert also wiederum eine Verbesserung der Lernbedingungen des K-Typs. Und abnehmendes  $k$  impliziert eine Verschiebung der kritischen Grenze  $c_2$  nach oben. Die Bedingung für globale Stabilität des inneren Nash-Gleichgewichtes wird wiederum weniger restriktiv. Zusammenfassend kann man also festhalten, daß das Modell auf Parametervariation absolut

plausibel reagiert. Je besser die Lernbedingungen für die K-Mutante *ceteris paribus* sind (je kleiner  $n$  bei gegebenem  $T$  oder je größer  $T$  bei gegebenem  $n$ ), desto größer ist die Parameterregion, für die ein global stabiles inneres Gleichgewicht existiert.

Es sei abschließend zur komparativen Statik angemerkt, daß die simplifizierende Annahme eines hinreichend großen  $n$  tatsächlich gerechtfertigt ist. Betrachtet sei

nochmals der Ausdruck  $k \equiv \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right)$ . Man könnte der Auffassung

sein, die Annahme „ $n$  groß“ führe zu trivialen Resultaten des Lernparameters  $k$ , da  $\left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right) = \left( 1 - \left( 1 - \frac{1}{n} \right)^{T+1} \right)$  gegen Null zu streben scheint. Dies ist jedoch

nicht richtig. Auch für große  $n$  gilt, daß der Parameter  $k$  für „vernünftige“ Werte von  $T$  nicht triviale Werte annimmt. Die folgende Tabelle mag dies verdeutlichen.

T	k(n,T), n=5000
500	0,952
1000	0,906
2000	0,824
3000	0,752
4000	0,688
5000	0,632
10000	0,432

Tabelle 10: Numerische Beispiele für den Parameter  $k$

Es war bereits angemerkt worden, daß das obige Modell eine Verallgemeinerung eines Modells von Amann und Yang (1994) darstellt. Amann/Yang führen ebenfalls eine lernfähige Mutante in das GD ein. Dieser Typ ist im Unterschied zu unserem K-Typ jedoch perfekt informiert. Man sollte daher erwarten, daß sich die Amann Yang-Resultate als Spezialfall unseres Modells reproduzieren lassen. Dies ist in der Tat der Fall. Perfekte Information ergibt sich in unserem Modell, wenn die Lerndauer für gegebene Subpopulationsgröße  $n$  gegen unendlich strebt. Formal:

$$\lim_{T \rightarrow \infty} \left( \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right) \right) = \lim_{T \rightarrow \infty} \frac{n}{T+1} - \lim_{T \rightarrow \infty} \frac{n}{T+1} \left( \frac{n-1}{n} \right)^{T+1} = 0 \quad (5-38)$$

Setzt man nun  $k=0$  ein, so erhält man tatsächlich alle Resultate aus Amann/Yang (1994).



Es konnte somit gezeigt werden, daß kooperatives Verhalten auch im Fall imperfekter Information über die Zusammensetzung der Population überlebensfähig ist. Die Stützung von Kooperation ist also nicht davon abhängig, daß der Typ eines Gegenspielers immer zuverlässig identifiziert werden kann. Es genügt, wenn dies in einer hinreichend großen Zahl von Fällen gelingt. Dies erfordert, daß in der Vergangenheit hinreichend viele Defektoren gelernt wurden. Insbesondere benötigen wir für unsere Resultate nicht die ad hoc Annahme von Signalen, die – seien sie nun perfekter oder imperfekter Natur – Aufschluß über den Typus eines Gegenspielers geben. Die Diskussion des Frank-Modells hatte ergeben, daß die Annahme von Signalen in theoretischer wie auch in empirischer Hinsicht fragwürdig ist. Der in dieser Arbeit entwickelte lerntheoretische Ansatz vermeidet von vornherein diese Schwierigkeiten. Es sei an dieser Stelle nochmals auf die Harrington-Kritik aufmerksam gemacht. Das Frank-Modell benötigt einen Signalwert-Bereich, in dem anhand eines Signals mit Wahrscheinlichkeit eins der Typ eines Spielers erkannt wird. Innerhalb des Frankschen Ansatzes erscheint die Annahme eines solchen Bereichs problematisch. Insbesondere die empirische psychologische Literatur legt nahe, daß keineswegs von perfekten Signalen ausgegangen werden kann. Gewissermaßen liefert nun unser lerntheoretischer Ansatz eine Rechtfertigung für die Existenz „perfekter Information“, wobei perfekt hier geeignet zu interpretieren ist. Bezüglich eines einmal getroffenen Spielers ist dessen Typinformation für einen K-Typ natürlich bekannt, und zwar mit Sicherheit. Wann immer genau dieser eine Spieler erneut getroffen wird, so ist bekannt, welchem Typus er angehört. In diesem Sinne ist die durch das Lernmodell generierte Information durchaus perfekt. Global betrachtet, das heißt also bezüglich der Zusammensetzung der Gesamtpopulation, ist die generierte Information im allgemeinen natürlich imperfekt, es sei denn, man kapriziert sich auf die oben erwähnten Spezialfälle, in denen das Lernmodell gegen perfekte Information konvergiert. Der lerntheoretische Zugang ermöglicht also in natürlicher Weise die Koexistenz von lokal perfekter und global imperfekter Information: Bekannt ist im allgemeinen immer nur ein Teil der übrigen Population, aber dieser bekannte Anteil kann perfekt identifiziert werden. Der Signalansatz ist hingegen gezwungen, einen Bereich eines perfekt funktionierenden Signals zu postulieren – wenigstens implizit – ohne diese Annahme jedoch recht plausibel machen zu können.

Erwähnenswert ist sicherlich auch die Tatsache, daß die kognitiven Anforderungen bezüglich des Lernprozesses sich sehr in Grenzen halten. Das hier implementierte Lernen ist von äußerst simpler Art, insofern es nämlich lediglich Information der Vergangenheit nutzt. Man benötigt also insbesondere keine antizipatorischen Elemente, wie es etwa in den bekannten bayesianischen Lerntheorien der Fall ist. Es ist völlig ausreichend, daß die kooperative Mutante ihr Verhalten auf

die gewonnenen Erfahrungen konditioniert. Diese Form des Erfahrungslernens setzt lediglich die Existenz eines Gedächtnisses voraus. Man kann daher davon ausgehen, daß derartige Lernprozesse bereits in sehr frühen Stadien der Entwicklungsgeschichte der Menschheit präsent gewesen sein dürften. Dies ist deshalb von Bedeutung, weil der Erfolg konditionierend-kooperativen Verhaltens sich am ehesten in kleinen und übersichtlichen Gesellschaftsverbänden einstellen wird. Wir hatten in der komparativen Statik des Modells ausführlich dargelegt, daß ein inneres Gleichgewicht mit der Koexistenz aller drei Verhaltenstypen umso eher zu erwarten ist, je erfolgreicher der K-Typ lernt. Für eine gegebene Interaktionshäufigkeit ist dies aber gerade für eine abnehmende Gruppengröße der Fall. Die Tatsache, daß die Menschheit sich zunächst in überschaubaren Verbänden organisierte – Horden, Stämmen und Sippen – dürfte die Entstehung von Kooperation also begünstigt haben.

Dies soll jedoch nicht nahelegen, das hier vorgestellte Modell sei ausschließlich auf den Fall kleiner und kleinster Gruppen anwendbar. Der Lernerfolg der kooperativen Mutante – reflektiert durch den Wert des Parameters  $k$  – hängt von zwei Variablen ab, der Gruppengröße sowie der Interaktionshäufigkeit. Eine Zunahme der Gruppengröße kann also durch eine entsprechende Zunahme der Interaktionshäufigkeit kompensiert werden (vgl. hierzu auch Tabelle 7). Die bloße Tatsache der bei weitem größeren Bevölkerung heutiger Industriegesellschaften gegenüber vormodernen Gesellschaftsformationen stellt auch aus einem weiteren Grund a priori kein Argument für eine Abnahme von Kooperation dar. Trotz der unbestreitbar größeren Komplexität moderner Gesellschaften interagieren dennoch die meisten Menschen in mehr oder weniger konstanten Zusammenhängen. Es gibt, auch für die große Mehrheit heutiger Menschen, immer noch relativ stabile Arbeitsbeziehungen und auch für Freundschaften und Nachbarschaften dürfte eine gewisse Konstanz angenommen werden. Relevant für den Erfolg kooperativen Verhaltens ist also nicht die absolute, globale Populationsgröße sondern die lokale, also die Größe der relevanten Bezugsgruppe eines Agenten. Unser Modell versucht diesem Umstand durch die Annahme einer in zahlreiche Subgruppen differenzierten Gesellschaft gerade Rechnung zu tragen. Man kann also nicht a priori davon ausgehen, daß Kooperation in modernen Gesellschaften allein schon wegen deren absoluter Größe gefährdet ist. Allerdings macht die Modellstruktur auch folgendes deutlich: Eine Abnahme der Konstanz von Gruppenzugehörigkeiten kann sich als Problem erweisen. Etwa kann die Stabilität von Gruppenbeziehungen durch die Zunahme räumlicher Mobilität gefährdet werden. Wenn Menschen beispielsweise aus beruflichen Gründen gezwungen sind, häufig ihren Arbeitsplatz und/oder Wohnort zu wechseln, dann ist dies in der Regel mit einem Wechsel der Bezugsgruppen verbunden. Die Entstehung kooperativen Verhaltens

wird damit erschwert, denn die Verweildauer innerhalb einer Gruppe reduziert sich, so daß der Erfolg konditionierendes Verhaltens gemindert wird (im Modell reduziert sich die Lerndauer  $T+1$ , daher reduziert sich auch der Lernerfolg der konditionierend-kooperativen Typen, weshalb weniger häufig als zuvor defektives Verhalten bestraft werden kann). Insofern also moderne Gesellschaften ihren Mitgliedern zunehmend mehr Mobilität abverlangen, kann das Fundament für Kooperation, nämlich stabile soziale Beziehungen, unterminiert werden.

Abschließend läßt sich festhalten, daß unser Modell durchaus geeignet ist, das stilisierte Faktum einer gemischten Dreitypen-Population zu erklären. Die experimentellen Befunde waren nun nicht ganz unzweideutig bezüglich der Existenz eines bedingungslos kooperierenden Typs. Man muß konzedieren, daß echte Tauben in den experimentellen Populationen ein recht seltenes Phänomen sind. So beobachteten Andreoni und Miller (1993) als auch Cooper et al. (1996) jeweils lediglich eine strikt kooperative Versuchsperson. Dies könnte zu der Spekulation Anlaß geben, daß wir uns real in der Nähe eines gemischten Zweitypen-Gleichgewichts mit Defektoren und bedingt kooperativen Typen befinden. Langfristig sollte dann bedingungslos kooperatives Verhalten überhaupt nicht mehr beobachtet werden. Allerdings sei darauf hingewiesen, daß der Selektionsdruck in der Nähe von Gleichgewichten schwach wird und es sehr lange dauern kann, bis ein Gleichgewicht exakt erreicht wird. Ein eigentlich (im formalen Sinne) transitorischer Zustand mit einem sehr kleinen Anteil von Tauben kann sich daher als ausgesprochen persistent erweisen. Für die Erklärungskraft des Modells ist die Interpretation der experimentellen Befunde ohnehin zweitrangig, da beide Interpretationen – echt gemischtes Dreitypen-Gleichgewicht einerseits, Zweitypen-Gleichgewicht mit "transitorischem" marginalen Anteil von Tauben andererseits – mit dem Modell kompatibel sind.

In diesem Zusammenhang sei daran erinnert, daß das Modell von Robson (1990) nicht in der Lage ist, eine (wie auch immer) gemischte Population mit Defektoren und Kooperatoren zu erklären. Im Basismodell des secret handshake ist die mutante Strategie die einzig evolutionär stabile Strategie. Die Ursache hierfür liegt natürlich klar auf der Hand: Die Annahme eines perfekt funktionierenden Signals führt dazu, daß die mutanten Kooperatoren immer unter sich bleiben und Ausbeutung durch Defektoren perfekt vermeiden können. Im hier vorliegenden Modell muß Information über Mitspieler erst erzeugt werden und liegt nicht von vornherein schon vor. Um mit einem bekannten Sprichwort zu formulieren, wird die K-Mutante nur durch Erfahrung klug – dies impliziert aber vor allem das Sammeln schlechter Erfahrungen. Defektoren können als solche nur erkannt werden, wenn der K-Typ in einer vorherigen Interaktion ausgebeutet worden ist.

Dennoch, und darin ist vielleicht der eigentliche Erkenntniswert des vorgestellten Modells zu sehen, erweist sich Kooperation als ein robustes Phänomen.

Einige abschließende Bemerkungen zu sehr grundsätzlichen konzeptionellen Fragen sind schließlich erforderlich. Die obigen Resultate wurden abgeleitet unter Verwendung der Replikator Dynamik. Diese Dynamik ist eine Differentialgleichung, die in der Biologie verwendet wird, um die asexuelle Reproduktion von Genen in einer Population zu beschreiben. Es erhebt sich daher die Frage, inwiefern es berechtigt sein sollte, diese Funktion in einem sozialwissenschaftlichen Kontext anzuwenden. Das Problem hierbei ist nicht, wie Börgers (1996) treffend festgestellt hat, daß menschliches Verhalten nicht (auch) genetisch determiniert sei. Vielmehr kann beim gegenwärtigen Stand der verhaltensgenetischen Forschung niemand seriös die Frage beantworten, **wie** das Verhältnis zwischen genetischem Material und beobachtbarem Verhalten aussieht. Es scheint, daß Verhaltensweisen nicht von einzelnen Genen (mit)determiniert sind, sondern durch ganze Komplexe von Genen, die in komplizierter Weise mit den Umweltbedingungen interagieren:

*“In economics these concepts [die Replikator Dynamik, C.V.] can only be understood metaphorically, not literally. This is probably uncontroversial, but it might be worthwhile to recapitulate the argument for this. The argument is not that human behavior is essentially different from other animals’ behavior and that it is not genetically determined. Rather, the crucial point is that it is not practically feasible today to derive predictions of human behavior by appealing to its genetic determination. Scientific understanding of the relation between human genes and human behaviour seems in its beginnings. Moreover, the way in which genetic codes determine behaviour seems very complicated. Genetic codes seem to endow human beings with programmes which interact with the environment, and it is only this interaction which determines ultimately behaviour.”* (Börgers (1996), S.1382)

Eine metaphorische Interpretationsweise der Replikator Dynamik besteht darin, sie als Resultat sozialer oder individueller Lernprozesse aufzufassen. In den letzten Jahren ist einiger Fortschritt in dieser Hinsicht erzielt worden. Karl Schlag (1998) hat ein Modell sozialen Lernens vorgestellt, in dem unter bestimmten Bedingungen der angenommene Lernprozess gegen die Replikator Dynamik konvergiert. Spieler einer Population können Aktionen anderer Spieler sowie die resultierenden Payoffs beobachten und imitieren erfolgreichere Strategien. Genau genommen nimmt Schlag an, daß die Imitationswahrscheinlichkeit proportional ist zur Differenz zwischen dem eigenen und dem beobachteten Payoff. Man kann

dann zeigen, daß in einer wachsenden Population die Trajektorien dieses stochastischen Imitationsprozesses gegen die der stetigen Replikator Dynamik konvergieren. Börgers und Sarin (1997) haben demgegenüber ein Modell individuellen Lernens vorgestellt, in denen Akteure nur ihre eigenen Aktionen und Payoffs beobachten können. Ihr Modell zählt zur Klasse der reinforcement-Modelle. Erfolgreichere Strategien werden in der Zukunft mit größerer Wahrscheinlichkeit gespielt. Läßt man die Zeitintervalle zwischen den Entscheidungszeitpunkten arbiträr klein werden, so erhält man im Grenzübergang wiederum die Replikator Dynamik.

Es zeigt sich also, daß die Replikator Dynamik durchaus einer lerntheoretischen Deutung zugänglich ist. Trotz der zitierten Resultate bleibt jedoch kritisch festzuhalten, daß die angenommenen Lernregeln in beiden Fällen letztlich spezieller Natur sind und sich insofern die Frage erhebt, inwiefern unter allgemeineren Bedingungen die Replikator-Dynamik lerntheoretisch reinterpretiert werden kann. Bis zur abschließenden Klärung dieser Frage und der Verfügbarkeit einer möglicherweise überzeugenderen Dynamik scheint die Verwendung der Replikator Dynamik aber gerechtfertigt zu sein.

## Appendix zu Kapitel 5.3

### Beweis des Lemmas 5.1

Der Beweis ist keineswegs schwierig, jedoch etwas umfänglich. Dies hängt damit zusammen, daß zunächst einmal die erwarteten Payoffs für die drei Typen Hawk, Dove und K-Typ bestimmt werden müssen. Die Modellstruktur macht dieses Unterfangen etwas mühselig. Wir beginnen mit dem einfachsten Fall, dem erwarteten Payoff für Dove.

Was ist der erwartete Payoff für eine Taube bezüglich eines Tripels  $Q_j$ ? Die Gruppengröße ist  $n+1$ . Bezeichne  $(D_j, H_j, K_j)$  die absoluten Anzahlen von Doves, Hawks und K-typen innerhalb einer Gruppe aus  $Q_j$ . Die Taube erzielt drei, wenn sie auf einen anderen naiven Kooperationspartner oder auf einen K-Typen trifft, während ihr Payoff Null beträgt, wenn sie durch einen Defektor ausgebeutet wird. Der erwartete Payoff bezüglich  $Q_j$  ist dann gegeben als

$$E\pi_{Q_j}^D = 3 \left( \frac{D_j - 1}{n} + \frac{K_j}{n} \right). \quad (5-39)$$

Mit  $\frac{D_j}{n+1} \equiv x_j^D, \frac{K_j}{n+1} \equiv x_j^K, \frac{H_j}{n+1} \equiv x_j^H$  kann (5-39) auch geschrieben werden als

$$\begin{aligned} E\pi_{Q_j}^D &= 3 \frac{1}{n} (n+1) \left( \frac{D_j}{n+1} + \frac{K_j}{n+1} - \frac{1}{n+1} \right) \Leftrightarrow \\ E\pi_{Q_j}^D &= 3 \frac{n+1}{n} (x_j^D + x_j^K) - 3 \frac{1}{n}. \end{aligned} \quad (5-40)$$

Der gesamte erwartete Payoff ergibt sich durch Summation über alle Tripel  $Q = \{Q_1, \dots, Q_q\}$ :

$$E\pi^D = \sum_{j=1}^q E\pi_{Q_j}^D f_{Q_j} = 3 \frac{n+1}{n} \sum_{j=1}^q (x_j^D + x_j^K) f_{Q_j} - 3 \sum_{j=1}^q \frac{1}{n} f_{Q_j} = 3 \frac{n+1}{n} (x_D + x_K) - 3 \frac{1}{n}$$

Da immer gilt  $x_D + x_K = 1 - x_H$ , erhält man

$$E\pi^D = 3(1 - x_H) + 3 \frac{1}{n} (1 - x_H) - 3 \frac{1}{n} = 3(1 - x_H) - 3 \frac{1}{n} x_H. \quad (5-41)$$

Wenden wir uns nun dem K-Typ zu. Bezeichne  $\Pr(H_K^j(t+1))$  die Wahrscheinlichkeit, im Zeitpunkt  $t+1$  einen bekannten Hawk in einer Gruppe aus  $Q_j$  zu treffen. Klar ist, daß diese Wahrscheinlichkeit sich im Zeitablauf ändert und damit auch der erwartete Payoff. Was also zu berechnen ist, ist der zeitliche Durchschnitt aller erwarteten Payoffs. Bezeichne nun  $\bar{\Pr}(H_K^j(T+1))$  die durchschnittliche Wahrscheinlichkeit einen bekannten Falken zu treffen, wenn es insgesamt  $T+1$  matching-Zeitpunkte gibt. D.h.,

$$\bar{\Pr}(H_K^j(T+1)) = \frac{1}{T+1} \sum_{t=0}^T \Pr(H_K^j(t+1)).$$

Wir bestimmen zunächst den erwarteten Payoff des K-Typs für einen beliebigen Zeitpunkt  $t+1$ . Der K-Typ erhält 3, wenn er andere kooperative Typen trifft (naiv oder lernfähig), er erzielt den Wert der Outside-option  $\lambda$ , wenn er einen bekannten Defektor trifft, er erzielt Null, wenn er einen unbekannt Defektor trifft, und er hat in allen Fällen Kosten in Höhe von  $-c$  zu tragen. Insgesamt ist der erwartete Payoff bezüglich  $Q_j$  gegeben als

$$\begin{aligned} E\pi_{Q_j}^K(t+1) &= 3 \frac{1}{n} (D_j + K_j - 1) + \lambda \frac{1}{n} (H_K^j(t+1)) - c \Leftrightarrow \\ E\pi_{Q_j}^K(t+1) &= 3 \frac{n+1}{n} \left( \frac{D_j + K_j - 1}{n+1} \right) + \lambda \frac{n+1}{n} \frac{(H_K^j(t+1))}{n+1} - c \end{aligned} \quad (5-42)$$

Die relative Häufigkeit, mit der in einem Zeitpunkt  $t+1$  ein bekannter Hawk getroffen wird, ist aus dem Lernmodell bekannt. Wir können (5-42) daher formulieren als

$$E\pi_{Q_j}^K(t+1) = 3 \frac{n+1}{n} (x_j^D + x_j^K) - 3 \frac{1}{n} + \lambda \frac{n+1}{n} \Pr(H_K^j(t+1)) - c \quad (5-43)$$

Der gesamte erwartete Payoff in  $t+1$  ergibt sich wieder durch Summation über alle  $Q_j$ :

$$\begin{aligned} E\pi^K(t+1) &= \sum_{j=1}^q E\pi_{Q_j}^K f_{Q_j} = 3 \frac{n+1}{n} \sum_{j=1}^q \left( x_j^D + x_j^K - \frac{1}{n+1} \right) f_{Q_j} + \\ &\lambda \frac{n+1}{n} \sum_{j=1}^q \Pr(H_K^j(t+1)) f_{Q_j} - c \sum_{j=1}^q f_{Q_j} \end{aligned} \quad (5-44)$$

Der zweite Term bedarf eingehenderer Untersuchung. Zunächst ist festzustellen, daß das Lernmodell natürlich direkt auf Subpopulationen anwendbar ist. Die Wahrscheinlichkeit, in einer Gruppe aus  $Q_j$  einen bekannten Hawk zu treffen, ist gegeben als

$$\Pr(H_K^j(t+1)) = x_j^H \left( 1 - \left( \frac{n-1}{n} \right)^t \right). \quad (5-45)$$

Daher:

$$\begin{aligned} \sum_{j=1}^q \Pr(H_K^j(t+1)) f_{Q_j} &= \sum_{j=1}^q x_j^H \left( 1 - \left( \frac{n-1}{n} \right)^t \right) f_{Q_j} = \left( 1 - \left( \frac{n-1}{n} \right)^t \right) \sum_{j=1}^q x_j^H f_{Q_j} \\ &= x_H \left( 1 - \left( \frac{n-1}{n} \right)^t \right) = \Pr(H_K(t+1)) \end{aligned} \quad (5-46)$$

Die Summierung über alle Tripel  $Q_j$  ergibt also gerade wieder unser ursprüngliches Resultat des Lernmodells für den Fall einer Population, die nicht als disaggregiert in verschiedene Subpopulationen gedacht wird. Damit ist der erwartete Payoff bezüglich eines Zeitpunktes  $t+1$

$$E\pi^K(t+1) = 3 \frac{n+1}{n} \left( x_D + x_K - \frac{1}{n+1} \right) + \lambda \frac{n+1}{n} \Pr(H_K(t+1)) - c. \quad (5-47)$$

Wenden wir uns nunmehr der Ermittlung des durchschnittlichen erwarteten Payoffs des K-Typs zu, d.h. wir summieren die erwarteten Payoffs über alle betrachteten Zeitpunkte  $t=0\dots T$  und dividieren durch  $T+1$ :

$$\begin{aligned} \frac{1}{T+1} \sum_{t=0}^T E\pi^K(t+1) &= \\ 3 \frac{1}{T+1} \frac{n+1}{n} \sum_{t=0}^T \left( x_D + x_K - \frac{1}{n+1} \right) &+ \lambda \frac{n+1}{n} \frac{1}{T+1} \sum_{t=0}^T \Pr(H_K(t+1)) - \frac{1}{T+1} \sum_{t=0}^T c \end{aligned} \quad (5-48)$$

Wiederum verdient der zweite Summand Aufmerksamkeit:

$$\frac{1}{T+1} \sum_{t=0}^T \Pr(H_K(t+1)) = \frac{1}{T+1} \sum_{t=0}^T x_H \left( 1 - \left( \frac{n-1}{n} \right)^t \right) = x_H - x_H \frac{1}{T+1} \sum_{t=0}^T \left( \frac{n-1}{n} \right)^t$$



Betrachtet sei nun  $\sum_{t=0}^T \left(\frac{n-1}{n}\right)^t = 1 + \left(\frac{n-1}{n}\right) + \dots + \left(\frac{n-1}{n}\right)^T$ . Die Summe dieser geometrischen Reihe läßt sich leicht bestimmen als  $n \left(1 - \left(\frac{n-1}{n}\right)^{T+1}\right)$ . Damit ergibt sich:

$$\begin{aligned} \frac{1}{T+1} \sum_{t=0}^T E\pi^K(t+1) = \\ 3 \frac{n+1}{n} \left(x_D + x_K - \frac{1}{n+1}\right) + \lambda \frac{n+1}{n} x_H \left(1 - \frac{n}{T+1} \left(1 - \left(\frac{n-1}{n}\right)^{T+1}\right)\right) - c \end{aligned} \quad (5-49)$$

Mit  $\frac{n}{T+1} \left(1 - \left(\frac{n-1}{n}\right)^{T+1}\right) \equiv k(n, T) \equiv k$  können wir schreiben

$$\begin{aligned} E\pi^K &= 3 \frac{n+1}{n} \left(x_D + x_K - \frac{1}{n+1}\right) + \lambda \frac{n+1}{n} x_H (1-k) - c \\ &= 3 \frac{n+1}{n} (1-x_H) - 3 \frac{1}{n} + \lambda \frac{n+1}{n} x_H (1-k) - c. \end{aligned} \quad (5-50)$$

Schließlich seien die Hawks betrachtet. Ein Hawk erzielt einen Payoff von sechs, wenn er einen naiven Kooperationspartner trifft oder aber einen K-Typ, für den er noch unbekannt ist. Trifft ein Hawk einen K-Typ, der ihn kennt, so erhält er den Wert der outside option. Beim Zusammentreffen mit einem anderen Defektor erhält er die Auszahlung eins. Die Wahrscheinlichkeit, für einen K-Typen bekannt zu sein, ist gegeben als  $\frac{\Pr(H_K^j(t+1))}{x_j^H}$ . Entsprechend ist die Wahrscheinlichkeit, unbekannt zu sein

$$\frac{x_j^H - \Pr(H_K^j(t+1))}{x_j^H} = 1 - \frac{\Pr(H_K^j(t+1))}{x_j^H}.$$

Für einen bekannten Falken ergibt sich als erwarteter Payoff in t+1 bezüglich des Tripels  $Q_j$

$$E\pi_{H_K}^j(t+1) = \left(6 \frac{D_j}{n} + \frac{H_j - 1}{n} + \lambda \frac{K_j}{n}\right) \frac{\Pr(H_K^j(t+1))}{x_j^H}, \quad (5-51)$$

für einen unbekanntem Falken

$$E\pi_{H_U}^j(t+1) = \left(6 \frac{D_j}{n} + \frac{H_j - 1}{n} + 6 \frac{K_j}{n}\right) \left(1 - \frac{\Pr(H_K^j(t+1))}{x_j^H}\right). \quad (5-52)$$

Addiert man beide Ausdrücke, so erhält man

$$\begin{aligned} E\pi_H^j(t+1) &= 6\frac{1}{n}(D_j + K_j) + \frac{1}{n}(H_j - 1) + (\lambda - 6)\frac{1}{n}K_j \frac{\Pr(H_K^j(t+1))}{x_j^H} \\ &= 6\frac{n+1}{n}(x_j^D + x_j^K) + \frac{n+1}{n}\left(x_j^H - \frac{1}{n+1}\right) + (\lambda - 6)\frac{n+1}{n}x_j^K \frac{\Pr(H_K^j(t+1))}{x_j^H}. \end{aligned} \quad (5-53)$$

Summation über alle Tripel liefert

$$\begin{aligned} E\pi_H(t+1) &= 6\frac{n+1}{n}\sum_{j=1}^q(x_j^D + x_j^K)f_{Q_j} + \frac{n+1}{n}\sum_{j=1}^q x_j^H f_{Q_j} - \frac{1}{n}\sum_{j=1}^q f_{Q_j} + \\ &(\lambda - 6)\frac{n+1}{n}\sum_{j=1}^q x_j^K \frac{\Pr(H_K^j(t+1))}{x_j^H} f_{Q_j}. \end{aligned} \quad (5-54)$$

Wir betrachten den letzten Summanden:

$$\begin{aligned} \sum_{j=1}^q x_j^K \frac{\Pr(H_K^j(t+1))}{x_j^H} f_{Q_j} &= \sum_{j=1}^q x_j^K \frac{1}{x_j^H} x_j^H \left(1 - \left(\frac{n-1}{n}\right)^t\right) f_{Q_j} \\ &= \left(1 - \left(\frac{n-1}{n}\right)^t\right) \sum_{j=1}^q x_j^K f_{Q_j} = x_K \left(1 - \left(\frac{n-1}{n}\right)^t\right) \end{aligned} \quad (5-55)$$

Somit:

$$E\pi_H(t+1) = 6\frac{n+1}{n}(x_D + x_K) + \frac{n+1}{n}x_H - \frac{1}{n} + (\lambda - 6)\frac{n+1}{n}x_K \left(1 - \left(\frac{n-1}{n}\right)^t\right) \quad (5-56)$$

Wir bilden den zeitlichen Durchschnitt unter Berücksichtigung der Tatsache, daß  $x_D + x_K = 1 - x_H$ , und erhalten

$$E\pi_H = 6 - 5x_H + \frac{5}{n}(1 - x_H) + (\lambda - 6)(1 - k)\frac{n+1}{n}x_K. \quad (5-57)$$

Nach diesen etwas umfänglichen, aber unvermeidlichen Vorarbeiten können wir uns nun der Frage der Existenz von Nash-Gleichgewichten zuwenden. Klar ist, daß die drei monomorphen Populationszustände Fixpunkte unter der Replikator Dynamik sind. Zwei der drei Fixpunkte können jedoch aus naheliegenden Gründen keine NGG sein: In einer reinen Dove-Population ist Kooperation keine beste

Antwort und in einer reinen K-Typ-Population ist es immer besser, ein naiver Kooperationspartner zu sein, da dieser keine Kosten zu tragen hat.

Im inneren Fixpunkt muß natürlich die Gleichheit aller erwarteten Payoffs gewährleistet sein. Außerdem muß offensichtlich  $x_H^* + x_K^* < 1$  gelten. Gleichsetzen von (5-41) and (5-57) liefert unverzüglich den gleichgewichtigen Populationsanteil der Falken

$$x_H^* = \frac{n}{n+1} \frac{c}{\lambda(1-k)} \quad (5-58)$$

Aus (5-50) = (5-41) erhält man:

$$x_K^* = \frac{2x_H^* - 3}{(\lambda - 6)(1 - k)} - \frac{2}{(n+1)(\lambda - 6)(1 - k)} \quad (5-59)$$

Einsetzen von 5-58 in 5-59:

$$\begin{aligned} x_K^* &= \frac{2 \frac{n}{n+1} \frac{c}{\lambda(1-k)} - 3}{(\lambda - 6)(1 - k)} - \frac{2}{(n+1)(\lambda - 6)(1 - k)} \\ &= \frac{2cn - 3(n+1)\lambda(1-k)}{(n+1)\lambda(\lambda - 6)(1 - k)^2} - \frac{2\lambda(1-k)}{(n+1)\lambda(\lambda - 6)(1 - k)^2} \\ &= \frac{n(2c - 3\lambda(1-k)) - 5\lambda(1-k)}{(n+1)\lambda(\lambda - 6)(1 - k)^2} \end{aligned} \quad (5-60)$$

Damit es sich um einen echt inneren Zustand in  $\Delta_3$  handelt, müssen sich  $x_K^*$  and  $x_H^*$  zu weniger als eins addieren:

$$\begin{aligned} x_K^* + x_H^* &< 1 \Leftrightarrow \\ \frac{cn}{(n+1)\lambda(1-k)} + \frac{n(2c - 3\lambda(1-k)) - 5\lambda(1-k)}{(n+1)\lambda(1-k)^2(\lambda - 6)} &< 1 \Leftrightarrow \\ \frac{cn(1-k)(\lambda - 6) + n(2c - 3\lambda(1-k)) - 5\lambda(1-k)}{(n+1)\lambda(1-k)^2(\lambda - 6)} &< 1 \Leftrightarrow \\ c &< \frac{(n+1)\lambda(1-k)^2(\lambda - 6) + \lambda(1-k)(5 + 3n)}{n(2 + (1-k)(\lambda - 6))} \end{aligned} \quad (5-61)$$

Selbstverständlich muß außerdem gelten, daß die Populationsanteile für Hawks und K-Typen strikt positiv sind. Für Hawks ist dies offensichtlich der Fall. Für K-Typen ergibt sich:

$$\begin{aligned}
 x_K^* &= \frac{n(2c - 3\lambda(1-k)) - 5\lambda(1-k)}{(n+1)\lambda(1-k)^2(\lambda-6)} > 0 \\
 \Leftrightarrow 2cn - 3\lambda n(1-k) - 5\lambda(1-k) &< 0 \\
 \Leftrightarrow c < (1-k)(5+3n)\frac{\lambda}{2n} & \quad (5-62)
 \end{aligned}$$

Es läßt sich leicht zeigen, daß (5-61) immer restriktiver ist als (5-62). Falls also (5-61) erfüllt ist, existiert der innere Fixpunkt.

Wenden wir uns als nächstes dem Zwei-Typen-Fixpunkt zu. Die Payoffs für Hawks und K-Typen, ausgewertet an der Stelle  $x_H = 1 - x_K$  ergeben sich als

$$E\pi_H|_{x_H=1-x_K} = 1 + \frac{1}{n}x_K(5(1+n) + (n+1)(1-k)(\lambda-6)) \quad (5-63)$$

$$E\pi_K|_{x_H=1-x_K} = \lambda(1-k)\frac{n+1}{n} - \frac{3}{n} - c + x_K\frac{1}{n}(3(n+1) - \lambda(n+1)(1-k)) \quad (5-64)$$

Gleichsetzen:

$$\begin{aligned}
 E\pi_H|_{x_H=1-x_K} &= E\pi_K|_{x_H=1-x_K} \Leftrightarrow \\
 x_K^* &= \frac{\lambda(1-k) - (1+c)\frac{n}{n+1} - \frac{3}{n+1}}{((1-k)(2\lambda-6)+2)} \quad (5-65)
 \end{aligned}$$

Dieser Fixpunkt existiert offensichtlich genau dann, wenn  $0 < x_K^* < 1$ .

$$x_K^* > 0 \Leftrightarrow \frac{\lambda(1-k) - (1+c)\frac{n}{n+1} - \frac{3}{n+1}}{((1-k)(2\lambda-6)+2)} > 0$$

Eine Fallunterscheidung ist erforderlich:

$$\text{i) } ((1-k)(2\lambda-6)+2) > 0$$

$$\begin{aligned}
 x_K^* > 0 &\Leftrightarrow \\
 c < \frac{n+1}{n}\lambda(1-k) - \frac{3+n}{n} & \quad (5-66)
 \end{aligned}$$

$$\begin{aligned}
 x_K^* < 1 &\Leftrightarrow \\
 c > \frac{n+1}{n}[(1-k)(6-\lambda) - 2] - \frac{3+n}{n} & \quad (5-67)
 \end{aligned}$$

$$\text{ii) } ((1-k)(2\lambda-6)+2) < 0$$

$$\begin{aligned} x_K^* > 0 &\Leftrightarrow \\ c > \frac{n+1}{n} \lambda(1-k) - \frac{3+n}{n} \end{aligned} \quad (5-68)$$

$$\begin{aligned} x_K^* < 1 &\Leftrightarrow \\ c < \frac{n+1}{n} [(1-k)(6-\lambda)-2] - \frac{3+n}{n} \end{aligned} \quad (5-69)$$

Als nächstes soll die Frage geklärt werden, ob der Zwei-Typen-Fixpunkt ein NGG sein kann. Zur Beantwortung dieser Frage muß geklärt werden, ob Hawk, bzw. die K-Strategie in einer Zwei-Typen-Population mit den gleichgewichtigen Populationsanteilen beste Antwort sind, oder aber die Strategie Dove in dieser Population einen höheren Payoff erzielen kann. Wir werten die Payoffs für Hawk und Dove im Zwei-Typen-Fixpunkt aus:

$$E\pi_H|_{x_K^*} = 1 + \frac{\lambda(1-k) - (1+c) \frac{n}{n+1} - \frac{3}{n+1}}{n((1-k)(2\lambda-6)+2)} (5(1+n) + (n+1)(1-k)(\lambda-6)) \quad (5-70)$$

$$E\pi_D|_{x_K^*} = 3 \frac{n+1}{n} \frac{\lambda(1-k) - (1+c) \frac{n}{n+1} - \frac{3}{n+1}}{n((1-k)(2\lambda-6)+2)} - 3 \frac{1}{n} \quad (5-71)$$

Ein Vergleich von (5-70) und (5-71) liefert nach einigen Umformungen:

$$\text{i) } ((1-k)(2\lambda-6)+2) > 0$$

$$\begin{aligned} E\pi_H|_{x_K^*} > E\pi_D|_{x_K^*} &\Leftrightarrow \\ c > \frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5] - 2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} \end{aligned} \quad (5-72)$$

$$\text{ii) } ((1-k)(2\lambda-6)+2) < 0$$

$$\begin{aligned} E\pi_H|_{x_K^*} > E\pi_D|_{x_K^*} &\Leftrightarrow \\ c < \frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5] - 2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} \end{aligned} \quad (5-73)$$

Der Zwei-Typen Fixpunkt existiert also und ist ein NGG, falls entweder (5-66) und (5-72) erfüllt sind, d.h.

$$\frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5]-2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]} < c < \frac{n+1}{n}\lambda(1-k) - \frac{3+n}{n},$$

oder aber (5-67) und (5-73), d.h.

$$\frac{n+1}{n}[(1-k)(6-\lambda)-2] - \frac{3+n}{n} < c < \frac{(n+1)\lambda(1-k)[(1-k)(\lambda-6)+5]-2\lambda n(1-k)}{n[(1-k)(\lambda-6)+2]}.$$

Wenden wir uns schließlich der monomorphen Ecklösung mit Hawks allein zu. Payoffs ausgewertet in (1,0,0) resultieren als 1 für Hawks, 0 für Doves und  $\lambda(1-k)\frac{n+1}{n} - \frac{3}{n} - c$  für K-Typen. Hawk ist also beste Antwort auf sich selbst, solange gilt:

$$\begin{aligned} E\pi_H|_{(1,0,0)} = 1 > \lambda(1-k)\frac{n+1}{n} - \frac{3}{n} - c = E\pi_K|_{(1,0,0)} \\ \Leftrightarrow c > \lambda(1-k)\frac{n+1}{n} - \frac{n+3}{n} \end{aligned} \quad (5-74)$$

Dies vervollständigt den Beweis von Lemma 5.1.

□

## 5.5 Eine alternative Variante des Basismodells

Das in den vorherigen Abschnitten dargestellte Modell macht von der Fiktion einer unendlichen Anzahl finiter Subpopulationen Gebrauch. Diese Fiktion war erforderlich geworden, um Lernmodell und evolutionäre Analyse miteinander in Einklang zu bringen. Die evolutionäre Spieltheorie operiert meistens mit der Annahme einer formal unendlichen Gesamtpopulation, um mit theoretischen Erwartungswerten rechnen zu können. Ohne den „Kunstgriff“ einer Zergliederung der Gesamtpopulation in finite Untereinheiten würde Lernen natürlich keine Effekte zeitigen.<sup>14</sup>

Wie schon früher in den konzeptionellen Vorüberlegungen angedeutet wurde, besteht ein weiterer Ausweg aus dem skizzierten Konflikt zwischen Lernmodell und dynamischer Analyse. Dieser Ausweg bedient sich wiederum eines Kunstgriffs, nämlich dem eines simultanen Grenzübergangs für die Größe der Gesamtpopulation sowie der Anzahl der Ziehungen aus der Population (Anzahl der matchings). Läßt man beide Modellgrößen simultan gegen Unendlich streben, so generiert das Lernmodell im Grenzübergang ebenfalls nicht triviale Effekte in einer formal tatsächlich unendlichen Population. Dieser Ansatz zeichnet sich durch eine gewisse formale Eleganz aus und generiert handlichere Terme, sieht sich allerdings gewisser Kritik ausgesetzt. Bevor wir uns jedoch der Diskussion dieser Technik widmen, soll sie zunächst dargestellt werden.

Bezeichne  $n$  nunmehr die Größe der Gesamtpopulation. Wir gehen von einer einheitlichen, also nicht weiter in Subpopulationen unterteilten Population aus. Interaktionen finden wie gehabt im random-matching-Verfahren statt. Bezeichne  $t$  einen solchen matching Zeitpunkt. Wir sind dann interessiert an der Frage, welche Resultate das Lernmodell generiert, wenn Populationsgröße  $n$  und Matching-Anzahl  $T$  gegen Unendlich streben. Um die Analyse nicht zu schwierig zu gestalten, nehmen wir an, daß beide Größen in einem festen Verhältnis streben. Entweder soll gelten

$$\Delta t = \gamma \Delta n \text{ oder } \Delta n = \gamma \Delta t, \gamma \in \mathbb{N}^+ . \quad (5-75)$$

Oder in einer äquivalenten Formulierung:

---

<sup>14</sup> Ein Nachteil dieser Modellierung besteht außerdem darin, daß die resultierenden algebraischen Ausdrücke recht unhandlich und unübersichtlich werden. Aus diesem Grunde war für die komparative Statik bereits von approximativen Werten der Terme ausgegangen worden.

$$\Delta t = \gamma \Delta n, \gamma = \begin{cases} 1/\alpha, & \text{falls } \Delta t < \Delta n \\ \alpha, & \text{falls } \Delta t \geq \Delta n \end{cases}, \alpha \in \mathbb{N}^+ \quad (5-76)$$

Ist der Zusammenhang zwischen den Änderungen von  $t$  und  $n$  durch die obige Gleichung definiert, so impliziert dies, daß die Beziehung zwischen  $t$  und  $n$  durch folgende Differenzengleichung gegeben ist:

$$t = t_0 + \gamma(n - n_0). \quad (5-77)$$

wobei  $t_0, n_0$  Anfangsbedingungen kennzeichnen. Die Anzahl der matchings gegen Unendlich streben zu lassen, hat nun zwei mögliche Implikationen für die dynamische Analyse. Einerseits kann  $t \rightarrow \infty$  bedeuten, daß auch die Lebensdauer einer Generation unendlich wird. Eine solche Implikation wird man als unrealistisch verwerfen wollen. Andererseits kann man an der Endlichkeit der Lebensdauer der Akteure festhalten. Die Annahme  $t \rightarrow \infty$  impliziert dann, daß es zu einem Kontinuum von matchings auf einem endlichen Lebenszeitintervall kommt.

Es ist vielleicht angebracht, hier einem möglichen Mißverständnis vorzubeugen: Die Vorstellung  $t \rightarrow \infty$  impliziert sicherlich, daß der Abstand zwischen zwei Matching-Zeitpunkten arbiträr klein wird. Dies könnte die Vermutung nahelegen,  $t$  werde nun zu einer reellwertigen Variablen. Dies ist aber nicht der Fall. Es muß hier sorgfältig unterschieden werden zwischen zwei verschiedenen Zeitbegriffen. Zum einen verwenden wir die Variable  $t$ , um die Anzahl der matchings zu messen.  $t$  hat hier also den Charakter einer Indexvariablen und dient lediglich dazu, verschiedene Ziehungszeitpunkte zu numerieren. Zum anderen existiert eine reale Lebenszeit der Akteure des Modells, wenn man so will die "echte", physikalische Zeit. Sei diese Realzeit mit  $\tau$  bezeichnet und sei diese Zeitdauer endlich. Läßt man nun  $t$  gegen unendlich streben und möchte man diese unendliche Anzahl von matchings auf ein endliches Realzeit-Intervall verteilen, so muß notwendigerweise der Abstand zwischen zwei matching-Zeitpunkten  $t$  und  $t+1$  marginal werden, aber dieser Abstand wird natürlich in der Realzeit  $\tau$  gemessen und nicht in der Indexzeit  $t$ .  $\tau$  ist notwendig eine reellwertige Variable,  $t$  jedoch kann völlig unproblematisch weiterhin als positiv ganzzahlig angenommen werden.

Nach dieser vorbeugenden Bemerkung wenden wir uns nun der Frage zu, welche Konvergenzeigenschaften das Lernmodell im simultanen Grenzübergang zeigt. Das zentrale Ergebnis wird im folgenden Lemma zusammengefaßt.

### **Lemma 5.3 (Grenzeigenschaft des Lernmodells):**



Sei  $\Delta t = \gamma \Delta n$ ,  $\gamma = \begin{cases} 1/\alpha, & \text{falls } \Delta t < \Delta n \\ \alpha & \Delta t \geq \Delta n \end{cases}$ ,  $\alpha \in \mathbb{N}_+$ .

Dann gilt:  $\lim_{n,t \rightarrow \infty} \Pr(H_K(t)) = x_H \left( 1 - \frac{1}{e^\gamma} \right)$ .

**Beweis:**

Unter Ausnutzung von (5-77) können wir zunächst schreiben:

$$\lim_{n,t \rightarrow \infty} \Pr(H_K(t+1)) = \lim_{n \rightarrow \infty} x_H \left( 1 - \left( \frac{n-1}{n} \right)^{t_0 + \gamma(n-n_0)} \right) \quad (5-78)$$

Es gilt

$$\lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^{t_0 + \gamma(n-n_0)} = \lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^{t_0} \left( \frac{n-1}{n} \right)^m \left( \frac{n-1}{n} \right)^{-m_0} = \lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^m, \quad (5-79)$$

da die Exponenten des ersten und dritten Binoms nicht von n abhängen.

Sei nun  $h \equiv \frac{1}{n}$ . Dann:  $\lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^m = \lim_{h \rightarrow 0} (1-h)^{\frac{1}{h}}$ .

Wir betrachten dann den Differenzenquotienten des natürlichen Logarithmus an der Stelle  $1/\gamma$ . Dieser kann definiert werden als

$$\frac{\ln\left(\frac{1}{\gamma}\right) - \ln\left(\frac{1}{\gamma} - \frac{1}{\gamma}h\right)}{\frac{1}{\gamma}h} = \frac{\ln\left(\frac{1/\gamma}{1 - h}\right)}{\frac{1}{\gamma}h} = \frac{\ln(1) - \ln(1-h)}{\frac{1}{\gamma}h} = -\ln\left((1-h)^{\frac{1}{h}}\right). \quad (5-80)$$

Wir betrachten nun den Limes des Differenzenquotienten. Dieser definiert die erste Ableitung des logarithmus naturalis. Daher:

$$\begin{aligned} \lim_{h \rightarrow 0} -\ln\left((1-h)^{\frac{1}{h}}\right) &\equiv \ln(x)'|_{x=1/\gamma} = \gamma \Leftrightarrow \\ \lim_{h \rightarrow 0} \ln\left((1-h)^{\frac{1}{h}}\right) &= -\gamma \end{aligned} \quad (5-81)$$

(5-81) impliziert natürlich

$$\lim_{h \rightarrow 0} (1-h)^{\frac{1}{h}} = e^{-1} = \frac{1}{e}. \quad (5-82)$$

□

Für die Berechnung der erwarteten Payoffs ist es erforderlich, auch den Grenzwert der durchschnittlichen Wahrscheinlichkeit des Lernmodells zu kennen. Diesen gibt das folgende Korollar an

:

**Korollar 5.1** (Grenzwert der durchschnittlichen Wahrscheinlichkeit):

Der Grenzwert  $\lim_{n, T \rightarrow \infty} \bar{Pr}(H_K(T+1))$  berechnet sich als  $x_H \left( 1 - \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) \right)$ .

**Beweis:**

Die über  $T+1$  matching-Zeitpunkte durchschnittliche Wahrscheinlichkeit, in einem Zeitpunkt  $t$  einen bekannten Defektor zu treffen, war gegeben als

$$\bar{Pr}(H_K(T+1)) = \frac{1}{T+1} \sum_{t=1}^{T+1} \Pr(H_K(t)) = x_H \left( 1 - \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^T \right) \right). \quad (5-83)$$

Mit (5-77) reduziert sich das Problem auf

$$\lim_{n \rightarrow \infty} x_H \left( 1 - \frac{n}{t_0 + \gamma(n-n_0)} \left( 1 - \left( \frac{n-1}{n} \right)^{t_0 + \gamma(n-n_0)} \right) \right) = x_H \left( \left( 1 - \frac{1}{\gamma} \right) - \frac{1}{\gamma} \lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^{t_0 + \gamma(n-n_0)} \right). \quad (5-84)$$

Dieser Limes ist jedoch bereits aus Lemma 5.3 bekannt. Mit

$$\lim_{n \rightarrow \infty} \left( \frac{n-1}{n} \right)^{t_0 + \gamma(n-n_0)} = \frac{1}{e^\gamma} \text{ erhalten wir dann das gewünschte Resultat.}$$

□

Abkürzend schreiben wir im folgenden  $\frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) \equiv k(\gamma) \equiv k$ .

Wie bereits zuvor reflektiert der Parameter  $k$  die Lernbedingungen. Man kann sich leicht klarmachen, daß  $k$  wiederum Werte zwischen 0 und 1 annimmt. Die Grenzwertbetrachtung liefert

$$\lim_{\gamma \rightarrow \infty} k = \lim_{\gamma \rightarrow \infty} \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) = 0, \quad (5-85)$$

bzw.

$$\begin{aligned} \lim_{\gamma \rightarrow 0} k &= \lim_{\gamma \rightarrow 0} \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) = \lim_{\gamma \rightarrow 0} \exp \left[ \ln \left( \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) \right) \right] = \\ &= \lim_{\gamma \rightarrow 0} \exp \left[ \ln(1) - \ln(\gamma) + \ln(e^\gamma - 1) - \ln(e^\gamma) \right] = \\ &= \exp[0 + \infty - \infty - 0] = \exp(0) = 1. \end{aligned} \quad (5-86)$$

Beide Fälle lassen sich leicht interpretieren. Strebt die multiplikative Konstante  $\gamma$  gegen Unendlich, so bedeutet dies gerade perfektes Lernen: Die Änderung der Interaktionshäufigkeit  $\Delta t$  ist für jede beliebige Änderung der Populationsgröße ja gerade “unendlich mal” größer als  $\Delta n$ . Das Lernmodell muß im Grenzwert also gegen perfekte Information konvergieren. Im umgekehrten Fall ist zu jeder denkbaren Änderung der Interaktionshäufigkeit die Änderung der Populationsgröße unendlich mal größer, so daß der Anteil gelernter Akteure an der Population notwendigerweise gegen Null strebt: In diesem Grenzübergang wird also überhaupt keine Information generiert. Beide Fälle haben ihre direkten Entsprechungen im Basismodell. Der Parameter  $k$  war dort gerade gegeben als  $k = \frac{n}{T+1} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right)$ . Der Grenzwert Null ergab sich dort, wenn die

matching-Anzahl  $T$  für gegebene Größe der Subpopulation gegen Unendlich strebte. Demgegenüber erhielt man  $k \rightarrow 1$  gerade für eine ins Unendliche strebende Populationsgröße bei gegebener, endlicher Lerndauer. In einer dynamisierten Betrachtungsweise des Parameters  $k$  ergeben sich also keine wesentlich neuen Eigenschaften. Auch die allgemeine komparative Statik fällt erwartungsgemäß aus. Betrachtet sei die partielle Ableitung des Parameters  $k$  nach  $\gamma$ :

$$\frac{\partial k}{\partial \gamma} = -\gamma^{-2} - [-\gamma^{-2} e^{-\gamma} + \gamma^{-1} e^{-\gamma}] = -\frac{1}{\gamma^2} \left[ 1 - \frac{1}{e^\gamma} \right] + \frac{1}{\gamma} \frac{1}{e^\gamma} \quad (5-87)$$

Zu vermuten ist natürlich:

$$\begin{aligned} \frac{\partial k}{\partial \gamma} &< 0 \Leftrightarrow \\ -\frac{1}{\gamma^2} \left[ 1 - \frac{1}{e^\gamma} \right] + \frac{1}{\gamma} \frac{1}{e^\gamma} &< 0 \Leftrightarrow \end{aligned}$$

$$\gamma < e^\gamma - 1 \quad (5-88)$$

Die Ungleichung ist explizit nicht lösbar. Man macht sich jedoch schnell klar, daß für alle positiven Werte von  $\gamma$  die Ungleichung erfüllt sein muß. Die Funktionen  $f(\gamma) = e^\gamma - 1$  und  $g(\gamma) = \gamma$  weisen einen Schnittpunkt im Ursprung auf.  $f(\gamma)$  weist jedoch für alle  $\gamma > 0$  die größere Steigung auf. Es ist

$$\frac{\partial f(\gamma)}{\partial \gamma} = e^\gamma, \frac{\partial g(\gamma)}{\partial \gamma} = 1, \quad (5-89)$$

und klarerweise gilt

$$\begin{aligned} \frac{\partial f(\gamma)}{\partial \gamma} > \frac{\partial g(\gamma)}{\partial \gamma} &\Leftrightarrow \\ e^\gamma > 1 &\Leftrightarrow \\ \gamma > 0. & \end{aligned} \quad (5-90)$$

Man kann nun, wie im Basismodell, die bekannte evolutionäre Analyse des modifizierten GD vornehmen. D.h., man ermittelt die Fixpunkte der Replikatorodynamik, eruiert, ob ggf. auftretende Randfixpunkte Nash-Gleichgewichte sein können und untersucht schließlich die lokale asymptotische Stabilität der ermittelten Nash-Gleichgewichte. Wir verzichten an dieser Stelle darauf, die entsprechenden Berechnungen vorzuführen, da sich gegenüber der Analyse im Basismodell praktisch keinerlei Änderungen ergeben. Stattdessen konstatieren wir direkt das Ergebnis bezüglich globaler Stabilität im modifizierten Basismodell:

**Satz 5.1' (Globale Stabilität):**

- i) Für  $c < \min \left\{ \left( \frac{\lambda(1-k)[(\lambda-6)(1-k)+3]}{(\lambda-6)(1-k)+2} \right), (2\lambda-3)(1-k) \right\}$  ist das innere Nash-Gleichgewicht  $(x_H^*, x_K^*, x_D^*) = \left\{ \frac{c}{\lambda(1-k)}, \frac{2c-3\lambda(1-k)}{\lambda(\lambda-6)(1-k)^2}, 1-x_H^*-x_K^* \right\}$  global stabil.
- ii) Für  $\frac{\lambda(1-k)[(1-k)(\lambda-6)+3]}{(\lambda-6)(1-k)+2} < c < (2\lambda-3)(1-k)$  ist das Zwei-Typen-Rand-NGG  $(x_H^*, x_K^*, x_D^*) = \left\{ \frac{3-\lambda-k(6-\lambda)-c}{4-2\lambda-k(6-\lambda)}, \frac{1+c-\lambda(1-k)}{4-2\lambda-k(6-\lambda)}, 0 \right\}$  global stabil.
- iii) Für  $c > \lambda(1-k)-1$  ist das Defektoren-NGG  $(x_H^*, x_K^*, x_D^*) = \{1,0,0\}$  global stabil.

$$\text{Dabei ist } k = \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right).$$

**Beweis:** analog Beweis Satz 5.1

Wie man sieht, erhält man faktisch identische Resultate zu denen des Basismodells. Die algebraischen Terme unterscheiden sich im wesentlichen lediglich hinsichtlich einer multiplikativen Konstante vom Typ  $\frac{n-1}{n}$ . Das

Auftauchen dieser Konstante im Basismodell ist natürlich gerade darauf zurückzuführen, daß der jeweils unter Betrachtung stehende Akteur aus der strikt endlichen Subpopulation "heraus gerechnet" werden mußte. In einer formal unendlich großen Population wird der eigene Populationsanteil  $1/n$  natürlich marginal und kann daher vernachlässigt werden.

## 5.6 Interpretation und Diskussion der Ergebnisse

Wie bereits im Basismodell ergibt sich auch in dieser Modellvariante, daß die Bedingungen für das Überleben von Kooperation umso weniger restriktiv sind, je kleiner der Wert des Parameters  $k$  ist – für gegebenes  $c$  und gegebenes  $\lambda$ . Dieses formal identische Resultat ist im vorliegenden Modellkontext wie folgt zu interpretieren: In einer Population mit Bevölkerungswachstum muß die Interaktionshäufigkeit relativ zur Bevölkerungszunahme hinreichend stark wachsen. Formal ergab sich, daß  $k(\gamma)$  eine in  $\gamma$  abnehmende Funktion ist. Höhere Werte von  $\gamma$  implizieren kleinere Werte für  $k(\gamma)$  und erweitern so die Parameterbereiche für Gleichgewichte mit Beteiligung von kooperativen Typen. Der hier verwendete simultane Grenzübergang für  $n$  und  $t$  kann durchaus kritisiert werden. Will man der Konsequenz ausweichen, wegen  $t \rightarrow \infty$  auch die Lebenszeit der Akteure ins Unendliche zu verlängern, so muß man akzeptieren, daß es zu einem Kontinuum von Interaktionen auf einem endlichen Lebenszeit-Intervall kommt. Man kann diese unvermeidliche Folge als ebenso unrealistisch bezeichnen wie die Annahme einer unendlichen Lebensdauer einer Generation. Jedoch ist die Forderung nach Realismus in evolutionären Modellen durchaus nicht unproblematisch. Die Standard-Analysemethoden der evolutionären Spieltheorie selbst basiert auf der „unrealistischen“ Annahme einer unendlich großen Population. Der Vorwurf mangelnden Realismus richtet sich somit automatisch gegen (fast) die gesamte evolutionäre Spieltheorie, denn eine unendliche Population ist und bleibt nun einmal eine Fiktion. Mehr noch: Die Tatsache, daß Lernen in evolutionären Standardmodellen keine Effekte zeitigt und

auch nicht zeitigen kann, ist eben dieser Basisfiktion geschuldet! Es handelt sich mithin um ein äußerst artifizielles Resultat, denn unbestreitbar spielt Lernen in realen Populationen eine Rolle. Es ist unmittelbar evident, daß Menschen aus ihren vergangenen Erfahrungen lernen. Standardmodelle sind jedoch nicht in der Lage, dieses offenkundige Faktum einzufangen. Daher soll an dieser Stelle die Technik des simultanen Grenzübergangs offensiv verteidigt werden, als ein Versuch, dieses künstliche Defizit zu beseitigen.

Will man den Einwand mangelnden Realismus evolutionärer Modelle wirklich ernst nehmen, so muß man sich in letzter Konsequenz von einem deterministischen Modellrahmen verabschieden und zu einer stochastischen Modellierung übergehen. Dieser Schritt wird in Abschnitt 7 dieser Arbeit vollzogen. Zunächst werden wir aber noch eine Weile im deterministischen Modellrahmen bleiben.

## 6. Robustheit der Resultate

Im folgenden soll der Frage nachgegangen werden, wie robust die bisher abgeleiteten Ergebnisse gegenüber der Einführung weiterer mutanter Typen sind. Die Diskussion des Modells von Robson (1990) hatte für den Fall des GD gezeigt, daß es zwar leicht ist, eine kooperative mutante Strategie zu formulieren, die einzige ESS ist. Jedoch zeigte sich auch, daß dieses Ergebnis ausgesprochen sensitiv auf die Einführung einer defektiven Mutante reagierte. Was im allgemeinen langfristig im secret-handshake-Modell mit vier Typen erwartet werden kann, sind Gleichgewichte mit beiden Typen von Defektoren. Vor diesem Hintergrund stellt sich die berechnete Frage, ob den bisher abgeleiteten Ergebnissen ein ähnliches Schicksal widerfahren könnte.

Grundsätzlich stellt sich in evolutionären Modellen das Problem, daß natürlich eine a priori unbegrenzte Zahl möglicher Verhaltensmutanten als Untersuchungsgegenstand in betracht kommt. Eine erschöpfende Behandlung aller möglichen Automaten ist nicht leistbar. Man muß, will man nicht zu Simulationen greifen, sondern analytische Aussagen ableiten, sich Selbstbeschränkung auferlegen und die Zahl weiterer mutanter Typen klein halten. Dennoch kann man versuchen, bei der Auswahl der zu betrachtenden Mutanten geschickt vorzugehen. In unserem Kontext sind wir an der Fragestellung interessiert, wie robust die Resultate des Drei-Typen-Modells sind. Eine überzeugende Bestätigung der bisherigen Ergebnisse wäre sicherlich dann gegeben, wenn diese sich in einem möglichst harten Test identisch, oder wenigstens ähnlich, reproduzieren ließen.

Ein harter Test wäre in unserem Fall sicherlich der folgende: Betrachtet sei eine weitere Mutante, die ebenfalls über Lernfähigkeit verfügt und genauso wie der K-Typ Akteure durch Erfahrung lernt, die jedoch im Gegensatz zum K-Typ nunmehr prinzipiell defektiv orientiert ist. Defektion wird nur dann nicht gespielt, wenn i) entsprechende Information für eine Verhaltensänderung vorliegt und ii) Abweichen von Defektion vorteilhaft ist. Dies heißt im einzelnen: Gegenüber einem naiven Kooperatortyp wird die defektive Mutante stets defektieren. Beim Zusammentreffen mit einem Defektor (Hawk) ist zu unterscheiden danach, ob der Akteur bekannt ist oder nicht. Wurde der Spieler früher bereits einmal getroffen, so wird die defektive Mutante die Outside-option wahrnehmen und beide Spieler bekommen  $\lambda$ . Ist der Hawk jedoch unbekannt, so wird einmütig defektiert. Trifft die defektive Mutante nun einen K-Typ, so geschieht folgendes: Kennt der K-Typ die defektive Mutante bereits, so wird er die Outside-option wählen. Ist die Mutante für den K-Typ hingegen noch unbekannt, so wird Mutante II defektieren und den K-Typ ausbeuten. Der K-Typ realisiert in diesem Fall also  $-c$ , während Mutante II den maximal möglichen Payoff erzielt. Schließlich bleibt der Fall zu unter

suchen, in dem zwei defektive Mutanten aufeinander treffen. Ist einer der beiden Spieler bekannt, so weicht der wissende Spieler aus, und beide erhalten den Wert der outside option. Beachtet man, daß natürlich auch die defektive Mutante Kosten zu tragen hat und sei Mutante II im folgenden als N-Typ (für nasty) bezeichnet, so läßt sich die Payoff-Struktur des modifizierten GD mit zwei mutantentypen wie folgt zusammenfassen. Die Superskripte K und U bezeichnen wieder die Zustände bekannt, bzw. unbekannt.

	D <sup>K</sup>	D <sup>U</sup>	H <sup>K</sup>	H <sup>U</sup>	K <sup>K</sup>	K <sup>U</sup>	N <sup>K</sup>	N <sup>U</sup>
D <sup>K</sup>	3	3	0	0	3	3	0	0
D <sup>U</sup>	3	3	0	0	3	3	0	0
H <sup>K</sup>	6	6	1	1	$\lambda$	$\lambda$	$\lambda$	$\lambda$
H <sup>U</sup>	6	6	1	1	6	6	1	1
K <sup>K</sup>	3-c	3-c	$\lambda$ -c	-c	3-c	3-c	$\lambda$ -c	-c
K <sup>U</sup>	3-c	3-c	$\lambda$ -c	-c	3-c	3-c	$\lambda$ -c	-c
N <sup>K</sup>	6-c	6-c	$\lambda$ -c	1-c	$\lambda$ -c	$\lambda$ -c	$\lambda$ -c	$\lambda$ -c
N <sup>U</sup>	6-c	6-c	$\lambda$ -c	1-c	6-c	6-c	$\lambda$ -c	1-c

Tabelle 11: Das GD mit zwei lernfähigen Mutanten

Man mache sich an dieser Stelle noch einmal klar, daß die so konstruierte defektive Mutante tatsächlich den eingangs behaupteten harten Test für das Basismodell darstellt. Gegeben seinen Informationsstand, realisiert der N-Typ nämlich den maximal möglichen Payoff, während für den K-Typ weitere Gelegenheiten entstehen, in denen er ausgebeutet wird. Es ist vorderhand also keineswegs klar, ob sich die bisherigen Resultate in einem Viertypen-Modell reproduzieren lassen. Es scheint vielmehr, daß der N-Typ durchaus gute Chancen hätte, sich evolutionär zu behaupten. Dies ist aber im wesentlichen nicht der Fall, wie die folgende Analyse zeigen wird.

Der Analyserahmen sei im folgenden der des eben entwickelten Modells mit simultanem Grenzübergang. Der Grund hierfür ist rein praktischer Natur. Die algebraischen Terme im Viertypen-Modell nehmen ohnehin, wie man sehen wird, teils beängstigende Ausmaße an. Eine Ableitung im Rahmen des Basismodells, mit einer in unendlich viele Subpopulationen zerfallende Population, bringt keinerlei Erkenntniszuwachs, verkompliziert die Rechenarbeit jedoch ungemein und läßt die Terme die Grenzen der Darstellbarkeit sprengen. Es sei darauf hingewiesen, daß alle Resultate, mit den bekannten algebraischen Verkomplizierungen, identisch im Rahmen des Basismodells hergeleitet werden können.



Es sei also angenommen, daß die zeitliche Entwicklung der Populationsanteile der beschriebenen vier Verhaltenstypen beschrieben wird durch die Replikatordynamik. Gelte also

$$\dot{x}_i = x_i(\bar{\pi}_i - \bar{\pi}), \quad i = H, D, K, N \quad (6-1)$$

Der folgende Satz gibt Aufschluß über zusätzliche Fixpunkte (die bisherigen Fixpunkte bleiben natürlich im Viertypenmodell erhalten).

### Satz 6.1: (Existenz II)

i) *Unter der Replikator-Dynamik existiert kein vollständig gemischter Fixpunkt in  $\text{int}\Delta_4$ .*

ii)  $\Delta_3$ : Für  $c < \frac{3\lambda(1-k)}{3-k(\lambda(1-\beta)+\beta)}$  existiert ein zusätzlicher Fixpunkt DKN

$$\{x_D^*, x_K^*, x_N^*\} = \left\{ \begin{array}{l} \frac{\lambda^2(1-k)^2 + \lambda(k(9-6k+2c-c\beta)-c-3) + c(3-k(6-\beta))}{(1-k)^2(\lambda-6)}, \\ \frac{\lambda k(3-c(1-\beta)) + c(3-k\beta) - 3\lambda}{\lambda(k-1)(k-1)(\lambda-6)}, 1-x_D^* - x_K^* \end{array} \right\}$$

iii) *In  $\Delta_2$ : Zusätzliche Fixpunkte KN*

$$\{x_K^*, x_N^*\} = \left\{ \frac{(\lambda(1-\beta)+\beta)k}{(\lambda(2-\beta)+\beta-6)+3-\lambda}, 1-x_K^* \right\} \text{ sowie HN}$$

$$\{x_H^*, x_N^*\} = \left\{ \frac{k(1-\beta)(\lambda-1)-c}{(k(2-\beta)-1)(\lambda-1)}, 1-x_H^* \right\} \text{ existieren für } k < \frac{3-\lambda}{6-\lambda}, \text{ bzw.}$$

$$(\lambda-1)(1-k) > c > k(1-\beta)(\lambda-1).$$

$$\text{Dabei bezeichnen } k = \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right), \quad \beta = \frac{e^{2\gamma} - 1}{2(e^{2\gamma} - e^\gamma)}.$$

**Beweis:** Der Beweis ist keineswegs schwierig, jedoch etwas mühselig und darum im Appendix zum Kapitel 6 nachzulesen.

Wir sind vorrangig wieder an der Frage der Stabilität von Fixpunkten interessiert. Insbesondere interessiert, ob die Fixpunkte mit Beteiligung kooperativer Typen aus dem Basismodell im Viertypenmodell stabil bleiben, oder ob etwa die defekti-ve Mutante in der Lage ist, das vollständig gemischte Gleichgewicht zu invadieren. Bemerkenswerterweise ist letzteres nicht der Fall, wie Teil (i) der folgenden Proposition zeigt.

**Satz 6.2: (Lokale Stabilität)**

i) Ist der Fixpunkt HKD:

$$(x_H^*, x_K^*, x_D^*) = \left\{ \frac{c}{\lambda(1-k)}, \frac{2c-3\lambda(1-k)}{\lambda(\lambda-6)(1-k)^2}, 1-x_H^*-x_K^* \right\} \text{ im Dreitypenmodell}$$

lokal stabil, so ist er es auch im Viertypenmodell.

ii) Der Fixpunkt HK:

$$(x_H^*, 1-x_H^*) = \left\{ \frac{3-\lambda-k(6-\lambda)-c}{4-2\lambda-k(6-\lambda)}, \frac{1+c-\lambda(1-k)}{4-2\lambda-k(6-\lambda)} \right\} \text{ ist lokal stabil, falls}$$

$$c < \frac{3\lambda(1-k)}{3-k(\lambda(1-\beta)+\beta)}$$

iii) Der Fixpunkt DKN ist niemals stabil.

iv) Der Fixpunkt KN ist niemals stabil

v) Der Fixpunkt HN ist immer lokal stabil.

vi) Der Fixpunkt H mit Hawks allein ist lokal stabil für  $c > (\lambda-1)(1-k)$ .

Es ergeben sich also im wesentlichen keine Änderungen gegenüber dem Basismodell. Das gemischte Gleichgewicht mit Tauben, K-Typen und Falken ist auch im Viertypenmodell lokal stabil, das heißt, es kann vom N-Typ nicht invadiert werden. Das Zweitypen-Gleichgewicht mit intelligenten Kooperativen und Falken ist ebenfalls weiterhin stabil, jedoch, wie sich zeigen wird, für einen etwas reduzierten Parameterbereich. Interessanterweise sind zwei der drei zusätzlichen Fixpunkte mit Beteiligung der defektiven Mutante immer instabil. Lediglich eine Mischung der beiden defektiven Typen erweist sich als lokal stabil. Der monomorphe Zustand mit Falken allein ist weiterhin stabil. Jedoch ergibt sich auch hier ein etwas zugunsten des neuen Fixpunktes aus (v) reduzierter Parameterbereich. Nun zum Beweis der Proposition.

**Beweis:**

Ad i) Um die lokale Stabilität des Dreitypen-Fixpunktes bezüglich des  $\Delta_4$  zu prüfen, reicht es aus, den transversalen Eigenwert bezüglich der charakteristischen Richtung zu berechnen. Ist dieser negativ, so ist der betreffende Fixpunkt stabil. Dieses Resultat wurde von Hofbauer und Sigmund (1988, S.166) bewiesen. Der transversale Eigenwert für unser Problem ist definiert als

$$\lambda_N \equiv \frac{\dot{x}_N}{x_N} \Big|_{HDK} = (\bar{\pi}_N - \bar{\pi}) \Big|_{HDK} = -\frac{c}{\lambda} < 0. \quad (6-2)$$

Offensichtlich ist der Eigenwert für alle zulässigen Werte von  $c, \lambda$  negativ, so daß der Dreitypen-Fixpunkt für den ursprünglichen Parameterbereich auch bezüglich des Simplex  $\Delta_4$  stabil ist. Die Forderung, daß für lokale Stabilität der transversale Eigenwert negativ sein muß, ist intuitiv sofort einsichtig. Sie besagt für den Fall der Replikatorodynamik nämlich nichts anderes, als daß der mutante N-Typ immer einen niedrigeren Payoff erzielt als der Populationsdurchschnitt im inneren Gleichgewicht. Dann ist klar, daß der N-Typ das innere Gleichgewicht niemals invadieren kann.

Ad ii) Wir berechnen wiederum den transversalen Eigenwert, ausgewertet im Zweitypen-Fixpunkt HK. Dieser ergibt sich als

$$\lambda_N \equiv \left. \frac{\dot{x}_N}{x_N} \right|_{HK} = (\lambda - 1)(1 - k)(1 - x_k^*) - c, \quad (6-3)$$

$$\text{wobei } x_k^* = \frac{1 - \lambda(1 - k) + c}{2(2 - \lambda(1 - k) - 3k)}$$

$$\lambda_N < 0 \Leftrightarrow$$

$$(\lambda - 1)(1 - k)(3 - \lambda(1 - k) - 6k) > c(3 - \lambda(1 - k) - 5k) \quad (6-4)$$

Offensichtlich ist eine Fallunterscheidung notwendig.

$$a) (3 - \lambda(1 - k) - 5k) > 0$$

Die einzig relevante Konstellation in diesem Fall ist offensichtlich, daß auch  $(3 - \lambda(1 - k) - 6k) > 0$ , da ansonsten von  $c$  verlangt würde, negative Werte anzunehmen. Also:

$$\lambda_N < 0 \Leftrightarrow c < \frac{(3 - \lambda(1 - k) - 6k)}{(3 - \lambda(1 - k) - 5k)} (\lambda - 1)(1 - k) \quad (6-5)$$

$$b) (3 - \lambda(1 - k) - 5k) < 0,$$

was natürlich impliziert  $(3 - \lambda(1 - k) - 6k) < 0$ . Damit :

$$\lambda_N < 0 \Leftrightarrow c > \frac{(3 - \lambda(1 - k) - 6k)}{(3 - \lambda(1 - k) - 5k)} (\lambda - 1)(1 - k) \quad (6-6)$$

Wir werden nun zeigen, daß Fall (b) keinerlei Relevanz besitzt, und zwar aus folgendem Grund: Lokale Stabilität von HK bezüglich  $\Delta_3$  erforderte  $c < \lambda(1 - k) - 1$ . Diese Bedingung ist nicht konsistent mit der Bedingung aus Fall (b). Dies zeigt man leicht wie folgt: Konsistenz würde verlangen

$$\begin{aligned}
\lambda(1-k)-1 &> \frac{(3-\lambda(1-k)-6k)}{(3-\lambda(1-k)-5k)}(\lambda-1)(1-k) \Leftrightarrow \\
(\lambda-1)(1-k) &\left[1 - \frac{(3-\lambda(1-k)-6k)}{(3-\lambda(1-k)-5k)}\right] > k \Leftrightarrow \\
\frac{k}{(3-\lambda(1-k)-5k)} &> \frac{k}{(\lambda-1)(1-k)} \Leftrightarrow \\
\frac{(\lambda-1)(1-k)}{(3-\lambda(1-k)-5k)} &> 1. \tag{6-7}
\end{aligned}$$

Klarerweise gilt  $\frac{(\lambda-1)(1-k)}{(3-\lambda(1-k)-5k)} < 0$ , solange  $(3-\lambda(1-k)-5k) < 0$ .

Die Forderung  $\frac{(\lambda-1)(1-k)}{(3-\lambda(1-k)-5k)} > 1$  kann im Fall (b) also niemals erfüllt werden.

Man kann sich außerdem leicht klar machen, daß die Stabilitätsbedingung bezüglich  $\Delta_4$  immer restriktiver ist als die Stabilitätsbedingung bezüglich  $\Delta_3$ .

Ad iii) Die Entwicklung der Populationsanteile im Viertypenmodell wird vollständig beschrieben durch ein System von drei nichtlinearen Differentialgleichungen. Der Populationsanteil des N-Typs wird dabei als Residuum zu 1 ermittelt. Für den Fixpunkt DKN gilt natürlich  $x_H^* = 0$ . Wie man sich leicht klar macht, muß die Jakobi-Matrix dann allgemein folgende Gestalt haben:

$$J = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad a_{ij} = \left. \frac{\partial \dot{x}_i}{\partial x_j} \right|_{DKN} \quad i, j = D, H, K \tag{6-8}$$

Dies folgt sofort aus folgender Überlegung: Der Populationsanteil der Hawks wird allgemein beschrieben durch

$$\dot{x}_H = x_H(\pi_H - \bar{\pi}). \tag{6-9}$$

Die Einträge der entsprechenden Zeile in der Jakobimatrix ergeben sich dann als

$$\begin{aligned}
\frac{\partial \dot{x}_H}{\partial x_D} &= x_H \frac{\partial \pi_H}{\partial x_D} - x_H \frac{\partial \bar{\pi}}{\partial x_D}, \\
\frac{\partial \dot{x}_H}{\partial x_H} &= \pi_H + x_H \frac{\partial \pi_H}{\partial x_H} - \bar{\pi} - x_H \frac{\partial \bar{\pi}}{\partial x_H}, \\
\frac{\partial \dot{x}_H}{\partial x_K} &= x_H \frac{\partial \pi_H}{\partial x_K} - x_H \frac{\partial \bar{\pi}}{\partial x_K}.
\end{aligned} \tag{6-10}$$

Für  $x_H^* = 0$  ergibt sich dann gerade

$$\frac{\partial \dot{x}_H}{\partial x_D} = 0, \quad \frac{\partial \dot{x}_H}{\partial x_H} = \pi_H - \bar{\pi} \equiv a_{22}, \quad \frac{\partial \dot{x}_H}{\partial x_K} = 0. \quad (6-11)$$

Dies hat komfortable Vereinfachungen für die Eigenwertberechnung zur Folge. Man zeigt leicht, daß  $a_{22}$  ein Eigenwert der Jakobimatrix sein muß:

$$\begin{aligned} DetJ &= a_{11} Det \begin{bmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{bmatrix} - a_{12} Det \begin{bmatrix} 0 & 0 \\ a_{31} & a_{33} \end{bmatrix} + a_{13} Det \begin{bmatrix} 0 & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \Leftrightarrow \\ DetJ &= a_{11} a_{22} a_{33} - a_{13} a_{22} a_{31} \Leftrightarrow \\ DetJ &= a_{22} (a_{11} a_{33} - a_{13} a_{31}) \end{aligned} \quad (6-12)$$

Definitionsgemäß läßt der Eigenwert einer Matrix diese singularär werden: Aus  $DetJ = 0$  folgt dann, daß  $a_{22}$  ein Eigenwert der Jakobimatrix ist.

Wertet man im Fixpunkt DKN aus, so ergibt sich  $a_{22}$  als:

$$a_{22} \equiv \left. \frac{\partial \dot{x}_H}{\partial x_H} \right|_{DKN} = \frac{c}{\lambda(1-k)} (\lambda(1+k(\beta-2)) + k(1-\beta)) \quad (6-13)$$

Man kann leicht zeigen, daß dieser Ausdruck immer positiv sein muß:

$$\begin{aligned} a_{22} > 0 &\Leftrightarrow \\ (\lambda(1+k(\beta-2)) + k(1-\beta)) &> 0 \end{aligned} \quad (6-14)$$

Wir wissen (vgl. (6-67) im Appendix 6.3), daß

$k < \frac{1}{2-\beta}$ , was äquivalent ist zu  $1+k(\beta-2) > 0$ . Also:

$$\begin{aligned} (\lambda(1+k(\beta-2)) + k(1-\beta)) &> 0 \Leftrightarrow \\ \lambda &> \frac{k(\beta-1)}{1+k(\beta-2)} \end{aligned} \quad (6-15)$$

Nun gilt einerseits  $1+k(\beta-2) > 0$ , andererseits kann  $\beta$  nur Werte zwischen 0 und 1 annehmen. Die rechte Seite der Ungleichung ist mithin negativ, weshalb die Ungleichung für alle zulässigen Werte des Parameters  $\lambda$  erfüllt ist. Ergo ist der

Eigenwert  $a_{22}$  notwendigerweise positiv und der betrachtete Fixpunkt DKN immer instabil.

Ad iv) Da  $x_H = x_D = 0$  ergibt sich die Jakobimatrix als

$$J = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \text{ mit } a_{ij} = \left. \frac{\partial \dot{x}_i}{\partial x_j} \right|_{KN} \quad i, j = D, H, K. \quad (6-16)$$

Wie man sich leicht klarmacht, sind die Eigenwerte daher durch die drei Elemente der Hauptdiagonalen gegeben. Denn:

$$\begin{aligned} \text{Det}J &= a_{11} \text{Det} \begin{bmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \text{Det} \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \text{Det} \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \Leftrightarrow \\ &\text{Det}J = a_{11} a_{22} a_{33}, \quad \text{da } a_{12} = a_{13} = 0. \end{aligned} \quad (6-17)$$

Betrachtet sei nun der dritte Eigenwert

$$a_{33} \equiv \left. \frac{\partial \dot{x}_K}{\partial x_K} \right|_{KN} = k \frac{\lambda^2(1-\beta)(1-k) + \lambda(k(6-7\beta) + 4\beta - 3) + 6k\beta - 3\beta}{-(k(\lambda(2-\beta) - 6 + \beta) + 3 - \lambda)}. \quad (6-18)$$

Man bemerkt zunächst, daß der Nenner des Terms immer negativ ist. Dies folgt unmittelbar aus dem Existenzkriterium

$$\begin{aligned} x_K^* &= (\lambda(1-\beta) + \beta) \frac{k}{k(\lambda(2-\beta) - 6 + \beta) + 3 - \lambda} > 0 \Leftrightarrow \\ &k(\lambda(2-\beta) - 6 + \beta) + 3 - \lambda > 0. \end{aligned} \quad (6-19)$$

Damit ergibt sich

$$\begin{aligned} a_{33} &> 0 \Leftrightarrow \\ \lambda^2(1-\beta)(1-k) + \lambda(k(6-7\beta) + 4\beta - 3) + 6k\beta - 3\beta &< 0. \end{aligned} \quad (6-20)$$

Nach einigen Umformungen findet man

$$a_{33} > 0 \Leftrightarrow k < \frac{3-\lambda}{6-\lambda}. \quad (6-21)$$

Diese Bedingung ist jedoch identisch mit derjenigen, die aus der Existenzforderung  $x_K^* < 1$  resultiert (vgl. Appendix 6.3). Damit ist gezeigt, daß der dritte Eigenwert der Jakobimatrix positiv ist, sobald der Fixpunkt KN existiert. KN kann also niemals stabil sein.

Ad v) Wenden wir uns schließlich dem Zweitypenfixpunkt HN zu. Wegen  $x_K = x_D = 0$  resultiert die Jakobimatrix als

$$J = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}, \text{ mit } a_{ij} = \left. \frac{\partial \dot{x}_i}{\partial x_j} \right|_{HN}, \quad i, j = D, H, K. \quad (6-22)$$

Also sind die Eigenwerte wiederum durch die drei Elemente auf der Hauptdiagonalen gegeben. Wir beginnen mit dem ersten Eigenwert. Dieser berechnet sich als

$$a_{11} \equiv \left. \frac{\partial \dot{x}_D}{\partial x_D} \right|_{HN} = \lambda \frac{(1-k)^2}{k(2-\beta)-1} - \frac{c(1-k) + k(k-\beta)}{k(2-\beta)-1}. \quad (6-23)$$

Da  $k(2-\beta)-1 < 0$ ,

$$a_{11} < 0 \Leftrightarrow c < \lambda(1-k) - k \frac{k-\beta}{1-k}. \quad (6-24)$$

Wir wissen:  $x_H^* < 1 \Leftrightarrow c < (\lambda-1)(1-k)$ .

Wir zeigen, daß wann immer der Fixpunkt existiert, der erste Eigenwert negativ sein muß:

$$\begin{aligned} (\lambda-1)(1-k) < \lambda(1-k) - k \frac{k-\beta}{1-k} &\Leftrightarrow \\ k(k-\beta) < (1-k)^2 & \quad (6-25) \end{aligned}$$

Da bekanntlich  $k-\beta < 0$  (Lemma 6.1, Appendix 6.3), ist die linke Seite der Ungleichung negativ und die Ungleichung somit immer erfüllt. Sobald die Existenzforderung erfüllt ist, ist der erste Eigenwert somit negativ.

Wir fahren fort mit dem Eigenwert Nummer 2:

$$a_{22} \equiv \left. \frac{\partial \dot{x}_H}{\partial x_H} \right|_{HN} =$$

$$-\frac{\lambda^2((\beta-1)(k^2-k)) + \lambda(2(k^2-k)(1-\beta) - c(1-k\beta)) - (1-\beta)(k^2-k) + c(1-k\beta) + c^2}{(\lambda-1)(k(2-\beta)-1)} \quad (6-26)$$

Zu fordern ist

$$\begin{aligned} a_{22} < 0 &\Leftrightarrow \\ \lambda^2((\beta-1)(k^2-k)) + \lambda(2(k^2-k)(1-\beta) - c(1-k\beta)) - (1-\beta)(k^2-k) + c(1-k\beta) + c^2 < 0 &\Leftrightarrow \\ c^2 + c((1-k\beta)(1-\lambda)) + \lambda^2((\beta-1)(k^2-k)) + \lambda(k^2-k)(2-\beta) - (1-\beta)(k^2-k) < 0. \end{aligned} \quad (6-27)$$

Diese Ungleichung ist erfüllt für

$$k(1-\beta)(\lambda-1) < c < (\lambda-1)(1-k). \quad (6-28)$$

Dies ist aber exakt der Existenzbereich des Fixpunktes (vgl. (6-82) und (6-81) im Appendix 6.3 ). Sobald HN also existiert, ist auch der zweite Eigenwert notwendig negativ.

Bleibt der dritte Eigenwert zu untersuchen:

$$a_{33} \equiv \left. \frac{\partial \dot{x}_K}{\partial x_K} \right|_{HN} = k\lambda \frac{(1-k)(1-\beta)}{k(2-\beta)-1} + k \frac{\beta-k-c(1-\beta)}{k(2-\beta)-1} \quad (6-29)$$

$$\begin{aligned} a_{33} < 0 &\Leftrightarrow \\ c < \frac{\lambda(1-k)(1-\beta) + \beta - k}{1-\beta} \end{aligned} \quad (6-30)$$

Wir zeigen wiederum, daß Bedingung (6-30) notwendig erfüllt ist, wenn der Fixpunkt existiert. Existenz verlangte  $c < (\lambda-1)(1-k)$ . Man zeigt leicht, daß die rechte Seite der Existenzforderung immer kleiner ist als die rechte Seite von (6-30):

$$\begin{aligned} (\lambda-1)(1-k) &< \frac{\lambda(1-k)(1-\beta) + \beta - k}{1-\beta} \Leftrightarrow \\ (\lambda-1)(1-k)(1-\beta) &< \lambda(1-k)(1-\beta) + \beta - k \Leftrightarrow \\ -(1-k)(1-\beta) &< \beta - k \end{aligned} \quad (6-31)$$



Nun ist die linke Seite der Ungleichung klarerweise negativ, während die rechte positiv ist (vgl. erneut Lemma 6.1). Wenn der Fixpunkt HN existiert, ist also automatisch die Forderung für lokale Stabilität erfüllt.

Es wurde gezeigt, daß im Parameterbereich der Existenz von HN alle drei Eigenwerte der Jakobimatrix negativ sind. Der Fixpunkt ist also immer lokal stabil.

Ad vi) Die Eigenwerte für den monomorphen Hawk-Fixpunkt ergeben sich als

$$\begin{aligned} a_{11} &\equiv \left. \frac{\partial \dot{x}_D}{\partial x_D} \right|_H = -1, & a_{22} &\equiv \left. \frac{\partial \dot{x}_H}{\partial x_H} \right|_H = (\lambda - 1)(1 - k) - c, \\ a_{33} &\equiv \left. \frac{\partial \dot{x}_K}{\partial x_K} \right|_H = \lambda(1 - k) - 1 - c. \end{aligned} \tag{6-32}$$

$a_{22} < 0$  ist restriktiver als  $a_{33} < 0$ . Daher: Für  $c > (\lambda - 1)(1 - k)$  ist der Fixpunkt lokal stabil.

Damit ist der Beweis zur lokalen Stabilität vollständig. □

## 6.2 Diskussion

Zunächst ist festzustellen, daß sich durch die Einführung einer lernfähigen, defektiven Mutante, dem N-Typ, nicht all zuviel zu ändern scheint. Einige neue Fixpunkte mit Beteiligung der N-Mutante tauchen auf. Die meisten von ihnen erweisen sich jedoch als instabil, so der neue Fixpunkt DKN im  $\Delta_3$  und der Fixpunkt KN im  $\Delta_2$ . Der vollständig gemischte Fixpunkt HDK aus dem Dreitypen-Modell behauptet seine lokale Stabilität für den ursprünglichen Parameterbereich auch im  $\Delta_4$ . Lediglich einer der neu auftauchenden Fixpunkte erweist sich als lokal stabil, nämlich der Zweitypen-Fixpunkt HN, mit “dummen” und “intelligenten”, lernfähigen Defektoren.

Bisher jedoch war die Rede nur von lokaler Stabilität. Die Selbstbeschränkung auf lokale Aussagen hat ihren triftigen Grund, und der ist in eben jener scheinbar harmlosen Novität des stabilen Fixpunktes HN zu sehen. Der Beweis der globalen Stabilität im Basismodell machte nämlich Gebrauch von dem komfortablen Umstand, daß es koexistierende lokal stabile NGG dort nicht gibt. Die Parameterbereiche für lokale Stabilität überlappen schlicht nicht. Dies kann, muß aber nicht der Fall sein im Viertypenmodell, wie das folgende Beispiel verdeutlicht.

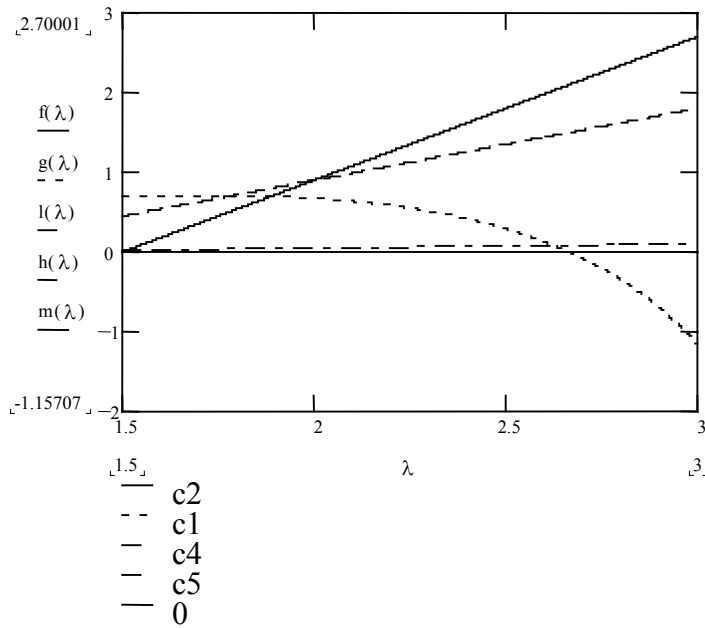


Abbildung 12: Verlauf der kritischen Werte für den Parameter  $c$

Die Abbildung zeigt verschiedene kritische Werte für den Parameter  $c$  als Funktion des Parameters  $\lambda$ . Es ist

$$f(\lambda) = (2\lambda - 3)(1 - k) \equiv c_2, \quad g(\lambda) = \frac{\lambda(1 - k)[(\lambda - 6)(1 - k) + 3]}{(\lambda - 6)(1 - k) + 2} \equiv c_1.$$

Diese beiden kritischen Grenzen definieren den Bereich lokaler Stabilität des Fixpunktes HDK (vgl. Satz 5.1'). Die zwei weiteren Funktionen

$$h(\lambda) = k(1 - \beta)(\lambda - 1) \equiv c_4, \quad l(\lambda) = (\lambda - 1)(1 - k) \equiv c_5$$

definieren den Stabilitätsbereich des Fixpunktes HN. Man sieht nun deutlich einen weiten Bereich des Parameters  $\lambda$ , etwa von 1.5 bis 2.6, in dem die Stabilitätsbedingungen für beide Fixpunkte ( $c < \min\{c_1, c_2\}$ ,  $c_4 < c < c_5$ ) simultan erfüllt sind. Man beachte, daß es sich bei der gewählten Parametrisierung ( $\gamma = 10$ ) im übrigen um ein ausgesprochen optimistisches Szenario handelt.

Man kann also im allgemeinen nicht damit rechnen, daß lokale Stabilität globale Stabilität impliziert. Natürlich gibt es auch weiterhin Parameterbereiche, in denen die beiden Fixpunkte nicht als lokal stabile Populationszustände koexistieren. In diesem Fall kann analog zum Beweis der Proposition 5.1 auch für das Viertypenmodell auf die globale Stabilität des Fixpunktes HDK geschlossen werden. Die Argumentation ist identisch: Es gibt keine konkurrierenden Kandidaten für Attraktoren des dynamischen Systems, nämlich weder lokal stabile Fixpunkte,

noch innere Grenzyklen (das Hofbauer-Sigmund-Argument greift ja für dynamische Systeme bis zu einer Dimensionalität von  $n=3$ ).

Man muß jedoch konzedieren, daß die Bedingungen für globale Stabilität restriktiver sind. Es sind zusätzliche Bedingungen an die Parameter zu stellen. Eine notwendige Bedingung für globale Stabilität ist die Überschneidungsfreiheit der Parameterbereiche für die lokale Stabilität der Fixpunkte HDK und HN. Für den Fall, daß  $c_2 < c_1$ , wäre diese gewährleistet sofern gilt  $c_4 > c_2$ . Wann aber kann am ehesten damit gerechnet werden, daß diese Bedingung erfüllt ist?

Natürlich kann man zunächst versuchen, diese Ungleichung nach einer Variablen, beispielsweise  $k$ , aufzulösen. Man erhält dann

$$c_4 > c_2 \Leftrightarrow k > \frac{2\lambda - 3}{(\lambda - 1)(3 - \beta) - 1}. \quad (6-33)$$

Eine solche Form der Darstellung ist für unsere Zwecke jedoch wenig dienlich. Man erinnere sich daran, daß sowohl  $k$  als auch  $\beta$  Funktionen des Parameters  $\gamma$  sind. Leider erweist es sich als unmöglich, die Ungleichung nach  $\gamma$  zu explizieren. Wir müssen daher einen indirekten Weg wählen, um Aufschluß über die oben gestellte Frage zu gewinnen.

Der Parameter  $c_4$  war gegeben als  $c_4 = k(1 - \beta)(\lambda - 1)$ . Man kann sich zunächst die Frage stellen, in welcher Weise der kritische Wert  $c_4$  auf partielle Variation von  $k$  und  $\beta$  reagiert. Man zeigt leicht, daß gilt

$$\frac{\partial k(\gamma)}{\partial \gamma} < 0, \quad \frac{\partial \beta(\gamma)}{\partial \gamma} < 0 \quad \forall \gamma > 0. \quad (6-34)$$

Damit aber ist klar, daß der Gesamteffekt der Reaktion von  $c_4$  bei Variation von  $\gamma$  offen ist. Wächst  $\gamma$ , so sinkt  $k$  und damit auch  $c_4$ .  $\beta$  jedoch sinkt ebenfalls, so daß  $c_4$  steigt. Welcher der Effekte überwiegt also?

Betrachtet sei das totale Differential von  $c_4$ , also

$$dc_4 = (1 - \beta)(\lambda - 1) dk - k(\lambda - 1)d\beta \quad (6-35)$$

Für  $dk, d\beta < 0$  ergibt sich:

$$dc_4 < 0 \Leftrightarrow \frac{dk}{d\beta} > \frac{k}{1 - \beta} \quad (6-36)$$

Natürlich können die Differentiale weiter zurückgeführt werden auf die Variation von  $\gamma$ :

$$dk = \frac{\partial k(\gamma)}{\partial \gamma} d\gamma = \frac{1}{\gamma^2} \left( \frac{1}{e^\gamma} (\gamma + 1) - 1 \right) d\gamma \quad (6-37)$$

$$d\beta = \frac{\partial \beta(\gamma)}{\partial \gamma} d\gamma = \frac{2e^{2\gamma} - e^\gamma - e^{2\gamma} e^\gamma}{2(e^\gamma - e^{2\gamma})^2} d\gamma \quad (6-38)$$

Division der Terme liefert nach einigen Umformungen:

$$\frac{dk}{d\beta} = 2 \frac{[e^{-\gamma} (\gamma + 1) - 1] [-e^{2\gamma} + e^\gamma]^2}{\gamma^2 (2e^{2\gamma} - e^{3\gamma} - e^\gamma)} \quad (6-39)$$

Die rechte Seite der Ungleichung (6-36) war gegeben als  $\frac{k}{1-\beta}$ . Einsetzen von  $k$  und  $\beta$  und Vereinfachen liefert

$$\frac{k}{1-\beta} = 2(e^{-\gamma} - 1) \frac{(e^{2\gamma} - e^\gamma)}{\gamma(2e^\gamma - e^{2\gamma} - 1)}. \quad (6-40)$$

Dann:

$$\begin{aligned} \frac{dk}{d\beta} &> \frac{k}{1-\beta} \Leftrightarrow \\ \frac{dk(1-\beta)}{kd\beta} &> 1 \Leftrightarrow \\ \frac{2 \frac{[e^{-\gamma} (\gamma + 1) - 1] [-e^{2\gamma} + e^\gamma]^2}{\gamma^2 [2e^{2\gamma} - e^{3\gamma} - e^\gamma]}}{2(e^{-\gamma} - 1) \frac{(e^{2\gamma} - e^\gamma)}{\gamma(2e^\gamma - e^{2\gamma} - 1)}} &> 1 \Leftrightarrow \\ \frac{[e^{-\gamma} (\gamma + 1) - 1] [-e^{2\gamma} + e^\gamma]^2 (2e^\gamma - e^{2\gamma} - 1)}{\gamma(e^{-\gamma} - 1)(e^{2\gamma} - e^\gamma)e^\gamma(2e^\gamma - e^{2\gamma} - 1)} &> 1 \Leftrightarrow \\ \frac{[e^{-\gamma} (\gamma + 1) - 1] [-e^{2\gamma} + e^\gamma]}{\gamma e^\gamma (-1)(e^{-\gamma} - 1)} &> 1 \Leftrightarrow \\ \frac{[e^{-\gamma} (\gamma + 1) - 1] (-1)(1 - e^\gamma)}{\gamma \left[ \frac{1 - e^\gamma}{e^\gamma} \right]} &> 1 \Leftrightarrow \\ -\frac{1}{\gamma} e^\gamma [e^{-\gamma} (\gamma + 1) - 1] &> 1, \end{aligned} \quad (6-41)$$

was sich weiter vereinfachen läßt zu

$$\begin{aligned} e^\gamma &> 2\gamma + 1 \Leftrightarrow \\ \gamma &> \ln(2\gamma + 1). \end{aligned} \quad (6-42)$$

Man erinnere sich nun daran, daß

$$\gamma = \begin{cases} 1/\alpha, & \text{falls } \Delta t < \Delta n \\ \alpha & \Delta t \geq \Delta n \end{cases}, \quad \alpha \in \mathbb{N}_+.$$

$\gamma$  ist also entweder rationalwertig im Intervall zwischen 0 und 1, oder aber positiv ganzzahlig. Wir zeigen kurz, daß obige Ungleichung (6-42) für alle  $0 < \gamma \leq 1$  nicht, während sie für alle  $1 < \gamma$  stets erfüllt ist. Bezeichne im folgenden  $f(\gamma) = \gamma$ ,  $g(\gamma) = \ln(2\gamma + 1)$ .

i) Beide Funktionen  $g(\gamma)$  und  $f(\gamma)$  entspringen dem Ursprung. ii) Außerdem ist bekanntlich der Logarithmus naturalis für alle  $\gamma > 0$  eine in seinem Argument streng monoton wachsende, konkave Funktion. iii) An der Stelle  $\gamma = 1$  gilt:  $g(\gamma) = \ln(3) \cong 1.099 > 1 = f(\gamma)$ . Also muß, (i),(ii) und (iii) zusammen genommen, im gesamten Intervall  $(0,1]$  gelten:  $g(\gamma) > f(\gamma)$ .

Für alle  $\gamma > \frac{1}{2}$  gilt natürlich, daß  $g(\gamma)$  eine Steigung kleiner 1 aufweist, denn:

$$\frac{\partial g(\gamma)}{\partial \gamma} < 1 \Leftrightarrow \frac{2}{2\gamma + 1} < 1 \Leftrightarrow \gamma > \frac{1}{2} \quad (6-43)$$

An der Stelle  $\gamma = 2$  gilt:  $g(\gamma) = \ln(5) \cong 1.609 < 2 = f(\gamma)$ . Daher gilt für alle  $\gamma > 1, \gamma \in \mathbb{N}_+$ :  $g(\gamma) < f(\gamma)$ .

Damit wurde gezeigt, daß der kritische Wert  $c_4$  für  $\gamma > 1, \gamma \in \mathbb{N}_+$  in  $\gamma$  abnimmt, denn: Für  $\gamma > 1, \gamma \in \mathbb{N}_+$  ist (6-42) erfüllt. Dann aber gilt auch (6-36), weshalb dann auch gilt  $dc_4 < 0$ .

Befinden wir uns also bereits im Bereich vergleichsweise günstiger Lernbedingungen – indiziert durch  $\gamma > 1, \gamma \in \mathbb{N}_+$  – dann führt eine weitere Verbesserung der Lernbedingungen ( $d\gamma > 0$ ) zu einem Sinken des kritischen Wertes  $c_4$ . Je besser also gelernt wird, desto eher wird man damit rechnen müssen, daß die Parameterbereiche für lokale Stabilität von HDK und HN überlappen und damit globale Stabilität nicht ohne weiteres zu erwarten ist. Das langfristige Konvergenzverhalten der Replikator-Dynamik ist in diesen Fällen nicht mehr unabhängig von den Startbedingungen.

Diese Resultate kontrastieren zu denen, die Amann/Yang (1994) für den Spezialfall perfekten Wissens ableiten. Im Amann/Yang-Fall findet man zwar auch einen zusätzlichen Fixpunkt HN, jedoch kann dieser unter perfektem Wissen niemals

stabil sein. Diese Differenz ist sicherlich bemerkenswert und verdient eine genauere Untersuchung.

Rein formal erschließt sich diese Differenz zunächst über die Betrachtung der Eigenwerte. Im allgemeinen Fall imperfekter Information galt notwendigerweise, daß alle drei Eigenwerte der Jakobimatrix negativ sein müssen. Für den Fall perfekten Wissens kann man leicht zeigen, daß auch hier zwei der drei Eigenwerte negativ sind. Formal wird perfektes Wissens reflektiert durch  $k=0$ . Betrachtet man nun den dritten Eigenwert für  $k=0$ , so stellt man fest, daß dieser zu Null “degeneriert”:

$$a_{33} \equiv \left. \frac{\partial \dot{x}_K}{\partial x_K} \right|_{HN} = k\lambda \frac{(1-k)(1-\beta)}{k(2-\beta)-1} + k \frac{\beta-k-c(1-\beta)}{k(2-\beta)-1} = 0 \text{ für } k=0 \quad (6-44)$$

Invadiert also eine K-Mutante eine HN-Population, so wird sie sich zwar nicht weiter ausbreiten, ihr Populationsanteil wird aber auch nicht abnehmen, sondern er bleibt konstant. Was ist die tiefere Ursache für dieses Phänomen?

Vergleichen wir zur Beantwortung dieser Frage einmal die in einer HN-Population resultierenden Payoffs der N-Typen und einer singulären und zufällig in der HN-Population auftauchenden K-Mutante. Unter der speziellen Annahme perfekter Information muß natürlich notwendigerweise gelten, daß deren Payoffs übereinstimmen, denn: Sowohl ein K-Typ als auch ein N-Typ identifizieren Defektoren und N-Typen mit absoluter Sicherheit, also werden sie beide immer die Outside-option realisieren, wie man der nachfolgenden Tabelle entnehmen kann:

	Hawk	N
N	$\lambda-c$	$\lambda-c$
K	$\lambda-c$	$\lambda-c$

Tab.12: Payoffs für K-Typ und N-Typ in einer reinen HN-Population

Wie stellt sich demgegenüber die Payoff-Situation in einer HN-Population unter imperfekter Information dar? Hier ist nun zu berücksichtigen, daß lediglich ein Teil der übrigen Spieler bekannt, weil in der Vergangenheit gelernt, ist.

	H <sup>K</sup>	H <sup>U</sup>	N <sup>K</sup>	N <sup>U</sup>
K <sup>K</sup>	$\lambda-c$	-c	$\lambda-c$	-c
K <sup>U</sup>	$\lambda-c$	-c	$\lambda-c$	-c
N <sup>K</sup>	$\lambda-c$	1-c	$\lambda-c$	$\lambda-c$
N <sup>U</sup>	$\lambda-c$	1-c	$\lambda-c$	1-c

Tab.13: Payoffs für K-Typ und N-Typ in einer HN-Population bei imperfekter Information

Man sieht klar: In den Fällen, in denen der Typ eines Gegenspielers bekannt ist, unterscheiden sich die Payoffs von N-Typ und K-Typ natürlich nicht. Beide realisieren weiterhin die Outside-option und erzielen den Payoff  $\lambda-c$ . In den übrigen Fällen jedoch differieren die Payoffs notwendigerweise. Während die grundsätzlich kooperative K-Mutante in jenen Fällen, in denen sie einen unbekanntem Spieler trifft, ausgebeutet wird und daher lediglich  $-c$  realisieren kann, erhält die defektive Mutante wenigstens den Defektionspayoff  $1-c$ . Damit ist klar, daß für K und N niemals Payoffgleichheit herrschen kann. Der HN-Fixpunkt kann im allgemeinen Fall imperfekten Wissens von einer einzelnen K-Mutante nicht erfolgreich invadiert werden und ist daher immer lokal stabil.

Ganz offensichtlich ist es in einer Population, die lediglich aus Defektoren (Hawks und Nasty mutants) besteht von Vorteil, zunächst einmal Defektion zu spielen und später, wenn Akteure gelernt wurden, die outside option zu wählen. Man vermeidet dadurch den Ausbeutungspayoff, den der K-Typ notwendigerweise tragen muß, solange er Spieler noch nicht kennt. Dieser Vorteil einer grundsätzlich defektiven Orientierung verschwindet lediglich dann, wenn man die unrealistische Annahme perfekter Information trifft. Dann verschwindet nämlich die Asymmetrie in den Auszahlungen für den Fall des Zusammentreffens mit unbekanntem Akteuren, weil es qua Annahme keine unbekanntem Akteure gibt. Damit aber ist klar: Die globale Stabilität des Dreitypen-Fixpunktes HDK auch im Viertypen-Modell mit einer N-Mutante muß letztlich als ein sehr spezielles Resultat beurteilt werden. Im allgemeinen ( $k > 0$ ) ist der HN-Fixpunkt stabil, und man kann sich, wie oben gezeigt wurde, nicht darauf verlassen, daß es nicht zu einer Koexistenz lokal stabiler Fixpunkte HN und HDK kommt. Die globale Stabilität des HDK-Fixpunktes ist im Viertypen-Modell mit imperfekter Information kein allgemeingültiges Resultat, sondern parameterabhängig. Als eine weitere, neue Lösung ergibt sich die Möglichkeit der Konvergenz gegen eine Gesellschaft, die lediglich aus den zwei defektiven Typen besteht.

Zusammenfassend kann also festgestellt werden: Die Einführung einer defektiven lernfähigen Mutante analog zum K-Typ zerstört die Ergebnisse des Basismodells nicht. Diese erweisen sich vielmehr als weitgehend robust. Die lokalen Stabilitätseigenschaften der Fixpunkte aus dem Basismodell bleiben überwiegend im Viertypen-Modell erhalten. Der Fixpunkt HDK ist weiterhin im ursprünglichen Parameterbereich lokal stabil. Allerdings taucht ein neuer Fixpunkt HN auf, der lokal stabil ist, sobald er existiert. Da eine Koexistenz lokal stabiler Fixpunkte HDK und HN nicht ausgeschlossen werden kann, folgt globale Stabilität nicht mehr zwingend aus der lokalen Stabilität des Fixpunktes HDK.

### 6.3 Appendix:

Bevor wir uns dem Beweis des Satzes 6.1 zuwenden ist es erforderlich, den Pay-off der N-Mutante zu bestimmen. Ist die N-Mutante ein bekannter Typ, so resultiert der erwartete Payoff als

$$E\pi_N^K = 6x_D + x_H + P_{H_K}(\lambda - 1) - \lambda(x_D + x_H - 1) - c \quad (6-45)$$

bzw.

$$E\pi_N^U = 6x_D + x_H + P_{H_K}(\lambda - 1) + 1 + 5x_K - x_D - x_H + P_{N_K}(\lambda - 1) - c \quad (6-46)$$

falls er ein unbekannter Spieler ist, wobei  $P_{H_K}, P_{N_K}$  die ex-ante Wahrscheinlichkeiten bezeichnen, einen bekannten Defektor bzw. einen bekannten N-Typ zu treffen.

Die Wahrscheinlichkeiten für den N-Typ, selbst bekannt bzw. unbekannt zu sein, sind gegeben als  $\frac{1}{x_N} P_{N_K}$  und  $\left(1 - \frac{1}{x_N} P_{N_K}\right)$ .

Der gesamte erwartete Payoff ermittelt sich damit als

$$E\pi_N = \left(6x_D + x_H + P_{H_K}(\lambda - 1) - \lambda(x_D + x_H - 1) - c\right) \left(\frac{1}{x_N} P_{N_K}\right) + \left(6x_D + x_H + P_{H_K}(\lambda - 1) + 1 + 5x_K - x_D - x_H + P_{N_K}(\lambda - 1) - c\right) \left(1 - \frac{1}{x_N} P_{N_K}\right), \quad (6-47)$$

was vereinfacht werden kann zu

$$E\pi_N = 5(x_D + x_K) + 1 + P_{H_K}(\lambda - 1) - c + P_{N_K}(\lambda - 1) - \frac{1}{x_N} P_{N_K}(\lambda(x_D + x_H - 1) + 1 + 5x_K - x_D - x_H) - \frac{1}{x_N} P_{N_K}^2(\lambda - 1). \quad (6-48)$$

Der durchschnittliche erwartete Payoff über die Zeit ergibt sich damit als



$$\begin{aligned} \bar{\pi}_N &= 5(x_D + x_K) + 1 + \bar{P}_{H_K}(\lambda - 1) - c + \bar{P}_{N_K}(\lambda - 1) - \\ &\frac{1}{x_N} \bar{P}_{N_K}(\lambda(x_D + x_H - 1) + 1 + 5x_K - x_D - x_H) - \frac{1}{x_N} \frac{1}{T} (\lambda - 1) \sum_1^T P_{N_K}^2. \end{aligned} \quad (6-49)$$

Die Wahrscheinlichkeiten  $P_H^K, \bar{P}_H^K, P_N^K, \bar{P}_N^K$  sind aus dem Lernmodell bekannt als

$$P_H^K = x_H \left( 1 - \left( \frac{n-1}{n} \right)^t \right), \bar{P}_H^K = x_H \left( 1 - \frac{n}{T} \left( 1 - \left( \frac{n-1}{n} \right)^T \right) \right), \quad (6-50)$$

$$P_N^K = x_N \left( 1 - \left( \frac{n-1}{n} \right)^t \right), \bar{P}_N^K = x_N \left( 1 - \frac{n}{T} \left( 1 - \left( \frac{n-1}{n} \right)^T \right) \right). \quad (6-51)$$

Mit  $\sum_1^T P_{N_K}^2 = \sum_1^T x_N^2 - \sum_1^T 2x_N^2 \left( \frac{n-1}{n} \right)^t + \sum_1^T x_N^2 \left( \frac{n-1}{n} \right)^{2t}$  kann der durchschnittliche Payoff auch geschrieben werden als

$$\begin{aligned} \bar{\pi}_N &= 5(x_D + x_K) + 1 + \bar{P}_{H_K}(\lambda - 1) - c + \bar{P}_{N_K}(\lambda - 1) - \frac{1}{x_N} \bar{P}_{N_K}(\lambda(x_D + x_H - 1) + 1 + 5x_K - x_D - x_H) - \\ &x_N(\lambda - 1) + 2x_N(\lambda - 1) \frac{1}{T} \sum_1^T \left( \frac{n-1}{n} \right)^t - x_N(\lambda - 1) \frac{1}{T} \sum_1^T \left( \frac{n-1}{n} \right)^{2t}. \end{aligned} \quad (6-52)$$

Wir führen nun den simultanen Grenzübergang für  $n, T$  durch. Aus Korollar 5.1 ist bekannt:  $\lim_{n, T \rightarrow \infty} \bar{P}_H^K = x_H(1 - k(\gamma))$ . Vollkommen analog ermittelt man

$$\lim_{n, T \rightarrow \infty} \bar{P}_N^K = x_N(1 - k(\gamma)). \text{ Weiterhin ist bekannt } \frac{1}{T} \sum_1^T \left( \frac{n-1}{n} \right)^t = \frac{n}{T} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right)$$

$$\text{sowie } \lim_{n, T \rightarrow \infty} \frac{n}{T} \left( 1 - \left( \frac{n-1}{n} \right)^{T+1} \right) = k(\gamma).$$

Wir betrachten nun den einzig neuen Term  $\frac{1}{T} \sum_1^T \left( \frac{n-1}{n} \right)^{2t}$ . Vollkommen analog zum Beweis des Lemmas 5.3 zeigt man leicht, daß

$$\lim_{n, T \rightarrow \infty} \frac{1}{T} \sum_1^T \left( \frac{n-1}{n} \right)^{2t} = \frac{1}{2\gamma} \left( 1 - \frac{1}{e^{2\gamma}} \right) \equiv k_2(\gamma). \quad (6-53)$$

Wir definieren nun

$$\beta(\gamma) \equiv \frac{k_2(\gamma)}{k(\gamma)} = \frac{\frac{1}{\gamma} \left( \frac{e^\gamma - 1}{e^\gamma} \right)}{\frac{1}{2\gamma} \left( \frac{e^{2\gamma} - 1}{e^{2\gamma}} \right)} = 2 \frac{e^\gamma (e^\gamma - 1)}{(e^{2\gamma} - 1)}. \quad (6-54)$$

Mit  $k_2(\gamma) = \beta(\gamma)k(\gamma)$  ergibt sich

$$\begin{aligned} \bar{\pi}_N^{lim} = & 1 + 5(x_D + x_K) - c + (\lambda - 1)(1 - k)(x_H + x_N) + \\ & (1 - k)((\lambda - 1)(1 - x_D - x_H) - 5x_K) - (\lambda - 1)x_N + (\lambda - 1)k2x_N - (\lambda - 1)kx_N\beta. \end{aligned} \quad (6-55)$$

Mit  $x_N = 1 - x_D - x_H - x_K$  erhalten wir schließlich

$$\begin{aligned} \bar{\pi}_N^{lim} = & x_D(6 - \lambda + \beta k(\lambda - 1)) + x_H(k(1 - \beta) + \beta k(\lambda - 1)) + \\ & x_K(k(6 - \lambda) + \beta k(\lambda - 1)) - \beta k(\lambda - 1) + \lambda - c. \end{aligned} \quad (6-56)$$

Das folgende Lemma enthält ein nützliches Resultat, von dem mehrfach in den Beweisen zu Existenz und Stabilität im 4-Typenmodell Gebrauch gemacht wird.

### Lemma 6.1:

Betrachte  $k = \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right)$ ,  $\beta = \frac{e^{2\gamma} - 1}{2(e^{2\gamma} - e^\gamma)}$ . Für alle  $\gamma > 0$ :  $k(\gamma) < \beta(\gamma)$ .

### Beweis:

$$\begin{aligned} \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) & < \frac{e^{2\gamma} - 1}{2(e^{2\gamma} - e^\gamma)} \Leftrightarrow \\ \gamma(e^{2\gamma} - 1) & > 2(e^{2\gamma} - e^\gamma) \left( 1 - \frac{1}{e^\gamma} \right) = 2e^\gamma(e^\gamma - 1) \left( 1 - \frac{1}{e^\gamma} \right) \Leftrightarrow \\ \gamma(e^\gamma - 1)(e^\gamma + 1) & > 2(e^\gamma - 1)(e^\gamma - 1) \Leftrightarrow \\ e^\gamma(\gamma - 2) & > -2 - \gamma \end{aligned} \quad (6-57)$$

i)  $\gamma < 2$ :

$$e^\gamma < \frac{2 + \gamma}{2 - \gamma} \quad (6-58)$$

ii)  $\gamma > 2$ :

$$e^\gamma > \frac{2 + \gamma}{2 - \gamma} \quad (6-59)$$

Wir betrachten zunächst (6-59). Klarerweise gilt  $\forall \gamma > 2: \frac{2+\gamma}{2-\gamma} < 0$ . Daher gilt

$$\text{immer: } e^\gamma > \frac{2+\gamma}{2-\gamma}.$$

Zu (6-58): Zunächst bemerkt man, daß beide Funktionen,  $k(\gamma)$  und  $\beta(\gamma)$  für  $\gamma \rightarrow 0$  gegen 1 streben. Wir beweisen diese Behauptung für  $k$ :

$$\begin{aligned} \lim_{\gamma \rightarrow 0} \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) &= \lim_{\gamma \rightarrow 0} \exp \left( \ln \left( \frac{1}{\gamma} \left( 1 - \frac{1}{e^\gamma} \right) \right) \right) = \lim_{\gamma \rightarrow 0} \exp \left( \ln \frac{1}{\gamma} + \ln(e^\gamma - 1) - \ln(e^\gamma) \right) \\ &= \lim_{\gamma \rightarrow 0} \exp(0 - \ln(\gamma) + \ln(e^\gamma - 1) - \gamma) \\ &= \lim_{\gamma \rightarrow 0} e^{-\gamma} e^{-\ln(\gamma)} e^{\ln(e^\gamma - 1)} \\ &= \lim_{\gamma \rightarrow 0} \frac{1}{e^\gamma} \frac{1}{e^{\ln(\gamma)}} e^{\ln(e^\gamma - 1)} \\ &= 1 \frac{\exp(\ln(0))}{\exp(\ln(0))} \\ &= 1 \end{aligned}$$

Vollkommen analog zeigt man  $\lim_{\gamma \rightarrow 0} \beta(\gamma) = 1$ . Beide Funktionen sind dann abnehmend in  $\gamma$ . Als nächstes zeigt man, daß für alle  $0 < \gamma < 2$   $k(\gamma)$  immer steiler verläuft als  $\beta(\gamma)$ , damit also im gesamten betrachteten Intervall gilt  $k(\gamma) < \beta(\gamma)$ . (6-58) war gegeben als

$$\begin{aligned} e^\gamma < \frac{2+\gamma}{2-\gamma} &\Leftrightarrow \\ \gamma < \ln \frac{2+\gamma}{2-\gamma} &\equiv f(\gamma). \end{aligned} \quad (6-60)$$

Differenzieren liefert

$$\begin{aligned} f'(\gamma) > 1 &\Leftrightarrow \\ \frac{2-\gamma}{2+\gamma} \left( \frac{4}{(2-\gamma)^2} \right) &= \frac{4}{(2+\gamma)(2-\gamma)} > 1 \Leftrightarrow \\ \gamma > 0. & \end{aligned} \quad (6-61)$$

□

## Beweis des Satzes 6.1

i) Auflösen von  $\dot{x}_i = 0$ ,  $i = D, H, K, N$  nach  $x_D, x_H, x_K$  liefert

$$x_H^* = \frac{\lambda(c(k(2-\beta)) - 1) + ck(\beta - 1)}{\lambda(\lambda - 1)(k^2(\beta - 2) + k(3 - \beta) - 1)}, \quad (6-62)$$

$$x_K^* = \frac{\lambda(k^2(3\beta-6)+k(9-3\beta)-3)+ck(k-6+2\beta)+3c}{\lambda((k^2(\beta-2)+k(3-\beta)-1)(k-1)\lambda+6(1-k))}, \quad (6-63)$$

$$x_D^* = \frac{\lambda^2(k^3(\beta-2)+k^2(5-2\beta)+k(\beta-4)+1)+\lambda(k^3(12-6\beta)+k^2(\beta(9+c)-2c-24))}{\lambda(k-1)((k-1)\lambda+6(1-k))(k(\beta-2)+1)} + \frac{\lambda(k((15-\beta(3+c)+3c)-c-3))+k^2c(11-6\beta)+kc(4\beta-12)+3c}{\lambda(k-1)((k-1)\lambda+6(1-k))(k(\beta-2)+1)}. \quad (6-64)$$

Die Summation der Populationsanteile ergibt

$$\frac{\lambda^2(k(\beta-2)+1)-\lambda(k(\beta-2)+1)+c}{\lambda^2(k(\beta-2)+1)-\lambda(k(\beta-2)+1)}, \quad (6-65)$$

was weiter vereinfacht werden kann zu

$$1 + \frac{c}{\lambda((k(\beta-2)+1)(\lambda-1))}. \quad (6-66)$$

Man beachte, daß  $\frac{c}{\lambda((k(\beta-2)+1)(\lambda-1))}$  immer positiv ist, da  $(k(\beta-2)+1)$  niemals negativ werden kann:

$$(k(\beta-2)+1) < 0 \Leftrightarrow k > \frac{1}{2-\beta} \quad (6-67)$$

Man beachte jedoch, daß immer  $\frac{1}{2-\beta} > \beta$  gilt, so daß  $k > \frac{1}{2-\beta}$  niemals erfüllt ist (vgl. Lemma 6.1). Damit gilt  $x_D^* + x_H^* + x_K^* > 1$ .

ii) Der Fixpunkt  $\{x_D^*, x_K^*, x_N^*\}$  berechnet sich als:

$$\{x_D^*, x_K^*, x_N^*\} = \left\{ \begin{array}{l} \frac{\lambda^2(1-k)^2 + \lambda(k(9-6k+2c-c\beta)-c-3)+c(3-k(6-\beta))}{(1-k)^2(\lambda-6)}, \\ \frac{\lambda k(3-c(1-\beta))+c(3-k\beta)-3\lambda}{\lambda(k-1)(k-1)(\lambda-6)}, 1-x_D^*-x_K^* \end{array} \right\} \quad (6-68)$$

Existenz erfordert  $x_D^*, x_K^*, x_N^* > 0$ . und  $x_D^* + x_K^* < 1$ .

$$a) x_D^* > 0$$

$$i) 3 - \lambda(1 - k(2 - \beta)) + k(\beta - 6) > 0$$

Dann:

$$x_D^* > 0 \Leftrightarrow c < \frac{\lambda(6k^2 - 9k + 3) - \lambda^2(1 - k)^2}{3 - \lambda(1 - k(2 - \beta)) + k(\beta - 6)} \quad (6-69)$$

$$ii) 3 - \lambda(1 - k(2 - \beta)) + k(\beta - 6) < 0$$

Dann:

$$x_D^* > 0 \Leftrightarrow c > \frac{\lambda(6k^2 - 9k + 3) - \lambda^2(1 - k)^2}{3 - \lambda(1 - k(2 - \beta)) + k(\beta - 6)} \quad (6-70)$$

$$b) x_K^* > 0 \Leftrightarrow$$

$$c(3 - k(\lambda(1 - \beta) + \beta)) < 3\lambda(1 - k)$$

Der Fall  $(3 - k(\lambda(1 - \beta) + \beta)) < 0$  ist irrelevant, da er  $\beta < 0$  impliziert. Dies zeigt man leicht:

$$(3 - k(\lambda(1 - \beta) + \beta)) < 0 \Leftrightarrow k > \frac{3}{\lambda(1 - \beta) + \beta} \quad (6-71)$$

$k$  kann nicht größer als 1 werden, also

$$\frac{3}{\lambda(1 - \beta) + \beta} < 1 \Leftrightarrow \lambda > \frac{3 - \beta}{1 - \beta}. \quad (6-72)$$

Klar:

$$\frac{3 - \beta}{1 - \beta} < 3 \Leftrightarrow \beta < 0 \quad (6-73)$$

Daher ist nur der Fall zu betrachten

$$(3 - k(\lambda(1 - \beta) + \beta)) > 0 \quad (6-74)$$

Damit:

$$x_K^* > 0 \Leftrightarrow c < \frac{3\lambda(1 - k)}{(3 - k(\lambda(1 - \beta) + \beta))} \quad (6-75)$$

$$c) x_D^* + x_K^* < 1$$

Es zeigt sich, dass diese Forderung für alle  $c, k, \lambda$  erfüllt ist, da

$$x_D^* + x_K^* < 1 \Leftrightarrow \frac{c}{\lambda(k - 1)} < 0. \quad (6-76)$$

Bezeichne nun  $\frac{\lambda(6k^2 - 9k + 3) - \lambda^2(1 - k)^2}{3 - \lambda(1 - k(2 - \beta)) + k(\beta - 6)} = c_1, \frac{3\lambda(1 - k)}{3 - k(\lambda(1 - \beta) + \beta)} = c_2.$

Ein Vergleich dieser beiden kritischen Werte zeigt dann:

$$\frac{3\lambda(1-k)}{3-k(\lambda(1-\beta)+\beta)} < \frac{\lambda(6k^2-9k+3)-\lambda^2(1-k)^2}{3-\lambda(1-k(2-\beta))+k(\beta-6)} \Leftrightarrow$$

$$\frac{\beta}{\beta-1} < \lambda < 6 \quad (6-77)$$

Daher  $x_D^* > 0$  und  $x_K^* > 0$  simultan dann und nur dann, wenn  $c < c_2 < c_1$ . Dies vervollständigt den Beweis zur Existenz des Fixpunktes  $\{x_D^*, x_K^*, x_N^*\}$ .

iii) Wir betrachten zunächst den Fixpunkt

$$\{x_K^*, x_N^*\} = \left\{ \frac{(\lambda(1-\beta)+\beta)k}{(\lambda(2-\beta)+\beta-6)+3-\lambda} \right\}.$$

a)  $x_K^* > 0$

Da  $(\lambda(1-\beta)+\beta)k > 0$ ,  $x_K^* > 0 \Leftrightarrow (\lambda(2-\beta)+\beta-6)k+3-\lambda > 0$ . Der Fall  $(\lambda(2-\beta)+\beta-6) > 0$  ist irrelevant, da er  $\lambda > 3$  impliziert. Dies zeigt man leicht:

$$(\lambda(2-\beta)+\beta-6) > 0 \Leftrightarrow \lambda > \frac{6-\beta}{2-\beta} > 3 \quad (6-78)$$

Daher:

$$x_K^* > 0 \Leftrightarrow k < \frac{3-\lambda}{6-\beta-\lambda(2-\beta)} \quad (6-79)$$

b)  $x_K^* < 1$ , was immer dann der Fall ist, wenn  $k < \frac{3-\lambda}{6-\lambda}$ . Es bleibt der Vergleich

der kritischen Werte:

$$\frac{3-\lambda}{6-\lambda} < \frac{3-\lambda}{6-\beta-\lambda(2-\beta)} \Leftrightarrow$$

$$6-\lambda > 6-\beta-\lambda(2-\beta) \Leftrightarrow$$

$$\lambda > -\frac{\beta}{1-\beta} < 0 \quad (6-80)$$

Damit ist gezeigt, daß  $x_K^* < 1$  immer restriktiver ist als  $x_K^* > 0$ .

Betrachtet sei nun der Fixpunkt

$$\{x_H^*, x_N^*\} = \left\{ \frac{k(1-\beta)(\lambda-1)-c}{(k(2-\beta)-1)(\lambda-1)}, 1-x_H^* \right\}.$$

a)  $x_H^* > 0$

Zunächst stellt man fest, dass der Fall  $(k(2-\beta)-1) > 0$  irrelevant ist, da er

$k > \frac{1}{2-\beta} > \beta$  impliziert, ein Widerspruch zu Lemma 6.1. Daher

$$x_H^* > 0 \Leftrightarrow k(1-\beta)(\lambda-1)-c < 0 \Leftrightarrow c > k(1-\beta)(\lambda-1) \quad (6-81)$$

b)  $x_H^* < 1$

$$\begin{aligned} \Leftrightarrow k(1-\beta)(\lambda-1)-c &> (k(2-\beta)-1)(\lambda-1) \\ \Leftrightarrow (\lambda-1)-c &> k(\lambda-1) \\ \Leftrightarrow c &< (\lambda-1)(1-k) \end{aligned} \quad (6-82)$$

## 7. Überleben von Kooperation in einer stochastischen Welt

### 7.1 Motivation

Die bisherige Analyse der Überlebensfähigkeit kooperativer Verhaltensweisen bewegte sich im Standard-Modellrahmen der evolutionären Spieltheorie, der Replikator-Dynamik, und damit im Rahmen eines gänzlich deterministischen Konzepts. Wir hatten bereits früher ausführlich darauf hingewiesen, daß die beiden zentralen Bausteine des Modells – das Lernmodell einerseits, das dynamische Modell der Replikator-Dynamik andererseits – möglicherweise miteinander konfliktieren. Soll Lernen meßbare Effekte zeitigen, so muß bei einer endlichen Lebensdauer einer Generation auch die Populationsgröße als endlich angenommen werden. Die dynamische Analyse jedoch verlangt – jedenfalls in ihrer formal strengen Variante – die Annahme einer unendlich großen Population. Trivialerweise kann Lernen begrenzter Dauer in einer infiniten Population keinerlei Effekte zeitigen – es wird kein verwertbares Wissen generiert. Wir hatten einigen Raum darauf verwendet plausibel zu machen, daß es dennoch möglich ist, auch im Rahmen des rein deterministischen Modellrahmens, beide Modellelemente zu vermitteln. Am überzeugendsten gelingt dies vielleicht mit dem Basismodell, in dem eine unendliche Anzahl finiter Subpopulationen angenommen wird.

Dennoch bleibt als unumstößliche Tatsache bestehen, daß reale Populationen endlich sind. Die Annahme einer infiniten (Gesamt-) Population ist – ungeachtet des Faktums, daß es sich um eine Standardannahme der evolutionären Spieltheorie handelt – eine fragwürdige Fiktion. Gerade im Hinblick auf eine „realistische“ Interpretation des Modells dürfte dieses Problem besonders an Gewicht gewinnen: Entwicklungsgeschichtlich dürfte die Herausbildung von Verhaltenstypen mit kooperativer Disposition eher früh zu verorten sein. Spätestens bei der Vorstellung archaischer Jäger- und Sammlergesellschaften, überschaubaren Sippen- und Stammesverbänden, erscheint die Annahme infiniten Populationen kaum hinnehmbar. Ein größerer Realismus unseres Modells kann nur erzielt werden durch die Analyse von explizit endlichen Populationen.

Neben dem Spezialproblem der Generierung meßbarer Effekte von Lernen in unendlich großen Gesamtpopulationen, ist der evolutionär- deterministische Modellrahmen jedoch auch aus weiteren Gründen heraus kritisiert worden. Insbesondere, so wird beanstandet, fange der deterministische Ansatz das für jeden realen evolutionären Prozeß charakteristische Phänomen der Mutation nicht, oder bestenfalls ungenügend, ein. In der Tat werden „Mutationen“ im Standardansatz lediglich im Rahmen der Stabilitätsanalyse thematisiert. Bei der Untersuchung der Frage lokal asymptotischer Stabilität von Fixpunkten wird analysiert, ob die Trajektorien des



betrachteten dynamischen Systems nach kleinen Auslenkungen aus dem Fixpunkt heraus in diesen zurückkehren. Hierbei handelt es sich jedoch bloß um isolierte, singuläre Schocks, die eine nur temporäre Auslenkung aus dem Gleichgewicht verursachen. Foster und Young (1990) zählen zu den ersten, die der Frage nachspürten, was eigentlich geschieht, wenn dynamische Systeme permanenten Schocks ausgesetzt werden. In der Arbeit wird der Nachweis geführt, daß die sogenannten stochastisch stabilen Zustände (stochastically stable states, SSS, d.h. die langfristig zu erwartenden Gleichgewichtszustände des gestörten dynamischen Systems) keineswegs mehr notwendigerweise mit den Gleichgewichtszuständen des ungestörten Systems koinzidieren, also mit den evolutionär stabilen Strategien (ESS), bzw. den Attraktoren (in der Regel Fixpunkte oder Grenzzyklen) des dynamischen Systems. Die Intuition für dieses Resultat ist simpel. Wie Foster und Young schreiben:

*„[...] if each successive perturbation dies out (or almost dies out) before the next one is felt, then obviously the system can never move very far from its „equilibrium“ position“.* (ebd., S. 220).

Andersherum gesagt: Eine Folge kontinuierlicher Schocks kann kumulieren und damit die Trajektorien eines dynamischen Systems sukzessive von einer Gleichgewichtsposition wegführen. A priori spricht nun nichts für die Annahme, daß reale evolutionäre Prozesse nicht permanenten Störungen unterliegen. Angesichts der möglicherweise weitreichenden Konsequenzen sollten diese daher in der Analyse des Langfrist-Verhaltens dynamischer Systeme berücksichtigt werden. Foster und Young modellieren den daraus resultierenden „noise“ direkt auf der aggregierten Ebene, und zwar als einen Wiener-Prozeß, der die deterministische Replikator-Dynamik überlagert. Die Verwendung eines Wiener Prozesses ist aber, wie Foster und Young einräumen, nur dann statthaft, wenn sowohl die Anzahl der Individuen, als auch die Anzahl der Interaktionen „groß“ sind (was intuitiv sofort einsichtig wird, wenn man bedenkt, daß der Wiener Prozeß ursprünglich verwendet wurde, um eine Brownsche Bewegung in einer Flüssigkeit abzubilden). Benötigt wird also wiederum die Annahme einer großen Population. Diese Annahme kann, wie bereits ausführlich erläutert wurde, in unserem Modellkontext durchaus verletzt sein, bzw. es erscheint wenig sinnvoll, diese Annahme zu treffen, wenn man ausdrücklich an der Analyse endlicher, möglicherweise auch kleiner Populationen interessiert ist.

Im folgenden soll daher ein Modell entwickelt werden, das sowohl dem Umstand endlicher Populationen Rechnung trägt, als auch das Phänomen von „Mutation“ adäquat berücksichtigt. Wir orientieren uns dabei insbesondere an der Pionierar

beit von Kandori, Mailath und Rob (KMR) (1993) sowie den Arbeiten von Vega-Redondo (VR) (1995, 1996). Im Gegensatz zu Foster und Young (1993) modellieren KMR noise auf der Mikroebene. Sie gehen davon aus, daß die Mitglieder der betrachteten Population mit einer gewissen, kleinen Wahrscheinlichkeit unabhängig voneinander „mutieren“. Grundsätzlich, so wird weiterhin angenommen, entwickeln sich die Populationsanteile der unterschiedlichen Strategien gemäß einer deterministischen Dynamik, die gewisse sinnvolle, d.h. „darwinistische“ Eigenschaften aufweist. Insbesondere soll niemals gegen die erfolgreichste Strategie selektiert werden. Beide Komponenten zusammen genommen, Selektion und Mutation, generieren dann einen stochastischen Prozeß auf einem endlichen Zustandsraum, über den, wie wir sehen werden, sich erstaunlich allgemeine Aussagen treffen lassen. KMR und VR haben die hier nur sehr grob skizzierte Modellierung für 2x2-Spiele, insbesondere Koordinierungsspiele entwickelt. Wir werden ihre Technik im folgenden auf unser 3x3 Kooperationsproblem anwenden.

## 7.2 Ein stochastisches Modell der Kooperation

Betrachtet sei im folgenden eine endliche Population der Gesamtgröße  $n$ , die sich wiederum aus den drei Typen Hawk, Dove und Knowledgeable Dove zusammensetzt. Wir nehmen an, daß die Population in einer Anzahl von  $t=1, \dots, m$  Zeitpunkten zufällig paarweise gematcht wird. Es bezeichnen:  $P_{HH}$  die Anzahl der Paarungen von Hawks untereinander, entsprechend  $P_{HD}$  die Anzahl der Paarungen von Hawks mit Doves. Es bezeichnen ferner  $P_{H,K}$  die Anzahl aller Paarungen, in denen K-Typen auf Hawks treffen und in denen die Hawks bekannt waren, sowie entsprechend  $P_{H,U,K}$  die Anzahl aller Paarungen von K-Typen mit unbekanntem Hawks. Ferner geben  $P_{KK}$  und  $P_{DD}$  die Zahl der Paarungen von K-Typen und D-Typen untereinander, sowie  $P_{KD}$  die Zahl der Paarungen von K-Typen mit Doves an. Die Payoff-Matrix ist gegeben durch Tabelle 9, die hier zur Erinnerung noch einmal wieder gegeben wird.

	Dove	Hawk (kn.)	Hawk (unkn.)	K-Typ
Dove	3	0	0	3
Hawk (known)	6	1	1	$\lambda$
Hawk (unknown)	6	1	1	6
K-Typ	3-c	$\lambda$ -c	-c	3-c

Tabelle 9: Gefangenen Dilemma mit lernfähiger Mutante

Wir können uns die Anzahl der matchings  $m$  wiederum vorstellen als die Lebensdauer einer Generation. Wir wollen nun weiterhin an der darwinistischen Annahme des survival of the fittest festhalten. Die Annahmen bezüglich des deterministischen Teils des evolutionären Prozesses sollen dabei möglichst allgemein gehalten werden. So soll lediglich gefordert werden, daß zwischen dem realisierten durchschnittlichen Payoff der Strategien und ihrem reproduktiven Erfolg eine monotone Beziehung existiert. Bezeichne  $I = \{H, D, K\}$  die Strategiemenge und  $i = \{h, d, k\}$  die Anzahlen der Spieler in der Population, die die Strategien H,D,K spielen. Dann soll gelten:

$$\Delta i \equiv i(t+1) - i(t) > \Delta i' > \Delta i'', \text{ genau dann wenn } \bar{\pi}_i > \bar{\pi}_{i'} > \bar{\pi}_{i''}. \quad (7-1)$$

Zusätzlich zu dieser deterministischen Selektionskomponente soll der evolutionäre Prozeß durch einen gewissen noise, also stochastische Einflüsse charakterisiert sein. Den Ansätzen von KMR und VR folgend sei angenommen, daß die Spieler der neuen Generation mit einer positiven Wahrscheinlichkeit mutieren. Dies ist nun freilich nicht in einem biologischen Sinne zu verstehen. Man stelle sich vielmehr vor, daß die Kinder der Elterngeneration den Erfolg der jeweiligen von ihren Eltern verfolgten Strategie beobachten können, bzw. daß Eltern die von ihnen gemachten Erfahrungen in Form gutgemeinter Ratschläge an die Folgegeneration weitergeben. Dann kann man sich vorstellen, daß zunächst einmal grundsätzlich die Entwicklung der Populationsanteile der Strategien, bzw. Typen, durch einen Prozeß wie (7-1) der Tendenz nach beschrieben werden kann. Nun stelle man sich vor, daß die Kindergeneration, wenn sie ihrerseits, nach dem Tod der Elterngeneration, in das Kooperationsspiel<sup>15</sup> eintritt, mit einer gewissen Wahrscheinlichkeit Fehler macht. Oder aber man stelle sich vor, daß Kinder den Ratschlägen ihrer Eltern nicht trauen und durch bewußtes Experimentieren mit Verhaltensweisen eigene Erfahrungen sammeln möchten. „Mutation“ von Spielern soll im folgenden in diesem Sinne, also entweder als Folge von fehlerhaften, irrtümlichen Entscheidungen oder als Folge bewußten Experimentierens aufgefaßt werden. Sei die unabhängige und für alle Spieler identische Mutationswahrscheinlichkeit mit  $\varepsilon$  bezeichnet und gelte  $\varepsilon > 0$ . Um die zeitliche Struktur des Modells noch einmal zusammen zu fassen: Nach den  $t=1, \dots, m$  endlichen matching-Zeitpunkten wird also der durchschnittlich realisierte Payoff der jeweiligen Stra-

---

<sup>15</sup> Das eigentliche 3x3 GD-Spiel wird immer nur von einer Generation gespielt. Zu überlappenden Generationen soll es in unserem Modell also nur etwa in dem Sinne kommen, daß Kinder eine zeitlang ihre Eltern und deren Erfolg oder Mißerfolg beobachten können. Kinder nehmen aber während der Lebenszeit der alten Generation nicht am GD teil.

tegien ermittelt. In einem Prozeß kultureller Vererbung werden die Strategien gemäß einer payoffmonotonen Dynamik selektiert. Nachdem die Nachfolgegengeneration „programmiert“ worden ist, mutieren die Spieler, d.h. sie begehen mit Wahrscheinlichkeit  $\epsilon > 0$  Fehler oder experimentieren.

Die Einführung von noise in der beschriebenen Weise hat erhebliche Konsequenzen für das langfristige Verhalten des evolutionären Prozesses. Zunächst muß geklärt werden, was unter einem gleichgewichtigen Zustand des evolutionären Prozesses verstanden werden soll. Betrachten wir zunächst nur den deterministischen Teil der Dynamik. Ohne Mutationen ist genau dann kein Anlaß mehr zu Änderungen der Strategiewahl gegeben, wenn keine Payoffdifferenzen mehr herrschen. Dies ist natürlich gerade in den drei monomorphen Populationszuständen der Fall, in denen alle Mitglieder der Population entweder H, K oder D spielen. Entscheidend ist, daß der Prozeß, wenn er einmal einen dieser Zustände erreicht hat, dort verharret. Natürlich kann die Payoffgleichheit der Strategien grundsätzlich auch gegeben sein für einen inneren, also gemischten Populationszustand. Allerdings ist hier folgendes zu beachten: Payoffgleichheit ist in diesem Fall nur dann zu erwarten, wenn zwischen den matching-Zufallsvariablen ganz bestimmte Beziehungen herrschen. Im allgemeinen kann man jedoch nicht erwarten, daß diese Beziehungen im Zeitablauf stabil sind, also für eine beliebig große Anzahl zukünftiger Zeitpunkte immer erfüllt sein werden. Dieses Ereignis ist extrem unwahrscheinlich. Wie später noch gezeigt werden wird, konvergiert die mutationsfreie Dynamik tatsächlich mit positiver Wahrscheinlichkeit gegen einen der monomorphen Populationszustände. Ohne Mutationen würde der evolutionäre Prozeß klarerweise in dem jeweils erreichten monomorphen Zustand verharren. Unter der Berücksichtigung der Möglichkeit von Mutationen ist es nun jedoch möglich, daß auch ein solcher zuvor absorbierender Zustand wieder verlassen werden kann.

Insgesamt handelt es sich bei dem resultierenden Prozeß um eine Markov-Kette auf einem diskreten, endlichen Zustandsraum. Dies folgt einfach aus der Tatsache, daß die Entwicklung der Populationsanteile lediglich vom unmittelbar vorherigen Erfolg der Strategien abhängen soll. Man kann jedoch noch weit mehr aussagen. Die Möglichkeit von Mutationen, eventuell auch mehrerer simultaner Mutationen, führt dazu, daß nunmehr jeder Übergang eine positive, wenn auch möglicherweise sehr kleine Wahrscheinlichkeit aufweist. Technisch gesprochen weist die Markov-Matrix aller Übergangswahrscheinlichkeiten nur strikt positive Einträge auf. Dies bedeutet, daß die Markovkette aperiodisch und irreduzibel ist, und dies wiederum hat die bedeutende analytische Konsequenz, daß das langfristige Verhalten unseres Prozesses eindeutig charakterisiert werden kann. Für aperiodische und irreduzible Prozesse ist nämlich wohlbekannt, daß diese eine eindeutige stationäre Ver

teilung aufweisen, die zudem von den Startbedingungen des Prozesses unabhängig ist. Diese hier sehr kompakt vorgetragene Argumentation soll im folgenden etwas detaillierter entwickelt werden.

Zunächst sind einige definitorische Vorarbeiten vonnöten. Die folgenden Definitionen orientieren sich eng an Fahrmeier, Kaufmann, Ost (1980, S. 14 ff.).  $t$  indiziere im folgenden die Zeit,  $X_t$  bezeichne die Realisation des Prozesses im Zeitpunkt  $t$ ,  $i$  und  $j$  bezeichnen Zustände, die der Prozeß annehmen kann.

**Definition (7.1):** (Markov-Kette)

Ein stochastischer Prozeß (SP)  $X = \{X_t, t \in IN_0\}$  mit abzählbarem Zustandsraum  $E$

heißt *MARKOV-KETTE* (MK):  $\Leftrightarrow$

$$\forall t \in IN_0, \forall j, i, i_{t-1}, \dots, i_0 \in E$$

$$P\{X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0\} = P\{X_{t+1} = j | X_t = i\}$$

„Eine MK ist also ein SP, bei dem der „nächste“ Zustand  $X_{t+1}$  bei bekanntem „gegenwärtigem“ Zustand  $X_t$  unabhängig von den „vergangenen“ Zuständen  $X_{t-1}, \dots, X_0$  ist.“ (Fahrmeier, Kaufmann, Ost (FKO)1980, S.14).

**Definition (7.2):** (Übergangswahrscheinlichkeit)

Die bedingte Wahrscheinlichkeit  $p_{ij}(t+1, t) := P\{X_{t+1} = j | X_t = i\}$

heißt einschrittige Übergangswahrscheinlichkeit (ÜW) von  $i$  nach  $j$ .

**Definition (7.3):** (Homogenität)

Eine MK  $X$  heißt *homogen*:  $\Leftrightarrow$  Die einschrittigen ÜW sind von  $t$  unabhängig, d.h.

$$p_{ij}(t+1, t) = p_{ij}. X \text{ heißt dann auch MK mit stationären ÜW } p_{ij}.$$

Bezeichnung: Die ÜW  $p_{ij}$  werden in einem quadratischen Feld  $P = (p_{ij})$

angeordnet.  $P$  heißt *Übergangsmatrix* (ÜM) der MK  $X$ . Für

$E = \{0, 1, 2, \dots, k\}$  ist

$$P = (p_{ij}) = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \dots & p_{0k} \\ p_{10} & p_{11} & p_{12} & \dots & p_{1k} \\ p_{20} & p_{21} & p_{22} & \dots & p_{2k} \\ \vdots & & & & \vdots \\ p_{k0} & p_{k1} & p_{k2} & \dots & p_{k0} \end{bmatrix}$$

$P$  besitzt folgende Eigenschaften:

$$\forall i, j \in E : p_{ij} \geq 0$$

$$\forall i \in E : \sum_{j \in E} p_{ij} = 1 \quad (\text{Zeilensumme 1}) \tag{7-2}$$

Im folgenden ist nur von homogenen Markov-Ketten die Rede.

**Definition (7.4):** (stochastische Matrix)

Jede  $|E| \times |E|$ -Matrix mit den Eigenschaften (7-2) heißt *stochastische Matrix* oder *Markov-Matrix*.

**Definition (7.5):** (t-schrittige Übergangswahrscheinlichkeiten)

Die (von s unabhängigen) bedingten Wahrscheinlichkeiten für einen Übergang von i nach j in t Zeitschritten  $p_{ij}^{(t)} = P\{X_{t+s} = j | X_s = i\}$ ,  $t \geq 1$  heißen t-schrittige Übergangswahrscheinlichkeiten (oder ÜW t-ter Ordnung).

**Definition (7.6)** (Zustandsverteilung):

$\pi_i(t) = P\{X_t = i\}$ ,  $i \in E$  heißen *Zustandswahrscheinlichkeiten* und der Zeilenvektor  $\pi(t) = (\pi_i(t))$  heißt *Zustandsverteilung*.

Bemerkung: Für  $t=0$  heißt die Zustandsverteilung *Anfangsverteilung*.

Die Zustandswahrscheinlichkeiten geben also die Gesamtwahrscheinlichkeit dafür an, daß der Prozeß sich im Zustand i aufhält.

Die folgende Klassifikation von Zuständen ist ausgesprochen hilfreich.

**Definition (7.7):** (Erreichbarkeit)

Ein Zustand j heißt vom Zustand i aus *erreichbar*, (in formaler Notation:  $i \rightarrow j$ ):

$$\Leftrightarrow \exists t \in \mathbb{N}_0 \text{ mit } p_{ij}^{(t)} > 0.$$

**Definition (7.8):** (Gegenseitige Erreichbarkeit)

Zwei Zustände i und j heißen *gegenseitig erreichbar*, (in formaler Notation:  $i \leftrightarrow j$ ):

$$\Leftrightarrow i \rightarrow j \wedge j \rightarrow i.$$

**Definition (7.9):** (Klassifikation gemäß Erreichbarkeit)

- Eine Menge C von Zuständen heißt *abgeschlossen*:  $\Leftrightarrow$  Kein Zustand aus  $E \setminus C$  ist von C aus erreichbar. Nicht abgeschlossene Mengen heißen offene Mengen.
- Ein Zustand i heißt *absorbierend*:  $\Leftrightarrow \{i\}$  abgeschlossen  $\Leftrightarrow p_{ii} = 1$ .
- Eine abgeschlossene Menge heißt *irreduzibel*:  $\Leftrightarrow$  Es gibt keine echte Teilmenge, die abgeschlossen ist.
- Eine MK heißt *irreduzibel*:  $\Leftrightarrow E$  irreduzibel, d.h. alle Zustände sind gegenseitig erreichbar.

**Definition (7.10):** ((Rückkehrzeit)

Die Zufallsvariable  $T_{ij}$  mit  $T_{ij}(\omega) = \begin{cases} \min\{s \in \mathbb{N} : X_s(\omega) = j | X_0(\omega)\} \\ \infty \text{ falls min nicht existiert} \end{cases}$

heißt für  $i=j$  Rückkehrzeit  $T_{ii}$  des Zustandes  $i$  und für  $i \neq j$  (erste) Übergangszeit von  $i$  nach  $j$ .

**Definition (7.11):** (Rückkehrwahrscheinlichkeit)

Die Wahrscheinlichkeit

$$f_{ij} := P\{T_{ij} < \infty | X_0 = i\} = P\{\exists t \text{ mit } X_t = j | X_0 = i\} = \sum_{t \in \mathbb{N}} f_{ij}^{(t)}$$

dafür, daß die in  $i$  startende MK in endlicher Zeit  $j$  erreicht, heißt *Prozeßübergangswahrscheinlichkeit*. Speziell für  $i=j$  heißt  $f_{ij}$  *Rückkehrwahrscheinlichkeit*.

Die vorstehende Definition erlaubt uns eine Klassifikation der Zustände nach dem Rückkehrverhalten der MK einzuführen:

**Definition (7.12):** (Klassifizierung gemäß Rückkehrverhalten)

Sei  $i \in E$  ein beliebiger Zustand,  $T := T_{ii}$ .

a)  $i$  heißt *rekurrent*:  $\Leftrightarrow P\{T < \infty | X_0 = i\} = f_{ii} = 1$ .

b)  $i$  heißt *transient*:  $\Leftrightarrow P\{T < \infty | X_0 = i\} = f_{ii} < 1$

c) Innerhalb der rekurrenten Zustände wird weiter differenziert in positiv-rekurrenente und null-rekurrenente Zustände.

$$i \text{ heißt } \textit{positiv-rekurrent}: \Leftrightarrow \mu_{ii} := E(T | X_0 = i) = \sum_{t=1}^{\infty} t f_{ii}^{(t)} < \infty$$

$$i \text{ heißt } \textit>null-rekurrent}: \Leftrightarrow \mu_{ii} := E(T | X_0 = i) = \infty$$

$\mu$  heißt *erwartete Rückkehrzeit* oder auch *mittlere Wiederkehrzeit*.

d)  $i$  heißt *periodisch mit Periode  $d$* :

$d \geq 2$  ist die größte natürliche Zahl, für die

$$\sum_{n=1}^{\infty} P\{T = nd | X_0 = i\} + P\{T = \infty | X_0 = i\} = 1$$

gilt. Andernfalls heißt  $i$  *aperiodisch*.

e)  $i$  heißt *ergodisch*:  $\Leftrightarrow i$  aperiodisch und positiv rekurrent.

Zur Charakterisierung des langfristigen Verhaltens von stochastischen Prozessen ist der Begriff der stationären Verteilung hilfreich.

**Definition (7.13):** (Stationäre Verteilung)

Sei  $P$  die Übergangsmatrix einer Markov-Kette. Der Zeilenvektor  $\pi = (\pi_j, j \in E)$  heißt *stationäre Verteilung* dieser MK:

$$\Leftrightarrow \pi = \pi P \text{ und } \pi(1,1,1,\dots,1) = 1$$

Der nicht unerhebliche definitorische Aufwand wird verständlich und gerechtfertigt durch das folgende, wohlbekanntes Resultat für ergodische Markov-Ketten.

**SATZ (7.1):** (Grenzwertsatz für ergodische Markov-Ketten)

Für eine irreduzible und aperiodische MK gilt:

Falls alle Zustände *positiv rekurrent* sind, dann besitzt das lineare Gleichungssystem

$$\begin{aligned} \pi_j &= \sum_{i \in E} \pi_i p_{ij}, \quad j \in E \Leftrightarrow \pi = \pi P, \quad \pi = (\pi_j), \quad j \in E \\ \sum_{j \in E} \pi_j &= 1 \Leftrightarrow \pi(1,1,1,\dots,1) = 1 \end{aligned} \tag{7-3}$$

eine eindeutige, strikt positive ( $\pi_j > 0 \forall j \in E$ ) Lösung und es gilt

$$\pi_j = \lim_{t \rightarrow \infty} p_{ij}^{(t)} = \frac{1}{\mu_{jj}}, \text{ sowie}$$

$$\lim_{t \rightarrow \infty} p(t) = \lim_{t \rightarrow \infty} p(0)P^t = p(0)P^\infty = \pi$$

für jede beliebige Anfangsverteilung  $p(0)$ .

**Beweis:** siehe jedes Standard-Lehrbuch über stochastische Prozesse, beispielsweise Karlin und Taylor (1996), S. 85 f.

Weil endliche und irreduzible MK nur positiv rekurrente Zustände aufweisen können, gilt folgendes

**Korollar 7.1**

Für eine irreduzible, aperiodische MK mit endlich vielen Zuständen hat das lineare Gleichungssystem (7-3) eine eindeutige, strikt positive Lösung.

Damit können wir ein erstes wichtiges Resultat formulieren.

**Lemma 7.1**

*Gegeben sei eine endliche Population, bestehend aus den Spielertypen Hawk, Dove und K-Typ, die das modifizierte GD im random-matching-Verfahren spielen. Sei die Entwicklung der Populationsanteile der Typen über die Generationen beschrieben durch eine payoff-monotone Dynamik vom Typ (7-1). Bestehe am Beginn einer jeden matching-Periode (Generation) für jeden Spieler unabhängig voneinander die Wahrscheinlichkeit  $\varepsilon$  zur Mutation (d.h. jeder Spieler wählt irrtümlich mit Wahrscheinlichkeit  $\varepsilon$  jede der beiden übrigen Strategien, die ihm nicht „vererbt“ wurden). Dann gilt: Die so induzierte Markov-Kette besitzt eine ein*



*deutige stationäre Verteilung, die strikt positiv und unabhängig von der Anfangsverteilung ist.*

**Beweis:**

i) Die Markov-Kette ist irreduzibel. Dies folgt aus der Tatsache, daß alle Zustände des gesamten Zustandsraumes gegenseitig erreichbar sind, denn: Jeder Zustand kann von jedem anderen Zustand aus durch eine geeignete Anzahl simultan auftretender Mutanten erreicht werden. Die Wahrscheinlichkeit für derartige Ereignisse ist zwar sehr klein, aber positiv. ii) Die Markov-Kette ist außerdem aperiodisch. Aufgrund der Möglichkeit geeigneter simultaner Mutationen kann jeder Zustand von jedem anderen Zustand aus in **einem Schritt** erreicht werden, d.h. die kleinste Periode der Kette ist  $d=1$ . iii) Da der Zustandsraum endlich ist, sind alle Zustände positiv rekurrent. Die Behauptung des Lemmas folgt dann aus der Anwendung des Korollars 7.1. □

Im allgemeinen gilt dabei, daß die stationäre Verteilung abhängig sein wird von der Höhe der Mutationswahrscheinlichkeit. Formal sei dies durch die Notationsweise  $\pi = \pi(\varepsilon)$  zum Ausdruck gebracht. Ohne weitere Annahmen bezüglich des Parameters  $\varepsilon$  kann die genaue Gestalt der stationären Verteilung nicht ermittelt werden. Man kann nun argumentieren, daß Mutationen (Irrtümer, Fehler) seltene Ereignisse sind, und daß deshalb der evolutionäre Prozeß im wesentlichen dominiert wird von einer Darwinschen Dynamik. Unter diesen Voraussetzungen läßt sich, wie wir sehen werden, ein ausgesprochen starkes Resultat bezüglich der stationären Verteilung ableiten. Wir definieren zunächst den Begriff der Grenzverteilung:

**Definition 7.14** (Grenzverteilung):

Unter der Grenzverteilung der Markov-Kette sei die stationäre Verteilung verstanden, die sich ergibt, wenn die Mutationswahrscheinlichkeit arbiträr klein wird, also

$$\pi^* \equiv \lim_{\varepsilon \rightarrow 0} \pi(\varepsilon)$$

Rein theoretisch könnte man nun das Lineare Gleichungssystem (7-3) lösen, die stationäre Verteilung und anschließend deren Grenzwert bestimmen. Dieses Unterfangen gerät jedoch schon recht bald zu einer unüberschaubaren Rechenarbeit. Bereits für relativ kleine Populationsgrößen wird der Zustandsraum des Prozesses recht groß, und man sähe sich gezwungen, hochdimensionale lineare Gleichungssysteme zu lösen. Als außerordentlich zweckreich erweist sich daher eine alterna

tive Form der Darstellung der stationären Verteilung. Freidlin und Wentzel (1984) haben gezeigt, daß diese auch graphentheoretisch charakterisiert werden kann.

**Definition 7.15** (i-Baum):

Betrachtet sei ein beliebiger Zustand  $i \in E$ . Ein i-Baum  $Y$  ist ein gerichteter Graph auf dem Zustandsraum  $E$  mit den folgenden Eigenschaften:

- i) Jeder Zustand  $i' \in E \setminus \{i\}$  ist der Ausgangspunkt für genau einen Pfeil; (ein Pfeil ist eine Verbindungslinie zwischen zwei Knoten des Graphen);
- ii) Für jeden Zustand  $i' \in E \setminus \{i\}$  existiert ein Pfad, der die Zustände  $i'$  und  $i$  miteinander verbindet, d.h. es existiert eine Abfolge von Pfeilen  $y = \{(i^0, i^1), (i^1, i^2), \dots, (i^{n-1}, i^n)\}$ , so daß  $i^0 = i', i^n = i$ .

**THEOREM** (Freidlin und Wentzel, 1984):

Bezeichne nun  $Y_i$  die Menge aller i-Bäume. Wir definieren ferner den Vektor  $q$  wie folgt:  $q \equiv (q_i)_{i \in E} \in \mathfrak{R}^{|E|}$  mit  $q_i = \sum_{Y \in Y_i} \prod_{(i', i'') \in Y} p_{i' i''}$ . Dann ist die stationäre Verteilung gegeben durch

$$\pi_i = \frac{q_i}{\sum_{i' \in E} q_{i'}} \quad \forall i \in E.$$

**Beweis:** Freidlin und Wentzel (1984), S.177f.

Wir sind nunmehr in der Lage, das Hauptresultat unseres stochastischen Modells angeben zu können. Wir bezeichnen im folgenden die monomorphen Populationszustände als ALLH (100% Hawks), ALLD (100% Doves), bzw. ALLK (100% K-Typen).

**SATZ 7.2** (Grenzverteilung):

Betrachtet sei die stationäre Verteilung aus Lemma (7.1). Bezeichnen  $\pi_1, \pi_2, \pi_3$  die Zustandswahrscheinlichkeiten für die monomorphen Populationszustände ALLH, ALLD und ALLK. Seien  $\tilde{n}_1 \equiv \frac{2(\lambda-1)-c}{(\lambda-1-c)}$ ,  $\tilde{n}_2 \equiv \frac{2\lambda-6+c}{(\lambda-3+c)}$  und

$$\tilde{m} \equiv \frac{(6-\lambda)(n-h) + \lambda h}{h(\lambda-3) + (n-h)(h(2-c) + 1 - \lambda)}.$$

Falls  $\lambda \in ]1+c, 3-c[$ ,  $n > \max\{\tilde{n}_1, \tilde{n}_2\}$

und  $m > \tilde{m}$ , dann gilt:  $\pi_i^* > 0$  für  $i = 1, 2, 3$  und  $\pi_i^* = 0 \quad \forall i > 3$ .

**Beweis:**

Der Beweis zerfällt in mehrere Schritte, die zunächst kurz skizziert sein sollen. Man zeigt zunächst leicht, daß wenn ein Zustand zur Grenzverteilung gehört, er

ebenfalls zu einer Grenzmenge der mutationsfreien Dynamik gehören muß. Im zweiten Beweisschritt bestimmt man sodann eben diese Grenzmenge. Es stellt sich heraus, daß diese gegeben ist durch die drei monomorphen Zustände. Damit werden gewissermaßen die Kandidaten für die Grenzverteilung ermittelt. Im dritten Beweisschritt wird gezeigt, daß für jeden der monomorphen Populationszustände mindestens ein i-Baum existiert mit der minimal möglichen Anzahl von zwei Mutationen. Zusammen mit der graphentheoretischen Charakterisierung der stationären Verteilung kann man dann schlußfolgern, daß tatsächlich alle drei monomorphen Zustände positive Wahrscheinlichkeit haben müssen. Nun zu den Details.

Zunächst zu Schritt eins. Unter der mutationsfreien Dynamik soll der stochastische Prozeß verstanden werden für den Fall, daß  $\varepsilon=0$  (nicht  $\varepsilon \rightarrow 0!$ ). Sei die mutationsfreie Dynamik als  $D_0$  notiert. Wir können dann den Begriff der Grenzmenge genauer bestimmen.

**Definition (7.16):**

Eine Menge von Zuständen  $K \subseteq E$  heiße Grenzmenge von  $D_0$ , falls:

- i)  $\forall i \in K, p_{i,i'} > 0 \Rightarrow i' \in K$  ;
- ii)  $\forall i, i' \in K, \exists m \in \mathbb{N}_+$  so daß  $p_{i,i'}^{(m)} > 0$ .

**Lemma 7.2 :**

Gelte  $i \in \text{supp}(\pi^*)$ . Dann gilt:  $i$  ist auch Element der Grenzmenge der mutationsfreien Dynamik  $D_0$ .

**Beweis:**

Vega-Redondo (1996, Lemma 3, S. 137 f.).

**Lemma 7.3:**

Gelte  $\lambda \in ]1 + c, 3 - c[$ ,  $n > \max\{\tilde{n}_1, \tilde{n}_2\}$ , sowie  $m > \tilde{m}$ . Dann gilt: Die Grenzmenge der mutationsfreien Dynamik ist gegeben durch die drei monomorphen Zustände ALLH, ALLK, ALLD.

**Beweis des Lemmas 7.3:**

Es genügt zu zeigen, daß von jedem inneren Punkt des Zustandsraums aus entweder ALLH oder ALLK oder ALLD in endlicher Zeit, mit positiver Wahrscheinlichkeit erreicht werden kann.

Wir zerlegen den Zustandsraum  $E$  zunächst gedanklich in zwei Bereiche, das Innere des Zustandsraums einerseits, welches im folgenden mit  $intE$  bezeichnet sei sowie die drei „Ränder“  $HK$ ,  $HD$  und  $KD$  andererseits. Unter  $intE$  sind dabei alle Zustände  $i$  zu verstehen, in denen alle drei Verhaltenstypen mit positiven Anzahlen in der Population vertreten sind. Formal:

$intE = \{i \mid h + d + k = n, h, d, k > 0\}$ . Entsprechend sind die Ränder definiert als:

$HK = \{i \mid h + k = n, h, k > 0\}$ ,  $HD = \{i \mid h + d = n, h, d > 0\}$  sowie

$KD = \{i \mid k + d = n, k, d > 0\}$ .

Sodann zerlege man  $intE$  vollständig in sechs Sektoren, nämlich:

Sektor I:  $h \leq k < d$

Sektor II:  $h \leq d \leq k$

Sektor III:  $k < h < d$

Sektor IV:  $k \leq d \leq h$

Sektor V:  $d < h < k$

Sektor VI:  $d < k < h$

- i) Wir betrachten zunächst  $KD$ . Unabhängig vom Resultat des matching-Prozesses gilt offensichtlich  $\bar{\pi}_D = 3 > 3 - c = \bar{\pi}_K$ . Doves realisiert immer den höheren durchschnittlichen Payoff, da die kooperative Strategie gegenüber  $K$  keine Kosten  $c$  zu tragen hat. Doves werden also nach maximal  $k-1$  Generationen sämtliche intelligenten Kooperateure verdrängt haben, so daß der monomorphe Zustand ALLD erreicht ist.
- ii) Man betrachte nun  $HD$ . Man kann  $HD$  weiter zergliedern in zwei Klassen von Zuständen, nämlich solche in denen  $h \leq d$  und jene, in denen  $h > d$  gilt. Für den Fall  $h \leq d$  nehme man nun an, daß alle  $h$  Hawks mit Doves gematcht werden, es also zu keinen cross-matchings zwischen Hawks kommt. Dann ist klar, daß  $\bar{\pi}_H = 6 > \bar{\pi}_D$ . Wiederhole sich nun dieses matching-Resultat hinreichend oft. Nach maximal  $d-h+1$  Generationen ist dann die zweite Klasse von  $HD$ -Zuständen erreicht, in denen  $h > d$  gilt. Wir nehmen nunmehr an, daß alle Doves mit Hawks gematcht werden, es also zu keinerlei cross-matchings unter Doves kommt. Klarerweise gilt  $\bar{\pi}_D = 0 < \bar{\pi}_H$ , so daß Hawks sich also weiterhin ausbreiten. Es sei wiederum angenommen, daß sich das vorstehende matching-Ergebnis hinreichend oft wiederholt. Nach maximal weiteren  $n-d$  Generationen wird dann der Zustand ALLH erreicht. (Bemerkung: Die Wahrscheinlichkeiten für die zugrunde gelegten matching-Resultate sind zwar möglicherweise sehr klein, in einer endlichen Population aber positiv. Es ist damit gezeigt, daß der Zustand ALLH mit positiver Wahrscheinlichkeit in endlicher Zeit erreicht wird).

- iii) Der dritte Beweisschritt besteht darin zu zeigen, daß von jedem Zustand  $i \in \text{int } E$  aus entweder der Rand HK oder der monomorphe Zustand ALLH erreicht werden kann. Betrachtet seien zunächst alle Sektoren, in denen  $d \leq h$  gilt (dies sind die Sektoren IV, V und VI). Es sei angenommen, daß alle  $d$  Doves in allen  $m$  matchings mit Hawks gematcht werden, also  $P_{HD} = md$ . Dieses matching-Resultat wiederhole sich hintereinander  $d$ -mal. Unter der getroffenen Annahme ist Dove eindeutig die Strategie mit der niedrigsten durchschnittlichen Auszahlung, so daß nach maximal  $d$  Generationen Doves ausgestorben sind. Wir betrachten nun die Sektoren gesondert. **Ad Sektor IV:** A priori ist unter der Annahme der payoff-monotonen Dynamik nicht auszuschließen, daß auch die Strategie K ausstirbt. Von Sektor IV aus kann also sowohl HK als auch ALLH erreicht werden. **Ad Sektor V:** Selbst im schlechtesten Fall für die Strategie K, nämlich  $\bar{\pi}_K < \bar{\pi}_H$  und  $\Delta k < 0$ , so gilt unter der Annahme der Payoffmonotonie doch  $|\Delta k| < |\Delta d|$ , so daß wegen anfangs  $k > d$  auch dann noch K-Typen existieren müssen, wenn  $d = 0$  erreicht wird. Vom Sektor V aus wird also mit positiver Wahrscheinlichkeit HK erreicht. **Ad Sektor VI:** analog zu Sektor V.

Als nächstes seien die übrigen Sektoren mit  $h < d$  betrachtet. Wir treffen die Annahme, daß alle  $h$  Hawks mit Doves gematcht werden. Damit gilt  $\bar{\pi}_H = 6$ , Hawk ist also die eindeutig erfolgreichste Strategie.

**Ad Sektor I:** Klar ist:  $\Delta h > 0$  sowie  $\Delta h > |\Delta k|, |\Delta d|$ . Klar ist ferner: Bei hinreichender Wiederholung des angenommenen matching-Resultates muß „irgendwann“ gelten  $h > d$ . **Rein kombinatorisch** ergeben sich drei mögliche Konstellationen unter Berücksichtigung von  $k$  und  $d$ , die mit der Reihung  $h > d$  konsistent sind, nämlich: a)  $h > d \geq k$ , (b)  $h > k > d$  und (c)  $k > h > d$ . Nun gehören alle Zustände aus (a) offensichtlich zum Sektor IV und (b) bzw. (c) konstituieren gerade die Sektoren VI und V. Es gilt also, daß unter der getroffenen matching-Annahme sowie unter einer payoff-monotonen Dynamik notwendig entweder ein Populationszustand aus Sektor IV, V oder VI erreicht wird.

Zu beachten ist jedoch: A priori ist auch der Fall nicht ausgeschlossen, daß K oder Dove ausstirbt. Im ersten Fall wird ein Zustand aus HD erreicht, von dem aus mit positiver Wahrscheinlichkeit ALLH erreicht wird (vgl. (ii)). Im zweiten Fall wird ein Zustand aus HK erreicht.

**Ad Sektor II:** Die Argumentation ist analog. Solange  $h \leq d$  nimmt die Anzahl der Hawks unter der getroffenen matching-Annahme unaufhörlich zu, bis schließlich ein Zustand mit  $h > d$  erreicht ist. Wiederum ergeben sich drei denkmögliche Konstellationen, nämlich (a)  $d < h < k$ , (b)  $k \leq d < h$ ,

oder (c)  $d < k < h$ . Unter den getroffenen Annahmen gilt also, daß die mutationsfreie Dynamik vom Sektor II aus notwendig gegen einen Zustand aus Sektor IV, V oder VI konvergiert. Schließlich besteht wieder die Möglichkeit, daß K oder Dove zuvor aussterben, also Zustände aus HD oder HK erreicht werden.

**Ad Sektor III:** Vom Sektor III aus können Zustände der Sektoren IV oder VI erreicht werden, allerdings keine Zustände des Sektors V. Wie bereits in den Fällen der Sektoren I und II nimmt  $h$  unaufhörlich zu, bis schließlich ein Zustand mit  $h > d$  erreicht sein muß. Eine Konstellation wie in Sektor V mit  $d < h < k$  ist jedoch nicht möglich, da die Startbedingungen in Sektor III charakterisiert sind durch  $k < h < d$ . Selbst wenn  $\Delta k > 0$ , so gilt doch auf jeden Fall:  $\Delta h > \Delta k$ , so daß  $k$  die Anzahl der Hawks niemals übersteigen kann. Vereinbar mit  $h > d$  sind aber weiterhin die Reihungen der Sektoren IV und VI.

Kurz zusammengefaßt wurde damit bisher folgendes gezeigt: Von den Sektoren I, II und III aus wird unter den getroffenen Annahmen ( $P_{HD} = mh$ , Payoffmonotonie) notwendig ein Zustand aus den Sektoren IV, V oder VI oder aus HD oder HK erreicht. Von den Sektoren IV, V und VI aus wird unter der matching Annahme  $P_{HD} = md$  notwendig entweder ein Zustand aus HK erreicht (Dove stirbt aus) oder aber der monomorphe Zustand ALLH. Für die Zustände der Menge HD existieren matching-Resultate mit positiver Wahrscheinlichkeit, unter denen notwendig ALLH erreicht wird. Für die Zustände der Menge KD gilt unabhängig vom matching-Resultat, daß ALLD erreicht wird. Soweit sind als Grenzmenge der mutationsfreien Dynamik also die monomorphen Zustände ALLH und ALLD identifiziert. Es verbleibt die Untersuchung der Zustände der Menge HK und der Nachweis, daß unter geeigneten Bedingungen ALLK erreicht werden kann.

- iv) Wir werden zeigen, daß unter geeigneten Bedingungen für jedes Element der Menge HK gilt, daß K gegenüber H die erfolgreichere Strategie ist, d.h. daß die durchschnittliche Auszahlung von K größer ist als die der Hawks.

O.B.d.A. sei angenommen, daß  $n$  gerade ist. Sei dann zunächst der Fall  $h$  gerade betrachtet. Klarerweise gilt dann auch  $k = n - h$  gerade. Angenommen, es komme zu folgendem matching-Resultat: In allen  $m$  matching-Zeitpunkten werden alle  $h$  Hawks untereinander gematcht und alle K-Typen ebenfalls untereinander. Die durchschnittlichen Auszahlungen beider Strategien sind offensichtlich  $\bar{\pi}_H = 1 < 3 - c = \bar{\pi}_K$ . K ist die eindeutig

erfolgreichste Strategie und unter payoffmonotonen Dynamiken muß die Zahl der K-Typen steigen.

Sei nun der schwierigere Fall  $h$  ungerade betrachtet. In diesem Fall muß es notwendigerweise zu einem cross-matching zwischen mindestens einem Hawk und einem K-Typen kommen. Die kumulierten Payoffs der Hawks sind gegeben als  $\pi_H = 6P_{HD} + 2P_{HH} + \lambda P_{H_KK} + 6P_{H_UK}$ . Natürlich gilt wegen  $d = 0$  auch  $P_{HD} = 0$ . Sei nun angenommen, daß  $h-1$  Hawks in allen  $m$  matching-Zeitpunkten untereinander gematcht werden, d.h. gelte  $2P_{HH} = m(h-1)$ . Bezüglich des cross-matchings des verbleibenden, letzten Hawks mit einem K-Typ sei nun die günstige Annahme getroffen, daß es sich bei dem Hawk immer um dasselbe Individuum handelt. D.h., der Hawk ist im ersten matching ein unbekannter Hawk und wird danach von demselben K-Typen immer wiedergetroffen. Ein solches Ereignis mag extrem unwahrscheinlich sein, dennoch gilt, daß für nicht-degenerierte Träger der Zufallsvariablen (wie es in einer endlichen Population mit endlicher Interaktionszahl der Fall ist) ein solches Ereignis positive Wahrscheinlichkeit hat. Gelte also  $P_{H_UK} = 1, P_{H_KK} = m-1$ . Der kumulierte Payoff der Hawks ergibt sich damit als

$$\pi_H = m(h-1) + \lambda(m-1) + 6, \quad (7-4)$$

bzw. der durchschnittliche Payoff resultiert als

$$\bar{\pi}_H = \frac{m(h-1) + \lambda(m-1) + 6}{mh}. \quad (7-5)$$

Der kumulierte Payoff der K-Typen ist allgemein gegeben als

$$\pi_K = (3-c)P_{DK} + (\lambda-c)P_{H_KK} + (3-c)2P_{KK} - cP_{H_UK}. \quad (7-6)$$

Wegen  $d=0$  gilt  $P_{DK} = 0$ . Ferner gilt unter der getroffenen matching-Annahme  $2P_{KK} = m(n-h-1)$  und natürlich weiterhin  $P_{H_KK} = m-1$ . Also erhält man

$$\pi_K = \lambda(m-1) + 3m(n-h-1) - cm(n-h), \quad (7-7)$$

bzw. als durchschnittliche Auszahlung

$$\bar{\pi}_K = \frac{\lambda(m-1) + 3m(n-h-1) - cm(n-h)}{m(n-h)}. \quad (7-8)$$

Damit K sich gegenüber H durchsetzt ist zu fordern:  $\bar{\pi}_K > \bar{\pi}_H$ . Ein Vergleich beider Payoffs liefert nach einigen algebraischen Umformungen

$$m[h(\lambda - 3) + (n - h)(h(2 - c) + 1 - \lambda)] > (6 - \lambda)(n - h) + \lambda h. \quad (7-9)$$

Zum ersten stellt man fest, daß die rechte Seite eindeutig positiv ist. Zum zweiten bemerkt man, daß daher bezüglich des Klammerausdrucks auf der linken Seite der Ungleichung zwingend zu fordern ist, daß dieser ebenfalls positiv sei.

$$\begin{aligned} h(\lambda - 3) + (n - h)(h(2 - c) + 1 - \lambda) > 0 &\Leftrightarrow \\ h^2(2 - c) - h(2\lambda - 4 + n(2 - c)) - n(1 - \lambda) < 0 &\Leftrightarrow \\ h_2 < h < h_1, \text{ wobei} & \\ h_{1,2} = \frac{2\lambda - 4 + n(2 - c)}{2(2 - c)} \pm \sqrt{\frac{(2\lambda - 4 + n(2 - c))^2}{4(2 - c)^2} + \frac{n(1 - \lambda)}{2 - c}} & \quad (7-10) \end{aligned}$$

Nun ist natürlich zu fordern, daß  $h_2 < h < h_1$  keine wirkliche Restriktion darstellen darf. Insbesondere müssen auch die (wichtigen und interessanten) Fälle  $h=n-1$  sowie  $h=1$  erlaubt sein. Es ist daher zu fordern: (i)  $h_1 > n-1$ , sowie (ii)  $h_2 < 1$ .

Ad (i): Nach einigen Umformungen erhält man schließlich

$$\begin{aligned} h_1 > n - 1 &\Leftrightarrow \\ n(\lambda - 1 - c) > 2(\lambda - 1) - c. & \quad (7-11) \end{aligned}$$

Eine Fallunterscheidung ist erforderlich:

a)  $(\lambda - 1 - c) > 0$  Dann:

$$\begin{aligned} h_1 > n - 1 &\Leftrightarrow \\ n > \frac{2(\lambda - 1) - c}{(\lambda - 1 - c)} \equiv \tilde{n}_1 & \quad (7-12) \end{aligned}$$

b)  $(\lambda - 1 - c) < 0$

Sinnvoll ist dann offensichtlich nur:  $2(\lambda - 1) - c < 0 \Leftrightarrow \lambda < 1 + \frac{c}{2}$ . Daher:

$$h_1 > n - 1 \Leftrightarrow$$



$$n < \tilde{n}_1 \quad (7-13)$$

Man stellt nun leicht fest, daß diese Restriktion nicht erfüllbar ist, da:

$$\tilde{n}_1 = \frac{2(\lambda-1)-c}{(\lambda-1-c)} = 1 + \frac{\lambda-1}{\lambda-1-c} < 1 \quad (7-14)$$

Relevant ist also nur der Fall (a).

Ad (ii): Man erhält nach einigen Umformungen endlich:

$$\begin{aligned} h_2 < 1 &\Leftrightarrow \\ n(\lambda-3+c) < 2\lambda-6+c &\quad (7-15) \end{aligned}$$

a)  $(\lambda-3+c) < 0$

Dann:

$$\begin{aligned} h_2 < 1 &\Leftrightarrow \\ n > \frac{2\lambda-6+c}{(\lambda-3+c)} &\equiv \tilde{n}_2 \quad (7-16) \end{aligned}$$

b)  $(\lambda-3+c) > 0$

$$\begin{aligned} h_2 < 1 &\Leftrightarrow \\ n < \tilde{n}_2 &\quad (7-17) \end{aligned}$$

Man zeigt leicht, daß diese Restriktion niemals erfüllt werden kann, denn:

$$\tilde{n}_2 = \frac{2\lambda-6+c}{(\lambda-3+c)} < 1 \Leftrightarrow \lambda < 3 \quad (7-18)$$

Wir halten fest: Es war zu fordern:  $h_1 > n-1$ . Diese Bedingung ist erfüllt, falls  $\lambda > 1+c$  und  $n > \frac{2(\lambda-1)-c}{(\lambda-1-c)} \equiv \tilde{n}_1$ . Ferner war zu fordern:  $h_2 < 1$ .

Diese Restriktion ist dann erfüllt, wenn  $\lambda < 3-c$  sowie  $n > \frac{2\lambda-6+c}{(\lambda-3+c)} \equiv \tilde{n}_2$ . Gegeben diese Bedingungen bezüglich  $\lambda, n$  sind erfüllt, so erhält man bezüglich der erforderlichen Anzahl der matchings

$$m[h(\lambda-3) + (n-h)(h(2-c)+1-\lambda)] > (6-\lambda)(n-h) + \lambda h \Leftrightarrow$$

$$m > \frac{(6-\lambda)(n-h)+\lambda h}{h(\lambda-3)+(n-h)(h(2-c)+1-\lambda)}. \quad (7-19)$$

Unter den genannten Bedingungen ist also K erfolgreicher als H und wird sich unter payoffmonotonen Dynamiken daher durchsetzen, bis schließlich ALLK erreicht ist. Damit ist der Beweis bezüglich der Grenzmenge der mutationsfreien Dynamik vollständig.

Aus Lemma 7.2 folgte, daß überhaupt nur solche Zustände mit positiver Häufigkeit zur Grenzverteilung gehören können, die auch Element der Grenzmenge der mutationsfreien Dynamik sind. Damit ist klar, daß alle Zustände  $i \in E \setminus \{ALLH, ALLK, ALLD\}$  nicht mit positiver Häufigkeit zur Grenzverteilung gehören.

Wir wenden uns nun dem letzten Beweisschritt zu, nämlich dem Nachweis, daß tatsächlich **alle** drei monomorphen Populationszustände positive Häufigkeit haben. Hierzu ist es erforderlich, sich die graphentheoretische Charakterisierung der stationären Verteilung nach Freidlin und Wentzel in Erinnerung zu rufen. Diese war gegeben als

$$\pi_i = \frac{q_i}{\sum_{i' \in E} q_{i'}} \quad \forall i \in E. \quad (7-20)$$

Dabei war  $q_i = \sum_{Y \in \mathcal{Y}_i} \prod_{(i', i'') \in Y} p_{i' i''}$ . Sowohl im Zähler, als auch im Nenner der stationären

Verteilung stehen also Summen über alle zu einem Zustand gehörigen i-Bäume. Zunächst gilt es sich klar zu machen, daß es keinen mutationsfreien i-Baum für die monomorphen Populationszustände geben kann. Dies wird sofort klar, wenn man sich Teil (ii) der Definition eines i-Baumes ins Gedächtnis ruft:

*Für jeden Zustand  $i' \in E \setminus \{i\}$  existiert ein Pfad, der die Zustände  $i'$  und  $i$  miteinander verbindet, d.h. es existiert eine Abfolge von Pfeilen  $y = \{(i^0, i^1), (i^1, i^2), \dots, (i^{n-1}, i^n)\}$ , so daß  $i^0 = i', i^n = i$ .*

Es muß also für jeden anderen Zustand außer dem in Betrachtung stehenden ein Pfad existieren, der zu diesem führt. Um es an einem Beispiel klar zu machen: Betrachten wir den monomorphen Zustand ALLH. Ein i-Baum für ALLH müßte dann unter anderem auch spezifizieren, auf welchen Pfaden ALLH von den übrigen beiden monomorphen Zuständen aus erreicht werden kann. Nun ist aber of

fensichtlich, daß ein solcher Pfad immer mindestens eine Mutation einschließen muß, da die monomorphen Zustände unter der mutationsfreien Dynamik absorbierend sind, d.h. sie werden ohne Mutation niemals verlassen.

Wir werden für jeden der drei monomorphen Populationszustände ALLK, ALLH und ALLD zeigen, daß die minimal erforderliche Anzahl von Mutationen in einem  $i$ -Baum jeweils zwei beträgt, so daß die resultierenden Produkte der Übergangswahrscheinlichkeiten innerhalb des Graphen von der Ordnung  $\varepsilon^2$  sein müssen.

Wir beginnen mit dem Zustand ALLK. Ein Graph mit minimal erforderlicher Anzahl von Mutationen muß definitionsgemäß für jeden anderen Zustand  $i \in E \setminus ALLK$  spezifizieren, auf welche Weise ALLK mit einem Minimum an Mutationen erreicht werden kann. a) Betrachten wir zunächst alle Zustände  $i \in \text{int } E$ . Für alle Zustände aus den Sektoren I, II und III war gezeigt worden, daß unter der mutationsfreien Dynamik mit positiver Wahrscheinlichkeit ein Zustand aus den Sektoren IV, V oder VI erreicht werden muß. Von IV, V und VI aus wird wiederum mit positiver Wahrscheinlichkeit entweder ein Zustand  $i \in \{HK\} \setminus ALLK$  oder aber direkt ALLH erreicht. Im ersten Fall ist offensichtlich gezeigt, daß von jedem Zustand aus dem Inneren des Ereignisraumes ALLK ohne jegliche Mutation erreicht werden kann. Im zweiten Fall gilt, daß für den verbleibenden Übergang von ALLH nach ALLK gerade eine Mutation erforderlich ist (und es müssen die Bedingungen des Satzes 7.2 erfüllt sein, damit der Übergang von ALLH nach ALLK mittels einer Mutante überhaupt möglich ist). b) Wenden wir uns nunmehr den Zuständen  $i \in \{HK\} \setminus ALLH$  zu. Sind die Bedingungen des Satzes 7.2 erfüllt, so erfolgt ein Übergang von jedem  $i \in \{HK\} \setminus ALLH$  mit positiver Wahrscheinlichkeit unter der mutationsfreien Dynamik allein. c) Betrachtet sei schließlich der monomorphe Zustand ALLH. Klar ist: Dieser kann mittels einer K-Mutation verlassen werden. Sodann befindet sich der Prozeß in einem  $i \in \{HK\} \setminus ALLH$ . Sind die Bedingungen des Satzes 7.2 erfüllt, so gilt das Argument aus b). ALLK wird von ALLH aus also mittels einer Mutation erreicht. d)  $i \in \{HD\} \setminus ALLD$ : Die mutationsfreie Dynamik konvergiert mit positiver Wahrscheinlichkeit gegen ALLH. Von dort aus wird ALLK mittels einer Mutation erreicht. e) ALLD: Mutiere genau ein Spieler zu Hawk. Danach befindet sich der Prozeß in einem  $i \in \{HD\} \setminus ALLD$  und konvergiert ohne weitere Mutationen gegen ALLH. Von ALLH aus wird mittels einer weiteren Mutation ALLK erreicht. ALLK wird von ALLD aus also mittels zweier Mutationen erreicht. Beachte: Die Alternative – der Weg über  $\text{int } E$  – würde ebenfalls mindestens zwei Mutationen erfordern. f)  $i \in \{KD\} \setminus ALLD$ : Hier können wir wie folgt argumentieren: Zunächst konvergiert die mutationsfreie Dynamik notwendig gegen ALLD. Sodann wird, wie oben für ALLD beschrieben,

ALLK mittels zweier Mutationen erreicht. Die Alternative wäre eine Mutation in das Innere des Ereignisraumes hinein. Dies impliziert mindestens eine H-Mutation. Von  $int E$  aus wird dann wieder unter der mutationsfreien Dynamik entweder ein  $i \in \{HK\}$  oder ALLH erreicht. Von ALLH aus ist zum Erreichen von ALLK eine weitere Mutation erforderlich. Da a priori keineswegs ausgeschlossen ist, daß tatsächlich ALLH erreicht wird, so wären auch auf diesem Weg mindestens zwei Mutationen zum Erreichen von ALLK erforderlich.

Die minimale Anzahl von Mutationen zum Erreichen von ALLK in einem  $i$ -Baum beträgt also zwei. Das Produkt der Übergangswahrscheinlichkeiten im Zähler der stationären Verteilung,  $\prod_{(i',i'') \in Y} p_{i'i''}$ , enthält die Mutationswahrscheinlichkeit  $\varepsilon$  damit

mindestens in der zweiten Potenz.

Wenden wir uns nun ALLD zu. Die obige Diskussion sollte auch folgendes deutlich gemacht haben: Für die Gesamtzahl der minimal erforderlichen Mutationen ist es gleichgültig, auf welchen der Ränder die mutationsfreie Dynamik von  $int E$  aus konvergiert. So wäre es etwa denkbar, für die Sektoren I,II und III matching-Resultate mit positiver Wahrscheinlichkeit zu konstruieren, für die eine Konvergenz gegen den Rand  $\{KD\}$  gezeigt werden könnte. ALLD würde dann von den genannten Sektoren aus ohne jegliche Mutation erreicht. Genausogut können wir aber annehmen, daß eine Konvergenz gegen  $\{HK\}$  erfolgt, so daß ALLD via einer Mutation von ALLK aus erreicht werden könnte. Für die insgesamt im  $i$ -Baum der minimal notwendigen Mutationen auftretende Zahl von Mutanten ist diese Unterscheidung jedoch ohne Belang, denn der Übergang von ALLK nach ALLD erfordert ohnehin mindestens eine Mutation. Die Zahl der minimal erforderlichen Mutationen bezüglich eines Baumes wird nämlich letztlich determiniert durch die Anzahl der Mutationen die notwendig sind, um die Übergänge von den anderen monomorphen Zuständen aus zu bewerkstelligen. a) Wie bereits gesagt: Der Übergang von ALLK nach ALLD erfordert eine Mutation. b) Der Übergang von ALLH nach ALLD erfordert mindestens zwei Mutationen, denn: Eine einzelne D-Mutante kann eine H-Population niemals invadieren. Erforderlich sind immer mindestens zwei Mutanten. Ferner: ALLD kann also erreicht werden durch Übergang von ALLH nach ALLK und schließlich von ALLK nach ALLD. Dies impliziert ebenfalls mindestens zwei Mutationen. Denkmöglich wäre auch noch eine Mutation ins Innere  $int E$ . Aber auch dieser Weg erfordert mindestens zwei Mutationen. Der  $i$ -Baum der minimal erforderlichen Mutationen für ALLD involviert also ebenfalls zwei Mutationen, so daß wiederum gilt: Das Produkt der Übergangswahrscheinlichkeiten im Zähler der stationären Verteilung,  $\prod_{(i',i'') \in Y} p_{i'i''}$ , enthält die Mutationswahrscheinlichkeit  $\varepsilon$  mindestens in der zweiten Potenz.

Schließlich ALLH: a) Klar ist: Eine singuläre H-Mutante invadiert eine D-Population immer erfolgreich. ALLH wird von ALLD aus also mittels einer Mutation erreicht. b) Für ALLK ist es wiederum gleichgültig, auf welchem Wege ALLH erreicht wird. Zunächst wäre denkbar, daß eine H-Mutante ALLK erfolgreich invadiert (man kann zeigen, daß dies unter gewissen Bedingungen, ähnlich zu denen für eine erfolgreiche Invadierung einer ALLH-Population durch eine K-Mutante, tatsächlich der Fall sein kann). Die minimal erforderliche Zahl der Mutanten im gesamten Baum betrüge dann zwei. Ebenso kann man sich aber vorstellen, daß ALLK durch eine D-Mutante invadiert wird. Mittels einer Mutation gelangt man so nach ALLD und von dort mittels einer weiteren Mutante nach ALLH. Auch dieser Weg impliziert zwei Mutationen, hat aber den Vorteil, daß er nicht an bestimmte Bedingungen geknüpft ist. Denkmöglich wäre schließlich auch, daß ALLH von ALLK über *intE* erreicht wird. Dies implizierte aber ebenfalls zwei Mutationen. Die minimal erforderliche Gesamtzahl an Mutationen, um ALLH zu erreichen, beträgt also wiederum zwei.

Das entscheidende Argument folgt nun: Man betrachte

$$\lim_{\varepsilon \rightarrow 0} \pi_i = \frac{q_i}{\sum_{i' \in E} q_{i'}} \quad \forall i \in E . \quad (7-21)$$

Dann ist klar, daß nur solche Terme in der Grenzverteilung übrig bleiben, die mit der minimal erforderlichen Anzahl von Mutationen auskommen. Alle i-Bäume mit drei oder mehr Mutationen, in denen also als ein Faktor des Produktes der Übergangswahrscheinlichkeiten notwendigerweise  $\varepsilon$  in der dritten oder aber noch höheren Potenz vorkommt, konvergieren – verglichen mit den 2-Mutationsbäumen – schneller gegen Null und verschwinden schließlich. Es bleiben lediglich diejenigen Terme übrig, die **am langsamsten** gegen Null konvergieren, und dies sind gerade die i-Bäume, in denen nur Pfade involviert sind, die lediglich zwei Mutationen benötigen.

Um das Argument etwas präziser zu machen, zergliedern wir die Menge aller i-Bäume in Teilmengen, und zwar dergestalt, daß alle Bäume mit gleicher Mutationszahl zusammengefaßt werden. Bezeichne also  $Y_i^2$  die Menge aller i-Bäume, in denen nur zwei Mutationen auftauchen, entsprechend  $Y_i^3, \dots, Y_i^n$  die Mengen der i-Bäume mit drei, vier, ..., n simultanen Mutationen. Wir können dann den Zähler der stationären Verteilung formulieren als

$$q_i = \sum_{Y \in Y_i^1(i', i'') \in Y} p_{i' i''} = \sum_{Y \in Y_i^2(i', i'') \in Y} p_{i' i''} + \sum_{Y \in Y_i^3(i', i'') \in Y} p_{i' i''} + \dots + \sum_{Y \in Y_i^n(i', i'') \in Y} p_{i' i''} . \quad (7-22)$$

Entsprechend kann der Nenner der Verteilung reformuliert werden als

$$\sum_{i' \in E} q_{i'} = \sum_{i' \in E} \left[ \sum_{Y \in Y_i^2(i', i'') \in Y} p_{i' i''} + \sum_{Y \in Y_i^3(i', i'') \in Y} p_{i' i''} + \dots + \sum_{Y \in Y_i^n(i', i'') \in Y} p_{i' i''} \right]. \quad (7-23)$$

Klar ist: Man kann aus jedem  $Y_i^2$ -Baum einen Faktor  $\varepsilon^2$  ausklammern, aus jedem  $Y_i^3$ -Baum einen Faktor  $\varepsilon^3$ , usw. Bezeichne nun  $\tilde{p}_{i' i''} = \frac{1}{\varepsilon^j} p_{i' i''}$ , wobei  $j = 2, 3, \dots, n$  die Anzahl der zum jeweiligen i-Baum gehörigen Mutationen angibt. Wir können dann schreiben:

$$\begin{aligned} \pi_i &= \frac{\varepsilon^2 \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon^3 \sum_{Y \in Y_i^3(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^n \sum_{Y \in Y_i^n(i', i'') \in Y} \tilde{p}_{i' i''}}{\sum_{i' \in E} \left[ \varepsilon^2 \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon^3 \sum_{Y \in Y_i^3(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^n \sum_{Y \in Y_i^n(i', i'') \in Y} \tilde{p}_{i' i''} \right]} \Leftrightarrow \\ \pi_i &= \frac{\varepsilon^2 \left[ \sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^{n-2} \sum_{Y \in Y_i^{n-2}(i', i'') \in Y} \tilde{p}_{i' i''} \right]}{\varepsilon^2 \sum_{i' \in E} \left[ \sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^{n-2} \sum_{Y \in Y_i^{n-2}(i', i'') \in Y} \tilde{p}_{i' i''} \right]} \Leftrightarrow \\ \pi_i &= \frac{\sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^{n-2} \sum_{Y \in Y_i^{n-2}(i', i'') \in Y} \tilde{p}_{i' i''}}{\sum_{i' \in E} \left[ \sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''} + \varepsilon \sum_{Y \in Y_i^2(i', i'') \in Y} \tilde{p}_{i' i''} + \dots + \varepsilon^{n-2} \sum_{Y \in Y_i^{n-2}(i', i'') \in Y} \tilde{p}_{i' i''} \right]} . \quad (7-24) \end{aligned}$$

Betrachte dann den Grenzübergang für  $\varepsilon \rightarrow 0$ . Offensichtlich gilt:

$$\lim_{\varepsilon \rightarrow 0} \pi_i = \frac{\sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''}}{\sum_{i' \in E} \left[ \sum_{Y \in Y_i^1(i', i'') \in Y} \tilde{p}_{i' i''} \right]} \quad (7-25)$$

Für die drei monomorphen Zustände war jedoch gezeigt worden, daß mindestens ein  $Y_i^2$ -Baum existiert. Also kann in der beschriebenen Weise ausgeklammert werden und die Häufigkeiten für ALLH, ALLK und ALLD in der Grenzverteilung sind strikt positiv.

Dies vervollständigt den Beweis des Satzes 7.2. □

### 7.3 Diskussion der Resultate

Kooperative Verhaltensweisen überleben also auch in einer stochastischen Welt, allerdings ist ihr Überleben, wie im deterministischen Modellkontext auch, an Bedingungen geknüpft. Die zentrale Bedingung betrifft die erforderliche Anzahl von matchings. Letztere muß nämlich hinreichend groß sein, damit die bedingt kooperative Strategie K sich gegenüber der defektiven Strategie behaupten kann. Abbildung 13 zeigt einen typischen Verlauf des kritischen Wertes für  $m$ .

$$\lambda := 1.75 \quad n := 100 \quad c := 0.5 \quad h := 1, 2, \dots, 99 \quad f(h) := \frac{(n-h) \cdot (6-\lambda) + \lambda \cdot h}{(h \cdot (\lambda - 3) + (n-h) \cdot (h \cdot (2-c) + 1 - \lambda))}$$

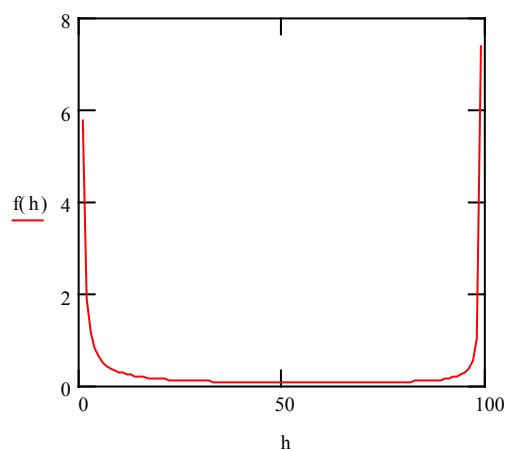


Abb. 13: Der kritische Wert  $\tilde{m}$  in Abhängigkeit von  $h$

Die Populationsgröße wurde mit  $n=100$  festgesetzt,  $h$  durchläuft den gesamten möglichen Wertebereich von eins bis 99. Es sollte nicht unerwähnt bleiben, daß der obige Funktionsverlauf ausgesprochen robust gegenüber Variation der Parameter  $c$  und  $\lambda$  ist, sofern die Bedingung bezüglich dieser Parameter –  $\lambda \in ]1+c, 3-c[$  – eingehalten wird.

Wie man sieht, ist die Bedingung bezüglich  $m$  im allgemeinen nicht sonderlich restriktiv. Für den überwiegenden Teil der  $h$ -Werte ( $1 < h < 99$ ) genügt es bereits, daß es zu mehr als einem matching kommt, damit K die erfolgreichere Strategie sein kann. Eine echte Restriktion stellt (7-19) lediglich für die Randfälle  $h=1$  und  $h=n-1$  dar. Im obigen Beispiel sieht man, daß  $m$  etwa die Werte 6 bzw. 7 übersteigen muß, damit  $\bar{\pi}_K > \bar{\pi}_H$  gilt. Dieses Resultat ist ausgesprochen plausibel. Im Falle nur eines Hawks, den es zu überwinden gilt, ist nämlich der Durchschnittspayoff des Hawks ausgesprochen hoch. In Ermangelung weiterer Defektoren kann es zu keinerlei cross-matchings innerhalb der Gruppe der defektiven Spieler kommen. Der einzige Falke streicht in  $m-1$  matchings den Wert der outside-option  $\lambda$  ein und realisiert in einem matching den Maximalpayoff von 6. Exi-

stieren mehrere Hawks, dann existiert eine positive Wahrscheinlichkeit dafür, daß es zu einer hinreichend großen Anzahl von Interaktionen zwischen Defektoren kommt, was die durchschnittliche Auszahlung erheblich reduziert.

Sehen wir uns Bedingung (7-19) für den Fall  $h=1$  einmal genauer an. Man erhält

$$\tilde{m}(h=1) = \frac{(n-1)(6-\lambda) + \lambda}{(n-1)(3-c) - 3 - \lambda(n-2)}$$

Klar ist:

$$\lim_{\lambda \rightarrow \frac{(n-1)(3-c)-3}{n-2}} \tilde{m}(h=1) = +\infty \tag{7-26}$$

Klar ist auch: Für „hinreichend große“  $n$  gilt

$$\frac{(n-1)(3-c) - 3}{n-2} \cong 3 - c \tag{7-27}$$

Falls also  $\lambda$  gegen die obere Grenze  $3-c$  strebt, wird (7-19) unerfüllbar, wie die folgende Graphik beispielhaft darlegt.

$n := 100$        $c := 0.5$        $\lambda := 1.51, 1.511.. 2.49$        $f(\lambda) := \frac{(n-1) \cdot (6-\lambda) + \lambda}{(\lambda - 3 + (n-1) \cdot (3-c-\lambda))}$

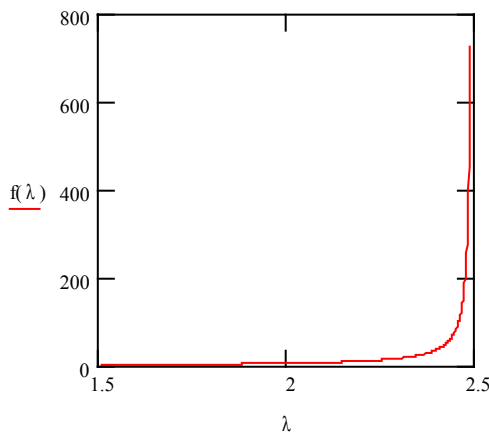


Abb. 14: Der kritische Wert  $\tilde{m}$  für  $h=1$  in Abhängigkeit von  $\lambda$

Dargestellt ist der Verlauf des kritischen Wertes für  $m$  für den Fall  $h=1$  in Abhängigkeit von  $\lambda$ . Der präzise Wert der Unendlichkeitsstelle ist gegeben durch  $\frac{(n-1)(3-c) - 3}{n-2} \cong 2.495$ , während der näherungsweise Wert natürlich gerade

durch  $3 - c = 2.5$  gegeben ist. Wie man sieht fällt die näherungsweise Koinzidenz der Werte bereits für relativ kleine  $n$  gut aus. Sofern also nur gilt  $\lambda < 3 - c$ , wie in Satz 7.2 gefordert, kann man davon ausgehen, daß auch für den Fall  $h=1$  Bedingung (7-19) nicht sonderlich restriktiv ausfällt. Im allgemeinen genügt eine hinreichend große, aber klar endliche Anzahl von matchings, um auch den letzten Hawk noch zum Verschwinden zu bringen.



Wenden wir uns nun dem zweiten Extremfall  $h=n-1$  zu. Auch hier wird (7-19) zu einer echten Restriktion. Die Intuition ist wiederum einfach. Zwar gibt es nun ausreichend Hawks (nämlich  $n-2$ ), die untereinander gematcht werden können, so daß entsprechende matching-Resultate die durchschnittliche Auszahlung reduzieren. Nunmehr entfällt aber die Möglichkeit von cross-matchings zwischen K-Typen (mit der Auszahlung  $3-c$ ), was die Durchschnittsauszahlung des einen K-Typs deutlich reduziert. Man erhält den kritischen Wert für  $h=n-1$  als

$$\tilde{m}(h = n - 1) = \frac{6 - \lambda + \lambda(n - 1)}{\lambda(n - 2) - (n - 1)(1 + c) + 1}. \tag{7-28}$$

Offensichtlich gilt

$$\lim_{\lambda \rightarrow \frac{(n-1)(1+c)-1}{n-2}} \tilde{m}(h = n - 1) = +\infty, \tag{7-29}$$

sowie

$$\frac{(n - 1)(1 + c) - 1}{n - 2} \cong 1 + c, \tag{7-30}$$

für  $n$  „hinreichend groß“.

Die folgende Graphik zeigt den Verlauf des kritischen Wertes für  $m$  und den Fall  $h=n-1$ , wenn, wie zuvor,  $n=100$  und  $c=0.5$ . Der präzise Wert der Unendlichkeitsstelle der Funktion ermittelt sich als  $\frac{(n-1)(1+c)+1}{n-2} = 1.505$ , was leicht oberhalb des approximativen Wertes 1.5 liegt.

$$n := 100 \quad c := 0.5 \quad \lambda := 1.51, 1.511, \dots, 3 \quad f(\lambda) := \frac{6 - \lambda + \lambda \cdot (n - 1)}{((n - 1) \cdot (\lambda - 3) + (n - 1) \cdot (2 - c) + 1 - \lambda)}$$

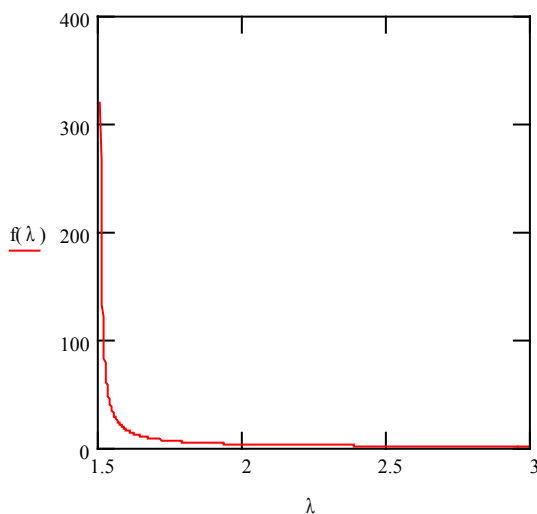


Abb.15: Der kritische Wert  $\tilde{m}$  für  $h=n-1$  in Abhängigkeit von  $\lambda$

Wie man sieht, wird Bedingung (7-19) im Fall  $h=n-1$  lediglich dann unerfüllbar, wenn  $\lambda \rightarrow 1+c$ .

Wenn man also von den beiden Unendlichkeitsstellen absieht, ist Bedingung (7-19) auch für die Randfälle  $h=1$  und  $h=n-1$  leicht erfüllbar. Damit der erste K-Typ sich gegen  $n-1$  Hawks durchsetzt, bzw. der letzte Hawk noch verdrängt werden kann reicht es, wenn es zu einer hinreichenden Anzahl von matchings kommt. Dann existiert eine positive Wahrscheinlichkeit dafür, daß der Hawk hinreichend oft als bekannter Hawk getroffen wird. Dieses Resultat ist direkt analog zu den Ergebnissen unserer deterministischen Modelle, in denen der durchschnittliche Lernerfolg, ausgedrückt durch den Parameter  $k = \frac{n}{T} \left( 1 - \left( \frac{n-1}{n} \right)^T \right)$ , hinreichend

groß sein mußte ( $k$  hinreichend klein), damit die kooperativen Strategien überleben konnten. Ein Unterschied beider Modelltypen besteht darin, daß es im stochastischen Modell ausreicht, wenn eine positive Wahrscheinlichkeit für ein bestimmtes matching-Resultat existiert, welches dann die gewünschte Selektionsrichtung liefert (wie etwa die Annahme  $P_{H,K} = m-1$ ). Ein derartiges Ereignis mag zwar extrem unwahrscheinlich sein, im „long run“ für  $t \rightarrow \infty$  geschieht es dennoch „irgendwann“ mit Sicherheit. Im deterministischen Modellrahmen hingegen mußten Durchschnittsgrößen bestimmte kritische Parameterwerte übersteigen oder unterschreiten. Insofern kann man sagen, daß die stochastische Modellierung weniger restriktiv ist und gegenüber deterministischen Modellen die Entstehung von Kooperation begünstigt.

Auch bezüglich des Parameters  $\lambda$  ergeben sich strukturelle Ähnlichkeiten zu den Ergebnissen des deterministischen Modellkomplexes. Die globale Stabilität des inneren Gleichgewichtes erforderte dort

$$c < \min\{c_1, c_2\}, c_1 = \frac{\lambda(1-k)[(1-k)(\lambda-6)+3]}{2+(1-k)(\lambda-6)}, c_2 = (1-k)(2\lambda-3).$$

Natürlich ist zu fordern:

i)

$$\begin{aligned} c_2 > 0 &\Leftrightarrow \\ (1-k)(2\lambda-3) > 0 &\Leftrightarrow \\ \lambda > \frac{3}{2} & \end{aligned} \quad (7-31)$$

ii)

$$\begin{aligned} c_1 > 0 &\Leftrightarrow \\ \lambda(1-k)[(1-k)(\lambda-6)+3] < 0 &\Leftrightarrow \end{aligned}$$

$$\lambda < 3 - \frac{3k}{1-k} \quad (7-32)$$

Auch dort erhält man also eine Ober- und eine Untergrenze für  $\lambda$ . Allerdings ist zu bemerken, daß die Untergrenze im deterministischen Modell bei 1.5 fixiert ist, während sie im stochastischen Modell von den Kosten des Lernens  $c$  abhängt. Insbesondere kann im stochastischen Modell die Untergrenze für hinreichend kleine  $c$  beliebig dem defektiven Payoff von eins angenähert werden. Auch wenn sich also der Wert der work-alone-Option kaum vom Payoff bei einmütiger Defektion unterscheidet, kann kooperatives Verhalten dennoch gedeihen, vorausgesetzt die Lernkosten sind hinreichend niedrig. Insofern erweist sich das stochastische Modell wiederum als weniger restriktiv.<sup>16</sup>

---

<sup>16</sup> Es ist vermutlich redlicher davon zu sprechen, das stochastische Modell sei flexibler. Denn immerhin ist ja die Untergrenze dort für  $c > 0.5$  größer als  $3/2$ , und insofern kann auch das stochastische Modell durchaus restriktiver sein als das deterministische. Gemeint ist, daß im stochastischen Modell Parameterkonstellationen **möglich** sind, für die selbst im worst case Kooperation noch Überlebenschancen hat. Dies ist im deterministischen Modell definitiv ausgeschlossen.

## 7.4 Ein Beispiel: Die Best-Reply-Dynamik

Die vorangegangene Argumentation verbleibt auf einem recht abstrakten Niveau. Ohne weitere Annahmen kann über Satz 7.2 hinaus mehr nicht ausgesagt werden, als daß unter den genannten Bedingungen in der Tat alle drei monomorphen Zustände mit positiven Häufigkeiten in der Grenzverteilung vertreten sind. Zur Veranschaulichung der Logik des stochastischen Modells und zur Ableitung etwas greifbarer Resultate ist es vielleicht angebracht, einmal beispielhaft eine Dynamik zu spezifizieren.

Es sei im folgenden davon ausgegangen, daß der Selektionsanteil des evolutionären Prozesses beschrieben werde durch eine Best-Reply-Dynamik. Betrachtet sei eine Population der absoluten und konstanten Größe  $n$ . Die Strategiemenge  $I$  bestehe wie gehabt aus den drei Strategien Hawk (H), Dove (D) und knowledgeable Dove (K):  $I = \{H, D, K\}$ . Bezeichne nun  $i(t)$  die Anzahl der Spieler, die in  $t$   $i \in I$  spielen. Dann:

$$i(t+1) = \begin{cases} n & \text{falls } \pi_i(t) > \pi_j(t) \quad \forall j \neq i \\ 0 & \text{falls } \pi_i(t) < \pi_j(t) \quad \text{für mind. ein } j \end{cases} \quad (7-33)$$

Sobald also Payoffdifferenzen auftreten, wechseln alle Spieler der Population in  $t+1$  ohne jegliche Anpassungsträgheit sofort zur erfolgreichsten Strategie. Wir nehmen dann weiterhin an, daß, nachdem eine Population in  $t+1$  in der beschriebenen Weise durch die Best-Reply-Dynamik „programmiert“ wurde, jeder Spieler unabhängig mit der Wahrscheinlichkeit  $\varepsilon$  „mutiert“ und eine der beiden anderen Strategien spielt.

Der resultierende Gesamtprozeß ist natürlich wiederum eine Markov-Kette, aber nun auf dem dramatisch reduzierten Zustandsraum  $E = \{ALLH, ALLD, ALLK\}$ . Dieser Umstand wird es uns erlauben, explizite Übergangswahrscheinlichkeiten anzugeben und explizit die Grenzverteilung zu berechnen.

- i) Beginnen wir mit der Frage, mit welcher Wahrscheinlichkeit der Zustand ALLD vom Zustand ALLH aus erreicht wird. Wir numerieren zur Vereinfachung der Notation im folgenden die Zustände durch. Sei ALLH als Zustand 1 bezeichnet, ALLK als Zustand 2 und ALLD als Zustand 3. Dann stellen wir uns nunmehr also die Frage nach der Übergangswahrscheinlichkeit  $p_{13} \equiv Pr(i_{t+1} = 3 | i_t = 1)$ . Klar ist, daß eine einzelne D-Mutation in einer ALLH-Population niemals erfolgreich sein kann, da sie notwendigerweise und unabhängig vom matching-Ergebnis immer eine durchschnittliche Auszahlung von 0 realisiert. Es werden also mindestens zwei Mutationen benötigt, um einen Übergang von ALLH nach ALLD herbei

zu führen. Hierbei sind nun zwei Möglichkeiten denkbar. Erstens kann es zum Auftreten zweier D-Mutanten kommen, und diese werden dann in einer hinreichenden Anzahl von matching-Zeitpunkten miteinander gematcht. Oder aber es kommt zum Auftreten einer D-Mutante und einer K-Mutante, und wiederum werden beide hinreichend oft gematcht. Es bezeichne  $m_1$  die erforderliche kritische Anzahl an matchings im ersten Fall, also die minimal erforderliche Anzahl von matchings des Ergebnisses Dove-Dove, so daß gilt  $\bar{\pi}_D > \bar{\pi}_H$ . Es bezeichne ferner  $\beta$  die Wahrscheinlichkeit für das Resultat Dove-Dove in einem matching-Zeitpunkt. Die Gesamtwahrscheinlichkeit, daß infolge zweier D-Mutanten ein Übergang zu ALLD erfolgt, ist dann offenbar gegeben durch  $\beta^{m_1} \varepsilon^2$ <sup>17</sup>. Bezeichne nun  $m_2$  die minimal erforderliche Anzahl von matchings zwischen einer K- und einer D-Mutante, damit gilt  $\bar{\pi}_D > \bar{\pi}_H, \bar{\pi}_K$ . Die Wahrscheinlichkeit für einen Übergang in diesem Fall ist dann  $\beta^{m_2} \varepsilon^2$ . Die Gesamtwahrscheinlichkeit für einen Übergang von ALLH zu ALLD, infolge des Auftretens der minimalen Anzahl erforderlicher Mutationen, ist dann einfach die Vereinigung der Wahrscheinlichkeiten aus beiden Fällen, also

$$p_{13} \equiv \Pr(i_{t+1} = 3 | i_t = 1) = (\beta^{m_1} + \beta^{m_2}) \varepsilon^2. \quad (7-34)$$

Mit  $(\beta^{m_1} + \beta^{m_2}) \equiv \beta_2$  können wir kurz schreiben

$$p_{13} = \beta_2 \varepsilon^2. \quad (7-35)$$

- ii) Betrachten wir nun die Frage eines Übergangs von ALLD nach ALLH. Klarerweise reicht hier bereits eine singuläre H-Mutante, um einen Übergang herbeizuführen. Ein einzelner Hawk erzielt in einer Dove-Population notwendig einen durchschnittlichen Payoff von  $\bar{\pi}_H = 6 > \bar{\pi}_D$ . Also gilt

$$p_{31} \equiv \Pr(i_{t+1} = 1 | i_t = 3) = \varepsilon. \quad (7-36)$$

---

<sup>17</sup> Schlimmstenfalls muß gelten  $m_1=m$ . Die beiden D-Mutanten werden dann in allen matching-Zeitpunkten miteinander gematcht. Mindestens unter diesem matching-Ergebnis gilt notwendig  $\bar{\pi}_D = 3 > 1 = \bar{\pi}_H$ . Aber selbst ein derartig extremes Resultat des matching-Prozesses hat immer noch positive Wahrscheinlichkeit.

- iii) Einfach gestaltet sich auch die Bestimmung der Übergangswahrscheinlichkeit von ALLK nach ALLD. Eine singuläre D-Mutante in einer K-Population erzielt natürlich die Auszahlung  $\bar{\pi}_D = 3 > 3 - c = \bar{\pi}_K$ , so daß auch hier gilt

$$p_{23} \equiv \Pr(i_{t+1} = 3 \mid i_t = 2) = \varepsilon. \quad (7-37)$$

- iv) Schwieriger gestaltet sich die Bestimmung der Wahrscheinlichkeit für einen Übergang von ALLD nach ALLK: Klar ist, daß eine einzelne K-Mutante immer einen niedrigeren Payoff erzielt als die Restpopulation der D-Spieler. Dies gilt selbst dann noch, wenn n-1 Spieler zu K-Typen mutieren. Ein Übergang durch K-Mutanten allein wäre also nur für den Fall möglich, daß die gesamte Population simultan mutiert. Allerdings sind auch Übergänge infolge simultanen Auftretens von K- und H-Mutanten denkbar. Wie sich zeigen wird, ist die genaue Kenntnis dieser Übergangswahrscheinlichkeit jedoch nicht erforderlich (insbesondere bleiben die Resultate auch für den worst case: n simultane K-Mutationen, erhalten). Sei daher  $\delta$  der kritische erforderliche Wert an K-Mutanten,  $\gamma$  ein korrespondierender kritischer Wert an H-Mutanten,  $\alpha$  die Summe der Mutanten beiden Typs und  $\bar{\alpha}$  ein minimaler kritischer Wert für die Summe aus  $\delta$  und  $\gamma$ . Bezeichne ferner  $\varphi$  die Wahrscheinlichkeit für ein entsprechendes matching-Resultat (es komme insbesondere zu einer hinreichenden Anzahl von matchings zwischen H-Mutanten und Tauben). Dann ist

$$p_{32} \equiv \Pr\{i_{t+1} = 2 \mid i_t = 3\} = \varphi \varepsilon^\alpha, \bar{\alpha} < \alpha \leq n. \quad (7-38)$$

- v) Als eigentlich kritisch für die Modellergebnisse, so wird sich wiederum herausstellen, ist die Frage, ob der Übergang von ALLH nach ALLK durch eine K-Mutante allein bewerkstelligt werden kann, oder ob hierzu mehrere Mutanten notwendig sind. Der durchschnittliche Payoff einer singulären K-Mutante in einer Population von n-1 Hawks ermittelt sich als

$$\bar{\pi}_K = (\lambda - c)P_{H_KK} - cP_{H_UK}. \quad (7-39)$$

Mit  $P_{H_KK} + P_{H_UK} = m \Leftrightarrow P_{H_KK} = m - P_{H_UK}$  ergibt sich

$$\bar{\pi}_K = \lambda P_{H_KK} - mc. \quad (7-40)$$

Der durchschnittliche Payoff der  $n-1$  Hawks ermittelt sich allgemein als

$$\bar{\pi}_H = \frac{1}{n-1} (2P_{HH} + \lambda P_{HK} + 6P_{HK}). \quad (7-41)$$

Mit  $2P_{HH} + P_{HK} + P_{HK} = m(n-1)$  und wiederum  $P_{HK} + P_{HK} = m$ , folgt

$$\bar{\pi}_H = m + \frac{1}{n-1} (P_{HK} (\lambda - 6) + 5m). \quad (7-42)$$

Die K-Mutante setzt sich durch, falls

$$\begin{aligned} \bar{\pi}_K > \bar{\pi}_H &\Leftrightarrow \\ P_{HK} > m \frac{n(1+c)+4-c}{\lambda(n-2)+6}. \end{aligned} \quad (7-43)$$

Falls also Bedingung (7-43) erfüllbar ist, das heißt: falls die Zufallsvariable  $P_{HK}$  Werte oberhalb von  $m \frac{n(1+c)+4-c}{\lambda(n-2)+6}$  annehmen kann, dann

existiert eine positive Wahrscheinlichkeit für matching-Resultate, so daß K die erfolgreichere Strategie ist. Der Übergang von ALLH nach ALLK kann dann mittels nur einer Mutante erfolgen. Ist Bedingung (7-43) hingegen nicht erfüllbar, so sind mindestens zwei Mutationen für den Übergang erforderlich. Die Fälle, in denen (7-43) erfüllbar bzw. nicht erfüllbar ist, werden im Appendix zu diesem Kapitel untersucht. Betrachtet sei also der Fall, daß eine Mutante ausreicht, um den Übergang herbeizuführen. Bezeichne  $\beta_1$  die Wahrscheinlichkeit für alle geeigneten matching-Resultate, so daß  $\bar{\pi}_K > \bar{\pi}_H$ . Die Übergangswahrscheinlichkeit ist dann gegeben durch

$$p_{12} \equiv Pr(i_{t+1} = 2 | i_t = 1) = \beta_1 \varepsilon. \quad (7-44)$$

Ist hingegen (7-43) nicht erfüllbar und sind also zwei Mutanten mindestens erforderlich, so gilt es wiederum zwei Möglichkeiten zu unterscheiden. Zum einen kann es zum simultanen Auftreten zweier K-Mutanten kommen. Werden diese hinreichend oft untereinander gematcht, so findet der Übergang statt. Bezeichne  $\beta_2$  die Wahrscheinlichkeit für alle geeigneten matching-Resultate. Dann ist die Übergangswahrscheinlichkeit gegeben durch  $\beta_2 \varepsilon^2$ . Zum zweiten kann es zum Auftreten einer K- sowie einer D-Mutante kommen. Komme es nun zu einer einerseits hinreichenden An

zahl von matchings zwischen K-Typ und Dove, so daß der Erfolg von K gewährleistet ist. Andererseits ist zu fordern, daß Dove nicht zu erfolgreich sein darf. Bezeichne  $\beta_3$  die Wahrscheinlichkeit für geeignete matching-Resultate<sup>18</sup>, so daß gilt  $\bar{\pi}_K > \bar{\pi}_H, \bar{\pi}_D$ . Die Übergangswahrscheinlichkeit ist dann gegeben durch  $\beta_3 \varepsilon^2$ , und die Gesamtübergangswahrscheinlichkeit ergibt sich als

$$p_{12} \equiv Pr(i_{t+1} = 2 | i_t = 1) = (\beta_2 + \beta_3) \varepsilon^2 \equiv \bar{\beta} \varepsilon^2. \quad (7-45)$$

- vi) Es verbleibt die Untersuchung des Übergangs von ALLK nach ALLH. Eine H-Mutante in einer Population von n-1 K-Spielern erzielt durchschnittlich

$$\bar{\pi}_H = \lambda P_{H_K K} + 6 P_{H_U K}. \quad (7-46)$$

Mit  $P_{H_U K} = m - P_{H_K K}$  erhält man

$$\bar{\pi}_H = (\lambda - 6) P_{H_K K} + 6m. \quad (7-47)$$

Für die K-Typen ermittelt man

$$\bar{\pi}_K = \frac{1}{n-1} (2P_{KK} (3-c) + P_{H_K K} (\lambda - c) - c P_{H_U K}). \quad (7-48)$$

Mit  $2P_{KK} + P_{H_K K} + P_{H_U K} = m(n-1)$  und  $P_{H_K K} + P_{H_U K} = m$  folgt

$$\bar{\pi}_K = (3-c)m + \frac{1}{n-1} (\lambda P_{H_K K} - 3m). \quad (7-49)$$

Die H-Mutante setzt sich durch, falls

$$\pi_H > \pi_K \Leftrightarrow$$

---

<sup>18</sup> Wir führen hier aus Platzgründen nicht den Nachweis, daß tatsächlich matching-Ergebnisse existieren, für die die geforderte Payoff-Relation gilt, so daß  $\beta_3$  also tatsächlich strikt positiv ist. Aber selbst wenn die Möglichkeit eines Übergangs infolge K- und D-Mutation ausgeschlossen wird, also gelten sollte  $\beta_3 = 0$ , so bleibt immer noch die Möglichkeit des Übergangs infolge zweier simultaner K-Mutationen, so daß die Gesamtübergangswahrscheinlichkeit also **immer** strikt positiv ist.



$$P_{H_{KK}} < m \frac{3n + c(n-1)}{2\lambda - 6 + n(6-\lambda)}. \quad (7-50)$$

Falls Bedingung (7-50) erfüllbar ist, reicht eine H-Mutante, um den Übergang herbeizuführen. Andernfalls ist als Minimalkonstellation vorstellbar, daß es zum Auftreten einer H-Mutante und einer D-Mutante kommt, und beide hinreichend oft miteinander gematcht werden. Man erhält also

$$q_{21} \equiv \Pr\{\omega_{t+1} = \omega_1 | \omega_t = \omega_2\} = \gamma\varepsilon, \quad (7-51)$$

falls (7-50) erfüllt werden kann und

$$q_{21} \equiv \Pr\{\omega_{t+1} = \omega_1 | \omega_t = \omega_2\} = \bar{\gamma}\varepsilon^2, \quad (7-52)$$

falls (7-50) nicht erfüllbar ist, wobei  $\gamma$ , bzw.  $\bar{\gamma}$  die Wahrscheinlichkeiten für geeignete matching-Resultate bezeichnen.

Damit sind nun alle erforderlichen Übergangswahrscheinlichkeiten beisammen. Es gilt nunmehr vier Fälle zu unterscheiden:

- 1) Pessimismus-Fall: Bedingung (7-43) nicht erfüllbar, Bedingung (7-50) erfüllbar.
- 2) Fall II: Beide Bedingungen erfüllbar.
- 3) Optimismus-Fall: Bedingung (7-43) erfüllbar, Bedingung (7-50) nicht erfüllbar.
- 4) Fall IV: Beide Bedingungen nicht erfüllbar.

Man zeigt leicht, daß die stationäre Verteilung einer Markovkette mit einem Zustandsraum von drei Zuständen allgemein gegeben ist als

$$\begin{aligned} \pi_1 &= \pi_2 \frac{p_{21} - p_{31}}{p_{12} + p_{13} + p_{31}} + \frac{p_{31}}{p_{12} + p_{13} + p_{31}}, \\ \pi_2 &= \frac{p_{32}(p_{12} + p_{13}) + p_{31}p_{12}}{p_{21}(p_{32} + p_{13} + p_{31}) + p_{31}p_{12} + p_{23}(p_{12} + p_{13} + p_{31}) + p_{32}(p_{12} + p_{13})}, \\ \pi_3 &= 1 - \pi_1 - \pi_2. \end{aligned} \quad (7-53)$$

Man beachte, daß für die Bestimmung der Grenzverteilung wiederum nur die Terme mit den maximalen Übergangswahrscheinlichkeiten relevant sind, also jene Übergänge, die mit einer minimal möglichen Anzahl von Mutanten auskom-

men. Alle Terme, die Übergänge mit mehr Mutanten widerspiegeln, konvergieren schneller gegen Null und müssen daher nicht in Betracht gezogen werden.

ad 1)

Die Markov-Matrix der maximalen Übergangswahrscheinlichkeiten in diesem Fall ist gegeben als:

$$M = \begin{bmatrix} 1 - \bar{\beta}_1 \varepsilon^2 - \beta_2 \varepsilon^2 & \bar{\beta} \varepsilon^2 & \beta_2 \varepsilon^2 \\ \gamma \varepsilon & 1 - \gamma \varepsilon - \varepsilon & \varepsilon \\ \varepsilon & \varphi \varepsilon^\alpha & 1 - \varepsilon - \varphi \varepsilon^\alpha \end{bmatrix} \quad (7-54)$$

Wir setzen in (7-53) ein und bestimmen den Grenzwert:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{p_{32}(p_{12} + p_{13}) + p_{31}p_{12}}{p_{21}(p_{32} + p_{13} + p_{31}) + p_{31}p_{12} + p_{23}(p_{12} + p_{13} + p_{31}) + p_{32}(p_{12} + p_{13})} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{\varphi \varepsilon^\alpha (\bar{\beta} \varepsilon^2 + \beta_2 \varepsilon^2) + \varepsilon \bar{\beta} \varepsilon^2}{\gamma \varepsilon (\varphi \varepsilon^\alpha + \beta_2 \varepsilon^2 + \varepsilon) + \varepsilon \bar{\beta} \varepsilon^2 + \varepsilon (\bar{\beta} \varepsilon^2 + \beta_2 \varepsilon^2 + \varepsilon) + \varphi \varepsilon^\alpha (\bar{\beta} \varepsilon^2 + \beta_2 \varepsilon^2)} \quad (7-55) \end{aligned}$$

Da  $\alpha > 1$  können wir wie folgt ausklammern:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2 [\varphi \varepsilon^\alpha (\bar{\beta} + \beta_2) + \varepsilon \bar{\beta}]}{\varepsilon^2 [\gamma \varphi \varepsilon^{\alpha-1} + \gamma \varepsilon \beta_2 + \gamma + \varepsilon \bar{\beta} + \varepsilon (\bar{\beta} + \beta_2) + 1 + \varphi \varepsilon^\alpha (\bar{\beta} + \beta_2)]} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{\varphi \varepsilon^\alpha (\bar{\beta} + \beta_2) + \varepsilon \bar{\beta}}{\varphi \varepsilon^{\alpha-1} [\gamma + \varepsilon (\bar{\beta} + \beta_2)] + 1 + \gamma + 2\varepsilon (\bar{\beta} + \beta_2)} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_2 &= 0 \quad (7-56) \end{aligned}$$

Für den Zustand ALLH ergibt sich damit:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_1 &= \lim_{\varepsilon \rightarrow 0} \pi_2 \frac{p_{21} - p_{31}}{p_{12} + p_{13} + p_{31}} + \lim_{\varepsilon \rightarrow 0} \frac{p_{31}}{p_{12} + p_{13} + p_{31}} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi &= 0 + \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\bar{\beta} \varepsilon^2 + \beta_2 \varepsilon^2 + \varepsilon} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\bar{\beta} \varepsilon + \beta_2 \varepsilon + 1} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi &= 1 \quad (7-57) \end{aligned}$$

Im Pessimismus-Fall hält sich der Prozeß also lediglich im monomorphen Zustand ALLH auf.

ad ii)

Die Markov-Matrix der maximalen Übergangswahrscheinlichkeiten in diesem Fall ist gegeben als:

$$M = \begin{bmatrix} 1 - \beta_1 \varepsilon - \beta_2 \varepsilon^2 & \beta_1 \varepsilon & \beta_2 \varepsilon^2 \\ \gamma \varepsilon & 1 - \gamma \varepsilon - \varepsilon & \varepsilon \\ \varepsilon & \varphi \varepsilon^\alpha & 1 - \varepsilon - \varphi \varepsilon^\alpha \end{bmatrix} \quad (7-58)$$

Man erhält für den Zustand ALLK:

$$\lim_{\varepsilon \rightarrow 0} \pi_2 = \lim_{\varepsilon \rightarrow 0} \frac{\varphi \varepsilon^\alpha (\beta_1 \varepsilon + \beta_2 \varepsilon^2) + \varepsilon \beta_1 \varepsilon}{\gamma \varepsilon (\varphi \varepsilon^\alpha + \beta_2 \varepsilon^2 + \varepsilon) + \beta_1 \varepsilon^2 + \varepsilon (\beta_1 \varepsilon + \beta_2 \varepsilon^2 + \varepsilon) + \varphi \varepsilon^\alpha (\beta_1 \varepsilon + \beta_2 \varepsilon^2)} \quad (7-59)$$

Wir können wiederum  $\varepsilon^2$  ausklammern:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2 [\varphi \varepsilon^{\alpha-1} (\beta_1 + \varepsilon \beta_2) + \beta_1]}{\varepsilon^2 [\gamma \varphi \varepsilon^{\alpha-1} + \beta_2 \gamma \varepsilon + \gamma + 2\beta_1 + \varepsilon \beta_2 + 1 + \varphi \varepsilon^{\alpha-1} (\beta_1 + \varepsilon \beta_2)]} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_2 &= \frac{\beta_1}{1 + \gamma + 2\beta_1} \end{aligned} \quad (7-60)$$

Für den Zustand ALLH ergibt sich:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_1 &= \lim_{\varepsilon \rightarrow 0} \left[ \pi_2 \frac{\gamma \varepsilon - \varepsilon}{\varepsilon \beta_1 + \varepsilon^2 \beta_2 + \varepsilon} \right] + \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\varepsilon \beta_1 + \varepsilon^2 \beta_2 + \varepsilon} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_1 &= \frac{\beta_1}{1 + \gamma + 2\beta_1} \frac{(\gamma - 1)}{(1 + \beta_1)} + \frac{1}{1 + \beta_1} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_1 &= \frac{\beta_1 (\gamma - 1) + 1 + \gamma + 2\beta_1}{(1 + \gamma + 2\beta_1)(1 + \beta_1)} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_1 &= \frac{1 + \gamma + \beta_1 (1 + \gamma)}{(1 + \gamma)(1 + \beta_1) + 2\beta_1 (1 + \beta_1)} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_1 &= \frac{(1 + \gamma)(1 + \beta_1)}{(1 + \beta_1)(1 + \gamma + 2\beta_1)} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_1 &= \frac{1 + \gamma}{(1 + \gamma + 2\beta_1)} \end{aligned} \quad (7-61)$$

Die Häufigkeit für den Zustand ALLD bestimmt sich residual:

$$\lim_{\varepsilon \rightarrow 0} \pi_3 = \frac{1 + \gamma + 2\beta_1 - \beta - (1 + \gamma)_1}{1 + \gamma + 2\beta_1} = \frac{\beta_1}{1 + \gamma + 2\beta_1} = \lim_{\varepsilon \rightarrow 0} \pi_2 \quad (7-62)$$

Im Fall II haben also alle drei Zustände positive Häufigkeiten in der Grenzverteilung.

ad iii)

Die Markov-Matrix der maximalen Übergangswahrscheinlichkeiten in diesem Fall ist gegeben als:

$$M = \begin{bmatrix} 1 - \beta_1 \varepsilon - \beta_2 \varepsilon^2 & \beta_1 \varepsilon & \beta_2 \varepsilon^2 \\ \bar{\gamma} \varepsilon^2 & 1 - \bar{\gamma} \varepsilon^2 - \varepsilon & \varepsilon \\ \varepsilon & \varphi \varepsilon^\alpha & 1 - \varepsilon - \varphi \varepsilon^\alpha \end{bmatrix} \quad (7-63)$$

Man erhält:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \pi_2 &= \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2 [\varphi \varepsilon^{\alpha-1} (\beta_1 + \varepsilon \beta_2) + \beta_1]}{\varepsilon^2 [\bar{\gamma} \varphi \varepsilon^\alpha + \beta_2 \bar{\gamma} \varepsilon^2 + \bar{\gamma} \varepsilon + 2\beta_1 + \varepsilon \beta_2 + 1 + \varphi \varepsilon^{\alpha-1} (\beta_1 + \varepsilon \beta_2)]} \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} \pi_2 &= \frac{\beta_1}{1 + 2\beta_1} \end{aligned} \quad (7-64)$$

Einsetzen für  $\lim_{\varepsilon \rightarrow 0} \pi_1$  und Vereinfachen führt dann zu:

$$\lim_{\varepsilon \rightarrow 0} \pi_1 = \frac{1}{1 + 2\beta_1} \quad (7-65)$$

Mithin:

$$\lim_{\varepsilon \rightarrow 0} \pi_3 = \frac{\beta_1}{1 + 2\beta_1} \quad (7-66)$$

Auch im Optimismusfall haben also alle drei Zustände positive Häufigkeiten.

ad iv)

Bleibt der Vollständigkeit halber der vierte denkmögliche Fall zu untersuchen, in dem weder Bedingung (7-43) noch (7-50) erfüllbar sind.

Die Markov-Matrix der maximalen Übergangswahrscheinlichkeiten in diesem Fall ist gegeben als:

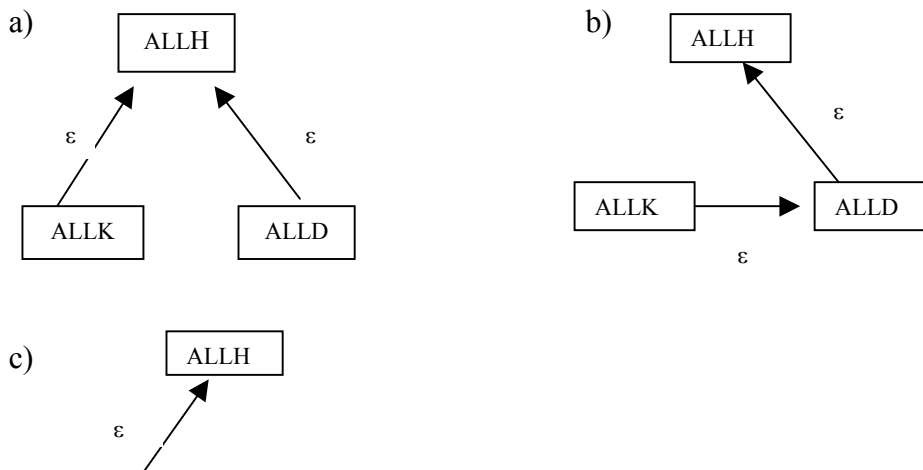
$$M = \begin{bmatrix} 1 - \bar{\beta}\varepsilon^2 - \beta_2\varepsilon^2 & \bar{\beta}\varepsilon^2 & \beta_2\varepsilon^2 \\ \bar{\gamma}\varepsilon^2 & 1 - \bar{\gamma}\varepsilon^2 - \varepsilon & \varepsilon \\ \varepsilon & \varphi\varepsilon^\alpha & 1 - \varepsilon - \varphi\varepsilon^\alpha \end{bmatrix} \quad (7-67)$$

Verfährt man wie in (i) – (iii) in der nunmehr hinlänglich bekannten Weise, so stellt man fest, daß, wie im pessimistischen Szenario, gilt

$$\lim_{\varepsilon \rightarrow 0} \pi_1 = 1. \quad (7-68)$$

### 7.5 Diskussion der Resultate

Das Überleben von Kooperation hängt, wie bereits im allgemeinen stochastischen Modell, entscheidend davon ab, ob eine singuläre K-Mutante eine Hawk-Population invadieren kann oder nicht. Lediglich in den beiden Fällen II und III (Optimismus), in denen Bedingung (7-43) erfüllbar ist, überleben die kooperativen Strategien. Die Intuition für dieses Ergebnis ist dieselbe wie bereits im allgemeinen Fall, hier aber weit offensichtlicher: Lediglich wenn eine einzelne K-Mutante ALLH invadieren kann, existieren auch für die kooperativen Zustände ALLD und ALLK Bäume, die mit der minimal möglichen Anzahl von zwei Mutationen auskommen. Im vorliegenden Fall der Best-Reply-Dynamik ist es möglich, sich die möglichen Bäume mit minimaler Mutationsanzahl (maximaler Übergangswahrscheinlichkeit) explizit zu überlegen. Wir beginnen mit Fall 1 (Pessimismusfall). Man erinnere sich, daß ein i-Baum genau dadurch definiert war, daß jeder Zustand  $i' \in E \setminus i$  Ausgangspunkt für genau einen Pfeil ist und für alle  $i' \in E \setminus i$  ein Pfad spezifiziert wird (eine Sequenz von Pfeilen), auf dem Zustand i erreicht werden kann. Bei einem übersichtlichen Ereignisraum wie im vorliegenden Fall kann man tatsächlich leicht alle möglichen i-Bäume bestimmen. Für den Zustand ALLH erhält man folgende Möglichkeiten (an den Pfeilen steht die Größenordnung der jeweiligen Übergangswahrscheinlichkeit):



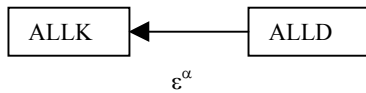


Abbildung 16: Bäume mit minimaler Mutationsanzahl für ALLH im Fall I

Man sieht, daß ALLH im Fall I offenbar genau über zwei  $Y_1^2$ -Bäume verfügt, also zwei Bäume, in denen alle vorkommenden Übergänge lediglich eine Mutation benötigen. Wenden wir uns nun den kooperativen Zuständen, zunächst ALLK, zu.

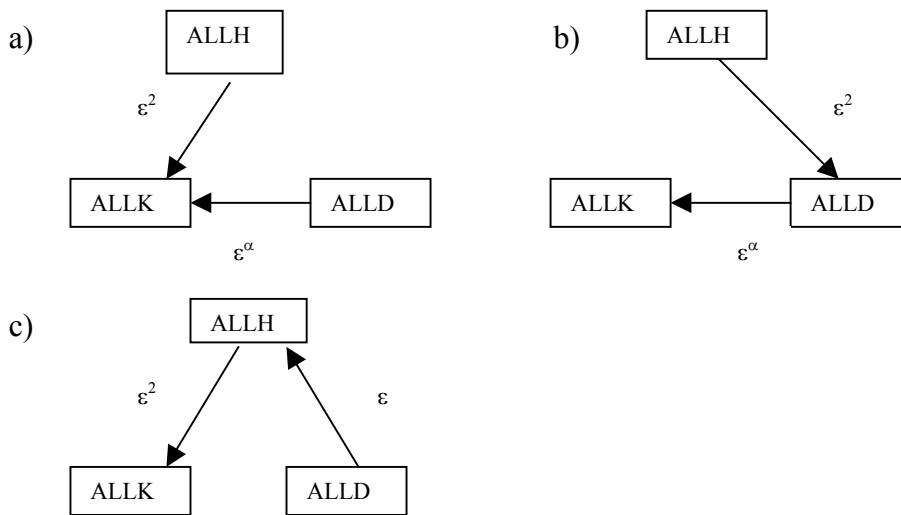


Abbildung 17: Bäume mit minimaler Mutationsanzahl für ALLK im Fall I

Wie man sieht, verfügt der Zustand ALLK über keinen einzigen  $Y_2^2$ -Baum. Jeder der drei möglichen Bäume impliziert mindestens drei Mutationen. Ähnlich stellt sich die Lage für ALLD dar:

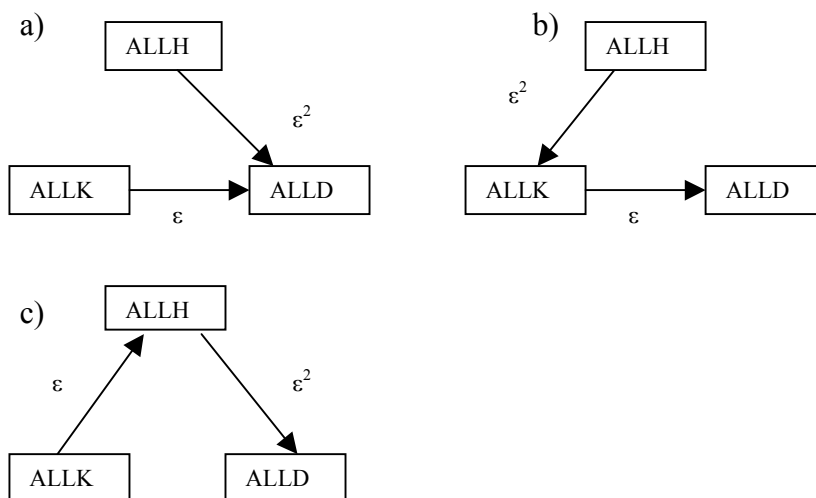


Abbildung 18: Bäume mit minimaler Mutationsanzahl für ALLD im Fall I

Wie ALLK, so verfügt auch ALLD über keinen  $Y_3^2$ -Baum. Die Folge ist, daß die Terme im Zähler der stationären Verteilung für ALLK und ALLD mindestens  $\epsilon$  in der dritten Potenz enthalten und deshalb bei der Grenzwertbetrachtung schneller

gegen Null konvergieren. Unter den Bedingungen von Fall I kann als einziger Zustand nur ALLH in der Grenzverteilung übrig bleiben.

Dies ändert sich jedoch schlagartig, wenn man annimmt, daß (7-43) erfüllbar ist. Betrachten wir Fall II, also (7-43) als auch (7-50) erfüllbar. Klar ist, daß ALLH auch weiterhin über zwei  $Y_1^2$ -Bäume verfügt. Wir wenden uns daher direkt ALLK zu:

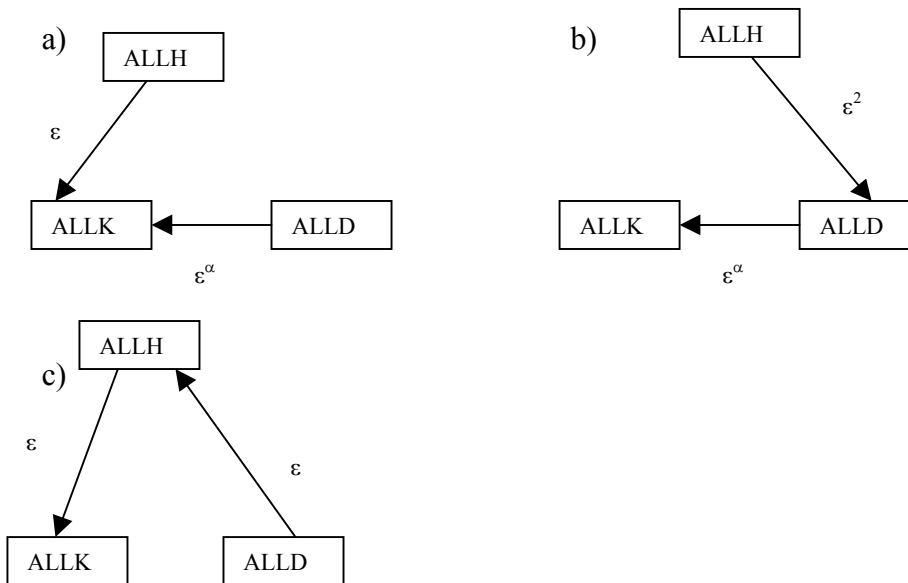


Abbildung 19: Bäume mit minimaler Mutationsanzahl für ALLK im Fall II

Die Tatsache, daß wenn (7-43) erfüllbar ist, eine singuläre K-Mutante ausreicht, eine Hawk-Population zu unterwandern, führt offensichtlich dazu, daß unter (c) nunmehr ein  $Y_2^2$ -Baum entsteht. Dies ist die tiefere Ursache dafür, daß ALLK nunmehr positive Häufigkeit in der Grenzverteilung aufweisen **muß**.

Für ALLD findet man, daß nunmehr im Unterfall (b) ein  $Y_3^2$ -Baum entsteht:

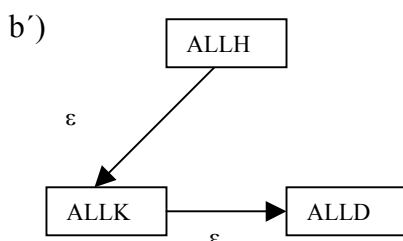
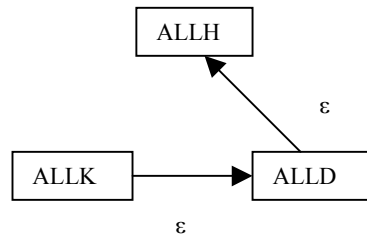


Abbildung 20:  $Y_3^2$ -Baum im Fall II

Bemerkenswert ist sicherlich auch, daß Defektion niemals gänzlich verschwindet, sondern stets mit positiver Häufigkeit in der Grenzverteilung vertreten ist. Bemerkenswert ist diese Tatsache deshalb, weil in den Fällen III und IV immerhin davon

ausgegangen wird, daß Bedingung (7-50) nicht erfüllbar ist, also mindestens zwei simultane Mutationen erforderlich sind, um den Zustand ALLK zu verlassen. Dies könnte zu der Vermutung verleiten, daß, zumindest im optimistischen Fall III – also dem Fall, daß wohl ALLH mittels einer K-Mutante invadiert werden kann, nicht aber umgekehrt – der Zustand ALLH nur noch Häufigkeit Null haben kann. Daß diese Vermutung fehlläuft, zeigt sofort eine nochmalige Betrachtung der möglichen Bäume für ALLH. Man stellt nämlich fest, daß der Unterfall (b)



sowohl im Fall III, als auch im Fall IV unverändert erhalten bleibt. Der Umstand, daß sowohl ALLK als auch ALLD jeweils durch eine einzelne Mutante invadiert werden können, sichert Defektion stets das Überleben. Im übrigen ist dieses Resultat strukturell analog zu den Ergebnissen der deterministischen Modelle. Auch dort war langfristig entweder Konvergenz gegen ein gemischtes Gleichgewicht mit Beteiligung der Hawks zu erwarten gewesen oder aber die Konvergenz gegen das reine Hawk-Gleichgewicht.

Man kann sich die Resultate unseres Beispiels auch noch wie folgt plausibel machen (vgl. Vega-Redondo 1996, S. 132). Die Grenzverteilung ermittelt diejenigen Zustände, die über maximales stochastisches Potential verfügen. Letzteres kann im vorliegenden Modellkontext einfach gedeutet werden als die Differenz zwischen der Anzahl von Mutationen, die mindestens erforderlich sind, um einen Zustand zu erreichen und der minimalen Zahl von Mutationen, durch die derselbe Zustand verlassen werden kann. Rekapitulieren wir die vier möglichen Konstellationen noch einmal:

Fall	7-43 erfüllbar ?	7-50 erfüllbar ?	Kooperation ?	Stochastische Potentiale
I	NEIN	JA	NEIN	SP(ALLH) = -1 SP(ALLK) = +1 SP3 (ALLD)= 0
II	JA	JA	JA	SP(ALLH)= 0 SP(ALLK)= 0 SP(ALLD)= 0
III	JA	NEIN	JA	SP(ALLH)= 0 SP(ALLK)= 0 SP(ALLD)= 0
IV	NEIN	NEIN	NEIN	SP(ALLH)= - 1 SP(ALLK)= + 1



				SP(ALLD)= 0
--	--	--	--	-------------

Tabelle 14: Stochastische Potentiale

Man sieht, daß in den Fällen II und III die Erfüllbarkeit von (7-43) zu einer Angleichung der stochastischen Potentiale führt. Alle drei Zustände werden nunmehr genauso leicht erreicht, wie sie auch verlassen werden können.

Betrachtet man die Fälle II und III, so würde man außerdem intuitiv erwarten, daß im Optimismus-Szenario die Wahrscheinlichkeit für die kooperativen Zustände größer ist als im Fall II. Diese Intuition läßt sich formal leicht bewahrheiten. Es war im Fall II:

$$\lim_{\varepsilon \rightarrow 0} \pi_2 = \lim_{\varepsilon \rightarrow 0} \pi_3 = \frac{\beta_1}{1 + \gamma + 2\beta_1}, \quad (7-69)$$

und im Fall III

$$\lim_{\varepsilon \rightarrow 0} \pi_2 = \lim_{\varepsilon \rightarrow 0} \pi_3 = \frac{\beta_1}{1 + 2\beta_1}. \quad (7-70)$$

Selbstverständlich gilt

$$\frac{\beta_1}{1 + 2\beta_1} > \frac{\beta_1}{1 + \gamma + 2\beta_1} \Leftrightarrow \gamma > 0. \quad (7-71)$$

## 7.6 Appendix: Erfüllbarkeit der Bedingungen (7-43) und (7-50)

Bedingung (7-43) war gegeben als

$$P_{H_K K} > m \frac{n(1+c)+4-c}{\lambda(n-2)+6} \equiv \tilde{m}_1. \quad (7-43)$$

Damit (7-43) erfüllt werden kann ist natürlich zu fordern:

$$\tilde{m}_1 < \max\{\text{supp}P_{H_K K}\} \quad (7-72)$$

Offensichtlich kann eine singuläre K-Mutante bei insgesamt  $m$  matchings maximal  $m-1$  mal einen bekannten Hawk treffen (nämlich genau den Hawk aus dem ersten matching und der dort notwendigerweise noch unbekannt ist). Es gilt also  $\max\{\text{supp}P_{H_K K}\} = m-1$  und wir fordern somit

$$\begin{aligned} m \frac{n(1+c)+4-c}{\lambda(n-2)+6} < m-1 &\Leftrightarrow \\ \lambda(n-2)+6 < m[n(\lambda-1-c)-2(\lambda-1)+c]. \end{aligned} \quad (7-73)$$

Offensichtlich gilt für alle  $n > 2$   $\lambda(n-2)+6 > 0$ , so daß zunächst zu fordern ist

$$\begin{aligned} [n(\lambda-1-c)-2(\lambda-1)+c] > 0 &\Leftrightarrow \\ n(\lambda-1-c) > 2(\lambda-1)-c. \end{aligned} \quad (7-74)$$

Man macht sich zunächst klar, daß der Fall  $(\lambda-1-c) < 0$  irrelevant ist, denn:

$$n < \frac{2(\lambda-1)-c}{(\lambda-1-c)} = 1 + \frac{(\lambda-1)}{(\lambda-1-c)} < 1 \quad (7-75)$$

Daher:  $(\lambda-1-c) > 0$

Dann ist also bezüglich  $n$  zu fordern:

$$n > \frac{2(\lambda-1)-c}{(\lambda-1-c)} \equiv \tilde{n}_1, \quad (7-76)$$

eine Bedingung, die uns als (7-12) bereits aus dem allgemeinen stochastischen Modell bekannt ist. Dann muß bezüglich  $m$  gelten:

$$m > \frac{\lambda(n-2)+6}{[n(\lambda-1-c)-2(\lambda-1)+c]} \equiv \tilde{m}_1. \quad (7-77)$$

(7-77) kann auch sofort aus (7-19) durch Einsetzen von  $h = n - 1$  abgeleitet werden. Für (7-77) wissen wir, daß sie lediglich in der Nähe der Unendlichkeitsstelle wirklich restriktiv wird. Die Unendlichkeitsstelle ist natürlich  $\lambda^\infty = 1 + c \frac{n-1}{n-2}$  und es ist klar, daß  $\lim_{\lambda \rightarrow \lambda^\infty} \tilde{m}_1 = +\infty$ . Ferner:

$$\lim_{\lambda \rightarrow 3} \tilde{m}_1 = \frac{3n}{(2-c)(n-2)-c} \quad (7-78)$$

Man zeigt leicht, daß für alle  $n > 3$  gilt  $\lim_{\lambda \rightarrow 3} \tilde{m}_1 < 3$ .

Außerdem ermittelt man:

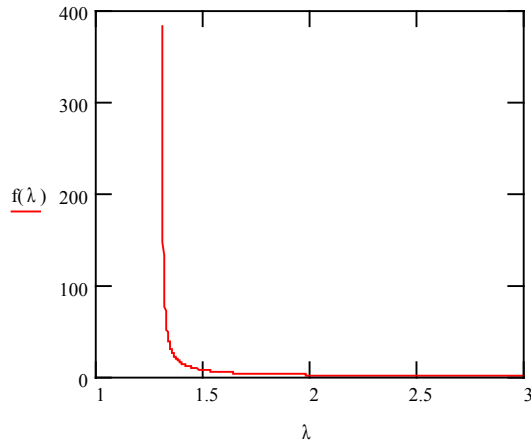
$$\frac{\partial}{\partial \lambda} \tilde{m}_1 = -(4+n+c(n-1)) \frac{(n-2)}{[(\lambda-1-c)(n-2)-c]^2} < 0 \quad \forall n > 0, \quad (7-79)$$

sowie

$$\frac{\partial^2}{\partial \lambda^2} \tilde{m}_1 = (4+n+c(n-1)) \frac{2(n-2)^2}{[(\lambda-1-c)(n-2)-c]^3} > 0. \quad (7-80)$$

Der kritische Wert  $\tilde{m}_1$  kommt also aus dem Unendlichen und konvergiert dann konvex gegen einen endlichen positiven Wert kleiner 3. Ein typischer Verlauf ist in der folgenden Abbildung wiedergegeben:

$$n := 50 \quad c := 0.3 \quad \lambda := 1.31, 1.311.. 3 \quad f(\lambda) := \frac{\lambda \cdot (n-2) + 6}{((\lambda-1-c) \cdot (n-2) - c)}$$


 Abb. 21: Verlauf von  $\tilde{m}_1$ 

Man sieht, daß für den überwiegenden Bereich des zulässigen Parameterraumes (7-43) kaum restriktiv ist. In der Regel genügt eine positive, aber durchaus kleine Anzahl von matchings, damit die Bedingung erfüllt werden kann. Lediglich in der Nähe der Unendlichkeitsstelle wird (7-43) zunehmend restriktiver.

Wenden wir uns nun Bedingung (7-50) zu. Diese war gegeben als

$$P_{H_K K} < m \frac{3n + c(n-1)}{2\lambda - 6 + n(6-\lambda)}. \quad (7-50)$$

Hier sind zwei Fälle zu unterscheiden, nämlich a)  $m \leq n-1$  sowie (b)  $m > n-1$ .

ad a) In diesem Fall ist denkbar, daß die singuläre H-Mutante in jedem Zeitpunkt einen K-Typen trifft, für den sie noch unbekannt ist. Es gilt also

$$\min\{\text{supp}P_{H_K K}\} = 0. \quad (7-81)$$

Zu fordern daher:

$$m \frac{3n + c(n-1)}{2\lambda - 6 + n(6-\lambda)} > 0 \Leftrightarrow n > \frac{c}{3+c} < 1 \quad (7-82)$$

(7-50) ist also immer erfüllbar.

ad b) Es gilt offensichtlich

$$\min\{\text{supp}P_{H_K K}\} = m - (n-1). \quad (7-83)$$

Zu fordern ist daher

$$m \frac{3n + c(n-1)}{2\lambda - 6 + n(6-\lambda)} > m - (n-1) \Leftrightarrow \\ (n-1)[(n-1)(\lambda-6) + \lambda] > m[(n-2)(3-\lambda) - c(n-1)]. \quad (7-84)$$

Offensichtlich gilt  $[(n-1)(\lambda-6) + \lambda] > 0$ . Falls nun  $[(n-2)(3-\lambda) - c(n-1)] < 0$ , so ist offenbar lediglich  $m > 0$  zu verlangen. Falls hingegen  $[(n-2)(3-\lambda) - c(n-1)] > 0 \Leftrightarrow \lambda < 3 - c \frac{n-1}{n-2}$ , so ist zu fordern

$$m \frac{3n + c(n-1)}{2\lambda - 6 + n(6-\lambda)} > m - (n-1) \Leftrightarrow \\ m < \frac{(n-1)[(n-1)(\lambda-6) + \lambda]}{[(n-2)(3-\lambda) - c(n-1)]} \equiv \tilde{m}_2. \quad (7-85)$$

Bezeichne  $\lambda_2^\infty = 3 - c \frac{n-1}{n-2}$  die Unendlichkeitsstelle. Dann ist klar:  $\lim_{\lambda \rightarrow \lambda_2^\infty} \tilde{m}_2 = +\infty$

Ferner ermittelt man

$$\frac{\partial}{\partial \lambda} \tilde{m}_2 = (n-1)(n-2) \frac{n(3+c) - c}{[(n-2)(3-\lambda) - c(n-1)]^2} > 0 \quad \forall n > 2, \quad (7-86)$$

$$\frac{\partial^2}{\partial \lambda^2} \tilde{m}_2 = 2(n-1)(n-2)^2 \frac{n(3+c) - c}{[(n-2)(3-\lambda) - c(n-1)]^3} > 0 \quad \forall n > 2, \quad (7-87)$$

sowie

$$\lim_{\lambda \rightarrow 1} \tilde{m}_2 = \frac{(n-1)(5n-4)}{(n-1)(2-c) - 2} > 0 \quad \forall n > 3. \quad (7-88)$$

Der kritische Wert  $\tilde{m}_2$  beginnt also bei dem positiven Wert  $\lim_{\lambda \rightarrow 1} \tilde{m}_2$  und steigt dann streng monoton und konvex an, um bei  $\lambda_2^\infty$  im Unendlichen zu verschwinden. Ein typischer Verlauf ist:

$$n := 50 \quad c := 0.3 \quad \lambda := 1, 1.001, \dots, 2.65 \quad f(\lambda) := \frac{(n-1) \cdot ((n-1) \cdot (6-\lambda) + \lambda)}{((3-\lambda) \cdot (n-2) - c \cdot (n-1))}$$

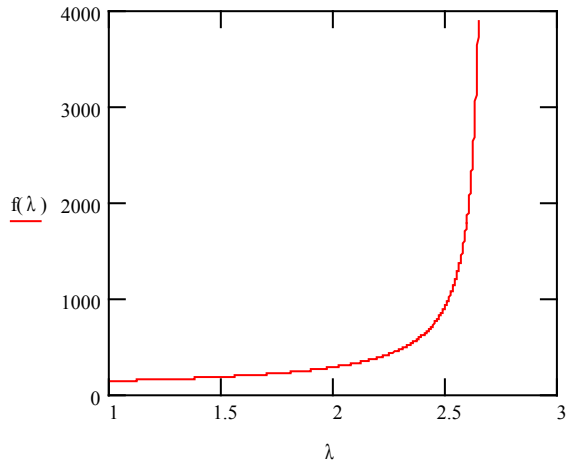


Abb. 22: Verlauf von  $\tilde{m}_2$

Insgesamt erhält man das folgende stilisiertes Bild:

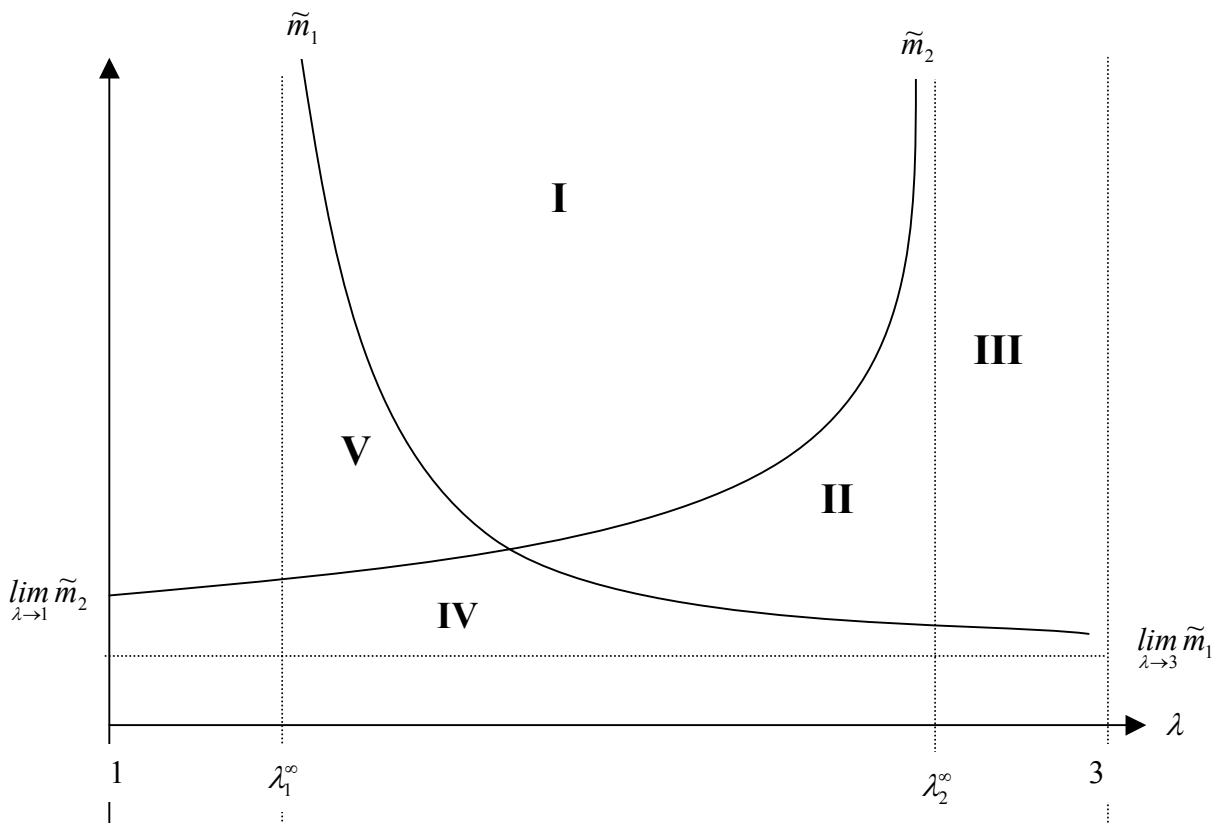


Abbildung 23: Erfüllbarkeit der Bedingungen (7-43) und (7-50)

Die unterschiedlichen Sektoren sind durch folgende Konstellationen charakterisiert:

Sektor	7-43	7-50	Kooperation
I	erfüllbar	nicht erfüllbar	Ja
II	erfüllbar	erfüllbar	Ja
III	erfüllbar	erfüllbar	Ja
IV	nicht erfüllbar	erfüllbar	Nein
V	nicht erfüllbar	nicht erfüllbar	Nein

Tabelle 15: Konstellationen im Best-Reply-Beispiel

## 8. Zusammenfassung

Die vorliegende Arbeit hat ihren Ausgangspunkt in dem experimentellen Befund, daß kooperatives Verhalten ein persistentes Phänomen in Dilemmaspielen ist, sowohl 2-Personen-GD-Spielen als auch N-Personen-GD-Spielen (öffentliche Gut Spiele). Eine disaggregierte Betrachtung der Daten enthüllt dabei die Existenz unterschiedlicher Spielertypen. Als stilisiertes Faktum wurde in Kapitel 2 die Koexistenz dreier Typen herausgearbeitet, nämlich des Defektors, des (unbedingten) Kooperationspartners und eines Typs, den man als bedingt kooperativ bezeichnen könnte. Die existierenden Modelle, die grundsätzlich geeignet erscheinen könnten, den oben skizzierten Befund zu erklären, sind letztlich signaltheoretischer Art. Das secret-handshake-Modell von Robson (1990) unterstellt, daß ein mutanter kooperativer Typ in der Lage ist, andere Mutanten anhand eines (perfekten) Signals zu identifizieren. So wird die selektive Bildung von „insider groups“ möglich, die Effizienzgewinne aus einmütiger Kooperation realisieren. Kapitel 3.4 dieser Arbeit machte deutlich, daß Robsons Modell jedoch nicht geeignet ist, den experimentellen Befund einer gemischten Population zu erklären. Die mutante Strategie im secret-handshake-Modell erweist sich als letztlich zu erfolgreich gegenüber den übrigen Strategien, Defektion und unbedingte Kooperation. Weit schwerer wiegt aber sicherlich das Problem der mangelnden Robustheit der Robsonschen Resultate. Eine intelligente defektive Mutante, die ebenfalls in der Lage ist, das Signal zu erkennen und ihr Verhalten hierauf zu konditionieren, erweist sich als superior zu der kooperativen Mutante, und es bleibt als einzig mögliche, gleichgewichtige Konstellation eine Population aus intelligenten und simplen Defektoren übrig.

Auf den ersten Blick befriedigendere Ergebnisse liefert das Modell von Frank (1988). Zentral für die Franksche Theorie ist die Existenz eines Signals, das in einem bestimmten Maße fälschungssicher sein muß. Dieses Signal liefert dann mit hinreichender Verlässlichkeit die Information über den Typus des jeweiligen Gegenspielers. Auf diese Information kann dann das Verhalten, insbesondere des kooperativen Typs, konditioniert werden. Hauptergebnis des Frank-Modells ist ein global stabiles, Zwei-Typen-Gleichgewicht, also ein Populationszustand, in dem Defektoren und konditional kooperative Typen koexistieren.

Das Frank-Modell wurde ausführlich in Kapitel 3 dieser Arbeit dargestellt und eingehender Kritik unterzogen. Vor dem Hintergrund der experimentellen psychologischen, aber auch ökonomischen Literatur (vgl. Kapitel 3.3.3) sind Zweifel berechtigt, ob Menschen tatsächlich in der Lage sind, Signale zu erkennen. Unstrittig ist dabei weitgehend, daß derartige Signale – wie willentlich unkontrollierbare Reaktionen des vegetativen Nervensystems beim Lügen – tatsächlich existie-



ren. Der experimentelle Befund spricht hier recht eindeutig für derartige Phänomene (vgl. Zuckerman et al. 1981). Weit strittiger ist jedoch die Frage, ob Probanden auch in signifikantem Ausmaß solche Signale wahrnehmen. Sowohl die Arbeit von Gilovich et al (1997), als auch die Arbeit von Ockenfels und Selten (1998) lassen hier Zweifel als durchaus berechtigt erscheinen.

Selbst wenn man dem experimentellen Befund den Status „unklar“ zuweist und im weiteren außer Acht läßt, so weist das Frank-Modell auch aus theoretischer Sicht erhebliche Probleme auf. Auf diesen Umstand hat bereits 1989 Harrington hingewiesen. Wie erläutert benötigt Frank die Annahme eines hinreichend fälschungssicheren Signals. „Hinreichend“, so wurde in Kapitel 3.3.1 gezeigt, bedeutet dabei genauer, daß die Signale über die beiden Typen (Defektor, Kooperator) sich **im Durchschnitt** unterscheiden. Technisch gesprochen mußten also die Mittelwerte der beiden Dichtefunktionen für beide Typen auseinanderfallen. Zusätzlich nimmt Frank an, daß beide Dichtefunktionen sich auch hinsichtlich ihres Trägers unterscheiden, ohne jedoch eine inhaltliche Begründung zu geben. Diese Annahme muß als eindeutig ad hoc qualifiziert werden. Problematisch hierbei ist, daß sie zentral für die Modellergebnisse ist, wie Harrington (1989) nachgewiesen hat. Nimmt man identische Träger der Dichten an, so fallen Franks Resultate weitgehend in sich zusammen.

Vor dem Hintergrund der geschilderten Probleme signaltheoretischer Ansätze schlägt die vorliegende Arbeit einen anderen Weg ein. Anknüpfend an Amann und Yang (1994) führen wir in das Standard-Zwei-Personen-GD einen sogenannten „knowledgeable type“ ein, der Akteure durch Interaktion lernt. Im Unterscheid zu Amann/Yang nehmen wir jedoch nicht an, daß der mutante Typ perfekte Information über den Typ seines Gegenübers erhält. Letztlich, so ist an dieser Stelle anzumerken, handelt es sich nämlich auch bei dem Modell von Amann/Yang um eine signaltheoretisches. Zwar wird durch die Bezeichnung „knowledgeable Dove“ eine lerntheoretische Assoziation geweckt. Faktisch jedoch erhält die Mutante vor jeder Interaktion die Information über den Typ ihres Gegenspielers, und wie anders als durch ein Signal sollte dies bewerkstelligt werden?

Im Gegensatz hierzu ist der K-Typ der vorliegenden Arbeit eine Mutante, die ihr Wissen erst durch erfolgte Interaktion mit anderen Mitgliedern der Population erhält. Mit anderen Worten, unser K-Typ wird erst aus Erfahrung klug. Dies impliziert insbesondere, daß zunächst Interaktionen mit Defektoren erfolgen müssen, bevor diese für den K-Typ bekannte Akteure sind. Wenn Defektoren gelernt wurden, so konditioniert der K-Typ im weiteren Spielverlauf sein Verhalten auf die erworbene Information. Im einzelnen wird angenommen, daß die Interaktion mit bekannten Defektoren verweigert und eine outside-option gewählt wird. In Kapitel 4 dieser Arbeit wird ein mathematisches Lernmodell für random-matching-

games in Populationen endlicher Größe entwickelt. Das Modell erlaubt die Berechnung der Wahrscheinlichkeit, in einem beliebigen matching-Zeitpunkt einen der bereits bekannten Defektoren zu treffen.

Das zentrale Ergebnis des Basismodells ist die Existenz eines global stabilen, vollständig gemischten Nash-Gleichgewichts, in dem alle drei Typen mit strikt positiven Populationsanteilen vertreten sind. Dieses Resultat mag als Erklärung für das stilisierte Faktum der experimentellen Evidenz aus Kapitel 2 gelten. Hierzu ist folgendes anzumerken: Die experimentelle Literatur zeigt recht deutlich die Koexistenz dreier Typen, eines mehr oder weniger strikten kooperativen, eines strikt defektiven Typen und eines dritten Typs, der zwischen Defektion und Kooperation mischt. Die vorliegende Arbeit behauptet nicht, daß dieser mischende Typ durch den K-Typ unseres Basismodells perfekt abgebildet wird. Bereits die offenkundige Diversität des individuellen Verhaltens innerhalb dieser Kategorie spricht gegen eine einfache Repräsentierbarkeit. Insbesondere wird mit dem hier vorgestellten Lernmodell auch kein deskriptiver Anspruch für real stattfindende Lernprozesse erhoben. Dies verbietet sich schon deshalb, weil sich Experimental- und Modellsituation in einem wichtigen Punkt unterscheiden. Das Lernmodell nimmt an, daß Mitspieler persönlich identifizierbar sind. Die typische Versuchsanordnung in Dilemmaexperimenten sieht genau dies jedoch nicht vor: Die Identität der Mitspieler wird üblicherweise nicht preisgegeben. Wohl scheint auch in Experimentalsituationen Erfahrungslernen eine gewisse Rolle zu spielen, jedoch nimmt dieses Lernen aufgrund der unterschiedlichen Bedingungen eine andere Form an. In wiederholten Spielen wird, wie im Lernmodell auch, Information über die Zusammensetzung der Population generiert, allerdings ohne daß das Verhalten von Vpn gegenüber einzelnen Mitspielern konditioniert werden könnte. Im Spielverlauf werden lediglich Häufigkeiten von Verhaltensweisen sukzessive bekannt, die dann als Schätzer benutzt werden können, um eventuell vorhandene a priori beliefs anzupassen. Modelle adaptiver Erwartungsanpassung, wie etwa die Lernrichtungstheorie (Selten und Stoecker (1986)) oder naives bayesianisches updating (Eichberger, Haller und Milne (1993)) scheinen daher grundsätzlich geeigneter zu sein, um diese Lernprozesse zu beschreiben.

Jedoch war die akkurate Beschreibung derartiger Lernprozesse auch nicht das Ziel dieser Arbeit. Erklärte Absicht war es vielmehr, die Effekte von Erfahrungslernen in random-matching games zu erfassen. Das hier entwickelte Lernmodell leistet dies in einfacher und intuitiver Weise. Die Simplizität des Modells kann dabei durchaus als Stärke aufgefaßt werden. Obwohl die modellierte Form des Lernens so ausgesprochen einfach strukturiert ist, reicht sie bereits, um das Überleben von Kooperation zu sichern. Hervorhebenswert ist die Tatsache, daß sich dabei nicht nur der bedingt kooperative K-Typ evolutionär behaupten kann, sondern unter

geeigneten Umständen auch der Typ des strikten Kooperators. Die einfachen Tauben profitieren von der Existenz des K-Typs indirekt durch die implizite Bestrafung der Defektoren bei Verweigerung der Interaktion. Als zentrale Determinante des evolutionären Erfolges der kooperativen Strategien erwies sich dabei der Lernerfolg der K-Typen, was intuitiv absolut plausibel ist. Für eine gegebene Populationsgröße muß die Anzahl der matchings hinreichend groß sein, damit ausreichend Wissen erworben werden kann, bzw. für eine gegebene Interaktionshäufigkeit darf die Population nicht zu groß werden. Dies bestätigt die übliche Intuition bezüglich des Zusammenhangs zwischen Gruppengröße und Erfolgchancen kooperativen Verhaltens. Allerdings gilt es hier zu beachten, daß ausschlaggebend die Größe der relevanten Bezugsgruppe ist und nicht zwingenderweise die Größe der Gesamtpopulation. So ist also nicht automatisch wegen der Größe moderner Industriegesellschaften davon auszugehen, daß die Chancen für Kooperation gegenüber vormodernen Gesellschaftsformen geringer sind. Menschen interagieren nicht mit der Gesamtpopulation, sondern mit kleinen Ausschnitten daraus. Das Basismodell versucht, diesem Umstand dadurch Rechnung zu tragen, daß eine infinite Anzahl finiter Subgruppen angenommen wird. Hinreichende Konstanz der sozialen Beziehungen vorausgesetzt, kann auch hier Kooperation gedeihen. Negativ formuliert: Die Zunahme von Mobilität reduziert die Konstanz von Gruppenzugehörigkeiten und vermindert die Chancen für kooperatives Verhalten. Im übrigen sei angemerkt, daß unser Modell damit genau jene Determinanten der Chancen von Kooperation identifiziert, die intuitiv bereits von Trivers (1973) in seiner bahnbrechenden Arbeit über reziproken Altruismus genannt wurden: Interaktionshäufigkeit und Konstanz der Interaktionsbeziehung.

Breiten Raum in dieser Arbeit nahm die Diskussion konzeptioneller Probleme in Anspruch. Standardmodelle der evolutionären Spieltheorie operieren auf der Basis der Fiktion einer unendlich großen Population. Diese Annahme muß naturgemäß mit dem Bestreben konfliktieren, meßbare Effekte von Lernen zu generieren. Die vorliegende Arbeit bietet zwei Auswege an. Zum einen wird im Basismodell angenommen, daß die Gesamtpopulation in eine infinite Anzahl finiter Subpopulationen zerfällt. Dieser modellierungstechnische Kunstgriff sichert nicht nur nicht-triviale Resultate des Lernmodells bei gleichzeitiger Anwendbarkeit der Replikator-Dynamik, sondern er kann auch begriffen werden als ein Schritt in die Richtung größerer Realitätsnähe (vgl. Berninghaus und Schwalbe (1993)). Zum anderen wird mit der Technik des simultanen Grenzübergangs (vgl. Kapitel 5.5) ein formal eleganter Ausweg aus dem oben skizzierten Dilemma gewiesen, der sich aber seinerseits gewisser Kritik ausgesetzt sieht. Wir hatten argumentiert, daß ursächlich für die Probleme der Integration von Lernen und evolutionärer Analyse die Basisfiktion der evolutionären Spieltheorie einer formal unendlich großen Po

pulation ist. Andererseits ist es evident, daß Lernen in realen und experimentellen Situationen eine Rolle spielt. Um diesem Faktum Rechnung tragen zu können, scheint es hinnehmbar, die inakzeptablen Konsequenzen einer realitätsfernen Annahme (einer unendlich großen Population) durch eine weitere realitätsferne Annahme (eines Kontinuums von Interaktionen auf dem endlichen Lebenszeitintervall eines Akteurs) zu korrigieren.

Die Ergebnisse erweisen sich in zweifacher Hinsicht als robust. Zum einen gegenüber der Einführung einer weiteren lernfähigen, diesmal prinzipiell defektiv orientierten Mutante. Zum anderen gegenüber dem Wechsel vom deterministischen zum stochastischen Modellrahmen. Die Beschränkung von Lernfähigkeit auf die kooperative Verhaltensdisposition muß willkürlich erscheinen. Naheliegender ist die Forderung, auch eine lernfähige defektive Mutante zuzulassen. In Kapitel 6 dieser Arbeit wurde gezeigt, daß die wesentlichen Modellresultate diese Erweiterung schadlos überstehen. Insbesondere bleibt das innere 3-Typen-Gleichgewicht aus dem Basismodell als lokal stabil erhalten. Gewisse Abstriche sind aber bei der globalen Stabilität zu machen. Diese folgt im Vier-Typen-Modell nicht mehr zwingend aus der lokalen Stabilität. Vielmehr ist nunmehr auch die Koexistenz lokal stabiler Fixpunkte denkbar mit der Folge, daß die langfristigen Konvergenzeigenschaften der Replikatorodynamik im allgemeinen von den Startbedingungen abhängen. Gleichwohl ist auch weiterhin denkbar, daß die Parameterbereiche für lokale Stabilität überlappungsfrei bleiben. In diesem Fall bleibt die globale Stabilität der betrachteten Gleichgewichte erhalten. Unsere Resultate unterscheiden sich damit zum einen von denen Amanns und Yangs (1994), die den Randfall perfekter Information untersucht haben. Bei ihnen bleibt der Zusammenhang von lokaler und globaler Stabilität immer erhalten. Zum anderen vermeiden wir die Probleme des Robson-Modells (1990), in dem eine intelligente, defektive Mutante die ursprüngliche, lernfähige kooperative Mutante dominiert. Bei Robson waren langfristig nur rein defektive Populationen zu erwarten.

Noch wichtiger scheint aber die Tatsache, daß auch der Wechsel des Modellrahmens das zentrale Ergebnis dieser Arbeit nicht zerstören kann. Eine stochastische Modellierung erscheint gleich aus mehreren Gründen überzeugend. Zum einen erlaubt sie es, sich von der hinreichend strapazierten Annahme einer unendlich großen Population zu verabschieden. Zum anderen ist der deterministische Standardmodellrahmen verschiedentlich kritisiert worden, da er das für reale Evolutionsprozesse vermutlich wichtige Element der Mutation (im vorliegenden Kontext geeignet zu interpretieren als Fehler oder bewußtes Experimentieren) nicht adäquat einfängt. „Mutation“ taucht dort nur sehr restringiert in Form singulärer Auslenkungen aus Gleichgewichtszuständen, also in der klassischen Stabilitätsanalyse auf. Foster und Young (1990) hatten aber gezeigt, daß kontinuierliche

Schocks das Konvergenzverhalten dynamischer Systeme grundlegend ändern können. Kapitel 7 knüpft an die bahnbrechenden Arbeiten von Kandori et al (1993) und Vega-Redondo (1995, 1997) an und wendet die dort entwickelten Techniken auf unser 3x3 GD an. Als zentral für die Ergebnisse erweist sich die Einführung einer positiven Mutationswahrscheinlichkeit für jeden Spieler. Dies generiert eine aperiodische und irreduzible Markov-Kette auf einem endlichen Zustandsraum, mit der Folge, daß eine eindeutige und strikt positive stationäre Verteilung für diesen Prozeß existiert. Läßt man die Mutationswahrscheinlichkeit arbiträr klein werden, so bleiben lediglich drei Zustände in der so ermittelten Grenzverteilung mit positiven Wahrscheinlichkeiten übrig, nämlich genau die drei monomorphen Zustände ALLH, ALLK und ALLD. Als zentral erweist sich wiederum eine Bedingung bezüglich des Lernerfolges des K-Typs, bzw. genauer gesagt bezüglich der Möglichkeit eines hinreichend großen Lernerfolges. Erforderlich ist nämlich, daß es zu einer hinreichend großen Anzahl an matchings innerhalb einer Generation kommt. Beide Modellkomplexe, sowohl der deterministische, als auch der stochastische, liefern damit strukturell identische Resultate unter strukturell identischen Bedingungen.

Insbesondere erweist sich die Frage als kritisch, ob eine singuläre K-Mutante eine Hawk-Population invadieren kann. Wenn dies der Fall ist, dann gilt, unabhängig von allen weiteren Faktoren, daß beide kooperativen Strategien überleben. Es läßt sich zwar ohne weitere Modellspezifikationen nicht sagen mit welchen Wahrscheinlichkeiten, aber es ist klar, daß dann beide Zustände mit positiver Wahrscheinlichkeit in der Grenzverteilung vertreten sind. Im stochastischen Modell zeigt sich damit auch klar, welche wichtige, ja unerläßliche Funktion der K-Typ für die Entstehung von Kooperation hat: Naive Kooperateure sind niemals in der Lage, eine Population aus Defektoren zu invadieren. Hierzu sind immer mindestens zwei Mutanten erforderlich. Umgekehrt kann aber ein Defektor immer eine Population aus naiven Kooperatoren invadieren. Ersteres ist also unwahrscheinlicher als letzteres, und dieser Umstand wird genau dadurch reflektiert, daß der Zustand ALLD in der Grenzverteilung Wahrscheinlichkeit Null aufweist. In einer Welt ohne intelligente, reziproke Kooperateure hat bedingungslos kooperatives Verhalten keinerlei Überlebenschance.

Das stochastische Modell weist zweifellos den Vorzug größeren Realismus bezüglich der Populationsgröße auf. Kritisch ist jedoch folgendes anzumerken: Das zentrale Resultat des Modells muß als einigermaßen artifiziell eingestuft werden. Der Prozeß hält sich, abgesehen von einigen transienten Zuständen, die aber zeitlich nicht ins Gewicht fallen, lediglich in den drei monomorphen Populationszuständen auf. Dies mag als analoges Resultat zum inneren Gleichgewicht des deterministischen Modells gedeutet werden. Indes, eine perfekte Analogie stellt es

nicht dar. Die Tatsache, daß ein echtes inneres Gleichgewicht im stochastischen Modell nicht existiert, ist auf folgenden Umstand zurückzuführen: Ein Populationszustand stellt genau dann ein Gleichgewicht dar, wenn die Gleichheit aller Payoffs gewährleistet ist. In einer Population endlicher Größe würde dies für einen gemischten Populationszustand jedoch typischerweise bedeuten, daß die gleichgewichtigen Anzahlen der beteiligten Typen nicht ganze Zahlen sind (vgl. auch Vega-Redondo (1996), S.145). Offenkundig hat dieses Ereignis in einer strikt endlichen Population die Wahrscheinlichkeit Null. Wollte man diesem Problem entgehen, so müßte man die Populationsgröße wiederum ins Unendliche streben lassen, mit den daraus resultierenden, bekannten Problemen. Es scheint also, als hätte jede Modellierungsvariante ihren Preis: Der größere Realismus des stochastischen Modells bezüglich der Populationsgröße wird mit einem artifiziellen Ergebnis bezahlt. Das vollständig gemischte Gleichgewicht im deterministischen Modell seinerseits wird erkaufte durch die Fiktion einer unendlich großen Population – eine Annahme, die erst durch geschickte Modellierung mit dem Lernmodell vereinbar wird.

## 9. Ausblick

Eine Reihe von Arbeiten sind in den letzten Jahren vorgestellt worden, die bemüht sind, daß beobachtbare Verhalten in einigen wichtigen ökonomischen Situationen zu erklären. Hierzu zählen das Comparative model von Bolton (1991), das Konzept des Fairneß-Gleichgewichts von Rabin (1993) sowie die ERC-Theorie (E=Equity, R=Reciprocity, C=Competition) von Bolton und Ockenfels (forthcoming). Der Hintergrund für Boltons Modell bilden „alternating-offer“-Spiele, wie etwa das Ultimatumspiel. Um das dort regelmäßig beobachtbare ablehnende Verhalten der Responder gegenüber positiven, aber offensichtlich als zu klein empfundenen Angeboten erklären zu können, führt Bolton eine Präferenz für „fair behandelt werden“ in ein neoklassisches Nutzenmaximierungskalkül ein. In Rabins Modell bewerten die Spieler die erwarteten Intentionen ihrer Mitspieler und verhalten sich demgemäß reziprok. Bolton und Ockenfels schließlich argumentieren, daß Spieler nicht nur ihren absoluten Payoff maximieren, sondern auch ihre relative Einkommensposition berücksichtigen. Allen drei Modellen ist gemeinsam, daß sie über eine gegenüber dem neoklassischen Standardmodell „reichere“ Motivationsstruktur verfügen. Menschen sind nicht ausschließlich „selfish“, sondern durch weitere Faktoren, die möglicherweise mit Fairneßabwägungen zu tun haben, motiviert. Unsere Arbeit liefert einen gewissen Anhaltspunkt dafür, daß derartige „abweichende“ Präferenzen evolutionär überlebensfähig sind. Schließlich ist auch das Verhalten unseres K-Typs durch Ausbeutungsaversion motiviert.

Es wäre eine reizvolle Aufgabe, im Rahmen eines evolutionären Ansatzes die Frage zu untersuchen, welche Motivationsstruktur genau sich evolutionär behaupten würde, um so klären zu können, welches der Modelle die menschliche Präferenzstruktur (vermutlich) zutreffend beschreibt.

**Literatur**

**Andreoni, J. and J. H. Miller**, 1993, Rational cooperation in the finitely repeated Prisoners' Dilemma: Experimental evidence, *Economic Journal* 103, 570-85

**Amann, E. and C.-L. Yang**, 1994, The robustness of a sophisticated mutant in an evolutionary model of cooperation, discussion paper No. 94-03, University of Dortmund

**Amann Erwin und Chun-Lei Yang**, 1998, Sophistication and the persistence of cooperation, *Journal Of Economic Behavior And Organization* 37 , 91-105

**Aumann, R. und L. Shapley**, 1976, Long Term Competition: A Game Theoretic Analysis, mimeo, Hebrew University

**Axelrod, R.**, 1981, The emergence of cooperation among egoists, *American Political Science Review* 75, 306-18

**Axelrod, R.**, 1984, *The evolution of cooperation* (Basic Books, New York)

**Berninghaus, S. und U.Schwalbe**, 1996, Evolution, Interaction and Nash Equilibria, *Journal of Economic Behavior and Organization* 29, 57-85

**Bolton, G.E.**, 1991, A Comparative Model of Bargaining: Theory and Evidence, *American Economic Review* 81, 1096-1136

**Bolton, G.E und A. Ockenfels**, ERC – A Theory of Equity, Reciprocity and Competition. *American Economic Review*, forthcoming

**Bomze, I.M.**, 1986, Non-cooperative, two-person games in biology: A classification, *International Journal of Game Theory* 15, 31-57

**Börgers, T.**, 1996, On the relevance of learning and evolution to economic theory, *Economic Journal* 106, 1374-85

**Börgers, T. und R. Sarin**, 1997, Learning Through Reinforcement and Replicator Dynamics, *Journal of Economic Theory* 77, 1-14



- Brosig, J.**, 2000, Is Cooperativeness identifiable? Some experimental results, University of Magdeburg, discussion paper
- Camerer, C. and K. Weigelt**, 1988, Experimental tests of the sequential equilibrium reputation model, *Econometrica* 56, 1-36
- Conlisk, J.**, 1980, Costly optimizers versus cheap imitators, *Journal of Economic Behavior and Organization* 1, 275-293
- Cooper, R., D.V. de Jong, R. Forsythe and T.W. Ross**, 1996, Cooperation without reputation: Experimental evidence from prisoner's dilemma games, *Games and Economic Behavior* 12, 187-218
- Dawes, R.**, 1980, Social Dilemmas, *Annual Review of Psychology* 31, 169-193
- Dawes, R. und R. Thaler**, 1988, Anomalies: Cooperation, *Journal of Economic Perspectives* 2, 187-197
- Eichberger, J., Haller, H. und F. Milne**, Naive Bayesian Learning in 2x2 Matrix Games, *Journal of Economic Behavior and Organization* 22, 69-90
- Ellison, G.**, 1994, Cooperation in the prisoners' dilemma with anonymous random matching, *Review of Economic Studies* 61, 567-588
- Fahrmeier, L., Kaufmann, H. L. und F. Ost**, 1981, *Stochastische Prozesse. Eine Einführung in Theorie und Anwendungen*, Carl Hanser Verlag, Wien
- Foster, D. und P. Young**, 1990, Stochastic Evolutionary Game Dynamics, *Theoretical Population Biology* 38, 219-232
- Frank, R.H.**, 1987, If homo economicus could choose his own utility function, would he want one with a conscience?, *American Economic Review* 77, 593-604
- Frank, R.H.**, 1988, *Passions within reason. The strategic role of the emotions* (Norton, New York)
- Frank, R.H., Gilovich, T. und D.T. Regan**, 1993, The Evolution of One-Shot Cooperation: An Experiment, *Ethology and Sociobiology* 14, 247-256

- Freidlin, M.I. und A.D. Wentzel**, 1984, Random Perturbations of Dynamical Systems, Springer, New-York
- Friedman, D.**, 1991, Evolutionary games in economics, *Econometrica* 59, 637-666
- Friedman, J.**, 1971, A Noncooperative Equilibrium for Supergames, *Review of Economic Studies* 38, 1-12
- Fudenberg, D. and E. Maskin**, 1986, The folk theorem in repeated games with discounting or with incomplete information, *Econometrica* 54, 533-554
- Fudenberg, D. und J. Tirole**, 1989, Noncooperative Game Theory for Industrial Organization. An Introduction and Overview, in: Schmalensee, R. und R.D. Willig (Hg.), *Handbook of Industrial Organization I*, 259-327
- Fudenberg, D. und J. Tirole**, 1991, *Game Theory*, Cambridge, Mass.
- Garcia, J. und R. Koelling**, 1966, Relation of Cue to Consequence in Avoidance Learning, *Psychonomic Science* 4, 123-124
- Gilovich, T., Medvec, V.H. und K. Savitsky**, 1998, The Illusion of Transparency: Biased Assessments of Other's Ability to Read One's Emotional States, *Journal of Personality and Social Psychology* 75, 332-346
- Güth, W.**, 1992, Spieltheorie und Industrieökonomik. Muß Liebe weh tun?, *Ifo-Studien*, 271-316
- Güth, W.**, 1997, Sequential versus independent commitment – An indirect evolutionary analysis of bargaining rules, Humboldt University Discussion Paper – Economics Series 106
- Güth, W. and H. Kliemt**, 1995, Evolutionarily Stable Co-operative Commitments, Humboldt University Discussion Paper – Economics Series 53
- Güth, W. and H. Kliemt**, 1996, Towards a completely indirect evolutionary approach – A note, Humboldt University Discussion Paper – Economics Series 82

**Güth, W. and H. Kliemt**, 1998, The Indirect Evolutionary Approach: Bridging the gap between Rationality and Adaptation, *Rationality and Society* 10, 377-399

**Güth, W., H. Kliemt and B. Peleg**, 1998, Co-evolution of preferences and information in simple games of trust, Humboldt University Discussion Paper – Economics Series 123

**Güth, W., R. Schmittberger and R. Schwarze**, 1982, An experimental study of ultimatum bargaining, *Journal of Economic Behavior and Organization* 3, 367-88

**Güth, W. and M. Yaari**, 1992, An Evolutionary Approach to Explaining Reciprocal Behavior in a Simple Strategic Game, in: Witt, U., ed., *Explaining Process and Change*, Ann Arbor: The University of Michigan Press

**Guttman, J.**, 1996, Rational actors, tit-for-tat types, and the evolution of cooperation, *Journal of Economic Behavior and Organization* 29, 27-56

**Hamilton, William D.**, 1964, The Genetical Evolution of Social Behavior, *Journal of theoretical biology* 7, 1-52

**Harrington, J.E.**, 1989, If Homo Economicus Could Choose his own Utility Function, Would He Want One with a Conscience? Comments, *American Economic Review* 79, 588-593

**Hey, J.D.**, 1991, *Experiments in Economics*, Oxford, Cambridge

**Hofbauer, J. and K. Sigmund**, 1988, *The theory of evolution and dynamical systems* (Cambridge University Press)

**Isaac, R.M. und J.M. Walker**, 1988, Group Size Effects in Public Good Provision: The Voluntary Contributions Mechanism, *Quarterly Journal of Economics*, 179-199

**Isaac, R.M., J.M. Walker und S.H. Thomas**, 1984, Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations, *Public Choice* 43, 113-149

- Isaac, R.M., J.M. Walker and A. Williams**, 1994, Group size and the voluntary provision of public goods: Experimental evidence utilizing very large groups, *Journal of Public Economics* 54, 1-36
- Kandori, M.**, 1992, Social norms and community enforcement, *Review of Economic Studies* 59, 63-80
- Kandori, M., Mailath, G. and R. Rob**, 1993, Learning, Mutation and Long-run Equilibria in Games, *Econometrica* 61, 29-56
- Karlin, S. und H.M. Taylor**, 1975, *A First Course in Stochastic Processes*, London, Academic Press
- Kreps, D., P. Milgrom, J. Roberts and R. Wilson**, 1982, Rational cooperation in the finitely repeated Prisoners' dilemma, *Journal of Economic Theory* 27, 245-52
- Ledyard, J.O.**, 1995, Public goods: A survey of experimental research, in: Kagel, J.H. and A.E. Roth, eds., *Handbook of experimental economics* (Princeton University Press)
- Marwell, G. und R.E. Ames**, 1981, Economists Free Ride, Does Anyone Else?, *Journal of Public Economics* 15, 295-310
- Milgrom, P., D. North and B. Weingast**, 1990, The role of institutions in the revival of trade: The law merchant, private judges and champagne fairs, *Economics and Politics* 2, 1-23
- Ochs, J. and A. Roth**, 1989, An experimental study of sequential bargaining, *American Economic Review* 79, 355-84
- Ockenfels, A. und R. Selten**, 1999, An Experiment on the Hypothesis of Involuntary Truth-Signalling in Bargaining, preprint, University of Magdeburg

- Ockenfels, A. und J. Weimann**, 1999, Types and Patterns: An Experimental East-West Comparison of Cooperation and Solidarity, *Journal of Public Economics* 71, 275-287
- Offerman, T., J. Sonnemans and A. Schram**, 1996, Value orientations, expectations and voluntary contributions in public goods, *Economic Journal* 106, 817-45
- Palfrey, T.R. and H. Rosenthal**, 1988, Private incentives in social dilemmas: The effects of incomplete information and altruism, *Journal of Public Economics* 35, 309-32
- Pflüger, M.**, 1997, Folk-Theoreme, *WiSt* 8, 412-416
- Rabin, M.**, 1993, Incorporating Fairness into Game Theory and Economics, *American Economic Review* 83, 1281-1302
- Rapoport, A. and A.M. Chammah**, 1965, Prisoners' Dilemma. (University of Michigan Press, Ann Arbor)
- Robson, A.**, 1990, Efficiency in evolutionary games: Darwin, Nash and the secret handshake, *Journal of Theoretical Biology* 144, 376-396
- Roth, A.E.**, 1988, Laboratory experimentation in economics: A methodological overview, *Economic Journal* 98, 974-1031
- Roth, A.E.**, 1995, Bargaining experiments, in: Kagel, J.H. and A.E. Roth, eds., *Handbook of experimental economics* (Princeton University Press)
- Rubinstein, A.**, 1979, Equilibrium in supergames with the overtaking criterion, *Journal of Economic Theory* 21, 1-9
- Samuelson, L.**, 1987, A Note on Uncertainty and Cooperation in the Finitely Repeated Prisoner's Dilemma, *International Journal of Game Theory* 16, 187-195
- Schelling, T.**, 1960, *The strategy of conflict*, (Harvard University Press, Cambridge)

**Schlag, K.H.**, 1998, Why Imitate, And If So, How?, *Journal of Economic Theory* 78, 130-156

**Schneider, F. und W.W. Pommerehne**, 1981, Free Riding and Collective Action: An Experiment in Public Microeconomics, *Quarterly Journal of Economics* 96, 689-704

**Selten, R.**, 1991, Evolution, learning and economic behavior, *Games and Economic Behavior* 3, 3-24

**Selten, R. und R. Stoecker**, 1986, End Behavior in Sequences of finite Prisoner's Dilemma Supergames, *Journal of Economic Behavior and Organization* 7, 47-70

**Sherratt, T. und Gilbert Roberts**, 1998, The Evolution of Generosity and Choosiness in Cooperative Exchanges, *Journal of Theoretical Biology* 193, 167-177

**Skinner, B.F.**, 1938, *The Behavior of Organisms*, New York, Appleton-Century-Crofts

**Tattersall, I.**, 2000, Wir waren nicht die Einzigen. Warum von allen Menschenarten nur der Homo sapiens überlebte, in: *Spektrum der Wissenschaft* 3, S. 46-53

**Townsend, R.**, 1980, Models of Money with Spatially Separated Agents, in: Kareken, J. und N. Wallace (eds.), *Monetary Economies*, Minneapolis, Federal Reserve Bank of Minneapolis

**Trivers, Robert L.**, 1971, The Evolution of Reciprocal Altruism, *Quarterly Review of Biology* 46, 35-57

**Vega-Redondo, F.**, 1996, *Evolution, Games and Economic Behavior*, Oxford University Press

- Vega-Redondo, F., 1997**, The Evolution of Walrasian Behavior, *Econometrica* 65, 375-384
- Vogt, C., 1998**, Endogenous learning mutants in an evolutionary prisoner's dilemma, mimeo
- Weibull, J., 1995**, *Evolutionary Game Theory* (MIT Press, Cambridge)
- Weimann, J., 1994**, Individual behavior in a free riding experiment, *Journal of Public Economics*, 54, 185-200
- Weimann, J., 1995**, Freifahrer im Test: Ein Überblick über 20 Jahre Freifahrerexperimente, in: *Ökonomie und Gesellschaft, Jahrbuch 12: Soziale Kooperation*, Campus, Frankfurt/Main, New York, 168-241
- Yang, C.-L., 1994**, Evolution of cooperation in prisoners' dilemma game with a knowledgeable mutant, discussion paper No. 93-09, University of Dortmund
- Zuckerman, M., de Paulo, B.M. und R. Rosenthal, 1981**, Verbal and Nonverbal Communication of Deception, *Advances in Experimental Social Psychology* 14, 1-59