Über eine alternative Totalauszahlung für stochastische Spiele

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Steffen Wekeck
aus Kassel

Göttingen 1999

D 7
Referent: Prof. Dr. Krengel
Korreferent: PD Dr. Gnedin
Tag der mündlichen Prüfung: 1. November 1999

Inhaltsverzeichnis

Einleitung

Gegenstand dieser Arbeit sind stochastische Zwei-Personen-Nullsummenspiele. Das zugrunde liegende Modell geht auf Shapley [Sha53] zurück. Man geht dabei von der folgenden Situation aus: Gegeben sind endlich viele Zustände, in denen beide Spieler über jeweils endlich viele Aktionen (Züge) verfügen. Ausgehend von einem Startzustand bestimmt die Wahl eines solchen Aktionspaares in jeder Spielrunde eine Auszahlung (von Spieler 2 an Spieler 1) und eine Übergangswahrscheinlichkeit, gemäß der sich das Spiel in den nächsten Zustand bewegt.
Auf diese Weise entsteht eine Folge von Auszahlungen, der man in geeigneter Weise eine Bewertung zuordnen möchte. Im Shapleyschen Modell wurden die Auszahlungen diskontiert, also auf den Spielbeginn abgezinst. Dies ist neben gewissen praktischen Anwendungen vor allem in der so erzwungenen Konvergenz der Summe der Auszahlungen begründet. Im Mittelpunkt der weiterführenden Untersuchungen stochastischer Spiele stand das erstmals von Gilette [Gil57] betrachtete Durchschnittsgewinnkriterium, welches in etwa den mittleren Gewinn (von Spieler 1) pro Spielrunde widerspiegelt. Wegen der Beschränktheit aller darin auftretenden Zufallsgrößen bereitet auch dieses Bewertungskriterium keine technischen Probleme.
Als Verfeinerung der Durchschnittsauszahlung führten Thuijsman und Vrieze in [TV87] das Kriterium der Totalauszahlung ein. Es eignet sich für die Untersuchung von Spielen, deren Wert bezüglich der Durchschnittsauszahlung Null beträgt. In dieser Situation betrachtet man mittels der Totalauszahlung grob gesprochen den minimalen mittleren Besitz von Spieler 1. Wegen der möglichen Unbeschränktheit dieser Größe ist dieses Kriterium bereits nicht mehr so leicht zu handhaben.

In dieser Arbeit soll es um eine Modifizierung dieser Totalauszahlung gehen, die durch gewisse praktische Beobachtungen motiviert ist: Die klassische Definition berücksichtigt nur Erwartungswerte des in *endlich* vielen Spielrunden akkumulierten Kapitals. Dadurch kann es passieren, dass die Möglichkeiten der Spieler, durch geschickte Strategien einen für sie günstigen Spielausgang zu erzwingen, nicht ausreichend beachtet werden. Durch Vertauschung von Limes- und Erwartungswertbildung wird diesem Phänomen Rechnung getragen. Der Preis dafür – und dies ist wohl auch der Hauptgrund für die ursprüngliche Definition – sind allerdings einige technische Schwierigkeiten, die die Handhabung des alternativen Kriteriums erschweren.

Wie in der Spieltheorie üblich werden in dieser Arbeit Konzepte wie Spielwert und Optimalität von Strategien für das neue Bewertungskriterium untersucht. Ferner werden hinreichende Bedingungen genannt, unter denen beide Versionen der Totalauszahlung übereinstimmen. Einen Schwerpunkt bildet außerdem die Analyse stationärer Strategien. Im Einzelnen ist die Arbeit wie folgt gegliedert:

Kapitel 1 führt in die Problemstellung ein und stellt die verschiedenen Bewertungskriterien vor. Ferner wird die Einführung der alternativen Totalauszahlung durch ein Beispiel motiviert. Im letzten Abschnitt wird näher auf die besonderen technischen Schwierigkeiten bei der Definition dieses Kriteriums eingegangen und eine mögliche Lösung aufgezeigt. Kapitel 2 ist der Beziehung zwischen klassischer und alternativer Definition der Totalauszahlung gewidmet.
In Kapitel 3 geht es um die Existenz des alternativen totalen Spielwerts. Anhand eines Beispiels wird gezeigt, dass die bei der klassischen Totalauszahlung für die Existenz des Spielwerts hinreichende Bedingung aus [TV96, FV97] hier nicht hinreichend ist. Ferner weisen wir nach, dass der alternative totale Spielwert umgekehrt auch dann existieren kann, wenn der gewöhnliche totale Spielwert nicht existiert. Kapitel 4 behandelt die Frage, ob optimale oder „einfache“ ε-optimale Strategien stets existieren.
Schließlich untersucht Kapitel 5 stationäre Strategien unter dem Kriterium der alternativen Totalauszahlung. Um eine allgemeine Formel herleiten zu können, beschäftigen wir uns zunächst mit dem asymptotischen Verhalten des Cesaro-Mittels von Funktionalen einer Markowkette. Die hierbei gewonnenen Resultate können somit auch außerhalb des spieltheoretischen Kontexts von Interesse sein. Ihre Übertragung auf stochastische Spiele beschließt dieses Kapitel.

Abschließend möchte ich mich bei Herrn Prof. Dr. Krengel für die freundliche Unterstützung bei dieser Arbeit bedanken. Herrn PD Dr. Gnedin danke ich für die Übernahme des Korreferats.

Kapitel 1

Problemstellung

1.1 Das Modell stochastischer Spiele

Wir betrachten stochastische Zwei-Personen-Nullsummenspiele mit endlichen Zustands- und Aktionsräumen. Ein solches Spiel ist gegeben durch:

(a) einen Zustandsraum $S = \{1, \dots, z\}$,

(b) einen Aktionsraum $A_s^k = \{1, \dots, n_s^k\}$ für jeden Zustand $s \in S$ und jeden Spieler $k = 1, 2$,

(c) eine Auszahlung $r(s, i, j)$ für jeden Zustand $s \in S$ und jedes Aktionspaar $(i, j) \in A_s := A_s^1 \times A_s^2$,

(d) Übergangswahrscheinlichkeiten $p(s, i, j) = (p(1|s, i, j), \dots, p(z|s, i, j))$ für jeden Zustand $s \in S$ und jedes Aktionspaar $(i, j) \in A_s$.

Mit $S_{abs} := \{s \in S \colon p(s|s, i, j) = 1,\, r(s, i, j) = 0\ \forall (i, j) \in A_s\}$ bezeichnen wir die Menge der absorbierenden Zustände eines Spiels. Ohne Einschränkung kann man $|S_{abs}| \leq 1$ voraussetzen.
In Runde 1 beginnt das Spiel in einem festen Zustand aus S. Befindet sich das Spiel zu einem Zeitpunkt n in Zustand s, wählen beide Spieler unabhängig voneinander Aktionen $(i, j) \in A_s$. Dann erhält Spieler 1 (von Spieler 2) die Auszahlung $r(s, i, j)$ und gemäß der Wahrscheinlichkeitsverteilung $p(s, i, j)$ wird der Zustand für Runde $n + 1$ bestimmt. Die Spieler dürfen gemischte Aktionen benutzen, d.h. über ihre Aktionen randomisieren.
Zur expliziten Angabe stochastischer Spiele verwenden wir die folgende Darstellung: Jeder Zustand $s \in S$ wird durch eine Matrix repräsentiert, deren Eintrag $(i, j) \in A_s$ die in Abbildung 1.1 links stehende Gestalt hat. Die

rechte Notation wird verwendet, falls $p(s,i,j)$ in einem Zustand $t \in S$ konzentriert ist. Im Fall $t \in S_{abs}$ schreibt man dann $*$ statt t. Ein Zustand aus S_{abs} wird selbst nicht explizit dargestellt.

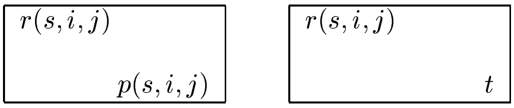


Abbildung 1.1: Graphische Darstellung stochastischer Spiele

Eine *Strategie* ist einen Vorschrift dafür, welche (gemischte) Aktion ein Spieler in jeder denkbaren Spielsituation („Historie") wählt. Den einfachsten Fall stellen *stationäre* Strategien dar: Bei jedem Besuch in einem Zustand s verwendet der Spieler dieselbe Aktion. Stationäre Strategien x, y von Spieler 1 und 2 sind also Vektoren $x = (x_1, \ldots, x_z)$ bzw. $y = (y_1, \ldots, y_z)$ mit $x_s \in \Delta^{n_s^1}$ bzw. $y_s \in \Delta^{n_s^2}$ für alle $s \in S$. Dabei sei Δ^n die Menge der W-Maße auf $\{1, \ldots, n\}$. Die Mengen stationärer Strategien werden mit X bzw. Y bezeichnet.

Bei *Markow-Strategien* hängt die Aktionswahl außer vom aktuellen Zustand nur von der Nummer der aktuellen Spielrunde ab. Solche Strategien f, g von Spieler 1 und 2 sind also Abbildungen $f\colon \mathbb{N} \to X$ bzw. $g\colon \mathbb{N} \to Y$. Mit $f_s^n = (f_s^n(1), \ldots, f_s^n(z))$ (analog g_s^n) sei dann abkürzend die von f bzw. g in Spielrunde n für Zustand s vorgeschriebene Aktion bezeichnet. Von *Semi-Markow-Strategien* spricht man, wenn darüber hinaus noch Abhängigkeit vom Startzustand besteht.

Der allgemeine Strategiebegriff kann nun wie folgt erklärt werden: Mit $H_s^n := \{h_n = (s_1, i_1, j_1, s_2, i_2, \ldots, s_n)\colon s_k \in S,\ (i_k, j_k) \in A_{s_k} (1 \le k < n),\ s_n = s\}$ bezeichnet man die Menge der Historien der Länge n, die in Zustand s enden. Man setzt noch $H^n := \bigcup_{s \in S} H_s^n$ und $H := \bigcup_{n=1}^{\infty} H^n$.
Strategien π, σ von Spieler 1 und 2 lassen sich dann allgemein als Abbildungen $\pi : H \to \bigcup_{s \in S} \Delta^{n_s^1}$ bzw. $\sigma : H \to \bigcup_{s \in S} \Delta^{n_s^2}$ beschreiben, die für festes $s \in S$, $n \in \mathbb{N}$ und $h \in H_s^n$ $\pi(h) \in \Delta^{n_s^1}$ bzw. $\sigma(h) \in \Delta^{n_s^2}$ erfüllen. Die allgemeinen Strategiemengen von Spieler 1 und 2 werden mit Π bzw. Σ bezeichnet. Will man betonen, dass eine gegebene Strategie nicht Markowsch ist, spricht man auch von *Verhaltensstrategien*.

Ein Startzustand $s \in S$ und ein Strategiepaar $(\pi, \sigma) \in \Pi \times \Sigma$ induzieren einen stochastischen Prozess auf S. Das zugrunde liegende W-Maß bezeichnen wir mit $P_{s\pi\sigma}$. Folgende Zufallsgrößen werden zur Beschreibung des Spielablaufs benutzt: $Z(n)$ sei der Zustand in Spielrunde n, $A(n) = (A^1(n), A^2(n))$ das in dieser Runde gewählte Aktionspaar und $R(n)$ die resultierende Auszahlung (an Spieler 1). $H(n)$ bezeichne die nach n Spielrunden beobachtete Historie.

Ferner sei $S(n) = \sum_{m=1}^{n} R(m)$ das Kapital (von Spieler 1) nach n Runden und $\overline{S}(N) = \frac{1}{N}\sum_{n=1}^{N} S(n)$ sein mittlerer Besitz in den ersten N Runden. Da in vielen Beispielen absorbierende Zustände auftreten, führen wir noch die Stoppzeit $\tau = \inf\{n \in \mathbb{N}\colon Z(n) \in S_{abs}\}$ ein.

$\gamma(s,\pi,\sigma) := \liminf_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} E_{s\pi\sigma}(R(m)) = \liminf_{n\to\infty} \frac{1}{n} E_{s\pi\sigma}(S(n))$ nennt man die *Durchschnittsauszahlung* unter s, π, σ. Man setzt $\gamma(\pi,\sigma) := (\gamma(1,\pi,\sigma),\ldots,\gamma(z,\pi,\sigma))$. Begriffe wie Spielwert und (ε-)Optimalität seien bezüglich dieses und der weiter unten genannten Kriterien wie üblich erklärt. Der Spielwert $v = (v(1),\ldots,v(z)) \in \mathbb{R}^z$ zu γ existiert immer (vgl. [MN81]). Dabei bezeichne $v(s)$ den Spielwert bei Startzustand s.

Im Fall $v = 0$ ist es sinnvoll, ein stochastisches Spiel mit Hilfe der *Totalauszahlung* genauer zu untersuchen. Die Totalauszahlung gibt den minimalen durchschnittlichen erwarteten Besitz (von Spieler 1) an. Zu einem Startzustand s und einem Strategiepaar (π,σ) ist die Totalauszahlung definiert als $\gamma_T(s,\pi,\sigma) := \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} E_{s\pi\sigma}(S(n)) = \liminf_{N\to\infty} E_{s\pi\sigma}(\overline{S}(N))$. Wie oben sei noch $\gamma_T(\pi,\sigma) := (\gamma_T(1,\pi,\sigma),\ldots,\gamma_T(z,\pi,\sigma))$. Der totale Spielwert werde mit $v_T = (v_T(1),\ldots,v_T(z))$ bezeichnet, ferner wollen wir unter v_T^{low} und v_T^{up} wie üblich den unteren bzw. oberen Spielwert verstehen. Der Wert v_T braucht selbst dann nicht zu existieren, wenn $v = 0$ gilt (vgl. [Thu92, FV97]). In einem ihrer jüngsten Artikel [TV96] konnten Thuijsman und Vrieze zeigen, dass der totale Spielwert zumindest dann existiert, wenn $v = 0$ gilt *und* beide Spieler über γ-optimale stationäre Strategien verfügen.

In dieser Arbeit sollen die Auswirkungen der Vertauschung von Erwartungswert- und Limesbildung in der Definition der Totalauszahlung untersucht werden. Zu einem Startzustand s und einem Strategiepaar (π,σ) definieren wir also $\tilde{\gamma}_T(s,\pi,\sigma) := E_{s\pi\sigma}(\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} S(n))$ und nennen $\tilde{\gamma}_T$ die *alternative Totalauszahlung*. Analog zu $\gamma_T(\pi,\sigma)$ sei auch $\tilde{\gamma}_T(\pi,\sigma)$ erklärt. Der zugehörige Spielwert werde mit $\tilde{v}_T$ bezeichnet, entsprechend verwenden wir $\tilde{v}_T^{low}$ und $\tilde{v}_T^{up}$. Man weiß, dass die Vertauschung von Erwartungswert und Limes den Spielwert unter dem Kriterium der Durchschnittsauszahlung nicht ändert (obwohl für nicht-stationäre Strategien unterschiedliche Bewertungen auftreten können). Im Fall der Totalauszahlung sind die Folgen dieser Vertauschung allerdings viel schwerwiegender.

1.2 Motivation der alternativen Totalauszahlung

Zuerst möchte ich die Behandlung der alternativen Totalauszahlung durch ein Beispiel motivieren, das einen Mangel der klassischen Definition aufzeigt.

Γ:

Zustand 1

1 3	−1 2
−1 2	1 3

Zustand 2

0 1	0 *

Zustand 3

0 1
0 *

Γ':

Zustand 1

1 2	−1 2
−1 2	1 2

Zustand 2

0 1	0 *

Abbildung 1.2: Die Spiele Γ und Γ'

Angenommen, Sie spielen regelmäßig das folgende Spiel: In jeder Runde wird eine Münze geworfen. Je nach Ausgang gewinnen oder verlieren Sie eine Mark und derjenige, der die Mark gewonnen hat, darf entscheiden, ob das Spiel fortgesetzt wird oder nicht. Vermutlich würden Sie dieses Spiel als „fair" ansehen. Doch nun unterbreitet Ihnen Ihr Kontrahent den Vorschlag, das Spiel so abzuändern, dass nach jeder Runde nur *er* entscheiden darf, ob weitergespielt wird. Sicher würden Sie dieser Variante nicht zustimmen.

Der entscheidende Punkt ist, dass Ihr Gegner keine Angst mehr zu haben bräuchte, zunächst viel Geld an Sie zu verlieren, da er sicher sein könnte, eines Tages wieder im Plus zu sein. Abbildung 1.2 zeigt dieses Spiel und seine offenbar ungünstige Variante in der üblichen Form. Dabei befinden Sie sich in der Rolle von Spieler 1. Dass Zustand 1 der Realisierung eines Münzwurfs entspricht, folgt aus der Tatsache, dass beide Spieler sinnvollerweise die gemischte Aktion $(\frac{1}{2}, \frac{1}{2})$ wählen, wann immer sich das Spiel in diesem Zustand befindet (sowohl in Γ als auch in Γ').

Somit haben wir $E_{s\pi\sigma}(R(n)) = 0 \; \forall n \in \mathbb{N}$ für jeden Startzustand s und jedes (sinnvolle) Strategiepaar (π, σ) in beiden Spielen. Also gilt $\gamma_T(s, \pi, \sigma) = 0$, sodass beide Spiele mit $v_T = 0$ bewertet werden. Dies suggeriert, dass keins von ihnen vorteilhaft für einen der Spieler ist.

Betrachten wir zuerst Γ. Das Spiel ist völlig symmetrisch im Hinblick auf die Rollen der Spieler. Es sei $x \in X$ mit $x_3 = (0,1)$ (bzw. $y \in Y$ mit $y_2 = (0,1)$). Mit $\tau = \inf\{n : Z(n) \in S_{abs}\}$ ist also trivialerweise $E_{sx\sigma}(\tau) < \infty$ für alle $\sigma \in \Sigma$ (und $E_{s\pi y}(\tau) < \infty$ für alle $\pi \in \Pi$). Im Vorgriff auf Lemma 2.2 folgt nun $\tilde{\gamma}_T(s,x,\sigma) = \gamma_T(s,x,\sigma) = 0$ (analog für (π,y)). Aus Korollar 2.4 schließen wir $\tilde{v}_T = v_T = 0$, und x, y sind somit $\tilde{\gamma}_T$-optimal. Es ist offensichtlich, dass die Möglichkeit, das Spiel zu stoppen (hier etwa nach Gewinn der ersten Mark), ein sehr mächtiges Instrument ist, weil in diesem Fall der andere Spieler keine Strategie besitzt, die ihm einen positiven erwarteten Gewinn sichert.
Betrachten wir nun Γ'. Wie bei Γ scheint die Bewertung $v_T = 0$ auszudrücken, dass das Spiel keinen Spieler bevorteilt. Doch wenn Spieler 2 etwa die Verhaltensstrategie σ_{-L} anwendet, die dann absorbiert (durch Wahl von Aktion 2 in Zustand 2), wenn sein Besitz zum ersten Mal $L \in \mathbb{N}$ beträgt, so beendet er das Spiel fast sicher mit einem Gewinn von L Mark. Spieler 1 besitzt keine Strategie, um dies zu verhindern. Somit ist die zugehörige alternative Totalauszahlung gleich $-L$. Der alternative totale Spielwert ist also $-\infty$, was den Umstand widerspiegelt, dass Spieler 2 sich (fast sicher) einen beliebig großen Gewinn garantieren kann.

Dieses Beispiel zeigt, dass die alternative Totalauszahlung zu völlig anderen Ergebnissen führen kann als die ursprüngliche Definition. Ferner sieht man, dass ein Mangel der klassischen Totalauszahlung darin besteht, dass nur Erwartungswerte von Einkommen aus *endlich* vielen Runden eingehen, obwohl in Spielen wie Γ' gerade Strategien von Bedeutung sind, die ein fast sicheres Ereignis abwarten, dessen Eintritt beliebig spät erfolgen kann.

Nun haben wir die Einführung der alternativen Totalauszahlung motiviert. Wie oben bereits erwähnt lautet die formale Definition

$$\tilde{\gamma}_T(s,\pi,\sigma) := E_{s\pi\sigma}\left(\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} S(n)\right), \tag{1.1}$$

wobei $s \in S$ ein Startzustand und $(\pi,\sigma) \in \Pi \times \Sigma$ ein Strategiepaar ist. Leider müssen wir uns mit dem Problem auseinandersetzen, dass das rechts stehende Integral nicht zu existieren braucht:

1.3 Allgemeine Definition von $\tilde{\gamma}_T$

Nehmen wir etwa die folgende Verhaltensstrategie $\hat{\sigma}$ von Spieler 2 in Γ': Wenn sich das Spiel in Zustand 2 befindet und der *Absolutbetrag* seines Besitzes zum ersten Mal gleich $k \in \mathbb{N}$ ist, wähle er Aktion 2 (Absorption) mit Wahrscheinlichkeit $\frac{1}{k+2}$. In allen anderen Situationen setze er das Spiel fort. Spieler 1 verfügt nur über *eine* (sinnvolle) Strategie x. Offenbar ist

$$P_{1x\hat{\sigma}}\left(\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}S(n)=k\right)=P_{1x\hat{\sigma}}\left(\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}S(n)=-k\right)$$
$$=\frac{1}{2}\cdot\frac{1}{k+2}\cdot\prod_{j=1}^{k-1}\frac{j+1}{j+2}=\frac{1}{(k+1)(k+2)}=:p_k$$

Die Nichtexistenz von $E_{1x\hat{\sigma}}(\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}S(n))$ folgt nun sofort aus

$$E_{1x\hat{\sigma}}\left(\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}S(n)\right)^{+}=E_{1x\hat{\sigma}}\left(\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}S(n)\right)^{-}=\sum_{k=1}^{\infty}p_k k=\infty.$$

Diese Problematik tritt nicht nur bei der hier betrachteten alternativen Totalauszahlung auf, sondern findet sich auch an vielen anderen Stellen in der Literatur. Oft wird die mögliche Nichtexistenz des Erwartungswertes (1.1) – oder einer ähnlich definierten Bewertung – durch restriktive Grundannahmen einfach ausgeklammert. In manchen Büchern über Markowsche Entscheidungsprobleme (die als stochastische Spiele mit nur einem Spieler anzusehen sind) wird etwa die Strategiemenge auf solche Handlungsvorschriften beschränkt, für die der erwartete Gesamtgewinn existiert – oder zumindest der Positiv- oder Negativteil dieses Integrals. Dazu siehe etwa van der Wal [vdW84].

Aus spieltheoretischer Sicht ist eine Einschränkung der Strategiemengen allerdings sehr problematisch, da die Existenz des Erwartungswertes (1.1) zumeist von den Strategien *beider* Spieler abhängt, es zwischen diesen aber keine Absprache gibt. Bei der Betrachtung des unteren Spielwerts kann man Spieler 1 zwar auf solche Strategien π beschränken, für die der Negativteil dieser Erwartung für *alle* $\sigma \in \Sigma$ endlich ist, doch könnten diese π (im Rahmen der Berechnung des oberen Spielwerts) unverzichtbar sein, um auf bestimmte *feste* σ zu antworten. Ein genereller Ausschluss solcher Strategien könnte Spieler 1 also stark beeinträchtigen. Entsprechendes gilt im umgekehrten Fall für Spieler 2. Um Konzepte wie Spielwert und Optimalität sinnvoll untersuchen zu können, muss man mittels $\tilde{\gamma}_T$ also *jedem* beliebigen Strategiepaar $(\pi,\sigma)\in\Pi\times\Sigma$ eine Bewertung aus $\overline{\mathbb{R}}$ zuweisen.

Eine andere in der Literatur vorkommende Methode zur Umgehung der Problematik ist die Beschränkung auf bestimmte Klassen von Spielen. In ihrem Übersichtsband [FV97] über Markowsche Entscheidungsprobleme und stochastische Spiele benutzen Filar und Vrieze ein der hier untersuchten Bewertung $\tilde{\gamma}_T$ entsprechendes Kriterium etwa nur für transiente oder positive Spiele. Im ersten Fall ist die erwartete Spieldauer unter *jedem* Strategiepaar endlich, im zweiten sind alle Auszahlungen positiv. Beide Annahmen bedeuten eine sehr starke Restriktion. Auch in Abhandlungen über Markowsche Entscheidungsprobleme werden die Auszahlungen des Öfteren als nicht-negativ vorausgesetzt, so etwa in [DN83]. Maitra und Sudderth wiederum erzwingen in [MS96] die Existenz des erwarteten Gesamtgewinns durch Hinzunahme geeigneter Nutzenfunktionen.

Um ein uneingeschränkt verwendbares Bewertungskriterium zu erhalten, muss $\tilde{\gamma}_T(s, \pi, \sigma)$ für alle Spiele und alle Strategiepaare erklärt werden. Will man eine plausible Lösung für den Sonderfall der Nichtexistenz des Erwartungswertes (1.1) finden, so ist es hilfreich, sich auf die Motivation der alternativen Totalauszahlung zu besinnen: Bei endlichem Spiel wird der Schlussgewinn $S(\tau)$ von Spieler 1 herangezogen, und als Verallgemeinerung dessen bei unendlichem Spiel sein langfristig gesehen minimaler mittlerer Besitz. Es erscheint also plausibel, dass Spieler 1 ein Spiel mit einem Strategiepaar (π, σ), für das $P_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) > 0$ gilt, mit $-\infty$ bewertet, da er es nicht riskieren möchte, mit positiver Wahrscheinlichkeit einen Verlust von $-\infty$ zu erleiden. Dies entspricht der üblichen worst-case-Sichtweise von Spieler 1.
Ist $\liminf_{N\to\infty} \overline{S}(N) > -\infty$ fast sicher und existiert der Erwartungswert $E_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N))$ dennoch *nicht*, so kann man der gegebenen Symmetrie $E_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N))^+ = E_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N))^- = \infty$ dadurch Rechnung tragen, dass man $\tilde{\gamma}_T(s, \pi, \sigma) = 0$ setzt. In dieser Situation braucht Spieler 1 keinen „unendlichen“ Verlust zu fürchten.

Insgesamt wollen wir also die folgende **Sonderfallregelung** verwenden: Existiert die Erwartung (1.1) *nicht*, so sei

$$\tilde{\gamma}_T(s, \pi, \sigma) := \begin{cases} 0 & \text{falls } P_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) = 0, \\ -\infty & \text{sonst.} \end{cases} \tag{1.2}$$

Als *sinnvolle* Sonderfallregelungen wollen wir generell alle Definitionen ansehen, bei denen im ersten Fall statt 0 ein anderer Wert aus $\overline{\mathbb{R}}$ zugewiesen wird und im zweiten Fall entweder derselbe Wert oder $-\infty$ Verwendung findet.

Da die meisten hier vorgestellten Resultate nicht von dieser technischen Frage berührt werden, verwenden wir $\tilde{\gamma}_T$ ohne weitere Fallunterscheidung, weisen aber gegebenenfalls auf Unterschiede hin. Von entscheidender Bedeutung wird diese Thematik lediglich beim Nachweis der Nichtexistenz des Spielwerts in Kapitel 3 sein. So ist es zum jetzigen Zeitpunkt noch unklar, ob das dortige Resultat für *alle* sinnvollen Sonderfallregelungen gezeigt werden kann.

Kapitel 2

Beziehungen zwischen $\tilde{\gamma}_T$ und γ_T

In diesem Abschnitt sollen einige einfache Bedingungen vorgestellt werden, unter denen beide Versionen der Totalauszahlung übereinstimmen. Ferner werden wir untersuchen, welche Verallgemeinerungen dieser Bedingungen möglich sind. Die betreffenden Aussagen werden im Weiteren häufiger Verwendung finden.

Wie oben betrachten wir die Stoppzeit $\tau = \inf\{n \in \mathbb{N} \colon Z(n) \in S_{abs}\}$. Im Fall $\tau < \infty$ erfolgt Absorption in Runde τ und es ist $\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} S(n) = S(\tau) = S(\tau - 1)$. Unter der Bedingung $\tau < \infty$ $P_{s\pi\sigma}$-fast sicher gilt also $\tilde{\gamma}_T(s,\pi,\sigma) = E_{s\pi\sigma}(S(\tau))$ (sofern dieser Erwartungswert erklärt ist). Zunächst zeigen wir ein grundlegendes Lemma.

Lemma 2.1 *Es seien $s \in S$, $(\pi,\sigma) \in \Pi \times \Sigma$ derart, dass $P_{s\pi\sigma}(\tau < \infty) = 1$. Ist außerdem $E_{s\pi\sigma}(S(\tau)) \in \mathbb{R}$, so gilt*

$$\gamma_T(s,\pi,\sigma) = \tilde{\gamma}_T(s,\pi,\sigma) + \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E_{s\pi\sigma}(S(n) 1_{\{\tau > n\}}). \tag{2.1}$$

Beweis: Es ist

$$\gamma_T(s,\pi,\sigma) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E_{s\pi\sigma}(S(n)) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E_{s\pi\sigma}(S(n \wedge \tau))$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Big(E_{s\pi\sigma}(S(\tau) 1_{\{\tau \le n\}}) + E_{s\pi\sigma}(S(n) 1_{\{\tau > n\}}) \Big). \tag{2.2}$$

Wegen $P_{s\pi\sigma}(\tau < \infty) = 1$ gilt $\lim_{n\to\infty} E_{s\pi\sigma}(S(\tau)1_{\{\tau\leq n\}}) = E_{s\pi\sigma}(S(\tau))$. Also ist (2.2) gleich der rechten Seite in (2.1). □

Es sei darauf hingewiesen, dass Gleichung (2.1) in der Situation von Lemma 2.1 allgemeiner auch dann gilt, wenn $E_{s\pi\sigma}(S(\tau))$ gleich $+\infty$ oder $-\infty$ und der liminf in (2.1) ungleich $-E_{s\pi\sigma}(S(\tau))$ ist.

Das nächste Lemma gibt uns zwei hinreichende Kriterien für die Gleichheit von $\tilde{\gamma}_T$ und γ_T an die Hand. Man beachte, dass für beliebige *stationäre* Strategien mit Satz 5.20 später eine noch stärkere Aussage möglich wird.

Lemma 2.2 *Es seien $s \in S$, $(\pi, \sigma) \in \Pi \times \Sigma$ derart, dass $P_{s\pi\sigma}(\tau < \infty) = 1$. Ferner sei eine der beiden folgenden Bedingungen erfüllt:*

(a) $E_{s\pi\sigma}(\tau) < \infty$,

(b) Es gibt $K \in \mathbb{R}^+$, sodass $P_{s\pi\sigma}(|S(n)| \leq K) = 1$ für alle $n \in \mathbb{N}$.

Dann gilt $\tilde{\gamma}_T(s, \pi, \sigma) = \gamma_T(s, \pi, \sigma)$.

Beweis: Ist (a) erfüllt, so gilt $E_{s\pi\sigma}(|S(n)1_{\{\tau>n\}}|) \leq nMP_{s\pi\sigma}(\tau > n)$, wobei $M := \max\{|r(s,i,j)| : s \in S,\ (i,j) \in A_s\}$. Angenommen, $nP_{s\pi\sigma}(\tau > n)$ konvergiert nicht gegen Null für $n \to \infty$. Dann gibt es $\varepsilon > 0$, sodass für jedes $k \in \mathbb{N}$ ein $n_k \geq k$ mit $n_k P_{s\pi\sigma}(\tau > n_k) \geq \varepsilon$ existiert. Ohne Einschränkung dürfen wir annehmen, dass $n_{k+1} \geq 2n_k$ für alle k. Sei noch $n_0 := 0$. Nun ist

$$\begin{aligned} E_{s\pi\sigma}(\tau) &= \sum_{n=0}^{\infty} P_{s\pi\sigma}(\tau > n) \geq \sum_{k=1}^{\infty} (n_k - n_{k-1}) P_{s\pi\sigma}(\tau > n_k) \\ &\geq \sum_{k=1}^{\infty} \frac{n_k - n_{k-1}}{n_k} \varepsilon \geq \sum_{k=1}^{\infty} \frac{1}{2} \varepsilon = \infty, \end{aligned}$$

was der Voraussetzung (a) widerspricht. Also gilt $E_{s\pi\sigma}(|S(n)1_{\{\tau>n\}}|) \to 0$ für $n \to \infty$. Außerdem existiert $E_{s\pi\sigma}(S(\tau))$ wegen $|S(\tau)| \leq M\tau \in \mathcal{L}^1(P_{s\pi\sigma})$. Ist Bedingung (b) erfüllt, so gilt $E_{s\pi\sigma}(|S(n)1_{\{\tau>n\}}|) \leq KP_{s\pi\sigma}(\tau > n) \to 0$ wegen $P_{s\pi\sigma}(\tau < \infty) = 1$. $E_{s\pi\sigma}(S(\tau))$ existiert aufgrund von $|S(\tau)| \leq K$. Somit ist in beiden Fällen $\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E_{s\pi\sigma}(S(n)1_{\{\tau>n\}}) = 0$ und $E_{s\pi\sigma}(S(\tau)) \in \mathbb{R}$. Mit Lemma 2.1 erhalten wir die Behauptung. □

Da die Bedingungen aus diesem Lemma sehr restriktiv sind, stellt sich die Frage, inwieweit sie abgeschwächt werden können:

Anmerkung 1: Beschränkung auf die Bedingung $P_{s\pi\sigma}(\tau < \infty) = 1$
Schon für Markow-Strategien lässt sich ein Beispiel angeben, in dem trotz $\tau < \infty$ $P_{s\pi\sigma}$-fast sicher $\tilde{\gamma}_T(s, \pi, \sigma) \neq \gamma_T(s, \pi, \sigma)$ gilt. Dazu betrachten wir

das Spiel Γ' aus Abbildung 1.2 und ändern lediglich $r(2,1,2) := c \in \mathbb{R}$. Spieler 1 verwende seine einzige (sinnvolle) Strategie $x \in X$ mit $x_1 = (\frac{1}{2}, \frac{1}{2})$. Es sei $I := \{2k^2 \colon k \in \mathbb{N}\}$. Spieler 2 spiele die Markow-Strategie g mit

$$g_1^n = \Big(\frac{1}{2}, \frac{1}{2}\Big), \quad g_2^n = \begin{cases} (\frac{k}{k+1}, \frac{1}{k+1}) & \text{falls } n = 2k^2 \in I, \\ (1,0) & \text{falls } n \notin I \end{cases} \qquad \forall n \in \mathbb{N}.$$

Dann gilt $P_{1xg}(\tau = 2n^2 + 1) = \big(\prod_{k=1}^{n-1} \frac{k}{k+1}\big) \frac{1}{n+1} = \frac{1}{n(n+1)}$ $\forall n \in \mathbb{N}$ und somit $P_{1xg}(\tau < \infty) = 1$. Es sei noch $y := (y_1, y_2) = ((\frac{1}{2}, \frac{1}{2}), (1,0)) \in Y$. Dann ist

$$\begin{aligned} E_{1xg}(S(\tau)^+) &= \sum_{n=1}^{\infty} E_{1xg}(S(2n^2)^+ 1_{\{\tau = 2n^2+1\}}) \\ &\geq \sum_{n=1}^{\infty} E_{1xy}(S(2n^2-1)^+ - |c|) P_{1xg}(\tau = 2n^2+1) && (2.3) \\ &= \sum_{n=1}^{\infty} \big(\sqrt{2n^2-1} - |c| + \Theta(1)\big) \frac{1}{n(n+1)} = \infty, && (2.4) \end{aligned}$$

wobei in (2.3) eingeht, dass x, g Markow-Strategien sind und $S(\tau - 1) = S(\tau - 2) + c$ gilt. In (2.4) verwendet man ein vom gewöhnlichen Random Walk bekanntes Resultat (siehe z.B. [Gut88]). Auf exakt dieselbe Weise zeigt man $E_{1xg}(S(\tau)^-) = \infty$. Also wird $\tilde{\gamma}_T(1, x, g)$ je nach der Sonderfallregelung aus der Definition von $\tilde{\gamma}_T$ (vgl. (1.2)) auf einen bestimmten Wert aus $\overline{\mathbb{R}}$ gesetzt – in unserem Modell z.B. auf 0.
Andererseits ist $E_{1xg}(R(n)) = cP_{1xg}(\tau = n + 1)$ für alle $n \in \mathbb{N}$, das heißt $E_{1xg}(S(n)) = cP_{1xg}(\tau \leq n + 1) \to c$ für $n \to \infty$. Somit gilt $\gamma_T(1, x, g) = c$. Wählt man also $c \neq \tilde{\gamma}_T(1, x, g)$, so stimmen beide Kriterien *nicht* überein.

Anmerkung 2: Beschränkung auf die Bedingungen (a) oder (b)
Es sei $P_{s\pi\sigma}(\tau < \infty) < 1$. Bedingung (a) ist in diesem Fall ohnehin nicht erfüllt, es gelte also (b). Erneut lässt sich schon für Markow-Strategien ein Beispiel angeben, in dem trotzdem $\tilde{\gamma}_T(s, \pi, \sigma) \neq \gamma_T(s, \pi, \sigma)$ gilt:
Dazu betrachten wir das Bad Match aus Kapitel 4 (siehe Abbildung 4.1). In diesem Spiel ist (b) sogar für *alle* Strategiepaare erfüllt. Wir fügen eine Kopie dieses Spiels hinzu (als Zustände 4, 5, und 6) und drehen darin mittels $r(s, i, j) = -r(s - 3, i, j)$ $(4 \leq s \leq 6)$ sämtliche Vorzeichen der Auszahlungen um. Schließlich ergänzen wir den Zustand 7 mit $A_7 = \{1\} \times \{1\}$, $r(7, 1, 1) = 0$ und $p(1|7, 1, 1) = p(4|7, 1, 1) = \frac{1}{2}$. Wir wählen Startzustand 7, sodass entweder das Bad Match oder seine mit Minuszeichen versehene Kopie gespielt wird. Dann ist stets $R(1) = 0$.
Spieler 1 verwende $x \in X$ mit $x_1 = x_4 = (1, 0)$, das heißt er wähle in den nicht-trivialen Zuständen stets Aktion 1. Dadurch ist $P_{7x\sigma}(\tau < \infty) = 0$ für jedes $\sigma \in \Sigma$. Nun konstruieren wir eine Markow-Strategie g von Spieler 2 wie folgt: Es sei $g_1^2 = g_4^2 = (1, 0)$, sodass mit $u := Z(2)(\omega)$ fast sicher

$R(2) = 2(-1)^u$ und $R(3) = -2(-1)^u$, also $\overline{S}(3) = \frac{2}{3}(-1)^u$ gilt. Weiter sei $g_1^n = g_4^n = (0,1)$ für $n = 4,\ldots,14$, sodass fast sicher $\overline{S}(15) = -\frac{2}{3}(-1)^u$. Danach sei wieder $g_1^n = g_4^n = (1,0)$ für $n = 16,\ldots,74$ und immer so weiter. Eine genaue Formalisierung dürfte sich erübrigen, denn unter x kann Spieler 2 für $n \geq 1$ jedes $R(2n)$ (und damit auch $R(2n+1)$) durch seine Aktion $A^2(2n)$ selbst bestimmen. Auf diese Weise gilt $\liminf_{N\to\infty} \overline{S}(N) = -\frac{2}{3}$ fast sicher, also $\tilde{\gamma}_T(7,x,g) = -\frac{2}{3}$.
Andererseits gilt $E_{7xg}(R(n)) = 0$ für alle n, denn auf den (gleich wahrscheinlichen) Mengen $\{Z(2) = 1\}$ bzw. $\{Z(2) = 4\}$ haben alle $R(n)$ jeweils entgegengesetzte Vorzeichen. Somit ist $\tilde{\gamma}_T(7,x,g) = -\frac{2}{3} \neq 0 = \gamma_T(7,x,g)$.

Anmerkung 3: Notwendige Bedingungen
Obwohl die hinreichenden Bedingungen aus Lemma 2.2 also kaum in allgemeiner Weise abgeschwächt werden können, sind sie erwartungsgemäß nicht *notwendig* für die Gleichheit der beiden Totalauszahlungsversionen. Dazu sei auf die Bemerkungen im Anschluss an Satz 3.5 und Lemma 4.2 verwiesen.

Aus Lemma 2.2 ergeben sich noch zwei Korollare:

Korollar 2.3 *Wenn Spieler 1 eine γ_T-optimale Strategie π^* besitzt, sodass die Voraussetzungen von Lemma 2.2 für jedes Tripel (s,π^*,σ) erfüllt sind, so ist $\tilde{v}_T^{low} = v_T^{low}$. Analog, wenn Spieler 2 eine γ_T-optimale Strategie σ^* besitzt, sodass die Voraussetzungen von Lemma 2.2 für jedes Tripel (s,π,σ^*) erfüllt sind, so ist $\tilde{v}_T^{up} = v_T^{up}$.*

Korollar 2.4 *Wenn beide Spieler γ_T-optimale Strategien π^* bzw. σ^* besitzen, sodass die Voraussetzungen von Lemma 2.2 für alle Tripel (s,π,σ^*) und (s,π^*,σ) erfüllt sind, und der totale Spielwert v_T existiert, so existiert auch der alternative totale Spielwert $\tilde{v}_T$ und ist gleich v_T.*

Kapitel 3

Existenz des Spielwerts

Bekanntlich braucht der totale Spielwert v_T allgemein nicht zu existieren – selbst im Fall $v = 0$ nicht (siehe Abschnitt 3.3). Jedoch konnten Thuijsman und Vrieze [TV96] zeigen, dass in stochastischen Spielen, die die folgende Eigenschaft erfüllen, der totale Spielwert v_T *immer* existiert.

Eigenschaft P: Es ist $v = 0$ und beide Spieler verfügen über stationäre γ-optimale Strategien.

Thuijsman und Vrieze schließen in [TV96] mit der Frage, ob Eigenschaft P auch hinreichend für die Existenz des *alternativen* totalen Spielwerts ist. In diesem Kapitel geben wir hierauf eine (negative) Antwort, indem wir zeigen, dass das Spiel Γ^* (siehe Abbildung 3.1) zwar die Eigenschaft P erfüllt, der Spielwert $\tilde{v}_T$ (unter unserer Sonderfallregelung (1.2)) jedoch nicht existiert. Schließlich zeigen wir an einem anderen Beispiel, dass $\tilde{v}_T$ aber auch dann existieren kann, wenn v_T nicht existiert – und dies sogar unabhängig von der verwendeten Sonderfallregelung.

3.1 Γ^* unter der klassischen Totalauszahlung

Lemma 3.1 *Γ^* besitzt die Eigenschaft P.*

Beweis: Es seien $x = (x_1, x_2) \in X$ und $y = (y_1, y_2) \in Y$ mit $x_2 = y_1 = (0, 0, 1)$. Offenbar endet das Spiel nach höchstens zwei Runden, wenn x gespielt wird – unabhängig von der von Spieler 2 verwendeten Strategie σ. Da die Durchschnittsauszahlung in diesem Fall Null beträgt, erhalten wir $v \geq 0$. Analog ist $\gamma(s, \pi, y) = 0$ für alle s, π, also $v \leq 0$. Somit gilt $v = 0$ und x, y sind γ-optimale Strategien. □

Zustand 1

-2 / 2	0 / 2	0 / $*$
0 / 2	-2 / 2	0 / $*$

Zustand 2

2 / 1	0 / 1
0 / 1	2 / 1
0 / $*$	0 / $*$

Abbildung 3.1: Das Spiel Γ^*

Nach dem Resultat von Thuijsman und Vrieze existiert demnach der totale Spielwert v_T für Γ^*. Wieder bezeichne τ die Runde, in der Absorption erfolgt (nach vorheriger Wahl von Aktion 3). Wir dürfen annehmen, dass beide Spieler ihre Aktionen 1 und 2 in jeder Spielsituation mit gleicher Wahrscheinlichkeit wählen. Also gilt für alle (sinnvollen) $(\pi, \sigma) \in \Pi \times \Sigma$ und $s = 1, 2$ $E_{s\pi\sigma}(R(n)) = (-1)^{n+s-1} P_{s\pi\sigma}(\tau > n+1)$ und so

$$E_{s\pi\sigma}(S(2n)) = \sum_{m=1}^{2n} (-1)^{m+s-1} P_{s\pi\sigma}(\tau > m+1)$$

$$= \sum_{m=1}^{2n} (-1)^{m+s-1} \left(\sum_{k=m+1}^{2n} P_{s\pi\sigma}(\tau = k+1) + P_{s\pi\sigma}(\tau > 2n+1) \right)$$

$$= \sum_{k=2}^{2n} P_{s\pi\sigma}(\tau = k+1) \sum_{m=1}^{k-1} (-1)^{m+s-1} + P_{s\pi\sigma}(\tau > 2n+1) \sum_{m=1}^{2n} (-1)^{m+s-1}$$

$$= (-1)^s \sum_{k=1}^{n} P_{s\pi\sigma}(\tau = 2k+1) = (-1)^s P_{s\pi\sigma}(\tau \leq 2n+1 \text{ ungerade}).$$

Außerdem ist

$$E_{s\pi\sigma}(S(2n-1)) = (-1)^s P_{s\pi\sigma}(\tau > 2n) + \sum_{m=1}^{2n-2} (-1)^{m+s-1} P_{s\pi\sigma}(\tau > m+1)$$

$$= (-1)^s P_{s\pi\sigma}(\tau > 2n) + (-1)^s P_{s\pi\sigma}(\tau \leq 2n-1 \text{ ungerade})$$

nach obiger Rechnung. Insgesamt erhalten wir

$$\gamma_T(s, \pi, \sigma) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E_{s\pi\sigma}(S(n))$$

$$= \liminf_{N\to\infty} (-1)^s \left(\frac{1}{N} \sum_{n=1}^{N} P_{s\pi\sigma}(\tau \leq n+1 \text{ ungerade}) + \frac{1}{N} \sum_{n=1}^{\lceil N/2 \rceil} P_{s\pi\sigma}(\tau > 2n) \right)$$

$$= (-1)^s \Big(P_{s\pi\sigma}(\tau < \infty \text{ ungerade}) + \frac{1}{2} P_{s\pi\sigma}(\tau = \infty) \Big). \tag{3.1}$$

Man beachte, dass obiger liminf sogar ein echter Limes ist. Wir betrachten $x^* = (x_1^*, x_2^*) = ((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}, 0)) \in X$. Aus (3.1) gewinnen wir $\gamma_T(1, x^*, \sigma) = -\frac{1}{2} P_{1x^*\sigma}(\tau = \infty) \geq -\frac{1}{2}$ für alle $\sigma \in \Sigma$, also $v_T(1) \geq -\frac{1}{2}$. Genauso gilt mit $y^* = (y_1^*, y_2^*) = (x_2^*, x_1^*) \in Y$ $\gamma_T(1, \pi, y^*) = \frac{1}{2} P_{1\pi y^*}(\tau = \infty) - 1 \leq -\frac{1}{2}$ für alle $\pi \in \Pi$. Also ist $v_T(1) = -\frac{1}{2}$ (analog $v_T(2) = \frac{1}{2}$) und x^*, y^* sind γ_T-optimal.

Es ist noch anzumerken, dass Γ^* nicht nur die Eigenschaft P erfüllt und damit einen totalen Spielwert besitzt, sondern in diesem Spiel beide Spieler sogar über total optimale stationäre Strategien verfügen. Dass Γ^* also ein sehr „nettes“ Spiel ist, verdeutlichen die Untersuchungen aus [TV87, Thu92] und Kapitel 4. Umso frappierender ist die Tatsache, dass der alternative totale Spielwert *nicht* existiert:

3.2 Γ^* unter der alternativen Totalauszahlung

Man sieht sofort, dass die total optimalen Strategien x^* und y^* nicht $\tilde{\gamma}_T$-optimal sind. Dazu betrachtet man die Strategien π_K von Spieler 1 bzw. σ_{-L} von Spieler 2, bei denen der betreffende Spieler stoppt, sobald er zum ersten Mal das Kapital K bzw. L besitzt. Verwendet Spieler 2 σ_{-L} (wobei $L \in \mathbb{N}$ gerade ist), so gilt $\lim_{n\to\infty} S(n) = -L$ $P_{sx^*\sigma_{-L}}$-fast sicher, also $\tilde{\gamma}_T(s, x^*, \sigma_{-L}) = -L$. Somit ist $\inf_{L\in 2\mathbb{N}} \tilde{\gamma}_T(s, x^*, \sigma_{-L}) = -\infty$. Analog ist $\tilde{\gamma}_T(s, \pi_K, y^*) = K$ für gerades $K \in \mathbb{N}$, also $\sup_{K\in 2\mathbb{N}} \tilde{\gamma}_T(s, \pi_K, y^*) = \infty$. Trotzdem werden x^* und y^* im Folgenden noch mehrfach benötigt.

Nun wollen wir die alternative Totalauszahlung für Γ^* genauer untersuchen. Das nächste Lemma ist der Schlüssel zum Beweis der Nichtexistenz des alternativen totalen Spielwerts.

Lemma 3.2 *Es sei $s \in \{1, 2\}$ der Startzustand. Dann gilt*

(a) $\inf_\sigma \tilde{\gamma}_T(s, \pi, \sigma) \leq s - 1 - P_{s\pi y^*}(\tau < \infty)$ *für alle* $\pi \in \Pi$ *und*

(b) $\sup_\pi \tilde{\gamma}_T(s, \pi, \sigma) \geq s - 2 + P_{sx^*\sigma}(\tau < \infty)$ *für alle* $\sigma \in \Sigma$.

Beweis: Es sei $\pi \in \Pi$ beliebig und $p := P_{s\pi y^*}(\tau < \infty)$. Für alle $\varepsilon > 0$ gibt es ein N^ε, sodass $P_{s\pi y^*}(\tau \leq N^\varepsilon) \geq p - \varepsilon$. Ohne Einschränkung sei N^ε ungerade. Sei σ^{N^ε} die Semi-Markow-Strategie von Spieler 2, die in Runde $N^\varepsilon + s - 1$ Aktion 3 wählt (und in Zustand 1 vorher stets die gemischte Aktion $(\frac{1}{2}, \frac{1}{2}, 0)$ verwendet). Unter $s, \pi, \sigma^{N^\varepsilon}$ ist dann $\tau \leq N^\varepsilon + 2$. Lemma 2.2 und (3.1) liefern

$$\tilde{\gamma}_T(s, \pi, \sigma^{N^\varepsilon}) = \gamma_T(s, \pi, \sigma^{N^\varepsilon}) = (-1)^s P_{s\pi\sigma^{N^\varepsilon}}(\tau < \infty \text{ ungerade}).$$

Für $s = 1$ ist die rechte Seite gleich $-P_{1\pi\sigma^{N^\varepsilon}}(\tau \leq N^\varepsilon) = -P_{1\pi y^*}(\tau \leq N^\varepsilon) \leq -p + \varepsilon$, denn beendet Spieler 1 das Spiel vor der Runde N^ε nicht, so sorgt Spieler 2 in dieser Runde für Absorption. Analog folgt im Fall $s = 2$, dass die rechte Seite gleich $1 - P_{1\pi\sigma^{N^\varepsilon}}(\tau \leq N_\varepsilon + 1) \leq 1 - p + \varepsilon$ ist. Da $\varepsilon > 0$ beliebig war, erhalten wir $\inf_\sigma \tilde{\gamma}_T(s, \pi, \sigma) \leq s - 1 - p$. Der Beweis der symmetrischen Aussage (b) ist vollkommen analog. □

Nun wollen wir zeigen, dass $\tilde{v}_T$ nicht existiert. In den Beweisen der beiden folgenden Lemmata geht die Sonderfallregelung (1.2) ein. Änderungen dieser Definition würden also einen anderen Beweis erfordern.

Lemma 3.3 *Es ist* $\tilde{v}_T^{up}(1) = \inf_\sigma \sup_\pi \tilde{\gamma}_T(1, \pi, \sigma) = 0$.

Beweis: Offenbar ist $\tilde{v}_T^{up}(1) \leq 0$, denn beendet Spieler 2 das Spiel sofort, so garantiert er sich die Auszahlung 0. Nun sei $\sigma \in \Sigma$ beliebig. Gilt $P_{1x^*\sigma}(\tau < \infty) = 1$, so haben wir $\sup_\pi \tilde{\gamma}_T(1, \pi, \sigma) \geq 0$ nach Lemma 3.2.
Gilt dagegen $P_{1x^*\sigma}(\tau = \infty) =: \delta > 0$, so sei π wie folgt definiert: Spiele Aktion 3 (in Zustand 2) mit Wahrscheinlichkeit $\frac{1}{n+1}$, wenn der aktuelle Gewinn zum ersten Mal $2n \in \mathbb{N}$ beträgt. Dann ist $P_{1\pi y^*}(S(\tau) = 2n) = \left(\prod_{k=1}^{n-1} \frac{k}{k+1}\right)\frac{1}{n+1} = \frac{1}{n(n+1)}$, sodass $P_{1\pi\sigma}(\tau < \infty) \geq P_{1\pi y^*}(\tau < \infty) = 1$. Ist $C := \{\tau = \infty\} \cup \{\tau \text{ ungerade}\}$ das Ereignis, dass Spieler 2 nicht stoppt, so gilt $P_{1\pi\sigma}(S(\tau) = 2n) \geq P_{1\pi\sigma}(C)P_{1\pi\sigma}(S(\tau) = 2n \mid C) \geq \delta P_{1\pi y^*}(S(\tau) = 2n)$, also

$$E_{1\pi\sigma}(S(\tau)^+) = \sum_{n=1}^{\infty} 2n P_{1\pi\sigma}(S(\tau) = 2n) \geq \sum_{n=1}^{\infty} \frac{2\delta}{n+1} = \infty.$$

Je nachdem, ob $E_{1\pi\sigma}(S(\tau)^-)$ endlich ist oder nicht, ist $\tilde{\gamma}_T(1, \pi, \sigma)$ also entweder $+\infty$ oder 0. In beiden Fällen gilt $\sup_\pi \tilde{\gamma}_T(1, \pi, \sigma) \geq 0$. □

Lemma 3.4 $\tilde{v}_T^{low}(1) = \sup_\pi \inf_\sigma \tilde{\gamma}_T(1, \pi, \sigma) = -1$.

Beweis: Offenbar ist $\tilde{v}_T^{low}(1) \geq -1$, denn beendet Spieler 1 das Spiel in der zweiten Runde, so garantiert er sich die erwartete Auszahlung -1. Nun sei $\pi \in \Pi$ beliebig. Im Fall $P_{1\pi y^*}(\tau < \infty) = 1$ folgt $\inf_\sigma \tilde{\gamma}_T(1, \pi, \sigma) \leq -1$ aus Lemma 3.2. Ist dagegen $P_{1\pi y^*}(\tau = \infty) =: \delta > 0$, so gilt

$$\begin{aligned}
&P_{1\pi y^*}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) \\
&= P_{1\pi y^*}(\tau = \infty) P_{1\pi y^*}(\liminf_{N\to\infty} \overline{S}(N) = -\infty \mid \tau = \infty) \\
&= \delta P_{1x^*y^*}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) = \delta > 0
\end{aligned}$$

unter Benutzung von Korollar 5.11. Für die hier relevanten x^*, y^* ist dessen Aussage allerdings auch unmittelbar klar. Somit gilt $\tilde{\gamma}_T(1, \pi, y^*) = -\infty$, also insgesamt $\inf_\sigma \tilde{\gamma}_T(1, \pi, \sigma) \leq -1$ für alle $\pi \in \Pi$. □

Somit ist gezeigt:

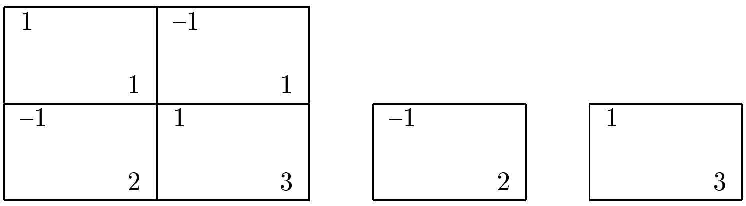
Satz 3.5 *Auch wenn Eigenschaft P für ein stochastisches Spiel erfüllt ist, braucht der alternative totale Spielwert $\tilde{v}_T$ nicht zu existieren.*

Obwohl $\tilde{v}_T$ im Gegensatz zu v_T für Γ^* nicht existiert, liefern beide Totalauszahlungskriterien dasselbe Resultat, wenn Markow-Strategien f, g gespielt werden, für die $\tau < \infty$ P_{sfg}-fast sicher und $E_{sfg}(S(\tau)) \in \mathbb{R}$: In diesem Fall gilt $E_{sfg}(S(n)1_{\{\tau>n\}}) = E_{sx^*y^*}(S(n))P_{sfg}(\tau > n+1) = -P_{sfg}(\tau > n+1)$ für ungerade n (und $= 0$ für gerade n). Wegen $P_{sfg}(\tau > n+1) \to 0$ liefert Lemma 2.1 $\tilde{\gamma}_T(s, f, g) = \gamma_T(s, f, g)$.

Damit ist noch gezeigt, dass die in Lemma 2.2 genannten hinreichenden Bedingungen für die Übereinstimmung von $\tilde{\gamma}_T$ und γ_T nicht *notwendig* sind. (Hier wird außer $P_{sfg}(\tau < \infty) = 1$ nur die Existenz von $E_{sfg}(S(\tau))$ verlangt. Es lassen sich aber leicht Paare (f, g) angeben, für die die Bedingungen (a) und (b) aus diesem Lemma trotzdem nicht erfüllt sind.)

3.3 Zusammenhang zwischen $\tilde{v}_T$ und v_T

Im letzten Abschnitt wurde gezeigt, dass die Eigenschaft P und die damit verbundene Existenz des gewöhnlichen totalen Spielwerts v_T nicht hinreichend für die Existenz des alternativen totalen Spielwerts $\tilde{v}_T$ ist. Darum stellt sich nun in natürlicher Weise die Frage, ob die Existenz von v_T zumindest eine *notwendige* Voraussetzung für die Existenz von $\tilde{v}_T$ ist. In diesem Abschnitt zeigen wir an einem Beispiel, dass auch dies nicht der Fall ist. Interessant dabei ist, dass dieses Resultat *nicht* von der Sonderfallregel (vgl. (1.2)) in der Definition von $\tilde{\gamma}_T$ abhängt.



Zustand 1 Zustand 2 Zustand 3

Abbildung 3.2: Das Spiel Γ^{**}

Γ^{**} ist eine Variante des bekannten Big Match von Blackwell und Ferguson (siehe [BF68]). Zunächst zeigen wir die Nichtexistenz von v_T.

Lemma 3.6 *Der totale Spielwert v_T existiert für Γ^{**} nicht.*

Beweis: Sei $\pi \in \Pi$ beliebig und $\eta := \inf\{n \colon A^1(n) = 2\}$ die erste Runde, in der Spieler 1 Aktion 2 wählt. Sei $y \in Y$ mit $y_1 = (0,1)$. Dann gilt im Fall $P_{1\pi y}(\eta < \infty) = 0$, dass $R(n) \equiv -1$ für alle n, also $\gamma_T(1, \pi, y) = -\infty$.
Sonst sei $m := \inf\{n \colon P_{1\pi y}(\eta = n) > 0\}$ und $\varepsilon := P_{1\pi y}(\eta = m)$. In diesem Fall verwende Spieler 2 die Markow-Strategie g mit $g_1^n = (0,1)$ für $n < m$, $g_1^m = (1,0)$ und $g_1^n = (\frac{1}{2}, \frac{1}{2})$ für $n > m$. Dann ist $R(n) \equiv -1$ für $n < m$, $E_{1\pi g}(R(m)) = 1 - 2\varepsilon$ und

$$\begin{aligned} E_{1\pi g}(R(n)) &= P_{1\pi g}(Z(n) = 3) - P_{1\pi g}(Z(n) = 2) \\ &= P_{1\pi g}(A^2(\eta) = 2,\, \eta < n) - P_{1\pi g}(A^2(\eta) = 1,\, \eta < n) \\ &= -P_{1\pi g}(A^2(m) = 1,\, \eta = m) = -\varepsilon \end{aligned}$$

für alle $n > m$ nach Konstruktion von g. Auch hier folgt $\gamma_T(1, \pi, g) = -\infty$. Somit gilt $v_T^{low}(1) = -\infty$. Dieser Teil des Beweises ähnelt der Argumentation aus [BF68].
Sei nun $\sigma \in \Sigma$ beliebig und die Antwort π dazu wie folgt konstruiert: Ist h eine beliebige Historie, so setzen wir

$$\pi(h) = \begin{cases} (1,0) & \text{falls } \sigma(h)(1) \geq \frac{1}{2}, \\ (0,1) & \text{falls } \sigma(h)(1) < \frac{1}{2}. \end{cases} \tag{3.2}$$

Spieler 1 spielt also genau dann Aktion 2, wenn Spieler 2 die zweite Spalte eher wählt als die erste. Es ist einleuchtend, dass er sich so in jeder Runde eine nicht-negative erwartete Auszahlung garantiert. Im Einzelnen sieht man dies so:
Es ist $E_{1\pi\sigma}(R(n)) = \sum_{h \in H^n} E_{1\pi\sigma}(R(n) \mid H(n) = h) P_{1\pi\sigma}(H(n) = h)$. Dabei erfolge die Summation nur über solche h, für die die hintere Wahrscheinlichkeit größer Null ist. Im Fall $h \in H_1^n$ ist

$$E_{1\pi\sigma}(R(n) \mid H(n) = h) = \Big(\pi(h)(1) - \pi(h)(2)\Big)\Big(\sigma(h)(1) - \sigma(h)(2)\Big) \geq 0$$

nach (3.2). Im Fall $h \in H_s^n$ mit $s \neq 1$ gilt $E_{1\pi\sigma}(R(n) \mid H(n) = h) = (-1)^{s-1}$.
So errechnet man

$$\begin{aligned} E_{1\pi\sigma}(R(n)) &\geq \sum_{h \in H^n \setminus H_1^n} E_{1\pi\sigma}(R(n) \mid H(n) = h) P_{1\pi\sigma}(H(n) = h) \\ &= P_{1\pi\sigma}(Z(n) = 3) - P_{1\pi\sigma}(Z(n) = 2) \\ &= \sum_{h \in H_1^k \colon k < n} \Big(P_{1\pi\sigma}(H(k+1) = (h, (2,2), 3)) - P_{1\pi\sigma}(H(k+1) = (h, (2,1), 2))\Big) \\ &= \sum_{h \in H_1^k \colon k < n} \Big(P_{1\pi\sigma}(A(k) = (2,2) \mid H(k) = h) - P_{1\pi\sigma}(A(k) = (2,1) \mid H(k) = h)\Big) \\ &\quad \cdot P_{1\pi\sigma}(H(k) = h) \end{aligned}$$

$$= \sum_{h \in H_1^k:\, k<n} \pi(h)(2)\Big(\sigma(h)(2) - \sigma(h)(1)\Big) P_{1\pi\sigma}(H(k) = h) \geq 0$$

nach (3.2). Wegen $E_{1\pi\sigma}(R(n)) \geq 0$ für alle n ist also $\gamma_T(1,\pi,\sigma) \geq 0$. Damit ist gezeigt, dass $v_T^{up}(1) \geq 0 > -\infty = v_T^{low}(1)$. □

Nun zum alternativen totalen Spielwert $\tilde{v}_T$:

Lemma 3.7 *Der alternative totale Spielwert $\tilde{v}_T$ existiert für Γ^{**}. Die Existenz gilt für alle (sinnvollen) Sonderfallregelungen in der Definition von $\tilde{\gamma}_T$.*

Beweis: Wenn im Fall, dass der Erwartungswert (1.1) nicht existiert, generell $\tilde{\gamma}_T(s,\pi,\sigma) = w$ gesetzt wird (wobei $w \in \overline{\mathbb{R}}$ beliebig ist), ist die Existenz von $\tilde{v}_T$ wie folgt zu sehen:
Es sei $\pi \in \Pi$ beliebig und $y \in Y$ mit $y_1 = (\frac{1}{2}, \frac{1}{2})$. Ist $P_{1\pi y}(A^1(n) = 2) = 0$ für alle n, so gilt $\liminf_{N\to\infty} \overline{S}(N) = -\infty$ fast sicher. (Streng genommen benutzt man hierfür wieder Korollar 5.11.) In diesem Fall ist also $\tilde{\gamma}_T(1,\pi,y) = -\infty$. Gilt $P_{1\pi y}(A^1(n) = 2) =: \varepsilon > 0$ für irgendein n, so ist $P_{1\pi y}(\liminf_{N\to\infty} \overline{S}(N) = \infty) \geq \frac{\varepsilon}{2}$ und $P_{1\pi y}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) \geq \frac{\varepsilon}{2}$, also $\tilde{\gamma}_T(1,\pi,y) = w$. Folglich ist $\tilde{v}_T^{up}(1) \leq w$.
Nun sei f die Markow-Strategie von Spieler 1 mit $f_1^n = \big(\frac{n(n+2)}{(n+1)^2}, \frac{1}{(n+1)^2}\big)$. Für jedes σ ist dann $P_{1f\sigma}(A^1(n) = 2) = \big(\prod_{k=1}^{n-1} f_1^k(1)\big) f_1^n(2) = \frac{1}{2n(n+1)}$ $\forall n$, d.h. es gilt $P_{1f\sigma}(A^1(n) = 2 \text{ für irgendein } n) = \sum_{n=1}^{\infty} \frac{1}{2n(n+1)} = \frac{1}{2}$.
Sei $\sigma \in \Sigma$ beliebig. Im Fall $P_{1f\sigma}(A^2(n) = 2) = 0$ $\forall n$ gilt $\liminf_{N\to\infty} \overline{S}(N) = \infty$ auf der Menge $\{A^1(n) = 1 \text{ für alle } n\}$ und $\liminf_{N\to\infty} \overline{S}(N) = -\infty$ sonst. Nach Konstruktion von f haben wir somit $P_{1f\sigma}(\liminf_{N\to\infty} \overline{S}(N) = \infty) = P_{1f\sigma}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) = \frac{1}{2}$, d.h. $\tilde{\gamma}_T(1,f,\sigma) = w$.
Ist dagegen $P_{1f\sigma}(A^2(n) = 2) =: \varepsilon > 0$ für irgendein n, so ergibt sich $P_{1f\sigma}(\liminf_{N\to\infty} \overline{S}(N) = \infty) \geq \varepsilon\big(\prod_{k=1}^{n-1} f_1^k(1)\big) f_1^n(2) > 0$. Demnach ist $\tilde{\gamma}_T(1,f,\sigma) \geq w$ und somit auch $\tilde{v}_T^{low}(1) \geq w$. Insgesamt folgt $\tilde{v}_T(1) = w$.
Fordert man im Sonderfall der Nichtexistenz der Erwartung (1.1) zusätzlich, dass bei $P_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) > 0$ stets $\tilde{\gamma}_T(s,\pi,\sigma) = -\infty$ gesetzt wird (wie es in unserem Modell der Fall ist), so folgt für Γ^{**} sofort $\tilde{v}_T(1) = -\infty$, denn mit der oben im Beweis verwendeten Strategie y kann Spieler 2 wie gesehen $P_{1\pi y}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) > 0$ erzwingen. □

Damit ist gezeigt:

Satz 3.8 *Die Existenz von v_T ist weder notwendig noch hinreichend für die Existenz von $\tilde{v}_T$.*

Kapitel 4

Optimale Strategien

In diesem Kapitel soll gezeigt werden, dass $\tilde{\gamma}_T$-optimale Strategien im Allgemeinen nicht zu existieren brauchen und Verhaltensstrategien unverzichtbar sein können, um ε-optimal zu spielen. Dazu betrachten wir das „Bad Match“, mit dem Thuijsman und Vrieze [TV87] dieselben Aussagen für die klassische Totalauszahlung bewiesen. Unser Ansatz wird der folgende sein: Zuerst zeigen wir, dass für dieses Spiel beide Kriterien im Fall von Markow-Strategien übereinstimmen. Im zweiten Schritt wird nachgewiesen, dass der alternative totale Spielwert $\tilde{v}_T$ für das Bad Match existiert und gleich v_T ist. Unter Benutzung der entsprechenden Resultate über γ_T ergeben sich die beiden zu zeigenden Aussagen dann ganz direkt.

4.1 Markow-Strategien im Bad Match

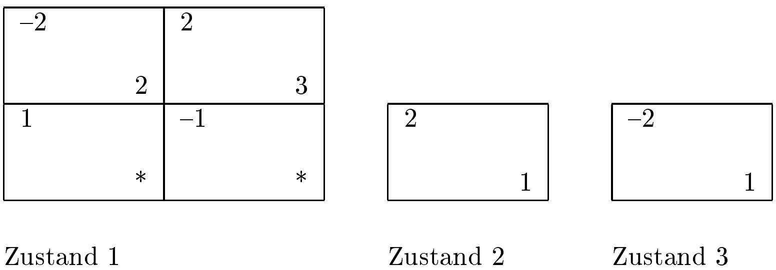


Abbildung 4.1: Das Bad Match

Zunächst ist festzuhalten, dass in diesem Spiel $|S(n)| \leq 2$ für alle $n \in \mathbb{N}$ gilt. Die Existenz des Erwartungswertes (1.1) bereitet hier also keine Schwierigkeiten – unabhängig vom Startzustand und den gewählten Strategien.

Es seien f, g Markow-Strategien von Spieler 1 bzw. 2. Um die Notation abzukürzen, sei $g^{(n)} := g_1^n(1)$. Dies ist die Wahrscheinlichkeit, dass Spieler 2 Aktion 1 wählt, wenn das Spiel in Runde n in Zustand 1 ist. Wir betrachten Startzustand 1. Dann befindet sich das Spiel in geraden Runden in einem trivialen Zustand, d.h. die $g^{(n)}$ interessieren nur für ungerade n. Wieder sei $\tau := \inf\{n \colon Z(n) \in S_{abs}\}$. Man beachte, dass τ niemals ungerade sein kann und das Kapital $S(n)$ für gerade $n < \tau$ stets Null beträgt. Also ist $S(n) = R(n)$ für ungerade $n < \tau$. Auf der Menge $\{\tau < \infty\}$ gilt offenbar $\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} S(n) = \lim_{n\to\infty} S(n) = R(\tau - 1) \in \{-1, 1\}$.

Im Fall $P_{1fg}(\tau < \infty) < 1$ gilt auf $\{\tau = \infty\}$ $\overline{S}(2N) = \frac{1}{2N}\sum_{n=1}^{2N} S(n) = \frac{1}{2N}\sum_{n=1}^{N} R(2n-1)$ und $\overline{S}(2N-1) = \frac{1}{2N-1}\sum_{n=1}^{N} R(2n-1) = \frac{2N}{2N-1}\overline{S}(2N)$, sodass dann

$$\liminf_{N\to\infty} \overline{S}(N) = \liminf_{N\to\infty} \overline{S}(2N) = \liminf_{N\to\infty} \frac{1}{2N}\sum_{n=1}^{N} R(2n-1).$$

Die $(R(2n-1))_{n\in\mathbb{N}}$ sind auf der Menge $\{\tau = \infty\}$ unabhängig und erfüllen $\mathrm{Var}_{1fg}(R(2n-1)) \leq 4\ \forall n \in \mathbb{N}$. Mit $\mu_{2n-1} := E_{1fg}(R(2n-1) \,|\, \tau = \infty) = 2 - 4g^{(2n-1)}$ ergibt das starke Gesetz $\frac{1}{N}\sum_{n=1}^{N}(R(2n-1) - \mu_{2n-1}) \to 0$ fast sicher auf $\{\tau = \infty\}$. Somit erhalten wir für fast alle $\omega \in \{\tau = \infty\}$ $\liminf_{N\to\infty} \overline{S}(N)(\omega) = \liminf_{N\to\infty}(1 - \frac{2}{N}\sum_{n=1}^{N} g^{(2n-1)})$. Es folgt

$$\begin{aligned} &\tilde{\gamma}_T(1, f, g) = E_{1fg}(\liminf_{N\to\infty} \overline{S}(N)) \\ &= E_{1fg}(\liminf_{N\to\infty} \overline{S}(N) 1_{\{\tau<\infty\}}) + E_{1fg}(\liminf_{N\to\infty} \overline{S}(N) 1_{\{\tau=\infty\}}) \qquad (4.1) \\ &= E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) + P_{1fg}(\tau = \infty) \liminf_{N\to\infty} \left(1 - \frac{2}{N}\sum_{n=1}^{N} g^{(2n-1)}\right). \end{aligned}$$

Nun wollen wir die klassische Totalauszahlung zum gegebenen Strategiepaar (f, g) berechnen. Für $n \in \mathbb{N}$ ist $E_{1fg}(S(2n)) = E_{1fg}(R(\tau-1)1_{\{\tau\leq 2n\}})$ und analog $E_{1fg}(S(2n-1)) = E_{1fg}(R(\tau-1)1_{\{\tau\leq 2n\}}) + E_{1fg}(R(2n-1)1_{\{\tau>2n\}})$, wobei der letzte Term gleich $P_{1fg}(\tau > 2n)\mu_{2n-1}$ ist. Somit gilt

$$\begin{aligned} &\gamma_T(1, f, g) = \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} E_{1fg}(S(n)) \\ &= \liminf_{N\to\infty} \frac{1}{N}\left(\sum_{n=1}^{N} E_{1fg}(R(\tau-1)1_{\{\tau\leq n+1\}}) + \sum_{n=1}^{\lceil \frac{N}{2}\rceil} P_{1fg}(\tau > 2n)\mu_{2n-1}\right) \\ &= E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) + \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{\lceil \frac{N}{2}\rceil} P_{1fg}(\tau > 2n)\mu_{2n-1}, \qquad (4.2) \end{aligned}$$

denn $E_{1fg}(R(\tau-1)1_{\{\tau\leq n+1\}}) \to E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) \in \mathbb{R}$.

Den verbleibenden liminf wollen wir mit Hilfe des folgenden Lemmas umformulieren.

Lemma 4.1 *Ist $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge mit Grenzwert $a \geq 0$ und $(b_n)_{n\in\mathbb{N}}$ eine beschränkte Folge, so gilt*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a_n b_n = a \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} b_n.$$

Beweis: Zu jedem $\varepsilon > 0$ gibt es ein n_ε, sodass $a \leq a_n \leq a + \varepsilon$ für alle $n \geq n_\varepsilon$. Wir erhalten für $N \geq n_\varepsilon$

$$\frac{1}{N} \sum_{n=1}^{N} a_n b_n \leq \frac{1}{N} \sum_{n=1}^{n_\varepsilon - 1} a_n b_n + \frac{1}{N} \sum_{n=n_\varepsilon}^{N} (a + \varepsilon_n) b_n,$$

wobei $\varepsilon_n := \varepsilon$, falls $b_n \geq 0$, und $\varepsilon_n := 0$ sonst. Somit folgt

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a_n b_n \leq \liminf_{N\to\infty} \left(\frac{1}{N} \sum_{n=1}^{N} a b_n + \frac{1}{N} \sum_{n_\varepsilon \leq n \leq N:\, b_n \geq 0} \varepsilon b_n \right) \tag{4.3}$$

$$\leq a \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} b_n + K\varepsilon,$$

da der zweite Summand in (4.3) nicht-negativ und durch $K\varepsilon$ beschränkt ist, wenn $K := \sup_{n\in\mathbb{N}} |b_n|$. Bei $\varepsilon \to 0$ ergibt sich $\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a_n b_n \leq a \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} b_n$, was die eine Hälfte der Behauptung ist. Der „$\geq$“-Teil läuft völlig analog. □

Nun kann Lemma 4.1 auf unser Problem wie folgt angewendet werden: Für $n \in \mathbb{N}$ sei $a_n := P_{1fg}(\tau > n+1)$, sodass $a := \lim_{n\to\infty} a_n = P_{1fg}(\tau = \infty)$. Weiter sei $b_n := \mu_n$ für ungerade n (und $b_n := 0$ für gerade n). Wegen $|b_n| \leq 2$ für alle n gilt also nach (4.2)

$$\gamma_T(1,f,g) = E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) + \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{\lceil \frac{N}{2} \rceil} P_{1fg}(\tau > 2n)\mu_{2n-1}$$

$$= E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) + P_{1fg}(\tau = \infty) \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{\lceil \frac{N}{2} \rceil} \mu_{2n-1}$$

$$= E_{1fg}(R(\tau-1)1_{\{\tau<\infty\}}) + P_{1fg}(\tau = \infty) \liminf_{N\to\infty} \left(1 - \frac{2}{N} \sum_{n=1}^{\lceil \frac{N}{2} \rceil} g^{(2n-1)}\right)$$

$$= \tilde{\gamma}_T(1,f,g)$$

nach (4.1). Damit ist gezeigt:

Lemma 4.2 *Für das Bad Match stimmen für alle Markow-Strategien f, g die klassische und die alternative Totalauszahlung überein: $\tilde{\gamma}_T(1, f, g) = \gamma_T(1, f, g)$.*

Wie in Kapitel 2 (Anmerkung 2) erwähnt verdeutlicht dieses Lemma erneut, dass die Gleichheit der beiden Versionen der Totalauszahlung auch unter schwächeren als den in Lemma 2.2 genannten Bedingungen gelten kann. (Hier gilt zwar die Bedingung (b) dieses Lemmas, doch wird nicht verlangt, dass P_{1fg}-fast sicher $\tau < \infty$ ist.)

4.2 $\tilde{\gamma}_T$-Optimalität im Bad Match

Um zu zeigen, dass der alternative totale Spielwert existiert und gleich Null ist (wie der klassische totale Spielwert), verwenden wir dieselbe Methode wie Thuijsman und Vrieze (vgl. [TV87, Thu92]).

Für $N \in \mathbb{N}$ betrachten wir die wie folgt definierte Strategie π_N von Spieler 1: Es sei $K(n) := \sum_{m=1}^{n}(1_{\{A^2(2m-1)=2\}} - 1_{\{A^2(2m-1)=1\}})$ für $n \geq 0$. Befindet sich das Spiel in Runde $2n+1$ in Zustand 1 ($n \geq 0$), so wähle Spieler 1 Aktion 2 mit der Wahrscheinlichkeit $(K(n)(\omega) + N + 1)^{-2}$.
Das Ziel ist zu zeigen, dass π_N Spieler 1 eine alternative Totalauszahlung von $-\frac{1}{N+1}$ garantiert. Sei $\sigma \in \Sigma$ beliebig und wie zuvor $\tau = \inf\{n\colon Z(n) \in S_{abs}\}$. τ ist niemals ungerade. In [TV87] wird gezeigt, dass für alle $m, N \in \mathbb{N}$

$$P_{1\pi_N\sigma}\big(\{\tau \geq 2m+2\} \cup \{\tau \leq 2m,\ A^2(\tau-1) = 1\}\big) \geq \frac{N}{2(N+1)}. \qquad (4.4)$$

Links steht die Wahrscheinlichkeit für das für Spieler 1 günstige Ereignis, dass bis einschl. Runde $2m-1$ in Zustand 1 noch *nicht* Reihe 2 gespielt wurde oder aber Spieler 2 in diesem Moment seine *erste* Aktion wählte. Auf den technischen, aber unschwierigen Beweis von (4.4) wollen wir hier verzichten.
Im Fall $P_{1\pi_N\sigma}(\tau = \infty) = 0$ strebt die linke Seite von (4.4) für $m \to \infty$ gegen $P_{1\pi_N\sigma}(A^2(\tau-1) = 1)$. Dann ergibt sich

$$\begin{aligned}\tilde{\gamma}_T(1, \pi_N, \sigma) &= P_{1\pi_N\sigma}(A^2(\tau-1) = 1) - P_{1\pi_N\sigma}(A^2(\tau-1) = 2)\\ &= 2P_{1\pi_N\sigma}(A^2(\tau-1) = 1) - 1 \geq \frac{N}{N+1} - 1 = -\frac{1}{N+1}.\end{aligned}$$

Andernfalls setzen wir $\lambda(m) := P_{1\pi_N\sigma}(\tau \leq 2m,\ A^2(\tau-1) = 1)$ und $\mu(m) := P_{1\pi_N\sigma}(\tau \leq 2m,\ A^2(\tau-1) = 2)$, sodass $\lambda := \lim_{m\to\infty}\lambda(m)$ (bzw. $\mu := \lim_{m\to\infty}\mu(m)$) die Wahrscheinlichkeit dafür ist, dass Spieler 1 jemals

Aktion 2 wählt und Spieler 2 sich in dieser Runde für Aktion 1 (bzw. 2) entscheidet. Also ist $E_{1\pi_N\sigma}(\liminf_{N\to\infty}\overline{S}(N)1_{\{\tau<\infty\}}) = \lambda-\mu$ und daher

$$\tilde{\gamma}_T(1,\pi_N,\sigma) = \lambda-\mu+E_{1\pi_N\sigma}(\liminf_{N\to\infty}\overline{S}(N)1_{\{\tau=\infty\}}).$$

Der letzte Term ist nicht-negativ, da im Fall $\tau(\omega)=\infty$ nach Definition von π_N $K(n)(\omega) > -N$ für alle $n\in\mathbb{N}$ ist, also $\sum_{m=1}^n S(m)(\omega) > -2N$ für jedes $n\in\mathbb{N}$ und so $\liminf_{n\to\infty}\overline{S}(n)(\omega)\geq 0$.
Es bleibt zu zeigen, dass $\lambda-\mu\geq-\frac{1}{N+1}$. Wir betrachten die Strategie σ_m von Spieler 2, die bis zum m-ten Besuch in Zustand 1 (Runde $2m-1$) mit σ übereinstimmt und dort ab Runde $2m+1$ stets die gemischte Aktion $(\frac{1}{2},\frac{1}{2})$ benutzt. Wird σ_m gespielt, so kann $\{\tau=\infty\}=\{K(n)>-N\ \forall n,\ \tau=\infty\}$ nur Nullmaß besitzen. Wie im ersten Fall gilt also $\tilde{\gamma}_T(1,\pi_N,\sigma_m)\geq-\frac{1}{N+1}$ für alle m. Nun ist zu beachten, dass $\tilde{\gamma}_T(1,\pi_N,\sigma_m)$ gleich $\lambda(m)-\mu(m)$ ist, da auf der Menge $\{\tau>2m\}$ der erwartete Gewinn Null beträgt. Also ist $\lambda(m)-\mu(m)\geq-\frac{1}{N+1}$ für alle m und mit $m\to\infty$ ergibt sich schließlich $\lambda-\mu\geq-\frac{1}{N+1}$ wie zu zeigen war. Somit gilt:

Lemma 4.3 *Für das Bad Match ist* $\tilde{v}_T^{low}(1)=\sup_\pi\inf_\sigma\tilde{\gamma}_T(1,\pi,\sigma)\geq 0$.

Nun können wir unser Hauptresultat formulieren:

Satz 4.4 *Für das Bad Match existiert der alternative totale Spielwert $\tilde{v}_T$ und es gilt $\tilde{v}_T(1)=0$. Bezüglich $\tilde{\gamma}_T$ verfügt Spieler 1 über keine optimale Strategie und keine ε-optimale Markow-Strategie (für alle $\varepsilon\in[0,1[$).*

Beweis: Lemma 4.3 liefert $\tilde{v}_T^{low}(1)\geq 0$. Wegen $|S(n)|\leq 2\ \forall n$ gilt

$$\gamma_T(1,\pi,\sigma)=\liminf_{N\to\infty}E_{1\pi\sigma}(\overline{S}(N))\geq E_{1\pi\sigma}(\liminf_{N\to\infty}\overline{S}(N))=\tilde{\gamma}_T(1,\pi,\sigma) \quad (4.5)$$

für alle $(\pi,\sigma)\in\Pi\times\Sigma$. Also ist $\tilde{v}_T^{up}(1)\leq v_T^{up}(1)=0$, wobei die hintere Gleichheit in [TV87] zu finden ist. Somit existiert $\tilde{v}_T(1)$ und ist gleich Null. Weil Spieler 1 keine γ_T-optimale Strategie besitzt (siehe [TV87]), also für alle π ein σ mit $\gamma_T(1,\pi,\sigma)<0$ existiert, folgt mit (4.5) $\tilde{\gamma}_T(1,\pi,\sigma)<0$. Somit hat Spieler 1 auch keine $\tilde{\gamma}_T$-optimale Strategie. Weil Spieler 1 (für $\varepsilon\in[0,1[$) auch über keine bzgl. γ_T ε-optimale Markow-Strategie verfügt (siehe wieder [TV87]), besitzt er nach Lemma 4.2 also auch keine bzgl. $\tilde{\gamma}_T$ ε-optimale Markow-Strategie. □

Kapitel 5

Stationäre Strategien

In diesem Kapitel sollen stationäre Strategien genauer untersucht werden. Das Ziel ist die Angabe einer expliziten Formel für die alternative Totalauszahlung unter einem Paar stationärer Strategien. Um dieses Resultat herleiten zu können, beschäftigen wir uns zunächst allgemein mit homogenen Markowketten.

5.1 Funktionale auf homogenen Markowketten

Wir betrachten eine homogene Markowkette $\{Z_n, n \geq 1\}$ auf einem W-Raum $(\Omega, \mathcal{A}, P)$. Der Zustandsraum S der Kette sei endlich. Das zum Startzustand $s \in S$ gehörende W-Maß bezeichnen wir mit P_s. Ohne Einschränkung nehmen wir an, dass die Markowkette aus einer einzigen rekurrenten Klasse von Zuständen besteht, denn die in den Abschnitten über Markowketten behandelten Fragestellungen setzen stets die Wahl eines festen rekurrenten Startzustands voraus. $\pi = (\pi_t)_{t \in S}$ stehe für das invariante Wahrscheinlichkeitsmaß auf S. Man beachte, dass $\pi_t > 0 \; \forall t \in S$.
Zu einer Auszahlungsfunktion $f : S \to \mathbb{R}$ setzen wir $S_n := \sum_{i=1}^n f(Z_i)$ und nennen S_n den *Besitz* oder das *Kapital* zum Zeitpunkt n. In [Chu67] kann man umfangreiche Resultate über die Asymptotik von S_n finden. Für unsere Zwecke benötigen wir jedoch Aussagen über die Asymptotik von $\overline{S}_N := \frac{1}{N} \sum_{n=1}^N S_n$. Um deren Herleitung soll es im Folgenden gehen.

Für einen beliebigen Zustand $s \in S$ bezeichnen wir mit $\tau_k(s)$ den Zeitpunkt des k-ten Besuchs in s. Formal sei also $\tau_1(s) := \inf\{n \in \mathbb{N} : Z_n = s\}$ und analog $\tau_k(s) := \inf\{n > \tau_{k-1}(s) : Z_n = s\}$ für $k > 1$. Wegen der Rekurrenzeigenschaft und der Endlichkeit von S bricht die Folge $(\tau_k(s))_{k \in \mathbb{N}}$

nie ab, d.h. es gilt fast sicher $\tau_k(s) < \infty$ für alle k. Wir setzen weiter $\rho_k(s) := \tau_{k+1}(s) - \tau_k(s)$ und $Y_k(s) := \sum_{n=\tau_k(s)}^{\tau_{k+1}(s)-1} f(Z_n)$. Die Größe $Y_k(s)$ beschreibt dann den Zugewinn im k-ten Rückkehrintervall von s. Offensichtlich sind $(\rho_k(s))_{k\in\mathbb{N}}$ und $(Y_k(s))_{k\in\mathbb{N}}$ iid-Variablen. Ferner sind die Verteilungen von $Y_1(s)$ und $\rho_1(s)$ unabhängig vom Startzustand. Daher verzichten wir für diese Größen auf die Indizierung des W-Maßes durch den Startzustand und schreiben etwa nur $E(Y_1(s))$ oder $P(\rho_1(s) > n)$.

Lemma 5.1 $E(Y_1(s)^m)$ *existiert für alle* $s \in S$ *und* $m \in \mathbb{N}$.

Beweis: Weil S endlich und die Kette rekurrent ist, existieren für $s \in S$ bekanntlich alle Momente von $\rho_1(s)$ (siehe z.B. [Chu67]). Daher gilt für alle $m \in \mathbb{N}$ $E(|Y_1(s)|^m) \leq E((\sum_{n=\tau_1(s)}^{\tau_2(s)-1} |f(Z_n)|)^m) \leq E((M\rho_1(s))^m) < \infty$, wobei $M := \max_{s\in S} |f(s)|$. □

Das nächste Lemma bestätigt folgende Überlegung: Der erwartete Gewinn zwischen zwei Besuchen in s ist gerade das Produkt aus der erwarteten Länge dieses Rückkehrintervalls und dem (im Limes) durchschnittlichen erwarteten Gewinn einer einzelnen Runde.

Lemma 5.2 *Es sei* $s \in S$. *Dann gilt* $\pi_s E(Y_1(s)) = \sum_{t\in S} f(t)\pi_t$. *Dies ist für jeden Startzustand* $s_1 \in S$ *gleich* $\lim_{n\to\infty} \frac{1}{n} E_{s_1}(S_n)$.

Beweis: Unter Beachtung von $Z(\tau_1(s)) = s$ kann man $Y_1(s)$ umformulieren: $Y_1(s) = f(s) + \sum_{n=1}^{\rho_1(s)-1} f(Z_{\tau_1(s)+n}) = f(s) + \sum_{n=1}^{\infty} f(Z_{\tau_1(s)+n}) 1_{\{\rho_1(s)>n\}}$. Da für $N \in \mathbb{N}$ $|\sum_{n=1}^{N} f(Z_{\tau_1(s)+n}) 1_{\{\rho_1(s)>n\}}| \leq \sum_{n=1}^{N} |f(Z_{\tau_1(s)+n})| 1_{\{\rho_1(s)>n\}} \leq M\rho_1(s) \in \mathcal{L}^1(P)$ gilt, sind Erwartungswert- und Limesbildung vertauschbar:

$$\begin{aligned}
E(Y_1(s)) &= f(s) + \sum_{n=1}^{\infty} E(f(Z_{\tau_1(s)+n}) 1_{\{\rho_1(s)>n\}}) \\
&= f(s) + \sum_{n=1}^{\infty} E\Bigg(\sum_{t\in S\setminus\{s\}} f(t) 1_{\{Z_{\tau_1(s)+n}=t,\, \rho_1(s)>n\}} \Bigg) \\
&= f(s) + \sum_{t\in S\setminus\{s\}} f(t) \sum_{n=1}^{\infty} P(Z_{\tau_1(s)+n} = t,\, \rho_1(s) > n) \\
&= f(s) + \sum_{t\in S\setminus\{s\}} f(t) \sum_{n=1}^{\infty} {}_s p_{st}^n.
\end{aligned}$$

Dabei ist ${}_s p_{st}^n$ die Wahrscheinlichkeit dafür, bei Start in s nach genau n Schritten Zustand t zu besuchen, ohne zuvor noch einmal s durchlaufen zu haben. Die Summe $\sum_{n=1}^{\infty} {}_s p_{st}^n$ ist dann die erwartete Anzahl der Besuche in

t vor der ersten Rückkehr nach s, wenn man in s startet. Dieser Erwartungswert ist bekanntlich gleich $\frac{\pi_t}{\pi_s}$ (siehe z.B. [Chu67]). Somit ist

$$E(Y_1(s)) = f(s) + \sum_{t \in S \setminus \{s\}} f(t) \frac{\pi_t}{\pi_s} = \frac{1}{\pi_s} \sum_{t \in S} f(t) \pi_t.$$

Ferner gilt für einen beliebigen Startzustand $s_1 \in S$

$$\begin{aligned} \frac{1}{n} E_{s_1}(S_n) &= \frac{1}{n} \sum_{i=1}^{n} E_{s_1}(f(Z_i)) = \frac{1}{n} \sum_{i=1}^{n} \sum_{t \in S} f(t) P_{s_1}(Z_i = t) \\ &= \sum_{t \in S} f(t) \frac{1}{n} \sum_{i=1}^{n} P_{s_1}(Z_i = t) \longrightarrow \sum_{t \in S} f(t) \pi_t \end{aligned}$$

für $n \to \infty$.Dies war zu zeigen. □

Korollar 5.3 *Die Größe $\pi_s E(Y_1(s))$ ist eine von s unabhängige Konstante. Insbesondere gilt: Ist für ein $s \in S$ $E(Y_1(s)) = 0$, so gilt $E(Y_1(t)) = 0$ für alle $t \in S$.*

Der *asymmetrische* Fall $\sum_{t \in S} f(t)\pi_t \neq 0$ ist uninteressant, denn dann gilt für $n \to \infty$ fast sicher $S_n \to +\infty$, falls $\sum_{t \in S} f(t)\pi_t > 0$, und analog $S_n \to -\infty$, falls $\sum_{t \in S} f(t)\pi_t < 0$, wie man sehr schnell zeigen kann (siehe etwa [Chu67]). Dasselbe gilt dann auch für $\overline{S}_N$ bei $N \to \infty$.
Uns wird im Folgenden der *symmetrische* Fall $\sum_{t \in S} f(t)\pi_t = 0$ beschäftigen. Wenn wir diese Voraussetzung stellen, bedeutet dies nach Korollar 5.3 also $E(Y_1(s)) = 0$ für *alle* $s \in S$. Zunächst betrachten wir eine „degenerierte" Situation:

Lemma 5.4 *Es sei $s \in S$. Gilt $Y_1(s) = 0$ fast sicher, so gilt auch $Y_1(t) = 0$ fast sicher für alle $t \in S$.*

Beweis: Sei $t \neq s$ beliebig. Weil S aus einer einzigen rekurrenten Klasse besteht, gibt es $n, m \geq 2$ und Folgen $s = u_1, u_2, \ldots, u_n = t$ und $t = v_0, v_1, \ldots, v_m = s$ aus S (mit $u_i \neq s$ für $1 < i < n$ und $v_i \neq s$ für $0 < i < m$), die positives Maß besitzen. Also gilt $F := \sum_{i=1}^{n-1} f(u_i) + \sum_{i=0}^{m-1} f(v_i) = 0$.
Sei nun $t = w_0, w_1, \ldots, w_l = t$ irgendeine Folge von Zuständen (mit $w_i \neq t$ für $0 < i < l$), die ebenfalls positives Maß besitzt. Dann hat auch die kombinierte Folge $s = u_1, \ldots, u_n = t = w_0, w_1, \ldots, w_l = t = v_0, v_1, \ldots, v_m = s$ positives Maß, woraus wieder $F + \sum_{i=0}^{l-1} f(w_i) = 0$ und damit $\sum_{i=0}^{l-1} f(w_i) = 0$ folgt. Somit ist auch $Y_1(t)$ fast sicher gleich Null. □

Korollar 5.5 *Unter der Voraussetzung von Lemma 5.4 ist S_{n-1} für $n \in \mathbb{N}$ $\sigma(Z_1, Z_n)$-messbar, das heißt es existieren Konstanten $\{c_s(t) : s, t \in S\}$, sodass $S_{n-1} = \sum_{s,t \in S} c_s(t) 1_{\{Z_1 = s,\, Z_n = t\}}$.*

Beweis: Nach Lemma 5.4 sind die $Y_1(t)$ fast sicher gleich Null. Für zwei Besuche $\tau_k(t)$ und $\tau_l(t)$ im Zustand t (mit $k < l$) gilt also $0 = \sum_{i=k}^{l-1} Y_i(t) = S_{\tau_l(t)-1} - S_{\tau_k(t)-1}$, d.h. das Kapital bei Besuchen in t ist konstant. Diese Konstante kann höchstens vom Startzustand s abhängen und werde mit $c_s(t)$ bezeichnet. □

Es ist noch anzumerken, dass wegen $S_n = S_{n-1} + f(Z_n)$ damit auch S_n $\sigma(Z_1, Z_n)$-messbar ist. Dieser Umstand wird allerdings nicht benötigt.
Lemma 5.4 rechtfertigt es, im Folgenden von einer degenerierten *Kette* zu sprechen, falls $Y_1(s) = 0$ fast sicher für ein (und damit für alle) $s \in S$ gilt. Für festes $s \in S$ kann $c_s(.)$ auch als Funktion $c_s \colon S \to \mathbb{R}$ aufgefasst werden. Nun können wir die Asymptotik von $\overline{S}_N$ im degenerierten Fall mühelos herleiten:

Satz 5.6 *Es sei $s_1 \in S$ ein Startzustand. Ist die Kette degeneriert, so gilt $\lim_{N\to\infty} \overline{S}_N = \sum_{t\in S} c_{s_1}(t)\pi_t$ P_{s_1}-fast sicher*

Beweis: Nach Lemma 5.5 gilt $S_n = \sum_{t\in S} c_{s_1}(t) 1_{\{Z_{n+1}=t\}}$. Zu einer Realisierung $\omega = (\omega_1, \omega_2, \ldots)$ der Markowkette setzen wir nun $g(\omega) := c_{s_1}(Z_2(\omega))$, sodass für $n \in \mathbb{N}_0$ $g(\sigma^n \omega) = c_{s_1}(Z_{n+2}(\omega))$ gilt, wenn σ der übliche Linksshift ist. Da σ maßtreu ist, ergibt sich

$$\begin{aligned}\overline{S}_N(\omega) &= \frac{1}{N} \sum_{n=1}^{N} \sum_{t\in S} c_{s_1}(t) 1_{\{Z_{n+1}(\omega)=t\}} = \frac{1}{N} \sum_{n=1}^{N} c_{s_1}(Z_{n+1}(\omega)) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} g(\sigma^n \omega) \longrightarrow \sum_{t\in S} c_{s_1}(t)\pi_t\end{aligned}$$

P_{s_1}-fast sicher nach dem Ergodensatz. □

Als nächstes wollen wir die Asymptotik von $\overline{S}_N$ im Fall untersuchen, dass die Kette nicht degeneriert ist, d.h. dass die $Y_1(s)$ nicht fast sicher gleich Null sind. Dies wird wesentlich mehr Aufwand erfordern als im soeben behandelten degenerierten Fall.

5.2 Vorbereitungen

Der akkumulierte Gewinn zum Zeitpunkt des K-ten Besuchs im Startzustand s_1 beträgt gerade $S_{\tau_K(s_1)-1}$. Zunächst betrachten wir $\overline{S}_N$ für die Teilfolge $N_K := \tau_K(s_1) - 1$, d.h. wir mitteln die S_n bis zum K-ten Besuch in s_1. Für die so gewählten Zeitpunkte erhält man die Zerlegung

$$\sum_{n=1}^{\tau_K(s_1)-1} S_n = \sum_{k=1}^{K-1} \rho_k(s_1) S_{\tau_k(s_1)-1} + \sum_{k=1}^{K-1} \sum_{j=1}^{\rho_k(s_1)} \left(S_{\tau_k(s_1)-1+j} - S_{\tau_k(s_1)-1}\right)$$

$$
\begin{aligned}
&= \sum_{k=1}^{K-1} \sum_{j=1}^{k-1} Y_j(s_1)\rho_k(s_1) + \sum_{k=1}^{K-1} X_k(s_1) \\
&= \sum_{k=1}^{K-1} \Big(\rho_k(s_1) \sum_{j=1}^{k-1} Y_j(s_1) + X_k(s_1) \Big), \qquad (5.1)
\end{aligned}
$$

wenn man $X_k(s_1) := \sum_{j=1}^{\rho_k(s_1)} (S_{\tau_k(s_1)-1+j} - S_{\tau_k(s_1)-1})$ setzt. Entscheidend ist nun, dass die $\psi_k(s_1) := (\rho_k(s_1), Y_k(s_1), X_k(s_1))$ für $k \in \mathbb{N}$ eine iid-Folge bilden. Das Tripel $\psi_k(s_1)$ gibt die *Länge* des k-ten Rückkehrintervalls, den darin erzielten *Gesamtgewinn* und die Summe der *Zugewinne* in diesem Zeitraum an.

Satz 5.7 *Es sei $s_1 \in S$ ein Startzustand. Dann ist* $\liminf_{K\to\infty} \overline{S}_{\tau_K(s_1)-1}$ *P_{s_1}-fast sicher konstant.*

Beweis: In diesem Beweis lassen wir das Argument s_1 an den Zufallsvariablen τ_k, ρ_k, Y_k und X_k der Einfachheit halber weg. Sei λ eine endliche Permutation von $\mathbb{N}$ mit $m := \max\{n \colon \lambda(n) \neq n\}$ und $K > m+1$. Dann gilt:

$$
\begin{aligned}
&\sum_{k=1}^{K-1} \Big(\rho_{\lambda(k)} \sum_{j=1}^{k-1} Y_{\lambda(j)} + X_{\lambda(k)} \Big) \\
&= \sum_{k=1}^{m} \rho_{\lambda(k)} \sum_{j=1}^{k-1} Y_{\lambda(j)} + \sum_{k=m+1}^{K-1} \rho_{\lambda(k)} \Big(\sum_{j=1}^{m} Y_{\lambda(j)} + \sum_{j=m+1}^{k-1} Y_{\lambda(j)} \Big) + \sum_{k=1}^{K-1} X_{\lambda(k)} \\
&= \sum_{k=1}^{m} \rho_{\lambda(k)} \sum_{j=1}^{k-1} Y_{\lambda(j)} + \sum_{k=m+1}^{K-1} \rho_{k} \Big(\sum_{j=1}^{m} Y_{j} + \sum_{j=m+1}^{k-1} Y_{j} \Big) + \sum_{k=1}^{K-1} X_{k} \\
&= \sum_{n=1}^{\tau_K-1} S_n + \sum_{k=1}^{m} \Big(\rho_{\lambda(k)} \sum_{j=1}^{k-1} Y_{\lambda(j)} - \rho_k \sum_{j=1}^{k-1} Y_j \Big). \qquad (5.2)
\end{aligned}
$$

Es sei $h_K(\psi) := \big(\sum_{k=1}^{K-1} \rho_k \big)^{-1} \sum_{k=1}^{K-1} \big(\rho_k \sum_{j=1}^{k-1} Y_j + X_k \big)$ für $\psi := (\psi_k)_{k\in\mathbb{N}}$. Da die in (5.2) hinten stehende Summe nicht mehr von K abhängt und fast sicher endlich ist, folgt

$$
\begin{aligned}
&\liminf_{K\to\infty} h_K(\lambda\psi) \\
&= \liminf_{K\to\infty} \frac{\sum_{n=1}^{\tau_K-1} S_n + \sum_{k=1}^{m} \big(\rho_{\lambda(k)} \sum_{j=1}^{k-1} Y_{\lambda(j)} - \rho_k \sum_{j=1}^{k-1} Y_j \big)}{\sum_{k=1}^{K-1} \rho_{\lambda(k)}} \\
&= \liminf_{K\to\infty} \frac{\sum_{n=1}^{\tau_K-1} S_n}{\sum_{k=1}^{K-1} \rho_k} = \liminf_{K\to\infty} h_K(\psi).
\end{aligned}
$$

Somit ist $\liminf_{K\to\infty} h_K$ endlich permutierbar bezüglich ψ. Fast sicher ist also $\liminf_{K\to\infty} h_K(\psi) = \liminf_{K\to\infty} \overline{S}_{\tau_K-1}$ konstant. □

Unser Ziel ist eine Aussage über die asymptotische Verteilung von $\overline{S}_N$, also des mittleren Besitzes bis zum Zeitpunkt N. In einem ersten Schritt wird nur der mittlere Besitz bei Besuchen im Startzustand s_1 betrachtet.

Lemma 5.8 *Es sei $s_1 \in S$ ein Startzustand und $\sum_{t\in S} f(t)\pi_t = 0$. Ist die Kette nicht degeneriert, so gilt*

$$\mathcal{L}\left(\frac{1}{\sigma\sqrt{K}}\frac{1}{K}\sum_{k=1}^{K} S_{\tau_k(s_1)-1}\middle| P_{s_1}\right) \longrightarrow \mathcal{N}(0,1)$$

für $K\to\infty$, wobei $\sigma := \sqrt{\frac{E(Y_1(s_1)^2)}{3}}$.

Beweis: Weil nach Voraussetzung $Y_1(s_1)$ nicht fast sicher konstant ist, gilt $\sigma > 0$. Wir setzen $\xi_{Kk} := \frac{K-k}{\sigma K\sqrt{K}} Y_k(s_1)$ für $K \in \mathbb{N}$ und $1 \le k < K$. Dann ist $\sum_{k=1}^{K-1}\xi_{Kk} = \frac{1}{\sigma K\sqrt{K}}\sum_{k=1}^{K-1}(K-k)Y_k(s_1) = \frac{1}{\sigma K\sqrt{K}}\sum_{k=1}^{K} S_{\tau_k(s_1)-1}$. Wegen Lemma 5.1 und Lemma 5.2 bilden die ξ_{Kk} ein quadratintegrierbares, zentriertes zeilenunabhängiges Schema. Es gilt:

$$\begin{aligned}\sum_{k=1}^{K-1}E(\xi_{Kk}^2) &= \frac{1}{\sigma^2K^3}\sum_{k=1}^{K-1}(K-k)^2E(Y_k(s_1)^2)\\ &= \frac{E(Y_1(s_1)^2)}{\sigma^2}\frac{(K-1)K(2K-1)}{6K^3} \longrightarrow \frac{E(Y_1(s_1)^2)}{3\sigma^2} = 1\end{aligned}$$

für $K\to\infty$, wenn σ gewählt wird wie oben. Seien nun $\varepsilon,\delta>0$ beliebig und K so groß, dass $P(|Y_1(s_1)| > \delta\sigma\sqrt{K}) \le \varepsilon^2$. Dann ist

$$\begin{aligned}&\sum_{k=1}^{K-1}E(\xi_{Kk}^2 1_{\{|\xi_{Kk}|>\delta\}}) = \sum_{k=1}^{K-1}E\left(\frac{(K-k)^2}{\sigma^2K^3}Y_k(s_1)^2 1_{\{|Y_k(s_1)|>\frac{\delta\sigma K\sqrt{K}}{K-k}\}}\right)\\ &\le \sum_{k=1}^{K-1}\frac{(K-k)^2}{\sigma^2K^3}\sqrt{E(Y_k(s_1)^4)P\left(|Y_k(s_1)| > \frac{\delta\sigma K\sqrt{K}}{K-k}\right)}\\ &\le \frac{\sqrt{E(Y_1(s_1)^4)}}{\sigma^2K^3}\sum_{k=1}^{K-1}(K-k)^2\sqrt{\varepsilon^2}\\ &= \frac{\varepsilon\sqrt{E(Y_1(s_1)^4)}}{\sigma^2}\frac{(K-1)K(2K-1)}{6K^3} \le \varepsilon\frac{\sqrt{E(Y_1(s_1)^4)}}{3\sigma^2}.\end{aligned}$$

Da $\varepsilon>0$ beliebig war, folgt $\sum_{k=1}^{K-1}E(\xi_{Kk}^2 1_{\{|\xi_{Kk}|>\delta\}}) \to 0$ für $K\to\infty$. Der Zentrale Grenzwertsatz liefert nun die Behauptung. □

Im nächsten Schritt stellen wir eine Beziehung zwischen dem soeben betrachteten und einem verwandten gewichteten Mittel her:

Lemma 5.9 *Es sei $s_1 \in S$ ein Startzustand und $\sum_{t \in S} f(t)\pi_t = 0$. Ist die Kette nicht degeneriert, so existiert für alle $\delta > 0$ ein $\alpha > 0$, sodass für alle $K \in \mathbb{N}$*

$$P_{s_1}\left(\left|\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1}\rho_k(s_1)S_{\tau_k(s_1)-1} - \frac{1}{K}\sum_{k=1}^{K-1}S_{\tau_k(s_1)-1}\right| \geq \alpha\right) < \delta.$$

Beweis: Auch in diesem Beweis werden wir die Abhängigkeit von ρ_k und anderen Zufallsgrößen von s_1 der Kürze wegen unterdrücken.
Es sei $\delta > 0$ gegeben und $\alpha > 0$ zunächst beliebig. Wegen $\tau_1 \equiv 1$ ist die zu zeigende Aussage im Fall $K \leq 2$ trivial. Sei also $K > 2$. Da alle erforderlichen Momente existieren, können wir die Tschebyschev-Ungleichung anwenden:

$$\begin{aligned}
&P_{s_1}\left(\left|\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1}\rho_k S_{\tau_k-1} - \frac{1}{K}\sum_{k=1}^{K-1}S_{\tau_k-1}\right| \geq \alpha\right)\\
&\leq \frac{\pi_{s_1}^2}{\alpha^2}E_{s_1}\left(\left[\frac{1}{K}\sum_{k=1}^{K-1}(\rho_k - \pi_{s_1}^{-1})S_{\tau_k-1}\right]^2\right)\\
&= \frac{\pi_{s_1}^2}{K^2\alpha^2}\sum_{k=1}^{K-1}\sum_{l=1}^{K-1}E_{s_1}\big((\rho_k - \pi_{s_1}^{-1})S_{\tau_k-1}(\rho_l - \pi_{s_1}^{-1})S_{\tau_l-1}\big)\\
&= \frac{\pi_{s_1}^2}{K^2\alpha^2}\sum_{k=1}^{K-1}E_{s_1}\big((\rho_k - \pi_{s_1}^{-1})^2 S_{\tau_k-1}^2\big),
\end{aligned}$$

denn in den gemischten Termen ist für $k < l$ die Größe $\rho_l = \tau_{l+1} - \tau_l$ unabhängig von allen anderen Faktoren und ihre Erwartung beträgt gerade $\pi_{s_1}^{-1}$. Indem wir die Unabhängigkeit von ρ_k und S_{τ_k-1} ausnutzen, erhalten wir

$$\begin{aligned}
&= \frac{\pi_{s_1}^2\mathrm{Var}(\rho_1)}{K^2\alpha^2}\sum_{k=1}^{K-1}E_{s_1}(S_{\tau_k-1}^2)\\
&= \frac{\pi_{s_1}^2\mathrm{Var}(\rho_1)}{K^2\alpha^2}\sum_{k=1}^{K-1}\sum_{j=1}^{k-1}\mathrm{Var}(Y_j) = \frac{\pi_{s_1}^2\mathrm{Var}(\rho_1)}{K^2\alpha^2}\sum_{k=1}^{K-1}(k-1)\mathrm{Var}(Y_1)\\
&= \frac{\pi_{s_1}^2\mathrm{Var}(\rho_1)\mathrm{Var}(Y_1)}{\alpha^2}\frac{(K-2)(K-1)}{2K^2} < \frac{\mathrm{Var}(\rho_1)\mathrm{Var}(Y_1)}{2\alpha^2}.
\end{aligned}$$

Setzt man also $\alpha := \sqrt{\frac{\mathrm{Var}(\rho_1)\mathrm{Var}(Y_1)}{2\delta}}$, so ist die Behauptung bewiesen. □

5.3 Asymptotik des Cesaro-Mittels $\overline{S}_N$

Im nächsten Satz gewinnen wir schließlich eine Aussage über die Verteilung von $\overline{S}_N$ (zu den Zeitpunkten $N_K = \tau_K(s_1) - 1$) im nicht-degenerierten Fall.

Satz 5.10 *Es sei $s_1 \in S$ ein Startzustand und $\sum_{t\in S} f(t)\pi_t = 0$. Ist die Kette nicht degeneriert, so gilt für $K \to \infty$*

$$\mathcal{L}\left(\frac{\overline{S}_{\tau_K(s_1)-1}}{\sigma\sqrt{\pi_{s_1}}\sqrt{\tau_K(s_1)-1}} \middle| P_{s_1} \right) \longrightarrow \mathcal{N}(0,1).$$

Beweis: Erneut verkürzen wir die Notation durch Unterdrückung der Abhängigkeit der verschiedenen Zufallsvariablen von s_1. Es seien $c \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Man errechnet

$$\begin{aligned}
P_{s_1}\left(\frac{\overline{S}_{\tau_K-1}}{\sqrt{\tau_K-1}} < c\right) &= P_{s_1}\left(\frac{\pi_{s_1}}{K}\sum_{n=1}^{\tau_K-1} S_n < c\pi_{s_1}\frac{(\tau_K-1)^{\frac{3}{2}}}{K}\right) \\
= P_{s_1}&\left(\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1}\rho_k S_{\tau_k-1} + \frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1} X_k < c\pi_{s_1}\frac{(\tau_K-1)^{\frac{3}{2}}}{K}\right) \\
\geq P_{s_1}&\left(\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1}\rho_k S_{\tau_k-1} < c\pi_{s_1}\frac{(\tau_K-1)^{\frac{3}{2}}}{K} - \pi_{s_1}(E(X_1)+1)\right) \\
&- P_{s_1}\left(\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1} X_k > \pi_{s_1}(E(X_1)+1)\right)
\end{aligned}$$

nach Zerlegung (5.1). Wegen $\frac{1}{K}\sum_{k=1}^{K} X_k \to E(X_1)$ fast sicher ist die zweite Wahrscheinlichkeit für hinreichend große K kleiner als ε. Dann ist die letzte Zeile

$$\begin{aligned}
\geq P_{s_1}&\left(\frac{1}{K}\sum_{k=1}^{K-1} S_{\tau_k-1} < c\pi_{s_1}\frac{(\tau_K-1)^{\frac{3}{2}}}{K} - \pi_{s_1}(E(X_1)+1) - \alpha\right) \\
&- P_{s_1}\left(\left|\frac{\pi_{s_1}}{K}\sum_{k=1}^{K-1}\rho_k S_{\tau_k-1} - \frac{1}{K}\sum_{k=1}^{K-1} S_{\tau_k-1}\right| \geq \alpha\right) - \varepsilon,
\end{aligned}$$

wobei man unter Benutzung von Lemma 5.9 α so wählt, dass die zweite Wahrscheinlichkeit kleiner als ε ist (für alle K). Setzt man abkürzend $\beta := \pi_{s_1}(E(X_1)+1)+\alpha$, ist Obiges

$$\begin{aligned}
&\geq P_{s_1}\left(\frac{1}{\sigma(K-1)^{\frac{3}{2}}}\sum_{k=1}^{K-1} S_{\tau_k-1} < \frac{K}{\sigma(K-1)^{\frac{3}{2}}}\left(c\pi_{s_1}\frac{(\tau_K-1)^{\frac{3}{2}}}{K} - \beta\right)\right) - 2\varepsilon \\
&\geq P_{s_1}\left(\frac{1}{\sigma(K-1)^{\frac{3}{2}}}\sum_{k=1}^{K-1} S_{\tau_k-1} < \frac{c\pi_{s_1}}{\sigma}\frac{(\tau_K-1)^{\frac{3}{2}}}{(K-1)^{\frac{3}{2}}} - \varepsilon\right) - 2\varepsilon,
\end{aligned}$$

wenn K so groß ist, dass $\frac{K}{\sigma(K-1)^{\frac{3}{2}}}\beta < \varepsilon$. Ferner gilt $\frac{\tau_K - 1}{K-1} = \frac{1}{K-1}\sum_{k=1}^{K-1}\rho_k \to E(\rho_1) = \pi_{s_1}^{-1}$ fast sicher. Damit strebt die in der letzten Wahrscheinlichkeit rechts stehende Zufallsgröße fast sicher gegen $\frac{c\pi_{s_1}}{\sigma}(E(\rho_1))^{\frac{3}{2}} - \varepsilon$. Also ist die letzte Zeile

$$\begin{aligned}
&\geq P_{s_1}\left(\frac{1}{\sigma(K-1)^{\frac{3}{2}}}\sum_{k=1}^{K-1} S_{\tau_k - 1} < \frac{c}{\sigma\sqrt{\pi_{s_1}}} - \varepsilon - \varepsilon\right) \\
&\quad - P_{s_1}\left(\frac{c\pi_{s_1}}{\sigma}\frac{(\tau_K - 1)^{\frac{3}{2}}}{(K-1)^{\frac{3}{2}}} < \frac{c}{\sigma\sqrt{\pi_{s_1}}} - \varepsilon\right) - 2\varepsilon \\
&\geq P_{s_1}\left(\frac{1}{\sigma(K-1)^{\frac{3}{2}}}\sum_{k=1}^{K-1} S_{\tau_k - 1} < \frac{c}{\sigma\sqrt{\pi_{s_1}}} - 2\varepsilon\right) - 3\varepsilon
\end{aligned}$$

für hinreichend große K. Diese Wahrscheinlichkeit strebt nach Lemma 5.8 gegen $\Phi(\frac{c}{\sigma\sqrt{\pi_{s_1}}} - 2\varepsilon)$. Für hinreichend große K ist die letzte Zeile also größer gleich $\Phi(\frac{c}{\sigma\sqrt{\pi_{s_1}}} - 2\varepsilon) - 4\varepsilon$. Somit gilt für alle $\varepsilon > 0$ und hinreichend große K

$$P_{s_1}\left(\frac{\overline{S}_{\tau_K - 1}}{\sigma\sqrt{\pi_{s_1}}\sqrt{\tau_K - 1}} < c\right) \geq \Phi(c - 2\varepsilon) - 4\varepsilon.$$

Lässt man ε gegen Null gehen, folgt

$$\liminf_{K\to\infty} P_{s_1}\left(\frac{\overline{S}_{\tau_K - 1}}{\sigma\sqrt{\pi_{s_1}}\sqrt{\tau_K - 1}} < c\right) \geq \Phi(c).$$

Eine völlig symmetrische Argumentation liefert, dass der limsup dieser Folge kleiner gleich $\Phi(c)$ ist. Zusammen folgt die Behauptung. □

Korollar 5.11 *Es sei $s_1 \in S$ ein Startzustand und $\sum_{t\in S} f(t)\pi_t = 0$. Ist die Kette nicht degeneriert, so gilt $\liminf_{N\to\infty}\overline{S}_N = -\infty$ P_{s_1}-fast sicher.*

Beweis: Nach Satz 5.7 ist $\liminf_{K\to\infty}\overline{S}_{\tau_K(s_1)-1}$ P_{s_1}-fast sicher konstant. Aus Satz 5.10 folgt nun, dass diese Konstante $-\infty$ sein muss. Also gilt $\liminf_{N\to\infty}\overline{S}_N \leq \liminf_{K\to\infty}\overline{S}_{\tau_K(s_1)-1} = -\infty$ P_{s_1}-fast sicher. □

Weil die Folge $(\tau_K(s_1) - 1)_{K\in\mathbb{N}}$ nicht abbricht, überrascht es nicht, dass aus Satz 5.10 sogar eine Aussage über die Verteilung von $\overline{S}_N$ gewonnen werden kann:

Satz 5.12 *Es sei $s_1 \in S$ ein Startzustand und $\sum_{t\in S} f(t)\pi_t = 0$. Ist die Kette nicht degeneriert, so gilt für $N \to \infty$*

$$\mathcal{L}\left(\frac{\overline{S}_N}{\sigma\sqrt{\pi_{s_1}}\sqrt{N}}\middle| P_{s_1}\right) \longrightarrow \mathcal{N}(0,1).$$

Beweis: Zum Startzustand s_1 und $N \in \mathbb{N}$ sei die Zufallsvariable $l_N(s_1)$ in Anlehnung an [Chu67] mittels $\tau_{l_N(s_1)}(s_1) \leq N < \tau_{l_N(s_1)+1}(s_1)$ definiert. Das Argument s_1 an dieser und anderen Größen lassen wir im Folgenden wieder weg. Wegen $\tau_1 \equiv 1$ ist $1 \leq l_N \leq N$ für alle N. Es gilt

$$\frac{\overline{S}_N}{\sqrt{N}} = \frac{\overline{S}_{\tau_{l_N}-1}}{\sqrt{\tau_{l_N}-1}} \left(\frac{\tau_{l_N}-1}{N}\right)^{3/2} + \frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} S_n. \tag{5.3}$$

Zuerst zeigen wir, dass $\frac{N-\tau_{l_N}+1}{\sqrt{N}}$ stochastisch gegen Null konvergiert. Für $m \in \mathbb{N}$ gilt

$$P_{s_1}(\tau_{l_N} \leq N-m) = P_{s_1}(Z_i \neq s_1 \text{ für } N-m < i \leq N) \leq \max_{t \in S} P_t(\tau_1 > m+1)$$

für alle $N \in \mathbb{N}$. Wegen $\sum_{m=0}^{\infty} P_t(\tau_1 > m) = E_t(\tau_1) < \infty$ für alle $t \in S$ strebt dieses Maximum für $m \to \infty$ gegen Null. Damit gilt für $m \to \infty$ auch $P_{s_1}(N - \tau_{l_N} \geq m) \to 0$ gleichmäßig in N. Wählt man also zu $\varepsilon, \delta > 0$ m_0 so groß, dass $P_{s_1}(N - \tau_{l_N} \geq m_0) < \varepsilon$ für alle N, und dazu N_0 so groß, dass $\delta\sqrt{N_0} - 1 \geq m_0$, so folgt sofort

$$P_{s_1}\left(\frac{N-\tau_{l_N}+1}{\sqrt{N}} > \delta\right) = P_{s_1}(N - \tau_{l_N} > \delta\sqrt{N} - 1) < \varepsilon \quad \forall N \geq N_0. \tag{5.4}$$

Betrachten wir nun Formel (5.3). Nach (5.4) gilt $\frac{\tau_{l_N}-1}{N} = 1 - \frac{N-\tau_{l_N}+1}{N} \to 1$ stochastisch. Wenn wir also zeigen können, dass der hintere Term in (5.3) stochastisch gegen Null strebt, so sind $\frac{\overline{S}_N}{\sqrt{N}}$ und $\frac{\overline{S}_{\tau_{l_N}-1}}{\sqrt{\tau_{l_N}-1}}$ asymptotisch identisch verteilt. Weil die zweite Größe nach Satz 5.10 in Verteilung gegen $\mathcal{N}(0, \sigma^2\pi_{s_1})$ konvergiert, gilt dies dann auch für $\frac{\overline{S}_N}{\sqrt{N}}$. Betrachten wir also den hinteren Summanden in (5.3):

$$\begin{aligned} \frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} S_n &= \frac{N-\tau_{l_N}+1}{\sqrt{N}N} S_{\tau_{l_N}-1} + \frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} (S_n - S_{\tau_{l_N}-1}) \\ &= \frac{N-\tau_{l_N}+1}{\sqrt{N}} \cdot \frac{l_N-1}{N} \cdot \frac{1}{l_N-1} \sum_{k=1}^{l_N-1} Y_k + \frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} (S_n - S_{\tau_{l_N}-1}). \end{aligned} \tag{5.5}$$

Es gilt $\frac{1}{l_N-1}\sum_{k=1}^{l_N-1} Y_k \to E(Y_1) = 0$ fast sicher unter Benutzung von Lemma 5.2. Ferner ist $\frac{l_N-1}{N} \leq 1$ für alle N und schließlich konvergiert $\frac{N-\tau_{l_N}+1}{\sqrt{N}}$ wie oben gesehen stochastisch gegen Null. Also strebt der erste Summand in (5.5) gegen Null. Für den zweiten Summanden errechnet man

$$\left|\frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} (S_n - S_{\tau_{l_N}-1})\right| \leq \frac{1}{\sqrt{N}N} \sum_{n=\tau_{l_N}}^{N} (n - \tau_{l_N} + 1)M$$

$$= \frac{M}{\sqrt{N}N} \sum_{n=1}^{N-\tau_{l_N}+1} n = \frac{(N-\tau_{l_N}+1)(N-\tau_{l_N}+2)M}{2\sqrt{N}N} \to 0$$

stochastisch wegen (5.4), wobei wieder $M = \max_{s\in S}|f(s)|$. Also konvergiert der hintere Term in (5.3) für $N \to \infty$ stochastisch gegen Null. Dies war zu zeigen. □

Damit haben wir insgesamt folgende Aussage über die Asymptotik von $\overline{S}_N$ im Fall einer aus einer einzigen rekurrenten Klasse bestehenden homogenen Markowkette bewiesen:

Satz 5.13 *Es sei $\{Z_n, n \geq 1\}$ eine homogene Markowkette mit endlichem Zustandsraum S, die nur aus einer rekurrenten Klasse besteht. Das invariante Maß auf S werde mit $(\pi_t)_{t\in S}$ bezeichnet. Zu einer Funktion $f\colon S \to \mathbb{R}$ sei $S_n = \sum_{i=1}^n f(Z_i)$ und $\overline{S}_N = \frac{1}{N}\sum_{n=1}^N S_n$. Für $s \in S$, $k \in \mathbb{N}$ sei $\tau_k(s)$ der Zeitpunkt des k-ten Besuchs in s und $Y_k(s) = \sum_{n=\tau_k(s)}^{\tau_{k+1}(s)-1} f(Z_n)$. Es sei $s_1 \in S$ ein Startzustand.*

(a) **Kette degeneriert:**
Ist $Y_1(s_1) = 0$ P_{s_1}-fast sicher, so gilt $S_n = \sum_{t\in S} c_{s_1}(t)1_{\{Z_{n+1}=t\}}$ für $n \in \mathbb{N}$ mit gewissen Konstanten $c_{s_1}(t)$ und

$$\overline{S}_N \to \sum_{t\in S} c_{s_1}(t)\pi_t$$

jeweils P_{s_1}-fast sicher.

(b) **Kette nicht degeneriert:**
Ist $Y_1(s_1) \neq 0$ mit positivem P_{s_1}-Maß und $\sum_{t\in S} f(t)\pi_t = 0$, so gilt

$$\mathcal{L}\left(\frac{\overline{S}_N}{\sqrt{\sigma^2 \pi_{s_1} N}}\,\middle|\, P_{s_1}\right) \to \mathcal{N}(0,1),$$

wobei $\sigma = \sqrt{\frac{E(Y_1(s_1)^2)}{3}}$. Es gilt $\liminf_{N\to\infty} \overline{S}_N = -\infty$ P_{s_1}-fast sicher.

(c) **Asymmetrischer Fall:**
Sonst (d.h. für $\sum_{t\in S} f(t)\pi_t \neq 0$) ist

$$\overline{S}_N \to \begin{cases} +\infty & \text{falls } \sum_{t\in S} f(t)\pi_t > 0, \\ -\infty & \text{falls } \sum_{t\in S} f(t)\pi_t < 0 \end{cases}$$

P_{s_1}-fast sicher.

5.4 Anwendung auf stochastische Spiele - Teil I

Nun wollen wir die im letzten Abschnitt gewonnenen Resultate auf stochastische Spiele anwenden. Es sei s_1 ein Startzustand und $(x, y) \in X \times Y$. Wie bisher bezeichnet $Z(n)$ den Zustand in Runde n und $S(n)$ den nach n Runden erzielten Gewinn von Spieler 1. Die Folge $(Z(n))_{n \in \mathbb{N}}$ ist aufgrund der Stationarität von (x, y) eine homogene Markowkette auf dem Zustandsraum S des Spiels. Das zugehörige invariante W-Maß auf S werde mit $(\pi_t)_{t \in S}$ bezeichnet.

Zunächst beschränken wir uns wie oben auf den Fall, dass S aus einer einzigen rekurrenten Klasse besteht. Die Resultate über homogene Markowketten sind jedoch auch unter dieser Voraussetzung nicht direkt anwendbar, weil die Auszahlungsfunktion r eines stochastischen Spiels nicht auf S selbst, sondern auf der Menge $\{(s, i, j) \colon s \in S,\ (i, j) \in A_s\}$ erklärt ist. Daher betrachten wir die zu (x, y) assoziierte Markowkette $\{(Z(n), A(n)), n \geq 1\}$ auf dem Raum $S^* := \{(s, i, j) \colon s \in S,\ (i, j) \in A_s,\ x_s(i) y_s(j) > 0\}$. Mit $(\pi^*_{t^*})_{t^* \in S^*}$ bezeichnen wir das zugehörige invariante Maß auf S^*. (Zur besseren Unterscheidung von dem nun auf S definierten W-Maß $(\pi_t)_{t \in S}$ verwenden wir dabei anders als in den letzten Abschnitten die $*$-Notation.) Für $t^* = (t, k, l)$ gilt offenbar $\pi^*_{t^*} = \pi_t x_t(k) y_t(l)$.

Setzen wir für $s^* = (s, i, j) \in S^*$ $f(s^*) := r(s, i, j)$, so befinden wir uns mit $S_n := \sum_{i=1}^{n} f((Z(i), A(i)))$ wieder in der Situation von Satz 5.13 – mit der Einschränkung, dass durch die Wahl eines Startzustands $s_1 \in S$ im stochastischen Spiel nicht eindeutig ein Startzustand $s^*_1 \in S^*$ für die assoziierte Markowkette bestimmt ist. Auf $\{Z(1) = s_1,\ A(1) = (k_1, l_1)\}$ gilt jedoch

$$
\begin{aligned}
S(n) &= R(1) + \sum_{i=2}^{n} R(i) \\
&= r(s_1, k_1, l_1) + \sum_{i=2}^{n} \sum_{(t,k,l) \in S^*} r(t, k, l) 1_{\{Z(i)=t,\ A(i)=(k,l)\}} \\
&= f((Z(1), A(1))) + \sum_{i=2}^{n} f((Z(i), A(i))) = S_n,
\end{aligned}
$$

sodass die Resultate über $\overline{S}_N$ auf $\overline{S}(N)$ übertragen werden können, wenn man die Problematik des Startzustands beachtet.

Dazu definieren wir die Zufallsgrößen $\tau_k(s)$, $\rho_k(s)$ und $Y_k(s)$ auch für Zustände $s \in S$. (Wie oben sind $(\rho_k(s))_{k \in \mathbb{N}}$ und $(Y_k(s))_{k \in \mathbb{N}}$ iid-Variablen, deren Verteilung unabhängig vom Startzustand ist.)

Genauso wie in den Lemmata 5.1 und 5.2 kann man zeigen, dass auch für $s \in S$ $E(Y_1(s)^m)$ für alle $m \in \mathbb{N}$ existiert und $\pi_s E(Y_1(s))$ konstant gleich $\sum_{t\in S} r(t,x,y)\pi_t$ ist (wobei $r(t,x,y) := \sum_{(k,l)\in A_t} r(t,k,l)x_t(k)y_t(l)$ die erwartete Auszahlung in Zustand t unter (x,y) ist).
Im Übrigen sieht man wie bei Lemma 5.4, dass mit $Y_1(s) = 0$ fast sicher (für festes $s \in S$) auch $Y_1(t) = 0$ fast sicher für *alle* $t \in S$ gilt. Die Beweise dieser Aussagen unterscheiden sich kaum oder gar nicht von denen der genannten Lemmata, sodass wir uns hier auf die Zusammenfassung der Ergebnisse beschränken. Zwischen S und S^* besteht die folgende Beziehung:

Lemma 5.14 *Für alle $s^* \in S^*$ und alle $s \in S$ ist*

$$\gamma(s,x,y) = \pi^*_{s^*} E(Y_1(s^*)) = \sum_{t^*\in S^*} f(t^*)\pi^*_{t^*} = \sum_{t\in S} r(t,x,y)\pi_t = \pi_s E(Y_1(s)).$$

Beweis: Die zweite Gleichheit stammt aus Lemma 5.2, die letzte aus der entsprechenden Version für $s \in S$. Ferner ist

$$\begin{aligned}\sum_{t^*\in S^*} f(t^*)\pi^*_{t^*} &= \sum_{(t,k,l)\in S^*} r(t,k,l)\pi_t x_t(k)y_t(l)\\ &= \sum_{t\in S}\pi_t \sum_{(k,l)\in A_t} r(t,k,l)x_t(k)y_t(l) = \sum_{t\in S} r(t,x,y)\pi_t\end{aligned}$$

sowie

$$\begin{aligned}&\sum_{(t,k,l)\in S^*} r(t,k,l)\pi_t x_t(k)y_t(l)\\ &= \sum_{(t,k,l)\in S^*} r(t,k,l)x_t(k)y_t(l) \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} P_{sxy}(Z(n)=t)\\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sum_{(t,k,l)\in S^*} r(t,k,l)P_{sxy}(Z(n)=t,\, A(n)=(k,l))\\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} E_{sxy}(R(n)) = \gamma(s,x,y)\end{aligned}$$

für jedes $s \in S$. □

Also ist $\gamma(s,x,y)$ auf rekurrenten Klassen konstant in s. Außerdem wird deutlich, dass die Voraussetzung $\sum_{t^*\in S^*} f(t^*)\pi^*_{t^*} = 0$ aus den Sätzen des Abschnitts 5.3 hier der Forderung $\gamma(x,y) = (\gamma(1,x,y),\ldots,\gamma(z,x,y)) = 0$ entspricht.

Wenden wir nun Satz 5.13 auf die assoziierte Markowkette an, so tritt für einen fest gewählten Startzustand s^*_1 neben dem asymmetrischen Sonderfall die Fallunterscheidung $Y_1(s_1^*) = 0$ fast sicher/nicht fast sicher auf. Auch diese Bedingung ist in die Situation des zugrunde liegenden Spiels zu übertragen.

Lemma 5.15 *Es sei $s_1^* = (s_1, i_1, j_1) \in S^*$. Dann gilt $Y_1(s_1) = 0$ fast sicher genau dann, wenn $Y_1(s_1^*) = 0$ fast sicher.*

Beweis: Ist $Y_1(s_1) = 0$ fast sicher, so gilt für alle Historien der Form $(s_1, k_1, l_1, s_2, \ldots, s_m = s_1)$ (mit $s_j \in S$ und $(k_j, l_j) \in A_{s_j}$), die positives P_{s_1xy}-Maß haben, dass $\sum_{j=1}^{m-1} r(s_j, k_j, l_j) = 0$. Weil das auf s_m folgende Aktionspaar dabei keine Rolle spielt, gilt in der assoziierten Markowkette ebenfalls für alle Historien der Gestalt $(s_1^* = (s_1, k_1, l_1), s_2^*, \ldots, s_m^* = s_1^*)$, dass $\sum_{j=1}^{m-1} f(s_j^*) = 0$. Für $(k_1, l_1) = (i_1, j_1)$ erhält man $Y_1(s_1^*) = 0$ fast sicher.
Für die Rückrichtung beachte man, dass nach Lemma 5.4 mit $Y_1(s_1^*) = 0$ f.s. auch $Y_1(s^*) = 0$ f.s. für alle $s^* \in S^*$ gilt. Ist nun $(s_1, k_1, l_1, s_2, \ldots, s_m = s_1)$ eine beliebige Historie, so folgt $\sum_{j=1}^{m-1} r(s_j, k_j, l_j) = 0$ analog zu eben wegen $Y_1((s_1, k_1, l_1)) = 0$ fast sicher. Damit gilt $Y_1(s_1) = 0$ fast sicher. □

Nun können wir die drei Fälle aus Satz 5.13 auf stochastische Spiele übertragen. Zuerst zeigen wir eine Aussage für die degenerierte Situation. (Man beachte, dass $Y_1(s_1) = 0$ f.s. nach Lemma 5.14 $\gamma(s_1, x, y) = 0$ impliziert.)

Lemma 5.16 *Besteht S nur aus einer einzigen rekurrenten Klasse und ist s_1 ein Startzustand mit $Y_1(s_1) = 0$ P_{s_1xy}-fast sicher, so gilt $\tilde{\gamma}_T(s_1, x, y) = \gamma_T(s_1, x, y)$. Es ist sogar $\lim_{N\to\infty} \overline{S}(N) = \gamma_T(s_1, x, y)$ P_{s_1xy}-fast sicher.*

Beweis: Satz 5.13 (a) besagt, dass bei Start in einem festen $s_1^* = (s_1, i_1, j_1)$ $(\overline{S}_N)_{N\in\mathbb{N}}$ betragsmäßig durch $M_{s_1}^* := \max_{t^*\in S^*} |c_{(s_1,i_1,j_1)}(t^*)|$ beschränkt ist und fast sicher gegen die Konstante $\sum_{t^*\in S^*} c_{(s_1,i_1,j_1)}(t^*)\pi_{t^*}^*$ konvergiert. Wegen $Y_1(s_1) = 0$ fast sicher hängen die $c_{(s_1,i_1,j_1)}(t^*)$ jedoch nicht von i_1, j_1 ab. Somit konvergiert $\overline{S}(N)$ auf ganz $\{Z(1) = s_1\}$ gegen dieselbe Konstante und ist auf dieser Menge dem Betrage nach durch $M_{s_1}^*$ beschränkt. Die Anwendung des Satzes von Lebesgue liefert daher:

$$\tilde{\gamma}_T(s_1, x, y) = E_{s_1xy}(\lim_{N\to\infty} \overline{S}(N)) = \lim_{N\to\infty} E_{s_1xy}(\overline{S}(N)) = \gamma_T(s_1, x, y).$$

Insbesondere gilt dann $\lim_{N\to\infty} \overline{S}(N) = \gamma_T(s_1, x, y)$ P_{s_1xy}-fast sicher. □

Nun kommen wir zum nicht-degenerierten symmetrischen Fall:

Lemma 5.17 *Besteht S nur aus einer einzigen rekurrenten Klasse und ist s_1 ein Startzustand mit $P_{s_1xy}(Y_1(s_1) \neq 0) > 0$ und $\gamma(s_1, x, y) = 0$, so gilt $\tilde{\gamma}_T(s_1, x, y) = -\infty$. Es ist sogar $\liminf_{N\to\infty} \overline{S}(N) = -\infty$ P_{s_1xy}-fast sicher.*

Beweis: Wenn $Y_1(s_1)$ nicht fast sicher gleich Null ist, so gilt dies nach Lemma 5.15 auch für alle $Y_1((s_1, i_1, j_1))$ (mit $(i_1, j_1) \in A_{s_1}$). Nach Lemma 5.14 ist $\sum_{t^* \in S^*} f(t^*)\pi^*_{t^*} = \gamma(s_1, x, y) = 0$.
Auf jeder der Mengen $\{Z(1) = s_1,\ A(1) = (i_1, j_1)\}$ gilt nach Satz 5.13 (b) also $\liminf_{N\to\infty} \overline{S}(N) = -\infty$ fast sicher. Somit gilt dies auf ganz $\{Z(1) = s_1\}$ und es folgt sofort $\tilde{\gamma}_T(s_1, x, y) = E_{s_1xy}(\liminf_{N\to\infty} \overline{S}(N)) = -\infty$. □

Schließlich noch eine Aussage zum trivialen (asymmetrischen) Fall:

Lemma 5.18 *Besteht S nur aus einer einzigen rekurrenten Klasse und ist s_1 ein Startzustand mit $\gamma(s_1, x, y) \neq 0$, so gilt*

$$\tilde{\gamma}_T(s_1, x, y) = \begin{cases} \infty & \textit{falls } \gamma(s_1, x, y) > 0, \\ -\infty & \textit{falls } \gamma(s_1, x, y) < 0. \end{cases}$$

Es ist sogar $\lim_{N\to\infty} \overline{S}(N) = \tilde{\gamma}_T(s_1, x, y)$ P_{s_1xy}*-fast sicher.*

Beweis: Nach Lemma 5.14 ist $\sum_{t^* \in S^*} f(t^*)\pi^*_{t^*} = \gamma(s_1, x, y) \neq 0$. Nach Satz 5.13 (c) gilt die Aussage also für jedes $(i_1, j_1) \in A_{s_1}$ auf der Menge $\{Z(1) = s_1,\ A(1) = (i_1, j_1)\}$. Damit gilt sie auch P_{s_1xy}-fast sicher. □

5.5 Anwendung auf stochastische Spiele - Teil II

Nun betrachten wir die allgemeine Situation: Der Zustandsraum S zerfällt unter (x, y) in verschiedene rekurrente Klassen (und transiente Zustände). Für die Menge der rekurrenten bzw. transienten Zustände schreiben wir S_{rek} bzw. S_{tr}. Abkürzend setzen wir $S^+_{rek} := \{s \in S_{rek} : \gamma(s, x, y) > 0\}$, $S^-_{rek} := \{s \in S_{rek} : \gamma(s, x, y) < 0\}$ und $S^0_{rek} := \{s \in S_{rek} : \gamma(s, x, y) = 0\}$.
Zur Unterscheidung des degenerierten und nicht-degenerierten Falls sei noch $S^0_{dg} := \{s \in S^0_{rek} : Y_1(s) = 0\ P_{sxy}\text{-f.s.}\}$ und $S^0_{ndg} := S^0_{rek} \backslash S^0_{dg}$.
Der Zeitpunkt des Eintreffens in einem rekurrenten Zustand werde mit $\varrho := \inf\{n : Z(n) \in S_{rek}\}$ bezeichnet. Sei nun $s_1 \in S$ ein *beliebiger* Startzustand.

Lemma 5.19 $E_{s_1xy}(\varrho^m)$ *existiert für alle* $m \in \mathbb{N}$.

Beweis: Im Fall $s_1 \in S_{rek}$ ist die Aussage klar. Sei also $s_1 \in S_{tr}$. Mit B_t werde die Anzahl der Besuche in einem Zustand $t \in S$ bezeichnet. Dann gilt

$$\varrho^m = \Big(\sum_{t \in S_{tr}} B_t + 1\Big)^m \leq (|S_{tr}| + 1)^m \Big(\sum_{t \in S_{tr}} B_t^m + 1^m\Big).$$

Also ist $E_{s_1xy}(\varrho^m) < \infty$, wenn $E_{s_1xy}(B_t^m) < \infty$ für jedes $t \in S_{tr}$. Im Fall $P_{s_1xy}(B_t \geq 1) = 0$ ist dies offensichtlich erfüllt. Andernfalls ist

$$\begin{aligned} E_{s_1xy}(B_t^m) &= \sum_{n=1}^{\infty} P_{s_1xy}(B_t^m \geq n) \\ &= P_{s_1xy}(B_t \geq 1) \sum_{n=1}^{\infty} P_{s_1xy}(B_t \geq n^{1/m} \mid B_t \geq 1) \\ &\leq \sum_{n=1}^{\infty} P_{txy}(B_t \geq n^{1/m} - 1) = \sum_{n=1}^{\infty} P_{txy}(B_t \geq \lceil n^{1/m} \rceil - 1) \quad (5.6) \\ &= \sum_{n=0}^{\infty} P_{txy}(B_t \geq n)((n+1)^m - n^m) \\ &= \sum_{n=0}^{\infty} P_{txy}(B_t \geq 1)^n((n+1)^m - n^m) < \infty. \quad (5.7) \end{aligned}$$

Dabei nutzen wir in (5.6) die Markoweigenschaft aus. In (5.7) geht ein, dass für transientes t $P_{txy}(B_t \geq 1) < 1$ gilt, die dort stehende Reihe also für jedes m konvergiert. □

Nun kommen wir schließlich zum Hauptsatz dieses Abschnitts. Für $T \subset S$ verstehen wir unter $s_1 \rightsquigarrow T$, dass $P_{s_1xy}(Z(n) \in T$ für irgendein $n) > 0$. Ferner müssen wir hier noch einmal die Sonderfallregel (vgl. (1.2)) aus der Definition von $\tilde{\gamma}_T$ berücksichtigen. Wir gehen davon aus, dass $\tilde{\gamma}_T(s, \pi, \sigma) := w \in \overline{\mathbb{R}}$ gesetzt wird, wenn der Erwartungswert (1.1) nicht exisitiert und außerdem $P_{s\pi\sigma}(\liminf_{N\to\infty} \overline{S}(N) = -\infty) > 0$ ist. (Gemäß (1.2) ist in unserem Modell $w = -\infty$.)

Die Fallunterscheidung im folgenden Satz erscheint auf den ersten Blick sehr technisch, doch zu einem gegebenen Strategiepaar (x, y) lassen sich die verschiedenen Teilmengen von S_{rek} leicht auseinander halten. Außerdem wird man stochastische Spiele, für die S_{rek}^+ und S_{rek}^- nicht beide leer sind, ohnehin eher mit der Durchschnittsauszahlung γ als den Totalauszahlungskriterien $\tilde{\gamma}_T$ bzw. γ_T untersuchen. Im Fall $S_{rek}^+ = S_{rek}^+ = \emptyset$ (also $\gamma(s, x, y) = 0$ für alle $s \in S$) bleiben in Satz 5.20 nur noch die ersten beiden Fälle (in vereinfachter Formulierung) übrig. Dann spielt auch die $\tilde{\gamma}_T$-Sonderfallregel keine Rolle mehr.

Satz 5.20 *Es sei $s_1 \in S$ irgendein Startzustand und $(x, y) \in X \times Y$. Dann gilt*

$$\tilde{\gamma}_T(s_1, x, y) = \begin{cases} \gamma_T(s_1, x, y) & \textit{falls } s_1 \not\rightsquigarrow S_{rek}^- \cup S_{ndg}^0, \\ -\infty & \textit{falls } s_1 \not\rightsquigarrow S_{rek}^+ \textit{ und } s_1 \rightsquigarrow S_{rek}^- \cup S_{ndg}^0, \\ w & \textit{sonst.} \end{cases} \quad (5.8)$$

Beweis: Für $N \in \mathbb{N}$ gilt

$$\overline{S}(N) = \frac{1}{N}\sum_{n=1}^{N} S(n) = \frac{1}{N}\sum_{n=1}^{N\wedge(\varrho-1)} S(n) + \frac{1}{N}\sum_{n=\varrho}^{N} S(n)$$
$$= \frac{1}{N}\sum_{n=1}^{N\wedge(\varrho-1)} S(n) + \frac{N-\varrho+1}{N} S(\varrho-1)1_{\{\varrho\leq N\}} + \frac{1}{N}\sum_{n=\varrho}^{N}(S(n)-S(\varrho-1)). \tag{5.9}$$

Wegen $|\frac{1}{N}\sum_{n=1}^{N\wedge(\varrho-1)} S(n)| \leq \frac{1}{N}\sum_{n=1}^{N\wedge(\varrho-1)} |S(n)| \leq \frac{1}{N}\sum_{n=1}^{N\wedge(\varrho-1)} nM \leq \frac{M\varrho^2}{2N}$ (wobei $M := \max_{s,i,j}|r(s,i,j)|$) strebt der erste Term in (5.9) fast sicher gegen Null, wenn $N \to \infty$. Der zweite Term strebt fast sicher gegen $S(\varrho-1)$. Außerdem können wir im Grenzübergang den Faktor $\frac{1}{N}$ vor dem dritten Term in (5.9) durch $\frac{1}{N-\varrho+1}$ ersetzen. Somit ist

$$\liminf_{N\to\infty} \overline{S}(N) = S(\varrho-1) + \liminf_{N\to\infty} \frac{1}{N-\varrho+1}\sum_{n=\varrho}^{N}(S(n)-S(\varrho-1))$$
$$= S(\varrho-1) + \sum_{s\in S_{rek}} 1_{\{Z(\varrho)=s\}} \liminf_{N\to\infty} \frac{1}{N-\varrho+1}\sum_{n=\varrho}^{N}(S(n)-S(\varrho-1)).$$

Wegen der Markow-Eigenschaft ist der hinten stehende liminf auf der Menge $\{Z(\varrho) = s\}$ genauso verteilt wie $\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} S(n)$ auf der Menge $\{Z(1) = s\}$. Die Situation eines Startzustand $s \in S_{rek}$ haben wir aber schon oben untersucht (siehe Lemmata 5.16 bis 5.18). Demnach ist

$$\liminf_{N\to\infty} \overline{S}(N) = \begin{cases} \gamma_T(s,x,y) & \text{f.s. auf } \{Z(\varrho) = s \in S^0_{dg}\}, \\ -\infty & \text{f.s. auf } \{Z(\varrho) \in S^0_{ndg}\}, \\ \infty & \text{f.s. auf } \{Z(\varrho) \in S^+_{rek}\}, \\ -\infty & \text{f.s. auf } \{Z(\varrho) \in S^-_{rek}\}. \end{cases} \tag{5.10}$$

Dabei handelt es sich in den Fällen 1,3 und 4 sogar um fast sichere Aussagen über $\lim_{N\to\infty} \overline{S}(N)$ (vgl. Lemma 5.16 und 5.18).
Im Fall $P_{s_1xy}(Z(\varrho) \in S^0_{dg}) = 1$ konvergiert $\overline{S}(N)$ also P_{s_1xy}-fast sicher (gegen $S(\varrho-1)+\sum_{s\in S^0_{dg}} \gamma_T(s,x,y)1_{\{Z(\varrho)=s\}}$). Ferner gilt mit (5.9), Lemma 5.16 und Lemma 5.19 $|\overline{S}(N)| \leq \frac{M\varrho^2}{2N} + M(\varrho-1) + \max_{s\in S^0_{dg}} M^*_s \in \mathcal{L}^1(P_{s_1xy})$, wobei M^*_s aus dem Beweis von Lemma 5.16 stammt. Im Fall $P_{s_1xy}(Z(\varrho) \in S^0_{dg}) = 1$ ist also

$$\tilde{\gamma}_T(s_1,x,y) = E_{s_1xy}(\lim_{N\to\infty} \overline{S}(N)) = \lim_{N\to\infty} E_{s_1xy}(\overline{S}(N)) = \gamma_T(s_1,x,y).$$

Nun betrachten wir den Fall $s_1 \rightsquigarrow S_{rek}^+$, aber $s_1 \not\rightsquigarrow S_{rek}^- \cup S_{ndg}^0$. Unter Benutzung von

$$E_{s_1xy}\left(\sum_{n=N+1}^{N+\varrho-1}\big(S(n)-S(\varrho-1)\big)\right) \le E_{s_1xy}\left(\sum_{n=1}^{\varrho-1} NM\right) = MNE_{s_1xy}(\varrho-1)$$

sowie (5.9) (inklusive der darauf folgenden Abschätzung) ist

$$\begin{aligned}
&E_{s_1xy}(\overline{S}(N))\\
&\ge -\frac{ME_{s_1xy}(\varrho^2)}{2N} - ME_{s_1xy}(\varrho) + E_{s_1xy}\left(\frac{1}{N}\sum_{n=\varrho}^{N}\big(S(n)-S(\varrho-1)\big)\right)\\
&\ge -\frac{ME_{s_1xy}(\varrho^2)}{2N} - 2ME_{s_1xy}(\varrho) + E_{s_1xy}\left(\frac{1}{N}\sum_{n=\varrho}^{N+\varrho-1}\big(S(n)-S(\varrho-1)\big)\right).
\end{aligned}$$

Mit $S'_{rek} := \{s \in S_{rek}\colon P_{s_1xy}(Z(\varrho)=s) > 0\}$ ist der letzte Term gleich

$$\begin{aligned}
&\sum_{s\in S'_{rek}} E_{s_1xy}\left(\frac{1}{N}\sum_{n=\varrho}^{N+\varrho-1}\big(S(n)-S(\varrho-1)\big)\,\middle|\, Z(\varrho)=s\right)P_{s_1xy}(Z(\varrho)=s)\\
&= \sum_{s\in S'_{rek}} E_{sxy}\left(\frac{1}{N}\sum_{n=1}^{N} S(n)\right)P_{s_1xy}(Z(\varrho)=s)
\end{aligned}$$

unter Ausnutzung der Markow-Eigenschaft. Für $s \in S_{dg}^0$ ist der innen stehende Erwartungswert wie oben gesehen beschränkt. Für $s \in S_{rek}^+$ dagegen konvergiert $\overline{S}(N)$ nach Lemma 5.18 P_{sxy}-fast sicher gegen ∞. Wie man leicht sieht ist dabei auch $E_{sxy}(\overline{S}(N)) \to \infty$. Also gilt im vorliegenden Fall insgesamt $\gamma_T(s_1,x,y) = \liminf_{N\to\infty} E_{s_1xy}(\overline{S}(N)) = \infty$. Da nach (5.10) in diesem Fall $\tilde{\gamma}_T(s_1,x,y) = \infty$ gilt, stimmen $\tilde{\gamma}_T$ und γ_T also auch hier überein. In den anderen Fällen kann man die Behauptung sofort ablesen. □

Es ist noch zu beachten, dass auch im zweiten Fall (d.h. $\tilde{\gamma}_T(s_1,x,y) = -\infty$) die Gleichheit $\tilde{\gamma}_T(s_1,x,y) = \gamma_T(s_1,x,y)$ bestehen kann. Dies gilt etwa dann, wenn $s_1 \rightsquigarrow S_{rek}^-$, aber $s_1 \not\rightsquigarrow S_{rek}^+ \cup S_{ndg}^0$, wie man durch eine zum Beweis des oben behandelten Falls $s_1 \rightsquigarrow S_{rek}^+$, $s_1 \not\rightsquigarrow S_{rek}^- \cup S_{ndg}^0$ symmetrische Argumentation sehen kann. Eine präzisere Fallunterscheidung wäre also möglich, ginge aber zu sehr auf Kosten der Übersichtlichkeit.

Damit haben wir eine Formel für die alternative Totalauszahlung unter einem stationären Strategiepaar gewonnen. Das Resultat macht deutlich, dass wir (abgesehen von asymmetrischen Situationen) nur in degenerierten Fällen mit einer anderen Bewertung als $-\infty$ rechnen können. Dies entspricht der Interpretation, dass Spieler 1 sich nicht auf Spiele einlassen möchte, in denen er auf lange Sicht immer wieder beliebig tief ins Minus gerät.

Literaturverzeichnis

[BF68] D. Blackwell and T.S. Ferguson. The Big Match. *Annals of Mathematical Statistics*, 39:159–163, 1968.

[Bla62] D. Blackwell. Discrete Dynamic Programming. *Annals of Mathematical Statistics*, 33:719–726, 1962.

[Chu67] K.L. Chung. *Markov Chains with Stationary Transition Probabilities (2nd Edition)*, volume 104 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen.* Springer-Verlag, Berlin, 1967.

[DN83] H.M. Dietz and V. Nollau. *Markov Decision Problems with Countable State Spaces*, volume 15 of *Mathematical Research.* Akademie-Verlag, Berlin, 1983.

[Fre71] D. Freedman. *Markov Chains.* Holden-Day, San Francisco, 1971.

[FV97] J. Filar and O.J. Vrieze. *Competitive Markov Decision Processes.* Springer-Verlag, New York, 1997.

[Gil57] D. Gilette. Stochastic games with zero stop probabilities. In A.W. Tucker M. Dresher and P. Wolfe, editors, *Contributions to the Theory of Games III*, number 39 in Annals of Mathematical Studies, pages 179–187. Princeton University Press, 1957.

[Gut88] A. Gut. *Stopped Random Walks: Limit Theorems and Applications*, volume 5 of *Applied Probability.* Springer-Verlag, New York, 1988.

[MN81] J.F. Mertens and A. Neyman. Stochastic Games. *International Journal of Game Theory*, 10:53–66, 1981.

[MS96] A.P. Maitra and W.D. Sudderth. *Discrete Gambling and Stochastic Games*, volume 32 of *Applications of Mathematics.* Springer-Verlag, New York, 1996.

[Pet95] V.V. Petrov. *Limit Theorems of Probability Theory*, volume 4 of *Oxford Studies in Probability.* Oxford University Press, Oxford, 1995.

[Sha53] L.S. Shapley. Stochastic Games. *Proceedings of the National Academy of Sciences U.S.A*, 39:1095–1100, 1953.

[Thu92] F. Thuijsman. *Optimality and Equilibria in Stochastic Games.* Number 82 in CWI Tracts. Centrum voor Wiskunde en Informatica, Amsterdam, 1992.

[TV87] F. Thuijsman and O.J. Vrieze. The Bad Match, a Total Reward Stochastic Game. *Operations Research Spektrum*, 9:93–99, 1987.

[TV96] F. Thuijsman and O.J. Vrieze. Total Reward Stochastic Games and Sensitive Average Reward Strategies. Technical Report M 96–08, Department of Mathematics, Maastricht University, 1996.

[vdW84] J. van der Wal. *Stochastic Dynamic Programming (2nd Edition).* Number 139 in Mathematical Centre Tracts. Mathematisch Centrum, Amsterdam, 1984.

Lebenslauf

20.10.1970	geboren in Kassel
1977–1980	Dorothea-Viehmann-Grundschule in Kassel
1980–1987	Leimborn-Gesamtschule (Gymnasialzweig) in Kassel
1987–1990	Oberstufengymnasium Oberzwehren in Kassel
11.6.1990	Abitur
1990–1991	Zivildienst beim Sozialen Friedensdienst in Kassel
1991–1994	Mathematikstudium (mit Nebenfach Informatik) an der Universität Gesamthochschule Kassel
18.10.1993	Vordiplom
1994–1997	Umzug nach Göttingen und Fortsetzung des Studiums an der Georg-August-Universität Göttingen
27.6.1997	Diplom
seit 1997	Promotionsstudium bei Prof. Krengel am Institut für Mathematische Stochastik der Universität Göttingen
seit 20.4.1998	Mitarbeiter des DFG-Projekts „Stochastische Spiele“
seit 16.10.1998	Wissenschaftlicher Angestellter am Institut für Mathematische Stochastik