



UNIVERSITÄT
BAYREUTH

FAKULTÄT FÜR MATHEMATIK UND PHYSIK
MATHEMATISCHES INSTITUT

Analyse eines dynamischen Spiels zur CO₂-Reduktion

Diplomarbeit

von

Nadine Löhner

Datum: 17. Februar 2009

Aufgabenstellung und Betreuung:
Prof. Dr. L. Grüne

Danksagung

An dieser Stelle möchte ich mich recht herzlich bei Herrn Prof. Dr. Lars Grüne für die interessante Themenstellung, sowie die hervorragende Betreuung während der Erstellung dieser Arbeit bedanken.

Inhaltsverzeichnis

Tabellenverzeichnis	iii
Abbildungsverzeichnis	v
1 Einleitung	1
2 Eine kurze Einführung in die Spieltheorie	5
3 Zeitabhängige, diskrete Systeme	11
4 Optimale Steuerung	13
4.1 Optimale Steuerung von Kontrollsystemen auf endlichem Horizont	14
4.2 Bellman'sches Optimalitätsprinzip auf endlichem Horizont	15
4.3 Optimale Steuerung bei dynamischen Spielen	19
4.3.1 Nicht-antizipierende Strategien	20
4.3.2 Bellman'sches Optimalitätsprinzip bei nicht-antizipierenden Strategien	25
5 Ein spieltheoretisches Modell zur CO₂-Reduktion	33
5.1 Motivation	33
5.2 Das ungesteuerte Modell	35
5.2.1 Das Wachstum der Mittelaufwendung	35
5.2.2 Die <i>em</i> -Matrix	37
5.2.3 Untersuchung auf einen attraktiven Fixpunkt	39
5.3 Das gesteuerte Modell	41
5.4 Das dynamische Spiel zur CO ₂ -Reduktion	42

6	Dynamische Programmierung	47
6.1	Spiel mit einem Spieler	47
6.1.1	Das maple-Programm	47
6.1.2	Ergebnisse des dynamischen Spiels zur CO ₂ -Reduktion . .	48
6.2	Spiel mit zwei Spielern	52
6.2.1	Das maple-Programm	52
6.2.2	Ergebnisse des dynamischen Spiels zur CO ₂ -Reduktion . .	54
7	Ein Greedy-Algorithmus	59
7.1	Spiel mit einem Spieler	59
7.1.1	Der Greedy-Algorithmus	59
7.1.2	Ergebnisse des Greedy-Algorithmus	62
7.2	Spiel mit zwei Spielern	65
7.2.1	Der Greedy-Algorithmus	65
7.2.2	Ergebnisse des Greedy-Algorithmus	69
8	Analyse und Vergleich der Algorithmen	73
8.1	Spiel mit einem Spieler	73
8.1.1	Abweichungen der Lösungstrajektorien von den Zielwerten	73
8.1.2	Gesamtkosten	75
8.2	Spiel mit zwei Spielern	75
8.2.1	Abweichungen der Lösungstrajektorien von den Zielwerten	75
8.2.2	Gesamtkosten	77
9	Das Stackelberg-Gleichgewicht	81
9.1	Theoretische Einführung	81
9.2	Anwendung auf Beispiele	83
10	Zusammenfassung	89
A	Inhalt der beiliegenden CD	I
	Literaturverzeichnis	II

Tabellenverzeichnis

6.1	Startparameter bei einem Spieler	49
6.2	Startparameter bei zwei Spielern bei nicht-antizipierender Spielsituation	54
6.3	Werte der <i>em</i> -Matrix	54
7.1	Startparameter bei einem Spieler	62
7.2	Startparameter bei zwei Spielern bei nicht-antizipierender Spielsituation	69
7.3	Werte der <i>em</i> -Matrix	69
9.1	Parameter für das Beispiel des Stackelberg-Gleichgewichts	88

Abbildungsverzeichnis

2.1	Stein-Schere-Papier	6
2.2	Gefangenendilemma	7
4.1	Bellman'sches Optimalitätsprinzip	16
4.2	Darstellung des Spiels mit nicht-antizipierender Spielsituation	23
5.1	Umweltdilemma	34
5.2	Darstellung des logistischen Wachstums	36
5.3	Darstellung des beschränkten Wachstums	37
5.4	Darstellung der em -Matrix bei zwei Spielern	38
5.5	Darstellung der em -Matrix bei zwei Spielern mit Steuerungen	42
6.1	Emissionsreduktion bei einem Spieler	49
6.2	Mittelaufwendung bei einem Spieler	50
6.3	Optimale Wertefunktion bei einem Spieler	50
6.4	Emissionsreduktion ohne untere Schranke der Mittelaufwendung	51
6.5	Mittelaufwendung ohne untere Schranke	51
6.6	Emissionsreduktion von Spieler 1	55
6.7	Emissionsreduktion von Spieler 2	55
6.8	Mittelaufwendung von Spieler 1	56
6.9	Mittelaufwendung von Spieler 2	56
6.10	Optimale Wertefunktion von Spieler 1	57
6.11	Optimale Wertefunktion von Spieler 2	57
7.1	Darstellung des Greedy-Algorithmus bei einem Spieler	61
7.2	Emissionsreduktion bei einem Spieler	62
7.3	Mittelaufwendung bei einem Spieler	63
7.4	Emissionsreduktion ohne untere Schranke der Mittelaufwendung	64

7.5	Mittelaufwendung ohne untere Schranke	64
7.6	Darstellung des Greedy-Algorithmus bei zwei Spielern	68
7.7	Emissionsreduktion von Spieler 1	70
7.8	Emissionsreduktion von Spieler 2	70
7.9	Mittelaufwendung von Spieler 1	71
7.10	Mittelaufwendung von Spieler 2	71
9.1	Kosten von Spieler 1	83
9.2	Kosten von Spieler 2	84

KAPITEL 1

Einleitung

Warum Spieltheorie?

Leute hören nicht auf zu spielen, weil sie alt werden, sie werden alt, weil sie aufhören zu spielen.

Oliver Wendell Holmes (1809-1894) [23]

Spiele beeinflussen unser ganzes Leben, schon als Kinder spielen wir. Zu Beginn vielleicht nur des Spiels wegen, später aber um selbiges zu gewinnen. Dafür versuchen wir spielerisch die für uns besten Strategien zu entwickeln. Dabei kann uns die Spieltheorie nützlich sein. Mit ihrer Hilfe ist es möglich nicht nur Gesellschaftsspiele unserer Kindheit zu entschlüsseln, nein, wir können vielmehr selbst neue Szenarien erschaffen und diese untersuchen. Ein Spiel in der Spieltheorie entspricht einer Entscheidungssituation mit mehreren Beteiligten, die einander in ihrer Entscheidung gegenseitig beeinflussen. Anwendung findet die Spieltheorie daher vor allem in der Ökonomie bei der Betrachtung von Märkten, der Politikwissenschaft um Konfliktsituationen zu untersuchen, aber auch in der Biologie bei der Betrachtung vom Wachstum einer Population. Aufgabe der Mathematik ist es nun das Ganze in einen formalen Rahmen zu bringen. Wie wir in dieser Arbeit sehen werden, hat sich die Spieltheorie von den einfachen statischen Spielen hin zu den dynamischen Spielen entwickelt. Hierfür sind dann auch die entsprechenden mathematischen Werkzeuge der optimalen Steuerung nötig. Dabei handelt es sich dann nicht mehr nur um kindliche Spielerei, sondern um einen sehr modernen Forschungszweig der Mathematik.

Ziel dieser Arbeit ist es dynamische Nicht-Nullsummenspiele mathematisch zu formalisieren und dann spieltheoretisch zu analysieren. Dazu muss zunächst der spieltheoretische und mathematische Hintergrund beleuchtet werden. Dann erst können wir dynamische Spiele in Beispielen behandeln. Im Mittelpunkt wird bei der praktischen Behandlung das CO₂-Reduktionsmodell aus [8] und [14] stehen. Dieses spieltheoretische Modell beschreibt ein Spiel, bei dem ein oder mehrere Spieler durch finanzielle Mittelaufwendung ihre CO₂-Emissionen verringern können.

Wie bereits erwähnt müssen wir uns zunächst mit den spieltheoretischen Rahmenbedingungen befassen. In Kapitel 2 wird deshalb mit einer kurzen Einführung in die Spieltheorie selbst begonnen. Dies erleichtert die Einordnung der dynamischen Spiele in den Hintergrund der Spieltheorie. Hier werden zu Beginn einfache Matrix-Spiele vorgestellt um den Unterschied von Nullsummenspielen und Nicht-Nullsummenspielen deutlich zu machen. Außerdem werden die sehr entscheidenden Begriffe des *Pareto-Optimums* und des *Nash-Gleichgewichts* vorgestellt.

Für die Behandlung von dynamischen Spielen sind außerdem mathematische Grundlagen zu schaffen, um die einzelnen Spiele korrekt definieren und analysieren zu können. Dies soll in den folgenden Kapiteln geschehen. In Kapitel 3 führen wir dynamische Systeme und Kontrollsysteme in diskreter Zeit ein. Diese Systeme werden dann auch in Kapitel 4 eine wesentliche Grundlage bei der Lösung der optimalen Steuerungsprobleme mit dem Bellman'schen Optimalitätsprinzip bilden. In diesem Kapitel der optimalen Steuerung verlassen wir dann die rein theoretische Sichtweise und betrachten die ersten dynamischen Spiele. Dabei werden wir besondere Strategien bei Spielen mit zwei gegnerischen Spielern kennenlernen, die sogenannten nicht-antizipierenden Strategien.

Da nun alle mathematischen Grundlagen gelegt sind, können wir in Kapitel 5 ein spieltheoretisches Modell zur CO₂-Reduktion vorstellen. Wir leiten dazu zunächst das ungesteuerte Modell bestehend aus einem System aus Differenzgleichungen her. Anschließend überführen wir dieses in ein gesteuertes Modell, das für unsere praktischen Betrachtungen entscheidend sein wird.

In den folgenden Kapiteln wollen wir zwei verschiedene Möglichkeiten vorstellen das dynamische Spiel der CO₂-Reduktion zu lösen. Zunächst werden wir in Kapitel 6 dazu das Bellman'sche Optimalitätsprinzip verwenden. Anschließend stellen wir als Alternative in Kapitel 7 einen Greedy-Algorithmus vor. Eine Analyse der beiden Methoden soll dann in Kapitel 8 folgen.

In Kapitel 9 soll schließlich ein alternatives Gleichgewicht zum Nash-Gleichgewicht bei der Behandlung von Spielen vorgestellt werden, das sogenannte Stackelberg-Gleichgewicht. Wie bei den nicht-antizipierenden Strategien, hat auch hier ein

Spieler einen gewichtigen Vorteil gegenüber seinem Gegner. Hier kann der *Leader* den *Follower* sogar zwingen eine bestimmte Strategie zu wählen.

In Kapitel 10 sollen die wichtigsten Ergebnisse nochmals zusammengefasst werden. Abschließend geben wir einen kurzen Ausblick auf weitere mögliche spieltheoretische Untersuchungen bezüglich dieses CO₂-Reduktionsmodells.

KAPITEL 2

Eine kurze Einführung in die Spieltheorie

In diesem Kapitel wird ein kurzer Abriss über die verschiedenen Arten von Spielen gegeben, damit die Einordnung des in dieser Arbeit behandelten Systems vereinfacht wird. Als Grundlage dafür wurden [7], [8], [10] und [15] verwendet.

Um nun eine Kategorisierung der Spiele im Allgemeinen vornehmen zu können, muss zunächst entschieden werden, ob die Summe der Auszahlungen aller Spieler am Ende des Spiels stets Null ist, oder nicht. Falls diese Summe gleich Null ist, so nennen wir die Spiele *Nullsummenspiele*, andernfalls *Nicht-Nullsummenspiele*. Im ersten Fall geht es lediglich darum Gewinne und Verluste untereinander aufzuteilen. Die ökonomisch interessanteren Spiele verhalten sich aber meist anders. In diesem Fall ist die Summe nicht Null, sie ist nicht einmal konstant. Es geht vielmehr um den Verbrauch, oder die Erzeugung von Gütern. In beiden Fällen nehmen wir an, dass die Spieler jeweils ihren eigenen Nutzen maximieren wollen, was einem rationalen Verhalten entspricht.

Ein weiteres Merkmal bei der Kategorisierung von Spielen ist die Zeitabhängigkeit. Die nun folgenden Beispiele sind *statische Spiele*, die unabhängig von jeglichem zeitlichen Einfluss gespielt werden. Im Gegensatz dazu werden die *dynamischen Spiele* zeitabhängig gespielt. Bei ihnen wird mit Hilfe kontrolltheoretischer Überlegungen ein optimales Steuerungsproblem von dynamischen Systemen gelöst. Dieses werden wir in Kapitel 4 ausführlich behandeln.

Definition 2.1 (Zwei-Personen-Spiel).

Ein *Zwei-Personen-Spiel* Γ besteht aus einem Paar nichtleerer Mengen S und T und zwei reellwertigen Funktionen Φ_1 und Φ_2 auf dem kartesischen Produkt

$S \times T$.

Definition 2.2 (Zwei-Personen-Nullsummenspiel).

Ein Zwei-Personen-Spiel heißt *Nullsummenspiel*, wenn gilt $\Phi_1 + \Phi_2 \equiv 0$, das heißt, wenn der Gewinn des einen Spielers gleich dem Verlust des anderen Spielers ist.

Um nun den Unterschied zwischen *Nullsummenspielen* und *Nicht-Nullsummenspielen* zu verdeutlichen, geben wir zwei Beispiele in Form von Matrizen an. Hier entspricht jede Auszahlung an einen Spieler einem Eintrag in einer Matrix. Der eine Spieler kann dabei aus den Reihen der Matrix, der andere Spieler aus den Spalten der Matrix wählen. Diese Wahlmöglichkeiten entsprechen ihren jeweiligen Strategien.

Als Beispiel eines Zwei-Personen-Nullsummenspiels soll uns an dieser Stelle das Kinderspiel *Stein-Schere-Papier* dienen.

Beispiel 2.3 (Stein-Schere-Papier).

Zwei Spieler können zwischen *Stein*, *Schere* und *Papier* wählen. Die Auszahlung des Gewinners ist 1, die des Verlierers -1 und bei Unentschieden ist sie jeweils 0. Dies führt unsere Spieler zu folgender Auszahlungsmatrix bei der die Summe dieser Auszahlungen immer Null ist.

	Stein	Schere	Papier
Stein	0,0	1,-1	-1,1
Schere	-1,1	0,0	1,-1
Papier	1,-1	-1,1	0,0

Abbildung 2.1: Stein-Schere-Papier

Für den etwas komplizierteren Fall eines *Zwei-Personen-Nicht-Nullsummenspiels* ist das sogenannte *Gefangenendilemma* wohl das bekannteste Beispiel.

Beispiel 2.4 (Gefangenendilemma).

Zwei Gefangene werden verdächtigt, gemeinsam eine Straftat begangen zu haben und werden getrennt von einander verhört. Man kann ihnen die zur Last gelegte Tat jedoch nicht zweifelsfrei nachweisen. Die beiden Gefangenen haben also die Wahl zu gestehen, oder nicht zu gestehen. Deshalb sind je nach Verhalten der beiden Gefangenen unterschiedliche Strafen denkbar. Gestehen beide nicht, können sie nur wegen kleinerer Delikte zu jeweils einem Jahr Haft verurteilt werden. Gestehen die zwei Gefangenen die Tat wird ihnen eine Strafminderung von 50 Prozent auf jeweils fünf Jahre Haft angeboten. Wenn aber einer der Gefangenen nicht gesteht, während der andere etwa als Kronzeuge auftritt und gesteht, geht der Kronzeuge straffrei aus und der andere Gefangene muss 10 Jahre ins Gefängnis. Dies führt uns zu folgender Auszahlungsmatrix, bei der die Summe eben nicht Null ist.

	Gestehen	Nicht Gestehen
Gestehen	5,5	0,10
Nicht Gestehen	10,0	1,1

Abbildung 2.2: Gefangenendilemma

Für den Ausgang dieses Spieles müssen wir folgendes Optimum betrachten.

Definition 2.5 (Pareto-Optimum).

Es ist nicht möglich einen Spieler besser zu stellen, ohne zugleich einen anderen Spieler schlechter zu stellen.

In unserem Beispiel würde dies der Situation entsprechen, dass beide Gefangenen gestehen.

Bei Spielen können aber durchaus mehr als zwei Spieler beteiligt sein. Deshalb

verallgemeinern wir die Definition eines *Zwei-Personen-Spiels* auf die eines *n-Personen-Spiels* und stellen dann das *Nash-Gleichgewicht* bei diesen verallgemeinerten Spielen vor.

Definition 2.6 (n-Personen-Spiel).

Ein *n-Personen-Spiel* für $n \geq 2$ besteht aus einem n -Tupel (S_1, \dots, S_n) (nicht-leerer Mengen) S_i , $i = 1, \dots, n$ und einem n -Tupel (Φ_1, \dots, Φ_n) reellwertiger Funktionen Φ_i , $i = 1, \dots, n$ auf dem kartesischen Produkt $S_1 \times \dots \times S_n$, wobei S_i für $i = 1, \dots, n$ der *Strategiemenge* des Spielers P_i und $\Phi_i(s_1, \dots, s_n)$ der *Auszahlung* an den Spieler P_i entspricht.

Jeder Spieler P_i versucht nun seine Auszahlung $\Phi_i(s_1, \dots, s_n)$ zu maximieren. Das ist im Allgemeinen aber nicht simultan möglich. Daher muss ein Kompromiss gefunden werden, der für alle Spieler in einem gewissen Sinne optimal ist. Dies führt uns zu folgendem Gleichgewicht.

Definition 2.7 (Nash-Gleichgewicht).

Ein Strategien- n -Tupel $(\hat{s}_1, \dots, \hat{s}_n) \in S_1 \times \dots \times S_n$ heißt ein *Nash-Gleichgewicht*, wenn für alle $i = 1, \dots, n$ gilt

$$\Phi_i(\hat{s}_1, \dots, \hat{s}_n) \geq \Phi_i(\hat{s}_1, \dots, \hat{s}_{i-1}, s_i, \hat{s}_{i+1}, \dots, \hat{s}_n)$$

für alle $s_i \in S_i$

Bemerkung 2.8 (Interpretation des Nash-Gleichgewichts).

Ein Nash-Gleichgewicht ist ein Strategien- n -Tupel derart, dass die Abweichung eines Spielers von diesem n -Tupel bei Festhalten der anderen Spieler an ihrer Strategie höchstens zu einer Verkleinerung der Auszahlung an den Abweichler führt.

Satz 2.9 (Existenz eines Nash-Gleichgewichts).

Es gelten folgende Annahmen:

1. Die Strategiemengen S_i für $i = 1, \dots, n$ sind konvexe und kompakte Teilmengen im \mathbb{R}^{m_i}
2. Die Auszahlungsfunktionen $\Phi_i : S_1 \times \dots \times S_n \rightarrow \mathbb{R}$ sind stetig.

Gilt zusätzlich, dass zu jedem n -Tupel $(s_1^*, \dots, s_n^*) \in S_1 \times \dots \times S_n$ und jedem $i = 1, \dots, n$ genau ein $\tilde{s}_i \in S_i$ existiert mit

$$\Phi_i(s_1^*, \dots, s_{i-1}^*, \tilde{s}_i, s_{i+1}^*, \dots, s_n^*) \geq \Phi_i(s_1^*, \dots, s_{i-1}^*, s_i, s_{i+1}^*, \dots, s_n^*) \quad (2.1)$$

für alle $s_i \in S_i$.

Dann existiert ein Nash-Gleichgewicht.

Beweis.

Für jedes $i = 1, \dots, n$ definieren wir eine Abbildung $A_i : S_1 \times \dots \times S_n \rightarrow S_i$ vermöge

$$A_i(s_1^*, \dots, s_n^*) = \tilde{s}_i$$

mit $\tilde{s}_i \in S_i$ nach (2.1).

Damit definieren wir eine Abbildung $A = (A_1, \dots, A_n) : S_1 \times \dots \times S_n \rightarrow S_1 \times \dots \times S_n$.

Ein n -Tupel $(\hat{s}_1, \dots, \hat{s}_n) \in S_1 \times \dots \times S_n$ ist genau dann ein Fixpunkt von A , d.h. es gilt

$$A(\hat{s}_1, \dots, \hat{s}_n) = (\hat{s}_1, \dots, \hat{s}_n)$$

wenn $(\hat{s}_1, \dots, \hat{s}_n)$ ein Nash-Gleichgewicht ist.

Da die Menge $S_1 \times \dots \times S_n$ in $\mathbb{R}^{m_1 + \dots + m_n}$ konvex und kompakt ist, folgt die Existenz eines Fixpunktes von A und damit eines Nash-Gleichgewichtes nach dem Brouwer'schen Fixpunktsatz, wenn wir noch zeigen, dass die Abbildung A stetig ist.

Dazu muss gezeigt werden, dass jede Abbildung A_i stetig ist. Zu dem Zweck wählen wir eine Folge $(s_1^k, \dots, s_n^k)_{k \in \mathbb{N}_0}$ in $S_1 \times \dots \times S_n$ mit

$$\lim_{k \rightarrow \infty} s_i^k = s_i^* \in S_i$$

für $i = 1, \dots, n$.

Für jedes $k \in \mathbb{N}_0$ gilt dann mit $s^k = (s_1^k, \dots, s_n^k)$ für jedes $i = 1, \dots, n$

$$\begin{aligned} & \Phi_i(s_1^k, \dots, s_{i-1}^k, (A_i(s^k))_i, s_{i+1}^k, \dots, s_n^k) \\ & \geq \Phi_i(s_1^k, \dots, s_{i-1}^k, s_i, s_{i+1}^k, \dots, s_n^k) \end{aligned} \quad (2.2)$$

für alle $s_i \in S_i$

Weiter ist für jedes $i = 1, \dots, n$

$$\begin{aligned} & \Phi_i(s_1^*, \dots, s_{i-1}^*, (A_i(s^*))_i, s_{i+1}^*, \dots, s_n^*) \\ & \geq \Phi_i(s_1^*, \dots, s_{i-1}^*, s_i, s_{i+1}^*, \dots, s_n^*) \end{aligned}$$

für alle $s_i \in S_i$ und $s^* = (s_1^*, \dots, s_n^*)$.

Nun sei $i \in \{1, \dots, n\}$ beliebig fest gewählt. Dann gibt es eine Teilfolge $(s^{k_l})_{l \in \mathbb{N}_0}$ und ein $\tilde{s}_i \in S_i$ mit

$$\lim_{l \rightarrow \infty} (A_i(s^{k_l}))_i = \tilde{s}_i.$$

Aus (2.2) folgt somit

$$\begin{aligned} & \Phi_i(s_1^*, \dots, s_{i-1}^*, \tilde{s}_i, s_{i+1}^*, \dots, s_n^*) \\ & \geq \Phi_i(s_1^*, \dots, s_{i-1}^*, s_i, s_{i+1}^*, \dots, s_n^*) \end{aligned}$$

für alle $s_i \in S_i$
und daraus $\tilde{s}_i = (A_i(s^*))_i$.

Auf die gleiche Weise zeigt man, dass zu jeder Teilfolge $(s^{k_l})_{l \in \mathbb{N}_0}$ eine Teilfolge $(s^{k_{l_m}})_{m \in \mathbb{N}_0}$ existiert mit

$$\lim_{m \rightarrow \infty} (A_i(s^{k_{l_m}}))_i = (A_i(s^*))_i$$

woraus folgt, dass

$$\lim_{k \rightarrow \infty} (A_i(s^k))_i = (A_i(s^*))_i$$

gilt, was den Beweis vollendet.

□

Zeitabhängige, diskrete Systeme

An dieser Stelle wollen wir zeitabhängige, diskrete Systeme einführen. Zu solchen Systemen zählen zum einen die *dynamischen Systeme* und zum anderen die *Kontrollsysteme*. Dabei stammen die Definitionen aus [4] und [5]. Dynamische Systeme sind mathematische Objekte, die Veränderungen eines Zustands über die Zeit beschreiben können. Kontrollsysteme sind ebensolche Systeme, die zudem noch von einem Parameter $u \in \mathbb{R}^m$ abhängen. Dabei sind zwei Interpretationsansätze dieses Parameters denkbar. Dieser kann als Steuergröße, mit Hilfe derer das System von außen aktiv beeinflusst wird, gesehen werden, oder aber als Störung, die passiv auf das System einwirkt. Hier wird es nur um ersteren Ansatz, der zeitabhängigen Beeinflussung von Systemen von außen gehen. Dazu müssen wir zunächst die Menge der zulässigen Zeiten betrachten.

Definition 3.1 (Zeitachse).

Wir definieren die *Zeitachse* \mathbb{K} als Menge der ganzen Zahlen, also

$$\mathbb{K} := \mathbb{Z}$$

Definition 3.2 (Diskretes dynamisches System).

Sei \mathbb{K} eine Zeitachse und R ein metrischer Raum. Ein *diskretes dynamisches System* ist eine stetige Abbildung $\Phi : \mathbb{K} \times R \rightarrow R$ mit der Eigenschaft

$$\Phi(0, x) = x \text{ für alle } x \in R$$

und der *Kozykluseigenschaft*

$$\Phi(k + s, x) = \Phi(k, \Phi(s, x)) \text{ für alle } k, s \in \mathbb{K}$$

Definition 3.3 (Kontrollsystem in diskreter Zeit).

- (i) Ein *Kontrollsystem in diskreter Zeit* $\mathbb{K} = \mathbb{Z}$ im \mathbb{R}^d , $d \in \mathbb{N}$, ist gegeben durch die Differenzengleichung

$$x(k+1) = f(x(k), u(k)), \quad (3.1)$$

wobei $f : \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ eine stetige Abbildung ist.

- (ii) Die Menge $U \subseteq \mathbb{R}^m$ heißt *Kontrollwertebereich*, und definiert die Werte, die $u(k)$ für $k \in \mathbb{Z}$ annehmen darf. U ist hier üblicherweise kompakt.

- (iii) Mit \mathcal{U} bezeichnen wir den *Raum der zulässigen Kontrollfunktionen*, also

$$\mathcal{U} := \{u : \mathbb{Z} \rightarrow U \mid u \text{ zulässig}\}$$

Bemerkung 3.4.

Da wir hier nur den zeitdiskreten Fall betrachten, ist die Wahl des Raumes der zulässigen Kontrollfunktionen einfach. Wir lassen für \mathcal{U} alle möglichen Funktionen u von \mathbb{Z} nach U zu, also

$$\mathcal{U} := \{u : \mathbb{Z} \rightarrow U\}.$$

Per Induktion sieht man leicht, dass dann für jeden Anfangswert $x_0 \in \mathbb{R}^d$ und jede Funktion $u \in \mathcal{U}$ eine eindeutige Lösung $\Phi(k, x_0, u)$ von (3.1) in positiver Zeitrichtung existiert, also eine Funktion $\Phi : \mathbb{N}_0 \times \mathbb{R}^d \times \mathcal{U}$ mit

$$\Phi(0, x_0, u) = x_0 \text{ und } \Phi(k+1, x_0, u) = f(\Phi(k, x_0, u), u).$$

KAPITEL 4

Optimale Steuerung

In Kapitel 3 haben wir Kontrollsysteme in diskreter Zeit eingeführt, in die von außen eingegriffen werden kann. Wie in [5] ist es auch bei uns das Ziel solche Systeme optimal zu steuern. Wir wollen nun Kosten- bzw. Ertragsfunktionen definieren, die unter Berücksichtigung eines solchen Kontrollsystems minimiert, bzw. maximiert werden sollen. Solche Funktionen sollen jedem Punkt $(x, u) \in \mathbb{R}^d \times U$ im kombinierten Zustandskontrollraum einen Wert zuweisen. Summiert man diese Funktionen entlang einer Trajektorie und der zugehörigen Kontrollfunktion auf, so erhält man einen Wert, der von dem Anfangswert x_0 und der Kontrollfunktion u abhängt. Ziel soll es nun sein die Kontrollfunktion $u \in \mathcal{U}$ zu jedem Zeitpunkt optimal zu wählen um diesen Wert zu minimieren bzw. maximieren. Solche mathematischen Probleme nennt man *optimale Steuerungsprobleme*. Diese können nun auf unendlichem Horizont, oder aber wie in dieser Arbeit auf endlichem Horizont definiert werden. In der Praxis bedeutet das, dass ein bestimmtes Ziel innerhalb eines festen Zeitraums erreicht werden muss. In dieser Arbeit werden wir ein praktisches Beispiel der Reduzierung der CO₂-Emissionen betrachten. Wie wir in Kapitel 5 genau erläutern werden, soll hier durch Einsatz von finanziellen Mitteln innerhalb eines bestimmten Zeitraums ein gewisses Ziel an CO₂-Reduktion erreicht werden. Nun wollen wir aber zunächst den theoretischen Hintergrund der optimalen Steuerung bilden. Für die nötigen Definitionen verwenden wir [12] und leiten mit [16] und [17] dann die Bellman'sche Rekursionsformel her. Danach betrachten wir das Ganze aus spieltheoretischer Sicht.

4.1 Optimale Steuerung von Kontrollsystemen auf endlichem Horizont

Definition 4.1 (Gütefunktional auf endlichem Zeithorizont).

Betrachte ein Kontrollsystem (3.1). Wir definieren das *Gütefunktional auf endlichem Zeithorizont in diskreter Zeit* als

$$J(x, u) := \vartheta(x(K)) + \sum_{k=0}^{K-1} (\Phi(x(k), u(k))) \quad (4.1)$$

Definition 4.2 (Optimales Steuerungsproblem).

Das *optimale Steuerungsproblem* besteht nun darin, das Gütefunktional

$$J(x, u) := \vartheta(x(K)) + \sum_{k=0}^{K-1} (\Phi(x(k), u(k)))$$

unter Berücksichtigung der Prozessnebenbedingungen

$$x(k+1) = f(x(k), u(k)) \quad (4.2)$$

mit der Anfangsbedingung

$$x(0) = x_0$$

und der Endbedingung

$$g(x(K)) = \hat{x} \quad (4.3)$$

zu minimieren und damit die *optimale Wertefunktion*

$$V(x) := \inf_{u \in \mathcal{U}} J(x, u) \quad (4.4)$$

zu berechnen.

Bemerkung 4.3.

- (i) Statt zu minimieren kann - völlig analog - das entsprechende Maximierungsproblem $V(x) := \sup_{u \in \mathcal{U}} J(x, u)$ betrachtet werden. Da allerdings im Folgenden nur Beispiele für Minimierungsprobleme betrachtet werden, beschränken wir uns bei den folgenden Definitionen auf diesen Fall.
- (ii) Wir setzen hier nicht voraus, dass optimale Kontrollfunktionen $u \in \mathcal{U}$ existieren, deshalb verwenden wir *inf* statt *min*.

4.2 Bellman'sches Optimalitätsprinzip auf endlichem Horizont

Das *Bellman'sche Optimalitätsprinzip*, oder auch *Prinzip der Dynamischen Programmierung* ist die Basis für die analytische Berechnung der optimalen Wertefunktion. Es besagt, dass Endstücke optimaler Trajektorien wieder optimale Trajektorien sind. Die klassische Definition

„An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.“

R.E. Bellman, 1957 [3]

soll nun formalisiert werden.

Wir betrachten also ein optimales Steuerungsproblem der optimalen Überführung eines Systemzustands $x(0) = x_0$ in eine Endbedingung $g(x(K)) = \hat{x}$ unter Berücksichtigung von Restriktionen. Seien $u^*(k)$ und $x^*(k)$, $k \in [0, K]$, die optimalen Steuer- und Zustandstrajektorien. Dazu nun das von R.E. Bellman formulierte Optimalitätsprinzip:

Satz 4.4 (Optimalitätsprinzip).

Jede Resttrajektorie $u^*(k)$, $k \in [k_1, K]$, $0 \leq k_1 \leq K$ der optimalen Trajektorie $u^*(k)$, $k \in [0, K]$, ist optimal im Sinne der Überführung des Zwischenzustands $x^*(k_1)$ in die Endbedingung $g(x(K)) = \hat{x}$

Zur Verdeutlichung soll folgendes Schaubild dienen. Hier wird die optimale Zustandstrajektorie $x^*(k)$ in zwei Zeitintervalle zerlegt:

- Intervall 1: $k \in [0, k_1]$
- Intervall 2: $k \in [k_1, K]$

Das Optimalitätsprinzip besagt nun, dass die Teilstrecke Nummer 2 in der Zeichnung, die optimale Trajektorie ist, um $x^*(k_1)$ in die Endbedingung $g(x(K)) = \hat{x}$ zu überführen. Wäre dies nicht der Fall, so müsste es eine Trajektorie 3 geben, die von $x^*(k_1)$ ausgehend mit geringeren Kosten als Teilstrecke 2 das Steuerungsziel $g(x(K)) = \hat{x}$ erreicht. Dies wäre aber ein Widerspruch zur Optimalität der aus 1 und 2 bestehenden Trajektorie, da die aus 1 und 3 bestehende Trajektorie weniger Kosten verursachen müsste.

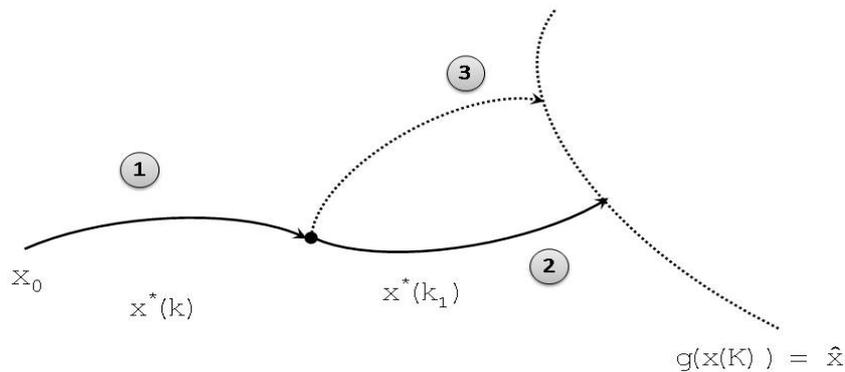


Abbildung 4.1: Bellman'sches Optimalitätsprinzip

Wir wollen nun bei der Berechnung der optimalen Wertefunktion auf eben dieses Prinzip der dynamischen Programmierung nach Bellman zurückgreifen. Dabei geht es um die Minimierung des bereits definierten zeitdiskreten Gütefunktional (4.1) unter Berücksichtigung eines Kontrollsystems (4.2) mit der Anfangsbedingung $x(0) = x_0$ und der Endbedingung $g(x(K)) = \hat{x}$, wobei K unser Endzeitpunkt ist. Für $k = 0, 1, \dots, K - 1$ kann diese Problemstellung als mehrstufiger Entscheidungsprozess gesehen werden.

Zur Herleitung der Bellman'schen Rekursionsformel benötigen wir den folgenden Satz. Dieser garantiert uns, ein Entscheidungsproblem mit zwei Entscheidungsvariablen in zwei Entscheidungsprobleme mit nur einer Variable aufteilen zu können.

Satz 4.5.

$$\inf_{a \in A, b \in B} \{f(a, b)\} = \inf_{a \in A} \left\{ \inf_{b \in B} \{f(a, b)\} \right\}$$

Beweis.

Der Beweis dazu steht in [16].

□

Definition 4.6 (Überführungskosten).

Wir definieren die *Überführungskosten* J_k um einen Zustand $x(k)$ in das geforderte Endziel $g(x(K)) = \hat{x}$ zu überführen wie folgt

$$J_k = \vartheta(x(K)) + \sum_{\kappa=k}^{K-1} \phi(x(\kappa), u(\kappa), \kappa)$$

Die minimalen Überführungskosten $J_k^* = \inf J_k$, die ausschließlich von dem zu überführenden Zustand $x(k)$ und dem Zeitpunkt k abhängen, bezeichnen wir mit $V(x(k), k)$. Sie lauten dann:

$$\begin{aligned} V(x(k), k) &= \inf J_k = \inf_{u(k), \dots, u(N-1)} \left\{ \vartheta(x(K)) + \sum_{\kappa=k}^{K-1} \phi(x(\kappa), u(\kappa), \kappa) \right\} \\ &= \inf_{u(k)} \left\{ \inf_{u(k+1), \dots, u(N-1)} \left\{ (\phi(x(k), u(k), k) + (x(K)) + \sum_{\kappa=k+1}^{K-1} \phi(x(\kappa), u(\kappa), \kappa)) \right\} \right\} \\ &= \inf_{u(k)} \left\{ \phi(x(k), u(k), k) + \inf_{u(k+1), \dots, u(N-1)} \left\{ (x(K)) + \sum_{\kappa=k+1}^{K-1} \phi(x(\kappa), u(\kappa), \kappa) \right\} \right\} \\ &= \inf_{u(k)} \{ \phi(x(k), u(k), k) + V(x(k+1), k+1) \} \end{aligned}$$

Da $x(k+1)$ aus einem Kontrollsystem (4.2) entsteht, ergibt sich

$$V(x(k), k) = \inf_{u(k)} \{ \phi(x(k), u(k), k) + V(f(x(k), u(k)), k+1) \} \quad (4.5)$$

Die rechte Seite dieser Gleichung bezeichnen wir als *Bellman'sche Rekursionsformel*. Diese ist nur von $u(k)$ abhängig, nicht aber von $u(\kappa)$, $\kappa = k+1, \dots, K-1$. Die Minimierung geschieht folglich nur über die Steuergrößen des Zeitpunktes k , $u(k)$. Wird diese einstufige Minimierung vom Endzeitpunkt angefangen für $k = K-1, K-2, \dots, 0$ nacheinander durchgeführt, so kann der ursprüngliche mehrstufige Entscheidungsprozess in K einstufige Entscheidungsprobleme zerlegt werden. Diese stufengebundene Vorgehensweise heißt *Dynamische Programmierung* und wird nun Schritt für Schritt erläutert.

• **Stufe $K-1$:**

Für alle Zustände $x(K-1)$ sollen die zugehörigen Steuerungen $u(K-1)$ bestimmt werden, die

$$J_{K-1} = \Phi(x(K-1), u(K-1), K-1) + \vartheta(x(K))$$

unter Berücksichtigung der Differenzgleichungen des zugehörigen Kontrollsystems

$$x(K) = f(x(K-1), u(K-1))$$

und der Endbedingung des Systems $g(x(K)) = \hat{x}$ minimieren. Den minimierenden Wert bezeichnen wir mit

$$u(K-1) = \mathcal{R}(x(K-1), K-1)$$

Die minimalen Werte von J_{K-1} seien nun definitionsgemäß mit

$$V(x(K-1), K-1)$$

bezeichnet.

- **Stufe $K-2$:**

Für alle Zustände $x(K-2)$ sollen nun die zugehörigen Steuerungen $u(K-2)$ bestimmt werden, die

$$\tilde{J}_{K-2} = \Phi(x(K-2), u(K-2), K-2) + V(x(K-1), K-1)$$

wieder unter Berücksichtigung des zugehörigen Kontrollsystems

$$x(K-1) = f(x(K-2), u(K-2))$$

minimieren. Den minimierenden Wert bezeichnen wir mit

$$u(K-2) = \mathcal{R}(x(K-2), K-2)$$

und die minimalen Werte von \tilde{J}_{K-2} mit

$$V(x(K-2), K-2)$$

⋮

- **Stufe 0:**

Im letzten Schritt soll für jeden Startzustand $x(0)$ das zugehörige $u(0)$ bestimmt werden, das

$$\tilde{J}_0 = \Phi(x(0), u(0), 0) + V(x(1), 1)$$

unter Berücksichtigung der Differenzgleichung des zugehörigen Kontrollsystems

$$x(1) = f(x(0), u(0))$$

minimiert. Somit ist unser letzter minimierender Wert

$$u(0) = \mathcal{R}(x(0), 0)$$

und die minimalen Werte von \tilde{J}_0 sind

$$V(x(0), 0)$$

Diese entsprechen zugleich der gewünschten optimalen Wertefunktion unseres optimalen Steuerungsproblems.

Am Ende dieser K -stufigen Prozedur erhalten wir also sowohl die optimalen Trajektorien der Überführung des Anfangszustands $x(0) = x_0$ in die geforderte Endbedingung $x(K) = \hat{x}$, als auch ein *optimales Regelgesetz*, das durch die einstufigen Minimierungsergebnisse

$$u(k) = \mathcal{R}(x(k), k)$$

für $k = 0, \dots, K - 1$ gegeben ist.

Bemerkung 4.7.

Mit der Information aus dem Regelgesetz können wir nicht nur einen festen Startzustand $x(0) = x_0$, sondern jeden beliebigen Punkt $x(k)$ für $k = 0, \dots, K - 1$ optimal in unser Ziel (4.3) überführen. In der Praxis bedeutet das, dass bei einer unvorhergesehenen Zustandsänderung nicht alle Optimierungsschritte erneut durchgeführt werden müssen, sondern, dass für jedes $x(k)$ für $k = 0, \dots, K - 1$ mittels des Regelgesetzes immer noch eine optimale Überführung in den Zielzustand gewährleistet ist.

4.3 Optimale Steuerung bei dynamischen Spielen

Im nun folgenden Abschnitt werden wir die eben gewonnenen Erkenntnisse der optimalen Steuerung und des Bellman'schen Optimalitätsprinzips auf spieltheoretische Fragestellungen anwenden. Dabei führen wir zunächst den Begriff der dynamischen Spiele ein. Wir werden dann den Begriff der nicht-antizipierenden Strategien aus [1] vorstellen und eine Version des Bellman'schen Optimalitätsprinzips für zwei Spieler bei ebensolcher Strategie angeben.

Zunächst allerdings zu den dynamischen Spielen selbst.

Bemerkung 4.8.

Den dynamischen Spielen liegen Gütefunktionale (4.1) zugrunde, deshalb müssen wir auch hier die optimalen Wertefunktionen (4.4) berechnen. Ziel eines Spielers kann es also auch hier wieder sein, entweder die eigenen Kosten zu minimieren, oder den eigenen Nutzen zu maximieren. Die folgenden dynamischen Spiele sind Minimierungsprobleme, deshalb werden die Definitionen auch für diese angegeben; allerdings ist auch hier wieder eine analoge Behandlung von Maximierungsproblemen möglich.

Definition 4.9 (Optimale Wertefunktion für einen Spieler).

Betrachte ein optimales Steuerungsproblem aus Definition 4.2. Mit unserem Gütefunktional für einen Spieler

$$J(x, u) := \vartheta(x(K)) + \sum_{k=0}^{K-1} (\Phi(x(k), u(k)))$$

definieren wir die *optimale Wertfunktion für einen Spieler* mit

$$V(x) := \inf_{u \in \mathcal{U}} J(x, u)$$

Definition 4.10 (Dynamisches Spiel).

Ein *Dynamisches Spiel* besteht nun darin, die optimale Wertfunktion des Spielers zu bestimmen und die zugehörigen optimalen Strategien zu berechnen.

4.3.1 Nicht-antizipierende Strategien

Gerade haben wir ein dynamisches Spiel mit nur einem Spieler kennengelernt, der alleine sein Ziel der Minimierung seines Gütefunktional verfolgt. Es ist aber gerade interessant zu sehen, welche Entscheidungen Spieler treffen können, die im Wettbewerb zueinander stehen, während sie gleichzeitig versuchen ihr eigenes Ziel zu verwirklichen. Dynamische Spiele mit zwei gegnerischen Spielern, sollen nun vorgestellt werden. Zudem soll dabei deren zugrunde liegende Informationsstruktur eine gewichtige Rolle spielen. Um möglichst realistische Szenarien betrachten zu können, sollen die Spieler ihre Entscheidungen treffen, ohne die zukünftigen Entscheidungen ihres Gegners zu kennen. Dafür soll es aber möglich sein, dass ein Spieler die Entscheidung seines Gegners kennt und darauf basierend seine eigene trifft. Diese spezielle Informationsstruktur liegt den sogenannten *nicht-antizipierenden Strategien* zugrunde. Spieler 1 hat dabei die *Strategie* μ_1 , Spieler 2 die *Kontrolle* u_2 zur Verfügung; diese stehen in folgender Beziehung.

Definition 4.11 (Nicht-antizipierende Strategie).

Sei ein Kontrollsystem in diskreter Zeit (3.1) mit zwei Spielern gegeben.

- (i) Die Strategie von Spieler 1 ist eine Abbildung $\mu_1 : \mathcal{U}_2 \rightarrow \mathcal{U}_1$. Sie heißt *nicht-antizipierend*, wenn für jedes $t > 0$ und alle $u_2, \tilde{u}_2 \in \mathcal{U}_2$, $u_2(s) = \tilde{u}_2(s)$ für alle $s \leq t$ die Gleichheit $\mu_1[u_2](s) = \mu_1[\tilde{u}_2](s)$ für alle $s \leq t$ impliziert.
- (ii) Die *Menge aller nicht-antizipierenden Strategien* von Spieler 1 bezeichnen wir mit Δ_1 . Formal lautet diese:

$$\Delta_1 := \{ \mu_1 : \mathcal{U}_2 \rightarrow \mathcal{U}_1 \mid u_2(s) = \tilde{u}_2(s) \forall s \leq t \Rightarrow \mu_1[u_2](s) = \mu_1[\tilde{u}_2](s) \forall s \leq t \}$$

Um diese Definition besser zu erläutern, hier zwei konträre Beispielstrategien. Zunächst wird eine nicht-antizipierende Strategie gegeben. Anschließend sehen wir eine Strategie, die obiger Definition gerade nicht genügt.

Beispiel 4.12 (Eine nicht-antizipierende Strategie).

Ein Beispiel für eine nicht-antizipierende Strategie ist gegeben durch

$$\mu_1[u_2](s) = \Phi(u_2(s))$$

für eine Abbildung $\Phi : U_2 \rightarrow U_1$.

Beispiel 4.13 (Keine nicht-antizipierende Strategie).

Ein Beispiel, das Definition 4.2 nicht erfüllt ist gegeben durch

$$\mu_1 [u_2] (s) = \Phi(u_2(s + \tau))$$

für ein festes $\tau > 0$ für eine Abbildung $\Phi : U_2 \rightarrow U_1$.

Nun soll es um die Optimierung dieser Strategien bzw. Kontrollen der beiden Spieler gehen. Dazu wählt Spieler 2 für jede Strategie μ_1 von Spieler 1 als optimale Antwort eine Kontrolle u_2 . Diese Menge der optimalen Antworten von Spieler 2 wollen wir nun definieren.

Definition 4.14 (Menge der optimalen Antworten von Spieler 2).

Betrachte ein dynamisches Spiel. Die Menge $\mathcal{U}_2^*(\mu_1) \in \mathcal{U}_2$ ist für jedes x_2 und jedes $\mu_1 \in \Delta_1$ die Menge der optimalen Antworten von Spieler 2 und wird definiert durch

$$\mathcal{U}_2^*(\mu_1) := \left\{ u_2 \in \mathcal{U}_2 \mid u_2 = \arg \min_{u_2 \in \mathcal{U}_2} J_2(x_2, \mu_1 [u_2], u_2) \right\}$$

Aus dieser Menge wählt Spieler 2 für jede Strategie μ_1 von Spieler 1 eine optimale Antwort $u_2^*(\mu_1)$ aus. Damit können wir nun die optimale Strategie von Spieler 1 angeben.

Definition 4.15 (Optimale Strategie von Spieler 1).

Betrachte ein dynamisches Spiel. Eine nicht-antizipierende Strategie $\mu_1 \in \Delta_1$ von Spieler 1 ist für jedes x_1 optimal, wenn sie das Funktional von Spieler 1 für jedes $u_2^*(\mu_1)$ minimiert, also

$$\mu_1^* = \mu_1^* [u_2^*(\mu_1^*)] = \arg \min_{\mu_1 \in \Delta_1} J_1(x_1, \mu_1 [u_2^*(\mu_1)], u_2^*(\mu_1))$$

Mit Definition 4.15 erhalten wir nun auch die optimale Kontrolle von Spieler 2.

Definition 4.16 (Optimale Kontrolle von Spieler 2).

Wir definieren die optimale Kontrolle von Spieler 2 als die optimale Antwort von Spieler 2 zur optimalen Strategie von Spieler 1. Sie lautet also $u_2^*(\mu_1^*)$.

Zum besseren Verständnis dieser sehr abstrakten Definition wollen wir nun ein einfaches Spiel mit nicht-antizipierender Spielsituation betrachten.

Beispiel 4.17 (Spiel mit nicht-antizipierender Spielsituation).

Gegeben sei ein Spiel mit zwei Spielern, wobei Spieler 1 eine nicht-antizipierende Strategie μ_1 und Spieler 2 eine Kontrolle u_2 zur Verfügung hat. Wir nehmen nun vereinfachend an, dass sich Spieler 1 zwischen $\mu_1 = a \cdot u_2$ und $\mu_1 = b \cdot u_2$ mit $a, b \in \mathbb{N}$ entscheiden kann, während Spieler 2 darauf mit $u_2 = 1$ oder $u_2 = -1$ reagieren kann.

Spieler 2 wählt also zunächst zu jeder Strategie μ_1 von Spieler 1 eine optimale Antwort aus:

- $u_2^*(\mu_1 = a \cdot u_2^*(\mu_1)) = 1$
- $u_2^*(\mu_1 = b \cdot u_2^*(\mu_1)) = -1$

Spieler 1 wählt nun die für sich optimale Strategie $\mu_1^* [u_2^*(\mu_1^*)] = b \cdot u_2^*(\mu_1^*)$. Damit ist dann die optimale Kontrolle von Spieler 2 $u_2^*(\mu_1^*) = -1$ und der Wert der optimalen Strategie von Spieler 1 ist $\mu_1^* [u_2^*(\mu_1^*)] = -b$.

Diesen Spielablauf bei nicht-antizipierender Strategie von Spieler 1 soll diese Grafik veranschaulichen.

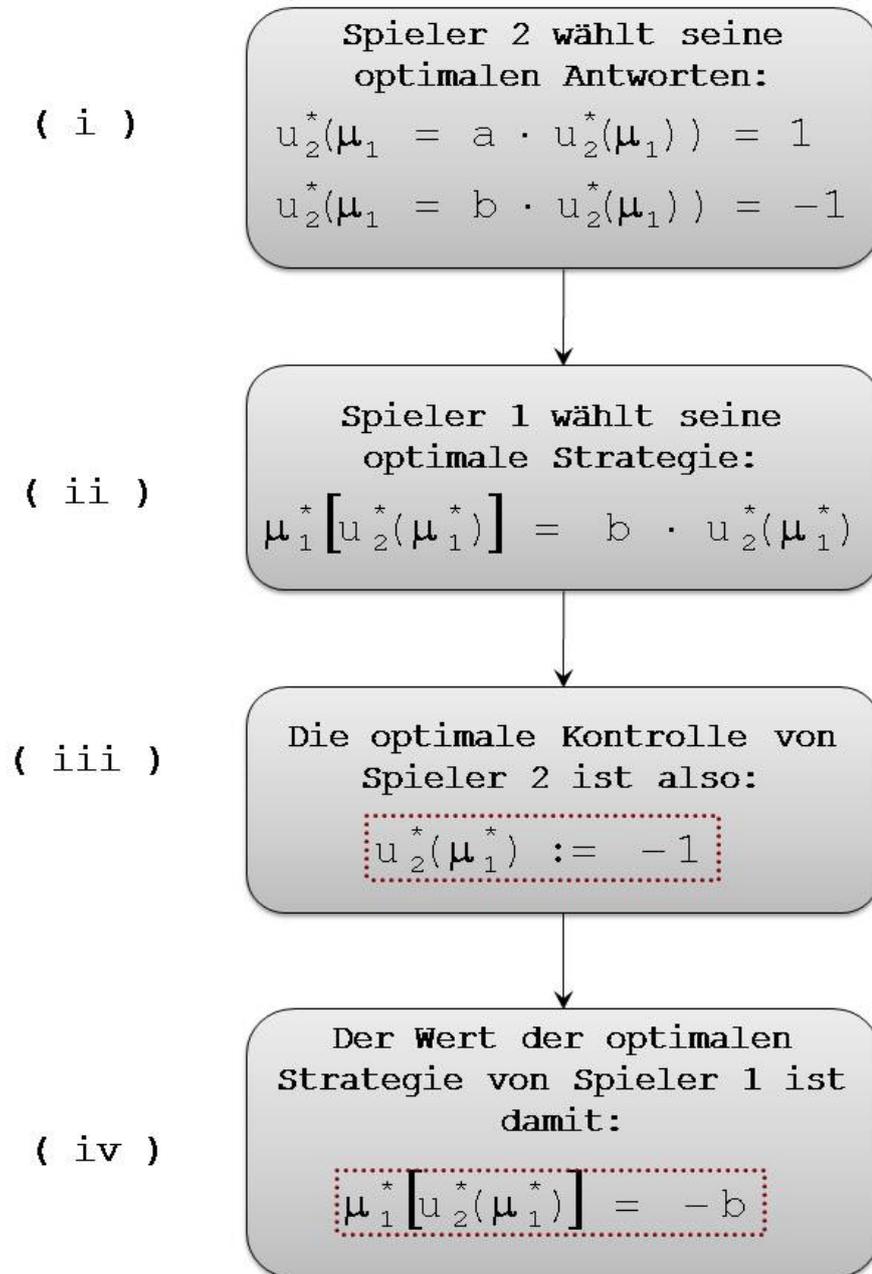


Abbildung 4.2: Darstellung des Spiels mit nicht-antizipierender Spielsituation

Bemerkung 4.18.

Aus Gründen der Übersichtlichkeit schreiben wir für die optimale Kontrolle von Spieler 2 im Folgenden statt $u_2^*(\mu_1^*)$ kurz u_2^* .

Wie in Definition 4.10 eingeführt, ist es das Ziel eines dynamischen Spiels die optimale Wertefunktion und die optimalen Trajektorien eines Spielers zu berechnen. Da zwei Spieler mit zwei Gütefunktionalen an unserem dynamischen Spiel beteiligt sind, existieren auch zwei optimale Wertefunktionen. Mit Definition 4.15 und Definition 4.16 ergeben sich also für den Fall der nicht-antizipierenden Spielsituation folgende optimale Wertefunktionen der Spieler.

Definition 4.19 (Optimale Wertefunktionen bei nicht-antizipierender Spielsituation).

Gegeben seien die Gütefunktionale der beiden Spieler

$$J_i(x_i, \mu_1[u_2], u_2) := \vartheta_i(x_i(K)) + \sum_{k=0}^{K-1} (\Phi_i(x_i(k), \mu_1[u_2](k), u_2(k)))$$

mit $i = 1, 2$

Bei nicht-antizipierender Spielsituation definieren wir

(i) die *optimale Wertefunktion von Spieler 1* durch

$$V_1(x_1) := J_1(x_1, \mu_1^*[u_2^*], u_2^*)$$

mit optimaler Strategie μ_1^* aus Definition 4.15 und optimaler Kontrolle u_2^* aus Definition 4.16

(ii) die *optimale Wertefunktion von Spieler 2* durch

$$V_2(x_2) := J_2(x_2, \mu_1^*[u_2^*], u_2^*)$$

mit optimaler Strategie μ_1^* aus Definition 4.15 und optimaler Kontrolle u_2^* aus Definition 4.16

Bemerkung 4.20.

Beachte, wenn μ_1^* und u_2^* in den Punkten x_1 für Spieler 1 bzw. x_2 für Spieler 2 optimal nach den Definitionen 4.15 und 4.16 optimal sind, gilt für die optimalen Wertefunktionen $V_1(x_1)$ und $V_2(x_2)$ der beiden Spieler

$$V_1(x_1) = J_1(x_1, \mu_1^*[u_2^*], u_2^*) \leq J_1(x_1, \mu_1[u_2^*], u_2^*)$$

und

$$V_2(x_2) = J_2(x_2, \mu_1^*[u_2^*], u_2^*) \leq J_2(x_2, \mu_1^*[u_2], u_2)$$

für alle $u_2 \in \mathcal{U}_2$ und $\mu_1 \in \Delta_1$.

Das bedeutet, dass Spieler 1 und Spieler 2 ihre minimalen Kosten $V_1(x_1)$ bzw. $V_2(x_2)$ durch eine einseitige Abweichung von ihrer optimalen Strategie, bzw. optimalen Kontrolle nicht weiter verringern können. Sie können sich nur schlechter stellen als zuvor. Das Paar (μ_1^*, u_2^*) bestehend aus der optimalen nicht-antizipierenden Strategie von Spieler 1 und der optimalen Kontrolle von Spieler 2 ist damit ein Nash-Gleichgewicht aus unserer anfänglichen Definition 2.7.

4.3.2 Bellman'sches Optimalitätsprinzip bei nicht-antizipierenden Strategien

Nun wollen wir das in Abschnitt 4.2 eingeführte Optimalitätsprinzip nach Bellman auf dynamische Spiele mit nicht-antizipierender Strategie anwenden. Das heißt es existieren zwei Funktionale auf die beide Spieler steuernd einwirken können und damit auch das System des jeweiligen Gegners beeinflussen können. Spieler 1 soll dabei wieder die Möglichkeit haben eine nicht-antizipierende Strategie $\mu_1[u_2]$ zu bestimmen, während Spieler 2 dagegen nur eine Kontrolle u_2 zu gegebener Strategie $\mu_1[u_2]$ wählen kann. Zur Herleitung der zugehörigen Bellman'schen Rekursionsformeln definieren wir die Überführungskosten bei nicht-antizipierender Strategie wie folgt.

Definition 4.21 (Überführungskosten bei nicht-antizipierender Strategie).

Wir definieren die *Überführungskosten bei nicht-antizipierender Strategie* J_i^k zur Überführung eines Zustandes $x_i(k)$ in das geforderte Endziel $g_i(x_i(K)) = \hat{x}_i$ für $i = 1, 2$ wie folgt für Spieler 1 als

$$J_1^k = \vartheta_1(x_1(K)) + \sum_{\kappa=k}^{K-1} \phi_1(x_1(\kappa), \mu_1[u_2](\kappa), u_2(\kappa), \kappa)$$

und für Spieler 2 als

$$J_2^k = \vartheta_2(x_2(K)) + \sum_{\kappa=k}^{K-1} \phi_2(x_2(\kappa), \mu_1[u_2](\kappa), u_2(\kappa), \kappa)$$

Wir wollen nun die Überführungskosten der beiden Spieler optimieren. Dafür benötigen wir die optimale Strategie von Spieler 1 aus Definition 4.15 und die optimale Kontrolle von Spieler 2 aus Definition 4.16. Die Berechnung der optimalen Strategie von Spieler 1 erfolgt dann in drei Teilschritten, die in der nächsten Definition angegeben werden.

Definition 4.22 (Optimierung bei nicht-antizipierender Spielsituation).

Wir definieren den dreistufigen Optimierungsprozess für alle Zustände ab dem Zeitpunkt k bestehend aus:

- (i) **Berechnung der, nach Definition 4.15 optimalen Strategie $\mu_1^*(k)$ von Spieler 1**

- (ii) **Berechnung der, nach Definition 4.16 optimalen Kontrolle $u_2^*(k)$ von Spieler 2**
 (iii) **Einsetzen der optimalen Kontrolle $u_2^*(k)$ um die optimale Strategie $\mu_1^*[u_2^*](k)$ von Spieler 1 zu erhalten**
 als *Optimierung bei nicht-antizipierender Spielsituation* und schreiben dafür $\text{opt}(\dots)$.

Bemerkung 4.23.

Bei dem Optimierungsprozess aus Definition 4.22 wird für alle Zustände ab dem Zeitpunkt k eine optimale Strategie $\mu_1^*(k)$ von Spieler 1 berechnet. Diese ist jeweils nur abhängig von der optimalen Kontrolle $u_2^*(k)$, nicht aber von zukünftigen optimalen Kontrollen von Spieler 2. Damit erfüllt unsere optimale Strategie $\mu_1^*[u_2^*](k)$ die Definition 4.11 einer nicht-antizipierenden Strategie.

Wir wollen nun für dieses Optimierungsproblem die Bellman'schen Rekursionsformeln aufstellen. Dazu bezeichnen wir die optimalen Kosten von Spieler 1 mit $V_1(x_1(k), k)$ und die von Spieler 2 mit $V_2(x_2(k), k)$. Diese sind jeweils wieder ausschließlich von dem zu überführenden Zustand $x_i(k)$ für $i = 1, 2$ und dem Zeitpunkt k abhängig.

Wir erhalten für Spieler 1

$$V_1(x_1(k), k) = \text{opt}(J_1^k)$$

Auch hier betrachten wir ein optimales Steuerungsproblem der optimalen Überführung eines Systemzustands $x_i(0) = x_{i0}$ in eine Endbedingung $g_i(x_i(K)) = \hat{x}_i$ für $i = 1, 2$ unter Berücksichtigung von Restriktionen. Dazu wollen das Optimalitätsprinzip von Bellman betrachten. Auch für unsere optimalen nicht-antizipierenden Strategien von Spieler 1 gilt dann:

Satz 4.24 (Optimalitätsprinzip für nicht-antizipierende Strategien).

Für $x_1(k)$ sei μ_1^* optimal auf dem Intervall $[0, K]$. Sei $\tilde{\mu}_1^*$ nun optimal für $x_1(k_1)$ mit $0 \leq k_1 \leq K$ auf dem Intervall $[k_1, K]$. Dann gilt $\mu_1^*|_{[k_1, K]} = \tilde{\mu}_1^*$.

Wenden wir dieses Prinzip an, so erhalten wir für Spieler 1

$$V_1(x_1(k), k) = \text{opt}_{\mu_1(k)} \{ \phi_1(x_1(k), \mu_1[u_2](k), u_2(k), k) + \\ + V_1(x_1(k+1), k+1) \}$$

Da $x_1(k+1)$ aus einem Kontrollsystem (4.2) entsteht, ergibt sich dann

$$V_1(x_1(k), k) = \text{opt}_{\mu_1(k)} \{ \phi_1(x_1(k), \mu_1[u_2](k), u_2(k), k) + \\ + V_1(f_1(x_1(k), \mu_1[u_2](k), u_2(k)), k+1) \} \quad (4.6)$$

Spieler 2 muss für sich hingegen ein Minimierungsproblem mit vorgegebener optimaler Strategie μ_1^* von Spieler 1 lösen. Analog zur Herleitung der Rekursionsformel (4.5) ergibt sich für das Minimierungsproblem von Spieler 2.

$$V_2(x(k), k) = \inf_{u_2(k)} \{ \phi_2(x_2(k), \mu_1^*[u_2](k), u_2(k), k) + V_2(x_2(k+1), k+1) \}$$

Da $x_2(k+1)$ aus einem Kontrollsystem (4.2) entsteht, ergibt sich dann

$$V_2(x_2(k), k) = \inf_{u_2(k)} \{ \phi_2(x_2(k), \mu_1^*[u_2](k), u_2(k), k) + V_2(f_2(x_2(k), \mu_1^*[u_2](k), u_2(k)), k+1) \} \quad (4.7)$$

Die rechten Seiten der beiden Gleichungen (4.6) und (4.7) bezeichnen wir als *Bellman'sche Rekursionsformeln für zwei Spieler bei nicht-antizipierender Spielsituation*. Mit diesen Bellman'schen Rekursionsformeln und der oben erläuterten Optimierung können wir nun die dynamische Programmierung durchführen.

• **Stufe $K - 1$:**

(i) **Berechnung der optimalen Strategie $\mu_1^*(K - 1)$ von Spieler 1:**

Spieler 1 möchte seine optimale Strategie $\mu_1^*(K - 1)$ für alle Zustände $x_1(K - 1)$ berechnen. Dazu muss er das Funktional

$$J_1^{K-1} = \Phi_1(x_1(K - 1), \mu_1[u_2](K - 1), u_2(K - 1), K - 1) + \vartheta_1(x_1(K))$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_1(K) = f_1(x_1(K - 1), \mu_1[u_2](K - 1), u_2(K - 1))$$

und der Endbedingung des Systems $x_1(K) = \hat{x}_1$ nach $\mu_1(K - 1)$ minimieren. Er erhält daraus dann seine optimale Strategie $\mu_1^*(K - 1)$. Diese entspricht einer Funktion, die von dem zu überführenden Zustand $x_1(K - 1)$, dem Zeitpunkt $K - 1$ und der Kontrolle $u_2(K - 1)$ von Spieler 2 abhängig ist.

(ii) **Berechnung der optimalen Kontrolle $u_2^*(K - 1)$ von Spieler 2**

Spieler 2 möchte nun seine optimale Kontrolle $u_2^*(K - 1)$ für alle Zustände $x_2(K - 1)$ berechnen, dabei muss er aber die eben berechnete optimale Strategie $\mu_1^*(K - 1)$ von Spieler 1 berücksichtigen. Er minimiert dazu das Funktional

$$J_2^{K-1} = \Phi_2(x_2(K - 1), \mu_1^*[u_2](K - 1), u_2(K - 1), K - 1) + \vartheta_2(x_2(K))$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_2(K) = f_2(x_2(K-1), \mu_1^*[u_2](K-1), u_2(K-1))$$

und der Endbedingung des Systems $x_2(K) = \hat{x}_2$ nach $u_2(K-1)$. Dies ergibt seine optimale Kontrolle $u_2^*(K-1)$. Diese Funktion ist dann nur von dem zu überführenden Zustand $x_2(K-1)$ und dem Zeitpunkt $K-1$ abhängig.

(iii) **Einsetzen der optimalen Kontrolle $u_2^*(K-1)$ um die optimale Strategie $\mu_1^*[u_2^*](K-1)$ von Spieler 1 zu erhalten**

Spieler 1 kann nun die optimale Kontrolle $u_2^*(K-1)$ von Spieler 2 in seine Funktion der optimalen Strategie $\mu_1^*(K-1)$ einsetzen und erhält dann eine Funktion $\mu_1^*[u_2^*](K-1)$ die nur noch von dem zu überführenden Zustand $x_1(K-1)$ und dem Zeitpunkt $K-1$ abhängig ist.

Die Ergebnisse aus dieser Optimierung bezeichnen wir für Spieler 1 mit

$$\mu_1^*[u_2^*](K-1) = \mathcal{R}_1(x_1(K-1), K-1)$$

und für Spieler 2 mit

$$u_2^*(K-1) = \mathcal{R}_2(x_2(K-1), K-1)$$

Die optimalen Werte von J_i^{K-1} für $i = 1, 2$ seien nun definitionsgemäß für Spieler 1 mit $V_1(x_1(K-1), K-1)$ und für Spieler 2 mit $V_2(x_2(K-1), K-1)$ bezeichnet.

• **Stufe $K-2$:**

(i) **Berechnung der optimalen Strategie $\mu_1^*(K-2)$ von Spieler 1:**

Spieler 1 möchte seine optimale Strategie $\mu_1^*(K-2)$ für alle Zustände $x_1(K-2)$ berechnen. Dazu muss er das Funktional

$$\begin{aligned} \tilde{J}_1^{K-2} = & \Phi_1(x_1(K-2), \mu_1[u_2](K-2), u_2(K-2), K-2) + \\ & + V_1(x_1(K-1), K-1) \end{aligned}$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_1(K-1) = f_1(x_1(K-2), \mu_1[u_2](K-2), u_2(K-2))$$

nach $\mu_1(K-2)$ minimieren. Er erhält daraus dann seine optimale Strategie $\mu_1^*(K-2)$. Diese entspricht einer Funktion, die von dem zu

überführenden Zustand $x_1(K-2)$, dem Zeitpunkt $K-2$ und der Kontrolle $u_2(K-2)$ von Spieler 2 abhängig ist.

(ii) **Berechnung der optimalen Kontrolle $u_2^*(K-2)$ von Spieler 2**

Spieler 2 möchte nun seine optimale Kontrolle $u_2^*(K-2)$ für alle Zustände $x_2(K-2)$ berechnen, dabei muss er aber die eben berechnete optimale Strategie $\mu_1^*(K-2)$ von Spieler 1 berücksichtigen. Er minimiert dazu das Funktional

$$\begin{aligned} \tilde{J}_2^{K-2} = & \Phi_2(x_2(K-2), \mu_1^*[u_2](K-2), u_2(K-2), K-2) + \\ & + V_2(x_2(K-2), K-2) \end{aligned}$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_2(K-1) = f_2(x_2(K-2), \mu_1^*[u_2](K-2), u_2(K-2))$$

nach $u_2(K-2)$. Das ergibt seine optimale Kontrolle $u_2^*(K-2)$. Diese Funktion ist nur von dem zu überführenden Zustand $x_2(K-2)$ und dem Zeitpunkt $K-2$ abhängig.

(iii) **Einsetzen der optimalen Kontrolle $u_2^*(K-2)$ um die optimale Strategie $\mu_1^*[u_2^*](K-2)$ von Spieler 1 zu erhalten**

Spieler 1 kann nun die optimale Kontrolle $u_2^*(K-2)$ von Spieler 2 in seine Funktion der optimalen Strategie $\mu_1^*(K-2)$ einsetzen und erhält dann eine Funktion $\mu_1^*[u_2^*](K-2)$ die nur noch von dem zu überführenden Zustand $x_1(K-1)$ und dem Zeitpunkt $K-1$ abhängig ist.

Die Ergebnisse aus dieser Optimierung bezeichnen wir für Spieler 1 mit

$$\mu_1^*[u_2^*](K-2) = \mathcal{R}_1(x_1(K-2), K-2)$$

und für Spieler 2 mit

$$u_2^*(K-2) = \mathcal{R}_2(x_2(K-2), K-2)$$

Die optimalen Werte von \tilde{J}_i^{K-2} für $i = 1, 2$ seien nun wieder definitionsgemäß für Spieler 1 mit $V_1(x_1(K-2), K-2)$ und für Spieler 2 mit $V_2(x_2(K-2), K-2)$ bezeichnet.

⋮

- **Stufe 0:**

- (i) **Berechnung der optimalen Strategie $\mu_1^*(0)$ von Spieler 1:**

Im letzten Schritt möchte Spieler 1 seine optimale Strategie $\mu_1^*(0)$ für jeden Startzustand $x_1(0)$ berechnen. Dazu muss er das Funktional

$$\tilde{J}_1^0 = \Phi_1(x_1(0), \mu_1[u_2](0), u_2(0), 0) + V_1(x_1(1), 1)$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_1(1) = f_1(x_1(0), \mu_1[u_2](0), u_2(0))$$

nach $\mu_1(0)$ minimieren. Er erhält daraus seine optimale Strategie $\mu_1^*(0)$. Diese ist von seinem Startzustand $x_1(0)$, dem Zeitpunkt 0 und der Kontrolle $u_2(0)$ von Spieler 2 abhängig.

- (ii) **Berechnung der optimalen Kontrolle $u_2^*(0)$ von Spieler 2**

Im letzten Schritt möchte Spieler 2 seine optimale Kontrolle $u_2^*(0)$ für jeden Startzustand $x_2(0)$ berechnen. Dazu muss er das Funktional

$$\tilde{J}_2^0 = \Phi_2(x_2(0), \mu_1^*[u_2](0), u_2(0), 0) + V_2(x_2(1), 1)$$

unter Berücksichtigung der Differenzgleichungen des Kontrollsystems

$$x_2(1) = f_2(x_2(0), \mu_1^*[u_2](0), u_2(0))$$

nach u_2 minimieren. Dies ergibt seine optimale Kontrolle $u_2^*(0)$. Diese Funktion ist nur vom Startzustand $x_2(0)$ und dem Zeitpunkt 0 abhängig.

- (iii) **Einsetzen der optimalen Kontrolle $u_2^*(0)$ um die optimale Strategie $\mu_1^*[u_2^*](0)$ von Spieler 1 zu erhalten**

Spieler 1 kann nun die optimale Kontrolle $u_2^*(0)$ von Spieler 2 in seine Funktion der optimalen Strategie $\mu_1^*(0)$ einsetzen und erhält dann seine optimale Strategie $\mu_1^*[u_2^*](0)$ zum Zeitpunkt 0, die nur noch von seinem Startwert $x_1(0)$ und dem Zeitpunkt 0 abhängt.

Die Ergebnisse der letzten Optimierung bezeichnen wir für Spieler 1 mit

$$\mu_1^*[u_2^*](0) = \mathcal{R}_1(x_1(0), 0)$$

und für Spieler 2 mit

$$u_2^*(0) = \mathcal{R}_2(x_2(0), 0)$$

Als optimale Werte von \tilde{J}_i^0 für $i = 1, 2$ erhalten wir mit $V_1(x_1(0), 0)$ die optimale Wertefunktion von Spieler 1 und mit $V_2(x_2(0), 0)$ die optimale Wertefunktion von Spieler 2, was der Lösung unseres dynamischen Spiels entspricht.

Wie im Fall des Minimierungsproblems, liefert uns die dynamische Programmierung auch hier nicht nur die Trajektorien zur Überführung eines Anfangszustands $x_i(0) = x_i^0$ in die geforderte Endbedingung $x_i(K) = \hat{x}_i$ für $i = 1, 2$, sondern auch das folgende *optimale Regelgesetz für Spieler 1*

$$\mu_1^*[u_2^*](k) = \mathcal{R}_1(x_1(k), k)$$

und das folgende *optimale Regelgesetz für Spieler 2*

$$u_2^*(k) = \mathcal{R}_2(x_2(k), k)$$

für $k = 0, \dots, K - 1$

Das gibt uns also auch hier die Möglichkeit nicht nur den Startzustand von Spieler 1 x_1^0 und den Startzustand von Spieler 2 x_2^0 , sondern jeden Punkt $x_i(k)$ für $k = 0, \dots, K - 1$ optimal in das jeweilige Ziel des Spielers \hat{x}_i für $i = 1, 2$ zu steuern.

Ein spieltheoretisches Modell zur CO₂-Reduktion

5.1 Motivation

In diesem Abschnitt sollen nun die theoretischen Erkenntnisse auf ein mehr denn je aktuelles Thema angewendet werden, der Reduzierung des Ausstoßes von Treibhausgasen wie zum Beispiel Kohlenstoffdioxid (CO₂). Dies stellt eine die Ländergrenzen überschreitende Problematik dar. Deshalb wurde am 11. Dezember 1997 im Rahmen der Klimakonferenz der Vereinten Nationen in Kyoto eine Vereinbarung [20] getroffen um die damit verbundenen negativen Folgen für die gesamte Umwelt einzudämmen. Ziel ist es die künftigen Emissionen der beteiligten Länder zu senken. Das sogenannte Kyoto-Protokoll sieht vor, den jährlichen Treibhausgas-Ausstoß der Industrieländer bis zum Jahr 2012 um mindestens 5 Prozent gegenüber 1990 zu reduzieren. Eben dieser Sachverhalt soll in der vorliegenden Arbeit anhand eines Modells zur CO₂-Reduktion mit Hilfe der Theorie der dynamischen Spiele behandelt werden.

Als kurze Einführung möchte ich zunächst das sogenannte Umweltdilemma aus [22] vorstellen. Es ist dem in Kapitel 2 vorgestellten Gefangenendilemma sehr ähnlich. Es stellt zugegebenermaßen stark vereinfacht den Konflikt zweier Länder dar, die vor der Entscheidung stehen, die Umwelt weiterhin zu verschmutzen, oder nicht.

Beispiel 5.1 (Umweltdilemma).

Es seien Land 1 und Land 2 an dem Spiel beteiligt. Ihre Möglichkeiten bestehen nun darin die Luft durch CO₂ zu verschmutzen oder nicht. Der Aufwand für die Reduktion der Schadstoffe beträgt 500 GE, die Umweltentlastung pro Land dagegen 800 GE.

Das führt zu diesem Schema:

	Verschmutzung	Keine Verschmutzung	
Verschmutzung	0	-100 ← 800/2 - 500	
Keine Verschmutzung	-100	300	

Abbildung 5.1: Umweltdilemma

Die sich daraus ergebenden Möglichkeiten sind:

1. Beide wählen Verschmutzen (0/0)
2. Land 1 verschmutzt/Land 2 verschmutzt nicht (400/-100)
3. Land 1 verschmutzt nicht/ Land 2 verschmutzt (-100/400)
4. Beide wählen Nichtverschmutzen (300/300)

Wie bereits in Beispiel 2.4 wird auch dieses Dilemma auf ein Pareto-Optimum hinauslaufen. Dies entspricht hier der ersten Alternative, nämlich dass beide Länder sich für Verschmutzen entscheiden.

5.2 Das ungesteuerte Modell

Um nun das Problem der Emissionsreduktion genauer betrachten zu können, wollen wir das sogenannte TEM-Modell (Technologie-Emissionen-Mittel-Modell) aus [8] und [14] verwenden. Es handelt sich hierbei um ein System aus Differenzgleichungen. Mit dessen Hilfe können die Abhängigkeiten von n Spielern bei der CO₂-Reduktion und den dazu aufzuwendenden finanziellen Mitteln dargestellt werden. Wir führen dazu zunächst das ungesteuerte Modell ein.

$$\begin{aligned} E_i(k+1) &= E_i(k) + \sum_{j=1}^n em_{ij} M_j(k) \\ M_i(k+1) &= M_i(k) - \lambda_i M_i(k) (M_i^* - M_i(k)) E_i(k) \end{aligned} \quad (5.1)$$

für $i = 1, \dots, n$ und $k \in \mathbb{N}_0$

Nun wollen wir die einzelnen Parameter näher betrachten. Bei unserem System bezeichnet:

- $E_i(k)$ die Größe der *Emissionsreduktion* von Spieler i
- $M_i(k)$ die *Mittelaufwendung* von Spieler i
- $M_i^* > 0$ eine obere Schranke für die Mittelaufwendung $M_i(k)$ von Spieler i mit $0 \leq M_i(k) \leq M_i^*$
- $\lambda_i > 0$ den *Reaktionskoeffizienten* von Spieler i , einen Wachstumsparameter
- die em -Matrix den Einfluss der verschiedenen Spieler auf deren Systeme

5.2.1 Das Wachstum der Mittelaufwendung

Für die Mittelaufwendung soll ein Wachstum angenommen werden, das wie in [21] beschrieben, so realistisch wie möglich deren Entwicklung abbildet. Ein solches ist das logistische Wachstum, das der logistischen Gleichung aus [18] genügt. Verwendung findet diese Art des Wachstums etwa bei der Modellierung der Vermehrung einer Population, die auch zunächst exponentiell wächst und dann auf Grund von natürlichen Feinden immer geringer ansteigt.

Formal bedeutet dies:

Definition 5.2 (Logistisches Wachstum).

Ein Wachstum heißt *logistisch*, wenn es folgender Differenzgleichung genügt.

$$x(k+1) = \lambda x(k)(1 - x(k))$$

mit $0 \leq x(0) \leq 1$ und dem Änderungsparameter $\lambda \in]1, 4]$.

Die nächste Abbildung stellt das logistische Wachstum mit $\lambda = 1.5$ dar.

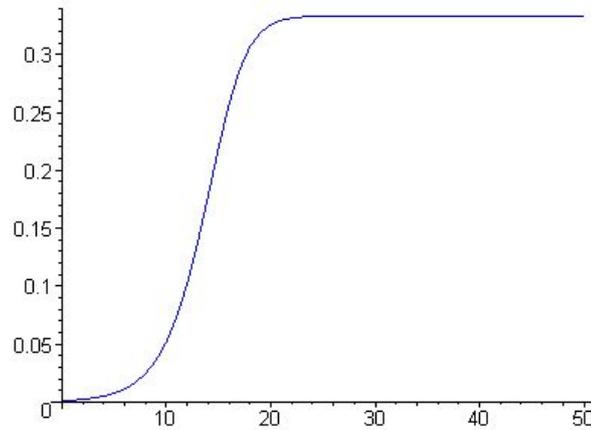


Abbildung 5.2: Darstellung des logistischen Wachstums

Für unsere Gleichung gilt:

$$M_i(k+1) = \begin{cases} \geq M_i(k) & , \text{ falls } E_i(k) \leq 0 \\ \leq M_i(k) & , \text{ falls } E_i(k) \geq 0 \end{cases}$$

Das bedeutet also, dass der Mittelaufwand steigt, falls die Emissionsreduktion negativ ist und fällt, falls diese positiv ist.

Ziel ist es unsere Erkenntnisse auch praktisch anzuwenden und das optimale Steuerungsproblem des CO₂-Reduktionsmodells, das im nächsten Abschnitt hergeleitet werden soll, analytisch zu lösen. Bei Verwendung des oben gegebenen Systems (5.3) ergeben sich Probleme bei der Berechnung der optimalen Wertefunktion mittels der Bellman'schen Rekursionsformel. Durch das hier verwendete Wachstum für die finanziellen Mittel, das sich am logistischen Wachstum orientiert, wächst die Dimension der Exponenten des Gütefunktional und damit auch die der optimalen Wertefunktion exponentiell an. Dies führt bereits bei einer sehr geringen Anzahl an Stufen der Bellman'schen Rekursionsformel zu großen Rechenproblemen. Deshalb möchte ich nun anstatt des logistischen Wachstums, was zweifelsohne das wohl realistischste Verhalten abbilden kann, das *beschränkte Wachstum* aus [19] verwenden um die Entwicklung der finanziellen Aufwendungen darzustellen. Auch hier können wir Vorgänge aus der Natur wie etwa ein Populationswachstum abbilden. Bei dieser Art des Wachstums gibt es eine Obergrenze, auch wird das Wachstum immer geringer, je näher man dieser kommt.

Definition 5.3 (Beschränktes Wachstum).

Ein Wachstum heißt *beschränkt*, wenn es folgender Differenzgleichung genügt.

$$x(k+1) = x(k) + \lambda(x^* - x(k))$$

mit oberer Schranke $x^* \in \mathbb{R}$ und Änderungsparameter $\lambda \in \mathbb{R}$.

Zur Veranschaulichung soll nun folgende Abbildung des beschränkten Wachstums mit $\lambda = \frac{1}{10}$ und $x^* = 1$ dienen.

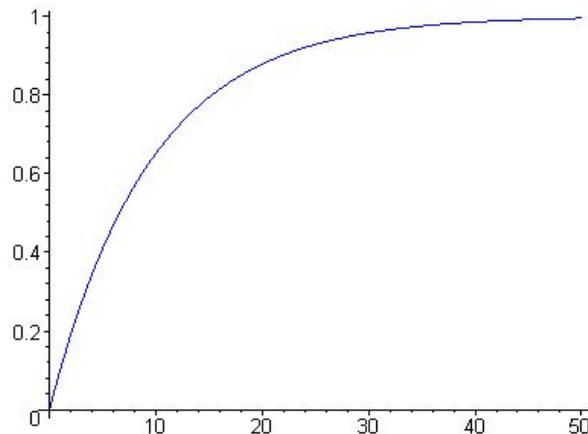


Abbildung 5.3: Darstellung des beschränkten Wachstums

Dies führt zu folgendem vereinfachten CO₂-Emissionsreduktionsmodell, auf das wir uns im weiteren Verlauf immer beziehen werden.

$$\begin{aligned} E_i(k+1) &= E_i(k) + \sum_{j=1}^n em_{ij} M_j(k) \\ M_i(k+1) &= M_i(k) - \lambda_i(M_i^* - M_i(k)) \end{aligned} \quad (5.2)$$

für $i = 1, \dots, n$ und $k \in \mathbb{N}_0$

5.2.2 Die *em*-Matrix

Ein gewichtiger Bestandteil unseres Systems ist die *em*-Matrix. Sie modelliert die Beeinflussung der Emissionsreduktion der einzelnen Spieler untereinander. Es wird stets $em_{ii} > 0$ für alle $i = 1, \dots, n$ gewählt, denn die eigenen finanziellen Aufwendungen wirken sich auf die eigene Emissionsreduktion positiv aus. em_{ij} zeigt die Veränderung bei der Emissionsreduktion von Spieler i , wenn der j -te

Spieler eine Geldeinheit aufwendet. Wird $em_{ij} < 0$ angenommen, so spielen unsere Spieler gegeneinander. In der Praxis könnte dies bedeuten, dass Spieler j ein Kohlekraftwerk im Land von Spieler i baut. Für $em_{ij} > 0$ hingegen würde dies eine gegenseitige Unterstützung, oder auch Kooperation bedeuten. Praktisch könnte dies dem Bau eines Wasserkraftwerks von Spieler j im Land von Spieler i entsprechen.

Diese gegenseitigen Beeinflussungen der einzelnen Spieler werden nun, anhand einer Grafik und eines Beispiels für zwei Spieler wie in [21] veranschaulicht.

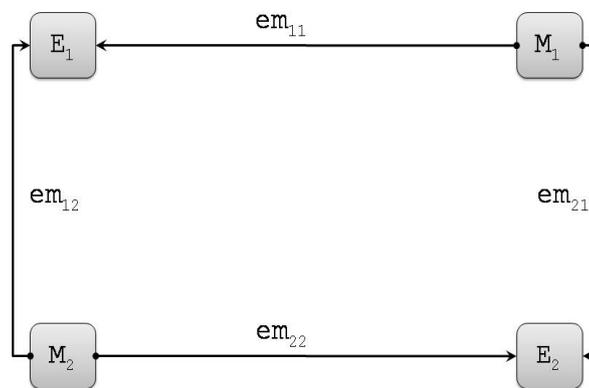


Abbildung 5.4: Darstellung der em -Matrix bei zwei Spielern

Mit dem nun folgenden Beispiel soll gezeigt werden, wie die einzelnen Werte dieser em -Matrix in der Praxis berechnet werden können.

Beispiel 5.4 (Bedeutung der em -Matrix in der Praxis).

Wir betrachten die Situation von zwei beteiligten Spielern, Land 1 und Land 2. Die zugehörigen Gleichungen lauten dann:

$$\begin{aligned} E_1(k+1) &= E_1(k) + em_{11}M_1(k) + em_{12}M_2(k) \\ M_1(k+1) &= M_1(k) - \lambda_1(M_1^* - M_1(k)) \end{aligned}$$

$$\begin{aligned} E_2(k+1) &= E_2(k) + em_{22}M_2(k) + em_{21}M_1(k) \\ M_2(k+1) &= M_2(k) - \lambda_2(M_2^* - M_2(k)) \end{aligned}$$

Weiterhin stellen wir uns folgendes Szenario vor:

Land 1 errichtet in Land 2 einen Windpark, und in Land 1 ein Wasserkraftwerk. Es verwendet also z Anteile im eigenen Land und $(1 - z)$ Anteile in Land 2.

Um eine Gewichtung der Effektivität von Wasserkraftwerken und Windparks abbilden zu können, führen wir die Parameter e_1, e_2 ein. Die Effektivität von Wasserkraftwerken soll e_1 betragen, die von einem Windpark e_2 . Dies bedeutet, dass eine eingesetzte Geldeinheit im Land 1 e_1 Reduktionseinheiten zur Folge hat.

Des weiteren müssen wir berücksichtigen in wieweit sich Emissionsreduktionen für das eigene Land anrechnen lassen, obwohl sie in einem anderen Land erbracht wurden. Dies wird auf Verhandlungen der einzelnen Länder hinauslaufen müssen. Hier einigen sich Land 1 und Land 2 auf folgende Anrechnung der Emissionsreduktion.

Land 1 darf sich auf eigenem Gebiet für die Errichtung des Wasserkraftwerks a Anteile und im Land 2 für die Errichtung des Windparks b Anteile anrechnen lassen. Land 2 ist es erlaubt $(1 - b)$ Anteile für sich zu verbuchen.

Die zugehörige em -Matrix der beiden Länder lautet also:

- $em_{11} = ze_1a + (1 - z)e_2b$
- $em_{12} = 0$
- $em_{21} = (1 - z)e_2(1 - b)$
- $em_{22} = 0$

5.2.3 Untersuchung auf einen attraktiven Fixpunkt

Es drängt sich nun die Frage auf, ob unser Modell der CO_2 -Reduktion ohne jeglichen Eingriff von außen einen gewissen Wert an Emissionsreduktion und finanzieller Mittelaufwendung erreichen kann, oder nicht. Mathematisch entspricht dies der Untersuchung, ob unser Differenzensystem einen attraktiven (anziehenden) Fixpunkt besitzt. Wir werden an dieser Stelle nun das zu (5.2) gehörige zeitdiskrete System mit Fixpunkt wie in [8] herleiten.

Wir definieren dazu für jedes $i = 1, \dots, n$

- Anfangsbedingungen $E_i(0) = E_{i0}$ und $M_i(0) = M_{i0}$

- eine Zustandsvektorfunktion $x_i : \mathbb{N}_0 \rightarrow \mathbb{R}^2$ mit $x_i(k) = (E_i(k), M_i(k))^T$, $k \in \mathbb{N}_0$
- eine Vektorfunktion $f_i : \mathbb{R}^{2n} \rightarrow \mathbb{R}^2$ mit

$$f_i(x) = \begin{pmatrix} E_i + \sum_{j=1}^n em_{ij} M_j \\ M_i - \lambda_i (M_i^* - M_i) \end{pmatrix}, x = (x_1^T, \dots, x_n^T)^T$$

Man kann (5.2) umschreiben in ein Differenzensystem der Form

$$x_i(k+1) = f_i(x(k))$$

für $i = 1, \dots, n$

In Vektorform, folgt für jedes $\hat{x} = (\hat{E}^T, \hat{M}^T)^T$ mit $\hat{E} \in \mathbb{R}^n$

$$\hat{x} = f(\hat{x})$$

\hat{x} ist also ein Fixpunkt unseres zeitdiskreten dynamischen Systems.

Für die Untersuchung auf einen attraktiven (anziehenden) Fixpunkt müssen wir nun zunächst die Jacobi-Matrix unserer Differenzgleichungen (5.2) berechnen und dann die zugehörigen Eigenwerte betrachten. Sind diese im Betrag < 1 , so ist der Fixpunkt attraktiv (anziehend). Für die genaue theoretische Grundlage sei zum Beispiel auf [9] verwiesen.

$$\begin{pmatrix} 1 & 0 & \dots & 0 & em_{11} & em_{12} & \dots & em_{1n} \\ 0 & 1 & \dots & \vdots & em_{21} & em_{22} & \dots & em_{2n} \\ \vdots & & \ddots & & & & & \\ 0 & 0 & \dots & 1 & em_{n1} & em_{n2} & \dots & em_{nn} \\ 0 & 0 & \dots & 0 & 1 + \lambda_1 & & & \\ \vdots & \vdots & & \vdots & & 1 + \lambda_2 & & \\ & & & & & & \ddots & \\ 0 & \dots & & 0 & 0 & & & 1 + \lambda_n \end{pmatrix}$$

Daraus sind die zugehörigen Eigenwerte abzulesen:

$$\chi_1 = \dots = \chi_n = 1$$

und

$$\chi_{n+i} = 1 + \lambda_i$$

für $i = 1, \dots, n$

Somit ist ersichtlich, dass die Eigenwerte im Betrag nicht wie gefordert alle < 1 sind. Unsere Bedingung für attraktive Fixpunkte ist also nicht erfüllt.

Das bedeutet für unser System (5.2), dass es einen gegebenen Anfangszustand $(E_0^T, M_0^T)^T$ nicht (auch nicht in unendlich vielen Zeitschritten) in einen Fixpunkt (\hat{E}^T, \hat{M}^T) überführt.

Wie W. Pickl in [14] bereits bemerkte, ist auch in der Realität „mit einer Selbstorganisation hinsichtlich der Klimaschutzrahmenkonvention im Allgemeinen nicht zu rechnen“. Um einen gewissen Emissionsreduktionswert mit bestimmten finanziellen Mitteln, also einen Fixpunkt unseres Systems zu erreichen, ist eine Koordinierung der verschiedenen Spielerinteressen nötig. In unser bisheriges ungesteuertes System (5.2) muss also von außen steuernd eingegriffen werden, dies geschieht im nächsten Abschnitt mit der Definition eines gesteuerten Systems zur CO₂-Reduktion.

5.3 Das gesteuerte Modell

An dieser Stelle wollen wir nun unser ungesteuertes System (5.2) in ein gesteuertes System überführen. Dazu verwenden wir eine Steuerung u_i für $i = 1, \dots, n$, die jeder Spieler zum Erreichen seines Ziels zur Verfügung hat.

Unser gesteuertes System ist dann gegeben durch:

$$\begin{aligned} E_i(k+1) &= E_i(k) + \sum_{j=1}^n em_{ij}(M_j(k) + u_j(k)) \\ M_i(k+1) &= M_i(k) + u_i(k) - \lambda_i(M_i^* - M_i(k)) \end{aligned} \quad (5.3)$$

für $i = 1, \dots, n$ und $k \in \mathbb{N}_0$

Der Teil, der die Emissionsreduktion beschreibt, wird also von den Steuerungen aller Spieler $j = 1, \dots, n$ beeinflusst. Die Mittelaufwendung ändert sich nur durch die eigene Steuerung. Dies erweitert die in Abbildung 5.4 dargestellte gegenseitige Beeinflussung der einzelnen Spieler enorm, so dass sich folgende Darstellung für zwei Spieler ergibt.

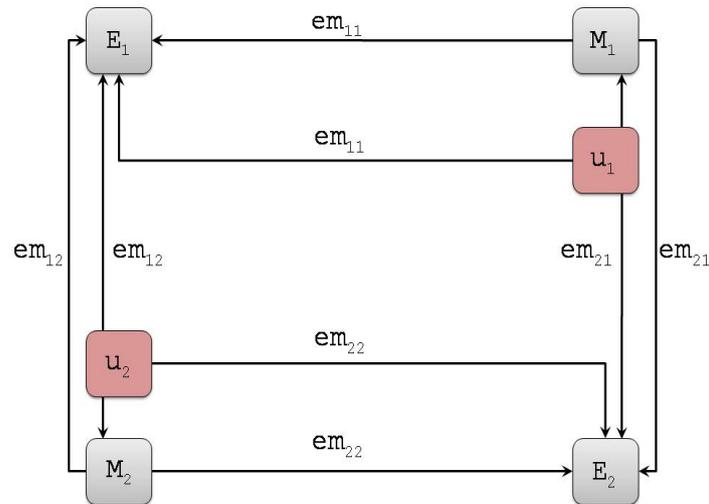


Abbildung 5.5: Darstellung der em -Matrix bei zwei Spielern mit Steuerungen

Auch mit diesen eben eingeführten Steuerungen, die die Spieler zur Verfügung haben, darf die Grenze der finanziellen Mittelaufwendung nicht überschritten werden. Die Steuerungsfunktionen $u_i : \mathbb{N}_0 \rightarrow \mathbb{R}$, $i = 1, \dots, n$ müssen also folgende Bedingungen erfüllen:

$$0 \leq M_i(k) + u_i(t) \leq M_i^* \quad (5.4)$$

für alle $i = 1, \dots, n$ und $k \in \mathbb{N}_0$

5.4 Das dynamische Spiel zur CO₂-Reduktion

Mit unserem gesteuerten Modell (5.3) haben wir den Grundstein für die Behandlung eines dynamischen Spiels aus Definition 4.10 gelegt. In diesem Abschnitt wollen wir nun das zugehörige Gütefunktional auf endlichem Zeithorizont definieren und das dynamische Spiel herleiten. Im praktischen Teil werden wir sowohl das Spiel mit einem Spieler, also auch das mit zwei Spielern bei nicht-antizipierender Spielsituation betrachten. Deshalb definieren wir nun zuerst das dynamische Spiel zur CO₂-Reduktion für einen Spieler, anschließend das für zwei gegnerische Spieler.

Definition 5.5 (Gütefunktional des CO₂-Reduktionsmodell mit einem Spieler).

Betrachte unser Kontrollsystem (5.3) für einen Spieler

$$\begin{aligned} E(k+1) &= E(k) + em(M(k) + u(k)) \\ M(k+1) &= M(k) + u(k) - \lambda(M^* - M(k)) \end{aligned} \quad (5.5)$$

Wir definieren das *Gütefunktional des CO₂-Reduktionsmodells mit einem Spieler* durch

$$\begin{aligned}
 J(x, u) = & \left\{ \left[E(k) + em(M(k) + u(k)) - \hat{E} \right]^2 + \right. \\
 & \left. + \left[M(k) + u(k) - \lambda(M^* - M(k)) - \hat{M} \right]^2 + u^2(k) \right\} + \\
 & + \vartheta(E(K), M(K))
 \end{aligned} \tag{5.6}$$

mit

$$\vartheta(E(K), M(K)) := (E(K) - \hat{E})^2 + (M(K) - \hat{M})^2$$

Dabei beschreibt $\left[E(k) + em(M(k) + u(k)) - \hat{E} \right]^2$ den Abstand der Emissionsreduktion zu dem gewünschten Zielwert und $\left[M(k) + u(k) - \lambda(M^* - M(k)) - \hat{M} \right]^2$ den Abstand zu dem gewünschten Ziel an finanzieller Mittelaufwendung. Die nächste Definition wird zeigen, dass es unser Ziel ist die optimale Wertefunktion zu berechnen. Das heißt wir müssen ein Minimierungsproblem lösen. Dabei steht uns die Steuerung $u(k)$ zur Verfügung um eben diese Abstände zu den jeweiligen Zielwerten zu minimieren.

Definition 5.6 (Das dynamische Spiel zur CO₂-Reduktion mit einem Spieler). *Das dynamische Spiel der CO₂-Reduktion mit einem Spieler* besteht nun darin, die zugehörige optimale Wertefunktion aus Definition 4.9 zu bestimmen und die optimalen Strategien des Spielers zu berechnen.

Für die Lösung dieses dynamischen Spiels wird, wie in Abschnitt 4.2 ausführlich erläutert, das Bellman'sche Optimalitätsprinzip verwendet. Für die Durchführung der dynamischen Programmierung wurde ein maple-Programm geschrieben. Die Ergebnisse werden in Kapitel 6 ausführlich dargestellt.

Auch für den Fall von zwei gegnerischen Spielern soll nun das dynamische Spiel der CO₂-Reduktion vorgestellt werden. Zuerst leiten wir aus unserem Kontrollsystem (5.3) das Modell für zwei Spieler mit nicht-antizipierender Spielsituation her und definieren dann damit die Gütefunktionale der beiden Spieler.

Definition 5.7 (CO₂-Reduktionsmodell für zwei Spieler mit nicht-antizipierender Spielsituation).

Betrachte unser Kontrollsystem (5.3).

Das *CO₂-Reduktionsmodell für zwei Spieler mit nicht-antizipierender Spielsituation* ist dann gegeben durch

$$\begin{aligned}
 E_1(k+1) &= E_1(k) + em_{11}(M_1(k) + \mu_1[u_2](k)) + em_{12}(M_2(k) + u_2(k)) \\
 M_1(k+1) &= M_1(k) + \mu_1[u_2](k) - \lambda(M_1^* - M_1(k))
 \end{aligned}$$

$$\begin{aligned} E_2(k+1) &= E_2(k) + em_{21}(M_1(k) + \mu_1[u_2](k)) + em_{22}(M_2(k) + u_2(k)) \\ M_2(k+1) &= M_2(k) + u_2 - \lambda(M_2^* - M_2(k)) \end{aligned}$$

Definition 5.8 (Gütefunktional des CO₂-Reduktionsmodells bei nicht-antizipierender Spielsituation).

Betrachte das Kontrollsystem aus Definition 5.7.

Das Gütefunktional des CO₂-Reduktionsmodells bei nicht-antizipierender Spielsituation für Spieler 1 ist dann gegeben durch

$$\begin{aligned} J_1(x_1, \mu_1[u_2], u_2) &= \\ &\left\{ \left[E_1(k) + em_{11}(M_1(k) + \mu_1[u_2](k)) + em_{12}(M_2(k) + u_2(k)) - \hat{E}_1 \right]^2 + \right. \\ &\left. + \left[M_1(k) + \mu_1[u_2](k) - \lambda_1(M_1^* - M_1(k)) - \hat{M}_1 \right]^2 + \mu_1[u_2]^2(k) \right\} + \\ &+ \vartheta_1(E_1(K), M_1(K)) \end{aligned} \quad (5.7)$$

mit

$$\vartheta_1(E_1(K), M_1(K)) := (E_1(K) - \hat{E}_1)^2 + (M_1(K) - \hat{M}_1)^2$$

und für Spieler 2 durch

$$\begin{aligned} J_2(x_2, \mu_1[u_2], u_2) &= \\ &\left\{ \left[E_2(k) + em_{21}(M_1(k) + \mu_1[u_2](k)) + em_{22}(M_2(k) + u_2(k)) - \hat{E}_2 \right]^2 + \right. \\ &\left. + \left[M_2(k) + u_2(k) - \lambda_2(M_2^* - M_2(k)) - \hat{M}_2 \right]^2 + u_2^2(k) \right\} + \\ &+ \vartheta_2(E_2(K), M_2(K)) \end{aligned} \quad (5.8)$$

mit

$$\vartheta_2(E_2(K), M_2(K)) := (E_2(K) - \hat{E}_2)^2 + (M_2(K) - \hat{M}_2)^2$$

Wie im Fall von nur einem Spieler wollen unsere beiden Spieler hier auch wieder die jeweiligen Abstände

$$\left[E_i(k) + em_{i1}(M_i(k) + \mu_1[u_2](k)) + em_{i2}(M_i(k) + u_2(k)) - \hat{E}_i \right]^2$$

für $i = 1, 2$ und

$$\left[M_1(k) + \mu_1[u_2](k) - \lambda_1(M_1^* - M_1(k)) - \hat{M}_1 \right]^2$$

beziehungsweise

$$\left[M_2(k) + u_2(k) - \lambda_2(M_2^* - M_2(k)) - \hat{M}_2 \right]^2$$

zu den Zielwerten der Emissionsreduktion und der finanziellen Mittelaufwendung minimieren. Dazu steht Spieler 1 die nicht-antizipierende Strategie $\mu_1(k)$ zur Verfügung, und Spieler 2 die Kontrolle $u_2(k)$. Mit diesen voneinander abhängigen Parametern ist eine gegenseitige Einflussnahme der beiden Spieler möglich. Aus Definition 4.19 ergeben sich folgende optimalen Wertefunktionen der beiden Spieler.

Definition 5.9 (Optimale Wertefunktion des CO₂-Reduktionsmodells bei nicht-antizipierender Spielsituation).

Gegeben seien die Gütefunktionale der beiden Spieler (5.7) und (5.8)

Bei nicht-antizipierender Spielsituation definieren wir

- (i) die *optimale Wertefunktion des CO₂-Reduktionsmodells bei nicht-antizipierender Spielsituation von Spieler 1* durch

$$\begin{aligned} V_1(x_1) &:= J_1(x_1, \mu_1^*[u_2^*], u_2^*) \\ &= \left\{ \left[E_1(k) + em_{11}(M_1(k) + \mu_1^*[u_2^*](k)) + em_{12}(M_2(k) + u_2^*(k) - \hat{E}_1) \right]^2 + \right. \\ &\quad \left. + \left[M_1(k) + \mu_1^*[u_2^*](k) - \lambda_1(M_1^* - M_1(k)) - \hat{M}_1 \right]^2 + (\mu_1^*[u_2^*](k))^2 \right\} + \\ &\quad + \vartheta_1(E_1(K), M_1(K)) \end{aligned}$$

mit optimaler Strategie μ_1^* aus Definition 4.15 und optimaler Kontrolle u_2^* aus Definition 4.16.

- (ii) die *optimale Wertefunktion des CO₂-Reduktionsmodells bei nicht-antizipierender Spielsituation von Spieler 2* durch

$$\begin{aligned} V_2(x_2) &:= J_2(x_2, \mu_1^*[u_2^*], u_2^*) \\ &\left\{ \left[E_2(k) + em_{21}(M_1(k) + \mu_1^*[u_2^*](k)) + em_{22}(M_2(k) + u_2^*(k) - \hat{E}_2) \right]^2 + \right. \\ &\quad \left. + \left[M_2(k) + u_2^*(k) - \lambda_2(M_2^* - M_2(k)) - \hat{M}_2 \right]^2 + (u_2^*(k))^2 \right\} + \\ &\quad + \vartheta_2(E_2(K), M_2(K)) \end{aligned}$$

mit optimaler Strategie μ_1^* aus Definition 4.15 und optimaler Kontrolle u_2^* aus Definition 4.16

Definition 5.10 (Dynamisches Spiel zur CO₂-Reduktion bei nicht-antizipierender Spielsituation).

Das dynamische Spiel der CO₂-Reduktion bei nicht-antizipierender Spielsituation besteht nun darin, die optimalen Wertefunktionen der beiden Spieler aus Definition 5.9 zu berechnen und die zugehörigen optimalen Strategien und optimalen Kontrollen zu berechnen.

Die Lösung dieses dynamischen Spiels zur CO₂-Reduktion erfolgt wieder mit Hilfe der dynamischen Programmierung. Dabei verwenden wir die Bellman'schen Rekursionsformeln für die beiden Spieler bei nicht-antizipierender Spielsituation (4.6) und (4.7). Wegen der besonderen Strategie von Spieler 1 sind zur Optimierung wieder die drei Teilschritte nötig, die in Unterabschnitt 4.3.2 ausführlich erläutert wurden. Um das Verhalten der beiden Spieler zu zeigen wurde für die praktische Behandlung dieses dynamischen Spiels ein maple-Programm geschrieben. Die daraus resultierenden Ergebnisse werden in Kapitel 6 gezeigt.

Dynamische Programmierung

In diesem Kapitel soll nun die Lösung unseres dynamischen Spiels zur CO₂-Reduktion mittels dynamischer Programmierung durchgeführt werden. Dazu verwenden wir maple-Programme, die die optimalen Wertefunktionen nach dem in Kapitel 4 beschriebenen Verfahren berechnen. Zunächst folgt die praktische Behandlung des Falls von nur einem Spieler wie in [13]. Im Anschluss daran betrachten wir dann den komplexeren Fall von zwei Spielern in einer nicht-antizipierenden Spielsituation.

6.1 Spiel mit einem Spieler

6.1.1 Das maple-Programm

Die Behandlung des dynamischen Spiels mit einem Spieler erfolgt mit dem maple-Programm `dynprog1spieler.mws`. Der Programmablauf soll hier stichpunktartig erläutert werden, für den genauen Quellcode sei auf die beiliegende CD verwiesen.

1. Definition der Variablen:

- K : Endzeitpunkt
- $EHat$: Zielwert der Emissionsreduktion
- $MHat$: Zielwert der Mittelaufwendung
- em : Wert der eindimensionalen em-Matrix
- $lambda$: Wachstumsparameter
- $MStern$: obere Grenze der finanziellen Mittel

2. Randbedingung der dynamischen Programmierung festlegen

$$V[K] := (E - \hat{E})^2 + (M - \hat{M})^2$$
3. Rekursive Berechnung der optimalen Wertefunktion für $k := K - 1, \dots, 0$
 - Definition folgender Überführungskosten
$$J[k](x, u) := (E + em(M + u) - \hat{E})^2 +$$

$$+ (M + u - \lambda(M^* - M) - \hat{M})^2 + u^2 +$$

$$+ V[k + 1]((E + em(M + u)), (M + u - \lambda(M^* - M)))$$
 - Berechnen der optimalen Strategie
 - Berechnen der optimalen Wertefunktion
4. Plotten der optimalen Wertefunktion
5. Festlegen der Startwerte für Emissionsreduktion und finanzielle Mittelaufwendung
6. Ausgabe des Werts der optimalen Wertefunktion
7. Vorwärtseinsetzen um Lösungstrajektorien für $k := 0, \dots, K - 1$ zu berechnen
 - Berechnung der Werte der Emissionsreduktion
 - Berechnung der Werte der finanziellen Mittelaufwendung unter Berücksichtigung der oberen und unteren Schranke
 - Ausgabe der Ergebnisse
8. Ausgabe der Lösungstrajektorien für Emissionsreduktion und finanzielle Mittelaufwendung
9. Arrays für die Plots definieren
10. Plotten der Kurve für die Emissionsreduktion
11. Plotten der Kurve der finanziellen Mittelaufwendung

6.1.2 Ergebnisse des dynamischen Spiels zur CO₂-Reduktion

Wir wollen nun die Ergebnisse unseres Spiels mit nur einem Spieler bei $K = 10$ Schritten betrachten. Dazu seien folgende Parameter vorgegeben:

$EHat := 1.0$	$MHat := 0$
$em := 1.0$	$lambda := 0.01$
$MStern := 1.0$	
$E := -1.0$	$M := 0.5$

Tabelle 6.1: Startparameter bei einem Spieler

Der Spieler möchte sein optimales Steuerungsproblem für die Startwerte der Emissionsreduktion $E = -1.0$ und des finanziellen Mittelaufwands $M = 0.5$ zu den Zielwerten $EHat = 1.0$, $MHat = 0$ hin lösen. Für den Aufwand der finanziellen Mittel wählt er das arithmetische Mittel aus der unteren und der oberen Grenze der finanziellen Mittel. Dies ist plausibel, da er im ersten Jahr nicht weiß, was er einsetzen muss, um sein Ziel zu erreichen. Der Wert $em = 1.0$ bedeutet lediglich, dass sich sein Mitteleinsatz in vollem Umfang positiv auf seine Emissionsreduktion auswirkt. Für den Wachstumsparameter soll $lambda = 0.01$ gelten. Wir setzen einen festen Endzeitpunkt, bis zu dem unser Ziel erreicht sein soll. Realistisch ist wohl ein Wert von $K = 10$ Jahren.

Die Entwicklung der Emissionsreduktion wird hier zunächst grafisch gezeigt, anschließend werden die Werte der Lösungstrajektorie angegeben.

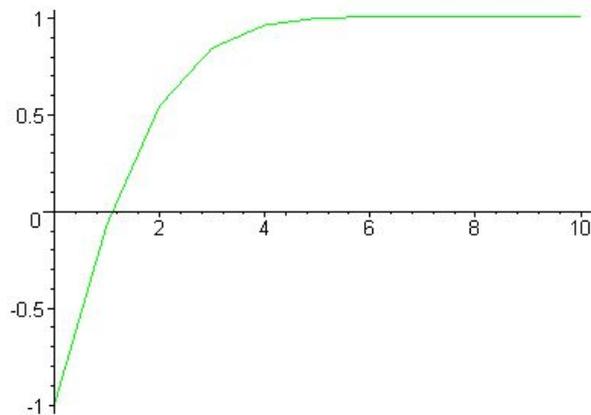


Abbildung 6.1: Emissionsreduktion bei einem Spieler

$$E = [-1, -0.0673202653, 0.5462470242, 0.8461221399, \\ 0.9643222110, 1.000493801, 1.006868456, 1.005977990, \\ 1.005808398, 1.006337709, 1.007802625]$$

Auch für die Mittelaufwendung geben wir die grafische und die numerische Lösung an.

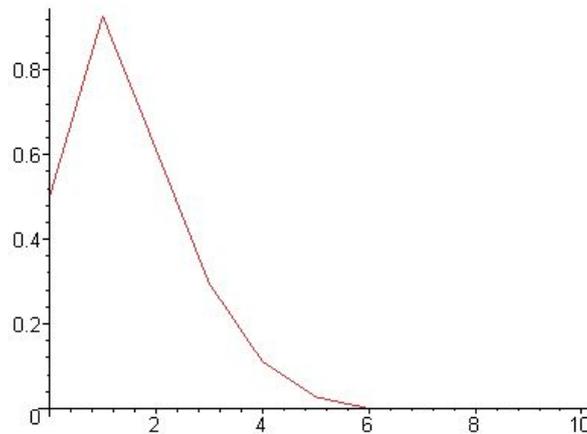


Abbildung 6.2: Mittelaufwendung bei einem Spieler

$$M = [0.5, 0.9276797347, 0.6128440868, 0.2960035566, 0.1111601072, \\ 0.02728319147, 0, 0, 0, 0, 0]$$

Die optimale Wertefunktion des Spielers bei dem Spiel der CO₂-Reduktion wird nun dargestellt.

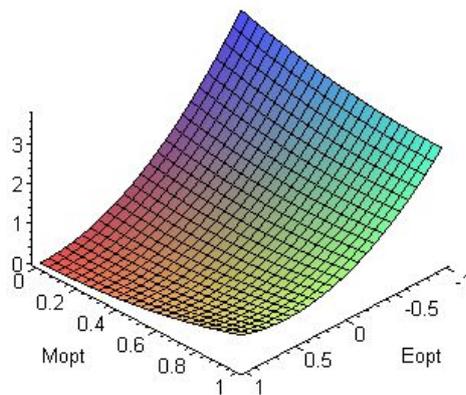


Abbildung 6.3: Optimale Wertefunktion bei einem Spieler

Für unsere gewählten Startwerte $E := -1$ und $M := 0.5$ ergibt sich dann folgender Wert der optimalen Wertefunktion:

$$V((E, M)) = 3.129464031$$

Es wäre hier eine Szenarioänderung denkbar. Würden wir hier für die Mittelaufwendung negative Werte zulassen, würde dies bedeuten, dass der Spieler bei Erreichen seines Emissionsreduktionsziels negative Kosten, also eine Art Bonus hätte auf den er bei seiner Optimierung zurückgreifen könnte. Hier ergäbe sich bei Verwendung der oben angegebenen Startparameter folgender Verlauf bei der Emissionsreduktion und der Mittelaufwendung.

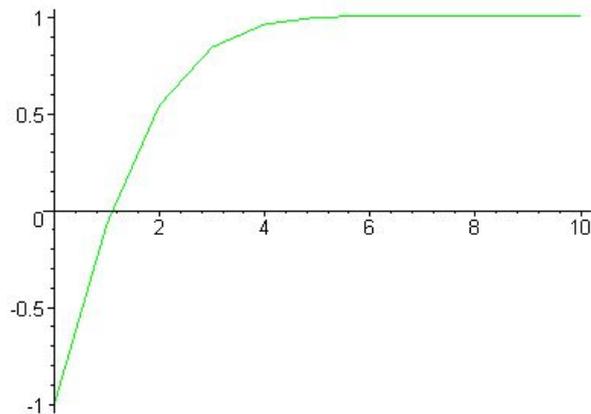


Abbildung 6.4: Emissionsreduktion ohne untere Schranke der Mittelaufwendung

$$E = [-1, -0.0673202653, 0.5462470242, 0.8461221399, \\ 0.9643222110, 1.000493801, 1.006868456, 1.005398653, \\ 1.003526514, 1.003052946, 1.003755453]$$

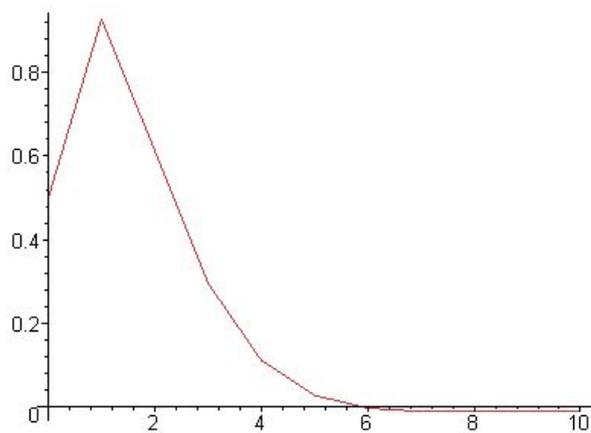


Abbildung 6.5: Mittelaufwendung ohne untere Schranke

$$M = [0.5, 0.9276797347, 0.6128440868, 0.2960035566, 0.1111601072, \\ 0.02728319147, -0.003352512715, -0.01150332744, -0.01198717281, \\ -0.01059344004, -0.00940342704]$$

6.2 Spiel mit zwei Spielern

6.2.1 Das maple-Programm

Für die praktische Behandlung des Spiels mit zwei Spielern bei nicht-antizipierender Spielsituation dient das maple-Programm `dynprog2spieler.mws`. Wie im Fall von einem Spieler soll auch hier der Ablauf zunächst stichpunktartig beschrieben werden. Für den Quellcode sei wieder auf die beiliegende CD verwiesen.

1. Definition der Variablen:

- K : Endzeitpunkt der dynamischen Programmierung
- $EHat1$: Zielwert der Emissionsreduktion von Spieler 1
- $MHat1$: Zielwert der finanziellen Mittelaufwendung von Spieler 1
- $EHat2$: Zielwert der Emissionsreduktion von Spieler 2
- $MHat2$: Zielwert der finanziellen Mittelaufwendung von Spieler 2
- $em11, em12, em21, em22$: Werte der em-Matrix
- $lambda1$: Wachstumsparameter von Spieler 1
- $lambda2$: Wachstumsparameter von Spieler 2
- $MStern1$: obere Grenze der finanziellen Mittel von Spieler 1
- $MStern2$: obere Grenze der finanziellen Mittel von Spieler 2

2. Randbedingungen für die dynamische Programmierung festlegen

$$V_1[K] := (E1 - EHat1)^2 + (M1 - MHat1)^2$$

$$V_2[K] := (E2 - EHat2)^2 + (M2 - MHat2)^2$$

3. Rekursive Berechnung der optimalen Wertefunktionen für $k := K - 1, \dots, 0$

- Definition der Überführungskosten von Spieler 1

$$J_1[k] := (E_1 + em_{11}(u_1 + M_1) + em_{12}(u_2 + M_2) - \hat{E}_1)^2 + \\ + (M_1 + u_1 - \lambda_1(M_1^* - M_1) - \hat{M}_1)^2 + (u_1)^2 + \\ + V_1[k + 1]((E_1 + em_{11}(u_1 + M_1) + em_{12}(u_2 + M_2)), \\ (M_1 + u_1 - \lambda_1(M_1^* - M_1)))$$

- Berechnung der optimalen Strategie von Spieler 1 (noch abhängig von u_2)
- Berechnung der optimalen Wertefunktion von Spieler 1 (noch abhängig von u_2)
- Definition der Überführungskosten von Spieler 2 unter Berücksichtigung der optimalen Strategie von Spieler 1

$$\begin{aligned}
 J_2[k] := & (E_2 + em_{21}(z_1 + M_1) + em_{22}(z_2 + M_2) - \hat{E}_2)^2 + \\
 & + (M_2 + z_2 - \lambda_2(M_2^* - M_2)) - \hat{M}_2)^2 + (z_2)^2 + \\
 & + V_2[k + 1]((E_2 + em_{21}(z_2 + M_2) + em_{22}(z_2 + M_2)), \\
 & (M_2 + z_2 - \lambda_2(M_2^* - M_2)))
 \end{aligned}$$

- Berechnung der optimalen Kontrolle von Spieler 2
 - Berechnung der optimalen Wertefunktion von Spieler 2
 - Berechnung der optimalen Strategie von Spieler 1 (unabhängig von u_2)
 - Berechnung der optimalen Wertefunktion von Spieler 1 (unabhängig von u_2)
4. Plotten der optimalen Wertefunktion von Spieler 1
 5. Plotten der optimalen Wertefunktion von Spieler 2
 6. Startwerte für Emissionsreduktion und finanzielle Mittelaufwendung von Spieler 1 festlegen
 7. Startwerte für Emissionsreduktion und finanzielle Mittelaufwendung von Spieler 2 festlegen
 8. Vowärtseinsetzen um Lösungstrajektorien zu berechnen für $k := 0, \dots, K - 1$
 - Berechnung der Emissionsreduktionswerte von Spieler 1
 - Berechnung der Emissionsreduktionswerte von Spieler 2
 - Berechnung der finanziellen Mittelaufwendung von Spieler 1 unter Berücksichtigung der oberen und unteren Schranke
 - Berechnung der finanziellen Mittelaufwendung von Spieler 2 unter Berücksichtigung der oberen und unteren Schranke
 - Ausgabe der Ergebnisse
 9. Ausgabe der Lösungstrajektorien

10. Arrays für die Plots definieren
11. Plotten der Kurve der Emissionsreduktion von Spieler 1
12. Plotten der Kurve der Emissionsreduktion von Spieler 2
13. Plotten der Kurve der finanziellen Mittelaufwendung von Spieler 1
14. Plotten der Kurve der finanziellen Mittelaufwendung von Spieler 2

6.2.2 Ergebnisse des dynamischen Spiels zur CO₂-Reduktion

Für den Fall mit zwei Spielern bei nicht-antizipierender Spielsituation seien folgende Parameter für die Berechnung von $K = 10$ Schritten gegeben:

$E1 := -1.0$	$EHat1 := 1.0$	
$E2 := -1.0$	$EHat2 := 1.0$	
$M1 := 0.5$	$MHat1 := 0$	$MStern1 := 1.0$
$M2 := 0.5$	$MHat2 := 0$	$MStern2 := 1.0$
$lambda1 := 0.01$	$lambda2 := 0.01$	

Tabelle 6.2: Startparameter bei zwei Spielern bei nicht-antizipierender Spielsituation

$em11 := 1.0$	$em12 := -0.5$
$em21 := -0.5$	$em22 := 1.0$

Tabelle 6.3: Werte der em -Matrix

Ziel ist also für beide Spieler ihr optimales Steuerungsproblem zu lösen. Sie wollen jeweils ihre Emissionsreduktion mit den Startwerten $E1 = -1.0$ und $E2 = -1.0$, und ihre finanziellen Mittelaufwendung mit den Startwerten $M1 = 0.5$ und $M2 = 0.5$ hin zu den Zielwerten $EHat1 = 1.0$, $EHat2 = 1.0$ und $MHat1 = 0$, $MHat2 = 0$ optimal zu steuern. Auch hier entspricht der Wert des ersten Einsatzes der finanziellen Mittel wieder dem arithmetischen Mittel aus oberer und unterer Grenze der möglichen Aufwendung. Für die im vorliegenden Fall gegebene em -Matrix gilt, dass die jeweils eigenen finanziellen Mittel sich in vollem Umfang positiv auf die eigene Reduktion auswirken. Die finanziellen Mittel des jeweiligen Gegners beeinflussen die eigene Emissionsreduktion dagegen zur Hälfte negativ. Das bedeutet, dass die Spieler in diesem dynamischen Spiel gegeneinander spielen. Für die Wachstumsparameter der beiden Spieler soll $lambda1 = 0.01$ und $lambda2 = 0.01$ gelten. Auch hier ist der feste Endzeitpunkt nach $K = 10$ Jahren wohl ein realistischer Wert.

Die Ergebnisse der Berechnungen sollen nun für beide Spieler dargestellt werden. Für die beiden Spieler ergeben sich folgende Kurvenverläufe und Werte der CO_2 -Reduktion.

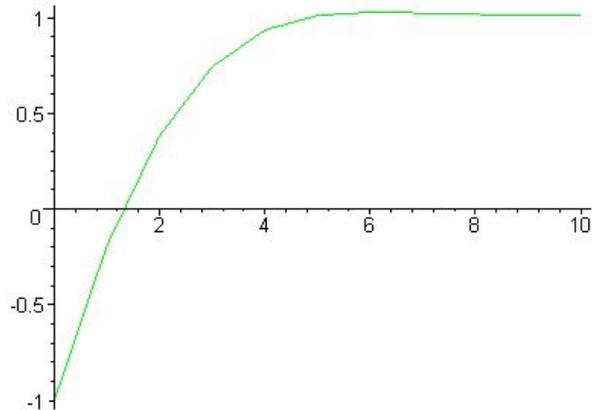


Abbildung 6.6: Emissionsreduktion von Spieler 1

$$E1 = [-1, -0.1905765625, 0.3832361191, 0.743927846, \\ 0.9352545786, 1.011442512, 1.028807765, 1.023993583, \\ 1.017247070, 1.013709888, 1.012656662]$$

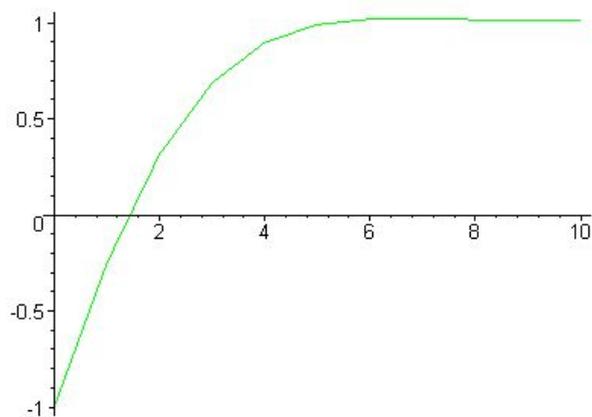


Abbildung 6.7: Emissionsreduktion von Spieler 2

$$E2 = [-1, -0.249883929, 0.3151493630, 0.6882199788, \\ 0.8983172417, 0.9910337160, 1.019544249, 1.020937278, \\ 1.016442969, 1.014054262, 1.013504052]$$

Als nächstes wollen wir die Ergebnisse der finanziellen Mittelaufwendung der beiden Spieler darstellen.

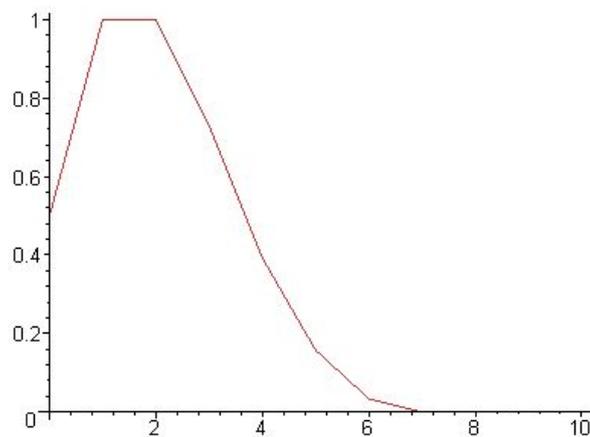


Abbildung 6.8: Mittelaufwendung von Spieler 1

$$M1 = [0.5, 1, 1, 0.7296360460, 0.3924635127, 0.1573195293, 0.03373388699, 0, 0, 0, 0]$$

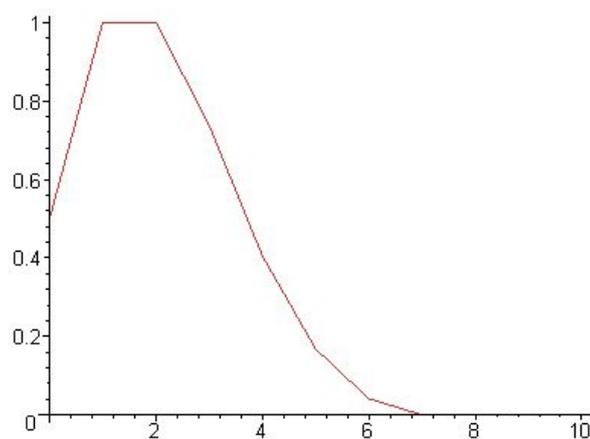


Abbildung 6.9: Mittelaufwendung von Spieler 2

$$M2 = [0.5, 1, 1, 0.7378886388, 0.4050597254, 0.1684645187, 0.04127552359, 0, 0, 0, 0]$$

Die folgenden Plots stellen die optimalen Wertefunktionen der beiden Spieler dar. Für die optimale Wertefunktion von Spieler 1 geben wir dabei $M2 := 0.5$, für die von Spieler 2 $M1 := 0.5$ vor.

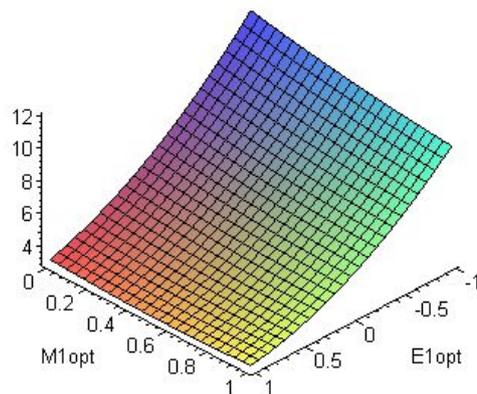


Abbildung 6.10: Optimale Wertefunktion von Spieler 1

Für unsere Startwerte ergibt sich für Spieler 1 folgender Wert der optimalen Wertefunktion:

$$V_1[0]((E1, M1)) = 10.88279996$$

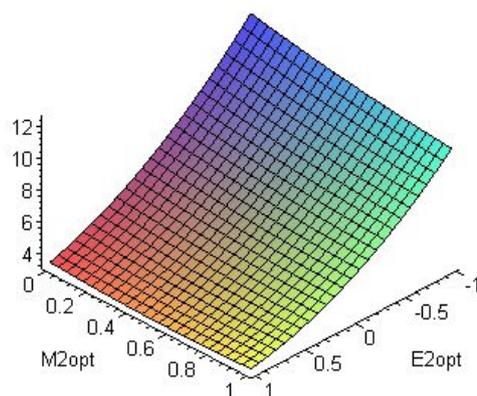


Abbildung 6.11: Optimale Wertefunktion von Spieler 2

Mit unseren Startwerten hat die optimale Wertefunktion von Spieler 2 diesen Wert:

$$V_2[0]((E2, M2)) = 11.34765285$$

An den Werten der optimalen Wertefunktionen wird deutlich, dass die Kosten von Spieler 1 geringer als die von Spieler 2 sind. Dieser Vorteil entspricht dem, den Spieler 1 durch seine nicht-antizipierende Strategie gegenüber Spieler 2 gewinnt. Vergleicht man jedoch die Werte der optimalen Trajektorien der beiden Spieler, wird deutlich, dass Spieler 1 nur einen sehr kleinen Vorteil gegenüber Spieler 2 hat und beide trotzdem gleichzeitig ihr gewünschtes Ziel erreichen.

Ein Greedy-Algorithmus

Als Alternative soll nun ein weiterer Ansatz zur Behandlung des CO₂-Reduktionsmodells gegeben werden. Die Berechnung ist an einen iterativen Algorithmus aus [14] angelehnt. Es handelt sich dabei um einen Algorithmus aus der Klasse der *Greedy*-, oder auch *gierigen Algorithmen*, die zum Beispiel in [11] oder [6] behandelt werden. Bei solch gierigen Algorithmen werden Entscheidungen getroffen, die für den Augenblick die bestmöglichen Ergebnisse liefern, allerdings nicht immer einem globalen Optimum entsprechen müssen. Sie stehen also ganz im Gegensatz zu der vorher behandelten dynamischen Programmierung, bei der wie beschrieben rekursiv berechnet wird, und dadurch ein globales Optimum erreicht wird.

7.1 Spiel mit einem Spieler

7.1.1 Der Greedy-Algorithmus

Zunächst wollen wir den verwendeten Greedy-Algorithmus für einen Spieler erläutern. Wir beginnen unsere Berechnungen zum Startzeitpunkt $k = 0$ mit gegebenen Startwerten für die CO₂-Reduktion $E(0)$ und die Mittelaufwendung $M(0)$. So wird dann die optimale Steuerung durch Minimierung des Funktionals

$$g(u(k)) = \left[E(k) + em(M(k) + u(k)) - \hat{E} \right]^2 + \left[M(k) + u(k) - \lambda(M^* - M(k)) - \hat{M} \right]^2 + (u(k))^2$$

berechnet.

Mit der berechneten Steuerung $u(k)$ können dann durch Einsetzen in die Differenzgleichungen (5.3) die neuen Werte für die Emissionsreduktion und die Mittelaufwendung für den Zeitpunkt $k = k + 1$ berechnet werden. Die iterative Minimierung wird solange wiederholt, bis der Wert für die finanzielle Mittelaufwendung den Zielwert erreicht hat. Allerdings wird hier auf die Annahme, dass dieses Ziel nach höchstens K Schritten erreicht sein muss, verzichtet. Eine Berücksichtigung der Zukunft, wie es bei der dynamischen Programmierung nach Bellman der Fall ist, ist bei diesem Algorithmus nicht möglich. Vielmehr wird sich hier auf die Berechnung des Nash-Gleichgewichts in jedem einzelnen Schritt beschränkt. Man kann so den Zeitpunkt ermitteln, wann der Spieler bei sukzessivem Minimieren seiner Funktion sein jeweiliges Ziel erreichen kann.

Das maple-Programm `greedy1spieler.mws` führt diesen Algorithmus praktisch aus. Der Ablauf soll hier mit Hilfe eines Verlaufsdiagramms dargestellt werden. Für den Quellcode sei wieder auf die beiliegende CD verwiesen.

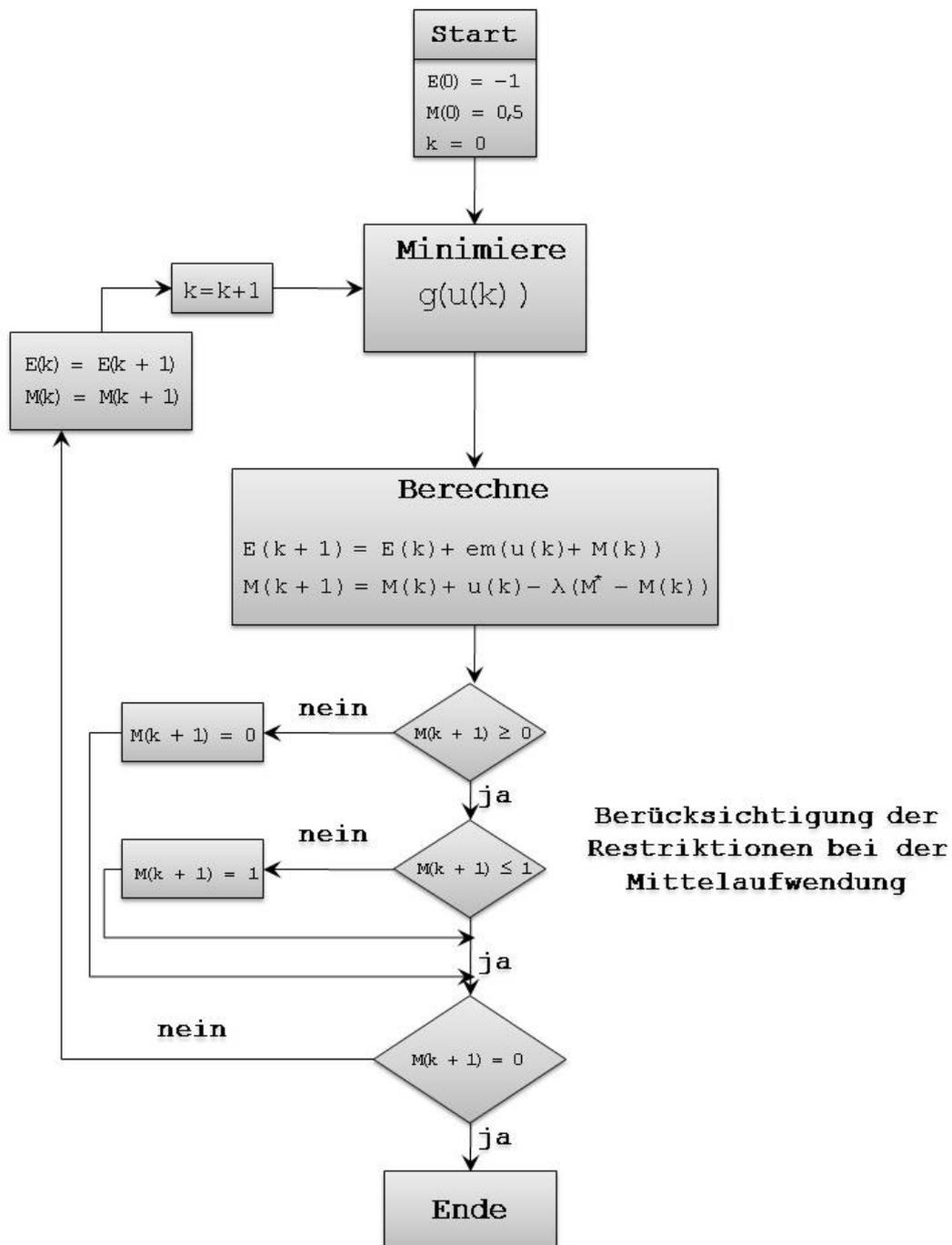


Abbildung 7.1: Darstellung des Greedy-Algorithmus bei einem Spieler

7.1.2 Ergebnisse des Greedy-Algorithmus

Für die praktische Behandlung unseres CO₂-Reduktionsmodells mit diesem Greedy-Algorithmus sind folgende Parameter gegeben.

$EHat := 1.0$	$MHat := 0$
$em := 1.0$	$lambda := 0.01$
$MStern := 1.0$	
$E(0) := -1.0$	$M(0) := 0.5$

Tabelle 7.1: Startparameter bei einem Spieler

Auch hier will unser Spieler von dem schlechtest möglichen Wert an Emissionsreduktion $E(0) = -1.0$ zu dem Zielwert von $EHat = 1.0$ gelangen. Dafür setzt er im ersten Jahr wie oben $M(0) = 0.5$ an finanziellen Mitteln ein. $em = 1.0$ bedeutet wieder, dass sich sein Mitteleinsatz in vollem Umfang positiv auf seine Emissionsreduktion auswirkt. Auch für den Wachstumsparameter soll wieder $lambda = 0.01$ gelten.

Der Verlauf der Emissionsreduktion und der Mittelaufwendung werden hier sowohl grafisch, als auch numerisch dargestellt:

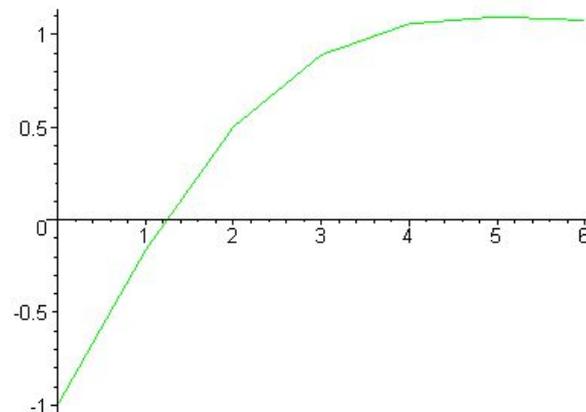


Abbildung 7.2: Emissionsreduktion bei einem Spieler

$$E = [-1.0, -0.1650000000, 0.5005666667, 0.8894537775, 1.056859358, 1.094455649, 1.075942703]$$

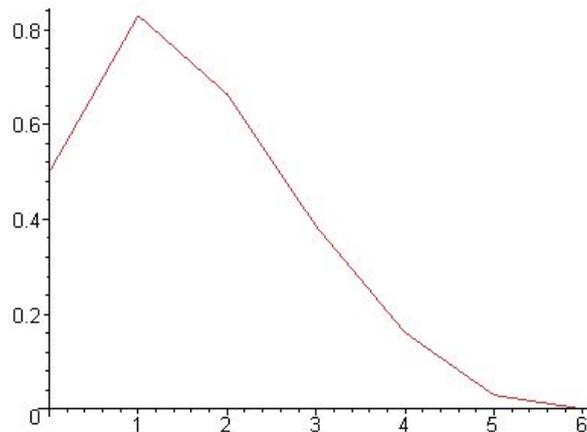


Abbildung 7.3: Mittelaufwendung bei einem Spieler

$$M = [0.5, 0.8300000000, 0.6638666667, 0.3855257779, \\ 0.1612608389, 0.0292088919, 0]$$

Sowohl an den Grafiken, als auch an den zugehörigen Werten der Trajektorien, wird deutlich, dass unser Spieler ab dem sechsten Jahr keine finanziellen Mittel mehr aufwenden muss, um sein Ziel der Emissionsreduktion zu halten. Er überschreitet unser angegebenes Ziel von $EHat = 1.0$ schon nach vier Jahren. Er muss in den darauffolgenden Jahren keine finanziellen Mittel mehr einsetzen, da er immer über dem geforderten Wert an Emissionsreduktion bleibt.

Es wäre auch hier, wie im Falle der dynamischen Programmierung eine Szenarioänderung denkbar. Wir könnten wieder negative Werte für die Mittelaufwendung zulassen. Wie sich die Emissionsreduktion und die finanzielle Mittelaufwendung ändert zeigen folgende Grafiken mit den zugehörigen Werten.

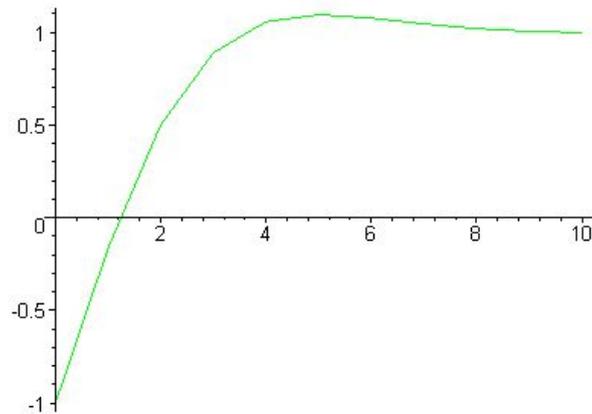


Abbildung 7.4: Emissionsreduktion ohne untere Schranke der Mittelaufwendung

$$E = [-1.0, -0.1650000000, 0.5005666667, 0.8894537775, \\ 1.056859358, 1.094455649, 1.075942703, 1.044648919, \\ 1.019379201, 1.004476593, 0.9979821060]$$

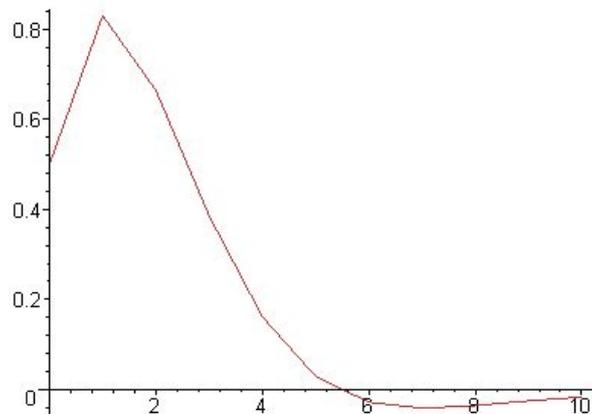


Abbildung 7.5: Mittelaufwendung ohne untere Schranke

$$M = [.5, .8300000000, .6638666667, .3855257779, .1612608389, \\ 0.02920889919, -0.02822085722, -0.04157599253, \\ -0.03568547698, -0.02525946250, -0.01674708159]$$

Auch hier überschreitet unser Spieler zunächst das Ziel unserer CO₂-Emissionsreduktion bereits nach vier Jahren, allerdings resultiert daraus eine negative Mittelaufwendung. Diese setzt sich weiter fort, so dass keine Mittelaufwendung mehr nötig ist, dem Spieler sogar jedes Jahr ein Bonus zugute kommt. Dass der Zielwert der Emissionsreduktion im letzten Jahr wieder unterschritten, ist nur einer temporäre Schwankung. Lässt man den Algorithmus länger als zehn Jahre weiter laufen, pendelt sich der Wert der Emissionsreduktion nach kleineren Schwankungen bei $E = 1$ ein. Dafür sei auf das Programm `greedy1spieler.mws` mit zum Beispiel $K := 50$ verwiesen.

7.2 Spiel mit zwei Spielern

7.2.1 Der Greedy-Algorithmus

Nun soll der Fall von zwei Spielern mit nicht-antizipierender Spielsituation bei dem Greedy-Algorithmus betrachtet werden. Der Startzeitpunkt der Berechnung ist wieder $k = 0$. Hier beginnt Spieler 1 damit sein Funktional

$$g_1(\mu_1[u_2](k), u_2(k)) = \\ \left[E_1(k) + em_{11}(M_1(k) + \mu_1[u_2](k)) + em_{12}(M_2(k) + u_2(k)) - \hat{E}_1 \right]^2 + \\ + [M_1(k) + \mu_1[u_2](k) - \lambda(M_1^* - M_1(k))]^2 + (\mu_1[u_2](k))^2$$

nach $\mu_1[u_2](k)$ abhängig von $u_2(k)$ zu minimieren, um seine optimale Strategie zu berechnen. Dies liefert uns sein optimales $\mu_1^*[u_2](k)$, das von $u_2(k)$ abhängt.

Nun berechnet Spieler 2 seine optimale Kontrolle, muss dafür aber die eben berechnete optimale Strategie von Spieler 1 berücksichtigen. Er minimiert nun also sein Funktional

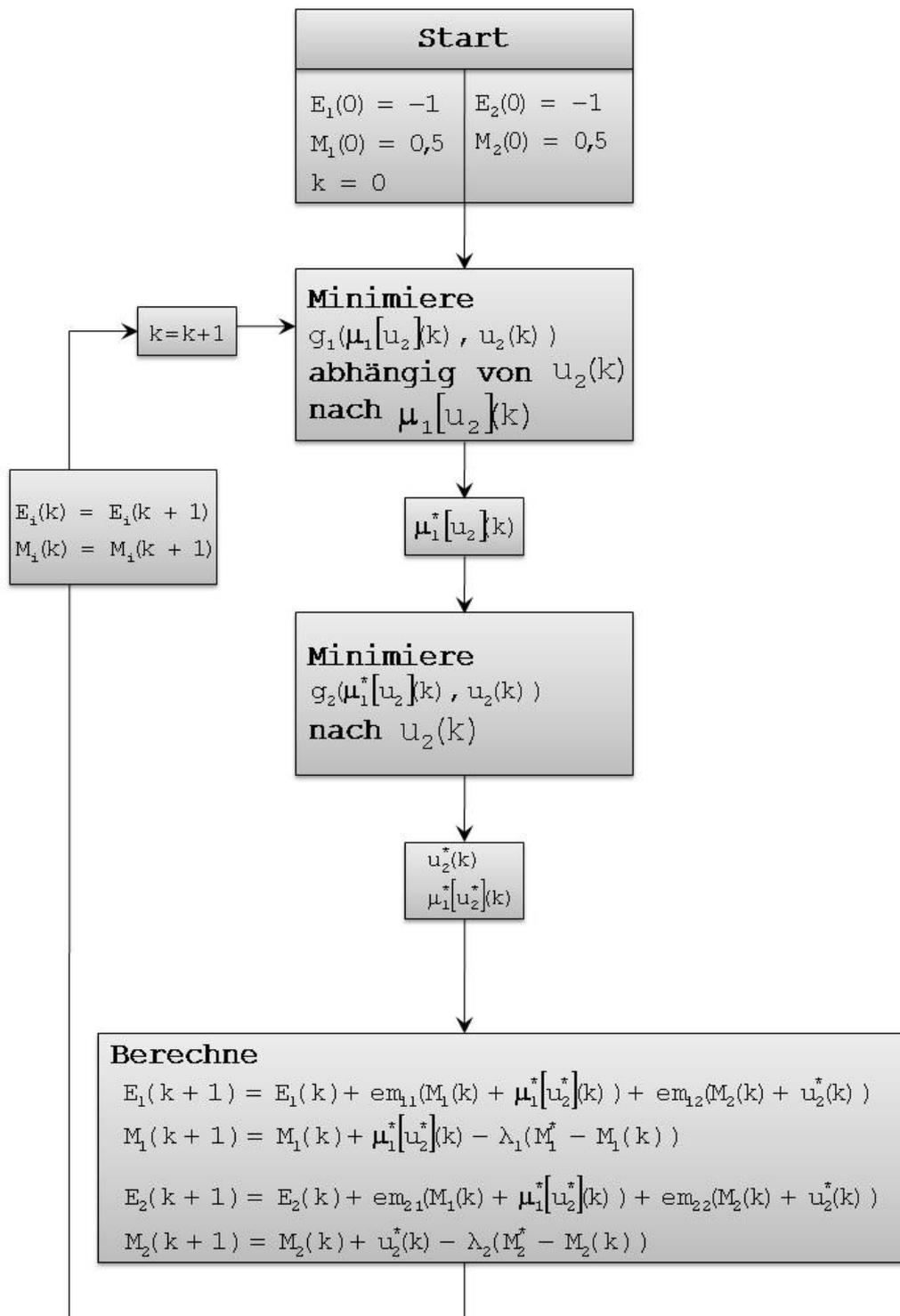
$$g_2(\mu_1^*[u_2](k), u_2(k)) = \\ \left[E_2(k) + em_{21}(M_1(k) + \mu_1^*[u_2](k)) + em_{22}(M_2(k) + u_2(k)) - \hat{E}_2 \right]^2 + \\ + [M_2(k) + u_2(k) - \lambda(M_2^* - M_2(k))]^2 + u_2(k)^2$$

und erhält seine optimale Kontrolle $u_2^*(k)$. Spieler 1 kann nun daraus den Wert seiner optimalen Strategie $\mu_1^*[u_2^*](k)$ berechnen.

Mit den beiden Steuerungen ist es nun möglich durch Einsetzen in die Differenzgleichungen die neuen Werte für Emissionsreduktion und finanzielle Mittelaufwendung für $k = k + 1$ zu berechnen. Diese werden dann in die Funktionale der beiden Spieler eingesetzt und die Minimierung kann mit den neuen Werten wieder durchgeführt werden. Diese iterative Berechnung wird solange wiederholt

bis die Werte der Mittelaufwendung von Spieler 2 den Zielwert erreichen. Da dieser einen Nachteil hat, wird Spieler 2 sein Ziel später erreichen als sein Gegner, der die nicht-antizipierende Strategie zur Verfügung hat.

Das maple-Programm `greedy2spieler.mws` führt diesen Algorithmus aus. Des-
sen Ablauf soll hier wieder mit Hilfe eines Verlaufsdiagramms dargestellt werden.
Für den Quellcode sei auch hier auf die beiliegende CD verwiesen.



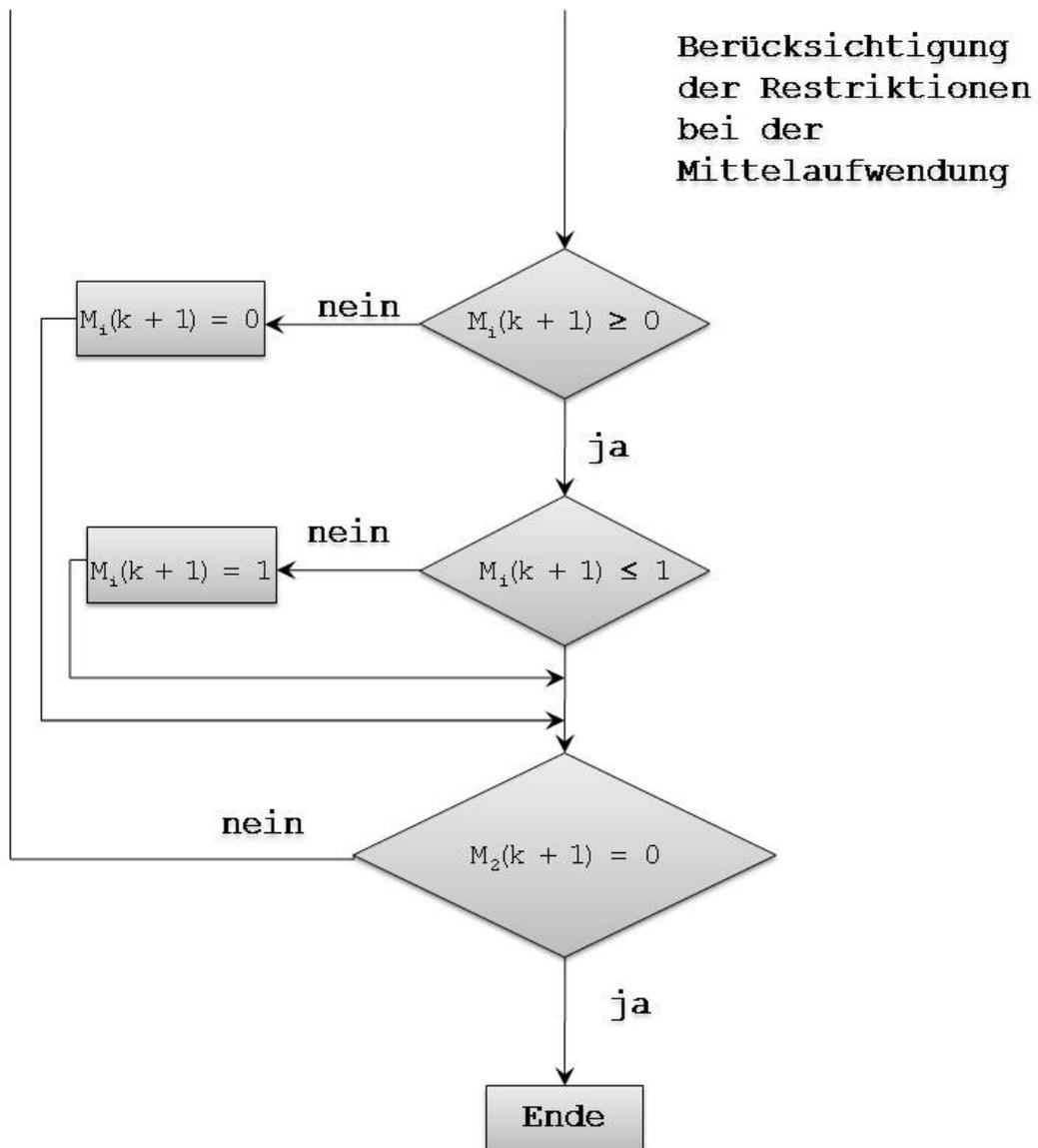


Abbildung 7.6: Darstellung des Greedy-Algorithmus bei zwei Spielern

7.2.2 Ergebnisse des Greedy-Algorithmus

Nun wollen wir diesen iterativen Algorithmus auf nicht-antizipierende Strategien von Spieler 1 anwenden. Für die Betrachtung werden folgende Parameter verwendet:

$E1(0) := -1.0$	$EHat1 := 1.0$	
$E2(0) := -1.0$	$EHat2 := 1.0$	
$M1(0) := 0.5$	$MHat1 := 0$	$MStern1 := 1.0$
$M2(0) := 0.5$	$MHat2 := 0$	$MStern2 := 1.0$
$lambda1 := 0.01$	$lambda2 := 0.01$	

Tabelle 7.2: Startparameter bei zwei Spielern bei nicht-antizipierender Spielsituation

Die Werte der em -Matrix lauten hier:

$em11 := 1.0$	$em12 := -0.5$
$em21 := -0.5$	$em22 := 1.0$

Tabelle 7.3: Werte der em -Matrix

Die Startwerte $E1(0) = -1.0$, bzw. $E2(0) = -1.0$ und $M1(0) = 0.5$, bzw. $M2(0) = 0.5$ sollen wieder zu den Endwerten $EHat1 = 1.0$, bzw. $EHat2 = 1.0$ und $MHat1 = 0.5$, bzw. $MHat2 = 0.5$ gesteuert werden. Dabei stehen den Spielern maximal die finanziellen Mittel $MStern1 = 1.0$ und $MStern2 = 1.0$ zur Verfügung. Für die Werte der em -Matrix und der Wachstumsparameter $lambda1$ und $lambda2$ gelten die gleichen Annahmen wie zuvor bei der dynamischen Programmierung.

Nun wollen wir wieder die Ergebnisse sowohl grafisch als auch numerisch betrachten. Die Emissionsreduktion verläuft bei den beiden Spielern wie folgt.

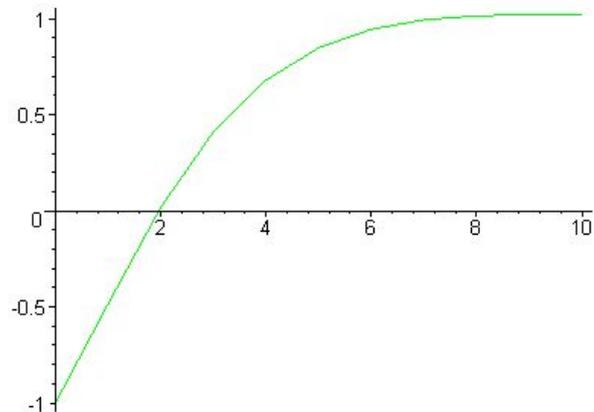


Abbildung 7.7: Emissionsreduktion von Spieler 1

$$E1 = [-1.0, -0.4843398534, 0.0185042070, 0.4109120096, \\ 0.6806570024, 0.8492087443, 0.9454035069, 0.9945912384, \\ 1.015662463, 1.021382535, 1.019741873]$$

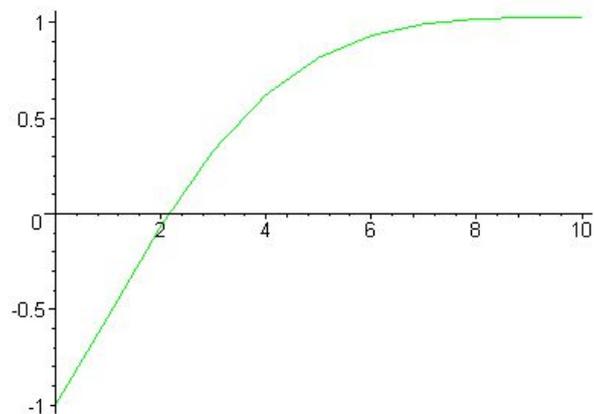


Abbildung 7.8: Emissionsreduktion von Spieler 2

$$E2 = [-1.0, -0.5393154030, -0.06655175461, 0.3297055546, \\ 0.6217464800, 0.8152644768, 0.9306552338, 0.9913559113, \\ 1.017759501, 1.025015673, 1.023092599]$$

Die finanzielle Mittelaufwendung verhält sich bei den beiden Spielern folgendermaßen.

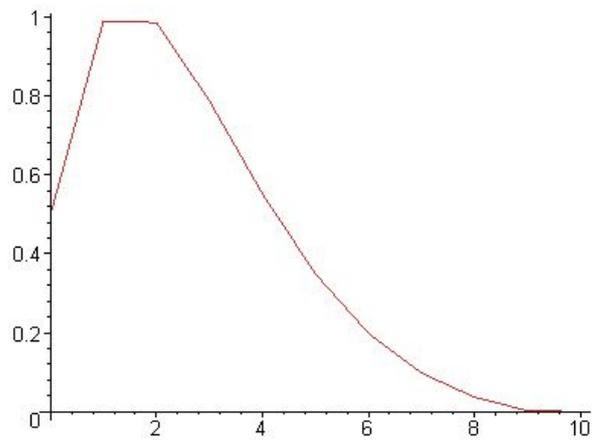


Abbildung 7.9: Mittelaufwendung von Spieler 1

$$M1 = [0.5, 0.9896699266, 0.9855312101, 0.7872372565, \\ 0.5522263134, 0.3492699163, 0.1986795539, 0.09803755554, \\ 0.03667773434, 0.002830988283, 0]$$

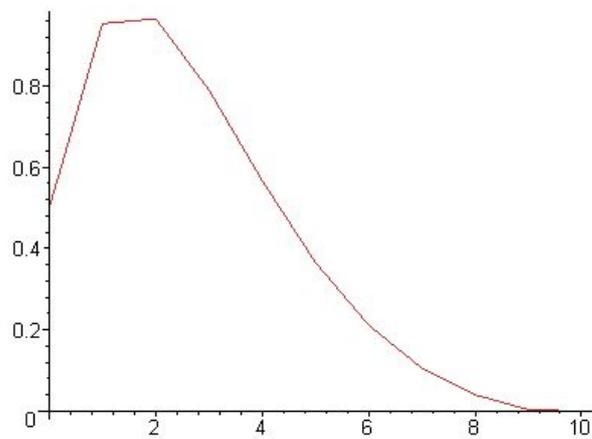


Abbildung 7.10: Mittelaufwendung von Spieler 2

$$M2 = [0.5, 0.9530195600, 0.9651110990, 0.7895993924, \\ 0.5671138897, 0.3660629624, 0.2116448138, 0.1058425052, \\ 0.04031069395, 0.003891383910, 0]$$

Bereits nach $k = 8$ Jahren überschreiten die beiden Spieler zum ersten Mal ihren Zielwert der Emissionsreduktion. Nach weiteren zwei Jahren ist kein finanzieller Mittelaufwand mehr nötig. Da dies das Ziel war, bricht der Greedy-Algorithmus also nach $k = 10$ Jahren ab.

Im nächsten Kapitel sollen nun die Ergebnisse des Greedy-Algorithmus mit denen der dynamischen Programmierung verglichen werden.

Analyse und Vergleich der Algorithmen

In den vergangenen beiden Kapiteln wurden zwei gänzlich verschiedene Arten vorgestellt, das CO₂-Reduktionsmodell zu behandeln. Zum einen wurden bei der dynamischen Programmierung bei der Berechnung der optimalen Strategien und der optimalen Wertefunktion auch die zukünftigen Kosten mit einbezogen. Bei dem iterativen Greedy-Algorithmus hingegen wurden nur die zu dem Zeitpunkt bestmöglichen Entscheidungen getroffen ohne die Zukunft zu berücksichtigen. In diesem Kapitel sollen nun die beiden Ansätze verglichen werden. Die verwendeten Werte und Berechnungen sind in `vergleichgreedydynprog.mws` auf der beiliegenden CD zu finden.

8.1 Spiel mit einem Spieler

8.1.1 Abweichungen der Lösungstrajektorien von den Zielwerten

Zum Vergleich der beiden Algorithmen wollen wir die jeweiligen Abweichungen der Lösungstrajektorien von den Zielwerten betrachten. Um allerdings eine Vergleichbarkeit zu schaffen, müssen wir zunächst ermitteln nach wievielen Schritten der Greedy-Algorithmus das Ziel erreicht und diesen Wert auch bei der dynamischen Programmierung verwenden.

Die Lösungstrajektorie des finanziellen Mittelaufwands

$$M^{Greedy} = [0.5, 0.8300000000, 0.6638666667, 0.3855257779, \\ 0.1612608389, 0.0292088991, 0]$$

zeigt, dass dies nach $k = 6$ Jahren der Fall ist.

Wir wählen also auch bei der dynamischen Programmierung als Endzeitpunkt $K = 6$. Dies liefert folgende Werte der finanziellen Mittelaufwendung.

$$M^{Dyn} = [0.5, 0.9276868655, 0.6129167563, 0.2962587967, \\ 0.1118201020, 0.02864979722, 0]$$

Wir wollen nun die Abweichung der Trajektorien von dem Zielwert berechnen. Dazu summieren wir die Beträge der Abweichungen der jeweiligen Einzelwerte wie folgt auf.

$$\sum_{k=0}^K |M_k - \hat{M}|$$

Damit ergibt sich für den Greedy-Algorithmus eine Gesamtabweichung von

$$\Delta^{Greedy}(M) = 2.569862183$$

und bei der dynamischen Programmierung eine Gesamtabweichung von

$$\Delta^{Dyn}(M) = 2.477332318$$

Mit der gleichen Methode wollen wir nun die Lösungstrajektorien der Emissionsreduktionen miteinander vergleichen.

Die Werte der Emissionsreduktion bei Verwendung des Greedy-Algorithmus lauten

$$E^{Greedy} = [-1.0, -0.1650000000, 0.5005666667, 0.8894537775, \\ 1.056859358, 1.094455649, 1.075942703]$$

die bei Verwendung der dynamischen Programmierung lauten

$$E^{Dyn} = [-1, -0.0673131345, 0.5463267531, 0.8464563818, \\ 0.9653138953, 1.002845491, 1.011322655]$$

Die Gesamtabweichungen erhalten wir mit Hilfe folgender Summe

$$\sum_{k=0}^K |E_k - \hat{E}|$$

Diese lauten dann

$$\Delta^{Greedy}(E) = 4.002237266$$

$$\Delta^{Dyn}(E) = 3.723384250$$

Es zeigt sich also, dass die Abweichung von den jeweiligen Zielwerten bei der dynamischen Programmierung immer kleiner ist.

8.1.2 Gesamtkosten

Bei dem Spiel mit einem Spieler handelt es sich, wie bereits ausführlich erläutert, um ein Minimierungsproblem. Ziel ist es also die geringsten Gesamtkosten zu erzielen. Wir zeigen nun, dass die dynamische Programmierung ein globales Minimum berechnet, während der Greedy-Algorithmus tatsächlich nur die Summe der jeweils lokalen Minima berechnet, die größer ist als dieses globale Optimum.

Für diesen Vergleich addieren wir zunächst für den Greedy-Algorithmus die Kosten jeder Stufe

$$\begin{aligned} \text{Kosten}^{\text{Greedy}} = & [2.158350000, 0.7171909267, 0.2364643485, \\ & 0.07681446517, 0.02506794989, 0.008841085298] \end{aligned}$$

und erhalten damit die Gesamtkosten

$$\text{Kosten}_{\text{gesamt}}^{\text{Greedy}} = 3.222728776$$

Anschließend berechnen wir den Wert der optimalen Wertefunktion bei der dynamischen Programmierung

$$\begin{aligned} V[0]((E, M)) = & 0.9253655529 - 1.877495163E - 0.8371491401M + \\ & 0.9532968327E^2 + 0.8332764541EM + 0.8300726316M^2 \end{aligned}$$

mit den Startwerten $E := -1$ und $M := 0.5$ und erhalten dann als Gesamtkosten

$$\text{Kosten}_{\text{gesamt}}^{\text{Dyn}} = 3.128462910$$

Die dynamische Programmierung hat uns also die tatsächlichen, minimalen Kosten geliefert.

8.2 Spiel mit zwei Spielern

8.2.1 Abweichungen der Lösungstrajektorien von den Zielwerten

Wie im Fall von einem Spieler, wollen wir für die nicht-antizipierende Spielsituation von zwei Spielern die Gesamtabweichungen der beiden Algorithmen vergleichen. Um die Vergleichbarkeit herzustellen, müssen wir auch hier ermitteln, nach wie vielen Schritten der Greedy-Algorithmus das Ziel der finanziellen Mittelaufwendung bei beiden Spielern erreicht. Dazu betrachten wir die Trajektorien von Spieler 1

$$M_1^{Greedy} = [0.5, 0.9896699266, 0.9855312101, 0.7872372565, \\ 0.5522263134, 0.3492699163, 0.1986795539, 0.09803755554, \\ 0.03667773434, 0.002830988283, 0]$$

$$E_1^{Greedy} = [-1.0, -0.4843398534, 0.0185042070, 0.4109120096, \\ 0.6806570024, 0.8492087443, 0.9454035069, 0.9945912384, \\ 1.015662463, 1.021382535, 1.019741873]$$

und von Spieler 2

$$M_2^{Greedy} = [0.5, 0.9530195600, 0.9651110990, 0.7895993924, \\ 0.5671138897, 0.3660629624, 0.2116448138, 0.1058425052, \\ 0.04031069395, 0.003891383910, 0]$$

$$E_2^{Greedy} = [-1.0, -0.5393154030, -0.06655175461, 0.3297055546, \\ 0.6217464800, 0.8152644768, 0.9306552338, 0.9913559113, \\ 1.017759501, 1.025015673, 1.023092599]$$

Wir sehen also, dass nach $k = 10$ Jahren sowohl bei Spieler 1, als auch bei Spieler 2 das Ziel erreicht wird. Wir führen deshalb die dynamische Programmierung auch mit Endzeitpunkt $K = 10$ durch. Dann erhalten wir folgende Trajektorien für Spieler 1

$$M1 = [0.5, 1, 1, 0.7296360460, 0.3924635127, 0.1573195293, 0.03373388699, 0, 0, 0, 0]$$

$$E_1^{Dyn} = [-1, -0.1905765625, 0.3832361191, 0.743927846, \\ 0.9352545786, 1.011442512, 1.028807765, 1.023993583, \\ 1.017247070, 1.013709888, 1.012656662]$$

und Spieler 2

$$M2 = [0.5, 1, 1, 0.7378886388, 0.4050597254, 0.1684645187, 0.04127552359, 0, 0, 0, 0]$$

$$E_2^{Dyn} = [-1, -0.249883929, 0.3151493630, 0.6882199788, \\ 0.8983172417, 0.9910337160, 1.019544249, 1.020937278, \\ 1.016442969, 1.014054262, 1.013504052]$$

Die Gesamtabweichungen werden wieder wie oben berechnet. Dabei ergibt sich für die finanzielle Mittelaufwendung von Spieler 1

$$\Delta^{Greedy}(M_1) = 4.500160456$$

und der von Spieler 2

$$\Delta^{Greedy}(M_2) = 4.502596300$$

Für die Gesamtabweichung bei der Emissionsreduktion gilt für Spieler 1

$$\Delta^{Greedy}(E_1) = 5.641850016$$

und für Spieler 2

$$\Delta^{Greedy}(E_2) = 5.983007274$$

Bei der dynamischen Programmierung ergibt sich folgende Gesamtabweichung bei der finanziellen Mittelaufwendung für Spieler 1

$$\Delta^{Dyn}(M_1) = 3.813152975$$

und für Spieler 2

$$\Delta^{Dyn}(M_2) = 3.852688407$$

Bei der Gesamtabweichung der Emissionsreduktion erhalten wir bei Spieler 1

$$\Delta^{Dyn}(E_1) = 4.236015498$$

und bei Spieler 2

$$\Delta^{Dyn}(E_2) = 4.441646439$$

Wie man deutlich sieht, sind die Gesamtabweichungen bei Verwendung der dynamischen Programmierung immer kleiner. Dieser Algorithmus steuert also besser zu unseren gewünschten Zielwerten als der Greedy-Algorithmus.

Desweiteren fällt bei Betrachtung der Lösungstrajektorien auf, dass bei der dynamischen Programmierung schon nach $k = 7$ keine finanzielle Mittelaufwendung mehr nötig ist, obwohl das Ziel erst nach $k = 10$ zu erreichen war.

8.2.2 Gesamtkosten

An dieser Stelle wollen wir auch bei dem Fall von zwei Spielern in nicht-anti-zipierender Spielsituation die beiden Algorithmen bezüglich ihrer Gesamtkosten vergleichen. Dazu addieren wir bei dem Greedy-Algorithmus die Kosten jeder Stufe auf und betrachten bei der dynamischen Programmierung die optimalen Wertefunktionen mit unseren gegebenen Startwerten.

Die Kosten von Spieler 1 des Greedy-Algorithmus lauten:

$$Kosten_1^{Greedy} = [3.427409699, 1.934622045, 1.006030290, 0.4611684900, \\ 0.18412125580.06321427084, 0.01822071046, 0.004330065139, \\ 0.001051522013, 0.0006101071460]$$

und ergeben folgende Gesamtkosten:

$$Kosten_{1\text{ gesamt}}^{Greedy} = 9.635345612$$

Die Kosten von Spieler 2 sind:

$$Kosten_2^{Greedy} = [3.487520109, 2.069129866, 1.103443857, 0.5132618933, \\ 0.2068288772, 0.07152954836, 0.02086543909, 0.005142806767, \\ 0.001360368817, 0.0007757354500]$$

und wir erhalten diese Gesamtkosten:

$$Kosten_{2\text{ gesamt}}^{Greedy} = 9.840329164$$

Die optimale Wertefunktion von Spieler 1 lautet:

$$V_1[0]((E1, M1)) = -1.789084450M1 - 4.447715046E1 + 6.313170217+ \\ + 0.8066727844M1^2 + 1.268299842E1^2 + 0.9070222437E1M1$$

Für die Startwerte $E1 = -1$, $M1 = 0.5$ ergibt sich:

$$V_1[0]((E1, M1)) = 10.88279995$$

Für Spieler 2 gilt:

$$V_2[0]((E2, M2)) = 6.693331625 - 4.539922830E2 - 1.813873173M2+ \\ + 0.8038555707M2^2 + 1.270613576E2^2 + .9004849714E2M2$$

und mit den Startwerten $E2 = -1$, $M1 = 0.5$ erhalten wir:

$$V_2[0]((E2, M2)) = 11.34765285$$

An diesen Werten wird deutlich, dass die dynamische Programmierung größere Gesamtkosten als der Greedy-Algorithmus liefert. Dies ist zunächst ungewöhnlich, hieß es doch, dass gerade durch die dynamische Programmierung immer die minimalen Kosten erreicht werden und damit ein globales Minimum erzielt wird. Bei dem Spiel mit nur einem Spieler war dies auch der Fall, da wir hier nur ein Minimierungsproblem zu untersuchen hatten. Im Fall von zwei Spielern geht es jedoch, wie in Kapitel 4 ausführlich erläutert um die Berechnung eines *Optimums*.

Dieses muss unter Berücksichtigung des Einflusses des jeweils anderen Spielers berechnet werden. Wir haben also kein reines Minimierungsproblem der Gesamtkosten. Wir haben viel mehr das Ziel *optimal* zu unseren Zielwerten hinzusteuern. Dies wird durch die dynamische Programmierung auch besser durchgeführt, wie der obige Vergleich der Gesamtabweichung zeigt. Unser Ziel hat damit die dynamische Programmierung trotz der höheren Gesamtkosten besser erreicht als der Greedy-Algorithmus.

Das Stackelberg-Gleichgewicht

Bisher wurde in dieser Arbeit der Fall betrachtet, dass Spieler 1 durch die nicht-antizipierende Strategie gegenüber Spieler 2 einen gewissen Vorteil hat. Spieler 2 durfte allerdings seine Kontrolle zu seinem eigenen Nutzen selbst wählen. In diesem Kapitel werden wir sehen, dass ein Spieler seinen Gegner aber auch zu einer bestimmten Strategie zwingen kann. Bei den nun folgenden Spielen sind immer ein *Leader* und ein *Follower* beteiligt. Das Ziel für den *Leader* ist es, dass das sogenannte *Stackelberg-Gleichgewicht* erreicht wird. Zunächst setzen wir den theoretischen Rahmen um das Stackelberg-Gleichgewicht in der Praxis behandeln zu können. Danach folgt ein Beispiel, an dem der Unterschied zu dem bisher betrachteten Nash-Gleichgewicht deutlich werden soll. Abschließend betrachten wir ein dynamisches Spiel, dem eine leicht veränderte Variante unseres CO₂-Reduktionsmodells zu Grunde liegt. Sowohl die theoretische Einführung als auch die Anwendung auf Beispiele basieren auf [2].

9.1 Theoretische Einführung

Wir bezeichnen die zulässigen Strategien des *Leaders* mit $u_L \in \mathcal{U}_L$ und die des *Followers* mit $u_F \in \mathcal{U}_F$. Die Kostenfunktionen der beiden Spieler bezeichnen wir wie in den vorhergehenden Kapiteln mit $J_L(x, u_L, u_F)$ bzw. $J_F(x, u_L, u_F)$.

Zunächst definieren wir wie in Kapitel 4 die Menge der optimalen Antworten des *Followers*.

Definition 9.1 (Menge der optimalen Antworten des *Followers*).

In einem endlichen Zwei-Personen Spiel, ist die Menge $\mathcal{U}_F^*(u_L) \subset \mathcal{U}_F$ für jedes

$u_L \in \mathcal{U}_L$ die Menge der optimalen Antworten des Followers und wird definiert durch

$$\mathcal{U}_F^*(u_L) = \{\xi \in \mathcal{U}_F : J_F(u_L, \xi) \leq J_F(u_L, u_F), \forall u_F \in \mathcal{U}_F\}$$

Definition 9.2 (Stackelberg-Strategie des Leaders).

In einem endlichen Zwei-Personen Spiel ist eine Strategie $u_L^* \in \mathcal{U}_L$ im Stackelberg-Gleichgewicht für den Leader, wenn

$$\sup_{u_F \in \mathcal{U}_F^*(u_L^*)} J_L(u_L^*, u_F) = \inf_{u_L \in \mathcal{U}_L} \sup_{u_F \in \mathcal{U}_F^*(u_L)} J_L(u_L, u_F) \hat{=} J_L^* \quad (9.1)$$

Dabei bezeichnen wir $u_L^* \in \mathcal{U}_L$ als Stackelberg-Strategie des Leaders und J_L^* als Stackelberg-Kosten des Leaders.

Satz 9.3.

Bei jedem endlichen Zwei-Personen-Spiel gibt es eine Stackelberg-Strategie für den Leader.

Beweis.

Da \mathcal{U}_L und \mathcal{U}_F endliche Mengen sind und $\mathcal{U}_F^*(u_L)$ eine Teilmenge von \mathcal{U}_F ist für jedes $u_L \in \mathcal{U}_L$, folgt dies aus (9.1) \square

Bemerkung 9.4.

Die Stackelberg-Strategie des Leaders ist nicht notwendigerweise eindeutig. Aber diese Uneindeutigkeit der Gleichgewichtsstrategie bereitet hier keine Probleme, da die Stackelberg-Kosten des Leaders eindeutig sind.

Bemerkung 9.5.

Wenn $\mathcal{U}_F^*(u_L)$ für jedes $u_L \in \mathcal{U}_L$ eine einelementige Menge ist, dann existiert eine Abbildung $T : \mathcal{U}_L \rightarrow \mathcal{U}_F$, so dass aus $u_F \in \mathcal{U}_F^*(u_L)$ folgt, dass $u_L = Tu_L$ gilt. Dies entspricht dem Fall, dass die optimale Antwort des Followers (gegeben durch die Abbildung T) für jede Strategie des Leaders eindeutig ist. Dies führt zu folgender vereinfachten Version von (9.1)

$$J_L(u_L^*, Tu_L^*) = \inf_{u_L \in \mathcal{U}_L} J_L(u_L, Tu_L) \hat{=} J_L^*$$

J_L^* sind hier nicht mehr nur die, dem Leader sicheren Stackelberg-Kosten aus Definition 9.2, sondern die Kosten, die tatsächlich erreicht werden.

Für den Follower ist die Stackelberg-Strategie in einem Stackelberg-Spiel jede optimale Antwort auf eine angekündigte Stackelberg-Strategie des Leaders. Dazu definieren wir folgende Begriffe.

Definition 9.6 (Stackelberg-Lösung und Stackelberg-Ergebnis).

Sei $u_L^* \in \mathcal{U}_L$ eine Stackelberg-Strategie für den Leader. Dann ist jedes Element $u_F^* \in \mathcal{U}_F^*(u_L^*)$ eine optimale Strategie des Followers, die im Gleichgewicht mit u_L^* ist. Das Paar $\{u_L^*, u_F^*\}$ ist dann eine Stackelberg-Lösung des Spiels und das Kostenpaar $(J_L(u_L^*, u_F^*), J_F(u_L^*, u_F^*))$ ist das zugehörige Stackelberg-Ergebnis.

Bemerkung 9.7.

Beachte, dass die Kosten des Stackelberg-Ergebnisses aus Definition 9.6 $J_L(u_L^*, u_F^*)$ kleiner als die Stackelberg-Kosten J_L^* aus Definition 9.2 sein können. Nur wenn $\mathcal{U}_F^*(u_L)$ für jedes $u_L \in \mathcal{U}_L$ eine einelementige Menge ist, stimmen diese überein.

9.2 Anwendung auf Beispiele

In diesem Abschnitt soll nun das Konzept der Stackelberg-Lösung anhand von Beispielen näher gebracht werden. Wir beginnen mit einem einfachen Matrix-Spiel mit zwei Spielern, anhand dessen der Unterschied des Nash-Gleichgewichts und des Stackelberg-Gleichgewichts deutlich werden soll. Anschließend wird ein dynamisches Spiel vorgestellt, das auf unserem CO₂-Reduktionsmodell in leicht veränderter Form basiert. Hier werden wir sehen, dass der *Leader* tatsächlich in der Lage ist, dem *Follower* eine Strategie vorzuschreiben.

Beispiel 9.8.

Gegeben sei folgendes Matrix-Spiel mit zwei Spielern. Beide Spieler versuchen dabei ihre Kosten zu minimieren. Aus Gründen der Übersichtlichkeit werden die Kosten in zwei Matrizen dargestellt. Die erste Matrix stellt die Kosten von Spieler 1, die zweite die Kosten von Spieler 2 dar.

		Spieler 2		
		A	B	C
Spieler 1	A	0 Stackelberg 1	2	1.5 Stackelberg 2
	B	1	1 Nash-GG	3
	C	-1	2	2

Abbildung 9.1: Kosten von Spieler 1

		Spieler 2		
		A	B	C
Spieler 1	A	-1 Stackelberg 1	1	-0.6 Stackelberg 2
	B	2	0 Nash-GG	1
	C	0	1	-0.5

Abbildung 9.2: Kosten von Spieler 2

Für dieses Spiel existiert ein Nash-Gleichgewicht der Form (B, B) . Ein Spieler könnte sich also nur schlechter stellen als zuvor, wenn er von dieser Strategie abweicht. Die Kosten betragen dann $(1, 0)$ für die beiden Spieler.

Außerdem können wir hier nach dem Stackelberg-Gleichgewicht suchen, das sich im Allgemeinen von dem eben genannten Nash-Gleichgewicht unterscheidet. Dazu wollen wir annehmen, dass die Rollen der beiden Spieler nicht mehr symmetrisch verteilt sind und Spieler 1 seine Strategie dem Gegner aufzwingen kann und somit der *Leader* des Spiels ist. Spieler 2 ist damit der *Follower* in diesem Spiel.

Bevor Spieler 1 seine Strategie bekannt gibt, zieht er alle möglichen Antworten von Spieler 2 bei seiner Entscheidung in Betracht. Er wird also die Strategie wählen, die für ihn die niedrigsten Kosten verspricht. Folgende Entscheidungsmöglichkeiten stehen Spieler 1 zur Verfügung.

- Spieler 1 wählt A , dann wählt Spieler 2 A , da er seine eigenen Kosten minimieren will
→ Kosten von Spieler 1 sind 0
- Spieler 1 wählt B , dann wählt Spieler 2 B , da er seine eigenen Kosten minimieren will
→ Kosten von Spieler 1 sind 1

- Spieler 1 wählt C , dann wählt Spieler 2 C , da er seine eigenen Kosten minimieren will
→ Kosten von Spieler 1 sind 2

Für Spieler 1 ist also die beste Strategie A zu wählen, da das die geringsten Kosten für ihn bedeutet. Diese Strategie nennen wir dann die *Stackelberg-Strategie von Spieler 1* und das zugehörige Gleichgewicht (A, A) das *Stackelberg-Gleichgewicht*. Im Fall von Spieler 1 als *Leader* des Spiels bedeutet dies Kosten von $(0, -1)$ für die beiden Spieler. Dieses stellt beide Spieler besser als das Nash-Gleichgewicht von oben mit Kosten $(1, 0)$ der beiden Spieler. Dass dies nicht allgemein gültig ist, sehen wir an folgender Veränderung der Spielsituation.

Wir nehmen nun an, dass die beiden Spieler die Rollen tauschen, also Spieler 2 der *Leader* des Spiels ist und Spieler 1 der *Follower*. Das führt bei analogen Überlegungen zu folgendem Stackelberg-Gleichgewicht (A, C) mit Kosten $(1.5, -0.6)$ der beiden Spieler. Das stellt Spieler 1 offensichtlich schlechter als im Fall des Nash-Gleichgewichts $(1, 0)$, während sich Spieler 2 wieder besser stellt.

Beispiel 9.9.

Wir betrachten nun ein zwei-stufiges dynamisches Spiel, das auf unserem Modell zur CO_2 -Reduktion basiert. In diesem Spiel soll die Situation dargestellt werden, dass es eine höhere Instanz gibt, die für die Einhaltung des Kyoto-Protokolls sorgen soll. Diese ist in diesem Spiel unser *Leader*. Das Land, das eine gewisse Emissionsreduktion erreichen muss ist dann der *Follower*.

Bei unserem Spiel handelt der *Leader* erst in Stufe 2 und kennt sowohl $\begin{pmatrix} E(1) \\ M(1) \end{pmatrix}$,
als auch $\begin{pmatrix} E(2) \\ M(2) \end{pmatrix}$.

Der *Follower* hat dabei nur Kenntniss über $\begin{pmatrix} E(1) \\ M(1) \end{pmatrix}$.

Eine zulässige Strategie u_L des *Leaders* ist also eine Abbildung aus dem $R \times R \rightarrow R$, die des *Followers* u_F eine aus $R \rightarrow R$. Die jeweiligen Strategie-Räume seien wie oben als \mathcal{U}_L bzw. \mathcal{U}_F definiert.

Für dieses zwei-stufige dynamische Spiel müssen wir unser System (5.3) modifizieren. Zunächst vereinfachen wir die Differenzgleichung der finanziellen Mittelaufwendung, dass diese nicht mehr durch die jeweilige Steuerung beeinflusst werden. Dies liefert uns anschaulichere Gleichungen und Werte bei der Berechnung mit maple. Außerdem verwenden der *Leader* und der *Follower* in diesem dynamischen Spiel das gleiche System an Differenzgleichungen, allerdings handeln sie zu verschiedenen Zeitpunkten. Es ergeben sich daher folgende Gleichungen für

die Handlung des *Followers* zum Zeitpunkt $k = 1$:

$$\begin{aligned} E(2) &= E(1) + em(M(1) + u_F) \\ M(2) &= M(1) - \lambda(M^* - M(1)) \end{aligned}$$

Dem *Leader* stehen zum Zeitpunkt $k = 2$ diese Gleichungen zur Verfügung:

$$\begin{aligned} E(3) &= E(2) + em(M(2) + u_L) \\ M(3) &= M(2) - \lambda(M^* - M(2)) \end{aligned}$$

Die Kostenfunktionale der beiden Spieler sehen dabei wie folgt aus:

$$\begin{aligned} J_L &= (E(3) - \hat{E})^2 + (M(3) - \hat{M})^2 + u_L^2 + u_F^2 \\ J_F &= (E(3) - \hat{E})^2 + (M(3) - \hat{M})^2 + u_F^2 \end{aligned}$$

Der *Leader* gibt schon im Voraus seine Strategie bekannt, der *Follower* hat also nur die Möglichkeit darauf zu reagieren, was bedeutet, seine optimale Strategie wäre:

$$\begin{aligned} \inf_{u_F \in R} J_F &= \left\{ \inf_{u_F \in R} (E(3) - \hat{E})^2 + (M(3) - \hat{M})^2 + u_F^2 \right\} \\ &= \inf_{u_F \in R} \left\{ ((E(1) + em(M(1) + u_F) + em(M(2) + u_L)) - \hat{E})^2 + \right. \\ &\quad \left. + ((M(1) - \lambda(M^* - M(1))) - \lambda(M^* - M(2))) - \hat{M})^2 + u_F^2 \right\} \end{aligned}$$

Die Strategie u_F des *Followers* ist also eine Abbildung $R \times \mathcal{U}_L \rightarrow R$. Dafür gibt es aber keine einfache Beschreibung, da sie vom *Leader* abhängt.

Wir betrachten deshalb nun den *Leader*, dieser will sein Optimierungsproblem lösen:

$$\begin{aligned} \inf_{u_L \in \mathcal{U}_L} J_L &= \left\{ \inf_{u_L \in \mathcal{U}_L} (E(3) - \hat{E})^2 + (M(3) - \hat{M})^2 + u_L^2 + u_F^2 \right\} \\ &= \inf_{u_L \in \mathcal{U}_L} \left\{ ((E(2) + em(M(2) + u_L) - \hat{E})^2 + \right. \\ &\quad \left. + (M(2) - \lambda(M^* - M(2))) - \hat{M})^2 + u_L^2 + u_F^2 \right\} \end{aligned}$$

Um die Stackelberg-Lösung zu finden, müssen wir zunächst die untere Schranke der Stackelberg-Kosten des *Leaders* bestimmen. Damit können wir dann seine geeignete Strategie finden. Die Lösung soll dann diese untere Schranke im Hinblick auf mögliche, rationale Antworten des *Followers* erfüllen.

Die kleinsten Kosten, die der *Leader* erreichen kann sind

$$\inf_{u_L \in \mathcal{U}_L} \inf_{u_F \in \mathcal{U}_F} J_L(u_L, u_F)$$

falls der *Follower* mit dem Leader kooperiert und somit die Kostenfunktion des *Leaders* ebenfalls minimiert.

Solch eine Lösung des Problems nennen wir *team-optimal*. Diese wollen wir nun für unser Beispiel berechnen. Die nötigen Berechnungen finden sich in `stackelberg.mws` auf der beiliegenden CD.

Es ergeben sich folgende team-optimale Strategien:

$$u_L^{Team}(x(2)) = -\frac{1}{2}E(2) - \frac{1}{2}M(2) + \frac{1}{2} \quad (9.2)$$

$$u_F^{Team}(x(1)) = -\frac{1}{3}E(1) - \frac{2}{3}M(1) + \frac{101}{300} \quad (9.3)$$

Die team-optimale Lösung ist damit dann

$$x^{Team}(2) = \begin{pmatrix} E(2) \\ M(2) \end{pmatrix} = \begin{pmatrix} \frac{2}{3}E(1) + \frac{33}{100}M(1) + \frac{101}{300} \\ 1\frac{1}{100}M(1) - \frac{1}{100} \end{pmatrix}$$

$$x^{Team}(3) = \begin{pmatrix} E(3) \\ M(3) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}E(2) + \frac{1}{2}M(2) + \frac{1}{2} \\ 1\frac{1}{100}M(2) - \frac{1}{100} \end{pmatrix}$$

und die minimalen Kosten des *Leaders* sind

$$J_L^{Team} = J_L(u_L^{Team}, u_F^{Team}) =$$

$$0.3333333333E(1)^2 + 1.34E(1)M(1) - 0.6733333333E(1)$$

$$+ 2.387304010M(1)^2 - 1.394408020M(1) + 0.3404373433$$

Für die numerische Berechnung der team-optimalen Lösungen verwenden wir diese Parameter.

$EHat := 1.0$	$MHat := 0$
$em := 1.0$	$lambda := 0.01$
$MStern := 1.0$	
$E(1) := -1.0$	$M(1) := 0.5$

Tabelle 9.1: Parameter für das Beispiel des Stackelberg-Gleichgewichts

Damit ergeben sich folgende team-optimale Lösungstrajektorien

$$x^{Team}(2) = \begin{pmatrix} E(2) \\ M(2) \end{pmatrix} = \begin{pmatrix} -0.165 \\ 0.495 \end{pmatrix}$$

und

$$x^{Team}(3) = \begin{pmatrix} E(3) \\ M(3) \end{pmatrix} = \begin{pmatrix} 0.665 \\ 0.48995 \end{pmatrix}$$

Die Stackelberg-Kosten für den *Leader* betragen dann:

$$J_L^{Team} = 0.5767260025$$

Beachte, dass dieses team-optimale Minimum nicht garantiert werden kann. Der *Follower* will schließlich wie oben beschrieben in erster Linie sein eigenes Kostenfunktional minimieren. Um diese team-optimale Lösung zu einem Stackelberg-Gleichgewicht zu machen, muss es eine optimale Strategie u_L^* des *Leaders* geben, mit der er die oben berechnete team-optimale Strategie von dem *Follower* erzwingen kann. Es muss also eine Beziehung folgender Form gelten:

$$u_F^{Team} = \arg \min_{u_F \in \mathcal{U}_F} J_F(u_L^*, u_F)$$

Der *Leader* wird nun eine Strategie wählen, die (9.2) vorschreibt, wenn der *Follower* (9.3) wählt und ihn ansonsten stark bestraft. Solch ein Strategie könnte dann wie folgt aussehen:

$$u_L^*(x(2), x(1)) = \begin{cases} u_L^{Team} & , \quad \text{falls } x(2) = x^{Team}(2) \\ -kx(1) & , \quad \text{sonst} \end{cases}$$

mit $k \in \mathbb{R}$. Dies führt zu folgender Ungleichung.

$$\inf_{u_F \in \mathcal{U}_F} J_F(u_L = u_L^{Team}, u_F) < \inf_{u_F \in \mathcal{U}_F} J_F(u_L = -kx(1), u_F)$$

KAPITEL 10

Zusammenfassung

Ziel dieser Arbeit war es ein dynamisches Spiel zur CO₂-Reduktion zu analysieren. Dies geschah zum einen für einen Spieler und zum anderen für zwei Spieler mit einer besonderen Informationsstruktur, der *nicht-antizipierenden* Spielsituation. Bei der praktischen Analyse unseres dynamischen Spiels wurden zwei unterschiedliche Verfahren betrachtet, die dynamische Programmierung nach Bellman und ein Greedy-Algorithmus. Bei der Auswertung der Ergebnisse wurde deutlich, dass nur mit Hilfe des Bellman'schen Optimalitätsprinzips jeweils das tatsächliche globale Optimum gefunden wird, während bei dem Greedy-Algorithmus nur lokale Optima erreicht werden können.

Bei unserer Betrachtung fiel auf, dass der Greedy-Algorithmus im Fall von zwei Spielern die geringeren Gesamtkosten liefert. Deshalb wäre es für die zukünftige Untersuchung interessant zu sehen, wie sich die Gesamtkosten verhalten, wenn nur einer der beiden Spieler greedy spielt und der andere die dynamische Programmierung wählt.

Da in dieser Arbeit nur ein, oder zwei Spieler an unseren dynamischen Spielen beteiligt waren, wurde die Möglichkeit der Koalitionsbildung mehrerer Spieler außer acht gelassen. Da unser CO₂-Reduktionsmodell aber auch für mehr als zwei Spieler definiert ist, ist ein Zusammenschluss unter den einzelnen Spielern möglich. Dies kann wie beschrieben durch Veränderung der *em*-Matrix geschehen. Betrachtungen diesbezüglich finden sich in [8] und [14].

ANHANG A

Inhalt der beiliegenden CD

- **Diplomarbeit:**
enthält die Diplomarbeit als PDF-Datei
- **Maple-Programme:**
enthält die Quellcodes der verwendeten Maple Worksheets:
 - `dynprog1spieler.mws`
 - `dynprog2spieler.mws`
 - `greedy1spieler.mws`
 - `greedy2spieler.mws`
 - `stackelberg.mws`
 - `vergleichgreedydynprog.mws`
- **Literatur:**
enthält die verwendete Literatur, die in elektronischer Form vorliegt als PDF-Dateien

Literaturverzeichnis

- [1] M. Bardi, I. Capuzzo-Dolcetta; *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkäuser Verlag Boston Basel Berlin, 1997.
- [2] T. Basar; G. J. Olsder *Dynamic Noncooperative Game Theory*, SIAM Society for Industrial and Applied Mathematics, 1999
- [3] R. Bellman; *Dynamic Programming*, Princeton University Press Princeton, 1957;
- [4] L.Grüne; *Numerik Dynamischer Systeme*, Vorlesungsskript, [http : //www.math.uni – bayreuth.de/ ~ lgruene/numdyn0506/](http://www.math.uni-bayreuth.de/~lgruene/numdyn0506/) .
- [5] L. Grüne; *Numerische Dynamik von Kontrollsystemen*, Vorlesungsskript, [http : //www.math.uni – bayreuth.de/ ~ lgruene/ndks04/](http://www.math.uni-bayreuth.de/~lgruene/ndks04/) .
- [6] T. Hagerup; *Datenstrukturen und Algorithmen*, Max-Planck-Institut für Informatik Saarbrücken, 1997;
- [7] Y.C. Ho, A.W. Starr; *Journal of Optimization Theory and Applications 1969: Nonzero-Sum Differential Games*, Plenum Publishing Corporation, 1969.
- [8] W. Krabs; *Spieltheorie Dynamische Behandlung von Spielen*; B. G. Teubner Verlag/GWV Fachverlage GmbH Wiesbaden, 2005.
- [9] W. Kaup; *Funktionentheorie*, Vorlesungsskript, [http : //www.mathematik.uni – tuebingen.de/ ~ kaup/futh.pdf](http://www.mathematik.uni-tuebingen.de/~kaup/futh.pdf)
- [10] O. Morgenstern, J. v. Neumann; *Spieltheorie und wirtschaftliches Verhalten*; Physika-Verlag Würzburg, 1973.

- [11] B. Oswick-Klewe; *Vorlesung zum Informatik-und Ingenieurstudium Band 5 Algorithmen und Datenstrukturen*, Dr. Bernd Wißner Augsburg, 1995;
- [12] M. Papageorgiou; *Optimimierung - Statische, dynamische, stochastische Verfahren für die Anwendung*, R. Oldenburg Verlag München Wien, 1996.
- [13] M. Parlar *Interactive Operations Research with Maple*, Birkhäuser Verlag Boston Basel Berlin, 2000.
- [14] W. Pickl; *Der τ -value als Kontrollparameter Modellierung und Analyse eines Joint-Implementation Programms mithilfe der dynamischen kooperativen Spieltheorie und der diskreten Optimierung*, Shaker Verlag Aachen, 1999.
- [15] B. Rauhut, N. Schmitz, E.-W. Zachow; *Spieltheorie*, B. G. Teubner Stuttgart, 1979.
- [16] M. Sniedovich; *Dynamic Programming*, Marcel Dekker, Inc. New York Basel Hong Kong, 1992.
- [17] P. Stahlecker, N. Hauenschild, M. Klintworth; *Optimierung und ökonomische Analyse*, Springer Berlin, 2003.
- [18] W.-H. Steeb, A. Kunick; *Chaos in dynamischen Systemen*, Wissenschaftsverlag Mannheim Wien Zürich, 1989.
- [19] Definition von beschränktem Wachstum:
<http://delphi.zsg-rottenburg.de/faqmath1.html>.
- [20] Kyotoprotokoll:
<http://un.fccc.int/resource/docs/convkp/kpger.pdf>.
- [21] TEM-Modell:
<http://www.zaik.uni-koeln.de/AFS/Projects/MultiMedia/TEMPPI/modell/mappe/mappe.pdf>.
- [22] Umweltdilemma:
http://www.wiwimaster.de/data/Dokumente/14/031124155718_Vorlesung_00-01.pdf.
- [23] Zitat von Oliver Wendell Holmes:
<http://www.taz.de/1/archiv/dossiers/dossier-buchmesse/artikel/1/denken-und-joggen/>.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Bayreuth, den 17. Februar 2009

.....
(Nadine Löhner)