

**Ein dual-basierter Ansatz zur Behandlung
linear-quadratischer
Optimalsteuerungsprobleme mit einem
a priori gegebenen unendlichen Zeithorizont**

von der Fakultät 1 - MINT - Mathematik, Informatik, Physik,
Elektro- und Informationstechnik
der Brandenburgischen Technischen Universität Cottbus-Senftenberg
genehmigte Dissertation zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von

Master of Science Katharina Kolo (geb. Heym)
geboren am 31.07.1986 in Dresden, Deutschland

Vorsitzender: Prof. Dr. Gerd Wachsmuth
Gutachterin: Prof. Dr. Sabine Pickenhain
Gutachter: Prof. Dr. Karl Worthmann
Gutachter: Prof. Dr. Christopher Schneider

Tag der mündlichen Prüfung: 07.04.2021

Zusammenfassung

Schwerpunkt der vorliegenden Arbeit ist die dual-basierte Behandlung linear-quadratischer Optimalsteuerungsprobleme mit einem a priori gegebenen unendlichen Zeithorizont unter der Berücksichtigung von Steuerungsbeschränkungen. Neben der theoretischen und analytischen Auseinandersetzung mit derartigen Aufgabentypen enthält die Arbeit zudem eine ausführliche Darstellung numerischer Experimente anhand von Anwendungsbeispielen aus der Biologie und Epidemiologie. Dafür betrachten wir neben dem Lotka-Volterra System auch eine Formulierung des SEIR Modells, welches die Ausbreitung einer Infektionserkrankung beschreibt. Für die Beschreibung der in das Modell eingehenden Parameter orientieren wir uns an Kennzahlen der Viruserkrankung Covid-19. Linear-quadratische Steuerungsaufgaben stellen eine bedeutende Aufgabenklasse der Optimalsteuerung dar, da sie sowohl in der Approximation nichtlinearer Aufgaben, als auch im Zuge der sequentiellen quadratischen Optimierung zum Einsatz kommen. Die Wahl eines a priori gegebenen unendlichen Zeithorizontes entsteht beispielsweise im Kontext der Quantenmechanik und der Theorie der Stabilisierung dynamischer Systeme. Motiviert durch derartige Anwendungen wird diese Aufgabenklasse genauer studiert. Die Optimalsteuerungsaufgabe wird dafür über gewichteten Funktionenräumen in Form gewichteter Sobolev- und Lebesgue-Räume formuliert und im Kontext der konvexen Hilbertraumtheorie studiert. Dabei differenzieren wir in unseren Ausführungen zwischen Problemen über Funktionenräumen mit echten Gewichtsfunktionen und Funktionenräumen mit Dichtefunktionen. Wir diskutieren zunächst die Existenz und Eindeutigkeit optimaler Lösungen der zugrundeliegenden Problemklasse und geben Optimalitätsbedingungen für Probleme über Funktionenräumen mit Dichtefunktionen an. Die noch ausstehende Formulierung notwendiger Optimalitätsbedingungen für Probleme über Funktionenräumen mit echter Gewichtsfunktion motiviert schließlich einen dual-basierten Zugang. Mit den Ideen der Lagrange-Dualität wird hierfür eine duale Aufgabe in Form eines Variationsproblems in gewichteten Funktionenräumen mit Dichtefunktionen hergeleitet. Für diese können schließlich Sattelpunktsbedingungen und notwendige Optimalitätsbedingungen in Form eines Pontryaginschen Maximumprinzips bewiesen werden. Dieses liefert neben den kanonischen Gleichungen und der Maximumbedingung auch die Existenz einer Adjungierten, welche einer natürlichen Transversalitätsbedingung genügt. Für die numerische Behandlung der so konstruierten dualen Aufgabe wird eine direkte Pseudospektralmethode vorgestellt. Diese basiert auf der Verwendung verallgemeinerter Laguerre-Polynome, welche ein vollständiges Orthogonalsystem auf dem durch die Dichtefunktion gewichteten Lebesgue-Raum bildet. Der erfolgreiche Einsatz dieser Methode wird anhand verschiedener Anwendungsbeispiele aus dem Bereich der Optimalen Steuerung und der asymptotischen Steuerbarkeit dynamischer Systeme veranschaulicht.

Abstract

The present work focuses on a dual-based treatment of linear-quadratic optimal control problems with an a priori given infinite time horizon and control constraints. In addition to the theoretical and analytical investigation of such types of problems, this work also includes a detailed presentation of numerical experiments based on applications from biology and epidemiology. For this purpose, in addition to the Lotka-Volterra system, we also consider a version of the SEIR model, which describes the spread of an infectious disease. Thereby the parameter sets involved in the model are based on the current statistical data of the viral disease Covid-19. Linear-quadratic control problems represent an important class in the theory of optimal control, since they are used in the approximation of non-linear problems, as well as in the field of sequential quadratic optimization. For example, the choice of an a priori given infinite time horizon arises in the context of quantum mechanics and the theory of stabilization of dynamical systems. Motivated by this wide range of applications this class of problems is considered in detail. The problem formulation over an infinite time horizon implies a weighted space setting. For the state space a weighted Sobolev space is chosen and for the control space a weighted Lebesgue space is considered. Therefore we distinguish between problems in functional spaces with proper weight functions and functional spaces with density functions. We discuss the existence and the uniqueness of optimal solutions of the underlying problem formulation and specify optimality conditions for problems in functional spaces with density functions. The missing formulation of necessary optimality conditions for problems in functional spaces with a proper weight function finally motivates a dual-based approach. Following this approach we present a construction of a dual problem based on the ideas of the Lagrange duality. This leads us to a dual program in the form of a variational problem in weighted functional spaces with density functions. We succeeded in providing saddle point conditions and necessary optimality conditions in form of a Pontryagin maximum principle. These necessary conditions provide in addition to the canonical equation and the maximum condition the existence of an adjoint satisfying a natural transversality condition. For the numerical treatment of the dual program we present a direct pseudospectral method based on the usage of generalized Laguerre polynomials which build a complete orthogonal set in the weighted Lebesgue space. The successful application of the proposed method is illustrated by means of various numerical examples from optimal control theory as well as from asymptotic controllability field.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen meinen Dank all denen entgegen zu bringen, welche mich in der Anfertigung meiner Dissertation unterstützten.

Ganz besonders danken möchte ich dabei meiner Doktormutter Frau Prof. Dr. Sabine Pickenhain für die langjährige und hervorragende Betreuung sowie die zahlreichen fachlichen Diskussionen, die zum Entstehen der Arbeit maßgeblich beigetragen haben.

Weiterhin danken möchte ich Herrn Prof. Dr. Karl Worthmann und Herrn Prof. Dr. Christopher Schneider sowohl für die Begutachtung meiner Arbeit als auch die zahlreichen interessanten Diskussionen im Rahmen verschiedenster Fachkonferenzen.

Außerdem danke ich meinem Kollegium vom Lehrstuhl Mathematik insbesondere Optimierung, vertreten durch Valeriya Lykina, Diana Wenzke, Angie Burtchen, Torsten Ziemann, Illia Dikariev, Inna Spivak und Henriette Dräger für das höchst angenehme Arbeitsklima, den wertvollen Austausch und die moralische Unterstützung.

Nicht zuletzt sei an dieser Stelle auch ein Dank an die Deutsche Forschungsgemeinschaft (DFG) gerichtet, deren finanzielle Förderung im Rahmen des Projektes „Optimalsteuerungsprobleme mit unendlichem Zeithorizont und ihre Anwendungen in der Epidemiologie: Dualität und hinreichende Optimalitätsbedingungen, dual-basierte Pseudospektralmethoden“ den Weg zur Erstellung der vorliegenden Arbeit ebnete.

Doch auch im privaten Bereich erfuhr ich zahlreiche Unterstützung für welche ich äußerst dankbar bin. Dabei seien zuallererst meine lieben Freunde genannt, welche mit größter Sorgfalt das Korrekturlesen meiner Arbeit übernahmen. In ganz besonderem Maße möchte ich mich jedoch bei meiner Familie – allen voran meinem lieben Partner und meinen Töchtern Pia und Marie – bedanken, welche mir insbesondere in den letzten Jahren den Rücken stärkten, stets an mich glaubten, stundenlangen mathematischen Monologen meinerseits ausgesetzt waren, mir großes Verständnis entgegen brachten, sich dabei sehr häufig zurück nahmen und mir damit den Raum schafften diese Phase meines Lebens erfolgreich und glücklich zu beschreiten.

Abschließend möchte ich auch all jenen danken, die mich in diesem Abschnitt meines Lebens begleitet und geprägt haben, die ich hier jedoch nicht vollständig erwähnen konnte. Danke.

Katharina Kolo

Inhaltsverzeichnis

Einleitung	1
1 Grundlegende Begriffe und Ergebnisse	7
1.1 Schwache Topologie	8
1.2 Klassische Lebesgue- und Sobolev-Räume	9
1.3 Gewichtete Lebesgue- und Sobolev-Räume	12
1.4 Differentialgleichungen in Sobolev-Räumen	15
1.5 Einführung wichtiger Differentiationsbegriffe	17
1.6 Asymptotische Steuerbarkeit dynamischer Systeme	19
2 Linear-quadratische Optimalsteuerungsprobleme mit unendlichem Zeit-	
 horizont	23
2.1 Formulierung der Aufgabe (P_ν)	23
2.2 Voraussetzungen	25
2.3 Formulierung als Optimierungsproblem im Hilbertraum	26
2.4 Existenz und Eindeutigkeit einer optimalen Lösung	30
2.5 Notwendige Optimalitätsbedingungen	40
3 Dualitätstheorie für linear-quadratische Optimalsteuerungsprobleme	45
3.1 Ein kurzer Einblick in die Dualitätstheorie	46
3.2 Anwendung des Konstruktionsschemas für Probleme der Klasse (P_ν)	48
3.3 Berechnung der dualen Aufgabe für (P_{ν_0})	52
3.3.1 Formulierung der untergeordneten Hilfsprobleme	52
3.3.2 Existenz und Eindeutigkeit einer Lösung der Hilfsprobleme	53
3.3.3 Optimalitätsbedingungen der Hilfsprobleme	59
3.4 Berechnung der dualen Aufgabe für (P_{ν_1})	62
3.4.1 Formulierung der untergeordneten Hilfsprobleme	62
3.4.2 Existenz und Eindeutigkeit einer Lösung der Hilfsprobleme	64
3.4.3 Optimalitätsbedingungen der Hilfsprobleme	65
3.5 Formulierung eines schwachen Dualitätssatzes	68
4 Die duale Aufgabe (D_ν)	69
4.1 Zusammenführung der dualen Aufgaben	70
4.2 Formulierung als Optimalsteuerungsproblem	73
4.3 Voraussetzungen	76
4.4 Formulierung als Optimierungsproblem im Hilbertraum	77
4.5 Notwendige Optimalitätsbedingungen	87
4.6 Formulierung eines starken Dualitätssatzes	100
4.7 Formulierung von Optimalitätsbedingungen für (P_{ν_i})	103

5 Numerische Behandlung	105
5.1 Vorstellung der Pseudospektralmethode	105
5.2 Anwendung einer direkten Pseudospektralmethode	110
5.2.1 Diskretisierung der Aufgabe	112
5.2.2 Kollokationsgleichungen	113
6 Beispiele	117
6.1 Das Beispiel von Garg, Hager und Rao	117
6.1.1 Einführung von Steuerungsbeschränkungen	118
6.1.2 Untersuchung der zulässigen Menge	119
6.1.3 Übergang zur dualen Aufgabe	120
6.1.4 Optimale Lösung	122
6.1.5 Anwendung der direkten Pseudospektralmethode	127
6.1.6 Auswertung der numerischen Ergebnisse	129
6.2 Das Lotka-Volterra Modell	139
6.2.1 Stabilitätseigenschaften des Gleichgewichtspunktes	140
6.2.2 Formulierung der primalen Aufgabe	141
6.2.3 Untersuchung der zulässigen Menge	142
6.2.4 Übergang zur dualen Aufgabe	144
6.2.5 Anwendung der direkten Pseudospektralmethode	147
6.2.6 Optimale Lösung	150
6.2.7 Auswertung der numerischen Ergebnisse	150
6.3 Das SEIR Modell	158
6.3.1 Formulierung der primalen Aufgabe	161
6.3.2 Steuerbarkeit des Systems	162
6.3.3 Übergang zur dualen Aufgabe	164
6.3.4 Anwendung der direkten Pseudospektralmethode	166
6.3.5 Auswertung der numerischen Ergebnisse	168
Abschließende Bemerkungen	181
A Wichtige Sätze und Definitionen	185
B Existenz und Eindeutigkeit der Hilfsprobleme	187
B.1 Existenz einer Lösung von (HP2)	187
B.2 Existenz einer Lösung von (HP3)	190
B.3 Existenz einer Lösung von (HP4)	194
C Quellcode für das Beispiel aus Abschnitt 6.1	199
C.1 LQR.m	199
C.2 Kollokation.m	200
C.3 LaguerreGenBeta.m	203

Abbildungsverzeichnis

6.1	Ansatz für die Schaltfunktion σ bei Box-Restriktionen $[-\bar{u}, \bar{u}]$ mit $\bar{u} > 0$ und $x_0 > -\bar{u}(1 - \sqrt{2})$ (entnommen aus [Bur19], Abbildung 6.17, S. 195)	123
6.2	Näherungslösung $x_N(t)$ des primalen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $x_*(t)$ (—)	131
6.3	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—)	131
6.4	Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—)	132
6.5	Näherungslösung $x_N(t)e^{-t}$ des primalen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $x_*(t)e^{-t}$ (—)	133
6.6	Näherungslösung $u_N(t)e^{-t}$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)e^{-t}$ (—)	133
6.7	Näherungslösung $y_N(t)e^{-t}$ des dualen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $y_*(t)e^{-t}$ (—)	134
6.8	Näherungslösung $x_N(t)$ des primalen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $x_*(t)$ (—)	136
6.9	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—)	136
6.10	Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—)	137
6.11	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—) auf $[0, 20]$	138
6.12	Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—) auf $[0, 20]$	138
6.13	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—) in der Nähe des Umschaltzeitpunktes \bar{t}	139
6.14	Mögliche Wahl der Anfangswerte im Falle normierter Steuerungsbereiche für $\rho_1, \rho_2 \in [0, 10]$	144
6.15	Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $y^*(t)$ (—)	152
6.16	Näherungslösung $p_N(t)$ der dualen Adjungierten für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $p^*(t)$ (—)	152

6.17	Näherungslösung $x_N(t)$ des rücktransformierten primalen Zustandes für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $x^*(t)$ (—)	153
6.18	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $u^*(t)$ (—)	153
6.19	Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)	154
6.20	Näherungslösung $p_N(t)$ der dualen Adjungierten für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)	154
6.21	Näherungslösung $x_N(t)$ des rücktransformierten primalen Zustandes für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)	155
6.22	Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)	155
6.23	Numerisch ermittelte Lösungen der Differentialgleichung (6.32) auf Basis der approximierten Schaltzeitpunkte $\bar{t}_N^{(1)}$ und $\bar{t}_N^{(2)}$. Dabei bezeichnet I die Funktion $y_1(t)$, II die Funktion $y_2(t)$, III die Funktion $y_1'(t)$ und IV die Funktion $y_2'(t)$	157
6.24	Das SEIR Modell	159
6.25	Näherungslösung der dualen Zustände $y_1(t)$ und $y_2(t)$ für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	171
6.26	Näherungslösung der dualen Zustände $y_3(t)$ und $y_4(t)$ für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	171
6.27	Näherungslösung der Komponenten S_N und E_N für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	172
6.28	Näherungslösung der Komponenten I_N und N_N für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	172
6.29	Näherungslösung der Steuerungen $u_1(t)$ und $u_2(t)$ für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	173
6.30	Näherungslösung der Steuerungen $u_3(t)$ und $u_4(t)$ für $\max_{i=1,\dots,4} x_{0i} - x_{zi} \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)	173
6.31	Näherungslösung der dualen Zustände $y_1(t)$ und $y_2(t)$ für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	175
6.32	Näherungslösung der dualen Zustände $y_3(t)$ und $y_4(t)$ für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	175
6.33	Näherungslösung der Komponenten S_N und E_N für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	176
6.34	Näherungslösung der Komponenten I_N und N_N für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	176
6.35	Näherungslösung der primalen Steuerungen $u_1(t)$ und $u_2(t)$ für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	177
6.36	Näherungslösung der primalen Steuerungen $u_3(t)$ und $u_4(t)$ für den Fall $\max_{i=1,\dots,4} x_{0i} - x_{zi} = 69$ für $N = 10$ (—) und $N = 11$ (—)	177

Tabellenverzeichnis

6.1	Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$	130
6.2	$L_{2,e^{-2t}}$ -Fehler des primalen Zustandes, des dualen Zustandes und der Funktion θ für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$	134
6.3	Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y^*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$	135
6.4	Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = \left[-\frac{11}{5}, \frac{11}{5}\right]$	136
6.5	Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y_*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = \left[-\frac{11}{5}, \frac{11}{5}\right]$	137
6.6	Vergleich des numerisch ermittelten Umschaltzeitpunktes \bar{t}_N zur exakten Lösung \bar{t} für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = \left[-\frac{11}{5}, \frac{11}{5}\right]$	139
6.7	Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$ und $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1 + \sqrt{2})$	151
6.8	Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y^*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$ und $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1 + \sqrt{2})$	153
6.9	Numerisch ermittelte Zielfunktionswerte, Anfangswerte und Schaltzeitpunkte für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	155
6.10	Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	157
6.11	Numerisch ermittelte Anfangswerte der dualen Zustände für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	171
6.12	Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	174
6.13	Numerisch ermittelte Anfangswerte der dualen Zustände für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	175
6.14	Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$	177

Einleitung

Was immer in der Welt passiert, in seinem Inneren hat es die Bedeutung von Maximum oder Minimum. Somit ist kein Zweifel, dass alle Naturphänomene über die Methode des Maximierens oder Minimierens erklärt werden können.

Leonhard Euler¹

Sowohl im Alltag, als auch in der Natur werden wir immer wieder mit den Mechanismen der Optimierung konfrontiert. Insbesondere in den Naturwissenschaften – wie der Physik oder der Chemie – findet sich häufig das „Prinzip der kleinsten Wirkung“ oder das „Prinzip der minimalen Energie“. Daher ist es wenig überraschend, dass die Optimierung als ein natürliches Denkschema erscheint und zahlreiche Anwendungsfelder offenbart. Ganz allgemein lässt sich die Optimierung in die Bereiche der endlich dimensionalen und der unendlich dimensionalen Optimierung unterteilen. Zu erster Kategorie zählen wir Probleme, deren Optimierungsvariablen Elemente eines endlich dimensionalen Raumes sind und deren der Aufgabe zugrundeliegenden Nebenbedingungen durch endlich viele Restriktionen beschrieben werden können. Dem gegenüber stehen die Probleme der unendlich dimensionalen Optimierung. Zu dieser Klasse zählen wir Optimierungsprobleme, in welchen die Optimierungsvariablen durch Elemente aus Funktionenräumen gegeben sind und/oder die Aufgaben unendlich vielen Restriktionen unterlegen sind. Ein vergleichsweise junges Teilgebiet der unendlich dimensionalen Optimierung ist die im Rahmen dieser Arbeit behandelte Optimalsteuerung. Aus den Bedürfnissen der Technik heraus entwickelte sich diese aus der im 17. Jahrhundert entstandenen Variationsrechnung. Den wesentlichen Impuls lieferten dabei Anwendungen aus dem Bereich der Luft- und Raumfahrt in den fünfziger und sechziger Jahren des vergangenen Jahrhunderts². Die zugrundeliegenden technischen Anwendungen erforderten nun die Betrachtung von Aufgaben, in denen die Zustands- und Steuerungsgrößen Beschränkungen unterworfen sind und motivierten somit die Erweiterung der durch die Variationsrechnung behandelbaren Klassen. Nachdem zunächst einzelne Problemstellungen gelöst waren, entwickelte sich recht zügig das Bedürfnis nach einer Formulierung von allgemeineren Lösungsmethoden, so dass bereits in den fünfziger

¹Dieses Zitat findet sich in Eulers Schrift „*Methodus inveniendi lineas curvas maximi minime proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti.*“ (Eine Methode, um Kurven zu finden, denen eine Eigenschaft im höchsten oder geringsten Grade zukommt oder die Lösung des Isoperimetrischen Problems, wenn es im weitesten Sinne des Wortes aufgefasst wird.). Sie stammt aus dem Jahre 1744 und gilt als erste systematische Behandlung der Variationsrechnung.

²Ein frühes Beispiel aus der Luft- und Raumfahrt ist das bekannte Problem von GODDARD aus dem Jahre 1919. In seiner Veröffentlichung „*Methods for Reaching Extreme Altitudes*“ (Methoden zum Erreichen extremer Höhen) schlussfolgerte er, dass Raketen genutzt werden könnten um Nutzlasten auf den Mond zu schicken. Tatsächlich wurde das von ihm formulierte „*Goddard-Problem*“ noch mit Methoden der Variationsrechnung gelöst.

und sechziger Jahren des zwanzigsten Jahrhunderts wegweisende Ergebnisse im Bereich der Optimalen Steuerung aus den Arbeitsgruppen um PONTRYAGIN, HESTENS, BELLMAN und ISAACS vorlagen. Der Fokus dieser ersten grundlegenden Ausarbeitungen lag dabei auf der Betrachtung von Problemen über einem endlichen Zeithorizont $[0, T]$ mit $T > 0$ (derartige Probleme nennen wir auch *klassische Steuerungsprobleme*). In den darauf folgenden Jahrzehnten wurden diese Ergebnisse aufgegriffen und für eine breite Klasse an Aufgabenstellungen studiert, so dass theoretische wie algorithmische Fragen für Probleme der optimalen Steuerung über einem endlichen Zeithorizont mittlerweile weitgehend geklärt sind.

Ziel dieser Arbeit ist daher nicht die Behandlung von Optimalsteuerungsproblemen über einem endlichen Zeithorizont, sondern die analytische und numerische Auseinandersetzung mit Optimalsteuerungsproblemen über dem unendlichen Zeithorizont.

Eine frühe Beschreibung eines derartigen Modells lieferte RAMSEY im Jahre 1928 in einer Anwendung aus den Wirtschaftswissenschaften. In [Ram28] motivierte dieser – basierend auf einer geschlossenen Marktwirtschaft, d.h. auf einem Markt ohne Im- und Export – ein sogenanntes Kapitalakkumulationsmodell. Nach den Annahmen der geschlossenen Marktwirtschaft wird das erwirtschaftete Einkommen Y entweder konsumiert (beschrieben durch die Konsumrate $C(t)$) oder investiert (beschrieben durch die Bruttoinvestitionen $I(t)$), so dass zu jedem Zeitpunkt t gilt $Y(t) = C(t) + I(t)$. In Abhängigkeit von der Höhe der getätigten Investitionen entwickelt sich der Kapitalbestand K der Marktwirtschaft, welcher beispielsweise für Innovationsentwicklungen oder für die Steigerung der Arbeitskraft Einsatz findet, und in diesem Rahmen Einfluss auf die Produktivität des Marktes nimmt. RAMSEY motiviert auf dieser Grundlage die Frage, durch welche Investitions- und Konsumptionsstrategie der Nutzen der Marktteilnehmer maximiert werden kann. Dabei ist davon auszugehen, dass sich der Nutzen über den getätigten Konsum misst. Eine der vielen Verallgemeinerungen des von RAMSEY formulierten Modells findet sich in [FH86] und lässt sich wie folgt beschreiben:

Maximiere

$$\int_0^{\infty} U(c(t)) e^{-\beta t} dt \quad (1)$$

unter den Nebenbedingungen

$$\begin{aligned} k'(t) &= f(k(t)) - \delta k(t) - c(t), \\ k(0) &= k_0, \\ 0 &\leq c(t) \leq f(k(t)). \end{aligned}$$

Die Beschreibung des Marktes geschieht hier auf der Grundlage einer vorgegebenen Nutzenfunktion $U(\cdot)$ und der Pro-Kopf-Produktion $f(\cdot)$ zu einem Pro-Kopf-Kapitalstock $k(\cdot)$ mit dem vorgegebenen Anfangsbestand k_0 unter Berücksichtigung der Abschreibungsrate $\delta > 0$. Die Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ beschreibt³ dabei die Pro-Kopf-Produktion in Abhängigkeit vom Kapital pro Arbeiter für jeden Zeitpunkt t und ist – wie bereits beschrieben – indirekt abhängig von der Entscheidung zum Konsumieren und Investieren. Wir haben

³Mit \mathbb{R}_+ bezeichnen wir das Intervall $[0, \infty)$.

daher Einfluss auf den beschriebenen Markt, indem wir zu jedem Zeitpunkt t für das erwirtschaftete Pro-Kopf-Einkommen eine Entscheidung zum Investieren oder Konsumieren (modelliert durch den Pro-Kopf-Konsum $c : \mathbb{R}_+ \rightarrow \mathbb{R}$) treffen. Die Ungleichungsnebenbedingungen sichern hierbei die Nicht-Negativität von Konsum und Produktion. Die Modellierung des Zielfunktional durch ein Integral über dem unbeschränkten Intervall $[0, \infty)$ erlaubt nun die Bewertung des Nutzens unter Berücksichtigung des Diskontierungsfaktors $e^{-\beta t}$ über den gesamten Planungshorizont. Die Einführung des unendlichen Zeithorizontes gestattete RAMSEY eine Modellierung des Marktes unter Berücksichtigung der Nachfolgenerationen. In ihrem Buch *Public Investment, the Rate of Return, and Optimal Fiscal Policy* merken ARROW und KURZ zu dieser Idee an⁴:

„The infinite horizon is an idealization of the fundamental point that the consequences of investment are very long-lived; any short horizon requires some methods of evaluating end-of-period capital stocks, and the only proper evaluation is their value in use in the subsequent future“

Die systematische Wahl eines unendlichen Planungsintervalls für die Berücksichtigung von Nachfolgenerationen setzte sich nachfolgend in einer breiten Klasse an Anwendungsfeldern durch. Exemplarisch dafür verweisen wir auf das Buch von FEICHTINGER und HARTL [FH86]⁵. In diesem findet sich ein breites Spektrum ökonomischer Anwendungen auf unendlichem Zeitintervall, wie beispielsweise aus dem Bereich der Produktions- und Lagerhaltung, dem Ressourcenmanagement oder auch dem Umweltschutz. Derartige ökonomische Anwendungen motivieren zudem häufig die Verwendung sogenannter *Dichtefunktionen* in Form eines Diskontierungsfaktors $e^{-\beta t}$ ($\beta > 0$).

Wie in der Beschreibung des Ramsey-Modells wird bei der Formulierung derartiger Steuerungsprobleme in der Regel ein Integralfunktional minimiert oder maximiert (vgl. (I)). Dabei zeigt sich, dass die Auswertung des Integrals in den zitierten Arbeiten häufig im Sinne eines uneigentlichen Riemann- oder Lebesgue-Integrals, d.h. durch die Betrachtung des Grenzwertüberganges

$$R/L - \int_0^{\infty} f(t)dt = \lim_{T \rightarrow \infty} R/L - \int_0^T f(t)dt \quad (2)$$

geschieht. Einige Anwendungen erlauben hierbei die Divergenz des Zielfunktional. Es gilt dann, potentielle Lösungskandidaten mit möglicherweise unendlich großem, beziehungsweise divergentem Zielfunktionalwert zu vergleichen und unter diesen eine optimale Lösung auszuwählen. Dieser Umstand erforderte die Einführung unterschiedlicher Optimalitätsbegriffe, wie beispielsweise „*overtaking optimality*“ (im Sinne von WEIZSÄCKER) oder auch „*catching up optimality*“ (im Sinne von GALE)⁶. Dabei ist leicht einzusehen, dass die Formulierung von Existenzsätzen und Optimalitätsbedingungen wesentlich von dem zugrundeliegenden Optimalitätsbegriff und dem verwendeten Integraltyp abhängt. Zudem besitzen die unterschiedlichen Integraltypen unterschiedliche Konvergenzeigenschaften. Daher ist im Rahmen einer analytischen Behandlung eine konsistente Interpretation von Inte-

⁴siehe dazu [AK70], S. xviii

⁵Weitere Anwendungsbeispiele über dem unendlichen Zeithorizont finden sich beispielsweise auch in [AK04, Arr68, AK70, Sol70].

⁶Die Wahl eines geeigneten Optimalitätsbegriffes ist dabei auch auf Grundlage der Anwendung zu tätigen. Eine Zusammenfassung verschiedenster Optimalitätsbegriffe befindet sich beispielsweise in [FH86], S. 186, Definition 7.2.

graltyp und Optimalitätsbegriff unabdingbar⁷.

Betrachten wir ökonomische Modelle, wie das obig eingeführte Ramsey-Modell, so erscheint eine Interpretation des Zeithorizontes als Grenzwertprozess klassischer Steuerungsprobleme mit einem beliebig großen endlichen Zeitintervall $[0, T]$ (mit $T > 0$), und somit eine Verwendung des uneigentlichen Riemann-Integrals, adäquat. Aufgabenklassen dieser Art stehen jedoch jenen Steuerungsproblemen gegenüber, in denen der unendliche Zeithorizont a priori in der Aufgabenstellung auftritt. Derartige Probleme finden wir insbesondere in technischen und physikalischen Anwendungen – wie beispielsweise dem harmonischen Oszillator in der Quantenmechanik⁸, in welchem die Verwendung des unendlichen Zeithorizontes auf natürliche Weise aus dem physikalischen Kontext der Aufgabe selbst entsteht – sowie in einer Vielzahl biologischer Modelle. Die Einführung des unendlichen Zeithorizontes wird bei letzteren häufig durch die Interpretation der Lebenszeit in Form einer exponentiell verteilten Zufallsgröße motiviert⁹. Ein weiteres Anwendungsbeispiel liefert die Stabilitäts- und Steuerbarkeitsanalyse. Aufbauend auf den Ausarbeitungen von KALMAN¹⁰, in denen er die Zusammenhänge zwischen der asymptotischen Steuerbarkeit eines dynamischen Systems und Aufgaben der optimalen Steuerung diskutierte, beziehen wir uns dafür auf eine in [PB18] präsentierte Philosophie, welche durch eine geschickte Wahl von gewichteten Funktionenräumen die exponentielle (und damit auch die asymptotische) Stabilität dynamischer Systeme im Sinne von LYAPUNOV garantiert. Anders als in den zuvor besprochenen Anwendungen liefert die in [PB18] präsentierte Herleitung jedoch eine Klasse von Problemen mit unendlichem Zeithorizont unter Verwendung einer *echten Gewichtsfunktion* $e^{\beta t}$ ($\beta > 0$). Ein uneigentliches Integral, wie es im zuvor beschriebenen Zugang verwendet wurde, ist für derartige Anwendungen nicht mehr passend. Anstelle dessen wird die Verwendung des Lebesgue-Integrals im Sinne von [LPW08] vorgeschlagen. Die derartige Interpretation sichert nun die Anwendbarkeit einer größeren Klasse an Konvergenzsätzen und ermöglicht zudem eine Formulierung der Optimierungsaufgabe über Banachräumen, was die analytische Behandlung der Aufgabe erheblich erleichtert.

In der vorliegenden Arbeit beschränken wir uns auf die Untersuchung von Optimalsteuerungsproblemen mit a priori gegebenem unendlichem Zeithorizont. Dabei beziehen wir uns explizit auf eine Formulierung der vorliegenden Problemklasse über gewichteten Sobolev- und Lebesgue-Räumen und bauen damit auf bereits bestehenden Ergebnissen aus [Lyk10, Zie19] auf. Den Autoren der zitierten Arbeiten gelang unter Ausnutzung der Raumeigenschaften die Formulierung von Existenzsätzen sowie der Nachweis von notwendigen und hinreichenden Optimalitätsbedingungen. Die Ergebnisse in den zitierten Arbeiten beziehen sich dabei auf Aufgabenklassen, die durch die Verwendung von Dichtefunktionen im Zielfunktional und in der Gewichtung der Räume geprägt werden. Hervorzuheben ist hierbei, dass die in [Lyk10, Zie19] formulierten Existenzaussagen und Optimalitätsbedingungen in den gleichen Funktionenräumen formuliert werden konnten und somit eine konsistente Untersuchung der zugrundeliegenden Problemklasse ermöglichen. Aufbauend auf diesen Ergebnissen soll die betrachtete Klasse nun auf Probleme erweitert werden, welche die Verwendung allgemeiner Gewichtsfunktionen der Form $e^{\beta t}$ ($\beta \neq 0$) erlauben. Die derartige

⁷Eine ausführliche Darlegung dieser Zusammenhänge findet sich beispielsweise in [Lyk10, LPW08].

⁸vgl. beispielsweise [PB19], S. 618 ff.

⁹Vgl. hierzu beispielsweise [LML10, Pic10b]. Die Interpretationen in den zitierten Arbeiten führt wiederum zu einer Verwendung von Dichtefunktionen der Form $e^{-\beta t}$ ($\beta > 0$) im Zielfunktional.

¹⁰vgl. beispielsweise [Kal59, Kal60]

Erweiterung der betrachteten Klasse wird uns damit den Zugang zur Stabilisierungs- und Steuerbarkeitstheorie im Kontext der Optimalen Steuerung ermöglichen. Dabei beschränken wir uns jedoch auf die Betrachtung linear-quadratischer Optimalsteuerungsprobleme. Der Bedarf einer Betrachtung dieser eingeschränkten Problemklasse lässt sich legitimieren, da sie sowohl häufig in der Approximation nichtlinearer Aufgaben, als auch im Zuge der sequentiellen quadratischen Optimierung¹¹ zum Einsatz kommen. Darüber hinaus spielen derartige Regulatormodelle eine entscheidende Rolle im Rahmen der Stabilitätstheorie, bei der man durch Linearisierung um einen Gleichgewichtspunkt oder entlang einer Referenztrajektorie eine lineare Systemdynamik erhält. Das quadratische Zielfunktional ergibt sich dabei aus der Definition einer exponentiell stabilen Lösung im Sinne von LYAPUNOV und der den Funktionenräumen zugrundeliegenden Norm.

Da das analytische Lösen von Optimalsteuerungsproblemen in der Regel nicht zu bewerkstelligen ist, soll zudem ein Zugang für die numerische Behandlung dieser Problemklasse vorgestellt werden. Tatsächlich zeigt sich hierbei, dass ein Großteil bestehender numerischer Verfahren auf dem Begriff des uneigentlichen Riemann-Integrals aufbauen. Da sowohl die Interpretation des Optimalitätsbegriffes als auch die analytische Behandlung der zugrundeliegenden Aufgabe auf der Verwendung des Lebesgue-Integrals basieren, scheint jedoch ein alternativer Zugang angebracht. In dieser Arbeit beziehen wir uns deshalb auf eine in [PB19] vorgestellte Pseudospektralmethode in direkter Form. Die vorgeschlagene Methode basiert auf den Konzepten der Fourier-Analyse unter Ausnutzung der Eigenschaften gewichteter Sobolev- und Lebesgue-Räume mit Dichtefunktionen. Das verwendete Verfahren erlaubt somit eine numerische Behandlung des Optimalsteuerungsproblems ohne „Abschneiden“ des unendlichen Zeithorizontes. Wesentlich hierbei ist die Verwendung der Laguerre-Polynome. Diese bilden eine Basis des gewichteten Lebesgue-Raumes $L_2(\mathbb{R}_{++}, e^{-\beta t})$ ($\beta > 0$)¹². Diese Eigenschaft ermöglicht die numerische Behandlung von Optimalsteuerungsproblemen in gewichteten Sobolev- und Lebesgue-Räumen mit Dichtefunktionen und wurde so in indirekter Form bereits für linear-quadratische Modelle in [Bur19] diskutiert. Für die Anwendbarkeit des vorgestellten Verfahrens auf Optimalsteuerungsprobleme mit allgemeinen Gewichtsfunktionen präsentieren wir in dieser Arbeit einen Zugang über die Dualitätstheorie im Sinne von Lagrange. Diese Vorgehensweise wird uns zur Formulierung einer dualen Aufgabe in Form eines Variations- bzw. Optimalsteuerungsproblems in gewichteten Sobolev- und Lebesgue-Räumen unter Verwendung von Dichtefunktionen führen und erlaubt somit die Anwendung der vorgestellten Pseudospektralmethode auf die duale Aufgabe. Der Nachweis einer starken Dualitätsbeziehung liefert letztlich Rückschlüsse auf die ursprünglich betrachtete Aufgabe.

Die Darlegung dieser Sachverhalte wird in dieser Arbeit wie folgt strukturiert: In Kapitel 1 führen wir die für die weitere Behandlung notwendigen Definitionen und Resultate aus den Bereichen der Analysis und der Funktionalanalysis ein. Dafür beginnen wir mit einem kurzen Einblick in das Konzept der schwachen Topologie, auf welchem die Ergebnisse der vorliegenden Arbeiten aufbauen. Wesentlich wird hierbei die Verwendung der gewichteten Lebesgue- und Sobolev-Räume, die wir in diesem Zusammenhang einführen. Mit einem Einblick in das Gebiet der Stabilisierung und Steuerbarkeit dynamischer Systeme zeigen wir zudem eine Anwendung der besprochenen Problemklasse auf. Die so eingeführten Grundlagen erlauben uns nun die Diskussion linear-quadratischer Opti-

¹¹kurz: SQP-Verfahren; engl. sequential quadratic programming

¹²Mit \mathbb{R}_{++} bezeichnen wir das Intervall $(0, \infty)$.

malsteuerungsmodelle mit a priori gegebenem unendlichen Zeithorizont unter Verwendung allgemeiner Gewichtsfunktionen. Der Blick auf verschiedene Anwendungen aus Wirtschaft, Technik und Biologie motiviert zudem die Einführung sogenannter *Box-Constraints* an die Steuerungsfunktion. Dabei widmen wir uns im Rahmen des [2](#) Kapitels zunächst der Analyse der vorgestellten Problemklasse im Hinblick auf die Existenz und Eindeutigkeit optimaler Lösungen. Aufbauend auf den Ideen aus [Lyk10](#), [Zie19](#) werden wir diese Aufgabenklasse über gewichteten Sobolev- und Lebesgue-Räumen definieren. Dies ermöglicht uns das Studium der Optimalsteuerungsprobleme im Kontext der Hilbertraumtheorie und führt uns schließlich zur Formulierung von Existenzsätzen für Probleme mit Gewichtsfunktionen der Form $e^{\beta t}$ ($\beta \neq 0$). Die so nachgewiesene Existenz einer Lösung für Probleme mit einer echten Gewichtsfunktion motiviert nun die Konstruktion einer dualen Aufgabe mit dem Ziel über diese zu einer Formulierung noch ausstehender Optimalitätsbedingungen zu gelangen.

Die Konstruktion der dualen Aufgabe geschieht dabei auf dem Konzept der Lagrangeschen Dualität und wird Gegenstand der Betrachtungen des [3](#) Kapitels. Einen besonderen Fokus legen wir hierbei auf das Studium der im Rahmen der Konstruktion zu lösenden untergeordneten Optimierungsprobleme. In Kapitel [4](#) setzen wir uns intensiv mit der Struktur der erhaltenen dualen Aufgabe auseinander und gelangen zu einer Formulierung dieser in Form eines Optimalsteuerungsproblems ohne Steuerungsbeschränkungen in gewichteten Funktionenräumen mit Dichtefunktionen. Wir sind in der Lage auch diese Aufgabe in den Kontext der Hilbertraumtheorie einzubetten und weisen unter Ausnutzung der Ideen aus Kapitel [2](#) notwendige Optimalitätsbedingungen vom Pontryagin-Typ für die duale Aufgabe nach. Über die so erhaltenen notwendigen Optimalitätsbedingungen gelingt schließlich der Nachweis der starken Dualitätsbeziehung, welche wiederum den Rückschluss auf notwendige und hinreichende Optimalitätsbedingungen für die primale Aufgabe liefert. Neben der Erweiterung der Formulierung notwendiger Optimalitätsbedingungen für Probleme mit echten Gewichtsfunktionen liefert die so erhaltene Formulierung – aufgrund fehlender Wachstumsbeschränkungen – im Vergleich zu bisherigen Ergebnissen zudem eine Erweiterung der behandelbaren Klasse von Problemen mit Dichtefunktion.

In Kapitel [5](#) präsentieren wir die Ideen der in [PBI9](#) vorgestellten Pseudospektralmethode und wenden diese auf die vorliegende duale Aufgabe an. Die Formulierung der dualen Aufgabe über gewichteten Funktionenräumen mit Dichtefunktion erlaubt uns die Anwendbarkeit dieses Verfahrens nun auch für Aufgabenklassen, die durch die Wahl echter Gewichtsfunktionen geprägt sind. Dafür präsentieren wir im Kapitel [6](#) drei verschiedene Anwendungsbeispiele. Wir beginnen mit der Diskussion eines Problems mit Dichtefunktion aus [GHR11](#). Aufgrund der in Kapitel [4](#) formulierten notwendigen Optimalitätsbedingungen und der vergleichsweise einfachen Struktur der Aufgabe, sind wir in der Lage die numerischen Ergebnisse anhand der exakten analytischen Lösung zu verifizieren. In den nachfolgenden Beispielen stellen wir den Einsatz der Methode zur asymptotischen Stabilisierung eines linearisierten dynamischen Systems vor und gelangen somit zur Betrachtung von Problemen mit echten Gewichtsfunktionen. Dafür beginnen wir mit dem bekannten Lotka-Volterra Modell, für welches wir ebenfalls eine analytische Lösung als Referenz stellen. Die vielversprechenden Ergebnisse motivieren letztlich die Untersuchung eines komplexeren epidemiologischen Modelles in Form des SEIR Modells.

Die Arbeit schließt mit einer Zusammenfassung der Ergebnisse und einem Ausblick auf zukünftige Fragestellungen.

Kapitel 1

Grundlegende Begriffe und Ergebnisse

Sei $(\Omega, \|\cdot\|)$ ein reeller normierter Raum, $M \subseteq \Omega$ nichtleer und $F : M \rightarrow \overline{\mathbb{R}}$ ein reelles Funktional¹, dann ist durch

$$(O) : \quad F(x) \rightarrow \inf! \\ x \in M$$

eine allgemeine Aufgabenstellung der unendlich dimensionalen Optimierung gegeben.

Dabei bezeichnet F das Zielfunktional und M die *zulässige Menge*. Ein Element x heißt *zulässig* zur Aufgabe (O) , falls x in der zulässigen Menge liegt. Im Zusammenhang mit derartigen Aufgabenstellungen sind aus Optimierersicht vor allen Dingen folgende drei Fragestellungen interessant:

- Existiert eine Lösung für das Problem (O) ?
- Ist diese Lösung eindeutig?
- Wie kann ich eine (eventuell eindeutige) Lösung berechnen?

Im Rahmen dieser Arbeit werden wir Problemklassen behandeln, in denen die normierten Räume $(\Omega, \|\cdot\|)$ durch separable Hilberträume gegeben sind. Im Kontext der schwachen Topologie erlaubt uns diese Setzung die Verwendung folgender allgemeiner Prinzipien zum Nachweis der Existenz einer optimalen Lösung:

1. Gemäß der Aussagen des *verallgemeinerten Satzes von Weierstraß*² genügt der Nachweis der schwachen Unterhalbstetigkeit des Zielfunktional F sowie der Nachweis der schwachen Kompaktheit der Menge M für die Sicherung der Existenz einer optimalen Lösung von (O) .
2. Alternativ dazu können wir durch den Nachweis der schwachen Kompaktheit der zugehörigen Niveaumengen

$$\mathcal{N}_M(F, \alpha) := \{x \in M \mid F(x) \leq \alpha\} \tag{1.1}$$

¹ $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$

²vgl. hierfür zum Beispiel Satz [7](#), S. [30](#)

ebenfalls auf die Existenz einer optimalen Lösung des Problems (O) schließen, falls das Zielfunktional F schwach unterhalbstetig ist.

Wesentlich hierbei ist der Begriff der *schwachen Topologie*, welchem wir uns im ersten Teil dieses Kapitels widmen. Die Unterkapitel [1.2](#) und [1.3](#) beinhalten Definitionen und Eigenschaften der im Rahmen der Arbeit verwendeten topologischen Räume. In den darauf folgenden Unterkapiteln wenden wir uns Differentiationsbegriffen zu. Dafür erörtern wir in Unterkapitel [1.4](#) zunächst den Begriff der Differentialgleichung in Sobolev-Räumen. Im Anschluss werden Definitionen für verallgemeinerte Differentiationsbegriffe auf unendlich dimensionalen Räumen angegeben. Das letzte Unterkapitel gibt schließlich einen kurzen Einblick in die asymptotische Steuerbarkeit dynamischer Systeme. Diese Problematik motiviert die später behandelte Problemklasse über die in Unterkapitel [1.3](#) eingeführten gewichteten Funktionenräume mit echter Dichtefunktion.

Alle Unterkapitel sind lediglich als kurzer Einblick zu verstehen. Fehlende Beweise sind den ausgewiesenen Quellen zu entnehmen. Grundlegende Begriffe werden als bekannt vorausgesetzt, können allerdings in verschiedenen Quellen wie beispielsweise [\[Wer07\]](#), [\[Els07\]](#), [\[Kuf81\]](#), [\[ITM09\]](#), [\[Bla82\]](#), [\[AF03\]](#) nachgeschlagen werden.

1.1 Schwache Topologie

Die schwache Topologie ist ein zentrales Konzept der Funktionalanalysis, welche unter anderem auf normierten Räumen definiert wird. Die Ausführungen dieses Unterkapitels beziehen sich auf die folgende, in [\[Bla82\]](#), [\[Wer07\]](#) eingeführte, Definition der *schwachen Topologie*^{[3](#)}:

Sei X ein normierter Raum und X^* sein topologischer Dualraum^{[4](#)}. Wir betrachten die Halbnormen $\{p_f | f \in X^*\}$

$$p_f(x) = |f(x)| \quad x \in X.$$

Diese Halbnormen erzeugen die *schwache Topologie* $\sigma(X, X^*)$ auf X .

Kompaktheit, Beschränktheit und Konvergenz bezüglich der schwachen Topologie $\sigma(X, X^*)$ eines normierten Raumes X mit Dualraum X^* werden dadurch ausgedrückt, dass wir von *schwacher Kompaktheit*, *schwacher Beschränktheit* und *schwacher Konvergenz* sprechen^{[5](#)}. Damit gelangen wir also zu folgender Definition der schwachen Konvergenz:

Definition 1 (schwache Konvergenz^{[6](#)}). Sei X ein normierter Raum und X^* sein Dualraum. Eine Folge $\{x_n\}$ aus X heißt schwach konvergent gegen x_0 (Schreibweise $x_n \rightharpoonup x_0$), falls für alle stetigen linearen Funktionale $f \in X^*$ gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

³vgl. [\[Bla82\]](#), S. 256 ff., [\[Wer07\]](#), S. 327 ff.

⁴Der Raum X^* aller linearen stetigen Funktionale auf dem normierten Raum X heißt (topologischer) Dualraum von X .

⁵vgl. [\[Bla82\]](#), S. 264. Reden wir hingegen von (*starker*) *Kompaktheit*, (*starker*) *Beschränktheit*, (*starker*) *Konvergenz* so ist die Kompaktheit, Beschränktheit und Konvergenz der durch die (des normierten Raumes zugrundeliegenden) Norm induzierten Topologie – auch *Normtopologie* oder *starke Topologie* genannt – zu verstehen.

⁶vgl. [\[Wer07\]](#), S. 106, Definition III.3.6

Über die Definition der schwachen Konvergenz gelangen wir nun zu Formulierungen von Abgeschlossenheit und Kompaktheit im Kontext der schwachen Topologie. Diese lassen sich wie folgt einführen:

Definition 2 (schwache Abgeschlossenheit). Eine Teilmenge M eines normierten Raumes X heißt schwach abgeschlossen, falls aus $x_n \in M$ mit $x_n \rightharpoonup x_0$ auch $x_0 \in M$ folgt.

Definition 3 (schwache Kompaktheit). Sei X ein normierter Raum. Eine Menge $M \in X$ heißt schwach kompakt, wenn jede unendliche Folge aus M eine schwach konvergente Teilfolge enthält, deren schwacher Grenzwert x_0 ebenfalls in M liegt.

Wesentlich für den Nachweis der Existenz einer optimalen Lösung wird der Nachweis der schwachen Kompaktheit der zulässigen Menge sein. Unsere Argumente werden dabei hauptsächlich auf den folgenden zwei Sätzen basieren⁷:

Satz 1 ([Bla82], S. 264). Eine Teilmenge M eines reflexiven Banachraumes X ist schwach kompakt genau dann, wenn M schwach abgeschlossen und beschränkt ist.

Satz 2 ([Alt12], Satz 6.13, S. 252). Eine konvexe Teilmenge M eines reflexiven Banachraumes ist schwach abgeschlossen, wenn M abgeschlossen bzgl. der Normtopologie ist⁸.

Die Prinzipien für den Nachweis der Existenz einer optimalen Lösung motivieren zudem die Angabe folgender Definition:

Definition 4 (schwache Unterhalbstetigkeit). Ein Funktional $f : X \rightarrow \mathbb{R}$ auf dem normierten Raum X heißt schwach unterhalbstetig in einem Punkt x_0 , wenn für jede gegen x_0 schwach konvergente Folge $\{x_n\} \subset X$ gilt:

$$\liminf_{n \rightarrow \infty} f(x_n) \geq f(x_0).$$

1.2 Klassische Lebesgue- und Sobolev-Räume

Im Rahmen dieses Unterkapitels sollen die für die vorliegende Arbeit wesentlichen Definitionen und Sätze der klassischen Lebesgue- und Sobolev-Räume festgehalten werden. Dabei setzen wir voraus, dass die wesentlichen Eigenschaften der hier behandelten Räume dem Leser bekannt seien. Für eine ausführlichere Übersicht verweisen wir auf [BZ18].

Sei im Folgenden Ω eine offene, zusammenhängende, nichtleere Menge mit $\Omega \subseteq \mathbb{R}$ und $\mathcal{M}^n(\Omega) = \mathcal{M}(\Omega; \mathbb{R}^n)$ die Menge aller Lebesgue-messbaren Funktionen, die auf der Menge Ω definiert sind und Werte im \mathbb{R}^n haben.

Definition 5 (Lebesgue-Raum (vgl. [AF03], S. 23 f., 2.1)). Sei $1 \leq p < \infty$. Dann bezeichnen wir mit $L_p^n(\Omega)$ den Raum der messbaren und im Sinne von Lebesgue p -fach integrierbaren Funktionen:

$$L_p^n(\Omega) := \left\{ x \in \mathcal{M}^n(\Omega) \mid \int_{\Omega} |x(t)|^p dt < \infty \right\}^9.$$

⁷Wir werden im späteren Verlauf zeigen, dass die von uns verwendeten Funktionalräume reflexive Banachräume sind.

⁸Wir nennen die Abgeschlossenheit bzgl. der Normtopologie im Folgenden auch (*starke*) *Abgeschlossenheit*.

⁹Dabei bezeichne $|\cdot|$ im gesamten Verlauf der Arbeit die euklidische Norm.

Bemerkung 1. Gilt in der Definition [5](#) der Fall $n = 1$, so schreiben wir kurz $L_p(\Omega)$. Wir identifizieren Funktionen des $L_p^n(\Omega)$ miteinander, wenn diese im Sinne des Lebesgue-Maßes fast überall^{[10](#)} in Ω übereinstimmen. Damit bildet der Raum $L_p^n(\Omega)$ den Raum der Äquivalenzklassen $[x]$ von Funktionen x . Für eine bessere Lesbarkeit einigen wir uns jedoch an dieser Stelle auf die Schreibweise $x \in L_p^n(\Omega)$ statt $[x] \in L_p^n(\Omega)$ und $x = 0$ in $L_p^n(\Omega)$ falls $x(t) = 0$ für fast alle $t \in \Omega$.

Definition 6 (vgl. [For09](#), S. 58). Eine Funktion $x : \Omega \rightarrow \mathbb{R}$ heißt lokal auf Ω Lebesgue-integrierbar, falls sie auf jeder offenen Menge U mit den Eigenschaften

(i) $\bar{U} \subset \Omega$ und

(ii) \bar{U} ist kompakte Teilmenge von \mathbb{R}

Lebesgue-integrierbar ist^{[11](#)}. Die Menge dieser Funktionen bezeichnen wir mit $L_{1,loc}(\Omega)$. Wir definieren weiterhin

$$L_{1,loc}^n(\Omega) := \{x : \Omega \rightarrow \mathbb{R}^n \mid x_i \in L_{1,loc}(\Omega) \text{ für alle } i = 1, \dots, n\}.$$

Definition 7 (vgl. [SAS13](#), S. 27). Der Raum

$$L_\infty^{n \times m}(\Omega) := \left\{ M \in \mathcal{M}^{n \times m}(\Omega) \mid \operatorname{ess\,sup}_{t \in \Omega} |M(t)| < \infty \right\}$$

heißt Raum der wesentlich beschränkten Matrixfunktionen.

Bemerkung 2. Gilt in der Definition [7](#) der Fall $m = 1$, so erhalten wir den Raum der wesentlich beschränkten Vektorfunktionen $L_\infty^n(\Omega)$.

Für die Einführung der Sobolev-Räume benötigen wir die Definition der *schwachen Ableitung*, welche eine Erweiterung des „klassischen“ Ableitungsbegriffes darstellt und auf der Verwendung von *Testfunktionen* basiert. Diese wiederum führen wir wie folgt ein:

Definition 8. Sei $x : \Omega \rightarrow \mathbb{R}^n$ eine Funktion, die fast überall auf Ω definiert ist. Dann bezeichnen wir mit

$$\operatorname{supp}(x) := \overline{\{t \in \Omega : x(t) \neq 0\}}$$

den Träger der Funktion x .

Die Menge der Testfunktionen bezeichnen wir mit $C_0^\infty(\Omega, \mathbb{R}^n)$ und definieren diese über die Menge^{[12](#)}

$$C_0^\infty(\Omega, \mathbb{R}^n) := \{x \in C^\infty(\Omega, \mathbb{R}^n) : \operatorname{supp}(x) \text{ ist kompakte Teilmenge von } \Omega\}.$$

Definition 9 (schwache Ableitung) (vgl. [Wer07](#), S. 163 ff.). Sei $x \in L_{1,loc}(\Omega)$. Existiert eine Funktion $y \in L_{1,loc}(\Omega)$, so dass

$$\int_\Omega x(t) D^\alpha \phi(t) dt = (-1)^{|\alpha|} \int_\Omega y(t) \phi(t) dt, \quad t \in \Omega,$$

¹⁰kurz: f.ü., in Zeichen (\forall)

¹¹Dabei bezeichnen wir mit \bar{U} den Abschluss der Menge U .

¹²Für $k \in \mathbb{N}_0$ bezeichnen wir mit $C^k(\Omega, \mathbb{R}^n)$ die Menge aller Funktionen $x : \Omega \rightarrow \mathbb{R}^n$, die zusammen mit ihren partiellen Ableitungen bis zur Ordnung k stetig auf Ω sind. Für $C^0(\Omega, \mathbb{R}^n)$ schreiben wir $C(\Omega, \mathbb{R}^n)$. Im Fall $n = 1$ setzen wir $C^k(\Omega) := C^k(\Omega, \mathbb{R})$.

mit dem Multiindex $\alpha := (\alpha_1, \dots, \alpha_s) \in \mathbb{N}_0^s$ für den gilt $|\alpha| := \sum_{i=1}^s \alpha_i \leq k$ und

$$\mathcal{D}^\alpha \phi := \frac{\partial^{|\alpha|}}{\partial t_1^{\alpha_1} \dots \partial t_s^{\alpha_s}} \phi$$

für alle Testfunktionen $\phi \in C_0^\infty(\Omega, \mathbb{R})$ erfüllt ist, dann heißt y die schwache partielle Ableitung von x und wird mit $\mathcal{D}^\alpha = y$ bezeichnet.

Auf Basis der letzten beiden Definitionen können wir unter Verwendung der α -ten schwachen Ableitung \mathcal{D}^α die klassischen Sobolev-Räume einführen.

Definition 10 (Sobolev-Raum) (vgl. [AF03], S. 59 f.). Sei $1 \leq p < \infty$ und $k \in \mathbb{N}_0$. Der Raum

$$W_p^{k,n}(\Omega) := \left\{ x \in L_p^n(\Omega) \mid \mathcal{D}^\alpha x \in L_p^n(\Omega) \text{ für alle } \alpha \text{ mit } |\alpha| \leq k \right\}$$

heißt Sobolev-Raum. Darüber hinaus definieren wir den Raum

$$W_{1,loc}^{k,n}(\Omega) := \left\{ x \in L_{1,loc}^n(\Omega) \mid \mathcal{D}^\alpha x \in L_{1,loc}^n(\Omega) \text{ für alle } \alpha \text{ mit } |\alpha| \leq k \right\}.$$

Bemerkung 3. Gilt in der Definition [10] der Fall $n = 1$, so schreiben wir kurz $W_p^k(\Omega)$, bzw. $W_{1,loc}^k(\Omega)$.

Für das Intervall $\Omega := I \subseteq \mathbb{R}$ liefert der nachfolgende Satz einen nützlichen Zusammenhang zwischen der schwachen Ableitung einer Funktion $x \in W_{1,loc}^1(I)$ und der punktweise fast überall existierenden „klassischen“ Ableitung ihres Repräsentanten $\tilde{x} \in AC_{loc}(I)$ [13], bezeichnet mit \tilde{x}' .

Satz 3. ([KJF77], S. 274, 5.6.3 Theorem) Sei $I \subseteq \mathbb{R}$ und $x \in W_{1,loc}^1(I)$. Dann existiert eine Funktion $\tilde{x} \in AC_{loc}(I)$, so dass $x = \tilde{x}$ und $\mathcal{D}x = \tilde{x}'$ fast überall auf I gelten.

Wesentlich für spätere Argumentationen im Rahmen der Betrachtungen von Lebesgue- und Sobolev-Räumen sind zudem die Aussagen der folgenden zwei Sätze [14]:

Satz 4 (Sobolevscher Einbettungssatz / Satz von Rellich-Kondrachov) (vgl. [Bla82], S. 274 ff. und [AF03], S. 144 Satz 6.2)). Es sei $\Omega \subset \mathbb{R}$ ein Gebiet und $\Omega_0 \subset \Omega$ ein beschränktes Teilgebiet. Ferner sei $1 \leq p < \infty$ und $k = 1, 2, \dots$

a) Es bestehen die folgenden stetigen Einbettungen:

$$\begin{aligned} i) \quad kp = 1: & \quad W_p^k(\Omega) \hookrightarrow L_q(\Omega), \quad p \leq q < \infty; \\ ii) \quad kp > 1: & \quad W_p^k(\Omega) \hookrightarrow C^0(\Omega). \end{aligned}$$

b) Die folgenden Einbettungen sind kompakt:

$$\begin{aligned} i) \quad kp = 1: & \quad W_p^k(\Omega) \hookrightarrow L_q(\Omega_0), \quad p \leq q < \infty; \\ ii) \quad kp > 1: & \quad W_p^k(\Omega) \hookrightarrow C^0(\Omega). \end{aligned}$$

¹³Mit $AC_{loc}(I)$ bezeichnen wir den Raum der auf I lokal absolut stetigen Funktionen.

¹⁴Die hier angeführte Version stellt eine Spezialisierung für die in dieser Arbeit behandelten Fälle dar. Für eine allgemeinere Formulierung dieses Satzes verweisen wir auf die angegebenen Quellen.

- c) Falls Ω beschränkt ist, existieren die stetigen Einbettungen i) und ii) in a) auch für die Indizes q , $1 \leq q < p$. Zudem können wir in b) für die Einbettungen i) $\Omega_0 = \Omega$ setzen.

Satz 5 (Satz von Rellich-Kondrachov Part III) (vgl. [AF03], S. 168, Theorem 6.3). Sei das beschränkte Gebiet $\Omega_0 = (a, b)$ mit $-\infty < a < b < \infty$ gegeben und seien weiter $j \geq 0$ und $k \geq 1$ ganzzahlig sowie $1 < p < \infty$. Dann ist die Einbettung

$$W_p^{j+k}(\Omega_0) \hookrightarrow C^j(\overline{\Omega_0})$$

kompakt, falls $kp > k$ gilt.

Bemerkung 4 [15]. Setzt man $k = 1$ in Satz [5], so folgt für $p > 1$ und für alle $j \geq 0$ die Kompaktheit der Einbettung

$$W_p^{j+1}((a, b)) \hookrightarrow C^j([a, b]).$$

Dies impliziert, dass für eine Sobolev-Funktion $x \in W_p^k((a, b))$ mit $k \geq 1$ und $p > 1$ sogar ein $(k - 1)$ -mal stetig differenzierbarer Repräsentant $\hat{x} \in C^{k-1}([a, b])$ existiert. Dadurch wird zusätzlich die punktweise Auswertung bis zur Ordnung $k - 1$ legitimiert und insbesondere können die Randwerte sowohl für x selbst als auch für $x', x'', \dots, x^{(k-1)}$ eindeutig festgelegt werden.

1.3 Gewichtete Lebesgue- und Sobolev-Räume

Für die Betrachtung von Optimalsteuerungsproblemen mit unendlichem Zeithorizont ist insbesondere der Fall $\Omega = \mathbb{R}_{++} = (0, \infty)$ interessant. Es ist leicht einzusehen, dass die bisher eingeführten klassischen Lebesgue- und Sobolev-Räume die Menge von Funktionen, welche den Definitionen [5] und [10] für $\Omega = (0, \infty)$ genügen, stark einschränkt. Dies motiviert die Einführung gewichteter Lebesgue- und Sobolev-Räume. Wesentlich dafür ist der Begriff der Gewichtsfunktion:

Definition 11 (Gewichtsfunktion / Dichtefunktion). Eine stetige Funktion $\nu : \Omega \rightarrow \mathbb{R}_{++}$ heißt Gewichtsfunktion, falls ν und $\nu^{-1/(p-1)} \in L_{1,loc}(\Omega)$ für $p > 1$. Die Menge aller Gewichtsfunktionen nennen wir $\mathcal{W}(\Omega)$. Gilt für eine Gewichtsfunktion ν zudem

$$\int_0^\infty \nu(t) dt < \infty \tag{1.2}$$

so nennen wir diese Dichtefunktion [16].

Definition 12 (Gewichteter Lebesgue-Raum). Sei $\nu \in \mathcal{W}(\Omega)$ und $1 \leq p < \infty$. Dann bezeichnen wir mit

$$L_p^n(\Omega, \nu) := \left\{ x \in \mathcal{M}^n(\Omega) \mid \int_\Omega |x(t)|^p \nu(t) dt < \infty \right\}$$

den gewichteten Lebesgue-Raum der Ordnung p .

¹⁵nachzulesen in [Bur19], S. 14

¹⁶Eine Gewichtsfunktion ν , welche der Bedingung (1.2) nicht genügt, nennen wir im Folgenden auch echte Gewichtsfunktion.

Bemerkung 5. Gilt in der Definition [12](#) der Fall $n = 1$, so schreiben wir kurz $L_p(\Omega, \nu)$.

Für Räume dieser Art gelten die folgenden Eigenschaften:

Lemma 1 ([\[DS88\]](#), S. 146, S. 243, S. 288^{[17](#)}). Sei $\nu \in \mathcal{W}(\Omega)$ und $1 \leq p < \infty$. Der Raum $L_p^n(\Omega, \nu)$ versehen mit der Norm

$$\|x\|_{(p), \nu} := \left(\int_{\Omega} |x(t)|^p \nu(t) dt \right)^{1/p}$$

genügt den folgenden Eigenschaften:

1. $L_p^n(\Omega, \nu)$ ist für $1 \leq p < \infty$ ein separabler Banachraum.
2. $L_p^n(\Omega, \nu)$ ist für $1 < p < \infty$ ein reflexiver Raum.
3. $L_2^n(\Omega, \nu)$ wird mit dem Skalarprodukt

$$\langle x, y \rangle_{(2), \nu} = \int_{\Omega} x^T(t) y(t) \nu(t) dt$$

ein Hilbertraum.

Im Hinblick auf die spätere Konstruktion der dualen Aufgabe wollen wir zudem auf folgende Eigenschaften des so eingeführten Skalarproduktes hinweisen:

Bemerkung 6 ([18](#)). Für $x \in L_p^n(\mathbb{R}_{++}, \nu)$ und $y \in L_q^n(\mathbb{R}_{++}, \nu^{1-q})$ mit $1/p + 1/q = 1$ definiert das Skalarprodukt $\langle x, y \rangle_{(2)}$ im $L_2^n(\mathbb{R}_{++})$ eine stetige Bilinearform, denn es gilt

$$\langle x, y \rangle_{(2)} \leq \int_0^{\infty} |x(t)| \nu^{1/p}(t) |y(t)| \nu^{-1+1/q}(t) dt \leq \|x\|_{(p), \nu} \|y\|_{(q), \nu^{1-q}}.$$

Mit dem Riesz'schen Darstellungssatz^{[19](#)} gilt für $p = 2$: $[L_2^n(\mathbb{R}_{++}, \nu)]^* = L_2^n(\mathbb{R}_{++}, \nu)$. Die Skalarprodukte in $L_2^n(\mathbb{R}_{++}, \nu)$ und $L_2^n(\mathbb{R}_{++})$ stehen also in folgender Relation zueinander: Für ein $x \in L_2^n(\mathbb{R}_{++}, \nu)$ und ein $y \in L_2^n(\mathbb{R}_{++}, \nu^{-1})$ existiert ein $\hat{y} \in L_2^n(\mathbb{R}_{++}, \nu)$ mit

$$\langle x, \hat{y} \rangle_{(2), \nu} = \langle x, y \rangle_{(2)} \quad \text{und} \quad \hat{y}(t) = y(t) \nu^{-1}(t) \quad \forall t \in \mathbb{R}_{++}.$$

Aufbauend auf den gewichteten Lebesgue-Räumen können wir nun auch die Definition der Sobolev-Räume verallgemeinern:

Definition 13 (Gewichteter Sobolev-Raum). Sei $\nu \in \mathcal{W}(\Omega)$ und $1 \leq p < \infty$ und $k \in \mathbb{N}_0$. Der Raum

$$W_p^{k,n}(\Omega, \nu) := \left\{ x \in L_p^n(\Omega, \nu) \mid \mathcal{D}^\alpha x \in L_p^n(\Omega, \nu) \text{ für alle } \alpha \text{ mit } |\alpha| \leq k \right\}$$

heißt gewichteter Sobolev-Raum zur Ordnung p .

Bemerkung 7. Gilt in der Definition [13](#) der Fall $n = 1$, so schreiben wir kurz $W_p^k(\Omega, \nu)$.

¹⁷Für die Verallgemeinerung auf den Fall vektorwertiger Räume siehe [\[Bur19\]](#), S. 34 Lemma 2.35.

¹⁸vgl. [\[Lyk10\]](#), S. 105

¹⁹vgl. [\[Els07\]](#), S. 279

Lemma 2 ([DS88], S. 146, S. 243, S. 288²⁰). Sei $\nu \in \mathcal{W}(\Omega)$ und $1 < p < \infty$ und $k \in \mathbb{N}_0$. Der Raum $W_p^{k,n}(\Omega, \nu)$ versehen mit der Norm

$$\|x\|_{k,(p),\nu} := \left(\|x\|_{(p),\nu}^p + \sum_{0 < |\alpha| \leq k} \|\mathcal{D}^\alpha x\|_{(p),\nu}^p \right)^{1/p}$$

ist

1. ein Banachraum.
2. für $p = 2$ mit dem Skalarprodukt

$$\langle x, y \rangle_{k,(2),\nu} := \langle x, y \rangle_{(2),\nu} + \sum_{0 < |\alpha| \leq k} \langle \mathcal{D}^\alpha x, \mathcal{D}^\alpha y \rangle_{(2),\nu}$$

ein Hilbertraum.

Bemerkung 8. Ist $\nu \in \mathcal{W}(\Omega)$ und $p > 1$, so gilt die Inklusion $L_p^n(\Omega, \nu) \subseteq L_{1,loc}^n(\Omega)$. Daraus folgt sofort $W_p^{k,n}(\Omega, \nu) \subseteq W_{1,loc}^{k,n}(\Omega)$.

Folgerung 1. Sei $\nu \in \mathcal{W}(\Omega)$, $1 \leq p < \infty$ und $k \in \mathbb{N}_0$. Ist $\Omega := I \subseteq \mathbb{R}$, so existiert zu $x \in W_p^{k,n}(I, \nu)$, $k \geq 1$ ein lokal absolut stetiger Repräsentant $\tilde{x} \in AC_{loc}(I, \mathbb{R}^n)$.

Zum Abschluss wollen wir noch die Definition der nicht uniform gewichteten Sobolev-Räume angeben. Auf diese werden wir im Kapitel 5 für die Herleitung des numerischen Verfahrens zurückgreifen. Dafür führen wir zunächst die folgende Bezeichnung ein: Sei $k \in \mathbb{N}_0$ und α ein Multiindex der Ordnung $|\alpha| \leq k$. Dann bezeichnen wir mit S_k die Menge der Gewichtsfunktionen definiert auf Ω , welche die Gestalt

$$S_k := \{\nu_\alpha \in \mathcal{W}(\Omega), \forall \alpha : 0 \leq |\alpha| \leq k\} \quad (1.3)$$

besitzt.

Definition 14. Für $1 \leq p < \infty$ definieren wir den gewichteten Sobolev-Raum $W_p^k(\Omega, S_k)$ als die Menge aller Funktionen

$$x \in L_p(\Omega, \nu_0) \cap L_{1,loc}(\Omega),$$

die für jeden Multiindex α mit $|\alpha| \leq k$ eine schwache partielle Ableitung $\mathcal{D}^\alpha x$ bis zur Ordnung k besitzen, so dass

$$\mathcal{D}^\alpha x \in L_p(\Omega, \nu_\alpha) \cap L_{1,loc}(\Omega)$$

gilt. Durch

$$\|x\|_{k,(p),S_k} := \left(\sum_{0 \leq |\alpha| \leq k} \|\mathcal{D}^\alpha x\|_{(p),\nu_\alpha}^p \right)^{1/p}$$

wird eine Norm auf dem $W_p^k(\Omega, S_k)$ eingeführt.

Lemma 3 ([Fau09], S. 13, Satz 3.21). Es sei S_k wie in (1.3) festgelegt.

²⁰Für die Verallgemeinerung der nachfolgenden Ergebnisse auf den Fall vektorwertiger Räume siehe [Bur19], S. 35.

1. Gilt $\nu_\alpha \in L_\infty(\Omega)$ für alle $\nu_\alpha \in S_k$, dann erhalten wir die Inklusion

$$W_p^k(\Omega) \subseteq W_p^k(\Omega, S_k).$$

2. Ist $\nu_\alpha(t) \geq \kappa > 0$ fast überall auf Ω für alle $\nu_\alpha \in S_k$ erfüllt, so gilt

$$W_p^k(\Omega, S_k) \subseteq W_p^k(\Omega).$$

Folgerung 2 ^[21]. Sei $k \geq 1$, $1 < p < \infty$, $-\infty < a < b < \infty$ und S_k eine Menge von Gewichtsfunktionen, so dass die Äquivalenz $W_p^k((a, b)) = W_p^k((a, b), S_k)$ gilt. Dann existiert für jede Äquivalenzklasse $[x]$ aus $W_p^k((a, b), S_k)$ ein $(k-1)$ -mal stetig differenzierbarer Repräsentant $\hat{x} \in C^{k-1}([a, b])$.

1.4 Differentialgleichungen in Sobolev-Räumen

Gemäß der Definition ^[13] ist der gewichtete Sobolev-Raum über den Begriff der schwachen Ableitung definiert. Ist $x \in W_p^{1,n}(\mathbb{R}_{++}, \nu)$, so können wir daher lediglich auf die Raumbühörigkeit $\mathcal{D}x \in L_p^n(\mathbb{R}_{++}, \nu)$ schließen. In Folge dessen können wir für Funktionen $x \in W_p^{1,n}(\mathbb{R}_{++}, \nu)$ zunächst nur Darstellungen einer schwach formulierten Differentialgleichung

$$\mathcal{D}x(\cdot) = f(\cdot, x(\cdot)) \quad \text{auf } (0, \infty)$$

angeben. Um die Wohldefiniertheit der später zu betrachtenden Optimalsteuerungsprobleme sicherzustellen, ist es allerdings notwendig, Anfangswertprobleme der Form

$$x'(t) = f(t, x(t)) \quad \text{für fast alle } t \in \mathbb{R}_{++}, \quad x(0) = x_0$$

sinnvoll zu interpretieren. In ^{[Bur19]^[22]} wird gezeigt, dass die Identifizierung der schwach formulierten Differentialgleichung und der punktwweisen Formulierung für Funktionen $x \in W_2^1(\mathbb{R}_{++}, \nu)$ mit $\nu(t) = e^{\beta t}$ mit $\beta < 0$ zulässig ist. Wir möchten in diesem Unterkapitel die dort getätigten Ausführungen skizzieren und den Fall einer Gewichtsfunktion $\nu(t) = e^{\beta t}$ mit $\beta > 0$ inkludieren. Dieses Unterkapitel ist daher als Auszug der genannten Quelle zu verstehen. Für eine ausführlichere Darstellung verweisen wir auf ^[Bur19]. Wesentlich für die Behandlungen in diesem Unterkapitel sind dabei die bereits dargestellten Eigenschaften der klassischen und gewichteten Lebesgue- und Sobolev-Räume ^[23].

Wir betrachten auf dem Raum der absolut stetigen Funktionen $AC^m([0, T])$ für ein $T \in (0, \infty)$ die zwei Formen der Differentialgleichung

$$\mathcal{D}x(\cdot) = f(\cdot, x(\cdot)) \quad \text{auf } (0, T) \tag{1.4}$$

$$x'(t) = f(t, x(t)) \quad \text{f.ü. auf } (0, T). \tag{1.5}$$

Bekanntermaßen ist die schwache Ableitung eindeutig bestimmt ^[24] und stimmt für eine absolut stetige Funktion fast überall mit ihrer klassischen Ableitung überein. Demnach sind die Gleichungen ^(1.4) und ^(1.5) äquivalent für Funktionen $x \in AC^m([0, T])$. Den Übergang zu Differentialgleichungen in Sobolev-Räumen erhalten wir nun über den folgenden

²¹nachzulesen beispielsweise in ^[Bur19], S. 34

²²siehe S. 77 ff.

²³Wesentlich sind hierbei insbesondere die Aussagen aus Satz ^[5], S. ^[12] und Lemma ^[3], S. ^[14].

²⁴vgl. ^[Wer07], S. 164

Zusammenhang²⁵:

Die Funktionenräume $W_1^{1,n}((0, T))$ und $AC^n([0, T])$ sind identisch in dem Sinne, dass für jedes $f \in W_1^{1,n}((0, T))$ eine Funktion $g \in AC^n([0, T])$ (und umgekehrt) mit $f(t) = g(t)$ fast überall auf $(0, T)$ existiert. Daher können wir jede Funktion $x \in W_1^{1,n}(0, T)$ mit ihrem absolut stetigen Repräsentanten identifizieren und auf $[0, T]$ fortsetzen. Dies erlaubt eine punktweise Auswertung an den Intervallrändern und rechtfertigt die sinnvolle Definition eines Anfangswertes $x(0) = x_0$. In diesem Sinne können wir die Gleichungen (1.4) und (1.5) auch für Funktionen im $W_1^{1,n}((0, T))$ als äquivalent betrachten. Genügt die Funktion f zudem der *Carathéodory-Bedingung*²⁶, so kann die Existenz einer Lösung für das zugehörige Anfangswertproblem mit $x(0) = x_0$ garantiert und durch die zu (1.4) und (1.5) äquivalenten Integralgleichung

$$x(t) = x(0) + \int_0^t f(\tau, x(\tau)) d\tau \quad (\forall) t \in (0, T)$$

angegeben werden²⁷.

Der Übergang zu gewichteten Sobolev-Räumen auf unendlichem Zeithorizont wird durch die folgende Überlegung motiviert: Wir betrachten die Einschränkung x_T einer Funktion $x \in W_2^{1,n}(\mathbb{R}_{++}, \nu)$ mit $\nu(t) = e^{\beta t}$, $\beta \neq 0$ auf das beschränkte Intervall $(0, T)$. Offenbar gilt dann auch $x_T \in W_2^{1,n}((0, T), \nu)$. Aufgrund der Gültigkeit der Inklusion

$$W_2^{1,n}((0, T), \nu) \subseteq W_2^{1,n}((0, T)) \subset W_1^{1,n}((0, T))$$

folgt sofort $x_T \in W_1^{1,n}((0, T))$. Mit den Ausführungen über das endliche Zeitintervall $(0, T)$ folgt – unter der Voraussetzung f erfülle die Carathéodory-Bedingungen – damit die Äquivalenz der Differentialgleichungen

$$\begin{aligned} \mathcal{D}x(\cdot) &= f(\cdot, x(\cdot)) && \text{auf } (0, \infty) \\ x'(t) &= f(t, x(t)) && \text{f.ü. auf } (0, \infty) \end{aligned}$$

für Funktionen $x \in W_2^{1,n}(\mathbb{R}_{++}, \nu)$ sowie die Möglichkeit eines Überganges zur äquivalenten Integraldarstellung

$$x(t) = x(0) + \int_0^t f(\tau, x(\tau)) d\tau \quad (\forall) t \in (0, \infty).$$

Damit ist das Anfangswertproblem

$$x'(t) = f(t, x(t)) \quad (\forall) t \in \mathbb{R}_{++}, \quad x(0) = x_0$$

im Raum $W_2^{1,n}(\mathbb{R}_{++}, \nu)$ wohldefiniert und die Existenz eines eindeutig bestimmten lokal absolut stetigen Repräsentanten erlaubt es zu einer punktweisen Darstellung überzugehen.

²⁵In [Dik16] erörtert Dikariev den Zusammenhang zwischen absolut stetigen Funktionen und Sobolev-Funktionen. Für eine ausführlichere Darstellung verweisen wir daher auf die angegebene Quelle.

²⁶vgl. Definition 24, S. 185

²⁷vgl. hierzu [Bur19], S. 79 ff.

1.5 Einführung wichtiger Differentiationsbegriffe

Im Rahmen der Diskussion von Optimalitätskriterien werden wir mit der Differentiation auf unendlich dimensionalen Räumen konfrontiert. In diesem Unterkapitel wollen wir die dafür verwendeten Differentiationsbegriffe vorstellen und einige Zusammenhänge und Eigenschaften der eingeführten Differentiationsbegriffe skizzieren.

Das *Gâteaux-Differential*, benannt nach René Gâteaux (1889-1914), stellt eine Verallgemeinerung des gewöhnlichen Differentiationsbegriffes dar, indem es die Richtungsableitung auch in unendlich dimensionalen Räumen definiert.

Definition 15 (Gâteaux-Differential, Gâteaux-Ableitung ([ITM09], S. 36)). Seien X und Y lineare topologische Räume, U eine Umgebung des Punktes $x \in X$ und $F : U \rightarrow Y$ eine Abbildung von U in Y . Existiert für alle $h \in X$ der Grenzwert

$$\delta F(x; h) := \lim_{\lambda \rightarrow 0} \frac{F(x + \lambda \cdot h) - F(x)}{\lambda} = \left. \frac{dF(x + \lambda \cdot h)}{d\lambda} \right|_{\lambda=0},$$

dann nennen wir $\delta F(x; h)$ das Gâteaux-Differential bzw. die erste Variation von F im Punkt x in Richtung h .

Existiert ein linearer stetiger Operator $\Lambda : X \rightarrow Y$ (d.h. ist die erste Variation eine lineare stetige Abbildung) mit $\Lambda h = \delta F(x; h)$, so wird der Operator Λ Gâteaux-Ableitung von F im Punkt x genannt und mit $F'_G(x)$ bezeichnet. Die Funktion F heißt dann Gâteaux-differenzierbar im Punkt x .

Im späteren Verlauf der Arbeit benötigen wir zudem das *einseitige Gâteaux-Differential*. Dieses erhalten wir über die einseitige Grenzwertbetrachtung der Form

$$\delta^+ F(x; h) := \lim_{\lambda \rightarrow 0+0} \frac{F(x + \lambda \cdot h) - F(x)}{\lambda} \quad \text{bzw.}$$

$$\delta^- F(x; h) := \lim_{\lambda \rightarrow 0-0} \frac{F(x + \lambda \cdot h) - F(x)}{\lambda}.$$

Maurice René Fréchet (1878-1973) lieferte durch die Einführung des *Fréchet-Differentials* darüber hinaus eine Verallgemeinerung der totalen Differenzierbarkeit auf normierten Räumen.

Definition 16 (Fréchet-Differential, Fréchet-Ableitung ([ITM09], S. 36)). Seien X und Y Banachräume, U eine Umgebung des Punktes $x \in X$ und $F : U \rightarrow Y$ eine Abbildung von U in Y . Die Abbildung F wird Fréchet-differenzierbar im Punkt x genannt, wenn ein linearer stetiger Operator $\Lambda : X \rightarrow Y$ existiert mit

$$F(x + h) = F(x) + \Lambda h + r(h),$$

wobei

$$\frac{\|r(h)\|_Y}{\|h\|_X} \rightarrow 0 \quad \text{für} \quad \|h\|_X \rightarrow 0$$

gilt. Der Operator Λ heißt dann Fréchet-Ableitung von F im Punkt x und wird mit $F'_F(x)$ bezeichnet. Λh bezeichnen wir als Fréchet-Differential.

Zwischen diesen Ableitungsbegriffen bestehen folgende Zusammenhänge:

Satz 6 ([ITM09], Hilfssatz 1, S. 37). *Es gelten folgende Aussagen:*

- (a) *die Operatoren $F'_G(x)$ und $F'_F(x)$ sind eindeutig bestimmte Abbildungen;*
- (b) *wenn die Abbildung F einer Umgebung des Punktes x des Banachraumes X in den Banachraum Y im Punkt x Fréchet-differenzierbar ist, so ist sie in diesem Punkt stetig und Gâteaux-differenzierbar, wobei gilt*

$$F'_G(x) = F'_F(x);$$

- (c) *wenn die Abbildung F einer Umgebung U des Punktes x des linearen topologischen Raumes X in den linearen topologischen Raum Y im Punkt x Gâteaux-differenzierbar ist, so ist in diesem Punkt die erste Variation dieser Abbildung definiert und es gilt*

$$F'_G(x)h = \delta F(x; h).$$

Gemäß der Aussagen des obigen Satzes liefert das Fréchet-Differential also Rückschlüsse auf das Gâteaux-Differential. Wir wollen Bedingungen angeben, unter denen wir auch von der Gâteaux-Differenzierbarkeit auf die Fréchet-Differenzierbarkeit schließen können und definieren dafür zunächst

Definition 17 (Gleichmäßige Operatortopologie ([ITM09], S. 36)). *Sind X und Y Banachräume, so nennt man die durch die Norm*

$$\Lambda = \sup_{x \in X} \left(\frac{\|\Lambda x\|_Y}{\|x\|_X} \right)$$

erzeugte Topologie die gleichmäßige Operatortopologie im Raum $\mathcal{L}(X, Y)$ der linearen stetigen Abbildungen aus X in Y .

Folgerung 3 ([ITM09] S. 41). *Es seien X ein Banachraum und F eine stetige Abbildung einer Umgebung U des Punktes $x_0 \in X$ in den Banachraum Y . Wir setzen voraus, die Abbildung F sei in jedem Punkt der Menge U Gâteaux-differenzierbar und die Abbildung $x \mapsto F'_G(x)$ aus U in den mit der gleichmäßigen Operatortopologie versehenen Raum $\mathcal{L}(X, Y)$ sei stetig. Dann ist die Abbildung F auf U Fréchet-differenzierbar und für alle $x \in U$ gilt*

$$F'_G(x) = F'_F(x).$$

Wir wollen weiterhin folgende – für den Verlauf der Arbeit grundlegende – Beispiele differenzierbarer Abbildungen und Funktionale zitieren:

Beispiel 1 (Affine Abbildungen [ITM09], S. 47). *Seien X und Y lineare Räume. Eine Abbildung $A : X \rightarrow Y$ wird affin genannt, wenn*

$$A(x) = \Lambda x + a$$

ist, wobei $a \in Y$ und Λ eine lineare Abbildung aus X in Y ist. Sind X und Y Banachräume und ist Λ eine lineare stetige Abbildung, so ist die Abbildung A überall Fréchet-differenzierbar und es ist

$$A'_F(x) = A'_G(x) = \Lambda.$$

Beispiel 2 (Quadratische Funktionen [ITM09], S. 48). Es seien X ein Banachraum, $B(x_1, x_2)$ eine stetige Bilinearform auf $X \times X$ und $Q(x) = B(x, x)$ die entsprechende quadratische Form. Laut Definition gilt

$$Q(x+h) = B(x, x) + B(x, h) + B(h, x) + B(h, h) = Q(x) + B(x, h) + B(h, x) + o(\|h\|).$$

Daher ist die Funktion Q Fréchet-differenzierbar:

$$Q'(x)h = B(x, h) + B(h, x).$$

Ist insbesondere X ein Hilbertraum, so hat jede quadratische Form die Gestalt

$$Q(x) = \frac{1}{2} \langle \Lambda x, x \rangle_X \quad \text{mit} \quad \Lambda \in \mathcal{L}(X, X), \Lambda^* = \Lambda$$

und es gilt

$$Q'_F(x) = Q'_G(x) = \Lambda x.$$

Als quadratisches Funktional im Hilbertraum bezeichnen wir die Summe aus einer quadratischen Form und einer affinen Funktion der Form

$$k(x) = \frac{1}{2} \langle \Lambda x, x \rangle_X + \langle x, a \rangle_X + \alpha.$$

Offensichtlich gilt

$$k'_F(x) = k'_G(x) = \Lambda x + a.$$

1.6 Asymptotische Steuerbarkeit dynamischer Systeme

Die Steuerungstheorie beschäftigt sich mit dynamischen Systemen, deren Verhalten durch sogenannte Steuerungsgrößen von außen beeinflusst werden können. Zwei wesentliche Aspekte dieses großen Teilgebietes sind dabei die Stabilitäts- und Steuerbarkeitsanalyse. Kern dieser Betrachtungen sind zum einen die Fragestellungen ob und wie ein gegebenes System durch vorliegende Steuermöglichkeiten in einen bestimmten Zustand überführt werden kann und zum anderen, ob und wie dieser erwünschte Zustand für den fortlaufenden Prozess stabil gehalten werden kann. Wir wollen in diesem Unterkapitel einen kurzen Auszug über den Zusammenhang der hier beschriebenen Problemklasse und den linearquadratischen Regulator-Modellen der Optimalsteuerung geben. Dafür orientieren wir uns an den Quellen [PB18] und [Bur19]. Die Inhalte dieses Unterkapitels sind daher als Auszug dieser Arbeiten zu verstehen.

Wir beginnen mit der Beschreibung der Stabilitätsanalyse im Sinne von Lyapunov²⁸ und betrachten dazu die folgende allgemeine nichtlineare autonome Systemdynamik

$$\mathcal{D}x(\cdot) = f(x(\cdot)) \tag{1.6}$$

zum gegebenen Anfangswert $x(0) = x_0$ mit vektorwertigem Zustand $x(t) \in \mathbb{R}^n$, der stetigen Abbildung $f : D \rightarrow \mathbb{R}^n$ und der offenen Menge $D \subset \mathbb{R}^n$, wobei $0 \in D$ sei. Im Folgenden bezeichne zudem x_G den sogenannten Gleichgewichtspunkt (bzw. Ruhelage) des Systems (1.6), d.h. den Punkt x_G für den gilt

$$0 = f(x_G).$$

Für derartige Gleichgewichtspunkte führen wir die folgende Definition ein:

²⁸vgl. [Lya92]

Definition 18 (vgl. z.B. [GJ08], S. 100, Definition 7.3).

1. Das Gleichgewicht x_G von (1.6) heißt stabil im Sinne von Lyapunov, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle Lösungen x von (1.6) mit $\|x(0) - x_G\| < \delta$ die Ungleichung $\|x(t) - x_G\| < \varepsilon$ für alle $t \geq 0$ gilt.
2. Das Gleichgewicht x_G von (1.6) heißt asymptotisch stabil im Sinne von Lyapunov, wenn es stabil im Sinne von Lyapunov ist und außerdem ein $\delta > 0$ mit der folgenden Eigenschaft existiert: Für jede Lösung x von (1.6) mit $\|x(0) - x_G\| < \delta$ gilt die Grenzwertbeziehung

$$\lim_{t \rightarrow \infty} \|x(t) - x_G\| = 0.$$

3. Das Gleichgewicht x_G von (1.6) heißt exponentiell stabil im Sinne von Lyapunov, falls Konstanten $\alpha, \beta > 0$ existieren, so dass für die Lösungen x von (1.6) mit $\|x(0) - x_G\| < \delta$ die Ungleichung

$$\|x(t) - x_G\| \leq \alpha \|x(0) - x_G\| e^{-\beta t} \quad \forall t \geq 0 \quad \text{erfüllt ist.}$$

In Abhängigkeit von der Existenz einer geeigneten Steuerungsfunktion lassen sich diese Begriffe auch auf den Fall gesteuerter Systeme übertragen. Dazu betrachten wir nun das allgemeine nichtlineare autonome System der Form

$$\mathcal{D}x(\cdot) = \tilde{f}(x(\cdot), u(\cdot)) \tag{1.7}$$

zum Anfangswert $x(0) = x_0$ mit dem vektorwertigen Zustand $x(t) \in \mathbb{R}^n$ und der vektorwertigen Steuerung $u(t) \in \mathbb{R}^m$. Dabei ist $\tilde{f} : \tilde{D} \rightarrow \mathbb{R}^n$ mit $\tilde{D} := D \times U$ eine auf \tilde{D} stetige Abbildung. Die Menge D sei dabei eine offene Teilmenge des \mathbb{R}^n und U eine abgeschlossene Teilmenge des \mathbb{R}^m . Sei nun x_G der Gleichgewichtspunkt des ungesteuerten Systems, d.h. der Punkt erfülle die Gleichung

$$0 = \tilde{f}(x_G, 0),$$

dann können wir die folgende Definition der Stabilisierbarkeit angeben:

Definition 19 (vgl. [Grü98], S. 1488, Definition 3.2).

1. Das Gleichgewicht x_G von (1.7) heißt stabilisierbar im Sinne von Lyapunov, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle Lösungen x von (1.7) mit $\|x_0 - x_G\| < \delta$ eine Steuerung $u_{x_0}(\cdot) \in U$ existiert, welche der Ungleichung

$$\|x(t, x_0, u_{x_0}(\cdot)) - x_G\| \leq \varepsilon \quad \forall t \geq 0 \quad \text{genügt.}$$

2. Das Gleichgewicht x_G von (1.7) heißt asymptotisch stabilisierbar im Sinne von Lyapunov, falls es ein $\delta > 0$ gibt, so dass für alle Lösungen x von (1.7) mit $\|x_0 - x_G\| < \delta$ eine Steuerung $u_{x_0}(\cdot) \in U$ existiert, so dass

$$\lim_{t \rightarrow \infty} \|x(t, x_0, u_{x_0}(\cdot)) - x_G\| = 0.$$

3. Das Gleichgewicht x_G von (1.7) heißt exponentiell stabilisierbar im Sinne von Lyapunov, falls es Konstanten $\alpha, \beta, \delta > 0$ gibt, so dass für die Lösungen x von (1.7) mit $\|x_0 - x_G\| < \delta$ eine Steuerung $u_{x_0}(\cdot) \in U$ existiert, die der Ungleichung

$$\|x(t, x_0, u_{x_0}(\cdot)) - x_G\| \leq \alpha \|x_0 - x_G\| e^{-\beta t} \quad \forall t \geq 0 \quad \text{genügt.}$$

Für die Berechnung einer stabilisierenden Steuerung nichtlinearer Systeme entwickelte Lyapunov die sogenannte *Direkte Methode*²⁹. Dazu führte er die sogenannte *Lyapunov-Funktion* ein. Basierend auf diesen Funktionen gab Lyapunov Kriterien an, mit denen sich Aussagen über die Stabilität des Gleichgewichtspunktes treffen lassen³⁰, ohne die tatsächliche Lösung des Differentialgleichungssystems auswerten zu müssen. Die Schwierigkeit dieser Methodik liegt an der Bestimmung der Lyapunov-Funktion selbst, da es kein allgemeines Konstruktionsschema gibt.

In den Publikationen [PB18, Bur19] motivieren die Autorinnen einen alternativen Ansatz durch die Einführung eines Optimalsteuerungsproblems über unendlichen Zeithorizont in gewichteten Funktionenräumen. Für diese Herangehensweise orientierten sie sich selbst an dem ursprünglichen Ansatz von Kalman³¹ aus dem Jahre 1960 und der Weiterentwicklung dieser Ideen durch Sontag³², welche Letztgenannten schließlich zu einer Konstruktion einer asymptotisch stabilisierenden Steuerung durch die Lösung des folgenden Optimalsteuerungsproblems über einen unendlichen Zeithorizont führte:

Minimiere

$$J_\infty(x, u) = L_{\lim} - \int_{t_0}^{\infty} L(x(t), u(t)) dt$$

unter Berücksichtigung der Restriktionen

$$\begin{aligned} (x, u) &\in AC_{loc}^n([t_0, \infty)) \times L_{1,loc}^m((t_0, \infty)), \\ x'(t) &= \tilde{f}(x(t), u(t)) \quad (\forall) t \geq t_0, \\ x(t_0) &= x_{t_0} \in \mathbb{R}^n, \\ u(t) &\in U \quad (\forall) t \geq t_0, \end{aligned}$$

wobei das Zielfunktional im Sinne des uneigentlichen Lebesgue-Integrals

$$L_{\lim} - \int_{t_0}^{\infty} \varphi(t) dt := \lim_{t_1 \rightarrow \infty} \int_{t_0}^{t_1} \varphi(t) dt$$

zu interpretieren ist. Problematisch an dieser Formulierung ist vor allen Dingen die Wahl der Funktionenräume in Form der lediglich lokal konvexen Hausdorff-Räume $AC_{loc}^n([t_0, \infty))$, welche Sontag durch die Erweiterung des durch Kalman zugrundegelegten Zustandsraumes

$$W_1^{1,n}([t_0, t_1]) = AC^n([t_0, t_1])$$

einführte. Die Wahl dieses Funktionenraumes erschwert die analytische Betrachtung der Aufgabenklasse enorm, so dass die in [PB18] motivierte Verwendung der gewichteten Funktionenräume $W_2^{1,n}((t_0, \infty), e^{\beta t})$ und $L_2^m((t_0, \infty), e^{\beta t})$ sowie die Festlegung des verwendeten Integraltyps in Form eines eigentlichen Lebesgue-Integrale im Sinne einer analytischen Auseinandersetzung praktikabler erscheint. Die so eingeführten Räume weisen aufgrund

²⁹Diese Methode wird auch die 2. Methode genannt. In einer ersten Methode traf Lyapunov Aussagen über lineare Systemgleichungen, vgl. hierzu [Lya92, Hah59].

³⁰vgl. [Hah59], S. 12 ff.

³¹vgl. [Kal60]

³²vgl. [Son83, Son98]; Für eine ausführlichere Darlegung siehe zum Beispiel [Bur19], S. 40 ff. oder [PB18].

ihrer Struktur in Form von Banachräumen wesentlich bessere Eigenschaften für die analytische Diskussion des Optimalsteuerungsproblems auf, als die durch Sontag eingeführten lokal konvexen Hausdorff-Räume.

Ausgehend von einer nichtlinearen Systemdynamik wird in [PB18] die Linearisierung der Dynamik um den ungesteuerten Gleichgewichtspunkt x_G vorgeschlagen. Eine anschließende lineare Transformation des Zustandes durch $\hat{x}(t) = x(t) - x_G$ für fast alle $t \geq t_0$ führt schließlich zu folgender Formulierung für die Bestimmung einer asymptotisch stabilisierenden Steuerung:

Minimiere

$$J_\infty(\hat{x}, u) = \frac{1}{2} \int_{t_0}^{\infty} \left\{ \hat{x}^T(t) W(t) \hat{x}(t) + u^T(t) R(t) u(t) \right\} e^{\beta t} dt$$

unter Berücksichtigung der Restriktionen

$$\begin{aligned} (\hat{x}, u) &\in W_2^{1,n}((t_0, \infty), e^{\beta t}) \times L_2^m((t_0, \infty), e^{\beta t}), \quad \beta > 0 \\ \hat{x}'(t) &= A(t)\hat{x}(t) + B(t)u(t) \quad (\forall) t \in (t_0, \infty), \\ \hat{x}(t_0) &= x_{t_0} - x_G \in \mathbb{R}^n, \\ u(t) &\in U \quad (\forall) t \in (t_0, \infty). \end{aligned}$$

Die Formulierung des Zielfunktional basiert dabei auf der Definition eines exponentiell stabilisierbaren Gleichgewichts³³. So erhalten wir für das exponentiell stabilisierbare Gleichgewicht x_G in einer hinreichend kleinen Umgebung $\|x(t_0) - x_G\| < \delta$ mit $\alpha, \beta, \delta > 0$ die Abschätzung

$$\begin{aligned} \int_{t_0}^{\infty} \|x(\tau, x_{t_0}, u_{x_{t_0}}(\cdot)) - x_G\|^2 e^{\beta t} dt &\leq \int_{t_0}^{\infty} (\alpha \|x(t_0) - x_G\| e^{-\beta t})^2 e^{\beta t} dt \\ &\leq \alpha^2 \|x(t_0) - x_G\|^2 \int_{t_0}^{\infty} e^{-\beta t} dt < \infty. \end{aligned}$$

Demnach sichert die derartige Modellierung des Zielfunktional für $\hat{x}(t) = x(t) - x_G$ (für fast alle $t \geq t_0$), die Endlichkeit des Zielfunktionalwertes für das exponentiell stabilisierbare Gleichgewicht x_G .

Für die Wahl der gewichteten Funktionenräume bauen die Ergebnisse aus [PB18] dabei insbesondere auf folgendem Resultat auf:

Lemma 4 ([PB18], S. 841, Lemma 2, Lemma 3).

1. Ist $x \in W_2^1(\mathbb{R}_{++}, \nu)$, $\nu(t) := e^{\beta t}$, $\beta > 0$, so folgt $x \in W_1^1(\mathbb{R}_{++})$.
2. Ist $x \in W_2^1(\mathbb{R}_{++}, \nu)$, $\nu(t) := e^{\beta t}$, $\beta > 0$ und $\|x\|_{1,(2),\nu} \leq 1$, dann gilt für alle $t \in \mathbb{R}_{++}$ die Ungleichung

$$|x(t)| \leq c\sqrt{t}e^{-\frac{\beta}{2}t}$$

mit einer Konstanten $c > 0$.

Wesentlich hierbei ist insbesondere die Folgerung, dass – unter den in Lemma 4 getätigten Voraussetzungen – für die Zustandsfunktionen der zugrundeliegenden Räume $\lim_{t \rightarrow \infty} x(t) = 0$ gültig ist und somit die exponentielle Stabilität im Sinne von Lyapunov für die Lösung $\hat{x} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ gewährleistet ist.

³³vgl. Definition 19, S. 20

Kapitel 2

Linear-quadratische Optimalsteuerungsprobleme mit unendlichem Zeithorizont

Ziel der Arbeit ist die dual-basierte Behandlung linear-quadratischer Optimalsteuerungsprobleme mit a priori gegebenem unendlichen Zeithorizont. Wesentlich dafür ist zunächst die Auseinandersetzung mit der zugrunde liegenden primalen Aufgabenstellung, insbesondere im Hinblick auf die Sicherung der Existenz einer optimalen Lösung und der Angabe von Optimalitätsbedingungen. Dabei werden wir feststellen, dass wir bisher nur für einen Teil der zu betrachtenden Probleme in der Lage sind (notwendige) Optimalitätsbedingungen zu formulieren. Dieser Umstand motiviert die spätere dual-basierte Behandlung.

In [\[Bur19\]](#) wird bereits eine ähnliche Problemklasse beschrieben. Zwecks Kontinuität lehnen wir uns in diesem Kapitel daher stark an die dortige Vorgehensweise an und erweitern die vorliegenden Ergebnisse auf den Fall von allgemeineren Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$ mit $\beta \neq 0$. Zudem betrachten wir im Gegensatz zu der Problemformulierung in [\[Bur19\]](#), in welcher lediglich die Abgeschlossenheit des Steuerbereiches gefordert ist, einen kompakten Steuerbereich \mathcal{U} . Die daraus entstehenden Unterschiede werden an entsprechender Stelle hervorgehoben.

2.1 Formulierung der Aufgabe (P_ν)

In diesem Unterkapitel soll die grundlegende Aufgabenstellung, auf welche sich alle weiteren Ergebnisse dieser Arbeit beziehen, vorgestellt werden. Wir betrachten dafür die folgende Aufgabenstellung (\mathbf{P}_ν):

Minimiere

$$J(x, u) = \int_0^\infty \{\varphi(t, x(t), u(t))\} \nu(t) dt \quad (2.1)$$

bezüglich aller Paare

$$(x, u) \in W_2^{1,n}(\mathbb{R}_{++}, \nu) \times L_2^m(\mathbb{R}_{++}, \nu), \quad \text{mit } \nu(t) = e^{\beta t}, \beta \neq 0, \quad (2.2)$$

für die gilt

$$x'(t) = A(t)x(t) + B(t)u(t) + C(t) \quad (\forall)t \in \mathbb{R}_{++}, \quad (2.3)$$

$$x(0) = x_0 \quad (2.4)$$

$$u(t) \in \mathcal{U} \subseteq \mathbb{R}^m \quad (\forall)t \in \mathbb{R}_{++}, \quad (2.5)$$

wobei der Integrand durch

$$\varphi(t, \xi, v) = \frac{1}{2}(\xi^T W(t)\xi + v^T R(t)v) + w^T(t)\xi + r^T(t)v, \quad (2.6)$$

mit den gegebenen Funktionen $W : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$, $R : \mathbb{R}_+ \rightarrow \mathbb{R}^{m \times m}$, $w : \mathbb{R}_+ \rightarrow \mathbb{R}^n$ und $r : \mathbb{R}_+ \rightarrow \mathbb{R}^m$ gegeben sei. Für die gegebenen Funktionen der rechten Seite der Differentialgleichung (2.3) gelte zudem $A : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$, $B : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times m}$ sowie $C : \mathbb{R}_+ \rightarrow \mathbb{R}^n$. Die so eingeführten Funktionen werden wir im Laufe der Betrachtungen noch genauer spezifizieren.

Das zu minimierende Zielfunktional J liegt in Form eines Integralfunktionals vor, das wir im Sinne von Lebesgue interpretieren. Damit gelangen wir zu folgendem Optimalitätsbegriff:

Definition 20. Die Menge aller zulässigen Paare (x, u) bezeichnen wir mit \mathcal{B} . Sie besteht aus allen Steuerungsprozessen (x, u) , welche den Bedingungen (2.2)–(2.5) genügen und für die das Lebesgue-Integral (2.1) existiert und endlich ist.

Seien die Steuerungsprozesse (x, u) und $(x_*, u_*) \in \mathcal{B}$ gegeben. Das Paar $(x_*, u_*) \in \mathcal{B}$ heißt global optimal für (P_ν) (**optimal nach Kriterium L**), falls für alle zulässigen Paare $(x, u) \in \mathcal{B}$ gilt

$$\int_0^\infty \varphi(t, x(t), u(t))\nu(t)dt - \int_0^\infty \varphi(t, x_*(t), u_*(t))\nu(t)dt \geq 0.$$

Für den Zustand x wählen wir den gewichteten Sobolev-Raum $W_2^{1,n}(\mathbb{R}_{++}, \nu)$ und für die Steuerungsfunktion u den gewichteten Lebesgue-Raum $L_2^m(\mathbb{R}_{++}, \nu)$.

Zudem wollen wir an dieser Stelle anmerken, dass das Gewicht im Zielfunktional (2.1) mit der Gewichtung der Funktionenräume übereinstimmt. Sollte sich in einer vorgegebenen Aufgabe das Gewicht des Zielfunktionals $\tilde{\nu}$ von der Gewichtung ν der Funktionenräume unterscheiden, so kann die Problemformulierung mittels der Transformation

$$\begin{aligned} J(x, u) &= \int_0^\infty \varphi(t, x(t), u(t))\tilde{\nu}(t)dt = \int_0^\infty \varphi(t, x(t), u(t))\tilde{\nu}(t)\frac{\nu(t)}{\nu(t)}dt \\ &= \int_0^\infty \tilde{\varphi}(t, x(t), u(t))\nu(t)dt, \end{aligned}$$

mit

$$\tilde{\varphi}(t, x(t), u(t)) := \varphi(t, x(t), u(t))\frac{\tilde{\nu}(t)}{\nu(t)}$$

in ein Problem der Klasse (P_ν) überführt werden. Damit stellt diese Annahme keine zusätzliche Restriktion dar.

Wir möchten im Rahmen dieses Kapitels die vorliegende Problemklasse hinsichtlich der Existenz und Eindeutigkeit einer optimalen Lösung untersuchen. Dabei werden wir die Methoden aus [Pic14, Pic15] übernehmen und auf den Fall eines kompakten Steuerbereiches ausweiten¹. In den zitierten Arbeiten wird die günstige Wahl der Funktionenräume in Form von gewichteten Sobolev- und Lebesgue-Räumen ausgenutzt, um die Aufgabenklasse in den Kontext der Hilbertraumtheorie einzubetten und entsprechend dieser Beweistechniken zu studieren.

In den Arbeiten [ZP18, Zie19] wurden zudem für eine allgemeine Klasse von Optimalsteuerungsproblemen mit a priori gegebenem unendlichen Zeithorizont und vektorwertigen Zustands- und Steuerungsgrößen, unter Verwendung einer Dichtefunktion, notwendige Optimalitätsbedingungen bewiesen. Im Unterkapitel 2.5 werden wir uns auf die Angabe der angepassten Voraussetzungen für die Formulierung notwendiger Bedingungen für linear-quadratische Probleme vom Typ (P_ν) beschränken. Das Studium der dualen Aufgabe wird uns letztlich eine Ausweitung auf den Fall von echten Gewichtsfunktionen ermöglichen. Die obig beschriebene Betrachtung der primalen Aufgabe (P_ν) setzt einige Grundannahmen voraus. Diese wollen wir im folgenden Unterkapitel angeben.

2.2 Voraussetzungen

Im Rahmen dieser Arbeit wollen wir nun die folgenden Voraussetzungen für das Problem (P_ν) postulieren:

(V1) Die Funktionen aus dem Zielfunktional (2.6) wählen wir gemäß

$$W \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), \quad W'_{i,j} \in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}) \quad i, j = 1, \dots, n,$$

$$R \in L_\infty^{m \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m}),$$

$$w \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}), \quad r \in L_2^m(\mathbb{R}_{++}, e^{\beta t}) \cap C(\mathbb{R}_+, \mathbb{R}^m)$$

und die Funktionen der Differentialgleichung (2.3) seien durch

$$A \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), \quad A'_{i,j} \in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}) \quad i, j = 1, \dots, n,$$

$$B \in L_\infty^{n \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times m}), \quad C \in L_\infty^n(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^n)$$

festgelegt, wobei $B(t) \neq 0$ auf ganz \mathbb{R}_+ erfüllt sei.

Gilt $\nu(t) = e^{\beta t}$, $\beta > 0$, dann sei zudem

$$C(t) \equiv 0 \quad \forall t \in \mathbb{R}_+.$$

¹Tatsächlich wurde in [Pic14, Pic15] der skalarwertige Fall $n = m = 1$ besprochen. [Pic14] lässt zunächst nur die Verwendung von Dichtefunktionen zu. In [Pic15] erfolgt die Erweiterung für Probleme mit echten Gewichtsfunktionen, allerdings nur für den Fall abgeschlossener, konvexer Steuerbereiche. Eine Erweiterung dieser Methoden auf den vektorwertigen Fall für Probleme mit abgeschlossenem, konvexem Steuerbereich und Dichtefunktionen lässt sich zum Beispiel in [Bur19] nachlesen.

(V2) Die Matrizen $W(t)$ und $R(t)$ seien symmetrisch und es gelte

$$\underline{W} > 0 \quad \text{und} \quad \underline{R} > 0 \quad \text{mit} \quad \underline{M} := \inf_{t \in \mathbb{R}_+} \lambda_{\min}(M)(t),$$

wobei $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M bezeichne. Das heißt W und R seien für alle $t \in \mathbb{R}_+$ symmetrisch und positiv definit.

(V3) $\mathcal{U} = [u, \bar{u}] \subseteq \mathbb{R}^m$ sei eine nichtleere, kompakte und konvexe Teilmenge des \mathbb{R}^m .

(V4) Wir bezeichnen die Menge aller Prozesse welche den Bedingungen (2.2)–(2.5) genügen mit \mathcal{B}_P und nehmen an, dass $\mathcal{B}_P \neq \emptyset$ und $\mathcal{B}_P \neq \{(x_*, u_*)\}$.

Bemerkung 9.

- Durch die Voraussetzungen (V1) und (V2) ist gewährleistet, dass das Zielfunktional (2.1) für jedes Paar $(x, u) \in W_2^{1,n}(\mathbb{R}_{++}, \nu) \times L_2^m(\mathbb{R}_{++}, \nu)$ einen endlichen Wert liefert. Damit liefert auch jedes Paar $(x, u) \in \mathcal{B}_P$ einen endlichen Zielfunktionalwert, so dass \mathcal{B}_P der Definition einer zulässigen Menge genügt². Damit können wir die Mengen \mathcal{B} und \mathcal{B}_P gleichsetzen.
- Für den Fall echter Gewichtsfunktionen ist zu beachten, dass Voraussetzung (V4) impliziert, dass $0 \in \mathcal{U}$. Würde dies nicht gelten, so gäbe es für fast alle $t \in \mathbb{R}_{++}$ eine Konstante c mit $|u(t)| \geq c > 0$ womit folgt

$$\int_0^\infty |u(t)|^2 e^{\beta t} dt \geq c^2 \int_0^\infty e^{\beta t} dt = \infty$$

und damit $u \notin L_2^m(\mathbb{R}_{++}, e^{\beta t})$ für $\beta > 0$.

Mit Hilfe der eingeführten Voraussetzungen (V1)–(V4), von deren Erfülltsein wir im Folgenden stets ausgehen, widmen wir uns nun der Analyse des linear-quadratischen Optimalsteuerungsproblems (P_ν) .

2.3 Formulierung als Optimierungsproblem im Hilbertraum

Für den Nachweis der Existenz und Eindeutigkeit einer Lösung für Probleme der Klasse (P_ν) wird in [Pic15] eine Einbettung des Optimalsteuerungsproblems in den Kontext der Hilbertraumtheorie vorgeschlagen. Wir übernehmen die in [Bur19] vorgeschlagene Notation und definieren den Raum

$$\mathcal{H} := W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^m(\mathbb{R}_{++}, e^{\beta t}),$$

welcher mittels der Norm

$$\|(x, u)\|_{\mathcal{H}} := \left(\|x\|_{1,(2),\nu}^2 + \|u\|_{(2),\nu}^2 \right)^{1/2}$$

²vgl. Definition 20, S. 24

zu einem Hilbertraum wird. Dabei bezeichne $\|\cdot\|_{1,(2),\nu}$ im folgenden Verlauf der Arbeit $\|\cdot\|_{W_2^{1,n}(\mathbb{R}_{++},\nu)}$. Mit $\|\cdot\|_{(2),\nu}$ bezeichnen wir $\|\cdot\|_{L_2^2(\mathbb{R}_{++},\nu)}$ (bzw. $\|\cdot\|_{L_2^m(\mathbb{R}_{++},\nu)}$ für Funktionen auf \mathbb{R}^m). Auf diesem Raum führen wir die Abbildung $Q : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ durch

$$Q((x_1, u_1), (x_2, u_2)) := \int_0^\infty \left\{ x_1^T(t)W(t)x_2(t) + u_1^T(t)R(t)u_2(t) \right\} e^{\beta t} dt \quad (2.7)$$

sowie die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$ durch

$$q(x, u) := \int_0^\infty \left\{ w^T(t)x(t) + r^T(t)u(t) \right\} e^{\beta t} dt \quad (2.8)$$

ein. Damit können wir das Zielfunktional darstellen mittels

$$J(x, u) = \frac{1}{2}Q((x, u), (x, u)) + q(x, u).$$

Analog zu den Betrachtungen in [Pic15, Bur19] können wir für diese Abbildungen einige nützliche Eigenschaften nachweisen, die wir aufgrund der Analogie an dieser Stelle jedoch nur zitieren und mit einem Abriss der Beweismethoden versehen wollen. Für detailliertere Ausführungen verweisen wir daher auf die entsprechenden Quellenangaben.

Für die Gültigkeit wesentlicher Eigenschaften an die Abbildung Q zitieren wir das folgende Lemma:

Lemma 5 ([Bur19] Lemma 4.4 S. 82³). *Durch die Abbildung $Q : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in (2.7), wird eine symmetrische Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt und damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Es lässt sich leicht nachrechnen, dass Q für alle $z_1, z_2, \hat{z}_1, \hat{z}_2 \in \mathcal{H}$ mit $z_1 = (x_1, u_1)$, $z_2 = (x_2, u_2)$, $\hat{z}_1 = (\hat{x}_1, \hat{u}_1)$, $\hat{z}_2 = (\hat{x}_2, \hat{u}_2)$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned} Q(z_1 + \hat{z}_1, z_2) &= Q(z_1, z_2) + Q(\hat{z}_1, z_2) \\ Q(z_1, z_2 + \hat{z}_2) &= Q(z_1, z_2) + Q(z_1, \hat{z}_2) \\ Q(\lambda z_1, z_2) &= \lambda Q(z_1, z_2). \end{aligned}$$

Aufgrund der Symmetrie der Matrizen W und R gilt zudem für alle $z_1, z_2 \in \mathcal{H}$

$$Q(z_1, z_2) = Q(z_2, z_1).$$

Damit folgt die Bilinearität des Funktional Q . Des Weiteren können wir aufgrund der Voraussetzung (V1) auf die Endlichkeit der Werte

$$\|W\|_\infty := \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |W(t)| \quad \text{und} \quad \|R\|_\infty := \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |R(t)|$$

³Die folgenden aus [Bur19] zitierten Lemmata basieren auf den skalarwertigen Ergebnissen in [Pic15], S. 174 ff.

schließen und erhalten mit diesen sowie mit den Symmetrieeigenschaften der Matrizen W und R , die folgende Abschätzung⁴ für alle $z_1, z_2 \in \mathcal{H}$:

$$\begin{aligned}
|Q(z_1, z_2)| &\leq \int_0^\infty \left| x_1^T(t)W(t)x_2(t) + u_1^T(t)R(t)u_2(t) \right| e^{\beta t} dt \\
&\leq \|W\|_\infty \int_0^\infty \left| x_1^T(t)x_2(t) \right| e^{\beta t} dt + \|R\|_\infty \int_0^\infty \left| u_1^T(t)u_2(t) \right| e^{\beta t} dt \\
&\leq \|W\|_\infty \left(\int_0^\infty |x_1(t)|^2 e^{\beta t} dt \right)^{1/2} \left(\int_0^\infty |x_2(t)|^2 e^{\beta t} dt \right)^{1/2} \\
&\quad + \|R\|_\infty \left(\int_0^\infty |u_1(t)|^2 e^{\beta t} dt \right)^{1/2} \left(\int_0^\infty |u_2(t)|^2 e^{\beta t} dt \right)^{1/2} \\
&\leq \|W\|_\infty \|x_1\|_{(2),e^{\beta t}} \|x_2\|_{(2),e^{\beta t}} + \|R\|_\infty \|u_1\|_{(2),e^{\beta t}} \|u_2\|_{(2),e^{\beta t}} \\
&\leq \|W\|_\infty \left\{ \|x_1\|_{(2),e^{\beta t}} + \|x_1'\|_{(2),e^{\beta t}} \right\} \left\{ \|x_2\|_{(2),e^{\beta t}} + \|x_2'\|_{(2),e^{\beta t}} \right\} \\
&\quad + \|R\|_\infty \|u_1\|_{(2),e^{\beta t}} \|u_2\|_{(2),e^{\beta t}} \\
&= \|W\|_\infty \|x_1\|_{1,(2),e^{\beta t}} \|x_2\|_{1,(2),e^{\beta t}} + \|R\|_\infty \|u_1\|_{(2),e^{\beta t}} \|u_2\|_{(2),e^{\beta t}} \\
&\leq c \left(\|x_1\|_{1,(2),e^{\beta t}}^2 \|x_2\|_{1,(2),e^{\beta t}}^2 + \|u_1\|_{(2),e^{\beta t}}^2 \|u_2\|_{(2),e^{\beta t}}^2 \right)^{1/2} \\
&\leq c \left(\|x_1\|_{1,(2),e^{\beta t}}^2 + \|u_1\|_{(2),e^{\beta t}}^2 \right)^{1/2} \left(\|x_2\|_{1,(2),e^{\beta t}}^2 + \|u_2\|_{(2),e^{\beta t}}^2 \right)^{1/2} \\
&\leq c \|(x_1, u_1)\|_{\mathcal{H}} \|(x_2, u_2)\|_{\mathcal{H}} = c \|z_1\|_{\mathcal{H}} \|z_2\|_{\mathcal{H}}.
\end{aligned}$$

Damit ist die Bilinearform Q beschränkt auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$. Da eine Bilinearform genau dann stetig bzgl. der Normtopologie des zugrundeliegenden normierten Raumes ist, wenn sie bzgl. dieser Norm beschränkt ist⁵, sind alle Behauptungen des Lemmas bewiesen. \square

Über Eigenschaften der Abbildung q gibt das nachfolgende Lemma Auskunft:

Lemma 6 ([Bur19] Lemma 4.5 S. 85). *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (2.8), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Offenbar ist der Operator q linear. Für den Nachweis der Beschränktheit betrachten wir für alle $z := (x, u) \in \mathcal{H}$ den Betrag $|q(z)|$ und gelangen aufgrund der Raumzugehörigkeiten der Funktionen w und r sowie mit Hilfe der Ungleichung von Cauchy-Schwarz zu der folgenden Abschätzung:

$$\begin{aligned}
|q(z)| &\leq \|w\|_{(2),e^{\beta t}} \|x\|_{(2),e^{\beta t}} + \|r\|_{(2),e^{\beta t}} \|u\|_{(2),e^{\beta t}} \\
&\leq c \left(\|x\|_{1,(2),e^{\beta t}}^2 + \|u\|_{(2),e^{\beta t}}^2 \right)^{1/2} = c \|z\|_{\mathcal{H}}
\end{aligned}$$

⁴Für die Abschätzung nutzen wir die Ungleichung von Cauchy-Schwarz, d.h. die Gültigkeit der Abschätzung $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle \forall x, y$ des Hilbertraumes X (vgl. hierfür beispielsweise [Wer07], S. 157, Satz V.1.2) sowie die Gültigkeit der Abschätzung $a + b \leq \sqrt{2} (a^2 + b^2)^{1/2}$ für alle $a, b \geq 0$ mit $a, b \in \mathbb{R}$.

⁵vgl. z.B. [ITM09], S. 39

□

Zur Beschreibung des Differentialgleichungssystems (2.3) führen wir den Raum

$$\mathcal{Y} : L_2^n(\mathbb{R}_{++}, e^{\beta t}) \times \mathbb{R}^n$$

ein, welcher mit der Norm

$$\|(w, v)\|_{\mathcal{Y}} := \left(\|w\|_{(2), e^{\beta t}}^2 + |v|^2 \right)^{1/2},$$

ebenfalls einen Hilbertraum bildet, wobei $|v| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$ bezeichne.

Auf diesem Raum führen wir mittels der Abbildung $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{Y}$ den Operator

$$\mathcal{A}(x, u) := \begin{pmatrix} A(\cdot)x(\cdot) + B(\cdot)u(\cdot) - x'(\cdot) \\ x(0) \end{pmatrix}, \quad (2.9)$$

sowie $\mathbf{b} \in \mathcal{Y}$ mit

$$\mathbf{b} := \begin{pmatrix} -C(\cdot) \\ x_0 \end{pmatrix}$$

ein. An dieser Stelle möchten wir daran erinnern, dass wir im Fall $\beta > 0$ von der Setzung $C(t) \equiv 0 \forall t \in \mathbb{R}_+$ ausgehen. Demnach betrachten wir im Falle einer echten Gewichtsfunktion tatsächlich

$$\mathbf{b} := \begin{pmatrix} 0 \\ x_0 \end{pmatrix}.$$

Die Eigenschaften des Operators \mathcal{A} bleiben von dieser Setzung unberührt, so dass auch hier die Zitation des nachfolgenden Lemmas genügt.

Lemma 7 ([Bur19] Lemma 4.6 S. 86). *Der Operator $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{Y}$, definiert in (2.9), ist ein linearer und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränkter Operator auf \mathcal{H} . Insbesondere ist \mathcal{A} demnach auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Die Linearität des Operators \mathcal{A} lässt sich leicht nachrechnen. Wir zeigen hier daher lediglich die Beschränktheit des Operators. Eine ausführlichere Darstellung findet sich in der angegebenen Quelle.

Wir erhalten⁶

$$\begin{aligned} \|\mathcal{A}(x, u)\|_{\mathcal{Y}}^2 &= \int_0^{\infty} (A(t)x(t) + B(t)u(t) - x'(t))^T (A(t)x(t) + B(t)u(t) - x'(t)) e^{\beta t} dt + |x(0)|^2 \\ &\leq 2 \int_0^{\infty} \left\{ (A(t)x(t) + B(t)u(t))^T (A(t)x(t) + B(t)u(t)) + [x'(t)]^T x'(t) \right\} e^{\beta t} dt + |x(0)|^2. \end{aligned}$$

Die wesentliche Beschränktheit von A und B gemäß (V1) liefert

$$\begin{aligned} \|\mathcal{A}(x, u)\|_{\mathcal{Y}}^2 &= |x(0)|^2 + c_1 \|x\|_{(2), e^{\beta t}}^2 + c_2 \|u\|_{(2), e^{\beta t}}^2 + c_3 \|x'\|_{(2), e^{\beta t}}^2 \\ &\leq c \left(\|x\|_{1, (2), e^{\beta t}}^2 + \|u\|_{(2), e^{\beta t}}^2 \right) = c \|(x, u)\|_{\mathcal{H}}^2. \end{aligned}$$

□

⁶Für diese Abschätzung nutzen wir die Gültigkeit der Ungleichung $(a + b)^2 \leq 2a^2 + 2b^2 \forall a, b \in \mathbb{R}$.

Mit Hilfe der eingeführten Operatoren und der Menge

$$\mathbf{U} := \left\{ u \in L_2^m(\mathbb{R}_{++}, e^{\beta t}) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_{++} \right\}$$

ist der zulässige Bereich gegeben durch

$$\mathcal{B}_P := \{(x, u) \in \mathcal{H} \mid u \in \mathbf{U}, \mathcal{A}(x, u) = \mathbf{b}\}.$$

Damit lässt sich das Problem (P_ν) äquivalent als quadratisches Optimierungsproblem im Hilbertraum \mathcal{H} in der Form

$$(P_\nu) \quad J(x, u) = \frac{1}{2}Q((x, u), (x, u)) + q(x, u) \rightarrow \inf!$$

bzgl. aller $(x, u) \in \mathcal{B}_P$

angeben.

2.4 Existenz und Eindeutigkeit einer optimalen Lösung

Für die Untersuchung der Existenz und Eindeutigkeit einer optimalen Lösung beziehen wir uns auf die folgende Formulierung einer verallgemeinerten Version des Satzes von Weierstraß in Banachräumen:

Satz 7 (Verallgemeinerter Satz von Weierstraß) (vgl. [Göp73], S. 93, Satz 26). *Sei Ω ein Banachraum und M eine nichtleere (schwach) kompakte Teilmenge von Ω . Des Weiteren sei das Funktional F auf M (schwach) unterhalbstetig. Dann besitzt die Aufgabe*

$$F(x) \rightarrow \text{Min!}$$

$$x \in M$$

eine optimale Lösung auf M , d.h. $\exists x_0 \in M : \text{Min} \{F(x), x \in M\} = F(x_0)$.

Zum Nachweis dieses Satzes ist somit sowohl die schwache Unterhalbstetigkeit des Zielfunktional J , als auch die schwache Kompaktheit der zulässigen Menge \mathcal{B}_P zu zeigen. Wir wollen daher an dieser Stelle eine gesonderte Betrachtung von Zielfunktional und zulässiger Menge durchführen.

Das Zielfunktional von (P_ν)

Die nach Voraussetzung (V2) geforderte positive Definitheit der Matrizen W und R sowie die linear-quadratische Struktur des Integranden φ liefern uns die strikte Konvexität des Integranden in ξ und v und erlauben uns somit unmittelbar auf dessen schwache Unterhalbstetigkeit zu schließen. Dazu zitieren wir wiederum aus [Bur19]:

Lemma 8 ([Bur19], S. 90). *Das Zielfunktional J aus (2.1) ist schwach unterhalbstetig auf \mathcal{H} .*

Beweis. Wir folgen der Beweisführung aus [Bur19] und betrachten dafür die schwach konvergenten Folgen $\{(x_n, u_n)\}_{n \in \mathbb{N}} \in \mathcal{H}$, d.h. Folgen $\{(x_n, u_n)\}_{n \in \mathbb{N}}$ mit

$$x_n \rightharpoonup x_0 \text{ in } W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}) \quad \text{und} \quad u_n \rightharpoonup u_0 \text{ in } L_2^m(\mathbb{R}_{++}, e^{\beta t}).$$

Dann folgt aus der Konvexität des Integranden in ξ und v sowie der Symmetrie der Matrizen W und R die Ungleichung

$$\begin{aligned}\varphi(t, x_n(t), u_n(t)) &= \frac{1}{2}x_n^T(t)W(t)x_n(t) + w^T(t)x_n(t) + \frac{1}{2}u_n^T(t)R(t)u_n(t) + r^T(t)u_n(t) \\ &\geq \frac{1}{2}\left\{x_0^T(t)W(t)x_0(t) + u_0^T(t)R(t)u_0(t)\right\} + w^T(t)x_n(t) + r^T(t)u_n(t) \\ &\quad + x_0^T(t)W(t)(x_n(t) - x_0(t)) + u_0^T(t)R(t)(u_n(t) - u_0(t)),\end{aligned}$$

so dass das Zielfunktional abgeschätzt werden kann mittels

$$\begin{aligned}\infty > J(x_n, u_n) &= \int_0^\infty \{\varphi(t, x_n(t), u_n(t))\} e^{\beta t} dt \\ &\geq \int_0^\infty \left\{ \frac{1}{2}x_0^T(t)W(t)x_0(t) + \frac{1}{2}u_0^T(t)R(t)u_0(t) + r^T(t)u_n(t) + w^T(t)x_n(t) \right\} e^{\beta t} dt \\ &\quad + \int_0^\infty \left\{ x_0^T(t)W(t)(x_n(t) - x_0(t)) + u_0^T(t)R(t)(u_n(t) - u_0(t)) \right\} e^{\beta t} dt.\end{aligned}$$

Gemäß der Voraussetzung (V1) können wir auf die Existenz der uneigentlichen Integrale der rechten Seite der obigen Ungleichung schließen⁷. Der Übergang zum Limes inferior liefert

$$\liminf_{n \rightarrow \infty} J(x_n, u_n) \geq l_1 + l_2,$$

wobei

$$\begin{aligned}l_1 &:= \liminf_{n \rightarrow \infty} \int_0^\infty \left\{ \frac{1}{2}x_0^T(t)W(t)x_0(t) + \frac{1}{2}u_0^T(t)R(t)u_0(t) + r^T(t)u_n(t) + w^T(t)x_n(t) \right\} e^{\beta t} dt \\ l_2 &:= \liminf_{n \rightarrow \infty} \int_0^\infty \left\{ x_0^T(t)W(t)(x_n(t) - x_0(t)) + u_0^T(t)R(t)(u_n(t) - u_0(t)) \right\} e^{\beta t} dt.\end{aligned}$$

Auf die Herleitung der Abschätzung der einzelnen Summanden verzichten wir an dieser Stelle und verweisen stattdessen auf die in [\[Lyk10\]](#), S. 73 ff. geführte Argumentation. Grundidee ist hierbei die Betrachtung derjenigen Teilfolgen $\{x_{n^k}\}_{n^k \in \mathbb{N}} \subseteq \{x_n\}_{n \in \mathbb{N}}$ und $\{u_{n^k}\}_{n^k \in \mathbb{N}} \subseteq \{u_n\}_{n \in \mathbb{N}}$, die den entsprechenden Limes inferior in l_1 bzw. l_2 realisieren. Für den Ausdruck l_1 gelingt somit der Nachweis der Gültigkeit der Ungleichung

$$\begin{aligned}l_1 &\geq \int_0^\infty \left\{ \frac{1}{2}x_0^T(t)W(t)x_0(t) + \frac{1}{2}u_0^T(t)R(t)u_0(t) + r^T(t)u_0(t) + w^T(t)x_0(t) \right\} e^{\beta t} dt \\ &= J(x_0, u_0).\end{aligned}$$

Für den zweiten Ausdruck ergibt sich $l_2 = 0$ und somit insgesamt

$$\liminf_{n \rightarrow \infty} J(x_n, u_n) \geq l_1 + l_2 \geq J(x_0, u_0)$$

für jede schwach konvergente Folge $\{(x_n, u_n)\}_{n \in \mathbb{N}} \in \mathcal{H}$ mit $\{(x_n, u_n)\} \xrightarrow{\mathcal{H}} (x_0, u_0) \in \mathcal{H}$. Damit ist das Zielfunktional J schwach unterhalbstetig auf \mathcal{H} . \square

⁷vgl. Lemma [5](#), S. [27](#) und Lemma [6](#), S. [28](#)

Der zulässige Bereich \mathcal{B}_P für Probleme mit Dichtefunktionen

Wir möchten nun eine gesonderte Diskussion für Problemklassen mit Dichtefunktionen und Problemklassen mit echten Gewichtsfunktionen führen und beginnen mit der Betrachtung von Problemklassen mit Dichtefunktionen:

Lemma 9. *Seien die Voraussetzungen (V1), (V3) und (V4) erfüllt und $\beta < 0$, dann ist der zulässige Bereich \mathcal{B}_P von (P_ν) nichtleer, konvex und schwach kompakt.*

Beweis. Die Voraussetzung (V4) garantiert $\mathcal{B}_P \neq \emptyset$. Den übrigen Beweis führen wir in drei Schritten. Dafür verwenden wir im wesentlichen die Aussagen von Satz 1 und Satz 2.

Schritt 1: Wir zeigen die Konvexität der Menge \mathcal{B}_P : Diese folgt leicht aus der Konvexität der Menge \mathcal{U} und der Linearität der Dynamik.

Schritt 2: Wir zeigen die schwache Abgeschlossenheit der Menge \mathcal{B}_P : Hierfür bedienen wir uns der in Schritt 1 gesicherten Konvexität der Menge \mathcal{B}_P sowie der Aussagen des Satzes 2, gemäß derer wir aus der (starken) Abgeschlossenheit konvexer Teilmengen M reflexiver Banachräume auf die schwache Abgeschlossenheit der Teilmenge M schließen können. Dazu beginnen wir mit dem Nachweis der (starken) Abgeschlossenheit von \mathcal{B}_P bezüglich der Menge \mathbf{U} , um im Anschluss den Nachweis der (starken) Abgeschlossenheit der gesamten Menge \mathcal{B}_P zu vollziehen. Unter Ausnutzung von Satz 2 erhalten wir durch die in Schritt 1 nachgewiesene Konvexität der Menge \mathcal{B}_P letztlich den Schluss auf die schwache Abgeschlossenheit der Menge \mathcal{B}_P .

Wir beginnen also mit dem Nachweis der starken Abgeschlossenheit von \mathcal{B}_P bzgl. \mathbf{U} . Für die Menge \mathbf{U} können wir das folgende Lemma angeben⁸:

Lemma 10. *Sei eine Funktionenfolge $\{u_n\}_{n \in \mathbb{N}}$ mit Elementen aus dem Funktionenraum $L_2^m(\mathbb{R}_{++}, e^{\beta t})$, $\beta \neq 0$ konvergent gegen die Funktion u_0 bezüglich des Raumes $L_2^m(\mathbb{R}_{++}, e^{\beta t})$. Weiter sei die Inklusion*

$$u_n(t) \in \mathcal{U} \quad (\forall) t \in \mathbb{R}_{++}$$

erfüllt für jede Zahl $n \in \mathbb{N}$. Dabei bezeichnet \mathcal{U} eine nichtleere kompakte konvexe Teilmenge des \mathbb{R}^m . Dann gilt die Inklusion

$$u_0(t) \in \mathcal{U}$$

fast überall auf \mathbb{R}_{++} .

Beweis. Das Lemma ist im Prinzip eine direkte Folgerung aus dem folgenden Korollar:

Korollar 1 (H. Weyl⁹ 1909, [Els07], S. 231). *Es sei $1 \leq p \leq \infty$, $\Omega \subseteq \mathbb{R}$ und ν ein σ -endliches Maß.*

- Zu jeder Cauchy-Folge $\{f_n\}_{n \geq 1}$ in $L_p(\Omega, \nu)$ gibt es eine Teilfolge $\{f_{n_k}\}_{k \geq 1}$ und ein $f \in L_p(\Omega, \nu)$, so dass $f_{n_k} \rightarrow f$ ν -f.ü.*
- Konvergiert die Folge $\{f_n\}_{n \geq 1}$ in $L_p(\Omega, \nu)$ gegen $f \in L_p(\Omega, \nu)$, so existiert eine Teilfolge $\{f_{n_k}\}_{k \geq 1}$, die ν -f.ü. gegen f konvergiert.*

⁸Kernpunkt der Aussage des Lemmas ist, dass die Konvergenz im Raum $L_2^m(\mathbb{R}_{++}, e^{\beta t})$ die punktweise Konvergenz f.ü. nach sich zieht.

⁹H. Weyl: Über die Konvergenz von Reihen, die nach Orthogonalfunktionen fortschreiten, Math. Ann. 67, 225-245 (1909) (=Gesammelte Abhandlungen I, S. 154-174).

Damit genügt es, die Anwendbarkeit des Satzes – und somit die σ -Endlichkeit unseres Maßes $\nu(t) = e^{\beta t}$, $\beta \neq 0$ – zu überprüfen. Dazu wiederum betrachten wir das nachstehende Lemma:

Lemma 11 ([Eis07], Lemma 5.3, S. 59). *Ein Inhalt $\nu : \mathfrak{h} \rightarrow \overline{\mathbb{R}}$ auf dem Halbring \mathfrak{h} über X ist genau dann σ -endlich, wenn es eine Folge $\{A_n\}_{n \geq 1}$ disjunkter Mengen aus \mathfrak{h} gibt mit $\nu(A_n) < \infty$ ($n \in \mathbb{N}$) und $\cup_{n=1}^{\infty} A_n = X$.*

Wählen wir $\mathfrak{h} = \mathbb{R}_{++}$ und betrachten als Folge $\{A_n\}_{n \geq 1}$ eine Folge disjunkter endlicher Intervalle, so sehen wir leicht ein, dass die Forderung der σ -Endlichkeit für unser betrachtetes Maß gültig und das Korollar [1] somit anwendbar ist. \square

Sei nun $\{(x_n, u_n)\}_{n \in \mathbb{N}}$ eine konvergente Folge aus \mathcal{B}_P mit $\{(x_n, u_n)\} \xrightarrow{\mathcal{H}} (x_0, u_0)$, mit dem Grenzelement $(x_0, u_0) \in W_2^{1,n}(\mathbb{R}_{++}, \nu) \times L_2^m(\mathbb{R}_{++}, \nu)$. Die Stetigkeit des Operators \mathcal{A} bzgl. der Normtopologie [10] des Raumes \mathcal{H} liefert dann

$$\lim_{n \rightarrow \infty} \mathcal{A}(x_n, u_n) = \mathcal{A}\left(\lim_{n \rightarrow \infty} x_n, \lim_{n \rightarrow \infty} u_n\right) = \mathcal{A}(x_0, u_0)$$

und damit $\mathcal{A}(x_0, u_0) = \mathbf{b}$ und somit $(x_0, u_0) \in \mathcal{B}_P$. Zusammengefasst erhalten wir damit also die (starke) Abgeschlossenheit der Menge \mathcal{B}_P .

Die in Schritt 1 nachgewiesenen Konvexität der Menge \mathcal{B}_P erlaubt uns nun die Anwendung des Satzes [21] und liefert damit den Schluss auf die schwache Abgeschlossenheit der Menge \mathcal{B}_P .

Schritt 3: Wir widmen uns nun dem Nachweis der Beschränktheit der Menge \mathcal{B}_P . Hier können wir wiederum auf die Ergebnisse aus [Lyk10], S. 91 zurückgreifen und erhalten in leicht abgewandelter Form:

Die Menge aller zulässigen Steuerungen ist beschränkt, da fast überall auf \mathbb{R}_{++} die Inklusion $u(t) \in \mathcal{U}$ gilt, wobei \mathcal{U} gemäß Voraussetzung (V3) eine nichtleere kompakte und konvexe Teilmenge des \mathbb{R}^m ist. Daher existiert eine Konstante $K > 0$, mit der die Ungleichung

$$|u(t)| \leq K \tag{2.10}$$

für fast alle $t \in \mathbb{R}_{++}$ gültig bleibt. Dies impliziert

$$\|u\|_{(2), e^{\beta t}}^2 \leq c_1, \quad \text{mit } c_1 \in \mathbb{R}_{++}. \tag{2.11}$$

Ebenso können wir mit der Gültigkeit der Ungleichung (2.10) die Ideen der sogenannten „natürlichen Zustandsbeschränkungen“ [22] anwenden, so dass offenbar auch die Ungleichung

$$|x(t)| \leq \alpha(t) \quad (\forall) t \in \mathbb{R}_{++},$$

mit $\alpha(t) = x_0 e^{Mt}$ und $M := \operatorname{ess\,sup}_{t \in \mathbb{R}_{++}, v \in U} (A(t) + B(t)v)$, $\alpha \in L_2^n(\mathbb{R}_{++}, e^{\beta t})$ gültig bleibt und damit

$$\|x\|_{(2), e^{\beta t}}^2 = \int_0^{\infty} |x(t)|^2 e^{\beta t} dt \leq \int_0^{\infty} |\alpha(t)|^2 e^{\beta t} dt = \|\alpha\|_{(2), e^{\beta t}}^2 =: c_2. \tag{2.12}$$

¹⁰vgl. Lemma [7], S. [29]

¹¹siehe S. [9]

¹²vgl. [Wen20], Lemma 2.1.1, S. 45

Die Untersuchung der Beschränktheit¹³ des Operators \mathcal{A} lieferte die Ungleichung

$$\|\mathcal{A}(x, u)\|_{\mathcal{Y}}^2 \leq c \left(\|x\|_{1,(2),e^{\beta t}}^2 + \|u\|_{(2),e^{\beta t}}^2 \right).$$

Für diese können wir die obigen Abschätzungen einsetzen und erhalten

$$\|\mathcal{A}(x, u)\|_{\mathcal{Y}}^2 \leq c(c_1 \cdot c_2) = \text{const.}$$

Auf diese Weise ist die Beschränktheit der Menge \mathcal{B}_P bewiesen. Die Gültigkeit von Satz 1¹ erlaubt uns nun, aus der nachgewiesenen schwachen Abgeschlossenheit und der Beschränktheit der Menge \mathcal{B}_P auf die schwache Kompaktheit der Menge \mathcal{B}_P zu schließen¹⁴. Der Beweis des Satzes ist somit ebenfalls abgeschlossen. \square

Damit können wir nun die nachstehende Aussage formulieren:

Satz 8. *Unter der Gültigkeit der Voraussetzungen (V1)-(V4) besitzt das Problem (P_ν) für $\beta < 0$ eine eindeutige Lösung.*

Beweis. Die Existenz einer Lösung folgt gemäß Satz 7⁷ unmittelbar aus der schwachen Unterhalbstetigkeit des Zielfunktional J sowie der schwachen Kompaktheit des nichtleeren zulässigen Bereiches \mathcal{B}_P . Die Eindeutigkeit folgt dann aus der strikten Konvexität des Zielfunktional J . Eine ausführliche Darstellung dieser Argumentation findet sich in [Bur19], S. 97. \square

Der zulässige Bereich \mathcal{B}_P für Probleme mit echten Gewichtsfunktionen

Wir betrachten nun den Fall von echten Gewichtsfunktionen und weisen dafür zunächst die folgende Eigenschaft nach:

Lemma 12. *Seien die Voraussetzungen (V1), (V3) und (V4) erfüllt und $\beta > 0$, dann ist der zulässige Bereich \mathcal{B}_P von (P_ν) nichtleer, konvex und schwach abgeschlossen.*

Beweis. Die Voraussetzung (V4) sichert $\mathcal{B}_P \neq \emptyset$. Die übrigen Behauptungen folgen unmittelbar aus der Beweisführung in Schritt 1 und 2 von Lemma 9⁹. \square

Für den Schluss auf die Existenz einer optimalen Lösung fehlt die Beschränktheit der Menge \mathcal{B}_P ¹⁵. Um die fehlende Beschränktheit dieser Menge zu kompensieren, gehen wir einen Umweg über die Formulierung von Niveaumengen der restringierten Aufgabe (P_ν) ¹⁶. In Vorbereitung dieser Diskussion stellen wir den Nachweis einiger weiterer Eigenschaften der Bilinearform Q voran. Dabei gehen wir auch hier stets von der Gültigkeit der Voraussetzungen (V1)-(V4) aus:

Lemma 13. *Die Bilinearform Q ist auf dem Kern der Abbildung \mathcal{A} koerzitiv.*

Beweis. Die Bilinearform Q heißt koerzitiv, falls für alle Elemente $z \in \ker \mathcal{A}$ eine positive Konstante c existiert, so dass $Q(z, z) \geq c \|z\|_{\mathcal{H}}^2$ erfüllt ist. Der Kern der Abbildung \mathcal{A} ist definiert durch

$$\ker \mathcal{A} := \{(x, u) \in \mathcal{H} \mid \mathcal{A}(x, u) = 0\}.$$

¹³vgl. Lemma 7⁷, S. 29

¹⁴vgl. Satz 1¹, S. 9

¹⁵Aufgrund der Verwendung der echten Gewichtsfunktion $\nu(t) = e^{\beta t}$ mit $\beta > 0$ können wir der Argumentation aus Schritt 3 unter Ausnutzung der Relationen (2.11) und (2.12) nicht folgen.

¹⁶vgl. hierfür die Ideen von Punkt 2 auf S. 7⁷

Analog zu den Betrachtungen in [Pic15, Bur19] untersuchen wir ein beliebiges Element $z := (\zeta, \vartheta) \in \ker \mathcal{A}$ und unterscheiden zwei Fälle:

Fall 1: Sei $A \equiv 0$. Dann ist $\zeta'(t) = B(t)\vartheta(t)$, $\zeta(0) = 0$ und es gilt¹⁷

$$\|\zeta'\|_{(2),e^{\beta t}}^2 = \int_0^\infty |B(t)\vartheta(t)|^2 e^{\beta t} dt \leq \|B\|_\infty^2 \int_0^\infty |\vartheta(t)|^2 e^{\beta t} dt = \|B\|_\infty^2 \|\vartheta\|_{(2),e^{\beta t}}^2$$

und damit

$$\|\vartheta\|_{(2),e^{\beta t}}^2 \geq \frac{1}{\|B\|_\infty^2} \|\zeta'\|_{(2),e^{\beta t}}^2.$$

Damit können wir, unter Gültigkeit der Voraussetzungen (V1)-(V4), die Bilinearform weiter abschätzen durch

$$\begin{aligned} Q((\zeta, \vartheta), (\zeta, \vartheta)) &= \int_0^\infty \zeta^T(t)W(t)\zeta(t) + \vartheta^T(t)R(t)\vartheta(t)e^{\beta t} dt \\ &\geq \underline{W} \int_0^\infty |\zeta(t)|^2 e^{\beta t} dt + \underline{R} \int_0^\infty |\vartheta(t)|^2 e^{\beta t} dt \\ &= \underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 \\ &\geq \underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 + \frac{\underline{R}}{2\|B\|_\infty^2} \|\zeta'\|_{(2),e^{\beta t}}^2 \\ &\geq c_1 \left\{ \|\zeta\|_{1,(2),e^{\beta t}}^2 + \|\vartheta\|_{(2),e^{\beta t}}^2 \right\} = c_1 \|z\|_{\mathcal{H}}^2, \end{aligned}$$

mit $c_1 := \min \left\{ \underline{W}, \frac{1}{2}\underline{R}, \frac{\underline{R}}{2\|B\|_\infty^2} \right\} > 0$.

Fall 2: Sei nun $A \not\equiv 0$. Dann gilt $\zeta'(t) = A(t)\zeta(t) + B(t)\vartheta(t)$, $\zeta(0) = 0$ auf $\ker \mathcal{A}$. Damit folgt

$$\begin{aligned} \|\zeta'\|_{(2),e^{\beta t}}^2 &= \int_0^\infty |A(t)\zeta(t) + B(t)\vartheta(t)|^2 e^{\beta t} dt \\ &\leq 2\|A\|_\infty^2 \int_0^\infty |\zeta(t)|^2 e^{\beta t} dt + 2\|B\|_\infty^2 \int_0^\infty |\vartheta(t)|^2 e^{\beta t} dt \\ &= 2\|A\|_\infty^2 \|\zeta\|_{(2),e^{\beta t}}^2 + 2\|B\|_\infty^2 \|\vartheta\|_{(2),e^{\beta t}}^2 \\ &\leq c_2 \left(\|\zeta\|_{(2),e^{\beta t}}^2 + \|\vartheta\|_{(2),e^{\beta t}}^2 \right), \end{aligned}$$

mit $c_2 := 2 \max \left\{ \|A\|_\infty^2, \|B\|_\infty^2 \right\} > 0$ und damit

$$\|\vartheta\|_{(2),e^{\beta t}}^2 \geq \frac{1}{c_2} \|\zeta'\|_{(2),e^{\beta t}}^2 - \|\zeta\|_{(2),e^{\beta t}}^2, \text{ bzw.} \quad (2.13)$$

$$\|\zeta\|_{(2),e^{\beta t}}^2 \geq \frac{1}{c_2} \|\zeta'\|_{(2),e^{\beta t}}^2 - \|\vartheta\|_{(2),e^{\beta t}}^2. \quad (2.14)$$

¹⁷Aufgrund der in (V1) geforderten Gültigkeit von $B(t) \not\equiv 0$ gilt $\|B\|_\infty > 0$.

Wir betrachten zunächst die obere Abschätzung (2.13) an $\|\vartheta\|_{(2),e^{\beta t}}^2$ und erhalten damit

$$\begin{aligned}
Q((\zeta, \vartheta), (\zeta, \vartheta)) &= \int_0^\infty \left\{ \zeta^T(t)W(t)\zeta(t) + \vartheta^T(t)R(t)\vartheta(t) \right\} e^{\beta t} dt \\
&\geq \underline{W} \int_0^\infty |\zeta(t)|^2 e^{\beta t} dt + \underline{R} \int_0^\infty |\vartheta(t)|^2 e^{\beta t} dt \\
&= \underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 \\
&\geq \underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{R} \left(\frac{1}{c_2} \|\zeta'\|_{(2),e^{\beta t}}^2 - \|\zeta\|_{(2),e^{\beta t}}^2 \right) \\
&= \left(\underline{W} - \frac{1}{2}\underline{R} \right) \|\zeta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{R}\frac{1}{c_2} \|\zeta'\|_{(2),e^{\beta t}}^2 \\
&\geq c_3 \left\{ \|\zeta\|_{1,(2),e^{\beta t}}^2 + \|\vartheta\|_{(2),e^{\beta t}}^2 \right\} = c_3 \|z\|_{\mathcal{H}}^2,
\end{aligned}$$

mit $c_3 := \min \left\{ \left(\underline{W} - \frac{1}{2}\underline{R} \right), \frac{1}{2}\underline{R}, \frac{1}{2c_2}\underline{R} \right\} > 0$ für $\left(\underline{W} - \frac{1}{2}\underline{R} \right) > 0$.

Betrachten wir nun die Abschätzung (2.14) an $\|\zeta\|_{(2),e^{\beta t}}^2$ dann erhalten wir mit einer ähnlichen Argumentation

$$\begin{aligned}
Q((\zeta, \vartheta), (\zeta, \vartheta)) &= \int_0^\infty \left\{ \zeta^T(t)W(t)\zeta(t) + \vartheta^T(t)R(t)\vartheta(t) \right\} e^{\beta t} dt \\
&\geq \underline{W} \int_0^\infty |\zeta(t)|^2 e^{\beta t} dt + \underline{R} \int_0^\infty |\vartheta(t)|^2 e^{\beta t} dt \\
&= \underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 \\
&\geq \frac{1}{2}\underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \frac{1}{2}\underline{W} \left(\frac{1}{c_2} \|\zeta'\|_{(2),e^{\beta t}}^2 - \|\vartheta\|_{(2),e^{\beta t}}^2 \right) + \underline{R} \|\vartheta\|_{(2),e^{\beta t}}^2 \\
&= \frac{1}{2}\underline{W} \|\zeta\|_{(2),e^{\beta t}}^2 + \frac{1}{2c_2}\underline{W} \|\zeta'\|_{(2),e^{\beta t}}^2 + \left(\underline{R} - \frac{1}{2}\underline{W} \right) \|\vartheta\|_{(2),e^{\beta t}}^2 \\
&\geq c_4 \left\{ \|\zeta\|_{1,(2),e^{\beta t}}^2 + \|\vartheta\|_{(2),e^{\beta t}}^2 \right\} = c_4 \|z\|_{\mathcal{H}}^2,
\end{aligned}$$

mit $c_4 := \min \left\{ \left(\underline{R} - \frac{1}{2}\underline{W} \right), \frac{1}{2}\underline{W}, \frac{1}{2c_2}\underline{W} \right\} > 0$ für $\left(\underline{R} - \frac{1}{2}\underline{W} \right) > 0$.

Insgesamt konnten wir damit die Koerzitivität der Bilinearform Q auf dem Kern der Abbildung für alle Matrixfunktionen W und R zeigen, für die gilt

$$\left(\underline{W} - \frac{1}{2}\underline{R} \right) > 0 \quad \text{oder} \quad \left(\underline{R} - \frac{1}{2}\underline{W} \right) > 0. \quad (2.15)$$

Wir wollen nachweisen, dass die Bedingung (2.15) für jede Wahl positiver definiten Matrixfunktionen W und R erfüllt ist. Dafür zeigen wir:

- 1. Fall:** Gilt $\underline{R} \geq \underline{W}$, dann impliziert dies die Gültigkeit der Abschätzung $\underline{R} > \frac{1}{2}\underline{W} > 0$ und liefert somit die Gültigkeit der Ungleichung $\underline{R} - \frac{1}{2}\underline{W} > 0$.
- 2. Fall:** Sei nun $\underline{R} < \underline{W}$, dann gilt insbesondere $\underline{W} > \frac{1}{2}\underline{R}$. Dies impliziert sofort die Gültigkeit der Ungleichung $\underline{W} - \frac{1}{2}\underline{R} > 0$.

Damit ist die Gleichung (2.15) für jede Wahl positiv definiten Matrixfunktionen W und R erfüllt und die Abbildung Q unter Gültigkeit der Voraussetzungen (V1)-(V4) koerzitiv auf $\ker \mathcal{A}$. \square

Bemerkung 10. Die zweifache Abschätzung unter Ausnutzung der Relationen (2.13) und (2.14) erlaubt uns den Nachweis des Lemmas ohne die Hinzunahme von Koerzitivitätsbedingungen. Im Vergleich zu den Ergebnissen aus [Pic15, Bur19] ermöglicht dies eine wesentlich schwächere Formulierung von Voraussetzungen für die Sicherung der Existenz einer optimalen Lösung.

Weiterhin gilt für Probleme mit echten Gewichtsfunktionen die Aussage des nachstehenden Lemmas:

Lemma 14. Ist $\beta > 0$, dann ist die Bilinearform Q auf dem Urbildbereich der Abbildung \mathcal{A} koerzitiv.

Beweis. Für den Fall $\beta > 0$ ist der Urbildbereich der Abbildung \mathcal{A} gegeben durch

$$\mathcal{A}^{-1}(\mathbf{b}) := \{(x, u) \in \mathcal{H} \mid \mathcal{A}(x, u) = \mathbf{b}\}, \quad \text{mit } \mathbf{b} = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}.$$

Damit können wir unter Verwendung von $\zeta(0) = x_0$ die gleichen Abschätzungen aus der Beweisführung von Lemma 13 treffen und erhalten damit

$$Q(z, z) \geq c \|z\|_{\mathcal{H}}^2,$$

mit $c > 0$ für alle $z = (x, u) \in \mathcal{A}^{-1}(\mathbf{b})$. \square

Lemma 15. Ist die Bilinearform Q auf dem Kern der Abbildung \mathcal{A} koerzitiv, so ist das Funktional J aus (2.1) strikt konvex auf der zulässigen Menge \mathcal{B}_P .

Beweis. Wir betrachten den Operator Q sowie beliebige Elemente $z_1, z_2 \in \mathcal{B}_P$ mit $z_1 \neq z_2$. Durch diese Setzung gilt $z_1 - z_2 \in \ker \mathcal{A}$. Aufgrund der Koerzitivität der symmetrischen Bilinearform Q auf dem Kern der Abbildung \mathcal{A} existiert eine Konstante $c > 0$, so dass

$$0 < c \|z_1 - z_2\|_{\mathcal{H}}^2 \leq Q(z_1 - z_2, z_1 - z_2) = Q(z_1, z_1) - 2Q(z_1, z_2) + Q(z_2, z_2)$$

gilt und damit

$$2Q(z_1, z_2) < Q(z_1, z_1) + Q(z_2, z_2).$$

Die Multiplikation der letzten Gleichung mit $\lambda(1 - \lambda)$, wobei $\lambda \in (0, 1)$ sei, liefert dann

$$2\lambda(1 - \lambda)Q(z_1, z_2) < \lambda(1 - \lambda)Q(z_1, z_1) + \lambda(1 - \lambda)Q(z_2, z_2).$$

Damit können wir für alle $z_1, z_2 \in \mathcal{B}_P$ die folgende Abschätzung finden:

$$\begin{aligned} Q((1 - \lambda)z_1 + \lambda z_2, (1 - \lambda)z_1 + \lambda z_2) &= (1 - \lambda)^2 Q(z_1, z_1) + 2\lambda(1 - \lambda)Q(z_1, z_2) + \lambda^2 Q(z_2, z_2) \\ &< (1 - \lambda)^2 Q(z_1, z_1) + \lambda(1 - \lambda)Q(z_1, z_1) + \lambda(1 - \lambda)Q(z_2, z_2) + \lambda^2 Q(z_2, z_2) \\ &= (1 - \lambda)Q(z_1, z_1) + \lambda Q(z_2, z_2). \end{aligned}$$

Da q linear ist erhalten wir durch $J(z) = \frac{1}{2}Q(z, z) + q(z)$ die strikte Konvexität des Zielfunktionals auf \mathcal{B}_P . \square

Wir betrachten damit ein strikt konvexes Funktional über der konvexen zulässigen Menge \mathcal{B}_P . Dies erlaubt uns die Verwendung der klassischen Methoden der konvexen Analysis.

Die Niveaumengen $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$

Für $\alpha \in \mathbb{R}$ sind die *Niveaumengen von J* durch

$$\mathcal{N}(J, \alpha) := \{(x, u) \in \mathcal{H} \mid J(x, u) \leq \alpha\}$$

definiert. Um die Nebenbedingungen der Problemklasse (P_ν) einzuarbeiten, betrachten wir anstelle dieser die Schnittmengen

$$\mathcal{N}_{\mathcal{B}_P}(J, \alpha) := \mathcal{N}(J, \alpha) \cap \mathcal{B}_P = \{(x, u) \in \mathcal{B}_P \mid J(x, u) \leq \alpha\},$$

mit $\alpha > 0$ und damit die *Niveaumengen der restringierten Optimierungsaufgabe (P_ν)* . Für den Nachweis der Existenz einer Lösung von (P_ν) zeigen wir zunächst die schwache Kompaktheit der Niveaumengen $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ der restringierten Aufgabe (P_ν) . Wir weisen dafür die Gültigkeit des nachfolgenden Lemmas nach:

Lemma 16. *Für jedes zulässige Paar $(x, u) \in \mathcal{B}_P$ ist die Niveaumenge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ konvex, abgeschlossen und beschränkt (bezüglich der Normtopologie von \mathcal{H}) und somit schwach kompakt.*

Beweis. Wir zeigen zunächst die Konvexität der Menge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ und wählen dazu Elemente $z_1 := (x_1, u_1) \in \mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ und $z_2 := (x_2, u_2) \in \mathcal{N}_{\mathcal{B}_P}(J, \alpha)$. Das Zielfunktional J ist strikt konvex (auf \mathcal{B}_P) und somit insbesondere konvex. Dies impliziert

$$J((1 - \lambda)z_1 + \lambda z_2) \leq (1 - \lambda)J(z_1) + \lambda J(z_2) = (1 - \lambda)J(x_1, u_1) + \lambda J(x_2, u_2)$$

für alle $z_1, z_2 \in \mathcal{N}_{\mathcal{B}_P}(J, \alpha)$.

Die Abgeschlossenheit der Menge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ folgt aus der Stetigkeit der Operatoren Q und q (bzgl. der Normtopologie¹⁸) auf \mathcal{H} , welche wiederum die Stetigkeit des Zielfunktional J auf \mathcal{H} bezüglich der Normtopologie des Raumes \mathcal{H} impliziert. Aufbauend auf diesem Wissen betrachten wir nun eine (stark) konvergente Funktionenfolge $\{(x_n, u_n)\}$ mit $\{(x_n, u_n)\} \xrightarrow{\mathcal{H}} (x_0, u_0)$. Aufgrund der Abgeschlossenheit der Menge \mathcal{B}_P ¹⁹ gilt $(x_0, u_0) \in \mathcal{B}_P$. Die Stetigkeit des Funktional J bzgl. der Normtopologie des Raumes \mathcal{H} erlaubt dann die Abschätzung

$$J(x_0, u_0) = J(\lim_{n \rightarrow \infty} x_n, \lim_{n \rightarrow \infty} u_n) = \lim_{n \rightarrow \infty} J(x_n, u_n) \leq \alpha.$$

Damit ist $(x_0, u_0) \in \mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ und somit die Niveaumenge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ (stark) abgeschlossen. Aufgrund der bereits nachgewiesenen Konvexität der Menge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ können wir nun unter Ausnutzung von Satz 2 auf die schwache Abgeschlossenheit der Menge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ schließen²⁰.

Wir untersuchen nun die Beschränktheit der Niveaumengen: Gemäß Lemma 14 ist die Bilinearform Q koerzitiv auf $\mathcal{A}^{-1}(\mathbf{b})$. Lemma 6 sichert zudem die Beschränktheit der Linearform q bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ auf ganz \mathcal{H} . Damit existiert für alle $z = (x, u) \in \mathcal{B}_P \subseteq \mathcal{A}^{-1}(\mathbf{b}) \subset \mathcal{H}$ eine positive Konstante c_1 und eine Konstante $c_2 \in \mathbb{R}$, so dass

$$Q(z, z) \geq 2c_1 \|z\|_{\mathcal{H}}^2 \quad \text{und} \quad q(z) \geq -c_2 \|z\|_{\mathcal{H}}.$$

¹⁸vgl. Lemma 5 und Lemma 6, S. 27 ff.

¹⁹vgl. Lemma 12, S. 34

²⁰siehe Satz 2, S. 9

Demzufolge gilt für beliebige, aber feste Elemente $z = (x, u) \in \mathcal{B}_P \subseteq \mathcal{A}^{-1}(\mathbf{b}) \subset \mathcal{H}$

$$J(x, u) = \frac{1}{2}Q(z, z) + q(z) \geq c_1 \|z\|_{\mathcal{H}}^2 - c_2 \|z\|_{\mathcal{H}} = c_1 \left\{ \left(\|z\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right)^2 - \left(\frac{c_2}{2c_1} \right)^2 \right\},$$

so dass wir nun für alle Elemente $(x, u) \in \mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ die Abschätzung

$$\alpha \geq J(x, u) \geq c_1 \left\{ \left(\|z\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right)^2 - \left(\frac{c_2}{2c_1} \right)^2 \right\}$$

erhalten, welche impliziert

$$\left| \|z\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right| \leq \sqrt{\frac{1}{c_1} \alpha + \left(\frac{c_2}{2c_1} \right)^2}.$$

Damit ist $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ beschränkt auf \mathcal{H} bezüglich der Normtopologie des Raumes \mathcal{H} .

Die Ausführungen des Beweises liefern: $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$ ist eine auf \mathcal{H} (norm-)beschränkte, konvexe und schwach abgeschlossene Menge. Nach Satz 2 liefert dies den Schluss auf die schwache Kompaktheit der Menge $\mathcal{N}_{\mathcal{B}_P}(J, \alpha)$. Damit sind alle Aussagen des Lemmas nachgewiesen. \square

Wir fassen die Ergebnisse dieses Abschnittes zusammen und erhalten die folgende Aussage über die Existenz und Eindeutigkeit einer optimalen Lösung für (P_ν) mit $\beta > 0$:

Satz 9. *Sei $\beta > 0$. Unter Gültigkeit der Voraussetzungen (V1)-(V4) besitzt das Problem (P_ν) eine eindeutige Lösung.*

Beweis. Wir argumentieren wie in der Beweisführung aus [Bur19], Satz 4.13, S. 96:

Die Ergebnisse der vorangegangenen Lemmata erlauben uns nun den Schluss der Existenz einer optimalen Lösung der Aufgabe (P_ν) über den verallgemeinerten Satz von Weierstraß. Dazu betrachten wir die entsprechenden Niveaumengen $\mathcal{N}_{\mathcal{B}_P}$ des Zielfunktionals J und wählen ein Paar $(x_1, u_1) \in \mathcal{B}_P$, deren Existenz aufgrund der Voraussetzung (V4) gesichert ist. Da (x_1, u_1) selbst ein Element der Niveaumenge $\mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1))$ ist, gilt $\mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1)) \neq \emptyset$. Nun nimmt das schwach unterhalbstetige Funktional J auf der schwach kompakten Menge $\mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1))$ seinen Minimalwert an, das heißt, es existiert ein Funktionenpaar $(x_*, u_*) \in \mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1))$ mit

$$J(x_*, u_*) \leq J(x, u) \quad \forall (x, u) \in \mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1)).$$

Für alle anderen in der Aufgabe (P_ν) zulässigen Paare $(x, u) \in \mathcal{B}_P \setminus \mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1))$ ist

$$J(x, u) > J(x_1, u_1) \geq J(x_*, u_*).$$

Daher gilt

$$J(x_*, u_*) = \inf_{(x, u) \in \mathcal{B}_P} J(x, u) \tag{2.16}$$

und (x_*, u_*) ist demnach eine Lösung der Aufgabe (P_ν) . Insbesondere besitzen (P_ν) und die Hilfsaufgabe

$$J(x, u) \rightarrow \inf!$$

$$\text{bzgl. } (x, u) \in \mathcal{N}_{\mathcal{B}_P}(J, J(x_1, u_1))$$

mit $(x_1, u_1) \in \mathcal{B}_P$ dasselbe optimale Paar (x_*, u_*) und denselben Optimalwert $J(x_*, u_*)$. Die strikte Konvexität von J garantiert schließlich die Eindeutigkeit der Lösung. Den Nachweis dieser Aussage zeigen wir mit Hilfe eines Widerspruchs. Dafür nehmen wir an, (\tilde{x}, \tilde{u}) sei eine weitere Lösung von (P_ν) , d.h. es gilt $J(x_*, u_*) = J(\tilde{x}, \tilde{u})$. Aus der Konvexität der Menge \mathcal{B}_P können wir schlussfolgern, dass mit (x_*, u_*) und (\tilde{x}, \tilde{u}) auch (x_2, u_2) mit $x_2 := \frac{1}{2}(x_* + \tilde{x})$, $u_2 := \frac{1}{2}(u_* + \tilde{u})$ Element der Menge \mathcal{B}_P ist. Angenommen es wäre $(x_*, u_*) \neq (\tilde{x}, \tilde{u})$, dann liefert die strikte Konvexität von J die Ungleichung

$$J(x_2, u_2) < \frac{1}{2}J(x_*, u_*) + \frac{1}{2}J(\tilde{x}, \tilde{u}) = J(x_*, u_*),$$

welche im Widerspruch zu (2.16) und damit zu der Optimalität von (x_*, u_*) steht. Daher war die Annahme falsch und es muss $(x_*, u_*) = (\tilde{x}, \tilde{u})$ gelten, womit die Eindeutigkeit der Lösung von (P_ν) folgt. \square

2.5 Notwendige Optimalitätsbedingungen

Für Probleme vom Typ (P_ν) mit Dichtefunktionen ν können wir auf den in [Zie19] formulierten notwendigen Bedingungen aufbauen. Die konkrete Struktur der in (P_ν) behandelten linearen Dynamik und des Integranden φ ermöglicht uns jedoch eine vereinfachte Formulierung der in [Zie19] dargestellten Ergebnisse. Vergleichen wir dazu die Ergebnisse aus [ZP18], S. 763 ff. für ein Optimalsteuerungsproblem mit einer Dynamik der Form

$$f(t, \xi, v) = A(t)\xi + B(t)v + C(t)$$

in gewichteten Räumen mit Dichtefunktionen, zusammen mit den bereits getätigten Überlegungen hinsichtlich der schwachen Unterhalbstetigkeit des Zielfunktional, so vereinfachen sich die dortigen Voraussetzungen wie folgt:

- (Z1) Die Funktion $\varphi = \varphi(t, \xi, v)$ sei stetig in all ihren Argumenten sowie stetig differenzierbar bzgl. ξ und v .
- (Z2) Die Menge $\mathcal{U} \subseteq \mathbb{R}^m$ sei konvex und kompakt.
- (Z3) Die Funktionen A, B, C des dynamischen Systems genügen den folgenden Bedingungen:

$$A \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times n}), B \in L_\infty^{n \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times m})$$

$$C \in L_\infty^n(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^n).$$

- (Z4) Die Funktionen W, R, w und r des Integranden genügen den folgenden Bedingungen

$$W \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times n}), R \in L_\infty^{m \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m}),$$

$$w \in L_2^n(\mathbb{R}_{++}, e^{\beta t}) \cap C(\mathbb{R}_+, \mathbb{R}^n), r \in L_2^m(\mathbb{R}_{++}, e^{\beta t}) \cap C(\mathbb{R}_+, \mathbb{R}^m)$$

(Z5) $\beta < 0$ sei derart gewählt, dass gilt: Es existiere ein $T > 0$, so dass

$$\beta < -2 \int_T^t |A(\tau)| d\tau$$

für alle $t \geq T$, wobei $|A| := \sqrt{\lambda_{\max}(A^T A)}$ und $\lambda_{\max}(A^T A)$ den größten Eigenwert der Matrix $A^T A$ bezeichne.

Offenbar liefert damit einzig Voraussetzung (Z5) eine zusätzliche Restriktion an die Aufgabenklasse (P_ν) . Unter diesen Voraussetzungen sind die in [Ziel19](#) formulierten notwendigen Bedingungen anwendbar und durch die strikte Konvexität des Zielfunktional auch hinreichend und können in folgender Form angegeben werden:

Satz 10 ([ZP18](#), S. 764). *Sei $(x_*, u_*) \in \mathcal{B}$ eine optimale Lösung von (P_ν) , $\nu(t) = e^{\beta t}$, $\beta < 0$, im Sinne des **Kriteriums L**²¹. Unter den Bedingungen (Z1) - (Z5) existieren Multiplikatoren (λ_0, y) , welche den folgenden Bedingungen genügen:*

$$\lambda_0 = 1, \tag{N}$$

$$y \in W_2^{1,n}(\mathbb{R}_{++}, \nu^{-1}) \tag{T}$$

$$H(t, x_*(t), u_*(t), y(t), \lambda_0) = \max_{v \in U} H(t, x_*(t), v, y(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{M}$$

$$y'(t) = -\nabla_\xi H(t, x_*(t), u_*(t), y(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{K}$$

wobei $H : \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ die Pontryaginsche Funktion mit

$$H(t, \xi, v, \eta, 1) = -\frac{1}{2} \left(\xi^T W(t) \xi + v^T R(t) v + w^T(t) \xi + r^T(t) v \right) e^{\beta t} + \eta^T (A(t) \xi + B(t) v + C(t))$$

ist.

Wir sind in der Lage, eine äquivalente Formulierung von Satz [10](#) mittels der Setzung von $\tilde{y} = y\nu^{-1}$ in der folgenden Form anzugeben:

Satz 11. *Sei $(x_*, u_*) \in \mathcal{B}$ eine optimale Lösung von (P_ν) , $\nu(t) = e^{\beta t}$, $\beta < 0$, im Sinne des **Kriteriums L**. Unter den Bedingungen (Z1) - (Z5) existieren Multiplikatoren (λ_0, \tilde{y}) mit $\tilde{y} := y\nu^{-1}$, welche den folgenden Bedingungen genügen:*

$$\lambda_0 = 1, \tag{\tilde{N}}$$

$$\tilde{y} \in W_2^{1,n}(\mathbb{R}_{++}, \nu) \tag{\tilde{T}}$$

$$\tilde{H}(t, x_*(t), u_*(t), \tilde{y}(t), \lambda_0) = \max_{v \in U} \tilde{H}(t, x_*(t), v, \tilde{y}(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{\tilde{M}}$$

$$\tilde{y}'(t) = -\nabla_\xi \tilde{H}(t, x_*(t), u_*(t), \tilde{y}(t), \lambda_0) - \beta \tilde{y}(t) \quad (\forall) t \in \mathbb{R}_{++}, \tag{\tilde{K}}$$

wobei $\tilde{H} : \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ die modifizierte Pontryaginsche Funktion mit

$$\tilde{H}(t, \xi, v, \tilde{\eta}, 1) = -\frac{1}{2} \left(\xi^T W(t) \xi + v^T R(t) v + w^T(t) \xi + r^T(t) v \right) + \tilde{\eta}^T (A(t) \xi + B(t) v + C(t))$$

ist.

²¹vgl. Definition [20](#)

Beweis. Offensichtlich gelangen wir mittels der Setzung $y = \tilde{y}\nu$ unmittelbar zur Formulierung von Satz [10](#). Der Beweis hierfür ist in [ZP18](#) in ausführlicher Form gegeben. \square

Für den Fall sogenannter Box-Constraints, d.h. für kompakte Steuerbereiche der Form $\mathcal{U} := [\underline{u}, \bar{u}]$, mit $\underline{u}, \bar{u} \in \mathbb{R}^m$, lassen sich aus obigem Satz einige Regularitätseigenschaften der auftretenden Funktionen ableiten. Im Hinblick auf spätere Untersuchungen wollen wir diese kurz diskutieren.

Über die Regularität der adjungierten Funktion

Wir betrachten nun den Fall eines kompakten Steuerbereiches in Form von Box-Constraints. Zusätzlich zu den Bedingungen (Z1)-(Z5) sei die Voraussetzung (V1) erfüllt. Weiterhin gelte

(V5) Für die Matrix- und Vektorfunktionen im Zielfunktional [\(2.6\)](#) gelte zusätzlich

$$R \in L_{\infty}^{m \times m}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{m \times m}), \quad R'_{k,l} \in L_{\infty}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}) \quad k, l = 1, \dots, m,$$

$$w \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}) \cap C^1(\mathbb{R}_+, \mathbb{R}^n) \quad \text{und} \quad r \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}).$$

Aus der Maximumbedingung (M) bzw. (\tilde{M}) können wir Regularitätseigenschaften der Lösung (x_*, u_*, y) bzw. (x_*, u_*, \tilde{y}) ableiten. Dazu möchten wir jedoch zunächst das folgende Lemma einführen, auf welches sich unsere Argumentation stützen wird:

Lemma 17 (vgl. [PB18](#), S. 7, Lemma 3). *Sei $\nu(t) = e^{\beta t}$, $\beta < 0$, $y(0) = 0$. Dann gilt:*

1. Falls $x \in W_2^1(\mathbb{R}_{++}, e^{\beta t})$, dann ist $xe^{\beta t} \in W_1^1(\mathbb{R}_{++})$ und $xe^{\beta t} \in W_2^1(\mathbb{R}_{++}, e^{-\beta t})$.
2. Falls $y \in W_2^1(\mathbb{R}_{++}, e^{-\beta t})$, dann ist $y \in W_1^1(\mathbb{R}_{++})$ und $ye^{-\beta t} \in W_2^1(\mathbb{R}_{++}, e^{\beta t})$.

Beweis. Wir zeigen exemplarisch die erste Gleichung der Behauptung. Unter Verwendung der Ungleichung von Cauchy-Schwarz erhalten wir

$$\int_0^{\infty} |x(t)e^{\beta t}| dt \leq \left(\int_0^{\infty} x^2(t)e^{\beta t} dt \right)^{1/2} \left(\int_0^{\infty} e^{\beta t} dt \right)^{1/2} < \infty$$

und

$$\int_0^{\infty} |\mathcal{D}x(t)e^{\beta t}| dt \leq \left(\int_0^{\infty} |\mathcal{D}x(t)|^2 e^{\beta t} dt \right)^{1/2} \left(\int_0^{\infty} e^{\beta t} dt \right)^{1/2} < \infty,$$

woraus folgt $xe^{\beta t} \in W_1^1(\mathbb{R}_{++})$. Weiterhin besitzt $x(t)e^{\beta t}$ eine verallgemeinerte Ableitung. Mit $\mathcal{D}(x(t)e^{\beta t}) = (\mathcal{D}x(t) + \beta x(t))e^{\beta t} = (x'(t) + \beta x(t))e^{\beta t}$ erhalten wir

$$\int_0^{\infty} x^2(t)e^{2\beta t}e^{-\beta t} dt = \int_0^{\infty} x^2(t)e^{\beta t} dt < \infty$$

und

$$\int_0^{\infty} [(x'(t)e^{\beta t})]^2 e^{-\beta t} dt \leq 2 \int_0^{\infty} [(x'(t))^2 + \beta^2 x^2(t)] e^{2\beta t} e^{-\beta t} dt < \infty,$$

d.h. es folgt $xe^{\beta t} \in W_2^1(\mathbb{R}_{++}, e^{-\beta t})$. \square

Wir kehren nun zurück zur Betrachtung der Regularität der Funktion y und zeigen dafür zunächst die Regularität der Funktion u_* , die wir aus der Maximumbedingung (M) mit

$$u_*(t) = \sigma_*(t, y(t)) := \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) \left[-r(t) + B^T(t)y(t)e^{-\beta t} \right] \right) \quad (\forall)t \in \mathbb{R}_{++} \quad (2.17)$$

bzw.

$$u_*(t) = \sigma_*(t, \tilde{y}(t)) := \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) \left[-r(t) + B^T(t)\tilde{y}(t) \right] \right) \quad (\forall)t \in \mathbb{R}_{++}$$

erhalten, wobei $\text{Proj}_{\mathcal{U}}(f(t))$ die Projektion der Funktion $f(t)$ auf die Menge \mathcal{U} bezeichnet, d.h.

$$\text{Proj}_{\mathcal{U}}(f(t)) := (\min \{ \bar{u}_1, \max \{ f_1(t), \underline{u}_1 \} \}, \dots, \min \{ \bar{u}_m, \max \{ f_m(t), \underline{u}_m \} \})^T.$$

Da die Funktion y dem Raum $W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$, $\beta < 0$ angehört, liefert Lemma [17](#), dass y ein Element der Menge $W_1^{1,n}(\mathbb{R}_{++})$ ist²². Unter Gültigkeit unserer Voraussetzungen erhalten wir gemäß [ZP18](#) (vgl. S. 779) zudem die Stetigkeit und Beschränktheit von σ_* sowie $\sigma_* \in L_2^n(\mathbb{R}_{++}, e^{\beta t}) \cap C(\mathbb{R}_+, \mathbb{R}^m)$.

Mit Hilfe dieser Erkenntnis analysieren wir nun die Regularitätseigenschaften von (x_*, y) bzw. (x_*, \tilde{y}) . Diese resultieren aus den kanonischen Gleichungen, wenn wir die Funktion u_* im kanonischen System durch [\(2.17\)](#) ersetzen.

Lemma 18. *Sei (x_*, y) eine Lösung des kanonischen Gleichungssystems*

$$\begin{aligned} x_*'(t) &= A(t)x_*(t) + B(t)\sigma_*(t, y(t)) + C(t) \\ y'(t) &= (W(t)x_*(t) + w(t))e^{\beta t} - A^T(t)y(t) \end{aligned} \quad (\forall)t \in \mathbb{R}_{++}. \quad (2.18)$$

Dann können wir mit $\tilde{y}(t) := y(t)e^{-\beta t}$ schließen:

$$\begin{aligned} (x_*, y) &\in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}) \times W_2^{2,n}(\mathbb{R}_{++}, e^{-\beta t}) \quad \text{sowie} \\ (x_*, \tilde{y}) &\in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t}) \times W_2^{2,n}(\mathbb{R}_{++}, e^{\beta t}). \end{aligned}$$

Beweis. Sei $x_* \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ und $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$. Gemäß Lemma [17](#) erhalten wir $x_*e^{\beta t} \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$ und $ye^{-\beta t} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$. Mittels der Voraussetzungen (V1) und (V5) gilt dann die kanonische Gleichung [\(2.18\)](#) sogar punktweise $\forall t \in \mathbb{R}_+$,

$$\begin{aligned} x_*'(t) &= A(t)x_*(t) + B(t)\sigma_*(t, y(t)) + C(t), \\ y'(t) &= (W^T(t)x_*(t) + w(t))e^{\beta t} - A^T(t)y(t), \end{aligned}$$

wobei aufgrund der Abschätzung

$$\begin{aligned} \int_0^\infty |y'(t)|^2 e^{-\beta t} dt &= \int_0^\infty \left| (W^T(t)x_*(t) + w(t))e^{\beta t} - A^T(t)y(t) \right|^2 e^{-\beta t} dt \\ &\leq 2 \int_0^\infty |W^T(t)x_*(t) + w(t)|^2 e^{\beta t} dt + 2 \int_0^\infty |A^T(t)y(t)|^2 e^{-\beta t} dt \\ &\leq 4 \int_0^\infty |W^T(t)x_*(t)|^2 e^{\beta t} dt + 4 \int_0^\infty |w(t)|^2 e^{\beta t} dt + 2\bar{A}^2 \int_0^\infty |y(t)|^2 e^{-\beta t} dt \\ &\leq 4\bar{W}^2 \|x\|_{(2), e^{\beta t}}^2 + 4\|w\|_{(2), e^{\beta t}}^2 + 2\bar{A}^2 \|y\|_{(2), e^{-\beta t}}^2 < \infty \end{aligned}$$

²²Die Forderung der Anfangsbedingung $y(0) = 0$ aus Lemma [17](#) lässt sich leicht durch eine lineare Transformation wegdiskutieren.

gilt $y' \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$. Aufgrund der Eigenschaften der eingehenden Funktionen gilt $y''(t) = \left[\beta(W^T(t)x_*(t) + w(t)) + (W'^T(t)x_*(t) + W^T(t)x'_*(t) + w'(t)) \right] e^{\beta t} - A'^T(t)y(t) - A^T(t)y'(t)$ für fast alle $t \in \mathbb{R}_{++}$. Für diese Funktionen können wir abschätzen

$$\begin{aligned} \int_0^\infty |y''(t)|^2 e^{-\beta t} dt &\leq \int_0^\infty \left| \left[\beta(W^T(t)x_*(t) + w(t)) + (W'^T(t)x_*(t) + W^T(t)x'_*(t) + w'(t)) \right] e^{\beta t} - A'^T(t)y(t) - A^T(t)y'(t) \right|^2 e^{-\beta t} dt \\ &\leq 4\beta^2 \int_0^\infty |W^T(t)x_*(t) + w(t)|^2 e^{\beta t} dt + 4 \int_0^\infty |W'^T(t)x_*(t) + W^T(t)x'_*(t) + w'(t)|^2 e^{\beta t} dt \\ &\quad + 2 \int_0^\infty |A'^T(t)y(t) + A^T(t)y'(t)|^2 e^{-\beta t} dt \\ &\leq 8\beta^2 \left(\bar{W}^2 \|x\|_{(2),e^{\beta t}}^2 + \|w\|_{(2),e^{\beta t}}^2 \right) + 8\bar{W}'^2 \|x\|_{(2),e^{\beta t}}^2 + 16\bar{W}^2 \|x'\|_{(2),e^{\beta t}}^2 \\ &\quad + 16\|w\|_{(2),e^{\beta t}}^2 + 4\bar{A}'^2 \|y\|_{(2),e^{-\beta t}}^2 + 4\bar{A} \|y'\|_{(2),e^{-\beta t}}^2 < \infty, \end{aligned}$$

so dass $y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\beta t})$ gilt. Offenbar gilt zudem

$$\int_0^\infty |\tilde{y}(t)|^2 e^{\beta t} dt = \int_0^\infty |y(t)e^{-\beta t}|^2 e^{\beta t} dt = \int_0^\infty |y(t)|^2 e^{-\beta t} dt < \infty$$

sowie

$$\begin{aligned} \int_0^\infty |\tilde{y}'(t)|^2 e^{\beta t} dt &= \int_0^\infty |y'(t)e^{-\beta t} - \beta y(t)e^{-\beta t}|^2 e^{\beta t} dt \\ &\leq 2 \left(\int_0^\infty |y'(t)|^2 e^{-\beta t} dt + \beta^2 \int_0^\infty |y(t)|^2 e^{-\beta t} dt \right) < \infty \end{aligned}$$

und

$$\begin{aligned} \int_0^\infty |\tilde{y}''(t)|^2 e^{\beta t} dt &= \int_0^\infty \left| \beta^2 y(t)e^{-\beta t} - 2\beta y'(t)e^{-\beta t} + y''(t)e^{-\beta t} \right|^2 e^{\beta t} dt \\ &\leq 2\beta^2 \int_0^\infty |y(t)|^2 e^{-\beta t} dt + 2 \int_0^\infty |2\beta y'(t)e^{-\beta t} + y''(t)e^{-\beta t}|^2 e^{\beta t} dt \\ &\leq 2\beta^2 \|y\|_{(2),e^{-\beta t}}^2 + 8\beta \|y'\|_{(2),e^{-\beta t}}^2 + 4 \|y''\|_{(2),e^{-\beta t}}^2 < \infty \end{aligned}$$

und damit $\tilde{y} \in W_2^{2,n}(\mathbb{R}_{++}, e^{\beta t})$. □

Bemerkung 11. Jede Funktion $y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\beta t})$ bzw. $\tilde{y} \in W_2^{2,n}(\mathbb{R}_{++}, e^{\beta t})$ besitzt einen stetig differenzierbaren Repräsentanten auf \mathbb{R}_+ . Der Wert $y'(0)$, bzw. $\tilde{y}'(0)$ ist also wohldefiniert.

Eine Formulierung von Optimalitätsbedingungen für Probleme vom Typ (P_ν) mit echten Gewichtsfunktionen ist zu diesem Zeitpunkt nicht möglich. Da wir aber in der Lage waren, die Existenz und Eindeutigkeit einer solchen optimalen Lösung nachzuweisen, soll uns nun die Formulierung einer zu (P_ν) dualen Aufgabe helfen weitere Aussagen zu tätigen.

Kapitel 3

Dualitätstheorie für linear-quadratische Optimalsteuerungsprobleme

Gemäß der Betrachtungen des vorangegangenen Kapitels sind wir in der Lage Voraussetzungen anzugeben, welche die Existenz und Eindeutigkeit einer Lösung für Optimalsteuerungsprobleme vom Typ (P_ν) garantieren. Ist das Problem (P_ν) über gewichteten Funktionenräumen mit der Dichtefunktion $\nu(t) = e^{\beta t}$ mit $\beta < 0$ gegeben, so konnten wir zudem auf die Anwendbarkeit der notwendigen Optimalitätsbedingungen gemäß Satz [10](#) schließen. Offen bleibt jedoch die Ausweitung derartiger Ergebnisse auf Optimalsteuerungsprobleme über gewichtete Funktionenräumen mit echten Gewichtsfunktionen. Hier sollen uns die Ideen der Dualitätstheorie weiterführen. Dazu beginnen wir mit einem allgemeinen Einblick in die Thematik, welcher uns einerseits die Relevanz aber auch die Komplexität dieses Gebietes vor Augen führen wird. Die Komplexität leitet sich dabei unter anderem aus der Existenz vieler Freiheitsgrade im Zuge der Konstruktion ab. In Folge dessen sind im Laufe der Jahre verschiedene Dualitätstheorien entstanden. Im Rahmen dieser Arbeit werden wir jedoch nur auf die *Lagrange-Dualität* eingehen. Für die Behandlung anderer Dualitätskonzeptionen verweisen wir beispielsweise auf [\[Gao13\]](#) und die darin angegebenen Quellen.

Unsere hier angeführten Ideen basieren auf den Ergebnissen aus [\[Lyk10\]](#), in denen ein dualer Zugang für eine allgemeinere Klasse an Optimalsteuerungsproblemen mit unendlichem Zeithorizont in gewichteten Funktionenräumen mit Dichtefunktionen vorgestellt wird. Die dortigen Ausführungen basieren auf den Ausarbeitungen von Klötzler [\[1\]](#) für Optimalsteuerungsprobleme mit endlichem Zeithorizont. Damit lehnen wir uns ebenfalls an diesen Ideen an. Im Gegensatz zu den zitierten Arbeiten, in welchen die Autoren durch die punktweise Auswertung der Hamilton-Jacobischen Ungleichung zu einer Formulierung der dualen Aufgabe in punktwieser Form gelangen, wird uns die hier vorgestellte Konstruktion jedoch zu einer Formulierung der dualen Aufgabe in Form eines Variationsproblems führen und gleichzeitig die Behandlung von Problemen des Typs (P_ν) unter Verwendung allgemeiner Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$ mit $\beta \neq 0$ ermöglichen.

¹vgl. beispielsweise [\[Kl79\]](#)

3.1 Ein kurzer Einblick in die Dualitätstheorie

Definition 21 (vgl. beispielsweise [Kl080]). Gegeben seien reelle Funktionale $F : X \rightarrow \mathbb{R}$ und $G : Y \rightarrow \mathbb{R}$. Die Aufgabe

$$(D) \quad G(y) \rightarrow \sup! \quad \text{bzgl. } y \in Y$$

heißt duale Aufgabe zu

$$(P) \quad F(x) \rightarrow \inf! \quad \text{bzgl. } x \in X,$$

falls

$$G(y) \leq F(x) \quad \forall x \in X, \quad \forall y \in Y$$

$$\Leftrightarrow$$

$$\sup(D) = \sup_{y \in Y} G(y) \leq \inf_{x \in X} F(x) = \inf(P). \quad (3.1)$$

Gilt für die Relation (3.1) die Beziehung “=”, so sprechen wir von starker Dualität, andernfalls sprechen wir von schwacher Dualität.

Bemerkung 12. Falls $X = \emptyset$, so setzen wir $\inf_{x \in X} F(x) = \infty$. Gilt $Y = \emptyset$, so setzen wir $\sup_{y \in Y} G(y) = -\infty$.

Aus dieser Definition ergibt sich die Relevanz der Betrachtung der Dualität, denn durch eine derartige Einführung einer dualen Aufgabe erhalten wir eine hinreichende Bedingung für die Optimalität einer Lösung $\hat{x}(\cdot)$ des Problems (P):

$$\text{Gilt für ein } \hat{x} \in X \quad \sup_{y \in Y} G(y) = F(\hat{x}), \text{ so ist } \hat{x} \text{ globale Minimalstelle von (P).} \quad (3.2)$$

Um eine hinreichende Optimalitätsbedingung für das Steuerungsproblem zu formulieren, ist es daher hilfreich, eine zugehörige duale Aufgabe zu betrachten.

Einen weiteren Vorteil, der insbesondere auch für die später durchgeführte algorithmische Behandlung von Relevanz ist, erhalten wir selbst aus der schwachen Dualitätsbeziehung. Denn gemäß obiger Definition sind wir durch die Gegenüberstellung der dualen und primalen Aufgabe in der Lage, ihre Optimalwerte gemäß der Beziehung²

$$\exists \hat{x} \in X, \hat{y} \in Y : G(\hat{y}) \leq \inf_{x \in X} F(x) \leq F(\hat{x})$$

abzuschätzen.

Das nachfolgende Konstruktionsschema, welches generell für Steuerungsprobleme mit Differentialgleichungen erster Ordnung anwendbar ist, und die Dualitätsbegriffe wurden [Pic09] entnommen und sind so auch in [Lyk10] zu finden. Der Vollständigkeit halber wollen wir diesen Abschnitt auch hier anführen:

Der Algorithmus der Aufstellung einer dualen Aufgabe erfolgt in drei Schritten, die nun ausführlicher beschrieben werden.

²vgl. [Lyk10], S. 96

Schritt 1: Man gebe eine geeignete Darstellung des zulässigen Bereiches X von (P) an: $X = X_0 \cap X_1$. Die Vorgabe von X_0 und X_1 wird oft durch die Optimierungsaufgabe vorgeprägt.

Schritt 2: Man definiere eine Menge Y und ein reelles Funktional $\Phi(\cdot, \cdot) : X_0 \times Y \rightarrow \overline{\mathbb{R}}$ mit folgender Eigenschaft, die *Äquivalenzrelation* genannt wird:

$$\inf_{x \in X} F(x) = \inf_{x \in X_0} \left[\sup_{y \in Y} \Phi(x, y) \right]. \quad (3.3)$$

Für Ausdrücke der Form (3.3) gilt das folgende Lemma:

Lemma 19 ([Pic09]). Für beliebige X_0 und Y und beliebige Funktionale $\Phi(\cdot, \cdot)$ ist

$$\inf_{x \in X_0} \left[\sup_{y \in Y} \Phi(x, y) \right] \geq \sup_{y \in Y} \left[\inf_{x \in X_0} \Phi(x, y) \right].$$

Beweis.

$$\begin{array}{llll} \sup_{y \in Y} \Phi(x, y) & \geq & \Phi(x, y) \quad \forall y \in Y, \quad \forall x \in X & \Big| \inf_{x \in X_0} \\ \inf_{x \in X_0} \sup_{y \in Y} \Phi(x, y) & \geq & \inf_{x \in X_0} \Phi(x, y) & \Big| \sup_{y \in Y} \\ \inf_{x \in X_0} \sup_{y \in Y} \Phi(x, y) & \geq & \sup_{y \in Y} \inf_{x \in X_0} \Phi(x, y). & \end{array}$$

□

Damit erhalten wir aus der Relation (3.3) die folgende Abschätzung

$$\inf_{x \in X_0} \sup_{y \in Y} \Phi(x, y) \geq \sup_{y \in Y} \left[\inf_{x \in X_0} \Phi(x, y) \right]. \quad (3.4)$$

Schritt 3: Man setze für ein fixiertes y

$$G(y) := \inf_{x \in X_0} \Phi(x, y). \quad (3.5)$$

Laut dieser Konstruktion gilt dann die folgende Abschätzungskette:

$$\inf_{x \in X} F(x) \stackrel{(3.3)}{=} \inf_{x \in X_0} \left[\sup_{y \in Y} \Phi(x, y) \right] \stackrel{(3.4)}{\geq} \sup_{y \in Y} \left[\inf_{x \in X_0} \Phi(x, y) \right] \stackrel{(3.5)}{=} \sup_{y \in Y} G(y),$$

woraus man schließt, dass jede Aufgabe vom Typ

$$(D) : \quad \tilde{G}(y) \rightarrow \text{sup! bzgl. } y \in Y_0 \subset Y \text{ mit } \tilde{G}(s) \leq G(s) \quad \forall s \in Y_0 \quad (3.6)$$

dual zur Aufgabe (P) ist.

Bemerkung 13.

- (a) Das Aufstellen der dualen Aufgabe erfordert das Lösen einer Optimierungsaufgabe. Die Menge Y und das Funktional $\Phi(x, y)$ sind so zu wählen, dass dies einfacher als das Lösen der Ausgangsaufgabe ist.
- (b) In der Auswahl der Mengen X_0, X_1, Y und des Funktionals Φ hat man viele Freiheitsgrade. Deshalb sind mehrere Dualitätstheorien entstanden.

3.2 Anwendung des Konstruktionsschemas für Probleme der Klasse (P_ν)

Für die Formulierung einer dualen Aufgabe wollen wir nun konkret zwischen den Fällen

- $\nu(t) = \nu_0(t) := e^{\beta_0 t}$ mit $\beta_0 < 0$ ist Dichtefunktion und
- $\nu(t) = \nu_1(t) := e^{\beta_1 t}$ mit $\beta_1 > 0$ ist echte Gewichtsfunktion

unterscheiden. Dazu betrachten wir fortan die Problemklassen (\mathbf{P}_{ν_i}) , $i = 0, 1$:

Minimiere

$$J_{\nu_i}(x, u) = \int_0^\infty \{\varphi(t, x(t), u(t))\} \nu_i(t) dt \quad (3.7)$$

bezüglich aller Paare

$$(x, u) \in W_2^{1,n}(\mathbb{R}_{++}, \nu_i) \times L_2^m(\mathbb{R}_{++}, \nu_i) \quad \text{wobei} \quad \nu_i(t) = e^{\beta_i t}, \quad \beta_0 < 0, \quad \beta_1 > 0, \quad (3.8)$$

für die gilt

$$x'(t) = A(t)x(t) + B(t)u(t) + C(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad (3.9)$$

$$x(0) = x_0 \neq 0 \quad (3.10)$$

$$u(t) \in \mathcal{U} \subseteq \mathbb{R}^m \quad (\forall) t \in \mathbb{R}_{++}, \quad (3.11)$$

sowie

$$\varphi(t, \xi, v) = \frac{1}{2}(\xi^T W(t)\xi + v^T R(t)v) + w^T(t)\xi + r^T(t)v,$$

mit den gegebenen Funktionen $W : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$, $R : \mathbb{R}_+ \rightarrow \mathbb{R}^{m \times m}$, $w : \mathbb{R}_+ \rightarrow \mathbb{R}^n$, $r : \mathbb{R}_+ \rightarrow \mathbb{R}^m$ und $A : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$, $B : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times m}$ sowie $C : \mathbb{R}_+ \rightarrow \mathbb{R}^n$. Zudem seien die Voraussetzungen (V1)-(V4) aus Kapitel 2 weiterhin gültig. Insbesondere folgt aus Voraussetzung (V2), welche die positive Definitheit der symmetrischen Matrizen W und R für alle $t \in \mathbb{R}_+$ fordert, auch die Invertierbarkeit der Matrizen W und R .

Für Optimalsteuerungsprobleme der Form (P_{ν_i}) wollen wir nun eine duale Aufgabe formulieren und folgen dazu dem allgemeinen Konstruktionsschema aus Kapitel 3.1:

Schritt 1: Gemäß Schritt 1 des Konstruktionsschemas zerlegen wir den zulässigen Bereich \mathcal{B}_P der Probleme (P_{ν_i}) als Schnitt zweier Mengen X_0 und X_1 . Dabei wählen wir die Menge X_0 nun derart, dass sie alle Paare (x, u) beinhaltet, die den Nebenbedingungen des Problems (P_{ν_i}) genügen – mit Ausnahme der Zustandsgleichung (3.9). Die Menge X_1 beinhaltet alle Paare (x, u) die innerhalb der Regularitätsklasse (3.8) liegen und der Zustandsgleichung genügen. Demnach setzen wir:

$$X_0 := \left\{ (x, u) \in W_2^{1,n}(\mathbb{R}_{++}, \nu_i) \times L_2^m(\mathbb{R}_{++}, \nu_i) \mid x(0) - x_0 = 0, u(t) \in \mathcal{U} \quad (\forall) t \in \mathbb{R}_{++}, \right\}.$$

Wir definieren³ nun die Operatoren $\mathcal{A}_1 : L_2^n(\mathbb{R}_{++}, \nu_i) \rightarrow L_2^n(\mathbb{R}_{++}, \nu_i)$ und $\mathbf{b}_1 \in L_2^n(\mathbb{R}_{++}, \nu_i)$ mittels

$$\mathcal{A}_1(x, u) := x'(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot) \quad \text{und} \quad \mathbf{b}_1 := C(\cdot)$$

³Zur Gültigkeit der hier angegebenen Definition des Operators \mathcal{A}_1 verweisen wir auf die Ausführungen aus Kapitel 1.4.

und führen unter Verwendung dieser Operatoren nun die Menge X_1 ein durch

$$X_1 := \left\{ (x, u) \in W_2^{1,n}(\mathbb{R}_{++}, \nu_i) \times L_2^m(\mathbb{R}_{++}, \nu_i) \mid \mathcal{A}_1(x, u) = \mathbf{b}_1 \right\}.$$

Schritt 2: Als Funktional Φ wählen wir die Lagrange-Funktion, welche – wie wir in diesem Schritt zeigen werden – der Äquivalenzbedingung genügt^[4]. Dabei unterscheiden wir nun zwischen den Fällen $\nu = \nu_0$ und $\nu = \nu_1$ und definieren

$$\Phi_0((x, u), y) := J_{\nu_0}(x, u) + \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{(2), \nu_0}$$

für Probleme (P_{ν_0}) und

$$\Phi_1((x, u), y) := J_{\nu_1}(x, u) + \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{(2)}$$

für Probleme (P_{ν_1}) . Mit $\langle \cdot, \cdot \rangle_{\nu_i}$ bezeichnen wir im Folgenden $\langle \cdot, \cdot \rangle_{(2), \nu_0}$ für $i = 0$ und $\langle \cdot, \cdot \rangle_{(2)}$ für $i = 1$, so dass wir beide Fälle zusammenfassen können zu

$$\Phi_i((x, u), y) = J_{\nu_i}(x, u) + \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{\nu_i}, \quad i = 0, 1.$$

Zudem definieren wir die Mengen

$$Y_0 := L_2^n(\mathbb{R}_{++}, \nu_0) \quad \text{und} \quad Y_1 := L_2^n(\mathbb{R}_{++}, \nu_1^{-1}).$$

Bemerkung 14. Die Definition der Mengen Y_i garantiert, dass die Funktionen y in beiden Fällen Elemente von gewichteten Lebesgue-Räumen mit Dichtefunktion sind. Diese Eigenschaft wird uns im späteren Verlauf, insbesondere in Kapitel [5](#), von Nutzen sein.

Wir wollen nun zeigen, dass die Funktionale Φ_i (mit $i = 0, 1$) der Äquivalenzbedingung genügen und betrachten dazu

$$\sup_{y \in Y_i} \Phi_i((x, u), y) = \begin{cases} J_{\nu_i}(x, u), & (x, u) \in X_1 \\ +\infty, & (x, u) \notin X_1 \end{cases},$$

denn falls $\langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{\nu_i} = \alpha \neq 0$ für ein beliebiges Element y der Menge Y_i gilt, dann ist auch $ky \in Y_i$ für alle $k \in \mathbb{Z}$ und es folgt

$$\langle \mathcal{A}_1(x, u) - \mathbf{b}_1, ky \rangle_{\nu_i} = k\alpha \quad \text{und somit} \quad \sup_{y \in Y_i} \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{\nu_i} = \infty.$$

Gilt hingegen $\langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y \rangle_{\nu_i} = 0 \quad \forall y \in Y_i$, dann entspricht dies

$$\langle x'(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot) - C(\cdot), y(\cdot) \rangle_{\nu_i} = 0 \quad \forall y \in Y_i. \quad (3.12)$$

Wir beginnen mit der Betrachtung des Falles $\nu = \nu_0$: Offensichtlich gilt durch die Voraussetzungen (V1) und (V2) und die Raumzugehörigkeiten der Funktionen x und u

$$g(\cdot) := x'(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot) - C(\cdot) \in Y_0.$$

Damit muss die Relation [\(3.12\)](#) insbesondere auch für die Funktion g erfüllt sein und damit

$$\langle g(\cdot), g(\cdot) \rangle_{\nu_0} = 0,$$

⁴vgl. Bedingung [\(3.3\)](#), S. [47](#)

was nach der Definition des Skalarproduktes $g(t) = 0$ für fast alle $t \in \mathbb{R}_{++}$ impliziert⁵ und damit

$$x'(t) - A(t)x(t) - B(t)u(t) - C(t) = 0 \quad (\forall)t \in \mathbb{R}_{++}.$$

Wir betrachten nun den Fall $\nu = \nu_1$ und führen die Funktion $\tilde{y}(t) := y(t)e^{-\beta_1 t}$ ein. Für Funktionen $y \in Y_1 = L_2^n(\mathbb{R}_{++}, \nu_1^{-1})$ finden wir schließlich die folgende Abschätzung

$$\int_0^\infty |y(t)|^2 e^{-\beta_1 t} dt = \int_0^\infty |\tilde{y}(t)e^{\beta_1 t}|^2 e^{-\beta_1 t} dt = \int_0^\infty |\tilde{y}(t)|^2 e^{\beta_1 t} dt < \infty,$$

welche uns die Raumzugehörigkeit $\tilde{y} \in L_2^n(\mathbb{R}_{++}, \nu_1)$ liefert. Bemerkung 6⁶ erlaubt uns schließlich die äquivalente Formulierung der Relation (3.12) in der Form⁷

$$\langle x'(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot), \tilde{y}(\cdot) \rangle_{(2), \nu_1} = 0 \quad \forall \tilde{y} \in L_2^n(\mathbb{R}_{++}, \nu_1). \quad (3.13)$$

Wir können nun wieder in der gleichen Weise wie zuvor argumentieren:

Die Voraussetzungen (V1) und (V2) sowie die Raumzugehörigkeiten der Funktionen x und u liefern

$$g(\cdot) := x'(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot) \in L_2^n(\mathbb{R}_{++}, \nu_1).$$

Damit muss die Relation (3.13) insbesondere auch für die Funktion g erfüllt sein und damit

$$\langle g(\cdot), g(\cdot) \rangle_{(2), \nu_1} = 0,$$

was nach der Definition des Skalarproduktes $g(t) = 0$ für fast alle $t \in \mathbb{R}_{++}$ impliziert und damit

$$x'(t) - A(t)x(t) - B(t)u(t) = 0 \quad (\forall)t \in \mathbb{R}_{++}.$$

Damit gilt in beiden Fällen $(x, u) \in X_1$. Die Gültigkeit der Äquivalenzbedingung für Mengen Y_i und die Funktionale Φ_i (für $i = 0, 1$) wurde gezeigt.

Schritt 3: Gemäß des allgemeinen Konstruktionsschemas ist die zu (P_{ν_i}) duale Aufgabe gegeben durch

$$(\mathbf{D}_i) : \quad G_i(y) := \inf_{(x,u) \in X_0} \Phi_i((x,u), y) \longrightarrow \sup! \\ \text{u.d.N. } y \in Y_i$$

wobei die schwache Dualitätsrelation

$$\begin{aligned} \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x,u) &= \inf_{(x,u) \in X_0} \left[\sup_{y \in Y_i} \Phi_i((x,u), y) \right] \\ &\geq \sup_{y \in Y_i} \left[\inf_{(x,u) \in X_0} \Phi_i((x,u), y) \right] = \sup_{y \in Y_i} G_i(y) \end{aligned} \quad (3.14)$$

gilt.

⁵Diesen Schluss erhalten wir dabei insbesondere aus Beziehung $\langle x, x \rangle = 0 \Leftrightarrow x = 0$. Für eine ausführliche Angabe der Definition des Skalarproduktes siehe beispielsweise [Wer07], S. 157 Definition V.1.1

⁶siehe S. 13

⁷Es sei daran erinnert dass in diesem Falle gemäß Voraussetzung (V1) $C(t) \equiv 0$ gilt.

Weiterhin ist zu beachten⁸, dass mit (D_i) auch jedes Problem vom Typ

$$(\tilde{\mathbf{D}}_i) : \tilde{G}_i(y) \rightarrow \sup! \text{ u.d.N. } y \in \tilde{Y}_i \subset Y_i \quad (3.15)$$

mit

$$\tilde{G}_i(y) \leq G_i(y) \quad \forall y \in \tilde{Y}_i$$

(schwach) dual zu (P_{ν_i}) ist.

Entsprechend der Setzung aus der Relation (3.14) ist im Zuge der Formulierung der dualen Aufgabe ein untergeordnetes Optimierungsproblem zu lösen. Dieses untergeordnete Optimierungsproblem ist dabei zunächst selbst hinsichtlich Existenz und Eindeutigkeit von Optimallösungen sowie zugehöriger Optimalitätsbedingungen zu diskutieren.

Für die Problemklasse (P_{ν_0}) legen die Betrachtungen aus Kapitel 2.5 die Regularität der Funktion y nahe. Wir wollen daher die duale Aufgabe für beide Fälle zunächst unter ähnlichen Regularitätsannahmen herleiten und nutzen dazu die Relation (3.15). Im Hinblick auf die nachfolgenden Betrachtungen wollen wir zudem nur Funktionen y zulassen, welche in festem Zusammenhang zu den Anfangswerten der primalen Aufgabe stehen. Die nachfolgenden Setzungen erschließen sich im späteren Verlauf durch die Konstruktion. Wir definieren die folgenden Teilmengen von Y_i :

Für die Problemklasse (P_{ν_0}) definieren wir nun die Menge⁹

$$Y_{\nu_0} := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\beta_0 t}) \mid x_0 = W^{-1}(0) \left(-w(0) + y'(0) + (A^T(0) + \beta_0 \cdot I_{n \times n})y(0) \right) \right\}^{10}$$

Analog dazu betrachten wir für die Problemklasse (P_{ν_1}) die Menge

$$Y_{\nu_1} := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\beta_1 t}) \mid x_0 = W^{-1}(0) \left(-w(0) + y'(0) + A^T(0)y(0) \right) \right\}.$$

Die so konstruierten Mengen bilden Teilmengen der zuvor eingeführten Mengen Y_0 bzw. Y_1 . Daher kommen wir unter Berücksichtigung der Relation (3.15) zu folgendem Schluss:

Die Probleme (D_{ν_i}) ($i=0,1$) mit

$$(\mathbf{D}_{\nu_i}) : G_{\nu_i}(y) := \inf_{(x,u) \in X_0} \Phi_{\nu_i}((x,u), y) \longrightarrow \sup! \\ \text{u.d.N. } y \in Y_{\nu_i},$$

sind ebenfalls dual zu (P_{ν_i}) , denn da $Y_{\nu_i} \subseteq Y_i$, gilt für $i = 0, 1$ die folgende Relation

$$\sup_{y \in Y_{\nu_i}} G_{\nu_i}(y) \leq \sup_{y \in Y_i} G_i(y) \leq \inf(P_{\nu_i}).$$

⁸vgl. Relation 3.6, S. 47

⁹Die Raumzugehörigkeit der auftretenden Funktionen sichert die Wohldefiniertheit des rechten Ausdrucks.

¹⁰dabei bezeichne $I_{n \times n}$ die n -dimensionale Einheitsmatrix

3.3 Berechnung der dualen Aufgabe für (P_{ν_0})

3.3.1 Formulierung der untergeordneten Hilfsprobleme

In diesem Unterabschnitt widmen wir uns der Lösung der – im Zuge der Konstruktion der dualen Aufgabe erhaltenen – untergeordneten Optimierungsprobleme. Die vorgeschlagene Konstruktionsvorschrift führte uns zu einer Formulierung der dualen Aufgabe (D_{ν_0}) in linear-quadratischer Form. Diese Struktur ermöglicht auf der ersten Ebene eine Separierung und führt auf diese Weise zu der Formulierung zweier Optimierungsprobleme.

Im Zuge der Betrachtung der Problemklasse (P_{ν_0}) erhalten wir auf diese Weise als erstes Optimierungsproblem

(HP1) :

$$J_{1,y}^{\nu_0}(u) = \int_0^{\infty} \left\{ \frac{1}{2} u^T(t) R(t) u(t) + r^T(t) u(t) - [B(t)u(t) + C(t)]^T y(t) \right\} e^{\beta_0 t} dt \longrightarrow \inf!$$

$$\text{u.d.N. } u \in \mathbf{U} := \left\{ u \in L_2^m(\mathbb{R}_{++}, e^{\beta_0 t}) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_{++} \right\}.$$

Hierbei ist $y \in Y_{\nu_0}$ ein gegebener Parameter. Das zweite Optimierungsproblem lässt sich angeben durch

$$J_{2,y}^{\nu_0}(x) = \int_0^{\infty} \left\{ \frac{1}{2} x^T(t) W(t) x(t) + w^T(t) x(t) + [x'(t) - A(t)x(t)]^T y(t) \right\} e^{\beta_0 t} dt \longrightarrow \inf!$$

$$\text{u.d.N. } x \in \widetilde{X}_0 = \left\{ x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta_0 t}) \mid x(0) = x_0 \right\}.$$

Die Regularität der Funktion y erlaubt uns nun den folgenden Schritt:

Für $y \in Y_{\nu_0}$ finden wir mittels partieller Integration eine äquivalente Formulierung des Ausdruckes

$$\int_0^{\infty} x^T(t) y(t) e^{\beta_0 t} dt = \lim_{T \rightarrow \infty} x^T(T) y(T) e^{\beta_0 T} - x^T(0) y(0) - \int_0^{\infty} x^T(t) (\beta_0 y(t) + y'(t)) e^{\beta_0 t} dt.$$

Für die weitere Berechnung führen wir das folgende Lemma an:

Lemma 20. Seien $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ und $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ mit $\beta < 0$ dann gilt

$$\lim_{T \rightarrow \infty} x^T(T) y(T) e^{\beta T} = 0. \quad (3.16)$$

Beweis. In der Beweisführung des Lemmas folgen wir den Ideen aus [\[Mag82\]](#) Lemma 3.1 bzw. [\[Pic10a\]](#) Lemma 1: Für den Nachweis der Gültigkeit der Gleichung [\(3.16\)](#) zeigen wir für Elemente $x, y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ mit $\beta < 0$ zunächst unter Ausnutzung der Ungleichung von Cauchy-Schwarz

$$\begin{aligned} \int_0^{\infty} |x^T(t) y(t) e^{\beta t}| dt &\leq \left(\int_0^{\infty} |x(t)|^2 e^{\beta t} dt \right)^{1/2} \left(\int_0^{\infty} |y(t)|^2 e^{\beta t} dt \right)^{1/2} \\ &\leq \|x\|_{(2), e^{\beta t}} \cdot \|y\|_{(2), e^{\beta t}} < \infty. \end{aligned}$$

Aus der so nachgewiesenen Konvergenz des Ausdruckes $\int_0^{\infty} |x^T(t)y(t)e^{\beta t}| dt$ erhalten wir schließlich die Gültigkeit der Gleichung (3.16), da

$$\begin{aligned} \lim_{T \rightarrow \infty} \int_0^T x^T(t)y(t)e^{\beta t} dt &= \lim_{T \rightarrow \infty} \left(\int_0^{T-1} x^T(t)y(t)e^{\beta t} dt + \int_{T-1}^T x^T(t)y(t)e^{\beta t} dt \right) \\ &= \lim_{T \rightarrow \infty} \int_0^T x^T(t)y(t)e^{\beta t} dt + \lim_{\tau \rightarrow \infty} x^T(\tau)y(\tau)e^{\beta \tau}, \end{aligned}$$

wobei τ ein Element aus $[T-1, T]$ ist. \square

Die Anwendung des obigen Lemmas erlaubt somit die Formulierung

$$\int_0^{\infty} x'^T(t)y(t)e^{\beta_0 t} dt = -x^T(0)y(0) - \int_0^{\infty} x^T(t) (\beta_0 y(t) + y'(t)) e^{\beta_0 t} dt$$

und ermöglicht somit eine äquivalente Darstellung von $J_{2,y}^{\nu_0}$ in Form des folgenden Hilfsproblems

(HP2) :

$$\begin{aligned} J_{2,y}^{\nu_0}(x) &= \int_0^{\infty} \left\{ \frac{1}{2} x^T(t)W(t)x(t) + w^T(t)x(t) - x^T(t)[y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \right\} e^{\beta_0 t} dt \\ &\quad - x_0^T y(0) \longrightarrow \inf! \quad \text{u.d.N. } x \in \widetilde{X}_0 = \left\{ x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta_0 t}) \mid x(0) = x_0 \right\}. \end{aligned}$$

3.3.2 Existenz und Eindeutigkeit einer Lösung der Hilfsprobleme

Wir wollen die Optimierungsprobleme (HP1) und (HP2) zunächst auf die Existenz und Eindeutigkeit einer Lösung untersuchen. Gemäß der Betrachtungen aus Kapitel 2.4 handelt es sich bei dem Problem (HP1) um ein Optimierungsproblem über einer schwach kompakten Menge. Demnach können wir die Existenz und Eindeutigkeit mittels der Anwendung des verallgemeinerten Satzes von Weierstraß¹¹ sichern. Wir formulieren dafür den folgenden Satz:

Satz 12. *Unter Gültigkeit der Voraussetzungen (V1), (V2) und (V3) besitzt das Problem (HP1) eine eindeutige optimale Lösung.*

Beweis. Wir betrachten ein Optimierungsproblem über der (offensichtlich) nichtleeren Menge \mathbf{U} . Gemäß der Bedingung (V3) ist diese schwach kompakt¹². Für die Argumentation mittels Satz 7 genügt daher der Nachweis der schwachen Unterhalbstetigkeit des Zielfunktional. Dafür wählen wir wieder eine Einbettung der Aufgabe (HP1) in den Kontext der Hilbertraummethode und definieren für $\mathcal{H} := L_2^m(\mathbb{R}_{++}, \nu_0)$ die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ mit

$$\mathcal{Q}(u_1, u_2) := \int_0^{\infty} u_1^T(t)R(t)u_2(t)\nu_0(t)dt, \quad (3.17)$$

¹¹vgl. Satz 7, S. 30

¹²Der Beweis folgt unmittelbar aus den Betrachtungen der schwachen Kompaktheit des zulässigen Bereiches der primalen Aufgabe (vgl. Kapitel 2.4 Lemma 9, S. 32).

sowie den Operator $q : \mathcal{H} \rightarrow \mathbb{R}$

$$q(u) := \int_0^{\infty} u^T(t)[r(t) - B^T(t)y(t)]\nu_0(t)dt \quad (3.18)$$

und das Absolutglied

$$q_0 := - \int_0^{\infty} C^T(t)y(t)\nu_0(t)dt.$$

Mittels dieser Setzung können wir durch

$$J_{1,y}^{\nu_0}(u) = \frac{1}{2}\mathcal{Q}(u, u) + q(u) + q_0 \rightarrow \inf!$$

$$\text{u.d.N. } u \in \mathbf{U} = \{u \in L_2^m(\mathbb{R}_{++}, \nu_0) \mid u(t) \in \mathcal{U} (\forall) t \in \mathbb{R}_{++}\},$$

eine äquivalente Darstellung der Aufgabe (HP1) als Optimierungsproblem im Kontext der Hilbertraummethoden angeben.

Die nachfolgenden Lemmata geben Auskunft über die Eigenschaften der eingeführten Operatoren.

Lemma 21. *Durch die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.17), wird eine symmetrische, koerzitive Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt und stetig auf \mathcal{H} bezüglich der Normtopologie des zugrundeliegenden Raumes.*

Beweis. Es lässt sich leicht nachrechnen, dass \mathcal{Q} für alle $u_1, u_2 \in \mathcal{H}$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned} \mathcal{Q}(u_1 + \hat{u}_1, u_2) &= \mathcal{Q}(u_1, u_2) + \mathcal{Q}(\hat{u}_1, u_2) \\ \mathcal{Q}(u_1, u_2 + \hat{u}_2) &= \mathcal{Q}(u_1, u_2) + \mathcal{Q}(u_1, \hat{u}_2) \\ \mathcal{Q}(\lambda u_1, u_2) &= \lambda \mathcal{Q}(u_1, u_2). \end{aligned}$$

Aufgrund der Symmetrie der Matrix R gilt zudem für alle $u_1, u_2 \in \mathcal{H}$

$$\mathcal{Q}(u_1, u_2) = \mathcal{Q}(u_2, u_1).$$

Damit ist \mathcal{Q} eine symmetrische Bilinearform.

Wir wollen nun zeigen, dass die so definierte Abbildung auch koerzitiv ist. Dafür erinnern wir uns, dass aufgrund der positiven Definitheit der Matrix R gilt $\underline{R} > 0$ ¹³. Mit diesem Wissen betrachten wir nun die folgende Abschätzung:

$$\mathcal{Q}(u, u) = \int_0^{\infty} u^T(t)R(t)u(t)\nu_0(t)dt \geq \underline{R} \int_0^{\infty} u^2(t)\nu_0(t)dt \geq c \|u\|_{\mathcal{H}}^2,$$

mit der positiven Konstante $c = \underline{R} > 0$.

¹³Es gilt $\underline{R} = \inf_{t \in \mathbb{R}_+} \lambda_{\min}(R)(t)$. Dabei bezeichnet $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M .

Zudem sichert uns die Raumzugehörigkeit der Matrix R mit $R \in L_{\infty}^{m \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m})$ die Endlichkeit des Wertes

$$\|R\|_{\infty} = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |R(t)|$$

und liefert somit die Abschätzung

$$\begin{aligned} |Q(u_1, u_2)| &= \left| \int_0^{\infty} u_1^T(t) R(t) u_2(t) \nu_0(t) dt \right| \leq \|R\|_{\infty} \int_0^{\infty} |u_1^T(t) u_2(t)| \nu_0(t) dt \\ &\leq \|R\|_{\infty} \left(\int_0^{\infty} |u_1(t)|^2 \nu_0(t) dt \right)^{1/2} \left(\int_0^{\infty} |u_2(t)|^2 \nu_0(t) dt \right)^{1/2} \\ &\leq \|R\|_{\infty} \|u_1\|_{\mathcal{H}} \|u_2\|_{\mathcal{H}}, \end{aligned}$$

womit die Beschränktheit von Q auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist Q eine auf \mathcal{H} beschränkte Bilinearform und somit auch stetig in der Normtopologie¹⁴ des Raumes \mathcal{H} . \square

Lemma 22. *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.18), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Die Linearität der Abbildung q lässt sich leicht nachrechnen. Wir zeigen daher nun die Beschränktheit der Linearform q und treffen dazu die folgende Abschätzung

$$\begin{aligned} |q(u)| &= \left| \int_0^{\infty} u^T(t) [r(t) - B^T(t)y(t)] \nu_0(t) dt \right| \\ &\leq \left| \int_0^{\infty} u^T(t) r(t) \nu_0(t) dt \right| + \left| \int_0^{\infty} u^T(t) B^T(t) y(t) \nu_0(t) dt \right|. \end{aligned}$$

Durch die Raumzugehörigkeit der Funktionen r und B erhalten wir weiter

$$\begin{aligned} |q(u)| &\leq \left(\int_0^{\infty} |u(t)|^2 \nu_0(t) dt \right)^{1/2} \left(\int_0^{\infty} |r(t)|^2 \nu_0(t) dt \right)^{1/2} \\ &\quad + \|B\|_{\infty} \left(\int_0^{\infty} |u(t)|^2 \nu_0(t) dt \right)^{1/2} \left(\int_0^{\infty} |y(t)|^2 \nu_0(t) dt \right)^{1/2} \\ &\leq \|u\|_{\mathcal{H}} \|r\|_{\mathcal{H}} + \|B\|_{\infty} \|u\|_{\mathcal{H}} \|y\|_{\mathcal{H}} \leq c \|u\|_{\mathcal{H}} \quad \text{mit } c > 0, \end{aligned}$$

womit die Beschränktheit von q auf \mathcal{H} bewiesen ist. Damit ist q ein auf \mathcal{H} lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional und somit auch stetig in der Normtopologie des Raumes \mathcal{H} . \square

Die so nachgewiesenen Eigenschaften lassen nun direkte Schlüsse auf das Zielfunktional zu. Diese wollen wir im folgenden Lemma festhalten.

Lemma 23. *Das Zielfunktional des Problems (HP1) ist strikt konvex und schwach unterhalbstetig auf \mathcal{H} .*

¹⁴vgl. z.B. [TM09], S. 39

Beweis. Wir zeigen zunächst die strikte Konvexität des Zielfunktional und betrachten dafür zunächst den Operator \mathcal{Q} sowie beliebige Elemente $u_1, u_2 \in \mathcal{H}$ mit $u_1 \neq u_2$. Aufgrund der Koerzitivität der symmetrischen Bilinearform \mathcal{Q} existiert eine Konstante $c > 0$, so dass

$$0 < c \|u_1 - u_2\|_{\mathcal{H}}^2 \leq \mathcal{Q}(u_1 - u_2, u_1 - u_2) = \mathcal{Q}(u_1, u_1) - 2\mathcal{Q}(u_1, u_2) + \mathcal{Q}(u_2, u_2)$$

gilt und damit

$$2\mathcal{Q}(u_1, u_2) < \mathcal{Q}(u_1, u_1) + \mathcal{Q}(u_2, u_2).$$

Die Multiplikation der letzten Gleichung mit $\lambda(1 - \lambda)$, wobei $\lambda \in (0, 1)$ sei, liefert dann

$$2\lambda(1 - \lambda)\mathcal{Q}(u_1, u_2) < \lambda(1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_2, u_2).$$

Damit können wir für alle $u_1, u_2 \in \mathcal{H}$ die folgende Abschätzung finden:

$$\begin{aligned} \mathcal{Q}((1 - \lambda)u_1 + \lambda u_2, (1 - \lambda)u_1 + \lambda u_2) &= (1 - \lambda)^2 \mathcal{Q}(u_1, u_1) + 2\lambda(1 - \lambda)\mathcal{Q}(u_1, u_2) + \lambda^2 \mathcal{Q}(u_2, u_2) \\ &< (1 - \lambda)^2 \mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_2, u_2) + \lambda^2 \mathcal{Q}(u_2, u_2) \\ &= (1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda\mathcal{Q}(u_2, u_2). \end{aligned}$$

Da q linear ist, erhalten wir durch $J_{1,y}^{\nu_0}(u) = \frac{1}{2}\mathcal{Q}(u, u) + q(u) + q_0$ die strikte Konvexität des Zielfunktional. Als Summe (bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$) stetiger Funktionale ist auch das Zielfunktional stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . Die schwache Unterhalbstetigkeit folgt dann gemäß der Beweisführung von Lemma 8 aus der strikten Konvexität. □

Damit betrachten wir also ein schwach unterhalbstetiges Funktional über einer schwach kompakten Menge. Die Existenz einer Lösung ist damit nach Satz 7 gesichert. Die strikte Konvexität des Zielfunktional sichert zudem die Eindeutigkeit der Lösung¹⁵. Damit ist die Behauptung des Satzes bewiesen. □

Die Existenz und Eindeutigkeit einer Lösung für (HP1) konnten wir aus dem verallgemeinerten Satz von Weierstraß sowie der strikten Konvexität des Zielfunktional schlussfolgern. Die Problemstellung (HP2) stellt allerdings ein Optimierungsproblem über einer nicht (schwach) kompakten Menge dar und erfordert damit wieder den Nachweis beschränkter Niveaumengen in der entsprechenden Topologie. Die in (HP2) geforderte Topologie des Raumes $W_2^{1,n}(\mathbb{R}_{++}, \nu_0)$ scheint im Hinblick auf das verwendete Zielfunktional jedoch zu stark. Daher wollen wir stattdessen die Existenz und Eindeutigkeit einer Lösung x_{*y} des Problems

(HP2) :

$$\begin{aligned} J_{2,y}^{\nu_0}(x) &= \int_0^{\infty} \left\{ \frac{1}{2}x^T(t)W(t)x(t) + w^T(t)x(t) - x^T(t)[y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \right\} e^{\beta_0 t} dt \\ &\quad - x_0^T y(0) \longrightarrow \inf! \quad \text{u.d.N. } x \in L_2^n(\mathbb{R}_{++}, e^{\beta_0 t}) \end{aligned} \quad (3.19)$$

¹⁵vgl. hierfür beispielsweise [Bur19], Satz 4.13, S. 96

in der Topologie des Raumes $L_2^n(\mathbb{R}_{++}, \nu_0)$ diskutieren¹⁶. Können wir im Anschluss nachweisen, dass die eindeutige Lösung x_{*y} von $(\overline{HP2})$ auch Element des Raumes $W_2^{1,n}(\mathbb{R}_{++}, \nu_0)$ ist und dort der Anfangsbedingung $x_{*y}(0) = x_0$ genügt, so liefert uns das die Existenz und Eindeutigkeit der Lösung x_{*y} für $(HP2)$.

Wir betrachten daher zunächst das Problem $(\overline{HP2})$ und formulieren dazu folgenden Satz:

Satz 13. *Unter Gültigkeit der Voraussetzungen (V1) und (V2) besitzt das Problem $(\overline{HP2})$ eine eindeutige optimale Lösung.*

Beweis. Wir behandeln die Aufgabe $(\overline{HP2})$ und wollen diese wieder in den Kontext der Hilbertraummethode einbetten. Dazu definieren wir auf dem Hilbertraum $\mathcal{H} = L_2^n(\mathbb{R}_{++}, \nu_0)$ die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ mit

$$\mathcal{Q}(x_1, x_2) := \int_0^\infty x_1^T(t)W(t)x_2(t)\nu_0(t)dt, \quad (3.20)$$

sowie $q : \mathcal{H} \rightarrow \mathbb{R}$ mit

$$q(x) := - \int_0^\infty x^T(t)[w(t) - y'(t) - (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)]\nu_0(t)dt. \quad (3.21)$$

Zudem sei für beliebige aber feste Werte $y(0)$ die Konstante

$$q_0 := x_0^T y(0)$$

gegeben. Mittels dieser Setzungen können wir das Zielfunktional von $(HP2)$ gemäß

$$J_{2,y}^{\nu_0}(x) = \frac{1}{2}\mathcal{Q}(x, x) - q(x) - q_0$$

beschreiben. Wir halten die Eigenschaften der eingeführten Abbildungen in den folgenden zwei Lemmata fest.

Lemma 24. *Durch die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.20), wird eine symmetrische, koerzitive Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt und stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Der Beweis verläuft analog zu der Beweisführung von Lemma 21 und wird daher an dieser Stelle ausgelassen. Bei Bedarf kann er in Kapitel B.1 nachvollzogen werden. \square

Lemma 25. *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.21), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Der Beweis verläuft analog zu der Beweisführung von Lemma 22 und wird daher an dieser Stelle ausgelassen. Bei Bedarf kann er in Kapitel B.1 nachvollzogen werden. \square

Die so nachgewiesenen Eigenschaften lassen nun direkte Schlüsse auf das Zielfunktional zu. Diese wollen wir im folgenden Lemma festhalten.

¹⁶Durch die derartige Raumwahl ist eine punktweise Auswertung der Funktion x nicht mehr möglich. Daher müssen wir auf die Verwendung der Anfangsbedingung $x(0) = x_0$ verzichten

Lemma 26. *Das Zielfunktional des Problems $(\overline{HP2})$ ist strikt konvex und schwach unterhalbstetig auf \mathcal{H} .*

Beweis. Der Beweis verläuft analog zu der Beweisführung von Lemma 23 und wird daher an dieser Stelle ausgelassen. Bei Bedarf kann er in Kapitel B.1 nachvollzogen werden. \square

Wir wollen nun zu einer äquivalenten Problemformulierung von $(\overline{HP2})$ über einer schwach kompakten Menge übergehen. Dazu betrachten wir die Niveaumengen

$$\mathcal{N}(J_{2,y}^{\nu_0}, \alpha) := \left\{ x \in L_2^n(\mathbb{R}_{++}, \nu_0) \mid J_{2,y}^{\nu_0}(x) - \mu \leq \alpha \right\}, \quad (3.22)$$

wobei

$$\mu := \inf_{x \in L_2^n(\mathbb{R}_{++}, \nu_0)} J_{2,y}^{\nu_0}(x) \quad (3.23)$$

und weisen für diese die folgende Eigenschaft nach:

Lemma 27. *Die Niveaumengen $\mathcal{N}(J_{2,y}^{\nu_0}, \alpha)$ ($\alpha > 0$), definiert in (3.22), sind beschränkt (bzgl. der Normtopologie des Raumes \mathcal{H}).*

Beweis. Die Aussagen von Lemma 24 und Lemma 25 sichern die Endlichkeit des Wertes μ aus (3.23), denn es gilt

$$\begin{aligned} J_{2,y}^{\nu_0}(x) &= \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0 \geq \frac{1}{2} c_1 \|x\|_{\mathcal{H}}^2 - c_2 \|x\|_{\mathcal{H}} - c_3 \\ &= \frac{1}{2} c_1 \left(\|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right)^2 - \frac{c_2^2}{2c_1} - c_3 \geq -\frac{c_2^2}{2c_1} - c_3 > -\infty, \end{aligned}$$

mit $c_1, c_2, c_3 > 0$. Damit können wir nun die folgende Abschätzung für alle Elemente x der Niveaumengen $\mathcal{N}(J_{2,y}^{\nu_0}, \alpha)$ treffen:

$$\alpha \geq J_{2,y}^{\nu_0}(x) - \mu \geq \frac{1}{2} c_1 \left(\|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right)^2 - \frac{c_2^2}{2c_1} - c_3 - \mu.$$

Dies impliziert insbesondere die Gültigkeit von

$$\left| \|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right| \leq \left[\frac{2}{c_1} (\alpha + \mu) + \left(\frac{c_1}{c_2} \right)^2 + \frac{2c_3}{c_1} \right]^{1/2}.$$

Damit ist $\mathcal{N}(J_{2,y}^{\nu_0}, \alpha)$ beschränkt auf \mathcal{H} bezüglich der Normtopologie des Raumes \mathcal{H} . \square

Aus der Beweisführung des obigen Lemmas können wir nun schlussfolgern: $\|x\|_{\mathcal{H}}$ ist beschränkt, wenn wir nur derartige x betrachten, deren Zielfunktionalwerte J in einer α -Umgebung von μ liegen. Die Ergebnisse der Lemmata lassen daher nun den Schluss zu, dass das Problem

$$\begin{aligned} (\overline{HP2}') : \quad J_{2,y}^{\nu_0}(x) &= \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0 \longrightarrow \inf ! \\ \text{u.d.N.} \quad x &\in L_2^n(\mathbb{R}_{++}, e^{\beta_0 t}) \\ &\|x\|_{\mathcal{H}} \leq \kappa, \kappa \text{ hinreichend groß,} \end{aligned}$$

äquivalent zu $(\overline{HP2})$ ist. Dabei existiert für $(\overline{HP2}')$, als ein Optimierungsproblem mit schwach unterhalbstetigem Zielfunktional über einer schwach kompakten Menge, nach Satz 7 eine optimale Lösung x_{*y} . Diese ist aufgrund der strikten Konvexität des Zielfunktionalen eindeutig. Damit ist x_{*y} auch eindeutige Lösung von $(\overline{HP2})$ und die Aussage des Satzes somit bewiesen. \square

3.3.3 Optimalitätsbedingungen der Hilfsprobleme

Lösung des Hilfsproblems (HP1)

Satz 12 sichert die Existenz und Eindeutigkeit einer optimalen Lösung für das Problem (HP1). Ziel ist nun die Diskussion notwendiger und hinreichender Optimalitätsbedingungen zur Bestimmung der eindeutigen optimalen Lösung u_{*y} . Für eine optimale Lösung der restringierten Aufgabe (HP1) gibt das folgende Lemma Auskunft:

Lemma 28 ([Trö05], Lemma 2.20, S. 50). *Es sei U ein reeller Banachraum, $C \subset U$ eine konvexe Menge und $F : C \rightarrow \mathbb{R}$ ein auf C Gâteaux-differenzierbares Funktional. Mit u_* sei eine Lösung der Aufgabe*

$$\min_{u \in C} F(u)$$

gegeben. Dann ist die folgende Variationsungleichung erfüllt:

$$\delta F(u; u - u_*) \geq 0 \quad \forall u \in C. \quad (3.24)$$

Bemerkung 15. Die Setzung $\vartheta := u - u_* \in C$ führt uns zu der in Definition 15 angegebenen Formulierung des Gâteaux-Differentials 17. Aufgrund der strikten Konvexität des Funktionals $J_{1,y}^{\nu_0}$ sind die Bedingungen des Lemmas 28 notwendig und hinreichend.

Die Auswertung der Variationsungleichung (3.24) erfordert die Berechnung des Gâteaux-Differentials des Zielfunktional von (HP1). Für dessen Angabe erinnern wir an die Betrachtungen aus Kapitel 1.5, Satz 6, gemäß derer wir das Gâteaux-Differential mit dem zugehörigen Fréchet-Differential identifizieren können. Dabei erlaubt die Struktur von $J_{1,y}^{\nu_0}$ als quadratisches Funktional im Hilbertraum 18 die Angabe des Fréchet- und (nach Satz 6 damit auch) des Gâteaux-Differentials durch die lineare und (bzgl. der Normtopologie des Raumes $L_2^m(\mathbb{R}_{++}, \nu_0)$) stetige Abbildung $\vartheta \mapsto \mathcal{Q}(u_{*y}, \vartheta) + q(\vartheta)$ der Form

$$\delta J_{1,y}^{\nu_0}(u; \vartheta) = \int_0^\infty \vartheta^T(t) \left\{ R(t)u(t) + r(t) - B^T(t)y(t) \right\} e^{\beta_0 t} dt$$

für alle $\vartheta \in L_2^m(\mathbb{R}_{++}, \nu_0)$.

Auswertung der Optimalitätsbedingungen

Gemäß unserer vorherigen Berechnungen ergibt sich für die Setzung $\vartheta(t) := u(t) - u_{*y}(t)$ für fast alle $t \in \mathbb{R}_{++}$ als notwendige und hinreichende Bedingung für eine optimale Lösung u_{*y} die Variationsungleichung

$$\int_0^\infty \left(R(t)u_{*y}(t) + r(t) - B^T(t)y(t) \right)^T (u(t) - u_{*y}(t)) e^{\beta_0 t} dt \geq 0 \quad \forall u \in \mathbf{U}. \quad (3.25)$$

Für diese können wir das folgende Lemma angeben:

Lemma 29. *Die Variationsungleichung (3.25) gilt genau dann, wenn für fast alle $t \in \mathbb{R}_{++}$ die punktweise Variationsungleichung in \mathbb{R} erfüllt ist:*

$$\left(R(t)u_{*y}(t) + r(t) - B^T(t)y(t) \right)^T (v - u_{*y}(t)) e^{\beta_0 t} \geq 0 \quad \forall v \in [\underline{u}, \bar{u}].$$

¹⁷ siehe Definition 15 S. 17

¹⁸ vgl. Kapitel 1.5, Beispiel 2, S. 19

Beweis. Wir setzen $z(t) = (R(t)u_{*y}(t) + r(t) - B^T(t)y(t)) e^{\beta_0 t}$ und folgen der Beweisführung aus [Tr05] zu Lemma 2.26, S. 54. Dafür ist zu zeigen, dass für fast alle $t \in \mathbb{R}_{++}$ die Relation $z^T(t)(v - u_{*y}(t)) \geq 0$ erfüllt ist und zwar für alle $v \in [\underline{u}, \bar{u}]$. Diesen Nachweis führen wir wie folgt:

Aufgrund der Raumzugehörigkeiten der eingehenden Funktionen¹⁹ ist die Funktion z messbar auf \mathbb{R}_{++} . Damit sind fast alle Punkte auf \mathbb{R}_{++} Lebesgue-Punkte von z , d.h. für fast alle $t_0 \in \mathbb{R}_{++}$ gilt²⁰

$$\lim_{\rho \rightarrow 0} \frac{1}{|B(t_0, \rho)|} \int_{B(t_0, \rho)} z(t) dt = z(t_0).$$

Für hinreichend kleines $\rho \in \mathbb{R}_{++}$ ist das Intervall $B(t_0, \rho)$ vollständig in \mathbb{R}_{++} enthalten. Damit ist das Integral wohldefiniert. Die gleiche Argumentation gilt für u_{*y} . Damit sind fast alle Punkte $t \in \mathbb{R}_{++}$ gemeinsame Lebesgue-Punkte von z und u_{*y} . Wir wählen einen gemeinsamen Lebesgue-Punkt $t_0 \in \mathbb{R}_{++}$ beliebig aus sowie $v \in [\underline{u}, \bar{u}]$. Außerdem definieren wir

$$u(t) := \begin{cases} v & \text{in } B(t_0, \rho) \\ u_{*y}(t) & \text{sonst.} \end{cases}$$

Damit gehört $u(\cdot)$ offenbar zu $[\underline{u}, \bar{u}]$ und die Variationsungleichung (3.25) liefert nach der Division durch $|B(t_0, \rho)|$

$$0 \leq \frac{1}{|B(t_0, \rho)|} \int_0^\infty z^T(t)(u(t) - u_{*y}(t)) dt = \frac{1}{|B(t_0, \rho)|} \int_{B(t_0, \rho)} z^T(t)(u(t) - u_{*y}(t)) dt.$$

Der Grenzübergang von $\rho \rightarrow 0$ ergibt dann $0 \leq z^T(t_0)(u(t_0) - u_{*y}(t_0))$. Dies gilt für fast alle $t_0 \in \mathbb{R}_{++}$. Damit ist die Behauptung des Satzes bewiesen. \square

Die Gültigkeit des obigen Lemmas erlaubt uns nun die Berechnung der optimalen Lösung u_{*y} durch die punktweise Auswertung der Variationsungleichung, d.h. wir berechnen für fast alle $t \in \mathbb{R}_{++}$

$$(R(t)u_{*y}(t) + r(t) - B^T(t)y(t))^T u_{*y}(t) = \min_{v \in [\underline{u}, \bar{u}]} \sum_{i=1}^m (R(t)u_{*y}(t) + r(t) - B^T(t)y(t))_i v_i.$$

Aus der obigen Gleichung erhalten wir die komponentenweise Bedingung

$$u_{i* y}(t) = \begin{cases} \bar{u}_i, & \text{falls } (R(t)u_{*y}(t) + r(t) - B^T(t)y(t))_i < 0 \\ \in [\underline{u}_i, \bar{u}_i], & \text{falls } (R(t)u_{*y}(t) + r(t) - B^T(t)y(t))_i = 0 \\ \underline{u}_i, & \text{falls } (R(t)u_{*y}(t) + r(t) - B^T(t)y(t))_i > 0 \end{cases} \quad (\forall) t \in \mathbb{R}_{++}.$$

und damit

$$u_{*y}(t) = \text{Proj}_{\mathcal{U}} \left(R^{-1}(t)(-r(t) + B^T(t)y(t)) \right) \quad (\forall) t \in \mathbb{R}_{++}.$$

¹⁹vgl. Voraussetzungen (V1), (V2), S. 25

²⁰ $B(t, \rho)$ bezeichnet das offene Intervall um t mit dem Radius ρ .

Lösung des Hilfsproblems $(\overline{HP2})$

Wir behandeln die Aufgabe $(\overline{HP2})$ gemäß ihrer Einbettung in den Kontext der Hilbertraummethoden in der Form

$$J_{2,y}^{\nu_0}(x) = \frac{1}{2} \mathcal{Q}(x, x) - q(x) - x_0^T y(0) \rightarrow \inf!$$

$$\text{u.d.N. } x \in L_2^n(\mathbb{R}_{++}, \nu_0).$$

Die Struktur der Aufgabe sichert nach Satz [13](#) die Existenz und Eindeutigkeit einer optimalen Lösung für das Problem $(\overline{HP2})$. Die positive Definitheit der Bilinearform \mathcal{Q} sichert zudem die strikte Konvexität des Zielfunktional und rechtfertigt damit die Betrachtung der notwendigen (und für unseren Fall durch die Konvexität auch hinreichenden) Bedingung in lokaler Form des folgenden Satzes:

Satz 14 ([\[Zei12\]](#), Satz 2.A., S. 45 (Auszug)). *Es sei X ein reeller normierter Raum. Sei das Funktional $F : U(u_0) \subseteq X \rightarrow \mathbb{R}$ in einer offenen Umgebung $U(u_0)$ des Punktes u_0 gegeben. Dann gelten die folgenden Aussagen:*

(i) **Notwendige Bedingung.** *Falls F ein lokales Minimum im Punkt u_0 besitzt, dann ist u_0 ein kritischer Punkt von F und es gilt*

$$\delta F(u_0; h) = 0 \quad \forall h \in X. \quad (3.26)$$

Zur Auswertung der Optimalitätsbedingungen [\(3.26\)](#) benötigen wir demnach das Gâteaux-Differential. Gemäß der Betrachtungen aus Kapitel [1.5](#) können wir dieses wieder durch das zugehörige Fréchet-Differential beschreiben^{[21](#)}. Die Struktur von $J_{2,y}^{\nu_0}$ als quadratisches Funktional im Hilbertraum liefert die Darstellung des (Fréchet- und) Gâteaux-Differentials durch die lineare und (bzgl. der Normtopologie des Raumes $L_2^n(\mathbb{R}_{++}, \nu_0)$) stetige Abbildung $\zeta \mapsto \mathcal{Q}(x_{*y}, \zeta) + q(\zeta)$ der Form

$$\delta J_{2,y}^{\nu_0}(x_{*y}; \zeta) = \int_0^\infty \zeta^T(t) \left\{ W(t)x_{*y}(t) + w(t) - [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \right\} e^{\beta_0 t} dt$$

für alle $\zeta \in L_2^n(\mathbb{R}_{++}, \nu_0)$.

Auswertung der Optimalitätsbedingungen

Gemäß Satz [14](#) setzen wir für die Berechnung der optimalen Lösung

$$\delta J_{2,y}^{\nu_0}(x_{*y}; \zeta) \stackrel{!}{=} 0 \quad \forall \zeta \in L_2^n(\mathbb{R}_{++}, \nu_0) \Leftrightarrow$$

$$\int_0^\infty \zeta^T(t) \left\{ W(t)x_{*y}(t) + w(t) - [y'(t) + (A^T(t) + \beta_0 I_{n \times n})y(t)] \right\} e^{\beta_0 t} dt \stackrel{!}{=} 0 \quad \forall \zeta \in L_2^n(\mathbb{R}_{++}, \nu_0).$$

Diese Bedingung ist äquivalent zu der Darstellung

$$\left\langle \zeta, Wx_{*y} + w - [y' + (A^T + \beta_0 I_{n \times n})y] \right\rangle_{\nu_0} = 0 \quad \forall \zeta \in L_2^n(\mathbb{R}_{++}, \nu_0). \quad (3.27)$$

²¹vgl. Satz [6](#), S. [18](#)

Die Raumzugehörigkeit der auftretenden Funktionen impliziert, dass Gleichung (3.27) insbesondere auch für Funktionen

$$\zeta(t) = W(t)x_{*y}(t) + w(t) - [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \quad (\forall)t \in \mathbb{R}_{++}$$

erfüllt sein muss und damit $\langle \zeta, \zeta \rangle_{\nu_0} = 0$. Die Definition des Skalarproduktes impliziert dann $\zeta(t) = 0$ für fast alle $t \in \mathbb{R}_{++}$ und somit

$$W(t)x_{*y}(t) + w(t) - [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] = 0 \quad (\forall)t \in \mathbb{R}_{++}. \quad (3.28)$$

Wir erhalten die eindeutige Lösung x_y^* des Hilfsproblems ($\overline{HP2}$) somit aus der Gleichung (3.28) und können diese angeben durch

$$x_{*y}(t) = W^{-1}(t) \left(-w(t) + [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \right) \quad (\forall)t \in \mathbb{R}_{++}. \quad (3.29)$$

Aus den Raumzugehörigkeiten der Funktionen W , A , w und y folgt, dass die rechte Seite der Gleichung (3.29) Element des Raumes $W_2^{1,n}(\mathbb{R}_{++}, \nu_0)$ ist. Damit ist Gleichung (3.29) auch punktweise erfüllt und es gilt $x_{*y} \in W_2^{1,n}(\mathbb{R}_{++}, \nu_0)$.

Weiterhin sichert die Bedingung $y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\beta_0 t})$ die Existenz der klassischen Ableitung²² von $y'(t)$ für alle $t \in \mathbb{R}_+$ und somit insbesondere auch für $t = 0$. Damit können wir Gleichung (3.29) auch an dieser Stelle auswerten und erhalten

$$x_{*y}(0) = W^{-1}(0) \left(-w(0) + y'(0) + (A^T(0) + \beta_0 \cdot I_{n \times n})y(0) \right). \quad (3.30)$$

Da $y \in Y_{\nu_0}$ ist, gilt weiterhin

$$x_0 = W^{-1}(0) \left(-w(0) + y'(0) + (A^T(0) + \beta_0 \cdot I_{n \times n})y(0) \right), \quad (3.31)$$

so dass die Bedingungen aus Gleichung (3.30) und (3.31) die Relation $x_{*y}(0) = x_0$ ergeben.

Insgesamt konnten wir zeigen, dass für beliebige, aber feste $y \in Y_{\nu_0}$ eine eindeutige Lösung $x_{*y} \in W_2^{1,n}(\mathbb{R}_{++}, \nu_0)$ für das Problem ($\overline{HP2}$) existiert, welche der Anfangsbedingung $x_{*y}(0) = x_0$ genügt. Damit ist x_{*y} auch eindeutige Lösung von ($HP2$).

Bemerkung 16. Bei genauerer Betrachtung der optimalen Lösung x_{*y} stellen wir fest, dass diese mit $y = \tilde{y}$ der kanonischen Gleichung (\tilde{K}) aus Satz 11 entspricht²³.

3.4 Berechnung der dualen Aufgabe für (P_{ν_1})

3.4.1 Formulierung der untergeordneten Hilfsprobleme

Wir wollen in einer ähnlichen Art und Weise zum vorherigen Abschnitt nun auch eine Formulierung einer zu (P_{ν_1}) dualen Aufgabe angeben:

Auch hier stellt die erhaltene duale Aufgabe (D_{ν_1}) ein mehrstufiges Optimierungsproblem dar, das auf der ersten Ebene separiert werden kann. Wir erhalten die folgende Darstellung der untergeordneten Hilfsprobleme

²²vgl. Bemerkung 4, S. 12

²³siehe Satz 11, S. 41

(HP3) :

$$J_{1,y}^{\nu_1}(u) = \int_0^{\infty} \left\{ \left(\frac{1}{2} u^T(t) R(t) u(t) + r^T(t) u(t) \right) e^{\beta_1 t} - [B(t)u(t)]^T y(t) \right\} dt \longrightarrow \inf!$$

u.d.N. $u \in \mathbf{U} := \left\{ u \in L_2^m(\mathbb{R}_{++}, e^{\beta_1 t}) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_{++} \right\}$.

Hierbei ist $y \in Y_{\nu_1}$ wieder ein beliebiger, aber fester Parameter. Das 2. Optimierungsproblem ist

$$J_{2,y}^{\nu_1}(x) = \int_0^{\infty} \left\{ \left(\frac{1}{2} x^T(t) W(t) x(t) + w^T(t) x(t) \right) e^{\beta_1 t} + [x'(t) - A(t)x(t)]^T y(t) \right\} dt \longrightarrow \inf!$$

u.d.N. $x \in \widetilde{X}_0 = \left\{ x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta_1 t}) \mid x(0) = x_0 \right\}$.

Für $y \in Y_{\nu_1}$ finden wir wieder mittels partieller Integration eine äquivalente Formulierung des Ausdruckes

$$\int_0^{\infty} x'^T(t) y(t) dt = \lim_{T \rightarrow \infty} x^T(T) y(T) - x^T(0) y(0) - \int_0^{\infty} x^T(t) y'(t) dt.$$

Für die weitere Berechnung führen wir das folgende Lemma an:

Lemma 30. Sei $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ und $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$ mit $\beta > 0$ dann gilt

$$\lim_{T \rightarrow \infty} x^T(T) y(T) = 0. \quad (3.32)$$

Beweis. Wir führen den Beweis wieder mit den Ideen aus [Mag82](#) Lemma 3.1 bzw. [Pic10a](#) Lemma 1: Für den Nachweis der Gültigkeit der Gleichung [\(3.32\)](#) zeigen wir für Elemente $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta t})$ und $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\beta t})$ mit $\beta > 0$ zunächst unter Ausnutzung der Ungleichung von Cauchy-Schwarz

$$\begin{aligned} \int_0^{\infty} |x^T(t) y(t)| dt &= \int_0^{\infty} |x^T(t) e^{\frac{1}{2}\beta t} y(t) e^{-\frac{1}{2}\beta t}| dt \\ &\leq \left(\int_0^{\infty} |x(t)|^2 e^{\beta t} dt \right)^{1/2} \left(\int_0^{\infty} |y(t)|^2 e^{-\beta t} dt \right)^{1/2} \\ &\leq \|x\|_{(2), e^{\beta t}} \cdot \|y\|_{(2), e^{-\beta t}} < \infty. \end{aligned}$$

Aus der so nachgewiesenen Konvergenz des Ausdruckes $\int_0^{\infty} |x^T(t) y(t)| dt$ erhalten wir schließlich die Gültigkeit der Gleichung [\(3.32\)](#), da

$$\begin{aligned} \lim_{T \rightarrow \infty} \int_0^T x^T(t) y(t) dt &= \lim_{T \rightarrow \infty} \left(\int_0^{T-1} x^T(t) y(t) dt + \int_{T-1}^T x^T(t) y(t) dt \right) \\ &= \lim_{T \rightarrow \infty} \int_0^T x^T(t) y(t) dt + \lim_{\tau \rightarrow \infty} x^T(\tau) y(\tau), \end{aligned}$$

wobei τ ein Element aus $[T-1, T]$ ist. □

Die Anwendung des obigen Lemmas erlaubt somit die Formulierung

$$\int_0^{\infty} x'^T(t)y(t)dt = -x^T(0)y(0) - \int_0^{\infty} x^T(t)y'(t)dt$$

und ermöglicht damit eine äquivalente Darstellung von $J_{2,y}^{\nu_1}$ in der Form

(HP4) :

$$J_{2,y}^{\nu_1}(x) = \int_0^{\infty} \left\{ \left(\frac{1}{2} x^T(t)W(t)x(t) + w^T(t)x(t) \right) e^{\beta_1 t} - x^T(t)[y'(t) + A^T(t)y(t)] \right\} dt$$

$$-x_0^T y(0) \longrightarrow \inf!$$

$$\text{u.d.N. } x \in \widetilde{X}_0 = \left\{ x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\beta_1 t}) \mid x(0) = x_0 \right\}.$$

3.4.2 Existenz und Eindeutigkeit einer Lösung der Hilfsprobleme

Wir wollen die Optimierungsprobleme (HP3) und (HP4) ebenfalls auf die Existenz und Eindeutigkeit ihrer Optimallösungen untersuchen. Offenbar sind sich die Hilfsaufgaben (HP1) und (HP3) sowie (HP2) und (HP4) in ihrer Struktur sehr ähnlich. Daher ist es wenig überraschend, dass wir eine ähnliche Argumentation verfolgen werden.

Im Gegensatz zu der Problemformulierung in (HP1) ist die schwache Kompaktheit der Menge \mathbf{U} jedoch nicht aus den Betrachtungen der primalen Aufgabe abzuleiten. Diese sichert uns zunächst nur die schwache Abgeschlossenheit²⁴ der Menge \mathbf{U} . Für den Nachweis der Existenz einer optimalen Lösung wählen wir daher auch hier den Weg über die Konstruktion schwach kompakter Niveaumengen und können letztlich die folgenden Schlüsse ziehen:

Satz 15. *Unter Gültigkeit der Voraussetzungen (V1), (V2) und (V3) besitzt das Problem (HP3) eine eindeutige optimale Lösung.*

Beweis. Siehe Kapitel [B.2](#). □

Wir betrachten nun das Hilfsproblem (HP4). Ähnlich wie im Falle der Diskussion von (HP2) führt unsere Argumentation bzgl. etwaiger Existenz- und Eindeutigkeitsaussagen über den Nachweis der Beschränktheit der zugehörigen Niveaumengen. Diesen Nachweis können wir wiederum mit der Koerzitivität des Zielfunktional in der entsprechenden Topologie führen. Auch hier scheint die Topologie des Raumes $W_2^{1,n}(\mathbb{R}_{++}, \nu_1)$ zu stark, so dass wir zu der schwächeren Formulierung im $L_2^n(\mathbb{R}_{++}, \nu_1)$ übergehen wollen. Wir erhalten damit eine Problemformulierung der Art

(HP4) :

$$J_{2,y}^{\nu_1}(x) = \int_0^{\infty} \left\{ \left(\frac{1}{2} x^T(t)W(t)x(t) + w^T(t)x(t) \right) e^{\beta_1 t} - x^T(t)[y'(t) + A^T(t)y(t)] \right\} dt$$

$$-x_0^T y(0) \longrightarrow \inf! \quad \text{u.d.N. } x \in L_2^n(\mathbb{R}_{++}, e^{\beta_1 t}).$$

²⁴vgl. Kapitel [2.4](#), Beweis zu Lemma [9](#), Schritt 1 und Schritt 2 S. [32](#)

Auch in diesem Fall erhalten wir letztendlich erst Rückschlüsse auf die Aufgabe $(HP4)$, wenn wir in der Lage sind nachzuweisen, dass eine eindeutige Lösung x_{*y} des Problems $(\overline{HP4})$ Element des Raumes $W_2^{1,n}(\mathbb{R}_{++}, \nu_1)$ ist und der Anfangsbedingung $x_{*y}(0) = x_0$ genügt. Für die Aufgabe $(\overline{HP4})$ liefert der folgende Satz Aussagen hinsichtlich Existenz und Eindeutigkeit der Lösung:

Satz 16. *Unter Gültigkeit der Voraussetzungen $(V1)$ und $(V2)$ besitzt das Problem $(\overline{HP4})$ eine eindeutige optimale Lösung.*

Beweis. Siehe Kapitel [B.3](#). □

3.4.3 Optimalitätsbedingungen der Hilfsprobleme

Lösung des Hilfsproblems $(HP3)$

Die Betrachtung dieses Hilfsproblems können wir ähnlich zur Behandlung des Problems $(HP1)$ führen. Daher betten wir auch das Problem $(HP3)$ in den Kontext der Hilbertraummethoden ein, mittels

$$J_{1,y}^{\nu_1}(u) = \frac{1}{2} \mathcal{Q}(u, u) + q(u) \rightarrow \inf!$$

$$\text{u.d.N. } u \in \mathbf{U} := \left\{ u \in L_2^m(\mathbb{R}_{++}, e^{\beta_1 t}) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_{++} \right\},$$

mit der positiv definiten, bzgl. der Norm $\|\cdot\|_{(2),\nu_1}$ beschränkten, symmetrischen und koerzitativen Bilinearform $\mathcal{Q} : L_2^m(\mathbb{R}_{++}, \nu_1) \times L_2^m(\mathbb{R}_{++}, \nu_1) \rightarrow \mathbb{R}$, mit

$$\mathcal{Q}(u_1, u_2) := \int_0^\infty u_1^T(t) R(t) u_2(t) \nu_1(t) dt$$

sowie der bzgl. der Norm $\|\cdot\|_{(2),\nu_1}$ beschränkten Linearform [25](#) $q : L_2^m(\mathbb{R}_{++}, \nu_1) \rightarrow \mathbb{R}$

$$q(u) := \int_0^\infty u^T(t) [r(t) \nu_1(t) - B^T(t) y(t)] dt.$$

Satz [15](#) sichert die Existenz und Eindeutigkeit einer optimalen Lösung für das Problem $(HP3)$. Wir möchten die Herleitung notwendiger – und durch die strikte Konvexität des Zielfunktionals auch hinreichender – Bedingungen auf eine ähnliche Art und Weise führen wie in der Behandlung des Hilfsproblems $(HP1)$. Für eine optimale Lösung nutzen wir daher wieder die Aussagen von Lemma [28](#), gemäß derer die Auswertung der Variationsungleichung

$$\delta J_{1,y}^{\nu_1}(u_{*y}; u - u_{*y}) \geq 0 \quad \forall u \in \mathbf{U}.$$

notwendig und hinreichend für die Berechnung einer optimalen Lösung ist. Die Auswertung der Variationsungleichung erfordert die Berechnung des Gâteaux-Differentials des Zielfunktionals von $(HP3)$. Dieses ergibt sich mit einer ähnlichen Argumentation wie im Fall der Betrachtung von $(HP1)$ durch die lineare und (bzgl. der Normtopologie des Raumes $L_2^m(\mathbb{R}_{++}, \nu_1)$) stetige Abbildung $\vartheta \mapsto \mathcal{Q}(u_{*y}, \vartheta) + q(\vartheta)$ der Form

$$\delta J_{1,y}^{\nu_1}(u; \vartheta) = \int_0^\infty \vartheta^T(t) \left\{ (R(t)u(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right\} dt$$

für alle $\vartheta \in L_2^m(\mathbb{R}_{++}, \nu_1)$.

²⁵Ein Nachweis dieser Eigenschaften findet sich in der Beweisführung von Satz [15](#) in Kapitel [B.2](#)

Auswertung der Optimalitätsbedingungen

Gemäß unserer vorherigen Berechnungen ergibt sich mit der Setzung $\vartheta := u - u_*$ als notwendige und hinreichende Bedingung für eine optimale Lösung u_{*y} die Variationsungleichung

$$\int_0^{\infty} \left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]^T (u(t) - u_{*y}(t)) dt \geq 0 \quad \forall u \in \mathbf{U}. \quad (3.33)$$

Auch hier können wir durch Anwendung des nachstehenden Lemmas zu einer punktweisen Betrachtung der Variationsungleichung übergehen:

Lemma 31. Die Variationsungleichung (3.33) gilt genau dann, wenn für fast alle $t \in \mathbb{R}_{++}$ die punktweise Variationsungleichung in \mathbb{R} erfüllt ist:

$$\left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]^T (v - u_{*y}(t)) \geq 0 \quad \forall v \in [\underline{u}, \bar{u}]$$

Beweis. Wir setzen $z(t) := (R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t)$ und führen die gleiche Argumentation wie in Lemma 29.

□

Gemäß dieses Lemmas erhalten wir die optimale Lösung u_{*y} aus der punktweisen Auswertung der Variationsungleichung, d.h. wir berechnen für fast alle $t \in \mathbb{R}_{++}$

$$\left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]^T u_{*y}(t) = \min_{v \in [\underline{u}, \bar{u}]} \sum_{i=1}^m \left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]_i v_i.$$

Die obige Gleichung liefert die komponentenweise Bedingung

$$u_{i* y}(t) = \begin{cases} \bar{u}_i, & \text{falls } \left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]_i < 0 \\ \in [\underline{u}_i, \bar{u}_i], & \text{falls } \left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]_i = 0 \\ \underline{u}_i, & \text{falls } \left[(R(t)u_{*y}(t) + r(t)) e^{\beta_1 t} - B^T(t)y(t) \right]_i > 0 \end{cases} \quad (\forall) t \in \mathbb{R}_{++}.$$

und damit

$$u_{*y}(t) = \text{Proj}_{\mathcal{U}} \left(R^{-1}(t)(-r(t) + B^T(t)y(t)e^{-\beta_1 t}) \right) \quad (\forall) t \in \mathbb{R}_{++}.$$

Lösung des Hilfsproblems ($\overline{HP4}$)

Die Betrachtung dieses Hilfsproblems verläuft analog zur Behandlung des Hilfsproblems ($\overline{HP2}$). Die Einbettung in den Kontext der Hilbertraumtheorie liefert eine äquivalente Darstellung der Form

$$J_{2,y}^{\nu_1}(x) = \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0 \rightarrow \inf!$$

u.d.N. $x \in L_2^n(\mathbb{R}_{++}, \nu_1)$,

mit der positiv definiten, bzgl. der Norm $\|\cdot\|_{(2), \nu_1}$ beschränkten, symmetrischen und koerzitativen Bilinearform $\mathcal{Q} : L_2^n(\mathbb{R}_{++}, \nu_1) \times L_2^n(\mathbb{R}_{++}, \nu_1) \rightarrow \mathbb{R}$, beschrieben durch

$$\mathcal{Q}(x_1, x_2) := \int_0^{\infty} x_1^T(t) W(t) x_2(t) \nu_1(t) dt$$

sowie der bzgl. der Norm $\|\cdot\|_{(2),\nu_1}$ beschränkten Linearform²⁶ $q : L_2^n(\mathbb{R}_{++}, \nu_1) \rightarrow \mathbb{R}$ mit

$$q(x) := - \int_0^{\infty} x^T(t) [w(t)\nu_1(t) - y'(t) - A^T(t)y(t)] dt$$

und der Konstanten

$$q_0 := x_0^T y(0).$$

Auch hier können wir wieder die notwendigen Bedingungen des Satzes [14](#) anwenden²⁷ (die auch hier durch die Konvexität der Lösung ebenfalls hinreichend sind und durch die strikte Konvexität in globaler Form gelten). Diese erfordern wieder die Berechnung des Gâteaux-Differentials. Mit einer ähnlichen Argumentation wie im Fall der Betrachtung von $(HP2)$ können wir das Gâteaux-Differential angeben durch die lineare und (bzgl. der Normtopologie des Raumes $L_2^m(\mathbb{R}_{++}, \nu_1)$) stetige Abbildung $\xi \mapsto Q(x_{*y}, \xi) + q(\xi)$ der Form

$$\delta J_{2,y}^{\nu_1}(x; \xi) = \int_0^{\infty} \xi^T(t) \left\{ (W(t)x(t) + w(t))e^{\beta_1 t} - [y'(t) + A^T(t)y(t)] \right\} dt.$$

Auswertung der Optimalitätsbedingungen

Wir betrachten die notwendige und hinreichende Bedingung gemäß Satz [14](#). Diese liefert: x_{*y} ist Optimallösung genau dann, wenn für alle $\xi \in L_2^n(\mathbb{R}_{++}, \nu_1)$ gilt:

$$\delta J_{2,y}^{\nu_1}(x_{*y}; \xi) = \int_0^{\infty} \xi^T(t) \left\{ (W(t)x_{*y}(t) + w(t))e^{\beta_1 t} - [y'(t) + A^T(t)y(t)] \right\} dt \stackrel{!}{=} 0.$$

Wir erhalten schließlich analog zur Berechnung von $(\overline{HP2})$

$$x_{*y}(t) = W^{-1}(t) \left(-w(t) + [y'(t) + A^T(t)y(t)]e^{-\beta_1 t} \right) \quad (\forall t \in \mathbb{R}_{++}). \quad (3.34)$$

Auch hier liefert die Raumzugehörigkeit der Funktionen der rechten Seite der Gleichung $x_{*y} \in W_2^{1,n}(\mathbb{R}_{++}, \nu_1)$. Zudem existiert durch die Wahl der Menge Y_{ν_1} auch hier die klassische Ableitung von $y'(t)$ für alle $t \in \mathbb{R}_+$. Daher können wir Gleichung [\(3.34\)](#) ebenfalls punktweise auswerten und somit insbesondere auch für $t = 0$. Damit ergibt sich der Zusammenhang

$$x_{*y}(0) = W^{-1}(0) \left(-w(0) + y'(0) + A^T(0)y(0) \right).$$

Die Wahl der Mengen Y_{ν_1} liefert schließlich den Zusammenhang $x_{*y}(0) = x_0$.

Wir konnten somit die Existenz der eindeutigen Lösung $x_{*y} \in W_2^{2,n}(\mathbb{R}_{++}, \nu_1)$ für $(\overline{HP4})$, mit $x_{*y}(0) = x_0$ nachweisen. Damit löst x_{*y} auch das Problem $(HP4)$ eindeutig.

²⁶Ein Nachweis dieser Eigenschaften findet sich in der Beweisführung für Satz [16](#) in Kapitel [B.3](#)

²⁷siehe Satz [14](#), S. [61](#)

3.5 Formulierung eines schwachen Dualitätssatzes

Wir kombinieren nun die Ergebnisse der Optimierungsaufgaben und erhalten

$$\begin{aligned} J_{\nu_i}(x_*, u_*) &= \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) \geq \sup_{y \in Y_{\nu_i}} G_{\nu_i}(y) = \sup_{y \in Y_{\nu_i}} \left\{ \inf_{(x,u) \in X_0} \Phi_i((x, u), y) \right\} \\ &= \sup_{y \in Y_{\nu_i}} \left\{ \inf_{(x,u) \in X_0} \left\{ J_{2,y}^{\nu_i}(x) + J_{1,y}^{\nu_i}(u) \right\} \right\} = \sup_{y \in Y_{\nu_i}} \left\{ J_{2,y}^{\nu_i}(x_{*y}) + J_{1,y}^{\nu_i}(u_{*y}) \right\}, \end{aligned}$$

was letztlich zu der dualen Aufgabe führt:

$$\begin{aligned} (\mathbf{D}_{\nu_i}) \quad G_{\nu_i}(y) &:= J_{1,y}^{\nu_i}(u_{*y}) + J_{2,y}^{\nu_i}(x_{*y}) \longrightarrow \sup! \\ &\text{u.d.N. } y \in Y_{\nu_i}. \end{aligned}$$

Die Erkenntnisse dieses Kapitels fassen wir in dem folgenden Satz zusammen:

Satz 17 (Schwacher Dualitätssatz). *Sei ein Steuerungsproblem vom Typ (P_{ν_i}) , mit $i = 0, 1$ gegeben. Dann gilt unter den Voraussetzungen $(V1)$ - $(V4)$ für die Probleme (D_{ν_i}) , mit $i = 0, 1$*

$$G_{\nu_i}(y) \rightarrow \sup! \text{ bzgl. aller } y \in Y_{\nu_i}$$

die schwache Dualitätsungleichung

$$\inf(P_{\nu_i}) \geq \sup(D_{\nu_i}).$$

Beweis. Der Beweis folgt unmittelbar aus der Konstruktion. □

Wir waren in diesem Kapitel also in der Lage, duale Aufgaben für Probleme vom Typ (P_{ν_i}) herzuleiten. Im Zuge der Konstruktion nahmen wir allerdings einige recht einschränkende Voraussetzungen²⁸ für die Wahl der Mengen Y_{ν_i} in Kauf. Die Existenz einer optimalen Lösung der dualen Aufgaben (D_{ν_i}) ist an dieser Stelle damit völlig unklar. Wir widmen uns daher im nachfolgenden Kapitel dem Studium der Aufgabenklassen (D_{ν_i}) und legen dabei den Fokus auf die folgenden Fragestellungen:

1. Existiert eine optimale Lösung der dualen Aufgabe?
2. Wie können wir derartige Lösungen berechnen?
3. Welche Rückschlüsse liefern die Lösungen der dualen Aufgaben (D_{ν_i}) für die primalen Aufgabenstellungen?

²⁸Dies betrifft die Wahl des Funktionenraumes sowie der Zusammenhang der Anfangswerte der Funktion y und der Zustandsfunktion x .

Kapitel 4

Die duale Aufgabe (D_{ν})

Ziel dieses Kapitels ist der Nachweis der Existenz einer optimalen Lösung der dualen Aufgaben (D_{ν_i}) sowie die Herleitung notwendiger Optimalitätsbedingungen. Sind wir zudem in der Lage, für die betrachteten Problemklassen (P_{ν_i}) und (D_{ν_i}) die starke Dualitätsbeziehung

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = \sup_{y \in Y_{\nu_i}} G_{\nu_i}(y) = G_{\nu_i}(y_*) \quad (4.1)$$

herzuleiten, dann erlaubt uns dies Rückschlüsse zwischen primaler und dualer Aufgabenstellung zu ziehen.

Ein Nachweis der Gültigkeit der Relation (4.1) liefert darüber hinaus – dank der in Kapitel 2 nachgewiesenen Existenz einer optimalen Lösung (x_*, u_*) der primalen Aufgaben – hinreichende Optimalitätsbedingungen an die primalen Aufgaben (P_{ν_i}). Die *Sattelpunktbeziehung* (4.1) liefert somit den Schluss auf die noch ausstehenden Optimalitätsbedingungen der Aufgabe (P_{ν_i}). Für diese Nachweise ist das vorliegende Kapitel wie folgt strukturiert:

Im Sinne einer übersichtlicheren Behandlung werden wir zunächst die Fälle ν_0 und ν_1 mit Hilfe geeigneter Setzungen zu einer Problemformulierung (D_{ν}) zusammenfassen. Dies ist möglich, da uns die Wahl der Räume Y_{ν_i} in beiden Fällen zu einer Formulierung der dualen Aufgaben in Form von Variationsproblemen in gewichteten Funktionenräumen mit Dichtefunktionen führte¹. In Unterkapitel 4.2 werden wir zu einer Darstellung der dualen Aufgaben in Form eines Optimalsteuerungsproblems ohne Steuerungsbeschränkungen übergehen. Dies erlaubt uns die Formulierung notwendiger Optimalitätsbedingungen in einer ähnlichen Form wie in Satz 10.2.

Mit Hilfe dieser notwendigen Bedingungen werden wir schließlich in Unterkapitel 4.6 – unter einer geeigneten Wahl an Voraussetzungen – den Nachweis einer starken Dualitätsbeziehung liefern. Das Kapitel schließt mit der Formulierung von Optimalitätsbedingungen an die primalen Aufgaben, was – aufgrund bisher fehlender Aussagen über notwendige Optimalitätsbedingungen – insbesondere für den Aufgabentyp (P_{ν_1}) interessant ist.

Für das gesamte Kapitel gelte nun $C(t) \equiv w(t) \equiv 0$.

¹vgl. Kapitel 3

²siehe S. 41

4.1 Zusammenführung der dualen Aufgaben

Ziel dieses Unterabschnittes ist die Zusammenführung der Fälle ν_0 und ν_1 zu einer gemeinsamen dualen Aufgabenstellung. Dazu wollen wir zeigen, dass wir in der Lage sind, mittels einer geschickten Setzung der auftretenden Funktionen die Probleme (D_{ν_0}) und (D_{ν_1}) zu einer Formulierung der Form (D_ν) zusammenzufassen.

Formulierung für (D_{ν_0})

Für die Problemklasse (D_{ν_0}) bezeichne im Folgenden für alle $t \in \mathbb{R}_+$

$$\gamma := \beta_0, \quad \hat{A}(t) := A(t) + \gamma \cdot I_{n \times n} \quad \text{und} \quad \hat{r}(t) := r(t)$$

$$\text{sowie } \hat{\sigma}(t, y(t)) := \text{Proj}_{\mathcal{U}} \left[R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y(t) \right) \right] \quad (\forall) t \in \mathbb{R}_{++}.$$

Im Zuge der Konstruktion der dualen Aufgabe erhielten wir durch die Lösung der untergeordneten Optimierungsprobleme für fast alle $t \in \mathbb{R}_{++}$ die Zusammenhänge

$$x_{*y}(t) := W^{-1}(t) \left[y'(t) + \left(A^T(t) + \beta_0 \cdot I_{n \times n} \right) y(t) \right] \quad \text{und}$$

$$u_{*y}(t) := \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) (-r(t) + B^T(t)y(t)) \right).$$

Mit den obigen Setzungen lassen sich diese fast überall auf \mathbb{R}_{++} darstellen durch

$$x_{*y}(t) = W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] \quad \text{und}$$

$$u_{*y}(t) = \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) (-\hat{r}(t) + B^T(t)y(t)) \right) = \hat{\sigma}(t, y(t)).$$

Mit Hilfe dieser Setzungen können wir das Zielfunktional von (D_{ν_0}) nun schreiben als

$$G_{\nu_0}(y) = J_{2,y}^{\nu_0}(x_{*y}) + J_{1,y}^{\nu_0}(u_{*y})$$

wobei der Integrand von $J_{2,y}^{\nu_0}(x_{*y})$ durch obige Setzung nun für fast alle $t \in \mathbb{R}_{++}$ in der Form

$$\begin{aligned} j_{2,y}^{\nu_0}(t, x_{*y}(t)) &:= \left\{ \frac{1}{2} x_{*y}^T(t) W(t) x_{*y}(t) - x_{*y}^T(t) [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)] \right\} e^{\beta_0 t} \\ &= \left\{ \frac{1}{2} \left(W^{-1}(t) [y'(t) + \hat{A}^T(t)y(t)] \right)^T W(t) W^{-1}(t) [y'(t) + \hat{A}^T(t)y(t)] \right. \\ &\quad \left. - \left(W^{-1}(t) [y'(t) + \hat{A}^T(t)y(t)] \right)^T [y'(t) + \hat{A}^T(t)y(t)] \right\} e^{\beta_0 t} \\ &= -\frac{1}{2} [y'(t) + \hat{A}^T(t)y(t)]^T W^{-1}(t) [y'(t) + \hat{A}^T(t)y(t)] e^{\gamma t} \end{aligned}$$

gegeben ist und der Integrand von $J_{1,y}^{\nu_0}(u_{*y})$ durch

$$\begin{aligned} j_{1,y}^{\nu_0}(t, u_{*y}(t)) &:= \left\{ \frac{1}{2} u_{*y}^T(t) R(t) u_{*y}(t) + r^T(t) u_{*y}(t) - [B(t)u_{*y}(t)]^T y(t) \right\} e^{\beta_0 t} \\ &= \left\{ \frac{1}{2} \hat{\sigma}^T(t, y(t)) R(t) \hat{\sigma}(t, y(t)) + \hat{\sigma}^T(t, y(t)) \hat{r}(t) - \hat{\sigma}^T(t, y(t)) B^T(t)y(t) \right\} e^{\beta_0 t} \\ &= \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right] e^{\gamma t} \quad (\forall) t \in \mathbb{R}_{++}. \end{aligned}$$

Gemäß der Voraussetzungen (V1) und (V2)³ haben die auftretenden Funktionen folgende Eigenschaften:

$$\begin{aligned} W &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), \quad W'_{i,j} \in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \quad i, j = 1, \dots, n, \\ R &\in L_\infty^{m \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m}), \\ \hat{A} &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), \quad \hat{A}'_{i,j} \in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \quad i, j = 1, \dots, n, \\ B &\in L_\infty^{n \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times m}) \quad \text{und} \quad \hat{r} \in L_2^m(\mathbb{R}_{++}, e^{\gamma t}) \cap C(\mathbb{R}_+, \mathbb{R}^m). \end{aligned}$$

Die Nebenbedingungen ergeben sich entsprechend zu

$$y \in Y_{\nu_0} := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\}.$$

Formulierung für (D_{ν_1})

Für die Problemklasse (D_{ν_1}) bezeichne im Folgenden für alle $t \in \mathbb{R}_+$

$$\gamma := -\beta_1, \quad \hat{A}(t) := A(t), \quad \hat{r}(t) := r(t)e^{-\gamma t} \quad \text{sowie}$$

$$\hat{\sigma}(t, y(t)) := \text{Proj}_{\mathcal{U}} \left[R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y(t) \right) e^{\gamma t} \right] e^{-\gamma t} \quad (\forall) t \in \mathbb{R}_{++}.$$

Im Zuge der Konstruktion der dualen Aufgabe erhielten wir durch die Lösung der untergeordneten Optimierungsprobleme für fast alle $t \in \mathbb{R}_{++}$ die Zusammenhänge

$$x_{*y}(t) := W^{-1}(t) \left[y'(t) + A^T(t)y(t) \right] e^{-\beta_1 t} \quad \text{und}$$

$$u_{*y}(t) := \text{Proj}_{\mathcal{U}} \left[R^{-1}(t) \left(-r(t)e^{\beta_1 t} + B^T(t)y(t) \right) e^{-\beta_1 t} \right].$$

Mit den obigen Setzungen lassen sich diese fast überall auf \mathbb{R}_{++} darstellen durch

$$x_{*y}(t) = W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] e^{\gamma t} \quad \text{und}$$

$$u_{*y}(t) = \text{Proj}_{\mathcal{U}} \left[R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y(t) \right) e^{\gamma t} \right] = \hat{\sigma}(t, y(t))e^{\gamma t}.$$

Für das Zielfunktional von (D_{ν_1}) gilt damit

$$G_{\nu_1}(y) = J_{2,y}^{\nu_1}(x_{*y}) + J_{1,y}^{\nu_1}(u_{*y}),$$

wobei der Integrand von $J_{2,y}^{\nu_1}(x_{*y})$ nun beschrieben werden kann mittels

$$\begin{aligned} j_{2,y}^{\nu_1}(t, x_{*y}(t)) &:= \frac{1}{2} \left\{ x_{*y}^T(t) W(t) x_{*y}(t) \right\} e^{\beta_1 t} - x_{*y}^T(t) [y'(t) + A^T(t)y(t)] \\ &= \frac{1}{2} \left\{ \left(W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] e^{\gamma t} \right)^T W(t) W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] e^{\gamma t} \right\} e^{-\gamma t} \\ &\quad - \left(W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] e^{\gamma t} \right)^T [y'(t) + A^T(t)y(t)] \\ &= -\frac{1}{2} \left[y'(t) + \hat{A}^T(t)y(t) \right]^T W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] e^{\gamma t} \quad (\forall) t \in \mathbb{R}_{++}. \end{aligned}$$

³siehe S. [25](#)

und der Integrand von $J_{1,y}^{\nu_1}(u_{*y})$ mittels

$$\begin{aligned}
 j_{1,y}^{\nu_1}(t, u_{*y}(t)) &:= \left\{ \frac{1}{2} u_{*y}^T(t) R(t) u_{*y}(t) + r^T(t) u_{*y}(t) \right\} e^{\beta_1 t} - [B(t) u_{*y}(t)]^T y(t) \\
 &= \left\{ \frac{1}{2} \hat{\sigma}^T(t, y(t)) e^{\gamma t} R(t) \hat{\sigma}(t, y(t)) e^{\gamma t} + \hat{\sigma}^T(t, y(t)) e^{\gamma t} \hat{r}(t) e^{\gamma t} \right\} e^{-\gamma t} \\
 &\quad - \hat{\sigma}^T(t, y(t)) e^{\gamma t} B^T(t) y(t) \\
 &= \left\{ \frac{1}{2} \hat{\sigma}^T(t, y(t)) R(t) \hat{\sigma}(t, y(t)) + \hat{\sigma}^T(t, y(t)) \hat{r}(t) \right. \\
 &\quad \left. - \hat{\sigma}^T(t, y(t)) B^T(t) y(t) \right\} e^{\gamma t} \\
 &= \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t) y(t) \right] e^{\gamma t} \quad (\forall) t \in \mathbb{R}_{++}.
 \end{aligned}$$

Gemäß der Voraussetzungen (V1) und (V2)⁴ haben die auftretenden Funktionen auch hier die folgende Eigenschaften:

$$\begin{aligned} W &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), & W'_{i,j} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), & i, j &= 1, \dots, n, \\ R &\in L_\infty^{m \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m}), \\ \hat{A} &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), & \hat{A}'_{i,j} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), & i, j &= 1, \dots, n, \\ B &\in L_\infty^{n \times m}(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}^{n \times m}) & \text{und } \hat{r} &\in L_2^m(\mathbb{R}_{++}, e^{\gamma t}) \cap C(\mathbb{R}_+, \mathbb{R}^m). \end{aligned}$$

Die Raumzugehörigkeit der Funktion \hat{r} begründet sich dabei auf der Raumzugehörigkeit der Funktion $r \in L_2^m(\mathbb{R}_{++}, e^{\beta_1 t}) = L_2^m(\mathbb{R}_{++}, e^{-\gamma t})$ und der Abschätzung

$$\int_0^\infty |\hat{r}(t)|^2 e^{\gamma t} dt = \int_0^\infty |r(t)e^{-\gamma t}|^2 e^{\gamma t} dt = \int_0^\infty |r(t)|^2 e^{-\gamma t} dt = \int_0^\infty |r(t)|^2 e^{\beta_1 t} dt = \|r\|_{(2), e^{\beta_1 t}}.$$

Die Nebenbedingungen ergeben sich auch hier zu

$$y \in Y_{\nu_1} := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\}.$$

Mit Hilfe dieser Setzungen können wir die zu (P_{ν_0}) und (P_{ν_1}) dualen Aufgaben nun allgemein formulieren mittels

$$\begin{aligned} (\mathbf{D}_\nu) : \quad G(y) &:= \int_0^\infty g(t, y(t), y'(t)) e^{\gamma t} dt - x_0^T y(0) \longrightarrow \sup! \\ &\text{u.d.N. } y \in Y, \end{aligned}$$

mit

$$\begin{aligned} g(t, y(t), y'(t)) &:= -\frac{1}{2} \left[y'(t) + \hat{A}^T(t)y(t) \right]^T W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] \\ &\quad + \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right] \quad (\forall) t \in \mathbb{R}_{++} \end{aligned}$$

und

$$Y := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\},$$

mit den Funktionen $W : \mathbb{R}_+ \times \mathbb{R}^{n \times n}$, $R : \mathbb{R}_+ \rightarrow \mathbb{R}^{m \times m}$, $\hat{A} : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$ und $B : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times m}$ sowie $\hat{r} : \mathbb{R}_+ \rightarrow \mathbb{R}^m$.

4.2 Formulierung als Optimalsteuerungsproblem

Offensichtlich handelt es sich bei der dualen Aufgabe (D_ν) um ein Variationsproblem. Dies ermöglicht uns, unter Beachtung der Raumeigenschaften der einfließenden Funktionen, zu einer Formulierung der dualen Aufgabe in Form eines Optimalsteuerungsproblems ohne Steuerungsbeschränkungen überzugehen. Dafür definieren wir die Funktion

$$q(t) := y'(t) + \hat{A}^T(t)y(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad (4.2)$$

⁴siehe S. 25

wobei

$$\begin{aligned} \int_0^\infty |q(t)|^2 e^{\gamma t} dt &= \int_0^\infty \left| y'(t) + \hat{A}^T(t)y(t) \right|^2 e^{\gamma t} dt \\ &\leq 2 \int_0^\infty |y'(t)|^2 e^{\gamma t} dt + 2 \|\hat{A}\|_\infty \int_0^\infty |y(t)|^2 e^{\gamma t} dt \\ &\leq c \left(\|y'\|_{(2),e^{\gamma t}} + \|y\|_{(2),e^{\gamma t}} \right) < \infty, \quad \text{mit } c > 0. \end{aligned}$$

Demnach gilt $q \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$. Zudem erlauben die vorliegenden Raumeigenschaften der eingehenden Funktionen die Angabe der verallgemeinerten Ableitung der Funktion q in der Form

$$q'(t) = y''(t) + \hat{A}^T(t)y'(t) + \left(\hat{A}^T(t) \right)' y(t) \quad (\forall) t \in \mathbb{R}_{++},$$

wobei

$$\begin{aligned} \int_0^\infty |q'(t)|^2 e^{\gamma t} dt &= \int_0^\infty \left| y''(t) + \hat{A}^T(t)y'(t) + \left(\hat{A}^T(t) \right)' y(t) \right|^2 e^{\gamma t} dt \\ &\leq 2 \int_0^\infty |y''(t)|^2 e^{\gamma t} dt + 4 \|\hat{A}\|_\infty \int_0^\infty |y'(t)|^2 e^{\gamma t} dt + 4 \|\hat{A}'\|_\infty \int_0^\infty |y(t)|^2 e^{\gamma t} dt \\ &\leq c \left(\|y''\|_{(2),e^{\gamma t}} + \|y'\|_{(2),e^{\gamma t}} + \|y\|_{(2),e^{\gamma t}} \right) < \infty, \quad \text{mit } c > 0 \end{aligned}$$

und somit $q \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$.

Durch die Setzung (4.2) erhalten wir nun eine äquivalente Formulierung von (D_ν) als Optimalsteuerungsproblem:

$$\begin{aligned} (\mathbf{D}_\nu) : \quad G^\infty(y, q) &:= -\frac{1}{2} \int_0^\infty q^T(t) W^{-1} q(t) e^{\gamma t} dt \\ &+ \int_0^\infty \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T y(t) \right] e^{\gamma t} dt - x_0^T y(0) \longrightarrow \sup! \end{aligned}$$

bezüglich aller Paare

$$(y, q) \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}) \times W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \quad \text{mit } \gamma < 0, \quad (4.3)$$

für die gilt

$$y'(t) = -\hat{A}^T(t)y(t) + q(t) \quad (\forall) t \in \mathbb{R}_{++} \quad \text{und} \quad (4.4)$$

$$q(0) = x_0. \quad (4.5)$$

Bemerkung 17. Wir bezeichnen mit \mathcal{B}_D die Menge aller zu (D_ν) zulässigen Paare (y, q) , d.h. die Menge aller Funktionenpaare (y, q) , welche den Bedingungen (4.3) - (4.5) genügen.

Die Herleitung aus Kapitel 3 führte uns somit zu der Formulierung eines Variations- bzw. Optimalsteuerungsproblems in gewichteten Funktionenräumen mit Dichtefunktion. Dieser Umstand wird insbesondere für die spätere approximative Behandlung von Vorteil sein.

Wir wollen jedoch zunächst die Aufgabe (D_ν) als eigenständige Aufgabe betrachten und für diese notwendige Optimalitätsbedingungen formulieren. Dabei wollen wir uns den Methoden aus [Zie19, Bur19] zur Formulierung notwendiger Optimalitätsbedingungen mit Hilfe der Trennung konvexer Mengen in Hilberträumen annehmen. Daher untersuchen wir im weiteren Verlauf des Kapitels statt der Aufgabe (D_ν) die Problemformulierung

$$(\hat{D}) \quad \hat{G}(y) := -G(y) \longrightarrow \inf! \\ \text{u.d.N. } y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}),$$

wobei wir für die Aufgabe (\hat{D}) wieder zu der Form eines Optimalsteuerungsproblems übergehen wollen. Die Erweiterung auf die Betrachtung der größeren Menge $W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ wird uns eine Formulierung in den Funktionenräumen der Klasse (P_{ν_0}) und somit die Übernahme der Methodik aus [Zie19, Bur19] ermöglichen.

Bemerkung 18. *Durch die Einbettung des Problems (D_ν) in den größeren Raum*

$$W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \supset Y = \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\}$$

geht an dieser Stelle zunächst die Dualitätsbeziehung verloren. Die Formulierungen der notwendigen Bedingungen in Kapitel 4.5 werden uns jedoch schließlich den Nachweis erbringen, dass jede Lösung von (\hat{D}) Element des Raumes $W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t})$ ist und zudem der Bedingung

$$x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right)$$

genügt. Diese Argumentation liefert dann den Schluss zu unserer ursprünglichen Aufgabe (D_ν) über der Menge Y , für welche die Dualitätsbeziehung gilt.

Für die Aufgabe (\hat{D}) gelangen wir wieder über die Definition der Funktion

$$q(t) := y'(t) + \hat{A}^T(t)y(t) \quad (\forall) t \in \mathbb{R}_{++} \quad (4.6)$$

zu einer Formulierung als Optimalsteuerungsproblem. Aufgrund der Einbettung des Problems in die Menge $W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ können wir in diesem Fall jedoch nur auf die Raumzugehörigkeit $q \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ schließen. Damit ist eine punktweise Auswertung der Funktion q an der Stelle $t = 0$ nicht mehr möglich.

Die Setzung (4.6) erlaubt uns nun die äquivalente Formulierung von (\hat{D}) als Optimalsteuerungsproblem der Form

$$(\hat{D}) : \quad \hat{G}^\infty(y, q) := \frac{1}{2} \int_0^\infty q^T(t) W^{-1} q(t) e^{\gamma t} dt \\ - \int_0^\infty \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T y(t) \right] e^{\gamma t} dt + x_0^T y(0) \longrightarrow \inf!$$

bezüglich aller Paare

$$(y, q) \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \times L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \quad \text{mit} \quad \gamma < 0, \quad (4.7)$$

für die gilt

$$y'(t) = -\hat{A}^T(t)y(t) + q(t) \quad (\forall) t \in \mathbb{R}_{++}. \quad (4.8)$$

Bemerkung 19. Wir bezeichnen mit $\hat{\mathcal{B}}_D$ die Menge aller zu (\hat{D}) zulässigen Paare (y, q) , d.h. die Menge aller Funktionenpaare (y, q) , welche den Bedingungen (4.7) - (4.8) genügen.

Dies ist eine Aufgabe im gewichteten Hilbertraum mit Dichtefunktion ohne Steuerungsbeschränkungen und Anfangsbedingungen. Für diese wollen wir nun – ähnlich wie im Kapitel $\mathbb{2}$ – eine Einbettung in den Kontext der Hilbertraumtheorie vollziehen. Dem voran stellen wir die für die Aufgabe (\hat{D}) benötigten Grundvoraussetzungen.

4.3 Voraussetzungen

Wir betrachten die Aufgabe (\hat{D}) zunächst als ein von der primalen Aufgabe losgelöstes Problem und geben für dieses die folgenden Voraussetzungen an:

(VD1) Die auftretenden Funktionen wählen wir mit $i, j = 1, \dots, n$ und $k, l = 1, \dots, m$ gemäß

$$\begin{aligned} W &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), & W'_{i,j} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \\ R^{-1} &\in L_\infty^{m \times m}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{m \times m}), & (R^{-1})'_{k,l} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \\ \hat{A} &\in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n}), & \hat{A}'_{i,j} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \\ B &\in L_\infty^{n \times m}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times m}), & B'_{i,k} &\in L_\infty(\mathbb{R}_{++}) \cap C(\mathbb{R}_+, \mathbb{R}), \end{aligned}$$

$$\hat{r} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}).$$

(VD2) Die Matrizen W und R seien symmetrisch und positiv definit, d.h. es gelte $\underline{W} > 0$ und $\underline{R} > 0$. Die Matrix R sei zudem eine Diagonalmatrix.

(VD3) $\mathcal{U} = [\underline{u}, \bar{u}] \subseteq \mathbb{R}^m$ sei eine nichtleere, konvexe und kompakte Teilmenge des \mathbb{R}^m . Es gelte zudem $0 \in \text{int } \mathcal{U}$.

(VD4) Für die Menge $\hat{\mathcal{B}}_D$ gelte $\hat{\mathcal{B}}_D \neq \emptyset$ und $\hat{\mathcal{B}}_D \neq \{(y_*, q_*)\}$.

Bemerkung 20. Aus Voraussetzung (VD2) folgt unmittelbar die Existenz der inversen Matrizen W^{-1} und R^{-1} mit

$$\underline{W}^{-1} = \inf_{t \in \mathbb{R}_+} \lambda_{\min}(W^{-1})(t) > 0 \quad \text{sowie} \quad \underline{R}^{-1} = \inf_{t \in \mathbb{R}_+} \lambda_{\min}(R^{-1})(t) > 0.$$

Mit Hilfe der eingeführten Voraussetzungen (VD1)-(VD4), von deren Erfülltsein wir im Folgenden stets ausgehen, widmen wir uns nun der Analyse des Optimalsteuerungsproblems (\hat{D}) .

⁵mit $\underline{M} := \inf_{t \in \mathbb{R}_+} \lambda_{\min}(M)(t)$, wobei $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M bezeichne.

⁶Die Gültigkeit der Bedingung (VD3) wird uns den Nachweis des Lemmas $\mathbb{35}$ (siehe Seite $\mathbb{83}$) sichern. Die im Lemma $\mathbb{35}$ nachgewiesenen Eigenschaften sind wesentlich für den Nachweis der Gültigkeit von Satz $\mathbb{18}$ (siehe Seite $\mathbb{87}$). Dabei bezeichne $\text{int } \mathcal{U}$ das Innere der Menge \mathcal{U} .

4.4 Formulierung als Optimierungsproblem im Hilbertraum

Wir wollen nun wieder eine Formulierung der Aufgabe (\hat{D}) über Operatoren im Hilbertraum angeben. Wir definieren dafür den Raum

$$\mathcal{H} := \mathcal{H}_1 \times \mathcal{H}_2 = W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \times L_2^n(\mathbb{R}_{++}, e^{\gamma t}),$$

der versehen mit der Norm

$$\|(y, q)\|_{\mathcal{H}} := \left(\|y\|_{1,(2),e^{\gamma t}}^2 + \|q\|_{(2),e^{\gamma t}}^2 \right)^{1/2}$$

zu einem Hilbertraum wird. Wir führen dafür auf dem Raum \mathcal{H}_2 die Abbildung $\mathcal{Q} : \mathcal{H}_2 \times \mathcal{H}_2 \rightarrow \mathbb{R}$ durch

$$\mathcal{Q}(q_1, q_2) := \int_0^{\infty} \left\{ q_1^T(t) W^{-1}(t) q_2(t) \right\} e^{\gamma t} dt, \quad (4.9)$$

sowie auf dem Raum \mathcal{H}_1 die Abbildungen $b_1(y)$ und $b_2(y)$ mit $b_1, b_2 : \mathcal{H}_1 \rightarrow \mathbb{R}$ durch

$$b_1(y) := \int_0^{\infty} \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t) y(t) \right] e^{\gamma t} dt \quad \text{und} \quad (4.10)$$

$$b_2(y) := x_0^T y(0) \quad (4.11)$$

ein. Mit diesen Abbildungen können wir das Zielfunktional von (\hat{D}) auf ganz \mathcal{H} angeben mittels

$$\hat{G}(y, q) = \frac{1}{2} \mathcal{Q}(q, q) - b_1(y) + b_2(y). \quad (4.12)$$

Wir halten die Eigenschaften der aufgeführten Operatoren in den nachfolgenden Lemmata fest:

Lemma 32. *Durch die Abbildung \mathcal{Q} , definiert in (4.9), wird eine symmetrische Bilinearform auf \mathcal{H}_2 definiert. Diese ist strikt konvex sowie beschränkt und stetig auf \mathcal{H}_2 bzgl. der Normtopologie des Raumes \mathcal{H}_2 .*

Beweis. Es lässt sich leicht nachrechnen, dass \mathcal{Q} für alle $q_1, q_2, \hat{q}_1, \hat{q}_2 \in \mathcal{H}_2$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned} \mathcal{Q}(q_1 + \hat{q}_1, q_2) &= \mathcal{Q}(q_1, q_2) + \mathcal{Q}(\hat{q}_1, q_2) \\ \mathcal{Q}(q_1, q_2 + \hat{q}_2) &= \mathcal{Q}(q_1, q_2) + \mathcal{Q}(q_1, \hat{q}_2) \\ \mathcal{Q}(\lambda q_1, q_2) &= \lambda \mathcal{Q}(q_1, q_2). \end{aligned}$$

Aufgrund der Symmetrie der Matrix W^{-1} gilt zudem für alle $q_1, q_2 \in \mathcal{H}_2$

$$\mathcal{Q}(q_1, q_2) = \mathcal{Q}(q_2, q_1).$$

Damit ist \mathcal{Q} eine symmetrische Bilinearform.

Wir diskutieren nun die Beschränktheit der Bilinearform \mathcal{Q} . Die Raumzugehörigkeit der Matrix $W^{-1} \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ sichert die Endlichkeit des Wertes

$$\|W^{-1}\|_\infty = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |W^{-1}(t)|$$

und erlaubt somit die Abschätzung

$$\begin{aligned} |\mathcal{Q}(q_1, q_2)| &= \left| \int_0^\infty \{q_1^T(t) W^{-1}(t) q_2(t)\} e^{\gamma t} dt \right| \leq \|W^{-1}\|_\infty \int_0^\infty |q_1^T(t) q_2(t)| e^{\gamma t} dt \\ &\leq \|W^{-1}\|_\infty \left(\int_0^\infty |q_1(t)|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |q_2(t)|^2 dt \right)^{1/2} \\ &\leq \|W^{-1}\|_\infty \|q_1\|_{\mathcal{H}_2} \|q_2\|_{\mathcal{H}_2}, \end{aligned}$$

womit die (Norm-)Beschränktheit von \mathcal{Q} auf \mathcal{H}_2 bewiesen ist. Aufgrund der positiven Definitheit der Matrix W ist die in (4.9) definierte Bilinearform zudem strikt konvex. Damit ist \mathcal{Q} eine auf \mathcal{H}_2 (norm-)beschränkte, strikt konvexe Bilinearform und als solche stetig⁷ auf \mathcal{H}_2 bzgl. der Normtopologie des Raumes \mathcal{H}_2 . □

Lemma 33. *Durch die Abbildung b_1 , definiert in (4.10), wird ein bzgl. der Norm $\|\cdot\|_{\mathcal{H}_1}$ beschränkter Operator auf \mathcal{H}_1 definiert.*

Beweis. Wir untersuchen die Beschränktheit des Operators b_1 . Dafür unterscheiden wir zwischen den Fällen $\nu = \nu_0$ und $\nu = \nu_1$ und beginnen mit der Diskussion von $\nu = \nu_0$:

Aufgrund der Box-Constraints gilt $|\hat{\sigma}(t, y(t))| \leq c_1 := \max\{|\underline{u}|, |\bar{u}|\}$ fast überall auf \mathbb{R}_{++} . Zudem sichern die Raumzugehörigkeiten der Matrizen R und B die Endlichkeit der Werte

$$\|R\|_\infty = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |R(t)| \quad \text{sowie} \quad \|B\|_\infty = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |B(t)|.$$

Mit diesen Überlegungen gelangen wir zu folgender Abschätzung:

$$\begin{aligned} |b_1(y)| &= \int_0^\infty \left| \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t) y(t) \right] \right| e^{\gamma t} dt \\ &\leq \frac{c_2}{2} \|R\|_\infty + \int_0^\infty |\hat{\sigma}^T(t, y(t)) \hat{r}(t)| e^{\gamma t} dt + \|B\|_\infty \int_0^\infty |\hat{\sigma}^T(t, y(t)) y(t)| e^{\gamma t} dt \\ &\leq \frac{c_2}{2} \|R\|_\infty + \left(\int_0^\infty |\hat{\sigma}(t, y(t))|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |\hat{r}(t)|^2 e^{\gamma t} dt \right)^{1/2} \\ &\quad + \|B\|_\infty \left(\int_0^\infty |\hat{\sigma}(t, y(t))|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |y(t)|^2 e^{\gamma t} dt \right)^{1/2} \\ &\leq \frac{c_2}{2} \|R\|_\infty + c_2 \|\hat{r}\|_{\mathcal{H}_2} + c_2 \|B\|_\infty \|y\|_{\mathcal{H}_2} \leq c_3 \cdot \left\{ \|y\|_{\mathcal{H}_2} + \|y'\|_{\mathcal{H}_2} \right\} = c_3 \|y\|_{\mathcal{H}_1}, \end{aligned}$$

⁷vgl. z.B. [TMM09], S. 39

für $c_2 := \int_0^\infty c_1^2 e^{\gamma t} dt$ und hinreichend großes $0 < c_3 < \infty$. Damit ist die (Norm-)Beschränktheit des Operators für den Fall $\nu = \nu_0$ gezeigt. Basiert die duale Aufgabe jedoch auf der Konstruktion mittels der echten Gewichtsfunktion ν_1 , so finden wir folgende Argumentation:

Wir betrachten die Funktion $s(\cdot, y(\cdot)) := R^{-1}(\cdot) \left(-\hat{r}(\cdot) + B^T(\cdot)y(\cdot) \right)$. Gemäß Voraussetzung (VD2) ist R eine positiv definite Diagonalmatrix. Dies erlaubt uns für $i = 1, \dots, n$ die komponentenweise Definition der Funktion⁸

$$s_i(t, y(t)) = R_{ii}^{-1}(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t)y_k(t) \right) \quad \forall t \in \mathbb{R}_+.$$

Mit dieser können wir die Funktion $\hat{\sigma}$ nun komponentenweise angeben durch

$$\hat{\sigma}_i(t, \eta) = \text{Proj}_{\mathcal{U}} \left[s_i(t, \eta) e^{\gamma t} \right] e^{-\gamma t} := \begin{cases} \underline{u}_i e^{-\gamma t}, & \underline{u}_i > s_i(t, \eta) e^{\gamma t} \\ s_i(t, \eta), & \underline{u}_i \leq s_i(t, \eta) e^{\gamma t} \leq \bar{u}_i \\ \bar{u}_i e^{-\gamma t}, & \bar{u}_i < s_i(t, \eta) e^{\gamma t} \end{cases} \quad (\forall) t \in \mathbb{R}_{++}.$$

Demnach gilt

$$\hat{\sigma}_i(t, \eta) = s_i(t, \eta) \quad \Leftrightarrow \quad \underline{u}_i \leq s_i(t, \eta) e^{\gamma t} \leq \bar{u}_i.$$

Die Raumzugehörigkeiten der Funktionen R , \hat{r} , B und y implizieren, dass die Funktion s – unter Gültigkeit der postulierten Voraussetzungen – Element des Raumes $W_2^{1,m}(\mathbb{R}_{++}, e^{\gamma t})$ ist, denn es gilt

$$\begin{aligned} \int_0^\infty |s(t, y(t))|^2 e^{\gamma t} dt &= \int_0^\infty \left| R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y(t) \right) \right|^2 e^{\gamma t} dt \\ &\leq 2 \int_0^\infty \left| R^{-1}(t)\hat{r}(t) \right|^2 e^{\gamma t} dt + 2 \int_0^\infty \left| R^{-1}(t)B^T(t)y(t) \right|^2 e^{\gamma t} dt \\ &\leq 2 \|R^{-1}\|_\infty^2 \|\hat{r}\|_{(2), e^{\gamma t}}^2 + 2 \|R^{-1}\|_\infty^2 \|B\|_\infty^2 \|y\|_{(2), e^{\gamma t}}^2, \end{aligned}$$

d.h. $s \in L_2^m(\mathbb{R}_{++}, e^{\gamma t})$. Die Gültigkeit der Voraussetzung (VD1) sichert zudem

⁸Die Raumzugehörigkeiten der auftretenden Funktionen erlauben die punktweise Definition der Funktion $s(\cdot, \eta)$ auf ganz \mathbb{R}_+ .

$$\begin{aligned}
\int_0^\infty |s'(t, y(t))|^2 e^{\gamma t} dt &= \int_0^\infty \left| (R^{-1})'(t) (-\hat{r}(t) + B^T(t)y(t)) \right. \\
&\quad \left. + R^{-1}(t) \left(-\hat{r}'(t) + (B^T(t))' y(t) + B^T(t)y'(t) \right) \right|^2 e^{\gamma t} dt \\
&\leq 8 \left\{ \int_0^\infty \left| (R^{-1})'(t) \hat{r}(t) \right|^2 e^{\gamma t} dt + \int_0^\infty \left| (R^{-1})'(t) B^T(t)y(t) \right|^2 e^{\gamma t} dt \right. \\
&\quad + \int_0^\infty \left| R^{-1}(t) \hat{r}'(t) \right|^2 e^{\gamma t} dt + \int_0^\infty \left| R^{-1}(t) (B'(t))^T y(t) \right|^2 e^{\gamma t} dt \\
&\quad \left. + \int_0^\infty \left| R^{-1}(t) B^T(t)y'(t) \right|^2 e^{\gamma t} dt \right\} \\
&\leq 8 \left\{ \left\| (R^{-1})' \right\|_\infty^2 \|\hat{r}\|_{(2), e^{\gamma t}}^2 + \left\| (R^{-1})' \right\|_\infty^2 \|B\|_\infty^2 \|y\|_{(2), e^{\gamma t}}^2 \right. \\
&\quad + \|R^{-1}\|_\infty^2 \|\hat{r}'\|_{(2), e^{\gamma t}}^2 + \|R^{-1}\|_\infty^2 \|B'\|_\infty^2 \|y\|_{(2), e^{\gamma t}}^2 \\
&\quad \left. + \|R^{-1}\|_\infty^2 \|B\|_\infty^2 \|y'\|_{(2), e^{\gamma t}}^2 \right\}
\end{aligned}$$

und somit $s \in W_2^{1,m}(\mathbb{R}_{++}, e^{\gamma t})$. Damit können wir für die Funktion $\hat{s} : \mathbb{R}_+ \rightarrow \mathbb{R}^m$ mit

$$\hat{s}(t, y(t)) := s(t, y(t))e^{\gamma t} \quad \forall t \in \mathbb{R}_+$$

(wobei $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$) auf die Inklusion $\hat{s} \in W_2^{1,m}(\mathbb{R}_{++}, e^{-\gamma t})$ schließen, denn es gilt

$$\int_0^\infty |s(t, y(t))|^2 e^{\gamma t} dt = \int_0^\infty |\hat{s}(t, y(t))e^{-\gamma t}|^2 e^{\gamma t} dt = \int_0^\infty |\hat{s}(t, y(t))|^2 e^{-\gamma t} dt < \infty.$$

Damit gilt insbesondere auch $\lim_{t \rightarrow \infty} s(t, y(t))e^{\gamma t} = 0$. Zudem impliziert die Voraussetzung (VD3), dass $0 \in \text{int } \mathcal{U}$ ist. Aus diesen beiden Eigenschaften folgt

$$\exists T \in \mathbb{R}_{++} : s(t, y(t))e^{\gamma t} \in \text{int } \hat{\mathcal{U}} \quad (\forall) t > T.$$

Mit dieser Argumentation können wir den Operator b_1 nun abschätzen durch

$$\begin{aligned}
|b_1(y)| &= \int_0^\infty \left| \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right] \right| e^{\gamma t} dt \\
&= \int_0^T \left| \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right] \right| e^{\gamma t} dt \\
&\quad + \int_T^\infty \left| s^T(t, y(t)) \left[\frac{1}{2} R(t) s(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right] \right| e^{\gamma t} dt \\
&\leq c_4 + \|s\|_{\mathcal{H}_2} \left\{ \|R\|_\infty \|s\|_{\mathcal{H}_2} + \|\hat{r}\|_{\mathcal{H}_2} + \|\hat{B}\|_\infty \|y\|_{\mathcal{H}_2} \right\} \\
&\leq c_5 \left\{ \|y\|_{\mathcal{H}_2} + \|y'\|_{\mathcal{H}_2} \right\} = c_5 \|y\|_{\mathcal{H}_1},
\end{aligned}$$

wobei $c_4 < \infty$ den Wert des endlichen Integrals und $0 < c_5 < \infty$ eine hinreichend große Konstante bezeichne.

Damit haben wir die Beschränktheit des Operators b_1 bzgl. der Norm $\|\cdot\|_{\mathcal{H}_1}$ für beide Fälle gezeigt und die Aussage des Satzes bewiesen. \square

Wir wollen nun die Differenzierbarkeitseigenschaften des Operators b_1 analysieren. Dafür definieren wir zunächst fast überall auf \mathbb{R}_{++} die Funktion

$$\theta(t, \eta) := -\hat{\sigma}^T(t, \eta) \left[\frac{1}{2} R(t) \hat{\sigma}(t, \eta) + \hat{r}(t) - B^T(t) \eta \right] \quad (4.13)$$

und halten ihre Eigenschaften im nachfolgenden Lemma fest:

Lemma 34. *Für die Funktion θ , definiert in (4.13), gilt: $\theta(t, \cdot)$ ist stetig differenzierbar bezüglich η .*

Beweis. Gemäß Voraussetzung (VD2) ist R eine positiv definite Diagonalmatrix. Dies erlaubt die Darstellung der Funktion θ in der Form

$$\theta(t, \eta) = - \sum_{i=1}^m \left\{ \hat{\sigma}_i(t, \eta) \left[\frac{1}{2} R_{ii}(t) \hat{\sigma}_i(t, \eta) + \hat{r}_i(t) - \sum_{k=1}^n B_{ki}(t) \eta_k \right] \right\} \quad (\forall) t \in \mathbb{R}_{++}.$$

Wir erinnern an die Definition der Funktion

$$s_i(t, \eta) := R_{ii}^{-1}(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t) \eta_k \right) \quad \forall t \in \mathbb{R}_+$$

und betrachten zunächst den Fall $\nu = \nu_0$: Mit den obigen Darstellungen können wir für fast alle $t \in \mathbb{R}_{++}$ die einzelnen Summanden angeben mittels

$$\theta_i(t, \eta) = \begin{cases} -\frac{1}{2} \underline{u}_i R_{ii}(t) \underline{u}_i + \underline{u}_i \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t) \eta_k \right), & \underline{u}_i > s_i(t, \eta) \\ \frac{1}{2} R_{ii}^{-1}(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t) \eta_k \right)^2, & \underline{u}_i \leq s_i(t, \eta) \leq \bar{u}_i \\ -\frac{1}{2} \bar{u}_i R_{ii}(t) \bar{u}_i + \bar{u}_i \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t) \eta_k \right), & \bar{u}_i < s_i(t, \eta). \end{cases}$$

Für den Fall $\nu = \nu_1$ gilt fast überall auf \mathbb{R}_{++}

$$\hat{\sigma}_i(t, \eta) = \text{Proj}_{\mathcal{U}} \left[s_i(t, \eta) e^{\gamma t} \right] e^{-\gamma t} := \begin{cases} \underline{u}_i e^{-\gamma t}, & \underline{u}_i > s_i(t, \eta) e^{\gamma t} \\ s_i(t, \eta), & \underline{u}_i \leq s_i(t, \eta) e^{\gamma t} \leq \bar{u}_i \\ \bar{u}_i e^{-\gamma t}, & \bar{u}_i < s_i(t, \eta) e^{\gamma t}. \end{cases}$$

Mittels der Definitionen

$$\hat{\underline{u}}_i(t) := \underline{u}_i e^{-\gamma t} \quad \text{und} \quad \hat{\bar{u}}_i(t) := \bar{u}_i e^{-\gamma t} \quad \forall t \in \mathbb{R}_+$$

können wir die einzelnen Summanden der Funktionen θ_i fast überall auf \mathbb{R}_{++} angeben durch

$$\theta_i(t, \eta) = \begin{cases} -\frac{1}{2}\hat{u}_i(t)R_{ii}(t)\hat{u}_i(t) + \hat{u}_i(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t)\eta_k \right), & \hat{u}_i(t) > s_i(t, \eta) \\ \frac{1}{2}R_{ii}^{-1}(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t)\eta_k \right)^2, & \hat{u}_i(t) \leq s_i(t, \eta) \leq \bar{u}_i(t) \\ -\frac{1}{2}\bar{u}_i(t)R_{ii}(t)\bar{u}_i(t) + \bar{u}_i(t) \left(-\hat{r}_i(t) + \sum_{k=1}^n B_{ki}(t)\eta_k \right), & \bar{u}_i(t) < s_i(t, \eta). \end{cases}$$

Offenbar gleichen sich die Ausdrücke, so dass wir nun zu der gemeinsamen Notation

$$\hat{u}_i(t) := \begin{cases} u_i, & \text{für } \nu = \nu_0 \\ \underline{u}_i e^{-\gamma t}, & \text{für } \nu = \nu_1 \end{cases} \quad \text{und} \quad \bar{u}_i(t) := \begin{cases} \bar{u}_i, & \text{für } \nu = \nu_0 \\ \bar{u}_i e^{-\gamma t}, & \text{für } \nu = \nu_1 \end{cases} \quad \forall t \in \mathbb{R}_+$$

übergehen. Wir weisen nun – ähnlich wie [\[AKS16\]](#) und in [\[BKM14\]](#)⁹ – die Aussagen des Lemmas für die einzelnen Summanden nach und nutzen dafür den folgenden Hilfssatz aus der Analysis:

Hilfssatz. *Sei die Funktion $f : U(\eta_0) \rightarrow \mathbb{R}$ definiert für alle $\eta \in U(\eta_0)$. Ist f stetig differenzierbar in $U(\eta_0) \setminus \{\eta_0\}$ und existiert der Grenzwert*

$$\lim_{\eta \rightarrow \eta_0} f'(\eta) = A \quad \left(\Leftrightarrow \lim_{\eta \rightarrow \eta_0+0} f'(\eta) = \lim_{\eta \rightarrow \eta_0-0} f'(\eta) = A \right)$$

dann ist f in η_0 differenzierbar und es gilt

$$f'(\eta_0) = A.$$

Für die Anwendung des Hilfssatzes betrachten wir die Funktion θ nun für beliebiges aber festes $\tau \in \mathbb{R}_{++}$ und erhalten die folgende Darstellung der ersten partiellen Ableitungen der Funktion θ nach η über den offenen Intervallen:

$$\frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \begin{cases} \sum_{k=1}^n B_{ki}(\tau)\hat{u}_i(\tau), & \hat{u}_i(\tau) > s_i(\tau, \eta) \\ \sum_{k=1}^n B_{ki}(\tau) \left\{ R_{ii}^{-1}(\tau) \left(-\hat{r}_i(\tau) + \sum_{j=1}^n B_{ji}(\tau)\eta_j \right) \right\}, & \hat{u}_i(\tau) < s_i(\tau, \eta) < \bar{u}_i(\tau) \\ \sum_{k=1}^n B_{ki}(\tau)\bar{u}_i(\tau), & \bar{u}_i(\tau) < s_i(\tau, \eta). \end{cases} \quad (4.14)$$

Unter Ausnutzung der Definition

$$s_i(\tau, \eta) := R_{ii}^{-1}(\tau) \left(-\hat{r}_i(\tau) + \sum_{j=1}^n B_{ji}(\tau)\eta_j \right)$$

gelingt uns die Darstellung von [\(4.14\)](#) in der Form

$$\frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \begin{cases} \sum_{k=1}^n B_{ki}(\tau)\hat{u}_i(\tau), & \hat{u}_i(\tau) > s_i(\tau, \eta) \\ \sum_{k=1}^n B_{ki}(\tau)s_i(\tau, \eta), & \hat{u}_i(\tau) < s_i(\tau, \eta) < \bar{u}_i(\tau) \\ \sum_{k=1}^n B_{ki}(\tau)\bar{u}_i(\tau), & \bar{u}_i(\tau) < s_i(\tau, \eta). \end{cases}$$

⁹vgl. S. 1432, Lemma 2

Offensichtlich sind die einzelnen Ausdrücke über den offenen Intervallen stetig bzgl. η . Interessant ist daher die Betrachtung der Übergänge, das heißt die Betrachtung der links- und rechtsseitigen Grenzwerte an den Stellen $s_i(\tau, \eta) = \hat{u}_i(\tau)$ und $s_i(\tau, \eta) = \bar{u}_i(\tau)$ für beliebige aber feste $\tau \in \mathbb{R}_{++}$. Dies führt uns zur Betrachtung der Grenzwerte

$$\lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) - 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) - 0} \sum_{k=1}^n B_{ki}(\tau) \hat{u}_i(\tau) = \sum_{k=1}^n B_{ki}(\tau) \hat{u}_i(\tau)$$

und

$$\lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) + 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) + 0} \sum_{k=1}^n B_{ki}(\tau) s_i(\tau, \eta).$$

Die Struktur der in die Funktion s_i eingehenden Funktionen $R_{ii}^{-1}(t)$, $\hat{r}_i(t)$ und $B_{ji}(t)$ erlaubt uns den Schluß auf die Stetigkeit der Funktion $s_i(\cdot, \cdot)$ bzgl. t und η . Damit können wir die Funktion $s_i(\tau, \cdot)$ an der Stelle $\hat{u}_i(\tau)$ mit $\hat{u}_i(\tau)$ identifizieren und erhalten

$$\lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) + 0} \sum_{k=1}^n B_{ki}(\tau) s_i(\tau, \eta) = \sum_{k=1}^n B_{ki}(\tau) \hat{u}_i(\tau)$$

und damit

$$\lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) + 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \lim_{s_i(\tau, \eta) \rightarrow \hat{u}_i(\tau) - 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \sum_{k=1}^n B_{ki}(\tau) \hat{u}_i(\tau).$$

Auf gleiche Weise gelingt uns die Betrachtung der rechts- und linksseitigen Grenzwerte an der Stelle $s_i(\tau, \eta) = \bar{u}_i(\tau)$ und wir erhalten

$$\lim_{s_i(\tau, \eta) \rightarrow \bar{u}_i(\tau) + 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \lim_{s_i(\tau, \eta) \rightarrow \bar{u}_i(\tau) - 0} \frac{\partial \theta_i(\tau, \eta)}{\partial \eta} = \sum_{k=1}^n B_{ki}(\tau) \bar{u}_i(\tau).$$

Damit folgt die stetige Differenzierbarkeit aus der Gleichheit der beidseitigen Grenzwerte und wir finden die Darstellung

$$\nabla_{\eta} \theta(t, \eta) := B^T(t) \cdot \text{Proj}_{\hat{U}(t)}(s(t, \eta)) \quad (\forall) t \in \mathbb{R}_{++},$$

mit

$$\hat{U}(t) := [\hat{u}(t), \bar{u}(t)].$$

□

Lemma 35. Die Funktionen $\nabla_{\eta} \theta(\cdot, y(\cdot))$ und $\hat{\sigma}(\cdot, y(\cdot))$ mit $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ sind Elemente des Raumes $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$.

Beweis. Die Beweisführung aus Lemma 33 lieferte die Inklusion $s \in W_2^{1,m}(\mathbb{R}_{++}, e^{\gamma t})$. Die Beweisführung von Lemma 34 lieferte zudem die Darstellung

$$\nabla_{\eta} \theta(t, y(t)) := B^T(t) \cdot \text{Proj}_{\hat{U}(t)}(s(t, y(t))) = B^T(t) \hat{\sigma}(t, y(t)) \quad (\forall) t \in \mathbb{R}_{++},$$

mit

$$\hat{U}(t) = [\hat{u}(t), \bar{u}(t)] = \begin{cases} [u, \bar{u}], & \text{für } \nu = \nu_0 \\ [ue^{-\gamma t}, \bar{u}e^{-\gamma t}], & \text{für } \nu = \nu_1 \end{cases} \quad \forall t \in \mathbb{R}_+.$$

Für die weitere Betrachtung unterscheiden wir wieder zwischen den Fällen $\nu = \nu_0$ und $\nu = \nu_1$.

Wir beginnen mit der Betrachtung des Falles $\nu = \nu_0$: Hier ist die Menge $\hat{\mathcal{U}}(t) \equiv \mathcal{U}$, d.h. unabhängig von der Variablen t . Damit gilt für fast alle $t \in \mathbb{R}_{++}$

$$\begin{aligned} |\nabla_\eta \theta(t, y(t))| &\leq \|B\|_\infty |\hat{\sigma}(t, y(t))| \leq \|B\|_\infty \text{ess sup} \left\{ |\hat{u}(t)|, |\bar{u}(t)| \right\} \\ &= \|B\|_\infty \max \{ |\underline{u}|, |\bar{u}| \} = c \|B\|_\infty, \end{aligned}$$

mit $c := \max \{ |\underline{u}|, |\bar{u}| \}$, womit weiterhin folgt

$$\int_0^\infty |\nabla_\eta \theta(t, y(t))|^2 e^{\gamma t} dt \leq \|B\|_\infty^2 \int_0^\infty |\hat{\sigma}(t, y(t))|^2 e^{\gamma t} dt \leq c^2 \|B\|_\infty^2 \int_0^\infty e^{\gamma t} dt < \infty.$$

Wir betrachten nun den Fall $\nu = \nu_1$: Hier gilt gemäß Voraussetzung (VD3) $0 \in \text{int } \hat{\mathcal{U}}(t)$. Daraus folgt, dass $\hat{\mathcal{U}}(t)$ für fast alle $t \in \mathbb{R}_{++}$ von oben beschränkt wird durch eine Funktion

$$\bar{u}(t) = \bar{u} e^{-\gamma t} = c_1 e^{-\gamma t}, \quad \text{mit } c_1 > 0 \text{ und } \gamma < 0$$

und von unten durch eine Funktion

$$\underline{u}(t) = \underline{u} e^{-\gamma t} = -c_2 e^{-\gamma t}, \quad \text{mit } c_2 > 0 \text{ und } \gamma < 0.$$

Aus der Raumzugehörigkeit der Funktion s sowie der Tatsache, dass die Schranken des Steuerbereichs durch die echten Gewichtsfunktionen

$$\bar{u}(t) = c_1 e^{-\gamma t} \quad \text{und} \quad \underline{u}(t) = -c_2 e^{-\gamma t} \quad \text{mit } c_1, c_2 > 0 \text{ und } \gamma < 0$$

gegeben sind, folgt

$$\exists T \in \mathbb{R}_{++} : s \in \text{int } \hat{\mathcal{U}}(t) \quad (\forall) t > T.$$

Damit können wir die Funktion $\nabla_\eta \theta(\cdot, y(\cdot))$ mit der Setzung $c_3 := \max \{ |c_1|, |c_2| \}$ sowie der Konstanten $c_4 := \frac{c_3}{\gamma} (1 - e^{-\gamma T})$ weiter abschätzen in der Form

$$\begin{aligned} \int_0^\infty |\nabla_\eta \theta(t, y(t))|^2 e^{\gamma t} dt &\leq \|B\|_\infty^2 \int_0^\infty |\hat{\sigma}(t, y(t))|^2 e^{\gamma t} dt \\ &\leq \|B\|_\infty^2 \left(\int_0^T |c_3 e^{-\gamma t}|^2 e^{\gamma t} dt + \int_T^\infty |s(t, y(t))|^2 e^{\gamma t} dt \right) \\ &\leq \|B\|_\infty^2 (c_4 + \|s\|_{(2), e^{\gamma t}}^2) < \infty. \end{aligned}$$

Damit gilt offenbar in beiden Fällen $\nabla_\eta \theta(\cdot, y(\cdot)), \hat{\sigma}(\cdot, y(\cdot)) \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$, womit die Behauptungen des Lemmas bewiesen sind. \square

Lemma 36. Sei $x(0) = x_0$. Durch die Abbildung b_2 , definiert in [\(4.11\)](#), wird ein linearer Operator auf \mathcal{H}_1 definiert. Dieser ist beschränkt und damit auch stetig auf \mathcal{H}_1 bzgl. der Normtopologie des Raumes \mathcal{H}_1 .

Beweis. Die Linearität des Operators ist offensichtlich. Wir diskutieren daher die Beschränktheit und wollen nun unter Ausnutzung einer geschickten Umformulierung geeignete Schranken für b_2 finden.

Die Funktion y ist gemäß unserer Setzungen aus dem gewichteten Sobolev-Raum mit Dichtefunktion $W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$. Für die Betrachtung der Funktion x unterscheiden wir wieder zwischen den Fällen $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$ (dieser Fall beschreibt den Fall $\nu = \nu_1$) und $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ (dieser Fall beschreibt den Fall $\nu = \nu_0$). Dafür beginnen wir mit der Betrachtung der Funktion $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$:

Sei $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$ mit $x(0) = x_0$ und $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$, dann gilt¹⁰

$$-b_2(y(0)) = -x_0^T y(0) = \lim_{T \rightarrow \infty} x^T(T)y(T) - x^T(0)y(0). \quad (4.15)$$

Nach Hauptsatz der Differential- und Integralrechnung können wir Gleichung (4.15) auch schreiben als

$$\int_0^{\infty} [x^T(t)y(t)]' dt = \lim_{T \rightarrow \infty} x^T(T)y(T) - x^T(0)y(0).$$

Demnach können wir für b_2 nun folgende Abschätzung finden:

$$\begin{aligned} |b_2(y)| &= \left| \int_0^{\infty} [x^T(t)y(t)]' dt \right| \leq \int_0^{\infty} |x^T(t)y'(t)| dt + \int_0^{\infty} |y^T(t)x'(t)| dt \\ &\leq \left(\int_0^{\infty} |x(t)|^2 e^{-\gamma t} dt \right)^{1/2} \left(\int_0^{\infty} |y'(t)|^2 e^{\gamma t} dt \right)^{1/2} \\ &\quad + \left(\int_0^{\infty} |y(t)|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^{\infty} |x'(t)|^2 e^{-\gamma t} dt \right)^{1/2} \\ &\leq c_1 \|y\|_{\mathcal{H}_2} + c_2 \|y'\|_{\mathcal{H}_2} = c_3 \|y\|_{\mathcal{H}_1}, \quad \text{mit } c_1, c_2, c_3 > 0. \end{aligned}$$

Wir betrachten nun den Fall $x \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ mit $x(0) = x_0$ und definieren dafür die Funktion

$$\hat{x}(t) = x(t)e^{\gamma t} \quad (\forall) t \in \mathbb{R}_{++}.$$

Für die so eingeführte Funktion können wir die verallgemeinerte Ableitung angeben durch

$$\hat{x}'(t) = \gamma x(t)e^{\gamma t} + x'(t)e^{\gamma t} \quad (\forall) t \in \mathbb{R}_{++}$$

und erhalten $\hat{x} \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$, denn

$$\begin{aligned} \int_0^{\infty} |\hat{x}(t)|^2 e^{-\gamma t} dt &= \int_0^{\infty} |x(t)e^{\gamma t}|^2 e^{-\gamma t} dt = \int_0^{\infty} |x(t)|^2 e^{\gamma t} dt < \infty \quad \text{sowie} \\ \int_0^{\infty} |\hat{x}'(t)|^2 e^{-\gamma t} dt &= \int_0^{\infty} |(\gamma x(t) + x'(t))e^{\gamma t}|^2 e^{-\gamma t} dt \\ &\leq 2\gamma^2 \int_0^{\infty} |x(t)|^2 e^{\gamma t} dt + 2 \int_0^{\infty} |x'(t)|^2 e^{\gamma t} dt < \infty. \end{aligned}$$

¹⁰vgl. Lemma 30, S. 63

Für diese Funktion gilt zudem

$$-x_0^T y(0) = -x_0^T e^{\gamma 0} y(0) = -\hat{x}_0^T y(0) = \lim_{T \rightarrow \infty} \hat{x}^T(T) y(T) - \hat{x}^T(0) y(0).$$

Damit können die obigen Abschätzungen also analog für die Funktion $\hat{x} \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$ vollzogen werden und wir erhalten auch hier

$$|b_2(y)| \leq c_3 \|y\|_{\mathcal{H}_1},$$

mit $c_3 > 0$. Damit ist der lineare Operator b_2 beschränkt und somit auch stetig bzgl. der Normtopologie des Raumes \mathcal{H}_1 . \square

Wir wollen nun auch die Systemdynamik (4.3) mit Hilfe einer Operatorengleichung beschreiben. Dafür definieren wir den Operator $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{H}_2$ und $\mathbf{b} \in \mathcal{H}_2$, mit

$$\mathcal{A}(y, q) := -\hat{A}^T(\cdot)y(\cdot) + q(\cdot) - y'(\cdot), \quad \mathbf{b} := 0. \quad (4.16)$$

Lemma 37. *Der Operator \mathcal{A} , definiert in (4.16), ist ein linearer und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränkter Operator. Insbesondere ist \mathcal{A} damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Die Linearität lässt sich leicht nachrechnen. Für die Beschränktheit haben wir wieder die Existenz einer Konstanten $c \in \mathbb{R}$ nachzuweisen, so dass gilt

$$\|\mathcal{A}(y, q)\|_{\mathcal{H}_2} \leq c \|(y, q)\|_{\mathcal{H}}.$$

Wir betrachten daher für alle $(y, q) \in \mathcal{H}$ die Abschätzung

$$\begin{aligned} \|\mathcal{A}(y, q)\|_{\mathcal{H}_2} &= \int_0^\infty \left| -\hat{A}^T(t)y(t) + q(t) - y'(t) \right|^2 e^{\gamma t} dt \\ &\leq 2 \int_0^\infty \left| -\hat{A}^T(t)y(t) + q(t) \right|^2 + |y'(t)|^2 e^{\gamma t} dt \\ &\leq 4 \left(\int_0^\infty \left| -\hat{A}^T(t)y(t) \right|^2 e^{\gamma t} dt + \int_0^\infty |q(t)|^2 e^{\gamma t} dt \right) + 2 \int_0^\infty |y'(t)|^2 e^{\gamma t} dt \end{aligned}$$

Die Raumzugehörigkeit von \hat{A} erlaubt dann die Abschätzung

$$\begin{aligned} \|\mathcal{A}(y, q)\|_{\mathcal{H}_2} &\leq 4 \left(\|\hat{A}\|_\infty \|y\|_{(2), e^{\gamma t}}^2 + \|q\|_{(2), e^{\gamma t}}^2 \right) + 2 \|y'\|_{(2), e^{\gamma t}}^2 \\ &\leq c \|(y, q)\|_{\mathcal{H}}^2, \quad \text{mit } c = 0. \end{aligned}$$

Damit sind die Aussagen des Satzes bewiesen. \square

Mit Hilfe der eingeführten Operatoren können wir die zulässige Menge der Aufgabe (\hat{D}) durch

$$\hat{\mathcal{B}}_D := \{(y, q) \in \mathcal{H} \mid \mathcal{A}(y, q) = \mathbf{b}\}$$

beschreiben. Auf diese Weise erhalten wir eine Formulierung der Aufgabe (\hat{D}) in der Art

$$\hat{G}(y, q) = \frac{1}{2} \mathcal{Q}(q, q) - b_1(y) + b_2(y) \rightarrow \inf! \quad \text{bzgl. } (y, q) \in \hat{\mathcal{B}}_D.$$

4.5 Notwendige Optimalitätsbedingungen

Wir wollen in diesem Unterkapitel den Nachweis der Gültigkeit der nachstehenden Formulierung notwendiger Optimalitätsbedingungen erbringen. Dazu betrachten wir die Aufgabe (\hat{D}) zunächst als ein von der primalen Aufgabe losgelöstes Optimalsteuerungsproblem mit der neuen Zustandsfunktion y , der Steuerungsfunktion q und der Adjungierten p_0 . Weiterhin wollen wir uns auf die folgende Setzung der Platzhaltervariablen einigen:

$$\xi_D := y(t) \quad v_D := q(t) \quad \eta_D := p_0(t).$$

Satz 18. *Seien die Voraussetzungen (VD1)-(VD4) erfüllt und $(y_*, q_*) \in \hat{\mathcal{B}}_D$ eine optimale Lösung von (\hat{D}) im Sinne des **Kriteriums L**. Dann existieren Multiplikatoren (λ_0, p_0) mit:*

$$\lambda_0 = 1, \tag{N}$$

$$p_0 \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t}), \quad \gamma < 0 \tag{T}$$

$$H(t, y_*(t), q_*(t), p_0(t), \lambda_0) = \max_{v_D \in \mathbb{R}^n} H(t, y_*(t), v_D, p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{M}$$

$$p'_0(t) = -\nabla_{\xi_D}^T H(t, y_*(t), q_*(t), p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{K}$$

$$\text{mit } p_0(0) = x_0,$$

wobei $H : \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ die Pontryaginsche Funktion

$$H(t, \xi_D, v_D, \eta_D, \lambda_0) = -\lambda_0 g(t, \xi_D, v_D) e^{\gamma t} + \eta_D^T (-\hat{A}^T(t) \xi_D + v_D) \quad \text{mit}$$

$$g(t, \xi_D, v_D) = \frac{1}{2} v_D^T W^{-1}(t) v_D - \hat{\sigma}^T(t, \xi_D) \left[\frac{1}{2} R(t) \hat{\sigma}(t, \xi_D) + \hat{r}(t) - B^T(t) \xi_D \right] \quad \text{ist.}$$

Beweis. Wir führen den Beweis mit den gleichen Ideen wie in [\[Pic15\]](#), [\[Bur19\]](#), [\[ZP18\]](#) und betrachten dazu nachfolgend 7 Schritte:

Schritt 1: (Berechnung des einseitigen Gâteaux-Differentials)

Sei $(y, q) \in \hat{\mathcal{B}}_D$ und $(y_*, q_*) \in \hat{\mathcal{B}}_D$ optimal im Sinne des **Kriteriums L**. Wir weisen nach, dass für alle $\zeta := y - y_* \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ und $\vartheta := q - q_* \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ die einseitigen Gâteaux-Differenziale 1. Ordnung von \hat{G} in (y_*, q_*) existieren und nicht negativ sind, d.h. wir zeigen, dass gilt

$$0 \leq \delta^+ \hat{G}((y_*, q_*); (\zeta, \vartheta)) := \lim_{\lambda \rightarrow 0+0} \frac{1}{\lambda} \left\{ \hat{G}(y_* + \lambda \zeta, q_* + \lambda \vartheta) - \hat{G}(y_*, q_*) \right\}.$$

Für die Berechnung des Gâteaux-Differentials erinnern wir an die Schreibweise¹¹

$$\begin{aligned} \hat{G}(y, q) &= \frac{1}{2} \mathcal{Q}(y, y) - b_1(y) + b_2(y) \\ &= \frac{1}{2} \int_0^\infty \left\{ q^T(t) W^{-1}(t) q(t) \right\} e^{\gamma t} dt - \int_0^\infty \left\{ -\theta(t, y(t)) \right\} e^{\gamma t} dt + x_0^T y(0) \end{aligned}$$

¹¹vgl. Relation [\(4.12\)](#), S. [77](#)

und berechnen die Gâteaux-Differentiale der einzelnen Operatoren:

Wir beginnen mit der Betrachtung des Operators $\mathcal{Q}(q, q)$. Gemäß Lemma 32 ist durch diesen Operator eine (bzgl. der Normtopologie des Raumes \mathcal{H}_2) stetige Bilinearform auf $\mathcal{H}_2 \times \mathcal{H}_2$ mit $\mathcal{H}_2 = L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ definiert¹². Demnach können wir das Gâteaux-Differential mit seinem Fréchet-Differential identifizieren und erhalten¹³

$$\delta \mathcal{Q}((q, q); (\vartheta, \vartheta)) = \int_0^\infty \left\{ 2q_*^T(t) W^{-1}(t) \vartheta(t) \right\} e^{\gamma t} dt.$$

Wir berechnen nun das Gâteaux-Differential für den Operator b_2 und untersuchen für alle $\zeta = y - y_* \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ den Grenzwert

$$\begin{aligned} \delta^+ b_2(y_*; \zeta) &= \lim_{\lambda \rightarrow 0+0} \left\{ \frac{b_2(y_* + \lambda \zeta) - b_2(y_*)}{\lambda} \right\} \\ &= \lim_{\lambda \rightarrow 0+0} \left\{ \frac{x_0^T y_*(0) + \lambda x_0^T \zeta(0) - x_0^T y_*(0)}{\lambda} \right\} = x_0^T \zeta(0). \end{aligned}$$

und erhalten damit

$$\delta^+ b_2(y; \zeta) = x_0^T \zeta(0).$$

Wir betrachten nun den Operator b_1 und setzen mit $\zeta = y - y_* \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$

$$\Phi(\lambda) = b_1(y_* + \lambda \zeta) = \int_0^\infty - \underbrace{\{\theta(t, y_*(t) + \lambda \zeta(t))\}}_{=: \psi(t, \lambda)} e^{\gamma t} dt.$$

Aufgrund der stetigen Differenzierbarkeit des Integranden bzgl. ξ_D ¹⁴ mit

$$\begin{aligned} \theta(t, y(t)) &= \sum_{i=1}^m \theta_i(t, y(t)) \\ &= - \sum_{i=1}^m \left\{ \frac{1}{2} \hat{\sigma}_i(t, y(t)) R_{ii}(t) \hat{\sigma}_i(t, y(t)) - \hat{\sigma}_i(t, y(t)) \left(\sum_{k=1}^n B_{ki}(t) y_k(t) - \hat{r}_i(t) \right) \right\} \end{aligned}$$

für fast alle $t \in \mathbb{R}_{++}$ ist ψ nach λ differenzierbar:

$$\frac{\partial}{\partial \lambda} \psi(t, \lambda) = \left\{ \nabla_{\xi_D}^T \theta(t, y_*(t) + \lambda \zeta) \zeta(t) \right\} e^{\gamma t}.$$

Gemäß des Beweises von Lemma 34 können wir die partielle Ableitung der Funktion θ nach ξ_D angeben durch

$$\nabla_{\xi_D} \theta(t, y(t) + \lambda \zeta(t)) = \sum_{i=1}^m \nabla_{\xi_D} \theta_i(t, y(t) + \lambda \zeta(t)) \quad (\forall) t \in \mathbb{R}_{++},$$

wobei

¹²siehe Lemma 32, S. 77

¹³vgl. Satz 6 und Beispiel 2, S. 18 ff.

¹⁴vgl. Lemma 34, S. 81

$$\nabla_{\xi_D} \theta_i(t, y(t) + \lambda \zeta(t)) = \begin{cases} \sum_{k=1}^n B_{ki}(t) \hat{u}_i(t), & \hat{u}_i(t) > s_i(t, y(t) + \lambda \zeta(t)) \\ \sum_{k=1}^n B_{ki}(t) s_i(t, y(t) + \lambda \zeta(t)), & \hat{u}_i(t) \leq s_i(t, y(t) + \lambda \zeta(t)) \leq \bar{u}_i(t) \\ \sum_{k=1}^n B_{ki}(t) \bar{u}_i(t), & \bar{u}_i(t) < s_i(t, y(t) + \lambda \zeta(t)) \end{cases}$$

fast überall auf \mathbb{R}_{++} und damit

$$\nabla_{\xi_D} \theta(t, y(t) + \lambda \zeta(t)) = B^T(t) \cdot \text{Proj}_{\hat{U}(t)}(s(t, y(t) + \lambda \zeta(t))) \quad (\forall) t \in \mathbb{R}_{++}.$$

Dies impliziert unter Gültigkeit der Bedingung (VD3)

$$|\nabla_{\xi_D} \theta(t, y(t) + \lambda \zeta(t))| \leq |B^T(t) s(t, y(t) + \lambda \zeta(t))| \quad (\forall) t \in \mathbb{R}_{++}.$$

Wir überprüfen nun die Voraussetzungen von Satz 29, welche uns die „Differentiation unter dem Integral“ erlauben¹⁵ und betrachten daher für alle $\lambda \in [0, 1]$ fast überall auf \mathbb{R}_{++} :

$$\begin{aligned} \left| \frac{\partial}{\partial \lambda} \psi(t, \lambda) \right| &= \left| \zeta^T(t) \{ \nabla_{\xi_D} \theta(t, y_*(t) + \lambda \zeta(t)) \} e^{\gamma t} \right| \\ &\leq \left| \zeta^T(t) \{ B^T(t) s(t, y_*(t) + \lambda \zeta(t)) \} e^{\gamma t} \right|. \end{aligned} \quad (4.17)$$

Aufgrund der linearen Struktur der Funktion $s(t, \cdot)$ bzgl. ξ_D erhalten wir eine obere Abschätzung der Gleichung (4.17) durch die Betrachtung der Ränder. Wir erhalten somit für alle $\lambda \in [0, 1]$ fast überall auf \mathbb{R}_{++} eine Abschätzung der Form

$$\left| \frac{\partial}{\partial \lambda} \psi(t, \lambda) \right| \leq g(t) := \max \{ g_1(t), g_2(t) \}, \text{ mit}$$

$$g_1(t) = \left| \zeta^T(t) \{ B^T(t) s(t, y_*(t)) \} e^{\gamma t} \right| \text{ und } g_2(t) = \left| \zeta^T(t) \{ B^T(t) s(t, y(t)) \} e^{\gamma t} \right| \quad (\forall) t \in \mathbb{R}_{++}.$$

Die Bedingungen des Satzes 29 sind erfüllt, falls $g \in L_1(\mathbb{R}_{++})$. Wir betrachten daher nun für alle $y \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ (damit sind beide Funktionen g_1 und g_2 diskurtiert)

$$\int_0^\infty |g(t)| dt = \int_0^\infty \left| \zeta^T(t) \{ B^T(t) s(t, y(t)) \} \right| e^{\gamma t} dt.$$

Mit der Ungleichung von Cauchy-Schwarz erhalten wir weiter

$$\begin{aligned} \int_0^\infty |g(t)| dt &\leq \left(\int_0^\infty |\zeta(t)|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |B^T(t) s(t, y(t))|^2 e^{\gamma t} dt \right)^{1/2} \\ &\leq \|\zeta\|_{(2), e^{\gamma t}} \left(\|B\|_\infty^2 \int_0^\infty |s(t, y(t))|^2 e^{\gamma t} dt \right)^{1/2}. \end{aligned}$$

¹⁵siehe Satz 29, S. 186

Gemäß der Betrachtungen aus Lemma 33 ist die Funktion $s(\cdot, y(\cdot))$ Element des Raumes $W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$, so dass weiter gilt

$$\int_0^\infty |g(t)| dt \leq \|\zeta\|_{(2), e^{\gamma t}} \|B\|_\infty \|s\|_{(2), e^{\gamma t}} < \infty.$$

Damit gilt $g \in L_1(\mathbb{R}_{++})$, so dass wir gemäß Satz 29 „unter dem Integral“ differenzieren dürfen und erhalten

$$\delta^+ b_1(y_*; \zeta) = \int_0^\infty \left\{ -\nabla_{\xi_D}^T \theta(t, y_*(t)) \zeta(t) \right\} e^{\gamma t} dt.$$

Damit haben wir die Gâteaux-Differentiale der einzelnen Operatoren berechnet und können das nichtnegative Gâteaux-Differential¹⁶ der Funktion \hat{G} nun angeben durch

$$\delta^+ \hat{G}((y_*, q_*); (\zeta, \vartheta)) = \int_0^\infty \left\{ \nabla_{\xi_D}^T \theta(t, y_*(t)) \zeta(t) + q_*^T(t) W^{-1}(t) \vartheta(t) \right\} e^{\gamma t} dt + x_0^T \zeta(0).$$

Schritt 2: (Definition der Menge M_D)

Für die Setzung

$$\begin{aligned} a_1(\vartheta) &:= \int_0^\infty q_*^T(t) W^{-1}(t) \vartheta(t) e^{\gamma t} dt & a_2(\vartheta) &:= -\vartheta \\ b_1(\zeta) &:= \int_0^\infty \nabla_{\xi_D}^T \theta(t, y_*(t)) \zeta(t) e^{\gamma t} dt + x_0^T \zeta(0) & b_2(\zeta) &:= \zeta' + \hat{A}^T(\cdot) \zeta \end{aligned}$$

gilt das folgende Lemma:

Lemma 38. *Unter den getroffenen Voraussetzungen gilt:*

1. $a_1 : L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \rightarrow \mathbb{R}$ ist linear und beschränkt.
2. $b_1 : W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \rightarrow \mathbb{R}$ ist linear und beschränkt.
3. $a_2 : L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \rightarrow L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ ist linear und beschränkt.
4. $b_2 : W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \rightarrow L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ ist linear und beschränkt.

Beweis. Die Linearität aller Operatoren lässt sich leicht nachrechnen. Wir diskutieren daher lediglich die Beschränktheit der eingeführten Operatoren und betrachten dafür zunächst

$$\begin{aligned} |a_1(\vartheta)| &\leq \int_0^\infty \left| q_*^T(t) W^{-1}(t) \vartheta(t) \right| e^{\gamma t} dt \\ &\leq \|W^{-1}\|_\infty \left(\int_0^\infty |q_*(t)|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |\vartheta(t)|^2 e^{\gamma t} dt \right)^{1/2} \\ &\leq \|W^{-1}\|_\infty \cdot \|q_*\|_{(2), e^{\gamma t}} \cdot \|\vartheta\|_{(2), e^{\gamma t}}. \end{aligned}$$

¹⁶diese Eigenschaft folgt aus der Optimalität von (y_*, q_*)

Die Beschränktheit des Operators b_1 erhalten wir mit Hilfe der in Lemma [35](#) nachgewiesenen Raumzugehörigkeit

$$\nabla_{\xi_D}^T \theta(\cdot, y(\cdot)) \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$$

sowie durch die Beweisführung des Lemmas [36](#) mittels der Abschätzung

$$\begin{aligned} |b_1(\zeta)| &\leq \int_0^\infty \left| \nabla_{\xi_D}^T \theta(t, y_*(t)) \zeta(t) \right| e^{\gamma t} dt + \left| x_0^T \zeta(0) \right| \\ &\leq \left(\int_0^\infty \left| \nabla_{\xi_D}^T \theta(t, y_*(t)) \right|^2 e^{\gamma t} dt \right)^{1/2} \left(\int_0^\infty |\zeta(t)|^2 e^{\gamma t} dt \right)^{1/2} + \int_0^\infty \left| [x^T(t) \zeta(t)]' \right| dt \\ &\leq \|\nabla_\eta \theta\|_{(2), e^{\gamma t}} \cdot \|\zeta\|_{(2), e^{\gamma t}} + c_1 \|\zeta\|_{(2), e^{\gamma t}} + c_2 \|\zeta'\|_{(2), e^{\gamma t}} \\ &\leq c_3 \left(\|\zeta\|_{(2), e^{\gamma t}} + \|\zeta'\|_{(2), e^{\gamma t}} \right) = c_3 \|\zeta\|_{1, (2), e^{\gamma t}}, \quad \text{mit } c_1, c_2, c_3 > 0. \end{aligned}$$

Die Beschränktheit des Operators a_2 gilt offensichtlich. Für den Operator b_2 finden wir auf Grundlage der Voraussetzungen mit $\hat{A} \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ die Abschätzung

$$\|b_2\|_{(2), e^{\gamma t}}^2 \leq \|\zeta' + \hat{A}^T \zeta\|_{(2), e^{\gamma t}}^2 \leq \|\zeta'\|_{(2), e^{\gamma t}}^2 + \|\hat{A}\|_\infty^2 \cdot \|\zeta\|_{(2), e^{\gamma t}}^2.$$

Damit sind alle Behauptungen des Lemmas bewiesen. \square

Wir definieren nun die Menge M_D der Variationen mit

$$\begin{aligned} M_D := \{ (z_0, z) \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \mid & z_0 = a_1(\vartheta) + b_1(\zeta), z = a_2(\vartheta) + b_2(\zeta), \\ & \zeta = y - y_* \in W_2^{1, n}(\mathbb{R}_{++}, e^{\gamma t}), \|\zeta\|_{1, (2), e^{\gamma t}} \leq 1 \\ & \vartheta = q - q_* \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}), \|\vartheta\|_{(2), e^{\gamma t}} \leq 1 \}. \end{aligned}$$

Schritt 3: (Die Menge M_D ist konvex und \overline{M}_D entspricht nicht dem gesamten Raum $\mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$)

Wir bezeichnen mit \overline{M}_D den Abschluss der Menge M_D und formulieren das folgende Lemma:

Lemma 39.

1. Die Menge M_D ist konvex.
2. Das Paar $(-\alpha, 0) \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ gehört nicht zu \overline{M}_D für $\alpha > 0$.

Beweis.

1. Wir betrachten zwei Elemente $(z_0, z), (\hat{z}_0, \hat{z}) \in M_D$ mit den Funktionen $(\zeta, \vartheta) \in W_2^{1, n}(\mathbb{R}_{++}, e^{\gamma t}) \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und $(\hat{\zeta}, \hat{\vartheta}) \in W_2^{1, n}(\mathbb{R}_{++}, e^{\gamma t}) \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und zeigen, dass deren Konvexkombinationen selbst Elemente der Menge M_D sind: Aufgrund der Linearität der Operatoren a_1, a_2, b_1, b_2 mit $\lambda \in [0, 1]$ gilt

$$\begin{aligned} (1 - \lambda)z_0 + \lambda\hat{z}_0 &= (1 - \lambda)a_1(\vartheta) + (1 - \lambda)b_1(\zeta) + \lambda a_1(\hat{\vartheta}) + \lambda b_1(\hat{\zeta}) \\ &= a_1((1 - \lambda)\vartheta + \lambda\hat{\vartheta}) + b_1((1 - \lambda)\zeta + \lambda\hat{\zeta}) \end{aligned}$$

und

$$\begin{aligned}(1-\lambda)z + \lambda\hat{z} &= (1-\lambda)a_2(\vartheta) + (1-\lambda)b_2(\zeta) + \lambda a_2(\hat{\vartheta}) + \lambda b_2(\hat{\zeta}) \\ &= a_2((1-\lambda)\vartheta + \lambda\hat{\vartheta}) + b_2((1-\lambda)\zeta + \lambda\hat{\zeta}).\end{aligned}$$

Weiterhin ist $(1-\lambda)\zeta + \lambda\hat{\zeta} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ und $(1-\lambda)\vartheta + \lambda\hat{\vartheta} \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und die folgenden Abschätzungen sind erfüllt:

$$\left\| (1-\lambda)\zeta + \lambda\hat{\zeta} \right\|_{1,(2),e^{\gamma t}} \leq (1-\lambda) \|\zeta\|_{1,(2),e^{\gamma t}} + \lambda \|\hat{\zeta}\|_{1,(2),e^{\gamma t}} \leq 1$$

$$\text{sowie } \left\| (1-\lambda)\vartheta + \lambda\hat{\vartheta} \right\|_{(2),e^{\gamma t}} \leq 1.$$

Damit ist $((1-\lambda)z_0 + \lambda\hat{z}_0, (1-\lambda)z + \lambda\hat{z})$ ein Element der Menge M_D und diese demnach konvex.

2. Wir zeigen nun, dass das Tupel $(-\alpha, 0) \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ für jedes $\alpha > 0$ nicht Element der Menge \overline{M}_D ist: Dafür konstruieren wir ein beliebiges Element $(\hat{z}_0, 0)$ der Menge \overline{M}_D und zeigen, dass für derartige Elemente stets $\hat{z}_0 \geq 0$ gilt. Aufgrund der Abgeschlossenheit der Menge \overline{M}_D folgt für $(\hat{z}_0, 0) \in \overline{M}_D$ die Existenz einer Folge $\{(z_0^N, z^N)\}_{N \in \mathbb{N}}$ mit $(z_0^N, z^N) \in M_D$, die gegen $(\hat{z}_0, 0)$ konvergiert. Die Elemente dieser Folge genügen den Gleichungen

$$z_0^N = a_1(\vartheta^N) + b_1(\zeta^N), \quad (4.18)$$

$$z^N = a_2(\vartheta^N) + b_2(\zeta^N). \quad (4.19)$$

Die zugehörigen Folgen

$$\begin{aligned}\{\zeta^N\}_{N \in \mathbb{N}}, \quad \zeta^N \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \quad & \|\zeta^N\|_{1,(2),e^{\gamma t}} \leq 1 \\ \{\vartheta^N\}_{N \in \mathbb{N}}, \quad \vartheta^N \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}), \quad & \|\vartheta^N\|_{(2),e^{\gamma t}} \leq 1\end{aligned}$$

haben die folgenden Eigenschaften:

Es existiert eine Teilfolge von $\{\zeta^N\}_{N \in \mathbb{N}}$, welche wir wiederum mit $\{\zeta^N\}_{N \in \mathbb{N}}$ bezeichnen, so dass

$$\{\zeta^N\} \rightharpoonup \tilde{\zeta}, \quad \|\tilde{\zeta}\|_{1,(2),e^{\gamma t}} \leq 1$$

sowie eine Teilfolge von $\{\vartheta^N\}_{N \in \mathbb{N}}$, ebenfalls bezeichnet mit $\{\vartheta^N\}_{N \in \mathbb{N}}$, so dass

$$\{\vartheta^N\} \rightharpoonup \tilde{\vartheta}, \quad \|\tilde{\vartheta}\|_{(2),e^{\gamma t}} \leq 1.$$

Dies folgt aus dem Satz von Eberlein-Schmuljan^[17], welcher besagt, dass die abgeschlossenen Einheitskugeln in reflexiven Banachräumen schwach kompakt sind. Unter Verwendung von Lemma 38 können wir in der Gleichung (4.18) den Grenzübergang vollziehen und erhalten

¹⁷vgl. [Göp73], S. 94; [Yos80], S. 141

$$\hat{z}_0 = a_1(\tilde{\vartheta}) + b_1(\tilde{\zeta}).$$

Wir betrachten nun die Gleichung (4.19) und multiplizieren diese mit einer beliebigen Testfunktion $\phi \in C_0^\infty(\mathbb{R}_+, \mathbb{R}^n)$ sowie der Dichtefunktion ν , mit $\nu(t) = e^{\gamma t}$ und $\gamma < 0$ auf ganz \mathbb{R}_+ . Die anschließende Integration über das Intervall $(0, \infty)$ führt uns zu

$$\begin{aligned} \int_0^\infty z^N(t) \phi(t) \nu(t) dt &= \int_0^\infty [a_2(\vartheta^N(t)) + b_2(\zeta^N(t))]^T \phi(t) e^{\gamma t} dt \\ &= - \int_0^\infty [\vartheta^N(t)]^T \phi(t) e^{\gamma t} dt + \int_0^\infty [(\zeta^N)'(t)]^T \phi(t) e^{\gamma t} dt \quad (4.20) \\ &\quad + \int_0^\infty [\hat{A}^T(t) \zeta^N(t)]^T \phi(t) e^{\gamma t} dt. \end{aligned}$$

Die Funktionen $\phi(\cdot)$ und $\hat{A}^T(\cdot)\phi(\cdot)$ sind aufgrund der Raumzugehörigkeit von $\phi \in C_0^\infty(\mathbb{R}_+, \mathbb{R}^n)$ und $\hat{A} \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ Elemente des Raumes $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$. Damit sind die in (4.20) auftretenden Integranden linear und beschränkt bzgl. der Norm $\|\cdot\|_{(2), e^{\gamma t}}$ und somit insbesondere auch stetig bzgl. der Normtopologie des Raumes $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$. Daher betrachten wir nun lineare und stetige Funktionale, für welche wir schließlich den schwachen Grenzwertprozess vollziehen können. Dieser liefert

$$\begin{aligned} 0 &= \lim_{N \rightarrow \infty} \int_0^\infty z^N(t) \phi(t) \nu(t) dt = \int_0^\infty \{a_2(\tilde{\vartheta}(t)) + b_2(\tilde{\zeta}(t))\} \phi(t) e^{\gamma t} dt \\ &= \langle a_2(\tilde{\vartheta}) + b_2(\tilde{\zeta}), \phi \rangle_{(2), e^{\gamma t}}. \end{aligned}$$

Der Raum $C_0^\infty(\mathbb{R}_+, \mathbb{R}^n)$ liegt dicht¹⁸ im Raum $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$. Damit erhalten wir aus der Variationsgleichung

$$\langle a_2(\tilde{\vartheta}) + b_2(\tilde{\zeta}), \phi \rangle_{(2), e^{\gamma t}} = 0 \quad \forall \phi \in C_0^\infty(\mathbb{R}_+, \mathbb{R}^n)$$

die Bedingung

$$0 \equiv a_2(\tilde{\vartheta}) + b_2(\tilde{\zeta})$$

im L_2 -Sinn. Damit ist $(\hat{z}_0, 0) \in \overline{M}_D$. Wir wollen nun zeigen, dass für derartige Paare $\hat{z}_0 \geq 0$ gilt. Wir beweisen diese Aussage mit Hilfe eines Widerspruchsbeweises:

Sei $(-\alpha, 0) \in \overline{M}_D$ und $\alpha > 0$. Dann können wir auf die Existenz eines Paares (ζ, ϑ) mit $\zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ und $\|\zeta\|_{1,(2), e^{\gamma t}} \leq 1$ sowie $\vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und $\|\vartheta\|_{(2), e^{\gamma t}} \leq 1$ schließen, für das gilt

$$a_2(\vartheta) + b_2(\zeta) = 0,$$

¹⁸vgl. beispielsweise [Zie19], Lemma 2.4.3, S. 38

wobei $\zeta = y - y_*$ und $\vartheta = q - q_*$ und $(y, q) \in \mathcal{B}_D$. Schritt 1 liefert aufgrund der Optimalität von (y_*, q_*) für $(\hat{D})^\infty$ zudem die Relation

$$-\alpha = a_1(\vartheta) + b_1(\zeta) = \delta^+ \hat{G}((y_*, q_*); (\zeta, \vartheta)) \geq 0.$$

Dies steht allerdings im Widerspruch zu $\alpha > 0$. Damit gilt offenbar $(-\alpha, 0) \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ ist kein Element der Menge \overline{M}_D . \overline{M}_D entspricht nicht dem gesamten Raum $\mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$.

□

Schritt 4: (Trennung der Mengen)

Wir nutzen den Trennungssatz von Hahn-Banach¹⁹, um die abgeschlossene und konvexe Menge \overline{M}_D von der kompakten Menge $\{(-\alpha, 0)\} \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$, $\alpha > 0$ zu trennen – welche gemäß Lemma 39 disjunkte Mengen im $\mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ sind – und erhalten somit die Existenz des stetigen und linearen Funktionals $g_\alpha^* \in \mathbb{R} \times L_2^n(\mathbb{R}_{++}, e^{\gamma t})$, so dass für alle $(z_0, z) \in \overline{M}_D$ und mit $\mu_{1,\alpha}, \mu_{2,\alpha} > 0$ die folgende Ungleichung gilt:

$$\langle g_\alpha^*, (-\alpha, 0) \rangle \leq \mu_{1,\alpha} \leq \mu_{2,\alpha} \leq \langle g_\alpha^*, (z_0, z) \rangle. \quad (4.21)$$

Mit $g_\alpha^* = (\lambda_\alpha, p_\alpha)$ liest sich die Ungleichung (4.21) wie folgt:

$$-\lambda_\alpha \alpha < \lambda_\alpha z_0 + \langle p_\alpha, z \rangle_{(2), e^{\gamma t}} \quad \forall (z_0, z) \in \overline{M}_D.$$

Natürlich gehört das Paar $(z_0 = 0, z = 0)$ zu \overline{M}_D , daher muss obige Ungleichung auch für dieses Element erfüllt sein, so dass $\lambda_\alpha > 0$ für alle $\alpha > 0$ folgt. Wir dividieren durch λ_α und bezeichnen im Folgenden $p_\alpha := \frac{p_\alpha}{\lambda_\alpha}$ und erhalten damit:

$$-\alpha < z_0 + \langle p_\alpha, z \rangle_{(2), e^{\gamma t}} \quad \forall (z_0, z) \in \overline{M}_D. \quad (4.22)$$

Die Formel (4.22) beinhaltet alle wesentlichen Informationen, um daraus die Transversalitätsbedingungen und die Maximumbedingung zu schlussfolgern. In der Terminologie der Operatoren a_1, a_2, b_1 und b_2 lesen sich diese wie folgt:

$$\begin{aligned} -\alpha < a_1(\vartheta) + b_1(\zeta) + \langle p_\alpha, a_2(\vartheta) + b_2(\zeta) \rangle_{(2), e^{\gamma t}} \quad \forall (z_0, z) \in \overline{M}_D. \quad (4.23) \\ \forall \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \|\zeta\|_{(1), 2, e^{\gamma t}} \leq 1 \\ \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}), \|\vartheta\|_{2, e^{\gamma t}} \leq 1 \\ \forall \alpha > 0. \end{aligned}$$

Schritt 5: (Maximumbedingung)

Da die obige Ungleichung (4.23) für alle $\alpha > 0$ erfüllt ist, gilt diese insbesondere auch mit der Setzung $\alpha := \frac{1}{n}$, $n \in \mathbb{N}$, $p_n := p_\alpha$. Ziel ist es zunächst, die Beschränktheit der Folge $\{p_n\}_{n \in \mathbb{N}}$ auf dem gesamten $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ nachzuweisen. Dafür halten wir zunächst $y = y_*$ fest und variieren ausschließlich über die Steuerung. Gelingt uns der Nachweis der Beschränktheit der Folge $\{p_n\}_{n \in \mathbb{N}}$ auf dem gesamten $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$, so können wir mit Hilfe des Satzes von Eberlein-Schmuljan auf die Existenz einer schwach konvergenten Teilfolge schließen, mit deren Hilfe wir letztlich wiederum auf Basis der Ungleichung (4.23) auf die

¹⁹vgl. [Wer07], S: 116, Theorem III.2.4

Maximumbedingung und die kanonische Gleichung schließen können.

Sei nun $y = y_*$, d.h. $\zeta = y - y_* = 0$, dann erhalten wir aus Gleichung (4.23)

$$-\alpha < a_1(\vartheta) + \langle p_\alpha, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \quad \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}), \|\vartheta\|_{(2), e^{\gamma t}} \leq 1. \quad (4.24)$$

Unter Berücksichtigung von $\|\vartheta\|_{(2), e^{\gamma t}} \leq 1$ führt uns dies zu der Abschätzung

$$-\alpha \|\vartheta\|_{(2), e^{\gamma t}} < a_1(\vartheta) + \langle p_\alpha, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \quad \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}).$$

Aufgrund der Linearität von a_1 und a_2 folgt dann

$$\left| a_1(\vartheta) + \langle p_\alpha, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \right| \leq \alpha \|\vartheta\|_{(2), e^{\gamma t}} \quad \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}),$$

woraus wir schließlich für alle $\vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und alle $\alpha > 0$ die Relation

$$\left| \langle p_\alpha, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \right| - |a_1(\vartheta)| \leq \left| a_1(\vartheta) + \langle p_\alpha, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \right| \leq \alpha \|\vartheta\|_{(2), e^{\gamma t}}$$

erhalten. Damit folgt aufgrund der Beschränktheit von a_1 mit der Setzung $\alpha := \frac{1}{n}$, mit $n \in \mathbb{N}$ und $p_n := p_\alpha$

$$\left| \langle p_n, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \right| = \left| \langle p_n, -\vartheta \rangle_{(2), e^{\gamma t}} \right| \leq c \|\vartheta\|_{2, e^{\gamma t}} \quad \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$$

und die Folge $\{p_n\}_{n \in \mathbb{N}}$ ist beschränkt auf

$$Z_D := \left\{ z \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \mid z = a_2(\vartheta), \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}) \right\}.$$

Da $a_2(\vartheta) = -\vartheta$, entspricht die Menge Z_D offensichtlich dem $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$. Demnach ist die Folge $\{p_n\}_{n \in \mathbb{N}}$ beschränkt auf $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ und besitzt nach dem Satz von Eberlein-Smuljan als beschränkte Folge im reflexiven Banachraum eine schwach konvergente Teilfolge, wiederum bezeichnet mit $\{p_n\}_{n \in \mathbb{N}}$, die gegen $\hat{p} \in L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ konvergiert. Der Grenzprozess liefert schließlich mit (4.24)

$$a_1(\vartheta) + \langle \hat{p}, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} = 0 \quad \forall \vartheta \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}). \quad (4.25)$$

Aufgrund der Konvexität von \hat{G} bzgl. q (dies folgt aus der positiven Definitheit der Matrix W^{-1}) erhalten wir mit $\vartheta := q - q_*$

$$a_1(\vartheta) \leq \int_0^\infty \frac{1}{2} \left\{ -q_*^T(t) W^{-1}(t) q_*(t) + q^T(t) W^{-1}(t) q(t) \right\} e^{\gamma t} dt.$$

Somit können wir die Gleichung (4.25) abschätzen und erhalten

$$\begin{aligned} 0 &= a_1(\vartheta) + \langle \hat{p}, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \\ &\leq \int_0^\infty \frac{1}{2} \left\{ -q_*^T(t) W^{-1}(t) q_*(t) + q^T(t) W^{-1}(t) q(t) \right\} e^{\gamma t} dt + \langle \hat{p}, a_2(\vartheta) \rangle_{(2), e^{\gamma t}} \end{aligned}$$

und damit

$$0 \leq \int_0^\infty \left\{ \frac{1}{2} \left[q_*^T(t) \hat{W}^{-1}(t) q_*(t) + q^T(t) \hat{W}^{-1}(t) q(t) \right] - (q(t) - q_*(t))^T \hat{p}(t) \right\} e^{\gamma t} dt.$$

Die letzte Gleichung ist die integrierte Maximumbedingung

$$\int_0^\infty H(t, y_*(t), q_*(t), \hat{p}(t), \lambda_0) dt \geq \int_0^\infty H(t, y_*(t), q(t), \hat{p}(t), \lambda_0) dt \quad \forall q \in L_2^n(\mathbb{R}_{++}, e^{\gamma t}).$$

Folgen wir anschließend den Techniken in [Wag09](#)²⁰ so folgt die punktweise Bedingung (M) für fast alle $t \in \mathbb{R}_{++}$.

Schritt 6: (Kanonische Gleichungen)

Wir betrachten wieder die Ungleichung [\(4.23\)](#) und variieren nun, bei festgehaltenem $q = q_*$, über den Zustand. Damit erhalten wir die Relation

$$-\alpha < b_1(\zeta) + \langle p_\alpha, b_2(\zeta) \rangle_{(2), e^{\gamma t}} \quad \forall \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \|\zeta\|_{1,(2), e^{\gamma t}} \leq 1.$$

Mit den gleichen Überlegungen wie im Beweisschritt 5 erhalten wir mit der Setzung $\alpha := \frac{1}{n}$, $n \in \mathbb{N}$ und $p_n := p_\alpha$ aufgrund der Linearität von b_1 und b_2 die Ungleichung

$$\left| b_1(\zeta) + \langle p_n, b_2(\zeta) \rangle_{(2), e^{\gamma t}} \right| \leq \frac{1}{n} \|\zeta\|_{1,(2), e^{\gamma t}}.$$

Der Grenzübergang liefert schließlich

$$b_1(\zeta) + \langle \hat{p}, b_2(\zeta) \rangle_{(2), e^{\gamma t}} = 0 \quad \forall \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}). \quad (4.26)$$

Die Gleichung [\(4.26\)](#) gilt damit insbesondere auch $\forall \zeta \in \overset{\circ}{W}_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$, wobei

$$\overset{\circ}{W}_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) := \left\{ \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \mid \zeta(0) = 0 \right\}.$$

Daher können wir die Gleichung [\(4.26\)](#) schreiben als

$$\begin{aligned} 0 &= \int_0^\infty \left\{ \nabla_{\xi_D}^T \theta(t, y_*(t)) \zeta(t) + \hat{p}^T(t) \left[\zeta'(t) + \hat{A}^T(t) \zeta(t) \right] \right\} e^{\gamma t} dt \\ &= \int_0^\infty \left\{ \left[\hat{p}(t) e^{\gamma t} \right]^T \zeta'(t) + \left[\nabla_{\xi_D} \theta(t, y_*(t)) e^{\gamma t} + \hat{A}(t) \hat{p}(t) e^{\gamma t} \right]^T \zeta(t) \right\} dt \end{aligned} \quad (4.27)$$

$$\forall \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \zeta(0) = 0.$$

Gleichung [\(4.27\)](#) ist eine Variationsgleichung, welche mittels der Definition zeigt, dass $\hat{p}e^{\gamma t}$ eine verallgemeinerte Ableitung besitzt²¹ mit

$$\mathcal{D}(\hat{p}(\cdot)e^{\gamma \cdot}) = \nabla_{\xi_D} \theta(\cdot, y_*(\cdot)) e^{\gamma \cdot} + \hat{A}(\cdot) \hat{p}(\cdot) e^{\gamma \cdot}. \quad (4.28)$$

Unter Nutzung der Definition der Pontryaginschen Funktion erhalten wir mit $p_0 := \hat{p}e^{\gamma t}$ die kanonische Gleichung

$$p_0'(t) = -\nabla_{\xi_D} H(t, y_*(t), q_*(t), p_0(t), 1) \quad (\forall) t \in \mathbb{R}_{++}.$$

²⁰vgl. S. 553-554, Beweis zu Theorem 2.3

²¹vgl. [Yos80](#) S. 49

Jetzt betrachten wir die Gleichung (4.26) noch einmal und variieren folgendermaßen:

$$\begin{aligned}\zeta(t) &= \hat{\zeta}(t) + D \quad \text{mit einer beliebigen Konstante } D, \\ \zeta(0) &= D, \quad \hat{\zeta}(0) = 0.\end{aligned}$$

Mit dieser Setzung folgt

$$b_1(\zeta) = \int_0^\infty \nabla_{\xi_D}^T \theta(t, y_*(t)) (\hat{\zeta}(t) + D) e^{\gamma t} dt + x_0^T D.$$

Gemäß (4.26) gilt zudem

$$b_1(\zeta) + \langle \hat{p}, b_2(\zeta) \rangle_{2, e^{\gamma t}} = 0 \quad \forall \zeta \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$$

und somit

$$\begin{aligned}\int_0^\infty \nabla_{\xi_D}^T \theta(t, y_*(t)) (\hat{\zeta}(t) + D) e^{\gamma t} dt + x_0^T D + \int_0^\infty \left\{ \hat{p}^T(t) \left[\hat{\zeta}'(t) + \hat{A}^T(t) (\hat{\zeta}(t) + D) \right] \right\} e^{\gamma t} dt &= 0 \\ \forall D \in \mathbb{R}^n, \forall \hat{\zeta} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \hat{\zeta}(0) = 0.\end{aligned}$$

Damit gilt für alle $D \in \mathbb{R}^n$ und für alle $\hat{\zeta} \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ mit $\hat{\zeta}(0) = 0$

$$\begin{aligned}\underbrace{\int_0^\infty \left\{ \nabla_{\xi_D}^T \theta(t, y_*(t)) \hat{\zeta}(t) + \hat{p}^T(t) \left[\hat{\zeta}'(t) + \hat{A}^T(t) \hat{\zeta}(t) \right] \right\} e^{\gamma t} dt + x_0^T D}_{=0} \\ + \int_0^\infty \left\{ \nabla_{\xi_D}^T \theta(t, y_*(t)) D + \hat{p}^T(t) \hat{A}^T(t) D \right\} e^{\gamma t} dt = 0.\end{aligned}$$

Für die Setzung $D = \mathbf{1}_n := (1, 1, \dots, 1)^T \in \mathbb{R}^n$ gilt somit:

$$\int_0^\infty \left[\nabla_{\xi_D} \theta(t, y_*(t)) + \hat{A}(t) \hat{p}(t) \right]^T \cdot \mathbf{1}_n e^{\gamma t} dt + x_0 \stackrel{(4.28)}{=} \int_0^\infty [\hat{p}^T(t) e^{\gamma t}]' \cdot \mathbf{1}_n dt + x_0 = x_0 - \hat{p}(0) = 0,$$

was schließlich die Bedingung

$$\hat{p}(0) = x_0 \quad \text{bzw.} \quad p_0(0) = \hat{p}(0) e^{\gamma \cdot 0} = x_0$$

liefert.

Schritt 7: (Transversalitätsbedingung)

Die adjungierte Funktion p_0 gehört zum $W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$, da

$$\begin{aligned}\int_0^\infty |p_0(t)|^2 e^{-\gamma t} dt &= \int_0^\infty |\hat{p}(t) e^{\gamma t}|^2 e^{-\gamma t} dt = \int_0^\infty |\hat{p}(t)|^2 e^{\gamma t} dt < \infty \quad \text{und} \\ \int_0^\infty |p_0'(t)|^2 e^{-\gamma t} dt &= \int_0^\infty \left| \nabla_{\xi_D} \theta(t, y_*(t)) + \hat{A}(t) \hat{p}(t) \right| e^{\gamma t} dt < \infty.\end{aligned}$$

Da $\gamma < 0$, ist $e^{\gamma t}$ eine echte Dichtefunktion. In dem Fall schließen wir

$$\|y_* p_0\|_{W_1^1(\mathbb{R}_{++})} \leq \|y_*\|_{W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})} \|p_0\|_{W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})}.$$

Damit liegt $y_* p_0$ im $W_1^1(\mathbb{R}_{++})$ und ²²

$$\lim_{t \rightarrow \infty} y_*(t) p_0(t) = 0, \quad \lim_{t \rightarrow \infty} p_0(t) = 0.$$

Damit sind alle Behauptungen des Satzes bewiesen. \square

Bemerkung 21. Die Maximumbedingung liefert die Bedingung

$$q_*(t) = W(t) p_0(t) e^{-\gamma t} \quad (\forall t \in \mathbb{R}_{++}), \quad (4.29)$$

wobei $W \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ und $p_0 \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t})$. Damit ist auch $q_* \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$ und eine punktweise Auswertung der Gleichung (4.29) zulässig, denn

$$\int_0^\infty |q_*(t)|^2 e^{\gamma t} dt = \int_0^\infty |W(t) p_0(t) e^{-\gamma t}|^2 e^{\gamma t} dt \leq \overline{W}^2 \int_0^\infty |p_0(t)|^2 e^{-\gamma t} dt < \infty.$$

Die Ableitungen der auftretenden Funktionen existieren ebenfalls gemäß der Voraussetzungen, so dass weiterhin gilt

$$\int_0^\infty |q'_*(t)|^2 e^{\gamma t} dt < \infty.$$

Damit ist aber auch $y'_*(\cdot) = -\hat{A}^T(\cdot) y_*(\cdot) + q_*(\cdot)$ mit $\hat{A} \in L_\infty^{n \times n}(\mathbb{R}_{++}) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ Element des Raumes $W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t})$, denn

$$\begin{aligned} \int_0^\infty |y'_*(t)|^2 e^{\gamma t} dt &= \int_0^\infty |-\hat{A}^T(t) y_*(t) + q_*(t)|^2 e^{\gamma t} dt \\ &\leq 2 \left(\overline{\hat{A}}^2 \int_0^\infty |y_*(t)|^2 e^{\gamma t} dt + \int_0^\infty |q_*(t)|^2 e^{\gamma t} dt \right) = 2 \left(\overline{\hat{A}}^2 \|y_*\|_{(2), e^{\gamma t}} + \|q_*\|_{(2), e^{\gamma t}} \right) < \infty. \end{aligned}$$

Auch hier existieren die einzelnen Ableitungen, so dass zudem gilt

$$\int_0^\infty |y''_*(t)|^2 e^{\gamma t} dt < \infty.$$

Damit ist $y_* \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t})$. Weiterhin folgt aus der Beziehung

$$q_*(t) = W(t) p_0(t) e^{-\gamma t} \quad \forall t \in \mathbb{R}_+$$

zusammen mit

$$q_*(t) = y'_*(t) + \hat{A}^T(t) y_*(t) \quad \forall t \in \mathbb{R}_+$$

und der Bedingung

$$p_0(0) = x_0$$

²²vgl. hierfür Lemma ³⁰, S. ⁶³

$$q_*(0) = y'_*(0) + \hat{A}^T(0)y_*(0) = W(0)p_0(0) = W(0)x_0$$

die Gültigkeit der Relation

$$x_0 = W^{-1}(0) \left[y'_*(0) + \hat{A}^T(0)y_*(0) \right].$$

Damit stellt die optimale Lösung (y_*, q_*) für das Teilproblem (\hat{D}) unter den Voraussetzungen (VD1)-(VD4) mit $\hat{\mathcal{B}}_D \supset \mathcal{B}_D$ also auch eine optimale Lösung für (D_ν) dar.

Wir fassen die Erkenntnisse zusammen und formulieren

Folgerung 4. Seien die Voraussetzungen (VD1)-(VD4) erfüllt, dann gilt: Jede Lösung $(y_*, q_*) \in \hat{\mathcal{B}}_D$ vom (\hat{D}) ist auch Element der zulässigen Menge

$$\mathcal{B}_D := \left\{ (y, q) \left| \begin{array}{l} y \in W_2^{2,n}(\mathbb{R}_{++}, e^{\gamma t}), \quad y'(t) = -\hat{A}^T(t)y(t) + q(t) \quad \forall t \in \mathbb{R}_+ \\ q \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}), \quad x_0 = W^{-1}(0) \left[y'(0) + \hat{A}^T(0)y(0) \right] \end{array} \right. \right\}.$$

von (D_ν) . Damit ist jede Lösung (y_*, q_*) von (\hat{D}) auch Lösung von (D_ν) .

Die Ausführungen dieses Kapitels ermöglichen nun zunächst die Formulierung des folgenden Satzes:

Satz 19. Sei ein Steuerungsproblem vom Typ (P_{ν_i}) , mit $i = 0, 1$ gegeben. Existiert eine optimale Lösung (x_*, u_*) der primalen Aufgabe und seien die Voraussetzungen (VD1)-(VD3) und $\mathcal{B}_D \neq \emptyset$ erfüllt. Dann gilt für die Probleme (D_{ν_i}) , mit $i = 0, 1$

$$G_{\nu_i}(y, q) \rightarrow \sup! \text{ bzgl. aller } (y, q) \in \mathcal{B}_D$$

die Beziehung

$$J(x_*, u_*) = \inf(P_{\nu_i}) = \sup(D_{\nu_i}).$$

Beweis. Wir betrachten die eindeutige optimale Lösung $(x_*, u_*) \in \mathcal{B}_P \subset X_0$ von (P_{ν_i}) . Aufgrund der nachgewiesenen Äquivalenzbedingung gilt

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = \inf_{(x,u) \in X_0} \left[\sup_{y \in Y_{\nu_i}} \Phi_i((x, u), y) \right].$$

Wir setzen nun für alle $t \in \mathbb{R}_+$

$$\hat{W}^{-1}(t) := \begin{cases} W^{-1}(t), & \text{für } \gamma = \beta_0 \\ W^{-1}(t)e^{\gamma t}, & \text{für } \gamma = -\beta_1 \end{cases}$$

und wählen ein beliebiges Paar (\hat{x}_y, \hat{u}_y) gemäß der Vorschrift

$$\hat{x}_y(t) := \hat{W}^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] \quad \text{und} \quad \hat{u}_y(t) := \begin{cases} \hat{\sigma}(t, y(t)), & \text{für } \gamma = \beta_0 \\ \hat{\sigma}(t, y(t))e^{\gamma t}, & \text{für } \gamma = -\beta_1 \end{cases}$$

für fast alle $t \in \mathbb{R}_+$, wobei y Element der nichtleeren Menge²³ Y_{ν_i} sei. Aufgrund der Raumeigenschaften der auftretenden Funktionen sowie der Erkenntnisse aus Lemma [35](#) gilt unter den getätigten Voraussetzungen für diese Setzung offenbar

$$(\hat{x}_y, \hat{u}_y) \in X_0 := \left\{ (x, u) \in W_2^{1,n}(\mathbb{R}_{++}, e^{\gamma t}) \times L_2^m(\mathbb{R}_{++}, e^{\gamma t}) \mid x(0) - x_0 = 0, u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_{++} \right\}$$

²³Dies folgt aus der Voraussetzung $\mathcal{B}_D \neq \emptyset$.

und somit

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in X_0} \left[\sup_{y \in Y_{\nu_i}} \Phi_i((x,u), y) \right] \leq \sup \{ \Phi_i((\hat{x}_y, \hat{u}_y), y) \mid y \in Y_{\nu_i} \}.$$

Mit dieser Setzung können wir nun weiter abschätzen

$$\begin{aligned} J_{\nu_i}(x_*, u_*) &\leq \sup \{ \Phi_i((\hat{x}_y, \hat{u}_y), y) \mid y \in Y_{\nu_i} \} = \sup \left\{ J_{2,y}^{\nu_i}(\hat{x}_y) + J_{1,y}^{\nu_i}(\hat{u}_y) \mid y \in Y_{\nu_i} \right\} \\ &= \sup_{y \in Y_{\nu_i}} \left\{ \int_0^\infty \left\{ -\frac{1}{2} [y'(t) + \hat{A}^T(t)y(t)]^T W^{-1}(t) [y'(t) + \hat{A}^T(t)y(t)] \right. \right. \\ &\quad \left. \left. + \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) - B^T(t)y(t) + \hat{r}(t) \right] \right\} e^{\gamma t} dt - x_0^T y(0) \right\} \\ &= \sup_{(y,q) \in \mathcal{B}_D} \left\{ \int_0^\infty \left\{ -\frac{1}{2} q^T(t) W^{-1}(t) q(t) \right. \right. \\ &\quad \left. \left. + \hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) - B^T(t)y(t) + \hat{r}(t) \right] \right\} e^{\gamma t} dt - x_0^T y(0) \right\} \\ &= \sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q). \end{aligned}$$

Damit erhalten wir die Abschätzung

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) \leq \sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q). \quad (4.30)$$

Gleichzeitig gilt nach Satz [17](#) die schwache Dualitätsbeziehung und somit

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) \geq \sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q). \quad (4.31)$$

Gleichungen [\(4.30\)](#) und [\(4.31\)](#) liefern zusammen die Beziehung

$$J_{\nu_i}(x_*, u_*) = \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = \sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q).$$

□

4.6 Formulierung eines starken Dualitätssatzes

Im vorangegangenen Unterkapitel waren wir in der Lage, notwendige Optimalitätsbedingungen für die duale Aufgabe anzugeben. Offen ist jedoch die Frage nach der Existenz einer optimalen Lösung der dualen Aufgabe. Ziel dieses Unterkapitels ist daher die Formulierung von Voraussetzungen, unter denen die Existenz einer derartigen Lösung gesichert werden kann. Dafür definieren wir zunächst die Mengen

$$\tilde{Y}_{\nu_0} := \left\{ (y, q) \in \mathcal{B}_D \left| \begin{array}{l} y''(t) = -\hat{W}(t) (\hat{W}^{-1})'(t) q(t) - \left[(\hat{A}'(t))^T y(t) + \hat{A}^T(t) y'(t) \right] \\ \quad \quad \quad + \hat{W}(t) A(t) \hat{W}^{-1}(t) q(t) + \hat{W}(t) B(t) \hat{\sigma}(t, y(t)) \\ (\forall) t \in \mathbb{R}_{++} \end{array} \right. \right\}$$

und

$$\tilde{Y}_{\nu_i} := \left\{ (y, q) \in \mathcal{B}_D \left| \begin{array}{l} y''(t) = -\hat{W}(t) \left(\hat{W}^{-1} \right)'(t) q(t) - \left[\left(\hat{A}'(t) \right)^T y(t) + \hat{A}^T(t) y'(t) \right] \\ \quad + \hat{W}(t) A(t) \hat{W}^{-1}(t) q(t) + \hat{W}(t) B(t) \hat{\sigma}(t, y(t)) e^{\gamma t} \\ (\forall) t \in \mathbb{R}_{++} \end{array} \right. \right\}$$

und formulieren auf Grundlage dieser den folgenden Satz:

Satz 20 (Starker Dualitätssatz). *Sei ein Steuerungsproblem vom Typ (P_{ν_i}) mit $i = 0, 1$ gegeben. Sind die Voraussetzungen (V1)-(V4) und (VD1)-(VD4) erfüllt und gilt $\tilde{Y}_{\nu_i} \neq \emptyset$, dann gilt für die Probleme (D_{ν_i}) , mit $i = 0, 1$*

$$G_{\nu_i}(y, q) \rightarrow \sup! \text{ bzgl. aller } (y, q) \in \mathcal{B}_D$$

die starke Dualitätsbeziehung, d.h. es existieren Paare $(x_*, u_*) \in \mathcal{B}_P$ und $(y_*, q_*) \in \mathcal{B}_D$, so dass gilt

$$\inf(P_{\nu_i}) = J_{\nu_i}(x_*, u_*) = G_{\nu_i}(y_*, q_*) = \sup(D_{\nu_i}).$$

Beweis. Sei durch (y_*, q_*) ein beliebiges Paar der nichtleeren Menge $\tilde{Y}_{\nu_i} \subset \mathcal{B}_D$ gegeben, dann gilt

$$\begin{aligned} \sup_{(y, q) \in \mathcal{B}_D} G_{\nu_i}(y, q) &\geq G_{\nu_i}(y_*, q_*) \\ &= \int_0^\infty \left\{ -\frac{1}{2} q_*^T(t) W^{-1}(t) q_*(t) \right. \\ &\quad \left. + \hat{\sigma}^T(t, y_*(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y_*(t)) - B^T(t) y_*(t) + \hat{r}(t) \right] \right\} e^{\gamma t} dt - x_0^T y_*(0) \\ &= \int_0^\infty \left\{ -\frac{1}{2} \left[y_*'(t) + \hat{A}^T(t) y_*(t) \right]^T W^{-1}(t) \left[y_*'(t) + \hat{A}^T(t) y_*(t) \right] \right. \\ &\quad \left. + \hat{\sigma}^T(t, y_*(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y_*(t)) - B^T(t) y_*(t) + \hat{r}(t) \right] \right\} e^{\gamma t} dt - x_0^T y_*(0). \end{aligned}$$

Wir erinnern an die Definition der Funktion \hat{W}^{-1} und führen die folgenden Substitutionen ein:

$$a_y(t) := \hat{W}^{-1}(t) \left[y'(t) + \hat{A}^T(t) y(t) \right] \Leftrightarrow \left[y'(t) + \hat{A}^T(t) y(t) \right] = \hat{W}(t) a_y(t) \quad (\forall) t \in \mathbb{R}_{++}$$

$$\text{und } b_y(t) := \begin{cases} \hat{\sigma}(t, y(t)), & \text{für } \gamma = \beta_0 \\ \hat{\sigma}(t, y(t)) e^{\gamma t}, & \text{für } \gamma = -\beta_1 \end{cases} \quad (\forall) t \in \mathbb{R}_{++}.$$

Aufgrund der Raumeigenschaften der auftretenden Funktionen sowie der Bedingung $(y_*, q_*) \in \tilde{Y}_{\nu_i} \subset \mathcal{B}_D$ gilt $(a_{y_*}, b_{y_*}) \in X_0$. Mit der eingeführten Substitution gilt zudem

$$a'_{y_*}(t) = A(t) a_{y_*}(t) + B(t) b_{y_*}(t) \quad (\forall) t \in \mathbb{R}_{++}.$$

Wir zeigen dies für den Fall $\gamma = \beta_0$: Durch die Wahl von $(y_*, q_*) \in \tilde{Y}_{\nu_0}$ gilt $(\forall)t \in \mathbb{R}_{++}$

$$\begin{aligned}
a'_{y_*}(t) &= \left(\hat{W}^{-1}\right)'(t) \left[y'_*(t) + \hat{A}^T(t)y_*(t) \right] + \hat{W}^{-1}(t) \left[y''_*(t) + \left(\hat{A}^T(t)\right)' y_*(t) + \hat{A}^T(t)y'_*(t) \right] \\
&= \left(\hat{W}^{-1}\right)'(t)q_*(t) + \hat{W}^{-1}(t) \left\{ -\hat{W}(t) \left(\hat{W}^{-1}\right)'(t)q_*(t) - \left[\left(\hat{A}'(t)\right)^T y_*(t) + \hat{A}^T(t)y'_*(t) \right] \right. \\
&\quad \left. + \hat{W}(t)A(t)\hat{W}^{-1}(t)q_*(t) + \hat{W}(t)B(t)\hat{\sigma}(t, y_*(t)) + \left(\hat{A}^T(t)\right)' y_*(t) + \hat{A}^T(t)y'_*(t) \right\} \\
&= \left(\hat{W}^{-1}\right)'(t)q_*(t) - \left(\hat{W}^{-1}\right)'(t)q_*(t) + A(t)\hat{W}^{-1}(t)q_*(t) + B(t)\hat{\sigma}(t, y_*(t)) \\
&= A(t)\hat{W}^{-1}(t) \left[y'_*(t) + \hat{A}^T(t)y_*(t) \right] + B(t)\hat{\sigma}(t, y_*(t)) = A(t)a_{y_*}(t) + B(t)u_{y_*}(t).
\end{aligned}$$

Wir betrachten nun den Fall $\gamma = -\beta_1$: Durch die Wahl von $(y_*, q_*) \in \tilde{Y}_{\nu_1}$ gilt $(\forall)t \in \mathbb{R}_{++}$

$$\begin{aligned}
a'_{y_*}(t) &= \left(\hat{W}^{-1}\right)'(t) \left[y'_*(t) + \hat{A}^T(t)y_*(t) \right] + \hat{W}^{-1}(t) \left[y''_*(t) + \left(\hat{A}^T(t)\right)' y_*(t) + \hat{A}^T(t)y'_*(t) \right] \\
&= \left(\hat{W}^{-1}\right)'(t)q_*(t) + \hat{W}^{-1}(t) \left\{ -\hat{W}(t) \left(\hat{W}^{-1}\right)'(t)q_*(t) - \left[\left(\hat{A}'(t)\right)^T y_*(t) + \hat{A}^T(t)y'_*(t) \right] \right. \\
&\quad \left. + \hat{W}(t)A(t)\hat{W}^{-1}(t)q_*(t) + \hat{W}(t)B(t)\hat{\sigma}(t, y_*(t))e^{\gamma t} + \left(\hat{A}^T(t)\right)' y_*(t) + \hat{A}^T(t)y'_*(t) \right\} \\
&= \left(\hat{W}^{-1}\right)'(t)q_*(t) - \left(\hat{W}^{-1}\right)'(t)q_*(t) + A(t)\hat{W}^{-1}(t)q_*(t) + B(t)\hat{\sigma}(t, y_*(t))e^{\gamma t} \\
&= A(t)\hat{W}^{-1}(t) \left[y'_*(t) + \hat{A}^T(t)y_*(t) \right] + B(t)\hat{\sigma}(t, y_*(t))e^{\gamma t} = A(t)a_{y_*}(t) + B(t)u_{y_*}(t).
\end{aligned}$$

Damit ist das Paar (a_{y_*}, b_{y_*}) für beide Fälle auch Element der Menge X_1 und es gilt $(a_{y_*}, b_{y_*}) \in X_0 \cap X_1 = \mathcal{B}_P$. Mit diesem Wissen gelangen wir zu der weiteren Abschätzungskette

$$\begin{aligned}
G_{\nu_i}(y_*, q_*) &= G(a_{y_*}, b_{y_*}) \\
&= \int_0^\infty \left\{ -\frac{1}{2} \left[\hat{W}(t)a_{y_*}(t) \right]^T W^{-1}(t) \left[\hat{W}(t)a_{y_*}(t) \right] \right. \\
&\quad \left. + b_{y_*}^T(t) \left[\frac{1}{2}R(t) + -B^T(t)y_*(t) + \hat{r}(t) \right] \right\} e^{\gamma t} dt - x_0^T y_*(0) \\
&= \Phi_{\nu_i}((a_{y_*}, b_{y_*}), y_*) \geq \inf_{(x,u) \in \mathcal{B}_P} \Phi_{\nu_i}((x, u), y_*) \\
&= \inf_{(x,u) \in \mathcal{B}_P} \left\{ J_{\nu_i}(x, u) + \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y_* \rangle_{\nu_i} \right\}.
\end{aligned}$$

Da $(x, u) \in \mathcal{B}_P$ gilt $\mathcal{A}_1(x, u) - \mathbf{b}_1 = 0$. Daher können wir weiter schließen:

$$\begin{aligned}
G_{\nu_i}(y_*, q_*) &\geq \inf_{(x,u) \in \mathcal{B}_P} \left\{ J_{\nu_i}(x, u) + \langle \mathcal{A}_1(x, u) - \mathbf{b}_1, y_* \rangle_{\nu_i} \right\} \\
&= \inf_{(x,u) \in \mathcal{B}_P} \left\{ J_{\nu_i}(x, u) + \langle 0, y_* \rangle_{\nu_i} \right\} = J_{\nu_i}(x_*, u_*)
\end{aligned}$$

und somit $\sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q) \geq G_{\nu_i}(y_*, q_*) \geq \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = J_{\nu_i}(x_*, u_*)$,

wobei die letzte Abschätzung aus der Existenz einer optimalen Lösung der primalen Aufgabe gemäß der Aussagen von Satz 8 und Satz 9 unter Gültigkeit der Bedingungen (V1)-(V4)²⁴ folgt. Gleichzeitig folgt aus Satz 17 die Gültigkeit der schwachen Dualitätsbeziehung²⁵, d.h. es gilt

$$\sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q) \leq \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = J_{\nu_i}(x_*, u_*)$$

und damit

$$G_{\nu_i}(y_*, q_*) \leq \sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q) \leq \inf_{(x,u) \in \mathcal{B}_P} J_{\nu_i}(x, u) = J_{\nu_i}(x_*, u_*) \leq G_{\nu_i}(y_*, q_*). \quad (4.32)$$

Offensichtlich ist die Gültigkeit der Relation (4.32) nur bei Gleichheit gegeben, so dass folgt

$$\sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q) = G_{\nu_i}(y_*, q_*).$$

Damit können wir schließen: Sind die Voraussetzungen (V1)-(V4) und (VD1)-(VD4) erfüllt und existiert (mindestens) ein Paar $(y, q) \in \tilde{Y}_{\nu_i}$, d.h. gilt $\tilde{Y}_{\nu_i} \neq \emptyset$, dann existiert eine optimale Lösung der dualen Aufgabe und die Zielfunktionalwerte der primalen und dualen Aufgabe stimmen überein. \square

Bemerkung 22. Die Herausforderung des Nachweises der Bedingung $\tilde{Y}_{\nu_i} \neq \emptyset$ liegt insbesondere in dem Nachweis der geforderten Raumzugehörigkeit der Lösung (y, q) der Differentialgleichung. Die Beweisführung des vorangegangenen Lemmas liefert zudem die Einsicht: Alle Elemente der Menge \tilde{Y}_{ν_i} sind optimale Lösungen der dualen Aufgabe (D_{ν_i}) .

4.7 Formulierung von Optimalitätsbedingungen für (P_{ν_i})

Ziel dieses Unterkapitels ist der Nachweis der folgenden Eigenschaft:

Die Lösungen der dualen Probleme (D_{ν_i}) definieren die (eindeutigen) Lösungen der primalen Aufgaben (P_{ν_i}) ($i = 0, 1$).

Dazu dient uns die – unter der Gültigkeit der in Satz 19 postulierten Voraussetzungen – nachgewiesene Beziehung von primaler und dualer Aufgabe der Form

$$\sup_{(y,q) \in \mathcal{B}_D} G_{\nu_i}(y, q) = J_{\nu_i}(x_*, u_*),$$

denn gemäß unserer vorherigen Ausführungen²⁶ liefert die obige Gleichung unmittelbar Rückschlüsse auf die hinreichenden Optimalitätsbedingungen an das Paar (x_*, u_*) der primalen Aufgabe.

In Kapitel 2 untersuchten wir die Problemklassen (P_{ν_i}) und konnten zeigen, dass für diese eindeutige Lösungen existieren. Zudem waren wir in der Lage, für die Problemklasse (P_{ν_0}) notwendige und – durch die strikte Konvexität des Integranden bzgl. der Argumente ξ und v – auch hinreichende Bedingungen zu formulieren. Daher ist die nochmalige Formulierung hinreichender Optimalitätsbedingungen für (P_{ν_0}) im Grunde überflüssig. Wir wollen diese dennoch als Folgerung in Form einer Sattelpunktbedingung angeben:

²⁴vgl. hierfür S. 34, S. 39 sowie S. 25

²⁵siehe Satz 17, S. 68

²⁶vgl. hierfür Bedingung (3.2) S. 46

Satz 21 (Hinreichende Bedingungen/ Sattelpunktbedingung). Seien (VD1)-(VD4) sowie (V_4) ²⁷ erfüllt. Die Menge \mathcal{U} sei in der Form $\mathcal{U} = [\underline{u}, \bar{u}] \subset \mathbb{R}^m$ gegeben. Ist (y_*, q_*) Lösung der dualen Aufgabe (D_{ν_0}), dann ist durch

$$x_*(t) = W^{-1}(t)q_*(t) \text{ und } u_*(t) = \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y_*(t) \right) \right)$$

die eindeutige optimale Lösung (x_*, u_*) von (P_{ν_0}) im Sinne des **Kriteriums L** gegeben.

Für das Problem (P_{ν_1}) konnten wir im Rahmen des **2.** Kapitels hingegen keine Optimalitätsbedingungen angeben. Dennoch ist auch hier die Existenz einer eindeutigen Lösung gesichert und die Betrachtungen der letzten zwei Unterkapitel ermöglichen uns die Schlussfolgerung des nachstehenden Satzes:

Satz 22 (Hinreichende Bedingungen/ Sattelpunktbedingung). Seien (VD1)-(VD4) sowie (V_4) erfüllt. Die Menge \mathcal{U} sei in der Form $\mathcal{U} = [\underline{u}, \bar{u}] \subset \mathbb{R}^m$ gegeben. Ist (y_*, q_*) Lösung der dualen Aufgaben (D_{ν_1}), dann ist durch

$$x_*(t) = W^{-1}(t)q_*(t)e^{-\beta_1 t} \text{ und } u_*(t) = \text{Proj}_{\mathcal{U}} \left(R^{-1}(t) \left(-\hat{r}(t) + B^T(t)y_*(t) \right) e^{-\beta_1 t} \right)$$

die eindeutige optimale Lösung (x_*, u_*) von (P_{ν_1}) im Sinne des **Kriteriums L** gegeben.

Beweis. Die Behauptungen der letzten beiden Sätze folgen unmittelbar aus Satz **18** und Satz **20**. \square

Die Ausführungen der letzten Unterkapitel erlauben uns nun mit den Setzungen

$$\hat{W}^{-1}(t) = \begin{cases} W^{-1}(t), & \text{für } \gamma = \beta_0 \\ W^{-1}(t)e^{\gamma t}, & \text{für } \gamma = -\beta_1 \end{cases} \text{ und } \hat{R}^{-1}(t) = \begin{cases} R^{-1}(t), & \text{für } \gamma = \beta_0 \\ R^{-1}(t)e^{\gamma t}, & \text{für } \gamma = -\beta_1 \end{cases} \quad \forall t \in \mathbb{R}_{++}$$

zudem die folgende Formulierung notwendiger und hinreichender Bedingungen an die primale Aufgabe (P_ν) :

Satz 23. Seien (VD1)-(VD4) sowie (V_4) erfüllt, $\tilde{Y}_{\nu_i} \neq \emptyset$ (für $i = 0, 1$) und die Menge \mathcal{U} sei in der Form $\mathcal{U} = [\underline{u}, \bar{u}] \subset \mathbb{R}^m$ gegeben. Das Paar $(x_*, u_*) \in \mathcal{B}_P$ ist optimale Lösung des Problems (P_ν) im Sinne des **Kriteriums L** genau dann, wenn gilt: Es existieren $(y_*, q_*, \lambda_0, p_0)$ mit

$$x_*(t) = \hat{W}^{-1}(t)q_*(t) \text{ und } u_*(t) = \text{Proj}_{\mathcal{U}} \left(\hat{R}^{-1}(t) \left(-\hat{r}(t) + B^T(t)y_*(t) \right) \right),$$

wobei

$$\lambda_0 = 1, \tag{N}$$

$$p_0 \in W_2^{1,n}(\mathbb{R}_{++}, e^{-\gamma t}), \quad \gamma < 0 \tag{T}$$

$$H(t, y_*(t), q_*(t), p_0(t), \lambda_0) = \max_{v_D \in \mathbb{R}^n} H(t, y_*(t), v_D, p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{M}$$

$$p_0'(t) = -\nabla_{\xi_D}^T H(t, y_*(t), q_*(t), p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++}, \tag{K}$$

$$p_0(0) = x_0,$$

wobei $H : \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ die Pontryaginsche Funktion

$$H(t, \xi_D, v_D, \eta_D, \lambda_0) = -\lambda_0 g(t, \xi_D, v_D) e^{\gamma t} + \eta_D^T (-\hat{A}^T(t)\xi_D + v_D) \quad \text{mit}$$

$$g(t, \xi_D, v_D) = \frac{1}{2} v_D W^{-1}(t) v_D - \hat{\sigma}^T(t, \xi_D) \left[\frac{1}{2} R(t) \hat{\sigma}(t, \xi_D) + \hat{r}(t) - B^T(t) \xi_D \right] \quad \text{ist.}$$

Beweis. Der Beweis folgt unmittelbar aus Satz **20**, Satz **21** und Satz **22**. \square

²⁷Die Gültigkeit der Bedingungen (VD1)-(VD4) implizieren die Gültigkeit der Bedingungen (V1)-(V3), siehe S. **25** und S. **76**, daher genügt die Forderung der Gültigkeit der angegebenen Voraussetzungen.

Kapitel 5

Numerische Behandlung

Für die numerische Behandlung linear-quadratischer Optimalsteuerungsprobleme mit unendlichem Zeithorizont in gewichteten Sobolev-Räumen mit Dichtefunktion wird in [Bur19] eine indirekte Pseudospektralmethode vorgeschlagen. Die vorgestellte Methode erlaubt eine numerische Behandlung des Optimalsteuerungsproblems ohne Abschneiden des unendlichen Zeithorizontes und liefert Konvergenzaussagen im Sinne der $L_2^n(\mathbb{R}_{++}, e^{\gamma t})$ -Norm. Da die vorliegende duale Aufgabe ebenfalls in die Klasse der in [Bur19] diskutierten Probleme fällt, wollen wir uns dieser Methodik annehmen.

Dafür wollen wir in diesem Kapitel einen kurzen Einblick in diese Pseudospektralmethode geben und die Grundideen des Verfahrens für die numerische Behandlung der dualen Aufgabe (D_ν) übernehmen. Für eine ausführlichere Betrachtung dieser Methodik verweisen wir auf [Bur19]. Das nachfolgende Unterkapitel ist daher als Extrakt der Dissertation [Bur19] zu verstehen. Nach der Vorstellung der grundlegenden Ideen der in [Bur19] behandelten Pseudospektralmethode werden wir zudem die bereits in [Bur19] und [PB19] angeführte Modifizierung des Verfahrens vorstellen, welche weitreichendere Konvergenzaussagen ermöglicht.

5.1 Vorstellung der Pseudospektralmethode

Die Ausführungen in [Bur19] motivieren die behandelte Pseudospektralmethode ausgehend von einer allgemeinen Differentialgleichung über dem Gebiet Ω der Form

$$\mathcal{L}\{x(t)\} = f(t), t \in \Omega, \quad (5.1)$$

wobei \mathcal{L} einen beliebigen Differentialoperator bezeichne. Die Approximation der gesuchten Lösung x geschieht dabei über einen Ansatz der Form

$$x(t) \approx x_N(t) = \sum_{k=1}^N a_k \phi_k(t), \quad (5.2)$$

mit den linear unabhängigen Ansatzfunktionen ϕ_k und den noch zu bestimmenden, unbekanntem Koeffizienten a_k . Ziel dieses Ansatzes ist eine möglichst gute Näherungslösung x_N . Dabei werden in [Bur19] *gewichtete Residuen* für die Bewertung der Güte der Approximation vorgeschlagen. Anstatt einer punkweisen Betrachtung des lokalen Fehlers der approximierten Lösung x_N der Form

$$\mathbf{R}_N(t) := \mathcal{L}\{x_N(t)\} - f(t) \neq 0, t \in \Omega,$$

geschieht die Bewertung der Approximation daher mit Hilfe der Testfunktionen ψ_j über das gewichtete Funktionalprodukt $\langle \cdot, \cdot \rangle_\Omega$ in der Setzung

$$\langle \mathbf{R}_N, \psi_j \rangle_\Omega := \int_{\Omega} \mathbf{R}_N \psi_j d\Omega \stackrel{!}{=} 0, \quad 0 \leq j \leq N.$$

Für die Wahl des Ansatzes (5.2) wird eine Kollokations- (im Sinne von Funaro [Fun92] auch Pseudospektral-) methode unter Verwendung des Galerkin-Verfahrens verwendet. Hinter diesen Begriffen verbergen sich folgende Setzungen für den Ansatz in (5.2):

Im Sinne des Galerkin-Verfahrens sind Ansatz und Testfunktionen identisch zu wählen, d.h. es gilt $\phi_k = \psi_j$. Weiterhin wird gefordert, dass die Näherungslösung x_N die Differentialgleichung (5.1) in vorgegebenen Kollokationspunkten $\{\tau_j\}$, $\tau_j \in \Omega$ exakt erfüllt und somit auch der lokale Fehler in diesen Punkten verschwindet. Die Wahl der Ansatz- und Testfunktionen ϕ_k sowie die Bestimmung der Kollokationspunkte basieren auf bekannten Ergebnissen der Fourier-Analyse und der Approximationstheorie in separablen Hilberträumen¹ und wird in [Bur19] in folgender Weise motiviert:

Wir nutzen die Struktur der in der Aufgabenklasse (P_ν) zugrundeliegenden Funktionenräume in Form von gewichteten Lebesgue-Räumen mit echter Dichtefunktion $\nu(t) = e^{-\rho t}$, $\rho > 0$ über dem Gebiet $\Omega = \mathbb{R}_{++}$. Hier bilden die *verallgemeinerten Laguerre-Polynome*

$$L_k^{(\omega, \rho)}(t) = \frac{1}{k!} t^{-\omega} e^{\rho t} \frac{d^k}{dt^k} \left(t^{k+\omega} e^{-\rho t} \right), \quad k = 0, 1, \dots$$

für allgemeine Gewichtsfunktionen

$$\nu_{\omega, \rho}(t) = t^\omega e^{-\rho t}, \quad \omega > -1, \rho > 0, t \in \mathbb{R}_{++} \quad (5.3)$$

ein vollständiges Orthogonalsystem², so dass sich jede Funktion $f \in L_2(\mathbb{R}_{++}, \nu_{\omega, \rho})$ durch die Fourier-Laguerresche Reihe

$$f = \sum_{k=0}^{\infty} \hat{f}_k L_k^{(\omega, \rho)} \quad (5.4)$$

mit den Fourierkoeffizienten

$$\hat{f}_k := \frac{\langle f, L_k^{(\omega, \rho)} \rangle_{(2), \nu_{\omega, \rho}}}{\|L_k^{(\omega, \rho)}\|_{(2), \nu_{\omega, \rho}}^2}, \quad k \geq 0 \quad (5.5)$$

darstellen lässt, wobei unter der Konvergenz der Reihe (5.4) die Normkonvergenz des $L_2(\mathbb{R}_{++}, \nu_{\omega, \rho})$ zu verstehen ist³.

Die für unsere Behandlung wesentlichen Eigenschaften der Laguerre-Polynome können den nachfolgenden Lemmata entnommen werden⁴.

¹Für eine umfassendere Einführung in die Approximationstheorie in Hilberträumen siehe beispielsweise [Dav75, Wer07, Isk18, FH07]

²vgl. [Wer07], S. 245, Satz V.3.4, [Tri70] S. 235 ff., [Bur19], S. 121, Lemma 5.1

³vgl. zum Beispiel [Bur19] S. 55 Satz 3.13 und S. 125

⁴Eine ausführliche Darlegung der Eigenschaften der verallgemeinerten Laguerre-Polynome für den Fall $\rho = 1$ findet sich beispielsweise in [Tri70, SSG39]. Eine Darlegung für den Fall von $\omega = 0$ und $\rho > 0$ kann [XyHf13] entnommen werden.

Lemma 40 (vgl. [Tri70], S. 214 und S. 235). Die verallgemeinerten Laguerre-Polynome $L_k^{(\omega, \rho)}$ bilden ein vollständiges Funktionensystem im $L_2(\mathbb{R}_{++}, \nu_{\omega, \rho})$ bezüglich der Gewichtsfunktion aus (5.3). Es gilt die Orthogonalitätsrelation

$$\left\langle L_k^{(\omega, \rho)}, L_m^{(\omega, \rho)} \right\rangle_{(2), \nu_{\omega, \rho}} := \int_0^\infty L_k^{(\omega, \rho)}(t) L_m^{(\omega, \rho)}(t) \nu_{\omega, \rho}(t) dt = \begin{cases} \gamma_k^{(\omega, \rho)}, & k = m \\ 0 & k \neq m \end{cases}$$

mit

$$\gamma_k^{(\omega, \rho)} = \frac{\Gamma(k + \omega + 1)}{\rho^{\omega+1} k!},$$

wobei $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ mit $z \in \mathbb{R}$, $z > 0$ die Eulersche Gammafunktion meint⁵.

Beweis. In der angegebenen Quelle werden die Eigenschaften für den Fall $\rho = 1$, d.h. für Gewichtsfunktionen $\nu_\omega(\tau) = \tau^\omega e^{-\tau}$, nachgewiesen. Die Substitution $t := h(\tau) = \frac{\tau}{\rho}$ liefert die Verallgemeinerung der Ergebnisse auf Gewichtsfunktionen der Form $\nu_{\omega, \rho}(t) = t^\omega e^{-\rho t}$ mit $\omega > -1$ und $\rho > 0$. Eine ausführliche Darlegung dieser Beweisidee findet sich beispielsweise in [Bur19], Lemma 5.1, S. 121. \square

Lemma 41 (vgl. [SSG39], S. 101 f., [Tri70], S. 215). Für die Ableitung der verallgemeinerten Laguerre-Polynome $L_k^{(\omega, \rho)}$ gelten die Gleichungen

$$\frac{d}{dt} L_k^{(\omega, \rho)}(t) = -\rho L_{k-1}^{(\omega+1, \rho)}(t) = -\rho \sum_{l=0}^{k-1} L_l^{(\omega, \rho)}(t), \quad k \geq 1$$

$$L_k^{(\omega, \rho)}(t) = \frac{1}{\rho} \left[\frac{d}{dt} L_k^{(\omega, \rho)}(t) - \frac{d}{dt} L_{k+1}^{(\omega, \rho)}(t) \right], \quad k \geq 1$$

und es ist

$$L_k^{(\omega, \rho)}(0) = \frac{\Gamma(k + \omega + 1)}{\Gamma(\omega + 1) \Gamma(k + 1)}. \quad (5.6)$$

Beweis. In der angegebenen Quelle werden die Eigenschaften für den Fall $\rho = 1$ nachgewiesen. Die Substitution $t := h(\tau) = \frac{\tau}{\rho}$ liefert wieder die Verallgemeinerung der Ergebnisse. Eine ausführliche Darlegung dieser Beweisidee findet sich beispielsweise in [Bur19], Lemma 5.2, S. 123. \square

Lemma 42 ([Bur19], Lemma 5.3, S. 124). Es sei $l \in \mathbb{N}$. Die l -ten Ableitungen der verallgemeinerten Laguerre-Polynome $\frac{d^l}{dt^l} L_k^{(\omega, \rho)}$, $k \geq l$ bilden ein vollständiges Funktionensystem im $L_2(\mathbb{R}_{++}, \nu_{\omega+l, \rho})$ bezüglich der Gewichtsfunktion $\nu_{\omega+l, \rho} = t^{\omega+l} e^{-\rho t}$, $\omega > -1$, $\rho > 0$.

Auf Grundlage dieser Eigenschaften wählen wir nun die Setzung

$$x_N(t) := \sum_{k=0}^N a_k L_k^{(\omega, \rho)}(t), \quad \text{mit} \quad a_k := \frac{\left\langle x, L_k^{(\omega, \rho)} \right\rangle_{(2), \nu_{\omega, \rho}}}{\left\| L_k^{(\omega, \rho)} \right\|_{(2), \nu_{\omega, \rho}}^2},$$

⁵Einen Überblick über die grundlegenden Eigenschaften der Gammafunktion findet man in [Fun92], S. 2ff., [Tri70], S. 143 ff.

für die Darstellung der Näherungslösung x_N . Im Sinne der *Gauß-Laguerre-Quadratur* werden die Kollokationspunkte $\tau_i^{(N)}$, $i = 1, \dots, N$ durch die Nullstellen des N -ten Laguerre-Polynoms $L_N^{(\omega, \rho)}$ vorgeschlagen⁶. Die Hinzunahme des weiteren, zusätzlichen Kollokationspunktes $\tau_0^{(N)} = 0$ führt letztlich für alle $i = 0, \dots, N$ zu den Ansätzen

$$x_N(\tau_i^{(N)}) = \sum_{k=0}^N a_k L_k^{(\omega, \rho)}(\tau_i^{(N)}).$$

Die Aussagen der vorangegangenen Lemmata erlauben zudem eine Darstellung der Ableitungen der Näherungsfunktionen x'_N in der Form⁷

$$x'_N(\tau_i^{(N)}) = \sum_{k=0}^N a_k \frac{d}{dt} L_k^{(\omega, \rho)}(\tau_i^{(N)}),$$

mit den gleichen Fourierkoeffizienten a_k . Mit dieser Setzung wird in [Bur19] schließlich die numerische Behandlung von Optimalsteuerungsproblemen der Form (P_ν) in indirekter Form motiviert und die Normkonvergenz des Verfahrens nachgewiesen. In [PB19] findet sich zudem eine Erweiterung dieser Ideen mit dem Ergebnis stärkerer Konvergenzaussagen. Dafür wird eine Modifikation der Methode vorgestellt. Ausgangspunkt dieser Modifikation ist die Frage, unter welchen Voraussetzungen eine Funktion f auf \mathbb{R}_{++} auch punktweise durch ihre formale Fourier-Laguerre-Reihe approximierbar ist, d.h wann Gleichung (5.4) auch in der Form

$$f(t) = \sum_{k=0}^{\infty} \hat{f}_k L_k^{(\omega, \rho)}(t) \quad (\forall) t \in \mathbb{R}_{++}$$

zulässig ist. Die Ergebnisse in [PB19] basieren dabei wieder auf den Ideen aus [Tri70] und wurden dort für Gewichte der Form $\nu_{\omega, \rho}(t) = t^\omega e^{-\rho t}$ mit $\omega > -1$ und $\rho > 0$ verallgemeinert. Die derartigen Verallgemeinerungen liefern die folgenden hinreichenden Bedingungen für die gleichmäßige Konvergenz:

Satz 24 ([PB19], Theorem 4). *Jede Funktion Φ , welche sich mit einer Funktion $f \in L_2(\mathbb{R}_{++}, \nu_{\omega, \rho})$ für alle $t > 0$ in der Form*

$$\Phi(t) = \rho e^{\rho t} \int_t^{\infty} f(\tau) e^{-\rho \tau} d\tau$$

darstellen lässt, kann in jedem endlichen Intervall $(0, b)$, $b > 0$ in die absolut gleichmäßig konvergente Fourier-Reihe

$$\Phi(t) = \sum_{k=0}^{\infty} \hat{\Phi}_k L_k^{(\omega, \rho)}(t) \tag{5.7}$$

mit den Koeffizienten

$$\hat{\Phi}_k := \hat{f}_k - \hat{f}_{k+1}, \quad k = 0, 1, \dots$$

entwickelt werden. Dabei sind \hat{f}_k die Fourier-Koeffizienten von f bzgl. der verallgemeinerten Laguerre-Polynome $L_k^{(\omega, \rho)}$, $k = 0, 1, \dots$ (vgl. (5.5)).

⁶Die Nullstellen orthogonaler Polynome auf dem Intervall (a, b) sind alle reell, einfach und im Inneren des Definitionsbereiches gelegen. Vgl. hierzu [FH07], S. 175, (3.6.10 Satz).

⁷vgl. [Bur19], Lemma 5.4, S. 127 ff.

Beweis. Für den Beweis verweisen wir auf [PB19], S. 18. \square

Bemerkung 23. Die Multiplikation der Reihe (5.7) mit dem Faktor $e^{-\rho t}$ liefert die gleichmäßige Konvergenz der Reihe auf ganz \mathbb{R}_{++} .

Da wir für die Behandlung der Problemklasse (D_ν) Funktionen der Räume $W_2^{2,n}(\mathbb{R}_{++}, \nu)$ behandeln, sind wir auch an den distributionellen Ableitungen der zu approximierenden Funktionen interessiert. Für diese Behandlung definieren wir die Menge von Gewichtsfunktionen

$$S_2 := \{\nu_{0,\rho}, \nu_{1,\rho}, \nu_{1,\rho}\} = \{e^{-\rho t}, te^{-\rho t}, te^{-\rho t}\}$$

und untersuchen die zu approximierenden Funktionen auf dem Raum $W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho}) \cap W_2^{2,n}(\mathbb{R}_{++}, S_2)$ mit

$$W_2^{2,n}(\mathbb{R}_{++}, S_2) = \left\{ f \in L_2^n(\mathbb{R}_{++}, \nu_{0,\rho}) \mid \mathcal{D}^1 f, \mathcal{D}^2 f \in L_2^n(\mathbb{R}_{++}, \nu_{1,\rho}) \right\}.$$

Gemäß der Betrachtungen in Kapitel 1.3 sichert auch diese Raumwahl die Existenz eines stetig differenzierbaren Repräsentanten, so dass wir auch hier die schwache Ableitung $\mathcal{D}^1 f$ der Sobolev-Funktion $f \in W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho}) \cap W_2^{2,n}(\mathbb{R}_{++}, S_2)$ mit der klassischen Ableitung ihres Repräsentanten identifizieren können. Dies ermöglicht zudem unter der Anwendung des Satzes von Rellich-Kondrachov⁸ die sinnvolle Definition der Anfangswerte $f(0)$ und $f'(0)$. Mit diesem Wissen können wir nun den folgenden Satz zitieren:

Satz 25 ([PB19], Theorem 5). *Es sei $x \in W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho}) \cap W_2^{2,n}(\mathbb{R}_{++}, S_2)$ mit $x(0) = x'(0) = 0$. Dann lässt sich x' in jedem Intervall $(0, b)$, $b > 0$ in die absolut gleichmäßig konvergente Fourier-Laguerre-Reihe*

$$x'(t) = \sum_{k=0}^{\infty} (\hat{f}_k - \hat{f}_{k+1}) L_k^{(1,\rho)}(t) \quad (5.8)$$

$$= \sum_{k=0}^{\infty} \hat{f}_k [L_k^{(1,\rho)}(t) - L_{k-1}^{(1,\rho)}(t)] \quad (5.9)$$

entwickeln, wobei \hat{f}_k die Fourier-Koeffizienten von $f := x' - \frac{1}{\rho} \mathcal{D}^1(x') \in L_2^n(\mathbb{R}_{++}, \nu_{1,\rho})$ bzgl. der verallgemeinerten Laguerre-Polynome $L_k^{(1,\rho)}$, $k = 0, 1, \dots$ sind und $L_{-1}^{(1,\rho)}(t) := 0$ für alle $t \in \mathbb{R}_{++}$.

Beweis. Für den Beweis verweisen wir wieder auf [PB19], S. 19. \square

Damit erhalten wir nun Aussagen hinsichtlich der gleichmäßigen Konvergenz der Fourier-Laguerre-Reihen (5.8) und (5.9) für alle Intervalle $(0, b)$, $b > 0$. Den noch fehlenden Schluss über Konvergenzaussagen dieser Reihe in $t = 0$ liefert das nachfolgende Lemma.

Lemma 43 ([PB19], Lemma 6). *Unter den Voraussetzungen des Satzes 25 konvergiert die Fourier-Laguerre-Reihe in (5.9) gleichmäßig in jedem Intervall $[0, b)$, $b > 0$ mit*

$$\hat{f}_0 = - \sum_{k=1}^{\infty} \hat{f}_k$$

und es gilt

$$x(t) = t \left[\hat{f}_0 + \sum_{k=1}^{\infty} \frac{\hat{f}_k}{k+1} L_k^{(1,\rho)}(t) \right].$$

Beweis. Für den Beweis verweisen wir auf [PB19], S. 20 ff. \square

⁸vgl. Satz 4, S. 11 sowie Satz 5, S. 12

5.2 Anwendung einer direkten Pseudospektralmethode

Zur Behandlung der direkten Pseudospektralmethode ziehen wir die Aufgabe (D_ν) in der Formulierung eines Variationsproblems heran, d.h. wir betrachten mit $\rho = -\gamma$

$$(\mathbf{D}_\nu) : \quad G^\infty(y) := \int_0^\infty g(t, y(t), y'(t)) e^{-\rho t} dt + x_0^T y(0) \longrightarrow \inf!$$

$$\text{u.d.N. } y \in Y := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\rho t}) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\},$$

mit

$$g(t, y(t), y'(t)) = \frac{1}{2} \left[y'(t) + \hat{A}^T(t)y(t) \right]^T W^{-1}(t) \left[y'(t) + \hat{A}^T(t)y(t) \right] + \theta(t, y(t)),$$

wobei

$$\theta(t, y(t)) = -\hat{\sigma}^T(t, y(t)) \left[\frac{1}{2} R(t) \hat{\sigma}(t, y(t)) + \hat{r}(t) - B^T(t)y(t) \right].$$

Im Hinblick auf eine gleichmäßige Konvergenz der Fourier-Reihen motiviert das vorangegangene Unterkapitel die Approximation der gesuchten Funktionen mit Hilfe von Ansatzfunktionen in Form verallgemeinerter Laguerre-Polynome für Funktionen y über der Teilmenge $Y_S \subset Y$ mit

$$Y_S := \left\{ y \in W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho}) \cap W_2^{2,n}(\mathbb{R}_{++}, S_2) \mid x_0 = W^{-1}(0) \left(y'(0) + \hat{A}^T(0)y(0) \right) \right\}.$$

Die Existenz einer optimalen Lösung y_* der dualen Aufgabe über den Raum $W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho})$ ist unter Gültigkeit der Voraussetzungen aus Satz 20 gesichert⁹. Mit der Wahl des Raumes Y_S nehmen wir allerdings in Kauf optimale Lösungen für (D_ν) – welche lediglich im Raum $W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho})$, aber nicht in $W_2^{2,n}(\mathbb{R}_{++}, S_2)$ liegen – nicht zu berücksichtigen, gewinnen dafür jedoch die stärkeren Konvergenzeigenschaften des Verfahrens.

Finden wir aber eine optimale Lösung in der Menge Y_S , dann ist diese eine (eventuell auch scharfe) untere Schranke für (D_ν) und liefert Rückschlüsse auf die eindeutige optimale Lösung für (P_ν) ¹⁰, denn aufgrund der nachgewiesenen starken Dualität gilt zusammen mit der Bedingung (3.15)¹¹

$$\inf_{(x,u) \in \mathcal{B}_P} J(x, u) = \sup_{y \in Y} G(y) \geq \sup_{y \in Y_S} G(y).$$

Bemerkung 24. Liefert das hier vorgestellte Verfahren keine Lösung, schlagen wir eine Verwendung der weniger restriktiven Methode mittels eines Ansatzes der Form

$$y_N(t) := \sum_{j=0}^N \mathbf{a}_j L_j^{(0,\rho)}(t) \quad \text{und} \quad y'_N(t) := \sum_{j=0}^N \mathbf{a}_j \frac{d}{dt} L_j^{(0,\rho)}(t),$$

mit den verallgemeinerten Laguerre-Polynomen $L_j^{(0,\rho)}$ und den noch zu bestimmenden Koeffizientenvektoren $\mathbf{a}_j = (a_j^0, \dots, a_j^n)^T \in \mathbb{R}^n$ mit $j \in \{1, \dots, N\}$ vor.

⁹vgl. S. 101

¹⁰vgl. Satz 21, S. 104 und Satz 22, S. 104

¹¹siehe S. 51

Erlauben die Voraussetzungen an die Problemstellung (P_ν) bzw. (D_ν) die Wahl eines hinreichend großen Intervalls für die Wahl des Gewichtungsparmeters ρ , so liefert das nachfolgende Lemma Aussagen über die Zugehörigkeit der Funktion y zur Menge Y_S :

Lemma 44. *Gilt $y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\hat{\rho}t})$ für alle $\hat{\rho} \in (\rho - \varepsilon, \rho + \varepsilon)$ für ein beliebiges $\varepsilon > 0$, dann ist $y \in W_2^{2,n}(\mathbb{R}_{++}, S_2)$.*

Beweis. Die Reihendarstellung der Funktion $e^{\varepsilon t}$ liefert die folgende Abschätzung

$$e^{\varepsilon t} = \sum_{k=0}^{\infty} \frac{(\varepsilon t)^k}{k!} > 1 + \varepsilon t.$$

Damit können wir für eine beliebige Funktion $\hat{y} \in L_2^n(\mathbb{R}_{++}, \nu_{0,\rho})$ leicht abschätzen

$$\int_0^{\infty} |\hat{y}(t)|^2 e^{-\rho t} dt = \int_0^{\infty} |\hat{y}(t)|^2 e^{(-\rho-\varepsilon)t} e^{\varepsilon t} dt \geq \int_0^{\infty} |\hat{y}(t)|^2 e^{-\hat{\rho}t} dt + \varepsilon \int_0^{\infty} |\hat{y}(t)|^2 t e^{-\hat{\rho}t} dt,$$

mit $\hat{\rho} = \rho + \varepsilon$. Ist $\hat{y} \in L_2^n(\mathbb{R}_{++}, \nu_{0,\rho})$, dann impliziert obige Abschätzung $\hat{y} \in L_2^n(\mathbb{R}_{++}, \nu_{1,\hat{\rho}})$. Da $y \in W_2^{2,n}(\mathbb{R}_{++}, e^{-\hat{\rho}t})$, gilt für ihre verallgemeinerten Ableitungen damit auch $\mathcal{D}^1 y, \mathcal{D}^2 y \in L_2^n(\mathbb{R}_{++}, t e^{-\rho t})$. Insgesamt ergibt sich damit

$$y \in \left\{ y \in L_2^n(\mathbb{R}_{++}, \nu_{0,\rho}) \mid \mathcal{D}^1 y, \mathcal{D}^2 y \in L_2^n(\mathbb{R}_{++}, t e^{-\rho t}) \right\} = W_2^{2,n}(\mathbb{R}_{++}, S_2).$$

□

Die in Unterkapitel [5.1](#) zitierten Lemmata liefern jedoch nur Aussagen für Funktionen $y \in Y_S$ mit $y(0) = y'(0) = 0$. Da dies im allgemeinen Fall nicht zu erwarten ist, folgen wir der Idee aus [PB19](#) und setzen

$$y(t) = \hat{y}(t) + E \cdot t + D, \quad \hat{y}(0) = \hat{y}'(0) = 0, \quad D, E \in \mathbb{R}^n.$$

Offensichtlich sind die so eingeführten Funktionen \hat{y} ebenfalls Elemente der Schnittmenge $W_2^{2,n}(\mathbb{R}_{++}, \nu_{0,\rho}) \cap W_2^{2,n}(\mathbb{R}_{++}, S_2)$, so dass für diese nun die Aussagen der vorangegangenen Lemma [40](#) bis Lemma [43](#) sowie von Satz [25](#) gelten.

Damit finden wir eine Darstellung der Funktionen \hat{y}, \hat{y}' durch

$$\begin{aligned} \hat{y}(t) &= \psi_1(F)(t) := t \left[\hat{f}_0 + \sum_{k=1}^{\infty} \frac{\hat{f}_k}{k+1} L_k^{(1,\rho)}(t) \right], \\ \hat{y}'(t) &= \psi_2(F)(t) := \sum_{k=0}^{\infty} \hat{f}_k \left(L_k^{(1,\rho)}(t) - L_{k-1}^{(1,\rho)}(t) \right), \end{aligned}$$

$$\text{mit } \hat{f}_0 = - \sum_{k=1}^{\infty} \hat{f}_k \quad \text{und} \quad F := \{ \hat{f}_k \}_{k \in \mathbb{N}_0}, \quad \hat{f}_k \in \mathbb{R}^n.$$

Diese Darstellungen lassen sich unter Ausnutzung der Definition $L_{-1}^{(1,\rho)}(t) := 0$ für alle $t \in \mathbb{R}_+$ auch komprimierter formulieren durch

$$\begin{aligned} \hat{y}(t) &= \psi_1(F)(t) := t \sum_{k=1}^{\infty} \hat{f}_k \left[-1 + \frac{1}{k+1} L_k^{(1,\rho)}(t) \right], \\ \hat{y}'(t) &= \psi_2(F)(t) := \sum_{k=1}^{\infty} \hat{f}_k \left[-1 + L_k^{(1,\rho)}(t) - L_{k-1}^{(1,\rho)}(t) \right], \end{aligned}$$

$$\text{mit } \hat{f}_0 = -\sum_{k=1}^{\infty} \hat{f}_k \quad \text{und} \quad F := \{\hat{f}_k\}_{k \in \mathbb{N}_0}, \quad \hat{f}_k \in \mathbb{R}^n.$$

Unter Ausnutzung der Relation (5.6) erhalten wir zudem den Zusammenhang

$$\psi_1(F)(0) = 0 \quad \text{und} \quad \psi_2(F)(0) = 0.$$

Mit diesen Setzungen können wir die duale Aufgabe (D_ν) über der Menge Y_S nun darstellen als

$$G(D, E, F) := \int_0^{\infty} \tilde{g}(t, D, E, F) e^{-\rho t} dt + x_0^T D \longrightarrow \text{inf!}$$

$$\text{u.d.N.} \quad x_0 = W^{-1}(0) \left(E + \hat{A}^T(0)D \right), \quad \rho > 0, \\ (D, E, F) \in \mathbb{R}^n \times \mathbb{R}^n \times l_2^n.$$

5.2.1 Diskretisierung der Aufgabe

Wir diskretisieren die Aufgabe (D_ν) nun mittels einer polynomialen Approximation des zulässigen Bereiches. Gemäß der Erkenntnisse der vorangegangenen Unterkapitel geschieht dies über die Darstellung der Funktionen \hat{y} , \hat{y}' durch ihre zugehörigen Partialsummen, das heißt durch die Wahl der Ansatzfunktionen, gemäß

$$\hat{y}_N(t) = \psi_1(F^N)(t) := t \sum_{k=1}^N \hat{f}_k \left[-1 + \frac{1}{k+1} L_k^{(1,\rho)}(t) \right], \\ \hat{y}'_N(t) = \psi_2(F^N)(t) := \sum_{k=1}^N \hat{f}_k \left[-1 + L_k^{(1,\rho)}(t) - L_{k-1}^{(1,\rho)}(t) \right], \\ \text{mit } \hat{f}_0 = -\sum_{k=1}^N \hat{f}_k \quad \text{und} \quad F^N := \{\hat{f}_k\}_{k=0}^N, \quad \hat{f}_k \in \mathbb{R}^n.$$

Durch diese Setzung erhalten wir eine Folge von endlich dimensionalen Problemen $\{(D^N)\}_{N \in \mathbb{N}_0}$ mit

$$(\mathbf{D}^N) : \quad G^N(D, E, F^N) := \int_0^{\infty} \tilde{g}_N(t, D, E, F^N) e^{-\rho t} dt + x_0^T D \longrightarrow \text{inf!}$$

$$\text{u.d.N.} \quad x_0 = W^{-1}(0) \left[E + \hat{A}^T(0)D \right], \quad \rho > 0,$$

$$(D, E, F^N) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times N}.$$

Die Anwendung der Karush-Kuhn-Tucker Bedingungen¹² erlaubt uns die Überführung des restringierten Optimierungsproblem durch die Verwendung der Lagrange-Funktion

$$\mathcal{L}^N(D, E, F^N, \lambda) := G^N(D, E, F^N) + \lambda^T \left(W^{-1}(0) \left[E + \hat{A}^T(0)D \right] - x_0 \right)$$

in das unrestringierte Problem

$$\mathcal{L}^N(D, E, F^N, \lambda) \longrightarrow \text{inf!}$$

¹²vgl. hierfür beispielsweise [GK02], Seite 55, Satz 2.42

$$(D, E, F^N, \lambda) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times N} \times \mathbb{R}^n.$$

Damit gilt für die optimale Lösung $D_*, E_*, F_*^N, \lambda_* \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times N} \times \mathbb{R}^n$ notwendig

$$\nabla_{(D, E, F^N, \lambda)} \mathcal{L}^N(D_*, E_*, F_*^N, \lambda_*) = 0. \quad (5.10)$$

Mit den Setzungen

$$S_k(t) := t \cdot \left(-1 + \frac{1}{k+1} L_k^{(1, \rho)}(t) \right) \quad \text{und} \quad T_k(t) := \left(-1 + L_k^{(1, \rho)}(t) - L_{k-1}^{(1, \rho)}(t) \right) \quad (5.11)$$

für alle $t \in \mathbb{R}_+$ und alle $k = 1, \dots, N$ stellen die Lösungen $(D_*, E_*, F_*^N, \lambda_*) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times N} \times \mathbb{R}^n$ des Gleichungssystems [\(5.10\)](#) den Zusammenhang

$$\begin{aligned} y_{N_*}(t) &= \sum_{k=1}^N \hat{f}_{k_*} S_k(t) + E_* \cdot t + D_*, \\ y'_{N_*} &= \sum_{k=1}^N \hat{f}_{k_*} T_k(t) + E_*, \\ \text{mit } \hat{f}_{0_*} &= - \sum_{k=1}^N \hat{f}_{k_*} \quad \text{her.} \end{aligned}$$

5.2.2 Kollokationsgleichungen

Für die Bestimmung der Kollokationsgleichungen werten wir zunächst die notwendigen Bedingungen [\(5.10\)](#) aus. Unter Nutzung der Definition

$$\sigma_N(t, D, E, F^N) := \min \left\{ \max \left\{ \hat{u}(t), R^{-1}(t) \left[-\hat{r}(t) + \hat{B}^T(t) \left(\sum_{k=1}^N \hat{f}_k S_k(t) + tE + D \right) \right] \right\}, \bar{u}(t) \right\}$$

für alle $(t, D, E, F^N) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times N}$ ergeben sich diese zu

$$\int_0^\infty \left\{ \left[(I_{n \times n})_{(\vartheta)} T_j(t) + \hat{A}_{(\vartheta)}^T(t) S_j(t) \right]^T W^{-1}(t) \left[\sum_{k=1}^N \hat{f}_k T_k(t) + E + \hat{A}^T(t) \left(\sum_{k=1}^N \hat{f}_k S_k(t) + tE + D \right) \right] - \left[\sigma_N^T(t, D, E, F^N) \hat{B}_{(\vartheta)}^T(t) \right] S_j(t) \right\} e^{-\rho t} dt \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n \quad \forall j = 1, \dots, N$$

$$\int_0^\infty \left\{ \left[(I_{n \times n})_{(\vartheta)} + \hat{A}_{(\vartheta)}^T(t) \cdot t \right]^T W^{-1}(t) \left[\sum_{k=1}^N \hat{f}_k T_k(t) + E + \hat{A}^T(t) \left(\sum_{k=1}^N \hat{f}_k S_k(t) + tE + D \right) \right] - \left[\sigma_N^T(t, D, E, F^N) \hat{B}_{(\vartheta)}^T(t) \right] \cdot t \right\} e^{-\rho t} dt + \lambda^T W_{(\vartheta)}^{-1}(0) \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n$$

$$\int_0^\infty \left\{ \left[\hat{A}_{(\vartheta)}^T(t) \right]^T W^{-1}(t) \left[\sum_{k=1}^N \hat{f}_k T_k(t) + E + \hat{A}^T(t) \left(\sum_{k=1}^N \hat{f}_k S_k(t) + tE + D \right) \right] - \sigma_N^T(t, D, E, F^N) \hat{B}_{(\vartheta)}^T(t) \right\} e^{-\rho t} dt + x_{0(\vartheta)} + \lambda^T \left[W^{-1}(0) \hat{A}^T(0) \right]_{(\vartheta)} \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n$$

$$W^{-1}(0) \left(E + \hat{A}^T(0) D \right) - x_0 \stackrel{!}{=} 0.$$

Dabei bezeichnen wir mit $A_{(\vartheta)}$ die ϑ -te Spalte der Matrix A . Es ergibt sich somit ein nichtlineares Gleichungssystem mit $N \cdot n + 3 \cdot n$ Unbekannten und $N \cdot n + 3 \cdot n$ Gleichungen.

Für die Näherungslösungen fordern wir, dass sie die obigen Gleichungen an fest vorgegebenen Kollokationspunkten exakt erfüllen. In Anlehnung an die Vorgehensweise aus [Bur19] wählen auch wir – im Sinne der Gauß-Laguerre-Quadratur¹³ – die N Nullstellen $\{\tau_i^{(N)}\}_{i=1}^N$ des N -ten Laguerre-Polynoms zu den Quadraturgewichten $\{\mathbf{w}_i^{(N)}\}_{i=1}^N$ und führen den zusätzlichen Kollokationspunkt $\tau_0^{(N)} = 0$ ein. Die zugehörigen Kollokationsgleichungen finden sich auf der nachfolgenden Seite.

¹³vg. [FH07], S. 184, (3.6.12) Satz

Kollokationsgleichungen der direkten Methode:

$$\sum_{i=1}^N \left\{ \left[(I_{n \times n})_{(\vartheta)} T_j(\tau_i^{(N)}) + \hat{A}_{(\vartheta)}^T(\tau_i^{(N)}) S_j(\tau_i^{(N)}) \right]^T W^{-1}(\tau_i^{(N)}) \left[\sum_{k=1}^N \hat{f}_k T_k(\tau_i^{(N)}) + E + \hat{A}^T(\tau_i^{(N)}) \left(\sum_{k=1}^N \hat{f}_k S_k(\tau_i^{(N)}) + \tau_i^{(N)} E + D \right) \right] \right. \\ \left. - \left[\sigma_N^T(\tau_i^{(N)}, D, E, F^N) \hat{B}_{(\vartheta)}^T(\tau_i^{(N)}) \right] S_j(\tau_i^{(N)}) \right\} \mathbf{w}_i + D^T W_{(\vartheta)}^{-1}(\tau_0^{(N)}) T_j(\tau_0^{(N)}) \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n \quad \forall j = 1, \dots, N$$

$$\sum_{i=1}^N \left\{ \left[(I_{n \times n})_{(\vartheta)} + \hat{A}_{(\vartheta)}^T(\tau_i^{(N)}) \cdot \tau_i^{(N)} \right]^T W^{-1}(\tau_i^{(N)}) \left[\sum_{k=1}^N \hat{f}_k T_k(\tau_i^{(N)}) + E + \hat{A}^T(\tau_i^{(N)}) \left(\sum_{k=1}^N \hat{f}_k S_k(\tau_i^{(N)}) + \tau_i^{(N)} E + D \right) \right] \right. \\ \left. - \left[\sigma_N^T(\tau_i^{(N)}, D, E, F^N) \hat{B}_{(\vartheta)}^T(\tau_i^{(N)}) \right] \cdot \tau_i^{(N)} \right\} \mathbf{w}_i + \lambda^T W_{(\vartheta)}^{-1}(\tau_0^{(N)}) \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n$$

$$\sum_{i=1}^N \left\{ \left[\hat{A}_{(\vartheta)}^T(\tau_i^{(N)}) \right]^T W^{-1}(\tau_i^{(N)}) \left[\sum_{k=1}^N \hat{f}_k T_k(\tau_i^{(N)}) + E + \hat{A}^T(\tau_i^{(N)}) \left(\sum_{k=1}^N \hat{f}_k S_k(\tau_i^{(N)}) + \tau_i^{(N)} E + D \right) \right] \right. \\ \left. - \sigma_N^T(\tau_i^{(N)}, D, E, F^N) \hat{B}_{(\vartheta)}^T(\tau_i^{(N)}) \right\} \mathbf{w}_i + x_{0(\vartheta)} + \lambda^T \left[W^{-1}(\tau_0^{(N)}) \hat{A}^T(\tau_0^{(N)}) \right]_{(\vartheta)} \stackrel{!}{=} 0 \quad \forall \vartheta = 1, \dots, n$$

$$W^{-1}(\tau_0^{(N)}) \left(E + \hat{A}^T(\tau_0^{(N)}) D \right) - x_0 \stackrel{!}{=} 0.$$

Kapitel 6

Beispiele

6.1 Das Beispiel von Garg, Hager und Rao

Wir beginnen mit der Betrachtung eines Beispiels aus [GHR11] von Garg, Hager und Rao. Gegenstand der Untersuchung ist das folgende nichtlineare Problem der Optimalsteuerung mit unendlichem Zeithorizont:

$$\begin{aligned} J(z, \tilde{u}) &= \frac{1}{2} \int_0^{\infty} \{ \ln^2 z(t) + \tilde{u}^2(t) \} dt \rightarrow \text{Min!} \\ z'(t) &= z(t) \ln z(t) + z(t) \tilde{u}(t), \quad z(0) = e^2, \tilde{u}(t) \in \mathbb{R}. \end{aligned}$$

Die Betrachtung dieses Beispiels war bereits Gegenstand der Ausarbeitung [Bur19]¹. Dort wurde die zitierte Aufgabe hinsichtlich Existenz- und Eindeutigkeitsaussagen sowie notwendiger Optimalitätsbedingungen untersucht und zudem die Anwendung einer indirekten Pseudospektralmethode diskutiert. Dazu wurde die in Kapitel 5 skizzierte Diskretisierung notwendiger Optimalitätsbedingungen mit Hilfe der Laguerre-Polynome $L_k^{(0,\rho)}$ verwendet. Die derartig in [Bur19] formulierten Ergebnisse sollen nun die Grundlage unserer Ausführungen bilden.

Zur Behandlung des obigen Beispiels wurde bereits in [PW12] eine Zustandstransformation der Form $\tilde{x}(t) := \ln z(t)$ vorgeschlagen. Die so eingeführte Transformation überführt die obige Aufgabe in ein linear-quadratisches Modell der Form

$$\begin{aligned} J(\tilde{x}, \tilde{u}) &= \frac{1}{2} \int_0^{\infty} \{ \tilde{x}^2(t) + \tilde{u}^2(t) \} dt \rightarrow \text{Min!} \\ \tilde{x}'(t) &= \tilde{x}(t) + \tilde{u}(t), \quad \tilde{x}(0) = 2, \tilde{u}(t) \in \mathbb{R}. \end{aligned}$$

Eine weitere Transformation des Zustandes $x(t) := \tilde{x}(t)e^t$ und der Steuerung $u(t) = \tilde{u}(t)e^t$ führt letztlich zu einem linear-quadratischen Optimalsteuerungsproblem der Form (P_ν) mit der Dichtefunktion $\nu(t) = e^{-2t}$. Es ergibt sich das folgende Problem der Optimalsteuerung

$$\begin{aligned} J(x, u) &= \frac{1}{2} \int_0^{\infty} \{ x^2(t) + u^2(t) \} e^{-2t} dt \rightarrow \text{Min!} \\ x'(t) &= 2x(t) + u(t), \quad \tilde{x}(0) = 2, u(t) \in \mathbb{R}. \end{aligned}$$

Für die weiteren Betrachtungen werden in [PW12] für die Zustandsfunktionen der Raum $W_2^1(\mathbb{R}_{++}, e^{-2t})$ und für die Steuerungsfunktionen der Raum $L_2(\mathbb{R}_{++}, e^{-2t})$ eingeführt. In

¹vgl. S. 157 ff.

[Bur19] wurde an diesen Ergebnissen angeknüpft und für das vorliegende linear-quadratische Problem die Existenz und Eindeutigkeit der optimalen Lösung (x_*, u_*) mit

$$x_*(t) = 2e^{(1-\sqrt{2})t} \quad \text{und} \quad u_*(t) = -2(1 + \sqrt{2})e^{(1-\sqrt{2})t} \quad (6.1)$$

nachgewiesen. Für die zugehörige adjungierte Funktion $y_* \in W_2^1(\mathbb{R}_{++}, e^{-2t})$ ergibt sich der Zusammenhang

$$y_*(t) = -2(1 + \sqrt{2})e^{(1-\sqrt{2})t}.$$

6.1.1 Einführung von Steuerungsbeschränkungen

Wir betrachten nun eine Modifizierung des Beispiels durch die Einführung von Steuerungsbeschränkungen. Bereits in [PW12] wurde folgende modifizierte Formulierung des Beispiels von Garg, Hager und Rao betrachtet:

$$\begin{aligned} J(z, \tilde{u}) &= \frac{1}{2} \int_0^\infty \{ \ln^2 z(t) + \tilde{u}^2(t) \} dt \rightarrow \text{Min!} \\ z'(t) &= z(t) \ln z(t) + z(t) \tilde{u}(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad z(0) = e^2, \\ |\tilde{u}(t)| &\leq g(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad g(t) = 6e^{-t} \quad \forall t \in \mathbb{R}_+. \end{aligned}$$

Die obig vorgeschlagenen Transformationen überführen das Problem in das folgende linear-quadratische Optimalsteuerungsproblem mit konvexem und kompaktem Steuerbereich U :

$$\begin{aligned} J(x, u) &= \frac{1}{2} \int_0^\infty \{ x^2(t) + u^2(t) \} e^{-2t} dt \rightarrow \text{Min!} \\ x'(t) &= 2x(t) + u(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad \tilde{x}(0) = 2, \\ |u(t)| &\leq 6 \quad (\forall) t \in \mathbb{R}_{++}. \end{aligned}$$

Auch hier legen wir die Funktionenräume

$$W_2^1(\mathbb{R}_{++}, e^{-2t}) \times L_2(\mathbb{R}_{++}, e^{-2t})$$

zugrunde. Offensichtlich ist die freie Lösung (6.1) auch zulässig für das Problem mit dem Steuerbereich $U = [-6, 6]$ und stellt somit auch die optimale Lösung für das Problem mit Steuerbeschränkungen dar². Wir wollen daher zu einer allgemeineren Betrachtung von Steuerungsbeschränkungen übergehen und untersuchen im Folgenden das Problem (P_ν) :

$$\begin{aligned} J(x, u) &= \frac{1}{2} \int_0^\infty \{ x^2(t) + u^2(t) \} e^{-2t} dt \rightarrow \text{Min!} \\ x'(t) &= 2x(t) + u(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad \tilde{x}(0) = x_0, \\ |u(t)| &\leq \bar{u} \quad (\forall) t \in \mathbb{R}_{++} \\ (x, u) &\in W_2^1(\mathbb{R}_{++}, e^{-2t}) \times L_2(\mathbb{R}_{++}, e^{-2t}), \end{aligned}$$

welches wir durch die Einführung der allgemeineren Funktion $g(t) = \bar{u}e^{-t}$ auf ganz \mathbb{R}_+ (mit $\bar{u} > 0$) und des beliebig aber fest wählbaren Anfangswertes $x(0) = x_0$ erhalten. Mit den Setzungen $W(t) \equiv R(t) \equiv B(t) \equiv 1$, $A(t) \equiv 2$ sowie $C(t) \equiv w(t) \equiv r(t) \equiv 0$ können wir dieses Problem der Klasse (P_ν) zuordnen. Für den Nachweis der Existenz einer optimalen Lösung genügt nach Satz 8 der Nachweis der Gültigkeit der Voraussetzungen (V1)-(V4)³. Unter den gegebenen Setzungen verifizieren wir leicht die Gültigkeit der Bedingungen (V1)-(V3). Auf den Nachweis der Bedingung (V4) gehen wir nun etwas ausführlicher ein.

²vgl. hierfür auch [Bur19], S. 193 ff.

³siehe Satz 8, S. 34 sowie Voraussetzungen (V1)-(V4) S. 25

6.1.2 Untersuchung der zulässigen Menge

Für den Nachweis der Gültigkeit der Voraussetzung (V4) ist zu zeigen, dass die Menge der zulässigen Lösungen nichtleer und nicht einelementig ist. Wir suchen daher Funktionen $(x, u) \in W_2^1(\mathbb{R}_{++}, e^{-2t}) \times L_2(\mathbb{R}_{++}, e^{-2t})$, die unter Einhaltung der Steuerungsbeschränkung $|u(t)| \leq \bar{u}$ ($\forall t \in \mathbb{R}_{++}$) der Differentialgleichung

$$x'(t) = 2x(t) + u(t)$$

zum gegebenen Anfangswert $x(0) = x_0$ fast überall auf \mathbb{R}_{++} genügen. Können wir zeigen, dass mindestens zwei zulässige Lösungen existieren, so ist die Gültigkeit der Voraussetzung (V4) gezeigt.

Die homogene Lösung der obigen Differentialgleichung lässt sich angeben durch den stetigen Repräsentanten

$$x_h(t) = ce^{2t} \quad \forall t \in \mathbb{R}_+.$$

Offensichtlich ist dieser für $c \neq 0$ nicht Element des Raumes $W_2^1(\mathbb{R}_{++}, e^{-2t})$ und somit nur zulässig, falls $x_0 = 0$. Diese Beobachtung impliziert zudem, dass für die Steuerung $u(t) \equiv 0$ kein zulässiger Kandidat existiert falls $x_0 \neq 0$.

Wir wählen daher den Ansatz $\hat{u}(t) = \alpha e^{\gamma t}$ für alle $t \in \mathbb{R}_+$ mit $\alpha \in [-\bar{u}, \bar{u}]$ und $\gamma < 0$. Offensichtlich handelt es sich bei der Funktion \hat{u} um eine stetige, monotone Funktion. Damit nimmt diese Funktion ihren betragsmäßig größten Wert für $t = 0$ an. Da der Parameter α in Abhängigkeit von den Steuerbeschränkungen gewählt wurde, ist gewährleistet, dass diese zum Zeitpunkt $t = 0$ den Restriktionen des Steuerbereiches \mathcal{U} genügen. Durch die Wahl des Parameters γ mit $\gamma < 0$ erreicht die Funktion für $t \rightarrow \infty$ zudem ihren betragsmäßig kleinsten Funktionswert, so dass die Funktion \hat{u} auch in diesem Fall in den Steuerbereich abbildet, denn es gilt

$$\lim_{T \rightarrow \infty} \hat{u}(T) = \lim_{T \rightarrow \infty} \alpha e^{\gamma T} = 0^+ \quad \text{für positive } \alpha \quad \text{und}$$

$$\lim_{T \rightarrow \infty} \hat{u}(T) = \lim_{T \rightarrow \infty} \alpha e^{\gamma T} = 0^- \quad \text{für negative } \alpha.$$

Zudem gilt mit dieser Setzung $\hat{u} \in L_2(\mathbb{R}_{++}, e^{-2t})$. Für diese Wahl der Steuerung ergibt sich als Lösung des Systems

$$\hat{x}(t) = \left(\frac{\alpha}{\gamma - 2} \right) e^{\gamma t}.$$

Für $\gamma < 1$ ist die Funktion \hat{x} Element des Raumes $W_2^1(\mathbb{R}_{++}, e^{-2t})$. Diese Bedingung ist durch die vorherige Setzung von $\gamma < 0$ stets erfüllt. Für den Nachweis der Zulässigkeit der so konstruierten Schar von Paaren (\hat{x}, \hat{u}) fordern wir nun die Gültigkeit der Anfangswertbedingungen. Diese Forderung liefert den Zusammenhang $x_0 = \frac{\alpha}{\gamma - 2}$. Gilt also $x_0(\gamma - 2) \in [-\bar{u}, \bar{u}]$, so sind durch (\hat{x}, \hat{u}) eine Schar von zulässigen Lösungen gegeben. Gilt für den Anfangswert x_0 des obigen Systems also

$$x_0 \in \left[- \left| \frac{\bar{u}}{\gamma - 2} \right|, \left| \frac{\bar{u}}{\gamma - 2} \right| \right]$$

so ist die Menge \mathcal{B} nichtleer und nicht einelementig und Voraussetzung (V4) ist erfüllt. Das größtmögliche Intervall erhalten wir offensichtlich für $\gamma \rightarrow 0$, so dass wir die Erkenntnisse dieses Unterabschnittes wie folgt zusammenfassen können:

Folgerung 5. *Gilt*

$$x_0 \in \left(-\frac{\bar{u}}{2}, \frac{\bar{u}}{2} \right),$$

dann existiert eine optimale Lösung des Problems (P_ν) .

Bemerkung 25. *Die in Folgerung 5 formulierten Ergebnisse stellen einen Zusammenhang zwischen der Zulässigkeit von Anfangswerten und Steuerungsbeschränkungen her. Dabei ist zu beachten, dass die Aussagen der Folgerung 5 nur hinreichend, aber nicht notwendig sind. Finden wir allerdings eine Paarung von Anfangswerten und Steuerungsbeschränkungen welche die Bedingungen der Folgerung 5 erfüllen, so ist die Existenz einer Optimallösung gesichert.*

6.1.3 Übergang zur dualen Aufgabe

Gemäß der Betrachtungen aus Kapitel 4 ist die zu (P_ν) duale Aufgabe (D_ν) gegeben durch

Minimiere

$$G(y, q) = \int_0^\infty \left\{ \frac{1}{2} q^2(t) + \theta(t, y(t)) \right\} e^{-2t} dt + \tilde{x}_0^T y(0)$$

bezüglich aller

$$(y, q) \in W_2^2(\mathbb{R}_{++}, e^{-2t}) \times W_2^1(\mathbb{R}_{++}, e^{-2t}),$$

für die gilt

$$y'(t) = q(t), \quad \text{f.ü. auf } \mathbb{R}_{++}, \quad q(0) = \tilde{x}_0,$$

wobei

$$\theta(t, y(t)) = -\frac{1}{2} \sigma^2(t, y(t)) + \sigma(t, y(t)) y(t) \quad \text{und} \quad \sigma(t, y(t)) = \min \{ \max \{ -\bar{u}, y(t) \}, \bar{u} \} \quad \text{ist.}$$

Sind die Bedingungen (V1)-(V4) und (VD1)-(VD4) erfüllt und gilt zudem $\tilde{Y}_{\nu_0} \neq \emptyset$, so können wir auf Grundlage von Satz 20 auf die Existenz einer optimalen Lösung der dualen Aufgabe und auf Grundlage von Satz 23 auf notwendige und hinreichende Optimalitätsbedingungen schließen⁴. Die Gültigkeit der Bedingungen (V1)-(V4) haben wir bereits im vorangegangenen Unterkapitel untersucht. Für die hier betrachtete duale Aufgabe lassen sich die Bedingungen (VD1)-(VD3) leicht verifizieren. Für den Nachweis der Gültigkeit von (VD4) genügt nun wieder die Angabe mindestens zweier Paare, welche zulässig für die duale Aufgabe sind. Diese lassen sich aufgrund der fehlenden Steuerungsbeschränkungen für die duale Steuerungsfunktion q sowie der vorliegenden Funktionenräume in Form gewichteter Sobolev-Räume mit echter Dichtefunktion leicht finden. Offensichtlich sind durch (\hat{y}, \hat{q}) und (\tilde{y}, \tilde{q}) zwei zulässige Lösungen gegeben, wobei

$$\hat{y} = x_0 t, \quad \hat{q} \equiv x_0 \quad \text{und} \quad \tilde{y} = x_0 t + t^2, \quad \tilde{q} = x_0 + t.$$

Damit sind die Voraussetzungen (VD1)-(VD4) erfüllt. Für die gegebenen Paare können wir zudem schließen $(\hat{y}, \hat{q}) \in W_2^2(\mathbb{R}_{++}, e^{-2t}) \times W_2^1(\mathbb{R}_{++}, e^{-2t})$, mit $\hat{q}(0) = x_0$, so dass zudem $(\hat{y}, \hat{q}) \in \mathcal{B}_D$ und damit $\mathcal{B}_D \neq \emptyset$ gelten. Wir zeigen nun die Gültigkeit der Bedingung

⁴siehe Voraussetzungen (V1)-(V4) auf S. 25, (VD1)-(VD4) S. 76, Satz 20, S. 101 und Satz 23, S. 104

$\tilde{Y}_{\nu_0} \neq \emptyset$: Für diesen Nachweis genügt die Angabe eines Paares $(y, q) \in \mathcal{B}_D$, welches der Differentialgleichung

$$\begin{aligned} y''(t) &= -\hat{W}(t) \left(\hat{W}^{-1} \right)'(t) q(t) - \left[\left(\hat{A}'(t) \right)^T y(t) + \hat{A}^T(t) y'(t) \right] \\ &\quad + \hat{W}(t) A(t) \hat{W}^{-1}(t) q(t) + \hat{W}(t) B(t) \hat{\sigma}(t, y(t)) \end{aligned}$$

fast überall auf \mathbb{R}_{++} genügt. Die obige Differentialgleichung ergibt sich für die Setzung von $\hat{W}(t) = \hat{R}(t) = B(t) \equiv 1$ und $\hat{A}(t) \equiv 2 - 2 = 0$ zu

$$y''(t) = 2q(t) + \hat{\sigma}(t, y(t)) = 2y'(t) + \hat{\sigma}(t, y(t)).$$

Wir betrachten die Funktion

$$y(t) = \left(\frac{x_0}{1 - \sqrt{2}} \right) e^{(1 - \sqrt{2})t}. \quad (6.2)$$

Gilt $x_0 \in [-(\sqrt{2} - 1)\bar{u}, (\sqrt{2} - 1)\bar{u}]$, so gilt für die Funktion y aus (6.2) offenbar $y(t) \in [-\bar{u}, \bar{u}] \forall t \in \mathbb{R}_+$ und damit

$$\hat{\sigma}(t, y(t)) = \text{Proj}_{[-\bar{u}, \bar{u}]} \left(R^{-1}(t) (-\hat{r}(t) + B^T(t) y(t)) \right) = \text{Proj}_{[-\bar{u}, \bar{u}]} (y(t)) = y(t).$$

Zudem gilt für die Funktion aus (6.2)

$$y'(t) = x_0 e^{(1 - \sqrt{2})t} \quad \text{mit} \quad y'(0) = q(0) = x_0 \quad \text{und} \quad y''(t) = (1 - \sqrt{2}) x_0 e^{(1 - \sqrt{2})t}.$$

Insgesamt erhalten wir damit die Aussage:

Folgerung 6. *Gilt*

$$x_0 \in [-(\sqrt{2} - 1)\bar{u}, (\sqrt{2} - 1)\bar{u}],$$

dann existiert eine optimale Lösung der dualen Aufgabe.

Beweis. Sei y aus (6.2) und $x_0 \in [-(\sqrt{2} - 1)\bar{u}, (\sqrt{2} - 1)\bar{u}]$ gegeben. Dann gilt

$$\begin{aligned} y''(t) &= (1 - \sqrt{2}) x_0 e^{(1 - \sqrt{2})t} = 2x_0 e^{(1 - \sqrt{2})t} + \left(\frac{x_0}{1 - \sqrt{2}} \right) e^{(1 - \sqrt{2})t} \\ &= 2y'(t) + y(t) = 2y'(t) + \hat{\sigma}(t, y(t)). \end{aligned}$$

Damit ist $y \in \tilde{Y}_{\nu_0}$ und somit $\tilde{Y}_{\nu_0} \neq \emptyset$. Die Gültigkeit der Bedingungen (V1)-(V4) sowie (VD1)-(VD4) wurde für diese Setzung an Anfangswerten bereits gezeigt. Damit folgt aus Satz 20 die Existenz einer optimalen Lösung der dualen Aufgabe. \square

Bemerkung 26. *Die Bedingungen der Folgerung 6 liefern die Existenz einer optimalen Lösung der dualen Aufgabe für ein kleineres Intervall an zulässigen Anfangswerten als die Bedingungen der Folgerung 5, welche die Existenz einer primalen Lösung garantieren. Die explizite Berechnung der optimalen Lösungen im nachfolgenden Unterkapitel wird uns jedoch den Nachweis liefern, dass auch für die Bedingung*

$$x_0 \in \left(-\frac{\bar{u}}{2}, \frac{\bar{u}}{2} \right)$$

die Menge \tilde{Y}_{ν_0} nichtleer und somit ebenfalls die Existenz einer Lösung der dualen Aufgabe unter dieser Bedingung gesichert ist.

6.1.4 Optimale Lösung

Für das vorliegende primale Problem ist eine Aufgabe mit dem Gewichtungparameter $\beta = -2$ und der zugehörigen Dynamik

$$x'(t) = Ax(t) + u(t)$$

mit $A = 2$ gegeben. Durch diese Setzungen ist das Maximumprinzip aus Satz 10 nicht anwendbar⁵. Die Voraussetzung (Z5) fordert $\beta < -4$ und steht somit im Widerspruch zu $\beta = -2$. Wir konnten jedoch zeigen, dass unter der Einhaltung der Bedingung aus Folgerung 5 eine optimale Lösung der primalen Aufgabe existiert. Zudem verifizierten wir die Gültigkeit der Bedingungen (VD1)-(VD4) und \tilde{Y}_{ν_0} für $x_0 \in [-(\sqrt{2}-1)\bar{u}, (\sqrt{2}-1)\bar{u}]$. Dies erlaubt den Schluss auf die Gültigkeit des Satzes 23 für eine derartige Wahl von Anfangswerten, welche auf Grundlage der so nachgewiesenen starken Dualität, notwendige und hinreichende Optimalitätsbedingungen für die primale Aufgabenstellung liefert.

Zur Berechnung der optimalen Lösung der vorliegenden primalen und dualen Problemformulierungen ziehen wir also die notwendigen und hinreichenden Bedingungen aus Satz 23 heran⁶. Demnach ist die Pontryaginsche Funktion der dualen Aufgabe gegeben durch

$$H(t, \xi_D, v_D, \eta_D, 1) = \left(-\frac{1}{2}v_D^2 + \frac{1}{2}\sigma^2(t, \xi_D) - \sigma(t, \xi)\xi \right) e^{-2t} + \eta_D \cdot v_D.$$

Zudem existieren für die optimale Lösung $(y_*, q_*) \in W_2^{2,1}(\mathbb{R}_{++}, e^{-2t}) \times W_2^{1,1}(\mathbb{R}_{++}, e^{-2t})$ Multiplikatoren $\lambda_0 = 1$ und $p_0 \in W_2^{1,1}(\mathbb{R}_{++}, e^{2t})$, so dass die Maximumbedingung

$$H(t, y_*(t), q_*(t), p_0(t), \lambda_0) = \max_{v_D \in \mathbb{R}} H(t, y_*(t), v_D, p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++} \quad (M)$$

und die kanonische Gleichung

$$p_0'(t) = -\nabla_{\xi_D} H(t, y_*(t), q_*(t), p_0(t), \lambda_0) \quad (\forall) t \in \mathbb{R}_{++} \quad (K)$$

zum Anfangswert

$$p_0(0) = x_0$$

erfüllt sind.

Aufgrund der fehlenden Steuerbeschränkungen auf dualer Seite lässt sich die Maximumbedingung leicht auswerten und wir erhalten den Zusammenhang

$$q_*(t) = p_0(t)e^{2t}.$$

In Folge der Struktur der Funktion $\sigma(t, \xi_D) = \min \{ \max \{ -\bar{u}, \xi_D \}, \bar{u} \}$ ergibt sich für die kanonische Gleichung die Formulierung:

$$p_0'(t) = \sigma(t, y_*(t))e^{-2t} = \begin{cases} -\bar{u}e^{-2t}, & -\bar{u} > y_*(t) \\ y_*(t)e^{-2t}, & y_*(t) \in [-\bar{u}, \bar{u}] \\ \bar{u}e^{-2t}, & \bar{u} < y_*(t) \end{cases}.$$

⁵vgl. Satz 10, S. 41

⁶vgl. Satz 23, S. 104

Für die weitere Berechnung scheint es sinnvoll, die Lösungsstruktur der zugrunde liegenden Aufgabe zu analysieren. Dafür orientieren wir uns zunächst an der freien Lösung (\tilde{y}, \tilde{p}_0) ⁷ des freien kanonischen Systems

$$\begin{aligned}\tilde{y}'(t) &= \tilde{p}_0(t)e^{2t} \\ \tilde{p}'_0(t) &= \tilde{y}(t)e^{-2t}\end{aligned}$$

unter Beachtung der Bedingungen $\tilde{p}'_0(0) = x_0$ und $\tilde{p}_0 \in W_2^{1,1}(\mathbb{R}_{++}, e^{2t})$, deren Auswertung auf die freie Lösung

$$\tilde{y}(t) = \frac{x_0}{(1 - \sqrt{2})} e^{(1 - \sqrt{2})t} \quad (6.3)$$

führt. Abbildung 6.1 skizziert den Verlauf der Schaltfunktion σ_* (-) für den Fall fehlender Steuerungsbeschränkungen und veranschaulicht die mögliche Strukturveränderung für die Behandlung aktiver Steuerungsbeschränkungen. Dies motiviert zunächst den folgenden Ansatz für die Schaltfunktion:

$$\sigma_*(t, y_*(t)) = \begin{cases} -\bar{u}, & t < \bar{t} \\ y_*(t), & t \geq \bar{t}. \end{cases} \quad (6.4)$$

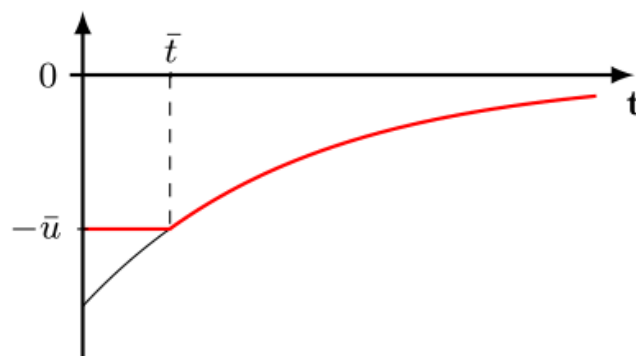


Abbildung 6.1: Ansatz für die Schaltfunktion σ bei Box-Restriktionen $[-\bar{u}, \bar{u}]$ mit $\bar{u} > 0$ und $x_0 > -\bar{u}(1 - \sqrt{2})$ (entnommen aus [Bur19], Abbildung 6.17, S. 195)

Im Folgenden setzen wir daher die Existenz eines Umschaltzeitpunktes $\bar{t} > 0$ voraus und unterstellen eine Lösungsstruktur der Schaltfunktion der Form (6.4). Wir wollen überprüfen, ob wir mit einer derartigen Wahl einen Lösungskandidaten finden, der den Bedingungen des Satzes 23 genügt. Dafür beginnen wir mit der Betrachtung des Intervalls $[0, \bar{t}]$:

Auswertung der Bedingungen auf $[0, \bar{t}]$:

Der Ansatz (6.4) liefert $\sigma(t, \xi_D) = -\bar{u}$. Unter Beachtung der aus der Maximumbedingung resultierenden Gleichung $q_*(t) = p_0(t)e^{2t}$ erhalten wir damit das kanonische System

$$\begin{aligned}y'(t) &= p_0(t)e^{2t} \\ p'_0(t) &= -\bar{u}e^{-2t}\end{aligned}$$

⁷Diese Annahme deckt sich mit der Annahme einer Schaltstruktur $\sigma(t, \xi_D) = \xi_D$ für alle $t \in [0, \infty)$.

zum Anfangswert $p_0(0)$. Dessen Auswertung liefert die Lösungen

$$\begin{aligned} y(t) &= \left(\frac{x_0}{2} - \frac{\bar{u}}{4}\right)e^{2t} + \frac{\bar{u}}{2}t + D \\ p_0(t) &= \left(x_0 - \frac{\bar{u}}{2}\right) + \frac{\bar{u}}{2}e^{-2t} \end{aligned}$$

mit der noch zu bestimmenden Konstante D .

Auswertung der Bedingungen auf $[\bar{t}, +\infty)$:

Der Ansatz (6.4) liefert auf dem betrachteten Intervall $\sigma(t, \xi_D) = \xi_D$. Unter Beachtung der aus der Maximumbedingung resultierenden Gleichung $q_*(t) = p_0(t)e^{2t}$ erhalten wir damit das kanonische System

$$\begin{aligned} y'(t) &= p_0(t)e^{2t} \\ p_0'(t) &= y(t)e^{-2t}. \end{aligned}$$

Dessen Auswertung liefert die Lösungen

$$\begin{aligned} y(t) &= C_1e^{(1-\sqrt{2})t} + C_2e^{(1+\sqrt{2})t} \\ p_0(t) &= C_1(1-\sqrt{2})e^{(-1-\sqrt{2})t} + C_2(1+\sqrt{2})e^{(1+\sqrt{2})t} \end{aligned}$$

mit den noch zu bestimmenden Konstanten C_1 und C_2 .

Die Transversalitätsbedingung $p_0 \in W^{1,2}(\mathbb{R}_{++}, e^{2t})$ impliziert den Zusammenhang

$$\lim_{T \rightarrow \infty} p_0(T) = 0.$$

Aufgrund des positiven Exponenten $(1 + \sqrt{2})t$ liefert die Transversalitätsbedingung den Schluss $C_2 = 0$. Damit erhalten wir mit der Setzung $C := C_1$ für das Intervall $[\bar{t}, +\infty)$ die Lösungen

$$\begin{aligned} y(t) &= Ce^{(1-\sqrt{2})t} \\ p_0(t) &= C(1-\sqrt{2})e^{(1-\sqrt{2})t}. \end{aligned}$$

Offen ist nun noch die Wahl der Konstanten C und D sowie die Berechnung des Umschaltzeitpunktes \bar{t} . Diese fehlenden Werte erhalten wir aufgrund der Regularitätsvoraussetzungen an die Funktionen y , p_0 und σ ⁸, nach welchen wir im Schaltpunkt \bar{t} von der Existenz eines stetigen Repräsentanten ausgehen können. Diese Überlegungen implizieren die folgenden Gleichungen:

$$\begin{aligned} \text{I} \quad & \left(\frac{x_0}{2} - \frac{\bar{u}}{4}\right)e^{2\bar{t}} + \frac{\bar{u}}{2}\bar{t} + D = -\bar{u} \\ \text{II} \quad & Ce^{(1-\sqrt{2})\bar{t}} = -\bar{u} \\ \text{III} \quad & \left(x_0 - \frac{\bar{u}}{2}\right) + \frac{\bar{u}}{2}e^{-2\bar{t}} = C(1-\sqrt{2})e^{(1-\sqrt{2})\bar{t}}. \end{aligned}$$

Das Lösen des Gleichungssystems liefert die noch fehlenden Werte C , D und \bar{t} . Unter Beachtung der in Satz 23 formulierten Beziehungen

$$x_*(t) = q_*(t) \quad \text{und} \quad u_*(t) = \text{Proj}_{[-\bar{u}, \bar{u}]}(y_*(t))$$

⁸für die Regularität der Funktion σ vgl. die Ausführungen von S. 42 ff.

können wir die Kandidaten für die optimalen Lösungen der primalen und dualen Aufgaben schließlich angeben durch

$$x_*(t) = \begin{cases} (x_0 - \frac{\bar{u}}{2})e^{2t} + \frac{\bar{u}}{2}, & \text{falls } t \leq \bar{t} \\ C(1 - \sqrt{2})e^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}, \quad u_*(t) = \begin{cases} -\bar{u}, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}, \quad (6.5)$$

$$y_*(t) = \begin{cases} D + (\frac{x_0}{2} - \frac{\bar{u}}{4})e^{2t} + \frac{\bar{u}}{2}t, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}, \quad (6.6)$$

wobei \bar{t} den Umschaltzeitpunkt

$$\bar{t} := \frac{1}{2} \ln \left(\frac{-3\bar{u} + 2\sqrt{2}\bar{u}}{2x_0 - \bar{u}} \right)$$

für $\bar{u}(\sqrt{2} - 1) < x_0 < \frac{\bar{u}}{2}$ bezeichne und die auftretenden Konstanten C und D berechnet werden können durch

$$C = -\bar{u}e^{(\sqrt{2}-1)\bar{t}} \quad \text{und} \quad D = \left(\frac{\bar{u}}{4} - \frac{x_0}{2}\right)e^{2\bar{t}} - \frac{\bar{u}}{2}\bar{t} - \bar{u}.$$

Dabei ist zu beachten, dass die sinnvolle Auswertung des Punktes \bar{t} die berechnete Lösungsstruktur auf den Fall $x_0 \in \left((-1 + \sqrt{2})\bar{u}, \frac{\bar{u}}{2}\right)$ eingrenzt. Damit liegt x_0 jedoch außerhalb des in Folgerung [6](#) angegebenen Intervalls, für welches wir die Existenz einer Lösung der dualen Aufgabe und damit die Gültigkeit von Satz [23](#) sichern können. Allerdings lässt sich leicht einsehen, dass für das hier berechnete y_* gilt $y_* \in \tilde{Y}_{\nu_0}$ und damit $\tilde{Y}_{\nu_0} \neq \emptyset$.

Zu zeigen ist: Für die in [\(6.6\)](#) angegebene Funktion

$$y_*(t) = \begin{cases} D + (\frac{x_0}{2} - \frac{\bar{u}}{4})e^{2t} + \frac{\bar{u}}{2}t, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}$$

gilt: y_* ist Element des Raumes $W_2^2(\mathbb{R}_{++}, e^{-2t})$ und löst die Differentialgleichung

$$y''(t) = 2y'(t) + \hat{\sigma}(t, y(t)) \quad (6.7)$$

zu der Anfangsbedingung $y'(0) = x_0$. Die Raumzugehörigkeit $y_* \in W_2^2(\mathbb{R}_{++}, e^{-2t})$ ist offensichtlich. Die verallgemeinerten Ableitungen der Funktion y_* lassen sich angeben mit

$$y'_*(t) = \begin{cases} (x_0 - \frac{\bar{u}}{2})e^{2t} + \frac{\bar{u}}{2}, & \text{falls } t \leq \bar{t} \\ (1 - \sqrt{2})Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases} \quad \text{und} \quad y''_*(t) = \begin{cases} (2x_0 - \bar{u})e^{2t}, & \text{falls } t \leq \bar{t} \\ (1 - \sqrt{2})^2Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}.$$

Für die Funktion $\hat{\sigma}(t, y(t)) = \text{Proj}_{[-\bar{u}, \bar{u}]}(y(t))$ gilt damit

$$\hat{\sigma}(t, y_*(t)) = \begin{cases} -\bar{u}, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}.$$

Wir betrachten nun das Intervall $[0, \bar{t}]$: Offensichtlich erfüllt y_* die Anfangsbedingung $y'_*(0) = x_0$. Weiterhin ergibt sich die Gleichung [\(6.7\)](#) zu

$$y''_*(t) = 2y'_*(t) - \bar{u} = 2 \left\{ (x_0 - \frac{\bar{u}}{2})e^{2t} + \frac{\bar{u}}{2} \right\} - \bar{u} = (2x_0 - \bar{u})e^{2t}.$$

Für das Intervall (\bar{t}, ∞) liefert die Gleichung (6.7)

$$\begin{aligned} y_*''(t) &= 2y_*'(t) + y_*(t) = 2(1 - \sqrt{2})Ce^{(1-\sqrt{2})t} + Ce^{(1-\sqrt{2})t} \\ &= (3 - 2\sqrt{2})Ce^{(1-\sqrt{2})t} = (1 - \sqrt{2})^2 Ce^{(1-\sqrt{2})t}. \end{aligned}$$

Damit gilt $y \in \tilde{Y}_{\nu_0}$. Die Existenz einer optimalen Lösung der dualen Aufgabe und somit die Anwendbarkeit von Satz 23 sind damit auch für $x_0 \in \left(-1 + \sqrt{2}\bar{u}, \frac{\bar{u}}{2}\right)$ gesichert. Insgesamt liefern die vorangegangenen Ausführungen den Schluss: Die Funktionen aus (6.5)-(6.6) erfüllen alle Bedingungen aus Satz 23. Das Paar (x_*, u_*) stellt somit die optimale Lösung der primalen Aufgabe und das Paar (y_*, q_*) die optimale Lösung der dualen Aufgabe dar. Der Zielfunktionalwert ergibt sich in diesem Fall zu

$$J(x_*, u_*) = \left(\frac{x_0}{2} - \frac{\bar{u}}{4}\right)^4 e^{2\bar{t} + \bar{u}} \left(x_0 - \frac{\bar{u}}{2}\right) \bar{t} - \frac{5}{8}\bar{u}^2 e^{-2\bar{t}} - \left(\frac{x_0}{2} - \frac{\bar{u}}{4}\right)^4 + \frac{5}{8}\bar{u}^2 - (1 - \sqrt{2})C^2 e^{-2\sqrt{2}\bar{t}},$$

wobei gilt

$$J(x_*, u_*) = G(y_*, q_*).$$

Bemerkung 27.

1. Die Folgerung 5 sichert die Existenz einer eindeutigen optimalen Lösung der primalen Aufgabe für alle $x_0 \in \left(-\frac{\bar{u}}{2}, \frac{\bar{u}}{2}\right)$.
2. Für $x_0 \in \left(-1 + \sqrt{2}\bar{u}, \frac{\bar{u}}{2}\right)$ konnten wir die eindeutige optimale Lösung der primalen Aufgabe durch (6.5) und (6.6) angeben.
3. Die Unterstellung einer Schaltstruktur der Form $\sigma(t, \xi_D) = \xi_D$ für alle $t \in [0, \infty)$ lieferte die Lösung (6.3). Aus dieser lassen sich mit Hilfe der Bedingungen aus Satz 23 die zugehörigen primalen Lösungen

$$x_*(t) = x_0 e^{(1-\sqrt{2})t} \quad \text{und} \quad u_*(t) = \frac{x_0}{(1-\sqrt{2})} e^{(1-\sqrt{2})t} \quad (6.8)$$

berechnen. Offensichtlich ist die Bedingung $\sigma(t, y_*(t)) = u_*(t) \in [-\bar{u}, \bar{u}]$ erfüllt für alle $x_0 \in \left[(1 - \sqrt{2})\bar{u}, -(1 - \sqrt{2})\bar{u}\right]$. Die Formeln (6.3) und (6.8) liefern für diesen Fall die eindeutigen optimalen Lösungen.

4. Unter der Annahme einer Schaltstruktur von

$$\sigma(t, \xi_D) = \begin{cases} \bar{u}, & t < \bar{t} \\ \xi_D, & t \geq \bar{t} \end{cases}$$

erhalten wir durch ein ähnliches Vorgehen wie im ausführlich geschilderten Fall die eindeutige optimale Lösung

$$x_*(t) = \begin{cases} (x_0 + \frac{\bar{u}}{2})e^{2t} - \frac{\bar{u}}{2}, & \text{falls } t \leq \bar{t} \\ C(1 - \sqrt{2})e^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}, \quad u_*(t) = \begin{cases} \bar{u}, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases}, \quad (6.9)$$

⁹Dabei gilt $q_*(t) = y_*'(t)$ für alle $t \in \mathbb{R}_+$.

$$y_*(t) = \begin{cases} D + \left(\frac{x_0}{2} + \frac{\bar{u}}{4}\right)e^{2t} - \frac{\bar{u}}{2}t, & \text{falls } t \leq \bar{t} \\ Ce^{(1-\sqrt{2})t}, & \text{falls } t > \bar{t} \end{cases} \quad (6.10)$$

$$\bar{t} := \frac{1}{2} \ln \left(\frac{3\bar{u} - 2\sqrt{2}\bar{u}}{2x_0 + \bar{u}} \right)$$

$$\text{sowie } C = \bar{u}e^{(\sqrt{2}-1)\bar{t}} \quad \text{und} \quad D = -\left(\frac{\bar{u}}{4} + \frac{x_0}{2}\right)e^{2\bar{t}} + \frac{\bar{u}}{2}\bar{t} + \bar{u}$$

für alle $x_0 \in \left(-\frac{\bar{u}}{2}, (1-\sqrt{2})\bar{u}\right)$. Auch hier lässt sich zeigen, dass die Funktion y_* Element der Menge \tilde{Y}_{ν_0} ist und damit eine optimale Lösung der dualen Aufgabe auch für diese Wahl der Anfangswerte existiert. Somit haben wir schließlich für alle in Folgerung 5 beschriebenen Fälle eine eindeutige optimale Lösung angeben können.

5. Die analytischen Lösungen liefern für alle betrachteten Fälle den Zusammenhang $y_* \in W_2^2(\mathbb{R}_{++}, te^{-2t})$. Damit ist die modifizierte Methode anwendbar, auch wenn die feste Wahl der Gewichtsfunktion $\nu(t) = e^{-2t}$ eine Argumentation im Sinne von Lemma 44 nicht zulässt¹⁰.

6.1.5 Anwendung der direkten Pseudospektralmethode

Für die numerische Behandlung mittels der in den Unterkapiteln 5.1 und 5.2 vorgestellten direkten (modifizierten) Pseudospektralmethode gehen wir mit Hilfe der Setzung

$$q(t) := y'(t) + \hat{A}y(t) = y'(t) + (2-2)y(t) = y'(t) \quad (\forall)t \in \mathbb{R}_{++}$$

über zu der Formulierung der dualen Aufgabe in Form eines Variationsproblems. Diese Setzung führt uns zu der Aufgabe

Minimiere

$$G(y, y') := \int_0^\infty \left\{ \frac{1}{2} (y'(t))^2 + \theta(t, y(t)) \right\} e^{-2t} dt + x_0^T y(0)$$

bezüglich aller

$$y \in W_2^2(\mathbb{R}_{++}, e^{-2t}) \quad \text{mit} \quad x_0 = y'(0)$$

wobei

$$\theta(t, y(t)) = -\frac{1}{2}\sigma^2(t, y(t)) + \sigma(t, y(t))y(t) \quad \text{und} \quad \sigma(t, y(t)) = \min \{ \max \{ -\bar{u}, y(t) \}, \bar{u} \}$$

ist.

Entsprechend des Vorgehens in Kapitel 5.2 wählen wir nun die Ansätze

$$y_N(t) = \sum_{k=1}^N a_k S_k(t) + E \cdot t + D \quad \text{und} \quad y'_N(t) = \sum_{k=1}^N a_k T_k(t) + E$$

und setzen diese in das Variationsproblem ein. Durch dieses Vorgehen erhalten wir das endlich dimensionale, konvexe Minimierungsproblem

$$G^N((a_1, \dots, a_N), E, D) := \int_0^\infty \tilde{g}_N(t, \mathbf{a}, E, D) e^{-2t} dt + x_0 D$$

¹⁰siehe Lemma 44, S. 111

bezüglich aller

$$(\mathbf{a}, E, D) \in \mathbb{R}^{N+2} \quad \text{mit} \quad x_0 = E,$$

wobei

$$\begin{aligned} \tilde{g}_N(t, \mathbf{a}, E, D) = & \frac{1}{2} \left(\sum_{k=1}^N a_k T_k(t) + E \right)^2 - \frac{1}{2} \sigma_N^2(t, \mathbf{a}, E, D) \\ & + \sigma_N(t, \mathbf{a}, E, D) \left(\sum_{k=1}^N a_k S_k(t) + E \cdot t + D \right) \end{aligned}$$

mit

$$\sigma_N(t, \mathbf{a}, E, D) = \min \left\{ \max \left\{ -\bar{u}, \sum_{k=1}^N a_k S_k(t) + E \cdot t + D \right\}, \bar{u} \right\} \quad \text{ist.}$$

Die Auswertung der Nebenbedingung erlaubt uns die Eliminierung der Variable E und führt uns zu dem unrestringierten, konvexen und endlich dimensionalen Optimierungsproblem:

Minimiere

$$\int_0^{\infty} \hat{g}_N(t, \mathbf{a}, D) e^{-2t} dt + x_0 D$$

bezüglich aller

$$(\mathbf{a}, D) \in \mathbb{R}^{N+1},$$

wobei

$$\begin{aligned} \hat{g}_N(t, \mathbf{a}, D) = & \frac{1}{2} \left(\sum_{k=1}^N a_k T_k(t) + x_0 \right)^2 - \frac{1}{2} \sigma_N^2(t, \mathbf{a}, x_0, D) \\ & + \sigma_N(t, \mathbf{a}, x_0, D) \left(\sum_{k=1}^N a_k S_k(t) + x_0 \cdot t + D \right) \end{aligned}$$

mit

$$\sigma_N(t, \mathbf{a}, x_0, D) = \min \left\{ \max \left\{ -\bar{u}, \sum_{k=1}^N a_k S_k(t) + x_0 \cdot t + D \right\}, \bar{u} \right\} \quad \text{ist.}$$

Aufgrund der Konvexität sind alle Lösungen dieses endlich dimensionalen Optimierungsproblems globale Lösungen und der Zielfunktionswert ist eindeutig bestimmt.

Gemäß der Betrachtungen aus Kapitel [5.2](#) erfüllt die Optimallösung die Kollokationsgleichungen der direkten Methode. Diese liefern das folgende System aus $N + 1$ Gleichungen mit $N + 1$ Variablen:

$$\begin{aligned} \sum_{i=1}^N \mathbf{w}_i \left\{ \sum_{k=1}^N a_k \left\{ T_j(\tau_i^{(N)}) T_k(\tau_i^{(N)}) \right\} + 2x_0 S_j(\tau_i^{(N)}) + S_j(\tau_i^{(N)}) \sigma_N(\tau_i^{(N)}, \mathbf{a}, D) \right\} & \stackrel{!}{=} 0 \quad \forall j = 1, \dots, N \\ \sum_{i=1}^N \mathbf{w}_i \left\{ \sigma_N(\tau_i^{(N)}, \mathbf{a}, D) + 2x_0 \right\} & \stackrel{!}{=} 0, \end{aligned}$$

wobei

$$\sigma_N(\tau_i^{(N)}, \mathbf{a}, D) = \min \left\{ \max \left\{ -\bar{u}, \sum_{k=1}^N a_k S_k(\tau_i^{(N)}) + x_0 \cdot \tau_i^{(N)} + D \right\}, \bar{u} \right\} \quad \text{ist.}$$

6.1.6 Auswertung der numerischen Ergebnisse

Wir widmen uns nun der numerischen Behandlung des so konstruierten Problems für eine unterschiedliche Wahl von Steuerungsbeschränkungen mit Hilfe der in Kapitel 5.2 vorgeschlagenen dual-basierten direkten Pseudospektralmethode. Zum Vergleich der in Kapitel 5 angesprochenen indirekten und direkten Pseudospektralverfahren richten wir uns dafür nach den in [Bur19] diskutierten Setzungen des eingeführten Problems und beginnen mit der Betrachtung des Steuerbereiches $\mathcal{U} = [-6, 6]$.

Betrachtung für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$

Wir betrachten das Beispiel von Garg, Hager und Rao unter Beachtung der Steuerungsbeschränkungen $\mathcal{U} = [-6, 6]$ zum Anfangswert $x_0 = 2$. In diesem Fall gilt

$$x_0 \in \left[(1 - \sqrt{2})\bar{u}, -(1 - \sqrt{2})\bar{u} \right] \approx (-2.485281374, 2.485281374).$$

Demnach können wir die optimale Lösung mit Hilfe der Formeln aus (6.8) und (6.3) angeben. Für den konkreten Fall ergeben sich diese zu

$$\begin{aligned} x_*(t) &= 2e^{(1-\sqrt{2})t}, & u_*(t) &= -2(1 + \sqrt{2})e^{(1-\sqrt{2})t} \quad \text{und} \\ y_*(t) &= -2(1 + \sqrt{2})e^{(1-\sqrt{2})t}. \end{aligned}$$

Der zugehörige Zielfunktionalwert berechnet sich damit zu

$$J(x_*, u_*) = G(y_*, q_*) = 2(1 + \sqrt{2}) \approx 4.828427124746190.$$

Wir präsentieren nun die numerischen Ergebnisse für eine unterschiedliche Anzahl N an Kollokationspunkten. Dabei werden wir feststellen, dass die verwendete Pseudospektralmethode bereits für $N = 4$, d.h. für die Verwendung von lediglich 4 Stützstellen, eine derart gute Approximation an die optimale Lösung liefert, dass der exakte Zielfunktionswert bereits bis auf 4 Nachkommastellen genau approximiert werden kann. Die Approximation geschieht dabei durch die Auswertung der Kollokationsgleichungen an den Punkten $\tau_i^{(4)}$, $1 \leq i \leq 4$ von $L_4^{(1,2)}$ zu den Ansätzen

$$y_4(t) = \sum_{k=1}^4 a_k S_k(t) + x_0 t + D, \quad y_4'(t) = \sum_{k=1}^4 a_k T_k(t) + x_0, \quad (6.11)$$

d.h. durch Polynome vom Grad 5. In diesem Fall ergeben sich die Kollokationspunkte zu

$$\begin{aligned} \tau_1^{(4)} &= 0.161273844809696 & \tau_2^{(4)} &= 0.872880550579173 \\ \tau_3^{(4)} &= 2.268310148460564 & \tau_4^{(4)} &= 4.697535456150566. \end{aligned}$$

Die Kollokationsmethode mit $N = 4$ reduziert das Lösen des Optimalsteuerungsproblems (P_ν) über dem unendlichen Zeithorizont $[0, \infty)$ damit auf das Lösen eines Systems mit 5 algebraischen Gleichungen und liefert die Koeffizienten

$$\begin{aligned} a_1 &= 0.279291566888336 & a_2 &= 0.046049031742107 \\ a_3 &= 0.001362323111680 & a_4 &= 0.021798433736409 & D &= -4.825749311355116 \end{aligned}$$

woraus sich durch Einsetzen in die Ansätze aus (6.11) die folgenden Näherungslösungen

$$\begin{aligned} y_4(t) &= -4.825749311355116 + 2 \cdot t + 0.279291566888336 \cdot S_1(t) \\ &\quad + 0.046049031742107 \cdot S_2(t) + 0.001362323111680 \cdot S_3(t) \\ &\quad + 0.021798433736409 \cdot S_4(t) \\ y_4'(t) &= x_0 + 0.279291566888336 \cdot T_1(t) \\ &\quad + 0.046049031742107 \cdot T_2(t) + 0.001362323111680 \cdot T_3(t) \\ &\quad + 0.021798433736409 \cdot T_4(t) \end{aligned}$$

mit

$$\begin{aligned} S_1(t) &= -t^2, & S_3(t) &= -2t^2 + t^3 - \frac{1}{3}t^3, \\ S_2(t) &= -2t^2 + \frac{4}{3}t^3, & S_4(t) &= -4t^2 + 4t^3 - \frac{4}{3}t^4 + \frac{2}{15}t^5 \end{aligned}$$

und

$$\begin{aligned} T_1(t) &= -2t, & T_3(t) &= -2t - \frac{4}{3}t^3, \\ T_2(t) &= -4t + 4t^2, & T_4(t) &= -12t + 16t^2 - \frac{16}{3}t^3 + \frac{2}{3}t^4 \end{aligned}$$

ergeben. Mit Hilfe der so ermittelten approximierten Lösungen erhalten wir durch die Setzung

$$q(t) = p(t) = y'(t)$$

die zugehörige approximierte duale Steuerung q_4 sowie die duale Adjungierte p_4 . Mittels dieser erhalten wir schließlich gemäß Satz 22 den Zusammenhang zu den approximierten Lösungen¹¹ x_4, u_4 durch

$$x_4(t) = q_4(t)e^{-2t} \text{ und } u_4(t) = \text{Proj}_{[-6,6]}(y_4(t)) = y_4(t).$$

Durch y_4 und y_4' berechnen wir den approximierten Zielfunktionswert mit Hilfe der Gaußschen Quadraturformel und erhalten

$$G(y_4, y_4') = 4.8284741144414159483497631.$$

Mit steigender Anzahl an Kollokationspunkten wird eine Verbesserung der Approximation erreicht. Dies lässt sich insbesondere an den Zielfunktionswerten erkennen, wie der folgenden Tabelle zu entnehmen ist:

N	$G(y_N, y_N')$	$ G(y_N, y_N') - J(x_*, u_*) $	$\frac{ G(y_N, y_N') - J(x_*, u_*) }{ J(x_*, u_*) }$
2	4.8548387096774197701	0.02641158493122992324	0.00547001834114215184
4	4.8284741144414159483	$4.6989695226101 \cdot 10^{-5}$	$9.7318845272147 \cdot 10^{-6}$
6	4.8284271806892761211	$5.5943086274313 \cdot 10^{-8}$	$1.1586192527914 \cdot 10^{-8}$
8	4.8284271248040951480	$5.790530114249 \cdot 10^{-11}$	$1.199258053326 \cdot 10^{-11}$
10	4.8284271247462459524	$5.610562390128 \cdot 10^{-14}$	$1.161985517265 \cdot 10^{-14}$
12	4.8284271247461893477	$4.990652451975 \cdot 10^{-16}$	$1.033597965349 \cdot 10^{-16}$

Tabelle 6.1: Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$.

¹¹vgl. Satz 22, S. 104

Im Vergleich dazu konnte in [Bur19] durch die Verwendung der indirekten Pseudospektralmethode eine Approximationsgenauigkeit für die Zielfunktionalwerte anhand des absoluten Fehlers für $N = 8$ Kollokationspunkten mit $3.350045470362772 \cdot 10^{-6}$, für $N = 10$ Kollokationspunkten mit $1.93537497139573 \cdot 10^{-8}$ und für $N = 12$ mit $9.075 \cdot 10^{-11}$ nachgewiesen werden. Für eine Anzahl N an Kollokationspunkten mit $N < 8$ sowie eine Betrachtung der relativen Fehler fand in der zitierten Arbeit keine Auswertung statt. Die in Tabelle 6.1 beschriebenen Werte lassen jedoch eine schnellere Konvergenz der Zielfunktionalwerte unter Verwendung der modifizierten, direkten Pseudospektralmethode vermuten. Für $N = 2, 4, 8, 10, 12$ sind die Näherungslösungen der approximierten Funktionen in den Abbildungen 6.2-6.4 grafisch in einem direkten Vergleich zu den exakten Lösungen dargestellt.

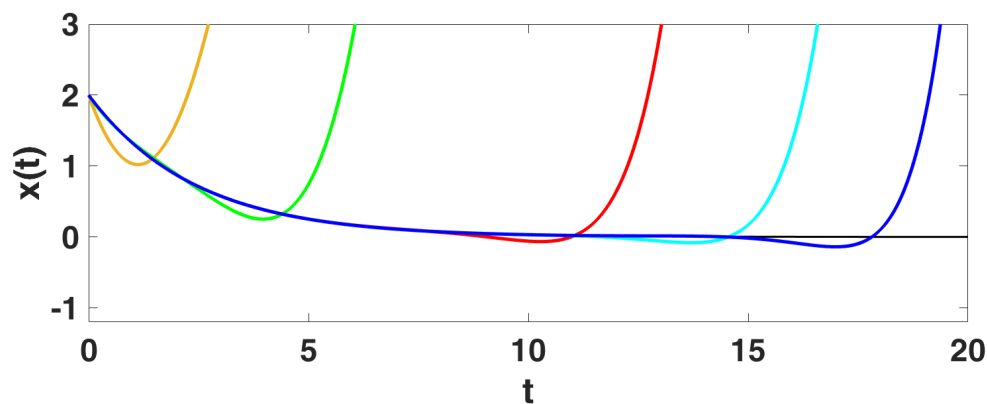


Abbildung 6.2: Näherungslösung $x_N(t)$ des primalen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $x_*(t)$ (—)

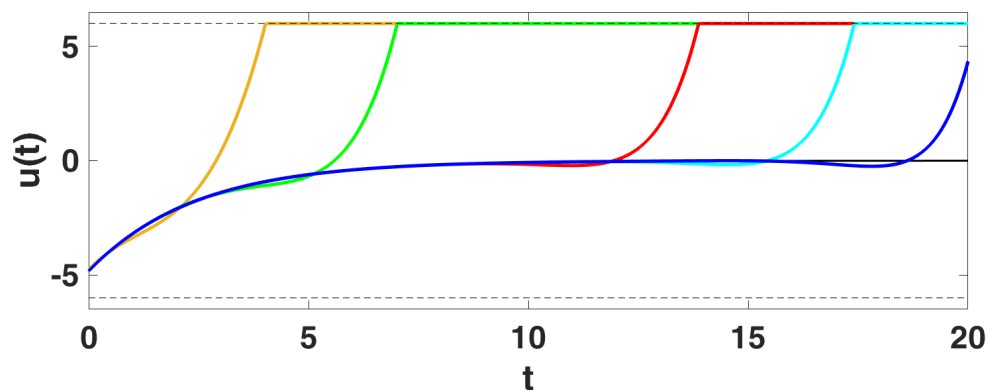


Abbildung 6.3: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—)

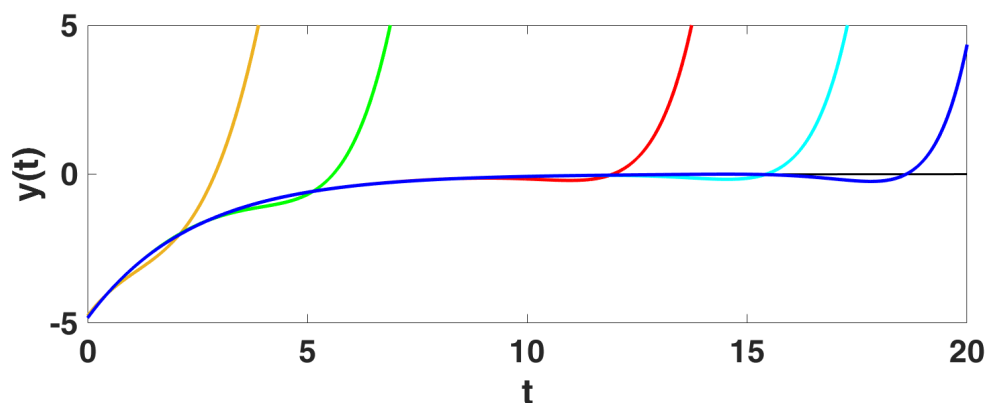


Abbildung 6.4: Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—)

Die jeweils letzten Kollokationspunkte sind durch

$$\tau_2^{(2)} = 1.707106781186547 \quad \tau_4^{(4)} = 4.697535456150566$$

$$\tau_8^{(8)} = 11.431565868444633 \quad \tau_{10}^{(10)} = 14.960348506136945 \quad \tau_{12}^{(12)} = 18.549560522233460$$

festgelegt. Die Abbildungen verdeutlichen, dass sich die approximierten Lösungen mit wachsendem N auf einem vergrößerten Intervall den exakten Lösungen annähern. Bei der Interpretation der Abbildungen ist allerdings auch zu beachten, dass die Genauigkeit der approximierten Lösungen in Abhängigkeit zur Anzahl N der verwendeten Kollokationspunkte steht. Daher ist die Interpretation der approximierten Lösung lediglich in einem von der Anzahl N der maximalen Kollokationspunkte abhängigen Intervall durchzuführen. Die Abbildungen 6.2 bis 6.4 verdeutlichen dies durch den Vergleich zur exakten Lösung. Demnach können wir bereits für eine Anzahl an $N = 10$ Kollokationspunkten eine gute Approximation für das Intervall $[0, 10]$ erkennen. Für eine Auswertung über das Intervall hinaus ist eine Anwendung des Verfahrens mit einer höheren Anzahl an Kollokationspunkten notwendig.

So suggeriert Abbildung 6.3 die Existenz eines Schaltzeitpunktes. Der Vergleich zur analytischen Lösung verdeutlicht jedoch, dass ein derartiger Schaltzeitpunkt für diese Konstellation an Anfangswerten und Steuerungsbeschränkungen nicht existiert. Die Struktur der primalen Steuerungsfunktion u ergibt sich dabei aus der Beziehung¹²

$$u_*(t) = \text{Proj}_{[-6,6]}(y_*(t)).$$

Für die approximierten Lösungen gilt damit

$$u_N(t) = \text{Proj}_{[-6,6]}(y_N(t)).$$

Der Approximationsfehler des dualen Zustandes im Intervall $(10, \infty)$ vererbt sich somit an die approximierte Steuerungsfunktion. Für das Intervall $[0, 10]$ liefert das Verfahren jedoch bereits für diese geringe Anzahl an Kollokationspunkten eine gute Annäherung. Für die Abschätzung des Fehlers der approximierten Funktionen mit N Kollokationspunkten können wir uns auf das folgende Resultat (über die Konvergenz im quadratischen Mittel) beziehen:

¹²Das dual-basierte Verfahren berechnet lediglich die duale Variable y . Alle weiteren Funktionen erhalten wir aus den Zusammenhängen aus Satz 23 S. 104 sowie durch die Setzung $q(t) = y'(t) + \hat{A}^T(t)y(t)$ (vgl. Relation 4.2 S. 73).

Lemma 45 ^[13]. Für $r \in \mathbb{N}$ sei $f \in W_2^r(\mathbb{R}_{++}, \nu_{0,\rho})$, $f' \in W_2^{r-1}(\mathbb{R}_{++}, \nu_{1,\rho})$, $f(0) = 0$ und $N \geq r$. Dann gilt

$$\left\| \overset{\circ}{P}_{N,\rho} f - f \right\|_{2,\nu_{0,\rho}} \leq c(\rho N)^{(1-r)/2} \left\| \frac{d^r}{dt^r} f \right\|_{2,\nu_{r-1,\rho}}$$

mit einer Konstanten $c > 0$.

Gemäß der obigen Aussagen bzgl. der Normkonvergenz minimiert das Verfahren den Wert

$$\int_0^\infty |f_N(t) - f_*(t)|^2 e^{-2t} dt = \int_0^\infty |f_N(t)e^{-t} - f_*(t)e^{-t}|^2 dt.$$

Wir interessieren uns daher für ein globales Fehlermaß, gemessen im gewichteten $L_{2,e^{-2t}}$ -Sinn. In den Abbildungen ^[6.5] bis ^[6.7] stellen wir dafür nun die Funktionen $x_*(t)e^{-t}$, $u_*(t)e^{-t}$, $y_*(t)e^{-t}$ mit den durch die e -Funktion gewichteten approximierten Lösungen $x_N(t)e^{-t}$, $u_N(t)e^{-t}$, $y_N(t)e^{-t}$ gegenüber.

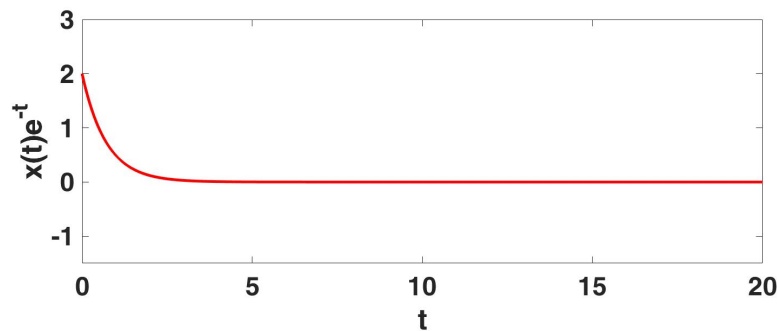


Abbildung 6.5: Näherungslösung $x_N(t)e^{-t}$ des primalen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $x_*(t)e^{-t}$ (—)

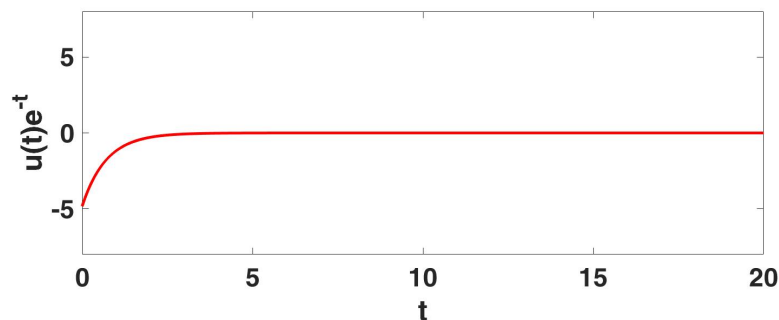


Abbildung 6.6: Näherungslösung $u_N(t)e^{-t}$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)e^{-t}$ (—)

¹³vgl. ^[XyHf13], S. 1119, Lemma 2.4 sowie ^[Bir19], S. 140, Lemma 5.7; dabei bezeichnet $\overset{\circ}{P}_{N,\rho} f$ die orthogonale Projektion $\overset{\circ}{P}_{N,\rho} : L_2(\mathbb{R}_{++}, \nu_{0,\rho}) \rightarrow \overset{\circ}{P}_N(\mathbb{R}_{++})$ der Form $\overset{\circ}{P}_{N,\rho} f(t) = \sum_{k=0}^N \hat{f}_k L^{(\omega,\rho)_k}(t)$; $\overset{\circ}{P}_N(\mathbb{R}_{++})$ ist die Menge aller Polynome auf \mathbb{R}_{++} mit Maximalgrad N und Nullrandwerten

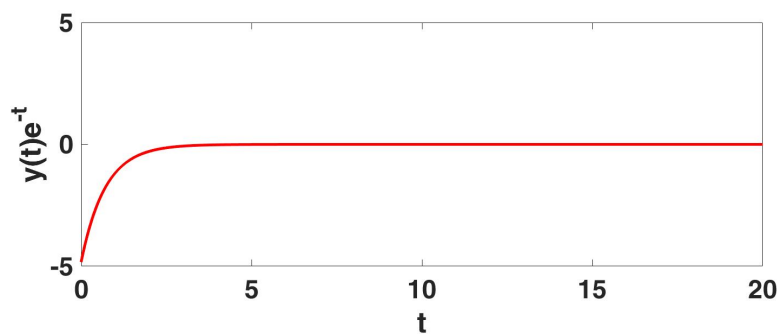


Abbildung 6.7: Näherungslösung $y_N(t)e^{-t}$ des dualen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $y_*(t)e^{-t}$ (—)

Die obigen Abbildungen deuten auf eine Normkonvergenz des Verfahrens hin. So sind hier bereits für $N = 8$ Kollokationspunkte exakte und approximierte Darstellungen optisch nicht mehr voneinander zu unterscheiden. Die Untersuchung des gewichteten $L_{2,e^{-2t}}$ -Fehlers unterstreicht diese Beobachtung. In Tabelle 6.2 sind die $L_{2,e^{-2t}}$ -Fehler der Näherungslösungen für $N = 8, 10, 11$ dargestellt.

N	$\ x_* - x_N\ _{2,e^{-2t}}$	$\ y_* - y_N\ _{2,e^{-2t}}$	$\ \theta_* - \theta_N\ _{2,e^{-2t}}$
8	$1.21750928042905 \cdot 10^{-11}$	$9.48201264933292 \cdot 10^{-12}$	$2.84632085426970 \cdot 10^{-12}$
10	$2.12910380624603 \cdot 10^{-14}$	$1.02166389413606 \cdot 10^{-14}$	$5.51710092715635 \cdot 10^{-16}$
11	$3.74515441157253 \cdot 10^{-15}$	$3.77423366556112 \cdot 10^{-16}$	$1.24359504132737 \cdot 10^{-17}$

Tabelle 6.2: $L_{2,e^{-2t}}$ -Fehler des primalen Zustandes, des dualen Zustandes und der Funktion θ für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$.

Offenbar werden die gewichteten $L_{2,e^{-2t}}$ -Fehler mit zunehmendem N kleiner. Insbesondere haben wir in diesem Beispiel bereits eine Approximationsgüte erreicht, bei welcher der gewichtete Fehler des primalen Zustandes bereits bei einer Größenordnung von $3.8 \cdot 10^{-15}$ liegt und sich der Fehler des dualen Zustandes in einer Größenordnung von $3.8 \cdot 10^{-16}$ bewegt. Der Fehler der Funktion

$$\theta(t, y(t)) = -\frac{1}{2}\sigma^2(t, y(t)) + \sigma(t, y(t))y(t)$$

beläuft sich für $N = 11$ Kollokationspunkte sogar auf etwa $1.2 \cdot 10^{-17}$. Aufgrund der quadratischen Struktur des Zielfunktional

$$G(y, q) := \frac{1}{2} \int_0^\infty q^2(t) e^{-2t} dt + \int_0^\infty \theta^2(t, y(t)) e^{-2t} dt + \tilde{x}_0^T y(0)$$

spiegelt sich dieses Verhalten ebenfalls in der bereits untersuchten Genauigkeit des Zielfunktionalwertes wieder¹⁴. Dies liegt unter anderem auch an der Genauigkeit des Anfangswertes der approximierten Funktion $y^{(N)}(0)$ im Vergleich zu dem Anfangswert der exakten Lösung $y_*(0) = -2(1 + \sqrt{2}) \approx -4.828427124746190$, welcher sowohl auf primaler als auch auf dualer Seite a priori nicht bekannt ist. Die Anfangswerte der dualen Zustandsfunktion

¹⁴dabei besteht der Zusammenhang $q_*(t) = y'_*(t) = x_*(t)$

ergeben sich dabei aus dem zu bestimmenden Koeffizienten D . Die nachfolgende Tabelle 6.3 veranschaulicht die Approximationsgüte der so bestimmten Anfangswerte.

N	$y_N(0)$	$ y_N(0) - y_*(0) $	$\frac{ y_N(0) - y_*(0) }{ y_*(0) }$
2	-4.758064525691831	0.070362599054359	0.014572571406068
4	-4.825749347032190	0.002677777714000	$5.545859230796382 \cdot 10^{-4}$
6	-4.828356767969400	$7.035677678945973 \cdot 10^{-5}$	$1.457136557552540 \cdot 10^{-5}$
8	-4.828425283592586	$1.841153603798773 \cdot 10^{-6}$	$3.813153965527759 \cdot 10^{-7}$
10	-4.828427084354846	$4.039134360311891 \cdot 10^{-8}$	8.36532116144180510^{-9}
12	-4.828427114098449	$1.064774046710681 \cdot 10^{-8}$	$2.205219255052239 \cdot 10^{-9}$

Tabelle 6.3: Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y^*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 2$ und $\mathcal{U} = [-6, 6]$.

Im Vergleich dazu konnte in [Bur19] durch die Verwendung der indirekten Pseudospektralmethode eine Approximationsgenauigkeit für die Anfangswerte der approximierten Funktion y_N anhand des absoluten Fehlers bei $N = 8$ Kollokationspunkten mit $8.523485028888445 \cdot 10^{-9}$, bei $N = 10$ Kollokationspunkten mit $1.913491587401950 \cdot 10^{-11}$ und bei $N = 12$ Kollokationspunkten mit $2.753353101070388 \cdot 10^{-14}$ nachgewiesen werden. Für eine kleinere Anzahl an Kollokationspunkten fand keine Auswertung statt, jedoch lässt sich für diesen Fall eine höhere Genauigkeit der Approximation unter Verwendung der in [Bur19] behandelten Methode vermuten.

Betrachtung für $x_0 = 1$ und $\mathcal{U} = \left[-\frac{11}{5}, \frac{11}{5}\right]$

Wir betrachten nun die Konstellation eines Anfangswertes von $x_0 = 1$ und dem Steuerbereich $\mathcal{U} = \left[-\frac{11}{5}, \frac{11}{5}\right]$. Mit dieser Setzung sind die Voraussetzungen aus Folgerung 5 erfüllt. Damit ist die Existenz einer optimalen Lösung gesichert. Zudem ist für diese Konstellation die Bedingung

$$x_0 \in \left((-1 + \sqrt{2})\bar{u}, \frac{\bar{u}}{2} \right) \approx \left(0.9112698372, \frac{11}{10} \right)$$

erfüllt. Demnach können wir die optimalen Lösungen (6.5)-(6.6) aus Abschnitt 6.1.4 als Referenz für die numerischen Lösungen zugrunde legen.

Der zugehörige Zielfunktionalwert berechnet sich damit zu

$$J(x_*, u_*) = G(y_*, q_*) \approx 1.209503362707199.$$

Wir diskutieren nun die numerischen Ergebnisse für eine unterschiedliche Anzahl N an Kollokationspunkten und beginnen wieder mit der Betrachtung der Zielfunktionalwerte.

N	$G(y_N, y'_N)$	$ G(y_N, y'_N) - J(x_*, u_*) $	$\frac{ G(y_N, y'_N) - J(x_*, u_*) }{ J(x_*, u_*) }$
2	1.213709677419354	0.0042063147121556070976	0.003477720560231205719
4	1.208681545382406	$8.218173247924 \cdot 10^{-4}$	$6.794667548116 \cdot 10^{-4}$
6	1.209803403759883	$3.000410526846 \cdot 10^{-4}$	$2.480696308384 \cdot 10^{-4}$
8	1.209863638781235	$3.602760740370 \cdot 10^{-4}$	$2.978710809291 \cdot 10^{-4}$
10	1.209689745167206	$1.863824600071 \cdot 10^{-4}$	$1.540983396606 \cdot 10^{-4}$
12	1.209477017099520	$2.634560767838 \cdot 10^{-5}$	$2.178216984813 \cdot 10^{-5}$

Tabelle 6.4: Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = [-\frac{11}{5}, \frac{11}{5}]$.

In [Bur19] konnte – für die gleiche Wahl an Anfangswerten und Steuerungsbeschränkungen – durch die Verwendung der indirekten Pseudospektralmethode eine Approximationsgenauigkeit für die Zielfunktionalwerte für $N = 8$ Kollokationspunkten mit $6.454879291 \cdot 10^{-4}$ und für $N = 9$ Kollokationspunkten mit $3.08723447210 \cdot 10^{-4}$ nachgewiesen werden. Für eine höhere Anzahl N an Kollokationspunkten fand in der zitierten Arbeit keine Auswertung statt. Die in Tabelle 6.4 beschriebenen Werte lassen jedoch eine leicht bessere Konvergenz der Zielfunktionalwerte unter Verwendung der modifizierten, direkten Pseudospektralmethode vermuten.

Für $N = 2, 4, 8, 10, 12$ sind die Näherungslösungen der approximierten Funktionen in den Abbildungen 6.8–6.10 grafisch in einem direkten Vergleich zu den exakten Lösungen dargestellt.

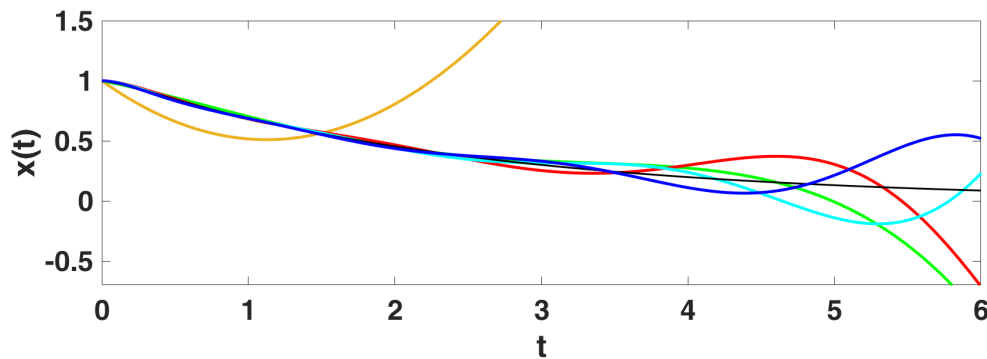


Abbildung 6.8: Näherungslösung $x_N(t)$ des primalen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $x_*(t)$ (—)

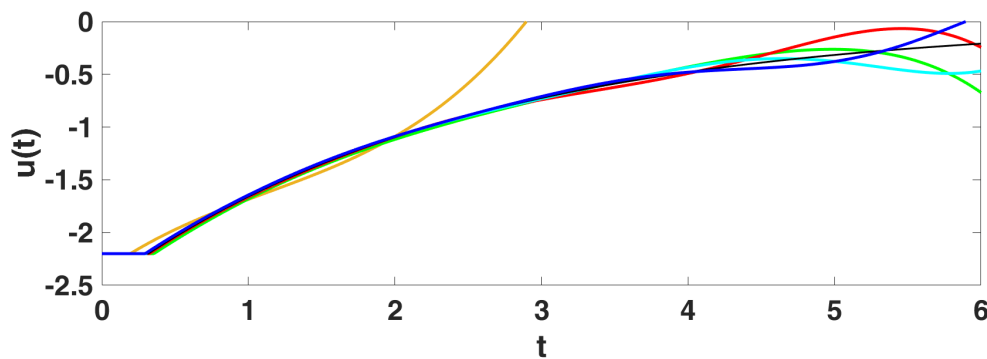


Abbildung 6.9: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—)

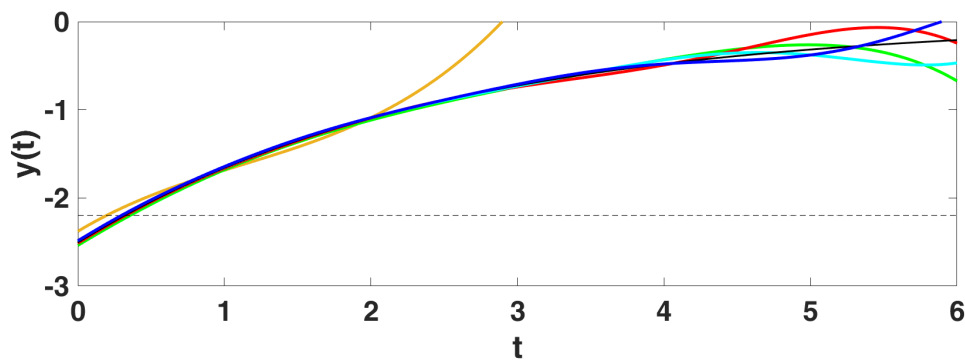


Abbildung 6.10: Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 2$ (—), $N = 4$ (—), $N = 8$ (—), $N = 10$ (—) und $N = 12$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—)

Die Abbildungen legen auch hier die Konvergenz der Folge der approximierten Lösungen nahe. Zudem wollen wir wieder einen Blick auf die Anfangswerte der dualen Zustandsfunktion legen. Die nachfolgende Tabelle 6.5 veranschaulicht die Approximationsgüte der so bestimmten Anfangswerte, indem sie die numerisch ermittelten Anfangswerte der analytischen Lösung

$$y_*(0) \approx -2.504966372928010$$

gegenüberstellt. Eine Betrachtung dieser Werte wurde in [Bur19] nicht durchgeführt. Die benannte Quelle kann an dieser Stelle daher nicht als Vergleich herangezogen werden.

N	$y_N(0)$	$ y_N(0) - y_*(0) $	$\frac{ y_N(0) - y_*(0) }{ y_*(0) }$
2	-2.379032260995189	0.125934111932820	0.050273773450147
4	-2.538693011276094	0.033726638348084	0.013463908622718
6	-2.533851991050683	0.028885618122673	0.011531339675793
8	-2.513573002580870	0.008606629652860	0.003435826423011
10	-2.497683483155568	0.007282889772442	0.002907380255140
28	-2.498991842778801	0.005974530149209	0.002385073992920

Tabelle 6.5: Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y_*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = [-\frac{11}{5}, \frac{11}{5}]$.

Interessant ist zudem ein Blick auf die erhaltene Schaltstruktur der optimalen Steuerungsfunktion u_* . Die theoretischen Ergebnisse dieser Arbeit ließen nun eine Formulierung notwendiger Optimalitätsbedingungen unter schwächeren Voraussetzungen als in Satz 10 zu¹⁵ und erlaubten die analytische Diskussion des vorliegenden Falles. Entgegen der Annahmen in [Bur19], in welchen die numerischen Ergebnisse eine Schaltstruktur mit 6 Schaltzeitpunkten nahelegten, können wir auf Grundlage des Satzes 23 auf eine optimale Lösung u_* mit genau einem Schaltzeitpunkt $\bar{t} = \frac{1}{2} \ln \left(\frac{-3\bar{u} + 2\sqrt{2\bar{u}}}{2x_0 - \bar{u}} \right) \approx 0.3175740526$ schließen.

Die Annahmen in [Bur19] basierten dabei auf der Betrachtung der – mittels der indirekten Pseudospektralmethode ermittelten – approximierten Lösung¹⁶ u_N mit $N = 9$. Für eine Anzahl von $N = 8$ Kollokationspunkten erhalten wir über die Verwendung der modifizierten, direkten Pseudospektralmethode eine ähnliche Abbildung (vgl. Abbildung 6.11).

¹⁵vgl. hierzu Satz 23, S. 104 und Satz 10, S. 41

¹⁶vgl. [Bur19], Abbildung 6.20, S. 202

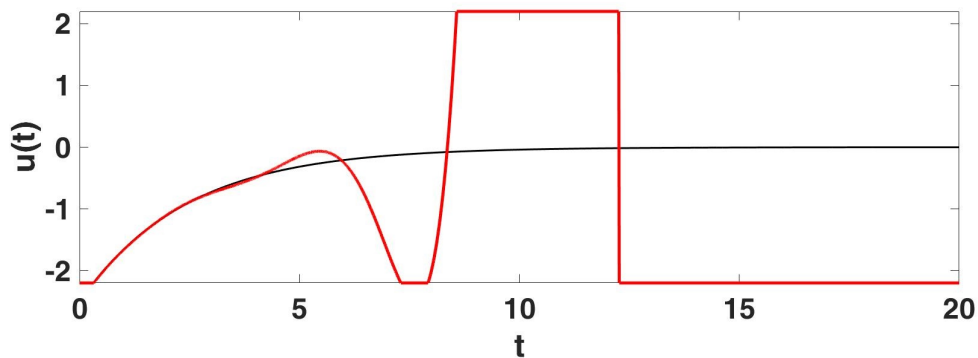


Abbildung 6.11: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—) auf $[0, 20]$

Die Abbildung der approximierten Steuerungsfunktion lässt sich wieder auf das von der Anzahl N an Kollokationspunkten abhängige Konvergenzverhalten zurückführen. Zudem ist zu berücksichtigen, dass das hier verwendete Verfahren zu keinem Zeitpunkt eine Approximation der primalen Funktionen berechnet. Diese erhalten wir lediglich aus dem in Satz 23 dargestellten Zusammenhang $x_*(t) = q_*(t)$ und $u_*(t) = \text{Proj}_{\mathcal{U}}(y_*(t))$ durch die approximierten Lösungen in der Setzung

$$x_N(t) = q_N(t) \quad \text{und} \quad u_N(t) = \text{Proj}_{[-\frac{11}{5}, \frac{11}{5}]}(y_N(t)). \quad (6.12)$$

Unter Berücksichtigung der Beziehung (6.12) verdeutlicht die Veranschaulichung der Approximation des zugehörigen dualen Zustandes in Abbildung 6.12 dieses Verhalten. Eine Interpretation der Ergebnisse einzelner Komponenten sollte daher nicht losgelöst, sondern unter Einbeziehung der Ergebnisse aller Komponenten geschehen.

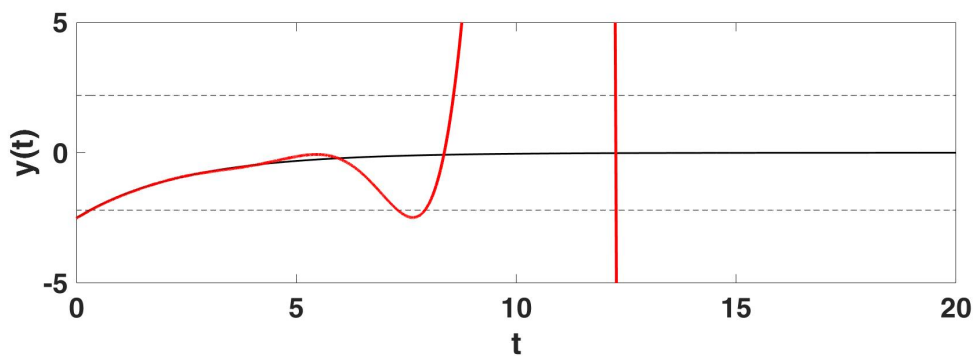


Abbildung 6.12: Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 8$ (—) im Vergleich zur exakten Lösung $y_*(t)$ (—) auf $[0, 20]$

Abschließend widmen wir uns der Untersuchung des Schaltzeitpunktes \bar{t} . Bereits die Betrachtung der Abbildung 6.11 lässt eine gute Annäherung an den Schaltzeitpunkt vermuten. Abbildung 6.13 veranschaulicht die dortige Approximation in der Nähe des Umschaltzeitpunktes und verdeutlicht diesen Aspekt.

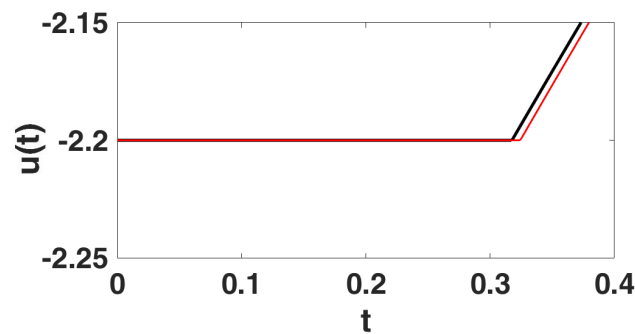


Abbildung 6.13: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—) im Vergleich zur exakten Lösung $u_*(t)$ (—) in der Nähe des Umschaltzeitpunktes \bar{t}

Die Tabelle 6.6 verdeutlicht die Approximationsgüte des Schaltzeitpunktes für eine unterschiedliche Anzahl an Kollokationspunkten.

N	\bar{t}_N	$ \bar{t}_N - \bar{t} $	$\frac{ \bar{t}_N - \bar{t} }{ \bar{t} }$
6	0.345779	0.0282049474	0.0888137764
8	0.324249	0.0066749474	0.0210185644
10	0.307901	0.0096730526	0.0304591933
18	0.326595	0.0090209506	0.0284058179

Tabelle 6.6: Vergleich des numerisch ermittelten Umschaltzeitpunktes \bar{t}_N zur exakten Lösung \bar{t} für eine unterschiedliche Anzahl von Stützstellen für $x_0 = 1$ und $\mathcal{U} = [-\frac{11}{5}, \frac{11}{5}]$.

6.2 Das Lotka-Volterra Modell

Anfang des 20. Jahrhunderts beschäftigten sich der Chemiker und Versicherungsmathematiker Alfred J. Lotka sowie der Mathematiker und Physiker Vito Volterra unabhängig voneinander mit dem Verhalten zweier konkurrierender biologischer Systeme. Die Erkenntnisse dieser Bemühungen sind seither als sogenannte „Lotka-Volterra-Gleichungen“ bekannt und stellen bis heute wesentliche Grundpfeiler der theoretischen Biologie, und dabei insbesondere der Epidemiologie, aber auch weiteren Anwendungsfeldern, wie zum Beispiel den Wirtschaftswissenschaften, dar.

Sowohl Lotka als auch Volterra beschrieben das System konkurrierender Spezies für ein beliebiges Zeitintervall \mathcal{I} durch

$$\begin{aligned}x_1'(t) &= l_1 x_1(t) - \alpha_1 x_1(t) x_2(t) \\x_2'(t) &= -l_2 x_2(t) + \alpha_2 x_1(t) x_2(t),\end{aligned}$$

mit

- x_1 Populationsgröße der Beute
- x_2 Populationsgröße der Räuber
- l_1 Reproduktionsrate der Beute
- l_2 Sterberate der Räuber wenn keine Beute vorhanden ist
- α_1 Fressrate der Räuber pro Beutelebewesen
- α_2 Reproduktionsrate der Räuber pro Beutelebewesen,

wobei l_1, l_2, α_1 und α_2 positive Konstanten seien.

In ihren Arbeiten stießen sie im Zuge der Untersuchungen von sogenannten *Gleichgewichtspunkten* bereits auf ein periodisches Verhalten der beiden konkurrierenden Populationen. Die Gleichgewichtspunkte eines dynamischen Systems bestimmen sich aus der Annahme, dass das System im Gleichgewicht (oder auch Equilibrium) stabil ist und somit auch für $\xi_1, \xi_2 \neq 0$ keine Änderung stattfindet. Aus diesen Überlegungen leitet sich die folgende Setzung für die Bestimmung des Gleichgewichtspunktes $(x_1^G, x_2^G) = (\xi_1, \xi_2)$ ab:

$$\begin{aligned} 0 &= \xi_1 [l_1 - \alpha_1 \xi_2], \\ 0 &= \xi_2 [-l_2 + \alpha_2 \xi_1], \end{aligned}$$

welche den 17 eindeutigen Gleichgewichtspunkt $(x_1^G, x_2^G) = (\frac{l_2}{\alpha_2}, \frac{l_1}{\alpha_1})$ liefert. Wir wollen im Folgenden die Stabilitätseigenschaften dieses Gleichgewichtspunktes untersuchen.

6.2.1 Stabilitätseigenschaften des Gleichgewichtspunktes

Für die Untersuchung der Stabilitätseigenschaften des Gleichgewichtspunktes $(x_1^G, x_2^G) = (\frac{l_2}{\alpha_2}, \frac{l_1}{\alpha_1})$ betrachten wir die Störung des Gleichgewichtspunktes um ein $\delta := (\delta_1, \delta_2)$. Dafür setzen wir

$$\delta_1(t) := x_1(t) - \frac{l_2}{\alpha_2} \quad \text{und} \quad \delta_2(t) := x_2(t) - \frac{l_1}{\alpha_1}.$$

Diese Setzung liefert das System (δ) :

$$\begin{aligned} \delta_1'(t) &= l_1(\delta_1(t) + \frac{l_2}{\alpha_2}) - \alpha_1(\delta_1(t) + \frac{l_2}{\alpha_2})(\delta_2(t) + \frac{l_1}{\alpha_1}) \\ \delta_2'(t) &= -l_2(\delta_2(t) + \frac{l_1}{\alpha_1}) + \alpha_2(\delta_1(t) + \frac{l_2}{\alpha_2})(\delta_2(t) + \frac{l_1}{\alpha_1}). \end{aligned}$$

Die Linearisierung des Systems (δ) um den Gleichgewichtspunkt in der Form

$$\delta'(t) = A\delta(t),$$

liefert eine Darstellung mit der *marginal stabilen* Matrix 18

$$A = \begin{pmatrix} 0 & -\frac{l_2\alpha_1}{\alpha_2} \\ \frac{l_1\alpha_2}{\alpha_1} & 0 \end{pmatrix},$$

mit den zugehörigen Eigenwerten

$$\lambda_{1,2} = \pm i\sqrt{l_1 l_2}.$$

Demnach ist der Gleichgewichtspunkt weder instabil noch asymptotisch stabil. Insbesondere ist der Gleichgewichtspunkt damit auch nicht exponentiell stabil.

¹⁷abgesehen von der trivialen Lösung $(0, 0)$

¹⁸Aus der Eigenschaft $Re(\lambda_{1,2}) = 0$ schließen wir auf die marginale Stabilität der Matrix A .

6.2.2 Formulierung der primalen Aufgabe

Ziel unserer Betrachtungen ist die Beantwortung der Frage, wie wir durch ein Eingreifen in das System das marginal stabile Verhalten in ein exponentiell stabiles System überführen können. Dafür gehen wir im Folgenden davon aus, dass wir von Außen in das System eingreifen können, beispielsweise durch Entfernen oder durch Hinzufügen eines Anteils der/zur entsprechenden Spezies. Dieses Eingreifen modellieren wir mit Hilfe einer Steuerungsfunktion (wobei negative Werte der Steuerungsfunktion das Entfernen, positive Werte der Steuerungsfunktion das Hinzufügen beschreiben) und führen dafür auf der Seite der Beute die Steuerungsfunktion $u_1 \in [\underline{u}_1, \bar{u}_1]$ sowie auf der Seite der Räuber die Steuerungsfunktion $u_2 \in [\underline{u}_2, \bar{u}_2]$, mit $0 \in \text{int} [\underline{u}_i, \bar{u}_i]$ für $i = 1, 2$, ein¹⁹. Dies führt uns zur Betrachtung des folgenden Systems

$$\begin{aligned} x_1'(t) &= l_1 x_1(t) - \alpha_1 x_1(t)x_2(t) + u_1(t) \\ x_2'(t) &= -l_2 x_2(t) + \alpha_2 x_1(t)x_2(t) + u_2(t) \end{aligned} \quad \text{mit } u_i \in [\underline{u}_i, \bar{u}_i], \quad \forall t \in \mathcal{I}.$$

Die Taylorpolynome 1. Grades um den Gleichgewichtspunkt liefern schließlich, zusammen mit der eingeführten Steuerung, die linearisierten Gleichungen der Form

$$\begin{aligned} x_1'(t) &= -\frac{\alpha_1 l_2}{\alpha_2} \left[x_2(t) - \frac{l_1}{\alpha_1} \right] + u_1(t), \\ x_2'(t) &= \frac{l_1 \alpha_2}{\alpha_1} \left[x_1(t) - \frac{l_2}{\alpha_2} \right] + u_2(t). \end{aligned}$$

Mittels der linearen Transformation

$$\tilde{x}_1(t) := x_1(t) - x_1^G \quad \text{sowie} \quad \tilde{x}_2(t) := x_2(t) - x_2^G$$

überführen wir obiges System in

$$\begin{aligned} \tilde{x}_1'(t) &= -\frac{\alpha_1 l_2}{\alpha_2} \tilde{x}_2(t) + u_1(t), \quad \tilde{x}_1(0) = x_1^0 - x_1^G \\ \tilde{x}_2'(t) &= \frac{l_1 \alpha_2}{\alpha_1} \tilde{x}_1(t) + u_2(t), \quad \tilde{x}_2(0) = x_2^0 - x_2^G. \end{aligned}$$

Die exponentielle Stabilisierung des Systems erreichen wir, gemäß der Betrachtungen aus Kapitel 1.6, zum einen durch die Wahl des Integranden $\varphi(t, \tilde{\xi}_1, \tilde{\xi}_2, v) = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + v_1^2 + v_2^2$, als auch durch die Einführung der Gewichtsfunktion $\nu(t) = e^{\beta t}$, mit $\beta > 0$. Als primale Aufgabe (P_ν) betrachten wir daher

Minimiere

$$J(\tilde{x}, u) = \frac{1}{2} \int_0^\infty \left\{ \tilde{x}_1^2(t) + \tilde{x}_2^2(t) + u_1^2(t) + u_2^2(t) \right\} e^{\beta t} dt$$

bezüglich aller

$$(\tilde{x}, u) \in W_2^{1,2}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^2(\mathbb{R}_{++}, e^{\beta t}) \quad \text{mit } \beta > 0,$$

für die gilt

$$\tilde{x}_1'(t) = -\frac{\alpha_1 l_2}{\alpha_2} \tilde{x}_2(t) + u_1(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad \tilde{x}_1(0) = x_1^0 - x_1^G, \quad (6.13)$$

$$\tilde{x}_2'(t) = \frac{l_1 \alpha_2}{\alpha_1} \tilde{x}_1(t) + u_2(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad \tilde{x}_2(0) = x_2^0 - x_2^G, \quad (6.14)$$

$$u(t) \in \mathcal{U} = [\underline{u}_1, \bar{u}_1] \times [\underline{u}_2, \bar{u}_2], \quad 0 \in \text{int } \mathcal{U} \quad (\forall) t \in \mathbb{R}_{++}. \quad (6.15)$$

¹⁹Wollen wir nur den Einfluss auf eine Population erlauben, ist auch die Einführung einer einzigen Steuerungsfunktion möglich. Diese Setzung spiegelt sich in der späteren Formulierung der Matrix B wider.

Mit den Setzungen

$$W \equiv R \equiv B(t) \equiv I_{2 \times 2} \quad \text{und} \quad A \equiv \begin{pmatrix} 0 & -\frac{\alpha_1 l_2}{\alpha_2} \\ \frac{l_1 \alpha_2}{\alpha_1} & 0 \end{pmatrix} \quad (6.16)$$

sowie $C \equiv w \equiv r(t) \equiv 0$ können wir dieses Problem der Klasse (P_ν) zuordnen und verifizieren leicht die Gültigkeit der Bedingungen (V1)-(V3)^[20]. Den Schluss auf die Existenz einer optimalen Lösung für das vorliegende Problem mit echter Gewichtsfunktion erhalten wir, gemäß der Aussagen von Satz 9, damit über den noch fehlenden Nachweis der Gültigkeit der Bedingung (V4)^[21]. Auf diesen Nachweis gehen wir wieder etwas ausführlicher ein. Da unser Hauptaugenmerk jedoch der Vorstellung der Methodik gilt, möchten wir uns für die weitere Untersuchung dieses Beispiels aus Gründen der Übersichtlichkeit auf die Betrachtung des Spezialfalles $l_1 = l_2 = \alpha_1 = \alpha_2 = 1$ beschränken.

6.2.3 Untersuchung der zulässigen Menge

Für den Nachweis der Gültigkeit der Voraussetzung (V4) ist zu zeigen, dass die Menge der zulässigen Lösungen nichtleer und nicht einelementig ist. Für das vorliegende Problem, in welchem wir eine Steuerung suchen, die das linearisierte Lotka-Volterra System exponentiell stabilisiert, betrachten wir Funktionen in gewichteten Sobolev-Räumen mit echten Gewichtsfunktionen. Wir suchen daher Funktionen $(\tilde{x}, u) \in W_2^{1,2}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^2(\mathbb{R}_{++}, e^{\beta t})$ die unter Einhaltung der Steuerungsbeschränkungen dem Differentialgleichungssystem

$$\begin{aligned} \tilde{x}'_1(t) &= -\tilde{x}_2(t) + u_1(t), \\ \tilde{x}'_2(t) &= \tilde{x}_1(t) + u_2(t), \end{aligned} \quad (6.17)$$

für fast alle $t \in \mathbb{R}_{++}$ zu den gegebenen Anfangswerten

$$\tilde{x}_1(0) = x_1^0 - 1, \quad \tilde{x}_2(0) = x_2^0 - 1$$

genügen. Können wir zeigen, dass mindestens zwei zulässige Lösungen existieren, so ist die Gültigkeit der Voraussetzung (V4) gezeigt.

Es lässt sich leicht berechnen, dass sich die homogene Lösung des obigen Systems beschreiben lässt durch

$$\begin{aligned} x_1(t) &= -C_1 \sin(t) + C_2 \cos(t) \\ x_2(t) &= C_1 \cos(t) + C_2 \sin(t). \end{aligned}$$

Die homogene Lösung ist damit nur mit $C_1 = C_2 = 0$ Element des Raumes $W_2^{1,2}(\mathbb{R}_{++}, e^{\beta t})$ und somit nur zulässige Lösung falls $x_0 = (0, 0)^T$ ^[22]. Diese Beobachtung impliziert zudem, dass für die Steuerung $u_1(t) = u_2(t) \equiv 0$ kein zulässiger Kandidat existiert, falls $x_0 \neq (0, 0)^T$.

Wir betrachten nun den Fall $\hat{u}_1(t) = C_3 e^{-\rho_1 t}$ und $\hat{u}_2(t) = C_4 e^{-\rho_2 t}$ mit $\rho_1, \rho_2 > \beta/2 > 0$, $C_3 \in [\underline{u}_1, \bar{u}_1]$ und $C_4 \in [\underline{u}_2, \bar{u}_2]$. Offensichtlich handelt es sich bei den Funktionen \hat{u}_1 und \hat{u}_2 um stetige, monotone Funktionen. Damit nehmen diese Funktionen ihren betragsmäßig

²⁰siehe S. 25

²¹siehe Satz 9, S. 39

²²Dies folgt auch sofort aus der Erkenntnis, dass (x_1^G, x_2^G) kein exponentiell stabiles Gleichgewicht des ungesteuerten Systems ist.

größten Wert für $t = 0$ an. Da C_3 und C_4 in Abhängigkeit der Steuerbeschränkungen gewählt wurden, ist gewährleistet, dass diese zum Zeitpunkt $t = 0$ den Restriktionen des Steuerbereiches \mathcal{U} genügen. Den betragsmäßig kleinsten Funktionswert erreichen die Funktionen für $t \rightarrow \infty$. Die Bedingung $0 \in \text{int } \mathcal{U}$ sichert, dass die Funktionen auch in diesem Fall in den Steuerbereich abbilden, denn es gilt

$$\lim_{T \rightarrow \infty} \hat{u}_i(T) = \lim_{T \rightarrow \infty} C_j e^{-\rho_1 T} = 0^+ \quad \text{für positive } C_j \quad \text{und}$$

$$\lim_{T \rightarrow \infty} \hat{u}_i(T) = \lim_{T \rightarrow \infty} C_j e^{-\rho_1 T} = 0^- \quad \text{für negative } C_j.$$

Zudem gilt mit dieser Setzung $\hat{u} \in L^2_2(\mathbb{R}_{++}, e^{\beta t})$. Für diese Wahl der Steuerung ergibt sich als Lösung des Systems (6.17)

$$\begin{aligned} \hat{x}_1(t) &= \frac{-\rho_1}{\rho_1^2 + 1} C_3 e^{-\rho_1 t} - \frac{1}{\rho_2^2 + 1} C_4 e^{-\rho_2 t} \\ \hat{x}_2(t) &= \frac{1}{\rho_1^2 + 1} C_3 e^{-\rho_1 t} - \frac{\rho_2}{\rho_2^2 + 1} C_4 e^{-\rho_2 t}. \end{aligned}$$

Für $\rho_1 > \beta/2 > 0$ und $\rho_2 > \beta/2 > 0$ sind die Funktionen \hat{x}_1 und \hat{x}_2 Elemente des Raumes $W^{1,2}_2(\mathbb{R}_{++}, e^{\beta t})$. Für den Nachweis der Zulässigkeit der so konstruierten Schar von Paaren (\hat{x}, \hat{u}) fordern wir nun die Gültigkeit der Anfangswertbedingungen. Diese Forderung liefert die folgende Setzung der Konstanten C_3, C_4 :

$$C_3 = \frac{(\rho_1^2 + 1)(x_2^0 - \rho_2 x_1^0)}{\rho_1 \rho_2 + 1} \quad \text{und} \quad C_4 = \frac{-(\rho_1^2 + 1)(x_1^0 + \rho_1 x_2^0)}{\rho_1 \rho_2 + 1}.$$

Gilt für diese Konstanten $C_3 \in [\underline{u}_1, \bar{u}_1]$ und $C_4 \in [\underline{u}_2, \bar{u}_2]$, so sind durch (\hat{x}, \hat{u}) eine Schar von zulässigen Lösungen gegeben. Die Menge \mathcal{B} ist somit nichtleer und nicht einelementig. Folglich ist Voraussetzung (V4) erfüllt.

Folgerung 7. *Gilt für die Anfangswerte \hat{x}_1^0, \hat{x}_2^0 des Systems (6.13)-(6.14)*

$$\hat{x}_1^0 \in \left[\frac{-\bar{u}_2}{\rho_2^2 + 1} - \frac{\rho_1 \bar{u}_1}{\rho_1^2 + 1}, \frac{-\underline{u}_2}{\rho_2^2 + 1} - \frac{\rho_1 \underline{u}_1}{\rho_1^2 + 1} \right] \quad \text{und} \quad \hat{x}_2^0 \in \left[\frac{-\rho_2 \bar{u}_2}{\rho_2^2 + 1} + \frac{\underline{u}_1}{\rho_1^2 + 1}, \frac{-\rho_2 \underline{u}_2}{\rho_2^2 + 1} + \frac{\bar{u}_1}{\rho_1^2 + 1} \right],$$

dann existiert eine optimale Lösung des Problems (P_ν) .

Bemerkung 28. *Die in Folgerung 7 formulierten Ergebnisse stellen einen Zusammenhang zwischen der Zulässigkeit von Anfangswerten, Steuerungsbeschränkungen und des zu verwendenden Gewichtungsparameters β her. Dabei ist zu beachten, dass die Aussagen der Folgerung 7 nur hinreichend und nicht notwendig sind. Finden wir allerdings eine Paarung von Anfangswerten, Steuerungsbeschränkungen und eines Gewichtungsparameters β , welche die Bedingungen der Folgerung 7 erfüllen, so ist die Existenz einer Optimallösung gesichert.*

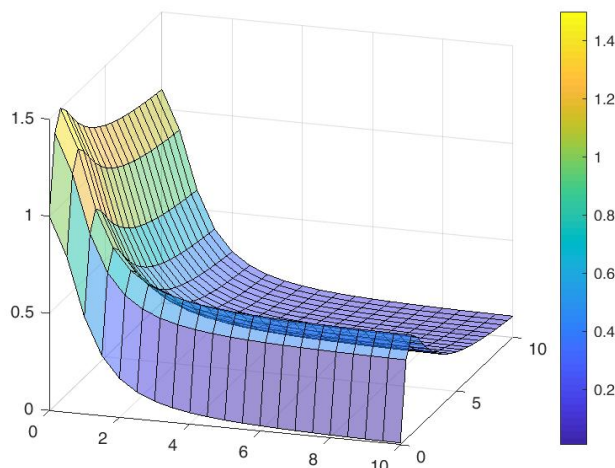


Abbildung 6.14: Mögliche Wahl der Anfangswerte im Falle normierter Steuerungsbereiche für $\rho_1, \rho_2 \in [0, 10]$.

Für $\underline{u}_1 = \underline{u}_2 = -1$ und $\bar{u}_1 = \bar{u}_2 = 1$ veranschaulicht Abbildung 6.14 die Wahl der betragsmäßig größten Anfangswerte für alle Kombinationen von $\rho_1, \rho_2 \in [0, 10]$. Offensichtlich sind die betragsmäßig größten Anfangswerte durch $\tilde{x}_1^0 = \tilde{x}_2^0 = 1.5$ gegeben. Dies gilt jedoch nur für den Fall $\rho_1 = \rho_2 = 0$, welcher durch die Bedingung $0 < \beta < 2\rho_i, i = 1, 2$ nicht zulässig ist. Damit ist die Existenz einer optimalen Lösung des Stabilisierungsproblems (P_ν) nur für $|\tilde{x}_1^0| < 1.5$ und $|\tilde{x}_2^0| < 1.5$ garantiert.

6.2.4 Übergang zur dualen Aufgabe

Gemäß der Betrachtungen aus Kapitel 4 ist die zu (P_ν) duale Aufgabe gegeben durch (D_ν) :

Minimiere

$$G(y, q) = \int_0^\infty \left\{ \frac{1}{2} (q_1^2(t) + q_2^2(t)) + \theta(t, y(t)) \right\} e^{-\beta t} dt + \tilde{x}_0^T y(0)$$

bezüglich aller

$$(y, q) \in W_2^{2,2}(\mathbb{R}_{++}, e^{-\beta t}) \times W_2^{1,2}(\mathbb{R}_{++}, e^{-\beta t}) \text{ mit } \beta > 0,$$

für die gilt

$$y_1'(t) = -y_2(t) + q_1(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad q_1(0) = \tilde{x}_1^0 \quad (6.18)$$

$$y_2'(t) = y_1(t) + q_2(t) \quad (\forall) t \in \mathbb{R}_{++}, \quad q_2(0) = \tilde{x}_2^0, \quad (6.19)$$

wobei

$$\theta(t, y(t)) = \sum_{i=1}^2 -\frac{1}{2} \sigma_i^T(t, y(t)) \sigma_i(t, y(t)) + \sigma_i^T(t, y(t)) y_i(t) \quad \text{und}$$

$$\sigma_i(t, y(t)) = \min \left\{ \max \left\{ \underline{u}_i, y_i(t) e^{-\beta t} \right\}, \bar{u}_i \right\} e^{\beta t} \quad \text{ist.}$$

Sind die Bedingungen (V1)-(V4) und (VD1)-(VD4) erfüllt und gilt zudem $\tilde{Y}_{\nu_0} \neq \emptyset$, so können wir auf Grundlage von Satz 20 auf die Existenz einer optimalen Lösung der dualen Aufgabe und auf Grundlage von Satz 23 auf notwendige und hinreichende Optimalitätsbedingungen schließen²³. Die Gültigkeit der Bedingungen (V1)-(V4) haben wir bereits im vorangegangenen Unterkapitel untersucht. Für die hier betrachtete duale Aufgabe lassen sich die Bedingungen (VD1)-(VD3) leicht verifizieren. Für den Nachweis der Gültigkeit von (VD4) genügt nun wieder die Angabe mindestens zweier Paare, welche zulässig für die duale Aufgabe sind. Diese lassen sich aufgrund der fehlenden Steuerungsbeschränkungen für die duale Steuerungsfunktion q sowie der vorliegenden Funktionenräume in Form gewichteter Sobolev-Räume mit echter Dichtefunktion leicht finden. Offensichtlich sind durch (\hat{y}, \hat{q}) und (\tilde{y}, \tilde{q}) zwei zulässige Lösungen gegeben, wobei

$$\begin{aligned}\hat{y} &= (-\tilde{x}_2^0, \tilde{x}_1^0)^T, & \hat{q} &= (\tilde{x}_1^0, \tilde{x}_2^0) \quad \text{und} \\ \tilde{y} &= (t - \tilde{x}_2^0 + 1, t + \tilde{x}_1^0 - 1)^T, & \tilde{q} &= (t + \tilde{x}_1^0, \tilde{x}_2^0 - t).\end{aligned}$$

Damit sind die Voraussetzungen (VD1)-(VD4) erfüllt. Für die gegebenen Paare können wir zudem schließen $(\hat{y}, \hat{q}) \in W_2^2(\mathbb{R}_{++}, e^{-\beta t}) \times W_2^1(\mathbb{R}_{++}, e^{-\beta t})$, mit $\hat{q}(0) = x_0$, so dass zudem $(\hat{y}, \hat{q}) \in \mathcal{B}_D$ und damit $\mathcal{B}_D \neq \emptyset$ gilt. Für den Nachweis der Gültigkeit der Bedingung $\tilde{Y}_{\nu_1} \neq \emptyset$ genügt nun die Angabe eines Paares $(y, q) \in \mathcal{B}_D$, welches der Differentialgleichung

$$\begin{aligned}y''(t) &= -\hat{W}(t) \left(\hat{W}^{-1} \right)'(t) q(t) - \left[\left(\hat{A}'(t) \right)^T y(t) + \hat{A}^T(t) y'(t) \right] \\ &\quad + \hat{W}(t) A(t) \hat{W}^{-1}(t) q(t) + \hat{W}(t) B(t) \hat{\sigma}(t, y(t)) e^{-\beta t}\end{aligned}$$

fast überall auf \mathbb{R}_{++} genügt. Die obige Differentialgleichung ergibt sich für die Setzung von

$$\hat{W}(t) = \hat{R}(t) = \begin{pmatrix} e^{\beta t} & 0 \\ 0 & e^{\beta t} \end{pmatrix}, \quad B \equiv I_{2 \times 2} \quad \text{und} \quad \hat{A} \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

zu

$$y''(t) = \begin{pmatrix} \beta & -2 \\ 2 & \beta \end{pmatrix} y'(t) + \begin{pmatrix} 1 & \beta \\ -\beta & 1 \end{pmatrix} y(t) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \hat{\sigma}(t, y(t)), \quad (6.20)$$

wobei $x_0 = y'(0) + \hat{A}^T y(0)$. Aufgrund der Komplexität des Differentialgleichungssystems soll uns nun die exemplarische Betrachtung für den Fall der festen Setzung des Gewichtungsparameters durch $\beta = 2$ genügen. Für diesen Fall betrachten wir die Funktion $y(\cdot) = (y_1(\cdot), y_2(\cdot))^T$ mit

$$y_1(t) = \left\{ \tilde{x}_2^0(1 + \sqrt{2}) \sin t - \tilde{x}_1^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t} \quad (6.21)$$

$$y_2(t) = \left\{ -\tilde{x}_1^0(1 + \sqrt{2}) \sin t - \tilde{x}_2^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t}. \quad (6.22)$$

Gilt

$$\tilde{x}_1^0 \in \left[-\frac{\bar{u}_1}{1 + \sqrt{2}}, -\frac{u_1}{1 + \sqrt{2}} \right] \quad \text{und} \quad \tilde{x}_2^0 \in \left[-\frac{\bar{u}_2}{1 + \sqrt{2}}, -\frac{u_2}{1 + \sqrt{2}} \right],$$

so gilt für die Funktion y aus (6.21)-(6.22) offenbar $ye^{-2t} \in [-\bar{u}, \bar{u}]$ und damit

$$\hat{\sigma}(t, y(t)) = \text{Proj}_{[-\bar{u}, \bar{u}]} \left(\hat{R}^{-1}(t) (-\hat{r}(t) + B^T(t)y(t)) \right) e^{2t} = \text{Proj}_{[-\bar{u}, \bar{u}]} \left(y(t) e^{-2t} \right) e^{2t} = y(t).$$

²³für (V1)-(V4) siehe S. 25, für (VD1)-(VD4) siehe S. 76, für Satz 20 siehe S. 101, für Satz 23 siehe S. 104

Zudem gilt für die Funktionen aus (6.21)-(6.22)

$$\begin{aligned} y_1'(t) &= \left\{ \left[(1 + \sqrt{2})\tilde{x}_1^0 - \tilde{x}_2^0 \right] \sin t + \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t}, \\ y_1''(t) &= \left\{ \left[-2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \sin t + \left[2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t}, \\ y_2'(t) &= \left\{ \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \sin t + \left[-(1 + \sqrt{2})\tilde{x}_1^0 + \tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t}, \\ y_2''(t) &= \left\{ \left[2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \sin t + \left[2\tilde{x}_1^0 + 2\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t}. \end{aligned}$$

Damit erhalten wir leicht die Gültigkeit der Bedingung $x_0 = y'(0) + \hat{A}^T y(0)$. Insgesamt erhalten wir damit die Aussage:

Folgerung 8. Sei $\beta = 2$ und gilt

$$\tilde{x}_1^0 \in \left[-\frac{\bar{u}_1}{1 + \sqrt{2}}, -\frac{u_1}{1 + \sqrt{2}} \right] \quad \text{und} \quad \tilde{x}_2^0 \in \left[-\frac{\bar{u}_2}{1 + \sqrt{2}}, -\frac{u_2}{1 + \sqrt{2}} \right],$$

dann existiert eine optimale Lösung der dualen Aufgabe (D_ν).

Beweis. Sei y aus (6.21)-(6.22) und x_0 entsprechend der Voraussetzungen der Folgerung gegeben. Dann gilt

$$\begin{aligned} y_1''(t) &\stackrel{(6.20)}{=} \{2y_1'(t) - 2y_2'(t) + 2y_1(t) + 2y_2(t)\} \\ &= 2 \left\{ \left[(1 + \sqrt{2})\tilde{x}_1^0 - \tilde{x}_2^0 \right] \sin t + \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad - 2 \left\{ \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \sin t + \left[-(1 + \sqrt{2})\tilde{x}_1^0 + \tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad + 2 \left\{ \tilde{x}_2^0(1 + \sqrt{2}) \sin t - \tilde{x}_1^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad + 2 \left\{ -\tilde{x}_1^0(1 + \sqrt{2}) \sin t - \tilde{x}_2^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t} \\ &= \left\{ \left[-2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \sin t + \left[2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t} \end{aligned}$$

und

$$\begin{aligned} y_2''(t) &\stackrel{(6.20)}{=} \{2y_1'(t) + 2y_2'(t) - 2y_1(t) + 2y_2(t)\} \\ &= 2 \left\{ \left[(1 + \sqrt{2})\tilde{x}_1^0 - \tilde{x}_2^0 \right] \sin t + \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad + 2 \left\{ \left[\tilde{x}_1^0 + (1 + \sqrt{2})\tilde{x}_2^0 \right] \sin t + \left[-(1 + \sqrt{2})\tilde{x}_1^0 + \tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad - 2 \left\{ \tilde{x}_2^0(1 + \sqrt{2}) \sin t - \tilde{x}_1^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t} \\ &\quad + 2 \left\{ -\tilde{x}_1^0(1 + \sqrt{2}) \sin t - \tilde{x}_2^0(1 + \sqrt{2}) \cos t \right\} e^{(1-\sqrt{2})t} \\ &= \left\{ \left[2\tilde{x}_1^0 - 2\tilde{x}_2^0 \right] \sin t + \left[2\tilde{x}_1^0 + 2\tilde{x}_2^0 \right] \cos t \right\} e^{(1-\sqrt{2})t}. \end{aligned}$$

Damit lösen die Funktionen $y_1', y_2' \in W_n^{2,1}(\mathbb{R}_{++}, e^{\gamma t})$ aus (6.21)-(6.22) offenbar die Differentialgleichung (6.20), d.h. für diese gilt $y \in \tilde{Y}_{\nu_1}$ und somit $\tilde{Y}_{\nu_1} \neq \emptyset$. Zudem haben wir für

Anfangswerte aus diesem Intervall bereits die Gültigkeit der Voraussetzungen (V1)-(V4) und (VD1)-(VD4) gezeigt. Damit folgt aus Satz 20 die Existenz einer optimalen Lösung der dualen Aufgabe. \square

Für den Fall $\beta = 2$ konnten wir somit die Gültigkeit der Bedingungen (V1)-(V4), (VD1)-(VD4) und $\tilde{Y}_{\nu_1} \neq \emptyset$ für Anfangswerte

$$\tilde{x}_1^0 \in \left[-\frac{\bar{u}_1}{1 + \sqrt{2}}, -\frac{u_1}{1 + \sqrt{2}} \right] \quad \text{und} \quad \tilde{x}_2^0 \in \left[-\frac{\bar{u}_2}{1 + \sqrt{2}}, -\frac{u_2}{1 + \sqrt{2}} \right]$$

zeigen. Damit ist auch die Existenz einer optimalen Lösung (y_*, q_*) der dualen Aufgabe für diese Wahl der Anfangswerte gesichert. Satz 23 liefert uns schließlich den Zusammenhang zur eindeutigen optimalen Lösung der primalen Aufgabe (P_ν) durch²⁴

$$x_*(t) = q_*(t)e^{-\beta t} \quad \text{und} \quad u_*(t) = \text{Proj}_{[\underline{u}, \bar{u}]} \left(B^T(t)y_*(t)e^{-\beta t} \right).$$

6.2.5 Anwendung der direkten Pseudospektralmethode

Für die numerische Behandlung mittels der in Kapitel 5.1 vorgestellten direkten, modifizierten Pseudospektralmethode gehen wir mit Hilfe der Setzung

$$q(t) := y'(t) + A^T y(t) \quad (\forall) t \in \mathbb{R}_{++}$$

über zu der Formulierung der dualen Aufgabe in Form eines Variationsproblems. Diese Setzung führt uns zu der Aufgabe (D_ν)

Minimiere

$$G(y, y') := \int_0^\infty \left\{ \frac{1}{2} [y'(t) + A^T y(t)]^T [y'(t) + \hat{A}^T y(t)] + \theta(t, y(t)) \right\} e^{-\beta t} dt + x_0^T y(0)$$

bezüglich aller

$$y \in W_2^{2,2}(\mathbb{R}_{++}, e^{-\beta t}) \quad \text{mit} \quad x_0 = y'(0) + A^T y(0) \quad \text{und} \quad \beta > 0,$$

wobei

$$\theta(t, y(t)) = \sum_{i=1}^2 -\frac{1}{2} \sigma_i^2(t, y(t)) + \sigma_i(t, y(t)) y_i(t) \quad \text{und}$$

$$\sigma_i(t, y(t)) = \min \left\{ \max \left\{ \underline{u}_i, y_i(t) e^{-\beta t} \right\}, \bar{u}_i \right\} e^{\beta t} \quad \text{ist.}$$

Entsprechend des Vorgehens in Kapitel 5.2 wählen wir nun die Ansätze

$$\begin{aligned} y_N^{(1)}(t) &= \sum_{k=1}^N a_k S_k(t) + E_1 \cdot t + D_1 & y_N^{(2)}(t) &= \sum_{k=1}^N b_k S_k(t) + E_2 \cdot t + D_2 \\ y_N'^{(1)}(t) &= \sum_{k=1}^N a_k T_k(t) + E_1 & y_N'^{(2)}(t) &= \sum_{k=1}^N b_k T_k(t) + E_2 \end{aligned}$$

und setzen diese in das Variationsproblem ein. Durch dieses Vorgehen erhalten wir das endlich dimensionale, konvexe Problem (D^N) :

²⁴siehe Satz 23, S. 104

Minimiere

$$G^N((a_1, \dots, a_N), (b_1, \dots, b_N), E_1, E_2, D_1, D_2) := \int_0^{\infty} \tilde{g}_N(t, \mathbf{a}, \mathbf{b}, \mathbf{E}, \mathbf{D}) e^{-\beta t} dt + \tilde{x}_1^0 D_1 + \tilde{x}_2^0 D_2$$

bezüglich aller

$$(\mathbf{a}, \mathbf{b}, \mathbf{E}, \mathbf{D}) \in \mathbb{R}^{2N+4},$$

für die gilt

$$\begin{aligned} \tilde{x}_1^0 &= E_1 + D_2 \\ \tilde{x}_2^0 &= E_2 - D_1. \end{aligned}$$

Die Auswertung der Nebenbedingung erlaubt uns die Eliminierung zweier Variablen und führt uns zu dem unrestringierten, konvexen und endlich dimensionalen Optimierungsproblem

Minimiere

$$\int_0^{\infty} \hat{g}_N(t, \mathbf{a}, \mathbf{b}, \mathbf{E}, E_2 - \tilde{x}_2^0, \tilde{x}_1^0 - E_1) e^{-\beta t} dt + \tilde{x}_1^0 (E_2 - \tilde{x}_2^0) + \tilde{x}_2^0 (\tilde{x}_1^0 - E_1)$$

bezüglich aller

$$(\mathbf{a}, \mathbf{b}, \mathbf{E}) \in \mathbb{R}^{2N+2}.$$

Aufgrund der Konvexität sind alle lokalen Lösungen dieses endlich dimensionalen Optimierungsproblems globale Lösungen und der Zielfunktionswert ist eindeutig bestimmt.

Gemäß der Betrachtungen aus Kapitel [5.2](#) erfüllt die Optimallösung die Kollokationsgleichungen der direkten Methode. Diese liefern ein System aus $2N+2$ Gleichungen mit $2N+2$ Variablen. Eine Darstellung des Systems findet sich auf der nachfolgenden Seite.

Kollokationsgleichungen der direkten Methode für das linearisierte Lotka-Volterra Modell²⁵

$$\begin{aligned} & \sum_{i=1}^N \mathbf{w}_i \left\{ \sum_{k=1}^N a_k \left\{ T_j(\tau_i^{(N)}) T_k(\tau_i^{(N)}) + S_j(\tau_i^{(N)}) S_k(\tau_i^{(N)}) \right\} + \sum_{k=1}^N b_k \left\{ -S_j(\tau_i^{(N)}) T_k(\tau_i^{(N)}) + T_j(\tau_i^{(N)}) S_k(\tau_i^{(N)}) + E_1 \left\{ \tau_i^{(N)} S_j(\tau_i^{(N)}) \right\} \right\} \right\} \\ & + \sum_{i=1}^N \mathbf{w}_i \left\{ E_2 \left\{ \tau_i^{(N)} T_j(\tau_i^{(N)}) \right\} + \tilde{x}_1^0 \left\{ \beta S_j(\tau_i^{(N)}) \right\} + \tilde{x}_2^0 \left\{ -S_j(\tau_i^{(N)}) \right\} + S_j(\tau_i^{(N)}) \sigma_{1N}(\tau_i^{(N)}, \mathbf{a}, E_1, E_2) \right\} \stackrel{!}{=} 0 \quad \forall j = 1, \dots, N \end{aligned}$$

$$\begin{aligned} & \sum_{i=1}^N \mathbf{w}_i \left\{ \sum_{k=1}^N a_k \left\{ S_j(\tau_i^{(N)}) T_k(\tau_i^{(N)}) - T_j(\tau_i^{(N)}) S_k(\tau_i^{(N)}) \right\} + \sum_{k=1}^N b_k \left\{ T_j(\tau_i^{(N)}) T_k(\tau_i^{(N)}) + S_j(\tau_i^{(N)}) S_k(\tau_i^{(N)}) \right\} + E_1 \left\{ -\tau_i^{(N)} T_j(\tau_i^{(N)}) \right\} \right\} \\ & + \sum_{i=1}^N \mathbf{w}_i \left\{ E_2 \left\{ \tau_i^{(N)} S_j(\tau_i^{(N)}) \right\} + \tilde{x}_1^0 \left\{ S_j(\tau_i^{(N)}) \right\} + \tilde{x}_2^0 \left\{ \beta S_j(\tau_i^{(N)}) \right\} S_j(\tau_i^{(N)}) \sigma_{2N}(\tau_i^{(N)}, \mathbf{b}, E_1, E_2) \right\} \stackrel{!}{=} 0 \quad \forall j = 1, \dots, N \end{aligned}$$

$$\begin{aligned} & \sum_{i=1}^N \mathbf{w}_i \left\{ \sum_{k=1}^N a_k \left\{ \tau_i^{(N)} S_k(\tau_i^{(N)}) \right\} + \sum_{k=1}^N b_k \left\{ -\tau_i^{(N)} T_k(\tau_i^{(N)}) \right\} + E_1 \left\{ \tau_i^{(N)2} \right\} + E_2 \{0\} + \tilde{x}_1^0 \left\{ -1 + \beta \tau_i^{(N)} \right\} + \tilde{x}_2^0 \left\{ -\tau_i^{(N)} - \beta \right\} \right\} \\ & + \sum_{i=1}^N \mathbf{w}_i \left\{ \tau_i^{(N)} \sigma_{1N}(\tau_i^{(N)}, \mathbf{a}, E_1, E_2) - \sigma_{2N}(\tau_i^{(N)}, \mathbf{b}, E_1, E_2) \right\} \stackrel{!}{=} 0 \end{aligned}$$

$$\begin{aligned} & \sum_{i=1}^N \mathbf{w}_i \left\{ \sum_{k=1}^N a_k \left\{ \tau_i^{(N)} T_k(\tau_i^{(N)}) \right\} + \sum_{k=1}^N b_k \left\{ \tau_i^{(N)} S_k(\tau_i^{(N)}) \right\} + E_1 \{0\} + E_2 \left\{ \tau_i^{(N)2} \right\} + \tilde{x}_1^0 \left\{ \tau_i^{(N)} + \beta \right\} + \tilde{x}_2^0 \left\{ -1 + \beta \tau_i^{(N)} \right\} \right\} \\ & + \sum_{i=1}^N \mathbf{w}_i \left\{ \sigma_{1N}(\tau_i^{(N)}, \mathbf{a}, E_1, E_2) + t \sigma_{2N}(\tau_i^{(N)}, \mathbf{b}, E_1, E_2) \right\} \stackrel{!}{=} 0 \end{aligned}$$

²⁵ $\sigma_{1N}(t, \mathbf{a}, E_1, E_2) = \min \left\{ \max \left\{ \underline{u}_1 e^{\beta t}, \sum_{k=1}^N a_k S_k(t) + E_1 \cdot t + E_2 - \tilde{x}_2^0 \right\}, \bar{u}_1 e^{\beta t} \right\}, \sigma_{2N}(t, \mathbf{b}, E_1, E_2) = \min \left\{ \max \left\{ \underline{u}_2 e^{\beta t}, \sum_{k=1}^N b_k S_k(t) + E_2 \cdot t + \tilde{x}_1^0 - E_1 \right\}, \bar{u}_2 e^{\beta t} \right\}$

6.2.6 Optimale Lösung

Zur Bewertung der numerischen Ergebnisse betrachten wir zunächst eine Formulierung des Problems (P_ν) ohne Steuerungsbeschränkungen. Für die Wahl des Gewichtungsparmeters $\beta = 2$ können wir eine analytische Lösung des gegebenen Stabilisierungsproblems durch die Auswertung der Maximumbedingung aus [PB19]²⁶ berechnen. Die Auswertung dieser Maximumbedingung liefert für das Intervall $[0, \infty)$ den Kandidaten

$$\tilde{x}_1^*(t) = \left\{ -\tilde{x}_2^0 \sin(t) + \tilde{x}_1^0 \cos(t) \right\} e^{-(1+\sqrt{2})t} \quad (6.23)$$

$$\tilde{x}_2^*(t) = \left\{ \tilde{x}_2^0 \cos(t) + \tilde{x}_1^0 \sin(t) \right\} e^{-(1+\sqrt{2})t} \quad (6.24)$$

$$u_1^*(t) = \left\{ \tilde{x}_2^0(1 + \sqrt{2}) \sin(t) - \tilde{x}_1^0(1 + \sqrt{2}) \cos(t) \right\} e^{-(1+\sqrt{2})t} \quad (6.25)$$

$$u_2^*(t) = \left\{ -\tilde{x}_2^0(1 + \sqrt{2}) \cos(t) - \tilde{x}_1^0(1 + \sqrt{2}) \sin(t) \right\} e^{-(1+\sqrt{2})t} \quad (6.26)$$

mit dem Zielfunktionalwert

$$J(x^*, u^*) = \frac{1 + \sqrt{2}}{4} \left\{ (\tilde{x}_1^0)^2 + (\tilde{x}_2^0)^2 \right\}$$

Unter Beachtung der Bedingungen aus Folgerung 7 ist auch der zulässige Bereich für das Problem ohne Steuerungsbeschränkungen nichtleer und nicht einelementig, so dass aufgrund der strikten Konvexität des Zielfunktionals der so berechnete Lösungskandidat die eindeutige optimale Lösung liefert²⁷. Für das freie Problem können wir nun a posteriori Steuerungsbeschränkungen angeben durch $u \in [-C, C] \times [-C, C]$, wobei $C \geq (1 + \sqrt{2}) \max \{ |\tilde{x}_1^0|, |\tilde{x}_2^0| \}$. Damit gelangen wir zu einer Problemformulierung der Form (P_ν) mit Steuerungsbeschränkungen, mit der gleichen optimalen Lösung (x^*, u^*) . Die Aussagen des Satzes 23 liefern uns dann den Zusammenhang zu den optimalen Lösungen der stark dualen Aufgabe (D_ν)

$$y_1^*(t) = \left\{ \tilde{x}_2^0(1 + \sqrt{2}) \sin(t) - \tilde{x}_1^0(1 + \sqrt{2}) \cos(t) \right\} e^{(1-\sqrt{2})t} \quad (6.27)$$

$$y_2^*(t) = \left\{ -\tilde{x}_2^0(1 + \sqrt{2}) \cos(t) - \tilde{x}_1^0(1 + \sqrt{2}) \sin(t) \right\} e^{(1-\sqrt{2})t} \quad (6.28)$$

$$p_1^*(t) = \left\{ -\tilde{x}_2^0 \sin(t) + \tilde{x}_1^0 \cos(t) \right\} e^{(1-\sqrt{2})t} \quad (6.29)$$

$$p_2^*(t) = \left\{ \tilde{x}_2^0 \cos(t) + \tilde{x}_1^0 \sin(t) \right\} e^{(1-\sqrt{2})t}, \quad (6.30)$$

sowie

$$q^*(t) = p^*(t) \quad \forall t \in \mathbb{R}_+.$$

Die so konstruierte Lösung soll nachfolgend als Referenz dienen.

6.2.7 Auswertung der numerischen Ergebnisse

Für die Auswertung der numerischen Ergebnisse betrachten wir nachfolgend unterschiedliche Konstellationen von Anfangswerten und Gewichtungsparmetern. Zur Darstellung der Effizienz des Verfahrens beginnen wir mit der Betrachtung der freien Steuerung, für welche wir eine exakte Lösung durch (x^*, u^*) aus den Gleichungen (6.23)-(6.30) als Vergleich heranziehen können.

²⁶vgl. dortiges Theorem 2, S. 10

²⁷siehe Folgerung 7, S. 143

Betrachtung für $\beta = 2$, $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1 + \sqrt{2})$

Wir starten mit der Betrachtung einer Konstellation aus Gewichtungparameter und Anfangswerten, die nach vorangegangener Betrachtung die freie Steuerung ermöglicht. Die Wahl von $\beta = 2$ erlaubt zudem den direkten Vergleich zur analytischen Lösung. Gemäß der Überlegungen aus Unterkapitel [6.2.6](#) erhielten wir für die freie Lösung a posteriori die Steuerungsbeschränkungen

$$u \in [-C, C] \times [-C, C] \quad \text{mit} \quad C \geq (1 + \sqrt{2}) \max \left\{ \left| \tilde{x}_1^0 \right|, \left| \tilde{x}_2^0 \right| \right\}.$$

Für den Fall normierter Steuerungsbeschränkungen der Form $u \in [-1, 1] \times [-1, 1]$ erhalten wir daher für $|\tilde{x}_1^0| \leq \frac{1}{1+\sqrt{2}}$ und $|\tilde{x}_2^0| \leq \frac{1}{1+\sqrt{2}}$ eine optimale Lösung der Form [\(6.23\)](#)-[\(6.30\)](#). Wir wählen daher $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1 + \sqrt{2})$. Die zugehörige analytische Lösung lautet dann

$$\begin{aligned} \tilde{x}_1^*(t) &= \left(-\frac{1}{(1+\sqrt{2})} \sin(t) + \frac{1}{(1+\sqrt{2})} \cos(t) \right) e^{-(1+\sqrt{2})t}, & u_1^*(t) &= (\sin(t) - \cos(t)) e^{-(1+\sqrt{2})t}, \\ \tilde{x}_2^*(t) &= \left(\frac{1}{(1+\sqrt{2})} \cos(t) + \frac{1}{(1+\sqrt{2})} \sin(t) \right) e^{-(1+\sqrt{2})t}, & u_2^*(t) &= (-\cos(t) - \sin(t)) e^{-(1+\sqrt{2})t} \end{aligned}$$

mit den zugehörigen dualen Lösungen

$$\begin{aligned} y_1^*(t) &= (\sin(t) - \cos(t)) e^{(1-\sqrt{2})t}, & q_1^*(t) &= \left(-\frac{1}{(1+\sqrt{2})} \sin(t) + \frac{1}{(1+\sqrt{2})} \cos(t) \right) e^{(1-\sqrt{2})t}, \\ y_2^*(t) &= (-\cos(t) - \sin(t)) e^{(1-\sqrt{2})t}, & q_2^*(t) &= \left(\frac{1}{(1+\sqrt{2})} \cos(t) + \frac{1}{(1+\sqrt{2})} \sin(t) \right) e^{(1-\sqrt{2})t} \end{aligned}$$

und dem Zielfunktionalwert

$$J(x^*, u^*) = G(y^*, q^*) = \frac{1}{2(1 + \sqrt{2})} \approx 0.207106781186548.$$

Wir präsentieren nun die numerischen Ergebnisse für eine unterschiedliche Anzahl N an Kollokationspunkten. Dabei werden wir feststellen, dass die verwendete Pseudospektralmethode bereits für $N = 4$, d.h. für die Verwendung von lediglich 4 Stützstellen, eine derart gute Approximation an die optimalen Lösungen liefert, dass der exakte Zielfunktionswert bereits bis auf 4 Nachkommastellen genau approximiert werden kann. Mit steigender Anzahl an Kollokationspunkten wird eine Verbesserung der Approximation erreicht. Dies lässt sich insbesondere an den Zielfunktionswerten erkennen, wie der folgenden Tabelle zu entnehmen ist:

N	$G(y_N, y'_N)$	$ G(y_N, y'_N) - J(x^*, u^*) $	$\frac{ G(y_N, y'_N) - J(x^*, u^*) }{ J(x^*, u^*) }$
2	0.210988265514820	0.003881484328273	0.018741464214911
4	0.207393276071427	$2.864948848796678 \cdot 10^{-4}$	0.001383319673254
6	0.207119651024813	$1.286983826537247 \cdot 10^{-5}$	$6.214107617162076 \cdot 10^{-5}$
8	0.207107259470989	$4.782844419404242 \cdot 10^{-7}$	$2.309361572809233 \cdot 10^{-6}$
10	0.207106797497157	$1.631060922435701 \cdot 10^{-8}$	$7.875458800002063 \cdot 10^{-8}$

Tabelle 6.7: Numerisch ermittelte Zielfunktionswerte für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$ und $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1 + \sqrt{2})$.

Für $N = 6, 8, 10$ sind die Näherungslösungen der approximierten Funktionen y_N und p_N in den Abbildungen [6.15](#)-[6.16](#) grafisch in einem direkten Vergleich zu den exakten Lösungen dargestellt.

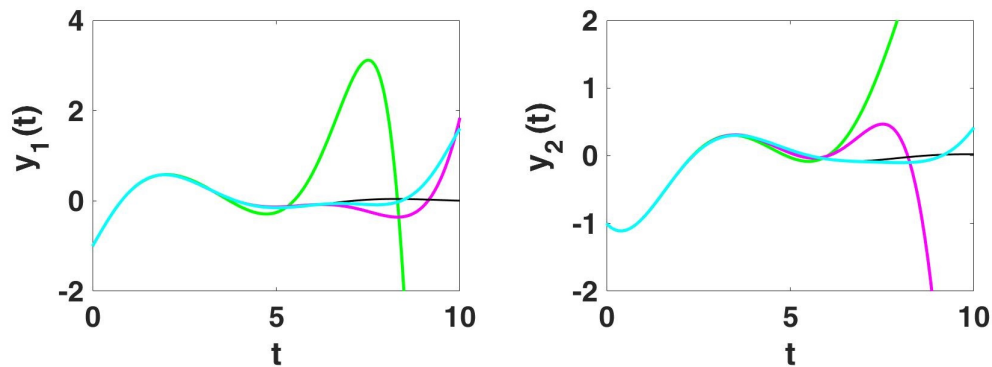


Abbildung 6.15: Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $y^*(t)$ (—)

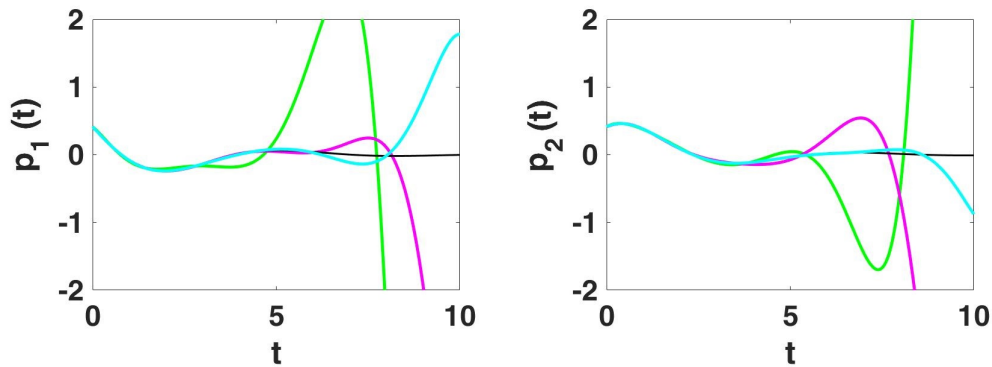


Abbildung 6.16: Näherungslösung $p_N(t)$ der dualen Adjungierten für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $p^*(t)$ (—)

Die jeweils letzten Kollokationspunkte sind durch

$$\tau_6^{(6)} = 7.991436990300850 \quad \tau_8^{(8)} = 11.431565868444633 \quad \tau_{10}^{(10)} = 14.960348506136945$$

festgelegt. Die Abbildungen verdeutlichen, dass sich die approximierten Lösungen mit wachsendem N auf einem vergrößerten Intervall den exakten Lösungen annähern. Bemerkenswert ist zudem die Genauigkeit der Anfangswerte der approximierten Funktionen $y_N(0)$ und $p_N(0)$ im Vergleich zu den Anfangswerten der exakten Lösung $y^*(0) = (-1, -1)$ und $p^*(0) = (\frac{1}{(1+\sqrt{2})}, \frac{1}{(1+\sqrt{2})})$. Insbesondere die Anfangswerte der dualen Zustandsfunktion sind sowohl auf primaler als auch auf dualer Seite a priori nicht bekannt. Die Anfangswerte der dualen Zustandsfunktion ergeben sich dabei aus den zu bestimmenden Koeffizienten E_1 und E_2 . Die nachfolgende Tabelle [6.8](#) veranschaulicht die Approximationsgüte der so bestimmten Anfangswerte durch die Angabe des absoluten Fehlers $|y_N^{(i)}(0) - y_i^*(0)|$.

N	$y_N^{(1)}(0) - y_1^*(0)$	$y_N^{(2)}(0) - y_2^*(0)$
2	0.048428302655103	0.003648458072925
4	0.003023207577894	0.012399637837796
6	0.001725808462099	$7.751414198051698 \cdot 10^{-4}$
8	$1.134107335014622 \cdot 10^{-4}$	$2.476685800714140 \cdot 10^{-4}$
10	$3.780172365575041 \cdot 10^{-5}$	$1.713957112814413 \cdot 10^{-5}$

Tabelle 6.8: Vergleich der numerisch ermittelten Anfangswerte der Funktion $y_N(t)$ zur exakten Lösung $y^*(t)$ für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$ und $\tilde{x}_1^0 = \tilde{x}_2^0 = 1/(1+\sqrt{2})$.

Mit Hilfe dieser direkt approximierten Funktionen erhalten wir durch die bereits beschriebenen Setzungen

$$x_N(t) = q_N(t)e^{-2t} \text{ und } u_N(t) = \text{Proj}_{[-1,1]}(y_N(t)e^{-2t})$$

die approximierten Lösungen der primalen Seite. Bemerkenswerter Weise erhalten wir derartig gute Approximationen an die exakte optimale Lösung (x^*, u^*) , dass diese grafisch nicht von den approximierten Lösungen zu unterscheiden ist (vgl. Abbildungen [6.17](#) und [6.18](#)). Dieser Umstand ist umso bemerkenswerter, wenn man sich verdeutlicht, dass sich die so ermittelten Lösungen nur indirekt aus den approximierten Lösungen $y_N(t)$, $p_N(t)$ ergeben.

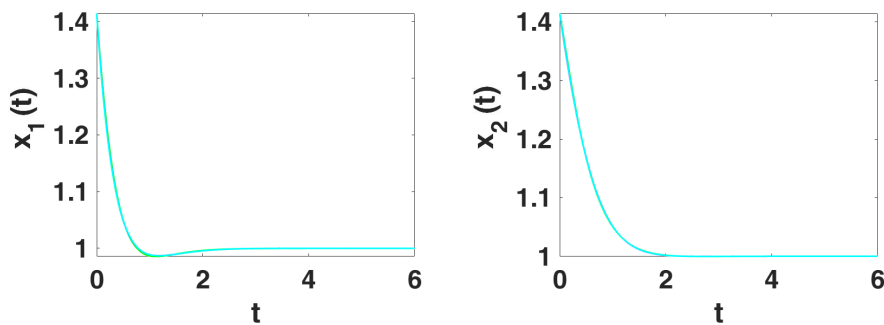


Abbildung 6.17: Näherungslösung $x_N(t)$ des rücktransformierten primalen Zustandes für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $x^*(t)$ (—)

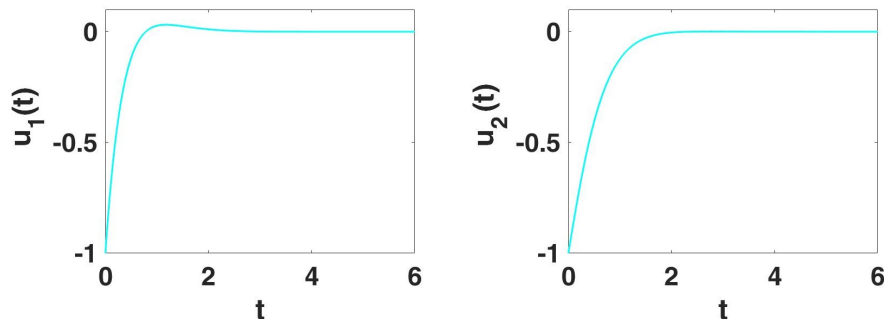


Abbildung 6.18: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 6$ (—), $N = 8$ (—) und $N = 10$ (—) im Vergleich zur exakten Lösung $u^*(t)$ (—)

Betrachtung für $\beta = 2$, $\tilde{x}_1^0 = -0.9$, $\tilde{x}_2^0 = 0.6$

Wir betrachten nun den Fall aktiver Steuerungsbeschränkungen²⁸. Mit den Setzungen $\beta = 2$, $\tilde{x}_1^0 = -0.9$ und $\tilde{x}_2^0 = 0.6$ sind die Bedingungen der Folgerung 7 für die normierten Steuerungsbeschränkungen $u \in [-1, 1] \times [-1, 1]$ erfüllt²⁹, denn mit $\rho_1 = \rho_2 = 1.01 > 0.5\beta$ gilt

$$\tilde{x}_1^0, \tilde{x}_2^0 \in \left[\frac{-(1+\rho)}{\rho^2+1}, \frac{1+\rho}{\rho^2+1} \right] = [-0.9950002475, 0.9950002475].$$

Damit ist die Existenz einer optimalen Lösung der primalen Aufgabe für diese Konstellation an Anfangswerten und Gewichtungsparmeter gesichert. Zudem sichern die vorangegangenen Untersuchungen die schwache Dualitätsbeziehung. Die erhaltenen numerischen Lösungen liefern somit zumindest untere Schranken an die optimale Lösung der primalen Aufgabe. Die Existenz einer Lösung der dualen Aufgabe können wir für diese Setzung der Anfangswerte jedoch nicht garantieren.

Die Anwendung des Verfahrens liefert die folgenden approximierten Lösungen der dualen Aufgaben:

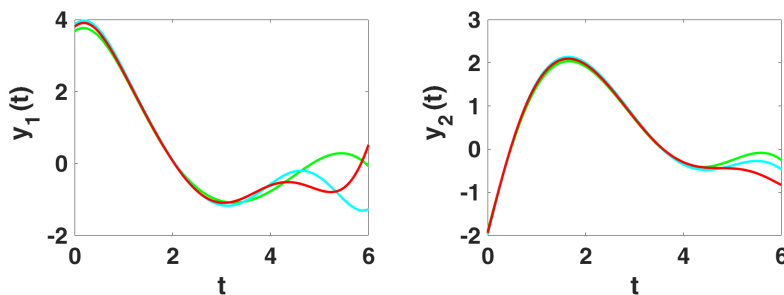


Abbildung 6.19: Näherungslösung $y_N(t)$ des dualen Zustandes für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)

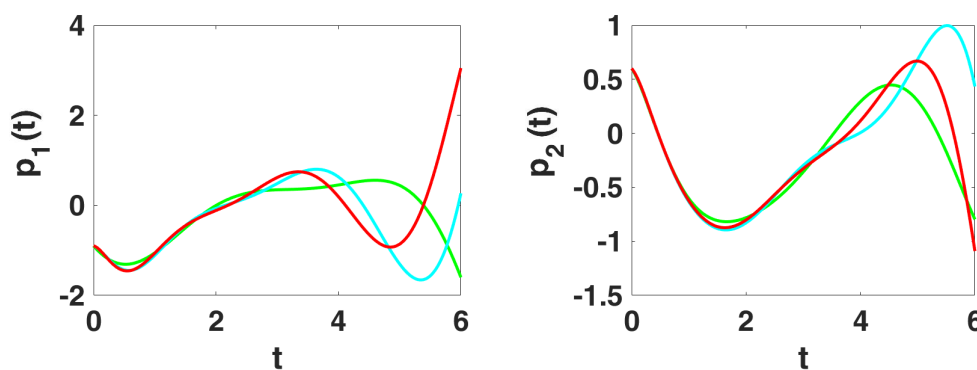


Abbildung 6.20: Näherungslösung $p_N(t)$ der dualen Adjungierten für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)

²⁸Gemäß der Betrachtungen aus Unterkapitel 6.2.6 sind die normierten Steuerungsbeschränkungen $u \in [-1, 1] \times [-1, 1]$ mit dieser Wahl der Anfangswerte aktiv.

²⁹vgl. Folgerung 7, S. 143

Unterstellen wir den Zusammenhang

$$\tilde{x}(t) = \hat{W}^{-1}(t)q(t) \quad \text{und} \quad u(t) = \text{Proj}_{\mathcal{U}} \left(\hat{R}^{-1}(t)B^T y(t) \right) \quad (6.31)$$

– von dessen Gültigkeit wir jedoch nicht ohne Weiteres ausgehen können – so ergeben sich die folgenden Approximationen an die primalen Lösungen:

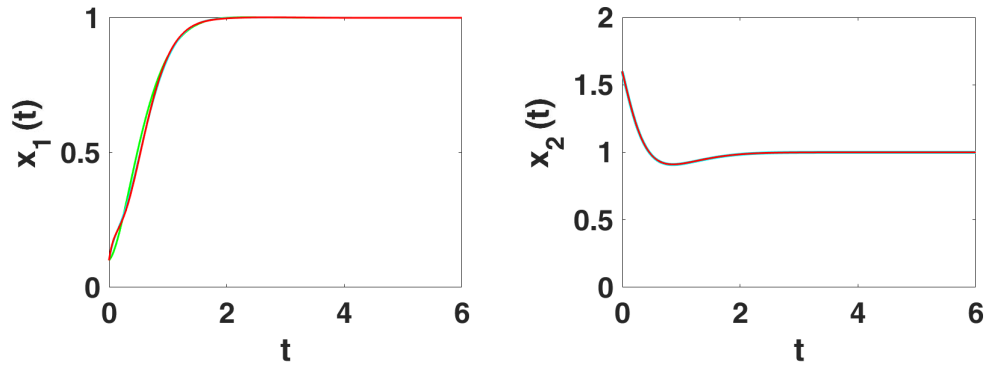


Abbildung 6.21: Näherungslösung $x_N(t)$ des rücktransformierten primalen Zustandes für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)

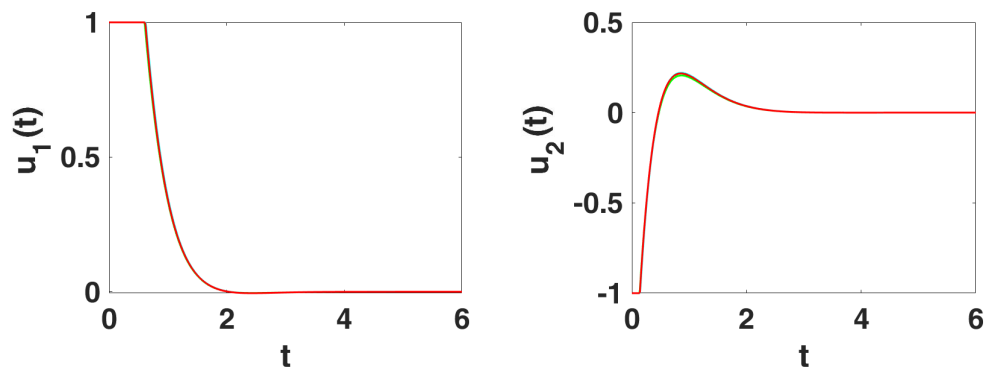


Abbildung 6.22: Näherungslösung $u_N(t)$ der primalen Steuerung für $N = 8$ (—), $N = 10$ (—) und $N = 11$ (—)

Die Abbildungen [6.19](#) bis [6.22](#) lassen eine Konvergenz des Verfahrens vermuten. Der Vergleich von Zielfunktionswerten, Anfangswerten der dualen Zustände und Schaltzeitpunkten scheint diese These zu unterstreichen. Dabei bezeichnet $\bar{t}_N^{(i)}$ den zur Komponente u_i gehörigen Schaltzeitpunkt. Die konkreten Werte können der Tabelle [6.9](#) entnommen werden.

N	Zielfunktionswert	$y_N^{(1)}(0)$	$y_N^{(2)}(0)$	$\bar{t}_N^{(1)}$	$\bar{t}_N^{(2)}$
8	0.8083439244	3.6767967251870	-1.8989950258459	0.767404	0.137007
9	0.8225601261	3.8544222679455	-1.9856625633094	0.838172	0.145616
10	0.8319618152	3.8635737246205	-1.9759051371537	0.855510	0.144003
11	0.8362073880	3.8063547689544	-1.9370968603347	0.853675	0.139661

Tabelle 6.9: Numerisch ermittelte Zielfunktionswerte, Anfangswerte und Schaltzeitpunkte für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

Da für das vorliegende Modell mit aktiven Steuerbeschränkungen keine analytische Lösung zum Vergleich vorliegt, nutzen wir den Vergleich der Differenzen der in der Tabelle 6.9 beschriebenen Werte. Auf Grundlage dieses Vorgehens zeichnet sich für die Zielfunktionswerte und für den Schaltzeitpunkt $\bar{t}_N^{(1)}$ eine Genauigkeit von 10^{-2} und für alle übrigen aufgelisteten Werte von 10^{-1} ab.

Wir wollen nun überprüfen, ob die derart ermittelte Lösung Element der Menge \tilde{Y}_{ν_1} ist. In diesem Fall wäre die Existenz einer optimalen Lösung der dualen Aufgabe und damit die Anwendbarkeit des Satzes 23 gesichert³⁰. Die erhaltenen numerischen Lösungen legen, zusammen mit den Bedingungen aus Satz 23, die folgenden Ansätze der Schaltstruktur nahe:

$$u_1^*(t) = \begin{cases} 1, & t < \bar{t}_N^{(1)} \\ y_1^*(t), & t \geq \bar{t}_N^{(1)} \end{cases} \quad \text{und} \quad u_2^*(t) = \begin{cases} -1, & t < \bar{t}_N^{(2)} \\ y_2^*(t), & t \geq \bar{t}_N^{(2)} \end{cases},$$

wobei $\bar{t}_N^{(2)} < \bar{t}_N^{(1)}$. Damit ergibt sich für die Funktion $\hat{\sigma}^*(\cdot, y^*(\cdot))$:

$$\hat{\sigma}(t, y^*(t)) := \begin{cases} (e^{2t}, -e^{2t})^T, & t < \bar{t}_N^{(2)} \\ (e^{2t}, y_2^*(t))^T, & \bar{t}_N^{(2)} \leq t < \bar{t}_N^{(1)} \\ (y_1^*(t), y_2^*(t))^T, & t \geq \bar{t}_N^{(1)} \end{cases}.$$

Ist die optimale Lösung y^* Element der Menge \tilde{Y}_{ν_1} , so genügt sie der Differentialgleichung

$$y''(t) = \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix} y'(t) + \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} y(t) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \hat{\sigma}(t, y(t)), \quad (6.32)$$

mit $q(0) = y'(0) + A^T y(0) = \tilde{x}_0$.

Für die obig unterstellte Schaltstruktur veranschaulicht Abbildung 6.23 die numerisch ermittelten Lösungen der Differentialgleichung (6.32) für die Intervalle $[0, \bar{t}_N^{(2)})$ und $[\bar{t}_N^{(2)}, \bar{t}_N^{(1)})$ sowie die analytische Lösung des Intervalls $[\bar{t}_N^{(1)}, +\infty)$. Dabei wurden die numerisch ermittelten Schaltzeitpunkte $\bar{t}_N^{(1)}$ und $\bar{t}_N^{(2)}$ für $N = 11$ Kollokationspunkte zugrunde gelegt.

Die analytische Lösung liefert für das Intervall $[\bar{t}_N^{(1)}, +\infty)$ die Struktur

$$\begin{aligned} y_N^{(1)}(t) &= \left\{ C_1 \left(\frac{(2+\sqrt{2})}{4} \cos t + \frac{\sqrt{2}}{4} \sin t \right) + C_2 \left(-\frac{\sqrt{2}}{4} \cos t + \frac{(2+\sqrt{2})}{4} \sin t \right) \right\} e^{(1-\sqrt{2})t} \\ y_{1N}^{(2)}(t) &= \left\{ C_1 \left(-\frac{\sqrt{2}}{4} \cos t + \frac{(2+\sqrt{2})}{4} \sin t \right) + C_2 \left(\frac{-(2+\sqrt{2})}{4} \cos t - \frac{\sqrt{2}}{4} \sin t \right) \right\} e^{(1-\sqrt{2})t} \\ y_N^{\prime(1)}(t) &= \left\{ -C_1 \sin t + \cos t \right\} e^{(1-\sqrt{2})t} \\ y_N^{\prime(2)}(t) &= \left\{ C_1 \cos t + \sin t \right\} e^{(1-\sqrt{2})t} \end{aligned}$$

und damit $y \in W_2^{2,2}((\bar{t}_N^{(1)}, +\infty), e^{-2t})$. Die Abbildung 6.23 legt zudem die stetige Differenzierbarkeit der numerisch ermittelten Lösungen $y_N^{(1)}$ und $y_N^{(2)}$ sowie die stückweise stetige

³⁰siehe Satz 23, S. 104

Differenzierbarkeit der numerisch ermittelten Lösungen $y_N^{(1)}$ und $y_N^{(2)}$ nahe. Damit lassen die vorangegangenen Ausführungen die Existenz einer Funktion $y \in \tilde{Y}_{\nu_i}$ vermuten. In diesem Fall wäre die optimale Lösung der dualen Aufgabe gesichert und Satz 23 anwendbar³¹. Gemäß der starken Dualitätsbeziehung gilt für die optimalen Lösungen (x^*, u^*) und (y^*, q^*) der primalen und dualen Aufgaben zudem $J(x^*, u^*) = G(y^*, q^*)$.

Der Vergleich der approximierten Zielfunktionalwerte der dualen Aufgabe und der primalen Aufgabe (vgl. Tabelle 6.10) unterstreicht die These der starken Dualitätsbeziehung:

N	$G^N(y_N, q_N)$	$J^N(\tilde{x}_N, u_N)$	$G^N(y_N, q_N) - J^N(\tilde{x}_N, u_N)$
8	0.8083439244	0.808484613585	$1.406890902826 \cdot 10^{-4}$
9	0.8225601261	0.821410943777	$1.149182399097 \cdot 10^{-3}$
10	0.8319618152	0.830569784810	$1.392030438526 \cdot 10^{-3}$
11	0.8362073880	0.834869062952	$1.338325111090 \cdot 10^{-3}$

Tabelle 6.10: Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

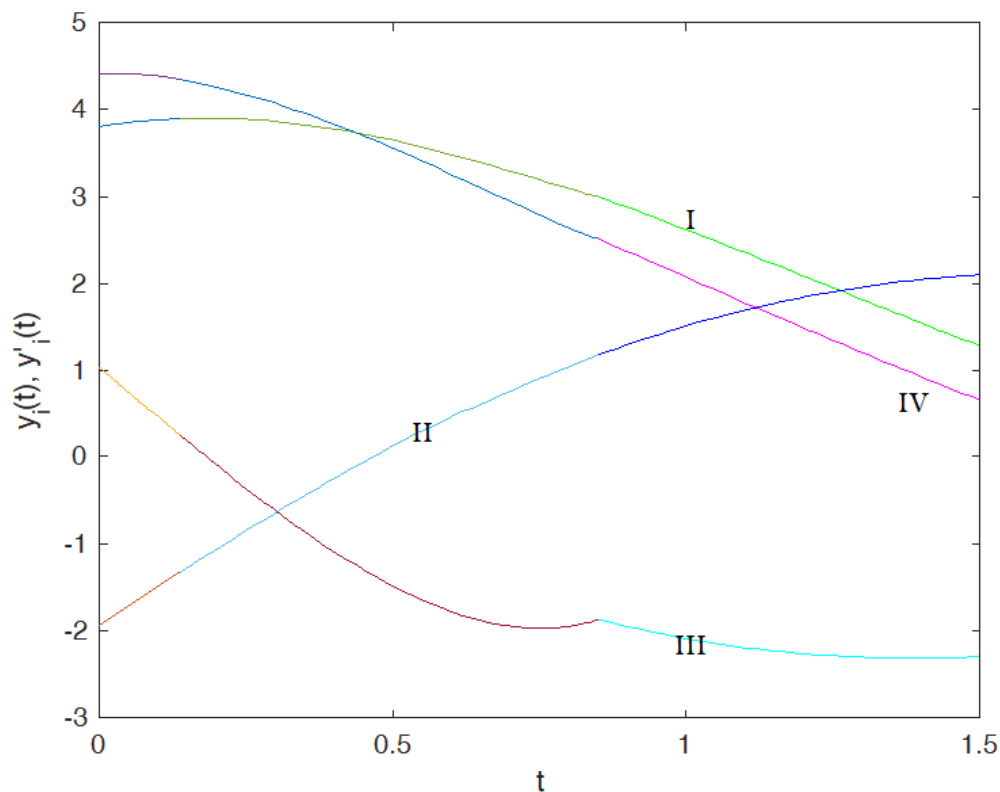


Abbildung 6.23: Numerisch ermittelte Lösungen der Differentialgleichung (6.32) auf Basis der approximierten Schaltzeitpunkte $\bar{t}_N^{(1)}$ und $\bar{t}_N^{(2)}$. Dabei bezeichnet I die Funktion $y_1(t)$, II die Funktion $y_2(t)$, III die Funktion $y_1'(t)$ und IV die Funktion $y_2'(t)$.

³¹Die Abbildungen 6.21 und 6.22 entsprechen in diesem Fall den approximierten Lösungen der primalen Aufgabe.

6.3 Das SEIR Modell

Dem jährlichen Infektionsbericht der WHO (World Health Organisation) ist zu entnehmen, dass die Ausbreitung von Infektionskrankheiten noch immer – weltweit – eines der größten Gesundheitsprobleme darstellt. Ein ganz aktuelles und globales Beispiel liefert der Fall der Viruserkrankung *Covid-19* (engl. für *coronavirus disease 2019*). Nach ihrer erstmaligen Beschreibung im Jahre 2019 in der chinesischen Provinzstadt Wuhan entwickelte sich das Virus innerhalb weniger Monate zu einer weltweiten Pandemie. Aufgrund der Neuartigkeit dieses Virus war eine medikamentöse Eindämmung der Pandemie zunächst nicht möglich. Zudem fehlten verlässliche Informationen über das Verhalten der Viruserkrankung, was zuverlässige Aussagen über geeignete Strategien zur Bekämpfung der Pandemie unmöglich machten. In Folge dessen entwickelten die betroffenen Staaten unterschiedliche und im Laufe der Zeit auch wechselnde Herangehensweisen im Rahmen des Umgangs mit dem neuen Virus.

Zu Beginn der Pandemie verfolgte die deutsche Bundesregierung beispielsweise den sogenannten Ansatz des „Flattening the curve“. Hier war die Idee, durch die Einführung von Isolationsmaßnahmen (wie Ausgangsbeschränkungen) die Verbreitung des Virus zu verlangsamen. Über diese Herangehensweise sollte sichergestellt werden, dass die aktiven Fallzahlen die Kapazitäten des Gesundheitssystems zu keinem Zeitpunkt überschreiten. Andere von der deutschen Bundesregierung genutzte Kennzahlen, welche als Indikatoren für notwendige Verschärfungen oder Lockerungen der Isolationsmaßnahmen dienten, waren beispielsweise die Verdopplungszeit und die Reproduktionsrate – auch hier mit dem Ziel der Verlangsamung bzw. Eindämmung der Verbreitung des Virus. Keine der Strategien basierte jedoch auf der Annahme einer möglichen Auslöschung des Virus zu einer festen oder endlichen Zeit. Stattdessen stützten sich alle Maßnahmen auf der Annahme des Fortbestandes des Erregers und decken sich somit mit unseren Ansätzen der Modellierung über den unendlichen Zeithorizont.

Die beschriebenen Ansätze der deutschen Bundesregierung zum Umgang mit der Covid-19 Pandemie basierten unter anderem auf der Analyse epidemiologischer Modelle. Eines dieser epidemiologischen Modelle ist das sogenannte *SEIR* Modell. Grundlage des Modells ist die Einteilung der Gesamtpopulation N in die vier Komponenten:

- $S(t)$ - Anteil der Anfälligen (*engl. susceptible*) zum Zeitpunkt t
- $E(t)$ - Anteil der Infizierten, welche jedoch (noch) nicht infektiös sind (*engl. exposed*) zum Zeitpunkt t
- $I(t)$ - Anteil der Infektiösen (*engl. infectious*) zum Zeitpunkt t
- $R(t)$ - Anteil der Genesenen (*engl. recovered*) zum Zeitpunkt t ,

so dass sich die Gesamtpopulation über den gesamten Betrachtungszeitraum Ω als Summe der vier Komponenten $N(t) = S(t) + E(t) + I(t) + R(t) \forall t \in \Omega$ darstellen lässt.

Die Behandlung des SEIR Modells findet sich bereits in zahlreichen Publikationen³² mit diversen Betrachtungsschwerpunkten wieder. Beispielsweise legen die Autoren in [LF06] den Fokus auf die Beschreibung stochastischer Modelle, während sich Publikationen wie [LM95]

³²vgl. beispielsweise [CPKM15], [BPDPI14], [LM95], [CTX08], [LF06]

mit den globalen Stabilitätseigenschaften des epidemischen Systems um den Gleichgewichtspunkt beschäftigen. Die Publikation [GTX08] behandelt hingegen die Problematik sogenannter *time-delayed-systems*, welche die Darstellung zeitverzögerter Effekte innerhalb des Systems erlauben. In den Publikationen [dPKM15] und [BPDP14] behandeln die Autoren gar die Einbettung des SEIR Modells in den Kontext der Optimalen Steuerung.

Wir folgen den Ideen der letztgenannten Autoren und beziehen uns auf die dort vorgestellten Notationen. Die der Arbeit [BPDP14] entnommene Abbildung 6.24 zeigt ein schematisches Diagramm des Infektionsverlaufes zwischen den einzelnen Individuen und lässt sich wie folgt interpretieren:

Wir nehmen an, dass jeder Angehörige der Population von Geburt an zur Gruppe S der Anfälligen gehöre und sich durch den Kontakt zu Infizierten, also zu Angehörigen der Gruppe I , zu einem gegebenen Ansteckungskoeffizienten c zur Rate cIS ansteckt. Dabei durchlebt der neu Infizierte zunächst eine gewisse Latenzzeit, so dass er nun Teil der Gruppe E wird. Mit Hilfe des Parameters e – welcher mit eE die Rate beschreibt, mit welcher ein neu Infizierter selbst infektiös wird – modellieren die Autoren den Anteil der Gruppe E , welcher zur Gruppe I übergeht. Es ist davon auszugehen, dass die Infektionen zu einem gegebenen Genesungsparameter g gesunden und somit schließlich Element der Gruppe R werden. Darüber hinaus definieren wir die natürliche Geburtenrate b , die natürliche Sterberate d sowie die infektionsbedingte Sterberate a .

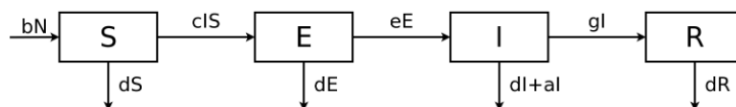


Abbildung 6.24: Das SEIR Modell

Wir wollen nun davon ausgehen, dass wir durch die Einführung einer Steuerungsfunktion in der Lage sind, von außen in das System einzugreifen. Im Gegensatz zu der Behandlung in [BPDP14] und [dPKM15], welche in ihrer Modellierung lediglich die Steuerbarkeit der Komponente S erlaubten, wollen wir im Rahmen dieses Kapitels zunächst die Möglichkeit der Steuerbarkeit aller Komponenten erlauben. Dafür nehmen wir an, dass verschiedene Behandlungsmöglichkeiten wie Impfstoffe, diverse Medikamente und Maßnahmen zur Steigerung des Immunsystems vorliegen. Dabei erscheint es sinnvoll, die Gabe eines Impfstoffes lediglich der Komponente S zur Verfügung zu stellen. Wir nehmen an, dass für diesen Impfstoff eine Effektivitätsrate von α_1 vorliegt ($\alpha_1 \in [0, 1]$) und modellieren dann mit $u_1(t)$ den Anteil der zum Zeitpunkt t geimpften Personen der Gruppe S . Für die Komponente E erlauben wir die Gabe eines Medikamentes mit der Effektivitätsrate α_2 . Zudem soll eine Gabe von Immunsystem stärkenden Mitteln (wie beispielsweise durch die Verabreichung von Vitaminen) mit der Effektivitätsrate α_4 den Ausbruch der Krankheit verhindern. Die Steuerungsfunktionen $u_2(t)$ und $u_4(t)$ beschreiben auch hier den Anteil der zum Zeitpunkt t durch die entsprechenden Maßnahmen behandelten Personen. Für die Komponente I erlauben wir schließlich ebenfalls die Behandlung mit einem weiteren Medikament, welches für die Elemente der Gruppe I die Effektivitätsrate α_3 habe. Auch hier beschreibt $u_3(t)$ den Anteil der zum Zeitpunkt t behandelten Individuen der Komponente I . Diese Überlegungen führen uns zu dem folgenden nichtlinearen Differentialgleichungs-

system:

$$\begin{aligned}
 \mathbf{S}'(\mathbf{t}) &= b\mathbf{N}(\mathbf{t}) - d\mathbf{S}(\mathbf{t}) - c\mathbf{S}(\mathbf{t})\mathbf{I}(\mathbf{t}) - \alpha_1\mathbf{u}_1(\mathbf{t})\mathbf{S}(\mathbf{t}), \\
 \mathbf{E}'(\mathbf{t}) &= c\mathbf{I}(\mathbf{t})\mathbf{S}(\mathbf{t}) - (e + d)\mathbf{E}(\mathbf{t}) - \alpha_2\mathbf{u}_2(\mathbf{t})\mathbf{E}(\mathbf{t}) - \alpha_4\mathbf{u}_4(\mathbf{t})\mathbf{E}(\mathbf{t}), \\
 \mathbf{I}'(\mathbf{t}) &= e\mathbf{E}(\mathbf{t}) - (g + a + d)\mathbf{I}(\mathbf{t}) - \alpha_3\mathbf{u}_3(\mathbf{t})\mathbf{I}(\mathbf{t}), \\
 \mathbf{R}'(\mathbf{t}) &= g\mathbf{I}(\mathbf{t}) - d\mathbf{R}(\mathbf{t}) + \alpha_1\mathbf{u}_1(\mathbf{t})\mathbf{S}(\mathbf{t}) + \alpha_2\mathbf{u}_2(\mathbf{t})\mathbf{E}(\mathbf{t}) + \alpha_4\mathbf{u}_4(\mathbf{t})\mathbf{E}(\mathbf{t}) + \alpha_3\mathbf{u}_3(\mathbf{t})\mathbf{I}(\mathbf{t}) \\
 &0 \leq \mathbf{u}_i(\mathbf{t}) \leq 1 \quad i = 1, 2, 3, 4,
 \end{aligned}$$

zu den Anfangswerten

$$S(0) = S_0, \quad E(0) = E_0, \quad I(0) = I_0 \quad \text{und} \quad R(0) = R_0.$$

Die Beziehung $N(t) = S(t) + E(t) + I(t) + R(t)$ erlaubt uns die folgende alternative Formulierung

$$(S) \left\{ \begin{array}{l}
 \mathbf{S}'(\mathbf{t}) = b\mathbf{N}(\mathbf{t}) - d\mathbf{S}(\mathbf{t}) - c\mathbf{S}(\mathbf{t})\mathbf{I}(\mathbf{t}) - \alpha_1\mathbf{u}_1(\mathbf{t})\mathbf{S}(\mathbf{t}), \\
 \mathbf{E}'(\mathbf{t}) = c\mathbf{I}(\mathbf{t})\mathbf{S}(\mathbf{t}) - (e + d)\mathbf{E}(\mathbf{t}) - \alpha_2\mathbf{u}_2(\mathbf{t})\mathbf{E}(\mathbf{t}) - \alpha_4\mathbf{u}_4(\mathbf{t})\mathbf{E}(\mathbf{t}), \\
 \mathbf{I}'(\mathbf{t}) = e\mathbf{E}(\mathbf{t}) - (g + a + d)\mathbf{I}(\mathbf{t}) - \alpha_3\mathbf{u}_3(\mathbf{t})\mathbf{I}(\mathbf{t}), \\
 \mathbf{N}'(\mathbf{t}) = (b - d)\mathbf{N}(\mathbf{t}) - a\mathbf{I}(\mathbf{t}) + \alpha_4\mathbf{u}_4(\mathbf{t})\mathbf{E}(\mathbf{t}), \\
 0 \leq \mathbf{u}_i(\mathbf{t}) \leq 1 \quad i = 1, 2, 3, 4,
 \end{array} \right.$$

zu den Anfangswerten

$$S(0) = S_0, \quad E(0) = E_0, \quad I(0) = I_0 \quad \text{und} \quad N(0) = N_0.$$

Das so eingeführte Modell soll nun die Grundlage für unsere weiteren Betrachtungen liefern. Im Gegensatz zu der Behandlung des Lotka-Volterra Modells soll unser Ziel nun aber nicht die Stabilisierung des Systems um den Gleichgewichtspunkt sein. Dies liegt in der Tatsache begründet, dass sich der Gleichgewichtspunkt des ungesteuerten Systems in Abhängigkeit zu den – dem Differentialgleichungssystem zugrundeliegenden – Parametern bestimmt und damit unabhängig von den Anfangswerten der zugrundeliegenden Population ist. Für das vorliegende Modell lässt sich der nicht-triviale Gleichgewichtspunkt beschreiben durch

$$\begin{aligned}
 S_G &= \frac{(e+d)(g+a+d)}{ce} & E_G &= \left(\frac{g+a+d}{e} \right) I_G \\
 I_G &= \frac{d(b-d)S_G}{ab-(b-d)S_G} & N_G &= \left(\frac{ad}{ab-c(b-d)} \right) I_G.
 \end{aligned}$$

In ungünstigen Konstellationen kann diese Setzung einerseits zu negativen Werten führen, andererseits kann beispielsweise auch der Wert N_G wesentlich kleiner als die betrachtete Ausgangspopulation N_0 sein und damit eine starke Reduzierung der Ausgangspopulation beschreiben. Wir nutzen dennoch die Ideen der bereits vorgestellten Stabilisierungstheorien, um das System in einen durch den Anwender festzulegenden Zielpunkt (S_z, E_z, I_z, N_z) zu steuern. Im Allgemeinen ist daher davon auszugehen, dass der anzusteuernde Punkt instabil ist. Dieses Vorgehen scheint sinnvoll, denn die eigentliche Problematik einer Epidemie besteht nicht in der Existenz der Infektion an sich, sondern in dem Maß der Ausbreitung. Sind wir in der Lage durch ein Eingreifen in das System die Infektion auf ein vertretbares Maß (welches beispielsweise die Kapazitätsgrenzen der Krankenhäuser und Intensivstationen deutlich unterschreitet) abzusenken und dort stabil zu halten, so können wir die Behandlung des epidemischen Verlaufes gegebenenfalls bereits als erfolgreich betrachten.

6.3.1 Formulierung der primalen Aufgabe

Ziel unserer Betrachtungen ist die Beantwortung der Frage, wie wir durch ein Eingreifen in das System einen beliebigen Punkt des Infektionsverhaltens exponentiell stabilisieren können. Dafür starten wir ähnlich wie bei der Betrachtung des Lotka-Volterra Modells mit einer Linearisierung um den Punkt (S_z, E_z, I_z, N_z) . Ziel der Steuerung ist damit nicht eine Auslöschung des Infektes, sondern die Steuerung in – und Stabilisierung um – ein akzeptables Maß an Infizierten. Im Sinne einer allgemeineren Behandlung des Problems wollen wir daher annehmen, dass die Steuerungsmaßnahmen wie Impfungen und Verabreichung von Immunsystem stärkenden Mitteln und Medikamenten auch dauerhaft zu dem durch den Anwender festzulegenden Maß $(u_1^z, u_2^z, u_3^z, u_4^z)$, mit $u_i^z \in [0, 1]$ erfolgen dürfen. Wir linearisieren daher die Dynamik (S) um das Tupel $(x_z, u_z) = ((S_z, E_z, I_z, N_z), (u_1^z, u_2^z, u_3^z, u_4^z))$, verschieben das erhaltene System schließlich durch die lineare Transformation

$$\tilde{S}(t) = S(t) - S_z, \quad \tilde{E}(t) = E(t) - E_z, \quad \tilde{I}(t) = I(t) - I_z, \quad \tilde{N}(t) = N(t) - N_z \quad \text{sowie}$$

$$\tilde{u}_1(t) = u_1(t) - u_1^z, \quad \tilde{u}_2(t) = u_2(t) - u_2^z, \quad \tilde{u}_3(t) = u_3(t) - u_3^z, \quad \tilde{u}_4(t) = u_4(t) - u_4^z$$

und erhalten damit das lineare System

$$\begin{aligned} \tilde{\mathbf{S}}'(t) &= -(d + cI_z + \alpha_1 u_1^z) \tilde{\mathbf{S}}(t) - cS_z \tilde{\mathbf{I}}(t) + b \tilde{\mathbf{N}}(t) - \alpha_1 S_z \tilde{\mathbf{u}}_1(t), \\ \tilde{\mathbf{E}}'(t) &= cI_z \tilde{\mathbf{S}}(t) - (e + d + \alpha_2 u_2^z + \alpha_4 u_4^z) \tilde{\mathbf{E}}(t) + cS_z \tilde{\mathbf{I}}(t) - \alpha_2 E_z \tilde{\mathbf{u}}_2(t) - \alpha_4 E_z \tilde{\mathbf{u}}_4(t), \\ \tilde{\mathbf{I}}'(t) &= e \tilde{\mathbf{E}}(t) - (g + a + d + \alpha_3 u_3^z) \tilde{\mathbf{I}}(t) - \alpha_3 I_z \tilde{\mathbf{u}}_3(t), \\ \tilde{\mathbf{N}}'(t) &= \alpha_4 u_4^z \tilde{\mathbf{E}}(t) + (b - d) \tilde{\mathbf{N}}(t) - a \tilde{\mathbf{I}}(t) + \alpha_4 E_z \tilde{\mathbf{u}}_4(t), \end{aligned}$$

$$\text{wobei} \quad -u_i^z \leq \tilde{\mathbf{u}}_i(t) \leq 1 - u_i^z \quad i = 1, 2, 3, 4.$$

Gemäß der Betrachtungen aus Kapitel [1.6](#) erreichen wir die exponentielle Stabilisierung des Systems in dem gewählten Punkt sowohl durch die Wahl des Integranden $\varphi(t, \tilde{\xi}, \tilde{v}) = \tilde{\xi}^T W \tilde{\xi} + \tilde{v}^T R \tilde{v}$ mit den Matrizen $W = R = I_{4 \times 4}$, als auch durch die Einführung der Gewichtsfunktion $\nu(t) = e^{\beta t}$, mit $\beta > 0$. Als primale Aufgabe (P_ν) betrachten wir daher mit den Setzungen $W \equiv R \equiv I_{4 \times 4}$,

$$A \equiv \begin{pmatrix} -(d + cI_z + \alpha_1 u_1^z) & 0 & -cS_z & b \\ cI_z & -(e + d + \alpha_2 u_2^z + \alpha_4 u_4^z) & cS_z & 0 \\ 0 & e & -(g + a + d + \alpha_3 u_3^z) & 0 \\ 0 & \alpha_4 u_4^z & -a & b - d \end{pmatrix}$$

und

$$B \equiv \begin{pmatrix} -\alpha_1 S_z & 0 & 0 & 0 \\ 0 & -\alpha_2 E_z & 0 & -\alpha_4 E_z \\ 0 & 0 & -\alpha_3 I_z & 0 \\ 0 & 0 & 0 & \alpha_4 E_z \end{pmatrix}$$

sowie $\tilde{x} = (\tilde{S}, \tilde{E}, \tilde{I}, \tilde{R})$ und $\tilde{u} = (\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4)$:

Minimiere

$$J(\tilde{x}, \tilde{u}) = \frac{1}{2} \int_0^\infty \left\{ \tilde{x}^T(t) W \tilde{x}(t) + \tilde{u}^T(t) R \tilde{u}(t) \right\} e^{\beta t} dt$$

bezüglich aller

$$(\tilde{x}, \tilde{u}) \in W_2^{1,4}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^4(\mathbb{R}_{++}, e^{\beta t}) \text{ mit } \beta > 0,$$

für die gilt

$$\tilde{x}'(t) = A\tilde{x}(t) + B\tilde{u}(t) \quad (\forall)t \in \mathbb{R}_{++}, \quad \tilde{x}(0) = \tilde{x}_0$$

und

$$\tilde{u}(t) \in \tilde{\mathcal{U}} = [-u_1^z, 1 - u_1^z] \times [-u_2^z, 1 - u_2^z] \times [-u_3^z, 1 - u_3^z] \times [-u_4^z, 1 - u_4^z] \quad (\forall)t \in \mathbb{R}_{++}.$$

Offensichtlich können wir – mit obiger Setzung – dieses Problem der Klasse (P_ν) zuordnen, so dass auch hier der Nachweis der Gültigkeit der Bedingungen (V1)-(V4) erlaubt, auf die Existenz einer optimalen Lösung für Probleme der Klasse (P_ν) zu schließen³³. Für das vorliegende Problem verifizieren wir leicht die Gültigkeit der Bedingungen (V1)-(V3). Aufgrund der Komplexität des Systems ist eine Verifizierung der Bedingung (V4) im Sinne der Betrachtung der vorangegangenen Beispiele nicht möglich. Dieser Umstand führt uns auf den Begriff der *Steuerbarkeit* nach Kalman. Wir wollen dies im Folgenden eingehender betrachten.

6.3.2 Steuerbarkeit des Systems

Für den Nachweis der Existenz einer optimalen Lösung der primalen Aufgabe fehlt uns der Nachweis der Gültigkeit der Voraussetzung (V4). Diese sichert, dass die zulässige Menge

$$\mathcal{B} := \left\{ (\tilde{x}, \tilde{u}) \in W_2^{1,4}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^4(\mathbb{R}_{++}, e^{\beta t}) \left| \begin{array}{l} \tilde{x}'(t) = A\tilde{x}(t) + B\tilde{u}(t) \quad (\forall)t \in \mathbb{R}_{++} \\ \tilde{x}(0) = \tilde{x}_0, \quad \tilde{u}(t) \in \tilde{\mathcal{U}} \quad (\forall)t \in \mathbb{R}_{++} \end{array} \right. \right\}$$

nichtleer und nicht einelementig ist. Da der Nachweis aufgrund der hohen Komplexität des Systems nicht so leicht zu führen ist, wollen wir zumindest Bedingungen angeben, welche sicherstellen, dass eine Steuerung existiert, die die beschriebene Systemdynamik in den angestrebten Zustand $\tilde{x}_z = (\tilde{S}_Z, \tilde{E}_Z, \tilde{I}_Z, \tilde{N}_Z)$ überführt.

Wir betrachten dafür das System Σ ³⁴ welches gegeben ist durch

$$x'(t) = f(x(t), u(t)) = Ax(t) + Bu(t) \quad \forall t \in \mathcal{T}, \quad \text{wobei } u(t) \in \mathcal{U} \subseteq \mathbb{R}^m \quad \forall t \in \mathcal{T}$$

mit $f: \mathbb{R}^n \times \mathcal{U} \rightarrow \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ und $\mathcal{T} \subseteq \mathbb{R}$ und führen zunächst den Begriff der Steuerbarkeit ein.

Definition 22 (Erreichbarkeit, Steuerbarkeit³⁵). Wir bezeichnen mit $(z, t) \in \mathbb{R}^n \times \mathcal{T}$ ein Ereignis und führen für dieses die folgenden Bezeichnungen ein:

Das Ereignis (z_2, τ) heißt erreichbar vom Ereignis (z_1, σ) (in Zeichen $(z_1, \sigma) \rightsquigarrow (z_2, \tau)$), wenn ein Pfad des Systems Σ auf dem Intervall $[\sigma, \tau] \subseteq \mathcal{T}$ mit dem Anfangswert z_1 und dem Endwert z_2 existiert, d.h. falls eine zulässige Steuerung $u(\cdot) \in \mathcal{U}$ existiert, so dass

$$z_2 = z_1 + \int_{\sigma}^{\tau} f(x(t), u(t)) dt.$$

Wir schreiben $z_1 \rightsquigarrow z_2$, wenn das Ereignis z_2 vom Ereignis z_1 für beliebige, nicht feste $\sigma, \tau \in \mathcal{T}$ erreicht werden kann. Gilt $z_1 \rightsquigarrow z_2$ für alle $z_1, z_2 \in \mathbb{R}^n$, dann nennen wir das System Σ (vollständig) steuerbar.

³³vgl. Voraussetzungen (V1)-(V4), S. 25 sowie Satz 9, S. 39

³⁴Für eine allgemeine Definition des Begriffs *System* vgl. [Son98], Definition 2.1.2, S. 26

³⁵vgl. [Son98], Definition 3.1.1., S. 79 und Definition 3.1.6, S. 81

Gilt $\mathcal{U} = \mathbb{R}^m$ und $f(\cdot, v) \in C^1$ für feste $v \in \mathcal{U}$, so genügt das System Σ der Definition³⁶ eines zeitstetigen, linearen Systems und ist zudem autonom. Für derartige Systeme können wir die folgende Bedingung der Steuerbarkeit formulieren:

Satz 26 ⁽³⁷⁾. *Das n -dimensionale, zeitstetige, lineare System Σ ist genau dann (vollständig) steuerbar, wenn für den Rang rg der Steuerbarkeitsmatrix*

$$\mathbf{R}(A, B) := [B, AB, A^2B, \dots, A^{n-1}B]$$

gilt $\text{rg}\mathbf{R}(A, B) = n$.

Wir überführen diese Ideen nun auf das vorliegende Problem. Statt der zulässigen Menge \mathcal{B} betrachten wir zunächst die (kleinere) Teilmenge

$$\hat{\mathcal{B}} := \left\{ (\tilde{x}, \tilde{u}) \in W_2^{2,4}(\mathbb{R}_{++}, e^{\beta t}) \times L_2^4(\mathbb{R}_{++}, e^{\beta t}) \mid (\tilde{x}, \tilde{u}) \in \mathcal{B} \right\}.$$

Die Raumzugehörigkeit der Funktion \tilde{x} in der Menge $\hat{\mathcal{B}}$ sichert die Existenz eines stetig differenzierbaren Repräsentanten. Damit genügt das der primalen Aufgabe zugrundeliegende System

$$\begin{aligned} \tilde{x}'(t) &= A\tilde{x}(t) + B\tilde{u}(t) \quad (\forall)t \in \mathbb{R}_{++}, \quad \tilde{x}(0) = \tilde{x}_0, \\ \tilde{u}(t) &\in \tilde{\mathcal{U}} \quad (\forall)t \in \mathbb{R}_{++}, \end{aligned}$$

über der Menge $\hat{\mathcal{B}}$ der Bedingung $f(\cdot, v) \in C^1$. Sei nun der Punkt x_0 hinreichend nah an dem Punkt $x_z = (S_Z, E_Z, I_Z, N_Z)$ ³⁸, so dass die Steuerungsbeschränkungen nicht aktiv werden. Für diesen Fall können wir zu einer Betrachtung des Systems über dem Steuerbereich $\hat{\mathcal{U}} = \mathbb{R}^4$ übergehen. Damit erfüllt das System die Definition eines zeitstetigen, linearen Systems. Wir können somit auf die Existenz einer Steuerungsfunktion schließen, die das System von einem Punkt x_0 , welcher hinreichend nah am Punkt x_z liegt, in den Punkt x_z überführt, falls

$$\text{rg}\mathbf{R}(A, B) := [B, AB, A^2B, A^3B] = 4.$$

Unter der getroffenen Annahme³⁹ sichert diese Bedingung zwar die Steuerbarkeit des Systems, liefert jedoch keine Aussagen über die Raumzugehörigkeiten der zugehörigen Trajektorien (x, u) und (\tilde{x}, \tilde{u}) . Da wir zudem – aufgrund der starken Differenzierbarkeitsbedingungen an die rechte Seite der Dynamik – jedoch nur Aussagen über die kleinere Teilmenge getroffen haben ist es möglich, dass Lösungen $(\tilde{x}, \tilde{u}) \in W_2^{1,4}(\mathbb{R}_{++}, e^{\beta t}) \times L_m^4(\mathbb{R}_{++}, e^{\beta t})$ existieren, welche eine Überführung des Systems in den Punkt $(\tilde{x}_z, \tilde{u}_z)$ bewirken, ohne dass die Rangbedingung aus Satz²⁶ erfüllt ist. Die Bedingung des Satzes²⁶ ist daher weder als notwendige noch als hinreichende Bedingung für die Existenz einer Lösung unserer primalen Aufgabe zu verstehen. Dennoch liefert diese Bedingung aufgrund des fehlenden Nachweises der Gültigkeit der Bedingung (V4) einen Indikator für die Existenz einer optimalen Lösung der primalen Aufgabe. Wir werden die Gültigkeit dieser Bedingung im Rahmen der numerischen Auswertung für eine feste Parametersetzung exemplarisch überprüfen. Den Nachweis der Existenz einer optimalen Lösung der primalen Aufgabe konnten wir jedoch durch den fehlenden Nachweis der Gültigkeit der Bedingung (V4) nicht erbringen.

³⁶ vgl. Definition²⁷ bis Definition²⁹, S. 186

³⁷ vgl. [Son98], Theorem 3, S. 87; diese Bedingung wird auch *Kalman-Bedingung* genannt

³⁸ d.h. $|\tilde{x}_0| = |x_0 - x_z| < \delta$ für hinreichend kleines δ

³⁹ x_0 sei hinreichend nah an x_z , bzw. \tilde{x}_0 sei hinreichend nah an 0

6.3.3 Übergang zur dualen Aufgabe

Unter der (nicht nachgewiesenen) Annahme der Gültigkeit der Voraussetzung (V4) können wir – gemäß der Betrachtungen aus Kapitel 4 – die zu (P_ν) duale Aufgabe angeben durch (D_ν) :

Minimiere

$$G(y, q) = \int_0^\infty \left\{ \frac{1}{2} q^T(t) W^{-1} q(t) + \theta(t, y(t)) \right\} e^{-\beta t} dt + \tilde{x}_0^T y(0)$$

bezüglich aller

$$(y, q) \in W_2^{2,4}(\mathbb{R}_{++}, e^{-\beta t}) \times W_2^{1,4}(\mathbb{R}_{++}, e^{-\beta t}) \text{ mit } \beta > 0,$$

für die gilt

$$y'(t) = -A^T y(t) + q(t) \quad (\forall) t \in \mathbb{R}_{++}, q(0) = \tilde{x}^0$$

wobei

$$\theta(t, y(t)) = \sum_{i=1}^4 -\frac{1}{2} \sigma_i^T(t, y(t)) \sigma_i(t, y(t)) + \sigma_i^T(t, y(t)) \sum_{k=1}^4 B_{ki} y_k(t) \quad \text{und}$$

$$\sigma_i(t, y(t)) = \min \left\{ \max \left\{ \tilde{u}_i, \sum_{k=1}^4 B_{ki} y_k(t) e^{-\beta t} \right\}, \bar{u}_i \right\} e^{\beta t} \quad \text{ist.}$$

Für das SEIR Modell ergeben sich damit die folgenden Schaltfunktionen σ_i :

$$\sigma_1(t, y(t)) = \min \left\{ \max \left\{ -u_1^z, -\alpha_1 S_z y_1(t) e^{-\beta t} \right\}, 1 - u_1^z \right\} e^{\beta t},$$

$$\sigma_2(t, y(t)) = \min \left\{ \max \left\{ -u_2^z, (-\alpha_2 E_z y_2(t) - \alpha_4 E_z y_4(t)) e^{-\beta t} \right\}, 1 - u_2^z \right\} e^{\beta t},$$

$$\sigma_3(t, y(t)) = \min \left\{ \max \left\{ -u_3^z, -\alpha_3 I_z y_3(t) e^{-\beta t} \right\}, 1 - u_3^z \right\} e^{\beta t},$$

$$\sigma_4(t, y(t)) = \min \left\{ \max \left\{ -u_4^z, \alpha_4 E_z y_4(t) e^{-\beta t} \right\}, 1 - u_4^z \right\} e^{\beta t}.$$

Können wir die Gültigkeit der Voraussetzungen (VD1)-(VD4) nachweisen, so liefert die Aussage des Satzes 22 – unter Annahme der Existenz einer optimalen Lösung des primalen Problems – hinreichende Optimalitätsbedingungen⁴⁰. Für die obige duale Aufgabe lassen sich die Bedingungen (VD1)-(VD3) leicht verifizieren. Für den Nachweis der Gültigkeit der Voraussetzung (VD4) genügt nun wieder die Angabe mindestens zweier zulässiger Lösungen. Die Existenz solcher Lösungen lässt sich im Falle einer regulären Matrix A aufgrund der fehlenden Steuerungsbeschränkungen für die duale Steuerungsfunktion q sowie der vorliegenden Funktionenräume in Form gewichteter Sobolev-Räume mit echter Dichtefunktion leicht garantieren:

Setzen wir $q \equiv (\tilde{S}_0, \tilde{E}_0, \tilde{I}_0, \tilde{N}_0)^T$ sowie $y' \equiv (0, 0, 0, 0)^T$ dann erhalten wir aufgrund der Regularität⁴¹ der Matrix A eine eindeutige Lösung $(c_1, c_2, c_3, c_4)^T$, mit $c_i \in \mathbb{R}$, durch Einsetzen des Ansatzes $y \equiv (c_1, c_2, c_3, c_4)^T$ in das System

$$y'(t) = -A^T y(t) + q(t).$$

⁴⁰vgl. hierfür Voraussetzungen (VD1)-(VD4), S. 76 sowie Satz 22, S. 104

⁴¹Diese Eigenschaft ist abhängig von den dem Modell zugrundeliegenden Parametern und ist im Vorfeld für das jeweilige Modell zu überprüfen.

Eine zweite zulässige Lösung erhalten wir schließlich durch die Ansätze

$$q = (t + \tilde{S}_0, t + \tilde{E}_0, t + \tilde{I}_0, t + \tilde{N}_0)^T, \quad y' = (d_1, d_2, d_3, d_4)^T$$

sowie

$$y = (d_1 \cdot t + c_1, d_2 \cdot t + c_2, d_3 \cdot t + c_3, d_4 \cdot t + c_4)^T, \quad \text{mit } c_i, d_i \in \mathbb{R}.$$

Damit sind die Voraussetzungen (VD1)-(VD4) erfüllt. Satz 22 liefert damit im Falle der Gültigkeit der Voraussetzung (V4) die Gültigkeit der Beziehung

$$J(\tilde{x}_*, \tilde{u}_*) = \sup_{(y,q) \in \mathcal{B}_D} G(y, q).$$

Gelingt zudem der Nachweis der Gültigkeit der Bedingung $\tilde{Y}_{\nu_i} \neq \emptyset$, so ist Satz 23 anwendbar und liefert in diesem Falle schließlich den Zusammenhang zur eindeutigen optimalen Lösung der primalen Aufgabe (P_ν) durch⁴²

$$\tilde{x}_*(t) = q_*(t)e^{-\beta t} \quad \text{und} \quad \tilde{u}_*(t) = \text{Proj}_{[\underline{u}, \bar{u}]} \left(B^T(t)y_*(t)e^{-\beta t} \right).$$

Für hinreichend kleine Anfangswerte \tilde{x}_0 ist anzunehmen, dass die Steuerungsbeschränkungen der primalen Aufgabe nicht aktiv werden. In diesem Fall genügt für den Nachweis der Bedingung $\tilde{Y}_{\nu_i} \neq \emptyset$ der Nachweis der Existenz einer Lösung y der Differentialgleichung

$$y''(t) = \beta q(t) - A^T(t)y'(t) + A(t)q(t) + B(t)\hat{\sigma}(t, y(t))e^{-\beta t} \quad (\forall) t \in \mathbb{R}_{++}$$

mit $q(t) = y'(t) + A^T y(t)$, welche der Bedingung $\tilde{x}_0 = q(0)$ genügt. Für das vorliegende Beispiel des SEIR Modells können wir die obige Differentialgleichung für hinreichend kleine Anfangswerte \tilde{x}_0 durch

$$y''(t) = \left(2 \cdot I_{4 \times 4} + A - A^T \right) y'(t) + \left[(2 \cdot I_{4 \times 4} + A) \cdot A^T + B \cdot B^T \right] y(t), \quad (6.33)$$

mit $\tilde{x}_0 = y'(0) + A^T y(0)$ angeben. Setzen wir

$$C := \left(2 \cdot I_{4 \times 4} + A - A^T \right) \quad \text{und} \quad D := \left[(2 \cdot I_{4 \times 4} + A) \cdot A^T + B \cdot B^T \right],$$

so können wir das lineare System 2. Ordnung (6.33) durch die Setzungen

$$z_1 := y_1, \quad z_2 := y_2, \quad z_3 := y_3, \quad z_4 := y_4, \quad z_5 := y'_1, \quad z_6 := y'_2, \quad z_7 := y'_3, \quad z_8 := y'_4,$$

in ein lineares System 1. Ordnung der Form

$$\mathbf{z}'(t) = \underbrace{\begin{pmatrix} 0 & I_{4 \times 4} \\ D & C \end{pmatrix}}_{=: \mathbf{M}} \cdot \mathbf{z}(t), \quad \text{mit} \quad \mathbf{z}(0) = \begin{pmatrix} \hat{\mathbf{z}}_0 \\ \tilde{x}_0 - A^T \hat{\mathbf{z}}_0 \end{pmatrix} \quad (6.34)$$

überführen. Dabei ist $\hat{\mathbf{z}}_0 = (z_1^0, z_2^0, z_3^0, z_4^0)^T$ ein Vektor der Dimension 4 mit den freien Variablen $z_1^0, z_2^0, z_3^0, z_4^0 \in \mathbb{R}$. Aufgrund der konstanten Koeffizienten sichert der Satz von Picard-Lindelöf die Existenz einer stetig differenzierbaren und globalen Lösung \mathbf{z} des Systems (6.34) auf ganz \mathbb{R}_+ . Diese Lösung \mathbf{z} ist damit für alle $T \in \mathbb{R}_{++}$ ein Element

⁴²vgl. Satz 23, S. 104

des Raumes $W_2^{1,8}(0, T)$. Die Lösungsstruktur des Differentialgleichungssystems (6.34) ist durch eine Darstellung der Form

$$\mathbf{z}(t) = \sum_{j=1}^k \left\{ \sum_{l_j=1}^{m_j} c_{l_j} \cdot e^{\lambda_j \cdot t} \cdot p_{l_j}(t) \right\}$$

gegeben, wobei $\lambda_j = a_j + b_j \cdot i \in \mathbb{C}$ mit $j = 1, \dots, k$ die k verschiedenen Eigenwerte der Matrix $\mathbf{M} \in \mathbb{R}^{8 \times 8}$ mit den algebraischen Vielfachheiten m_j beschreiben, $p_{l_j}(\cdot)$ für $l_j = 1, \dots, m_j$ Polynome vom Grad $\leq m_j - 1$ und c_{l_j} noch zu bestimmende reelle Koeffizienten sind.

Existieren für die Matrix \mathbf{M} vier verschiedene $\lambda_j = a_j + b_j \cdot i$ (wobei wir λ_j und $\bar{\lambda}_j = a_j - b_j \cdot i$ an dieser Stelle als zwei verschiedene Eigenwerte interpretieren) mit $2a_j - \beta < 0$, dann können wir weiterhin auf die Inklusion $z \in W_2^{1,8}(\mathbb{R}_{++}, e^{-\beta t})$ schließen, denn:

Seien o.B.d.A. durch $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ die Eigenwerte gegeben, welche der Bedingung $(2a_j - \beta) < 0$ genügen. In diesem Falle setzen wir $c_{l_j} = 0$ für alle $l_j \geq 5$. Damit ergibt sich die Lösungsstruktur des Systems (6.34) zu

$$\mathbf{z}(t) = \sum_{j=1}^4 \left\{ \sum_{l_j=1}^{m_j} c_{l_j} \cdot e^{\lambda_j \cdot t} \cdot p_{l_j}(t) \right\}.$$

Die übrigen vier Koeffizienten c_{l_j} erhalten wir nun aus der Anfangsbedingung

$$\mathbf{z}(0) = \sum_{j=1}^4 \left\{ \sum_{l_j=1}^{m_j} c_{l_j} \cdot v_{l_j}(0) \right\} = \begin{pmatrix} \hat{\mathbf{z}}_0 \\ x_0 - A^T \hat{\mathbf{z}}_0 \end{pmatrix},$$

wobei v_{l_j} die zum Eigenwert λ_j zugehörigen Eigenvektoren bezeichnen. Dabei genügt die Auswertung der untersten vier Gleichungen des entstehenden Gleichungssystems⁴³, welche aufgrund der linearen Unabhängigkeit der Eigenvektoren v_{l_j} eindeutig lösbar sind.

Gilt $z \in W_2^{1,8}(\mathbb{R}_{++}, e^{-\beta t})$, so impliziert dies die Existenz einer Lösung $y \in W_2^{1,4}(\mathbb{R}_{++}, e^{-\beta t})$ des Systems (6.33) und damit die Existenz eines Paares $(y, q) \in \tilde{Y}_{\nu_1}$ und wir können schließen: Existieren für die Matrix \mathbf{M} vier verschiedene $\lambda_j = a_j + b_j \cdot i$ mit $2a_j - \beta < 0$, dann ist $\tilde{Y}_{\nu_1} \neq \emptyset$.

Bemerkung 29. Sind die Anfangswerte nicht in einer hinreichend kleinen Umgebung des Zielpunktes, so schlagen wir beispielsweise ein vorgeschaltetes Optimalsteuerungsproblem über dem endlichen Zeithorizont mit festgelegten Endmengen vor.

6.3.4 Anwendung der direkten Pseudospektralmethode

Für die numerische Behandlung mittels der in Kapitel 5.1 vorgestellten direkten, modifizierten Pseudospektralmethode gehen wir mit Hilfe der Setzung

$$q(t) := y'(t) + A^T y(t) \quad (\forall) t \in \mathbb{R}_{++}$$

⁴³Der Vektor $\hat{\mathbf{z}}_0$ ist frei wählbar, da die Menge \tilde{Y}_{ν_1} kein klassisches Anfangswertproblem darstellt, sondern Freiheitsgrade besitzt.

über zu der Formulierung der dualen Aufgabe in Form eines Variationsproblems. Diese Setzung führt uns zu der Aufgabe (D_ν)

Minimiere

$$G(y, y') := \int_0^\infty \left\{ \frac{1}{2} [y'(t) + A^T y(t)]^T [y'(t) + A^T y(t)] + \theta(t, y(t)) \right\} e^{-\beta t} dt + x_0^T y(0)$$

bezüglich aller

$$y \in W_2^{2,4}(\mathbb{R}_{++}, e^{-\beta t}) \quad \text{mit} \quad x_0 = y'(0) + A^T y(0) \quad \text{und} \quad \beta > 0,$$

wobei

$$\theta(t, y(t)) = \sum_{i=1}^4 -\frac{1}{2} \sigma_i^T(t, y(t)) \sigma_i(t, y(t)) + \sigma_i^T(t, y(t)) \sum_{k=1}^4 B_{ki} y_k(t) \quad \text{und}$$

$$\sigma_i(t, y(t)) = \min \left\{ \max \left\{ \tilde{u}_i, \sum_{k=1}^4 B_{ki} y_k(t) e^{-\beta t} \right\}, \bar{u}_i \right\} e^{\beta t} \quad \text{ist.}$$

Entsprechend des Vorgehens in Kapitel [5.2](#) wählen wir nun die Ansätze

$$\begin{aligned} y_N^{(1)}(t) &= \sum_{k=1}^N a_k S_k(t) + E_1 \cdot t + D_1 & y_N'^{(1)}(t) &= \sum_{k=1}^N a_k T_k(t) + E_1 \\ y_N^{(2)}(t) &= \sum_{k=1}^N b_k S_k(t) + E_2 \cdot t + D_2 & y_N'^{(2)}(t) &= \sum_{k=1}^N b_k T_k(t) + E_2 \\ y_N^{(3)}(t) &= \sum_{k=1}^N c_k S_k(t) + E_3 \cdot t + D_3 & y_N'^{(3)}(t) &= \sum_{k=1}^N c_k T_k(t) + E_3 \\ y_N^{(4)}(t) &= \sum_{k=1}^N d_k S_k(t) + E_4 \cdot t + D_4 & y_N'^{(4)}(t) &= \sum_{k=1}^N d_k T_k(t) + E_4 \end{aligned}$$

und setzen diese in das Variationsproblem ein. Durch dieses Vorgehen erhalten wir das endlich dimensionale, konvexe Minimierungsproblem

$$G^N(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{E}, \mathbf{D}) := \int_0^\infty \tilde{g}_N(t, \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{E}, \mathbf{D}) e^{-\beta t} dt + x_0^T \mathbf{D}$$

bezüglich aller

$$(\mathbf{a}, \mathbf{b}, \mathbf{E}, \mathbf{D}) \in \mathbb{R}^{4N+8},$$

für die gilt

$$x_0 = \mathbf{E} + A^T \mathbf{D}.$$

Wir überführen die restringierte Aufgabe durch einen Lagrange-Ansatz in ein unrestringiertes, konvexes und endlich dimensionales Optimierungsproblem der Form

Minimiere

$$\int_0^\infty \tilde{g}_N(t, \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{E}, \mathbf{D}) e^{-\beta t} dt + x_0^T \mathbf{D} + \lambda^T [x_0 - \mathbf{E} - A^T \mathbf{D}]$$

bezüglich aller

$$(\mathbf{a}, \mathbf{b}, \mathbf{E}, \mathbf{D}, \lambda) \in \mathbb{R}^{4N+12}.$$

Aufgrund der Konvexität sind alle Lösungen dieses endlich dimensionalen Optimierungsproblems globale Lösungen und der Zielfunktionswert ist eindeutig bestimmt.

Gemäß der Betrachtungen aus Kapitel 5.2 erfüllt die Optimallösung die Kollokationsgleichungen der direkten Methode. Diese liefern ein System aus $4N + 12$ Gleichungen mit $4N + 12$ Variablen.

6.3.5 Auswertung der numerischen Ergebnisse

Wir wollen die direkte Pseudospektralmethode für das oben eingeführte Modell exemplarisch anhand einer festen Konstellation von Parametern anwenden. Dazu ziehen wir das Beispiel der Ausbreitung der Covid-19 Infektion in Deutschland heran und legen folgende Daten zu Grunde⁴⁴:

Laut Informationen des jährlichen Berichtes [SBD20] des Statistischen Bundesamtes lag die Zahl der Lebendgeborenen im Jahr 2018 bei 787 523 und die Anzahl der Verstorbenen bei 954 874. Dies entspricht einer durchschnittlichen Anzahl von 196 880.75 Geburten und 238 718.5 Todesfällen pro 3 Monaten⁴⁵. Dem gegenüber wird eine Gesamtbevölkerungszahl von 83 019 200 angegeben. Aus diesen Kennzahlen erhalten wir eine grobe Schätzung für die natürliche Geburtenrate $b \approx 0.002371$ und die natürliche Sterberate $d \approx 0.002875$ in 3 Monaten.

Für die Parameter, welche den Krankheitsverlauf beschreiben, orientieren wir uns an den durch die *Tagesschau* veröffentlichten Daten mit dem Stand vom 27.04.2020⁴⁶. Gemäß dieser Quelle beläuft sich die Anzahl der bestätigten Infektionen am 27.04.2020 auf 157 770, wovon 5 976 Personen zu diesem Zeitpunkt bereits verstorben sind⁴⁷ und 114 500 als genesen geschätzt werden. Damit ist von einem Wert von 37 294 aktiven Fällen, d.h. von einem Prozentsatz von etwa 0.045 der Gesamtbevölkerung auszugehen. Unter der Annahme, dass nach einer Infektion mit dem Erreger Covid-19 von einer Immunität auszugehen ist ergeben sich auf Grundlage der zitierten Daten folgende Ansätze für die Wahl der Parameter $a \approx 0.0496$ und $g \approx 0.95$ ⁴⁸. Für unsere Modellierung wollen wir weiterhin annehmen, dass das Virus hoch ansteckend ist und etwa die Hälfte aller Infizierten infektiös werden. Daher setzen wir den Parameter c auf den Wert 0.9 und den Parameter e auf den Wert 0.5. Insgesamt ergibt sich damit die folgende Wahl der Parameter:

⁴⁴Die derartig bestimmten Parameter dienen nur einer Annäherung an realistische Verhältnisse, um ein erstes Gefühl für das Ausbreitungsverhalten einer Epidemie zu erhalten und sind ausdrücklich nicht als tatsächlich verwertbare Parameter zu verstehen.

⁴⁵Am 27.01.2020 wurde in Deutschland der erste Corona-Fall erfasst. Die Zahlen zur Beschreibung der Covid-19 Pandemie beziehen sich auf die Erhebung der Daten vom 27.04.2020 und daher aus deutscher Sicht auf einen Zeitraum von 3 Monaten. Aus diesem Grund passen wir die Parameter der Geburten- und natürlichen Sterberate auf den selben Zeitraum an.

⁴⁶vgl. [Tag20]

⁴⁷Anmerkung: Diese Zahl spiegelt die Zahl der gestorbenen Personen wider, welche den Erreger Covid-19 nachweislich zum Todeszeitpunkt in sich trugen.

⁴⁸Als Grundlage für die Berechnung dieser Werte setzen wir die Zahl der nicht mehr aktiven Fälle (in Höhe von 120 476) an.

Parameter	Beschreibung	Wert
c	Ansteckungskoeffizient	0.9
e	Rate der Neu-Infektiösen	0.5
g	Immunsierungsrate	0.95
b	natürliche Geburtenrate	0.002371
d	natürliche Sterberate	0.002875
a	krankheitsbedingte Sterberate	0.0496

Bemerkung 30. Für die Wahl der gegebenen Parameter berechnet sich der Gleichgewichtspunkt des ungesteuerten Systems zu

$$\begin{aligned} S_G &= 1.120265812500000 & E_G &= -0.004770582250950 \\ I_G &= -0.002379402105264 & N_G &= -5.940157384200490 \cdot 10^{-4} \end{aligned}$$

und ist somit als unrealistisch sowie wenig wünschenswert anzunehmen.

Für die Beschreibung der einzelnen Komponenten S , E , I , N setzen wir prozentuale Werte an. Daher sei der Anfangsbestand der Gesamtpopulation auf $N_0 = 100$ normiert. Ausgangspunkt unserer Steuerung sei ein Zeitpunkt $t = 0$, an welchem sowohl ein Impfstoff, als auch wirksame Medikamente existieren. Die Effektivitätsrate aller Medikamente sei der Einfachheit halber mit $\alpha_i = 1$ anzunehmen. Als Ziel definieren wir die Stabilisierung des Systems um den Punkt

$$((S_z, E_z, I_z, N_z), (u_1^z, u_2^z, u_3^z, u_4^z)) = ((30, 0.001, 0.044, 100), (0.1, 0.1, 0.1, 0.1)).$$

Damit gehen wir davon aus, dass die medikamentöse Situation soweit ausgebildet sei, dass das Gesundheitssystem der Belastung von einem Anteil an Infizierten von 0.045% der Gesamtbevölkerung dauerhaft standhält und eine Immunität von 70% der Bevölkerung gefordert sei. Zudem erlauben wir mit dieser Setzung einen dauerhaften Einsatz der Medikamente zu einer Rate von 0.1 und fordern eine konstante Gesamtbevölkerung. Diese Zielsetzung erfolgt dabei auf Grundlage der Tatsache, dass das deutsche Gesundheitssystem zur Zeit des 27.04.2020 deutlich unter der Belastungsgrenze stand und demnach auch dauerhaft einer derartigen Belastung standhalten könnte. Zudem deuten erste Studien zum Covid-19 Erreger auf eine Herdenimmunität bei einem Immunitätsanteil von 50 – 67% der Bevölkerung⁴⁹ hin. Damit liegt unser Hauptaugenmerk auf der Sicherung der Herdenimmunität. Die Setzung von u_z sichert darüber hinaus, dass für das transformierte System die Bedingung $0 \in \text{int } \tilde{U}$ und somit die Gültigkeit der Bedingung (VD3)⁵⁰ gesichert ist.

Für die derartige Setzung der Parameter lässt sich leicht nachweisen, dass die Rangbedingungen des Satzes ²⁶ erfüllt sind⁵¹ und die dem Modell zugrunde liegende Matrix A regulär ist. Für hinreichend kleine $\delta_i > |x_{0i} - x_{zi}|$ mit $i = 1, 2, 3, 4$ können wir damit davon ausgehen, dass das System Σ steuerbar ist und somit eine Mindestanforderung für die Existenz einer asymptotisch stabilisierenden Steuerung gegeben ist.

⁴⁹vgl. beispielsweise ^{Töd20}, S. 11

⁵⁰siehe S. ⁷⁶

⁵¹siehe Satz ²⁶, S. ¹⁶³

Das freie System

Wir beginnen mit der Betrachtung eines Szenarios, in welchem wir den Anfangswert x_0 derart nah an dem Zielpunkt x_z modellieren, dass die Steuerungsbeschränkungen nicht aktiv werden. Wir betrachten dafür folgende Konstellation an Anfangswerten:

Parameter	Beschreibung	Wert
S_0	Anfangsbestand S	30.01
E_0	Anfangsbestand E	0.001
I_0	Anfangsbestand I	0.044
N_0	Anfangsbestand N	100

Die (im Vergleich zu den zuvor betrachteten Systemen) höhere Komplexität des zu lösenden Gleichungssystems erfordert für eine sinnvolle Auswertung eine höhere Anzahl an Kollokationspunkten. Für die numerische Behandlung der zuvor betrachteten Beispiele gelang eine Auswertung für $N > 11$ Kollokationspunkte jedoch nur durch die Verwendung der Funktion *vpa* (*Variable-precision arithmetic*), welche eine Erhöhung der voreingestellten Rechengenauigkeit in Matlab erlaubt. Im Rahmen der numerischen Experimente stellte sich jedoch heraus, dass die Laufzeit des verwendeten Matlab-Solvers 'fsolve' bei einer Erhöhung der Rechengenauigkeit für das vorliegende Modell deutlich anstieg. Daher soll für das vorliegende Modell eine Betrachtung unter Nutzung der voreingestellten doppelten Rechengenauigkeit für Fließkommazahlen (64 Bit) genügen. Diese Einschränkung erlaubt jedoch lediglich die Betrachtung von maximal $N = 11$ Kollokationspunkten. Dies begründet sich in der Darstellung der verwendeten Gewichte

$$\begin{aligned}
 w_1^{(11)} &= 0.142466606447100 & w_2^{(11)} &= 0.194860444763925 & w_3^{(11)} &= 0.116390915924496 \\
 w_4^{(11)} &= 0.038282226773098 & w_5^{(11)} &= 0.007196641383675 & w_6^{(11)} &= 0.000759440423242 \\
 w_7^{(11)} &= 0.000042565612177 & w_8^{(11)} &= 0.000001146201940 & w_9^{(11)} &= 0.000000012431769 \\
 w_{10}^{(11)} &= 0.000000000038563 & w_{11}^{(11)} &= 0.000000000000014.
 \end{aligned}$$

Die Erhöhung um einen weiteren Kollokationspunkt lieferte unter Nutzung der voreingestellten Rechengenauigkeit als letzten Gewichtungparameter $w_{12}^{(12)} = 0.000000000000000$. Erhöhen wir die Rechengenauigkeit, so liefert Matlab hingegen den Wert

$$w_{12}^{(12)} = 0.00000000000000040740387337131208412365934\dots$$

Eine Auswertung der numerischen Ergebnisse für $N \geq 12$ Kollokationspunkte erscheint für die Berechnung unter der voreingestellten Rechengenauigkeit daher nicht sinnvoll. In den nachfolgenden Abbildungen sind die numerischen Lösungen der direkten Pseudospektral­methode für die Wahl des Gewichtungparameters $\beta = 2$ für $N = 10$ (—) und $N = 11$ (—) Kollokationspunkten dargestellt:

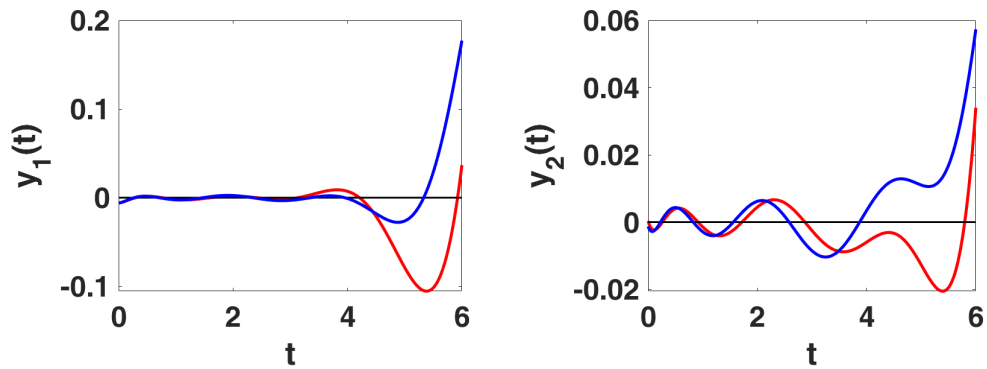


Abbildung 6.25: Näherungslösung der dualen Zustände $y_1(t)$ und $y_2(t)$ für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

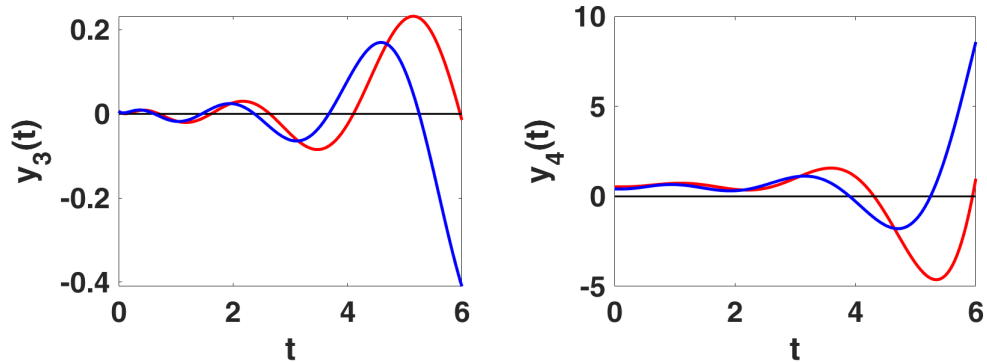


Abbildung 6.26: Näherungslösung der dualen Zustände $y_3(t)$ und $y_4(t)$ für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

Eine vergleichende Darstellung der Anfangswerte der dualen Zustandsfunktionen ist der Tabelle 6.11 zu entnehmen. Die größte Differenz beim Vergleich der Anwendung des Verfahrens mit $N = 10$ und $N = 11$ Kollokationspunkten findet sich in der Auswertung des Anfangswertes des dualen Zustandes y_4 mit einer Differenz von weniger als 0.121. Die scheinbar größte Genauigkeit zwischen zwei approximierten Lösungen lieferte die Auswertung des Anfangswertes des dualen Zustandes y_1 mit einer Differenz von weniger als 10^{-4} .

N	$y_N^{(1)}(0)$	$y_N^{(2)}(0)$	$y_N^{(3)}(0)$	$y_N^{(4)}(0)$
10	-0.006102119084	$3.398340281440 \cdot 10^{-4}$	0.007140541183	0.527839897874
11	-0.006052277422	-0.001272636153	0.006157620053	0.406860243082

Tabelle 6.11: Numerisch ermittelte Anfangswerte der dualen Zustände für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

Unterstellen wir die Gültigkeit der Bedingungen des Satzes 23, so erhalten wir Rückschlüsse

auf die Lösungen der primalen Aufgabe über die Setzungen⁵²

$$\tilde{x}(t) = \hat{W}^{-1}(t) \left[y'_N(t) + A^T y_N(t) \right] \quad \text{und} \quad u(t) = \text{Proj}_{\tilde{U}} \left(B^T y_N(t) e^{-\beta t} \right) \quad (6.35)$$

und können somit auf die nachstehenden Abbildungen 6.27 bis 6.30 für die approximierten Lösungen der primalen Zustands- und Steuerungsfunktionen für eine wachsende Anzahl an Kollokationspunkten schließen. Eine Zeiteinheit t entspricht aufgrund der Modellierung der Dauer von 3 Monaten.

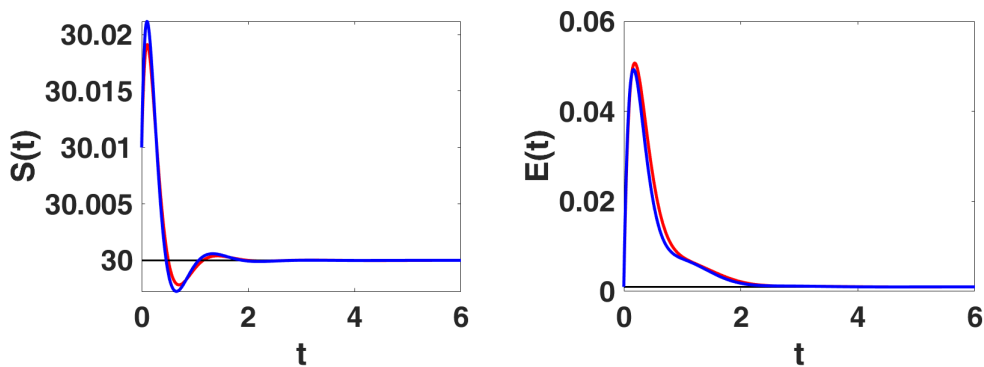


Abbildung 6.27: Näherungslösung der Komponenten S_N und E_N für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

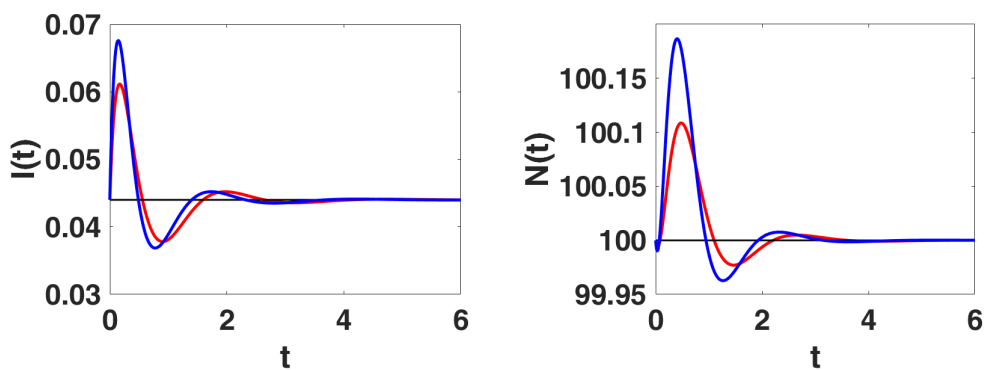


Abbildung 6.28: Näherungslösung der Komponenten I_N und N_N für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

⁵²vgl. Satz 23, S. 104

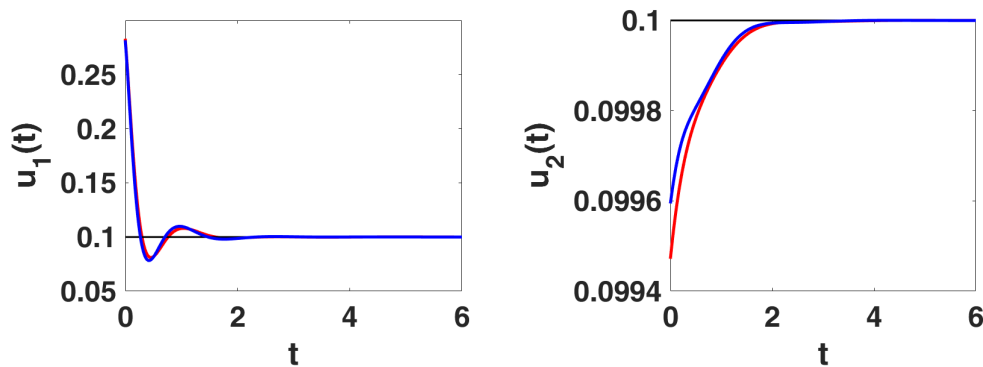


Abbildung 6.29: Näherungslösung der Steuerungen $u_1(t)$ und $u_2(t)$ für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

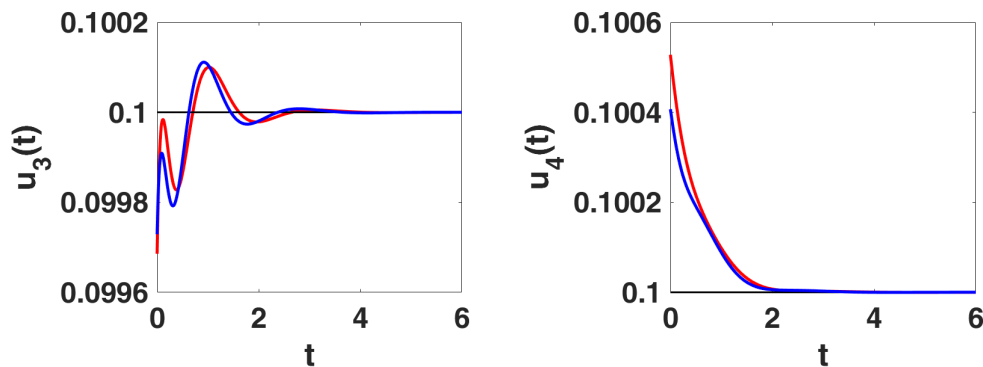


Abbildung 6.30: Näherungslösung der Steuerungen $u_3(t)$ und $u_4(t)$ für $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| \leq 0.01$ für $N = 10$ (—) und $N = 11$ (—)

Die obigen Abbildungen lassen eine Konvergenz des Verfahrens vermuten. Diese Annahme spiegelt sich zudem in den Zielfunktionalwerten wider. Für $N = 10$ liefert der Algorithmus einen Zielfunktionalwert von $G(y_{N*}, y'_{N*}) = 0.012788171984766$. Eine Erhöhung um einen Kollokationspunkt zu $N = 11$ liefert den Wert $G(y_{N*}, y'_{N*}) = 0.015333688014868$ und stellt somit eine Differenz der Zielfunktionalwerte für die verschiedenen Kollokationspunkte von weniger als 10^{-2} dar. Aufgrund des fehlenden Nachweises der Voraussetzungen aus Satz 23 ist jedoch die Existenz einer optimalen Lösung der primalen und dualen Aufgaben sowie die Gültigkeit der starken Dualitätsbeziehung – und damit die Gültigkeit des Zusammenhanges (6.35) – unklar. Gemäß unserer Konstruktion können wir im Fall der Existenz einer Lösung (x_*, u_*) der primalen Aufgabe jedoch auf die Gültigkeit der Relation

$$\sup_{y \in Y_{\nu_1}} G_{\nu_1}(y, y') \leq \inf_{(x, u) \in \mathcal{B}_P} J(x, u) = J(x_*, u_*)$$

schließen⁵³. Der Vergleich der approximierten Zielfunktionalwerte der dualen Aufgabe und der primalen Aufgabe⁵⁴ legt die Vermutung nahe, dass im Falle der Existenz einer Dualitätslücke, diese jedoch nicht größer als $5 \cdot 10^{-3}$ ist und die erhaltenen approximierten

⁵³vgl. Satz 17, S. 68

⁵⁴vgl. Tabelle 6.12

Lösungen (y_N, y'_N) somit eine gute obere Schranke für die optimalen Lösungen der primalen Aufgaben liefern.

N	$G^N(y_N, y'_N)$	$J^N(\tilde{x}_N, u_N)$	$G^N(y_N, q_N) - J^N(\tilde{x}_N, u_N)$
10	0.012788171984766	0.017221821975220	0.004433649990453
11	0.015333688014868	0.019338204882494	0.004004516867625

Tabelle 6.12: Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

Das beschränkte System

Wir betrachten nun ein Szenario, bei welchem die Steuerungsbeschränkungen aktiv werden. Damit ist eine Auswertung der Bedingung aus Satz 26 nicht mehr möglich⁵⁵. Dennoch interessiert uns das derzeitige Szenario, bei welchem der Anteil der Immunen deutlich von der gewünschten Herdenimmunität von 50 – 67% abweicht. Ausgehend von der kumulativen Anzahl an nachweislich Infizierten in Höhe von 157 770 Personen (mit Stand des 27.04.2020) können wir bei einer Gesamtbevölkerung in Höhe von 83 019 200 Personen derzeit von einer Immunität von 0.19% der Bevölkerung (innerhalb von Deutschland) ausgehen. Im Rahmen unseres Modells wollen wir daher für den Zeitpunkt $t = 0$, an welchem sowohl Impfstoff als auch wirksame Medikamente zur Bekämpfung des Covid-19 Erregers vorliegen, von einem (optimistisch angesetzten) Anfangsbestand der Komponente S in Höhe von 99% der Gesamtpopulation ausgehen. Die weiteren Werte sollen unverändert bleiben, so dass sich die betrachteten Anfangswerte wie folgt ergeben:

Parameter	Beschreibung	Wert
S_0	Anfangsbestand S	99
E_0	Anfangsbestand E	0.001
I_0	Anfangsbestand I	0.044
N_0	Anfangsbestand N	100

Die nachfolgenden Abbildungen stellen (unter Verwendung des Gewichtungsparameters $\beta = 2$) die Näherungslösungen der dualen Zustände des eingeführten Stabilisierungsproblems für $N = 10$ und $N = 11$ Kollokationspunkte dar:

⁵⁵vgl. Satz 26, S. 163

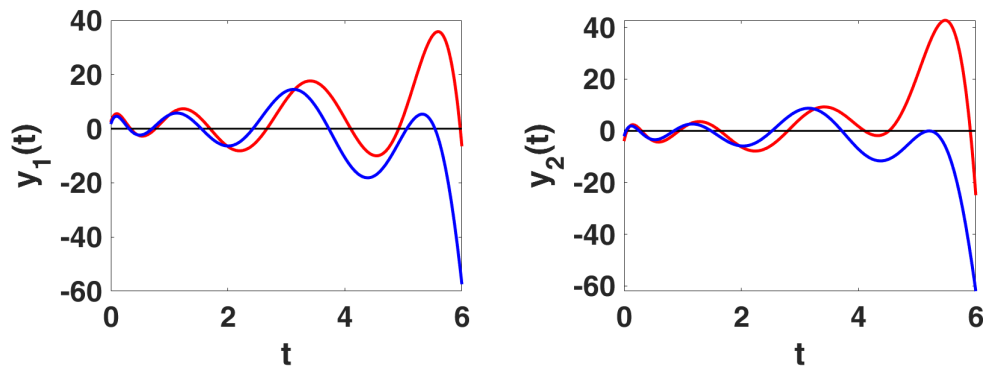


Abbildung 6.31: Näherungslösung der dualen Zustände $y_1(t)$ und $y_2(t)$ für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

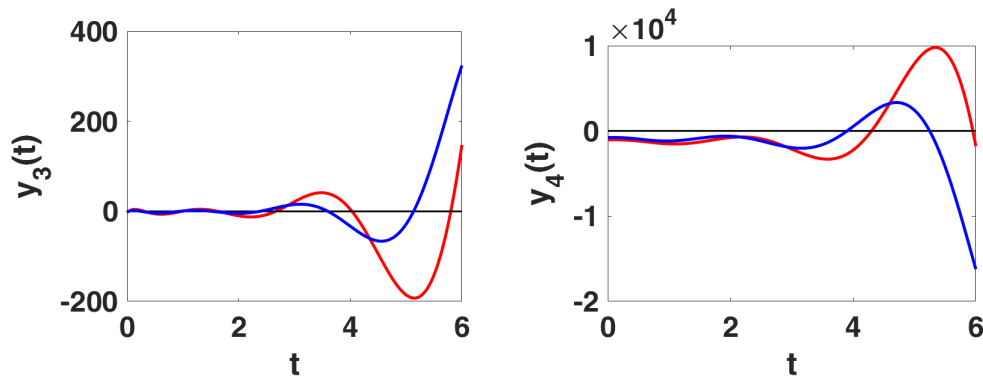


Abbildung 6.32: Näherungslösung der dualen Zustände $y_3(t)$ und $y_4(t)$ für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

Eine vergleichende Darstellung der Anfangswerte der dualen Zustandsfunktionen ist der Tabelle 6.13 zu entnehmen. Die größte Differenz zwischen der Anwendung des Verfahrens mit $N = 10$ und $N = 11$ Kollokationspunkten findet sich dabei in der Auswertung des Anfangswertes des dualen Zustandes y_4 mit einer Differenz von über 10^2 . Die scheinbar größte Genauigkeit zwischen zwei approximierten Lösungen lieferte auch für diese Setzung die Auswertung des Anfangswertes des dualen Zustandes y_1 mit einer Differenz von weniger als $8 \cdot 10^{-1}$.

N	$y_N^{(1)}(0)$	$y_N^{(2)}(0)$	$y_N^{(3)}(0)$	$y_N^{(4)}(0)$
10	2.445701565930	-3.976115623217	-2.809558869249	$-1.037707952605 \cdot 10^3$
11	1.690452706928	-2.088655033162	-1.373645426461	$-7.705007987755 \cdot 10^2$

Tabelle 6.13: Numerisch ermittelte Anfangswerte der dualen Zustände für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

Unterstellen wir die Gültigkeit der Bedingungen des Satzes [23](#), so erhalten wir wieder Rückschlüsse auf die Lösungen der primalen Aufgabe über die Setzungen^{[56](#)}

$$\tilde{x}(t) = \hat{W}^{-1}(t) \left[y'_N(t) + A^T y_N(t) \right] \quad \text{und} \quad u(t) = \text{Proj}_{\tilde{U}} \left(B^T y_N(t) e^{-\beta t} \right). \quad (6.36)$$

Unter Ausnutzung dieser Setzung erhalten wir die folgenden Abbildungen für die approximierten Lösungen der primalen Zustands- und Steuerungsfunktionen für eine wachsende Anzahl an Kollokationspunkten. Eine Zeiteinheit t entspricht aufgrund der Modellierung der Dauer von 3 Monaten.

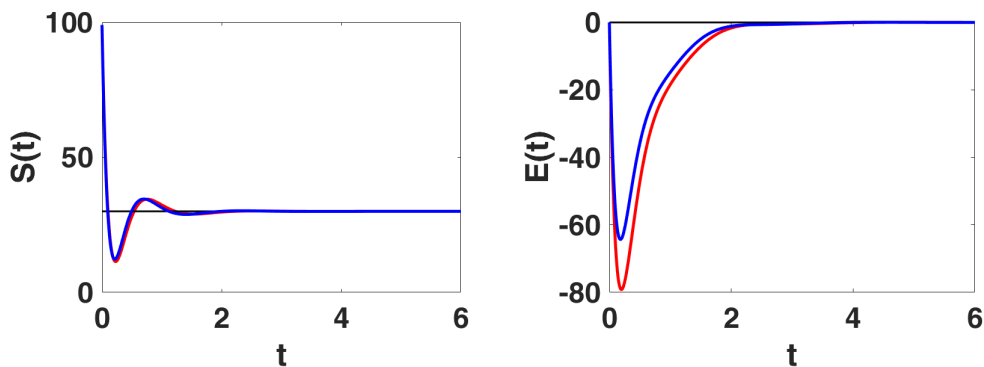


Abbildung 6.33: Näherungslösung der Komponenten S_N und E_N für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

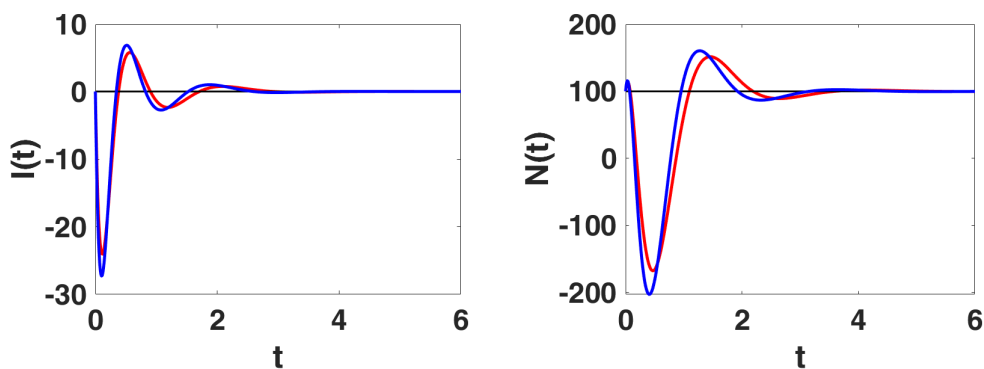


Abbildung 6.34: Näherungslösung der Komponenten I_N und N_N für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

⁵⁶vgl. Satz [23](#) S. [104](#)

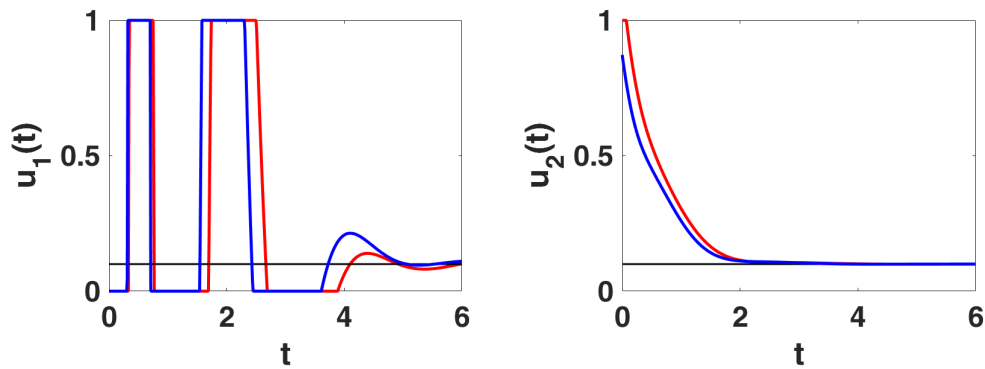


Abbildung 6.35: Näherungslösung der primalen Steuerungen $u_1(t)$ und $u_2(t)$ für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

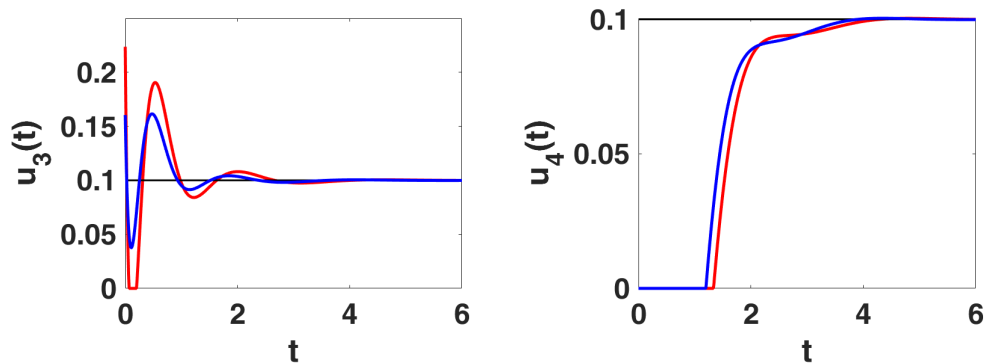


Abbildung 6.36: Näherungslösung der primalen Steuerungen $u_3(t)$ und $u_4(t)$ für den Fall $\max_{i=1,\dots,4} |x_{0i} - x_{zi}| = 69$ für $N = 10$ (—) und $N = 11$ (—)

Die Darstellungen der approximierten Lösungen der Steuerungsfunktionen u_1 und u_4 in den Abbildungen [6.35](#) und [6.36](#) bestätigen die Annahme, dass durch die Betrachtung eines größeren $\delta_i > |x_{0i} - x_{zi}|$ die Steuerungsbeschränkungen aktiv werden. Gleichzeitig zeigen sich in den Abbildungen [6.33](#) und [6.34](#) die Grenzen der verwendeten Modellierung: Die Interpretation der Abbildungen lässt innerhalb eines Teilintervalls auf eine negative Anzahl an Personen der zugehörigen Komponenten schließen. Dieser Umstand offenbart den Bedarf der Erweiterung des eingeführten SEIR Modells durch Zustandsbeschränkungen.

Aufgrund des fehlenden Nachweises der Gültigkeit der Voraussetzungen aus Satz [23](#) ist zudem unklar, ob eine Setzung der Form [\(6.36\)](#) zulässig ist. Wir wollen daher – ähnlich wie im zuvor betrachteten Fall der freien Steuerung des SEIR Modells die Funktionalwerte – der primalen und dualen Aufgabe für die ermittelten approximierten Lösungen vergleichen. Eine entsprechende Darstellung findet sich in der nachfolgenden Tabelle:

N	$G^N(y_N, y'_N)$	$J^N(\tilde{x}_N, u_N)$	$\frac{ G^N(y_N, q_N) - J^N(\tilde{x}_N, u_N) }{ G^N(y_N, q_N) }$
10	$5.609794163429842 \cdot 10^4$	$5.599878110408496 \cdot 10^4$	0.001767632239698
11	$4.170669465720381 \cdot 10^4$	$4.163851993518903 \cdot 10^4$	0.001634622992187

Tabelle 6.14: Vergleich der numerisch ermittelten Zielfunktionswerte von dualer und primaler Aufgabe für eine unterschiedliche Anzahl von Stützstellen für $\beta = 2$

Bemerkung 31. Die exemplarische Untersuchung alternativer Gewichtungsparemetrier lieferte die folgende Erkenntnis: Für kleinere β stieg die Anzahl der auswertbaren Kollokationspunkte, allerdings zeigte sich ein langsames Konvergenzverhalten. Für größere β war die Auswertung auch für eine geringere Anzahl an Kollokationspunkten zielführend. Jedoch sank die Anzahl der auswertbaren Kollokationspunkte, so dass auch in diesen Fällen keine stärkeren Konvergenzeigenschaften nachgewiesen werden konnten. Die Schaltstruktur der Komponenten u_i blieb auch für eine unterschiedliche Wahl der Gewichtungsparemetrier β erhalten. Diese Beobachtungen decken sich mit den in [Bur19]⁵⁷ dargelegten Erkenntnissen über den Einfluss des Gewichtungsparemetrier β .

Folgen wir den Ausführungen aus [Kac01], so wird deutlich, dass der obig dargestellte Verlauf der Zustandsfunktionen – und somit die Notwendigkeit der Einführung von Zustandsbeschränkungen – wenig überraschend ist. In [Kac01] definiert der Autor für das System⁵⁸

$$x'(t) = Ax(t) + Bu(t) \quad (6.37)$$

mit den Matrizen $A \in \mathbb{R}^{n \times n}$ und $B \in \mathbb{R}^{n \times m}$ den Begriff der *internen Positivität* durch:

Definition 23 (Interne Positivität ([Kac01])). Das System (6.37) heißt intern positiv, falls für jeden Anfangswert $x_0 \in \mathbb{R}_+^n$ und alle $u(t) \in \mathbb{R}_+^m$ der Zustandsvektor $x(t) \in \mathbb{R}_+^n$ für alle $t \geq t_0$ ist.

Zudem wird in [Kac01] die Gültigkeit des folgenden Satzes bewiesen:

Satz 27 ([Kac01]). Das System (6.37) ist genau dann intern positiv, wenn

- für die Matrix $A := [a_{ij}] \in \mathbb{R}^{n \times n}$ für alle Komponenten a_{ij} mit $i \neq j$ gilt: $a_{i,j} \geq 0$.
- $B \in \mathbb{R}_+^{n \times m}$.

Für die im SEIR Modell betrachtete Matrix A mit

$$A = \begin{pmatrix} -(d + cI_z + \alpha_1 u_1^z) & 0 & -cS_z & b \\ cI_z & -(e + d + \alpha_2 u_2^z + \alpha_4 u_4^z) & cS_z & 0 \\ 0 & e & -(g + a + d + \alpha_3 u_3^z) & 0 \\ 0 & \alpha_4 u_4^z & -a & b - d \end{pmatrix}$$

ist die Bedingung des Satzes 27 offensichtlich nur für die Setzung $a = c \cdot S_z = 0$ und somit für $a = 0$ und $c = 0$ oder $S_z = 0$ erfüllt, da $a_{13} = -a_{23}$.⁵⁹

Für die Matrix B mit

$$B = \begin{pmatrix} -\alpha_1 S_z & 0 & 0 & 0 \\ 0 & -\alpha_2 E_z & 0 & -\alpha_4 E_z \\ 0 & 0 & -\alpha_3 I_z & 0 \\ 0 & 0 & 0 & \alpha_4 E_z \end{pmatrix}$$

⁵⁷ vgl. hierfür [Bur19], S. 199

⁵⁸ Die Definition des Systems (6.37) sowie die Ausführungen von Definition 23 und Satz 27 entsprechen einer Spezialisierung der in [Kac01] formulierten Ergebnisse auf eine für unseren Fall des SEIR Modells angepasste Formulierung mit den zeitunabhängigen Matrixfunktionen A und B und einer nicht existenten Outputfunktion mittels der Setzung $C \equiv D \equiv 0$. Für eine allgemeinere Formulierung verweisen wir auf die ausgewiesene Quelle.

⁵⁹ Den Fall $c = 0$ können wir jedoch ausschließen, da dies den Ansteckungskoeffizienten beschreibt. Ist dieser Parameter also 0, so ist die Epidemie nicht existent und bedarf demnach auch keiner Behandlung.

könnten wir zwar die Setzungen $\hat{\alpha}_i := -\alpha_i \in \mathbb{R}_+$ und $\hat{u}_i(\cdot) := -\tilde{u}_i(\cdot)$ vorschlagen, jedoch erzwingt auch dann die Komponente $b_{44} = -\hat{\alpha}_4 E_z$ die Setzung $\hat{\alpha}_4 \cdot E_z = 0$. In diesem Fall ist die Kalman-Bedingung aus Satz [26](#) jedoch nicht mehr erfüllt^{[60](#)}.

Unabhängig von diesen Ausführungen lässt sich weiter einsehen, dass selbst unter der Gültigkeit der Bedingungen aus Satz [27](#) die Einführung von Zustandsbeschränkungen für das hier betrachtete SEIR Modell sinnvoll erscheint. So können wir ohne derartige Beschränkungen auch keine oberen Schranken an die Anzahl der Infizierten (und damit beispielsweise auch nicht die Einhaltung der Kapazitätsgrenzen der Intensivstationen) garantieren.

⁶⁰vgl. Satz [26](#), S. [163](#)

Abschließende Bemerkungen

Gegenstand der vorliegenden Arbeit war die Untersuchung linear-quadratischer Optimalsteuerungsprobleme mit einem a priori gegebenen unendlichen Zeithorizont $[0, \infty)$ unter der Berücksichtigung von Steuerungsbeschränkungen und Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta \neq 0$. Die behandelte Aufgabenklasse ist für die Approximation nichtlinearer Aufgaben von großer Bedeutung und tritt bei Verfahren der sequentiellen quadratischen Optimierung auf. Zudem gelangen wir im Kontext der asymptotischen Steuerbarkeit dynamischer Systeme zu einer Beschreibung linear-quadratischer Optimalsteuerungsprobleme, welche durch die Verwendung echter Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta > 0$ geprägt sind. Für die Behandlung derartiger Steuerungsprobleme wurde ein Zugang mittels speziell gewichteter Funktionenräume gewählt und der Lebesguesche Integralbegriff zugrunde gelegt. Dafür wurden Zustände und Steuerungen betrachtet, welche Elemente eines gewichteten Sobolev- bzw. eines gewichteten Lebesgue-Raumes sind. Diese Raumwahl ermöglichte die Einbettung der linear-quadratischen Optimalsteuerungsaufgaben über $[0, \infty)$ in eine Hilbertraumtheorie und eine Behandlung der Problemklasse mit den Methoden der konvexen Analysis. Auf diese Weise gelang der Nachweis der Existenz und Eindeutigkeit einer optimalen Lösung. Zur Bestimmung dieser konnten wir für Probleme, welche durch die Verwendung von Dichtefunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta < 0$ geprägt sind, auf bestehende Optimalitätsbedingungen verweisen. Ausstehende Optimalitätsbedingungen für Probleme mit echten Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta > 0$ motivierten schließlich die Formulierung einer dualen Aufgabe.

Für die Formulierung der dualen Aufgabe wählten wir einen Zugang im Sinne von Lagrange. Die geschickte Wahl eingehender Mengen im Rahmen der Konstruktion führte uns schließlich zu einer Formulierung der dualen Aufgabe in Form eines Variationsproblems in gewichteten Funktionenräumen mit Dichtefunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta < 0$. Die derartige Formulierung ermöglichte zunächst die Überführung der Variationsaufgabe in die Form eines Optimalsteuerungsproblems in gewichteten Funktionenräumen mit echter Dichtefunktion und ohne Steuerungsbeschränkungen. Für diese Aufgabenklasse gelang der Nachweis der Gültigkeit eines Maximumprinzips vom Pontryagin-Typ. Dieses liefert die kanonischen Gleichungen, die Maximumbedingung aber auch die Existenz einer adjungierten Funktion, welche einer natürlichen Transversalitätsbedingung genügt. Aufbauend auf den durch das Maximumprinzip gegebenen notwendigen Optimalitätsbedingungen konnte schließlich die starke Dualitätsbeziehung nachgewiesen werden. Als Folge dessen erhielten wir durch die notwendigen Optimalitätsbedingungen der dualen Aufgaben nun auch hinreichende – und auf Grund der linear-quadratischen Struktur zudem auch notwendig geltende – Optimalitätsbedingungen für die zugrundeliegende primale Aufgabenklasse. Damit gelingt nun die Angabe bisher noch fehlender notwendiger und hinreichender Optimalitätsbedingungen für linear-quadratische Optimalsteuerungsprobleme mit einem a priori

gegebenen unendlichen Zeithorizont $[0, \infty)$ unter der Berücksichtigung von Steuerungsbeschränkungen und echten Gewichtsfunktionen der Form $\nu(t) = e^{\beta t}$, mit $\beta > 0$. Zudem erlaubte dieser Zugang schwächere Voraussetzungen an die primale Aufgabe, womit die erhaltenen Optimalitätsbedingungen auch für Probleme mit Dichtefunktionen eine Erweiterung darstellen.

Für die numerische Behandlung wurde ein direktes Spektralverfahren zur Lösung linear-quadratischer Optimalsteuerungsaufgaben über $[0, \infty)$ vorgestellt. Die genutzten Approximationsansätze basieren auf Verwendung verallgemeinerter Laguerre-Polynome. Diese bilden ein vollständiges Orthogonalsystem auf dem durch die Dichtefunktion $\nu(t) = e^{\beta t}$, mit $\beta < 0$ gewichteten Lebesgue-Raum. Die Formulierung der dualen Aufgabe in Form eines Variationsproblems über gewichteten Funktionenräumen mit Dichtefunktionen ermöglichte nun eine dualbasierte numerische Behandlung linear-quadratischer Optimalsteuerungsprobleme mit Gewichtsfunktionen und somit insbesondere auch die Behandlung von Aufgaben der asymptotischen Steuerbarkeit. Die Anwendung des direkten Spektralverfahrens wurde anhand verschiedener Beispiele im Detail erläutert. So betrachteten wir neben einem Beispiel mit einer Dichtefunktion von Garg, Hager und Rao [GHR11] auch zwei Anwendungen aus dem Bereich der asymptotischen Steuerbarkeit. Neben der Behandlung des Lotka-Volterra Modells untersuchten wir dafür ein Beispiel aus der Epidemiologie in Form des SEIR Modells. Für die Beschreibung der in das Modell eingehenden Parameter, orientierten wir uns an Kennzahlen der Viruserkrankung Covid-19 und berechneten Behandlungs- und Vorsorgestrategien mit den Mitteln der Steuerungstheorie. Neben der ausführlichen Herleitung der Kollokationsgleichungen beinhaltet diese Arbeit im Anhang auch den Quellcode für eine Implementierung in Matlab. Im Rahmen der numerischen Experimente wurden die Approximationsergebnisse in Abhängigkeit des gewählten Polynomgrades N bzw. der Anzahl der verwendeten Kollokationspunkte ausgewertet und unter verschiedenen Gesichtspunkten betrachtet.

Die in dieser Arbeit präsentierten numerischen Experimente führten zu zufriedenstellenden Approximationsergebnissen und motivieren eine weitere Untersuchung behandelbarer Klassen. Dabei zeigte sich insbesondere die Notwendigkeit der Beachtung von Zustandsbeschränkungen. Tatsächlich ist zu erwarten, dass die hier vorgeschlagene Philosophie für die Konstruktion der dualen Aufgabe auch für linear-quadratische Probleme mit Zustandsbeschränkungen der Form $|x(t)| \leq \alpha$ anwendbar ist. Hierzu schlagen wir eine Zerlegung des zulässigen Bereiches $\mathcal{B} = X_0 \cap X_1$ im Sinne der in Kapitel 3 geführten Konstruktion vor und schränken die Menge X_0 nun durch die Hinzunahme der Zustandsbeschränkungen ein. Die derartige Konstruktion führt schließlich auf zwei untergeordnete Hilfsprobleme über schwach kompakten Mengen, welche nachfolgend hinsichtlich der Existenz und Eindeutigkeit optimaler Lösungen und Optimalitätsbedingungen zu untersuchen sind.

Betrachten wir allgemeinere Optimalsteuerungsprobleme mit bezüglich ξ und v nicht separierbaren Integranden und Zustandsgleichungen sowie allgemeineren Zustandsbeschränkungen der Form $|x(t)| \leq \alpha(t)$, so ist ein ähnliches Schema für die Konstruktion der dualen Aufgabe denkbar. Die Zerlegung des zulässigen Bereiches \mathcal{B} wäre in diesem Falle beispielsweise durch die Bildung der Mengen X_0 und X_1 denkbar, wobei die Menge X_0 alle Paare (x, u) beinhaltet, die den Nebenbedingungen der primalen Aufgabe genügen, außer den Zustandsbeschränkungen. Die Menge X_1 ließe sich durch alle Paare (x, u) , die Elemente der – der Aufgabe zugrundeliegenden – gewichteten Funktionenräume sind und der Zustandsgleichung genügen, konstruieren. Als Funktional Φ schlagen wir auch hier

die Verwendung der erweiterten Lagrangefunktion vor. Eine derartige Setzung führt uns schließlich zur Formulierung eines untergeordneten Hilfsproblems in Form eines Integralfunktionals, welches die Pontryaginsche Funktion enthält. Im Gegensatz zu der Behandlung linear-quadratischer Modelle ist in diesem Fall jedoch davon auszugehen, dass eine derartige Konstruktion zu einem untergeordneten Hilfsproblem führt, welches bezüglich ξ und v nicht separierbar ist. Daher schlagen wir statt dessen eine Abschätzung der Pontryaginschen Funktion durch die Hamiltonfunktion im Sinne der Behandlung von [Klö81] und [Lyk10] vor. Diese Herangehensweise erlaubt schließlich die Eliminierung der Steuerungsvariablen und führt somit auf die Formulierung eines untergeordneten Hilfsproblems über einer schwach kompakten Menge. Dieses ist nun wiederum selbst auf die Existenz und Eindeutigkeit optimaler Lösungen sowie zugehöriger Optimalitätsbedingungen zu untersuchen.

Zudem offenbarte die im Zuge der numerischen Untersuchung zu diskutierende Existenz der optimalen Lösung der dualen Aufgabe einige Herausforderungen. Aufgrund der Struktur der dualen Aufgabe, als Steuerungsproblem in gewichteten Funktionenräumen mit Dichtefunktion und ohne Steuerungsbeschränkungen, könnte eine weitere Untersuchung, im Sinne der Existenzbetrachtungen auf primaler Seite, zu einer Formulierung leichter verifizierbarer Voraussetzungen an die Existenz einer Lösung der dualen Aufgabe führen und somit eine derartige Diskussion erleichtern.

Anhang A

Wichtige Sätze und Definitionen

Definition 24 ([AF88], S. 3 ff.). Sei $I \subset \mathbb{R}$ ein Intervall, $D \subset \mathbb{R}^n$. Die Funktion $f : I \times D \rightarrow \mathbb{R}^n$ erfüllt die Carathéodory-Bedingungen auf $I \times D$, falls sie die folgenden Eigenschaften besitzt:

1. Die Funktion $f = f(t, \xi)$ ist für fast alle $t \in I$ stetig in ξ .
2. Die Funktion $f = f(t, \xi)$ ist für jedes $\xi \in D$ messbar in t .
3. Es existiert eine nichtnegative, messbare Funktion $m : I \rightarrow \mathbb{R}$, so dass $|f(t, \xi)| \leq m(t)$ für alle $(t, \xi) \in I \times D$ erfüllt ist. In dem Fall dass t in I unbeschränkt ist, soll diese Bedingung auf jedem endlichen Teilintervall erfüllt sein.

Definition 25 (Carathéodory-Lösung). Die Differentialgleichung $x' = f(t, x)$ heißt Carathéodory-Gleichung, falls f die Carathéodory-Bedingungen erfüllt. Eine Funktion $x : I \rightarrow \mathbb{R}^n$ heißt Lösung der Carathéodory-Gleichung, falls sie auf jedem Teilintervall $[a, b] \subset I$ absolut stetig ist und die Gleichung

$$x'(t) = f(t, x(t)) \text{ für fast alle } t \in I$$

erfüllt. Gilt zusätzlich $x(t_0) = x_0$, so heißt x Carathéodory-Lösung des Anfangswertproblems

$$x' = f(t, x), \quad x(t_0) = x_0. \tag{A.1}$$

Satz 28 ([Wa100], S. 128, XVIII. Existenzsatz). Es sei $I = [t_0, t_0 + a]$, $a > 0$. Erfüllt die Funktion $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ die Carathéodory-Bedingungen, so besitzt die Aufgabe (A.1) auf dem Intervall I mindestens eine Lösung im Sinne von Carathéodory.

Lemma 46 ([PB19], Lemma 2, S. 6). Sei $x \in W_2^1(\mathbb{R}_{++}, e^{\beta t})$, mit $\beta \neq 0$, dann gilt

1. Für alle x welche auf dem eingeschränkten Intervall $[0, t]$ definiert sind, gilt die Inklusion $x \in W_1^1([0, t])$, $\forall t \in \mathbb{R}_+$.
2. Es gilt die folgende Ungleichung vom Poincaré Typ:

$$\|x - x(0)\|_{2, e^{\beta t}} \leq c \|Dx\|_{2, e^{\beta t}}, \quad c > 0.$$

Satz 29 ([Heu04] S. 102). $[a, b]$ sei ein kompaktes und I ein völlig beliebiges Intervall. Für jedes feste $t \in I$ sei die Funktion $\psi(t, \lambda)$ partiell nach λ differenzierbar und es gebe ein $g \in L_1(I)$ mit

$$\left| \frac{\partial}{\partial \lambda} \psi(t, \lambda) \right| \leq g(t) \quad \forall \lambda \in [a, b], t \in I. \quad (\text{A.2})$$

Dann ist die Funktion $\frac{\partial}{\partial \lambda} \psi(t, \lambda)$ für jedes feste $\lambda \in [a, b]$ Lebesgue-integrierbar auf I , die Funktion $\Psi(\lambda) = \int_I \psi(t, \lambda) dt$ ist differenzierbar auf $[a, b]$ und ihre Ableitung kann „durch Differentiation unter dem Integral“ gewonnen werden, kurz: es gilt

$$\frac{d}{d\lambda} \int_I \psi(t, \lambda) dt = \int_I \frac{\partial}{\partial \lambda} \psi(t, \lambda) dt.$$

Relevante Definitionen im Rahmen der Steuerbarkeitsanalyse

Definition 26 (rechte Seite (vgl. [Son98], Def. 2.6.1, S.43)). Sei \mathcal{X} eine offene Teilmenge des \mathbb{R}^n und sei \mathcal{U} ein metrischer Raum. Als rechte Seite unter Berücksichtigung von \mathcal{X} und \mathcal{U} bezeichnen wir eine Funktion

$$f : \mathbb{R} \times \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}^n$$

welche in der folgenden Weise beschrieben werden kann: Es existiert ein anderer metrischer Raum \mathcal{S} und Abbildungen

$$\tilde{f} : \mathcal{S} \times \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}^n \quad \text{und} \quad \pi : \mathbb{R} \rightarrow \mathcal{S}$$

so dass

$$f(t, x, u) := \tilde{f}(\pi(t), x, u)$$

und die folgenden Bedingungen erfüllt sind:

1. $\tilde{f}(s, \cdot, u)$ ist Element der Klasse C^1 für alle festen s, u ;
2. \tilde{f} und ihre partielle Ableitung \tilde{f}_x sind stetig auf $\mathcal{S} \times \mathcal{X} \times \mathcal{U}$;
3. π ist eine messbare und lokal wesentlich beschränkte Funktion.

Definition 27 (zeitstetiges System (vgl. [Son98], Def. 2.6.4, S. 45)). Ein zeitstetiges System (über \mathbb{R}) ist ein System vom Typ Σ_f , wobei f eine rechte Seite ist.

Definition 28 (zeitstetiges System der Klasse C^k (vgl. [Son98], Def. 2.6.6, S. 45)). Sei $k = 1, 2, \dots$. Ein zeitstetiges System der Klasse C^k ist ein System Σ in welchem \mathcal{U} eine offene Teilmenge des \mathbb{R}^m ($m \in \mathbb{N}$, $m > 0$) und für die \mathcal{S} und \tilde{f} aus Definition 26 derart gewählt werden können, so dass \mathcal{S} eine offene Teilmenge des euklidischen Raumes und $\tilde{f} \in C^k$ sei.

Definition 29 (lineares System (vgl. [Son98], Def. 2.7.2, S. 47)). Das zeitstetige System Σ heißt linear (über \mathbb{R}) wenn es von der Klasse C^1 ist und:

- $\mathcal{X} = \mathbb{R}^n$ und $\mathcal{U} = \mathbb{R}^m$; und
- $f(t, \cdot, \cdot)$ linear für alle $t \in \mathbb{R}$ ist.

Anhang B

Existenz und Eindeutigkeit der Hilfsprobleme

B.1 Existenz einer Lösung von $(\overline{HP2})$

$(\overline{HP2})$:

$$J_{2,y}^{\nu_0}(x) = \int_0^\infty \left\{ \frac{1}{2} x^T(t) W(t) x(t) + w^T(t) x(t) - x^T(t) [y'(t) + (A^T(t) + \beta_0 \cdot I_{n \times n}) y(t)] \right\} e^{\beta_0 t} dt$$
$$-x_0^T y(0) \longrightarrow \inf! \quad \text{u.d.N. } x \in L_2^n(\mathbb{R}_+, e^{\beta_0 t}) \quad (3.19)$$

Wir behandeln die Aufgabe $(\overline{HP2})$ unter Gültigkeit der Voraussetzungen (V1) und (V2) und wollen diese wieder in den Kontext der Hilbertraummethode einbetten. Dazu definieren wir auf dem Hilbertraum $\mathcal{H} = L_2^n(\mathbb{R}_+, \nu_0)$ die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ mit

$$\mathcal{Q}(x_1, x_2) := \int_0^\infty x_1^T(t) W(t) x_2(t) \nu_0(t) dt, \quad (3.20)$$

sowie $q : \mathcal{H} \rightarrow \mathbb{R}$ mit

$$q(x) := - \int_0^\infty x^T(t) [w(t) - y'(t) - (A^T(t) + \beta_0 \cdot I_{n \times n}) y(t)] \nu_0(t) dt. \quad (3.21)$$

Zudem sei für beliebige aber feste Werte $y(0)$ die Konstante

$$q_0 := x_0^T y(0)$$

gegeben. Mittels dieser Setzungen können wir das Zielfunktional von $(\overline{HP2})$ gemäß

$$J_{2,y}^{\nu_0}(x) = \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0$$

beschreiben. Wir halten die Eigenschaften der eingeführten Abbildungen in den folgenden zwei Lemmata fest.

Lemma 47. *Durch die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.20), wird eine symmetrische, koerzitive Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt und damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Es lässt sich leicht nachrechnen, dass \mathcal{Q} für alle $x_1, x_2, \hat{x}_1, \hat{x}_2 \in \mathcal{H}$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned}\mathcal{Q}(x_1 + \hat{x}_1, x_2) &= \mathcal{Q}(x_1, x_2) + \mathcal{Q}(\hat{x}_1, x_2) \\ \mathcal{Q}(x_1, x_2 + \hat{x}_2) &= \mathcal{Q}(x_1, x_2) + \mathcal{Q}(x_1, \hat{x}_2) \\ \mathcal{Q}(\lambda x_1, x_2) &= \lambda \mathcal{Q}(x_1, x_2).\end{aligned}$$

Aufgrund der Symmetrie der Matrix W gilt zudem für alle $x_1, x_2 \in \mathcal{H}$

$$\mathcal{Q}(x_1, x_2) = \mathcal{Q}(x_2, x_1).$$

Damit ist \mathcal{Q} eine symmetrische Bilinearform. Wir wollen nun zeigen, dass die so definierte Abbildung auch koerzitiv ist. Dafür erinnern wir uns, dass aufgrund der positiven Definitheit der Matrix W gilt¹ $\underline{W} = \inf_{t \in \mathbb{R}} \lambda_{\min}(W)(t) > 0$. Mit diesem Wissen betrachten wir nun die folgende Abschätzung

$$\mathcal{Q}(x, x) = \int_0^{\infty} x^T(t)W(t)x(t)\nu_0(t)dt \geq \underline{W} \int_0^{\infty} x^2(t)\nu_0(t)dt \geq c \|x\|_{\mathcal{H}}^2,$$

mit der positiven Konstante $c = \underline{W} > 0$.

Zudem sichert uns die Raumzugehörigkeit der Matrix W mit $W \in L_{\infty}^{n \times n}(\mathbb{R}_+) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ die Endlichkeit des Wertes

$$\|W\|_{\infty} = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |W(t)|$$

und sichert somit die Abschätzung

$$\begin{aligned}|\mathcal{Q}(x_1, x_2)| &= \left| \int_0^{\infty} x_1^T(t)W(t)x_2(t)\nu_0(t)dt \right| \leq \|W\|_{\infty} \int_0^{\infty} |x_1^T(t)x_2(t)|\nu_0(t)dt \\ &\leq \|W\|_{\infty} \left(\int_0^{\infty} |x_1(t)|^2 \nu_0(t)dt \right)^{1/2} \left(\int_0^{\infty} |x_2(t)|^2 \nu_0(t)dt \right)^{1/2} \\ &\leq \|W\|_{\infty} \|x_1\|_{\mathcal{H}} \|x_2\|_{\mathcal{H}},\end{aligned}$$

womit die Beschränktheit von \mathcal{Q} auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist \mathcal{Q} ein auf \mathcal{H} lineares und (norm-)beschränktes Funktional und somit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . \square

Lemma 48. *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (3.21), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Die Linearität der Abbildung q lässt sich leicht nachrechnen. Wir zeigen daher nun

¹Dabei bezeichnet $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M .

die (Norm-)Beschränktheit der Linearform q und treffen dazu die folgende Abschätzung

$$\begin{aligned} |q(x)| &= \left| \int_0^\infty x^T(t)[w(t) - y'(t) - (A^T(t) + \beta_0 \cdot I_{n \times n})y(t)]\nu_0(t)dt \right| \\ &\leq \left| \int_0^\infty x^T(t)w(t)\nu_0(t)dt \right| + \left| \int_0^\infty x^T(t)y'(t)\nu_0(t)dt \right| \\ &\quad + \left| \int_0^\infty x^T(t)(A^T(t) + \beta_0 \cdot I_{n \times n})y(t)\nu_0(t)dt \right|. \end{aligned}$$

Durch die Raumzugehörigkeit² der Funktionen w und A erhalten wir weiter

$$\begin{aligned} |q(x)| &\leq \left(\int_0^\infty |x(t)|^2 \nu_0(t)dt \right)^{1/2} \left(\int_0^\infty |w(t)|^2 \nu_0(t)dt \right)^{1/2} \\ &\quad + \left(\int_0^\infty |x(t)|^2 \nu_0(t)dt \right)^{1/2} \left(\int_0^\infty |y'(t)|^2 \nu_0(t)dt \right)^{1/2} \\ &\quad + \|A + \beta_0 \cdot I_{n \times n}\|_\infty \left| \int_0^\infty x^T(t)y(t)\nu_0(t)dt \right| \\ &\leq \|x\|_{\mathcal{H}} \|w\|_{\mathcal{H}} + \|x\|_{\mathcal{H}} \|y'\|_{\mathcal{H}} \\ &\quad + \|A + \beta_0 \cdot I_{n \times n}\|_\infty \left(\int_0^\infty |x(t)|^2 \nu_0(t)dt \right)^{1/2} \left(\int_0^\infty |y(t)|^2 \nu_0(t)dt \right)^{1/2} \\ &\leq \|x\|_{\mathcal{H}} (\|w\|_{\mathcal{H}} + \|y'\|_{\mathcal{H}} + \|A + \beta_0 \cdot I_{n \times n}\|_\infty \|y\|_{\mathcal{H}}) \leq c \|x\|_{\mathcal{H}}, \end{aligned}$$

womit die Beschränktheit von q auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist q ein auf \mathcal{H} lineares und (norm-)beschränktes Funktional und somit auch stetig bzgl. der Normtopologie des Raumes \mathcal{H} . \square

Die so nachgewiesenen Eigenschaften lassen nun direkte Schlüsse auf das Zielfunktional zu. Diese wollen wir im folgenden Lemma festhalten.

Lemma 49. *Das Zielfunktional des Problems $(\overline{HP2})$ ist strikt konvex und schwach unterhalbstetig auf \mathcal{H} .*

Beweis. Wir zeigen zunächst die strikte Konvexität des Zielfunctionals und betrachten dafür zunächst den Operator \mathcal{Q} sowie beliebige Elemente $x_1, x_2 \in \mathcal{H}$ mit $x_1 \neq x_2$. Aufgrund der Koerzitivität der symmetrischen Bilinearform \mathcal{Q} existiert eine Konstante $c > 0$, so dass

$$0 < c \|x_1 - x_2\|^2 \leq \mathcal{Q}(x_1 - x_2, x_1 - x_2) = \mathcal{Q}(x_1, x_1) - 2\mathcal{Q}(x_1, x_2) + \mathcal{Q}(x_2, x_2)$$

gilt und damit

$$2\mathcal{Q}(x_1, x_2) < \mathcal{Q}(x_1, x_1) + \mathcal{Q}(x_2, x_2).$$

Die Multiplikation der letzten Gleichung mit $\lambda(1 - \lambda)$, wobei $\lambda \in (0, 1)$ sei, liefert dann

$$2\lambda(1 - \lambda)\mathcal{Q}(x_1, x_2) < \lambda(1 - \lambda)\mathcal{Q}(x_1, x_1) + \lambda(1 - \lambda)\mathcal{Q}(x_2, x_2).$$

²vgl. Voraussetzung (V1), S. 25

Damit können wir für alle $x_1, x_2 \in \mathcal{H}$ die folgende Abschätzung finden:

$$\begin{aligned} \mathcal{Q}((1-\lambda)x_1 + \lambda x_2, (1-\lambda)x_1 + \lambda x_2) &= (1-\lambda)^2 \mathcal{Q}(x_1, x_1) + 2\lambda(1-\lambda) \mathcal{Q}(x_1, x_2) + \lambda^2 \mathcal{Q}(x_2, x_2) \\ &< (1-\lambda)^2 \mathcal{Q}(x_1, x_1) + \lambda(1-\lambda) \mathcal{Q}(x_1, x_1) + \lambda(1-\lambda) \mathcal{Q}(x_2, x_2) + \lambda^2 \mathcal{Q}(x_2, x_2) \\ &= (1-\lambda) \mathcal{Q}(x_1, x_1) + \lambda \mathcal{Q}(x_2, x_2). \end{aligned}$$

Da q linear ist erhalten wir durch $J_{2,y}^{\nu_0}(x) = \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0$ die strikte Konvexität des Zielfunktional. Als Summe (bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$) stetiger Funktionale ist auch das Zielfunktional stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . Die schwache Unterhalbstetigkeit folgt dann gemäß Lemma [8](#) aus der strikten Konvexität³. □

B.2 Existenz einer Lösung von (HP3)

(HP3) :

$$J_{1,y}^{\nu_1}(u) = \int_0^{\infty} \left\{ \frac{1}{2} (u^T(t) R(t) u(t) + r^T(t) u(t)) e^{\beta_1 t} - [B(t) u(t)]^T y(t) \right\} dt \longrightarrow \inf!$$

$$\text{u.d.N. } u \in \mathbf{U} := \left\{ u \in L_2^n(\mathbb{R}_+, e^{\beta_1 t}) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_+ \right\}.$$

Satz 30. *Unter Gültigkeit der Voraussetzungen (V1), (V2) und (V3) besitzt das Problem (HP3) eine eindeutige optimale Lösung.*

Beweis. Wir behandeln die Aufgabe (HP3) und wollen sie wieder in den Kontext der Hilbertraummethoden einbetten. Wir definieren für $\mathcal{H} := L_2^m(\mathbb{R}_+, \nu_1)$ die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ mit

$$\mathcal{Q}(u_1, u_2) := \int_0^{\infty} u_1^T(t) R(t) u_2(t) \nu_1(t) dt, \quad (\text{B.1})$$

sowie den Operator $q : \mathcal{H} \rightarrow \mathbb{R}$

$$q(u) := \int_0^{\infty} u^T(t) [r(t) e^{\beta_1 t} - B^T(t) y(t)] dt. \quad (\text{B.2})$$

Mittels dieser Setzung können wir durch

$$J_{1,y}^{\nu_1}(u) = \frac{1}{2} \mathcal{Q}(u, u) + q(u) \rightarrow \inf!$$

$$\text{u.d.N. } u \in \mathbf{U} = \left\{ u \in L_2^n(\mathbb{R}_+, \nu_1) \mid u(t) \in \mathcal{U} \ (\forall) t \in \mathbb{R}_+ \right\},$$

eine äquivalente Darstellung der Aufgabe (HP3) als konvexes Optimierungsproblem im Kontext der Hilbertraummethoden angeben.

Die nachfolgenden Lemmata geben Auskunft über die Eigenschaften der eingeführten Operatoren.

³vgl. Lemma [8](#), S. [30](#)

Lemma 50. *Durch die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in (B.1), wird eine symmetrische, koerzitive Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt und damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Es lässt sich leicht nachrechnen, dass \mathcal{Q} für alle $u_1, u_2, \hat{u}_1, \hat{u}_2 \in \mathcal{H}$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned} \mathcal{Q}(u_1 + \hat{u}_1, u_2) &= \mathcal{Q}(u_1, u_2) + \mathcal{Q}(\hat{u}_1, u_2) \\ \mathcal{Q}(u_1, u_2 + \hat{u}_2) &= \mathcal{Q}(u_1, u_2) + \mathcal{Q}(u_1, \hat{u}_2) \\ \mathcal{Q}(\lambda u_1, u_2) &= \lambda \mathcal{Q}(u_1, u_2). \end{aligned}$$

Aufgrund der Symmetrie der Matrix R gilt zudem für alle $u_1, u_2 \in \mathcal{H}$

$$\mathcal{Q}(u_1, u_2) = \mathcal{Q}(u_2, u_1).$$

Damit ist \mathcal{Q} eine symmetrische Bilinearform.

Wir wollen nun zeigen, dass die so definierte Abbildung auch koerzitiv ist. Dafür erinnern wir uns, dass aufgrund der positiven Definitheit der Matrix R gilt $\underline{R} = \inf_{t \in \mathbb{R}} \lambda_{\min}(R)(t) > 0$. Mit diesem Wissen betrachten wir nun die folgende Abschätzung

$$\mathcal{Q}(u, u) = \int_0^\infty u^T(t)R(t)u(t)\nu_1(t)dt \geq \underline{R} \int_0^\infty u^2(t)\nu_1(t)dt \geq c \|u\|_{\mathcal{H}}^2,$$

mit der positiven Konstante $c = \underline{R} > 0$.

Zudem sichert uns die Raumzugehörigkeit der Matrix $R \in L_\infty^{m \times m}(\mathbb{R}_+) \cap C(\mathbb{R}_+, \mathbb{R}^{m \times m})$ die Endlichkeit des Wertes

$$\|R\|_\infty = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |R(t)|$$

und liefert somit die Abschätzung

$$\begin{aligned} |\mathcal{Q}(u_1, u_2)| &= \left| \int_0^\infty u_1^T(t)R(t)u_2(t)\nu_1(t)dt \right| \leq \|R\|_\infty \int_0^\infty |u_1^T(t)u_2(t)|\nu_1(t)dt \\ &\leq \|R\|_\infty \left(\int_0^\infty |u_1(t)|^2 \nu_1(t)dt \right)^{1/2} \left(\int_0^\infty |u_2(t)|^2 \nu_1(t)dt \right)^{1/2} \\ &\leq \|R\|_\infty \|u_1\|_{\mathcal{H}} \|u_2\|_{\mathcal{H}}, \end{aligned}$$

womit die (Norm-)Beschränktheit von \mathcal{Q} auf \mathcal{H} bewiesen ist. Damit ist \mathcal{Q} ein auf \mathcal{H} lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional und somit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . \square

Lemma 51. *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (B.2), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

⁴Dabei bezeichnet $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M .

Beweis. Die Linearität der Abbildung q lässt sich leicht nachrechnen. Wir zeigen daher nun die Beschränktheit der Linearform q und treffen dazu die folgende Abschätzung

$$\begin{aligned} |q(u)| &= \left| \int_0^\infty u^T(t)[r(t)\nu_1(t) - B^T(t)y(t)]dt \right| \\ &\leq \left| \int_0^\infty u^T(t)r(t)\nu_1(t)dt \right| + \left| \int_0^\infty u^T(t)B^T(t)y(t)dt \right|. \end{aligned}$$

Durch die Raumzugehörigkeit⁵ der Funktionen r und B erhalten wir weiter

$$\begin{aligned} |q(u)| &\leq \left(\int_0^\infty |u(t)|^2 \nu_1(t)dt \right)^{1/2} \left(\int_0^\infty |r(t)|^2 \nu_1(t)dt \right)^{1/2} \\ &\quad + \|B\|_\infty \left(\int_0^\infty |u(t)|^2 \nu_1(t)dt \right)^{1/2} \left(\int_0^\infty |y'(t)|^2 \nu_1^{-1}(t)dt \right)^{1/2} \\ &\leq \|u\|_{\mathcal{H}} \|r\|_{\mathcal{H}} + \|B\|_\infty \|u\|_{\mathcal{H}} \|y'\|_{L_2^n(\mathbb{R}_+, \nu_1^{-1})} \leq c \|u\|_{\mathcal{H}}, \end{aligned}$$

womit die Beschränktheit von q auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist q ein auf \mathcal{H} lineares und (norm-)beschränktes Funktional und somit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . \square

Diese so nachgewiesenen Eigenschaften lassen nun direkte Schlüsse auf das Zielfunktional zu. Diese wollen wir im folgenden Lemma festhalten.

Lemma 52. *Das Zielfunktional des Problems (HP3) ist strikt konvex und schwach unterhalbstetig auf \mathcal{H} .*

Beweis. Wir zeigen zunächst die strikte Konvexität des Zielfunktional und betrachten dafür zunächst den Operator \mathcal{Q} sowie beliebige Elemente $u_1, u_2 \in \mathcal{H}$ mit $u_1 \neq u_2$. Aufgrund der Koerzitivität der symmetrischen Bilinearform \mathcal{Q} existiert eine Konstante $c > 0$, so dass

$$0 < c \|u_1 - u_2\|^2 \leq \mathcal{Q}(u_1 - u_2, u_1 - u_2) = \mathcal{Q}(u_1, u_1) - 2\mathcal{Q}(u_1, u_2) + \mathcal{Q}(u_2, u_2)$$

gilt und damit

$$2\mathcal{Q}(u_1, u_2) < \mathcal{Q}(u_1, u_1) + \mathcal{Q}(u_2, u_2).$$

Die Multiplikation der letzten Gleichung mit $\lambda(1 - \lambda)$, wobei $\lambda \in (0, 1)$ sei, liefert dann

$$2\lambda(1 - \lambda)\mathcal{Q}(u_1, u_2) < \lambda(1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_2, u_2).$$

Damit können wir für alle $u_1, u_2 \in \mathcal{H}$ die folgende Abschätzung finden:

$$\begin{aligned} \mathcal{Q}((1 - \lambda)u_1 + \lambda u_2, (1 - \lambda)u_1 + \lambda u_2) &= (1 - \lambda)^2 \mathcal{Q}(u_1, u_1) + 2\lambda(1 - \lambda)\mathcal{Q}(u_1, u_2) + \lambda^2 \mathcal{Q}(u_2, u_2) \\ &< (1 - \lambda)^2 \mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda(1 - \lambda)\mathcal{Q}(u_2, u_2) + \lambda^2 \mathcal{Q}(u_2, u_2) \\ &= (1 - \lambda)\mathcal{Q}(u_1, u_1) + \lambda\mathcal{Q}(u_2, u_2). \end{aligned}$$

⁵vgl. Voraussetzung (V1), S. 25

Da q linear ist erhalten wir durch $J_{1,y}^{\nu_1}(u) = \frac{1}{2}Q(u, u) + q(u)$ die strikte Konvexität des Zielfunktional. Als Summe (bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$) stetiger Funktionale ist auch das Zielfunktional stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . Die schwache Unterhalbstetigkeit folgt dann gemäß Lemma 8 aus der strikten Konvexität⁶. \square

Für den Schluss auf die Existenz einer optimalen Lösung fehlt die Beschränktheit der Menge \mathbf{U} . Um die fehlende Beschränktheit dieser Menge zu kompensieren, gehen wir für den Nachweis der Existenz einen Umweg über die Formulierung von Niveaumengen der restringierten Aufgabe (HP3) und geben deren Definition an durch

$$\mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha) := \left\{ u \in \mathbf{U} \mid J_{1,y}^{\nu_1}(u) \leq \alpha \right\} \quad (\text{B.3})$$

mit $\alpha > 0$ und weisen für diese die folgende Eigenschaft nach:

Lemma 53. *Für jedes zulässige Paar $u \in \mathbf{U}$ ist die Niveaumenge $\mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ konvex, abgeschlossen und beschränkt (bezüglich der Normtopologie von \mathcal{H}) und somit schwach kompakt.*

Beweis. Wir zeigen zunächst die Konvexität der Menge $\mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ und wählen dazu Elemente $u_1 \in \mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ und $u_2 \in \mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$. Das Zielfunktional $J_{1,y}^{\nu_1}$ ist strikt konvex (auf $\mathcal{H} \supset \mathbf{U}$) und somit insbesondere konvex⁷. Dies impliziert

$$J_{1,y}^{\nu_1}((1 - \lambda)u_1 + \lambda u_2) \leq (1 - \lambda)J_{1,y}^{\nu_1}(u_1) + \lambda J_{1,y}^{\nu_1}(u_2) = (1 - \lambda)J_{1,y}^{\nu_1}(u_1) + \lambda J_{1,y}^{\nu_1}(u_2)$$

für alle $u_1, u_2 \in \mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$.

Die Abgeschlossenheit der Menge $\mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ folgt aus der Stetigkeit der Operatoren Q und q (bzgl. der Normtopologie⁸) auf \mathcal{H} , welche wiederum die Stetigkeit des Zielfunktional $J_{1,y}^{\nu_1}$ auf \mathcal{H} bezüglich der Normtopologie impliziert. Aufbauend auf diesem Wissen betrachten wir nun eine konvergente Funktionenfolge $\{u_n\}$ mit $\{u_n\} \xrightarrow{\mathcal{H}} u_0$. Aufgrund der Abgeschlossenheit⁹ der Menge \mathbf{U} gilt $u_0 \in \mathbf{U}$. Die Stetigkeit des Funktional $J_{1,y}^{\nu_1}$ in der Normtopologie erlaubt dann die Abschätzung

$$J_{1,y}^{\nu_1}(u_0) = J_{1,y}^{\nu_1}(\lim_{n \rightarrow \infty} u_n) = \lim_{n \rightarrow \infty} J_{1,y}^{\nu_1}(u_n) \leq \alpha.$$

Damit ist $u_0 \in \mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ und somit die Niveaumenge $\mathcal{N}_{\mathbf{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ abgeschlossen.

Wir untersuchen nun die Beschränktheit der Niveaumengen: Gemäß Lemma 50 ist die Bilinearform Q koerzitiv auf \mathcal{H} und die Linearform q (norm-)beschränkt¹⁰ auf ganz \mathcal{H} . Damit existiert für alle $u \in \mathbf{U} \subset \mathcal{H}$ eine positive Konstante c_1 und eine Konstante $c_2 \in \mathbb{R}$, so dass

$$Q(u, u) \geq 2c_1 \|u\|_{\mathcal{H}}^2 \quad \text{und} \quad q(u) \geq -c_2 \|u\|_{\mathcal{H}}.$$

Demzufolge gilt für beliebige, aber feste Elemente $u \in \mathbf{U} \subset \mathcal{H}$

$$J_{1,y}^{\nu_1}(u) = \frac{1}{2}Q(u, u) + q(u) \geq c_1 \|u\|_{\mathcal{H}}^2 - c_2 \|u\|_{\mathcal{H}} = c_1 \left\{ \left(\|u\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right)^2 - \left(\frac{c_2}{2c_1} \right)^2 \right\},$$

⁶vgl. Lemma 8, S. 30

⁷vgl. Lemma 52, S. 192

⁸vgl. Lemma 50 und Lemma 51, S. 191 ff.

⁹vgl. Lemma 10, S. 32

¹⁰vgl. Lemma 51, S. 191

so dass wir nun für alle Elemente $u \in \mathcal{N}_{\mathcal{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ die Abschätzung

$$\alpha \geq J_{1,y}^{\nu_1}(u) \geq c_1 \left\{ \left(\|u\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right)^2 - \left(\frac{c_2}{2c_1} \right)^2 \right\}$$

erhalten, welche impliziert

$$\left| \|u\|_{\mathcal{H}} - \frac{c_2}{2c_1} \right| \leq \sqrt{\frac{1}{c_1} \alpha + \left(\frac{c_2}{2c_1} \right)^2}.$$

Damit ist $\mathcal{N}_{\mathcal{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ beschränkt auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$.

Damit ist $\mathcal{N}_{\mathcal{U}}(J_{1,y}^{\nu_1}(u), \alpha)$ eine auf \mathcal{H} (norm-)beschränkte, konvexe und abgeschlossene Menge. Nach Satz 2 und Satz 1 liefert dies den Schluss auf die schwache Kompaktheit der Menge $\mathcal{N}_{\mathcal{U}}(J_{1,y}^{\nu_1}(u), \alpha)$. Damit sind alle Aussagen des Satzes nachgewiesen. \square

Für den noch fehlenden Schluss auf die Existenz einer Lösung der Aufgabe (HP3) folgen wir nun der Argumentation aus der Beweisführung von Satz 9, gemäß derer ein strikt konvexes, schwach unterhalbstetiges Funktional auf der schwach kompakten Niveaumenge sein eindeutiges Minimum annimmt¹¹. \square

B.3 Existenz einer Lösung von $(\overline{HP4})$

$(\overline{HP4})$:

$$J_{2,y}^{\nu_1}(x) = \int_0^{\infty} \left\{ \frac{1}{2} (x^T(t)W(t)x(t) + w^T(t)x(t))e^{\beta_1 t} - x^T(t)[y'(t) + A^T(t)y(t)] \right\} dt$$

$$-x_0^T y(0) \longrightarrow \inf! \quad \text{u.d.N. } x \in x \in L_2^n(\mathbb{R}_+, e^{\beta_1 t}).$$

Satz 31. *Unter Gültigkeit der Voraussetzungen (V1) und (V2) besitzt das Problem $(\overline{HP4})$ eine eindeutige optimale Lösung.*

Beweis. Wir behandeln die Aufgabe $(\overline{HP4})$ und wollen diese wieder in den Kontext der Hilbertraummethode einbetten. Dazu definieren wir auf dem Hilbertraum $\mathcal{H} = L_2^n(\mathbb{R}_+, \nu_1)$ die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ mit

$$\mathcal{Q}(x_1, x_2) := \int_0^{\infty} x_1^T(t)W(t)x_2(t)\nu_1(t)dt, \quad (\text{B.4})$$

sowie $q : \mathcal{H} \rightarrow \mathbb{R}$ mit

$$q(x) := - \int_0^{\infty} x^T(t)[w(t)\nu_1(t) - y'(t) - A^T(t)y(t)]dt. \quad (\text{B.5})$$

¹¹vgl. Satz 9, S. 39

Zudem sei für beliebige aber feste Werte $y(0)$ die Konstante

$$q_0 := x_0^T y(0)$$

gegeben. Mittels dieser Setzungen können wir das Zielfunktional von $(\overline{HP4})$ gemäß

$$J_{2,y}^{\nu_1}(x) = \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0$$

beschreiben. Wir halten die Eigenschaften der eingeführten Abbildungen in den folgenden zwei Lemmata fest.

Lemma 54. *Durch die Abbildung $\mathcal{Q} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$, definiert in $(\overline{B.4})$, wird eine symmetrische, koerzitive Bilinearform auf \mathcal{H} definiert. Diese ist beschränkt bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ und damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Es lässt sich leicht nachrechnen, dass \mathcal{Q} für alle $x_1, x_2, \hat{x}_1, \hat{x}_2 \in \mathcal{H}$ und $\lambda \in \mathbb{R}$ folgenden Bedingungen genügt:

$$\begin{aligned} \mathcal{Q}(x_1 + \hat{x}_1, x_2) &= \mathcal{Q}(x_1, x_2) + \mathcal{Q}(\hat{x}_1, x_2) \\ \mathcal{Q}(x_1, x_2 + \hat{x}_2) &= \mathcal{Q}(x_1, x_2) + \mathcal{Q}(x_1, \hat{x}_2) \\ \mathcal{Q}(\lambda x_1, x_2) &= \lambda \mathcal{Q}(x_1, x_2). \end{aligned}$$

Aufgrund der Symmetrie der Matrix W gilt zudem für alle $x_1, x_2 \in \mathcal{H}$

$$\mathcal{Q}(x_1, x_2) = \mathcal{Q}(x_2, x_1).$$

Damit ist \mathcal{Q} eine symmetrische Bilinearform. Wir wollen nun zeigen, dass die so definierte Abbildung auch koerzitiv ist. Dafür erinnern wir uns, dass aufgrund der positiven Definitheit der Matrix W gilt $\underline{W} = \inf_{t \in \mathbb{R}} \lambda_{\min}(W)(t) > 0$. Mit diesem Wissen betrachten wir nun die folgende Abschätzung

$$\mathcal{Q}(x, x) = \int_0^{\infty} x^T(t) W(t) x(t) \nu_1(t) dt \geq \underline{W} \int_0^{\infty} x^2(t) \nu_1(t) dt \geq c \|x\|_{\mathcal{H}}^2,$$

mit der positiven Konstante $c = \underline{W} > 0$.

Zudem sichert uns die Raumzugehörigkeit der Matrix W mit $W \in L_{\infty}^{n \times n}(\mathbb{R}_+) \cap C^1(\mathbb{R}_+, \mathbb{R}^{n \times n})$ die Endlichkeit des Wertes

$$\|W\|_{\infty} = \operatorname{ess\,sup}_{t \in \mathbb{R}_+} |W(t)|$$

und sichert somit die Abschätzung

$$\begin{aligned} |\mathcal{Q}(x_1, x_2)| &= \left| \int_0^{\infty} x_1^T(t) W(t) x_2(t) \nu_1(t) dt \right| \leq \|W\|_{\infty} \int_0^{\infty} |x_1^T(t) x_2(t)| \nu_1(t) dt \\ &\leq \|W\|_{\infty} \left(\int_0^{\infty} |x_1(t)|^2 \nu_1(t) dt \right)^{1/2} \left(\int_0^{\infty} |x_2(t)|^2 \nu_1(t) dt \right)^{1/2} \\ &\leq \|W\|_{\infty} \|x_1\|_{\mathcal{H}} \|x_2\|_{\mathcal{H}}, \end{aligned}$$

womit die Beschränktheit von \mathcal{Q} auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist \mathcal{Q} ein auf \mathcal{H} lineares und (norm-)beschränktes Funktional und somit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . \square

¹²Dabei bezeichnet $\lambda_{\min}(M)(t)$ den vom Parameter t abhängigen kleinsten Eigenwert der symmetrischen Matrix M .

Lemma 55. *Durch die Abbildung $q : \mathcal{H} \rightarrow \mathbb{R}$, definiert in (B.5), wird ein lineares und bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ beschränktes Funktional auf \mathcal{H} definiert. Insbesondere ist q damit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} .*

Beweis. Die Linearität der Abbildung q lässt sich leicht nachrechnen. Wir zeigen daher nun die Beschränktheit der Linearform q bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ und treffen dazu die folgende Abschätzung

$$\begin{aligned} |q(x)| &= \left| \int_0^{\infty} x^T(t) [w(t)\nu_1(t) - y'(t) - A^T(t)y(t)] dt \right| \\ &\leq \left| \int_0^{\infty} x^T(t) w(t) \nu_1(t) dt \right| + \left| \int_0^{\infty} x^T(t) y'(t) dt \right| + \left| \int_0^{\infty} x^T(t) A^T(t) y(t) dt \right|. \end{aligned}$$

Durch die Raumzugehörigkeit¹³ der Funktionen w und A erhalten wir weiter

$$\begin{aligned} |q(x)| &\leq \left(\int_0^{\infty} |x(t)|^2 \nu_1(t) dt \right)^{1/2} \left(\int_0^{\infty} |w(t)|^2 \nu_1(t) dt \right)^{1/2} \\ &\quad + \left(\int_0^{\infty} |x(t)|^2 \nu_1(t) dt \right)^{1/2} \left(\int_0^{\infty} |y'(t)|^2 \nu_1^{-1}(t) dt \right)^{1/2} + \|A\|_{\infty} \left| \int_0^{\infty} x^T(t) y(t) dt \right| \\ &\leq \|x\|_{\mathcal{H}} \|w\|_{\mathcal{H}} + \|x\|_{\mathcal{H}} \|y'\|_{L_2^n(\mathbb{R}_+, \nu_1^{-1})} \\ &\quad + \|A\|_{\infty} \left(\int_0^{\infty} |x(t)|^2 \nu_1(t) dt \right)^{1/2} \left(\int_0^{\infty} |y(t)|^2 \nu_1^{-1}(t) dt \right)^{1/2} \\ &\leq \|x\|_{\mathcal{H}} \left(\|w\|_{\mathcal{H}} + \|y'\|_{L_2^n(\mathbb{R}_+, \nu_1^{-1})} + \|A\|_{\infty} \|y\|_{L_2^n(\mathbb{R}_+, \nu_1^{-1})} \right) \leq c \|x\|_{\mathcal{H}}, \end{aligned}$$

womit die Beschränktheit von q auf \mathcal{H} bzgl. der Norm $\|\cdot\|_{\mathcal{H}}$ bewiesen ist. Damit ist q ein auf \mathcal{H} lineares und (norm-)beschränktes Funktional und somit auch stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . \square

Diese so nachgewiesenen Eigenschaften lassen nun direkte Schlüsse auf das Zielfunktional zu. Diese wollen wir im folgenden Lemma festhalten.

Lemma 56. *Das Zielfunktional des Problems ($\overline{HP4}$) ist strikt konvex und schwach unterhalbstetig auf \mathcal{H} .*

Beweis. Wir zeigen zunächst die strikte Konvexität des Zielfunktionals und betrachten dafür zunächst den Operator \mathcal{Q} sowie beliebige Elemente $x_1, x_2 \in \mathcal{H}$ mit $x_1 \neq x_2$. Aufgrund der Koerzitivität der symmetrischen Bilinearform \mathcal{Q} existiert eine Konstante $c > 0$, so dass

$$0 < c \|x_1 - x_2\|^2 \leq \mathcal{Q}(x_1 - x_2, x_1 - x_2) = \mathcal{Q}(x_1, x_1) - 2\mathcal{Q}(x_1, x_2) + \mathcal{Q}(x_2, x_2)$$

gilt und damit

$$2\mathcal{Q}(x_1, x_2) < \mathcal{Q}(x_1, x_1) + \mathcal{Q}(x_2, x_2).$$

¹³vgl. Voraussetzung (V1), S. 25

Die Multiplikation der letzten Gleichung mit $\lambda(1 - \lambda)$, wobei $\lambda \in (0, 1)$ sei, liefert dann

$$2\lambda(1 - \lambda)\mathcal{Q}(x_1, x_2) < \lambda(1 - \lambda)\mathcal{Q}(x_1, x_1) + \lambda(1 - \lambda)\mathcal{Q}(x_2, x_2).$$

Damit können wir für alle $x_1, x_2 \in \mathcal{H}$ die folgende Abschätzung finden:

$$\begin{aligned} \mathcal{Q}((1 - \lambda)x_1 + \lambda x_2, (1 - \lambda)x_1 + \lambda x_2) &= (1 - \lambda)^2\mathcal{Q}(x_1, x_1) + 2\lambda(1 - \lambda)\mathcal{Q}(x_1, x_2) + \lambda^2\mathcal{Q}(x_2, x_2) \\ &< (1 - \lambda)^2\mathcal{Q}(x_1, x_1) + \lambda(1 - \lambda)\mathcal{Q}(x_1, x_1) + \lambda(1 - \lambda)\mathcal{Q}(x_2, x_2) + \lambda^2\mathcal{Q}(x_2, x_2) \\ &= (1 - \lambda)\mathcal{Q}(x_1, x_1) + \lambda\mathcal{Q}(x_2, x_2). \end{aligned}$$

Da q linear ist erhalten wir durch $J_{2,y}^{\nu_1}(x) = \frac{1}{2}\mathcal{Q}(x, x) - q(x) - q_0$ die strikte Konvexität des Zielfunktionals. Als Summe (bzgl. der Normtopologie des Raumes \mathcal{H}) stetiger Funktionale ist auch das Zielfunktional stetig auf \mathcal{H} bzgl. der Normtopologie des Raumes \mathcal{H} . Die schwache Unterhalbstetigkeit folgt dann gemäß Lemma [8](#) aus der strikten Konvexität^{[14](#)}. \square

Wir wollen nun zu einer äquivalenten Problemformulierung von $(\overline{HP4})$ über einer schwach kompakten Menge übergehen. Dazu betrachten wir die Niveaumengen

$$\mathcal{N}(J_{2,y}^{\nu_1}, \alpha) := \left\{ x \in L_2^n(\mathbb{R}_+, \nu_1) \mid J_{2,y}^{\nu_1}(x) - \mu \leq \alpha \right\} \quad (\text{B.6})$$

wobei

$$\mu := \inf_{x \in L_2^n(\mathbb{R}_+, \nu_0)} J_{2,y}^{\nu_0}(x) \quad (\text{B.7})$$

und weisen für diese die folgende Eigenschaft nach:

Lemma 57. Die Niveaumengen $\mathcal{N}(J_{2,y}^{\nu_1}, \alpha)$ ($\alpha > 0$), definiert in [\(B.6\)](#), sind beschränkt (bzgl. der Normtopologie des Raumes \mathcal{H}).

Beweis. Die Aussagen von Lemma [54](#) und Lemma [55](#) sichern die Endlichkeit des Wertes μ aus [\(B.7\)](#), denn es gilt

$$\begin{aligned} J_{2,y}^{\nu_1}(x) &= \frac{1}{2}\mathcal{Q}(x, x) - q(x) - q_0 \geq \frac{1}{2}c_1 \|x\|_{\mathcal{H}}^2 - c_2 \|x\|_{\mathcal{H}} - c_3 \\ &= \frac{1}{2}c_1 \left(\|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right)^2 - \frac{c_2^2}{2c_1} - c_3 \geq -\frac{c_2^2}{2c_1} - c_3 > -\infty, \end{aligned}$$

mit $c_1, c_2, c_3 > 0$. Damit können wir nun die folgende Abschätzung für alle Elemente x der Niveaumengen $\mathcal{N}(J_{2,y}^{\nu_1}, \alpha)$ treffen:

$$\alpha \geq J_{2,y}^{\nu_1}(x) - \mu \geq \frac{1}{2}c_1 \left(\|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right)^2 - \frac{c_2^2}{2c_1} - c_3 - \mu.$$

Dies impliziert insbesondere die Gültigkeit von

$$\left| \|x\|_{\mathcal{H}} - \frac{c_2}{c_1} \right| \leq \left[\frac{2}{c_1}(\alpha + \mu) + \left(\frac{c_1}{c_2} \right)^2 + \frac{2c_3}{c_1} \right]^{1/2}.$$

Damit ist $\mathcal{N}(J_{2,y}^{\nu_1}, \alpha)$ eine auf \mathcal{H} beschränkte Menge (bzgl. der Normtopologie des Raumes \mathcal{H}). \square

¹⁴vgl. Lemma [8](#), S. [30](#)

Aus der Gültigkeit des obigen Lemmas können wir nun schlussfolgern: $\|x\|_{\mathcal{H}}$ ist beschränkt, wenn wir nur solche x betrachten, deren Zielfunktionalwerte in einer α -Umgebung von μ liegen. Die Ergebnisse der Lemmata lassen daher nun den Schluss zu, dass das Problem

$$\begin{aligned} (\overline{\text{HP4}}') : \quad J_{2,y}^{\nu_1}(x) &= \frac{1}{2} \mathcal{Q}(x, x) - q(x) - q_0 \longrightarrow \inf ! \\ \text{u. d. N.} \quad x &\in L_2^n(\mathbb{R}_+, e^{\beta_1 t}) \\ &\|x\|_{\mathcal{H}} \leq \kappa, \kappa \text{ hinreichend groß,} \end{aligned}$$

äquivalent zu $(\overline{\text{HP4}})$ ist. Dabei existiert für $(\overline{\text{HP4}}')$ als ein Optimierungsproblem mit schwach unterhalbstetigem Zielfunktional über einer schwach kompakten Menge nach Satz [7](#) eine optimale Lösung [15](#) x_{y^*} . Diese ist aufgrund der strikten Konvexität des Zielfunktionalen eindeutig. Damit ist x_{y^*} auch eindeutige Lösung von $(\overline{\text{HP4}})$ und die Aussage des Satzes somit bewiesen. □

¹⁵vgl. Satz [7](#), S. [30](#)

Anhang C

Quellcode für das Beispiel aus Abschnitt 6.1

Durch den Aufruf der Funktion $LQR(\text{Polynomgrad}, \beta)$ übergibt der Anwender die gewünschte Anzahl an Kollokationspunkten N sowie den Gewichtungsparmeter β aus der Gewichtsfunktion. Nach der Berechnung der Momente der Gewichtsfunktion werden diese gemeinsam mit den Eingabeparametern β und N an *Kollokation.m* übergeben. Dort werden mit Hilfe der übergebenen Parameter die Kollokationspunkte (Zeile 37) und die Quadraturgewichte (Zeile 39-42) bestimmt. Für die Berechnung der verallgemeinerten Laguerre-Polynome wird die Funktion *LaguerreGenBeta(varargin)* von Matthias Trampusch aufgerufen. Der zugehörige M-File ist unter

<https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/26218/versions/5/previews/LaguerreGen.m/index.html>

zu finden. In den Zeilen 44 bis 55 werden Matrizen S und T erzeugt, welche die Werte der Funktionen S_k und T_k aus (5.11) an den zuvor berechneten Kollokationspunkten beinhalten. In den Zeilen 99 bis 116 erfolgt die Implementierung der Kollokationsgleichungen des zugehörigen freien Systems. Eine Lösung des freien Systems wird in Zeile 128 durch den Aufruf der Funktion *fsolve(funfrei,z0)* in der Form eines Newtonverfahrens ermittelt. Die so ermittelte Lösung des freien Problems dient schließlich als Startwert für das in Zeile 170 aufgerufene Newtonverfahren für die vorher in den Zeilen 140 bis 166 implementierten Kollokationsgleichungen des beschränkten Systems.

C.1 LQR.m

```
1 function [weights , nodes] = LQR(Polynomgrad , beta)
2 format long ;
3 old = 2 * Polynomgrad+1;
4 global N;
5 N=Polynomgrad;
6 momente = factorial(0:old-1)./beta.^(1:old);
7
8 [weights , nodes] = Kollokation(momente , beta ,N)
9
10
11 end
```

C.2 Kollokation.m

```

1 function [ weights , tau , kn , beta1 ] = Kollokation ( momente , beta , N)
2
3 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4 % Initialisierung der Modellparameter %
5 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
6
7 umax=11/5;
8 umin=-11/5;
9
10 aw=1;
11
12 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
13 % Berechnung der Quadraturstellen und -gewichte %
14 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
15 global TN;
16 TN=zeros (N,N) ;
17
18 Hauptdiagonale = vpa(1/beta .* [1:2:2*N-1]) ;
19 Nebendiagonale = vpa(1/beta .* [1:1:N-1]) ;
20
21 TN = diag (Hauptdiagonale) ;
22 TN = TN + diag (Nebendiagonale ,1) ;
23 TN = TN + diag (Nebendiagonale ,-1) ;
24
25 global LagPoly ;
26
27 for k = 1:N+1
28     LagPoly{k} = LaguerreGenBeta (k-1,1,beta) ; % Orthogonale Polynome
29 end
30
31 for j= 1:N+1
32     LagPoly{j}=flipud (LagPoly{j})
33     AblLagPoly{j}=polyder (LagPoly{j}) ; %Ableitung der Orthog. L-Polynome
34 end
35
36 [V,D] = eig (TN) ; % Eigenwerte bestimmen
37 global tau ; % Nullstellen von L_N
38 tau = diag (D) % Quad.-stellen sind Eigenw. von T
39
40 weights = vpa (zeros (N,1)) ; % Quadraturgewichte sind skalierte
41 for k = 1:N % Skalarprodukte der Eigenvektoren
42     weights (k) = vpa (momente (1) * V (1,k)^2 / (V (:,k)' * V (:,k)))
43 end
44
45 tau = double (tau) ;
46
47 global T ;
48 global S ;
49
50 S=zeros (N,N) ;
51 T=zeros (N,N) ;
52
53 for k=1:N
54     for j=1:N
55         S (j , k)=tau (j) * (-1+(1/(k+1)) * polyval (LagPoly {k+1}, tau (j))) ;
56         T (j , k)=(-1+polyval (LagPoly {k+1}, tau (j)))-polyval (LagPoly {k}, tau (j)) ;
57     end

```

```

58 end
59
60
61 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
62 % einige Unterfunktionen zur Behandlung der auftretenden Polynome %
63 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
64
65 % Polynome miteinander multiplizieren
66 function p = mult(s,t)
67     p = zeros(length(s)+length(t)-1,1);
68
69     for j = 1:length(s);
70         p(j:length(t)+j-1) = p(j:length(t)+j-1) + s(j) * t(:);
71     end
72
73 end
74 % Polynome addieren
75 function r = add(p,q)
76     if length(q) >= length(p)
77         r = q(:);
78         r(1:size(p,1)) = r(1:size(p,1)) + p(:);
79     else
80         r = p(:);
81         r(1:size(q,1)) = r(1:size(q,1)) + q(:);
82     end
83 end
84 % Integral auswerten
85 function r = integ(p,momente)
86     r = (p(:))'*momente(1:length(p))';
87 end
88
89 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
90 % Ende %
91 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
92
93 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
94 % Berechnung der freien Lsg. zur Ermittlung eines Startwertes %
95 % des Newtonverfahrens %
96 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
97
98 m= N+1;
99
100 global z1;
101 z1=zeros(m,1);
102 % Gleichungssystem F(z1)=0
103
104 function Frei = myfunfrei(z1)
105
106     for k=1:N
107         ak(k)=z1(k);
108     end
109
110     A=z1(N+1);
111
112     for i=1:N+1 % Nullvektor wird erzeugt
113         Frei(i)=0;
114     end
115
116 % partielle Ableitungen nach Koeffizienten a

```

```

117 for j=1:N
118     for i=1:N
119         Frei(j)=Frei(j)+weights(i).*sum(ak(1:N).*(T(i,1:N)*T(i,j)))+weights(i)*aw*
120             S(i,j)*beta+S(i,j)*weights(i)*(sum(ak(1:N).*S(i,1:N))+aw*tau(i)+A);
121     end
122 end
123 % partielle Ableitungen nach Koeffizienten A
124 j=N+1;
125 for i=1:N
126     Frei(j)=Frei(j)+weights(i)*aw*beta+weights(i)*(sum(ak(1:N).*S(i,1:N))+aw*
127         tau(i)+A);
128 end
129 end
130
131 funfrei=@myfunfrei;
132 z0=ones(m,1);
133 [z1,fval]=fsolve(funfrei,z0) % Loesen des freien Systems
134
135 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
136 % Newtonverfahren des beschraenkten Systems %
137 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
138
139 global z;
140 z=zeros(m,1);
141
142 % Gleichungssystem F(z)=0
143
144
145 function F = myfun(z)
146
147     for k=1:N
148         ak(k)=z(k);
149     end
150
151     A=z(N+1);
152
153
154     for i=1:N+1 % Nullvektor wird erzeugt
155         F(i)=0;
156     end
157
158 % partielle Ableitungen nach Koeffizienten a
159 for j=1:N
160     for i=1:N
161         F(j)=F(j)+weights(i).*sum(ak(1:N).*(T(i,1:N)*T(i,j)))+weights(i)*aw*S(i,j)
162             *beta+S(i,j)*weights(i)*min(max(umin,sum(ak(1:N).*S(i,1:N))+aw*tau(i)
163             +A),umax);
164     end
165 end
166
167 % partielle Ableitungen nach Koeffizienten A
168 j=N+1;
169 for i=1:N
170     F(j)=F(j)+weights(i)*aw*beta+weights(i)*min(max(umin,sum(ak(1:N).*S(i,1:N)
171         +aw*tau(i)+A),umax);
172 end

```

```

171 end
172
173 fun=@myfun;
174
175 [z , fval]=fsolve(fun , z1) % Loesen des beschraenkten Systems
176
177 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
178 % Ende Newtonverfahren %
179 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
180 end

```

C.3 LaguerreGenBeta.m

```

1 function y = LaguerreGenBeta(varargin)
2
3 %LaguerreGen calculates the generalized Laguerre polynomial L{n,alpha,beta}
4 %
5 % The function computes the generalized Laguerre polynomial L{n,alpha,beta}
6 % and is based on LaguerreGen(varargin).m from Matthias Trampisch.
7 % Here we considered the Laguerre-weight function including the additional
8 % parameter beta: t^{alpha}* e^{-beta t}
9
10
11 % Input:
12 % - N = nonnegative integer as degree level
13 % - alpha >= -1 real number (input is optional)
14 % - beta >0 real number
15 %
16 % The output is formatted as a polynomial vector of length (n+1).
17 % [a_0 a_1 a_2 ... a_N]
18 % The lowest coefficient is the first element.
19 % Ln = a_0+a_1*t+a_2*t^2+...+a_N*t^N
20 %
21
22 %%% ===== %
23 % set default parameters and rename input
24 % ===== %
25 if (nargin == 2) %only one parameter "n" supplied
26     N = varargin{1};
27     alpha = 0; %set default value for alpha
28     beta = 0;
29 elseif (nargin == 3) %at least two parameters supplied
30     N = varargin{1};
31     alpha = varargin{2};
32     beta = varargin{3};
33 end;
34 %%% ===== %
35 % error checking of input parameters
36 % ===== %
37 %if (nargin == 0) || (nargin > 2) || (n~=abs(round(n))) || (alpha<-1)
38 % error('n must be integer, and (optional) alpha >= -1');
39 %end;
40 %%% ===== %
41 % Recursive calculation of generalized Laguerre polynomial
42 % ===== %
43 L=zeros(N+1); %reserve memory for faster storage
44 switch N

```

```

45 case 0
46     L(1,:) = 1;
47 otherwise %n>1 so we need to do recursion
48     L(1,:) = [zeros(1,N), 1];
49     L(2,:) = [zeros(1, N-1), -1, (alpha+1)];
50     for i=3:N+1
51         A1 = 1/(i-1) * (conv([zeros(1, N-1), -1, (2*(i-1)+alpha-1)], L(i-1,:)));
52         A2 = 1/(i-1) * (conv([zeros(1, N), ((i-1)+alpha-1)], L(i-2,:)));
53         B1=A1(length(A1)-N:1:length(A1));
54         B2=A2(length(A2)-N:1:length(A2));
55         L(i,:) = B1-B2; % i-th row corresponds to L{i-1, alpha}
56     end;
57 end;
58 %% ===== %
59 %                                     Define output
60 % ===== %
61 p=fliplr(L(N+1,:)); %last row is the gen. Laguerre polynomial L{n, alpha}
62 Lbeta=zeros(1,N+1);
63
64 for i=1:N+1
65     j=i-1;
66     betaPotenz=beta^j;
67     Lbeta(i)=p(i)*betaPotenz;
68 end;
69
70 LbetaN=fliplr(Lbeta);
71 y=LbetaN();

```


Literaturverzeichnis

- [AF88] ARSCOTT, F.M.; FILIPPOV, A.F.: *Differential Equations with Discontinuous Righthand Sides: Control Systems*. Mathematics and its Applications. Springer Netherlands, 1988.
- [AF03] ADAMS R.A.; FOURNIER, J.J.F.: *Sobolev Spaces*. ISSN. Elsevier Science, 2003.
- [AK70] ARROW, K.J.; KURZ, M.: *Public Investment, the Rate of Return, and Optimal Fiscal Policy*. Resources for the future. Resources for the Future, 1970.
- [AK04] ASEEV, S.; KRYAZHIMSKIY, A.: The pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. *SIAM J. Control and Optimization*, 43:1094–1119, 01 2004.
- [AKS16] ALT, W.; KAYA, C.; SCHNEIDER, C.: Dualization and discretization of linear-quadratic control problems with bang–bang solutions. *EURO Journal on Computational Optimization*, 4:47–77, 02 2016.
- [Alt12] ALT, W.: *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Masterclass. Springer Berlin Heidelberg, 2012.
- [Arr68] ARROW, K.J.: Applications of control theory to economic growth. mathematics of the decision sciences part 2, 85-119. *American Mathematical Society, Providence*, 1968.
- [BKM14] BURACHIK, R.; KAYA, C.; MAJEED, S.: A duality approach for solving control-constrained linear-quadratic optimal control problems. *SIAM Journal on Control and Optimization*, 52:1423–1456, 05 2014.
- [Bla82] BLANCHARD, E.; BRÜNING, P.: *Direkte Methoden der Variationsrechnung*. Springer-Verlag, Wien, New York, 1982.
- [BPDP14] BISWAS, M.H.A.; PAIVA, L.; DE PINHO, M.D.R.: A seir model for control of infectious diseases with constraints. *Mathematical biosciences and engineering: MBE*, 11, 08 2014.
- [Bur19] BURTCHEM, A.: *Ein Beitrag zur Behandlung von Optimalsteuerungsaufgaben mit unendlichem Zeithorizont*. Dissertation, Brandenburgische Technische Universität Cottbus-Senftenberg, 2019.
- [BZ18] BURTCHEM, A.; ZIEMANN, T.: *Analytische Hilfsmittel der Steuerungstheorie*. Reihe Mathematik: Preprint, 01/2018. Brandenburgische Technische Universität, Mathematisches Institut, 2018.

- [Dav75] DAVIS, P.J.: *Interpolation and approximation*. Dover books on advanced mathematics. Dover Publications, 1975.
- [Dik16] DIKARIEV, I.: Ein Beitrag zur Existenz optimaler Lösungen bei Steuerungsproblemen mit unendlichem Zeithorizont. Master's thesis, BTU Cottbus-Senftenberg, 2016.
- [dPKM15] DE PINHO, M.D.R.; KORNIENKO, I.; MAURER, H.: Optimal control of a seir model with mixed constraints and l1 cost. In António Paulo Moreira, Aníbal Matos, and Germano Veiga, editors, *CONTROLO'2014 – Proceedings of the 11th Portuguese Conference on Automatic Control*, pages 135–145, Cham, 2015. Springer International Publishing.
- [DS88] DUNFORD, N.; SCHWARTZ, J.T.: *Linear Operators Part I: General Theory*. Wiley Classics Library Edition, 1988.
- [Els07] ELSTRODT, J.: *Maß- und Integrationstheorie*. Grundwissen Mathematik. Springer London, Limited, 2007.
- [Fau09] FAUSTMANN, M.: *Gewichtete Sobolevräume und nicht "klassische" lineare elliptische Randwertprobleme*. Bachelorarbeit. Technische Universität Wien, 2009.
- [FH86] FEICHTINGER, G.; HARTL, R.F.: *Optimale Kontrolle ökonomischer Prozesse: Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften*. Walter de Gruyter, 1986.
- [FH07] FREUND, R.W.; HOPPE, R.W.: *Stoer/Bulirsch: Numerische Mathematik 1*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2007.
- [For09] FORSTER, O.: *Analysis 3*. Analysis. Vieweg + Teubner, 2009.
- [Fun92] FUNARO, D.: *Polynomial Approximation of Differential Equations*. Number Bd. 8 in Lecture notes in physics: Monographs. Springer-Verlag, 1992.
- [Gao13] GAO, D.Y.: *Duality principles in nonconvex systems: theory, methods and applications*, volume 39. Springer Science & Business Media, 2013.
- [GHR11] GARG, D.; HAGER, W.W.; RAO, A.V.: Pseudospectral methods for solving infinite-horizon optimal control problems. *Automatica*, 47(4):829 – 837, 2011.
- [GJ08] GRÜNE, L.; JUNGE, O.: *Gewöhnliche Differentialgleichungen: Eine Einführung aus der Perspektive der dynamischen Systeme*. Bachelorkurs Mathematik. Vieweg+Teubner Verlag, 2008.
- [GK02] GEIGER, C.; KANZOW, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Lehrbuch Masterclass. Springer Berlin Heidelberg, 2002.
- [Göp73] GÖPFERT, A.: *Mathematische Optimierung in allgemeinen Vektorräumen*. Mathematisch-naturwissenschaftliche Bibliothek. Teubner, BSB, 1973.

- [Grü98] GRÜNE, L.: Asymptotic controllability and exponential stabilization of nonlinear control systems at singular points. *SIAM J. Control Optim.*, 36(5):1485–1503, September 1998.
- [GTX08] GAO, S.; TENG, Z.; XIE, D.: The effects of pulse vaccination on seir model with two time delays. *Applied Mathematics and Computation*, 201(1):282 – 292, 2008.
- [Hah59] HAHN, W.: *Theorie und Anwendungen der direkten Methode von Ljapunov*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge. Springer-Verlag Berlin Heidelberg, 1959.
- [Heu04] HEUSER, H.: *Lehrbuch der Analysis. 2. Mit 633 Aufgaben, zum Teil mit Lösungen*. Lehrbuch der Analysis / von Harro Heuser. Teubner, 2004.
- [Isk18] ISKE, A.: *Approximation*. Masterclass. Springer Berlin Heidelberg, 2018.
- [ITM09] IOFFE, A.D.; TIKHOMIROV, V.V.M.; MAKOWSKI, K.: *Theory of extremal problems*. Studies in Mathematics and its Applications. Elsevier Science, 2009.
- [Kac01] KACZOREK, T.: Externally and internally positive time-varying linear systems. *International Journal of Applied Mathematics and Computer Science*, 11:957–964, 2001.
- [Kal59] KALMAN, R.: On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):110–110, 1959.
- [Kal60] KALMAN, R.: Contributions to the theory of optimal control. In *Boletín de la Sociedad Matemática Mexicana* 5, pages 102–105, 1960.
- [KJF77] KUFNER, A.; JOHN, O.; FUCIK, S.: *Function Spaces*. Mechanics: Analysis. Springer Netherlands, 1977.
- [Klö79] KLÖTZLER, R.: On a general conception of duality in optimal control. *Equadiff IV. Proceedings of the Czechoslovak Conference on Differential Equations and their Applications held in Prague, August 22 – 26, 1977*. Springer, Berlin (*Lecture Notes in Mathematics* **703**, pages 189 – 196, 1979.
- [Klö80] KLÖTZLER, R.: Starke Dualität in der Steuerungstheorie. *Mathematische Nachrichten*, 95(1):253–263, 1980.
- [Klö81] KLÖTZLER, R.: Dualität bei diskreten Steuerungsproblemen. *Mathematische Operationsforschung und Statistik.*, 12:411–420, 1981.
- [Kuf81] AKUFNER, A.: *Weighted Sobolev spaces*. Teubner-Texte zur Mathematik. Teubner, 1981.
- [LF06] LEKONE, P.E.; FINKENSTÄDT, B.F.: Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study. *Biometrics*, 62(4):1170–1177, 2006.
- [LM95] LI, M.Y.; MULDOWNY, J.S.: Global stability for the seir model in epidemiology. *Mathematical Biosciences*, 125(2):155 – 164, 1995.

- [LML10] LIN, F.; MUTHURAMAN, K.; LAWLEY, M.: An optimal control theory approach to non-pharmaceutical interventions. *BMC infectious diseases*, 10:32, 02 2010.
- [LPW08] LYKINA, V.; PICKENHAIN, S.; WAGNER, M.: Different interpretations of the improper integral objective in an infinite horizon control problem. *Journal of Mathematical Analysis and Applications - J MATH ANAL APPL*, 340:498–510, 04 2008.
- [Lya92] LYAPUNOV, A.M. : *General Problem of the Stability Of Motion*. Control Theory and Applications Series. Taylor & Francis, 1992.
- [Lyk10] LYKINA, V.: *Beiträge zur Theorie der Optimalsteuerungsprobleme mit unendlichem Zeithorizont*. Dissertation, Brandenburgische Technische Universität Cottbus, 2010.
- [Mag82] MAGILL, M.J.P.: Pricing infinite horizon programs. *Journal of Mathematical Analysis and Applications*, 88(2):398–421, 1982.
- [PB18] PICKENHAIN, S.; BURTCHEM, A.: Regulator problems on unbounded domains: Stationarity–optimal control–asymptotic controllability. *Vietnam Journal of Mathematics*, 46, 09 2018.
- [PB19] PICKENHAIN, S.; BURTCHEM, A.: Problems in the calculus of variations on unbounded intervals—fourier–laguerre analysis and approximations. *Vietnam Journal of Mathematics*, 04 2019.
- [Pic09] PICKENHAIN, S.: Duality in optimal control with first order differential equations. In *Encyclopedia of Optimization, Second Edition*, pages 805–811. Springer, 2009.
- [Pic10a] PICKENHAIN, S.: On adequate transversality conditions for infinite horizon optimal control problems—a famous example of halkan. In *Dynamic systems, economic growth, and the environment*, pages 3–22. Springer, 2010.
- [Pic10b] PICKENHAIN, S.: On infinite horizon optimal control of a lotka-voltera-system. In Moritz Diehl, Francois Glineur, Elias Jarlebring, and Wim Michiels, editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 277–286, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [Pic14] PICKENHAIN, S.: Hilbert space treatment of optimal control problems with infinite horizon. In Hans Georg Bock, Xuan Phu Hoang, Rolf Rannacher, and Johannes P. Schlöder, editors, *Modeling, Simulation and Optimization of Complex Processes - HPSC 2012*, pages 169–182, Cham, 2014. Springer International Publishing.
- [Pic15] PICKENHAIN, S.: Infinite horizon optimal control problems in the light of convex analysis in hilbert spaces. *Set-Valued and Variational Analysis*, 23(1):169–189, Mar 2015.
- [PW12] PICKENHAIN, S.; WENZKE, D.: Transformation of infinite horizon optimal-control problems. *Preprint Reihe Mathematik.*, M-02/2012, 2012.

- [Ram28] RAMSEY, F.P.: A mathematical theory of saving. *Econom. J.*, 28:543–559, 1928.
- [SAS13] SCHNEIDER, C.; ALT, W.; SEYDENSCHWANZ, M.: *EAGLE-STARHILFE Optimale Steuerung: Theorie und numerische Verfahren*. Edition am Gutenbergplatz Leipzig, 2013.
- [SBD20] 2020 STATISTISCHES BUNDESAMT (DESTATIS): Bevölkerung und Erwerbstätigkeit 1946-2018. *DESTATIS*, page 2, 2020.
- [Sol70] SOLOW, R.M.: *Growth theory. An exposition*. Oxford: Clarendon Press., 1970.
- [Son83] SONTAG, E.: A lyapunov-like characterization of asymptotic controllability. *Siam Journal on Control and Optimization - SIAM J CONTR OPTIMIZAT*, 21, 05 1983.
- [Son98] SONTAG, E.: *Mathematical Control Theory: Deterministic Finite Dimensional Systems (2nd Ed.)*. Springer-Verlag, Berlin, Heidelberg, 1998.
- [SSG39] SZEGŐ, G.: *Orthogonal Polynomials*. Bd. XXIII. American Mathematical Society, 1939.
- [Tag20] TAGESSCHAU: Verbreitung des Coronavirus, April 2020. zuletzt aufgerufen am 27.04.2020.
- [Töd20] TÖDTER, K.H.: Ein SIRD-Modell zur Infektionsdynamik mit endogener Behandlungskapazität und Lehren für Corona-Statistiken. Technical report, IMFS Working Paper Series, 2020.
- [Tri70] TRICOMI, F.G.: *Vorlesungen über Orthogonalreihen*. Springer, Berlin, Heidelberg, 1970.
- [Trö05] TRÖLTZSCH, F.: *Optimale Steuerung partieller Differentialgleichungen: Theorie, Verfahren und Anwendungen*. Vieweg, 2005.
- [Wag09] WAGNER, M.: Pontryagin's maximum principle for multidimensional control problems in image processing. *Journal of Optimization Theory and Applications*, 140:543–576, 03 2009.
- [Wal00] WALTER, W.: *Gewöhnliche Differentialgleichungen: Eine Einführung*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2000.
- [Wen20] WENZKE, D.: *Beiträge zu Problemen der Optimalsteuerung mit unendlichem Zeithorizont - Transformationseigenschaften* -. Dissertation, Brandenburgische Technische Universität Cottbus-Senftenberg, 2020.
- [Wer07] WERNER, D.: *Funktionalanalysis*. Springer-Lehrbuch. Springer London, Limited, 2007.
- [XyHf13] XIAO-YONG, Z.; HUAFENG, W.: Generalized laguerre approximation and its application to ordinary differential equation. In *Applied Mathematics and Information Sciences* 7, 2013.

- [Yos80] YOSIDA, K.: *Functional Analysis*. Springer, Berlin - Heidelberg - New York, 6th ed., 1980.
- [Zei12] ZEIDLER, E.: *Applied Functional Analysis: Applications to Mathematical Physics*. Springer Publishing Company, Incorporated, 2012.
- [Zie19] ZIEMANN, T.: *Notwendige Optimalitätskriterien für eine Klasse von Steuerungsaufgaben mit einem a priori gegebenen unendlichem Zeithorizont*. Dissertation, Brandenburgische Technische Universität Cottbus-Senftenberg, 2019.
- [ZP18] ZIEMANN, T.; PICKENHAIN, S.: First-order necessary optimality conditions for infinite horizon optimal control problems with linear dynamics and convex objective. *Optimization*, 67(6):757–782, 2018.