

Skript zur Vorlesung  
MATHEMATISCHE  
SYSTEM- UND REGELUNGSTHEORIE  
Sommersemester 2009, TU Chemnitz

Peter Benner  
Professur Mathematik in Industrie und Technik  
Fakultät für Mathematik  
Technische Universität Chemnitz  
09107 Chemnitz  
Germany  
eMail: [benner@mathematik.tu-chemnitz.de](mailto:benner@mathematik.tu-chemnitz.de)  
URI: <http://www.tu-chemnitz.de/~benner>





# Vorwort



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Analyse von Regelungssystemen</b>	<b>9</b>
2.1	Steuerbarkeit . . . . .	9
2.2	Stabilisierbarkeit . . . . .	19
2.3	Beobachtbarkeit und Entdeckbarkeit . . . . .	21
2.4	Die Staircase-Form der Systemtheorie . . . . .	24
<b>3</b>	<b>Stabilisierung, Lyapunovgleichungen und Polvorgabe</b>	<b>31</b>
3.1	Stabilitätstheorie nach Lyapunov . . . . .	32
3.2	Stabilisierung mit Lyapunovgleichungen . . . . .	37
3.3	Numerische Lösung von Lyapunovgleichungen . . . . .	40
3.4	Stabilisierung durch Polvorgabe . . . . .	43
<b>4</b>	<b>Optimale Steuerung</b>	<b>51</b>
<b>5</b>	<b>Beobachterentwurf, Kalman-Filter und LQG Regelung</b>	<b>77</b>
	<b>Literatur</b>	<b>83</b>



# Abbildungsverzeichnis

1.1	Systembeschreibung als Black-Box. . . . .	1
1.2	Nichtlineares System als Black-Box. . . . .	2
1.3	Pendel als mathematisches Modell eines eingelenkigen Rotqtionsarms. . . . .	6
1.4	Inverses Pendel: Steuerung in aufrechte Position. . . . .	6
4.1	Typisches Spektrum einer Hamiltonischen Matrix . . . . .	64
4.2	Spektrum einer Hamiltonischen Matrix, einmal berechnet mit dem unstrukturierten QR Algorithmus und einmal mit einem strukturerhaltenden Algorithmus. . . . .	71





# Kapitel 1

## Einleitung

Wir werden dynamische Systeme betrachten, die einen physikalischen, technischen oder ökonomischen Prozeß beschreiben. Diese sollen mit Hilfe von Eingangsgrößen so beeinflußt werden, daß gewisse Ausgangsgrößen ein gewünschtes Verhalten zeigen. Schematisch kann man sich das wie in Abbildung 1.1 vorstellen.

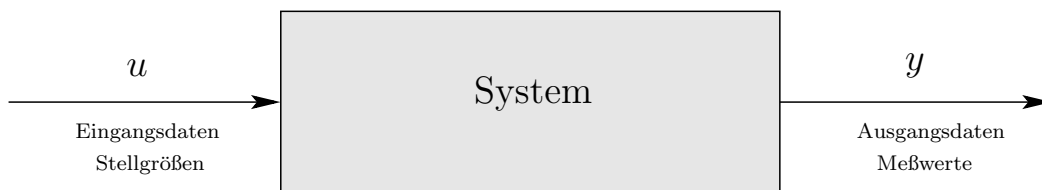


Abbildung 1.1: Systembeschreibung als Black-Box.

**Beispiel 1.1** *Verändert die Europäische Zentralbank den Leitzinssatz, so beeinflusst dies die Entwicklung des deutschen Aktien-, Devisen- und Kapitalmarkts. Z.B. steigen oder fallen der Wechselkurs zwischen € und US-\$ und die Aktienkurse.*

*Betrachtet man den Aktien-, Devisen- und Kapitalmarkt als dynamisches System, dann kann man die Höhe des Leitzinssatzes als eine Eingangsgröße  $u$  und den deutschen Aktienindex DAX als eine Ausgangsgröße  $y$  betrachten. Abbildung 1.1 liefert dann eine sehr grobe Beschreibung der Märkte (die natürlich überhaupt nicht beim Spekulieren hilft...) Man beachte, daß in diesem Beispiel das System selbst nicht durch mathematische Gleichungen beschrieben ist (und auch bis heute in seiner ganzen Komplexität nicht durch ein mathematisches Modell beschreibbar ist). Das System ist also eine "Black-Box" — man kennt den Systemzustand und seine Veränderung nicht, man kann aber durchaus Aussagen über die Beziehungen zwischen Ein- und Ausgangsgrößen gewinnen.*

Wir werden hier instationäre, also zeitabhängige, Prozesse betrachten, d.h. Eingangs- und Ausgangsgrößen sind Funktionen der Zeit  $t$ , also  $u(t)$  und  $y(t)$ . Die möglicherweise nicht genau bekannten Zustände des Systems zum Zeitpunkt  $t$  bezeichnen wir mit  $x(t)$ .

Wir werden davon ausgehen, daß das dynamische Verhalten der Systeme durch ein System von Differentialgleichungen erster Ordnung beschrieben wird. (Erinnere: Differentialgleichungen höherer Ordnung können auf Systeme erster Ordnung zurückgeführt werden.) Recht allgemein kann man die Systeme, die wir hier betrachten, wie in folgender Definition beschreiben.

**Definition 1.2** Ein (nichtlineares) Regelungssystem (oder gesteuertes System) genügt für  $t \in [t_0, t_f]$ ,  $t_0 < t_f \leq \infty$ , den folgenden Gleichungen:

$$\dot{x} = f(t, x, u) \quad (\text{Zustandsgleichung}), \quad (1.1)$$

$$x(t_0) = x^0 \in \mathcal{X} \quad (\text{Anfangsbedingung/ -zustand}), \quad (1.2)$$

$$y = g(t, x, u) \quad (\text{Ausgangsgleichung}). \quad (1.3)$$

Dabei sind

$$x(t) : [t_0, t_f] \rightarrow \mathcal{X} \text{ der Zustand(svektor),}$$

$$u(t) : [t_0, t_f] \rightarrow \mathcal{U} \text{ die Steuerung oder Stellgrößen,}$$

$$y(t) : [t_0, t_f] \rightarrow \mathcal{Y} \text{ der Ausgang(svektor).}$$

und

$$\mathcal{X} \subset \mathbb{R}^n \quad \text{der Zustandsraum,}$$

$$\mathcal{U} \subset \mathbb{R}^m \quad \text{der Eingangsraum,}$$

$$\mathcal{Y} \subset \mathbb{R}^p \quad \text{der Ausgangsraum.}$$

$n$  ist die Ordnung des Systems (auch: Zustandsraumdimension, falls  $\mathcal{X} = \mathbb{R}^n$ ). Das System heißt autonom (zeit-invariant), falls

$$f(t, x, u) \equiv f(x, u) \quad \text{und} \quad g(t, x, u) \equiv g(x, u),$$

d.h. für  $u(t) \equiv 0$  ist  $\dot{x} = f(x)$  eine autonome Differentialgleichung.

Für solche Systeme kann man also die Abbildung 1.1, wie in Abbildung 1.2 geschehen, ergänzen:

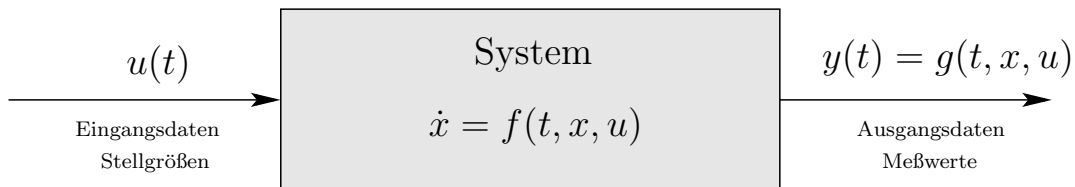


Abbildung 1.2: Nichtlineares System als Black-Box.

Versucht man nun, einen physikalischen, technischen oder ökonomischen Prozeß durch Gleichungen der Form (1.1)–(1.3) zu modellieren, sind folgende Aspekte von Bedeutung:

1. Was sind die “freien” Eingangsparameter (Steuer-/Stellgrößen)?
2. Was sind die Zustandsgrößen?
3. Welche Größen kann ich messen bzw. beobachten? (Einige Zustandsgrößen oder alle oder nur davon abgeleitete Größen?)
4. Was ist der funktionale Zusammenhang?
5. Ist eine (zeit-)kontinuierliche Modellierung wie in (1.1)–(1.3) angemessen oder benötigt man ein diskretes Modell, d.h. eine Beschreibung der Dynamik durch Differenzgleichungen?

Oft sind auch gemischte Modelle (sogenannte *Hybridsysteme*) nötig, da sich einige Modellgrößen kontinuierlich, andere nur diskret beschreiben lassen.

6. Verhalten sich die Modellgrößen deterministisch oder stochastisch?

Im Folgenden werden wir meist davon ausgehen, daß wir es mit einer kontinuierlichen Zeitskala zu tun haben und daß ein deterministisches Modell zugrunde liegt. Später werden wir auch Möglichkeiten untersuchen, wie man gewisse stochastische Einflüsse berücksichtigen kann. Außerdem werden wir durchgehend einen einfacheren funktionalen Zusammenhang als in (1.1)–(1.3) zugrunde legen: Wir nehmen an, daß  $f$  und  $g$  linear-affine Funktionen sind.

**Definition 1.3** *Ein lineares Regelungssystem liegt vor, falls  $\mathcal{X} = \mathbb{R}^n$ ,  $\mathcal{U} = \mathbb{R}^m$ ,  $\mathcal{Y} = \mathbb{R}^p$  und*

$$\begin{aligned} f(t, x, u) &= A(t)x(t) + B(t)u(t), \\ g(t, x, u) &= C(t)x(t) + D(t)u(t), \end{aligned}$$

wobei  $A(t) : [t_0, \infty] \rightarrow \mathbb{R}^{n \times n}$ ,  $B(t) : [t_0, \infty] \rightarrow \mathbb{R}^{n \times m}$ ,  $C(t) : [t_0, \infty] \rightarrow \mathbb{R}^{p \times n}$ ,  $D(t) : [t_0, \infty] \rightarrow \mathbb{R}^{p \times m}$  matrixwertige und hinreichend glatte Funktionen sind.

Für autonome Systeme gilt  $A(t) \equiv A$ ,  $B(t) \equiv B$ , etc. Speziell sprechen wir von linearen zeit-invarianten Systemen (LTI Systeme, von engl. “linear time-invariant”), wenn das System folgenden Gleichungen genügt:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(t_0) = x^0 \in \mathbb{R}^n, \quad (1.4)$$

$$y(t) = Cx(t) + Du(t). \quad (1.5)$$

Ein zeitvariantes System (LTV System, von engl. “linear time-varying”) ist gegeben durch

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x^0 \in \mathbb{R}^n, \quad (1.6)$$

$$y(t) = C(t)x(t) + D(t)u(t). \quad (1.7)$$

Analog kann man *lineare (zeit-)diskrete Systeme* für  $k = 0, 1, 2, \dots$  wie folgt beschreiben:

$$x_{k+1} = A_k x_k + B_k u_k, \quad x_0 = x^0 \in \mathbb{R}^n, \quad (1.8)$$

$$y_k = C_k x_k + D_k u_k, \quad (1.9)$$

wobei  $A_k \in \mathbb{R}^{n \times n}$ ,  $B_k \in \mathbb{R}^{n \times m}$ ,  $C_k \in \mathbb{R}^{p \times n}$ ,  $D_k \in \mathbb{R}^{p \times m}$ . Im zeitinvarianten Fall gilt  $A_k \equiv A, B_k \equiv B$ , usw.

**Bemerkung 1.4** Für autonome Systeme können wir o.B.d.A. annehmen, daß  $t_0 = 0$ : Bewegen wir uns von  $x^0$  nach  $x^1$  im Zeitraum  $[t_0, t_1]$  mit der Steuerungsfunktion  $u(t)$ , dann können wir genauso von  $x^0$  nach  $x^1$  im Zeitraum  $[0, t_1 - t_0]$  gelangen, wenn wir die Steuerungsfunktion  $\tilde{u}(t) := u(t - t_0)$  und Lösungstrajektorie  $\tilde{x}(t) := x(t - t_0)$  betrachten.

Wir werden im Folgenden davon ausgehen, daß die Steuerungsfunktion  $u(t)$  in einem Funktionenraum  $U_{ad}$  zulässiger Steuerungen enthalten ist. Wir gehen davon aus, daß  $u$  keinen Beschränkungen wie z.B.  $a(t) \leq u(t) \leq b(t)$  (komponentenweise) unterliegt. Dies würde zu Fragestellungen der linearen oder nichtlinearen Optimierung führen und soll hier nicht weiter untersucht werden. Der Funktionenraum  $U_{ad}$  wird hier i.d.R. die Menge der vektorwertigen, quadratisch-integrierbaren Funktionen  $L_2([t_0, \infty]; \mathcal{U})$  oder der stückweise stetigen Funktionen  $PC([t_0, \infty]; \mathcal{U})$  sein. Integration wird im Lebesgueschen Sinn verstanden.

Um die Abhängigkeit der Lösungstrajektorie der Differentialgleichungen (1.1), (1.4) oder (1.6) von der Steuerungsfunktion  $u(t)$  zu verdeutlichen, schreiben wir

$$x(t) = x(t; u),$$

wobei wir annehmen, daß die Lösung der jeweiligen Anfangswertaufgabe auf dem Intervall  $[t_0, t_f]$  für alle  $u \in U_{ad}$  existiert und eindeutig ist.

Eine zentrale Frage in der Regelungs- und Steuerungstheorie lautet:

Sind ein Anfangszustand  $x^0$  und ein Ziel  $x^1$  gegeben, können wir ein  $\hat{u} \in U_{ad}$  finden, so daß für ein  $t_f \geq t_1 \geq t_0$  gilt:  $x(t_1; \hat{u}) = x^1$ ?

Eine Verschärfung lautet:

Sind ein Anfangszustand  $x^0$ , ein Ziel  $x^1$ , sowie  $t_1 \leq t_f$  gegeben, können wir ein  $\hat{u} \in U_{ad}$  finden, so daß  $x(t_1; \hat{u}) = x^1$  gilt?

Oft kann man das Problem so formulieren, daß das Ziel  $x^1 = 0$  ist, d.h.  $x$  beschreibt die Abweichung von einem nominalen Pfad (einer vorgegebenen Trajektorie). Eine schwächere Forderung lautet dann, eine *asymptotisch stabilisierende* Steuerung  $\hat{u} \in U_{ad}$  zu finden, d.h.  $\lim_{t \rightarrow \infty} x(t; \hat{u}) = 0$ . Leicht modifiziert lautet die Frage, ob man in endlicher Zeit in eine (beliebige) Umgebung der Null gelangen kann.

Neben der Existenz solcher Steuerungsfunktionen spielt auch die Frage der Optimalität oft eine wichtige oder entscheidende Rolle. Mögliche Zielfunktionale lauten bei gegebenem  $x_1 \in \mathcal{X}$  oder gegebener Referenztrajektorie  $x_{ref}(t)$  (z.B.  $x_{ref}(t) \equiv 0$ , falls die Zustandsgröße die Abweichung von einem Nominalzustand beschreibt)

$$\min_{u \in U_{ad}} \{ t_1 \in [t_0, t_f] \mid x(t_1; u) = x^1 \} \quad \text{zeitoptimale Steuerung,} \quad (1.10)$$

$$\min_{u \in U_{ad}} \int_{t_0}^{t_f} \|x(t) - x_{ref}(t)\| dt \quad \text{Steuerung minimaler Abweichung,} \quad (1.11)$$

$$\min_{u \in U_{ad}, x(t; u) = x^1} \int_{t_0}^{t_f} \|u(t)\| dt \quad \text{energieminimale Steuerung.} \quad (1.12)$$

Dabei ist  $\|\cdot\|$  eine geeignete Vektornorm, z.B. die (gewichtete) Euklidische Norm, aber auch die 1- und  $\infty$ -Normen können sinnvoll sein. Mischungen dieser Kostenfunktionale treten häufig auf, wir werden insbesondere Kombinationen von (1.11) und (1.12) betrachten, während (1.10) Gegenstand der *Optimalsteuerungstheorie* ist. Im Sinne der Systemtheorie bei Fokussierung auf Eingangs-/Ausgangsverhalten, d.h. auf  $u(t)$  und  $y(t)$  statt auf Zustände  $x(t)$  werden obige Kostenfunktionale auch häufig für  $y$  statt  $x$  formuliert, insbesondere bei der Verfolgung vorgegebener Trajektorien ist oft  $y_{ref}(t)$  statt  $x_{ref}(t)$  vorgegeben.

**Bemerkung 1.5** Oft findet man in der Literatur auch das Kostenfunktional

$$\min_{u \in U_{ad}} \int_{t_0}^{t_f} \|x(t) - x^1\| dt, \quad (1.13)$$

bzw. in Kombination mit (1.12)

$$\min_{u \in U_{ad}} \int_{t_0}^{t_f} \|x(t) - x^1\| + \|u(t)\| dt, \quad (1.14)$$

Dies führt aber oft zu physikalisch schwer zu realisierenden oder die Mechanik/Elektronik des Systems stark beanspruchenden Verläufen der Steuerungsfunktion, da die Steuerung dazu neigt, erst am Intervallende starken Einfluß auf die Systemdynamik zu nehmen. Dem kann zwar durch die Energieminimierung entgegen gewirkt werden, aber meist erhält man vernünftigeren Steuerungen, wenn man eine Referenztrajektorie  $x_{ref}(t)$  mit  $x_{ref}(t_f) = x^1$  angibt.

Wir werden uns hier meistens mit Steuerungsfunktionen  $u(t)$  beschäftigen, die als *Regelung* (engl. “feedback control”, “closed-loop control”) auftreten. Dabei wird die Kenntnis über den gegenwärtigen Zustand oder Ausgang verwendet, um das System in einen gewünschten Zustand zu bringen bzw. die Abweichung vom gewünschten Zustand zu korrigieren. Dabei unterscheidet man

- *Zustandsrückführung*:  $u(t) = u(t, x(t))$ , im linearen Fall  $u(t) = F(t)x(t)$  bzw. im zeitinvarianten Fall  $u(t) = Fx(t)$  mit  $F, F(t) \in \mathbb{R}^{m \times n}$ .

- *Ausgangsrückführung*:  $u(t) = u(t, y(t))$ , im linearen Fall  $u(t) = F(t)y(t)$  bzw. im zeitinvarianten Fall  $u(t) = Fy(t)$  mit  $F, F(t) \in \mathbb{R}^{m \times p}$ .

Die dabei auftretende Matrix  $F$  heißt *Verstärkungs- oder Korrekturmatri*x (engl. “feedback matrix/gain”) und ist geeignet zu wählen, um das vorgegebene Steuerungsziel (optimal) zu erreichen. D.h., wir werden uns in der Vorlesung i.W. der Frage widmen, ob solch eine Matrix  $F$  existiert und wie sie ggf. bestimmt werden kann. Im hier betrachteten linearen Fall führt das Einsetzen der Rückführung in (1.6) (bzw. analog für (1.4)) zur folgenden Darstellung des *geschlossenen Regelkreises*:

- Bei Zustandsrückführung:  $\dot{x}(t) = A(t)x(t) + B(t)u(t) = (A(t) + B(t)F(t))x(t)$ .
- Bei Ausgangsrückführung:  $\dot{x}(t) = A(t)x(t) + B(t)u(t) = A(t)x(t) + B(t)F(t)y(t)$ .

Die Ausgangsrückführung liefert einen *geschlossenen Regelkreis* wie in Abbildung ?? dargestellt.

Das folgende Beispiel illustriert die bisher definierten Begriffe und zeigt die Fragestellungen und Schwierigkeiten ihrer Lösung auf. Es wird uns auch in den weiteren Kapiteln weiter begleiten, um systemtheoretische Begriffe zu veranschaulichen.

**Beispiel 1.6** *Text fehlt noch!*

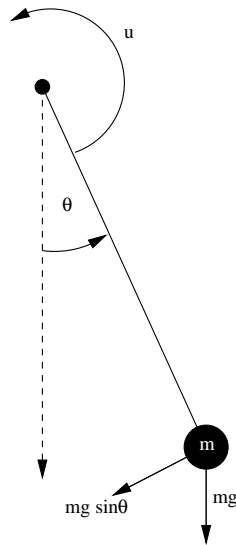


Abbildung 1.3: Pendel als mathematisches Modell eines eingelenkigen Rotationsarms.

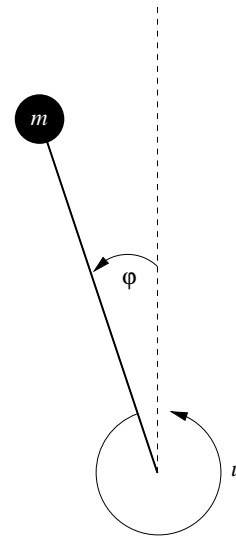


Abbildung 1.4: Inverses Pendel: Steuerung in aufrechte Position.

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \Theta(t) \\ \dot{\Theta}(t) \end{bmatrix},$$

$$\mathbf{f}(t, \mathbf{x}, \mathbf{u}) = \begin{bmatrix} x_2 \\ -mg \sin(x_1) \end{bmatrix}, \quad \mathbf{g}(t, \mathbf{x}, \mathbf{u}) = x_1,$$

**Definition 1.7** stabil

**Satz 1.8** stabiles LTI System

**Bemerkung 1.9** *offener/geschlossener Regelkreis*





# Kapitel 2

## Analyse von Regelungssystemen

### 2.1 Steuerbarkeit

Zunächst stellen wir die Frage, ob für einen gegebenen Anfangszustand  $x^0 \in \mathbb{R}^n$  ein gegebenes Ziel  $x^1 \in \mathbb{R}^n$  mit Hilfe einer Steuerungsfunktion  $u \in U_{ad}$  erreicht werden kann. Da Steuerbarkeit nur von der Zustandsgleichung (1.1) abhängt, werden wir die Ausgangsgleichung zunächst ignorieren.

**Definition 2.1 (Steuerbarkeit)** Sei  $x^1 \in \mathcal{X} \subset \mathbb{R}^n$ .

- a) Das Regelungssystem (1.1) mit Anfangsbedingung  $x(t_0) = x^0 \in \mathcal{X} \subset \mathbb{R}^n$  ist steuerbar nach  $x^1$  in der Zeit  $t_1 > t_0$ , falls ein  $u \in U_{ad}$  existiert, so daß  $x(t_1; u) = x^1$ . Das Paar  $(t_1, x^1)$  heißt dann erreichbar von  $(t_0, x^0)$ .
- b) Das Regelungssystem (1.1) mit Anfangsbedingung  $x(t_0) = x^0 \in \mathcal{X} \subset \mathbb{R}^n$  ist steuerbar nach  $x^1$  falls ein  $t_1 > t_0$  existiert, so daß  $(t_1, x^1)$  erreichbar von  $(t_0, x^0)$  ist.
- c) Ist  $(t_0, x^0)$  für alle  $x^0 \in \mathcal{X}$  nach  $x^1$  steuerbar für alle  $x^1 \in \mathcal{X}$ , so heißt das Regelungssystem (1.1) (vollständig) steuerbar.
- d) Die Steuerbarkeitsmenge bzgl.  $x^1$  ist definiert als

$$\mathcal{C}(t_0) := \bigcup_{t_1 > t_0} \mathcal{C}(t_0, t_1),$$

wobei  $\mathcal{C}(t_0, t_1) := \{x^0 \in \mathcal{X} \mid \exists u \in U_{ad} : x(t_1; u) = x^1\}$ .

Analog definiert man die Erreichbarkeitsmengen bzgl.  $x^0$ ,

$$\mathcal{R}(t_0, t_1) := \{x^1 \in \mathcal{X} \mid \exists u \in U_{ad} : x(t_1; u) = x^1\}$$

und  $\mathcal{R}(t_0) := \bigcup_{t_1 > t_0} \mathcal{R}(t_1)$ .

Die Steuerbarkeitsmenge enthält also alle Anfangszustände, die nach  $x^1$  gesteuert werden können, während die Erreichbarkeitsmenge alle Zustände enthält, die man von einem gegebenen  $x^0$  aus ansteuern (erreichen) kann.

Im Folgenden werden wir uns auf lineare Systeme beschränken. Wir werden sehen, daß für LTI Systeme alle Steuerbarkeitsbegriffe zusammenfallen und  $\mathcal{C} := \mathcal{C}(0) = \mathbb{R}^n$  äquivalent zur Steuerbarkeit ist. Dazu benötigen wir zunächst die Lösungen der Anfangswertaufgaben (1.4) und (1.6). Hierbei interessiert uns insbesondere die *Input-to-State*-Abbildung

$$\mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times U_{ad} \rightarrow \mathbb{R}^n : (t_0, x^0, t, u) \rightarrow x(t),$$

die durch folgendes bekannte Resultat aus der Theorie gewöhnlicher Differentialgleichungen gegeben ist.

**Satz 2.2** a) Sei  $\Phi$  die Fundamentallösung zu  $\dot{x}(t) = A(t)x(t)$ , d.h. Lösung der linearen homogenen Matrix-Differentialgleichung

$$\frac{\partial}{\partial t} \Phi(t, s) = A(t)\Phi(t, s), \quad \Phi(s, s) = I_n. \quad (2.1)$$

Dann gilt für die eindeutige Lösung von (1.6)

$$x(t) = \Phi(t, t_0)x^0 + \int_{t_0}^t \Phi(t, s)B(s)u(s)ds. \quad (2.2)$$

b) Die eindeutige Lösung von (1.4) erfüllt  $\Phi(t, s) = e^{A(t-s)}$  und daher

$$x(t) = e^{At}x^0 + \int_0^t e^{A(t-s)}Bu(s)ds = e^{At}\left(x^0 + \int_0^t e^{-As}Bu(s)ds\right). \quad (2.3)$$

**Beweis:** Durch Ausdifferenzieren und Verwendung der Leibniz-Formel in der vereinfachten Form,

$$\frac{d}{dt} \int_{t_0}^t F(t, s) dt = f(t, t) + \int_{t_0}^t \frac{\partial}{\partial t} F(t, s) ds. \quad \square$$

Als Konsequenz aus Satz 2.2 ergibt sich sofort eine Darstellung der *Input-to-Output*-Abbildung

$$\mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times U_{ad} \rightarrow \mathbb{R}^p : (t_0, x^0, t, u) \rightarrow y(t).$$

**Folgerung 2.3** a) Die eindeutige Lösung von (1.7) erfüllt

$$y(t) = C(t)\Phi(t, t_0)x^0 + C(t) \int_{t_0}^t \Phi(t, s)B(s)u(s)ds. \quad (2.4)$$

b) Die eindeutige Lösung von (1.5) ist gegeben durch

$$y(t) = Ce^{At}x^0 + \int_0^t Ce^{A(t-s)}Bu(s)ds = e^{At}\left(x^0 + \int_0^t e^{-As}Bu(s)ds\right). \quad (2.5)$$

Im Folgenden werden wir verwenden, daß die Fundamentallösung  $\Phi$  die *Halbgruppeneigenschaft*

$$\Phi(t, t) = I_n, \quad \Phi(t, s) = \Phi(t, \tau)\Phi(\tau, s) \quad (2.6)$$

für alle  $t, s, \tau \in \mathbb{R}$  besitzt. Außerdem ist  $\Phi$  für alle  $s, t \in \mathbb{R}$  invertierbar und es gilt:

$$\Phi(t, s)^{-1} = \Phi(s, t). \quad (2.7)$$

Zur Vereinfachung der Notation gehen wir im Folgenden von  $U_{ad} = PC([t_0, \infty); \mathbb{R}^m)$  aus. Die Aussagen gelten analog für  $U_{ad} = L_2([t_0, \infty); \mathbb{R}^m)$ , statt "für alle" muß dann meist nur "f.ü." geschrieben werden.

Zunächst betrachten wir das Ziel  $x^1 = 0$  und die entsprechenden Steuerbarkeitsmengen bzgl.  $x^1 = 0$ . Dies wird sich bei linearen Systemen nicht als einschränkend erweisen.

**Lemma 2.4**  $x^0 \in \mathcal{C}(t_0, t_1)$  genau dann, wenn ein  $u \in U_{ad}$  existiert mit

$$x^0 = - \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s)ds.$$

**Beweis:** Nach Satz 2.2 ist  $x^0 \in \mathcal{C}(t_0, t_1)$  äquivalent zu

$$\begin{aligned} 0 = x(t_1) &= \Phi(t_1, t_0)x^0 + \int_{t_0}^{t_1} \Phi(t_1, s)B(s)u(s)ds \\ &= \Phi(t_1, t_0) \left( x^0 + \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s)ds \right), \end{aligned}$$

wobei wir die Halbgruppeneigenschaft von  $\Phi$  verwendet haben. Mit der Invertierbarkeit von  $\Phi$  folgt damit

$$0 = x^0 + \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s)ds$$

für ein  $u \in U_{ad}$  und damit die Behauptung.  $\square$

Der Begriff der Gramschen Matrix wird im weiteren Verlauf eine herausragende Rolle spielen. Diese Matrix wird zunächst wie folgt definiert.

**Definition 2.5** Für  $G(t) \in PC((-\infty, \infty); \mathbb{R}^{n \times m})$  heißt

$$P(t_0, t_1) = \int_{t_0}^{t_1} G(t)G(t)^T dt$$

die  $(t_0, t_1)$ -Gramsche (Matrix) zu  $G$ .

Offensichtlich ist die Gramsche positiv semidefinit. Weitere Eigenschaften, die wir im Folgenden benötigen werden, enthalten die folgenden Lemmata.

**Lemma 2.6**

$$\ker P(t_0, t_1) = \{x \in \mathbb{R}^n \mid G(t)^T x \equiv 0 \text{ auf } [t_0, t_1]\}.$$

**Beweis:** Für beliebiges  $x \in \mathbb{R}^n$  gilt

$$x^T P(t_0, t_1)x = x^T \int_{t_0}^{t_1} G(t)G(t)^T dt x = \int_{t_0}^{t_1} \underbrace{(G(t)^T x)^T (G(t)^T x)}_{\geq 0 \forall t} dt \geq 0$$

Damit ist  $P(t_0, t_1)x = 0$  genau dann, wenn  $G(t)^T x \equiv 0$  auf  $[t_0, t_1]$ . □

**Lemma 2.7** Sei  $G$  wie in Definition 2.5. Dann sind folgende Aussagen äquivalent:

a) Es existiert  $u \in U_{ad}$ , so daß  $x = \int_{t_0}^{t_1} G(t)u(t)dt$ .

b)  $x \in \text{range}(P(t_0, t_1))$ , d.h. es existiert  $z \in \mathbb{R}^n$  mit  $x = P(t_0, t_1)z$ .

**Beweis:** Zunächst definiere

$$\mathcal{L} := \{x \in \mathbb{R}^n \mid \exists u \in U_{ad} \text{ with } x = \int_{t_0}^{t_1} G(t)u(t)dt\}.$$

$\mathcal{L}$  ist aufgrund der Linearität des Integrals und der Vektorraumeigenschaften von  $U_{ad}$  ein Unterraum des  $\mathbb{R}^n$ , also insbesondere selbst ein Vektorraum.

Zu zeigen ist also  $\mathcal{L} = \text{range}(P(t_0, t_1))$ .

Sofort klar ist  $\text{range}(P(t_0, t_1)) \subset \mathcal{L}$ . (Dazu setze  $u(t) = G(t)^T z$  für  $x = P(t_0, t_1)z$ .)

Sei nun  $x \in \mathcal{L} \cap \ker P(t_0, t_1)$ . Dann gilt wegen  $x \in \mathcal{L}$  und Lemma 2.6

$$x^T x = \int_{t_0}^t \underbrace{x^T G(t)}_{\substack{=0, \text{ da} \\ x \in \ker P(t_0, t_1)}} u(t)dt = 0$$

woraus sofort  $x = 0$  folgt. Also erhält man  $\dim \mathcal{L} \cap \ker P(t_0, t_1) = \{0\}$  und daher mit Hilfe der Dimensionsformel

$$\begin{aligned} n &\geq \dim(\mathcal{L} + \ker P(t_0, t_1)) = \dim(\mathcal{L}) + \dim(\ker P(t_0, t_1)) \\ &\geq \dim(\text{range}(P(t_0, t_1))) + \dim(\ker P(t_0, t_1)) = n. \end{aligned}$$

Damit ergibt sich sofort  $\dim \mathcal{L} = \dim \text{range}(P(t_0, t_1))$ , also insgesamt  $\mathcal{L} = \text{range}(P(t_0, t_1))$ .  $\square$

Setzt man nun  $G(t) = \Phi(t_0, t)B(t)$ , dann nennt man

$$P(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_0, t)B(t)B(t)^T\Phi(t_0, t)^T dt \quad (2.8)$$

die  $(t_0, t_1)$ -Steuerbarkeits-Gramsche des linearen Systems (1.6)–(1.7).

Damit erhält man eine erste Charakterisierung der Steuerbarkeitsmenge.

**Satz 2.8** Sei  $x^1 = 0$  und betrachte das LTV System (1.6)–(1.7) mit  $P(t_0, t_1)$  wie in (2.8). Dann gilt:

- a)  $\mathcal{C}(t_0, t_1) = \text{range}(P(t_0, t_1))$ .
- b)  $P(t_0, t_1)x = 0 \iff x^T\Phi(t_0, t)B(t) \equiv 0$  auf  $[t_0, t_1]$ .

**Beweis:**

a) Lemma 2.4 und Lemma 2.7.

b) Lemma 2.6.  $\square$

Eine im weiteren sehr nützliche Charakterisierung der vollständigen Steuerbarkeit von LTV Systemen erhält man nun über die genauere Untersuchung der Eigenschaften der  $(t_0, t_1)$ -Steuerbarkeits-Gramschen  $P(t_0, t_1)$ . Dazu erinnern wir uns zunächst aus der Theorie für lineare Differentialgleichungen an die folgende Eigenschaft der zu  $\dot{x} = Ax$  adjungierten Gleichung

$$\dot{z}(t) = -A(t)^T z(t). \quad (2.9)$$

Ist nämlich  $\Phi(t, s)$  die Fundamentallösung zu  $\dot{x}(t) = A(t)x(t)$ , d.h. Lösung der linearen homogenen Matrix-Differentialgleichung (2.1), dann ist

$$\Phi(t, s)^{-T} = \Phi(s, t)^T$$

Fundamentallösung von (2.9)<sup>1</sup> und insbesondere kann jede Lösung des Anfangswertproblems zu (2.9) mit  $z(t_0) = z_0$  geschrieben werden als

$$z(t) = \Phi(t_0, t)^T z_0. \quad (2.10)$$

**Satz 2.9** Die folgenden Aussagen sind äquivalent:

---

<sup>1</sup>**Beweis:** Sei  $\Psi$  Fundamentallösung von (2.9), d.h.  $\frac{\partial}{\partial t}\Psi(t, s) = -A(t)^T\Psi(t, s)$ ,  $\Psi(s, s) = I_n$ . Dann gilt  $(\frac{\partial}{\partial t}\Psi(t, s)^T)\Phi(t, s) = -\Psi(t, s)^T A(t)\Phi(t, s) = -\Psi(t, s)^T(\frac{\partial}{\partial t}\Phi(t, s))$ , also  $0 = (\frac{\partial}{\partial t}\Psi(t, s)^T)\Phi(t, s) + \Psi(t, s)^T(\frac{\partial}{\partial t}\Phi(t, s)) = \frac{\partial}{\partial t}\Psi(t, s)^T\Phi(t, s)$ . Also ist  $\Psi(t, s)^T\Phi(t, s)$  konstant und wegen der Anfangsbedingung gilt  $\Psi(t, s)^T\Phi(t, s) = I_n$ .

a) Das LTV System (1.6) ist vollständig steuerbar.

b) Jede Lösung der adjungierten Gleichung (2.9) hat die Eigenschaft

$$z(t)^T B(t) \equiv 0 \text{ auf } [t_0, \infty) \text{ für ein } t_0 \in \mathbb{R} \implies z(t) \equiv 0. \quad (2.11)$$

c) Für alle  $t_0 \in \mathbb{R}$  existiert ein  $t_1 \in \mathbb{R}$ , so daß  $P(t_0, t_1)$  positiv definit ist.

**Beweis:** Der Beweis erfolgt durch Ringschluß.

**a)  $\Rightarrow$  b):** Angenommen, es existiert eine nichttriviale Lösung von (2.9) mit  $z(t)^T B(t) \equiv 0$  auf  $[t_0, \infty)$  für ein  $t_0 \in \mathbb{R}$ , aber  $z(\hat{t}) \neq 0$  für mindestens ein  $\hat{t} \in \mathbb{R}$ . Dann ist  $z(t_0) \neq 0$ , da mit (2.10)  $z(\hat{t}) = \Phi(t_0, \hat{t})^T z(t_0)$  gilt und  $\Phi(t_0, t)$  invertierbar ist.

Nun wählen wir  $x^0 \in \mathbb{R}^n$  so, daß  $(x^0)^T z(t_0) \neq 0$  gilt. Da (1.6) vollständig steuerbar ist, existieren nach Definition 2.1 ein  $t_1 > t_0$  und  $u \in U_{ad}$  so daß  $x(t_1) \equiv x(t_1; u) = 0$  Lösung von (1.6) mit Anfangsbedingung  $x(t_0) = x^0$  ist. Damit folgt:

$$\begin{aligned} \frac{d}{dt} (x(t)^T z(t)) &= \dot{x}(t)^T z(t) + x(t)^T \dot{z}(t) \\ &= x(t)^T A(t)^T z(t) + u(t)^T \underbrace{B(t)^T z(t)}_{\equiv 0 \text{ auf } [t_0, \infty)} - x(t)^T A(t)^T z(t) \\ &= 0. \end{aligned}$$

Also ist  $x(t)^T z(t)$  konstant und wegen der Anfangsbedingungen gilt

$$x(t_1)^T z(t_1) = x(t_0)^T z(t_0) = (x^0)^T z(t) \neq 0,$$

was im Widerspruch zu  $x(t_1) = 0$  steht.

**b)  $\Rightarrow$  c):** Dieser Schritt erfolgt in zwei Teilen. Zunächst zeigen wir folgende Aussage:

Für alle  $t_0 \in \mathbb{R}$  existiert ein  $t_1 \in \mathbb{R}$ , so daß jede nichttriviale Lösung der adjungierten Gleichung (2.9) die Eigenschaft

$$z(t)^T B(t) \neq 0 \text{ auf } [t_0, t_1] \quad (2.12)$$

hat.

Angenommen, dies wäre nicht der Fall. Dies würde bedeuten, daß es eine Folge  $(t_k)_{k=1}^{\infty}$  mit  $t_k \rightarrow \infty$  für  $k \rightarrow \infty$  und eine Folge von Lösungen  $z_k(t)$  von (2.9) mit Anfangswerten  $\|z_k(t_0)\| = 1$  gibt, so daß

$$z_k(t)^T B(t) \equiv 0 \text{ auf } [t_0, t_k]. \quad (2.13)$$

O.B.d.A. nehmen wir an, daß  $(z_k(t_0))_{k=1}^{\infty}$  konvergent ist — sonst können wir eine konvergente Teilfolge finden, da  $\{z \in \mathbb{R}^n \mid \|z\| = 1\}$  kompakt ist.

Sei nun  $\hat{z}_0 := \lim_{k \rightarrow \infty} z_k(t_0)$  und  $\hat{z}$  Lösung von (2.9) mit Anfangswert  $\hat{z}(t_0) = \hat{z}_0$ . Dann ist  $\hat{z}(t) \neq 0$  (da  $\|\hat{z}(t_0)\| = 1$ ) und wegen (2.11) gilt  $\hat{z}(t)^T B(t) \neq 0$  auf  $[t_0, \infty)$ . Also existiert  $\hat{t} > t_0$  mit  $\hat{z}(\hat{t})^T B(\hat{t}) \neq 0$ .

Da die Lösung von (2.9) stetig vom Anfangswert abhängt, konvergiert  $(z_k(t))_{k=1}^\infty$  gleichmäßig gegen  $\hat{z}(t)$ . Dann gilt aber  $z_k(\hat{t})^T B(\hat{t}) \neq 0$  für hinreichend grosses  $k$ , was mit  $t_k \rightarrow \infty$  zum Widerspruch zu (2.13) führt.

Im zweiten Schritt zeigen wir:

Wenn jede nichttriviale Lösung von (2.9) die Eigenschaft (2.12) besitzt, dann ist  $P(t_0, t_1) > 0$ .

Da wir bereits  $P(t_0, t_1) \geq 0$  haben, reicht zu zeigen, daß  $\ker P(t_0, t_1) = \{0\}$ . Nach Satz 2.8 ist  $z_0 \in \ker P(t_0, t_1)$  äquivalent zu

$$z_0^T \Phi(t_0, t) B(t) \equiv 0 \text{ auf } [t_0, t_1].$$

Es ist wegen (2.10)  $z(t) = \Phi(t_0, t)^T z_0$  Lösung von (2.9) mit  $z(t_0) = z_0$ . Also gilt  $z(t)^T B(t) \equiv 0$  auf  $[t_0, t_1]$  und aufgrund des ersten Schrittes damit  $z(t) \equiv 0$  auf  $[t_0, t_1]$ . Dies impliziert  $z_0 = 0$ , womit  $\ker P(t_0, t_1) = \{0\}$  gezeigt ist.

**c)  $\Rightarrow$  a):** Sei  $t_1$  so gewählt, daß  $P(t_0, t_1) > 0$ . Dann kann jedes Paar  $(t_0, x^0)$  nach  $x^1 \in \mathbb{R}^n$  (beliebig) in der Zeit  $t_1$  gesteuert werden mit

$$u(t) := B(t)^T \Phi(t_0, t)^T c,$$

wobei  $c \in \mathbb{R}^n$  als Lösung der Gleichung

$$\begin{aligned} x^0 = x(t_0) &= \Phi(t_0, t_1) x^1 + \int_{t_1}^{t_0} \Phi(t_0, s) B(s) B(s)^T \Phi(t_0, s)^T c \, ds \\ &= \Phi(t_0, t_1) x^1 - P(t_0, t_1) c \end{aligned}$$

bestimmt werden kann. Diese Gleichung folgt aus (2.2) durch Vertauschen von Anfangs- und Endzeitpunkt und besitzt eine eindeutige Lösung wegen der positiven Definitheit von  $P(t_0, t_1)$ . □

Für LTI Systeme, wo wir o.B.d.A.  $t_0 = 0$  annehmen können, kann man noch mehr aussagen. Zunächst folgt aus der expliziten Formel für die Fundamentallösung zu  $\dot{x} = Ax$ , daß sich die  $(0, t_1)$ -Steuerbarkeits-Gram'sche schreiben läßt als

$$P(0, t_1) = \int_0^{t_1} e^{-At} B B^T e^{-A^T t} dt.$$

Daraus folgt sofort, daß  $x \in \ker P(0, t_1)$  genau dann, wenn

$$B^T e^{-A^T t} x \equiv 0 \quad \text{in } [0, t_1]. \quad (2.14)$$

Im weiteren werden wir Steuerbarkeit über Eigenschaften der folgenden Matrix charakterisieren.

**Definition 2.10 (Steuerbarkeitsmatrix)** Die Steuerbarkeitsmatrix eines LTI Systems ist

$$\mathcal{K}(A, B) := [B, AB, A^2B, \dots, A^{n-1}B] \in \mathbb{R}^{n \times nm}.$$

Damit erhalten wir eine Charakterisierung der Steuerbarkeitsmengen  $\mathcal{C}(0, t)$  für LTI Systeme.

**Satz 2.11** Für LTI Systeme gilt:  $\mathcal{C}(0, t) = \text{range}(\mathcal{K}(A, B))$  für alle  $t > 0$ .

**Beweis:** Wir zeigen die Aussage des Beweises indirekt, indem wir  $\mathcal{C}(0, t)^\perp = \text{range}(\mathcal{K}(A, B))^\perp$  für alle  $t > 0$  beweisen. Aus Satz 2.8a) folgt mit  $P(0, t) = P(0, t)^T \geq 0$ , daß

$$\mathcal{C}(0, t)^\perp = \text{range}(P(0, t))^\perp = \ker P(0, t).$$

Damit bleibt also zu zeigen, daß

$$\ker P(0, t) = \text{range}(\mathcal{K}(A, B))^\perp = \ker \mathcal{K}(A, B)^T, \quad (2.15)$$

oder, mit anderen Worten,  $P(0, t)x = 0$  genau dann, wenn  $x^T \mathcal{K}(A, B) = 0$ .

Aus Satz 2.8b) bzw. (2.14) kennen wir bereits eine Eigenschaft der Elemente des Kerns von  $P(0, t)$ . Diese wollen wir nun ausnutzen. Dazu zunächst einige Vorüberlegungen. Es sei  $\phi_A(x) = \sum_{j=0}^n \alpha_j x^j$  das charakteristische Polynom von  $A$ . Der Satz von Cayley-Hamilton aus der linearen Algebra besagt, daß  $\phi_A(A) = 0$ . Daraus folgt wegen  $\alpha_n = 1$  mit  $\beta_j = -\alpha_j$

$$A^n = \sum_{j=0}^{n-1} \beta_j A^j. \quad (2.16)$$

Also gilt  $x^T A^n B = \sum_{j=0}^{n-1} \beta_j x^T A^j B$ . Durch wiederholte Anwendung von (2.16) sowie Zusammenfassung aller Koeffizienten von  $x^T A^j B$  zu  $\beta_j^{(\nu)}$  erhält man die Darstellung

$$x^T A^{n+\nu} B = \sum_{j=0}^{n-1} \beta_j^{(\nu)} x^T A^j B \quad \forall \nu \in \mathbb{N}_0. \quad (2.17)$$

Daraus ergibt sich die folgende Kette von Äquivalenzen:

$$\begin{aligned} x^T \mathcal{K}(A, B) = 0 &\iff x^T A^j B = 0, \quad j = 0, 1, \dots, n-1 \\ &\stackrel{(2.17)}{\iff} x^T A^\nu B = 0 \quad \forall \nu \in \mathbb{N}_0 \\ &\iff 0 = \sum_{j=0}^{\infty} \frac{(-\tau)^j}{j!} x^T A^j B = x^T e^{-A\tau} B \quad \forall \tau \in [0, t] \\ &\stackrel{(2.14)}{\iff} P(0, t)x = 0. \end{aligned}$$



Damit folgt (2.15) und somit die Aussage des Satzes.  $\square$

Theorem 2.11 zeigt, daß für ein steuerbares LTI System gilt:  $\mathcal{C}(0, t_1) = \mathcal{C}(0, t_2)$  für alle  $t_1, t_2 > 0$ . D.h. insbesondere, daß  $\mathcal{C}(t) \equiv \mathcal{C}(0, t)$  für alle  $t > 0$  und damit alle Steuerbarkeitskonzepte für LTI Systeme zusammenfallen. Abkürzend schreiben wir daher im Folgenden einfach  $\mathcal{C}$  für  $\mathcal{C}(t)$  oder  $\mathcal{C}(0, t)$ . Weiter folgt zusammen mit Satz 2.8, daß  $\text{range}(\mathcal{K}(A, B)) \equiv \text{range}(P(0, t))$ . Für ein steuerbares LTI System gilt also offenbar nach Satz 2.9, daß  $P(0, t) > 0$  für alle  $t > 0$ .

Eine für die Praxis äußerst nützlich Charakterisierung der Steuerbarkeit liefert der sogenannte Hautus-Test.

**Satz 2.12 (Hautus-Popov-Lemma)**

Sei  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ . Die folgenden Aussagen sind äquivalent:

- a)  $(A, B)$  is steuerbar.
- b)  $\text{rank}(\mathcal{K}(A, B)) = n$ .
- c) Falls  $z$  Linkseigenvektor von  $A$  ist, so muß  $z^*B \neq 0$  gelten.
- d)  $\text{rank} \left( \begin{bmatrix} A - \lambda I & B \end{bmatrix} \right) = n$  für alle  $\lambda \in \mathbb{C}$ .

**Beweis:**

a)  $\Leftrightarrow$  b) Dies folgt sofort aus Satz 2.11 und der Äquivalenz von Steuerbarkeit und  $\mathcal{C} = \mathbb{R}^n$ .

c)  $\Leftrightarrow$  d)  $z^* \begin{bmatrix} A - \lambda I & B \end{bmatrix} = 0$  gilt genau dann, wenn  $z^*A = \lambda z^*$  und  $z^*B = 0$ . Also ist  $\text{rank} \left( \begin{bmatrix} A - \lambda I & B \end{bmatrix} \right) < n$  genau dann, wenn ein Links-Eigenvektor  $z$  von  $A$  existiert, der  $z^*B = 0$  erfüllt.

b)  $\Rightarrow$  d) Angenommen, es gelte  $\text{rank} \left( \begin{bmatrix} A - \lambda I & B \end{bmatrix} \right) < n$ . Dann existiert ein  $z \neq 0$  mit  $z^* \begin{bmatrix} A - \lambda I & B \end{bmatrix} = 0$ , d.h.  $z^*A = \lambda z^*$ ,  $z^*B = 0$ . Nun gilt

$$z^*A^jB = \lambda^j z^*B = 0 \quad \text{für alle } j \in \mathbb{N}_0.$$

Daraus folgt  $z^*\mathcal{K}(A, B) = 0$  im Widerspruch zu b).

d)  $\Rightarrow$  b) Angenommen, es gelte  $\text{rank}(\mathcal{K}(A, B)) = r < n$ . Dann existiert eine Orthonormalbasis  $\{v_1, \dots, v_r\}$  von  $\mathcal{K} := \text{range}(\mathcal{K}(A, B))$ . Diese Basis werde zu einer Orthonormalbasis des  $\mathbb{R}^n$  ergänzt durch  $\{v_{r+1}, \dots, v_n\}$ . Dabei gilt

$$\text{range}(\mathcal{K}(A, B))^\perp = \text{span} \{v_{r+1}, \dots, v_n\}.$$

Definiere  $V := [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$ . Da die Spalten von  $V$  orthonormal sind, gilt  $VV^T = I_n = V^TV$ . Außerdem gilt  $v_{r+j}^T \mathcal{K}(A, B) = 0$ ,  $j = 1, \dots, n - r$  und damit insbesondere  $v_{r+j}^T B = 0$ ,  $j = 1, \dots, n - r$ , d.h.  $V^TB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$ .

Mit dem Satz von Cayley-Hamilton (siehe Beweis von Satz 2.11 bzw. (2.16)) folgt, daß  $\mathcal{AK}(A, B) \subset \mathcal{K}(A, B)$ , d.h.  $\mathcal{K}$  ist ein  $A$ -invarianter Unterraum des  $\mathbb{R}^n$ . Da die Spalten von  $V_1 := [v_1, \dots, v_r]$  eine Basis für diesen  $A$ -invarianten Unterraum bilden, existiert  $A_1 \in \mathbb{R}^{r \times r}$  mit  $\Lambda(A_1) \subset \Lambda(A)$  und  $AV_1 = V_1A_1$ . Damit folgt

$$AV = V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}.$$

Sei nun  $\tilde{z} \neq 0$  Links-Eigenvektor von  $A_3$ , d.h.  $\tilde{z}^*A_3 = \lambda\tilde{z}^*$  für ein  $\lambda \in \Lambda(A_3) \subset \Lambda(A)$ . Definiert man nun  $z := V \cdot \begin{bmatrix} 0 \\ \tilde{z} \end{bmatrix}$ , so ist  $z \neq 0$  (da  $V$  orthogonal und  $\tilde{z} \neq 0$ ) und erfüllt

$$\begin{aligned} z^*B &= [0 \quad \tilde{z}^*] V^T B = [0 \quad \tilde{z}^*] \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = 0, \\ z^*A &= [0 \quad \tilde{z}^*] V^T A = [0 \quad \tilde{z}^*] V^T AV V^T = [0 \quad \tilde{z}^*] \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T \\ &= \lambda [0 \quad \tilde{z}^*] V^T = \lambda z^*. \end{aligned}$$

Damit gilt  $z^* [A - \lambda I \quad B] = 0$  im Widerspruch zu d). □

Teil c) des Satzes liefert einen praktikablen Test für die Steuerbarkeit eines LTI Systems: Berechne alle Eigenwerte und Links-Eigenvektoren  $z_j$  von  $A$  und teste ob  $z_j^*B = 0$ . Für numerische Berechnungen ist dieser Test allerdings weniger tauglich als die sogenannte Staircase-Form (siehe Satz 2.27 bzw. Satz 2.29), da die Genauigkeit der berechneten Eigenvektoren sehr sensitiv gegenüber Rundungsfehlern sein kann, insbesondere falls  $A$  (fast) defektiv ist. Außerdem ist die Entscheidung  $z_j^*B = 0$  numerisch sehr schwierig.

Die im Beweis von Satz 2.12 verwendete Zerlegung von  $(A, B)$ ,

$$A = V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T, \quad B = V \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad (2.18)$$

mit  $V$  orthogonal und  $(A_1, B_1)$  steuerbar heißt (*orthogonale*) *Kalman-Zerlegung* von  $(A, B)$ . Für ein LTI System erhält man durch den Basiswechsel  $\tilde{x} := V^T x$  im Zustandsraum das folgende äquivalente System:

$$\begin{aligned} \dot{\tilde{x}}_1 &= A_1 \tilde{x}_1 + A_2 \tilde{x}_2 + B_1 u, \\ \dot{\tilde{x}}_2 &= A_3 \tilde{x}_2. \end{aligned}$$

Die Komponenten von  $\tilde{x}_2$  sind also durch den Anfangswert  $\tilde{x}_2(0)$  festgelegt und nicht durch eine Steuerung beeinflussbar; es gilt  $\tilde{x}_2(t) = e^{A_3 t} \tilde{x}_2(0)$ . Daher nennt man die Komponenten von  $\tilde{x}_2$  auch die *unsteuerbaren Zustände*, die Eigenwerte von  $A_3$  heißen die *unsteuerbaren Moden* des LTI Systems.

**Beispiel 1.6, fortgesetzt.**

Nach Linearisierung und Reduktion auf ein System erster Ordnung erhält man ein LTI System mit Zustandsraum  $\mathcal{X} = \mathbb{R}^2$  und

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Damit ist  $\mathcal{K}(A, B) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ , also  $\text{rank}(\mathcal{K}(A, B)) = 2$  und damit ist das System nach Satz 2.12a),b) steuerbar.

Alternativ kann man den Hautus-Test anwenden. Es ist  $\Lambda(A) = \{\pm 1\}$ , der Links-Eigenvektor zu  $\lambda_1 = 1$  ist  $z_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  und  $z_1^* B = 1 \neq 0$ , der Links-Eigenvektor zu  $\lambda_2 = -1$  ist  $z_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$  und  $z_2^* B = -1 \neq 0$ , womit ebenfalls die Steuerbarkeit des Systems gezeigt ist.

## 2.2 Stabilisierbarkeit

Nun wollen wir uns der bereits in der Einleitung angesprochenen und gegenüber der Steuerbarkeit etwas abgeschwächten Forderung nach der asymptotischen Erreichung eines Ziels zuwenden. Als Ziel nehmen wir wieder  $x^1 = 0$  als anzusteuernde Gleichgewichtslage unseres Systems an.

Zunächst betrachten wir LTV Systeme.

**Definition 2.13** Das LTV System (1.6)–(1.7) heißt (asymptotisch) stabilisierbar, falls zu jedem Anfangszustand  $x^0 \in \mathbb{R}^n$  ein  $u \in U_{ad}$  existiert, so daß die Lösung von (1.6) mit dieser Steuerung  $u$

$$\lim_{t \rightarrow \infty} x(t; u) = 0$$

erfüllt.

Eine notwendige Bedingung für die Stabilisierbarkeit von LTV Systemen liefert folgendes Ergebnis.

**Satz 2.14** Falls das LTV System (1.6) stabilisierbar und  $z$  eine für  $t \rightarrow \infty$  beschränkte, nichttriviale Lösung der adjungierten Gleichung (2.9) ist, so gilt

$$z(t)^T B(t) \neq 0 \text{ auf } [t_0, \infty).$$

**Beweis:** Angenommen, es existiert eine Lösung der adjungierten Gleichung (2.9), so daß

$$z(t)^T B(t) = 0 \quad \forall t \in [t_0, \infty) \quad \text{und} \quad \lim_{t \rightarrow \infty} \|z(t)\| < \infty.$$

Da  $z$  nach Voraussetzung nicht trivial ist, gilt  $z(t_0) \neq 0$ , so daß wir einen Anfangszustand  $x^0 \in \mathbb{R}^n$  finden können mit

$$(x^0)^T z(t_0) \neq 0.$$

Analog zum Beweis von Satz 2.11, a)  $\Rightarrow$  b), folgt

$$x(t)^T z(t) \equiv (x^0)^T z(t_0) \neq 0 \text{ auf } [t_0, \infty). \quad (2.19)$$

Sei nun  $u \in U_{ad}$  eine für die Lösung von (1.6) mit  $x(t_0) = x^0$  stabilisierende Steuerung. Dann gilt mit  $\lim_{t \rightarrow \infty} x(t; u) = 0$  auch

$$\lim_{t \rightarrow \infty} \|x(t; u)\| = 0.$$

Da  $\|z(t)\|$  für  $t \rightarrow \infty$  beschränkt ist, gibt es eine Folge  $\{t_j\}_{j=1}^{\infty}$  mit  $t_j \rightarrow \infty$  und

$$\lim_{j \rightarrow \infty} x(t_j)^T z(t_j) = 0.$$

(Beachte: Wegen der Cauchy-Schwarz-Ungleichung gilt:  $|x(t_j)^T z(t_j)| \leq \|x(t_j)\| \|z(t_j)\|$ .)  
Damit haben wir einen Widerspruch zu (2.19) konstruiert.  $\square$

Für die Stabilisierbarkeit eines LTI Systems bzw. eines Matrix-Paars  $(A, B)$  existieren ähnliche Charakterisierungen wie in Satz 2.12.

**Satz 2.15 (*Hautus-Test für Stabilisierbarkeit*)**

Sei  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ . Die folgenden Aussagen sind äquivalent:

- a)  $(A, B)$  ist stabilisierbar.
- b) Es existiert  $F \in \mathbb{R}^{m \times n}$  mit  $\Lambda(A + BF) \subset \mathbb{C}^-$ .
- c) In der Kalman-Zerlegung (2.18) gilt  $\Lambda(A_3) \subset \mathbb{C}^-$ .
- d) Falls  $z \neq 0$  ein Links-Eigenvektor von  $A$  zu einem Eigenwert  $\lambda$  mit  $\operatorname{Re}(\lambda) \geq 0$  ist, so muß  $z^* B \neq 0$  gelten.
- e)  $\operatorname{rank} \left( \begin{bmatrix} A - \lambda I & B \end{bmatrix} \right) = n$  für alle  $\lambda \in \mathbb{C}$  mit  $\operatorname{Re}(\lambda) \geq 0$ .

**Beweis:** Hausaufgabe 4, H12.

Ringschluß a)  $\Rightarrow$  d)  $\Rightarrow$  c)  $\Rightarrow$  e)  $\Rightarrow$  b)  $\Rightarrow$  a), ähnlich zum Beweis von Satz 2.12.

Folgende andere Implikationen können gezeigt werden:

c)  $\Leftrightarrow$  e)  $\Leftrightarrow$  d),

a)  $\Rightarrow$  d)  $\Rightarrow$  e)  $\Rightarrow$  b)  $\Rightarrow$  a).

$\square$

**Beispiel 1.6, fortgesetzt.**

Hier bietet es sich an, den Hautus-Test anzuwenden. Für die Systemmatrizen

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

ergibt sich — wie bereits oben festgestellt —  $\Lambda(A) = \{\pm 1\}$ , wobei  $z_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  ein Links-Eigenvektor zum einzigen Eigenwert mit nichtnegativem Realteil,  $\lambda_1 = 1$ , ist. Es gilt  $z_1^T B = 1 \neq 0$ , woraus die Stabilisierbarkeit mit Satz 2.15 folgt. Dabei ist zu beachten, daß es ausreicht,  $z_1^T B$  zu testen, da der Eigenraum zu  $\lambda_1$  eindimensional und damit jeder Eigenvektor zu  $\lambda_1$  ein skalares Vielfaches von  $z_1$  ist.

Einfacher hätte man es sich allerdings mit folgender einfacher Konsequenz aus den Sätzen 2.12 und 2.15 machen können.

**Folgerung 2.16** *Ein steuerbares LTI System ist stabilisierbar.*

## 2.3 Beobachtbarkeit und Entdeckbarkeit

Wir betrachten zunächst wieder LTV Systeme der Form (1.6)–(1.7) und stellen uns die Frage, wieviel Information über den Zustand des Systems aus der Beobachtungsgleichung (1.7) gezogen werden kann. Dies ist eine in der Praxis sehr relevante Fragestellung, da man meist nicht den gesamten Zustand für den Entwurf einer Regelung oder Steuerung zur Verfügung hat, sondern nur beobachtete bzw. gemessene Größen. Dies können einige der Zustandsvariablen sein oder aber davon abgeleitete Größen. Im Beispiel 1.6 konnten wir z.B. nur den Ort (erste Komponente des Zustandsvektors) messen, nicht aber die Winkelgeschwindigkeit (zweite Komponente des Zustandsvektors)!

**Definition 2.17 (Beobachtbarkeit)** *Ein LTV System heißt rekonstruierbar (beobachtbar), falls folgende Bedingung erfüllt ist:*

*Sind  $x, \tilde{x}$  Lösungen von (1.6) für dieselbe Steuerungsfunktion  $u \in U_{ad}$  und gilt*

$$C(t)x(t) = C(t)\tilde{x}(t) \quad \forall t \leq t_0 \quad (t \geq t_0),$$

*dann folgt*

$$x(t) = \tilde{x}(t) \quad \forall t \leq t_0 \quad (t \geq t_0).$$

Rekonstruierbarkeit bedeutet also, daß Systeme mit gleichen Eingängen und gleichen Ausgängen in der Vergangenheit auch gleiche Zustände in der Vergangenheit hatten. Beobachtbarkeit liefert dieselbe Aussage für die Zukunft, wobei als Bezugszeitpunkt immer  $t_0$  zugrunde liegt. Wir werden sehen, daß für LTI Systeme beide Konzepte äquivalent sind.

Aussagen über Rekonstruierbarkeit und Beobachtbarkeit lassen sich ziemlich leicht mit Aussagen über Steuerbarkeit eines dualen Systems zeigen. Das folgende Dualitätsprinzip ist auch in vielen weiteren Bereichen der System- und Regelungstheorie hilfreich.

**Satz 2.18 (Dualität)** *Ein LTV System ist rekonstruierbar genau dann, wenn*

$$\dot{x}(t) = A(-t)^T x(t) + C(-t)^T u(t) \tag{2.20}$$

*steuerbar ist.*

**Beweis:** Definiert man  $z(t) := \tilde{x}(t) - x(t)$ , so bedeutet Rekonstruierbarkeit nichts anderes als

$$C(t)z(t) = 0 \quad \forall t \leq t_0 \quad \implies \quad z(t) = 0 \quad \forall t \leq t_0.$$

Dies ist äquivalent zur Aussage:

$$z \neq 0 \text{ Lösung von } \dot{z}(t) = A(t)z(t) \quad \implies \quad C(t)z(t) \neq 0 \quad \text{auf } (-\infty, t_0].$$

Ersetzt man  $t$  durch  $-t$ , so wird diese Aussage zu

$$z \neq 0 \text{ Lösung von } \dot{z}(t) = -A(-t)z(t) \quad \implies \quad C(-t)z(t) \neq 0 \quad \text{auf } [-t_0, \infty).$$

Dies ist aber mit Satz 2.9 äquivalent zur Steuerbarkeit von (2.20).  $\square$

Damit läßt sich nun leicht die Rekonstruierbarkeit von LTV Systemen charakterisieren.

**Satz 2.19** *Ein LTV System ist rekonstruierbar genau dann, wenn zu jedem  $t_1 \in \mathbb{R}$  ein  $t_0 < t_1$  existiert, so daß die  $(t_0, t_1)$ -Rekonstruierbarkeits-Gramsche*

$$Q(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t, t_1)^T C(t)^T C(t) \Phi(t, t_1) dt \quad (2.21)$$

*positiv definit ist.*

**Beweis:** Dies ist eine Konsequenz aus Satz 2.9, angewendet auf das LTV System (2.20), und dem Dualitätsprinzip, Satz 2.19.  $\square$

Für LTI Systeme erhält man als Konsequenz aus dem Dualitätsprinzip und dem Hautus-Popov-Lemma (Satz 2.12) folgende Charakterisierungen von Beobachtbarkeit und Rekonstruierbarkeit. Da beide Begriffe nur über  $A$  und  $C$  definiert werden, sprechen wir bei Beobachtbarkeit und Rekonstruierbarkeit von LTI Systemen der Form (1.4)–(1.5) auch von den entsprechenden Begriffen für Matrixpaare  $(A, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{p \times n}$ .

**Folgerung 2.20 (Hautus-Test)** *Sei  $A \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{p \times n}$ . Die folgenden Aussagen sind äquivalent:*

- a)  $(A, C)$  ist rekonstruierbar.
- b)  $(A, C)$  ist beobachtbar.
- c) Für die Beobachtbarkeitsmatrix

$$\mathcal{O}(A, C) := \mathcal{K}(A^T, C^T)^T = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{np \times n} \quad (2.22)$$

*gilt  $\text{rank}(\mathcal{O}(A, C)) = n$ .*

d) Falls  $z \neq 0$  Rechtseigenvektor von  $A$  ist, so muß  $Cz \neq 0$  gelten.

$$d) \operatorname{rank} \left( \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} \right) = n \text{ für alle } \lambda \in \mathbb{C}.$$

Hat man nur den Ausgang  $y(t)$  beim Entwurf einer Regelung zur Verfügung, so stellt sich das Problem, eine *Ausgangsrückführung*

$$u(t) = Fy(t), \quad F \in \mathbb{R}^{m \times p},$$

zu bestimmen, so daß das vorgegebene Ziel erfüllt wird.

Aufgrund der Äquivalenz von Beobachtbarkeit und Rekonstruierbarkeit bei LTI Systemen verwendet man hier meist nur das Beobachtbarkeitskonzept.

**Bemerkung 2.21** Wendet man die Kalman-Zerlegung (2.18) auf das LTI System  $\dot{x} = A^T x + C^T u$  an, so erhält man eine orthogonale Matrix  $W \in \mathbb{R}^{n \times n}$  mit

$$W^T A^T W = \begin{bmatrix} A_1^T & A_2^T \\ 0 & A_3^T \end{bmatrix}, \quad W^T C^T = \begin{bmatrix} C_1^T \\ 0 \end{bmatrix},$$

wobei  $A_{11} \in \mathbb{R}^{r \times r}$ ,  $C_1 \in \mathbb{R}^{p \times r}$  und  $r = \dim(\mathcal{O}(A, C))$ . Dies ergibt die (orthogonale) Beobachtbarkeits-Kalman-Zerlegung

$$W^T A W = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}, \quad C W = \begin{bmatrix} C_1 \\ 0 \end{bmatrix}. \quad (2.23)$$

Mit dem Basiswechsel  $\tilde{x} := W^T x$  und Partitionierung analog zu (2.23) ergibt sich das System

$$\begin{aligned} \dot{\tilde{x}}_1(t) &= A_1 \tilde{x}_1(t), \\ \dot{\tilde{x}}_2(t) &= A_2 \tilde{x}_1(t) + A_3 \tilde{x}_2(t), \\ y(t) &= C_1 \tilde{x}_1(t). \end{aligned}$$

Die Zustandsvariablen in  $\tilde{x}_2$  haben also keinen Einfluß auf den Ausgang, man nennt sie daher nicht-beobachtbare Zustände, die Eigenwerte von  $A_3$  die unbeobachtbaren Moden.

Auch die Beobachtbarkeit läßt sich abschwächen, analog zur Abschwächung der Steuerbarkeit zur Stabilisierbarkeit. Das zur Stabilisierbarkeit duale Konzept ist wie folgt definiert:

**Definition 2.22 (Entdeckbarkeit)** Ein LTI System heißt entdeckbar, falls für jede Lösung  $z(t)$  von  $\dot{z} = Az$  mit  $Cz(t) \equiv 0$  gilt:  $\lim_{t \rightarrow \infty} z(t) = 0$ .

Setzt man in der Definition  $z = x - \tilde{x}$  für zwei Lösungen  $x, \tilde{x}$  von (1.4) zur selben Eingangsfunktion  $u$ , so kann man Entdeckbarkeit wie folgt interpretieren: Aus  $Cx(t) \equiv C\tilde{x}(t)$  kann man zwar nicht  $x(t) \equiv \tilde{x}(t)$  folgern, aber  $\lim_{t \rightarrow \infty} (x(t) - \tilde{x}(t)) = 0$ . In anderen Worten, der nicht-beobachtbare Anteil des Zustands ist zwar nicht bekannt, aber wir kennen dessen asymptotisches Verhalten.

Mit Satz 2.19 und der Beobachtbarkeits-Kalman-Zerlegung (2.23) erhält man folgende Variante des Dualitätsprinzips:

**Folgerung 2.23** *Das LTI System (1.4)–(1.5) ist entdeckbar genau dann, wenn*

$$\dot{x}(t) = A^T x(t) + C^T u(t)$$

*stabilisierbar ist.*

Analog zum Hautus-Test für die Stabilisierbarkeit erhält man damit nun folgende Aussage:

**Folgerung 2.24 (Hautus-Test für Entdeckbarkeit)** *Sei  $A \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{p \times n}$ . Die folgenden Aussagen sind äquivalent:*

a)  $(A, C)$  ist entdeckbar.

b) Es existiert  $G \in \mathbb{R}^{n \times p}$  mit  $\Lambda(A + GC) \subset \mathbb{C}^-$ .

c) In der Beobachtbarkeits-Kalman-Zerlegung (2.23) gilt  $\Lambda(A_3) \subset \mathbb{C}^-$ .

d) Falls  $z$  ein Eigenvektor von  $A$  zu einem Eigenwert  $\lambda$  mit  $\operatorname{Re}(\lambda) \geq 0$  ist, so muß  $Cz \neq 0$  gelten.

e)  $\operatorname{rank} \left( \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} \right) = n$  für alle  $\lambda \in \mathbb{C}$  mit  $\operatorname{Re}(\lambda) \geq 0$ .

Will man ein LTI System stabilisieren und hat nur den Ausgang  $y(t)$  beim Entwurf der Regelung zur Verfügung, so stellt sich also das Problem, eine Ausgangsrückführung  $u(t) = Fy(t)$  mit  $F \in \mathbb{R}^{m \times p}$  zu bestimmen, so daß  $\lim_{t \rightarrow \infty} x(t; u) = 0$ .

## 2.4 Die Staircase-Form der Systemtheorie

In diesem Abschnitt widmen wir uns der Frage, wie man die Eigenschaften von LTI Systemen numerisch überprüfen kann. Zwar liefert der Hautus-Test dazu gewisse Möglichkeiten, aber die Zuverlässigkeit der numerisch zu treffenden Entscheidungen dort hängt stark von der Kondition des Eigenwertproblems für  $A$  ab. In diesem Abschnitt werden wir eine davon unabhängige Möglichkeit kennenlernen.

**Definition 2.25** *Folgende Transformationen sind Zustandsraumtransformationen:*



a) Basiswechsel:

$$\begin{aligned} x &\rightarrow Px && \text{für } P \in \mathbb{R}^{n \times n} && \text{regulär,} \\ u &\rightarrow Qu && \text{für } Q \in \mathbb{R}^{m \times m} && \text{regulär,} \\ y &\rightarrow Ry && \text{für } R \in \mathbb{R}^{p \times p} && \text{regulär.} \end{aligned}$$

b) Lineare Zustandsrückführung:  $u \rightarrow Fx + u$  mit  $F \in \mathbb{R}^{m \times n}$ .

c) Lineare Ausgangsrückführung:  $u \rightarrow Gy + u$  mit  $G \in \mathbb{R}^{m \times p}$ .

Z.B. führt ein Basiswechsel in  $\mathcal{X} = \mathbb{R}^n$  zum äquivalenten LTI System:

$$\dot{\tilde{x}} = \tilde{A}x + \tilde{B}u, \quad y = \tilde{C}\tilde{x},$$

mit  $\tilde{A} := PAP^{-1}$ ,  $\tilde{B} := PB$  und  $\tilde{C} := CP^{-1}$ .

Allgemeiner kann man sich die Wirkung von Zustandsraumtransformationen mit Hilfe der einem LTI System zugeordneten Blockmatrix  $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{(n+p) \times (n+m)}$  (auch *Systemmatrix* genannt) klarmachen. Basiswechsel und lineare Zustandsrückführung führen zum neuen LTI System

$$\begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} := \begin{bmatrix} P & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ F & Q^{-1} \end{bmatrix} = \begin{bmatrix} PAP^{-1} + PBF & PBQ^{-1} \\ RCP^{-1} + RDF & RDQ^{-1} \end{bmatrix}, \quad (2.24)$$

während man bei Basiswechsel und linearer Ausgangsrückführung

$$\begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} := \begin{bmatrix} P & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ GC & Q^{-1} \end{bmatrix} = \begin{bmatrix} PAP^{-1} + PBGC & PBQ^{-1} \\ RCP^{-1} + RDGC & RDQ^{-1} \end{bmatrix} \quad (2.25)$$

erhält. Hierbei sind natürlich  $P = I_n$ ,  $Q = I_m$ ,  $R = I_p$ ,  $G = 0$  und  $F = 0$  zugelassen.

Mit Zustandsraumtransformationen kann man nun versuchen, ein LTI System so zu transformieren, daß man seine Systemeigenschaften leichter analysieren kann. Dies beruht auf folgendem Satz.

**Satz 2.26** a) *Das LTI System (1.4)–(1.5) ist steuerbar bzw. stabilisierbar genau dann, wenn die durch Zustandsraumtransformationen erhaltenen LTI Systeme (2.24) bzw. (2.25) steuerbar bzw. stabilisierbar sind.*

b) *Das LTI System (1.4)–(1.5) ist beobachtbar bzw. entdeckbar genau dann, wenn das durch Zustandsraumtransformationen erhaltene LTI Systeme (2.25) beobachtbar bzw. entdeckbar ist.*

**Beweis:** Alle Aussagen kann man mit dem für die jeweilige Systemeigenschaft geltenden Hautus-Test beweisen, in dem man  $\text{rank} \left( \begin{bmatrix} \tilde{A} - \lambda I & \tilde{B} \end{bmatrix} \right) = \text{rank} \left( \begin{bmatrix} A - \lambda I & B \end{bmatrix} \right)$  bzw.  $\text{rank} \left( \begin{bmatrix} \tilde{A} - \lambda I \\ \tilde{C} \end{bmatrix} \right) = \text{rank} \left( \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} \right)$  zeigt.  $\square$

Beachte: Bei Zustandsrückführung kann Beobachtbarkeit bzw. Entdeckbarkeit verlorengehen, wie das Beispiel

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = [0 \quad 1], \quad F = -A,$$

zeigt. Bei Zustandsrückführungen ist allerdings die Frage nach Beobachtbarkeit bzw. Entdeckbarkeit irrelevant, da man ohnehin davon ausgeht, daß man den gesamten Zustand zur Verfügung hat.

Da wir aus numerischen Stabilitätsgründen nur orthogonale Transformationsmatrizen zulassen wollen, können wir z.B.  $A$  nicht auf Jordan-Normalform transformieren. Insgesamt zeigt sich aber, daß wir mit Hilfe orthogonaler Basiswechsel im Zustands- und Eingangsraum (bzw. Zustands- und Ausgangsraum) eine Darstellung von  $(A, B)$  (bzw.  $(A, C)$ ) erhalten können, die es uns erlaubt, die Systemeigenschaften abzulesen. Diese Form wird Staircase-Form genannt:

**Satz 2.27 (Staircase-Form, Rosenbrock 1968, Van Dooren 1979)**

Für alle  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  existieren orthogonale Matrizen  $P \in \mathbb{R}^{n \times n}$  und  $Q \in \mathbb{R}^{m \times m}$ , so daß

$$\begin{aligned}
 PAP^T &= \left[ \begin{array}{cccc|cc}
 A_{11} & \dots & \dots & \dots & A_{1,s-1} & A_{1,s} \\
 A_{21} & \ddots & & & \vdots & \vdots \\
 0 & \ddots & \ddots & & \vdots & \vdots \\
 \vdots & \ddots & & & \vdots & \vdots \\
 0 & \dots & 0 & A_{s-2,s-2} & A_{s-2,s-1} & A_{s-2,s} \\
 0 & \dots & 0 & A_{s-1,s-2} & A_{s-1,s-1} & A_{s-1,s} \\
 \hline
 0 & \dots & \dots & \dots & 0 & A_{ss} \\
 n_1 & \dots & \dots & n_{s-2} & n_{s-1} & n_s
 \end{array} \right] \begin{array}{l} n_1 \\ n_2 \\ \vdots \\ n_{s-2} \\ n_{s-1} \\ n_s \end{array}, \\
 PBQ &= \left[ \begin{array}{cc}
 B_1 & 0 \\
 0 & 0 \\
 \vdots & \vdots \\
 0 & 0 \\
 n_1 & m - n_1
 \end{array} \right] \begin{array}{l} n_1 \\ n_2 \\ \vdots \\ n_s \end{array},
 \end{aligned} \tag{2.26}$$

mit  $\text{rank}(B) = n_1 \geq n_2 \geq \dots \geq n_{s-1} \geq n_s \geq 0$ ,  $n_{s-1} > 0$  und  $A_{i,i-1} = [\Sigma_{i,i-1} \quad 0] \in \mathbb{R}^{n_i \times n_{i-1}}$ . Die  $\Sigma_{i,i-1} \in \mathbb{R}^{n_i \times n_i}$  für  $i = 1, \dots, s-1$  sowie  $B_1 \in \mathbb{R}^{n_1 \times n_1}$  sind regulär und  $\Sigma_{s-1,s-2}$  ist diagonal.

**Beweis:** Der Staircase-Algorithmus (Algorithmus 2.28) ist ein konstruktiver Weg zur Berechnung der Staircase-Form für beliebige  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ .  $\square$

Der Staircase-Algorithmus auf Seite 28 verwendet in allen Berechnungsschritten die Singulärwertzerlegung (kurz SVD, von engl., *singular value decomposition*), die für jedes  $M \in \mathbb{R}^{n \times m}$  eine Zerlegung  $M = U\Sigma V^T$  mit orthogonalen  $U \in \mathbb{R}^{n \times n}$  und  $V \in \mathbb{R}^{m \times m}$

berechnet, so daß  $\Sigma$  eine Diagonalmatrix mit nichtnegativen Diagonalelementen  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(n,m)} \geq 0$  ist. Mit  $\sigma_r > 0$ ,  $\sigma_{r+1} = 0$  gilt  $r = \text{rank}(A)$ . (Siehe z.B. [13, Section 2.5.3].)

Der Staircase-Algorithmus bricht entweder mit  $n_s = 0$  oder  $A_{s,s-1} = 0$  ab, daher ergibt sich die Staircase-Form (2.26). Da offensichtlich  $n_i \leq n_{i-1}$  für alle  $i = 1, 2, \dots$  gilt und der Restblock, von dem eine SVD berechnet wird, immer mindestens eine Zeile weniger hat als im vorangegangenen Schritt, muß eine der Bedingungen nach höchstens  $n$  Schritten eintreten, da  $\text{rank}(A_{i,i-1}) \geq 1$  solange  $A_{i,i-1} \neq 0$ .

Da ausschließlich Singulärwertzerlegungen verwendet werden, ist Algorithmus 2.28 numerisch rückwärts stabil, solange keine falschen Rangentscheidungen getroffen werden. Dies ist natürlich der kritische Teil des Algorithmus, da die akkumulierten Rundungsfehler zu numerischen Rängen führen können, die nicht den exakten Rängen entsprechen. Der Aufwand des Staircase-Algorithmus beträgt schlimmstenfalls  $\mathcal{O}(n^4)$ , da höchstens  $n$  Schritte erforderlich sind und jeder Schritt höchstens  $\mathcal{O}(n^3)$  Flops kostet.

Wenden wir den Staircase-Algorithmus auf  $(A, B)$  bzw.  $(A^T, C^T)$  an, wobei  $A, B, C$  die zu (1.4)–(1.5) gehörenden Systemmatrizen sind, so können wir alle Systemeigenschaften aus den resultierenden Staircase-Formen ablesen.

**Satz 2.29** a) *Das LTI System (1.4)–(1.5) ist steuerbar genau dann, wenn  $n_s = 0$  in der Staircase-Form von  $(A, B)$ .*

b) *Das LTI System (1.4)–(1.5) ist beobachtbar genau dann, wenn  $n_s = 0$  in der Staircase-Form von  $(A^T, C^T)$ .*

c) *Das LTI System (1.4)–(1.5) ist stabilisierbar genau dann, wenn in der Staircase-Form von  $(A, B)$  gilt:  $\Lambda(A_{s,s}) \subset \mathbb{C}^-$ , d.h.  $A_{s,s}$  ist Hurwitz.*

d) *Das LTI System (1.4)–(1.5) ist entdeckbar genau dann, wenn in der Staircase-Form von  $(A^T, C^T)$  gilt:  $\Lambda(A_{s,s}) \subset \mathbb{C}^-$ , d.h.  $A_{s,s}$  ist Hurwitz.*

**Beweis:**

a) Der Hautus-Test (Satz 2.12, Teil d)), angewendet auf die Staircase-Form von  $(A, B)$ , und Satz 2.26, besagen, daß  $(A, B)$  steuerbar ist genau dann, wenn für alle  $\lambda \in \mathbb{C}$

$$n = \text{rank} \left( \left[ \begin{array}{cccccc|cc} \lambda I_{n_1} - A_{11} & -A_{12} & \dots & \dots & -A_{1,s} & B_1 & 0 \\ -A_{21} & \lambda I_{n_2} - A_{22} & \ddots & & \vdots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & -A_{s-1,s-2} & \lambda I_{n_{s-1}} - A_{s-1,s-1} & -A_{s-1,s} & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \lambda I_{n_s} - A_{s,s} & 0 & 0 \end{array} \right] \right)$$

gilt. Da nach Konstruktion  $B_1, A_{21}, \dots, A_{s-1,s-2}$  alle vollen Rang haben, folgt die Äquivalenz der Steuerbarkeit zu

$$\text{rank}(\lambda I_{n_s} - A_{s,s}) = n_s \quad \forall \lambda \in \mathbb{C}.$$

**Algorithmus 2.28 (Staircase-Algorithmus)**INPUT:  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ .OUTPUT:  $\tilde{A} := PAP^T, \tilde{B} := PBQ$  in Staircase-Form.

1: Berechne die SVD von B:

$$B = U_B \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix} V_B^T \quad \text{mit } \Sigma_B \in \mathbb{R}^{n_1 \times n_1} \text{ regulär.}$$

Definiere  $P := U_B^T, Q := V_B$  und

$$A := U_B^T A U_B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{11} \in \mathbb{R}^{n_1 \times n_1},$$

$$B := U_B^T B V_B = \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 \in \mathbb{R}^{n_1 \times n_1}.$$

2: Berechne die SVD von  $A_{21} \in \mathbb{R}^{n-n_1 \times n_1}$ :

$$A_{21} = U_{21} \begin{bmatrix} \Sigma_{21} & 0 \\ 0 & 0 \end{bmatrix} V_{21}^T \quad \text{mit } \Sigma_{21} \in \mathbb{R}^{n_2 \times n_2} \text{ regulär (Beachte: } n_2 \leq n_1).$$

Definiere  $P_2 := V_{21}^T \oplus U_{21}^T, P := P_2 P$  und

$$A := P_2 A P_2^T = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix}, \quad A_{21} = \begin{bmatrix} \Sigma_{21} & 0 \end{bmatrix} \in \mathbb{R}^{n_2 \times n_1},$$

$$B := P_2 B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 := V_{21}^T \quad \text{mit } \Sigma_B \in \mathbb{R}^{n_1 \times n_1} \text{ regulär.}$$

3: Setze  $i := 3$ .4: **while**  $n_{i-1} > 0$  **and**  $A_{i,i-1} \neq 0$  **do**5: Berechne die SVD von  $A_{i,i-1} \in \mathbb{R}^{(n-n_1-\dots-n_{i-1}) \times n_{i-1}}$ :

$$A_{i,i-1} = U_{i,i-1} \begin{bmatrix} \Sigma_{i,i-1} & 0 \\ 0 & 0 \end{bmatrix} V_{i,i-1}^T \quad \text{mit } \Sigma_{i,i-1} \in \mathbb{R}^{n_i \times n_i} \text{ regulär (Beachte: } n_i \leq n_{i-1}).$$

6: Setze  $P_i := I_{n_1} \oplus \dots \oplus I_{n_{i-2}} \oplus V_{i,i-1}^T \oplus U_{i,i-1}^T, P := P_i P$  und

$$A := P_i A P_i^T = \begin{bmatrix} A_{11} & \dots & \dots & & A_{1,i+1} \\ A_{21} & A_{22} & & & A_{2,i+1} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & A_{i,i-1} & A_{i,i} & \vdots \\ 0 & \dots & 0 & A_{i+1,i} & A_{i+1,i+1} \end{bmatrix},$$

$$A_{i,i-1} = \begin{bmatrix} \Sigma_{i,i-1} & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times n_{i-1}}.$$

7: Setze  $i := i + 1$ .8: **end while**9: Setze  $s := i$ .

Dies kann aber nur dann der Fall sein wenn  $n_s = 0$ , da für alle Eigenwerte  $\lambda$  von  $A_{s,s}$   $\text{rank}(\lambda I_{n_s} - A_{s,s}) < n_s$  gilt.

- b) Dies folgt mit a) und der Dualität (Satz 2.19) sowie Folgerung 2.20.
- c) Analog zu a) mit Satz 2.15, Teil e).
- d) Dies folgt mit c) und der Dualität (Folgerung 2.23) sowie Folgerung 2.24.

□

Ein Spezialfall der Staircase-Form ergibt sich für Single-Input ( $m = 1$ ) bzw. Single-Output ( $p = 1$ ) Systeme. (Im Fall von  $m = p = 1$  spricht man auch von *Eingrößen-Regelungssystemen*, engl. *single-input-single-output systems*, kurz SISO Systeme, während für  $m > 1$  und  $p > 1$  die Abkürzung MIMO System (von engl., *multi-input-multi-output system*) gebräuchlich ist. In der Regelungstheorie spricht man in diesem Fall von *Mehrgrößen-Regelungssystemen*.)

**Folgerung 2.30 (System-Hessenberg-Form)** Für alle  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$  existiert eine orthogonale Matrix  $P \in \mathbb{R}^{n \times n}$ , so daß der von  $P$  induzierte Basiswechsel im Zustandsraum  $(A, B)$  in System-Hessenberg-Form

$$\hat{A} := PAP^T = \begin{bmatrix} a_{11} & & & a_{1,n} \\ a_{21} & & & \vdots \\ & \ddots & & \vdots \\ & & a_{n,n-1} & a_{n,n} \end{bmatrix}, \hat{B} := PB = \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.27)$$

transformiert.

**Beweis:** Dies folgt mit dem Staircase-Algorithmus sowie  $n_1 \leq m = 1$ . Falls  $n_s \neq 0$  muß nur noch eine Hessenberg-Reduktion  $U_s A_{s,s} U_s^T = \begin{bmatrix} \diagdown \end{bmatrix}$  mit  $U_s \in \mathbb{R}^{n_s \times n_s}$  orthogonal berechnet sowie mit  $P_s := I_{n-n_s} \oplus U_s$  die Ähnlichkeitstransformation  $A := P_s A P_s^T$  und ggf.  $P := P_s P$  durchgeführt werden. □

Damit erhält man aus Satz 2.29 direkt Kriterien für Steuerbarkeit und Beobachtbarkeit von Single-Input bzw. Single-Output Systemen.

**Folgerung 2.31 a)** Ein Single-Input System ist steuerbar genau dann, wenn in (2.27)  $PAP^T$  eine unreduzierte Hessenberg-Matrix ist (d.h.,  $a_{i,i-1} \neq 0$  für alle  $i = 1, \dots, n-1$ ) und  $b_1 \neq 0$ .

b) Ein Single-Output System ist beobachtbar genau dann, wenn in (2.27), angewendet auf  $(A^T, C^T)$ ,  $\hat{A}$  eine unreduzierte Hessenberg-Matrix ist und  $b_1 \neq 0$ .



# Kapitel 3

## Stabilisierung, Lyapunovgleichungen und Polvorgabe

In diesem Kapitel gehen wir der Frage nach, wie eine stabilisierende Feedback-Matrix für LTI Systeme berechnet werden kann. Aus Satz 2.15 wissen wir, daß die Berechnung einer stabilisierenden Steuerungsfunktion  $u(t)$  mit Hilfe von Zustandsrückführung möglich ist. Dazu benötigen wir  $F \in \mathbb{R}^{m \times n}$ , so daß  $\Lambda(A + BF) \subset \mathbb{C}^-$ . Mit  $u(t) := Fx(t)$  folgt, daß die Lösungstrajektorie von

$$\dot{x} = Ax + Bu = Ax + BFx = (A + BF)x(t)$$

genau dann asymptotisch stabil ist, wenn  $F$  stabilisierend ist, d.h.  $\Lambda(A + BF) \subset \mathbb{C}^-$  erfüllt.

Eine stabilisierende Zustandsrückführung erhält man unter bestimmten Voraussetzungen durch die Lösung von Optimalsteuerungsproblemen, siehe Kapitel 4.

Hier wollen wir zunächst zwei einfachere Methoden kennenlernen.

**Direkte Methode nach Lyapunov** Mit Hilfe der Lyapunovschen Stabilitätstheorie und der Lösung eines linearen Gleichungssystem kann eine stabilisierende Feedback-Matrix direkt berechnet werden.

Diese Methode ist ein Spezialfall einer allgemeineren Theorie für nichtlineare Systeme, die auf der Berechnung von *Lyapunovfunktionen* für nichtlineare Systeme der Form  $\dot{x} = f(x)$  (mit  $f(0) = 0$ , d.h.  $x = 0$  ist eine Gleichgewichtslage des dynamischen Systems) beruht. Dabei sucht man eine in einer Umgebung von  $x = 0$  positiv definite, total differenzierbare Funktion  $V : \mathcal{X} \rightarrow \mathbb{R}$ , so daß die Richtungsableitung (auch: *Lie-Ableitung*) von  $V$  in Richtung  $f$ , bezeichnet mit  $L_f V$  und für  $x \in \mathcal{X}$  berechenbar als

$$L_f V(x) = \nabla V \cdot f(x),$$

$V(0) = 0$  und  $L_f V(x) < 0$  für alle  $x \neq 0$  erfüllt. Kann man solche eine Lyapunovfunktion finden, dann sagt der Satz von Lyapunov, daß  $x = 0$  asymptotisch stabil ist (siehe z.B. [12, Chapter 2, §4]). I.a. ist es leider nicht immer einfach, eine Lyapunovfunktion zu bestimmen. Der Begriff Lyapunovfunktion wird hier (also bei LTI Systemen)

keine besondere Rolle spielen; es sei hier nur erwähnt, daß bei der hier vorgestellten Methode implizit verwendet wird, daß für asymptotisch stabile und steuerbare LTI Systeme  $V(x) = x^T P x$ ,  $x \in \mathbb{R}^n$  eine (quadratische) Lyapunovfunktion des freien (ungesteuerten, d.h.  $u \equiv 0$ ) Systems ist, wobei  $P$  Lösung der Lyapunovgleichung  $AP + PA^T + I_n = 0$  ist.

**Polvorgabe** Im allgemeinen besteht das Polvorgabe-Problem darin, für eine gegebene Menge  $\mathcal{L} := \{\mu_1, \dots, \mu_n\} \subset \mathbb{C}$  eine Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$  so zu bestimmen, daß  $\Lambda(A + BF) = \mathcal{L}$ . Wählt man  $\mathcal{L} \subset \mathbb{C}^-$ , so hat man damit das System stabilisiert.

Dieser Ansatz läßt sich nicht ohne weiteres auf nichtlineare Systeme verallgemeinern. (Erinnere: Für LTV Systeme ist  $\Lambda(A(t) + B(t)F(t)) \in \mathbb{C}^-$  weder hinreichend noch notwendig für Stabilität. Lokal können nichtlineare Systeme durch LTV Systeme genähert werden, d.h., zur Stabilisierung mit Polvorgabe müßte man zumindest LTV Systeme stabilisieren können, aber selbst das ist nicht ausreichend. Um mit Hilfe lokaler Stabilisierungen der Linearisierungen insgesamt Stabilisierung eines nichtlinearen Systems zu erreichen, benötigt man zusätzliche Annahmen. Eine Technik, die dies liefert, ist *modellprädiktive Regelung* (kurz MPC, von engl. *model predictive control*), siehe z.B. [9].)

Zunächst werden wir uns mit der Methode nach Lyapunov beschäftigen. Die Darstellung des nächsten Abschnitts orientiert sich an [19, § 13] und [14].

### 3.1 Stabilitätstheorie nach Lyapunov

In diesem Abschnitt werden lineare Matrixgleichungen eine wesentliche Rolle spielen. Daher werden wir uns zunächst mit einigen Eigenschaften solcher Gleichungen befassen. Betrachte dazu die *Sylvestergleichung*

$$AX + XB = W \tag{3.1}$$

mit  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{m \times m}$ ,  $W \in \mathbb{R}^{n \times m}$  und der unbekannt Matrix  $X \in \mathbb{R}^{n \times m}$ . Bei dieser Gleichung handelt es sich um eine lineare Gleichung in den  $n \cdot m$  Unbekannten  $x_{ij}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . Also existiert eine Darstellung von (3.1) in der üblichen Form  $Mx = w$  eines linearen Gleichungssystems in  $\mathbb{R}^{mn}$ . Mit Hilfe dieser Darstellung können wir sofort Aussagen über die (eindeutige) Lösbarkeit von Sylvestergleichungen erhalten. Dazu benötigen wir ein Kalkül, welches auf der folgenden Definition beruht.

**Definition 3.1** Es seien  $A \in \mathbb{R}^{n \times p}$  und  $B \in \mathbb{R}^{m \times q}$ . Dann ist das Kroneckerprodukt (Tensorprodukt) von  $A$  und  $B$  definiert durch

$$A \otimes B := \begin{bmatrix} a_{11}B & \dots & a_{1,p}B \\ \vdots & & \vdots \\ a_{n,1}B & \dots & a_{n,p}B \end{bmatrix} \in \mathbb{R}^{nm \times pq}.$$



Außerdem wird der  $\text{vec}$ -Operator definiert durch

$$\text{vec} : \mathbb{R}^{n \times p} \rightarrow \mathbb{R}^{n \cdot p} : A \rightarrow [a_{11}, \dots, a_{n,1}, a_{12}, \dots, a_{n,2}, \dots, a_{n,p}]^T.$$

Folgende Eigenschaften des Kroneckerprodukts ergeben sich direkt aus der Definition 3.1:

- a)  $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$  für alle  $\alpha \in \mathbb{R}$ ,
- b)  $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$ ,
- b)  $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ ,
- d)  $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ ,
- e)  $(A \otimes B)^T = A^T \otimes B^T$ ,
- f)  $(A \otimes B)(C \otimes D) = AC \otimes BD$ ,
- g)  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ , falls  $A$  und  $B$  invertierbar sind.

Eine weitere wichtige Eigenschaft verbindet Kroneckerprodukt und den  $\text{vec}$ -Operator, mit dieser Eigenschaft lassen sich dann Sylvestergleichungen “vektorisieren”.

**Lemma 3.2** Für  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{m \times m}$  und  $X \in \mathbb{R}^{n \times m}$  gilt:

$$\text{vec}(AXB) = (B^T \otimes A) \text{vec}(X).$$

**Beweis:** Übung. □

Damit erhält man sofort die gewünschte vektorisierte Darstellung der Sylvestergleichung (3.1).

**Folgerung 3.3** Die Sylvestergleichung (3.1) ist äquivalent zu

$$\left( (I_m \otimes A) + (B^T \otimes I_n) \right) \text{vec}(X) = \text{vec}(W), \quad (3.2)$$

d.h.  $X$  löst (3.1) genau dann, wenn  $\text{vec}(X)$  (3.2) löst.

Definiert man  $M := (I_m \otimes A) + (B^T \otimes I_n)$ , so ist sofort klar, daß die Sylvestergleichung genau dann eine eindeutige Lösung hat, wenn  $M$  regulär ist. Ein hinreichendes und notwendiges Kriterium dafür ist bekanntermaßen, daß  $M$  keinen Eigenwert  $\lambda = 0$  hat. Da man mit Hilfe des Kroneckerprodukt-Kalküls die Eigenwerte von  $M$  explizit angeben kann, erhält man das gewünschte Ergebnis über die Lösbarkeit der Sylvestergleichung (3.1).

Die Beziehung zwischen Eigenwerten von  $A$ ,  $B$  und  $M$  erhält man mit folgendem Ergebnis:

**Satz 3.4 (Satz von Stephanos)** Es sei  $p(x, y)$  ein komplexes Polynom in zwei Variablen, d.h.  $p(x, y) = \sum_{j,k=1}^{\ell} \alpha_{jk} x^j y^k$  mit  $x, y, \alpha_{jk} \in \mathbb{C}$ . Für  $A \in \mathbb{C}^{n \times n}$  und  $B \in \mathbb{C}^{m \times m}$  definiere damit ein matrixwertiges Polynom, indem die Multiplikation durch das Kronecker-Produkt ersetzt wird:

$$p(A, B) := \sum_{j,k=1}^{\ell} \alpha_{jk} (A^j \otimes B^k).$$

Sind die Spektren von  $A$  und  $B$  gegeben durch  $\Lambda(A) = \{\lambda_1, \dots, \lambda_n\}$ ,  $\Lambda(B) = \{\mu_1, \dots, \mu_m\}$ , dann gilt

$$\Lambda(p(A, B)) = \{p(\lambda_j, \mu_k), j = 1, \dots, n, k = 1, \dots, m\}.$$

**Beweis:** Übung. □

Damit können wir jetzt Aussagen zur Lösbarkeit von (3.1) treffen.

**Satz 3.5** Betrachte die Sylvestergleichung (3.1) und sei  $\Lambda(A) = \{\lambda_1, \dots, \lambda_n\}$ ,  $\Lambda(B) = \{\mu_1, \dots, \mu_m\}$ . Dann gilt:

a)  $\Lambda(M) = \Lambda((I_m \otimes A) + (B^T \otimes I_n)) = \{\lambda_j + \mu_k, j = 1, \dots, n, k = 1, \dots, m\}$ .

b) Die Sylvestergleichung (3.1) und damit das lineare Gleichungssystem (3.2) haben eine eindeutige Lösung genau dann, wenn  $\Lambda(A) \cap \Lambda(-B) = \emptyset$ .

**Beweis:**

a) Folgt aus dem Satz von Stephanos/Übung.

b)  $0 \in \Lambda(M) \iff$  Es existiert  $\lambda_j \in \Lambda(A)$  und  $\mu_k \in \Lambda(B)$  mit  $\lambda_j = -\mu_k$ .  
 $\iff \Lambda(A) \cap \Lambda(-B) \neq \emptyset.$  □

Betrachte nun den Spezialfall von (3.1) mit  $m = n$ ,  $B = A^T$  und  $W = W^T$ . Damit erhält man die *Lyapunovgleichung*

$$AX + XA^T = W. \tag{3.3}$$

Da die Lyapunovgleichung symmetrisch ist, folgt sofort daß auch  $X^T$  Lösung von (3.3) ist. Falls die Lösung eindeutig ist, d.h. nach Satz 3.5b), falls  $\Lambda(A) \cap \Lambda(-A) = \emptyset$ , dann ist diese eindeutige Lösung symmetrisch. Eine hinreichende Bedingung für die Existenz der eindeutigen Lösbarkeit ist, daß  $\Lambda(A) \subset \mathbb{C}^-$ , also, daß  $A$  Hurwitz<sup>1</sup> ist. Man erhält in diesem Fall sogar eine explizite Lösungsformel, die auch allgemeiner für (3.1) gültig ist, falls  $A$  und  $B$  stabil sind.

---

<sup>1</sup>Das bedeutet, daß  $A$  ist *asymptotisch* stabil ist, siehe Definition ???. Allerdings ist der Begriff "stabil" für Matrizen mit Spektrum in der offenen linken Halbebene gebräuchlich in der Matrixtheorie und wird deshalb auch im Folgenden derart verwendet.

**Satz 3.6** Sind  $\Lambda(A), \Lambda(B) \subset \mathbb{C}^-$ , dann hat (3.1) eine eindeutige Lösung, die durch

$$X = - \int_0^\infty e^{At} W e^{Bt} dt \quad (3.4)$$

gegeben ist.

**Beweis:** Die Eindeutigkeit der Lösung folgt sofort aus Satz 3.5b).

Definiere nun  $Z : [0, \infty) \rightarrow \mathbb{R}^{n \times m}$  als Lösung der linearen matrixwertigen Differentialgleichung

$$\dot{Z}(t) = AZ(t) + Z(t)B \quad (3.5)$$

für die Anfangsbedingung  $Z(0) = W$ . Aus der Theorie linearer homogener Differentialgleichungen folgt, daß diese Anfangswertaufgabe für die *Sylvester-Differentialgleichung* (3.5) eine eindeutige Lösung auf ganz  $[0, \infty)$  besitzt. Diese Lösung ist  $Z(t) = e^{At} W e^{Bt}$ , wie man leicht nachrechnet:  $Z(0) = W$  und

$$\dot{Z}(t) = A e^{At} W e^{Bt} + e^{At} W B e^{Bt} = A e^{At} W e^{Bt} + e^{At} W e^{Bt} B = AZ(t) + Z(t)B,$$

wobei man verwendet, daß  $B$  und  $e^{Bt}$  kommutieren. Da  $A, B$  stabil sind, gilt  $\lim_{t \rightarrow \infty} e^{At} = 0$ ,  $\lim_{t \rightarrow \infty} e^{Bt} = 0$  und damit

$$Z_\infty := \lim_{t \rightarrow \infty} Z(t) = \lim_{t \rightarrow \infty} e^{At} W e^{Bt} = 0.$$

Integration von (3.5) über  $[0, \infty)$  liefert

$$Z_\infty - Z(0) = A \int_0^\infty Z(t) dt + \int_0^\infty Z(t) dt B,$$

also

$$A \int_0^\infty Z(t) dt + \int_0^\infty Z(t) dt B = -W.$$

Damit ist  $-\int_0^\infty Z(t) dt$  eine Lösung der Sylvestergleichung (3.1). Aus der Eindeutigkeit der Lösung folgt dann  $X = -\int_0^\infty Z(t) dt = -\int_0^\infty e^{At} W e^{Bt} dt$ .  $\square$

Betrachtet man nun ein asymptotisch stabiles LTI System  $\dot{x} = Ax + Bu$  und setzt  $W = -BB^T$  in (3.3), dann erhält man aus Satz 3.6 sofort, daß die *Steuerbarkeits-Gramsche Matrix*

$$P = \int_0^\infty e^{At} B B^T e^{A^T t} dt$$

die eindeutige Lösung der *stabilen Lyapunovgleichung*

$$AP + PA^T + BB^T = 0 \quad (3.6)$$

ist. Daraus ergibt sich ein weiteres Kriterium für die Steuerbarkeit von LTI Systemen.

**Folgerung 3.7** Sei  $\Lambda(A) \subset \mathbb{C}^-$ . Dann ist das durch  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  definierte LTI System steuerbar genau dann, wenn die eindeutige Lösung der Lyapunovgleichung (3.6) positiv definit ist.

**Bemerkung 3.8** Man beachte, daß  $P \neq \lim_{\tau \rightarrow \infty} P(0, \tau)$ , da dieser Grenzwert für  $A$  Hurwitz nicht definiert ist. Damit ist  $P$  nicht eine “[0,  $\infty$ ]-Steuerbarkeits-Gramsche”. Die Bezeichnung Steuerbarkeits-Gramsche leitet sich aus obiger Steuerbarkeitscharakterisierung (Folgerung 3.7) ab.

Analog erhält man eine Charakterisierung der Beobachtbarkeit für asymptotisch stabile LTI Systeme über die positive Definitheit der *Beobachtbarkeits-Gramschen Matrix*, die als Lösung von

$$A^T Q + Q A + C^T C = 0$$

gegeben ist.

Im Folgenden zitieren wir die beiden grundlegenden Sätze aus der Lyapunovschen Stabilitätstheorie, die im weiteren benötigt werden.

**Satz 3.9 (Satz von Lyapunov, 1897)** Seien  $A, W \in \mathbb{R}^{n \times n}$  mit  $W = W^T$  negativ definit. Dann gilt:

- Ist  $\Lambda(A) \subset \mathbb{C}^-$ , dann hat die Lyapunovgleichung (3.3) eine eindeutige Lösung  $X$ . Außerdem gilt  $X = X^T > 0$ .
- Falls (3.3) eine Lösung  $X > 0$  hat, dann ist  $A$  stabil.

**Beweis:** Teil a) folgt aus Satz 3.6. Für Teil b) siehe z.B. [14, 17, 19]. □

Aus Satz 3.9b) folgt ein Test auf asymptotische Stabilität bei linearen dynamischen Systemen, den man als *direkte Methode nach Lyapunov*<sup>2</sup> bezeichnet. Dazu löst man die Lyapunovgleichung  $AX + XA^T = -\alpha I_n$  für ein  $\alpha < 0$ . Ist  $X > 0$  (was man z.B. mit Hilfe der Cholesky-Zerlegung von  $X$  überprüfen kann), so sind alle Lösungen von  $\dot{x} = Ax$  asymptotisch stabil. Daher bezeichnet man asymptotische Stabilität auch als *Lyapunov-Stabilität*. Wie bereits oben erwähnt, erhält man auch für nichtlineare Systeme eine hinreichende Bedingung mittels Lyapunovscher Stabilitätstheorie, indem man eine Lyapunovfunktion für  $\dot{x} = f(x)$  angibt. (Im linearen Fall ist eine solche Lyapunovfunktion durch  $V(x) = x^T X x$  gegeben.) Allerdings ist die Bestimmung einer Lyapunovfunktion für nichtlineare Systeme oft sehr schwierig und i.a. nicht durch einen einfachen Algorithmus möglich.

In der Regelungs- und Systemtheorie ist die folgende Version von Satz 3.9 wesentlich, die auf Chen (1973) und Wimmer (1974) zurückgeht und bei der die Definitheit der rechten Seite der Lyapunovgleichung abgeschwächt werden kann unter der zusätzlichen Annahme der Steuerbarkeit.

---

<sup>2</sup>Der Name “direkte Methode” bezieht sich darauf, daß keine Lösungstrajektorie berechnet werden muß, um Stabilität zu überprüfen.

**Satz 3.10** *Es sei das durch  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  definierte LTI System steuerbar.*

- a) *Ist  $\Lambda(A) \subset \mathbb{C}^-$ , dann hat die Lyapunovgleichung (3.6) eine eindeutige Lösung  $P$ . Außerdem gilt  $P = P^T > 0$ .*
- b) *Falls (3.6) eine Lösung  $P > 0$  hat, dann ist  $A$  stabil.*

**Beweis:** Teil a) ist nichts anderes als Folgerung 3.7. Für Teil b) siehe z.B. [19, Chapter 13].  $\square$

## 3.2 Stabilisierung mit Lyapunovgleichungen

Der folgende Satz, der auf Kleinman (1970) und Armstrong (1975) zurückgeht und frühere Ideen von Bass aufgreift, liefert ein erstes Stabilisierungsverfahren. Im Folgenden bezeichnet  $M^+$  die (Moore-Penrose) Pseudoinverse von  $M$ , d.h. die eindeutige Matrix, für die die Moore-Penrose Bedingungen

$$\begin{aligned} (i) \quad MM^+M &= M, & (ii) \quad M^+MM^+ &= M^+, \\ (iii) \quad (MM^+)^T &= MM^+, & (iv) \quad (M^+M)^T &= M^+M \end{aligned}$$

erfüllt sind.

**Satz 3.11** *Es sei  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  stabilisierbar und  $\beta \in \mathbb{R}$  mit  $\beta > \rho(A)$ , wobei  $\rho(A) = \max\{|\lambda| \mid \lambda \in \Lambda(A)\}$  der Spektralradius von  $A$  ist. Ist  $X$  die eindeutige Lösung der Lyapunovgleichung*

$$(A + \beta I_n)X + X(A + \beta I_n)^T = 2BB^T, \quad (3.7)$$

*dann ist  $F := -B^T X^+$  eine stabilisierende Feedback-Matrix für  $(A, B)$ .*

**Beweis:** Sei zunächst  $(A, B)$  steuerbar (und damit natürlich auch stabilisierbar). Da  $\beta > \rho(A)$  gilt  $\Lambda(A + \beta I_n) \subset \mathbb{C}^+$ . Um Satz 3.10a) anwenden zu können, muß  $(-(A + \beta I_n), \sqrt{2}B)$  steuerbar sein. Dies folgt aber sofort mit dem Hautus-Test, da

$$\begin{aligned} n &= \text{rank}([A - \lambda I_n, B]) \quad \forall \lambda \in \mathbb{C} \\ &\iff \\ n &= \text{rank} \left( \begin{bmatrix} (A + \beta I_n) + \tilde{\lambda} I_n & B \end{bmatrix} \begin{bmatrix} -I_n & 0 \\ 0 & \sqrt{2}I_m \end{bmatrix} \right) \\ &= \text{rank} \left( \begin{bmatrix} -(A + \beta I_n) - \tilde{\lambda} I_n & \sqrt{2}B \end{bmatrix} \right) \quad \forall \tilde{\lambda} \in \mathbb{C}. \end{aligned}$$

Also folgt mit Satz 3.10b), daß (3.7) eine eindeutige Lösung  $X > 0$  besitzt. Da  $X$  invertierbar ist, ergibt sich die Äquivalenz von (3.7) und

$$X^{-1}(A + \beta I_n) + (A + \beta I_n)^T X^{-1} = 2X^{-1}BB^T X^{-1}.$$

Durch Umstellen erhält man daraus

$$X^{-1}(A - BB^T X^{-1}) + (A - BB^T X^{-1})^T X^{-1} = -2\beta X^{-1}.$$

Da  $X$  und damit  $X^{-1}$  positiv definit sind, ist die rechte Seite dieser Lyapunovgleichung negativ definit. Damit folgt aus Satz 3.9b), daß  $A - BB^T X^{-1}$  stabil ist, d.h.  $F = -B^T X^+$  ist eine stabilisierende Feedback-Matrix, da für invertierbare Matrizen  $X^+ = X^{-1}$ .

Sei nun  $(A, B)$  stabilisierbar. Aufgrund von Satz 3.6 wissen wir, daß (3.7) eine eindeutige Lösung besitzt, die aufgrund der Darstellung (3.4) positiv semidefinit sein muß. Weiterhin wissen wir aus Satz 2.15, daß  $(A, B)$  eine Kalman-Zerlegung der Form

$$A = V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T, \quad B = V \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad (A_1, B_1) \text{ steuerbar, } A_3 \text{ stabil}$$

mit  $V \in \mathbb{R}^{n \times n}$  orthogonal besitzt. Multipliziert man (3.7) von links mit  $V^T$ , von rechts mit  $V$ , setzt  $\hat{X} = V^T X V = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix}$  und partitioniert analog zur Aufteilung in der Kalman-Zerlegung, so ergibt sich

$$\begin{bmatrix} A_1 + \beta I & A_2 \\ 0 & A_3 + \beta I \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} + \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} (A_1 + \beta I)^T & 0 \\ A_2^T & (A_3 + \beta I)^T \end{bmatrix} = \begin{bmatrix} 2B_1 B_1^T & 0 \\ 0 & 0 \end{bmatrix}.$$

Daraus erhält man die folgenden linearen Matrixgleichungen

$$(A_1 + \beta I)X_1 + X_1(A_1 + \beta I)^T + A_2 X_2^T + X_2 A_2^T = 2B_1 B_1^T \quad (3.8)$$

$$(A_3 + \beta I)X_3 + X_3(A_3 + \beta I)^T = 0. \quad (3.9)$$

Die homogene Gleichung (3.9) hat wegen Satz 3.5b) die eindeutige Lösung  $X_3 = 0$ . Da  $X$  und damit auch  $\hat{X}$  positiv semidefinit ist, muß  $X_2 = 0$  gelten. Aus der Steuerbarkeit von  $(A_1, B_1)$  folgt, daß (3.8) eine eindeutige Lösung  $X_1 > 0$  besitzt und  $F_1 = -B_1^T X_1^{-1}$  eine stabilisierende Feedback-Matrix für  $(A_1, B_1)$  ist. Setzt man  $F := \begin{bmatrix} F_1 & 0 \end{bmatrix} V^T$ , so gilt

$$V^T(A + BF)V = \begin{bmatrix} A_1 + B_1 F_1 & A_2 \\ 0 & A_3 \end{bmatrix},$$

d.h.  $\Lambda(A + BF) = \Lambda(A_1 + B_1 F_1) \cup \Lambda(A_3) \subset \mathbb{C}^-$ . Also ist  $F$  eine stabilisierende Feedback-Matrix für  $(A, B)$ . Außerdem gilt

$$F = \begin{bmatrix} -B_1^T X_1^{-1} & 0 \end{bmatrix} V^T = - \begin{bmatrix} B_1^T & 0 \end{bmatrix} \begin{bmatrix} X_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T = -B^T V \begin{bmatrix} X_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T.$$

Durch Nachrechnen überzeugt man sich, daß  $V \begin{bmatrix} X_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T$  die Moore-Penrose Bedingungen bzgl.  $X = V \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} V^T$  erfüllt. Damit gilt also letztlich, daß  $F = -B^T X^+$ .

□

Der obige Beweis verwendet im Wesentlichen, daß die nicht-steuerbaren Moden des LTI Systems nicht stabilisiert werden müssen, so daß das Stabilisierungsproblem auf den steuerbaren Fall zurückgeführt werden kann. Damit erhält man einen vollständigen Beweis von Satz 2.15, da Satz 3.11 den Schritt “c)  $\Rightarrow$  b)” liefert unter der Voraussetzung, daß “a)  $\Rightarrow$  c)” bewiesen ist.

---

**Algorithmus 3.12 (Bass Algorithmus)**


---

INPUT:  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  stabilisierbar.

OUTPUT: Stabilisierende Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$ , d.h.,  $\Lambda(A + BF) \subset \mathbb{C}^-$ .

- 1: Setze  $\beta = 2\|A\|_p$  für eine einfach zu berechnende Norm, z.B.  $p = 1, \infty, F$ .
  - 2: Löse (3.7).
  - 3: Berechne  $X^+$  und setze  $F = -B^T X^+$ .
- 

Der Faktor 2 in Zeile 1 ist ein Sicherheitsfaktor, der bewirken soll, daß die Eigenwerte von  $A + \beta I_n$  hinreichend weit von der imaginären Achse entfernt sind. Die Berechnung der Pseudoinversen  $X^+$  kann z.B. mit Hilfe einer Eigenwertzerlegung der positiv semidefiniten Matrix  $X$  erfolgen.

Eine Alternative zur Stabilisierung mit Hilfe der Lyapunovgleichung (3.7) liefert die sogenannte *algebraische Bernoulligleichung (ABE)*

$$A^T X + X A - X B B^T X = 0. \quad (3.10)$$

Die folgenden Ergebnisse konnten in [?] gezeigt werden:

**Satz 3.13** *Ist  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  stabilisierbar und  $\Lambda(A) \cap i\mathbb{R} = \emptyset$ , dann gelten folgende Aussagen:*

- a) *Die ABE (3.10) hat eine eindeutige stabilisierende positiv semidefinite Lösung  $X_*$ , d.h.  $X_* \geq 0$  und  $\Lambda(A - B B^T X_*) \subset \mathbb{C}^-$ .*
- b)  *$\text{rank}(X_*) = k$ , wobei  $k$  die Anzahl instabiler Eigenwerte von  $A$  ist. Damit gilt  $X = Z Z^T$  mit  $Z \in \mathbb{R}^{n \times k}$ .*
- c)  *$\Lambda(A - B B^T X_*) = (\Lambda(A) \cap \mathbb{C}^-) \cup -(\Lambda(A) \cap \mathbb{C}^+)$ .*

Damit kann auch die ABE zur Stabilisierung von linearen, zeit-invarianten Systemen verwendet werden. Der in [?] angegebene Algorithmus zur Berechnung von  $X$  bzw.  $Z$  ist vom Aufwand her mit dem Bartels-Stewart Algorithmus zur Lösung von Lyapunovgleichungen vergleichbar. Die Stabilisierungseigenschaften sind oft besser als beim Bass-Algorithmus, da das Spektrum von  $A - B B^T X_*$  oft weniger sensitiv gegenüber Störungen ist. Allerdings beruht diese Aussage bisher nur auf empirischen Untersuchungen, eine genaue Störungsanalyse, die dies bestätigt, fehlt noch.

### 3.3 Numerische Lösung von Lyapunovgleichungen

Die Lösung des zur Lyapunovgleichung nach Folgerung 3.3 äquivalenten linearen Gleichungssystems

$$((A \otimes I_n) + (I_n \otimes A)) \operatorname{vec}(X) = \operatorname{vec}(W) \quad (3.11)$$

mit dem Gaußschen Eliminationsverfahren würde  $\frac{2}{3}(n^2)^3 = \frac{2}{3}n^6$  Flops kosten und auf einen Speicherbedarf für  $n^4$  reelle Zahlen hinauslaufen. Dies ist bereits für  $n = 100$  nicht mehr vertretbar. (Es würden bereits mehr als 800MB Hauptspeicher benötigt!) Unser Ziel ist es, einen Algorithmus mit Rechenaufwand  $\mathcal{O}(n^3)$  und Speicherbedarf  $\mathcal{O}(n^2)$  zu entwickeln. Dies gelingt, da (3.11) ein Gleichungssystem mit sehr spezieller Struktur in  $\mathbb{R}^{n^2 \times n^2}$  ist und durch  $\mathcal{O}(n^2)$  Daten definiert ist.

Die Idee ist, die Lyapunovgleichung (3.3) in eine Gestalt zu transformieren, so daß das Gleichungssystem (3.11) durch einfaches Rückwärtseinsetzen mit einem Aufwand von  $\mathcal{O}(n^3)$  gelöst werden kann.

Dazu verwenden wir ein Ergebnis aus der numerischen linearen Algebra.

**Satz 3.14 (Reelle Schur-Zerlegung)** *Sei  $A \in \mathbb{R}^{n \times n}$ . Dann existiert  $Q \in \mathbb{R}^{n \times n}$  orthogonal, so daß*

$$A = QTQ^T, \quad T = \begin{bmatrix} T_{11} & \dots & T_{1,r} \\ & \ddots & \vdots \\ & & T_{rr} \end{bmatrix} \quad \text{mit } T_{jj} \in \mathbb{R}^{n_j \times n_j}, n_j \in \{1, 2\}, \quad \sum_{j=1}^r n_j = n, \quad (3.12)$$

wobei für  $n_j = 1$ ,  $T_{jj}$  ein reeller Eigenwert von  $A$  ist und für  $n_j = 2$  sind die Eigenwerte von  $T_{jj}$  ein komplex-konjugiertes Eigenwertpaar von  $A$ .

Außerdem kann  $Q$  so gewählt werden, daß die Eigenwerte von  $A$  in beliebiger Reihenfolge auf der Diagonale von  $T$  erscheinen.

**Beweis:** Siehe z.B. [13, 29]. □

Die Zerlegung von  $A$  in (3.12) heißt *Schur-Zerlegung* von  $A$ ,  $T$  heißt dann *Schurform* von  $A$  und die Spalten von  $Q$  sind die *Schurvektoren* von  $A$ . Beachte, daß die ersten  $n_1 + \dots + n_k$  Spalten von  $Q$  einen  $A$ -invarianten Unterraum bilden für alle  $k = 1, \dots, r$ . Die Schur-Zerlegung von  $A$  kann z.B. mit dem QR Algorithmus numerisch rückwärts stabil berechnet werden, siehe z.B. [13, 29].

Multipliziert man die Lyapunovgleichung (3.3) von rechts mit der Schurvektor-Matrix  $Q$  von  $A$  und von links mit  $Q^T$ , so ergibt sich die äquivalente Lyapunovgleichung

$$\begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} + \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} A_1^T & 0 \\ A_2^T & A_3^T \end{bmatrix} = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix}$$

mit  $A_3 \in \mathbb{R}^{n_r \times n_r}$ . Unter Ausnutzung der Symmetrie zerfällt diese Lyapunovgleichung in die folgenden drei linearen Matrixgleichungen:

$$A_1 X_1 + X_1 A_1^T = W_1 - A_2 X_2^T - X_2 A_2^T =: \tilde{W}_1 = \tilde{W}_1^T, \quad (3.13)$$

$$A_1 X_2 + X_2 A_3^T = W_2 - A_2 X_3 =: \tilde{W}_2, \quad (3.14)$$

$$A_3 X_3 + X_3 A_3^T = W_3. \quad (3.15)$$



Die Gleichung (3.15) kann nun explizit aufgelöst werden. Falls  $n_r = 1$ , dann ist (3.15) skalar und

$$X_3 = \frac{W_3}{2A_3},$$

da  $A_3 \neq 0$  (andernfalls wäre die Lyapunovgleichung nicht eindeutig lösbar). Im Fall  $n_r = 2$  verwendet man explizit die vektorisierte Darstellung von (3.15). Unter Ausnutzung der Symmetrien ergibt sich das Gleichungssystem

$$\begin{bmatrix} a_{n-1,n-1} & a_{n-1,n} & 0 \\ a_{n,n-1} & a_{n-1,n-1} + a_{n,n} & a_{n-1,n} \\ 0 & a_{n,n-1} & a_{n,n} \end{bmatrix} \begin{bmatrix} x_{n-1,n-1} \\ x_{n-1,n} \\ x_{n,n} \end{bmatrix} = \begin{bmatrix} w_{n-1,n-1}/2 \\ w_{n-1,n} \\ w_{n,n}/2 \end{bmatrix}. \quad (3.16)$$

Die Lösung von (3.16) sollte mit einer LU Zerlegung mit vollständiger Pivotisierung erfolgen, um möglichst hohe Genauigkeit zu erzielen.

Die Lösung  $X_3$  wird nun in (3.14) eingesetzt. Damit ergibt sich eine Sylvestergleichung, die eindeutig lösbar ist (wegen Satz 3.5 und  $\Lambda(A_1) \cap \Lambda(-A_3^T) = \emptyset$ , falls  $\Lambda(A) \cap \Lambda(-A^T) = \emptyset$ ). Aufgrund der sehr speziellen Struktur dieser Sylvestergleichung kann man diese Lösung einfach durch Rückwärtseinsetzen berechnen. Partitioniere dazu analog zu (3.12),

$$A_1 = \begin{bmatrix} A_{11} & \cdots & A_{1,r-1} \\ & \ddots & \vdots \\ & & A_{r-1,r-1} \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_1 \\ \vdots \\ x_{r-1} \end{bmatrix}, \quad \tilde{W}_2 = \begin{bmatrix} w_1 \\ \vdots \\ w_{r-1} \end{bmatrix}, \quad x_j, w_j \in \mathbb{R}^{n_j \times n_r}.$$

Dann kann man die  $x_j$  rückwärts berechnen aus

$$A_{jj}x_j + x_j A_3^T = w_j - \sum_{i=j+1}^{r-1} A_{j,i}x_i =: \tilde{w}_j, \quad j = r-1, r-2, \dots, 1.$$

Bei der Lösung dieser Gleichung muß man vier Fälle unterscheiden.

$n_j = 1, n_r = 1$ : Die Gleichung ist skalar und  $x_j = \frac{\tilde{w}_j}{A_{jj} + A_3}$ .

$n_j = 2, n_r = 1$ : Man erhält ein Gleichungssystem im  $\mathbb{R}^2$  mit eindeutiger Lösung,

$$(A_{jj} + A_3 \cdot I_2)x_j = \tilde{w}_j.$$

$n_j = 1, n_r = 2$ : Man erhält ein Gleichungssystem im  $\mathbb{R}^2$  mit eindeutiger Lösung,

$$(A_{jj} \cdot I_2 + A_3)x_j^T = \tilde{w}_j^T.$$

$n_j = 2, n_r = 2$ : Man erhält ein Gleichungssystem im  $\mathbb{R}^4$  mit eindeutiger Lösung,

$$((I_2 \otimes A_{jj}) + (A_3^T \otimes I_2)) \text{vec}(x_j) = \text{vec}(\tilde{w}_j).$$

Die Gleichungssysteme in den letzten drei Fällen werden mit LU Zerlegung und vollständiger Pivotisierung gelöst.

Setzt man nun die so berechnete Lösung von (3.14) in (3.13), so erhält man eine Lyapunovgleichung in  $\mathbb{R}^{n-n_r \times n-n_r}$  mit Koeffizientenmatrix in Schurform. Damit läßt sich diese Gleichung wieder aufteilen wie in (3.13)–(3.15), so daß sich das oben beschriebene Vorgehen rekursiv anwenden läßt bis (3.13) eine Gleichung im  $\mathbb{R}^{n_1 \times n_1}$  ist, die direkt aufgelöst werden kann.

Zum Schluß muß man noch die berechnete Lösung in das ursprüngliche Koordinatensystem zurücktransformieren. Damit erhält man insgesamt den Algorithmus 3.15.

---

**Algorithmus 3.15 (Bartels-Stewart Algorithmus (1972))**


---

INPUT:  $A, W \in \mathbb{R}^{n \times n}$ ,  $W = W^T$ .

OUTPUT: Lösung  $X = X^T$  der Lyapunovgleichung (3.3).

- 1: Berechne die reelle Schurform von  $A$  wie in (3.12) mit Hilfe des QR Algorithmus.
  - 2: **if**  $\Lambda(A) \cap \Lambda(-A) \neq \emptyset$  **then**
  - 3:   STOP; keine eindeutige Lösung.
  - 4: **end if**
  - 5:  $\begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} := Q^T W Q$ ,  $W_3 \in \mathbb{R}^{n_r \times n_r}$ .
  - 6:  $k := r$
  - 7: **while**  $k > 1$  **do**
  - 8:   Löse (3.15) mit  $A_3 = A_{kk}$ .
  - 9:   Löse (3.14) mit  $A_1 = \begin{bmatrix} A_{11} & \dots & A_{1,k-1} \\ & \ddots & \vdots \\ & & A_{k-1,k-1} \end{bmatrix}$ .
  - 10:  $W_1 := W_1 - A_2 X_2^T - X_2 A_2^T$ .
  - 11:  $k := k - 1$
  - 12: **end while**
  - 13: Löse (3.13) als lineares Gleichungssystem (unter Ausnutzung der Symmetrie) im  $\mathbb{R}^3$ .
  - 14:  $X := Q \tilde{X} Q^T$
- 

Der Bartels-Stewart Algorithmus benötigt ca.  $32n^3$  elementare Rechenoperationen, dabei entfallen auf den QR Algorithmus ca.  $25n^3$  Flops und auf die Schritte 5 und 14 je  $3n^3$ . Der gesamte Prozeß des Rückwärtseinsetzens (die **while**-Schleife) benötigt nur  $n^3$  Operationen. Da sowohl der QR Algorithmus als auch das Rückwärtseinsetzen numerisch rückwärts stabil sind und darüberhinaus nur orthogonale Ähnlichkeitstransformationen verwendet werden, kann der Bartels-Stewart Algorithmus als numerisch rückwärts stabil angesehen werden.

Der Bartels-Stewart Algorithmus zur Lösung von Lyapunovgleichungen ist z.B. in der MATLAB Control Toolbox Funktion `lyap` implementiert.

### 3.4 Stabilisierung durch Polvorgabe

Zunächst werden wir zeigen, daß ein System steuerbar genau dann ist, wenn es zu jeder Menge  $\mathcal{L} := \{\mu_1, \dots, \mu_n\} \subset \mathbb{C}$ , die abgeschlossen bzgl. komplexer Konjugation ist, eine Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$  mit  $\Lambda(A + BF) = \mathcal{L}$  gibt. Dazu führen wir zunächst zwei Normalformen für Single-Input Systeme ein, die weitere Charakterisierungen der Steuerbarkeit erlauben. Man beachte, daß sich diese Normalformen nicht numerisch stabil berechnen lassen und daher eher von theoretischem Interesse sind.

Im Folgenden sei

$$\phi_A(x) = x^n + \alpha_{n-1}x^{n-1} + \dots + \alpha_1x + \alpha_0 \quad (3.17)$$

das charakteristische Polynom von  $A$ . Desweiteren sagen wir, daß  $(A, B)$  und  $(\tilde{A}, \tilde{B})$  *ähnlich* sind, wenn  $(\tilde{A}, \tilde{B})$  durch einen Basiswechsel im Zustandsraum gemäß Definition 2.25 aus  $(A, B)$  hervorgeht, d.h. falls eine invertierbare Matrix  $T \in \mathbb{R}^{n \times n}$  existiert, so daß  $(\tilde{A}, \tilde{B}) = (TAT^{-1}, TB)$ .

**Lemma 3.16** *Es sei  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$  und*

$$A_s = \begin{bmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & \dots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -\alpha_{n-1} \end{bmatrix}, \quad B_s = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (3.18)$$

Dann gilt

$$AK(A, B) = \mathcal{K}(A, B)A_s, \quad B = \mathcal{K}(A, B)B_s, \quad (3.19)$$

wobei  $\mathcal{K}(A, B)$  die Steuerbarkeitsmatrix von  $(A, B)$  ist. Insbesondere ist  $(A, B)$  steuerbar genau dann wenn  $(A, B)$  und  $(A_s, B_s)$  ähnlich sind.

**Beweis:** Mit dem Satz von Cayley-Hamilton folgt (siehe auch Beweis von Satz 2.11)

$$A^n = - \sum_{j=0}^{n-1} \alpha_j A^j.$$

Also gilt

$$\begin{aligned} AK(A, B) &= \begin{bmatrix} AB & A^2B & \dots & A^{n-1}B & - \sum_{j=0}^{n-1} \alpha_j A^j B \end{bmatrix} \\ &= \mathcal{K}(A, B)A_s. \end{aligned}$$

Die zweite Gleichung in (3.19) ist sofort klar.

Nun ist  $(A_s, B_s)$  in System-Hessenberg-Form mit  $A_s$  unreduziert und damit ist  $(A_s, B_s)$  steuerbar nach Folgerung 2.30. Ist  $(A_s, B_s)$  ähnlich zu  $(A, B)$ , dann ist  $(A, B)$  steuerbar nach Satz 2.26. Ist andererseits  $(A, B)$  steuerbar, dann ist  $\mathcal{K}(A, B)$  regulär nach Satz 2.12 und somit sind wegen (3.19)  $(A, B)$  und  $(A_s, B_s)$  ähnlich mit  $T := \mathcal{K}(A, B)^{-1}$ .  $\square$

**Definition 3.17** Die Regler-Normalform eines Paares  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$  ist gegeben durch

$$A_{RNF} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix}, \quad B_{RNF} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \quad (3.20)$$

Man beachte, daß nicht jedes Paar  $(A, B)$  ähnlich zu seiner Regler-Normalform ist, wie der nächste Satz zeigt.

**Satz 3.18**  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$  ist steuerbar genau dann, wenn es eine reguläre Matrix  $S \in \mathbb{R}^{n \times n}$  gibt, so daß  $(A_{RNF}, B_{RNF}) = (SAS^{-1}, SB)$ .

**Beweis:** Zunächst beachte, daß  $\phi_A$  das charakteristische Polynom von  $A_{RNF}$  ist. (Entwickle  $\det(A_{RNF} - \lambda I_n)$  nach der letzten Zeile!) Damit gilt nach Lemma 3.16 bzw. (3.18)

$$A_{RNF} \mathcal{K}(A_{RNF}, B_{RNF}) = \mathcal{K}(A_{RNF}, B_{RNF}) A_s, \quad B_{RNF} = \mathcal{K}(A_{RNF}, B_{RNF}) B_s.$$

Nun ist

$$\mathcal{K}(A_{RNF}, B_{RNF}) = \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & \ddots & \ddots & * \\ 0 & \ddots & \ddots & \vdots \\ 1 & * & \dots & * \end{bmatrix}$$

und damit invertierbar. Also ist  $(A_{RNF}, B_{RNF})$  ähnlich zu  $(A_s, B_s)$ . Damit folgt die Behauptung mit Lemma 3.16.  $\square$

Desweiteren werden wir noch eine Eigenschaft des Teilraums  $\text{range}(\mathcal{K}(A, B))$  benötigen.

**Lemma 3.19** Sei  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  und  $b_j = Be_j$ ,  $j = 1, \dots, m$ . Dann gilt

$$\text{range}(\mathcal{K}(A, B)) = \text{span}\{A^k b_j \mid k \in \mathbb{N}_0, j = 1, \dots, m\} =: \mathcal{K}.$$

D.h.,  $\text{range}(\mathcal{K}(A, B))$  ist der kleinste  $A$ -invariante Unterraum, der  $\text{range}(B)$  enthält.

**Beweis:** Die Aussage folgt mit dem Satz von Cayley-Hamilton wie im Beweis von Satz 2.11, da

$$A^{n+\nu} = \sum_{j=0}^{n-1} \beta_j^{(\nu)} A^j \quad \forall \nu \in \mathbb{N}_0.$$

Da jeder  $A$ -invariante Unterraum, der  $\text{range}(B)$  enthält, auch  $\mathcal{K}$  enthalten muß, ergibt sich die Interpretation von  $\text{range}(\mathcal{K}(A, B))$  als kleinster  $A$ -invarianter Unterraum der  $\text{range}(B)$  enthält.  $\square$

Beachtet man noch, daß für ähnliche Matrixpaare  $(A, B)$  und  $(\tilde{A}, \tilde{B}) = (TAT^{-1}, TB)$  gilt  $\Lambda(A + BF) = \Lambda(\tilde{A} + \tilde{B}\tilde{F}) = \mathcal{L}$  für  $\tilde{F} = FT^{-1}$ , so haben wir nun die notwendigen Vorarbeiten geleistet, um den Satz von der Polvorgabe zu beweisen. Der Beweis folgt im Wesentlichen dem Beweis von Theorem 13 in [28].

**Satz 3.20 (Polvorgabe)** *Es sei  $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$  und die unsteuerbaren Moden von  $(A, B)$  seien  $\{\lambda_{k+1}, \dots, \lambda_n\}$ . Dann existiert eine Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$  so, daß  $\Lambda(A + BF) = \mathcal{L}$  genau dann, wenn  $\mathcal{L} = \{\mu_1, \dots, \mu_k, \lambda_{k+1}, \dots, \lambda_n\}$ , wobei  $\{\mu_1, \dots, \mu_k\} \subset \mathbb{C}$  beliebig gewählt werden kann, solange  $\{\mu_1, \dots, \mu_k\}$  abgeschlossen unter komplexer Konjugation ist. Im Fall  $m = 1$  ist  $F$  eindeutig.*

*Insbesondere ist  $(A, B)$  steuerbar genau dann, wenn es für jede Menge  $\mathcal{L} = \{\mu_1, \dots, \mu_n\}$  mit  $\mathcal{L} = \overline{\mathcal{L}}$  ein  $F \in \mathbb{R}^{m \times n}$  mit  $\Lambda(A + BF) = \mathcal{L}$  gibt.*

**Beweis:** Sei zunächst  $(A, B)$  nicht steuerbar. O.B.d.A. können wir annehmen, daß  $(A, B)$  in Kalman-Form vorliegt, d.h.

$$A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad (A_1, B_1) \text{ steuerbar,}$$

wobei  $A_1 \in \mathbb{R}^{k \times k}$  usw. Für ein beliebiges  $F = \begin{bmatrix} F_1 & F_2 \end{bmatrix}$  gilt also

$$\Lambda(A + BF) = \Lambda(A_1 + B_1 F_1) \cup \Lambda(A_3) = \Lambda(A_1 + B_1 F_1) \cup \{\lambda_{k+1}, \dots, \lambda_n\}.$$

Da  $\Lambda(A_3)$  die unkontrollierbaren Moden sind, ist klar, daß  $\mathcal{L}$  die im Satz angegebene Form annehmen muß und daß  $(A, B)$  steuerbar ist, genau dann wenn  $\mathcal{L}$  beliebig gewählt werden kann.

Es bleibt jetzt zu zeigen, daß  $\Lambda(A_1 + B_1 F_1)$  beliebig gewählt werden kann durch Wahl eines geeigneten  $F$ . Kann man für eine vorgegebene Menge  $\mathcal{L}_1 := \{\mu_1, \dots, \mu_k\}$  ein  $F_1 \in \mathbb{R}^{m \times k}$  finden mit  $\Lambda(A_1 + B_1 F_1) = \mathcal{L}_1$ , dann liefert  $F = \begin{bmatrix} F_1 & 0 \end{bmatrix}$  das Gewünschte.

Nach diesen Vorüberlegungen können wir uns also auf den Fall, daß  $(A, B)$  steuerbar ist, zurückziehen. Im Fall  $m = 1$  können wir wegen Satz 3.18 davon ausgehen, daß  $(A, B)$  in Regler-Normalform vorliegt. Dann sieht man sofort, daß mit  $F = \begin{bmatrix} f_1 & \dots & f_n \end{bmatrix}$ ,  $f_j \in \mathbb{R}$

$$A + BF = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 + f_1 & -\alpha_1 + f_2 & -\alpha_2 + f_3 & \dots & -\alpha_{n-1} + f_n \end{bmatrix}.$$

Da

$$\phi_{A+BF}(x) = x^n + (\alpha_{n-1} - f_n)x^{n-1} + \dots + (\alpha_1 - f_2)x + (\alpha_0 - f_1),$$

ist  $f$  eindeutig festgelegt durch  $f_j = \alpha_{j-1} - \beta_{j-1}$ , wobei

$$(x - \mu_1) \cdots (x - \mu_n) =: x^n + \beta_{n-1}x^{n-1} + \dots + \beta_1x + \beta_0.$$

Sei nun  $m$  beliebig. Wir werden dies auf den Fall  $m = 1$  zurückführen. Dazu sei  $v \in \mathbb{R}^m$  mit  $b := Bv \neq 0$ . Wir zeigen nun, daß ein  $G \in \mathbb{R}^{m \times n}$  existiert, so daß  $(A + BG, b)$  steuerbar ist. Ist dann  $f \in \mathbb{R}^{1 \times n}$  der eindeutig bestimmte Vektor so daß  $\Lambda(A + BG + bf) = \mathcal{L}$ , dann ist  $F := G + vf$  die gesuchte Feedback-Matrix.

Es bleibt also die Existenz von  $G$  zu zeigen. Definiere nun eine maximale Menge linear unabhängiger Vektoren  $\mathcal{R} = \{x_1, \dots, x_\ell\} \subset \mathbb{R}^n$  mit

$$\begin{aligned} x_1 &:= b = Bv, \\ x_j - Ax_{j-1} &\in \text{range}(B), \end{aligned} \quad (3.21)$$

wobei wir vereinbaren, daß  $x_0 = 0$ . Es ist klar, daß  $\mathcal{R} \neq \emptyset$ , da  $b \in \mathcal{R}$  und  $\dim \text{span} \mathcal{R} = \ell \leq n$  und damit “ $\ell$  maximal” wohldefiniert ist. Beachte, daß (3.21) auch wie folgt formuliert werden kann:

$$x_j = Ax_{j-1} + Bu \text{ für ein } u \in \mathbb{R}^m. \quad (3.22)$$

Wir zeigen jetzt, daß  $\ell = n$ . Da  $\ell$  maximal gewählt ist, folgt

$$Ax_\ell + Bu \in \text{span} \{x_1, \dots, x_\ell\} \quad \forall u \in \mathbb{R}^m. \quad (3.23)$$

(Sonst erhält man mit  $x_{\ell+1} := Ax_\ell + Bu$  eine größere Menge, die die Bedingungen erfüllt, was der Maximalität von  $\ell$  widerspricht.)

Insbesondere ergibt sich mit  $u = 0$ , daß

$$Ax_\ell \in \text{span} \{x_1, \dots, x_\ell\}.$$

Damit folgt aus (3.23)

$$\text{range}(B) \subset \text{span} \{x_1, \dots, x_\ell\} - Ax_\ell = \text{span} \{x_1, \dots, x_\ell\}$$

und mit (3.21)

$$Ax_k \in \text{span} \{x_1, \dots, x_\ell\}, \quad k = 1, \dots, \ell.$$

Damit ist  $\text{span} \{x_1, \dots, x_\ell\}$  ein  $A$ -invarianter Unterraum, der  $\text{range}(B)$  enthält. Nun ist aber nach Lemma 3.19  $\text{range}(\mathcal{K}(A, B))$  der kleinste Teilraum von  $\mathbb{R}^n$ , der diese Eigenschaft erfüllt. Also gilt mit der Steuerbarkeit von  $(A, B)$

$$n = \dim \text{range}(\mathcal{K}(A, B)) \leq \dim \text{span} \{x_1, \dots, x_\ell\}$$

und damit  $\dim \text{span} \{x_1, \dots, x_\ell\} = \ell = n$ .

Sei nun  $u_k \in \mathbb{R}^m$ ,  $k = 1, \dots, n-1$ , eine Folge von Vektoren, die  $\mathcal{R}$  wie in (3.22) erzeugt, d.h.

$$x_k - Ax_{k-1} = Bu_{k-1}, \quad k = 2, \dots, n.$$

Wählt man noch  $u_n \in \mathbb{R}^m$  beliebig und definiert

$$X := [x_1, \dots, x_n] \in \mathbb{R}^{n \times n}, \quad U := [u_1, \dots, u_n] \in \mathbb{R}^{m \times n},$$

so ist  $G := UX^{-1}$  wohldefiniert, da  $X$  regulär wegen der linearen Unabhängigkeit der  $x_j$  ist.

Die Matrix  $G$  erfüllt

$$Gx_k := u_k, \quad k = 1, \dots, n,$$

und mit (3.21) erhält man damit

$$\mathcal{K}(A + BG, b) = [x_1 \quad \dots \quad x_n].$$

Also ist  $(A + BG, b)$  steuerbar und die Behauptung gezeigt.  $\square$

Der Beweis von Satz 3.18 motiviert den Begriff *Regler-Normalform*, da die den Regler bestimmende Matrix  $F$  im Fall  $m = 1$  direkt aus  $A_{RNF}$  abgelesen werden kann.

**Bemerkung 3.21** *Bei der Berechnung der Feedback-Matrix  $F$  muß man im Fall  $m > 1$  die Eindeutigkeit durch Zusatzbedingungen erzwingen. Diese Freiheit sollte man ausnutzen, die Pole des rückgekoppelten Systems so robust wie möglich gegenüber Störungen zu machen. Ein für die numerische Rechnung geeignetes Kriterium ist,  $F$  so zu wählen, daß  $A + BF$  diagonalisierbar und die Eigenvektormatrix  $X = X(F)$  minimale Kondition hat. Es ergibt sich das Minimierungsproblem*

$$\min_{F \in \mathbb{R}^{m \times n}} \text{cond}(X(F))$$

so daß  $(A + BF)X(F) = X(F) \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{bmatrix}$ .

Das Problem der robusten Polvorgabe wurde 1985 von Kautsky, Nichols und Van Dooren formuliert und gelöst [15], wobei neben obigem Kriterium noch drei weitere untersucht werden, mit deren Hilfe die Sensitivität der Pole gegenüber Störungen gemessen werden kann.

Hier wollen wir nur einen Algorithmus für den Fall  $m = 1$  besprechen. Da wir von einem steuerbaren System ausgehen, können wir nach Folgerung 2.30 annehmen, daß  $(A, B)$  in System-Hessenbergform vorliegt und  $A$  unreduzierte Hessenberg-Matrix ist.

Der Algorithmus geht auf Miminis und Paige (1982) zurück [23] und verwendet im Prinzip den QR Algorithmus zur Berechnung der Eigenwerte einer nicht-symmetrischen Matrix (siehe z.B. [13, Kapitel 7]). Beachte, daß es sich hier um ein *inverses Eigenwertproblem* handelt: Gegeben ist das Spektrum und zu berechnen ist eine Matrix in Hessenbergform mit diesem Spektrum. Anstelle des QR Schritts im QR Algorithmus wird hier ein RQ Schritt verwendet. Die Idee ist dabei die Folgende.

Wir verwenden die RQ Zerlegung einer Matrix  $M \in \mathbb{R}^{n \times n}$ :

$$M = RQ = \begin{bmatrix} \diagdown & & \\ & \ddots & \\ & & \end{bmatrix} Q, \quad Q^T Q = I_n,$$

deren Existenz aus der Existenz der QR Zerlegung folgt. Weiterhin benötigen wir eine Aussage, welche in ähnlicher Form beim QR Algorithmus die Wahl von Shifts motiviert: Ist  $\lambda \in \Lambda(M)$ ,  $M$  eine unreduzierte Hessenbergmatrix und  $M - \lambda I = RQ$  eine RQ Zerlegung, dann ist  $R_{jj} \neq 0$ ,  $j = 2, \dots, n$  und  $r_{11} = 0$ . Dies folgt, da  $M$  unreduziert ist und deshalb  $\text{rank}(M - \lambda I) \geq n - 1$ , andererseits  $\text{rank}(M - \lambda I) \leq n - 1$  (da  $\lambda \in \Lambda(M)$ ). Da nun

$Q$  obere Hessenberg-Matrix ist und  $M$  unreduziert, folgt aus  $m_{j,j-1} = r_{jj}q_{j,j-1} \neq 0$ , daß  $r_{jj} \neq 0$  für  $j = 2, \dots, n$  und aus der Rangbedingung damit  $r_{11} = 0$ .

Wählt man nun  $F \in \mathbb{R}^{1 \times n}$  so, daß in der RQ Zerlegung von

$$A + BF - \mu_1 I = A + \beta e_1 F - \mu_1 I$$

$r_{11} = 0$  gilt, dann folgt

$$QR + \mu_1 I = Q(A + BF - \mu_1 I)Q^T + \mu_1 I = Q(A + BF)Q^T = \begin{bmatrix} \mu_1 & * & \dots & * \\ & * & \dots & * \\ & & \ddots & \vdots \\ & & & * \dots * \end{bmatrix}.$$

Also ist  $\mu_1 \in \Lambda(A + BF)$ .

Die Matrix  $F$  läßt sich damit rekursiv berechnen. Es sei

$$A + \beta e_1 F - \mu_1 I = RQ, \quad Q =: \begin{bmatrix} y^T \\ \tilde{Q} \end{bmatrix}. \quad (3.24)$$

Dann folgt  $(A + \beta e_1 F - \mu_1 I)y = 0$  und

$$e_1^T (A - \mu_1 I)y = -\beta Fy \quad (3.25)$$

Damit hat man eine Gleichung für die  $n$  unbekanntten Einträge von  $F$ . Allerdings kann man  $y$  nicht wie in (3.24) berechnen, da  $F$  ja noch unbekannt ist. Beachte aber, daß

$$\begin{aligned} A - \mu_1 I &= RQ - \beta e_1 F = RQ - \beta e_1 F Q^T Q \\ &= \underbrace{(R - \beta e_1 F Q^T)}_{=: \tilde{R}} Q. \end{aligned}$$

Es ist also  $\tilde{R}Q$  eine RQ Zerlegung von  $A - \mu_1 I$  und wir erhalten

$$\tilde{r}_{11} = r_{11} - \beta Fy = -\beta Fy.$$

Damit können wir die rechte Seite von (3.25) berechnen, ohne  $F$  zu kennen.

Da  $Q$  aus der RQ Zerlegung von  $A - \mu_1 I$  bekannt ist, können wir auch

$$Q(A + \beta e_1 F)Q^T = QR + \mu_1 I = \begin{bmatrix} \mu_1 & * \\ 0 & \tilde{Q}(A + \beta e_1 F)\tilde{Q}^T \end{bmatrix}$$

berechnen, ohne  $F$  zu kennen. Setze nun

$$\tilde{A} := \tilde{Q}A\tilde{Q}^T \in \mathbb{R}^{n-1 \times n-1}, \quad \tilde{\beta} = e_2^T \beta Q e_1 = q_{21} \beta, \quad \tilde{F} := F\tilde{Q}^T.$$

Aus  $q_{21} \neq 0$  (da  $A$  unreduziert) folgt  $\tilde{\beta} \neq 0$  und damit ist  $(\tilde{A}, \tilde{B}) := (\tilde{A}, \tilde{\beta} e_1)$  in System-Hessenbergform, wobei  $\tilde{A}$  unreduziert ist. Damit können wir den RQ Schritt mit  $(\tilde{A}, \tilde{B})$



**Algorithmus 3.22 (Polvorgabe nach Miminis-Paige (1982))**


---

 INPUT:  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^n$ ,  $\mathcal{L} = \{\mu_1, \dots, \mu_n\}$  mit  $(A, B)$  steuerbar,  $\mathcal{L} = \overline{\mathcal{L}}$ .
 

---

 OUTPUT: Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$  mit  $\Lambda(A + BF) = \mathcal{L}$ .

- 1: Berechne die System-Hessenbergform von
- $(A, B)$
- , d.h. berechne
- $P \in \mathbb{R}^{n \times n}$
- orthogonal so, daß

$$A_1 := P^T A P = \begin{bmatrix} \square & & \\ & \square & \\ & & \square \end{bmatrix}, \quad \beta_1 e_1 := P^T B,$$

 wobei  $A_1$  unreduzierte Hessenbergmatrix ist.

- 2:
- for**
- $k = 1, \dots, n - 1$
- do**
- 
- 3: Berechne die RQ Zerlegung

$$(A_k - \mu_k I_{n-k+1}) = R_k Q_k =: [r_{ij}^{(k)}]_{i,j=1}^n \cdot [q_{ij}^{(k)}]_{i,j=1}^n.$$

- 4: Partitioniere
- $Q_k =: \begin{bmatrix} y_k^T \\ \tilde{Q}_k \end{bmatrix}$
- ,
- $y_k \in \mathbb{R}^{n-k+1}$
- .

- 5: Setze
- $\tau_k := -\frac{r_{11}^{(k)}}{\beta_k}$
- (
- $\equiv -F_k y_k$
- ).

- 6: Berechne
- $A_{k+1} := \tilde{Q}_k A_k \tilde{Q}_k^T$
- ,
- $\beta_{k+1} := \beta_k q_{21}^{(k)}$
- .

- 7:
- end for**

- 8:
- $F_n := -\frac{A_n - \mu_n}{\beta_n}$
- .

- 9:
- for**
- $k = n - 1, n - 2, \dots, 1$
- do**

- 10:
- $F_k := \begin{bmatrix} \tau_k & F_{k+1} \end{bmatrix} Q_k$
- .

- 11:
- end for**

- 12:
- $F := F_1 P$
- .
- 

und  $\mu_2$  wiederholen und erhalten wiederum eine Gleichung der Form (3.25). Wiederholt man den RQ Schritt  $n - 1$  mal, bis man bei  $\mu_n$  und  $(\tilde{A}, \tilde{B}) \in \mathbb{R}^{1 \times 1} \times \mathbb{R}^{1 \times 1}$  angekommen ist, dann kann man die Gleichung (3.25) nach  $\tilde{F}$  auflösen und daraus dann rückwärts den gesamten Zeilenvektor  $F$  bestimmen.

Insgesamt ergibt sich damit der Algorithmus 3.22.

Das Vorgehen bei diesem Algorithmus erläutern wir an einem kleinen Beispiel.

**Beispiel 3.23** Seien

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathcal{L} = \{-1, -2\}.$$

Es ist  $P = I_2$ , da  $(A, B)$  bereits in System-Hessenbergform vorliegt, und daher  $A_1 = A$ ,  $\beta_1 = 1$ .

Berechne nun die RQ Zerlegung

$$A_1 - \mu_1 I_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = R_1 Q_1 = \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \right) \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \right).$$

Damit ergeben sich

$$\tau_1 = 0, \quad \beta_2 = \frac{1}{\sqrt{2}}, \quad \tilde{Q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \tilde{A}_2 = \tilde{Q}_1 A_1 \tilde{Q}_1^T = 1.$$

Die Feedback-Matrix  $F$  erhalten wir nun durch

$$F_2 = -\frac{A_2 - \mu_2}{\beta_2} = \frac{1 - (-2)}{\frac{1}{\sqrt{2}}} = -3\sqrt{2},$$

$$F_1 = \begin{bmatrix} \tau_1 & F_2 \end{bmatrix} Q_1 = \begin{bmatrix} 0 & -3\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = -\begin{bmatrix} 3 & 3 \end{bmatrix} = F.$$

Also ist  $A + BF = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 3 & 3 \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}$  und  $\Lambda(A + BF) = \mathcal{L}$ .

**Bemerkung 3.24** Algorithmus 3.22 verwendet im Wesentlichen numerisch stabile Operationen. Allerdings kann es bei der Berechnung von  $G_n$  zu Auslöschungseffekten kommen.

Prinzipiell ist das Polvorgabe-Problem ein inverses Problem. Solche Probleme sind meist inhärent schlecht konditioniert, so daß man potentiell mit großer Sensitivität der Pole gegenüber kleinen Störungen rechnen muß. In der Praxis kann dies katastrophale Folgen haben, wenn der Algorithmus z.B. zur Stabilisierung eingesetzt wird. Hier können kleine Störungen  $\delta F$  in der Feedback-Matrix dazu führen, daß einer oder mehrere der Eigenwerte  $\Lambda(A + B(F + \delta F))$  in der rechten Halbebene liegen. Damit erreicht man dann keine Stabilität des geschlossenen Regelkreises. In der Praxis können solche Systeme durch äußere Einflüsse leicht "aus der Bahn" geworfen werden.

# Kapitel 4

## Optimale Steuerung

Für ein allgemeines nichtlineares System wie in Definition 1.2, also

$$\begin{aligned}\dot{x} &= f(t, x, u), & x(t_0) &= x^0, & t &\in [t_0, t_f] \\ y &= g(t, x, u)\end{aligned}\tag{4.1}$$

suchen wir eine *optimale Steuerung*  $u \in \mathcal{U}_{ad}$ , die das Kostenfunktional  $\mathcal{J} : \mathcal{U}_{ad} \rightarrow \mathbb{R}$ ,

$$\mathcal{J}(u) = h_f(x(t_f)) + \int_{t_0}^{t_f} h(t, x, y, u) dt,\tag{4.2}$$

minimiert. Dabei ist  $h$  eine geeignet zu wählende Funktion, die den Zustands-, Ausgangs- und Eingangsgrößen Kosten zuweist, während  $h_f$  die Abweichung vom gewünschten Endzustand  $\hat{x}$  bewertet.

Um auch hier zuzulassen, daß wir mit einem asymptotischen Erreichen des Ziels zufrieden sind, also als Aufgabenstellung eine Stabilisierung des Systems vorgeben, wird  $t_f = \infty$  zugelassen. Zunächst betrachten wir jedoch den Fall  $t_f < \infty$ , der Fall eines unendlichen Zeithorizonts ergibt sich dann aus einer asymptotischen Betrachtung.

**Bemerkung 4.1** *Wählt man  $h \equiv 1$ ,  $h_f \equiv 0$ , so ergibt sich das Problem,  $t_f$  zu minimieren. Diese Aufgabe nennt man zeitoptimale Steuerung, es gilt also, die Steuerung zu finden, die das System in kürzest möglicher Zeit in den gewünschten Zustand bringt. Diese Aufgabenstellung wird hier nicht weiter betrachtet.*

Der allgemeine Zugang zu Problemen der optimalen Steuerung beruht auf dem Lagrange-Formalismus: Möchte man ein Problem der Form

$$\begin{aligned}\min_{x \in \mathbb{R}^n} & h(x) \\ \text{so daß} & f(x) = 0\end{aligned}$$

lösen, so definiert man die *Lagrange-Funktion*

$$\mathcal{L}(x, \lambda) = -h(x) + \lambda^T f(x),$$

mit den *Lagrange-Multiplikatoren*  $\lambda \in \mathbb{R}^m$  und entwickelt notwendige Optimalitätsbedingungen aus

$$\mathcal{L}_x = 0, \quad \mathcal{L}_\lambda = 0.$$

Analog verwendet man bei dynamischen Nebenbedingungen das *Hamilton-Prinzip*. Dazu definiert man in Analogie zu (4.1) eine *Hamilton-Funktion*. Für autonome Systeme mit  $h(x, y, u) \equiv h(x, u)$  ist diese definiert durch

$$\mathcal{H}(x, u, \mu) = -h(x, u) + \mu^T f(x, u), \quad (4.3)$$

wobei  $\mu : [t_0, t_f] \rightarrow \mathbb{R}^n$  die den Lagrange-Multiplikatoren entsprechende *Kozustandsfunktion* bezeichnet. Beachte, daß die dynamischen Nebenbedingungen damit als

$$\mathcal{H}_\mu = \dot{x}$$

geschrieben werden können.

Die notwendigen Optimalitätsbedingungen ergeben sich dann aus dem folgenden Satz, der auf Pontryagin [26] zurückgeht und hier in der Form eines “Minimumsprinzips” wie in [25] angegeben ist, wobei zunächst  $h_f \equiv 0$  vorausgesetzt ist.

**Satz 4.2 (Pontryagin’sches Maximumprinzip)** *Sei  $u_* \in \mathcal{U}_{ad}$  und  $x_*(t) := x(t; u_*)$  die zugehörige Lösungstrajektorie von (4.1). Ist  $u_*$  optimal für (4.2), so erfüllt  $u_*$  die notwendigen Optimalitätsbedingungen*

$$(i) \quad \mathcal{H}(x_*, u_*, \mu) = \inf_{u \in \mathcal{U}_{ad}} \mathcal{H}(x, u, \mu) \text{ für alle } t \in [t_0, t_f].$$

(ii) Die Kozustandsfunktion erfüllt die adjungierte Gleichung

$$\dot{\mu}(t) = -\mathcal{H}_x, \quad \text{d.h. } \dot{\mu}_j(t) = -\frac{\partial \mathcal{H}}{\partial x_j}, \quad j = 1, \dots, n.$$

(iii)  $\mu(t_f) = 0$  (Transversalitätsbedingung).

**Beweis:** Siehe [22, 25, 26]. □

Der Fall  $h_f \neq 0$  kann auf obigen Satz zurückgeführt werden. Dazu verwendet man, daß

$$h_f(x(t_f)) - h_f(x(t_0)) = \int_{t_0}^{t_f} \nabla h_f(x(t)) \cdot \dot{x}(t) dt = \int_{t_0}^{t_f} \nabla h_f(x(t)) \cdot f(x(t), u(t)) dt,$$

woraus sich eine neue Kostenfunktion

$$\tilde{h}(x, u) = h(x, u) + \nabla h_f(x) \cdot f(x, u)$$

ergibt. Damit ist also  $\mathcal{J}(u) = h_f(x(t_0)) + \int_{t_0}^{t_f} \tilde{h}(x, u) dt$ . Da  $h_f(x(t_0))$  konstant ist, kann dieser Term in der Optimierung vernachlässigt werden, d.h. man arbeitet dann mit dem modifizierten Kostenfunktional

$$\tilde{\mathcal{J}}(u) = \int_{t_0}^{t_f} \tilde{h}(x, u) dt.$$

Ersetzt man nun  $h$  durch  $\tilde{h}$  in der Hamilton-Funktion, so kann man Satz 4.2 auf  $\tilde{\mathcal{J}}$  anwenden. Zu beachten ist dabei lediglich, daß sich die Transversalitätsbedingung zu

$$(iii)' \quad \mu(t_f) = \nabla h_f(x(t_f))$$

ändert. Außerdem sind natürlich die erforderlichen Glattheitsanforderungen an  $h_f$  zu überprüfen.

Im Folgenden betrachten wir wieder LTI Systeme wie in (1.4)–(1.5), d.h.

$$\begin{aligned} \dot{x} &= Ax + Bu, & x(0) &= x^0, & t &\in [0, t_f]. \\ y &= Cx, \end{aligned} \quad (4.4)$$

Hierbei ist also

$$\begin{aligned} f(t, x, u) &= Ax + Bu, \\ g(t, x, u) &= Cx. \end{aligned}$$

Da  $y = Cx$  und damit  $h(t, x, y, u) = h(t, x, Cx, u) \equiv h(t, x, u)$ , werden wir hier o.B.d.A.  $C = I_n$  annehmen. Setzen wir dann

$$\begin{aligned} h(t, x, u) &= \frac{1}{2} \begin{bmatrix} x(t)^T & u(t)^T \end{bmatrix} \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \\ h_f(x) &= \frac{1}{2} x^T M x, \end{aligned}$$

so erhalten wir ein quadratisches Kostenfunktional  $\mathcal{J}$  und damit die folgende Aufgabenstellung.

### Definition 4.3 (Linear-quadratisches Optimalsteuerungsproblem)

*Das Minimierungsproblem*

$$\begin{aligned} \min \mathcal{J}(u) &= \frac{1}{2} \left( x(t_f)^T M x(t_f) + \int_0^{t_f} \{ x(t)^T Q x(t) + 2x(t)^T S u(t) + u(t)^T R u(t) \} dt \right) \\ &\text{unter der Nebenbedingung} \end{aligned} \quad (4.5)$$

$$\dot{x} = Ax + Bu, \quad x(0) = x^0,$$

heißt linear-quadratisches Optimalsteuerungsproblem.

In der Regelungstheorie nennt man das linear-quadratische Optimalsteuerungsproblem auch *linear-quadratisches Regelungsproblem* oder kurz *LQR Problem*. Im Folgenden werden wir diese Abkürzung ebenfalls verwenden.

**Bemerkung 4.4** *Wie man hier sofort sieht, ist die oben getroffene Vereinbarung  $C = I$  tatsächlich keine Einschränkung. Falls die Ausgänge im Kostenfunktional gewichtet werden sollen, hätte man bei einem quadratischen Zielfunktional Terme der Form*

$$y^T Q_y y + x^T Q_x x = x^T C^T Q_y C x + x^T Q_x x = x^T (C^T Q_y C + Q_x) x$$

und würde mit  $Q := C^T Q_y C + Q_x$  also genau ein Kostenfunktional der Form wie in (4.5) erhalten.

Das LQR Kostenfunktional bewertet also folgende Größen:

- Die Abweichung des Endzustandes  $x(t_f)$  vom Ziel  $\hat{x}$  mit Hilfe des Terms  $x(t_f)^T M x(t_f)$ .
- Das Einschwingverhalten des Zustands durch  $\int_0^{t_f} x^T Q x dt$ .
- Die (Energie-)Kosten, die für  $u$  aufgewendet werden, durch den Term  $\int_0^{t_f} u^T R u dt$ .

Im Beispiel 1.6 würde also  $M$  die Abweichung von  $\varphi(t_f) = \pi$  und  $\dot{\varphi}(t) = 0$  gewichten, der erste Term im Kostenfunktional kann dazu dienen, möglichst ein oszillatorisches Einschwingverhalten zu verhindern, während der letzte Term im Kostenfunktional die Energiekosten für das angelegte Trägheitsmoment bewertet.

Der gemischte Term  $x^T S u$  ist meist nicht vorhanden. Er tritt z.B. dann auf, wenn die Ausgangsgleichung einen Durchgriffsterm  $Du$  wie in (1.5). enthält. Dann hat man wie in Bemerkung 4.4

$$\begin{aligned} y^T Q_y y + x^T Q_x x &= x^T C^T Q_y C x + x^T Q_x x + u^T D^T Q_y C x + x^T C^T Q_y D u \\ &= x^T (C^T Q_y C + Q_x) x + 2x^T C^T Q_y D u \\ &=: x^T Q x + 2x^T S u. \end{aligned}$$

Durch Anwendung des Pontryagin'schen Maximumprinzips erhält man nun die notwendigen Optimalitätsbedingungen. Der Satz kann direkt bewiesen werden ohne Satz 4.2 zu verwenden, wobei die Beweisstruktur dem allgemeineren Beweis folgt. Im Folgenden sei  $\mathcal{U}_{ad}$  die Menge der auf  $[0, t_f]$  stückweise stetigen Funktionen, der Beweis für  $\mathcal{U}_{ad} = \mathcal{L}_2([0, t_f], \mathbb{R}^m)$  geht analog.

**Satz 4.5** *Betrachte das Optimalsteuerungsproblem (4.5). Sei  $u_* \in \mathcal{U}_{ad}$  eine optimale Steuerung und sei  $x_*(t) = x(t, u_*)$  die zugehörige Lösungstrajektorie.*

*Dann gibt es eine Kozustandsfunktion  $\mu(t) : [0, t_f] \rightarrow \mathbb{R}^n$ , so daß  $x_*(t), \mu(t), u_*(t)$  das lineare Randwertproblem*

$$\begin{bmatrix} A & 0 & B \\ Q & A^T & S \\ S^T & B^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ \mu(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} I_n & 0 & 0 \\ 0 & -I_n & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}(t) \\ \dot{\mu}(t) \\ \dot{u}(t) \end{bmatrix} \quad (4.6)$$

mit Randbedingungen

$$x(0) = x^0, \quad \mu(t_f) = M x(t_f) \quad (4.7)$$

lösen.

Beachte, daß  $\dot{u}$  hier nur formal eingeht,  $u$  also nicht differenzierbar vorausgesetzt werden muß.

**Beweis:** Der Beweis folgt im Wesentlichen dem Ansatz des "Verschwindens der ersten Variation" aus der Variationsrechnung. Sei dazu  $u_*$  die Optimalsteuerung. Betrachte eine Störung erster Ordnung

$$u(t) = u_*(t) + \varepsilon v(t)$$

mit  $u(t) \in \mathcal{U}_{ad}$ ,  $\varepsilon \in \mathbb{R}$ . Damit wird die Nebenbedingung in (4.5) zu

$$\dot{x}(t) = Ax(t) + Bu_*(t) + \varepsilon Bv(t)$$

mit Lösungstrajektorie (siehe Kapitel 2)

$$\begin{aligned} x(t) &= e^{At}x^0 + \int_0^t e^{A(t-s)}B(u_*(s) + \varepsilon v(s)) ds \\ &= x_*(t) + \varepsilon \underbrace{\int_0^t (e^{A(t-s)}Bv(s)) ds}_{=:z(t)} \\ &= x_*(t) + \varepsilon z(t). \end{aligned}$$

Dabei erfüllt  $z(t)$  die lineare, inhomogene Differentialgleichung

$$\dot{z}(t) = Az(t) + Bv(t), \quad z(0) = 0. \quad (4.8)$$

Wir führen nun den Vektor  $\mu(t) \in \mathbb{R}^n$  und die Hamilton-Funktion  $\mathcal{H}(x, u, \mu)$  ein durch

$$\mathcal{H}(x, u, \mu) = \frac{1}{2} (x^T(t)Qx(t) + 2x(t)^T Su(t) + u(t)^T Ru(t)) + \mu(t)^T (Ax(t) + Bu(t))$$

Damit schreiben wird das Kostenfunktional um zu

$$\mathcal{J}(u) = \frac{1}{2} x^T(t_f) M x(t_f) + \int_0^{t_f} (\mathcal{H}(x, u, \mu) - \mu^T \dot{x}) dt.$$

Analog erhalten wir für  $u_*$ ,  $x_*$

$$\mathcal{J}(u_*) = \frac{1}{2} x_*^T(t_f) M x_*(t_f) + \int_0^{t_f} (\mathcal{H}(x_*, u_*, \mu) - \mu^T \dot{x}_*) dt.$$

Subtraktion der beiden Gleichungen ergibt

$$\begin{aligned} \mathcal{J}(u) - \mathcal{J}(u_*) &= \frac{1}{2} (x(t)^T M x(t) - x_*(t)^T M x_*(t)) \Big|_{t=t_f} \\ &\quad + \int_0^{t_f} (H(x, u, \mu) - H(x_*, u_*, \mu)) dt \\ &\quad + \int_0^{t_f} \mu^T \underbrace{(\dot{x}_* - \dot{x})}_{=-\varepsilon \dot{z}} dt \end{aligned} \quad (4.9)$$

Wir betrachten nun die drei Terme der rechten Seite von (4.9) separat. Für den ersten Term erhält man durch Einsetzen von  $x = x_* + \varepsilon z$

$$x^T M x - x_*^T M x_* = x_*^T M x_* + 2\varepsilon x_*^T M z + \varepsilon^2 z^T M z - x_*^T M x_* = 2\varepsilon x_*^T M z + \mathcal{O}(\varepsilon^2).$$

Die aufwendigste Behandlung erfordert der zweite Term.

$$\begin{aligned}\mathcal{H}(x, u, \mu) - \mathcal{H}(x_*, u_*, \mu) &= \frac{1}{2} (x^T Q x + 2x^T S u + u^T R u) + \mu^T (A x + B u) \\ &\quad - \frac{1}{2} (x_*^T Q x_* + 2x_*^T S u_* + u_*^T R u_*) - \mu^T (A x_* + B u_*).\end{aligned}$$

Wir setzen jetzt

$$u = u_* + \varepsilon v, \quad x = x_* + \varepsilon z$$

ein und erhalten

$$\begin{aligned}\mathcal{H}(x, u, \mu) - \mathcal{H}(x_*, u_*, \mu) &= \frac{1}{2} \{ x_*^T Q x_* + 2\varepsilon x_*^T Q z + \varepsilon^2 z^T Q z - x_*^T Q x_* \\ &\quad + 2(\varepsilon^2 z^T S v + x_*^T S u_* + \varepsilon(x_*^T S v + z^T S u_*) - x_*^T S u_*) \\ &\quad + \varepsilon^2 v^T R v + u_*^T R u_* + 2\varepsilon(u_*^T R v) - u_*^T R u_* \} \\ &\quad + \mu^T A x_* + \varepsilon \mu^T A z + \mu^T B u_* + \varepsilon \mu^T B v - \mu^T A x_* - \mu^T B u_* \\ &= \varepsilon(x_*^T Q z + x_*^T S v + u_*^T S^T z + u_*^T R v + \mu^T A z + \mu^T B v) + \mathcal{O}(\varepsilon^2) \\ &= \varepsilon(\{x_*^T Q + u_*^T S^T + \mu^T A\}z + \{x_*^T S + u_*^T R + \mu^T B\}v) + \mathcal{O}(\varepsilon^2).\end{aligned}$$

Weiter gilt für den letzten Term in (4.9) mit partieller Integration

$$- \int_0^{t_f} \varepsilon \mu^T \dot{z} dt = -\varepsilon \mu^T z \Big|_0^{t_f} + \varepsilon \int_0^{t_f} \dot{\mu}^T z dt = -\varepsilon \mu^T(t_f) z(t_f) + \varepsilon \int_0^{t_f} \dot{\mu}^T z dt$$

wobei wir verwenden, daß  $z(0) = 0$  wegen (4.8).

Zusammen erhalten wir also

$$\begin{aligned}\mathcal{J}(u) - \mathcal{J}(u_*) &= \varepsilon \left( \int_0^{t_f} (\{x_*^T Q + u_*^T S^T + \mu^T A + \dot{\mu}^T\}z + \{x_*^T S + u_*^T R + \mu^T B\}v) dt \right. \\ &\quad \left. - \mu^T(t_f) z(t_f) + x_*^T(t_f) M z(t_f) \right) + \mathcal{O}(\varepsilon^2)\end{aligned}\tag{4.10}$$

Notwendig für ein Minimum von  $\mathcal{J}$  ist nun das Verschwinden aller Richtungsableitungen von  $\mathcal{J}$  von  $u_*$  aus, d.h.,

$$0 = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\mathcal{J}(u_* + \varepsilon v) - \mathcal{J}(u_*)) \quad \text{für alle } v \in \mathcal{U}_{ad}.$$

Wählt man  $\mu$  als Lösung der linearen inhomogenen Differentialgleichung

$$\dot{\mu}(t) = -(A^T \mu(t) + Q x_* + S u_*)\tag{4.11}$$

mit „Endbedingung“

$$\mu(t_f) = M x_*(t_f)\tag{4.12}$$

dann ergibt sich aus (4.10) notwendigerweise

$$x_*^T S + u_*^T R + \mu^T B \equiv 0 \quad \forall t \in [0, t_f]\tag{4.13}$$

Fassen wir nun die Gleichungen (4.11), (4.12), (4.13) und die erste Gleichung in (4.4) zusammen, so erhalten wir das Zwei-Punkt-Randwertproblem (4.6),(4.7).  $\square$



**Bemerkung 4.6** Das Randwertproblem (4.6) ergibt sich auch, wenn man das Pontryagin'sche Maximumprinzip auf das LQR Problem anwendet. Die erste Zeile von (4.6) entspricht der Nebenbedingung, also  $\mathcal{H}_\mu = \dot{x}$ , die zweite Zeile ergibt sich aus der adjungierten Gleichung (Satz 4.2(ii)), während die letzte Zeile von (4.6) aus der notwendigen Bedingung für ein Minimum,

$$\mathcal{H}_u = 0,$$

resultiert. (Beachte dabei, daß  $u$  hier als unbeschränkt angenommen wird.)

Die Randbedingungen sind gerade der Anfangswert der dynamischen Nebenbedingung in (4.5) und die Transversalitätsbedingung in der Form (iii)' für nicht-verschwindende Funktion  $h_f$ .

Man beachte, daß zur Herleitung der notwendigen Optimalitätsbedingungen keine Bedingungen an die das Kostenfunktional definierenden Matrizen  $M, Q, R, S$  notwendig waren. Um nun hinreichende Optimalitätsbedingungen zu erhalten, verwendet man im Wesentlichen, daß  $\mathcal{J}_{uu} \geq 0$  für ein Minimum gelten muß. Daraus ergibt sich, daß  $M$  und  $\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}$  zumindest positiv semidefinit sein sollten.

Als hinreichende Optimalitätsbedingung erhält man dann den folgenden Satz.

**Satz 4.7** Seien  $x_*, \mu, u_*$  so gewählt, daß  $\begin{bmatrix} x_* \\ \mu \\ u_* \end{bmatrix}$  das lineare Randwertproblem (4.6), (4.7) löst. Weiterhin gelte, daß  $\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}$ ,  $M$  positiv semidefinit sind.

Dann gilt

$$\mathcal{J}(u) \geq \mathcal{J}(u_*) \quad (4.14)$$

für alle  $u \in \mathcal{U}_{ad}$ .

**Beweis:** Wir gehen hier vor wie bei konvexer Optimierung. Definiere dazu

$$\Phi(s) = \mathcal{J}(su_*(t) + (1-s)v(t)).$$

Beachte, daß sich aus der Linearität von  $\dot{x} = Ax + Bu$  durch Einsetzen von  $u = su_*(t) + (1-s)v$  die Lösungstrajektorie  $x = sx_* + (1-s)z$  ergibt, wobei  $z$  die zu  $v$  gehörende Lösung von (4.4) ist.

Die Behauptung des Satzes ist äquivalent zu der Aussage, daß  $\Phi(s)$  sein Minimum bei  $s = 1$  annimmt für alle  $x(t), u(x)$  welche (1.4) erfüllen. Da  $\Phi(s)$  quadratisch in  $s$  ist, hat also  $\Phi(s)$  ein Minimum für  $s = 1$  genau dann, wenn

$$\left. \frac{d\Phi}{ds} \right|_{s=1} = 0, \quad \left. \frac{d^2\Phi}{ds^2} \right|_{s=1} \geq 0.$$

Aus der Bedingung an die erste Ableitung folgt mit der für jede quadratische Form mit  $K = K^T$  gültigen Identität

$$\begin{aligned} & \left. \frac{d}{ds} \left( \frac{1}{2} (sq + (1-s)p)^T K (sq + (1-s)p) \right) \right|_{s=1} \\ &= \left. (sq^T Kq - sp^T Kq + (1-s)p^T Kq - (1-s)p^T Kp) \right|_{s=1} \\ &= q^T Kq - p^T Kq = (q-p)^T Kq \end{aligned}$$

der folgende Ausdruck:

$$\begin{aligned} \left. \frac{d\Phi}{ds} \right|_{s=1} &= (x_* - z)^T M x_* \Big|_{t=t_f} \\ &+ \int_0^{t_f} \{ (x_* - z)^T Q x_* + u_*^T S^T (x_* - z) + (u_* - v)^T S^T x_* + (u_* - v)^T R u_* \} dt \end{aligned} \quad (4.15)$$

Multipliziert man die zweite Gleichung von (4.6) von links mit  $x_*^T$ , setzt dann die erste und anschließend die dritte Gleichung von (4.6) ein, so erhält man

$$\begin{aligned} x_*^T Q x_* &= -x_*^T A^T \mu - x_*^T S u_* - x_*^T \dot{\mu} \\ &= u_*^T B^T \mu - \dot{x}_*^T \mu - x_*^T S u_* - x_*^T \dot{\mu} \\ &= -u_*^T S^T x_* - u_*^T R u_* - \dot{x}_*^T \mu - x_*^T S u_* - x_*^T \dot{\mu} \end{aligned} \quad (4.16)$$

Analog ergibt sich nach Multiplikation mit  $z^T$

$$\begin{aligned} z^T Q x_* &= -z^T A^T \mu - z^T S u_* - z^T \dot{\mu} \\ &= v^T B^T \mu - \dot{z}^T \mu - z^T S u_* - z^T \dot{\mu} \\ &= -v^T S^T x_* - v^T R u_* - \dot{z}^T \mu - z^T S u_* - z^T \dot{\mu} \end{aligned} \quad (4.17)$$

Einsetzen von (4.16), (4.17) in (4.15) ergibt

$$\begin{aligned} \left. \frac{d\Phi}{ds} \right|_{s=1} &= (x_* - z)^T M x_* \Big|_{t=t_f} + \int_0^{t_f} (z^T \dot{\mu} + \dot{z}^T \mu - x_*^T \dot{\mu} - \dot{x}_*^T \mu) dt \\ &= (x_*(t) - z(t))^T M x_*(t) \Big|_{t=t_f} + z^T(t) \mu(t) \Big|_{t=0}^{t=t_f} - x_*^T(t) \mu(t) \Big|_{t=0}^{t=t_f} \end{aligned}$$

Nun gilt nach (4.7)  $z(0) = x^0 = x_*(0)$  und  $\mu(t_f) = M x_*(t_f)$ , also  $\left. \frac{d\Phi}{ds} \right|_{s=1} = 0$ .

Mit der Identität

$$\frac{d^2}{ds^2} \left( \frac{1}{2} (sq + (1-s)p)^T K (sq + (1-s)p) \right) = (q-p)^T K (q-p)$$

für  $K = K^T$  gilt für die zweite Ableitung von  $\Phi$

$$\begin{aligned} \left. \frac{d^2\Phi}{ds^2} \right|_{s=1} &= (x_* - z)^T M (x_* - z) \Big|_{t=t_f} \\ &+ \int_0^{t_f} [(x_* - z)^T, (u_* - v)^T] \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x_* - z \\ u_* - v \end{bmatrix} dt \\ &\geq 0, \end{aligned}$$

wobei die positiv Semidefinitheit folgt, da  $M, \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}$  nach Voraussetzung positiv semidefinit sind.  $\square$

Wir haben also den Zusammenhang hergestellt zwischen der Lösung des optimalen Steuerungsproblems und der Lösung eines linearen 2-Punkt Randwertproblems. Im Prinzip könnten wir also die optimale Steuerung  $u_*$  des LQR Problems durch Lösen von (4.6), (4.7) erhalten.

Numerisch wesentlich einfacher und günstiger ist jedoch der im Folgenden beschriebene Weg.

## Lösung des LQR Problems durch Riccatigleichungen

Die Voraussetzungen von Satz 4.7 implizieren, daß auf jeden Fall  $R \geq 0$  im Kostenfunktional gewählt werden sollte. In der Regel hat man sogar  $R > 0$ . Sonst gibt es kostenfreie Steuerungsparameter, was meist physikalisch nicht sinnvoll ist. Wir wollen uns im Folgenden also auf den Fall einer positiv definiten Gewichtsmatrix für die Steuerung beschränken. In diesem Fall ist  $R$  invertierbar und die dritte Gleichung in (4.6) kann nach  $u$  aufgelöst werden. Man erhält damit

$$u(t) = -R^{-1}(S^T x(t) + B^T \mu(t)). \quad (4.18)$$

Damit kann man  $u$  aus dem Randwertproblem (4.6) eliminieren und erhält

$$\begin{bmatrix} \dot{x} \\ \dot{\mu} \end{bmatrix} = H \begin{bmatrix} x \\ \mu \end{bmatrix}, \quad x(0) = x^0, \quad \mu(t_f) = Mx(t_f), \quad (4.19)$$

wobei

$$H = \begin{bmatrix} A - BR^{-1}S^T & -BR^{-1}B^T \\ -(Q - SR^{-1}S^T) & -(A - BR^{-1}S^T)^T \end{bmatrix}. \quad (4.20)$$

Im Folgenden setzen wir zur Vereinfachung der Schreibweise

$$A := A - BR^{-1}S^T, \quad G := BR^{-1}B^T, \quad Q := Q - SR^{-1}S^T, \quad (4.21)$$

so daß  $Ax + Bu$  in dieser Schreibweise mit (4.18) zu  $Ax - G\mu$  wird und wir  $H = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix}$  schreiben können.

**Definition 4.8**  $H \in \mathbb{R}^{2n \times 2n}$  heißt Hamiltonisch, falls

$$(HJ) = (HJ)^T, \quad \text{wobei } J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \quad (4.22)$$

Aus (4.22) ergibt sich sofort, daß Hamiltonische Matrizen eine explizite Blockstruktur besitzen.

**Lemma 4.9**  $H \in \mathbb{R}^{2n \times 2n}$  ist Hamiltonisch genau dann, wenn

$$H = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \quad \text{wobei } G = G^T, \quad Q = Q^T.$$

D.h. also, daß (4.19) ein Randwertproblem für eine lineare Differentialgleichung mit Hamiltonischer Koeffizientenmatrix darstellt. Einige Eigenschaften Hamiltonischer Matrizen werden wir später noch näher untersuchen.

Mit dem Ansatz  $\mu(t) = X(t)x(t)$  erhält man aus (4.19) und mit

$$\begin{aligned} \dot{\mu} &= \dot{X}x + X\dot{x}, \\ \mu(t_f) &= X(t_f)x(t_f) \end{aligned}$$

die “Endbedingung”  $X(t_f) = M$  sowie

$$\begin{aligned}\dot{x} &= Ax - G\mu = Ax - GXx = (A - GX)x, \\ \dot{\mu} &= -Qx - A^T\mu = -Qx - A^T Xx \\ &= \dot{X}x + X\dot{x} = \dot{X}x + XAx - XGXx.\end{aligned}$$

Aus der letzten Gleichung erhalten wir

$$\dot{X}x = -(Q + A^T X + XA - XGX)x.$$

Da diese Gleichung für alle  $t \in [0, t_f]$  erfüllt sein muß, ergibt sich aus der Variation von  $x$ , daß  $X$  die *Matrix-Riccati-Differentialgleichung*

$$\dot{X}(t) = -\mathcal{R}(X(t)) := -(Q + A^T X(t) + X(t)A - X(t)GX(t)), \quad t \in [0, t_f] \quad (4.23)$$

mit Endbedingung

$$X(t_f) = M \quad (4.24)$$

erfüllen muß.

Man kann zeigen, daß (4.23), (4.24) auf  $[0, t_f]$  eine eindeutige Lösung besitzt. Da mit  $X$  auch  $X^T$  Lösung ist, folgt aus der Eindeutigkeit, daß  $X(t) = X(t)^T$  für alle  $t \in [0, t_f]$ . Weiter kann man zeigen, daß  $X(t)$  auf dem gesamten Intervall positiv semidefinit ist. (Die Beweise findet man z.B. in [17].)

Insgesamt erhält man das folgende Resultat, welches man auch für LTV Systeme zeigen kann (siehe [17]).

**Satz 4.10** *Falls  $R > 0$ , so gilt unter den Voraussetzungen von Satz 4.7, daß die optimale Steuerung  $u_*$ , die das LQR Problem (4.5) löst, gegeben ist durch*

$$u_*(t) = -R^{-1}(S^T + B^T X_*(t))x_*(t) \quad \forall t \in [0, t_f], \quad (4.25)$$

wobei  $X_*$  die eindeutige Lösung der Riccati-Differentialgleichung (4.23) mit Endbedingung (4.24) ist.

Die “optimalen Kosten” betragen

$$\mathcal{J}(u_*) = \frac{1}{2}(x^0)^T X_*(0)x^0. \quad (4.26)$$

**Beweis:** Die Form der optimalen Steuerung (4.25) ergibt sich direkt aus dem Randwertproblem (4.6)–(4.7) und dem Einsetzen von  $\mu = X_*x_*$  in (4.18).

Die optimalen Kosten erhält man durch Betrachtung der “Wertefunktion”<sup>1</sup>

$$V(x(t)) := x(t)^T X(t)x(t),$$

---

<sup>1</sup>Der Begriff Wertefunktion kommt aus einem alternativen Ansatz zur Lösung von dynamischen Optimierungsproblemen, dem “Bellmannschen Optimalitätsprinzip”, welches hier auf dieselbe Lösung (4.25) führt, aber auch im Fall nichtlinearer Systeme anwendbar ist.

wobei  $X$  die Lösung der Riccati-Differentialgleichung (4.23) ist. Es gilt durch Einsetzen von  $\dot{x} = Ax + Bu$ , (4.23) und (4.25)

$$\begin{aligned}
\frac{d}{dt}V(x_*) &= \dot{x}_*^T X_* x_* + x_*^T \dot{X}_* x_* + x_*^T X_* \dot{x}_* = 2\dot{x}_*^T X_* x_* + x_*^T \dot{X}_* x_* \\
&= 2x_*^T A^T X_* x_* + 2u_*^T B^T X_* x_* - x_*^T (Q + A^T X_* + X_* A - X_* G X_*) x_* \\
&= -2X_* B R^{-1} B^T X_* - x_*^T Q x_* + x_*^T G x_* \\
&= -x_*^T Q x_* - x_*^T G x_*.
\end{aligned}$$

Damit kann man das Kostenfunktional für die optimale Steuerung schreiben als

$$\begin{aligned}
\mathcal{J}(u_*) &= \frac{1}{2} x_*^T(t_f)^T M x_*(t_f) - \frac{1}{2} \int_0^{t_f} \frac{d}{dt} V(x_*)(t) dt \\
&= \frac{1}{2} \left( x_*^T(t_f)^T M x_*(t_f) - V(x_*(t)) \Big|_{t=0}^{t=t_f} \right) \\
&= \frac{1}{2} \left( x_*^T(t_f)^T M x_*(t_f) - x_*^T(t_f)^T X_*(t_f) x_*(t_f) + x_*^T(0)^T X_*(0) x_*(0) \right) \\
&= \frac{1}{2} x_*^T(0)^T X_*(0) x_*(0),
\end{aligned}$$

wobei zum Schluß noch die Endbedingung  $X(t_f) = M$  verwendet wurde.  $\square$

**Bemerkung 4.11** *Man beachte, daß die optimale Steuerung  $u_*$  in (4.25) als lineare Zustandsrückführung bzw. Feedback-Steuerung vorliegt. Man erhält also einen geschlossenen Regelkreis, obwohl dies aus dem Randwertproblem (4.6)–(4.7) nicht ersichtlich ist. Der Ansatz  $\mu = Xx$  erklärt sich allerdings aus der Zielstellung, eine Feedback-Steuerung zu erhalten. Das bedeutet, daß man im Fall einer bzgl. des LQR Problems optimalen Steuerung eine Regelung erhält (siehe Bemerkung 1.9).*

Damit hat man zur Lösung des LQR Problems nun eine Alternative, indem man das “Endwertproblem” für die Riccati-Differentialgleichung löst. Dazu kann man die Gleichung mit Hilfe des vec-Operators und des Kroneckerprodukts vektorisieren und Standardverfahren zur Lösung von Anfangswertaufgaben heranziehen, wobei man durch die Transformation  $t \rightarrow t_f - t$  aus dem Endwertproblem ein Anfangswertproblem machen kann. Sinnvoller ist hier allerdings der Einsatz von speziellen Verfahren für Riccati-Differentialgleichungen, die die gegebene Struktur ausnutzen [10, 11, 16].

**Unendlicher Zeithorizont:**  $t_f = \infty$

Wie bereits in der Einleitung diskutiert, reicht häufig eine asymptotische Erreichung des Steuerungsziels aus. Dies führt zu der Frage nach einer optimalen Stabilisierung hinsichtlich des Kostenfunktionals in (4.5) mit  $t_f = \infty$ , wobei wir jetzt  $M = 0$  setzen. Um eine Stabilisierung des Systems überhaupt erreichen zu können, müssen wir natürlich die Stabilisierbarkeit voraussetzen. Da wir auch hier wieder eine Lösung als Feedback-Steuerung

anstreben, legt (4.18) den Ansatz  $\mu(t) = Xx(t)$  für eine konstante Matrix  $X = X^T \in \mathbb{R}^{n \times n}$  nahe. Mit diesem Ansatz erhält man analog zum endlichen Fall

$$\begin{aligned}\dot{x} &= (A - GX)x, \\ \dot{\mu} &= -Qx - A^T\mu = -Qx - A^TXx \\ &= X\dot{x} = X(A - GX)x.\end{aligned}\tag{4.27}$$

Die letzte Gleichung ist nun äquivalent zu

$$(Q + A^TX + XA - XGX)x(t) = 0 \quad \forall t \in [0, \infty).$$

Durch Variation von  $x$  erhält man als notwendige Bedingung, daß  $X$  die *algebraische Riccatigleichung* (kurz: ARE von *algebraic Riccati equation*)

$$0 = \mathcal{R}(X) := Q + A^TX + XA - XGX\tag{4.28}$$

erfüllen muß. Man beachte, daß im Gegensatz zur Riccati-Differentialgleichung, (4.28) unendlich viele Lösungen besitzen kann und sogar unsymmetrische Lösungen möglich sind. Die Struktur der Lösungsmenge von (4.28) ist Gegenstand zahlreicher Forschungsarbeiten und bisher am ausführlichsten in [18] beschrieben.

Wir werden im Folgenden sehen, daß wir eine bestimmte Lösung der ARE benötigen. Da die Lösungstrajektorie einer jeden Steuerung, die aus unserem Ansatz generiert wird, die lineare homogene Differentialgleichung (4.27) erfüllt, so daß sich die Lösung schreiben läßt als  $x(t) = e^{(A-GX)t}x^0$ , muß notwendigerweise (siehe Satz 1.8)

$$\Lambda(A - GX) \subset \mathbb{C}^-$$

gelten, da sonst keine Stabilisierung des Systems erreicht würde. Dies motiviert folgende Definition.

**Definition 4.12** *Ein Lösung  $X \in \mathbb{R}^{n \times n}$  der ARE (4.28) heißt stabilisierend falls  $\Lambda(A - GX) \subset \mathbb{C}^-$ .*

Um also eine stabilisierende Feedback-Matrix  $F \in \mathbb{R}^{m \times n}$  mit Hilfe des LQR Problems zu berechnen, benötigen wir eine stabilisierende Lösung der ARE, da

$$A - GX = A - BR^{-1}B^TX = A + BF, \quad \text{mit } F := -R^{-1}B^TX.$$

Es bleibt also zu zeigen, wie man die stabilisierende Lösung, im Folgenden mit  $X_*$  bezeichnet, berechnen kann. Sei dazu  $X$  eine (zunächst beliebige) Lösung der ARE (4.28) und  $T := \begin{bmatrix} I_n & 0 \\ X & I_n \end{bmatrix}$ . Dann gilt für die Hamiltonische Matrix  $H$  aus (4.20) mit den Abkürzungen (4.21)

$$\begin{aligned}T^{-1}HT &= \begin{bmatrix} I_n & 0 \\ -X & I_n \end{bmatrix} \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} I_n & 0 \\ X & I_n \end{bmatrix} \\ &= \begin{bmatrix} A - GX & -G \\ -\mathcal{R}(X) & -(A - GX)^T \end{bmatrix} = \begin{bmatrix} A - GX & -G \\ 0 & -(A - GX)^T \end{bmatrix}\end{aligned}$$

Daraus folgt

$$H \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} (A - GX), \quad (4.29)$$

d.h.  $\Lambda(A - GX) \subset \Lambda(H)$  und die Spalten von  $\begin{bmatrix} I \\ X \end{bmatrix}$  spannen einen  $H$ -invarianten Unterraum zu diesen Spalten auf. Dies gilt für jede Lösung der ARE, für die stabilisierende Lösung benötigen wir also einen  $n$ -dimensionalen  $H$ -invarianten Unterraum zu den Eigenwerten in der linken Halbebene. Zunächst ist die Frage zu klären, ob solch ein Unterraum überhaupt existiert. Dazu benötigen wir einige Eigenschaften des Spektrums Hamiltonischer Matrizen.

**Lemma 4.13** *Ist  $H \in \mathbb{R}^{2n \times 2n}$  Hamiltonisch und  $\lambda \in \Lambda(H)$  mit zugehörigem Rechts-Eigenvektor  $x \in \mathbb{C}^{2n}$ , dann ist  $-\bar{\lambda} \in \Lambda(H)$  mit zugehörigem Links-Eigenvektor  $Jx$ , wobei*

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

**Beweis:** Es gilt nach Definition 4.8

$$\lambda x = Hx = HJJ^T x = (HJ)^T J^T x = J^T H^T J^T x$$

Multiplikation von links mit  $J^T$  und komplex-konjugiert Transponieren ergibt

$$-\bar{\lambda}(Jx)^* = (Jx)^* H,$$

womit die Behauptung folgt. □

Da bei reellen Matrizen mit  $\lambda$  auch  $\bar{\lambda}$  im Spektrum erhalten ist, treten Eigenwerte Hamiltonischer Matrizen immer in Quadrupeln  $\lambda, \bar{\lambda}, -\lambda, \bar{-\lambda}$  auf, es sei denn, es handelt sich um rein reelle oder imaginäre Eigenwerte, die in Paaren  $\lambda, -\lambda$  bzw.  $\omega, -\omega$  ( $\omega \in \mathbb{R}$ ) auftreten. Insgesamt können wir also das Spektrum Hamiltonischer Matrizen wie folgt schreiben:

$$\Lambda(H) = \{\lambda_1, \dots, \lambda_n\} \cup \{-\lambda_1, \dots, -\lambda_n\}, \quad \operatorname{Re}(\lambda_j) \leq 0 \quad \forall j \in \{1, \dots, n\}. \quad (4.30)$$

Ein typisches Spektrum einer Hamiltonischen Matrix ist in Abbildung 4.1 dargestellt.

Damit eine stabilisierende Lösung der ARE existieren kann, folgt also aus (4.29), daß die zugehörige Hamiltonische Matrix keine Eigenwerte auf der imaginären Achse haben darf, da dann exakt  $n$  Eigenwerte von  $H$  in der linken Halbebene liegen und ein  $n$ -dimensionaler  $H$ -invarianten Unterraum zu diesen Eigenwerten existiert.

Dies kann man aber schon mit den minimalen Forderungen an ein geregeltes System sicherstellen, wie der folgende Satz zeigt.

**Satz 4.14** *Sei  $H = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$  Hamiltonisch, wobei  $(A, G)$  stabilisierbar,  $(A, Q)$  entdeckbar und  $G, Q \geq 0$  seien. Dann gilt*

$$\operatorname{Re}(\lambda) \neq 0 \quad \text{für alle } \lambda \in \Lambda(H).$$

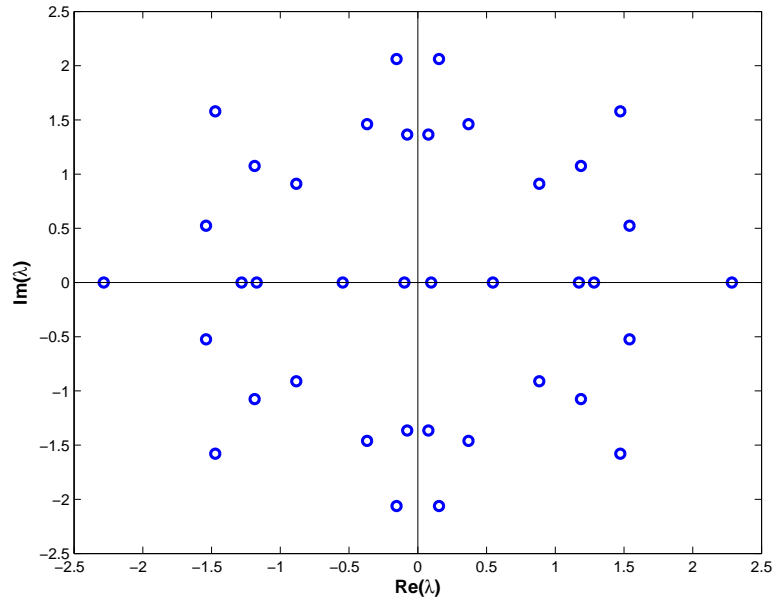


Abbildung 4.1: Typisches Spektrum einer Hamiltonischen Matrix

**Beweis:** Wir führen die Annahme, daß  $\lambda = \omega \in \Lambda(H)$  für ein  $\omega \in \mathbb{R}$  zum Widerspruch.

Wegen Lemma 4.13 können wir o.B.d.A. annehmen, daß  $\omega \geq 0$ . Sei  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0$  mit  $x_1, x_2 \in \mathbb{C}^n$  ein Eigenvektor zu  $\omega$ , dann gilt

$$H \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \omega \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (4.31)$$

Multiplikation von links mit  $\begin{bmatrix} x_2 \\ x_1 \end{bmatrix}^*$  führt zu

$$\begin{aligned} \begin{bmatrix} x_2^* & x_1^* \end{bmatrix} H \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \underbrace{x_2^* A x_1}_{:=\alpha} - \underbrace{x_2^* G x_2}_{:=\gamma} - \underbrace{x_1^* Q x_1}_{:=\beta} - \underbrace{x_2^* A x_1}_{=-\alpha} \\ &= \begin{bmatrix} x_2^* & x_1^* \end{bmatrix} \omega \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \omega (\underbrace{x_2^* x_1}_{=:\zeta} + \underbrace{x_1^* x_2}_{=\bar{\zeta}}) = 2\omega \operatorname{Re}(\zeta). \end{aligned}$$

Da  $\beta, \gamma \in \mathbb{R}$  ist der Realteil des ersten Ausdrucks

$$\gamma + \beta + \operatorname{Re}(\alpha - \bar{\alpha}) = \gamma + \beta,$$

während der Realteil des zweiten (rein-imaginären) Ausdrucks Null ist. Mit  $\beta, \gamma \geq 0$  wegen  $G, Q \geq 0$  folgt also  $\beta = \gamma = 0$  und daher

$$x_1^* Q = 0, \quad Q x_1 = 0 \quad \text{und} \quad x_2^* G = 0, \quad G x_2 = 0.$$



Aus der ersten Gleichung in (4.31) erhält man

$$\omega x_1 = Ax_1 - Gx_2 = Ax_1$$

und damit

$$\begin{bmatrix} A - \omega I \\ Q \end{bmatrix} x_1 = 0. \quad (4.32)$$

Analog erhält man mit Hilfe der zweiten Gleichung aus (4.31)

$$x_2^* [A - \omega I, G] = 0. \quad (4.33)$$

Da zumindest entweder  $x_1 \neq 0$  oder  $x_2 \neq 0$  gelten muß, führt entweder (4.32) zum Widerspruch gegen die angenommene Entdeckbarkeit oder (4.33) zum Widerspruch gegen die Stabilisierbarkeit.  $\square$

Die stabilisierende Lösung der ARE erhält man nun wie folgt.

**Lemma 4.15** *Seien  $U = [u_1, \dots, u_n]$ ,  $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$  so, daß  $\text{span} \left\{ \begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \dots, \begin{bmatrix} u_n \\ v_n \end{bmatrix} \right\}$  der  $n$ -dimensionale  $H$ -invariante Unterraum zu  $\{\lambda_1, \dots, \lambda_n\}$  ist (mit der Notation wie in (4.30)) und es sei  $\text{Re}(\lambda_j) < 0$  für alle  $j = 1, \dots, n$ . Ist  $U$  invertierbar, dann ist  $X_* = VU^{-1}$  die stabilisierende Lösung der ARE (4.28).*

**Beweis:** Nach Voraussetzung gilt

$$\begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} T, \quad \Lambda(T) = \{\lambda_1, \dots, \lambda_n\}.$$

Multiplikation der ersten Blockzeile dieser Gleichung mit  $U^{-1}$  von links ergibt

$$U^{-1}AU - U^{-1}GV = T.$$

Damit erhält man aus der zweiten Blockzeile

$$-QU - A^T V = VT = VU^{-1}AU - VU^{-1}GV.$$

Multipliziert man diese Gleichung von rechts mit  $U^{-1}$ , so folgt daß  $X_* = VU^{-1}$  die ARE (4.28) löst.

Weiter gilt

$$\text{range} \left( \begin{bmatrix} U \\ V \end{bmatrix} \right) = \text{range} \left( \begin{bmatrix} U \\ V \end{bmatrix} U^{-1} \right) = \text{range} \left( \begin{bmatrix} I \\ X_* \end{bmatrix} \right).$$

Also ist  $X_*$  die stabilisierende Lösung der ARE.  $\square$

Die folgenden Resultate zeigen einige weitere Eigenschaften der stabilisierenden Lösung  $X_*$  der ARE auf. Zunächst zeigen wir eine Symmetrieeigenschaft, die für das folgende hilfreich sein wird.

**Lemma 4.16** Seien  $U = [u_1, \dots, u_n]$ ,  $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$  so, daß  $\text{span} \left\{ \begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \dots, \begin{bmatrix} u_n \\ v_n \end{bmatrix} \right\}$  der  $n$ -dimensionale  $H$ -invariante Unterraum zu  $\{\lambda_1, \dots, \lambda_n\}$  ist (mit der Notation wie in (4.30)) und es sei  $\text{Re}(\lambda_j) < 0$  für alle  $j = 1, \dots, n$ . Dann ist  $V^T U$  symmetrisch. Gilt zusätzlich, daß  $Q$  und  $G$  positiv semidefinit sind, so folgt  $V^T U \geq 0$ .

**Beweis:** Nach Voraussetzung gilt

$$H \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} Z \quad (4.34)$$

mit  $\Lambda(Z) = \{\lambda_1, \dots, \lambda_n\}$ . Multipliziert man die erste Blockzeile von (4.34) mit  $V^T$  von links, erhält man

$$V^T A U - V^T G V = V^T U Z. \quad (4.35)$$

Transponieren der zweiten Blockzeile von (4.34) und Multiplikation mit  $U$  ergibt

$$-U^T Q U - V^T A U = Z^T V^T U. \quad (4.36)$$

Addition von (4.35) und (4.36) ergibt

$$Z^T V^T U + V^T U Z = -V^T G V - U^T Q U. \quad (4.37)$$

Dies ist eine Lyapunovgleichung für die "Unbekannte"  $V^T U$ . Da  $Z$  nach Voraussetzung stabil ist, folgt mit Satz 3.6, daß (4.37) eine eindeutige Lösung besitzt und

$$V^T U = \int_0^\infty e^{Z^T t} (V^T G V + U^T Q U) e^{Z t} dt.$$

Da wegen  $e^{Z^T t} = (e^{Z t})^T$  und  $Q = Q^T$ ,  $G = G^T$  der Integrand symmetrisch ist, gilt dies auch für  $V^T U$ . Sind  $G$  und  $Q$  positiv semidefinit, so folgt dies aufgrund des Sylvesterschen Trägheitssatzes auch für den Integranden, damit für die gesamte rechte Seite<sup>2</sup> und somit für  $V^T U$ .  $\square$

Das nächste Lemma garantiert unter den für das LQR Problem gültigen Annahmen die in Lemma 4.15 noch vorausgesetzte Invertierbarkeit von  $U$ .

**Lemma 4.17** Sind  $U, V$  wie in Lemma 4.16,  $Q, G$  positiv semidefinit, und  $(A, G)$  stabilisierbar, so ist  $U$  invertierbar.

**Beweis:** Angenommen,  $U$  wäre singulär. Dann existiert ein Vektor  $z \neq 0$  mit  $U z = 0$ . Multipliziert man nun die erste Blockzeile von (4.34) mit  $(V z)^T$  von links und  $z$  von rechts, so erhält man

$$z^T V^T A \underbrace{U z}_{=0} - z^T V^T G V z = z^T V^T U Z z. \quad (4.38)$$

<sup>2</sup>Dabei verwenden wir die Eigenschaft, daß Integration die positive Semidefinitheit nicht verändert. — Ist das klar??? — Wurde schon in Kap. 2 stillschweigend verwendet.

Wegen Lemma 4.16 folgt

$$z^T V^T G V z = -z^T V^T U Z z = -(U z)^T V Z z = 0.$$

Da  $G$  symmetrisch positiv semidefinit ist, folgt  $G V z = 0$ . Also ergibt die erste Blockzeile von (4.34) multipliziert mit  $z$  von rechts  $U Z z = 0$ . Da  $z \in \text{Kern}(U)$  beliebig gewählt war, folgt  $Z z \in \text{Kern}(U)$  für alle  $z \in \text{Kern}(U)$  und damit daß  $\text{Kern}(U)$   $Z$ -invariant ist. Also existiert ein Eigenwert von  $Z$ , d.h. ein  $\lambda_j$  ( $1 \leq j \leq n$ ), und ein Vektor  $0 \neq z_j \in \text{Kern}(U)$  mit  $Z z_j = \lambda_j z_j$ . Multiplikation der zweiten Blockzeile von (4.34) mit  $z_j$  ergibt

$$-Q \underbrace{U z_j}_{=0} - A^T V z_j = V Z z_j = \lambda_j V z_j,$$

also  $(\lambda_j I_n + A^T) V z_j = 0$ . Wir hatten schon  $G V z = 0$  für beliebige  $z \in \text{Kern}(U)$  gezeigt. Dies gilt also insbesondere auch für  $z_j$ . Damit folgt

$$(V z_j)^T [\lambda_j I_n + A, G] = 0$$

und wegen der Stabilisierbarkeit von  $(A, G)$  mit Satz 2.15 (Hautus-Test)  $V z_j = 0$ . Also haben wir

$$\begin{bmatrix} U \\ V \end{bmatrix} z_j = 0.$$

Da  $z_j \neq 0$  folgt  $\text{rank} \left( \begin{bmatrix} U \\ V \end{bmatrix} \right) < n$ , was im Widerspruch dazu steht, daß die Spalten von  $\begin{bmatrix} U \\ V \end{bmatrix}$  einen  $n$ -dimensionalen  $H$ -invarianten Unterraum aufspannen!  $\square$

Damit können wir nun den folgenden Satz über die stabilisierende Lösung der ARE formulieren.

**Satz 4.18** *Betrachte die ARE*

$$0 = Q + A^T X + X A - X G X \quad (4.39)$$

mit  $G = G^T \geq 0$ ,  $Q = Q^T \geq 0$ , und  $(A, G)$  stabilisierbar. Die zugehörige Hamiltonische Matrix  $H = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix}$  mit Spektrum wie in (4.30) habe keine rein imaginären Eigenwerte und der  $H$ -invariante Unterraum bzgl.  $\{\lambda_1, \dots, \lambda_n\}$  werde aufgespannt von den Spalten von  $\begin{bmatrix} U \\ V \end{bmatrix}$ ,  $U, V \in \mathbb{R}^{n \times n}$ . Dann besitzt die ARE (4.39) eine eindeutige stabilisierende Lösung  $X_*$ . Diese ist symmetrisch und positiv semidefinit.

**Beweis:** Zunächst gilt mit Lemma 4.17, daß  $U$  invertierbar ist. Also existiert  $X_* := V U^{-1}$  und die erste Blockzeile von (4.34) multipliziert mit  $U^{-1}$  liefert

$$A - G X_* = U Z U^{-1}.$$

Damit ist  $X_*$  nach Definition 4.12 stabilisierend, da  $\Lambda(U Z U^{-1}) = \Lambda(Z) = \{\lambda_1, \dots, \lambda_n\}$ . Die Symmetrie folgt mit Lemma 4.16, denn mit  $V^T U = U^T V$  erhalten wir

$$X_* = V U^{-1} = U^{-T} U^T V U^{-1} = U^{-T} V^T U U^{-1} = (V U^{-1})^T = X_*^T.$$

Unter den gegebenen Voraussetzungen ist Lemma 4.16  $V^T U \geq 0$  und aufgrund der Kongruenz

$$U^T X_* U = V^T U$$

gilt dann auch  $X_* \geq 0$  nach Trägheitssatz von Sylvester.

Es bleibt noch die Eindeutigkeit zu zeigen. Dazu nehme an, daß  $X_*$  und  $\tilde{X}_*$  zwei stabilisierende Lösungen der ARE (4.39) sind, also

$$\begin{aligned} 0 &= Q + A^T X_* + X_* A - X_* G X_*, \\ 0 &= Q + A^T \tilde{X}_* + \tilde{X}_* A - \tilde{X}_* G \tilde{X}_*. \end{aligned}$$

Subtraktion beider Gleichungen ergibt

$$0 = (A - G X_*)^T (X_* - \tilde{X}_*) + (X_* - \tilde{X}_*) (A - G \tilde{X}_*).$$

Dies ist eine homogene Sylvestergleichung und da nach Annahme  $\Lambda(A - G X_*) \cup \Lambda(A - G \tilde{X}_*) \subset \mathbb{C}^-$ , folgt nach Satz 3.6  $X_* - \tilde{X}_* = 0$ , also die Eindeutigkeit der stabilisierenden Lösung.  $\square$

Beachte: Neben der stabilisierende Lösung existieren unter den gegebenen Voraussetzungen noch endlich viele weitere Lösungen der ARE.

Positive Definitheit von  $X_*$  erhält man unter etwas stärkeren Voraussetzungen:

**Folgerung 4.19** *Gilt zusätzlich zu den Voraussetzungen von Satz 4.18 noch, daß  $(A, Q)$  beobachtbar ist, so ist die stabilisierende Lösung der ARE (4.39) positiv definit.*

**Beweis:** Man kann die ARE als Lyapunovgleichung

$$(A - G X_*)^T X_* + X_* (A - G X_*) = -Q - X_* G X_*$$

schreiben. Wenn wir zeigen können, daß mit  $(A, Q)$  auch  $(A - G X_*, Q + X_* G X_*)$  beobachtbar ist, dann folgt die Behauptung mit Satz 3.10 a), wenn man noch verwendet, daß dann  $((A - G X_*)^T, (Q + X_* G X_*)^T)$  steuerbar ist.

Nehmen wir dazu an,  $(A - G X_*, Q + X_* G X_*)$  wäre nicht beobachtbar. Dann existiert nach Folgerung 2.20 ein  $0 \neq z \in \mathbb{C}^n$  mit

$$(A - G X_*)z = \lambda z \quad \text{und} \quad (Q + X_* G X_*)z = 0.$$

Die letzte Gleichung impliziert

$$z^H (Q + X_* G X_*) z = 0,$$

und, da sowohl  $Q$  als auch  $X_* G X_*$  symmetrisch positiv semidefinit sind, auch

$$z^H Q z = 0 \quad \text{und} \quad z^H X_* G X_* z = 0.$$

Aufgrund der Semidefinitheit beider Matrizen gilt dann auch<sup>3</sup>

$$Qz = 0 \quad \text{und} \quad GX_*z = 0.$$

Damit erhält man

$$(A - GX_*)z = Az = \lambda z \quad \text{und} \quad Qz = 0,$$

was wegen Folgerung 2.20 im Widerspruch zu Beobachtbarkeit von  $(A, Q)$  steht.  $\square$

**Bemerkung 4.20** *Man beachte, daß man wegen Satz 4.14 die Folgerung 4.19 ohne die Annahme  $\operatorname{Re}(\lambda_j) < 0$  für alle  $j = 1, \dots, n$  formulieren kann. Ebenso kann man Satz 4.18 ohne diese Annahme beweisen, wenn man zusätzlich Entdeckbarkeit von  $(A, Q)$  voraussetzt.*

Insgesamt erhält man folgenden Satz über die Lösung des LQR Problems für  $t_f = \infty$ .

**Satz 4.21** *Seien  $Q - SR^{-1}S^T$ ,  $G = BR^{-1}B^T$  positiv semidefinit,  $(A, B)$  stabilisierbar und  $(A - BR^{-1}S^T, Q - SR^{-1}S^T)$  entdeckbar. Dann gelten die folgenden Aussagen:*

a) *Die ARE*

$$Q + A^T X + XA - (S + XB)R^{-1}(S + XB)^T = 0$$

*hat eine eindeutige stabilisierende Lösung  $X_* = X_*^T \geq 0$ .*

b) *Das LQR Problem (4.5) mit  $M = 0$ ,  $t_f = \infty$  hat eine eindeutige Lösung, die durch die Feedback-Steuerung*

$$u_*(t) = -R^{-1}(S^T + B^T X_*)x_*(t) \tag{4.40}$$

*gegeben ist. Die optimalen Kosten betragen*

$$\mathcal{J}(u_*) = \frac{1}{2}(x^0)^T X_* x^0.$$

**Beweis:** Es ist lediglich zu beachten, daß  $(A, B)$  stabilisierbar ist genau dann, wenn  $(A - BR^{-1}S^T, BR^{-1}B^T)$  stabilisierbar ist, was aus Satz 2.26 folgt. Dann folgt Teil a) mit Satz 4.14 sowie den o.e. Zusatzüberlegungen, Teil b) folgt analog wie im Fall  $t_f < \infty$ .  $\square$

**Bemerkung 4.22** *Wählt man  $t_f = \infty$ ,  $M = 0$ ,  $S = 0$ , in (4.5), so liefert jede Wahl  $Q > 0$ ,  $R > 0$  für ein stabilisierbares System eine stabilisierende Feedback-Matrix*

$$F_* = F_*(Q, R) = -R^{-1}B^T X_*(Q, R),$$

*wobei  $X_*(Q, R)$  die stabilisierende Lösung der zugehörigen ARE (4.28) ist. Bei geeigneter Wahl von  $Q, R$  erhält man in der Regel geschlossene Regelkreise, die weit weniger sensitiv gegenüber Störungen sind als solche, die über die Lösung des Polvorgabe-Problems oder mit dem Bass Algorithmus (Algorithmus 3.12) berechnet werden.*

Um das LQR Problem zu lösen, benötigen wir also Verfahren zur numerischen Lösung der ARE (4.28).

---

<sup>3</sup>Dabei verwende jeweils: Ist  $M = M^T \geq 0$ , so existiert  $L$  mit  $M = L^T L$ . Aus  $0 = z^H M z = z^H L^T L z = (Lz)^H (Lz)$  folgt damit  $Lz = 0$  und schließlich  $0 = L^T L z = Mz$ .

## Numerische Lösung von algebraischen Riccatigleichungen

Der erste Ansatz folgt direkt aus dem Satz 4.14. Ist  $H$  diagonalisierbar, so könnte man  $U, V$  aus den Eigenvektoren zu  $\{\lambda_1, \dots, \lambda_n\}$  konstruieren und das lineare Gleichungssystem

$$XU = V \quad (4.41)$$

lösen. Dieser Ansatz ist numerisch allerdings mit Vorsicht zu verwenden, insbesondere wenn  $H$  fast defektive Eigenwerte besitzt, da dann die Eigenvektorberechnung schlecht konditioniert ist.

Numerisch stabiler ist die in Algorithmus 4.23 beschriebene Methode, die auf Laub (1979) [20] zurückgeht.

---

### Algorithmus 4.23 (Schurvektormethode)

---

INPUT:  $A, G, Q \in \mathbb{R}^{n \times n}$ ,  $G = G^T$ ,  $Q = Q^T$ .

OUTPUT: Stabilisierende Lösung  $X_*$  der ARE (4.28).

1: Berechne die Schur-Zerlegung

$$\begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} W = W \begin{bmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{bmatrix}, \quad \Lambda(H_{11}) = \{\lambda_1, \dots, \lambda_n\}$$

mit der Sortierung wie in (4.30) (existiert nach Satz 3.14).

2: **if**  $\exists j \in \{1, \dots, n\}$  mit  $\operatorname{Re}(\lambda_j) = 0$  **then**

3: STOP; es existiert keine stabilisierende Lösung von (4.28).

4: **end if**

5: Partitioniere  $W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$  mit  $W_{ij} \in \mathbb{R}^{n \times n}$ .

6:  $X_* := W_{21}W_{11}^{-1}$ .

---

Die Schurvektormethode ist numerisch rückwärts stabil, falls  $\|X_*\|_2 \approx 1$ . Dies kann man meist durch eine geeignete Skalierung  $\tilde{\mathcal{R}}(X) = \mathcal{R}(\gamma X)$ ,  $\gamma \in \mathbb{R}^+$ , erreichen. Die Kosten betragen ca.  $205n^3$  Flops.

Der Nachteil der Methode besteht darin, daß sie die Hamiltonische Struktur von  $H$  ignoriert. In endlicher Arithmetik gilt daher (4.30) i.a. nicht mehr. D.h., wenn  $\tilde{\lambda}$  ein numerisch berechneter Eigenwert von  $H$  ist, so ist  $-\tilde{\lambda}$  i.a. nicht im berechneten Spektrum enthalten, siehe z.B. Abbildung 4.2.

Damit besteht die Möglichkeit daß mehr oder weniger als  $n$  berechnete Eigenwerte in der linken Halbebene liegen und damit eine stabilisierende Lösung mit der Schurvektormethode nicht berechnet werden kann.

Dieses Problem könnte vermieden werden, wenn  $W$  in der Schurzerlegung von  $H$  so gewählt werden könnte, daß die Schurform wieder Hamiltonisch ist, d.h.

$$H_{22} = -H_{11}^T, \quad H_{12} = H_{12}^T,$$

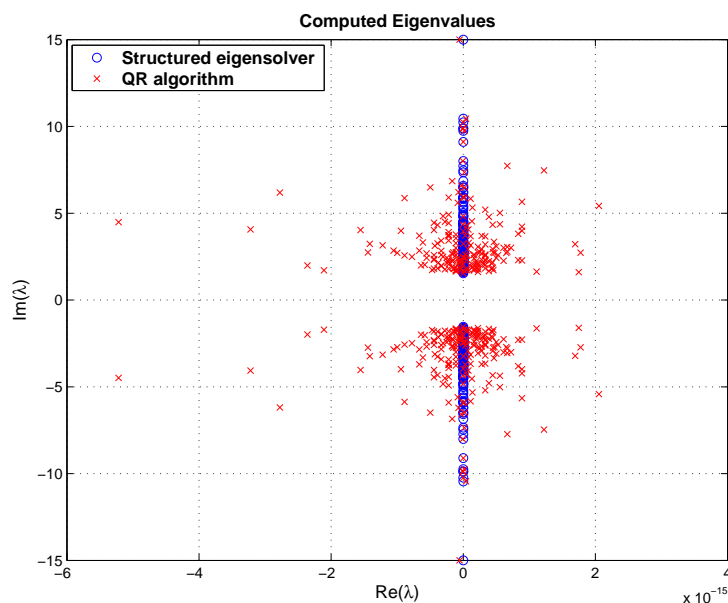


Abbildung 4.2: Spektrum einer Hamiltonischen Matrix, einmal berechnet mit dem unstrukturierten QR Algorithmus und einmal mit einem strukturerhaltenden Algorithmus.

wobei  $H_{11}$  reelle Schurform hat. Eine solche Schurform heißt *Hamiltonische Schurform* und man kann zeigen, daß eine solche Form selbst unter schwächeren Voraussetzungen als den hier vorliegenden existiert [24, 21]. Allerdings gibt es bislang keinen numerisch stabilen Algorithmus zur Berechnung der Hamiltonischen Schurform allgemeiner Hamiltonischer Matrizen, der mit einem vertretbaren numerischen Aufwand auskommt. Einige Algorithmen, die versuchen, die Symmetrie des Spektrums Hamiltonischer Algorithmen zu erhalten, sind im Folgenden kurz beschrieben.

**SR Algorithmus** [5] Berechnet eine kondensiertere Form als die Hamiltonische Schurform, verwendet aber nicht-orthogonale Ähnlichkeitstransformationen und ist deshalb nicht numerisch rückwärts stabil.

**Hamiltonischer QR Algorithmus** [6] Berechnet die Hamiltonische Schurform mit einem strukturerhaltenden QR Algorithmus, ist aber nur effizient, (d.h.  $\mathcal{O}(n^3)$  Rechenoperationen), wenn  $\text{rank}(G) = 1$  oder  $\text{rank}(Q) = 1$ .

**Multishift-Algorithmus** [2, 1] Berechnet eine Hamiltonische Block-Schurform, d.h.  $H_{11}$  ist nicht in Schurform, aber  $\Lambda(H_{11}) = \{\lambda_1, \dots, \lambda_n\}$ . Allerdings gibt es hier Probleme mit der Konvergenz für steigendes  $n$ , so daß der Algorithmus nur für  $n < 100$  sinnvoll eingesetzt werden kann.

**Einbettungsalgorithmus** [3] Überführt das Hamiltonische Eigenwertproblem in ein Hamiltonisches Eigenwertproblem doppelter Größe mit zusätzlicher Struktur, für die die

Hamiltonische Schurform strukturerhaltend und effizient ausgerechnet werden kann. Aus dem stabilen invarianten Unterraum der Hamiltonischen Matrix in  $\mathbb{R}^{4n \times 4n}$  kann man die Lösung der ARE berechnen.

**Hamiltonischer Jacobi-Algorithmus** [8, 4] Berechnet die Hamiltonsche Schurform, allerdings ist die Konvergenz der Jacobi-Iteration meist viel zu langsam, so daß der Algorithmus ebenfalls nicht effizient ist.

**Signumfunktionsmethode** [27, 7] Die Idee ist, iterativ einen Projektor auf den gesuchten  $H$ -invarianten Unterraum zu berechnen und daraus die Lösung der ARE zu extrahieren. Wir hatten bereits gesehen<sup>4</sup>, dass die Signumfunktion einer Matrix zur Lösung von Matrixgleichungen herangezogen werden kann. Dies trifft auch auf AREs zu. Da

$$P_- := \frac{1}{2}(I_{2n} - \text{sign}(H))$$

ein Projektor auf den stabilen  $H$ -invarianten Unterraum  $\mathcal{V}_-$  ist, ist

$$P_+ := I - P_- = \frac{1}{2}(I_{2n} + \text{sign}(H))$$

der Projektor auf den komplementären *antistabilen*  $H$ -invarianten Unterraum  $\mathcal{V}_+$  d.h. den  $H$ -invarianten Unterraum zu  $\{-\lambda_1, \dots, -\lambda_n\}$ . Da  $\mathcal{V}_- \cap \mathcal{V}_+ = \{0\}$  und nach Satz 4.18 und den Vorüberlegungen die Spalten von  $\begin{bmatrix} I_n \\ X_* \end{bmatrix}$  den Unterraum  $\mathcal{V}_-$  aufspannen, gilt  $\begin{bmatrix} I_n \\ X_* \end{bmatrix} \in \text{Kern}(P_+)$ , also

$$0 = P_+ \begin{bmatrix} I_n \\ X_* \end{bmatrix} = \frac{1}{2}(I_{2n} + \text{sign}(H)) \begin{bmatrix} I_n \\ X_* \end{bmatrix}. \quad (4.42)$$

Dies ist ein überbestimmtes, aber konsistentes lineares Gleichungssystem. Mit der Partitionierung  $\text{sign}(H) = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$  erhält man  $X_*$  nach Multiplikation von (4.42) mit 2 als Lösung des linearen Ausgleichsproblems

$$\begin{bmatrix} W_{12} \\ W_{22} + I_n \end{bmatrix} X_* = - \begin{bmatrix} W_{11} + I_n \\ W_{21} \end{bmatrix}. \quad (4.43)$$

$\text{sign}(H)$  ergibt sich als Grenzwert der Newton-Iteration

$$H_0 \leftarrow H, \quad H_{j+1} \leftarrow \frac{1}{2\gamma_j}(H_j + \gamma_j^2 H_j^{-1}), \quad j = 0, 1, \dots \quad (4.44)$$

Dabei ist  $\gamma_j$  ein Parameter, mit dem die Konvergenz der Iteration beschleunigt werden kann. Dazu schlägt Byers in [7]

$$\gamma_j = |\det(Z_j)|^{\frac{1}{n}} \quad (4.45)$$

---

<sup>4</sup>WS08/09: Ü4, H17b).



vor, was sich für die Lösung von AREs bewährt hat. Da die Summe Hamiltonischer Matrizen Hamiltonisch ist und die Inverse einer Hamiltonischen Matrix (falls sie existiert) ebenfalls, sind die Iterierten  $H_j$  in (4.44) alle Hamiltonisch. In endlicher Arithmetik gilt dies aufgrund der Rundungsfehler jedoch i.a. nicht mehr. Hier hilft ein simpler Trick: Da  $H_j J$  nach Definition 4.8 symmetrisch ist, schreibt man

$$H_j^{-1} = (H_j J J^T)^{-1} = J(H_j J)^{-1}.$$

Verwendet man zur Berechnung der Inversen  $(H_j J)^{-1}$  eine Methode, die die Symmetrie der Inversen garantiert, z.B. die  $LDL^T$ -Zerlegung, so kann man die Hamiltonische Struktur von  $H_j^{-1}$  auch bei numerischer Rechnung exakt garantieren, was natürlich auch für die Summe gilt. Damit wird die Iteration (4.44) zu einer strukturerhaltenden Methode.

Allerdings gibt es Konvergenzprobleme, falls  $H$  Eigenwerte nahe der imaginären Achse besitzt, da die Signumfunktion von  $H$  nicht definiert ist, wenn  $H$  rein-imaginäre Eigenwerte hat.

Ein völlig anderer Ansatz liegt darin, die ARE (4.28) als ein Nullstellenproblem für die nichtlineare Funktion  $\mathcal{R}(X)$  aufzufassen. Dazu bietet es sich an, das Newtonverfahren anzuwenden, wobei man sicherstellen muß, daß man gegen die richtige Lösung konvergiert. Dies kann man aber recht leicht erreichen, wie wir im Folgenden sehen werden. Das Newtonverfahren für ein gegebenes  $X_0 = X_0^T \in \mathbb{R}^{n \times n}$  besteht aus der Iteration

$$X_{k+1} = X_k - (\mathcal{R}'_{X_k})^{-1}(\mathcal{R}(X_k)), \quad k = 0, 1, \dots \quad (4.46)$$

Hierbei ist  $\mathcal{R}'_X$  die Fréchet-Ableitung von  $\mathcal{R}$  in Richtung  $X$ . Diese ist gegeben durch

$$\begin{aligned} \mathcal{R}'_X(Z) &= \lim_{h \rightarrow \infty} \frac{1}{h} (\mathcal{R}(X + hZ) - \mathcal{R}(X)) \\ &= \lim_{h \rightarrow \infty} \frac{1}{h} (Q + A^T(X + hZ) + (X + hZ)A^T - (X + hZ)G(X + hZ) - \\ &\quad - (Q + A^T X + X A - X G X)) \\ &= \lim_{h \rightarrow \infty} \frac{1}{h} (h A^T Z + h Z A - h X G Z - h Z G X - h^2 Z G Z) \\ &= (A - G X)^T Z + Z(A - G X). \end{aligned}$$

Definiert man  $N_k := -(\mathcal{R}'_{X_k})^{-1}(\mathcal{R}(X_k))$ , so folgt, daß  $N_k$  die Lyapunovgleichung

$$(A - G X_k)^T N_k + N_k(A - G X_k) = -\mathcal{R}(X_k) \quad (4.47)$$

löst. Die Lösung von (4.47) kann man z.B. mit dem Bartels-Stewart Algorithmus (Algorithmus 3.15) berechnen.

Unter der Annahme, daß die Iteration (4.46) gegen  $\hat{X}$  konvergiert und daß alle Iterierten  $\Lambda(A - G X_k) \subset \mathbb{C}^-$  erfüllen, folgt  $\Lambda(A - G \hat{X}) \subset \mathbb{C}^- \cup i\mathbb{R}$ . Hat die zugehörige Hamiltonische

Matrix aber keine Eigenwerte auf der imaginären Achse, so muß  $\Lambda(A - G\hat{X}) \subset \mathbb{C}^-$  gelten, d.h.  $\hat{X} = X_*$  muß die stabilisierende Lösung sein. Die Konvergenz des Newtonverfahrens für die ARE (4.28) ist in folgendem Satz zusammengefaßt.

**Satz 4.24** *Es sei  $G \geq 0$ ,  $(A, G)$  stabilisierbar und es existiere die stabilisierende Lösung  $X_*$  der ARE (4.28). Ist  $X_0 = X_0^T$  so gewählt, daß  $\Lambda(A - GX_0) \subset \mathbb{C}^-$ , dann gilt für die durch (4.46) erzeugte Folge  $\{X_k\}_{k=0}^\infty$ :*

(i) *Alle  $X_k$  sind stabilisierend, d.h.  $\Lambda(A - GX_k) \subset \mathbb{C}^-$  für alle  $k \in \mathbb{N}_0$ .*

(ii)  *$X_* \leq \dots \leq X_{k+1} \leq X_k \leq \dots \leq X_1$ .*

(iii)  *$\lim_{k \rightarrow \infty} X_k = X_*$ .*

(iv) *Es existiert eine Konstante  $\gamma > 0$ , so daß*

$$\|X_{k+1} - X_*\| \leq \gamma \|X_k - X_*\|^2, \quad k \geq 1,$$

*d.h. die  $X_k$  konvergieren quadratisch.*

Für den Beweis folgen wir i.W. [18]. Zunächst beweisen wir jedoch ein Lemma, welches Teil (i) des Satzes liefert.

**Lemma 4.25** *Sei  $N_k$  die Lösung der Lyapunovgleichung (4.47). Ist  $G \geq 0$  und  $A - GX_k$  stabil<sup>5</sup>, dann gilt für alle  $t \in [0, 2]$ , dass  $A - G(X_k + tN_k)$  stabil ist.*

**Beweis:** Mit  $A_k := A - GX_k$  ist die Lyapunovgleichung (4.47) äquivalent zu

$$A_k^T(X_k + N_k) + (X_k + N_k)A_k = -Q - X_kGX_k. \quad (4.48)$$

Subtrahiert man (4.48) von  $\mathcal{R}_c(X_*) = 0$  und subtrahiert  $X_kGX_* + X_*GX_k$  auf beiden Seiten, ergibt sich

$$A_k^T(X_* - (X_k + N_k)) + (X_* - (X_k + N_k))A_k = (X_* - X_k)G(X_* - X_k). \quad (4.49)$$

Da  $A_k$  stabil ist und  $G \geq 0$ , folgt aus dieser Lyapunovgleichung mit Satz 3.6, dass

$$X_* - (X_k + N_k) \leq 0. \quad (4.50)$$

Subtrahiert man nun auf beiden Seiten von (4.49)

$$tN_kG(X_* - (X_k + N_k)) + (X_* - (X_k + N_k))GtN_k,$$

so erhalten wir

$$\begin{aligned} & (A - G(X_k + tN_k))^T(X_* - (X_k + N_k)) + (X_* - (X_k + N_k))(A - G(X_k + tN_k)) = \\ & = (X_* - (X_k + tN_k))G(X_* - (X_k + tN_k)) + t(2-t)N_kGN_k \\ & =: W. \end{aligned} \quad (4.51)$$

---

<sup>5</sup>D.h.,  $\Lambda(A - GX_k) \subset \mathbb{C}^-$ , also eigentlich asymptotisch stabil.

Da  $t \in [0, 2]$ , ist  $W$  in (4.51) positiv semidefinit.

Nehmen wir nun an,  $A - G(X_k + tN_k)$  hätte einen Eigenwert  $\lambda$  mit  $\operatorname{Re}(\lambda) \geq 0$  und zugehörigem Eigenvektor  $z \neq 0$ , d.h.

$$(A - G(X_k + tGN_k))z = \lambda z \text{ und } z^H(A - G(X_k + tGN_k))^T = \bar{\lambda}z^H. \quad (4.52)$$

Multiplikation von (4.51) von links mit  $z^H$  und von rechts mit  $z$  ergibt

$$2 \cdot \operatorname{Re}(\lambda)z^H(X_* - (X_k + N_k))z = z^HWz. \quad (4.53)$$

Die linke Seite von (4.53) ist nichtpositiv, da nach (4.50)  $X_* - (X_k + N_k) \leq 0$  und  $\operatorname{Re}(\lambda) \geq 0$ . Da  $W \geq 0$ , ist die rechte Seite von (4.53) nicht-negativ und es folgt  $z^HWz = 0$  sowie

$$z^H(X_* - (X_k + tN_k))G(X_* - (X_k + tN_k))z = 0.$$

Wegen  $G = G^T \geq 0$  ist  $G(X_* - (X_k + tN_k))z = 0$ , bzw.

$$GX_*z = G(X_k + tN_k)z. \quad (4.54)$$

Aus (4.52) und (4.54) erhalten wir somit

$$\lambda z = (A - G(X_k + tN_k))z = (A - GX_*)z.$$

Also muss  $\lambda$  ein Eigenwert von  $A - GX_*$  sein, was der Stabilität von  $A - GX_*$  widerspricht.  $\square$

Eigentlich benötigen wir das Lemma nur für  $t = 1$ , was dem Newton-Schritt  $X_{k+1} = X_k + N_k$  entspricht. Mit obigem Lemma lässt sich aber ein gedämpftes Newton-Verfahren bzw. ein Liniensuchverfahren zur Lösung von AREs definieren.

Eine direkte Konsequenz des vorangegangenen Lemmas ist

**Folgerung 4.26** *Ist  $A - GX_0$  stabil so ist unter den Voraussetzungen von Lemma 4.25 der Lyapunov-Operator  $\mathcal{R}'_k$  invertierbar für alle  $k \in \mathbb{N}_0$ .*

**Beweis:** von Satz 4.24 Teil (i) haben wir mit Lemma 4.25 bewiesen. Teil (ii) ergibt sich einerseits aus (4.50), denn damit folgt  $X_* \leq X_k$  für all  $k \geq 1$ . Die weiteren Ungleichungen folgen wie im Beweis von Lemma 4.25, wenn man zwei aufeinanderfolgende Newtonschritte in der Form (4.48) (für  $k$  und  $k - 1$ ) voneinander abzieht. Damit erhält man

$$\begin{aligned} A_k^T(X_{k+1} - X_k) + (X_{k+1} - X_k)A_k &= -X_kGX_k + X_{k-1}GX_{k-1} + N_{k-1}GX_k + X_kGN_{k-1} \\ &= N_{k-1}GN_{k-1}. \end{aligned} \quad (4.55)$$

Da  $A_k$  nach Lemma 4.25 stabil ist für alle  $k$  und  $G \geq 0$  folgt mit Satz 3.6, dass  $X_{k+1} - X_k \leq 0$ , also  $X_{k+1} \leq X_k$  für alle  $k \geq 1$ , also (ii). (Beachte: (4.55) kann nur für  $k \geq 1$  gebildet werden.)

Nun ist  $\{X_k\}_{k=1}^\infty$  nach (ii) eine monoton fallende und nach unten beschränkte Folge. Damit existiert der Grenzwert  $\hat{X} := \lim_{k \rightarrow \infty} X_k$ . Da  $A_k = A - GX_k$  stabil für alle  $k$  nach Lemma 4.25 und die Eigenwerte einer Matrix stetig von den Matrixeinträgen abhängen, gilt  $\Lambda(A - G\hat{X}) \subset \mathbb{C}^- \cup i\mathbb{R}$ . Bildet man den Grenzwert auf beiden Seiten in (4.48), so folgt dass  $\hat{X}$  die ARE (4.28) löst. Also bilden die Spalten von  $\begin{bmatrix} I_n \\ \hat{X} \end{bmatrix}$  einen  $H$ -invarianten Unterraum zu  $n$  Eigenwerten in der abgeschlossenen linken komplexen Halbebene. Aufgrund der Existenz der stabilisierenden Lösung und der Tatsache, dass  $\Lambda(A - GX_*) \subset \Lambda(H)$  folgt, dass  $H$  keine rein-imaginären Eigenwerte haben kann. Also ist  $\Lambda(A - G\hat{X}) \subset \mathbb{C}^-$  und  $\hat{X}$  stabilisierend. Da die stabilisierende Lösung, falls sie existiert, immer eindeutig ist, muss  $\hat{X} = X_*$  gelten und damit ist (iii) bewiesen. Dabei beachte, dass der Eindeutigkeitsbeweis in Satz 4.18 auch ohne die Voraussetzungen dieses Satz funktioniert, wenn wie hier die Existenz der stabilisierenden Lösung vorausgesetzt wird.

Um (iv), also die quadratische Konvergenz zu zeigen, betrachte (4.49). Wegen Folgerung 4.26 können wir diese Identität auch schreiben als

$$X_* - X_{k+1} = (\mathcal{R}'_{X_k})^{-1}((X_* - X_k)G(X_* - X_k)).$$

Nun sei  $\|\cdot\|$  eine konsistente Norm. Damit folgt

$$\|X_* - X_{k+1}\| \leq \|(\mathcal{R}'_{X_k})^{-1}\| \|G\| \|X_* - X_k\|^2.$$

Mit  $\lim_{k \rightarrow \infty} A_k = A - GX_*$  existiert auch der Grenzwert  $\lim_{k \rightarrow \infty} \mathcal{R}'_{X_k} = \mathcal{R}'_{X_*}$  und damit  $\lim_{k \rightarrow \infty} (\mathcal{R}'_{X_k})^{-1} = (\mathcal{R}'_{X_*})^{-1}$ . Sei nun  $\rho_* := \|(\mathcal{R}'_{X_*})^{-1}\|$ ,  $\rho_k := \|(\mathcal{R}'_{X_k})^{-1}\|$ , dann gilt auch  $\lim_{k \rightarrow \infty} \rho_k = \rho_*$ . Mithin existiert eine Konstante  $\delta \geq 0$  und ein  $k_0 = k_0(\delta)$  so dass  $|\rho_* - \rho_k| < \delta$  für alle  $k \geq k_0$ . Damit folgt (iv) mit

$$\gamma = \|G\| \max\{\rho_* + \delta, \max_{0 \leq k < k_0} \rho_k\}.$$

Die Allgemeinheit der Aussage für beliebige Matrixnormen folgt mit der Normäquivalenz auf  $\mathbb{R}^{n \times n}$ .  $\square$

Die Kosten des Newtonverfahrens hängen natürlich von der Anzahl der benötigten Iterationen ab. Man kann hier verschiedene Abbruchkriterien verwenden, z.B.

$$\begin{aligned} \|N_k\| = \|X_{k+1} - X_k\| &\leq \tau \cdot \|X_k\|, \\ \|\mathcal{R}(X_k)\| &\leq \tau \end{aligned}$$

für eine vorgegebene Fehlertoleranz  $\tau$ . Bei 6–7 Iterationen entsprechen die Kosten in etwa den Kosten der Schurvektormethode. Oft werden aber mehr Iterationen benötigt, so daß man das Newtonverfahren meist nur zur Verbesserung einer mit einem anderen Verfahren berechneten Näherungslösung einsetzt. Mindestens ein Schritt des Newtonverfahrens sollte immer angewendet werden, da man mit dem Newtonverfahren den kleinstmöglichen relativen Fehler

$$\frac{\|X_* - \tilde{X}\|}{\|X_*\|} \approx \mathbf{u} \cdot \kappa_{ARE}$$

erreichen kann. Hierbei ist  $\tilde{X}$  die mit dem Newtonverfahren berechnete Näherungslösung und  $\kappa_{ARE}$  die Kondition der ARE.

# Kapitel 5

## Beobachterentwurf, Kalman-Filter und LQG Regelung

In diesem Kapitel geht es darum, den möglicherweise bei einem Regelungssystem nicht messbaren Zustand  $x(t)$  mit Hilfe des gemessenen Ausgangs  $y(t)$  zu schätzen. Ein solcher Schätzer wird allgemein Beobachter genannt. Da in der Praxis gemessene Signale in aller Regel mit einem Rauschen behaftet sind, ist es sinnvoll, dieses gleich mit zu modellieren und den Beobachter so zu entwerfen, dass der Erwartungswert des Schätzfehlers minimiert wird. Dies führt dann zum Kalman(-Bucy)-Filter.

Da der Entwurf von Beobachtern sowie des Kalman-Bucy-Filters für LTI Systeme als Spezialfall beim LTV Fall abfällt, wollen wir hier diesen allgemeineren Fall betrachten.

Zunächst wollen wir definieren, was wir unter einem Beobachter verstehen. Dazu betrachten wir zunächst das LTV System

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad (5.1a)$$

$$y(t) = C(t)x(t), \quad (5.1b)$$

d.h. die Durchgriffsmatrix erfüllt  $D(t) = 0$ .

**Definition 5.1** Ein (Zustands-)Beobachter für (5.1) ist eine Funktion

$$\hat{x} : [0, \infty) \mapsto \mathbb{R}^n,$$

für die der Schätzfehler  $e(t) = x(t) - \hat{x}(t)$

$$\lim_{t \rightarrow \infty} e(t) = 0 \quad \text{für alle } u \in \mathcal{U}_{ad} \text{ und alle } x^0 \in \mathbb{R}^n$$

erfüllt.

Da der Beobachter  $\hat{x}(t)$  die Trajektorie eines dynamischen Systems verfolgen soll, bietet es sich an, den Beobachter selbst als solches zu definieren. Dies wurde von Luenberger 1966 in [?] vorgeschlagen und der resultierende Beobachter gilt als der klassische Beobachter in der System- und Regelungstheorie, der vielfach angewendet und verallgemeinert wurde.

**Definition 5.2** Der (vollständige) **Luenberger-Beobachter** ist definiert als die Lösung des dynamischen Systems

$$\dot{\hat{x}}(t) = H(t)\hat{x}(t) + F(t)y(t) + G(t)u(t), \quad \hat{x}(0) = \hat{x}^0, \quad (5.2)$$

wobei  $H : [0, \infty) \mapsto \mathbb{R}^{n \times n}$ ,  $F : [0, \infty) \mapsto \mathbb{R}^{n \times p}$ , und  $G : [0, \infty) \mapsto \mathbb{R}^{n \times m}$  die folgenden Bedingungen erfüllen:

$$H(t) = A(t) - L(t)C(t), \quad F(t) = L(t), \quad G(t) = B(t).$$

Dabei wird  $L : [0, \infty) \mapsto \mathbb{R}^{n \times p}$  so gewählt, dass  $z(t) \equiv 0$  eine asymptotisch stabile Trajektorie für  $\dot{z}(t) = H(t)z(t)$  ist.

**Satz 5.3** Der Luenberger-Beobachter ist ein Zustandsbeobachter.

**Beweis:** Die Beobachtergleichung für den Luenberger-Beobachter lautet

$$\dot{\hat{x}}(t) = (A(t) - L(t)C(t))\hat{x}(t) + L(t)y(t) + B(t)u(t) \quad (5.3a)$$

$$= A(t)\hat{x}(t) + B(t)u(t) + L(t)(y(t) - C(t)\hat{x}(t)), \quad \hat{x}(0) = \hat{x}^0. \quad (5.3b)$$

Für den Schätzfehler  $e(t) = x(t) - \hat{x}(t)$  gilt damit:

$$\begin{aligned} \dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\ &= A(t)x(t) + B(t)u(t) - ((A(t) - L(t)C(t))\hat{x}(t) + L(t)y(t) + B(t)u(t)) \\ &= (A(t) - L(t)C(t))(x(t) - \hat{x}(t)) + \underbrace{L(t)C(t)x(t) - L(t)y(t)}_{=0} \\ &= (A(t) - L(t)C(t))e(t). \end{aligned}$$

Da  $L(t)$  so gewählt war, dass diese lineare, homogene Differentialgleichung asymptotisch stabil ist, gilt  $\lim_{t \rightarrow \infty} e(t) = 0$ .  $\square$

Die Schwierigkeit im zeitvarianten Fall besteht darin, die stabilisierende Funktion  $L(t)$  anzugeben. Es wird dabei oft übersehen, dass es nicht ausreicht,  $L$  so zu wählen, daß  $\Lambda(A(t) - L(t)C(t)) \subset \mathbb{C}^-$ , da bekanntermaßen für nichtautonome, lineare homogene Differentialgleichungen die punktweise Stabilität der Koeffizientenmatrix nicht hinreichend für asymptotische Stabilität ist.

Im LTI Fall ist es jedoch ausreichend,  $L$  stabilisierend zu wählen, wie folgender Satz zeigt.

**Satz 5.4** Ist  $(A, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{p \times n}$  beobachtbar, dann existiert ein Beobachter in der Form

$$\dot{\hat{x}}(t) = (A - LC)\hat{x}(t) + Ly(t) + Bu(t), \quad (5.4)$$

wobei  $L \in \mathbb{R}^{n \times p}$  so gewählt wird, daß  $A - LC$  Hurwitz ist.

**Beweis:** Nach Folgerung 2.20 folgt aus der Beobachtbarkeit von  $(A, C)$  die Steuerbarkeit von  $(A^T, C^T)$ . Damit existiert nach Satz 3.20 ein  $F \in \mathbb{R}^{p \times n}$ , so daß  $\Lambda(A^T + C^T F) \subset \mathbb{C}^-$ . Also gilt mit  $L := -F^T$ , daß  $\Lambda(A - LC) \subset \mathbb{C}^-$ .

Für den Schätzfehler folgt damit

$$\begin{aligned}\dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\ &= Ax(t) + Bu(t) - (A - LC)\hat{x}(t) - Ly(t) - Bu(t) \\ &= (A - LC)e(t)\end{aligned}$$

und damit  $e(t) = e^{(A-LC)t}e(0)$ , also  $\lim_{t \rightarrow \infty} e(t) = 0$  für beliebige  $u \in \mathcal{U}_{ad}$  und alle Startwerte  $x^0, \hat{x}^0$ .  $\square$

Ein konstruktiver Weg, einen Beobachter, den man im LTI Fall ebenfalls als *Luenberger-Beobachter* bezeichnet, zu konstruieren, liefert folgender Satz.

**Satz 5.5** Die Funktion  $\hat{x}(t) = Zz(t)$  mit  $z$  Lösung der Differentialgleichung

$$\dot{z}(t) = Hz(t) + Ly(t) + Gu(t) \quad (5.5)$$

ist ein Beobachter im Sinne von Definition 5.1, falls

(i)  $Z$  invertierbar ist und  $X := Z^{-1}$  die Sylvester-Beobachtergleichung

$$XA - HX = LC \quad (5.6)$$

erfüllt;

(ii)  $G = XB$ ;

(iii)  $H$  Hurwitz ist.

**Beweis:** Betrachte  $\tilde{e}(t) := Xe(t) = Xx(t) - z(t)$ . Diese erfüllt

$$\begin{aligned}\dot{\tilde{e}}(t) &= X\dot{x}(t) - \dot{z}(t) \\ &= XAx(t) + XBu(t) - Hz(t) - Ly(t) - Gu(t) \\ &= \underbrace{(XA - HX - LC)}_{\stackrel{(i)}{=}0} x(t) + H(Xx(t) - z(t)) + \underbrace{(XB - G)}_{\stackrel{(ii)}{=}0} u(t) \\ &= H\tilde{e}(t).\end{aligned}$$

Wegen (iii) folgt wie im Beweis von Satz 5.4  $\lim_{t \rightarrow \infty} \tilde{e}(t) = 0$  und damit auch

$$\lim_{t \rightarrow \infty} e(t) = \lim_{t \rightarrow \infty} Z\tilde{e}(t) = Z \lim_{t \rightarrow \infty} \tilde{e}(t) = 0.$$

$\square$

Dieser Satz erlaubt die Konstruktion von Beobachtern wie z.B. in folgendem Beispiel.

**Beispiel 5.6 (Beispiel 1.6, fortgesetzt)** Das inverse Pendel aus der Einleitung führte nach Linearisierung auf ein LTI System mit

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \ 0], \quad D = 0.$$

Wählt man in (5.5)  $H = \text{diag}(-2, -\frac{1}{2})$  und  $L = [2 \ 1]^T$ , so ergibt die Lösung von (5.6)

$$X = \frac{1}{3} \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}.$$

Da  $X$  invertierbar ist, erhält man dann den Luenberger-Beobachter mit

$$G = XB = \frac{1}{3} \begin{bmatrix} -2 \\ 4 \end{bmatrix}.$$

Da die Qualität des Beobachters direkt vom gemessenen Ausgangssignal  $y(t)$  sowie dem vorgegeben Eingangssignal und  $u(t)$  abhängt und kleine Abweichungen möglicherweise zu einem großen Abweichen des geschätzten Zustands führen kann, ist es robuster und der Realität angemessener, Störungen in den Eingangs- und Ausgangssignalen gleich mit zu modellieren. Dies geschieht durch Hinzunahme von Rauschtermen in der Zustandsgleichung (5.1a) und der Ausgangsgleichung (5.1b). Es bietet sich dann an, auch den Startwert als verrauscht anzusehen, da auch dieser in der Praxis meist als Messung vorliegt. D.h., wir betrachten folgendes System

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + F(t)v(t), \quad x(t_0) = x^0 + \eta, \quad (5.7a)$$

$$y(t) = C(t)x(t) + w(t), \quad (5.7b)$$

wobei  $v$  und  $w$  als *Weißes Rauschen* angenommen werden, d.h. als stationäre stochastische Prozesse von unkorrelierten Zufallsvariablen mit Erwartungswert Null, es gilt also für alle  $t, s \in \mathbb{R}$

$$\begin{aligned} \mathcal{E}[v(t)] &= 0, & \mathcal{E}[w(t)] &= 0, \\ \mathcal{E}[v(t)v^T(s)] &= V\delta(t-s), \\ \mathcal{E}[w(t)w^T(s)] &= W\delta(t-s), \end{aligned}$$

wobei  $V = V^T > 0$  und  $W = W^T > 0$  die Kovarianzmatrizen von  $v$  bzw.  $w$  sind und  $\delta(t-s)$  die Diracsche  $\delta$ -Distribution bezeichnet. Außerdem nehmen wir an, dass  $v$  und  $w$  unkorreliert sind, also

$$\mathcal{E}[v(t)w^T(s)] = 0 \quad \forall t, s.$$

Die Kovarianzmatrix der normalverteilten Zufallsvariable  $\eta$  mit Erwartungswert Null sei  $\Sigma_0 \in \mathbb{R}^{n \times n}$ , also

$$\mathcal{E}[\eta\eta^T] = \Sigma_0.$$



Verwendet man nun einen Beobachter in der bereits zuvor definierten Form (siehe (5.7))

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + B(t)u(t) + L(t)(y(t) - C(t)\hat{x}(t)), \quad (5.8)$$

so ist das Ziel, den Erwartungswert des Schätzfehlers zu minimieren,  $L$  zu bestimmen durch

$$\min_L \lim_{t \rightarrow \infty} \mathcal{E}[e(t)^T e(t)] \quad \text{with} \quad e(t) = x(t) - \hat{x}(t). \quad (5.9)$$

Die Lösung dieser Aufgabe liefert der Kalman-(Bucy)-Filter:

**Satz 5.7** *Unter den o.g. Voraussetzungen an die stochastischen Prozesse  $v(t), w(t)$  und die Zufallsvariable  $\eta$  wird das Minimum in (5.9) angenommen für  $L(t)$  definiert durch*

$$L(t) = \Sigma_*(t)C(t)^T W^{-1}, \quad (5.10)$$

wobei  $\Sigma_*(t)$  die positiv definite Lösung der (Filter)-Riccati-Differentialgleichung (F)RDE

$$\dot{\Sigma}(t) = A(t)\Sigma(t) + \Sigma(t)A(t)^T - \Sigma(t)C(t)^T W^{-1} C(t)\Sigma(t) + F(t)VF(t)^T, \quad (5.11)$$

$$\Sigma(0) = \Sigma_0 \quad (5.12)$$

ist.

**Beweis:** Fehlt noch. □

Filter-ARE im LTI-Fall:

$$A\Sigma + \Sigma A^T - \Sigma C^T W^{-1} C \Sigma + F V F^T,$$

LQG Entwurf (für  $S = 0, D = 0$ ): G steht dabei für Gauß, da die stochastischen Prozesse i.d.R. als Gaußsches Weißes Rauschen angenommen werden, d.h. die Zufallsvariablen genügen der Gaußverteilung.

$$u(t) = -R^{-1}B(t)^T X_*(t)\hat{x}(t),$$

wobei  $\hat{x}$  in (5.8) definiert ist und  $L(t)$  wie in (5.10), also

$$\dot{\hat{x}}(t) = (A(t) - B(t)R^{-1}B(t)^T X_*(t))\hat{x}(t) + \Sigma_*(t)C(t)^T W^{-1} (y(t) - C(t)\hat{x}(t)),$$



# Literaturverzeichnis

- [1] G. AMMAR, P. BENNER, AND V. MEHRMANN, *A multishift algorithm for the numerical solution of algebraic Riccati equations*, *Electr. Trans. Num. Anal.*, 1 (1993), pp. 33–48.
- [2] G. AMMAR AND V. MEHRMANN, *On Hamiltonian and symplectic Hessenberg forms*, *Linear Algebra Appl.*, 149 (1991), pp. 55–72.
- [3] P. BENNER, V. MEHRMANN, AND H. XU, *A new method for computing the stable invariant subspace of a real Hamiltonian matrix*, *J. Comput. Appl. Math.*, 86 (1997), pp. 17–43.
- [4] A. BUNSE-GERSTNER AND H. FASSBENDER, *A Jacobi-like method for solving algebraic Riccati equations on parallel computers*, *IEEE Trans. Automat. Control*, 42 (1997), pp. 1071–1084.
- [5] A. BUNSE-GERSTNER AND V. MEHRMANN, *A symplectic QR-like algorithm for the solution of the real algebraic Riccati equation*, *IEEE Trans. Automat. Control*, AC-31 (1986), pp. 1104–1113.
- [6] R. BYERS, *A Hamiltonian QR-algorithm*, *SIAM J. Sci. Statist. Comput.*, 7 (1986), pp. 212–229.
- [7] ———, *Solving the algebraic Riccati equation with the matrix sign function*, *Linear Algebra Appl.*, 85 (1987), pp. 267–279.
- [8] ———, *A Hamiltonian-Jacobi algorithm*, *IEEE Trans. Automat. Control*, 35 (1990), pp. 566–570.
- [9] E. CAMACHO AND C. BORDONS, *Model Predictive Control*, *Advanced Textbooks in Control and Signal Processing*, Springer-Verlag, London, 2nd ed., 2004.
- [10] C. CHOI AND A. LAUB, *Efficient matrix-valued algorithms for solving stiff Riccati differential equations*, *IEEE Trans. Automat. Control*, 35 (1990), pp. 770–776.
- [11] L. DIECI, *Numerical integration of the differential Riccati equation and some related issues*, *SIAM J. Numer. Anal.*, 29 (1992), pp. 781–815.

- [12] S. GODUNOV, *Ordinary Differential Equations with Constant Coefficient*, vol. 169 of Translations of Mathematical Monographs, AMS, Providence, RI, 1997.
- [13] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, third ed., 1996.
- [14] R. HORN AND C. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [15] J. KAUTSKY, N. NICHOLS, AND P. VAN DOOREN, *Robust pole assignment in linear state feedback*, Internat. J. Control, 41 (1985), pp. 1129–1155.
- [16] C. KENNEY AND R. LEIPNIK, *Numerical integration of the differential matrix Riccati equation*, IEEE Trans. Automat. Control, AC-30 (1985), pp. 962–970.
- [17] H. KNOBLOCH AND H. KWAKERNAAK, *Lineare Kontrolltheorie*, Springer-Verlag, Berlin, 1985. In German.
- [18] P. LANCASTER AND L. RODMAN, *The Algebraic Riccati Equation*, Oxford University Press, Oxford, 1995.
- [19] P. LANCASTER AND M. TISMENETSKY, *The Theory of Matrices*, Academic Press, Orlando, 2nd ed., 1985.
- [20] A. LAUB, *A Schur method for solving algebraic Riccati equations*, IEEE Trans. Automat. Control, AC-24 (1979), pp. 913–921.
- [21] W.-W. LIN, V. MEHRMANN, AND H. XU, *Canonical forms for Hamiltonian and symplectic matrices and pencils*, Linear Algebra Appl., 301–303 (1999), pp. 469–533.
- [22] J. MACKI AND A. STRAUSS, *Introduction to Optimal Control Theory*, Springer-Verlag, 1982.
- [23] G. MIMINIS AND C. PAIGE, *An algorithm for pole assignment of time invariant linear systems*, Internat. J. Control, 35 (1982), pp. 341–354.
- [24] C. PAIGE AND C. VAN LOAN, *A Schur decomposition for Hamiltonian matrices*, Linear Algebra Appl., 41 (1981), pp. 11–32.
- [25] E. PINCH, *Optimal Control and the Calculus of Variations*, Oxford University Press, Oxford, UK, 1993.
- [26] L. PONTRYAGIN, V. BOLTYANSKII, R. GAMKRELIDZE, AND E. MISHENKO, *The Mathematical Theory of Optimal Processes*, Interscience, New York, 1962.
- [27] J. ROBERTS, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687. (Reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971).

- [28] E. SONTAG, *Mathematical Control Theory*, Springer-Verlag, New York, NY, 2nd ed., 1998.
- [29] J. STOER AND R. BULIRSCH, *Einführung in die Numerische Mathematik II*, Springer-Verlag, Berlin, 3. ed., 1990. In German.