



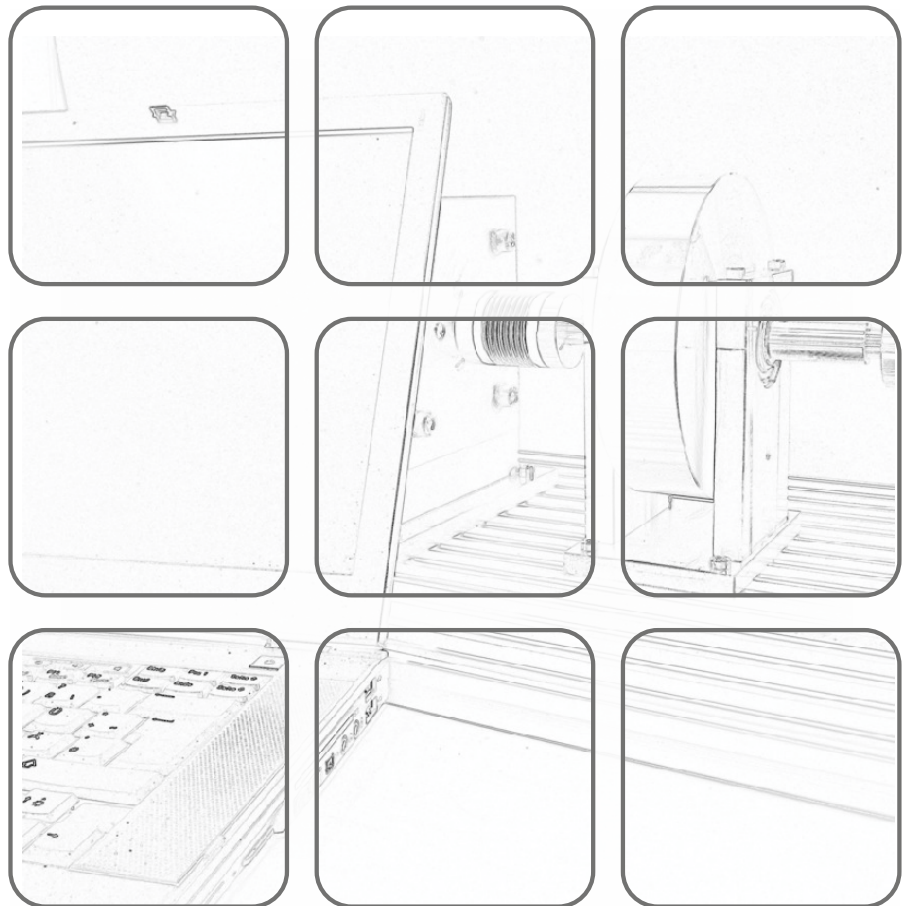
TECHNISCHE  
UNIVERSITÄT  
WIEN



# FORTGESCHRITTENE METHODEN DER NICHTLINEAREN REGELUNG

Vorlesung und Übung  
Wintersemester 2021/2022

Andreas Deutschmann-Olek, Andreas Kugi, Andreas Steinböck



## **Fortgeschrittene Methoden der nichtlinearen Regelung**

Vorlesung und Übung  
Wintersemester 2021/2022

Andreas Deutschmann-Olek, Andreas Kugi, Andreas Steinböck

TU Wien  
Institut für Automatisierungs- und Regelungstechnik  
Gruppe für komplexe dynamische Systeme

Gußhausstraße 27–29  
1040 Wien  
Telefon: +43 1 58801 – 37615  
Internet: <https://www.acin.tuwien.ac.at>

© Institut für Automatisierungs- und Regelungstechnik, TU Wien

# Inhaltsverzeichnis

<b>1</b>	<b>Iterativ Lernende Regelung</b>	<b>1</b>
1.1	Fixpunktiterationen . . . . .	2
1.2	Lifted-System-Darstellung . . . . .	5
1.3	Grundidee der ILR . . . . .	9
1.4	Stabilitäts- und Konvergenzanalyse . . . . .	12
1.5	ILR Entwurfsverfahren . . . . .	15
1.5.1	P-Type und PD-Type ILR . . . . .	15
1.5.2	Inversionsbasierte ILR . . . . .	23
1.5.3	Normoptimale ILR . . . . .	25
1.6	Literatur . . . . .	27
<b>2</b>	<b>Modellprädiktive Regelung</b>	<b>29</b>
2.1	Bestandteile und Grundidee von MPC . . . . .	30
2.1.1	Modell . . . . .	30
2.1.2	Horizonte . . . . .	31
2.1.3	Beschränkungen . . . . .	33
2.1.4	Skalares Gütemaß . . . . .	34
2.1.5	Optimierung . . . . .	35
2.1.6	Annahmen . . . . .	36
2.1.7	Regelgesetz . . . . .	36
2.2	Stabilität . . . . .	37
2.2.1	Prädiktionshorizont mit unendlicher Länge . . . . .	37
2.2.2	Endlicher Prädiktionshorizont mit vorgeschriebenem Endzustand . . . . .	40
2.2.3	Endlicher Prädiktionshorizont mit vorgeschriebenem Endgebiet und Endkostenterm . . . . .	44
2.2.4	Endlicher Prädiktionshorizont mit Endkostenterm . . . . .	47
2.2.5	Endlicher Prädiktionshorizont mit vorgeschriebenem Endgebiet . . . . .	51
2.3	Implementierung . . . . .	54
2.3.1	Entwurf eines stabilisierenden Zustandsreglers für ein Endgebiet . . . . .	54
2.3.2	Methoden zur Lösung von Optimalsteuerungsaufgaben . . . . .	55
2.4	Literatur . . . . .	58
<b>3</b>	<b>Zustandsschätzung auf bewegten Horizonten</b>	<b>63</b>
3.1	Bestandteile von MHE . . . . .	64
3.1.1	Modell . . . . .	64
3.1.2	Horizont . . . . .	64
3.1.3	Beschränkungen . . . . .	66
3.1.4	Skalares Gütemaß . . . . .	66

3.1.5	Optimierung . . . . .	66
3.1.6	Annahmen . . . . .	67
3.2	Stabilität von Zustandsschätzern . . . . .	68
3.3	Zustandsschätzung mit vollständiger Information . . . . .	69
3.4	Zustandsschätzung auf bewegtem Horizont . . . . .	73
3.4.1	Anfangskostenterm für vollständige Information . . . . .	74
3.4.2	Kein Anfangskostenterm . . . . .	75
3.4.3	Approximation der Anlaufkosten . . . . .	78
3.5	Maximum-a-posteriori Zustandsschätzung . . . . .	81
3.6	Zustands- und Parameterschätzung . . . . .	83
3.7	Literatur . . . . .	85
<b>4</b>	<b>Pfadfolgeregelung</b>	<b>87</b>
4.1	Definition von Pfaden . . . . .	88
4.1.1	Parametrierte Darstellung von Pfaden . . . . .	89
4.1.2	Implizite Definition von Pfaden . . . . .	90
4.1.3	Vergleich und Umrechnung der Darstellungen . . . . .	90
4.2	Begriffe und Ziele der Pfadfolgeregelung . . . . .	93
4.3	Pfadfolgeregelung basierend auf exakter Linearisierung . . . . .	94
4.3.1	Erweiterung der exakten Linearisierung für AI-Systeme . . . . .	94
4.3.2	Implizit definierte Pfade . . . . .	99
4.3.3	Parametrierte Pfade . . . . .	108
4.4	Modellprädiktive Pfadfolgeregelung . . . . .	117
4.5	Literatur . . . . .	126
<b>5</b>	<b>Dissipativität und Passivität</b>	<b>127</b>
5.1	Glühsimulator . . . . .	127
5.2	Einfaches Elektromagnetventil . . . . .	129
5.3	Systemtheoretisches Konzept . . . . .	130
5.3.1	Dissipativität . . . . .	130
5.3.2	Passivität . . . . .	131
5.3.3	Eigenschaften Passiver Systeme . . . . .	133
5.3.4	Passivität und Lyapunov-Stabilität . . . . .	135
5.4	Lineare passive Systeme . . . . .	136
5.5	Positive Reellheit . . . . .	139
5.6	Kanonische Form Passiver Systeme . . . . .	142
5.6.1	Hamiltonsche Systeme . . . . .	142
5.6.2	Port-Hamiltonsche Systeme . . . . .	144
5.7	Passivitätsbasierter Reglerentwurf . . . . .	146
5.8	Literatur . . . . .	153

# 1 Iterativ Lernende Regelung

*Iterativ Lernende Regelung* (ILR) (engl. Iterative Learning Control (ILC)) basiert auf der Idee, dass die Performance eines sich wiederholenden Prozesses verbessert werden kann, indem aus vorangegangenen Versuchen (Iterationen) gelernt wird. Die ILR nutzt Information vorangegangener Iterationen gezielt aus, um den Ausgang eines System einer vorgegebenen Trajektorie folgen zu lassen. Für diese Art von Prozessen würde ein nicht-lernender Regler in jeder Iteration immer den gleichen Regelfehler aufweisen.

Um die ILR anzuwenden, muss ein Prozess folgende Bedingungen erfüllen:

- (i.) Jede Iteration endet innerhalb einer festen Iterationsdauer, d. h.  $t \in [0, t_f]$ .
- (ii.) Jede Iteration startet vom (nahezu) gleichen Anfangszustand.
- (iii.) Für einen gegebenen Sollausgang  $y_d(t)$  existiert eine eindeutige Steuerung  $u_d(t)$ .

Bei der ILR wird in jeder Iteration  $j = 0, 1, \dots$  eine Steuerung  $u_{j+1}(t)$  auf Basis des Ausgangsfehlers  $e_j(t) = y_d(t) - y_j(t)$  berechnet und auf das System geschaltet. Ein zur Berechnung der Steuerung vielfach eingesetztes ILR-Gesetz lautet für  $L > 0$

$$u_{j+1}(t) = u_j(t) + L(y_d(t) - y_j(t)) . \quad (1.1)$$

Das ILR-Gesetz (1.1) stellt eine Fixpunktiteration für die Steuerung  $u_j(t)$  dar. Für  $j \rightarrow \infty$  resultiert eine Steuerung  $u_\infty(t)$ , die den Ausgangsfehler  $e_\infty(t) = y_d(t) - y_\infty(t) = 0$  zu Null regelt.

Die ILR zeichnet sich insbesondere dadurch aus, dass eine *hohe Regelgüte* erreicht werden kann, sie *robust gegenüber Modellunsicherheiten* ist und *sich wiederholende Störungen unterdrücken lassen*. Die ILR unterscheidet sich von anderen lernenden Regler Algorithmen, wie z. B. der Adaptiven Regelung. Bei der Adaptiven Regelung werden zum Erreichen der Regelgüte die Regelparameter modifiziert - bei der ILR wird eine Steuerung modifiziert. Man kann die ILR daher auch als ein Verfahren zur Adaption einer Steuerung auffassen. Die ILR weist einige Vorteile gegenüber einer klassischen Regelung auf. Mithilfe einer Regelung ist es im Allgemeinen schwierig Totzeiten zu kompensieren. Im Rahmen der ILR ist dies einfach möglich, da sie von Iteration zu Iteration arbeitet. Aus diesem Grund ist es auch möglich, akasale Filter für den ILR-Entwurf einzusetzen.

In der Literatur [1.1–1.5] werden zwei Arten der Lernenden Regelung unterschieden: *Repetitive Regelung* und *Iterativ Lernende Regelung*. Prinzipiell sind die Verfahren gleich, allerdings mit einem Unterschied.

- Bei der *Iterativ Lernenden Regelung* sind alle Iterationen vollständig unabhängig voneinander. Der Anfangszustand ist in jeder Iteration gleich.
- Bei der *Repetitiven Regelung* sind die Iteration abhängig voneinander. Der Endzustand des Systems am Ende einer Iteration bildet den Anfangszustand des Systems für die nächste Iteration. Der Anfangszustand der Iterationen variiert also von Iteration zu Iteration.

## 1.1 Fixpunktiterationen

Die Nullstellensuche einer Funktion

$$\gamma(\mathbf{z}) = \mathbf{0} \quad (1.2)$$

mit  $\gamma(\mathbf{z}) : \mathbb{R}^N \rightarrow \mathbb{R}^N$  und  $\mathbf{z}^T = [z_1 \ \dots \ z_N]$  kann in der Form

$$\psi(\mathbf{z}) = \mathbf{z} \quad (1.3)$$

angeschrieben werden. In dieser Formulierung sucht man nach einem *Fixpunkt*, d. h. einem Vektor  $\mathbf{z}_\infty$ , der

$$\psi(\mathbf{z}_\infty) = \mathbf{z}_\infty \quad (1.4)$$

erfüllt. Bei einer *Fixpunktiteration* definiert man eine Folge  $\mathbf{z}_0, \mathbf{z}_1, \dots$  durch die Vorschrift

$$\mathbf{z}_{j+1} = \psi(\mathbf{z}_j), \quad j = 0, 1, 2, \dots \quad (1.5)$$

Die Konvergenzeigenschaften der Folge hängen von der Wahl der Funktion  $\psi$  ab. Es ist wichtig zu verstehen, dass für ein gegebenes Problem  $\gamma(\mathbf{z}) = \mathbf{0}$  mehrere Funktionen  $\psi$  definiert werden können. Z. B. können die Funktionen

- $\psi(\mathbf{z}) = \mathbf{z} - \gamma(\mathbf{z})$
- $\psi(\mathbf{z}) = \mathbf{z} + 2\gamma(\mathbf{z})$
- $\psi(\mathbf{z}) = \mathbf{z} - \left(\frac{\partial}{\partial \mathbf{z}} \gamma\right)^{-1}(\mathbf{z})\gamma(\mathbf{z})$

gewählt werden. Die Wahl der Funktion  $\psi$  ist entscheidend für die Konvergenzeigenschaften der Fixpunktiteration. Um diese darzustellen, sind einige Definitionen notwendig.

**Definition 1.1 (Konvergenz).** Die Iteration (1.5) ist

- *lokal konvergent (LK)* gegen  $\mathbf{z}_\infty$ , falls es ein  $\delta > 0$  gibt, so dass wenn  $\|\mathbf{z}_0 - \mathbf{z}_\infty\| < \delta$  gilt, die Iteration (1.5) existiert und gegen  $\mathbf{z}_\infty$  konvergiert,

- *global konvergent (GK)*, falls die Iteration (1.5) für alle  $\mathbf{z}_0$  gegen  $\mathbf{z}_\infty$  konvergiert.

**Definition 1.2 (Stabilität).** Der Fixpunkt (1.4) ist

- *stabil* (im Sinne von Lyapunov), falls es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  gibt, so dass wenn  $\|\mathbf{z}_0 - \mathbf{z}_\infty\| < \delta$  gilt, die Folge  $\{\mathbf{z}_j\}$  von (1.5) existiert und  $\|\mathbf{z}_j - \mathbf{z}_\infty\| < \varepsilon$  für alle  $j \geq 1$  gilt,
- *attraktiv (A)*, falls es ein  $\delta > 0$  gibt, so dass wenn  $\|\mathbf{z}_0 - \mathbf{z}_\infty\| < \delta$  gilt, die Folge  $\{\mathbf{z}_j\}$  von (1.5) existiert und  $\lim_{j \rightarrow \infty} \mathbf{z}_j = \mathbf{z}_\infty$ , *global attraktiv (GA)* falls  $\delta = \infty$ ,
- *asymptotisch stabil (AS)*, falls er stabil und attraktiv ist und *global asymptotisch stabil (GAS)*, falls er stabil und global attraktiv ist.

Attraktivität und Konvergenz sind äquivalente Konzepte. Die folgenden Äquivalenzen gelten [1.6]

$$GAS \implies GA \iff GK \implies A \iff AS . \quad (1.6)$$

Für eine lineare Iteration

$$\mathbf{z}_{j+1} = \Psi \mathbf{z}_j , \quad j = 0, 1, 2, \dots \quad (1.7)$$

können zusätzlich folgende Aussagen getroffen werden.

**Definition 1.3 (Stabilität und asymptotische Stabilität einer linearen Iteration).** Eine lineare Iteration (1.7) heißt stabil, falls

$$\sup_{j \geq 1} \|\Psi^j\| < \infty , \quad (1.8)$$

sie heißt asymptotisch stabil, falls

$$\lim_{j \rightarrow \infty} \|\Psi^j\| = 0 . \quad (1.9)$$

**Definition 1.4 (Spektralradius).** Als Spektrum einer Matrix  $\mathbf{\Gamma}$  bezeichnet man die Menge aller Eigenwerte von  $\mathbf{\Gamma}$ , d. h.

$$\sigma(\mathbf{\Gamma}) = \{\lambda \in \mathbb{C} \mid \det(\lambda \mathbf{E} - \mathbf{\Gamma}) = 0\} \quad (1.10)$$

und

$$\rho(\mathbf{\Gamma}) = \max_{\lambda \in \sigma(\mathbf{\Gamma})} |\lambda| \quad (1.11)$$

ist der Spektralradius von  $\mathbf{\Gamma}$ .

**Satz 1.1.** Eine lineare Iteration (1.7) ist genau dann stabil, wenn für den Spektralradius  $\rho(\Psi) \leq 1$  gilt und alle Eigenwerte bei 1 nur einfach vorkommen. Sie ist genau dann asymptotisch stabil, wenn für den Spektralradius  $\rho(\Psi) < 1$  gilt.

Für den Beweis sei auf die angeführte Literatur verwiesen.

**Definition 1.5 (BIBO-Stabilität).** Eine lineare Iteration

$$\mathbf{z}_{j+1} = \Psi \mathbf{z}_j + \Lambda \mathbf{v}_j, \quad \mathbf{z}_0 = \mathbf{0} \quad (1.12)$$

heißt BIBO-stabil, wenn zu jeder beschränkten Eingangsfolge  $\{\mathbf{v}_j\}$  eine beschränkte Ausgangsfolge  $\{\mathbf{z}_j\}$  gehört.

**Satz 1.2.** Eine lineare Iteration  $\mathbf{z}_{j+1} = \Psi \mathbf{z}_j + \Lambda \mathbf{v}_j, \mathbf{z}_0 = \mathbf{0}$  ist genau dann BIBO-stabil, wenn für den Spektralradius  $\rho(\Psi) < 1$  gilt.

**Lemma 1.1.** Ist  $\{\mathbf{z}_j\}$  eine beschränkte Folge reeller Vektoren mit  $\rho > 0, \varepsilon > 0 \in \mathbb{R}$  für die

$$\|\mathbf{z}_{j+1}\| \leq \rho \|\mathbf{z}_j\| + \varepsilon, \quad 0 \leq \rho < 1 \quad (1.13)$$

gilt, dann folgt

$$\limsup_{j \rightarrow \infty} \|\mathbf{z}_j\| \leq \frac{1}{1 - \rho} \varepsilon. \quad (1.14)$$

*Beweis.* Die Iteration (1.13) lässt auf

$$\begin{aligned} \|\mathbf{z}_1\| &\leq \rho \|\mathbf{z}_0\| + \varepsilon \\ \|\mathbf{z}_2\| &\leq \rho^2 \|\mathbf{z}_0\| + (1 + \rho)\varepsilon \\ &\vdots \\ \|\mathbf{z}_j\| &\leq \rho^j \|\mathbf{z}_0\| + \sum_{k=0}^{j-1} \rho^k \varepsilon = \rho^j \|\mathbf{z}_0\| + \frac{1 - \rho^j}{1 - \rho} \varepsilon \end{aligned} \quad (1.15)$$

schließen. Für  $j \rightarrow \infty$  folgt  $\rho^j \rightarrow 0$  für  $0 \leq \rho < 1$  und damit (1.14).  $\square$

Außerdem können die Monotonieeigenschaften der linearen Iteration (1.7) charakterisiert werden.



**Definition 1.6** (Maximaler Singulärwert). Als maximalen Singulärwert einer Matrix  $\Psi$  bezeichnet man

$$\bar{\sigma}(\Psi) = \sqrt{\rho(\Psi^T \Psi)} . \quad (1.16)$$

Der maximale Singulärwert  $\bar{\sigma}(\Psi)$  gibt eine obere Schranke für die Verstärkung einer Abbildung  $\Psi$  an, denn es gilt

$$\|\Psi \mathbf{z}_j\|_2 \leq \|\Psi\|_2 \|\mathbf{z}_j\|_2 = \bar{\sigma}(\Psi) \|\mathbf{z}_j\|_2 . \quad (1.17)$$

**Satz 1.3** (Monotone Konvergenz einer linearen Iteration). Die lineare Iteration  $\mathbf{z}_{j+1} = \Psi \mathbf{z}_j$  ist monoton konvergent gegen  $\mathbf{0}$  in der  $l_2$ -Norm, d. h. es gilt

$$\|\mathbf{z}_{j+1}\|_2 \leq \beta \|\mathbf{z}_j\|_2 \quad \text{bzw.} \quad \|\mathbf{z}_{j+1}\|_2 \leq \beta^j \|\mathbf{z}_0\|_2 \quad (1.18)$$

für  $0 \leq \beta < 1$ , falls

$$\bar{\sigma}(\Psi) < 1 . \quad (1.19)$$

Für den Beweis wird auf die angeführte Literatur verwiesen. Da  $\rho(\Psi) \leq \bar{\sigma}(\Psi)$  gilt, ist jede monoton konvergente Iteration auch konvergent.

## 1.2 Lifted-System-Darstellung

Betrachtet wird ein lineares, zeitinvariantes SISO-System der Form

$$\dot{\mathbf{x}}_j(t) = \mathbf{A} \mathbf{x}_j(t) + \mathbf{b} u_j(t) , \quad \mathbf{x}_j(0) = \mathbf{x}_0 \quad (1.20a)$$

$$y_j(t) = \mathbf{c}^T \mathbf{x}_j(t) + d u_j(t) + v_j(t) . \quad (1.20b)$$

Es bezeichnet dabei  $t$  die Zeit und  $j = 0, 1, \dots$  den Iterationsindex. Ferner ist mit  $\mathbf{x}_j^T(t) = [x_1(t) \ \dots \ x_n(t)] \in \mathbb{R}^n$  der Zustandsvektor der Dimension  $n$ , mit  $\mathbf{x}_0$  die Anfangsbedingung, mit  $u(t)$  der Eingang, mit  $y(t)$  der Ausgang und mit  $v(t)$  das Messrauschen gegeben. Außerdem ist  $\mathbf{A}$  die Systemmatrix,  $\mathbf{b}$  der Eingangsvektor,  $\mathbf{c}^T$  der Ausgangsvektor und  $d$  der Durchgriffsparameter.

Verwendet man ein Abtast/Halteglied Nullter Ordnung, um (1.20) zu diskretisieren (ZOH-Diskretisierung (engl. Zero-Order-Hold-Discretization)), so ergibt sich mit dem Zeitindex  $k = 0, 1, \dots$  und dem Zustandsvektor  $\mathbf{x}_j[k] = \mathbf{x}_j(kT_a) = \mathbf{x}_j(t)$ , dem Eingang  $u_j[k] = u_j(kT_a) = u_j(t)$  und dem Ausgang  $y_j[k] = y_j(kT_a) = y_j(t)$  für  $kT_a \leq t \leq (k+1)T_a$

$$\mathbf{x}_j[k+1] = \Phi \mathbf{x}_j[k] + \gamma u_j[k] , \quad \mathbf{x}_j[0] = \mathbf{x}_0 \quad (1.21a)$$

$$y_j[k] = \mathbf{c}^T \mathbf{x}_j[k] + d u_j[k] + v_j[k] \quad (1.21b)$$

mit

$$\Phi = \exp(\mathbf{A}T_a) \quad \text{und} \quad \gamma = \int_0^{T_a} \exp(\mathbf{A}\tau) d\tau \mathbf{b} = (\exp(\mathbf{A}T_a) - \mathbf{E}) \mathbf{A}^{-1} \mathbf{b} . \quad (1.22)$$

Die äquivalente Eingangs-Ausgangs-Darstellung lautet

$$y_j[0] = \mathbf{c}^T \mathbf{x}_0 + du_j[0] + v_j[0] , \quad (1.23a)$$

$$y_j[k] = \mathbf{c}^T \Phi^k \mathbf{x}_0 + \mathbf{c}^T \sum_{m=0}^{k-1} (\Phi^{k-m-1} \gamma u_j[m]) + du_j[k] + v_j[k] , \quad k = 1, 2, \dots . \quad (1.23b)$$

Definiert man die Markov-Parameter<sup>1</sup>

$$g[k] = \begin{cases} d & \text{für } k = 0 \\ \mathbf{c}^T \Phi^{k-1} \gamma & \text{für } k = 1, 2, \dots \end{cases} , \quad (1.24)$$

so findet man eine weitere Darstellung von (1.23)

$$y_j[0] = \mathbf{c}^T \mathbf{x}_0 + du_j[0] + v_j[0] , \quad (1.25a)$$

$$y_j[k] = \mathbf{c}^T \Phi^k \mathbf{x}_0 + \sum_{m=0}^k (g[m] u_j[k-m]) + v_j[k] , \quad k = 1, 2, \dots . \quad (1.25b)$$

Anzumerken ist, dass der zweite Term in (1.25b) die diskrete Faltung  $u_j[k] * g[k] = \sum_{m=\mathcal{D}} g[m] u_j[k-m]$  darstellt, wobei  $\mathcal{D}$  der diskrete Definitionsbereich ist. Definiert man den Schiebeoperator  $\delta : \delta(z_j[k]) = z_j[k+1]$ , dann kann für  $\mathbf{x}_0 = \mathbf{0}$  die Übertragungoperator

$$\begin{aligned} G(\delta) &= \frac{y_j[k]}{u_j[k]} = d + \mathbf{c}^T (\delta \mathbf{E} - \Phi)^{-1} \gamma = d + \delta^{-1} \mathbf{c}^T (\mathbf{E} - \delta^{-1} \Phi)^{-1} \gamma \\ &= d + \delta^{-1} \mathbf{c}^T (\mathbf{E} + \delta^{-1} \Phi^1 + \delta^{-2} \Phi^2 + \dots) \gamma = d + \delta^{-1} \mathbf{c}^T \sum_{k=0}^{\infty} (\delta^{-k} \Phi^k) \gamma \\ &= d + \sum_{k=1}^{\infty} \mathbf{c}^T \Phi^{k-1} \gamma \delta^{-k} = \sum_{k=0}^{\infty} g[k] \delta^{-k} \end{aligned} \quad (1.26)$$

eingeführt werden.

**Bemerkung 1.1.** In der Herleitung eines ILC-Gesetzes spielt der relative Grad des Systems eine wichtige Rolle. Dieser ist für ein kontinuierliches System wie folgt definiert:

**Definition 1.7.** Das kontinuierliche System (1.20) für  $d = 0$  und  $v_j(t) = 0$  besitzt den relativen Grad  $r$ , wenn

$$(A) \quad \mathbf{c}^T \mathbf{A}^k \mathbf{b} = 0 , \quad k = 0, 1, \dots, r-2$$

$$(B) \quad \mathbf{c}^T \mathbf{A}^{r-1} \mathbf{b} \neq 0 .$$

<sup>1</sup>Aus der Vorlesung Automatisierung ist bekannt, dass die Markov-Parameter direkt die Impulsantwortfolge darstellen.

Man überzeugt sich leicht, dass der relative Grad eines kontinuierlichen Systems der Anzahl an Zeitableitungen des Ausgangs  $y_j(t)$  entspricht, bis erstmalig der Eingang  $u_j(t)$  auftritt. Dies wird deutlich, wenn man fortlaufend die Zeitableitung des Ausgangs bildet:

$$y_j(t) = \mathbf{c}^T \mathbf{x}_j(t) \quad (1.27a)$$

$$\dot{y}_j(t) = \mathbf{c}^T \mathbf{A} \mathbf{x}_j(t) + \underbrace{\mathbf{c}^T \mathbf{b}}_{=0} u_j(t) \quad (1.27b)$$

$$\ddot{y}_j(t) = \mathbf{c}^T \mathbf{A}^2 \mathbf{x}_j(t) + \underbrace{\mathbf{c}^T \mathbf{A} \mathbf{b}}_{=0} u_j(t) \quad (1.27c)$$

⋮

$$y_j^{(r-1)}(t) = \mathbf{c}^T \mathbf{A}^{r-1} \mathbf{x}_j(t) + \underbrace{\mathbf{c}^T \mathbf{A}^{r-2} \mathbf{b}}_{=0} u_j(t) \quad (1.27d)$$

$$y_j^{(r)}(t) = \mathbf{c}^T \mathbf{A}^r \mathbf{x}_j(t) + \underbrace{\mathbf{c}^T \mathbf{A}^{r-1} \mathbf{b}}_{\neq 0} u_j(t) \quad (1.27e)$$

Ist  $r = n$ , dann besitzt das System vollen relativen Grad. Für ein diskretes System gilt folgende Definition für den relativen Grad:

*Definition 1.8.* Das diskrete System (1.21) für  $d = 0$  und  $v_j[k] = 0$  besitzt den relativen Grad  $r$ , wenn

$$(A) \quad \mathbf{c}^T \Phi^k \boldsymbol{\gamma} = 0, \quad k = 0, 1, \dots, r-2$$

$$(B) \quad \mathbf{c}^T \Phi^{r-1} \boldsymbol{\gamma} \neq 0.$$

Der relative Grad eines diskreten Systems entspricht dem Zeitindex  $k$  des Ausgangs  $y_j[k]$  für den erstmalig der Eingang  $u_j[0]$  auftritt. Dies wird deutlich, wenn man in Analogie zu (1.27) die Ausgangsfolge betrachtet:

$$y_j[0] = \mathbf{c}^T \mathbf{x}_0 \quad (1.28a)$$

$$y_j[1] = \mathbf{c}^T \Phi \mathbf{x}_0 + \underbrace{\mathbf{c}^T \boldsymbol{\gamma}}_{=0} u_j[0] \quad (1.28b)$$

$$y_j[2] = \mathbf{c}^T \Phi^2 \mathbf{x}_0 + \underbrace{\mathbf{c}^T \Phi \boldsymbol{\gamma}}_{=0} u_j[0] + \underbrace{\mathbf{c}^T \boldsymbol{\gamma}}_{=0} u_j[1] \quad (1.28c)$$

⋮

$$y_j[r-1] = \mathbf{c}^T \Phi^{r-1} \mathbf{x}_0 + \underbrace{\mathbf{c}^T \sum_{m=0}^{r-2} \Phi^{r-m-1} \boldsymbol{\gamma} u_j[m]}_{=0} \quad (1.28d)$$

$$y_j[r] = \mathbf{c}^T \Phi^r \mathbf{x}_0 + \underbrace{\mathbf{c}^T \sum_{m=0}^{r-1} \Phi^{r-m-1} \boldsymbol{\gamma} u_j[m]}_{\neq 0} \quad (1.28e)$$

Es stellt sich die Frage, ob das diskrete System (1.21) dann ebenso einen vollen relativen Grad besitzt. Diese Frage kann beantwortet werden, indem in (1.22) die Exponentialmatrix  $\exp(\mathbf{A}\tau)$  durch ihre Reihendarstellung ersetzt, mit  $\mathbf{c}^T$  multipliziert und die Integration explizit ausgeführt wird:

$$\mathbf{c}^T \boldsymbol{\gamma} = \mathbf{c}^T \int_0^{T_a} \sum_{m=0}^{\infty} \frac{(\mathbf{A}\tau)^m}{m!} \mathbf{b} d\tau \quad (1.29a)$$

$$= \mathbf{c}^T \int_0^{T_a} \mathbf{b} + \mathbf{A}\mathbf{b}\tau + \dots + \frac{1}{(n-1)!} \mathbf{A}^{n-1} \mathbf{b}\tau^{n-1} + \mathcal{O}(T_a^n) d\tau \quad (1.29b)$$

$$= \mathbf{c}^T \mathbf{b} T_a + \frac{1}{2!} \mathbf{c}^T \mathbf{A} \mathbf{b} T_a^2 + \dots + \frac{1}{n!} \mathbf{c}^T \mathbf{A}^{n-1} \mathbf{b} T_a^n + \mathcal{O}(T_a^{n+1}) \quad (1.29c)$$

Dies zeigt, dass unter Verwendung einer ZOH-Diskretisierung für (1.20) (mit  $d = 0$  und  $v_j(t) = 0$ ) das diskrete System (1.21) immer einen relativen Grad von  $r = 1$  aufweist, da  $\mathbf{c}^T \boldsymbol{\gamma} \neq 0$ , vgl. (1.28b).

Der relative Grad eines diskreten Systems entspricht dem Zeitindex  $r$  für den der Eingang  $u_j[0]$  erstmalig direkt auf den Ausgang  $y_j[r]$  wirkt. Der relative Grad kann daher auch als Verzögerung verstanden werden. In einer realen Anwendung kann zusätzlich zur Verzögerung durch den relativen Grad  $r$  eine Verzögerung durch die Digital-Analog- bzw. Analog-Digital-Wandlung auftreten. Benötigt die Wandlung  $w$  Abtastschritte, dann kann eine Gesamtverzögerung  $m = w + r$  eingeführt und die Wandlungszeit in den Modellgleichungen mitberücksichtigt werden. Es wird nun eine Eingangs- und Ausgangsfolge für  $N$  Zeitschritte

$$u_j[k], \quad k = 0, 1, \dots, N-1 \quad (1.30a)$$

$$v_j[k], \quad k = 0, 1, \dots, N-1 \quad (1.30b)$$

$$y_j[k], \quad k = m, m+1, \dots, N+m-1 \quad (1.30c)$$

$$y_0[k] = \mathbf{c}^T \boldsymbol{\Phi}^k \mathbf{x}_0, \quad k = m, m+1, \dots, N+m-1 \quad (1.30d)$$

als auch eine Sollausgangsfolge

$$y_d[k], \quad k = m, m+1, \dots, N+m-1 \quad (1.31a)$$

betrachtet. Der Ausgangsfehler lautet sinngemäß

$$e_j[k] = y_d[k] - y_j[k], \quad k = m, m+1, \dots, N+m-1. \quad (1.32)$$

Definiert man die Vektoren

$$\mathbf{u}_j^T = [u_j[0] \quad u_j[1] \quad \dots \quad u_j[N-1]] \in \mathbb{R}^N \quad (1.33a)$$

$$\mathbf{v}_j^T = [v_j[0] \quad v_j[1] \quad \dots \quad v_j[N-1]] \in \mathbb{R}^N \quad (1.33b)$$

$$\mathbf{y}_j^T = [y_j[m] \quad y_j[m+1] \quad \dots \quad y_j[m+N-1]] \in \mathbb{R}^N \quad (1.33c)$$

$$\mathbf{y}_d^T = [y_d[m] \quad y_d[m+1] \quad \dots \quad y_d[m+N-1]] \in \mathbb{R}^N \quad (1.33d)$$

$$\mathbf{y}_0^T = [y_0[m] \quad y_0[m+1] \quad \dots \quad y_0[m+N-1]] \in \mathbb{R}^N \quad (1.33e)$$

$$\mathbf{e}_j^T = \mathbf{y}_d^T - \mathbf{y}_j^T = [e_j[m] \quad e_j[m+1] \quad \dots \quad e_j[m+N-1]] \in \mathbb{R}^N, \quad (1.33f)$$

so ergibt sich aus (1.25) die sogenannte Lifted-System-Darstellung

$$\mathbf{y}_j = \mathbf{y}_0 + \mathbf{G}\mathbf{u}_j + \mathbf{v}_j \quad (1.34)$$

mit der Matrix

$$\mathbf{G} = \begin{bmatrix} g[m] & 0 & \dots & 0 \\ g[m+1] & g[m] & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ g[m+N-1] & g[m+N-2] & \dots & g[m] \end{bmatrix} \in \mathbb{R}^{N \times N}. \quad (1.35)$$

Zu beachten ist, dass zum einen stets  $g[m] \neq 0$  gilt und zum anderen die Matrix  $\mathbf{G}$  Toeplitz ist. Eine Matrix  $\mathbf{G}$  wird Toeplitz-Matrix genannt wenn die Einträge  $\mathbf{G}_{ij}$  nur von der Differenz  $i - j$  der Indizes abhängen. Es sei angemerkt, dass lineare, zeitvariante Systeme ebenso in einer Lifted-System-Darstellung darstellbar sind, siehe z. B. [1.7].

### 1.3 Grundidee der ILR

Die Grundidee der ILR besteht darin, in jeder Iteration  $j = 0, 1, 2, \dots$  eine Steuerung  $\mathbf{u}_{j+1}$  mithilfe einer Iteration

$$\mathbf{u}_{j+1} = \psi(\mathbf{u}_j, \mathbf{e}_j(\mathbf{u}_j)), \quad j = 0, 1, 2, \dots \quad (1.36)$$

zu berechnen. Dazu muss in jeder Iteration  $j = 0, 1, 2, \dots$  der Ausgang  $\mathbf{y}_j$  und die Steuerung  $\mathbf{u}_j$  zwischengespeichert, der Ausgangsfehler  $\mathbf{e}_j$  gebildet und eine Steuerung  $\mathbf{u}_{j+1}$  berechnet werden. Abbildung 1.1 veranschaulicht die Grundidee der ILR. Um die Gestalt von (1.36) zu motivieren und näher zu spezifizieren, wird die Minimierungsaufgabe

$$\min_{\mathbf{u}} \frac{1}{2} \mathbf{e}^T \mathbf{P} \mathbf{e} + \frac{1}{2} \mathbf{u}^T \mathbf{W} \mathbf{u} + \mathbf{u}^T \mathbf{F} \mathbf{e} \quad (1.37a)$$

$$\text{u.B.v. } \mathbf{e} = \mathbf{y}_d - \mathbf{G} \mathbf{u}, \quad (1.37b)$$

mit den symmetrischen, positiv (semi-) definiten Gewichtungsmatrizen  $\mathbf{P}$  und  $\mathbf{W}$  betrachtet. Ohne Beschränkung der Allgemeinheit wird angenommen, dass  $\mathbf{F}\mathbf{G}$  schiefsymmetrisch

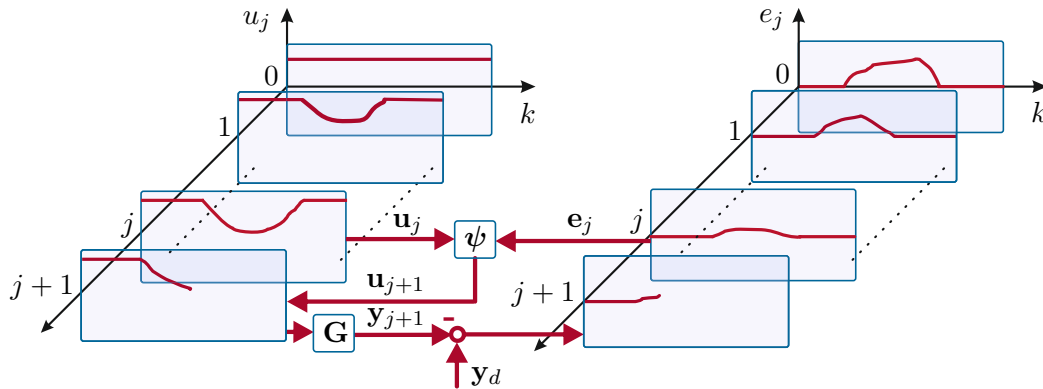


Abbildung 1.1: Grundidee der ILR.

ist. Dies ist zulässig, da der symmetrische Anteil von  $\mathbf{FG}$  immer in die Gewichtungsmatrix  $\mathbf{W}$  verschoben werden kann. Deutlich wird das, wenn (1.37b) in der Zielfunktion (1.37a) substituiert wird, wodurch sich die äquivalente Minimierungsaufgabe

$$\min_{\mathbf{u}} J(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T \bar{\mathbf{A}} \mathbf{u} + \mathbf{u}^T \bar{\mathbf{b}} + \bar{c} \quad (1.38)$$

mit

$$\bar{\mathbf{A}} = \mathbf{G}^T \mathbf{P} \mathbf{G} + \mathbf{W} \quad (1.39a)$$

$$\bar{\mathbf{b}} = (\mathbf{F} - \mathbf{G}^T \mathbf{P}) \mathbf{y}_d \quad (1.39b)$$

$$\bar{c} = \frac{1}{2} \mathbf{y}_d^T \mathbf{P} \mathbf{y}_d \quad (1.39c)$$

ergibt. Die quadratische Form  $\mathbf{u}^T \mathbf{F} \mathbf{G} \mathbf{u}$  verschwindet, da angenommen wurde, dass  $\mathbf{F} \mathbf{G}$  schiefsymmetrisch ist. Zudem wurde angenommen, dass  $\mathbf{P}$  und  $\mathbf{W}$  symmetrische, positiv (semi-) definite Matrizen sind. Daher ist auch  $\bar{\mathbf{A}}$  symmetrisch, positiv (semi-) definit.

Der Gradient der Zielfunktion  $J(\mathbf{u})$  in (1.38) lautet  $\nabla J(\mathbf{u}) = \bar{\mathbf{A}} \mathbf{u} + \bar{\mathbf{b}}$ . Falls  $\bar{\mathbf{A}}$  symmetrisch, positiv definit ist, ist  $J(\mathbf{u})$  in (1.38) strikt konvex und weist ein eindeutiges globales Minimum bei  $\mathbf{u}_\infty = -\bar{\mathbf{A}}^{-1} \bar{\mathbf{b}}$  auf, welches die notwendige und hinreichende Bedingung für Optimalität  $\nabla J(\mathbf{u}_\infty) = \mathbf{0}$  erfüllt. Die Minimierungsaufgabe (1.38) kann numerisch mithilfe des Gradientenverfahrens mit konstanter Schrittweite  $\alpha$  der Form

$$\mathbf{u}_{j+1} = \mathbf{u}_j - \alpha \nabla J(\mathbf{u}_j) = (\mathbf{E} - \alpha \bar{\mathbf{A}}) \mathbf{u}_j - \alpha \bar{\mathbf{b}}, \quad (1.40)$$

gelöst werden, wobei  $0 < \alpha < 2/\|\bar{\mathbf{A}}\|_2$  (monotone) Konvergenz garantiert, vgl. Theorem 2.1.13 in [1.8]. Ersetzen von  $\mathbf{y}_d$  in (1.39b) anhand der Nebenbedingung (1.37b) im Gradientenverfahren (1.40) liefert das ILR-Gesetz

$$\mathbf{u}_{j+1} = \mathbf{Q}(\mathbf{u}_j + \mathbf{L}e_j) \quad (1.41)$$

mit den Matrizen

$$\mathbf{Q} = \mathbf{E} - \alpha(\mathbf{W} + \mathbf{F}\mathbf{G}) \quad (1.42a)$$

$$\mathbf{Q}\mathbf{L} = \alpha(\mathbf{G}^T\mathbf{P} - \mathbf{F}) . \quad (1.42b)$$

Im Allgemeinen ist die  $Q$ -Filtermatrix  $\mathbf{Q}$  gegeben durch

$$\mathbf{Q} = \begin{bmatrix} q[0] & q[-1] & \cdots & q[-(N-1)] \\ q[1] & q[0] & \cdots & q[-(N-2)] \\ \vdots & \vdots & \ddots & \vdots \\ q[N-1] & q[N-2] & \cdots & q[0] \end{bmatrix} \in \mathbb{R}^{N \times N} \quad (1.43)$$

und die Verstärkungsmatrix  $\mathbf{L}$  durch

$$\mathbf{L} = \begin{bmatrix} l[0] & l[-1] & \cdots & l[-(N-1)] \\ l[1] & l[0] & \cdots & l[-(N-2)] \\ \vdots & \vdots & \ddots & \vdots \\ l[N-1] & l[N-2] & \cdots & l[0] \end{bmatrix} \in \mathbb{R}^{N \times N} . \quad (1.44)$$

Das ILR-Gesetz (1.41) kann auch in der Form

$$u_{j+1}[k] = Q(\delta)(u_j[k] + L(\delta)e_j[k+m]) \quad (1.45)$$

bzw.

$$u_{j+1}[k] = q[k] * (u_j[k] + l[k] * e_j[k+m]) \quad (1.46)$$

angegeben werden. Damit entsprechen die  $Q$ -Filterung und die Multiplikation mit der Verstärkungsmatrix nicht-kausalen Filterungen mit den Übertragungsoperatoren

$$Q(\delta) = \cdots + q[-2]\delta^2 + q[-1]\delta^1 + q[0] + q[1]\delta^{-1} + q[2]\delta^{-2} + \cdots \quad (1.47a)$$

$$L(\delta) = \cdots + l[-2]\delta^2 + l[-1]\delta^1 + l[0] + l[1]\delta^{-1} + l[2]\delta^{-2} + \cdots \quad (1.47b)$$

und den Impulsantwortfolgen

$$q[k] \quad \text{für } k = -(N-1), -(N-2), \dots, (N-2), (N-1) \quad (1.48a)$$

$$l[k] \quad \text{für } k = -(N-1), -(N-2), \dots, (N-2), (N-1) . \quad (1.48b)$$

Das ILR-Gesetz (1.45) kann unterschiedlich interpretiert werden. Zum einen stellt für  $Q(\delta) = L(\delta) = 1$  das ILR-Gesetz (1.41) einen Integrator im Iterationsbereich dar. Aus der Vorlesung Automatisierung ist bekannt, dass ein Integrator einen konstanten Ausgangsfehler  $e_j$  zu Null regeln kann. Zum anderen kann (1.41) als eine Fixpunktiteration für die Nullstellensuche der Funktion  $e_j(u_j)$  verstanden werden. Eine weitere Interpretation der ILR ergibt sich, wenn man die ILR als das Online-Lösen der Optimierungsaufgabe (1.111) mit einem Einzelschießverfahren interpretiert. Das Aufschalten der Steuerung entspricht einem Einzelschießen (Lösen der Anfangswertaufgabe) und das ILR-Gesetz (1.41) einer überlagerten Nullstellensuche. Im Unterschied zu einem Offline-Verfahren muss ein Online-Verfahren robust gegenüber Messrauschen und Variation der Anfangsbedingung sein. Wie später gezeigt wird, unterdrückt die  $Q$ -Filterung den Einfluss des Messrauschens und den von sich nicht wiederholenden Störungen.

## 1.4 Stabilitäts- und Konvergenzanalyse

Es soll nun die Stabilität des ILR-Gesetzes (1.41)

$$\mathbf{u}_{j+1} = \mathbf{Q}(\mathbf{u}_j + \mathbf{L}\mathbf{e}_j) , \quad (1.49)$$

angewendet auf die Lifted-System-Darstellung (1.34)

$$\mathbf{y}_j = \mathbf{y}_0 + \mathbf{G}\mathbf{u}_j + \mathbf{v}_j , \quad (1.50)$$

gezeigt werden. Dazu sind folgende Annahmen zu treffen:

- (i.) Jede Iteration endet innerhalb einer festen Iterationsdauer  $t_f = NT_a$ .
- (ii.) Der Prozess startet immer von der gleichen Anfangsbedingung, d. h.  $\mathbf{y}_0 = \text{konst.}$
- (iii.) Für einen gegebenen Sollausgang  $\mathbf{y}_d$  existiert eine eindeutige Steuerung  $\mathbf{u}_d$ , d. h. es gilt  $\mathbf{y}_d = \mathbf{y}_0 + \mathbf{G}\mathbf{u}_d$ .
- (iv.) Die Messung  $\mathbf{y}_j$  ist ideal, das heißt, dass  $\mathbf{v}_j = \mathbf{0}$  gilt.

Mit dem Ausgangsfehler  $\mathbf{e}_j = \mathbf{y}_d - \mathbf{y}_j$  nach (1.33f), der Lifted-System-Darstellung (1.34) und  $\tilde{\mathbf{y}}_d = \mathbf{y}_d - \mathbf{y}_0$ , findet man mit

$$\mathbf{e}_j = \tilde{\mathbf{y}}_d - \mathbf{G}\mathbf{u}_j \quad (1.51)$$

anhand von (1.41) und die lineare Iteration

$$\mathbf{u}_{j+1} = \Psi\mathbf{u}_j + \Lambda\tilde{\mathbf{y}}_d , \quad (1.52)$$

wobei  $\Psi = \mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})$  und  $\Lambda = \mathbf{Q}\mathbf{L}$  gilt. Der Fixpunkt  $\mathbf{u}_{j+1} = \mathbf{u}_j = \mathbf{u}_\infty$  der Iteration errechnet sich zu

$$\mathbf{u}_\infty = (\mathbf{E} - \Psi)^{-1} \Lambda\tilde{\mathbf{y}}_d . \quad (1.53)$$

Führt man den Eingangsfehler  $\bar{\mathbf{u}}_j = \mathbf{u}_j - \mathbf{u}_\infty$  ein, so ergibt sich die Eingangsfehleriteration

$$\bar{\mathbf{u}}_{j+1} = \Psi\bar{\mathbf{u}}_j , \quad (1.54)$$

für welche Satz 1.1 folgende Stabilitätsaussage liefert:

**Satz 1.4 (Asymptotische Stabilität des ILR-Gesetzes).** Die Eingangsfehleriteration (1.54) des ILR-Gesetzes (1.41), angewendet auf die Lifted-System-Darstellung (1.34), ist asymptotisch stabil, falls

$$\rho(\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})) < 1 . \quad (1.55)$$



Anhand von  $\mathbf{u}_j = \mathbf{G}^{-1}(\tilde{\mathbf{y}}_d - \mathbf{e}_j)$  gemäß (1.51) folgt aus der Eingangssiteration (1.52) die Ausgangssiteration

$$\mathbf{e}_{j+1} = \mathbf{G}\Psi\mathbf{G}^{-1}\mathbf{e}_j + (\mathbf{E} - \mathbf{G}\mathbf{Q}\mathbf{G}^{-1})\tilde{\mathbf{y}}_d. \quad (1.56)$$

Für dessen Fixpunkt  $\mathbf{e}_{j+1} = \mathbf{e}_j = \mathbf{e}_\infty$  gilt

$$\begin{aligned} \mathbf{e}_\infty &= \tilde{\mathbf{y}}_d - \mathbf{G}\mathbf{u}_\infty \\ &= (\mathbf{E} - \mathbf{G}(\mathbf{E} - \Psi)^{-1}\mathbf{A})\tilde{\mathbf{y}}_d. \end{aligned} \quad (1.57)$$

Man erkennt anhand von (1.56), dass nur für  $\mathbf{Q} = \mathbf{E}$  der Ausgangsfehler verschwindet, d. h.

$$\mathbf{e}_\infty = \mathbf{0}. \quad (1.58)$$

Mit  $\bar{\mathbf{e}}_j = \mathbf{e}_j - \mathbf{e}_\infty$  ergibt sich

$$\bar{\mathbf{e}}_{j+1} = \mathbf{G}\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\mathbf{G}^{-1}\bar{\mathbf{e}}_j. \quad (1.59)$$

Mit dem Satz 1.1 folgt somit für die Ausgangsfehlerdynamik:

**Satz 1.5 (Stabilität der Ausgangsfehlerdynamik).** Die Ausgangsfehlerdynamik (1.59) der Lifted-System-Darstellung (1.34), (1.41) ist asymptotisch stabil, falls

$$\rho(\mathbf{G}\Psi\mathbf{G}^{-1}) < 1. \quad (1.60)$$

Ein Vergleich der Eingangssiteration (1.52) und der Ausgangssiteration (1.56) zeigt, dass die Dynamikmatrizen die selben Eigenwerte aufweisen, d. h.

$$\rho(\mathbf{G}\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\mathbf{G}^{-1}) = \rho(\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})) \quad (1.61)$$

gilt. Damit folgt aus der asymptotischen Stabilität der Eingangssiteration automatisch die asymptotische Stabilität der Ausgangssiteration und umgekehrt. In einer Norm lautet (1.56)

$$\|\mathbf{e}_{j+1}\| \leq \underbrace{\|\mathbf{G}\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\mathbf{G}^{-1}\|}_{=\sigma} \|\mathbf{e}_j\| + \underbrace{\|\mathbf{E} - \mathbf{G}\mathbf{Q}\mathbf{G}^{-1}\|}_{=\gamma} \underbrace{\|\tilde{\mathbf{y}}_d\|}_{=\varepsilon}, \quad 0 \leq \sigma < 1. \quad (1.62)$$

Lemma 1.1 liefert dann die asymptotische Aussage

$$\limsup_{j \rightarrow \infty} \|\mathbf{e}_j\| \leq \frac{1}{1 - \sigma} \gamma \varepsilon. \quad (1.63)$$

In der Literatur findet man häufig ILR-Gesetze, die kein  $Q$ -Filter verwenden, d. h.  $\mathbf{Q} = \mathbf{E}$ . Die vorangegangenen Überlegungen zeigen, dass dies notwendig ist, um perfektes Folgeverhalten ( $\mathbf{e}_\infty = \mathbf{0}$ ) zu gewährleisten. Die Verwendung eines  $Q$ -Filters verbessert allerdings die Robustheit gegenüber Variationen der Anfangsbedingungen  $\mathbf{y}_0$  und gegenüber Messrauschen  $\mathbf{v}_j$ . Um dies zu plausibilisieren, nimmt man an, dass die Anfangsbedingung

nicht in jeder Iteration gleich ist, d. h.  $\mathbf{y}_0 = \mathbf{y}_{0,j}$ , und dass die Messung nicht ideal ist, d. h. ein Rauschprozess  $\mathbf{v}_j$  wirkt. Dann lautet (1.52)

$$\mathbf{u}_{j+1} = \mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\mathbf{u}_j + \mathbf{Q}\mathbf{L}(\mathbf{y}_d - \mathbf{y}_{0,j} - \mathbf{v}_j) . \quad (1.64)$$

Ohne die Verwendung eines  $Q$ -Filters ( $\mathbf{Q} = \mathbf{E}$ ) würden das Messrauschen und Variationen der Anfangsbedingung direkt verstärkt und in die Steuerung für die nächste Iteration eingehen. Es ist unmittelbar einsichtig, dass z. B. ein Tiefpassfilter hochfrequentes Rauschen unterdrücken kann. Da das  $Q$ -Filter auf der ganzen rechten Seite von (1.64) wirkt, schaltet man damit das Lernen für hochfrequente Signalanteile aus.

Um große Transienten während der Iterationen zu vermeiden, muss monotone Konvergenz sichergestellt werden. Satz 1.3 liefert nunmehr die Aussagen:

**Satz 1.6 (Monotone Konvergenz der Eingangsfehleriteration).** *Die Eingangsfehleriteration (1.54) der Lifted-System-Darstellung (1.34), (1.41) ist monoton konvergent gegen  $\mathbf{u}_\infty$  in der  $l_2$ -Norm, d. h. es gilt*

$$\|\mathbf{u}_{j+1} - \mathbf{u}_\infty\|_2 \leq \alpha \|\mathbf{u}_j - \mathbf{u}_\infty\|_2 \quad (1.65)$$

für  $0 \leq \alpha < 1$ , falls

$$\bar{\sigma}(\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})) = \alpha < 1 . \quad (1.66)$$

**Satz 1.7 (Monotone Konvergenz der Ausgangsfehleriteration).** *Die Ausgangsfehleriteration (1.59) der Lifted-System-Darstellung (1.34), (1.41) ist monoton konvergent gegen  $\mathbf{e}_\infty$  in der  $l_2$ -Norm, d. h. es gilt*

$$\|\mathbf{e}_{j+1} - \mathbf{e}_\infty\|_2 \leq \beta \|\mathbf{e}_j - \mathbf{e}_\infty\|_2 \quad (1.67)$$

für  $0 \leq \beta < 1$ , falls

$$\bar{\sigma}(\mathbf{G}\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\mathbf{G}^{-1}) = \beta < 1 . \quad (1.68)$$

Das ILR-Gesetz (1.41) mit (1.42) ist monoton konvergent, falls  $\|\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})\|_2 = \|\mathbf{E} - \alpha\bar{\mathbf{A}}\|_2 < 1$  gilt. Dies ist in Einklang mit dem Konvergenzkriterium  $0 < \alpha < 2/\|\bar{\mathbf{A}}\|_2$  des Gradientenverfahrens mit fester Schrittweite, da anhand der umgekehrten Dreiecksungleichung  $|1 - \|\alpha\bar{\mathbf{A}}\|_2| \leq \|\mathbf{I} - \alpha\bar{\mathbf{A}}\|_2 < 1$  gilt.

## 1.5 ILR Entwurfsverfahren

Bisher wurde lediglich die Stabilität und Konvergenz des ILR-Gesetzes (1.41) untersucht. Offen ist die Frage, wie das  $Q$ -Filter  $\mathbf{Q}$  und die Verstärkungsmatrix  $\mathbf{L}$  entworfen werden. Aus der Vielzahl von Entwurfsverfahren sollen drei vorgestellt werden:

- (i.) P-Type und PD-Type ILR
- (ii.) Inversionsbasierte ILR
- (iii.) Normoptimale ILR

### 1.5.1 P-Type und PD-Type ILR

Bei der P-Type und PD-Type ILR [1.9, 1.10] bedient man sich bekannter Konzepte der PID-Ausgangsregelung. Die Rückführung beinhaltet also je nach Typ einen Proportional- und/ oder einen Differentialanteil des Ausgangsfehlers  $e_j$ . Anzumerken ist, dass meist kein Integralterm verwendet wird, da das ILR-Gesetz (1.41) ohnehin schon integrierendes Verhalten aufweist. Das Verfahren ist ein Trial-and-Error Verfahren, das in zwei Schritten erfolgt. Im ersten Schritt wird eine Verstärkungsmatrix  $\mathbf{L}$  entworfen und die Einstellparameter so eingestellt, dass die Stabilität nach Satz 1.5 gewährleistet wird und sich eine hinreichend gute Regelgüte einstellt. Im zweiten Schritt wird das ILR-Gesetz um ein  $Q$ -Filter erweitert um Robustheit gegenüber Messrauschen zu sichern. Um das Konzept der P-Type und PD-Type ILR zu erklären, wird zunächst der Entwurf der Verstärkungsmatrix  $\mathbf{L}$  aufgezeigt ( $L$ -Filterentwurf).

#### $L$ -Filterentwurf

Die einfachste Möglichkeit die Verstärkungsmatrix  $\mathbf{L}$  zu entwerfen, besteht in der Wahl einer diagonalen Verstärkungsmatrix mit einem Freiheitsgrad, d. h. einem P-Type ILR-Gesetz in der Form

$$\mathbf{u}_{j+1} = \mathbf{Q}(\mathbf{u}_j + k_p \mathbf{E} e_j) . \quad (1.69)$$

Dabei bezeichnet  $k_p > 0$  einen Einstellparameter. Bei der PD-Type ILR wird der Ausgangsfehler und zusätzlich eine numerische Approximation der Zeitableitung des Ausgangsfehlers verwendet. Ein PD-Type ILR-Gesetz, welches den Zentralen Differenzenquotienten nutzt, lautet für  $k = 1, 2, \dots, N - 2$

$$u_{j+1}[k] = q[k] * (u_j[k] + k_p e_j[k + r] + \frac{k_d}{2T_a} (e_j[k + r + 1] - e_j[k + r - 1])) . \quad (1.70)$$

$k_p > 0$  und  $k_d > 0$  sind wiederum Einstellparameter. Für die Randiteration bei  $k = 0$  und  $k = N - 1$  kann z. B. ein Vorwärts- bzw. ein Rückwärts-Differenzenquotient angesetzt

werden. Führt man die Ableitungsmatrix

$$\mathbf{D} = \frac{1}{T_a} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1/2 & 0 & 1/2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & -1/2 & 0 & 1/2 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad (1.71)$$

ein, so folgt für den PD-Type die zu (1.70) äquivalente Lifted-System-Darstellung

$$\mathbf{u}_{j+1} = \mathbf{Q} \left( \mathbf{u}_j + \underbrace{(k_p \mathbf{E} + k_d \mathbf{D})}_{\mathbf{L}} \mathbf{e}_j \right). \quad (1.72)$$

Die Gleichung (1.72) verdeutlicht, dass die Stabilität des PD-Typs wieder anhand von Satz 1.5 charakterisiert werden kann.

### Q-Filterentwurf

Der Einsatz eines Q-Filters verhindert das Lernen von sich nicht wiederholenden Störungen und Messrauschen. Da eine akausale Filterung im Rahmen der ILR möglich ist, bietet es sich an, einen gleitenden Mittelwert zu verwenden. Es werden im Folgenden zwei Entwurfsmethoden vorgestellt:

- (i.) Gauß-Filter
- (ii.) Savitzky-Golay-Filter

Der Entwurf eines Savitzky-Golay-Filters erlaubt es zusätzlich Ableitungsmatrix höherer Ordnung zu entwerfen. Diese Ableitungsmatrix, die im Grunde Differenzenquotienten darstellen, können zur Konstruktion einer Ableitungsmatrix  $\mathbf{D}$  verwendet werden, vgl. (1.71).

### Gauß-Filter

Ein einfacher Entwurf eines akausalen Glättungsfilters (Tiefpassfilters) ist mithilfe eines Gauß-Filters [1.11] möglich. Dazu betrachtet man die Gauß-Verteilung

$$q_G(t) = \frac{1}{\sigma_q \sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma_q^2}\right) \quad (1.73)$$

mit der Standardabweichung  $\sigma_q$ , die mit der 3-dB Bandbreite  $f_c$  über

$$\sigma_q = \frac{\sqrt{\ln(2)}}{2\pi f_c} \quad (1.74)$$

im Zusammenhang steht. Mit  $q_G[k] = q_G(kT_a) = q_G(t)$  für  $kT_a \leq t \leq (k+1)T_a$  findet man

$$q_G[k] = \frac{1}{\sigma_q \sqrt{2\pi}} \exp\left(-\frac{(kT_a)^2}{2\sigma_q^2}\right), \quad k \in \mathbb{Z} \quad (1.75)$$

und eine Fensterung mit einem Rechteckfenster der Länge  $M_q = 2N_q + 1$  und Normierung  $\sum_{m=-N_q}^{N_q} q[m] = 1$  ergibt

$$q[k] = \frac{q_G[k]}{\sum_{m=-N_q}^{N_q} q_G[m]}, \quad k \in \mathbb{Z}. \quad (1.76)$$

Für  $M_q = 2N_q + 1$  Datenpunkte  $y[k+m]$ ,  $m = -N_q, \dots, N_q$  erhält man das gefilterte Signal

$$\hat{y}[k] = q[k] * y[k] = \sum_{m=-N_q}^{N_q} q[m] y[k-m]. \quad (1.77)$$

**Bemerkung 1.2.** Durch die Verwendung eines Gauß-Filters erhält man eine  $Q$ -Filtermatrix, die eine Bandstruktur aufweist. Es stellt sich allerdings die Frage, wie die  $Q$ -Filtermatrix an den Rändern aufzubauen ist. Dazu ein Beispiel:

**Beispiel 1.1.** Einfaches Einsetzen der Filterkoeffizienten (1.76) in (1.43) liefert für  $N = 5$  und  $N_q = 1$  die  $Q$ -Filtermatrix

$$\mathbf{Q} = \begin{bmatrix} q[0] & q[-1] & 0 & 0 & 0 \\ q[1] & q[0] & q[-1] & 0 & 0 \\ 0 & q[1] & q[0] & q[-1] & 0 \\ 0 & 0 & q[1] & q[0] & q[-1] \\ 0 & 0 & 0 & q[1] & q[0] \end{bmatrix} \in \mathbb{R}^{5 \times 5}. \quad (1.78)$$

Man erkennt, dass die Filterverstärkung in der zweiten, dritten und vierten Zeile eins ist, d. h.  $q[1] + q[0] + q[-1] = 1$ . In den Randzeilen ist sie ungleich eins, d. h.  $q[0] + q[-1] \neq 1$  und  $q[1] + q[0] \neq 1$ . Um dies zu vermeiden, kann man die Randzeilen jeweils getrennt normieren. Das Filter wird dann allerdings für die Randpunkte nicht mehr die gewünschte 3-dB Bandbreite  $f_c$  einhalten.

**Beispiel 1.2 (Gauß-Filter).** Abbildung 1.2 zeigt die Impulsantwortfolge  $q[k]$  und den diskreten Frequenzgang  $G(\exp(j\omega T_a))$  eines Gauß-Filters für  $T_a = 1$  ms,  $f_c = 20, 100, 200$  Hz und  $N_q = 10$ . Die Abbildung zeigt deutlich, dass für hohe Frequenzen  $f_c$  die Impulsantwortfolge  $q[k]$  sehr spitz wird und damit im Grenzfall  $f_c = 1/(2T_a)$  keine Filterung mehr stattfindet. Für niedrige Frequenzen ( $f_c \lesssim 1/(40T_a)$ ) wird die Impulsantwortfolge  $q[k]$  immer breiter. In diesem Grenzfall gehen alle Datenpunkte gleich in die Mittlung ein. Außerdem verringert sich die Abschneidelänge, was die

Ausbildung einer Welligkeit im Sperrband zur Folge hat. Grundsätzlich muss damit beim Entwurf darauf geachtet werden, dass die Impulsantwortfolge nicht zu spitz und nicht zu flach gewählt wird.

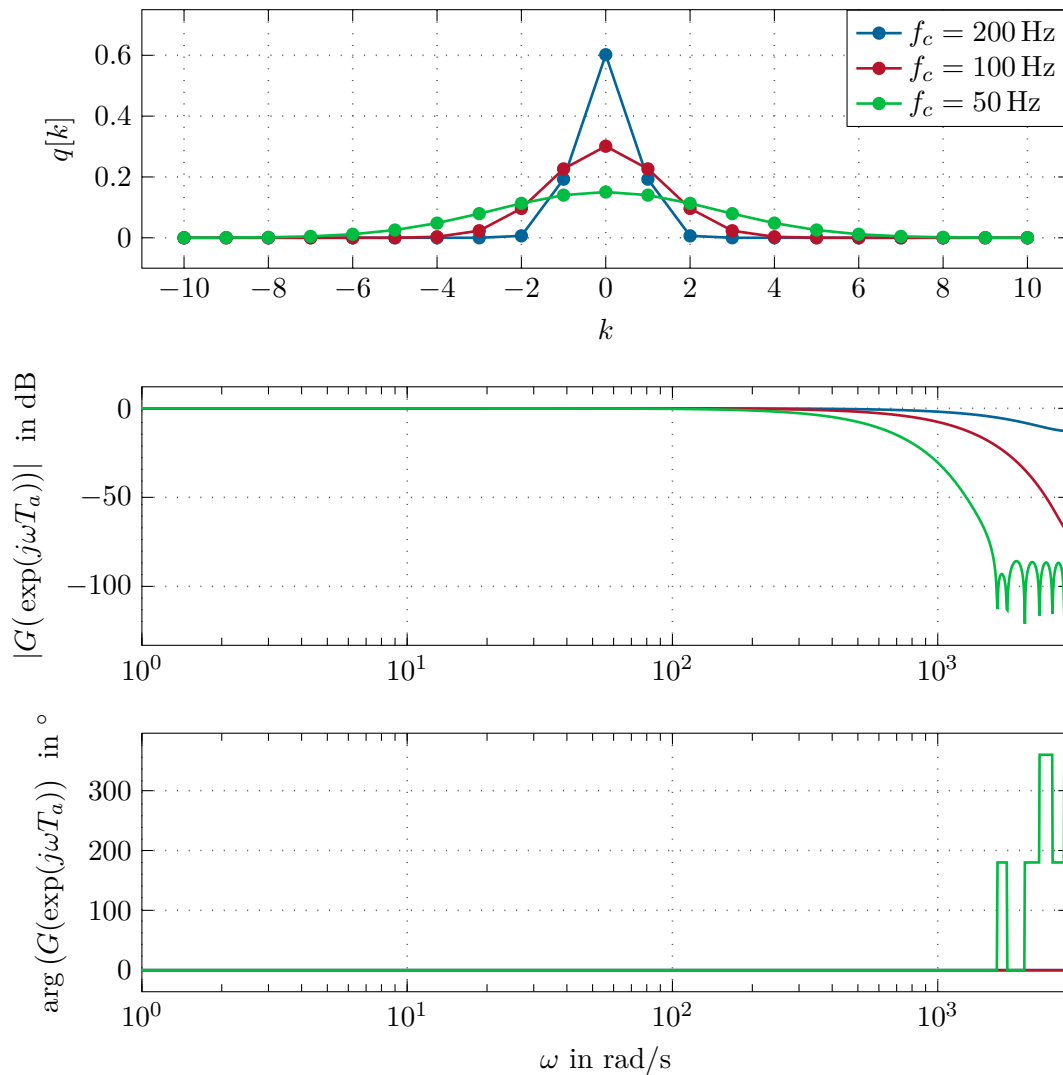


Abbildung 1.2: Impulsantwortfolge  $q[k]$  und diskreter Frequenzgang  $G(\exp(j\omega T_a))$  eines Gauß-Filters für  $T_a = 1$  ms und  $N_q = 10$ .

### Savitzky-Golay-Filter

Das Savitzky-Golay-Filter [1.12, 1.13] ist ebenso wie das Gauß-Filter ein Glättungsfilter. Im Wesentlichen wird dazu eine polynomiale Regression über eine Folge von Datenpunkten durchgeführt. Zusätzlich ist es im Rahmen des Entwurfs des Savitzky-Golay-Filters möglich ein Ableitungsfilter zu erstellen. So ein Ableitungsfilter kann zur numerischen Approximation der Zeitableitung des Ausgangsfehlers verwendet werden. Im Gegensatz

zum Entwurf eines Gauß-Filters ist die 3-dB Bandbreite eines Savitzky-Golay-Filters nicht direkt vorgebar und im Unterschied zum Gauß-Filter erhält das Savitzky-Golay-Filter hochfrequente Signalanteile des Nutzsignals auf Kosten einer schlechteren Rauschunterdrückung.

Es werden dazu  $M_q = 2N_q + 1$  Datenpunkte  $y[k + m]$ ,  $m = -N_q, \dots, N_q$  betrachtet für welche eine polynomiale Regression anhand eines Polynoms  $d$ -ten Grades der Form

$$p_m[k] = \sum_{i=0}^d c_i[k] (T_a m)^i, \quad (1.79)$$

mit den Koeffizienten  $c_i[k]$ ,  $i = 0, 1, \dots, d$ , ausgeführt wird. Die Koeffizienten  $c_i[k]$  werden dementsprechend anhand der Minimierungsaufgabe

$$\min_{c_0[k], c_1[k], \dots, c_d[k]} \sum_{m=-N_q}^{N_q} (y[k + m] - p_m[k])^2 \quad (1.80)$$

berechnet. Um die Minimierungsaufgabe zu lösen, fasst man die Datenpunkte  $y[k + m]$ ,  $m = -N_q, \dots, N_q$  in einem Vektor

$$\mathbf{y}^T[k] = [y[k - N_q] \quad \dots \quad y[k - 1] \quad y[k] \quad y[k + 1] \quad \dots \quad y[k + N_q]] \quad (1.81)$$

zusammen. Außerdem wird der Parametervektor

$$\mathbf{c}^T[k] = [c_0[k] \quad c_1[k] \quad \dots \quad c_d[k]] \quad (1.82)$$

und ein Vektor

$$\hat{\mathbf{y}}^T[k] = [p_{-N_q}[k] \quad \dots \quad p_{-1}[k] \quad p_0[k] \quad p_1[k] \quad \dots \quad p_{N_q}[k]] \quad (1.83)$$

definiert. Die Minimierungsaufgabe (1.80) lautet damit

$$\min_{\mathbf{c}[k]} (\mathbf{y}[k] - \hat{\mathbf{y}}[k])^T (\mathbf{y}[k] - \hat{\mathbf{y}}[k]). \quad (1.84)$$

Da (1.79) linear in den Koeffizienten  $c_i[k]$ ,  $i = 0, 1, \dots, d$  ist, kann die Beziehung zu

$$\hat{\mathbf{y}}[k] = \mathbf{S}\mathbf{c}[k] \quad (1.85)$$

angegeben werden. Dazu definiert man  $d + 1$  Vektoren  $\mathbf{s}_i$ ,  $i = 0, 1, \dots, d$ , welche die Einträge

$$\mathbf{s}_i(m + N_q + 1) = (T_a m)^i, \quad m = -N_q, \dots, N_q \quad (1.86)$$

aufweisen und die Matrix

$$\mathbf{S} = [\mathbf{s}_0 \quad \mathbf{s}_1 \quad \dots \quad \mathbf{s}_d]. \quad (1.87)$$

Bekannterweise lautet die Lösung der Minimierungsaufgabe

$$\mathbf{c}[k] = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{y}[k] \quad (1.88)$$

und damit folgt für (1.85)

$$\hat{\mathbf{y}}[k] = \mathbf{S} (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{y}[k]. \quad (1.89)$$

Definiert man die Matrizen

$$\mathbf{H}^T = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T = [\mathbf{h}_0 \quad \mathbf{h}_1 \quad \dots \quad \mathbf{h}_d]^T \quad (1.90a)$$

$$\mathbf{B} = \mathbf{S} \mathbf{H}^T = [\mathbf{b}_{-N_q} \quad \dots \quad \mathbf{b}_{-1} \quad \mathbf{b}_0 \quad \mathbf{b}_1 \quad \dots \quad \mathbf{b}_{N_q}] \quad (1.90b)$$

mit den Spaltenvektoren  $\mathbf{h}_i$ ,  $i = 0, 1, \dots, d$  und  $\mathbf{b}_m$ ,  $m = -N_q, \dots, N_q$ , dann gilt für die Koeffizienten

$$\mathbf{c}[k] = \mathbf{H}^T \mathbf{y}[k] \iff c_i[k] = \mathbf{h}_i^T \mathbf{y}[k], \quad i = 0, 1, \dots, d \quad (1.91)$$

und für das gefilterte Signal

$$\hat{\mathbf{y}}[k] = \mathbf{B} \mathbf{y}[k] \iff p_m[k] = \mathbf{b}_m^T \mathbf{y}[k], \quad m = -N_q, \dots, N_q. \quad (1.92)$$

D. h., das gefilterte Signal am Mittelpunkt  $\bar{m} = mT_a = 0$  ist durch

$$\hat{y}[k] = p_0[k] = \mathbf{b}_0^T \mathbf{y}[k] \quad (1.93)$$

bzw.

$$\hat{y}[k] = \sum_{m=-N_q}^{N_q} \mathbf{b}_0(m + N_q + 1) y[k + m] = \sum_{m=-N_q}^{N_q} \mathbf{b}_0(N_q - m + 1) y[k - m] \quad (1.94)$$

gegeben, wobei  $\mathbf{b}_0(m + N_q + 1)$  der  $m + N_q + 1$ -te Eintrag von  $\mathbf{b}_0$  ist.

### Savitzky-Golay-Ableitungsfilter

Anhand der  $d + 1$  Spalten der  $M_q \times (d + 1)$ -dimensionalen Matrix  $\mathbf{H}$  kann ein Savitzky-Golay-Ableitungsfilter für die  $i = 1, 2, \dots, d$ -te Ableitung des Signals  $\hat{y}[k]$  aufgebaut werden. Differenziert man (1.79)  $i$ -mal und setzt  $\bar{m} = mT_a = 0$ , erhält man

$$\frac{d^i}{d\bar{m}^i} \hat{y}[k] = \left. \frac{d^i}{d\bar{m}^i} p_m[k] \right|_{\bar{m}=0} = i! c_i[k] = i! \mathbf{h}_i^T \mathbf{y}[k] \quad (1.95)$$

bzw.

$$\frac{d^i}{d\bar{m}^i} \hat{y}[k] = i! \sum_{m=-N_q}^{N_q} \mathbf{h}_i(m + N_q + 1) y[k + m], \quad (1.96)$$

wobei  $\mathbf{h}_i(m + N_q + 1)$  der  $m + N_q + 1$ -te Eintrag von  $\mathbf{h}_i$  ist.



*Bemerkung 1.3.* Die Behandlung der Randpunkte beim Aufbau einer  $Q$ -Filtermatrix kann auf Basis eines Savitzky-Golay-Filters einfacher als auf Basis eines Gauß-Filters erfolgen, da für jede Zeile mit der gleichen Filterlänge  $M_q$  gearbeitet werden kann und lediglich für die Randzeilen der Filterpunkt  $m$  verschoben werden muss. Gleiches gilt für den Aufbau einer Ableitungsmatrix anhand eines Savitzky-Golay-Ableitungsfilters.

*Beispiel 1.3 (Savitzky-Golay-Filter).* Für  $d = 2$  lautet das Polynom (1.79)

$$p_m[k] = c_0[k] + c_1[k]T_a m + c_2[k](T_a m)^2, \quad (1.97)$$

und für  $N_q = 2$ , d. h.  $y[k + m]$ ,  $m = -2, \dots, 2$  Datenpunkte, lautet der Vektor (1.81)

$$\mathbf{y}^T[k] = [y[k-2] \quad y[k-1] \quad y[k] \quad y[k+1] \quad y[k+2]]. \quad (1.98)$$

Der Parametervektor (1.82) schreibt sich folglich als

$$\mathbf{c}^T[k] = [c_0[k] \quad c_1[k] \quad c_2[k]] \quad (1.99)$$

und der Vektor (1.83) ist gegeben durch

$$\hat{\mathbf{y}}^T[k] = [p_{-2}[k] \quad p_{-1}[k] \quad p_0[k] \quad p_1[k] \quad p_2[k]]. \quad (1.100)$$

Die  $d + 1$  Vektoren  $\mathbf{s}_i$ ,  $i = 0, 1, \dots, d$  lauten

$$\mathbf{s}_0^T = [1 \quad 1 \quad 1 \quad 1 \quad 1] \quad (1.101a)$$

$$\mathbf{s}_1^T = [-2 \quad -1 \quad 0 \quad 1 \quad 2]T_a \quad (1.101b)$$

$$\mathbf{s}_2^T = [4 \quad 1 \quad 0 \quad 1 \quad 4]T_a^2 \quad (1.101c)$$

und können in der Matrix

$$\mathbf{S} = [\mathbf{s}_0 \quad \mathbf{s}_1 \quad \mathbf{s}_2] \quad (1.102)$$

angeordnet werden. Die Matrizen (1.90) ergeben sich zu

$$\mathbf{H}^T = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T = [\mathbf{h}_0 \quad \mathbf{h}_1 \quad \mathbf{h}_2]^T = \frac{1}{35} \begin{bmatrix} -3 & -7/T_a & 5/T_a^2 \\ 12 & -3.5/T_a & -2.5/T_a^2 \\ 17 & 0 & -5/T_a^2 \\ 12 & 3.5/T_a & -2.5/T_a^2 \\ -3 & 7/T_a & 5/T_a^2 \end{bmatrix}^T \quad (1.103a)$$

$$\mathbf{B} = \mathbf{S} \mathbf{H}^T = [\mathbf{b}_{-2} \quad \mathbf{b}_{-1} \quad \mathbf{b}_0 \quad \mathbf{b}_1 \quad \mathbf{b}_2] = \frac{1}{35} \begin{bmatrix} 31 & 9 & -3 & -5 & 3 \\ 9 & 13 & 12 & 6 & -5 \\ -3 & 12 & 17 & 12 & -3 \\ -5 & 6 & 12 & 13 & 9 \\ 3 & -5 & -3 & 9 & 31 \end{bmatrix}. \quad (1.103b)$$

D. h., das gefilterte Signal am Mittelpunkt  $m = 0$  folgt der Rechenvorschrift

$$\begin{aligned} \hat{y}[k] &= p_0[k] = \mathbf{b}_0^T \mathbf{y}[k] \\ &= \frac{1}{35} (-3y[k-2] + 12y[k-1] + 17y[k] + 12y[k+1] - 3y[k+2]) \end{aligned} \quad (1.104)$$

und das Savitzky-Golay-Ableitungsfiler für die erste und zweite Ableitung am Mittelpunkt  $\bar{m} = T_a m = 0$  lautet

$$\begin{aligned} \frac{d}{d\bar{m}} \hat{y}[k] &= \frac{d}{d\bar{m}} (c_0[k] + c_1[k]\bar{m} + c_2[k]\bar{m}^2) \Big|_{\bar{m}=0} = c_1[k] = \mathbf{h}_1^T \mathbf{y}[k] \\ &= \frac{1}{35T_a} (-7y[k-2] - 3.5y[k-1] + 3.5y[k+1] + 7y[k+2]) \end{aligned} \quad (1.105)$$

bzw.

$$\begin{aligned} \frac{d^2}{d\bar{m}^2} \hat{y}[k] &= \frac{d^2}{d\bar{m}^2} (c_0[k] + c_1[k]\bar{m} + c_2[k]\bar{m}^2) \Big|_{\bar{m}=0} = 2c_2[k] = 2\mathbf{h}_2^T \mathbf{y}[k] \\ &= \frac{2}{35T_a^2} (5y[k-2] - 2.5y[k-1] - 5y[k] - 2.5y[k+1] + 5y[k+2]). \end{aligned} \quad (1.106)$$

Abbildung 1.3 zeigt die Impulsantwortfolgen  $q[k]$  und die diskreten Frequenzgänge  $G(\exp(j\omega T_a))$  eines Savitzky-Golay-Filters (S-G-F), eines Savitzky-Golay-Ableitungsfilters erster Ordnung (1. S-G-D-F) und eines Savitzky-Golay-Ableitungsfilters zweiter Ordnung (2. S-G-D-F) für  $T_a = 1$  ms,  $d = 2$  und  $N_q = 2$ .

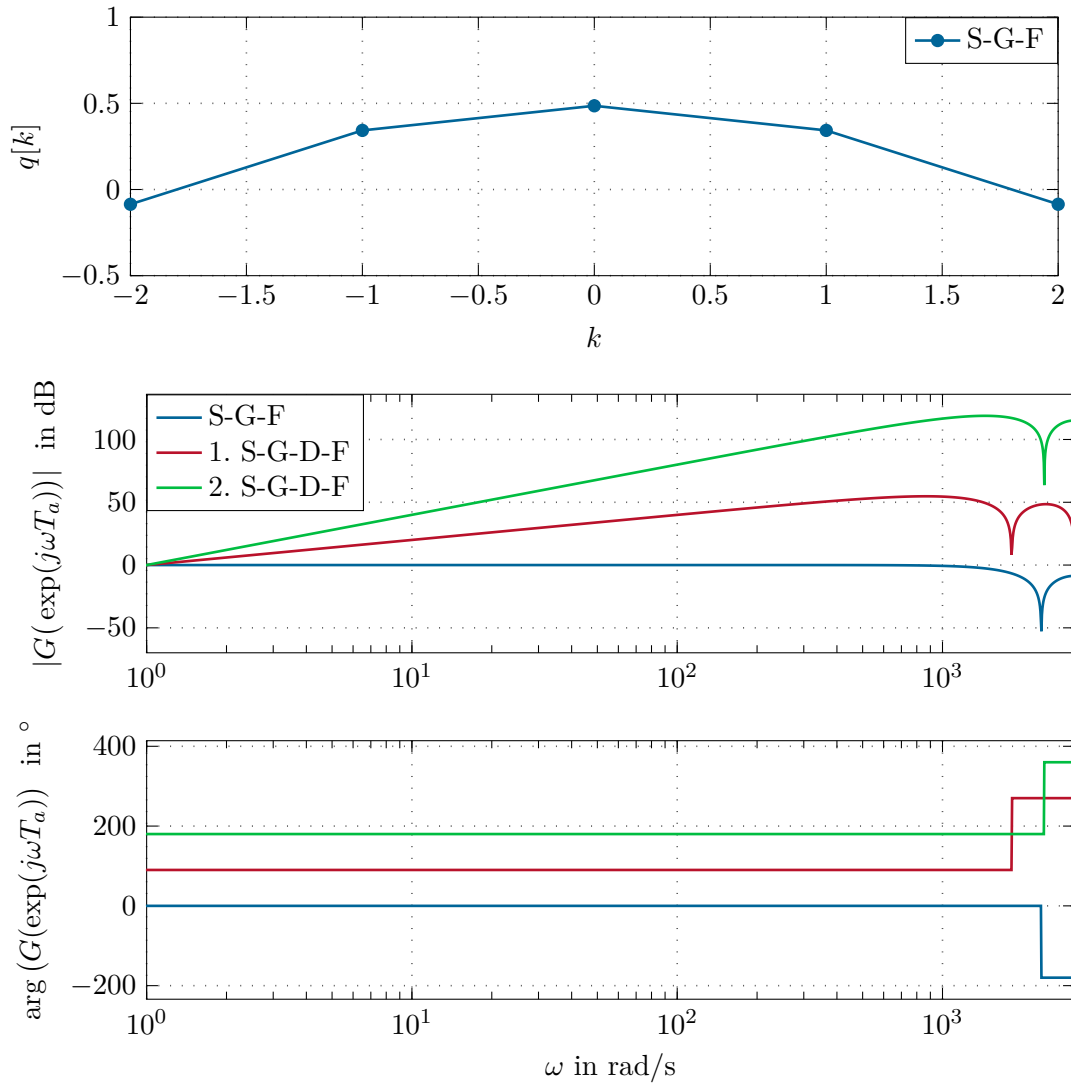


Abbildung 1.3: Impulsantwortfolge  $q[k]$  und diskreter Frequenzgang  $G(\exp(j\omega T_a))$  eines Savitzky-Golay-Filters (S-G-F), eines Savitzky-Golay-Differentiationsfilters erster Ordnung (1. S-G-D-F) und eines Savitzky-Golay-Differentiationsfilters zweiter Ordnung (2. S-G-D-F) für  $T_a = 1$  ms,  $d = 2$  und  $N_q = 2$ .

### 1.5.2 Inversionsbasierte ILR

Bei der inversionsbasierten ILR [1.14] wird die Verstärkungsmatrix zu  $\mathbf{L} = \mathbf{G}^{-1}$  gewählt, da dann

$$\mathbf{u}_{j+1} = \mathbf{Q}(\mathbf{u}_j + \mathbf{G}^{-1}\mathbf{e}_j) = \mathbf{Q}(\mathbf{u}_j + \mathbf{G}^{-1}(\mathbf{y}_d - \mathbf{G}\mathbf{u}_j - \mathbf{y}_0)) = \mathbf{Q}(\mathbf{G}^{-1}(\mathbf{y}_d - \mathbf{y}_0)) \quad (1.107)$$

gilt. Mit dieser Wahl ist Dead-Beat-Verhalten erzielbar. D. h., der Ausgangsfehler wird in einer Iteration zu  $\mathbf{e}_\infty$ :

$$\mathbf{e}_{j+1} = \mathbf{y}_d - \mathbf{G}\mathbf{Q}\left(\mathbf{G}^{-1}(\mathbf{y}_d - \mathbf{y}_0)\right) - \mathbf{y}_0 = \mathbf{e}_\infty . \quad (1.108)$$

Für  $\mathbf{Q} = \mathbf{E}$  gilt zudem  $\mathbf{e}_\infty = \mathbf{0}$ . Selbstverständlich ist dieses Ergebnis von theoretischer Natur. Zum einen ist in einer realen Anwendung  $\mathbf{G}$  nie exakt bekannt und die Verwendung eines  $Q$ -Filters unabdingbar. Zum anderen kann gezeigt werden, dass die Matrixinversion von  $\mathbf{G}$  für nicht-minimalphasige Systeme schlecht konditioniert ist. Einen weiteren Zugang zu dieser Wahl verschafft ein Blick auf die Lifted-System-Darstellung (1.34). Für  $\mathbf{y}_0 = \mathbf{0}$ ,  $\mathbf{v}_j = \mathbf{0}$ ,  $\mathbf{y}_j = \mathbf{y}_d$  und  $\mathbf{u}_j = \mathbf{u}_d$  folgt für die Steuerung

$$\mathbf{u}_d = \mathbf{G}^{-1}\mathbf{y}_d . \quad (1.109)$$

Gleichung (1.109) kann als Parametrierung des Eingangs  $\mathbf{u}_d$  durch den Ausgang  $\mathbf{y}_d$  verstanden werden. Ein Konzept, welches aus der Theorie der differentiellen Flachheit bekannt ist. Aus dieser Theorie ist bekannt, dass ein kontinuierliches, lineares System bezüglich eines Ausgangs  $\mathbf{z}$  differentiell flach ist, wenn der relative Grad  $r$  (die Anzahl von Zeitableitungen des Ausgangs bis zum ersten Mal explizit der Eingang auftritt) der Systemordnung  $n$  entspricht. Ist  $r < n$ , dann besitzt (1.20) Nullstellen. Ist (1.20) stabil und befindet sich eine dieser Nullstellen in der rechten offenen Halbebene, dann ist das System nicht-minimalphasig. Durch die Inversion werden dann die Nullstellen zu Polstellen und damit die einzelnen Einträge der Verstärkungsmatrix  $\mathbf{L} = \mathbf{G}^{-1}$  sehr groß.

*Beispiel 1.4.* Die Übertragungsfunktion  $G(z) = (z - a)/z$  besitzt eine Polstelle bei  $z = 0$  und eine Nullstelle bei  $z = a$ . Für  $|a| > 1$  ist das System nicht-minimalphasig. Die zugehörige Lifted-System-Darstellung bzw. deren Inverse lautet [1.15]

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ -a & 1 & 0 & 0 & \cdots \\ 0 & -a & 1 & 0 & \cdots \\ 0 & 0 & -a & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbf{G}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ a & 1 & 0 & 0 & \cdots \\ a^2 & a & 1 & 0 & \cdots \\ a^3 & a^2 & a & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.110)$$

Offensichtlich wachsen die Matrixeinträge in  $\mathbf{G}^{-1}$  und damit  $\mathbf{u}_d = \mathbf{G}^{-1}\mathbf{y}_d$  für  $a \gg 1$  sehr stark an. Dies wird auch deutlich, wenn man sich vor Augen führt, dass die erste Spalte von  $\mathbf{G}^{-1}$  der Impulsantwortfolge entspricht.

Dieser Umstand kann umgangen werden, indem (1.20) anhand einer regulären Zustands- transformation  $\mathbf{z} = \mathbf{T}\mathbf{y}$  auf Regelungsnormalform transformiert wird und für den neuen, flachen Ausgang  $\mathbf{z}$  eine Solltrajektorie  $\mathbf{z}_d$  geplant wird. In Regelungsnormalform besitzt das System vollen relativen Grad ( $n = r$ ) und damit weist (1.20) keine instabilen Nullstellen auf.

### 1.5.3 Normoptimale ILR

Bei der Normoptimalen ILR [1.7, 1.16] wird die  $Q$ -Filtermatrix bzw. Verstärkungsmatrix nicht direkt entworfen. Vielmehr löst man eine Minimierungsaufgabe der Form

$$\min_{\mathbf{u}_{j+1}} J(\mathbf{u}_{j+1}) = \underbrace{\frac{1}{2} \mathbf{e}_{j+1}^T \mathbf{V} \mathbf{e}_{j+1}}_{J_1(\mathbf{u}_{j+1})} + \underbrace{\frac{1}{2} \mathbf{u}_{j+1}^T \mathbf{S} \mathbf{u}_{j+1} + \frac{1}{2} (\mathbf{u}_{j+1} - \mathbf{u}_j)^T \mathbf{R} (\mathbf{u}_{j+1} - \mathbf{u}_j)}_{J_2(\mathbf{u}_{j+1})} \quad (1.111)$$

$$\text{u.B.v.} \quad \mathbf{e}_{j+1} = \mathbf{y}_d - \mathbf{y}_{j+1} = \mathbf{e}_j + \mathbf{G} \mathbf{u}_j - \mathbf{G} \mathbf{u}_{j+1}$$

im Iterationsraum. Dabei wird gefordert, dass  $\mathbf{V}$ ,  $\mathbf{S}$  und  $\mathbf{R}$  symmetrisch, positiv semi-definiten Matrizen sind und  $\mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{R}$  symmetrisch, positiv definit ist. Setzt man in (1.111) den Ausgangsfehler  $\mathbf{e}_{j+1}$  in die Zielfunktion ein, dann folgt

$$\begin{aligned} J_1(\mathbf{u}_{j+1}) &= \frac{1}{2} (\mathbf{e}_j + \mathbf{G} \mathbf{u}_j - \mathbf{G} \mathbf{u}_{j+1})^T \mathbf{V} (\mathbf{e}_j + \mathbf{G} \mathbf{u}_j - \mathbf{G} \mathbf{u}_{j+1}) \\ &= \frac{1}{2} (\mathbf{e}_j^T \mathbf{V} \mathbf{e}_j + \mathbf{u}_j^T \mathbf{G}^T \mathbf{V} \mathbf{e}_j - 2 \mathbf{u}_{j+1}^T \mathbf{G}^T \mathbf{V} \mathbf{e}_j \\ &\quad + \mathbf{e}_j^T \mathbf{V} \mathbf{G} \mathbf{u}_j + \mathbf{u}_j^T \mathbf{G}^T \mathbf{V} \mathbf{G} \mathbf{u}_j - 2 \mathbf{u}_{j+1}^T \mathbf{G}^T \mathbf{V} \mathbf{G} \mathbf{u}_j + \mathbf{u}_{j+1}^T \mathbf{G}^T \mathbf{V} \mathbf{G} \mathbf{u}_{j+1}) \end{aligned} \quad (1.112)$$

bzw.

$$\begin{aligned} J_2(\mathbf{u}_{j+1}) &= \frac{1}{2} \mathbf{u}_{j+1}^T \mathbf{S} \mathbf{u}_{j+1} + \frac{1}{2} (\mathbf{u}_{j+1}^T - \mathbf{u}_j^T) \mathbf{R} (\mathbf{u}_{j+1} - \mathbf{u}_j) \\ &= \frac{1}{2} (\mathbf{u}_{j+1}^T \mathbf{S} \mathbf{u}_{j+1} + \mathbf{u}_{j+1}^T \mathbf{R} \mathbf{u}_{j+1} - 2 \mathbf{u}_{j+1}^T \mathbf{R} \mathbf{u}_j + \mathbf{u}_j^T \mathbf{R} \mathbf{u}_j) . \end{aligned} \quad (1.113)$$

Damit  $\mathbf{u}_{j+1}$  ein Minimum des statischen Optimierungsproblems (1.111) ist, muss die notwendige Bedingung erster Ordnung

$$\left( \frac{\partial}{\partial \mathbf{u}_{j+1}} J \right) (\mathbf{u}_{j+1}) = \mathbf{0} \quad (1.114)$$

erfüllt sein. Anwendung der Optimalitätsbedingung (1.114) resultiert in

$$\left( \mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{S} + \mathbf{R} \right) \mathbf{u}_{j+1} = \left( \mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{R} \right) \mathbf{u}_j + \mathbf{G}^T \mathbf{V} \mathbf{e}_j \quad (1.115)$$

bzw. in die  $Q$ -Filtermatrix bzw. Verstärkungsmatrix

$$\mathbf{Q} = \left( \mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{S} + \mathbf{R} \right)^{-1} \left( \mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{R} \right) \quad (1.116a)$$

$$\mathbf{L} = \left( \mathbf{G}^T \mathbf{V} \mathbf{G} + \mathbf{R} \right)^{-1} \mathbf{G}^T \mathbf{V} . \quad (1.116b)$$

Laut Satz 1.4 ist die Eingangsfehleriteration (1.54) des ILR-Gesetzes (1.41), angewendet auf die Lifted-System-Darstellung (1.34), asymptotisch stabil, falls

$$\rho(\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G})) < 1 . \quad (1.117)$$

Einsetzen von (1.116) führt auf

$$\begin{aligned}\mathbf{Q}(\mathbf{E} - \mathbf{L}\mathbf{G}) &= (\mathbf{G}^T\mathbf{V}\mathbf{G} + \mathbf{S} + \mathbf{R})^{-1}(\mathbf{G}^T\mathbf{V}\mathbf{G} + \mathbf{R})\left(\mathbf{E} - (\mathbf{G}^T\mathbf{V}\mathbf{G} + \mathbf{R})^{-1}\mathbf{G}^T\mathbf{V}\mathbf{G}\right) \\ &= (\mathbf{G}^T\mathbf{V}\mathbf{G} + \mathbf{S} + \mathbf{R})^{-1}\mathbf{R}.\end{aligned}\quad (1.118)$$

Es kann gezeigt werden, dass unter den gegebenen Annahmen

$$\rho\left(\left(\mathbf{G}^T\mathbf{V}\mathbf{G} + \mathbf{S} + \mathbf{R}\right)^{-1}\mathbf{R}\right) < 1 \quad (1.119)$$

gilt. Damit ist die Stabilität der Normoptimalen ILR immer gewährleistet. Es bleibt die Frage offen, wie die Matrizen  $\mathbf{V}$ ,  $\mathbf{S}$  und  $\mathbf{R}$  gewählt werden sollen. Eine Antwort darauf gibt ein Vergleich der Normoptimalen ILR mit der Linear Quadratischen Regelung (LQR). In der Regel wird die Anzahl der Entwurfswfreiheitsgrade eingeschränkt, indem diagonale Matrizen angesetzt werden, d. h.  $\mathbf{V} = v\mathbf{E} > 0$ ,  $\mathbf{S} = s\mathbf{E} > 0$  und  $\mathbf{R} = r\mathbf{E} > 0$ . Im Einzelnen heißt das:

- Für ein großes  $v$  wird der Ausgangsfehler  $\mathbf{e}_j$  stärker bestraft und die bleibende Regelabweichung  $\mathbf{e}_\infty$  wird kleiner.
- Für ein großes  $s$  wird die Stellgröße  $\mathbf{u}_j$  stärker bestraft und die bleibende Regelabweichung  $\mathbf{e}_\infty$  wird größer. Für  $s \rightarrow 0$  gilt  $\mathbf{Q} \rightarrow \mathbf{E}$  und damit  $\mathbf{e}_\infty \rightarrow \mathbf{0}$ .
- Für ein großes  $r$  wird die Änderung der Stellgröße  $\mathbf{u}_{j+1} - \mathbf{u}_j$  stärker bestraft. Damit wird eine bessere Rauschunterdrückung erreicht. Allerdings auf Kosten der Konvergenzgeschwindigkeit. Für  $r \rightarrow 0$  stellt sich Dead-Beat-Verhalten ein.

## 1.6 Literatur

- [1.1] Z. Bien und J.-X. Xu, *Iterative learning control - analysis, design, integration and applications*. London: Springer, 1998.
- [1.2] D. Owens und J. Hätönen, „Iterative learning control - an optimization paradigm,“ *Annual Reviews in Control*, Jg. 29, Nr. 1, S. 57–70, 2005.
- [1.3] D. A. Bristow, M. Tharayil und A. Alleyne, „A survey of iterative learning control,“ *Control Systems, IEEE*, Jg. 26, Nr. 3, S. 96–114, 2006.
- [1.4] H.-S. Ahn, Y.-Q. Chen und K. Moore, „Iterative Learning Control: Brief Survey and Categorization,“ *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, Jg. 37, Nr. 6, S. 1099–1121, 2007.
- [1.5] H.-S. Ahn, K. L. Moore und Y. Chen, *Iterative learning control - robustness and monotonic convergence for interval system*. London: Springer, 2007.
- [1.6] J. M. Ortega, „Stability of Difference Equations and Convergence of Iterative Processes,“ *SIAM Journal on Numerical Analysis*, Jg. 10, Nr. 2, S. 268–282, 1973.
- [1.7] J. H. Lee, K. S. Lee und W. C. Kim, „Model-based iterative learning control with a quadratic criterion for time-varying linear systems,“ *Automatica*, Jg. 36, Nr. 5, S. 641–657, 2000.
- [1.8] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*. Dordrecht: Kluwer Academic Publishers, 2004.
- [1.9] S. Arimoto, S. Kawamura und F. Miyazaki, „Bettering operation of Robots by learning,“ *Journal of Robotic Systems*, Jg. 1, Nr. 2, S. 123–140, 1984.
- [1.10] R. W. Longman, „Iterative learning control and repetitive control for engineering practice,“ *International Journal of Control*, Jg. 73, Nr. 10, S. 930–954, 2000.
- [1.11] D. A. Bristow, „Iterative learning control for precision motion of microscale and nanoscale tracking systems,“ Diss., University of Illinois at Urban-Champaign, 2007.
- [1.12] A. Savitzky und M. Golay, „Smoothing and differentiation of data by simplified least-squares procedures,“ *Anal. Chem.*, Jg. 36, Nr. 8, S. 1627–1639, 1964.
- [1.13] R. W. Schafer, „What is a Savitzky-Golay-Filter,“ *IEEE Signal Processing Letters*, S. 111–117, 2011.
- [1.14] T. J. Harte, J. Hätönen und D. H. Owens, „Discrete-time inverse model-based iterative learning control: stability, monotonicity and robustness,“ *International Journal of Control*, Jg. 78, Nr. 8, S. 577–586, 2005.
- [1.15] S. Hashemi und J. Hammond, „The interpretation of singular vales in the inversion of minimum and non-minimum phase systems,“ *Mechanical Systems and Signal Processing*, Jg. 10, Nr. 3, S. 225–240, 1996, ISSN: 0888-3270.
- [1.16] K. Barton und A. Alleyne, „A Norm Optimal Approach to Time-Varying ILC With Application to a Multi-Axis Robotic Testbed,“ *Control Systems Technology, IEEE Transactions on*, Jg. 19, Nr. 1, S. 166–180, Jan. 2011.





## 2 Modellprädiktive Regelung

In diesem Abschnitt soll die Methode der modellprädiktiven Regelung kurz vorgestellt werden. Es werden dazu die Bestandteile und die Grundidee dieser Regelungsmethode erläutert und es werden Ansätze zum Nachweis der Stabilität von geschlossenen Regelkreisen diskutiert.

Der Begriff modellprädiktive Regelung wird im Englischen oft als *model predictive control (MPC)* bezeichnet. Mit MPC werden die folgenden Vorgehensweisen in Verbindung gebracht:

- Die Eingangsgröße  $\mathbf{u}(t)$  eines dynamischen Systems wird basierend auf einem mit einem mathematischen Modell in die Zukunft prädizierten Systemverhalten gewählt.
- In vielen Fällen beruht die Wahl von  $\mathbf{u}(t)$  auf der Lösung einer dynamischen Optimierungsaufgabe.
- Die Wahl von  $\mathbf{u}(t)$  wird kontinuierlich oder zu diskreten Zeitpunkten wiederkehrend neu durchgeführt. So können am realen System aktuell gemessene oder beobachtete Größen in der Prädiktion berücksichtigt werden. Durch diese Rückkopplung wird der Regelkreis geschlossen.

Die modellbasierte Prädiktion des Systemverhaltens und die Formulierung der zu lösenden Optimierungsaufgabe erfolgt im Allgemeinen für einen Zeithorizont, der zum aktuellen Zeitpunkt  $t$  beginnt und in die Zukunft ragt. Für die wiederkehrende Wahl von  $\mathbf{u}(t)$  muss dieser Horizont zeitlich fortgeschoben werden. Aus diesem Grund wird MPC im Englischen auch synonym als *receding horizon control (RHC)* oder *moving horizon control* bezeichnet [2.1].

Wesentliche Vorzüge von MPC sind:

- Der Methode wohnt insofern ein *intuitives, natürliches Handlungsmuster* inne, als in die Entscheidung über die aktuelle Stellgröße deren zukünftige Konsequenzen systematisch einfließen. Oft folgt auch menschliches Handeln (bewusst oder unbewusst) diesem Muster, z. B. beim Lenken eines Fahrzeuges im Straßenverkehr oder bei rollierenden Planungen in der Unternehmensführung.
- Mit der Methode können *bekannte zukünftige Ereignisse, Störungen oder Verläufe von Sollgrößen und Beschränkungen* systematisch berücksichtigt werden. Die Methode kann daher auch als *antizipatorischer* Ansatz verstanden werden.
- MPC ist eine Form der *optimalen Regelung*, denn die mit MPC berechnete Stellgröße geht meist aus der Lösung einer anwendungsspezifisch gestaltbaren, dynamischen Optimierungsaufgabe hervor. In diesem Fall kann es also (im Sinne dieser Optimierungsaufgabe) im aktuellen Zeithorizont keine bessere Stellgröße geben als jene, die MPC liefert.

- *Beschränkungen* von Eingangs-, Zustands- oder Ausgangsgrößen oder deren Zeitableitungen können systematisch berücksichtigt werden. D. h. das System kann nahe oder exakt an diesen Beschränkungen betrieben werden, was für die bestmögliche Erreichung der Regelungsziele förderlich sein kann.
- Es handelt sich um einen *relativ allgemeinen Ansatz*, der auf viele Klassen von dynamischen Systemen (z. B. auch Mehrgrößensysteme und Systeme mit Totzeitverhalten) anwendbar ist.
- Die Methode weist eine relativ gute *Übertragbarkeit* auf. Ist der Regelungsalgorithmus entwickelt und implementiert, so kann er mit relativ geringen Modifikationen und daher geringem Aufwand auf andere Systeme übertragen werden.

Während bei linearen dynamischen Systemen auch mit anderen Methoden, z. B. LQR-Regler (siehe [2.2]), ähnliche Regelungsziele erreicht werden können, sind die genannten Vorzüge in Summe im Bereich der nichtlinearen Regelung derzeit ein Alleinstellungsmerkmal von MPC. So nicht anders angeführt, wird hier der allgemeinere Fall der nichtlinearen modellprädiktiven Regelung (NMPC) behandelt.

Als Preis für die genannten Vorzüge sind zu nennen:

- MPC erfordert im Allgemeinen einen hohen *Rechenaufwand*, da meist eine dynamische Optimierungsaufgabe gelöst werden muss.
- Da die Lösung der dynamischen Optimierungsaufgabe grundsätzlich in *Echtzeit* erfolgen muss, sind je nach Systemdynamik eine effiziente Algorithmik und Programmierung sowie der Einsatz leistungsfähiger Rechner erforderlich.
- Der *Nachweis der Stabilität* des geschlossenen Kreises kann schwierig sein.
- Die erstmalige Entwicklung und Implementierung des Regelungsalgorithmus kann *aufwendig* sein.

Gängige Bücher zum Thema MPC sind z. B. [2.3–2.13], wobei in [2.12, 2.13] speziell auf NMPC eingegangen wird. Weitere Informationen finden sich auch in den Sammelbänden [2.14–2.17]. Darüber hinaus bieten die Übersichtsaufsätze [2.18–2.21] gute Einstiegspunkte in das Thema.

## 2.1 Bestandteile und Grundidee von MPC

### 2.1.1 Modell

Zur Prädiktion des zukünftigen Systemverhaltens wird ein mathematisches Modell verwendet. In vielen Fällen wird ein dynamisches Modell in Zustandsraumdarstellung herangezogen. Im zeitkontinuierlichen Fall lautet es

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad \forall t > 0, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (2.1)$$

mit den Zustandsgrößen  $\mathbf{x}(t) \in \mathbb{R}^l$ , dem Anfangszustand  $\mathbf{x}_0$ , den Eingangsgrößen  $\mathbf{u}(t) \in \mathbb{R}^m$  und der Funktion  $\mathbf{f} : \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}^l$ . Ziel der Regelung soll die Stabilisierung dieses Systems an einer Ruhelage sein. Für diese durch

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_R, \mathbf{u}_R) \quad (2.2)$$

definierte Ruhelage soll  $\mathbf{x}_R = \mathbf{0}$  und  $\mathbf{u}_R = \mathbf{0}$  gelten<sup>1</sup>. Für MPC wird häufig auch die zeitdiskrete Systemdarstellung

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k) \quad \forall k \geq 0 \quad (2.3)$$

mit dem Zeitgitter  $t_k = kT_c$ , dem Zeitindex  $k = 0, 1, 2, \dots$ , der Abtastzeit  $T_c$ , der (exakten oder näherungsweise) Repräsentation  $\mathbf{x}_k$  des Zustandes  $\mathbf{x}(t_k)$ , dem Anfangszustand  $\mathbf{x}_0$  und den Eingangsparametern  $\mathbf{u}_k \in \mathbb{R}^M$  verwendet. Die Differenzengleichung (2.3) kann als Restriktion der (exakten oder näherungsweise) Lösung von (2.1) auf das Zeitgitter interpretiert werden, wobei die Eingangsparameter  $\mathbf{u}_k$  den Eingang  $\mathbf{u}(t)$  im Intervall  $[t_k, t_{k+1})$  definieren<sup>2</sup>, d. h. es existiert eine Abbildung  $\mathbf{U} : [0, T_c) \times \mathbb{R}^M \rightarrow \mathbb{R}^m$ , so dass

$$\mathbf{u}(t) = \mathbf{U}(t - t_k, \mathbf{u}_k) \quad \forall t \in [t_k, t_{k+1}) \quad (2.4)$$

und  $\mathbf{0} = \mathbf{U}(\tau, \mathbf{0}) \quad \forall \tau \in [0, T_c)$ . Mit dieser Form der Eingangsparametrierung kann erreicht werden, dass der Suchraum der für MPC zu lösenden Optimierungsaufgabe eine finite und niedrige Dimension besitzt. Die durch (2.2) definierte Ruhelage des zeitkontinuierlichen Systems soll auch eine Ruhelage von (2.3) sein, d. h. es soll  $\mathbf{0} = \mathbf{F}(\mathbf{0}, \mathbf{0})$  gelten.

Für lineare MPC werden häufig auch Übertragungsfunktionen, Impuls- oder Sprungantwortmodelle [2.22], polynomiale Modelle [2.23] oder andere aus der Systemidentifikation [2.2, 2.24] bekannte Modelldarstellungen verwendet. Ferner ist es möglich, durch Modellerweiterungen den Einfluss von Rauschen, Störungen [2.25] sowie Parameterunsicherheiten und -variationen [2.26] in MPC Formulierungen systematisch zu berücksichtigen.

Der Einfachheit halber werden in dieser Vorlesung die nominellen und ungestörten Modelle (2.1) und (2.3) verwendet. Man beachte, dass es sich hierbei um zeitinvariante Modelle handelt, was den Reglerentwurf und die Stabilitätsanalyse vereinfachen kann. Außerdem wird auf die Verwendung einer Ausgangsgleichung des dynamischen Systems verzichtet.

## 2.1.2 Horizonte

Wie in Abbildung 2.1a angedeutet, treten bei der Realisierung von MPC zumeist zwei Zeithorizonte auf: ein *Prädiktionshorizont*  $[t, t + T]$  (Englisch: *prediction horizon*) und ein *Steuerungshorizont*  $[t, t + T_c]$  (Englisch: *control horizon*) mit  $T \geq T_c > 0$ . So nicht anders erwähnt, werden hier der Einfachheit halber Horizonte mit konstanter Länge  $T$  bzw.  $T_c$  verwendet. Das MPC Regelgesetz wird zum Zeitpunkt  $t$  ausgeführt.

Der Prädiktionshorizont  $[t, t + T]$  ist jenes zukünftige Intervall für das das Systemverhalten basierend auf dem mathematischen Modell (2.1) und dem aktuellen Zustand

<sup>1</sup>Dies stellt keine Einschränkung dar, da andere Ruhelagen sofort in den Ursprung transformiert werden können.

<sup>2</sup>Im Falle einer einfachen Abtastregelung gilt natürlich  $M = m$  und  $\mathbf{u}(t) = \mathbf{u}_k \quad \forall t \in [t_k, t_{k+1})$ .

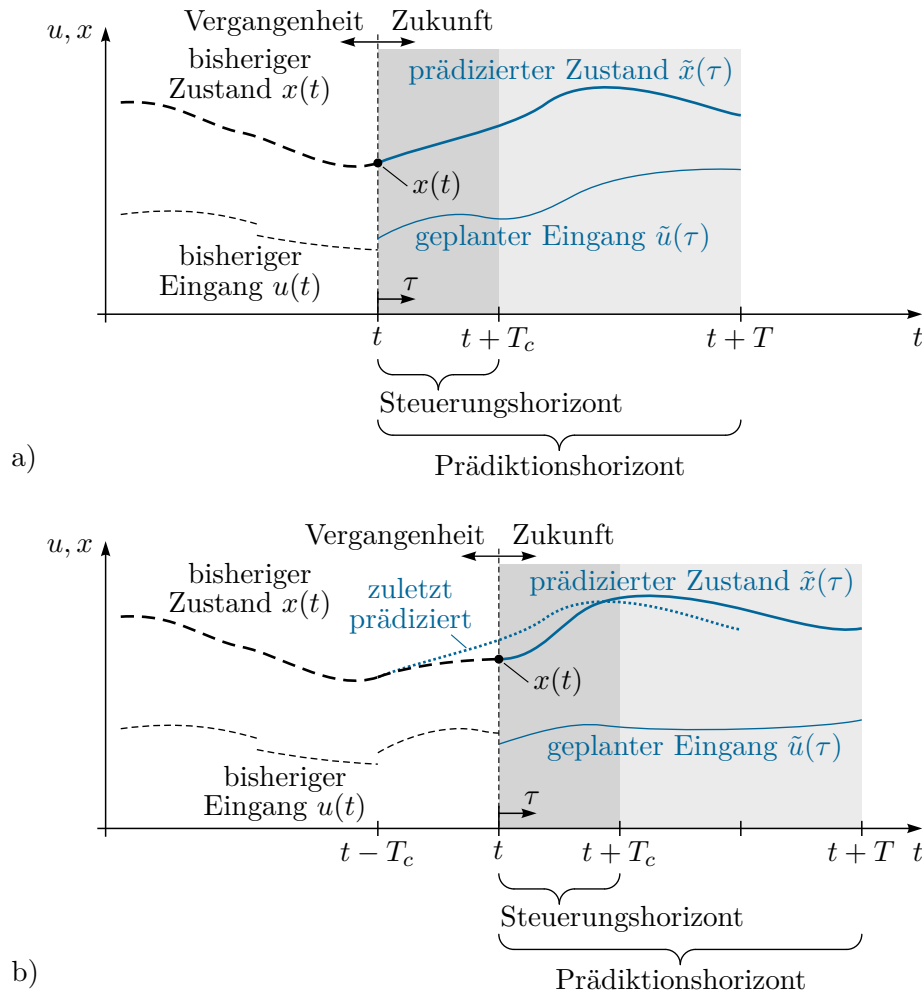


Abbildung 2.1: Horizonte und Signale einer MPC Regelung für ein System mit einem Eingang und einem Zustand, a) aktueller Horizont, b) nächstfolgender Horizont.

$\mathbf{x}(t)$  prädiziert wird. Zur Unterscheidung von den tatsächlich am System auftretenden Trajektorien  $\mathbf{u}(t)$  und  $\mathbf{x}(t)$  werden fortan die für das Intervall  $\tau \in [0, T]$  geplante Eingangstrajektorie mit  $\tilde{\mathbf{u}}(\tau)$  und die zugehörige mit (2.1) prädizierte Zustandstrajektorie mit  $\tilde{\mathbf{x}}(\tau)$  bezeichnet, wobei für den Anfangszustand  $\tilde{\mathbf{x}}(0) = \mathbf{x}(t)$  gilt. Zumeist wird für den Prädiktionshorizont eine lokale Zeitachse  $\tau$  verwendet und am aktuellen Zeitpunkt  $t$  gilt  $\tau = 0$ .

Die geplante Eingangstrajektorie  $\tilde{\mathbf{u}}(\tau)$  wird nun nicht im gesamten Prädiktionshorizont  $[t, t + T]$  auf die Strecke aufgeschaltet sondern nur im ersten Abschnitt  $[t, t + T_c]$  desselben. Dieser Abschnitt wird Steuerungshorizont genannt, denn offensichtlich ist der Regelkreis im Intervall  $(t, t + T_c)$  offen.

Am Ende des Steuerungshorizonts muss das Regelgesetz (Lösung der Optimierungsaufgabe zur Bestimmung einer neuen Steuertrajektorie  $\tilde{\mathbf{u}}(\tau)$ ) mit nun zeitlich verschobenen

Horizonten (siehe Abbildung 2.1b) erneut ausgeführt werden. Genau dann fließt der aktuell gemessene oder geschätzte Systemzustand  $\mathbf{x}(t)$  in die Berechnung von  $\tilde{\mathbf{u}}(\tau)$  ein, was den Regelkreis schließt.

Abbildung 2.1 zeigt das MPC Prinzip für ein System mit einem Eingang und einer skalaren Zustandsgröße anhand von zwei aufeinanderfolgenden Horizonten. In der Darstellung weicht die tatsächliche Zustandstrajektorie ( $\mathbf{x}(\tau)$  für  $\tau \in [t - T_c, t]$  in Abbildung 2.1b) während des ersten Steuerungshorizonts von der ursprünglich prädizierten Trajektorie ( $\tilde{\mathbf{x}}(\tau)$  für  $\tau \in [0, T_c]$  in Abbildung 2.1a) ab. Diese Abweichung im Steuerungshorizont kann beispielsweise durch Modellfehler oder unbekannte Störungen verursacht werden. Im übrigen Prädiktionshorizont (außerhalb des Steuerungshorizonts) treten solche Abweichungen im Allgemeinen immer auf, d. h. auch für das nominelle System ohne Modellfehler und Störungen. Der Grund hierfür ist, dass die prädizierten Zustandstrajektorien aus Optimierungsaufgaben mit jeweils unterschiedlichen Horizonten hervorgehen.

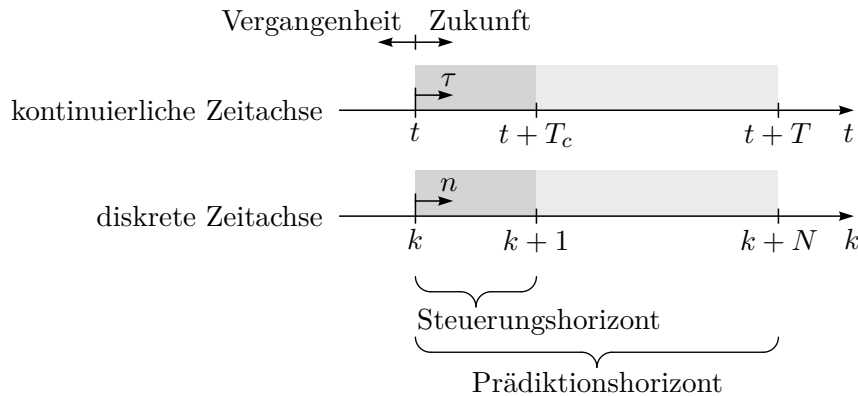


Abbildung 2.2: Horizonte bei zeitkontinuierlicher und zeitdiskreter Formulierung.

Wie in Abbildung 2.2 angedeutet, können analog zur zeitkontinuierlichen Formulierung auch im zeitdiskreten Fall ein Prädiktionshorizont  $[t_k, t_{k+N}]$  und ein Steuerungshorizont  $[t_k, t_{k+1}]$  verwendet werden. Der Prädiktionshorizont umfasse  $N$  Abtastintervalle und  $n$  sei ein lokaler Zeitindex, so dass  $n = 0$  den Beginn  $t_k$  des Prädiktionshorizonts markiere. Die geplante Eingangsfolge ( $\tilde{\mathbf{u}}_n$ ) muss nur für die Gitterpunkte  $n = 0, 1, \dots, N - 1$  (nicht aber für  $n = N$ ) definiert werden, da  $\tilde{\mathbf{u}}_N$  keinen Einfluss auf die (prädizierte) Zustandsfolge ( $\tilde{\mathbf{x}}_n$ ) im Prädiktionshorizont hat. Der Einfachheit halber soll hier der Steuerungshorizont immer genau die Länge  $T_c$  eines Abtastintervalls haben.

### 2.1.3 Beschränkungen

Bei vielen praktischen Anwendungen sind Eingangs-, Zustands- oder Ausgangsgrößen oder auch deren Zeitableitungen beschränkt. In dieser Vorlesung werden folgende Beschränkungen berücksichtigt:

$$\mathbf{x}(t) \in X, \quad \mathbf{u}(t) \in U, \quad \forall t \geq 0. \quad (2.5)$$

Hierbei soll für die Menge  $X$  der zulässigen Zustände und die Menge  $U$  der zulässigen Eingänge

$$\{\mathbf{0}\} \subset X \subseteq \mathbb{R}^l, \quad \{\mathbf{0}\} \subset U \subseteq \mathbb{R}^m \quad (2.6)$$

gelten. Zusätzlich kann es für MPC Formulierungen erforderlich sein, am Ende des Prädiktionshorizontes restriktivere Zustandsbeschränkungen in der Form

$$\tilde{\mathbf{x}}(T) \in X_T \quad (2.7)$$

zu verwenden, wobei

$$\{\mathbf{0}\} \subseteq X_T \subseteq X \quad (2.8)$$

gelten soll. Man beachte, dass die Beschränkung (2.7) an den Prädiktionshorizont gebunden ist und daher mit diesem wiederkehrend zeitlich fortgeschoben wird.

Im zeitdiskreten Fall werden statt (2.5) und (2.7) die Beschränkungen

$$\mathbf{x}_k \in X, \quad \mathbf{u}_k \in U_d, \quad \forall k = 0, 1, 2, \dots \quad (2.9a)$$

$$\tilde{\mathbf{x}}_N \in X_T \quad (2.9b)$$

verwendet, wobei die Menge  $U_d$

$$\{\mathbf{0}\} \subset U_d \subseteq \mathbb{R}^M \quad (2.10)$$

und

$$\mathbf{u} \in U_d \Rightarrow \mathbf{U}(t, \mathbf{u}) \in U \quad \forall t \in [0, T_c) \quad (2.11)$$

erfüllen soll.

### 2.1.4 Skalares Gütemaß

Zur Beurteilung von zukünftigen Eingangstrajektorien und prädizierten Zustandstrajektorien wird im zeitkontinuierlichen Fall ein skalares Gütemaß (Gütefunktional) der Form

$$J_T(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = C(\tilde{\mathbf{x}}(T)) + \int_0^T c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.12)$$

mit den Funktionen  $C : \mathbb{R}^l \rightarrow \mathbb{R}_{\geq 0}$  und  $c : \mathbb{R}_{\geq 0} \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$  verwendet. Hierbei sei  $C(\mathbf{x})$  positiv definit und  $c(t, \mathbf{x}, \mathbf{u})$  sei positiv definit bezüglich der Argumente  $\mathbf{x}$  und  $\mathbf{u}$ . Ziel bei der Wahl von  $\tilde{\mathbf{u}}(\cdot)$  ist es,  $J_T$  zu minimieren. Die Ziele der Regelung fließen in die Gestaltung von  $C$  und  $c$  ein. Da im vorliegenden Fall die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  stabilisiert werden soll und gemäß (2.2) auch  $\mathbf{u}_R = \mathbf{0}$  gilt, ist es sinnvoll  $C$  und  $c$  als positiv definite Funktionen vorzugeben. Die Gestaltung einer Gütefunktion für andere Regelungsziele, wie etwa die Stabilisierung um eine Solltrajektorie, wird z. B. in [2.13] besprochen.

Im zeitdiskreten Fall wird ein skalares Gütemaß (Gütefunktion) der Form

$$J_{d,N}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n)) = D(\tilde{\mathbf{x}}_N) + \sum_{n=0}^{N-1} d_{k+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \quad (2.13)$$

mit den positiv definiten Funktionen  $D : \mathbb{R}^l \rightarrow \mathbb{R}_{\geq 0}$  und  $d_k : \mathbb{R}^l \times \mathbb{R}^M \rightarrow \mathbb{R}_{\geq 0}$  verwendet. Die Summe in (2.13) endet beim Index  $N - 1$ , da  $\tilde{\mathbf{x}}_N$  bereits vom Endkostenterm  $D$  erfasst wird und  $\tilde{\mathbf{u}}_N$  keine Optimierungsgröße ist.

Wenn  $\tilde{\mathbf{x}}_n = \tilde{\mathbf{x}}(nT_c) \forall n = 0, 1, \dots, N$ ,  $\tilde{\mathbf{u}}(\tau + nT_c) = \mathbf{U}(\tau, \tilde{\mathbf{u}}_n) \forall (\tau, n) \in [0, T_c) \times \{0, 1, \dots, N-1\}$ ,  $T = NT_c$ ,  $D(\mathbf{x}) = C(\mathbf{x})$  und

$$d_{k+n}(\tilde{\mathbf{x}}(nT_c), \tilde{\mathbf{u}}_n) = \int_{nT_c}^{(n+1)T_c} c(t_k + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad \forall n = 0, 1, \dots, N-1, \quad (2.14)$$

dann liefern (2.12) und (2.13) identische Werte. Dies ist eine naheliegende Möglichkeit zur Gestaltung von  $D$  und  $d$  ausgehend von  $C$  und  $c$ . Man beachte, dass eine Auswertung von (2.14) implizit auch die Berechnung der Lösung von (2.1) auf dem Intervall  $\tau \in [nT_c, (n+1)T_c]$  erfordert.

### 2.1.5 Optimierung

Um im zeitkontinuierlichen Fall die Stellgröße  $\tilde{\mathbf{u}}(\tau)$  im Prädiktionshorizont  $\tau \in [0, T]$  zu wählen, wird die beschränkte dynamische Optimierungsaufgabe

$$\tilde{\mathbf{u}}^*(\cdot) = \arg \min_{\tilde{\mathbf{u}}(\cdot)} J_T(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = C(\tilde{\mathbf{x}}(T)) + \int_0^T c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.15a)$$

$$\text{u.B.v. } \dot{\tilde{\mathbf{x}}}(\tau) = \mathbf{f}(\tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)), \quad \tilde{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.15b)$$

$$\tilde{\mathbf{x}}(\tau) \in X, \quad \tilde{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, T] \quad (2.15c)$$

$$\tilde{\mathbf{x}}(T) \in X_T \quad (2.15d)$$

gelöst. In ihr sind das dynamische Modell, die Beschränkungen und das skalare Gütemaß aus den vorhergehenden Abschnitten zusammengefasst. Allgemein sei mit  $J_{t|T}^*(\mathbf{x}(t)) = J_T(t, \mathbf{x}(t), \tilde{\mathbf{u}}^*(\cdot))$  der optimale Wert des Gütefunktions in (2.15) für den Anfangszustand  $\mathbf{x}(t)$  bezeichnet. Offensichtlich ist  $J_{t|T}^*(\mathbf{x}(t))$  eine positiv definite Funktion.

Um im zeitdiskreten Fall die Stellgröße  $\tilde{\mathbf{u}}_n$  im Prädiktionshorizont  $n = 0, 1, \dots, N-1$  zu wählen, wird die beschränkte statische Optimierungsaufgabe

$$(\tilde{\mathbf{u}}_n^*) = \arg \min_{(\tilde{\mathbf{u}}_n)} J_{d,N}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n)) = D(\tilde{\mathbf{x}}_N) + \sum_{n=0}^{N-1} d_{k+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \quad (2.16a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{n+1} = \mathbf{F}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n), \quad \tilde{\mathbf{x}}_0 = \mathbf{x}_k \quad (2.16b)$$

$$\tilde{\mathbf{x}}_n \in X, \quad \tilde{\mathbf{u}}_n \in U_d, \quad \forall n = 0, 1, \dots, N-1 \quad (2.16c)$$

$$\tilde{\mathbf{x}}_N \in X_T \quad (2.16d)$$

gelöst. Allgemein sei mit  $J_{d,k|N}^*(\mathbf{x}_k) = J_{d,N}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n^*))$  der optimale Wert der Gütefunktion in (2.16) für den Anfangszustand  $\mathbf{x}_k$  bezeichnet. Offensichtlich ist  $J_{d,k|N}^*(\mathbf{x}_k)$  eine positiv definite Funktion.

Die Lösung der Optimierungsaufgaben (2.15) und (2.16) ist im Allgemeinen mit hohem numerischem Aufwand verbunden und stellt meist eine zentrale Herausforderung bei MPC Reglern dar. Einige Lösungsmethoden wurden bereits in der Vorlesung *Optimierung* [2.27] besprochen. Weitere Methoden finden sich z. B. in den Büchern [2.28–2.39]. Es existieren darüber hinaus MPC Formulierungen, bei denen auf die Lösung oder zumindest die exakte Lösung von Optimierungsaufgaben verzichtet werden kann (siehe z. B. [2.40–2.43]). Diese Formulierungen werden gelegentlich als *suboptimal MPC* bezeichnet. Der zu dieser Klasse gehörende Ansatz [2.44] wird in Abschnitt 2.2.5 diskutiert.

### 2.1.6 Annahmen

Es werden die folgenden Annahmen getroffen:

- A1) Auf dem Gebiet  $X \times U$  sei die Abbildung  $\mathbf{f}$  aus (2.1) Lipschitz-stetig in ihren Argumenten [2.45].
- A2) Auf den Gebieten  $X$  bzw.  $X \times U$  seien die Kostenfunktionen  $C$  und  $c$  aus (2.12) Lipschitz-stetig in ihren Argumenten [2.45].
- A3) Für die Kostenfunktionen  $C$  und  $c$  aus (2.12) existieren untere und obere Schranken der Form

$$\underline{C}\|\mathbf{x}\|_2^2 \leq C(\mathbf{x}) \leq \overline{C}\|\mathbf{x}\|_2^2 \quad (2.17a)$$

$$\underline{c}(\|\mathbf{x}\|_2^2 + \|\mathbf{u}\|_2^2) \leq c(t, \mathbf{x}, \mathbf{u}) \leq \overline{c}(\|\mathbf{x}\|_2^2 + \|\mathbf{u}\|_2^2) \quad \forall t \geq 0 \quad (2.17b)$$

mit geeignet gewählten Konstanten  $\underline{C}$ ,  $\overline{C}$ ,  $\underline{c}$  und  $\overline{c} \in \mathbb{R}_{>0}$ .

- A4) Auf dem Gebiet  $X \times U_d$  sei die Abbildung  $\mathbf{F}$  aus (2.3) Lipschitz-stetig in ihren Argumenten [2.45].
- A5) Auf den Gebieten  $X$  bzw.  $X \times U_d$  seien die Kostenfunktionen  $D$  und  $d_n$  aus (2.13) Lipschitz-stetig in ihren Argumenten [2.45].
- A6) Für die Kostenfunktionen  $D$  und  $d_n$  aus (2.13) existieren untere und obere Schranken der Form

$$\underline{D}\|\mathbf{x}\|_2^2 \leq D(\mathbf{x}) \leq \overline{D}\|\mathbf{x}\|_2^2 \quad (2.18a)$$

$$\underline{d}(\|\mathbf{x}\|_2^2 + \|\mathbf{u}\|_2^2) \leq d_k(\mathbf{x}, \mathbf{u}) \leq \overline{d}(\|\mathbf{x}\|_2^2 + \|\mathbf{u}\|_2^2) \quad \forall k = 0, 1, 2, \dots \quad (2.18b)$$

mit geeignet gewählten Konstanten  $\underline{D}$ ,  $\overline{D}$ ,  $\underline{d}$  und  $\overline{d} \in \mathbb{R}_{>0}$ .

Aus der Annahme A1 folgt, dass das System (2.1) für jeden beschränkten Anfangszustand  $\mathbf{x}_0 \in X$  und jede beschränkte Eingangstrajektorie  $\mathbf{u}(t) \in U$  auf dem Intervall  $[0, T]$  (für finites  $T$ ) eine eindeutige und beschränkte Lösung besitzt. Aus der Annahme A4 folgt, dass das System (2.3) für jeden beschränkten Zustand  $\mathbf{x}_k \in X$  und jeden beschränkten Eingang  $\mathbf{u}_k \in U_d$  eine eindeutige und beschränkte Lösung  $\mathbf{x}_{k+1}$  besitzt.

### 2.1.7 Regelgesetz

Die bisher genannten Bestandteile von MPC können nun zu einem Regelgesetz zusammengefasst werden. In der zeitkontinuierlichen Formulierung lautet es:

Zum Zeitpunkt  $t$  wird auf dem Prädiktionshorizont  $[t, t + T]$  die beschränkte dynamische Optimierungsaufgabe (2.15) gelöst. Hierbei sei  $\mathbf{x}(t)$  der gemessene oder geschätzte aktuelle Systemzustand. Während des Steuerungshorizonts  $[t, t + T_c)$  wird die optimierte Stellgröße

$$\mathbf{u}(\tau) = \tilde{\mathbf{u}}^*(\tau - t) \quad \forall \tau \in [t, t + T_c) \quad (2.19)$$

mit  $\tilde{\mathbf{u}}^*(\cdot)$  gemäß (2.15) auf die Strecke (2.1) aufgeschaltet. Zum Zeitpunkt  $t + T_c$



wird das Regelgesetz von Neuem ausgeführt.

Im üblichen Fall  $T_c > 0$  wird das Regelgesetz also zu zeitdiskreten Zeitpunkten ausgeführt (diskrete Regelung). Im Fall  $T_c \rightarrow 0$  (kontinuierliche Regelung) gilt für die Stellgröße  $\mathbf{u}(t) = \tilde{\mathbf{u}}^*(0)$  und (2.15) muss kontinuierlich gelöst werden.

In der zeitdiskreten Formulierung lautet das Regelgesetz wie folgt:

Zum Zeitpunkt  $t_k$  wird auf dem Prädiktionshorizont  $k, \dots, k + N$  die beschränkte Optimierungsaufgabe (2.16) gelöst. Hierbei sei  $\mathbf{x}_k$  der gemessene oder geschätzte aktuelle Systemzustand. Für das zeitdiskrete System (2.3) gilt nun, dass zum Zeitpunkt  $t_k$  die optimierte Stellgröße

$$\mathbf{u}_k = \tilde{\mathbf{u}}_0^* \quad (2.20)$$

mit  $(\tilde{\mathbf{u}}_n^*)$  gemäß (2.16) auf die Strecke aufgeschaltet wird. Für ein zeitkontinuierliches System (2.1) gilt, dass im Abtastintervall  $[t_k, t_{k+1})$  (zugleich Steuerungshorizont) die Stellgröße

$$\mathbf{u}(\tau) = \mathbf{U}(\tau - t_k, \tilde{\mathbf{u}}_0^*) \quad \forall \tau \in [t_k, t_{k+1}) \quad (2.21)$$

mit  $(\tilde{\mathbf{u}}_n^*)$  gemäß (2.16) auf die Strecke aufgeschaltet wird. Zum Zeitpunkt  $t_{k+1}$  wird das Regelgesetz von Neuem ausgeführt.

## 2.2 Stabilität

Es werden verschiedene Varianten zum Nachweis der Stabilität eines mit MPC geschlossenen Regelkreises diskutiert.

### 2.2.1 Prädiktionshorizont mit unendlicher Länge

Es wird die Stabilität des geschlossenen nominellen Regelkreises im Falle eines Prädiktionshorizontes mit unendlicher Länge analysiert. Die Ausführungen folgen im zeitkontinuierlichen Fall [2.46] und im zeitdiskreten Fall [2.13].

Für einen Prädiktionshorizont mit unendlicher Länge kann die zeitkontinuierliche Optimierungsaufgabe (2.15) in die Form

$$\tilde{\mathbf{u}}^*(\cdot) = \arg \min_{\tilde{\mathbf{u}}(\cdot)} J_\infty(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = \int_0^\infty c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.22a)$$

$$\text{u.B.v. } \dot{\tilde{\mathbf{x}}}(\tau) = \mathbf{f}(\tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)), \quad \tilde{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.22b)$$

$$\tilde{\mathbf{x}}(\tau) \in X, \quad \tilde{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, \infty) \quad (2.22c)$$

umgeschrieben werden. Dass hier die Beschränkung (2.15d) und ein allfälliger Endgewichtungsterm  $C$  keinen Einfluss auf die Lösung haben und somit entfallen können, folgt direkt aus dem Beweis des nachfolgenden Satzes.

#### Satz 2.1 (Stabilität bei Prädiktionshorizont mit unendlicher Länge, zeitkontinuierlich).

Es seien die Annahmen A1-A3 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.22) für  $\forall \mathbf{x}(t) \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}(t) \in \mathbb{R}^l \setminus X_0$ ) und der zugehörige Wert des Gütefunctionals  $J_{t|\infty}^*(\mathbf{x}(t)) < \infty$  erfüllt. Dann ist die

Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.1) lokal asymptotisch stabil mit dem Einzugsbereich  $X_0$ .

*Beweis.* Die Funktion  $J_{t|\infty}^*(\mathbf{x}(t))$  ist positiv definit und ein Kandidat für eine Lyapunovfunktion. Da sie zeitabhängig ist, ist die Stabilitätsanalyse für ein nichtautonomes System vorzunehmen [2.45]. Es sind also zwei positiv definite, zeitinvariante Funktionen  $\underline{J}_\infty^*(\mathbf{x}(t))$  und  $\bar{J}_\infty^*(\mathbf{x}(t))$  zu suchen, die

$$\underline{J}_\infty^*(\mathbf{x}) \leq J_{t|\infty}^*(\mathbf{x}) \leq \bar{J}_\infty^*(\mathbf{x}) \quad \forall \mathbf{x} \in X_0, t \geq 0 \quad (2.23)$$

erfüllen. Offensichtlich eignen sich die beiden Ersatzprobleme

$$\underline{J}_\infty^*(\mathbf{x}(t)) = \min_{\underline{\mathbf{u}}(\cdot)} \quad \underline{J}_\infty(\mathbf{x}(t), \underline{\mathbf{u}}(\cdot)) = \int_0^\infty \underline{c}(\|\underline{\mathbf{x}}(\tau)\|_2^2 + \|\underline{\mathbf{u}}(\tau)\|_2^2) d\tau \quad (2.24a)$$

$$\text{u.B.v.} \quad \dot{\underline{\mathbf{x}}}(\tau) = \mathbf{f}(\underline{\mathbf{x}}(\tau), \underline{\mathbf{u}}(\tau)), \quad \underline{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.24b)$$

$$\underline{\mathbf{x}}(\tau) \in X, \quad \underline{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, \infty) \quad (2.24c)$$

und

$$\bar{J}_\infty^*(\mathbf{x}(t)) = \min_{\bar{\mathbf{u}}(\cdot)} \quad \bar{J}_\infty(\mathbf{x}(t), \bar{\mathbf{u}}(\cdot)) = \int_0^\infty \bar{c}(\|\bar{\mathbf{x}}(\tau)\|_2^2 + \|\bar{\mathbf{u}}(\tau)\|_2^2) d\tau \quad (2.25a)$$

$$\text{u.B.v.} \quad \dot{\bar{\mathbf{x}}}(\tau) = \mathbf{f}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau)), \quad \bar{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.25b)$$

$$\bar{\mathbf{x}}(\tau) \in X, \quad \bar{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, \infty). \quad (2.25c)$$

Wegen (2.17b) und da die optimale Lösung  $\tilde{\mathbf{u}}^*(\cdot)$  von (2.22) eine zulässige aber im Allgemeinen nicht optimale Lösung von (2.24) und (2.25) ist, gilt

$$J_{t|\infty}^*(\mathbf{x}(t)) < \infty \quad \Rightarrow \quad \bar{J}_\infty^*(\mathbf{x}(t)) \leq \bar{J}_\infty(\mathbf{x}(t), \tilde{\mathbf{u}}^*(\cdot)) < \infty \quad (2.26)$$

und

$$\underline{J}_\infty^*(\mathbf{x}(t)) \leq \underline{J}_\infty(\mathbf{x}(t), \tilde{\mathbf{u}}^*(\cdot)) \leq J_{t|\infty}^*(\mathbf{x}(t)). \quad (2.27)$$

Da ferner die optimale Lösung  $\bar{\mathbf{u}}^*(\cdot)$  von (2.25) eine zulässige aber im Allgemeinen nicht optimale Lösung von (2.22) ist, gilt

$$J_{t|\infty}^*(\mathbf{x}(t)) \leq J_\infty(t, \mathbf{x}(t), \bar{\mathbf{u}}^*(\cdot)) \leq \bar{J}_\infty^*(\mathbf{x}(t)). \quad (2.28)$$

Damit ist (2.23) gezeigt.

Gemäß dem Optimalitätsprinzip nach Bellman [2.2] gilt für das geregelte System

$$J_{t+T_c|\infty}^*(\mathbf{x}(t + \Delta t)) = J_{t|\infty}^*(\mathbf{x}(t)) - \int_t^{t+T_c} c(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau \quad (2.29)$$

mit beliebigem  $T_c \in [0, T]$ . Die Optimierungsaufgabe (2.22) hat also auch zum Zeitpunkt  $t + T_c$  eine Lösung, es gilt  $\mathbf{x}(t + T_c) \in X_0$  und  $X_0$  ist eine positiv invariante Menge. Mit (2.17b) folgt aus (2.29)

$$J_{t+T_c|\infty}^*(\mathbf{x}(t + \Delta t)) \leq J_{t|\infty}^*(\mathbf{x}(t)) - \int_t^{t+T_c} \underline{c} \|\mathbf{x}(\tau)\|_2^2 d\tau . \quad (2.30)$$

Im Fall  $T_c \rightarrow 0$  (kontinuierliche Regelung), welcher hier gemäß den in Abschnitt 2.1.6 getroffenen Stetigkeitsannahmen möglich ist, ergibt sich der Grenzwert

$$\begin{aligned} J_{t|\infty}^*(\mathbf{x}(t)) &= \lim_{T_c \rightarrow 0} \frac{J_{t+T_c|\infty}^*(\mathbf{x}(t + T_c)) - J_{t|\infty}^*(\mathbf{x}(t))}{T_c} \\ &= -c(t, \mathbf{x}(t), \mathbf{u}(t)) \leq -\underline{c} \|\mathbf{x}(t)\|_2^2 . \end{aligned} \quad (2.31)$$

$J_{t|\infty}^*(\mathbf{x}(t))$  ist also eine Lyapunovfunktion und aus der direkten Methode nach Lyapunov [2.45] folgt die lokale asymptotische Stabilität der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ .

Für  $T_c \in (0, T]$  (diskrete Regelung mit  $T_c > 0$ ) kann der Regelkreis als zeitdiskretes System auf dem Zeitgitter  $t, t + T_c, t + 2T_c, \dots$  aufgefasst werden und aus (2.30) folgt

$$J_{t+T_c|\infty}^*(\mathbf{x}(t + \Delta t)) - J_{t|\infty}^*(\mathbf{x}(t)) \leq - \int_t^{t+T_c} \underline{c} \|\mathbf{x}(\tau)\|_2^2 d\tau \quad \forall t \geq 0, \mathbf{x}(t) \in X_0 . \quad (2.32)$$

Die rechte Seite dieser Ungleichung ist aufgrund der Stetigkeit von  $\mathbf{x}(t)$  negativ definit.  $J_{t|T}^*(\mathbf{x}(t))$  ist also eine Lyapunovfunktion des gedachten zeitdiskreten Systems und aus der direkten Methode nach Lyapunov für zeitdiskrete Systeme [2.47] folgt die lokale asymptotische Stabilität der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ .  $\square$

Für einen Prädiktionshorizont mit unendlicher Länge kann die zeitdiskrete Optimierungsaufgabe (2.16) in die Form

$$(\tilde{\mathbf{u}}_n^*) = \arg \min_{(\tilde{\mathbf{u}}_n)} J_{d,\infty}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n)) = \sum_{n=0}^{\infty} d_{k+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \quad (2.33a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{n+1} = \mathbf{F}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n), \quad \tilde{\mathbf{x}}_0 = \mathbf{x}_k \quad (2.33b)$$

$$\tilde{\mathbf{x}}_n \in X, \quad \tilde{\mathbf{u}}_n \in U_d, \quad \forall n = 0, 1, \dots, \infty . \quad (2.33c)$$

umgeschrieben werden. Die Beschränkung (2.16d) und ein allfälliger Endgewichtungsterm  $D$  haben wiederum keinen Einfluss auf die Lösung und können daher entfallen.

**Satz 2.2 (Stabilität bei Prädiktionshorizont mit unendlicher Länge, zeitdiskret).** *Es seien die Annahmen A4-A6 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.33) für  $\forall \mathbf{x}_k \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}_k \in \mathbb{R}^l \setminus X_0$ ) und der zugehörige Wert des Gütefunctionals  $J_{d,k|\infty}^*(\mathbf{x}_k) < \infty$  erfüllt. Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.3) lokal asymptotisch stabil mit dem Einzugsbereich  $X_0$ .*

*Beweis.* In Aufgabe 2.1 wird gezeigt, dass positiv definite, zeitinvariante Funktionen  $\underline{J}_{d,\infty}^*(\mathbf{x}_k)$  und  $\bar{J}_{d,\infty}^*(\mathbf{x}_k)$  existieren, die

$$\underline{J}_{d,\infty}^*(\mathbf{x}) \leq J_{d,k|\infty}^*(\mathbf{x}) \leq \bar{J}_{d,\infty}^*(\mathbf{x}) \quad \forall \mathbf{x} \in X_0, k \in \mathbb{N}_{\geq 0} \quad (2.34)$$

erfüllen.

Gemäß dem Optimalitätsprinzip nach Bellman [2.2] gilt für das geregelte System

$$J_{d,k+1|\infty}^*(\mathbf{x}_{k+1}) = J_{d,k|\infty}^*(\mathbf{x}_k) - d_k(\mathbf{x}_k, \mathbf{u}_k) . \quad (2.35)$$

Die Optimierungsaufgabe (2.33) hat also auch zum Zeitpunkt  $t_{k+1}$  eine Lösung, es gilt  $\mathbf{x}_{k+1} \in X_0$  und  $X_0$  ist eine positiv invariante Menge. Mit (2.18b) folgt aus (2.35)

$$J_{d,k+1|\infty}^*(\mathbf{x}_{k+1}) - J_{d,k|\infty}^*(\mathbf{x}_k) = -d_k(\mathbf{x}_k, \mathbf{u}_k) \leq -\underline{d} \|\mathbf{x}_k\|_2^2 . \quad (2.36)$$

$J_{d,k|\infty}^*(\mathbf{x}_k)$  stellt also eine Lyapunovfunktion dar und aus der direkten Methode nach Lyapunov für zeitdiskrete Systeme [2.47] folgt die lokale asymptotische Stabilität der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ .  $\square$

**Aufgabe 2.1.** Zeigen Sie, dass positiv definite, zeitinvariante Funktionen  $\underline{J}_{d,\infty}^*(\mathbf{x}_k)$  und  $\bar{J}_{d,\infty}^*(\mathbf{x}_k)$  existieren, die

$$\underline{J}_{d,\infty}^*(\mathbf{x}) \leq J_{d,k|\infty}^*(\mathbf{x}) \leq \bar{J}_{d,\infty}^*(\mathbf{x}) \quad \forall \mathbf{x} \in X_0, k \in \mathbb{N}_{\geq 0} \quad (2.37)$$

erfüllen.

Bei MPC in zeitkontinuierlicher Formulierung ist eine infinit-dimensionale Optimierungsaufgabe zu lösen. Dies gilt auch für die zeitdiskrete Formulierung von MPC, wenn ein unendlich langer Prädiktionshorizont verwendet wird. Der mit der Lösung einer infinit-dimensionalen Optimierungsaufgabe verbundene Rechenaufwand ist im Allgemeinen ebenfalls unbeschränkt, weshalb grundsätzlich nur näherungsweise Lösungen bestimmt werden können. Alle nachfolgenden Methoden zum Stabilitätsnachweis verwenden daher Prädiktionshorizonte mit endlicher Länge, so dass zumindest in der zeitdiskreten Formulierung nur eine finit-dimensionale Optimierungsaufgabe auftritt.

## 2.2.2 Endlicher Prädiktionshorizont mit vorgeschriebenem Endzustand

Es wird die Stabilität des geschlossenen nominellen Regelkreises im Falle eines endlichen Prädiktionshorizontes mit vorgeschriebenem Endzustand analysiert. Der zu erreichende Endzustand entspricht der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ . Die Ausführungen folgen im zeitkontinuierlichen Fall [2.46] und im zeitdiskreten Fall [2.13].

Für einen endlichen Prädiktionshorizont mit vorgeschriebenem Endzustand kann die

zeitkontinuierliche Optimierungsaufgabe (2.15) in die Form

$$\tilde{\mathbf{u}}^*(\cdot) = \arg \min_{\tilde{\mathbf{u}}(\cdot)} J_T(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = \int_0^T c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.38a)$$

$$\text{u.B.v. } \dot{\tilde{\mathbf{x}}}(\tau) = \mathbf{f}(\tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)), \quad \tilde{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.38b)$$

$$\tilde{\mathbf{x}}(\tau) \in X, \quad \tilde{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, T] \quad (2.38c)$$

$$\tilde{\mathbf{x}}(T) = \mathbf{0} \quad (2.38d)$$

umgeschrieben werden.

**Satz 2.3** (Stabilität bei endlichem Prädiktionshorizont mit vorgeschriebenem Endzustand, zeitkontinuierlich). *Es seien die Annahmen A1-A3 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.38) für  $\forall \mathbf{x}(t) \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}(t) \in \mathbb{R}^l \setminus X_0$ ). Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.1) lokal exponentiell stabil mit dem Einzugsbereich  $X_0$ .*

*Beweis.* Es sei  $\tilde{\mathbf{u}}^*(\cdot)$  die optimale Lösung von (2.38) und  $\tilde{\mathbf{x}}^*(\cdot)$  die zugehörige optimale Zustandstrajektorie. Ferner wird die zur Optimierungsaufgabe (2.38) gehörige Lösungsfunktion  $\tilde{\mathbf{u}}^*(\tau) = \mathbf{k}(t + \tau, \tilde{\mathbf{x}}^*(\tau))$  definiert. Wegen der Annahmen A1-A3 ist  $\mathbf{k} : \mathbb{R}_{\geq 0} \times X_0 \rightarrow U$  im Gebiet  $X_0$  Lipschitz-stetig, d. h. es existiert eine Konstante  $L_{\mathbf{k}}$ , so dass  $\|\mathbf{k}(t, \mathbf{x}) - \mathbf{k}(t, \mathbf{y})\|_2 \leq L_{\mathbf{k}}\|\mathbf{x} - \mathbf{y}\|_2 \quad \forall \mathbf{x}, \mathbf{y} \in X_0$ . Es gilt also auch

$$\|\tilde{\mathbf{u}}^*(\tau)\|_2 \leq L_{\mathbf{k}}\|\tilde{\mathbf{x}}^*(\tau)\|_2 \quad \forall \tau \in [0, T]. \quad (2.39)$$

Weiters seien  $L_{\mathbf{x}}$  und  $L_{\mathbf{u}}$  die Lipschitzkonstanten der Funktion  $\mathbf{f}$  aus (2.1) bezüglich ihres ersten und zweiten Arguments. Es gilt daher

$$\|\dot{\tilde{\mathbf{x}}}^*(\tau)\|_2 \leq \|\tilde{\mathbf{x}}^*(\tau)\|_2(L_{\mathbf{x}} + L_{\mathbf{u}}L_{\mathbf{k}}) \quad (2.40)$$

und folglich

$$\|\mathbf{x}(t)\|_2^2 e^{-2(L_{\mathbf{x}} + L_{\mathbf{u}}L_{\mathbf{k}})\tau} \leq \|\tilde{\mathbf{x}}^*(\tau)\|_2^2 \leq \|\mathbf{x}(t)\|_2^2 e^{2(L_{\mathbf{x}} + L_{\mathbf{u}}L_{\mathbf{k}})\tau} \quad \forall \tau \in [0, T]. \quad (2.41)$$

Mit den Ungleichungen (2.17b), (2.39) und (2.41) können sofort Faktoren  $\bar{J} \geq \underline{J} > 0$  berechnet werden, so dass

$$\underline{J}\|\mathbf{x}(t)\|_2^2 \leq J_{t|T}^*(\mathbf{x}(t)) \leq \bar{J}\|\mathbf{x}(t)\|_2^2 < \infty \quad \forall t \geq 0, \mathbf{x}(t) \in X_0 \quad (2.42)$$

erfüllt ist.

An die optimale Lösung  $\tilde{\mathbf{u}}^*(\cdot)$  von (2.38) kann im Zeitintervall  $(t+T, t+T+T_c]$  die zur Ruhelage  $\mathbf{x}_R = \mathbf{0}$  gehörige Stellgröße  $\mathbf{u}_R = \mathbf{0}$  angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\mathbf{u}}(\tau) = \begin{cases} \tilde{\mathbf{u}}^*(T_c + \tau) & \text{falls } \tau \in [-T_c, T - T_c] \\ \mathbf{0} & \text{falls } \tau \in [T - T_c, T] \end{cases} \quad (2.43)$$

und die zugehörige Zustandstrajektorie mit

$$\tilde{\mathbf{x}}(\tau) = \tilde{\mathbf{x}}^*(\min\{\tau + T_c, T\}), \quad \forall \tau \in [-T_c, T] \quad (2.44)$$

bezeichnet werden. Wird dem MPC Regelgesetz (2.19) entsprechend im Steuerungshorizont  $[t, t + T_c)$  die Stellgröße  $\mathbf{u}(\tau) = \tilde{\mathbf{u}}^*(\tau - t) \forall \tau \in [t, t + T_c)$  aufgeschaltet, dann ist  $\tilde{\mathbf{u}}(\cdot)$  eine zulässige aber im Allgemeinen nicht optimale Eingangstrajektorie der Optimierungsaufgabe (2.38) für den nachfolgenden Zeithorizont  $[t + T_c, t + T + T_c]$  mit dem Anfangszustand  $\mathbf{x}(t + T_c) = \tilde{\mathbf{x}}^*(T_c)$ . Folglich gilt für die optimalen Werte des Gütefunktional

$$\begin{aligned} J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) &\leq \int_0^{T-T_c} c(t + T_c + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \\ &= J_{t|T}^*(\mathbf{x}(t)) - \int_t^{t+T_c} c(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau \\ &\leq J_{t|T}^*(\mathbf{x}(t)) - \int_t^{t+T_c} \underline{c} \|\mathbf{x}(\tau)\|_2^2 d\tau. \end{aligned} \quad (2.45)$$

Für  $T_c \rightarrow 0$  (kontinuierliche Regelung) kann damit (ähnlich zum Beweis von Satz 2.1) der Grenzwert

$$\begin{aligned} j_{t|T}^*(\mathbf{x}(t)) &= \lim_{T_c \rightarrow 0} \frac{J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) - J_{t|T}^*(\mathbf{x}(t))}{T_c} \\ &= -c(t, \mathbf{x}(t), \mathbf{u}(t)) \leq -\underline{c} \|\mathbf{x}(t)\|_2^2 \end{aligned} \quad (2.46)$$

berechnet werden.  $J_{t|T}^*(\mathbf{x}(t))$  stellt also wieder eine Lyapunovfunktion dar und es folgt direkt die lokale exponentielle Stabilität der Ruhelage  $\mathbf{x}_R = \mathbf{0}$  (vgl. [2.45]). Für  $T_c \in (0, T]$  kann der Regelkreis als zeitdiskretes System auf dem Zeitgitter  $t, t + T_c, t + 2T_c, \dots$  aufgefasst werden und mit den Ungleichungen (2.17b), (2.39) und (2.41) lässt sich eine Schranke  $\Delta J > 0$  berechnen, so dass

$$J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) - J_{t|T}^*(\mathbf{x}(t)) \leq -\Delta J \|\mathbf{x}(t)\|_2^2 \quad \forall t \geq 0, \mathbf{x}(t) \in X_0. \quad (2.47)$$

$J_{t|T}^*(\mathbf{x}(t))$  ist also eine Lyapunovfunktion des gedachten zeitdiskreten Systems und es folgt direkt die lokale exponentielle Stabilität der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ .  $\square$

Für einen endlichen Prädiktionshorizont mit vorgeschriebenem Endzustand kann die zeitdiskrete Optimierungsaufgabe (2.16) in die Form

$$(\tilde{\mathbf{u}}_n^*) = \arg \min_{(\tilde{\mathbf{u}}_n)} J_{d,N}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n)) = \sum_{n=0}^{N-1} d_{k+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \quad (2.48a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{n+1} = \mathbf{F}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n), \quad \tilde{\mathbf{x}}_0 = \mathbf{x}_k \quad (2.48b)$$

$$\tilde{\mathbf{x}}_n \in X, \quad \tilde{\mathbf{u}}_n \in U_d, \quad \forall n = 0, 1, \dots, N-1 \quad (2.48c)$$

$$\tilde{\mathbf{x}}_N = \mathbf{0} \quad (2.48d)$$

umgeschrieben werden.

**Satz 2.4** (Stabilität bei endlichem Prädiktionshorizont mit vorgeschriebenem Endzustand, zeitdiskret). *Es seien die Annahmen A4-A6 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.48) für  $\forall \mathbf{x}_k \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}_k \in \mathbb{R}^l \setminus X_0$ ). Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.3) lokal exponentiell stabil mit dem Einzugsbereich  $X_0$ .*

*Beweis.* Es sei  $(\tilde{\mathbf{u}}_n^*)$  die optimale Lösung von (2.48) und  $(\tilde{\mathbf{x}}_n^*)$  die zugehörige optimale Zustandsfolge. Wegen der Annahmen A4-A6 können analog zu den Ungleichungen (2.39) und (2.41) auch im zeitdiskreten Fall Abschätzungen für  $(\tilde{\mathbf{u}}_n^*)$  und  $(\tilde{\mathbf{x}}_n^*)$  bestimmt werden. Diese Abschätzungen erlauben die Berechnung von Faktoren  $\bar{J} \geq \underline{J} > 0$ , so dass

$$\underline{J} \|\mathbf{x}_k\|_2^2 \leq J_{d,k|N}^*(\mathbf{x}_k) \leq \bar{J} \|\mathbf{x}_k\|_2^2 < \infty \quad \forall \mathbf{x}_k \in X_0, k \in \mathbb{N}_{\geq 0} \quad (2.49)$$

erfüllt ist.

An die optimale Lösung  $(\tilde{\mathbf{u}}_n^*)$  von (2.48) kann am Gitterpunkt  $k + N$  die zur Ruhelage  $\mathbf{x}_R = \mathbf{0}$  gehörige Stellgröße  $\mathbf{u}_R = \mathbf{0}$  angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\tilde{\mathbf{u}}}_n = \begin{cases} \tilde{\mathbf{u}}_{n+1}^* & \text{falls } n \in \{-1, 0, 1, \dots, N-2\} \\ \mathbf{0} & \text{falls } n = N-1 \end{cases} \quad (2.50)$$

und die zugehörige Zustandsfolge mit

$$\tilde{\tilde{\mathbf{x}}}_n = \tilde{\mathbf{x}}_{\min\{n+1, N\}}^* \quad \forall n = -1, 0, 1, \dots, N \quad (2.51)$$

bezeichnet werden. Wird dem MPC Regelgesetz (2.20) entsprechend zum Zeitpunkt  $t_k$  die Stellgröße  $\mathbf{u}_k = \tilde{\tilde{\mathbf{u}}}_0^*$  aufgeschaltet, dann ist  $(\tilde{\tilde{\mathbf{u}}}_n)$  eine zulässige aber im Allgemeinen nicht optimale Eingangsfolge der Optimierungsaufgabe (2.48) für den nachfolgenden Zeithorizont  $k+1, \dots, k+N+1$  mit dem Anfangszustand  $\mathbf{x}_{k+1} = \tilde{\tilde{\mathbf{x}}}_1^*$ . Folglich gilt für die optimalen Werte der Gütefunktion

$$\begin{aligned} J_{d,k+1|N}^*(\mathbf{x}_{k+1}) &\leq \sum_{n=0}^{N-2} d_{k+1+n}(\tilde{\tilde{\mathbf{x}}}_n, \tilde{\tilde{\mathbf{u}}}_n) \\ &= J_{d,k|N}^*(\mathbf{x}_k) - d_k(\mathbf{x}_k, \mathbf{u}_k) \\ &\leq J_{d,k|N}^*(\mathbf{x}_k) - \underline{d} \|\mathbf{x}_k\|_2^2. \end{aligned} \quad (2.52)$$

Aus (2.52) folgt

$$J_{d,k+1|N}^*(\mathbf{x}_{k+1}) - J_{d,k|N}^*(\mathbf{x}_k) \leq -\underline{d} \|\mathbf{x}_k\|_2^2. \quad (2.53)$$

$J_{d,k|N}^*(\mathbf{x}_k)$  stellt also eine Lypunovfunktion für das zeitdiskrete System dar und es folgt direkt die lokale exponentielle Stabilität [2.47] der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ .  $\square$

Die Bedingungen (2.38d) und (2.48d) erzwingen, dass der prädizierte Endzustand zum Zeitpunkt  $t + T$  bzw.  $t_{k+N}$  im Ursprung liegt. Man beachte aber, dass dies im Allgemeinen nicht für die tatsächliche Zustandstrajektorie bzw. Zustandsfolge gilt. Diese nähern sich

dem Ursprung nur asymptotisch.

Die Verwendung eines endlichen Prädiktionshorizonts mit vorgeschriebenem Endzustand kann folgende Nachteile mit sich bringen: Die Methode ist nicht geeignet für Systeme, die zwar stabilisierbar aber nicht vollständig steuerbar sind. Bei der Lösung der Optimierungsaufgabe kann die exakte Einhaltung einer Gleichungsbedingung am Ende des Prädiktionshorizontes hohen Rechenaufwand erfordern. Bei kurzen Prädiktionshorizonten, welche aus Sicht des Rechenaufwands wünschenswert wären, kann das Vorschreiben eines Endzustandes zu kleinen Einzugsbereichen  $X_0$  führen. Dies beeinträchtigt meist die Robustheit der Regelung. Bei den nachfolgenden Methoden zum Stabilitätsnachweis wird daher auf einen vorgeschriebenen Endzustand verzichtet.

### 2.2.3 Endlicher Prädiktionshorizont mit vorgeschriebenem Endgebiet und Endkostenterm

Es wird die Stabilität des geschlossenen nominellen Regelkreises im Falle eines endlichen Prädiktionshorizontes mit vorgeschriebenem Endgebiet und Endkostenterm analysiert. Die Ausführungen sind im zeitkontinuierlichen Fall an [2.46] und im zeitdiskreten Fall an [2.13] angelehnt.

**Satz 2.5** (Stabilität bei endlichem Prädiktionshorizont mit vorgeschriebenem Endgebiet und Endkostenterm, zeitkontinuierlich). *Es seien die Annahmen A1-A3 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.15) für  $\forall \mathbf{x}(t) \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}(t) \in \mathbb{R}^l \setminus X_0$ ). Es existiere ein Zustandsregelgesetz*

$$\mathbf{u}(t) = \boldsymbol{\kappa}(t, \mathbf{x}(t)) \quad (2.54)$$

mit  $\boldsymbol{\kappa} : \mathbb{R} \times X_T \rightarrow U$ , so dass die Zustandstrajektorie  $\bar{\mathbf{x}}(\tau)$  mit  $\tau \in [0, T_c]$  des mit  $\boldsymbol{\kappa}(\tau+t, \bar{\mathbf{x}}(\tau))$  geregelten Systems (2.1) für beliebige Anfangszustände  $\bar{\mathbf{x}}(0) = \mathbf{x}(t) \in X_T$

$$C(\bar{\mathbf{x}}(T_c)) - C(\mathbf{x}(t)) \leq - \int_0^{T_c} c(t + \tau, \bar{\mathbf{x}}(\tau), \boldsymbol{\kappa}(t + \tau, \bar{\mathbf{x}}(\tau))) d\tau, \quad (2.55)$$

$\bar{\mathbf{x}}(\tau) \in X \forall \tau \in [0, T_c]$  und  $\bar{\mathbf{x}}(T_c) \in X_T$  erfüllt. Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.1) lokal exponentiell stabil mit dem Einzugsbereich  $X_0$ .

*Beweis.* An die optimale Lösung  $\tilde{\mathbf{u}}^*(\cdot)$  von (2.15), wobei  $\tilde{\mathbf{x}}^*(\cdot)$  die zugehörige optimale Zustandstrajektorie sei, kann im Zeitintervall  $(t+T, t+T_c+T]$  die Eingangstrajektorie des mit dem Zustandsregelgesetz (2.54) geregelten Systems angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\mathbf{u}}(\tau) = \begin{cases} \tilde{\mathbf{u}}^*(T_c + \tau) & \text{falls } \tau \in [-T_c, T - T_c] \\ \boldsymbol{\kappa}(t + T_c + \tau, \bar{\mathbf{x}}(\tau)) & \text{falls } \tau \in [T - T_c, T] \end{cases} \quad (2.56)$$



und die zugehörige Zustandstrajektorie mit

$$\tilde{\mathbf{x}}(\tau) = \begin{cases} \tilde{\mathbf{x}}^*(T_c + \tau) & \text{falls } \tau \in [-T_c, T - T_c] \\ \bar{\mathbf{x}}(\tau) & \text{falls } \tau \in [T - T_c, T] \end{cases} \quad (2.57)$$

bezeichnet werden, wobei  $\bar{\mathbf{x}}(\tau)$  mit  $\tau \in [T - T_c, T]$  die Zustandstrajektorie des mit  $\kappa(t + T_c + \tau, \bar{\mathbf{x}}(\tau))$  geregelten Systems (2.1) für den Anfangszustand  $\bar{\mathbf{x}}(T - T_c) = \tilde{\mathbf{x}}^*(T)$  ist. Wird dem MPC Regelgesetz (2.19) entsprechend im Steuerungshorizont  $[t, t + T_c]$  die Stellgröße  $\mathbf{u}(\tau) = \tilde{\mathbf{u}}^*(\tau - t) \forall \tau \in [t, t + T_c]$  aufgeschaltet, dann ist  $\tilde{\mathbf{u}}(\cdot)$  eine zulässige aber im Allgemeinen nicht optimale Eingangstrajektorie der Optimierungsaufgabe (2.15) für den nachfolgenden Zeithorizont  $[t + T_c, t + T_c + T]$  mit dem Anfangszustand  $\mathbf{x}(t + T_c) = \tilde{\mathbf{x}}^*(T_c)$ . Unter Berücksichtigung von (2.55) gilt folglich für die optimalen Werte des Gütefunktional

$$\begin{aligned} J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) &\leq C(\tilde{\mathbf{x}}(T - T_c)) + \int_0^{T-T_c} c(t + T_c + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \\ &= J_{t|T}^*(\mathbf{x}(t)) - \int_t^{t+T_c} c(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau . \end{aligned} \quad (2.58)$$

Der Rest des Beweises erfolgt völlig analog zum Beweis von Satz 2.3.  $\square$

Bevor die zeitdiskrete Variante dieser Methode diskutiert wird, sollen kurz verschiedene Interpretationen des Ansatzes gegeben werden.

- Aus dem Zustandsregelgesetz (2.54) und der Ungleichung (2.55) folgt die Existenz einer Eingangstrajektorie, die sicherstellt, dass bei einer Verlängerung des Prädiktionshorizonts um  $T_c$  der Kostenzuwachs zufolge des integralen Kostenterms  $c$  durch eine Kostenreduktion zufolge des geänderten Endkostenterms  $C$  zumindest wettgemacht wird.
- Der Ansatz wird gelegentlich als *MPC mit quasi-unendlichem Horizont* bezeichnet [2.48], da gilt

$$C(\mathbf{x}(t)) \geq \int_t^\infty c(\tau, \mathbf{x}(\tau), \kappa(\tau, \mathbf{x}(\tau))) d\tau , \quad (2.59)$$

sofern  $\mathbf{x}(t) \in X_T$  und das System im Horizont  $[t, \infty)$  mit dem Zustandsregelgesetz (2.54) betrieben wird. D. h. der Endkostenterm  $C(\mathbf{x}(t))$  ist eine obere Schranke für den integralen Kostenterm in der gesamten Zukunft. Um (2.59) zu zeigen, kann man einfach (2.55) rekursiv einsetzen. Wegen (2.59) und  $J_{t|T}^*(\mathbf{x}(t)) < \infty \forall \mathbf{x}(t) \in X_0$  kann der Stabilitätsbeweis auch analog zu jenem von Satz 2.1 für den unendlichen Prädiktionshorizont erfolgen.

- Man beachte, dass die Forderung (2.55) im Zusammenspiel mit den Annahmen A1 und A3 impliziert, dass der mit dem Zustandsregelgesetz  $\kappa(t, \mathbf{x}(t))$  und dem System (2.1) geschlossene Kreis im Sinne eines zeitdiskreten Systems mit dem Abtastintervall  $T_c$  lokal exponentiell stabil mit dem Einzugsbereich  $X_T$  sein muss. Allgemein kann der Entwurf eines stabilisierenden Regelgesetzes, das einen für die jeweilige Anwendung hinreichend großen Einzugsbereich erlaubt, schwierig sein. Dies

kann grundsätzlich auch für den Entwurf des Zustandsregelgesetzes  $\kappa(t, \mathbf{x}(t))$ , das den in Satz 2.5 genannten Anforderungen zu genügen hat und daher einen Einzugsbereich  $X_T$  sicherstellen muss, gelten. Wenn der Einzugsbereich  $X_T$  für die Anwendung ausreicht und auch sonst alle Regelungsziele bereits mit dem Zustandsregelgesetz  $\kappa(t, \mathbf{x}(t))$  erreicht werden, kann auf die Verwendung von MPC verzichtet werden. Wenn jedoch  $X_T$  keine zufriedenstellende Ausdehnung besitzt, kann MPC als Strategie verstanden werden um den vom Zustandsregler erreichten Einzugsbereich  $X_T$  auf  $X_0$  auszuweiten.

- Mit der speziellen Wahl  $X_T = \{\mathbf{0}\}$  erhält man sofort die in Abschnitt 2.2.2 beschriebene Formulierung. D. h. letztere ist ein Spezialfall des hier vorgestellten Ansatzes.

**Satz 2.6** (Stabilität bei endlichem Prädiktionshorizont mit vorgeschriebenem Endgebiet und Endkostenterm, zeitdiskret). *Es seien die Annahmen A4-A6 erfüllt.  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.16) für  $\forall \mathbf{x}_k \in X_0$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}_k \in \mathbb{R}^l \setminus X_0$ ). Es existiere ein Zustandsregelgesetz*

$$\mathbf{u}_k = \kappa_k(\mathbf{x}_k) \quad (2.60)$$

mit  $\kappa_k : X_T \rightarrow U_d$ , so dass für beliebige Anfangszustände  $\mathbf{x}_k \in X_T$

$$D(\mathbf{F}(\mathbf{x}_k, \kappa_k(\mathbf{x}_k))) - D(\mathbf{x}_k) \leq -d_k(\mathbf{x}_k, \kappa_k(\mathbf{x}_k)) \quad (2.61)$$

und  $\mathbf{F}(\mathbf{x}_k, \kappa_k(\mathbf{x}_k)) \in X_T$  erfüllt sind. Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des geregelten Systems (2.3) lokal exponentiell stabil mit dem Einzugsbereich  $X_0$ .

*Beweis.* An die optimale Lösung  $(\tilde{\mathbf{u}}_n^*)$  von (2.16), wobei  $(\tilde{\mathbf{x}}_n^*)$  die zugehörige optimale Zustandsfolge sei, kann am Gitterpunkt  $k + N$  die Stellgröße  $\kappa_{k+N}(\tilde{\mathbf{x}}_N^*)$  gemäß dem Zustandsregelgesetz (2.60) angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\tilde{\mathbf{u}}}_n = \begin{cases} \tilde{\mathbf{u}}_{n+1}^* & \text{falls } n \in \{-1, 0, 1, \dots, N-2\} \\ \kappa_{k+N}(\tilde{\mathbf{x}}_N^*) & \text{falls } n = N-1 \end{cases} \quad (2.62)$$

und die zugehörige Zustandsfolge mit

$$\tilde{\tilde{\mathbf{x}}}_n = \begin{cases} \tilde{\mathbf{x}}_{n+1}^* & \text{falls } n \in \{-1, 0, 1, \dots, N-1\} \\ F(\tilde{\mathbf{x}}_N^*, \kappa_{k+N}(\tilde{\mathbf{x}}_N^*)) & \text{falls } n = N \end{cases} \quad (2.63)$$

bezeichnet werden. Wird dem MPC Regelgesetz (2.20) entsprechend zum Zeitpunkt  $t_k$  die Stellgröße  $\mathbf{u}_k = \tilde{\tilde{\mathbf{u}}}_0^*$  aufgeschaltet, dann ist  $(\tilde{\tilde{\mathbf{u}}}_n)$  eine zulässige aber im Allgemeinen nicht optimale Eingangsfolge der Optimierungsaufgabe (2.16) für den nachfolgenden Zeithorizont  $k+1, \dots, k+N+1$  mit dem Anfangszustand  $\mathbf{x}_{k+1} = \tilde{\tilde{\mathbf{x}}}_1^*$ .

Unter Berücksichtigung von (2.61) gilt folglich für die optimalen Werte der Güte-

funktion

$$\begin{aligned} J_{d,k+1|N}^*(\mathbf{x}_{k+1}) &\leq D(\tilde{\mathbf{x}}_{N-1}) + \sum_{n=0}^{N-2} d_{k+1+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \\ &= J_{d,k|N}^*(\mathbf{x}_k) - d_k(\mathbf{x}_k, \mathbf{u}_k) . \end{aligned} \quad (2.64)$$

Der Rest des Beweises erfolgt völlig analog zum Beweis von Satz 2.4.  $\square$

## 2.2.4 Endlicher Prädiktionshorizont mit Endkostenterm

Es wird ein zu der in Abschnitt 2.2.3 beschriebenen MPC Variante verwandter Ansatz vorgestellt, der kein vorgeschriebenes Endgebiet benötigt. Dies vereinfacht in der Regel die zugrunde liegende dynamische Optimierungsaufgabe erheblich (freier Endzustand). Der hier vorgestellte Ansatz wurde ursprünglich in [2.49] vorgeschlagen. Die nachfolgende Stabilitätsanalyse des geschlossenen nominellen Regelkreises ist an [2.10] angelehnt.

Für einen endlichen Prädiktionshorizont ohne vorgeschriebenem Endzustand entfällt in der zeitkontinuierlichen Optimierungsaufgabe (2.15) die Bedingung (2.15d). Im Allgemeinen kann der Einzugsbereich beeinflusst werden, indem der Endkostenterm  $C$  um einen konstanten Faktor  $\gamma \geq 1$  vergrößert wird. Die Optimierungsaufgabe (2.15) kann dann in die Form

$$\tilde{\mathbf{u}}^*(\cdot) = \arg \min_{\tilde{\mathbf{u}}(\cdot)} J_T(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = \gamma C(\tilde{\mathbf{x}}(T)) + \int_0^T c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.65a)$$

$$\text{u.B.v. } \dot{\tilde{\mathbf{x}}}(\tau) = \mathbf{f}(\tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) , \quad \tilde{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.65b)$$

$$\tilde{\mathbf{x}}(\tau) \in X , \quad \tilde{\mathbf{u}}(\tau) \in U , \quad \forall \tau \in [0, T] \quad (2.65c)$$

umgeschrieben werden. Es wird von der Existenz eines Zustandsregelgesetzes  $\kappa(t, \mathbf{x}(t))$ , welches das System für alle Anfangszustände aus einem abgeschlossenen Gebiet  $X_\Gamma$  stabilisiert, ausgegangen. Hierbei sei  $X_\Gamma$  eine Niveaumenge des Endkostenterms  $C(\tilde{\mathbf{x}}(T))$  mit dem Niveau  $\Gamma > 0$ , d. h.

$$X_\Gamma = \{\mathbf{x} \in X | C(\mathbf{x}) \leq \Gamma\} \supset \{\mathbf{0}\} . \quad (2.66)$$

Anstatt nun  $X_\Gamma$  als Endgebiet in der Optimierungsaufgabe vorzuschreiben, wird eine Niveaumenge  $\underline{X}_0$  von  $J_{t|T}^*(\mathbf{x}(t))$  gesucht, so dass die MPC Regelung für alle Anfangszustände aus  $\underline{X}_0$  sicherstellt, dass die prädizierte Trajektorie  $\tilde{\mathbf{x}}(\tau)$  innerhalb des Prädiktionshorizonts mit der Länge  $T$  in das Gebiet  $X_\Gamma$  einläuft.

**Satz 2.7** (Stabilität bei endlichem Prädiktionshorizont mit Endkostenterm und ohne vorgeschriebenem Endgebiet, zeitkontinuierlich). *Es seien die Annahmen A1-A3 erfüllt. Es existiere ein Zustandsregelgesetz*

$$\mathbf{u}(t) = \kappa(t, \mathbf{x}(t)) \quad (2.67)$$

mit  $\kappa : \mathbb{R} \times X_\Gamma \rightarrow U$  und eine zugehörige Konstante  $\Gamma > 0$ , so dass

$$\frac{dC(\mathbf{x})}{d\mathbf{x}} \mathbf{f}(\mathbf{x}, \kappa(t, \mathbf{x})) \leq -c(t, \mathbf{x}, \kappa(t, \mathbf{x})) \quad \forall \mathbf{x} \in X_\Gamma, t \geq 0. \quad (2.68)$$

Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des mit dem MPC Regelgesetz (2.19) entsprechend der Optimierungsaufgabe (2.65) geregelten Systems (2.1) lokal exponentiell stabil mit einem Einzugsbereich, der

$$\underline{X}_0 = \left\{ \mathbf{x} \in X \mid J_{t|T}^*(\mathbf{x}) \leq \Gamma \left( \gamma + T \frac{c}{C} \right) \quad \forall t \geq 0 \right\} \supset \{\mathbf{0}\} \quad (2.69)$$

beinhaltet.

*Beweis.* Es sei  $\tilde{\mathbf{u}}^*(\cdot)$  die optimale Lösung von (2.65) und  $\tilde{\mathbf{x}}^*(\cdot)$  die zugehörige optimale Zustandstrajektorie, wobei  $\tilde{\mathbf{x}}^*(0) = \mathbf{x}(t) \in \underline{X}_0$  gelten soll. Aus (2.17a) folgt die Implikation

$$\|\mathbf{x}\|_2^2 \leq \frac{\Gamma}{C} \quad \Rightarrow \quad \mathbf{x} \in X_\Gamma. \quad (2.70)$$

Wegen der Ungleichung

$$\begin{aligned} \Gamma \left( \gamma + T \frac{c}{C} \right) &\geq J_{t|T}^*(\mathbf{x}(t)) = \gamma C(\tilde{\mathbf{x}}^*(T)) + \int_0^T c(t + \tau, \tilde{\mathbf{x}}^*(\tau), \tilde{\mathbf{u}}^*(\tau)) d\tau \\ &\geq \gamma C(\tilde{\mathbf{x}}^*(T)) + \int_0^T c \|\tilde{\mathbf{x}}^*(\tau)\|_2^2 d\tau \end{aligned} \quad (2.71)$$

(vgl. (2.17b)) muss es also einen Zeitpunkt  $\bar{\tau} \in [0, T]$  geben, so dass  $\tilde{\mathbf{x}}^*(\bar{\tau}) \in X_\Gamma$ . Es sei nun  $\bar{\mathbf{x}}(\tau)$  mit  $\tau \in [\bar{\tau}, T]$  die Zustandstrajektorie des mit  $\kappa(t + \tau, \bar{\mathbf{x}}(\tau))$  geregelten Systems (2.1) für den Anfangszustand  $\bar{\mathbf{x}}(\bar{\tau}) = \tilde{\mathbf{x}}^*(\bar{\tau})$ . Wegen (2.68) und der Suboptimalität der Eingangstrajektorie  $\kappa(t + \tau, \bar{\mathbf{x}}(\tau))$  für  $\tau \in [\bar{\tau}, T]$  gilt

$$\begin{aligned} \Gamma &\geq C(\tilde{\mathbf{x}}^*(\bar{\tau})) \geq C(\bar{\mathbf{x}}(T)) + \int_{\bar{\tau}}^T c(t + \tau, \bar{\mathbf{x}}(\tau), \kappa(t + \tau, \bar{\mathbf{x}}(\tau))) d\tau \\ &\geq C(\tilde{\mathbf{x}}^*(T)) + \int_{\bar{\tau}}^T c(t + \tau, \tilde{\mathbf{x}}^*(\tau), \tilde{\mathbf{u}}^*(\tau)) d\tau \geq C(\tilde{\mathbf{x}}^*(T)). \end{aligned} \quad (2.72)$$

Daraus folgt  $\tilde{\mathbf{x}}^*(T) \in X_\Gamma$ . An die optimale Lösung  $\tilde{\mathbf{u}}^*(\cdot)$  von (2.65) kann folglich im Zeitintervall  $(t+T, t+T_c+T]$  die Eingangstrajektorie des mit dem Zustandsregelgesetz (2.67) geregelten Systems angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\tilde{\mathbf{u}}}(\tau) = \begin{cases} \tilde{\mathbf{u}}^*(T_c + \tau) & \text{falls } \tau \in [-T_c, T - T_c] \\ \kappa(t + T_c + \tau, \bar{\mathbf{x}}(\tau)) & \text{falls } \tau \in [T - T_c, T] \end{cases} \quad (2.73)$$

und die zugehörige Zustandstrajektorie mit

$$\tilde{\mathbf{x}}(\tau) = \begin{cases} \tilde{\mathbf{x}}^*(T_c + \tau) & \text{falls } \tau \in [-T_c, T - T_c] \\ \bar{\mathbf{x}}(\tau) & \text{falls } \tau \in [T - T_c, T] \end{cases} \quad (2.74)$$

bezeichnet werden, wobei  $\bar{\mathbf{x}}(\tau)$  mit  $\tau \in [T - T_c, T]$  die Zustandstrajektorie des mit  $\kappa(t + T_c + \tau, \bar{\mathbf{x}}(\tau))$  geregelten Systems (2.1) für den Anfangszustand  $\bar{\mathbf{x}}(T - T_c) = \tilde{\mathbf{x}}^*(T)$  ist. Wird dem MPC Regelgesetz (2.19) entsprechend im Steuerungshorizont  $[t, t + T_c]$  die Stellgröße  $\mathbf{u}(\tau) = \tilde{\mathbf{u}}^*(\tau - t) \forall \tau \in [t, t + T_c]$  aufgeschaltet, dann ist  $\tilde{\mathbf{u}}(\cdot)$  eine zulässige aber im Allgemeinen nicht optimale Eingangstrajektorie der Optimierungsaufgabe (2.65) für den nachfolgenden Zeithorizont  $[t + T_c, t + T_c + T]$  mit dem Anfangszustand  $\mathbf{x}(t + T_c) = \tilde{\mathbf{x}}^*(T_c)$ . Unter Berücksichtigung von (2.68) gilt folglich für die optimalen Werte des Gütefunktional

$$\begin{aligned} J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) &\leq \gamma C(\tilde{\mathbf{x}}(T - T_c)) + \int_0^{T-T_c} c(t + T_c + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \\ &= J_{t|T}^*(\mathbf{x}(t)) - \int_t^{t+T_c} c(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau . \end{aligned} \quad (2.75)$$

Daraus folgt nun noch

$$\Gamma\left(\gamma + T \frac{c}{C}\right) \geq J_{t|T}^*(\mathbf{x}(t)) \geq J_{t+T_c|T}^*(\mathbf{x}(t + T_c)) \quad (2.76)$$

und somit  $\mathbf{x}(t + T_c) \in \underline{X}_0$ . Der Rest des Beweises erfolgt völlig analog zum Beweis von Satz 2.3.  $\square$

Anhand der Definition (2.69) wird klar, welche Parameter des Entwurfs zu verändern sind um einen größeren Einzugsbereich zu erhalten. Für eine möglichst große Menge  $\underline{X}_0$  sollte der Niveauewert  $\Gamma$  maximal sein.  $\Gamma$  wird jedoch meist durch das verwendete Zustandsregelgesetz  $\kappa(t, \bar{\mathbf{x}}(t))$  limitiert. Die Länge  $T$  des Prädiktionshorizonts wirkt sich direkt proportional auf das Niveau von  $\underline{X}_0$  aus. Lange Prädiktionshorizonte sind aufgrund des damit verbundenen numerischen Aufwands beim Lösen der zugrundeliegenden Optimierungsaufgabe oft nicht erwünscht. Ferner bewirkt eine Erhöhung des Verstärkungsfaktors  $\gamma$  eine Vergrößerung von  $\underline{X}_0$ . Natürlich wirkt sich dies aber auch auf das Regelverhalten aus, da der integrale Kostenterm  $c$  gegenüber dem Endkostenterm  $\gamma C$  an Bedeutung verliert.

Für einen endlichen Prädiktionshorizont ohne vorgeschriebenem Endzustand und mit erhöhter Gewichtung des Endkostenterms kann die zeitdiskrete Optimierungsaufgabe (2.16) in die Form

$$(\tilde{\mathbf{u}}_n^*) = \arg \min_{(\tilde{\mathbf{u}}_n)} J_{d,N}(k, \mathbf{x}_k, (\tilde{\mathbf{u}}_n)) = \gamma D(\tilde{\mathbf{x}}_N) + \sum_{n=0}^{N-1} d_{k+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \quad (2.77a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{n+1} = \mathbf{F}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n), \quad \tilde{\mathbf{x}}_0 = \mathbf{x}_k \quad (2.77b)$$

$$\tilde{\mathbf{x}}_n \in X, \quad \tilde{\mathbf{u}}_n \in U_d, \quad \forall n = 0, 1, \dots, N-1 \quad (2.77c)$$

umgeschrieben werden, wobei wieder  $\gamma \geq 1$  gilt. Es wird wieder eine Niveaumenge

$$X_\Gamma = \{\mathbf{x} \in X | D(\mathbf{x}) \leq \Gamma\} \supset \{\mathbf{0}\} \quad (2.78)$$

definiert.

**Satz 2.8** (Stabilität bei endlichem Prädiktionshorizont mit Endkostenterm und ohne vorgeschriebenem Endgebiet, zeitdiskret). *Es seien die Annahmen A4-A6 erfüllt. Es existiere ein Zustandsregelgesetz*

$$\mathbf{u}_k = \boldsymbol{\kappa}_k(\mathbf{x}_k) \quad (2.79)$$

mit  $\boldsymbol{\kappa}_k : X_\Gamma \rightarrow U_d$  und eine zugehörige Konstante  $\Gamma$ , so dass

$$D(\mathbf{F}(\mathbf{x}_k, \boldsymbol{\kappa}_k(\mathbf{x}_k))) - D(\mathbf{x}_k) \leq -d_k(\mathbf{x}_k, \boldsymbol{\kappa}_k(\mathbf{x}_k)) \quad \forall \mathbf{x}_k \in X_\Gamma, k \in \mathbb{N}_{\geq 0} . \quad (2.80)$$

Dann ist die Ruhelage  $\bar{\mathbf{x}}_R = \mathbf{0}$  des mit dem MPC Regelgesetz (2.20) entsprechend der Optimierungsaufgabe (2.77) geregelten Systems (2.3) lokal exponentiell stabil mit einem Einzugsbereich, der

$$\underline{X}_0 = \left\{ \mathbf{x} \in X \mid J_{d,k|N}^*(\mathbf{x}) \leq \Gamma \left( \gamma + N \frac{d}{D} \right) \quad \forall k \in \mathbb{N}_{\geq 0} \right\} \supset \{ \mathbf{0} \} \quad (2.81)$$

beinhaltet.

*Beweis.* Es sei  $(\bar{\mathbf{u}}_n^*)$  die optimale Lösung von (2.77) und  $(\bar{\mathbf{x}}_n^*)$  die zugehörige optimale Zustandsfolge, wobei  $\bar{\mathbf{x}}^*(0) = \mathbf{x}_k \in \underline{X}_0$  gelten soll. Aus (2.18a) folgt die Implikation

$$\|\mathbf{x}\|_2^2 \leq \frac{\Gamma}{D} \quad \Rightarrow \quad \mathbf{x} \in X_\Gamma . \quad (2.82)$$

Wegen der Ungleichung

$$\begin{aligned} \Gamma \left( \gamma + N \frac{d}{D} \right) &\geq J_{d,k|N}^*(\mathbf{x}_k) = \gamma D(\bar{\mathbf{x}}_N^*) + \sum_{n=0}^{N-1} d_{k+n}(\bar{\mathbf{x}}_n^*, \bar{\mathbf{u}}_n^*) \\ &\geq \gamma D(\bar{\mathbf{x}}_N^*) + \sum_{n=0}^{N-1} \underline{d} \|\bar{\mathbf{x}}_n^*\|_2^2 \end{aligned} \quad (2.83)$$

(vgl. (2.18b)) muss es also einen Zeitindex  $\bar{n} \in \{0, 1, \dots, N\}$  geben, so dass  $\bar{\mathbf{x}}_{\bar{n}}^* \in X_\Gamma$ . Es sei nun  $(\bar{x}_n)$  mit  $n = \bar{n}, \bar{n} + 1, \dots, N$  die Zustandsfolge des mit  $\boldsymbol{\kappa}_{k+n}(\bar{\mathbf{x}}_n)$  geregelten Systems (2.3) für den Anfangszustand  $\bar{\mathbf{x}}_{\bar{n}} = \bar{\mathbf{x}}_{\bar{n}}^*$ . Wegen (2.80) und der Suboptimalität der Eingangsfolge  $(\boldsymbol{\kappa}_{k+n}(\bar{\mathbf{x}}_n))$  für  $n = \bar{n}, \bar{n} + 1, \dots, N - 1$  gilt

$$\begin{aligned} \Gamma &\geq D(\bar{\mathbf{x}}_{\bar{n}}^*) \geq D(\bar{\mathbf{x}}_N) + \sum_{n=\bar{n}}^{N-1} d_{k+n}(\bar{\mathbf{x}}_n, \boldsymbol{\kappa}_{k+n}(\bar{\mathbf{x}}_n)) \\ &\geq D(\bar{\mathbf{x}}_N^*) + \sum_{n=\bar{n}}^{N-1} d_{k+n}(\bar{\mathbf{x}}_n^*, \bar{\mathbf{u}}_n^*) \geq D(\bar{\mathbf{x}}_N^*) . \end{aligned} \quad (2.84)$$

Daraus folgt  $\bar{\mathbf{x}}_N^* \in X_\Gamma$ . An die optimale Lösung  $(\bar{\mathbf{u}}_n^*)$  von (2.77) kann folglich am Gitterpunkt  $k + N$  die Stellgröße  $\boldsymbol{\kappa}_{k+N}(\bar{\mathbf{x}}_{k+N}^*)$  gemäß dem Zustandsregelgesetz (2.79)

angefügt werden. Die so erweiterte Lösung soll mit

$$\tilde{\mathbf{u}}_n = \begin{cases} \tilde{\mathbf{u}}_{n+1}^* & \text{falls } n \in \{-1, 0, 1, \dots, N-2\} \\ \boldsymbol{\kappa}_{k+N}(\tilde{\mathbf{x}}_N^*) & \text{falls } n = N-1 \end{cases} \quad (2.85)$$

und die zugehörige Zustandsfolge mit

$$\tilde{\mathbf{x}}_n = \begin{cases} \tilde{\mathbf{x}}_{n+1}^* & \text{falls } n \in \{-1, 0, 1, \dots, N-1\} \\ F(\tilde{\mathbf{x}}_N^*, \boldsymbol{\kappa}_{k+N}(\tilde{\mathbf{x}}_N^*)) & \text{falls } n = N \end{cases} \quad (2.86)$$

bezeichnet werden. Wird dem MPC Regelgesetz (2.20) entsprechend zum Zeitpunkt  $t_k$  die Stellgröße  $\mathbf{u}_k = \tilde{\mathbf{u}}_0^*$  aufgeschaltet, dann ist  $(\tilde{\mathbf{u}}_n)$  eine zulässige aber im Allgemeinen nicht optimale Eingangstrajektorie der Optimierungsaufgabe (2.77) für den nachfolgenden Zeithorizont  $k+1, \dots, k+N+1$  mit dem Anfangszustand  $\mathbf{x}_{k+1} = \tilde{\mathbf{x}}_1^*$ . Unter Berücksichtigung von (2.80) gilt folglich für die optimalen Werte der Gütefunktion

$$\begin{aligned} J_{d,k+1|N}^*(\mathbf{x}_{k+1}) &\leq \gamma D(\tilde{\mathbf{x}}_{N-1}) + \sum_{n=0}^{N-2} d_{k+1+n}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{u}}_n) \\ &= J_{d,k|N}^*(\mathbf{x}_k) - d_k(\mathbf{x}_k, \mathbf{u}_k) . \end{aligned} \quad (2.87)$$

Daraus folgt nun noch

$$\Gamma\left(\gamma + N \frac{d}{D}\right) \geq J_{d,k|N}^*(\mathbf{x}_k) \geq J_{d,k+1|N}^*(\mathbf{x}_{k+1}) \quad (2.88)$$

und somit  $\mathbf{x}_{k+1} \in \underline{X}_0$ . Der Rest des Beweises erfolgt völlig analog zum Beweis von Satz 2.4.  $\square$

In diesem Abschnitt wurde eine MPC Variante besprochen, die zwar kein vorgeschriebenes Endgebiet aber einen Endkostenterm benötigt. Es wurde beobachtet, dass zwischen der Gewichtung  $\gamma$  des Endkostenterms, der Horizontlänge  $T$  bzw.  $N$  und der Größe der Menge  $\underline{X}_0$ , die im Einzugsbereich liegt, ein Zusammenhang besteht. Da der Endkostenterm unerwünschte Auswirkungen auf das Verhalten des geschlossenen Regelkreises haben kann, wurden MPC Varianten entwickelt, die ohne einen Endkostenterm auskommen. Nachdem MPC mit unendlich langem Prädiktionshorizont ebenfalls weder einen Endkostenterm noch einen vorgeschriebenes Endgebiet benötigt, ist es nicht verwunderlich, dass eine Mindestlänge für den Prädiktionshorizont existiert, so dass auch bei Verzicht auf einen Endkostenterm und ein vorgeschriebenes Endgebiet exponentielle oder zumindest asymptotische Stabilität nachgewiesen werden kann. Dies wurde für den zeitkontinuierlichen Fall in [2.50] und für den zeitdiskreten Fall in [2.51] gezeigt und wird in dieser Vorlesung nicht weiter besprochen.

## 2.2.5 Endlicher Prädiktionshorizont mit vorgeschriebenem Endgebiet

Es wird ein zu der in Abschnitt 2.2.3 beschriebenen MPC Variante verwandter Ansatz vorgestellt, der ebenfalls ein vorgeschriebenes Endgebiet benötigt aber keinen Endkos-

tenterm  $C$ . Die Länge  $T$  des Prädiktionshorizonts ist nun keine feste Größe mehr und wird durch die Optimierungsvariable  $\tilde{T}$  ersetzt, d. h. es werden Prädiktionshorizonte mit variabler Länge zugelassen. Die Idee für diesen Ansatz entstammt [2.44]<sup>3</sup>. Der Einfachheit halber wird hier nur die zeitkontinuierliche Variante dieses MPC Ansatzes vorgestellt. Die Umsetzung der gleichen Idee in zeitdiskreter Form ist einfach möglich, erfordert aber weitere Überlegungen, denn wegen der zusätzlichen Optimierungsvariable  $\tilde{N} \in \mathbb{N}_{>0}$  für die variable Horizontlänge tritt eine gemischt-ganzzahlige Optimierungsaufgabe auf.

Mit der zusätzlichen Optimierungsvariable  $\tilde{T}$  kann die zeitkontinuierliche Optimierungsaufgabe (2.15) in die Form

$$(\tilde{\mathbf{u}}^*(\cdot), \tilde{T}^*) = \arg \min_{(\tilde{\mathbf{u}}(\cdot), \tilde{T})} J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) = \int_0^{\tilde{T}} c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau \quad (2.89a)$$

$$\text{u.B.v. } \dot{\tilde{\mathbf{x}}}(\tau) = \mathbf{f}(\tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)), \quad \tilde{\mathbf{x}}(0) = \mathbf{x}(t) \quad (2.89b)$$

$$\tilde{\mathbf{x}}(\tau) \in X, \quad \tilde{\mathbf{u}}(\tau) \in U, \quad \forall \tau \in [0, \tilde{T}] \quad (2.89c)$$

$$\tilde{\mathbf{x}}(\tilde{T}) \in X_T \quad (2.89d)$$

$$\tilde{T} \in (0, \infty) \quad (2.89e)$$

umgeschrieben werden. Ferner wird ein Zustandsregelgesetz

$$\mathbf{u}(t) = \boldsymbol{\kappa}(t, \mathbf{x}(t)) \quad (2.90)$$

mit  $\boldsymbol{\kappa} : \mathbb{R} \times X_T \rightarrow U$  benötigt, das folgenden Anforderungen genügt: Das mit dem Zustandsregelgesetz  $\boldsymbol{\kappa}(\tau + t, \bar{\mathbf{x}}(\tau))$  geregelte System (2.1) ist lokal asymptotisch (exponentiell) stabil mit dem Einzugsbereich  $X_T$ . Für die mit  $\bar{\mathbf{x}}(\tau)$  bezeichneten Zustandstrajektorien dieses geregelten Systems gilt für beliebige Anfangszustände  $\bar{\mathbf{x}}(0) = \mathbf{x}(t) \in X_T$ , dass  $\bar{\mathbf{x}}(\tau) \in X \forall \tau \geq 0$ .

Der in [2.44] vorgeschlagene MPC Algorithmus wird als *Zwei-Phasen MPC Regler* (Englisch: *dual-mode MPC*) bezeichnet und ist in Tabelle 2.1 zusammengefasst. Die Schritte 1 und 2 realisieren eine permanent auszuführende Schleife. Schritt 3 wird wiederkehrend auf einem Zeitgitter mit dem Abtastintervall  $T_c$  ausgeführt.

Der Regler gemäß Tabelle 2.1 weist einige Eigenschaften auf, die ihn von den bisher beschriebenen MPC Formulierungen unterscheiden:

- Es ist zu keinem Zeitpunkt die Lösung einer Optimierungsaufgabe nötig. Es werden lediglich Paare  $(\tilde{\mathbf{u}}(\cdot), \tilde{T})$  benötigt, die im Sinne von (2.89) zulässig sind und die Bedingung  $J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) < \infty$  erfüllen. Man spricht daher auch von einer *suboptimalen* MPC Formulierung.
- Das Finden eines Paares  $(\tilde{\mathbf{u}}(\cdot), \tilde{T})$  in Schritt 3 des Algorithmus ist trivial.
- Sobald der Zustand  $\mathbf{x}(t)$  das Gebiet  $X_T$  erreicht, wird auf das Zustandsregelgesetz (2.90) umgeschaltet. Ab diesem Umschaltzeitpunkt erfolgt keine weitere Optimierung der Stellgröße im Sinne des Gütefunktional (2.89a). Dieser eventuell nachteilige Umstand sollte auch bei der Festlegung von  $X_T$  berücksichtigt werden. Wie der nachfolgende Satz zeigt, wird dieser Umschaltzeitpunkt in endlicher Zeit erreicht.

<sup>3</sup>Ein anderer Ansatz, der auch bei Prädiktionshorizonten mit fester Länge keinen Endkostenterm benötigt, findet sich in [2.40].



<b>Initialisierung</b>	Wenn $\mathbf{x}(t) \notin X_T$ , wähle ein Paar $(\tilde{\mathbf{u}}(\cdot), \tilde{T})$ , so dass (2.89b)-(2.89e) und $J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) < \infty$ erfüllt sind, und setze $t_c = t$ .
<b>Schritt 1</b>	Wenn $\mathbf{x}(t) \in X_T$ , gehe zu <b>Ende</b> .
<b>Schritt 2</b>	Wenn $t < t_c + T_c$ , verwende $\mathbf{u}(t) = \tilde{\mathbf{u}}(t - t_c)$ und gehe zu <b>Schritt 1</b> .
<b>Schritt 3</b>	Setze $\tilde{\tilde{T}} = \tilde{T} - T_c$ und $\tilde{\tilde{\mathbf{u}}}(\tau) = \tilde{\mathbf{u}}(\tau + T_c) \forall \tau \in [0, \tilde{\tilde{T}}]$ . Wähle ein Paar $(\tilde{\tilde{\mathbf{u}}}(\cdot), \tilde{\tilde{T}})$ , so dass (2.89b)-(2.89e) und
	$J_{\tilde{\tilde{T}}}(t, \mathbf{x}(t), \tilde{\tilde{\mathbf{u}}}(\cdot)) \leq J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) \quad (2.91)$
	erfüllt sind, setze $t_c = t$ und gehe zu <b>Schritt 1</b> .
<b>Ende</b>	Verwende fortan das asymptotisch (exponentiell) stabilisierende Regelgesetz (2.90).

Tabelle 2.1: Zwei-Phasen MPC Regler gemäß [2.44].

- Auch der Fall  $\tilde{T} < T_c$  ist hier zulässig. Er impliziert, dass innerhalb des aktuellen Abtastintervalls auf das Regelgesetz (2.90) umgeschaltet wird und der Algorithmus terminiert.

**Satz 2.9 (Stabilität des Zwei-Phasen MPC Reglers, zeitkontinuierlich).** *Es seien die Annahmen A1 und A3 erfüllt. Es gelte  $T_c > 0$  und es existiere ein  $\alpha > 0$ , so dass  $\{\mathbf{x} \in X \mid \|\mathbf{x}\|_2^2 \leq \alpha\} \subseteq X_T$ .  $X_0 \subseteq X$  sei eine nichtleere Menge, genau so dass (2.89) für  $\forall \mathbf{x}(t) \in X_0 \setminus X_T$  eine Lösung besitzt (nicht jedoch für  $\forall \mathbf{x}(t) \in \mathbb{R}^l \setminus X_0$ ). Dann ist die Ruhelage  $\mathbf{x}_R = \mathbf{0}$  des mit dem Zwei-Phasen MPC Regler gemäß Tabelle 2.1 geregelten Systems (2.1) lokal asymptotisch (exponentiell) stabil mit dem Einzugsbereich  $X_0$  und die Zustandstrajektorie erreicht die Menge  $X_T$  in endlicher Zeit.*

*Beweis.* Wenn für den Anfangszustand  $\mathbf{x}(t) \in X_T$  gilt, ist nichts zu zeigen. Es müssen daher nur noch die Fälle  $\mathbf{x}(t) \in X_0 \setminus X_T$  untersucht werden. Falls  $\tilde{T} \leq T_c$ , so erreicht die Zustandstrajektorie das Gebiet  $X_T$  innerhalb des aktuellen Abtastintervalls und es ist nichts weiter zu zeigen.

Für die übrigen Fälle  $\tilde{T} > T_c$  wird das System (2.1) im Steuerungshorizont  $[t, t + T_c]$  mit der geplanten Trajektorie  $\tilde{\mathbf{u}}(\cdot)$  betrieben und es gilt im betrachteten nominellen Fall  $\mathbf{x}(t + T_c) = \tilde{\mathbf{x}}(T_c)$ . Ferner gilt wegen  $T_c > 0$ , (2.17b),  $\mathbf{x} \notin X_T \Rightarrow \|\mathbf{x}\|_2^2 > \alpha$  und  $\tilde{\mathbf{x}}(\tau) \notin X_T \forall \tau \in [0, T_c]$ , dass

$$\begin{aligned} J_{\tilde{\tilde{T}}}(t + T_c, \mathbf{x}(t + T_c), \tilde{\tilde{\mathbf{u}}}(\cdot)) &= J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) - \int_0^{T_c} c(t + \tau, \tilde{\mathbf{x}}(\tau), \tilde{\mathbf{u}}(\tau)) d\tau, \\ &\leq J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) - T_c c\alpha, \end{aligned} \quad (2.92)$$

wobei  $\tilde{\tilde{T}}$  und  $\tilde{\tilde{\mathbf{u}}}(\cdot)$  in Schritt 3 des Algorithmus (siehe Tabelle 2.1) definiert sind. Es sei

nun das Paar  $(\tilde{\mathbf{u}}(\cdot), \tilde{T})$  jenes, das zum Zeitpunkt  $t$  in Schritt 3 des Algorithmus gewählt wurde, und  $(\tilde{\mathbf{u}}'(\cdot), \tilde{T}')$  jenes, das zum Zeitpunkt  $t + T_c$  gewählt wurde. In Schritt 3 des Algorithmus kann (2.91) trivial durch die mögliche Wahl  $(\tilde{\mathbf{u}}(\cdot), \tilde{T}) = (\tilde{\tilde{\mathbf{u}}}(\cdot), \tilde{\tilde{T}})$  erfüllt werden. Aus (2.91) und (2.92) folgt

$$J_{\tilde{T}'}(t + T_c, \mathbf{x}(t + T_c), \tilde{\mathbf{u}}'(\cdot)) \leq J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) - T_c c \alpha . \quad (2.93)$$

Es gilt

$$J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) \leq T_c c \alpha \quad \Rightarrow \quad \exists \tau \in [0, T_c]: \tilde{\mathbf{x}}(\tau) \in X_T . \quad (2.94)$$

Aus einem Anfangswert  $J_{\tilde{T}}(t, \mathbf{x}(t), \tilde{\mathbf{u}}(\cdot)) < \infty$  des Gütefunktional am Beginn der Regelung (vgl. Tabelle 2.1), der Mindestreduktion  $T_c c \alpha$  des Gütefunktional je Abtastschritt gemäß (2.93) und der Abbruchbedingung (2.94) kann sofort eine obere Schranke für jene Zeitspanne berechnet werden, die vergeht bis die Zustandstrajektorie das Gebiet  $X_T$  erreicht.  $\square$

## 2.3 Implementierung

Es werden kurz zwei Aspekte der Implementierung betrachtet.

### 2.3.1 Entwurf eines stabilisierenden Zustandsreglers für ein Endgebiet

In den Abschnitten 2.2.3-2.2.5 wurde für ein vorgeschriebenes oder automatisch erreichtes Endgebiet die Existenz eines stabilisierenden Zustandsreglers vorausgesetzt. Zusätzlich wurde gefordert, dass der mit diesem Regler geschlossene Regelkreis zu einer gewissen Mindestreduktion des Endkostenterms  $C$  bzw.  $D$  führt (siehe z. B. die Bedingungen (2.55) und (2.61)). Zum Entwurf eines solchen Zustandsregelgesetzes können zahlreiche der in den Vorlesungen *Regelungssysteme 1* [2.2] und *Regelungssysteme 2* [2.45] besprochenen Methoden verwendet werden. Im Falle einer stabilisierbaren, linearen, zeitinvarianten Strecke genügt wegen (2.17) bzw. (2.18) meist schon der Entwurf eines LQR-Reglers (siehe [2.2]). Für den Fall einer nichtlinearen Strecke (2.1), deren Linearisierung an der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ ,  $\mathbf{u}_R = \mathbf{0}$  stabilisierbar ist, wurde in [2.48] eine Methode zur Konstruktion eines linearen Zustandsreglers vorgeschlagen, der bei quadratischem Gütefunktional die geforderte Mindestreduktion des Endkostenterms lokal sicherstellt. Diese Methode wird hier kurz vorgestellt. Es sei

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{u} \quad (2.95)$$

die Linearisierung von (2.1) an der Ruhelage  $\mathbf{x}_R = \mathbf{0}$ ,  $\mathbf{u}_R = \mathbf{0}$ . Das System (2.95) sei stabilisierbar. Die Funktion  $\lambda_{\max}(\cdot)$  liefere den größten Realteil der Eigenwerte einer Matrix.

**Lemma 2.1** (Konstruktion eines linearen Regelgesetzes für das Endgebiet). *Es sei*

$$\mathbf{u} = \mathbf{K}\boldsymbol{\xi} \quad (2.96)$$

ein Zustandsregelgesetz für das System (2.95), so dass  $\mathbf{A} + \mathbf{BK}$  eine Hurwitzmatrix ist. Mit einer Konstante  $\kappa \in [0, -\lambda_{\max}(\mathbf{A} + \mathbf{BK})]$  und den symmetrisch, positiv definiten Matrizen  $\mathbf{Q} \in \mathbb{R}^{l \times l}$  und  $\mathbf{R} \in \mathbb{R}^{m \times m}$  hat die Lyapunov-Gleichung

$$(\mathbf{A} + \mathbf{BK} + \kappa \mathbf{E})^T \mathbf{P} + \mathbf{P}(\mathbf{A} + \mathbf{BK} + \kappa \mathbf{E}) + \mathbf{Q} + \mathbf{K}^T \mathbf{R} \mathbf{K} = \mathbf{0} \quad (2.97)$$

daher eine eindeutige und symmetrisch, positiv definite Lösung  $\mathbf{P}$ . Ferner existiert eine Konstante  $\Gamma \in \mathbb{R}_{>0}$ , so dass auf dem Gebiet

$$X_\Gamma = \{\mathbf{x} \in X \mid \mathbf{x}^T \mathbf{P} \mathbf{x} \leq \Gamma\} \supset \{\mathbf{0}\} \quad (2.98)$$

Folgendes gilt:

- Der Regler hält die Stellgrößenbeschränkungen ein, d. h.  $\mathbf{K} \boldsymbol{\xi} \in U \quad \forall \boldsymbol{\xi} \in X_\Gamma$ .
- $X_\Gamma$  ist eine positiv invariante Menge des mit  $\mathbf{u} = \mathbf{K} \mathbf{x}$  geregelten nichtlinearen Systems (2.1).
- Die Trajektorien  $\bar{\mathbf{x}}(\tau)$  und  $\bar{\mathbf{u}}(\tau)$  mit  $\tau \geq 0$  des gemäß  $\bar{\mathbf{u}}(\tau) = \mathbf{K} \bar{\mathbf{x}}(\tau)$  geregelten nichtlinearen Systems (2.1) erfüllen für jeden Anfangszustand  $\bar{\mathbf{x}}(0) = \mathbf{x}(t) \in X_\Gamma$

$$\mathbf{x}^T(t) \mathbf{P} \mathbf{x}(t) \geq \int_0^\infty \bar{\mathbf{x}}^T(\tau) \mathbf{Q} \bar{\mathbf{x}}(\tau) + \bar{\mathbf{u}}^T(\tau) \mathbf{R} \bar{\mathbf{u}}(\tau) d\tau \quad (2.99)$$

und weiters

$$\begin{aligned} \frac{d}{dt} \bar{\mathbf{x}}^T(t) \mathbf{P} \bar{\mathbf{x}}(t) &\leq \bar{\mathbf{x}}^T(t) ((\mathbf{A} + \mathbf{BK} + \kappa \mathbf{E})^T \mathbf{P} + \mathbf{P}(\mathbf{A} + \mathbf{BK} + \kappa \mathbf{E})) \bar{\mathbf{x}}(t) \\ &= -\bar{\mathbf{x}}^T(t) (\mathbf{Q} + \mathbf{K}^T \mathbf{R} \mathbf{K}) \bar{\mathbf{x}}(t) . \end{aligned} \quad (2.100)$$

Der Nachweis dieses Lemmas ist in [2.48] zu finden. Aus dem Lemma folgt direkt, dass im Falle  $C(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$  und  $c(t, \mathbf{x}, \mathbf{u}) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}$  das Regelgesetz (2.96) die Anforderungen der Sätze 2.5, 2.7 und 2.9 erfüllt. Der zeitdiskrete Fall kann analog zu Lemma 2.1 behandelt werden.

### 2.3.2 Methoden zur Lösung von Optimalsteuerungsaufgaben

Die Lösung einer dynamischen Optimierungsaufgabe, z. B. (2.15) oder (2.16), stellt meist eine zentrale Herausforderung bei der echtzeitfähigen Implementierung eines MPC Reglers dar. Ob eine solche Lösung in der zur Verfügung stehenden Zeit gelingt und mit welcher Methode dies erfolgen soll, hängt auch wesentlich von der Problemformulierung (Wahl des Gütefunktionalen oder der Gütefunktion, Festlegung von Beschränkungen, etc.) ab. Es werden kurz einige gängige Lösungsstrategien besprochen.

*Zeitkontinuierliche* Formulierungen müssen im Regelfall in eine zeitdiskrete Formulierung transformiert werden, um auf einem Rechner implementiert zu werden. Hierbei wird zwischen *direkten* und *indirekten* Verfahren unterschieden:

- Wie in der Vorlesung *Optimierung* [2.27] ausführlich besprochen, wird die dynamische Optimierungsaufgabe bei *indirekten Verfahren* zunächst mittels Variationsrechnung

oder dem Minimumsprinzip von Pontryagin in zeitkontinuierliche Optimalitätsbedingungen in Form eines Randwertproblems umgeschrieben. Das erhaltene Randwertproblem kann meist mit gängigen numerischen Methoden [2.52–2.54], wie z. B. Einfach-Schießverfahren, Mehrfach-Schießverfahren und Kollokationsverfahren, gelöst werden. Gelegentlich ist eine Lösung nicht möglich, z. B. wenn adjungierte Variablen zufolge von Zustandsbeschränkungen unstetig sind.

- Bei *direkten Verfahren* wird die zeitkontinuierliche Optimierungsaufgabe mittels numerischer Integrationsverfahren [2.52–2.54] direkt diskretisiert. Dies betrifft vor allem integrale Kostenanteile (vgl. (2.15a)) im Gütefunktional und Nebenbedingungen in Form von Differenzialgleichungen (vgl. (2.15b)). Das Resultat lässt sich im Allgemeinen in der Form (2.16) darstellen. Werden implizite Zeitintegrationsverfahren verwendet, so ist (2.16b) durch eine implizite Differenzgleichung zu ersetzen. Die gewählte Diskretisierungsschrittweite sowie die Art der Parametrierung der Eingangsgröße (vgl. (2.4)) haben einen direkten Einfluss auf die erzielte Genauigkeit und auf die Dimension der resultierenden zeitdiskreten Optimierungsaufgabe. In [2.13] wird studiert welche Auswirkungen Diskretisierungsfehler auf die Stabilität und Konvergenz von MPC Reglern haben können.

Zeitdiskrete Formulierungen können (außer bei unendlich langem Prädiktionshorizont) als finit-dimensionale, statische Optimierungsaufgaben aufgefasst werden. D. h. sie können in der Form

$$\min_{\mathbf{z} \in \mathbb{R}^o} J(\mathbf{z}) \quad (2.101a)$$

$$\text{u.B.v. } \mathbf{g}(\mathbf{z}) = \mathbf{0} \quad (2.101b)$$

$$\mathbf{h}(\mathbf{z}) \leq \mathbf{0} \quad (2.101c)$$

angeschrieben werden. Hierbei hat der Suchraum die Dimension  $o$ ,  $J : \mathbb{R}^o \rightarrow \mathbb{R}_{\geq 0}$  ist eine entsprechende Kostenfunktion und alle Beschränkungen sind in  $\mathbf{g} : \mathbb{R}^o \rightarrow \mathbb{R}^p$  und  $\mathbf{h} : \mathbb{R}^o \rightarrow \mathbb{R}^q$  zusammengefasst.

Zur Lösung der Standardaufgabe (2.101) stehen zahlreiche numerische Algorithmen zur Verfügung (siehe [2.28–2.39]). Viele wurden auch bereits in der Vorlesung *Optimierung* [2.27] besprochen, so dass hier auf eine weitere Vertiefung verzichtet werden kann. Es verbleibt die Frage, wie die zeitdiskrete Optimierungsaufgabe (2.16) bzw. ihre Abwandlungen (2.33), (2.48) und (2.77) in die Form (2.101) umgeschrieben werden. Nachfolgende aus [2.13] entnommene Varianten zeigen, dass die Antwort auf diese Frage keineswegs eindeutig ist.

- **Volldiskretisierung:** In der Optimierungsaufgabe (2.16) sind die Größen  $\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_N$  und  $\tilde{\mathbf{u}}_0, \tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_{N-1}$  unbekannt. Es ist daher naheliegend, sie alle im Vektor der Optimierungsvariablen

$$\mathbf{z} = \left[ \tilde{\mathbf{x}}_0^T \quad \tilde{\mathbf{x}}_1^T \quad \dots \quad \tilde{\mathbf{x}}_N^T \quad \tilde{\mathbf{u}}_0^T \quad \tilde{\mathbf{u}}_1^T \quad \dots \quad \tilde{\mathbf{u}}_{N-1}^T \right]^T \quad (2.102)$$

zusammenzufassen. Damit hat die Optimierungsaufgabe die Dimension  $o = (N + 1)l + NM$ . Die Funktionen  $J$ ,  $\mathbf{g}$  und  $\mathbf{h}$  folgen direkt aus (2.16). Diese Variante hat den

Nachteil, dass sie im Allgemeinen zu einer hochdimensionalen Optimierungsaufgabe führt. Sie hat den Vorteil, dass Ableitungen von  $J$ ,  $\mathbf{g}$  und  $\mathbf{h}$  bezüglich  $\mathbf{z}$ , wie sie oft von numerischen Lösungsverfahren benötigt werden, sehr einfach analytisch berechnet werden können. Man beachte, dass die unbekannt Zustände  $\tilde{\mathbf{x}}_n$  und die gesuchten Eingangsparemeter  $\tilde{\mathbf{u}}_n$  grundsätzlich mit der gleichen Genauigkeit bestimmt werden, was insbesondere dann von Bedeutung ist, wenn ein iteratives Lösungsverfahren in einer MPC Implementierung frühzeitig abgebrochen wird.

- **Unterlagerte Zeitintegration:** Diese Variante wird auch Teildiskretisierung oder reduzierte Diskretisierung genannt [2.55]. Entsprechend (2.16b) lassen sich die Zustandsgrößen  $\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_N$  (zumindest formal) als Funktionen der Eingangsparemeter  $\tilde{\mathbf{u}}_0, \tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_{N-1}$  und des bekannten Anfangszustands  $\mathbf{x}_k$  ausdrücken. D. h. es existiert eine eindeutige Abbildung

$$\begin{bmatrix} \mathbf{x}_k^T & \tilde{\mathbf{u}}_0^T & \tilde{\mathbf{u}}_1^T & \dots & \tilde{\mathbf{u}}_{N-1}^T \end{bmatrix}^T \mapsto \begin{bmatrix} \tilde{\mathbf{x}}_0^T & \tilde{\mathbf{x}}_1^T & \dots & \tilde{\mathbf{x}}_N^T \end{bmatrix}^T, \quad (2.103)$$

mit der alle Zustandsvektoren  $\tilde{\mathbf{x}}_n$  in (2.16) eliminiert werden können. Wird die verbleibende Optimierungsaufgabe in die Form (2.101) umgeschrieben, so lauten die Optimierungsvariablen

$$\mathbf{z} = \begin{bmatrix} \tilde{\mathbf{u}}_0^T & \tilde{\mathbf{u}}_1^T & \dots & \tilde{\mathbf{u}}_{N-1}^T \end{bmatrix}^T. \quad (2.104)$$

Damit hat die Optimierungsaufgabe die Dimension  $o = NM$ . Die Funktionen  $J$ ,  $\mathbf{g}$  und  $\mathbf{h}$  folgen wieder direkt aus (2.16), wobei die in (2.103) verwendeten Gleichungsbeschränkungen (2.16b) nun nicht mehr in  $\mathbf{g}$  aufzunehmen sind. Diese Variante hat den Vorteil, dass sie im Allgemeinen zu einer niedrigdimensionalen Optimierungsaufgabe führt. Sie hat aber den Nachteil, dass totale Ableitungen von  $J$ ,  $\mathbf{g}$  und  $\mathbf{h}$  bezüglich  $\mathbf{z}$  grundsätzlich auch das meist numerisch sensitive Nachdifferenzieren der Abbildung (2.103) erfordern. Bei der Implementierung von MPC Algorithmen wird (2.103) oft als unterlagerte Zeitintegrationsroutine realisiert. Daher ist es auch möglich, die Genauigkeit der Zeitintegration des dynamischen Systems unabhängig von einer Genauigkeitsanforderung oder einem allfälligen frühzeitigen Abbruch des iterativen Optimierungsverfahrens (suboptimale MPC) vorzugeben.

- **Mehrfach-Schießverfahren:** Im Mehrfach-Schießverfahren werden die Grundideen der *Volldiskretisierung* und der *unterlagerten Zeitintegration* vereint. Es werden daher nicht alle sondern nur einige der unbekannt Zustandsvektoren  $\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_N$  zu den Optimierungsvariablen  $\mathbf{z}$  hinzugenommen. Werden die Zustandsvektoren  $\tilde{\mathbf{x}}_{n_i}$  mit  $n_i \in \{0, 1, \dots, N\}$  und  $0 \leq n_1 < n_2 < \dots < n_I \leq N$  als Optimierungsvariablen in den Vektor

$$\mathbf{z} = \begin{bmatrix} \tilde{\mathbf{x}}_{n_1}^T & \tilde{\mathbf{x}}_{n_2}^T & \dots & \tilde{\mathbf{x}}_{n_I}^T & \tilde{\mathbf{u}}_0^T & \tilde{\mathbf{u}}_1^T & \dots & \tilde{\mathbf{u}}_{N-1}^T \end{bmatrix}^T \quad (2.105)$$

aufgenommen, so hat dieser die Dimension  $o = Il + NM$  und von den Gleichungsbeschränkungen (2.16b) sind nur jene in  $\mathbf{g}$  aufzunehmen, die auf ihrer linken Seite ein unbekanntes  $\tilde{\mathbf{x}}_{n_i}$  stehen haben.

## 2.4 Literatur

- [2.1] F. Allgöwer, T. Badgwell, J. Qin, J. Rawlings und S. Wright, „Nonlinear Predictive Control and Moving Horizon Estimation: An Introductory Overview,“ in *Advances in Control*, Springer London, 1999, S. 391–449.
- [2.2] W. Kemmetmüller, *Skriptum zur VO Regelungssysteme 1 (WS 2018/2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2018. Adresse: <https://www.acin.tuwien.ac.at/master/regelungssysteme-1/>.
- [2.3] M. Alamir, *A Pragmatic Story of Model Predictive Control: Self-Contained Algorithms and Case-Studies*. CreateSpace Independent Publishing Platform, 2013.
- [2.4] F. Borrelli, A. Bemporad und M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 2015, in Druck.
- [2.5] E.F. Camacho und C. Bordons, *Model Predictive Control*, 2. Aufl., Ser. Advanced Textbooks in Control and Signal Processing. Springer-Verlag, 2004.
- [2.6] R. Dittmar und B. Pfeifer, *Modellbasierte prädiktive Regelung: Eine Einführung für Ingenieure*. München: Oldenbourg, 2004.
- [2.7] W. Kwon und S. Han, *Receding Horizon Control*, Ser. Advanced Textbooks in Control and Signal Processing. Springer, 2005.
- [2.8] J.M. Maciejowski, *Predictive Control with Constraints*. Prentice Hall, 2002.
- [2.9] J. Maestre und R. Negenborn, Hrsg., *Distributed Model Predictive Control Made Easy*, Ser. Intelligent Systems, Control and Automation: Science and Engineering. Springer, 2014, Bd. 69.
- [2.10] J.B. Rawlings und D.Q. Mayne, *Model Predictive Control: Theory and Design*. Madison, Wisconsin: Nob Hill Publishing, 2009.
- [2.11] J.A. Rossiter, *Model-Based Predictive Control: A Practical Approach*. Boca Raton, Florida: CRC Press, 2003.
- [2.12] M. Alamir, *Stabilization of Nonlinear Systems Using Receding-horizon Control Schemes: A Parametrized Approach for Fast Systems*, Ser. Lecture Notes in Control and Information Sciences. Springer, 2006, Bd. 339.
- [2.13] L. Grüne und J. Pannek, *Nonlinear Model Predictive Control*. London: Springer, 2011.
- [2.14] F. Allgöwer und A. Zheng, Hrsg., *Nonlinear Model Predictive Control*, Bd. 26, Progress in Systems and Control Theory, Basel: Birkhäuser, 2000.
- [2.15] R. Findeisen, F. Allgöwer und L.B. Biegler, Hrsg., *Assessment and Future Directions of Nonlinear Model Predictive Control*, Bd. 358, Lecture Notes in Control and Information Sciences, Berlin: Springer, 2007.
- [2.16] B. Kouvaritakis und M. Cannon, Hrsg., *Nonlinear Predictive Control: Theory and Practice*, Bd. 61, IEE Control Engineering Series, Herts, UK: Institution of Electrical Engineers, 2001.

- [2.17] L. Magni, D.M. Raimondo und F. Allgöwer, Hrsg., *Nonlinear Model Predictive Control*, Bd. 384, Lecture Notes in Control and Information Science, Berlin, Heidelberg: Springer, 2009.
- [2.18] C.E. Garcia, D.M. Prett und M. Morari, „Model Predictive Control: Theory and Practice - a Survey“, *Automatica*, Jg. 25, Nr. 3, S. 335–348, 1989.
- [2.19] D.Q. Mayne, J.B. Rawlings, C.V. Rao und P.O.M. Scokaert, „Constrained model predictive control: Stability and optimality“, *Automatica*, Jg. 36, Nr. 6, S. 789–814, 2000.
- [2.20] S. Qin und T. Badgwell, „A survey of industrial model predictive control technology“, *Control Engineering Practice*, Jg. 11, Nr. 7, S. 733–764, 2003.
- [2.21] D. Mayne, „Model predictive control: Recent developments and future promise“, *Automatica*, Jg. 50, Nr. 12, S. 2967–2986, 2014.
- [2.22] J. Richalet, A. Rault, J. Testud und J. Papon, „Model predictive heuristic control: Applications to industrial processes“, *Automatica*, Jg. 14, Nr. 5, S. 413–428, 1978.
- [2.23] D. Clarke, C. Mohtadi und P. Tuffs, „Generalized predictive control: Part I. The basic algorithm“, *Automatica*, Jg. 23, Nr. 2, S. 137–148, 1987.
- [2.24] L. Ljung, *System Identification: Theory for the User*, 2. Aufl. Upper Saddle River, New Jersey: Prentice Hall, 1999.
- [2.25] P. Scokaert und D. Mayne, „Min-max feedback model predictive control for constrained linear systems“, *IEEE Transactions on Automatic Control*, Jg. 43, Nr. 8, S. 1136–1142, Aug. 1998.
- [2.26] K. Tae-Hyoung, H. Fukushima und T. Sugie, „Robust adaptive model predictive control based on comparison model“, in *Proceedings of the 43<sup>rd</sup> IEEE Conference on Decision and Control (CDC)*, Bd. 2, Dez. 2004, S. 2041–2046.
- [2.27] A. Steinböck, *Skriptum zur VU Optimierung (WS 2018/2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2018. Adresse: <https://www.acin.tuwien.ac.at/master/optimierung/>.
- [2.28] D.P. Bertsekas, *Nonlinear Programming*, 2. Aufl. Belmont, Massachusetts: Athena Scientific, 1999.
- [2.29] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*, Ser. Advances in Design and Control. Philadelphia, USA: SIAM - Society for Industrial und Applied Mathematics, 2001.
- [2.30] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal und C.A. Sagastizábal, *Numerical Optimization - Theoretical and Practical Aspects*, 2. Aufl. Berlin, Heidelberg: Springer, 2006.
- [2.31] S. Boyd und L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [2.32] A.E. Bryson und Y.-C. Ho, *Applied Optimal Control*. New York: John Wiley & Sons, 1975.

- [2.33] P.E. Gill, W. Murray und M.H. Wright, *Practical Optimization*. London: Academic Press, 1981.
- [2.34] C.T. Kelley, *Iterative Methods for Optimization*, Ser. Frontiers in Applied Mathematics 18. Philadelphia: Society for Industrial und Applied Mathematics, 1999.
- [2.35] D.G. Luenberger und Y. Ye, *Linear and Nonlinear Programming*, 3. Aufl., Ser. International Series in Operations Research and Management Science. New York: Springer, 2008.
- [2.36] J. Nocedal und S.J. Wright, *Numerical Optimization*, 2. Aufl., Ser. Springer Series in Operations Research. New York: Springer, 2006.
- [2.37] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Ser. Applied Mathematical Sciences 124. New York: Springer, 1997.
- [2.38] J.L. Speyer und D.H. Jacobson, *Primer on Optimal Control Theory*, Ser. Advances in Design and Control. Philadelphia: Siam, 2010.
- [2.39] K.L. Teo, C.J. Goh und K.H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Ser. Pitman Monographs and Surveys in Pure and Applied Mathematics. New York: John Wiley & Sons, 1991.
- [2.40] P. Scokaert, D. Mayne und J. Rawlings, „Suboptimal model predictive control (feasibility implies stability),“ *IEEE Transactions on Automatic Control*, Jg. 44, Nr. 3, S. 648–654, März 1999.
- [2.41] K. Graichen und A. Kugi, „Stability of incremental model predictive control without terminal constraints,“ *IEEE Transactions on Automatic Control*, Jg. 55, Nr. 11, S. 2576–2580, 2010.
- [2.42] M. Diehl, R. Findeisen, F. Allgöwer, H.G. Bock und J.P. Schlöder, „Nominal stability of real-time iteration scheme for nonlinear model predictive control,“ *IEE Proceedings Control Theory and Applications*, Jg. 152, Nr. 3, S. 296–308, Mai 2005.
- [2.43] D. DeHaan und M. Guay, „A real-time framework for model-predictive control of continuous-time nonlinear systems,“ *IEEE Transactions on Automatic Control*, Jg. 52, Nr. 11, S. 2047–2057, Nov. 2007.
- [2.44] H. Michalska und D.Q. Mayne, „Robust receding horizon control of constrained nonlinear systems,“ *IEEE Transactions on Automatic Control*, Jg. 38, Nr. 11, S. 1623–1633, 1993.
- [2.45] A. Kugi, *Skriptum zur VO Regelungssysteme 2 (SS 2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2019. Adresse: <https://www.acin.tuwien.ac.at/master/regelungssysteme-2/>.
- [2.46] K. Graichen, *Skriptum zur VO Methoden der Optimierung und optimalen Steuerung (WS 2014/2015)*, Institut für Mess-, Regel- und Mikrotechnik, Universität Ulm, 2014.
- [2.47] M. Vidyasagar, *Nonlinear Systems Analysis*, 2. Aufl., Ser. Classics in Applied Mathematics 42. Philadelphia: SIAM, 1992.



- [2.48] H. Chen und F. Allgöwer, „A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability,“ *Automatica*, Jg. 34, Nr. 10, S. 1205–1217, 1998.
- [2.49] D. Limon, T. Alamo, F. Salas und E. Camacho, „On the stability of constrained MPC without terminal constraint,“ *IEEE Transactions on Automatic Control*, Jg. 51, Nr. 5, S. 832–836, Mai 2006.
- [2.50] A. Jadbabaie und J. Hauser, „On the Stability of Receding Horizon Control With a General Terminal Cost,“ *IEEE Transactions on Automatic Control*, Jg. 50, Nr. 5, S. 674–678, 2005.
- [2.51] A. Boccia, L. Grüne und K. Worthmann, „Stability and feasibility of state constrained MPC without stabilizing terminal constraints,“ *Systems & Control Letters*, Jg. 72, Nr. 0, S. 14–21, 2014.
- [2.52] H.R. Schwarz und N. Köckler, *Numerische Mathematik*, 6. Aufl. Wiesbaden: B.G. Teubner, 2006.
- [2.53] M. Hermann, *Numerik gewöhnlicher Differentialgleichungen: Anfangs- und Randwertprobleme*. München: Oldenbourg, 2004.
- [2.54] J. Stoer und R. Bulirsch, *Introduction to Numerical Analysis*, 3. Aufl., Ser. Texts in Applied Mathematics 12. New York, Berlin: Springer, 2002.
- [2.55] M. Gerds, *Optimal Control of ODEs and DAEs*. Berlin, Boston: De Gruyter, 2012.



## 3 Zustandsschätzung auf bewegten Horizonten

In diesem Abschnitt wird die Methode der Zustandsschätzung auf bewegten Horizonten kurz vorgestellt. Es werden dazu die Bestandteile dieser Schätzmethode und Möglichkeiten zur Berücksichtigung von Informationen, die vor dem aktuellen Horizont gesammelt wurden, diskutiert. Abschließend wird eine wahrscheinlichkeitstheoretische Interpretation der Methode gegeben und skizziert, wie neben Zuständen auch Systemparameter geschätzt werden können.

Der Begriff Schätzung auf bewegtem Horizont wird im Englischen oft als *moving horizon estimation (MHE)* oder *receding horizon estimation* bezeichnet. Hier wird der allgemeine Fall der *nichtlinearen modellbasierten* Zustandsschätzung auf bewegten Horizonten behandelt. Mit MHE werden die folgenden Vorgehensweisen in Verbindung gebracht:

- Für einen bestimmten Zeithorizont werden der Anfangszustand eines dynamischen Systems und die auf das System wirkenden Störungen geschätzt.
- Für die Schätzung werden Messwerte des Systemausgangs, ein mathematisches Modell zur Beschreibung des Systemverhaltens und meist Informationen aus früheren Schätzungen verwendet. Die Schätzung beruht auf der Minimierung der Diskrepanz zwischen den Messwerten des Systemausgangs und den mit dem Modell berechneten Ausgangswerten.
- Im Allgemeinen erfordert diese Schätzung die Lösung einer dynamischen Optimierungsaufgabe.
- Die Schätzung wird zu diskreten Zeitpunkten wiederkehrend durchgeführt.

Die modellbasierte Berechnung des Systemverhaltens und die Formulierung der zu lösenden Optimierungsaufgabe erfolgt im Allgemeinen für einen Zeithorizont, der in der Vergangenheit beginnt und zum aktuellen Zeitpunkt endet. Für die wiederkehrende Schätzung muss dieser Horizont zeitlich fortgeschoben werden.

MHE eignet sich gut zur Zustandsschätzung bei Systemen, die

- Beschränkungen unterworfen sind,
- nichtlinear sind,
- nur selten oder zu unregelmäßigen Zeitpunkten Messungen zulassen und
- Störungen oder Rauschen mit nicht näher spezifizierten stochastischen Eigenschaften ausgesetzt sind.

Als potentielle mit MHE verbundene Schwierigkeiten sind zu nennen:

- Die Methode erfordert mitunter einen hohen *Rechenaufwand*, da eine dynamische Optimierungsaufgabe gelöst werden muss.
- Die erstmalige Entwicklung und Implementierung des Schätzverfahrens samt einem Lösungsalgorithmus für die zugehörige dynamische Optimierungsaufgabe kann *aufwendig* sein.

Einen Überblick über die Methode und das Forschungsgebiet MHE bieten die Arbeiten [3.1–3.4]. Darüber hinaus bieten die Beiträge [3.5–3.8] gute Einstiegspunkte in das Thema. In [3.7, 3.9] wird MHE mit dem Extended Kalman-Filter [3.10] verglichen. Das vorliegende Skriptum orientiert sich an [3.2, 3.11].

## 3.1 Bestandteile von MHE

### 3.1.1 Modell

Zur Beschreibung des Systemverhaltens dient grundsätzlich ein mathematisches Modell, dessen Eingänge zum Teil Zufallsvariablen sein können. In diesem Abschnitt wird das folgende zeitvariante zeitdiskrete dynamische Modell in Zustandsraumdarstellung verwendet:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k) \quad \forall k \in \mathbb{N}_0 \quad (3.1a)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \quad \forall k \in \mathbb{N}_0 . \quad (3.1b)$$

Hierbei ist  $k \in \mathbb{N}_0$  der Zeitindex,  $t_k \in \mathbb{R}_{\geq 0}$  die zugehörige Zeit,  $\mathbf{x}_k \in \mathbb{R}^n$  der exakte Systemzustand,  $\mathbf{x}_0$  der exakte Anfangszustand,  $\mathbf{w}_k \in \mathbb{R}^p$  eine Prozessstörung (Prozessrauschen),  $\mathbf{v}_k \in \mathbb{R}^q$  eine Messstörung (Messrauschen) und  $\mathbf{y}_k \in \mathbb{R}^q$  der Messwert des Systemausgangs. Allfällige zusätzliche Systemeingänge, z. B. Stelleingänge, seien bekannt und bereits in  $\mathbf{f}_k$  enthalten. Das Modell (3.1) kann aus der Abtastung eines zeitkontinuierlichen dynamischen Systems hervorgehen, wobei die Abtastzeit variabel sein kann.

Im Allgemeinen entstammen die Störungen  $\mathbf{w}_k$  und  $\mathbf{v}_k$  stationären stochastischen Prozessen [3.10]. Der Einfachheit halber werden sie hier als unbekannte, beschränkte, mittelwertfreie Zufallsvariablen aufgefasst. Auch der Anfangszustand  $\mathbf{x}_0$  sei eine unbekannte, beschränkte Zufallsvariable.

Es ist nun zu unterscheiden zwischen den Systemvariablen  $(\mathbf{x}_k, \mathbf{w}_k, \mathbf{v}_k, \mathbf{y}_k)$ , den korrespondierenden Größen  $(\tilde{\mathbf{x}}_k, \tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k, \mathbf{y}_k)$  in einer Schätzaufgabe und deren (optimalen) Schätzwerten  $(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k, \hat{\mathbf{v}}_k, \mathbf{y}_k)$ , die als Lösung der Schätzaufgabe von einem Beobachter bestimmt werden. Diese Variablen und ihre Zusammenhänge sind in Tabelle 3.1 zusammengefasst. Es ist zu beachten, dass die gemessene Ausgangsgröße  $\mathbf{y}_k$  im realen System und in der Schätzaufgabe identisch ist.

### 3.1.2 Horizont

In den Abschnitten 3.3 und 3.4 wird erläutert, warum zur Zustandsschätzung im Allgemeinen nicht die gesamte Information, die in allen in der Vergangenheit aufgetretenen

	Systemvariablen	Variablen im Schätzproblem	Optimale Schätzwerte
Zustand	$\mathbf{x}_k$	$\tilde{\mathbf{x}}_k$	$\hat{\mathbf{x}}_k$
Prozessstörung	$\mathbf{w}_k$	$\tilde{\mathbf{w}}_k$	$\hat{\mathbf{w}}_k$
Systemdynamik	$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k)$	$\tilde{\mathbf{x}}_{k+1} = \mathbf{f}_k(\tilde{\mathbf{x}}_k, \tilde{\mathbf{w}}_k)$	$\hat{\mathbf{x}}_{k+1} = \mathbf{f}_k(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k)$
Messstörung	$\mathbf{v}_k$	$\tilde{\mathbf{v}}_k$	$\hat{\mathbf{v}}_k$
Nominelle Ausgangsgröße	$\mathbf{h}_k(\mathbf{x}_k)$	$\mathbf{h}_k(\tilde{\mathbf{x}}_k)$	$\mathbf{h}_k(\hat{\mathbf{x}}_k)$
Gemessene Ausgangsgröße	$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k$	$\mathbf{y}_k = \mathbf{h}_k(\tilde{\mathbf{x}}_k) + \tilde{\mathbf{v}}_k$	$\mathbf{y}_k = \mathbf{h}_k(\hat{\mathbf{x}}_k) + \hat{\mathbf{v}}_k$

Tabelle 3.1: Variablen des Systems und des Beobachters.

Messwerten  $\mathbf{y}_k$  enthalten ist, exakt ausgenutzt werden kann. Aus diesem Grund werden bei MHE nur Messwerte aus einem Zeithorizont verwendet, der in der Vergangenheit beginnt, zum aktuellen Zeitpunkt endet und eine finite Länge besitzt. Informationen aus Messwerten, die vor dem aktuellen Horizont generiert wurden, können aber näherungsweise in der aktuellen Schätzung berücksichtigt werden. Um neue hinzukommende Messwerte zu berücksichtigen, wird der Horizont zeitlich fortbewegt und die Schätzung wiederholt. D. h. mit jedem Abtastschritt wird ein neuer Messwert zur Schätzung hinzugenommen und der älteste bisher berücksichtigte Messwert verworfen.

Es soll nun  $N \geq 1$  die Länge des Horizonts und  $K$  der aktuelle Zeitindex sein, d. h.  $t_K$  ist der aktuelle Zeitpunkt. Es werden die Messwerte  $\mathbf{y}_{K-N}, \mathbf{y}_{K-N+1}, \dots, \mathbf{y}_{K-1}$  zur Schätzung herangezogen. Sie werden in der Folge  $\mathbf{y}_{K-N|K-1} = (\mathbf{y}_{K-N}, \dots, \mathbf{y}_{K-1})$  zusammengefasst. Der aktuelle Messwert  $\mathbf{y}_K$  soll keine Berücksichtigung mehr finden, damit das Zeitintervall  $(t_{K-1}, t_K)$  zur Berechnung der Lösung der Schätzaufgabe zur Verfügung steht. Wird unter Verwendung der Messwerte  $\mathbf{y}_{K-N|K-1}$  der Zustand  $\mathbf{x}_k$  geschätzt, so handelt es sich im Fall  $k \geq K$  strenggenommen um eine *Prädiktion*, im Fall  $k = K - 1$  um eine *Filterung* und im Fall  $k \in \{K - N, K - N + 1, \dots, K - 2\}$  um eine *Glättung*. Die hier vorgestellte MHE Variante vereint grundsätzlich diese Vorgehensweisen:

Zum Zeitpunkt  $K$  werden die Zustandsgröße  $\mathbf{x}_{K-N}$  und die zu einer Folge zusammengefassten Prozessstörungen  $\mathbf{w}_{K-N|K-1} = (\mathbf{w}_{K-N}, \dots, \mathbf{w}_{K-1})$  geschätzt. Daraus können mithilfe von (3.1a) sofort Schätzwerte für die Zustände  $\mathbf{x}_{K-N+1}, \mathbf{x}_{K-N+2}, \dots, \mathbf{x}_K$  berechnet werden.

Für eine kompaktere Schreibweise soll nun  $\check{\mathbf{x}}_k(l, \mathbf{x}_l, \mathbf{w}_{l|k-1})$  mit  $l \leq k$  die Lösung von (3.1a) für den Anfangszustand  $\mathbf{x}_l$  zum Zeitpunkt  $t_l$  und die Störfolge  $\mathbf{w}_{l|k-1}$  sein. Im speziellen Fall  $l = k$  sei  $\mathbf{w}_{l|k-1}$  eine leere Folge und  $\check{\mathbf{x}}_k(l, \mathbf{x}_l, \mathbf{w}_{l|k-1}) = \mathbf{x}_l$ . In ähnlicher Weise wird auch für die Folge der Messstörungen die abgekürzte Schreibweise  $\mathbf{v}_{K-N|K-1} = (\mathbf{v}_{K-N}, \dots, \mathbf{v}_{K-1})$  verwendet.

Wird der MHE Beobachter neu eingeschaltet, so gilt zunächst für die Horizontlänge  $N = K$ , d. h. alle verfügbaren Messwerte werden genutzt. Bei MHE wird ab einem gewissen Zeitindex  $N$  festgehalten und es werden fortan nur noch die Messwerte aus einem Horizont mit fester, finiter Länge  $N$  zur Schätzung verwendet.

### 3.1.3 Beschränkungen

Bei vielen praktischen Anwendungen sind Stör-, Zustands- oder Ausgangsgrößen beschränkt. Hier werden die folgenden Beschränkungen berücksichtigt:

$$\mathbf{x}_k \in X_k, \quad \mathbf{w}_k \in W_k, \quad \mathbf{v}_k \in V_k, \quad \forall k \in \mathbb{N}_0. \quad (3.2)$$

Hierbei seien  $X_k$ ,  $W_k$  und  $V_k$  zeitvariante Mengen, die

$$X_k \subseteq \mathbb{R}^n, \quad \mathbf{0} \in W_k \subseteq \mathbb{R}^p, \quad \mathbf{0} \in V_k \subseteq \mathbb{R}^q \quad (3.3)$$

erfüllen. Die Mengen  $W_k$  und  $V_k$  bieten eine einfache Möglichkeit, die Beschränktheit von Störungen zu modellieren, die sich z. B. auch in Form von lokal verschwindenden Wahrscheinlichkeitsdichten manifestiert. Bei der Interpretation und Festlegung der Menge  $X_k$  ist Sorgfalt angebracht: Anders als ein Regler, kann ein Zustandsschätzer die Einhaltung von Schranken in der realen Strecke nicht erzwingen. Ist ein realer Systemzustand tatsächlich physikalischen Schranken ausgesetzt, so sollten diese impliziert durch das Modell (3.1) abgebildet sein. Ist dies, z. B. aufgrund von Modellungenauigkeiten, nicht der Fall, so können solche Schranken durch eine entsprechende Wahl von  $X_k$  erzwungen werden. Durch die Festlegung von  $X_k$  können also Modellfehler kompensiert oder Modellvereinfachungen erzielt werden. Eine falsche Wahl der Schranken  $X_k$  kann aber zu (unphysikalischen) Akausalitäten führen [3.1] oder die Konvergenz von Schätzwerten gegen ihre wahren Werte verhindern [3.2].

### 3.1.4 Skalares Gütemaß

Zur Beurteilung der Qualität von Schätzwerten  $\tilde{\mathbf{x}}_{K-N}$  und  $\tilde{\mathbf{w}}_{K-N|K-1}$  wird ein skalares Gütemaß der Form

$$J_{K|N}(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) = B_{K-N}(\tilde{\mathbf{x}}_{K-N}) + \sum_{k=K-N}^{K-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) \quad (3.4)$$

mit den Funktionen  $B_k : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  und  $b_k : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}_{\geq 0}$  minimiert. Die Anfangskostenfunktion  $B_{K-N} \geq 0$  enthält bekannte oder zu früheren Zeitpunkten geschätzte Informationen über den Anfangszustand  $\mathbf{x}_{K-N}$ . Da von mittelwertfreien Störgrößen ausgegangen wird, soll  $b_k(\mathbf{w}, \mathbf{v})$  positiv definit bezüglich beider Argumente  $\mathbf{w}$  und  $\mathbf{v}$  sein. Gemäß Tabelle 3.1 werden die Schätzwerte  $\tilde{\mathbf{v}}_k$  für  $\forall k = K - N, \dots, K - 1$  mittels

$$\tilde{\mathbf{v}}_k = \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k(K - N, \tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N,k-1})) \quad (3.5)$$

berechnet.

### 3.1.5 Optimierung

Zur Bestimmung der optimalen Schätzgrößen wird die beschränkte statische Optimierungsaufgabe

$$(\hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) = \arg \min_{\substack{(\tilde{\mathbf{x}}_{K-N}, \\ \tilde{\mathbf{w}}_{K-N|K-1})}} J_{K|N}(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) \quad (3.6a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{k+1} = \mathbf{f}_k(\tilde{\mathbf{x}}_k, \tilde{\mathbf{w}}_k) \quad \forall k = K-N, \dots, K-1 \quad (3.6b)$$

$$\tilde{\mathbf{v}}_k = \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k) \quad \forall k = K-N, \dots, K-1 \quad (3.6c)$$

$$\tilde{\mathbf{x}}_k \in X_k \quad \forall k = K-N, \dots, K \quad (3.6d)$$

$$\tilde{\mathbf{w}}_k \in W_k, \quad \tilde{\mathbf{v}}_k \in V_k, \quad \forall k = K-N, \dots, K-1 \quad (3.6e)$$

gelöst. In ihr sind das dynamische Modell, die Beschränkungen und das skalare Gütemaß aus den vorhergehenden Abschnitten zusammengefasst. Der optimale Wert der Kostenfunktion von (3.6) wird in der Form  $\hat{J}_{K|N} = J_{K|N}(\hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1})$  abgekürzt.

Die Lösung der Optimierungsaufgabe (3.6) ist im Allgemeinen mit hohem numerischen Aufwand verbunden, was eine der zentralen Herausforderungen bei MHE Beobachtern sein kann. Da (3.6) strukturell ähnlich zu den bei der modellprädiktiven Regelung auftretenden Optimierungsproblemen ist, wird auf die in Abschnitt 2 angeführten Literaturverweise für Lösungsmethoden verwiesen.

### 3.1.6 Annahmen

**Definition 3.1 (Vergleichsfunktionen).** Es werden die folgenden Klassen von Funktionen definiert:

$$\mathcal{K} = \{\sigma : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \mid \sigma(0) = 0 \text{ und } \sigma \text{ ist stetig und streng monoton steigend}\}$$

$$\mathcal{K}_{\infty} = \{\sigma \in \mathcal{K} \mid \lim_{t \rightarrow \infty} \sigma(t) = \infty\}$$

$$\mathcal{L} = \{\gamma : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \mid \lim_{t \rightarrow \infty} \gamma(t) = 0 \text{ und } \gamma \text{ ist stetig und nicht steigend}\}$$

$$\mathcal{KL} = \{\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \mid \beta(\cdot, t) \in \mathcal{K} \text{ und } \beta(s, \cdot) \in \mathcal{L}\}$$

Abbildung 3.1 zeigt ein Beispiel einer Funktion der Klasse  $\mathcal{KL}$ .

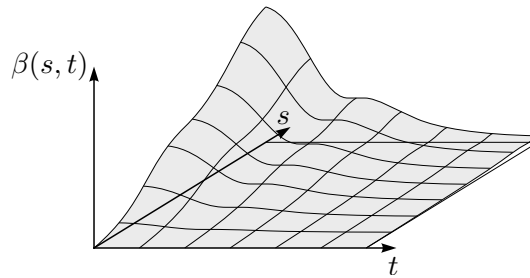


Abbildung 3.1: Beispiel für eine Funktion der Klasse  $\mathcal{KL}$ .

Es werden die folgenden Annahmen getroffen:

- A1) Für jeden beschränkten Zustand  $\mathbf{x}_k \in X_k$  und jede beschränkte Störung  $\mathbf{w}_k \in W_k$  besitze (3.1a) eine eindeutige und beschränkte Lösung  $\mathbf{x}_{k+1}$ . Die Funktion  $\mathbf{f}_k$  ist Lipschitz-stetig.

A2) Für jeden beschränkten Zustand  $\mathbf{x}_k \in X_k$  und jede beschränkte Störung  $\mathbf{v}_k \in V_k$  besitze (3.1b) eine eindeutige und beschränkte Lösung  $\mathbf{y}_k$ . Die Funktion  $\mathbf{h}_k$  ist stetig.

A3) Die Störungen  $\mathbf{w}_k$  und  $\mathbf{v}_k$  mit  $k \in \mathbb{N}_0$  sind beschränkt und es gilt

$$\lim_{k \rightarrow \infty} \mathbf{w}_k = \mathbf{0}, \quad \lim_{k \rightarrow \infty} \mathbf{v}_k = \mathbf{0}. \quad (3.7)$$

Das folgende Lemma ist eine Konsequenz der Annahme A3. Der Beweis dazu findet sich in [3.12].

**Lemma 3.1** (Beschränktheit der Summe von konvergenten Störungen). *Ist die Annahme A3 erfüllt, dann existiert eine Funktion  $\bar{\gamma}_b \in \mathcal{K}_\infty$ , so dass*

$$\sum_{k=0}^{\infty} \bar{\gamma}_b(\|\mathbf{w}_k\|_2 + \|\mathbf{v}_k\|_2) < \infty. \quad (3.8)$$

A4) Der Vektor  $\bar{\mathbf{x}}_k$  soll nun entweder dem bekannten Zustand  $\mathbf{x}_k$  oder einem früher gefundenen Schätzwert für  $\mathbf{x}_k$  entsprechen. Fortan wird  $\bar{\mathbf{x}}_k$  daher als *A-priori-Schätzung des Zustands  $\mathbf{x}_k$*  bezeichnet. Einem MHE Beobachter mit dem Horizont  $K - N, \dots, K$  steht der Wert  $\bar{\mathbf{x}}_{K-N}$  zur Verfügung, im Allgemeinen jedoch nicht der tatsächliche Zustand  $\mathbf{x}_{K-N}$ .

A5) Die Kostenfunktionen  $B_k$  und  $b_k$  aus (3.4) sind stetig und für beliebige  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{w} \in \mathbb{R}^p$  und  $\mathbf{v} \in \mathbb{R}^q$  durch

$$\underline{\gamma}_B(\|\mathbf{x} - \bar{\mathbf{x}}_k\|_2) \leq B_k(\mathbf{x}) - B_k(\bar{\mathbf{x}}_k) \leq \bar{\gamma}_B(\|\mathbf{x} - \bar{\mathbf{x}}_k\|_2) \quad \forall k \in \mathbb{N}_{\geq 0} \quad (3.9a)$$

$$\underline{\gamma}_b(\|\mathbf{w}\|_2 + \|\mathbf{v}\|_2) \leq b_k(\mathbf{w}, \mathbf{v}) \leq \bar{\gamma}_b(\|\mathbf{w}\|_2 + \|\mathbf{v}\|_2) \quad \forall k \in \mathbb{N}_0 \quad (3.9b)$$

beschränkt, wobei  $\underline{\gamma}_B, \bar{\gamma}_B, \underline{\gamma}_b, \bar{\gamma}_b \in \mathcal{K}_\infty$  und  $\bar{\gamma}_b$  in Lemma 3.1 definiert wurde.

Aus Annahme A5 folgt natürlich  $B_k(\mathbf{x}) \geq B_k(\bar{\mathbf{x}}_k)$ , wobei das Gleichheitszeichen nur im Fall  $\mathbf{x} = \bar{\mathbf{x}}_k$  gilt.

## 3.2 Stabilität von Zustandsschätzern

Bevor verschiedene Varianten von Zustandsbeobachtern basierend auf der Optimierungsaufgabe (3.6) besprochen werden, wird der Begriff der *Stabilität des Schätzfehlers* in Anlehnung an [3.2] definiert. Ist das System (3.1) keinen Störungen  $\mathbf{w}_k$  und  $\mathbf{v}_k$  ausgesetzt und ist der Anfangszustand exakt bekannt, d. h.  $\bar{\mathbf{x}}_{K-N} = \mathbf{x}_{K-N}$ , dann gilt für die optimale Lösung von (3.6) mit  $N \leq K$  natürlich

$$\hat{\mathbf{x}}_{K-N} = \mathbf{x}_{K-N} \quad (3.10a)$$

$$\hat{\mathbf{w}}_{K-N|K-1} = (\mathbf{0}, \dots, \mathbf{0}) \quad (3.10b)$$

$$\hat{\mathbf{v}}_k = \mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k(K-N), \hat{\mathbf{x}}_{K-N}, (\mathbf{0}, \dots, \mathbf{0})) = \mathbf{0} \quad \forall k = K-N, \dots, K-1 \quad (3.10c)$$

und  $\hat{J}_{K|N} = B_{K-N}(\mathbf{x}_{K-N})$ . Abweichungen von diesem Idealzustand können verschiedene Ursachen haben:



- Ungenaue Informationen über den Anfangszustand,  $\bar{\mathbf{x}}_{K-N} \neq \mathbf{x}_{K-N}$
- Nicht verschwindende Prozessstörungen,  $\mathbf{w}_k \neq \mathbf{0}$
- Nicht verschwindende Messstörungen,  $\mathbf{v}_k \neq \mathbf{0}$

Die folgenden Stabilitätsdefinitionen klären, wie der Schätzfehler auf derartige Abweichungen reagiert.

**Definition 3.2** (Nominell global asymptotisch stabiler Schätzfehler (NGAS)). Eine Zustandsschätzung für das System (3.1) beginnend zum Zeitpunkt  $k \geq 0$  ist *nominell global asymptotisch stabil*, wenn im Falle von verschwindenden Störungen, d. h.  $\mathbf{w}_l = \mathbf{0}$  und  $\mathbf{v}_l = \mathbf{0}$  für  $\forall l = k, \dots, K-1$  mit  $k \leq K$ , eine Funktion  $\beta(\cdot, \cdot) \in \mathcal{KL}$  existiert, so dass für beliebige  $\mathbf{x}_k, \bar{\mathbf{x}}_k \in \mathbb{R}^n$  gemäß Annahme A4 und für  $\forall K \in \mathbb{N}_0$

$$\| \underbrace{\check{\mathbf{x}}_K(k, \mathbf{x}_k, (\mathbf{0}, \dots, \mathbf{0}))}_{= \mathbf{x}_K} - \check{\mathbf{x}}_K(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|K-1}) \|_2 \leq \beta(\|\mathbf{x}_k - \bar{\mathbf{x}}_k\|_2, K - k) \quad (3.11)$$

erfüllt ist.

**Definition 3.3** (Robust global asymptotisch stabiler Schätzfehler (RGAS)). Eine Zustandsschätzung für das System (3.1) beginnend zum Zeitpunkt  $k \geq 0$  ist *robust global asymptotisch stabil*, wenn bei gegebenen Funktionen  $\bar{\gamma}_B(\cdot), \bar{\gamma}_b(\cdot) \in \mathcal{K}_\infty$  für jedes  $\varepsilon > 0$  ein  $\delta(\varepsilon) > 0$  existiert, so dass für beliebige  $\mathbf{x}_k, \bar{\mathbf{x}}_k \in \mathbb{R}^n$  gemäß Annahme A4 und beliebige Störungen  $\mathbf{w}_l$  und  $\mathbf{v}_l$  für  $\forall l = k, \dots, K-1$ , die der Annahme A3 genügen,

$$\begin{aligned} \bar{\gamma}_B(\|\mathbf{x}_k - \bar{\mathbf{x}}_k\|_2) + \sum_{l=k}^{\infty} \bar{\gamma}_b(\|\mathbf{w}_l\|_2 + \|\mathbf{v}_l\|_2) \leq \delta(\varepsilon) \quad \Rightarrow \\ \| \underbrace{\check{\mathbf{x}}_K(k, \mathbf{x}_k, \mathbf{w}_{k|K-1})}_{= \mathbf{x}_K} - \check{\mathbf{x}}_K(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|K-1}) \|_2 \leq \varepsilon \quad \forall K \in \mathbb{N}_0 \end{aligned} \quad (3.12)$$

mit  $K \geq k$  erfüllt ist und darüber hinaus

$$\lim_{K \rightarrow \infty} \underbrace{(\check{\mathbf{x}}_K(k, \mathbf{x}_k, \mathbf{w}_{k|K-1}) - \check{\mathbf{x}}_K(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|K-1}))}_{= \mathbf{x}_K} = \mathbf{0} \quad (3.13)$$

gilt.

Die Definition 3.3 verlangt also, dass der Beobachtungsfehler beschränkt bleibt und für  $K \rightarrow \infty$  gegen  $\mathbf{0}$  konvergiert. Jeder RGAS Schätzer ist auch NGAS.

### 3.3 Zustandsschätzung mit vollständiger Information

Wird die Zustandsschätzung mit der Horizontlänge  $N = K$  durchgeführt, so spricht man von *Schätzung mit vollständiger Information*. Diese Beobachtersvariante wird hier kurz studiert, da sie aus theoretischer Sicht sehr gute Stabilitäts- und Konvergenzeigenschaften

besitzt. Ihr Nachteil ist, dass der Rechenaufwand im Allgemeinen zumindest proportional zu  $K$  und daher ohne Schranken wächst.

Im vorliegenden Abschnitt wird von einer Zustandsschätzung durch Lösen der Optimierungsaufgabe (3.6) mit  $N = K$  ausgegangen. Sind das Modell (3.1) und die Beschränkungen (3.2) bekannt, so sind beim Entwurf des Schätzers lediglich die Kostenfunktionen  $B_0$  und  $b_k$  für  $\forall k = 0, \dots, K-1$  zu wählen.

Es stellt sich nun die Frage für welche Klasse von Systemen der Fehler einer Zustandsschätzung mit  $N = K$  überhaupt stabil sein kann. Ähnlich zur *Detektierbarkeit* bei linearen Systemen kann für nichtlineare zeitdiskrete Systeme die Eigenschaft der *inkrementellen Eingangs/Ausgangs-Zustands-Stabilität* (siehe [3.11, 3.13]) als notwendige Voraussetzung für einen stabilen Beobachtungsfehler verwendet werden.

**Definition 3.4 (Inkrementelle Eingangs/Ausgangs-Zustands-Stabilität (IIOSS)).** Das System (3.1) ist *inkrementell Eingangs/Ausgangs-Zustands-stabil*, wenn Funktionen  $\beta(\cdot, \cdot) \in \mathcal{KL}$  und  $\gamma_1(\cdot), \gamma_2(\cdot) \in \mathcal{K}$  existieren, so dass für beliebige Anfangszustände  $\mathbf{x}_a, \mathbf{x}_b$  zu einem Zeitpunkt  $k \in \mathbb{N}_{\geq 0}$ , beliebige Störfolgen  $\mathbf{w}_{a,k|K-1}, \mathbf{w}_{b,k|K-1}$  und für  $\forall k \in \mathbb{N}_0$  mit  $k \leq K$

$$\begin{aligned} & \|\check{\mathbf{x}}_K(k, \mathbf{x}_a, \mathbf{w}_{a,k|K-1}) - \check{\mathbf{x}}_K(k, \mathbf{x}_b, \mathbf{w}_{b,k|K-1})\|_2 \\ & \leq \beta(\|\mathbf{x}_a - \mathbf{x}_b\|_2, K - k) + \gamma_1\left(\max_{l=k, \dots, K-1} \{\|\mathbf{w}_{a,l} - \mathbf{w}_{b,l}\|_2\}\right) \\ & \quad + \gamma_2\left(\max_{l=k, \dots, K-1} \{\|\mathbf{h}_l(\check{\mathbf{x}}_l(k, \mathbf{x}_a, \mathbf{w}_{a,k|l-1})) - \mathbf{h}_l(\check{\mathbf{x}}_l(k, \mathbf{x}_b, \mathbf{w}_{b,k|l-1}))\|_2\}\right) \end{aligned} \quad (3.14)$$

gilt.

In dieser Definition steht der Begriff *inkrementell* für den Vergleich zweier beliebiger Zustandsfolgen. Die *Eingangs/Ausgangs-Zustands-Stabilität* (IOSS) [3.13] ist eine schwächere Eigenschaft, betrachtet nur eine einzelne Zustandsfolge und gibt an, wann diese Zustandsfolge in den Ursprung konvergiert. Für Systeme mit  $\mathbf{0} = \mathbf{f}_k(\mathbf{0}, \mathbf{0})$  und  $\mathbf{0} = \mathbf{h}_k(\mathbf{0})$  impliziert die Eigenschaft IIOSS die Eigenschaft IOSS [3.13]. Die Umkehrung gilt nicht. Für lineare Systeme sind IIOSS und IOSS äquivalente Eigenschaften.

**Lemma 3.2 (Konvergenz des Zustandes von IIOSS-Systemen).** Wenn das System (3.1) die IIOSS-Eigenschaft gemäß Definition 3.4 besitzt sowie für  $\lim_{K \rightarrow \infty} (\mathbf{w}_{a,K} - \mathbf{w}_{b,K}) = \mathbf{0}$  und beliebige Anfangszustände  $\mathbf{x}_a, \mathbf{x}_b$  zum Zeitpunkt  $k \in \mathbb{N}_0$   $\lim_{K \rightarrow \infty} (\mathbf{h}_K(\check{\mathbf{x}}_K(k, \mathbf{x}_a, \mathbf{w}_{a,k|K-1})) - \mathbf{h}_K(\check{\mathbf{x}}_K(k, \mathbf{x}_b, \mathbf{w}_{b,k|K-1}))) = \mathbf{0}$  mit  $K \geq k$  gilt, dann folgt daraus

$$\lim_{K \rightarrow \infty} (\check{\mathbf{x}}_K(k, \mathbf{x}_a, \mathbf{w}_{a,k|K-1}) - \check{\mathbf{x}}_K(k, \mathbf{x}_b, \mathbf{w}_{b,k|K-1})) = \mathbf{0} . \quad (3.15)$$

**Aufgabe 3.1.** Beweisen Sie Lemma 3.2.

**Satz 3.1 (RGAS der Zustandsschätzung mit vollständiger Information).** Für ein System (3.1), das die IIOSS-Eigenschaft besitzt, und Störungen, die der Annahme A3

genügen, führt der Zustandsschätzer (3.6) mit  $N = K$  (Schätzung mit vollständiger Information) und Kostenfunktionen  $B_0$  und  $b_k$  gemäß der Annahme A5 zu einem robust global asymptotisch stabilen Schätzfehler.

*Beweis.* Es sei  $(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|\infty}) = (\mathbf{x}_0, \mathbf{w}_{0|\infty})$  eine zulässige, wenngleich nicht notwendigerweise optimale und bekannte Lösung der Optimierungsaufgabe (3.6) für  $N = K \rightarrow \infty$ . Aus Lemma 3.1 und der Annahme A5 erhält man

$$\begin{aligned} J_{\infty|\infty}(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|\infty}) &= B_0(\mathbf{x}_0) + \sum_{k=0}^{\infty} b_k(\mathbf{w}_k, \mathbf{v}_k) - B_0(\bar{\mathbf{x}}_0) + B_0(\bar{\mathbf{x}}_0) \\ &\leq \bar{\gamma}_B(\|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_2) + \sum_{k=0}^{\infty} \bar{\gamma}_b(\|\mathbf{w}_k\|_2 + \|\mathbf{v}_k\|_2) + B_0(\bar{\mathbf{x}}_0) = \bar{J}, \end{aligned} \quad (3.16)$$

wobei  $\bar{J}$  eine Konstante ist. Daraus folgt direkt

$$\hat{J}_{K|K} \leq \bar{J} \quad \forall K \in \mathbb{N}_0. \quad (3.17)$$

Diese Beschränktheit des optimalen Gütefunktionswert  $\hat{J}_{K|K}$  und die Annahme A5 implizieren die Existenz einer eindeutigen Lösung der Optimierungsaufgabe (3.6) mit  $N = K$  für  $\forall K \in \mathbb{N}_0$ .

Es sei  $(\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1})$  das zu einem beliebigen Zeitpunkt  $K$  gefundene optimale Schätzergebnis. Sicher ist dieses Schätzergebnis auch eine zulässige aber im Allgemeinen nicht optimale Lösung der Optimierungsaufgabe (3.6) mit  $N = K - 1$  zum Zeitpunkt  $K - 1$ . Es gilt daher

$$\begin{aligned} J_{K-1|K-1}(\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-2}) &= \hat{J}_{K|K} - b_{K-1}(\hat{\mathbf{w}}_{K-1}, \underbrace{\mathbf{y}_{K-1} - \mathbf{h}_{K-1}(\check{\mathbf{x}}_{K-1}(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-2}))}_{= \hat{\mathbf{v}}_{K-1}}) \end{aligned} \quad (3.18)$$

und folglich

$$\hat{J}_{K|K} \geq \hat{J}_{K-1|K-1} + b_{K-1}(\hat{\mathbf{w}}_{K-1}, \hat{\mathbf{v}}_{K-1}). \quad (3.19)$$

Gemäß (3.17) und (3.19) ist die Folge  $(\hat{J}_{K|K})$  nach oben durch  $\bar{J}$  beschränkt und nicht fallend. Aus (3.19), das für beliebige  $K$  gilt, folgt daher

$$\lim_{K \rightarrow \infty} b_K(\hat{\mathbf{w}}_K, \hat{\mathbf{v}}_K) = 0, \quad (3.20)$$

was wegen (3.9b) und (3.5)

$$\lim_{K \rightarrow \infty} \hat{\mathbf{w}}_K = \mathbf{0} \quad (3.21a)$$

$$\lim_{K \rightarrow \infty} \hat{\mathbf{v}}_K = \lim_{K \rightarrow \infty} (\mathbf{y}_K - \mathbf{h}_K(\check{\mathbf{x}}_K(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1}))) = \mathbf{0} \quad (3.21b)$$

impliziert. Weil gemäß Annahme A3 auch

$$\lim_{K \rightarrow \infty} \mathbf{w}_K = \mathbf{0} \quad (3.22a)$$

$$\lim_{K \rightarrow \infty} \mathbf{v}_K = \lim_{K \rightarrow \infty} (\mathbf{y}_K - \mathbf{h}_K(\underbrace{\check{\mathbf{x}}_K(0, \mathbf{x}_0, \mathbf{w}_{0|K-1})}_{=\mathbf{x}_K})) = \mathbf{0} \quad (3.22b)$$

gilt, erhält man

$$\lim_{K \rightarrow \infty} (\hat{\mathbf{w}}_K - \mathbf{w}_K) = \mathbf{0} \quad (3.23a)$$

$$\lim_{K \rightarrow \infty} (\mathbf{h}_K(\check{\mathbf{x}}_K(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1})) - \mathbf{h}_K(\mathbf{x}_K)) = \mathbf{0} . \quad (3.23b)$$

Da das System die HIOSS-Eigenschaft besitzt, ist die Existenz von Funktionen  $\beta(\cdot, \cdot) \in \mathcal{KL}$  und  $\gamma_1(\cdot), \gamma_2(\cdot) \in \mathcal{K}$  gesichert, so dass

$$\begin{aligned} & \|\check{\mathbf{x}}_K(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|K-1}) - \underbrace{\check{\mathbf{x}}_K(k, \mathbf{x}_k, \mathbf{w}_{k|K-1})}_{=\mathbf{x}_K}\|_2 \\ & \leq \beta(\|\hat{\mathbf{x}}_k - \mathbf{x}_k\|_2, K - k) + \gamma_1\left(\max_{l=k, \dots, K-1} \{\|\hat{\mathbf{w}}_l - \mathbf{w}_l\|_2\}\right) \\ & \quad + \gamma_2\left(\max_{l=k, \dots, K-1} \{\|\mathbf{h}_l(\check{\mathbf{x}}_l(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|l-1})) - \underbrace{\mathbf{h}_l(\check{\mathbf{x}}_l(k, \mathbf{x}_k, \mathbf{w}_{k|l-1}))}_{=\mathbf{x}_l}\|_2\}\right) \end{aligned} \quad (3.24)$$

für  $\forall k, K \in \mathbb{N}_0$  mit  $k \leq K$  gilt. Analog zu Lemma 3.2 folgt aus (3.23) und (3.24)

$$\lim_{K \rightarrow \infty} (\check{\mathbf{x}}_K(k, \hat{\mathbf{x}}_k, \hat{\mathbf{w}}_{k|K-1}) - \underbrace{\check{\mathbf{x}}_K(k, \mathbf{x}_k, \mathbf{w}_{k|K-1})}_{=\mathbf{x}_K}) = \mathbf{0} \quad (3.25)$$

womit die in (3.13) geforderte Konvergenz des Schätzfehlers gezeigt ist.

Wählt man nun ein festes  $\delta \geq 0$  und fordert  $\bar{J} = \delta$  für  $\bar{J}$  aus (3.16), so folgt unter Berücksichtigung von (3.9a) und  $B_0(\bar{\mathbf{x}}_0) \geq 0$  aus (3.16) und (3.17), dass

$$\bar{\gamma}_B(\|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_2) \leq \delta, \quad \underline{\gamma}_B(\|\hat{\mathbf{x}}_0 - \bar{\mathbf{x}}_0\|_2) \leq \delta, \quad (3.26)$$

wobei hier  $\hat{\mathbf{x}}_0$  der zu einem beliebigen Zeitpunkt  $K$  ermittelte optimale Schätzwert des Anfangszustands  $\mathbf{x}_0$  ist. Umformung von (3.26) führt mithilfe der Dreiecksungleichung auf

$$\|\hat{\mathbf{x}}_0 - \mathbf{x}_0\|_2 \leq \bar{\gamma}_B^{-1}(\delta) + \underline{\gamma}_B^{-1}(\delta), \quad (3.27)$$

womit die Beschränktheit des Anfangsschätzfehlers für Schätzungen zu beliebigen Zeitpunkten  $K$  gesichert ist. Mit der Forderung  $\bar{J} = \delta$  folgt unter Berücksichtigung von (3.9b) und  $B_0(\bar{\mathbf{x}}_0) \geq 0$  aus (3.16) und (3.17) ferner, dass

$$\bar{\gamma}_b(\|\mathbf{w}_l\|_2) \leq \delta, \quad \underline{\gamma}_b(\|\hat{\mathbf{w}}_l\|_2) \leq \delta, \quad \forall l \in \mathbb{N}_0 \quad (3.28a)$$

$$\bar{\gamma}_b(\|\mathbf{v}_l\|_2) \leq \delta, \quad \underline{\gamma}_b(\|\hat{\mathbf{v}}_l\|_2) \leq \delta, \quad \forall l \in \mathbb{N}_0, \quad (3.28b)$$

wobei hier  $\hat{\mathbf{w}}_l$  und  $\hat{\mathbf{v}}_l$  die zu einem beliebigen Zeitpunkt  $K > l$  berechneten optimalen Schätzwerte für die Störungen  $\mathbf{w}_l$  und  $\mathbf{v}_l$  sind. Unter Verwendung von (3.5) und der Dreiecksungleichung lässt sich (3.28) in die Form

$$\|\hat{\mathbf{w}}_l - \mathbf{w}_l\|_2 \leq \bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta) \quad \forall l \in \mathbb{N}_0 \quad (3.29a)$$

$$\begin{aligned} \|\hat{\mathbf{v}}_l - \mathbf{v}_l\|_2 &= \|\mathbf{h}_l(\check{\mathbf{x}}_l(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|l-1})) - \mathbf{h}_l(\mathbf{x}_l)\|_2 \\ &\leq \bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta) \quad \forall l \in \mathbb{N}_0 \end{aligned} \quad (3.29b)$$

umschreiben. Die Abschätzungen (3.27) und (3.29) können nun gemeinsam mit  $k = 0$  in (3.24) eingesetzt werden und man erhält unter Berücksichtigung von  $\beta(\cdot, l) \leq \beta(\cdot, 0) \forall l \geq 0$

$$\begin{aligned} \|\check{\mathbf{x}}_K(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1}) - \mathbf{x}_K\|_2 &\leq \beta(\bar{\gamma}_B^{-1}(\delta) + \underline{\gamma}_B^{-1}(\delta), 0) + \gamma_1(\bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta)) \\ &\quad + \gamma_2(\bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta)) \quad \forall K \in \mathbb{N}_0, \end{aligned} \quad (3.30)$$

wobei  $K$  wieder der Zeitpunkt der Schätzung ist. Die gesamte rechte Seite von (3.30) ist eine  $\mathcal{K}$ -Funktion in  $\delta$  und daher nach  $\delta$  auflösbar. Es existiert also zu jedem gegebenen  $\varepsilon \geq 0$  ein  $\delta(\varepsilon) \geq 0$  wobei  $\delta(0) = 0$ , so dass

$$\|\check{\mathbf{x}}_K(0, \hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1}) - \mathbf{x}_K\|_2 \leq \varepsilon \quad \forall K \in \mathbb{N}_0. \quad (3.31)$$

Damit ist auch die Existenz einer Beziehung  $\delta(\varepsilon)$ , die die Erfüllung der Implikation (3.12) und somit die Beschränktheit des Schätzfehlers sichert, gezeigt.  $\square$

### 3.4 Zustandsschätzung auf bewegtem Horizont

Erfolgt zum Zeitindex  $K$  eine MHE Zustandsschätzung durch Lösen der Optimierungsaufgabe (3.6), so werden Messwerte  $\mathbf{y}_{K-N|K-1}$  aus dem aktuellen Zeithorizont  $K - N, \dots, K$  verwendet. Dieser umfasst  $N \in \mathbb{N}_0$  Abtastintervalle, wobei  $N$  eine feste finite Zahl ist. Im vorliegenden Abschnitt wird nur noch der Fall  $K > N$  explizit betrachtet. Im Fall  $K \leq N$ , also wenn der Beobachter neu eingeschaltet wird, verwendet man einfach die Zustandsschätzung mit vollständiger Information gemäß Abschnitt 3.3.

Sind das Modell (3.1) und die Beschränkungen (3.2) bekannt, so sind beim Entwurf des MHE Beobachters lediglich die Horizontlänge  $N$  sowie die Kostenfunktionen  $B_{K-N}$  und  $b_k$  für  $\forall k = K - N, \dots, K - 1$  zu wählen. Diese Wahl ist natürlich entscheidend für die Stabilität und Konvergenz des Schätzfehlers. Ein Vergleich der Gütefunktion

$$J_{K|K}(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|K-1}) = B_0(\tilde{\mathbf{x}}_0) + \sum_{k=0}^{K-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) \quad (3.32)$$

aus der Zustandsschätzung mit vollständiger Information mit der Gütefunktion

$$J_{K|N}(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) = B_{K-N}(\tilde{\mathbf{x}}_{K-N}) + \sum_{k=K-N}^{K-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) \quad (3.33)$$

für MHE zeigt, dass ein wesentlicher Punkt des Entwurfes die Einbeziehung vergangener Informationen und Schätzergebnisse in den Anfangskostenterm  $B_k(\cdot)$  (meist mit  $k = K - N$ ) ist.  $B_k(\tilde{\mathbf{x}}_k)$  soll daher Abweichungen zwischen  $\tilde{\mathbf{x}}_k$  und der A-priori-Schätzung  $\bar{\mathbf{x}}_k$  bestrafen. Weist der Schätzwert  $\tilde{\mathbf{x}}_k$  eine hohe (geringe) *statistische Zuverlässigkeit* auf, so sollte die Abweichung  $\tilde{\mathbf{x}}_k - \bar{\mathbf{x}}_k$  durch  $B_k(\cdot)$  entsprechend hoch (gering) bestraft werden [3.8].

### 3.4.1 Anfangskostenterm für vollständige Information

Eine naheliegende Frage ist nun, wie der Anfangskostenterm  $B_k(\cdot)$  zu wählen ist, damit der MHE Beobachter und die Zustandsschätzung mit vollständiger Information (siehe Abschnitt 3.3) die gleichen Ergebnisse liefern. Um diese Frage zu beantworten (siehe [3.8]), wird zunächst (3.32) in die Form

$$J_{K|K}(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|K-1}) = \underbrace{B_0(\tilde{\mathbf{x}}_0) + \sum_{k=0}^{K-N-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k)}_{\text{Anfangskostenterm}} + \sum_{k=K-N}^{K-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) \quad (3.34)$$

umgeschrieben. Offensichtlich hängt die letzte Summe in (3.34) nur von  $\tilde{\mathbf{x}}_{K-N}$ ,  $\tilde{\mathbf{w}}_{K-N|K-1}$  und den gegebenen Messwerten  $\mathbf{y}_{K-N|K-1}$  ab. Die Unabhängigkeit des Systemverhaltens im Horizont  $K - N, \dots, K$  von Zuständen vor dem Zeitpunkt  $K - N$  wird auch als *Markov-Eigenschaft* des zugrunde liegenden stochastischen Prozesses bezeichnet. Gemäß dem Optimalitätsprinzip nach Bellman [3.10] kann daher der als *Anfangskostenterm* bezeichnete Term in (3.34) in ein separates Optimierungsproblem als Funktion von  $\tilde{\mathbf{x}}_{K-N}$  ausgelagert werden. Dieses lautet für einen allgemeinen Zeitpunkt  $k \in \mathbb{N}_0$

$$\hat{J}_{k|k}^B(\mathbf{x}) = \min_{(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|k-1})} J_{k|k}(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|k-1}) \quad (3.35a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{l+1} = \mathbf{f}_l(\tilde{\mathbf{x}}_l, \tilde{\mathbf{w}}_l) \quad \forall l = 0, \dots, k-1 \quad (3.35b)$$

$$\tilde{\mathbf{v}}_l = \mathbf{y}_l - \mathbf{h}_l(\tilde{\mathbf{x}}_l) \quad \forall l = 0, \dots, k-1 \quad (3.35c)$$

$$\tilde{\mathbf{x}}_l \in X_l \quad \forall l = 0, \dots, k-1 \quad (3.35d)$$

$$\tilde{\mathbf{w}}_l \in W_l, \quad \tilde{\mathbf{v}}_l \in V_l, \quad \forall l = 0, \dots, k-1 \quad (3.35e)$$

$$\tilde{\mathbf{x}}_k = \mathbf{x}. \quad (3.35f)$$

Die Funktion  $\hat{J}_{k|k}^B(\mathbf{x})$  liefert die minimalen Kosten, die entstehen wenn die Zustandsfolge zum Zeitpunkt  $k$  am Punkt  $\mathbf{x} \in X_k$  ankommt. Die Funktion wird daher auch als *Ankunftskostenterm* (Englisch: *arrival cost*) bezeichnet (vgl. [3.2]). Sie sorgt für die Berücksichtigung von Informationen, die vor dem Zeitpunkt  $k$  gesammelt wurden (z. B. die Messwerte  $\mathbf{y}_{0|k-1}$ ) und daher nicht explizit in einer MHE Optimierungsaufgabe auftreten, deren Horizont zum Zeitpunkt  $k$  beginnt. Der einzige Unterschied zwischen (3.35) und der ursprünglichen Optimierungsaufgabe (3.6) ist die Endbedingung (3.35f). Offensichtlich gilt aufgrund von dieser Endbedingung

$$\hat{J}_{k|k}^B(\mathbf{x}) \geq \hat{J}_{k|k} \quad \forall \mathbf{x} \in X_k, k \in \mathbb{N}_0. \quad (3.36)$$

**Lemma 3.3** (Bedingung für Äquivalenz zwischen MHE und Zustandsschätzung mit vollständiger Information). Wenn die MHE Optimierungsaufgabe (3.6) für  $K \geq N$  unter Verwendung des speziellen Anfangskostenterms

$$B_{K-N}(\tilde{\mathbf{x}}_{K-N}) = \hat{J}_{K-N}^B(\tilde{\mathbf{x}}_{K-N}) \quad (3.37)$$

mit  $\hat{J}_{K-N}^B(\cdot)$  gemäß (3.35) formuliert und gelöst wird, so stimmt ihre Lösung im Intervall  $K-N, \dots, K$  mit jener der Zustandsschätzung mit vollständiger Information gemäß Abschnitt 3.3 exakt überein.

Der Beweis dieses Lemmas folgt konstruktiv aus dem Vorangegangenen.

Lemma 3.3 zeigt, dass eine MHE Formulierung gefunden werden kann, die zur Zustandsschätzung mit vollständiger Information äquivalent ist und daher die gleichen guten Stabilitäts- und Konvergenzeigenschaften besitzt. Im Allgemeinen ist mit dieser MHE Formulierung aber auch der Rechenaufwand äquivalent zu jenem der Zustandsschätzung mit vollständiger Information. Der hohe Rechenaufwand ist dem Anfangskostenterm (3.37) geschuldet und limitiert den praktischen Nutzen dieser MHE Formulierung. Im Folgenden werden daher die Stabilitäts- und Konvergenzeigenschaften von MHE Varianten mit alternativen Anfangskostentermen  $B_k(\cdot)$  untersucht.

### 3.4.2 Kein Anfangskostenterm

In der MHE Formulierung kann auch einfach auf den Anfangskostenterm  $B_{K-N}$  verzichtet werden, d. h.  $B_k(\cdot) = 0 \forall k \in \mathbb{N}_0$  und die Optimierungsaufgabe lautet

$$\hat{J}_{K|N}^0 = \min_{\substack{(\tilde{\mathbf{x}}_{K-N}, \\ \tilde{\mathbf{w}}_{K-N|K-1})}} J_{K|N}(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) = \sum_{k=K-N}^{K-1} b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) \quad (3.38a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{k+1} = \mathbf{f}_k(\tilde{\mathbf{x}}_k, \tilde{\mathbf{w}}_k) \quad \forall k = K-N, \dots, K-1 \quad (3.38b)$$

$$\tilde{\mathbf{v}}_k = \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k) \quad \forall k = K-N, \dots, K-1 \quad (3.38c)$$

$$\tilde{\mathbf{x}}_k \in X_k \quad \forall k = K-N, \dots, K \quad (3.38d)$$

$$\tilde{\mathbf{w}}_k \in W_k, \quad \tilde{\mathbf{v}}_k \in V_k, \quad \forall k = K-N, \dots, K-1. \quad (3.38e)$$

Allerdings reicht dann die IIOSS-Eigenschaft des Systems (3.1) nicht mehr aus, um die Existenz eines optimalen MHE Schätzwertes und die (asymptotische) Stabilität des Schätzfehlers zu garantieren. Wegen der Wahl  $B_k(\cdot) = 0$  ist natürlich die Bedingung (3.9a) aus Annahme A5 in diesem Abschnitt nicht erfüllt. Aus diesem Grund ist die Existenz einer beschränkten optimalen Lösung der MHE Optimierungsaufgabe in diesem Fall nicht mehr gesichert. Gemäß [3.2] ist die nachfolgend definierte Beobachtbarkeitseigenschaft des Systems (3.1), welche stärker ist als die IIOSS-Eigenschaft, für die Existenz und Eindeutigkeit der Lösung von (3.38) hinreichend.

**Definition 3.5** (Beobachtbarkeit des Anfangszustandes). Der Anfangszustand des Systems (3.1) ist *beobachtbar*, wenn eine Zahl  $\underline{N} \in \mathbb{N}$  und Funktionen  $\gamma_1(\cdot), \gamma_2(\cdot) \in \mathcal{K}$  existieren, so dass für beliebige Anfangszustände  $\mathbf{x}_a, \mathbf{x}_b$  zum Zeitpunkt  $K-N$ , beliebige

Störfolgen  $\mathbf{w}_{a,K-N|K-1}$ ,  $\mathbf{w}_{b,K-N|K-1}$  und  $\forall K, N \in \mathbb{N}_0$  mit  $K \geq N \geq \underline{N}$

$$\begin{aligned} \|\mathbf{x}_a - \mathbf{x}_b\|_2 \leq & \gamma_1 \left( \max_{k=K-N, \dots, K-1} \{\|\mathbf{w}_{a,k} - \mathbf{w}_{b,k}\|_2\} \right) \\ & + \gamma_2 \left( \max_{k=K-N, \dots, K-1} \{ \|\mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \mathbf{x}_a, \mathbf{w}_{a,K-N|k-1})) \right. \\ & \quad \left. - \mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \mathbf{x}_b, \mathbf{w}_{b,K-N|k-1}))\|_2 \} \right) \end{aligned} \quad (3.39)$$

gilt.

**Definition 3.6** (Beobachbarkeit des Endzustandes). Der Endzustand des Systems (3.1) ist *beobachtbar*, wenn eine Zahl  $\underline{N} \in \mathbb{N}$  und Funktionen  $\bar{\gamma}_1(\cdot)$ ,  $\bar{\gamma}_2(\cdot) \in \mathcal{K}$  existieren, so dass für beliebige Anfangszustände  $\mathbf{x}_a$ ,  $\mathbf{x}_b$  zum Zeitpunkt  $K - N$ , beliebige Störfolgen  $\mathbf{w}_{a,K-N|K-1}$ ,  $\mathbf{w}_{b,K-N|K-1}$  und  $\forall K, N \in \mathbb{N}_0$  mit  $K \geq N \geq \underline{N}$

$$\begin{aligned} & \|\check{\mathbf{x}}_K(K-N, \mathbf{x}_a, \mathbf{w}_{a,K-N|K-1}) - \check{\mathbf{x}}_K(K-N, \mathbf{x}_b, \mathbf{w}_{b,K-N|K-1})\|_2 \\ & \leq \bar{\gamma}_1 \left( \max_{k=K-N, \dots, K-1} \{\|\mathbf{w}_{a,k} - \mathbf{w}_{b,k}\|_2\} \right) \\ & \quad + \bar{\gamma}_2 \left( \max_{k=K-N, \dots, K-1} \{ \|\mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \mathbf{x}_a, \mathbf{w}_{a,K-N|k-1})) \right. \\ & \quad \quad \left. - \mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \mathbf{x}_b, \mathbf{w}_{b,K-N|k-1}))\|_2 \} \right) \end{aligned} \quad (3.40)$$

gilt.

*Bemerkung 3.1.* Die Eigenschaft *Beobachbarkeit des Anfangszustandes* sichert also bei gleichen Störungen  $\mathbf{w}_{a,K-N|K-1} = \mathbf{w}_{b,K-N|K-1}$  und  $\mathbf{v}_{a,K-N|K-1} = \mathbf{v}_{b,K-N|K-1}$  und gleichen Messwerten  $\mathbf{y}_{a,K-N|K-1} = \mathbf{y}_{b,K-N|K-1}$  für  $\forall N \geq \underline{N}$  die Äquivalenz der Anfangszustände, d. h.  $\mathbf{x}_a = \mathbf{x}_b$ . Die Eigenschaft *Beobachbarkeit des Endzustandes* hingegen sichert bei gleichen Störungen  $\mathbf{w}_{a,K-N|K-1} = \mathbf{w}_{b,K-N|K-1}$  und  $\mathbf{v}_{a,K-N|K-1} = \mathbf{v}_{b,K-N|K-1}$  und gleichen Messwerten  $\mathbf{y}_{a,K-N|K-1} = \mathbf{y}_{b,K-N|K-1}$  für  $\forall N \geq \underline{N}$  für beliebige Anfangszustände  $\mathbf{x}_a$ ,  $\mathbf{x}_b$  die Äquivalenz der Endzustände, d. h.  $\check{\mathbf{x}}_K(K-N, \mathbf{x}_a, \mathbf{w}_{a,K-N|K-1}) = \check{\mathbf{x}}_K(K-N, \mathbf{x}_b, \mathbf{w}_{b,K-N|K-1})$ . Wegen der Annahme **A1** über die Lipschitz-Stetigkeit der Funktion  $\mathbf{f}_k$  impliziert die Beobachbarkeit des Anfangszustandes die Beobachbarkeit des Endzustandes. Ein Vergleich der Definitionen 3.4 und 3.6 zeigt ferner, dass (abgesehen von der Forderung  $N \geq \underline{N}$ ) die Beobachbarkeit des Endzustandes die Eigenschaft **IIOSS** impliziert.

**Satz 3.2** (RGAS der MHE Zustandsschätzung ohne Anfangskostenterm). Für ein System (3.1), dessen Anfangszustand beobachtbar im Sinne der Definition 3.5 ist, mit Störungen, die der Annahme **A3** genügen, führt der MHE Zustandsschätzer (3.38) (Schätzung ohne Anfangskostenterm) mit festem  $N$  und  $b_k$  gemäß der Annahme **A5** zu einem robust global asymptotisch stabilen Schätzfehler.



*Beweis.* Der Beweis erfolgt ähnlich zum Beweis des Satzes 3.1. Für die Optimierungsaufgabe (3.38) sei nun  $(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) = (\mathbf{x}_{K-N}, \mathbf{w}_{K-N|K-1})$  eine zulässige, wengleich nicht notwendigerweise optimale und bekannte Lösung. Folglich gilt mit (3.9b) und dem in (3.16) definierten Wert  $\bar{J}$

$$\begin{aligned} \sum_{k=K-N}^{K-1} \underline{\gamma}_b(\|\hat{\mathbf{w}}_k\|_2 + \|\hat{\mathbf{v}}_k\|_2) &\leq \hat{J}_{K|N}^0 \leq J_{K|N}(\tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|K-1}) \\ &= \sum_{k=K-N}^{K-1} b_k(\mathbf{w}_k, \mathbf{v}_k) \leq \sum_{k=K-N}^{K-1} \bar{\gamma}_b(\|\mathbf{w}_k\|_2 + \|\mathbf{v}_k\|_2) \leq \bar{J}. \end{aligned} \quad (3.41)$$

Um aus dieser Beschränktheit des optimalen Gütefunktionswertes  $\hat{J}_{K|N}^0$  auf die Existenz einer Lösung  $(\hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1})$  der Optimierungsaufgabe (3.38) mit finitem  $\hat{\mathbf{x}}_{K-N}$  schließen zu können, wird die Eigenschaft Beobachtbarkeit des Anfangszustandes benötigt. Unter Verwendung der Annahme A3 folgen aus (3.41) die Konvergenzresultate

$$\lim_{K \rightarrow \infty} \hat{J}_{K|N}^0 = 0 \quad (3.42a)$$

$$\lim_{K \rightarrow \infty} \hat{\mathbf{w}}_{K-k} = \mathbf{0} \quad \forall k = 1, \dots, N \quad (3.42b)$$

$$\begin{aligned} \lim_{K \rightarrow \infty} \hat{\mathbf{v}}_{K-k} &= \lim_{K \rightarrow \infty} (\mathbf{y}_{K-k} - \mathbf{h}_{K-k}(\check{\mathbf{x}}_{K-k}(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-k-1}))) \\ &= \mathbf{0} \quad \forall k = 1, \dots, N, \end{aligned} \quad (3.42c)$$

wobei in (3.42c) wieder (3.5) verwendet wurde. Weil gemäß der Annahme A3 für einen finiten Wert  $N$  auch

$$\lim_{K \rightarrow \infty} \mathbf{w}_{K-k} = \mathbf{0} \quad \forall k = 1, \dots, N \quad (3.43a)$$

$$\lim_{K \rightarrow \infty} \mathbf{v}_{K-k} = \lim_{K \rightarrow \infty} (\mathbf{y}_{K-k} - \mathbf{h}_{K-k}(\mathbf{x}_{K-k})) = \mathbf{0} \quad \forall k = 1, \dots, N \quad (3.43b)$$

gilt, erhält man

$$\lim_{K \rightarrow \infty} (\hat{\mathbf{w}}_{K-k} - \mathbf{w}_{K-k}) = \mathbf{0} \quad \forall k = 1, \dots, N \quad (3.44a)$$

$$\begin{aligned} \lim_{K \rightarrow \infty} (\mathbf{h}_{K-k}(\check{\mathbf{x}}_{K-k}(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-k-1})) - \mathbf{h}_{K-k}(\mathbf{x}_{K-k})) \\ = \mathbf{0} \quad \forall k = 1, \dots, N. \end{aligned} \quad (3.44b)$$

Da das System die Eigenschaft Beobachtbarkeit des Anfangszustandes besitzt und diese die Beobachtbarkeit des Endzustandes impliziert, ist die Existenz von

Funktionen  $\bar{\gamma}_1(\cdot), \bar{\gamma}_2(\cdot) \in \mathcal{K}$  und einer Zahl  $\underline{N} \in \mathbb{N}$  gesichert, so dass

$$\begin{aligned} & \|\check{\mathbf{x}}_K(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) - \mathbf{x}_K\|_2 \\ & \leq \bar{\gamma}_1\left(\max_{k=K-N, \dots, K-1} \{\|\hat{\mathbf{w}}_k - \mathbf{w}_k\|_2\}\right) \\ & \quad + \bar{\gamma}_2\left(\max_{k=K-N, \dots, K-1} \{\|\mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|k-1})) - \mathbf{h}_k(\mathbf{x}_k)\|_2\}\right) \end{aligned} \quad (3.45)$$

für  $\forall K, N \in \mathbb{N}_0$  mit  $K \geq N \geq \underline{N}$  und finitem  $N$  gilt. Aus (3.44) und (3.45) folgt für finites  $N \geq \underline{N}$

$$\lim_{K \rightarrow \infty} (\check{\mathbf{x}}_K(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) - \mathbf{x}_K) = \mathbf{0}, \quad (3.46)$$

womit die in (3.13) geforderte Konvergenz des Schätzfehlers gezeigt ist.

Wählt man nun ein festes  $\delta \geq 0$  und fordert  $\bar{J} = \delta$  für  $\bar{J}$  aus (3.16) bzw. (3.41), so folgt in einer zum Beweis des Satzes 3.1 analogen Weise, dass

$$\|\hat{\mathbf{w}}_k - \mathbf{w}_k\|_2 \leq \bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta) \quad \forall k = K-N, \dots, K-1 \quad (3.47a)$$

$$\begin{aligned} \|\mathbf{h}_k(\check{\mathbf{x}}_k(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|k-1})) - \mathbf{h}_k(\mathbf{x}_k)\|_2 & \leq \bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta) \\ & \quad \forall k = K-N, \dots, K-1 \end{aligned} \quad (3.47b)$$

gilt. Einsetzen der Abschätzung (3.47) in (3.45) liefert

$$\begin{aligned} & \|\check{\mathbf{x}}_K(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) - \mathbf{x}_K\|_2 \\ & \leq \bar{\gamma}_1(\bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta)) + \bar{\gamma}_2(\bar{\gamma}_b^{-1}(\delta) + \underline{\gamma}_b^{-1}(\delta)) \end{aligned} \quad (3.48)$$

für  $\forall K, N \in \mathbb{N}_0$  mit  $K \geq N \geq \underline{N}$  und finitem  $N$ . Die gesamte rechte Seite von (3.48) ist eine  $\mathcal{K}$ -Funktion in  $\delta$  und daher nach  $\delta$  auflösbar. Es existiert also zu jedem gegebenen  $\varepsilon \geq 0$  ein  $\delta(\varepsilon) \geq 0$  wobei  $\delta(0) = 0$ , so dass

$$\|\check{\mathbf{x}}_K(K-N, \hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) - \mathbf{x}_K\|_2 \leq \varepsilon \quad (3.49)$$

für  $\forall K, N \in \mathbb{N}_0$  mit  $K \geq N \geq \underline{N}$  und finitem  $N$ . Damit ist auch die Existenz einer Beziehung  $\delta(\varepsilon)$ , die die Erfüllung der Implikation (3.12) und somit die Beschränktheit des Schätzfehlers sichert, gezeigt.  $\square$

Nachteilig beim MHE Zustandsbeobachter ohne Anfangskostenterm ist, dass die Strecke beobachtbar gemäß Definition 3.5 sein muss und dass dies eine möglicherweise große Mindesthorizontlänge  $\underline{N}$  erfordert. Nachfolgend werden daher Bedingungen für einen von Null verschiedenen Anfangskostenterm formuliert, so dass diese Beobachtbarkeitsbedingung nicht mehr erforderlich ist.

### 3.4.3 Approximation der Ankunfts-kosten

Der vorliegende Abschnitt beschreibt eine MHE Formulierung mit einem Anfangskostenterm  $B_k(\cdot)$ , der selbst aus der Betrachtung eines finiten mitbewegten Horizonts hervorgeht.

Es wird versucht, die guten Stabilitäts- und Konvergenzeigenschaften der Zustandsschätzung mit vollständiger Information (siehe die Abschnitte 3.3 und 3.4.1) zu erreichen und trotzdem den Rechenaufwand in einem vertretbaren Bereich zu halten.

Es wird zunächst in Anlehnung an den Ankunfts-kostenterm (3.35) für vollständige Information, der sogenannte *MHE Ankunfts-kostenterm*

$$\hat{J}_{k|N}^B(\mathbf{x}) = \min_{\substack{(\tilde{\mathbf{x}}_{k-N}, \\ \tilde{\mathbf{w}}_{k-N|k-1})}} J_{k|N}(\tilde{\mathbf{x}}_{k-N}, \tilde{\mathbf{w}}_{k-N|k-1}) \quad (3.50a)$$

$$\text{u.B.v. } \tilde{\mathbf{x}}_{l+1} = \mathbf{f}_l(\tilde{\mathbf{x}}_l, \tilde{\mathbf{w}}_l) \quad \forall l = k - N, \dots, k - 1 \quad (3.50b)$$

$$\tilde{\mathbf{v}}_l = \mathbf{y}_l - \mathbf{h}_l(\tilde{\mathbf{x}}_l) \quad \forall l = k - N, \dots, k - 1 \quad (3.50c)$$

$$\tilde{\mathbf{x}}_l \in X_l \quad \forall l = k - N, \dots, k - 1 \quad (3.50d)$$

$$\tilde{\mathbf{w}}_l \in W_l, \quad \tilde{\mathbf{v}}_l \in V_l, \quad \forall l = k - N, \dots, k - 1 \quad (3.50e)$$

$$\tilde{\mathbf{x}}_k = \mathbf{x} \quad (3.50f)$$

definiert, der den finiten Horizont  $k - N, \dots, k$  mit  $k > N$  abdeckt. Für die Gütefunktion gilt in üblicher Art

$$J_{k|N}(\tilde{\mathbf{x}}_{k-N}, \tilde{\mathbf{w}}_{k-N|k-1}) = B_{k-N}(\tilde{\mathbf{x}}_{k-N}) + \sum_{l=k-N}^{k-1} b_l(\tilde{\mathbf{w}}_l, \tilde{\mathbf{v}}_l). \quad (3.51)$$

Für den Fall  $k \leq N$  wird  $\hat{J}_{k|N}^B(\cdot) = \hat{J}_{k|k}^B(\cdot)$  mit  $\hat{J}_{k|k}^B(\cdot)$  gemäß (3.35) verwendet.

Mit rekursiven Definitionen  $B_k(\cdot) = \hat{J}_{k|N}^B(\cdot)$ ,  $B_{k-N}(\cdot) = \hat{J}_{k-N|N}^B(\cdot)$ , ... würde man also wieder den *Ankunfts-kostenterm für vollständige Information* gemäß (3.35) erhalten. Es stellt sich daher die Frage, wie die Funktion  $\hat{J}_{k|N}^B(\cdot)$  sinnvoll durch  $B_k(\cdot)$  approximiert werden kann. Es kann gezeigt werden (siehe [3.2]), dass für Stabilität und Konvergenz des MHE Schätzfehlers die Einhaltung der folgenden Ungleichungen von Bedeutung ist:

$$\hat{J}_{k|N}^B(\mathbf{x}) \geq B_k(\mathbf{x}) \geq \hat{J}_{k|N} + \underline{\gamma}_B(\|\mathbf{x} - \bar{\mathbf{x}}_k\|_2) \quad \forall \mathbf{x} \in X_k, k > N. \quad (3.52)$$

Hierbei ist  $\underline{\gamma}_B(\cdot) \in \mathcal{K}_\infty$  die bereits in Annahme A5 verwendete Vergleichsfunktion und

$$\bar{\mathbf{x}}_k = \arg \min_{\mathbf{x} \in X_k} \hat{J}_{k|N}^B(\mathbf{x}) = \bar{\mathbf{x}}_k(k - N, \hat{\mathbf{x}}_{k-N}, \hat{\mathbf{w}}_{k-N|k-1}) \quad (3.53)$$

die A-priori-Schätzung gemäß Annahme A4, wobei  $(\hat{\mathbf{x}}_{k-N}, \hat{\mathbf{w}}_{k-N|k-1})$  natürlich einfach die Lösung der ursprünglichen Optimierungsaufgabe (3.6) (ohne die Endgleichungsbeschränkung (3.50f)) ist und

$$\hat{J}_{k|N} = B_k(\bar{\mathbf{x}}_k) = \min_{\mathbf{x} \in X_k} \hat{J}_{k|N}^B(\mathbf{x}) \quad (3.54)$$

der zugehörige optimale Gütefunktionswert. Für einen skalaren Zustand  $x$  ist in Abbildung 3.2 ein Beispiel für  $B_k(\cdot)$  dargestellt, welches (3.52) bis (3.54) erfüllt.

Eine Voraussetzung für die Konvergenz der MHE Zustandsschätzung ist die im folgenden Satz postulierte Beschränktheit des Anfangskostenterms.

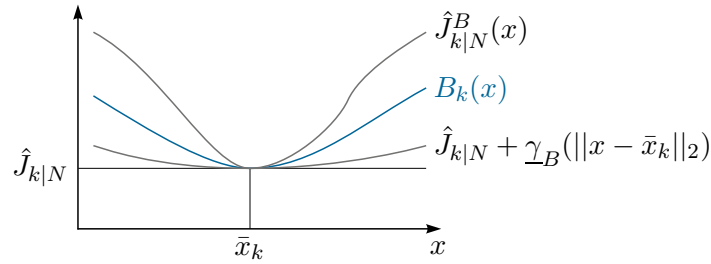


Abbildung 3.2: Beispiel für einen möglichen Anfangskostenterm  $B_k(x)$ , der (3.52) und (3.54) erfüllt.

**Satz 3.3** (Beschränktheit der MHE Gütefunktionswerte). Wird die Ungleichung (3.52) eingehalten, so gilt

$$\hat{J}_{k|N}^B(\mathbf{x}) \leq \hat{J}_{k|k}^B(\mathbf{x}), \quad \hat{J}_{k|N} \leq \hat{J}_{k|k}, \quad \forall \mathbf{x} \in X_k, k \geq 0. \quad (3.55)$$

**Aufgabe 3.2.** Beweisen Sie Satz 3.3. Dies gelingt besonders einfach mittels vollständiger Induktion.

**Definition 3.7** (MHE Detektierbarkeit). Das System (3.1) ist *MHE detektierbar*, wenn das erweiterte Modell

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_{1,k}) + \mathbf{w}_{2,k} \quad \forall k \in \mathbb{N}_0 \quad (3.56a)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \quad \forall k \in \mathbb{N}_0 \quad (3.56b)$$

mit der erweiterten Prozessstörung  $\mathbf{w}_k = [\mathbf{w}_{1,k}^\top, \mathbf{w}_{2,k}^\top]^\top$  die IIOSS-Eigenschaft gemäß Definition 3.4 besitzt.

**Bemerkung 3.2.** Die Eigenschaft *MHE Detektierbarkeit* ist geringfügig restriktiver als die Eigenschaft IIOSS aber weniger restriktiv als die Eigenschaft Beobachtbarkeit des Endzustands gemäß Definition 3.6. Bei vielen IIOSS Systemen tritt bereits die ursprüngliche Störung  $\mathbf{w}_k$  in (3.1a) nur additiv auf. Diese Systeme sind natürlich automatisch MHE detektierbar.

**Satz 3.4** (RGAS des MHE Zustandsbeobachters mit Approximation der MHE Ankunfts-kosten). Für ein System (3.1), das MHE detektierbar im Sinne der Definition 3.7 ist, mit Störungen, die der Annahme A3 genügen, führt der MHE Zustandsschätzer (3.6) mit festem  $N$  sowie einem Anfangskostenterm  $B_k$ , der die Ungleichung (3.52) erfüllt (Schätzung mit Approximation der MHE Ankunfts-kosten), und einer Kostenfunktion  $b_k$  gemäß der Annahme A5 zu einem robust global asymptotisch stabilen Schätzfehler.

Der Beweis dieses Satzes ist in [3.2] zu finden. Mit der Ungleichung (3.52) kennt man nun Bedingungen, die der Anfangskostenterm  $B_k(\cdot)$  erfüllen muss, um einen RGAS Beobachter

zu erhalten. Der tatsächliche Entwurf der Funktion  $B_k(\cdot)$  bleibt aber eine anspruchsvolle Aufgabe [3.2].

### 3.5 Maximum-a-posteriori Zustandsschätzung

Wird das System (3.1) als stochastischer Prozess aufgefasst, so liefert die Wahrscheinlichkeitstheorie konkrete Hinweise auf eine sinnvolle Wahl der Funktionen  $B_k(\cdot)$  und  $b_k(\cdot, \cdot)$ . Sind bestimmte Wahrscheinlichkeitsdichtefunktionen bekannt, so kann mit dem *Maximum-a-posteriori (MAP) Schätzer* sogar ein optimaler Beobachter konstruiert werden. Die nachfolgenden Ausführungen basieren zum Teil auf [3.7, 3.14]. Einige Grundlagen der Stochastik wurden auch bereits in der Vorlesung *Regelungssysteme 1* [3.10] besprochen. Zunächst wird der Begriff des Modalwerts definiert:

**Definition 3.8** (Modalwert einer Wahrscheinlichkeitsdichtefunktion). Für eine stetige Wahrscheinlichkeitsdichtefunktion  $P_{\mathbf{z}}(\mathbf{z})$  ist der Modalwert  $\hat{\mathbf{z}}$  definiert als

$$\hat{\mathbf{z}} = \arg \max_{\mathbf{z}} P_{\mathbf{z}}(\mathbf{z}) . \quad (3.57)$$

Bei der Realisierung einer Zufallsvariable  $\mathbf{z}$  tritt also der Modalwert  $\hat{\mathbf{z}}$  mit der höchsten Wahrscheinlichkeit auf. Man beachte, dass (3.57) immer eine Lösung besitzt, diese muss aber nicht eindeutig sein.

Gemäß dem Modell (3.1) liegen statistische Zusammenhänge zwischen den Messgrößen  $\mathbf{y}_{0|K-1}$  und den unbekanntenen Größen  $\mathbf{x}_0$  und  $\mathbf{w}_{0|K-1}$  vor. Für das Schätzproblem mit vollständiger Information können diese Zusammenhänge durch die *bedingte Wahrscheinlichkeitsdichte*

$$P(\mathbf{x}_0, \mathbf{w}_{0|K-1} | \mathbf{y}_{0|K-1}) \quad (3.58)$$

beschrieben werden.

**Definition 3.9** (Maximum-a-posteriori (MAP) Zustandsschätzer). Ein MAP Zustandsschätzer für das System (3.1) liefert für gegebene Messwerte  $\mathbf{y}_{0|K-1}$ , d. h. bei Berücksichtigung der vollständigen Information, als Schätzwert für die unbekanntenen Größen  $\mathbf{x}_0$  und  $\mathbf{w}_{0|K-1}$  zum Zeitpunkt  $K \in \mathbb{N}_0$  den Modalwert

$$(\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1}) = \arg \max_{(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|K-1})} P(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|K-1} | \mathbf{y}_{0|K-1}) . \quad (3.59)$$

Diese Optimierungsaufgabe kann um die Beschränkungen (3.2) erweitert werden. Derartige Beschränkungen können aber auch bereits in der Formulierung von (3.58) berücksichtigt werden, indem die Wahrscheinlichkeitsdichte für nicht zulässige Werte auf Null gesetzt wird. Der Einfachheit halber wird in diesem Abschnitt auf Beschränkungen der Art (3.2) verzichtet.

Im nächsten Schritt soll die bedingte Wahrscheinlichkeitsdichte (3.58) konkreter formuliert werden. Da es sich bei  $\mathbf{x}_0$  und  $\mathbf{w}_{0|K-1}$  um stochastisch unabhängige Zufallsvariablen handelt, gilt

$$P(\mathbf{x}_0, \mathbf{w}_{0|K-1}) = P_{\mathbf{x}_0}(\mathbf{x}_0) \prod_{k=0}^{K-1} P_{\mathbf{w}_k}(\mathbf{w}_k) . \quad (3.60)$$

Berücksichtigt man (3.1) und die stochastische Unabhängigkeit der Zufallsvariablen  $\mathbf{v}_{0|K-1}$ , so folgt für die bedingte Wahrscheinlichkeit der Messgrößen

$$P(\mathbf{y}_{0|K-1}|\mathbf{x}_0, \mathbf{w}_{0|K-1}) = \prod_{k=0}^{K-1} P_{\mathbf{v}_k}(\mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k(0, \mathbf{x}_0, \mathbf{w}_{0|k-1}))) = \prod_{k=0}^{K-1} P_{\mathbf{v}_k}(\mathbf{v}_k). \quad (3.61)$$

Mit dem Satz von Bayes folgt aus (3.60) und (3.61) die bedingte Wahrscheinlichkeitsdichte

$$\begin{aligned} P(\mathbf{x}_0, \mathbf{w}_{0|K-1}|\mathbf{y}_{0|K-1}) &= \frac{P(\mathbf{y}_{0|K-1}|\mathbf{x}_0, \mathbf{w}_{0|K-1})P(\mathbf{x}_0, \mathbf{w}_{0|K-1})}{P(\mathbf{y}_{0|K-1})} \\ &= \frac{P_{\mathbf{x}_0}(\mathbf{x}_0) \prod_{k=0}^{K-1} P_{\mathbf{w}_k}(\mathbf{w}_k) P_{\mathbf{v}_k}(\mathbf{v}_k)}{P(\mathbf{y}_{0|K-1})}, \end{aligned} \quad (3.62)$$

welche direkt in das Schätzgesetz (3.59) eingesetzt werden könnte. Berücksichtigt man (3.1b), die Monotonizität der Logarithmusfunktion  $\ln(\cdot)$  und dass für gegebene Messwerte  $\mathbf{y}_{0|K-1}$  der Nenner von (3.62) eine feste Zahl ist, so kann die vereinfachte Optimierungsaufgabe

$$(\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_{0|K-1}) = \arg \min_{(\tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|K-1})} \left( B_0(\tilde{\mathbf{x}}_0) + \sum_{k=0}^{K-1} b_k(\tilde{\mathbf{w}}_k, \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k(0, \tilde{\mathbf{x}}_0, \tilde{\mathbf{w}}_{0|k-1}))) \right) \quad (3.63)$$

mit den Kostenfunktionen

$$B_0(\tilde{\mathbf{x}}_0) = -\ln(P_{\mathbf{x}_0}(\tilde{\mathbf{x}}_0)) \quad (3.64a)$$

$$b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) = -\ln(P_{\mathbf{w}_k}(\tilde{\mathbf{w}}_k)) - \ln(P_{\mathbf{v}_k}(\tilde{\mathbf{v}}_k)) \quad (3.64b)$$

formuliert werden. Offensichtlich liefert sie das gleiche Ergebnis wie (3.59).

Da (3.63) der ursprünglichen Optimierungsaufgabe (3.6) (ohne Beschränkungen) mit  $N = K$  für die Zustandsschätzung mit vollständiger Information entspricht, ist mit (3.64) die aus wahrscheinlichkeitstheoretischer Sicht optimale Formulierung der Kostenfunktionen  $B_0(\cdot)$  und  $b_k(\cdot, \cdot)$  gefunden.

Analog zu Definition 3.9 kann natürlich sofort ein MHE MAP Zustandsschätzer formuliert werden. Es tritt dann statt (3.58) die bedingte Wahrscheinlichkeitsdichte  $P(\mathbf{x}_{K-N}, \mathbf{w}_{K-N|K-1}|\mathbf{y}_{K-N|K-1})$  mit  $K > N$  und festem  $N$  auf. Bei der Formulierung der Wahrscheinlichkeitsdichte von  $\mathbf{x}_{K-N}$  sind bei Verwertung der vollständigen Information auch die Messwerte  $\mathbf{y}_{0|K-N-1}$  zu berücksichtigen. In diesem Fall lautet die der Optimierungsaufgabe zugrunde liegende bedingte Wahrscheinlichkeitsdichte also

$$\begin{aligned} P(\mathbf{x}_{K-N}, \mathbf{w}_{K-N|K-1}|\mathbf{y}_{0|K-1}) \\ = P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})P(\mathbf{w}_{K-N|K-1}|\mathbf{x}_{K-N}, \mathbf{y}_{K-N|K-1}). \end{aligned} \quad (3.65)$$

Die hier vorgenommene Aufspaltung in zwei Faktoren ist wegen der Markov-Eigenschaft des stochastischen Prozesses zulässig.

Die Wahrscheinlichkeitsdichte  $P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})$  modelliert alle Informationen, die vor dem Zeitpunkt  $K-N$  gesammelt wurden. Sie kann gemäß dem Satz der totalen Wahrscheinlichkeit durch Integration der Wahrscheinlichkeitsdichte  $P(\mathbf{x}_0, \mathbf{w}_{0|K-N-1}|\mathbf{y}_{0|K-N-1})$  über die Variablen  $\mathbf{x}_0, \mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{K-N-1}$  unter Berücksichtigung der Bedingung  $\mathbf{x}_{K-N} =$

$\check{\mathbf{x}}_{K-N}(0, \mathbf{x}_0, \mathbf{w}_{0|K-N-1})$  berechnet werden. Natürlich korrespondiert  $P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})$  mit dem Ankunfts-kostenterm (3.35) für vollständige Information. In vielen praktischen Fällen wird man aufgrund des Rechenaufwands  $P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})$  durch eine leichter berechenbare Approximation ersetzen. Wird z. B. von einer gleichverteilten Zufallsvariable  $\mathbf{x}_{K-N}$  ohne Berücksichtigung der Messwerte  $\mathbf{y}_{0|K-N-1}$  ausgegangen, so ist  $P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})$  konstant bezüglich all seiner Argumente und man erhält eine MHE Zustandsschätzung ohne Anfangskostenterm (vgl. Abschnitt 3.4.2). In [3.15] werden verschiedene weitere Möglichkeiten zur Approximation von  $P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1})$  mittels Filter (extended Kalman-Filter, unscented Kalman-Filter, Partikelfilter, etc.), die entlang der mit MHE geschätzten Folge  $\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_{K-N}$  mitgerechnet werden, vorgestellt.

Fasst man nun für gegebene, feste Messwerte  $\mathbf{y}_{0|K-N-1}$  den Zustand  $\mathbf{x}_{K-N}$  als Zufallsvariable mit der Wahrscheinlichkeitsdichte

$$P_{\mathbf{x}_{K-N}}(\mathbf{x}_{K-N}) = P(\mathbf{x}_{K-N}|\mathbf{y}_{0|K-N-1}) \quad (3.66)$$

auf, so erhält man in völlig analoger Weise zur Herleitung von (3.63)

$$\begin{aligned} & (\hat{\mathbf{x}}_{K-N}, \hat{\mathbf{w}}_{K-N|K-1}) \\ &= \arg \min_{\substack{(\tilde{\mathbf{x}}_{K-N}, \\ \tilde{\mathbf{w}}_{K-N|K-1})}} \left( B_{K-N}(\tilde{\mathbf{x}}_{K-N}) \right. \\ & \quad \left. + \sum_{k=K-N}^{K-1} b_k(\tilde{\mathbf{w}}_k, \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}_k(K-N), \tilde{\mathbf{x}}_{K-N}, \tilde{\mathbf{w}}_{K-N|k-1})) \right) \end{aligned} \quad (3.67)$$

mit den aus wahrscheinlichkeitstheoretischer Sicht optimalen Kostenfunktionen

$$B_k(\tilde{\mathbf{x}}_k) = -\ln(P_{\mathbf{x}_k}(\tilde{\mathbf{x}}_k)) = -\ln(P(\tilde{\mathbf{x}}_k|\mathbf{y}_{0|k-1})) \quad (3.68a)$$

$$b_k(\tilde{\mathbf{w}}_k, \tilde{\mathbf{v}}_k) = -\ln(P_{\mathbf{w}_k}(\tilde{\mathbf{w}}_k)) - \ln(P_{\mathbf{v}_k}(\tilde{\mathbf{v}}_k)) . \quad (3.68b)$$

### 3.6 Zustands- und Parameterschätzung

In vielen Aufgabenstellungen sollen nicht nur die Systemzustände  $\mathbf{x}_k$  sondern auch unbekannte Systemparameter  $\boldsymbol{\rho} \in R \subseteq \mathbb{R}^r$  geschätzt werden [3.3, 3.16], wobei  $R$  die Menge der zulässigen Systemparameter ist. Nachfolgend wird skizziert, wie dies mit den bisher in diesem Abschnitt beschriebenen Verfahren einfach möglich ist.

Ist bekannt, dass die Parameter  $\boldsymbol{\rho}$  zumindest innerhalb des Schätzhorizonts  $K-N, \dots, K$  konstant sind, so kann das Modell (3.1) in der Form

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k, \boldsymbol{\rho}) \quad \forall k \in \mathbb{N}_0 \quad (3.69a)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\rho}) + \mathbf{v}_k \quad \forall k \in \mathbb{N}_0 \quad (3.69b)$$

angeschrieben werden. Zusätzlich zu  $\tilde{\mathbf{x}}_{K-N}$  und  $\tilde{\mathbf{w}}_{K-N|K-1}$  wird dann in der Optimierungsaufgabe  $\tilde{\boldsymbol{\rho}} \in R$  als (beschränkte) Optimierungsvariable verwendet. Die Kostenfunktionen  $B_k$  und  $b_k$  sind entsprechend anzupassen. Mit jedem Aufruf des MHE Schätzers wird ein neuer Schätzwert  $\hat{\boldsymbol{\rho}}$  für  $\boldsymbol{\rho}$  ermittelt.

Ist hingegen bekannt, dass die Parameter  $\boldsymbol{\rho}$  einer (zumeist langsamen) Dynamik unterliegen und ist die Länge  $N$  des Schätzhorizonts so groß, dass dynamische Änderungen von  $\boldsymbol{\rho}$  relevant sein können, so kann  $\boldsymbol{\rho}_k \in R_k \subseteq \mathbb{R}^r$  als weiterer Systemzustand aufgefasst werden und das Modell (3.69) in die Form

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k, \boldsymbol{\rho}_k) \quad \forall k \in \mathbb{N}_0 \quad (3.70a)$$

$$\boldsymbol{\rho}_{k+1} = \mathbf{g}_k(\mathbf{x}_k, \mathbf{w}_k, \boldsymbol{\rho}_k) \quad \forall k \in \mathbb{N}_0 \quad (3.70b)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\rho}_k) + \mathbf{v}_k \quad \forall k \in \mathbb{N}_0 \quad (3.70c)$$

umgeschrieben werden. Die Funktion  $\mathbf{g}_k$  ist nun entweder ein bekanntes Modell der Parameteränderungen oder, falls ein solches nicht vorhanden ist, ein *random-walk Modell* der Form

$$\mathbf{g}_k(\mathbf{x}_k, \mathbf{w}_k, \boldsymbol{\rho}_k) = \boldsymbol{\rho}_k + \mathbf{w}_k^\rho \quad (3.71)$$

mit zufälligen Störungen  $\mathbf{w}_k^\rho$ , die Teil der Prozessstörung  $\mathbf{w}_k$  sind. Je nach Art der Parameteränderungen kann  $\mathbf{w}_k^\rho$  stark beschränkt oder in der Kostenfunktion  $b_k$  stark bestraft werden.



### 3.7 Literatur

- [3.1] C. Rao, „Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems,“ Diss., University of Wisconsin-Madison, Wisconsin, 2000.
- [3.2] J.B. Rawlings und D.Q. Mayne, *Model Predictive Control: Theory and Design*. Madison, Wisconsin: Nob Hill Publishing, 2009.
- [3.3] N. Haverbeke, „Efficient Numerical Methods for Moving Horizon Estimation,“ Diss., Katholieke Universiteit Leuven, Heverlee, Belgium, 2011.
- [3.4] P. Philipp, „Centralized and Distributed Moving Horizon Strategies for State Estimation of Networked Control Systems,“ Diss., Technische Universität München, München, 2014.
- [3.5] F. Allgöwer, T. Badgwell, J. Qin, J. Rawlings und S. Wright, „Nonlinear Predictive Control and Moving Horizon Estimation: An Introductory Overview,“ in *Advances in Control*, Springer London, 1999, S. 391–449.
- [3.6] M. Alamir, „Nonlinear Moving Horizon Observers: Theory and Real-Time Implementation,“ in *Nonlinear Observers and Applications*, Ser. Lecture Notes in Control and Information Sciences, Bd. 363, Berlin Heidelberg: Springer, 2007, S. 139–179.
- [3.7] E. Haseltine und J. Rawlings, „Critical evaluation of extended Kalman filtering and moving-horizon estimation,“ *Industrial and Engineering Chemistry Research*, Jg. 44, Nr. 8, S. 2451–2460, 2005.
- [3.8] C. Rao, J. Rawlings und D. Mayne, „Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations,“ *IEEE Transactions on Automatic Control*, Jg. 48, Nr. 2, S. 246–258, Feb. 2003.
- [3.9] D. Simon, „Kalman filtering with state constraints: a survey of linear and nonlinear algorithms,“ *Control Theory Applications, IET*, Jg. 4, Nr. 8, S. 1303–1318, 2010.
- [3.10] W. Kemmetmüller, *Skriptum zur VO Regelungssysteme 1 (WS 2018/2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2018. Adresse: <https://www.acin.tuwien.ac.at/master/regelungssysteme-1/>.
- [3.11] J. Rawlings und L. Ji, „Optimization-based state estimation: Current status and some new results,“ *Journal of Process Control*, Jg. 22, Nr. 8, S. 1439–1444, 2012.
- [3.12] E. Sontag, „Comments on integral variants of ISS,“ *Systems and Control Letters*, Jg. 34, S. 93–100, 1998.
- [3.13] E. Sontag und Y. Wang, „Output-to-state stability and detectability of nonlinear systems,“ *Systems and Control Letters*, Jg. 29, S. 279–290, 1997.
- [3.14] C. Rao und J. Rawlings, „Nonlinear Moving Horizon State Estimation,“ in *Nonlinear Model Predictive Control*, Ser. Progress in Systems and Control Theory, Bd. 26, Birkhäuser Basel, 2000, S. 45–69.

- 
- [3.15] S. Ungarala, „Computing arrival cost parameters in moving horizon estimation using sampling based filters,“ *Journal of Process Control*, Jg. 19, Nr. 9, S. 1576–1588, 2009.
- [3.16] T. Johansen, „Introduction to Nonlinear Model Predictive Control and Moving Horizon Estimation,“ in *Selected Topics on Constrained and Nonlinear Control*, Bratislav, Trondheim: STU/NTNU, 2011, S. 1–53.

## 4 Pfadfolgeregelung

Die Aufgabenstellungen in der Regelungstechnik lassen sich unter anderem in *Arbeitspunktstabilisierung*, *Trajektorienfolgeregelung* und *Pfadfolgeregelung* einteilen. Während die ersten beiden Aufgabenstellungen bereits für sehr viele Systemklassen erforscht und systematisch gelöst sind, stellt die Pfadfolgeregelung ein vergleichsweise junges Forschungsfeld dar.

Das Ziel der Arbeitspunktstabilisierung ist die Stabilisierung eines konstanten oder eines (im Vergleich zur Systemdynamik) nur sehr langsam veränderlichen Arbeitspunktes. Im Gegensatz dazu beschäftigt sich die Trajektorienfolgeregelung mit der Stabilisierung einer explizit zeitabhängigen Solltrajektorie. Typischerweise werden als Regelgrößen die Ausgangsgrößen des Systems verwendet. Die Solltrajektorie gibt nicht nur einen geometrischen Verlauf im Ausgangsraum vor, sondern auch zu welchem Zeitpunkt sich die Ausgangsgrößen des Systems an welchem Punkt des geometrischen Verlaufs befinden sollen.

Im Gegensatz dazu geht es bei der Pfadfolgeregelung darum, dass a priori nur der geometrische Verlauf ohne zeitlicher Information (im Weiteren als Pfad bezeichnet) definiert ist. In diesem Kapitel soll stets der Fall betrachtet werden, dass der Pfad für die Ausgangsgrößen des Systems vorgegeben ist. Das primäre Ziel des Pfadfolgereglers besteht darin, den Pfad (asymptotisch) zu stabilisieren, d. h. die Ausgangsgrößen des Systems zum Pfad zu bewegen. Die genaue Stelle entlang des Pfades ist von vornherein nicht definiert. Darüber hinaus soll auch (sofern es die Systemstruktur erlaubt) die Bewegung entlang des Pfades gezielt beeinflusst werden können.

Durch diese Strategie kann ein Manko der Trajektorienfolgeregelung behoben werden. Um dies zu erläutern, wird ein System, das sich zu einem gewissen Zeitpunkt  $\tilde{t}$  genau auf der Solltrajektorie befindet, betrachtet. Selbst im idealen Fall ist bei der Trajektorienfolgeregelung nicht gesichert, dass das System auch auf der Solltrajektorie bleibt. Dazu muss nämlich auch die zeitliche Zuordnung stimmen, d. h. gemäß der Solltrajektorie muss sich das System zum Zeitpunkt  $\tilde{t}$  genau am richtigen Punkt aufhalten. Stimmt diese Zuordnung nicht, d. h. sollte sich das System zum Zeitpunkt  $\tilde{t}$  eigentlich an einem anderen Ort der Solltrajektorie befinden, dann wird diese im Allgemeinen verlassen und dann wieder asymptotisch angefahren. Dadurch, dass es bei der Pfadfolgeregelung a priori keine zeitliche Parametrierung des Pfades gibt, wird, falls das System genau auf dem Pfad startet (bzw. auf einer entsprechenden Teilmenge des Zustandsraumes, siehe Abschnitt 4.3), dieser auch nicht mehr verlassen. Diese Eigenschaft bezeichnet man als *Invarianzeigenschaft* eines Pfadfolgereglers, d. h. sobald sich das System einmal auf dem Pfad befindet, gilt dies auch in Zukunft (im nominellen, ungestörten Fall).

Eine typische Anwendung einer Pfadfolgeregelung stellen CNC<sup>1</sup>-Fräsmaschinen dar.

---

<sup>1</sup>Computerized Numerical Control.

Mit diesen können üblicherweise komplexe geometrische Formen gefräst werden, die z. B. durch Splines dargestellt sind. Diese definieren eine geometrische Kurve, der das Fräs Werkzeug folgen soll. Um eine geforderte Genauigkeit zu erreichen, ist es das primäre Ziel, der geometrischen Kurve (dem Pfad) möglichst gut zu folgen. Erst an zweiter Stelle steht die Geschwindigkeit dieser Verfolgung. Diese wird in diesem Zusammenhang als Vorschubgeschwindigkeit bezeichnet und soll natürlich auch möglichst genau eingehalten werden. Selbst eine ideal eingehaltene Vorschubgeschwindigkeit nützt aber nichts, wenn die Kontur des fertigen Werkstückes die geforderten Genauigkeitsanforderungen nicht erfüllt. Aus dieser Argumentation leitet sich die Priorisierung der Ziele ab. Darüber hinaus möchte man (im nominellen, ungestörten Fall) nicht, dass die ideale Kontur wieder verlassen wird, sobald sich das Werkzeug einmal auf ihr befindet. Dies soll auch gelten, wenn bspw. durch Inhomogenitäten im Material die aktuelle Vorschubgeschwindigkeit nicht exakt der Vorgabe entspricht. Bei einer Trajektorienfolgeregelung würde dies nach sich ziehen, dass die zeitliche Zuordnung entlang des Pfades nicht erfüllt ist, womit die Kontur wieder verlassen werden würde. Die Pfadfolgeregelung behebt diesen Nachteil und eignet sich damit besser für diese Anwendung.

Ein Pfadfolgeregler kann auf Grundlage von unterschiedlichsten Regelungsverfahren entwickelt werden. In diesem Kapitel werden Konzepte zur Pfadfolgeregelung basierend auf der exakten Eingangs-Ausgangs- und Eingangs-Zustandslinearisierung (vgl. [4.1]) sowie der modellprädiktiven Regelung vorgestellt. Es werden Systeme mit dem Zustand  $\mathbf{x} \in \mathbb{R}^n$  und Ausgangsgrößen der Form  $\mathbf{y} = \mathbf{h}(\mathbf{x}) \in \mathbb{R}^p$  mit  $p \geq 2$  betrachtet. Wie bereits erwähnt, soll die Pfadfolgeregelung stets nur im Ausgangsraum  $\mathbb{R}^p$  des Systems, d. h. für die Ausgangsgrößen  $\mathbf{y}$  des Systems, durchgeführt werden.

## 4.1 Definition von Pfaden

Als *Pfad*  $\mathcal{P}$  wird eine *Kurve im Ausgangsraum des Systems* bezeichnet, die entweder offen oder geschlossen sein kann, siehe Abbildung 4.1.

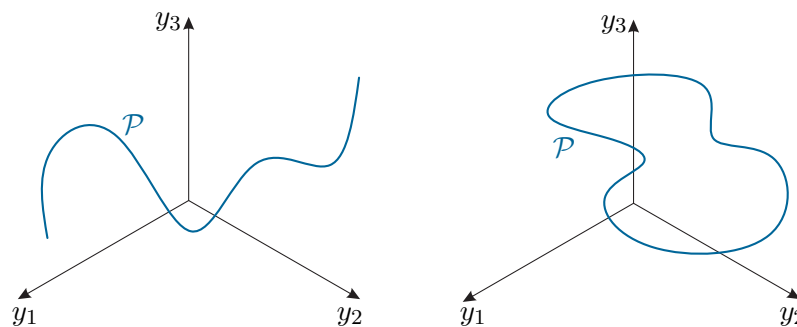


Abbildung 4.1: Offene und geschlossene Kurve im Ausgangsraum.

Dieser Pfad soll in weiterer Folge vom Pfadfolgeregler stabilisiert werden, d. h. die Ausgangsgrößen  $\mathbf{y}$  des Systems sollen sich zu der Kurve bewegen bzw. sich auf dieser aufhalten. Es gibt die Möglichkeiten der *parametrierten* und *impliziten* Definition von Pfaden, die im Folgenden näher beleuchtet werden sollen.

### 4.1.1 Parametrierte Darstellung von Pfaden

Bei der Parameterdarstellung eines Pfades wird die Kurve über einen reellen Kurven- oder *Pfadparameter*  $\theta$  dargestellt. Der Kurvenparameter ist Element einer Menge  $\mathcal{T}$ , die abgeschlossen (bspw.  $\mathcal{T} = [\theta_0, \theta_1]$ ) oder offen (bspw.  $\mathcal{T} = \mathbb{R}$ ) sein kann. Die Abbildung

$$\sigma(\theta) : \mathcal{T} \rightarrow \mathbb{R}^p \quad (4.1)$$

weist jedem Element aus  $\mathcal{T}$  explizit ein Element des Ausgangsraumes zu. Mit Hilfe von (4.1) kann nun der Pfad  $\mathcal{P}$  in der Form

$$\mathcal{P} = \{\bar{y} \in \mathbb{R}^p \mid \bar{y} = \sigma(\theta), \theta \in \mathcal{T}\} \quad (4.2)$$

definiert werden. In weiterer Folge wird angenommen, dass die Komponenten  $\sigma_i(\theta)$  stets genügend oft stetig differenzierbar sind und  $\mathcal{P}$  gemäß (4.2) eine reguläre Kurve [4.2] darstellt, d. h.

$$\sigma'(\bar{\theta}) = \frac{\partial \sigma}{\partial \theta}(\bar{\theta}) \neq \mathbf{0} \quad \forall \bar{\theta} \in \mathcal{T}. \quad (4.3)$$

Der Pfad  $\mathcal{P}$  gemäß (4.2) kann je nach Wahl von  $\sigma(\theta)$  und  $\mathcal{T}$  sowohl offen als auch geschlossen sein.

*Beispiel 4.1.* Als ein einfaches Beispiel betrachte man für  $p = 2$  einen elliptischen Pfad definiert durch

$$\sigma_e(\theta) = \begin{bmatrix} y_{10} + a \cos(\theta) \\ y_{20} + b \sin(\theta) \end{bmatrix}. \quad (4.4)$$

Wie in Abbildung 4.2 dargestellt, ist durch  $y_{10}$  und  $y_{20}$  der Versatz der Ellipse vom Koordinatenursprung gegeben. Die Parameter  $a > 0$  und  $b > 0$  stellen die Längen der großen und kleinen Halbachsen dar.

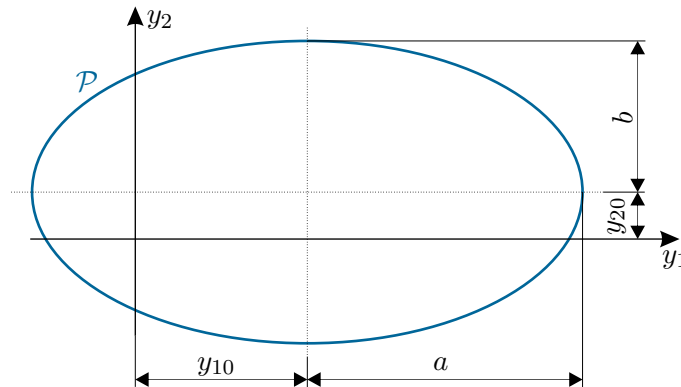


Abbildung 4.2: Elliptischer Pfad.

Der Definitionsbereich  $\mathcal{T}$  des Pfadparameters kann entweder durch  $\mathcal{T} = \mathbb{R}$  oder bspw. durch  $\mathcal{T} = [0, 2\pi)$  gegeben sein, wobei in beiden Fällen ein geschlossener Pfad resultiert. Für z. B.  $\mathcal{T} = [0, \pi]$  resultiert ein offener Pfad. Die partielle Ableitung folgt

zu

$$\frac{\partial \sigma_e}{\partial \theta} = \begin{bmatrix} -a \sin(\theta) \\ b \cos(\theta) \end{bmatrix}. \quad (4.5)$$

Da die beiden Komponenten von (4.5) für beliebige Werte von  $\theta$  nicht gleichzeitig verschwinden, definiert (4.4) eine reguläre Kurve.

### 4.1.2 Implizite Definition von Pfaden

Die implizite Definition eines Pfades basiert auf einem Gleichungssystem in Abhängigkeit der Ausgangsgrößen  $\mathbf{y}$  des Systems

$$\mathbf{s}(\mathbf{y}) = \mathbf{0} \quad (4.6)$$

mit einer stetigen Abbildung  $\mathbf{s} : \mathbb{R}^p \rightarrow \mathbb{R}^{p-1}$ . Der Pfad folgt als die Menge

$$\mathcal{P} = \{ \bar{\mathbf{y}} \in \mathbb{R}^p \mid \mathbf{s}(\bar{\mathbf{y}}) = \mathbf{0} \}, \quad (4.7)$$

wobei speziell im Hinblick auf die Pfadfolgeregelung gefordert wird, dass die Funktion  $\mathbf{s}$

$$\text{rang} \left( \frac{\partial \mathbf{s}}{\partial \mathbf{y}}(\bar{\mathbf{y}}) \right) = p - 1 \quad \forall \bar{\mathbf{y}} \in \mathcal{P} \quad (4.8)$$

erfüllt. Der Einfachheit halber sei weiters vorausgesetzt, dass die Komponenten  $s_i, i = 1, \dots, p - 1$  bis zur jeweils erforderlichen Ordnung stetig differenzierbar sind. Auch hier gilt, dass der Pfad  $\mathcal{P}$  gemäß (4.7) je nach Wahl von  $\mathbf{s}(\mathbf{y})$  sowohl offen als auch geschlossen sein kann.

*Beispiel 4.2.* Als Illustration dient wieder ein elliptischer Pfad für  $p = 2$  definiert durch

$$s_e(\mathbf{y}) = \frac{(y_1 - y_{10})^2}{a^2} + \frac{(y_2 - y_{20})^2}{b^2} - 1, \quad (4.9)$$

der in Abbildung 4.2 dargestellt ist. Für die partielle Ableitung folgt

$$\frac{\partial s_e}{\partial \mathbf{y}} = \begin{bmatrix} \frac{2(y_1 - y_{10})}{a^2} & \frac{2(y_2 - y_{20})}{b^2} \end{bmatrix}. \quad (4.10)$$

Da auf dem Pfad für  $y_1 = y_{10}$  gilt, dass  $y_2 \neq y_{20}$  ist (und vice versa), folgt, dass auf  $\mathcal{P}$  nie beide Komponenten von (4.10) simultan verschwinden, womit (4.8) erfüllt ist.

### 4.1.3 Vergleich und Umrechnung der Darstellungen

Die Parameterdarstellung und implizite Definition von Pfaden sind natürlich bis zu einem gewissen Grad äquivalent. Dennoch kann eine der beiden Darstellungen in manchen Situationen geeigneter sein als die andere. Als Beispiel dazu betrachte man einen 8-förmigen Pfad gegeben durch die so genannte Lemniskate von Geronon, siehe Abbildung 4.3. Diese ist definiert durch die Parameterdarstellung

$$\sigma_l(\theta) = \begin{bmatrix} \cos(\theta) \\ \cos(\theta) \sin(\theta) \end{bmatrix} \quad (4.11)$$

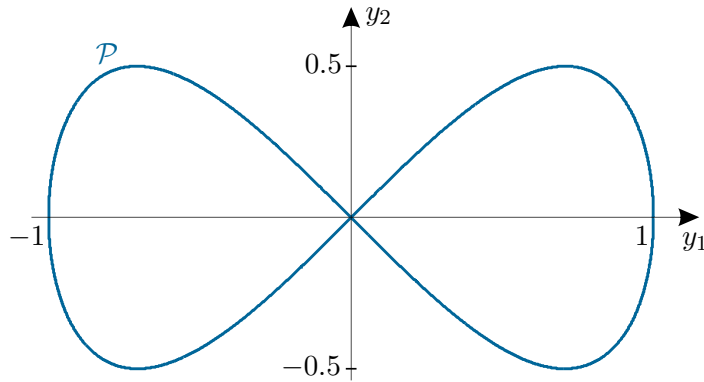


Abbildung 4.3: Lemniskate von Gerono.

mit  $\mathcal{T} = [0, 2\pi)$  bzw. die implizite Darstellung

$$s_l(\mathbf{y}) = y_1^4 - y_1^2 + y_2^2. \quad (4.12)$$

Die partiellen Ableitungen folgen als

$$\frac{\partial \sigma_l}{\partial \theta}(\theta) = \begin{bmatrix} -\sin(\theta) \\ \underbrace{-\sin^2(\theta) + \cos^2(\theta)}_{\cos(2\theta)} \end{bmatrix} \quad (4.13a)$$

$$\frac{\partial s_l}{\partial \mathbf{y}} = \begin{bmatrix} 4y_1^3 - 2y_1 & 2y_2 \end{bmatrix}. \quad (4.13b)$$

Offensichtlich ist für  $\mathbf{y} = \mathbf{0} \in \mathcal{P}$  die Bedingung (4.8) verletzt, womit sich (4.12) nicht für die Pfadfolgeregelung entlang der gesamten Lemniskate eignet. Diese Problematik tritt bei der Parameterdarstellung gemäß (4.11) nicht auf, weil die beiden Komponenten von  $\frac{\partial \sigma_l}{\partial \theta}(\theta)$  gemäß (4.13a) für alle  $\theta \in \mathcal{T}$  nie gleichzeitig verschwinden. Damit eignet sich die Parameterdarstellung von  $\mathcal{P}$  besser für die Pfadfolgeregelung entlang einer Lemniskate.

In vielen Fällen können die Parameterdarstellung und implizite Definition eines Pfades  $\mathcal{P}$  auch ineinander umgerechnet werden. Nachfolgend sind die beiden Richtungen erläutert.

1. Parameterdarstellung  $\rightarrow$  implizite Definition: Diese Umrechnung beruht auf der Elimination des Pfadparameters. Gemäß dem *Satz über implizite Funktionen* kann auf Grund von Bedingung (4.3) lokal stets  $\theta$  aus einer der Gleichungen  $\sigma_i(\theta) = y_i$ ,  $i \in \{1, \dots, p\}$  errechnet werden, für die  $\frac{\partial \sigma_i}{\partial \theta} \neq 0$  gilt. Daraus folgt  $\theta = \sigma_i^{-1}(y_i)$ , was eingesetzt in alle anderen Gleichungen  $\sigma_j(\theta) = y_j$ ,  $j \in \{1, 2, \dots, p\} \setminus i$  die implizite Definition

$$\sigma_j(\sigma_i^{-1}(y_i)) - y_j = 0, \quad j \in \{1, 2, \dots, p\} \setminus i \quad (4.14)$$

des Pfades in der Form (4.6) liefert.

2. implizite Definition  $\rightarrow$  Parameterdarstellung: Eine Möglichkeit besteht darin,  $p - 1$  Ausgangsgrößen aus den Gleichungen  $\mathbf{s}(\mathbf{y}) = \mathbf{0}$  in Abhängigkeit einer verbleibenden Ausgangsgröße  $y_i$ ,  $i \in \{1, \dots, p\}$  zu berechnen, was wiederum gemäß dem Satz über

implizite Funktionen und Bedingung (4.8) möglich ist. Damit erhält man  $\tilde{\mathbf{y}} = \mathbf{s}^{-1}(y_i)$ , wobei  $\tilde{\mathbf{y}} \in \mathbb{R}^{p-1}$  alle Ausgangsgrößen bis auf  $y_i$  enthält. Anschließend setzt man  $\theta = y_i$ , womit unmittelbar die Parameterdarstellung des Pfades in der Form (4.1)

$$\begin{bmatrix} \tilde{\mathbf{y}} \\ y_i \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{-1}(\theta) \\ \theta \end{bmatrix} \quad (4.15)$$

folgt.

**Beispiel 4.3.** Zur Illustration betrachte man wieder den elliptischen Pfad mit den Definitionen gemäß (4.4) bzw. (4.9). Der Einfachheit halber gelte  $y_{10} = y_{20} = 0$ .

1. Parameterdarstellung  $\rightarrow$  implizite Definition: Die Berechnung von  $\theta$  aus der ersten Gleichung von (4.4) liefert  $\theta = \arccos\left(\frac{y_1}{a}\right)$ , was eingesetzt in die zweite Gleichung von (4.4) die gesuchte implizite Darstellung

$$y_2 - b \sin(\theta) = y_2 - b \sin\left(\arccos\left(\frac{y_1}{a}\right)\right) = y_2 - b \sqrt{1 - \frac{y_1^2}{a^2}} = 0 \quad (4.16)$$

ergibt.

2. implizite Definition  $\rightarrow$  Parameterdarstellung: Ein möglicher Ansatz besteht darin, aus  $s_e(\mathbf{y}) = 0$  die Ausgangsgröße  $y_1$  in der Form

$$y_1 = \pm a \sqrt{1 - \frac{y_2^2}{b^2}} \quad (4.17)$$

zu berechnen. Über Setzen von  $\theta = y_2$  ergibt sich eine Parameterdarstellung zu

$$\boldsymbol{\sigma}_e(\theta) = \begin{bmatrix} \pm a \sqrt{1 - \frac{\theta^2}{b^2}} \\ \theta \end{bmatrix} \quad (4.18)$$

mit  $\mathcal{T} = [-b, b]$ . Durch Ausnutzen von  $\sqrt{1 - x^2} = \cos(\arcsin(x))$  erhält man

$$\boldsymbol{\sigma}_e(\theta) = \begin{bmatrix} \pm a \cos\left(\arcsin\left(\frac{\theta}{b}\right)\right) \\ \theta \end{bmatrix}. \quad (4.19)$$

Substituiert man  $\arcsin\left(\frac{\theta}{b}\right) = \tilde{\theta}$  folgt die Beziehung  $\theta = b \sin(\tilde{\theta})$  und damit die alternative Parameterdarstellung

$$\tilde{\boldsymbol{\sigma}}_e(\tilde{\theta}) = \begin{bmatrix} \pm a \cos(\tilde{\theta}) \\ b \sin(\tilde{\theta}) \end{bmatrix}, \quad (4.20)$$

die bis auf die Wahl des Vorzeichens identisch zu (4.4) ist.



## 4.2 Begriffe und Ziele der Pfadfolgeregelung

Für alle folgenden Ausführungen werden AI-Systeme (affine input-Systeme)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u} \quad (4.21a)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad (4.21b)$$

mit  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{u} \in \mathbb{R}^m$  und  $\mathbf{y} \in \mathbb{R}^p$  betrachtet. Es gelte  $m \geq p$  womit im Allgemeinen mehr Eingangs- als Ausgangsgrößen erlaubt sind. Für das System (4.21) sei wie in [4.1] vorausgesetzt, dass  $\mathbf{f}(\mathbf{x})$  und  $\mathbf{g}_j(\mathbf{x})$ , die die Spalten von  $\mathbf{G}(\mathbf{x}) \in \mathbb{R}^{n \times m}$  darstellen, glatte Vektorfelder sind. Weiters sind die Komponenten von  $\mathbf{h}(\mathbf{x})$  glatte Funktionen.

Für die Erläuterung der grundlegenden Konzepte und Ziele der Pfadfolgeregelung wird von einem Pfad  $\mathcal{P}$  im Ausgangsraum des Systems ausgegangen. Aus geometrischer Sicht kann jede Bewegung des Ausgangs  $\mathbf{y}$  des Systems in *einen transversalen und einen tangentialen* Teil im Hinblick auf den Pfad  $\mathcal{P}$  zerlegt werden. Der transversale Teil beschreibt den Anteil der Bewegung, der quer zum Pfad verläuft. Ist dieser Teil identisch null, befindet sich der Ausgang  $\mathbf{y}$  des Systems exakt am Pfad. Der tangentielle Teil der Bewegung ist jener, der in Pfadrichtung verläuft. Grob gesprochen kann er als Position des Ausgangs entlang des Pfades gedeutet werden.

Für die Formulierung der Ziele der Pfadfolgeregelung wird die Abbildung  $\|\mathbf{y}\|_{\mathcal{P}} : \mathbb{R}^p \rightarrow \mathbb{R}_{\geq 0}$  benötigt, die jedem Punkt  $\mathbf{y}$  im Ausgangsraum eine nichtnegative reelle Zahl zuordnet, die den kürzesten Abstand zum Pfad angibt, d. h.  $\|\mathbf{y}\|_{\mathcal{P}} = \min_{\bar{\mathbf{y}} \in \mathcal{P}} \|\bar{\mathbf{y}} - \mathbf{y}\|$ . Weiters wird noch folgende Definition benötigt.

**Definition 4.1** (Gesteuert positiv invariante Menge [4.3]). Eine Teilmenge  $\mathcal{M}$  des Zustandsraumes von (4.21) wird *gesteuert positiv invariant* genannt, falls es eine Abbildung  $\mathbf{k} : \mathcal{M} \rightarrow \mathbb{R}^m$  so gibt, dass  $\mathcal{M}$  eine positiv invariante Menge des autonomen Systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{k}(\mathbf{x}) \quad (4.22)$$

ist.

In der englischsprachigen Literatur wird eine gesteuert positiv invariante Menge als *controlled invariant set* bezeichnet. Sie stellt die Verallgemeinerung der positiv invarianten Menge gemäß [4.1] auf Systeme mit Eingangsgrößen dar.

Damit können die Ziele der Pfadfolgeregelung für einen Pfad  $\mathcal{P}$  im Ausgangsraum formuliert werden.

(Z1) *Asymptotische Konvergenz zu  $\mathcal{P}$* : Der Ausgang  $\mathbf{y}$  des Systems soll asymptotisch zum Pfad  $\mathcal{P}$  konvergieren. Das heißt,  $\|\mathbf{y}(t)\|_{\mathcal{P}} \rightarrow 0$  für  $t \rightarrow \infty$ .

(Z2) *Invarianzeigenschaft*: Befindet sich der Zustand des Systems  $\mathbf{x}(t_0)$  zu einem Zeitpunkt  $t_0$  in einer gesteuert positiv invarianten Teilmenge  $\Gamma^*$  von

$$\Gamma = \{\bar{\mathbf{x}} \in \mathbb{R}^n \mid \mathbf{h}(\bar{\mathbf{x}}) \in \mathcal{P}\}, \quad (4.23)$$

dann fordert die Invarianzeigenschaft, dass für alle  $t \geq t_0$  die Eigenschaft  $\|\mathbf{y}(t)\|_{\mathcal{P}} = 0$  gelten soll.

(Z3) *Tangentiale Bewegung*: Es soll eine applikationsspezifische Bewegung des Systems auf dem Pfad  $\mathcal{P}$  erzielt werden.

Zum Ziel (Z2) sei erwähnt, dass die Menge  $\Gamma$  im Allgemeinen keine gesteuert positiv invariante Menge darstellt. Das Ziel (Z3) kann erst dann genauer spezifiziert werden, wenn der Pfad festgelegt ist und hängt damit sehr stark von der betrachteten Anwendung ab. Darüber hinaus spielen unter anderem auch die Systemstruktur und die Anzahl der Eingangsgrößen eine Rolle. Diese bestimmen auch, welche Bewegung auf dem Pfad überhaupt realisiert werden kann.

## 4.3 Pfadfolgeregelung basierend auf exakter Linearisierung

Ein Ansatz für den systematischen Entwurf von Pfadfolgereglern beruht auf der Verwendung von Konzepten der exakten Eingangs-Ausgangs- und Eingangs-Zustandslinearisierung. Das System wird dabei in eine spezifische Normalform transformiert. In den transformierten Koordinaten ist der weitere Entwurf eines (Pfadfolge-)Reglers für die Bewegung zum/auf dem Pfad wesentlich vereinfacht. Bezüglich der Herangehensweise sind gewisse Unterschiede für parametrisierte und implizit definierte Pfade gegeben. Die gesuchte bzw. resultierende Normalform ist aber qualitativ die selbe.

### 4.3.1 Erweiterung der exakten Linearisierung für AI-Systeme

In [4.1] ist die exakte Eingangs-Ausgangs- und Eingangs-Zustandslinearisierung für AI-Systeme

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u} \quad (4.24a)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad (4.24b)$$

mit  $p = m$ , d. h. gleich vielen Eingangs- wie Ausgangsgrößen, erläutert. Im Folgenden soll das Konzept der exakten Linearisierung auf Systeme der Form (4.24) mit  $m \geq p$  erweitert werden. Um auf die Erweiterung hinzuweisen, werden in weiterer Folge Systeme der Form (4.24) mit mehr Eingangs- als Ausgangsgrößen (d. h.  $m > p$ ) als *nichtquadratische Systeme* bezeichnet. Alle kommenden Resultate gelten aber auch für den Spezialfall  $p = m$  (quadratisches System) mit den gleichen Resultaten wie in [4.1].

**Definition 4.2 (Relativer Grad eines nichtquadratischen AI-Systems).** Das nichtquadratische System (4.24) hat den vektoriiellen relativen Grad  $\{r_1, r_2, \dots, r_p\}$  mit  $r = \sum_{j=1}^p r_j \leq n$  an der Stelle  $\bar{\mathbf{x}}$ , wenn

$$(A) \quad L_{\mathbf{g}_j} L_{\mathbf{f}}^k h_i(\mathbf{x}) = 0, \quad j = 1, \dots, m, \quad i = 1, \dots, p, \quad k = 0, \dots, r_i - 2 \quad \text{für alle } \mathbf{x} \text{ in einer Umgebung } \mathcal{U} \text{ von } \bar{\mathbf{x}} \text{ und}$$

(B) die  $(p \times m)$ -Entkopplungsmatrix

$$\mathbf{D}(\mathbf{x}) = \begin{bmatrix} \mathbf{L}_{\mathbf{g}_1} \mathbf{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) & \mathbf{L}_{\mathbf{g}_2} \mathbf{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) & \cdots & \mathbf{L}_{\mathbf{g}_m} \mathbf{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) \\ \mathbf{L}_{\mathbf{g}_1} \mathbf{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) & \mathbf{L}_{\mathbf{g}_2} \mathbf{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) & \cdots & \mathbf{L}_{\mathbf{g}_m} \mathbf{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{L}_{\mathbf{g}_1} \mathbf{L}_{\mathbf{f}}^{r_p-1} h_p(\mathbf{x}) & \mathbf{L}_{\mathbf{g}_2} \mathbf{L}_{\mathbf{f}}^{r_p-1} h_p(\mathbf{x}) & \cdots & \mathbf{L}_{\mathbf{g}_m} \mathbf{L}_{\mathbf{f}}^{r_p-1} h_p(\mathbf{x}) \end{bmatrix} \quad (4.25)$$

vollen Rang an der Stelle  $\mathbf{x} = \bar{\mathbf{x}}$  hat, d. h.  $\text{rang}(\mathbf{D}(\bar{\mathbf{x}})) = p$ .

Besitzt ein nichtquadratisches System (4.24) den vektoriellen relativen Grad  $\{r_1, r_2, \dots, r_p\}$  dann kommen in den zeitlichen Ableitungen der Ausgangsgrößen  $y_j^{(k)}$ ,  $j = 1, \dots, p$ ,  $k = 0, \dots, r_j - 1$ , keine Eingangsgrößen  $u_i$ ,  $i = 1, \dots, m$  vor. Darüber hinaus gilt

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_{p-1}^{(r_{p-1})} \\ y_p^{(r_p)} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{L}_{\mathbf{f}}^{r_1} h_1(\mathbf{x}) \\ \vdots \\ \mathbf{L}_{\mathbf{f}}^{r_{p-1}} h_{p-1}(\mathbf{x}) \\ \mathbf{L}_{\mathbf{f}}^{r_p} h_p(\mathbf{x}) \end{bmatrix}}_{\mathbf{b}(\mathbf{x})} + \mathbf{D}(\mathbf{x}) \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_{m-1} \\ u_m \end{bmatrix}}_{\mathbf{u}}. \quad (4.26)$$

In einer Umgebung von  $\bar{\mathbf{x}}$  kann damit mit Hilfe des Zustandsregelgesetzes

$$\mathbf{u} = \mathbf{D}^T(\mathbf{x}) \left( \mathbf{D}(\mathbf{x}) \mathbf{D}^T(\mathbf{x}) \right)^{-1} (\mathbf{v} - \mathbf{b}(\mathbf{x})) \quad (4.27)$$

ein *exakt lineares Eingangs-Ausgangsverhalten* vom neuen Eingang  $\mathbf{v}^T = [v_1 \ \cdots \ v_p]$  zum Ausgang  $\mathbf{y}^T = [y_1 \ \cdots \ y_p]$  in Form von  $p$  Integratorketten der Länge  $r_j$ ,  $j = 1, \dots, p$ ,

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_{p-1}^{(r_{p-1})} \\ y_p^{(r_p)} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_{p-1} \\ v_p \end{bmatrix} \quad (4.28)$$

erzeugt werden. Die Wahl eines Diffeomorphismus in der Form

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \xi_{1,1} \\ \xi_{1,2} \\ \vdots \\ \xi_{1,r_1} \\ \xi_{2,1} \\ \vdots \\ \xi_{2,r_2} \\ \vdots \\ \xi_{p,1} \\ \vdots \\ \xi_{p,r_p} \\ \eta_1 \\ \vdots \\ \eta_{m-r} \end{bmatrix} = \boldsymbol{\Phi}(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ L_{\mathbf{f}}h_1(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}) \\ h_2(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x}) \\ \vdots \\ h_p(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{r_p-1}h_p(\mathbf{x}) \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} \quad (4.29)$$

und die Zustandsrückführung (4.27) transformieren das System (4.24) in die Byrnes-Isidori Normalform, die im Allgemeinen die Form

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{v} \quad (4.30a)$$

$$\dot{\boldsymbol{\eta}} = \tilde{\mathbf{q}}(\boldsymbol{\eta}, \boldsymbol{\xi}) + \tilde{\mathbf{P}}(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v} \quad (4.30b)$$

aufweist. Das Teilsystem (4.30a) liegt in Brunovsky Normalform vor.

In [4.1] sind die Bedingungen angeführt unter denen die Methode der exakten Eingangs-Ausgangslinearisierung zu einem stabilen geschlossenen Kreis führt. Unter analogen Anforderungen kann auch in der Normalform (4.30) ein Regler entworfen werden, der einen stabilen geschlossenen Kreis liefert. Allerdings weist die Verwendung der Zustandsrückführung (4.27) einen Nachteil auf. Das System (4.24) in den ursprünglichen Koordinaten weist  $m$  Freiheitsgrade in Form der Eingangsgrößen  $\mathbf{u} \in \mathbb{R}^m$  auf. Wählt man die Zustandsrückführung (4.27), dann gilt dies für das System (4.30) in transformierten Koordinaten nicht mehr. Es stehen nur mehr  $p$  neue Eingangsgrößen  $\mathbf{v} \in \mathbb{R}^p$  gemäß der Anzahl der zu regelnden Ausgangsgrößen  $\mathbf{y} \in \mathbb{R}^p$  zur Verfügung. Damit verliert man Freiheitsgrade, die man nutzen könnte, um bspw. das System (4.30b) gezielt zu beeinflussen. Nachfolgend wird gezeigt, dass dieser Nachteil behoben werden kann und eine Normalform existiert, die für die Pfadfolgeregelung geeigneter ist. Für die Formulierung des Ergebnisses wird noch nachstehende Definition benötigt.

**Definition 4.3** (Äquivalenz unter statischer Rückführung). Zwei dynamische Systeme der Form

$$\Sigma : \quad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u} \quad (4.31a)$$

$$\hat{\Sigma} : \quad \dot{\hat{\mathbf{x}}} = \hat{\mathbf{f}}(\hat{\mathbf{x}}) + \hat{\mathbf{G}}(\hat{\mathbf{x}})\hat{\mathbf{u}} \quad (4.31b)$$

mit  $\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}^n$  und  $\mathbf{u}, \hat{\mathbf{u}} \in \mathbb{R}^m$  sind lokal äquivalent unter statischer Rückführung an einer Stelle  $\bar{\mathbf{x}}$  falls in einer Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}}$

(A) eine reguläre Zustandsrückführung

$$\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x}) + \boldsymbol{\Lambda}(\mathbf{x})\hat{\mathbf{u}} \quad (4.32)$$

mit glatten Abbildungen  $\boldsymbol{\alpha}$  und  $\boldsymbol{\Lambda}$  sowie  $\boldsymbol{\Lambda}(\mathbf{x})$  invertierbar für alle  $\mathbf{x}$  in der Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}}$  und

(B) ein Diffeomorphismus  $\hat{\mathbf{x}} = \boldsymbol{\Phi}(\mathbf{x})$

so existieren, dass

$$\hat{\mathbf{f}}(\hat{\mathbf{x}}) = \left. \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}}(\mathbf{x})(\mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x})) \right|_{\mathbf{x}=\boldsymbol{\Phi}^{-1}(\hat{\mathbf{x}})} \quad (4.33a)$$

$$\hat{\mathbf{G}}(\hat{\mathbf{x}}) = \left. \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}}(\mathbf{x})\mathbf{G}(\mathbf{x})\boldsymbol{\Lambda}(\mathbf{x}) \right|_{\mathbf{x}=\boldsymbol{\Phi}^{-1}(\hat{\mathbf{x}})}. \quad (4.33b)$$

Man sagt, die beiden Systeme  $\Sigma$  und  $\hat{\Sigma}$  sind äquivalent unter statischer Rückführung auf  $\mathcal{U}$ .

**Aufgabe 4.1.** Rechnen Sie die Beziehungen (4.33) nach.

Der folgende Satz aus [4.4] liefert nun die gewünschte Normalform.

**Satz 4.1** (Exakte Eingangs-Ausgangslinearisierung für nichtquadratische AI-Systeme). Angenommen, das System (4.24) hat vektoriellen relativen Grad  $\{r_1, r_2, \dots, r_p\}$  mit  $r = \sum_{j=1}^p r_j \leq n$  an der Stelle  $\bar{\mathbf{x}}$ . Dann gilt, dass (4.24) in einer Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}}$  äquivalent unter statischer Rückführung zu einem System der Form

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{v}_1 \quad (4.34a)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \boldsymbol{\xi}) + \mathbf{G}_1(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v}_1 + \mathbf{G}_2(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v}_2 \quad (4.34b)$$

ist. Dabei gilt  $\boldsymbol{\xi} \in \mathbb{R}^r$ ,  $\boldsymbol{\eta} \in \mathbb{R}^{n-r}$ ,  $\mathbf{v}_1 \in \mathbb{R}^p$  und  $\mathbf{v}_2 \in \mathbb{R}^{m-p}$ . Das Teilsystem (4.34a) liegt in Brunovsky Normalform vor. In den neuen Koordinaten gilt mit einer Partitionierung von  $\boldsymbol{\xi}$  gemäß (4.29), dass  $y_i = \xi_{i,1}$ ,  $i = 1, \dots, p$ .

*Beweis.* Analog zu [4.1] gilt auch hier, dass ein Diffeomorphismus  $\mathbf{z} = \boldsymbol{\Phi}(\mathbf{x})$  gemäß (4.29) in einer Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}}$  existiert. Damit folgt die Systemdynamik in

transformierten Koordinaten zu

$$\begin{bmatrix} \dot{\xi}_{1,r_1} \\ \dot{\xi}_{2,r_2} \\ \vdots \\ \dot{\xi}_{p,r_p} \end{bmatrix} = \begin{bmatrix} L_{\mathbf{f}}^{r_1} h_1(\mathbf{x}) \\ L_{\mathbf{f}}^{r_2} h_2(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{r_p} h_p(\mathbf{x}) \end{bmatrix}_{\mathbf{x}=\Phi^{-1}(\mathbf{z})} + \mathbf{D}(\mathbf{x})|_{\mathbf{x}=\Phi^{-1}(\mathbf{z})} \mathbf{u} \quad (4.35a)$$

$$\dot{\xi}_{i,j} = \xi_{i,j+1}, \quad i = 1, \dots, p, \quad j = 1, \dots, r_i - 1 \quad (4.35b)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{q}(\boldsymbol{\eta}, \boldsymbol{\xi}) + \mathbf{P}(\boldsymbol{\eta}, \boldsymbol{\xi}) \mathbf{u} \quad (4.35c)$$

mit

$$q_i(\boldsymbol{\eta}, \boldsymbol{\xi}) = L_{\mathbf{f}} \phi_{r+i}(\mathbf{x})|_{\mathbf{x}=\Phi^{-1}(\mathbf{z})}, \quad i = 1, \dots, n - r \quad (4.35d)$$

$$P_{i,j}(\boldsymbol{\eta}, \boldsymbol{\xi}) = L_{\mathbf{g}_j} \phi_{r+i}(\mathbf{x})|_{\mathbf{x}=\Phi^{-1}(\mathbf{z})}, \quad i = 1, \dots, n - r, \quad j = 1, \dots, m. \quad (4.35e)$$

Gemäß der Annahme folgt, dass die Entkopplungsmatrix  $\mathbf{D}$  in einer Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}} = \Phi^{-1}(\bar{\mathbf{z}})$  vollen Rang besitzt. Sei  $\mathbf{N}(\mathbf{x}) \in \mathbb{R}^{m \times (m-p)}$  eine Matrix, deren Spalten den Nullraum von  $\mathbf{D}(\mathbf{x})$  aufspannen. Damit lässt sich die in der Umgebung  $\mathcal{U}$  von  $\bar{\mathbf{x}}$  reguläre Matrix  $\boldsymbol{\Lambda}(\mathbf{x}) = \begin{bmatrix} \mathbf{M}(\mathbf{x}) & \mathbf{N}(\mathbf{x}) \end{bmatrix}$  mit  $\mathbf{M}(\mathbf{x}) = \mathbf{D}^T(\mathbf{x}) (\mathbf{D}(\mathbf{x}) \mathbf{D}^T(\mathbf{x}))^{-1} \in \mathbb{R}^{m \times p}$  definieren. Mit

$$\boldsymbol{\alpha}(\mathbf{x}) = -\boldsymbol{\Lambda}(\mathbf{x}) \begin{bmatrix} L_{\mathbf{f}}^{r_1} h_1(\mathbf{x}) \\ L_{\mathbf{f}}^{r_2} h_2(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{r_p} h_p(\mathbf{x}) \\ \mathbf{0} \end{bmatrix}, \quad (4.36)$$

wobei der Nullvektor in (4.36)  $m - p$  Zeilen aufweist, folgt die gesuchte Zustandsrückführung

$$\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x})|_{\mathbf{x}=\Phi^{-1}(\mathbf{z})} + \boldsymbol{\Lambda}(\mathbf{x})|_{\mathbf{x}=\Phi^{-1}(\mathbf{z})} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}, \quad (4.37)$$

die die gewünschte Normalform (4.34) in  $(\boldsymbol{\xi}, \boldsymbol{\eta})$ -Koordinaten liefert.  $\square$

**Aufgabe 4.2.** Rechnen Sie nach, dass die im Beweis von Satz 4.1 angegebene Rückführung tatsächlich auf die gewünschte Normalform führt.

In der Normalform (4.34) kommen im Gegensatz zu (4.30) nicht nur die neuen Eingangsgrößen  $\mathbf{v}_1 \in \mathbb{R}^p$  sondern auch  $\mathbf{v}_2 \in \mathbb{R}^{m-p}$  vor. Das bedeutet, dass die zusätzlichen Freiheitsgrade im Fall  $m > p$  nicht verloren gehen. Sie können in Form von  $\mathbf{v}_2$  ggf. dafür genutzt werden, das Teilsystem (4.34b) zu beeinflussen. Basierend auf der Normalform (4.34) werden nachfolgend Pfadfolgeregler für implizit definierte und parametrisierte Pfade entworfen.

### 4.3.2 Implizit definierte Pfade

Ein möglicher Ansatz zum Entwurf von Pfadfolgereglern für implizit definierte Pfade besteht darin, neue fiktive Ausgangsgrößen

$$\hat{\mathbf{y}} = \hat{\mathbf{h}}(\mathbf{x}) = \mathbf{s}(\mathbf{h}(\mathbf{x})) \quad (4.38)$$

mit  $\hat{\mathbf{y}} \in \mathbb{R}^{p-1}$  zu definieren. Für die nachfolgenden Ausführungen sind zwei Annahmen essenziell.

1. Es wird vorausgesetzt, dass das System  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$  mindestens  $p - 1$  Eingangsgrößen aufweist, d. h.  $m \geq p - 1$ ,  $p > 1$ .
2. Das System  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$  mit der Ausgangsgröße  $\hat{\mathbf{y}}$  gemäß (4.38) hat in  $\Gamma^*$  einen wohldefinierten vektoriellen relativen Grad gemäß Definition 4.2.

Unter diesen Annahmen kann Satz 4.1 angewandt und das System  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$  mit dem Ausgang (4.38) auf die Normalform (4.34) transformiert werden. Diese Normalform stellt sich im gegebenen Zusammenhang in der Form

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{v}^{\hat{\eta}} \quad (4.39a)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \boldsymbol{\xi}) + \mathbf{G}^{\hat{\eta}}(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v}^{\hat{\eta}} + \mathbf{G}^{\parallel}(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v}^{\parallel} \quad (4.39b)$$

mit  $\mathbf{v}^{\hat{\eta}} \in \mathbb{R}^{p-1}$  und  $\mathbf{v}^{\parallel} \in \mathbb{R}^{m-(p-1)}$  dar. Die Darstellung (4.39) wird in der englischsprachigen Literatur als *Transverse Normal Form* bezeichnet und soll hier analog dazu *transversale Normalform* genannt werden. Die zugehörige Zustandsrückführung stellt sich gemäß (4.32) in der Form

$$\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x}) + \boldsymbol{\Lambda}(\mathbf{x}) \begin{bmatrix} \mathbf{v}^{\hat{\eta}} \\ \mathbf{v}^{\parallel} \end{bmatrix} \quad (4.40)$$

dar.

Das Differenzialgleichungssystem (4.39a) stellt die *transversale Dynamik*, d. h. die Dynamik quer zum Pfad  $\mathcal{P}$ , dar. Dies begründet sich in der Tatsache, dass durch die Koordinaten  $\xi_{i,1}$ ,  $i = 1, \dots, p - 1$  gemäß (4.35a) und (4.35b) direkt die fiktiven Ausgangsgrößen  $\hat{\mathbf{y}}$  in der Form

$$\begin{bmatrix} \xi_{1,1} \\ \xi_{2,1} \\ \vdots \\ \xi_{p-1,1} \end{bmatrix} = \hat{\mathbf{h}}(\mathbf{x}) \quad (4.41)$$

gegeben sind. Falls

$$\xi_{i,1} = 0, \quad i = 1, \dots, p - 1 \quad (4.42)$$

gilt, befindet sich der Ausgang  $\mathbf{y}$  des Systems exakt am Pfad  $\mathcal{P}$ . Ist (4.42) nicht für alle  $\xi_{i,1}$ ,  $i = 1, \dots, p - 1$  erfüllt, bewegt sich der Ausgang  $\mathbf{y}$  abseits vom Pfad. Die verbleibenden Koordinaten in  $\boldsymbol{\xi}$  stellen entsprechend dem vektoriellen relativen Grad die Ableitungen von  $\hat{\mathbf{h}}(\mathbf{x})$  dar. Damit ist unmittelbar einsichtig, dass die Koordinaten  $\boldsymbol{\xi}$  der

Dynamik quer (transversal) zum Pfad zugeordnet sind. Die transversale Dynamik (4.39a) liegt entsprechend Satz 4.1 in Brunovsky Normalform vor, d. h. das System (4.39a) ist vollständig erreichbar. Damit ist das Ziel (Z1) der Pfadfolgeregelung leicht durch den Entwurf eines asymptotisch stabilisierenden linearen Zustandsreglers in der Form

$$v_i^{\uparrow} = -k_{i,0} \int_{t_0}^t \xi_{i,1} dt - \sum_{j=1}^{r_i} k_{i,j} \xi_{i,j} \quad i = 1, \dots, p-1 \quad (4.43)$$

mit geeigneten Koeffizienten  $k_{i,j} > 0$  erfüllbar. Das Regelgesetz (4.43) beinhaltet einen Integralanteil, um z. B. mit Modellabweichungen besser umgehen zu können und damit besser für einen praktischen Einsatz geeignet zu sein.

Um die Dynamik (4.39b) zu erhalten, müssen die Funktionen  $\phi_{r+i}(\mathbf{x})$ ,  $i = 1, \dots, n-r$  zur Komplettierung des Diffeomorphismus

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \Phi(\mathbf{x}) = \begin{bmatrix} \Phi_{\xi}(\mathbf{x}) \\ \Phi_{\eta}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \Phi_{\xi}(\mathbf{x}) \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} \quad (4.44)$$

gewählt werden. Da für  $\xi = \mathbf{0}$  unabhängig vom Wert von  $\eta$  die Ausgangsgrößen  $\mathbf{y}$  exakt am Pfad liegen, stellt (4.39b) die *tangentiale Dynamik* dar. Die Koordinaten  $\eta_i = \phi_{r+i}(\mathbf{x})$ ,  $i = 1, \dots, n-r$  werden entsprechend dem vorliegenden Pfad und der applikationsspezifischen Ziele (Bewegung auf  $\mathcal{P}$ ) gewählt. Dabei ist noch die zusätzliche Nebenbedingung zu beachten, dass durch (4.44) tatsächlich ein Diffeomorphismus gegeben ist. Im Fall  $m > p-1$  sind die Eingangsgrößen  $\mathbf{v}^{\parallel}$  in (4.39b) vorhanden und damit können geeignete Regler entworfen werden, um (Z3) zu erfüllen.

Das noch offene Ziel der Pfadfolgeregelung ist (Z2), die Invarianzeigenschaft. Befindet sich der Zustand des Systems zum Zeitpunkt  $t_0$  in der Menge

$$\Gamma = \{ \bar{\mathbf{x}} \in \mathbb{R}^n \mid \mathbf{s}(\mathbf{h}(\bar{\mathbf{x}})) = \mathbf{0} \}, \quad (4.45)$$

d. h. die Ausgangsgrößen befinden sich exakt am Pfad  $\mathcal{P}$ , ist es im Allgemeinen nicht garantiert, dass dies für alle  $t > t_0$  der Fall ist. Dafür müssen auch alle Ableitungen von  $\hat{\mathbf{h}}(\mathbf{x})$  zum Zeitpunkt  $t_0$  null sein. Aus dieser Überlegung folgt die Definition der Menge

$$\Gamma^* = \{ \bar{\mathbf{x}} \in \mathbb{R}^n \mid \Phi_{\xi}(\bar{\mathbf{x}}) = \mathbf{0} \} \subset \Gamma. \quad (4.46)$$

Gilt nun, dass  $\mathbf{x}(t_0) \in \Gamma^*$ , dann kann durch die Wahl  $\mathbf{v}^{\uparrow} \equiv \mathbf{0}$  erreicht werden, dass  $\xi \equiv \mathbf{0} \forall t > t_0$  wie leicht an (4.39a) erkennbar ist. Damit gilt, dass die Menge  $\Gamma^*$  gesteuert positiv invariant ist und der Pfad  $\mathcal{P}$  für alle Zeiten  $t > t_0$  nicht mehr verlassen wird, was die Erfüllung der Invarianzeigenschaft (Z2) impliziert.

Damit ist gezeigt, dass basierend auf der Normalform (4.39), die unter den getroffenen Annahmen stets systematisch errechenbar ist, einfach ein Pfadfolgeregler für die Erfüllung der Ziele (Z1)–(Z3) entworfen werden kann.



**Aufgabe 4.3.** Zeigen Sie, dass mit der gezeigten Vorgehensweise die Erfüllung der Ziele (Z1)–(Z3) nicht widersprüchlich ist, d. h. das Erreichen eines Zieles beeinflusst nicht die anderen beiden.

Die Menge  $\Gamma^*$  gemäß (4.46) stellt eine Untermannigfaltigkeit des Zustandsraumes dar. In der englischsprachigen Literatur wird sie gelegentlich als *zero path error manifold* oder auch *path following manifold* bezeichnet. Die Dynamik des Systems auf dieser Mannigfaltigkeit ist durch

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \mathbf{0}) + \mathbf{G}^{\parallel}(\boldsymbol{\eta}, \mathbf{0})\mathbf{v}^{\parallel} \quad (4.47)$$

gegeben. Sie entspricht der internen Dynamik des Systems auf dem Pfad  $\mathcal{P}$ .

**Beispiel 4.4 (Fahrzeug mit variabler Vorwärtsgeschwindigkeit).** Die gezeigte Theorie soll für die Pfadfolgeregelung eines Fahrzeuges, beschrieben durch das mathematische Modell

$$\dot{\mathbf{x}} = \frac{d}{dt} \begin{bmatrix} x \\ y \\ \varphi \\ v_l \end{bmatrix} = \underbrace{\begin{bmatrix} v_l \cos \varphi \\ v_l \sin \varphi \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{f}_v(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ v_l & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{G}_v(\mathbf{x})} \mathbf{u} \quad (4.48a)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) = \begin{bmatrix} x \\ y \end{bmatrix}, \quad (4.48b)$$

angewandt werden. Die Stellgröße  $u_1$  kann mit dem Lenkwinkel in Verbindung gebracht werden während  $u_2$  direkt die zeitliche Ableitung der Vorwärtsgeschwindigkeit  $v_l \neq \text{konst.}$  angibt. Der Winkel  $\varphi$  beschreibt die Orientierung des Fahrzeuges relativ zur  $x$ -Achse. Der Ausgang  $\mathbf{y}$  ist durch die  $x$ - und  $y$ -Position des Fahrzeuges gegeben.

Im konkreten Fall soll der Pfadfolgeregler das Fahrzeug zu und entlang einer Ellipse gegeben durch die implizite Gleichung

$$s_e(\mathbf{y}) = \frac{(y_1 - y_{10})^2}{a^2} + \frac{(y_2 - y_{20})^2}{b^2} - 1 \quad (4.49)$$

führen. Als Ziel (Z3) wird für diese Anwendung gewählt, dass die Vorwärtsgeschwindigkeit auf dem Pfad einer hinreichend oft stetig differenzierbaren Solltrajektorie  $v_{l,d}(t)$  folgen soll.

Der erste Schritt besteht darin, die am Beginn von Abschnitt 4.3.2 getroffenen Annahmen zu verifizieren. Die Voraussetzung  $m \geq p - 1 = 1$  ist offensichtlich erfüllt. Für die Überprüfung der zweiten Annahme wird die fiktive Ausgangsgröße

$$\hat{y} = s_e(\mathbf{h}(\mathbf{x})) = \frac{(x - y_{10})^2}{a^2} + \frac{(y - y_{20})^2}{b^2} - 1 = \hat{h}(\mathbf{x}) \quad (4.50)$$

benötigt. Nun muss überprüft werden, ob das System (4.48a) mit der Ausgangsgröße (4.50) einen wohldefinierten relativen Grad in  $\Gamma^*$  aufweist. Dies soll gemeinsam mit

der Berechnung von  $\Gamma^*$  gemäß (4.46) im nächsten Schritt durchgeführt werden. Dazu berechnet man zuerst die Ableitungen von (4.50) in der Form

$$\dot{\hat{y}} = \mathbf{L}_{\mathbf{f}_v} \hat{h}(\mathbf{x}) + \underbrace{\mathbf{L}_{\mathbf{g}_{v,1}} \hat{h}(\mathbf{x})}_0 u_1 + \underbrace{\mathbf{L}_{\mathbf{g}_{v,2}} \hat{h}(\mathbf{x})}_0 u_2 = 2v_l \left( \frac{(x - y_{10}) \cos \varphi}{a^2} + \frac{(y - y_{20}) \sin \varphi}{b^2} \right) \quad (4.51a)$$

$$\ddot{\hat{y}} = \mathbf{L}_{\mathbf{f}_v}^2 \hat{h}(\mathbf{x}) + \mathbf{L}_{\mathbf{g}_{v,1}} \mathbf{L}_{\mathbf{f}_v} \hat{h}(\mathbf{x}) u_1 + \mathbf{L}_{\mathbf{g}_{v,2}} \mathbf{L}_{\mathbf{f}_v} \hat{h}(\mathbf{x}) u_2 = \quad (4.51b)$$

$$= 2v_l^2 \left( \frac{\cos^2 \varphi}{a^2} + \frac{\sin^2 \varphi}{b^2} \right) + \underbrace{\begin{bmatrix} 2v_l^2 \left( \frac{(y - y_{20}) \cos \varphi}{b^2} - \frac{(x - y_{10}) \sin \varphi}{a^2} \right) \\ 2 \left( \frac{(x - y_{10}) \cos \varphi}{a^2} + \frac{(y - y_{20}) \sin \varphi}{b^2} \right) \end{bmatrix}^T}_{\mathbf{D}(\mathbf{x})} \mathbf{u}, \quad (4.51c)$$

wobei in  $\ddot{\hat{y}}$  der Eingang  $\mathbf{u}$  erscheint. Damit ist  $\Gamma^*$  durch

$$\Gamma^* = \left\{ \bar{\mathbf{x}} \in \mathbb{R}^4 \mid \hat{h}(\bar{\mathbf{x}}) = \mathbf{L}_{\mathbf{f}_v} \hat{h}(\bar{\mathbf{x}}) = 0 \right\} \quad (4.52)$$

gegeben. Sei  $\bar{\mathbf{x}} = [\bar{x} \ \bar{y} \ \bar{\varphi} \ \bar{v}_l]^T$  ein Punkt, der in  $\Gamma^*$  liegt. Es muss nun noch nachgewiesen werden, dass  $\mathbf{D}(\bar{\mathbf{x}})$  vollen Rang für alle  $\bar{\mathbf{x}} \in \Gamma^*$  aufweist. Die Bedingung  $\mathbf{L}_{\mathbf{f}_v} \hat{h}(\bar{\mathbf{x}}) = 0$  gemäß (4.51a) impliziert, dass sich das Fahrzeug tangential zum Pfad bewegt. Dies kann verifiziert werden, indem der Winkel der Tangente an die Ellipse, der durch

$$\bar{\varphi} = \arctan \left( \frac{(\bar{x} - y_{10}) b^2}{-(\bar{y} - y_{20}) a^2} \right) \quad (4.53)$$

gegeben ist, in (4.51a) eingesetzt wird. Drückt man nun beispielsweise aus  $\hat{h}(\bar{\mathbf{x}}) = 0$   $\bar{x}$  in Abhängigkeit von  $\bar{y}$  aus und setzt diese Beziehung zusammen mit (4.53) in  $\mathbf{D}$  ein, dann erhält man

$$\mathbf{D}(\bar{\mathbf{x}}) = \begin{bmatrix} -\frac{2\bar{v}_l^2 \sqrt{a^2(\bar{y} - y_{20})^2 + b^4 - b^2(\bar{y} - y_{20})^2}}{ab^2} & 0 \end{bmatrix} \quad (4.54)$$

für alle  $\bar{\mathbf{x}} \in \Gamma^*$ . Damit ist gezeigt, dass die Entkopplungsmatrix für alle  $\bar{v}_l \neq 0$  und  $y_{20} - b \leq \bar{y} \leq y_{20} + b$  Rang 1 aufweist und damit das System (4.48a) mit der Ausgangsgröße (4.50) einen wohldefinierten relativen Grad  $r = 2$  in  $\Gamma^*$  aufweist.

Damit kann direkt Satz 4.1 angewandt werden. Der erste Teil des Diffeomorphismus für die Transformation in die transversale Normalform lautet

$$\Phi_{\xi}(\mathbf{x}) = \begin{bmatrix} \hat{h}(\mathbf{x}) \\ \mathbf{L}_{\mathbf{f}_v} \hat{h}(\mathbf{x}) \end{bmatrix}. \quad (4.55a)$$

Im Hinblick auf das Ziel (Z3) wird die Zustandstransformation mit

$$\Phi_{\eta}(\mathbf{x}) = \begin{bmatrix} v_l \\ \varphi \end{bmatrix} \quad (4.55b)$$

zu

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \Phi(\mathbf{x}) = \begin{bmatrix} \Phi_{\xi}(\mathbf{x}) \\ \Phi_{\eta}(\mathbf{x}) \end{bmatrix} \quad (4.55c)$$

ergänzt.

**Aufgabe 4.4.** Verifizieren Sie, dass es sich bei (4.55) tatsächlich um einen Diffeomorphismus in  $\Gamma^*$  handelt.

Für die Berechnung der exakt linearisierenden Zustandsrückführung (4.37) wird zuerst die Matrix  $\mathbf{N}(\mathbf{x})$  in der Form

$$\mathbf{N}(\mathbf{x}) = \begin{bmatrix} -\frac{2\left(\frac{(x-y_{10})\cos\varphi}{a^2} + \frac{(y-y_{20})\sin\varphi}{b^2}\right)}{2v_l^2\left(\frac{(y-y_{20})\cos\varphi}{b^2} - \frac{(x-y_{10})\sin\varphi}{a^2}\right)} \\ 1 \end{bmatrix} \quad (4.56)$$

gewählt. Aus  $\mathbf{D}(\mathbf{x})$  gemäß (4.51c) lassen sich direkt  $\mathbf{M}(\mathbf{x}) = \mathbf{D}^T(\mathbf{x})(\mathbf{D}(\mathbf{x})\mathbf{D}^T(\mathbf{x}))^{-1} \in \mathbb{R}^{2 \times 1}$  und in weiterer Folge  $\Lambda(\mathbf{x}) = \begin{bmatrix} \mathbf{M}(\mathbf{x}) & \mathbf{N}(\mathbf{x}) \end{bmatrix}$  errechnen. Damit folgt die Zustandsrückführung zu

$$\mathbf{u} = \alpha(\mathbf{x}) + \Lambda(\mathbf{x}) \begin{bmatrix} v^{\uparrow} \\ v^{\parallel} \end{bmatrix} \quad (4.57)$$

mit

$$\alpha(\mathbf{x}) = -\Lambda(\mathbf{x}) \begin{bmatrix} \mathbf{L}_{f_v}^2 \hat{h}(\mathbf{x}) \\ 0 \end{bmatrix}. \quad (4.58)$$

Die transversale Normalform ergibt sich im vorliegenden Fall zu

$$\dot{\xi} = \begin{bmatrix} \xi_2 \\ v^{\uparrow} \end{bmatrix} \quad (4.59a)$$

$$\dot{\eta} = \mathbf{f}^0(\eta, \xi) + \mathbf{g}^{\uparrow}(\eta, \xi)v^{\uparrow} + \mathbf{g}^{\parallel}(\eta, \xi)v^{\parallel} \quad (4.59b)$$

wobei die Terme in  $\mathbf{f}^0$ ,  $\mathbf{g}^{\uparrow}$  und  $\mathbf{g}^{\parallel}$  bereits sehr umfangreich sind. Im Hinblick auf (Z3) ist vor allem die Einschränkung von (4.59b) auf  $\Gamma^*$  entsprechend (4.47) interessant. Diese interne Dynamik auf der Ellipse ergibt sich zu

$$\dot{\eta} = \begin{bmatrix} v^{\parallel} \\ \pm \frac{\eta_1^2 (a^4 (\cos \eta_2)^4 - 2a^2 b^2 (\cos \eta_2)^4 + b^4 (\cos \eta_2)^4 - 2a^4 (\cos \eta_2)^2 + 2a^2 b^2 (\cos \eta_2)^2 + a^4) \sin \eta_2}{b^2 \sqrt{a^4 \eta_1^2 (\sin \eta_2)^2 (a^2 (\sin \eta_2)^2 - b^2 (\sin \eta_2)^2 + b^2)}} \end{bmatrix} \quad (4.60)$$

wobei das Vorzeichen von  $\dot{\eta}_2$  davon abhängig ist, in welcher Richtung die Ellipse abgefahren wird. Die erste Zeile von (4.60) lautet  $\dot{v}_l = v^{\parallel}$  womit sich (Z3) mit dem Regelgesetz

$$v^{\parallel} = \dot{v}_{l,d}(t) - k^{\parallel}(v_l - v_{l,d}(t)) \quad k^{\parallel} > 0 \quad (4.61)$$

erzielen lässt. Die transversale Dynamik wird mit dem Regelgesetz

$$v^{\perp} = -k_1^{\perp}\xi_1 - k_2^{\perp}\xi_2 \quad k_1^{\perp}, k_2^{\perp} > 0 \quad (4.62)$$

stabilisiert. Es sei noch angemerkt, dass man durch das Erreichen des Zieles (Z3) zumindest auf der Ellipse sicherstellen kann, dass  $v_l \neq 0$  gilt. Damit ist gewährleistet, dass die Entkopplungsmatrix  $\mathbf{D}(\mathbf{x})$  stets vollen Rang in  $\Gamma^*$  aufweist.

Für die im Folgenden dargestellten Ergebnisse werden die Parameter  $a = 4$ ,  $b = 2.5$ ,  $y_{10} = 1$ ,  $y_{20} = -1.5$ ,  $k_1^{\perp} = 0.25$ ,  $k_2^{\perp} = 1$  und  $k^{\parallel} = 0.5$  verwendet. Weiters gelte  $v_{l,d}(t) = 0.5 + (\sin(\frac{t}{2}))^2$ . In Abbildung 4.4 ist die Trajektorie des Fahrzeuges für den Anfangswert  $\mathbf{x}_0 = [5.5 \ 1 \ \frac{3\pi}{2} \ 0.45]^T$  dargestellt. An der Startposition des Fahrzeuges ist ein Pfeil eingezeichnet, der die anfängliche Orientierung zusätzlich verdeutlichen soll.

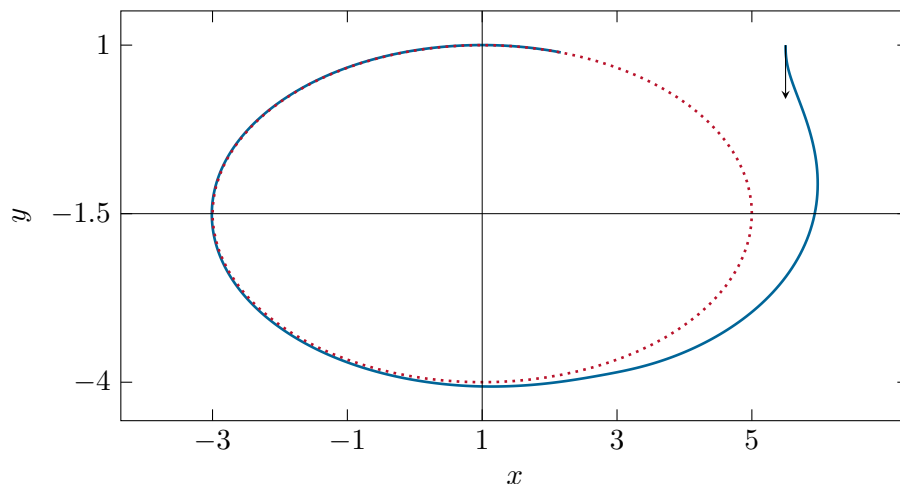


Abbildung 4.4: Trajektorie des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die Pfadfolgeregelung einer Ellipse.

In Abbildung 4.5 sind die zugehörigen Zeitverläufe im geschlossenen Kreis ersichtlich. Die Zustände  $\xi$  der transversalen Dynamik konvergieren zu null womit der Zustand  $\mathbf{x}$  in die Menge  $\Gamma^*$  gezwungen wird. Gleichzeitig ist ersichtlich, dass  $\eta_1 = v_l$  der Solltrajektorie  $v_{l,d}(t)$  nachgeführt wird.

Abbildung 4.6 zeigt Trajektorien des Fahrzeuges für verschiedene Anfangswerte. Wenn  $\mathbf{x}_0$  in  $\Gamma^*$  liegt (das Fahrzeug befindet sich auf der Ellipse und ist tangential dazu ausgerichtet) bleibt das Fahrzeug für alle zukünftigen Zeiten auf der Ellipse. Dies zeigt, dass die Invarianzeigenschaft (Z2) erfüllt ist. Je nach Startposition und -orientierung ist die Richtung, in welcher die Ellipse abgefahren wird, eine andere.

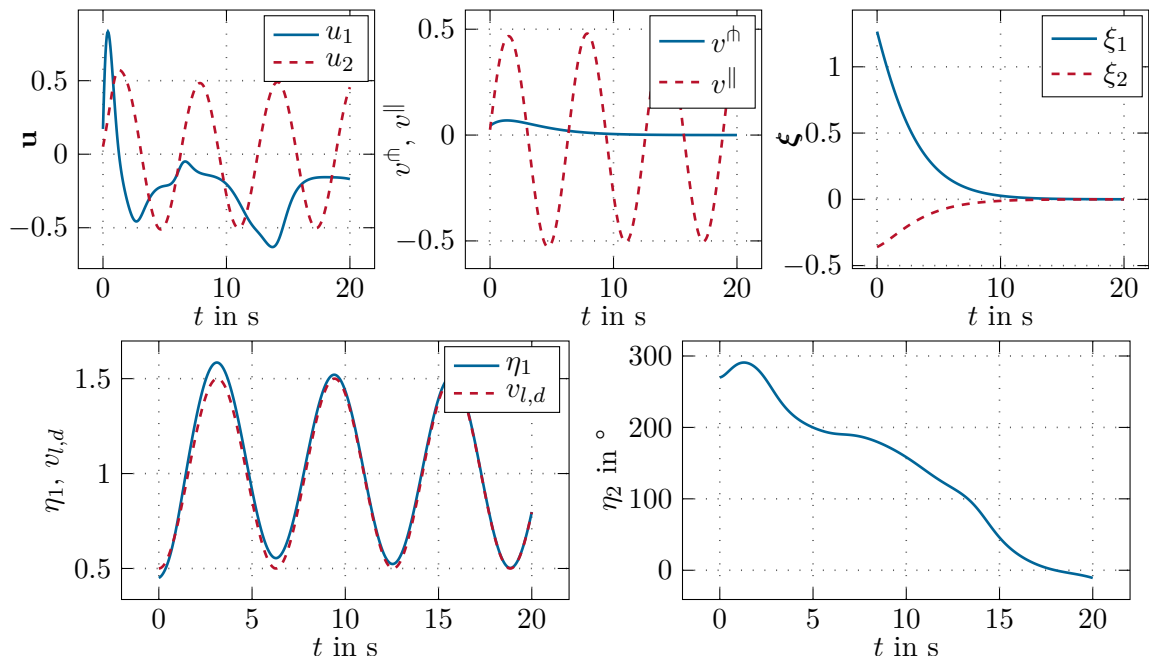


Abbildung 4.5: Zeitverläufe im geschlossenen Kreis bei der Pfadfolgeregelung einer Ellipse für das Fahrzeug mit  $v_l \neq \text{konst.}$

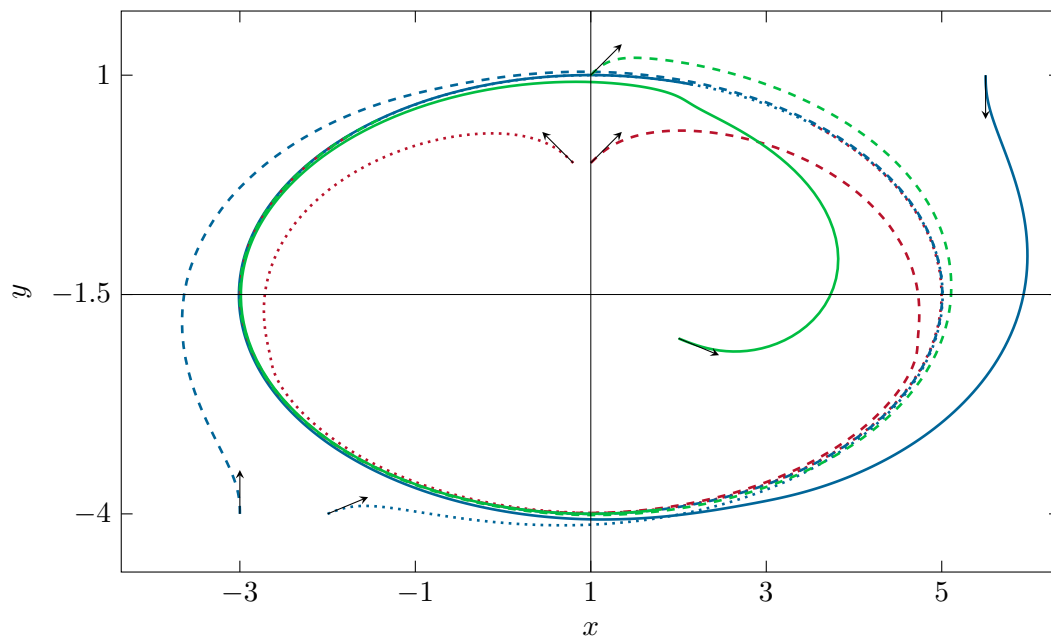


Abbildung 4.6: Trajektorien des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die Pfadfolgeregelung einer Ellipse und verschiedenen Anfangswerten.

**Beispiel 4.5 (Fahrzeug mit konstanter Vorwärtsgeschwindigkeit).** Im Gegensatz zu Beispiel 4.4 soll hier das Fahrzeug mit konstanter Vorwärtsgeschwindigkeit  $v_l = \text{konst.} > 0$  beschrieben durch die Differentialgleichungen

$$\dot{\mathbf{x}} = \frac{d}{dt} \begin{bmatrix} x \\ y \\ \varphi \end{bmatrix} = \underbrace{\begin{bmatrix} v_l \cos \varphi \\ v_l \sin \varphi \\ 0 \end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ v_l \end{bmatrix}}_{\mathbf{g}} u \quad (4.63a)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) = \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.63b)$$

betrachtet werden. Berechnet man wieder die Ableitungen von (4.50) in der Form

$$\dot{\hat{y}} = L_{\mathbf{f}} \hat{h}(\mathbf{x}) + \underbrace{L_{\mathbf{g}} \hat{h}(\mathbf{x})}_0 u = 2v_l \left( \frac{(x - y_{10}) \cos \varphi}{a^2} + \frac{(y - y_{20}) \sin \varphi}{b^2} \right) \quad (4.64a)$$

$$\ddot{\hat{y}} = L_{\mathbf{f}}^2 \hat{h}(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}} \hat{h}(\mathbf{x}) u = \quad (4.64b)$$

$$= 2v_l^2 \left( \frac{\cos^2 \varphi}{a^2} + \frac{\sin^2 \varphi}{b^2} \right) + \underbrace{\left( 2v_l^2 \left( \frac{(y - y_{20}) \cos \varphi}{b^2} - \frac{(x - y_{10}) \sin \varphi}{a^2} \right) \right)}_{d(\mathbf{x})} u \quad (4.64c)$$

sieht man, dass

1. die erste Ableitung von  $\hat{y}$  ident zu (4.51a) ist und
2. der Vorfaktor von  $u$  durch den ersten Eintrag der Entkopplungsmatrix  $\mathbf{D}(\mathbf{x})$  in (4.51c) gegeben ist.

Damit bleibt  $\Gamma^*$  qualitativ unverändert und der Ausgang  $\hat{y}$  besitzt wieder einen wohldefinierten relativen Grad in  $\Gamma^*$ .

Damit kann wieder direkt Satz 4.1 angewandt werden, dessen Aussagen auf Grund von  $m = p - 1 = 1$  zu den Ergebnissen in [4.1] ident sind. Die Transformation in die transversale Normalform kann wieder zu

$$\Phi(\mathbf{x}) = \begin{bmatrix} \hat{h}(\mathbf{x}) \\ L_{\mathbf{f}} \hat{h}(\mathbf{x}) \\ \varphi \end{bmatrix} \quad (4.65)$$

gewählt werden. Auf Grund von  $m = p - 1$  ist kein  $v^{\parallel}$  vorhanden. Daher wird auch die Matrix  $\mathbf{N}$  zur leeren Matrix und  $\mathbf{M}$  reduziert sich zu  $m(\mathbf{x}) = \frac{d(\mathbf{x})}{d(\mathbf{x})d(\mathbf{x})} = \frac{1}{d(\mathbf{x})} = \lambda(\mathbf{x})$ .

Damit folgt die exakt linearisierende Zustandsrückführung zu

$$u = \alpha(\mathbf{x}) + \lambda(\mathbf{x})v^{\text{fl}} = \frac{2v_l^2 \left( a^2 (\cos \varphi)^2 - (\cos \varphi)^2 b^2 - a^2 \right) + a^2 b^2 v^{\text{fl}}}{2v_l^2 (-\sin(\varphi)b^2 x + \sin(\varphi)b^2 y_{10} + \cos(\varphi)a^2 y - \cos(\varphi)a^2 y_{20})}. \quad (4.66)$$

Die transversale Normalform lautet

$$\dot{\boldsymbol{\xi}} = \begin{bmatrix} \xi_2 \\ v^{\text{fl}} \end{bmatrix} \quad (4.67a)$$

$$\dot{\eta} = f^0(\eta, \boldsymbol{\xi}) + g^{\text{fl}}(\eta, \boldsymbol{\xi})v^{\text{fl}} \quad (4.67b)$$

wobei die tangentielle Dynamik (4.67b) nicht beeinflussbar ist. Damit können in diesem Fall keine Ziele (Z3) berücksichtigt werden. Die Trajektorien des Fahrzeuges im geschlossenen Kreis sehen qualitativ sehr ähnlich zu denen von Beispiel 4.4 aus.

*Beispiel 4.6.* Wieder soll das Fahrzeug mit konstanter Vorwärtsgeschwindigkeit  $v_l$  in der Form (4.63) betrachtet werden. Diesmal soll ein Pfadfolgeregler für die Verfolgung einer Lemniskate gegeben durch die implizite Gleichung

$$s_l(\mathbf{y}) = y_1^4 - y_1^2 + y_2^2 \quad (4.68)$$

entworfen werden. Dazu wird wieder die fiktive Ausgangsgröße

$$\hat{y} = s_l(\mathbf{h}(\mathbf{x})) = x^4 - x^2 + y^2 = \hat{h}(\mathbf{x}) \quad (4.69)$$

angeschrieben und festgestellt, ob diese einen relativen Grad in  $\Gamma^*$  aufweist. Dazu werden die Ableitungen von (4.69)

$$\dot{\hat{y}} = \mathbf{L}_{\mathbf{f}} \hat{h}(\mathbf{x}) + \underbrace{\mathbf{L}_{\mathbf{g}} \hat{h}(\mathbf{x})}_0 u = (4x^3 - 2x)v_l \cos(\varphi) + 2yv_l \sin(\varphi) \quad (4.70a)$$

$$\ddot{\hat{y}} = \mathbf{L}_{\mathbf{f}}^2 \hat{h}(\mathbf{x}) + \mathbf{L}_{\mathbf{g}} \mathbf{L}_{\mathbf{f}} \hat{h}(\mathbf{x}) u = \quad (4.70b)$$

$$= (12x^2 - 2)v_l^2 (\cos \varphi)^2 + 2v_l^2 (\sin \varphi)^2 + \underbrace{2v_l^2 \left( (x - 2x^3) \sin(\varphi) + y \cos(\varphi) \right)}_{d(\mathbf{x})} u \quad (4.70c)$$

berechnet. Wie man unmittelbar erkennt, verschwindet der Vorfaktor  $d(\mathbf{x})$  von  $u$  für beliebige  $\begin{bmatrix} 0 & 0 & \varphi \end{bmatrix}^T \in \Gamma^*$  womit die Voraussetzung verletzt ist, dass das System mit der Ausgangsgröße  $\hat{y}$  einen relativen Grad in  $\Gamma^*$  aufweist. Mit dem resultierenden Regelgesetz könnte nicht der gesamte Pfad abgefahren werden. Wie im folgenden Abschnitt 4.3.3 gezeigt wird, kann basierend auf der parametrisierten Darstellung der Lemniskate aber trotzdem ein Pfadfolgeregler entworfen werden, der es ermöglicht, die gesamte Lemniskate abzufahren.

### 4.3.3 Parametrisierte Pfade

Das Ziel dieses Abschnitts ist der Entwurf von Pfadfolgereglern für parametrisierte Pfade definiert durch eine Abbildung  $\sigma(\theta) : \mathcal{T} \rightarrow \mathbb{R}^p$ . Dazu sollen auch hier ähnliche Annahmen wie in Abschnitt 4.3.2 getroffen werden, d. h.

1. für das System  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$  gilt  $m \geq p - 1$ ,  $p > 1$ , und
2. das System  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$  mit der Ausgangsgröße  $\mathbf{y} = \mathbf{h}(\mathbf{x})$  hat in  $\Gamma^*$  einen wohldefinierten vektoriellen relativen Grad  $\{r_1, r_2, \dots, r_p\}$ . Da  $m \geq p - 1$  erlaubt ist, muss im Folgenden eine Unterscheidung getroffen werden.
  - a) Falls  $m \geq p$  gilt, kann direkt die Definition 4.2 des vektoriellen relativen Grades verwendet werden.
  - b) Für den Fall  $m = p - 1$  bleibt diese Definition unverändert bis auf die Tatsache, dass  $\text{rang}(\mathbf{D}(\bar{\mathbf{x}})) = m$  gelten muss.

Um die zeitliche Evolution des Pfadparameters  $\theta$  beschreiben zu können, wird ein *Hilfssystem* der Form

$$\theta^{(\hat{r})} = v_\theta \quad (4.71)$$

mit einer neuen *fiktiven Eingangsgröße*  $v_\theta$  und  $\hat{r} = \max\{r_1, r_2, \dots, r_p\}$  definiert. Im Folgenden werden der Pfadparameter und seine zeitlichen Ableitungen zu einem Vektor  $\zeta = [\theta \ \dot{\theta} \ \ddot{\theta} \ \dots \ \theta^{(\hat{r}-1)}]^\top$  zusammengefasst und die Darstellung von (4.71) als ein lineares, zeitinvariantes System

$$\dot{\zeta} = \mathbf{F}\zeta + \mathbf{h}v_\theta \quad (4.72a)$$

mit

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{\hat{r} \times \hat{r}} \quad \text{und} \quad \mathbf{h} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{\hat{r} \times 1} \quad (4.72b)$$

verwendet.

Für den Fall  $m = p - 1$  werden die oben getroffenen Annahmen noch erweitert. Dazu definiert man die Indexmenge  $\mathcal{I} = \{i \mid r_i = \hat{r}\}$ . Basierend auf dieser Indexmenge kann der Vektor  $\mathbf{d}_\theta \in \mathbb{R}^p$  aufgebaut werden, für den gilt, dass

$$d_{\theta,i} = \begin{cases} -\sigma'_i(\theta) & i \in \mathcal{I} \\ 0 & i \notin \mathcal{I} \end{cases} \quad (4.73)$$

Es soll nun nicht nur gelten, dass die Entkopplungsmatrix  $\mathbf{D}$  zum Ausgang  $\mathbf{y}$  in  $\Gamma^*$  den Rang  $m = p - 1$  aufweist, sondern auch, dass die Matrix

$$\begin{bmatrix} \mathbf{D} & \mathbf{d}_\theta \end{bmatrix} \in \mathbb{R}^{(p \times p)} \quad (4.74)$$



den Rang  $p$  besitzt.

Für den Entwurf eines Pfadfolgereglers wird das erweiterte System

$$\dot{\boldsymbol{\chi}} = \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\boldsymbol{\zeta}} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u} \\ \mathbf{F}\boldsymbol{\zeta} + \mathbf{h}v_\theta \end{bmatrix} = \mathbf{f}_e(\boldsymbol{\chi}) + \mathbf{G}_e(\boldsymbol{\chi})\boldsymbol{\nu} \quad (4.75a)$$

mit  $\boldsymbol{\nu}^T = \begin{bmatrix} \mathbf{u}^T & v_\theta \end{bmatrix}$  und dem Fehlerausgang

$$\mathbf{e} = \mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}(\boldsymbol{\zeta}_1) = \mathbf{h}_e(\boldsymbol{\chi}) \quad (4.75b)$$

betrachtet.

**Lemma 4.1.** *Das System (4.75) weist unter den getroffenen Annahmen einen wohldefinierten vektoriellen relativen Grad  $\{r_1, r_2, \dots, r_p\}$  auf.*

**Aufgabe 4.5.** Zeigen Sie Lemma 4.1.

Für den Fall  $m \geq p$  kann unter den getroffenen Annahmen stets eine spezielle Normalform gefunden werden, die den Entwurf von Pfadfolgereglern für parametrisierte Pfade wesentlich vereinfacht. Dieses Ergebnis ist im nachfolgenden Satz formuliert und beruht auf Lemma 4.1 und Satz 4.1.

**Satz 4.2.** *Es sei  $m \geq p$  und alle getroffenen Annahmen seien erfüllt. Dann gilt, dass das System (4.75) äquivalent unter statischer Rückführung zu einem System der Form*

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{v}_P \quad (4.76a)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_\theta) + \mathbf{G}^P(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_\theta)\mathbf{v}_P + \mathbf{G}^R(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_\theta)\mathbf{v}_R \quad (4.76b)$$

$$\dot{\boldsymbol{\eta}}_\theta = \mathbf{F}\boldsymbol{\eta}_\theta + \begin{bmatrix} \mathbf{0} & \mathbf{h} \end{bmatrix} \mathbf{v}_R \quad (4.76c)$$

ist. Dabei gilt  $\boldsymbol{\xi} \in \mathbb{R}^r$ ,  $\boldsymbol{\eta} \in \mathbb{R}^{n-r}$ ,  $\boldsymbol{\eta}_\theta \in \mathbb{R}^{\hat{r}}$ ,  $\mathbf{v}_P \in \mathbb{R}^p$  und  $\mathbf{v}_R \in \mathbb{R}^{m-p+1}$  mit  $r = \sum_{j=1}^p r_j$ . Das Teilsystem (4.76a) liegt in Brunovsky Normalform vor.

*Beweis.* Die Herleitung beruht auf Satz 4.1 und liefert gleichzeitig die Vorschrift, wie die Transformation in die Normalform (4.76) zu errechnen ist. Es wird die spezielle Struktur des Systems (4.75) genutzt, nämlich dass die Dynamiken des Pfadparameters und des ursprünglichen Systems nur über den Fehlerausgang  $\mathbf{e}$  verknüpft sind. Um die Normalform (4.76) zu erhalten, wird dem System (4.75) eine weitere skalare Ausgangsgröße

$$y_\theta = \zeta_1 = h_\theta(\boldsymbol{\chi}) \quad (4.77)$$

hinzugefügt. Es gilt, dass basierend auf Lemma 4.1 auch das System

$$\dot{\boldsymbol{\chi}} = \mathbf{f}_e(\boldsymbol{\chi}) + \mathbf{G}_e(\boldsymbol{\chi})\boldsymbol{\nu} \quad (4.78a)$$

$$\mathbf{e} = \mathbf{h}_e(\boldsymbol{\chi}) \quad (4.78b)$$

$$y_\theta = h_\theta(\boldsymbol{\chi}) \quad (4.78c)$$

einen wohldefinierten vektoriellen relativen Grad  $\{r_1, \dots, r_p, \hat{r}\}$  besitzt.

**Aufgabe 4.6.** Zeigen Sie diese Aussage.

Damit kann direkt Satz 4.1 auf das System (4.78) angewandt werden. Der erste Teil der Zustandstransformation in die Normalform (4.76) lautet

$$\boldsymbol{\xi} = \boldsymbol{\Phi}_{\boldsymbol{\xi}} = \begin{bmatrix} h_{e,1}(\boldsymbol{\chi}) \\ L_{\mathbf{f}_e} h_{e,1}(\boldsymbol{\chi}) \\ \vdots \\ L_{\mathbf{f}_e}^{r_1-1} h_{e,1}(\boldsymbol{\chi}) \\ h_{e,2}(\boldsymbol{\chi}) \\ \vdots \\ L_{\mathbf{f}_e}^{r_2-1} h_{e,2}(\boldsymbol{\chi}) \\ \vdots \\ h_{e,p}(\boldsymbol{\chi}) \\ \vdots \\ L_{\mathbf{f}_e}^{r_p-1} h_{e,p}(\boldsymbol{\chi}) \end{bmatrix} \quad (4.79a)$$

$$\boldsymbol{\eta}_{\theta} = \boldsymbol{\Phi}_{\boldsymbol{\eta}_{\theta}} = \begin{bmatrix} h_{\theta}(\boldsymbol{\chi}) \\ L_{\mathbf{f}_e} h_{\theta}(\boldsymbol{\chi}) \\ \vdots \\ L_{\mathbf{f}_e}^{\hat{r}-1} h_{\theta}(\boldsymbol{\chi}) \end{bmatrix} = \boldsymbol{\zeta}, \quad (4.79b)$$

die mit den Funktionen

$$\boldsymbol{\eta} = \boldsymbol{\Phi}_{\boldsymbol{\eta}} = \begin{bmatrix} \phi_{\eta,1}(\boldsymbol{\chi}) \\ \vdots \\ \phi_{\eta,n-r}(\boldsymbol{\chi}) \end{bmatrix} \quad (4.79c)$$

zu einem Diffeomorphismus komplettiert wird. Die zeitliche Ableitung von (4.79) kann gemäß Satz 4.1 in der Form

$$\frac{d}{dt} \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta}_{\theta} \end{bmatrix} = \bar{\mathbf{A}} \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta}_{\theta} \end{bmatrix} + \bar{\mathbf{B}} \begin{bmatrix} \bar{\mathbf{v}}_1 \\ v_{\theta} \end{bmatrix} \quad (4.80a)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_{\theta}) + \mathbf{G}_1(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_{\theta}) \begin{bmatrix} \bar{\mathbf{v}}_1 \\ v_{\theta} \end{bmatrix} + \mathbf{G}_2(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\eta}_{\theta}) \mathbf{v}_2 \quad (4.80b)$$

mit  $\bar{\mathbf{A}}$  und  $\bar{\mathbf{B}}$  in Brunovsky Normalform dargestellt werden. Durch Setzen von  $\mathbf{v}_p = \bar{\mathbf{v}}_1$  und  $\mathbf{v}_R = \begin{bmatrix} \mathbf{v}_2^T & v_{\theta} \end{bmatrix}^T$  folgt die gewünschte Normalform (4.76). Die Zustandsrückfüh-

rung stellt sich in der Form

$$\boldsymbol{\nu} = \boldsymbol{\alpha}(\boldsymbol{\chi}) + \boldsymbol{\Lambda}(\boldsymbol{\chi}) \begin{bmatrix} \bar{\mathbf{v}}_1 \\ v_\theta \\ \mathbf{v}_2 \end{bmatrix} = \boldsymbol{\alpha}(\boldsymbol{\chi}) + \bar{\boldsymbol{\Lambda}}(\boldsymbol{\chi}) \begin{bmatrix} \mathbf{v}_P \\ \mathbf{v}_R \end{bmatrix} \quad (4.81)$$

dar, wobei die letzte Zeile in (4.81)  $v_\theta = v_\theta$  lautet.  $\square$

Im Fall  $m = p - 1$  kann ebenfalls eine Normalform analog zu (4.76) erhalten werden. Dazu wendet man basierend auf Lemma 4.1 Satz 4.1 direkt auf das System (4.75) an und erhält

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi} + \mathbf{B}\mathbf{v}_P \quad (4.82a)$$

$$\dot{\boldsymbol{\eta}} = \mathbf{f}^0(\boldsymbol{\eta}, \boldsymbol{\xi}) + \mathbf{G}^P(\boldsymbol{\eta}, \boldsymbol{\xi})\mathbf{v}_P \quad (4.82b)$$

mit  $\boldsymbol{\xi} \in \mathbb{R}^r$ ,  $\boldsymbol{\eta} \in \mathbb{R}^{n+\hat{r}-r}$  und  $\mathbf{v}_P \in \mathbb{R}^p$ . Die Normalform (4.82) weist im Gegensatz zu (4.76) weniger Struktur auf. Die interne Dynamik mit den Koordinaten  $\boldsymbol{\eta}$  kann im Allgemeinen nicht mehr in zwei Teile aufgespalten werden, von denen einer direkt dem Hilfssystem entspricht. Beide Normalformen (4.76) und (4.82) können analog zu (4.39) als transversale Normalformen bezeichnet werden. Die Koordinaten  $\boldsymbol{\xi}$  sind wieder der transversalen Dynamik zugeordnet. Dementsprechend repräsentieren die Koordinaten  $\boldsymbol{\eta}$  die tangentielle Dynamik.

Im Fall  $m \geq p$  erkennt man anhand von (4.76), dass sich die Trajektorie des Hilfssystems (und damit des Pfadparameters  $\theta$ ) stets gezielt vorgeben lässt. Da für  $\boldsymbol{\xi} = \mathbf{0}$  unabhängig vom Wert von  $\boldsymbol{\zeta}$  die Ausgangsgrößen  $\mathbf{y}$  exakt am Pfad liegen, ist es sinnvoll die Dynamik des Pfadparameters der tangentialen Dynamik zuzuordnen.

Analog zu Abschnitt 4.3.2 kann auch hier die Menge  $\Gamma^*$  in der Form

$$\Gamma^* = \left\{ \bar{\boldsymbol{\chi}} \in \mathbb{R}^{n+\hat{r}} \mid \boldsymbol{\Phi}_\xi(\bar{\boldsymbol{\chi}}) = \mathbf{0} \right\} \subset \Gamma \quad (4.83)$$

definiert werden. Die Ziele (Z1)–(Z3) können ebenfalls so wie in Abschnitt 4.3.2 durch den Entwurf von Reglern für die transversale und (falls möglich) tangentielle Dynamik erreicht werden.

Bei der Implementierung des resultierenden Pfadfolgereglers muss die Dynamik (4.71) des Pfadparameters im Digitalrechner mitintegriert werden. Dafür sind auch geeignete Anfangswerte  $\boldsymbol{\zeta}(t_0)$  erforderlich. In der Literatur [4.5] wird häufig empfohlen,  $\dot{\theta}(t_0) = \ddot{\theta}(t_0) = \dots = \theta^{(\hat{r}-1)}(t_0) = 0$  und

$$\theta(t_0) = \arg \min_{\theta \in \mathcal{T}} \|\mathbf{h}(\mathbf{x}(t_0)) - \boldsymbol{\sigma}(\theta)\| \quad (4.84)$$

zu wählen. Man beachte aber, dass im Allgemeinen die Wahl von  $\boldsymbol{\zeta}(t_0)$  maßgeblich beeinflusst, ob die Invarianzeigenschaft (Z2) erfüllt ist, siehe dazu auch das Beispiel 4.7.

**Beispiel 4.7 (Fahrzeug mit variabler Vorwärtsgeschwindigkeit).** Als Beispiel dient zunächst wieder das Fahrzeug mit variabler Vorwärtsgeschwindigkeit mit dem mathematischen Modell (4.48). Für dieses gilt  $m = p = 2$ . Es soll ein Pfadfolgeregler für

die Verfolgung einer Lemniskate gegeben durch die parametrisierte Darstellung

$$\boldsymbol{\sigma}_l(\theta) = \begin{bmatrix} \cos(\theta) \\ \cos(\theta) \sin(\theta) \end{bmatrix} \quad (4.85)$$

entworfen werden. Das Ziel (Z3) wird im konkreten Fall so spezifiziert, dass die Lemniskate alle 20 Sekunden komplett abgefahren werden soll. Zunächst wird der vektorielle relative Grad des Systems (4.48) ermittelt. Durch Bilden der Ableitungen

$$\dot{\mathbf{y}} = \begin{bmatrix} v_l \cos(\varphi) \\ v_l \sin(\varphi) \end{bmatrix} \quad (4.86a)$$

$$\ddot{\mathbf{y}} = \underbrace{\begin{bmatrix} -v_l^2 \sin(\varphi) & \cos(\varphi) \\ v_l^2 \cos(\varphi) & \sin(\varphi) \end{bmatrix}}_{\mathbf{D}(\mathbf{x})} \mathbf{u} \quad (4.86b)$$

erhält man einen vektoriellen relativen Grad von  $\{2, 2\}$ . Die Entkopplungsmatrix ist für alle Punkte im Zustandsraum mit  $v_l \neq 0$  regulär. Damit folgen  $\hat{r} = 2$  und das Hilfssystem für den Pfadparameter in der Form

$$\dot{\boldsymbol{\zeta}} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\mathbf{F}} \boldsymbol{\zeta} + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\mathbf{h}} v_\theta \quad (4.87)$$

mit  $\boldsymbol{\zeta} = [\theta \quad \dot{\theta}]^T$ .

Grundlage des Entwurfes ist das erweiterte System

$$\begin{aligned} \dot{\boldsymbol{\chi}} &= \mathbf{f}_e(\boldsymbol{\chi}) + \mathbf{G}_e(\boldsymbol{\chi}) \boldsymbol{\nu} = \mathbf{f}_e(\boldsymbol{\chi}) + \begin{bmatrix} \mathbf{g}_{e,1}(\boldsymbol{\chi}) & \mathbf{g}_{e,2}(\boldsymbol{\chi}) & \mathbf{g}_{e,3}(\boldsymbol{\chi}) \end{bmatrix} \boldsymbol{\nu} \\ &= \begin{bmatrix} v_l \cos \varphi \\ v_l \sin \varphi \\ 0 \\ 0 \\ \zeta_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ v_l & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ v_\theta \end{bmatrix} \end{aligned} \quad (4.88a)$$

mit den Ausgangsgrößen

$$\mathbf{e} = \mathbf{h}_e(\boldsymbol{\chi}) = \mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}_l(\zeta_1) = \begin{bmatrix} x - \cos(\zeta_1) \\ y - \cos(\zeta_1) \sin(\zeta_1) \end{bmatrix} \quad (4.88b)$$

$$y_\theta = h_\theta(\boldsymbol{\chi}) = \zeta_1. \quad (4.88c)$$

Da gleich viele Eingangs- wie Ausgangsgrößen vorliegen und für den vektoriellen relativen Grad  $r_1 + r_2 + \hat{r} = 6$  gilt, reduziert sich die weitere Vorgehensweise auf eine Eingangs-Zustandslinearisierung gemäß [4.1]. Die Koordinaten der transversalen Normalform ergeben sich gemäß (4.79) zu

$$\xi = \begin{bmatrix} x - \cos(\zeta_1) \\ v_l \cos(\varphi) + \sin(\zeta_1)\zeta_2 \\ y - \cos(\zeta_1) \sin(\zeta_1) \\ v_l \sin(\varphi) + \left( (\sin \zeta_1)^2 - (\cos \zeta_1)^2 \right) \zeta_2 \end{bmatrix} \quad \eta_\theta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}. \quad (4.89)$$

Die exakt linearisierende Zustandsrückführung lautet

$$\begin{aligned} \nu &= \begin{bmatrix} L_{g_{e,1}} L_{f_e} h_{e,1}(\chi) & L_{g_{e,2}} L_{f_e} h_{e,1}(\chi) & L_{g_{e,3}} L_{f_e} h_{e,1}(\chi) \\ L_{g_{e,1}} L_{f_e} h_{e,2}(\chi) & L_{g_{e,2}} L_{f_e} h_{e,2}(\chi) & L_{g_{e,3}} L_{f_e} h_{e,2}(\chi) \\ L_{g_{e,1}} L_{f_e} h_\theta(\chi) & L_{g_{e,2}} L_{f_e} h_\theta(\chi) & L_{g_{e,3}} L_{f_e} h_\theta(\chi) \end{bmatrix}^{-1} \left( \begin{bmatrix} \mathbf{v}_P \\ v_\theta \end{bmatrix} - \begin{bmatrix} L_{f_e}^2 h_{e,1}(\chi) \\ L_{f_e}^2 h_{e,2}(\chi) \\ L_{f_e}^2 h_\theta(\chi) \end{bmatrix} \right) \\ &= \begin{bmatrix} \frac{\sin \varphi}{v_l^2} & \frac{\cos \varphi}{v_l^2} & \frac{\cos(\varphi)(2(\cos \zeta_1)^2 - 1) + \sin(\varphi) \sin(\zeta_1)}{v_l^2} \\ \cos \varphi & \sin \varphi & \sin(\varphi)(2(\cos \zeta_1)^2 - 1) - \cos(\varphi) \sin(\zeta_1) \\ 0 & 0 & 1 \end{bmatrix} \left( \begin{bmatrix} \mathbf{v}_P \\ v_\theta \end{bmatrix} - \begin{bmatrix} \cos(\zeta_1)\zeta_2^2 \\ 4 \cos(\zeta_1) \sin(\zeta_1)\zeta_2^2 \\ 0 \end{bmatrix} \right). \end{aligned} \quad (4.90)$$

Diese transformiert gemeinsam mit (4.89) das System (4.88) in die Normalform (4.76), die sich im vorliegenden Fall zu

$$\dot{\xi} = \begin{bmatrix} \xi_2 \\ v_{P,1} \\ \xi_4 \\ v_{P,2} \end{bmatrix} \quad \dot{\eta}_\theta = \begin{bmatrix} \eta_{\theta,2} \\ v_\theta \end{bmatrix} \quad (4.91)$$

ergibt. Da  $r_1 + r_2 = r = 4 = n$  gilt, gibt es keinen Zustand  $\eta$  bzw. keine Dynamik (4.76b). Durch die Verwendung der Regelgesetze

$$v_{P,1} = -k_{P,11}\xi_1 - k_{P,12}\xi_2 \quad (4.92a)$$

$$v_{P,2} = -k_{P,21}\xi_3 - k_{P,22}\xi_4 \quad (4.92b)$$

$$v_\theta = -k_\theta(\eta_{\theta,2} - \dot{\theta}_d) \quad (4.92c)$$

mit  $k_{P,ij}, k_\theta > 0$  kann der Pfad stabilisiert und gleichzeitig mit  $\dot{\theta}_d = \pm \frac{2\pi}{20}$  das Ziel (Z3) erreicht werden.

Für die nachfolgenden Ergebnisse werden die Parameterwerte  $k_{P,11} = k_{P,21} = 0.25$ ,  $k_{P,12} = k_{P,22} = 1$  und  $k_\theta = 0.5$  verwendet. Am Anfangszeitpunkt  $t_0 = 0$  wird  $\zeta_1(0)$  aus (4.84) bestimmt. Damit die Invarianzeigenschaft erfüllt ist, muss zusätzlich nicht

nur das Fahrzeug auf dem Pfad und tangential dazu starten, sondern es muss auch  $\zeta_2(0)$  entsprechend (4.89) so gewählt werden, dass die Gleichungen

$$v_l(0) \cos(\varphi(0)) + \sin(\zeta_1(0))\zeta_2(0) = 0 \quad (4.93a)$$

$$v_l(0) \sin(\varphi(0)) + \left( (\sin \zeta_1(0))^2 - (\cos \zeta_1(0))^2 \right) \zeta_2(0) = 0 \quad (4.93b)$$

erfüllt sind. Da bei den nachfolgend gezeigten Simulationsergebnissen das Fahrzeug ohnehin nicht am Pfad startet, wird, sofern nicht anders angegeben,  $\zeta_2(0) = 0$  gewählt.

In Abbildung 4.7 ist die Trajektorie des Fahrzeuges für den Anfangswert  $\mathbf{x}_0 = \left[ 5.5 \ 1 \ \frac{3\pi}{2} \ 0.45 \right]^T$  dargestellt.

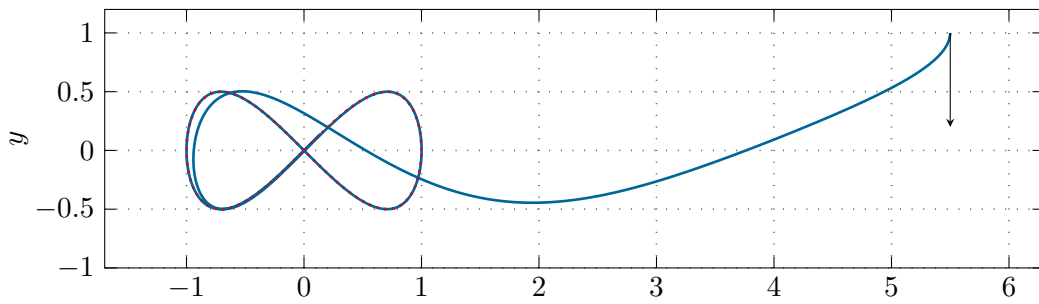


Abbildung 4.7: Trajektorie des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die Pfadfolgeregelung einer Lemniskate.

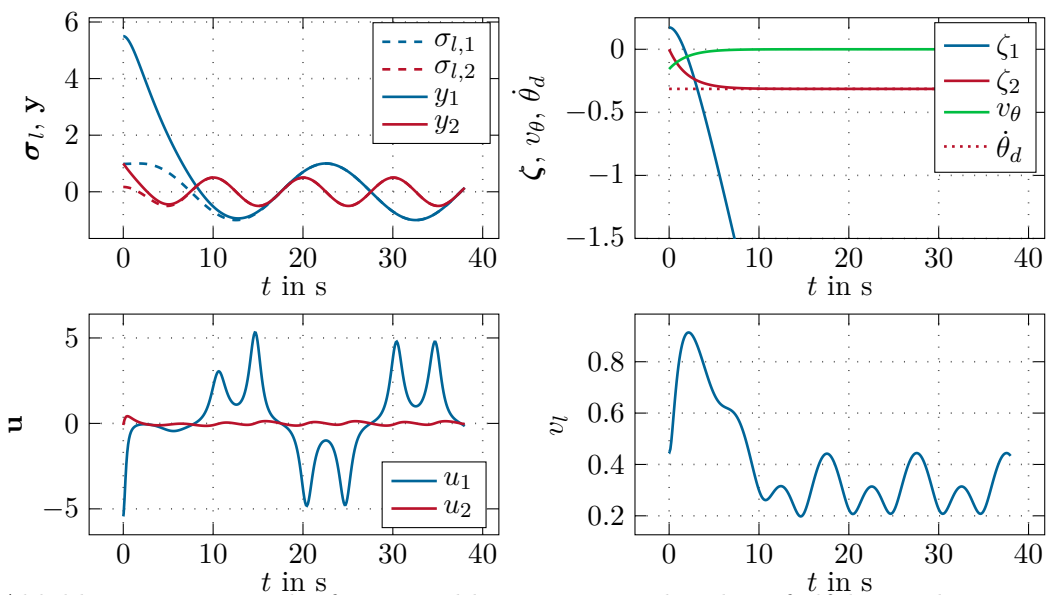


Abbildung 4.8: Zeitverläufe im geschlossenen Kreis bei der Pfadfolgeregelung einer Lemniskate für das Fahrzeug mit  $v_l \neq \text{konst.}$

Die zugehörigen Zeitverläufe im geschlossenen Kreis sind in Abbildung 4.8 ersichtlich. Der Fehlerausgang wird zu null geregelt was durch die Konvergenz von  $\mathbf{y}$  zu  $\boldsymbol{\sigma}_l$  ersichtlich ist. Die zeitliche Ableitung des Pfadparameters  $\dot{\theta} = \zeta_2$  konvergiert zum

gewünschten Wert  $\dot{\theta}_d$ . Dadurch wächst im Laufe der Pfadverfolgung der Wert des Pfadparameters  $\theta = \zeta_1$  über alle Grenzen. Bedingt durch die Tatsache, dass  $\dot{\theta}$  konstant gehalten wird, schwankt die Längsgeschwindigkeit  $v_l$  des Fahrzeuges auch bei exakter Verfolgung der Lemniskate, siehe dazu auch Beispiel 4.8.

Abbildung 4.9 zeigt Trajektorien des Fahrzeuges für verschiedene Anfangswerte.

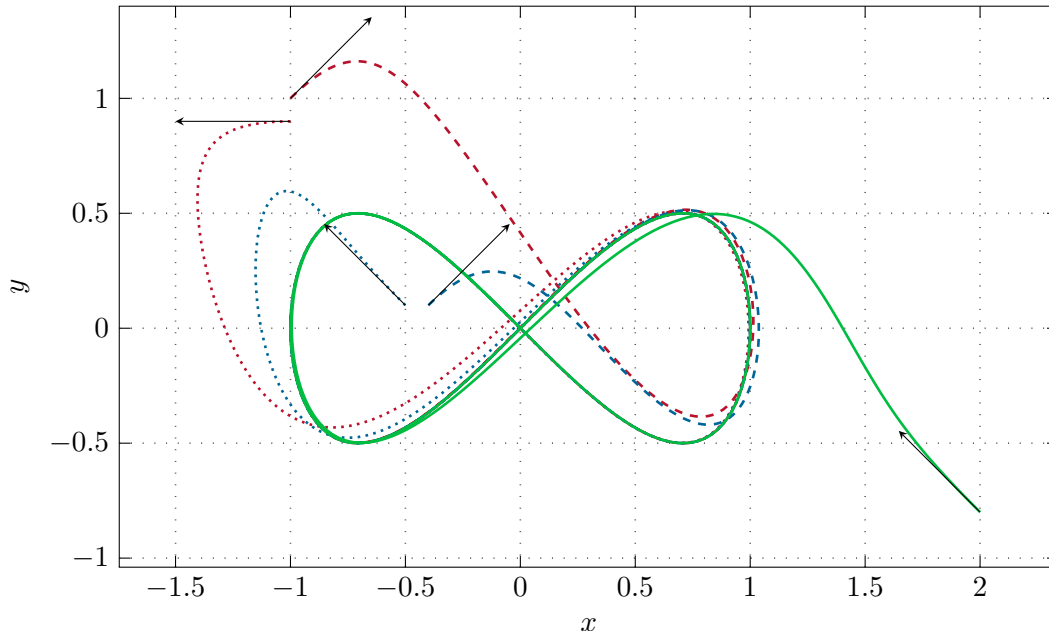


Abbildung 4.9: Trajektorien des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die Pfadfolgeregelung einer Lemniskate und verschiedenen Anfangswerten.

**Beispiel 4.8 (Fahrzeug mit konstanter Vorwärtsgeschwindigkeit).** Zur Illustration des Falles  $m = p - 1$  soll das Fahrzeug mit konstanter Vorwärtsgeschwindigkeit (4.63) herangezogen werden. Das Ziel ist der Entwurf eines Pfadfolgereglers für die Verfolgung einer parametrisierten Lemniskate. Für die Ableitungen der Ausgangsgröße  $\mathbf{y}$  gilt

$$\dot{\mathbf{y}} = \begin{bmatrix} v_l \cos \varphi \\ v_l \sin \varphi \end{bmatrix} \quad (4.94a)$$

$$\ddot{\mathbf{y}} = \underbrace{\begin{bmatrix} -v_l^2 \sin \varphi \\ v_l^2 \cos \varphi \end{bmatrix}}_{\mathbf{d}(\mathbf{x})} \mathbf{u}. \quad (4.94b)$$

Für  $v_l \neq 0$  und beliebige Werte  $\varphi$  gilt, dass nicht beide Komponenten von  $\mathbf{d}$  gleichzeitig verschwinden. Damit gilt  $\text{rang}(\mathbf{d}) = m = 1$  und  $r_1 = r_2 = 2$ . Daraus folgt wie in Beispiel 4.7  $\hat{r} = 2$ . Da der Fall  $m = p - 1$  vorliegt, müssen noch die erweiterten Annahmen aus Abschnitt 4.3.3 überprüft werden. Mit der Indexmenge  $\mathcal{I} = \{1, 2\}$  gilt

$\mathbf{d}_\theta = -\boldsymbol{\sigma}'_l(\theta)$ . Es ist gefordert, dass die erweiterte Matrix

$$\begin{bmatrix} -v_l^2 \sin \varphi & \sin \theta \\ v_l^2 \cos \varphi & (\sin \theta)^2 - (\cos \theta)^2 \end{bmatrix} \quad (4.95)$$

in der Menge  $\Gamma^*$  Rang  $p = 2$  besitzt. In  $\Gamma^*$  gilt, dass sich das Fahrzeug tangential zur Lemniskate bewegen muss, d. h.

$$\varphi = \arctan\left(\frac{\sigma'_{l,2}(\theta)}{\sigma'_{l,1}(\theta)}\right) = \arctan\left(\frac{-(\sin \theta)^2 + (\cos \theta)^2}{-\sin \theta}\right). \quad (4.96)$$

Setzt man diese Beziehung in (4.95) ein, ergibt sich die Determinante der so entstandenen Matrix zu

$$v_l^2 \sqrt{4(\cos \theta)^4 - 5(\cos \theta)^2 + 2}. \quad (4.97)$$

Diese ist für alle  $\theta \in \mathbb{R}$  ungleich null, womit die Regularität der erweiterten Matrix (4.95) entlang des gesamten Pfades nachgewiesen ist.

Das erweiterte System für den Reglerentwurf lautet

$$\dot{\boldsymbol{\chi}} = \mathbf{f}_e(\boldsymbol{\chi}) + \mathbf{G}_e(\boldsymbol{\chi})\boldsymbol{\nu} = \begin{bmatrix} v_l \cos \varphi \\ v_l \sin \varphi \\ 0 \\ \zeta_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ v_l & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v_\theta \end{bmatrix} \quad (4.98a)$$

mit den Ausgangsgrößen

$$\mathbf{e} = \mathbf{h}_e(\boldsymbol{\chi}) = \mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}_l(\zeta_1) = \begin{bmatrix} x - \cos(\zeta_1) \\ y - \cos(\zeta_1) \sin(\zeta_1) \end{bmatrix}. \quad (4.98b)$$

Um den Fehlerausgang  $\mathbf{e}$  auf null zu regeln, wird für (4.98) in weiterer Folge exakte Eingangs-Ausgangslinearisierung vorgenommen. Es gilt

$$\ddot{\mathbf{e}} = \begin{bmatrix} \cos(\zeta_1)\zeta_2^2 \\ 4 \cos(\zeta_1) \sin(\zeta_1)\zeta_2^2 \end{bmatrix} + \begin{bmatrix} -v_l^2 \sin \varphi & \sin \zeta_1 \\ v_l^2 \cos \varphi & (\sin \zeta_1)^2 - (\cos \zeta_1)^2 \end{bmatrix} \boldsymbol{\nu}, \quad (4.99a)$$

wobei die Entkopplungsmatrix natürlich mit der erweiterten Matrix (4.95) übereinstimmt. Deren Regularität wurde bereits nachgewiesen womit die exakt linearisierende Rückführung

$$\boldsymbol{\nu} = \begin{bmatrix} -v_l^2 \sin \varphi & \sin \zeta_1 \\ v_l^2 \cos \varphi & (\sin \zeta_1)^2 - (\cos \zeta_1)^2 \end{bmatrix}^{-1} \left( \mathbf{v}_P - \begin{bmatrix} \cos(\zeta_1)\zeta_2^2 \\ 4 \cos(\zeta_1) \sin(\zeta_1)\zeta_2^2 \end{bmatrix} \right) \quad (4.100)$$



berechnet werden kann. Mit der Wahl  $\eta = \zeta_1$  ergibt sich die Zustandstransformation

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \Phi(\chi) = \begin{bmatrix} x - \cos(\zeta_1) \\ v_l \cos(\varphi) + \sin(\zeta_1)\zeta_2 \\ y - \cos(\zeta_1) \sin(\zeta_1) \\ v_l \sin(\varphi) + ((\sin \zeta_1)^2 - (\cos \zeta_1)^2)\zeta_2 \\ \zeta_1 \end{bmatrix} \quad (4.101)$$

und die transversale Normalform (4.82) zu

$$\dot{\xi}_1 = \xi_2 \quad (4.102a)$$

$$\dot{\xi}_2 = v_{\mathcal{P},1} \quad (4.102b)$$

$$\dot{\xi}_3 = \xi_4 \quad (4.102c)$$

$$\dot{\xi}_4 = v_{\mathcal{P},2} \quad (4.102d)$$

$$\dot{\eta} = f^0(\eta, \xi) . \quad (4.102e)$$

Die interne Dynamik am Pfad lautet

$$\dot{\eta} = f^0(\eta, \mathbf{0}) = \pm \frac{|v_l|}{\sqrt{4(\cos \eta)^4 - 5(\cos \eta)^2 + 2}} . \quad (4.103)$$

Aus (4.103) sieht man unmittelbar, dass am Pfad für  $v_l = \text{konst.}$  die zeitliche Ableitung des Pfadparameters  $\dot{\eta} = \dot{\theta}$  nicht konstant sein kann. Umgekehrt erklärt dies auch, wieso bei konstantem  $\dot{\theta}$  die Längsgeschwindigkeit  $v_l$  des Fahrzeuges schwanken muss, wie es in Beispiel 4.7 ersichtlich ist.

Mit Reglern analog zu (4.92) kann der Pfad stabilisiert werden. Die Trajektorien des Fahrzeuges im geschlossenen Kreis sehen qualitativ sehr ähnlich zu denen von Beispiel 4.7 aus.

## 4.4 Modellprädiktive Pfadfolgeregelung

In diesem Abschnitt soll ein Pfadfolgeregler basierend auf der *modellprädiktiven Regelung* (MPC) entwickelt werden. Grundsätzlich gibt es sehr viele Möglichkeiten, wie MPC für die Pfadfolgeregelung eingesetzt werden kann. Welche Lösung im konkreten Fall gewählt wird, hängt unter anderem von der Systemstruktur ab und ob der Pfad parametrisiert oder implizit definiert ist.

In diesem Abschnitt wird die Formulierung des modellprädiktiven Reglers für den folgenden Fall vorgestellt. Die Systemstruktur sei durch

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad (4.104a)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad (4.104b)$$

mit  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{u} \in \mathbb{R}^m$  und  $\mathbf{y} \in \mathbb{R}^p$  gegeben. Der Einfachheit halber gelte  $m = p$ . Weiters

sollen parametrisierte Pfade  $\mathcal{P}$  der Form

$$\mathcal{P} = \{\bar{\mathbf{y}} \in \mathbb{R}^p \mid \bar{\mathbf{y}} = \boldsymbol{\sigma}(\theta), \theta \in \mathcal{T}\} \quad (4.105)$$

verfolgt werden. Die *Stellgrößen* unterliegen *Beschränkungen*, die durch

$$\mathbf{u}_{\min} \leq \mathbf{u} \leq \mathbf{u}_{\max} \quad (4.106)$$

gegeben sind. Alle Zustände  $\mathbf{x}$  seien direkt messbar bzw. bekannt. Bezüglich der Dynamik des Pfadparameters  $\theta$  wird wieder angenommen, dass dieser über das Hilfssystem

$$\dot{\boldsymbol{\zeta}} = \mathbf{F}\boldsymbol{\zeta} + \mathbf{h}v_\theta \quad (4.107)$$

der Ordnung  $\hat{r}$  vollkommen analog zu (4.72) beschrieben wird. Das Ziel (Z3) der Bewegung auf dem Pfad sei so formulierbar, dass gewisse Komponenten  $\zeta_i$  einen gewünschten Wert  $\zeta_{d,i} = \text{konst.}$  annehmen sollen. Die Ordnung  $\hat{r}$  soll vorerst einfach passend zu dieser Anforderung gewählt sein. Es sei darauf hingewiesen, dass der Schwerpunkt in diesem Abschnitt auf der *Formulierung* des modellprädiktiven Reglers liegt. Es werden *keine Stabilitätsbetrachtungen* durchgeführt.

Als eine Möglichkeit, den Pfad zu stabilisieren, bietet sich folgende MPC-Formulierung an. In jedem Abtastschritt  $t_k = kT_a$  wird das Optimalsteuerungsproblem

$$\min_{\mathbf{u}(\cdot), v_\theta(\cdot)} J(t_k, \mathbf{u}(\cdot), v_\theta(\cdot)) \quad (4.108a)$$

$$\text{u.B.v.} \quad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad \mathbf{x}(t_k) = \mathbf{x}_k \quad (4.108b)$$

$$\dot{\boldsymbol{\zeta}} = \mathbf{F}\boldsymbol{\zeta} + \mathbf{h}v_\theta \quad \boldsymbol{\zeta}(t_k) = \boldsymbol{\zeta}_k \quad (4.108c)$$

$$\mathbf{u}_{\min} \leq \mathbf{u}(t) \leq \mathbf{u}_{\max} \quad \forall t \in [t_k, t_k + T] \quad (4.108d)$$

$$\zeta_1(t) \in \mathcal{T} \quad \forall t \in [t_k, t_k + T] \quad (4.108e)$$

mit

$$J(t_k, \mathbf{u}(\cdot), v_\theta(\cdot)) = \int_{t_k}^{t_k+T} l(\mathbf{x}, \boldsymbol{\zeta}, \mathbf{u}, v_\theta) dt + V(\mathbf{x}(t_k + T), \boldsymbol{\zeta}(t_k + T)) \quad (4.108f)$$

und

$$l(\mathbf{x}, \boldsymbol{\zeta}, \mathbf{u}, v_\theta) = \frac{1}{2} \left[ \left( \mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}(\zeta_1) \right)_{\mathbf{Q}_e} + \left( \boldsymbol{\zeta} - \boldsymbol{\zeta}_d \right)_{\mathbf{Q}_\zeta} + \left( \mathbf{u} \right)_{\mathbf{R}_u} + r_\theta v_\theta^2 \right] \quad (4.108g)$$

$$V(\mathbf{x}, \boldsymbol{\zeta}) = \frac{1}{2} \left[ \left( \mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}(\zeta_1) \right)_{\mathbf{S}_e} + \left( \boldsymbol{\zeta} - \boldsymbol{\zeta}_d \right)_{\mathbf{S}_\zeta} \right] \quad (4.108h)$$

gelöst. Dabei ist  $\left( \mathbf{x} \right)_{\mathbf{Q}}$  eine Kurzschreibweise für die quadratische Form  $\left( \mathbf{x} \right)_{\mathbf{Q}} = \mathbf{x}^T \mathbf{Q} \mathbf{x}$ . Die Matrizen  $\mathbf{Q}_e$ ,  $\mathbf{R}_u$  und  $\mathbf{S}_e$  seien positiv definit. Da wie oben ausgeführt nur gewisse Komponenten  $\zeta_i$  von  $\boldsymbol{\zeta}$  einen konstanten Sollwert  $\zeta_{d,i}$  annehmen sollen, sind die Matrizen  $\mathbf{Q}_\zeta$  und  $\mathbf{S}_\zeta$  im Allgemeinen nur positiv semidefinit. Die restlichen Komponenten von  $\boldsymbol{\zeta}_d$  können beliebig festgesetzt werden. Weiters gelte  $r_\theta > 0$ . Der Optimierungshorizont ist durch  $T \geq T_a$  gegeben. Wie schon in Kapitel 2 werden im Folgenden optimale Lösungen

eines Optimalsteuerungsproblems mit  $*$  gekennzeichnet. Zusätzlich wird im Argument der entsprechenden Funktionen der Abtastzeitpunkt hinzugefügt, an dem sie ermittelt wurden. Dementsprechend lautet die optimale Lösung von (4.108)  $\mathbf{u}^*(\cdot, t_k)$  und  $v_\theta^*(\cdot, t_k)$  mit den zugehörigen Zustandstrajektorien  $\mathbf{x}^*(\cdot, t_k)$  und  $\boldsymbol{\zeta}^*(\cdot, t_k)$ . Der optimale Wert des Kostenfunctionals  $J$  ergibt sich zu  $J^*(t_k) = J(t_k, \mathbf{u}^*(\cdot, t_k), v_\theta^*(\cdot, t_k))$ . Das Optimalsteuerungsproblem (4.108) wird mit den Anfangswerten  $\mathbf{x}_k$  und  $\boldsymbol{\zeta}_k$  gelöst. Der Anfangswert  $\mathbf{x}_k$  entspricht direkt dem Wert des Zustandes des Systems (4.104a) zum Zeitpunkt  $t_k$ . Für das Hilfssystem wird die optimale Lösung aus dem letzten Abtastschritt  $t_{k-1}$ , ausgewertet zum Zeitpunkt  $t_k$ , herangezogen, d. h.  $\boldsymbol{\zeta}_k = \boldsymbol{\zeta}^*(t_k, t_{k-1})$ . Der Anfangswert für  $\boldsymbol{\zeta}$  im ersten Zeitschritt  $t_0$  kann so bestimmt werden wie am Ende von Abschnitt 4.3.3 beschrieben.

Intuitiv ist der modellprädiktive Regler basierend auf (4.108) durch die positiv definite Gewichtung des Fehlers  $\mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}(\zeta_1)$  in der Lage, das Ziel (Z1) zu erfüllen. Darüber hinaus können durch die Gewichtung von  $\boldsymbol{\zeta} - \boldsymbol{\zeta}_d$  die Ziele (Z3) erreicht werden. Das Problem bei der obigen MPC-Formulierung besteht darin, dass die Invarianzeigenschaft (Z2) im Allgemeinen nicht gegeben ist. Dies liegt an der absoluten Gewichtung der Stellgrößen  $\mathbf{u}$ . Eine Gewichtung der Stellgrößen ist für gewöhnlich sinnvoll, um unnötig große Amplituden der Stellgrößen zu vermeiden bzw. um ein gewünschtes Verhalten im geschlossenen Kreis einzustellen. Allerdings ist für eine exakte Verfolgung des Pfades, d. h. für  $\mathbf{h}(\mathbf{x}) \equiv \boldsymbol{\sigma}(\zeta_1)$ , üblicherweise eine Stellgröße  $\mathbf{u} \neq \mathbf{0}$  erforderlich, die nicht konstant ist. Dies bewirkt, dass dadurch unvermeidliche Kosten in  $J$  entstehen, die auch bei stationärer Verfolgung des Pfades nicht verschwinden. Grundsätzlich kann es Fälle geben, für die eine nicht exakte Verfolgung des Pfades ( $\mathbf{h}(\mathbf{x}) \neq \boldsymbol{\sigma}(\zeta_1)$ ) bewirkt, dass die absoluten Stellgrößenamplituden kleiner werden. Damit sinkt der Anteil der Kosten  $\left(\mathbf{u}\right)_{\mathbf{R}_u}$  während die Kosten  $\left(\mathbf{h}(\mathbf{x}) - \boldsymbol{\sigma}(\zeta_1)\right)_{\mathbf{Q}_e}$  steigen. Man kann sich nun vorstellen, dass es einen Kompromiss zwischen diesen beiden Anteilen geben kann, sodass der Wert des Kostenfunctionals kleiner wird. Es kann also Fälle geben, für die es optimaler ist, eine Abweichung vom Pfad in Kauf zu nehmen. Dies bewirkt aber automatisch ein Verletzen der Invarianzeigenschaft (Z2).

Aus obiger Argumentation kann abgeleitet werden, dass es günstiger ist, die Stellgrößen  $\mathbf{u}$  nicht absolut zu gewichten, sondern nur deren Differenz zu den für exakte Verfolgung des Pfades benötigten Stellgrößen. Diese zu berechnen, stellt sich für allgemeine Systemklassen als herausfordernde Aufgabe dar. Daher soll im Folgenden eine *Einschränkung auf flache Systeme* der Form (4.104) vorgenommen werden, für die diese Aufgabe systematisch gelöst werden kann. Dies resultiert in einer MPC-Formulierung, die auch die Invarianzbedingung (Z2) sicherstellt.

Ein flaches System (4.104) zeichnet sich unter anderem dadurch aus, dass stets Parametrisierungen der Zustands- und Stellgrößen in der Form

$$\mathbf{x} = \boldsymbol{\psi}_x(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}, \dots, \mathbf{y}^{(\beta-1)}) \quad (4.109a)$$

$$\mathbf{u} = \boldsymbol{\psi}_u(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}, \dots, \mathbf{y}^{(\beta)}) \quad (4.109b)$$

mit endlichem  $\beta \in \mathbb{N}$  gefunden werden können. Der nächste Schritt besteht darin, den Pfad  $\boldsymbol{\sigma}(\theta)$  mit den zeitlichen Ableitungen

$$\mathbf{y} = \boldsymbol{\sigma}(\theta) \quad (4.110a)$$

$$\dot{\mathbf{y}} = \dot{\boldsymbol{\sigma}}(\theta, \dot{\theta}) = \boldsymbol{\sigma}'(\theta)\dot{\theta} \quad (4.110b)$$

$$\ddot{\mathbf{y}} = \ddot{\boldsymbol{\sigma}}(\theta, \dot{\theta}, \ddot{\theta}) = \boldsymbol{\sigma}''(\theta)\dot{\theta}^2 + \boldsymbol{\sigma}'(\theta)\ddot{\theta} \quad (4.110c)$$

$$\vdots \quad (4.110d)$$

$$\mathbf{y}^{(\beta)} = \boldsymbol{\sigma}^{(\beta)}(\theta, \dot{\theta}, \ddot{\theta}, \dots, \theta^{(\beta)}) \quad (4.110e)$$

in (4.109) einzusetzen. Dies ergibt

$$\mathbf{x} = \boldsymbol{\psi}_x(\boldsymbol{\sigma}(\theta), \boldsymbol{\sigma}'(\theta)\dot{\theta}, \boldsymbol{\sigma}''(\theta)\dot{\theta}^2 + \boldsymbol{\sigma}'(\theta)\ddot{\theta}, \dots, \boldsymbol{\sigma}^{(\beta-1)}(\theta, \dot{\theta}, \ddot{\theta}, \dots, \theta^{(\beta-1)})) \quad (4.111a)$$

$$\mathbf{u} = \boldsymbol{\psi}_u(\boldsymbol{\sigma}(\theta), \boldsymbol{\sigma}'(\theta)\dot{\theta}, \boldsymbol{\sigma}''(\theta)\dot{\theta}^2 + \boldsymbol{\sigma}'(\theta)\ddot{\theta}, \dots, \boldsymbol{\sigma}^{(\beta)}(\theta, \dot{\theta}, \ddot{\theta}, \dots, \theta^{(\beta)})) \quad (4.111b)$$

und motiviert die Wahl  $\hat{r} = \beta + 1$ . Daraus resultiert  $\boldsymbol{\zeta} = [\theta \ \dot{\theta} \ \ddot{\theta} \ \dots \ \theta^{(\beta)}]^\top$  sowie gemäß (4.71)

$$\theta^{(\beta+1)} = v_\theta. \quad (4.112)$$

Durch Substituieren des Pfadparameters und seiner Ableitungen mit den Komponenten von  $\boldsymbol{\zeta}$  erhält man aus (4.111) die neuen Abbildungen

$$\mathbf{x} = \mathbf{p}_x(\boldsymbol{\zeta}) \quad (4.113a)$$

$$\mathbf{u} = \mathbf{p}_u(\boldsymbol{\zeta}). \quad (4.113b)$$

Wenn die Beziehungen (4.113) erfüllt sind, befindet sich das System für beliebige Werte des Pfadparameters und seiner Ableitungen exakt am Pfad  $\mathcal{P}$ . Man beachte, dass (4.113a) die parametrisierte Form der Menge  $\Gamma^*$  darstellt, d. h. es gilt

$$\Gamma^* = \{ \bar{\mathbf{x}} \in \mathbb{R}^n \mid \bar{\mathbf{x}} = \mathbf{p}_x(\boldsymbol{\zeta}), \boldsymbol{\zeta} \in \mathcal{T} \times \mathbb{R}^\beta \}. \quad (4.114)$$

Die Abbildungen (4.113) liefern in Abhängigkeit von  $\boldsymbol{\zeta}$  die Werte der Zustände und der Stellgrößen für die exakte Verfolgung des Pfades im Ausgangsraum. Damit kann eine neue MPC-Formulierung basierend auf dem Optimalsteuerungsproblem

$$\min_{\mathbf{u}(\cdot), v_\theta(\cdot)} J(t_k, \mathbf{u}(\cdot), v_\theta(\cdot)) \quad (4.115a)$$

$$\text{u.B.v.} \quad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad \mathbf{x}(t_k) = \mathbf{x}_k \quad (4.115b)$$

$$\dot{\boldsymbol{\zeta}} = \mathbf{F}\boldsymbol{\zeta} + \mathbf{h}v_\theta \quad \boldsymbol{\zeta}(t_k) = \boldsymbol{\zeta}_k \quad (4.115c)$$

$$\mathbf{u}_{\min} \leq \mathbf{u}(t) \leq \mathbf{u}_{\max} \quad \forall t \in [t_k, t_k + T] \quad (4.115d)$$

$$\boldsymbol{\zeta}_1(t) \in \mathcal{T} \quad \forall t \in [t_k, t_k + T] \quad (4.115e)$$

mit

$$J(t_k, \mathbf{u}(\cdot), v_\theta(\cdot)) = \int_{t_k}^{t_k+T} l(\mathbf{x}, \boldsymbol{\zeta}, \mathbf{u}, v_\theta) dt + V(\mathbf{x}(t_k + T), \boldsymbol{\zeta}(t_k + T)) \quad (4.115f)$$

und

$$l(\mathbf{x}, \boldsymbol{\zeta}, \mathbf{u}, v_\theta) = \frac{1}{2} \left[ \left( \mathbf{x} - \mathbf{p}_x(\boldsymbol{\zeta}) \right)_{\mathbf{Q}_e} + \left( \boldsymbol{\zeta} - \boldsymbol{\zeta}_d \right)_{\mathbf{Q}_\zeta} + \left( \mathbf{u} - \mathbf{p}_u(\boldsymbol{\zeta}) \right)_{\mathbf{R}_u} + r_\theta v_\theta^2 \right] \quad (4.115g)$$

$$V(\mathbf{x}, \zeta) = \frac{1}{2} \left[ \left( \mathbf{x} - \mathbf{p}_x(\zeta) \right)_{\mathbf{S}_e} + \left( \zeta - \zeta_d \right)_{\mathbf{S}_\zeta} \right] \quad (4.115h)$$

entwickelt werden.

Wenn der modellprädiktive Regler basierend auf (4.115) einen asymptotisch stabilen geschlossenen Kreis bewirkt, dann gilt  $J^*(t_k) \rightarrow 0$  für  $k \rightarrow \infty$ . Das bedeutet, dass durch die positiv definite Gewichtung die Differenz  $\mathbf{x} - \mathbf{p}_x(\zeta)$  verschwindet und zumindest die relevanten Komponenten von  $\zeta - \zeta_d$  gegen Null konvergieren. Daher werden auch mit dem modellprädiktiven Regler basierend auf (4.115) die Ziele (Z1) und (Z3) berücksichtigt. Da nun keine absolute Gewichtung von  $\mathbf{u}$  sondern eine relative Gewichtung zu  $\mathbf{p}_u(\zeta)$  verwendet wird, ist im Allgemeinen auch die Invarianzbedingung (Z2) erfüllt. Voraussetzung dafür ist allerdings, dass  $\mathbf{u}_{\min} \leq \mathbf{p}_u(\zeta^*(t, t_k)) \leq \mathbf{u}_{\max} \forall t \in [t_k, t_k + T]$  erfüllt ist, weil nur dann  $\mathbf{u}^*(\cdot, t_k) \equiv \mathbf{p}_u(\zeta^*(\cdot, t_k))$  gelten kann. Grob gesprochen hängt die Erfüllung der Bedingung  $\mathbf{u}_{\min} \leq \mathbf{p}_u(\zeta^*(\cdot, t_k)) \leq \mathbf{u}_{\max}$  von der Größe der Komponenten von  $\zeta^*(\cdot, t_k)$  ab. Daher müssen sowohl die Gewichtungen  $\mathbf{Q}_\zeta$ ,  $\mathbf{S}_\zeta$  und  $r_\theta$  als auch  $\zeta_d$  (d. h. das Ziel (Z3)) so gewählt werden, dass  $\mathbf{u}^*(\cdot, t_k) \equiv \mathbf{p}_u(\zeta^*(\cdot, t_k))$  erfüllbar ist. Für die Erfüllung der Invarianzbedingung muss weiters gelten, dass  $\mathbf{x}(t_0) = \mathbf{p}_x(\zeta(t_0))$  gilt, d. h. insbesondere ist die Wahl von  $\zeta(t_0)$  auch hier wieder entscheidend.

**Beispiel 4.9 (Fahrzeug mit variabler Vorwärtsgeschwindigkeit).** Als Anwendungsbeispiel für die modellprädiktive Pfadfolgeregelung soll wieder das Fahrzeug mit variabler Vorwärtsgeschwindigkeit (4.48) dienen. Für dieses ist die Annahme  $m = p$  erfüllt. Analog zu Beispiel 4.7 soll der Pfadfolgeregler sicherstellen, dass das Fahrzeug einer Lemniskate folgt und diese alle 20 Sekunden komplett abfährt (Ziel (Z3)).

Durch Anschreiben des Systemausganges und Bilden der Ableitungen

$$\mathbf{y} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.116a)$$

$$\dot{\mathbf{y}} = \begin{bmatrix} v_l \cos(\varphi) \\ v_l \sin(\varphi) \end{bmatrix} \quad (4.116b)$$

$$\ddot{\mathbf{y}} = \underbrace{\begin{bmatrix} -v_l^2 \sin(\varphi) & \cos(\varphi) \\ v_l^2 \cos(\varphi) & \sin(\varphi) \end{bmatrix}}_{\mathbf{D}(\mathbf{x})} \mathbf{u} \quad (4.116c)$$

ergibt sich, dass das mathematische Modell des Fahrzeugs (4.48) eingangs-zustands-linearisierbar und damit flach ist. Aus (4.116a) und (4.116b) erhält man zwei Lösungen für die Parametrierung der Zustände

$$\psi_x^+(\mathbf{y}, \dot{\mathbf{y}}) = \begin{bmatrix} y_1 \\ y_2 \\ \arctan\left(\frac{\dot{y}_2}{\dot{y}_1}\right) \\ \sqrt{\dot{y}_1^2 + \dot{y}_2^2} \end{bmatrix} \quad \psi_x^-(\mathbf{y}, \dot{\mathbf{y}}) = \begin{bmatrix} y_1 \\ y_2 \\ \arctan\left(\frac{-\dot{y}_2}{-\dot{y}_1}\right) \\ -\sqrt{\dot{y}_1^2 + \dot{y}_2^2} \end{bmatrix}. \quad (4.117)$$

Aus (4.116c) folgt die Parametrierung der Stellgrößen

$$\boldsymbol{\psi}_u^\pm(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}) = \mathbf{D}^{-1}(\mathbf{x})|_{\mathbf{x}=\boldsymbol{\psi}_x^\pm(\mathbf{y}, \dot{\mathbf{y}})} \ddot{\mathbf{y}}, \quad (4.118)$$

für die nach Einsetzen von (4.117) ebenfalls zwei Lösungen

$$\boldsymbol{\psi}_u^+(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}) = \begin{bmatrix} \frac{\dot{y}_1 \ddot{y}_2 - \dot{y}_2 \ddot{y}_1}{(y_1^2 + y_2^2)^{\frac{3}{2}}} \\ \frac{\dot{y}_1 \ddot{y}_1 + \dot{y}_2 \ddot{y}_2}{\sqrt{y_1^2 + y_2^2}} \end{bmatrix} \quad \boldsymbol{\psi}_u^-(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}) = \begin{bmatrix} -\frac{\dot{y}_1 \ddot{y}_2 - \dot{y}_2 \ddot{y}_1}{(y_1^2 + y_2^2)^{\frac{3}{2}}} \\ -\frac{\dot{y}_1 \ddot{y}_1 + \dot{y}_2 \ddot{y}_2}{\sqrt{y_1^2 + y_2^2}} \end{bmatrix} \quad (4.119)$$

existieren. Aus (4.119) folgt  $\beta = 2$  und damit  $\hat{r} = \beta + 1 = 3$  mit dem Hilfssystem  $\theta^{(3)} = v_\theta$ . Durch Einsetzen des Pfades mit den zugehörigen Ableitungen

$$\boldsymbol{\sigma}_l(\theta) = \begin{bmatrix} \cos(\theta) \\ \cos(\theta) \sin(\theta) \end{bmatrix} \quad \dot{\boldsymbol{\sigma}}_l(\theta, \dot{\theta}) = \dot{\theta} \begin{bmatrix} -\sin \theta \\ -(\sin \theta)^2 + (\cos \theta)^2 \end{bmatrix} \quad (4.120a)$$

$$\ddot{\boldsymbol{\sigma}}_l(\theta, \dot{\theta}, \ddot{\theta}) = \begin{bmatrix} -\dot{\theta}^2 \cos(\theta) - \ddot{\theta} \sin(\theta) \\ -4\dot{\theta}^2 \cos(\theta) \sin(\theta) + \ddot{\theta} (-(\sin \theta)^2 + (\cos \theta)^2) \end{bmatrix} \quad (4.120b)$$

und  $\theta = \zeta_1$ ,  $\dot{\theta} = \zeta_2$ ,  $\ddot{\theta} = \zeta_3$  in (4.117) und (4.119) erhält man

$$\mathbf{p}_x^+(\boldsymbol{\zeta}) = \begin{bmatrix} \cos \zeta_1 \\ \cos(\zeta_1) \sin(\zeta_1) \\ \arctan\left(\frac{\zeta_2(2(\cos \zeta_1)^2 - 1)}{-\sin(\zeta_1)\zeta_2}\right) \\ \sqrt{\zeta_2^2(4(\cos \zeta_1)^4 - 5(\cos \zeta_1)^2 + 2)} \end{bmatrix} \quad (4.121a)$$

$$\mathbf{p}_x^-(\boldsymbol{\zeta}) = \begin{bmatrix} \cos \zeta_1 \\ \cos(\zeta_1) \sin(\zeta_1) \\ \arctan\left(\frac{-\zeta_2(2(\cos \zeta_1)^2 - 1)}{\sin(\zeta_1)\zeta_2}\right) \\ -\sqrt{\zeta_2^2(4(\cos \zeta_1)^4 - 5(\cos \zeta_1)^2 + 2)} \end{bmatrix} \quad (4.121b)$$

sowie  $\mathbf{p}_u^+(\boldsymbol{\zeta})$  und  $\mathbf{p}_u^-(\boldsymbol{\zeta})$ , die auf Grund ihres Umfangs nicht explizit angegeben sind.

Damit das Ziel (Z3) erreicht werden kann, wird

$$\boldsymbol{\zeta}_d = \begin{bmatrix} \text{beliebig} \\ \pm \frac{2\pi}{20} \\ 0 \end{bmatrix} \quad (4.122)$$

gesetzt. Die Lösungen (4.121) für  $\mathbf{p}_x(\boldsymbol{\zeta})$  unterscheiden sich in der Orientierung des Fahrzeuges und dem Vorzeichen der Vorwärtsgeschwindigkeit. Das heißt, der Unterschied besteht darin, ob das Fahrzeug den Pfad vor- oder rückwärts abfährt.

Hinsichtlich einer Verwendung von  $\mathbf{p}_x(\zeta)$  und  $\mathbf{p}_u(\zeta)$  im Optimalsteuerungsproblem stellt sich die Frage, welche Lösung wann verwendet wird. Grundsätzlich soll gefordert werden, dass das Fahrzeug den Pfad in Vorwärtsrichtung, d. h. mit  $v_l > 0$ , abfahren soll. Wenn nun  $\zeta_{d,2} > 0$  gilt, dann ergibt sich im geschlossenen Kreis im eingeschwungenen Zustand ebenfalls  $\zeta_2 > 0$ . Für diesen eingeschwungenen Zustand soll  $v_l > 0$  gelten, womit für  $\zeta_2 > 0$  die Lösungen  $\mathbf{p}_x^+(\zeta)$  und  $\mathbf{p}_u^+(\zeta)$  verwendet werden müssen. Für  $\zeta_2 \leq 0$  kommen  $\mathbf{p}_x^-(\zeta)$  und  $\mathbf{p}_u^-(\zeta)$  zum Einsatz. Im Fall  $\zeta_{d,2} < 0$  gilt im eingeschwungenen Zustand  $\zeta_2 < 0$ . Dafür soll aber ebenfalls  $v_l > 0$  gelten. Daher sind die Verhältnisse für  $\zeta_{d,2} < 0$  genau umgekehrt. Falls  $\zeta_2 < 0$  wird  $\mathbf{p}_x^+(\zeta)$  und  $\mathbf{p}_u^+(\zeta)$  verwendet, ansonsten  $\mathbf{p}_x^-(\zeta)$  und  $\mathbf{p}_u^-(\zeta)$ .

Den im Folgenden gezeigten Simulationsergebnissen liegen die Parameterwerte

$$\mathbf{Q}_e = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix} \quad \mathbf{S}_e = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix} \quad \mathbf{R}_u = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.123a)$$

$$\mathbf{Q}_\zeta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{S}_\zeta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad r_\theta = 1 \quad (4.123b)$$

zugrunde. Man beachte, dass  $\mathbf{Q}_\zeta$  und  $\mathbf{S}_\zeta$  passend zu (Z3) gewählt sind. Die Beschränkungen der Stellgrößen werden zu

$$\mathbf{u}_{\min} = \begin{bmatrix} -5 \\ -3 \end{bmatrix} \quad \mathbf{u}_{\max} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \quad (4.124)$$

festgelegt. Da  $\theta = \zeta_1 \in \mathbb{R}$ , ist die Beschränkung (4.115e) hinfällig. Der Optimierungshorizont sei  $T = 4\text{ s}$  und die Abtastzeit wird zu  $T_a = 50\text{ ms}$  gewählt. Das Optimalsteuerungsproblem (4.115) wird durch ein direktes Verfahren und anschließende Volldiskretisierung [4.6] in ein statisches Optimierungsproblem überführt, vgl. Abschnitt 2.3.2. Dieses statische Optimierungsproblem wird in weiterer Folge mit dem SQP-Verfahren des Softwarepakets SNOPT [4.7] gelöst. Die resultierenden optimalen Stellgrößen werden konstant über das Abtastintervall aufgeschaltet, d. h.  $\mathbf{u}(t) = \mathbf{u}^*(t_k, t_k)$ ,  $v_\theta(t) = v_\theta^*(t_k, t_k)$ ,  $t \in [t_k, t_k + T_a)$ . Die erste Komponente des Anfangswerts  $\zeta_0$  im Abtastschritt  $t_0 = 0$  wird gemäß (4.84) bestimmt. Die zweite und dritte Komponente wird jeweils mit der zweiten und dritten Komponente von  $\zeta_d$  gemäß (4.122) gleichgesetzt.

Abbildung 4.10 zeigt die Zeitverläufe im geschlossenen Kreis für den Anfangswert  $\mathbf{x}_0 = \left[ 3 \quad 1 \quad \frac{3\pi}{2} \quad 0.45 \right]^T$ . Aus der anfänglichen Abweichung konvergieren sowohl die Stellgrößen als auch die Zustände schnell zu  $\mathbf{p}_u$  und  $\mathbf{p}_x$ . Die Stellgrößenbeschränkungen werden mehrmals aktiv. Der optimale Wert des Kostenfunktional  $J^*(\cdot)$  wird von Abtastschritt zu Abtastschritt kleiner.

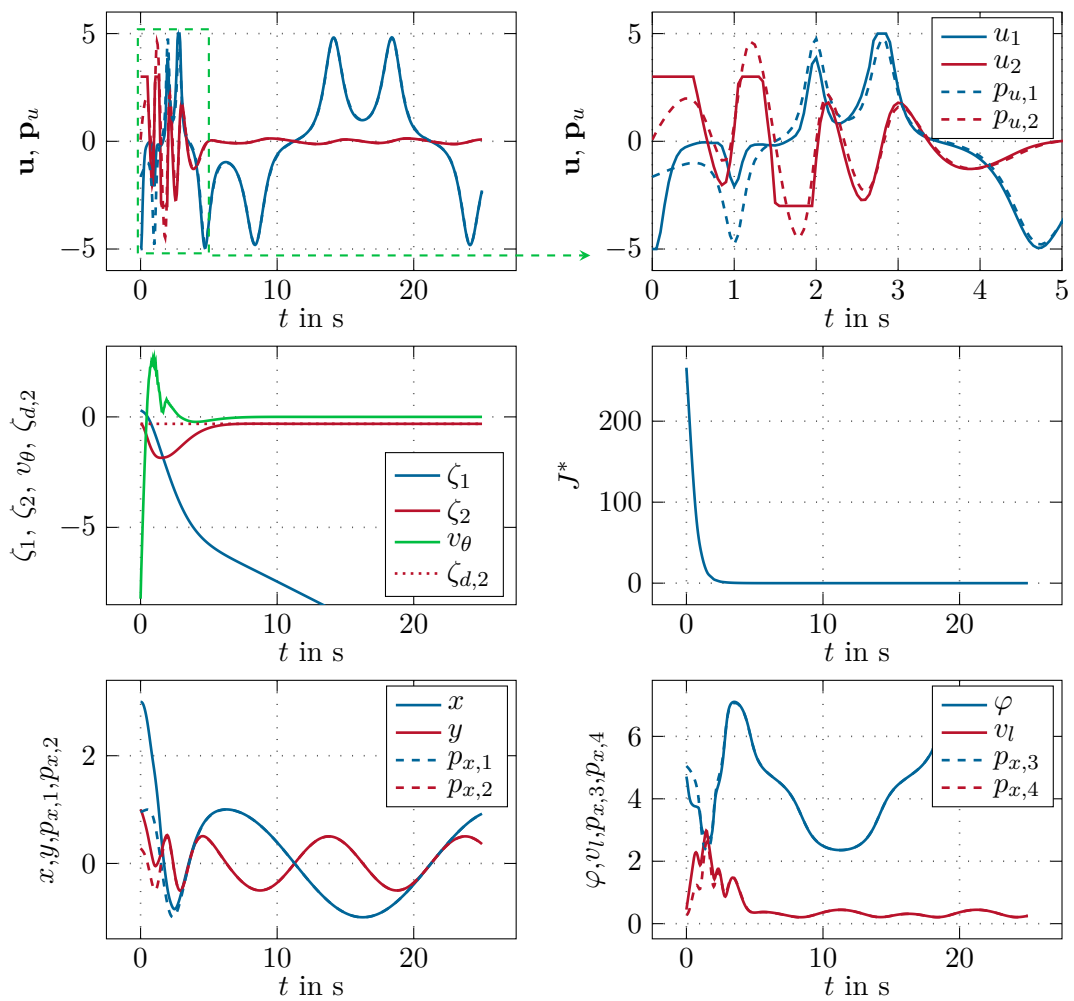


Abbildung 4.10: Zeitverläufe im geschlossenen Kreis bei der modellprädiktiven Pfadfolge-  
regelung einer Lemniskate für das Fahrzeug mit  $v_l \neq \text{konst.}$

In Abbildung 4.11 ist die zugehörige Trajektorie des Fahrzeuges dargestellt.

Die Trajektorien des Fahrzeuges für mehrere verschiedene Anfangswerte sind in  
Abbildung 4.12 zu sehen.



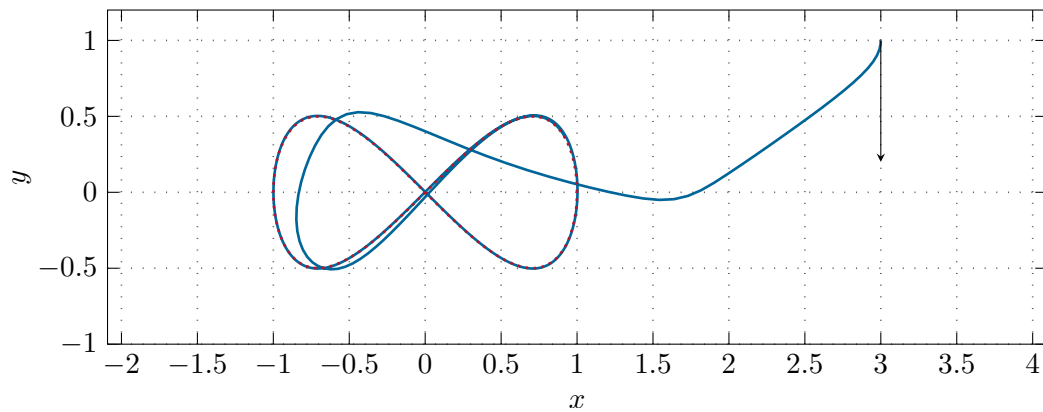


Abbildung 4.11: Trajektorie des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die modellprädiktive Pfadfolgeregelung einer Lemniskate.

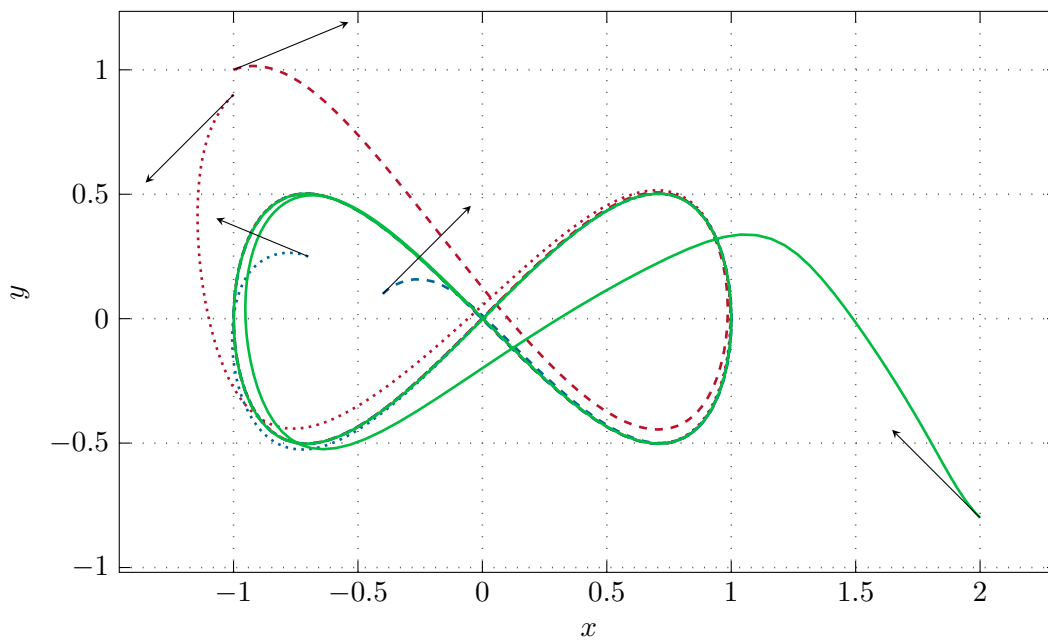


Abbildung 4.12: Trajektorien des Fahrzeuges mit  $v_l \neq \text{konst.}$  für die modellprädiktive Pfadfolgeregelung einer Lemniskate und verschiedenen Anfangswerten.

## 4.5 Literatur

- [4.1] A. Kugi, *Skriptum zur VO Regelungssysteme 2 (SS 2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2019. Adresse: <https://www.acin.tuwien.ac.at/master/regelungssysteme-2/>.
- [4.2] V. A. Toponogov, *Differential Geometry of Curves and Surfaces - A Concise Guide*. Boston: Birkhäuser, 2006.
- [4.3] A. Isidori, *Nonlinear Control Systems*, 3. Aufl. London: Springer, 1995.
- [4.4] C. Nielsen, „Set Stabilization Using Transverse Feedback Linearization,“ Diss., University of Toronto, 2009.
- [4.5] T. Faulwasser, „Optimization-based Solutions to Constrained Trajectory-tracking and Path-following Problems,“ Diss., Otto-von-Guericke-Universität Magdeburg, 2012.
- [4.6] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*, Ser. Advances in Design and Control. Philadelphia, USA: SIAM - Society for Industrial und Applied Mathematics, 2001.
- [4.7] P. E. Gill, W. Murray und M. A. Saunders, *User's Guide for SNOPT Version 7: Software for Large-Scale Nonlinear Programming*, Department of Mathematics, University of California, San Diego, CA, USA, Juni 2008.

## 5 Dissipativität und Passivität

Vereinfachend gesprochen, ist das Konzept der Dissipativität und Passivität die systemtheoretische Verallgemeinerung des Energieerhaltungsprinzips, welches besagt, dass in einem abgeschlossenen System Energie weder erzeugt noch vernichtet werden kann. Eine nähere Betrachtung des systemtheoretischen Konzeptes der Dissipativität wird jedoch zeigen, dass dies a priori mit dem Prinzip der Energieerhaltung nichts zu tun hat und lediglich bei gewissen physikalischen Systemen analoge Aussagen zulässt. Diese Analogie zu physikalischen Systemen trägt aber sicherlich zum Verständnis dieser Konzepte bei, weshalb im Folgenden zwei physikalische Systeme, ein Wärmeübertragungssystem und ein elektromechanisches System, diskutiert werden.

### 5.1 Glüh Simulator

Abbildung 5.1 zeigt die schematische Darstellung eines so genannten Glüh Simulators, der dazu verwendet wird, durch Ohmsches Erwärmen und freie bzw. erzwungene Konvektion (Pressluft oder Ventilator) für Metallproben vorgegebene Temperaturprofile abzufahren.

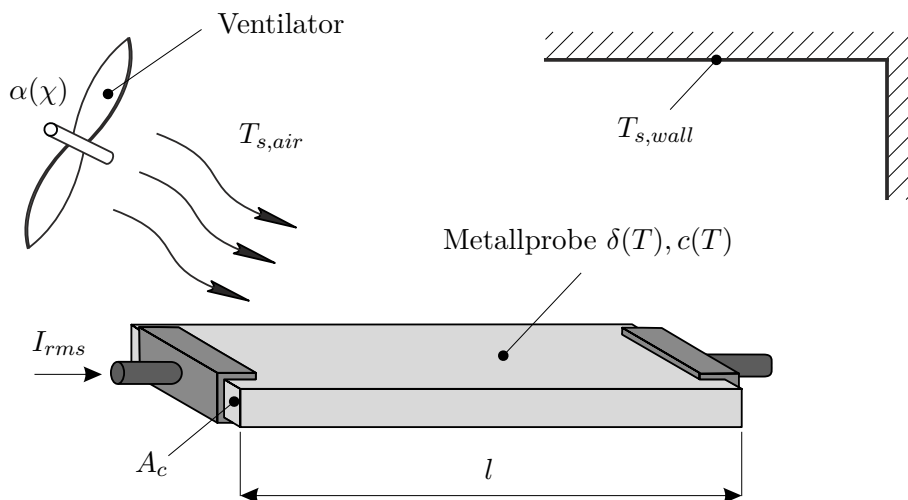


Abbildung 5.1: Glüh Simulator.

Es ist naheliegend für dieses System die elektromechanischen Effekte zu vernachlässigen und die Änderung der im System gespeicherten Energie allein durch die Änderung der thermisch gespeicherten Energie zu erfassen. Das Energieerhaltungsprinzip besagt dann, dass die Änderung der thermisch gespeicherten Energie  $V$  der Beziehung

$$\frac{d}{dt}V = p_{in} - p_{out} \quad (5.1)$$

genügt, wobei  $p_{in}$  und  $p_{out}$  die Energieflüsse in das System und vom System beschreiben. Es wird angenommen, dass die Temperatur  $T$  in der Metallprobe zu jedem Zeitpunkt  $t$  gleichförmig verteilt ist, dass die Oberfläche der Probe sehr klein verglichen mit den umgebenden Wänden ist, und dass die Wärmeleitung vernachlässigt werden kann. Die in der Probe gespeicherte thermische Energie  $V$  lautet

$$V(T) = c(T)mT \quad (5.2)$$

mit der konstanten Probenmasse  $m$  und der spezifischen Wärmekapazität  $c(T)$ . Mit Hilfe des Ohmschen Gesetzes errechnet sich der Energiefluss in die Probe zu

$$p_{in} = I_{rms}^2 \delta(T) \frac{l}{A_c} \quad (5.3)$$

mit dem Effektivwert des durch die Probe fließenden Stromes  $I_{rms}$ , dem spezifischen Widerstand  $\delta(T)$ , der Länge der Probe  $l$  und der Probenquerschnittsfläche  $A_c$ . Die Energieflüsse von der Probe in die Umgebung werden einerseits durch die freie und erzwungene Konvektion

$$p_{out,1} = \alpha(\chi)A_s(T - T_{s,air}) \quad (5.4)$$

und andererseits durch die Wärmestrahlung

$$p_{out,2} = \varepsilon\sigma A_s(T^4 - T_{s,wall}^4) \quad (5.5)$$

verursacht. Dabei bezeichnen  $A_s$  die Oberfläche der Metallprobe,  $T_{s,air}$  und  $T_{s,wall}$  die Temperaturen der umgebenden Luft und Wände,  $\varepsilon$  ist der Emissionsgrad,  $\sigma = 5.67 \cdot 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}$  die Stefan-Boltzmann Konstante und  $\alpha(\chi)$  ist der Konvektionskoeffizient, wobei  $\chi$  im Falle eines Lüfters für die Drehwinkelgeschwindigkeit des Lüfters und im Falle von Druckluft für den Druck steht. Bei freier Konvektion ist  $\alpha(\chi)$  konstant und liegt im Bereich von  $2 - 25 \text{ Wm}^{-2}\text{K}^{-1}$ . Das mathematische Modell des Glühsimulators erhält man einfach durch Einsetzen von (5.2) - (5.5) in (5.1) mit der Zustandsgröße  $T$  und den Eingangsgrößen  $\mathbf{u}^T = [I_{rms}, \chi, T_{s,air}, T_{s,wall}]$ . Integriert man (5.1) entlang einer Lösungskurve vom Zeitpunkt  $t_0 = 0$  zum Zeitpunkt  $t$  für gegebene Eingangsgrößen  $\mathbf{u}(\tau)$ ,  $0 \leq \tau \leq t$ , dann erhält man

$$V(T(t)) - V(T(0)) = \int_0^t s(I_{rms}, \chi, T_{s,air}, T_{s,wall}, T) d\tau \quad (5.6)$$

mit

$$s(I_{rms}, \chi, T_{s,air}, T_{s,wall}, T) = I_{rms}^2 \delta(T) \frac{l}{A_c} - \alpha(\chi)A_s(T - T_{s,air}) - \varepsilon\sigma A_s(T^4 - T_{s,wall}^4) \quad (5.7)$$

Gleichung (5.6) besagt, dass die zum Zeitpunkt  $t$  im System gespeicherte thermische Energie  $V$  gleich der zum Zeitpunkt  $t_0 = 0$  gespeicherten Energie plus oder minus der in dieser Zeit mit der so genannten Versorgungsrate  $s(I_{rms}, \chi, T_{s,air}, T_{s,wall}, T)$  dem System zu- oder abgeführten Energie ist.

## 5.2 Einfaches Elektromagnetventil

Abbildung 5.2 zeigt das Elektromagnetventil mit einem zylindrischen Gehäuse und einem zylindrischen Stößel mit der Masse  $m$  und dem Durchmesser  $D$ . Die aus  $N$  Windungen bestehende Spule mit einem gesamten Innenwiderstand  $R$  wird mit einer Spannung  $U_0$  versorgt. Es wird angenommen, dass der magnetische Widerstand des Gehäuses und des Stößels Null ist, dass die Gleithülse die gleiche Permeabilität wie Luft besitzt und dass für die geometrischen Abmessungen gilt  $h \ll D$  und  $\delta \ll b$  (keine Streuflüsse).

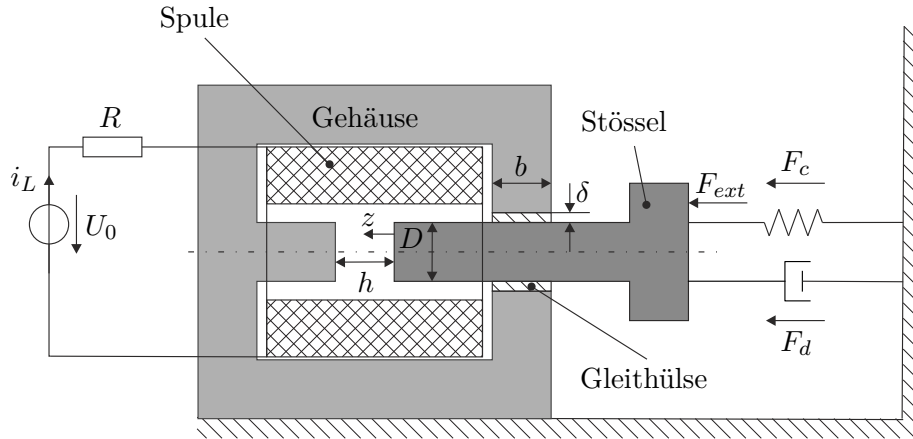


Abbildung 5.2: Einfaches Elektromagnetventil.

Auf analoge Art und Weise zu (5.1) gilt für die Änderung der im System gespeicherten Energie  $V$  die Beziehung

$$\frac{d}{dt}V = p_{in} - p_{out} - p_{diss} \quad (5.8)$$

mit den Energieflüssen  $p_{in}$  und  $p_{out}$ , die über die Systemgrenzen in das System bzw. vom System fließen und mit der in Wärme dissipierten Leistung  $p_{diss}$ .

Unter den obigen Voraussetzungen errechnet sich die im Magnetkreis gespeicherte Koenergie in der Form

$$\check{w}_L = \frac{1}{2}L(z)i_L^2 \quad (5.9)$$

mit der Ersatzinduktivität des magnetischen Kreises

$$L(z) = \frac{\mu_0 N^2 D^2 \pi (D + \delta) \pi b}{4(h - z)(D + \delta) \pi b + \delta D^2 \pi} \quad (5.10)$$

und der Permeabilität von Luft  $\mu_0 = 4\pi \cdot 10^{-7} \text{ VsA}^{-1}\text{m}^{-1}$ .

**Aufgabe 5.1.** Rechnen Sie die Beziehung für die Induktivität  $L(z)$  von (5.10) nach.

Da das betrachtete Elektromagnetventil magnetisch linear ist, sind die Ausdrücke für Energie  $\hat{w}_L$  und Koenergie  $\check{w}_L$  identisch. Die auf den Stößel wirkende Magnetkraft errechnet sich zu

$$F_{mag} = \frac{\partial}{\partial z} \check{w}_L = \frac{1}{2} \frac{\partial L(z)}{\partial z} i_L^2. \quad (5.11)$$

Wie in Abbildung 5.2 gezeichnet, wirkt der Stößel gegen ein lineares Feder-Dämpfer System mit der Dämpfungskraft  $F_d = dv$ ,  $v = \dot{z}$ ,  $d > 0$ , der Federkraft  $F_c = cz(t)$ ,  $c > 0$  und einer externen Kraft  $F_{ext}$ . Das mathematische Modell des Elektromagnetventils lautet dann

$$\frac{d}{dt}z = v \quad (5.12)$$

$$\frac{d}{dt}v = \frac{1}{m} \left( \frac{1}{2} \frac{\partial L(z)}{\partial z} i_L^2 - cz - dv + F_{ext} \right) \quad (5.13)$$

$$\frac{d}{dt}i_L = \frac{1}{L(z)} \left( U_0 - Ri_L - \frac{\partial L(z)}{\partial z} i_L v \right) \quad (5.14)$$

mit den Zustandsgrößen  $\mathbf{x}^T = [z, v, i_L]$  und den Eingangsgrößen  $\mathbf{u}^T = [U_0, F_{ext}]$ .

Die im System gespeicherte Energie setzt sich nun aus der magnetischen Energie (5.9), der kinetischen Energie des Stößels und der potenziellen Energie der Feder

$$V = \frac{1}{2} \left( L(z) i_L^2 + mv^2 + cz^2 \right) \quad (5.15)$$

zusammen. Die Änderung der gespeicherten Energie  $V$  entlang einer Lösungskurve ergibt sich in der Form

$$\frac{d}{dt}V = \underbrace{U_0 i_L + F_{ext} v}_{p_{in} - p_{out}} - \underbrace{\left( dv^2 + Ri_L^2 \right)}_{p_{diss}}. \quad (5.16)$$

Integriert man nun wieder (5.16) entlang einer Lösungskurve vom Zeitpunkt  $t_0 = 0$  zum Zeitpunkt  $t$  für gegebene Eingangsgrößen  $\mathbf{u}(\tau)$ ,  $0 \leq \tau \leq t$ , dann erhält man wegen  $p_{diss} \geq 0$

$$V(\mathbf{x}(t)) - V(\mathbf{x}(t_0)) \leq \int_{t_0}^t s(U_0, F_{ext}, i_L, v) d\tau \quad (5.17)$$

mit der Versorgungsrate

$$s(U_0, F_{ext}, i_L, v) = U_0 i_L + F_{ext} v. \quad (5.18)$$

## 5.3 Systemtheoretisches Konzept

### 5.3.1 Dissipativität

Den nachfolgenden Betrachtungen liege ein nichtlineares dynamisches System der Form

$$\begin{aligned} \frac{d}{dt}\mathbf{x} &= \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= \mathbf{h}(\mathbf{x}, \mathbf{u}) \end{aligned} \quad (5.19)$$

mit dem Zustand  $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^n$ , dem Stelleingang  $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$  und dem Ausgang  $\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^p$  zu Grunde. Es sei angenommen, dass der Zustand  $\mathbf{x}(t)$  zu jedem Zeitpunkt  $t$  eindeutig durch die Wahl der Eingangsgröße  $\mathbf{u}(t)$  und des Anfangszustandes  $\mathbf{x}(0) = \mathbf{x}_0$ , bestimmt ist. Dies erlaubt es, die so genannte Versorgungsrate  $s(\mathbf{u}, \mathbf{y}) : \mathcal{U} \times \mathcal{Y} \rightarrow \mathbb{R}$ , eine

reellwertige Funktion, die für alle Anfangswerte  $\mathbf{x}_0 \in \mathcal{X}$  und alle Eingangsgrößen  $\mathbf{u}$  die Bedingung

$$\int_0^t |s(\mathbf{u}, \mathbf{y})| d\tau < \infty \quad (5.20)$$

für alle Zeiten  $t \geq 0$  erfüllt, einzuführen.

**Definition 5.1.** Das System (5.19) heißt *dissipativ bezüglich der Versorgungsrate  $s$* , wenn eine nichtnegative Funktion  $V(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}$  so existiert, dass die so genannte *integrale Dissipativitätsungleichung*

$$V(\mathbf{x}(t)) - V(\mathbf{x}(0)) \leq \int_0^t s(\mathbf{u}(\tau), \mathbf{y}(\tau)) d\tau \quad (5.21)$$

für alle Anfangswerte  $\mathbf{x}(0) \in \mathcal{X}$  und alle Eingangsgrößen  $\mathbf{u}(t)$  für alle Zeiten  $t \geq 0$  erfüllt ist. Die Funktion  $V(\mathbf{x})$  wird als *Speicherfunktion* bezeichnet. Falls in (5.21) das Gleichheitszeichen gilt, nennt man das System (5.19) *verlustlos bezüglich der Versorgungsrate  $s$* .

Im Sinne dieser Definition ist der Glühsimulator von Abbildung 5.1 verlustlos bezüglich der Versorgungsrate (5.7) und das Elektromagnetventil von Abbildung 5.2 ist dissipativ bezüglich der Versorgungsrate (5.18). Wenn die Speicherfunktion  $V(\mathbf{x})$  bezüglich  $\mathbf{x}$  stetig differenzierbar ist, dann kann man die Änderung von  $V(\mathbf{x})$  entlang einer Lösungskurve von (5.19) berechnen und man erhält die so genannte differenzielle Dissipativitätsungleichung

$$\frac{d}{dt} V(\mathbf{x}) \leq s(\mathbf{u}(t), \mathbf{y}(t)) \quad (5.22)$$

für alle Zeiten  $t \geq 0$ .

### 5.3.2 Passivität

Die Passivität kann als Spezialfall der Dissipativität aufgefasst werden. Zur Definition betrachte man wiederum das System (5.19), wobei nun die Dimension des Systemeingangs  $m$  gleich der Dimension des Ausgangs  $p$  ist.

**Definition 5.2.** Das System (5.19) mit  $m = p$  nennt man *passiv*, wenn eine Konstante  $\delta$  so existiert, dass die Ungleichung

$$\int_0^t \mathbf{y}^T \mathbf{u} d\tau \geq \delta \quad (5.23)$$

für alle zulässigen Eingangsgrößen  $\mathbf{u}(t)$  und alle  $t \geq 0$  erfüllt ist.

Wenn darüberhinaus für geeignete reelle Konstanten  $\alpha, \beta$  die Ungleichung

$$\int_0^t \mathbf{y}^T \mathbf{u} d\tau \geq \delta + \alpha \int_0^t \mathbf{u}^T \mathbf{u} d\tau \quad \text{bzw.} \quad \int_0^t \mathbf{y}^T \mathbf{u} d\tau \geq \delta + \beta \int_0^t \mathbf{y}^T \mathbf{y} d\tau \quad (5.24)$$

für alle zulässigen Eingangsgrößen  $\mathbf{u}(t)$  und alle  $t \geq 0$  erfüllt ist, dann nennt man das System  $\alpha$ -eingangspassiv bzw.  $\beta$ -ausgangspassiv.

Offensichtlich muss  $\delta \leq 0$  gelten, denn die Ungleichung (5.23) muss auch für die Eingangsgröße  $\mathbf{u}(t) = \mathbf{0}$  gültig sein.

**Satz 5.1 (Verbindung Passivität und Dissipativität).** *Existiert nun für das System (5.19) mit  $m = p$  eine nichtnegative Funktion  $V(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}$  so, dass gilt (integrale Passivitätsungleichung)*

$$V(\mathbf{x}(t)) - V(\mathbf{x}(0)) \leq \int_0^t \mathbf{y}^T \mathbf{u} d\tau \quad (5.25)$$

für alle zulässigen Eingangsgrößen  $\mathbf{u}(t)$ , alle  $\mathbf{x}(0)$  und alle  $t \geq 0$ , dann ist das System (5.19) vom Eingang  $\mathbf{u}$  zum Ausgang  $\mathbf{y}$  passiv. Offensichtlich ist dies gemäß Definition 5.1 äquivalent dazu, dass das System (5.19) bezüglich der speziellen bilinearen Versorgungsrate  $s(\mathbf{u}, \mathbf{y}) = \langle \mathbf{y}, \mathbf{u} \rangle = \mathbf{y}^T \mathbf{u}$  dissipativ ist. Ist darüberhinaus das System (5.19) bezüglich der Versorgungsrate  $s(\mathbf{u}, \mathbf{y}) = \mathbf{y}^T \mathbf{u} - \alpha \|\mathbf{u}\|^2$  bzw.  $s(\mathbf{u}, \mathbf{y}) = \mathbf{y}^T \mathbf{u} - \beta \|\mathbf{y}\|^2$  für geeignete reelle Konstanten  $\alpha, \beta$  dissipativ, so ist (5.19)  $\alpha$ -eingangspassiv bzw.  $\beta$ -ausgangspassiv. Ein verlustloses passives System nennt man in diesem Zusammenhang auch ein konservatives System.

*Beweis.* Der Beweis des Satzes ist trivial, da wegen  $V(\mathbf{x}) \geq 0$  aus (5.25) unmittelbar folgt

$$\int_0^t \mathbf{y}^T \mathbf{u} d\tau \geq -V(\mathbf{x}(0)) = \delta. \quad (5.26)$$

□

Mit dieser Definition erkennt man unmittelbar, dass das Elektromagnetventil von Abbildung 5.2 mit dem Eingang  $\mathbf{u}^T = [U_0, F_{ext}]$  und dem Ausgang  $\mathbf{y}^T = [i_L, v]$  passiv, ja sogar  $\beta$ -ausgangspassiv mit  $0 < \beta < \min(d, R)$  ist, da für die dissipierte Leistung von (5.16) gilt  $p_{diss} = dv^2 + Ri_L^2 \geq \beta \|\mathbf{y}\|^2$ .

Die physikalische Interpretation der Passivitätsungleichung (5.25) lautet nun wie folgt: Gibt der Ausdruck  $\mathbf{y}^T \mathbf{u}$  eine Leistung an (z.B. geeignete Paare von Strömen und Spannungen bei elektrischen Systemen oder kollokierte Geschwindigkeiten und Kräfte bei mechanischen Systemen) und ist  $V(\mathbf{x})$  die im System gespeicherte Energie, so besagt die Passivitätsungleichung (5.25), dass die Zunahme der im System gespeicherten Energie kleiner oder gleich der dem System zugeführten Energie ist.

**Aufgabe 5.2.** Zeigen Sie, dass der Integrator mit der Zustandsdarstellung

$$\begin{aligned} \frac{d}{dt} x &= u \\ y &= x \end{aligned} \quad (5.27)$$

passiv ist.



**Aufgabe 5.3.** Unter welchen Voraussetzungen an die Parameter  $\sigma_0, \sigma_1, \sigma_2, r_C, r_H$  und  $v_0$  beschreibt das *LuGre-Reibmodell* (siehe z. B. Skriptum zur VO Regelungssysteme 2 [5.1]) ein passives System vom Eingang  $\Delta v$  zum Ausgang  $F_R$ . Zur Wiederholung soll das LuGre-Reibmodell nochmals in der Form

$$\begin{aligned} \frac{d}{dt}z &= \Delta v - \frac{\text{abs}(\Delta v)}{\chi(\Delta v)}\sigma_0 z \\ F_R &= \sigma_0 z + \sigma_1 \frac{d}{dt}z + \sigma_2 \Delta v \end{aligned} \quad (5.28)$$

mit

$$\chi(\Delta v) = r_C + (r_H - r_C) \exp\left(-\left(\frac{\Delta v}{v_0}\right)^2\right) \quad (5.29)$$

angeschrieben werden.

**Aufgabe 5.4.** Zeigen Sie, dass eine nichtlineare Kennlinie  $y = \psi(u)$ , die die Sektorbedingung  $k_1 u^2 \leq \psi(u)u \leq k_2 u^2$  erfüllt,  $k_1$ -eingangspassiv und  $\left(\frac{1}{k_2}\right)$ -ausgangspassiv gemäß Definition 5.2 ist.

### 5.3.3 Eigenschaften Passiver Systeme

Passive Systeme haben nun die bemerkenswerte Eigenschaft, dass die Parallelschaltung und die Rückkopplung passiver Systeme, wie in Abbildung 5.3 dargestellt, wiederum passiv ist.

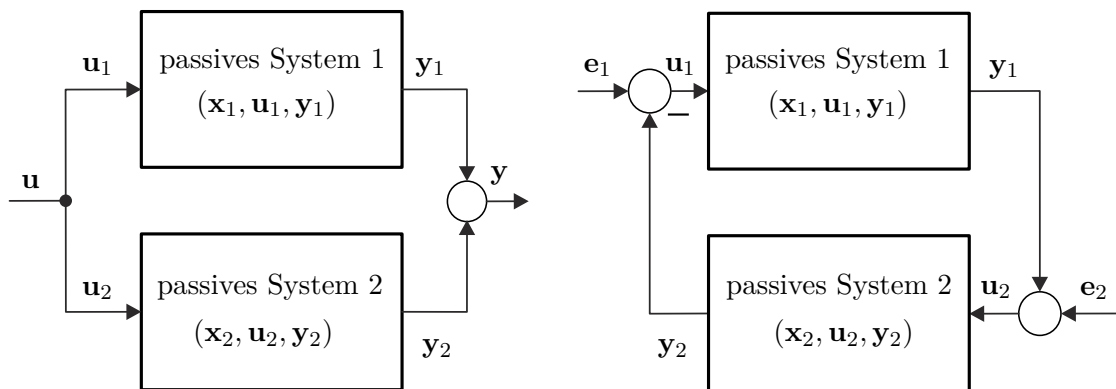


Abbildung 5.3: Parallelschaltung und Rückkopplung zweier passiver Systeme.

*Beweis.* Um dies zu zeigen, nimmt man zwei passive Systeme der Form (5.19) mit  $m = p$  an. Für diese existieren dann zwei nichtnegative Speicherfunktionen  $V_1(\mathbf{x}_1)$

und  $V_2(\mathbf{x}_2)$ , die den Passivitätsungleichungen

$$\begin{aligned} V_1(\mathbf{x}_1(t)) - V_1(\mathbf{x}_1(0)) &\leq \int_0^t \mathbf{y}_1^T \mathbf{u}_1 \, d\tau \\ V_2(\mathbf{x}_2(t)) - V_2(\mathbf{x}_2(0)) &\leq \int_0^t \mathbf{y}_2^T \mathbf{u}_2 \, d\tau \end{aligned} \quad (5.30)$$

genügen. Für die Parallelschaltung nach Abbildung 5.3 gilt  $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{u}$ ,  $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$  und damit

$$V_1(\mathbf{x}_1(t)) + V_2(\mathbf{x}_2(t)) - V_1(\mathbf{x}_1(0)) - V_2(\mathbf{x}_2(0)) \leq \int_0^t (\mathbf{y}_1^T + \mathbf{y}_2^T) \mathbf{u} \, d\tau \quad (5.31)$$

bzw.

$$V(\mathbf{x}(t)) - V(\mathbf{x}(0)) \leq \int_0^t \mathbf{y}^T \mathbf{u} \, d\tau \quad (5.32)$$

mit der nichtnegativen Speicherfunktion  $V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2)$  und dem Zustand  $\mathbf{x}^T = [\mathbf{x}_1^T, \mathbf{x}_2^T]$ .  $\square$

**Aufgabe 5.5.** Zeigen Sie, dass der geschlossene Kreis der Rückkopplung zweier passiver Systeme (siehe Abbildung 5.3, rechtes Bild) vom Eingang  $(\mathbf{e}_1, \mathbf{e}_2)$  zum Ausgang  $(\mathbf{y}_1, \mathbf{y}_2)$  passiv ist.

Darüberhinaus ist auch die Hintereinanderschaltung zweier passiver Systeme gemäß Abbildung 5.4 passiv, sofern das Verbindungssystem energieerhaltend ist, d.h. folgende Zusammenschaltungsbedingung

$$\int_0^t (\mathbf{y}_1^T \mathbf{u}_I + \mathbf{y}_2^T \mathbf{y}_I) \, d\tau = 0 \quad (5.33)$$

erfüllt ist.

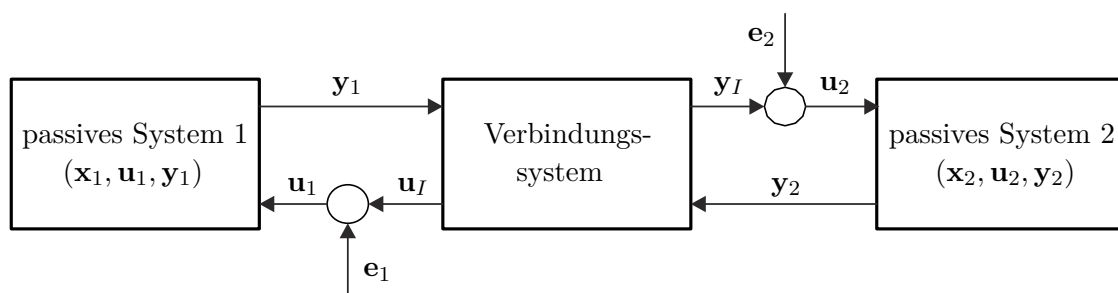


Abbildung 5.4: Hintereinanderschaltung passiver Systeme.

Man überzeugt sich leicht, dass dies der Fall ist, da die nachfolgende Passivitätsungleichung

$$V(\mathbf{x}(t)) - V(\mathbf{x}(0)) \leq \int_0^t (\mathbf{y}_1^T \mathbf{e}_1 + \mathbf{y}_2^T \mathbf{e}_2) \, d\tau \quad (5.34)$$

mit  $V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2)$  und  $\mathbf{x}^T = [\mathbf{x}_1^T, \mathbf{x}_2^T]$  gilt. Gerade diese Eigenschaft wird bei gewissen passivitätsbasierten Reglerentwurfsverfahren genutzt, wobei das System 1 einer passiven Strecke und das System 2 einem passiven Regler entspricht. Für das Verbindungssystem wird in diesem Fall ein System der Form

$$\begin{bmatrix} \mathbf{u}_I \\ \mathbf{y}_I \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{U}_I(\mathbf{x}) \\ -\mathbf{U}_I^T(\mathbf{x}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad (5.35)$$

mit einer vorerst beliebigen quadratischen Matrix  $\mathbf{U}_I(\mathbf{x})$  gewählt.

**Aufgabe 5.6.** Zeigen Sie, dass (5.35) die Zusammenschaltungsbedingung (5.33) erfüllt.

### 5.3.4 Passivität und Lyapunov-Stabilität

Es sei angenommen, dass das System (5.19) passiv mit einer stetig differenzierbaren, positiv definiten Speicherfunktion  $V(\mathbf{x})$  ist. Dann folgt unmittelbar aus der Passivitätsungleichung (5.25) in ihrer differentiellen Form

$$\frac{d}{dt}V(\mathbf{x}) \leq \mathbf{y}^T \mathbf{u}, \quad (5.36)$$

dass die Ruhelage  $\mathbf{x} = \mathbf{0}$  des freien Systems (5.19), also für  $\mathbf{u} = \mathbf{0}$ , stabil im Sinne von Lyapunov ist mit der Lyapunovfunktion  $V(\mathbf{x})$ . Ob die Ruhelage asymptotisch stabil ist, muss von Fall zu Fall mit Hilfe des Invarianzprinzips von Krassovskii-LaSalle untersucht werden.

Für die Rückkopplung zweier passiver Systeme, wie sie im rechten Teil von Abbildung 5.3 gezeigt ist, kann die asymptotische Stabilität der Ruhelage des freien geschlossenen Kreises, also für  $\mathbf{e}_1 = \mathbf{e}_2 = \mathbf{0}$ , auf Eigenschaften der Teilsysteme zurückgeführt werden.

**Satz 5.2.** *Angenommen, die Ruhelage  $\mathbf{x}_1 = \mathbf{0}$  des Teilsystems 1 ist asymptotisch stabil und  $\alpha$ -eingangspassiv gemäß Definition 5.2 mit einer stetig differenzierbaren, positiv definiten Speicherfunktion  $V_1(\mathbf{x}_1)$ . Weiters sei das Teilsystem 2 nullzustandsermittelbar und  $\beta$ -ausgangspassiv gemäß Definition 5.2 mit einer stetig differenzierbaren, positiv definiten Speicherfunktion  $V_2(\mathbf{x}_2)$ . Die Ruhelage des geschlossenen Kreises  $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{0}, \mathbf{0})$  ist dann asymptotisch stabil, wenn  $\alpha + \beta > 0$  gilt.*

Bevor dieser Satz gezeigt wird, sollen noch die Begriffe der Nullzustandsermittelbarkeit und Nullzustandsbeobachtbarkeit definiert werden.

**Definition 5.3.** Das System (5.19) heißt nullzustandsermittelbar (nullzustandsbeobachtbar), wenn aus  $\mathbf{u}(t) = \mathbf{0}$  und  $\mathbf{y}(t) = \mathbf{0}$  für alle Zeiten  $t \geq 0$  folgt  $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{0}$  ( $\mathbf{x}(t) = \mathbf{0}$  für alle Zeiten  $t \geq 0$ ).

*Beweis.* Zum Beweis von Satz 5.2 wähle man als Lyapunovfunktion des geschlossenen Kreises  $V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2)$  und bilde deren zeitliche Ableitung

$$\frac{d}{dt}V(\mathbf{x}) \leq -(\alpha + \beta)\|\mathbf{y}_2\|^2. \quad (5.37)$$

Da aber nach Satz 5.2  $\alpha + \beta > 0$  ist, folgt unmittelbar, dass die Ruhelage des geschlossenen Kreises  $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{0}, \mathbf{0})$  stabil im Sinne von Lyapunov ist. Aufgrund

der Nullzustandsermittelbarkeit des Teilsystems 2 und der asymptotischen Stabilität der Ruhelage  $\mathbf{x}_1 = \mathbf{0}$  des Teilsystems 1 kann man zeigen, dass die größte positiv invariante Menge, die in  $\mathcal{H} = \left\{ \mathbf{x} \in \mathcal{X} \mid \frac{d}{dt} V(\mathbf{x}) = 0 \right\}$  enthalten ist, der Ursprung  $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{0}, \mathbf{0})$  ist. Damit ist aber nach dem Invarianzprinzip von Krassovskii-LaSalle die Ruhelage des geschlossenen Kreises  $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{0}, \mathbf{0})$  asymptotisch stabil.  $\square$

Satz 5.2 wird im Zusammenhang mit dem Begriff der *absoluten Stabilität* benötigt, insbesondere zur Herleitung des *Kreis- und Popov-Kriteriums*.

## 5.4 Lineare passive Systeme

Für ein lineares zeitinvariantes System der Form

$$\begin{aligned} \frac{d}{dt} \mathbf{x} &= \mathbf{A} \mathbf{x} + \mathbf{b} u \\ y &= \mathbf{c}^T \mathbf{x} + d u \end{aligned} \quad (5.38)$$

lässt sich die Eigenschaft der Passivität auch an Hand der zugehörigen Übertragungsfunktion

$$G(s) = \frac{\hat{y}(s)}{\hat{u}(s)} = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d \quad (5.39)$$

beurteilen. Ohne Einschränkung der Allgemeinheit werden hier nur Eingrößensysteme behandelt, für Mehrgrößensysteme sei auf die am Ende angeführte Literatur verwiesen. Gemäß Definition 5.2 ist das System (5.38) genau dann passiv, wenn folgende Ungleichung

$$\int_0^t y u d\tau \geq 0 \quad (5.40)$$

erfüllt ist. Damit lässt sich folgender Satz für die Passivität linearer zeitinvarianter Eingrößensysteme angeben:

**Satz 5.3.** *Das lineare zeitinvariante System (5.38) mit der Übertragungsfunktion  $G(s)$  von (5.39) ist*

(1) *genau dann passiv, wenn gilt*

$$\operatorname{Re}(G(j\omega)) \geq 0 \quad \text{für alle } \omega, \quad (5.41)$$

(2) *genau dann  $\alpha$ -eingangspassiv mit  $\alpha > 0$ , wenn gilt*

$$\operatorname{Re}(G(j\omega)) \geq \alpha > 0 \quad \text{für alle } \omega \quad (5.42)$$

(3) *und genau dann  $\beta$ -ausgangspassiv mit  $\beta > 0$ , wenn gilt*

$$\operatorname{Re}(G(j\omega)) \geq \beta |G(j\omega)|^2 > 0 \quad \text{für alle } \omega. \quad (5.43)$$

*Man beachte, dass die Überprüfung der Bedingungen (5.41) - (5.43) sehr einfach an Hand der Nyquist-Ortskurve von  $G(s)$  möglich ist.*

*Beweis.* Zum Beweis dieses Satzes benötigt man das so genannte *Theorem von Parseval*. Bezeichnen  $x(t)$  und  $y(t)$  zwei quadratisch integrierbare Zeitfunktionen, also  $x(t), y(t) \in L_2(-\infty, \infty)$ , und

$$\hat{x}(\omega) = \int_{-\infty}^{\infty} x(t) \exp(-I\omega t) dt \quad \text{bzw.} \quad \hat{y}(\omega) = \int_{-\infty}^{\infty} y(t) \exp(-I\omega t) dt \quad (5.44)$$

seien die zugehörigen Fouriertransformierten, dann gilt für das innere Produkt

$$\int_{-\infty}^{\infty} x(t)y(t)dt = \langle x, y \rangle = \langle \hat{x}, \hat{y} \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{x}(\omega)\hat{y}^*(\omega)d\omega . \quad (5.45)$$

Aus (5.45) folgt dann unmittelbar die Beziehung

$$\|x\|_2 = \|\hat{x}\|_2 . \quad (5.46)$$

Um das Theorem von Parseval für den Beweis von Satz 5.3 anwenden zu können, wird der Abschneideoperator  $(\ )_T$  in der Form

$$u_T(t) = \begin{cases} u(t) & \text{für } t \leq T \\ 0 & \text{für } t > T \end{cases} \quad (5.47)$$

eingeführt. Weiters wird angenommen, dass die Zeitfunktionen  $u(t)$  und  $y(t)$  kausal sind, d.h.  $u(t) = 0$  und  $y(t) = 0$  für  $t < 0$ . Damit erhält man

$$\int_0^T u(t)y(t)dt = \int_{-\infty}^{\infty} u_T(t)y(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}_T(\omega)\hat{y}^*(\omega)d\omega \quad (5.48)$$

bzw. mit  $\hat{y}(\omega) = G(I\omega)\hat{u}_T(\omega)$  ergibt sich

$$\begin{aligned} \int_0^T u(t)y(t) dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G^*(I\omega)\hat{u}_T(\omega)\hat{u}_T^*(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} (\text{Re}(G(I\omega)) - I\text{Im}(G(I\omega)))|\hat{u}_T(\omega)|^2 d\omega . \end{aligned} \quad (5.49)$$

Da die linke Seite von (5.49) rein reell ist, muss der Imaginärteil auf der rechten Seite verschwinden, und es gilt

$$\int_0^T u(t)y(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{Re}(G(I\omega))|\hat{u}_T(\omega)|^2 d\omega . \quad (5.50)$$

” $\Leftarrow$ ”: Setzt man nun voraus, dass (5.42) gilt, dann folgt

$$\int_0^T u(t)y(t)dt \geq \frac{\alpha}{2\pi} \int_{-\infty}^{\infty} |\hat{u}_T(\omega)|^2 d\omega = \alpha \int_0^T u^2(t)dt \quad (5.51)$$

und damit nach Definition 5.2 die  $\alpha$ -Eingangspassivität von (5.38).

” $\Rightarrow$ ”: Umgekehrt, wenn das System (5.38)  $\alpha$ -eingangspassiv ist, dann existiert ein  $\alpha > 0$  so, dass die Ungleichung

$$\int_0^T u(t)y(t)dt \geq \alpha \int_0^T u^2(t)dt \quad (5.52)$$

erfüllt ist, bzw. mit Hilfe des Theorems von Parseval erhält man

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} (\operatorname{Re}(G(\mathbf{I}\omega)) - \alpha) |\hat{u}_T(\omega)|^2 d\omega \geq 0. \quad (5.53)$$

Die Ungleichung (5.53) ist aber nur dann für alle Eingangsgrößen  $u(t)$  gültig, wenn für alle  $\omega$  gilt  $\operatorname{Re}(G(\mathbf{I}\omega)) \geq \alpha$ . Angenommen, es existiert ein  $\omega_0$  so, dass  $\operatorname{Re}(G(\mathbf{I}\omega_0)) < \alpha$  ist, dann sieht man, dass für die Eingangsgröße  $u(t) = U \sin(\omega_0 t)$  und hinreichend großes  $T$  die Ungleichung (5.53) nicht erfüllt ist. Damit ist aber Punkt (2) und für  $\alpha = 0$  auch Punkt (1) von Satz 5.3 bewiesen.

**Aufgabe 5.7.** Beweisen Sie Punkt (3) von Satz 5.3. □

Als einfaches Anwendungsbeispiel soll gezeigt werden, dass der PID-Regler

$$R(s) = V \frac{1 + T_I s}{s} \frac{1 + T_D s}{1 + \alpha T_D s} \quad (5.54)$$

mit den positiven Parametern  $V$ ,  $T_I$ ,  $T_D$  und  $0 < \alpha < 1$  passiv ist. Dazu berechne man einfach

$$\operatorname{Re}(R(\mathbf{I}\omega)) = \frac{V(T_I + T_D(1 - \alpha) + \alpha T_D^2 T_I \omega^2)}{1 + \alpha^2 T_D^2 \omega^2} > 0. \quad (5.55)$$

**Aufgabe 5.8.** Zeigen Sie, dass ein PI-Regler passiv ist.

**Aufgabe 5.9.** Zeigen Sie, dass das lineare zeitinvariante System (5.38) mit der Übertragungsfunktion  $G(s)$  von (5.39) passiv ist, wenn

$$|\arg(G(\mathbf{I}\omega))| \leq \frac{\pi}{2}. \quad (5.56)$$

**Aufgabe 5.10.** Betrachten Sie einen einschleifigen Standardregelkreis mit einer passiven Strecke  $G(s)$  und einem  $\alpha$ -eingangspassiven Regler  $R(s)$  mit  $\alpha > 0$ . Zeigen Sie, dass der geschlossene Kreis BIBO-stabil ist.

**Hinweis:** Verwenden Sie dazu das Nyquistkriterium.

**Aufgabe 5.11.** Der Zusammenhang zwischen Strom  $\hat{i}(x, s)$  und Spannung  $\hat{u}(x, s)$  an der Stelle  $x = 0$  und an der Stelle  $x = l$  einer langen elektrischen Leitung mit dem Kapazitätsbelag  $c$ , dem Induktivitätsbelag  $l$ , dem Widerstandsbelag  $r$  und dem

Leitwertsbelag  $g$  lautet

$$\begin{bmatrix} \hat{u}(0, s) \\ \hat{i}(0, s) \end{bmatrix} = \begin{bmatrix} \cosh(\gamma(s)l) & Z_0(s) \sinh(\gamma(s)l) \\ \frac{1}{Z_0(s)} \sinh(\gamma(s)l) & \cosh(\gamma(s)l) \end{bmatrix} \begin{bmatrix} \hat{u}(l, s) \\ \hat{i}(l, s) \end{bmatrix}, \quad (5.57)$$

wobei  $Z_0(s)$  den Wellenwiderstand und  $\gamma(s)$  den Ausbreitungskoeffizienten

$$Z_0(s) = \sqrt{\frac{r + sl}{g + sc}} \quad \text{und} \quad \gamma(s) = \sqrt{(r + sl)(g + sc)} \quad (5.58)$$

bezeichnen. überprüfen Sie für verschiedene Lastimpedanzen  $Z_L(s)$  mit

$$\hat{u}(l, s) = Z_L(s) \hat{i}(l, s) \quad (5.59)$$

die Passivität der Übertragungsfunktion  $G(s) = \frac{\hat{u}(0, s)}{\hat{i}(0, s)}$ .

## 5.5 Positive Reellheit

Bei linearen zeitinvarianten Systemen (5.38) wird an Stelle der Passivität sehr oft der Begriff der positiven Reellheit der zugehörigen Übertragungsfunktion (5.39) verwendet. Ohne Beweis sei angemerkt, dass das System (5.38) genau dann passiv ist, wenn (5.39) positiv reell ist.

**Satz 5.4.** Eine Übertragungsfunktion  $G(s)$  ist genau dann positiv reell, wenn

- (1)  $G(s)$  keine Pole in der rechten offenen  $s$ -Halbebene besitzt,
- (2)  $\operatorname{Re}(G(j\omega)) \geq 0$  ist für alle  $\omega$ , für die gilt,  $j\omega$  ist kein Pol von  $G(s)$  und
- (3) wenn  $s = j\omega_0$  ein Pol von  $G(s)$  ist, dann ist dieser einfach und für endliches  $\omega_0$  muss das Residuum

$$\lim_{s \rightarrow j\omega_0} (s - j\omega_0)G(s) \quad (5.60)$$

positiv und reell sein. Ist  $\omega_0$  unendlich, dann muss der Grenzwert

$$\lim_{\omega \rightarrow \infty} \frac{G(j\omega)}{j\omega} \quad (5.61)$$

positiv und reell sein.

Man nennt  $G(s)$  streng positiv reell, wenn  $G(s - \delta)$  für ein geeignetes  $\delta > 0$  positiv reell ist.

**Aufgabe 5.12.** Zeigen Sie, dass die Bedingungen

- (1) die Graddifferenz zwischen Zähler- und Nennerpolynom von  $G(s)$  sind  $-1, 0$  oder  $1$  und
- (2)  $G(s)$  hat keine Nullstellen in der rechten offenen  $s$ -Halbebene

notwendig dafür sind, dass  $G(s)$  positiv reell ist.

**Aufgabe 5.13.** Sind die nachfolgenden Übertragungsfunktionen

$$G_1(s) = -(s - 3), G_2(s) = \frac{1}{s^2 + 2s + 1}, G_3(s) = \frac{s + 1}{s^2 + 1}, G_4(s) = \frac{s + 10}{(s + 1)(s + 2)} \quad (5.62)$$

positiv reell?

Wie im nachfolgenden Satz gezeigt wird, hängt die positive Reellheit einer Übertragungsfunktion  $G(s)$  eng mit der Lösbarkeit eines Gleichungssystems zusammen. Für den Beweis dieses Satzes sei auf die am Ende angeführte Literatur verwiesen.

**Satz 5.5 (Kalman-Yakubovich-Popov (KYP)-Lemma).** Gegeben ist das System (5.38), wobei angenommen wird, dass das Paar  $(\mathbf{A}, \mathbf{b})$  erreichbar und das Paar  $(\mathbf{c}^T, \mathbf{A})$  beobachtbar ist. Die Übertragungsfunktion (5.39) ist genau dann positiv reell (passiv), wenn ein Skalar  $w$ , ein Vektor  $\mathbf{m}$  und eine positiv definite Matrix  $\mathbf{P}$  so existieren, dass nachfolgende Bedingungen

$$\begin{aligned} \mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} &= -\mathbf{m}\mathbf{m}^T \\ \mathbf{P}\mathbf{b} &= \mathbf{c} - \mathbf{m}w \\ w^2 &= 2d \end{aligned} \quad (5.63)$$

erfüllt sind. Die Übertragungsfunktion (5.39) ist darüberhinaus genau dann streng positiv reell nach Satz 5.4, wenn Skalare  $w$  und  $\varepsilon > 0$ , ein Vektor  $\mathbf{m}$  und eine positiv definite Matrix  $\mathbf{P}$  so existieren, dass nachfolgende Bedingungen

$$\begin{aligned} \mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} &= -\mathbf{m}\mathbf{m}^T - \varepsilon\mathbf{P} \\ \mathbf{P}\mathbf{b} &= \mathbf{c} - \mathbf{m}w \\ w^2 &= 2d \end{aligned} \quad (5.64)$$

erfüllt sind.

**Aufgabe 5.14.** Angenommen  $w$ ,  $\mathbf{m}$ ,  $\mathbf{P} > \mathbf{0}$  und  $\varepsilon > 0$  seien Lösungen von (5.64). Zeigen Sie, dass dann im Falle  $d \neq 0$  die Riccati-Gleichung

$$\mathbf{P}\left(\frac{\varepsilon}{2}\mathbf{E} + \mathbf{A}\right) + \left(\frac{\varepsilon}{2}\mathbf{E} + \mathbf{A}^T\right)\mathbf{P} + (\mathbf{c} - \mathbf{P}\mathbf{b})\frac{1}{2d}(\mathbf{c}^T - \mathbf{b}^T\mathbf{P}) = \mathbf{0} \quad (5.65)$$

erfüllt ist.



Als Anwendung des KYP Lemmas betrachte man den geschlossenen Regelkreis von Abbildung 5.5 mit der nichtlinearen passiven Strecke im Vorwärtszweig und dem streng positiv reellen Regler im Rückwärtszweig.

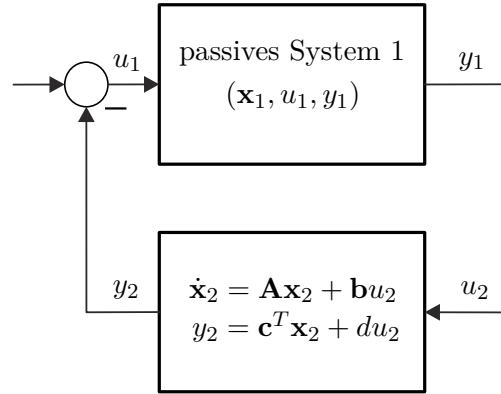


Abbildung 5.5: Passives System mit linearem Regler.

Angenommen das passive nichtlineare System habe eine stetig differenzierbare, positiv definite Speicherfunktion  $V_1(\mathbf{x}_1)$ , die der differentiellen Passivitätsungleichung (siehe (5.36))

$$\frac{d}{dt}V_1(\mathbf{x}_1) = -W_1(\mathbf{x}_1) + y_1 u_1 \leq y_1 u_1, \quad (5.66)$$

mit der positiv semidefiniten Funktion  $W_1(\mathbf{x}_1)$  genügt. Für das Weitere sei der streng positiv reelle Regler durch folgende Minimalrealisierung

$$\begin{aligned} \frac{d}{dt}\mathbf{x}_2 &= \mathbf{A}\mathbf{x}_2 + \mathbf{b}u_2 \\ y_2 &= \mathbf{c}^T\mathbf{x}_2 + du_2 \end{aligned} \quad (5.67)$$

beschrieben. Aufgrund des KYP Lemmas Satz 5.5 findet man für das System (5.67) Skalare  $w$  und  $\varepsilon > 0$ , einen Vektor  $\mathbf{m}$  und eine positiv definite Matrix  $\mathbf{P}$  so, dass (5.64) erfüllt ist. Damit ergibt sich die *Lyapunov-Funktion des geschlossenen Kreises* von Abbildung 5.5 zu

$$V_e(\mathbf{x}_1, \mathbf{x}_2) = V_1(\mathbf{x}_1) + \frac{1}{2}\mathbf{x}_2^T \mathbf{P} \mathbf{x}_2. \quad (5.68)$$

Um dies zu zeigen, berechnet man die zeitliche Änderung von (5.68) entlang der Lösungskurve und berücksichtigt die Zusammenschaltungsbedingung  $u_1 = -y_2$  und  $u_2 = y_1$  gemeinsam mit (5.64) und (5.66)

$$\begin{aligned}
\frac{d}{dt}V_e(\mathbf{x}_1, \mathbf{x}_2) &= -W_1(\mathbf{x}_1) + y_1 u_1 + \frac{1}{2} \underbrace{\dot{\mathbf{x}}_2^T \mathbf{P} \mathbf{x}_2}_{(\mathbf{x}_2^T \mathbf{A}^T + u_2 \mathbf{b}^T) \mathbf{P} \mathbf{x}_2} + \frac{1}{2} \underbrace{\mathbf{x}_2^T \mathbf{P} \dot{\mathbf{x}}_2}_{\mathbf{x}_2^T \mathbf{P} (\mathbf{A} \mathbf{x}_2 + \mathbf{b} u_2)} \\
&= -W_1(\mathbf{x}_1) + y_1 u_1 + \frac{1}{2} \mathbf{x}_2^T \underbrace{(\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A})}_{-\mathbf{m} \mathbf{m}^T - \varepsilon \mathbf{P}} \mathbf{x}_2 + \mathbf{x}_2^T \underbrace{\mathbf{P} \mathbf{b}}_{\mathbf{c} - \mathbf{m} w} u_2 \\
&= -W_1(\mathbf{x}_1) - \underbrace{y_1 \mathbf{c}^T \mathbf{x}_2}_{=} - d y_1^2 - \frac{1}{2} \mathbf{x}_2^T \mathbf{m} \mathbf{m}^T \mathbf{x}_2 - \frac{1}{2} \varepsilon \mathbf{x}_2^T \mathbf{P} \mathbf{x}_2 + \underbrace{\mathbf{x}_2^T \mathbf{c} y_1}_{=} - \mathbf{x}_2^T \mathbf{m} w y_1 \quad (5.69) \\
&= -W_1(\mathbf{x}_1) - \frac{1}{2} \varepsilon \mathbf{x}_2^T \mathbf{P} \mathbf{x}_2 - \frac{1}{2} y_1^2 \underbrace{(2d)}_{w^2} - \frac{1}{2} \mathbf{x}_2^T \mathbf{m} \mathbf{m}^T \mathbf{x}_2 - \mathbf{x}_2^T \mathbf{m} w y_1 \\
&= -W_1(\mathbf{x}_1) - \frac{1}{2} \varepsilon \mathbf{x}_2^T \mathbf{P} \mathbf{x}_2 - \frac{1}{2} (\mathbf{m}^T \mathbf{x}_2 + w y_1)^T (\mathbf{m}^T \mathbf{x}_2 + w y_1) \leq 0.
\end{aligned}$$

Dies zeigt unmittelbar die Stabilität des geschlossenen Kreises von Abbildung 5.5.

## 5.6 Kanonische Form Passiver Systeme

Bevor eine kanonische Form für passive Systeme vorgestellt wird, soll gezeigt werden, dass die wohlbekannten Euler-Lagrange Gleichungen passiv sind.

### 5.6.1 Hamiltonsche Systeme

Betrachtet man ein endlich-dimensionales Lagrangesches System mit  $n$  Freiheitsgraden und den generalisierten Koordinaten  $\mathbf{q} \in \mathbb{R}^n$ , dann folgen bekannterweise die Bewegungsgleichungen aus den Euler-Lagrange Gleichungen in der Form

$$\frac{d}{dt} \left( \frac{\partial L}{\partial v_k} \right) - \frac{\partial L}{\partial q_k} = \tau_k, \quad k = 1, \dots, n \quad (5.70)$$

mit der Lagrangefunktion  $L(\mathbf{q}, \mathbf{v})$ , den generalisierten Geschwindigkeiten  $\frac{d}{dt} \mathbf{q} = \mathbf{v}$  und den generalisierten Kräften  $\tau_k$ ,  $k = 1, \dots, n$ . Bei einfachen Lagrangeschen Systemen entspricht die Lagrangefunktion der Differenz aus kinetischer und potenzieller Energie

$$L(\mathbf{q}, \mathbf{v}) = T(\mathbf{q}, \mathbf{v}) - V(\mathbf{q}). \quad (5.71)$$

Es sei angenommen, dass sich die generalisierten Kräfte  $\boldsymbol{\tau}$  aus externen Kräften  $\boldsymbol{\tau}_e$  (Stell- und Störeingänge im regelungstechnischen Sinne) und dissipativen Kräften  $\boldsymbol{\tau}_d^T = -\left(\frac{\partial}{\partial \mathbf{v}} R\right)(\mathbf{v})$  mit der *Rayleighsche Dissipationsfunktion*  $R(\mathbf{v})$  und

$$\left(\frac{\partial}{\partial \mathbf{v}} R\right)(\mathbf{v}) \mathbf{v} \geq 0 \quad (5.72)$$

zusammensetzen. Damit ergibt sich (5.70) zu

$$\frac{d}{dt} \left( \frac{\partial L}{\partial v_k} \right) - \frac{\partial L}{\partial q_k} + \frac{\partial}{\partial v_k} R = \tau_{e,k}, \quad k = 1, \dots, n. \quad (5.73)$$

**Definition 5.4.** Man bezeichnet das Lagrangesche System (5.73) *voll gedämpft*, wenn die Rayleighsche Dissipationsfunktion  $R(\mathbf{v})$  folgender Ungleichung

$$\left(\frac{\partial}{\partial \mathbf{v}} R\right)(\mathbf{v})\mathbf{v} \geq \sum_{k=1}^n \beta_k v_k^2, \quad \beta_k > 0, \quad k = 1, \dots, n \quad (5.74)$$

genügt. Ist ein  $\beta_k = 0$ , dann spricht man auch von einem *nicht voll gedämpften* Lagrangeschen System.

Mit Hilfe der generalisierten Impulskoordinaten

$$p_k = \frac{\partial L}{\partial v_k}, \quad k = 1, \dots, n \quad (5.75)$$

und der *Legendre-Transformation*  $(q_k, v_k) \rightarrow (q_k, p_k)$  erhält man direkt aus den Euler-Lagrange Gleichungen (5.70) die *äquivalenten Hamiltonschen Gleichungen*

$$\begin{aligned} \frac{d}{dt} q_k &= \frac{\partial H}{\partial p_k} \\ \frac{d}{dt} p_k &= -\frac{\partial H}{\partial q_k} + \tau_k, \quad k = 1, \dots, n \end{aligned} \quad (5.76)$$

mit der Hamiltonfunktion

$$H(\mathbf{q}, \mathbf{p}) = \sum_{k=1}^n p_k v_k - L(\mathbf{q}, \mathbf{v}). \quad (5.77)$$

Der Satz über implizite Funktionen besagt, dass die generalisierten Geschwindigkeiten  $v_k$  aus (5.75) genau dann lokal berechnet werden können, wenn die Matrix  $\left[\frac{\partial^2}{\partial v_i \partial v_j} L\right]$  regulär ist. Man spricht dann auch von einer *nichtdegenerierten Lagrangefunktion*  $L$ .

*Beweis.* Zum Beweis betrachte man die kurzen Ableitungen

$$\frac{\partial H}{\partial p_k} = v_k + \sum_{j=1}^n \left( p_j \frac{\partial v_j}{\partial p_k} - \underbrace{\frac{\partial L}{\partial v_j}}_{=p_j} \frac{\partial v_j}{\partial p_k} \right) = v_k = \frac{d}{dt} q_k \quad (5.78)$$

und

$$\frac{\partial H}{\partial q_k} = \sum_{j=1}^n \left( p_j \frac{\partial v_j}{\partial q_k} - \underbrace{\frac{\partial L}{\partial v_j}}_{=p_j} \frac{\partial v_j}{\partial q_k} \right) - \frac{\partial L}{\partial q_k} = \tau_k - \frac{d}{dt} \left( \frac{\partial L}{\partial v_k} \right) = \tau_k - \frac{d}{dt} p_k. \quad (5.79)$$

□

Wenn die kinetische Energie  $T(\mathbf{q}, \mathbf{v})$  in (5.71) die Form

$$T(\mathbf{q}, \mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbf{D}(\mathbf{q}) \mathbf{v} \quad (5.80)$$

mit der positiv definiten Massenmatrix  $\mathbf{D}(\mathbf{q})$  hat, dann entspricht die Hamiltonfunktion

$$H(\mathbf{q}, \mathbf{p}) = \sum_{k=1}^n p_k v_k - \frac{1}{2} \mathbf{v}^T \mathbf{D}(\mathbf{q}) \mathbf{v} + V(\mathbf{q}) = \frac{1}{2} \mathbf{v}^T \mathbf{D}(\mathbf{q}) \mathbf{v} + V(\mathbf{q}) \quad (5.81)$$

der im System gespeicherten Energie. Berechnet man die zeitliche Änderung der Hamiltonfunktion (5.81)

$$\frac{d}{dt} H(\mathbf{q}, \mathbf{p}) = \sum_{k=1}^n \left[ \frac{\partial H}{\partial q_k} \frac{\partial H}{\partial p_k} + \underbrace{\frac{\partial H}{\partial p_k}}_{v_k} \left( -\frac{\partial H}{\partial q_k} - \frac{\partial}{\partial v_k} R + \tau_{e,k} \right) \right] \leq \sum_{k=1}^n v_k \tau_{e,k}, \quad (5.82)$$

dann sieht man, dass das Lagrangesche System gemäß Definition 5.2 passiv ist mit der Eingangsgröße  $\tau_e$ , der Ausgangsgröße  $\mathbf{v} = \frac{d}{dt} \mathbf{q}$  und der Speicherfunktion  $H(\mathbf{q}, \mathbf{p})$ . Ist darüberhinaus das Lagrangesche System gemäß Definition 5.4 voll gedämpft, dann ist das Lagrangesche System wegen (5.74) sogar  $\beta$ -ausgangspassiv mit  $\beta = \min_k(\beta_k)$ ,  $k = 1, \dots, n$ , da gilt

$$\frac{d}{dt} H(\mathbf{q}, \mathbf{p}) \leq \sum_{k=1}^n v_k \tau_{e,k} - \sum_{k=1}^n \beta_k v_k^2 \leq \sum_{k=1}^n v_k \tau_{e,k} - \min_k(\beta_k) \|\mathbf{v}\|_2^2. \quad (5.83)$$

Man sagt dann auch, dass  $v_k$  der zur generalisierten Kraft  $\tau_{e,k}$  *kollokierte Ausgang* ist. D.h., die Paarung  $(\tau_{e,k}, v_k)$  beschreibt einen Energieeingang in das System, wie z.B. zusammengehörende Strom und Spannungen, Kräfte und Geschwindigkeiten oder Momente und Drehwinkelgeschwindigkeiten. Im Rahmen der Netzwerkstheorie werden solche Paarungen von Strom und Spannung, die einen Energieeingang bilden, auch als *Tor* (im Englischen port) bezeichnet. Die Generalisierung der Hamiltonschen Gleichungen (5.76) in Kombination mit dem Torkonzept führt direkt zur Klasse der *Port-Hamiltonschen Systeme*.

### 5.6.2 Port-Hamiltonsche Systeme

Ein finit-dimensionales Port-Hamiltonsches System lässt sich in der Form

$$\frac{d}{dt} \mathbf{x} = (\mathbf{J}(\mathbf{x}) - \mathbf{S}(\mathbf{x})) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \mathbf{G}_u(\mathbf{x}) \mathbf{u} \quad (5.84)$$

mit dem Zustand  $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^n$  und dem Stelleingang  $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$  formulieren. Dabei bezeichnet  $V(\mathbf{x})$ ,  $V(\mathbf{0}) = 0$ , eine stetig differenzierbare positiv definite Speicherfunktion und die Einträge der Matrizen  $\mathbf{G}_u(\mathbf{x})$ ,  $\mathbf{J}(\mathbf{x}) = -\mathbf{J}^T(\mathbf{x})$  und  $\mathbf{S}(\mathbf{x}) = \mathbf{S}^T(\mathbf{x}) \geq \mathbf{0}$  seien glatte Funktionen in  $\mathbf{x}$ . Wählt man als Ausgang  $\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^m$  den *kollokierten Ausgang*

$$\mathbf{y} = \mathbf{G}_u^T(\mathbf{x}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T, \quad (5.85)$$

dann erkennt man unmittelbar aus der differentiellen Passivitätsungleichung

$$\frac{d}{dt} V = \mathbf{y}^T \mathbf{u} - \left( \frac{\partial V}{\partial \mathbf{x}} \right) \mathbf{S}(\mathbf{x}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T \leq \mathbf{y}^T \mathbf{u}, \quad (5.86)$$

dass das System (5.84) passiv ist mit der Speicherfunktion  $V(\mathbf{x})$ . Die Darstellung in der Form von (5.84) erlaubt mehr als nur die einfache Feststellung der Passivität – sie ermöglicht, falls die Speicherfunktion  $V(\mathbf{x})$  gleich der im System gespeicherten Gesamtenergie ist, einen tieferen Einblick in die Energieflüsse des Systems im Inneren und mit der Systemumgebung: Die schiefsymmetrische Matrix  $\mathbf{J}(\mathbf{x})$  ist nämlich mit den Energieflüssen im Systeminneren verbunden, die symmetrische, positiv semidefinite Matrix  $\mathbf{S}(\mathbf{x})$  umfasst das Verhalten der dissipativen Effekte und  $\mathbf{G}_u(\mathbf{x})$  beschreibt den Energieaustausch des Systems mit der Systemumgebung über die Systemtore. Wenn (5.84) keine dissipativen Elemente enthält, also  $\mathbf{S}(\mathbf{x}) = \mathbf{0}$  ist, dann ist das System verlustlos bezüglich der Versorgungsrate  $\mathbf{y}^T \mathbf{u}$ .

Eine perfekte Aktuator/Sensor Kollokation bringt den Vorteil mit sich, dass eine einfache (zustandsabhängige) Rückführung des kollokierten Ausgangs (5.85) der Form

$$\mathbf{u} = -\mathbf{K}(\mathbf{x})\mathbf{y} = -\mathbf{K}(\mathbf{x})\mathbf{G}_u^T(\mathbf{x})\left(\frac{\partial V}{\partial \mathbf{x}}\right)^T, \quad (5.87)$$

mit der positiv definiten Matrix  $\mathbf{K}(\mathbf{x}) > 0$  für alle  $\mathbf{x} \in \mathcal{X}$  bei stabilen Strecken die Stabilität im geschlossenen Kreis erhält, da gilt

$$\frac{d}{dt}V = -\left(\frac{\partial V}{\partial \mathbf{x}}\right)\left(\mathbf{S}(\mathbf{x}) + \mathbf{G}_u(\mathbf{x})\mathbf{K}(\mathbf{x})\mathbf{G}_u^T(\mathbf{x})\right)\left(\frac{\partial V}{\partial \mathbf{x}}\right)^T \leq 0. \quad (5.88)$$

In der Literatur wird diese Art der Rückführung (5.87) im Zusammenhang mit Port-Hamiltonschen Systemen als *damping injection* bezeichnet oder bei allgemeinen nichtlinearen Systemen mit affinem Eingang als *Jurdjevic-Quinn Rückführung*.

**Beispiel 5.1** (Port-Hamiltonsche Darstellung des Elektromagnetventils (5.14)). Um das mathematische Modell des Elektromagnetventils (5.14) in Port-Hamiltonsche Darstellung (5.84) zu bringen, führt man die neuen Zustandsgrößen  $\mathbf{x}^T = [z, p = mv, \psi_L = L(z)i_L]$  ein. Die im Magnetventil gespeicherte Energie gemäß (5.15) formuliert im neuen Zustand  $[z, p, \psi_L]$

$$V = \frac{1}{2}\left(\frac{1}{L(z)}\psi_L^2 + \frac{1}{m}p^2 + cz^2\right) \quad (5.89)$$

wird in weiterer Folge als Speicherfunktion verwendet. Mit

$$\frac{\partial V}{\partial \mathbf{x}} = \begin{bmatrix} cz - \frac{1}{2} \frac{\partial L(z)}{\partial z} \frac{\psi_L^2}{L^2(z)} & \frac{p}{m} & \frac{\psi_L}{L(z)} \end{bmatrix} \quad (5.90)$$

und den Systemgleichungen (5.14) im transformierten Zustand

$$\begin{aligned} \frac{d}{dt}z &= \frac{p}{m} \\ \frac{d}{dt}p &= \left(\frac{1}{2} \frac{\partial L(z)}{\partial z} \frac{\psi_L^2}{L^2(z)} - cz - d\frac{p}{m} + F_{ext}\right) \\ \frac{d}{dt}\psi_L &= U_0 - R\frac{\psi_L}{L(z)} \end{aligned} \quad (5.91)$$

ergibt sich unmittelbar die Port-Hamiltonsche Darstellung (5.84) zu

$$\frac{d}{dt} \begin{bmatrix} z \\ p \\ \psi_L \end{bmatrix} = \left( \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{J}(\mathbf{x})} - \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & R \end{bmatrix}}_{\mathbf{S}(\mathbf{x})} \right) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{G}_u(\mathbf{x})} \underbrace{\begin{bmatrix} U_0 \\ F_{ext} \end{bmatrix}}_{\mathbf{u}}. \quad (5.92)$$

Der zugehörige kollokierte Ausgang gemäß (5.85) lautet

$$\mathbf{y} = \mathbf{G}_u^T(\mathbf{x}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T = \begin{bmatrix} \psi_L \\ L(z) \\ p \\ m \end{bmatrix} = \begin{bmatrix} i_L \\ v \end{bmatrix}. \quad (5.93)$$

**Aufgabe 5.15.** Stellen Sie die mathematischen Modelle des Balkens mit rollender Kugel und des Krans mit einem Schwenkarm aus dem Skriptum zur VO Regelungssysteme 2 [5.1] als Port-Hamiltonsche Systeme dar.

**Aufgabe 5.16.** Stellen Sie die unterschiedlichen Gleichstrommaschinen vom Abschnitt 1.7 aus dem Skriptum zur VO Regelungssysteme 2 [5.1] als Port-Hamiltonsche Systeme dar.

## 5.7 Passivitätsbasierter Reglerentwurf

Ein mit der Port-Hamiltonschen Struktur (5.84) unmittelbar verbundenes Reglerentwurfsverfahren ist die so genannte *IDA-PBC* (*Interconnection and Damping Assignment Passivity-Based Control*). Dazu sei folgender Satz formuliert:

**Satz 5.6 (IDA-PBC).** Gegeben ist das nichtlineare System

$$\frac{d}{dt} \mathbf{x} = \mathbf{f}(\mathbf{x}) + \mathbf{G}_u(\mathbf{x}) \mathbf{u} \quad (5.94)$$

mit dem Zustand  $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^n$  und dem Stelleingang  $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$  mit  $m < n$ . Von der Matrix  $\mathbf{G}_u(\mathbf{x})$  wird vorausgesetzt, dass diese für alle  $\mathbf{x} \in \mathcal{X}$  spaltenregulär ist, d.h.  $\text{rang}(\mathbf{G}_u(\mathbf{x})) = m$ . Im Weiteren bezeichne  $\mathbf{G}_u^\perp(\mathbf{x})$  den Linksannihilator von  $\mathbf{G}_u(\mathbf{x})$ , d.h.  $\mathbf{G}_u^\perp(\mathbf{x}) \mathbf{G}_u(\mathbf{x}) = \mathbf{0}$ , und  $V_d(\mathbf{x})$  sei die Speicherfunktion des geschlossenen Kreises und habe an der gewünschten Ruhelage  $\mathbf{x} = \mathbf{x}_d$  ein striktes Minimum, d.h.

$$V_d(\mathbf{x}) > V_d(\mathbf{x}_d) \quad \text{für alle } \mathbf{x} \neq \mathbf{x}_d, \quad \left( \frac{\partial V_d}{\partial \mathbf{x}} \right) (\mathbf{x}_d) = \mathbf{0} \quad \text{und} \quad \left( \frac{\partial^2 V_d}{\partial \mathbf{x}^2} \right) (\mathbf{x}_d) > 0. \quad (5.95)$$

Damit ist  $V_d(\mathbf{x}) - V_d(\mathbf{x}_d)$  positiv definit und eignet sich als Lyapunovfunktion für den geschlossenen Kreis. Angenommen die Matrizen  $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^T(\mathbf{x})$ ,  $\mathbf{S}_d(\mathbf{x}) = \mathbf{S}_d^T(\mathbf{x}) \geq 0$ , der Linksannihilator  $\mathbf{G}_u^\perp(\mathbf{x})$  und die Speicherfunktion  $V_d(\mathbf{x})$  genügen der Bedingung (PBC matching equation)

$$\mathbf{G}_u^\perp(\mathbf{x})\mathbf{f}(\mathbf{x}) = \mathbf{G}_u^\perp(\mathbf{x})(\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T, \quad (5.96)$$

dann ergibt sich mit der Zustandsrückführung

$$\mathbf{u} = \beta(\mathbf{x}) = \left(\mathbf{G}_u^T(\mathbf{x})\mathbf{G}_u(\mathbf{x})\right)^{-1}\mathbf{G}_u^T(\mathbf{x})\left\{(\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T - \mathbf{f}(\mathbf{x})\right\} \quad (5.97)$$

eingesetzt in (5.94) ein geschlossener Kreis in Port-Hamiltonscher Form

$$\frac{d}{dt}\mathbf{x} = (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T \quad (5.98)$$

mit der stabilen gewünschten Ruhelage des geschlossenen Kreises  $\mathbf{x} = \mathbf{x}_d$ . Wenn die Menge  $\{\mathbf{x}_d\}$  die größte positiv invariante Menge von

$$\left\{\mathbf{x} \in \mathbb{R}^n \mid \left(\frac{\partial V_d}{\partial \mathbf{x}}\right)\mathbf{S}_d(\mathbf{x})\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T = 0\right\} \quad (5.99)$$

ist, dann ist  $\mathbf{x} = \mathbf{x}_d$  sogar asymptotisch stabil.

*Beweis.* Setzt man die rechten Seiten von (5.98) und (5.94) mit (5.97) gleich, d.h.

$$\mathbf{f}(\mathbf{x}) + \mathbf{G}_u(\mathbf{x})\beta(\mathbf{x}) = (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T, \quad (5.100)$$

und multipliziert man mit  $\mathbf{G}_u^\perp(\mathbf{x})$  von links, so erhält man unmittelbar die PBC matching equation (5.96). Die Zustandsrückführung (5.97) folgt direkt aus (5.100) durch Multiplikation mit der Pseudoinversen  $\left(\mathbf{G}_u^T(\mathbf{x})\mathbf{G}_u(\mathbf{x})\right)^{-1}\mathbf{G}_u^T(\mathbf{x})$  von links. Man beachte, dass die zuvor angenommene Spaltenregularität von  $\mathbf{G}_u(\mathbf{x})$  die Regularität der Pseudoinversen garantiert. Im nächsten Schritt soll gezeigt werden, dass (5.94) mit (5.97) tatsächlich (5.98) entspricht, also gilt

$$\begin{aligned} \Psi = \mathbf{f}(\mathbf{x}) + \mathbf{G}_u(\mathbf{x})\left(\mathbf{G}_u^T(\mathbf{x})\mathbf{G}_u(\mathbf{x})\right)^{-1}\mathbf{G}_u^T(\mathbf{x})\left\{(\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T - \mathbf{f}(\mathbf{x})\right\} \\ - (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d(\mathbf{x}))\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T = \mathbf{0}. \end{aligned} \quad (5.101)$$

Dazu multipliziert man (5.101) mit der regulären Matrix

$$\mathbf{T}(\mathbf{x}) = \begin{bmatrix} \mathbf{G}_u^T(\mathbf{x}) \\ \mathbf{G}_u^\perp(\mathbf{x}) \end{bmatrix} \quad (5.102)$$

und zufolge der PBC matching condition (5.96) folgt  $\mathbf{T}(\mathbf{x})\boldsymbol{\Psi} = \mathbf{0}$  und damit unmittelbar  $\boldsymbol{\Psi} = \mathbf{0}$ .  $\square$

Die Schwierigkeit dieser Reglerentwurfsmethode besteht offensichtlich darin, die PBC matching equation (5.96), welche ein *System partieller Differentialgleichungen* darstellt, zu lösen.

Dazu sei erwähnt, dass

- die Matrizen  $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^T(\mathbf{x})$  und  $\mathbf{S}_d(\mathbf{x}) = \mathbf{S}_d^T(\mathbf{x}) \geq 0$  frei zu wählen sind,
- die Speicherfunktion des geschlossenen Kreises  $V_d(\mathbf{x})$  abgesehen von der Bedingung (5.95) ebenfalls frei gewählt werden kann,
- und der Linksannihilator  $\mathbf{G}_u^\perp(\mathbf{x})$  mit jeder regulären  $(n-m) \times (n-m)$  Matrix  $\boldsymbol{\Lambda}(\mathbf{x})$  von links multipliziert werden kann, d.h.  $\tilde{\mathbf{G}}_u^\perp(\mathbf{x}) = \boldsymbol{\Lambda}(\mathbf{x})\mathbf{G}_u^\perp(\mathbf{x})$ , ohne die PBC matching equation (5.96) zu ändern. Die Matrix  $\boldsymbol{\Lambda}(\mathbf{x})$  stellt somit einen weiteren Entwurfsfreiheitsgrad dar.

In den letzten Jahren haben sich im Wesentlichen folgende Varianten des IDA-PBC Entwurfsverfahrens durchgesetzt:

- *Non-Parametrized IDA-PBC*: In diesem Fall wird die Struktur der Zusammenschaltung in Form der Matrizen  $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^T(\mathbf{x})$  und  $\mathbf{S}_d(\mathbf{x}) = \mathbf{S}_d^T(\mathbf{x}) \geq 0$  vorgegeben. Mit bekanntem  $\mathbf{G}_u^\perp(\mathbf{x})$  resultiert die PBC matching equation (5.96) zu einer partiellen Differentialgleichung für die Speicherfunktion  $V_d(\mathbf{x})$ . Aus der Familie aller Lösungen müssen dann jene extrahiert werden, die die Bedingung (5.95) erfüllen. In der Literatur, siehe beispielsweise [5.2], findet man auch Bedingungen für die Existenz einer Lösung der zugrundeliegenden partiellen Differentialgleichung (5.96).
- *Algebraic IDA-PBC*: In diesem Fall wird die Speicherfunktion  $V_d(\mathbf{x})$  unter der Bedingung (5.95) festgelegt und die PBC matching equation (5.96) degeneriert zu einer algebraischen Gleichung für die Bestimmung der Matrizen  $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^T(\mathbf{x})$  und  $\mathbf{S}_d(\mathbf{x}) = \mathbf{S}_d^T(\mathbf{x}) \geq 0$ .
- *Parametrized IDA-PBC*: Hier wird die Speicherfunktion  $V_d(\mathbf{x})$  auf eine bestimmte Klasse eingeschränkt, beispielsweise bei mechanischen Systemen, dass die gewünschte potenzielle Energie nur von den generalisierten Lagekoordinaten abhängt und die gewünschte kinetische Energie eine quadratische Form in den generalisierten Geschwindigkeiten ist. Diese spezielle Form von  $V_d(\mathbf{x})$  impliziert eine neue PBC matching equation mit Einschränkungen bezüglich der Wahl von  $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^T(\mathbf{x})$  und  $\mathbf{S}_d(\mathbf{x}) = \mathbf{S}_d^T(\mathbf{x}) \geq 0$ .



*Beispiel 5.2.* Als Anwendungsbeispiel betrachte man eine *permanentmagnetisch erregte Synchronmaschine* in  $dq$ -Darstellung

$$\begin{aligned} L_d \frac{d}{dt} i_d &= -R_s i_d + \omega L_q i_q + u_d \\ L_q \frac{d}{dt} i_q &= -R_s i_q - \omega(L_d i_d + \Phi) + u_q \\ J \frac{d}{dt} \omega &= p((L_d - L_q) i_d i_q + \Phi i_q) - \tau_l \end{aligned} \quad (5.103)$$

mit den Statorströmen  $i_d$  und  $i_q$  sowie der Drehwinkelgeschwindigkeit des Rotors  $\omega$  als Zustandsgrößen, den Statorspannungen  $u_d$  und  $u_q$  als Stellgrößen und dem Lastmoment  $\tau_l$ . Im Weiteren bezeichnet  $J$  das Trägheitsmoment,  $R_s$  den Statorwicklungswiderstand,  $L_d$  und  $L_q$  die Statorinduktivitäten,  $p$  die Polpaarzahl und  $\Phi$  den Fluss des Permanentmagneten im Rotor. Es sei an dieser Stelle erwähnt, dass für den Fall eines *gleichförmigen Luftspaltes* gilt  $L_d = L_q = L$  und sich damit das mathematische Modell (5.103) entsprechend vereinfacht.

Wählt man nun als Zustandsgrößen  $\mathbf{x}^T = [x_1, x_2, x_3] = [L_d i_d, L_q i_q, J\omega/p]$ , dann lässt sich (5.103) in Form eines Port-Hamiltonschen Systems

$$\frac{d}{dt} \mathbf{x} = (\mathbf{J}(\mathbf{x}) - \mathbf{S}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \mathbf{G}_u \mathbf{u} + \mathbf{g}_d \tau_l \quad (5.104)$$

mit der Energiefunktion als Speicherfunktion

$$V(\mathbf{x}) = \frac{1}{2L_d} x_1^2 + \frac{1}{2L_q} x_2^2 + \frac{p}{2J} x_3^2 \quad (5.105)$$

und

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} 0 & 0 & x_2 \\ 0 & 0 & -(x_1 + \Phi) \\ -x_2 & x_1 + \Phi & 0 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} R_s & 0 & 0 \\ 0 & R_s & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5.106)$$

sowie

$$\mathbf{G}_u = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{g}_d = \begin{bmatrix} 0 \\ 0 \\ -1/p \end{bmatrix} \quad \text{und} \quad \mathbf{u} = \begin{bmatrix} u_d \\ u_q \end{bmatrix} \quad (5.107)$$

schreiben.

*Aufgabe 5.17.* Zeigen Sie die Gültigkeit von (5.104).

Es soll nun mit Hilfe der *Non-Parametrized IDA-PBC* eine Zustandsrückführung gemäß

Satz 5.6 so entworfen werden, dass der stationäre Arbeitspunkt

$$\mathbf{x}_d^T = [0, x_{2,d}, x_{3,d}] \quad \text{mit} \quad x_{2,d} = \frac{\bar{\tau}_l L_q}{\Phi p} \quad (5.108)$$

für ein konstantes Moment  $\bar{\tau}_l$  und eine gewünschte Drehwinkelgeschwindigkeit  $\omega_d = x_{3,d} p / J$  stabilisiert wird. Die Struktur des geschlossenen Kreises  $\mathbf{J}_d(\mathbf{x})$  und  $\mathbf{S}_d$  wird nun entsprechend einer Maschine mit gleichförmigem Luftspalt gewählt, d.h., es gilt  $L_d = L_q = L$ .

**Aufgabe 5.18.** Zeigen Sie, dass für  $L_d = L_q = L$  die Matrizen  $\mathbf{J}_d(\mathbf{x})$  und  $\mathbf{S}_d$  des zu (5.103) zugehörigen Port-Hamiltonschen Systems folgende Struktur aufweisen

$$\mathbf{J}_d(\mathbf{x}) = \begin{bmatrix} 0 & \frac{Lp}{J} x_3 & 0 \\ -\frac{Lp}{J} x_3 & 0 & -\Phi \\ 0 & \Phi & 0 \end{bmatrix} \quad \text{und} \quad \mathbf{S}_d = \mathbf{S} . \quad (5.109)$$

Die PBC matching equation (5.96) lautet dann

$$(\mathbf{J}(\mathbf{x}) - \mathbf{S}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \mathbf{G}_u \boldsymbol{\beta}(\mathbf{x}) + \mathbf{g}_d \bar{\tau}_l = (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d) \left( \frac{\partial V_d}{\partial \mathbf{x}} \right)^T \quad (5.110)$$

bzw. mit dem Linksannihilator von  $\mathbf{G}_u$

$$\mathbf{G}_u^\perp = [0, 0, 1] \quad (5.111)$$

und den Größen  $V_a(\mathbf{x}) = V_d(\mathbf{x}) - V(\mathbf{x})$  sowie

$$\mathbf{J}_a(\mathbf{x}) = \mathbf{J}_d(\mathbf{x}) - \mathbf{J}(\mathbf{x}) = \begin{bmatrix} 0 & \frac{Lp}{J} x_3 & -x_2 \\ -\frac{Lp}{J} x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{bmatrix} \quad (5.112)$$

ergibt sich

$$\mathbf{G}_u^\perp (\mathbf{J}(\mathbf{x}) - \mathbf{S}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \mathbf{G}_u^\perp \mathbf{g}_d \bar{\tau}_l = \mathbf{G}_u^\perp (\mathbf{J}(\mathbf{x}) + \mathbf{J}_a(\mathbf{x}) - \mathbf{S}) \left( \left( \frac{\partial V_a}{\partial \mathbf{x}} \right)^T + \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T \right) \quad (5.113)$$

bzw.

$$-\mathbf{G}_u^\perp \mathbf{J}_a(\mathbf{x}) \left( \frac{\partial V}{\partial \mathbf{x}} \right)^T + \mathbf{G}_u^\perp \mathbf{g}_d \bar{\tau}_l = \mathbf{G}_u^\perp (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}) \left( \frac{\partial V_a}{\partial \mathbf{x}} \right)^T . \quad (5.114)$$

Die Auswertung von (5.114) resultiert in folgender partieller Differentialgleichung

$$-\frac{x_2 x_1}{L_d} + \frac{x_2 x_1}{L_q} - \frac{1}{p} \bar{\tau}_l = \Phi \frac{\partial V_a}{\partial x_2} , \quad (5.115)$$

deren allgemeine Lösung sich wie folgt

$$V_a(x_1, x_2, x_3) = \alpha_1 \left( \frac{1}{2} x_2^2 x_1 \left( \frac{L_d - L_q}{L_d L_q \Phi} \right) - \frac{x_2}{\Phi p} \bar{\tau}_l \right) + \psi(x_1, x_3) \quad (5.116)$$

mit dem positiven Parameter  $\alpha_1$  und einer noch zu wählenden Funktion  $\psi(x_1, x_3)$  darstellen lässt. Damit besitzt die Speicherfunktion des geschlossenen Kreises  $V_d = V + V_a$  folgende Struktur

$$V_d = \frac{1}{2L_d}x_1^2 + \frac{1}{2L_q}x_2^2 + \frac{p}{2J}x_3^2 + \frac{1}{2}\alpha_1x_2^2x_1\left(\frac{L_d - L_q}{L_dL_q\Phi}\right) - \alpha_1\frac{x_2}{\Phi p}\bar{\tau}_l + \psi(x_1, x_3). \quad (5.117)$$

Die Aufgabe besteht nun darin, die Funktion  $\psi(x_1, x_3)$  so festzulegen, dass die Bedingungen (5.95) erfüllt werden. Man kann sich nun einfach überzeugen, dass der Ansatz

$$\psi(x_1, x_3) = -\frac{1}{2}\alpha_1\left(\frac{L_d - L_q}{L_dL_q\Phi}\right)x_1x_{2,d}^2 + \frac{\alpha_2}{2}x_1^2 - \frac{p}{2J}x_3^2 + \frac{\alpha_3}{2}(x_3 - x_{3,d})^2 - \frac{1}{2L_q}x_{2,d}^2 \quad (5.118)$$

mit den positiven Entwurfparametern  $\alpha_1$ ,  $\alpha_2$  und  $\alpha_3$  diese Bedingungen erfüllt. Dazu berechnet man für

$$V_d = \left(\frac{1}{2L_d} + \frac{\alpha_2}{2}\right)x_1^2 + \left(\frac{1}{2L_q} + \frac{\alpha_1}{2}x_1\left(\frac{L_d - L_q}{L_dL_q\Phi}\right)\right)(x_2^2 - x_{2,d}^2) - \frac{\alpha_1}{L_q}x_2x_{2,d} + \frac{\alpha_3}{2}(x_3 - x_{3,d})^2 \quad (5.119)$$

vorerst den Gradienten und wertet diesen an der Stelle  $\mathbf{x} = \mathbf{x}_d$  (siehe (5.108)) aus

$$\left(\frac{\partial}{\partial \mathbf{x}}V_d\right)^T(\mathbf{x}_d) = \begin{bmatrix} \left(\frac{1}{L_d} + \alpha_2\right)x_{1,d} \\ \left(\frac{1}{L_q} + \alpha_1x_{1,d}\left(\frac{L_d - L_q}{L_dL_q\Phi}\right)\right)x_{2,d} - \frac{\alpha_1}{L_q}x_{2,d} \\ 0 \end{bmatrix}. \quad (5.120)$$

Offensichtlich ist für  $\alpha_1 = 1$  die Forderung  $\left(\frac{\partial}{\partial \mathbf{x}}V_d\right)^T(\mathbf{x}_d) = \mathbf{0}$  erfüllt. Um zu gewährleisten, dass  $\mathbf{x}_d$  ein striktes lokales Minimum von  $V_d$  ist, muss im Weiteren

$$\left(\frac{\partial^2}{\partial \mathbf{x}^2}V_d\right)(\mathbf{x}_d) = \begin{bmatrix} \frac{1}{L_d} + \alpha_2 & \left(\frac{L_d - L_q}{L_dL_q\Phi}\right)x_{2,d} & 0 \\ \left(\frac{L_d - L_q}{L_dL_q\Phi}\right)x_{2,d} & \frac{1}{L_q} & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} \quad (5.121)$$

positiv definit sein, was durch geeignete Wahl der Parameter  $\alpha_2 > 0$  und  $\alpha_3 > 0$  mit

$$\frac{1}{L_d} + \alpha_2 > 0 \quad \text{und} \quad \left(\frac{1}{L_d} + \alpha_2\right)\frac{1}{L_q} - \left(\frac{L_d - L_q}{L_dL_q\Phi}\right)^2 x_{2,d}^2 > 0 \quad (5.122)$$

sichergestellt wird. Die Zustandsrückführung errechnet sich dann gemäß (5.97) in der Form

$$\beta(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \left\{ (\mathbf{J}_d(\mathbf{x}) - \mathbf{S}_d)\left(\frac{\partial V_d}{\partial \mathbf{x}}\right)^T - (\mathbf{J}(\mathbf{x}) - \mathbf{S})\left(\frac{\partial V}{\partial \mathbf{x}}\right)^T - \mathbf{g}_d\Phi p x_{2,d} \right\}. \quad (5.123)$$

**Aufgabe 5.19.** Bestimmen Sie die expliziten Ausdrücke des Zustandsregelgesetzes (5.123).

## 5.8 Literatur

- [5.1] A. Kugi, *Skriptum zur VO Regelungssysteme 2 (SS 2019)*, Institut für Automatisierungs- und Regelungstechnik, TU Wien, 2019. Adresse: <https://www.acin.tuwien.ac.at/master/regelungssysteme-2/>.
- [5.2] P. Tabuada und G. Pappas, „From Nonlinear to Hamiltonian via Feedback,“ *IEEE Transactions on Automatic Control*, Jg. 48, Nr. 8, S. 1439–1442, 2003.
- [5.3] O. Föllinger, *Nichtlineare Regelung I + II*. München: Oldenbourg, 1993.
- [5.4] H. K. Khalil, *Nonlinear Systems (3rd Edition)*. New Jersey: Prentice Hall, 2002.
- [5.5] A. Kugi, *Non-linear Control Based on Physical Models*, Ser. Lecture Notes in Control and Information Sciences 260. London: Springer, 2001.
- [5.6] A. Kugi und K. Schlacher, „Analyse und Synthese nichtlinearer dissipativer Systeme: Ein Überblick (Teil 1),“ *at – Automatisierungstechnik*, Jg. 50, S. 63–69, 2002.
- [5.7] ———, „Analyse und Synthese nichtlinearer dissipativer Systeme: Ein Überblick (Teil 2),“ *at – Automatisierungstechnik*, Jg. 50, S. 103–111, 2002.
- [5.8] R. Lozano, B. Brogliato, O. Egeland und B. Maschke, *Dissipative Systems Analysis and Control*. London: Springer, 2000.
- [5.9] R. Ortega und E. García-Canseco, „Interconnection and Damping Assignment Passivity-Based Control: A Survey,“ *European Journal of Control*, Jg. 10, S. 432–450, 2004.
- [5.10] R. Ortega, A. van der Schaft, F. Castaños und A. Astolfi, „Control by Interconnection and Standard Passivity-Based Control of Port-Hamiltonian Systems,“ *IEEE Transactions on Automatic Control*, Jg. 53, Nr. 11, S. 2527–2542, 2008.
- [5.11] V. Petrović, R. Ortega und A. Stanković, „Interconnection and Damping Assignment Approach to Control of PM Synchronous Motors,“ *IEEE Transactions on Control Systems Technology*, Jg. 9, Nr. 6, S. 811–820, 2001.
- [5.12] R. Sepulchre, M. Janković und P. Kokotović, *Constructive Nonlinear Control*. London: Springer, 1997.
- [5.13] E. Slotine und W. Li, *Applied Nonlinear Control*. New Jersey: Prentice Hall, 1991.
- [5.14] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.