

Kai Michels

Regelungstechnik

Vorlesungsmanuskript

29. Juni 2022

Universität Bremen

Die meisten Probleme entstehen bei ihrer Lösung.

Leonardo da Vinci

Vorwort

Während ingenieurtechnische Fachrichtungen wie „Fahrzeugtechnik“, „Energietechnik“ oder „Softwareentwicklung“ sofort Assoziationen wecken und jeder Studienanfänger mit diesen Fächern gewisse Vorstellungen verbindet, erscheint die Regelungstechnik in dieser Hinsicht etwas farblos. Daher taucht sie auch bei Befragungen unter Studienanfängern, welche Fachrichtung sie später vertiefen möchten, so gut wie nie auf.

Dies ist umso erstaunlicher, als dass kaum eine Vertiefungsrichtung so breit aufgestellt und so vielseitig anwendbar ist wie die Regelungstechnik. Regler wirken in Kraftwerken, Flugzeugen und Automobilen genauso wie in Mobilfunkgeräten, und die Konzeption sowie die Auslegung solcher Regler sind Aufgabe des Regelungstechniklers. Selbst auf ökonomische oder biologische Prozesse sind die Methoden der Regelungstechnik anwendbar.

Um dieser umfassenden Aufgabe gewachsen zu sein, benötigt der Regelungstechniker nicht nur ein breites Wissen über die Methoden und Lösungsansätze der Regelungstechnik, sondern er ist mindestens ebenso stark angewiesen auf fundierte Kenntnisse des Systems, in dem der Regler wirken soll. Damit ist er ein Allrounder, der sich in einen Kraftwerksprozess ebenso einarbeiten können muss wie in die Funktionsweise eines Flugzeugs.

Entsprechend diesem Anforderungsprofil werden in diesem Buch zunächst die Analyse und strukturierte Darstellung bzw. Modellierung von dynamischen Systemen behandelt, und erst danach folgt die Darstellung regelungstechnischer Entwurfsverfahren. Auch diese folgt dabei den Erfordernissen der Praxis. Angesichts der Tatsache, dass nach wie vor der überwältigende Anteil aller industriellen Regler vom PID-Typ sind, bildet die klassische Entwurfsmethodik dieser Regler den Beginn der Darstellung. Anschließend wird auf Zustandsdarstellung und -regelung bis hin zu normoptimalen Regelungen eingegangen, die heutzutage die Grundlage aller Forschungsaktivitäten bilden und ohne die der Entwurf moderner Mehrgößenregelungen gar nicht möglich wäre.

Im nachfolgenden Kapitel geht es dann um diskrete Systeme, in dem die bis dahin gewonnenen Methoden auf die in der Realität üblichen Mikroprozessoren übertragen werden. Da diese nicht mehr, wie vorher angenommen, kontinuierlich, sondern nur noch zeitdiskret in festen Zeitabständen Algorithmen und Programme berechnen können und die Ausgangsgrößen derartiger

Regler keinen kontinuierlichen, sondern einen diskontinuierlichen, sprunghaften Verlauf aufweisen, hat dies auch Konsequenzen auf das Verhalten des gesamten Regelkreises. Die Digitalisierung bringt aus regelungstechnischer Sicht jedoch nicht nur neue Fragestellungen, sondern schafft auch neue Möglichkeiten, wie im darauf folgenden Kapitel über adaptive Regelungen gezeigt wird. Dabei handelt es sich um Regler, die sich über die Zeit an das zu regelnde System anpassen können und damit auch für Systeme geeignet sind, die sich im Laufe der Zeit verändern, beispielsweise durch Verschleiß, Verschmutzung oder den Verbrauch des Tankinhaltes bei Flugzeugen.

Während bis zu diesem Punkt ausschließlich lineare Systeme behandelt worden sind, d.h. Systeme, deren Verhalten durch lineare Differentialgleichungen beschrieben werden kann und die sich unter anderem dadurch auszeichnen, dass sich bei Veränderung ihrer Eingangsgröße um einen bestimmten Faktor die Ausgangsgröße um denselben Faktor verändert, geht es in den nachfolgenden Kapiteln um nichtlineare Systeme, die genau diese Eigenschaft nicht mehr besitzen und dadurch in ihrer Behandlung deutlich komplexer sind. In der Praxis trifft man leider in praktisch jedem System auf irgendwie geartete Nichtlinearitäten, mit denen man geeignet umgehen muss. Hierzu werden einige Methoden und Ansätze aufgezeigt.

Den Abschluss bildet die Behandlung von Fuzzy-Systemen und Fuzzy-Reglern. Hierbei handelt es sich um eine völlig andere methodische Herangehensweise an den Entwurf einer Regelung, die auf eine Idee aus den sechziger Jahren des letzten Jahrhunderts zurück geht und in den achtziger und neunziger Jahren sehr populär wurde. Mittlerweile hat sich die Aufregung um diese Methodik gelegt, aber sie hat ihren Platz in der Regelungstechnik gefunden.

Um die Vorstellung der derzeit in unserem Institut eingesetzten Verfahren zu komplettieren, fehlen noch Darstellungen zu modellprädiktiven Regelungen und Neuronalen Netzen, für die ich in den nächsten Jahren hoffentlich noch Zeit finden werde. Außerdem liegen die Kapitel über diskrete Systeme und adaptive Regelungen bisher nur auf Englisch vor. Auch deren Übersetzungen werden zu einem späteren Zeitpunkt noch erfolgen.

Das Buch richtet sich an Studierende der Ingenieurwissenschaften ab dem 5. Semester. Es ist so geschrieben, dass es prinzipiell zum Selbststudium geeignet ist. Genutzt wird es als Grundlage für meine Vorlesungen, um den Studierenden die Möglichkeit zu geben, sich voll und ganz auf den Vortrag zu konzentrieren ohne mitschreiben zu müssen.

Inhaltsverzeichnis

1. Grundbegriffe	1
2. Modell der Strecke	5
2.1 Problemstellung	5
2.2 Normierung	6
2.3 Elementare lineare Übertragungsglieder	7
2.4 Elementare nichtlineare Übertragungsglieder	11
2.5 Verzögerungsglieder erster und zweiter Ordnung	13
2.6 Anwendungsbereich	22
2.7 Linearisierung	23
2.8 Abschließende Bemerkungen	26
3. Übertragungsfunktion	27
3.1 Berechnung von Übertragungsfunktionen	27
3.2 Interpretation der Übertragungsfunktion	30
3.3 Berechnung der Sprungantwort	33
3.4 Vereinfachung einer Übertragungsfunktion	37
4. Frequenzgang	41
4.1 Einführung des Frequenzganges	41
4.2 Bode-Diagramm	42
4.2.1 Herleitung	42
4.2.2 Negativ reelle Polstelle	44
4.2.3 Positiv reelle Polstelle	46
4.2.4 Negativ reelle Nullstelle	47
4.2.5 Positiv reelle Nullstelle	48
4.2.6 Integrator	49
4.2.7 Differenzierer	51
4.2.8 Laufzeitglied	51
4.2.9 Polstellenpaar	53
4.2.10 Überlagerung von Teilfunktionen	56
4.3 Ortskurve	59
4.3.1 Verzögerungsglied erster Ordnung	60
4.3.2 Verzögerungsglied zweiter Ordnung	62

4.3.3	Integrator	63
4.3.4	Laufzeitglied	64
4.3.5	Integrator mit Verzögerungsglied	64
4.3.6	Zusammengesetzte Übertragungsfunktion	65
5.	Stabilität linearer Eingrößensysteme	69
5.1	Definition der Stabilität	69
5.2	Stabilität einer Übertragungsfunktion	72
5.3	Stabilität eines Regelkreises	73
5.4	Kriterium von Cremer, Leonhard und Michailow	75
5.5	Nyquist-Kriterium	77
5.6	Systeme mit nicht-minimaler Phase	84
6.	Regelung linearer Eingrößensysteme	91
6.1	Anforderungen an einen Regler	91
6.2	Regler vom PID-Typ und ihr Entwurf	95
6.2.1	P-Regler	95
6.2.2	I-Regler	97
6.2.3	PI-Regler	101
6.2.4	PID-Regler	113
6.2.5	Realisierung mit Operationsverstärkern	118
6.3	Weitere Reglerentwurfsverfahren	119
6.3.1	Wurzelortskurven	119
6.3.2	Reglerentwurf nach Ziegler-Nichols	121
6.3.3	Kompensationsregler und Youla-Parametrierung	121
6.3.4	Smith-Prädiktor	125
6.4	Strukturerweiterung	127
6.4.1	Vorfilter	127
6.4.2	Störgrößenaufschaltung	128
6.4.3	Ergänzende Rückführung	130
6.4.4	Kaskadenschaltung	132
6.4.5	Statische Vorsteuerung	134
6.4.6	Dynamische Vorsteuerung	136
6.4.7	Entkopplung	139
7.	Grundlagen der Zustandsdarstellung	143
7.1	Definition von Zustandsgrößen	143
7.2	Eigenschaften von Zustandsgrößen	145
7.3	Linearisierung	148
7.4	Zustandsdarstellung linearer Systeme	149
7.5	Normalformen	151
7.5.1	Regelungs- und Beobachtungs-Normalform	151
7.5.2	Jordansche Normalform	155
7.6	Koordinatentransformation	158
7.6.1	Grundlagen der Koordinatentransformation	158

7.6.2	Transformation auf Jordansche Normalform	160
7.6.3	Transformation auf Regulationsnormalform	162
7.7	Allgemeine Lösung der linearen Zustandsgleichung	165
7.8	Der Ljapunovsche Stabilitätsbegriff für lineare Systeme	167
7.9	Dominanzmaß von Eigenwerten	171
7.10	Steuerbarkeit und Beobachtbarkeit	174
7.10.1	Steuerbarkeit	175
7.10.2	Beobachtbarkeit	181
7.10.3	Abschließende Bemerkungen	185
8.	Zustandsregelung	189
8.1	Konzept der Zustandsregelung	191
8.2	Stationäre Genauigkeit von Zustandsreglern	192
8.3	Beobachter	200
8.4	Entwurf nach dem Polvorgabeverfahren	204
8.4.1	Herleitung	204
8.4.2	Anmerkungen	209
8.5	Falb-Wolovich-Entwurf	211
8.5.1	Differenzengrad	211
8.5.2	Regler-Entwurf	216
8.6	Riccati-Entwurf	221
8.6.1	Herleitung	221
8.6.2	Anmerkungen	228
8.6.3	Wahl der Gewichtungsmatrizen	230
8.6.4	Kalman-Filter	233
9.	Nullstellen	235
9.1	Einführung	235
9.2	Zusammenhang mit Steuer- und Beobachtbarkeit	237
9.3	Invariante Nullstellen	239
9.4	Anwendungen	242
9.4.1	Berechnung eines Vorfilters	242
9.4.2	Falb-Wolovich-Entwurf	243
10.	Normoptimale Regelung	245
10.1	Beziehungen im geschlossenen Kreis	245
10.2	Normen für Signale und Systeme	247
10.2.1	Grundbegriffe	247
10.2.2	Signalnormen	248
10.2.3	Normen von Eingrößensystemen	252
10.2.4	Beziehungen zwischen Systemgrößen	257
10.2.5	Normen von Mehrgrößensystemen	262
10.3	2-Norm-optimaler Regler für Eingrößensysteme	269
10.3.1	Koprime Faktorisierung	269
10.3.2	Die Menge aller G stabilisierenden Regler	273

10.3.3	Darstellung des Optimierungsproblems	278
10.3.4	Berechnung des optimalen $Q(s)$	280
10.3.5	Beispiele	284
10.4	Berechnung normoptimaler Regler im Zustandsraum	288
10.4.1	Formulierung der Entwurfsaufgabe	288
10.4.2	2-Norm-optimaler Regler	293
10.4.3	∞ -Norm-optimaler Regler	295
10.5	Modellunsicherheiten und Robustheit	298
10.5.1	Grundlagen	298
10.5.2	Zusammenhang mit normoptimalen Regelungen	303
10.5.3	Der strukturierte singuläre Wert	305
11.	Diskrete Systeme	311
11.1	Systemstruktur	311
11.2	Abtasttheorem	316
11.3	Differenzengleichung	319
11.4	Diskretes Zustandsraummodell	322
11.4.1	Einleitung	322
11.4.2	Normalformen	323
11.4.3	Diskretes und kontinuierliches Zustandsmodell	325
11.4.4	Eigenwerte von diskreten Systemen	327
11.4.5	Totzeiten in diskreter Zustandsdarstellung	328
11.4.6	Lösung der diskreten Zustandsgleichung	329
11.4.7	Stabilität diskreter Systeme	330
11.4.8	Steuerbarkeit und Beobachtbarkeit	331
11.5	Impuls- und Sprungantwort von diskreten Systemen	334
11.6	z-Transformation	338
11.6.1	Introduction	338
11.6.2	Theorems about z-Transformation	340
11.6.3	Discrete Transfer Function	343
11.6.4	Discrete Eigenvalues and Shannon Theorem	344
11.6.5	Damping of Discrete Systems	346
11.6.6	Discrete and Continuous Transfer Functions	348
11.6.7	Discrete Controller Design	352
12.	Adaptive Control	355
12.1	Motivation and Overview	355
12.2	Identification	358
12.2.1	Least Squares Method	358
12.2.2	Recursive Least Squares Method	359
12.3	Direct Adaptation	363
12.3.1	SISO systems	363
12.3.2	MIMO systems	367

13. Analyse nichtlinearer Systeme	373
13.1 Eigenschaften nichtlinearer Systeme	373
13.2 Schaltende Übertragungsglieder	374
13.2.1 Ideales Zweipunktglied	374
13.2.2 Zweipunktglied mit Hysterese	376
13.2.3 Dreipunktglied	378
13.2.4 Vorzeitiges Umschalten und Sliding Mode	380
13.2.5 Zeitoptimale Regelung	382
13.2.6 Pulsweitenmodulation	384
13.3 Definition der Stabilität bei nichtlinearen Systemen	386
13.3.1 Ruhelage	386
13.3.2 Stabilitätsdefinition nach Ljapunov	386
13.3.3 Stabilität von Trajektorien	388
13.3.4 Stabilität von linearen Systemen	390
13.3.5 Steuerbarkeit und Beobachtbarkeit	392
13.3.6 Definition der Ruhelage Null	392
13.4 Direkte Methode von Ljapunov	394
13.5 Harmonische Balance / Beschreibungsfunktion	398
13.5.1 Herleitung	398
13.5.2 Beispiele	402
13.5.3 Erweiterungen	408
13.6 Popov-Kriterium	411
13.7 Kreiskriterium	418
13.8 Hyperstabilität	425
13.8.1 Positiv reelle Systeme	425
13.8.2 Herleitung	426
14. Entwurf nichtlinearer Regelungen	439
14.1 Sliding Mode-Regler	439
14.2 Gain-Scheduling-Regler	444
14.2.1 GS-Regler und GS-Modelle	444
14.2.2 Stabilitätsanalyse nach Ljapunov	447
14.2.3 Normenbasierte Stabilitätsanalyse	454
15. Fuzzy-Systeme	459
15.1 Fuzzy-Mengen	459
15.2 Fuzzy-Logik	463
15.3 Operationen auf Fuzzy-Mengen	470
15.4 Fuzzy-Relationen	472
15.5 Fuzzy-Regler	476
15.6 Diskussion der Fuzzy-Regler	481
15.7 Praktische Aspekte	483
16. Neural Networks	485
16.1 Basics	485

A. Anhang	491
A.1 Laplace-Transformation	491
A.2 Korrespondenztabel zur Laplace-Transformation	494
A.3 Matrizen-e-Funktion	495
A.4 Theorem von Caley-Hamilton	498
A.5 Die Ljapunov-Gleichung	499
A.6 Statistische Grundbegriffe	501
A.7 Lineare Matrixungleichungen	504
Literaturverzeichnis	505
Index	511

1. Grundbegriffe

Die Regelungstechnik beschäftigt sich mit der Beeinflussung von Systemen, um bestimmten Ausgangsgrößen einen gewünschten zeitlichen Verlauf aufzuprägen. Dies können technische Systeme sein wie eine Raumheizung mit der Ausgangsgröße *Temperatur*, ein Schiff mit den Ausgangsgrößen *Kurs* und *Geschwindigkeit* oder ein Kraftwerk mit der Ausgangsgröße *abgegebene elektrische Leistung*. Es können aber auch soziale, chemische oder biologische Systeme sein, wie zum Beispiel das System *Volkswirtschaft* mit der Ausgangsgröße *Inflationsrate*. Die Natur der Systeme spielt keine Rolle. Lediglich ihr dynamisches Verhalten ist für den Regelungstechniker von Interesse. Dieses kann durch Differentialgleichungen, Differenzengleichungen oder andere Funktionalbeziehungen beschrieben werden. In der klassischen Regelungstechnik, die sich vorwiegend mit technischen Systemen beschäftigt, wird das zu beeinflussende System als *Strecke* bezeichnet.

Wie kann die Beeinflussung der Strecke erfolgen? Jede Strecke hat nicht nur Ausgangs-, sondern auch Eingangsgrößen. Bei der Raumheizung ist dies zum Beispiel die Stellung des Heizungsventils, beim Schiff die Leistung des Schiffsmotors und die Ruderstellung. Diese Eingangsgrößen sind so zu verstellen, dass die Ausgangsgrößen den gewünschten Verlauf aufweisen. Sie werden deshalb auch als *Stellgrößen* bezeichnet. Neben den Stellgrößen wirken auf die Strecke aber auch *Störgrößen* ein. Bei der Raumheizung wird die Temperatur zum Beispiel noch durch die Anzahl der Personen im Raum oder durch das Öffnen der Fenster beeinflusst, während beim Schiff Strömungen auftreten können, die den Kurs beeinflussen.

Der gewünschte zeitliche Verlauf der Ausgangsgrößen wird durch die *Sollgrößen* oder *Sollwerte* definiert. Diese können von Menschen festgelegt werden oder aber auch von einem völlig anderen System stammen. Ihre Entstehung soll hier nicht diskutiert werden, sie werden als gegeben hingenommen. Zu beachten ist, dass ein Sollwert nicht unbedingt einen konstanten Wert aufweisen muss. Er kann auch durchaus ein zeitveränderliches Signal sein.

Welche Information ist nun erforderlich, um Stellgrößen zu berechnen, die die Strecke so beeinflussen, dass die Ausgangsgrößen gleich den Sollwerten sind? Offensichtlich müssen die einzuhaltenden Sollwerte für die Ausgangsgrößen, das Verhalten der Strecke und der zeitliche Verlauf der Störgrößen bekannt sein. Damit lassen sich, zumindest theoretisch, Stellgrößen erzeugen,

die wiederum das System gerade so beeinflussen, dass die Ausgangsgrößen den vorgeschriebenen Verlauf haben. Dies ist das Prinzip einer *Steuerung* (Abb. 1.1). Eingangsgröße der Steuerung ist der Sollwert w , ihre Ausgangsgröße ist die Stellgröße u . Diese ist wiederum - zusammen mit der Störgröße d - eine Eingangsgröße der Strecke. y ist die Ausgangsgröße des Systems.

Der Nachteil des Verfahrens liegt auf der Hand. Entspricht das Verhalten der Strecke nicht den gemachten Annahmen oder treten unvorhergesehene Störungen auf, so werden die Ausgangsgrößen nicht mehr dem gewünschten Verlauf entsprechen. Eine Steuerung kann auf diese Abweichung aber nicht reagieren, da ihr die Ausgangsgrößen der Strecke gar nicht bekannt sind.

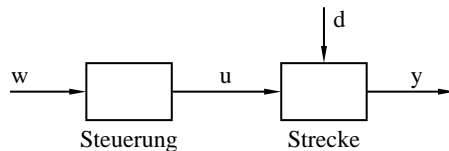


Abb. 1.1. Prinzip einer Steuerung

Als Verbesserung ergibt sich sofort das Prinzip einer *Regelung* (Abb. 1.2). Der Sollwert w (*Führungsgröße*) wird im Regler mit der gemessenen Ausgangsgröße der Strecke y (Istwert, *Regelgröße*) verglichen, und im *Regelglied* wird aus der Differenz e (*Regelabweichung* oder *Regelfehler*) eine geeignete Regler-Ausgangsgröße u berechnet. Früher ist das Regelglied selbst als Regler bezeichnet worden, doch weisen moderne Regler eine Struktur auf, in der sich Differenzbildung und Rechnen des Regelalgorithmus nicht mehr auf die gezeichnete Art und Weise trennen lassen. Deshalb geht man heute dazu über, den Block, in dem aus Führungs- und gemessener Regelgröße eine Regler-Ausgangsgröße erzeugt wird, als Regler zu bezeichnen.

Die Größe u liegt normalerweise als Signal mit niedriger Leistung, heutzutage meist als digitales Signal, vor. Mit niedriger Leistung ist aber eine Beeinflussung des physikalischen Prozesses nicht zu erreichen. Wie will man beispielsweise mit einem digital errechneten Ruderwinkel, also einer Folge aus Nullen und Einsen bei einer Spannung von 5 Volt, ein Schiff dazu bringen, den Kurs zu ändern? Da dies nicht auf direktem Wege möglich ist, sind beispielsweise noch ein Stromrichter und eine elektrische Rudermaschine erforderlich, die ihrerseits erst die Ruderstellung und damit auch den Kurs des Schiffes beeinflussen kann. Fasst man die Ruderstellung als Stellgröße des Systems auf, so bilden Stromrichter, Rudermaschine und Ruder zusammengefasst das *Stellglied*, in dem ein Signal niedriger Leistung, nämlich die Regler-Ausgangsgröße, in ein Signal hoher Leistung, die Ruderstellung, umgewandelt wird.

Man könnte aber beispielsweise auch die Ausgangsgröße des Stromrichters, also die Ankerspannung bzw. den Ankerstrom der Rudermaschine, schon als Stellgröße auffassen. In dem Fall bestände das Stellglied nur noch aus dem Stromrichter, während das dynamische Verhalten der Rudermaschine und

des Ruders selbst dem der Strecke hinzuzurechnen wäre. Daran wird deutlich, dass eine allgemein gültige Abgrenzung zwischen Stellglied und Strecke nicht möglich ist, auch wenn dies in der Literatur teilweise anders gesehen wird. Letztendlich ist die Aufteilung aber auch gar nicht erforderlich, denn für die Auslegung eines Reglers muss sowieso das gesamte Übertragungsverhalten von der Ausgangsgröße des Reglers bis zur Regelgröße berücksichtigt werden. Das Stellglied wird daher von nun an als Teil der Strecke betrachtet und die Regler-Ausgangsgröße im Folgenden als Stellgröße bezeichnet.

Für die Rückführung der Regelgröße zum Regler stellt sich in umgekehrter Richtung dieselbe Aufgabe wie für das Stellglied. Ein Signal hoher Leistung ist in ein Signal niedriger Leistung umzuformen. Dies geschieht im *Messglied*, dessen dynamische Eigenschaften entweder zu vernachlässigen oder, wie schon beim Stellglied, der Strecke hinzuzurechnen sind.

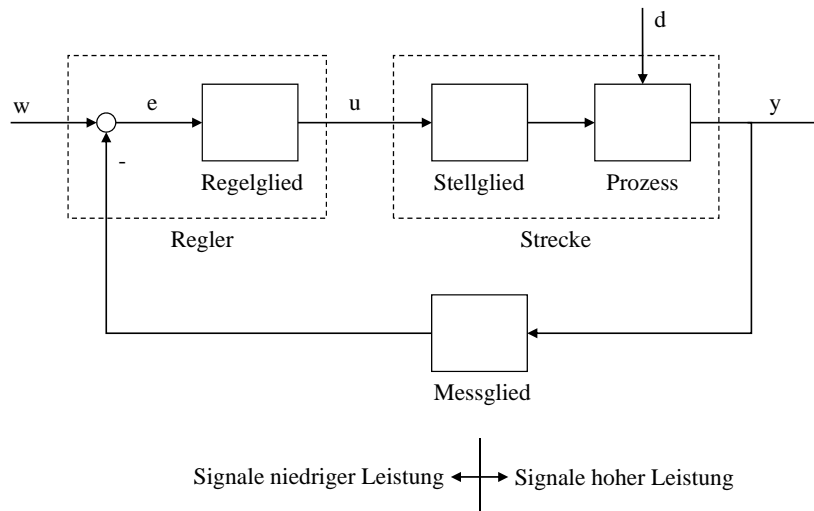


Abb. 1.2. Elemente eines Regelkreises

Durch die Rückkopplung entsteht ein entscheidendes Problem, das durch ein Beispiel verdeutlicht werden soll (Abb. 1.3). Die Regelstrategie für den Kursregler eines Schiffes könnte lauten: Je größer die Kursabweichung, desto stärker muss das Ruder in entgegengesetzter Richtung ausgelenkt werden. Oberflächlich betrachtet wirkt diese Strategie vernünftig. Falls aus irgendeinem Grund eine Kursabweichung vorliegt, wird das Ruder verstellt. Durch die Auslenkung des Ruders entsteht eine Drehbeschleunigung des Schiffes in Richtung des Sollkurses. Die Kursabweichung verringert sich, bis sie schließlich verschwindet. Die Drehgeschwindigkeit des Schiffes ist zu diesem Zeitpunkt aber nicht verschwunden, sie kann nur durch entgegengesetztes Auslenken des Ruders wieder zu Null gemacht werden. Im vorliegenden Beispiel wird das Schiff als Folge seiner Drehgeschwindigkeit nach Erreichen des Sollkurses

eine Kursabweichung zur anderen Seite erfahren. Erst dann wird die Drehgeschwindigkeit durch entgegengesetztes Auslenken des Ruders verschwinden. Da nun aber wieder eine Kursabweichung vorliegt, beginnt der ganze Vorgang mit anderem Vorzeichen von neuem. Die entstandene Kursabweichung ist möglicherweise sogar noch größer als die vorhergehende. Das Schiff wird sich in einem Schlingerkurs bewegen, dessen einzelne Auslenkungen im ungünstigsten Fall immer größer werden. Diesen Fall bezeichnet man als *Instabilität*. Bei gleichbleibenden Schwingungsamplituden spricht man vom *Stabilitätsgrenzfall*, nur bei abnehmenden Amplituden ist das System *stabil*. Um eine akzeptable Regelung zu erhalten, hätte man im vorliegenden Fall die Dynamik der Strecke bei der Auswahl der Regelstrategie berücksichtigen müssen. Ein geeigneter Regler erzeugt rechtzeitig eine Gegenauslenkung des Ruders, damit bei Erreichen des vorgegebenen Kurses auch die Drehgeschwindigkeit des Schiffes verschwunden ist.

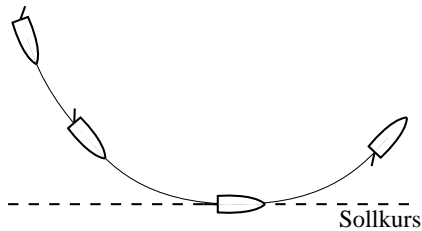


Abb. 1.3. Kursregelung eines Schiffes

Anhand dieses Beispiels werden auch die Anforderungen an eine Regelung klar. Eine Forderung ist die *Genauigkeit*, d.h. die Regelabweichung soll im stationären Zustand nach Beendigung aller Einschwingvorgänge möglichst klein sein. Eine weitere Forderung ist die *Schnelligkeit*, d.h. im Falle einer Führungsgrößenänderung oder einer Störung soll die entstandene Regelabweichung möglichst schnell wieder eliminiert werden. Man spricht in diesem Fall vom *Führungs-* bzw. *Störverhalten* des geschlossenen Regelkreises. Die dritte und wichtigste Forderung ist die nach der *Stabilität* des Gesamtsystems. Es wird sich noch zeigen, dass diese Forderungen einander teilweise widersprechen, so dass jeder Regler immer nur einen Kompromiss hinsichtlich dieser Forderungen darstellen kann.

2. Modell der Strecke

2.1 Problemstellung

Der Entwurf einer Regelung besteht aus zwei Schritten. Im ersten Schritt wird die Regelstrecke analysiert und ein sie beschreibendes, mathematisches Modell erstellt. Im zweiten Schritt wird auf der Basis dieses Modells ein Regler entwickelt.

Zur Beschreibung eines dynamischen Systems bieten sich Differential- oder Differenzengleichungen an. Es wird versucht, das dynamische Verhalten durch einen Satz möglichst einfacher Gleichungen zu beschreiben. Diese können dann durch graphische Symbole veranschaulicht und zu einem Strukturbild, dem sogenannten *Blockschaltbild* zusammengefasst werden. Die graphische Darstellung hat den Vorteil, dass die Wirkung der einzelnen Größen aufeinander relativ einfach zu überschauen ist, was später die Auslegung einer Regelung natürlich erleichtert.

An einem einfachen Beispiel soll die Vorgehensweise erläutert werden (Abb. 2.1): Gegeben sei ein Körper der Masse m , der über eine Feder an einer Wand befestigt und auf einer ebenen Unterlage in einer Richtung frei beweglich ist. Auf diesen Körper wirken die Antriebskraft f_a sowie die durch Reibung auf der Unterlage entstehende Kraft f_r und die Rückstellkraft der Feder f_f . Reibungs- und Rückstellkraft wirken der Antriebskraft entgegen gerichtet. Die Größe l gibt die Auslenkung des Körpers aus der Ruhelage an. Gesucht ist ein Modell, das den dynamischen Zusammenhang zwischen Antriebskraft und Auslenkung des Körpers beschreibt.

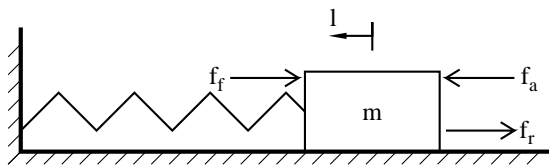


Abb. 2.1. Aufbau eines Feder-Masse-Systems

Nun sind die das System beschreibenden Gleichungen aufzustellen. Als erstes wird die Newtonsche Bewegungsgleichung betrachtet, die den Zusammenhang zwischen den am Körper angreifenden Kräften und der resultierenden

den Beschleunigung a angibt:

$$a(t) = \frac{1}{m}(f_a(t) - f_f(t) - f_r(t)) \quad (2.1)$$

Weiterhin gelten die durch Differentiation oder Integration definierten Zusammenhänge zwischen Beschleunigung und Geschwindigkeit v bzw. zwischen Geschwindigkeit und Weglänge l :

$$a(t) = \frac{dv(t)}{dt} \quad \text{bzw.} \quad v(t) = \int_{\tau=0}^t a(\tau) d\tau + v(0) \quad (2.2)$$

$$v(t) = \frac{dl(t)}{dt} \quad \text{bzw.} \quad l(t) = \int_{\tau=0}^t v(\tau) d\tau + l(0) \quad (2.3)$$

Die Rückstellkraft der Feder sei proportional zu ihrer Auslenkung, ebenso sei die durch die Reibung entstehende Kraft proportional zur Geschwindigkeit des Körpers:

$$f_f(t) = c_f l(t) \quad (2.4)$$

$$f_r(t) = c_r v(t) \quad (2.5)$$

2.2 Normierung

Die weiteren Betrachtungen werden deutlich vereinfacht, wenn alle Größen in den Systemgleichungen dimensionslos sind. Dazu müssen sie nun zunächst normiert werden. Die Normierung erfolgt, indem jede zeitveränderliche Größe durch eine konstante Normierungsgröße derselben Dimension dividiert wird.

Für die Normierung von Gleichung (2.1) sind eine konstante Beschleunigung a_0 , eine konstante Masse m_0 und eine konstante Kraft f_0 festzulegen. Dann wird die Gleichung durch a_0 und f_0 dividiert und mit m_0 multipliziert. Es ergibt sich

$$\frac{m_0}{f_0} \frac{a(t)}{a_0} = \frac{1}{a_0} \frac{1}{\frac{m}{m_0}} \frac{1}{f_0} (f_a(t) - f_f(t) - f_r(t)) \quad (2.6)$$

oder in vereinfachter Schreibweise

$$\frac{m_0 a_0}{f_0} a'(t) = \frac{1}{m'} (f'_a(t) - f'_f(t) - f'_r(t)) \quad (2.7)$$

mit den dimensionslosen Größen $a'(t) = \frac{a(t)}{a_0}$, $m'(t) = \frac{m(t)}{m_0}$ und $f'(t) = \frac{f(t)}{f_0}$. Wenn man darüber hinaus bei der Festlegung der Normierungsgrößen den Zusammenhang $f_0 = m_0 a_0$ eingehalten hat (was jedoch nicht notwendig ist), dann vereinfacht sich diese Gleichung zu

$$a'(t) = \frac{1}{m'}(f'_a(t) - f'_f(t) - f'_r(t)) \quad (2.8)$$

Im einfachsten Fall haben die Normierungsgrößen jeweils den Wert Eins, also $a_0 = 1 \frac{m}{s^2}$, $m_0 = 1 kg$ und $f_0 = 1 \frac{kgm}{s^2}$. Dann weist die normierte Systemgleichung (2.8) dieselben Zahlenwerte auf wie die nicht normierte Originalgleichung (2.1), und der Unterschied besteht lediglich in den weggelassenen Einheiten.

Es kann aber auch Fälle geben, in denen es sinnvoll ist, einzelne Größen nicht mit dem Wert Eins, sondern mit anderen Werten zu normieren. Beispielsweise, wenn die in eine Gleichung eingehenden Größen völlig verschiedene Größenordnungen aufweisen und es so zu numerischen Problemen kommen kann, oder auch, wenn ein Wertebereich auf das Einheitsintervall beschränkt werden soll.

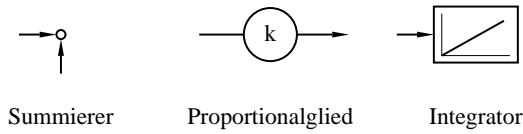
In der Praxis erfordern insbesondere Messaufnehmer und Stellglieder erhöhte Aufmerksamkeit hinsichtlich der Normierung. In den seltensten Fällen liefert ein Messaufnehmer den Wert der gemessenen Größe in SI-Einheiten. Üblich ist z.B. in der Anlagentechnik sogar eher, dass der Messbereich und das Umrechnungsverhältnis des Messaufnehmers einstellbar sind. Diese Umrechnungsfaktoren innerhalb von Messgliedern und analog in umgekehrter Richtung auch in Stellgliedern stellen in der Praxis eine der häufigsten Fehlerquellen bei der Auslegung von Regelungen dar.

Eine grundsätzliche Ausnahme ist die Zeit t . Sie sollte immer mit $t_0 = 1s$ normiert werden, um auch nach der Normierung noch eine Abschätzung der zeitlichen Abläufe zu ermöglichen.

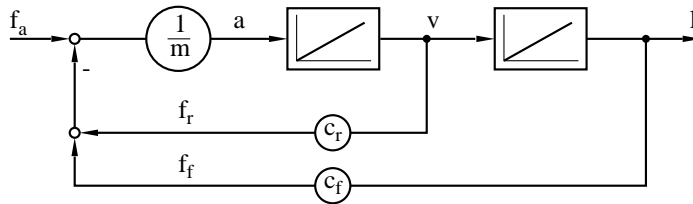
Im folgenden wird davon ausgegangen, dass alle auftretenden physikalischen Größen geeignet normiert sind, weshalb auf eine besondere Kennzeichnung verzichtet werden soll. Als Grundlage der weiteren Betrachtungen kann man damit wieder zu den Gleichungen (2.1) bis (2.5) zurückkehren und stillschweigend voraussetzen, dass es sich bei allen beteiligten Größen um dimensionslose Größen handelt.

2.3 Elementare lineare Übertragungsglieder

Die einzelnen Gleichungen müssen nun durch geeignete Symbole repräsentiert werden. Benötigt werden demnach graphische Darstellungen für die additive Überlagerung verschiedener Signale, die Multiplikation mit einem konstanten Faktor und die Integration einer Größe. Diese sogenannten *Übertragungsglieder* sind in Abb. 2.2 dargestellt. Die Ausgangsgröße des *Summierers* entspricht der Summe der beiden Eingangsgrößen, die Ausgangsgröße des *Proportionalgliedes* der mit k multiplizierten Eingangsgröße, und die Ausgangsgröße des *Integrators* dem über die Zeit t integrierten Eingangssignal, wobei die Ausgangsgröße zu einem Zeitpunkt $t = 0$ normalerweise als Null angenommen wird.

**Abb. 2.2.** Elemente von Blockschaltbildern

Mit diesen Elementen lässt sich das Blockschaltbild 2.3 der Strecke angeben. Die Summierer und das Proportionalglied $\frac{1}{m}$ stellen die Newtonsche Bewegungsgleichung (2.1) dar, wobei das negative Vorzeichen von f_f und f_r durch ein Minuszeichen am oberen Summierglied berücksichtigt ist. Der erste Integrator repräsentiert Gleichung (2.2), der zweite Integrator Gleichung (2.3). Die Proportionalglieder mit den Koeffizienten c_f und c_r stehen für die Gleichungen (2.4) und (2.5).

**Abb. 2.3.** Blockschaltbild eines Feder-Masse-Systems

Anhand dieses Blockschaltbildes wird auch klar, warum die Verwendung von Integrationsblöcken sinnvoller ist als die Verwendung von Differentiationsblöcken. So ist beispielsweise beim ersten Integrator die Beschleunigung die Eingangs- und die Geschwindigkeit die Ausgangsgröße. Dies entspricht auch den physikalischen Gegebenheiten, denn die Geschwindigkeit resultiert aus der Beschleunigung bzw. der Antriebskraft und nicht umgekehrt. Bei Verwendung eines Differenzierers müssten aber Ein- und Ausgangsgröße vertauscht werden, womit der Signalfluss im Blockschaltbild nicht mehr der Anschauung entspräche.

Dargestellt wird der Integrator durch einen Block, in dem ein Funktionsverlauf eingezeichnet ist. Dieser ist charakteristisch für den Integrator, es ist die sogenannte *Sprungantwort*. Springt das Eingangssignal eines Integrators zum Zeitpunkt $t = 0$ von Null auf Eins, so hat das Ausgangssignal des Integrators, wenn es vorher den Wert Null hatte, wegen der Integration den in Abb. 2.4 gezeigten, rampenförmigen Verlauf:

$$\frac{1}{T} \int_0^t 1 d\tau = \frac{t}{T} \quad (2.9)$$

In diese Gleichung wurde die sogenannte *Integrationszeitkonstante* T eingefügt, was in zweifacher Hinsicht sinnvoll ist. Zum einen werden beide Seiten der Gleichung dimensionslos, was bei der späteren Verwendung der Laplace-

Transformation und in Bode-Diagrammen zwingend erforderlich ist, zum anderen bietet diese Zeitkonstante als Parameter die Möglichkeit, eine bestimmte Integrationsgeschwindigkeit vorzugeben, d.h. einzustellen, wie schnell z.B. eine Sprungantwort ansteigt.

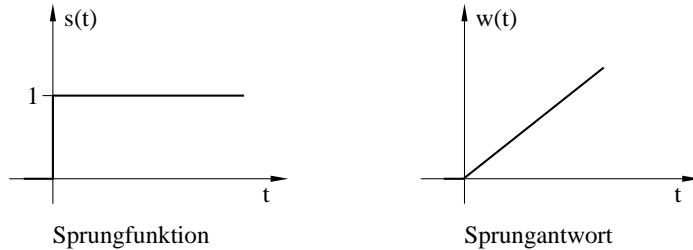


Abb. 2.4. Sprungantwort des Integrators

Es ist in der regelungstechnischen Literatur weit verbreitet, lineare Übertragungsglieder, d.h. Übertragungsglieder, die durch eine lineare Differentialgleichung beschrieben werden können, durch einen Block mit der zugehörigen Sprungantwort zu kennzeichnen. Auch im Folgenden soll diese Konvention, mit Ausnahme der ebenfalls linearen Übertragungsglieder *Summierer* und *Proportionalglied*, eingehalten werden.

Zu Integratoren ist noch eine weitere Überlegung interessant: Jeder Integrator lässt sich als Speicher auffassen, dessen Inhalt durch eine Eingangsgröße verändert wird. Sofern diese Eingangsgröße nicht unendlich groß wird, kann sich der Speicherinhalt nur stetig verändern.

Insbesondere kann man einen Integrator als Energiespeicher auffassen. Im Feder-Masse-System ist die Ausgangsgröße des ersten Integrators die Geschwindigkeit, sie repräsentiert die kinetische Energie der Masse, und die Ausgangsgröße des zweiten Integrators ist die Position, die die potentielle Energie der Feder repräsentiert.

Springt die Antriebskraft f_a von Null auf einen positiven Wert, so wird die Masse zunächst so lange nach links beschleunigt, bis die Reibungskraft f_r und die Rückstellkraft der Feder f_f diese Antriebskraft kompensieren. Der Körper bewegt sich aber weiter nach links, da die Geschwindigkeit als Ausgangsgröße des ersten Integrators noch positiv ist. Die im ersten Integrator enthaltene Energie treibt die Bewegung also weiter an. Dadurch wächst wiederum die der Antriebskraft entgegengerichtete Federkraft, wodurch die Summe der angreifenden Kräfte nun negativ wird. Die Beschleunigung der Masse, also die Eingangsgröße des ersten Integrators, wird negativ und bremst die Masse ab. Dies führt schließlich zum Stillstand der Masse. In diesem Extrempunkt enthält der erste Integrator keine Energie mehr ($v = 0$), während jetzt der zweite Integrator seinen Maximalwert erreicht hat, bei dem die Feder stark zusammengedrückt ist. Die Summe der Kräfte am Eingang des ersten Integrators ist nun negativ, und der gleiche Bewegungsablauf erfolgt in umgekehrter Richtung, wenn auch nicht mehr mit der gleichen Amplitude, da das System

durch die Reibung ja schon Energie verloren hat. Nach mehreren Schwingungen wird das System schließlich in einer Position zur Ruhe kommen, in der sich die Rückstellkraft der Feder und die (konstante) Antriebskraft f_a die Waage halten.

Grundsätzlich gilt, dass ein System so lange nicht zur Ruhe kommt, wie es noch Eingangsgrößen von Integratoren gibt, die von Null verschieden sind. Eine Ruheposition ist erst dann erreicht, wenn alle Eingangsgrößen aller Integratoren den Wert Null angenommen haben.

Abschließend sei für den weiteren Verlauf definiert, dass sämtliche Signale im System für $t < 0$ den Wert Null haben. Da die Ausgangsgrößen der Integratoren nur stetig veränderlich sind, folgt daraus, dass diese auch zum Zeitpunkt $t = 0$ noch den Wert $y(0) = 0$ aufweisen.

Lineare Übertragungsglieder zeichnen sich durch zwei besonders angenehme Eigenschaften aus, die durch die folgenden beiden Sätze charakterisiert werden:

Satz 2.1 (*Überlagerungssatz*) *Ein Übertragungsglied erzeuge aus dem Eingangssignal $x_1(t)$ das Ausgangssignal $y_1(t)$ und aus dem Eingangssignal $x_2(t)$ das Ausgangssignal $y_2(t)$. Es ist genau dann linear, wenn es aus dem Eingangssignal $a_1x_1(t) + a_2x_2(t)$ das Ausgangssignal $a_1y_1(t) + a_2y_2(t)$ erzeugt.*

Satz 2.2 *Entsteht ein Übertragungsglied durch Verknüpfung linearer Übertragungsglieder, so ist es ebenfalls linear.*

Summierer, Proportionalglied und Integrator sind lineare Übertragungsglieder. Die Bedeutung von Satz 2.2 ist weitreichend: Jede beliebige Kombination von linearen Übertragungsgliedern stellt wieder ein lineares Übertragungsglied dar. Darüber hinaus lässt sich nun auch erklären, warum die Sprungantwort zur Kennzeichnung linearer Glieder gewählt wurde. Bei einem linearen Übertragungsglied unterscheiden sich nämlich nach Satz 2.1 bei Eingangssprüngen verschiedener Amplitude die Systemantworten ebenfalls nur um eben diese Amplitude, nicht aber im prinzipiellen Verlauf. Insofern stellt die Sprungantwort schon eine relativ allgemeingültige Systemeigenschaft dar, und ihre Verwendung zur Charakterisierung des Systems erscheint gerechtfertigt.

Neben den drei in Abb. 2.2 vorgestellten Übertragungsgliedern existieren noch vier weitere elementare Bausteine von Blockschaltbildern. Der erste ist das Tot- oder *Laufzeitglied*. Es ist die graphische Repräsentation der Gleichung $y(t) = x(t - T_L)$, d.h. der zeitliche Verlauf des Ausgangssignales entspricht dem um die Laufzeit T_L verschobenen Verlauf des Eingangssignales (Abb. 2.5). Auch das Laufzeitglied ist ein lineares Übertragungsglied und wird durch seine Sprungantwort gekennzeichnet. Laufzeiten treten beispielsweise bei der Kommunikation über weite Entfernungen auf. So darf ein Mensch, der einen Roboter im All von der Erde aus steuert, nicht ungeduldig werden,

wenn der Roboter aus seiner Sicht nur zögerlich reagiert. Denn die zögerliche Reaktion ist nicht das Resultat einer schlecht konstruierten Maschine, sondern die Folge der langen Signallaufzeiten zwischen Mensch und Roboter. Ebenso muss ein automatischer Regler die in der Strecke vorhandenen Laufzeiten berücksichtigen.

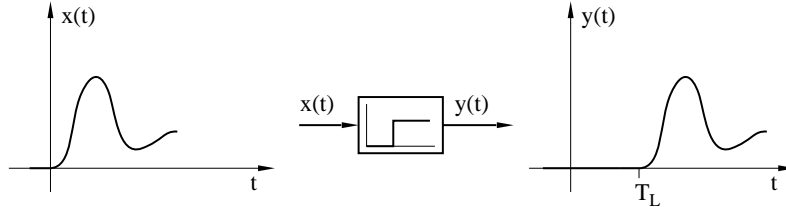


Abb. 2.5. Laufzeitglied

Ein etwas “erdnäheres” Beispiel ist das Einstellen der Wassertemperatur beim Duschen. Wenn man die Wassertemperatur verändern will, muss man an der Armatur drehen und dann einen gewissen Moment abwarten, da das Wasser ja erst den Weg von der Armatur zum Duschkopf zurücklegen muss, bevor man die Temperaturänderung spürt. Auch hier liegt also eine Laufzeit vor, und der Mensch als Regler handelt intuitiv richtig, wenn er zunächst abwartet, bevor er weitere Drehungen an der Armatur ausführt.

2.4 Elementare nichtlineare Übertragungsglieder

Weitere Bausteine sind *Multiplizierer* und *Dividierer* (Abb. 2.6). Im Gegensatz zum Proportionalglied wird beim Multiplizierer nicht nur ein Signal mit einem konstanten Faktor, sondern zwei zeitveränderliche Signale werden miteinander multipliziert. Ein Beispiel ist die Entstehung des Drehmomentes T_a in einer Gleichstrommaschine. Dieses ist proportional zum Produkt aus Ankerstrom i_a und Erregerfluss Φ_e , die beide unabhängig voneinander geregelt werden können:

$$T_a(t) = c i_a(t) \Phi_e(t) \quad (2.10)$$

Analog zum Multiplizierer werden beim Dividierer zwei zeitveränderliche Signale durcheinander dividiert, wobei eine Division durch Null natürlich auszuschließen ist. Ein Beispiel ist die Berechnung der Winkelgeschwindigkeit ω um die Hauptachse eines Roboters. In Analogie zur Newtonschen Bewegungsgleichung (2.1) gilt für rotatorische Bewegungen der Zusammenhang

$$\frac{d\omega}{dt} = \frac{1}{J(t)} T_a(t) \quad (2.11)$$

Dabei ist $T_a(t)$ das Antriebsmoment an der Hauptachse und $J(t)$ das Trägheitsmoment des rotierenden Körpers. Während einer Drehung des Roboters um seine Hauptachse kann sich die Stellung der anderen Drehgelenke,

also gewissermaßen die Armhaltung und damit auch das Trägheitsmoment des rotierenden Körpers, verändern. $J(t)$ ist deshalb in dieser Gleichung als zeitabhängige Größe anzusetzen, die aber auf jeden Fall von Null verschieden ist.

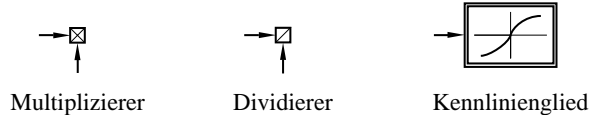


Abb. 2.6. Weitere elementare Übertragungsglieder

Die Gleichungen (2.10) und (2.11) sind wegen der Multiplikation bzw. Division zweier zeitveränderlicher Größen nichtlineare Differentialgleichungen, und die zugehörigen Übertragungsglieder sind damit im Gegensatz zu den vorher behandelten Gliedern nichtlineare Übertragungsglieder.

Dies gilt auch für den letzten Baustein, das *Kennlinienglied*. Mit Hilfe dieses Bausteins können beliebige statische Zusammenhänge repräsentiert werden. So steht der in Abb. 2.6 dargestellte Block für den Zusammenhang $y(t) = \sin x(t)$. Die Sinusfunktion tritt zum Beispiel auf, wenn die Bewegungsgleichung für ein frei aufgehängtes Pendel aufgestellt werden soll (Abb. 2.7).

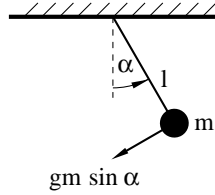


Abb. 2.7. Pendel

Für ein infinitesimal kleines Wegstück in tangentialer Bewegungsrichtung gilt zunächst

$$l \sin \alpha(t) \approx l \alpha(t) \quad (2.12)$$

und damit für die Beschleunigung

$$l \frac{d^2 \alpha(t)}{dt^2} \quad (2.13)$$

Der in tangentialer Bewegungsrichtung wirkende Anteil der Gewichtskraft ist

$$gm \sin \alpha(t) \quad (2.14)$$

und die Gleichung für das Kräftegleichgewicht in tangentialer Bewegungsrichtung lautet:

$$g m \sin \alpha(t) + ml \frac{d^2 \alpha(t)}{dt^2} = 0 \quad (2.15)$$

oder

$$\frac{d^2\alpha(t)}{dt^2} = \frac{-g}{l} \sin \alpha(t) \quad (2.16)$$

Dabei ist g die Erdbeschleunigung, l die Seillänge des Pendels und m seine punktförmig gedachte Masse. Das zugehörige Blockschaltbild zeigt Abb. 2.8. Das negative Vorzeichen für $\frac{g}{l}$ bedeutet, dass der Körper für positive Winkel wieder abgebremst und in entgegengesetzter Richtung beschleunigt wird. Weiterhin sind keine Reibungseffekte berücksichtigt, so dass das Modell ein ideales Pendel beschreibt, das, einmal angestoßen, ohne äußere Einwirkung unendlich lange weiterschwingt.

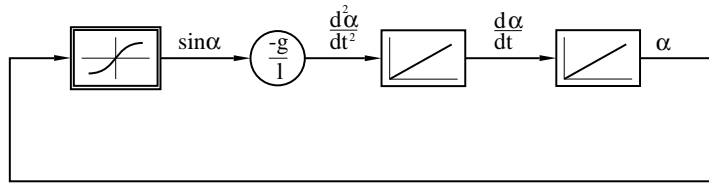


Abb. 2.8. Blockschaltbild des Pendels

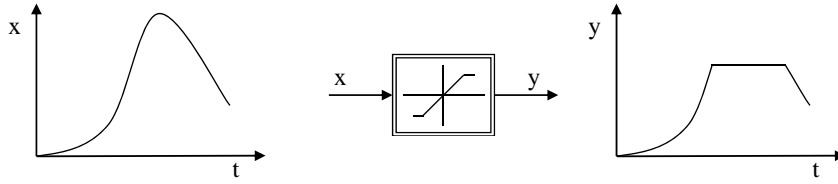
Im Gegensatz zu den linearen Übertragungsgliedern wird das Kennlinienglied durch einen doppelt umrandeten Block gekennzeichnet, in dessen Mitte die jeweilige Kennlinie eingezeichnet ist. Durch die doppelte Umrandung ist damit sofort klar, dass es sich bei der eingezeichneten Funktion nicht um eine Sprungantwort, sondern um eine Kennlinie handelt.

Mit Hilfe von Kennliniengliedern lassen sich auch die in der Praxis immer wieder auftretenden Begrenzungen von bestimmten Größen im Blockschaltbild darstellen. So kann beispielsweise das Ruder bei einem Schiff aus konstruktiven Gründen nur einen bestimmten Maximalwinkel α_{max} annehmen. Ist dieser erreicht, so kann der Winkel auch bei weiterem Ankerstrom auf die Rudermaschine nicht mehr wachsen. Der weitgehend lineare Zusammenhang zwischen Motorstrom und Ruderwinkel muss durch einen nichtlinearen Zusammenhang ersetzt werden, bei dem der Ruderwinkel unabhängig vom weiter wachsenden Ankerstrom auf dem Maximalwert verharret.

Dieses Verhalten wird durch die *Begrenzung* (Abb. 2.9) repräsentiert. Es ist offensichtlich, dass der Zusammenhang zwischen Eingangs- und Ausgangsgröße innerhalb gewisser Grenzen linear ist, während die Ausgangsgröße außerhalb der Grenzen unabhängig von der Eingangsgröße nur den Maximal- bzw. Minimalwert annehmen kann.

2.5 Verzögerungsglieder erster und zweiter Ordnung

Oft werden bestimmte, aus elementaren Übertragungsgliedern bestehende Strukturen zu einem einzigen Übertragungsglied zusammengefasst, um

**Abb. 2.9.** Begrenzung bzw. Begrenzungsglied

die Übersichtlichkeit des Blockschaltbildes zu erhöhen. Zwei dieser nicht-elementaren Übertragungsglieder kommen dabei besonders häufig vor, nämlich die Verzögerungsglieder erster und zweiter Ordnung, das PT_1 - und das PT_2 -Glied, die im Folgenden beschrieben werden.

Denkt man sich im Feder-Masse-System die Feder weggelassen, so besteht das System aus einem Körper auf einer reibungsbehafteten Unterlage, der von einer Antriebskraft beschleunigt wird. Untersucht werden soll der Zusammenhang zwischen Antriebskraft f_a und Geschwindigkeit v . Weglassen der Federkraft f_f in Gleichung (2.1) und Einsetzen der Gleichungen (2.2) und (2.5) liefert die gesuchte Differentialgleichung

$$\frac{dv(t)}{dt} + \frac{c_r}{m}v(t) = \frac{1}{m}f_a(t) \quad (2.17)$$

Mit $T = \frac{m}{c_r}$, $V = \frac{1}{c_r}$, $y(t) = v(t)$ und $x(t) = f_a(t)$ ergibt sich

$$T \frac{dy(t)}{dt} + y(t) = Vx(t) \quad (2.18)$$

bzw.

$$\begin{aligned} \frac{dy(t)}{dt} &= \frac{1}{T}(Vx(t) - y(t)) \\ y(t) &= \int_{\tau=0}^t \frac{1}{T}(Vx(\tau) - y(\tau))d\tau + y(0) \end{aligned} \quad (2.19)$$

Da alle Signale für $t < 0$ den Wert Null haben und das Ausgangssignal eines Integrators darüber hinaus stetig ist, gilt: $y(0) = 0$

Ebenso wie z.B. $y(t) = kx(t)$ das Übertragungsverhalten eines Proportionalgliedes beschreibt, so kennzeichnet diese Gleichung das Übertragungsverhalten eines sogenannten *Verzögerungsgliedes erster Ordnung* oder auch PT_1 -Gliedes.

Abb. 2.10 zeigt das zugehörige Blockschaltbild, welches unmittelbar aus Gleichung (2.19) hervorgeht. Dieses Blockschaltbild soll jetzt durch ein einziges, wegen Satz 2.2 lineares Übertragungsglied ersetzt werden, welches dann das PT_1 -Glied charakterisiert. Dazu ist die Berechnung der Sprungantwort aus der Differentialgleichung erforderlich:

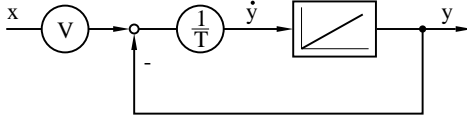


Abb. 2.10. Normiertes Blockschaltbild eines PT_1 -Gliedes

Aus der homogenen Differentialgleichung (vgl. (2.18))

$$T \frac{dy(t)}{dt} + y(t) = 0 \quad (2.20)$$

mit der charakteristischen Gleichung $Ts + 1 = 0$ und deren Lösung $s = -\frac{1}{T}$ ergibt sich zunächst als Lösung der homogenen Differentialgleichung

$$y_{hom}(t) = ae^{-\frac{t}{T}} \quad (2.21)$$

Nun ist die inhomogene Differentialgleichung zu betrachten. Da die Sprungantwort gesucht ist, wird die Eingangsgröße als Einheitssprung definiert ($x(t) = 1$ für $t \geq 0$ und $x(t) = 0$ für $t < 0$), und (2.18) wird zu

$$T \frac{dy(t)}{dt} + y(t) = V \quad (2.22)$$

Aufgrund der konstanten Anregung muss die Lösung der homogenen Differentialgleichung lediglich um einen konstanten Summanden b erweitert werden, um den Lösungsansatz für diese inhomogene Differentialgleichung zu erhalten:

$$y(t) = y_{hom}(t) + b = ae^{-\frac{t}{T}} + b \quad (2.23)$$

Dieser Lösungsansatz muss (2.22) erfüllen:

$$T(-\frac{a}{T}e^{-\frac{t}{T}}) + ae^{-\frac{t}{T}} + b = V \quad (2.24)$$

Da sich die ersten beiden Summanden auf der linken Seite der Gleichung aufheben, erhält man unmittelbar $b = V$.

Zur Bestimmung von a nutzt man die Festlegung aus, dass sich das System für $t < 0$ im Ruhezustand befinden soll und daher alle Signale für $t < 0$ den Wert Null haben. Da sich die Ausgangsgröße eines Integrators nur stetig verändern kann, muss seine Ausgangsgröße daher auch für $t = 0$ noch den Wert Null haben. Man erhält damit die Randbedingung $y(0) = 0$, die zur Bestimmung von a verwendet werden kann:

$$0 = y(0) = ae^{\frac{0}{T}} + b = a + V \quad (2.25)$$

Es ergibt sich $a = -V$, und damit für die Sprungantwort des Verzögerungsgliedes:

$$y(t) = V(1 - e^{-\frac{t}{T}}) \quad (2.26)$$

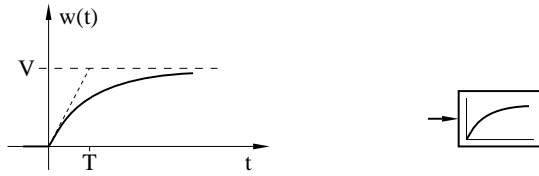


Abb. 2.11. Sprungantwort des PT_1 -Gliedes

Ihr Verlauf ist in Abb. 2.11 gezeigt, ebenso wie das gesuchte Übertragungsglied. Dieses Übertragungsglied mit den beiden Parametern V und T ersetzt das gesamte Blockschaltbild 2.10.

Die Sprungantwort zeigt, dass sich die Kurve umso schneller dem Endwert V nähert, je kürzer die Verzögerungszeit T ist. Erreicht wird der Endwert jedoch erst nach unendlich langer Zeit. Dieser Endwert ist mit dem Faktor V proportional zur Eingangsgröße, die bei der Sprungantwort ja konstant Eins ist.

Der Verlauf der Sprungantwort lässt sich für den beschleunigten Körper sehr einfach erklären. Setzt man die Antriebskraft f_a auf den konstanten Wert Eins, so wird die Geschwindigkeit ansteigen. Dadurch steigt aber die der Antriebskraft entgegengesetzte Reibungskraft f_r ebenfalls an, so dass die Summe aller Kräfte und damit die Beschleunigung bzw. der Anstieg der Geschwindigkeit immer kleiner wird. Nach einer gewissen Zeit (größer als T) ist der Endwert näherungsweise erreicht, und der Name *Verzögerungsglied* daher gerechtfertigt.

Für die Einführung des PT_2 -Gliedes (*Verzögerungsglied zweiter Ordnung*) wird das gesamte Feder-Masse-System betrachtet. Durch Einsetzen der Gleichungen (2.2) bis (2.5) in Gleichung (2.1) lassen sich die Größen v und a eliminieren, und es ergibt sich eine Differentialgleichung zweiter Ordnung für den Weg l :

$$\frac{m}{c_f} \frac{d^2 l(t)}{dt^2} + \frac{c_r}{c_f} \frac{dl(t)}{dt} + l(t) = \frac{1}{c_f} f_a(t) \quad (2.27)$$

Mit $\omega_0 = \sqrt{\frac{c_f}{m}}$, $D = \frac{\omega_0 c_r}{2c_f}$, $V = \frac{1}{c_f}$, $y(t) = l(t)$ und $x(t) = f_a(t)$ erhält man die Normalform

$$\frac{1}{\omega_0^2} \frac{d^2 y(t)}{dt^2} + \frac{2D}{\omega_0} \frac{dy(t)}{dt} + y(t) = Vx(t) \quad (2.28)$$

bzw.

$$\frac{d^2 y(t)}{dt^2} = \omega_0^2 \left[Vx(t) - \frac{2D}{\omega_0} \frac{dy(t)}{dt} - y(t) \right] \quad (2.29)$$

und aus (2.29) wieder unmittelbar das normierte Blockschaltbild 2.12. Da gegenüber den ursprünglichen Differentialgleichungen des Feder-Masse-Systems lediglich einige Umbenennungen erfolgt sind, weist dieses Blockschaltbild natürlich eine ähnliche Struktur auf wie das in Abb. 2.3.

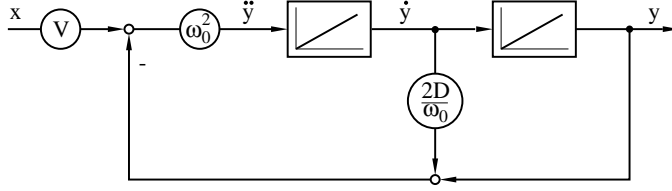


Abb. 2.12. Normiertes Blockschaltbild eines PT_2 -Gliedes

Wie schon beim PT_1 -Glied soll dieses Blockschaltbild nun durch ein einziges, lineares Übertragungsglied, das PT_2 -Glied, ersetzt werden. Um die graphische Darstellung zu erhalten, aber auch, um einige Erkenntnisse über das dynamische Verhalten des PT_2 -Gliedes zu gewinnen, ist wieder die Kenntnis der Sprungantwort erforderlich.

Die homogene Differentialgleichung lautet (vgl. (2.28))

$$\frac{1}{\omega_0^2} \frac{d^2 y(t)}{dt^2} + \frac{2D}{\omega_0} \frac{dy(t)}{dt} + y(t) = 0 \quad (2.30)$$

und ihr charakteristisches Polynom

$$\frac{1}{\omega_0^2} s^2 + \frac{2D}{\omega_0} s + 1 \quad (2.31)$$

hat die Nullstellen

$$s_{1,2} = \omega_0 \left[-D \pm \sqrt{D^2 - 1} \right] \quad (2.32)$$

Sofern die Differentialgleichung ein reales System beschreibt, gilt $\omega_0 > 0$ und $D \geq 0$. Für $D = 1$ weisen beide Nullstellen den gleichen negativ reellen Wert $s_{1,2} = -\omega_0 D$ auf. Für $D > 1$ ergeben sich zwei verschiedene Werte, die aber ebenfalls beide negativ reell sind. Für $0 < D < 1$ erhält man ein konjugiert-komplexes Nullstellenpaar mit negativem Realteil. Und für $D = 0$ das konjugiert-komplexe Nullstellenpaar $s_{1,2} = \pm j\omega_0$ auf der imaginären Achse.

Für die Lösung der Differentialgleichung sind zunächst nur die beiden Fälle $D = 1$ (Doppelnulstelle) und $D \neq 1$ (zwei einfache Nullstellen) zu unterscheiden.

Für $D \neq 1$ ergibt sich als Lösung der homogenen Differentialgleichung

$$y_{hom}(t) = ae^{s_1 t} + be^{s_2 t} \quad (2.33)$$

Der Fall $D = 1$ wird anschließend noch behandelt. Nun ist analog zur Vorgehensweise beim PT_1 -Glied wieder die inhomogene Differentialgleichung zu betrachten. Mit der Eingangsgröße $x(t) = 1$ für $t \geq 0$ ergibt sie sich aus (2.28) zu

$$\frac{1}{\omega_0^2} \frac{d^2 y(t)}{dt^2} + \frac{2D}{\omega_0} \frac{dy(t)}{dt} + y(t) = V \quad (2.34)$$

Aufgrund der konstanten Anregung muss die Lösung der homogenen Differentialgleichung wieder lediglich um einen konstanten Summanden (in diesem Fall c) erweitert werden, um den Lösungsansatz für diese inhomogene Differentialgleichung zu erhalten:

$$y(t) = y_{hom}(t) + c = ae^{s_1 t} + be^{s_2 t} + c \quad (2.35)$$

Setzt man diesen Ansatz in (2.34) ein, so ergibt sich wegen

$$\frac{dy(t)}{dt} = \frac{dy_{hom}(t)}{dt} \quad \text{und} \quad \frac{d^2 y(t)}{dt^2} = \frac{d^2 y_{hom}(t)}{dt^2} \quad (2.36)$$

der Zusammenhang

$$\frac{1}{\omega_0^2} \frac{d^2 y_{hom}(t)}{dt^2} + \frac{2D}{\omega_0} \frac{dy_{hom}(t)}{dt} + y_{hom}(t) + c = V \quad (2.37)$$

Da y_{hom} aber gerade die Lösung der homogenen Differentialgleichung (2.30) ist, muss die Summe der ersten drei Summanden verschwinden, und es bleibt $c = V$.

Zur Bestimmung von a und b nutzt man wie schon beim PT_1 -Glied die Festlegung aus, dass das System für $t < 0$ im Ruhezustand ist und die Ausgangsgrößen aller Integratoren zum Zeitpunkt $t = 0$ damit den Wert Null haben müssen. Die Ausgangsgröße des zweiten Integrators ist $y(t)$. Seine Eingangsgröße muss $\frac{dy}{dt}$ sein, und dies ist auch gerade die Ausgangsgröße des ersten Integrators. Damit ergeben sich für die beiden Integratoren des PT_2 -Gliedes die Randbedingungen $y(0) = 0$ und $\frac{dy}{dt}(0) = 0$.

Aus $\frac{dy}{dt}(0) = 0$ folgt

$$\begin{aligned} 0 &= \frac{dy}{dt}(0) = as_1 e^{s_1 0} + bs_2 e^{s_2 0} \\ a &= -\frac{s_2}{s_1} b \end{aligned} \quad (2.38)$$

und aus $y(0) = 0$ anschließend

$$0 = y(0) = ae^{s_1 0} + be^{s_2 0} + V = (1 - \frac{s_2}{s_1})b + V \quad (2.39)$$

Man erhält

$$a = \frac{s_2}{s_1 - s_2} V \quad b = \frac{s_1}{s_2 - s_1} V \quad (2.40)$$

und damit für die Sprungantwort

$$y(t) = V(1 + \frac{s_2}{s_1 - s_2} e^{s_1 t} + \frac{s_1}{s_2 - s_1} e^{s_2 t}) \quad (2.41)$$

Dieser Ausdruck gilt jedoch nur für $D \neq 1$. Für $D = 1$ sind s_1 und s_2 gemäß Gleichung (2.32) gleich, und man erhält eine einzige Nullstelle $s_1 = -\omega_0$ mit

der Vielfachheit zwei. Der Lösungsansatz für die homogene Differentialgleichung lautet jetzt

$$y_{hom}(t) = ae^{s_1 t} + bte^{s_1 t} \quad (2.42)$$

und für die inhomogene Differentialgleichung (2.34)

$$y(t) = y_{hom}(t) + c = ae^{s_1 t} + bte^{s_1 t} + c \quad (2.43)$$

Auch hier gilt offenbar der Zusammenhang (2.36), so dass sich nach Einsetzen dieses Ansatzes in (2.34) wieder (2.37) und demnach $c = V$ ergibt.

Zur Bestimmung von a und b werden wieder die Randbedingungen $y(0) = 0$ und $\frac{dy}{dt}(0) = 0$ genutzt:

$$\begin{aligned} 0 &= y(0) = a + V \\ 0 &= \frac{dy}{dt}(0) = as_1 e^{s_1 0} + b(e^{s_1 0} + 0s_1 e^{s_1 0}) = as_1 + b \end{aligned} \quad (2.44)$$

Man erhält $a = -V$ und $b = Vs_1$ und damit für $D = 1$ die Sprungantwort

$$y(t) = V [1 - (1 - s_1 t)e^{s_1 t}] \quad (2.45)$$

Zusammengefasst ergibt sich für die gesuchte Sprungantwort des PT_2 -Gliedes

$$y(t) = \begin{cases} V(1 + \frac{s_2}{s_1 - s_2} e^{s_1 t} + \frac{s_1}{s_2 - s_1} e^{s_2 t}) & : D \neq 1 \\ V[1 - (1 - s_1 t)e^{s_1 t}] & : D = 1 \end{cases} \quad (2.46)$$

so dass das PT_2 -Glied nun dargestellt werden kann (Abb. 2.13).

Zunächst ist festzustellen, dass die Anfangssteigung der Kurve, anders als beim PT_1 -Glied, Null ist. Dies ist anschaulich anhand des Blockschaltbildes 2.12 leicht zu erklären. Ein Sprung der Eingangsgröße erscheint (multipliziert mit $V\omega_0^2$) direkt am Eingang des ersten Integrators. Dementsprechend steigt seine Ausgangsgröße im ersten Augenblick wie die Sprungantwort eines Integrators rampenförmig an. Der zweite Integrator bekommt keinen Sprung, sondern - als Ausgangsgröße des ersten Integrators - eine Rampe als Eingangsgröße. Das Integral einer Rampenfunktion ist jedoch eine Parabel, und daher ist die Steigung der Ausgangsgröße dieses Integrators, die ja auch die Ausgangsgröße des PT_2 -Gliedes ist, im ersten Augenblick Null.

Der sogenannte *Dämpfungsfaktor* D bestimmt den weiteren Verlauf des Einschwingvorganges. Der Fall $D < 0$ ist, wie schon gesagt, physikalisch nicht möglich und muss daher nicht betrachtet werden. Für $D > 0$ weisen s_1 und s_2 einen negativen Realteil auf. Damit konvergieren die Exponentialfunktionen in der Sprungantwort immer gegen Null. Je größer D ist, desto größer wird auch der Betrag von s_1 und s_2 , und damit konvergieren auch die Exponentialfunktionen umso schneller gegen Null. Die Sprungantwort konvergiert damit für $D > 0$ immer gegen den Endwert V , und zwar umso schneller, je größer D ist.

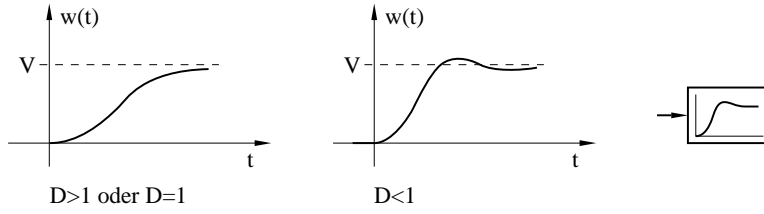


Abb. 2.13. Sprungantwort des PT_2 -Gliedes für $D \geq 0$

Für $D \geq 1$ sind s_1 und s_2 sogar rein reell, und die Sprungantwort sieht in diesem Fall, abgesehen von der Anfangssteigung, ähnlich aus wie beim PT_1 -Glieder. Man spricht vom *aperiodischen Fall*.

Für $0 \leq D < 1$ lässt sich schreiben: $s_{1,2} = \omega_0 [-D \pm j\sqrt{1-D^2}]$. Die Exponenten in (2.46) sind jetzt komplex. In diesem Fall bieten sich einige Umformungen an:

$$\begin{aligned}
 y(t) &= V \left[1 + \frac{s_2}{s_1 - s_2} e^{s_1 t} + \frac{s_1}{s_2 - s_1} e^{s_2 t} \right] \\
 &= V \left[1 + \frac{\omega_0 (-D - j\sqrt{1-D^2})}{2\omega_0 j\sqrt{1-D^2}} e^{-\omega_0 D t} e^{\omega_0 j\sqrt{1-D^2} t} \right. \\
 &\quad \left. + \frac{\omega_0 (-D + j\sqrt{1-D^2})}{-2\omega_0 j\sqrt{1-D^2}} e^{-\omega_0 D t} e^{-\omega_0 j\sqrt{1-D^2} t} \right]
 \end{aligned}$$

Mit der Abkürzung $\alpha = \sqrt{1-D^2}\omega_0 t$ wird daraus

$$\begin{aligned}
 y(t) &= V \left[1 + \frac{e^{-\omega_0 D t}}{2\sqrt{1-D^2}} \left((jD - \sqrt{1-D^2})e^{j\alpha} - (jD + \sqrt{1-D^2})e^{-j\alpha} \right) \right] \\
 &= V \left[1 + \frac{e^{-\omega_0 D t}}{2\sqrt{1-D^2}} \left((jD - \sqrt{1-D^2})(\cos \alpha + j \sin \alpha) \right. \right. \\
 &\quad \left. \left. - (jD + \sqrt{1-D^2})(\cos \alpha - j \sin \alpha) \right) \right] \\
 &= V \left[1 + \frac{e^{-\omega_0 D t}}{2\sqrt{1-D^2}} \left(-2D \sin \alpha - 2\sqrt{1-D^2} \cos \alpha \right) \right] \\
 &= V \left[1 - \frac{e^{-\omega_0 D t}}{\sqrt{1-D^2}} \left(D \sin \alpha + \sqrt{1-D^2} \cos \alpha \right) \right] \\
 &= V \left[1 - \frac{e^{-\omega_0 D t}}{\sqrt{1-D^2}} \left(D \sin \alpha + \sqrt{1 - (\cos \arccos D)^2} \cos \alpha \right) \right] \\
 &= V \left[1 - \frac{e^{-\omega_0 D t}}{\sqrt{1-D^2}} \left((\cos \arccos D) \sin \alpha + (\sin \arccos D) \cos \alpha \right) \right] \\
 &= V \left[1 - \frac{e^{-\omega_0 D t}}{\sqrt{1-D^2}} \left(\sin(\alpha + \arccos D) \right) \right]
 \end{aligned}$$

$$= V \left[1 - \frac{e^{-\omega_0 D t}}{\sqrt{1-D^2}} \sin(\sqrt{1-D^2} \omega_0 t + \arccos D) \right] \quad (2.47)$$

Das System ist jetzt schwingungsfähig, wie man an der Sinusfunktion erkennen kann. Wie schnell die Schwingungen nach einer Anregung, also beispielsweise einem Sprung der Eingangsgröße, abklingen, hängt auch hier vom Exponenten der e-Funktion und damit von D ab.

Der Parameter ω_0 ist die *Eigenkreisfrequenz* des Systems. Aus Gleichung (2.47) sieht man, dass für $D = 0$ die Sprungantwort des Systems eine Sinusschwingung mit konstanter Amplitude und eben dieser Frequenz ω_0 ist (Bild 2.14). Bei $0 < D < 1$ ergibt sich eine abklingende Schwingung mit der - etwas kleineren - *natürlichen Kreisfrequenz* $\omega_n = \omega_0 \sqrt{1-D^2}$.

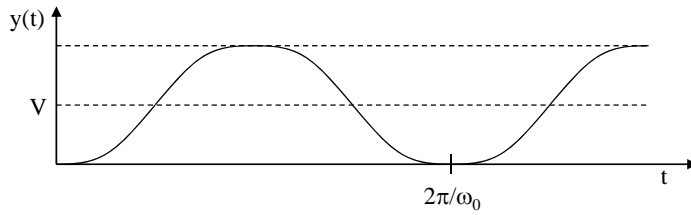


Abb. 2.14. Sprungantwort des PT_2 -Gliedes für $D = 0$

Interessanterweise besteht ein sehr einfacher geometrischer Zusammenhang zwischen der das Einschwingverhalten des Systems bestimmenden Größe D und der Lage der Nullstellen $s_{1,2}$ in der komplexen Ebene (Abb. 2.15):

$$\cos \alpha_0 = \frac{|\operatorname{Re}(s_1)|}{\sqrt{(\operatorname{Re}(s_1))^2 + (\operatorname{Im}(s_1))^2}} = \frac{\omega_0 D}{\omega_0 \sqrt{D^2 + 1 - D^2}} = D \quad (2.48)$$

Je größer also der Winkel α_0 wird, desto kleiner wird die Dämpfung. Für $\alpha_0 = 0$ ($D = 1$) weist das System ein aperiodisches Einschwingverhalten auf, während für $\alpha_0 = \pi/2$ ($D = 0$) die Schwingungen überhaupt nicht mehr abklingen. Dieser Zusammenhang wird im weiteren Verlauf der Vorlesung noch genutzt.

Nach Einführung des PT_2 -Gliedes kann das gesamte Blockschaltbild 2.3 durch ein einziges PT_2 -Glieder mit der Eingangsgröße f_a und der Ausgangsgröße l ersetzt werden. Die inneren Größen *Beschleunigung* und *Geschwindigkeit* treten dann aber nicht mehr explizit auf.

PT_1 - und PT_2 -Glieder sind nach Satz 2.2 offensichtlich lineare Übertragungsglieder. Sie werden häufig auch benutzt, um kompliziertere Strukturen näherungsweise zu beschreiben. Wenn beispielsweise für das Feder-Masse-System am Anfang dieses Kapitels eine Regelung entworfen werden soll, so muss auch die Entstehung der Antriebskraft f_a in einer Maschine berücksichtigt werden. Hier bietet es sich an, die maschineninternen Vorgänge nicht

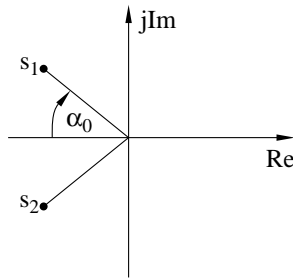


Abb. 2.15. Pole des PT_2 -Gliedes

explizit zu modellieren, sondern das Übertragungsverhalten von der Stellgröße des Reglers (also der Eingangsgröße der Maschine) bis zur mechanischen Kraft durch ein Verzögerungsglied angenähert zu beschreiben. Wegen der Schnelligkeit der maschineninternen Ausgleichsvorgänge im Vergleich zu den dynamischen Vorgängen im Feder-Masse-System ist der dadurch entstehende Modellfehler relativ gering. Wie eine solche Vereinfachung im einzelnen vorzunehmen ist, wird später noch beschrieben.

Abschließend sollen noch einmal die Unterschiede zwischen PT_1 - und PT_2 -Glied herausgestellt werden: Der erste Unterschied besteht in der Anfangssteigung bei der Sprungantwort. Diese ist beim PT_1 -Glied von Null verschieden, während sie beim PT_2 -Glied Null ist. Der zweite Unterschied besteht in der Schwingungsfähigkeit des Systems. Das PT_2 -Glied ist für $D < 1$ schwingungsfähig, während das PT_1 -Glied grundsätzlich nie schwingungsfähig ist. Damit ist das PT_1 -Glied aus regelungstechnischer Sicht deutlich angenehmer, da es einerseits schneller reagiert und andererseits ohne Schwingungen gegen den Endwert konvergiert.

2.6 Anwendungsbereich

Obwohl die Beispiele in diesem Kapitel ausschließlich mechanischer Natur waren, so lassen sich mit den jetzt bekannten Übertragungsgliedern auch dynamische Vorgänge z.B. in der Elektrotechnik oder Hydromechanik beschreiben. Prinzipiell kann jedes System beschrieben werden, das die folgenden Eigenschaften erfüllt:

- Das System ist *zeitinvariant*, die Parameter der Strecke bleiben im Laufe der Zeit konstant. Ein Beispiel für ein zeitvariantes System ist ein Flugzeug, dessen Tankinhalt während des Fluges verbraucht wird. Dadurch ändert sich auch das Gewicht, also ein Streckenparameter, und damit das dynamische Verhalten des Flugzeuges.
- Das System ist *kontinuierlich*, der Verlauf der Signale ist für jeden Zeitpunkt gegeben. Im Gegensatz dazu stehen *zeitdiskrete* Systeme, bei denen

der Wert eines Signales nur zu bestimmten Zeitpunkten bekannt ist. Werden für die Regelung zum Beispiel Mikroprozessoren eingesetzt, so tritt durch die notwendige Analog/Digital-Wandlung eine Zeitdiskretisierung auf (Abb. 2.16). Da Information verloren geht, muss diese Tatsache bei der Auslegung des Reglers berücksichtigt werden. Hierzu existiert eine umfangreiche Theorie [1, 19, 20, 35, 64], die in Kapitel 11 behandelt wird. Wenn die Abtastrate des A/D-Wandlers allerdings hoch ist im Vergleich zu den Frequenzen der Streckensignale, kann man die Zeitdiskretisierung vernachlässigen und das System als kontinuierlich behandeln, und genau davon soll im Folgenden auch ausgegangen werden, auch wenn ggf. ein digitaler Regler im Einsatz ist.

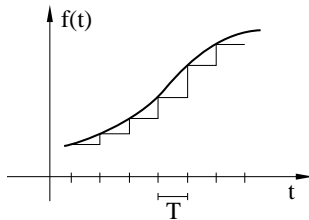


Abb. 2.16. Zeitdiskretisierung eines Signals

- Die Parameter des Systems sind *konzentriert*. So ist zum Beispiel die Temperatur in einem Raum nicht nur zeit-, sondern auch ortsabhängig. Zur Beschreibung der Vorgänge sind daher partielle Differentialgleichungen notwendig. Jedes Volumenelement stellt einen kleinen Energiespeicher dar und tritt mit den benachbarten Volumenelementen in Wechselwirkung. Würde man versuchen, das System durch ein Blockschaltbild mit den vorgestellten Übertragungsgliedern zu modellieren, so bräuchte man unendlich viele Bausteine, weil jedes Volumenelement einzeln modelliert werden muss. Die Parameter dieses Systems sind nicht konzentriert. In solchen Fällen wird auf eine exakte Modellierung verzichtet und das System näherungsweise durch einige wenige Energiespeicher und damit eine endliche Anzahl an Übertragungsgliedern angenähert.

2.7 Linearisierung

Sowohl lineare als auch nichtlineare Differentialgleichungen lassen sich durch Übertragungsglieder bzw. Blockschaltbilder repräsentieren. Da die weitergehende Behandlung, und zwar insbesondere der Reglerentwurf, für lineare Systeme aber wesentlich einfacher ist, versucht ein Regelungstechniker grundsätzlich zunächst einmal, sein zu behandelndes System durch lineare Differentialgleichungen zu beschreiben.

Andererseits kommen aber in den meisten realen Regelstrecken Nichtlinearitäten vor, die nicht einfach vernachlässigt werden dürfen. Hier behilft

man sich damit, nichtlineare Strecken zu *linearisieren*: Es wird ein *Arbeitspunkt* definiert, in dessen Umgebung man eine nichtlineare Funktion durch eine lineare Funktion annähert. Dies geschieht durch Entwicklung in eine Taylorreihe, die nach dem ersten Glied abgebrochen wird. Gegeben sei beispielsweise ein nichtlineares Glied mit der Übertragungsfunktion $y = f(x)$ und dem Arbeitspunkt $(x_0, f(x_0))$. Für die Abweichungen vom Arbeitspunkt gilt $\Delta x = x - x_0$ und $\Delta y = \Delta f = f(x) - f(x_0)$. Die Entwicklung der Funktion $f(x)$ in eine Taylorreihe um den Arbeitspunkt ergibt:

$$f(x) = f(x_0) + \frac{\partial f}{\partial x}(x_0) \Delta x + r(x) \quad (2.49)$$

$r(x)$ stellt dabei ein Restglied mit den höheren Ableitungen der Funktion $f(x)$ dar und soll vernachlässigt werden. Betrachtet man nun anstelle der Größen f und x nur noch ihre Abweichungen vom Arbeitspunkt Δf und Δx , so ergibt sich näherungsweise ein linearer Zusammenhang zwischen Eingangs- und Ausgangsgröße des Übertragungsgliedes:

$$\Delta f = f(x) - f(x_0) \approx \frac{\partial f}{\partial x}(x_0) \Delta x = k \Delta x \quad (2.50)$$

Ein anschauliches Beispiel stellt wieder das Pendel dar (Abb. 2.7). Das Verhalten des Systems soll am Arbeitspunkt $\alpha_0 = 0$ linearisiert werden. Dazu ist die Sinusfunktion in Gleichung (2.16) durch ein lineares Übertragungsglied zu ersetzen. Mit $f(\alpha) = \sin \alpha$ gilt

$$\Delta f \approx \frac{\partial f}{\partial \alpha}(\alpha_0) \Delta \alpha = \cos 0 (\alpha - 0) = 1 \alpha \quad (2.51)$$

Demnach kann die Sinusfunktion in einer Umgebung des Arbeitspunktes auch durch ein Proportionalglied mit dem Faktor Eins ersetzt werden. Dies bedeutet aber, dass das Kennlinienglied im Blockschaltbild 2.8 auch völlig entfallen kann, womit das System dann linear ist (Abb. 2.17).

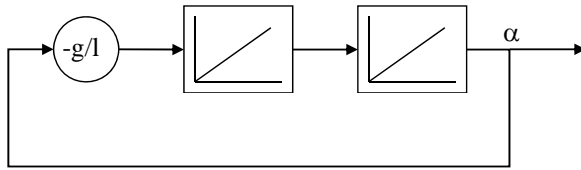


Abb. 2.17. Linearisiertes Blockschaltbild des Pendels

Ein Beispiel für die praktische Anwendung der Linearisierung ist das sogenannte *Invertierte Pendel* (Abb. 2.18). Es besteht aus einem Wagen, der durch eine Kraft f_a sowohl vorwärts als auch rückwärts bewegt werden kann. An einem Drehgelenk ist ein Stab mit einer Masse befestigt. Der Stab kann in Bewegungsrichtung des Wagens kippen. Die Aufgabe der Regelung besteht nun darin, den Stab durch geeignetes Bewegen des Wagens mit Hilfe

der Stellgröße f_a in der senkrechten Position ($\alpha = 0$) zu halten. Da die waagrecht wirkende Antriebskraft f_a in der Bewegungsgleichung für die Masse an der Spitze des Stabes mit dem Faktor $\cos(\alpha)$ auftaucht, ist das System nichtlinear.

Hier bietet es sich an, für betragsmäßig große Winkel eine (primitive) nichtlineare Regelung zu implementieren, die für $\alpha > 0$ die maximale negative Stellgröße und für $\alpha < 0$ die maximale positive Stellgröße ausgibt. Auf diese Art und Weise wird der Stab - sofern er nicht schon zu weit gekippt ist - zunächst schnellstmöglich in eine annähernd aufrechte Position gebracht. Einen solchen Regler, der nur entweder den positiven oder negativen Maximalwert als Stellgröße ausgibt, bezeichnet man als *Zweipunktregler* oder *Bang Bang Control*.

Es ist offensichtlich, dass dieser Regler zum Aufrichten des Pendels bei großen Auslenkungen eine sehr gute Wahl ist. Um das Pendel dann aber in einer aufrechten Position zu halten, ist er ungeeignet. Er würde bei einer sehr kleinen positiven Auslenkung $\alpha > 0$ schon mit der maximalen negativen Stellgröße reagieren, und entsprechend mit der maximalen positiven Stellgröße auch bei kleinsten negativen Winkeln. Der Wagen würde dauernd mit maximaler Kraft hin und her bewegt, was natürlich keine optimale Regelstrategie sein kann.

Hier bietet es sich an, für den Bereich kleiner Winkel eine andere Regelung zu entwerfen, die auf eine Auslenkung des Stabes mit einer angemessenen Stellgröße reagiert. Diese Aufgabe kann eine lineare Zustandsregelung übernehmen (vgl. Kap. 8). Zu ihrer Auslegung wird das Streckenmodell im Bereich kleiner Winkel um den Arbeitspunkt $\alpha = 0$ linearisiert und anschließend auf Basis des linearen Modells ein linearer Regler entworfen.

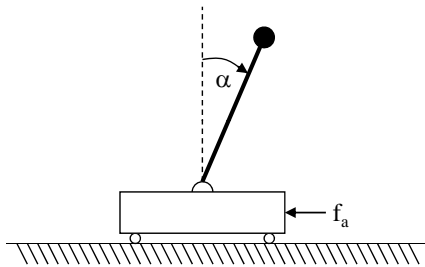


Abb. 2.18. Invertiertes Pendel

Abschließend soll aber noch einmal vor der bedenkenlosen Anwendung von linearisierten Modellen gewarnt werden, da ein durch eine Linearisierung gewonnenes Streckenmodell nur in einem begrenzten Bereich um den Arbeitspunkt Gültigkeit besitzt.

2.8 Abschließende Bemerkungen

Es ist leicht einzusehen, dass es für ein und dieselbe Strecke verschiedene Blockschaltbilder geben kann. Je nach Wahl der Zwischengrößen und Übertragungsblöcke entstehen völlig unterschiedliche Strukturen, die aber zueinander äquivalent sind. Entscheidend für die Aufteilung des Blockschaltbildes ist meistens, welche Größen tatsächlich messbar sind oder eine bestimmte physikalische Bedeutung haben. In diesem Fall spiegelt das Blockschaltbild den realen Aufbau eines Systems recht gut wider, und die Verknüpfung und gegenseitige Beeinflussung der einzelnen physikalischen Größen ist wesentlich einfacher zu überblicken als bei einem Gleichungssystem. Dies ist für den Regelungstechniker hilfreich, denn trotz aller Möglichkeiten, die sich durch den Einsatz von Computern im Bereich der Regelungstechnik eröffnet haben, ist immer noch ein gewisses Maß an Intuition und Übersicht bei der Auslegung einer Regelung erforderlich.

Zur Bestimmung von Blockschaltbildern stehen statistische Verfahren zur Verfügung, mit deren Hilfe aus den Messwerten der Strecke Informationen über ihre Struktur gewonnen werden können [21, 22, 33, 37, 65]. Der Einsatz dieser Verfahren ist unerlässlich, wenn die physikalischen Zusammenhänge der Strecke nicht mehr so einfach zu überschauen sind wie in den Beispielen dieses Kapitels, oder auch, wenn zwar die Struktur, nicht aber die einzelnen Parameter der Differentialgleichungen bekannt sind. Wenn das Blockschaltbild erstellt ist, existiert darüber hinaus die Möglichkeit einer Simulation der Strecke mit Hilfe numerischer Integrationsverfahren. Dadurch können weitere Einblicke in das Streckenverhalten gewonnen und nicht zuletzt auch Regler überprüft werden.

Prinzipiell sind alle in diesem Kapitel vorgestellten Gedanken auch auf Übertragungsglieder mit mehreren Ein- und Ausgangsgrößen erweiterbar. Auf diese Erweiterung soll aber zunächst noch verzichtet werden, um das Verständnis zu erleichtern.

3. Übertragungsfunktion

Für dieses und die folgenden Kapitel muss die Klasse der betrachteten Übertragungsglieder bzw. Systeme eingeschränkt werden, und zwar auf die rein linearen, zeitinvarianten Systeme mit einer Ein- und einer Ausgangsgröße (SISO-Systeme: single input single output). Vor einer Anwendung der vorgestellten Verfahren ist also gegebenenfalls eine Linearisierung der nichtlinearen Übertragungsglieder sowie eine Vernachlässigung oder Umrechnung von Zeitvarianzen vorzunehmen.

3.1 Berechnung von Übertragungsfunktionen

Gegeben sei das Problem, dass in einem Regelkreis der Verlauf des Eingangssignales einer Strecke gegeben ist und das zugehörige Ausgangssignal berechnet werden soll. Prinzipiell ist eine Lösung dieses Problems mit Hilfe der Differentialgleichung der Strecke möglich, erfordert aber einen außerordentlich hohen Aufwand. Hier bietet sich der Einsatz der Laplace-Transformation (Anhang A.1) an. Zunächst wird das Eingangssignal nach Gleichung (A.2) oder besser mit Hilfe der Korrespondenztabelle im Anhang A.2 in den Frequenz- oder Bildbereich transformiert. Dann werden im Frequenzbereich mit Hilfe der im Anhang A.1 genannten Sätze die Ausgangssignale der einzelnen Übertragungsglieder in der Reihenfolge berechnet, wie sie vom Eingangssignal durchlaufen werden. Das Ausgangssignal des letzten Übertragungsgliedes ist das Ausgangssignal der Strecke, das schließlich in den Zeitbereich zurücktransformiert wird. Auch dafür steht wieder die Korrespondenztabelle im Anhang A.2 zur Verfügung, so dass sich das Problem auf die Berechnung des Ausgangssignales im Frequenzbereich reduziert.

Diese Berechnung ist aber bei linearen Übertragungsgliedern sehr einfach. Die Integration eines Signales reduziert sich im Frequenzbereich wegen des Integrationssatzes (Gleichung (A.8)) auf eine Multiplikation mit $\frac{1}{Ts}$.

Für die einfache Differentiation ergibt sich bei verschwindendem Eingangswert ($f(0) = 0$) gemäß Gleichung (A.10) die einfache Multiplikation mit Ts .

Die Summation und Multiplikation mit einem konstanten Faktor bleiben wegen des Additionssatzes (Gleichung (A.7)) erhalten, und ein Laufzeitglied

wird nach dem Verschiebungssatz (Gleichung (A.11)) durch den Faktor $e^{-T_L s}$ berücksichtigt.

In allen Fällen wird ein transformiertes Eingangssignal $x(s)$ mit einer von s abhängigen Funktion $G(s)$ multipliziert, um das Ausgangssignal $y(s)$ zu erhalten. $G(s)$ wird als Übertragungsfunktion bezeichnet (Abb. 3.1):

$$y(s) = G(s)x(s) \quad (3.1)$$

Von einer aufwändigen Lösung von Differentialgleichungen im Zeitbereich hat sich das Problem damit auf das Aufstellen einer Übertragungsfunktion und die Multiplikation mit dem Eingangssignal im Frequenzbereich reduziert.

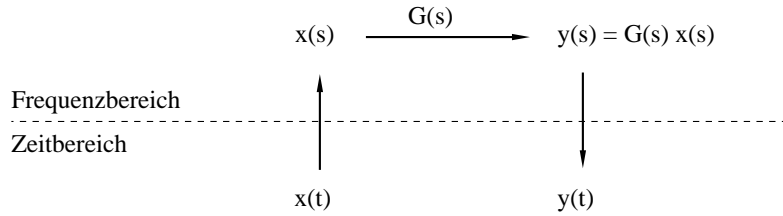


Abb. 3.1. Anwendung der Laplace-Transformation

Für den Integrator lautet die Übertragungsfunktion

$$G(s) = \frac{1}{Ts} \quad (3.2)$$

für das Laufzeitglied

$$G(s) = e^{-T_L s} \quad (3.3)$$

und für das Proportionalglied

$$G(s) = k \quad (3.4)$$

Auch für das PT_1 - und PT_2 -Glied lassen sich Übertragungsfunktionen angeben: Aus Gleichung (2.18) wird durch Anwendung des Differentiationssatzes bei verschwindendem Anfangswert von $y(t)$

$$Tsy(s) + y(s) = Vx(s) \quad (3.5)$$

und damit für die Übertragungsfunktion des PT_1 -Gliedes

$$G(s) = \frac{y(s)}{x(s)} = \frac{V}{Ts + 1} \quad (3.6)$$

Alternativ lässt sich diese Übertragungsfunktion auch anhand des Blockschaltbildes 2.10 berechnen: Der Integrator entspricht einer Multiplikation mit $\frac{1}{s}$, und die Multiplikationen mit den konstanten Faktoren V und $\frac{1}{T}$ bleiben erhalten. Ausgehend vom Eingangssignal $x(s)$ kann man dann von links

nach rechts im Blockschaltbild jeweils die Faktoren, Subtraktionen oder Additionen in der Gleichung hinzufügen, bis sich das Ausgangssignal auf der rechten Seite der Gleichung ergibt:

$$(x(s)V - y(s))\frac{1}{T}\frac{1}{s} = y(s) \quad (3.7)$$

Nach Auflösen dieser Gleichung erhält man wieder die Übertragungsfunktion des PT_1 -Gliedes.

Analog ergibt sich für das PT_2 -Glied aus Gleichung (2.28)

$$\frac{1}{\omega_0^2}s^2y(s) + \frac{2D}{\omega_0}sy(s) + y(s) = Vx(s) \quad (3.8)$$

und die Übertragungsfunktion

$$G(s) = \frac{y(s)}{x(s)} = \frac{V}{\frac{1}{\omega_0^2}s^2 + \frac{2D}{\omega_0}s + 1} \quad (3.9)$$

Auch hier kann alternativ die Berechnung gemäß Blockschaltbild 2.12 erfolgen, wobei hier zunächst eine kurze Vorbetrachtung erforderlich ist. Und zwar soll für die Kalkulation die erste Ableitung von y von vornherein mit sy und die zweite Ableitung mit s^2y bezeichnet werden. Dann erhält man, wenn man, ausgehend vom Eingangssignal $x(s)$, über den Faktor V , die Subtraktionsstelle und den Faktor ω_0^2 das Eingangssignal des ersten Integrators berechnet, gerade

$$(x(s)V - (\frac{2D}{\omega_0}sy(s) + y(s)))\omega_0^2 = s^2y(s) \quad (3.10)$$

und nach Ausmultiplikation und Umstellen wieder die Übertragungsfunktion (3.9).

Die Koeffizienten der Differentialgleichungen finden sich direkt in der Übertragungsfunktion wieder. Der Nenner der Übertragungsfunktion entspricht in beiden Fällen gerade dem charakteristischen Polynom der homogenen Differentialgleichung. Dies ist ein bemerkenswerter Zusammenhang, der im Folgenden noch benötigt wird.

Besteht ein Blockschaltbild nur aus Integratoren, Summierern und Proportionalgliedern, so entsteht durch Zusammenfassen der einzelnen Terme immer eine rein rationale Übertragungsfunktion

$$G(s) = \frac{y(s)}{x(s)} = \frac{b_ms^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{a_ns^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} \quad m \leq n \quad (3.11)$$

bei der der Grad des Nenners grundsätzlich größer oder gleich dem Grad des Zählers ist. Es sei aber nochmals darauf hingewiesen, dass eine solche Übertragungsfunktion nur dann entsteht, wenn die Anfangswerte der einzelnen Signale und gegebenenfalls ihrer Ableitungen verschwinden. Andernfalls

würden durch die Anwendung des Differentiationssatzes zusätzliche Terme entstehen. Im Folgenden soll diese Tatsache ohne weitere Erwähnung vorausgesetzt werden.

Weiterhin muss man sich darüber im Klaren sein, dass nur für lineare Übertragungsglieder Übertragungsfunktionen angegeben werden können. Für nichtlineare Übertragungsglieder ist dies nicht möglich. Es ist sogar Vorsicht geboten, denn eine Multiplikation oder Division im Zeitbereich entspricht nicht einer Multiplikation oder Division im Frequenzbereich. Auch eine Kennlinie darf nicht direkt vom Zeit- in den Frequenzbereich übertragen werden. Nichtlineare Übertragungsglieder müssen deshalb nach wie vor im Zeitbereich behandelt werden.

3.2 Interpretation der Übertragungsfunktion

Im Folgenden sollen nun einige Betrachtungen angestellt werden, die die Interpretation des Begriffes *Übertragungsfunktion* erleichtern. Dazu wird zunächst die *Impulsfunktion* bzw. der *Dirac-Impuls* $\delta(t)$ (Abb. 3.2) eingeführt, der definiert ist durch:

$$\delta(t) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (s(t) - s(t - \varepsilon)) \quad (3.12)$$

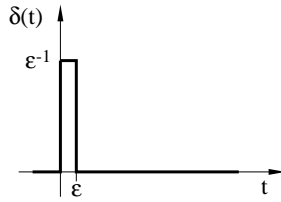


Abb. 3.2. Impulsfunktion

$s(t)$ sei dabei die Sprungfunktion. Man kann daher die Impulsfunktion auch als Ableitung der Sprungfunktion auffassen. Die Fläche unter einer Impulsfunktion ist gerade Eins:

$$\begin{aligned} \int_0^{\infty} \delta(t) dt &= \int_0^{\infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (s(t) - s(t - \varepsilon)) dt \\ &= \lim_{\varepsilon \rightarrow 0} \int_0^{\infty} \frac{1}{\varepsilon} (s(t) - s(t - \varepsilon)) dt \\ &= \lim_{\varepsilon \rightarrow 0} \int_0^{\varepsilon} \frac{1}{\varepsilon} dt = \lim_{\varepsilon \rightarrow 0} 1 = 1 \end{aligned} \quad (3.13)$$

Für die Laplace-Transformierte der Impulsfunktion ergibt sich

$$\begin{aligned}
 \mathcal{L}\{\delta(t)\} &= \int_0^{\infty} e^{-st} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (s(t) - s(t - \varepsilon)) dt \\
 &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^{\infty} e^{-st} (s(t) - s(t - \varepsilon)) dt \\
 &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^{\varepsilon} e^{-st} dt = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \frac{1}{s} (1 - e^{-s\varepsilon}) \quad (3.14)
 \end{aligned}$$

Zähler und Nenner des letzten Terms konvergieren jeweils gegen Null. Damit ist einer der Grenzwertsätze von de l'Hospital anwendbar. Demnach gilt für $f(x_0) = 0$, $g(x_0) = 0$ und $g'(x) \neq 0$ in einer Umgebung von x_0 :

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} \quad (3.15)$$

und im vorliegenden Fall

$$\mathcal{L}\{\delta(t)\} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \frac{1}{s} (1 - e^{-s\varepsilon}) = \lim_{\varepsilon \rightarrow 0} \frac{se^{-s\varepsilon}}{s} = 1 \quad (3.16)$$

Damit kann man schreiben:

$$G(s) = G(s) 1 = G(s) \mathcal{L}\{\delta(t)\} \quad (3.17)$$

Ein Vergleich mit Gleichung (3.1) zeigt, dass die Übertragungsfunktion $G(s)$ auch als Laplace-Transformierte des Ausgangssignales bei Anregung der Strecke durch einen Impuls interpretiert werden kann. $G(s)$ ist damit die Laplace-Transformierte einer gedachten *Impulsantwort* $g(t)$.

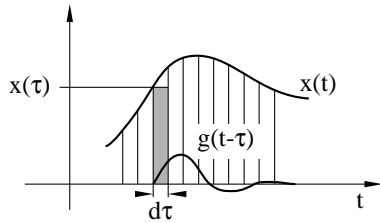


Abb. 3.3. Anschauliche Erklärung des Faltungsintegrals

Diese Interpretation führt wiederum auf eine sehr anschauliche Erklärung mit Hilfe des Faltungssatzes. Um das Ausgangssignal im Frequenzbereich zu berechnen, hat man nach (3.1) die Laplace-Transformierte des Eingangssignales mit der Übertragungsfunktion zu multiplizieren. Mit dem Faltungssatz entspricht diese Operation einer Faltung im Zeitbereich:

$$y(s) = G(s)x(s) \longleftrightarrow y(t) = \int_0^t g(t-\tau)x(\tau)d\tau \quad (3.18)$$

Die Funktion $g(t)$ ist dabei gerade die Impulsantwort. Abb. 3.3 veranschaulicht diese Formel. Man kann sich das Eingangssignal näherungsweise als eine Folge von Impulsen mit der Höhe $x(\tau)$ und der Breite $d\tau$ vorstellen, von denen jeder am Ausgang eine Impulsantwort hervorruft. Wegen der Linearität der Strecke beeinflussen sich diese Antworten nicht gegenseitig, so dass der Momentanwert am Ausgang eine Überlagerung aller Impulsantworten darstellt. Der zum Zeitpunkt τ am Eingang anliegende Impuls mit der Fläche $x(\tau)d\tau$ verursacht die Impulsantwort $g(t-\tau)x(\tau)d\tau$. Dies ist auch gerade der Beitrag dieser Impulsantwort zum Momentanwert $y(t)$.

Nachdem die Bedeutung der Impulsantwort damit geklärt wurde, sollen noch einige für die Praxis brauchbare Zusammenhänge angegeben werden.

Eingangs dieses Kapitels wurde schon gezeigt, dass die Impulsfunktion als Ableitung der Sprungfunktion aufgefasst werden kann. Die Sprungfunktion wiederum ist offenbar die Ableitung der sogenannten *Rampenfunktion* oder *Anstiegsfunktion* $r(t)$, die in Abbildung 3.4 gezeigt ist.

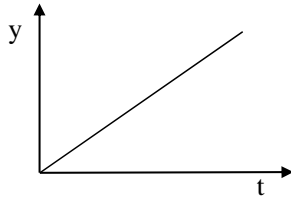


Abb. 3.4. Rampen- oder Anstiegsfunktion

Dieser Zusammenhang spiegelt sich auch in den Laplace-Transformierten dieser drei Funktionen. Sie lauten gemäß Korrespondenztabelle im Anhang A.2

$$\begin{aligned} \mathcal{L}\{\delta(t)\} &= 1 \\ \mathcal{L}\{s(t)\} &= \frac{1}{s} \\ \mathcal{L}\{r(t)\} &= \frac{1}{s^2} \end{aligned} \quad (3.19)$$

Der Faktor $\frac{1}{s}$ entspricht einer Integration. Wenn daher 1 die Laplacetransformierte der Impulsfunktion ist, dann muss die Sprungfunktion $\frac{1}{s}$ das Integral der Impulsfunktion sein, und die Anstiegsfunktion $\frac{1}{s^2}$ wiederum das Integral der Sprungfunktion. Umgekehrt ist die Sprungfunktion die Ableitung der Anstiegsfunktion, und die Impulsfunktion die Ableitung der Sprungfunktion, was ja aufgrund des Verlaufs der Funktionen im Zeitbereich schon offensichtlich war.

Für eine gegebene Strecke mit der Übertragungsfunktion $G(s)$ ergeben sich die Systemantworten für die drei Signale aus einer Multiplikation der Laplace-Transformierten des jeweiligen Eingangssignales mit der Übertragungsfunktion:

$$\begin{aligned} G(s) &: \text{Impulsantwort} \\ \frac{1}{s}G(s) &: \text{Sprungantwort} \\ \frac{1}{s^2}G(s) &: \text{Anstiegsantwort} \end{aligned} \quad (3.20)$$

Auch hier lässt sich der Faktor $\frac{1}{s}$ als Integration auffassen. Demnach ist die Sprungantwort das Integral der Impulsantwort, und die Anstiegsantwort ist das Integral der Sprungantwort. Umgekehrt muss die Sprungantwort die Ableitung der Anstiegsantwort sein, und die Impulsantwort, also die Übertragungsfunktion, die Ableitung der Sprungantwort.

3.3 Berechnung der Sprungantwort

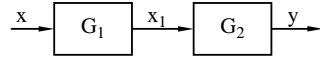
Mit Hilfe der Übertragungsfunktionen gelangt man auch zu Regeln, wie man Blockschaltbilder linearer Strecken umstrukturieren kann, ohne das Übertragungsverhalten des Modells zu verändern. In Abb. 3.5 sind jeweils zwei zueinander äquivalente Blockschaltbilder nebeneinander aufgeführt. Die Äquivalenz lässt sich sofort durch Berechnung der jeweiligen Übertragungsfunktionen nachweisen, wobei für die aufgestellten Gleichungen natürlich alle Rechenregeln wie z.B. Kommutativität oder Assoziativität Gültigkeit haben. Der erste Fall zeigt beispielsweise, dass lineare Übertragungsglieder in ihrer Reihenfolge beliebig vertauscht werden dürfen. Zu beachten ist allerdings, dass durch solche Umformungen in der Regel die Zuordnung zu internen, realen physikalischen Größen verloren geht.

Anhand eines einfachen Beispiels soll nun demonstriert werden, wie sich bei rationalen Übertragungsfunktionen Ausgangssignalverläufe bestimmen lassen. Mit Hilfe der Korrespondenztabelle im Anhang A.2 ist dies kein besonderes Problem. Die Laplace-Transformierte der Sprungfunktion ist beispielsweise $\frac{1}{s}$. Mit der allgemeinen Formel $y(s) = G(s)x(s)$ ergibt sich daher für die Sprungantwort eines linearen Übertragungsgliedes:

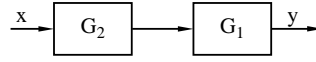
$$y(s) = G(s) \frac{1}{s} \quad (3.21)$$

Setzt man in diese Gleichung die Übertragungsfunktion des PT_1 -Gliedes ein, so erhält man als Sprungantwort

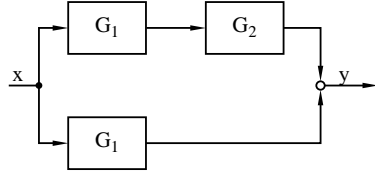
$$y(s) = \frac{V}{Ts + 1} \frac{1}{s} \quad (3.22)$$



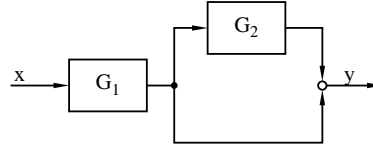
$$y(s) = G_2(s) x_1(s) = G_2(s) G_1(s) x(s)$$



$$y(s) = G_1(s) G_2(s) x(s)$$



$$y(s) = (G_2(s) G_1(s) + G_1(s)) x(s)$$



$$y(s) = (G_2(s) + 1) G_1(s) x(s)$$

Abb. 3.5. Äquivalente Umformungen bei linearen Übertragungsgliedern

Dieser Ausdruck ist nun nur noch in den Zeitbereich zurückzutransformieren. Er findet sich aber nicht in der Korrespondenztabelle. Führt man jedoch eine Partialbruchzerlegung durch, so ergibt sich

$$y(s) = V \left[\frac{1}{s} - \frac{1}{s + \frac{1}{T}} \right] \quad (3.23)$$

Wegen des Additionssatzes dürfen beide Summanden in der Klammer einzeln in den Zeitbereich zurücktransformiert werden. $\frac{1}{s}$ ist die Laplace-Transformierte der Sprungfunktion $s(t) = 1$ und $\frac{1}{s + \frac{1}{T}}$ die Transformierte der Funktion $s(t)e^{-t/T} = e^{-t/T}$, wie sich aus der Korrespondenztabelle ablesen lässt. Man erhält damit den schon bekannten Funktionsverlauf

$$y(t) = V \left[1 - e^{-\frac{t}{T}} \right] \quad \text{für } t \geq 0 \quad (3.24)$$

Für die Sprungantwort eines Übertragungsgliedes mit einer rationalen Übertragungsfunktion lässt sich ein allgemeiner Ausdruck angeben. Zunächst sind Zähler- und Nennerpolynom aus (3.11) in Linearfaktoren zu zerlegen:

$$G(s) = \frac{b_m}{a_n} \frac{\prod_{\mu=1}^m (s - n_{\mu})}{\prod_{\nu=1}^n (s - p_{\nu})} \quad \text{mit } m \leq n \quad (3.25)$$

Da $G(s)$ nur reelle Koeffizienten hat, sind alle Pole p_{ν} bzw. Nullstellen n_{μ} entweder reell oder paarweise komplex konjugiert.

Für die Sprungantwort $y(s)$ gilt wegen $y(s) = \frac{1}{s} G(s)$:

$$y(s) = \frac{b_m}{a_n} \frac{\prod_{\mu=1}^m (s - n_{\mu})}{\prod_{\nu=1}^{n+1} (s - p_{\nu})} \quad \text{mit } p_{n+1} = 0 \quad (3.26)$$

Für die folgende Betrachtung sei angenommen, dass Zähler- und Nennerpolynom von $y(s)$ teilerfremd sind. Ansonsten sind sie vorher entsprechend gegeneinander zu kürzen. Die Ordnung des Zählers von $y(s)$ ist wegen der zusätzlichen Polstelle $p_{n+1} = 0$ auf jeden Fall kleiner als die Ordnung des Nenners.

Damit lässt sich eine Partialbruchzerlegung durchführen. Um dies formal darstellen zu können, müssen die Pole neu sortiert werden, was zunächst anhand eines einfachen Beispiels erklärt werden soll. Gegeben seien 7 Pole p_1 bis p_7 mit $p_1 = p_2 = p_3$ und $p_4 = p_5$. Diese werden nun umbenannt in s_1 bis s_4 , wobei der Pol s_1 dreifach und der Pol s_2 doppelt vorkommt:

$$\begin{array}{ccccccc} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ || & || & || & || & || & || & || \\ s_1 & s_1 & s_1 & s_2 & s_2 & s_3 & s_4 \end{array} \quad (3.27)$$

Für die formale Darstellung dieser Polzuweisung sei angenommen, dass unter den $n+1$ Polen p_ν insgesamt i verschiedene Pole existieren. Diese Pole werden jetzt mit s_λ bezeichnet, wobei λ Werte von 1 bis i annehmen kann. Die Vielfachheit des Pols s_λ beträgt n_λ . Offensichtlich existieren zu jedem s_λ ein oder mehrere Pole p_ν mit $s_\lambda = p_\nu$, während zu jedem p_ν genau ein s_λ existiert mit $s_\lambda = p_\nu$. Darüber hinaus muss gelten:

$$n+1 = \sum_{\lambda=1}^i n_\lambda \quad (3.28)$$

Für den Pol s_λ mit seiner Vielfachheit n_λ gilt die Partialbruchzerlegung

$$\frac{r_{\lambda 1}}{(s-s_\lambda)^1} + \frac{r_{\lambda 2}}{(s-s_\lambda)^2} + \frac{r_{\lambda 3}}{(s-s_\lambda)^3} + \dots + \frac{r_{\lambda n_\lambda}}{(s-s_\lambda)^{n_\lambda}} \quad (3.29)$$

Damit ergibt sich für die gesamte Sprungantwort:

$$y(s) = \sum_{\lambda=1}^i \sum_{\nu=1}^{n_\lambda} \frac{r_{\lambda \nu}}{(s-s_\lambda)^\nu} \quad \text{mit } n+1 = \sum_{\lambda=1}^i n_\lambda \quad (3.30)$$

Mit Hilfe der Korrespondenztabelle im Anhang A.2 lässt sich nun die Rücktransformation in den Zeitbereich durchführen. Man erhält

$$\begin{array}{l} \frac{1}{s^{n+1}} \longleftrightarrow \frac{t^n}{n!} \\ \frac{1}{(s-a)^{n+1}} \longleftrightarrow \frac{t^n}{n!} e^{at} \\ \frac{r_{\lambda \nu}}{(s-s_\lambda)^\nu} \longleftrightarrow \frac{r_{\lambda \nu}}{(\nu-1)!} t^{\nu-1} e^{s_\lambda t} \end{array} \quad (3.31)$$

Jeder Pol s_λ liefert damit entsprechend seiner Vielfachheit n_λ zur Sprungantwort im Zeitbereich den Beitrag

$$\sum_{\nu=1}^{n_\lambda} \frac{r_{\lambda\nu}}{(\nu-1)!} t^{\nu-1} e^{s_\lambda t} = h_\lambda(t) e^{s_\lambda t} \quad (3.32)$$

also das Produkt aus einem Polynom $h_\lambda(t)$ vom Grad $n_\lambda - 1$ und einer Exponentialfunktion. Insgesamt folgt für die Sprungantwort

$$y(t) = \sum_{\lambda=1}^i h_\lambda(t) e^{s_\lambda t} \quad (3.33)$$

Für rein reelle, negative Polstellen verschwindet der Beitrag $h_\lambda(t) e^{s_\lambda t}$ mit wachsendem t , denn die Exponentialfunktion konvergiert schneller gegen Null als jede endliche Potenz von t anwächst. Umgekehrt wächst dagegen für rein reelle, positive Polstellen der Ausdruck über alle Maßen.

Komplexe Polstellen können aufgrund der rein reellen Koeffizienten der Übertragungsfunktion nur als konjugiert komplexe Polpaare auftreten und kennzeichnen damit einen schwingungsfähigen Anteil des Systems. Wie schon beim PT_2 -Glied gezeigt, hängt es aber auch hier vom Realteil der Polstellen ab, ob die zugehörige Schwingung auf- oder abklingend ist.

Für Pole $s_\lambda = 0$ nimmt die Exponentialfunktion den Wert Eins an und kann deshalb entfallen. Übrig bleibt in dem Fall nur das Polynom. Hat die Strecke selbst keinen Pol bei Null, so enthält $y(s)$ wegen der Sprungfunktion $\frac{1}{s}$ nur einen einfachen Pol an dieser Stelle. Das zugehörige Polynom $h_\lambda(t)$ ist demnach vom Grad Null, d.h. konstant. Der zugehörige Ausdruck $h_\lambda(t) e^{s_\lambda t}$ ist damit ebenfalls konstant. Falls sonst nur Pole mit negativem Realteil vorliegen, deren Beitrag mit wachsendem t verschwindet, bildet dieser konstante Wert den Endwert der Sprungantwort.

Falls dagegen die Strecke selbst mindestens einen Pol bei $s_\lambda = 0$ enthält, so steigt die Vielfachheit dieses Poles mindestens auf zwei, und der Grad von $h_\lambda(t)$ beträgt mindestens eins. In dem Fall wächst der zugehörige Ausdruck mit t über alle Maßen. Dies ist sofort einsichtig, denn ein Pol der Strecke bei $s = 0$ steht für einen Integrator in der Strecke, der nicht wie beim PT_1 - oder PT_2 -Glied zurückgekoppelt, sondern zum Rest der Strecke quasi in Reihe geschaltet ist. Da es bei Strecken, die aus linearen Übertragungsgliedern bestehen, nicht auf die Reihenfolge der einzelnen Übertragungsblöcke ankommt, kann man diesen Integrator am Anfang der Strecke annehmen. Dort erhält er aber die am Eingang anliegende Sprungfunktion als Eingangssignal, die er rampenförmig immer weiter aufintegriert. Diese Rampenfunktion kann dann vom Rest der Strecke nicht mehr kompensiert werden, so dass die Ausgangsgröße der Strecke mit t immer weiter ansteigt.

Ein Polpaar mit $\text{Re}(s_\lambda) = 0$ und $\text{Im}(s_\lambda) \neq 0$ erzeugt gemäß Gleichung (2.47) eine Schwingung mit konstanter Amplitude. Wenn es in größerer Vielfachheit als Eins auftritt, wird der Grad des Polynoms $h_\lambda(t)$ größer als Null, und das Produkt $h_\lambda(t) e^{s_\lambda t}$ wächst dann auch hier über alle Maßen.

Offensichtlich wird das Einschwingverhalten des Systems vollständig durch die Pole der Übertragungsfunktion beschrieben. Zusammenfassend lässt sich

sagen, dass die Sprungantwort immer gegen einen endlichen Wert konvergiert, wenn alle Pole der Übertragungsfunktion einen negativen Realteil aufweisen.

Interessant ist, dass sich Anfangs- und Endwert der Sprungantwort auch mit Hilfe der Grenzwertsätze der Laplace-Transformation berechnen lassen, sofern die Grenzwerte existieren. Für den Endwert der Sprungantwort gilt mit dem Grenzwertsatz der Laplace-Transformation (A.13), der allgemeinen Übertragungsfunktion (3.11) und der Formel für die Sprungantwort eines linearen Übertragungsgliedes (3.21):

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sy(s) = \lim_{s \rightarrow 0} s \frac{1}{s} G(s) = \lim_{s \rightarrow 0} G(s) = \frac{b_0}{a_0} \quad (3.34)$$

und analog dazu für den Anfangswert

$$\lim_{t \rightarrow 0} y(t) = \lim_{s \rightarrow \infty} G(s) = \begin{cases} \frac{b_m}{a_n} & \text{für } m = n \\ 0 & \text{für } m < n \end{cases} \quad (3.35)$$

Anfangs- und Endwert der Sprungantwort lassen sich demnach auch sofort aus den Koeffizienten der Übertragungsfunktion ablesen.

Die späteren Kapitel werden zeigen, dass die Berechnung von Signalverläufen für die Auslegung von Reglern und die Analyse von Regelkreisen gar nicht erforderlich ist, denn eine Analyse der Übertragungsfunktion liefert bereits alle notwendigen Informationen über Stabilität und Einschwingverhalten der Strecke.

3.4 Vereinfachung einer Übertragungsfunktion

Zum Abschluss dieses Kapitels soll noch die Möglichkeit der Vereinfachung einer Übertragungsfunktion diskutiert werden. Insbesondere wenn diese zur Auslegung eines Reglers herangezogen werden soll, kann die Approximation eines gegebenen Übertragungsverhaltens durch eine möglichst einfache Funktion sinnvoll sein. Gegeben sei die folgende Übertragungsfunktion, die aus einem rationalen Anteil und einem Laufzeitglied besteht. Der rationale Anteil besitze ausschließlich Pole mit negativem Realteil, d.h. die Sprungantwort habe einen endlichen Endwert:

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} e^{-T_L s} \quad \text{mit } m < n \quad (3.36)$$

Eine Zerlegung in Linearfaktoren liefert

$$G(s) = \frac{b_0}{a_0} \frac{\prod_{\mu=1}^m (T_{z\mu} s + 1)}{\prod_{\nu=1}^n (T_{n\nu} s + 1)} e^{-T_L s} \quad (3.37)$$

Nun können Zählerfaktoren und Exponentialfunktion in Potenzreihen entwickelt werden:

$$\begin{aligned} (T_{z\mu}s + 1) &= \frac{1}{1 + \sum_{i=1}^{\infty} (-T_{z\mu}s)^i} \quad \text{für } |T_{z\mu}s| < 1 \\ e^{T_L s} &= 1 + \sum_{\lambda=1}^{\infty} \frac{(T_L s)^\lambda}{\lambda!} \end{aligned} \quad (3.38)$$

Damit ergibt sich

$$G(s) = \frac{\frac{b_0}{a_0}}{\left[\prod_{\mu=1}^m (1 + \sum_{i=1}^{\infty} (-T_{z\mu}s)^i) \right] \left[\prod_{\nu=1}^n (T_{n\nu}s + 1) \right] \left[1 + \sum_{\lambda=1}^{\infty} \frac{(T_L s)^\lambda}{\lambda!} \right]} \quad (3.39)$$

Ausmultiplizieren ergibt eine neue Reihe im Nenner

$$G(s) = \frac{b_0}{a_0} \frac{1}{1 + \left(- \sum_{\mu=1}^m T_{z\mu} + \sum_{\nu=1}^n T_{n\nu} + T_L \right) s + \dots} \quad (3.40)$$

Bricht man diese Reihenentwicklung nach dem ersten Glied ab, so ergibt sich ein PT_1 -Glied mit der *Ersatzzeitkonstanten* T_e :

$$G(s) \approx \frac{b_0}{a_0} \frac{1}{1 + T_e s} \quad (3.41)$$

mit

$$T_e = \sum_{\nu=1}^n T_{n\nu} - \sum_{\mu=1}^m T_{z\mu} + T_L \quad (3.42)$$

Das Ausmultiplizieren der Linearfaktoren der Übertragungsfunktion aus Gleichung (3.37) und der anschließende Koeffizientenvergleich mit Gleichung (3.36) liefern wiederum die Zusammenhänge

$$a_0 \sum_{\nu=1}^n T_{n\nu} = a_1 \quad \text{und} \quad b_0 \sum_{\mu=1}^m T_{z\mu} = b_1 \quad (3.43)$$

so dass sich für die Ersatzzeitkonstante der sehr einfache Zusammenhang

$$T_e = \frac{a_1}{a_0} - \frac{b_1}{b_0} + T_L \quad (3.44)$$

ergibt. Zu beachten ist allerdings die Einschränkung $|T_{z\mu}s| < 1$ aus der Reihenentwicklung. Je nach Betrag von $T_{z\mu}$ gilt die Näherung nur für kleine Werte von s . Sofern der Grad des Zählerpolynoms jedoch gleich Eins ist, entfällt

diese Einschränkung. Glücklicherweise kommt auch dieser Fall in der Praxis noch häufig genug vor. Wenn beispielsweise die Reihenschaltung von mehreren PT_1 - und PT_2 -Gliedern zu einem einzigen PT_1 -Glieder zusammengefasst werden soll, so muss man nach Gleichung (3.42) nur die Verzögerungszeiten der einzelnen Glieder aufaddieren, und man erhält die Ersatzzeitkonstante T_e .

Anhand der Sprungantwort kann man dieses approximierende PT_1 -Glieder recht gut mit der Originalstrecke vergleichen (Abb. 3.6). Es lässt sich zeigen, dass die sogenannte *Regelfläche*

$$\int_0^{\infty} \left| \frac{b_0}{a_0} - y(t) \right| dt \quad \text{mit } y(t) = \text{Sprungantwort} \quad (3.45)$$

also die Fläche zwischen der Sprungantwort und ihrem Endwert, in beiden Fällen gleich groß ist. Aufgrund dieser Tatsache besteht auch die Möglichkeit, eine Ersatzfunktion auf graphischem Wege zu konstruieren, wenn von der Originalstrecke keine Übertragungsfunktion, sondern lediglich eine gemessene Sprungantwort vorliegt.

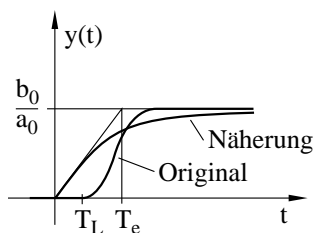


Abb. 3.6. Zur Ersatzzeitkonstanten

Je nach Anforderungen an die Genauigkeit der Approximation kann die Reihe in Gleichung (3.40) auch nach einem späteren Glied abgebrochen werden. Zu bedenken ist aber, dass es sich immer nur um eine Näherung handelt. Die dynamischen Vorgänge in einer Strecke n -ter Ordnung können durch eine Übertragungsfunktion niedrigerer Ordnung nicht vollständig beschrieben werden. Insbesondere im Hinblick auf das Stabilitätsverhalten kann dies schwerwiegende Konsequenzen haben. Während ein PT_1 -Glieder keine Probleme bereitet, kann sich die Strecke, die sich hinter dieser Näherung verbirgt, dicht an der Grenze zur Instabilität befinden. Und ein Regler, der mit einem PT_1 -Glieder als Strecke hervorragend funktionieren würde, kann zusammen mit der tatsächlichen Strecke ein instabiles System bilden. Die vorgestellte Näherung ist also prinzipiell mit Vorsicht zu genießen. Grundsätzlich kann eine Strecke niedriger Ordnung eine Strecke höherer Ordnung nur im Bereich tiefer Signalfrequenzen gut approximieren. Je höher die Frequenzen werden, desto ungenauer ist die Näherung. Liegt allerdings der Nutzfrequenzbereich eines Regelungssystems eher bei niedrigen Frequenzen, so ist diese Näherung

durchaus angebracht, um zu einer übersichtlicheren Übertragungsfunktion zu gelangen.

4. Frequenzgang

4.1 Einführung des Frequenzganges

Der *Frequenzgang* lässt sich am einfachsten definieren als Übertragungsfunktion eines linearen Übertragungsgliedes für rein imaginäre Werte von s . Die komplexe Variable s in der Übertragungsfunktion wird demnach lediglich durch eine rein imaginäre Variable $j\omega$ ersetzt: $G(j\omega) = G(s)|_{s=j\omega}$. Damit ist der Frequenzgang eine komplexe Funktion des Parameters ω . Wegen der Beschränkung auf rein imaginäre Werte von s stellt der Frequenzgang nur einen Ausschnitt der Übertragungsfunktion dar und enthält somit weniger Information als die Übertragungsfunktion. Am Frequenzgang lässt sich aber direkt ablesen, wie sich das System in Abhängigkeit von der Frequenz des Eingangssignales verhält:

Satz 4.1 *Besitzt ein lineares Übertragungsglied den Frequenzgang $G(j\omega)$, so antwortet es auf das Eingangssignal $x(t) = a \sin \omega t$ nach Abklingen der Einschwingvorgänge mit dem Ausgangssignal*

$$y(t) = a |G(j\omega)| \sin(\omega t + \varphi(G(j\omega))) \quad (4.1)$$

sofern gilt:

$$\int_0^{\infty} |g(t)| dt < \infty \quad (4.2)$$

$\varphi(G(j\omega))$ ist die Phase der komplexen Größe $G(j\omega)$ und $g(t)$ die Impulsantwort der Strecke. Konvergiert das Integral aus (4.2) nicht, so ist die rechte Seite von (4.1) um einen Term $r(t)$ zu ergänzen, der auch für $t \rightarrow \infty$ nicht verschwindet.

Der Beweis zu diesem Satz findet sich in [11].

Im Gegensatz zu den Koeffizienten einer Übertragungsfunktion sind Betrag und Phase des Frequenzganges direkt messbar: Die Strecke wird durch ein sinusförmiges Eingangssignal mit einer bestimmten Frequenz und Amplitude angeregt. Nach Abklingen der Einschwingvorgänge (wie durch Gleichung (4.2) vorgeschrieben) wird sich am Ausgang ein ebenfalls sinusförmiges Ausgangssignal einstellen, das sich gegenüber dem Eingangssignal aber

in Amplitude und Phasenlage unterscheidet. Beide Größen sind messbar, und aus ihnen lassen sich nach Gleichung (4.1) auch sofort Betrag und Phase des Frequenzganges $G(j\omega)$ berechnen. So lässt sich für verschiedene Frequenzen eine Wertetabelle bestimmen, die den prinzipiellen Verlauf des Frequenzganges skizziert. Die Messung mit negativen Werten von ω , d.h. mit negativen Frequenzen, ist natürlich nicht möglich, aber auch nicht notwendig. Denn für rationale Übertragungsfunktionen mit reellen Koeffizienten und auch für Laufzeitglieder ist $G(j\omega)$ konjugiert komplex zu $G(-j\omega)$. Da die Funktion $G(j\omega)$ für $\omega \geq 0$ also bereits die gesamte Information enthält, erübrigt sich demnach eine Betrachtung negativer Werte von ω .

Grundsätzlich ist mit Hilfe von Satz 4.1 auch die Ermittlung des Ausgangssignales bei komplexeren Eingangssignalen, die aus mehreren unterschiedlichen Frequenzen zusammengesetzt sind, möglich. Denn aufgrund der vorausgesetzten Linearität des Übertragungsgliedes beeinflussen sich die Wirkungen, die durch die einzelnen Frequenzanteile hervorgerufen werden, nicht gegenseitig, und daher kann das Ausgangssignal als additive Überlagerung der Reaktionen auf die einzelnen Frequenzanteile des Eingangssignales berechnet werden.

Die Berechnung einzelner Signalverläufe ist aber in der Regelungstechnik weder üblich noch erforderlich. Man ist vielmehr an generellen Aussagen über das Stabilitätsverhalten des Systems interessiert, und genau dazu eignet sich der Frequenzgang als Hilfsmittel hervorragend, wie das nachfolgende Kapitel 5 über die Stabilität linearer Systeme zeigen wird.

Besonders einfach wird die Stabilitätsanalyse, wenn der Frequenzgang in graphischer Darstellung vorliegt. Sofern Messwerte des Frequenzganges ermittelt wurden, ist eine derartige Darstellung trivial. Deutlich häufiger ist jedoch der Fall, dass eine Übertragungsfunktion gegeben ist, aus der die graphische Darstellung des Frequenzganges zum Zweck einer Stabilitätsanalyse entwickelt werden muss.

Selbstverständlich kann diese Aufgabe heutzutage von einer regelungstechnischen Software quasi per Knopfdruck erledigt werden. Eine derartige Software steht jedoch nicht immer zur Verfügung, und sie ist auch gar nicht erforderlich. Denn die - für eine Stabilitätsanalyse völlig ausreichende - prinzipielle Darstellung des Frequenzganges lässt sich sehr schnell und einfach von Hand entwickeln. Die entsprechende Vorgehensweise wird im Folgenden erläutert.

4.2 Bode-Diagramm

4.2.1 Herleitung

Beim *Bode-Diagramm* [4] werden Betrag und Phase des Frequenzganges in zwei Diagrammen getrennt voneinander jeweils über der Frequenz aufgetragen.

Es soll nun gezeigt werden, wie die prinzipiellen Verläufe von Betrag und Phase direkt aus einer Übertragungsfunktion entwickelt werden können. Diese sei in der aus Gleichung (3.37) bereits bekannten Form

$$G(s) = \frac{b_0 \prod_{\mu=1}^m (T_{z\mu}s + 1)}{a_0 \prod_{\nu=1}^n (T_{n\nu}s + 1)} \quad (4.3)$$

gegeben. Ersetzen von s durch $j\omega$ liefert den Frequenzgang

$$G(j\omega) = \frac{b_0 \prod_{\mu=1}^m (jT_{z\mu}\omega + 1)}{a_0 \prod_{\nu=1}^n (jT_{n\nu}\omega + 1)} = V \left(\prod_{\mu=1}^m G_{z\mu}(j\omega) \right) \left(\prod_{\nu=1}^n G_{n\nu}(j\omega) \right) \quad (4.4)$$

mit

$$V = \frac{b_0}{a_0} \quad G_{z\mu}(j\omega) = jT_{z\mu}\omega + 1 \quad G_{n\nu}(j\omega) = \frac{1}{jT_{n\nu}\omega + 1} \quad (4.5)$$

Für die Beträge der Teil-Frequenzgänge $G_{z\mu}$ und $G_{n\nu}$ gilt

$$|G_{z\mu}(j\omega)| = \sqrt{T_{z\mu}^2\omega^2 + 1} \quad |G_{n\nu}(j\omega)| = \frac{1}{\sqrt{T_{n\nu}^2\omega^2 + 1}} \quad (4.6)$$

und für die Phasen

$$\varphi(G_{z\mu}(j\omega)) = \arctan\left(\frac{\operatorname{Im}(G_{z\mu})}{\operatorname{Re}(G_{z\mu})}\right) = \arctan(T_{z\mu}\omega) \quad (4.7)$$

sowie mit

$$G_{n\nu}(j\omega) = \frac{1}{jT_{n\nu}\omega + 1} = \frac{1 - jT_{n\nu}\omega}{1 + T_{n\nu}^2\omega^2} \quad (4.8)$$

für die Terme im Nenner

$$\varphi(G_{n\nu}(j\omega)) = \arctan\left(\frac{\operatorname{Im}(G_{n\nu})}{\operatorname{Re}(G_{n\nu})}\right) = \arctan(-T_{n\nu}\omega) = -\arctan(T_{n\nu}\omega) \quad (4.9)$$

Damit ergibt sich für den gesamten Frequenzgang

$$G(j\omega) = |G(j\omega)| e^{j\varphi(G(j\omega))} \quad (4.10)$$

mit

$$\begin{aligned} |G(j\omega)| &= |V| \prod_{\mu=1}^m |G_{z\mu}(j\omega)| \prod_{\nu=1}^n |G_{n\nu}(j\omega)| \\ \varphi(G(j\omega)) &= \sum_{\mu=1}^m \varphi(G_{z\mu}(j\omega)) + \sum_{\nu=1}^n \varphi(G_{n\nu}(j\omega)) \end{aligned} \quad (4.11)$$

Die Addition der Phasenwinkel der Teil-Frequenzgänge ergibt den Phasenwinkel des Gesamt-Frequenzganges. Diese Addition lässt sich auch auf graphischem Wege sehr einfach durchführen, wie in Kapitel 4.2.10 noch gezeigt wird.

Bei den Beträgen ergibt sich der Betrag des Gesamt-Frequenzganges durch Multiplikation der Beträge der Teil-Frequenzgänge. Da eine Multiplikation auf graphischem Wege jedoch nicht möglich ist, geht man hier zu den Logarithmen der Beträge über:

$$\log |G(j\omega)| = \log |V| + \sum_{\mu=1}^m \log |G_{z\mu}(j\omega)| + \sum_{\nu=1}^n \log |G_{n\nu}(j\omega)| \quad (4.12)$$

In der logarithmischen Darstellung können offenbar auch die Beträge der Teil-Frequenzgänge additiv zum Betrag des Gesamt-Frequenzganges überlagert werden.

Die Gleichungen (4.11) und (4.12) zeigen, dass sich sowohl der Phasen- als auch der Betragsverlauf eines Frequenzganges graphisch ermitteln lassen, indem man zunächst die Betrags- und Phasenverläufe der Teil-Frequenzgänge $G_{z\mu}$ und $G_{n\nu}$ zeichnet und diese dann additiv überlagert, wobei für den Betrag eine logarithmische Darstellung zu wählen ist.

Wie eine solche Überlagerung zu erfolgen hat, wird in Kapitel 4.2.10 anhand eines Beispiels gezeigt. Zunächst soll aber noch geklärt werden, wie Betrags- und Phasenverlauf der Teil-Frequenzgänge überhaupt aussehen. Dabei wird sich zeigen, dass die Darstellung besonders einfach wird, wenn man nicht nur den Betrag, sondern auch die Frequenz ω im logarithmischen Maßstab aufträgt. Die Phase bleibt jedoch im linearen Maßstab.

4.2.2 Negativ reelle Polstelle

Zunächst sei eine Nenner-Teilfunktion $G_{n\nu}$ mit negativ reeller Polstelle betrachtet, wobei nun auf die Indizes verzichtet werden kann:

$$G(s) = \frac{1}{Ts + 1} \quad \text{mit} \quad T > 0 \quad (4.13)$$

Diese Funktion, ergänzt um einen Verstärkungsfaktor, entspräche gerade der Übertragungsfunktion eines PT_1 -Gliedes. Unter Verwendung von (4.6) und (4.9) lautet der Frequenzgang

$$G(j\omega) = \frac{1}{\sqrt{T^2\omega^2 + 1}} e^{-j \arctan(T\omega)} \quad (4.14)$$

Um den Verlauf von Betrag und Phase über der Frequenz skizzieren zu können, werden die Asymptoten beider Funktionen für sehr kleine und sehr große Frequenzen benötigt, die sich unmittelbar aus (4.14) ablesen lassen:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	1	$\frac{1}{\omega T}$
$\log G(j\omega) $	0	$\log \frac{1}{\omega T} = -\log(\omega T)$ $= -\log \omega - \log T$ $\approx -\log \omega$
$\varphi(G(j\omega))$	0	$-\frac{\pi}{2}$

An der vorletzten Zeile dieser Tabelle erkennt man, dass die Hochfrequenz-Asymptote für den Betrag bei logarithmischer Darstellung von Frequenz und Betrag wegen $\log |G(j\omega)| \approx -\log \omega$ eine Gerade mit der Steigung -1 ist.

Am Schnittpunkt beider Betrags-Asymptoten müssen beide den gleichen Wert haben. Aus der ersten Zeile der Tabelle ergibt sich durch Gleichsetzen der beiden Asymptoten: $1 = \frac{1}{\omega T}$. Damit schneiden sich beide Asymptoten bei der Frequenz $\omega = \frac{1}{T}$. Der Betrag hat bei dieser Frequenz gemäß (4.14) den Wert $|G(\frac{1}{T})| = \frac{1}{\sqrt{2}}$, und die Phase beträgt $\varphi(G(\frac{1}{T})) = -\arctan(1) = -\frac{\pi}{4}$.

Damit lässt sich der Verlauf von Betrag und Phase für diese Teilfunktion nun in einem Bode-Diagramm skizzieren, wie in Bild 4.1 dargestellt.

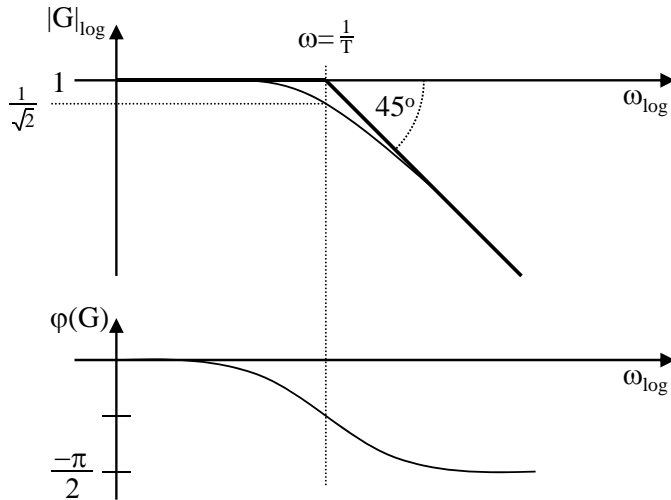


Abb. 4.1. Bode-Diagramm der Übertragungsfunktion $G(s) = \frac{1}{Ts+1}$

Zu beachten ist, dass die Skalen für den Betrag und die Frequenz zwar logarithmisch sind, die Beschriftung aber mit den nicht logarithmierten Werten erfolgt. Dies hat den Vorteil, dass diese Werte für die Diagramme nicht erst umgerechnet werden müssen.

Etwas irritierend mag wirken, dass die Kurvenform für den Winkel φ in diesem Bild nicht dem bekannten Verlauf des Arcustangens entspricht. Aber auch dies ist natürlich durch den logarithmischen Maßstab zu erklären, in dem die Frequenz ω aufgetragen wird.

Ändert sich der Wert der Verzögerungszeit T , so werden die Kurven lediglich in Richtung der ω -Achse verschoben. Bei Berücksichtigung eines Verstärkungsfaktors V ändert sich sogar nur die Beschriftung der $|G|$ -Achse: 1 wird durch V ersetzt, und $\frac{1}{\sqrt{2}}$ durch $\frac{V}{\sqrt{2}}$.

Ablesen lässt sich, dass die Verstärkung für Gleichsignale gleich Eins und für niedrige Frequenzen immer noch nahezu Eins beträgt, während sie für höhere Frequenzen abnimmt und schließlich gegen Null konvergiert. Die negativ reelle Polstelle stellt damit einen Tiefpass dar.

Die Phasenverzögerung nimmt, ausgehend von Null, immer weiter ab und konvergiert gegen $-\frac{\pi}{2}$. Dies ergibt sich natürlich auch sofort bei einer Betrachtung von Gleichung (4.14).

4.2.3 Positiv reelle Polstelle

Als zweiter Fall soll eine Nenner-Teilfunktion G_{nv} mit positiv reeller Polstelle betrachtet werden, wobei auch hier wieder auf die Indizes verzichtet wird:

$$G(s) = \frac{1}{-Ts + 1} \quad \text{mit} \quad T > 0 \quad (4.15)$$

Der Frequenzgang lautet

$$G(j\omega) = \frac{1}{\sqrt{T^2\omega^2 + 1}} e^{j \arctan(T\omega)} \quad (4.16)$$

Für die Asymptoten ergibt sich hier:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	1	$\frac{1}{\omega T}$
$\log G(j\omega) $	0	$-\log \omega$
$\varphi(G(j\omega))$	0	$+\frac{\pi}{2}$

Das Bodediagramm für diese Teilfunktion ist in Bild 4.2 dargestellt. Gegenüber Bild 4.1 hat sich lediglich der Phasenverlauf geändert, der jetzt nicht gegen $-\frac{\pi}{2}$, sondern gegen $+\frac{\pi}{2}$ konvergiert.

Auch die positiv reelle Polstelle ist damit im Prinzip ein Tiefpass. Da hier die Phase für hohe Frequenzen aber positiv und nicht negativ ist, was bedeutet, dass die Schwingung am Ausgang der Eingangsschwingung vorausseilt,

würde man hier nicht direkt von einem Tiefpass sprechen, da ein “echter” Tiefpass üblicherweise mit einer Phasennacheilung einher geht.

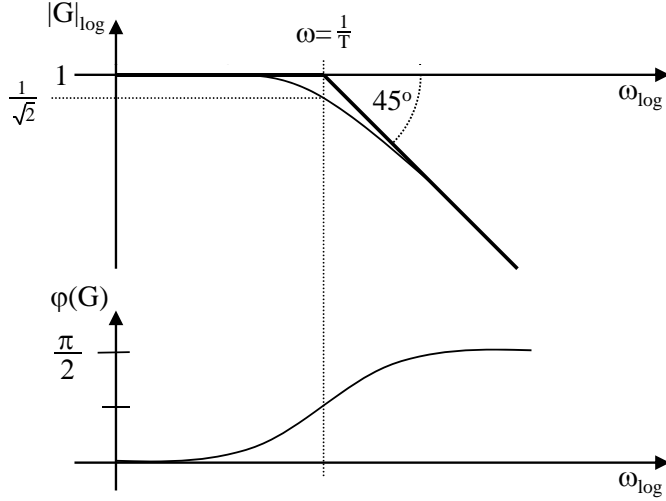


Abb. 4.2. Bode-Diagramm der Übertragungsfunktion $G(s) = \frac{1}{-Ts+1}$

4.2.4 Negativ reelle Nullstelle

Nun sollen die möglichen Zähler-Teilfunktionen betrachtet werden, und zwar zunächst der Fall einer negativ reellen Nullstelle:

$$G(s) = Ts + 1 \quad \text{mit} \quad T > 0 \quad (4.17)$$

Der Frequenzgang lautet

$$G(j\omega) = \sqrt{T^2\omega^2 + 1} e^{j \arctan(T\omega)} \quad (4.18)$$

Für die Asymptoten ergibt sich hier:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	1	ωT
$\log G(j\omega) $	0	$\log \omega$
$\varphi(G(j\omega))$	0	$+\frac{\pi}{2}$

Bild 4.3 zeigt das Bodediagramm für diese Teilfunktion. Der Betrag wächst nun mit ω ins Unendliche, und die Phase konvergiert gegen $+\frac{\pi}{2}$. Die negativ reelle Nullstelle wirkt damit ähnlich zu einem Hochpass.

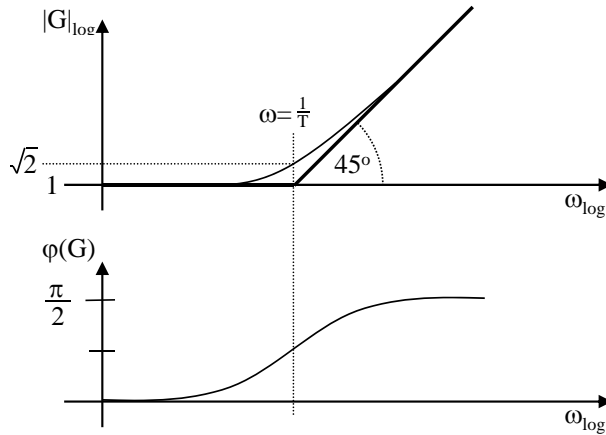


Abb. 4.3. Bode-Diagramm der Übertragungsfunktion $G(s) = Ts + 1$

4.2.5 Positiv reelle Nullstelle

Als letzte Variante bleibt die positiv reelle Nullstelle:

$$G(s) = -Ts + 1 \quad \text{mit} \quad T > 0 \quad (4.19)$$

Der Frequenzgang lautet

$$G(j\omega) = \sqrt{T^2\omega^2 + 1} e^{-j \arctan(T\omega)} \quad (4.20)$$

und für die Asymptoten ergibt sich:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	1	ωT
$\log G(j\omega) $	0	$\log \omega$
$\varphi(G(j\omega))$	0	$-\frac{\pi}{2}$

Bild 4.4 zeigt das Bodediagramm für diese Teilfunktion. Der Betrag wächst wieder mit ω ins Unendliche, aber die Phase konvergiert jetzt gegen $-\frac{\pi}{2}$. Auch die positiv reelle Polstelle wirkt damit ähnlich zu einem Hochpass.

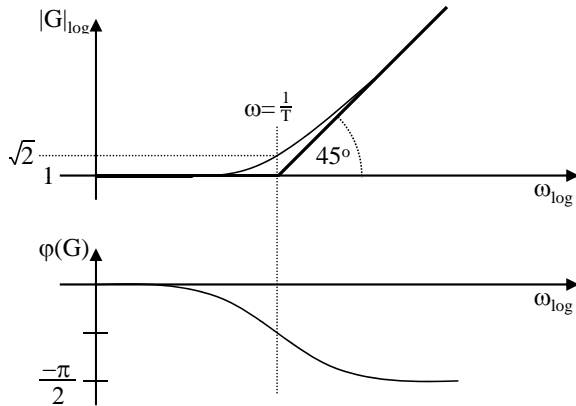


Abb. 4.4. Bode-Diagramm der Übertragungsfunktion $G(s) = -Ts + 1$

4.2.6 Integrator

Neben den Teilfunktionen der rationalen Übertragungsfunktion aus Gleichung (4.3) ist als weiteres elementares Übertragungsglied der Integrator mit der Übertragungsfunktion

$$G(s) = \frac{1}{Ts} \quad \text{mit} \quad T > 0 \quad (4.21)$$

zu betrachten. Der Frequenzgang lautet

$$G(j\omega) = \frac{-j}{T\omega} = \frac{1}{T\omega} e^{-j\frac{\pi}{2}} \quad (4.22)$$

Die Asymptoten sind für niedrige und hohe Frequenzen gleich:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	$\frac{1}{T\omega}$	$\frac{1}{T\omega}$
$\log G(j\omega) $	$-\log \omega$	$-\log \omega$
$\varphi(G(j\omega))$	$-\frac{\pi}{2}$	$-\frac{\pi}{2}$

Entsprechend einfach gestaltet sich das Bode-Diagramm des Integrators (Bild 4.5). Die Phase hat den konstanten Wert $-\frac{\pi}{2}$, während der Betrag monoton fällt. Für den Wert $\omega = \frac{1}{T}$ nimmt er den Wert Eins an, so dass sich bei dieser Frequenz auch der Schnittpunkt mit der ω -Achse befindet.

Der Betrag des Frequenzganges ist für Gleichsignale unendlich groß. Dies ist einsichtig, da der Integrator in diesem Fall immer weiter aufintegriert und es keinen festen Endwert gibt. Das Verhältnis von der Ausgangs- zur Eingangsamplitude wird damit unendlich groß, und dementsprechend muss auch der Betrag des Frequenzganges für die entsprechende Frequenz unendlich groß sein.

Je höher die Frequenz des Eingangssignales, desto mehr macht sich die Trägheit des Integrators bemerkbar, so daß der Betrag für hohe Frequenzen gegen Null konvergiert. Der Integrator kann damit als Tiefpaß aufgefasst werden, obwohl er keine ausgeprägte Grenzfrequenz aufweist.

Die konstante Phasenverzögerung von $-\pi/2$ ergibt sich auch, wenn man das Ausgangssignal des Integrators mit einer Sinusschwingung am Eingang im Zeitbereich berechnet:

$$y(t) = \int_0^t \sin(\omega\tau) d\tau = -\frac{1}{\omega} \cos(\omega t) + \frac{1}{\omega} = \frac{1}{\omega} \sin(\omega t - \frac{\pi}{2}) + \frac{1}{\omega} \quad (4.23)$$

Man sieht, dass das Ausgangssignal nicht nur die verzögerte Sinusschwingung, sondern auch einen konstanten Term $\frac{1}{\omega}$ enthält, obwohl doch nach Satz 4.1 auch am Ausgang eine reine Sinusschwingung auftreten müsste. Dies liegt daran, dass der Integrator als einziges der hier genannten Beispiele die Voraussetzung des Satzes nicht erfüllt, d.h. dass das Integral seiner Impulsantwort nicht konvergiert. $\frac{1}{\omega}$ stellt damit das in Satz 4.1 erwähnte Restglied $r(t)$ dar. Da dies im vorliegenden Fall für eine feste Frequenz aber konstant ist und keine weiteren Auswirkungen hat, kann es auch vernachlässigt werden.

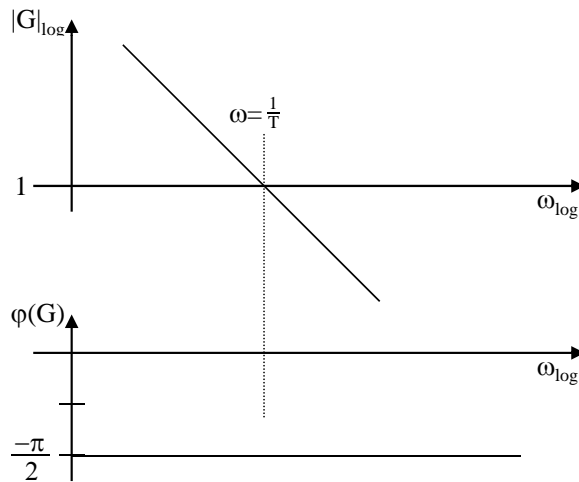


Abb. 4.5. Bode-Diagramm für einen Integrator: $G(s) = \frac{1}{Ts}$

4.2.7 Differenzierer

Genau das gegenteilige Übertragungsverhalten weist ein Differenzierer mit der Übertragungsfunktion

$$G(s) = Ts \quad \text{mit} \quad T > 0 \quad (4.24)$$

auf. Sein Frequenzgang lautet

$$G(j\omega) = jT\omega = T\omega e^{j\frac{\pi}{2}} \quad (4.25)$$

Auch hier sind die Asymptoten für niedrige und hohe Frequenzen gleich:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	$T\omega$	$T\omega$
$\log G(j\omega) $	$\log \omega$	$\log \omega$
$\varphi(G(j\omega))$	$+\frac{\pi}{2}$	$+\frac{\pi}{2}$

Am Betragsverlauf (Bild 4.6) lässt sich deutlich ablesen, dass der Betrag für niedrige Frequenzen gegen Null geht. Dies ist offensichtlich richtig, da die Ableitung eines Gleichsignales (konstante Funktion) Null ergibt. Andersherum steigt der Betrag für hohe Frequenzen hier ins Unendliche. Deshalb sollte man eine Differentiation, und sei sie auch nur näherungsweise realisierbar, in einem Regelkreis vermeiden, denn in dem Fall würden gerade die hochfrequenten Rauschsignale besonders hoch verstärkt.

4.2.8 Laufzeitglied

Relativ einfach sind die Verhältnisse beim Laufzeitglied mit der Übertragungsfunktion

$$G(s) = e^{-T_L s} \quad T_L > 0 \quad (4.26)$$

und dem Frequenzgang

$$G(j\omega) = e^{-jT_L \omega} \quad (4.27)$$

Der Betrag ist offensichtlich konstant gleich Eins. Die Phase $\varphi(j\omega) = -T_L \omega$ wächst negativ linear mit ω . Wenn ω nun, wie im Bode-Diagramm üblich, in einer logarithmischen Skala aufgetragen wird, dann ergibt sich für φ offenbar eine negative Exponentialfunktion (Bild 4.7). Die wachsende Phasennachteilung für steigende Frequenzen ist leicht zu erklären: Die Verzögerung des Eingangssignales ist beim Laufzeitglied mit T_L immer gleich, während der Phasenwinkel $\varphi(j\omega)$ die Verzögerung des Eingangssignales in Abhängigkeit von dessen Frequenz angibt. Je höher diese Frequenz ist, umso mehr Perioden dauert es, bevor das Signal am Ausgang des Laufzeitgliedes erscheint.

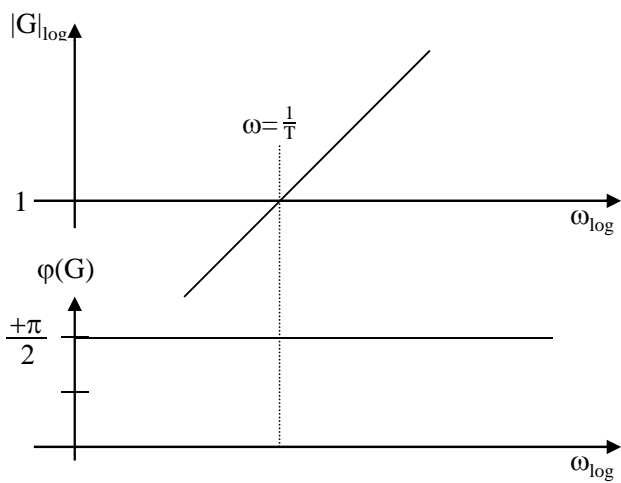


Abb. 4.6. Bode-Diagramm für einen Differenzierer: $G(s) = Ts$

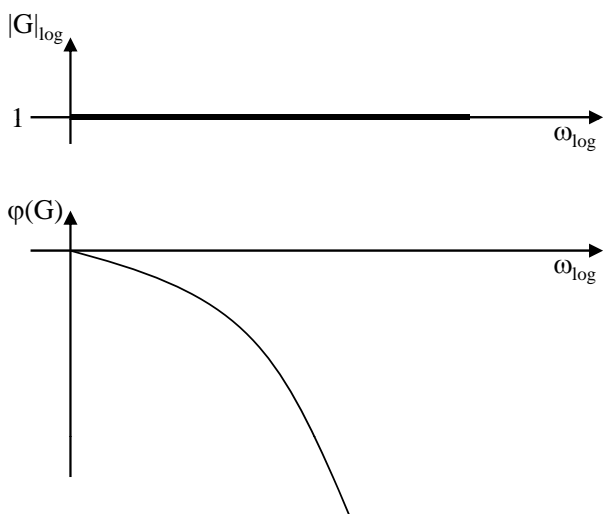


Abb. 4.7. Bode-Diagramm für ein Laufzeitglied: $G(s) = e^{-T_L s}$

4.2.9 Polstellenpaar

Die letzte zu betrachtende Teilfunktion ist ein Polstellenpaar mit

$$G(s) = \frac{1}{\frac{1}{\omega_0^2}s^2 + \frac{2D}{\omega_0}s + 1} = \frac{\omega_0^2}{s^2 + 2D\omega_0s + \omega_0^2} = \frac{\omega_0^2}{(s - s_1)(s - s_2)} \quad (4.28)$$

dessen Übertragungsfunktion, abgesehen von einem Verstärkungsfaktor, dem PT_2 -Glied entspricht. s_1 und s_2 sind komplex-konjugiert zueinander und in (2.32) bereits angegeben.

Für den Frequenzgang ergibt sich

$$G(j\omega) = \frac{1}{1 - (\frac{\omega}{\omega_0})^2 + j2D\frac{\omega}{\omega_0}} \quad (4.29)$$

und für dessen Betrag

$$|G(j\omega)| = \frac{1}{\sqrt{(1 - (\frac{\omega}{\omega_0})^2)^2 + 4D^2(\frac{\omega}{\omega_0})^2}} \quad (4.30)$$

Wenn ω gegen Null konvergiert, dann konvergiert der Betrag gegen Eins. Wird ω dagegen unendlich groß, dann dominiert der Ausdruck mit dem größten Exponenten, also ω^4 , unter der Wurzel im Nenner, so dass sich als Hochfrequenzasymptote $\frac{1}{\omega^2}$ ergibt. Diese Werte sind in der weiter unten stehenden Tabelle noch einmal aufgeführt.

Der Betragsverlauf ist hier allerdings nicht für alle Werte von D monoton fallend. Um dies genauer zu analysieren, soll für den Term unter der Wurzel im Nenner eine Extremwertbestimmung durchgeführt werden:

$$\begin{aligned} f(\omega) &= (1 - (\frac{\omega}{\omega_0})^2)^2 + 4D^2(\frac{\omega}{\omega_0})^2 \\ 0 &= f'(\omega) = 2(1 - (\frac{\omega}{\omega_0})^2)(-2\frac{\omega}{\omega_0^2}) + 8D^2\frac{\omega}{\omega_0^2} \\ 0 &= -4\frac{\omega}{\omega_0} + 4(\frac{\omega}{\omega_0})^3 + 8D^2\frac{\omega}{\omega_0} \end{aligned} \quad (4.31)$$

Da $\omega = 0$ eine Lösung dieser Gleichung ist, liegt für $\omega = 0$ ein potentieller Extremwert vor, welcher allerdings keine Auswirkungen hätte, da er sich am Rand des betrachteten Frequenz-Intervalls $[0, \infty)$ befindet. Für die weitere Rechnung kann daher von $\omega \neq 0$ ausgegangen werden, und die Division durch ω ist zulässig. Aus (4.31) wird damit

$$0 = -1 + (\frac{\omega}{\omega_0})^2 + 2D^2 \quad (4.32)$$

In Abhängigkeit von der Dämpfung D ergeben sich die folgenden Lösungen:

$$\omega = \begin{cases} \omega_0 \sqrt{1 - 2D^2} = \omega_n & : D < \frac{1}{\sqrt{2}} \\ 0 & : D = \frac{1}{\sqrt{2}} \\ \text{keine Lösung} & : D > \frac{1}{\sqrt{2}} \end{cases} \quad (4.33)$$

Für $D > \frac{1}{\sqrt{2}}$ gibt es keine Lösung, und für $D = \frac{1}{\sqrt{2}}$ tritt der Extremwert wieder nur am Rand des Intervalls $[0, \infty)$ bei $\omega = 0$ auf. In beiden Fällen ist der Betragsverlauf daher wie beim PT_1 -Glied monoton fallend.

Sofern jedoch die Dämpfung $D < \frac{1}{\sqrt{2}}$ ist, nimmt der Nenner der Betragsfunktion bei der Frequenz $\omega_0 \sqrt{1 - 2D^2}$ ein Minimum und die Betragsfunktion selber ein Maximum an, was sich leicht nachvollziehen lässt. Es kommt zu einer sogenannten *Resonanzüberhöhung* (Abbildung 4.8). Diese ausgezeichnete Frequenz entspricht aber gerade der in Kapitel 2.5 schon eingeführten *natürlichen Kreisfrequenz* ω_n .

Setzt man diese Frequenz in Gleichung (4.30) ein, so erhält man für den Wert des Betragsmaximums:

$$|G(j\omega)|_{\max} = |G(j\omega_n)| = \frac{1}{2D\sqrt{1 - 2D^2}} \quad \text{für } D < \frac{1}{\sqrt{2}} \quad (4.34)$$

Man erkennt sofort, dass die Resonanzüberhöhung mit $D \rightarrow 0$ gegen ∞ konvergiert. Dies bedeutet, dass für kleine Werte von D selbst eine sehr kleine Anregung eine Schwingung mit sehr großer Amplitude am Ausgang verursachen kann. Für $D \rightarrow 0$ nähert sich dann auch die natürliche Kreisfrequenz, also der Ort des Betragsmaximums, der Resonanzfrequenz ω_0 an.

Der hier ermittelte "Grenzwert" $D = \frac{1}{\sqrt{2}}$ für das Auftreten einer Resonanzüberhöhung ist im Übrigen nicht zu verwechseln mit dem in Kapitel 2.5 eingeführten "Grenzwert" $D = 1$ für die Schwingungsfähigkeit des Systems. Für $D \geq 1$ ist das System nicht schwingungsfähig, beide Pole liegen auf der negativ reellen Achse (vgl. Gleichung (2.32)), das System entspricht damit der Hintereinanderschaltung von zwei PT_1 -Gliedern, und die Sprungantwort weist ein asymptotisches Einschwingverhalten ohne Überspringen auf. Verkleinert man die Dämpfung auf Werte $D < 1$, so wird das Polstellenpaar konjugiert komplex und das System zunächst einmal schwingungsfähig, was man am Überspringen bei der Sprungantwort erkennen kann. Bei weiterer Verkleinerung von D auf Werte $D < \frac{1}{\sqrt{2}}$ tritt dann noch die Resonanzüberhöhung hinzu, d.h. für Frequenzen im Bereich um ω_n ist der Betrag des Frequenzganges größer als Eins und kann für kleine Werte von D sogar sehr große Werte annehmen. Dies bedeutet, dass bei einer Schwingung am Eingang des Übertragungsgliedes die Schwingung am Ausgang eine größere Amplitude aufweist als die Eingangsschwingung. Dies kann katastrophale Folgen haben, wie die berühmten Bilder von jener großen Hängebrücke aus Stahl und Beton zeigen, die bei einem Sturm in Schwingungen im Bereich ihrer Resonanzfrequenz versetzt wurde und schließlich - nach eindrucksvollen Bewegungen - in sich zusammenstürzte.

In der nachfolgenden Tabelle sind die Betrachtungen noch einmal zusammengefasst:

Dämpfung	Pole	Bemerkung
$D \geq 1$	neg. reell	aperiodisches Einschwingen
$1 > D \geq \frac{1}{\sqrt{2}}$	konj. komplex	Sprungantw. mit Überschw.
$\frac{1}{\sqrt{2}} > D \geq 0$	konj. komplex	Überschw. und Resonanzüberh.

Von daher sollte man bei der Auslegung eines Regelkreises darauf achten, dass die Dämpfung im Falle von konjugiert-komplexen Polstellenpaaren nie kleiner als $D = \frac{1}{\sqrt{2}}$ wird. Ob darüber hinaus die Dämpfung sogar größer als Eins sein sollte, hängt von den Gegebenheiten ab. Eine höhere Dämpfung bedeutet nämlich auch eine Verlangsamung des Einschwingverhaltens. Aus dem Grund nimmt man oft ein leichtes Überspringen in Kauf, um den Regelkreis nicht allzu langsam zu machen. Nur in speziellen Fällen, wie beispielsweise der Positionsregelung von Aufzügen oder Fräsmaschinen, in denen ein Überspringen völlig inakzeptabel wäre, wird man den Wert der Dämpfung auf $D \geq 1$ festlegen.

Nach dem Betragsverlauf ist nun noch der Phasenverlauf zu betrachten. Hierzu wird die letzte der Darstellungen in Gleichung (4.28) betrachtet und der Frequenzgang in zwei Faktoren entsprechend den Polstellen s_1 und s_2 der Übertragungsfunktion zerlegt. Diese Polstellen lauten gemäß Gleichung (2.32)

$$s_{1,2} = \begin{cases} \omega_0 [-D \pm \sqrt{D^2 - 1}] & : D \geq 1 \\ \omega_0 [-D \pm j\sqrt{1 - D^2}] & : D < 1 \end{cases} \quad (4.35)$$

Damit ergibt sich für den Frequenzgang

$$\begin{aligned} G(j\omega) &= \frac{\omega_0^2}{(j\omega - s_1)(j\omega - s_2)} \\ &= \begin{cases} \frac{\omega_0}{j\omega + \omega_0[D - \sqrt{D^2 - 1}]} \frac{\omega_0}{j\omega + \omega_0[D + \sqrt{D^2 - 1}]} & : D \geq 1 \\ \frac{\omega_0}{\omega_0 D + j(\omega - \omega_0 \sqrt{1 - D^2})} \frac{\omega_0}{\omega_0 D + j(\omega + \omega_0 \sqrt{1 - D^2})} & : D < 1 \end{cases} \quad (4.36) \end{aligned}$$

und für die Phase

$$\varphi(G(j\omega)) = \begin{cases} -\arctan \frac{\omega}{\omega_0[D - \sqrt{D^2 - 1}]} - \arctan \frac{\omega}{\omega_0[D + \sqrt{D^2 - 1}]} & : D \geq 1 \\ -\arctan \frac{\omega - \omega_0 \sqrt{1 - D^2}}{\omega_0 D} - \arctan \frac{\omega + \omega_0 \sqrt{1 - D^2}}{\omega_0 D} & : D < 1 \end{cases} \quad (4.37)$$

Für niedrige Frequenzen folgt

$$\lim_{\omega \rightarrow 0} \varphi(G(j\omega)) = \begin{cases} 0 & : D \geq 1 \\ -\arctan \frac{-\sqrt{1-D^2}}{D} - \arctan \frac{\sqrt{1-D^2}}{D} = 0 & : D < 1 \end{cases} \quad (4.38)$$

und für hohe Frequenzen

$$\lim_{\omega \rightarrow \infty} \varphi(G(j\omega)) = \begin{cases} -\arctan \infty - \arctan \infty = -\pi & : D \geq 1 \\ -\arctan \infty - \arctan \infty = -\pi & : D < 1 \end{cases} \quad (4.39)$$

Zusammengefasst ergeben sich für Betrag und Phase die folgenden Asymptoten:

	$\omega T \ll 1$	$\omega T \gg 1$
$ G(j\omega) $	1	$\frac{1}{\omega^2}$
$\log G(j\omega) $	0	$-2 \log \omega$
$\varphi(G(j\omega))$	0	$-\pi$

Bild 4.8 zeigt den Betrags- und Phasenverlauf. Die Phase fällt monoton von 0 auf $-\pi$. Der Betrag geht von Eins gegen Null mit der Hochfrequenz-Asymptote $-2 \log \omega$, wobei hier aber, in Abhängigkeit von der Dämpfung D , eine Resonanzüberhöhung auftreten kann, welche umso größer wird, je kleiner der Wert der Dämpfung D ist.

4.2.10 Überlagerung von Teilfunktionen

Nachdem in den letzten Kapiteln die Bode-Diagramme von möglichen Teilfunktionen entwickelt wurden, soll nun gezeigt werden, wie Betrags- und Phasenverlauf einer Übertragungsfunktion durch Überlagerung der Betrags- und Phasenverläufe der verschiedenen Teilfunktionen sehr einfach und schnell konstruiert werden können.

In Kapitel 4.2.1 sind die mathematischen Grundlagen dazu bereits hergeleitet worden: Die Phase der Gesamtfunktion ist die Summe der Phasen der Teilfunktionen (4.11), so dass sie direkt durch eine additive Überlagerung der einzelnen Phasenverläufe der Teilfunktionen gewonnen werden kann. Für die Beträge gilt das Gleiche, sofern man deren logarithmierte Werte betrachtet (4.12).

Anhand eines Beispiels soll die Vorgehensweise nun demonstriert werden. Gegeben sei die Übertragungsfunktion

$$G(s) = V \frac{1}{T_1 s} \frac{T_2 s + 1}{(T_3 s + 1)(T_4 s + 1)} \quad \text{mit} \quad T_1 > T_2 > T_3 > T_4 > 0 \quad (4.40)$$

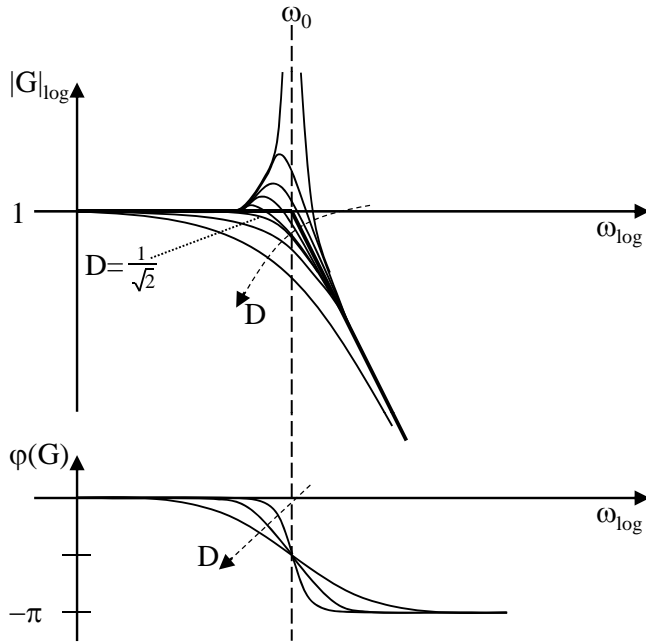


Abb. 4.8. Bode-Diagramm eines konjugiert-komplexen Polstellenpaares (PT_2 -Glied)

Diese Funktion besteht aus einem Integrator mit der Teilfunktion $\frac{1}{T_1 s}$, einer negativ reellen Nullstelle mit der Teilfunktion $T_2 s + 1$ und zwei negativ reellen Polstellen mit den Teilfunktionen $\frac{1}{T_3 s + 1}$ und $\frac{1}{T_4 s + 1}$. Die Asymptoten der Phasen- und Betragsverläufe aller vier Teilfunktionen sind in Abbildung 4.9 eingetragen. Im Betragsdiagramm sieht man die Betragskurve des Integrators als Gerade mit einem ω -Achsenschnittpunkt bei $\frac{1}{T_1}$ (vgl. Kap. 4.2.6) sowie die Hochfrequenzasymptoten für die Nullstelle mit dem Achsenschnittpunkt bei $\frac{1}{T_2}$ (vgl. Kap. 4.2.4) und für die beiden Polstellen mit den Achsenschnittpunkten bei $\frac{1}{T_3}$ und $\frac{1}{T_4}$ (vgl. Kap. 4.2.2). Für niedrige Frequenzen sind die Beträge von Null- und Polstellen jeweils annähernd gleich Eins, so dass sie den Betrag der Gesamtfunktion praktisch nicht verändern. Von daher sind die Niederfrequenzasymptoten dieser drei Teilfunktionen nicht im Diagramm eingetragen.

Ohne Integrator wäre der Betrag der Gesamtfunktion für niedrige Frequenzen bzw. Gleichsignale gerade gleich V . Dies ist durch die Beschriftung der $|G|_{\log}$ -Achse an der Stelle des Achsenschnittpunktes berücksichtigt. Da nur die Skalierung der $|G|_{\log}$ -Achse logarithmisch ist, nicht jedoch ihre Bezifferung, ist dieser Punkt mit V und nicht mit $\log V$ bezeichnet. Eine andere Verstärkung ändert den Verlauf der Kurven nicht, sondern lediglich die Beschriftung an der $|G|_{\log}$ -Achse.

Im Phasendiagramm finden sich die angedeuteten Phasenverläufe der Teilfunktionen, allerdings nicht der des Integrators. Dieser ist konstant gleich $-\frac{\pi}{2}$ und wegen seiner Einfachheit weggelassen. In der Überlagerung der einzelnen Funktionen ist dieser konstante Wert aber zu berücksichtigen. Die Phasenverläufe der übrigen Teilfunktionen lassen sich direkt den jeweiligen Zeitkonstanten zuordnen.

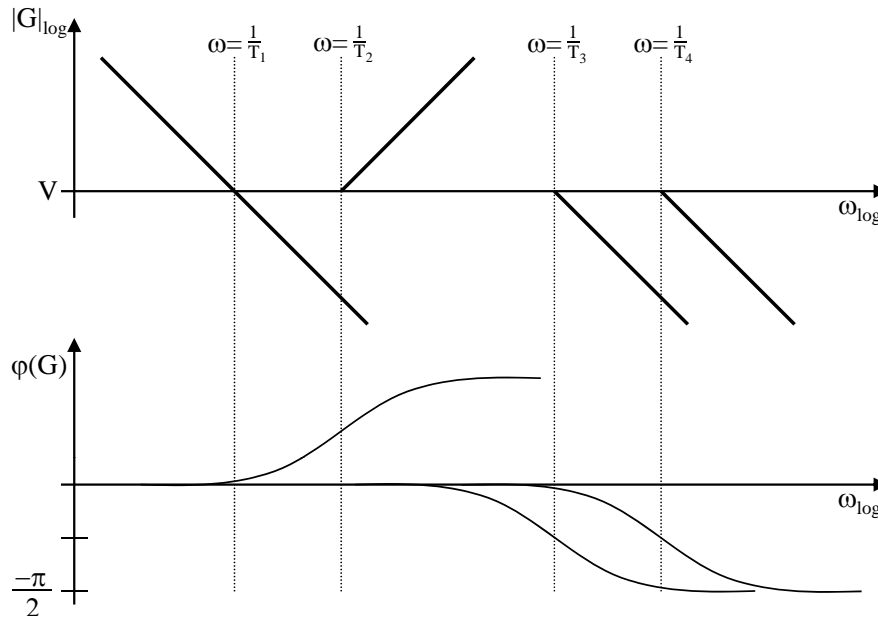


Abb. 4.9. Betrags- und Phasenverläufe der Teil-Übertragungsfunktionen

Die Überlagerung der Teilfunktionen wird in Abbildung 4.10 gezeigt. Der Betragsverlauf wird zunächst ausschließlich durch den Betragsverlauf des Integrators bestimmt, da die Beträge der übrigen Teilfunktionen alle noch annähernd gleich Eins sind. Erst ab $\frac{1}{T_2}$ heben sich die negative Steigung des Betragsverlaufes des Integrators und die positive Steigung des Betragsverlaufes der Nullstelle gegenseitig auf, so daß die Betragskurve zwischen $\frac{1}{T_2}$ und $\frac{1}{T_3}$ einen konstanten Verlauf hat. Ab $\frac{1}{T_3}$ tritt dann die negative Steigung des Betragsverlaufes der ersten Polstelle und ab $\frac{1}{T_4}$ die negative Steigung des Betragsverlaufes der zweiten Polstelle hinzu, so daß die Betragskurve bei diesen Frequenzen jeweils einen Knick nach unten erfährt.

Auch der Phasenverlauf wird zunächst durch den Integrator bestimmt. Die Kurve startet bei $-\frac{\pi}{2}$, der konstanten Phase des Integrators. Mit Annäherung an die Frequenz $\frac{1}{T_2}$ wirkt sich der Phasenverlauf der Nullstelle aus und hebt die Kurve an. Ohne weitere Teilfunktionen würde die Phase um $\frac{\pi}{2}$ bis auf Null ansteigen. Bevor dieser Wert erreicht wird, beginnt sich jedoch zwischen

$\frac{1}{T_2}$ und $\frac{1}{T_3}$ die erste der beiden Polstellen auszuwirken, und die Phase senkt sich wieder ab. Zwischen $\frac{1}{T_3}$ und $\frac{1}{T_4}$ erfolgt dann eine weitere Absenkung wegen der zweiten Polstelle. Der Endwert der Phase für sehr hohe Frequenzen beträgt $-\pi$ als Summe der Phasenanstiegen oder -absenkungen $-\frac{\pi}{2}$ (Integrator), $+\frac{\pi}{2}$ (Nullstelle) und $-2\frac{\pi}{2}$ (Zwei Polstellen).

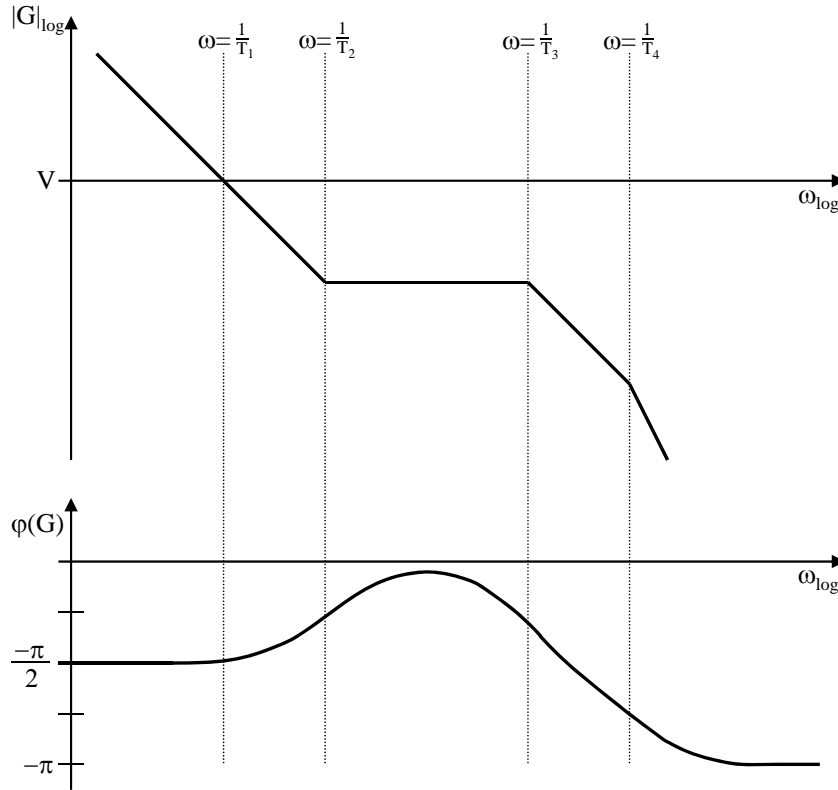


Abb. 4.10. Betrags- und Phasenverlauf der Gesamt-Übertragungsfunktion

4.3 Ortskurve

Während beim Bode-Diagramm Betrag und Phase des Frequenzganges in zwei getrennten Diagrammen gezeichnet werden, kann die Darstellung von $G(j\omega)$ auch als frequenzabhängige Kurve in der komplexen Ebene erfolgen. Dies ist die sogenannte *Ortskurve*. Da $G(j\omega)$ zu $G(-j\omega)$ konjugiert komplex ist, enthält die Kurve für positive Werte von ω bereits die gesamte Information des Frequenzganges, und man kann sich auf die Darstellung von $G(j\omega)$ für $\omega \in [0, \infty)$ beschränken.

Ortskurven sind ein hervorragendes Hilfsmittel, um die Stabilität eines Systems oder die notwendigen Maßnahmen, um Stabilität zu erzielen, schnell und anschaulich bestimmen zu können. Dies wird im nachfolgenden Kapitel behandelt. Hier soll zunächst anhand einiger Beispiele gezeigt werden, wie Ortskurven ohne Zuhilfenahme eines Computers von Hand skizziert werden können. Die Ausgangssituation ist dabei die gleiche wie bei den Bode-Diagrammen: Gegeben sei eine Übertragungsfunktion, deren Frequenzgang in Form einer Ortskurve dargestellt werden soll, um anhand dieser Darstellung eine Stabilitätsanalyse zu ermöglichen. In einem Zwischenschritt wird dabei zunächst das Bode-Diagramm entwickelt, und anschließend aus dem Bode-Diagramm die Ortskurve.

4.3.1 Verzögerungsglied erster Ordnung

Das erste Beispiel sei ein PT_1 -Glied mit der Übertragungsfunktion (vgl. Gleichung 3.6)

$$G(s) = \frac{V}{Ts + 1} \quad (4.41)$$

und dem Frequenzgang

$$G(j\omega) = \frac{V}{Tj\omega + 1} = \frac{V}{\sqrt{\omega^2 T^2 + 1}} e^{-j \arctan \omega T} \quad (4.42)$$

Das Bode-Diagramm entspricht dem einer negativ reellen Polstelle, wie in Abbildung 4.1 gezeigt. Lediglich an der Betragsachse ist die Eins durch V und $\frac{1}{\sqrt{2}}$ durch $\frac{V}{\sqrt{2}}$ zu ersetzen.

Will man nun den Frequenzgang als Kurve in der komplexen Ebene zeichnen, so stellt jeder Punkt dieser Kurve für eine bestimmte Frequenz ω_1 den komplexen Wert $G(j\omega_1)$ dar. Ein Punkt in der komplexen Ebene kann aber auch als Endpunkt eines Vektors interpretiert werden, der im Ursprung beginnt. Für jeden Punkt der Ortskurve entspricht dann der Betrag dieses Vektors gerade dem Betrag des Frequenzganges $|G(j\omega_1)|$ für die jeweilige Frequenz ω_1 , und der Winkel des Vektors entspricht der Phase $\varphi(G(j\omega_1))$.

Mit Hilfe dieser Interpretation ist nun das Zeichnen der Ortskurve relativ einfach. Für das vorliegende Verzögerungsglied lässt sich am Bode-Diagramm ablesen, dass der Frequenzgang für $\omega = 0$ den Betrag V und die Phase $\varphi = 0$ aufweist. Der entsprechende Vektor, der im Ursprung beginnt, hat also eine Länge von V und weist genau in Richtung der reellen Achse. Als Anfangspunkt der Ortskurve ergibt sich damit der Punkt V auf der positiv reellen Achse.

Anschließend verkleinert sich laut Bode-Diagramm der Betrag bis auf den Wert Null, während die Phase negative Werte annimmt und sich von Null auf $-\frac{\pi}{2}$ verkleinert. Die Vektoren, die zu den Punkten der Ortskurve führen, werden also immer kürzer und drehen sich von Null auf $-\frac{\pi}{2}$. Damit ergibt

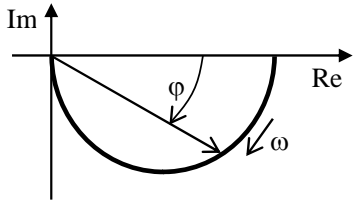


Abb. 4.11. Ortskurve des PT_1 -Gliedes

sich die in Abbildung 4.11 gezeichnete Ortskurve, die senkrecht von unten in den Ursprung läuft.

Wie beim Bode-Diagramm, so lassen sich auch anhand der Ortskurve bestimmte Eigenschaften des Übertragungsgliedes ablesen. Der Anfangspunkt der Ortskurve bezeichnet beispielsweise immer das Übertragungsverhalten für Signale mit der Frequenz $\omega = 0$, also für Gleichsignale. Im vorliegenden Fall lässt sich ablesen, dass die Verstärkung für Gleichsignale den Wert V hat. Dies ist gleichzeitig auch der Endwert der Sprungantwort. Denn auch bei einem Sprung mit Sprungantwort handelt es sich nach Abklingen der Einschwingvorgänge um Gleichsignale am Ein- und Ausgang. Und da der Sprung die Amplitude Eins aufweist, muss das Ausgangssignal gerade die Amplitude $1V = V$ besitzen.

Geht man etwas weiter auf der Ortskurve, so sieht man, dass die Verstärkung für niederfrequente Signale noch nahezu gleich V ist. Je höher die Frequenz wird, desto geringer wird die Verstärkung und desto mehr werden demnach die Signale durch das Verzögerungsglied gedämpft. Das Verzögerungsglied ist daher ein Tiefpass.

Mit zunehmender Frequenz nimmt aber auch die Phasenverzögerung durch das PT_1 -Glied immer weiter zu. Für $\omega = 0$ ist $G(j\omega)$ rein reell und die Phasenverzögerung damit Null, während für wachsende Frequenzen die Phasenverzögerung gegen den Wert $-\frac{\pi}{2}$ konvergiert. Eine hochfrequente Sinusschwingung wird durch das Verzögerungsglied also nicht nur gedämpft, sondern auch um nahezu eine Viertelperiode verzögert.

Trotz dieser Möglichkeiten muss festgestellt werden, dass die Ortskurve definitiv weniger Information enthält als ein Bode-Diagramm. Denn die Ortskurve entsteht ja aus dem Bode-Diagramm durch Elimination der ω -Achse und Verschmelzung von Betrags- und Phaseninformation in einer einzigen Kurve. Damit ist bei der Ortskurve im Gegensatz zum Bode-Diagramm eine Zuordnung von Betrag und Phase zu einer bestimmten Frequenz nicht mehr möglich. Will man also bestimmen, wie das Übertragungsglied auf die Anregung durch ein Signal einer bestimmten Frequenz reagiert, so kann die Beantwortung dieser Frage nur anhand des Bode-Diagramms, nicht aber anhand der Ortskurve erfolgen.

Dieser Mangel ist für Fragestellungen im Zusammenhang mit der Stabilität eines Regelkreises jedoch nicht weiter relevant, während auf der anderen

Seite die Informationsreduktion zu einer deutlich besseren Übersichtlichkeit führt. Dies wird in den nachfolgenden Kapiteln noch demonstriert.

4.3.2 Verzögerungsglied zweiter Ordnung

Das zweite Beispiel sei ein Verzögerungsglied zweiter Ordnung bzw. ein PT_2 -Glied mit der Übertragungsfunktion

$$G(s) = \frac{V}{\frac{1}{\omega_0^2}s^2 + \frac{2D}{\omega_0}s + 1} \quad (4.43)$$

dessen Frequenzgang und Bode-Diagramm im Abschnitt 4.2.9 ausführlich analysiert worden sind, wobei dort der Verstärkungsfaktor gleich Eins war. Das Hinzufügen eines von Eins verschiedenen Verstärkungsfaktors V erfordert jedoch lediglich eine Änderung der Beschriftung der Betragsachse im Bode-Diagramm 4.8, d.h. die 1 muss dort durch V ersetzt werden.

Am Bode-Diagramm lässt sich ablesen, dass der Frequenzgang wie beim PT_1 -Glied für $\omega = 0$ mit dem Betrag V und der Phase $\varphi = 0$ beginnt. Sofern keine Resonanzüberhöhung vorliegt, verkleinert sich der Betrag wieder bis auf den Wert Null, während sich die Phase von Null jetzt auf $-\pi$ verringert, wobei sie bei der Kreisfrequenz ω_0 den Wert $-\frac{\pi}{2}$ annimmt. Die Vektoren zu den Punkten der Ortskurve haben also zunächst den Betrag V und werden für höhere Frequenzen auch hier immer kürzer, während sie sich von 0 auf $-\pi$ drehen.

Damit zeigen die Vektoren zu den Punkten der Ortskurve für hohe Frequenzen hier nicht mehr wie beim PT_1 -Glied nach unten, sondern nach links, und die Ortskurve läuft dementsprechend nicht von unten, sondern von links in den Ursprung. Dies entspricht einer Verzögerung von hochfrequenten Eingangsschwingungen um eine halbe Periode.

Für kleiner werdende Werte von D tritt ab $D < \frac{1}{\sqrt{2}}$ eine Resonanzüberhöhung auf. Der Betrag der Übertragungsfunktion wird insbesondere in der Nähe der Resonanzfrequenz ω_n vergrößert. Dementsprechend werden die Vektoren zu den Punkten der Ortskurve länger, und die Ortskurve wird gedehnt. Der maximale Betrag, d.h. die größte Entfernung zum Ursprung, tritt bei der Frequenz ω_n auf.

Es ergeben sich die in Abbildung 4.12 gezeichneten Ortskurven, wobei hier zwei Varianten in Abhängigkeit von der Dämpfung D gezeichnet sind. Jede dieser Ortskurven beginnt wie beim PT_1 -Glied auf der reellen Achse, schneidet bei der Frequenz ω_0 die imaginäre Achse und läuft unter einem Winkel von $-\pi$ von links in den Ursprung.

Die Phasendrehung des PT_2 -Gliedes von $-\pi$ für hohe Frequenzen entspricht der Erkenntnis, die schon in Kapitel 4.2.10 über den Endwert des Phasenverlaufes gewonnen wurde. Nullstellen und Differenzierer, d.h. Teilfunktionen, deren Terme im Zähler der Übertragungsfunktion auftauchen, verdrehen die Phase jeweils um $+\frac{\pi}{2}$, während Polstellen und Integratoren,

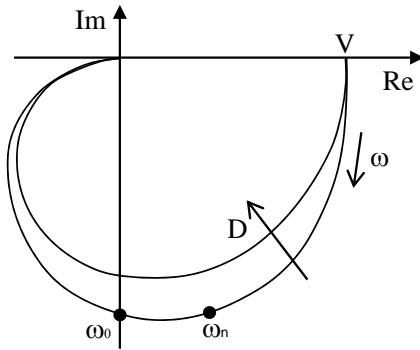


Abb. 4.12. Ortskurve des PT_2 -Gliedes

deren Terme sich im Nenner der Übertragungsfunktion finden, die Phase um jeweils $-\frac{\pi}{2}$ verdrehen. Die gesamte Phasendrehung und damit der Winkel, unter dem die Ortskurve in den Ursprung läuft, lässt sich dann sofort durch simples Abzählen der Zähler- und Nenner-Teilfunktionen ermitteln.

4.3.3 Integrator

Die Darstellung der Ortskurve eines Integrators ist mit den Ergebnissen aus 4.2.6 trivial. Der Betragsverlauf ist monoton fallend von Unendlich bis auf Null, während die Phase konstant gleich $-\frac{\pi}{2}$ ist. Damit sind alle Vektoren, die zu Punkten auf der Ortskurve zeigen, senkrecht nach unten gerichtet. Ihr Betrag ist zunächst unendlich groß und konvergiert dann gegen Null. Somit verläuft die Ortskurve auf der imaginären Achse von $-\infty$ gegen Null, wie in Abbildung 4.13 zu sehen ist.

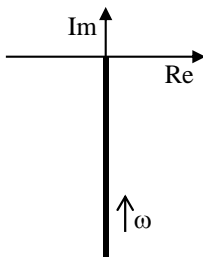


Abb. 4.13. Ortskurve des Integrators

Aus dem Verlauf der Kurve lässt sich folgern, dass die Signalverstärkung für Gleichsignale entsprechend dem Anfangspunkt der Ortskurve unendlich groß ist. Dies ist aber auch plausibel, da der Ausgang des Integrators bei Anlegen eines Gleichsignales am Eingang immer weiter aufintegriert.

Die Phasenverzögerung beträgt beim Integrator immer $-\frac{\pi}{2}$. Auch dies ist sofort plausibel, wenn man sich vergegenwärtigt, dass sich für das Ausgangssignal des Integrators bei Anlegen einer Sinusschwingung gerade $\int \sin t dt = -\cos t$ ergibt und $-\cos t$ gegenüber $\sin t$ eine Viertelperiode nacheilt.

4.3.4 Laufzeitglied

Ähnlich trivial sind die Verhältnisse beim Laufzeitglied mit dem Frequenzgang $G(j\omega) = e^{-jT_L\omega}$. Wie schon in Abschnitt 4.2.8 beschrieben, ist der Betrag für alle Frequenzen konstant gleich Eins, während die Phase $\varphi(j\omega) = -T_L\omega$ negativ linear mit ω wächst.

Demnach sind die Vektoren vom Ursprung zur Ortskurve für alle Frequenzen gleich lang, während der Winkel für wachsende Frequenzen in negativer Richtung bis ins Unendliche ansteigt. Die Ortskurve muss daher einen Kreis beschreiben, der in negativer Richtung immer weiter durchlaufen wird (Abbildung 4.14).

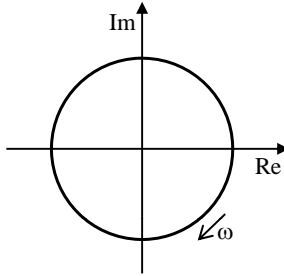


Abb. 4.14. Ortskurve des Laufzeitgliedes

4.3.5 Integrator mit Verzögerungsglied

Relativ häufig kommt in der Praxis eine Hintereinanderschaltung von Integrator und Verzögerungsglied, das sogenannte IT_1 -Glied, vor. Die Übertragungsfunktion lautet

$$G(s) = \frac{1}{T_1 s} \frac{V}{T_2 s + 1} \quad (4.44)$$

und der Frequenzgang

$$G(j\omega) = \frac{1}{j\omega T_1} \frac{V}{T_2 j\omega + 1} = \frac{-VT_2}{T_1(\omega^2 T_2^2 + 1)} + j \frac{-V}{\omega T_1(\omega^2 T_2^2 + 1)} \quad (4.45)$$

Für $\omega \rightarrow 0$ konvergiert der Realteil gegen $-\frac{VT_2}{T_1}$ und der Imaginärteil gegen $-\infty$. Die Ortskurve beginnt also bei negativ unendlichen Imaginärteilen links von der imaginären Achse mit dem Realteil $Re(G) = -\frac{VT_2}{T_1}$. Für $\omega \rightarrow \infty$

gehen Real- und Imaginärteil gegen Null, d.h. die Ortskurve endet auch hier im Ursprung der komplexen Ebene.

Zu klären ist noch der Phasenverlauf bei Annäherung an den Ursprung. Diesen ermittelt man am besten anhand des Bode-Diagramms. Die Phasenverläufe von Integrator und PT_1 -Glied sind analog zum Abschnitt 4.2.10 zu überlagern. Die Phase des Integrators beträgt konstant $-\frac{\pi}{2}$, und der Phasenverlauf des PT_1 -Gliedes beginnt bei Null und endet bei $-\frac{\pi}{2}$. Für die Überlagerung ergibt sich daher der in Abbildung 4.15 im linken Diagramm gezeigte Verlauf. Die Phase des IT_1 -Gliedes beginnt bei $-\frac{\pi}{2}$ und endet bei $-\pi$. Die Vektoren zu den Punkten der Ortskurve haben demnach für hohe Frequenzen den Winkel $-\pi$ und zeigen nach links, so dass die Ortskurve von links in den Ursprung hinein laufen muss.

Ein Widerspruch ergibt sich aber beim Startwinkel der Ortskurve. Aus (4.45) wurde ermittelt, dass die Ortskurve links von der imaginären Achse beginnt. Andererseits beträgt der Winkel der Vektoren zu den Punkten der Ortskurve laut Bode-Diagramm für niedrigen Frequenzen $-\frac{\pi}{2}$. Die Vektoren zeigen nach unten, und demnach müsste die Ortskurve auf der imaginären Achse beginnen, was im Widerspruch zu (4.45) steht. Dies ist aber nicht ganz korrekt. Tatsächlich wirkt sich die Phasenabsenkung durch das PT_1 -Glied auch schon bei kleinen Frequenzen aus, so dass der Phasenverlauf im Bode-Diagramm nicht bei $-\frac{\pi}{2}$, sondern leicht darunter beginnt. Der Winkel der Vektoren zu den Anfangspunkten der Ortskurve beträgt daher nicht $-\frac{\pi}{2}$, sondern $-\frac{\pi}{2} - \varepsilon$. Und damit zeigen die Vektoren nicht genau nach unten, sondern schräg nach links unten, womit Bode-Diagramm und Gleichung (4.45) doch die gleiche Aussage liefern.

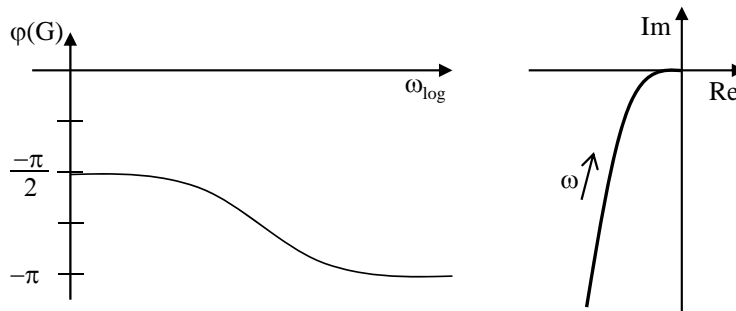


Abb. 4.15. Phasenverlauf und Ortskurve des IT_1 -Gliedes

4.3.6 Zusammengesetzte Übertragungsfunktion

Abschließend soll noch die Ortskurve der in Abschnitt 4.2.10 betrachteten Übertragungsfunktion

$$G(s) = V \frac{1}{T_1 s} \frac{T_2 s + 1}{(T_3 s + 1)(T_4 s + 1)} \quad \text{mit} \quad T_1 > T_2 > T_3 > T_4 > 0 \quad (4.46)$$

entwickelt werden.

Aus dem Bode-Diagramm in Abbildung 4.10 sieht man zunächst, dass der Betrag bei unendlichen Werten beginnt und gegen Null konvergiert. Die Kurve muss also in unendlicher Entfernung vom Ursprung der komplexen Ebene beginnen und im Ursprung enden.

Am Phasenverlauf erkennt man, dass der Winkel zunächst $-\frac{\pi}{2}$ beträgt. Die Vektoren zu den Punkten der Ortskurve zeigen also zunächst nach unten, und daher muss die Ortskurve aus Richtung der negativ imaginären Achse, also von unten, kommen.

Laut Bode-Diagramm vergrößert sich dann die Phase, bevor sie wieder abfällt und schließlich gegen $-\pi$ konvergiert. Die Vektoren zu den Punkten der Ortskurve drehen sich also zunächst in positiver Richtung nach rechts, bevor sie sich für hohe Frequenzen nach links verdrehen, bis sie exakt nach links zeigen. Damit verläuft die Ortskurve, aus Richtung der negativ imaginären Achse kommend, zunächst noch im vierten Quadranten, bevor die Winkel der Vektoren wieder kleiner werden, die Ortskurve die imaginäre Achse kreuzt und schließlich von links unter dem Winkel $-\pi$ in den Ursprung hinein läuft (Abbildung 4.16).

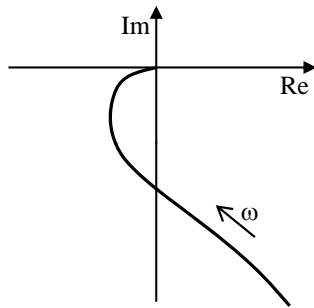


Abb. 4.16. Ortskurve einer zusammengesetzten Übertragungsfunktion

Nach diesen Beispielen lassen sich für Übertragungsfunktionen gemäß Gleichung (3.36), d.h. rationale Übertragungsfunktionen mit Laufzeit, einige allgemeine Aussagen machen: Ist die Ordnung des Zählerpolynoms kleiner als die Ordnung des Nennerpolynoms, so enden die Ortskurven immer im Ursprung der komplexen Ebene, da der Betrag der Übertragungsfunktion offensichtlich gegen Null konvergiert. Falls keine Polstelle bei Null vorliegt, liegt der Anfangspunkt immer auf einem endlichen reellen Wert, wie es beim PT_1 - oder PT_2 -Glieder der Fall ist. Ist mindestens ein Pol bei Null vorhanden, so beginnt die Ortskurve im Unendlichen, und zwar unter einem Winkel $-k\frac{\pi}{2}$, wobei k die Ordnung dieses Pols ist. Phasen- und Betragsverlauf müssen aber keinesfalls immer monoton fallen, wie schon die Betrachtung des

PT_2 –Gliedes gezeigt hat. Der Verlauf der Ortskurve hängt von der Verteilung der Pol- und Nullstellen der Übertragungsfunktion ab.

5. Stabilität linearer Eingrößensysteme

5.1 Definition der Stabilität

In diesem Kapitel soll für lineare Systeme mit einer Ein- und einer Ausgangsgröße der Begriff der Stabilität erläutert werden, die durch die Übertragungsfunktion

$$\begin{aligned} G(s) &= \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} e^{-T_L s} \quad \text{mit } m \leq n \\ &= \frac{b_m}{a_n} \frac{\prod_{\mu=1}^m (s - n_\mu)}{\prod_{\nu=1}^n (s - p_\nu)} e^{-T_L s} = V \frac{\prod_{\mu=1}^m (\frac{-s}{n_\mu} + 1)}{\prod_{\nu=1}^n (\frac{-s}{p_\nu} + 1)} e^{-T_L s} \end{aligned} \quad (5.1)$$

mit dem Verstärkungsfaktor

$$V = \frac{b_0}{a_0} \quad (5.2)$$

gegeben sind.

Zur Einführung sei dabei noch einmal auf das Beispiel vom Ende des Kapitels 1 hingewiesen. Das zu regelnde System ist ein Schiff, und die Regelgröße ist der Kurs des Schiffes. Im Beispiel wird erläutert, wie es bei Auswahl einer falschen Regelstrategie in Folge einer Abweichung vom Sollkurs zu Schwingungen kommen kann, die im schlimmsten Fall sogar aufklingend sind. Wie die Abweichung vom Sollkurs und damit die Anregung des Systems entstanden ist, also entweder durch eine Veränderung des Sollkurses oder durch eine Störung von außen wie zum Beispiel eine Querströmung, ist dabei nicht relevant.

Generell können drei Arten von Systemantworten auf eine bestimmte Anregung auftreten. Im günstigsten Fall kommt das System nach einer Anregung, gegebenenfalls nach einigen Schwingungen, wieder zur Ruhe. In diesem Fall spricht man von einem stabilen System. Dabei ist für die Frage der Stabilität nicht relevant, ob die Regelgröße im Ruhezustand dem Sollwert entspricht. Dies wäre die Frage nach der Regelgenauigkeit, die hier aber noch nicht behandelt werden soll.

Im zweiten Fall wird das System durch eine Anregung in Schwingungen versetzt, die nicht mehr abklingen, so dass das System theoretisch unendlich lange weiter schwingt. Diesen Fall bezeichnet man als Stabilitätsgrenzfall.

Im letzten und ungünstigsten Fall schließlich sind die Schwingungen, die in Folge einer Anregung auftreten, aufklingend. Diesen Fall bezeichnet man als Instabilität.

Um zu einer mathematisch präzisen Formulierung zu kommen, wird als Anregung ein Einheitssprung der Eingangsgröße definiert. Eine solche Anregung ist sicherlich ein guter Prüfstein für die Stabilität des Systems. Man kann davon ausgehen, dass ein System, das nach einer so heftigen Anregung wieder zur Ruhe kommt, sicherlich auch auf andere Anregungen nicht mit aufklingenden Schwingungen reagiert.

Die Tatsache, dass für die Amplitude des Eingangssprunges Eins gewählt wird, ist dabei keine Einschränkung. Da hier ausschließlich lineare Systeme betrachtet werden, würde sich bei einer anderen Sprungamplitude das gleiche Verhalten am Ausgang mit entsprechend veränderter Amplitude einstellen, was hinsichtlich der Stabilitätsaussage keinen Unterschied bedeuten würde.

Als noch härterer Prüfstein für die Stabilität wäre ein Sprung mit kurz darauf folgendem Rücksprung als Anregung denkbar. Aber auch dies würde aufgrund der Linearität des Systems zur gleichen Stabilitätsaussage führen. Denn wegen der Linearität wäre die Systemantwort auf einen Rücksprung die gleiche wie auf den Hinsprung, nur zeitlich versetzt und mit negativem Vorzeichen. Die Antworten auf den Hin- und auf den Rücksprung würden sich dann gemäß der Definition von Linearität am Ausgang des Systems überlagern, ohne sich gegenseitig zu beeinflussen. Kommt demnach das System nach einem einfachen Sprung als Anregung zur Ruhe, so kommt es auch nach Sprung und Rücksprung wieder zur Ruhe.

Tatsächlich lässt sich zeigen, dass die Forderung nach Einnahme eines Ruhezustandes nach einer sprungförmigen Anregung tatsächlich zu einer allgemein gültigen Stabilitätsaussage für das System führt.

Nach diesen Vorüberlegungen soll die Stabilität präzise definiert werden. Dabei existieren verschiedene Definitionsmöglichkeiten, von denen an dieser Stelle zwei betrachtet werden sollen. Eine dritte Definition nach dem russischen Mathematiker Ljapunov wird dann später noch erläutert. Die erste Definition entspricht den obigen Vorüberlegungen:

Definition 5.1 *Wenn die Sprungantwort eines Systems für $t \rightarrow \infty$ einem endlichen Wert zustrebt, so heißt das System stabil. Andernfalls heiße es instabil.*

Es lässt sich leicht nachvollziehen, dass PT_1 - und PT_2 -Glieder nach dieser Definition stabil sind und ein Integrator instabil ist. Auch ein Differenzierer ist nach dieser Definition stabil.

Die zweite Definition berücksichtigt, dass die Eingangsgröße eines Systems ständigen Schwankungen unterworfen sein kann:

Definition 5.2 *Ein lineares System heie stabil, wenn bei einer Eingangsgre mit beschrnkter Amplitude auch die Amplitude der Ausgangsgre beschrnkt ist. Dies ist die BIBO-Stabilitt (bounded input - bounded output).*

Auch nach dieser Definition sind PT_1 - und PT_2 -Glieder stabil und ein Integrator instabil. Ein Differenzierer ist nach dieser Definition jetzt aber instabil. Denn eine sprunghmige Anregung ist sicherlich eine Eingangsgre mit beschrnkter Amplitude (Sprunghhe), aber die Sprungantwort des Differenzierers weist im ersten Augenblick einen unendlich groen Wert auf, so dass das Ausgangssignal nicht beschrnkt ist.

Der grundstzliche Zusammenhang zwischen beiden Definitionen ist relativ einfach herzustellen: Eine sprunghmige Anregung ist durch den Wert ihrer Sprunghhe zuzglich einem beliebigen Wert ε beschrnkt. Wenn die Sprungantwort einem endlichen Wert zustrebt und vorher keine unendlich groen Werte annimmt, so lsst sich fr sie ebenfalls eine Schranke angeben. Sofern wir also ausschlieen, dass die Strecke Differenzierer enthlt, ist Definition 5.1 in Definition 5.2 enthalten und Definition 5.2 damit die umfassendere von beiden Definitionen.

Es wrde sich nun anbieten, Stabilitt immer im Sinne der BIBO-Stabilitt aufzufassen, weil dies die umfassendere Definition ist und somit keine weitere Unterscheidung notwendig wre. Andererseits werden die folgenden Betrachtungen wesentlich vereinfacht, wenn man Stabilitt im Sinne einer endlichen Sprungantwort auffasst. Zudem lsst sich zeigen, dass fr rein rationale bertragungsfunktionen, deren Nenner eine hhere Ordnung als der Zhler aufweist (und dies ist der in der Praxis blicherweise vorkommende Fall), ohnehin beide Definitionen quivalent sind. Fr die Stabilitt gilt daher von nun an immer Definition 5.1.

Aus dieser Definition lsst sich zunchst eine Bedingung fr die Impulsantwort ableiten. Fr die Sprungantwort eines Systems gilt im Frequenzbereich (vgl. (3.21))

$$y(s) = G(s) \frac{1}{s}$$

Fasst man den Faktor $1/s$ nicht als Laplace-Transformierte des Sprungsignals, sondern als Integration auf, so ergibt sich im Zeitbereich mit $y(0) = 0$

$$y(t) = \int_{\tau=0}^t g(\tau) d\tau \quad (5.3)$$

$y(t)$ kann nur dann einem endlichen Wert zustreben, wenn die Impulsantwort $g(t)$ gegen Null konvergiert.

Andersherum gilt aber nicht, dass ein System stabil ist, wenn seine Impulsantwort gegen Null konvergiert, wie man sofort anhand der Funktion $g(t) = \frac{1}{t+1}$ nachweisen kann. Diese Funktion konvergiert fr $t \rightarrow \infty$ sicherlich gegen Null. Fr die Sprungantwort ergibt sich jedoch

$$y(t) = \int_{\tau=0}^t \frac{1}{\tau+1} d\tau = \ln(1+t) \quad (5.4)$$

Sie konvergiert also nicht gegen einen endlichen Wert, und das System ist trotz einer verschwindenden Impulsantwort nicht stabil.

5.2 Stabilität einer Übertragungsfunktion

Um zu vermeiden, dass für den Nachweis der Stabilität die Sprungantwort des Systems explizit zu berechnen ist, bietet es sich an, die Übertragungsfunktion des Systems direkt zu betrachten und die Bedingungen abzuleiten, unter denen das System stabil ist. Dies ist nach den Überlegungen, die bereits zur Sprungantwort einer rationalen Übertragungsfunktion gemacht wurden, relativ einfach. Es gilt der folgende Satz:

Satz 5.3 *Ein Übertragungsglied mit einer rationalen Übertragungsfunktion ist genau dann stabil im Sinne von Definition 5.1, wenn alle Pole der Übertragungsfunktion einen negativen Realteil aufweisen.*

Für den Beweis soll zunächst gezeigt werden, dass bei negativem Realteil aller Pole der Endwert der Sprungantwort endlich ist. Entsprechend Definition 5.1 ist damit das System dann auch stabil.

Nach Gleichung (3.33) lautet die Sprungantwort eines rationalen Übertragungsgliedes:

$$y(t) = \sum_{\lambda=1}^i h_{\lambda}(t) e^{s_{\lambda} t} \quad (5.5)$$

Zu jedem n_{λ} -fachen Pol s_{λ} gehört ein Summand $h_{\lambda}(t) e^{s_{\lambda} t}$ mit einem Polynom $h_{\lambda}(t)$ vom Grade $n_{\lambda} - 1$. Weist der Pol einen negativen Realteil auf, so verschwindet dieser Summand mit wachsendem t , da die Exponentialfunktion schneller gegen Null konvergiert als das Polynom $h_{\lambda}(t)$ wachsen kann. Wenn alle Pole der Übertragungsfunktion einen negativen Realteil aufweisen, so verschwinden alle zugehörigen Summanden. Übrig bleibt nur der durch die Sprungfunktion verursachte Summand $h_i(t) e^{s_i t}$ mit dem einfachen Pol $s_i = 0$. Das Polynom $h_i(t)$ ist vom Grade $n_i - 1 = 0$, d.h. konstant, und die Exponentialfunktion reduziert sich ebenfalls auf eine Konstante. Damit bildet dieser Summand gerade den endlichen Endwert der Sprungfunktion, und das System ist stabil.

Der Beweis der Umkehrung, dass nämlich bei mindestens einem Pol mit nicht negativem Realteil ein instabiles System vorliegt, erfolgt analog. Wenn ein Pol s_{λ} mit positivem Realteil vorliegt, dann wächst die zugehörige Exponentialfunktion $e^{s_{\lambda} t}$ und damit auch der Summand $h_{\lambda}(t) e^{s_{\lambda} t}$ unabhängig vom Polynom $h_{\lambda}(t)$ über alle Maßen, und die Sprungantwort $y(t)$ konvergiert

nicht gegen einen endlichen Endwert. Falls mindestens ein Pol $s_\lambda = 0$ vorliegt, so ist das zur Polstelle Null gehörende Polynom mindestens vom Grad 1, da ja auch der Eingangssprung schon einen Pol bei $s_\lambda = 0$ verursacht. Das Polynom und damit auch der zugehörige Summand $h_\lambda(t)e^{s_\lambda t}$ wächst über alle Maßen, und auch hier konvergiert die Sprungantwort nicht mehr gegen einen endlichen Endwert.

Im allgemeinen ist neben der Tatsache der Stabilität auch die Form der Einschwingvorgänge nach einer äußeren Anregung interessant. Weist die Strecke unter anderem ein konjugiert komplexes Polpaar $s_\lambda, \bar{s}_\lambda$ auf, so ist nach Gleichung (2.48) das Verhältnis $|\operatorname{Re}(s_\lambda)|/\sqrt{\operatorname{Re}(s_\lambda)^2 + \operatorname{Im}(s_\lambda)^2}$ gerade gleich der Dämpfung D und somit für die Form des zu diesem Polpaar gehörenden Einschwingvorgangs verantwortlich. Man wird daher in der Praxis nicht nur darauf achten, dass die Pole eines Systems einen negativen Realteil aufweisen, sondern auch darauf, dass die Dämpfung D einen ausreichend großen Wert hat, d.h. dass ein konjugiert komplexes Polpaar ausreichend weit von der imaginären Achse entfernt liegt.

5.3 Stabilität eines Regelkreises

Das System, dessen Stabilität beurteilt werden soll, ist in den meisten Fällen ein geschlossener Regelkreis, wie er in Abb. 1.2 dargestellt ist. Eine vereinfachte Struktur gibt Abb. 5.1 wieder. Das Regelglied habe die Übertragungsfunktion $K(s)$, die Strecke ist durch $G(s)$ und das Messglied durch $M(s)$ gegeben. Um die weiteren Herleitungen nicht unnötig zu erschweren, wird aber $M(s) = 1$ gesetzt, d.h. das dynamische Verhalten des Messgliedes wird vernachlässigt. Es ist aber kein Problem, im Einzelfall ein Messglied zu berücksichtigen.

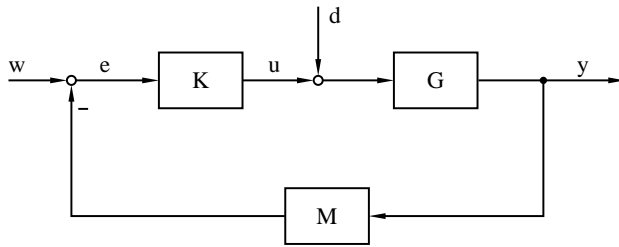


Abb. 5.1. Regelkreis

Die Störgrößen, die prinzipiell an beliebigen Stellen des Regelkreises angreifen können, werden zu einer einzigen Störgröße d zusammengefasst, die direkt am Eingang der Strecke aufgeschaltet ist. Auch diese Maßnahme vereinfacht die Theorie, ohne dass die Verhältnisse aus Sicht des Reglers einfacher werden, als sie in der Praxis sind. Der Angriffspunkt am Eingang der Strecke ist nämlich die für eine Regelung denkbar ungünstigste Stelle, da die

Störgröße zunächst unbehelligt auf die Strecke einwirken kann, während der Regler erst nach einer Veränderung der Ausgangsgröße Gegenmaßnahmen einleiten kann.

Zur Anwendung der Stabilitätskriterien auf dieses System ist zunächst die Übertragungsfunktion aufzustellen, die das Übertragungsverhalten von der Eingangsgröße w zur Ausgangsgröße y beschreibt. Dies ist die Übertragungsfunktion des geschlossenen Kreises und wird auch als *Führungs-Übertragungsfunktion* bezeichnet. Zur Berechnung wird die Störgröße d zunächst gleich Null gesetzt. Im Frequenzbereich gilt:

$$\begin{aligned} y(s) &= G(s)u(s) = G(s)K(s)(w(s) - y(s)) \\ T(s) &= \frac{y(s)}{w(s)} = \frac{G(s)K(s)}{G(s)K(s) + 1} \end{aligned} \quad (5.6)$$

Analog dazu lässt sich auch eine *Stör-Übertragungsfunktion* berechnen, die das Übertragungsverhalten des Systems von der Störgröße d zur Ausgangsgröße y beschreibt:

$$S(s) = \frac{y(s)}{d(s)} = \frac{G(s)}{G(s)K(s) + 1} \quad (5.7)$$

Der Term $G(s)K(s)$ hat eine besondere Bedeutung: Entfernt man nämlich die Rückführung, so bildet er gerade die Übertragungsfunktion des dann vorliegenden, offenen Kreises. Er wird auch als *Kreisübertragungsfunktion* bezeichnet. Der Verstärkungsfaktor V dieser Funktion (vgl. Gleichung (5.1)) wird als *Kreisverstärkung* bezeichnet.

Man sieht, dass sowohl Führungs- als auch Stör-Übertragungsfunktion denselben Nenner $G(s)K(s) + 1$ aufweisen. Andererseits ist es aber nach Satz 5.3 gerade der Nenner der Übertragungsfunktion, der für die Stabilität verantwortlich ist. Daran lässt sich erkennen, dass für die Stabilität eines Systems nur die Kreisübertragungsfunktion, nicht aber der Angriffspunkt der Eingangsgröße relevant ist. Eine Stabilitätsuntersuchung kann sich daher auf die Untersuchung von $G(s)K(s) + 1$ beschränken. Da Zähler und Nenner der beiden Übertragungsfunktionen $T(s)$ und $S(s)$ offensichtlich nie die gleichen Nullstellen haben können, ist auch ein Kürzen von Pol- gegen Nullstellen bei beiden Funktionen nicht möglich. Daher entsprechen die Nullstellen von $G(s)K(s) + 1$ gerade den Polen der beiden Übertragungsfunktionen, und es ergibt sich als direkte Folgerung aus Satz 5.3:

Satz 5.4 *Ein geschlossener Kreis mit der Kreisübertragungsfunktion $G(s)K(s)$ ist genau dann stabil, wenn alle Lösungen der charakteristischen Gleichung*

$$G(s)K(s) + 1 = 0 \quad (5.8)$$

einen negativen Realteil aufweisen.

Eine Berechnung dieser Nullstellen ist aber auf analytischem Wege nicht mehr möglich, wenn die Ordnung der Strecke größer als zwei ist oder die

Kreisübertragungsfunktion eine Exponentialfunktion enthält. Die exakte Lage der Nullstellen muss aber für eine Stabilitätsuntersuchung auch gar nicht bekannt sein. Wichtig ist lediglich die Tatsache, ob sie einen positiven oder negativen Realteil aufweisen. Aus diesem Grund sind in der Vergangenheit Stabilitätskriterien entwickelt worden, mit denen ohne aufwändige Rechnung genau dies überprüft werden kann.

5.4 Kriterium von Cremer, Leonhard und Michailow

Als erstes soll auf ein Kriterium eingegangen werden, das von Cremer [10], Leonhard [32] und Michailow [40] in den Jahren 1938 bis 1947 unabhängig voneinander herausgefunden wurde und gewöhnlich auch nach diesen Forschern benannt wird. Gegenstand der Betrachtungen ist die Phasendrehung der Ortskurve eines Polynoms in Abhängigkeit von der Lage seiner Nullstellen. Gegeben sei ein Polynom der Form

$$P(s) = s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0 = \prod_{\nu=1}^n (s - s_\nu) \quad (5.9)$$

Mit $s = j\omega$ wird daraus

$$\begin{aligned} P(j\omega) &= \prod_{\nu=1}^n (j\omega - s_\nu) = \prod_{\nu=1}^n (|j\omega - s_\nu| e^{j\varphi_\nu(\omega)}) \\ &= \left(\prod_{\nu=1}^n |j\omega - s_\nu| \right) e^{j \sum_{\nu=1}^n \varphi_\nu(\omega)} = |P(j\omega)| e^{j\varphi(\omega)} \end{aligned} \quad (5.10)$$

Der Frequenzgang $P(j\omega)$ ist also das Produkt der Vektoren $(j\omega - s_\nu)$, wobei die Phase $\varphi(\omega)$ gerade die Summe der Winkel $\varphi_\nu(\omega)$ dieser Vektoren ist. Abb. 5.2 zeigt die Verhältnisse bei einem konjugiert komplexen Nullstellenpaar mit negativem und einer einzelnen Nullstelle mit positivem Realteil.

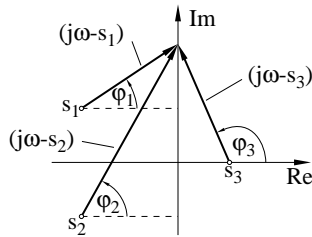


Abb. 5.2. Zum Kriterium von Cremer-Leonhard-Michailow

Durchläuft der Parameter ω das Intervall $(-\infty, \infty)$, so wandert der Endpunkt der Vektoren $(j\omega - s_\nu)$ einmal längs der imaginären Achse in positiver

Richtung. Für Nullstellen mit negativem Realteil durchläuft der zugehörige Winkel φ_ν das Intervall von $-\frac{\pi}{2}$ bis $+\frac{\pi}{2}$, für Nullstellen mit positivem Realteil das Intervall von $+\frac{3\pi}{2}$ bis $+\frac{\pi}{2}$. Für Nullstellen auf der imaginären Achse hat der zugehörige Winkel φ_ν zunächst den Wert $-\frac{\pi}{2}$ und springt dann bei $j\omega = s_\nu$ auf den Wert $+\frac{\pi}{2}$.

Nun soll die Phasendrehung des Frequenzganges $P(j\omega)$ betrachtet werden, also der gesamte Verlauf des Winkels $\varphi(\omega)$. Dieser Winkel ist aber gerade die Summe der Winkel $\varphi_\nu(\omega)$. Daher trägt jede Nullstelle mit negativem Realteil zur Phasendrehung des Frequenzganges den Winkel $+\pi$ bei, und jede Nullstelle mit positivem Realteil den Winkel $-\pi$. Für Nullstellen auf der imaginären Achse lässt sich wegen des unstetigen Phasenverlaufes keine Aussage machen. Ob solche Nullstellen vorliegen, kann man aber sofort anhand der Ortskurve des Polynoms $P(s)$ erkennen. Wenn das Polynom eine rein imaginäre Nullstelle $s = s_\nu$ hat, so muss die Ortskurve für die Frequenz $\omega = |s_\nu|$ durch den Ursprung gehen. Damit ergibt sich der folgende Satz:

Satz 5.5 *Ein Polynom $P(s)$ vom Grad n mit reellen Koeffizienten weist genau dann nur Nullstellen mit negativem Realteil auf, wenn seine Ortskurve nicht durch den Ursprung der komplexen Ebene geht und die Phasendrehung $\Delta\varphi$ des Frequenzganges für $-\infty < \omega < +\infty$ gerade $n\pi$ beträgt. Durchläuft ω nur den Bereich $0 \leq \omega < +\infty$, so beträgt die notwendige Phasendrehung $\frac{n}{2}\pi$.*

Die Tatsache, dass für $0 \leq \omega < +\infty$ die notwendige Phasendrehung nur noch $\frac{n}{2}\pi$ und damit gerade die Hälfte beträgt, ist leicht zu beweisen:

Für Nullstellen auf der reellen Achse ist es offensichtlich, dass ihr Beitrag zur Phasendrehung nur noch halb so groß ist, wenn ω nur die halbe imaginäre Achse von 0 bis ∞ durchläuft. Interessanter sind die Nullstellen, deren Imaginärteil von Null verschieden ist. Diese können aber wegen der reellen Koeffizienten des Polynoms immer nur als komplex konjugiertes Polpaar auftreten. Abb. 5.3 zeigt ein solches Polpaar, $s_1 = \bar{s}_2$ und $\alpha_1 = -\alpha_2$. Für $-\infty < \omega < +\infty$ ist der Beitrag dieses Polpaars zur Phasendrehung 2π . Für $0 \leq \omega < +\infty$ beträgt der Beitrag von s_1 gerade $\frac{\pi}{2} + |\alpha_1|$, für s_2 ist er $\frac{\pi}{2} - |\alpha_1|$. Insgesamt ist der Beitrag dieses Polpaars damit π , auch hier hat sich also die Phasendrehung auf die Hälfte reduziert.

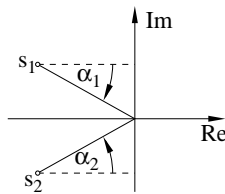


Abb. 5.3. Zur Phasendrehung bei einem konjugiert komplexen Polpaar

Abb. 5.4 zeigt als Beispiel zwei Ortskurven von Polynomen fünfter Ordnung. Da die Ortskurve die graphische Repräsentation des Frequenzganges ist, lässt sich die jeweilige Phasendrehung auch direkt an der Ortskurve ablesen. Dazu ist ein Vektor vom Ursprung zur Ortskurve einzuzeichnen, wie dies für Kurve 1 zu sehen ist. Dann hat man die Anzahl der Umdrehungen um den Ursprung zu ermitteln, die dieser Vektor für den gesamten Verlauf der Ortskurve vollzieht. Der so gewonnene Winkel entspricht der gesuchten Phasendrehung des Frequenzganges.

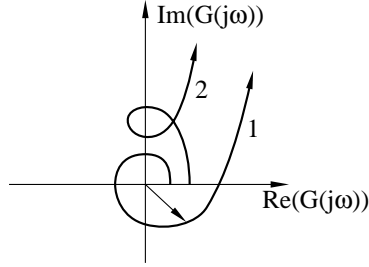


Abb. 5.4. Ortskurven von Polynomen 5. Ordnung

Kurve 1 weist eine Gesamt-Phasendrehung $\Delta\varphi$ von $5\frac{\pi}{2}$ auf, das zugehörige Polynom hat also nur Nullstellen mit negativem Realteil. Die Phasendrehung von Kurve 2 beträgt dagegen nur $\frac{\pi}{2}$, obwohl Anfangs- und Endwinkel denen von Kurve 1 entsprechen. Das dieser Kurve zu Grunde liegende Polynom besitzt also Nullstellen, deren Realteil nicht negativ ist.

5.5 Nyquist-Kriterium

Aus den Sätzen 5.4 und 5.5 lässt sich nun ein sehr elegantes Stabilitätskriterium ableiten, das *Nyquist-Kriterium*. Beim Nyquist-Kriterium [46] wird direkt die Ortskurve der Kreisübertragungsfunktion $G(s)K(s)$ betrachtet, die auch einfach gemessen werden kann, falls die Funktion nicht in analytischer Form vorliegt. Es sei

$$\frac{Z_k(s)}{N_k(s)} = G(s)K(s)$$

mit zwei teilerfremden Polynomen $Z_k(s)$ und $N_k(s)$. Außerdem sei der Grad m von $Z_k(s)$ höchstens gleich dem Grad n von $N_k(s)$, was aber für physikalisch realisierbare Systeme immer erfüllt ist. Wegen

$$T(s) = \frac{G(s)K(s)}{1 + G(s)K(s)} = \frac{\frac{Z_k(s)}{N_k(s)}}{1 + \frac{Z_k(s)}{N_k(s)}} = \frac{Z_k(s)}{Z_k(s) + N_k(s)} \quad (5.11)$$

ist $N_g(s) = Z_k(s) + N_k(s)$ der Nenner der Übertragungsfunktion des geschlossenen Kreises $T(s)$ mit dem Grad

$$\text{Grad}(N_g(s)) = \max(\text{Grad}(Z_k(s)), \text{Grad}(N_k(s))) = \max(m, n) = n \quad (5.12)$$

Damit gilt:

$$1 + G(s)K(s) = 1 + \frac{Z_k(s)}{N_k(s)} = \frac{N_g(s)}{N_k(s)} \quad (5.13)$$

Der Term $1 + G(s)K(s)$ lässt sich also auch als Quotient aus dem Nenner der Übertragungsfunktion des geschlossenen Kreises und dem Nenner der Kreisübertragungsfunktion ausdrücken.

Die Phase des Frequenzganges $1 + G(j\omega)K(j\omega)$ ist damit die Differenz der Phasen der beiden Nennerpolynome:

$$\varphi_{1+GK}(\omega) = \varphi_{N_g}(\omega) - \varphi_{N_k}(\omega) \quad (5.14)$$

Und der gleiche Zusammenhang muss auch für die Phasendrehung gelten:

$$\Delta\varphi_{1+GK} = \Delta\varphi_{N_g} - \Delta\varphi_{N_k} \quad (5.15)$$

An dieser Stelle sei noch einmal auf den Unterschied zwischen Phase und Phasendrehung eingegangen: $(1+GK)(\omega)$ ist eine komplexe Zahl, die auch als Vektor in der komplexen Ebene aufgefasst werden kann. Der Vektor ist von der Frequenz ω abhängig. Je nach Frequenz wird sich also ein anderer Vektor ergeben, und die Endpunkte aller Vektoren beschreiben in Abhängigkeit von ω eine Kurve in der komplexen Ebene, die Ortskurve. Die Phase $\varphi_{1+GK}(\omega)$ ist der Winkel des Vektors $(1+GK)$ für eine bestimmte Frequenz ω . Die Phasendrehung $\Delta\varphi_{1+GK}$ beschreibt dagegen den Drehwinkel, den ein Vektor ausführen würde, dessen Anfangspunkt im Ursprung der komplexen Ebene liegt und dessen Spitze die Ortskurve entlang läuft.

Zur Berechnung der Phasendrehungen $\Delta\varphi_{N_g}$ und $\Delta\varphi_{N_k}$ muss nach Satz 5.5 die Verteilung der Nullstellen der Polynome $N_g(s)$ und $N_k(s)$ bekannt sein. Die Nullstellen von $N_g(s)$ sind die Polstellen des geschlossenen Kreises. Von diesen n Polstellen mögen r_g in der rechten Hälfte der s -Ebene, i_g auf der imaginären Achse und $n - r_g - i_g$ in der linken Hälfte liegen. Entsprechend sind die Nullstellen von $N_k(s)$ gerade die Polstellen der Kreisübertragungsfunktion. Von diesen ebenfalls n Polstellen mögen r_k rechts von der imaginären Achse, i_k auf und $n - r_k - i_k$ links von ihr liegen.

Da sowohl i_g als auch i_k von Null verschieden sein können, weisen die Phasengänge $\varphi_{N_g}(\omega)$ und $\varphi_{N_k}(\omega)$ einen möglicherweise unstetigen Verlauf auf, wie schon in der Herleitung von Satz 5.5 erklärt wurde. Um Schwierigkeiten zu vermeiden, soll nur der stetige Anteil der Phasendrehungen betrachtet werden. Nach Satz 5.5 steuert zur Phasendrehung einer Ortskurve mit $0 < \omega < \infty$ jede Nullstelle mit negativem Realteil die Phasendrehung $\frac{\pi}{2}$, jede mit positivem Realteil die Phasendrehung $-\frac{\pi}{2}$ bei:

$$\begin{aligned} \Delta\varphi_{N_g, \text{stetig}} &= [(n - r_g - i_g) - r_g] \frac{\pi}{2} \\ \Delta\varphi_{N_k, \text{stetig}} &= [(n - r_k - i_k) - r_k] \frac{\pi}{2} \end{aligned} \quad (5.16)$$

Für den stetigen Anteil der Phasendrehung von $1 + G(j\omega)K(j\omega)$ ergibt sich wegen (5.15):

$$\begin{aligned}\Delta\varphi_{1+GK,\text{stetig}} &= [(n - r_g - i_g) - r_g] \frac{\pi}{2} - [(n - r_k - i_k) - r_k] \frac{\pi}{2} \\ &= [2(r_k - r_g) + i_k - i_g] \frac{\pi}{2}\end{aligned}\quad (5.17)$$

Fordert man nun Stabilität des geschlossenen Kreises, so darf dieser nur Polstellen mit negativem Realteil aufweisen. Es muss $r_g = i_g = 0$ gelten und damit

$$\Delta\varphi_{1+GK,\text{stetig}} = r_k\pi + i_k \frac{\pi}{2} \quad (5.18)$$

Ob die Kreisübertragungsfunktion dabei stabil oder instabil ist, spielt keine Rolle. Lediglich die Anzahl ihrer Pole auf und rechts neben der imaginären Achse muss bekannt sein.

Allerdings ist in dieser Formel nur der stetige Anteil der Phasendrehung behandelt worden. Nullstellen von $N_g(s)$ auf der imaginären Achse verursachen aber unstetige Phasenänderungen. Demnach kann durch eine Analyse der stetigen Phasendrehung nach Gleichung (5.18) zwar ausgeschlossen werden, dass $N_g(s)$ Nullstellen mit positivem Realteil besitzt, nicht aber, dass rein imaginäre Nullstellen auftreten. Wegen Gleichung (5.13) entsprechen die Nullstellen von $N_g(s)$ den Nullstellen von $G(s)K(s) + 1$. Eine rein imaginäre Nullstelle von $N_g(s)$ hat demnach zur Folge, dass auch der Frequenzgang $G(j\omega)K(j\omega) + 1$, dessen Argument $j\omega$ rein imaginär ist, eine Nullstelle bei der entsprechenden Frequenz aufweist. Das bedeutet aber wiederum, dass die Ortskurve von $G(j\omega)K(j\omega) + 1$ durch den Ursprung geht. Damit ergibt sich, dass für einen stabilen Regelkreis nicht nur Gleichung (5.18) gelten muss, sondern die Ortskurve $G(j\omega)K(j\omega) + 1$ auch nicht durch den Ursprung laufen darf.

Statt der Ortskurve $1 + G(j\omega)K(j\omega)$ kann man auch die - direkt messbare - Ortskurve der Kreisübertragungsfunktion $G(j\omega)K(j\omega)$ betrachten. Sämtliche Überlegungen beziehen sich dann nicht mehr auf den Ursprung der komplexen Ebene, sondern auf den Punkt -1 , wie aus Abb. 5.5 ersichtlich ist. Dies führt zum folgenden Satz.

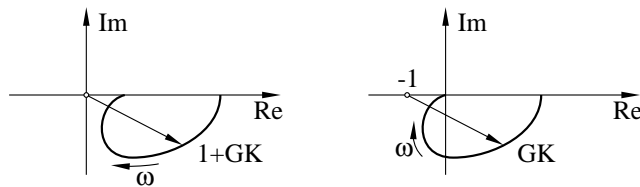


Abb. 5.5. Übergang von der Kurve $(1+GK)$ auf (GK)

Satz 5.6 (Nyquist-Kriterium) *Ein geschlossener Kreis ist genau dann stabil, wenn die stetige Phasendrehung der Ortskurve seiner Kreisübertragungs-*

funktion $G(s)K(s)$ um den Punkt -1 gerade

$$\Delta\varphi_{GK, \text{stetig}} = r_k\pi + i_k\frac{\pi}{2} \quad (5.19)$$

beträgt und die Kurve nicht durch den Punkt -1 läuft. Dabei ist i_k die Anzahl der Polstellen der Kreisübertragungsfunktion auf der imaginären Achse der s -Ebene und r_k die Anzahl der Polstellen rechts von ihr.

Wichtig für die Anwendung des Nyquist-Kriteriums ist, dass r_k und i_k bekannt sein müssen. Weiterhin sei angemerkt, dass das Nyquist-Kriterium auch für Laufzeiten in der Kreisübertragungsfunktion gilt. Auf den Beweis hierzu soll aber verzichtet werden.

Da das Nyquist-Kriterium nicht ganz einfach zu verstehen ist, soll es nachfolgend anhand einiger Beispiele erläutert werden.

Das erste Beispiel sei ein rückgekoppelter Integrator (Abb. 5.6). Die Kreisübertragungsfunktion besteht hier nur aus dem Integrator und damit aus einem instabilen Übertragungsglied: $GK(s) = \frac{1}{s}$. Für die Anzahl ihrer Polstellen auf und rechts von der imaginären Achse gilt bei einem Integrator: $i_k = 1$ und $r_k = 0$. Die erforderliche Phasendrehung der Ortskurve um den Punkt -1 beträgt wegen (5.19) damit $\Delta\varphi = +\frac{\pi}{2}$. Nun ist anhand der Ortskurve zu prüfen, wie groß die Phasendrehung tatsächlich ist.

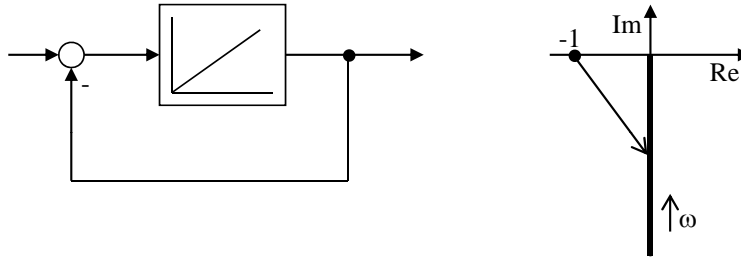


Abb. 5.6. Anwendung des Nyquist-Kriterium auf ein PT_1 -Glieder

Die ebenfalls in Abbildung 5.6 eingezeichnete Ortskurve der Kreisübertragungsfunktion ist die des Integrators und verläuft auf der negativ imaginären Achse. Zur Ermittlung der Phasendrehung um den Punkt -1 hat man sich nun einen Vektor vorzustellen, der im Punkt -1 beginnt und zu den Punkten der Ortskurve führt. Wenn die Spitze des Vektors einmal der gesamten Ortskurve folgt, wird sich der Vektor um einen gewissen Winkel drehen. Dieser ist gerade die gesuchte Phasendrehung.

Im vorliegenden Beispiel beginnt die Ortskurve im negativ Unendlichen auf der imaginären Achse. Der Vektor zeigt damit praktisch senkrecht nach unten. Anschließend wird die Ortskurve durchlaufen, und der Vektor dreht sich allmählich nach rechts, bis er im Endpunkt der Ortskurve genau nach rechts zeigt. Die Phasendrehung beträgt $\Delta\varphi = +\frac{\pi}{2}$ und entspricht damit der

laut Nyquist-Kriterium erforderlichen Phasendrehung. Zudem durchläuft die Ortskurve auch nicht den Punkt -1 . Somit ist das Nyquist-Kriterium erfüllt, und der geschlossene Kreis ist stabil. Dies entspricht auch unserem Vorwissen, denn ein rückgekoppelter Integrator ist doch gerade ein PT_1 -Glieder. Und dieses ist natürlich ein stabiles Übertragungsglied.

Das nächste Beispiel bezieht sich auf das in Kapitel 4.3.5 schon behandelte IT_1 -Glieder. Die Kreisübertragungsfunktion besteht jetzt aus einem Integrator und einem PT_1 -Glieder. Wegen des Integrators ist auch dieses Übertragungsglied instabil. Seine Kreisübertragungsfunktion lautet:

$$GK(s) = \frac{1}{T_1 s(T_2 s + 1)} \quad (5.20)$$

Für die Anzahl der Polstellen auf und rechts von der imaginären Achse gilt: $i_k = 1$ und $r_k = 0$. Die erforderliche Phasendrehung um den Punkt -1 beträgt damit auch hier $\Delta\varphi = +\frac{\pi}{2}$. Dies ist nun wiederum anhand der Ortskurve zu verifizieren.

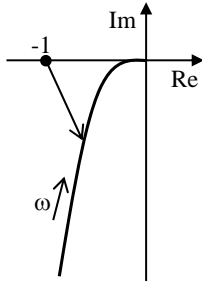


Abb. 5.7. Anwendung des Nyquist-Kriterium auf ein IT_1 -Glieder

Die Ortskurve beginnt im negativ Unendlichen links neben der imaginären Achse (Abbildung 5.7). Man sieht, dass der Zeiger vom Punkt -1 zur Ortskurve zunächst nach unten zeigt und sich dann wie im vorhergehenden Beispiel nach rechts dreht. Als Phasendrehung ergibt sich wieder $+\frac{\pi}{2}$, und auch hier entspricht die tatsächliche Phasendrehung derjenigen, die laut Nyquist-Kriterium erforderlich ist. Darüber hinaus wird der Punkt -1 auch nicht von der Ortskurve durchlaufen. Ein geschlossener Kreis mit einem Integrator und einem PT_1 -Glieder als Kreisübertragungsfunktion wäre daher stabil. Würde man die Kreisverstärkung V verändern, so würde die Ortskurve gestreckt oder gestaucht. Ihr prinzipieller Verlauf bliebe aber erhalten und damit auch die Phasendrehung, d.h. auch bei einer Veränderung von V bleibt das System stabil. Dies gilt nicht für alle Systeme, wie die nächsten Beispiele zeigen werden.

Im folgenden Beispiel soll die Kreisübertragungsfunktion

$$G(s)K(s) = V \frac{(T_1 s + 1)^2}{(T_2 s + 1)(T_i s)^2} \quad T_i > T_2 > T_1 \quad (5.21)$$

untersucht werden. Wegen $i_k = 2$ beträgt die für Stabilität erforderliche Phasendrehung bezüglich -1 gerade π . Für die Untersuchung der Ortskurve ist zunächst das Bode-Diagramm zu zeichnen. Abbildung 5.8 zeigt die prinzipiellen Betrags- und Phasenverläufe der einzelnen Teil-Übertragungsfunktionen.

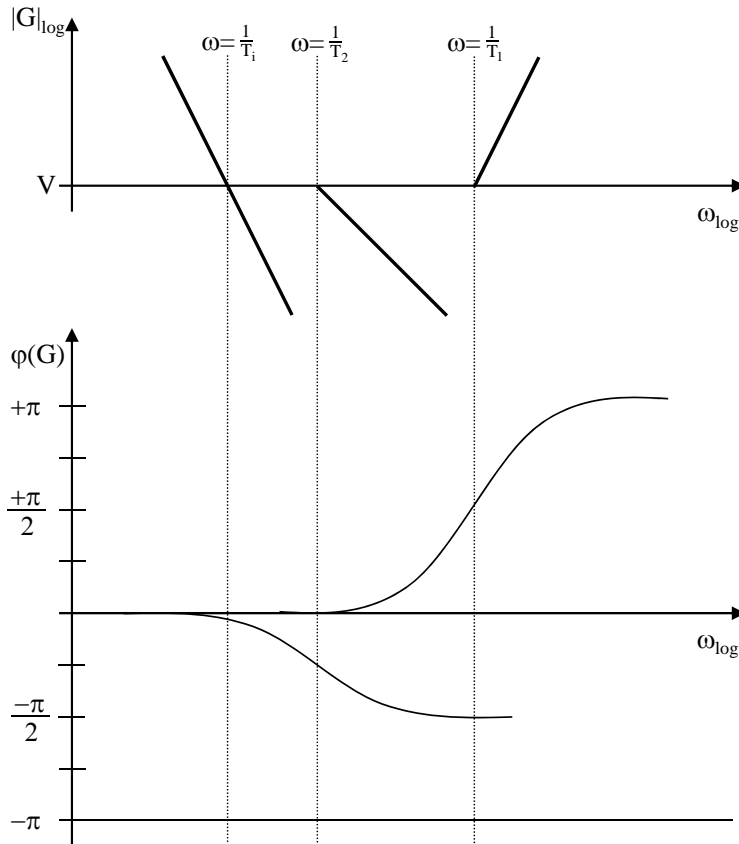


Abb. 5.8. Bode-Diagramme der Teil-Übertragungsfunktionen

Der doppelte Integrator hat einen linear abnehmenden Betragsverlauf mit der Steigung -2 , der die ω -Achse bei $\frac{1}{T_i}$ kreuzt, und eine konstante Phase von $-2\frac{\pi}{2} = -\pi$. Die Polstelle bei $-\frac{1}{T_2}$ weist einen bei $\frac{1}{T_2}$ nach unten abknickenden Betragsverlauf und einen von Null auf $-\frac{\pi}{2}$ abnehmenden Phasenverlauf auf. Die doppelte Nullstelle wiederum hat einen bei $\frac{1}{T_1}$ nach oben abknickenden Betragsverlauf mit doppelter Steigung, und der Phasenverlauf nimmt von Null auf $+2\frac{\pi}{2} = +\pi$ zu.

Der Betragsverlauf der gesamten Übertragungsfunktion hat damit immer eine negative Steigung und nimmt von ∞ auf Null ab. Der Phasenverlauf als Überlagerung der Teil-Phasenverläufe ist in Abbildung 5.9 gezeigt. Zunächst

wirkt sich nur die Phase des doppelten Integrators aus, so dass der Phasenverlauf bei $-\pi$ beginnt. Mit Annäherung an die Frequenz $\frac{1}{T_2}$ beginnt sich die Polstelle auszuwirken, und die Phase nimmt ab. Bevor sie den Wert $-\frac{3}{2}\pi$ erreicht, beginnt sich dann die doppelte Nullstelle auszuwirken, so dass die Phase wieder zunimmt und schließlich gegen den Endwert $-\frac{\pi}{2}$ konvergiert.

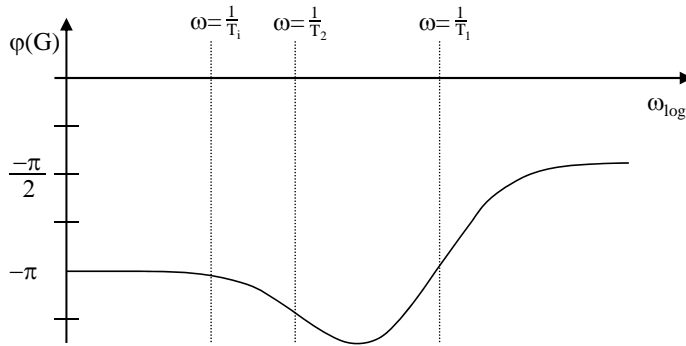


Abb. 5.9. Phasenverlauf der gesamten Übertragungsfunktion

Die zugehörige Ortskurve zeigt Abbildung 5.10. Sie beginnt unter einem Winkel von $-\pi$ im Unendlichen. Der Winkel nimmt zunächst ab, so dass die Ortskurve im zweiten Quadranten verläuft. Mit zunehmender Frequenz steigt die Phase, so dass die Kurve die negativ reelle Achse schneidet und schließlich unter einem Winkel von $-\frac{\pi}{2}$ in den Ursprung hinein läuft.

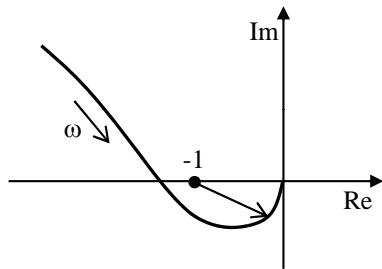


Abb. 5.10. Ortskurve der gesamten Übertragungsfunktion

Da die Ortskurve aus dem negativ reell Unendlichen kommt, zeigt der Zeiger vom Punkt -1 zur Ortskurve zunächst nach links und dreht sich dann im mathematisch positiven Sinn nach rechts, was einer Phasendrehung von $+\pi$ entspricht. Dieser Wert entspricht gerade der Vorgabe des Nyquist-Kriteriums, so dass ein geschlossener Kreis mit einer solchen Kreisübertragungsfunktion stabil wäre.

Verkleinert man jetzt aber V , so wird die Kurve gestaucht, bis der Schnittpunkt mit der reellen Achse rechts vom Punkt -1 liegt. Für die Phasendre-

hung ergibt sich nun folgendes: Der Vektor vom Punkt -1 zur Ortskurve zeigt wieder zunächst nach links, dreht sich jetzt aber im mathematisch negativen Sinn nach rechts. Die Phasendrehung beträgt dann statt $+\pi$ jetzt $-\pi$, und der geschlossene Kreis wäre instabil.

Mit diesen Beispielen wird klar, dass man das Nyquist-Kriterium auch in einer vereinfachten, anschaulicheren Form formulieren kann:

Satz 5.7 *Hat die Kreisübertragungsfunktion $G(s)K(s)$ keine Pole in der rechten Halbebene, so ist der geschlossene Kreis genau dann stabil, wenn die Ortskurve der Kreisübertragungsfunktion den Punkt -1 von sich aus gesehen rechts passiert.*

Mit Hilfe der Ortskurven lassen sich auch Aussagen über die Dämpfung des geschlossenen Regelkreises machen. Zunächst gilt: Das Einschwingverhalten eines Systems wird durch die Lage der Pole seiner Übertragungsfunktion bestimmt, und je weiter ein konjugiert komplexes Polpaar von der imaginären Achse entfernt ist, desto größer ist die Dämpfung der zugehörigen Schwingung (vgl. Gl. (2.48)). Weiterhin sind die Pole des geschlossenen Kreises gerade die Nullstellen der Gleichung $G(s)K(s) + 1 = 0$. Alle Pole werden demnach durch die Abbildung $G(s)K(s)$ in den Punkt -1 abgebildet. Dagegen wird ein Punkt der imaginären Achse $s = j\omega$ durch die Abbildung $G(s)K(s)$ auf $G(j\omega)K(j\omega)$ abgebildet. Dies ist aber gerade die Ortskurve der Kreisübertragungsfunktion. Die imaginäre Achse der komplexen Ebene wird also auf die Ortskurve $G(j\omega)K(j\omega)$ abgebildet. Wenn aber -1 das Abbild aller Pole ist und die Ortskurve das Abbild der imaginären Achse, so ist bei Stetigkeit der Abbildung der Abstand der Ortskurve vom Punkt -1 auch ein Maß für den Abstand der Pole von der imaginären Achse und somit für die Dämpfung des geschlossenen Kreises. Es ergibt sich der folgende Satz:

Satz 5.8 *Der Abstand der Ortskurve $G(j\omega)K(j\omega)$ vom kritischen Punkt -1 ist ein Maß für die Dämpfung des geschlossenen Kreises.*

Zwei weitere Stabilitätskriterien sollen hier nur kurz erwähnt werden, es sind die Kriterien von Hurwitz [18] und Routh [54]. Beide beziehen sich auf die Koeffizienten des Nenners der Übertragungsfunktion und sind gewissermaßen numerische Kriterien. Derartige Kriterien sind aber durch die heutige Möglichkeit, Nullstellen von Polynomen vom Computer numerisch berechnen zu lassen, praktisch nicht mehr relevant.

5.6 Systeme mit nicht-minimaler Phase

Nachdem mit dem Nyquistkriterium die Grundlage zur Stabilitätsanalyse eines Eingrößensystems bekannt ist, könnte man direkt auf den Entwurf von Reglern zu sprechen kommen. Vorher soll jedoch noch eine spezielle Klasse von Systemen vorgestellt werden, die einige sehr unangenehme Eigenschaften

aufweist, über deren Existenz man sich im Klaren sein sollte, bevor man eine Regelung entwirft. Es handelt sich um die Systeme mit nicht-minimaler Phase.

Um derartige Systeme besser definieren zu können, soll zunächst ein spezieller Typ von Übertragungsfunktionen eingeführt werden, der *Allpass*. Dieser ist dadurch charakterisiert, dass der Betrag seines Frequenzganges für alle Frequenzen gleich ist. Keine Frequenz wird unterdrückt oder gefiltert, so dass alle Frequenzen das Übertragungsglied gleich gut passieren können (daher der Name).

Ein Laufzeitglied wäre damit ein besonders einfaches Beispiel für einen Allpass, doch werden normalerweise nur rein rationale, stabile Übertragungsfunktionen der Form

$$G(s) = \frac{b_m \prod_{\mu=1}^m (s - n_{\mu})}{a_n \prod_{\nu=1}^n (s - p_{\nu})} \quad (5.22)$$

deren Frequenzgang den konstanten Betrag Eins hat, als Allpass bezeichnet. Im Folgenden sollen nun die Eigenschaften dieser Übertragungsglieder näher betrachtet werden. Schreibt man den Frequenzgang in der Form

$$G(j\omega) = \frac{b_m \prod_{\mu=1}^m (j\omega - n_{\mu})}{a_n \prod_{\nu=1}^n (j\omega - p_{\nu})} \quad (5.23)$$

so lassen sich Zähler- und Nennerpolynom als Produkt von komplexen Vektoren auffassen. Ein Vektor $(j\omega - n_{\mu})$ geht dabei in der komplexen Ebene vom Punkt n_{μ} zum Punkt $j\omega$, und das gleiche gilt analog für die Vektoren $(j\omega - p_{\nu})$ im Nenner. Betrag und Phase der Vektoren ändern sich dabei mit der Frequenz ω (vgl. Abb. 5.11).

Offensichtlich kann der Betrag des Frequenzganges nur dann für alle Frequenzen konstant Eins sein, wenn Zähler- und Nennerpolynom vom selben Grad sind und zu jedem Vektor im Nenner ein Vektor gleichen Betrages im Zähler existiert. Setzt man teilerfremde Polynome voraus, so ist dies nur erfüllt für $n_{\nu} = -\bar{p}_{\nu}$, wenn also Pole und Nullstellen spiegelbildlich zur imaginären Achse angeordnet sind. Wegen der vorausgesetzten Stabilität der Übertragungsfunktion müssen dabei die Pole auf der linken Seite der imaginären Achse liegen (vgl. Abb. 5.11).

Als Beispiel sei ein *Allpass erster Ordnung* betrachtet. Die Polstelle muss in der linken Hälfte der komplexen Ebene liegen. Somit gilt für die Übertragungsfunktion

$$G(s) = \frac{-s + p_1}{s + p_1} \quad p_1 > 0 \quad (5.24)$$

und den Frequenzgang

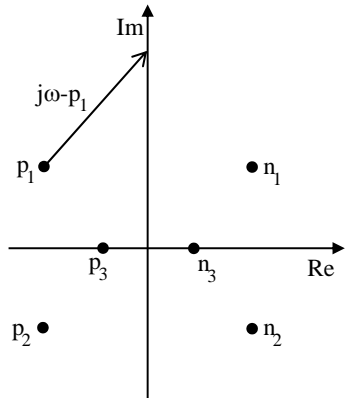


Abb. 5.11. Pole und Nullstellen bei einem Allpass

$$G(j\omega) = \frac{-j\omega + p_1}{j\omega + p_1} = e^{-j2 \arctan \frac{\omega}{p_1}} \quad (5.25)$$

Das Bode-Diagramm ist in Abbildung 5.12 gezeigt. Der Betragsverlauf für die Polstelle ist bei der Frequenz p_1 nach unten abgeknickt, und der Betragsverlauf für die Nullstelle an der gleichen Stelle nach oben. Offensichtlich heben sich beide Effekte gegenseitig auf, so dass der Betragsverlauf des gesamten Übertragungsgliedes konstant Eins ist.

Die Phase senkt sich sowohl bei der Pol- als auch bei der Nullstelle von Null auf $-\frac{\pi}{2}$ ab, so dass sich für den Phasenverlauf des Übertragungsgliedes die gestrichelt eingezeichnete Kurve ergibt, d.h. die Phase senkt sich von Null auf $-\pi$ ab.

Die zugehörige Ortskurve zeigt Abb. 5.13. Der Betrag ist konstant. Daher sind die Vektoren, die vom Ursprung zu den Punkten der Ortskurve führen, immer gleich lang. Die Vektoren haben gemäß Bode-Diagramm zunächst den Winkel Null und zeigen daher nach rechts. Für höhere Frequenzen verringert sich der Winkel bis auf $-\pi$, und die Vektoren zeigen nach links. Damit ergibt sich der gezeichnete Halbkreis.

Für die Sprungantwort des Allpass ergibt sich mit einer einfachen Partialbruchzerlegung

$$y(s) = \frac{1}{s} G(s) = \frac{1}{s} \frac{-s + p_1}{s + p_1} = \frac{1}{s} - \frac{2}{s + p_1} \quad (5.26)$$

und mit Hilfe der Korrespondenztabelle in Anhang A.2 im Zeitbereich

$$y(t) = 1 - 2e^{-p_1 t} \quad (5.27)$$

Der Verlauf der Sprungantwort ist ebenfalls in Abbildung 5.13 zu sehen. Man sieht, dass der Anfangswert ein dem Endwert entgegengesetztes Vorzeichen hat. Dieses Verhalten ist typisch für einen Allpass und macht ihn so schwer zu

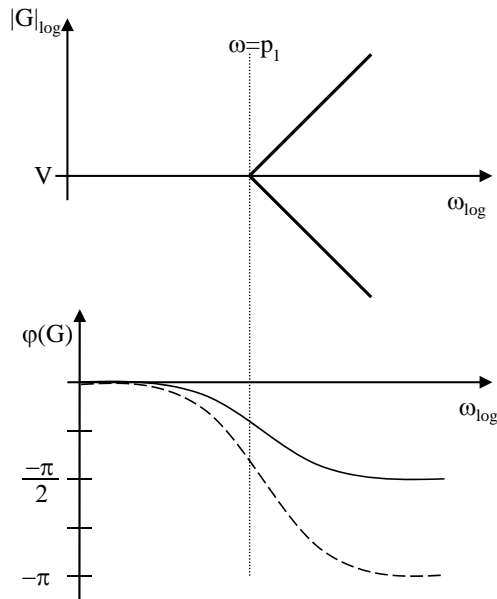


Abb. 5.12. Bode-Diagramm eines Allpass

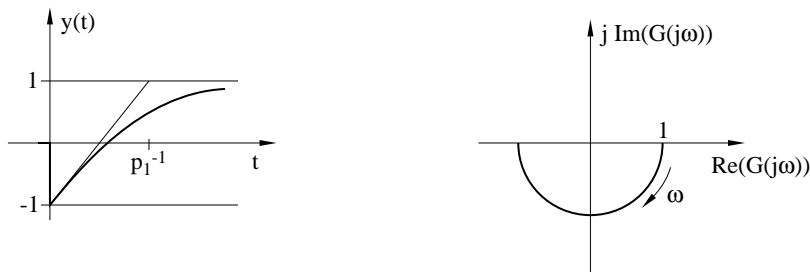


Abb. 5.13. Sprungantwort und Ortskurve eines Allpass erster Ordnung

regeln. Man muss sich nur einen Menschen vorstellen, der einen Allpass regeln soll. Wenn er in der Anfangsphase der Sprungantwort feststellt, dass seine Stellgröße anscheinend genau die falsche Reaktion der Strecke hervorruft, wird er vermutlich die Stellgröße in die entgegengesetzte Richtung verändern, obwohl er doch eigentlich nur hätte abwarten müssen. Daran sieht man schon, dass ein Regler die Dynamik des Systems kennen muss, um es erfolgreich regeln zu können.

Darüber hinaus gefährdet der Allpass, da er bei konstantem Betragsverlauf einen negativen Phasenverlauf hat, die Stabilität des geschlossenen Kreises, wie das nachfolgende Beispiel zeigt.

Gegeben sei die Hintereinanderschaltung eines PT_1 -Gliedes und des oben behandelten Allpass erster Ordnung mit der Kreisübertragungsfunktion

$$G(s)K(s) = \frac{V}{Ts + 1} \frac{-s + p_1}{s + p_1} \quad (5.28)$$

Zunächst soll das PT_1 -Glied für sich betrachtet werden. Seine Ortskurve ist schon in Kapitel 4.3.1 hergeleitet worden und in Abbildung 5.14 noch einmal dargestellt. Um anhand des Nyquist-Kriteriums (Satz 5.6) zu untersuchen, ob die Rückkopplung eines PT_1 -Gliedes stabil ist, ist zunächst die Anzahl seiner Polstellen auf und rechts von der imaginären Achse zu bestimmen. Diese ist Null, und nach Gleichung (5.19) beträgt die notwendige Phasendrehung um den kritischen Punkt -1 damit Null.

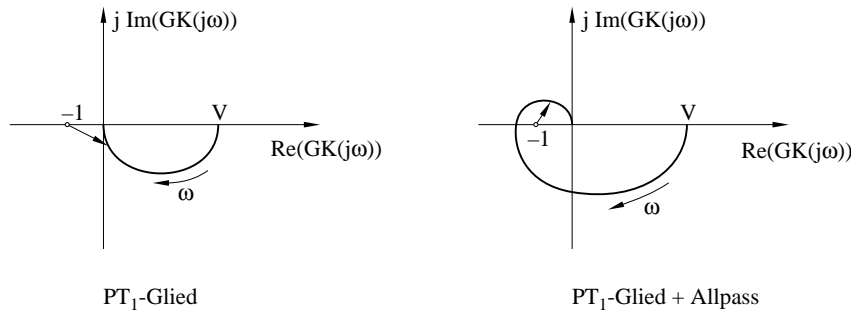


Abb. 5.14. Ortskurven von PT_1 -Glied ohne und mit Allpass

In Abbildung 5.14 ist ein Vektor vom Punkt -1 zur Ortskurve eingezeichnet. Wenn dieser Vektor der Ortskurve einmal vom Anfang bis zum Ende folgt, so wird er zunächst nach rechts zeigen, sich dann ins Negative verdrehen und am Ende wieder nach rechts zeigen. Die resultierende Phasendrehung ist damit Null und entspricht der Forderung des Nyquist-Kriteriums. Die Rückkopplung eines PT_1 -Gliedes wäre damit stabil.

Nun wird das PT_1 -Glied um einen Allpass ergänzt, und es ergibt sich die Kreisübertragungsfunktion (5.28). Die Anzahl der Pole und Nullstellen auf und rechts von der imaginären Achse bleibt weiterhin Null, so dass auch hier die notwendige Phasendrehung um den kritischen Punkt -1 Null beträgt.

Nun ist für diese Hintereinanderschaltung die Ortskurve zu konstruieren. Dazu wird zunächst das Bode-Diagramm gezeichnet. Abbildung 5.15 zeigt die Betrags- und Phasenverläufe von Allpass und PT_1 -Glied. Der Betragsverlauf der Polstelle des PT_1 -Gliedes knickt bei $\frac{1}{T}$ nach unten ab, ebenso der Betragsverlauf der Polstelle des Allpass bei p_1 . Der Betragsverlauf der Nullstelle des Allpass knickt bei p_1 nach oben ab. Da sich die Betragsverläufe von Null- und Polstelle des Allpass gegenseitig kompensieren, entspricht der Betragsverlauf der Hintereinanderschaltung genau dem Betragsverlauf des PT_1 -Gliedes: Er fällt monoton von V auf Null. Aber dies war angesichts des konstanten Betragsverlaufes eines Allpass auch zu erwarten.

Der Phasenverlauf des Allpass beginnt bei Null und endet bei $-\pi$, der des PT_1 -Gliedes beginnt bei Null und endet bei $-\frac{\pi}{2}$. Überlagert man diese Pha-

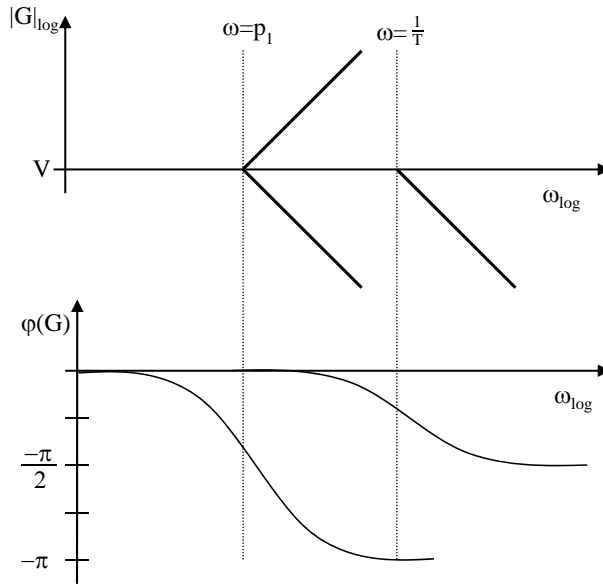


Abb. 5.15. Bode-Diagramm von PT_1 -Glieder und Allpass

senverläufe, so ergibt sich für die Hintereinanderschaltung ein Phasenverlauf, der bei Null beginnt und bei $-\frac{3}{2}\pi$ endet.

Die zugehörige Ortskurve ist in Abbildung 5.14 zu sehen. Der gedachte Vektor vom Ursprung zur Ortskurve, der in der Abbildung nicht eingezeichnet ist, beginnt laut Bode-Diagramm mit dem Betrag V und dem Winkel Null. Damit liegt der Anfangswert der Ortskurve auf der reellen Achse beim Wert V . Anschließend verkürzt sich der Vektor vom Ursprung zur Ortskurve laut Bode-Diagramm bis auf Null und verdreht sich dabei in negativer Richtung bis auf $-\frac{3}{2}\pi$. Die Ortskurve läuft damit durch den vierten, dritten und zweiten Quadranten von oben in den Ursprung der komplexen Ebene.

Je nachdem, welchen Wert V hat, umläuft die Ortskurve dabei den Punkt -1 oder nicht. Für beide Fälle soll nun die Phasendrehung um den kritischen Punkt -1 ermittelt und mit der Forderung des Nyquist-Kriteriums verglichen werden. In beiden Fällen zeigt der Vektor vom kritischen Punkt -1 zur Ortskurve zunächst nach rechts. Wenn der Vektor dann der Ortskurve folgt, zeigt er auch beim Endpunkt der Ortskurve in beiden Fällen wieder nach rechts.

Sofern der Punkt -1 nicht umlaufen wird, zeigt der Vektor zwischenzeitlich nach unten, dreht sich dann nach oben in den positiven Bereich und erreicht schließlich seine Endposition mit dem Winkel Null. In dem Fall ist die resultierende Phasendrehung Null und entspricht damit genau der Forderung des Nyquist-Kriteriums. Die Bildung eines geschlossenen Kreises mit der

Kreisübertragungsfunktion (5.28) würde demnach einen stabilen geschlossenen Kreis ergeben.

Anders ist dies, falls der Punkt -1 umlaufen wird. In dem Fall dreht sich der Vektor, der vom Punkt -1 zur Ortskurve führt, mit dem Verlauf der Ortskurve einmal in negativer Richtung um seine eigene Achse. Die resultierende Phasendrehung beträgt -2π und entspricht nicht der Forderung des Nyquist-Kriteriums, so dass hier die Rückkopplung auf einen instabilen geschlossenen Kreis führen würde.

Anhand dieses Beispiels wird deutlich, wie der Allpass durch Absenken der Phase ohne gleichzeitiges Absenken des Betrages das Stabilitätsverhalten eines geschlossenen Kreises negativ beeinflussen kann.

Grundsätzlich lässt sich feststellen: Durch das Hinzufügen des Allpass wird nicht der Betragsverlauf, wohl aber der Phasenverlauf des Frequenzganges verändert, und zwar wird der Absolutwert der Phase vergrößert. Das PT_1 -Glied ohne Allpass weist also bei gleichem Betragsverlauf des Frequenzganges kleinere Absolutwerte der Phase auf als das PT_1 -Glied mit Allpass. Diese Vergrößerung der Absolutwerte der Phase bei Hinzufügen eines Allpasses gilt offenbar für jedes stabile, rationale Übertragungsglied. Andererseits lässt sich zeigen, dass ein stabiles, rationales Übertragungsglied ohne Allpass die für einen gegebenen Betragsverlauf kleinstmöglichen Absolutwerte der Phase aufweist. Ein solches Übertragungsglied bezeichnet man daher als *Minimalphasenglied*, während ein Übertragungsglied mit Allpass als *Nicht-Minimalphasenglied* bezeichnet wird.

Offenbar können nur Übertragungsglieder, die weder Null- noch Polstellen mit positivem Realteil enthalten, Minimalphasenglieder sein. Denn von jeder anderen, stabilen, rationalen Übertragungsfunktion lässt sich noch ein Allpass abspalten, wie das folgende, einfache Beispiel zeigt:

$$G(s) = \frac{s-2}{s+1} = \frac{s+2}{s+1} \frac{s-2}{s+2} = G_{\min}(s)G_{\text{all}}(s) \quad (5.29)$$

Eine rationale Übertragungsfunktion, die Nullstellen mit positivem Realteil aufweist, enthält also einen Allpass und ist demnach ein Nicht-Minimalphasenglied.

In der Praxis treten allpasshaltige Strecken wesentlich häufiger auf, als man zunächst annimmt. Ein Beispiel für eine allpasshaltige Strecke ist ein Motorradfahrer, der eine Linkskurve fahren will. Er muss zunächst eine kleine Bewegung nach rechts machen, um die erforderliche Schräglage nach links zu bekommen. Ein anderes Beispiel ist das Heck eines Flugzeuges, das beim Übergang vom Horizontal- in den Steigflug zunächst nach unten absackt, bevor es sich dann schließlich nach oben bewegt.

6. Regelung linearer Eingrößensysteme

6.1 Anforderungen an einen Regler

Nachdem in den bisherigen Kapiteln die nötigen Kenntnisse zur Analyse dynamischer Systeme vermittelt wurden, soll in diesem Kapitel auf den eigentlichen Entwurf von Reglern eingegangen werden. Zur Erinnerung sei dazu noch einmal die Standardkonfiguration eines Regelkreises skizziert (Abb. 6.1), wobei im Vergleich zu Abb. 5.1 das Messglied von vornherein vernachlässigt wird. Führungs- und Störübertragungsfunktion lauten:

$$T(s) = \frac{y(s)}{w(s)} = \frac{G(s)K(s)}{G(s)K(s) + 1} \quad (6.1)$$

$$S(s) = \frac{y(s)}{d(s)} = \frac{G(s)}{G(s)K(s) + 1} \quad (6.2)$$

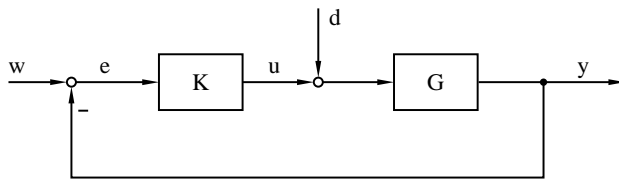


Abb. 6.1. Regelkreis

Zu einer gegebenen Strecke $G(s)$ soll nun ein geeigneter Regler $K(s)$ gefunden werden. Aber welche Forderungen sind überhaupt zu erfüllen? Grundlegend ist natürlich die Forderung nach Stabilität. $K(s)$ ist so zu wählen, dass der entstehende geschlossene Kreis stabil ist und damit beispielsweise das Nyquist-Kriterium (als eines von verschiedenen möglichen Stabilitätskriterien) erfüllt.

Optimal wäre darüber hinaus offensichtlich $T(j\omega) = 1$ und $S(j\omega) = 0$. Die erste Forderung bedeutet, dass die Regelgröße y unabhängig von der Frequenz des Eingangssignales immer gleich der Führungsgröße w ist. Damit wäre das System natürlich auch BIBO-stabil. Die zweite Forderung entspricht einer vollständigen Unterdrückung des Einflusses der Störgröße d auf die Regelgröße. Insgesamt stehen die beiden Forderungen demnach für die Forderung nach *Genauigkeit* der Regelung.

Leider ist dieser für einen Regelungstechniker paradiesische Zustand nicht zu verwirklichen. Nach Gleichung (6.1) kann nämlich für einen gegebenen Frequenzgang der Strecke $G(j\omega)$ die Funktion $T(j\omega)$ nur dann konstant Eins werden, wenn der Frequenzgang des Reglers $K(j\omega)$ für alle Frequenzen unendlich große Werte annimmt. Auf dasselbe Ergebnis führt auch die Forderung nach vollständiger Unterdrückung der Störgröße. Ein solcher Regler ist aber nicht zu realisieren, und seine Ausgangsgröße würde auch die Möglichkeiten jedes Stellgliedes übersteigen.

Andererseits ist es bei praktischen Strecken auch gar nicht notwendig, dass die oben genannten Forderungen für alle Frequenzen erfüllt werden. Vielmehr reicht es aus, wenn Genauigkeit im *Nutzfrequenzbereich* erzielt wird, d.h. im Bereich derjenigen Frequenzen, die im Eingangssignal auch tatsächlich enthalten sind. Dies sind aber normalerweise die niedrigen Frequenzen einschließlich der Frequenz Null, was einem Gleichsignal entspricht. Dabei kann für Gleichsignale auf Genauigkeit am wenigsten verzichtet werden, denn gerade bei einem konstanten Eingangssignal sollte man von einem geregelten System nach Beendigung aller Einschwingvorgänge erwarten dürfen, dass seine Ausgangsgröße denselben Wert wie die Eingangs- bzw. Sollgröße annimmt. Die Anforderungen an eine Regelung werden deshalb so weit zurückgenommen, dass die Optimalforderungen nur noch für Gleichsignale (Frequenz $s = 0$) erhoben werden:

$$\lim_{s \rightarrow 0} T(s) \stackrel{!}{=} 1 \quad \text{und} \quad \lim_{s \rightarrow 0} S(s) \stackrel{!}{=} 0 \quad (6.3)$$

Wegen der Stetigkeit der beiden Übertragungsfunktionen $T(s)$ und $S(s)$ sind dann auch für kleine Werte von s bzw. ω und damit im Nutzfrequenzbereich die Forderungen zumindest noch näherungsweise erfüllt. Bei Zutreffen der Gleichungen (6.3) spricht man auch von *stationärer Genauigkeit*. Ein stationär genaues System ist auch auf jeden Fall stabil im Sinne von Def. 5.1, d.h. es weist eine endliche Sprungantwort auf. Es ergibt sich nämlich für die Sprungantwort

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sy(s) = \lim_{s \rightarrow 0} s \frac{1}{s} T(s) = \lim_{s \rightarrow 0} T(s) = 1 \quad (6.4)$$

d.h. die Ausgangsgröße des geregelten Systems weist bei einem Eingangssprung den konstanten Endwert Eins auf.

Für den Regler folgt aus der Forderung nach Genauigkeit mit Gleichung (6.1)

$$\lim_{s \rightarrow 0} K(s) = \infty \quad (6.5)$$

Setzt man voraus, dass $K(s)$ eine rationale Funktion ist, so führt dies auf die notwendige Bedingung, dass $K(s)$ einen Pol bei $s = 0$ aufweisen muss. Sofern $G(s)$ keine Nullstelle bei $s = 0$ hat, wird das Produkt $G(s)K(s)$ für $s = 0$ unendlich groß, und $T(s)$ konvergiert gegen Eins. Besitzt $G(s)$ dagegen eine solche Nullstelle, so nimmt $\lim_{s \rightarrow 0} G(s)K(s)$ einen endlichen Wert an, und

$\lim_{s \rightarrow 0} T(s)$ konvergiert nicht gegen Eins. Offensichtlich muss die Ordnung des Pols von $K(s)$ die Ordnung der Nullstelle von $G(s)$ bei $s = 0$ um mindestens Eins übersteigen.

Ein Sonderfall soll hier nicht unerwähnt bleiben: Wenn die Strecke integrierende Wirkung hat, kann man die Übertragungsfunktion in der Form $G(s) = \frac{1}{s} \tilde{G}(s)$ mit $\tilde{G}(0) \neq 0$ schreiben, was einer Hintereinanderschaltung von Integrator und dem Streckenteil $\tilde{G}(s)$ entspricht. Wenn außerdem die Störgröße d erst hinter dem Integrator angreift (Abb. 6.2), so ergibt sich für $T(s)$ und $S(s)$:

$$\begin{aligned} T(s) &= \frac{\tilde{G}(s)K(s)}{\tilde{G}(s)K(s) + s} \\ S(s) &= \frac{s\tilde{G}(s)}{\tilde{G}(s)K(s) + s} \end{aligned} \quad (6.6)$$

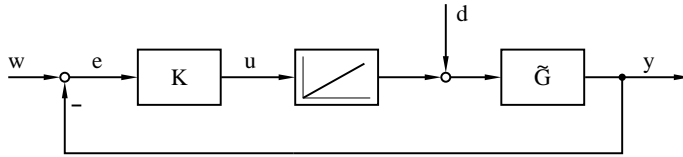


Abb. 6.2. Abspalten eines Integrators von der Strecke

Die geforderten Grenzwerte zur Erzielung stationärer Genauigkeit (vgl. (6.3)) werden jetzt für jeden beliebigen Regler mit $K(0) \neq 0$ erreicht. Man kann $K(s) = 1$ setzen und somit im Prinzip auf den Regler verzichten. Oder anders ausgedrückt, man kann den Integrator als Teil des Reglers auffassen, so dass $\lim_{s \rightarrow 0} K(s) = \infty$ gegeben ist. Diese günstige Konstellation kann vor allem dann entstehen, wenn das Stellglied, das aus Sicht des Reglers Teil der Strecke ist, integrierende Wirkung hat.

Ein Beispiel für ein solches Stellglied ist ein durch einen Motor angetriebenes Ventil, mit dem der Durchfluss durch ein Rohr geregelt werden soll. Der Regler kann in diesem Beispiel gleich Eins gesetzt werden: $K(s) = 1$. Der Motor, dessen interne Ausgleichsvorgänge vernachlässigt werden sollen (PT_1 -Glieder mit der Ersatzzeitkonstanten $T_e \ll 1s$), wird mit der Stellgröße u des Reglers angesteuert. Solange diese ungleich Null ist, wird sich der Motor drehen und den Öffnungsquerschnitt A des Ventils stetig verändern. Das Übertragungsverhalten von der Motordrehzahl ω zum Ventilöffnungsquerschnitt A entspricht damit dem eines Integrators (Abb. 6.3). Diese Veränderung von A erfolgt so lange, bis die Ausgangsgröße y , also der Durchfluss durch das Rohr, gleich dem Sollwert w ist. Erst dann wird die Regeldifferenz e und, leicht verzögert durch die Motor-internen Einschwingvorgänge, auch die Drehzahl ω zu Null, und der dann eingestellte Öffnungsquerschnitt hört auf, sich zu verändern.

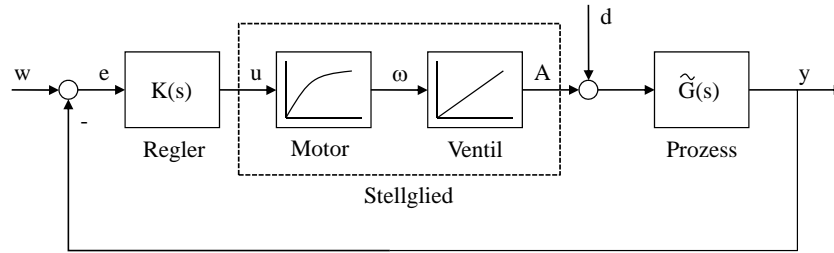


Abb. 6.3. Beispiel für ein integrierendes Stellglied

Die Erfüllung der Gleichungen (6.3), d.h. stationäre Genauigkeit, ist nach der Gewährleistung der Stabilität fast immer die erste Anforderung an eine Regelung. Darüber hinaus gibt es aber noch weitere Kriterien, die nicht zu vernachlässigen sind. Zum einen ist dies die Forderung nach einer ausreichenden *Regelgeschwindigkeit*. So kann man es beispielsweise den Fahrgästen in einem Aufzug nicht zumuten, minutenlang auf das Erreichen des nächsten Stockwerks zu warten. Wie die späteren Beispiele zeigen werden, ist dies der Forderung nach stationärer Genauigkeit und damit nach Stabilität oft entgegengerichtet, so dass eine Regelung hier immer nur einen Kompromiss darstellen kann.

Ein weiteres Kriterium ist eine ausreichend große Dämpfung des Systems. So ist beim Aufzug ein Überschwingen des Lageregelkreises ebenfalls nicht akzeptabel, denn soll ein bestimmtes Stockwerk erreicht werden, so darf der Aufzug nicht erst etwas zu weit fahren und sich dann auf die richtige Stelle einpendeln. Hier ist ein aperiodisches Einschwingverhalten gefordert, d.h. die Dämpfung D muss größer als Eins sein. Bei Systemen, in denen ein leichtes Überschwingen nicht so kritisch ist, strebt man meist eine Dämpfung $D = \frac{1}{\sqrt{2}}$ an, weil dies die kleinstmögliche Dämpfung ist, bei der noch keine Resonanzüberhöhung auftritt (vgl. Kap. 4.2.9).

Im Einzelfall lassen sich noch weitere Kriterien definieren. Je nach Anwendungsfall kann zum Beispiel die Amplitude des Überschingers bei der Sprungantwort relevant sein, oder die Zeit, die benötigt wird, um einen vorgegebenen Toleranzbereich um den Endwert der Sprungantwort zu erreichen. Dementsprechend kann man auch die verschiedensten Gütemaße für eine Regelung definieren. Ein oft verwendetes, zu minimierendes Gütemaß lautet beispielsweise:

$$Q = \int_0^{\infty} [(e(t))^2 + k(u(t))^2] dt \quad \text{mit } k > 0 \quad (6.7)$$

Damit Q möglichst klein ist, muss somit einerseits der mittlere quadratische Regelfehler e und damit die Abweichung zwischen Soll- und Istwert klein sein. Der zweite Summand gewährleistet dagegen, dass dieses Ziel mit einer kleinen

Stellgröße erreicht werden kann, um so die Stelleinrichtung des Regelkreises zu schonen.

Das Verhalten des Stellgliedes ist beim Reglerentwurf ohnehin in zweierlei Hinsicht zu berücksichtigen. Zum einen als Teil der zu regelnden Strecke, wenn es um Stabilität, Dämpfung und Regelgeschwindigkeit geht. Zum anderen ist zu beachten, dass es aufgrund seiner technischen Ausführung nur Signale mit einer bestimmten maximalen Amplitude und Frequenz übertragen kann. Es nützt also nichts, wenn die Ausgangsgröße des Reglers Signale von hoher Frequenz oder Amplitude enthält, die vom Stellglied gar nicht an die Strecke weitergegeben werden können. Es besteht dann sogar die Gefahr, dass das Stellglied übersteuert und ein nichtlineares Verhalten aufweist, womit der gesamte Reglerentwurf, der von einem linearen Verhalten der einzelnen Übertragungsglieder ausgeht, wieder in Frage gestellt wird. Ein einfaches Beispiel hierfür ist das Ruder bei einem Schiff. Über einen durch technische Randbedingungen vorgegebenen Maximalwinkel hinaus kann es nicht verstellt werden. Während im Normalbetrieb der Ruderwinkel proportional zur Eingangsgröße der Rudereinrichtung ist, kann nach Erreichen der maximalen Auslenkung auf eine weitere Erhöhung der Eingangsgröße nicht mehr reagiert werden. Aus dem linearen Übertragungsglied mit proportionalem Verhalten ist ein nichtlineares Glied geworden, dessen Kennlinie in Abb. 6.4 zu sehen ist. Eine solche Kennlinie ist typisch für viele Stellglieder.

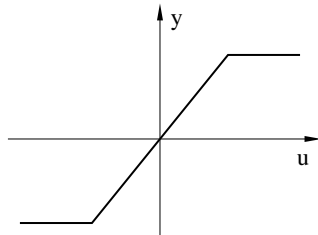


Abb. 6.4. Kennlinie eines Stellgliedes

Nachdem jetzt eine Vorstellung über die Anforderungen an einen Regler besteht, sollen im Folgenden einige Standardregler behandelt werden, die einfach zu verstehen, zu realisieren und vor allem zu dimensionieren sind. Aus diesem Grund wird auch der weitaus größte Teil aller in der Praxis vorkommenden Regelungen mit diesen Reglern verwirklicht.

6.2 Regler vom PID-Typ und ihr Entwurf

6.2.1 P-Regler

Der *Proportionalregler* (*P-Regler*) stellt sicherlich den einfachsten Ansatz dar. Da die Eingangsgröße des Reglers die Regelabweichung ist und seine

Ausgangsgröße die Stellgröße, erzeugt dieser Regler mit der Übertragungsfunktion

$$K(s) = P \quad (6.8)$$

eine zur Regelabweichung proportionale Stellgröße

$$u = Pe = P(w - y) \quad (6.9)$$

Je größer die Abweichung zwischen Ist- und Sollwert ist, desto größer ist die Stellgröße des Reglers. Als einzigen einstellbaren Parameter hat dieser Regler seinen *Verstärkungsfaktor* P .

Stationäre Genauigkeit und die vollständige Ausregelung einer Störung können mit dem P-Regler nicht erzielt werden, denn dazu müsste $K(0) = P$ nach den obigen Betrachtungen unendlich groß werden, sofern nicht die Konstellation aus Abbildung 6.2 vorliegt.

Dies ist auch plausibel, wie die folgende Erklärung zeigt: Gegeben sei eine proportional wirkende, lineare Regelstrecke, die nur aus Verzögerungsgliedern besteht, ein von Null verschiedener Sollwert w und zunächst eine Regelabweichung $e = w - y > 0$. Der Regler wird eine zur Regelabweichung proportionale Stellgröße ausgeben. Sie wird den Regelfehler verringern. Damit verringert sich aber auch wiederum die Stellgröße. Wenn der Regelfehler Null wird, dann wird auch die Stellgröße und damit die Eingangsgröße der Strecke Null. Wenn aber bei einer proportional wirkenden Strecke die Eingangsgröße u Null wird, dann wird, mit einer gewissen Verzögerung, auch die Ausgangsgröße y Null, und damit wäre der Regelfehler $e = w - y$ wieder da, denn der Sollwert war ja laut Voraussetzung ungleich Null. Letztendlich führt diese Erklärung auf einen Widerspruch, der zeigt, dass der P-Regler nicht ausreicht, um eine Strecke stationär genau zu regeln.

Damit der Regler zumindest näherungsweise dem Idealwert $P = K(s) = \infty$ entspricht und damit die Regelabweichung zwar nicht verschwindet, aber zumindest möglichst klein wird, sollte P möglichst groß gewählt werden. Neben der Genauigkeit steigt dadurch auch die Regelgeschwindigkeit, denn bei gegebener Regelabweichung führt eine Vergrößerung von P offensichtlich zu einer Vergrößerung der Stellgröße. Und aus dieser stärkeren Anregung der Strecke resultiert natürlich eine schnellere Annäherung der Regelgröße an den Sollwert, auch wenn dieser zum Schluss nicht genau erreicht wird.

Einer Erhöhung von P sind aber aus Stabilitätsgründen Grenzen gesetzt. An zwei Beispielen soll dies verdeutlicht werden. Gegeben seien zwei Tiefpassstrecken zweiter und dritter Ordnung, d.h. Hintereinanderschaltungen von zwei bzw. drei PT_1 -Gliedern, deren Ortskurven aus Abb. 6.5 ersichtlich sind. Bei Regelung mit einem P-Regler ergibt sich für die Kreisübertragungsfunktionen des geschlossenen Kreises jeweils: $G(s)K(s) = G(s)P$. Die zugehörigen Ortskurven gewinnt man durch Multiplikation der gegebenen Ortskurven mit P , was für $P > 1$ gleichbedeutend mit einer Dehnung ist. Und damit wird auch die Gefahr deutlich, die eine Erhöhung von P beinhaltet: Nach dem Nyquist-Kriterium ist die zulässige Phasendrehung bezüglich

-1 in beiden Fällen Null, anschaulich gesehen dürfen die Ortskurven den Punkt -1 also nicht links umfahren und sollten sich ihm im Interesse einer ausreichenden Dämpfung auch nicht zu sehr nähern. Der Kreis mit der Strecke dritter Ordnung wird demnach instabil, wenn man P immer weiter vergrößert. Der Kreis mit der Strecke zweiter Ordnung kann zwar durch Erhöhen von P nicht instabil werden, die Ortskurve kommt dem Punkt -1 aber immer näher, so dass der geschlossene Kreis eine unzumutbar kleine Dämpfung aufweist (vgl. Satz 5.8).

In der Praxis sind die durch Stabilitäts- bzw. Dämpfungsanforderungen gegebenen Obergrenzen für die Verstärkung des P-Reglers normalerweise so niedrig, dass mit den zulässigen Werten für P eine stationäre Genauigkeit nicht einmal näherungsweise realisiert werden kann. Dennoch gibt es genügend Anwendungsfälle, in denen stationäre Genauigkeit nicht wichtig ist und stattdessen das Kostenargument zugunsten des P-Reglers entscheidet. Und nicht zuletzt kann ein P-Regler immer dann eingesetzt werden, wenn, wie oben erläutert, die Strecke bzw. das Stellglied integrierende Wirkung hat.

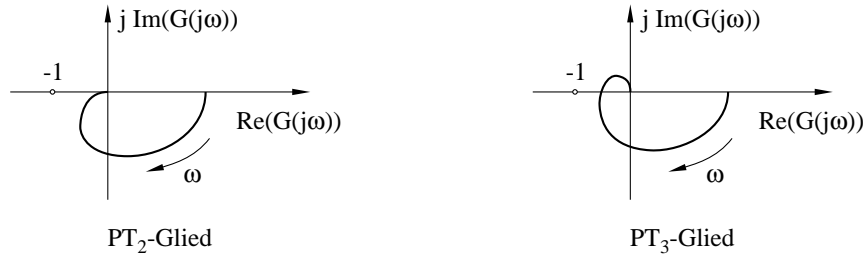


Abb. 6.5. Ortskurven von Tiefpassstrecken zweiter und dritter Ordnung

Um das Verhalten des P-Reglers und auch anderer, im Folgenden noch vorgestellter Regler besser einschätzen zu können, zeigt Abb. 6.6 die Sprungantworten eines geschlossenen Kreises mit verschiedenen Reglern und einem Tiefpass dritter Ordnung als Strecke:

$$G(s) = \frac{1}{(T_1 s + 1)(T_2 s + 1)(T_3 s + 1)} \quad (6.10)$$

Die mit P bezeichnete Kurve kennzeichnet die Sprungantwort bei Regelung mit einem P-Regler. Deutlich ist zu sehen, dass nach Beendigung des Einschwingvorganges eine stationäre Regelabweichung zurückbleibt. Würde man die Reglerverstärkung erhöhen, so könnte man zwar die Regelabweichung verkleinern, müsste aber gleichzeitig noch größere Schwingungen am Anfang und schließlich sogar Instabilität in Kauf nehmen.

6.2.2 I-Regler

Hinsichtlich der stationären Genauigkeit lassen sich deutlich bessere Regelergebnisse mit einem *Integralregler* (*I-Regler*) erzielen:

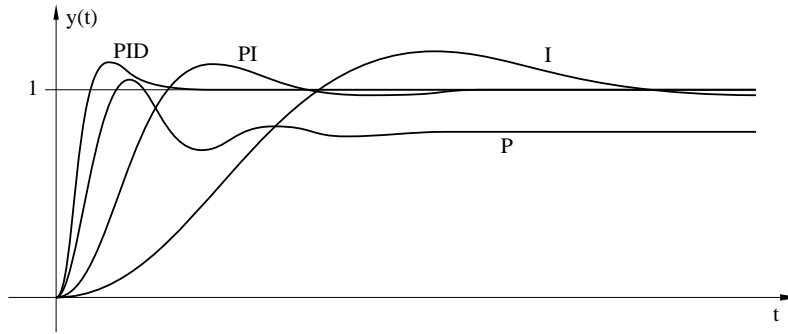


Abb. 6.6. Vergleich verschiedener Reglertypen

$$K(s) = I \frac{1}{s} = \frac{1}{T_i s} \quad (6.11)$$

Dieser ist wegen (6.5) offensichtlich ein stationär genauer Regler, sofern die Strecke keine Nullstelle bei $s = 0$ hat. Doch von diesem Fall soll hier abgesehen werden. Die stationäre Genauigkeit lässt sich auch anschaulich begründen: Solange die Eingangsgröße e des Reglers ungleich Null ist, wird sich wegen der Integration auch die Ausgangsgröße u des Reglers immer weiter verändern. Erst wenn $e = 0$ gilt, ändert sich auch die Stellgröße u nicht mehr. $e = 0$ bedeutet aber, dass die Regelgröße y gleich der Führungsgröße w ist.

Der Vollständigkeit sei zu dieser anschaulichen Erklärung angemerkt, dass mit $e = 0$ noch nicht unbedingt der stationäre Endzustand des Systems erreicht sein muss. Denn falls in der Strecke weitere Integratoren vorhanden sind, so kann deren Eingangsgröße durchaus noch von Null verschieden sein, so dass sie weiter aufintegrieren und sich auch die Ausgangsgröße y des Systems wieder von w entfernt. In dem Fall muss der Regler die neue Abweichung wieder mit einer Stellgrößenänderung bekämpfen. Es ist eine Frage der Stabilität des geschlossenen Regelkreises, ob jede neu entstehende Abweichung kleiner ist als die vorhergehende und das System damit irgendwann zur Ruhe kommt. Klar ist aber, dass ein Ruhezustand nur dann erreicht sein kann, wenn auch $e = 0$ ist.

Der Parameter T_i wird Integrierzeit genannt. Je kürzer diese Integrierzeit ist, desto schneller ändert sich die Stellgröße bei gegebener Regelabweichung. Im Interesse einer hohen Regelgeschwindigkeit sollte man T_i also möglichst klein wählen.

Auch bei diesem Regler steht dem aber die Forderung nach Stabilität im Wege. Als Beispiel soll ein PT_2 -Glied mit einem Integralregler geregelt werden. Die Kreisübertragungsfunktion lautet

$$G(s)K(s) = \frac{V_S}{\frac{s^2}{\omega_0^2} + \frac{2D}{\omega_0}s + 1} \frac{1}{T_i s} = \frac{V_S}{T_i} \frac{1}{\frac{s^3}{\omega_0^2} + \frac{2D}{\omega_0}s^2 + s} \quad (6.12)$$

mit der Kreisverstärkung $\frac{V_s}{T_i}$. Das zugehörige Bode-Diagramm ist in Abbildung 6.7 gezeigt. Im Betragsdiagramm sieht man den linear fallenden Verlauf des Integrators sowie den Verlauf des PT_2 -Gliedes, hier der Einfachheit halber ohne Resonanzüberhöhung. In Summe (gestrichelte Linie) ergibt sich damit offenbar ein fallender Betragsverlauf, der von ∞ gegen Null konvergiert.

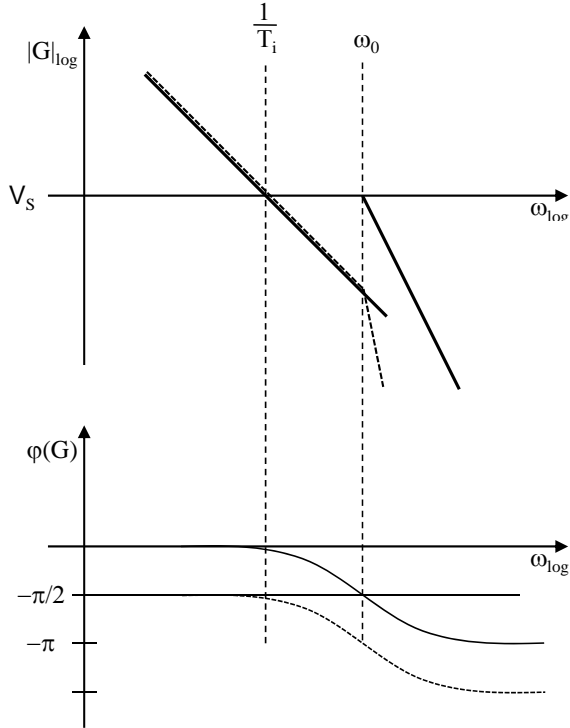


Abb. 6.7. Bode-Diagramm für den I-Regler mit PT_2 -Glied

Der Phasenverlauf setzt sich zusammen aus der Phase des Integrators, die für alle Frequenzen den Wert $-\frac{\pi}{2}$ aufweist, sowie die von Null auf $-\pi$ fallende Phase des PT_2 -Gliedes. Für die Überlagerung ergibt sich damit der gestrichelt gezeichnete Phasenverlauf, der aussieht wie der des PT_2 -Gliedes, lediglich um $-\frac{\pi}{2}$ nach unten versetzt.

Die zugehörige Ortskurve zeigt Abb. 6.8 (linke Kurve). Sie beginnt im Unendlichen unter dem Winkel $-\frac{\pi}{2}$, kommt also von unten, und läuft unter dem Winkel $-\frac{3}{2}\pi$ in den Ursprung hinein.

Die erforderliche Phasendrehung der Ortskurve um den Punkt -1 beträgt nach Satz 5.6 wegen des Integrators in der Kreisübertragungsfunktion $\frac{\pi}{2}$. Falls die Ortskurve der Kreisübertragungsfunktion also wie eingezeichnet verläuft, entspricht die Phasendrehung um den kritischen Punkt -1 genau dem erforderlichen Wert, und der geschlossene Kreis ist stabil. Verkleinert man aber die

Integrierzeit T_i zur Erhöhung der Regelgeschwindigkeit, so steigt die Kreisverstärkung und die Ortskurve wird gedehnt, bis sie den Punkt -1 links umfährt. In dem Fall würde die Phasendrehung $-\frac{3}{2}\pi$ betragen, und der geschlossene Kreis wäre instabil.



Abb. 6.8. Ortskurve eines PT_2 -Gliedes mit I- und PI-Regler

Sehr ähnliche Verhältnisse treten auch beim PT_3 -Glied auf, wie die Abbildungen 6.9 und 6.10 verdeutlichen. Der Betragsverlauf ist im Bode-Diagramm nicht mit eingezeichnet, da er offensichtlich wieder von Unendlich gegen Null geht. Beim Phasenverlauf müssen die Verläufe von drei PT_1 -Gliedern und dem Integrator überlagert werden. Der Gesamt-Phasenverlauf beginnt daher wegen des Integrators bei $-\frac{\pi}{2}$ und endet bei -2π . Die Ortskurve kommt damit auch hier wieder aus dem Unendlichen von unten, läuft jetzt aber von rechts in den Ursprung.

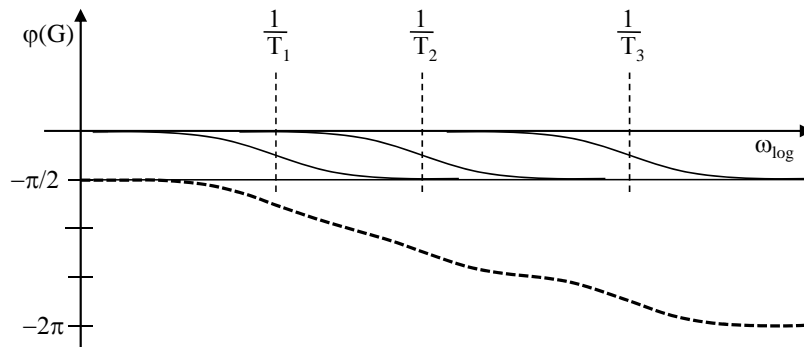


Abb. 6.9. Bode-Diagramm eines PT_3 -Gliedes mit I-Regler

Die Überlegungen hinsichtlich der Stabilität sind die gleichen wie im vorhergehenden Beispiel. Wegen des Integrators beträgt die erforderliche Phasendrehung der Ortskurve um den Punkt -1 wieder $\frac{\pi}{2}$, und entsprechend wird auch hier die Verkleinerung der Integrierzeit des Integralreglers durch das Nyquist-Kriterium begrenzt.

Im Hinblick auf die Stabilität ist der Integralregler, wie man an den Beispielen erkennen konnte, nicht besonders günstig. Um die Kreisübertragungsfunktion zu erhalten, muss die Streckenübertragungsfunktion mit dem Faktor $\frac{1}{T_i s}$ multipliziert werden. Die Phase der Kreisübertragungsfunktion ergibt

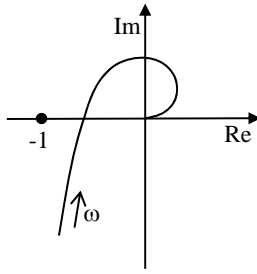


Abb. 6.10. Ortskurve eines PT_3 -Gliedes mit I-Regler

sich demnach als Summe aus der Phase der Streckenübertragungsfunktion $G(s)$ und dem konstanten Winkel $-\frac{\pi}{2}$. Dies bedeutet, dass die Phase der Kreisübertragungsfunktion gegenüber der Streckenübertragungsfunktion um den konstanten Winkel $-\frac{\pi}{2}$ abgesenkt und die Ortskurve im mathematisch negativen Sinn, näher zum Punkt -1 hin, verdreht wird. Dadurch steigt offensichtlich die Gefahr einer Verletzung des Nyquist-Kriteriums und damit der Instabilität.

Weiterhin ist der Integralregler, unabhängig von der Integrierzeit, schon von seiner Konfiguration her ein langsamer Regler. Tritt beispielsweise eine plötzliche Regelabweichung auf, so muss der Regler diese Größe erst aufintegrieren, bevor die Stellgröße einen nennenswerten Betrag erreicht. Dieser Effekt ist auch deutlich in Abb. 6.6 zu erkennen. Die Sprungantwort des mit dem I-Regler geregelten PT_3 -Gliedes erreicht zwar den richtigen Endwert, steigt aber zu Anfang nur sehr langsam an.

6.2.3 PI-Regler

Die Nachteile werden behoben, wenn man P- und I-Regler miteinander kombiniert. Man gelangt dann zu dem am häufigsten eingesetzten Regler überhaupt, dem *PI-Regler*. Seine Übertragungsfunktion lautet:

$$K(s) = P + I \frac{1}{s} = V_R \frac{T_{pi}s + 1}{T_{pi}s} \quad (6.13)$$

Man kann den PI-Regler also entweder als Parallelschaltung aus P- und I-Regler oder als Hintereinanderschaltung aus Vorhalt $(T_{pi}s + 1)$ und Integrator $\frac{1}{T_{pi}s}$ auffassen. Die zweite Darstellungsart bietet sich zur Aufstellung von Kreisübertragungsfunktionen $G(s)K(s)$ an, da sich nur bei faktorieller Darstellung Phase und Betrag der einzelnen Teilübertragungsfunktionen graphisch überlagern lassen. Anhand der ersten Darstellungsart können dagegen leichter die Sprungantwort und die Ortskurve des Reglers skizziert werden.

Für die Sprungantwort ergibt sich:

$$y(s) = \frac{1}{s} K(s) = P \frac{1}{s} + I \frac{1}{s^2} \quad \leftrightarrow \quad y(t) = P + It \quad (6.14)$$

Sie ist in Abbildung 6.11 dargestellt. Anhand des Verlaufes der Sprungantwort wird sofort deutlich, warum dieser Regler schneller ist als ein I-Regler: Auf eine sprungförmige Regelabweichung reagiert der Regler von vornherein mit einer von Null verschiedenen Stellgröße, die dann durch den Integrator nur noch nachgebessert wird. Die Stellgröße muss im Gegensatz zum I-Regler nicht erst langsam aufintegriert werden. Das verbesserte Regelverhalten ist auch in Abb. 6.6 zu erkennen.



Abb. 6.11. Sprungantwort und Ortskurve eines PI-Reglers

Die Ortskurve erhält man leicht anhand folgender Überlegung: Der Term I_s^1 für sich genommen hätte die Ortskurve eines Integrators, die auf der negativ imaginären Achse aus dem Unendlichen kommend in den Ursprung läuft. Durch das Hinzunehmen des Proportionalanteils wird diese Kurve um P nach rechts verschoben (siehe Abbildung 6.11).

Inwiefern der PI-Regler auch im Hinblick auf die Stabilität deutlich günstiger als der I-Regler ist, lässt sich anhand des Beispiels zeigen, das schon für den I-Regler angeführt wurde. Gegeben sei wieder eine Strecke in Form eines PT_2 -Gliedes, das nun mit einem PI-Regler geregelt werden soll. Für diese Hintereinanderschaltung ergibt sich das Bode-Diagramm gemäß Abbildung 6.12.

Im Betragsdiagramm findet sich zunächst der Betragsverlauf des Integrators, der als lineare Funktion die ω -Achse bei $\frac{1}{T_{pi}}$ schneidet. Dazu kommt der Betragsverlauf der Regler-Nullstelle, der bei $\frac{1}{T_{pi}}$ nach oben abknickt, und schließlich der Betragsverlauf der Strecke, den man an der Resonanzüberhöhung in der Nähe der Eigenkreisfrequenz ω_0 gut erkennen kann. Die Überlagerung dieser Verläufe (gestrichelte Linie) liefert für niedrige Frequenzen einen fallenden Betragsverlauf entsprechend dem des Integrators. Für Frequenzen zwischen $\frac{1}{T_{pi}}$ und ω_0 kompensieren sich dann die Betragsverläufe der Nullstelle und des Integrators, während sich das PT_2 -Glied noch nicht auswirkt. Ab der Resonanzüberhöhung folgt der Betragsverlauf dann dem des PT_2 -Gliedes.

Anders der Phasenverlauf. Zu sehen ist der Phasenverlauf des Integrators mit dem konstanten Wert $-\frac{\pi}{2}$, der Phasenverlauf des Vorhaltes, der von Null ausgehend gegen $\frac{\pi}{2}$ konvergiert, sowie der Verlauf des PT_2 -Gliedes, der von Null gegen $-\pi$ konvergiert. Der Phasenverlauf der Kreisübertragungsfunktion (gestrichelte Linie) beginnt demnach bei $-\frac{\pi}{2}$, wird dann durch den Vorhalt angehoben und am Ende durch das PT_2 -Glied auf $-\pi$ gedrückt.

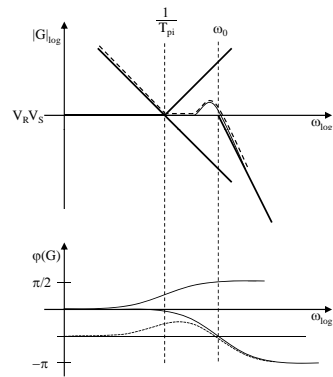


Abb. 6.12. Bode-Diagramm für den PI-Regler mit PT_2 -Glieder

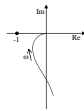


Abb. 6.13. Ortskurve für den PI-Regler mit PT_2 -Glieder

Damit beginnt die Ortskurve (Abb. 6.13) im Unendlichen unter dem Winkel $-\frac{\pi}{2}$, kommt also von unten. Die Werte der Phase liegen zunächst zwischen $-\frac{\pi}{2}$ und Null. Damit verläuft die Kurve zunächst im vierten Quadranten, schneidet dann aber beim Winkel $-\frac{\pi}{2}$ die imaginäre Achse und läuft schließlich unter dem Winkel $-\pi$ von links in den Ursprung.

Laut Nyquist-Kriterium beträgt die erforderliche Phasendrehung für die Hintereinanderschaltung von PI-Regler und PT_2 -Glied wie schon beim I-Regler wieder $\frac{\pi}{2}$. Denkt man sich wieder den Vektor, der im kritischen Punkt seinen Anfangspunkt hat und dessen Endpunkt dem Verlauf der Ortskurve folgt, so vollzieht dieser Vektor eine Drehung um $\frac{\pi}{2}$, und das Nyquist-Kriterium ist erfüllt. Man erkennt auch sofort, dass der prinzipielle Verlauf bezüglich -1 auch bei anderen Verstärkungsfaktoren nicht verändert wird. Unabhängig von der Wahl der Reglerv Verstärkung ist der geschlossene Kreis mit PT_2 -Glied und PI-Regler damit immer stabil.

Dies ist eine Folge des Vorhaltes, der beim PI-Regler gegenüber dem I-Regler hinzugetreten ist. Da die Ortskurve vieler realer Strecken für hohe Frequenzen in die Nähe des kritischen Punktes -1 kommt, ist es besonders wichtig, dass der Vorhalt des Reglers in diesem Frequenzbereich mit seiner positiven Phase die negative Phase des Integrators kompensiert, damit der Regler nicht noch eine zusätzliche negative Phasendrehung zur sowieso schon vorhandenen negativen Phase der Strecke verursacht.

Da die Reglerv Verstärkung im Hinblick auf die Stabilität offenbar beliebig gewählt werden kann, stellt sich nun die Frage, wie die Parameter des Reglers denn nun optimal bestimmt werden können. Ein besonders einfacher Entwurf ist dabei möglich, wenn man beim PT_2 -Glied von zwei reellen Polen ausgeht:

$$G(s) = \frac{V}{(T_1s + 1)(T_2s + 1)} \quad \text{mit} \quad T_2 > T_1 \quad (6.15)$$

Dann kann man den Regler so dimensionieren, dass sich sein Vorhalt gegen eine Polstelle der Strecke kürzen lässt und ein IT_1 -Glied entsteht (vgl. Kap. 4.3.5). Diese Maßnahme vereinfacht zunächst einmal die Kreisübertragungsfunktion

$$\begin{aligned} G(s)K(s) &= \frac{V}{(T_1s + 1)(T_2s + 1)} V_R \frac{T_{pi}s + 1}{T_{pi}s} \\ &= VV_R \frac{1}{T_2s(T_1s + 1)} \quad \text{mit} \quad T_{pi} = T_2 \end{aligned} \quad (6.16)$$

Im Interesse einer hohen Regelgeschwindigkeit ist die größere von beiden Zeitkonstanten gekürzt worden. Denn der zugehörige Einschwingvorgang $e^{-\frac{t}{T_2}}$ verläuft langsamer und sollte deshalb eliminiert werden. Das Kürzen von Zeitkonstanten macht aber natürlich nur dann Sinn, wenn die Zeitkonstante in der Übertragungsfunktion auch eine Entsprechung in der realen Strecke hat. Insbesondere bei einer Ersatzzeitkonstanten lässt sich eine Polkürzung

nicht durchführen, denn dort stellt der durch die Polstelle repräsentierte Einschwingvorgang nur eine vereinfachte Näherung für den tatsächlichen Einschwingvorgang dar.

Ein (kleines) Problem tritt auf, wenn die Zeitkonstanten der Strecke nicht exakt bekannt sind, was in der Praxis sogar der Normalfall ist. Dann kann man den Vorhalt des Reglers nicht so bestimmen, dass eine Kompensation mit dem gewünschten Pol erfolgt. Dies ist jedoch nicht weiter schlimm, solange es sich um einen stabilen Pol mit negativem Realteil handelt. Wie man leicht nachvollziehen kann, unterscheidet sich das Bode-Diagramm für den Fall $T_{pi} = T_2$ nur unwesentlich vom Bode-Diagramm für den Fall $T_{pi} \approx T_2$, und damit ist auch das Verhalten des geschlossenen Kreises annähernd gleich. Die auf der Polkürzung basierenden Ergebnisse werden auch für unvollständige Kompensation noch annähernd richtig sein.

Dies ist jedoch vollkommen anders bei einem instabilen Pol mit positivem Realteil. Wenn man beim Reglerentwurf einen derartigen Pol durch eine Nullstelle kompensiert, so mag das Ergebnis in der Theorie zwar gut aussehen, aber in der Praxis kann es zu großen Problemen führen. Denn eine Polstelle ist, wie oben schon gesagt, in den seltensten Fällen exakt bekannt. Die in der Theorie erfolgreiche Kompensation ist in der Praxis vermutlich unvollständig, so dass der instabile Pol in der realen Kreisübertragungsfunktion nach wie vor enthalten ist. Dies verändert dann aber die erforderliche Phasendrehung beim Nyquist-Kriterium über den Parameter r_k in Gleichung (5.19) massiv, und es ist davon auszugehen, dass die unter der Voraussetzung einer erfolgreichen Kompensation entworfene Kreisübertragungsfunktion diese Bedingung dann nicht erfüllt und der geschlossene Kreis in der Realität instabil ist. Von daher ist von der Kompensation einer instabilen Polstelle beim Reglerentwurf grundsätzlich abzuraten.

Nach der Festlegung $T_{pi} = T_2$ ist nun noch der verbliebene Reglerparameter V_R zu bestimmen. Dazu stellt man die Übertragungsfunktion des geschlossenen Kreises auf:

$$T(s) = \frac{y(s)}{w(s)} = \frac{G(s)K(s)}{G(s)K(s) + 1} = \frac{1}{\frac{T_1 T_2}{V V_R} s^2 + \frac{T_2}{V V_R} s + 1} \quad (6.17)$$

Der geschlossene Kreis ist offenbar ein PT_2 -Glieder, dessen Dämpfung D durch den Parameter V_R eingestellt werden kann. Ein Koeffizientenvergleich mit Gleichung (3.9) liefert

$$\frac{1}{\omega_0^2} = \frac{T_1 T_2}{V V_R} \quad \text{und} \quad \frac{2D}{\omega_0} = \frac{T_2}{V V_R} \quad (6.18)$$

Eliminiert man ω_0 , so ergibt sich

$$V_R = \frac{T_2}{4V T_1 D^2} \quad (6.19)$$

Durch die Wahl einer gewünschten Dämpfung D wird damit auch der zweite Reglerparameter festgelegt, und der geschlossene Kreis verhält sich wie ein

gewöhnliches PT_2 -Glied mit vorgegebener Dämpfung. Bezogen auf die Ortskurve der Kreisübertragungsfunktion bedeutet die Berechnung von V_R aus D , dass die Ortskurve einen bestimmten Abstand zum kritischen Punkt -1 einhält. Plausibel ist der Zusammenhang (6.19) auch: Je größer die Dämpfung D sein soll, desto kleiner muss die Reglerverstärkung V_R gewählt werden.

Als weiteres Beispiel sei nun angenommen, dass die Strecke selbst ein IT_1 -Glied (vgl. Kap. 4.3.5) ist, welches stationär genau geregelt werden soll und bei dem die Störgrößen an beliebigen Stellen angreifen können. Eine Verwendung des Integralanteils der Strecke zur Erzielung stationärer Genauigkeit (vgl. Abbildung 6.2) ist damit nicht möglich. Der ausgewählte Regler muss einen Integralanteil besitzen. In Frage kommen der I- oder der PI-Regler.

Um den Regler so einfach wie möglich zu gestalten, soll zunächst der I-Regler betrachtet werden. Mit diesem ergibt sich die Kreisübertragungsfunktion

$$G(s)K(s) = \frac{1}{T_1 s} \frac{1}{T_2 s + 1} \frac{1}{T_i s} \quad (6.20)$$

Bode-Diagramm und Ortskurve sind in Abbildung 6.14 gezeigt. Der Betragsverlauf ist im Bode-Diagramm nicht eingetragen. Wegen der beiden Integratoren geht er offensichtlich von ∞ gegen Null. Der Phasenverlauf entspricht dem einer einfachen Polstelle, wegen der beiden Integratoren aber um $-\pi$ nach unten versetzt. Für die Ortskurve ergibt sich ein Verlauf, der unter dem Winkel $-\pi$, also von links, aus dem Unendlichen kommt und unter dem Winkel $-\frac{3}{2}\pi$ in den Ursprung läuft.

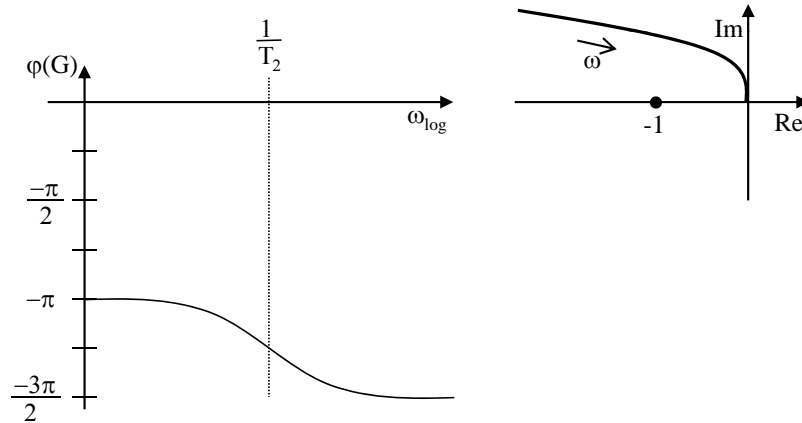


Abb. 6.14. Bode-Diagramm und Ortskurve eines IT_1 -Gliederes mit I-Regler

Laut Nyquist-Kriterium beträgt die erforderliche Phasendrehung um den Punkt -1 wegen der beiden Integratoren $+\pi$. Denkt man sich einen Vektor vom kritischen Punkt -1 zur Ortskurve, so wird dieser, dem Verlauf der Ortskurve folgend, zunächst nach links zeigen und sich dann im mathematisch

negativen Sinn nach rechts drehen. Die Phasendrehung beträgt damit $-\pi$, so dass der geschlossene Kreis instabil wäre. Offensichtlich ändert sich daran auch nichts, wenn man den Reglerparameter T_i verändert. Die grundsätzliche Form der Ortskurve und damit die Phasendrehung um -1 bleibt immer erhalten. Mit einem I-Regler kann die Strecke daher nicht stabilisiert werden.

Daher soll jetzt der PI-Regler untersucht werden. Als Kreisübertragungsfunktion ergibt sich mit diesem Regler

$$G(s)K(s) = \frac{1}{T_1 s} \frac{1}{T_2 s + 1} V_R \frac{T_R s + 1}{T_R s} = \frac{V_R}{T_1 T_R s^2} \frac{T_R s + 1}{T_2 s + 1} \quad (6.21)$$

Für den Regler sind nun zwei freie Parameter zu bestimmen, V_R und T_R . Wählt man zunächst $T_R < T_2$, so ergeben sich Bode-Diagramm und Ortskurve gemäß Abbildung 6.15. Die Phase der Nullstelle steigt in der Umgebung von $\frac{1}{T_R}$ von Null auf $+\frac{\pi}{2}$, und analog fällt die Phase der Polstelle in der Umgebung von $\frac{1}{T_2}$ von Null auf $-\frac{\pi}{2}$. Die beiden Integratoren steuern jeweils die konstante Phase $-\frac{\pi}{2}$ bei, so dass der Phasenverlauf durch die beiden Integratoren insgesamt um $-\pi$ abgesenkt wird. Der resultierende Phasenverlauf der Hintereinanderschaltung aller vier Komponenten ist gestrichelt eingezeichnet.

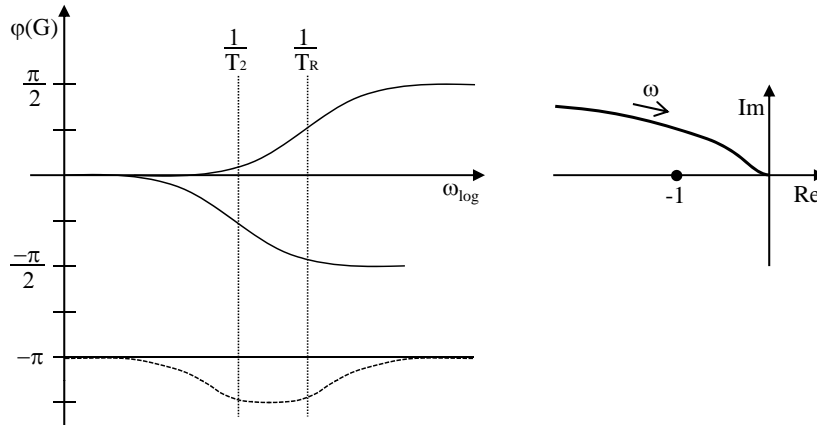


Abb. 6.15. Bode-Diagramm und Ortskurve eines IT_1 -Gliedes mit PI-Regler und $T_R < T_2$

Der Betragsverlauf beginnt wegen der Integratoren bei ∞ und endet bei Null. Damit beginnt die Ortskurve im Unendlichen unter einem Winkel von $-\pi$, kommt also von links, und läuft auch unter diesem Winkel von links in den Ursprung der komplexen Ebene. Durch die Kombination aus Null- und Polstelle erfolgt eine zwischenzeitliche Phasenabsenkung, die eine Verdrehung der Ortskurve im mathematisch negativen Sinn zur Folge hat. Bei einem reinen doppelten Integrator würde die Ortskurve auf der negativ reellen Achse verlaufen. Aufgrund der Phasenabsenkung verläuft sie nun aber im zweiten Quadranten.

Da die Zahl der Integratoren gegenüber dem Integralregler unverändert ist, beträgt die laut Nyquist-Kriterium erforderliche Phasendrehung wieder $+\pi$. Denkt man sich den Vektor, der vom kritischen Punkt zur Ortskurve führt, so dreht sich dieser Vektor mit dem Verlauf der Ortskurve wieder im mathematisch negativen Sinn von links nach rechts. Entsprechend beträgt auch hier die Phasendrehung um den Punkt -1 gerade $-\pi$. Eine Veränderung der Reglerverstärkung als letzten freien Parameter würde die Ortskurve strecken oder stauchen, am prinzipiellen Verlauf jedoch nichts ändern. Somit wäre auch diese Konfiguration im geschlossenen Kreis instabil.

Die durch Pol- und Nullstelle erfolgte, zwischenzeitliche Absenkung des Phasenverlaufes legt aber einen anderen Gedanken nahe: Aufgrund der Tatsache, dass T_R kleiner als T_2 gewählt wurde, wurde der Phasenverlauf der Kreisübertragungsfunktion zunächst durch die Polstelle bei $\omega = \frac{1}{T_2}$ ins Negative abgesenkt und erst danach durch die Regler-Nullstelle bei $\omega = \frac{1}{T_R}$ wieder auf den ursprünglichen Wert angehoben. Wenn man nun $T_R > T_2$ wählt, müsste sich genau der gegenteilige Effekt einstellen, womit sich dann die Ortskurve von der negativ reellen Achse ins Positive verdreht und im dritten Quadranten anstelle des zweiten verläuft.

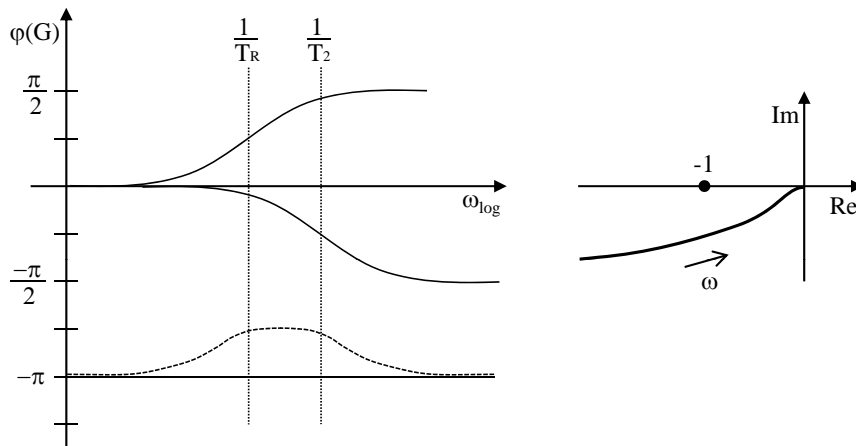


Abb. 6.16. Bode-Diagramm und Ortskurve eines IT_1 -Gliedes mit PI-Regler und $T_R > T_2$

Abbildung 6.16 zeigt die Verhältnisse. Der resultierende Phasenverlauf, der im linken Diagramm gestrichelt eingezeichnet ist, beginnt und endet bei $-\pi$. Zwischenzeitlich steigt die Phase an, fällt dann aber wieder ab. Der Betragsverlauf beginnt wegen der Integratoren bei ∞ und endet bei Null. Damit kommt die Ortskurve auch hier unter dem Winkel $-\pi$ aus dem Unendlichen von links und läuft unter dem Winkel $-\pi$ in den Ursprung. Zwischenzeitlich wird die Phase aber ins Positive verdreht, so dass die Kurve jetzt im dritten Quadranten verläuft.

Vergegenwärtigt man sich die Drehung des Vektors, der vom kritischen Punkt zur Ortskurve zeigt und deren Verlauf folgt, so wird klar, dass dieser Vektor sich zwar wieder von links nach rechts verdreht, diese Drehung jetzt aber in mathematisch positiver Richtung erfolgt. Damit beträgt die Phasendrehung um den kritischen Punkt nicht mehr $-\pi$, sondern $+\pi$, was genau dem geforderten Wert des Nyquist-Kriteriums entspricht.

Mit der Festlegung $T_R > T_2$ ist nun zunächst einmal die grundsätzliche Bedingung für die Stabilität des geschlossenen Kreises ermittelt worden. Im weiteren Verlauf können jetzt weitere Kriterien hinzugezogen werden, um T_R und V_R optimal im Sinne dieser Kriterien festzulegen.

Die nachfolgenden Betrachtungen werden zeigen, dass die Stabilitätsverhältnisse bei diesem System so kritisch sind, dass eine gute Dämpfung $D = 1$ oder $D = \frac{1}{\sqrt{2}}$ in der Regel nicht erzielt werden kann und man sich mit kleineren Dämpfungswerten zufrieden geben muss, die dann aber wenigstens möglichst groß sein sollten.

Das Kriterium für den nachfolgend vorgestellten Entwurf ist daher, die Dämpfung möglichst groß zu machen. Dies ist gemäß Satz 5.8 dann der Fall, wenn der Abstand der Ortskurve vom kritischen Punkt möglichst groß wird. Doch wie kann man diese Bedingung formal beschreiben, so dass sich V_R und T_R berechnen lassen? Hierzu werden zunächst einige neue Begriffe benötigt:

Definition 6.1 Die Mittenfrequenz ω_m zweier Frequenzen ω_1 und ω_2 bezeichnet den geometrischen Mittelwert $\omega_m = \sqrt{\omega_1 \omega_2}$ dieser beiden Frequenzen.

Die Durchtrittsfrequenz ω_d einer Kreisübertragungsfunktion $K(s)G(s)$ bezeichnet diejenige Frequenz, bei der der Betrag des Frequenzganges den Wert Eins annimmt: $|K(j\omega)G(j\omega)| = 1$

Der Durchtrittswinkel (oder auch die Phasenreserve) α_d ist definiert durch

$$\alpha_d = \varphi(G(j\omega_d)K(j\omega_d)) + \pi \quad (6.22)$$

Dabei ist $\varphi(G(j\omega_d)K(j\omega_d))$ die Phase des Frequenzganges bei der Durchtrittsfrequenz. α_d ist damit der Winkel zwischen der Ortskurve bei der Durchtrittsfrequenz und der negativ reellen Achse.

Die Bedeutung dieser Definitionen veranschaulicht Abbildung 6.17 anhand des hier untersuchten Systems. Die Mittenfrequenz ω_m ist hier das geometrische Mittel der Frequenzen $\frac{1}{T_R}$ und $\frac{1}{T_2}$. Diese Frequenz befindet sich im Bode-Diagramm aufgrund der logarithmischen Darstellung auch räumlich genau zwischen $\frac{1}{T_R}$ und $\frac{1}{T_2}$. Der Phasenverlauf der Kreisübertragungsfunktion ist offensichtlich symmetrisch zu ω_m . Und aus dem Grund kennzeichnet ω_m hier diejenige Frequenz, bei der die Phase des Frequenzganges ihren Maximalwert annimmt.

Die Bedeutung der Durchtrittsfrequenz ist aus der rechten Zeichnung zu entnehmen. Bei der Durchtrittsfrequenz ω_d ist der Betrag des Frequenzganges gleich Eins. Damit muss die Ortskurve bei dieser Frequenz zum Ursprung des

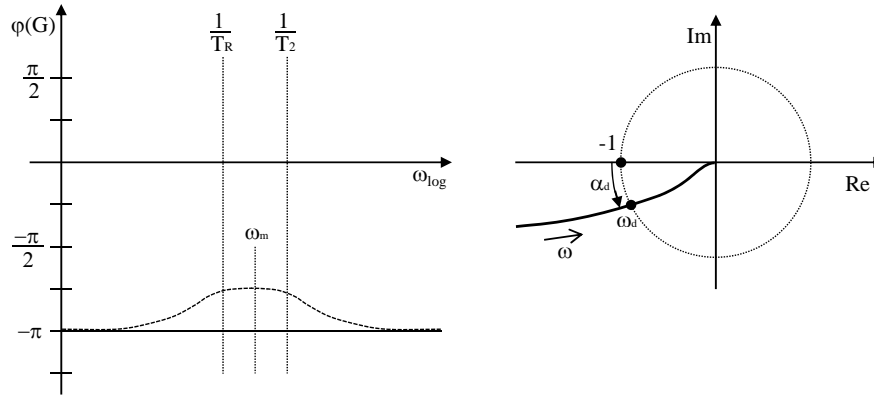


Abb. 6.17. Anschauliche Darstellung von Mitten- und Durchtrittsfrequenz

Koordinatensystems gerade den Abstand Eins haben. Zeichnet man einen Einheitskreis um den Ursprung, so schneidet die Ortskurve bei der Frequenz ω_d diesen Kreis. α_d ist gemäß der Definition gerade der eingezeichnete Winkel zwischen Ortskurve und negativ reeller Achse.

Die Entwurfsidee ist nun, den Regler so zu dimensionieren, dass die Mittenfrequenz ω_m , bei der die maximale Phase auftritt, gerade mit der Durchtrittsfrequenz ω_d zusammenfällt. Damit hat die Phase der Ortskurve gerade dann ihren maximalen Wert und damit die Ortskurve selbst einen besonders großen Abstand zur reellen Achse, wenn sie in der Nähe des kritischen Punktes -1 ist.

Hierzu wird angesetzt:

$$T_R = a^2 T_2 \quad (6.23)$$

mit dem zunächst frei wählbaren Entwurfparameter a . Daraus folgt entsprechend der Definition von ω_m

$$\omega_m = \sqrt{\frac{1}{T_R} \frac{1}{T_2}} = \sqrt{\frac{1}{a^2 T_2} \frac{1}{T_2}} = \frac{1}{a T_2} \quad (6.24)$$

Für die Durchtrittsfrequenz gilt

$$1 = |G(j\omega_d)K(j\omega_d)| = \frac{V_R}{T_1 T_R \omega_d^2} \frac{\sqrt{T_R^2 \omega_d^2 + 1}}{\sqrt{T_2^2 \omega_d^2 + 1}} \quad (6.25)$$

Mit $\omega_d = \omega_m$ und (6.24) wird daraus

$$1 = \frac{V_R}{T_1 T_R (\frac{1}{a T_2})^2} \frac{\sqrt{T_R^2 (\frac{1}{a T_2})^2 + 1}}{\sqrt{T_2^2 (\frac{1}{a T_2})^2 + 1}} = \frac{V_R a^2 T_2^2}{T_1 T_R} \frac{\sqrt{(\frac{T_R}{a T_2})^2 + 1}}{\sqrt{\frac{1}{a^2} + 1}} \quad (6.26)$$

Einsetzen von (6.23) liefert schließlich

$$1 = \frac{V_R T_2}{T_1} \frac{\sqrt{a^2 + 1}}{\sqrt{\frac{1}{a^2} + 1}} = \frac{V_R T_2 a}{T_1} \quad (6.27)$$

und daher

$$V_R = \frac{T_1}{T_2 a} \quad (6.28)$$

Mit den Festlegungen (6.23) für T_R und (6.28) für V_R ist sichergestellt, dass $\omega_d = \omega_m$ ist.

Den Entwurfsparameter a kann man nun dazu nutzen, einen vorgegebenen Wert für die Dämpfung einzustellen. Zunächst ergibt sich mit den Festlegungen für T_R und V_R aus (6.21) die Kreisübertragungsfunktion

$$G(s)K(s) = \frac{T_1}{T_2 a} \frac{1}{T_1 a^2 T_2 s^2} \frac{a^2 T_2 s + 1}{T_2 s + 1} = \frac{1}{a^3 (T_2 s)^2} \frac{a^2 T_2 s + 1}{T_2 s + 1} \quad (6.29)$$

und mit Einführung einer normierten komplexen Frequenz $q = a T_2 s$ als abkürzende Schreibweise

$$G(s)K(s) = G(q)K(q) = \frac{1}{a q^2} \frac{a q + 1}{\frac{q}{a} + 1} \quad (6.30)$$

Damit ergibt sich als Übertragungsfunktion des geschlossenen Kreises

$$T(q) = \frac{G(q)K(q)}{1 + G(q)K(q)} = \frac{a q + 1}{q^3 + a q^2 + a q + 1} \quad (6.31)$$

Offensichtlich ist $q = -1$ eine Polstelle des Systems, so dass sich im Nenner der Faktor $(q + 1)$ abspalten lässt:

$$T(q) = \frac{a q + 1}{(q + 1)(q^2 + (a - 1)q + 1)} \quad (6.32)$$

Der Pol bei -1 ist nicht schwingungsfähig und spielt damit für die optimale Dämpfung keine Rolle. Für das verbleibende Polynom zweiter Ordnung lässt sich ein Koeffizientenvergleich mit der Normalform des Nenners eines PT_2 -Gliedes durchführen:

$$q^2 + (a - 1)q + 1 = \frac{1}{\omega_0^2} s^2 + \frac{2D}{\omega_0} s + 1 \quad (6.33)$$

Es folgt $\omega_0 = 1$ und $2D = a - 1$ und damit für den Entwurfsparameter a

$$a = 2D + 1 \quad (6.34)$$

Zusammengefasst mit (6.23) und (6.28) ergibt sich als Resultat des Reglerentwurfs für die Reglerkoeffizienten bei Vorgabe einer Dämpfung D :

$$\begin{aligned}
a &= 2D + 1 \\
T_R &= a^2 T_2 \\
V_R &= \frac{T_1}{T_2 a}
\end{aligned} \tag{6.35}$$

Damit ist der PI-Regler vollständig dimensioniert. Da die Grundidee dieses Entwurfes auf dem zu ω_m symmetrischen Phasenverlauf basiert, bezeichnet man dieses Verfahren auch als *Symmetrisches Optimum*.

Als letztes Beispiel soll auch hier wieder das PT_3 -Glieder betrachtet werden, um den Vergleich mit den anderen Reglertypen zu ermöglichen. Die Kreisübertragungsfunktion lautet:

$$G(s)K(s) = \frac{1}{T_1 s + 1} \frac{1}{T_2 s + 1} \frac{1}{T_3 s + 1} V_R \frac{T_R s + 1}{T_R s} \quad T_1 > T_2 > T_3 \tag{6.36}$$

Natürlich ist es im ersten Ansatz naheliegend, mit dem Reglervorhalt die größte der drei Zeitkonstanten zu kürzen ($T_R = T_1$). Übrig bliebe ein PT_2 -Glieder mit einem Integralregler, dessen Integrierzeit mit V_R eingestellt werden kann:

$$G(s)K(s) = \frac{1}{T_2 s + 1} \frac{1}{T_3 s + 1} \frac{1}{\frac{T_1}{V_R} s} \tag{6.37}$$

Da die Ordnung der Strecke um Eins reduziert ist, wird man mit diesem Ansatz auf jeden Fall schon einmal ein besseres Ergebnis erzielen als mit dem I-Regler, da dieser ja eine Strecke dritter Ordnung regeln muss.

Falls die Zeitkonstanten der PT_1 -Glieder aber eine sehr unterschiedliche Größenordnung aufweisen, wird eine andere Vorgehensweise zu besseren Ergebnissen führen. Dazu sind zunächst zwei Abschätzungen erforderlich. Zum einen nähert sich im Bereich hoher Frequenzen oder sehr großer Zeitkonstanten das Verhalten eines PT_1 -Gliedes dem Verhalten eines Integrators an:

$$\frac{1}{Ts + 1} \approx \frac{1}{Ts} \quad \text{für} \quad Ts \gg 1 \tag{6.38}$$

Analog gilt für niedrige Frequenzen oder sehr kleine Zeitkonstanten:

$$\frac{1}{Ts + 1} \approx 1 \quad \text{für} \quad Ts \ll 1 \tag{6.39}$$

Nun ist bei der Auslegung eines Reglers vor allem der Frequenzbereich interessant, bei dem die Ortskurve der Kreisübertragungsfunktion dem kritischen Punkt -1 nahe kommt. Für genau diesen Frequenzbereich ist zu prüfen, ob T_1 groß genug ist, damit die Abschätzung (6.38) zulässig ist, und ob T_3 klein genug ist, damit die Abschätzung (6.39) gilt.

Wenn dies der Fall ist, dann lassen sich die PT_1 -Glieder mit T_1 und T_3 entsprechend vereinfachen, und es ergibt sich für die Kreisübertragungsfunktion

$$G(s)K(s) = \frac{1}{T_1 s} \frac{1}{T_2 s + 1} V_R \frac{T_R s + 1}{T_R s} \quad (6.40)$$

Damit hat man das Problem auf den Fall eines PI-Reglers mit einem IT_1 -Glied zurückgeführt, für den es mit dem symmetrischen Optimum eine hervorragende und einfach zu berechnende Lösung gibt.

6.2.4 PID-Regler

Die Stellgröße des PI-Reglers setzt sich aus zwei Anteilen zusammen, einem Integralanteil für die Genauigkeit und einem Proportionalanteil zur Erhöhung der Regelgeschwindigkeit. Eine weitere Verbesserung des Regelverhaltens ist zu erwarten, wenn eine Regelabweichung nicht erst dann bekämpft wird, wenn sie schon existiert, wie es durch den Proportionalanteil geschieht, sondern am besten schon dann, wenn sie im Entstehen ist. Zu diesem Zweck kann man den PI-Regler um einen Differentialanteil erweitern, und man erhält einen PID-Regler:

$$K(s) = P + I \frac{1}{s} + Ds \quad (6.41)$$

Ein idealer Differenzierer mit der Übertragungsfunktion s ist aber weder realisierbar noch erwünscht. Denn ein Faktor s in einem Summanden bedeutet, dass der Summand umso größere Werte annimmt, je höher die Frequenz ist. Wegen dieser Hochpasseigenschaft verstärkt ein idealer Differenzierer daher die in der Praxis immer vorhandenen hochfrequenten Rauschsignale, was natürlich vermieden werden sollte. Bei einem realen PID-Regler ist deshalb der D-Anteil mit der (sehr kleinen) Zeitkonstanten T_v verzögert:

$$K(s) = P + I \frac{1}{s} + D \frac{s}{T_v s + 1} = V_R \frac{T_1 s + 1}{T_1 s} \frac{T_2 s + 1}{T_v s + 1} \quad (6.42)$$

Wie man sieht, lässt sich der PID-Regler auch als Reihenschaltung von PI-Regler und einem rationalen Übertragungsglied erster Ordnung, einem sogenannten DT_1 -Glied, auffassen.

Die Vorteile des PID-Reglers gegenüber dem PI-Regler lassen sich anhand seiner Sprungantwort und der Ortskurve erklären. Die Sprungantwort lautet für den idealen PID-Regler

$$y(s) = \frac{1}{s} K(s) = P \frac{1}{s} + I \frac{1}{s^2} + D \quad \leftrightarrow \quad y(t) = P + It + D\delta(t) \quad (6.43)$$

und für den realen PID-Regler

$$y(s) = P \frac{1}{s} + I \frac{1}{s^2} + D \frac{1}{T_v s + 1} \quad \leftrightarrow \quad y(t) = P + It + D \frac{1}{T_v} e^{-\frac{t}{T_v}} \quad (6.44)$$

Beide Verläufe sind in Abbildung 6.18 links dargestellt, wobei die Sprungantwort des idealen PID-Regler gestrichelt gezeichnet ist. Beim idealen PID-Regler ergibt sich bei $t = 0$ ein unendlich hoher Impuls als Ableitung der

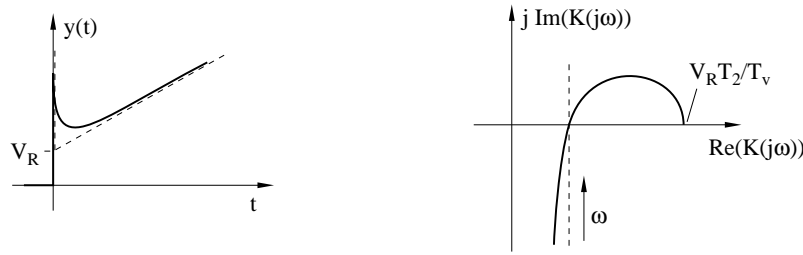


Abb. 6.18. Sprungantwort und Ortskurve eines PID-Reglers

sprungförmigen Eingangsgröße. Beim realen PID-Regler ist dieser Impuls zu einer abklingenden Exponentialfunktion mit dem Anfangswert $\frac{D}{T_v}$ geworden.

Die Sprungantwort zeigt, dass der Regler genau das geforderte Verhalten aufweist: Eine (hier sprungförmige) Regelabweichung wird durch den D-Anteil in ihrer Anfangsphase sehr heftig bekämpft, während sich der Regler im weiteren Verlauf wie ein PI-Regler verhält.

Die Ortskurve des idealen PID-Reglers kann ebenfalls aus der Summen-darstellung entwickelt werden. Für den Frequenzgang ergibt sich

$$K(j\omega) = P - jI\frac{1}{\omega} + jD\omega = P + j(D\omega - I\frac{1}{\omega}) \quad (6.45)$$

Der Realteil ist konstant gleich P , während der Imaginärteil von $-\infty$ bis $+\infty$ läuft. Es ergibt sich die gestrichelte Linie in der rechten Zeichnung in Abbildung 6.18.

Für die Ortskurve des realen PID-Reglers ist zunächst das Bode-Diagramm zu zeichnen (Abb. 6.19). Im Betragsdiagramm sieht man die fallende Betragskurve des Integralanteils, der die Frequenzachse bei $\frac{1}{T_1}$ schneidet, die beiden nach oben abknickenden Verläufe für die beiden Vorhalte sowie den nach unten abknickenden Verlauf für die Polstelle bei $-\frac{1}{T_v}$. In Summe ergibt sich zunächst ein aus dem unendlichen kommender, fallender Verlauf, dessen Abfallen bei $\frac{1}{T_1}$ durch den ersten Vorhalt kompensiert wird. Beim zweiten Vorhalt beginnt die Betragskurve zu steigen, und dieser Anstieg wird dann durch die Polstelle kompensiert. Damit ergibt sich ein von Null verschiedener Endwert.

Im Phasendiagramm sieht man den konstanten Phasenverlauf des Integralanteils bei $-\frac{\pi}{2}$, die ansteigenden Phasenverläufe der beiden Vorhalte sowie den abfallenden Phasenverlauf der Polstelle. Für den gesamten Phasenverlauf ergibt sich ein Beginn bei $-\frac{\pi}{2}$. Durch die beiden Vorhalte steigt dann die Phase auf positive Werte, bevor der Anstieg durch die Polstelle wieder rückgängig gemacht wird, so dass der Endwert gleich Null ist.

Für die Ortskurve in Abbildung 6.18 folgt daraus, dass sie im Unendlichen unter einem Winkel von $-\frac{\pi}{2}$ beginnen muss, also von unten kommt. Die Phase ist größer als $-\frac{\pi}{2}$, so dass die Kurve im vierten Quadranten verläuft. Sobald die Werte der Phase positiv werden, schneidet die Ortskurve die reelle Achse, verläuft dann mit steigender Phase und steigendem Betrag im ersten

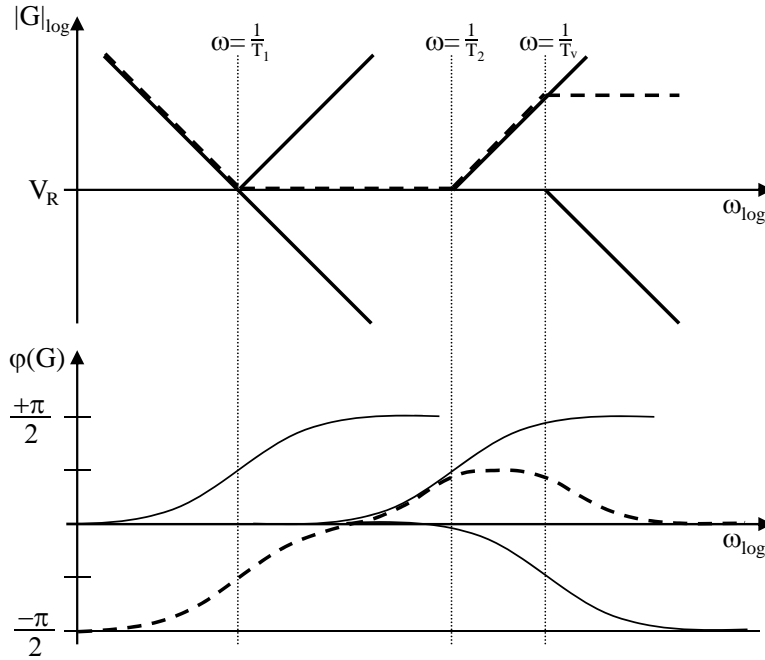


Abb. 6.19. Bode-Diagramm eines PID-Reglers

Quadranten und konvergiert schließlich mit der Phase Null gegen einen reellen Endwert.

Dieser Endwert lässt sich durch Bilden des Grenzwertes leicht aus der Übertragungsfunktion ermitteln:

$$K(\infty) = \lim_{\omega \rightarrow \infty} V_R \frac{T_1 j\omega + 1}{T_1 j\omega} \frac{T_2 j\omega + 1}{T_v j\omega + 1} = V_R \frac{T_2}{T_v} \quad (6.46)$$

Lohnenswert ist eine separate Betrachtung desjenigen Anteils, der beim PID-Regler gegenüber dem PI-Regler hinzutritt:

$$\frac{T_2 s + 1}{T_v s + 1} \quad \text{mit} \quad T_2 > T_v \quad (6.47)$$

Man bezeichnet diesen Anteil als *Phasenkorrekturglied*. Die Einführung eines zusätzlichen Verstärkungsfaktors führt auf den sogenannten *PD-Regler* mit der Übertragungsfunktion

$$K(s) = V_R \frac{T_2 s + 1}{T_v s + 1} = P + D \frac{s}{T_v s + 1} \quad (6.48)$$

Der PD-Regler besteht aus einem Proportional- und einem Differentialanteil, wobei der Differentialanteil nicht ideal, sondern nur mit einer Verzögerung

$\frac{1}{T_v s + 1}$ realisiert werden kann. Stationäre Genauigkeit kann mit diesem Regler wegen $K(0) \neq \infty$ nur bei integrierenden Strecken erzielt werden. Wegen des Differentialanteils reagiert er auf eine Änderung der Regelabweichung stärker als ein P-Regler. Sein Einsatz bietet sich daher an, wenn Genauigkeit nicht so wichtig oder durch einen Integrator in Strecke bzw. Stellglied gewährleistet und die mit einem P-Regler erreichbare Regelgeschwindigkeit nicht groß genug ist.

Das Bode-Diagramm des Phasenkorrekturgliedes bzw. des PD-Reglers ist in Abbildung 6.20 gezeigt. Die Phase wird durch den Vorhalt angehoben und durch die Polstelle wieder abgesenkt. Dieser Effekt begründet den Vorteil des PID-Reglers gegenüber dem PI-Regler. Durch die Phasenhebung in einem bestimmten Frequenzbereich ist es nämlich möglich, den Phasenverlauf der Kreisübertragungsfunktion in eben diesem Frequenzbereich anzuheben und damit die Ortskurve im mathematisch positiven Sinn vom kritischen Punkt -1 wegzudrehen.

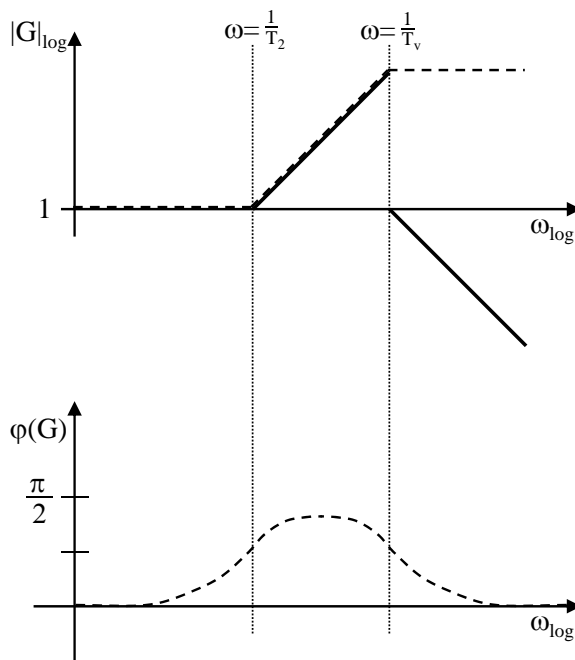


Abb. 6.20. Bode-Diagramm eines Phasenkorrekturgliedes bzw. eines PD-Reglers

Der Nachteil ergibt sich aus dem Betragsverlauf. Man sieht, dass durch die Nullstelle eine Betragsanhebung erfolgt, die dann erst durch die Polstelle wieder kompensiert wird. Dadurch, dass die Anhebung früher erfolgt als die Kompensation, ergibt sich insgesamt für höhere Frequenzen durch das Phasenkorrekturglied eine Betragsanhebung.

Diese kann für die Stabilität des Systems kritisch werden. Denn durch die Betragsanhebung für hohe Frequenzen wird die Ortskurve der Kreisübertragungsfunktion in diesem Frequenzbereich gedehnt und damit näher an den kritischen Punkt -1 gebracht. Die Gefahr, dass der kritische Punkt auf der falschen Seite passiert wird, steigt.

Man sieht, dass das Phasenkorrekturglied mit Vorsicht einzusetzen ist. Dem positiven Effekt einer Phasenhebung steht der negative Effekt einer Betragsanhebung gegenüber. Letztendlich muss man durch geeignete Dimensionierung der beiden Parameter T_2 und T_v dafür sorgen, dass ein stabiles und ausreichend gedämpftes System entsteht.

Angemerkt sei, dass die geeignete Dimensionierung von T_2 und T_v sogar so aussehen kann, dass $T_2 < T_v$ zu wählen ist. Man nimmt dann eine Phasenabsenkung in Kauf, weil der Nutzen der gleichzeitigen Betragsabsenkung größer ist als der Schaden, den die Phasenabsenkung verursacht.

Wesentlich ist, dass man beim PID-Regler durch das Phasenkorrekturglied gegenüber dem PI-Regler zwei weitere, einstellbare Parameter T_2 und T_v gewinnt, die hilfreich sein können, um ein gegebenes System zu stabilisieren oder weitergehende Forderungen an die Regelung zu realisieren. Aufgrund dieser beiden zusätzlichen Parameter ist der PID-Regler leistungsfähiger als ein PI-Regler, aber natürlich auch schwieriger zu dimensionieren. Während PI-Regler häufig ohne Rechnung durch Ausprobieren eingestellt werden, ist dies bei einem PID-Regler kaum möglich, vor allem nicht, wenn eine optimale Einstellung angestrebt wird, die die Möglichkeiten des Reglers voll ausschöpft.

Zum Abschluss soll auch hier wieder das PT_3 -Glied betrachtet werden. Die Kreisübertragungsfunktion lautet jetzt:

$$G(s)K(s) = \frac{1}{T_1s + 1} \frac{1}{T_2s + 1} \frac{1}{T_3s + 1} V_R \frac{T_{R1}s + 1}{T_{R1}s} \frac{T_{R2}s + 1}{T_{Rv}s + 1} \quad (6.49)$$

mit $T_1 > T_2 > T_3$. Der letzte Term, also das Phasenkorrekturglied, wird jetzt dazu genutzt, um die größte Zeitkonstante zu eliminieren ($T_{R2} = T_1$). Man erhält

$$G(s)K(s) = \frac{1}{T_{Rv}s + 1} \frac{1}{T_2s + 1} \frac{1}{T_3s + 1} V_R \frac{T_{R1}s + 1}{T_{R1}s} \quad (6.50)$$

Damit ist das Problem auf die Auslegung eines PI-Reglers mit PT_3 -Glied zurückgeführt, wobei sich die Verhältnisse durch die verkleinerten Zeitkonstanten vereinfacht haben und somit bessere Regelergebnisse erzielt werden können.

Dies ist auch in Abb. 6.6 zu erkennen. Der PID-Regler ist von den Regelergebnissen her der beste der vorgestellten Regler. Prinzipiell lässt sich sagen, dass mit zunehmender Komplexität des Reglers die Regelergebnisse immer besser werden. Dies ist nicht verwunderlich, da mehr Freiheitsgrade zur Verfügung stehen, um die gegensätzlichen Forderungen nach Stabilität und ausreichender Dämpfung einerseits sowie Schnelligkeit andererseits zu erfüllen.

An Gleichung (6.41) bzw. (6.42) ist zu erkennen, dass der PID-Regler als Sonderfälle auch die anderen vorgestellten Regler enthält. Je nachdem, ob I , P oder D zu Null gesetzt werden, erhält man einen P-, I- oder PI-Regler. Aus diesem Grund wird der Ausdruck *PID-Regler* häufig als Sammelbegriff für alle hier vorgestellten Regler verwendet.

6.2.5 Realisierung mit Operationsverstärkern

Obwohl heutzutage die meisten Regler mit Hilfe von Software und Signalprozessoren realisiert werden, soll hier kurz gezeigt werden, wie sich ein Regler auch deutlich billiger mit Hilfe von Operationsverstärkern realisieren lässt.

Abbildung 6.21 zeigt zunächst die grundsätzliche Beschaltung eines Operationsverstärkers.

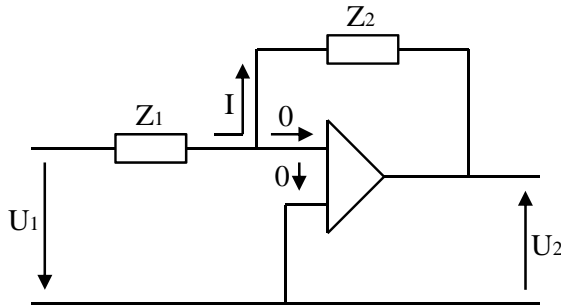


Abb. 6.21. Grundsätzliche Beschaltung eines Operationsverstärkers

Eingangsstrom und Eingangsspannung des Verstärkers können näherungsweise zu Null angenommen werden. Damit fließt der Strom I vollständig durch die Impedanzen Z_1 und Z_2 , und das Potential auf beiden Verstärkereingängen ist gleich. Berücksichtigt man, dass die Spannung U_2 am Ausgang von unten nach oben eingezeichnet ist, so gilt

$$\frac{U_1}{Z_1} = I = \frac{U_2}{Z_2} \quad \rightarrow \quad \frac{U_2}{U_1} = \frac{Z_2}{Z_1} \quad (6.51)$$

Auf dieser Grundlage lässt sich nun auch sofort eine mögliche Schaltung für einen PID-Regler angeben, wie sie in Abbildung 6.22 gezeigt ist.

Für die beiden Verstärker ergibt sich zunächst separat

$$\begin{aligned} \frac{U_H}{U_1} &= \frac{\frac{1}{sC_1} + R_2}{R_1} = \frac{1 + sC_1R_2}{sC_1R_1} \\ \frac{U_2}{U_H} &= \frac{\frac{1}{sC_2} + R_4}{\frac{1}{sC_2} + R_3} = \frac{1 + sC_2R_4}{1 + sC_2R_3} \end{aligned} \quad (6.52)$$

und damit für das gesamte Übertragungsverhalten

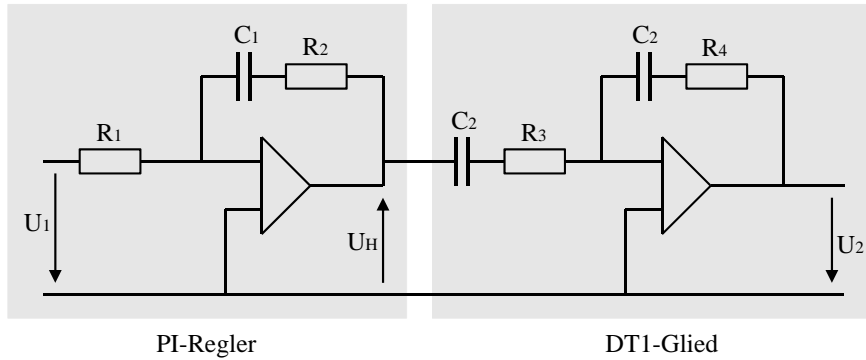


Abb. 6.22. PID-Regler mit Operationsverstärkern

$$\frac{U_2}{U_1} = \frac{1 + sC_1R_2}{sC_1R_1} \frac{1 + sC_2R_4}{1 + sC_2R_3} \quad (6.53)$$

Ein Vergleich mit der Definitionsgleichung (6.42) des PID-Reglers

$$K(s) = V_R \frac{T_1s + 1}{T_1s} \frac{T_2s + 1}{T_vs + 1} \quad (6.54)$$

zeigt sofort den Zusammenhang zwischen den Schaltungselementen und den Reglerparametern:

$$\begin{aligned} T_1 &= C_1R_2 \\ T_2 &= C_2R_4 \\ T_v &= C_2R_3 \\ V_R &= \frac{R_2}{R_1} \end{aligned} \quad (6.55)$$

6.3 Weitere Reglerentwurfsverfahren

In der Praxis erfolgt die Berechnung der Reglerparameter nach entsprechender Vereinfachung der Streckenübertragungsfunktion oft anhand einfacher Überlegungen mit Hilfe von Bode-Diagramm und Ortskurve, wie dies in den vorangegangenen Kapiteln gezeigt wurde. Es gibt aber natürlich noch andere Entwurfsverfahren, von denen einige in diesem Kapitel vorgestellt werden sollen.

6.3.1 Wurzelortskurven

Beim Wurzelortungsverfahren muss zunächst anhand der gleichen Überlegungen, wie sie auch in den vorangegangenen Kapiteln skizziert wurden, der geeignete

Reglertyp ausgewählt werden. Auf Bode-Diagramm und Ortskurve ist also auch hier nicht zu verzichten. Anschließend erfolgt durch das Wurzelortsverfahren die Auswahl einer optimalen Parameterkonfiguration, wobei dies unter Vorbehalt zu sehen ist, wie nachfolgend noch erläutert wird.

Zunächst werden diejenigen Reglerparameter festgelegt, die sich aufgrund bestimmter Anforderungen sofort ergeben, also beispielsweise durch das Kürzen einer stabilen Polstelle. Hier gelten grundsätzlich dieselben Überlegungen, wie sie auch in den vorangegangenen Kapiteln zu den einzelnen Reglertypen skizziert wurden. Nur ein frei wählbarer Parameter darf übrig bleiben, meist die Reglerverstärkung.

Dann wird die Lage der Pole des geschlossenen Kreises in Abhängigkeit von diesem einstellbaren Reglerparameter angegeben. Es ergibt sich für jede Polstelle in Abhängigkeit von dem Reglerparameter eine Kurve in der komplexen Ebene, die sogenannte *Wurzelortskurve*. Aus diesen Kurven lässt sich für jeden Wert des Reglerparameters eine bestimmte Polkonfiguration ablesen. Nach bestimmten Kriterien wie z.B. vorgegebener Dämpfung wird dann eine dieser Polkonfigurationen ausgewählt, und damit ist der Wert des Reglerparameters festgelegt.

Nachteilig an diesem Verfahren ist, dass es die Festlegung des Reglers auf die Auswahl eines einzigen frei wählbaren Parameters reduziert. Die anderen Parameter werden nach gewissen Kriterien vorab festgelegt. Sofern diese Festlegung nicht optimal war, kann auch der am Ende entstandene Regler nicht optimal sein. Hier ist es offenbar deutlich sinnvoller, mit Bode-Diagramm und Ortskurve der Kreisübertragungsfunktion zu arbeiten, da man dabei die Wirkung aller Reglerparameter im Blick behält.

Dazu kommt der notwendige Rechenaufwand zur Bestimmung der Wurzelortskurven für jeden einzelnen Pol. Angesichts der Tatsache, dass es heutzutage mit Hilfe von MATLAB oder anderen Software-Werkzeugen möglich ist, einen optimalen Zustandsregler quasi auf Knopfdruck zu berechnen, verlieren klassische Verfahren, die zwar ebenfalls den Einsatz eines Computers erfordern, am Ende aber nur einen suboptimalen Regler liefern, ihre Existenzberechtigung. Unter diesem Gesichtspunkt ist das Wurzelortsverfahren als überholt anzusehen und wird deshalb hier nicht näher dargestellt.

Der Reglerentwurf mittels Bode-Diagramm und Ortskurve ist dagegen ohne Computer noch von Hand durchführbar. Er bietet die Möglichkeit, schnell auf einem Blatt Papier den richtigen Reglertyp und auch eine brauchbare Parameterkonfiguration zu ermitteln. Dies ist in Situationen, in denen schnelle Entscheidungen notwendig sind, sehr hilfreich. Derartige Situationen treten beispielsweise bei der Inbetriebnahme von verfahrenstechnischen Anlagen in der Chemie oder in der Energieerzeugung auf. Hier sind innerhalb kürzester Zeit hunderte oder sogar tausende von Reglern einzustellen, und es wäre unmöglich, für jeden einzelnen Regler eine computergestützte Analyse durchzuführen.

6.3.2 Reglerentwurf nach Ziegler-Nichols

Der Reglerentwurf nach *Ziegler-Nichols* ist ein Beispiel für ein vollständig schematisiertes Entwurfsverfahren, von denen es auch noch andere gibt (z.B. nach Chien, Hrones, Reswick). Anhand der gemessenen Sprungantwort der Strecke werden gewisse Kenndaten wie z.B. die Höhe des ersten Überschingers oder die Zeit, die verstreicht, bis der Wert der Ausgangsgröße 10, 50, oder 90 Prozent vom Sollwert erreicht hat, ermittelt, aus denen dann mit Hilfe von vorgegebenen Tabellen die Reglerparameter festgelegt werden. Die regelungstechnischen Überlegungen, die zu diesen Tabellen führten, sind nicht ohne weiteres nachzuvollziehen, so dass der Anwender im Falle eines Misserfolges kaum weiß, wie die Parameter zu modifizieren sind.

Von daher kann man im praktischen Anwendungsfall mit diesem Verfahren durchaus einen ersten Versuch zur Auslegung eines geeigneten Reglers durchführen. Verlassen sollte man sich auf dieses Verfahren jedoch nicht. Denn wenn der ermittelte Regler keine ausreichend gute Regelqualität liefert, bieten die Ziegler-Nichols-Formeln keinerlei Anhalt für eine mögliche Modifikation der Reglerparameter.

6.3.3 Kompensationsregler und Youla-Parametrierung

Alle bisher vorgestellten Auslegungsverfahren haben gemeinsam, dass die Struktur des Reglers festgelegt werden muss und sich nur seine Parameter aus den jeweiligen Verfahren ergeben. Und es ist offensichtlich, dass nicht mit jedem Regler jede Strecke stabilisiert werden kann. So lässt sich beispielsweise eine aus zwei hintereinandergeschalteten Integratoren bestehende Strecke prinzipiell nur mit dem PID-, nicht aber mit einem P-, I- oder PI-Regler regeln. Vor der Dimensionierung eines Reglers steht also in allen Fällen die Analyse der Strecke und die Festlegung eines geeigneten Reglertyps.

Dieser Schritt entfällt bei den sogenannten analytischen Verfahren. Hier liefert das Verfahren nicht nur die Reglerparameter, sondern auch die Struktur des Reglers, also die gesamte Regler-Übertragungsfunktion, die sich deutlich von der Struktur eines PID-Reglers unterscheiden kann.

Beim *Kompensationsregler nach Truxal-Guillemain* (vgl. [63]) wird für den geschlossenen Kreis zunächst eine Modell-Übertragungsfunktion $M(s)$ vorgegeben, die sich ihrerseits aus bestimmten Anforderungen an das Einschwingverhalten des geschlossenen Kreises ergibt:

$$T(s) = \frac{G(s)K(s)}{1 + G(s)K(s)} \stackrel{!}{=} M(s) \quad (6.56)$$

Für den Regler ergibt sich damit sofort

$$K(s) = \frac{1}{G(s)} \frac{M(s)}{1 - M(s)} \quad (6.57)$$

Der Regler enthält also die invertierte Streckenübertragungsfunktion, so dass der Einfluss der Strecke in der Kreisübertragungsfunktion $G(s)K(s)$ und damit auch in der Übertragungsfunktion des geschlossenen Kreises vollständig eliminiert ist.

$M(s)$ muss allerdings so gewählt werden, dass ein realisierbarer Regler entsteht, bei dem die Ordnung des Nennerpolynoms größer oder gleich der Ordnung des Zählerpolynoms ist. Wenn bei $G(s)$ also die Ordnung des Nenners um k größer ist als die Ordnung des Zählers, dann muss bei $M(s)$ wegen des Kehrwertes von $G(s)$ in Gleichung (6.57) die Ordnung des Nenners die des Zählers ebenfalls um mindestens k übersteigen.

Dies macht auch physikalisch Sinn. Denn die Forderung bedeutet doch zum Beispiel, dass für ein Verzögerungsglied der Ordnung n als Strecke auch die Modellfunktion ein Verzögerungsglied derselben Ordnung sein muss. Man kann damit vom geschlossenen Kreis nicht mehr fordern als die Strecke selbst zulässt. Ein Schiff kann eben nicht mit derselben Geschwindigkeit einen Kurswechsel durchführen wie ein Kraftfahrzeug.

An Gleichung (6.57) kann man weiterhin erkennen, dass die Grundidee des Verfahrens darauf beruht, dass der Regler mit $\frac{1}{\tilde{G}(s)}$ die Strecke vollständig kompensiert. Das kann jedoch nur dann gelingen, wenn $G(s)$ exakt bekannt ist. Dies ist allerdings in der Praxis selten der Fall. Es ist eher die Regel, dass die Lage der Pole und Nullstellen von $G(s)$ nicht exakt bestimmt werden kann. Für den Regler lautet damit die Übertragungsfunktion

$$K(s) = \frac{1}{\tilde{G}(s)} \frac{M(s)}{1 - M(s)} \quad (6.58)$$

mit dem von der realen Strecke $G(s)$ abweichenden Modell $\tilde{G}(s)$, und für den geschlossenen Kreis ergibt sich

$$\begin{aligned} T(s) &= \frac{G(s)K(s)}{1 + G(s)K(s)} = \frac{G(s) \frac{1}{\tilde{G}(s)} \frac{M(s)}{1 - M(s)}}{1 + G(s) \frac{1}{\tilde{G}(s)} \frac{M(s)}{1 - M(s)}} \\ &= \frac{G(s)M(s)}{\tilde{G}(s) + (G(s) - \tilde{G}(s))M(s)} \end{aligned} \quad (6.59)$$

Die Kompensation ist damit unvollständig.

Solange $G(s)$ nur Pole und Nullstellen mit negativem Realteil aufweist, ist dies zwar ärgerlich, aber noch kein Problem. Schlimmstenfalls treten (abklingende) Schwingungen auf, die eigentlich hätten eliminiert werden sollen.

Sofern bei $G(s)$ aber Pole mit positivem Realteil auftreten, die durch $K(s)$ wegen eines fehlerhaften Streckenmodells nicht kompensiert werden, enthält die Kreisübertragungsfunktion $G(s)K(s)$ wider Erwarten Pole mit positivem Realteil. Das bedeutet wiederum eine dramatische Veränderung der Bedingungen beim Nyquist-Kriterium, die leicht zur Instabilität des Systems führen können.

Diese Probleme sind aber hier relativ einfach zu vermeiden. $M(s)$ muss lediglich so festgelegt werden, dass $1 - M(s)$ die einem instabilen Pol von $G(s)$ entsprechende Nullstelle enthält. Dadurch kürzen sich Pol und Nullstelle auf der rechten Seite von Gleichung (6.57), so dass der Regler $K(s)$ die Nullstelle nicht mehr enthält. Das Verfahren erwartet damit vom Regler gar nicht mehr die Kompensation der Polstelle der Strecke im geschlossenen Kreis, sondern löst das Problem anderweitig.

Dies soll anhand eines Beispiels erläutert werden. Gegeben sei eine instabile Strecke mit der Übertragungsfunktion

$$G(s) = \frac{1}{Ts - 1} \quad \text{mit} \quad T > 0 \quad (6.60)$$

Als Modellfunktion für das Übertragungsverhalten des geschlossenen Kreises wird ein Verzögerungsglied erster Ordnung gewählt, dessen Koeffizienten erst im Verlauf des Verfahrens festgelegt werden sollen:

$$M(s) = \frac{a}{bs + 1} \quad \text{mit} \quad b > 0 \quad (6.61)$$

Für $1 - M(s)$ erhält man

$$1 - M(s) = 1 - \frac{a}{bs + 1} = \frac{bs + 1 - a}{bs + 1} \quad (6.62)$$

Damit sich die Nullstelle dieser Funktion gegen die Polstelle von $G(s)$ kürzen lässt, muss $b = T$ und $a = 2$ gelten. Für den Regler ergibt sich daraus mit Gleichung (6.57)

$$K(s) = \frac{1}{G(s)} \frac{M(s)}{1 - M(s)} = \frac{Ts - 1}{1} \frac{2}{Ts + 1} \frac{Ts + 1}{Ts - 1} = 2 \quad (6.63)$$

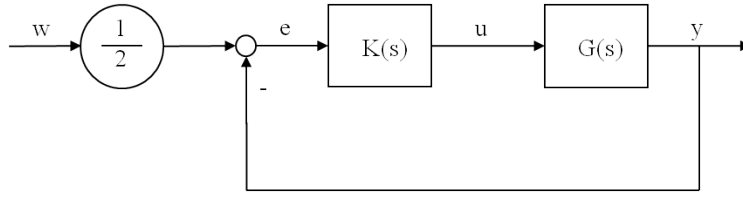
also ein reiner Proportionalregler. Mit diesem ergibt sich für den geschlossenen Kreis

$$T(s) = \frac{G(s)K(s)}{1 + G(s)K(s)} = \frac{\frac{1}{Ts-1} 2}{1 + \frac{1}{Ts-1} 2} = \frac{2}{Ts + 1} \quad (6.64)$$

Diese Berechnung zeigt, dass der geschlossene Kreis auf jeden Fall stabil ist, unabhängig davon, ob der genaue Wert von T bekannt ist oder nicht. Das Problem einer möglichen Nicht-Kompensation der instabilen Polstelle ist durch die geschickte Festlegung von $1 - M(s)$ einfach umgangen worden.

Zu beachten ist lediglich, dass der Verstärkungsfaktor von $T(s)$ in diesem Beispiel ungleich eins ist. Die Ausgangsgröße wäre nach Beendigung des Einschwingvorgangs immer doppelt so groß wie der Sollwert. Dieses Problem kann aber leicht dadurch umgangen werden, dass der Faktor 2 bei der Berechnung des Sollwertes gemäß Abbildung 6.23 berücksichtigt wird.

Falls $G(s)$ eine Nullstelle mit positivem Realteil enthält, tritt dasselbe Problem auf wie bei der instabilen Polstelle. Da der Term $\frac{1}{G(s)}$ im Regler

**Abb. 6.23.** Sollwertkorrektur

auftaucht, würde nun $K(s)$ einen instabilen Pol bekommen, und die Kreisübertragungsfunktion wäre bei fehlerhafter Kompensation wieder instabil. Die Lösung des Problems besteht nun darin, dass $M(s)$ die gleiche Nullstelle enthalten muss wie $G(s)$, denn dadurch kürzen sich beide Nullstellen in der Bestimmungsgleichung (6.57) des Reglers.

Abschließend sei angemerkt, dass die Ordnung des Reglers durch die Ordnung der Modellfunktion $M(s)$ bestimmt wird. Eine höhere Ordnung von $M(s)$ hat offenbar eine höhere Reglerordnung und dementsprechend auch mehr frei wählbare Parameter zur Folge, mit denen dann noch andere Forderungen an das Verhalten des Regelkreises erfüllt werden können.

Interessant ist ein Vergleich des hier vorgestellten Verfahrens mit der sogenannten *Youla-Parametrierung*. Diese geht auf Youla, Jabr und Bongiorno ([68] und [69]) zurück. Ihre Grundidee ist dem oben vorgestellten Ansatz sehr ähnlich, wie nachfolgend gezeigt wird.

Gegeben sei das Übertragungsverhalten eines geschlossenen Kreises:

$$T(s) = \frac{G(s)K(s)}{1 + G(s)K(s)} \quad (6.65)$$

Sucht man nun einen Regler, um ein bestimmtes Übertragungsverhalten erzielen, wird zunächst die Hilfsgröße

$$Q(s) = \frac{K(s)}{1 + G(s)K(s)} \quad (6.66)$$

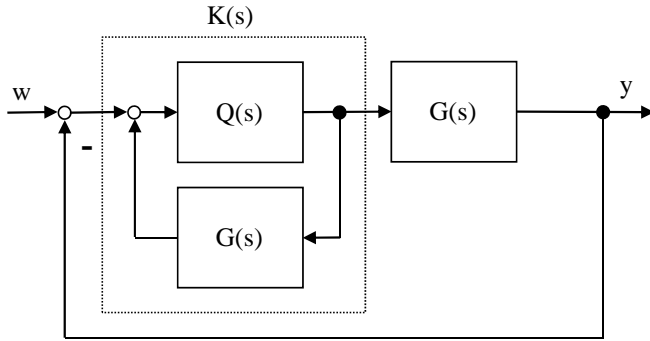
eingeführt. Mit dieser ergibt sich für den geschlossenen Kreis

$$T(s) = G(s)Q(s) \quad (6.67)$$

Für den Regler folgt durch Umstellen von (6.66)

$$K(s) = \frac{Q(s)}{1 - G(s)Q(s)} \quad (6.68)$$

Das entsprechende Blockschaltbild zeigt Abbildung 6.24. Der Regler enthält in einer Rückkopplung das Modell $G(s)$ der Strecke und entspricht damit der Übertragungsfunktion in Gleichung (6.68). Man erkennt sofort, dass die Wirkung der Strecke durch das im Regler rückgekoppelte Streckenmodell gerade kompensiert wird.

**Abb. 6.24.** Youla-Parametrierung

Wenn $G(s)$ stabil ist, dann kann man für $Q(s)$ wegen (6.67) jede beliebige stabile Übertragungsfunktion verwenden, und es wird sich immer ein stabiler geschlossener Kreis ergeben. Dies ist die Essenz der Youla-Parametrierung.

Für einen Vergleich der Youla-Parametrierung mit dem vorhergehenden Verfahren soll nun eine Modellfunktion $M(s)$ vorgegeben und nach der Youla-Parametrierung ein Regler entworfen werden.

Wegen $T(s) = M(s)$ gilt für $Q(s)$

$$Q(s) = \frac{M(s)}{G(s)} \quad (6.69)$$

Damit ergibt sich für den Regler gemäß (6.68)

$$K(s) = \frac{Q(s)}{1 - G(s)Q(s)} = \frac{\frac{M(s)}{G(s)}}{1 - G(s)\frac{M(s)}{G(s)}} = \frac{1}{G(s)} \frac{M(s)}{1 - M(s)} \quad (6.70)$$

Die Berechnungsvorschrift für den Regler ist exakt die gleiche wie in (6.57). Damit sind beide Auslegungsverfahren äquivalent.

6.3.4 Smith-Prädiktor

Abschließend soll der sogenannte *Smith-Prädiktor* vorgestellt werden, der auch durchaus noch mit zu den Kompensationsreglern gezählt werden kann. Dieses Verfahren ist besonders für Regelstrecken geeignet, deren dynamisches Verhalten sehr von den in ihnen enthaltenen Laufzeiten geprägt ist und bei denen diese Laufzeiten relativ genau bekannt sind. Typische Beispiele sind Heizungs- oder Kühlkreisläufe, Druckmaschinen und Walzwerke.

So wird beispielsweise in einem Walzwerk der Anpressdruck zweier Rollen geregelt, um eine bestimmte Dicke des gewalzten Stahls zu erzielen. Die Dicke des Stahls wird aber erst in einiger Entfernung der Rollen gemessen, so dass der Stahl von den Rollen bis zum Messpunkt eine bestimmte Transportzeit

benötigt. Diese Zeit darf bei der Einstellung des Reglers nicht vernachlässigt werden. Erleichternd für die Reglereinstellung ist aber, dass die Laufzeit anhand der Rotationsgeschwindigkeit der Rollen sehr genau berechnet werden kann.

Aufgrund der vorausgesetzten Linearität der gesamten Strecke lässt sich das Streckenmodell in einen rationalen Teil $G_R(s)$ ohne Laufzeiten und eine Laufzeit e^{-T_L} aufteilen, wobei die Laufzeit durchaus die Summe verschiedener Laufzeiten sein kann:

$$G(s) = G_R(s) e^{-T_L} \quad (6.71)$$

Wegen der Laufzeit kann die Wirkung eines Stelleingriffs zum Zeitpunkt t frühestens zum Zeitpunkt $t + T_L$ am Ausgang der Regelstrecke erkannt werden. Es kann daher bei falscher Einstellung des Reglers leicht passieren, dass der Regler beispielsweise bei einem positiven Regelfehler seine Stellgröße immer weiter aufintegriert, weil sich das Ausgangssignal der Strecke und damit der Regelfehler während der Laufzeit T_L nicht verändern. Wenn dann nach Ablauf der Laufzeit das Ausgangssignal endlich höhere Werte annimmt und der Regler seine Stellgröße wieder zurück nimmt, so erfolgt diese Rücknahme möglicherweise zu spät, und die Ausgangsgröße wird einen Überschwinger aufweisen. Im ungünstigsten Fall kann dieser Effekt sogar zu aufklingenden Schwingungen und damit zur Instabilität des Systems führen.

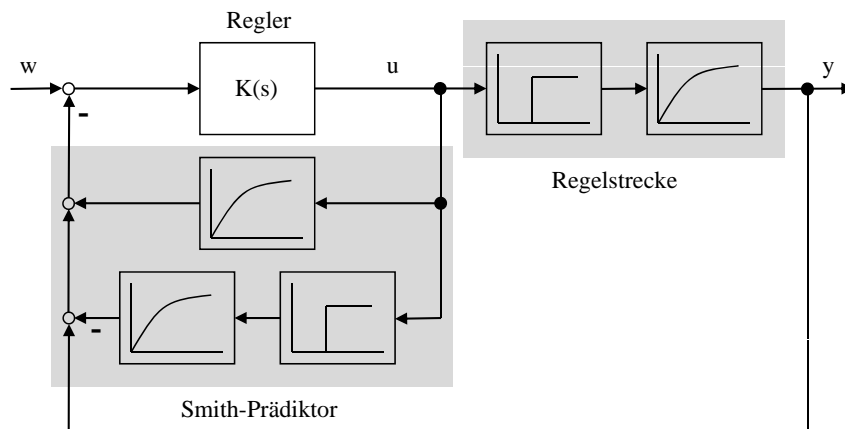


Abb. 6.25. Smith-Prädiktor für Strecken mit Laufzeit

Daher wird beim Smith-Prädiktor die Wirkung eines Stelleingriffs mit Hilfe eines im Regler enthaltenen Streckenmodells *ohne Laufzeit* vorhergesagt (oberer Teil des Smith-Prädiktors in Abb. 6.25). Die Ausgangsgröße dieses Modells wird als Pseudo-Messwert zum Regler zurückgeführt. Der Regler kann damit die Wirkung seiner Stellgröße auf den rationalen Streckenteil sofort erkennen und seine Stellgröße entsprechend zurücknehmen, sobald der Sollwert erreicht ist.

Der Regler schickt damit eine Stellgrößensequenz in die laufzeitbehaftete Strecke, die, bei ausreichender Genauigkeit des Modells, genau die richtigen Werte beinhaltet, um die Ausgangsgröße zum Sollwert zu führen. Im Prinzip ist damit die Kombination aus Reglerfunktion $K(s)$ und Modell des rationalen Systemteils eher eine Steuerung als eine Regelung, denn der reale Messwert wird dabei ja gar nicht verwendet.

In der Tat wird dieser durch den unteren Teil des Smith-Prädiktors durch das vollständige Streckenmodell *mit Laufzeit* kompensiert. Denn ansonsten würde der verspätet eintreffende, reale Messwert ja den Regler zu erneuter Stellgrößenaktivität veranlassen und damit die oben beschriebene Stellgrößensequenz wieder zunichte machen.

Es stellt sich die Frage, warum der reale Messwert überhaupt noch zurückgekoppelt wird. Dies ist aber unbedingt notwendig, denn ohne Rückkopplung eines realen Messwertes könnten Störungen nicht erkannt werden. Wenn eine Störung auftritt, so wird diese die reale Ausgangsgröße y verändern, nicht aber den Pseudo-Messwert des Smith-Prädiktors. Eine Kompensation erfolgt in diesem Falle nicht, aber das ist ja auch gewollt. Denn nun sieht der Regler eine Regelabweichung und kann, im Zusammenspiel mit dem Streckenmodell *ohne Laufzeit*, die für die Ausregelung der Störung optimale Stellgrößensequenz generieren.

6.4 Strukturserweiterung

6.4.1 Vorfilter

Noch weitergehende Möglichkeiten eröffnen sich, wenn man die in Abb. 6.1 gezeichnete Struktur hinter sich lässt und zusätzliche Elemente und Verbindungen in den Regelkreis einfügt. Eine sehr einfache Maßnahme ist die Verwendung eines *Vorfilters* (Abb. 6.26). Mit diesem ergibt sich für die Führungsübertragungsfunktion

$$T(s) = \frac{y(s)}{w(s)} = F(s) \frac{G(s)K(s)}{1 + G(s)K(s)} \quad (6.72)$$

während die Stör-Übertragungsfunktion $S(s)$ unverändert ist, da das Filter außerhalb des geschlossenen Kreises liegt. Aus dem Grund hat es natürlich auch keine Auswirkungen auf die Stabilität des Systems. Damit erhält man die Möglichkeit, Führungs- und Störübertragungsfunktion unabhängig voneinander zu gestalten. Zunächst wird der Regler $K(s)$ für ein optimales Störverhalten dimensioniert und anschließend das Vorfilter $F(s)$ für das Führungsverhalten.

Optimales Störverhalten bedeutet meist eine schnelle Ausregelung von Störungen. Dabei muss zwar die Stabilität berücksichtigt werden, eine hohe Dämpfung ist aber nicht unbedingt erforderlich. Bei der Auslegung ergibt

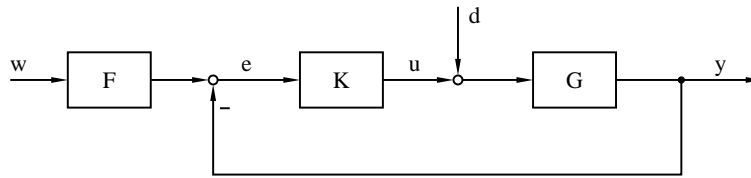


Abb. 6.26. Vorfiler

sich daher ein Regler mit großer Verstärkung und als Folge davon ein schlecht gedämpftes System. Dagegen ist beim Führungsverhalten eine ausreichende Mindestdämpfung von großem Interesse. Ein gut gedämpftes Führungsverhalten lässt sich aber nun durch das Vorfiler erzielen. Verwendet man hier beispielsweise einen Tiefpass, so gelangt eine Sollwertveränderung nur noch verzögert, d.h. mit stetigem Verlauf auf den geschlossenen Kreis und kann diesen trotz seiner geringen Dämpfung nicht mehr zu Schwingungen anregen. Aus Sicht der Eingangsgröße stellt sich das System damit als gut gedämpft dar. Durch Einbau eines Vorfilters wird es also möglich, die gegensätzlichen Forderungen nach schneller Ausregelung von Störungen und einem gut gedämpften Führungsverhalten besser in Einklang zu bringen.

Als Beispiel in der Praxis sei eine Lageregelung genannt. Der Regler wird für eine Dämpfung von $D = \frac{1}{\sqrt{2}}$ ausgelegt. Damit würde sich bei Eingabe eines neuen Lagesollwertes gemäß Abbildung 2.13 ein leichtes Überschwingen ergeben. Wählt man nun aber als Vorfiler ein Verzögerungsglied erster Ordnung $\frac{1}{Ts+1}$, dann ist durch dieses Vorfiler sichergestellt, dass keine sprunghafte Eingangsgröße mehr das System erreicht, und dementsprechend wird es bei einer Sollwertveränderung auch kein Überschwingen der Regelgröße mehr geben. Eventuelle Störungen werden durch den Regler aber schneller ausgegelt, als dies bei einer Dämpfung $D \geq 1$ der Fall wäre.

In der Praxis wird statt eines Verzögerungsgliedes oft auch die in Abbildung 6.27 gezeigte Struktur gewählt, die sicherstellt, dass ein neuer Sollwert nicht sprung- sondern rampenförmig erreicht wird.

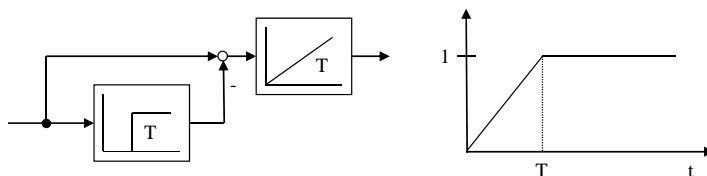


Abb. 6.27. Vorfiler für rampenförmigen Sollwertanstieg

6.4.2 Störgrößenaufschaltung

Eine andere einfache Strukturweiterung ist die *Störgrößenaufschaltung*, die ebenfalls nicht im geschlossenen Kreis erfolgt und somit keine Auswirkung auf

die Stabilität des Systems hat. In Abb. 6.28 ist die entsprechende Struktur gezeichnet.

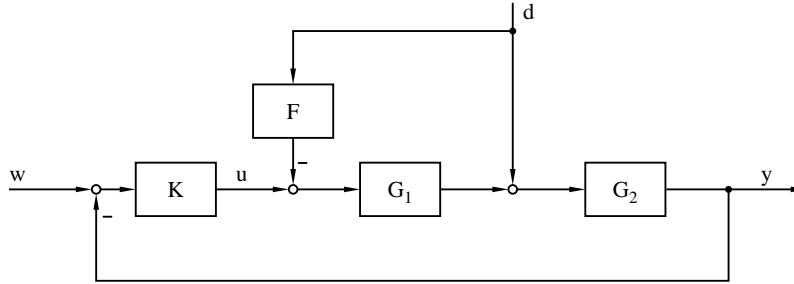


Abb. 6.28. Störgrößenaufschaltung

Dabei sei angenommen, dass die Störung zwischen zwei Streckenteilen $G_1(s)$ und $G_2(s)$ angreift. Ziel ist, die Störung zu kompensieren, bevor sie sich auf die Strecke auswirken kann. Voraussetzung ist natürlich, dass die Störgröße messbar ist. Die einfachste Idee ist, an der Angriffsstelle der Störung ein gleich großes Signal mit entgegengesetztem Vorzeichen aufzuschalten. Dies ist aber normalerweise nicht möglich, da die Störung oft an einer Stelle auf die Strecke einwirkt, an der der Mensch mit einem Stellglied überhaupt keinen Einfluss nehmen kann. So wird beispielsweise der Kurs eines Schiffes durch eine seitlich angreifende Strömung beeinflusst, ohne dass sich die dadurch auf das Schiff wirkende Kraft durch eine an derselben Stelle wirkende Gegenkraft kompensieren lässt. Abhilfe bietet hier nur ein rechtzeitiges Gegenauslenken des Ruders, also ein Eingriff mit der Stellgröße selbst. Entsprechend erfolgt eine Störgrößenaufschaltung normalerweise direkt nach dem Regler, so dass die zusätzliche Information in die Stellgröße mit eingehen kann.

Für eine vollständige Kompensation an der Angriffsstelle der Störung muss gelten:

$$F(s) = \frac{1}{G_1(s)} \quad (6.73)$$

Weil diese Funktion aber im allgemeinen nicht exakt zu realisieren ist, hat man sich meist mit einer unvollständigen Kompensation zu begnügen. Im Hinblick auf die Stabilität oder Genauigkeit der Regelung spielt dies jedoch keine Rolle. Denn aus Sicht des Reglers stellt die Aufschaltung nur eine weitere Störung dar, die mit geregelt wird. Dass diese zweite Störung der ersten entgegengerichtet ist, macht die Aufgabe für den Regler kleiner, und das ist ja auch der Sinn einer Störgrößenaufschaltung.

Eine Störgrößenaufschaltung ist sinnvoll, wenn der Regler aus Stabilitätsgründen, also um das Nyquist-Kriterium nicht zu verletzen, nur eine kleine Verstärkung (bzw. eine große Integratorzeitkonstante) haben darf und daher sehr langsam ist. Wenn bei einer solchen Regelung eine Störung auftritt, so

wird es entsprechend lange dauern, bis diese wieder ausgeregelt ist. In solchen Fällen dient die Störgrößenaufschaltung dazu, die Störung noch außerhalb des geschlossenen Kreises schon bei ihrem Entstehen zumindest teilweise zu kompensieren.

Ein schönes Beispiel für eine Störgrößenkompensation ist die Regelung des Energieversorgungsnetzes bei einem Fußballländerspiel. Es ist bekannt, dass in der Pause eines Länderspiels Millionen von Kühlschränken geöffnet werden und der Stromverbrauch entsprechend ansteigt. Aus dem Grund wird die Kraftwerksleistung kurz vorher erhöht, um ein größeres Leistungsdefizit im Netz zu vermeiden.

6.4.3 Ergänzende Rückführung

Eine weitere Möglichkeit zur Verbesserung der Regelung besteht in einer *ergänzenden Rückführung*. Diese wird im Gegensatz zu den beiden bisher vorgestellten strukturerweiternden Maßnahmen in den geschlossenen Kreis eingefügt und hat damit Auswirkungen auf die Stabilität des Systems (Abb. 6.29).

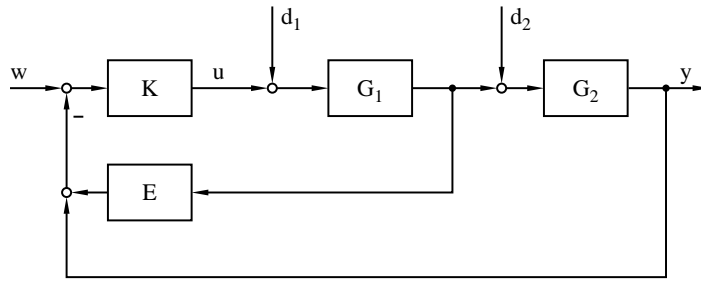


Abb. 6.29. Ergänzende Rückführung

Die Idee ist die folgende: Bei tiefpasshaltigen Strecken, wie sie meistens in der Praxis vorliegen, ist die Ausgangsgröße gegenüber internen Größen verzögert. Wenn es daher möglich ist, eine oder mehrere der internen Größen zu messen, kann dem Regler im Falle einer Störung Information über eine bevorstehende Änderung der Ausgangsgröße schon zugeführt werden, wenn sich die Ausgangsgröße selbst noch gar nicht verändert hat. Entsprechend früher kann der Regler auch Gegenmaßnahmen einleiten, was zu einem verbesserten Störverhalten führt. Dies ist bei der Störgröße d_1 der Fall.

Zu berücksichtigen ist natürlich, dass sich durch die Rückführung das Übertragungsverhalten des Systems ändert. Bei verschwindendem Sollwert w und verschwindenden Störgrößen d_1 und d_2 ergibt sich für die Eingangsgröße e des Reglers:

$$e(s) = -(E(s)G_1(s)u(s) + G_2(s)G_1(s)u(s)) = -(E(s) + G_2(s))G_1(s)u(s) \quad (6.74)$$

Ohne Rückführung gilt der Zusammenhang

$$e(s) = -G(s)u(s) \quad \text{mit} \quad G(s) = G_2(s)G_1(s) \quad (6.75)$$

Der Vergleich der beiden Gleichungen zeigt, dass die neue Übertragungsfunktion der Strecke aus Sicht des Reglers nun

$$G(s) = (E(s) + G_2(s))G_1(s) \quad (6.76)$$

lautet und bei der Auslegung des Reglers entsprechend zu berücksichtigen ist.

Der Vorteil eines verbesserten Störverhaltens besteht nicht, wenn die Störung hinter der Abgriffsstelle angreift, wie es für die eingezeichnete Störgröße d_2 der Fall ist. Hier tritt sogar ein Problem auf, das sich zeigt, wenn man die Störübertragungsfunktion von d_2 nach y berechnet.

Für diese Berechnung bietet es sich an, eine Hilfsgröße h am Ausgang von G_1 einzuführen. Es gilt für den linken Teil des Regelkreises

$$h(s) = -G_1(s)K(s)(E(s)h(s) + y(s)) \quad (6.77)$$

und für den rechten Teil

$$G_2(s)(h(s) + d_2(s)) = y(s) \quad \rightarrow \quad h(s) = \frac{1}{G_2(s)}y(s) - d_2(s) \quad (6.78)$$

Einsetzen von (6.78) in (6.77) liefert

$$\frac{1}{G_2(s)}y(s) - d_2(s) = -G_1(s)K(s)(E(s)(\frac{1}{G_2(s)}y(s) - d_2(s)) + y(s)) \quad (6.79)$$

Daraus folgt für die Störübertragungsfunktion von d_2 nach y

$$\frac{y(s)}{d_2(s)} = \frac{G_2(s)(1 + G_1(s)K(s)E(s))}{1 + G_1(s)K(s)(E(s) + G_2(s))} \quad (6.80)$$

Für stationäre Genauigkeit muss diese Funktion für $s \rightarrow 0$ gegen Null konvergieren (vgl. Gl. (6.3)). Bei einem integrierenden Regler gilt $K(s \rightarrow 0) = \infty$. Bei einer nicht differenzierenden Strecke konvergiert daher der Nenner gegen ∞ , ebenso aber auch der Term $G_1(s)K(s)E(s)$ im Zähler. Damit würde die Gesamtfunktion nicht mehr gegen Null konvergieren.

Um zu vermeiden, dass der Zähler gegen ∞ konvergiert, muss $E(s)$ die Polstelle des Reglers bei $s = 0$ kompensieren. $E(s)$ muss daher eine Nullstelle bei $s = 0$ enthalten, also beispielsweise einen verzögerten Differenzierer $\frac{s}{T_v s + 1}$. Nur dann ist die Störübertragungsfunktion im stationären Zustand gleich Null, und der Regler arbeitet stationär genau.

6.4.4 Kaskadenschaltung

Eine spezielle und in der Praxis sehr häufig eingesetzte Form der ergänzenden Rückführung ist die *Kaskadenschaltung*. In Abb. 6.30 ist ein Beispiel für eine zweischleifige Schaltung gezeigt. Dabei können im Prinzip beliebig viele Schleifen auftreten. Eine Kaskadenschaltung bietet sich an, wenn die Strecke als Reihenschaltung verschiedener Übertragungsglieder mit Tiefpasswirkung dargestellt werden kann.

Das System wird als eine Folge ineinander geschachtelter Regelkreise behandelt. Jeder Regler ist dabei für den in seiner Rückkopplungsschleife liegenden Streckenteil zuständig. In Abb. 6.30 regelt also Regler 2 die interne Größe y_2 und bekommt als Sollwert die Ausgangsgröße u_1 des Reglers 1. Der geschlossene, innere Kreis ist für den Regler 1 wiederum Teil der von ihm zu regelnden Strecke. Seine Regelgröße ist die Ausgangsgröße des Systems y_1 .

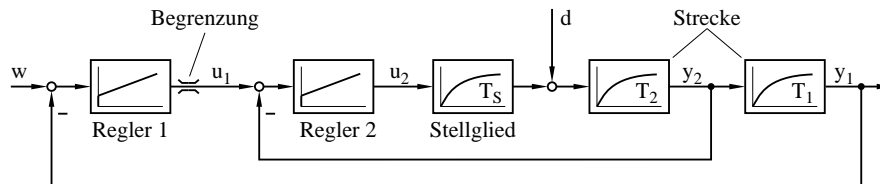


Abb. 6.30. Kaskadenschaltung

Ein Beispiel für eine Kaskadenregelung ist die Regelung einer elektrischen Maschine. Grundsätzlich besteht die Strecke zunächst aus einem Stromrichter. Aus dem Strom entsteht, je nach Typ der Maschine, ein Antriebsmoment. Die Division des Antriebsmomentes durch die Rotationsmasse liefert die Winkelbeschleunigung. Die Integration der Winkelbeschleunigung liefert die Winkelgeschwindigkeit, und die Integration der Winkelgeschwindigkeit liefert wiederum den Lagewinkel. Die Regelstrecke besteht damit aus einem in der Regel stark nichtlinearen Systemteil, der den Stromrichter beschreibt, einem weiteren, ebenfalls nichtlinearen Systemteil für die Entstehung des Drehmomentes aus dem Strom, und zwei Integratoren für den Zusammenhang zwischen Winkelbeschleunigung, Winkelgeschwindigkeit und Lagewinkel. Als Messgrößen stehen der Ausgangsstrom des Stromrichters, die Winkelgeschwindigkeit und der Lagewinkel zur Verfügung.

Diese Strecke wird mit einer Dreifach-Reglerkaskade geregelt. Der innerste Regelkreis ist der Stromregelkreis, der aus einem Stromregler und dem Umrichter besteht. Durch diesen Regelkreis wird gewährleistet, dass der tatsächliche Strom dem erwarteten Strom entspricht. Der nächstäußere Regelkreis ist der Drehzahlregelkreis. Dieser besteht aus einem Drehzahlregler, während seine Strecke aus dem (geschlossenen) Stromregelkreis, der Entstehung des Drehmomentes und dem Integrator von der Beschleunigung zur Drehzahl entspricht. Der äußerste Regelkreis schließlich ist der Lageregelkreis. Dieser besteht aus dem Lageregler mit einer Strecke, die aus dem

Drehzahlregelkreis und einem Integrator besteht, der den Zusammenhang zwischen Drehzahl und Lage repräsentiert: $\alpha(s) = \frac{1}{s}\omega(s)$. Der Lageregler gibt dem Drehzahlregler einen Drehzahl-Sollwert vor, und dieser wiederum gibt dem Stromregler einen Strom-Sollwert vor.

Ein Vorteil der Kaskadenregelung ist zunächst wie bei der ergänzenden Rückführung die schnellere Ausregelung von Störungen. Entsteht beispielsweise an der in Abb. 6.30 eingezeichneten Stelle eine Störung d , so kann diese vom Regler 2 bereits bekämpft werden, sobald sich y_2 ändert. Die Auslenkung der eigentlichen Regelgröße y_1 wird dann natürlich weniger stark ausfallen.

Ein weiterer Vorteil ist die Möglichkeit, interne Größen zu begrenzen. In Abb. 6.30 ist am Ausgang des Reglers 1 eine solche Begrenzung eingezeichnet, die in ihrer Funktionsweise dem nichtlinearen Übertragungsglied in Abb. 6.4 entspricht. Diese Begrenzung wirkt zwar auf die Größe u_1 ein, soll aber eigentlich eine Begrenzung für die interne Größe y_2 darstellen. Wenn nämlich Regler 2 schnell und genau genug arbeitet, kann man davon ausgehen, dass y_2 in etwa dem durch u_1 vorgegebenen Verlauf entspricht und damit auch innerhalb der gegebenen Grenzen bleibt.

Schließlich wird die Auswirkung von nichtlinearen Gliedern auf den Regelkreis beschränkt, in dem sie enthalten sind. Würde der innere Regelkreis beispielsweise ein nichtlineares Übertragungsglied enthalten, so wäre von dieser Nichtlinearität bei hinreichend schnellem und genauem Regler 2 im äußeren Regelkreis kaum etwas wahrzunehmen. Denn durch die Regelung ist gewährleistet, dass y_2 in etwa dem Signal u_1 folgt, was einem verzögerten, proportionalen und damit linearen Übertragungsverhalten entspricht.

Der für den Praktiker interessanteste Vorteil ist aber die leichte Inbetriebnahme einer Kaskadenschaltung. Zunächst wird Regler 2 für den inneren Kreis dimensioniert. Der geschlossene innere Kreis kann dann, bei hinreichend schneller Regelung, durch ein PT_1 -Glied angenähert werden. Mit dieser Vereinfachung ist es anschließend auch möglich, Regler 1 für den äußeren Kreis zu berechnen. Grundsätzlich dimensioniert man bei einer Kaskadenschaltung die Regler sukzessive von innen nach außen, wobei jeweils der innere Kreis durch ein einfaches Übertragungsglied angenähert wird.

Diese Vorgehensweise soll anhand des Beispiels in Abb. 6.30 kurz erläutert werden. Gegeben sei eine aus drei hintereinandergeschalteten PT_1 -Gliedern bestehende Strecke, wobei das erste PT_1 -Glied als Näherung für das dynamische Verhalten des Stellgliedes zu verstehen ist. Alle drei PT_1 -Glieder haben die Verstärkung Eins, was die Rechnung erleichtert.

Der von Regler 2 zu regelnde Streckenteil besteht aus zwei PT_1 -Gliedern. Dieser Fall ist aber bereits behandelt worden (vgl. Gleichung (6.15)). Als Regler ist ein PI-Regler zu wählen, dessen Zeitkonstante T_{pi} man der größeren der beiden Streckenzeitkonstanten gleichzusetzen hat, um die entsprechende Polstelle zu kompensieren. Hier ist aber eine der beiden Zeitkonstanten, nämlich T_S , sowieso nur eine Ersatzzeitkonstante, d.h. es existiert in der realen Strecke kein entsprechender Pol, der durch den Regler kompensiert werden könnte.

Von daher erübrigt sich die Frage nach der größeren Zeitkonstanten, und man setzt $T_{pi} = T_2$. Für die Reglerverstärkung ergibt sich nach Gleichung (6.19)

$$V_R = \frac{T_2}{4T_S D^2} \quad (6.81)$$

Nun ist noch eine geeignete Dämpfung D zu wählen. Der innere Kreis wird nach außen als PT_2 -Glieder mit gerade dieser Dämpfung erscheinen. Wählt man $D < 1$, so ist dieses PT_2 -Glieder schwingungsfähig, was wiederum die Regelung des äußeren Regelkreises erschwert. Aus dem Grund kommt nur ein Wert $D \geq 1$ in Frage, wobei für alle diese Werte der innere Kreis ein aperiodisches Einschwingverhalten aufweist, das mit größer werdendem D immer langsamer wird. Für eine optimale Regelgeschwindigkeit bei aperiodischem Einschwingverhalten ist deshalb $D = 1$ die richtige Wahl für den Dämpfungsfaktor. Für das Übertragungsverhalten des inneren Kreises folgt daraus

$$\frac{y_2(s)}{u_1(s)} = \frac{1}{4T_S^2 s^2 + 4T_S s + 1} \quad (6.82)$$

Die Annäherung dieser Übertragungsfunktion nach Gleichung (3.42) durch ein PT_1 -Glieder ergibt

$$\frac{y_2(s)}{u_1(s)} \approx \frac{1}{4T_S s + 1} \quad (6.83)$$

Damit besteht dann aber auch die vom Regler 1 zu regelnde Strecke aus zwei PT_1 -Gliedern, und bei Vernachlässigung der Begrenzung ergeben sich hier nach völlig analogen Überlegungen die Reglerparameter zu $T_{pi} = T_1$ und

$$V_R = \frac{T_1}{16T_S D^2} \quad (6.84)$$

6.4.5 Statische Vorsteuerung

Neben den genannten Vorteilen hat die Kaskadenschaltung allerdings den Nachteil eines schlechten Führungsverhaltens. Der Grund ist leicht einzusehen. Wird am Eingang der Schaltung für den äußersten Kreis ein neuer Sollwert vorgegeben, so muss diese Anregung erst die gesamte Reglerkaskade durchlaufen, ehe die Strecke selbst angeregt wird und sich die Ausgangsgröße verändert. Abhilfe bietet hier eine *Vorsteuerung*. Sie verbessert das Führungsverhalten und wird daher oft in Verbindung mit einer Kaskadenregelung verwendet. Eine sogenannte statische Vorsteuerung ist in Abb. 6.31 zu sehen.

Die Führungs-Übertragungsfunktion des Systems lautet:

$$T(s) = \frac{y(s)}{w(s)} = \frac{G(s)(K(s) + V)}{G(s)K(s) + 1} \quad (6.85)$$

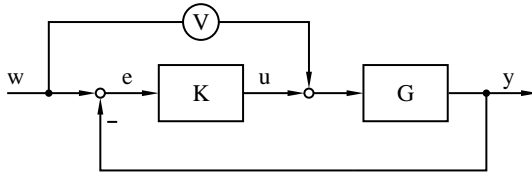


Abb. 6.31. Statische Vorsteuerung

Da die Vorsteuerung außerhalb des geschlossenen Kreises liegt, kann der Faktor V ohne Rücksicht auf die Stabilität festgelegt werden, was man auch daran erkennt, dass V nicht im Nenner der Übertragungsfunktion auftaucht. Die stationäre Genauigkeit der Regelung ist ebenfalls nicht gefährdet, denn falls $K(s)$ einen Integralanteil enthält, d.h. $\lim_{s \rightarrow 0} K(s) = \infty$, gilt weiterhin $\lim_{s \rightarrow 0} T(s) = 1$. Die Wirkung einer Vorsteuerung lässt sich folgendermaßen erklären: Falls der Regler beispielsweise ein PI-Regler ist, so wird bei einer Sollgrößenänderung Δw dieser Sprung, multipliziert mit dem Proportionalanteil P des Reglers, an die Strecke weitergegeben. Gleichzeitig gelangt aber auch ein Sprung $V\Delta w$ über den Vorsteuerkanal auf die Strecke. Bei einer Veränderung der Sollgröße vergrößert also gewissermaßen der Vorsteuerkanal den Proportionalanteil des Reglers. Das rückgekoppelte Signal y läuft dagegen nicht über den Vorsteuerkanal, d.h. im weiteren Verlauf ist der Regler auf sich allein gestellt. Der Vorsteuerkanal bleibt konstant auf dem anfangs erreichten Wert, und der Regler wird, falls er einen Integrator enthält, seine Stellgröße so lange weiter verändern, bis die Regelabweichung e verschwunden ist. In der Anfangsphase wird durch die Vorsteuerung also die Anregung auf die Strecke vergrößert, was zu einer erhöhten Regelgeschwindigkeit führt, während im weiteren Verlauf der Vorsteuerkanal die Ausregelung der Regelabweichung nicht mehr beeinflusst.

Bei einer Kaskadenschaltung lässt sich im Prinzip für jeden Regelkreis eine statische Vorsteuerung gemäß dem oberen Blockschaltbild in Abbildung 6.32 einfügen.

Bei dieser Struktur wird jeder einzelne Regelkreis durch die Vorsteuerung beschleunigt, aber der prinzipielle Nachteil der Kaskadenregelung, dass nämlich ein neuer Sollwert erst die Reglerkaskade durchlaufen muss, ehe er auf die Strecke wirken kann, wird dadurch nicht behoben. Dazu kommt, dass die Vorsteuerung V_2 des inneren Kreises aus Sicht des äußeren Reglers K_1 mit zur Strecke gerechnet werden muss und sich damit auf das Stabilitätsverhalten des Gesamtsystems auswirkt.

Sinnvoller erscheint unter diesen Aspekten die Struktur, die im unteren Blockschaltbild von Abbildung 6.32 dargestellt ist. Ein neuer Sollwert wird über verschiedene Vorsteuerfaktoren direkt auf die einzelnen Regelkreise gegeben, so dass keine Verzögerung mehr durch die Reglerkaskade auftreten kann. Außerdem arbeitet die Vorsteuerung in dieser Struktur komplett au-

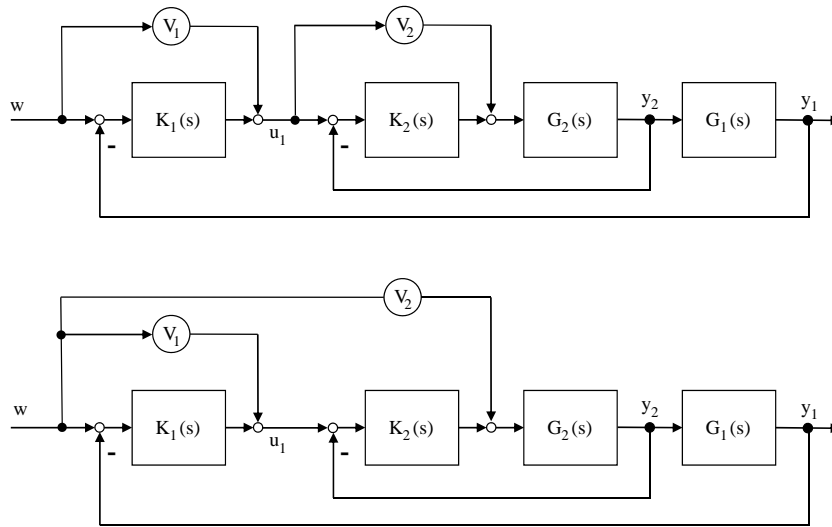


Abb. 6.32. Statische Vorsteuerung bei einer Kaskadenregelung

ßerhalb des geschlossenen Kreises, so dass hier auch keine Stabilitätsprobleme auftreten können.

Dafür entsteht jetzt aber ein anderer Nachteil: Die Vorsteuerung greift an verschiedenen Stellen in das System ein. Dadurch kann es zu der Situation kommen, dass dem Regler K_2 in der Anfangsphase des Sollwertsprungs über V_1 ein Sollwert vorgegeben wird, der nicht zu dem Wert passt, den V_2 über den Streckenteil G_2 bei der internen Ausgangsgröße y_2 hervorruft. Der Regler 2 wird mit Hilfe seiner Stellgröße versuchen, y_2 an den ihm vorgegebenen Sollwert anzupassen, und arbeitet damit im Prinzip gegen den über V_2 eingespeisten Vorsteuerwert. Der gleiche Effekt kann auch für den Regler 1 auftreten, wenn der über V_2 eingespeiste Vorsteuerwert schließlich auch bei y_1 eine Veränderung hervorruft, die nicht zum Sollwert des Reglers 1 passt. Es wird zu Einschwingvorgängen kommen, die nicht durch schlecht gedämpfte Polstellen im geschlossenen Kreis, sondern durch schlecht eingestellte Vorsteuerwerte hervorgerufen werden.

6.4.6 Dynamische Vorsteuerung

Gelöst wird dieses Problem mit Hilfe einer dynamischen Vorsteuerung, die auch als *Führungsgrößengenerator* (FGG) bezeichnet wird. Das Prinzip der Vorsteuerung, nämlich durch eine zusätzliche Aufschaltung das Führungsverhalten des Regelkreises zu verbessern, bleibt erhalten. Durch die Struktur des FGG wird jedoch gewährleistet, dass die einzelnen aufgeschalteten Größen zueinander passen und die Regler der Kaskadenschaltung nicht gegeneinander arbeiten.

Führungsgrößengeneratoren werden heutzutage üblicherweise digital in Form einer Software-Routine auf einem Signalprozessor realisiert. In Abhängigkeit vom zu erreichenden Endwert der Ausgangsgröße kann der optimale dynamische Verlauf der Sollwerte für die einzelnen Regler der Kaskadenregelung mit wenigen Zeilen Programmcode berechnet werden. Welche Zusammenhänge dabei zu beachten sind, kann man jedoch am besten anhand der klassischen, analogen Realisierung erläutern (Abb. 6.33).

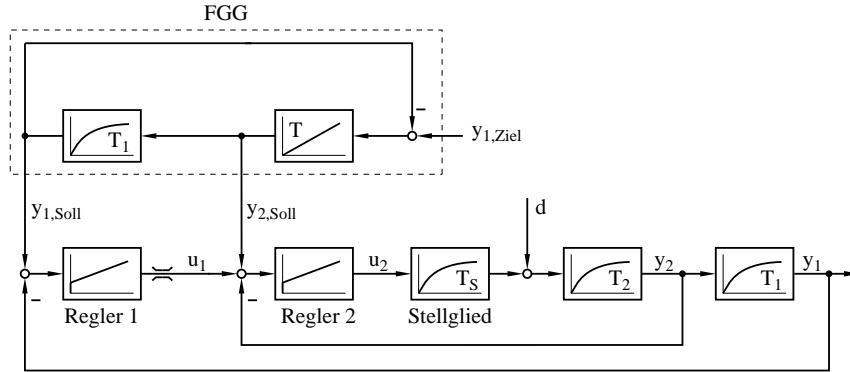


Abb. 6.33. Dynamische Vorsteuerung

Der FGG erzeugt passend zum Sollwert $y_{1,Ziel}$ für die Hauptregelgröße y_1 den Sollwert $y_{1,Soll}$ und für die innere Regelgröße y_2 den Sollwert $y_{2,Soll}$. Der innere Regler 2 regelt dann seine Regelgröße y_2 auf den geforderten Wert, und die Hauptregelgröße y_1 stellt sich in der Folge im Idealfall ohne Mitwirkung von Regler 1 auf den richtigen Wert ein.

Die Berechnung des Sollwertes $y_{2,Soll}$ erfolgt im FGG mittels eines Streckenmodells. Zwischen $y_{2,Soll}$ und $y_{1,Soll}$ muss der gleiche dynamische Zusammenhang wie zwischen y_2 und y_1 bestehen, damit die einzelnen Regler nicht gegeneinander arbeiten. Im vorliegenden Fall ist dieser Zusammenhang ein PT_1 -Glieder mit der Verzögerungszeit T_1 . Dieses PT_1 -Glieder ist im FGG nachgebildet und stellt sicher, dass $y_{1,Soll}$ und $y_{2,Soll}$ zueinander passen.

Wenn nun ein neuer Sollwert $y_{1,Ziel}$ für die Hauptregelgröße vorgegeben wird, so wird sich der Ausgang des Integrators im FGG so lange verändern, bis sein Eingang gleich Null ist, also bis $y_{1,Soll} = y_{1,Ziel}$ gilt. In dem Fall hat aber wegen des oben erwähnten Zusammenhanges auch $y_{2,Soll}$ gerade den zu $y_{1,Ziel}$ passenden Wert erreicht.

Durch den Integrator im FGG wird darüber hinaus auch gewährleistet, dass $y_{2,Soll}$ einen stetigen Verlauf aufweist. Denn da sich y_2 als Ausgang eines Verzögerungsgliedes nur stetig verändern kann, könnte der Regler 2 mit seiner Regelgröße y_2 einem unstetigen Sollwertverlauf gar nicht folgen.

Die einzelnen Teile des FGG sind damit erklärt: Der Integrator wird benötigt, damit sich die internen Größen des FGG so lange verändern, bis $y_{1,Soll} = y_{1,Ziel}$ gilt. Er gewährleistet damit die stationäre Genauigkeit des

FGG. Und das PT_1 -Glied gewährleistet den korrekten dynamischen Zusammenhang zwischen $y_{1,Soll}$ und $y_{2,Soll}$.

Zu bedenken ist allerdings, dass der FGG für sich genommen einen geschlossenen Kreis darstellt, dessen Stabilität sicherzustellen ist. Für das Übertragungsverhalten von $y_{1,Ziel}$ nach $y_{1,Soll}$ ergibt sich

$$\frac{y_{1,Soll}(s)}{y_{1,Ziel}(s)} = \frac{1}{\frac{T_1 T}{V_1} s^2 + \frac{T_1 T}{V_1} s + 1} \quad (6.86)$$

Dies entspricht einem PT_2 -Glied. Ein Koeffizientenvergleich mit der Normalform (3.9) liefert den Zusammenhang zwischen der Zeitkonstante T des Integrators und der Dämpfung D :

$$T = 4D^2 \frac{V_1}{T_1} \quad (6.87)$$

Da der FGG die Sollwerte für das eigentliche Regelsystem vorgeben soll, ist ein aperiodisches Einschwingen unbedingt erforderlich, damit sich die Sollwerte monoton verändern. Somit ist $D = 1$ zu setzen. Man erhält

$$T = 4 \frac{V_1}{T_1} \quad (6.88)$$

Wenn das Übertragungsverhalten von y_2 nach y_1 nicht durch ein Verzögerungsglied, sondern durch einen Integrator mit der Integrierzeit T_1 charakterisiert wäre, müsste das PT_1 -Glied im FGG, das ja genau diesen Zusammenhang widerspiegeln soll, durch einen Integrator mit der Integrierzeit T_1 ersetzt werden. Für das Übertragungsverhalten von $y_{1,Ziel}$ nach $y_{1,Soll}$ ergäbe sich damit

$$\frac{y_{1,Soll}(s)}{y_{1,Ziel}(s)} = \frac{1}{T_1 T s^2 + 1} \quad (6.89)$$

also ein Übertragungsverhalten, das unabhängig von der Wahl von T immer instabil wäre, so dass das Schema eines FGG in der hier behandelten, analogen Form nicht mehr anwendbar wäre.

Für eine Software-Realisierung besteht aber kein Problem. Die Rückkopplung von $y_{1,Soll}$ nach $y_{1,Ziel}$ ist hier überhaupt nicht erforderlich. Diese Rückkopplung und auch der Integrator mit der Integrierzeit T sollen ja nur sicherstellen, dass $y_{1,Soll}$ nach Abklingen der Einschwingvorgänge im FGG $y_{1,Ziel}$ entspricht. Dies lässt sich aber in einem Programm auch durch eine simple IF-THEN-Abfrage erledigen. Bei einer Veränderung des Sollwertes $y_{1,Ziel}$ lässt man hier also $y_{2,Soll}$ mit jedem Abtastschritt ansteigen. Aus $y_{2,Soll}$ wird gemäß dem Streckenmodell $y_{1,Soll}$ berechnet. Und sobald $y_{1,Soll} = y_{1,Ziel}$ ist, bleibt der FGG auf den Werten stehen, die er zu dem Zeitpunkt gerade erreicht hat.

Abschließend bleibt die Frage, warum der äußere Regelkreis überhaupt noch benötigt wird, wenn der FGG $y_{2,Soll}$ schon passend zu $y_{1,Soll}$ bzw.

$y_{1,Ziel}$ berechnet und sich damit aus $y_2 = y_{2,Soll}$ auch $y_1 = y_{1,Ziel}$ automatisch ergeben wird. Man muss aber in der Praxis immer damit rechnen, dass das Streckenmodell im FGG nicht exakt ist oder Störungen in der Strecke zwischen y_2 und y_1 angreifen, die vom Regler 2 natürlich nicht erkannt werden können. Daher ist ein äußerer Regelkreis für die Hauptregelgröße unerlässlich. Bei relativ exaktem Streckenmodell im FGG und geringen Störungen im Streckenteil zwischen y_2 und y_1 bleibt für den Regler 1 aber in der Tat wenig zu tun.

Wie bei der statischen Vorsteuerung verursacht auch der Führungsgrößen-generator im Regelkreis keine Stabilitätsprobleme, sofern seine interne Übertragungsfunktion nicht instabil ist, da auch er außerhalb des geschlossenen Kreises arbeitet. Damit ist er natürlich auch nur bei einer Änderung der Führungsgröße wirksam. Die schnelle Ausregelung von Störungen ist aber durch die Kaskadenschaltung ohnehin gewährleistet.

6.4.7 Entkopplung

Bisher nicht behandelt wurden bislang die Mehrgrößensysteme. Dabei ist es in der Realität der Normalfall, dass auf ein System mehrere Größen einwirken und andererseits auch mehrere Ausgangsgrößen von Interesse sind, wobei jede Ausgangsgröße von mehreren Eingangsgrößen abhängen kann. Man spricht hier von einer *Verkopplung* der einzelnen Größen.

In manchen Fällen lässt sich zu jeder Ausgangsgröße genau eine Eingangsgröße angeben, die einen wesentlichen Einfluss auf die Ausgangsgröße hat, während der Einfluss aller anderen Eingangsgrößen eher gering ist. Dann kann man versuchen, die Übertragungsfunktion von jeder Eingangs- zur zugehörigen Ausgangsgröße zu bestimmen und für dieses Teilsystem eine Regelung wie für ein gewöhnliches Eingrößensystem auszulegen. Der Einfluss der anderen Teilsysteme wird als Störgröße behandelt. Das Mehrgrößensystem wird also in mehrere Teilsysteme zerlegt, die als voneinander unabhängige Eingrößensysteme behandelt werden. Voraussetzung für den Erfolg dieser Vorgehensweise ist offensichtlich, dass die Kopplungen zwischen den Teilsystemen ausreichend schwach sind.

Bei zu starker Verkopplung der Teilsysteme kann diese Methode aber fatale Folgen haben. Wird beispielsweise durch einen Regler eine Eingangsgröße der Mehrgrößenstrecke verändert, so verändern sich mehrere Ausgangsgrößen, was Reaktionen anderer Regler hervorruft, die sich wiederum auf die Regelgröße des ersten Reglers auswirken können. Dieser leitet daraufhin entsprechende Gegenmaßnahmen ein, was wiederum Gegenmaßnahmen anderer Regler hervorruft. Es ist unwahrscheinlich, dass das System jemals in einen stationären Endzustand kommt. Schlimmstenfalls können die Schwingungen sogar aufklingen.

Man kann aber weiterhin für jede Größe einen eigenen Regelkreis auslegen, falls sich der Einfluss der anderen Größen auf diesen Kreis kompensieren lässt. Abb. 6.34 zeigt als Beispiel eine Zweigrößenstrecke, bei der sich die

Stellgrößen u_1 und u_2 über die Übertragungsglieder G_{ij} auf beide Ausgangsgrößen auswirken. Damit jeder der beiden Regelkreise für sich dimensioniert werden kann, sind die Größen d_1 und d_2 zu kompensieren. Dies geschieht mit den Signalen a_1 und a_2 , die über die *Entkopplungsglieder* E_{ij} aus den Stellgrößen hervorgehen. Beispielsweise muss für eine Kompensation von d_2 durch a_2 gelten:

$$G_{11}(s)a_2(s) \stackrel{!}{=} -d_2(s) \quad (6.90)$$

Die Entkopplung entspricht damit im Prinzip einer Störgrößenaufschaltung. Aus (6.90) folgt mit $d_2(s) = G_{12}(s)u_2(s)$ und $a_2(s) = E_{12}(s)u_2(s)$ für die Berechnung des Entkopplungsgliedes:

$$\begin{aligned} G_{11}(s)E_{12}(s)u_2(s) &\stackrel{!}{=} -G_{12}(s)u_2(s) \\ \Rightarrow E_{12}(s) &= -\frac{G_{12}(s)}{G_{11}(s)} \end{aligned} \quad (6.91)$$

Analog dazu gilt für das zweite Entkopplungsglied:

$$E_{21}(s) = -\frac{G_{21}(s)}{G_{22}(s)} \quad (6.92)$$

Da bei diesen Formeln Übertragungsfunktionen im Nenner auftauchen, gibt es natürlich Beschränkungen in Bezug auf Nullstellen mit positivem Realteil und auch in Bezug auf die Ordnung von Zähler- und Nennerpolynom, so dass eine solche Entkopplung in vielen Fällen nur eingeschränkt oder gar nicht möglich ist.

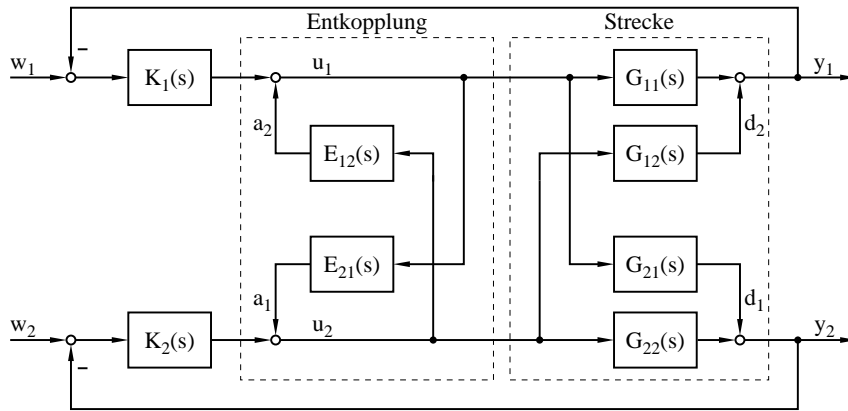


Abb. 6.34. Entkopplung bei einem Zweigrößensystem

Die bisher vorgestellten Methoden eignen sich hervorragend zur Regelung linearer Eingrößensysteme, aber auch nichtlineare Strecken (nach einer Linearisierung) und sogar Mehrgrößensysteme lassen sich teilweise noch mit diesen

einfachen Mitteln regeln. Daher besteht der überwältigende Anteil aller industriell eingesetzten Regler aus Reglern vom PID-Typ mit den verschiedensten Strukturweiterungen. In der Forschung wird dagegen seit den sechziger Jahren ein anderer Ansatz verfolgt, der eine wesentlich tiefere Einsicht in das Systemverhalten mit sich bringt und dadurch eine Mehrgrößenregelung von Strecken beliebiger Ordnung in einem geschlossenen Ansatz ermöglicht. Es ist das im nächsten Kapitel vorgestellte Prinzip der Zustandsregelung.

7. Grundlagen der Zustandsdarstellung

7.1 Definition von Zustandsgrößen

Die Methodik, die den klassischen regelungstechnischen Verfahren zu Grunde liegt, lässt sich folgendermaßen skizzieren: Die Differentialgleichungen der Strecke werden in den Frequenzbereich transformiert und zu einer Übertragungsfunktion zusammengefasst, für die dann - ebenfalls im Frequenzbereich - ein geeigneter Regler gesucht wird. Diese Vorgehensweise ist in der Mitte des letzten Jahrhunderts entwickelt worden, während sich die Regelungstechnik als eigenständige Wissenschaft etablierte, und hat sich über den langen Zeitraum bis heute in vielfältigen Anwendungen bewährt, so dass auch heute noch der überwiegende Teil aller industriellen Regelungen nach dieser Methode entworfen wird.

An die Grenze stößt man mit dieser Methodik jedoch bei Systemen von zunehmender Komplexität, insbesondere bei Systemen mit mehreren Ein- und Ausgangsgrößen, den sogenannten *MIMO-Systemen* (Multiple Input - Multiple Output). Bis zu einem gewissen Grad lässt sich die Aufgabenstellung mit Hilfe von Entkopplungsstrategien (vgl. Kap. 6.4.7) auf den SISO-Fall (Single Input - Single Output) zurückführen, doch sind die Möglichkeiten, mit dieser Methodik zu einer guten Regelung zu gelangen, aufgrund der sehr engen strukturellen Vorgaben bei der Entkopplung außerordentlich begrenzt.

Als Beispiel für ein MIMO-System möge ein Flugzeug dienen, wie es in Abbildung 7.1 gezeigt ist.

Wenn der Rollwinkel verändert wird, d.h. wenn das Flugzeug aus seiner waagerechten Flugposition herausgedreht wird, dann verringert sich gleichzeitig der Auftrieb an den Tragflächen und das Flugzeug geht in den Sink- bzw. Sturzflug über. Dieses Verhalten kann man gut in Filmaufnahmen aus dem 2. Weltkrieg beobachten. Wenn Sturzkampfbomber zum Angriff in den Sturzflug übergehen, kippen Sie immer über eine Tragfläche ab. Beeinflusst wird über den Rollwinkel also auch der Steigungswinkel, ohne dass am Höhenruder ein Stelleingriff erfolgt wäre.

Eine Veränderung des Rotationswinkels würde dazu führen, dass sich das Flugzeug quer zur Flugrichtung stellt. Die dann nach vorn zeigende Tragfläche würde durch den Luftwiderstand nach oben oder unten gedrückt, was offenbar den Rollwinkel beeinflusst und damit am Ende auch den Steigungswinkel.

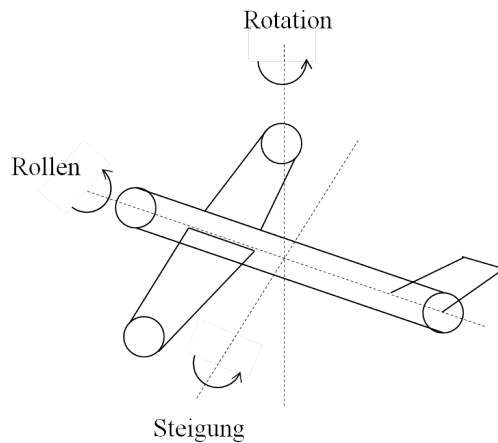


Abb. 7.1. Drehachsen beim Flugzeug

Offensichtlich beeinflussen sich die verschiedenen Größen gegenseitig. Ein Stelleingriff an einer einzigen Größe hat auch Auswirkungen auf die anderen Größen, so dass hier, wollte man einen klassischen Eingrößen-Reglerentwurf für jede Drehachse vornehmen, zunächst aufwendige Entkopplungsmaßnahmen durchzuführen wären.

Seit Beginn der sechziger Jahre ist an die Seite des klassischen Reglerentwurfs eine völlig andere Methodik getreten, die im Folgenden vorgestellt werden soll. Sie wird als Zustandsraummethodik bezeichnet und geht zu großen Teilen auf Rudolf Kalman zurück [24, 70].

Der entscheidende Unterschied zur bisherigen Vorgehensweise liegt in einer Betrachtung der internen Größen des Systems. Versuchte man früher, die internen Größen zu eliminieren und durch die Übertragungsfunktion einen direkten Zusammenhang zwischen Ein- und Ausgangsgrößen herzustellen, so wird bei der Zustandsraummethodik gerade das gegenteilige Ziel verfolgt. Hier wird auf das Verhalten der internen Systemgrößen besonderes Augenmerk gelegt, während die Ausgangsgrößen nur am Rande betrachtet werden. Es wird sich zeigen, dass die Betrachtung der internen Größen zu wesentlich besseren Einsichten in das Systemverhalten führt und den Entwurf eines Mehrgrößenreglers deutlich vereinfacht.

Das zu Grunde liegende Prinzip ist relativ einfach. Die Differentialgleichungen der Strecke werden nicht mehr in den Frequenzbereich transformiert und zu einer Übertragungsfunktion zusammengefasst, sondern im Zeitbereich so zerlegt, dass ein System aus Differentialgleichungen erster Ordnung entsteht. Dies ist immer möglich, denn jede Differentialgleichung k -ter Ordnung lässt sich in k Gleichungen erster Ordnung zerlegen, sofern man genügend interne Hilfsgrößen, die sogenannten *Zustandsgrößen* einführt. Diese Zustandsgrößen oder *Zustandsvariablen* können dabei durchaus realen physikalischen Größen entsprechen. Am Ende erhält man für eine Strecke n -ter Ordnung

ein System von n Differentialgleichungen erster Ordnung, wodurch die n Zustandsgrößen festgelegt sind. Die Ausgangsgrößen des Systems können dann durch gewöhnliche Funktionen der Eingangs- und Zustandsgrößen beschrieben werden. Es ergibt sich

$$\begin{aligned}\dot{x}_i &= f_i(x_1, \dots, x_n, u_1, \dots, u_p, t) & i \in \{1, \dots, n\} \\ y_j &= g_j(x_1, \dots, x_n, u_1, \dots, u_p, t) & j \in \{1, \dots, q\}\end{aligned}\quad (7.1)$$

bzw. in vektorieller Schreibweise

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}, \mathbf{u}, t)\end{aligned}\quad (7.2)$$

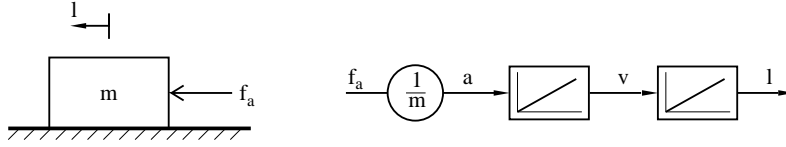
mit $\mathbf{x} = [x_1, \dots, x_n]^T$, $\mathbf{u} = [u_1, \dots, u_p]^T$, $\mathbf{y} = [y_1, \dots, y_q]^T$, $\mathbf{f} = [f_1, \dots, f_n]^T$ und $\mathbf{g} = [g_1, \dots, g_q]^T$.

Diese Gleichungen bezeichnet man als die *Zustandsdarstellung* des Systems. Dabei sind die u_i die Eingangsgrößen des Systems, die x_i die Zustandsgrößen, die y_i die Ausgangsgrößen und die f_i und g_i zunächst beliebige, skalare Funktionen der Zustands- und Eingangsgrößen sowie der Zeit t . Die Ausgangsgrößen müssen nicht unbedingt von den Zustandsgrößen verschieden sein. Bei Gleichheit einer Zustandsgröße und einer Ausgangsgröße wird die zugehörige Funktion g_j natürlich trivial: $y_j = g_j(\mathbf{x}, \mathbf{u}, t) = x_j$. Allgemein ist aber davon auszugehen, dass die Anzahl der Ein- und Ausgangsgrößen p, q kleiner ist als die Anzahl der Zustandsgrößen n . Da \mathbf{u} und \mathbf{y} als Vektoren definiert werden, sind Mehrgrößensysteme in diesem Ansatz offensichtlich von vornherein enthalten. Bei Bedarf kann aus diesen Gleichungen auch ein direkter Zusammenhang zwischen Ein- und Ausgangsgrößen ermittelt werden, wie später noch gezeigt wird. Eine Beschränkung auf lineare und zeitinvariante Systeme, wie sie bisher durch die Anwendung der Laplace-Transformation erforderlich war, ist für die Zustandsdarstellung grundsätzlich nicht erforderlich.

7.2 Eigenschaften von Zustandsgrößen

Anhand eines sehr einfachen Beispiels soll nun ein Gefühl dafür vermittelt werden, was eigentlich eine Zustandsgröße ist. Gegeben sei ein Körper der Masse m , auf den eine Kraft f_a einwirkt (Abb. 7.2), wobei die Reibung vernachlässigt werden soll. Die zugehörigen Gleichungen, die sich auch im Blockschaltbild in Abb. 7.2 widerspiegeln, lauten:

$$\begin{aligned}f_a(t) &= m a(t) \\ a(t) &= \frac{dv(t)}{dt} \\ v(t) &= \frac{dl(t)}{dt}\end{aligned}\quad (7.3)$$

**Abb. 7.2.** Beschleunigter Körper

Ausgehend von einem festen Zeitpunkt t_0 soll nun die Lage l des Körpers zu einem späteren Zeitpunkt $t_1 > t_0$ ermittelt werden. Eine Umformung der Gleichungen liefert

$$\begin{aligned}
 a(t) &= \frac{1}{m} f_a(t) \\
 v(t) &= \int_{t_0}^t a(\tau) d\tau + v(t_0) \\
 l(t) &= \int_{t_0}^t v(\tau) d\tau + l(t_0)
 \end{aligned} \tag{7.4}$$

und Einsetzen in die letzte Gleichung schließlich

$$l(t_1) = \int_{t_0}^{t_1} \left[\int_{t_0}^{\tau} \frac{1}{m} f_a(\sigma) d\sigma + v(t_0) \right] d\tau + l(t_0) \tag{7.5}$$

Folgendes ist ersichtlich: Die Lage $l(t_1)$ zu einem bestimmten Zeitpunkt kann nur dann berechnet werden, wenn die Anfangswerte $l(t_0)$ und $v(t_0)$ sowie der Verlauf der Eingangsgröße des Systems $f_a(t)$ im Zeitintervall $t \in [t_0, t_1]$ bekannt sind. Frühere Vorgänge für $t < t_0$ spielen keine Rolle. Daraus folgt, dass $l(t_0)$ und $v(t_0)$ offenbar den Zustand des Systems zum Zeitpunkt t_0 vollständig charakterisieren. Mit Kenntnis dieses Zustandes und der von diesem Zeitpunkt an angreifenden Eingangsgröße $f_a(t)$ lässt sich jeder Folgezustand berechnen. $a(t_0)$ ist dafür nicht erforderlich, da zur Berechnung von $a(t)$ aus der Eingangsgröße im Gegensatz zu $v(t)$ und $l(t)$ auch keine Integration erforderlich ist. Die Beschleunigung kann sogar ohne weiteres eliminiert werden. Man erhält dann eine Zustandsdarstellung des Systems gemäß Gleichung (7.1) mit den Zustandsgrößen $x_1 = v$ und $x_2 = l$, der Ausgangsgröße $y = l$ und der Eingangsgröße $u = f_a$:

$$\begin{aligned}
 \dot{x}_1 &= \frac{dv(t)}{dt} = \frac{1}{m} f_a(t) = f_1(u) \\
 \dot{x}_2 &= \frac{dl(t)}{dt} = v(t) = f_2(x_1) \\
 y &= l(t) = g(x_2)
 \end{aligned} \tag{7.6}$$

Anhand der Form der beiden Gleichungen (7.2) lässt sich auch allgemein beweisen, dass das Systemverhalten durch die Werte der Zustandsgrößen zu einem bestimmten Zeitpunkt und den weiteren Verlauf der Eingangsgrößen eindeutig bestimmt ist. Ebenso ergibt sich durch Umstellen der Zustandsgleichungen sofort, dass die Zustandsgrößen immer durch Integration aus anderen Größen hervorgehen und daher in einem Blockschaltbild Ausgangsgrößen von Integratoren sein müssen (vgl. auch Abb. 7.2):

$$\mathbf{x}(t) = \int_{\tau=0}^t \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau + \mathbf{x}(0) \quad (7.7)$$

Da weiterhin als Argumente des Vektors \mathbf{f} keine Ableitungen auftreten, sind die Zustandsgrößen immer das Ergebnis einer Integration endlicher Größen und damit grundsätzlich stetig. Die Integratoren können daher wegen ihres nur stetig veränderlichen Inhaltes auch als Speicher interpretiert werden, was in vielen Fällen die Anschaulichkeit der Zustandsdarstellung erhöht. Als Speichergrößen kommen stetig veränderliche Größen wie Masse an Flüssigkeit in einem Behälter oder Energie in Frage. Die Zustandsgrößen repräsentieren dann beispielsweise den Energieinhalt des Systems.

Fasst man die Zustandsgrößen zu einem Zustandsvektor $\mathbf{x} = [x_1, \dots, x_n]^T$ zusammen, so beschreibt dieser Vektor einen Punkt im n -dimensionalen *Zustandsraum*. Wegen der Stetigkeit der einzelnen Komponenten bilden diese Punkte im zeitlichen Verlauf eine Trajektorie oder *Zustandskurve*. In Abb. 7.3 ist eine solche Kurve für die oben beschriebene Strecke dargestellt. Ausgehend vom Anfangszustand $l(0) = v(0) = 0$ nehmen Lage und Geschwindigkeit bei konstanter positiver Beschleunigung zunächst zu. Da die Geschwindigkeit in der Anfangsphase der Bewegung stärker ansteigt als die Lage, ergibt sich eine parabolische Kurvenform. Anschaulich gesehen muss erst eine Geschwindigkeit vorhanden sein, bevor eine Lageänderung eintritt. Zum Zeitpunkt t_1 wird die Kraft bzw. die Beschleunigung auf einen negativen Wert umgeschaltet. Die Geschwindigkeit verringert sich wieder, bis die Endposition zum Zeitpunkt t_2 erreicht ist. Berechnen lässt sich eine solche Zustandskurve, wenn man in den Zustandsgleichungen die Zeit eliminiert und v direkt in Abhängigkeit von l angibt.

Da auf die hier aufgeführten Eigenschaften von Zustandsgrößen später immer wieder Bezug genommen wird, sollen sie noch einmal zusammengefasst werden:

- Jede Zustandsgröße ist das Resultat einer Integration.
- Sofern im System nur endliche Größen auftreten, ist der Verlauf einer Zustandsgröße immer stetig.
- Zustandsgrößen können als Inhalt von Speichern aufgefasst werden, insbesondere als Inhalt von Energiespeichern.
- Die Zustandsgrößen eines Systems charakterisieren dessen Zustand vollständig. Bei Kenntnis aller Zustandsgrößen zu einem Zeitpunkt und Kennt-

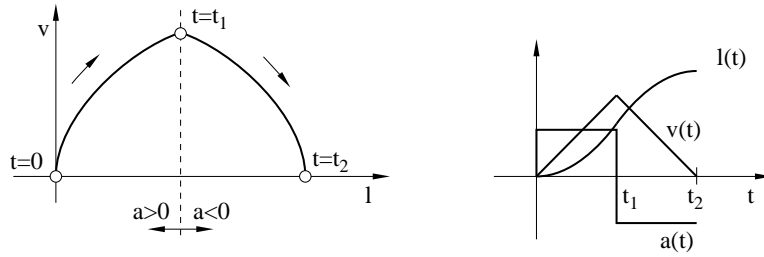


Abb. 7.3. Verstellvorgang mit konstanter Beschleunigung

nis der Eingangsgrößen ab diesem Zeitpunkt lässt sich jeder nachfolgende Systemzustand aus diesen Größen berechnen.

7.3 Linearisierung

Nachdem in Kapitel 2.7 bereits die Linearisierung für ein Übertragungselement mit einer Ein- und einer Ausgangsgröße vorgestellt wurde, soll dies hier nun auch für eine Mehrgrößenstrecke in Zustandsdarstellung noch einmal gezeigt werden, da die Linearisierung für die praktische Anwendung von hoher Bedeutung ist. Praktisch jede Strecke enthält Nichtlinearitäten, und man versucht grundsätzlich immer, die Strecke durch ein lineares Modell zu beschreiben, da dies den Reglerentwurf signifikant vereinfacht.

Für eine Linearisierung wird zunächst ein *Arbeitspunkt* festgelegt und dann das Verhalten der Strecke an diesem Arbeitspunkt durch ein lineares Modell beschrieben. Dafür ist die allgemeine Zustandsdarstellung in eine Taylorreihe um diesen gegebenen Arbeitspunkt zu entwickeln. Beispielsweise habe das nichtlineare Eingrößensystem

$$\dot{x} = f(x, u) \quad (7.8)$$

den Arbeitspunkt $x = x_0$, $u = u_0$ und $f(x_0, u_0) = 0$. Für die Abweichungen von diesem Arbeitspunkt gilt

$$\begin{aligned} \Delta x &= x - x_0 \\ \Delta u &= u - u_0 \\ \dot{\Delta x} &= \dot{x} = f(x, u) \end{aligned} \quad (7.9)$$

Die Entwicklung von $f(x, u)$ in eine Taylorreihe um den Arbeitspunkt liefert dann

$$f(x, u) = f(x_0, u_0) + \frac{\partial f}{\partial x}(x_0, u_0)\Delta x + \frac{\partial f}{\partial u}(x_0, u_0)\Delta u + r(x, u) \quad (7.10)$$

Das Restglied $r(x, u)$ enthält dabei die höheren Ableitungen und soll vernachlässigt werden. Aus (7.9) und (7.10) ergibt sich mit $f(x_0, u_0) = 0$

$$\Delta x = \frac{\partial f}{\partial x}(x_0, u_0)\Delta x + \frac{\partial f}{\partial u}(x_0, u_0)\Delta u = a\Delta x + b\Delta u \quad (7.11)$$

also eine lineare Differentialgleichung für die Abweichungen vom Arbeitspunkt mit den Koeffizienten a und b . In entsprechender Weise ist die Ausgangsgleichung $y = g(x, u)$ zu linearisieren. Damit ist das nichtlineare System am Arbeitspunkt durch ein lineares Modell dargestellt und kann nun mit Methoden der linearen Regelungstechnik behandelt werden.

Im Mehrgrößenfall ist die Vorgehensweise analog. Für

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad (7.12)$$

erhält man als Ergebnis der Linearisierung für $\mathbf{f}(\mathbf{x}_0, \mathbf{u}_0) = \mathbf{0}$

$$\dot{\Delta \mathbf{x}} = \mathbf{F}_x(\mathbf{x}_0, \mathbf{u}_0)\Delta \mathbf{x} + \mathbf{F}_u(\mathbf{x}_0, \mathbf{u}_0)\Delta \mathbf{u} = \mathbf{A}\Delta \mathbf{x} + \mathbf{B}\Delta \mathbf{u} \quad (7.13)$$

wobei die einzelnen Elemente der Jacobimatrizen \mathbf{F}_x und \mathbf{F}_u definiert sind durch

$$[\mathbf{F}_x]_{i,j} = \frac{\partial f_i}{\partial x_j} \quad \text{und} \quad [\mathbf{F}_u]_{i,j} = \frac{\partial f_i}{\partial u_j} \quad (7.14)$$

mit $\mathbf{f} = [f_1, \dots, f_n]^T$.

Die Linearisierung ist ein häufig eingesetztes Mittel, um die Werkzeuge und Entwurfsverfahren der linearen Regelungstechnik auf eine nichtlineare Strecke anwenden zu können. Die nichtlineare Strecke wird linearisiert, und für das lineare Modell wird dann ein linearer Regler entworfen. Dabei ist aber zu berücksichtigen, dass die Abweichungen zwischen Modell und realer Strecke mit zunehmender Entfernung vom Arbeitspunkt immer größer werden. Der lineare Regler muss deshalb eine ausreichende Robustheit (siehe späteres Kapitel) aufweisen, d.h. er muss auch im Zusammenspiel mit einer Strecke, die sich von dem Modell unterscheidet, für das er entworfen wurde, noch Stabilität garantieren.

Hohe Robustheit bedeutet aber in der Regel einen Verzicht auf Regelgeschwindigkeit. Wenn daher die Nichtlinearitäten der Strecke so beschaffen sind, dass das Verhalten der realen Strecke schon in relativ kleiner Entfernung vom Arbeitspunkt stark vom linearisierten Modell abweicht, wäre beim Regler eine sehr hohe Robustheit erforderlich, was wiederum in einer nicht mehr akzeptablen Regelgeschwindigkeit resultiert. In einem solchen Fall bringt eine Linearisierung daher keinen Nutzen mehr.

7.4 Zustandsdarstellung linearer Systeme

Für die weiteren Ausführungen sei angenommen, dass ein lineares Modell der Strecke vorliegt, welches entweder eine lineare Strecke repräsentiert oder aus einer Linearisierung hervorgegangen ist. Man beschränkt sich also auf lineare und zeitinvariante Systeme ohne Totzeiten. Die Vektorfunktionen $\mathbf{f} =$

$[f_1, \dots, f_n]^T$ und $\mathbf{g} = [g_1, \dots, g_m]^T$ werden dadurch zu linearen Funktionen von den Zustands- und Eingangsgrößen. Damit lassen sich die Gleichungen (7.2) auch schreiben als

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\end{aligned}\quad (7.15)$$

\mathbf{A} , \mathbf{B} , \mathbf{C} und \mathbf{D} sind Matrizen mit konstanten Koeffizienten. \mathbf{A} bezeichnet man als *Systemmatrix*, \mathbf{B} als *Eingangsmatrix*, \mathbf{C} als *Ausgangsmatrix* und \mathbf{D} als *Durchgangsmatrix*. Abb. 7.4 skizziert die Zusammenhänge. Der Integrator steht dabei für eine komponentenweise Integration des Vektors $\dot{\mathbf{x}}$.

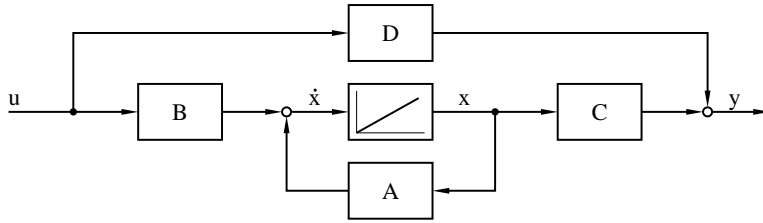


Abb. 7.4. Zustandsdarstellung einer linearen Strecke

Im Eingrößenfall hat das System zwar nur eine Ein- und Ausgangsgröße, d.h. u und y sind Skalare, aber weiterhin beliebig viele Zustandsgrößen. \mathbf{A} ist daher immer von der Ordnung $n \times n$, wobei n die Streckenordnung kennzeichnet. Dagegen wird \mathbf{B} im Eingrößenfall zu einer $n \times 1$ -Matrix (Spaltenvektor), \mathbf{C} zu einer $1 \times n$ -Matrix (Zeilenvektor) und \mathbf{D} zu einem Skalar.

Zur Verdeutlichung soll das oben angegebene Beispiel um eine Feder und eine geschwindigkeitsabhängige Reibung erweitert werden, so dass sich das schon in Abb. 2.1 dargestellte System ergibt. Umformen der Gleichungen (2.1) bis (2.5) und Elimination der Beschleunigung liefert die Zustandsgleichungen

$$\begin{aligned}\frac{dv(t)}{dt} &= \frac{1}{m}f_a(t) - \frac{c_r}{m}v(t) - \frac{c_f}{m}l(t) \\ \frac{dl(t)}{dt} &= v(t)\end{aligned}\quad (7.16)$$

und die triviale Ausgangsgleichung

$$l(t) = l(t) \quad (7.17)$$

In Matrizenschreibweise lauten die Gleichungen

$$\begin{aligned}\dot{\mathbf{x}} = \begin{bmatrix} \dot{v} \\ \dot{l} \end{bmatrix} &= \begin{bmatrix} -\frac{c_r}{m} & -\frac{c_f}{m} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ l \end{bmatrix} + \begin{bmatrix} \frac{1}{m} \\ 0 \end{bmatrix} f_a = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} = l &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ l \end{bmatrix} + 0 f_a = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\end{aligned}\quad (7.18)$$

Hier bietet sich eine Deutung der Zustandsgrößen als Energieinhalt des Systems an. Und zwar repräsentiert die Lage die Auslenkung der Feder und damit die potentielle Energie des Systems, während die Geschwindigkeit ein Maß für die kinetische Energie darstellt.

Wegen $\mathbf{D} = \mathbf{0}$ ist in diesem Beispiel die Ausgangsgröße ausschließlich eine Linearkombination der Zustandsgrößen und nicht direkt von der Eingangsgröße abhängig. Damit kann sich die Ausgangsgröße aber ebenso wie die Zustandsgrößen auch bei sprungförmiger Eingangsgröße nur stetig verändern. Dies gilt für alle Strecken, bei denen der Übertragungskanal vom Ein- zum Ausgang einen Tiefpass wie z.B. einen Integrator oder ein Verzögerungsglied enthält, denn die Ausgangsgröße eines solchen Übertragungsgliedes und damit natürlich auch die Ausgangsgröße des Systems kann grundsätzlich nur stetig verlaufen. Da weiterhin fast alle realen Strecken tiefpasshaltig sind, ist der Fall $\mathbf{D} = \mathbf{0}$ der Normalfall in der Praxis und wird auch als Voraussetzung für manche Reglerauslegungsverfahren benötigt.

Zu beachten ist außerdem, dass Laufzeitglieder, obwohl sie lineare Übertragungsglieder sind, in einer Zustandsgleichung nicht dargestellt werden können. Dies widerspricht auch dem Gedanken, dass sich aus dem momentanen Zustand der Strecke alle folgenden Vorgänge im System berechnen lassen können, denn zur Berechnung der Ausgangsgröße eines Laufzeitgliedes wäre die Kenntnis eines vergangenen Zustandes notwendig. Bei Laufzeitgliedern bleibt demnach nur die Approximation durch Verzögerungsglieder erster oder höherer Ordnung, die dann wiederum problemlos in Zustandsform darstellbar sind.

7.5 Normalformen

In den beiden vorangegangenen Beispielen waren die gegebenen Differentialgleichungen der Strecke von vornherein erster Ordnung. Dies hatte zur Folge, dass keine zusätzlichen Größen als Zustandsgrößen eingeführt werden mussten. Sie ergaben sich aus der Struktur der Strecke und hatten deshalb auch eine reale physikalische Entsprechung in den Größen *Geschwindigkeit* und *Lage*. Dies muss nicht immer so sein. Grundsätzlich können die Zustandsgrößen nach beliebigen Kriterien festgelegt werden, beispielsweise, um der Systemmatrix \mathbf{A} eine bestimmte Form zu verleihen.

7.5.1 Regelungs- und Beobachtungs-Normalform

Einige spezielle Formen, die für spätere Betrachtungen noch benötigt werden, sollen im Folgenden vorgestellt und hergeleitet werden. Den Beginn macht die sogenannte *Regelungsnormalform*.

Gegeben sei ein Eingrößensystem mit der Übertragungsfunktion

$$G(s) = \frac{b_n s^n + b_{n-1} s^{n-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0} = \frac{y(s)}{u(s)} \quad (7.19)$$

die so normiert wurde, dass der Vorfaktor a_n von s^n gerade gleich Eins ist und somit entfallen kann.

Zunächst soll versucht werden, ein zu dieser Übertragungsfunktion gehörendes Blockschaltbild zu entwickeln, das nur Integratoren, Multiplikationen mit konstanten Faktoren sowie Addition und Subtraktion enthält. Da es sich hier um lineare Operationen handelt und diese, wie schon in Kapitel 3.3 gezeigt, in ihrer Reihenfolge beliebig vertauscht werden können, gibt es offensichtlich verschiedene Möglichkeiten, ein solches Blockschaltbild zu konstruieren.

Zunächst wird eine Hilfs-Übertragungsfunktion eingeführt:

$$G_1(s) := \frac{1}{s^n + a_{n-1}s^{n-1} + \dots + a_0} \quad (7.20)$$

Mit dieser ergibt sich

$$y(s) = b_n s^n G_1(s)u(s) + b_{n-1} s^{n-1} G_1(s)u(s) + \dots + b_0 G_1(s)u(s) \quad (7.21)$$

Nun werden neue Variablen definiert:

$$x_1(s) := G_1(s)u(s) \quad (7.22)$$

$$x_2(s) := sG_1(s)u(s) \quad (7.23)$$

$$x_3(s) := s^2 G_1(s)u(s) \quad (7.24)$$

...

$$x_n(s) := s^{n-1} G_1(s)u(s) \quad (7.25)$$

Offensichtlich gilt

$$x_i(s) = s x_{i-1}(s) \quad (7.26)$$

$$x_i(s) = s^{i-1} x_1(s) \quad (7.27)$$

und im Zeitbereich

$$x_i(t) = \dot{x}_{i-1}(t) \quad (7.28)$$

$$x_i(t) = \left(\frac{d}{dt}\right)^{i-1} x_1(t) \quad (7.29)$$

Aus (7.22) und (7.20) folgt

$$x_1(s) = G_1(s)u(s) = \frac{1}{s^n + a_{n-1}s^{n-1} + \dots + a_0} u(s) \quad (7.30)$$

$$s^n x_1(s) + a_{n-1} s^{n-1} x_1(s) + \dots + a_0 x_1(s) = u(s) \quad (7.31)$$

Unter Verwendung von (7.27) wird daraus

$$s x_n(s) + a_{n-1} x_n(s) + a_{n-2} x_{n-1}(s) + \dots + a_0 x_1(s) = u(s) \quad (7.32)$$

und nach Umstellen und Transformation in den Zeitbereich

$$\dot{x}_n(t) = u(t) - a_{n-1}x_n(t) - a_{n-2}x_{n-1}(t) - \dots - a_0x_1(t) \quad (7.33)$$

Aus (7.21) wird mit (7.22) bis (7.25)

$$y(s) = b_n s x_n(s) + b_{n-1}x_n(s) + b_{n-2}x_{n-1}(s) + \dots + b_0x_1(s) \quad (7.34)$$

und im Zeitbereich

$$y(t) = b_n \dot{x}_n(t) + b_{n-1}x_n(t) + b_{n-2}x_{n-1}(t) + \dots + b_0x_1(t) \quad (7.35)$$

Aus (7.28), (7.33) und (7.35) lässt sich direkt das Blockschaltbild in Regelungsnormalform entwickeln, das in Abb. 7.5 oben gezeigt ist.

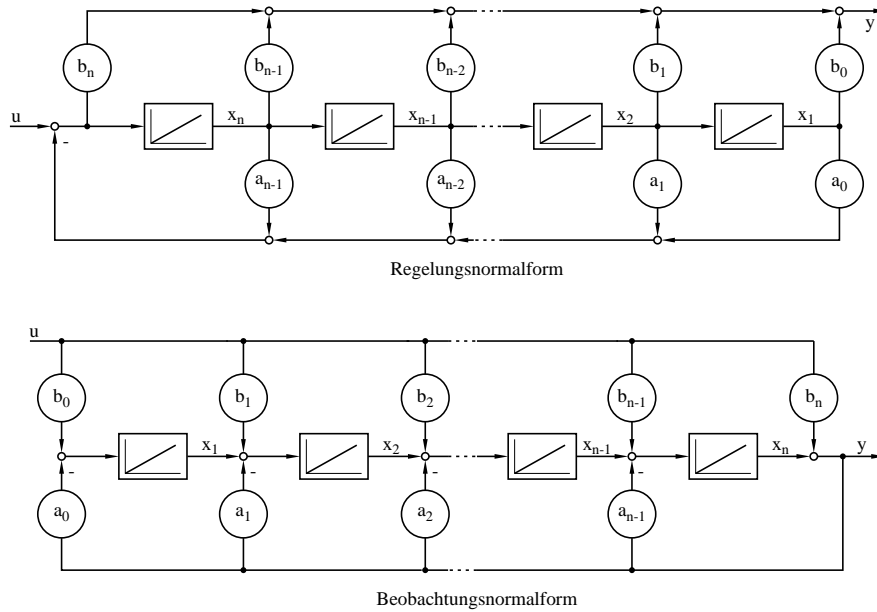


Abb. 7.5. Regelungsnormalform und Beobachtungsnormalform

Zunächst gewinnt man aus (7.28) die Information, dass jede Zustandsgröße die Ableitung der vorhergehenden Zustandsgröße ist, oder andersherum, dass jede Zustandsgröße durch Integration aus der Zustandsgröße mit dem nächsthöheren Index hervorgeht. Daraus ergibt sich sofort eine Kette von Integratoren, wobei jede Zustandsgröße sowohl Ausgangsgröße des einen Integrators als auch Eingangsgröße des nächsten Integrators sein muss.

Als nächstes nutzt man Gleichung (7.33), wobei man sich für $\dot{x}_n(t)$ nur klar machen muss, dass $\dot{x}_n(t)$ die Eingangsgröße desjenigen Integrators sein muss, an dessen Ausgang $x_n(t)$ erscheint. Damit ergibt sich sofort die Rückkopplungsstruktur mit den Koeffizienten a_i .

Als letztes schließlich verwendet man die Ausgangsgleichung (7.35). Hier hat man die Zustandsgrößen aus der Kette von Integratoren abzugreifen und

jeweils mit dem entsprechenden Koeffizienten b_i zu multiplizieren, dann die Ergebnisse der einzelnen Multiplikationen aufzusummieren, und als Ergebnis erhält man $y(t)$.

Ähnlich wie die Regelungsnormalform lässt sich auch die *Beobachtungsnormalform* herleiten, die in Abb. 7.5 unten gezeigt ist. Aufgrund der Ähnlichkeit soll auf die Herleitung hier aber verzichtet werden.

Beide Normalformen sind äquivalent zueinander, wie man durch Bilden der Übertragungsfunktion aus dem jeweiligen Blockschaltbild leicht nachvollziehen kann. Sie zeichnen sich dadurch aus, dass die Koeffizienten der Übertragungsfunktion auch direkt als Faktoren im Blockschaltbild auftauchen.

Beide Darstellungen weisen n Integratoren auf, deren Ausgänge jeweils die n Zustandsgrößen sind. Entsprechend handelt es sich hier um ein System der Ordnung n , was sich auch sofort aus der Ordnung des Nenners der Übertragungsfunktion ergibt.

Für jede einzelne Zustandsgröße lässt sich nun anhand des Blockschaltbildes die zugehörige Differentialgleichung erster Ordnung angeben. Für die Regelungsnormalform sind dies die Gleichungen (7.28) und (7.33), und für die Beobachtungsnormalform lautet beispielsweise die Gleichung für die Zustandsgröße x_n :

$$x_n(t) = \int_{\tau=0}^t b_{n-1}u(\tau) + x_{n-1}(\tau) - a_{n-1}(x_n(\tau) + b_n u(\tau))d\tau + x_n(0) \quad (7.36)$$

bzw.

$$\dot{x}_n = b_{n-1}u + x_{n-1} - a_{n-1}(x_n + b_n u) \quad (7.37)$$

Stellt man für jede der n Zustandsgrößen eine derartige Differentialgleichung auf und fasst anschließend alle Gleichungen zusammen, so ergibt sich für die Zustandsdarstellung des Systems in Regelungsnormalform

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ \cdots \\ \cdots \\ \cdots \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [b_0 - b_n a_0, b_1 - b_n a_1, \dots, b_{n-1} - b_n a_{n-1}] \mathbf{x} + b_n u \quad (7.38)$$

und in Beobachtungsnormalform:

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_0 - b_n a_0 \\ b_1 - b_n a_1 \\ b_2 - b_n a_2 \\ \cdots \\ b_{n-1} - b_n a_{n-1} \end{bmatrix} u$$

$$y = [0, \dots, 0, 1] \mathbf{x} + b_n u \quad (7.39)$$

Man sieht, dass sich aufgrund anders gewählter Zustandsgrößen für beide Normalformen unterschiedliche Matrizen ergeben. Dennoch handelt es sich natürlich in beiden Fällen um ein- und dieselbe Strecke.

Angemerkt sei, dass beide Normalformen nur durch die Form ihrer Systemmatrizen definiert sind, während die übrigen Matrizen nicht unbedingt die Form aus (7.38) bzw. (7.39) aufweisen müssen. So können auch Mehrgrößensysteme unter bestimmten Voraussetzungen in einer der beiden Normalformen beschrieben werden, wobei dann natürlich die Matrizen **B**, **C** und **D** eine völlig andere Gestalt haben.

Interessant ist in diesem Zusammenhang eine kurze Diskussion des Koeffizienten b_n . Für $b_n \neq 0$ hat der Zähler der Übertragungsfunktion die gleiche Ordnung wie der Nenner, was bedeutet, dass das System auf eine sprungförmige Anregung zum Zeitpunkt $t = 0$ mit einem Sprung der Ausgangsgröße zum Zeitpunkt $t = 0$ reagiert. Diese direkte Wirkung vom Ein- zum Ausgang erkennt man auch an den Blockschaltbildern der beiden Normalformen: Der direkte Weg vom Ein- zum Ausgang führt in beiden Darstellungen über den Koeffizienten b_n . Und bei den Zustandsgleichungen erkennt man die direkte Wirkung vom Ein- zum Ausgang daran, dass die Durchgangsmatrix **D**, die im vorliegenden Fall nur aus einem Skalar (b_n) besteht, ungleich Null ist.

Falls $b_n = 0$ gilt, so existiert keine direkte Wirkung vom Ein- auf den Ausgang. Die Übertragungsfunktion weist im Nenner eine höhere Ordnung als im Zähler auf, der Koeffizient b_n verschwindet in den Blockschaltbildern, die Zustandsgleichungen vereinfachen sich, und vor allem ist dort **D** = **0**. All diese Eigenschaften zeigen an, dass die reale Strecke auf eine sprungförmige Eingangsgröße nur verzögert reagieren wird und somit Tiefpassverhalten aufweist.

7.5.2 Jordansche Normalform

Die dritte Normalform, die hier vorgestellt werden soll, ist die *Jordansche Normalform*. Sie zeichnet sich dadurch aus, dass die Systemmatrix **A** eine reine Diagonalmatrix ist, wobei die Elemente dieser Diagonalmatrix gerade ihre Eigenwerte sind. Nur im Fall eines komplex-konjugierten Pol- bzw. Eigenwertpaares entsteht eine Abweichung von der Diagonalform, wie nachfolgend noch gezeigt wird. Die übrigen Matrizen **B**, **C** und **D** haben keine spezielle Form, auch wenn sich in der nachfolgenden Herleitung, die von einer Eingrößenstrecke ausgeht, für die Eingangsmatrix ein Vektor aus lauter Einsen ergibt.

Für die Herleitung dieser Normalform soll die Übertragungsfunktion (7.19) zunächst in Partialbrüche zerlegt werden (vgl. (3.30)):

$$G(s) = \sum_{\lambda=1}^i \sum_{\nu=1}^{n_\lambda} \frac{r_{\lambda\nu}}{(s - s_\lambda)^\nu} + r_0 \quad (7.40)$$

i ist die Anzahl unterschiedlicher Pole, und n_λ die Vielfachheit des Poles s_λ . Offenbar gilt $n = \sum_{\lambda=1}^i n_\lambda$. Zunächst sei nur von Einzelpolen ausgegangen, so dass sich die Darstellung zu

$$G(s) = \sum_{\lambda=1}^n \frac{r_\lambda}{s - s_\lambda} + r_0 \quad (7.41)$$

vereinfacht. Nun werden die Zustandsgrößen wie folgt definiert:

$$\begin{aligned} x_\lambda(s) &= \frac{1}{s - s_\lambda} u(s) \\ s x_\lambda(s) &= s_\lambda x_\lambda(s) + u(s) \\ \dot{x}_\lambda(t) &= s_\lambda x_\lambda(t) + u(t) \end{aligned} \quad (7.42)$$

Diese Definitionsgleichungen bilden schon die Zustandsgleichungen des Systems. Ausgehend von (7.41) erhält man die Ausgangsgleichung

$$\begin{aligned} y(s) &= G(s)u(s) \\ y(s) &= \sum_{\lambda=1}^n \frac{r_\lambda}{s - s_\lambda} u(s) + r_0 u(s) \\ y(s) &= \sum_{\lambda=1}^n r_\lambda x_\lambda(s) + r_0 u(s) \\ y(t) &= \sum_{\lambda=1}^n r_\lambda x_\lambda(t) + r_0 u(t) \end{aligned} \quad (7.43)$$

Fasst man diese Gleichungen in Matrizenform zusammen, so ergibt sich

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & 0 & \dots & 0 \\ 0 & 0 & s_3 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & s_n \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} u \\ y &= [r_1, \dots, r_n] \mathbf{x} + r_0 u \end{aligned} \quad (7.44)$$

Die Systemmatrix ist also eine reine Diagonalmatrix, und die Pole der Übertragungsfunktion erscheinen als Elemente der Hauptdiagonalen. Wenn die Ordnung des Zählers der Übertragungsfunktion gleich der ihres Nenners ist, dann ist $r_0 \neq 0$ und damit auch $\mathbf{D} \neq \mathbf{0}$.

Nun soll der Fall betrachtet werden, dass der erste Pol m -fach und alle andere Pole einfach vorkommen. Für die Partialbruchzerlegung ergibt sich

$$G(s) = \frac{r_1}{s - s_1} + \frac{r_2}{(s - s_1)^2} + \dots + \frac{r_m}{(s - s_1)^m} + \sum_{\lambda=m+1}^n \frac{r_\lambda}{s - s_\lambda} + r_0 \quad (7.45)$$

Die Definition der Zustandsgrößen erfolgt ähnlich zum vorhergehenden Fall:

$$\begin{aligned}
 x_1(s) &= \frac{1}{s - s_1} u(s) \\
 x_2(s) &= \frac{1}{(s - s_1)^2} u(s) = \frac{1}{s - s_1} x_1(s) \\
 x_m(s) &= \frac{1}{(s - s_1)^m} u(s) = \frac{1}{s - s_1} x_{m-1}(s) \\
 x_i(s) &= \frac{1}{s - s_i} u(s) \quad \text{für } i > m
 \end{aligned} \tag{7.46}$$

Im Zeitbereich ergibt sich

$$\begin{aligned}
 \dot{x}_1(t) &= s_1 x_1(t) + u(t) \\
 \dot{x}_2(t) &= s_1 x_2(t) + x_1(t) \\
 \dot{x}_m(t) &= s_1 x_m(t) + x_{m-1}(t) \\
 \dot{x}_i(t) &= s_i x_i(t) + u(t) \quad \text{für } i > m
 \end{aligned} \tag{7.47}$$

In Matrizenform erhält man

$$\dot{\mathbf{x}} = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 1 & s_1 & 0 & \dots & 0 \\ 0 & 1 & s_1 & 0 & \dots & 0 \\ 0 & \dots & 0 & s_{m+1} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & s_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_m \\ x_{m+1} \\ \dots \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ \dots \\ 1 \end{bmatrix} u$$

$$y = [r_1, \dots, r_n] \mathbf{x} + r_0 u \tag{7.48}$$

Falls mehrere Pole mit einer Vielfachheit größer als eins auftauchen, ist offensichtlich, wie die Systemmatrix dann aussehen würde, so dass dieser Fall hier nicht mehr betrachtet werden muss.

Es bleibt als letzter zu betrachtender Fall, dass ein komplex-konjugiertes Polpaar auftaucht. Hierzu sei angenommen, dass die Pole s_1 und s_2 komplex-konjugiert und alle anderen Pole einfach sind. Die Zustandsgrößen werden genau wie im Fall einfacher Polstellen definiert, und für die beiden ersten Zustandsgleichungen ergibt sich:

$$\dot{x}_1 = s_1 x_1 + u \tag{7.49}$$

$$\dot{x}_2 = s_2 x_2 + u \tag{7.50}$$

$$s_1 = a + jb \quad s_2 = a - jb \tag{7.51}$$

Daraus folgt, dass auch x_1 und x_2 komplex-konjugiert zueinander sein müssen:

$$x_1 = r + jq \quad x_2 = r - jq \tag{7.52}$$

Einsetzen von (7.51) und (7.52) in (7.49) liefert

$$\dot{x}_1 = (ar - bq) + j(br + aq) + u \quad (7.53)$$

Andererseits ergibt sich aus (7.52) aber auch

$$\dot{x}_1 = \dot{r} + j\dot{q} \quad (7.54)$$

Ein Vergleich der Real- und Imaginärteile liefert die beiden Gleichungen

$$\begin{aligned} \dot{r} &= ar - bq + u \\ \dot{q} &= br + aq \end{aligned} \quad (7.55)$$

Diese beiden Gleichungen enthalten offenbar dieselbe Information und beschreiben denselben dynamischen Zusammenhang wie die beiden Zustandsgleichungen (7.49) und (7.50). Von daher können (7.49) und (7.50) auch durch die Gleichungen (7.55) ersetzt werden, wobei natürlich auch die Zustandsgrößen (x_1, x_2) durch (r, q) bzw. $(\text{Re}(x_1), \text{Im}(x_1))$ zu ersetzen sind.

Für das Gesamtsystem erhält man damit

$$\begin{aligned} \dot{\mathbf{x}} = \begin{bmatrix} \text{Re}(\dot{x}_1) \\ \text{Im}(\dot{x}_1) \\ \dot{x}_3 \\ \vdots \\ \dot{x}_n \end{bmatrix} &= \begin{bmatrix} \text{Re}(s_1) & -\text{Im}(s_1) & 0 & \cdots & 0 \\ \text{Im}(s_1) & \text{Re}(s_1) & 0 & \cdots & 0 \\ 0 & 0 & s_3 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & & & 0 & s_n \end{bmatrix} \begin{bmatrix} \text{Re}(x_1) \\ \text{Im}(x_1) \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} u \\ y &= [r_1, \dots, r_n] \mathbf{x} + r_0 u \end{aligned} \quad (7.56)$$

wobei s_2 komplex-konjugiert zu s_1 ist.

7.6 Koordinatentransformation

7.6.1 Grundlagen der Koordinatentransformation

Allgemein lässt sich die Festlegung von Zustandsgrößen auch als Definition eines Koordinatensystems im n -dimensionalen Zustandsraum interpretieren. Ein gegebener Systemzustand zu einem bestimmten Zeitpunkt t_1 ist ein Punkt in diesem Raum. Dieser Punkt ist gerade durch den Wert jeder einzelnen Zustandsgröße zum Zeitpunkt t_1 definiert. Bei anders gewählten Zustandsgrößen ergeben sich zwar andere Werte, der Punkt und damit der Systemzustand muss aber immer derselbe bleiben. Eine andere Wahl der Zustandsgrößen entspricht damit lediglich einer Änderung bzw. Transformation des Koordinatensystems.

Abbildung 7.6 zeigt die Verhältnisse im zweidimensionalen Raum. Der durch den Punkt markierte Zustand hat in b-Koordinaten die Darstellung $(b_1, b_2) = (1, 2)$ und in t-Koordinaten die Darstellung $(t_1, t_2) \approx (1.67, 1.53)$.

**Abb. 7.6.** Basistransformation

Eine Koordinatentransformation lässt sich durch eine reguläre und deshalb auch invertierbare Transformationsmatrix \mathbf{T} beschreiben. Sie besteht aus den Vektoren der neuen Basis, ausgedrückt in den Koordinaten der alten Basis:

$$\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n] \quad \text{mit} \quad \mathbf{t}_i = t_{i1}\mathbf{b}_1 + \dots + t_{in}\mathbf{b}_n = \sum_{k=1}^n t_{ki}\mathbf{b}_k \quad (7.57)$$

Gegeben sei nun ein Punkt im Raum, der in Koordinaten der alten Basis durch $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ und in Koordinaten der neuen Basis durch $\mathbf{z} = (z_1, z_2, \dots, z_n)^T$ beschrieben wird. Definitionsgemäß handelt es sich dabei nur um eine abgekürzte Schreibweise. Ausgeschrieben ergibt sich z.B. für \mathbf{x} (mit den Basisvektoren \mathbf{b}_i der Ausgangsbasis):

$$\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \dots + x_n\mathbf{b}_n \quad (7.58)$$

Analog dazu lässt sich auch für den Vektor \mathbf{z} , der ja denselben Punkt im Raum beschreibt, eine entsprechende Vektorsumme mit den neuen Basisvektoren angeben:

$$\begin{aligned} z_1\mathbf{t}_1 + z_2\mathbf{t}_2 + \dots + z_n\mathbf{t}_n &= \\ z_1\left(\sum_{k=1}^n t_{k1}\mathbf{b}_k\right) + z_2\left(\sum_{k=1}^n t_{k2}\mathbf{b}_k\right) + \dots + z_n\left(\sum_{k=1}^n t_{kn}\mathbf{b}_k\right) &= \\ \left(\sum_{k=1}^n t_{1k}z_k\right)\mathbf{b}_1 + \left(\sum_{k=1}^n t_{2k}z_k\right)\mathbf{b}_2 + \dots + \left(\sum_{k=1}^n t_{nk}z_k\right)\mathbf{b}_n & \quad (7.59) \end{aligned}$$

Weiterhin gilt

$$\mathbf{T}\mathbf{z} = \begin{bmatrix} \sum_{k=1}^n t_{1k}z_k \\ \sum_{k=1}^n t_{2k}z_k \\ \dots \\ \sum_{k=1}^n t_{nk}z_k \end{bmatrix} \quad (7.60)$$

Man erkennt, dass der Vorfaktor von \mathbf{b}_i in Gleichung (7.59) gerade die i -te Komponente von $\mathbf{T}\mathbf{z}$ ist. Damit liefert der Vergleich von (7.58) und (7.59) den Zusammenhang

$$\mathbf{Tz} = \mathbf{x} \quad \text{bzw.} \quad \mathbf{z} = \mathbf{T}^{-1}\mathbf{x} \quad (7.61)$$

Einsetzen von $\mathbf{x} = \mathbf{Tz}$ und $\dot{\mathbf{x}} = \mathbf{T}\dot{\mathbf{z}}$ in (7.15) liefert die Zustandsdarstellung für die neuen Zustandsgrößen

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{T}^{-1}\mathbf{ATz} + \mathbf{T}^{-1}\mathbf{Bu} \\ \mathbf{y} &= \mathbf{CTz} + \mathbf{Du} \end{aligned} \quad (7.62)$$

mit den neuen Matrizen

$$\mathbf{A}' = \mathbf{T}^{-1}\mathbf{AT} \quad \mathbf{B}' = \mathbf{T}^{-1}\mathbf{B} \quad \mathbf{C}' = \mathbf{CT} \quad \mathbf{D}' = \mathbf{D} \quad (7.63)$$

Da das neue System aber lediglich durch eine Koordinatentransformation aus dem alten hervorgegangen ist, müssen beide Darstellungen vollständig äquivalent sein.

In der deutschsprachigen Literatur wird die Transformationsmatrix einer Basistransformation teilweise auch durch $\mathbf{z} = \mathbf{T}\mathbf{x}$ definiert, was zur Folge hat, dass in den Gleichungen (7.62) und (7.63) jeweils \mathbf{T} durch \mathbf{T}^{-1} ersetzt werden muss. International ist jedoch die oben eingeführte Definition (7.61) gebräuchlicher, die darüber hinaus der Definition einer Transformationsmatrix in der linearen Algebra entspricht und in diesem Buch daher auch beibehalten werden soll.

Weiterhin sind die Eigenwerte der Systemmatrix \mathbf{A} als die Nullstellen s_λ der Determinante

$$|s\mathbf{I} - \mathbf{A}| \quad (7.64)$$

definiert. Außerdem gilt für quadratische Matrizen \mathbf{X} , \mathbf{Y} und \mathbf{Z}

$$|\mathbf{XYZ}| = |\mathbf{X}| |\mathbf{Y}| |\mathbf{Z}| \quad (7.65)$$

Deshalb entsprechen wegen

$$\begin{aligned} |s\mathbf{I} - \mathbf{T}^{-1}\mathbf{AT}| &= |s\mathbf{T}^{-1}\mathbf{IT} - \mathbf{T}^{-1}\mathbf{AT}| = |\mathbf{T}^{-1}(s\mathbf{I} - \mathbf{A})\mathbf{T}| \\ &= |\mathbf{T}^{-1}| |s\mathbf{I} - \mathbf{A}| |\mathbf{T}| = |s\mathbf{I} - \mathbf{A}| \end{aligned} \quad (7.66)$$

die Eigenwerte der neuen Systemmatrix gerade den Eigenwerten der alten. Diese Tatsache wird für die folgenden Herleitungen noch benötigt.

7.6.2 Transformation auf Jordansche Normalform

Für bestimmte Regler-Auslegungsverfahren wird vorausgesetzt, dass das System in einer der oben vorgestellten Normalformen gegeben ist. Von daher soll nachfolgend gezeigt werden, wie ein System, das in allgemeiner Zustandsdarstellung durch die Matrizen $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ gegeben ist, auf eine der Normalformen transformiert werden kann.

Den Anfang macht die Jordansche Normalform. Die Aufgabenstellung lautet wie folgt: Gegeben sei ein System in Zustandsdarstellung mit der Systemmatrix \mathbf{A} . Gesucht ist eine Transformationsmatrix \mathbf{T} , mit der das System auf Jordansche Normalform transformiert wird, so dass gilt

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix} = \mathbf{L} \quad (7.67)$$

bzw.

$$\mathbf{A}\mathbf{T} = \mathbf{T}\mathbf{L} \quad (7.68)$$

Für die Transformationsmatrix gilt wegen (7.57)

$$\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n] \quad (7.69)$$

und damit

$$\mathbf{A}\mathbf{T} = [\mathbf{A}\mathbf{t}_1, \mathbf{A}\mathbf{t}_2, \dots, \mathbf{A}\mathbf{t}_n] \quad (7.70)$$

Für das Produkt $\mathbf{T}\mathbf{L}$ ergibt sich mit \mathbf{L} gemäß (7.67) und \mathbf{T} gemäß (7.69)

$$\mathbf{T}\mathbf{L} = [\lambda_1\mathbf{t}_1, \lambda_2\mathbf{t}_2, \dots, \lambda_n\mathbf{t}_n] \quad (7.71)$$

Da $\mathbf{A}\mathbf{T} = \mathbf{T}\mathbf{L}$ gefordert ist, ergibt der komponentenweise Vergleich nach diesen beiden Gleichungen für die Spalten \mathbf{t}_i von \mathbf{T} die Forderung

$$\begin{aligned} \mathbf{A}\mathbf{t}_i &= \lambda_i\mathbf{t}_i = \lambda_i\mathbf{I}\mathbf{t}_i \\ (\lambda_i\mathbf{I} - \mathbf{A})\mathbf{t}_i &= \mathbf{0} \end{aligned} \quad (7.72)$$

$\mathbf{t}_i = \mathbf{0}$ ist die triviale Lösung dieser Gleichung, die uns aber hier nicht weiter hilft, da $\mathbf{T} = \mathbf{0}$ keine Transformationsmatrix wäre. Eine nicht-triviale Lösung $\mathbf{t}_i \neq \mathbf{0}$ existiert genau dann, wenn für jedes λ_i

$$\text{Rang}(\lambda_i\mathbf{I} - \mathbf{A}) < n \quad \text{bzw.} \quad \det(\lambda_i\mathbf{I} - \mathbf{A}) = 0 \quad (7.73)$$

gilt. Dies ist jedoch gerade die Definitionsgleichung für einen Eigenwert der Matrix \mathbf{A} . Und wenn λ_i ein Eigenwert von \mathbf{A} ist, dann ergibt sich aus

$$(\lambda_i\mathbf{I} - \mathbf{A})\mathbf{t}_i = \mathbf{0} \quad (7.74)$$

gerade der zu λ_i gehörende Eigenvektor.

Zusammenfassend ergibt sich damit, dass die Transformationsmatrix $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n]$, mit deren Hilfe eine Systemmatrix \mathbf{A} bzw. ein System nach (7.15) auf die Jordansche Normalform transformiert werden kann, gerade aus den Eigenvektoren von \mathbf{A} besteht. Voraussetzung dafür, dass \mathbf{T} tatsächlich Matrix einer Basistransformation ist, ist jedoch, dass sie den Rang n aufweist, d.h. sämtliche Eigenvektoren \mathbf{t}_i müssen voneinander linear unabhängig sein. In diesem Fall bezeichnet man \mathbf{A} als *diagonalähnlich*.

Neben einigen anderen Fällen ist \mathbf{A} dann diagonalähnlich, wenn \mathbf{A} symmetrisch ist oder sämtliche Eigenwerte von \mathbf{A} verschieden sind. Wenn \mathbf{A} nicht diagonalähnlich ist, dann kann das gegebene System nicht auf die Jordansche Normalform transformiert werden.

Sofern ein System in Jordanscher Normalform vorliegt, bezeichnet man die dem zugehörigen Zustandsvektor \mathbf{z} zugrunde liegende Basis als *Modalkoordinaten*.

7.6.3 Transformation auf Regelungsnormalform

Nun soll für SISO-Systeme gezeigt werden, wie man ein System auf die Regelungsnormalform (vgl. (7.38)) mit der Systemmatrix und dem Eingangsvektor

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \quad \text{und} \quad \mathbf{b}_c = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (7.75)$$

transformiert. Zwischen dem ursprünglichen System und dem transformierten System besteht nach (7.63) der Zusammenhang

$$\mathbf{A}_c = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} \quad \text{und} \quad \mathbf{b}_c = \mathbf{T}^{-1} \mathbf{b} \quad (7.76)$$

Um die nachfolgende Rechnung zu vereinfachen, wird eine Hilfsmatrix \mathbf{H} als Inverse der Transformationsmatrix definiert:

$$\mathbf{H} = \mathbf{T}^{-1} = \begin{bmatrix} \mathbf{h}_1^T \\ \cdots \\ \mathbf{h}_n^T \end{bmatrix} \quad (7.77)$$

Aus (7.76) wird damit

$$\mathbf{A}_c = \mathbf{H} \mathbf{A} \mathbf{H}^{-1} \quad \text{bzw.} \quad \mathbf{A}_c \mathbf{H} = \mathbf{H} \mathbf{A} \quad \text{und} \quad \mathbf{b}_c = \mathbf{H} \mathbf{b} \quad (7.78)$$

Wie schon bei der Herleitung zur Jordanschen Normalform soll auch hier ein komponentenweiser Vergleich dieser Matrizengleichungen auf eine Lösung für die Transformationsmatrix führen.

Mit den Zeilen \mathbf{h}_i^T der Matrix \mathbf{H} gilt

$$\mathbf{H} \mathbf{A} = \begin{bmatrix} \mathbf{h}_1^T \mathbf{A} \\ \cdots \\ \mathbf{h}_n^T \mathbf{A} \end{bmatrix} \quad (7.79)$$

und

$$\mathbf{A}_c \mathbf{H} = \begin{bmatrix} \mathbf{h}_2^T \\ \mathbf{h}_3^T \\ \dots \\ \mathbf{h}_n^T \\ -a_0 \mathbf{h}_1^T - a_1 \mathbf{h}_2^T - \dots - a_{n-1} \mathbf{h}_n^T \end{bmatrix} \quad (7.80)$$

Wegen der Forderung $\mathbf{A}_c \mathbf{H} = \mathbf{H} \mathbf{A}$ ergibt der komponentenweise Vergleich nach diesen beiden Gleichungen die Zusammenhänge

$$\begin{aligned} \mathbf{h}_2^T &= \mathbf{h}_1^T \mathbf{A} \\ \mathbf{h}_3^T &= \mathbf{h}_2^T \mathbf{A} = \mathbf{h}_1^T \mathbf{A}^2 \\ \mathbf{h}_4^T &= \mathbf{h}_3^T \mathbf{A} = \mathbf{h}_1^T \mathbf{A}^3 \\ \mathbf{h}_n^T &= \mathbf{h}_{n-1}^T \mathbf{A} = \mathbf{h}_1^T \mathbf{A}^{n-1} \\ \mathbf{h}_n^T \mathbf{A} &= -a_0 \mathbf{h}_1^T - a_1 \mathbf{h}_2^T - \dots - a_{n-1} \mathbf{h}_n^T \end{aligned} \quad (7.81)$$

Damit lässt sich die inverse Transformationsmatrix zunächst wie folgt angeben:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_1^T \mathbf{A} \\ \mathbf{h}_1^T \mathbf{A}^2 \\ \dots \\ \mathbf{h}_1^T \mathbf{A}^{n-1} \end{bmatrix} \quad (7.82)$$

Unbekannt ist jedoch noch \mathbf{h}_1 . Dieser Vektor lässt sich aber aus $\mathbf{b}_c = \mathbf{H} \mathbf{b}$ ermitteln (vgl. (7.78)). Schreibt man diese Vektorgleichung in Zeilendarstellung, also in transponierter Form, so ergibt sich

$$\begin{aligned} [0, \dots, 0, 1] &= [\mathbf{h}_1^T \mathbf{b}, \mathbf{h}_1^T \mathbf{A} \mathbf{b}, \mathbf{h}_1^T \mathbf{A}^2 \mathbf{b}, \dots, \mathbf{h}_1^T \mathbf{A}^{n-1} \mathbf{b}] \\ &= \mathbf{h}_1^T [\mathbf{b}, \mathbf{A} \mathbf{b}, \mathbf{A}^2 \mathbf{b}, \dots, \mathbf{A}^{n-1} \mathbf{b}] \\ &= \mathbf{h}_1^T \mathbf{Q}_s \end{aligned} \quad (7.83)$$

mit der sogenannten *Steuerbarkeitsmatrix*

$$\mathbf{Q}_s = [\mathbf{b}, \mathbf{A} \mathbf{b}, \mathbf{A}^2 \mathbf{b}, \dots, \mathbf{A}^{n-1} \mathbf{b}] \quad (7.84)$$

Wenn diese regulär und damit invertierbar ist, dann kann man \mathbf{h} aus

$$\mathbf{h}_1^T = [0, \dots, 0, 1] \mathbf{Q}_s^{-1} \quad (7.85)$$

berechnen, wobei auf der rechten Seite der Gleichung ausschließlich die Matrizen des gegebenen Systems auftauchen.

Neben der Forderung nach Invertierbarkeit von \mathbf{Q}_s , die sich direkt aus der obigen Herleitung ergibt, bleibt abschließend zu prüfen, ob die resultierende (inverse) Transformationsmatrix \mathbf{H} regulär ist, da sie (bzw. \mathbf{T}) sonst nicht Matrix einer Basistransformation sein könnte.

Die Regularität einer Matrix ist gegeben, wenn alle Zeilen dieser Matrix linear unabhängig sind, d.h. die einzige Lösung $\mathbf{c} = [c_1, \dots, c_n]$ der Gleichung

$$c_1 \mathbf{h}_1^T + \dots + c_n \mathbf{h}_n^T = \mathbf{0}^T \quad (7.86)$$

darf der Nullvektor $\mathbf{c} = \mathbf{0}^T$ sein. Denn sonst ließe sich jeweils ein Vektor \mathbf{h}_i^T durch eine Linearkombination der anderen beschreiben.

Aus (7.86) folgt unter Verwendung von (7.81)

$$\mathbf{0}^T = c_1 \mathbf{h}_1^T + c_2 \mathbf{h}_1^T \mathbf{A} + \dots + c_n \mathbf{h}_1^T \mathbf{A}^{n-1} \quad (7.87)$$

und nach einer Multiplikation mit \mathbf{b}

$$0 = c_1 \mathbf{h}_1^T \mathbf{b} + c_2 \mathbf{h}_1^T \mathbf{A} \mathbf{b} + \dots + c_n \mathbf{h}_1^T \mathbf{A}^{n-1} \mathbf{b} \quad (7.88)$$

Wegen (7.83) sind aber alle Terme $\mathbf{h}_1^T \mathbf{A}^i \mathbf{b}$ für $i < n-1$ gleich Null, nur für den letzten Term gilt $\mathbf{h}_1^T \mathbf{A}^{n-1} \mathbf{b} = 1$. Damit reduziert sich die Gleichung auf

$$0 = c_n \quad (7.89)$$

Mit diesem Ergebnis wird aus (7.87)

$$\mathbf{0}^T = c_1 \mathbf{h}_1^T + c_2 \mathbf{h}_1^T \mathbf{A} + \dots + c_{n-1} \mathbf{h}_1^T \mathbf{A}^{n-2} \quad (7.90)$$

Diese Gleichung wird nun mit $\mathbf{A} \mathbf{b}$ multipliziert:

$$0 = c_1 \mathbf{h}_1^T \mathbf{A} \mathbf{b} + c_2 \mathbf{h}_1^T \mathbf{A}^2 \mathbf{b} + \dots + c_{n-1} \mathbf{h}_1^T \mathbf{A}^{n-1} \mathbf{b} \quad (7.91)$$

Wieder macht man sich zu Nutzen, dass alle Terme $\mathbf{h}_1^T \mathbf{A}^i \mathbf{b}$ für $i < n-1$ gleich Null sind und nur für den letzten Term $\mathbf{h}_1^T \mathbf{A}^{n-1} \mathbf{b} = 1$ gilt. Es bleibt

$$0 = c_{n-1} \quad (7.92)$$

In analoger Weise lässt sich Schritt für Schritt zeigen, dass $c_i = 0$ für alle c_i gilt. Die einzige Lösung von (7.86) ist damit der Nullvektor, und so ist gezeigt, dass die \mathbf{h}_i^T linear unabhängig sowie \mathbf{H} und $\mathbf{T} = \mathbf{H}^{-1}$ regulär und damit Transformationsmatrizen sind.

Zusammenfassend lässt sich über die Transformation auf Regelungsnormalform sagen:

Satz 7.1 *Ein System $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u$ lässt sich auf Regelungsnormalform transformieren, wenn die Steuerbarkeitsmatrix $\mathbf{Q}_s = [\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}]$ regulär ist. In diesem Fall gilt für die Transformationsmatrix*

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{t}_1^T \\ \mathbf{t}_1^T \mathbf{A} \\ \vdots \\ \mathbf{t}_1^T \mathbf{A}^{n-1} \end{bmatrix} \quad (7.93)$$

mit $\mathbf{t}_1^T = [0, \dots, 0, 1] \mathbf{Q}_s^{-1}$ bzw. $\mathbf{t}_1^T \mathbf{Q}_s = [0, \dots, 0, 1]$.

Anmerkung: Hier sind die \mathbf{t}_i^T die Zeilen der inversen Transformationsmatrix, während die \mathbf{t}_i in der ursprünglichen Definition (7.57) die Spalten der Transformationsmatrix \mathbf{T} sind. Es gilt daher auch nicht, dass man lediglich die hier vorliegenden Zeilenvektoren transformieren muss, um die Spaltenvektoren von \mathbf{T} zu erhalten. Dazu wäre eine vollständige Inversion der gesamten Matrix erforderlich.

7.7 Allgemeine Lösung der linearen Zustandsgleichung

Ebenfalls als Grundlage für spätere Rechnungen wird die allgemeine Lösung der Zustandsgleichung benötigt, die direkt den Zusammenhang zwischen Ein- und Ausgangsgrößen des Systems angibt.

Ausgangspunkt ist die Zustandsgleichung, die unter Verwendung der in Anhang A.3 hergeleiteten Rechenregeln für die Matrizen-e-Funktion umgeformt wird:

$$\begin{aligned}
 \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\
 e^{-\mathbf{A}t}\dot{\mathbf{x}}(t) &= e^{-\mathbf{A}t}\mathbf{A}\mathbf{x}(t) + e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t) \\
 e^{-\mathbf{A}t}\dot{\mathbf{x}}(t) - e^{-\mathbf{A}t}\mathbf{A}\mathbf{x}(t) &= e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t) \\
 \frac{d}{dt}(e^{-\mathbf{A}t}\mathbf{x}(t)) &= e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t) \\
 \int_{t_0}^t \frac{d}{d\tau}(e^{-\mathbf{A}\tau}\mathbf{x}(\tau))d\tau &= \int_{t_0}^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)d\tau \\
 e^{-\mathbf{A}t}\mathbf{x}(t) - e^{-\mathbf{A}t_0}\mathbf{x}(t_0) &= \int_{t_0}^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)d\tau \\
 \mathbf{x}(t) - e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) &= \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \\
 \mathbf{x}(t) &= e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \quad (7.94)
 \end{aligned}$$

Dies ist eine Lösung für die Zustandsgleichung. An ihr lässt sich ablesen, dass ein bestimmter Zustand immer von einem Anfangszustand $\mathbf{x}(t_0)$ und vom anschließenden Verlauf der Anregung $\mathbf{u}(t)$ abhängig ist.

Ist keine äußere Anregung vorhanden ($\mathbf{u}(t) = \mathbf{0}$), so reduziert sich die Lösung auf

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) \quad (7.95)$$

Der Anfangszustand $\mathbf{x}(t_0)$ ist also mit der Matrix $e^{\mathbf{A}(t-t_0)}$ zu multiplizieren, um den aktuellen Zustandsvektor zu erhalten. Diese Matrix wird deshalb auch als *Transitionsmatrix* bezeichnet.

Nun seien $t_0 < t_1 < t_2$ drei Zeitpunkte. Offenbar gilt

$$\begin{aligned}
 \mathbf{x}(t_1) &= e^{\mathbf{A}(t_1-t_0)}\mathbf{x}(t_0) \\
 \mathbf{x}(t_2) &= e^{\mathbf{A}(t_2-t_1)}\mathbf{x}(t_1) = e^{\mathbf{A}(t_2-t_1)}e^{\mathbf{A}(t_1-t_0)}\mathbf{x}(t_0) \quad (7.96)
 \end{aligned}$$

Für jeweils 2 Zeitpunkte lässt sich also eine Transitionsmatrix angeben, die den Zustand zum ersten Zeitpunkt in den Zustand zum zweiten Zeitpunkt

überführt. Und für zwei hintereinander liegende Zeitintervalle muss man lediglich die Transitionsmatrizen multiplizieren, um die Transitionsmatrix für das gesamte Zeitintervall zu erhalten.

Wegen (A.24) lautet die Inverse einer Transitionsmatrix

$$\left(e^{\mathbf{A}(t_1-t_0)}\right)^{-1} = e^{\mathbf{A}(t_0-t_1)} \quad (7.97)$$

Diese beschreibt offenbar die Rückkehr vom Endzustand in den Anfangszustand. Außerdem folgt, dass jede Transitionsmatrix regulär sein muss.

Der Zusammenhang zwischen Zustands- und Ausgangsgrößen ergibt sich, wenn man die Lösung der Zustandsgleichung in die Ausgangsgleichung einsetzt:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \\ &= \mathbf{C}e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau + \mathbf{D}\mathbf{u}(t) \end{aligned} \quad (7.98)$$

Abschließend sollen die gewonnenen Erkenntnisse in den Frequenzbereich übertragen werden. Für die Laplace-Transformierte von $\dot{\mathbf{x}}$ gilt bei komponentenweiser Laplace-Transformation dieses Vektors

$$\mathcal{L}\{\dot{\mathbf{x}}(t)\} = s\mathbf{x}(s) - \mathbf{x}(0) = \begin{bmatrix} sx_1(s) - x_1(0) \\ sx_2(s) - x_2(0) \\ \dots \end{bmatrix} \quad (7.99)$$

und damit für die Laplace-Transformation der Zustandsgleichungen

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \rightarrow s\mathbf{x}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{x}(s) + \mathbf{B}\mathbf{u}(s) \quad (7.100)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \rightarrow \mathbf{y}(s) = \mathbf{C}\mathbf{x}(s) + \mathbf{D}\mathbf{u}(s) \quad (7.101)$$

Aus (7.100) folgt

$$\begin{aligned} (s\mathbf{I} - \mathbf{A})\mathbf{x}(s) &= \mathbf{B}\mathbf{u}(s) + \mathbf{x}(0) \\ \mathbf{x}(s) &= (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s) + (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \end{aligned} \quad (7.102)$$

Für die homogene Differentialgleichung (ohne äußere Anregung) ergibt sich

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \quad (7.103)$$

Vergleicht man diesen Ausdruck mit der Lösung der homogenen Zustandsgleichung im Zeitbereich (7.95), so ergibt sich für $t_0 = 0$ aus den jeweiligen Vorfaktoren von $\mathbf{x}(0)$ sofort der Zusammenhang

$$\mathcal{L}\{e^{\mathbf{A}t}\} = (s\mathbf{I} - \mathbf{A})^{-1} \quad (7.104)$$

Dieser deckt sich mit dem aus der Korrespondenztabelle im Anhang zu gewinnenden Zusammenhang für skalare Funktionen

$$\mathcal{L}\{e^{at}\} = \frac{1}{s-a} \quad (7.105)$$

Setzt man (7.102) in (7.101) ein, so erhält man einen geschlossenen Ausdruck für die Ausgangsgröße \mathbf{y} in Abhängigkeit von der Eingangsgröße \mathbf{u} :

$$\begin{aligned} \mathbf{y}(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) + \mathbf{D}\mathbf{u}(s) \\ &= [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}] \mathbf{u}(s) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \\ &= \mathbf{G}(s)\mathbf{u}(s) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \end{aligned} \quad (7.106)$$

mit der *Übertragungsmatrix* oder *Transfermatrix*

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (7.107)$$

bei der ein Koeffizient $G_{ik}(s)$ die Übertragungsfunktion von der Komponente $u_k(s)$ des Eingangsvektors zur Komponente $y_i(s)$ des Ausgangsvektors bildet.

Die Übertragungsmatrix ist zu unterscheiden von der Transitionsmatrix. Die Übertragungsmatrix kennzeichnet den Zusammenhang zwischen Eingangs- und Ausgangsvektor im Frequenzbereich, während die Transitionsmatrix den Zusammenhang zwischen einem Anfangszustand und einem späteren Systemzustand im Zeitbereich beschreibt.

7.8 Der Ljapunovsche Stabilitätsbegriff für lineare Systeme

Mit den jetzigen Kenntnissen ist klar, dass die bisher verwendeten Stabilitätsdefinitionen (vgl. Kap. 5.1) im Prinzip recht unvollständig waren, weil sie sich nur auf das Verhalten der Ausgangsgrößen des Systems bezogen. Deshalb soll an dieser Stelle die Stabilitätsdefinition von M.A. Ljapunov [36] für lineare Systeme vorgestellt werden:

Definition 7.2 *Ein lineares System ist genau dann asymptotisch stabil, wenn seine Zustandsgrößen ohne äußere Anregung aus jedem beliebigen Anfangszustand gegen Null streben:*

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{0} \quad \text{mit} \quad \mathbf{u}(t) = \mathbf{0} \quad (7.108)$$

Ein stabiles System kommt also von alleine aus jedem beliebigen Anfangszustand wieder zur Ruhe. Zu einem späteren Zeitpunkt wird diese Definition verallgemeinert und auch der Unterschied zwischen einfacher und asymptotischer Stabilität erläutert, der jedoch nur für nichtlineare Systeme relevant

ist. Spricht man bei linearen Systemen von Stabilität, so ist normalerweise die asymptotische Stabilität gemeint, weshalb hier auf eine Unterscheidung verzichtet werden kann.

Im Gegensatz zu den früher behandelten Stabilitätsdefinitionen 5.1 (endliche Sprungantwort) und 5.2 (begrenzter Ausgang bei begrenztem Eingang) wird bei dieser Definition nicht die Reaktion der Ausgangs- auf eine Eingangsgröße betrachtet, sondern das auf einen Anfangszustand folgende, interne Verhalten des Systems.

Die Frage ist nun, ob sich die Stabilität schon aus den Matrizen der Zustandsdarstellung ablesen lässt. Da der Ljapunovsche Stabilitätsbegriff von einem System ohne äußere Anregung ausgeht, vereinfacht sich die zu betrachtende Zustandsgleichung mit $\mathbf{u} = \mathbf{0}$ zu einer homogenen Vektor-Differentialgleichung:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad (7.109)$$

Für die Stabilitätsdefinition ist nun zu untersuchen, unter welchen Bedingungen die Lösung dieser Gleichung für beliebige Anfangswerte gegen Null strebt. Dabei gestaltet sich die weitere Betrachtung am einfachsten, wenn man die Gleichung einer Laplace-Transformation unterzieht. Gemäß (7.103) ist damit der Ausgangspunkt der Überlegungen

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \quad (7.110)$$

Zur Berechnung der inversen Matrix bietet sich die Cramersche Regel an. Sie lautet für eine reguläre quadratische Matrix \mathbf{M} :

$$\mathbf{M}^{-1} = \frac{1}{\det(\mathbf{M})} \begin{bmatrix} M_{11} & M_{12} & \cdots \\ M_{21} & M_{22} & \cdots \\ \dots & \dots & \dots \end{bmatrix}^T \quad (7.111)$$

mit den Adjunkten

$$M_{ik} = (-1)^{i+k} \det \begin{bmatrix} m_{11} & \cdots & m_{1,k-1} & m_{1,k+1} & \cdots & m_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ m_{i-1,1} & \dots & \dots & \dots & \dots & \dots \\ m_{i+1,1} & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ m_{n,1} & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (7.112)$$

Dabei sind die m_{ij} die Elemente von \mathbf{M} .

Im vorliegenden Fall gilt $\mathbf{M} = s\mathbf{I} - \mathbf{A}$, d.h. s taucht in jedem Diagonalelement von \mathbf{M} auf. Damit ist $M_{ik} = p_{ik}(s)$ ein von s abhängiges Polynom vom Grad $n - 1$. Für die Inverse ergibt sich aus (7.111)

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{\det(s\mathbf{I} - \mathbf{A})} \begin{bmatrix} p_{11}(s) & p_{12}(s) & \cdots \\ p_{21}(s) & p_{22}(s) & \cdots \\ \dots & \dots & \dots \end{bmatrix}^T = \frac{\mathbf{P}(s)}{\det(s\mathbf{I} - \mathbf{A})} \quad (7.113)$$

$\mathbf{P}(s)$ ist eine Polynommatrix, d.h. ihre einzelnen Elemente sind von s abhängige Polynome.

Aus (7.110) wird damit

$$\mathbf{x}(s) = \frac{\mathbf{P}(s)}{\det(s\mathbf{I} - \mathbf{A})} \mathbf{x}_0 \quad (7.114)$$

Diese Schreibweise ist natürlich nur möglich, wenn die Inverse überhaupt existiert bzw. die Determinante $\det(s\mathbf{I} - \mathbf{A})$ von Null verschieden ist. Für diese Determinante wiederum gilt mit den Eigenwerten λ_i von \mathbf{A}

$$\det(s\mathbf{I} - \mathbf{A}) = \prod_{i=1}^n (s - \lambda_i) \quad (7.115)$$

Zusammenfassend hat damit jedes Element von $(s\mathbf{I} - \mathbf{A})^{-1}$ die Form

$$\frac{p(s)}{\prod_{i=1}^n (s - \lambda_i)} \quad (7.116)$$

wobei $p(s)$ ein von s abhängiges Polynom vom Grad $n - 1$ ist. Damit lässt sich für jedes einzelne Element von $(s\mathbf{I} - \mathbf{A})^{-1}$ eine Partialbruchzerlegung durchführen, und es ergibt sich

$$\frac{p(s)}{\prod_{i=1}^n (s - \lambda_i)} = \sum_{\mu=1}^l \sum_{\nu=1}^{r_\mu} \frac{k_{\mu\nu}}{(s - \lambda_\mu)^\nu} \quad (7.117)$$

Dabei ist l die Anzahl verschiedener Eigenwerte des Systems und r_μ die Vielfachheit des Eigenwertes λ_μ . $k_{\mu\nu}$ ist ein konstanter Wert.

In Matrizenschreibweise wird daraus

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{x}_0 = \frac{\mathbf{P}(s)}{\prod_{i=1}^n (s - \lambda_i)} \mathbf{x}_0 = \sum_{\mu=1}^l \sum_{\nu=1}^{r_\mu} \frac{\mathbf{K}_{\mu\nu}}{(s - \lambda_\mu)^\nu} \mathbf{x}_0 \quad (7.118)$$

$\mathbf{K}_{\mu\nu}$ ist eine Matrix mit konstanten Koeffizienten.

Für die Rücktransformation in den Zeitbereich ist nur der Term $\frac{1}{(s - \lambda_\mu)^\nu}$ zu betrachten, da alle anderen Terme nicht von s abhängig sind. Mit Hilfe der Korrespondenztabelle im Anhang erhält man

$$\frac{1}{(s - \lambda_\mu)^\nu} \leftrightarrow e^{\lambda_\mu t} \frac{t^{\nu-1}}{(\nu-1)!} \quad (7.119)$$

Damit gilt für die Gesamtlösung

$$\mathbf{x}(t) = \sum_{\mu=1}^l e^{\lambda_\mu t} \sum_{\nu=1}^{r_\mu} \frac{t^{\nu-1}}{(\nu-1)!} \mathbf{K}_{\mu\nu} \mathbf{x}_0 \quad (7.120)$$

Jede Komponente von $\mathbf{x}(t)$ enthält damit Produkte aus Exponentialfunktionen und Polynomen in t . In einem solchen Produkt ist die Exponentialfunktion immer der ausschlaggebende Term. Ist der Realteil ihres Exponenten positiv, so wächst das Produkt unabhängig vom Polynom über alle Maßen, während die Exponentialfunktion bei einem negativen Realteil ihres Exponenten so schnell gegen Null konvergiert, dass das Polynom ebenfalls keine Rolle mehr spielt. Der Vektor $\mathbf{x}(t)$ strebt damit genau dann gegen Null, wenn der Realteil aller Koeffizienten λ_μ negativ ist.

Satz 7.3 *Ein lineares, zeitinvariantes System ist genau dann asymptotisch stabil im Sinne der Definition von Ljapunov, wenn alle Eigenwerte der Systemmatrix \mathbf{A} einen negativen Realteil aufweisen.*

Dabei entscheiden die Eigenwerte natürlich nicht nur über die Stabilität des Systems, sondern auch über die Form der Einschwingvorgänge, wie man an Gleichung (7.120) unschwer erkennen kann. Entsprechend dem Absolutwert der Realteile wird das System schneller oder langsamer gegen den Nullzustand konvergieren, und im Falle eines konjugiert komplexen Eigenwertpaares kommt es wie im skalaren Fall bei einem konjugiert komplexen Polpaar zu Schwingungen. Insbesondere gilt, dass das System umso schneller gegen Null konvergiert, je größer der Absolutwert der Eigenwerte ist, sofern diese denn einen negativen Realteil haben.

Von der Wahl der Zustandsgrößen ist die Stabilität natürlich unabhängig, da die Eigenwerte einer Matrix \mathbf{A} durch eine Basistransformation $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ nicht verändert werden (vgl. (7.66)).

Es stellt sich noch die Frage, inwieweit der neue Stabilitätsbegriff mit den beiden alten Definitionen in Zusammenhang gebracht werden kann. Hierzu greifen wir auf Gleichung (7.106)

$$\mathbf{y}(s) = \mathbf{G}(s)\mathbf{u}(s) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \quad (7.121)$$

zurück mit der Transfermatrix

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (7.122)$$

$\mathbf{G}(s)$ beschreibt das Übertragungsverhalten vom Ein- zum Ausgang, wobei ein Koeffizient $G_{ik}(s)$ die Übertragungsfunktion von der Komponente $u_k(s)$ des Eingangsvektors zur Komponente $y_i(s)$ des Ausgangsvektors bildet. Der von \mathbf{x}_0 abhängige Term stellt den Einfluss eines Anfangszustandes auf die Ausgangsgröße dar.

Mit (7.118) erhält man für die Transfermatrix:

$$\mathbf{G}(s) = \mathbf{C} \frac{\mathbf{P}(s)}{\prod_{i=1}^n (s - \lambda_i)} \mathbf{B} + \mathbf{D} \quad (7.123)$$

Da der hier auftretende Nenner $\prod_{i=1}^n (s - \lambda_i)$ durch die Multiplikation mit den konstanten Matrizen \mathbf{B} und \mathbf{C} nicht verändert wird, ist er der (gemeinsame) Nenner aller Übertragungsfunktionen $G_{ik}(s)$ in $\mathbf{G}(s)$. Die Nullstellen dieses Nenners, also die λ_i , bilden daher die Polstellen der skalaren Übertragungsfunktionen $G_{ik}(s)$. Ist das System stabil im Sinne der Definition von Ljapunov, d.h. weisen alle Eigenwerte λ_i der Systemmatrix einen negativen Realteil auf, so gilt dies auch für alle Pole der Übertragungsfunktionen $G_{ik}(s)$, und die Übertragungsfunktionen sind nach Satz 5.3 stabil. Wenn aber alle Elemente $G_{ik}(s)$ von $\mathbf{G}(s)$ stabile skalare Übertragungsfunktionen sind, so ist das Gesamtsystem ebenfalls stabil im Sinne der Definitionen 5.1 und 5.2. Aus der Ljapunov-Stabilität folgt also auch die Übertragungsstabilität eines Systems.

Andersherum gilt diese Folgerung aber nicht, denn da sich die Pol- und Nullstellen der Übertragungsfunktionen gegeneinander kürzen lassen, müssen nicht alle Eigenwerte der Systemmatrix auch tatsächlich als Polstellen der Übertragungsfunktionen in Erscheinung treten. Wenn also alle Polstellen einen negativen Realteil aufweisen, so muss dies nicht unbedingt auch für alle Eigenwerte der Systemmatrix gelten. Der Ljapunovsche Stabilitätsbegriff ist damit umfassender als die bisher behandelten Stabilitätsdefinitionen, was nicht verwunderlich ist, denn wenn sämtlichen internen Systemgrößen einen stabilen Verlauf aufweisen, kann es keine Ausgangsgröße mit einem instabilen Verlauf geben. Andererseits kann ein System nach außen hin durchaus als stabil erscheinen, während interne Vorgänge instabil werden und nur deshalb nicht bemerkt werden, weil sie sich gegenseitig kompensieren oder die entsprechenden Ausgangsgrößen nicht gemessen werden.

7.9 Dominanzmaß von Eigenwerten

Nachdem im letzten Kapitel die entscheidende Bedeutung der Eigenwerte für das dynamische Verhalten des Systems herausgearbeitet wurde, soll in diesem Abschnitt ein praktischer Aspekt, der die Eigenwerte betrifft, behandelt werden.

Der Grundgedanke ist dabei wie folgt: Oft steht man bei der Auslegung einer Regelung vor dem Problem, dass das zu regelnde System bzw. dessen Modell außerordentlich komplex ist, was dazu führt, dass auch der Reglerentwurf und der entstehende Regler sehr komplex, für die praktische Anwendung sogar zu komplex werden. In Kapitel 3.4 wurde bereits für Übertragungsfunktionen gezeigt, wie sich diese vereinfachen lassen, bevor man mit dem Reglerentwurf beginnt. Ähnliches ist auch bei Systemmatrizen möglich. Das hier vorgestellte Verfahren ist dabei nur eine von verschiedenen Möglichkeiten, die darüber hinaus auch nicht immer zum Erfolg führen muss. Ebenso wie das in Kapitel 3.4 vorgestellte Verfahren ist es mit Augenmaß und Überlegung

einzusetzen, damit man bei der Vereinfachung des Streckenmodells nicht aus Versehen wesentliche Teile weglässt.

Das Verfahren versucht jedem Eigenwert der Systemmatrix ein Dominanzmaß zuzuweisen, das umso größer ist, je wichtiger der Eigenwert für die Beschreibung des Systemverhaltens ist. Ausgangspunkt der Überlegungen ist die Zustandsgleichung (7.15) der Strecke

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x}\end{aligned}\tag{7.124}$$

wobei die Durchgangsmatrix von vornherein zu Null angenommen wird.

Um die Darstellung hier nicht zu kompliziert werden zu lassen, wird von ausschließlich einfachen Eigenwerten ausgegangen. Gemäß Kapitel 7.6.2 ergibt sich dann bei der Transformation auf Jordansche Normalform für die Systemmatrix \mathbf{A}' eine Diagonalmatrix

$$\mathbf{A}' = \text{diag}(\lambda_1, \dots, \lambda_n)\tag{7.125}$$

Die transformierten Systemmatrizen werden zur Bildung der Übertragungsmatrix gemäß (7.107) herangezogen:

$$\mathbf{G}(s) = \mathbf{C}'(s\mathbf{I} - \mathbf{A}')^{-1}\mathbf{B}'\tag{7.126}$$

Da \mathbf{A}' eine Diagonalmatrix ist, gilt

$$s\mathbf{I} - \mathbf{A}' = \begin{bmatrix} s - \lambda_1 & & & \\ & s - \lambda_2 & & \\ & & \dots & \\ & & & s - \lambda_n \end{bmatrix}\tag{7.127}$$

und

$$(s\mathbf{I} - \mathbf{A}')^{-1} = \begin{bmatrix} \frac{1}{s - \lambda_1} & & & \\ & \frac{1}{s - \lambda_2} & & \\ & & \dots & \\ & & & \frac{1}{s - \lambda_n} \end{bmatrix}\tag{7.128}$$

Für das Element (i, j) der Übertragungsmatrix folgt daraus

$$\begin{aligned}G_{i,j}(s) &= [c'_{i1}, \dots, c'_{in}] \begin{bmatrix} \frac{1}{s - \lambda_1} & & & \\ & \frac{1}{s - \lambda_2} & & \\ & & \dots & \\ & & & \frac{1}{s - \lambda_n} \end{bmatrix} \begin{bmatrix} b'_{1j} \\ \vdots \\ \vdots \\ b'_{nj} \end{bmatrix} \\ &= \sum_{k=1}^n \frac{c'_{ik} b'_{kj}}{s - \lambda_k}\end{aligned}\tag{7.129}$$

Dies ist die Übertragungsfunktion von u_j nach y_i , also von der j -ten Komponente des Eingangsvektors zur i -ten Komponente des Ausgangsvektors.

Die Impulsantwort dieser Übertragungsfunktion ergibt sich durch Rücktransformation vom Frequenz- in den Zeitbereich (vgl. Korrespondenztabelle im Anhang A.2):

$$g_{i,j}(t) = \sum_{k=1}^n e^{\lambda_k t} c'_{ik} b'_{kj} \quad (7.130)$$

Hiermit lässt sich die Sprungantwort der Ausgangskomponente y_i als Reaktion auf einen Sprung der Eingangskomponente u_j berechnen. Im Frequenzbereich lautet der Zusammenhang

$$y_i(s) = G_{i,j}(s) \frac{1}{s} \quad (7.131)$$

Die Multiplikation mit dem Faktor $\frac{1}{s}$ kann auch als Integration interpretiert werden, so dass für den Zusammenhang im Zeitbereich folgt

$$\begin{aligned} y_i(t) &= \int_0^t g_{i,j}(\tau) d\tau \\ &= \int_0^t \sum_{k=1}^n e^{\lambda_k \tau} c'_{ik} b'_{kj} d\tau \\ &= \sum_{k=1}^n \frac{c'_{ik} b'_{kj}}{\lambda_k} e^{\lambda_k \tau} \Big|_0^t \\ &= \sum_{k=1}^n \frac{c'_{ik} b'_{kj}}{\lambda_k} (e^{\lambda_k t} - 1) \end{aligned} \quad (7.132)$$

Diese Formel zeigt, dass es naheliegend ist, als Maß für den Einfluss des Eigenwertes λ_k auf den Signalpfad (i, j) die Zahl

$$\left| \frac{c'_{ik} b'_{kj}}{\lambda_k} \right| \quad (7.133)$$

zu definieren. Der Maximalwert über alle Signalpfade

$$\max_{i,j} \left| \frac{c'_{ik} b'_{kj}}{\lambda_k} \right| \quad (7.134)$$

könnte damit ein sinnvolles Maß für die Wichtigkeit, d.h. für das Dominanzmaß dieses Eigenwertes sein.

Ein wesentlicher Punkt fehlt jedoch noch. Denn wenn die Eingangsgröße u_j nur sehr kleine Werte annimmt, dann ist der gesamte Signalpfad eher unwichtig. In den obigen Ausdruck muss daher die Größenordnung der Signale mit eingehen.

Das Problem lässt sich am einfachsten lösen, indem man sämtliche Größen des Systems vor Berechnung des Dominanzmaßes normiert, so dass im normierten System alle Ein- und Ausgangsgrößen dieselbe Größenordnung aufweisen. Für ein Eingrößensystem $y = Gu$ bedeutet dieser Schritt, dass man von den Größen u, y auf die normierten Größen $\tilde{u} = \frac{u}{u_0}$ bzw. $\tilde{y} = \frac{y}{y_0}$ übergeht. Es ergibt sich

$$\begin{aligned} y &= Gu \\ \tilde{y}y_0 &= G\tilde{u}u_0 \\ \tilde{y} &= \frac{u_0}{y_0}G\tilde{u} \end{aligned} \quad (7.135)$$

Entsprechend ist jede Teil-Übertragungsfunktion $G_{i,j}(s)$ vor Beginn der Dominanzmaßberechnung mit dem Faktor $\frac{u_{j0}}{y_{i0}}$ zu multiplizieren.

Für das Dominanzmaß des Eigenwertes λ_k ergibt sich damit der endgültige Ausdruck

$$D_k = \max_{i,j} \left| \frac{u_{j0}}{y_{i0}} \frac{c'_{ik} b'_{kj}}{\lambda_k} \right| \quad (7.136)$$

Dieser Ausdruck ist plausibel. Zunächst wird der Einfluss des Eigenwertes auf den Signalpfad berechnet. Das Ergebnis wird dann mit dem Normierungsfaktor, d.h. der Größenordnung der Eingangsgröße des Signalpfades gewichtet. Sofern die Eingangsgröße klein ist, wird sich ein kleines Resultat ergeben, denn dann ist auch die Wichtigkeit des Signalpfades entsprechend gering. Genau andersherum wirkt sich die Größenordnung der Ausgangsgröße aus. Die Auswirkung des Eigenwertes auf eine eher kleine Ausgangsgröße muss höher gewichtet werden als die zahlenmäßig gleiche Auswirkung auf eine große Ausgangsgröße. Daher ist auch die Division durch y_{i0} absolut plausibel.

Am Ende erfolgt hier die Maximumbildung über alle Teil-Einflüsse auf die verschiedenen Signalpfade, um das endgültige Dominanzmaß zu ermitteln. Hier wäre sicherlich auch z.B. eine Summenbildung denkbar, aber die Maximumbildung hat sich in der Praxis bewährt.

7.10 Steuerbarkeit und Beobachtbarkeit

Ziel einer Regelung ist es, eine Strecke so zu beeinflussen, dass bestimmte (Ausgangs-)größen der Strecke einen gewünschten Verlauf aufweisen. Nach den Ausführungen des vorangegangenen Kapitels stellt sich jedoch die Frage, ob es nicht sinnvoller wäre, das Augenmerk auf die Beeinflussung der Zustandsgrößen der Strecke zu legen. Denn wenn man die Zustandsgrößen gezielt beeinflussen kann, dann stellt sich der gewünschte Verlauf der Ausgangsgrößen von allein ein, da diese ja lediglich Linearkombinationen der Zustandsgrößen sind, wie man leicht anhand der Ausgangsgleichung der Zustandsdarstellung nachvollziehen kann. Fixiert man sich jedoch auf die Ausgangsgrößen, so kann es, wie das vorangegangene Kapitel gezeigt hat, interne

Vorgänge geben, die man nicht regelt, aber die aufgrund ihrer Instabilität an anderer Stelle zu Problemen führen.

Die Beeinflussung der Zustandsgrößen führt direkt auf das Konzept der sogenannten *Zustandsregelung*. Durch die Regelung der Zustandsgrößen unterliegen sämtliche internen Vorgänge einer ständigen Kontrolle, was natürlich eine wesentlich bessere Beherrschung des Systems ermöglicht, als dies bei einer klassischen Regelung der Ausgangsgrößen der Fall wäre. Zwei Eigenschaften muss ein System jedoch aufweisen, um eine Zustandsregelung zu ermöglichen, und zwar die *Steuerbarkeit* und die *Beobachtbarkeit*. Beide Begriffe wurden von Rudolf Kalman 1960 eingeführt [24] und sollen hier erläutert werden, bevor dann im nachfolgenden Kapitel das Konzept der Zustandsregelung vorgestellt werden soll.

Bei der Steuerbarkeit geht es um die Frage, ob es die Systemstruktur überhaupt zulässt, die Zustandsgrößen mit Hilfe der vorhandenen Stellgrößen in der gewünschten Weise zu beeinflussen. Ist dies der Fall, so bezeichnet man das System als steuerbar.

Eine geeignete Beeinflussung der Zustandsgrößen ist aber wiederum nur möglich, wenn ihr Verlauf auch bekannt ist. Da normalerweise nur die Ausgangsgrößen des Systems als Messgrößen zur Verfügung stehen, muss gewährleistet sein, dass jede Zustandsgröße in einer bestimmten Art und Weise auf die Ausgangsgrößen einwirkt, damit man aus diesen den Verlauf der Zustandsgrößen ermitteln kann. Ein solches System bezeichnet man dann als beobachtbar.

Ganz offensichtlich sollten sowohl die Steuerbarkeit als auch die Beobachtbarkeit vor Beginn des eigentlichen Reglerentwurfs untersucht und sichergestellt werden. Dabei hängen beide Eigenschaften ausschließlich von der Struktur der Strecke ab und nicht von der Art der eingesetzten Regelung. Eine nicht steuer- oder beobachtbare Strecke kann daher auch nur durch Änderung ihrer Struktur und nicht durch Wahl eines anderen Regelalgorithmus in eine steuer- bzw. beobachtbare Strecke umgewandelt werden.

Bei nicht vorhandener Steuerbarkeit sind Art oder Anzahl der Stellgrößen zu ändern, bei nicht vorhandener Beobachtbarkeit Art oder Anzahl der gemessenen Ausgangsgrößen. Sowohl Steuerbarkeit als auch Beobachtbarkeit sind kein spezielles Problem der Zustandsregelung, sondern grundlegende Systemeigenschaften, deren systematische Behandlung durch die Zustandsdarstellung überhaupt erst ermöglicht wurde.

7.10.1 Steuerbarkeit

Zunächst soll auf die Steuerbarkeit näher eingegangen werden. Dabei lässt sich ein gewisses Verständnis für diese Eigenschaft am besten anhand von zwei Beispielen vermitteln. Die Zustandsdarstellung des ersten Beispiels lautet:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ b_2 \end{bmatrix} u \quad (7.137)$$

Da die Zustandsgröße x_1 offenbar nur durch sich selbst angeregt wird, kann auf sie durch die Stellgröße kein Einfluss ausgeübt werden, und das System ist nicht vollständig steuerbar.

Als nicht steuerbar wird ein System aber auch dann bezeichnet, wenn seine Zustandsgrößen wie im folgenden Fall nicht unabhängig voneinander beeinflusst werden können.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} u \quad (7.138)$$

Hier wirkt nicht nur die Eingangsgröße u auf beide Zustandsgrößen gleichermaßen ein, sondern auch die internen Rückkopplungen sind für beide Größen gleich. Die Folge ist, dass dieses System beispielsweise nicht aus dem Anfangszustand $[x_1, x_2] = [1, 2]$ in den Nullzustand $[0, 0]$ überführt werden kann. Eine anschauliche Definition der Steuerbarkeit lautet damit:

Definition 7.4 Ein System heie (zustands-)steuerbar, wenn es durch geeignete Wahl der Eingangsgrößen nach endlicher Zeit aus jedem beliebigen Anfangszustand in den Endzustand $\mathbf{0}$ überführt werden kann.

Der Endzustand $\mathbf{0}$ bedeutet dabei keine besondere Einschränkung, da durch eine Koordinatenverschiebung jeder beliebige Punkt zum Nullpunkt gemacht werden kann.

Die Frage ist nun, ob man bereits anhand der Matrizen der Zustandsdarstellung ablesen kann, ob das System steuerbar ist oder nicht. Wie die folgende Herleitung zeigt, ist dies der Fall.

Steuerbarkeitskriterium nach Kalman. Ausgangspunkt ist die allgemeine Lösung der Zustandsgleichung (7.94)

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{B} \mathbf{u}(\tau) d\tau \quad (7.139)$$

Nach einer endlichen Zeit t_1 soll der Nullzustand erreicht sein:

$$\mathbf{0} = \mathbf{x}(t_1) = e^{\mathbf{A}t_1} \mathbf{x}_0 + \int_0^{t_1} e^{\mathbf{A}(t_1-\tau)} \mathbf{B} \mathbf{u}(\tau) d\tau \quad (7.140)$$

Dabei ist \mathbf{x}_0 ein beliebiger Anfangszustand.

Aus (7.140) folgt

$$-\mathbf{x}_0 = \int_0^{t_1} e^{-\mathbf{A}\tau} \mathbf{B} \mathbf{u}(\tau) d\tau \quad (7.141)$$

Das System ist steuerbar, wenn die Matrizen \mathbf{A} und \mathbf{B} so beschaffen sind, daß für jedes beliebige \mathbf{x}_0 ein Stellgrößenverlauf $\mathbf{u}(\tau)$ existiert, mit dem diese Gleichung erfüllt ist.

Einsetzen der Definition der Matrizen-e-Funktion (vgl. Anhang A.3) führt auf

$$\begin{aligned}
 -\mathbf{x}_0 &= \int_0^{t_1} \sum_{k=0}^{\infty} \mathbf{A}^k \frac{(-\tau)^k}{k!} \mathbf{B} \mathbf{u}(\tau) d\tau \\
 &= \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{B} \int_0^{t_1} \frac{(-\tau)^k}{k!} \mathbf{u}(\tau) d\tau \\
 &= \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{B} \mathbf{c}_k
 \end{aligned} \tag{7.142}$$

mit

$$\mathbf{c}_k = \int_0^{t_1} \frac{(-\tau)^k}{k!} \mathbf{u}(\tau) d\tau \tag{7.143}$$

Die Koeffizienten dieses Vektors lassen sich durch den Verlauf des Stellgrößenvektors beliebig festlegen. Damit kann man die rechte Seite von (7.142) als Linearkombination der Spalten der Matrizen $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}, \dots)$ auffassen, wobei die Vorfaktoren dieser Spaltenvektoren die Koeffizienten der Vektoren \mathbf{c}_k sind. Der Verlauf der Stellgrößen muss so gestaltet werden, dass mit den resultierenden Koeffizienten der \mathbf{c}_k die Linearkombination der Spaltenvektoren auf der rechten Seite von Gleichung (7.142) gerade gleich $-\mathbf{x}_0$ ist.

Damit eine solche Lösung für beliebige \mathbf{x}_0 existiert, muss aber sichergestellt sein, dass sich unter den Spalten der Matrizen $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}, \dots)$ n linear unabhängige Vektoren befinden, die den gesamten n -dimensionalen Raum aufspannen. Das System kann daher nur dann steuerbar sein, wenn die Matrix $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}, \dots)$ n linear unabhängige Spalten aufweist, also den Rang n hat.

Da diese Matrix unendlich viele Spalten aufweist, kann die Bedingung nicht direkt geprüft werden. Abhilfe bietet hier jedoch das Theorem von Caley-Hamilton (vgl. A.4), aus dem folgt, dass sich alle höheren Potenzen von \mathbf{A} (mit $k \geq n$) als Linearkombinationen der ersten $(n-1)$ Potenzen von \mathbf{A} darstellen lassen.

Daraus folgt aber auch, dass $\mathbf{A}^n, \mathbf{A}^{n+1}, \mathbf{A}^{n+2}$, usw. keine Spalten haben können, die von den Spalten der ersten $(n-1)$ Potenzen von \mathbf{A} linear unabhängig sind. Und ebenso können $\mathbf{A}^n\mathbf{B}, \mathbf{A}^{n+1}\mathbf{B}, \mathbf{A}^{n+2}\mathbf{B}$, usw. keine Spalten haben, die von den Spalten von $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B})$ linear unabhängig sind.

Für den Nachweis der Steuerbarkeit reduziert sich die Suche nach n linear unabhängigen Spalten damit auf die Matrizen $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B})$:

Satz 7.5 Kalman-Kriterium: Ein System ist genau dann vollständig steuerbar, wenn die Steuerbarkeitsmatrix

$$\mathbf{Q}_s = [\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}] \quad (7.144)$$

n linear unabhängige Spaltenvektoren enthält, d.h. wenn sie den Rang n hat.

Dabei wurde bisher nur gezeigt, dass aus dem Höchstrang von \mathbf{Q}_s die Steuerbarkeit folgt, während der Satz auch die umgekehrte Behauptung enthält. Auf deren Beweis soll an dieser Stelle aber verzichtet werden.

Für ein SISO-System reduziert sich die Matrix \mathbf{Q}_s auf die schon in Gleichung (7.84) eingeführte Form

$$\mathbf{Q}_s = [\mathbf{b}, \mathbf{Ab}, \mathbf{A}^2\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}] \quad (7.145)$$

Bei einem SISO-System hat \mathbf{Q}_s demnach die Dimension $n \times n$, was bedeutet, dass zur Erfüllung des Steuerbarkeitskriteriums sämtliche Spalten von \mathbf{Q}_s linear unabhängig sein müssen. Bei einem MIMO-System mit p Eingangsgrößen ist $\dim(\mathbf{Q}_s) = n \times (np)$, so dass zur Erfüllung des Steuerbarkeitskriteriums also nur n von np Spalten von \mathbf{Q}_s linear unabhängig sein müssen. Durch eine größere Anzahl an Stellgrößen steigen die Möglichkeiten, das System und die Zustandsgrößen zu beeinflussen, und damit auch die Wahrscheinlichkeit, dass das System steuerbar ist.

Der hier vorgestellte Begriff der Steuerbarkeit wird auch als *Zustandssteuerbarkeit* bezeichnet, um ihn von der sogenannten *Ausgangssteuerbarkeit* zu unterscheiden, die sich auf die Beeinflussbarkeit der Ausgangsgrößen bezieht und in Abschnitt 7.10.3 kurz behandelt wird.

Steuerbarkeitskriterium nach Hautus. Aus dem Kalmanschen Steuerbarkeitskriterium lässt sich direkt ein weiteres Kriterium für die Steuerbarkeit ableiten, das Steuerbarkeitskriterium von Hautus.

Laut Kalman-Kriterium (Satz 7.5) muss die Steuerbarkeitsmatrix \mathbf{Q}_s den Rang n aufweisen, damit das System steuerbar ist. Dies bedeutet, dass sie sowohl n linear unabhängige Spalten als auch n linear unabhängige Zeilen aufweisen muss. Für ein nicht-steuerbares System lässt sich demnach eine Linearkombination der Zeilen von \mathbf{Q}_s angeben, deren Summe Null ergibt:

$$\mathbf{q}^T \mathbf{Q}_s = \mathbf{0}^T \quad \text{mit} \quad \mathbf{q}^T = (q_1, \dots, q_n) \neq \mathbf{0}^T \quad (7.146)$$

Einsetzen der Definition von \mathbf{Q}_s liefert

$$\mathbf{q}^T [\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}] = \mathbf{0}^T \quad (7.147)$$

was im einzelnen bedeutet, dass

$$\mathbf{q}^T \mathbf{B} = \mathbf{0}^T, \quad \mathbf{q}^T \mathbf{AB} = \mathbf{0}^T, \dots, \quad \mathbf{q}^T \mathbf{A}^{n-1} \mathbf{B} = \mathbf{0}^T \quad (7.148)$$

gelten muss.

Zum Verständnis des eben durchgeführten Schrittes ist es sicher sinnvoll, kurz auf die Dimension der Gleichungen einzugehen. Offensichtlich hat \mathbf{q} die

Dimension n . \mathbf{B} hat die Dimension $n \times p$ (bei p Eingangsgrößen des Systems). Daher muss der Nullvektor in den Gleichungen (7.148) die Dimension p haben, und der Nullvektor in Gleichung (7.147) als Hintereinanderreihung von n Gleichungen die Dimension np .

Die folgende Argumentation soll für Systeme mit nur einer Eingangsgröße durchgeführt werden, da dieser Grenzfall der kritische Fall ist und die Erweiterung auf Systeme mit mehreren Eingangsgrößen unmittelbar auf der Hand liegt.

Da nach dem Caley-Hamilton-Theorem (Anhang A.4) die Matrizen \mathbf{A} bis \mathbf{A}^{n-1} linear unabhängig sind, müssen auch die Vektoren $\mathbf{A}\mathbf{b}$ bis $\mathbf{A}^{n-1}\mathbf{b}$ linear unabhängig sein. Daraus folgt, dass das auf eine Eingangsgröße reduzierte Gleichungssystem (7.148)

$$\mathbf{q}^T \mathbf{b} = 0, \mathbf{q}^T \mathbf{A} \mathbf{b} = 0, \dots, \mathbf{q}^T \mathbf{A}^{n-1} \mathbf{b} = 0 \quad (7.149)$$

aus n voneinander linear unabhängigen Gleichungen besteht. Die einzige Lösung für ein solches Gleichungssystem ist aber der Nullvektor $\mathbf{q} = \mathbf{0}$. Eine vom Nullvektor verschiedene, nicht-triviale Lösung ($\mathbf{q}^T \neq \mathbf{0}^T$) ist nur dann möglich, wenn der Einfluss der Matrizen \mathbf{A} bis \mathbf{A}^{n-1} im Gleichungssystem verschwindet. Und das gleiche Resultat erhält man offenbar auch für Systeme mit mehreren Eingangsgrößen. Damit kann im weiteren Verlauf wieder der Mehrgrößenfall betrachtet werden.

Der Einfluss der Matrizen \mathbf{A} bis \mathbf{A}^{n-1} im Gleichungssystem verschwindet genau dann, wenn \mathbf{q}^T ein Linkseigenvektor von \mathbf{A} ist. Denn dann gilt mit einem Eigenwert λ von \mathbf{A}

$$\begin{aligned} \mathbf{q}^T \mathbf{A} &= \lambda \mathbf{q}^T \\ \mathbf{q}^T \mathbf{A}^2 &= \lambda \mathbf{q}^T \mathbf{A} = \lambda^2 \mathbf{q}^T \\ \mathbf{q}^T \mathbf{A}^3 &= \lambda^3 \mathbf{q}^T \\ &\text{usw.} \end{aligned} \quad (7.150)$$

so dass sich das Gleichungssystem (7.148) vereinfacht zu

$$\mathbf{q}^T \mathbf{B} = \mathbf{0}^T, \lambda \mathbf{q}^T \mathbf{B} = \mathbf{0}^T, \dots, \lambda^{n-1} \mathbf{q}^T \mathbf{B} = \mathbf{0}^T \quad (7.151)$$

womit nur noch die Gleichung

$$\mathbf{q}^T \mathbf{B} = \mathbf{0}^T \quad (7.152)$$

übrig bleibt. Wenn also ein Vektor $\mathbf{q}^T \neq \mathbf{0}^T$ existiert, der diese Gleichung erfüllt und gleichzeitig Linkseigenvektor von \mathbf{A} ist, dann ist das System nicht steuerbar. Damit ergeben sich als Bedingungen für die Nicht-Steuerbarkeit die beiden Gleichungen

$$\begin{aligned} 1. \quad & \mathbf{q}^T \mathbf{B} = \mathbf{0}^T \\ 2. \quad & \mathbf{q}^T \mathbf{A} = \lambda \mathbf{q}^T \quad \text{bzw.} \quad \mathbf{q}^T (\lambda \mathbf{I} - \mathbf{A}) = \mathbf{0}^T \end{aligned} \quad (7.153)$$

oder, in einer Matrixengleichung zusammengefasst,

$$\mathbf{q}^T(\lambda \mathbf{I} - \mathbf{A}, \mathbf{B}) = \mathbf{0}^T \quad (7.154)$$

Das System ist nicht steuerbar, wenn eine Lösung $\mathbf{q}^T \neq \mathbf{0}^T$ für diese Gleichung existiert. Wenn der Rang der Matrix $(\lambda \mathbf{I} - \mathbf{A}, \mathbf{B})$ gleich n ist, dann ist die einzige Lösung die triviale Lösung $\mathbf{q}^T = \mathbf{0}^T$. Das System wäre nicht nicht-steuerbar, also steuerbar.

Satz 7.6 Hautus-Kriterium: *Ein System ist genau dann vollständig steuerbar, wenn für jeden Eigenwert λ_i von \mathbf{A} die Hautus-Matrix*

$$(\lambda_i \mathbf{I} - \mathbf{A}, \mathbf{B}) \quad (7.155)$$

den Rang n aufweist.

Sofern für einen Eigenwert λ_i dessen Hautus-Matrix den Rang n aufweist, spricht man von einem steuerbaren Eigenwert, andernfalls von einem nicht-steuerbaren.

Interessant ist das Hautus-Kriterium im Vergleich zum Kalman-Kriterium durch die Tatsache, dass hier für jeden Eigenwert einzeln die Steuerbarkeit angegeben werden kann. So kann man bei nicht vollständig steuerbaren Systemen zwischen steuerbaren und nicht steuerbaren Eigenwerten unterscheiden. Diese Möglichkeit wird später noch genutzt.

Steuerbarkeitskriterium nach Gilbert. Dieses Steuerbarkeitskriterium ist im Gegensatz zu den beiden vorhergehenden Kriterien relativ einfach herzuleiten, da es von einem System in Jordanscher Normalform ausgeht, bei dem die Systemmatrix \mathbf{A} gemäß Kapitel 7.5.2 eine reine Diagonalmatrix ist (vgl. (7.44)):

$$\dot{\mathbf{x}} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \mathbf{x} + \mathbf{B} \mathbf{u} \quad (7.156)$$

Für jede einzelne Zustandsvariable gilt damit

$$\dot{x}_i = \lambda_i x_i + \mathbf{b}_i^T \mathbf{u} \quad (7.157)$$

wobei \mathbf{b}_i^T die i -te Zeile der Eingangsmatrix \mathbf{B} ist.

Man erkennt sofort, dass die Zustandsgröße x_i nur dann beeinflusst werden kann, wenn der Vektor \mathbf{b}_i^T ungleich Null ist. Dies ist die erste Bedingung des Gilbert-Kriteriums.

Darüber hinaus ist es aber auch notwendig, dass sämtliche Zustandsgrößen unabhängig voneinander beeinflusst werden können. Sofern alle Eigenwerte voneinander verschieden sind, ist dies gewährleistet. Bei zwei gleichen

Eigenwerten λ_i und λ_j (mit $i \neq j$) muss aber offensichtlich die Zusatzforderung erhoben werden, dass die beiden Vektoren \mathbf{b}_i^T und \mathbf{b}_j^T voneinander linear unabhängig sind.

Dies ist aber immer noch nicht ausreichend. Denn bei Eingrößensystemen sind \mathbf{b}_i^T und \mathbf{b}_j^T Skalare, die gar nicht voneinander linear unabhängig sein können. Für Eingrößensysteme gilt daher die Forderung, dass sämtliche Eigenwerte voneinander verschieden sein müssen.

Satz 7.7 *Gilbert-Kriterium: Ein System, dessen Zustandsgleichung in Jordanscher Normalform vorliegt, ist nur dann vollständig steuerbar, wenn die Eingangsmatrix \mathbf{B} keine Nullzeile besitzt.*

Darüber hinaus müssen in einem Eingrößensystem alle Eigenwerte voneinander verschieden sein, und in einem MIMO-System müssen diejenigen Zeilen von \mathbf{B} , die zu den Zustandsgrößen eines mehrfachen Eigenwertes gehören, voneinander linear unabhängig sein.

Das Gilbert-Kriterium bietet damit - wie das Hautus-Kriterium - die Möglichkeit, zwischen steuerbaren und nicht-steuerbaren Eigenwerten bzw. Zustandsgrößen zu unterscheiden, sofern das Gesamtsystem nicht vollständig steuerbar ist. Während beim Hautus-Kriterium die Hautus-Matrix die entsprechenden Hinweise gibt, sind es hier, sofern das System in Jordanscher Normalform gegeben ist, die Zeilen der Eingangsmatrix \mathbf{B} .

7.10.2 Beobachtbarkeit

In Analogie zur Steuerbarkeit gilt die Definition:

Definition 7.8 *Ein System ist beobachtbar, wenn man aus den über eine endliche Zeitspanne $t \in [t_0, t_1]$ gemessenen Eingangs- und Ausgangsgrößen $\mathbf{u}(t), \mathbf{y}(t)$ jeden beliebigen Anfangs-Zustandsvektor $\mathbf{x}(t_0)$ rekonstruieren kann.*

Praxisnäher wäre diese Definition, wenn aus den bisher gemessenen Größen nicht der längst vergangene Anfangsvektor, sondern der aktuelle Zustandsvektor $\mathbf{x}(t_1)$ berechnet werden könnte. Diese Eigenschaft gibt es natürlich auch, sie wird als *Rekonstruierbarkeit* bezeichnet. Bei linearen, zeit-invarianten Systemen sind Rekonstruierbarkeit und Beobachtbarkeit äquivalente Eigenschaften.

Wie bei der Steuerbarkeit sollen nun Kriterien abgeleitet werden, mit deren Hilfe man bereits an den Systemmatrizen erkennen kann, ob ein System beobachtbar ist oder nicht. Ausgangspunkt der Überlegungen ist Gleichung (7.98), die den Zusammenhang zwischen einem Anfangszustand, dem Verlauf der Stellgröße und dem resultierenden Ausgangsvektor beschreibt:

$$\mathbf{y}(t) = \mathbf{C}e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau + \mathbf{D}\mathbf{u}(t) \quad (7.158)$$

Im vorliegenden Fall sind der Verlauf von Stellgröße und Ausgangsvektor bekannt, die einzige Unbekannte ist der Anfangszustand $\mathbf{x}(t_0)$. Umstellen der Gleichung liefert

$$\mathbf{y}(t) - \int_{t_0}^t \mathbf{C}e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau - \mathbf{D}\mathbf{u}(t) = \mathbf{C}e^{\mathbf{A}(t-t_0)} \mathbf{x}(t_0) \quad (7.159)$$

wobei sich die linke Seite durch einen resultierenden Ausgangsvektor

$$\mathbf{y}_r(t) := \mathbf{y}(t) - \int_{t_0}^t \mathbf{C}e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau - \mathbf{D}\mathbf{u}(t) \quad (7.160)$$

ersetzen lässt, dessen Verlauf bekannt ist. Gleichung (7.159) reduziert sich damit auf

$$\mathbf{y}_r(t) = \mathbf{C}e^{\mathbf{A}(t-t_0)} \mathbf{x}(t_0) \quad (7.161)$$

Die Lösbarkeit dieser Gleichung hängt, wie man sieht, von den Matrizen \mathbf{A} und \mathbf{C} ab. Die Frage lautet also: Wie müssen diese beiden Matrizen beschaffen sein, damit man aus einem gegebenen zeitlichen Verlauf von $\mathbf{y}_r(t)$ in einem Zeitintervall $t \in [t_0, t_1]$ den gesuchten Vektor $\mathbf{x}(t_0)$ berechnen kann?

Um eine Aussage über die notwendige Struktur von \mathbf{A} und \mathbf{C} ableiten zu können, soll die Matrix $\mathbf{C}e^{\mathbf{A}(t_i-t_0)}$ für einen bestimmten Zeitpunkt t_i etwas genauer analysiert werden. Unter Verwendung der Definition der Matrizen-e-Funktion (Anhang A.3) ergibt sich zunächst

$$\begin{aligned} \mathbf{C}e^{\mathbf{A}(t_i-t_0)} &= \mathbf{C} + \mathbf{C}\mathbf{A} \frac{t_i-t_0}{1!} + \mathbf{C}\mathbf{A}^2 \frac{(t_i-t_0)^2}{2!} + \mathbf{C}\mathbf{A}^3 \frac{(t_i-t_0)^3}{3!} + \dots \\ &= \sum_{k=0}^{\infty} \mathbf{C}\mathbf{A}^k \frac{(t_i-t_0)^k}{k!} \end{aligned} \quad (7.162)$$

Aufgrund des Caley-Hamilton-Theorems (Anhang A.4) können sämtliche Matrizen $\mathbf{C}\mathbf{A}^k$ für $k > n$ als Linearkombinationen der ersten n Matrizen $[\mathbf{C}, \mathbf{C}\mathbf{A}, \mathbf{C}\mathbf{A}^2, \dots, \mathbf{C}\mathbf{A}^{n-1}]$ dargestellt werden. Damit ist aber auch die Matrix $\mathbf{C}e^{\mathbf{A}(t_i-t_0)}$ darstellbar als Linearkombination dieser ersten n Matrizen, wobei die Koeffizienten h_k wegen der Vorfaktoren $\frac{(t_i-t_0)^k}{k!}$ vom jeweiligen Zeitpunkt t_i abhängig sind:

$$\mathbf{C}e^{\mathbf{A}(t_i-t_0)} = h_0(t_i)\mathbf{C} + h_1(t_i)\mathbf{C}\mathbf{A} + h_2(t_i)\mathbf{C}\mathbf{A}^2 + \dots + h_{n-1}(t_i)\mathbf{C}\mathbf{A}^{n-1} \quad (7.163)$$

Wie die Abhängigkeit $h_k(t_i)$ genau aussieht, kann nicht angegeben werden, denn ein Koeffizient h_k beinhaltet ja nicht nur den Ausdruck $\frac{(t_i-t_0)^k}{k!}$, also den direkten Vorfaktor von $\mathbf{C}\mathbf{A}^k$, sondern auch das weitere Auftauchen von $\mathbf{C}\mathbf{A}^k$ in den Linearkombinationen der Terme $\mathbf{C}\mathbf{A}^j$ mit $j \geq n$. Die Abhängigkeit ist aber im Grunde auch gar nicht relevant. Wichtig ist die Erkenntnis, dass die rechte Seite der Gleichung unabhängig vom Zeitpunkt t_i immer eine Linearkombination der Matrizen $[\mathbf{C}, \mathbf{C}\mathbf{A}, \mathbf{C}\mathbf{A}^2, \dots, \mathbf{C}\mathbf{A}^{n-1}]$ ist.

Um die weitere Erklärung zu vereinfachen, sei nun zunächst angenommen, dass das System nur eine Ausgangsgröße besitzt. Gleichung (7.161) vereinfacht sich damit zu

$$y_r(t) = \mathbf{c}^T e^{\mathbf{A}(t-t_0)} \mathbf{x}(t_0) \quad (7.164)$$

Dies ist eine Gleichung für die n Komponenten von $\mathbf{x}(t_0)$, die man aber für beliebig viele Zeitpunkte t_i aus dem Intervall $[t_0, t_1]$ aufstellen kann, so dass man beliebig viele Gleichungen erhält. Eine eindeutige Lösung für die n unbekannten Komponenten von $\mathbf{x}(t_0)$ kann man jedoch nur dann erhalten, wenn von diesen Gleichungen n linear unabhängig sind. Und das ist genau dann der Fall, wenn die Koeffizientenvektoren $\mathbf{c}^T e^{\mathbf{A}(t-t_0)}$ in den unterschiedlichen Gleichungen voneinander linear unabhängig sind.

Einsetzen von Gleichung (7.163) in Gleichung (7.164) liefert für den Fall einer Ausgangsgröße

$$y_r(t) = [h_0(t_i)\mathbf{c}^T + h_1(t_i)\mathbf{c}^T \mathbf{A} + \dots + h_{n-1}(t_i)\mathbf{c}^T \mathbf{A}^{n-1}] \mathbf{x}(t_0) \quad (7.165)$$

Der Term in den eckigen Klammern ist der Koeffizientenvektor. Er ist - unabhängig vom Zeitpunkt t_i - immer eine Linearkombination der Zeilenvektoren $[\mathbf{c}^T, \mathbf{c}^T \mathbf{A}, \mathbf{c}^T \mathbf{A}^2, \dots, \mathbf{c}^T \mathbf{A}^{n-1}]$. Je nach Zeitpunkt variieren nur die Vorfaktoren $h_k(t_i)$. n linear unabhängige Koeffizientenvektoren und damit n linear unabhängige Gleichungen zur Berechnung der n Komponenten von $\mathbf{x}(t_0)$ kann man nur dann erhalten, wenn diese n Zeilenvektoren voneinander linear unabhängig sind.

Damit ist die notwendige Bedingung für die Beobachtbarkeit von Systemen mit einer Ausgangsgröße gefunden. Diese kann nun relativ einfach für Systeme mit mehreren Ausgangsgrößen erweitert werden. Dazu kehren wir zum mehrdimensionalen Fall (7.161) mit (7.163) zurück:

$$\mathbf{y}_r(t) = [h_0(t_i)\mathbf{C} + h_1(t_i)\mathbf{C}\mathbf{A} + \dots + h_{n-1}(t_i)\mathbf{C}\mathbf{A}^{n-1}] \mathbf{x}(t_0) \quad (7.166)$$

Für jeden Zeitpunkt t_i erhält man bei q Ausgangsgrößen nun q Gleichungen, aber entscheidend für die Existenz einer eindeutigen Lösung ist nach wie vor die rechte Seite: Benötigt werden n linear unabhängige Gleichungen, und die ergeben sich nur dann, wenn sich unter den Zeilen der Matrizen $[\mathbf{C}, \mathbf{C}\mathbf{A}, \mathbf{C}\mathbf{A}^2, \dots, \mathbf{C}\mathbf{A}^{n-1}]$ n voneinander linear unabhängige Zeilen befinden.

Satz 7.9 Kalman-Kriterium: Ein System ist genau dann vollständig beobachtbar, wenn die Matrix

$$\mathbf{Q}_b = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \dots \\ \mathbf{C}\mathbf{A}^{n-1} \end{bmatrix} \quad (7.167)$$

n linear unabhängige Zeilenvektoren enthält, d.h. wenn sie den Rang n hat.

Wie schon bei der Steuerbarkeit wurde auch hier nur hergeleitet, dass aus dem Rang n der Matrix die Beobachtbarkeit folgt, nicht jedoch umgekehrt. Auf den dazu nötigen Beweis soll hier jedoch verzichtet werden.

Interessant ist nun der Vergleich zwischen dem Steuerbarkeits- und dem Beobachtbarkeitskriterium von Kalman. Wenn man die Beobachtbarkeitsmatrix \mathbf{Q}_b transponiert, so erhält man

$$\mathbf{Q}_b^T = [\mathbf{C}^T, \mathbf{A}^T \mathbf{C}^T, (\mathbf{A}^T)^2 \mathbf{C}^T, \dots, (\mathbf{A}^T)^{n-1} \mathbf{C}^T] \quad (7.168)$$

Bei dieser Matrix müssen nun n linear unabhängige Spalten auftreten. Damit entspricht diese Form aber genau dem Steuerbarkeitskriterium von Kalman, es wurde lediglich ersetzt:

$$\mathbf{A} \rightarrow \mathbf{A}^T \quad \text{und} \quad \mathbf{B} \rightarrow \mathbf{C}^T \quad (7.169)$$

Aufgrund dieses Zusammenhangs bezeichnet man Steuerbarkeit und Beobachtbarkeit als zueinander *duale* Eigenschaften. Mit Hilfe dieses Zusammenhangs können alle Sätze, die für die Steuerbarkeit entwickelt wurden, direkt auch für die Beobachtbarkeit angegeben werden. Und das gleiche gilt in den nachfolgenden Kapiteln auch für den Entwurf von Reglern und Beobachtern.

Das Hautus-Kriterium für die Beobachtbarkeit lautet beispielsweise:

Satz 7.10 *Hautus-Kriterium: Ein System ist genau dann vollständig beobachtbar, wenn für jeden Eigenwert λ_i von \mathbf{A} die Matrix*

$$\begin{bmatrix} \lambda_i \mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix} \quad (7.170)$$

den Rang n aufweist.

Sofern für einen Eigenwert λ_i dessen Hautus-Matrix den Rang n aufweist, spricht man von einem beobachtbaren Eigenwert, andernfalls von einem nicht-beobachtbaren.

Für den Beweis geht man vom Steuerbarkeitskriterium von Hautus (Satz 7.6) aus. Dies besagt, dass ein System genau dann steuerbar ist, wenn für jeden Eigenwert λ_i die Matrix

$$[\lambda_i \mathbf{I} - \mathbf{A}, \mathbf{B}] \quad (7.171)$$

den Rang n aufweist. In diesem Ausdruck ist nun lediglich gemäß (7.169) zu ersetzen, und man erhält

$$[\lambda_i \mathbf{I} - \mathbf{A}^T, \mathbf{C}^T] \quad (7.172)$$

Diese Matrix muss für jeden Eigenwert λ_i den Rang n aufweisen, damit das System beobachtbar ist. Eine Matrix verändert jedoch ihren Rang nicht, wenn man sie transponiert:

$$[\lambda_i \mathbf{I} - \mathbf{A}^T, \mathbf{C}^T]^T = \begin{bmatrix} \lambda_i \mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix} \quad (7.173)$$

Damit ist Satz 7.10 bewiesen.

Das Gilbert-Kriterium für die Beobachtbarkeit lautet:

Satz 7.11 *Gilbert-Kriterium: Ein System, dessen Zustandsgleichung in Jordanscher Normalform vorliegt, ist nur dann vollständig beobachtbar, wenn die Ausgangsmatrix \mathbf{C} keine Nullspalte besitzt.*

Darüber hinaus müssen in einem Eingrößensystem alle Eigenwerte voneinander verschieden sein, und in einem MIMO-System müssen diejenigen Spalten von \mathbf{C} , die zu den Zustandsgrößen eines mehrfachen Eigenwertes gehören, voneinander linear unabhängig sein.

Der Beweis ist noch einfacher als für das Hautus-Kriterium, denn wenn man in Satz 7.7 gemäß (7.169) \mathbf{B} durch \mathbf{C}^T ersetzt, so ergibt sich Satz 7.11 direkt.

7.10.3 Abschließende Bemerkungen

Nicht vollständig steuerbare oder beobachtbare Systeme. Immer wieder kann es vorkommen, dass bei einem System nur ein Teil der Zustandsgrößen bzw. der Eigenwerte steuerbar ist. Als Beispiel sei das System

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 0 & 0 \\ 2 & -2 & -1 \\ 1 & 2 & -3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{u} \quad (7.174)$$

angenommen.

Man sieht schon ohne Rechnung, dass die erste Zustandsgröße nicht beeinflusst werden kann. Entsprechend ergibt sich für die Steuerbarkeitsmatrix

$$\mathbf{Q}_s = [\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 5 \\ 1 & -3 & 7 \end{bmatrix} \quad (7.175)$$

Diese Matrix weist in der ersten Zeile einen Nullvektor auf. Die drei Vektoren $\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}$ liegen damit in der $x_2 - x_3$ -Ebene, was bedeutet, dass nur in dieser Ebene Zustände beliebig angefahren werden können. Diese Aussage deckt sich daher mit der Anschauung.

Ein analoges Ergebnis liefert das Hautus-Kriterium. Die Eigenwerte des Systems sind -1 sowie $\frac{1}{2}(-5 \pm j\sqrt{7})$. Für die Eigenwerte $\frac{1}{2}(-5 \pm j\sqrt{7})$ hat die Hautus-Matrix jeweils den Rang 3, während sich für den Eigenwert -1 die Hautus-Matrix

$$(\lambda_i \mathbf{I} - \mathbf{A}, \mathbf{B}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 0 \\ -1 & -2 & 2 & 1 \end{bmatrix} \quad (7.176)$$

ergibt, die offensichtlich nur den Rang 2 aufweist. Der Eigenwert -1 ist damit nicht steuerbar.

Geht man nun noch einmal zurück zur Zustandsgleichung, so folgt aus dieser für die erste Zustandsvariable

$$\dot{x}_1 = -x_1 \quad (7.177)$$

Das charakteristische Polynom dieser Gleichung lautet offenbar $s + 1 = 0$, und dessen Nullstelle bzw. der Eigenwert dieser Gleichung ist -1 . Daraus folgt, dass der Eigenwert -1 gerade zur ersten Zustandsgröße gehört. Damit deckt sich dann aber auch das Ergebnis des Hautus-Kriteriums mit dem des Kalman-Kriteriums.

Das Beispiel zeigt, dass in einem nicht vollständig steuerbaren System relativ einfach ermittelt werden kann, welche Eigenwerte bzw. Zustandsgrößen nicht steuerbar sind. Damit hat man die Möglichkeit zu entscheiden, ob man mit diesem Mangel leben kann (wenn es sich z.B. um einen stabilen Eigenwert und eine für das Systemverhalten nicht entscheidende Zustandsgröße handelt) oder ob Abhilfe in Form einer zusätzlichen Stellgröße zu schaffen ist.

Grundsätzlich sollte man versuchen, bei nicht vollständig steuerbaren Systemen durch Basistransformation oder schlichtes Umsortieren der Zustandsgrößen das System in die Form

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \end{bmatrix} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{0} \end{bmatrix} \mathbf{u} \\ \mathbf{y} &= [\mathbf{C}_1, \mathbf{C}_2] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \end{aligned} \quad (7.178)$$

zu bringen, wobei der Zustandsvektor $\mathbf{x}^T = (\mathbf{x}_1^T, \mathbf{x}_2^T)$ aus einem steuerbaren Anteil \mathbf{x}_1 und einem nicht steuerbaren Anteil \mathbf{x}_2 besteht. Die Nicht-Steuerbarkeit von \mathbf{x}_2 kann man an der Zustandsgleichung klar erkennen.

Da die Zustandsgrößen des Vektors \mathbf{x}_2 nicht steuerbar sind, können sie aus Sicht der Regelung nur als Störgrößen betrachtet werden, die aber durchaus beobachtbar sein können und deren Auswirkungen damit durch eine Vorsteuerung effektiv bekämpft werden können (vgl. Kap. 6.4.5).

Die Regelung wird dann anhand desjenigen Teils der Strecke ausgelegt, dessen dynamisches Verhalten auf den steuerbaren Zustandsgrößen basiert (vgl. (7.107)):

$$\mathbf{G}_{red}(s) = \mathbf{C}_1(s\mathbf{I} - \mathbf{A}_{11})^{-1}\mathbf{B}_1 \quad (7.179)$$

In analoger Weise sind auch nicht vollständig beobachtbare Systeme zu behandeln, worauf hier aber nicht näher eingegangen werden soll, da die Überlegungen praktisch die gleichen sind.

Ausgangssteuerbarkeit. Anstelle der (Zustands-)Steuerbarkeit, wie sie bisher betrachtet wurde, kann man auch direkt die Frage nach der Steuerbarkeit der Ausgangsgrößen stellen, was für praktische Anwendungen oftmals sogar naheliegender ist.

Diese Betrachtung hilft auch bei der Beantwortung der bei nicht vollständig (zustands-)steuerbaren Systemen auftretenden Frage, ob die nicht steuerbaren Zustandsgrößen denn für das Systemverhalten relevant oder irrelevant sind.

Man sollte sich aber nicht dazu verleiten lassen, auf die Betrachtung der (Zustands-)Steuerbarkeit komplett zu verzichten, wenn sämtliche Ausgangsgrößen steuerbar sind. Denn es ist ja durchaus denkbar, dass ein System, welches vollständig ausgangssteuerbar ist, interne Zustandsgrößen aufweist, die zu instabilen Eigenwerten gehören und weder steuerbar noch beobachtbar sind. Über derartige Zustandsgrößen sollte man zumindest im Bilde sein, auch wenn sie keine Auswirkung auf die Ausgangsgrößen haben.

Die Formeln für die Ausgangssteuerbarkeit sind aufwendiger als für die Zustandssteuerbarkeit, da die Ausgangs- und die Durchgangsmatrix nun Berücksichtigung finden müssen. Das Kalman-Kriterium, welches sich für die Ausgangssteuerbarkeit mit denselben Überlegungen herleiten lässt wie für die Zustandssteuerbarkeit, führt hier auf die (Ausgangs)steuerbarkeitsmatrix

$$\mathbf{Q}_{sa} = [\mathbf{CB} \ , \ \mathbf{CAB} \ , \ \mathbf{CA}^2\mathbf{B} \ , \ \dots, \ \mathbf{CA}^{n-1}\mathbf{B} \ , \ \mathbf{D}] \quad (7.180)$$

welche nun den Rang q haben muss, wobei q die Dimension des Ausgangsvektors \mathbf{y} ist.

Zukünftig wird die Ausgangssteuerbarkeit nicht weiter diskutiert, so dass, wenn von Steuerbarkeit die Rede ist, immer die Zustandssteuerbarkeit gemeint ist.

8. Zustandsregelung

Nachdem die Grundlagen der Zustandsdarstellung im letzten Kapitel ausführlich behandelt wurden, soll nun auf den Entwurf linearer Zustandsregler eingegangen werden.

Für die weiteren Ausführungen sei vorausgesetzt, dass $\mathbf{D} = \mathbf{0}$ ist und \mathbf{B} sowie \mathbf{C} Höchststrang aufweisen, d.h. die Spalten von \mathbf{B} bzw. die Zeilen von \mathbf{C} sind linear unabhängig. Doch was bedeuten diese Einschränkungen?

Relativ einfach ist die Einschränkung für \mathbf{D} zu verstehen. \mathbf{D} stellt ja den direkten Durchgriff vom Eingang zum Ausgang dar. Wenn sich eine Eingangsgröße sprungförmig ändert, so wird sich für $\mathbf{D} \neq \mathbf{0}$ auch mindestens eine Ausgangsgröße sprungförmig ändern. Dies kann aber für Strecken mit Tiefpasscharakter, wie sie in der Praxis üblich sind, nicht vorkommen. Von daher ist die Voraussetzung $\mathbf{D} = \mathbf{0}$ gerechtfertigt und stellt keine besondere Einschränkung hinsichtlich praktischer Anwendungen dar. Darüber hinaus wird der Reglerentwurf deutlich vereinfacht.

Um die Einschränkung bzgl. \mathbf{B} zu verstehen, muss man die Zustandsgleichung betrachten. Es gilt

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} = \mathbf{A}\mathbf{x} + [\mathbf{b}_1, \dots, \mathbf{b}_p] \begin{bmatrix} u_1 \\ \vdots \\ u_p \end{bmatrix} \\ &= \mathbf{A}\mathbf{x} + \mathbf{b}_1 u_1 + \dots + \mathbf{b}_p u_p\end{aligned}\tag{8.1}$$

Dabei sind die \mathbf{b}_i die Spalten von \mathbf{B} . Der Ausdruck $[\mathbf{b}_1 u_1 + \dots + \mathbf{b}_p u_p]$ ist eine Linearkombination der Spalten von \mathbf{B} und beschreibt die Wirkung des Eingangsvektors \mathbf{u} auf die Veränderung des Zustandsvektors $\dot{\mathbf{x}}$. Falls die \mathbf{b}_i nicht voneinander linear unabhängig sind, so lässt sich einer dieser Vektoren (z.B. \mathbf{b}_p) als Linearkombination der übrigen Vektoren angeben:

$$\mathbf{b}_p = k_1 \mathbf{b}_1 + \dots + k_{p-1} \mathbf{b}_{p-1}\tag{8.2}$$

Entsprechend gilt

$$u_p \mathbf{b}_p = k_1 u_p \mathbf{b}_1 + \dots + k_{p-1} u_p \mathbf{b}_{p-1}\tag{8.3}$$

und damit für die Zustandsgleichung

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + u_1\mathbf{b}_1 + \dots + u_{p-1}\mathbf{b}_{p-1} + k_1u_p\mathbf{b}_1 + \dots + k_{p-1}u_p\mathbf{b}_{p-1} \\
&= \mathbf{A}\mathbf{x} + (u_1 + k_1u_p)\mathbf{b}_1 + \dots + (u_{p-1} + k_{p-1}u_p)\mathbf{b}_{p-1} \\
&= \mathbf{A}\mathbf{x} + [\mathbf{b}_1, \dots, \mathbf{b}_{p-1}] \begin{bmatrix} u_1 + k_1u_p \\ \vdots \\ u_{p-1} + k_{p-1}u_p \end{bmatrix}
\end{aligned} \tag{8.4}$$

Man sieht, dass die letzte Spalte von \mathbf{B} gar nicht mehr auftaucht, und \mathbf{u} hat sich um eine Komponente verkürzt, obwohl Gleichung (8.4) zur Ausgangsgleichung (8.1) natürlich äquivalent ist. Die letzte Stellgrößenkomponente ist demnach überflüssig.

Damit hat man im ursprünglichen System also mehr Stellgrößen zur Verfügung, als man eigentlich bräuchte. Das bedeutet aber wiederum, dass ein Regler nicht nur ermitteln muss, was zu tun ist, um die Strecke in einer bestimmten Art und Weise zu beeinflussen, sondern auch noch entscheiden muss, mit welchen Stellgrößen er die Wirkung erzielen will. Natürlich sind solche überbestimmten Systeme in der Praxis möglich und technisch hochinteressant, da man mit den zusätzlichen Freiheitsgraden auch zusätzliche Anforderungen an die Regelung erfüllen kann, aber die Theorie wird deutlich erschwert, so dass hier darauf verzichtet werden soll. Wir können davon ausgehen, dass - sofern das System steuerbar ist - genau die notwendige Anzahl an Stellgrößen zur Verfügung steht.

Zur Erläuterung der Einschränkung für \mathbf{C} wird die Ausgangsgleichung betrachtet:

$$\begin{aligned}
\mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} = \mathbf{C}\mathbf{x} \\
\begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix} &= \begin{bmatrix} \mathbf{c}_1^T \\ \vdots \\ \mathbf{c}_q^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{c}_1^T \mathbf{x} \\ \vdots \\ \mathbf{c}_q^T \mathbf{x} \end{bmatrix}
\end{aligned} \tag{8.5}$$

Jede Komponente des Ausgangsvektors ist das Produkt aus einer Zeile von \mathbf{C} und dem Zustandsvektor \mathbf{x} . Wenn nun die Zeilen von \mathbf{C} linear abhängig sind, so lässt sich eine dieser Zeilen (z.B. \mathbf{c}_q^T) als Linearkombination der übrigen Zeilen angeben:

$$\mathbf{c}_q^T = k_1\mathbf{c}_1^T + \dots + k_{q-1}\mathbf{c}_{q-1}^T \tag{8.6}$$

Entsprechend gilt

$$\begin{aligned}
\mathbf{c}_q^T \mathbf{x} &= k_1\mathbf{c}_1^T \mathbf{x} + \dots + k_{q-1}\mathbf{c}_{q-1}^T \mathbf{x} \\
y_q &= k_1y_1 + \dots + k_{q-1}y_{q-1}
\end{aligned} \tag{8.7}$$

Das bedeutet aber, dass sich auch die entsprechende Komponente y_q des Ausgangsvektors als Linearkombination der übrigen Komponenten des Ausgangsvektors darstellen lässt. Sie könnte damit entfallen, weil man ihren Wert rechnerisch leicht aus den übrigen Ausgangsgrößen ermitteln kann.

Der Fall, dass man mehr Ausgangsgrößen, d.h. Messgrößen, zur Verfügung hat, als man eigentlich bräuchte, kommt in der Praxis, vor allem in sicherheitsrelevanten Bereichen, oft vor. Dort ist man daran interessiert, auf mehr

als einem Weg Informationen über eine bestimmte Systemgröße zu erhalten, damit man bei Ausfall eines einzelnen Sensors trotzdem noch weiß, was in der Strecke passiert. Die - im Normalfall - gleichen Messwerte werden in einer vorgeschalteten Auswahllogik beispielsweise nach dem 2-aus-3-Prinzip (3 Messwerte derselben Größe liegen vor. Wenn 2 davon den gleichen Wert aufweisen, wird dieser Wert als richtig angenommen) ausgewertet, und der als zutreffend angenommene Messwert wird dann an den Regler weitergegeben. Aus Sicht des Reglers ist die Auswahllogik irrelevant. Er bekommt den Messwert, den er braucht, und die Tatsache, dass dieser Messwert zuvor aus mehreren Messwerten ermittelt wurde, hat auf die Dynamik des Regelkreises und damit für die Auslegung des Reglers keine Bedeutung und kann vernachlässigt werden. Ohne Einschränkung der Allgemeinheit kann deshalb von nun an angenommen werden, dass alle dem Regler zugeführten Messwerte, d.h. Ausgangsgrößen y_i , voneinander linear unabhängig sind und dies damit auch für die Zeilen der Matrix \mathbf{C} gilt.

8.1 Konzept der Zustandsregelung

Bei Kenntnis der internen Zustandsgrößen ist eine Systembeeinflussung sehr einfach und elegant möglich, wie Abb. 8.1 zeigt. Der Zustandsvektor wird lediglich mit einer konstanten Matrix \mathbf{F} multipliziert und auf den Eingang zurückgeführt.

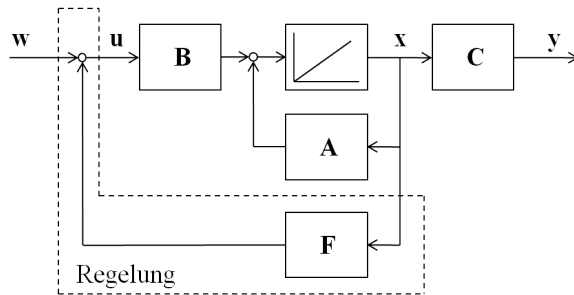


Abb. 8.1. Grundstruktur einer Zustandsregelung

Die Zustandsdarstellung dieses Systems lautet:

$$\begin{aligned}\dot{\mathbf{x}} &= (\mathbf{A} + \mathbf{BF})\mathbf{x} + \mathbf{B}\mathbf{w} \\ \mathbf{y} &= \mathbf{C}\mathbf{x}\end{aligned}\tag{8.8}$$

Man sieht, dass sich durch die Regelung ein neues System mit der Systemmatrix $(\mathbf{A} + \mathbf{BF})$ ergibt. Wenn \mathbf{F} so gewählt wird, dass sämtliche Eigenwerte von $(\mathbf{A} + \mathbf{BF})$ einen negativen Realteil haben, dann ist das System auf jeden Fall stabil im Sinne von Ljapunov (vgl. Satz 7.3). Durch die Wahl von \mathbf{F} können

die Eigenwerte darüber hinaus so gelegt werden, dass auch die Dämpfung des Systems einem gewünschten Wert entspricht.

Stabilität nach Ljapunov bedeutet, dass das System ohne äußere Anregung aus jedem beliebigen Anfangszustand gegen den Ruhezustand $\mathbf{x} = \mathbf{0}$ konvergiert (vgl. Def. 7.2)). Da aus der Stabilität nach Ljapunov auch die Übertragungsstabilität folgt, ist gewährleistet, dass das Übertragungsverhalten vom Sollwert \mathbf{w} zur Ausgangsgröße \mathbf{y} auf jeden Fall stabil ist. Noch nichts gesagt ist damit aber über die Genauigkeit der Regelung, d.h. ob die Ausgangsgröße - zumindest im stationären Zustand - dem Sollwert entspricht. Diese Frage wird im folgenden Unterkapitel geklärt.

Zuvor sei noch ein grundsätzlicher Punkt angesprochen: Offensichtlich ist ein Zustandsregler in Form der Matrix \mathbf{F} nichts weiter als ein simples mehrdimensionalen Proportionalglied, da ja lediglich die Zustandsgrößen mit konstanten Faktoren multipliziert und wieder auf die Strecke zurückgeführt werden. Im Gegensatz dazu sind Regler vom PID-Typ außerordentlich kompliziert mit ihren Vorhalten und Integratoren. Die Einfachheit des Zustandsreglers ist darauf zurückzuführen, dass einem PID-Regler nur die Ausgangsgrößen der Strecke als Informationsquelle zur Verfügung stehen, während ein Zustandsregler ständig auf die Information über den gesamten Zustand der Strecke zugreifen kann. Dieses Informationsdefizit muss im PID-Regler durch aufwändigere reglerinterne Berechnungen ausgeglichen werden. Dafür treten beim Zustandsregler aber andere Probleme auf, die Erweiterungen des Konzepts erforderlich machen und nachfolgend diskutiert werden. Erst anschließend wenden wir uns dann den unterschiedliche Entwurfsverfahren für die Berechnung der Matrix \mathbf{F} zu.

8.2 Stationäre Genauigkeit von Zustandsreglern

Angeichts der Tatsache, dass die oben eingeführte Grundform des Zustandsreglers nur ein mehrdimensionaler Proportionalregler ist, stellt sich sofort die Frage nach der stationären Genauigkeit eines solchen Reglers. In der Tat ist diese nicht einmal im ungestörten Fall gewährleistet.

Stationäre Genauigkeit bedeutet, dass die durch $\mathbf{y}(s) = \mathbf{T}(s)\mathbf{w}(s)$ definierte Übertragungsmatrix des geschlossenen Kreises (vgl. (6.3)) die Bedingung

$$\lim_{s \rightarrow 0} \mathbf{T}(s) \stackrel{!}{=} \mathbf{I} \quad (8.9)$$

erfüllt, was bedeutet, dass für Gleichsignale der Ausgangsvektor \mathbf{y} gleich dem Sollwert-Eingangsvektor \mathbf{w} sein muss.

Mit der Systemmatrix $(\mathbf{A} + \mathbf{BF})$ des geschlossenen Kreises gemäß (8.8) folgt aus (7.107) für die Übertragungsmatrix des geschlossenen Kreises

$$\mathbf{T}(s) = \mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{B} \quad (8.10)$$

und damit die Forderung

$$\lim_{s \rightarrow 0} \mathbf{T}(s) = \lim_{s \rightarrow 0} [\mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{B}] \stackrel{!}{=} \mathbf{I} \quad (8.11)$$

Üblicherweise erfolgt die Parametrierung der Rückkopplungsmatrix \mathbf{F} im Hinblick auf Stabilität und Dämpfung des geschlossenen Kreises. Weitere Anforderungen, wie z.B. die stationäre Genauigkeit, können dann normalerweise nicht mehr berücksichtigt werden, da die Anzahl der freien Parameter (also der Koeffizienten von \mathbf{F}) dazu nicht ausreicht. Aus dem Grund wird in den Referenzkanal die Multiplikation mit einer weiteren Matrix, dem *Vorfilter* \mathbf{R} , eingefügt, um zusätzliche freie Parameter zu gewinnen und diese für die Gewährleistung der stationären Genauigkeit zu nutzen (Abb. 8.2).

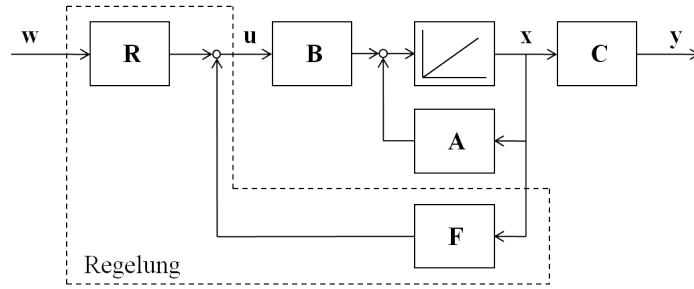


Abb. 8.2. Zustandsregelung mit Referenz-Filtermatrix \mathbf{R}

Die Regelung besteht nun aus der Rückkopplung \mathbf{F} und dem Vorfilter \mathbf{R} . Die Grundstruktur aus Abbildung 6.1 haben wir damit hinter uns gelassen.

Durch Hinzufügen des Vorfilters erweitert sich die Übertragungsmatrix des Gesamtsystems durch die Multiplikation von \mathbf{w} mit \mathbf{R} gemäß

$$\mathbf{y}(s) = \mathbf{T}(s)\mathbf{w}(s) = \mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{BR}\mathbf{w}(s) \quad (8.12)$$

und die Forderung nach stationärer Genauigkeit wird zu

$$\lim_{s \rightarrow 0} \mathbf{T}(s) = \lim_{s \rightarrow 0} [\mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{BR}] \stackrel{!}{=} \mathbf{I} \quad (8.13)$$

Wenn man s gegen Null gehen lässt, so verschwindet lediglich der Term $s\mathbf{I}$ und man erhält

$$-\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{BR} = \mathbf{I} \quad (8.14)$$

und daraus die Bestimmungsgleichung für \mathbf{R} :

$$\mathbf{R} = -[\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}]^{-1} \quad (8.15)$$

Gemäß dieser Gleichung kann man die Vorfiltermatrix \mathbf{R} berechnen, sofern die Rückkopplungsmatrix \mathbf{F} schon vorliegt. Diese Berechnungsreihenfolge ist durchaus sinnvoll, denn erst wird \mathbf{F} so dimensioniert, dass die wesentlichste

Anforderung an einen Regler, nämlich die Stabilität des geschlossenen Kreises, erfüllt ist. Und erst dann wird darauf aufbauend \mathbf{R} für die Einhaltung der Genauigkeit berechnet.

Voraussetzung für die Berechenbarkeit von \mathbf{R} ist zunächst die Invertierbarkeit von $(\mathbf{A} + \mathbf{BF})$. Da \mathbf{F} aber so ausgelegt wird, dass das Gesamtsystem stabil ist, liegen sämtliche Eigenwerte λ_i von $(\mathbf{A} + \mathbf{BF})$ in der linken Hälfte der komplexen Ebene, und es gilt:

$$\det(\mathbf{A} + \mathbf{BF}) = \prod_{i=1}^n \lambda_i \neq 0 \quad (8.16)$$

Damit ist gewährleistet, dass $(\mathbf{A} + \mathbf{BF})$ invertierbar ist.

Weiterhin muss aber auch der Gesamtterm $[\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}]$ invertierbar sein, und dies ist nicht zwangsläufig gegeben. Auf genauere Ausführungen soll hier aber zunächst verzichtet werden. Die Invertierbarkeit dieses Terms wird später im Zusammenhang mit der Betrachtung der Nullstellen eines Systems (Kapitel 9.4.1) noch einmal diskutiert.

Auch das Einfügen von \mathbf{R} ist jedoch noch nicht die endgültige Lösung des Problems. Dies verdeutlicht die Untersuchung der Stör-Übertragungsfunktion für eine vor der Regelstrecke angreifende Störung \mathbf{z} gemäß Abbildung 8.3 (vgl. auch Abb. 6.1).

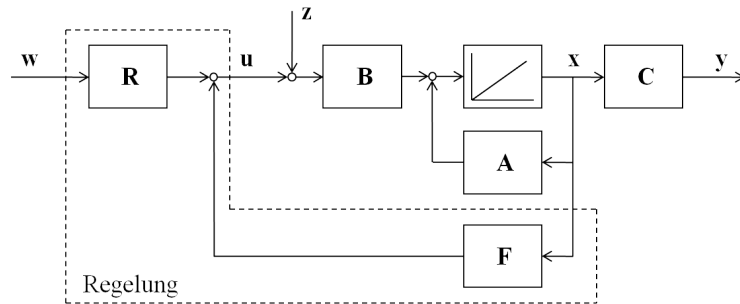


Abb. 8.3. Gestörtes System mit Zustandsregelung

Setzt man, wie bei der Berechnung der Stör-Übertragungsfunktion üblich, $\mathbf{w} = \mathbf{0}$, so ergibt sich aus dem Blockschaltbild:

$$\mathbf{B}\mathbf{z} + \mathbf{B}\mathbf{F}\mathbf{x} + \mathbf{A}\mathbf{x} = s\mathbf{x} \quad (8.17)$$

Umstellen liefert

$$\begin{aligned} \mathbf{B}\mathbf{z} &= (s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))\mathbf{x} \quad \text{bzw.} \\ \mathbf{x} &= (s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{B}\mathbf{z} \end{aligned} \quad (8.18)$$

und daraus für das Übertragungsverhalten von der Störung zum Ausgang

$$\mathbf{y} = \mathbf{C}\mathbf{x} = \mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{F}))^{-1}\mathbf{B}\mathbf{z} = \mathbf{S}(s)\mathbf{z} \quad (8.19)$$

Für Gleichsignale, d.h. $s = 0$, sollte die Stör-Übertragungsfunktion $\mathbf{S}(s)$ gemäß Gleichung (6.3) verschwinden. Tatsächlich bleibt aber hier

$$\mathbf{S}(0) = -\mathbf{C}(\mathbf{A} + \mathbf{B}\mathbf{F})^{-1}\mathbf{B} \quad (8.20)$$

Daraus folgt, dass eine stationäre Störung durch die aus \mathbf{F} und \mathbf{R} bestehende Reglerstruktur nicht ausgeregelt wird. Dies ist auch anschaulich sofort einsichtig. \mathbf{R} hat offensichtlich überhaupt keinen Einfluss auf die Stör-Übertragungsfunktion, da die Störung hinter \mathbf{R} angreift und von \mathbf{R} gar nicht bemerkt werden kann. \mathbf{R} beeinflusst - als Führungsgrößenfilter - lediglich die Führungs-Übertragungsfunktion, nicht jedoch die Ausregelung von Störungen.

Es ist also eine andere Reglerstruktur erforderlich, um nicht nur den geschlossenen Kreis zu stabilisieren, sondern um stationäre Genauigkeit auch in Anwesenheit von stationären Störungen zu gewährleisten. Diese Struktur ist in Abbildung 8.4 gezeigt.

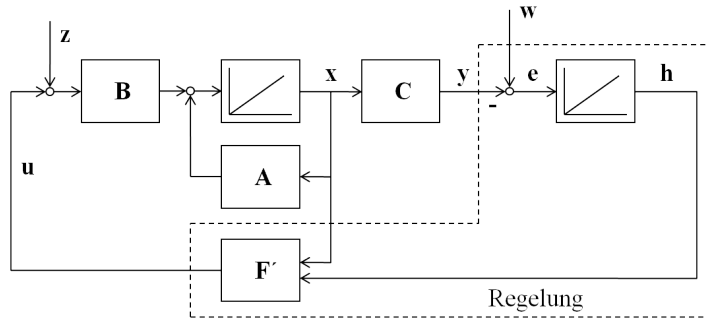


Abb. 8.4. Erweiterte Zustandsregelung für stationäre Genauigkeit

Der Angriffspunkt der Sollgröße hat sich nicht verändert, er ist nur anders eingezeichnet. Der Regelfehler $\mathbf{e} = \mathbf{w} - \mathbf{y}$ wird jetzt integriert, und der resultierende Ausgangsvektor \mathbf{h} des Integrators enthält in seinen Komponenten zusätzliche Zustandsgrößen des Systems, da es sich hier ja um Ausgänge von Integratoren handelt. Der neue Zustandsvektor, der auch der Eingangsvektor der neuen, erweiterten Rückführmatrix \mathbf{F}' ist, lautet

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} \quad (8.21)$$

Für die Stabilitätsbetrachtung nach Ljapunov sind die externen Größen nicht interessant, so dass $\mathbf{w} = \mathbf{z} = \mathbf{0}$ gesetzt werden kann. Damit lässt sich das System beschreiben durch die Systemgleichung

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{h}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{u} \quad (8.22)$$

und die Regler-Gleichung

$$\mathbf{u} = \mathbf{F}' \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} \quad (8.23)$$

Während in der Grundstruktur $\dim(\mathbf{F}) = p \times n$ galt (p = Anzahl der Stellgrößen, n = Anzahl der Zustandsgrößen), so gilt für \mathbf{F}' wegen des erweiterten Zustandsvektors $\dim(\mathbf{F}') = p \times (n + q)$ (q = Anzahl der Ausgangsgrößen).

Einsetzen der Regler-Gleichung in die Systemgleichung liefert die Gleichung des geschlossenen, von außen nicht angeregten Kreises:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{h}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{F}' \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{F}' \right) \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} \quad (8.24)$$

Für dieses erweiterte System ist nun \mathbf{F}' so auszulegen, dass alle Eigenwerte der neuen Systemmatrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{F}' \quad (8.25)$$

einen negativen Realteil besitzen (vgl. Satz 7.3). Damit ist sichergestellt, dass das System aufgrund seiner Stabilität einen stationären Endzustand erreicht. Dies ist gleichbedeutend damit, dass die Eingangsgrößen aller Integratoren verschwinden müssen, also auch der Fehlervektor \mathbf{e} . Damit ist stationäre Genauigkeit erreicht.

Der Nachteil der erweiterten Struktur gemäß Abbildung 8.4 besteht in der niedrigen Regelgeschwindigkeit. Gegenüber der Grundstruktur in Abbildung 8.1 muss hier der Sollwert \mathbf{w} erst einen Integrator durchlaufen, bevor er - über die beiden proportional wirkenden Matrizen \mathbf{F}' und \mathbf{B} - auf die Strecke wirkt. Damit ergibt sich der gleiche Geschwindigkeitsnachteil wie bei einem klassischen Eingrößen-Integralregler.

Aus dem Grund erweitert man das Schema genau so, wie man auch den Eingrößen-Integralregler um einen Proportionalanteil zum PI-Regler erweitert hat (Abb. 8.5).

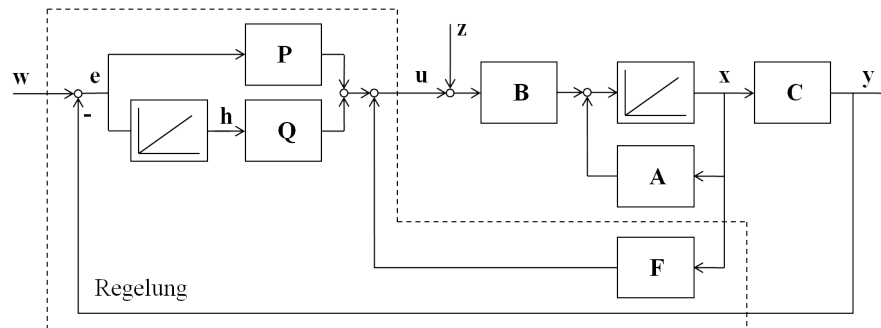


Abb. 8.5. PI-Zustandsregelung für stationäre Genauigkeit

Der Angriffspunkt des Sollwertes ist nun wieder wie bei der ursprünglichen Konstellation links eingezeichnet. Der Regelfehler $\mathbf{e} = \mathbf{w} - \mathbf{y}$ wird nicht - wie oben - integriert und anschließend als zusätzliche Zustandsgröße auf die Rückführmatrix \mathbf{F} gegeben, sondern er wird jetzt über eine PI-Struktur direkt auf die Strecke geleitet.

Dennoch müssen die Zustandsgleichungen gegenüber dem gerade beschriebenen Fall im Grunde nicht verändert werden, wenn man Strecke und Regelung geeignet aufteilt. Der Zustandsvektor setzt sich natürlich auch hier aus den Komponenten der Vektoren \mathbf{x} und \mathbf{h} zusammen:

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} \quad (8.26)$$

Für die Betrachtung der Stabilität im Ljapunovschen Sinn werden die externen Größen zu Null gesetzt ($\mathbf{w} = \mathbf{z} = \mathbf{0}$). Darüber hinaus wird zum Aufstellen der Systemgleichungen die Aufteilung zwischen Regler und Strecke aufgegeben, und die Gleichungen werden stattdessen so formuliert, wie es für den späteren Reglerentwurf am sinnvollsten ist. Dies ist problemlos möglich, solange nur die Gleichungen die dynamischen Zusammenhänge des Systems korrekt und vollständig beschreiben.

Mit dieser Zielsetzung ergibt sich für die Systemgleichung, in die auch die Integratoren der Regelung mit einbezogen sind

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{h}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{u} \quad (8.27)$$

und für die Regelung

$$\mathbf{u} = [\mathbf{F} - \mathbf{P}\mathbf{C}, \mathbf{Q}] \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} = \mathbf{F}' \begin{bmatrix} \mathbf{x} \\ \mathbf{h} \end{bmatrix} \quad (8.28)$$

Man sieht, dass die Integration der Regelabweichung noch der Strecke zugeordnet wird, obwohl sie doch eigentlich im Regler abläuft, und dafür die Streckenmatrix \mathbf{C} in der Regler-Rückführmatrix auftaucht. Dadurch entsteht einerseits eine Rückkopplungsmatrix \mathbf{F}' ohne interne Integratoren und andererseits eine Systemgleichung, die der Standard-Form genügt. Auf eine Ausgangsgleichung kann man verzichten, da der Zusammenhang $\mathbf{y} = \mathbf{C}\mathbf{x}$ bereits in der Reglergleichung enthalten ist.

Auch hier hat die erweiterte Rückführmatrix $\mathbf{F}' = [\mathbf{F} - \mathbf{P}\mathbf{C}, \mathbf{Q}]$ wegen des erweiterten Zustandsvektors die Dimension $p \times (n + q)$ (p = Anzahl der Stell- bzw. Eingangsgrößen, q = Anzahl der Ausgangsgrößen).

Für dieses erweiterte System ist nun \mathbf{F}' auch hier so auszulegen, dass alle Eigenwerte der neuen Systemmatrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{F}' \quad (8.29)$$

einen negativen Realteil besitzen. Die entsprechenden Verfahren werden in den nachfolgenden Kapiteln vorgestellt.

Setzt man nun voraus, dass \mathbf{F}' mit Hilfe eines geeigneten Verfahrens bestimmt werden konnte, so bleibt die Frage, wie daraus die einzelnen Teilmatrizen des Reglers \mathbf{F} , \mathbf{P} und \mathbf{Q} abzuleiten sind. Hierzu ist der Zusammenhang

$$\mathbf{F}' = [\mathbf{F} - \mathbf{P}\mathbf{C}, \mathbf{Q}] \quad (8.30)$$

heranzuziehen. Wenn \mathbf{F}' festliegt, so ist \mathbf{Q} , und damit der Integralanteil des PI-Reglers, eindeutig definiert, denn \mathbf{Q} entspricht dem rechten Teil von \mathbf{F}' . Für den linken Teil der Matrix gibt es aber unendlich viele Möglichkeiten, wie \mathbf{F} und \mathbf{P} gewählt werden können, um diese Gleichung zu erfüllen.

Eine einfache Lösung wäre, entweder \mathbf{P} oder \mathbf{F} zu Null zu setzen. Setzt man \mathbf{F} gleich Null, so bleibt für den linken Teil der Regler-Rückführung der Zusammenhang

$$\mathbf{F}'_l = -\mathbf{P}\mathbf{C} \quad (8.31)$$

Je nach Struktur von \mathbf{F}' und \mathbf{C} wird im Normalfall für diese Gleichung keine Lösung existieren. Dies ist nicht verwunderlich, denn \mathbf{F}' wurde unter der Voraussetzung entworfen, dass man eine vollwertige Zustandsregelung zur Verfügung hat, und dazu ist \mathbf{F} offenbar unverzichtbar, wie die Nicht-Lösbarkeit von (8.31) zeigt. Würde man auf die Zustandsrückführung \mathbf{F} verzichten, so würde man bei der Regelung auf die interne Systeminformation verzichten, die ja im Zustandsvektor steckt, und stattdessen nur auf die - weniger informativen - Ausgangsgrößen \mathbf{y} der Strecke zurückgreifen. Dass eine solche Vorgehensweise normalerweise keine gleichwertigen Ergebnisse liefern kann, ist offensichtlich.

Das Nullsetzen der Zustandsrückführung \mathbf{F} scheidet als genereller Ansatz damit aus, und die Frage, warum man angesichts eines kompletten, mehrdimensionalen PI-Reglers überhaupt eine zusätzliche Zustandsrückführung \mathbf{F} braucht, ist damit beantwortet.

Im Gegensatz zum Nullsetzen von \mathbf{F} ist das Nullsetzen von \mathbf{P} jedoch definitiv möglich. In dem Fall muss man nur \mathbf{F} gleich der linken Hälfte von \mathbf{F}' setzen. Damit hat man aber wieder exakt die gleiche Struktur wie in Abbildung 8.4 und auch das gleiche Problem, nämlich dass das Führungsverhalten wegen des reinen I-Reglers relativ langsam ist. Dennoch ist diese Konstellation ein guter Ausgangspunkt, um zu einer guten Lösung zu kommen. Denn man muss jetzt nur langsam die Koeffizienten von \mathbf{P} von Null ausgehend erhöhen, um zu einer besseren Regelgeschwindigkeit zu gelangen, wobei man jedoch nicht vergessen darf, die Koeffizienten von \mathbf{F} entsprechend nachzuführen, damit Gleichung (8.30) nach wie vor erfüllt ist.

Damit kommen wir zur letzten offenen Frage, nämlich ob das erweiterte System überhaupt steuerbar ist. Schließlich taucht ja nun die Ausgangsmatrix \mathbf{C} in der Systemmatrix auf, und es wird sicherlich strukturelle Anforderungen an das Zusammenspiel aus \mathbf{A} , \mathbf{B} und \mathbf{C} geben, damit Steuerbarkeit gilt. Die Untersuchung soll mit Hilfe des Hautus-Kriteriums (Satz 7.6) geführt werden.

Zunächst sind die Eigenwerte λ_i der neuen, erweiterten Systemmatrix zu ermitteln. Die Eigenwerte einer Matrix \mathbf{M} sind als Lösungen der Gleichung $\det(\lambda_i \mathbf{I} - \mathbf{M}) = 0$ definiert. Dies bedeutet hier:

$$\det \begin{bmatrix} \lambda_i \mathbf{I}_n - \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \lambda_i \mathbf{I}_q \end{bmatrix} = 0 \quad (8.32)$$

Dabei ist \mathbf{I}_n eine Einheitsmatrix der Dimension $n \times n$ und \mathbf{I}_q eine Einheitsmatrix der Dimension $q \times q$. Die Analyse der Determinante ergibt

$$\det \begin{bmatrix} \lambda_i \mathbf{I}_n - \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \lambda_i \mathbf{I}_q \end{bmatrix} = \det(\lambda_i \mathbf{I} - \mathbf{A}) \lambda_i^q = 0 \quad (8.33)$$

Die Lösungen dieser Gleichung und damit die Eigenwerte der erweiterten Systemmatrix entsprechen also den Eigenwerten von \mathbf{A} zuzüglich einem q -fachen Eigenwert bei Null.

Das Hautus-Kriterium fordert nun, dass für jeden dieser Eigenwerte die Hautus-Matrix

$$\begin{bmatrix} \lambda_i \mathbf{I}_n - \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \lambda_i \mathbf{I}_q & \mathbf{0} \end{bmatrix} \quad (8.34)$$

den Rang $(n+q)$ aufweist. Zunächst einmal werden die Spalten dieser Matrix vertauscht, was den Rang der Matrix nicht verändert:

$$\text{Rang} \begin{bmatrix} \lambda_i \mathbf{I}_n - \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \lambda_i \mathbf{I}_q & \mathbf{0} \end{bmatrix} = \text{Rang} \begin{bmatrix} \lambda_i \mathbf{I}_n - \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \lambda_i \mathbf{I}_q \end{bmatrix} \quad (8.35)$$

Nun soll der Rang dieser Matrix zunächst für den Eigenwert $\lambda_i = 0$ überprüft werden, wobei es keine Rolle spielt, ob dieser Eigenwert aus dem q -fachen Eigenwert bei 0 entsteht oder ob es sich um einen Eigenwert der Systemmatrix \mathbf{A} handelt. Einsetzen dieses Eigenwertes in die Matrix ergibt

$$\text{Rang} \begin{bmatrix} -\mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \text{Rang} \begin{bmatrix} -\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = \text{Rang} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} \quad (8.36)$$

Für den q -fachen Eigenwert $\lambda_i = 0$ hat also die letzte Matrix genau den gleichen Rang wie die Hautus-Matrix aus (8.34). Wenn diese Matrix den Rang $(n+q)$ aufweist, so gilt dies auch für die Hautus-Matrix, und das Hautus-Kriterium ist für den Eigenwert $\lambda_i = 0$ erfüllt.

Nun ist noch zu prüfen, wie sich die Hautus-Matrix für die übrigen Eigenwerte $\lambda_i \neq 0$ verhält. Wegen der Diagonalmatrix $\lambda_i \mathbf{I}_q$ sind die letzten q Zeilen der Hautus-Matrix mit Sicherheit linear unabhängig. Es bleibt die Frage, wie es um die lineare Unabhängigkeit der ersten n Zeilen steht. Diese sind genau dann linear unabhängig, wenn die Matrix

$$[\lambda_i \mathbf{I}_n - \mathbf{A}, \mathbf{B}] \quad (8.37)$$

ausschließlich linear unabhängige Zeilen enthält, also den Rang n aufweist. Dies ist aber gerade das Kriterium für die Steuerbarkeit des nicht erweiterten Systems (\mathbf{A}, \mathbf{B}) .

Zusammengefasst gilt damit: Durch die Forderung nach Steuerbarkeit des nicht erweiterten Systems (\mathbf{A}, \mathbf{B}) ist sichergestellt, dass die Hautus-Matrix des erweiterten Systems für alle Eigenwerte $\lambda_i \neq 0$ den Rang $(n+q)$ aufweist. Und durch die Forderung

$$\text{Rang} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} = n + q \quad (8.38)$$

wird dies auch für den q -fachen Eigenwert $\lambda_i = 0$ erfüllt.

Satz 8.1 *Das erweiterte System gemäß Abbildung 8.4 ist genau dann steuerbar, wenn das ursprüngliche System (\mathbf{A}, \mathbf{B}) steuerbar ist und die Matrix*

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix} \quad (8.39)$$

den Rang $(n+q)$ aufweist, wobei n die Ordnung des ursprünglichen Systems und q die Anzahl seiner Ausgangsgrößen ist.

Interessant ist eine Untersuchung der Dimension der Matrix in (8.39), sie ist $(n+q) \times (n+p)$, wobei q die Anzahl der Ausgangs- und p die Anzahl der Eingangsgrößen ist. Offensichtlich kann der Rang dieser Matrix nur dann $(n+q)$ betragen, wenn $p \geq q$ gilt, also die Anzahl der Eingangsgrößen mindestens genau so groß ist wie die Anzahl der Ausgangsgrößen.

8.3 Beobachter

In diesem Kapitel soll nun das zweite wesentliche Problem behandelt werden, das im Zusammenhang mit einer Zustandsregelung auftritt, nämlich die Frage, woher man eigentlich die aktuellen Werte der Zustandsgrößen als Eingangsgrößen der Rückführmatrix \mathbf{F} bekommt, wenn diese doch normalerweise nicht direkt gemessen werden können.

Die einfachste Lösung wäre sicherlich, den Zustandsvektor \mathbf{x} direkt aus dem - messbaren - Ausgangsvektor \mathbf{y} zu berechnen: $\mathbf{x} = \mathbf{C}^{-1}\mathbf{y}$. Wegen der normalerweise unterschiedlichen Anzahl von Ausgangs- und Zustandsgrößen ist \mathbf{C} aber im allgemeinen nicht quadratisch und somit nicht invertierbar. Diese Lösung kommt daher nicht in Frage.

Auf D.G.Luenberger ([38],[39]) geht die Idee zurück, den Zustandsvektor mit einem Streckenmodell zu schätzen. Dieses Streckenmodell wird parallel zur realen Strecke mitgerechnet und erhält dieselben Eingangsgrößen wie die Strecke (Abb. 8.6). Im Modell werden dann ein Zustandsvektor $\hat{\mathbf{x}}$ und ein Ausgangsvektor $\hat{\mathbf{y}}$ berechnet, die natürlich nicht unbedingt den realen Größen \mathbf{x} und \mathbf{y} entsprechen müssen. Die Abweichung zwischen dem Ausgangsvektor $\hat{\mathbf{y}}$ des Modells und dem Ausgangsvektor \mathbf{y} der Strecke wird deshalb zur Verbesserung der Schätzung als Korrekturterm über eine Matrix \mathbf{H} wieder in das Modell eingespeist.

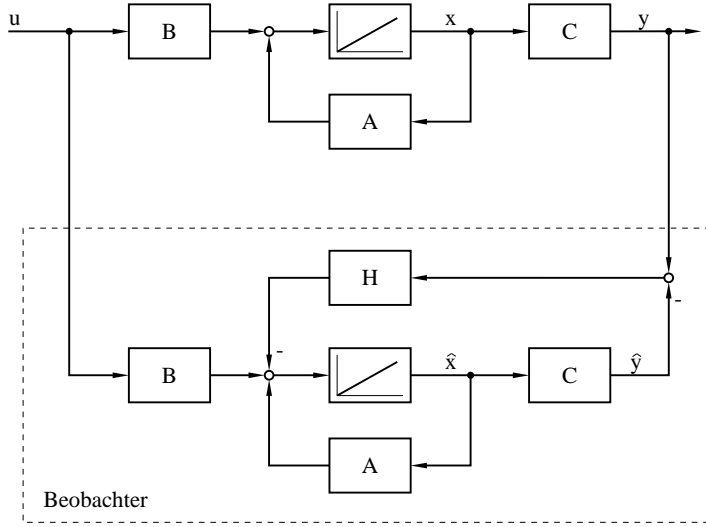


Abb. 8.6. Aufbau eines Beobachters

Aus dem Blockschaltbild lässt sich für den Schätzfehler $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$ ablesen:

$$\begin{aligned}
 \dot{\tilde{\mathbf{x}}} &= \dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} - [\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{H}\hat{\mathbf{y}} - \mathbf{H}\mathbf{y}] \\
 &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} - [\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{H}\mathbf{C}\hat{\mathbf{x}} - \mathbf{H}\mathbf{C}\mathbf{x}] \\
 &= [\mathbf{A} + \mathbf{H}\mathbf{C}](\mathbf{x} - \hat{\mathbf{x}}) = [\mathbf{A} + \mathbf{H}\mathbf{C}]\tilde{\mathbf{x}}
 \end{aligned} \tag{8.40}$$

Wenn die Matrix $(\mathbf{A} + \mathbf{H}\mathbf{C})$ nur Eigenwerte mit negativem Realteil aufweist, konvergiert der Schätzfehler im stationären Zustand gegen Null. Da \mathbf{A} und \mathbf{C} durch die Strecke vorgegeben sind, ist also eine geeignete Rückführmatrix \mathbf{H} zu finden.

Offensichtlich taucht beim Entwurf eines Beobachters ein ähnliches Problem auf wie beim Entwurf eines Zustandsreglers. Während beim Reglerentwurf eine Matrix \mathbf{F} so zu bestimmen ist, dass die Systemmatrix $(\mathbf{A} + \mathbf{B}\mathbf{F})$ ausschließlich Eigenwerte mit negativem Realteil aufweist, muss jetzt eine Matrix \mathbf{H} so bestimmt werden, dass dies für die Matrix $(\mathbf{A} + \mathbf{H}\mathbf{C})$ gilt. Da die Matrizen \mathbf{C} und \mathbf{H} gegenüber \mathbf{B} und \mathbf{F} in ihrer Reihenfolge aber vertauscht sind, ist das Problem nicht vollständig äquivalent. Dennoch lässt sich der Entwurf eines Beobachters auf den Entwurf eines Zustandsreglers zurückführen, wie die folgende Betrachtung zeigt.

Die Eigenwerte der Matrix $(\mathbf{A} + \mathbf{H}\mathbf{C})$ bestimmen die Stabilität des Beobachters und die Konvergenz des Schätzfehlers gegen Null. Sie sind definiert als Lösung der Gleichung

$$0 = \det[\lambda\mathbf{I} - (\mathbf{A} + \mathbf{H}\mathbf{C})] \tag{8.41}$$

Da sich eine Determinante durch Transposition nicht verändert, lässt sich diese Gleichung auch schreiben als

$$0 = \det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{H}\mathbf{C})]^T = \det[\lambda \mathbf{I} - \mathbf{A}^T - \mathbf{C}^T \mathbf{H}^T] \quad (8.42)$$

Die gleiche Betrachtung lässt sich auch für die Systemmatrix beim Reglerentwurf $(\mathbf{A} + \mathbf{B}\mathbf{F})$ durchführen. Hier berechnen sich die Eigenwerte aus

$$0 = \det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{F})] = \det[\lambda \mathbf{I} - \mathbf{A} - \mathbf{B}\mathbf{F}] \quad (8.43)$$

Ein Vergleich zeigt, dass die Problematik exakt dieselbe ist, wenn man folgendermaßen ersetzt:

$$\mathbf{A} \rightarrow \mathbf{A}^T \quad \mathbf{B} \rightarrow \mathbf{C}^T \quad \mathbf{F} \rightarrow \mathbf{H}^T \quad (8.44)$$

Der Entwurf einer Zustandsrückführung \mathbf{F} und einer Beobachter-Korrekturmatrix \mathbf{H} sind also ebenso dual zueinander wie Steuerbarkeit und Beobachtbarkeit (vgl. (7.169)). Von daher wird in den nachfolgenden Kapiteln jeweils der Entwurf eines Zustandsreglers ausführlich behandelt, während auf den Beobachterentwurf nur am Ende des jeweiligen Kapitels kurz eingegangen wird, da es ja im Prinzip das gleiche Verfahren ist.

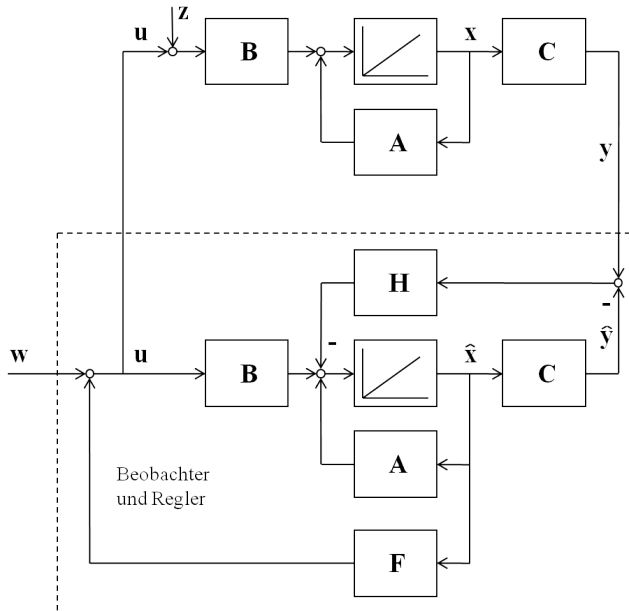


Abb. 8.7. Zustandsregelung mit einem Beobachter

Im Normalfall ergibt sich die in Abb. 8.7 gezeigte Gesamtstruktur aus Strecke, Beobachter und Regler. Beobachter und Regelung sind bei entsprechender Auslegung jeweils für sich genommen stabil. Es drängt sich aber die Frage auf, ob dies auch noch für das Gesamtsystem gilt, da beide gemeinsam in einem großen geschlossenen Kreis wirken.

Zur Beantwortung dieser Frage geht man von der Zustandsdarstellung des Gesamtsystems aus:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\hat{\mathbf{x}}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{BF} \\ -\mathbf{HC} & \mathbf{A} + \mathbf{BF} + \mathbf{HC} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \hat{\mathbf{x}} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{B} \end{bmatrix} \mathbf{w} \quad (8.45)$$

Um die Eigenwerte der Systemmatrix zu ermitteln, wird zunächst eine Basistransformation durchgeführt, die ja die Eigenwerte nicht verändert. Die hier verwendete Transformationsmatrix ist gleichzeitig ihre eigene Inverse:

$$\mathbf{T} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} = \mathbf{T}^{-1} \quad (8.46)$$

Damit ergibt sich für die transformierte Systemmatrix

$$\begin{aligned} \mathbf{T}^{-1} \mathbf{A}_{gesamt} \mathbf{T} &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{BF} \\ -\mathbf{HC} & \mathbf{A} + \mathbf{BF} + \mathbf{HC} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{BF} \\ \mathbf{A} + \mathbf{HC} & \mathbf{BF} - \mathbf{A} - \mathbf{BF} - \mathbf{HC} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{BF} \\ \mathbf{A} + \mathbf{HC} & -\mathbf{A} - \mathbf{HC} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} + \mathbf{BF} & -\mathbf{BF} \\ \mathbf{0} & \mathbf{A} + \mathbf{HC} \end{bmatrix} \end{aligned} \quad (8.47)$$

Die Eigenwerte dieser Matrix ergeben sich aus

$$\begin{aligned} 0 &= \det[\lambda \mathbf{I} - \mathbf{A}_{gesamt}] \\ &= \det[\lambda \mathbf{I} - \mathbf{T}^{-1} \mathbf{A}_{gesamt} \mathbf{T}] \\ &= \det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{BF})] \det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{HC})] \end{aligned} \quad (8.48)$$

Die Lösungen dieser Gleichung sind die Nullstellen von $\det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{BF})]$ sowie die Nullstellen von $\det[\lambda \mathbf{I} - (\mathbf{A} + \mathbf{HC})]$. Dies sind aber gerade die Eigenwerte von $(\mathbf{A} + \mathbf{BF})$ sowie die Eigenwerte von $(\mathbf{A} + \mathbf{HC})$. Daraus folgt das *Separationstheorem*:

Satz 8.2 *Gegeben sei ein System gemäß Abbildung 8.7, bestehend aus einer Strecke, einem Beobachter und einem Regler. Dann setzen sich die Eigenwerte des Gesamtsystems zusammen aus den Eigenwerten der Matrix $(\mathbf{A} + \mathbf{BF})$ sowie den Eigenwerten der Matrix $(\mathbf{A} + \mathbf{HC})$.*

Damit können Regler und Beobachter getrennt voneinander ausgelegt werden. Die Regler-Rückführmatrix \mathbf{F} wird so ausgelegt, dass die Systemmatrix $(\mathbf{A} + \mathbf{BF})$ ausschließlich Eigenwerte mit negativem Realteil enthält, und die Rückführmatrix \mathbf{H} des Beobachters wird so ausgelegt, dass dies für die Systemmatrix $(\mathbf{A} + \mathbf{HC})$ ebenfalls gilt. Damit ist sichergestellt, dass sämtliche Eigenwerte des Gesamtsystems einen negativen Realteil aufweisen und das Gesamtsystem somit stabil ist.

Voraussetzung für das Theorem ist allerdings, dass das Modell exakt mit der Strecke übereinstimmt, damit die Matrizen $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ im Beobachter auch tatsächlich den jeweiligen Matrizen in der Strecke entsprechen. Dies ist in der Praxis jedoch selten gegeben. Andererseits hat sich herausgestellt, dass auch bei ungenauem Modell eine weitgehend entkoppelte Auslegung von Beobachter und Regler möglich ist, ohne die Stabilität des Gesamtsystems zu gefährden.

8.4 Entwurf nach dem Polvorgabeverfahren

8.4.1 Herleitung

Ein Standardverfahren für den Entwurf eines Zustandsreglers für Eingrößensysteme ist das *Polvorgabeverfahren*, bei dem die Eigenwerte der Systemmatrix $(\mathbf{A} + \mathbf{B}\mathbf{F})$ vorgegeben werden und daraus die entsprechende Reglermatrix \mathbf{F} berechnet wird.

Zu berücksichtigen ist dabei, dass es sich, sofern eine PI-Reglerstruktur gemäß Abbildung 8.5 gewählt wird, bei \mathbf{A} und \mathbf{B} nicht um die Systemmatrizen der Strecke, sondern um die erweiterten Systemmatrizen gemäß Gleichung (8.27) handelt. Das Resultat des Polvorgabeverfahrens ist dann dementsprechend nicht die Rückkopplungsmatrix \mathbf{F} der Standard-Struktur (Abb. 8.1), sondern die erweiterte Matrix \mathbf{F}' gemäß Gleichung (8.30). Aus \mathbf{F}' müssen dann die einzelnen Matrizen der PI-Struktur $(\mathbf{P}, \mathbf{Q}, \mathbf{F})$ anschließend noch ermittelt werden. Die dazu notwendigen Schritte sind in Kapitel 8.2 aber schon diskutiert worden.

Für die Herleitung des Polvorgabeverfahrens sei zunächst angenommen, dass das System in Regelungsnormalform vorliegt (vgl. (7.38)):

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ \cdots \\ \cdots \\ \cdots \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [b_0 - b_n a_0, b_1 - b_n a_1, \dots, b_{n-1} - b_n a_{n-1}] \mathbf{x} + b_n u \quad (8.49)$$

Für dieses System soll nun eine Rückkopplungsmatrix \mathbf{F} entworfen werden, mit der sich die Systemmatrix $(\mathbf{A} + \mathbf{B}\mathbf{F})$ ergibt. Da wir von einem Eingrößensystem ausgehen, reduziert sich \mathbf{B} auf einen Spaltenvektor \mathbf{b} und \mathbf{F} auf einen Zeilenvektor \mathbf{f}^T . Für das Produkt $\mathbf{B}\mathbf{F}$ gilt

$$\mathbf{B}\mathbf{F} = \mathbf{b}\mathbf{f}^T = \begin{bmatrix} 0 \\ \cdots \\ 0 \\ 1 \end{bmatrix} [f_1, \dots, f_n] = \begin{bmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \\ f_1 & \cdots & f_n \end{bmatrix} \quad (8.50)$$

und damit

$$\mathbf{A} + \mathbf{B}\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ f_1 - a_0 & f_2 - a_1 & f_3 - a_2 & \cdots & f_{n-1} - a_{n-2} & f_n - a_{n-1} \end{bmatrix} \quad (8.51)$$

Die Eigenwerte $(\lambda_1, \dots, \lambda_n)$ des geschlossenen Kreises werden vorgegeben. Damit lassen sich auch die Koeffizienten p_i des charakteristischen Polynoms berechnen:

$$\det(s\mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{F})) = \prod_{i=1}^n (s - \lambda_i) = s^n + p_{n-1}s^{n-1} + \dots + p_1s + p_0 \quad (8.52)$$

Mit diesen Koeffizienten muss für die Systemmatrix des geschlossenen Kreises in Regelungsnormalform gelten:

$$\mathbf{A} + \mathbf{B}\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-2} & -p_{n-1} \end{bmatrix} \quad (8.53)$$

Der Vergleich von (8.51) und (8.53) liefert die Bestimmungsgleichungen für die n Koeffizienten der Rückkopplung \mathbf{f}^T :

$$\begin{aligned} f_1 - a_0 &= -p_0 & \rightarrow & f_1 = a_0 - p_0 \\ f_2 - a_1 &= -p_1 & \rightarrow & f_2 = a_1 - p_1 \\ & \dots & & \\ f_n - a_{n-1} &= -p_{n-1} & \rightarrow & f_n = a_{n-1} - p_{n-1} \end{aligned} \quad (8.54)$$

Die so bestimmte Rückkopplung \mathbf{f}^T stabilisiert den geschlossenen Kreis, so dass dessen Systemmatrix die vorgegebenen Eigenwerte λ_i aufweist.

Anhand eines Beispiels lässt sich zeigen, wie einfach und plausibel die hergeleitete Formel für die Rückführung ist. Gegeben sei eine Eingrößenstrecke dritter Ordnung in Regelungsnormalform (vgl. Abb. 7.5). Fügt man dieser Strecke die Rückführung \mathbf{f}^T hinzu, so erhält man die im Blockschaltbild 8.8 gezeigte Struktur. Man sieht, dass durch Hinzufügen der Rückführung jeder einzelne Streckenkoeffizient a_i eliminiert und durch den Koeffizienten p_i des gewünschten charakteristischen Polynoms des geschlossenen Kreises ersetzt wird. Diese besonders einfache Vorgehensweise resultiert aus der speziellen Struktur der Regelungsnormalform, die gerade deshalb ihren Namen zu Recht trägt.

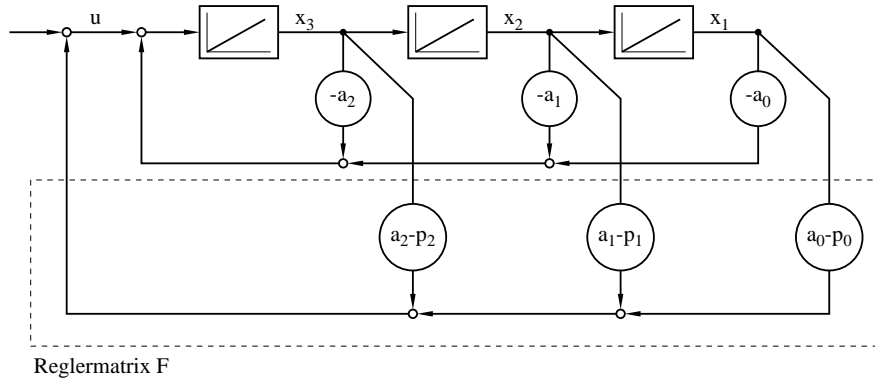


Abb. 8.8. Zustandsregelung einer Strecke in Regelungsnormform

Nun soll der Algorithmus für Systeme, die nicht in Regelungsnormform gegeben sind, verallgemeinert werden. Diese sind zunächst einer Basistransformation zu unterziehen, um die Regelungsnormform zu erhalten. Die entsprechenden Gleichungen (vgl. (7.61), (7.62) und (7.63)) lauten

$$\mathbf{T}\mathbf{z} = \mathbf{x} \quad \text{bzw.} \quad \mathbf{z} = \mathbf{T}^{-1}\mathbf{x} \quad (8.55)$$

und

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{z} + \mathbf{T}^{-1}\mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{T}\mathbf{z} + \mathbf{D}\mathbf{u} \end{aligned} \quad (8.56)$$

Wenn man für dieses System einen Zustandsregler entwirft, dann ergibt sich das obere Blockschaltbild in Abbildung 8.9. Es zeigt im Prinzip die Grundstruktur einer Zustandsregelung, wobei hier lediglich die Systemmatrizen \mathbf{A} , \mathbf{B} und \mathbf{C} durch ihre transformierten Matrizen ersetzt sind. Zu beachten ist, dass man auch den alten Zustandsvektor \mathbf{x} einzeichnen kann, nämlich gemäß (8.55) als Resultat der Multiplikation $\mathbf{T}\mathbf{z}$.

Für den Regler ergibt sich aus dem Blockschaltbild für $\mathbf{w} = \mathbf{0}$ das Regelgesetz

$$\mathbf{u} = \mathbf{F}_z \mathbf{z} \quad (8.57)$$

Für die Praxis ist dieses Regelgesetz jedoch nicht anwendbar, denn der z.B. von einem Beobachter zur Verfügung gestellte Zustandsvektor ist ja nicht der aus einer rein theoretischen Basistransformation hervorgegangene Zustandsvektor \mathbf{z} , sondern der nicht transformierte Original-Zustandsvektor \mathbf{x} .

Aus dem Grund muss ein Regelgesetz gefunden werden, das den Zustandsvektor \mathbf{x} als Eingangsgröße besitzt. Dieses erhält man durch Einsetzen von (8.55) in (8.57):

$$\mathbf{u} = \mathbf{F}_z \mathbf{z} = \mathbf{F}_z \mathbf{T}^{-1} \mathbf{x} = \mathbf{F}_x \mathbf{x} \quad (8.58)$$

Die Rückführmatrix $\mathbf{F}_x = \mathbf{F}_z \mathbf{T}^{-1}$ kann nun direkt für das nicht transformierte Originalsystem verwendet werden.

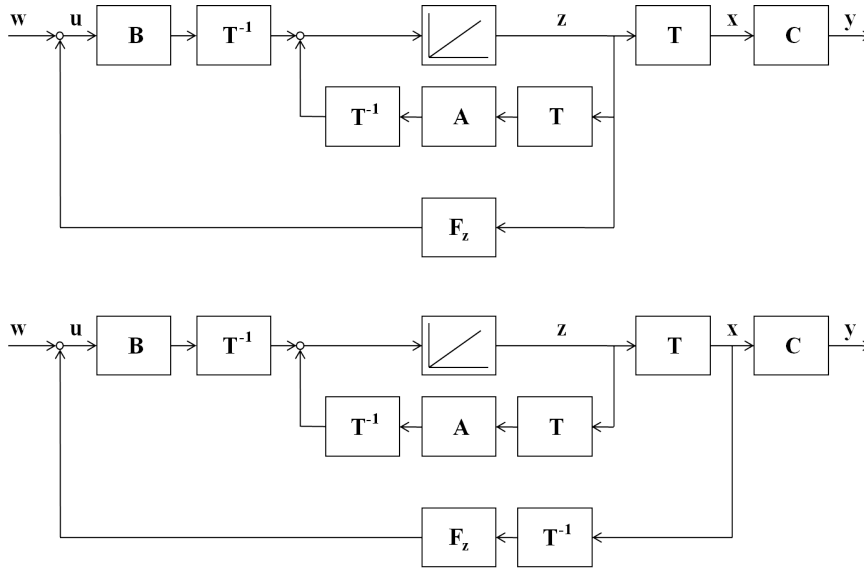


Abb. 8.9. Zustandsregelung für ein transformiertes System

Den Zusammenhang zwischen \mathbf{F}_x und \mathbf{F}_z zeigt das untere Blockschaltbild in Abbildung 8.9. Offensichtlich ist dieses Blockschaltbild zum oberen Blockschaltbild äquivalent. Der Vektor \mathbf{z} durchläuft die Matrizen \mathbf{T} und \mathbf{T}^{-1} , die sich in ihrer Wirkung gegenseitig aufheben, bevor er die für ihn gedachte Rückführung \mathbf{F}_z erreicht. Somit wird auch hier \mathbf{z} mit \mathbf{F}_z multipliziert. Diese Struktur kann man aber auch so interpretieren, dass der Original-Zustandsvektor \mathbf{x} mit der Rückführung $\mathbf{F}_z \mathbf{T}^{-1}$ auf die Strecke zurückgeführt wird. Dies entspricht der Grundstruktur einer Zustandsregelung (Abb. 8.1), wo ja auch der Zustandsvektor direkt vor der Ausgangsmatrix \mathbf{C} abgegriffen und zurückgeführt wird.

Zusammenfassend seien noch einmal die einzelnen Schritte zur Berechnung eines Zustandsreglers nach dem Polvorgabeverfahren für ein beliebiges, nicht in Regelungsnormalform gegebenes Eingrößensystem angegeben:

1. Das System ist auf Regelungsnormalform zu transformieren. Für ein steuerbares System ist dies immer möglich.
2. Aus der Systemmatrix der Regelungsnormalform lassen sich die Koeffizienten des charakteristischen Polynoms ablesen: $[a_0, a_1, \dots, a_{n-1}]$
3. Die Eigenwerte des geschlossenen Kreises sind festzulegen (Polvorgabe): $[\lambda_1, \lambda_2, \dots, \lambda_n]$
4. Mit Hilfe der Eigenwerte lassen sich aus (8.52) die Koeffizienten des charakteristischen Polynoms des geschlossenen Kreises berechnen: $[p_0, p_1, \dots, p_{n-1}]$
5. Für die Koeffizienten der Rückführung \mathbf{f}_z^T gilt: $f_i = a_{i-1} - p_{i-1}$. Die Reglermatrix besteht aus einem Zeilenvektor.

6. Den Zustandsregler für das nicht transformierte Originalsystem erhält man aus: $\mathbf{f}_x = \mathbf{f}_z \mathbf{T}^{-1}$

Diese Vorgehensweise lässt sich noch optimieren, wenn man auf Satz 7.1 zurückgreift. Für die Rückführung gilt mit (7.93)

$$\begin{aligned} \mathbf{F}_x &= \mathbf{f}_x^T = \mathbf{f}_z^T \mathbf{T}^{-1} = [a_0 - p_0, a_1 - p_1, \dots, a_{n-1} - p_{n-1}] \begin{bmatrix} \mathbf{t}_1^T \\ \mathbf{t}_1^T \mathbf{A} \\ \dots \\ \mathbf{t}_1^T \mathbf{A}^{n-1} \end{bmatrix} \\ &= (a_0 - p_0) \mathbf{t}_1^T + (a_1 - p_1) \mathbf{t}_1^T \mathbf{A} + \dots + (a_{n-1} - p_{n-1}) \mathbf{t}_1^T \mathbf{A}^{n-1} \\ &= -p_0 \mathbf{t}_1^T - p_1 \mathbf{t}_1^T \mathbf{A} - \dots - p_{n-1} \mathbf{t}_1^T \mathbf{A}^{n-1} \\ &\quad + a_0 \mathbf{t}_1^T + a_1 \mathbf{t}_1^T \mathbf{A} + \dots + a_{n-1} \mathbf{t}_1^T \mathbf{A}^{n-1} \end{aligned} \quad (8.59)$$

Weiterhin gilt gemäß Gleichung (7.81) für die Zeilen der inversen Transformationsmatrix (wobei hier \mathbf{h} durch \mathbf{t} ersetzt ist):

$$\begin{aligned} \mathbf{t}_2^T &= \mathbf{t}_1^T \mathbf{A} \\ \mathbf{t}_3^T &= \mathbf{t}_2^T \mathbf{A} = \mathbf{t}_1^T \mathbf{A}^2 \\ \mathbf{t}_4^T &= \mathbf{t}_3^T \mathbf{A} = \mathbf{t}_1^T \mathbf{A}^3 \\ \mathbf{t}_n^T &= \mathbf{t}_{n-1}^T \mathbf{A} = \mathbf{t}_1^T \mathbf{A}^{n-1} \\ \mathbf{t}_n^T \mathbf{A} &= -a_0 \mathbf{t}_1^T - a_1 \mathbf{t}_2^T - \dots - a_{n-1} \mathbf{t}_n^T \end{aligned} \quad (8.60)$$

Durch Einsetzen der ersten Gleichungen in die letzte wird daraus

$$\mathbf{t}_n^T \mathbf{A} = -a_0 \mathbf{t}_1^T - a_1 \mathbf{t}_1^T \mathbf{A} - \dots - a_{n-1} \mathbf{t}_1^T \mathbf{A}^{n-1} \quad (8.61)$$

Diese Gleichung wird in (8.59) eingesetzt:

$$\begin{aligned} \mathbf{f}_x^T &= -p_0 \mathbf{t}_1^T - p_1 \mathbf{t}_1^T \mathbf{A} - \dots - p_{n-1} \mathbf{t}_1^T \mathbf{A}^{n-1} - \mathbf{t}_n^T \mathbf{A} \\ &= -p_0 \mathbf{t}_1^T - p_1 \mathbf{t}_1^T \mathbf{A} - \dots - p_{n-1} \mathbf{t}_1^T \mathbf{A}^{n-1} - \mathbf{t}_1^T \mathbf{A}^n \\ &= -\mathbf{t}_1^T (p_0 + p_1 \mathbf{A} + \dots + p_{n-1} \mathbf{A}^{n-1} + \mathbf{A}^n) \end{aligned} \quad (8.62)$$

Dies ist die sogenannte *Ackermann-Formel* zur Berechnung einer Zustandsrückführung nach dem Polvorgabe-Verfahren. Der Algorithmus lautet damit wie folgt:

1. Die Eigenwerte des geschlossenen Kreises sind festzulegen (Polvorgabe): $[\lambda_1, \lambda_2, \dots, \lambda_n]$
2. Mit Hilfe der Eigenwerte lassen sich die Koeffizienten p_i des charakteristischen Polynoms des geschlossenen Kreises berechnen:

$$\prod_{i=1}^n (s - \lambda_i) = s^n + p_{n-1} s^{n-1} + \dots + p_1 s + p_0 \quad (8.63)$$

3. Die Steuerbarkeitsmatrix $\mathbf{Q}_s = [\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}]$ ist zu berechnen. Wenn diese Matrix regulär bzw. invertierbar ist, dann ist das System steuerbar.
4. Die letzte Zeile der Inversen der Steuerbarkeitsmatrix ist zu berechnen: $\mathbf{t}_1^T = [0, \dots, 0, 1]\mathbf{Q}_s^{-1}$
5. Die Rückführung ergibt sich aus der Ackermann-Formel:

$$\mathbf{f}_x^T = -\mathbf{t}_1^T(p_0 + p_1\mathbf{A} + \dots + p_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n) \quad (8.64)$$

Wenn man in diesem Algorithmus die Ersetzungen gemäß (8.44) vornimmt, so ergibt sich der folgende Algorithmus für den Entwurf eines Beobachters:

1. Die Eigenwerte der Beobachter-Systemmatrix $(\mathbf{A} + \mathbf{H}\mathbf{C})$ sind festzulegen (Polvorgabe): $[\lambda_1, \lambda_2, \dots, \lambda_n]$
2. Mit Hilfe der Eigenwerte lassen sich die Koeffizienten p_i des charakteristischen Polynoms für den Beobachter berechnen:

$$\prod_{i=1}^n (s - \lambda_i) = s^n + p_{n-1}s^{n-1} + \dots + p_1s + p_0 \quad (8.65)$$

3. Die Beobachtbarkeitsmatrix

$$\mathbf{Q}_b = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T \mathbf{A} \\ \mathbf{c}^T \mathbf{A}^2 \\ \dots \\ \mathbf{c}^T \mathbf{A}^{n-1} \end{bmatrix} \quad (8.66)$$

ist zu berechnen. Wenn diese Matrix regulär bzw. invertierbar ist, dann ist das System beobachtbar.

4. Die letzte Spalte der Inversen der Beobachtbarkeitsmatrix ist zu berechnen: $\mathbf{t}_1 = \mathbf{Q}_b^{-1}[0, \dots, 0, 1]^T$
5. Die Rückführung ergibt sich aus der Ackermann-Formel:

$$\mathbf{h}_x = -(p_0 + p_1\mathbf{A} + \dots + p_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n)\mathbf{t}_1 \quad (8.67)$$

Offensichtlich ist die Beobachter-Rückführung im Eingrößenfall ein Spaltenvektor der Dimension n , während die Regler-Rückführung ein Zeilenvektor der Dimension n ist. Dies ist sofort einsichtig, wenn man sich klar macht, dass beim Beobachter die skalare Differenz der Ausgangsgrößen zunächst mit \mathbf{h}_x multipliziert wird und dann in die n -dimensionale Ableitung des Zustandsvektors eingeht. Von daher muss \mathbf{h}_x ein Spaltenvektor der Dimension n sein.

8.4.2 Anmerkungen

Interessant ist beim Polvorgabe-Verfahren, dass die Eigenschaft der Steuerbarkeit bzw. Beobachtbarkeit so offensichtlich für dieses Verfahren benötigt

wird, da sonst die Inverse der Steuerbarkeits- bzw. Beobachtbarkeitsmatrix nicht gebildet werden könnte. Bei den nachfolgend vorgestellten Verfahren tritt diese Voraussetzung nicht so offen zutage, so dass man zur Vermutung gelangen könnte, dass nur das Polvorgabe-Verfahren diese Eigenschaft benötigt. Diese Vermutung ist jedoch nur bedingt richtig. Auch bei den nachfolgend vorgestellten Verfahren ist die Voraussetzung der Steuerbarkeit oder zumindest der Stabilisierbarkeit (Def. 8.9) unbedingt erforderlich, wenngleich sie dort im Algorithmus nicht so offen zutage tritt.

Wenn das System steuer- bzw. beobachtbar ist, dann können die Eigenwerte des geschlossenen Kreises beim Polvorgabeverfahren beliebig gewählt werden. Dies ist einerseits ein Vorteil, da man dem geschlossenen Kreis somit ein beliebiges Einschwingverhalten einprägen kann, andererseits aber auch ein Nachteil, da die Festlegung geeigneter Eigenwerte ein gewisses Maß an Intuition und Erfahrung erfordert, so dass bei diesem Verfahren normalerweise einige Versuche erforderlich sind, um einen der Problemstellung angepassten Regler zu finden.

Einige grundsätzliche Überlegungen mögen dabei hilfreich sein. Zunächst muss man sich über die Wirkung eines Eigenwertes im Klaren sein, die man anhand von Gleichung (7.120) sehr gut erläutern kann:

$$\mathbf{x}(t) = \sum_{\mu=1}^l e^{\lambda_{\mu} t} \sum_{\nu=1}^{r_{\mu}} \frac{t^{\nu-1}}{(\nu-1)!} \mathbf{K}_{\mu\nu} \mathbf{x}_0 \quad (8.68)$$

Bestimmend für die Konvergenzgeschwindigkeit eines Einschwingvorgangs jedes Terms ist seine Exponentialfunktion und daher der Eigenwert λ_i . Wenn dieser einen positiven Realteil hat, dann konvergiert der jeweilige Summand nicht gegen Null, und das System ist instabil. Bei einem negativen Realteil des Eigenwertes konvergiert der Summand gegen Null, und zwar umso schneller, je größer der Betrag des Realteiles ist.

Auf dieser Basis lassen sich nun Überlegungen hinsichtlich der optimalen Wahl der Eigenwerte anstellen. Die erste bezieht sich auf die Wahl der Eigenwerte für die Regelung. Wenn man die Eigenwerte des geschlossenen Kreises gleich den Eigenwerten der Strecke setzt, so bleibt offensichtlich für den Regler nichts mehr zu tun, denn er kann die Strecke praktisch sich selbst überlassen. Entsprechend klein werden die Amplituden der Stellgröße sein. Daraus folgt, dass diese Amplituden umso größer werden, je weiter die Eigenwerte des geschlossenen Kreises von denen der Strecke entfernt liegen, d.h. wenn man das System zu einem völlig anderen Verhalten zwingt, als es sich aus der Strecke selbst heraus ergeben würde.

Wenn man beispielsweise annimmt, dass zwei Eigenwerte der Strecke in der linken Hälfte der komplexen Halbebene bei $(-a \pm jb)$ liegen, so hat dies bei einer Veränderung der Eingangsgröße ein bestimmtes Einschwingverhalten zur Folge. Wenn man den Einschwingvorgang des geschlossenen Kreises beschleunigen und darüber hinaus die Schwingungen beseitigen will, so wird man die entsprechenden Eigenwerte des geschlossenen Kreises weiter nach

links auf den Wert $-c$ legen, wobei $c > a$ gilt. Je größer man c wählt, desto schneller wird der geschlossene Kreis seinen Endwert erreichen, aber diese Schnelligkeit muss mit großen Stellgrößenamplituden bezahlt werden. Und offenbar werden die Stellgrößenamplituden umso größer, je größer man c wählt. Ein Verschieben der Eigenwerte nach links hat also eine Beschleunigung der Einschwingvorgänge und damit des Regelkreises, aber eben auch eine Zunahme der Stellgrößenamplituden zur Folge. Darüber hinaus kann man dies bei einer realen Strecke sowieso nicht beliebig fortsetzen, da es für die Stellgrößenamplitude grundsätzlich praktische Beschränkungen gibt, die zu berücksichtigen sind.

Die zweite Überlegung bezieht sich auf den Beobachter. Ein Regler, dessen Eingangsgröße der geschätzte Zustandsvektor aus dem Beobachter ist, kann nur dann gute Ergebnisse erzielen, wenn dieser Schätzvektor möglichst genau dem wahren Zustandsvektor entspricht. Bevor die Schätzwerte des Beobachters nicht gegen die realen Werte konvergiert sind, kann die auf den Schätzwerten aufbauende Regelung zu keinem guten Ergebnis kommen. Von daher sollten die Eigenwerte des Beobachters betragsmäßig größere Realteile aufweisen (also weiter links liegen) als die Eigenwerte des Regelkreises. Denn dann laufen die Einschwingvorgänge im Beobachter schneller ab als diejenigen im Regelkreis, und genau das ist ja auch notwendig.

8.5 Falb-Wolovich-Entwurf

Beim Falb-Wolovich-Entwurf wird das Mehrgrößensystem durch geschickte Entkopplung in voneinander unabhängige Eingrößensysteme zerlegt, deren Eingrößen-Regler dann jeder für sich unabhängig voneinander ausgelegt werden können.

8.5.1 Differenzengrad

Um die hierzu notwendigen Schritte verstehen zu können, muss vorab der Begriff des *Differenzengrades* erläutert werden. Hierzu sei zunächst ein Eingrößensystem mit der Übertragungsfunktion

$$G(s) = \frac{y(s)}{u(s)} = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0} \quad (8.69)$$

gegeben, wobei immer $m \leq n$ gilt. Der Differenzengrad ist für ein derartiges System gerade die Graddifferenz von Zähler- und Nennerpolynom $\delta = n - m$. Offensichtlich gilt immer $0 \leq \delta \leq n$.

Um diese Definition für Mehrgrößensysteme weiter entwickeln zu können, soll eine Grenzwertbetrachtung durchgeführt werden:

$$\begin{aligned}
\lim_{s \rightarrow \infty} s^k G(s) &= \lim_{s \rightarrow \infty} \frac{b_m s^{m+k} + \dots + b_1 s^{1+k} + b_0 s^k}{a_n s^n + \dots + a_1 s + a_0} \\
&= \lim_{s \rightarrow \infty} \frac{b_m \frac{s^k}{s^{n-m}} + \dots + b_1 \frac{s^k}{s^{n-1}} + b_0 \frac{s^k}{s^n}}{a_n + \dots + a_1 \frac{1}{s^{n-1}} + a_0 \frac{1}{s^n}} \\
&= \begin{cases} 0 & \text{für } k < n - m = \delta \\ \frac{b_m}{a_n} & \text{für } k = n - m = \delta \\ \infty & \text{für } k > n - m = \delta \end{cases} \quad (8.70)
\end{aligned}$$

Der Differenzengrad lässt sich also auch wie folgt definieren:

Definition 8.3 *Derjenige Wert von k , für den der Grenzwert $\lim_{s \rightarrow \infty} s^k G(s)$ einen endlichen, von Null verschiedenen Wert liefert, ist der Differenzengrad δ des durch $G(s)$ gegebenen Systems.*

Im folgenden Schritt soll diese Definition auf ein SISO-System, welches in Zustandsdarstellung gegeben und dessen Durchgangsmatrix \mathbf{D} gleich Null ist, angewendet werden. Das System ist jetzt gegeben durch seine Zustandsgleichungen

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u \\
y &= \mathbf{c}^T \mathbf{x}
\end{aligned} \quad (8.71)$$

Mit (7.107) gilt hier

$$G(s) = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \quad (8.72)$$

Weiterhin gilt mit (7.104) und der Definition der Matrizen-Exponentialfunktion (A.15) sowie unter Zuhilfenahme der Korrespondenztabelle in Anhang A.2

$$\begin{aligned}
(s\mathbf{I} - \mathbf{A})^{-1} &= \mathcal{L} \{ e^{\mathbf{A}t} \} \\
&= \mathcal{L} \left\{ \sum_{\nu=0}^{\infty} \mathbf{A}^{\nu} \frac{t^{\nu}}{\nu!} \right\} \\
&= \sum_{\nu=0}^{\infty} \mathbf{A}^{\nu} \mathcal{L} \left\{ \frac{t^{\nu}}{\nu!} \right\} = \sum_{\nu=0}^{\infty} \mathbf{A}^{\nu} \frac{1}{s^{\nu+1}} \quad (8.73)
\end{aligned}$$

und damit für die Übertragungsfunktion

$$G(s) = \sum_{\nu=0}^{\infty} \mathbf{c}^T \mathbf{A}^{\nu} \mathbf{b} \frac{1}{s^{\nu+1}} \quad (8.74)$$

Für diesen Ausdruck wird nun wieder der Grenzwert wie in (8.70) gebildet, jetzt aber von vornherein mit $k = \delta$:

$$\lim_{s \rightarrow \infty} s^{\delta} G(s) = \lim_{s \rightarrow \infty} \sum_{\nu=0}^{\infty} \mathbf{c}^T \mathbf{A}^{\nu} \mathbf{b} \frac{1}{s^{\nu+1-\delta}} \quad (8.75)$$

Für $\nu > \delta - 1$ sind alle Summanden der Reihe offensichtlich Null, so dass man auch schreiben kann

$$\begin{aligned} \lim_{s \rightarrow \infty} s^\delta G(s) &= \lim_{s \rightarrow \infty} \sum_{\nu=0}^{\delta-1} \mathbf{c}^T \mathbf{A}^\nu \mathbf{b} \frac{1}{s^{\nu+1-\delta}} \\ &= \lim_{s \rightarrow \infty} (s^{\delta-1} \mathbf{c}^T \mathbf{b} + s^{\delta-2} \mathbf{c}^T \mathbf{A} \mathbf{b} + \dots + \mathbf{c}^T \mathbf{A}^{\delta-1} \mathbf{b}) \end{aligned} \quad (8.76)$$

Dieser Ausdruck muss nun gemäß Def. 8.3 einen endlichen, von Null verschiedenen Wert liefern. Da s gegen ∞ konvergiert, kann der Gesamtausdruck nur dann endlich werden, wenn alle Terme, in denen s vorkommt, verschwinden:

$$\begin{aligned} \mathbf{c}^T \mathbf{b} &= 0 \\ \mathbf{c}^T \mathbf{A} \mathbf{b} &= 0 \\ \mathbf{c}^T \mathbf{A}^2 \mathbf{b} &= 0 \\ &\dots \\ \mathbf{c}^T \mathbf{A}^{\delta-2} \mathbf{b} &= 0 \\ \mathbf{c}^T \mathbf{A}^{\delta-1} \mathbf{b} &\neq 0 \end{aligned} \quad (8.77)$$

Dieses Gleichungssystem stellt eine weitere Definition für den Differenzengrad dar:

Definition 8.4 Der Differenzengrad δ eines SISO-Systems ist dadurch definiert, dass er das Gleichungssystem (8.77) erfüllt.

Mit diesem Ergebnis kann nun zu MIMO-Systemen übergegangen werden, die durch die allgemeine Form der linearen Zustandsgleichungen beschrieben werden, wobei auch hier wieder von $\mathbf{D} = \mathbf{0}$ ausgegangen wird:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{aligned} \quad (8.78)$$

\mathbf{b}_k sei die k -te Spalte von \mathbf{B} , die mit der k -ten Komponente u_k des Stellgrößenvektors \mathbf{u} multipliziert wird und demnach die Wirkung dieser Komponente auf das System beschreibt. Analog dazu sei \mathbf{c}_i^T die i -te Zeile von \mathbf{C} , die die Ausgangsgleichung für die i -te Komponente von \mathbf{y} beschreibt: $y_i = \mathbf{c}_i^T \mathbf{x}$. Damit folgt aus (8.72) für die Übertragungsfunktion G_{ik} von u_k nach y_i :

$$G_{ik}(s) = \mathbf{c}_i^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b}_k \quad (8.79)$$

Der Differenzengrad dieser Übertragungsfunktion sei δ_{ik} , und er ist gemäß Def. 8.4 durch das Gleichungssystem

$$\begin{aligned} \mathbf{c}_i^T \mathbf{b}_k &= 0 \\ \mathbf{c}_i^T \mathbf{A} \mathbf{b}_k &= 0 \\ \mathbf{c}_i^T \mathbf{A}^2 \mathbf{b}_k &= 0 \end{aligned}$$

$$\begin{aligned}
& \dots \\
& \mathbf{c}_i^T \mathbf{A}^{\delta_{ik}-2} \mathbf{b}_k = \mathbf{0} \\
& \mathbf{c}_i^T \mathbf{A}^{\delta_{ik}-1} \mathbf{b}_k \neq \mathbf{0}
\end{aligned} \tag{8.80}$$

definiert. δ_{ik} ist der Differenzengrad des Übertragungsverhaltens von u_k nach y_i . Daher erscheint es durchaus plausibel, zur Berechnung des Differenzengrades δ_i des Übertragungsverhaltens vom gesamten Eingangsvektor \mathbf{u} nach y_i den Vektor \mathbf{b}_k durch die gesamte Eingangsmatrix \mathbf{B} zu ersetzen. Für δ_i ergibt sich damit das Gleichungssystem

$$\begin{aligned}
& \mathbf{c}_i^T \mathbf{B} = \mathbf{0} \\
& \mathbf{c}_i^T \mathbf{A} \mathbf{B} = \mathbf{0} \\
& \mathbf{c}_i^T \mathbf{A}^2 \mathbf{B} = \mathbf{0} \\
& \dots \\
& \mathbf{c}_i^T \mathbf{A}^{\delta_i-2} \mathbf{B} = \mathbf{0} \\
& \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \neq \mathbf{0}
\end{aligned} \tag{8.81}$$

Wie stehen nun δ_i und die verschiedenen δ_{ik} (für $1 \leq k \leq p$) (p ist die Anzahl der Eingangsgrößen) zueinander? Fest steht, dass ein Ausdruck $\mathbf{c}_i^T \mathbf{A}^\nu \mathbf{B}$ auf der linken Seite der Gleichungen nur dann Null ergeben kann, wenn alle Ausdrücke $\mathbf{c}_i^T \mathbf{A}^\nu \mathbf{b}_k$ (für $1 \leq k \leq p$) ebenfalls Null ergeben. Für den kleinsten Differenzengrad $\nu = \min_k \delta_{ik}$ wird für eben dieses k der Ausdruck $\mathbf{c}_i^T \mathbf{A}^\nu \mathbf{b}_k$ ungleich Null und damit auch $\mathbf{c}_i^T \mathbf{A}^\nu \mathbf{B}$. Damit gilt:

Definition 8.5 Bei einem MIMO-System ist der Differenzengrad δ_i des Übertragungsverhaltens vom Eingangsvektor \mathbf{u} zur Ausgangsgröße y_i gegeben durch

$$\delta_i := \min(\delta_{i1}, \dots, \delta_{ip}) \tag{8.82}$$

Dabei sind die δ_{ik} die Differenzengrade der Übertragungsfunktionen G_{ik} von der Eingangsgröße u_k zur Ausgangsgröße y_i (lt. Def. 8.4), und p ist die Anzahl der Eingangsgrößen.

Darüber hinaus lässt sich auch ein Differenzengrad des Gesamtsystems angeben, dessen Bedeutung erst am Ende dieses Kapitels deutlich wird:

Definition 8.6 Der Differenzengrad δ eines MIMO-Systems ist gegeben durch

$$\delta := \sum_{i=1}^q \delta_i \tag{8.83}$$

Dabei sind die δ_i gegeben durch Def. 8.5, und q ist die Anzahl der Ausgangsgrößen.

Auf Basis der Differenzengrade kann man nun das dynamische Verhalten des Systems in einer anderen Form angeben. Ausgangspunkt der Überlegungen ist die Ausgangsgleichung für eine Ausgangsgröße

$$y_i = \mathbf{c}_i^T \mathbf{x} \quad (8.84)$$

Für die Ableitung dieser Gleichung gilt mit Einsetzen der Zustandsgleichung

$$\dot{y}_i = \mathbf{c}_i^T \dot{\mathbf{x}} = \mathbf{c}_i^T \mathbf{A} \mathbf{x} + \mathbf{c}_i^T \mathbf{B} \mathbf{u} \quad (8.85)$$

Für $\delta_i > 1$ gilt gemäß (8.81) $\mathbf{c}_i^T \mathbf{B} = \mathbf{0}^T$, und es bleibt

$$\dot{y}_i = \mathbf{c}_i^T \mathbf{A} \mathbf{x} \quad (8.86)$$

Erneutes Ableiten liefert mit den gleichen Rechenschritten

$$\ddot{y}_i = \mathbf{c}_i^T \mathbf{A} \dot{\mathbf{x}} = \mathbf{c}_i^T \mathbf{A}^2 \mathbf{x} + \mathbf{c}_i^T \mathbf{A} \mathbf{B} \mathbf{u} \quad (8.87)$$

Für $\delta_i > 2$ gilt gemäß (8.81) $\mathbf{c}_i^T \mathbf{A} \mathbf{B} = \mathbf{0}^T$, und es bleibt

$$\ddot{y}_i = \mathbf{c}_i^T \mathbf{A}^2 \mathbf{x} \quad (8.88)$$

Dieses Vorgehen lässt sich bis zum Differenzengrad δ_i fortsetzen, so dass man das folgende Gleichungssystem erhält:

$$\begin{aligned} y_i &= \mathbf{c}_i^T \mathbf{x} \\ \dot{y}_i &= \mathbf{c}_i^T \mathbf{A} \mathbf{x} \\ \ddot{y}_i &= \mathbf{c}_i^T \mathbf{A}^2 \mathbf{x} \\ &\dots \\ \frac{d}{dt^{\delta_i-1}} y_i &= \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{x} \\ \frac{d}{dt^{\delta_i}} y_i &= \mathbf{c}_i^T \mathbf{A}^{\delta_i} \mathbf{x} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{u} \end{aligned} \quad (8.89)$$

Dieses Gleichungssystem gestattet eine neue Interpretation des Differenzengrades: δ_i ist die niedrigste Ableitung der Ausgangsgröße y_i , auf die der Eingangsvektor \mathbf{u} einen direkten Einfluss hat. Auf alle niedrigeren Ableitungen von y_i hat \mathbf{u} nur einen indirekten Einfluss über den Zustandsvektor.

Nun wird ein neuer, künstlicher Ausgangsvektor definiert:

$$\tilde{\mathbf{y}} = \begin{bmatrix} \frac{d}{dt^{\delta_1}} y_1 \\ \frac{d}{dt^{\delta_2}} y_2 \\ \dots \\ \frac{d}{dt^{\delta_q}} y_q \end{bmatrix} \quad (8.90)$$

Mit (8.89) ergibt sich für diesen neuen Ausgangsvektor die Zustandsdarstellung

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \tilde{\mathbf{y}} &= \tilde{\mathbf{C}}\mathbf{x} + \tilde{\mathbf{D}}\mathbf{u}\end{aligned}\quad (8.91)$$

mit

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1} \\ \mathbf{c}_2^T \mathbf{A}^{\delta_2} \\ \dots \\ \mathbf{c}_q^T \mathbf{A}^{\delta_q} \end{bmatrix} \quad \text{und} \quad \tilde{\mathbf{D}} = \begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1-1} \mathbf{B} \\ \mathbf{c}_2^T \mathbf{A}^{\delta_2-1} \mathbf{B} \\ \dots \\ \mathbf{c}_q^T \mathbf{A}^{\delta_q-1} \mathbf{B} \end{bmatrix} \quad (8.92)$$

8.5.2 Regler-Entwurf

Der Entwurf nach Falb-Wolovich basiert auf einer gewöhnlichen Regelungsstruktur nach dem Blockschaltbild in Abbildung 8.2. Im Gegensatz zum Ansatz in Kapitel 8.2 wird hier jedoch nicht erst die Rückführmatrix \mathbf{F} im Hinblick auf Stabilität und anschließend die Filtermatrix \mathbf{R} im Hinblick auf stationäre Genauigkeit ausgelegt, sondern beide Matrizen werden zunächst dazu verwendet, um das Mehrgrößensystem zu entkoppeln, so dass jede Ausgangsgröße y_i ausschließlich durch ihren Sollwert w_i beeinflusst wird. Erst dann werden die restlichen Freiheitsgrade bei der Wahl der Koeffizienten von \mathbf{F} und \mathbf{R} dazu genutzt, dass jede Ausgangsgröße y_i ihrem Sollwert w_i möglichst schnell und genau folgt. Voraussetzung ist natürlich, dass \mathbf{w} und \mathbf{y} die gleiche Dimension haben ($p = q$).

Mit $\mathbf{u} = \mathbf{R}\mathbf{w} + \mathbf{F}\mathbf{x}$ folgt aus (8.90) bis (8.92) für jede Ausgangsgröße

$$\begin{aligned}\frac{d}{dt^{\delta_i}} y_i &= \mathbf{c}_i^T \mathbf{A}^{\delta_i} \mathbf{x} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} [\mathbf{R}\mathbf{w} + \mathbf{F}\mathbf{x}] \\ &= [\mathbf{c}_i^T \mathbf{A}^{\delta_i} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B}\mathbf{F}] \mathbf{x} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B}\mathbf{R}\mathbf{w}\end{aligned}\quad (8.93)$$

\mathbf{R} wird dann so gewählt, dass für jedes i gilt:

$$\mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B}\mathbf{R}\mathbf{w} = [0, \dots, 0, k_i, 0, \dots, 0] \mathbf{w} = k_i w_i \quad (8.94)$$

Dabei ist k_i eine beliebige, reelle Zahl. Damit ist erst einmal sichergestellt, dass der Sollwert w_i der einzige ist, der einen direkten Einfluss auf y_i hat.

Schreibt man die Bedingung (8.94) für alle i in Matrizenform, so ergibt sich

$$\begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1-1} \mathbf{B} \\ \mathbf{c}_2^T \mathbf{A}^{\delta_2-1} \mathbf{B} \\ \dots \\ \mathbf{c}_q^T \mathbf{A}^{\delta_q-1} \mathbf{B} \end{bmatrix} \mathbf{R}\mathbf{w} = \begin{bmatrix} k_1 & & \\ & k_2 & \\ & & \dots \\ & & & k_q \end{bmatrix} \mathbf{w} \quad (8.95)$$

bzw.

$$\tilde{\mathbf{D}}\mathbf{R}\mathbf{w} = \text{diag}(k_1, \dots, k_q) \mathbf{w} \quad (8.96)$$

und damit für \mathbf{R} die Bestimmungsgleichung

$$\mathbf{R} = \tilde{\mathbf{D}}^{-1} \text{diag}(k_1, \dots, k_q) \quad (8.97)$$

Eine Bedingung für dieses Verfahren ist damit die Invertierbarkeit von $\tilde{\mathbf{D}}$ bzw. $\det(\tilde{\mathbf{D}}) \neq 0$.

Sofern diese Bedingung erfüllt ist und \mathbf{R} entsprechend gewählt wurde, wird aus Gleichung (8.93)

$$\frac{d}{dt^{\delta_i}} y_i = [\mathbf{c}_i^T \mathbf{A}^{\delta_i} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{F}] \mathbf{x} + k_i w_i \quad (8.98)$$

Der einzige Sollwert, der in dieser Gleichung auftaucht und damit einen direkten Einfluss auf y_i hat, ist w_i . Die anderen Sollwerte können sich aber noch über den Zustandsvektor \mathbf{x} auf y_i auswirken. Um auch diese Quereinflüsse noch zu eliminieren, soll diese Gleichung nun durch geeignete Wahl von \mathbf{F} auf die Form

$$\frac{d}{dt^{\delta_i}} y_i + q_{i,\delta_i-1} \frac{d}{dt^{\delta_i-1}} y_i + \dots + q_{i,1} \frac{d}{dt} y_i + q_{i,0} y_i = k_i w_i \quad (8.99)$$

gebracht werden, also auf eine gewöhnliche Eingrößen-Differentialgleichung der Ordnung δ_i bzw. auf die Übertragungsfunktion

$$\frac{y_i}{w_i} = \frac{k_i}{s^{\delta_i} + q_{i,\delta_i-1} s^{\delta_i-1} + \dots + q_{i,1} s + q_{i,0}} \quad (8.100)$$

Daraus folgt für ein i die Bedingung

$$[\mathbf{c}_i^T \mathbf{A}^{\delta_i} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{F}] \mathbf{x} = - \sum_{\nu=0}^{\delta_i-1} q_{i,\nu} \frac{d}{dt^{\nu}} y_i \quad (8.101)$$

und nach Einsetzen der Beziehungen aus (8.89)

$$[\mathbf{c}_i^T \mathbf{A}^{\delta_i} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{F}] \mathbf{x} = - \sum_{\nu=0}^{\delta_i-1} q_{i,\nu} \mathbf{c}_i^T \mathbf{A}^{\nu} \mathbf{x} \quad (8.102)$$

\mathbf{x} kann eliminiert werden, und es verbleibt die Beziehung

$$\mathbf{c}_i^T \mathbf{A}^{\delta_i} + \mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{F} = - \sum_{\nu=0}^{\delta_i-1} q_{i,\nu} \mathbf{c}_i^T \mathbf{A}^{\nu} \quad (8.103)$$

Umstellen liefert

$$\mathbf{c}_i^T \mathbf{A}^{\delta_i-1} \mathbf{B} \mathbf{F} = -\mathbf{c}_i^T \mathbf{A}^{\delta_i} - \sum_{\nu=0}^{\delta_i-1} q_{i,\nu} \mathbf{c}_i^T \mathbf{A}^{\nu} \quad (8.104)$$

Fasst man diese Gleichung für alle i in Matrizenschreibweise zusammen, so erhält man

$$\tilde{\mathbf{D}}\mathbf{F} = \begin{bmatrix} -\mathbf{c}_1^T \mathbf{A}^{\delta_1} - \sum_{\nu=0}^{\delta_1-1} q_{1,\nu} \mathbf{c}_1^T \mathbf{A}^\nu \\ \vdots \\ -\mathbf{c}_p^T \mathbf{A}^{\delta_p} - \sum_{\nu=0}^{\delta_p-1} q_{p,\nu} \mathbf{c}_p^T \mathbf{A}^\nu \end{bmatrix} \quad (8.105)$$

woraus sich die gesuchte Bestimmungsgleichung für die Rückkopplungsmatrix \mathbf{F} ergibt:

$$\mathbf{F} = -\tilde{\mathbf{D}}^{-1} \begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1} + \sum_{\nu=0}^{\delta_1-1} q_{1,\nu} \mathbf{c}_1^T \mathbf{A}^\nu \\ \vdots \\ \mathbf{c}_p^T \mathbf{A}^{\delta_p} + \sum_{\nu=0}^{\delta_p-1} q_{p,\nu} \mathbf{c}_p^T \mathbf{A}^\nu \end{bmatrix} \quad (8.106)$$

Offensichtlich ist auch für die Berechnung von \mathbf{F} die Invertierbarkeit von $\tilde{\mathbf{D}}$ die Voraussetzung.

Zusammengefasst ergibt sich der folgende Satz:

Satz 8.7 Gegeben sei die Strecke

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \quad \text{bzw.} \quad y_i = \mathbf{c}_i^T \mathbf{x} \end{aligned} \quad (8.107)$$

mit der gleichen Anzahl an Ein- und Ausgangsgrößen $p = q$. Der Regler für diese Strecke sei gegeben durch

$$\mathbf{u} = \mathbf{F}\mathbf{x} + \mathbf{R}\mathbf{w} \quad (8.108)$$

mit dem Sollvektor \mathbf{w} .

Das System kann entkoppelt werden, wenn die Entkopplungsbedingung

$$\det \tilde{\mathbf{D}} = \det \begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1-1} \mathbf{B} \\ \mathbf{c}_2^T \mathbf{A}^{\delta_2-1} \mathbf{B} \\ \vdots \\ \mathbf{c}_q^T \mathbf{A}^{\delta_q-1} \mathbf{B} \end{bmatrix} \neq 0 \quad (8.109)$$

erfüllt ist. In dem Fall lassen sich die Regler-Matrizen berechnen aus

$$\begin{aligned} \mathbf{R} &= \tilde{\mathbf{D}}^{-1} \text{diag}(k_1, \dots, k_q) \\ \mathbf{F} &= -\tilde{\mathbf{D}}^{-1} \begin{bmatrix} \mathbf{c}_1^T \mathbf{A}^{\delta_1} + \sum_{\nu=0}^{\delta_1-1} q_{1,\nu} \mathbf{c}_1^T \mathbf{A}^\nu \\ \vdots \\ \mathbf{c}_p^T \mathbf{A}^{\delta_p} + \sum_{\nu=0}^{\delta_p-1} q_{p,\nu} \mathbf{c}_p^T \mathbf{A}^\nu \end{bmatrix} \end{aligned} \quad (8.110)$$

und es ergibt sich ein entkoppeltes System, in dem jede Ausgangsgröße gemäß

$$y_i(s) = \frac{k_i}{s^{\delta_i} + q_{i,\delta_i-1}s^{\delta_i-1} + \dots + q_{i,1}s + q_{i,0}} w_i(s) = T_i(s)w_i(s) \quad (8.111)$$

von der jeweiligen Eingangsgröße abhängig ist.

Die freien Parameter bei diesem Entwurfsverfahren sind die k_i und die $q_{i,\nu}$. Mit diesen Parametern kann die Übertragungsfunktion in Gleichung (8.111) beliebig gestaltet werden.

Da die $q_{i,\nu}$ die Koeffizienten des charakteristischen Polynoms darstellen, kann man hier wie beim Polvorgabeverfahren zunächst die Pole der Übertragungsfunktion vorgeben und damit nach Gleichung (8.63) die Koeffizienten $q_{i,\nu}$ berechnen. Anschließend ist ein Wert für k_i zu wählen. Dieser sollte so festgelegt werden, dass die Übertragungsfunktion im stationären Fall den Wert Eins aufweist, damit dann $y_i = w_i$ gilt:

$$\lim_{s \rightarrow 0} T_i(s) = 1 \quad (8.112)$$

Daraus folgt für k_i :

$$k_i = q_{i,0} \quad (8.113)$$

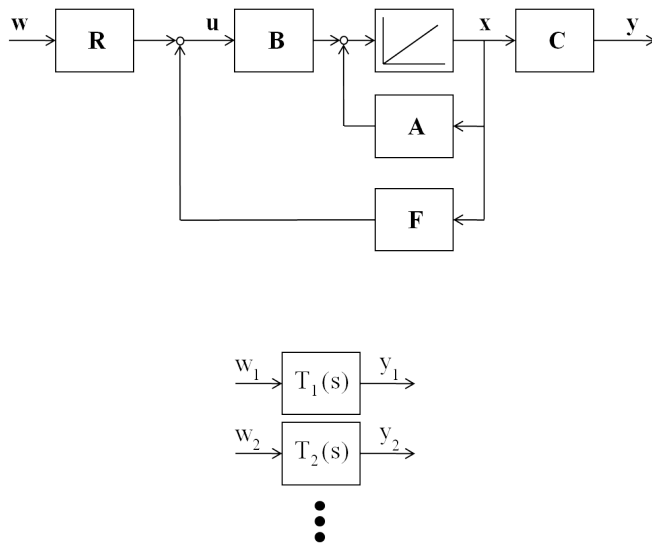


Abb. 8.10. Prinzip des Falb-Wolovich-Entwurfs

Abbildung 8.10 verdeutlicht noch einmal die Grundidee des Falb-Wolovich-Entwurfs: Im Prinzip besteht die Regelung aus einer Rückkopplungsmatrix **F** und einer Referenz-Filtermatrix **R**. Durch Wahl der Parameter dieser beiden Matrizen gemäß Satz 8.7 entsteht aber ein System, das sich genau so verhält wie das System in der unteren Hälfte der Abbildung, also wie eine Parallelschaltung aus p Eingrößensystemen.

Die freie Wahl der Polstellen in diesen Eingrößensystemen führt nun zu einer interessanten Überlegung. Offensichtlich können im i -ten Eingrößensystem gemäß Gleichung (8.111) δ_i Polstellen frei vorgegeben werden. Für das Gesamtsystem sind dies $\sum \delta_i$ frei wählbare Polstellen, was nach Definition 8.6 gerade dem Gesamt-Differenzengrad δ des MIMO-Systems entspricht.

Andererseits ergibt sich aus (7.123), dass als Polstellen nur Eigenwerte des Systems auftreten können. Durch eine Regelung können zwar die Werte der Eigenwerte festgelegt werden, was ja hier auch der Fall ist, nicht jedoch deren Anzahl. Ein System n -ter Ordnung hat, ob geregelt oder nicht, maximal n verschiedene Eigenwerte und damit auch maximal n verschiedene Polstellen. Daraus folgt $\delta \leq n$.

Wenn der maximal mögliche Wert n für δ nicht erreicht wird, dann bedeutet dies, dass ein oder mehrere Eigenwerte offenbar durch Nullstellen kompensiert werden.

In der Theorie hat dies keine Konsequenzen. Ein kompensierter Eigenwert hat dort keine Auswirkungen mehr auf das Übertragungsverhalten des Systems. In der Praxis muss man aber davon ausgehen, dass das Entwurfsverfahren auf einem nicht perfekten Streckenmodell basiert und dementsprechend auch die implizite Polstellenkompensation nicht perfekt ist. Der Eigenwert, der beim Entwurf in $T_i(s)$ überhaupt nicht zu sehen und daher durch die $q_{i,\nu}$ auch nicht zu beeinflussen ist, ist in der Praxis aufgrund mangelhafter Kompensation wahrscheinlich dennoch vorhanden.

Hier gibt es nun zwei Möglichkeiten, die im Prinzip auch schon in Kapitel 6.2.3 erläutert worden sind. Wenn der kompensierte Eigenwert einen negativen Realteil aufweist und damit einen stabilen Einschwingvorgang repräsentiert, so hat eine unvollständige Kompensation praktisch keine Auswirkungen. Nullstelle und Eigenwert liegen nun zwar nicht exakt an der gleichen Stelle, aber immer noch relativ dicht beieinander, so dass sich ihre Wirkungen auf den Phasen- und Amplitudenverlauf der Übertragungsfunktion immer noch nahezu kompensieren und damit die Stabilität des geschlossenen Kreises nicht gefährden.

Bei einem Eigenwert mit positivem Realteil sind die Verhältnisse jedoch nicht mehr so einfach. Ein derartiger Eigenwert, der dann ja wider Erwarten als Polstelle in der Kreisübertragungsfunktion des realen, entkoppelten Eingrößensystems auftaucht, verändert die laut Nyquist-Kriterium erforderliche Phasendrehung über den Parameter r_k in Gleichung (5.19), und die unter der Voraussetzung einer erfolgreichen Kompensation entstandene Kreisübertragungsfunktion aus Strecke und Regler kann diese Phasendrehung natürlich nicht aufweisen. Der Regelkreis wird damit instabil, und der Reglerentwurf ist gescheitert.

Daher ist es nach Durchführung des Falb-Wolovich-Entwurfs für den Fall $\delta < n$ erforderlich, sämtliche Eigenwerte des geschlossenen Kreises zu ermitteln und zu prüfen, ob diese alle in der negativen komplexen Halbebene liegen. Falls dies nicht der Fall ist, muss der Regler nach einem anderen Verfahren entworfen werden.

In Kapitel 9.4.2 wird diese Problematik noch einmal aufgegriffen und auch eine einfachere Methode zur Überprüfung der Stabilität des geschlossenen Regelkreises gezeigt.

8.6 Riccati-Entwurf

8.6.1 Herleitung

Dem *Riccati-Entwurf* ([44]) liegt die Idee zugrunde, einen *optimalen* Regler zu finden. Dabei ist zunächst zu definieren, was man unter *optimal* versteht. Hier soll dies derjenige Regler sein, der den von außen nicht angeregten geschlossenen Kreis aus einem beliebigen Anfangszustand unter Minimierung des Funktional

$$J = \int_0^{\infty} (\mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) + \mathbf{u}^T(t) \mathbf{R} \mathbf{u}(t)) dt \quad (8.114)$$

in den Ruhezustand überführt. Wenn es gelingt, den Regler so auszulegen, dass $J < \infty$ gilt, so folgt daraus, dass der Integrand gegen Null konvergieren muss, d.h. sowohl \mathbf{x} als auch \mathbf{u} müssen gegen Null konvergieren. Mit der Konvergenz von \mathbf{x} ist dann aber auch Stabilität im Sinne von Ljapunov implizit gewährleistet.

Das Funktional wird dann klein, wenn einerseits der Zustandsvektor schnell gegen den Nullvektor konvergiert und dies andererseits mit kleinen Stellgrößen erreicht wird. \mathbf{Q} und \mathbf{R} stellen im Prinzip Gewichtungsfaktoren für den Verlauf von Stell- und Zustandsgrößen dar. Wenn man mehr Wert auf eine schnelle Konvergenz des Zustandsvektors gegen den Ruhezustand legt, so wird man den Verlauf des Zustandsvektors höher gewichten, im Prinzip also "bestrafen", und die Koeffizienten von \mathbf{Q} entsprechend größer wählen als die Koeffizienten von \mathbf{R} . Legt man dagegen mehr Wert auf kleine Stellgrößen und damit geringeren Energieverbrauch bei der Regelung, so sollten die Koeffizienten von \mathbf{R} größer gewählt werden.

Dabei können die Komponenten von Zustands- und Stellgrößenvektor sogar einzeln gewichtet werden. Wählt man für \mathbf{Q} und \mathbf{R} Diagonalmatrizen, was aus Gründen der Übersichtlichkeit im Normalfall auch so gemacht wird, so erhält man als Gütefunktional

$$J = \int_0^{\infty} \left(\sum_{i=1}^n q_{ii} x_i^2 + \sum_{i=1}^p r_{ii} u_i^2 \right) dt \quad (8.115)$$

d.h. das i -te Diagonalelement einer Matrix gewichtet gerade die i -te Komponente des entsprechenden Vektors. Dies ist eine sehr interessante Möglichkeit, die der Riccati-Entwurf bietet. Man kann einzelne Stell- oder Zustandsgrößen "bestrafen" und damit klein halten.

Es ist offensichtlich, dass die Wahl von \mathbf{R} und \mathbf{Q} entscheidenden Einfluss auf den resultierenden Regler hat. Die Freiheitsgrade, die beim Polvorgabeverfahren darin bestanden, die Pole bzw. Eigenwerte des geschlossenen Kreises frei vorgeben zu können, bestehen jetzt darin, den Verlauf der einzelnen Komponenten von \mathbf{x} und \mathbf{u} unterschiedlich stark gewichten zu können.

Völlig frei können \mathbf{R} und \mathbf{Q} jedoch nicht gewählt werden. Sie sollen symmetrisch und positiv definit sein, und zwar aus folgendem Grund: Positive Definitheit von \mathbf{Q} bedeutet, dass das Produkt $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ für beliebige Vektoren $\mathbf{x} \neq \mathbf{0}$ immer größer als Null ist. Da dies auch für Vektoren \mathbf{x} gilt, bei denen nur eine einzige Komponente von Null verschieden ist, folgt daraus, dass jede Komponente jedes beliebigen Vektors \mathbf{x} einen positiven Beitrag zum Produkt $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ leistet. Genau dies ist aber auch gewollt. Schließlich soll der Verlauf aller Zustandsgrößen zu jedem Zeitpunkt in das Gütefunktional eingehen, wobei der Wert von J umso größer werden soll, je größer die Werte der Zustandsgrößen sind. Analog dazu ergibt sich die Forderung nach positiver Definitheit von \mathbf{R} , denn sämtliche Stellgrößen sollen natürlich ebenfalls in das Funktional eingehen.

Im nachfolgenden Kapitel 8.6.2 wird gezeigt, dass sich die Forderung nach positiver Definitheit von \mathbf{Q} auch abschwächen lässt, doch soll für die nun folgende Herleitung und auch für den abschließend formulierten Satz von positiver Definitheit ausgegangen werden, da diese Forderung in der Praxis relativ einfach zu überprüfen und auch einzuhalten ist.

Die Forderung nach positiver Definitheit hat zudem den angenehmen Nebeneffekt, dass eine positiv definite Matrix invertierbar ist. Denn ihre sämtlichen Eigenwerte sind positiv und damit ebenso ihre Determinante, woraus direkt die Invertierbarkeit folgt. Diese Eigenschaft wird in der nachfolgenden Herleitung für \mathbf{R} noch ausgenutzt.

Für die Herleitung wird zunächst von der unregelmäßigen Strecke ohne äußere Anregung ausgegangen. Gleichung (8.114) reduziert sich damit auf

$$J = \int_0^{\infty} \mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) dt \quad (8.116)$$

Einsetzen der allgemeinen Lösung der homogenen Zustandsgleichung (7.95) mit dem Anfangszeitpunkt $t = 0$ liefert

$$\begin{aligned} J &= \int_0^{\infty} \mathbf{x}^T(0) e^{\mathbf{A}^T t} \mathbf{Q} e^{\mathbf{A} t} \mathbf{x}(0) dt \\ &= \mathbf{x}^T(0) \int_0^{\infty} e^{\mathbf{A}^T t} \mathbf{Q} e^{\mathbf{A} t} dt \mathbf{x}(0) \\ &= \mathbf{x}^T(0) \mathbf{P} \mathbf{x}(0) \end{aligned} \quad (8.117)$$

mit

$$\mathbf{P} = \int_0^{\infty} e^{\mathbf{A}^T t} \mathbf{Q} e^{\mathbf{A} t} dt \quad (8.118)$$

Unter Zuhilfenahme des Schemas der partiellen Integration

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx \quad (8.119)$$

und Gleichung (A.25) im Anhang lässt sich Gleichung (8.118) umformen zu

$$\mathbf{P} = e^{\mathbf{A}^T t} \mathbf{Q} \mathbf{A}^{-1} e^{\mathbf{A} t} \Big|_0^\infty - \int_0^\infty \mathbf{A}^T e^{\mathbf{A}^T t} \mathbf{Q} \mathbf{A}^{-1} e^{\mathbf{A} t} dt \quad (8.120)$$

wobei noch die Frage nach der Existenz der inversen Systemmatrix \mathbf{A}^{-1} auftaucht. Hierzu wird jedoch vorausgesetzt, dass das System stabil ist, denn dann hat \mathbf{A} nur Eigenwerte mit negativem Realteil, womit die Determinante von \mathbf{A} ungleich Null ist und daher \mathbf{A} invertierbar. Aufgrund der angenommenen Stabilität von \mathbf{A} folgt auch, dass der Systemzustand aus jedem Anfangszustand gegen Null konvergiert

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \lim_{t \rightarrow \infty} e^{\mathbf{A} t} \mathbf{x}(0) = \mathbf{0} \quad (8.121)$$

und damit

$$\lim_{t \rightarrow \infty} e^{\mathbf{A} t} = \mathbf{0} \quad (8.122)$$

gilt. Deshalb verschwindet der erste Term in Gleichung (8.120) für die obere Integrationsgrenze, und es bleibt

$$\mathbf{P} = -\mathbf{Q} \mathbf{A}^{-1} - \int_0^\infty \mathbf{A}^T e^{\mathbf{A}^T t} \mathbf{Q} \mathbf{A}^{-1} e^{\mathbf{A} t} dt \quad (8.123)$$

Auch das Integral lässt sich unter Verwendung von (A.28) noch weiter umformen, und man erhält

$$\begin{aligned} \mathbf{P} &= -\mathbf{Q} \mathbf{A}^{-1} - \mathbf{A}^T \int_0^\infty e^{\mathbf{A}^T t} \mathbf{Q} e^{\mathbf{A} t} dt \mathbf{A}^{-1} \\ &= -\mathbf{Q} \mathbf{A}^{-1} - \mathbf{A}^T \mathbf{P} \mathbf{A}^{-1} \end{aligned} \quad (8.124)$$

Multiplikation mit \mathbf{A} von rechts und Umstellen liefert die *Ljapunov-Gleichung*

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q} \quad (8.125)$$

aus der sich \mathbf{P} bei gegebenem \mathbf{Q} und bekannter Systemmatrix \mathbf{A} berechnen lässt. Aufgrund der symmetrischen Struktur der Gleichung muss \mathbf{P} auf jeden Fall symmetrisch sein, und bei einem stabilen System \mathbf{A} ist \mathbf{P} darüber hinaus sogar positiv definit (Beweis siehe Anhang A.5).

Nachdem dieses Ergebnis für das unregelte System gefunden wurde, soll nun die Herleitung für den geschlossenen Kreis mit Regelung, jedoch weiterhin ohne äußere Anregung $\mathbf{w} = \mathbf{0}$ erfolgen. Dabei gilt hier das gleiche,

was auch schon zu Beginn des letzten Kapitels gesagt wurde: Sofern eine PI-Reglerstruktur gewählt wird, sind im weiteren Verlauf für \mathbf{A} und \mathbf{B} nicht die Systemmatrizen der Strecke, sondern die erweiterten Systemmatrizen gemäß Gleichung (8.27) anzusetzen. Und entsprechend handelt es sich beim Regler nicht um die Rückkopplungsmatrix \mathbf{F} der Standard-Struktur (Abb. 8.1), sondern um die erweiterte Matrix \mathbf{F}' gemäß Gleichung (8.30), aus der dann die einzelnen Matrizen der PI-Struktur $(\mathbf{P}, \mathbf{Q}, \mathbf{F})$ nach dem Riccati-Entwurf noch ermittelt werden müssen. Die dazu notwendigen Schritte sind in Kapitel 8.2 schon diskutiert worden. Der Einfachheit halber soll im Folgenden aber nur von \mathbf{A} , \mathbf{B} und \mathbf{F} gesprochen werden.

Die Systemmatrix des geschlossenen Kreises lautet $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{B}\mathbf{F}$, und das Regelgesetz, welches den Zustandsregler definiert, $\mathbf{u} = \mathbf{F}\mathbf{x}$. Damit ergibt sich für das vollständige Gütefunktional nach (8.114)

$$\begin{aligned}
 J &= \int_0^{\infty} (\mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) + \mathbf{u}^T(t) \mathbf{R} \mathbf{u}(t)) dt \\
 &= \int_0^{\infty} (\mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) + \mathbf{x}^T(t) \mathbf{F}^T \mathbf{R} \mathbf{F} \mathbf{x}(t)) dt \\
 &= \int_0^{\infty} \mathbf{x}^T(t) (\mathbf{Q} + \mathbf{F}^T \mathbf{R} \mathbf{F}) \mathbf{x}(t) dt \\
 &= \int_0^{\infty} \mathbf{x}^T(t) \tilde{\mathbf{Q}} \mathbf{x}(t) dt
 \end{aligned} \tag{8.126}$$

mit

$$\tilde{\mathbf{Q}} = \mathbf{Q} + \mathbf{F}^T \mathbf{R} \mathbf{F} \tag{8.127}$$

Damit ist das Gütefunktional für den geregelten Kreis genauso definiert wie für die ungeregelte Strecke (8.116), wobei hier lediglich \mathbf{Q} durch $\tilde{\mathbf{Q}}$ ersetzt ist. Nach den gleichen Rechenschritten ergibt sich auch hier eine Matrix \mathbf{P} , mit der

$$J = \mathbf{x}^T(0) \mathbf{P} \mathbf{x}(0) \tag{8.128}$$

gilt und die als Lösung aus der Gleichung

$$\tilde{\mathbf{A}}^T \mathbf{P} + \mathbf{P} \tilde{\mathbf{A}} = -\tilde{\mathbf{Q}} \tag{8.129}$$

hervorgeht. Die Rechnung weist jetzt sogar den Vorteil auf, dass für die Systemmatrix $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{B}\mathbf{F}$ die Stabilität nicht wie oben vorausgesetzt werden muss, sondern dass sie hier sichergestellt ist, da es sich um die Systemmatrix des geregelten Systems handelt.

Da die unbekannte Regler-Rückführung \mathbf{F} sowohl in $\tilde{\mathbf{A}}$ als auch in $\tilde{\mathbf{Q}}$ auftaucht, lässt sich \mathbf{P} aber aus dieser Gleichung nicht direkt berechnen.

Daher ist zunächst ein Ansatz für \mathbf{F} zu finden, den man hier einsetzen kann, um \mathbf{F} zu eliminieren.

Hierzu soll ausgehend von Gleichung (8.128) das Minimum der Gütefunktion J gefunden werden. Die notwendige Bedingung ist, dass die Ableitung von J nach jeder Komponente f_{ij} der Reglermatrix \mathbf{F} Null sein muss:

$$0 = \frac{\delta J}{\delta f_{ij}} \quad \text{für alle } i, j \quad (8.130)$$

Daraus wird mit (8.128)

$$0 = \mathbf{x}^T(0) \frac{\delta \mathbf{P}}{\delta f_{ij}} \mathbf{x}(0) \quad \text{für alle } i, j \quad (8.131)$$

Da diese Bedingung für alle Anfangszustände $\mathbf{x}(0)$ gelten muss, folgt die Bedingung

$$\frac{\delta \mathbf{P}}{\delta f_{ij}} = \mathbf{0} \quad \text{für alle } i, j \quad (8.132)$$

Nun soll die Ljapunov-Gleichung (8.129) nach den Elementen von \mathbf{F} abgeleitet werden. Für jedes Wertepaar (i, j) ergibt sich nach der Produktregel

$$\frac{\delta \tilde{\mathbf{A}}^T}{\delta f_{ij}} \mathbf{P} + \tilde{\mathbf{A}}^T \frac{\delta \mathbf{P}}{\delta f_{ij}} + \frac{\delta \mathbf{P}}{\delta f_{ij}} \tilde{\mathbf{A}} + \mathbf{P} \frac{\delta \tilde{\mathbf{A}}}{\delta f_{ij}} = -\frac{\delta \tilde{\mathbf{Q}}}{\delta f_{ij}} \quad (8.133)$$

Im Optimum gilt (8.132), d.h. die Ableitungen von \mathbf{P} verschwinden, und es bleibt

$$\frac{\delta \tilde{\mathbf{A}}^T}{\delta f_{ij}} \mathbf{P} + \mathbf{P} \frac{\delta \tilde{\mathbf{A}}}{\delta f_{ij}} = -\frac{\delta \tilde{\mathbf{Q}}}{\delta f_{ij}} \quad (8.134)$$

Nun sollen die Ableitungen von $\tilde{\mathbf{A}}$ und $\tilde{\mathbf{Q}}$ näher untersucht werden. Einsetzen der Systemmatrizen liefert für die Ableitung von $\tilde{\mathbf{A}}$

$$\frac{\delta \tilde{\mathbf{A}}}{\delta f_{ij}} = \frac{\delta}{\delta f_{ij}} (\mathbf{A} + \mathbf{B}\mathbf{F}) = \mathbf{B} \frac{\delta \mathbf{F}}{\delta f_{ij}} \quad (8.135)$$

und für die Ableitung von $\tilde{\mathbf{Q}}$ mit (8.127)

$$\frac{\delta \tilde{\mathbf{Q}}}{\delta f_{ij}} = \frac{\delta}{\delta f_{ij}} (\mathbf{Q} + \mathbf{F}^T \mathbf{R} \mathbf{F}) = \frac{\delta \mathbf{F}^T}{\delta f_{ij}} \mathbf{R} \mathbf{F} + \mathbf{F}^T \mathbf{R} \frac{\delta \mathbf{F}}{\delta f_{ij}} \quad (8.136)$$

Setzt man diese Resultate in (8.134) ein, so erhält man

$$\frac{\delta \mathbf{F}^T}{\delta f_{ij}} \mathbf{B}^T \mathbf{P} + \mathbf{P} \mathbf{B} \frac{\delta \mathbf{F}}{\delta f_{ij}} = -\frac{\delta \mathbf{F}^T}{\delta f_{ij}} \mathbf{R} \mathbf{F} - \mathbf{F}^T \mathbf{R} \frac{\delta \mathbf{F}}{\delta f_{ij}} \quad (8.137)$$

Umstellen liefert

$$\begin{aligned} \frac{\delta \mathbf{F}^T}{\delta f_{ij}} (\mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F}) + (\mathbf{P} \mathbf{B} + \mathbf{F}^T \mathbf{R}) \frac{\delta \mathbf{F}}{\delta f_{ij}} &= \mathbf{0} \\ \frac{\delta \mathbf{F}^T}{\delta f_{ij}} (\mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F}) + (\mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F})^T \frac{\delta \mathbf{F}}{\delta f_{ij}} &= \mathbf{0} \end{aligned} \quad (8.138)$$

Hierbei wurde die Tatsache ausgenutzt, dass \mathbf{R} eine symmetrische Matrix ist und daher $\mathbf{R} = \mathbf{R}^T$ gilt.

Als nächstes sollen die Ableitungen von \mathbf{F} bzw. \mathbf{F}^T analysiert werden. Offensichtlich bleibt bei der Ableitung von \mathbf{F} nach einem der Koeffizienten f_{ij} genau an der Stelle (i, j) eine 1, während an allen anderen Stellen die Ableitung Null liefert:

$$\frac{\delta \mathbf{F}}{\delta f_{ij}} = i \begin{bmatrix} & j \\ & 1 \\ & \end{bmatrix} \quad (8.139)$$

Entsprechend ergibt sich für die Ableitung der Matrix \mathbf{F}^T die 1 an der Stelle (j, i) .

Für die weiteren Überlegungen wird die Matrix

$$\mathbf{M} := \mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F} \quad (8.140)$$

mit den Elementen m_{ij} definiert. Aus Gleichung (8.138) wird damit

$$\frac{\delta \mathbf{F}^T}{\delta f_{ij}} \mathbf{M} + \mathbf{M}^T \frac{\delta \mathbf{F}}{\delta f_{ij}} = \mathbf{0} \quad (8.141)$$

Da die Ableitungen von \mathbf{F} bzw. \mathbf{F}^T jeweils nur ein einziges von Null verschiedenes Element haben, ergibt sich für die beiden Matrizenprodukte

$$\begin{aligned} \begin{bmatrix} m_{i1} & m_{i2} & m_{i3} & \cdots & \cdots \end{bmatrix} + \begin{bmatrix} m_{i1} \\ m_{i2} \\ m_{i3} \\ \cdots \\ \cdots \end{bmatrix} &= \mathbf{0} \\ \begin{bmatrix} m_{i1} & m_{i2} & \cdots & 2m_{ij} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} &= \mathbf{0} \end{aligned} \quad (8.142)$$

Das erste Produkt liefert eine Matrix, deren j -te Zeile der i -ten Zeile von \mathbf{M} entspricht, während das zweite Produkt eine Matrix liefert, deren j -te Spalte der i -ten Spalte von \mathbf{M}^T entspricht, wobei sich in dieser Spalte wegen der

Transponieren von \mathbf{M} auch hier die Elemente der i -ten Zeile von \mathbf{M} wiederfinden. Die Gesamtsumme liefert demnach eine Matrix, in der genau die j -te Zeile und die j -te Spalte von Null verschieden sind und an der Stelle (j, j) das Element m_{ij} zweifach auftaucht.

Diese Matrix muss nun gleich der Nullmatrix sein. Daraus folgt sofort, dass alle Koeffizienten der i -ten Zeile von \mathbf{M} und damit von $\mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F}$ gleich Null sein müssen. Und da die Bedingung (8.130) im Optimum für alle Wertepaare (i, j) und damit für alle Indizes i gelten muss, ergibt sich, dass sämtliche Koeffizienten dieser Matrix gleich Null sein müssen:

$$\mathbf{B}^T \mathbf{P} + \mathbf{R} \mathbf{F} = \mathbf{0} \quad (8.143)$$

Für die Reglermatrix erhält man daraus die Bestimmungsgleichung

$$\mathbf{F} = -\mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} \quad (8.144)$$

Die Invertierbarkeit von \mathbf{R} ist durch deren positive Definitheit sichergestellt, wie anfangs schon erläutert wurde.

Da \mathbf{R} bei der Definition des Gütefunktional vorgegeben wird und \mathbf{B} als Eingangsmatrix der Strecke bekannt ist, sind die beiden Unbekannten in dieser Gleichung \mathbf{F} und \mathbf{P} . Zusammen mit der Ljapunov-Gleichung (8.129) für das geregelte System hat man nun aber zwei Gleichungen für diese beiden Unbekannten.

Zur Lösung dieses Gleichungssystems wird (8.144) in (8.129) eingesetzt und die Symmetrie von \mathbf{P} und \mathbf{R} ausgenutzt ($\mathbf{P} = \mathbf{P}^T$ und $\mathbf{R} = \mathbf{R}^T$):

$$\tilde{\mathbf{A}}^T \mathbf{P} + \mathbf{P} \tilde{\mathbf{A}} = -\tilde{\mathbf{Q}} \quad (8.145)$$

$$(\mathbf{A} + \mathbf{B} \mathbf{F})^T \mathbf{P} + \mathbf{P} (\mathbf{A} + \mathbf{B} \mathbf{F}) = -(\mathbf{Q} + \mathbf{F}^T \mathbf{R} \mathbf{F}) \quad (8.146)$$

$$(\mathbf{A} - \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P})^T \mathbf{P} + \mathbf{P} (\mathbf{A} - \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P}) = -(\mathbf{Q} + \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{R} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P}) \quad (8.147)$$

$$(\mathbf{A}^T - \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T) \mathbf{P} + \mathbf{P} (\mathbf{A} - \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P}) = -\mathbf{Q} - \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} \quad (8.148)$$

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} - \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} = -\mathbf{Q} \quad (8.149)$$

Dies ist die sogenannte *Matrix-Riccati-Gleichung*, aus der \mathbf{P} bestimmt werden kann, wobei \mathbf{P} immer positiv definit ist. Da es sich hier um ein Standardproblem der Regelungstechnik handelt, liegen z.B. in MATLAB leistungsfähige Algorithmen vor, die dieses Problem lösen können. Voraussetzung für die Lösbarkeit dieser Gleichung ist die positive Semidefinitheit von \mathbf{Q} , aber da oben bereits positive Definitheit von \mathbf{Q} gefordert wurde, ist dies sicher erfüllt.

Ein Punkt ist jedoch noch offen: Die Bedingung (8.130) ist nur notwendig, nicht jedoch hinreichend dafür, dass J einen minimalen Wert annimmt, denn die Gradienten einer Funktion können sowohl im Minimum als auch im Maximum der Funktion Null sein. Es ist also noch sicherzustellen, dass es sich im vorliegenden Fall tatsächlich um ein Minimum und nicht um ein Maximum handelt.

Hierzu reicht aber eine einfache Plausibilitätsbetrachtung aus: Grundsätzlich besteht immer die Möglichkeit, die Koeffizienten des Reglers f_{ij} noch etwas “schlechter“ zu wählen, d.h. den Verlauf von \mathbf{x} und \mathbf{u} immer noch etwas ungünstiger zu gestalten, so dass \mathbf{x} und \mathbf{u} und damit auch J immer größere Werte annehmen, und sei es, indem die Vorzeichen der f_{ij} so gewählt werden, dass der Regelkreis instabil wird, und man dann den Betrag der Reglerkoeffizienten immer weiter wachsen lässt. Daraus folgt aber, dass ein endliches Maximum für J gar nicht auftreten kann. Von daher muss es sich im obigen Fall um ein Minimum von J handeln.

Damit kann nun der entsprechende Satz formuliert werden:

Satz 8.8 *Riccati-Regler: Gegeben seien ein System in Zustandsdarstellung*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (8.150)$$

sowie eine Gütefunktion

$$J = \int_0^{\infty} (\mathbf{x}^T(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}\mathbf{u}(t)) dt \quad (8.151)$$

wobei die Matrizen \mathbf{R} und \mathbf{Q} symmetrisch und positiv definit zu wählen sind. Dann ist J minimal für einen Regler der Form

$$\mathbf{u} = \mathbf{F}\mathbf{x}(t) \quad (8.152)$$

mit

$$\mathbf{F} = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} \quad (8.153)$$

wobei \mathbf{P} die symmetrische, positiv definite Lösung der Matrix-Riccati-Gleichung

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} = -\mathbf{Q} \quad (8.154)$$

ist.

8.6.2 Anmerkungen

Eine relativ einfache Betrachtung zeigt, dass die Gleichung (8.153) für den Regler durchaus sinnvoll ist. Wenn man die Koeffizienten von \mathbf{R} nämlich groß wählt, um große Stellgrößen zu “bestrafen“, dann werden die Koeffizienten von \mathbf{F} wegen der Abhängigkeit von \mathbf{R}^{-1} umso kleiner. Und dies ist genau das, was erforderlich ist, um kleine Stellgrößen zu erhalten.

Als nächster Punkt soll die Steuerbarkeit als Voraussetzung für den Reglerentwurf diskutiert werden. Hierzu ist zunächst der Begriff der *Stabilisierbarkeit* einzuführen:

Definition 8.9 *Eine Strecke wird als stabilisierbar bezeichnet, wenn ein Regler existiert, mit dem der geschlossene Kreis stabil im Sinne von Ljapunov ist.*

Ein steuerbares System ist damit immer auch stabilisierbar, da ja eine Regelung hinzugefügt werden kann, die den geschlossenen Kreis stabil macht. Aber auch ein nicht steuerbares System kann stabilisierbar sein, wenn nämlich die nicht steuerbaren Zustandsgrößen auch ohne Regelung gegen Null konvergieren.

Für den Riccati-Regler gilt nun, dass die Stabilisierbarkeit der Strecke Voraussetzung für die Existenz einer Lösung ist. Wäre das System nämlich nicht stabilisierbar, so würde mindestens eine Zustandsgröße existieren, die nicht mit t gegen Null konvergiert. Dann könnte aber auch das Funktional J keinen endlichen Wert mehr annehmen, und eine Optimierung wäre nicht mehr möglich. Aus demselben Grund ist der geschlossene Kreis mit dem gefundenen Regler sicher stabil, d.h. alle Zustandsgrößen konvergieren gegen Null, denn sonst könnte J ebenfalls keinen endlichen Wert aufweisen.

Als nächstes soll die zu Beginn von Kapitel 8.6.1 bereits angesprochene, mögliche Abschwächung der Forderung nach positiver Definitheit von \mathbf{Q} diskutiert werden. Wie dort bereits erläutert wurde, wird durch die Forderung nach positiver Definitheit sichergestellt, dass sämtliche Zustandsgrößen zu jedem Zeitpunkt einen Beitrag zum Term $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ und damit zum Gütefunktional J leisten. Diese Forderung kann maximal dahingehend abgeschwächt werden, dass \mathbf{Q} nur noch positiv semidefinit sein muss, denn diese Eigenschaft wird für die Lösbarkeit der Matrix-Riccati-Gleichung benötigt.

Wenn \mathbf{Q} aber nur noch positiv semidefinit ist, dann ist für beliebige Zustandsvektoren \mathbf{x} nur noch $\mathbf{x}^T \mathbf{Q} \mathbf{x} \geq 0$ und nicht mehr $\mathbf{x}^T \mathbf{Q} \mathbf{x} > 0$ sichergestellt. Damit kann es vorkommen, dass bestimmte Zustandsgrößen überhaupt keinen Effekt mehr auf den Term $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ und damit auf das Gütefunktional J haben. Dies ist jedoch nicht akzeptabel. Der resultierende Regler kann nur dann brauchbar sein, wenn alle Zustandsgrößen in die Optimierung eingehen, um sicherzustellen, dass im geschlossenen Kreis alle Zustandsgrößen einen zufriedenstellenden Verlauf aufweisen. Es ist demnach eine Bedingung zu ermitteln, die trotz positiver Semidefinitheit von \mathbf{Q} sicherstellt, dass sämtliche Zustandsgrößen einen Effekt auf den Term $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ haben.

Hierzu wird zunächst folgende Zerlegung durchgeführt:

$$\mathbf{Q} = (\mathbf{Q}^{\frac{1}{2}})^T \mathbf{Q}^{\frac{1}{2}} \quad (8.155)$$

Damit lässt sich eine Hilfsgröße $\tilde{\mathbf{y}} := \mathbf{Q}^{\frac{1}{2}} \mathbf{x}$ definieren, mit der gilt

$$\mathbf{x}^T \mathbf{Q} \mathbf{x} = \mathbf{x}^T (\mathbf{Q}^{\frac{1}{2}})^T \mathbf{Q}^{\frac{1}{2}} \mathbf{x} = \tilde{\mathbf{y}}^T \tilde{\mathbf{y}} \quad (8.156)$$

Wenn nun das Paar $(\mathbf{A}, \mathbf{Q}^{\frac{1}{2}})$ beobachtbar ist, dann ist sichergestellt, dass sämtliche Zustandsgrößen einen Effekt auf die Hilfsgröße $\tilde{\mathbf{y}}$ haben und damit auch auf den Term $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ und das Gütefunktional J .

Satz 8.10 *Die Forderung nach positiver Definitheit von \mathbf{Q} in Satz 8.8 lässt sich auch ersetzen durch die Forderung nach positiver Semidefinitheit von \mathbf{Q} und Beobachtbarkeit von $(\mathbf{A}, \mathbf{Q}^{\frac{1}{2}})$.*

8.6.3 Wahl der Gewichtungsmatrizen

In diesem Kapitel sollen einige Ansätze diskutiert werden, um die Matrizen \mathbf{R} und \mathbf{Q} geeignet festlegen zu können.

Die erste Überlegung überhaupt hat den Wertebereichen der Zustandsgrößen zu gelten. Angenommen, man hat zwei Zustandsgrößen x_1 und x_2 , und die Werte von x_1 sind ca. zehnmal so groß wie die Werte von x_2 . Weiterhin sei angenommen, dass man \mathbf{Q} als Diagonalmatrix gewählt hat, deren Diagonalelemente alle den gleichen Wert q aufweisen, um alle Zustandsgrößen im Gütefunktional gleich stark zu gewichten. Tatsächlich ist dann aber der Beitrag der ersten Zustandsgröße $x_1 q x_1$ zum Gütefunktional hundertmal so groß wie der Beitrag $x_2 q x_2$ der zweiten Zustandsgröße.

Daraus folgt, dass man sich vor Festlegung von \mathbf{R} und \mathbf{Q} intensiv mit den Wertebereichen der einzelnen Zustands- und Stellgrößen auseinandersetzen muss, damit die Gewichtung beim Reglerentwurf nicht unbeabsichtigt auf den völlig falschen Größen liegt. Sinnvoll ist es, die Zustandsgrößen alle auf denselben Wertebereich zu normieren (vgl. Kap. 2.2), bevor der Regler entworfen wird.

Nachdem die Wertebereiche identifiziert und die einzelnen Größen geeignet normiert sind, wird man die Suche nach einem geeigneten Regler üblicherweise mit einer Start-Einstellung für \mathbf{R} und \mathbf{Q} beginnen, wobei sich hier empfiehlt, beide Matrizen als Diagonalmatrizen zu wählen, deren Diagonalelemente alle gleich Eins sind. Den mit dieser Einstellung berechneten Regler kann man in der Simulation testen. Wenn die Regelung zu langsam ist, können die Koeffizienten von \mathbf{Q} vergrößert werden, um die von Null verschiedenen Werte der Zustandsgrößen zu bestrafen. Wenn dagegen die Amplituden der Stellgrößen zu stark waren, dann sollten stattdessen die Koeffizienten von \mathbf{R} vergrößert werden. Auf diese Art und Weise kann man sich sukzessive an den "richtigen" Regler herantasten.

Sehr einfach kann man auch statt der Zustandsgrößen, die ja als interne Größen möglicherweise sehr unanschaulich sind, die Ausgangsgrößen gewichten, die sich dem Anwender in der Regel viel eher erschließen. Hierzu ist

$$\mathbf{Q} = \mathbf{C}^T \text{diag}(q_{ii}) \mathbf{C} \quad (8.157)$$

zu wählen. Mit dieser Wahl ergibt sich

$$\begin{aligned} \mathbf{x}^T \mathbf{Q} \mathbf{x} &= \mathbf{x}^T \mathbf{C}^T \text{diag}(q_{ii}) \mathbf{C} \mathbf{x} = \mathbf{y}^T \text{diag}(q_{ii}) \mathbf{y} \\ &= \sum_{i=1}^q q_{ii} y_i^2 \end{aligned} \quad (8.158)$$

Dabei ist q die Anzahl der Ausgangsgrößen des Systems. Man sieht, dass nun die durch die Diagonalelemente q_{ii} gewichteten Ausgangsgrößen in den Riccati-Entwurf eingehen.

Der resultierende Regler wird gewährleisten, dass alle Ausgangsgrößen im von außen nicht angeregten Fall gegen Null konvergieren. Da für Stabilität

im Ljapunovschen Sinne aber die Zustandsgrößen gegen Null konvergieren müssen, muss hier zusätzlich sichergestellt werden, dass das System (\mathbf{A}, \mathbf{C}) beobachtbar ist. Denn damit ist sichergestellt, dass alle Zustandsgrößen eine Wirkung auf die Ausgangsgrößen haben. Wenn dann die Ausgangsgrößen gegen Null konvergieren, so konvergieren auch die Zustandsgrößen gegen Null.

Ein weiterer interessanter Aspekt ist die Robustheit des aus einem Riccati-Entwurf hervorgegangenen Reglers. Wie schon gesagt, ist der geschlossene Kreis mit einem Riccati-Regler stabil. Aber die Frage ist, ob die Stabilität auch im Zusammenspiel aus Regler und realer Strecke noch gewährleistet ist, wenn die reale Systemmatrix \mathbf{A} von der für den Entwurf verwendeten Systemmatrix \mathbf{A} abweicht.

Das Problem lässt sich auch in Bezug auf die Eigenwerte des geschlossenen Kreises formulieren: Da die Stabilität des geschlossenen Kreises durch den Riccati-Entwurf implizit gewährleistet ist, liegen alle Eigenwerte des geschlossenen Kreises zwar sicher in der negativen Hälfte der komplexen Ebene, möglicherweise aber nicht allzu weit von der imaginären Achse entfernt. Wenn der Regler nun auf die reale Strecke trifft, deren Systemmatrix von der Modell-Systemmatrix abweicht, könnten diese Eigenwerte in die rechte Halbebene wandern, und das System wäre instabil.

Um dies zu verhindern, kann beim Riccati-Entwurf durch eine Erweiterung sichergestellt werden, dass alle Eigenwerte des geschlossenen Kreises einen gewissen Mindestabstand zur imaginären Achse aufweisen:

$$\operatorname{Re}(\lambda_i) < -a \quad (8.159)$$

Dies lässt sich erreichen, indem der Verlauf jeder einzelnen Stell- und Zustandsgröße im Gütefunktional mit dem "Straffaktor" e^{at} versehen wird:

$$J = \int_0^{\infty} (e^{at} \mathbf{x}^T(t) \mathbf{Q} e^{at} \mathbf{x}(t) + e^{at} \mathbf{u}^T(t) \mathbf{R} e^{at} \mathbf{u}(t)) dt \quad (8.160)$$

J kann nur dann einen endlichen Wert annehmen, wenn sämtliche zeitabhängigen Terme im Integral gegen Null konvergieren. Für die mit dem Straffaktor versehenen Zustandsgrößen bedeutet dies, dass sie schneller gegen Null konvergieren müssen als die Funktion e^{-at} , da ansonsten das Produkt $e^{at} x_i(t)$ nicht gegen Null konvergieren würde. Laut Gleichung (7.120) wird aber die Konvergenzgeschwindigkeit jedes Einschwingvorgangs durch den zugehörigen Eigenwert definiert. Wenn also durch den Riccati-Entwurf sichergestellt wird, dass alle Einschwingvorgänge des geschlossenen Kreises mindestens mit e^{-at} gegen Null konvergieren, so folgt daraus, dass der Realteil jedes Eigenwertes kleiner als $-a$ sein muss. Damit ist das Entwurfsziel erreicht und der Straffaktor für die Zustandsgrößen erklärt.

Der Straffaktor für die Stellgrößen ist dann nur noch als Konsequenz aus den Straffaktoren für die Zustandsgrößen zu sehen. Wenn die Zustandsgrößen

mit einer bestimmten Geschwindigkeit gegen Null konvergieren, so folgt daraus, dass die Stellgrößen mit der gleichen Geschwindigkeit gegen Null konvergieren, da sie ja im Nullzustand ebenfalls Null sind. Von daher würde es im Entwurfsverfahren zu Inkonsistenzen führen, wenn man die Zustandsgrößen nicht mit dem gleichen Straffaktor belegt wie die Stellgrößen.

Wie wirken sich nun die Straffaktoren auf das Riccati-Verfahren aus? Um dies zu ermitteln, werden zunächst die Hilfsgrößen

$$\tilde{\mathbf{x}}(t) = e^{at}\mathbf{x}(t) \quad \text{und} \quad \tilde{\mathbf{u}}(t) = e^{at}\mathbf{u}(t) \quad (8.161)$$

eingeführt. Mit diesen ergibt sich für das Gütefunktional

$$J = \int_0^\infty (\tilde{\mathbf{x}}^T(t)\mathbf{Q}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{u}}^T(t)\mathbf{R}\tilde{\mathbf{u}}(t)) dt \quad (8.162)$$

Die Zustandsgleichung des auf den neuen Zustandsgrößen basierenden Systems gewinnt man aus der Ableitung des neuen Zustandsvektors:

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \frac{d}{dt}(e^{at}\mathbf{x}(t)) = e^{at}\dot{\mathbf{x}}(t) + ae^{at}\mathbf{x}(t) \\ &= e^{at}(\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)) + ae^{at}\mathbf{x}(t) \\ &= (\mathbf{A} + a\mathbf{I})e^{at}\mathbf{x}(t) + \mathbf{B}e^{at}\mathbf{u}(t) \\ &= (\mathbf{A} + a\mathbf{I})\tilde{\mathbf{x}}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) \\ &= \tilde{\mathbf{A}}\tilde{\mathbf{x}}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) \quad \text{mit} \quad \tilde{\mathbf{A}} = \mathbf{A} + a\mathbf{I} \end{aligned} \quad (8.163)$$

Damit sind Zustandsgleichung und Gütefunktional für das erweiterte System mit dem Zustandsvektor $\tilde{\mathbf{x}}$ und dem Stellgrößenvektor $\tilde{\mathbf{u}}$ gegeben, und der Riccati-Entwurf kann durchgeführt werden. \mathbf{P} gewinnt man aus der Matrix-Riccati-Gleichung

$$\tilde{\mathbf{A}}^T\mathbf{P} + \mathbf{P}\tilde{\mathbf{A}} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} = -\mathbf{Q} \quad (8.164)$$

und für das Regelgesetz gilt mit der so ermittelten Matrix \mathbf{P}

$$\tilde{\mathbf{u}} = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\tilde{\mathbf{x}} \quad (8.165)$$

Setzt man die Definitionen für die Hilfsvektoren $\tilde{\mathbf{x}}$ und $\tilde{\mathbf{u}}$ ein, so erhält man

$$\begin{aligned} e^{at}\mathbf{u} &= -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}e^{at}\mathbf{x} \\ \mathbf{u} &= -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x} \end{aligned} \quad (8.166)$$

also genau das gleiche Regelgesetz wie beim normalen Riccati-Entwurf. Daraus folgt:

Satz 8.11 *Ersetzt man beim Riccati-Entwurf die Systemmatrix \mathbf{A} durch $\tilde{\mathbf{A}} = \mathbf{A} + a\mathbf{I}$, so ist sichergestellt, dass die Eigenwerte des geschlossenen Kreises in der negativen Halbebene mindestens einen Abstand von a zur imaginären Achse aufweisen.*

8.6.4 Kalman-Filter

Zum Abschluss der Ausführungen zum Riccati-Entwurf soll auf den Entwurf von Beobachtern nach dieser Methode eingegangen werden. Mit den Ersetzungen nach (8.44) wird aus der Matrix-Riccati-Gleichung

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T - \mathbf{P}\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\mathbf{P} = -\mathbf{Q} \quad (8.167)$$

und aus der Bestimmungsgleichung für \mathbf{F} jetzt die Bestimmungsgleichung für \mathbf{H} :

$$\mathbf{H} = -\mathbf{P}\mathbf{C}^T\mathbf{R}^{-1} \quad (8.168)$$

Damit kann die Beobachter-Rückführung ermittelt werden. Es stellt sich aber die Frage nach der Bedeutung von \mathbf{R} und \mathbf{Q} , denn auch das Gütefunktional J wird sich durch die Ersetzungen nach (8.44) verändern. Zudem würde eine Minimierung von \mathbf{x} und \mathbf{u} beim Beobachterentwurf überhaupt keinen Sinn ergeben.

In der Tat gingen die Gleichungen (8.167) und (8.168) ursprünglich auch nicht, wie hier dargestellt, durch eine Analogieüberlegung aus dem Riccati-Reglerentwurf hervor, sondern sind auf eine völlig andere Art und Weise von Rudolf Kalman hergeleitet worden, weshalb ein nach diesen Gleichungen entworfener Beobachter auch *Kalman-Filter* genannt wird.

Ausgangspunkt der Überlegungen von Kalman waren die um Störungen erweiterten Systemgleichungen

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{s}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{m}(t) \end{aligned} \quad (8.169)$$

Dabei kann man \mathbf{s} als "Systemrauschen" verstehen und \mathbf{m} als Messrauschen, wobei für jede einzelne Komponente der beiden Störvektoren von weißem Rauschen ausgegangen wird. Damit lassen sich für beide Rauschprozesse die Autokorrelationsmatrizen entsprechend Gleichung (A.58) im Anhang A.6 angeben:

$$\begin{aligned} \psi_{\mathbf{ss}}(\tau) &= E(\mathbf{s}(t) \mathbf{s}^T(t + \tau)) = \mathbf{Q} \delta(\tau) \\ \psi_{\mathbf{mm}}(\tau) &= E(\mathbf{m}(t) \mathbf{m}^T(t + \tau)) = \mathbf{R} \delta(\tau) \end{aligned} \quad (8.170)$$

Dabei sind \mathbf{Q} und \mathbf{R} symmetrische Matrizen. Falls sie reine Diagonalmatrizen sind, bedeutet dies, dass jede Komponente der beiden Rauschsignalvektoren ein eigenständiges weißes Rauschen darstellt. Falls dagegen z.B. eine Komponente $r_{ij} = r_{ji}$ ungleich Null ist, so bedeutet dies, dass das Rauschen der Komponenten m_i und m_j korreliert ist. Dies kann auftreten, wenn die beiden Messaufnehmer der entsprechenden Komponenten des Ausgangsvektors derselben Störquelle ausgesetzt sind.

Für das Gütefunktional gilt beim Kalman-Filter

$$J = \sum_{i=1}^n E(\tilde{x}_i^2) = \sum_{i=1}^n E((x_i(t) - \hat{x}_i)^2) \quad (8.171)$$

Dabei kennzeichnet $E()$ den Erwartungswert einer Funktion (vgl. Anhang A.6). Die Minimierung dieses Gütefunktional bedeutet, dass bei diesem Beobachterentwurf versucht wird, die Differenz zwischen tatsächlichem Zustandsvektor und geschätztem Zustandsvektor trotz zweier Rauschquellen möglichst klein zu halten. Aus der Minimierung dieses Gütefunktional gehen dann nach aufwendiger Rechnung die Gleichungen (8.167) und (8.168) für den Beobachter hervor.

Ein Problem stellt bei diesem Beobachterentwurf die Festlegung der Matrizen \mathbf{Q} und \mathbf{R} dar. Während sie beim Riccati-Reglerentwurf direkt im Gütefunktional auftauchen und jeder einzelne Koeffizient einen Gewichtungsfaktor für eine Zustands- oder Stellgröße darstellt, so sind die Verhältnisse hier deutlich schwieriger. \mathbf{Q} und \mathbf{R} kennzeichnen jetzt die (angenommenen) Rauschprozesse, über die man in der Regel keinerlei Informationen hat, eben weil sie ja nur angenommen sind. Von daher bleibt einem nichts anderes übrig, als beide Matrizen zunächst als Einheitsmatrizen anzusetzen, den daraus berechneten Beobachter zu testen, und dann durch Modifikationen von \mathbf{Q} und \mathbf{R} zu versuchen, einen besseren Beobachter zu finden.

9. Nullstellen

9.1 Einführung

Im Verlauf dieses Kapitels soll auf den Zusammenhang zwischen Eigenwerten, Pol- und Nullstellen eines Systems eingegangen werden, wobei auch zu definieren ist, was bei einem Mehrgrößensystem unter Pol bzw. Nullstelle verstanden werden soll. Um den Einstieg zu erleichtern, sollen diese Begriffe zunächst noch einmal für Eingrößensysteme rekapituliert und in Zusammenhang gebracht werden.

Die Eigenwerte λ_i der Systemmatrix \mathbf{A} werden als Eigenwerte des Systems bezeichnet. Sie sind definiert als Nullstellen des charakteristischen Polynoms von \mathbf{A} und damit Lösungen der Gleichung

$$\det(s\mathbf{I} - \mathbf{A}) = 0 = \prod_{i=1}^n (s - \lambda_i) \quad (9.1)$$

Eigenwerte charakterisieren die Einschwingvorgänge, die das System ausführt, wenn es durch eine Anfangsauslenkung oder eine Eingangsgröße angeregt wird (vgl. (7.120)), und bestimmen auch das Stabilitätsverhalten des Systems.

Pole s_i sind bei einem Eingrößensystem die Nullstellen des Nenners der Übertragungsfunktion:

$$G(s) = \frac{Z(s)}{N(s)} = \frac{Z(s)}{(s - s_1)(s - s_2) \cdots (s - s_n)} \quad (9.2)$$

Das charakteristische Polynom der Systemmatrix bildet den Nenner der Übertragungsfunktion, so dass die Eigenwerte der Systemmatrix im Prinzip auch die Pole der Übertragungsfunktion sind. Dabei ist aber zu berücksichtigen, dass einzelne Nullstellen des Nennerpolynoms durch Nullstellen des Zählerpolynoms gekürzt werden können, so dass nicht jeder Eigenwert tatsächlich auch als Pol der Übertragungsfunktion auftaucht.

Nullstellen des Systems sind bei einem Eingrößensystem die Nullstellen des Zählers $Z(s)$ der Übertragungsfunktion. Sie beeinflussen die Form von Einschwingvorgängen, jedoch nicht das Stabilitätsverhalten des Systems.

Diese Begriffe sollen nun auf Mehrgrößensysteme übertragen werden. Für Polstellen ist dies relativ einfach:

Definition 9.1 Eine komplexe Zahl s_i heißt Pol des Mehrgrößensystems, wenn mindestens ein Element $G_{ij}(s)$ der Übertragungsmatrix $\mathbf{G}(s)$ einen Pol bei s_i besitzt.

Die Übertragung des Nullstellenbegriffs auf Mehrgrößensysteme ist schwieriger, und es gibt hierfür auch unterschiedliche Ansätze. Zunächst muss man sich klar machen, was eine Nullstelle in einem Eingrößensystem bedeutet. Nullstellen sind dort diejenigen (komplexen) Frequenzen s_i , für die die Übertragungsfunktion gleich Null ist. Bei einer Anregung des Systems durch die (komplexe) Funktion $e^{s_i t}$ ist die stationäre Ausgangsgröße nach Beendigung des Einschwingvorgangs gleich Null, d.h. das Eingangssignal wird durch das System nicht übertragen.

Diese Überlegung kann auf Mehrgrößensysteme übertragen werden. Die Frage lautet dann, ob es (komplexe) Frequenzen gibt, die vom Mehrgrößensystem nach Beendigung der Einschwingvorgänge nicht übertragen werden, d.h. im stationären Zustand muss für den Eingangsvektor $\mathbf{u}(t) = \mathbf{u}_0 e^{s_i t}$ der Ausgangsvektor verschwinden: $\mathbf{y} = \mathbf{0}$. Damit ergibt sich die folgende Definition:

Definition 9.2 Gegeben sei ein System mit der Übertragungsmatrix $\mathbf{G}(s)$, das mindestens ebenso viele Ausgangs- wie Eingangsgrößen aufweist. Dann ist eine komplexe Zahl s_i genau dann eine Übertragungsnullstelle des Systems, wenn ein Eingangsvektor $\mathbf{u}(s) \neq \mathbf{0}$ existiert, für den

$$\mathbf{G}(s_i)\mathbf{u}(s) = \mathbf{0} \quad (9.3)$$

gilt.

Aus dieser Definition lässt sich eine relativ leicht zu überprüfende Bedingung ableiten. Offenbar hat $\mathbf{G}(s_i)$ die Dimension $q \times p$, wobei q die Anzahl der Ausgangs- und p die Anzahl der Eingangsgrößen ist. Wenn $\mathbf{G}(s_i)$ nun p linear unabhängige Zeilen aufweist, dann existiert als Lösung der Gleichung (9.3) nur die triviale Lösung $\mathbf{u} = \mathbf{0}$, und damit wäre s_i keine Nullstelle.

Sofern der Rang von $\mathbf{G}(s_i)$ daher kleiner als p ist, existiert eine nicht-triviale Lösung, und s_i ist eine Nullstelle des Systems.

Wäre die Anzahl der Ausgangsgrößen kleiner als die Anzahl der Eingangsgrößen ($q < p$), dann könnte der Rang von $\mathbf{G}(s_i)$ maximal gleich q sein und wäre damit immer kleiner als p . In dem Fall wäre jede komplexe Zahl s_i eine Nullstelle des Systems, was natürlich sinnlos ist. Daher ist die Bedingung einer mindestens gleich großen Anzahl an Ausgangs- wie Eingangsgrößen in obiger Definition notwendig.

Satz 9.3 Gegeben sei ein System mit der Übertragungsmatrix $\mathbf{G}(s)$, das mindestens ebenso viele Ausgangs- wie Eingangsgrößen aufweist. Dann ist eine komplexe Zahl s_i genau dann eine Übertragungsnullstelle des Systems, wenn

$$\text{Rang}(\mathbf{G}(s_i)) < p \quad (9.4)$$

gilt. Dabei ist p die Anzahl der Eingangsgrößen. Für ein System mit gleich vielen Ein- und Ausgangsgrößen ist diese Bedingung gleichbedeutend mit

$$\det(\mathbf{G}(s_i)) = 0 \quad (9.5)$$

Angemerkt sei, dass die Übertragungsnullstellen eines Systems nicht unbedingt mit den Nullstellen der Teil-Übertragungsfunktionen $G_{ij}(s)$ übereinstimmen müssen. Daher können Übertragungsnullstellen durchaus dieselben Werte wie Pole aufweisen, da keine Kürzung stattfindet. Dies zeigt das folgende Beispiel:

Gegeben sei ein System mit der Übertragungsmatrix

$$\mathbf{G}(s) = \begin{bmatrix} \frac{s+1}{s+2} & 1 \\ 1 & \frac{s+2}{s+1} \end{bmatrix} \quad (9.6)$$

Da dieses System gleich viele Ein- und Ausgangsgrößen aufweist, können die Übertragungsnullstellen mit Hilfe der Determinante berechnet werden:

$$0 = \det(\mathbf{G}(s)) = 1 - \frac{1}{s+2} \rightarrow s = -1 \quad (9.7)$$

$s = -1$ ist also die (einzige) Übertragungsnullstelle, während, wie man leicht an den Teil-Übertragungsfunktionen ablesen kann, $s = -2$ sowie $s = -1$ die Polstellen des Systems sind.

9.2 Zusammenhang mit Steuer- und Beobachtbarkeit

Ausgangspunkt der Überlegungen ist ein System, welches in Jordanscher Normalform gegeben ist (vgl. Kap. 7.5.2). Der Einfachheit halber sei angenommen, dass sämtliche Eigenwerte verschieden sind, so dass die Systemmatrix \mathbf{A} eine reine Diagonalmatrix ist, deren Diagonalelemente die Eigenwerte des Systems sind. Die nachfolgenden Rechenschritte sind praktisch die gleichen, die auch schon im Zusammenhang mit dem Dominanzmaß von Eigenwerten durchgeführt worden sind (Kap. 7.9).

Die Systemmatrix lautet

$$\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n) \quad (9.8)$$

Die Übertragungsmatrix lautet gemäß (7.107):

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \quad (9.9)$$

Da \mathbf{A} eine Diagonalmatrix ist, gilt

$$s\mathbf{I} - \mathbf{A} = \begin{bmatrix} s - \lambda_1 & & & \\ & s - \lambda_2 & & \\ & & \dots & \\ & & & s - \lambda_n \end{bmatrix} \quad (9.10)$$

und

$$(s\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \frac{1}{s - \lambda_1} & & & \\ & \frac{1}{s - \lambda_2} & & \\ & & \dots & \\ & & & \frac{1}{s - \lambda_n} \end{bmatrix} \quad (9.11)$$

Für die Übertragungsmatrix folgt daraus

$$\begin{aligned} \mathbf{G}(s) &= [\mathbf{c}_1, \dots, \mathbf{c}_n] \begin{bmatrix} \frac{1}{s - \lambda_1} & & & \\ & \frac{1}{s - \lambda_2} & & \\ & & \dots & \\ & & & \frac{1}{s - \lambda_n} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \vdots \\ \mathbf{b}_n^T \end{bmatrix} \\ &= \sum_{i=1}^n \frac{1}{s - \lambda_i} \mathbf{c}_i \mathbf{b}_i^T \end{aligned} \quad (9.12)$$

Dabei sind die \mathbf{c}_i die Spaltenvektoren der Ausgangsmatrix \mathbf{C} , und die \mathbf{b}_i^T die Zeilenvektoren der Eingangsmatrix \mathbf{B} . Entsprechend ist $\mathbf{c}_i \mathbf{b}_i^T$ eine Matrix.

Nun folgt aber aus dem Gilbert-Kriterium für Steuerbarkeit (Satz 7.7), dass ein Eigenwert λ_i nicht steuerbar ist, wenn die zugehörige Zeile der Eingangsmatrix \mathbf{b}_i^T - in Jordanscher Normalform - gleich Null ist.

Analog dazu besagt das Gilbert-Kriterium für Beobachtbarkeit (Satz 7.11), dass ein Eigenwert λ_i nicht beobachtbar ist, wenn die zugehörige Spalte der Ausgangsmatrix \mathbf{c}_i gleich Null ist.

Daraus folgt aber doch, dass ein Summand

$$\frac{1}{s - \lambda_i} \mathbf{c}_i \mathbf{b}_i^T \quad (9.13)$$

in Gleichung (9.12) verschwindet, wenn der Eigenwert λ_i entweder nicht steuerbar oder nicht beobachtbar ist. Oder anders ausgedrückt, der entsprechende Summand ist nur dann von Null verschieden, wenn der Eigenwert sowohl steuerbar als auch beobachtbar ist.

Wenn dies der Fall ist, dann kommt in mindestens einer Teil-Übertragungsfunktion G_{ij} der Ausdruck $\frac{1}{s - \lambda_i}$ vor, d.h. der Eigenwert λ_i taucht als Pol in einer Teil-Übertragungsfunktion auf.

Satz 9.4 *Ein Eigenwert ist genau dann ein Pol, wenn er sowohl steuer- als auch beobachtbar ist.*

9.3 Invariante Nullstellen

Definition 9.5 *Verliert die Rosenbrock-Matrix*

$$\mathbf{P}(s) = \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (9.14)$$

an der Stelle $s = s_i$ ihren Höchststrang, so heißt s_i invariante Nullstelle des Systems.

Die Frage ist jetzt, was diese Definition mit Nullstellen zu tun hat. Dies soll für Systeme mit gleich vielen Ein- und Ausgängen sowie unter der Annahme, dass die s_i ungleich den Eigenwerten von \mathbf{A} sind, erläutert werden. Damit ist die Erläuterung zwar nicht vollständig, da sie nicht für alle Systeme gilt, aber ein vollständiger Nachweis würde hier den Rahmen sprengen. Entscheidend ist, dass die Beziehung zwischen der Rosenbrock-Matrix und den Übertragungsnullstellen eines Systems aufgezeigt wird.

Da für einen Eigenwert die Definitionsgleichung $\det(\lambda_i \mathbf{I} - \mathbf{A}) = 0$ gilt, folgt hier für die invarianten Nullstellen wegen ihrer Verschiedenheit von den Eigenwerten $\det(s_i \mathbf{I} - \mathbf{A}) \neq 0$. Daraus folgt aber, dass die Matrix $(s_i \mathbf{I} - \mathbf{A})$ auf jeden Fall invertierbar ist. Damit kann dann aber auch die Determinante

$$\det \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}(s_i \mathbf{I} - \mathbf{A})^{-1} & \mathbf{I} \end{bmatrix} \quad (9.15)$$

gebildet werden, deren Wert offensichtlich gleich Eins ist.

Nun kehren wir wieder zur Rosenbrock-Matrix zurück. Wenn die Rosenbrock-Matrix $\mathbf{P}(s)$ für $s = s_i$ keinen Höchststrang aufweist, so gilt

$$0 = \det(\mathbf{P}(s_i)) = \det \begin{bmatrix} s_i \mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (9.16)$$

Die rechte Seite wird nun mit der Determinanten (9.15) multipliziert. Da dies eine Multiplikation mit Eins ist, ändert sich die Gleichung dadurch nicht:

$$0 = \det \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}(s_i \mathbf{I} - \mathbf{A})^{-1} & \mathbf{I} \end{bmatrix} \det \begin{bmatrix} s_i \mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (9.17)$$

Für die Multiplikation zweier quadratischer Matrizen \mathbf{M} und \mathbf{N} gilt: $\det(\mathbf{MN}) = \det(\mathbf{M})\det(\mathbf{N})$. Daher entspricht das Produkt der Determinanten auf der rechten Seite von Gleichung (9.17) gerade der Determinanten des Matrizenprodukts:

$$\begin{aligned} 0 &= \det \left(\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}(s_i \mathbf{I} - \mathbf{A})^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} s_i \mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \right) \\ &= \det \begin{bmatrix} s_i \mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{0} & \mathbf{C}(s_i \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \end{bmatrix} \\ &= \det(s_i \mathbf{I} - \mathbf{A}) \det(\mathbf{C}(s_i \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}) \\ &= \det(s_i \mathbf{I} - \mathbf{A}) \det(\mathbf{G}(s_i)) \end{aligned} \quad (9.18)$$

Da die erste Determinante in diesem Ausdruck wegen $s_i \neq \lambda_i$ sicher ungleich Null ist, muss die zweite Determinante gleich Null sein, damit die Gleichung erfüllt ist. Es gilt also $0 = \det(\mathbf{G}(s_i))$, woraus folgt, dass s_i eine Übertragungsnullstelle des Systems ist.

Die Bezeichnung *Nullstelle* für diejenigen Werte $s = s_i$, für die die Rosenbrock-Matrix keinen Höchststrang aufweist, ist also gerechtfertigt. Doch warum *invariant*? Dies soll im folgenden für nicht sprungfähige Systeme ($\mathbf{D} = \mathbf{0}$) mit technisch sinnvoller Regelung geklärt werden.

Zunächst sei dazu die Matrix

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{F} & \mathbf{R} \end{bmatrix} \quad (9.19)$$

betrachtet, wobei \mathbf{F} die Regler-Rückführung und \mathbf{R} das Vorfilter ist (vgl. Kap. 8.2). Aus dem gleichen Grund, aus dem die Spalten der Eingangsmatrix \mathbf{B} eines Systems linear unabhängige Spalten aufweisen (vgl. Ausführungen zu Beginn des Kapitels 8), kann man auch davon ausgehen, dass die Spalten der Vorfiltermatrix \mathbf{R} linear unabhängig sind. Weiterhin kann man davon ausgehen, dass \mathbf{R} quadratisch ist, da jeder Sollwert üblicherweise auch an den Regelkreis weitergereicht wird. Damit sind aber auch die Zeilen von \mathbf{R} und somit die Zeilen in der unteren Hälfte der Matrix (9.19) linear unabhängig voneinander. Die Zeilen in der oberen Hälfte der Matrix (9.19) sind ihrerseits wegen der Identität in der linken oberen Hälfte linear unabhängig voneinander, so dass sich insgesamt ergibt, dass alle Zeilen dieser Matrix linear unabhängig voneinander sind und die Matrix damit regulär ist.

Wenn man eine Matrix mit einer regulären Matrix multipliziert, so ändert sich ihr Rang nicht. Dies soll nun für die Rosenbrock-Matrix (mit $\mathbf{D} = \mathbf{0}$) durchgeführt werden:

$$\begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{F} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} s\mathbf{I} - (\mathbf{A} + \mathbf{BF}) & -\mathbf{BR} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \quad (9.20)$$

Aufgrund des vorher Gesagten muss die Matrix auf der rechten Seite den gleichen Rang haben wie die Rosenbrock-Matrix. Die Matrix auf der rechten Seite ist aber gerade die Rosenbrock-Matrix des geregelten Systems mit Zustandsrückführung und Vorfilter. Denn an Abbildung 8.2 kann man erkennen, dass die Eingangsmatrix des so erweiterten Systems gerade das Produkt \mathbf{BR} ist und seine Systemmatrix $\mathbf{A} + \mathbf{BF}$. Daraus folgt, dass das geregelte System die gleichen invarianten Nullstellen besitzt wie das ungeregelte System. Damit ist die Bezeichnung *invariante Nullstelle* gerechtfertigt.

Es bleibt die Frage zu klären, was passiert, wenn eine invariante Nullstelle mit einem Eigenwert zusammenfällt. Die Antwort lässt sich mit Hilfe des Hautus-Kriteriums geben (Sätze 7.6 und 7.10).

Bei Nicht-Steuerbarkeit des Eigenwertes λ_i gilt lt. Hautus-Kriterium

$$\text{Rang}[\lambda_i \mathbf{I} - \mathbf{A}, \mathbf{B}] < n \quad (9.21)$$

d.h. die Matrix weist weniger als n linear unabhängige Zeilen auf. Damit sind aber offenbar auch die ersten n Zeilen der Rosenbrock-Matrix nicht mehr linear unabhängig, so dass diese ihren Höchstrang verliert. Daher muss λ_i auch eine invariante Nullstelle des Systems sein.

Jeder nicht steuerbare Eigenwert des Systems fällt daher mit einer invarianten Nullstelle an gleicher Stelle zusammen. Und da ein nicht steuerbarer Eigenwert wegen Satz 9.4 nicht als Pol auftaucht, bedeutet dies, dass die invariante Nullstelle den Eigenwert kompensiert.

In analoger Weise kann man auch das Beobachtbarkeitskriterium nach Hautus betrachten. Bei Nicht-Beobachtbarkeit des Eigenwertes λ_i gilt hier

$$\text{Rang} \begin{bmatrix} \lambda_i \mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix} < n \quad (9.22)$$

d.h. die Matrix weist weniger als n linear unabhängige Spalten auf. Damit sind aber offenbar auch die ersten n Spalten der Rosenbrock-Matrix nicht mehr linear unabhängig, so dass diese wieder ihren Höchstrang verliert. Daher muss λ_i auch hier eine invariante Nullstelle des Systems sein.

Jeder nicht beobachtbare Eigenwert des Systems fällt daher mit einer invarianten Nullstelle an gleicher Stelle zusammen. Und da ein nicht beobachtbarer Eigenwert wegen Satz 9.4 nicht als Pol auftaucht, bedeutet dies, dass die invariante Nullstelle den Eigenwert kompensiert.

Zusammenfassend lässt sich also sagen:

1. Ein System n -ter Ordnung habe n Eigenwerte und m Polstellen. Es gilt immer $n \geq m$.
2. Alle Polstellen sind auch Eigenwerte des Systems, aber nicht jeder Eigenwert ist auch Polstelle.
3. Jeder steuerbare und beobachtbare Eigenwert taucht als Pol in mindestens einer Teil-Übertragungsfunktion auf. Alle übrigen Eigenwerte tauchen nicht als Pole auf, werden also durch invariante Nullstellen kompensiert.
4. Diejenigen invarianten Nullstellen, die nicht durch Polstellen kompensiert werden, tauchen als Übertragungsnullstellen auf. Dabei können Übertragungsnullstellen und Pole durchaus auf denselben Wert fallen, da Übertragungsnullstellen nicht explizit im Zähler aller einzelnen Übertragungsfunktionen auftreten müssen und somit nicht jeder Pol gekürzt werden muss.
5. Diejenigen invarianten Nullstellen, die durch Polstellen kompensiert werden, bezeichnet man als *Entkopplungsnullstellen*.
6. Die Menge der invarianten Nullstellen setzt sich zusammen aus der Menge der Übertragungsnullstellen und der Menge der Entkopplungsnullstellen.
7. Ein vollständig steuer- und beobachtbares System besitzt keine Entkopplungsnullstellen, da alle Eigenwerte als Pole auftauchen und somit keine Kompensation stattfindet.

8. Die Entkopplungsnullstellen können nur mit Hilfe der Rosenbrock-Matrix berechnet werden, nicht jedoch aus $\det(\mathbf{G}(s_i)) = 0$. Diese Determinante dient ausschließlich zur Berechnung von Übertragungsnullstellen.
9. Eine invariante Nullstelle ist genau dann Übertragungsnullstelle, wenn sie mit keinem Eigenwert zusammenfällt oder wenn der Eigenwert, mit dem sie zusammenfällt, sowohl steuer- als auch beobachtbar ist.

9.4 Anwendungen

9.4.1 Berechnung eines Vorfilters

In Kapitel 8.2 wurde die Vorfiltermatrix \mathbf{R} zur Gewährleistung stationärer Genauigkeit des Führungsverhaltens eingeführt. Die Berechnung von \mathbf{R} erfolgt gemäß (8.15):

$$\mathbf{R} = -[\mathbf{C}(\mathbf{A} + \mathbf{B}\mathbf{F})^{-1}\mathbf{B}]^{-1} \quad (9.23)$$

Voraussetzung für die Berechenbarkeit von \mathbf{R} ist zunächst die Invertierbarkeit von $(\mathbf{A} + \mathbf{B}\mathbf{F})$, also der Systemmatrix des mit einer Zustandsrückführung geschlossenen Kreises. Da die Regler-Rückführung \mathbf{F} aber so ausgelegt wird, dass das Gesamtsystem stabil ist, liegen sämtliche Eigenwerte λ_i von $(\mathbf{A} + \mathbf{B}\mathbf{F})$ in der linken Hälfte der komplexen Ebene, und es gilt:

$$\det(\mathbf{A} + \mathbf{B}\mathbf{F}) = \prod_{i=1}^n \lambda_i \neq 0 \quad (9.24)$$

Damit ist gewährleistet, dass $(\mathbf{A} + \mathbf{B}\mathbf{F})$ invertierbar ist.

Weiterhin muss aber auch der Gesamtterm $[\mathbf{C}(\mathbf{A} + \mathbf{B}\mathbf{F})^{-1}\mathbf{B}]$ invertierbar sein. Die hierzu notwendige Bedingung liefert der folgende Satz:

Satz 9.6 *Unter der Voraussetzung, dass die Eigenwerte des mit Hilfe einer Rückführung \mathbf{F} geregelten Systems ungleich Null sind, existiert das Vorfilter \mathbf{R} genau dann, wenn keine invariante Nullstelle in Null liegt.*

Ausgangspunkt des Beweises ist die Determinante der Rosenbrock-Matrix, wobei die Durchgangsmatrix wieder Null sei:

$$\det(\mathbf{P}(s_i)) = \det \begin{bmatrix} s_i \mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \quad (9.25)$$

Nun sei angenommen, dass s_i keinem Eigenwert des geregelten Systems entspricht. Diese Annahme ist nicht weiter einschränkend, da uns weiter unten nur noch der Fall $s_i = 0$ interessiert, und Null ist sicherlich kein Eigenwert eines vernünftig geregelten Systems. Damit gilt

$$\det(s_i \mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{F})) \neq 0 \quad (9.26)$$

und die Inverse dieser Matrix kann gebildet werden.

Die Determinante der Rosenbrock-Matrix wird jetzt von links und rechts mit zwei Determinanten multipliziert, deren Wert jeweils Eins beträgt und die somit die Gleichung nicht beeinflussen:

$$\det(\mathbf{P}(s_i)) = \det \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}(s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1} & \mathbf{I} \end{bmatrix} \det \begin{bmatrix} s_i\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \det \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{F} & \mathbf{I} \end{bmatrix} \quad (9.27)$$

Das Produkt der Determinanten ist gleich der Determinante des Matrizenprodukts. Entsprechend werden zunächst die beiden hinteren Determinanten zusammengefasst, und anschließend die Resultierende mit der ersten Matrix:

$$\begin{aligned} \det(\mathbf{P}(s_i)) &= \det \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}(s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1} & \mathbf{I} \end{bmatrix} \det \begin{bmatrix} s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}) & -\mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \\ &= \det \begin{bmatrix} s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}) & -\mathbf{B} \\ \mathbf{0} & \mathbf{C}(s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{B} \end{bmatrix} \\ &= \det [s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF})] \det [\mathbf{C}(s_i\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{B}] \quad (9.28) \end{aligned}$$

Nun sei angenommen, dass $s_i = 0$ eine invariante Nullstelle ist. Nach Definition 9.5 verliert in dem Fall die Rosenbrock-Matrix ihren Höchststrang, d.h. ihre Determinante wird Null. Dann gilt mit (9.28)

$$0 = \det(\mathbf{P}(0)) = \det [-(\mathbf{A} + \mathbf{BF})] \det [-\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}] \quad (9.29)$$

Da alle Eigenwerte des geregelten Systems von Null verschieden sind, ist die erste Determinante sicherlich ungleich Null. Damit die Gleichung erfüllt ist, muss also die zweite Determinante gleich Null sein. Daraus folgt aber sofort, dass $[\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}]$ nicht invertierbar ist und somit die Vorfiltermatrix \mathbf{R} nicht existiert.

Falls $s_i = 0$ keine invariante Nullstelle ist, ist die Determinante der Rosenbrock-Matrix ungleich Null, und es gilt

$$0 \neq \det(\mathbf{P}(0)) = \det [-(\mathbf{A} + \mathbf{BF})] \det [-\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}] \quad (9.30)$$

Dieser Ausdruck trifft nur dann zu, wenn beide Determinanten auf der rechten Seite von Null verschieden sind. Für die erste Determinante wurde dies schon nachgewiesen. Und aus der Verschiedenheit der zweiten Determinante von Null folgt sofort, dass der Ausdruck $[\mathbf{C}(\mathbf{A} + \mathbf{BF})^{-1}\mathbf{B}]$ invertierbar ist und somit die Vorfiltermatrix \mathbf{R} existiert.

Damit ist der Satz bewiesen.

9.4.2 Falb-Wolovich-Entwurf

Hier sollen die Überlegungen vom Ende des Kapitels 8.5.2 noch einmal aufgegriffen und mit dem Begriff der invarianten Nullstellen verknüpft werden.

Beim Entwurf nach Falb-Wolovich werden die Rückführmatrix \mathbf{F} und das Vorfilter \mathbf{R} so berechnet, dass jede Ausgangsgröße ausschließlich von der ihr zugeordneten Eingangsgröße abhängig ist, dass also das geregelte Gesamtsystem aus lauter voneinander entkoppelten Eingrößensystemen besteht.

Die Übertragungsmatrix des geregelten Systems (vgl. (8.10))

$$\mathbf{T}(s) = \mathbf{C}(s\mathbf{I} - (\mathbf{A} + \mathbf{BF}))^{-1}\mathbf{BR} \quad (9.31)$$

ist damit diagonalförmig, und für jedes Paar von Ein- und Ausgangsgrößen gilt gemäß (8.111):

$$y_i(s) = \frac{k_i}{s^{\delta_i} + q_{i,\delta_i-1}s^{\delta_i-1} + \dots + q_{i,1}s + q_{i,0}} w_i(s) = T_i(s)w_i(s) \quad (9.32)$$

Mit der Differenzenordnung $\delta = \sum \delta_i$ hat das Gesamtsystem maximal δ Pole und keine Übertragungsnullstellen, da im Zähler jeweils nur eine konstante Zahl auftaucht. Falls δ kleiner als die Systemordnung n ist, dann müssen Eigenwerte existieren, die keine Pole sind, also in keiner der Übertragungsfunktionen lt. (9.32) auftauchen und somit offenbar unbeobachtbar sind. Entsprechend Kapitel 9.3 werden diese Eigenwerte durch invariante Nullstellen kompensiert.

Um sicherzustellen, dass die unbeobachtbaren Eigenwerte nicht instabil sind, hat man damit zwei Möglichkeiten: Entweder berechnet man, wie schon in Kapitel 8.5.2 skizziert, die Eigenwerte des geschlossenen Kreises und prüft, ob alle einen negativen Realteil haben. Oder man berechnet die invarianten Nullstellen des Systems und prüft, ob bei diesen ausschließlich negative Realteile vorliegen.

10. Normoptimale Regelung

Dieses Kapitel folgt sehr eng der didaktisch hervorragenden Darstellung in [43], mit dessen Autor, Prof. Dr.-Ing. Kai Müller, ich viele anregende Gespräche zu normoptimalen Regelungen und anderen Themen führen konnte und hoffentlich auch weiterhin führen werde.

10.1 Beziehungen im geschlossenen Kreis

In diesem Abschnitt werden einige Beziehungen für Mehrgrößensysteme hergeleitet, die in den nachfolgenden Abschnitten noch verwendet werden. Grundlage ist der aus Regler und Strecke bestehende Standard-Regelkreis entsprechend Abbildung 6.1 bzw. 10.1, nur dass jetzt von einem Mehrgrößensystem ausgegangen wird.

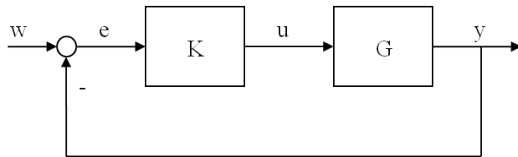


Abb. 10.1. Regelkreis

Damit werden Regler $\mathbf{K}(s)$ und Strecke $\mathbf{G}(s)$ nicht mehr durch Übertragungsfunktionen, sondern durch frequenzabhängige Übertragungsmatrizen repräsentiert (vgl. (7.107)), und die Systemgrößen \mathbf{w} , \mathbf{e} und \mathbf{y} sind keine skalaren Signale, sondern Signalvektoren. Dies hat zur Folge, dass auf die Reihenfolge der Multiplikationen geachtet werden muss, denn die Matrizen- und Vektormultiplikation ist nicht kommutativ. Das Produkt \mathbf{GK} ist etwas anderes als das Produkt \mathbf{KG} .

Unter Berücksichtigung dieser Tatsache gilt im Frequenzbereich:

$$\begin{aligned}\mathbf{y} &= \mathbf{GK}\mathbf{e} = \mathbf{GK}(\mathbf{w} - \mathbf{y}) \\ (\mathbf{I} + \mathbf{GK})\mathbf{y} &= \mathbf{GK}\mathbf{w} \\ \mathbf{y} &= (\mathbf{I} + \mathbf{GK})^{-1}\mathbf{GK}\mathbf{w} = \mathbf{T}\mathbf{w}\end{aligned}\tag{10.1}$$

Dabei ist

$$\mathbf{T} = (\mathbf{I} + \mathbf{GK})^{-1} \mathbf{GK} \quad (10.2)$$

die schon aus Kapitel 6.1 bekannte *Führungs-Übertragungsfunktion*, die sich im Eingrößenfall auf

$$T(s) = \frac{G(s)K(s)}{1 + G(s)K(s)} \quad (10.3)$$

reduziert. In den Gleichungen (10.1) und (10.2) sowie im weiteren Verlauf dieses Kapitels wird die Frequenzabhängigkeit der Übertragungsmatrizen ($\mathbf{T}(s)$) nicht immer explizit erwähnt, aber natürlich immer vorausgesetzt.

Den Ausdruck

$$\mathbf{S} = (\mathbf{I} + \mathbf{GK})^{-1} \quad (10.4)$$

bezeichnet man als *Empfindlichkeitsfunktion*. Diese kann auch als Übertragungsfunktion von einer Störung am Ausgang zur Ausgangsgröße gedeutet werden, was recht einfach herzuleiten ist. Man muss sich lediglich vorstellen, dass eine Störung \mathbf{d} nicht vor, sondern unmittelbar hinter der Strecke, aber noch vor der Rückkopplung, angreift. Dann gilt

$$\begin{aligned} \mathbf{y} &= \mathbf{d} - \mathbf{GK}\mathbf{y} \\ (\mathbf{I} + \mathbf{GK})\mathbf{y} &= \mathbf{d} \\ \mathbf{y} &= (\mathbf{I} + \mathbf{GK})^{-1} \mathbf{d} = \mathbf{S}\mathbf{d} \end{aligned} \quad (10.5)$$

Für den Zusammenhang zwischen \mathbf{T} und \mathbf{S} folgt aus (10.2):

$$\begin{aligned} \mathbf{T} &= \mathbf{SGK} = (\mathbf{I} + \mathbf{GK})^{-1} \mathbf{GK} = (\mathbf{I} + \mathbf{GK})^{-1} (\mathbf{I} + \mathbf{GK} - \mathbf{I}) \\ &= \mathbf{I} - (\mathbf{I} + \mathbf{GK})^{-1} = \mathbf{I} - \mathbf{S} \end{aligned} \quad (10.6)$$

bzw.

$$\mathbf{T} + \mathbf{S} = \mathbf{I} \quad (10.7)$$

Daher wird \mathbf{T} auch *komplementäre Empfindlichkeitsfunktion* genannt. Vergegenwärtigt man sich, dass das Ziel einer guten Regelung darin bestehen sollte, dass der Ausgangsvektor \mathbf{y} gleich dem Eingangsvektor \mathbf{w} und damit die Übertragungsmatrix $\mathbf{T}(s)$ zumindest für Gleichsignale ($s = 0$) der Identität entsprechen sollte (vgl. Kap. 8.2), so folgt daraus mit (10.7) zwangsläufig, dass die Empfindlichkeitsmatrix \mathbf{S} möglichst klein, am besten Null, sein muss.

Alternativ lässt sich der Zusammenhang zwischen \mathbf{T} und \mathbf{S} auch wie folgt berechnen:

$$\begin{aligned} \mathbf{e} &= \mathbf{w} - \mathbf{GK}\mathbf{e} \\ (\mathbf{I} + \mathbf{GK})\mathbf{e} &= \mathbf{w} \\ \mathbf{e} &= (\mathbf{I} + \mathbf{GK})^{-1} \mathbf{w} = \mathbf{S}\mathbf{w} \\ \mathbf{y} &= \mathbf{GK}\mathbf{e} = \mathbf{GKS}\mathbf{w} \end{aligned} \quad (10.8)$$

Damit folgt zusammenfassend aus den Gleichungen (10.2), (10.7) und (10.8)

$$\mathbf{T} = \mathbf{GKS} = \mathbf{SGK} = \mathbf{I} - \mathbf{S} \quad (10.9)$$

Grundsätzlich kann man \mathbf{T} auch mit vertauschter Reihenfolge von \mathbf{K} und \mathbf{G} berechnen, wenn man die Rechnung bzgl. der Stellgröße \mathbf{u} durchführt:

$$\begin{aligned} \mathbf{u} &= \mathbf{K}\mathbf{w} - \mathbf{KG}\mathbf{u} \\ (\mathbf{I} + \mathbf{KG})\mathbf{u} &= \mathbf{K}\mathbf{w} \\ \mathbf{u} &= (\mathbf{I} + \mathbf{KG})^{-1}\mathbf{K}\mathbf{w} \\ \mathbf{y} &= \mathbf{G}\mathbf{u} = \mathbf{G}(\mathbf{I} + \mathbf{KG})^{-1}\mathbf{K}\mathbf{w} \end{aligned} \quad (10.10)$$

Somit ergibt sich hier:

$$\mathbf{T} = \mathbf{G}(\mathbf{I} + \mathbf{KG})^{-1}\mathbf{K} \quad (10.11)$$

Eine weitere Interpretation bietet sich für den Ausdruck $(\mathbf{I} + \mathbf{KG})$ an. Hierzu denkt man sich den Regelkreis in der Rückführung aufgetrennt (Abb. 10.2).

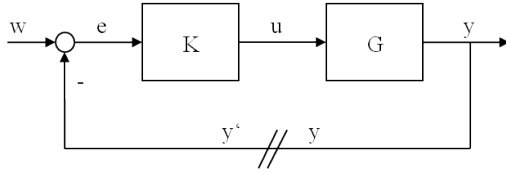


Abb. 10.2. Zur Rückführdifferenz

Für $\mathbf{w} = \mathbf{0}$ gilt

$$\begin{aligned} \mathbf{y} &= -\mathbf{GK}\mathbf{y}' \\ \mathbf{y}' - \mathbf{y} &= \mathbf{y}' + \mathbf{GK}\mathbf{y}' = (\mathbf{I} + \mathbf{GK})\mathbf{y}' = \mathbf{F}\mathbf{y}' \end{aligned} \quad (10.12)$$

mit der Rückführdifferenzmatrix

$$\mathbf{F} = \mathbf{I} + \mathbf{GK} \quad (10.13)$$

die bei SISO-Systemen gerade den Nenner der Übertragungsfunktion darstellt.

10.2 Normen für Signale und Systeme

10.2.1 Grundbegriffe

Mit den hier angegebenen Definitionen soll lediglich das Grundgerüst vermittelt werden, das zum Verständnis der Normen von Übertragungsgliedern notwendig ist. Eine wesentlich ausführlichere Darstellung findet sich z.B. im Anhang von [5].

Definition 10.1 Eine Menge X heißt Linearer Raum, wenn eine Addition und eine Skalarmultiplikation definiert sind, für die folgende Eigenschaften gelten (dabei sind $x, y, z \in X$ und a, b komplexe Zahlen):

$$\begin{aligned}
 x + y &= y + x && (\text{Kommutativität}) \\
 x + (y + z) &= (x + y) + z && (\text{Assoziativität}) \\
 0 + x &= x && (\text{Existenz eines Nullelementes}) \\
 x + \hat{x} &= x + (-x) = 0 && (\text{Existenz eines inversen Elementes}) \\
 a(x + y) &= ax + ay && (\text{Distributivität}) \\
 (a + b)x &= ax + bx && (\text{Distributivität}) \\
 (ab)x &= a(bx) && (\text{Assoziativität}) \\
 1x &= x && (\text{Neutrales Element})
 \end{aligned} \tag{10.14}$$

Definition 10.2 Ein linearer Raum X heisst normiert, wenn eine reellwertige Norm $\|\bullet\|$ mit den folgenden Eigenschaften definiert ist (dabei sind $x, y \in X$ und a eine komplexe Zahl):

$$\begin{aligned}
 \|x\| &> 0 && \text{für } x \neq 0 \\
 \|x\| &= 0 && \text{für } x = 0 \\
 \|x + y\| &\leq \|x\| + \|y\| && (\text{Dreiecksungleichung}) \\
 \|ax\| &= |a| \|x\|
 \end{aligned} \tag{10.15}$$

Ein einfaches Beispiel für eine Norm ist die Vektornorm, die für Vektoren reellwertiger Zahlen definiert ist:

$$\|\mathbf{x}\|_k := \left(\sum_{i=1}^n |x_i|^k \right)^{\frac{1}{k}} \tag{10.16}$$

Für $k = 2$ ergibt sich die euklidische Vektornorm, die, um Verwechslungsgefahr vorzubeugen, wie üblich mit einfachen Betragsstrichen gekennzeichnet werden soll:

$$|\mathbf{x}| := \sqrt{\sum_{i=1}^n |x_i|^2} \tag{10.17}$$

10.2.2 Signalnormen

Die Menge der für $-\infty < t < \infty$ in der p -ten Potenz $1 \leq p < \infty$ absolut integrierbaren Funktionen $f(t)$ bildet offenbar einen linearen Raum, und zwar den Raum L_p . Normiert wird dieser Raum durch die sogenannte p -Norm einer zeitveränderlichen Funktion:

$$\|f\|_p := \left(\int_{-\infty}^{\infty} |f(t)|^p dt \right)^{\frac{1}{p}} \quad (10.18)$$

Für einige spezielle Werte von p ist die p -Norm von praktischem Interesse. So ergibt sich beispielsweise für $p \rightarrow \infty$ die sogenannte ∞ -Norm des Signales $f(t)$:

$$\|f\|_{\infty} := \sup_t |f(t)| \quad (10.19)$$

Dies ist ein sehr einfaches Ergebnis, nämlich nichts weiter als die kleinste obere Schranke des Absolutwertes des Signals, die man unter praktischen Gesichtspunkten durchaus als Maximalwert verstehen kann. Erklären kann man sich dieses Ergebnis wie folgt: Für $p \rightarrow \infty$ wird zunächst der Integrand $|f(t)|$ mit ∞ potenziert. Bei dieser Operation ergibt sich für den Maximalwert von $|f(t)|$ ein Resultat, das so viel größer ist als die Resultate für alle anderen Funktionswerte, dass man diese vernachlässigen kann und nur noch den Maximalwert für die nächste Operation zu berücksichtigen braucht. Dies ist die Integration, wobei eine Integration über einen einzigen Wert genau diesen Wert als Ergebnis liefert, also weiterhin den mit ∞ potenzierten Maximalwert von $|f(t)|$. Dieser wird abschließend mit $\frac{1}{p}$ bzw. mit $\frac{1}{\infty}$ potenziert, womit sich beide Potenzierungen kompensieren, so dass man das obige Ergebnis erhält. Dass in der Formel das Supremum und nicht das Maximum auftritt, ist in den Unterschieden zwischen der hier aufgeführten anschaulichen Erklärung und der exakten theoretischen Herleitung begründet.

Von weiterem Interesse sind die 1- und die 2-Norm. Für die 1-Norm ergibt sich mit $p = 1$ aus Gleichung (10.18) das ebenfalls sehr einfache Resultat

$$\|f\|_1 := \int_{-\infty}^{\infty} |f(t)| dt \quad (10.20)$$

und für die 2-Norm mit $p = 2$:

$$\|f\|_2 := \sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt} \quad (10.21)$$

Falls $f(t)$ ein Signal am Ein- oder Ausgang eines Systems ist, so lässt sich die 2-Norm als ein Maß für den Energieinhalt des Signales deuten, also für die Energie, die in das System hinein- oder aus ihm herausgeführt wird. Als Beispiel sei $f(t) = u(t)$ die Spannung an einem elektrischen Widerstand R . Dann ergibt sich für die 2-Norm

$$\|u\|_2 = \sqrt{\int_{-\infty}^{\infty} |u(t)|^2 dt} = \sqrt{R \int_{-\infty}^{\infty} \frac{1}{R} u^2(t) dt}$$

$$= \sqrt{R \int_{-\infty}^{\infty} P(t) dt} = \sqrt{RW} \quad (10.22)$$

mit der im Widerstand umgesetzten elektrischen Leistung P und der Energie W . Die 2-Norm ist hier demnach proportional zur Wurzel aus der elektrischen Energie.

Falls der Energieinhalt eines Signales, d.h. seine 2-Norm, unendlich groß ist, wie z.B. bei einer stationären Sinusschwingung, kann statt der 2-Norm auch die mittlere Leistung zur Charakterisierung des Signals herangezogen werden:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t))^2 dt \quad (10.23)$$

und eine zur 2-Norm vergleichbare Größe

$$pow(f(t)) := \sqrt{\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t))^2 dt} \quad (10.24)$$

definiert werden. Bei dieser Größe handelt es sich jedoch nicht um eine Norm, da sie den Wert Null auch für $f(t) \neq 0$ annehmen kann, z.B. für alle Signale, die nur für endliche Zeit einen von Null verschiedenen Wert aufweisen und ansonsten Null sind. Für diese Signale ergibt sich aufgrund der Division durch $2T$ bei immer weiter wachsendem T der Wert $pow(f(t)) = 0$.

Die Absolutstriche im Integranden können hier weggelassen werden, da $pow(f)$ im Gegensatz zur 2-Norm grundsätzlich nur auf reellwertige Funktionen angewandt wird.

Offensichtlich gilt für alle Normen, dass unterschiedliche Signale durchaus den gleichen Wert für Ihre Norm aufweisen können. Dies ermöglicht die Zusammenfassung verschiedener Signale zu einer Klasse, die durch den Wert der gewählten Norm charakterisiert ist. Sinn macht eine derartige Klassifizierung allerdings nur dann, wenn die Norm einen endlichen Wert aufweist. Von daher sollen im folgenden einige Aussagen dazu hergeleitet werden, die dann später aber auch noch Verwendung finden.

Satz 10.3 *Gegeben sei ein Signal $f(t)$. Aus $\|f\|_2 < \infty$ folgt $pow(f) = 0$*

Beweis: Für das Quadrat der 2-Norm gilt für endliche Werte von T :

$$\|f\|_2^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt \geq \int_{-T}^T |f(t)|^2 dt \quad (10.25)$$

Nun dividiert man auf beiden Seiten der Ungleichung durch $2T$ und lässt T gegen ∞ laufen:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \|f\|_2^2 \geq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |f(t)|^2 dt = (\text{pow}(f))^2 \quad (10.26)$$

Da die 2-Norm und damit auch ihr Quadrat einen endlichen Wert haben, wird die linke Seite der Ungleichung aufgrund des immer größer werdenden Nenners $2T$ gegen Null konvergieren. Damit ist aber auch $\text{pow}(f) = 0$.

Der Satz ist auch anschaulich sofort einsichtig. Wenn die Energie eines Signales endlich ist, dann kann die mittlere Leistung bei Betrachtung eines unendlich langen Zeitintervalls nur Null sein.

Satz 10.4 *Gegeben sei ein Signal $f(t)$. Wenn $\|f\|_\infty < \infty$ und $\text{pow}(f) < \infty$, dann gilt $\text{pow}(f) \leq \|f\|_\infty$*

Beweis: Aufgrund ihrer Definition ist der Wert der ∞ -Norm eines Signales immer größer oder gleich dem größten absoluten Wert des Signales. Daher gilt

$$\int_{-T}^T (f(t))^2 dt \leq \int_{-T}^T \|f\|_\infty^2 dt = \int_{-T}^T dt \|f\|_\infty^2 = 2T \|f\|_\infty^2 \quad (10.27)$$

Nun dividiert man wieder auf beiden Seiten der Ungleichung durch $2T$ und bildet den Grenzwert:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t))^2 dt \leq \lim_{T \rightarrow \infty} \frac{1}{2T} 2T \|f\|_\infty^2 \quad (10.28)$$

Es ergibt sich mit der Definition für $\text{pow}(f)$:

$$(\text{pow}(f))^2 \leq \|f\|_\infty^2 \quad (10.29)$$

Auch dieser Satz ist plausibel. Natürlich kann die Wurzel aus dem mittleren quadratischen Wert eines Signales ($= \text{pow}(f)$) nicht größer sein als sein Maximalwert.

Satz 10.5 *Gegeben sei ein Signal $f(t)$. Wenn $\|f\|_\infty < \infty$ und $\|f\|_1 < \infty$, dann gilt $\|f\|_2 \leq \sqrt{\|f\|_\infty \|f\|_1}$*

Beweis:

$$\begin{aligned} \|f\|_2^2 &= \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |f(t)| |f(t)| dt \\ &\leq \int_{-\infty}^{\infty} \|f\|_\infty |f(t)| dt = \|f\|_\infty \int_{-\infty}^{\infty} |f(t)| dt = \|f\|_\infty \|f\|_1 \quad (10.30) \end{aligned}$$

Zum Abschluss dieses Kapitels sollen die Definitionen (10.19) und (10.21) auf den Mehrgrößenfall erweitert werden: Die Vektorfunktionen

$$\mathbf{f} = [f_1(t), \dots, f_n(t)]^T \quad (10.31)$$

mit $f_i \in L_p$ bilden den Raum L_p^n . Für $1 \leq p < \infty$ lautet eine mögliche Norm für diesen Raum:

$$\|\mathbf{f}\|_p := \left(\sum_{i=1}^n (\|f_i\|_p)^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^n \int_{-\infty}^{\infty} |f_i(t)|^p dt \right)^{\frac{1}{p}} \quad (10.32)$$

und entsprechend für $p = \infty$

$$\|\mathbf{f}\|_{\infty} := \max_{i=1, \dots, n} \|f_i\|_{\infty} \quad (10.33)$$

10.2.3 Normen von Eingrößensystemen

Die bisherigen Definitionen für Signale lassen sich leicht auf Systeme übertragen, indem man zunächst nur lineare Eingrößensysteme betrachtet, die durch ihre Übertragungsfunktion $G(s)$ bzw. ihren Frequenzgang $G(j\omega) = G(j2\pi f)$ gegeben sind. Der Parameter dieser Funktionen ist jetzt nicht mehr die Zeit t , sondern die Frequenz f bzw. die Kreisfrequenz ω .

In völliger Analogie der Definition (10.21) für die 2-Norm eines Signales ergibt sich für die 2-Norm eines linearen Eingrößensystems:

$$\|G\|_2 := \sqrt{\int_{-\infty}^{\infty} |G(jf)|^2 df} = \sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 d\omega} \quad (10.34)$$

und in entsprechender Analogie für die ∞ -Norm:

$$\|G\|_{\infty} := \sup_{\omega} |G(j\omega)| \quad (10.35)$$

Ein interessanter Aspekt hinsichtlich der 2-Norm ergibt sich aus dem Parseval'schen Theorem, welches besagt, dass die Energie eines Signales im Frequenz- und Zeitbereich gleich sein muss. Da die Übertragungsfunktion $G(s)$ gerade die Laplace-Transformierte der Impulsantwort $g(t)$ ist (vgl. Kap. 3.2), handelt es sich dabei um dasselbe Signal, einmal im Frequenz- und einmal im Zeitbereich. Daher muss die Energie von $G(s)$ bzw. $G(j\omega)$ die gleiche sein wie von $g(t)$:

$$\|G\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 d\omega} = \sqrt{\int_{-\infty}^{\infty} |g(t)|^2 dt} = \|g\|_2 \quad (10.36)$$

Die Konsequenz daraus ist, dass sich die 2-Norm eines Signales bzw. Systems alternativ sowohl im Frequenz- als auch im Zeitbereich berechnen lässt, was für keine andere Norm gilt.

Zur Erläuterung der ∞ -Norm sei auf Satz 4.1 mit Gleichung (4.1) verwiesen, die das Ausgangssignal für ein Eingangssignal der Form $x(t) = a \sin \omega t$ beschreibt:

$$y(t) = a |G(j\omega)| \sin(\omega t + \varphi(G(j\omega))) \quad (10.37)$$

Man erkennt, dass $|G(j\omega)|$ den Verstärkungsfaktor des Systems für Schwingungen der Frequenz ω darstellt. Damit ist die ∞ -Norm offensichtlich der maximal mögliche Verstärkungsfaktor vom Eingangs- zum Ausgangssignal hinsichtlich der Amplituden bzw. bei komplexer Schreibweise hinsichtlich der Länge der komplexen Ein- und Ausgangsvektoren.

Zudem ist die ∞ -Norm sub-multiplikativ:

$$\|GH\|_\infty \leq \|G\|_\infty \|H\|_\infty \quad (10.38)$$

wie man sich mit Hilfe eines Bode-Diagramms einer zusammengesetzten Übertragungsfunktion leicht selbst klarmachen kann.

Auch für Systemnormen ist deren Endlichkeit eine Voraussetzung dafür, dass mit Ihnen überhaupt gearbeitet werden kann. Um diese diskutieren zu können, ist zunächst die folgende Definition erforderlich:

Definition 10.6 $G(s)$ bezeichnet man als begrenzt, wenn $G(j\infty) < \infty$ gilt, und als streng begrenzt, wenn $G(j\infty) = 0$ gilt.

Bezogen auf die Ortskurve einer Übertragungsfunktion bedeutet dies, dass die Ortskurve einer begrenzten Funktion nicht im Unendlichen enden darf und die Ortskurve einer streng begrenzten Funktion im Nullpunkt enden muss. Sämtliche Übertragungsfunktionen, deren Nennerpolynom einen höheren Grad aufweist als ihr Zählerpolynom, sind demnach streng begrenzt. Einfach begrenzt ist die Übertragungsfunktion des PI-Reglers, bei dessen Übertragungsfunktion Nenner- und Zählerpolynom den gleichen Grad aufweisen. Und unbegrenzt ist die Übertragungsfunktion eines idealen PID-Reglers, da sein Zählerpolynom einen höheren Grad aufweist als sein Nennerpolynom.

Mit diesen Definitionen lassen sich nun zwei Sätze zur Endlichkeit von 2-Norm und ∞ -Norm von Übertragungsfunktionen angeben:

Satz 10.7 Die 2-Norm ist genau dann endlich, wenn G streng begrenzt ist und keine Pole auf der imaginären Achse aufweist.

Beweis: Wie Abbildung 10.3 verdeutlicht, lässt sich zu jeder streng begrenzten Strecke $G(s)$ ohne Pole auf der imaginären Achse, deren Ortskurve demnach auf der reellen Achse beginnt und im Ursprung endet, ein PT_1 -Glied mit der Übertragungsfunktion $V/(Ts + 1)$ und hinreichend großem V und hinreichend kleinem T finden, so dass für jeden Wert von ω gilt:

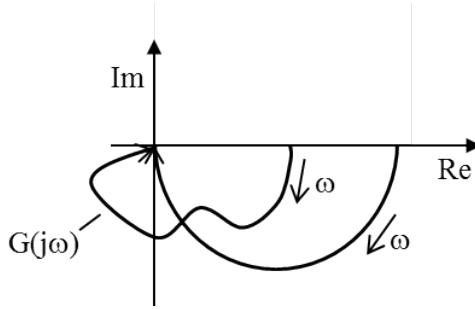


Abb. 10.3. Zum Beweis für die Endlichkeit der 2-Norm

$$\left| \frac{V}{Tj\omega + 1} \right| > |G(j\omega)| \quad (10.39)$$

Die 2-Norm des PT_1 -Gliedes soll gemäß Gleichung (10.36) im Zeitbereich mit Hilfe der Impulsantwort berechnet werden. Diese lautet für das PT_1 -Glied:

$$y(t) = \frac{V}{T} e^{-\frac{t}{T}} \quad (10.40)$$

Diese wird in die Definitionsgleichung eingesetzt, wobei die untere Integrationsgrenze zu Null gesetzt werden kann, da die Impulsantwort für $t < 0$ gleich Null ist:

$$\begin{aligned} \|y\|_2 &= \sqrt{\int_{-\infty}^{\infty} |y(t)|^2 dt} = \sqrt{\int_0^{\infty} \left(\frac{V^2}{T^2} e^{-\frac{2t}{T}} \right) dt} \\ &= \sqrt{\left(\frac{-V^2}{2T} e^{-\frac{2t}{T}} \right) \Big|_0^{\infty}} = \sqrt{\frac{V^2}{2T}} = \frac{V}{\sqrt{2T}} \end{aligned} \quad (10.41)$$

Die 2-Norm der Impulsantwort und damit wegen der Gleichheit auch die 2-Norm des PT_1 -Gliedes weisen demnach einen endlichen Wert auf. Wegen (10.39) muss dies auch für die gegebene beschränkte Strecke $G(s)$ ohne Pole auf der imaginären Achse gelten.

Ist andererseits $G(s)$ nicht streng begrenzt, so konvergiert $G(j\omega)$ für wachsendes ω nicht gegen Null. Bei der Berechnung der 2-Norm wird daher ein von Null verschiedener Wert bis ins Unendliche integriert, weshalb das Resultat nicht mehr endlich sein kann. Ein Pol auf der imaginären Achse führt bei der Berechnung der Norm zu einem unendlich großen Wert des Integranden, so dass auch in diesem Fall das Resultat nicht endlich sein kann. Damit ist der Satz bewiesen.

Der analoge Satz für die ∞ -Norm lautet:

Satz 10.8 *Die ∞ -Norm ist genau dann endlich, wenn G begrenzt ist und keine Pole auf der imaginären Achse aufweist.*

Beweis: Da die ∞ -Norm identisch mit dem Maximum des Betrages der Übertragungsfunktion ist, muss lediglich geprüft werden, ob $|G(j\omega)|$ für irgendeine Frequenz unendlich werden kann. Sofern G begrenzt ist, folgt die Endlichkeit von $|G(j\omega)|$ für $\omega = \infty$. Besitzt $G(s)$ auch keine Pole auf der imaginären Achse, so kann $|G(j\omega)|$ auch für alle anderen Frequenzen nicht unendlich groß werden.

Die Berechnung der 2-Norm mit Hilfe der Impulsantwort wie in Gleichung (10.41) ist durchaus möglich, aber bei komplexeren Übertragungsfunktionen und entsprechend komplexen Impulsantworten kann die Integration zum Problem werden. Daher soll im folgenden noch eine alternative Berechnungsmöglichkeit entwickelt werden.

Ausgangspunkt ist die quadrierte Definitionsgleichung für die 2-Norm (10.34):

$$\|G\|_2^2 := \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 d\omega \quad (10.42)$$

Für den quadrierten Absolutwert einer komplexen Zahl c gilt mit den reellen Zahlen a und b :

$$\begin{aligned} c &= a + jb \\ |c|^2 &= |a + jb|^2 = a^2 + b^2 = (a + jb)(a - jb) = c \bar{c} \end{aligned} \quad (10.43)$$

Dabei ist \bar{c} der komplex konjugierte Wert von c . Weiterhin gilt für rationale Übertragungsfunktionen mit reellwertigen Koeffizienten:

$$\overline{G(s)} = G(-s) \quad (10.44)$$

Einsetzen dieser Ergebnisse in (10.42) liefert

$$\|G\|_2^2 = \frac{1}{2\pi} \int_{-j\infty}^{j\infty} G(s)G(-s)ds \quad (10.45)$$

Neben der genannten Ersetzung des Integranden wurde die Integrationsvariable ω durch s ersetzt, was keine äquivalente Ersetzung ist, da s im Gegensatz zu ω eine komplexe Zahl ist. Aber dadurch, dass das Integrationsintervall ebenfalls verändert wurde, wird nach wie vor entlang der imaginären Achse von $-\infty$ bis ∞ integriert, so dass die Äquivalenz wieder hergestellt wurde.

Nun sei angenommen, dass der Wert der 2-Norm endlich ist. Daraus folgt, dass $G(s)$ streng begrenzt ist, weshalb alle Werte von $G(s)$ im Unendlichen Null sein müssen. Daher kann der Integrationsweg gemäß Abbildung 10.4 um einen Halbkreis im Unendlichen ergänzt werden, ohne den Wert des Integrales zu verändern.

Damit umschließt der Integrationsweg nun die linke Hälfte der komplexen Ebene, und aus dem Integral wird ein Ringintegral:

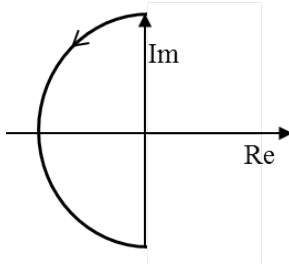


Abb. 10.4. Erweiterung des Integrationsweges

$$\|G\|_2^2 = \frac{1}{2\pi} \oint_{l.H.} G(s)G(-s)ds \quad (10.46)$$

Nach dem Residuensatz ist aber der Wert eines Ringintegrals gleich der Summe der Residuen aller Pole, die in der umschlossenen Fläche liegen. Damit gilt

$$\|G\|_2^2 = \sum_{\text{Pole in l.H.}} \text{Res}(G(s)G(-s)) \quad (10.47)$$

bzw. für die nicht-quadierte 2-Norm:

$$\|G\|_2 = \sqrt{\sum_{\text{Pole in l.H.}} \text{Res}(G(s)G(-s))} \quad (10.48)$$

Die Anwendung dieses Satzes soll an einem einfachen Beispiel verdeutlicht werden, und zwar an einem PT_1 -Glieder mit der Übertragungsfunktion $\frac{V}{(Ts+1)}$:

$$G(s)G(-s) = \frac{V}{Ts+1} \frac{V}{-Ts+1} = \frac{\frac{V^2}{2T}}{s + \frac{1}{T}} + \frac{\frac{V^2}{2T}}{-s + \frac{1}{T}} \quad (10.49)$$

Der einzige Pol in der linken Halbebene ist $s = -1/T$ und sein Residuum ist $V^2/2T$. Damit lautet das Resultat

$$\|G\|_2 = \sqrt{\frac{V^2}{2T}} = \frac{V}{\sqrt{2T}} \quad (10.50)$$

und man erhält die gleiche Lösung wie in (10.41).

Auch für die Berechnung der ∞ -Norm soll ein einfaches Beispiel vorgeführt werden. Grundsätzlich sind für den Frequenzgang $G(j\omega)$ entsprechend der Definition 10.35 seine maximalen Werte zu bestimmen und, sofern deren Bestimmung nicht exakt möglich ist, die kleinste obere Schranke festzulegen. Die entsprechenden Frequenzen findet man durch Ableitung und Nullsetzen des Frequenzganges:

$$\frac{d|G(j\omega)|}{d\omega} = 0 \quad (10.51)$$

In der Praxis findet man den Wert für die Norm aber in der Regel schneller durch Betrachtung des Betragsverlaufes im Bode-Diagramm. So ergibt sich für die Strecke

$$\frac{bs + 1}{as + 1} \quad (10.52)$$

mit $b > a$ der in Abbildung 10.5 schematisch gezeigte Betragsverlauf.

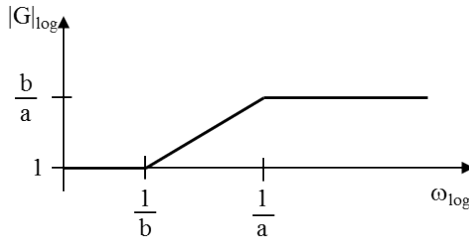


Abb. 10.5. Betragsverlauf für ein Übertragungsglied

Offensichtlich gilt $\|G\|_\infty = b/a$. Wäre $b < a$, dann hätte die ∞ -Norm den Wert Eins.

10.2.4 Beziehungen zwischen Systemgrößen

Äußerst relevant für den Reglerentwurf sind die Beziehungen zwischen den Normen der Ein- und Ausgangssignale und den Normen des Systems, die in diesem Kapitel der Einfachheit halber jedoch nur für Eingrößensysteme behandelt werden sollen. Dies ist allerdings auch hinreichend, um die für den Reglerentwurf wesentlichen Aspekte zu diskutieren.

Basis der Überlegungen ist die nachfolgende Tabelle:

	$\ u\ _2$	$\ u\ _\infty$	$pow(u)$
$\ y\ _2$	$\ G\ _\infty$	∞	∞
$\ y\ _\infty$	$\ G\ _2$	$\ g(t)\ _1$	∞
$pow(y)$	0	$\leq \ G\ _\infty$	$\ G\ _\infty$

Diese Tabelle stellt jeweils den Zusammenhang zwischen einer Norm des Eingangssignales u und einer Norm des Ausgangssignales y her. Dieser Zusammenhang lautet beispielsweise für den ersten Tabelleneintrag:

$$\|G\|_\infty = \sup_{u \neq 0} \frac{\|y\|_2}{\|u\|_2} \quad (10.53)$$

$\|G\|_\infty$ als erster Tabelleneintrag ist demnach der maximale Verstärkungsfaktor von der Energie des Eingangssignales $\|u\|_2$ zur Energie des Ausgangssignales $\|y\|_2$, während der darunter liegende Eintrag $\|G\|_2$ den maximalen

Verstärkungsfaktor von der Energie des Eingangssignales $\|u\|_2$ zur maximalen Amplitude des Ausgangssignales $\|y\|_\infty$ darstellt.

Eine typische Aufgabe einer Regelung ist die Unterdrückung von unbekannten Störgrößen, die, obwohl sie unbekannt sind, zumindest durch ihre Norm charakterisiert werden können. So gibt ja beispielsweise die ∞ -Norm eines Signales dessen maximale Amplitude an. Wenn nun G das Übertragungsverhalten von einer Störung zu einem Regelfehler beschreibt, ist die Norm dieser Übertragungsfunktion ein quantitativer Index für die erreichte Regelqualität.

Unter diesem Aspekt ist $\|G\|_\infty$ ein Maß dafür, wie stark sich eine dauerhafte Störung, die durch ihre Energie charakterisiert wird, auf die Energie des Regelfehlers und somit auf dessen Dauerverhalten auswirkt. Dagegen steht $\|G\|_2$ dafür, welche maximale Amplitude diese Störung am Ausgang hervorruft. Demnach charakterisiert $\|G\|_\infty$ eher die Qualität der Regelung bei Dauersignalverhalten und $\|G\|_2$ eher die Regelqualität bei kurzfristigem, transientem Verhalten. Dieser Aspekt wird später wichtig, wenn es um die Auswahl einer zu minimierenden Norm geht.

Nun sollen die Tabelleneinträge in kurzer Form bewiesen werden, wobei allerdings nicht gezeigt wird, dass die durch den Tabelleneintrag gegebene, obere Grenze tatsächlich auftritt, sondern nur, dass sie eingehalten wird. Bei der Numerierung wird die Tabelle als Matrix aufgefasst. Der erste Tabelleneintrag sei (1,1), der letzte (3,3).

Beweis Tabelleneintrag (1,1): Startpunkt der Überlegungen ist die 2-Norm des Ausgangssignales $\|y\|_2$, die wegen des Parseval'schen Theorems (10.36) in Zeit- und Frequenzbereich den gleichen Wert haben muss. Unter Anwendung von $y(s) = G(s)u(s)$ bzw. $y(j\omega) = G(j\omega)u(j\omega)$ ergibt sich:

$$\begin{aligned}
 \|y\|_2^2 &= \|y(t)\|_2^2 = \|y(j\omega)\|_2^2 \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |y(j\omega)|^2 d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)u(j\omega)|^2 d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 |u(j\omega)|^2 d\omega
 \end{aligned} \tag{10.54}$$

Aufgrund der Definition (10.35) der ∞ -Norm

$$\|G\|_\infty := \sup_{\omega} |G(j\omega)| \tag{10.55}$$

lässt sich abschätzen:

$$\|y\|_2^2 \leq \|G\|_\infty^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} |u(j\omega)|^2 d\omega = \|G\|_\infty^2 \|u\|_2^2 \quad (10.56)$$

Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (2,1): Hier ist zunächst eine Vorbetrachtung notwendig: Entsprechend Definition 10.1 bilden die integrierbaren Funktionen einen linearen Raum. Auf diesem Raum ist ein Skalarprodukt definiert:

$$(f, g) := \frac{1}{2T} \int_{-T}^T f(t) g(t) dt \quad (10.57)$$

Weiterhin gilt die Cauchy-Schwarz'sche Ungleichung für Skalarprodukte:

$$|(f, g)| \leq \sqrt{(f, f)} \sqrt{(g, g)} \quad (10.58)$$

Diese Ungleichung kann man beispielsweise für das Skalarprodukt einfacher Vektoren sofort nachvollziehen:

$$|\mathbf{x}\mathbf{y}| \leq \sqrt{\mathbf{x}\mathbf{x}} \sqrt{\mathbf{y}\mathbf{y}} = \sqrt{|\mathbf{x}|^2} \sqrt{|\mathbf{y}|^2} = |\mathbf{x}| |\mathbf{y}| \quad (10.59)$$

Für den Zusammenhang $y(s) = G(s)u(s)$ zwischen Ein- und Ausgangssignal im Frequenzbereich gilt im Zeitbereich gemäß Gleichung (3.18) das Faltungsintegral:

$$y(s) = G(s)x(s) \longleftrightarrow y(t) = \int_0^t g(t-\tau)u(\tau)d\tau \quad (10.60)$$

Unter Anwendung von (10.58) lässt sich schreiben:

$$\begin{aligned} |y(t)| &= \left| \int_0^t g(t-\tau)u(\tau)d\tau \right| \\ &\leq \sqrt{\int_0^t (g(t-\tau))^2 d\tau} \sqrt{\int_0^t (u(\tau))^2 d\tau} \\ &\leq \sqrt{\int_{-\infty}^{\infty} (g(\tau))^2 d\tau} \sqrt{\int_{-\infty}^{\infty} (u(\tau))^2 d\tau} \\ &= \|G\|_2 \|u\|_2 \end{aligned} \quad (10.61)$$

Wenn aber $|y(t)|$ für alle Werte von t kleiner als das Produkt der beiden 2-Normen ist, dann muss auch $\|y\|_\infty$ kleiner als dieses Produkt sein. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (3,1): Unter der Voraussetzung, dass die Systemnorm $\|G\|_\infty$ endlich ist, ergibt sich bei endlicher Energie des Eingangssignales $\|u\|_2 < \infty$ aus Tabelleneintrag (1,1), dass die Energie des Ausgangssignales endlich ist, da $\|G\|_\infty$ der maximale Verstärkungsfaktor von der Energie des Eingangssignales zur Energie des Ausgangssignales ist. Wenn aber $\|y\|_2 < \infty$ gilt, dann muss nach Satz 10.3 $\text{pow}(y) = 0$ gelten. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (1,2): Ein Signal mit endlicher ∞ -Norm ist beispielsweise $u(t) = \sin(\omega t)$, denn seine maximale Amplitude ist gleich Eins, und somit gilt $\|u\|_\infty = 1$. Eine Sinusschwingung am Eingang eines linearen Systems erzeugt am Ausgang ebenfalls eine Sinusschwingung. Die 2-Norm einer Sinusschwingung ist aber unendlich groß, da eine Sinusschwingung auch im Unendlichen nicht gegen Null konvergiert. Daher kann für den maximalen Verstärkungsfaktor von $\|u\|_\infty$ nach $\|y\|_2$ kein endlicher Wert angegeben werden. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (2,2): Mit dem Faltungsintegral gilt:

$$\begin{aligned}
 |y(t)| &= \left| \int_0^t g(t-\tau)u(\tau)d\tau \right| \\
 &\leq \int_0^t |g(t-\tau)| |u(\tau)|d\tau \\
 &\leq \int_0^t |g(t-\tau)|d\tau \|u\|_\infty \\
 &\leq \int_{-\infty}^{\infty} |g(\tau)|d\tau \|u\|_\infty \\
 &= \|g\|_1 \|u\|_\infty
 \end{aligned} \tag{10.62}$$

Wenn aber $|y(t)|$ für alle Werte von t kleiner als das Produkt der beiden Normen ist, dann muss auch $\|y\|_\infty$ kleiner als dieses Produkt sein. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (3,2): Der Beweis dieses Tabelleneintrages ergibt sich aus Eintrag (3,3), der nachfolgend noch bewiesen wird. Da $\|u\|_\infty \geq \text{pow}(u)$ gilt, muss, da die Ausgangsgröße in beiden Fällen $\text{pow}(y)$ und damit gleich ist, der Verstärkungsfaktor in Spalte 3 mindestens so groß sein wie der in Spalte 2. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (1,3): Ein Signal mit endlicher mittlerer Leistung $\text{pow}(u)$ ist beispielsweise $u(t) = \sin(\omega t)$. Eine Sinusschwingung am Eingang eines linearen Systems erzeugt am Ausgang ebenfalls eine Sinusschwingung. Die 2-Norm einer Sinusschwingung ist aber unendlich groß, da eine Sinusschwingung auch im Unendlichen nicht gegen Null konvergiert. Daher kann

für den maximalen Verstärkungsfaktor von $pow(u)$ nach $\|y\|_2$ kein endlicher Wert angegeben werden. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (2,3): Betrachtet sei als Beispiel ein Dirac-Impuls, dessen Fläche unter der Kurve gerade den Wert Eins aufweist (vgl. Gleichung (3.16)) und dessen Energie daher endlich sein muss. Dieser erzeugt beispielsweise am Ausgang eines einfachen Proportionalgliedes ebenfalls einen Dirac-Impuls, dessen ∞ -Norm aufgrund seiner unendlichen Amplitude unendlich groß ist. Daher kann für den maximalen Verstärkungsfaktor von $pow(u)$ nach $\|y\|_\infty$ kein endlicher Wert angegeben werden. Damit ist die Aussage bewiesen.

Beweis Tabelleneintrag (3,3): Die ∞ -Norm kennzeichnet das Maximum von $|G(j\omega)|$. Die größte mittlere Leistung am Ausgang ist bei Anregung der Strecke mit derjenigen Frequenz ω_m zu erwarten, bei der das Maximum von $|G(j\omega)|$ erreicht wird. Daher sei angenommen, dass die Strecke mit

$$u(t) = \sqrt{2} \sin(\omega_m t) \quad (10.63)$$

angeregt wird. Dieses Signal erzeugt gemäß Satz 4.1 am Ausgang des Systems die Schwingung

$$y(t) = |G(j\omega_m)| \sqrt{2} \sin(\omega_m t + \phi) \quad (10.64)$$

Für die mittlere Leistung des Eingangssignales gilt:

$$\begin{aligned} (pow(u))^2 &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (u(t))^2 dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (\sqrt{2} \sin(\omega_m t))^2 dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T \sin^2(\omega_m t) dt \end{aligned} \quad (10.65)$$

Beim Integral erkennt man, dass die Funktion $\sin^2(\omega_m t)$ über ein gewisses Intervall integriert und dann das Ergebnis durch die halbe Intervalllänge dividiert wird. Aufgrund der Periodizität des Signales erhält man das gleiche Ergebnis, wenn man den quadratischen Sinus nur über eine Halbperiode π integriert und anschließend durch $\frac{\pi}{2}$ dividiert, womit dann auch die Grenzwertbildung entfallen kann. Man erhält:

$$(pow(u))^2 = \frac{2}{\pi} \int_0^{\pi} \sin^2(\Theta) d\Theta = \frac{2}{\pi} \frac{\pi}{2} = 1 \quad (10.66)$$

Für die mittlere Leistung des Ausgangssignales ergibt sich:

$$\begin{aligned}
(pow(y))^2 &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (y(t))^2 dt \\
&= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (|G(j\omega_m)| \sqrt{2} \sin(\omega_m t + \phi))^2 dt \\
&= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |G(j\omega_m)|^2 2 \sin^2(\omega_m t + \phi) dt \\
&= |G(j\omega_m)|^2 \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T \sin^2(\omega_m t + \phi) dt \quad (10.67)
\end{aligned}$$

Aufgrund der Annahmen zu Beginn dieses Beweises gilt $|G(j\omega_m)| = \|G\|_\infty$. Und beim Integral sind die Verhältnisse exakt die gleichen wie in Gleichung (10.65), so dass sich auch hier für den zweiten Term (einschließlich der Grenzwertbildung) der Wert Eins ergibt. Letztendlich erhält man das Resultat

$$(pow(y))^2 = \|G\|_\infty^2 \quad (10.68)$$

Aus (10.66) und (10.68) ergibt sich für den Fall maximaler Systemverstärkung (bei der Frequenz ω_m) als Verstärkungsfaktor von der mittleren Leistung des Eingangssignales $pow(u)$ zur mittleren Leistung des Ausgangssignales $pow(y)$ gerade $\|G\|_\infty$. Damit ist die Aussage bewiesen.

10.2.5 Normen von Mehrgrößensystemen

Im Mehrgrößenfall wird aus der Übertragungsfunktion eine Übertragungsmatrix $\mathbf{G}(s)$ (vgl. Gleichung (7.107)), deren einzelne Koeffizienten aus Eingrößen-Übertragungsfunktionen $G_{ij}(s)$ bestehen.

2-Norm von Mehrgrößensystemen. Für die 2-Norm gibt es mehrere Möglichkeiten, die Definition (10.34) auf Mehrgrößensysteme zu erweitern. Dabei gilt auch hier das Parseval'sche Theorem (10.36), nach dem die Berechnung der 2-Norm sowohl im Zeit- als auch im Frequenzbereich den gleichen Wert liefern muss. Von daher kann man die Norm anhand der Eingrößen-Übertragungsfunktionen $G_{ij}(s)$ oder anhand der korrespondierenden Impulsantworten $g_{ij}(t)$ berechnen. Im zweiten Fall verwendet man statt der Übertragungsmatrix $\mathbf{G}(s)$ die Matrix der Impulsantworten $\mathbf{G}(t)$.

Mögliche Definitionen sind:

$$\|\mathbf{G}\|_2 := \sqrt{\text{Spur} \left[\int_{-\infty}^{\infty} \mathbf{G}(t) \mathbf{G}^T(t) dt \right]}$$

$$\begin{aligned}
\|\mathbf{G}\|_2 &:= \sqrt{\lambda_{max} \left[\int_{-\infty}^{\infty} \mathbf{G}(t) \mathbf{G}^T(t) dt \right]} \\
\|\mathbf{G}\|_2 &:= \sqrt{d_{max} \left[\int_{-\infty}^{\infty} \mathbf{G}(t) \mathbf{G}^T(t) dt \right]} \quad (10.69)
\end{aligned}$$

Dabei bezeichnet die Spur die Summe aller Diagonalelemente einer Matrix, λ_{max} ihren größten Eigenwert und d_{max} das maximale Element ihrer Hauptdiagonalen.

Da die erste Definition für den Entwurf normoptimaler Regelungen aus rechen-technischen Gründen am besten geeignet ist, soll sie im weiteren Verlauf verwendet werden. Für eine nähere Analyse dieser Definition wird zunächst das Matrizenprodukt betrachtet:

$$\begin{aligned}
\mathbf{G}\mathbf{G}^T &= \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots \\ g_{21} & g_{22} & g_{23} & \cdots \\ g_{31} & g_{32} & g_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} g_{11} & g_{21} & g_{31} & \cdots \\ g_{12} & g_{22} & g_{32} & \cdots \\ g_{13} & g_{23} & g_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \\
&= \begin{bmatrix} g_{11}^2 + g_{12}^2 + g_{13}^2 + \cdots & \cdots & \cdots \\ \cdots & g_{21}^2 + g_{22}^2 + g_{23}^2 + \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \quad (10.70)
\end{aligned}$$

Damit folgt für die Definition

$$\|\mathbf{G}\|_2 = \sqrt{\int_{-\infty}^{\infty} \sum_{ij} (g_{ij}(t))^2 dt} \quad (10.71)$$

und man erkennt, dass es sich bei dieser Definition um eine sehr plausible Erweiterung der 2-Norm für Eingrößensysteme (10.36) handelt, die den Eingrößenfall natürlich mit einschließt.

Ausgangspunkt der weiteren Überlegungen ist die Übertragungsmatrix nach Gleichung (7.107), wobei hier die Durchgangsmatrix \mathbf{D} zu Null angenommen wird:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \quad (10.72)$$

Da $(s\mathbf{I} - \mathbf{A})^{-1}$ die Laplace-Transformierte der Transitionsmatrix $e^{\mathbf{A}t}$ ist (vgl. Gleichung (7.104)), gilt für die Matrix der Impulsantworten

$$\mathbf{G}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B} \quad (10.73)$$

und für deren Transponierte

$$\mathbf{G}^T(t) = (\mathbf{C}e^{\mathbf{A}t}\mathbf{B})^T = \mathbf{B}^T e^{\mathbf{A}^T t} \mathbf{C}^T \quad (10.74)$$

Damit kann Gleichung (10.69) wie folgt geschrieben werden:

$$\begin{aligned}
\|\mathbf{G}\|_2 &= \sqrt{\text{Spur} \left[\int_{-\infty}^{\infty} \mathbf{G}(t) \mathbf{G}^T(t) dt \right]} \\
&= \sqrt{\text{Spur} \left[\int_{-\infty}^{\infty} \mathbf{C} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} \mathbf{C}^T dt \right]} \\
&= \sqrt{\text{Spur} \left[\mathbf{C} \int_{-\infty}^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt \mathbf{C}^T \right]} \\
&= \sqrt{\text{Spur} [\mathbf{C} \mathbf{G}_S \mathbf{C}^T]} \tag{10.75}
\end{aligned}$$

mit der sogenannten Steuerbarkeits-Gramschen

$$\mathbf{G}_S := \int_{-\infty}^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt \tag{10.76}$$

Es gilt

$$\begin{aligned}
\mathbf{A} \mathbf{G}_S + \mathbf{G}_S \mathbf{A}^T &= \\
\mathbf{A} \int_{-\infty}^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt + \int_{-\infty}^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt \mathbf{A}^T &= \\
\int_{-\infty}^{\infty} \mathbf{A} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt + \int_{-\infty}^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} \mathbf{A}^T dt &= \\
\int_{-\infty}^{\infty} (\mathbf{A} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} + e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} \mathbf{A}^T) dt &= \\
\int_{-\infty}^{\infty} \frac{d}{dt} (e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t}) dt &= e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} \Big|_{-\infty}^{\infty} \tag{10.77}
\end{aligned}$$

Gemäß der Stabilitätsdefinition 7.2 konvergiert bei stabilen Systemen ohne äußere Anregung der Zustandsvektor für $t \rightarrow \infty$ gegen Null. Wegen $\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}(0)$ muss daher für $t \rightarrow \infty$ auch die Transitionsmatrix $e^{\mathbf{A}t}$ gegen Null konvergieren, da der Anfangszustand ja von Null verschieden sein kann. Andererseits konvergiert wegen der Definition der Matrix-Exponentialfunktion die Transitionsmatrix für $t \rightarrow -\infty$ gegen die Einheitsmatrix. Damit ergibt sich

$$\mathbf{A}\mathbf{G}_S + \mathbf{G}_S\mathbf{A}^T = -\mathbf{B}\mathbf{B}^T \quad (10.78)$$

Dies ist gerade die Ljapunov-Gleichung (vgl. Satz A.2). Die Steuerbarkeits-Gramsche kann (für stabile Systeme) numerisch leicht aus dieser Ljapunov-Gleichung berechnet werden, und damit auch die 2-Norm gemäß Gleichung (10.75).

Die Spur von $\mathbf{G}(t)\mathbf{G}^T(t)$ ist mit der Spur von $\mathbf{G}^T(t)\mathbf{G}(t)$ identisch. Somit existiert eine duale Form zu (10.75) bzw. (10.76):

$$\begin{aligned} \|\mathbf{G}\|_2 &= \sqrt{\text{Spur} \left[\int_{-\infty}^{\infty} \mathbf{G}^T(t)\mathbf{G}(t)dt \right]} \\ &= \sqrt{\text{Spur} \left[\int_{-\infty}^{\infty} \mathbf{B}^T e^{\mathbf{A}^T t} \mathbf{C}^T \mathbf{C} e^{\mathbf{A} t} \mathbf{B} dt \right]} \\ &= \sqrt{\text{Spur} \left[\mathbf{B}^T \int_{-\infty}^{\infty} e^{\mathbf{A}^T t} \mathbf{C}^T \mathbf{C} e^{\mathbf{A} t} dt \mathbf{B} \right]} \\ &= \sqrt{\text{Spur} [\mathbf{B}^T \mathbf{G}_B \mathbf{B}]} \end{aligned} \quad (10.79)$$

mit der sogenannten Beobachtbarkeits-Gramschen

$$\mathbf{G}_B := \int_{-\infty}^{\infty} e^{\mathbf{A}^T t} \mathbf{C}^T \mathbf{C} e^{\mathbf{A} t} dt \quad (10.80)$$

Hier ergibt sich dann die Ljapunov-Gleichung

$$\mathbf{A}^T \mathbf{G}_B + \mathbf{G}_B \mathbf{A} = -\mathbf{C}^T \mathbf{C} \quad (10.81)$$

Um die 2-Norm einer Übertragungsmatrix zu berechnen, wird daher zunächst die Lösung einer der beiden Lyapunov-Gleichungen bestimmt (\mathbf{G}_S oder \mathbf{G}_B) und diese dann in (10.75) bzw. (10.79) eingesetzt.

∞ -Norm von Mehrgrößensystemen. Ausgangspunkt der Überlegungen sei die folgende Definition:

Definition 10.9 Als Singulärwertzerlegung einer komplexen $q \times p$ -Matrix \mathbf{M} mit Rang r bezeichnet man ein Produkt der Gestalt

$$\mathbf{M} = \mathbf{U} \Sigma \mathbf{V}^* \quad (10.82)$$

wobei \mathbf{U} eine unitäre $q \times q$ -Matrix ist ($\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$), \mathbf{V}^* die Adjungierte (= komplex-konjugiert Transponierte) einer unitären $p \times p$ -Matrix \mathbf{V} und

$$\Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \cdots & \\ & & & \sigma_r \\ \hline & & & 0 \\ & 0 & & 0 \end{array} \right) \quad (10.83)$$

eine reelle $p \times q$ -Matrix mit $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ ist. Die positiven Diagonaleinträge σ_i von Σ heißen Singulärwerte von \mathbf{M} . Die Singulärwerte und damit auch die Matrix Σ sind durch \mathbf{M} eindeutig bestimmt, nicht jedoch \mathbf{U} und \mathbf{V} , hier kann es mehrere Lösungen geben. Der maximale singuläre Wert σ_1 wird im folgenden mit $\bar{\sigma}$ bezeichnet.

Anzumerken ist, dass die Matrix Σ bei vollem Rang $r = \min(p, q)$ nicht gleichzeitig Nullspalten rechts und Nullzeilen unten haben kann. Je nachdem, ob $p > q$ oder $p < q$ gilt, tritt entweder nur die Nullspalte oder nur die Nullzeile auf. Ist \mathbf{M} quadratisch und regulär, sind alle Hauptdiagonalelemente von Σ von Null verschieden.

Nun sei \mathbf{M} die Übertragungsmatrix eines Systems, \mathbf{u} der Eingangsvektor und \mathbf{y} der zugehörige Ausgangsvektor. Damit gilt (unter Ausnutzung von $\mathbf{U}^* \mathbf{U} = \mathbf{I}$):

$$\begin{aligned} \mathbf{y} &= \mathbf{M} \mathbf{u} \\ &= \mathbf{U} \Sigma \mathbf{V}^* \mathbf{u} \\ \mathbf{y}^* \mathbf{y} &= \mathbf{u}^* \mathbf{V} \Sigma^* \mathbf{U}^* \mathbf{U} \Sigma \mathbf{V}^* \mathbf{u} \\ |\mathbf{y}|^2 &= \mathbf{u}^* \mathbf{V} \Sigma^* \Sigma \mathbf{V}^* \mathbf{u} \end{aligned} \quad (10.84)$$

Mit der abkürzenden Schreibweise $\mathbf{h} = \mathbf{V}^* \mathbf{u}$ wird daraus

$$\begin{aligned} |\mathbf{y}|^2 &= \mathbf{h}^* \Sigma^* \Sigma \mathbf{h} \\ &= \mathbf{h}^* \left(\begin{array}{ccc|c} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \cdots & \\ & & & \sigma_r^2 \\ \hline & & & 0 \\ & 0 & & 0 \end{array} \right) \mathbf{h} \\ &= h_1^2 \sigma_1^2 + h_2^2 \sigma_2^2 + \dots + h_r^2 \sigma_r^2 \\ &\leq h_1^2 \bar{\sigma}^2 + h_2^2 \bar{\sigma}^2 + \dots + h_r^2 \bar{\sigma}^2 \\ &= \bar{\sigma}^2 (h_1^2 + h_2^2 + \dots + h_r^2) \\ &\leq \bar{\sigma}^2 (h_1^2 + \dots + h_r^2 + \dots + h_n^2) \\ &= \bar{\sigma}^2 \mathbf{h}^* \mathbf{h} = \bar{\sigma}^2 \mathbf{u}^* \mathbf{V} \mathbf{V}^* \mathbf{u} = \bar{\sigma}^2 |\mathbf{u}|^2 \end{aligned} \quad (10.85)$$

$\bar{\sigma}$ ist offenbar der maximal mögliche Übertragungsfaktor des Systems \mathbf{M} vom Absolutwert des Eingangsvektors zum Absolutwert des Ausgangsvektors. Damit ist $\bar{\sigma}$ die konsequente Erweiterung der Definition (10.35) auf Mehrgrößensysteme, wobei zu berücksichtigen ist, dass sich für unterschiedliche Frequenzen auch unterschiedliche Singulärwerte ergeben:

$$\|\mathbf{M}\|_{\infty} := \sup_{\omega} \bar{\sigma}(\mathbf{M}(j\omega)) \quad (10.86)$$

Weiterhin gilt (auch hier wieder unter Ausnutzung von $\mathbf{U}^*\mathbf{U} = \mathbf{I}$)

$$\mathbf{M}^*\mathbf{M} = \mathbf{V}\Sigma^*\mathbf{U}^*\mathbf{U}\Sigma\mathbf{V}^* = \mathbf{V}\Sigma^*\Sigma\mathbf{V}^* \quad (10.87)$$

Vergleicht man diese Gleichung mit dem Zusammenhang bei einer Basis-transformation $\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ (Gleichung (7.63)), so stellt man fest, dass die Struktur hier dieselbe ist, denn da \mathbf{V} eine unitäre Matrix ist, muss \mathbf{V}^* ihre Inverse sein. Da es sich in beiden Fällen um eine Ähnlichkeitstransformation handelt, sind die Eigenwerte von \mathbf{A}' und \mathbf{A} gleich, und ebenso die Eigenwerte von $\mathbf{M}^*\mathbf{M}$ und $\Sigma^*\Sigma$.

Da $\Sigma^*\Sigma$ eine Diagonalmatrix ist, sind ihre Eigenwerte offensichtlich gerade $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$, und somit gilt

$$\begin{aligned} \sigma_i^2 &= \lambda_i(\Sigma^*\Sigma) = \lambda_i(\mathbf{M}^*\mathbf{M}) \\ \bar{\sigma} &= \sqrt{\lambda_{max}(\mathbf{M}^*\mathbf{M})} \end{aligned} \quad (10.88)$$

Auch für diese Norm soll abschließend noch eine mögliche Berechnungsmethode erläutert werden: Als abkürzende Schreibweise soll dabei im folgenden

$$\gamma := \|\mathbf{M}\|_{\infty} = \sup_{\omega} \bar{\sigma}(\mathbf{M}(j\omega)) \quad (10.89)$$

verwendet werden. Der gesuchte, zunächst noch unbekannte Wert ist also γ .

Aus Definition (10.86) folgt, dass eine Frequenz ω_1 existieren muss, für die der maximale singuläre Wert von $\mathbf{M}(j\omega)$ sein Maximum annimmt, der dem Wert von $\|\mathbf{M}\|_{\infty}$ entspricht:

$$\gamma = \bar{\sigma}(\mathbf{M}(j\omega_1)) \quad (10.90)$$

Weiterhin folgt aus (10.88)

$$\gamma = \sqrt{\lambda_{max}(\mathbf{M}^*(j\omega_1)\mathbf{M}(j\omega_1))} \quad (10.91)$$

Zum Eigenwert λ_{max} muss ein Eigenvektor \mathbf{v} existieren, für den der Zusammenhang

$$\lambda_{max}\mathbf{v} = (\mathbf{M}^*(j\omega_1)\mathbf{M}(j\omega_1))\mathbf{v} \quad (10.92)$$

aus Eigenvektor und Eigenwert bzw.

$$\gamma^2\mathbf{v} = (\mathbf{M}^*(j\omega_1)\mathbf{M}(j\omega_1))\mathbf{v} \quad (10.93)$$

gilt. Diese Gleichung lässt sich leicht umformen zu

$$(\gamma^2 \mathbf{I} - \mathbf{M}^*(j\omega_1) \mathbf{M}(j\omega_1)) \mathbf{v} = \mathbf{0} \quad (10.94)$$

Man erhält ein lineares Gleichungssystem für den Vektor \mathbf{v} , welches, sofern die Koeffizientenmatrix $\gamma^2 \mathbf{I} - \mathbf{M}^*(j\omega_1) \mathbf{M}(j\omega_1)$ regulär ist, nur die triviale Lösung $\mathbf{v} = \mathbf{0}$ besitzt. Da \mathbf{v} aber der zu λ_{max} gehörende Eigenvektor ist und damit von Null verschieden sein muss, kann die Koeffizientenmatrix nicht regulär sein, d.h. ihre Determinante ist gleich Null:

$$\det(\gamma^2 \mathbf{I} - \mathbf{M}^*(j\omega_1) \mathbf{M}(j\omega_1)) = 0 \quad (10.95)$$

Die Berechnung der Inversen einer frequenzabhängigen Übertragungsmatrix $\mathbf{H}(s)$ kann nach der Cramer'schen Regel (7.111) erfolgen:

$$\mathbf{H}^{-1}(s) = \frac{1}{\det(\mathbf{H}(s))} \begin{bmatrix} H_{11}(s) & H_{12}(s) & \cdots \\ H_{21}(s) & H_{22}(s) & \cdots \\ \dots\dots\dots \end{bmatrix}^T \quad (10.96)$$

Dabei sind die Funktionen $H_{ij}(s)$ die Adjunkten von $\mathbf{H}(s)$ gemäß (7.112).

Man erkennt, dass alle Übertragungsfunktionen der inversen Matrix $\mathbf{H}^{-1}(s)$ als Nenner die Determinante von $\mathbf{H}(s)$ aufweisen. Wenn diese Determinante für eine bestimmte Frequenz ω_1 Null wird, dann müssen $\mathbf{H}^{-1}(s)$ bzw. ihre Übertragungsfunktionen an der Stelle $s = j\omega_1$, d.h. auf der imaginären Achse, eine Polstelle und damit einen Eigenwert haben.

Überträgt man das eben Gesagte auf die Matrix $(\gamma^2 \mathbf{I} - \mathbf{M}^*(s) \mathbf{M}(s))$, so muss deren Inverse einen Eigenwert auf der imaginären Achse haben, und zwar bei $s = j\omega_1$.

Nun sei die Übertragungsmatrix $\mathbf{M}(s)$ durch ihre Zustandsmatrizen $[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}]$ gegeben. Dann ergibt sich nach einer längeren Rechnung für die Systemmatrix von $(\gamma^2 \mathbf{I} - \mathbf{M}^*(s) \mathbf{M}(s))^{-1}$:

$$\begin{bmatrix} \mathbf{A} + \mathbf{B} \mathbf{R}^{-1} \mathbf{D}^T \mathbf{C} & \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \\ -\mathbf{C}^T \mathbf{C} - \mathbf{C}^T \mathbf{D} \mathbf{R}^{-1} \mathbf{D}^T \mathbf{C} & -\mathbf{A}^T - \mathbf{C}^T \mathbf{D} \mathbf{R}^{-1} \mathbf{B}^T \end{bmatrix} \quad (10.97)$$

mit $\mathbf{R} = \gamma^2 \mathbf{I} - \mathbf{D}^T \mathbf{D}$.

Aus dieser Systemmatrix lässt sich nun der gesuchte Wert von γ und damit die ∞ -Norm von $\mathbf{M}(s)$ iterativ bestimmen, und zwar folgendermaßen: Zunächst ist klar, dass die Werte von \mathbf{R} unendlich groß werden müssen, wenn γ unendlich groß wird. Entsprechend konvergiert \mathbf{R}^{-1} in dem Fall gegen die Nullmatrix.

Daraus folgt aber wiederum, dass die Systemmatrix aus (10.97) für unendlich große Werte von γ gegen

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C}^T \mathbf{C} & -\mathbf{A}^T \end{bmatrix} \quad (10.98)$$

konvergieren muss.

Zur Berechnung der Eigenwerte dieser Matrix berechnet man die Nullstellen der Gleichung

$$\begin{aligned} 0 &= \det \left(s\mathbf{I} - \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C}^T \mathbf{C} & -\mathbf{A}^T \end{bmatrix} \right) \\ &= \det(s\mathbf{I} - \mathbf{A}) \det(s\mathbf{I} + \mathbf{A}^T) \end{aligned} \quad (10.99)$$

Da $-\mathbf{A}^T$ die gleichen Eigenwerte aufweist wie $-\mathbf{A}$, entsprechen die Eigenwerte der Matrix (10.98) denen von \mathbf{A} und $-\mathbf{A}$. Setzt man nun voraus, dass $\mathbf{M}(s)$ ein geregeltes System ist, so müssen alle Eigenwerte seiner Systemmatrix \mathbf{A} in der linken Hälfte der komplexen Ebene liegen und alle Eigenwerte von $-\mathbf{A}$ entsprechend in der rechten Hälfte. Kein Eigenwert der Matrix (10.98) liegt also auf der imaginären Achse.

Nun handelt es sich bei dieser Matrix aber gemäß Herleitung (10.97) um die Systemmatrix von $(\gamma^2 \mathbf{I} - \mathbf{M}^*(s)\mathbf{M}(s))^{-1}$ für unendlich große Werte von γ . Verkleinert man γ , so muss nach den vorhergehenden Betrachtungen irgendwann ein Eigenwert auf die imaginäre Achse wandern.

Genau diese Rechnung wird numerisch durchgeführt: Man berechnet die Systemmatrix von $(\gamma^2 \mathbf{I} - \mathbf{M}^*(s)\mathbf{M}(s))^{-1}$ und nimmt dabei zunächst einen sehr großen Wert für γ an. Dann berechnet man die Eigenwerte und prüft, ob von diesen keiner auf der imaginären Achse liegt. In den nachfolgenden Iterationen wird γ sukzessive verkleinert, und die Eigenwerte der Systemmatrix werden mit diesem Wert berechnet. Dies wird so lange durchgeführt, bis zum ersten Mal ein Eigenwert auf der imaginären Achse auftritt. Dann bildet der erreichte Wert von γ gerade den gesuchten Wert der ∞ -Norm von $\mathbf{M}(s)$ entsprechend Gleichung (10.89), und der Eigenwert selber bezeichnet diejenige Frequenz ω_1 , bei der der maximale singuläre Wert auftritt.

10.3 2-Norm-optimaler Regler für Eingrößensysteme

Anhand des Entwurfes von 2-Norm-optimalen Reglern für Eingrößensysteme sollen nun die Überlegungen und Schritte gezeigt werden, die für den Entwurf einer normoptimalen Regelung grundsätzlich erforderlich sind.

10.3.1 Koprime Faktorisierung

Definition 10.10 Die Menge ϕ sei die Menge aller stabilen, reell rationalen, begrenzten Übertragungsfunktionen.

Das Nennerpolynom einer solchen Übertragungsfunktion hat also mindestens die gleiche Ordnung wie das Zählerpolynom, und seine sämtlichen Nullstellen müssen in der linken Halbebene liegen.

ϕ ist offensichtlich abgeschlossen bezüglich der Addition und Multiplikation:

$$N, M \in \phi \rightarrow (N + M), (NM) \in \phi \quad (10.100)$$

Satz 10.11 *Jede reell rationale Übertragungsfunktion $G(s)$, die nicht notwendigerweise stabil oder begrenzt sein muss, lässt sich durch zwei Übertragungsfunktionen $N(s), M(s) \in \phi$ darstellen:*

$$G(s) = \frac{N(s)}{M(s)} \text{ mit } N(s), M(s) \in \phi \quad (10.101)$$

Der Beweis ist trivial: Beispielsweise kann die instabile Übertragungsfunktion $G(s) = \frac{s-2}{s-1}$ als Quotient der Übertragungsfunktionen $N(s) = \frac{s-2}{s+1}$ und $M(s) = \frac{s-1}{s+1}$ dargestellt werden, die beide Elemente von ϕ sind. Nach demselben Schema lässt sich offenbar jede reell rationale Übertragungsfunktion als Quotient zweier stabiler Übertragungsfunktionen darstellen.

Wählt man für den Regler die gleiche Darstellung

$$K(s) = \frac{X(s)}{Y(s)} \text{ mit } X(s), Y(s) \in \phi \quad (10.102)$$

so ergibt sich beispielsweise für die Empfindlichkeitsfunktion als eine mögliche Übertragungsfunktion im geschlossenen Kreis gemäß Gleichung (10.4)

$$S(s) = \frac{1}{1 + G(s)K(s)} = \frac{1}{1 + \frac{N(s)}{M(s)} \frac{X(s)}{Y(s)}} = \frac{M(s)Y(s)}{N(s)X(s) + M(s)Y(s)} \quad (10.103)$$

Da alle Übertragungsfunktionen im geschlossenen Kreis den gleichen Nenner $1 + GK$ bzw. $NX + MY$ aufweisen, ergeben sich denkbar einfache Verhältnisse, wenn X und Y so konstruiert werden können, dass

$$N(s)X(s) + M(s)Y(s) = 1 \quad (10.104)$$

gilt, denn dann kann der Nenner in all diesen Übertragungsfunktionen keine Instabilitäten mehr verursachen. Übrig bleibt nur der Zähler, der aber als Produkt zweier Funktionen aus ϕ auf jeden Fall stabil ist. Und dies gilt nicht nur für die hier gewählte Empfindlichkeitsfunktion, sondern für alle denkbaren Übertragungsfunktionen im geschlossenen Kreis, wie im nachfolgenden Kapitel noch gezeigt wird.

Damit ist aber auch jede Übertragungsfunktion im geschlossenen Kreis stabil. Gelingt es also, zu gegebenen Übertragungsfunktionen $N(s)$ und $M(s)$ zwei Funktionen $X(s)$ und $Y(s)$ aus ϕ zu finden, so dass Gleichung (10.104) gilt, dann ist der Regler $K(s)$ aus Gleichung (10.102) ein stabilisierender Regler für $G(s)$.

Definition 10.12 Man bezeichnet zwei Funktionen $N(s)$ und $M(s)$ als *koprim*, wenn es zwei Funktionen $X(s)$ und $Y(s)$ gibt, so dass $N(s)X(s) + M(s)Y(s) = 1$ gilt. Daraus folgt, dass $X(s)$ und $Y(s)$ ebenfalls *koprim* sind.

Koprime Funktionen, also $N(s)$ und $M(s)$ bzw. $X(s)$ und $Y(s)$, haben keine gemeinsamen Nullstellen, da anderenfalls die Bedingung (10.104) nicht erfüllt wäre:

$$N(s_0)X(s_0) + M(s_0)Y(s_0) = 0 \neq 1 \quad (10.105)$$

Koprime Funktionen sind also immer teilerfremd!

Weiterhin sei angemerkt, dass Definition 10.12 nicht voraussetzt, dass die genannten Funktionen aus ϕ sind. Die Koprim-Eigenschaft wird im folgenden nicht nur für Funktionen aus ϕ , sondern auch für Polynome, die ja wegen fehlender Begrenztheit nicht in ϕ enthalten sind, genutzt.

Die Frage ist nun, wie zu gegebenen Funktionen $N(s)$ und $M(s)$ die entsprechenden Funktionen $X(s)$ und $Y(s)$ berechnet werden können. Hierzu bietet sich der Algorithmus von Euklid an. Ausgangspunkt sind zwei Polynome (also keine Übertragungsfunktionen) $n(\lambda)$ und $m(\lambda)$, die keine gemeinsamen Nullstellen aufweisen. Dann berechnet man im ersten Schritt durch eine Division des größeren Polynoms $n(\lambda)$ durch das kleinere Polynom $m(\lambda)$ die Polynome $q_1(\lambda)$ und $r_1(\lambda)$. Anschließend ermittelt man aus einer weiteren Polynomdivision von $m(\lambda)$ durch $r_1(\lambda)$ die Polynome $q_2(\lambda)$ und $r_2(\lambda)$ usw., wie aus den folgenden Gleichungen nachzuvollziehen ist:

$$\begin{aligned} n(\lambda) &= m(\lambda)q_1(\lambda) + r_1(\lambda) \\ m(\lambda) &= r_1(\lambda)q_2(\lambda) + r_2(\lambda) \\ r_1(\lambda) &= r_2(\lambda)q_3(\lambda) + r_3(\lambda) \\ r_2(\lambda) &= r_3(\lambda)q_4(\lambda) + r_4(\lambda) \\ &\dots \end{aligned} \quad (10.106)$$

Der Algorithmus wird abgebrochen, sobald sich aus einer Division ein Restpolynom $r_k(\lambda) = \text{const.}$ ergibt.

In Matrizenschreibweise lässt sich das Gleichungssystem beschreiben durch:

$$\begin{bmatrix} 1 & 0 & \dots & & 0 \\ q_2 & 1 & 0 & \dots & 0 \\ -1 & q_3 & 1 & 0 & \dots & 0 \\ 0 & -1 & q_4 & 1 & 0 & \dots & 0 \\ & & & \dots & & & \\ 0 & \dots & & -1 & q_k & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ \dots \\ r_k \end{bmatrix} = \begin{bmatrix} 1 & -q_1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ \dots & \dots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} n \\ m \end{bmatrix} \quad (10.107)$$

Das Gleichungssystem lässt sich nach r_k auflösen, und man erhält:

$$r_k = (\dots)n + (\dots)m \quad (10.108)$$

Da r_k im vorliegenden Fall eine von Null verschiedene Konstante ist, ergibt die Division durch r_k :

$$\begin{aligned}
1 &= \frac{1}{r_k}(\dots)n + \frac{1}{r_k}(\dots)m \\
1 &= x(\lambda)n(\lambda) + y(\lambda)m(\lambda)
\end{aligned} \tag{10.109}$$

Damit sind zwei Polynome $x(\lambda)$ und $y(\lambda)$ gefunden, mit denen Gleichung (10.109) die gesuchte Struktur aus Gleichung (10.104) aufweist.

Nun ist dieser Algorithmus auf stabile, begrenzte Funktionen $N(s), M(s) \in \phi$ zu übertragen. Dies kann geschehen, indem man das Argument λ in $n(\lambda)$ und $m(\lambda)$ durch $\frac{1}{s+1}$ ersetzt. Dadurch wird aus jedem Polynom eine Übertragungsfunktion mit Polstellen bei $s = -1$ und gleicher Ordnung im Nenner und im Zähler. Somit entsteht eine Funktion aus ϕ .

Um vorher aus $G(s)$ die nach dem Algorithmus von Euklid zu dividierenden Polynome zu gewinnen, muss die inverse Ersetzung $s \rightarrow \frac{1-\lambda}{\lambda}$ erfolgen.

Es ergibt sich der folgende Algorithmus für die kopprime Faktorisierung:

1. Falls $G(s)$ eine stabile Funktion ist, dann setze $N := G$, $M := 1$, $X := 0$, $Y := 1$ und fahre fort mit Schritt 5.
2. Ersetze in $G(s)$ den Parameter s durch $\frac{1-\lambda}{\lambda}$. Man erhält $G(\lambda) = \frac{n(\lambda)}{m(\lambda)}$.
3. Berechne mit Hilfe des Algorithmus von Euklid aus $n(\lambda)$ und $m(\lambda)$ die gesuchten Polynome $x(\lambda)$ und $y(\lambda)$.
4. Ersetze in $x(\lambda)$ und $y(\lambda)$ den Parameter λ durch $\frac{1}{s+1}$. Es ergeben sich rationale, stabile Übertragungsfunktionen $X(s)$ und $Y(s)$, deren Nennerpolynome mindestens den gleichen Grad aufweisen wie ihre Zählerpolynome und die damit aus ϕ sind.
5. Ein stabilisierender Regler ergibt sich aus $K(s) = \frac{X(s)}{Y(s)}$. Bei $X(s) = 0$ (aus Punkt 1) bleibt die Strecke ungeregelt, aber da sie von sich aus schon stabil ist, ist auch kein stabilisierender Regler erforderlich.

Es stellt sich die Frage, warum man überhaupt den Umweg über λ bzw. die Polynome $n(\lambda)$ und $m(\lambda)$ gehen muss, anstatt die Polynomdivision direkt mit dem Zähler und Nenner von $G(s)$ durchzuführen. Die Antwort ist recht einfach: Man würde dann ja auch für den Regler nur das entsprechende Zähler- und Nennerpolynom erhalten. Polynome sind aber nicht begrenzt und damit auch keine Funktionen aus ϕ , was im folgenden jedoch benötigt wird, um endliche Werte für die zu optimierenden 2-Normen zu erhalten.

Nachfolgend soll ein Beispiel zur Berechnung eines stabilisierenden Reglers angegeben werden. Die Übertragungsfunktion der Strecke lautet:

$$G(s) = \frac{1}{(s-1)(s-2)} \tag{10.110}$$

Ersetzen von s durch $\frac{1-\lambda}{\lambda}$ führt auf

$$\begin{aligned}
G(\lambda) &= \frac{1}{\left(\frac{1-\lambda}{\lambda} - 1\right)\left(\frac{1-\lambda}{\lambda} - 2\right)} \\
&= \frac{\lambda^2}{(1-2\lambda)(1-3\lambda)}
\end{aligned}$$

$$= \frac{\lambda^2}{6\lambda^2 - 5\lambda + 1} = \frac{n(\lambda)}{m(\lambda)} \quad (10.111)$$

Nun erfolgen zwei Polynomdivisionen:

$$\begin{aligned} n &= \lambda^2 = (6\lambda^2 - 5\lambda + 1) \frac{1}{6} + \left(\frac{5}{6}\lambda - \frac{1}{6}\right) = mq_1 + r_1 \\ m &= (6\lambda^2 - 5\lambda + 1) = \left(\frac{5}{6}\lambda - \frac{1}{6}\right) \left(\frac{36}{5}\lambda - \frac{114}{25}\right) + \frac{6}{25} = r_1q_2 + r_2 \end{aligned} \quad (10.112)$$

Man erkennt, dass wegen $r_2 = \text{const.}$ die Rechnung schon beendet ist und man das Gleichungssystem nach r_2 auflösen kann:

$$r_2 = m - r_1q_2 = m - (n - mq_1)q_2 = (1 + q_1q_2)m - q_2n \quad (10.113)$$

Daraus folgen die Lösungen

$$\begin{aligned} x(\lambda) &= -\frac{q_2}{r_2} = -30\lambda + 19 \\ y(\lambda) &= \frac{1 + q_1q_2}{r_2} = 5\lambda + 1 \end{aligned} \quad (10.114)$$

Ersetzt man nun in n, m, x und y das Argument λ durch $\frac{1}{s+1}$, so erhält man die gesuchte koprime Faktorisierung:

$$\begin{aligned} N(s) &= \frac{1}{(s+1)^2} & M(s) &= \frac{(s-1)(s-2)}{(s+1)^2} \\ X(s) &= \frac{19s-11}{s+1} & Y(s) &= \frac{s+6}{s+1} \end{aligned} \quad (10.115)$$

Die Gleichung $N(s)X(s) + M(s)Y(s) = 1$ ist für beliebige s immer erfüllt, und der Regler

$$K(s) = \frac{X(s)}{Y(s)} = \frac{19s-11}{s+6} \quad (10.116)$$

ist in der Lage, die Strecke $G(s)$ zu stabilisieren.

10.3.2 Die Menge aller G stabilisierenden Regler

Im vorhergehenden Kapitel ist die koprime Faktorisierung anhand der Empfindlichkeitsfunktion motiviert worden, um einen stabilisierenden Regler berechnen zu können. In diesem Kapitel soll nun eine Verallgemeinerung in mehrfacher Hinsicht erfolgen. Zunächst soll der Begriff der internen Stabilität hergeleitet werden.

Grundlage der Überlegungen ist der geschlossene Kreis, ergänzt um die Dynamik des Messgliedes sowie alle denkbaren Eingangsgrößen (Abbildung 10.6). Der Systematik halber sind einige interne Größen mit h_i bezeichnet.

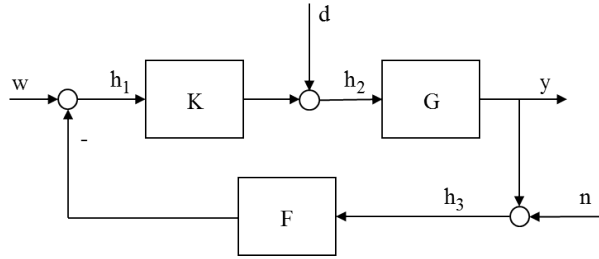


Abb. 10.6. Regelkreis mit Messglied und externen Eingängen

Die denkbaren Eingangsgrößen sind der Sollwert w , eine Streckenstörung d sowie ein Messrauschen n .

Für die internen Größen gilt

$$\begin{aligned} h_1 &= w - Fh_3 \\ h_2 &= d + Kh_1 \\ h_3 &= n + Gh_2 \end{aligned} \quad (10.117)$$

bzw. in Matrizenform

$$\begin{bmatrix} 1 & 0 & F \\ -K & 1 & 0 \\ 0 & -G & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = \begin{bmatrix} w \\ d \\ n \end{bmatrix} \quad (10.118)$$

Auflösen des Gleichungssystem nach den internen Größen liefert deren Abhängigkeiten von den externen Eingangsgrößen:

$$\begin{aligned} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & F \\ -K & 1 & 0 \\ 0 & -G & 1 \end{bmatrix}^{-1} \begin{bmatrix} w \\ d \\ n \end{bmatrix} \\ &= \frac{1}{1 + FGK} \begin{bmatrix} 1 & -GF & -F \\ K & 1 & -KF \\ GK & G & 1 \end{bmatrix} \begin{bmatrix} w \\ d \\ n \end{bmatrix} \end{aligned} \quad (10.119)$$

Mit dieser Gleichung kann nun die interne Stabilität definiert werden:

Definition 10.13 *Ein System ist intern stabil, wenn die 9 Übertragungsfunktionen in Gleichung (10.119) stabil sind.*

Man erkennt, dass in Gleichung (10.119) alle bisher verwendeten Übertragungsfunktionen enthalten sind, wie z.B. die Führungs-Übertragungsfunktion von w nach $h_3 (= y)$ oder die Stör-Übertragungsfunktion von d nach h_3 . Man sieht, dass alle Übertragungsfunktionen den gleichen Nenner haben, und man erkennt auch, dass die drei Fälle in der Hauptdiagonalen exakt die gleiche Übertragungsfunktion aufweisen, nämlich jeweils die Empfindlichkeitsfunktion. Diese ist demnach immer genau die Übertragungsfunktion von einer

externen Eingangsgröße zu der direkt an der Angriffsstelle liegenden internen Systemgröße.

Im weiteren Verlauf soll die Dynamik des Messgliedes vernachlässigt werden, d.h. $F(s) := 1$.

Mit der Definition für interne Stabilität lässt sich nun ein wichtiger Satz angeben:

Satz 10.14 *Mit der Strecke $G(s) = \frac{N(s)}{M(s)}$ und den Funktionen $N(s), M(s), X(s), Y(s) \in \phi$ mit der koprimen Faktorisierung $NX + MY = 1$ und beliebig gewähltem $Q(s) \in \phi$ lautet die Menge aller Regler für interne Stabilität:*

$$K(s) = \frac{X(s) + M(s)Q(s)}{Y(s) - N(s)Q(s)} \quad (10.120)$$

Zunächst soll bewiesen werden, dass Regler gemäß Gleichung (10.120) interne Stabilität gewährleisten. Dies kann einfach dadurch erfolgen, indem (10.120) in (10.119) eingesetzt wird, wobei es aus Gründen der Übersichtlichkeit Sinn macht, erst einmal nur den Term $\frac{1}{1+GK}$ zu betrachten. Unter Ausnutzung der Definitionsgleichung für die koprimen Faktorisierung erhält man:

$$\begin{aligned} \frac{1}{1+GK} &= \frac{1}{1 + \frac{N}{M} \frac{X+MQ}{Y-NQ}} \\ &= \frac{M(Y-NQ)}{M(Y-NQ) + N(X+MQ)} \\ &= \frac{M(Y-NQ)}{MY + NX} = M(Y-NQ) \end{aligned} \quad (10.121)$$

Mit diesem Resultat ergibt sich beim Einsetzen von (10.120) in (10.119):

$$\begin{aligned} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} &= \frac{1}{1+GK} \begin{bmatrix} 1 & -G & -1 \\ K & 1 & -K \\ GK & G & 1 \end{bmatrix} \begin{bmatrix} w \\ d \\ n \end{bmatrix} \\ &= M(Y-NQ) \begin{bmatrix} 1 & -\frac{N}{M} & -1 \\ \frac{X+MQ}{Y-NQ} & 1 & -\frac{X+MQ}{Y-NQ} \\ \frac{N}{M} \frac{X+MQ}{Y-NQ} & \frac{N}{M} & 1 \end{bmatrix} \begin{bmatrix} w \\ d \\ n \end{bmatrix} \\ &= \begin{bmatrix} M(Y-NQ) & -N(Y-NQ) & -M(Y-NQ) \\ M(X+MQ) & M(Y-NQ) & -M(X+MQ) \\ N(X+MQ) & N(Y-NQ) & M(Y-NQ) \end{bmatrix} \begin{bmatrix} w \\ d \\ n \end{bmatrix} \end{aligned} \quad (10.122)$$

Man erkennt, dass sämtliche internen Übertragungsfunktionen aus Summen oder Hintereinanderschaltungen von Funktionen aus ϕ bestehen und

damit stabil sind. Dies gilt allerdings nur dann, wenn neben den Funktionen N , M , X und Y auch Q stabil ist, womit begründet ist, warum auch $Q \in \phi$ in Satz 10.14 vorausgesetzt werden muss.

Weiterhin ist zu zeigen, dass es keine anderen stabilisierenden Regler geben kann außer denen, die durch (10.120) gegeben sind. Hierzu sei angenommen, dass der Regler durch $K(s) = \frac{N_k(s)}{M_k(s)}$ gegeben ist, wobei N_k und M_k teilerfremd und Elemente aus ϕ seien, ansonsten aber beliebig gewählt werden können. Damit ist wegen Satz 10.11 zunächst einmal jeder beliebige Regler möglich.

Einsetzen dieser Definition sowie von $G = \frac{N}{M}$ führt auf

$$\frac{1}{GK + 1} = \frac{1}{\frac{N}{M} \frac{N_k}{M_k} + 1} = \frac{MM_k}{NN_k + MM_k} \quad (10.123)$$

Da $NN_k + MM_k$ der gemeinsame Nenner aller Übertragungsfunktionen im geschlossenen Kreis ist, kann interne Stabilität nur dann gelten, wenn

$$V := \frac{1}{NN_k + MM_k} \quad (10.124)$$

eine stabile Funktion ist. Aus dieser Gleichung folgt

$$NN_kV + MM_kV = 1 \quad (10.125)$$

Weiterhin folgt aus $NX + MY = 1$:

$$\begin{aligned} 1 &= NX + MY = NX + MY + NMQ - NMQ \\ &= N(X + MQ) + M(Y - NQ) \end{aligned} \quad (10.126)$$

Nun sei Q die Lösung aus $M_kV = Y - NQ$, d.h. M_k ist eine beliebig gewählte Funktion aus ϕ , und Q ergibt sich daraus. Dann erhält man aus (10.125)

$$NN_kV + M(Y - NQ) = 1 \quad (10.127)$$

Der Vergleich von (10.125) und (10.126) liefert damit zwangsweise, dass $N_kV = X + MQ$ sein muss. Eine andere Lösung existiert nicht. Damit lautet der Regler

$$K(s) = \frac{N_k(s)}{M_k(s)} = \frac{N_k(s)V(s)}{M_k(s)V(s)} = \frac{X(s) + M(s)Q(s)}{Y(s) - N(s)Q(s)} \quad (10.128)$$

Die Frage ist, welchen Vorteil die Verallgemeinerung des Reglers von (10.102) auf (10.120) gebracht hat. Die Antwort ist relativ einfach: Man hat nun einen freien Parameter bzw. eine freie Funktion $Q(s) \in \phi$ zur Verfügung, die die Möglichkeit bietet, dem Regler und damit dem geschlossenen Kreis außer der immer notwendigen Stabilität weitere Eigenschaften aufzuprägen. Da (10.120) die allgemeine Formel darstellt und somit alle möglichen Lösungen beinhaltet, kann man sogar Optimalität anstreben und erreichen.

Hinzu kommt, dass die Abhängigkeit aller im geschlossenen Kreis möglichen Übertragungsfunktionen von eben diesem Parameter linear ist, wie Gleichung (10.122) zeigt. Das macht die Aufgabenstellung zum Auffinden einer optimalen Lösung für eine oder sogar mehrere gegebene Anforderungen deutlich übersichtlicher, als wenn man z.B. direkt versuchen würde, für eine Übertragungsfunktion der Gestalt

$$\frac{G(s)K(s)}{G(s)K(s) + 1} \quad (10.129)$$

einen optimalen Regler $K(s)$ zu finden, da die Abhängigkeit der Übertragungsfunktion vom zu optimierenden Parameter $K(s)$ hier eindeutig komplexer ist.

Der Algorithmus zur Berechnung eines geeigneten Reglers unter Verwendung der allgemeinen Form nach Satz 10.14 lautet wie folgt:

1. Falls $G(s)$ eine stabile Funktion ist, dann setze $N := G$, $M := 1$, $X := 0$, $Y := 1$ und fahre fort mit Schritt 5.
2. Ersetze in $G(s)$ den Parameter s durch $\frac{1-\lambda}{\lambda}$. Man erhält $G(\lambda) = \frac{n(\lambda)}{m(\lambda)}$.
3. Berechne mit Hilfe des Algorithmus von Euklid aus $n(\lambda)$ und $m(\lambda)$ die gesuchten Polynome $x(\lambda)$ und $y(\lambda)$.
4. Ersetze in $n(\lambda)$, $m(\lambda)$, $x(\lambda)$ und $y(\lambda)$ den Parameter λ durch $\frac{1}{s+1}$. Es ergeben sich rationale, stabile Übertragungsfunktionen $N(s)$, $M(s)$, $X(s)$ und $Y(s) \in \phi$.
5. Lege diejenige Übertragungsfunktion des geschlossenen Kreises fest, für die bestimmte Eigenschaften erfüllt sein müssen oder optimiert werden sollen. Diese Übertragungsfunktion beinhaltet, wie jede Übertragungsfunktion, $Q(s)$ als linearen Parameter.
6. Entwerfe $Q(s) \in \phi$ so, dass die geforderten Eigenschaften bzw. Optimalität der Übertragungsfunktion aus Schritt 5 erreicht werden. Dieser Schritt wird im nachfolgenden Kapitel erläutert. $Q(s)$ kann auch 0 sein.
7. Der Regler ergibt sich aus

$$K(s) = \frac{X(s) + M(s)Q(s)}{Y(s) - N(s)Q(s)} \quad (10.130)$$

Es kann die Frage aufkommen, ob bei diesem Regler-Entwurfsverfahren bei ungünstig gewähltem $Q(s)$ das gleiche Problem auftreten kann wie beim Verfahren nach Falb-Wolovich, dass nämlich eine instabile Polstelle in der Übertragungsfunktion $G(s)$ implizit gekürzt wird. Das ist aber nicht der Fall. Da es sich hier nur um Eingrößensysteme handelt, kann eine Kürzung nur erfolgen, indem die entsprechende Nullstelle im Regler entsteht. Es kann aber leicht nachgewiesen werden, dass die Nullstellen des Reglers grundsätzlich von den Polstellen der Strecke verschieden sind, und zwar aus folgendem Grund: Zunächst ist festzustellen, dass die Polstellen der Strecke die Nullstellen von $M(s)$ und die Nullstellen der Strecke die Nullstellen von $N(s)$ sind. Weiterhin

können $M(s)$ und $N(s)$ nicht die gleichen Nullstellen haben, da sie koprim sind. Wegen $NX + MY = 1$ müssen aber auch $X(s)$ und $M(s)$ unterschiedliche Nullstellen besitzen, da die linke Seite der Gleichung für jede Nullstelle sonst den Wert 0 annehmen würde. Damit kann aber auch der Zähler des Reglers nicht die gleichen Nullstellen wie $M(s)$ aufweisen.

10.3.3 Darstellung des Optimierungsproblems

Zielsetzung beim Entwurf eines normoptimalen Reglers für ein Eingrößensystem ist es, eine der neun Übertragungsfunktionen aus Gleichung (10.122) auszuwählen und diese im Sinne einer Norm durch geeignete Wahl von $Q(s)$ so klein wie möglich zu machen. Diese Minimierung macht nicht für alle dieser Funktionen Sinn, aber für die folgenden, die anhand des Blockschaltbildes 10.6 erläutert werden:

1. Die Übertragungsfunktion vom Sollwert w zum Regelfehler h_1 . Denn dies bedeutet, dass der Regler den Regelfehler in Abhängigkeit vom Sollwert immer so gering wie möglich hält, was bedeutet, dass der Istwert dem Sollwert weitestgehend folgt.
2. Die Übertragungsfunktion von einer der beiden Störgrößen d oder n zur Größe h_3 bzw. zur Ausgangsgröße y . Das bedeutet, dass die Wirkung einer solchen Störung auf den Ausgang minimiert wird.
3. Die Übertragungsfunktion von einer der beiden Störgrößen d oder n zum Regelfehler h_1 . Das bedeutet, dass die Wirkung einer solchen Störung auf den Regelfehler minimiert wird, was gleichbedeutend damit ist, dass die Ausgangsgröße auch in Anwesenheit von Störungen dem Sollwert weitestgehend folgt.
4. Die Übertragungsfunktion von einer externen Größe w , d oder n zur Größe h_2 bzw. zur nicht eingezeichneten Stellgröße u . Das bedeutet, dass der Regler die Stellgröße immer so klein wie möglich hält, was sinnvoll sein kann, wenn die Minimierung des Energieverbrauchs im Vordergrund steht.

Im weiteren Verlauf ist festzustellen, dass es bei Fokussierung auf eine einzige Übertragungsfunktion, wie es in diesem Kapitel durchgeführt werden soll, im Entwurfsverfahren zu Extrem Lösungen kommen kann, die für die praktische Anwendung völlig ungeeignet sind. Beispielsweise kann eine Fokussierung auf einen möglichst kleinen Regelfehler dazu führen, dass ein Regler entsteht, der mit sehr großen Stellgrößen operiert, was völlig plausibel, aber für die praktische Anwendung ungeeignet ist.

Von daher muss man im Grunde fast immer mehrere Übertragungsfunktionen gleichzeitig minimieren und diese dann durch Gewichtungsfaktoren oder -funktionen gegeneinander gewichten. Damit wird aus dem Eingrößen- aber ein Mehrgrößenproblem, so dass der Entwurf eines normoptimalen Reglers eigentlich immer ein MIMO-Reglerentwurf ist. Dennoch soll in diesem

Kapitel zunächst die reine Eingrößenlösung vorgestellt werden, weil an dieser die Entwurfsidee am besten erläutert werden kann. Der Übergang zu Mehrgrößensystemen erfolgt anschließend.

Neben der reinen Auswahl einer Übertragungsfunktion zur Minimierung kann man als weiteren Entwurfparameter auch im Eingrößenfall eine Gewichtungsfunktion einführen, die hier jedoch nicht zur Gewichtung verschiedener Übertragungsfunktionen, sondern zur Gewichtung unterschiedlicher Frequenzbereiche innerhalb einer zu minimierenden Übertragungsfunktion dienen.

Abbildung 10.7 zeigt eine solche Möglichkeit. Minimiert werden sollte z.B. die Übertragungsfunktion von d nach y , wobei diese Aufgabe jetzt durch die Minimierung der Übertragungsfunktion von d nach v ersetzt wird. Die Gewichtungsfunktion $W(s)$ ist jetzt praktisch der Entwurfparameter für das Optimierungsverfahren und kann weitgehend frei gewählt werden. Sinnvoll ist beispielsweise die Wahl eines Tiefpasses für diese Gewichtungsfunktion. Das hätte zur Folge, dass die tiefen Frequenzen in der zu minimierenden Übertragungsfunktion höher gewichtet werden, was sehr sinnvoll ist, denn man will ja vor allem, dass niederfrequente Störungen gut ausgeglichen werden, während man auf die Ausregelung von hochfrequenten Störungen ab einer gewissen Frequenz besser verzichten sollte, da dies einen sehr schnell reagierenden Regler mit sehr schnell veränderlicher Stellgröße zur Folge hätte, was das Stellglied zu stark belasten würde.

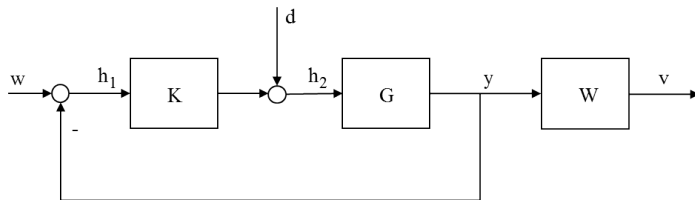


Abb. 10.7. Regelkreis mit Gewichtungsfunktion

Im Extremfall kann man sogar einen Integrator als Gewichtungsfunktion verwenden, bei dem die Frequenz $s = 0$, d.h. die Ausregelung von stationären Störungen, unendlich hoch bewertet würde, was natürlich in der Regel genau das ist, was man haben möchte. In dem Fall tritt allerdings das Problem auf, dass die Norm einer Übertragungsfunktion mit einem Pol auf der imaginären Achse minimiert werden muss. Diese ist aber nach Satz 10.7 nicht endlich, so dass die Rechnung zunächst nicht möglich ist. Man kann sich hier aber mit einem Trick behelfen, der in einem nachfolgenden Beispiel noch gezeigt wird.

Im folgenden soll aber der Einfachheit halber davon ausgegangen werden, dass $W(s)$ stabil und begrenzt und damit in ϕ enthalten ist.

Für die Berechnung ist es zunächst hilfreich, die Struktur der neun Übertragungsfunktionen aus Gleichung (10.122) zu betrachten. Es ist festzustellen, dass diese in allen Fällen durch

$$T(s) - U(s)Q(s) \quad (10.131)$$

beschrieben werden kann, wobei $T, U \in \phi$ gilt, und zwar auch, wenn noch eine Gewichtungsfunktion W hinzugefügt wird. Für die Übertragungsfunktion von d nach v aus Abbildung 10.7 ergibt sich unter Nutzung von (10.122) beispielsweise

$$\frac{v}{d} = W \frac{G}{1 + GK} = WN(Y - NQ) = WNY - WN^2Q = T - UQ \quad (10.132)$$

Nun soll $Q(s) \in \phi$ so bestimmt werden, dass die 2-Norm dieser Übertragungsfunktion minimal wird, d.h.

$$\|T - UQ\|_2 \rightarrow \min \quad (10.133)$$

Mit dem optimalen $Q(s)$ ergibt sich als kleinstmöglicher Wert der 2-Norm:

$$\gamma_{opt} = \min_Q \|T - UQ\|_2 \quad (10.134)$$

10.3.4 Berechnung des optimalen $Q(s)$

Der triviale Fall tritt ein, wenn $\frac{T}{U} \in \phi$ gilt, weil man dann $Q := \frac{T}{U}$ setzen kann und sich für die 2-Norm der bestmögliche Wert 0 ergibt. Leider ist aber bei fast allen technischen Problemen $\frac{T}{U}$ nicht Element von ϕ , weil U mindestens eine Nullstelle in der rechten Halbebene enthält. In dem Fall kann auch $\gamma_{opt} = 0$ nicht mehr erreicht werden.

Im nicht-trivialen Fall soll die Betrachtung auf streng begrenzte Übertragungsfunktionen $T - UQ$ ohne Pole auf der imaginären Achse beschränkt werden, weil ansonsten die 2-Norm wg. Satz 10.7 nicht endlich und der Versuch einer Minimierung damit sinnlos wäre. Pole können also nur in der rechten oder linken Halbebene auftreten. Folgende Mengen lassen sich definieren:

Definition 10.15 ϕ_0 sei die Menge aller streng begrenzten, stabilen Übertragungsfunktionen, d.h. deren Pole ausschließlich in der linken Halbebene liegen. ϕ_0^i sei die Menge aller streng begrenzten Übertragungsfunktionen, deren Pole ausschließlich in der rechten Halbebene liegen.

Die Vereinigungsmenge $\phi_0 \cup \phi_0^i$ ist damit die Menge aller Übertragungsfunktionen, deren 2-Norm einen endlichen Wert aufweist. Jede Funktion $H(s) \in \phi_0 \cup \phi_0^i$ kann offensichtlich mit Hilfe einer Partialbruchzerlegung in einen stabilen und einen instabilen Anteil zerlegt werden:

$$H(s) = F(s) + G(s) \quad \text{mit } F(s) \in \phi_0, G(s) \in \phi_0^i \quad (10.135)$$

Bezüglich der Addition von zwei Funktionen der Mengen ϕ_0 und ϕ_0^i gilt folgender Satz:

Satz 10.16 Für eine Funktion $F(s) \in \phi_0$ und eine Funktion $G(s) \in \phi_0^i$ gilt:

$$\|F + G\|_2^2 = \|F\|_2^2 + \|G\|_2^2 \quad (10.136)$$

Beweis: Zerlegt man F und G in Real- und Imaginärteil, so ergibt sich:

$$\begin{aligned} F(j\omega) &= f_r + jf_i \quad \text{und} \quad G(j\omega) = g_r + jg_i \\ 2\operatorname{Re}\{\overline{F}G\} &= 2\operatorname{Re}\{(f_r - jf_i)(g_r + jg_i)\} \\ &= 2\operatorname{Re}\{f_r g_r - jf_i g_r + jf_r g_i + f_i g_i\} \\ &= 2f_r g_r + 2f_i g_i \\ |F + G|^2 &= |f_r + jf_i + g_r + jg_i|^2 \\ &= (f_r + g_r)^2 + (f_i + g_i)^2 \\ &= f_r^2 + 2f_r g_r + g_r^2 + f_i^2 + 2f_i g_i + g_i^2 \\ &= f_r^2 + f_i^2 + g_r^2 + g_i^2 + 2f_r g_r + 2f_i g_i \\ &= |F|^2 + |G|^2 + 2\operatorname{Re}\{\overline{F}G\} \end{aligned} \quad (10.137)$$

Unter Verwendung der Definition der 2-Norm (10.34) gilt

$$\|F + G\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(j\omega) + G(j\omega)|^2 d\omega \quad (10.138)$$

Mit (10.137) wird daraus

$$\begin{aligned} \|F + G\|_2^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(j\omega)|^2 d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 d\omega \\ &\quad + 2\operatorname{Re}\left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{F(j\omega)} G(j\omega) d\omega \right\} \\ &= \|F\|_2^2 + \|G\|_2^2 + 2\operatorname{Re}\left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{F(j\omega)} G(j\omega) d\omega \right\} \end{aligned} \quad (10.139)$$

Wenn gezeigt werden kann, dass der letzte Term Null ist, ist der Beweis erbracht. Hierzu sollen dieselben Schritte ausgeführt werden wie bei der Herleitung einer alternativen Berechnung der 2-Norm mit Hilfe der Residuen: Da die Funktionen F und G streng begrenzt sind, muss ihr Wert im Unendlichen Null sein. Daher kann der bisherige Integrationsweg entlang der imaginären Achse gemäß Abbildung 10.4 um einen Halbkreis im Unendlichen ergänzt werden, ohne den Wert des Integrales zu verändern. Damit umschließt der

Integrationsweg nun die linke Hälfte der komplexen Ebene, und aus dem Integral wird ein Ringintegral:

$$\begin{aligned}
 \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{F(j\omega)} G(j\omega) d\omega &= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \overline{F(j\omega)} G(j\omega) dj\omega \\
 &= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \overline{F(s)} G(s) ds \\
 &= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} F(-s) G(s) ds \\
 &= \frac{1}{2\pi j} \oint_{l.H.} F(-s) G(s) ds \quad (10.140)
 \end{aligned}$$

Nach dem Residuensatz entspricht der Wert des Ringintegrals der Summe der Residuen aller Pole in dem umschlossenen Gebiet, d.h. in der linken Halbebene. Da $F(s)$ definitionsgemäß nur stabile Pole hat, liegen alle Pole von $F(-s)$ in der rechten Halbebene, und $G(s)$ hat definitionsgemäß sowieso nur Pole in der rechten Halbebene. Daher weist der Integrand des Ringintegrals $F(-s)G(s)$ keine Pole in dem umschlossenen Gebiet auf, und der Wert des Integrals muss Null sein. Somit ist der Beweis erbracht.

Mit Hilfe des gerade bewiesenen Satzes kann nun eine Formel zur Berechnung desjenigen $Q(s)$ angegeben werden, mit dem $\|T(s) - U(s)Q(s)\|_2$ minimal wird. Zunächst wird die Funktion $U(s)$ in einen Allpass und eine Minimalphasenfunktion zerlegt (vgl. Kapitel 5.6). Ein Allpass ist eine stabile Übertragungsfunktion, die zu jedem Pol eine Nullstelle mit entgegengesetztem Vorzeichen besitzt:

$$G(s) = \frac{(s + \lambda_1)(s + \lambda_2)(s + \lambda_3)\dots}{(s - \lambda_1)(s - \lambda_2)(s - \lambda_3)\dots} \quad (10.141)$$

Dabei liegen alle λ_i in der linken Halbebene. Der Frequenzgang eines Allpass ist konstant gleich Eins, wie man sich anhand eines Bode-Diagramms leicht klarmachen kann. Daher rührt der Name *Allpass*, weil ein solches Übertragungsglied alle Frequenzen gleichermaßen durchlässt.

Eine Minimalphasenfunktion ist eine stabile Übertragungsfunktion, deren Nullstellen ebenfalls ausschließlich in der linken Halbebene liegen. Warum diese Funktionen Minimalphasenfunktionen genannt werden, wird in Kapitel 5.6 erläutert, ist an dieser Stelle aber nicht relevant.

Offenbar lässt sich jede stabile rationale Übertragungsfunktion in eine Hintereinanderschaltung aus einem Allpass und einer Minimalphasenfunktion zerlegen, wie das folgende, einfache Beispiel zeigt:

$$U(s) = \frac{s-1}{(s+2)(s+3)} = \frac{s-1}{s+1} \frac{s+1}{(s+2)(s+3)} = U_{ap}(s)U_{mp}(s) \quad (10.142)$$

Die folgenden Umformungen nutzen Satz 10.16 sowie die soeben erläuterte Zerlegung:

$$\begin{aligned} \|T - UQ\|_2^2 &= \|T - U_{ap}U_{mp}Q\|_2^2 \\ &= \|U_{ap}(U_{ap}^{-1}T - U_{mp}Q)\|_2^2 \end{aligned} \quad (10.143)$$

Bei der Berechnung der 2-Norm wird ja gemäß der Definition (10.34) der Betrag der Funktion zunächst quadriert, dann integriert und zuletzt die Wurzel gezogen. Da der Frequenzgang eines Allpasses über alle Frequenzen konstant gleich Eins ist, ändert der Faktor U_{ap} nichts am Wert des Betrages im Integranden, so dass der Faktor bei der Berechnung der Norm entfallen kann:

$$\|T - UQ\|_2^2 = \|U_{ap}^{-1}T - U_{mp}Q\|_2^2 \quad (10.144)$$

Nun wird die Funktion $U_{ap}^{-1}T$ mit einer Partialbruchzerlegung in einen stabilen und einen instabilen Anteil zerlegt:

$$\|T - UQ\|_2^2 = \|[U_{ap}^{-1}T]_{instabil} + [U_{ap}^{-1}T]_{stabil} - U_{mp}Q\|_2^2 \quad (10.145)$$

Unter Ausnutzung der Tatsache, dass $U_{mp}Q$ sicherlich stabil ist, lässt sich mit Hilfe von Satz 10.16 die rechte Seite wie folgt zerlegen:

$$\|T - UQ\|_2^2 = \|[U_{ap}^{-1}T]_{instabil}\|_2^2 + \|[U_{ap}^{-1}T]_{stabil} - U_{mp}Q\|_2^2 \quad (10.146)$$

Da Q im ersten Term auf der rechten Seite nicht enthalten ist, kann man nur noch versuchen, den zweiten Term zu optimieren. Dies ist nun aber sehr einfach, weil die Invertierung von U_{mp} wieder auf eine stabile Funktion führt. Für

$$Q_{opt} = U_{mp}^{-1}[U_{ap}^{-1}T]_{stabil} \quad (10.147)$$

wird der zweite Ausdruck Null und $\|T - UQ\|_2^2$ und damit auch $\|T - UQ\|_2$ minimal.

Mit dieser Lösung muss jedoch der Lösungsweg noch einmal hinterfragt werden. Kritisch ist nämlich die strenge Begrenztheit von $[U_{ap}^{-1}T]_{instabil}$ und $([U_{ap}^{-1}T]_{stabil} - U_{mp}Q)$, da sie nämlich die Voraussetzung für die Anwendung von Satz 10.16 und damit für den Schritt von Gleichung (10.145) nach (10.146) ist.

Voraussetzung zur Minimierung der 2-Norm nach Gleichung (10.133) und (10.134) ist, dass die entsprechende Funktion und damit sowohl U als auch T streng begrenzt sind (vgl. Def. 10.15), weil die Norm sonst unendliche Werte annehmen würde und eine Minimierung damit sinnlos wäre. Dann ist aber auch U_{mp} streng begrenzt, während U_{ap} wie jeder Allpass nur einfach begrenzt ist. Wegen der strengen Begrenztheit von T sind auch $[U_{ap}^{-1}T]_{instabil}$ und $[U_{ap}^{-1}T]_{stabil}$ streng begrenzt. Um sicherzustellen, dass auch $U_{mp}Q$ streng begrenzt ist, muss daher Q zumindest begrenzt sein.

Da U_{mp} eine höhere Zähler- als Nennerordnung haben kann und in Gleichung 10.147 invertiert auftaucht, ist die Begrenztheit bei Q jedoch häufig nicht gegeben. Man kann aber stets die optimale Lösung approximieren, indem man

$$Q_{sub}(s) = \frac{Q_{opt}}{(\tau s + 1)^r} \quad (10.148)$$

setzt. Dabei ist r die Differenz zwischen Zähler- und Nennergrad von Q_{opt} . Je kleiner τ gewählt wird, desto näher liegt die Approximation an der optimalen Lösung.

10.3.5 Beispiele

Für die Strecke

$$G(s) = \frac{1-s}{s^2+s+2} \quad (10.149)$$

soll ein Regler gefunden werden, der die 2-Norm der Stör-Übertragungsfunktion von einer Streckenstörung d zum Ausgang y minimiert. Die Stör-Übertragungsfunktion soll durch die Wichtungsfunktion $W = \frac{1}{s}$ gewichtet werden, um Gleichsignale unendlich hoch zu gewichten und dadurch sicherzustellen, dass konstante Störungen komplett ausgegeregelt werden. Die zu minimierende Funktion lautet damit

$$W \frac{G}{GK+1} = WGS \quad (10.150)$$

mit der Empfindlichkeitsfunktion S .

Da es sich bei G um eine stabile Übertragungsfunktion handelt, kann eine koprime Zerlegung ohne Rechnung angegeben werden:

$$N = G, \quad M = 1, \quad X = 0, \quad Y = 1 \quad (10.151)$$

Einsetzen dieser Zerlegung in (10.120) liefert für den Regler

$$K = \frac{Q}{1-GQ} \quad (10.152)$$

Da W und damit WGS einen Pol auf der imaginären Achse aufweisen, kann $\|WGS\|_2$ nur dann einen endlichen Wert annehmen, wenn dieser Pol durch eine Nullstelle in der Funktion GS bei $s = 0$ kompensiert wird. Dies ist eine zusätzliche Nebenbedingung, welche die gesuchte Funktion Q neben der Forderung, $\|WGS\|_2$ zu minimieren, erfüllen muss.

Aufgrund dieser Nebenbedingung muss an der Stelle $s = 0$ folglich gelten:

$$\begin{aligned} 0 &= G(0)S(0) = \frac{G(0)}{1+G(0)K(0)} \\ &= \frac{G(0)}{1+G(0)\frac{Q(0)}{1-G(0)Q(0)}} \\ &= G(0)(1-G(0)Q(0)) \end{aligned} \quad (10.153)$$

Die Forderung ist erfüllt für

$$Q(0) = \frac{1}{G(0)} = 2 \quad (10.154)$$

Diese Forderung lässt sich leicht durch

$$Q(s) := 2 + sQ_1(s) \quad (10.155)$$

erfüllen. Um für $Q(s)$ Stabilität und Begrenztheit sicherzustellen, muss der Grad des Nennerpolynoms von $Q_1(s)$ den seines Zählerpolynoms um mindestens Eins übersteigen, was später noch zu berücksichtigen sein wird. Damit wird dann aber durch (10.155) die Gesamtheit aller stabilen, begrenzten Übertragungsfunktionen beschrieben, die für $s = 0$ den Wert 2 annehmen.

Einsetzen der bisherigen Ergebnisse liefert

$$\begin{aligned} WGS &= WG \frac{1}{1 + GK} \\ &= WG \frac{1}{1 + G \frac{Q}{1 - GQ}} \\ &= WG(1 - GQ) \\ &= WG(1 - G(2 + sQ_1)) \\ &= WG(1 - 2G) - WG^2 sQ_1 \\ &= WG(1 - 2G) - G^2 Q_1 = T - UQ_1 \end{aligned} \quad (10.156)$$

Mit den nun bekannten Funktionen T und U , deren strenge Begrenztheit aber noch nachzuweisen ist, kann die weitere Berechnung nach den im letzten Kapitel hergeleiteten Formeln erfolgen. Als Lösung erhält man dann jedoch nicht die endgültige Lösung Q , sondern Q_1 . Diese Funktion wird abschließend in (10.155) eingesetzt, um die endgültige Lösung Q zu bestimmen.

Durch Einsetzen von W und G erhält man für T und U

$$\begin{aligned} T(s) &= WG(1 - 2G) \\ &= \frac{1}{s} \frac{1 - s}{s^2 + s + 2} \left(1 - 2 \frac{1 - s}{s^2 + s + 2}\right) \\ &= \frac{1}{s} \frac{1 - s}{s^2 + s + 2} \frac{s^2 + 3s}{s^2 + s + 2} \\ &= \frac{(1 - s)(s + 3)}{(s^2 + s + 2)^2} \\ U(s) &= G^2 = \frac{(1 - s)^2}{(s^2 + s + 2)^2} \end{aligned} \quad (10.157)$$

Man erkennt leicht, dass sowohl U als auch T eine höhere Nenner- als Zählerordnung haben, damit also streng begrenzt sind und die nachfolgenden Schritte daher zulässig.

Die Zerlegung von U in Allpass und Minimalphasenfunktion liefert

$$U_{ap} = \frac{(1-s)^2}{(1+s)^2} \quad U_{mp} = \frac{(1+s)^2}{(s^2+s+2)^2} \quad (10.158)$$

Weiterhin ist nach Gleichung (10.147) $[U_{ap}^{-1}T]_{stabil}$ zu bestimmen:

$$U_{ap}^{-1}T = \frac{(1+s)^2}{(1-s)^2} \frac{(1-s)(s+3)}{(s^2+s+2)^2} = \frac{(1+s)^2(s+3)}{(1-s)(s^2+s+2)^2} \quad (10.159)$$

Ansetzen einer Partialbruchzerlegung für diesen Term ergibt unter Verwendung der Zuhalteregel für den ersten Term

$$\frac{(1+s)^2(s+3)}{(1-s)(s^2+s+2)^2} = \frac{1}{1-s} + \dots \quad (10.160)$$

Der Term $\frac{1}{1-s}$ auf der rechten Seite stellt den instabilen Anteil von $[U_{ap}^{-1}T]$ dar. Alle weiteren Terme müssen zum stabilen Anteil gehören. Daher lässt sich der stabile Anteil sehr einfach berechnen, indem man $\frac{1}{1-s}$ von der linken Seite der Gleichung subtrahiert:

$$\begin{aligned} [U_{ap}^{-1}T]_{stabil} &= \frac{(1+s)^2(s+3)}{(1-s)(s^2+s+2)^2} - \frac{1}{1-s} \\ &= \frac{(s^2+2s+1)(s+3) - (s^2+s+2)(s^2+s+2)}{(1-s)(s^2+s+2)^2} \\ &= \frac{(s^3+5s^2+7s+3) - (s^4+2s^3+5s^2+4s+4)}{(1-s)(s^2+s+2)^2} \\ &= \frac{s^4+s^3-3s+1}{(s-1)(s^2+s+2)^2} \\ &= \frac{(s-1)(s^3+2s^2+2s-1)}{(s-1)(s^2+s+2)^2} \\ &= \frac{s^3+2s^2+2s-1}{(s^2+s+2)^2} \end{aligned} \quad (10.161)$$

Damit lässt sich die Lösung für $Q_{1,opt}$ angeben:

$$\begin{aligned} Q_{1,opt} &= U_{mp}^{-1}[U_{ap}^{-1}T]_{stabil} \\ &= \frac{(s^2+s+2)^2}{(1+s)^2} \frac{s^3+2s^2+2s-1}{(s^2+s+2)^2} \\ &= \frac{s^3+2s^2+2s-1}{(1+s)^2} \end{aligned} \quad (10.162)$$

Diese Lösung muss noch entsprechend Gleichung (10.148) erweitert werden, damit die Ordnung des Nenners diejenige des Zählers um mindestens

Eins übersteigt, so dass sich am Ende mit (10.155) eine begrenzte Funktion für $Q(s)$ ergibt:

$$Q_{1,sub} = \frac{s^3 + 2s^2 + 2s - 1}{(1+s)^2(1+\tau s)^2} \quad (10.163)$$

Als finaler Schritt erfolgt nun das Einsetzen in Gleichung (10.155)

$$Q_{sub} = 2 + sQ_{1,sub} = 2 + \frac{s^4 + 2s^3 + 2s^2 - s}{(1+s)^2(1+\tau s)^2} \quad (10.164)$$

Der Regler ergibt sich dann aus Gleichung (10.152).

Das zweite Beispiel ist deutlich kürzer und nicht aufgrund des Rechenweges, sondern aufgrund seines Ergebnisses interessant:

Für ein PT_1 -Glied mit der Übertragungsfunktion $G = \frac{V}{Ts+1}$ soll derjenige Regler gefunden werden, der die 2-Norm der Übertragungsfunktion vom Sollwert w zum Regelfehler e minimiert.

Da die Übertragungsfunktion der Strecke stabil ist, kann die kopprime Faktorisierung auch hier sofort angegeben werden:

$$N = G, \quad M = 1, \quad X = 0, \quad Y = 1 \quad (10.165)$$

Einsetzen dieser Zerlegung in (10.120) liefert für den Regler wieder

$$K = \frac{Q}{1 - GQ} \quad (10.166)$$

Die zu minimierende Funktion ist gerade die Empfindlichkeitsfunktion:

$$\frac{e}{w} = S = \frac{1}{1 + GK} = \frac{1}{1 + G \frac{Q}{1 - GQ}} = 1 - GQ \quad (10.167)$$

Der Vergleich mit Gleichung (10.131) liefert

$$T = 1 \quad \text{und} \quad U = G = \frac{V}{Ts + 1} \quad (10.168)$$

Da U keine Nullstellen in der rechten Halbebene besitzt, kann das optimale Q sofort angegeben werden:

$$Q_{opt}(s) = \frac{T}{U} = \frac{1}{G} = \frac{Ts + 1}{V} \quad (10.169)$$

Festzustellen ist, dass in dieser Berechnung T nicht streng begrenzt war. Dies ist hier jedoch kein Problem, da Satz 10.16 hier gar nicht zur Anwendung kam. Dennoch sollte auch hier die Begrenztheit von Q angestrebt werden, um überhaupt einen realisierbaren Regler zu erhalten. Daher ist der Nenner von $Q_{opt}(s)$ auch hier wieder entsprechend Gleichung (10.148) zu erweitern:

$$Q_{sub}(s) = \frac{Q_{opt}(s)}{\tau s + 1} = \frac{1}{V} \frac{Ts + 1}{\tau s + 1} \quad (10.170)$$

Der resultierende Regler ist mit (10.166)

$$\begin{aligned}
 K(s) &= \frac{Q_{sub}(s)}{1 - G(s)Q_{sub}(s)} = \frac{\frac{1}{V} \frac{T_s+1}{\tau s+1}}{1 - \frac{V}{T_s+1} \frac{1}{V} \frac{T_s+1}{\tau s+1}} \\
 &= \frac{T_s+1}{V(\tau s+1) - V} = \frac{1}{V} \frac{T_s+1}{\tau s+1-1} = \frac{1}{V} \frac{T_s+1}{\tau s} \\
 &= V' \frac{T_s+1}{T_s} \text{ mit } V' = \frac{T}{V\tau}
 \end{aligned} \tag{10.171}$$

Dies ist gerade ein PI-Regler.

10.4 Berechnung normoptimaler Regler im Zustandsraum

Wie schon gesagt, kann bei Eingrößensystemen die Fokussierung auf nur eine einzige Übertragungsfunktion und die Minimierung von deren Norm zu Reglern führen, die zwar im Sinne der Norm optimal, für die Praxis jedoch nicht zu gebrauchen sind, da entweder Regler mit unendlich großen oder verschwindenden Stellgrößen entstehen können. Sinnvoller ist es, mehrere Übertragungsfunktionen gleichzeitig im Blick zu behalten und diese zu minimieren. Damit wird aus dem Eingrößen- aber ein Mehrgrößenproblem. Von daher war die Betrachtung im vorhergehenden Kapitel für das Verständnis der grundsätzlichen Idee einer normoptimalen Regelung zwar hilfreich, für die Praxis aber nicht ausreichend.

Daher sollen in diesem Kapitel die Mehrgrößensysteme behandelt werden, die sowohl die Systeme mit mehreren Ein- und Ausgangsgrößen als auch den Fall eines Eingrößensystems mit mehreren zu optimierenden Übertragungsfunktionen beinhalten.

10.4.1 Formulierung der Entwurfsaufgabe

Die Entwurfsaufgabe kann anhand von Abbildung 10.8 erläutert werden: Dabei bezeichnet \mathbf{z} den Vektor von Eingangsgrößen in das System, also Sollwerte oder Störsignale. \mathbf{v} bezeichnet den Vektor von Größen, die möglichst kleine Werte annehmen sollen, also Regelfehler, Stellgrößen oder Ausgangsgrößen, ggf. modifiziert durch Wichtungsfunktionen. \mathbf{u} ist der Stellgrößenvektor als Ausgang des Reglers \mathbf{K} , und \mathbf{y} der Ausgangsvektor der Strecke, der auf den Regler zurückgeführt wird.

Gesucht ist der Regler \mathbf{K} , der die Auswirkungen des Eingangsvektors \mathbf{z} auf den Ausgangsvektor \mathbf{v} bezüglich einer gewählten Norm minimiert.

Offenbar enthält \mathbf{P} neben dem Modell der Strecke auch die ggf. vorhandenen Gewichtungsfunktionen, um die Anforderungen an die Regelung näher

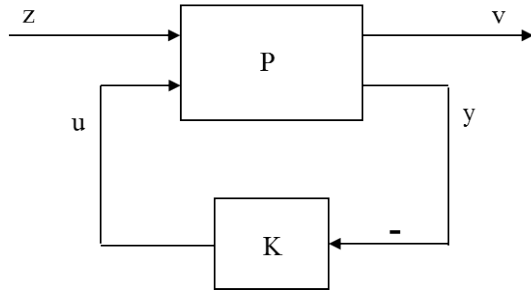


Abb. 10.8. Zur Formulierung der Entwurfsaufgabe bei normoptimalen Regelungen

zu spezifizieren. Zusammenfassend enthält \mathbf{P} also alles, was gegeben ist, und \mathbf{K} das, was berechnet werden soll.

Als Beispiel für die Strukturierung der Entwurfsaufgabe soll der Regelkreis in Abbildung 10.9 dienen. Unter der Voraussetzung, dass der Sollwert gleich Null ist, sollen in Anwesenheit einer Störung am Ausgang der Strecke sowohl die Ausgangsgröße als auch die Stellgröße so klein wie möglich sein, wobei beide Größen durch Gewichtungsfunktionen W_i gewichtet werden. Dabei wird W_1 als Tiefpass gewählt, um vor allem Abweichungen der Ausgangsgröße vom Sollwert Null bei niedrigen Frequenzen hoch zu gewichten, während W_2 ein Hochpass ist, um hochfrequente Aktivitäten der Stellgröße hoch zu gewichten und damit nach Möglichkeit zu verhindern.

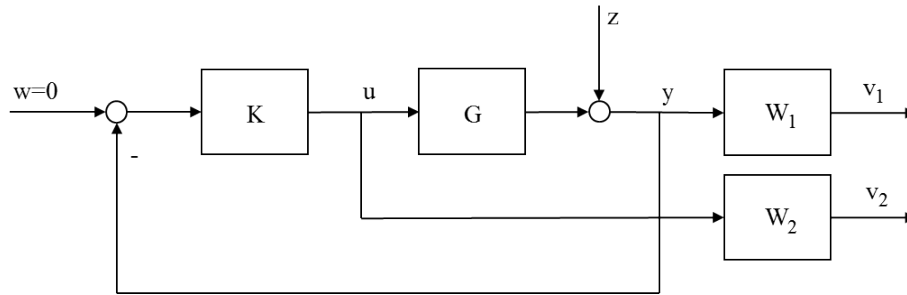


Abb. 10.9. Eingrößensystem mit zwei Gewichtungsfunktionen

Umstrukturieren dieses Blockschaltbildes gemäß der Struktur in Abbildung 10.8 führt auf Abbildung 10.10. Man erkennt jetzt deutlich den aus einer Größe bestehenden Eingangsvektor z sowie die beiden Ausgangsgrößen v_1 und v_2 , die gemeinsam den zu minimierenden Ausgangsvektor bilden. Zur Übertragungsmatrix \mathbf{P} gehören die Übertragungsfunktion der Strecke sowie die beiden Gewichtungsfunktionen.

Man kann \mathbf{P} in vier Blöcke \mathbf{P}_{11} , \mathbf{P}_{12} , \mathbf{P}_{21} und \mathbf{P}_{22} zerlegen, die die Übertragungsfunktionen von z und u nach v und y beschreiben:

$$\mathbf{v} = \mathbf{P}_{11}\mathbf{z} + \mathbf{P}_{12}\mathbf{u}$$

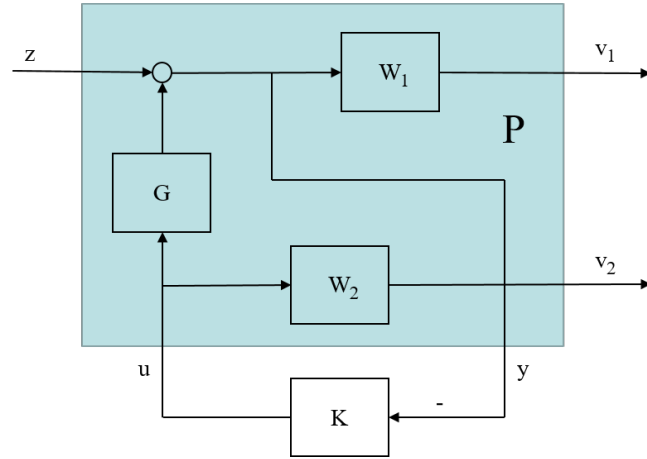


Abb. 10.10. Eingrößensystem mit zwei Gewichtungsfunktionen, strukturiert gemäß der Entwurfsaufgabe

$$\mathbf{y} = \mathbf{P}_{21}\mathbf{z} + \mathbf{P}_{22}\mathbf{u} \quad (10.172)$$

Mit $\mathbf{u} = -\mathbf{K}\mathbf{y}$ wird daraus

$$\begin{aligned} \mathbf{v} &= \mathbf{P}_{11}\mathbf{z} - \mathbf{P}_{12}\mathbf{K}\mathbf{y} \\ \mathbf{y} &= \mathbf{P}_{21}\mathbf{z} - \mathbf{P}_{22}\mathbf{K}\mathbf{y} \end{aligned} \quad (10.173)$$

Umformen der zweiten Gleichung liefert

$$\begin{aligned} (\mathbf{I} + \mathbf{P}_{22}\mathbf{K})\mathbf{y} &= \mathbf{P}_{21}\mathbf{z} \\ \mathbf{y} &= (\mathbf{I} + \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\mathbf{z} \end{aligned} \quad (10.174)$$

Und durch Einsetzen dieser Gleichung in (10.173) erhält man

$$\mathbf{v} = \mathbf{P}_{11}\mathbf{z} - \mathbf{P}_{12}\mathbf{K}(\mathbf{I} + \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\mathbf{z} \quad (10.175)$$

Die resultierende Übertragungsfunktion von \mathbf{z} nach \mathbf{v} lautet demnach:

$$\mathbf{T}_{zv} = \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{K}(\mathbf{I} + \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21} \quad (10.176)$$

Ziel des Reglerentwurfes ist es daher, diese Übertragungsfunktion im Sinne einer gewählten Norm zu minimieren.

Die Übertragungsmatrix \mathbf{P} wird üblicherweise in der Zustandsdarstellung

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B} \begin{bmatrix} \mathbf{z} \\ \mathbf{u} \end{bmatrix} \\ \begin{bmatrix} \mathbf{v} \\ \mathbf{y} \end{bmatrix} &= \mathbf{C}\mathbf{x} + \mathbf{D} \begin{bmatrix} \mathbf{z} \\ \mathbf{u} \end{bmatrix} \end{aligned} \quad (10.177)$$

angegeben, wobei die Matrizen \mathbf{B} , \mathbf{C} und \mathbf{D} in Einzelteile zerlegt werden:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + [\mathbf{B}_1, \mathbf{B}_2] \begin{bmatrix} \mathbf{z} \\ \mathbf{u} \end{bmatrix} \\ \begin{bmatrix} \mathbf{v} \\ \mathbf{y} \end{bmatrix} &= \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{u} \end{bmatrix}\end{aligned}\quad (10.178)$$

Daraus folgt für den Zusammenhang zwischen der Zustandsdarstellung und den vier Teil-Übertragungsfunktionen \mathbf{P}_{ii} :

$$\begin{aligned}\mathbf{P}_{11}(s) &= \mathbf{C}_1(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}_1 + \mathbf{D}_{11} \\ \mathbf{P}_{12}(s) &= \mathbf{C}_1(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}_2 + \mathbf{D}_{12} \\ \mathbf{P}_{21}(s) &= \mathbf{C}_2(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}_1 + \mathbf{D}_{21} \\ \mathbf{P}_{22}(s) &= \mathbf{C}_2(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}_2 + \mathbf{D}_{22}\end{aligned}\quad (10.179)$$

oder in verkürzter Schreibweise, wie sie im folgenden verwendet werden soll:

$$\begin{aligned}\mathbf{P}_{11} &= [\mathbf{A}, \mathbf{B}_1, \mathbf{C}_1, \mathbf{D}_{11}] \\ \mathbf{P}_{12} &= [\mathbf{A}, \mathbf{B}_2, \mathbf{C}_1, \mathbf{D}_{12}] \\ \mathbf{P}_{21} &= [\mathbf{A}, \mathbf{B}_1, \mathbf{C}_2, \mathbf{D}_{21}] \\ \mathbf{P}_{22} &= [\mathbf{A}, \mathbf{B}_2, \mathbf{C}_2, \mathbf{D}_{22}]\end{aligned}\quad (10.180)$$

Für \mathbf{P} ergibt sich in dieser Schreibweise (vgl. (10.178)):

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} = \left[\mathbf{A}, [\mathbf{B}_1, \mathbf{B}_2], \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \right] \quad (10.181)$$

Mit Hilfe der Matrizen in dieser Gleichung lassen sich die 2-Norm- und ∞ -Norm-optimalen Regler sowie notwendige und hinreichende Bedingungen für die Existenz von Lösungen angeben.

In Fortsetzung des Beispiels aus den Abbildungen 10.9 bzw. 10.10 soll nun die Berechnung von \mathbf{P} bzw. seiner Teilmatrizen verdeutlicht werden.

Die Strecke sei ein PT_1 -Glieder mit

$$G(s) = \frac{1}{s+1} \quad (10.182)$$

das gemäß Abbildung 2.10 als rückgekoppelter Integrator mit der Verstärkung $V = 1$ und der Zeitkonstanten $T = 1$ beschrieben werden kann. In der Zustandsdarstellung ergibt sich daher

$$G(s) = \frac{1}{s+1} = [\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}] = [-1, 1, 1, 0] \quad (10.183)$$

Ebenso wird die Wichtungsfunktion W_1 als PT_1 -Glieder angenommen:

$$W_1(s) = \frac{20}{10s+1} = [\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}] = [-0.1, 2, 1, 0] \quad (10.184)$$

Die zweite Wichtungsfunktion W_2 soll ein Hochpass sein, der hier in sehr einfacher Form als Differenz zwischen einem konstanten Anteil und einem PT_1 -Glied beschrieben wird:

$$W_2(s) = 6 - \frac{4}{4s + 1} = [\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}] = [-0.25, 1, -1, 6] \quad (10.185)$$

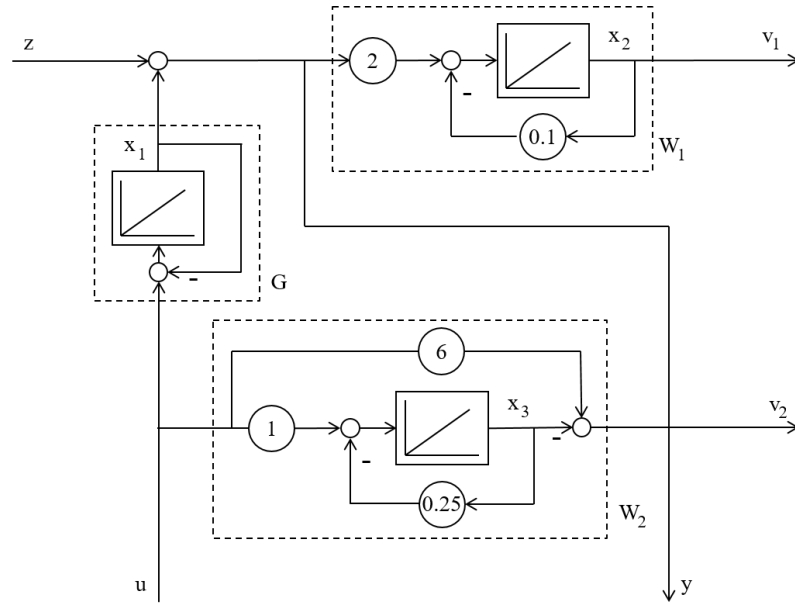


Abb. 10.11. Eingrößensystem mit zwei Gewichtungsfunktionen, strukturiert gemäß der Entwurfsaufgabe

Die Zustandsdarstellung von \mathbf{P} lässt sich aus dem Blockschaltbild 10.11 ablesen. Definiert man die Zustandsgrößen wie im Blockschaltbild eingetragen, so ergibt sich

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} = \left[\begin{bmatrix} -1 & 0 & 0 \\ 2 & -0.1 & 0 \\ 0 & 0 & -0.25 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 6 \\ 1 & 0 \end{bmatrix} \right] \quad (10.186)$$

und die Teilmatrizen lauten

$$\begin{aligned} \mathbf{B}_1 &= \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} & \mathbf{B}_2 &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{C}_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} & \mathbf{C}_2 &= [1 \ 0 \ 0] \\ \mathbf{D}_{11} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \mathbf{D}_{12} &= \begin{bmatrix} 0 \\ 6 \end{bmatrix} \end{aligned}$$

$$\mathbf{D}_{21} = [1] \quad \mathbf{D}_{22} = [0] \quad (10.187)$$

10.4.2 2-Norm-optimaler Regler

Da die Herleitung der Gleichungen zur Berechnung des Reglers im Zustandsraum zwar elegant, aber sehr aufwendig ist, sollen an dieser Stelle nur die Ergebnisse vorgestellt werden. Der 2-Norm-optimale Regler lautet:

$$\mathbf{K} = [\mathbf{A} + \mathbf{B}_2\mathbf{F} + \mathbf{H}\mathbf{C}_2, \mathbf{H}, \mathbf{F}, 0] \quad (10.188)$$

Dabei sind \mathbf{A} , \mathbf{B}_2 und \mathbf{C}_2 die entsprechenden Systemmatrizen aus der Matrix \mathbf{P} . Die untere Hälfte von Abbildung 10.12 verdeutlicht die Struktur dieses Reglers. Es handelt sich um die klassische Struktur aus Beobachter und Zustandsrückführung (vgl. Abbildung 8.7), wobei hier lediglich zu berücksichtigen ist, dass $(-\mathbf{y})$ das Eingangssignal des Reglers und \mathbf{u} sein Ausgangssignal ist. Daher bilden \mathbf{H} hier die Eingangsmatrix und \mathbf{F} die Ausgangsmatrix des Systems.

Die zu regelnde Strecke aus Sicht des Reglers hat, wie ein Blick auf den oberen Teil des Blockschaltbildes mit der Systemmatrix \mathbf{P} zeigt (dargestellt hier ohne Durchgangsmatrix \mathbf{D}), tatsächlich die Eingangsmatrix \mathbf{B}_2 und die Ausgangsmatrix \mathbf{C}_2 , so dass das Beobachtermodell im Regler die im geschlossenen Kreis zu regelnde Strecke korrekt widerspiegelt. Die Größen \mathbf{z} und \mathbf{v} mit den Blöcken \mathbf{B}_1 und \mathbf{C}_1 liegen außerhalb des geschlossenen Kreises.

Die Matrizen \mathbf{F} und \mathbf{H} berechnen sich aus den Lösungen \mathbf{X} und \mathbf{Y} zweier algebraischer Matrix-Riccati-Gleichungen (vgl. Satz 8.8):

$$\begin{aligned} \mathbf{F} &= -\mathbf{B}_2^T \mathbf{X} \\ \mathbf{X} : \mathbf{A}^T \mathbf{X} + \mathbf{X} \mathbf{A} - \mathbf{X} \mathbf{B}_2 \mathbf{B}_2^T \mathbf{X} &= -\mathbf{C}_1^T \mathbf{C}_1 \end{aligned} \quad (10.189)$$

und

$$\begin{aligned} \mathbf{H} &= -\mathbf{Y} \mathbf{C}_2^T \\ \mathbf{Y} : \mathbf{A} \mathbf{Y} + \mathbf{Y} \mathbf{A}^T - \mathbf{Y} \mathbf{C}_2^T \mathbf{C}_2 \mathbf{Y} &= -\mathbf{B}_1 \mathbf{B}_1^T \end{aligned} \quad (10.190)$$

Der Vergleich von Gleichung (10.189) mit dem Riccati-Regler (8.153)

$$\mathbf{F} = -\mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} \quad (10.191)$$

wobei \mathbf{P} die symmetrische, positiv definite Lösung der Matrix-Riccati-Gleichung (8.154)

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} - \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} = -\mathbf{Q} \quad (10.192)$$

ist, zeigt sofort, dass es sich beim 2-Norm-optimalen Regler um einen Riccati-Regler handelt, wobei $\mathbf{R} = \mathbf{I}$ ist und die positiv definite Gewichtungsmatrix \mathbf{Q} aus dem Riccati-Entwurf hier ersetzt wird durch das Produkt $\mathbf{C}_1^T \mathbf{C}_1$.

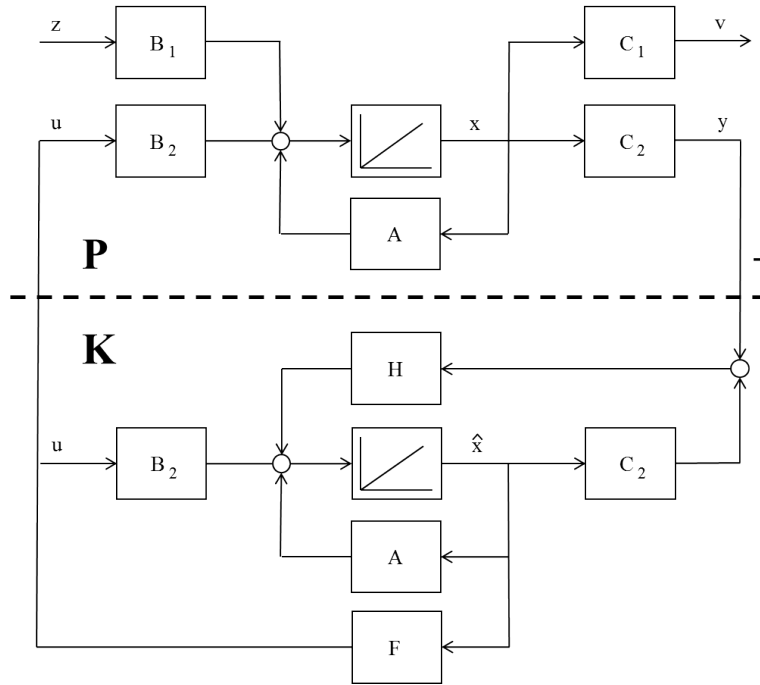


Abb. 10.12. P-Matrix mit 2-Norm-optimalem Regler

Dies macht aber Sinn aus Sicht des 2-Norm-optimalen Reglerentwurfs. Im Riccati-Entwurf bildet \mathbf{Q} im zu minimierenden Gütefunktional (8.114) den Gewichtungsfaktor für die Zustandsgrößen. Minimiert wird im Riccati-Entwurf (unter anderem) der Term $\mathbf{x}^T \mathbf{Q} \mathbf{x}$. Es gilt aber, wie man anhand eines Blickes auf den oberen Teil der Abbildung 10.12 leicht feststellen kann:

$$\mathbf{x}^T \mathbf{Q} \mathbf{x} = \mathbf{x}^T \mathbf{C}_1^T \mathbf{C}_1 \mathbf{x} = \mathbf{v}^T \mathbf{v} \quad (10.193)$$

so dass im vorliegenden Fall mit der so gewählten Matrix $\mathbf{C}_1^T \mathbf{C}_1$ gerade $\mathbf{v}^T \mathbf{v}$ minimiert wird, was exakt der im vorangegangenen Kapitel vorgestellten Entwurfsaufgabe entspricht.

Das Gleiche lässt sich auch für den Beobachterentwurf nach Gleichung (10.190) sagen. Die Formeln (8.167) und (8.168) für einen Riccati-Beobachter bzw. ein Kalman-Filter erhält man aus (10.192), indem man entsprechend Gleichung (8.44) wie folgt ersetzt:

$$\mathbf{A} \rightarrow \mathbf{A}^T \quad \mathbf{B} \rightarrow \mathbf{C}^T \quad \mathbf{F} \rightarrow \mathbf{H}^T \quad (10.194)$$

Damit ergibt sich als Matrix-Riccati-Gleichung für den Beobachter

$$\mathbf{A} \mathbf{P} + \mathbf{P} \mathbf{A}^T - \mathbf{P} \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \mathbf{P} = -\mathbf{Q} \quad (10.195)$$

und als Bestimmungsgleichung für \mathbf{H} :

$$\mathbf{H} = -\mathbf{P}\mathbf{C}^T\mathbf{R}^{-1} \quad (10.196)$$

Der Zusammenhang mit Gleichung (10.190) ist auch hier sofort zu erkennen. Wieder ist $\mathbf{R} = \mathbf{I}$, und \mathbf{Q} ist hier $\mathbf{B}_1\mathbf{B}_1^T$. Auch dies macht Sinn, denn gemäß den Ausführungen in Kapitel 8.6.4 beschreibt \mathbf{Q} beim Entwurf eines Kalman-Filters das Systemrauschen, also diejenige unbekannte Störung, die unmittelbar vor dem Integrationsblock angreift. An eben dieser Stelle greift aber beim 2-Norm-optimalen Reglerentwurf der Eingangsvektor \mathbf{z} über die Eingangsmatrix \mathbf{B}_1 an. Und während beim Entwurf des Kalman-Filters das Ziel ist, den Einfluss eben dieses Systemrauschens zu minimieren, ist es beim 2-Norm-optimalen Reglerentwurf der Einfluss des Eingangsvektors \mathbf{z} .

Es stellt sich die Frage, warum man überhaupt den 2-Norm-optimalen Reglerentwurf benötigt, wenn er doch nur ein Spezialfall des Riccati-Reglers ist. Die Antwort dazu lautet wie folgt: Beim Riccati-Regler hat man für den Entwurf nur zwei konstante Gewichtungsmatrizen \mathbf{Q} und \mathbf{R} als Entwurfsparameter zur Verfügung, deren Bedeutung zwar sehr klar ist, deren genaue Parameter aber schwer festzulegen sind. Im Grunde kann man nur definieren, welche Größen wichtiger sind als andere und daher durch den entsprechenden Koeffizienten in der Matrix höher zu gewichten sind. Die Größenverhältnisse der Koeffizienten untereinander kann man demnach gut festlegen, die absoluten Zahlenwerte jedoch nicht.

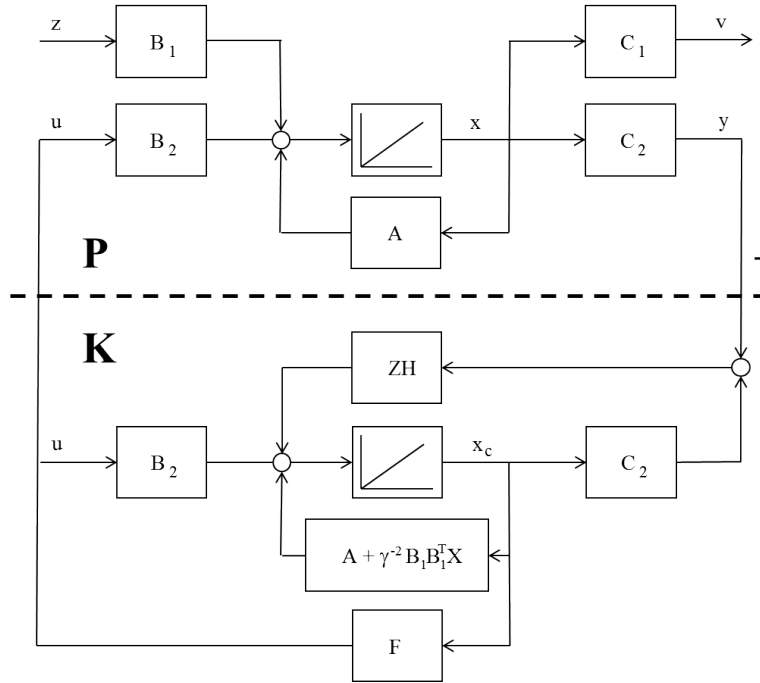
Dagegen kann man beim normoptimalen Reglerentwurf für jede zu berücksichtigende Ausgangsgröße eine (fast) beliebige, frequenzabhängige Wichtungsfunktion $\mathbf{W}_i(s)$ definieren, die diese Ausgangsgröße nicht nur mit einem Verstärkungsfaktor, sondern bei Bedarf auch noch frequenzabhängig gewichtet. Das bedeutet nicht nur mehr Flexibilität, sondern auch eine deutlich verbesserte und leichter zu verstehende Möglichkeit der Spezifikation der Reglereigenschaften.

10.4.3 ∞ -Norm-optimaler Regler

Auch für diesen Regler sollen nur die Ergebnisse vorgestellt werden. Der ∞ -Norm-optimale Regler lautet:

$$\mathbf{K} = [\mathbf{A} + \gamma^{-2}\mathbf{B}_1\mathbf{B}_1^T\mathbf{X} + \mathbf{B}_2\mathbf{F} + \mathbf{Z}\mathbf{H}\mathbf{C}_2, \mathbf{Z}\mathbf{H}, \mathbf{F}, \mathbf{0}] \quad (10.197)$$

Dabei sind auch hier, wie beim 2-Norm-optimalen Regler, \mathbf{A} , \mathbf{B}_2 und \mathbf{C}_2 die entsprechenden Systemmatrizen aus der Matrix \mathbf{P} . Die untere Hälfte von Abbildung 10.13 verdeutlicht die Struktur dieses Reglers. Offensichtlich ist dieser Regler gegenüber dem 2-Norm-optimalen Regler um die Matrix \mathbf{Z} in der Beobachter-Rückführung sowie um den Term $\gamma^{-2}\mathbf{B}_1\mathbf{B}_1^T\mathbf{X}$ in der Systemmatrix verändert. Da dieser Regler nicht mehr wie der 2-Norm-optimale Regler die klassische Struktur aus Beobachter und Regler-Rückführung aufweist, ist der reglerinterne Zustandsvektor hier nicht mehr mit $\hat{\mathbf{x}}$, sondern mit \mathbf{x}_c bezeichnet.

Abb. 10.13. P-Matrix mit ∞ -Norm-optimalem Regler

Zur Berechnung der Reglermatrizen müssen auch hier zunächst zwei algebraische Matrix-Riccati-Gleichungen (vgl. Satz 8.8) gelöst werden:

$$\begin{aligned} \mathbf{A}^T \mathbf{X} + \mathbf{X} \mathbf{A} - \mathbf{X} (\mathbf{B}_2 \mathbf{B}_2^T - \gamma^{-2} \mathbf{B}_1 \mathbf{B}_1^T) \mathbf{X} &= -\mathbf{C}_1^T \mathbf{C}_1 \\ \mathbf{A} \mathbf{Y} + \mathbf{Y} \mathbf{A}^T - \mathbf{Y} (\mathbf{C}_2^T \mathbf{C}_2 - \gamma^{-2} \mathbf{C}_1^T \mathbf{C}_1) \mathbf{Y} &= -\mathbf{B}_1 \mathbf{B}_1^T \end{aligned} \quad (10.198)$$

Man erhält daraus die Lösungen \mathbf{X} und \mathbf{Y} . Mit diesen ergeben sich die Regler-Matrizen wie folgt:

$$\begin{aligned} \mathbf{F} &= -\mathbf{B}_2^T \mathbf{X} \\ \mathbf{H} &= -\mathbf{Y} \mathbf{C}_2^T \\ \mathbf{Z} &= (\mathbf{I} - \gamma^{-2} \mathbf{Y} \mathbf{X})^{-1} \end{aligned} \quad (10.199)$$

Lässt man γ gegen unendlich konvergieren, dann geht der ∞ -Norm-optimale Regler offensichtlich in den 2-Norm-optimalen Regler über. Er kann damit als Erweiterung des 2-Norm-optimalen Reglers aufgefasst werden. Umgekehrt gewinnen der Term $\gamma^{-2} \mathbf{B}_1 \mathbf{B}_1^T \mathbf{X}$ und die Matrix \mathbf{Z} mit kleiner werdendem γ immer mehr an Einfluss.

Dabei ist γ der Wert der ∞ -Norm der Übertragungsfunktion \mathbf{T}_{zv} aus Gleichung (10.176), der zunächst unbekannt ist. Deshalb kann der ∞ -Norm-optimale Regler wie auch die ∞ -Norm (vgl. Kapitel 10.2.5) nur iterativ be-

rechnet werden. Die Berechnung erfolgt anhand der Prüfung der Lösbarkeitsbedingungen für die beiden Riccati-Gleichungen (10.198) und anhand der Prüfung der Invertierbarkeit von $(\mathbf{I} - \gamma^{-2}\mathbf{Y}\mathbf{X})$ aus Gleichung (10.199).

Zu Beginn wird ein ausreichend großer Startwert für γ gewählt. Mit diesem Wert können die den Riccati-Gleichungen für \mathbf{X} und \mathbf{Y} zugeordneten Hamilton-Matrizen

$$\begin{aligned}\mathbf{H}_X &= \begin{bmatrix} \mathbf{A} & \mathbf{B}_2\mathbf{B}_2^T - \gamma^{-2}\mathbf{B}_1\mathbf{B}_1^T \\ \mathbf{C}_1^T\mathbf{C}_1 & -\mathbf{A}^T \end{bmatrix} \\ \mathbf{H}_Y &= \begin{bmatrix} \mathbf{A}^T & \mathbf{C}_2^T\mathbf{C}_2 - \gamma^{-2}\mathbf{C}_1^T\mathbf{C}_1 \\ \mathbf{B}_1\mathbf{B}_1^T & -\mathbf{A} \end{bmatrix}\end{aligned}\quad (10.200)$$

aufgestellt werden. Die Riccati-Gleichungen sind lösbar, wenn die zugeordneten Hamilton-Matrizen keine Eigenwerte auf der imaginären Achse aufweisen, was numerisch leicht überprüft werden kann. In dem Fall werden die Riccati-Gleichungen berechnet, und man erhält Matrizen \mathbf{X} und \mathbf{Y} . Auch dies ist numerisch unproblematisch.

Nun ist die Invertierbarkeit der Matrix $\mathbf{U} := (\mathbf{I} - \gamma^{-2}\mathbf{Y}\mathbf{X})$ zu prüfen, die für die Berechnung der Matrix \mathbf{Z} gemäß Gleichung (10.199) benötigt wird. Für die nachfolgenden Überlegungen wird aber zunächst vom gegenteiligen Fall ausgegangen: Falls die Matrix \mathbf{U} nicht invertierbar ist, muss ihre Determinante $\det(\mathbf{U}) = 0$ sein. Da die Determinante dem Produkt der Eigenwerte von \mathbf{U} entspricht ($\det(\mathbf{U}) = \prod \lambda_i$), folgt daraus wiederum, dass mindestens ein Eigenwert λ_i von \mathbf{U} gleich Null sein muss.

Zu diesem Eigenwert existiert ein Eigenvektor \mathbf{x}_e , der die Eigenwertgleichung $\lambda_i\mathbf{x}_e = \mathbf{U}\mathbf{x}_e$ erfüllt. Damit gilt:

$$\begin{aligned}0 &= 0\mathbf{x}_e = \lambda_i\mathbf{x}_e = \mathbf{U}\mathbf{x}_e \\ 0 &= (\mathbf{I} - \gamma^{-2}\mathbf{Y}\mathbf{X})\mathbf{x}_e = \mathbf{x}_e - \gamma^{-2}\mathbf{Y}\mathbf{X}\mathbf{x}_e \\ \mathbf{x}_e &= \gamma^{-2}\mathbf{Y}\mathbf{X}\mathbf{x}_e \\ \gamma^2\mathbf{x}_e &= \mathbf{Y}\mathbf{X}\mathbf{x}_e\end{aligned}\quad (10.201)$$

Aus der Struktur der letzten Gleichung folgt sofort, dass \mathbf{x}_e auch ein Eigenvektor von $\mathbf{Y}\mathbf{X}$ sein muss und γ^2 der zugehörige Eigenwert.

Daher werden, nachdem \mathbf{X} und \mathbf{Y} aus den Riccati-Gleichungen (10.198) berechnet worden sind, im nächsten Schritt das Matrizenprodukt $\mathbf{Y}\mathbf{X}$ und seine Eigenwerte berechnet. Wenn der größte dieser Eigenwerte kleiner ist als γ^2 , dann gibt es für Gleichung (10.201) keine Lösung \mathbf{x}_e , und aufgrund der vorangegangenen Argumentationskette muss \mathbf{U} invertierbar sein.

In dem Fall wird γ verkleinert, und die ganze Berechnung wird erneut durchgeführt. Diese Iteration wird so lange fortgesetzt, bis \mathbf{U} nicht mehr invertierbar ist. Der kleinste Wert von γ , für den \mathbf{U} noch invertierbar war, ist dann der gesuchte Wert der ∞ -Norm der Übertragungsfunktion \mathbf{T}_{zv} aus Gleichung (10.176), und mit diesem können die Reglermatrizen anhand der Gleichung (10.199) berechnet werden.

Normoptimale Regler stehen auch in einem engen Zusammenhang mit dem Begriff der Robustheit. Daher werden im nachfolgenden Kapitel zunächst die Grundlagen zu diesem Thema erläutert, bevor dann am Ende des Kapitels der Zusammenhang mit den normoptimalen Reglern hergestellt wird.

10.5 Modellunsicherheiten und Robustheit

10.5.1 Grundlagen

Alle klassischen, analytischen Regler-Entwurfsverfahren basieren auf einem Streckenmodell, von dem man annimmt, dass es das Verhalten der realen Strecke ausreichend genau widerspiegelt, so dass der entworfene Regler auch im Zusammenspiel mit der realen Strecke Stabilität bei ausreichender Regelgeschwindigkeit garantiert.

Weiterhin gilt fast immer, dass eine Erhöhung der Regelgeschwindigkeit durch Anheben der Reglerverstärkung den stabilen, geschlossenen Kreis näher an die Stabilitätsgrenze bringt, d.h. an den Punkt, von dem ab das System instabil wird. Dies ergibt sich unmittelbar aus dem Small Gain Theorem, das gleich noch vorgestellt wird, aber z.B. auch aus dem Nyquist-Kriterium (Satz 5.6). Denn dort wird die Ortskurve des offenen Kreises durch ein Anheben der Reglerverstärkung gestreckt und damit meistens näher an den kritischen Punkt -1 gebracht.

Wenn man nun davon ausgeht, dass das nominale Modell, welches man zum Reglerentwurf verwendet, die reale Strecke nicht exakt widerspiegelt, muss man auch davon ausgehen, dass der geschlossene Kreis aus Regler und realer Strecke möglicherweise näher an der Stabilitätsgrenze liegt als der aus Regler und Streckenmodell bestehende Kreis. Um nicht Gefahr zu laufen, dass der reale geschlossene Kreis aufgrund der Abweichungen zwischen Modell und realer Strecke jenseits der Stabilitätsgrenze liegt und instabil wird, ist man daher gezwungen, beim auf einem Modell basierenden Reglerentwurf aus Sicherheitsgründen einen gewissen Abstand zur Stabilitätsgrenze zu wahren. Nach dem oben Gesagten bedeutet dies aber normalerweise einen Verzicht auf Regelgeschwindigkeit.

Ziel sollte nun zunächst sein, die *Modellunsicherheiten*, also die potentiellen Abweichungen zwischen Modell und realer Strecke, einzugrenzen, d.h. in geeigneter Form, z.B. durch einen Maximalwert für ihre Norm, zu beschreiben. Dies ist durchaus möglich, auch ohne die Unsicherheiten im Detail zu kennen. Dann kann man in einem zweiten Schritt mit einem geeigneten Regler-Entwurfsverfahren einen Regler so entwerfen, dass dieser das System auch dann stabilisieren kann, wenn tatsächlich Abweichungen zwischen Strecke und Modell in diesem Umfang auftreten. Einen solchen Regler bezeichnet man als *robust*.

Besonders interessant wird ein solcher Reglerentwurf, wenn er die gegebenen Stabilitätsgrenzen maximal ausnutzt und die maximal mögliche Regelge-

schwindigkeit liefert bei gleichzeitiger Robustheit gegenüber den potentiellen Abweichungen zwischen Modell und realer Strecke. Solch ein Reglerentwurf ist mit normoptimalen Reglern möglich.

Nun sollen aber zunächst einige der Möglichkeiten aufgezeigt werden, Modellunsicherheiten mathematisch zu beschreiben. Der erste Fall ist in Abbildung 10.14 gezeigt und wird als additive Modellunsicherheit bezeichnet. Die Übertragungsfunktion der realen Strecke $\mathbf{G}(s)$ ist die Summe aus dem nominalen Streckenmodell \mathbf{G}_0 und einer unbekannten Übertragungsfunktion Δ_A , die - wenn überhaupt - nur durch ihre Norm abgeschätzt werden kann:

$$\mathbf{G}(s) = \mathbf{G}_0(s) + \Delta_A(s) \quad (10.202)$$

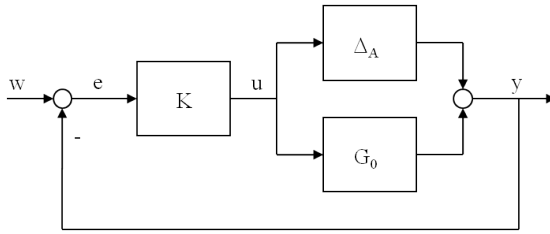


Abb. 10.14. Additive Modellunsicherheit

Auch die multiplikative Formulierung der Modellunsicherheit ist denkbar (Abbildung 10.15). Hier setzt sich die Übertragungsfunktion der realen Strecke aus einem Produkt aus dem nominalen Streckenmodell \mathbf{G}_0 und einem Term zusammen, der die multiplikative Modellunsicherheit Δ_M enthält:

$$\mathbf{G}(s) = (\mathbf{I} + \Delta_M(s))\mathbf{G}_0(s) \quad (10.203)$$

Offensichtlich gilt der Zusammenhang

$$\Delta_A(s) = \Delta_M(s)\mathbf{G}_0(s) \quad (10.204)$$

Je nach Definition können die beiden Produktterme in ihrer Reihenfolge auch vertauscht sein:

$$\mathbf{G}(s) = \mathbf{G}_0(s)(\mathbf{I} + \Delta_M(s)) \quad (10.205)$$

Daraus folgt der alternative Zusammenhang:

$$\Delta_A(s) = \mathbf{G}_0(s)\Delta_M(s) \quad (10.206)$$

Umstrukturieren des Blockschaltbildes 10.14 liefert das Blockschaltbild 10.16.

Man erkennt einen inneren geschlossenen Kreis, bestehend aus Regler und nominalem Streckenmodell, und einen äußeren Kreis, der die additive Modellunsicherheit enthält. Der innere Kreis kann als sicher stabil angenommen werden, weil der Regler ja für das nominale Modell entworfen wird. Es gilt:

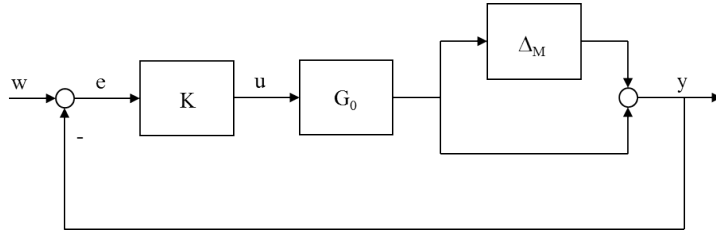


Abb. 10.15. Multiplikative Modellunsicherheit

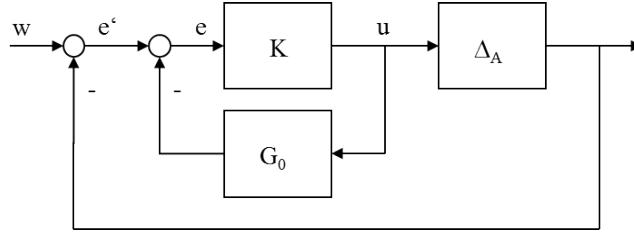


Abb. 10.16. Umstrukturierter Regelkreis mit additiver Modellunsicherheit

$$\begin{aligned}
 \mathbf{e}(s) &= \mathbf{e}'(s) - \mathbf{G}_0(s)\mathbf{K}(s)\mathbf{e}(s) \\
 (\mathbf{I} + \mathbf{G}_0(s)\mathbf{K}(s))\mathbf{e}(s) &= \mathbf{e}'(s) \\
 \mathbf{e}(s) &= (\mathbf{I} + \mathbf{G}_0(s)\mathbf{K}(s))^{-1}\mathbf{e}'(s) \\
 \mathbf{u}(s) &= \mathbf{K}(s)\mathbf{e}(s) \\
 &= \mathbf{K}(s)(\mathbf{I} + \mathbf{G}_0(s)\mathbf{K}(s))^{-1}\mathbf{e}'(s) \\
 &= \mathbf{K}(s)\mathbf{S}_0(s)\mathbf{e}'(s)
 \end{aligned} \tag{10.207}$$

Dabei ist \mathbf{S}_0 die Empfindlichkeitsfunktion des nominalen Regelkreises. Diesen Zusammenhang spiegelt Abbildung 10.17 wider.

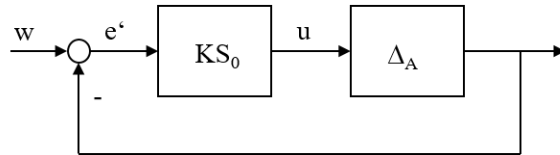


Abb. 10.17. Umstrukturierter Regelkreis mit additiver Modellunsicherheit

Weiterhin gilt der folgende Satz:

Satz 10.17 *Small Gain Theorem: Wenn die Norm des offenen Regelkreises (d.h. der Kreisübertragungsfunktion) kleiner als Eins ist, dann ist der geschlossene Kreis stabil im Sinne der Stabilitätsdefinition von Lyapunov 7.2.*

Der Beweis ist trivial: Die Norm des offenen Regelkreises ist sein Verstärkungsfaktor für die korrespondierende Norm eines Signales, das ihn durchläuft. Wenn die Norm des offenen Regelkreises kleiner als Eins ist, dann bedeutet dies, dass die korrespondierende Norm eines ihn durchlaufenden Signales

beim Durchlauf verkleinert wird. Der offene Regelkreis stellt das Übertragungsverhalten aller Übertragungsglieder des geschlossenen Kreises dar. Ein Signal, das den offenen Regelkreis durchläuft, würde in einem geschlossenen Kreis genau an seinem Ausgangspunkt wieder ankommen, und zwar laut dem eben Gesagten kleiner als vorher, wobei sich *kleiner* auf die jeweilige, korrespondierende Signalnorm bezieht. Geht man davon aus, dass sämtliche Eingangssignale des geschlossenen Kreises gleich Null sind, wie dies in der Stabilitätsdefinition von Lyapunov vorausgesetzt wird, wird sich das betrachtete Signal mit jedem Umlauf durch den geschlossenen Kreis immer weiter verkleinern und letztendlich gegen Null konvergieren. Da dies für alle Signale im geschlossenen Kreis gleichermaßen gilt, müssen alle Zustandsgrößen im geschlossenen Kreis gegen Null konvergieren. Damit ist der geschlossene Kreis stabil im Sinne von Lyapunov.

So einfach und elegant, wie das Small Gain Theorem ist, hat es jedoch einen gravierenden Nachteil. Es ist außerordentlich konservativ, wie seine Anwendung auf ein lineares Eingrößensystem zeigt. Verwendet man beispielsweise die ∞ -Norm, so bedeutet dies mit Definition (10.19), dass der Betrag der Kreisübertragungsfunktion immer kleiner als Eins sein muss, d.h. die gesamte Ortskurve der Kreisübertragungsfunktion muss innerhalb des Einheitskreises verlaufen. Wie schon das einfache Beispiel eines IT_1 -Gliedes in Kapitel 5.5 zeigt, gibt es jedoch viele Fälle, in denen die Ortskurve der Kreisübertragungsfunktion außerhalb des Einheitskreises verläuft und das System trotzdem stabil ist. Das Small Gain Theorem stellt daher eine oft viel zu hohe Anforderung dar, die zur Einhaltung der Stabilität gar nicht erforderlich wäre.

Die Anwendung des Small Gain Theorem auf Abbildung 10.17 liefert als Bedingung für Stabilität:

$$\|\Delta_A \mathbf{K} \mathbf{S}_0(s)\| < 1 \quad (10.208)$$

Wegen

$$\|\Delta_A \mathbf{K} \mathbf{S}_0(s)\| \leq \|\Delta_A(s)\| \|\mathbf{K} \mathbf{S}_0(s)\| \quad (10.209)$$

ist Gleichung (10.208) sicher erfüllt, wenn

$$\|\Delta_A(s)\| \|\mathbf{K} \mathbf{S}_0(s)\| < 1 \quad (10.210)$$

gilt bzw.

$$\|\Delta_A(s)\| < \frac{1}{\|\mathbf{K} \mathbf{S}_0(s)\|} \quad (10.211)$$

Damit hat man eine obere Schranke für die maximal zulässige additive Modellunsicherheit, die in Kauf genommen werden kann, ohne die Stabilität des geschlossenen Kreises zu gefährden, dessen Regler auf Basis des nominalen Streckenmodells entwickelt wurde. Diese Ungleichung ist äußerst plausibel: Eine Vergrößerung der Übertragungsfunktion bzw. der Verstärkung des Reglers $\mathbf{K}(s)$ führt ebenso zu einer Reduktion dieser oberen Schranke wie eine

Vergrößerung der Empfindlichkeit des nominalen geschlossenen Kreises $\mathbf{S}_0(s)$, in der der Regler im Übrigen auch noch einmal mit enthalten ist.

Mit Gleichung (10.204) lässt sich Entsprechendes auch für die multiplikative Modellunsicherheit ermitteln. Zunächst gilt

$$\|\Delta_M \mathbf{G}_0 \mathbf{K} \mathbf{S}_0(s)\| < 1 \quad (10.212)$$

und nach äquivalenten Überlegungen unter Berücksichtigung von Gleichung (10.9) erhält man hier

$$\|\Delta_M(s)\| < \frac{1}{\|\mathbf{G}_0 \mathbf{K} \mathbf{S}_0(s)\|} = \frac{1}{\|\mathbf{T}_0(s)\|} \quad (10.213)$$

wobei $\mathbf{T}_0(s)$ die Übertragungsfunktion des geschlossenen Kreises mit dem nominalen Modell bezeichnet.

Wenn man nicht den Regelfehler, sondern die Stellgröße als Ausgangspunkt der Berechnungen nimmt, erhält man zum Blockschaltbild 10.16 nach ähnlichen Überlegungen die folgenden Gleichungen:

$$\begin{aligned} \mathbf{u}(s) &= -\mathbf{K}(s) \mathbf{G}_0(s) \mathbf{u}(s) + \mathbf{K}(s) \mathbf{e}'(s) \\ (\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s)) \mathbf{u}(s) &= \mathbf{K}(s) \mathbf{e}'(s) \\ \mathbf{u}(s) &= (\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s) \mathbf{e}'(s) \\ \mathbf{e}'(s) &= -\Delta_A(s) \mathbf{u}(s) \\ \mathbf{u}(s) &= -(\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s) \Delta_A(s) \mathbf{u}(s) \end{aligned} \quad (10.214)$$

Für die Einhaltung des Small Gain Theorem muss hiermit

$$\|(\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s) \Delta_A(s)\| < 1 \quad (10.215)$$

gelten oder alternativ unter Verwendung einer durch (10.206) gegebenen multiplikativen Modellunsicherheit

$$\|(\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s) \mathbf{G}_0(s) \Delta_M(s)\| < 1 \quad (10.216)$$

Da eine Kreisübertragungsfunktion an jedem beliebigen Punkt des geschlossenen Kreises beginnen kann, lassen sich die einzelnen Blöcke in den Gleichungen des Small Gain Theorem auch zyklisch vertauschen. Statt der beiden Gleichungen (10.215) und (10.216) kann man damit z.B. auch schreiben:

$$\begin{aligned} \|\Delta_A(s) (\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s)\| &< 1 \\ \|\Delta_M(s) (\mathbf{I} + \mathbf{K}(s) \mathbf{G}_0(s))^{-1} \mathbf{K}(s) \mathbf{G}_0(s)\| &< 1 \end{aligned} \quad (10.217)$$

10.5.2 Zusammenhang mit normoptimalen Regelungen

Als Beispielsystem für die nachfolgenden Überlegungen soll das System in Abbildung 10.18 dienen. Der Sollwert w sei Null, es greifen zwei Störgrößen jeweils vor und hinter der Strecke an. Stell- und Ausgangsgröße sollen minimiert werden, beide gewichtet über die Gewichtungsfunktionen W_i .

Die Übertragungsfunktion von z_1 nach v_1 ergibt sich aus

$$\begin{aligned} u &= -KG(u + z_1) \\ (I + KG)u &= -KGz_1 \\ u &= (I + KG)^{-1}KGz_1 \\ v_1 &= W_1u = W_1(I + KG)^{-1}KGz_1 \end{aligned} \quad (10.218)$$

Die Übertragungsfunktion von z_2 nach v_1 ergibt sich aus

$$\begin{aligned} y &= z_2 - GK y \\ (I + GK)y &= z_2 \\ y &= (I + GK)^{-1}z_2 = Sz_2 \\ v_1 &= -W_1Ky = -W_1KSz_2 \end{aligned} \quad (10.219)$$

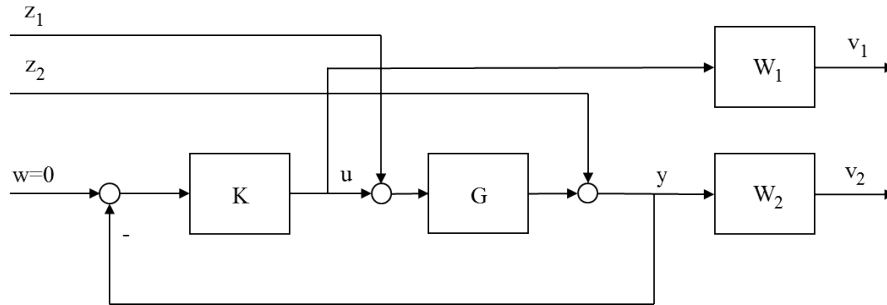


Abb. 10.18. Beispielsystem mit zwei Ein- und Ausgängen

Die Übertragungsfunktion von z_1 nach v_2 ergibt sich aus

$$\begin{aligned} y &= G(z_1 - Ky) = Gz_1 - GK y \\ (I + GK)y &= Gz_1 \\ y &= (I + GK)^{-1}Gz_1 = SGz_1 \\ v_2 &= W_2y = W_2SGz_1 \end{aligned} \quad (10.220)$$

Und die Übertragungsfunktion von z_2 nach v_2 ist offensichtlich W_2S . Damit lautet die zu minimierende Übertragungsmatrix

$$\mathbf{T}_{vz} = \begin{bmatrix} W_1(I + KG)^{-1}KG & -W_1KS \\ W_2SG & W_2S \end{bmatrix} \quad (10.221)$$

Als Lösung ergibt sich ein normoptimaler Regler, für den $\|\mathbf{T}_{vz}\|$ einen bestimmten, minimalen Wert γ annimmt. Aus $\|\mathbf{T}_{vz}\| < \gamma$ folgt aber auch für alle Teil-Übertragungsfunktionen

$$\begin{aligned} \|W_1(I + KG)^{-1}KG\| &< \gamma \\ \|W_1KS\| &< \gamma \\ \|W_2SG\| &< \gamma \\ \|W_2S\| &< \gamma \end{aligned} \quad (10.222)$$

Vergleicht man die erste dieser vier Ungleichungen mit Gleichung (10.217) und die zweite mit Gleichung (10.208), so stellt man fest, dass W_1 im ersten Fall die multiplikative und im zweiten Fall die additive Modellunsicherheit ersetzt. Aufgrund dieser Äquivalenz ergeben sich die nachfolgenden zwei möglichen Vorgehensweisen:

Für den Fall, dass eine additive Modellunsicherheit gegeben ist, kann man für das hier gegebene System $W_1 = \Delta_A$ setzen und prüfen, ob sich am Ende für die Norm ein Wert $\gamma < 1$ ergibt. Damit wäre dann nicht nur ein Regler gefunden, der die Norm $\|\mathbf{T}_{vz}\|$ minimiert, sondern gleichzeitig auch die Robustheit wegen der Äquivalenz der zweiten Gleichung aus (10.220) mit Gleichung (10.208) gewährleistet. Entsprechendes gilt für eine multiplikative Modellunsicherheit wegen Gleichung (10.217).

Alternativ kann man W_1 wie gewohnt im Hinblick auf gewünschte Regelziele festlegen und den normoptimalen Reglerentwurf durchführen. Für die Norm $\|\mathbf{T}_{vz}\|$ wird sich ein Wert $\gamma \neq 1$ ergeben. Dann ist $\frac{1}{\gamma}W_1$ die zulässige additive oder multiplikative Modellunsicherheit, für die mit dem berechneten Regler Robustheit garantiert werden kann.

Problematisch ist allerdings die Frage, wie man denn nun, in Abhängigkeit von einzelnen unbekannten Modellparametern und deren Toleranzbereichen, die additive oder multiplikative Modellunsicherheit als separate Funktion $\Delta_A(s)$ bzw. $\Delta_M(s)$ beschreiben kann. Dies ist in der Tat der Schwachpunkt dieses Ansatzes, da die Abhängigkeit dieser beiden Funktionen von einzelnen Modellparametern sehr komplex ist. Realistisch gesehen bleibt nur die Möglichkeit einer groben Schätzung der Unsicherheit in Form einer frequenzabhängigen Funktion, indem man also z.B. formuliert, dass die multiplikative Modellunsicherheit bezogen auf das Gesamtmodell 10% beträgt und für höhere Frequenzen auf 100% anwächst. Dann ergibt sich

$$\Delta_M(s) = 0.1 \frac{0.1s + 1}{0.01s + 1} \quad (10.223)$$

Ein wesentlich präziserer Umgang mit Unsicherheiten einzelner Parameter lässt sich mit dem strukturierten singulären Wert μ erreichen, der im folgenden Kapitel vorgestellt wird.

10.5.3 Der strukturierte singuläre Wert

Es sei Δ eine quadratische Matrix und Element der Menge

$$\mathcal{P} := \{\text{diag}[\delta_1 \mathbf{I}_{r_1}, \dots, \delta_j \mathbf{I}_{r_j}, \Delta_1, \dots, \Delta_k]\} \quad (10.224)$$

mit $\delta_i \in \mathbf{C}$ und $\Delta_i \in \mathbf{C}^{m_i \times m_i}$, wobei \mathbf{C} hier die Menge der komplexen Zahlen bezeichnet.

Eine solche Matrix definiert die Parameterunsicherheiten und wird als Perturbation bezeichnet. Sie sieht beispielsweise mit den Werten $r_1 = 2$, $r_2 = 1$ und $m_1 = 2$ wie folgt aus:

$$\Delta = \begin{bmatrix} \delta_1 & 0 & 0 & 0 & 0 \\ 0 & \delta_1 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 & 0 & 0 \\ 0 & 0 & 0 & \Delta_{11} & \Delta_{12} \\ 0 & 0 & 0 & \Delta_{21} & \Delta_{22} \end{bmatrix} \quad (10.225)$$

Für die Dimension der quadratischen Matrix Δ gilt offensichtlich

$$\dim(\Delta) = \sum_{i=1}^j r_i + \sum_{i=1}^k m_i \quad (10.226)$$

Mit einer quadratischen Matrix \mathbf{M} , die die gleiche Dimension wie Δ aufweist, kann eine Funktion $\mu_\Delta(\mathbf{M})$ wie folgt definiert werden:

$$\mu_\Delta(\mathbf{M}) := \frac{1}{\min_{\Delta} \{\bar{\sigma}(\Delta) : \Delta \in \mathcal{P}, \det(\mathbf{I} - \mathbf{M}\Delta) = 0\}} \quad (10.227)$$

μ_Δ ist somit der Kehrwert der ∞ -Norm der (im Sinne der ∞ -Norm) kleinsten Perturbation Δ , die $(\mathbf{I} - \mathbf{M}\Delta)$ singulär macht. Er ist der sogenannte *strukturierte singuläre Wert* der Matrix \mathbf{M} und ergibt sich aus ihrer Struktur und ihren Parametern, d.h. es handelt sich um eine \mathbf{M} zugeordnete Größe, die mittels geeigneter numerischer Methoden berechnet werden kann.

Doch was bedeutet die Definition?

Ausgangspunkt der Überlegungen ist der geschlossene Kreis in Abbildung 10.19, wobei davon auszugehen ist, dass $\mathbf{M}(s)$ eine frequenzabhängige Übertragungsmatrix ist. Aufgrund der positiven Rückkopplung lautet die Empfindlichkeitsfunktion bzw. -matrix des geschlossenen Kreises $\mathbf{S} = (\mathbf{I} - \mathbf{M}\Delta)^{-1}$. Deren Polstellen wiederum sind die Nullstellen von $\det(\mathbf{I} - \mathbf{M}\Delta)$. Wenn mindestens eine davon auf der imaginären Achse oder sogar in der rechten Halbebene liegt, dann ist das System aus Abbildung 10.19 nicht mehr stabil.

Basierend auf diesen Überlegungen lässt sich nun die Definition in (10.227) erklären: Für $\Delta = \mathbf{0}$ ist $\det(\mathbf{I} - \mathbf{M}\Delta) = 1$ und damit sicher ungleich Null. Dann wird Δ vergrößert (d.h. seine ∞ -Norm wächst), und zwar so lange, bis $\det(\mathbf{I} - \mathbf{M}\Delta) = 0$ wird. Damit ist die kleinste Perturbation gefunden, für die

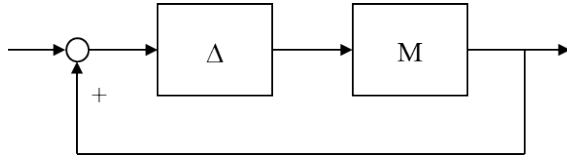


Abb. 10.19. Zur Erläuterung von $\det(\mathbf{I} - \mathbf{M}\Delta) = 0$

das System in Abbildung 10.19 instabil wird. Für kleinere Perturbationen ist das System gerade eben noch stabil. Der Kehrwert der ∞ -Norm dieser Perturbation ist dann μ_Δ .

Ein großer Wert für μ_Δ bedeutet demnach, dass schon eine kleine Perturbation (=Unsicherheit) ausreicht, um das System in Abbildung 10.19 instabil zu machen. Beschränkt man die Menge der Perturbationen Δ auf die Menge

$$\mathcal{P}_b := \{\Delta \in \mathcal{P} : \bar{\sigma}(\Delta) \leq 1\} \quad (10.228)$$

dann folgt daraus offensichtlich:

Satz 10.18 *Ein System gemäß Abbildung 10.19 ist für alle Perturbationen $\Delta \in \mathcal{P}_b$ genau dann stabil, wenn*

$$\mu_\Delta(\mathbf{M}) < 1 \quad (10.229)$$

gilt.

Wie lassen sich diese Erkenntnisse auf ein regelungstechnisches Problem anwenden? Ausgangspunkt der nun folgenden Überlegungen ist das Blockschaltbild 10.20. Das System \mathbf{M} besteht, wie ein Vergleich mit Abbildung 10.8 zeigt, aus der Matrix \mathbf{P} mit Strecke und Wichtungsfunktionen sowie dem Regler \mathbf{K} . Die Matrix Δ (Perturbation) enthält die Parameterunsicherheiten, die, um Satz 10.18 anwenden zu können, so formuliert werden müssen, dass $\bar{\sigma}(\Delta) \leq 1$ gilt, d.h. ihre ∞ -Norm kleiner als Eins ist.

Dies lässt sich durch eine geschickte Formulierung relativ leicht erreichen. Man muss nur einen Teil der Verstärkungsfaktoren der Perturbationen innerhalb von \mathbf{M} abbilden. Dies wird nachfolgend in einem Beispiel gezeigt.

Diese Teile von \mathbf{M} haben jedoch keinen Einfluss auf das Zusammenspiel aus \mathbf{P} und \mathbf{K} , so dass im ersten Schritt ein klassischer normoptimaler Reglerentwurf entsprechend Kapitel 10.4.2 oder 10.4.3 ohne Berücksichtigung der Perturbationen erfolgen kann. Sobald der Regler berechnet ist, steht damit auch \mathbf{M} fest, wobei \mathbf{M} als Übertragungsverhalten von \mathbf{z}_Δ nach \mathbf{v}_Δ beschreiben muss, um den geschlossenen Kreis entsprechend Abb. 10.19 zu bilden.

Dann wird $\mu_\Delta(\mathbf{M})$ aus \mathbf{M} numerisch berechnet. Falls dieser Wert kleiner als Eins ist, ist die entworfene Regelung für die gegebenen Perturbationen stabil.

Als Beispiel für den Umgang mit Perturbationen soll der Feder-Masse-Schwinger aus Abbildung 2.1 dienen. Die Newtonsche Bewegungsgleichung lautet

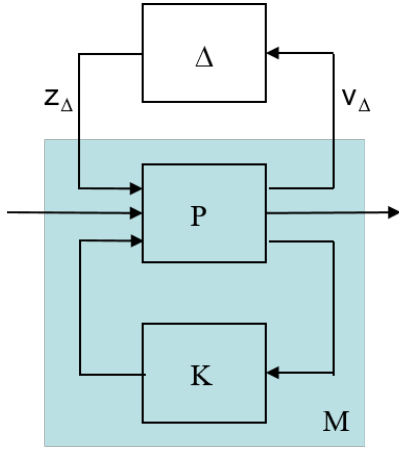


Abb. 10.20. Zur Anwendung des strukturierten singulären Wertes

$$f_a - f_f - f_r = a m \quad (10.230)$$

Für die Ableitungen von Geschwindigkeit und Weg gilt

$$a = \frac{dv}{dt} \quad \text{und} \quad v = \frac{dl}{dt} \quad (10.231)$$

Und für die Feder-Rückstellkraft und die Reibungskraft seien

$$f_f = c_f l \quad \text{und} \quad f_r = c_r v \quad (10.232)$$

angenommen. Mit den Zustandsgrößen v und l lauten die Zustandsgleichungen:

$$\begin{aligned} \begin{bmatrix} \dot{v} \\ \dot{l} \end{bmatrix} &= \begin{bmatrix} -\frac{c_r}{m} & -\frac{c_f}{m} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ l \end{bmatrix} + \begin{bmatrix} \frac{1}{m} \\ 0 \end{bmatrix} f_a \\ l &= [0, 1] \begin{bmatrix} v \\ l \end{bmatrix} + 0 f_a \end{aligned} \quad (10.233)$$

Nun seien die Parameter des Systems mit Unsicherheiten behaftet:

$$\begin{aligned} 0.07 &\leq c_r \leq 0.13 \\ 0.7 &\leq c_f \leq 1.3 \\ 0.07 &\leq m \leq 0.13 \end{aligned} \quad (10.234)$$

Um eine Perturbationsmatrix Δ mit $\bar{\sigma}(\Delta) \leq 1$ zu erhalten, werden diese wie folgt formuliert:

$$\begin{aligned} c_r &\in \{0.1 + 0.03\delta_1 : |\delta_1| \leq 1\} \\ c_f &\in \{1 + 0.3\delta_2 : |\delta_2| \leq 1\} \\ m &\in \{0.1 + 0.03\delta_3 : |\delta_3| \leq 1\} \end{aligned} \quad (10.235)$$

Da die Masse m im Nenner der Zustandsmatrizen auftritt, ist es günstiger, die Größen

$$\tilde{c}_r := \frac{c_r}{m} \quad \tilde{c}_f := \frac{c_f}{m} \quad \tilde{m} := \frac{1}{m} \quad (10.236)$$

zu verwenden. Damit erhält man die neuen Zustandsgleichungen

$$\begin{aligned} \begin{bmatrix} \dot{v} \\ \dot{l} \end{bmatrix} &= \begin{bmatrix} -\tilde{c}_r & -\tilde{c}_f \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ l \end{bmatrix} + \begin{bmatrix} \tilde{m} \\ 0 \end{bmatrix} f_a \\ l &= [0, 1] \begin{bmatrix} v \\ l \end{bmatrix} + 0 f_a \end{aligned} \quad (10.237)$$

Die Intervalle für die neuen Koeffizienten ergeben sich, wenn man jeweils die maximalen und minimalen Werte von Zähler und Nenner miteinander kombiniert:

$$\begin{aligned} \tilde{c}_{r,max} &= \frac{c_{r,max}}{m_{min}} = \frac{0.13}{0.07} = 1.857 \\ \tilde{c}_{r,min} &= \frac{c_{r,min}}{m_{max}} = \frac{0.07}{0.13} = 0.539 \\ \tilde{c}_{f,max} &= \frac{c_{f,max}}{m_{min}} = \frac{1.3}{0.07} = 18.57 \\ \tilde{c}_{f,min} &= \frac{c_{f,min}}{m_{max}} = \frac{0.7}{0.13} = 5.39 \\ \tilde{m}_{max} &= \frac{1}{m_{min}} = \frac{1}{0.07} = 14.286 \\ \tilde{m}_{min} &= \frac{1}{m_{max}} = \frac{1}{0.13} = 7.692 \end{aligned} \quad (10.238)$$

Um auch für diese Parameter sicherzustellen, dass $\bar{\sigma}(\Delta) \leq 1$ gilt, muss

$$\begin{aligned} \tilde{c}_r &\in \{1.198 + 0.659\delta_1 : |\delta_1| \leq 1\} \\ \tilde{c}_f &\in \{11.978 + 6.593\delta_2 : |\delta_2| \leq 1\} \\ \tilde{m} &\in \{10.989 + 3.297\delta_3 : |\delta_3| \leq 1\} \end{aligned} \quad (10.239)$$

gewählt werden. Die Matrix Δ lautet damit

$$\Delta = \begin{bmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & \delta_3 \end{bmatrix} \quad (10.240)$$

Die Einbindung der Perturbation in den geschlossenen Regelkreis zeigt Abbildung 10.21. Ihr Eingangsvektor ist $\mathbf{v}_\Delta = [v, l, f_a]^T$. Dies ergibt sich sofort aus den Matrizen \mathbf{C}_3 und \mathbf{D}_3 . Die Matrix \mathbf{B}_3 beschreibt die Wirkung der Perturbation auf die Ableitung des Zustandsvektors, die sich aus den Gleichungen (10.237) und (10.239) ergibt:

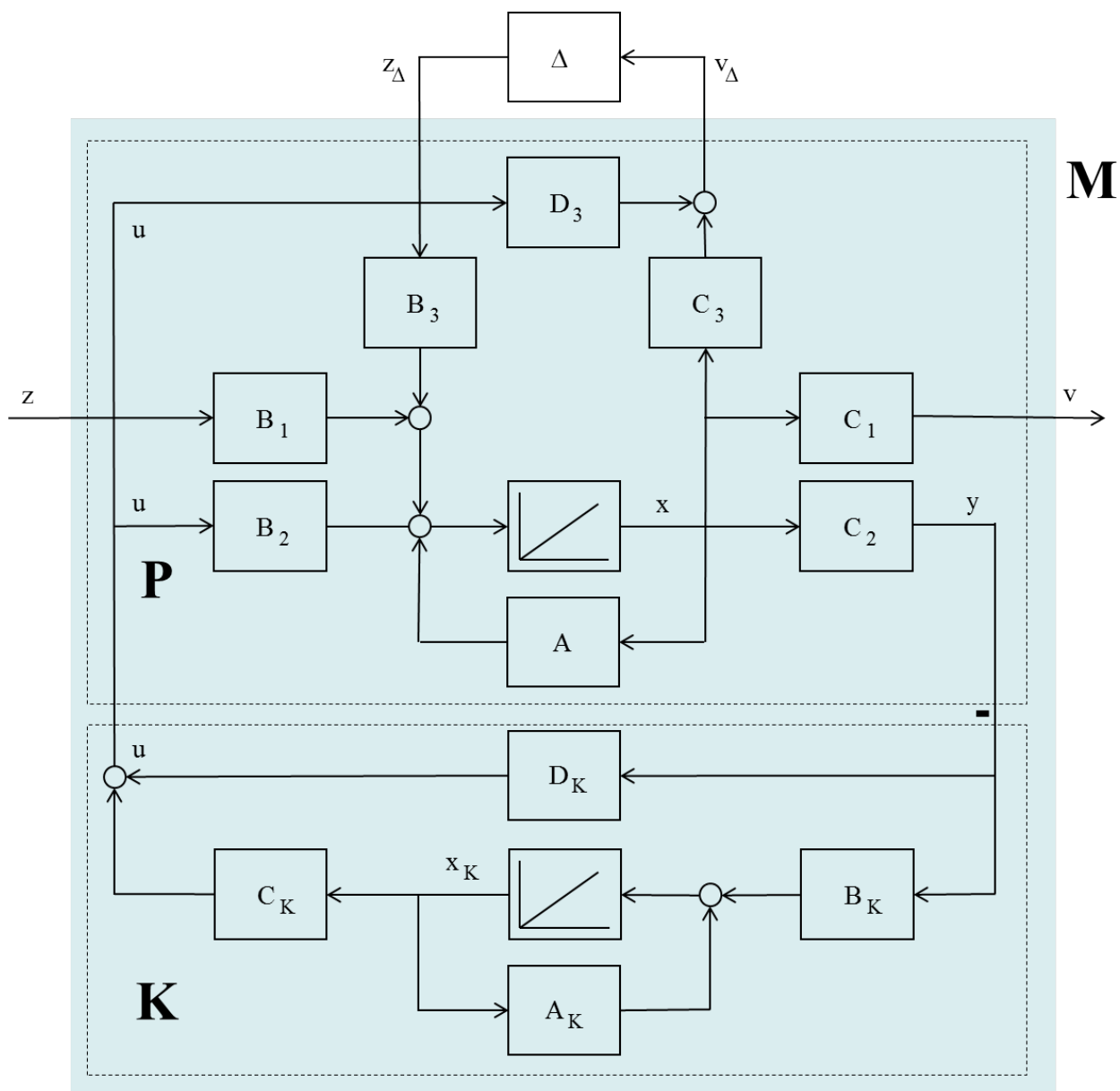


Abb. 10.21. Zur Einbindung von Perturbationen in den geschlossenen Regelkreis

$$\begin{aligned}
\mathbf{B}_3 &= \begin{bmatrix} -0.659 & -6.593 & 3.297 \\ 0 & 0 & 0 \end{bmatrix} \\
\mathbf{C}_3 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\
\mathbf{D}_3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned} \tag{10.241}$$

Die Matrizen \mathbf{B}_2 , \mathbf{C}_2 , und \mathbf{A} ergeben sich aus den Zustandsgleichungen wie folgt:

$$\begin{aligned}
\mathbf{B}_2 &= \begin{bmatrix} 10.989 \\ 0 \end{bmatrix} \\
\mathbf{C}_2 &= [0 \ 1] \\
\mathbf{A} &= \begin{bmatrix} -1.198 & -11.978 \\ 1 & 0 \end{bmatrix}
\end{aligned} \tag{10.242}$$

Für die Matrix \mathbf{M} ergibt sich, wenn man den Ausgang der Perturbation als Eingangsgröße und die Eingangsgröße der Perturbation als Ausgangsgröße von \mathbf{M} ansieht:

$$\mathbf{M} = \left[\begin{bmatrix} \mathbf{A} - \mathbf{B}_2 \mathbf{D}_K \mathbf{C}_2 & \mathbf{B}_2 \mathbf{C}_K \\ -\mathbf{B}_K \mathbf{C}_2 & \mathbf{A}_K \end{bmatrix}, \begin{bmatrix} \mathbf{B}_3 \\ \mathbf{0} \end{bmatrix}, [\mathbf{C}_3 - \mathbf{D}_3 \mathbf{D}_K \mathbf{C}_2, \mathbf{D}_3 \mathbf{C}_K], \mathbf{0} \right] \tag{10.243}$$

11. Diskrete Systeme

11.1 Systemstruktur

Diskrete Systeme enthalten als Regler einen Mikroprozessor, was bei den heutigen Regelungssystemen der Normalfall ist (fig. 11.1).

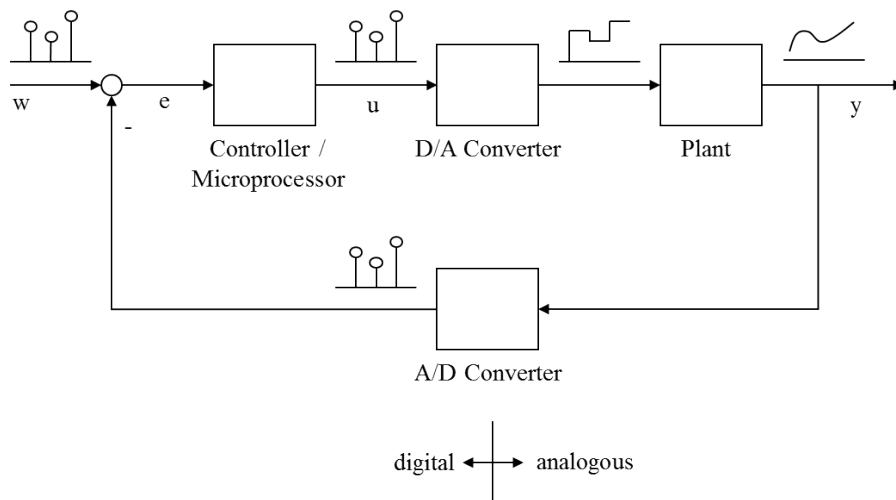


Abb. 11.1. Geschlossener Regelkreis mit Mikroprozessor als Regler

Der Mikroprozessor erhält Folgen von digitalen Werten als Eingangssignale und erzeugt Folgen von digitalen Werten als Ausgangssignale. Da die Strecke immer noch ein kontinuierliches und analoges physikalisches System ist, werden Wandler benötigt. Der D/A-Wandler wandelt eine von der Regelung kommende Folge von Digitalwerten in eine analoge Funktion um. Da jedoch zwischen zwei Werten der eingehenden Folge immer eine gewisse Zeit liegt, ist das Ausgangssignal des Wandlers eine sprungförmig verlaufende Funktion. Der A/D-Wandler hingegen wandelt das analoge und kontinuierliche Ausgangssignal der Strecke in eine Folge von digitalen Werten um, die vom Mikroprozessor verarbeitet werden können. Auch der Sollwert muss als digitaler Wert vorgegeben werden.

Die Zeit zwischen zwei Wandlungen des D/A- oder A/D-Wandlers muss immer gleich sein, sie wird Abtastzeit genannt, und im Folgenden wird davon ausgegangen, dass die Wandlungen in den Wandlern immer zur gleichen Zeit erfolgen, d. h. es gibt einen gemeinsamen Abtastzeitpunkt für das gesamte System. Dies kann natürlich nur eine theoretische Annahme sein, da beide Wandler in der Praxis unabhängig voneinander arbeiten. Die Abweichung zwischen Theorie und Praxis aufgrund dieser Annahme ist aber vernachlässigbar, während der Vorteil in der theoretischen Behandlung sehr groß ist.

Wegen der Notwendigkeit der Umsetzungen kann der Mikroprozessorregler normalerweise nicht mehr als stetiger Regler $K(s)$ angesehen werden. Stattdessen führen die Umrechnungen zu einem Informations- und Leistungsverlust, da

- Ausgangssignale nur in diskreten Zeitschritten gemessen werden (Zeitdiskretisierung),
- Reglerausgänge sich nur in diskreten Zeitschritten ändern können (Zeitdiskretisierung), und
- beliebige Werte je nach prozessorinterner Wertedarstellung nur mit einer begrenzten Genauigkeit dargestellt werden können (Quantifizierung).

Die folgenden Abbildungen zeigen die Auswirkungen von Zeitdiskretisierung und Quantifizierung. Wenn die Zeitdiskretisierung in sehr kleinen Zeitschritten im Vergleich zu den im Regelkreis auftretenden Signalen erfolgt, kann der Informationsverlust vernachlässigt werden, und der Regler verhält sich praktisch wie ein kontinuierlicher Regler. Der Quantifizierungseffekt war in den Anfängen der digitalen Mikroprozessoren von wesentlicher Bedeutung, aber heutzutage, bei der Verwendung von 32- oder 64-Bit-Systemen, kann der Informationsverlust durch die Quantifizierung vernachlässigt werden.

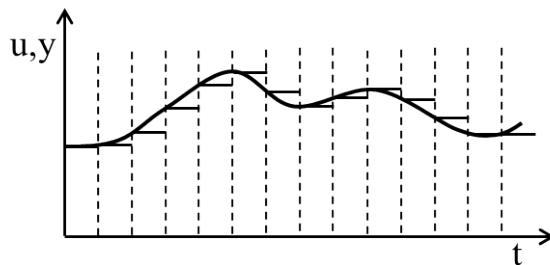


Abb. 11.2. Zeitdiskretisierung eines kontinuierlichen Signals

Im folgenden soll der Zusammenhang zwischen einem Regelkreis nach Abb. 11.1 und einem konventionellen Regelkreis mit analoger Strecke und Regler aufgestellt werden.

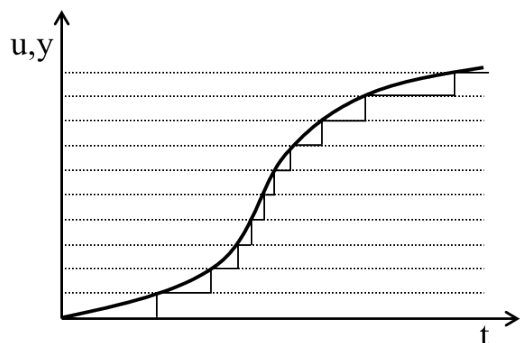


Abb. 11.3. Quantifizierung eines kontinuierlichen Signals

Zunächst benötigt man eine Einheit, die die Zeitdiskretisierung eines kontinuierlichen Signals durchführt, wie in Abb. 11.2 dargestellt. Man kann die Aufgabe in zwei Schritte aufteilen: Der erste Schritt besteht darin, zu jedem Abtastzeitpunkt einen einzelnen Wert aus dem kontinuierlichen Signal auszuwählen. Dieser Schritt wird *Sampling* genannt. Der zweite Schritt besteht darin, diesen Wert ein Abtastzeitintervall lang zu halten, bis der nächste abgetastete Wert kommt. Dieser Schritt wird *Hold* genannt. Beide Blöcke zusammen bilden den Block *Sample+Hold*.

Die Aufgabe besteht nun darin, die Funktion des Sample+Hold-Blocks durch konventionelle analoge Blöcke darzustellen. Die Lösung ist in Abb. 11.4 dargestellt. Die Abtastung kann durch die Multiplikation des eingehenden kontinuierlichen Signals mit einer Folge von Dirac-Impulsen $\delta(t)$ mit unendlicher Amplitude und einer Breite von Null, aber einem wohldefinierten Integralwert von Eins ausgedrückt werden:

$$\int_0^{\infty} \delta(\tau) d\tau = 1 \quad (11.1)$$

Der Ausgang des Sample-Blocks ist immer noch eine Folge von Dirac-Impulsen, wobei jeder dieser Impulse mit dem entsprechenden abgetasteten Wert des kontinuierlichen Signals multipliziert wird. Es stellt sich dabei jedoch die Frage, worin der Unterschied besteht zwischen einem Impuls mit unendlicher Amplitude und demselben Impuls, multipliziert mit einem Faktor. Die Antwort besteht darin, dass die Multiplikation mit einem reellen Wert den Integralwert des Dirac-Impulses ändert, nicht die Amplitude.

Diese Impulse mit unendlicher Amplitude sind notwendig. Impulse mit endlicher Amplitude und einer Breite von Null würden keine Energie enthalten, um eine Wirkung auf nachfolgende kontinuierliche Blöcke zu verursachen. Der Dirac-Impuls mit seinem von Null verschiedenen Integral jedoch enthält

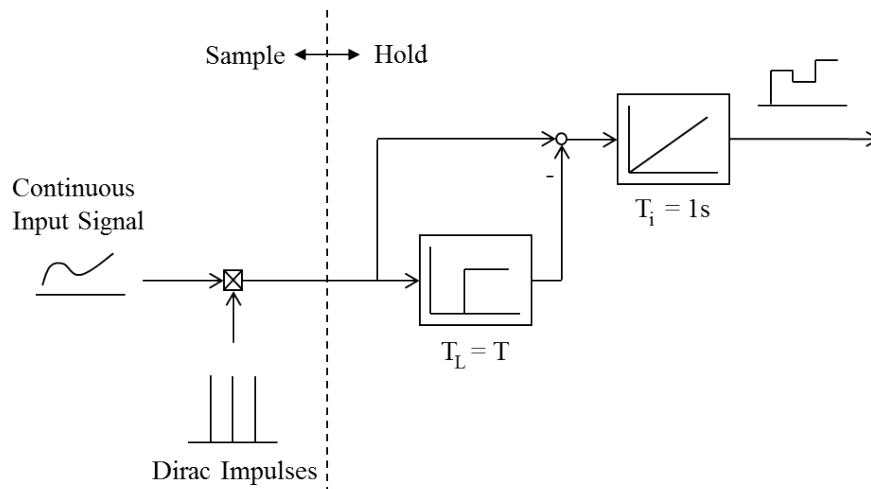


Abb. 11.4. Sample + Hold

Energie, so dass er eine Wirkung auf einen nachfolgenden kontinuierlichen Block haben kann, nämlich die bekannte Impulsantwort dieses Blocks.

Der Hold-Block ist viel einfacher zu verstehen. Jeder Dirac-Impuls aus dem Sample-Block bildet eine Impulsantwort am Integrator, die eine einfache Sprungfunktion ist. Da der Dirac-Impuls mit dem entsprechenden Wert des kontinuierlichen Eingangssignals multipliziert wurde, wird die Sprungfunktion mit demselben Wert multipliziert, so dass die Stufenhöhe gerade dem abgetasteten Wert des kontinuierlichen Signals entspricht. Einen Abtastschritt später erscheint aus dem Verzögerungsglied derselbe Impuls erneut am Integrator, nun mit negativem Vorzeichen. Die Ausgangsgröße des Integrators wird damit wieder auf Null gesetzt. Somit bildet der Hold-Block aus jedem Dirac-Impuls eine Rechteckfunktion mit der Dauer des Abtastintervalls und einer Amplitude, die dem abgetasteten Eingangssignalwert entspricht. Dies ist genau das, was benötigt wird, wie in Abb. 11.2 zu sehen ist.

In einem zweiten Schritt soll der Algorithmus des Mikroprozessor-Reglers mit herkömmlichen Blöcken beschrieben werden. Zu diesem Zweck ersetzt man den Mikroprozessorregler einfach durch eine klassische Reglerübertragungsfunktion. Da nun sowohl der Regler als auch die Strecke kontinuierliche Blöcke sind, könnte man den Ausgang des Reglers an die Strecke geben und den Ausgang der Strecke an den Regler zurückführen, was zu einem klassischen, analogen Regelkreis führen würde. Es muss aber auch der Informationsverlust modelliert werden, der in der realen Anlage durch die zeitliche Diskretisierung der D/A- und A/D-Wandler verursacht wird. Deshalb müssen in beiden Richtungen Sample+Hold-Blöcke zwischen Regler und Strecke geschaltet werden. Das Ergebnis ist in Abb. 11.5 zu sehen.

Der wesentliche Unterschied zu Abb. 11.1 ist, dass in diesem Blockdiagramm keine digitalen Werte mehr vorkommen, da Folgen von Dirac-

Impulsen keine digitalen Werte sind. Sie sind analoge Signale, wie alle anderen Signale in diesem Blockdiagramm auch. Und das ist der Grund, warum man nun die klassische Regelungstheorie auf dieses Blockdiagramm anwenden kann.

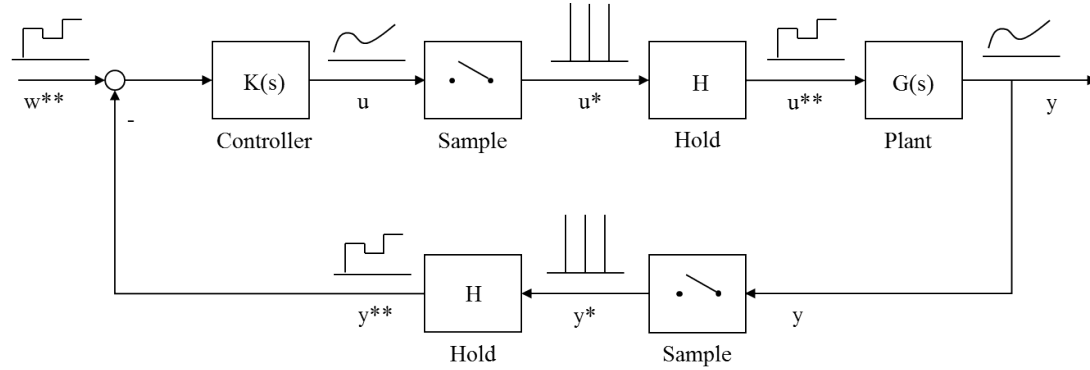


Abb. 11.5. Theoretische Darstellung eines geschlossenen Regelkreises mit Mikroprozessorregler

Zu beachten ist aber, dass der Sollwert im neuen Blockschaltbild als Stufenfunktion und nicht mehr als Folge von digitalen Werten anzugeben ist. Für die Realität spielt dies jedoch keine Rolle, da die gesamte Struktur sowieso rein theoretisch ist. Und zweitens ersetzen die Sample+Hold-Einheiten nicht direkt die D/A- bzw. A/D-Wandler, die vorher an den gleichen Stellen waren. Da nun sowohl der Regler als auch die Strecke kontinuierliche Übertragungsfunktionen sind, werden keine Wandler mehr benötigt. Die Sample+Hold-Blöcke stellen lediglich den Informationsverlust dar, der durch die Wandler im realen Regelkreis verursacht wird, aber sie arbeiten nicht mehr mit digitalen Werten.

Weiterhin ist auf die Signalbezeichnungen in Abb. 11.5 zu achten. Das kontinuierliche Ausgangssignal des Reglers ist $u(t)$, wie es aus den Grundlagen der Regelungstechnik bekannt ist. Nach der Abtastung ist dieses Signal eine Folge von Dirac-Impulsen, nämlich u^* , und nach dem Hold-Block ist es eine stufenförmige Funktion, nämlich u^{**} . Das Gleiche gilt für das Ausgangssignal des Systems y .

Da alle Blöcke in 11.5 lineare Blöcke sind, können Übertragungsfunktionen für dieses System definiert werden, z. B. $\frac{y(s)}{u(s)}$. Diese Übertragungsfunktion in einem kontinuierlichen System ist wohlbekannt, es ist die Übertragungsfunktion der Strecke. Aber hier umfasst diese Übertragungsfunktion auch den Sample- und den Hold-Block, so dass gilt:

$$\frac{y(s)}{u(s)} = G(s) H(s) S \quad (11.2)$$

mit dem Hold-Block $H(s)$ und dem Sample-Block S . Auf die gleiche Weise kann man die Übertragungsfunktion von einer Impulsfolge zu einer anderen definieren:

$$G^*(s) := \frac{y^*(s)}{u^*(s)} = S G(s) H(s) \quad (11.3)$$

Und schließlich kann man auch die Übertragungsfunktion von einer Stufenfunktion zu einer anderen definieren:

$$G^{**}(s) := \frac{y^{**}(s)}{u^{**}(s)} = H(s) S G(s) \quad (11.4)$$

Die Linearität aller Blöcke auf der rechten Seite der Gleichungen erlaubt die Anwendung des Kommutativgesetzes, so dass gilt:

$$G^{**}(s) = G^*(s) = G(s) H(s) S = \frac{y(s)}{u(s)} \quad (11.5)$$

Dieses interessante Resultat wird später noch verwendet.

11.2 Abtasttheorem

In diesem Abschnitt sollen die zeitliche Diskretisierung und das daraus resultierende Abtastproblem genauer untersucht werden. Wie bereits erwähnt, führt die Diskretisierung zu einem Informationsverlust, da die Messwerte nur zu bestimmten Zeitpunkten abgetastet werden, während der Rest des kontinuierlichen Eingangssignals übersprungen wird. Dies kann dazu führen, dass zwei unterschiedliche, kontinuierliche Eingangssignale am Ende die gleichen Ausgangssequenzen von Abtastwerten haben, was natürlich in einem geschlossenen Regelkreis ein großes Problem wäre.

Angenommen sei ein kontinuierliches Sinussignal $y_1(t) := \sin(\omega_1 t)$, das zu jedem Zeitpunkt mit $t = kT$ abgetastet wird, wobei $k = 0, 1, 2, 3, \dots$ und T die Abtastzeit ist. Die Abtastfrequenz kann definiert werden als $f_T := 1/T$, und die Abtastkreisfrequenz als $\omega_T := 2\pi f_T = 2\pi/T$. Während das ursprüngliche kontinuierliche Signal $y_1(t)$ eine zeitabhängige, kontinuierliche Funktion ist, ist das abgetastete Signal $y_1(kT)$ eine Folge von diskreten Werten.

Nun sei ein weiteres Sinussignal $y_2(t) := \sin(\omega_2 t)$ angenommen, und sein Abtastsignal soll mit dem Abtastsignal des ersten Sinussignals verglichen werden:

$$\begin{aligned} y_1(kT) &= \sin(\omega_1 kT) \\ y_2(kT) &= \sin(\omega_2 kT) \end{aligned} \quad (11.6)$$

Für

$$\omega_2 = \omega_1 + l \frac{2\pi}{T} \quad (11.7)$$

mit $l = 1, 2, 3, \dots$ erhält man

$$\sin(\omega_2 kT) = \sin(\omega_1 kT + kl2\pi) = \sin(\omega_1 kT) \quad (11.8)$$

was bedeutet, dass zwei verschiedene kontinuierliche Signale die gleichen Folgen von Abtastwerten haben. Dies ist der so genannte *Aliasing-Effekt*: Ein kontinuierliches Signal mit einer bestimmten (hohen) Frequenz kann nach der Abtastung als ein anderes Signal mit einer anderen (niedrigen) Frequenz erscheinen (Abb. 11.6). Dieser Effekt hat natürlich dramatische Folgen für einen Regler, wenn dieser nicht die echte Frequenz, sondern eine viel niedrigere Frequenz sieht und auf die virtuelle niedrige Frequenz reagiert.

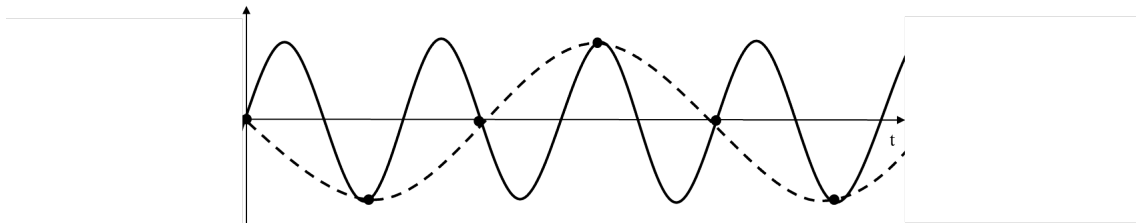


Abb. 11.6. Aliasing-Effekt bei zwei Sinussignalen

Der Aliasing-Effekt kann sogar ein virtuelles Nullsignal erzeugen, wie Abb. 11.7 zeigt. Es ist auch möglich, dass hochfrequentes Rauschen auf den Messungen zu Abtastwerten führt, die als virtuelle Sinuswelle interpretiert werden könnten, auch wenn das ursprüngliche Ausgangssignal der Strecke (ohne Rauschen) Null ist.

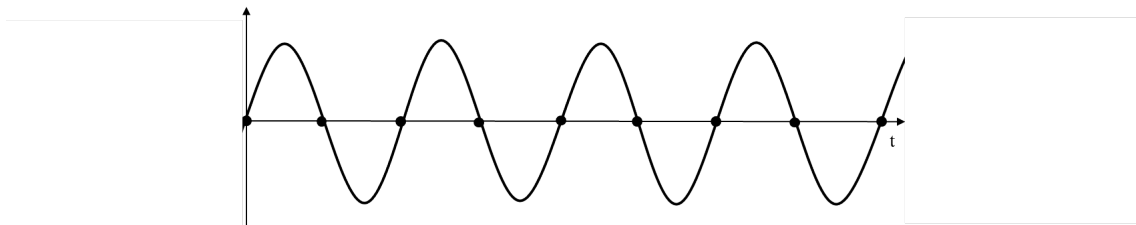


Abb. 11.7. Virtuelles Nullsignal durch Aliasing-Effekt

Die Beispiele zeigen, dass Probleme auftreten, wenn die Abtastfrequenz nicht hoch genug ist, um genügend Informationen aus dem kontinuierlichen Signal zu gewinnen. Zwei Werte pro Periode sind nicht genug, wie Abb. 11.6 zeigt. In der Tat ist dies nur der Grenzfall, wie mathematisch bewiesen werden kann. Es lässt sich ein Theorem formulieren, das nach Claude Elwood Shannon benannt ist, der dieses Theorem im Jahr 1948 formulierte. Vladimir Kotelnikov formulierte es bereits 1933, aber es war bis in die 80er Jahre des 20. Jahrhunderts in den westlichen Ländern unbekannt. Heutzutage wird es nach beiden Wissenschaftlern benannt:

Satz 11.1 *Satz von Kotelnikov-Shannon: Ein Signal, das verschiedene Frequenzen enthält, kann nur dann aus seinem abgetasteten Signal rekonstruiert werden, wenn die Abtastfrequenz mehr als das Zweifache der höchsten Frequenz im Signal beträgt:*

$$f_T > 2f_{max} \quad (11.9)$$

Die einzige Lösung für dieses Problem ist das Einfügen eines Anti-Aliasing-Filters in den geschlossenen Regelkreis, d.h. eines Tiefpassfilters mit einer Grenzfrequenz kleiner als $f_A = f_T/2$, um alle hochfrequenten Signale herauszufiltern, die durch Aliasing-Effekte Probleme verursachen würden (Abb. 11.8). Das einfachste Tiefpassfilter könnte ein Verzögerungsglied erster Ordnung mit einer Zeitkonstante $T_{lp} = 1/f_A$ sein (Abb. 11.9).

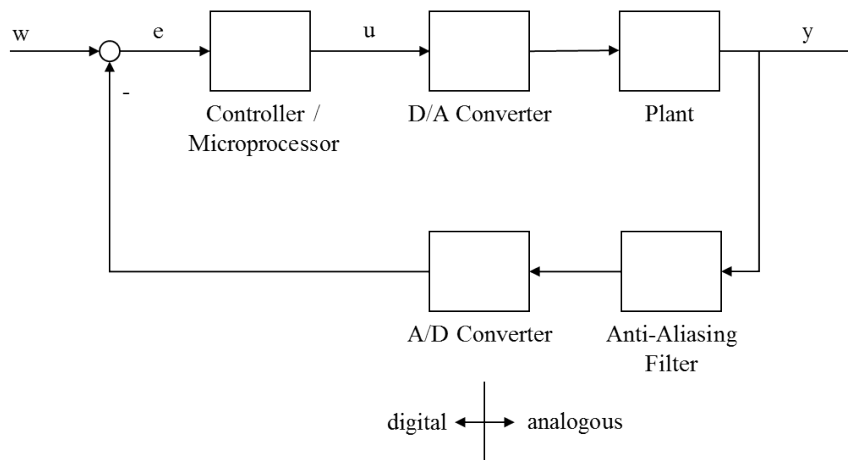


Abb. 11.8. Geschlossener Regelkreis mit Anti-Aliasing-Filter

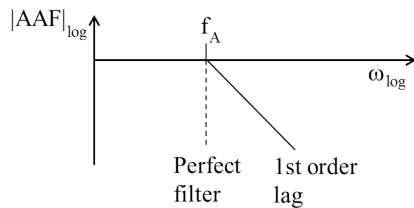


Abb. 11.9. Bode-Diagramm eines Anti-Aliasing-Filters

Aus Abb. 11.9 ist zu erkennen, dass die reale Filterkurve im Gegensatz zu einem theoretisch perfekten Filter hohe Frequenzen immer noch durchlässt, wenn auch gedämpft. Um eine bessere Unterdrückung des Aliasing-Effektes zu erzielen, wird die Grenzfrequenz des Aliasing-Filters in der Praxis daher mit $f_A = f_T/5$ oder sogar $f_A = f_T/10$ gewählt.

Das Anti-Aliasing-Filter löscht die hochfrequenten Signale. Daraus folgt, dass Signale mit zu hohen Frequenzen von einer Mikroprozessorregelung mit einer für diese Signale zu geringen Abtastrate einfach nicht geregelt werden können. Hochfrequente Strecken benötigen daher leistungsfähige Mikroprozessoren mit ausreichender Abtastrate als Regler. Andererseits können für langsame Prozesse einfachere Mikroprozessoren mit niedriger Abtastrate gewählt werden.

Ein weiterer zu beachtender Punkt des Anti-Aliasing-Filters ist die zusätzliche Phasenverschiebung der Kreisübertragungsfunktion, die durch das Tiefpassverhalten verursacht wird. Die zusätzliche Phasenverschiebung verringert die Stabilitätsreserve, wie anhand des Nyquist-Kriteriums leicht zu erkennen ist, so dass die Reglerverstärkung auf einen niedrigeren Wert eingestellt werden muss. Dies verringert jedoch die Geschwindigkeit des geschlossenen Regelkreises, was ein weiterer Grund für die Verwendung eines Hochleistungsmikroprozessors sein kann. Denn in diesem Fall kann die Zeitkonstante des Anti-Aliasing-Filters viel kleiner gewählt werden, die Phasenverschiebung wird verringert und die Reglerverstärkung kann erhöht werden.

11.3 Differenzengleichung

Die Frage ist nun, wie man einen kontinuierlichen, analogen Regelalgorithmus, z.B. einen PID-Regler, mit Hilfe eines Mikroprozessors zumindest näherungsweise realisieren kann.

Ausgangspunkt der Überlegungen ist die Übertragungsfunktion des kontinuierlichen Reglers:

$$K(s) = \frac{u(s)}{e(s)} = \frac{b_{c,n}s^n + b_{c,n-1}s^{n-1} + \dots + b_{c,1}s + b_{c,0}}{a_{c,n}s^n + a_{c,n-1}s^{n-1} + \dots + a_{c,1}s + a_{c,0}} \quad (11.10)$$

Der Index c zeigt an, dass die Koeffizienten das kontinuierliche Systemverhalten beschreiben.

Um die folgenden Berechnungen einfach zu halten, sollen sie nur für ein System 2. Ordnung durchgeführt werden, aber die Überlegungen für Systeme n -ter Ordnung sind genau die gleichen:

$$K(s) = \frac{u(s)}{e(s)} = \frac{b_{c,2}s^2 + b_{c,1}s + b_{c,0}}{a_{c,2}s^2 + a_{c,1}s + a_{c,0}} \quad (11.11)$$

Durch Multiplikation der Nenner der beiden Brüche auf der rechten und linken Seite der Gleichung erhält man

$$b_{c,2}s^2e(s) + b_{c,1}se(s) + b_{c,0}e(s) = a_{c,2}s^2u(s) + a_{c,1}su(s) + a_{c,0}u(s) \quad (11.12)$$

Dies ist eine Differentialgleichung im Frequenzbereich, die zunächst in den Zeitbereich transformiert werden soll. Unter Verwendung von Gleichung

(A.10) aus Abschnitt A.1 im Anhang erhält man unter der Annahme, dass sämtliche Anfangswerte gleich Null sind,

$$\mathcal{L} \left\{ \frac{d^n f(t)}{dt^n} \right\} = s^n \mathcal{L} \{f(t)\} \quad (11.13)$$

Diese Beziehung kann nun verwendet werden, um Gleichung (11.12) in den Zeitbereich zu transformieren:

$$\begin{aligned} b_{c,2} \left(\frac{d}{dt} \right)^2 e(t) + b_{c,1} \left(\frac{d}{dt} \right) e(t) + b_{c,0} e(t) \\ = a_{c,2} \left(\frac{d}{dt} \right)^2 u(t) + a_{c,1} \left(\frac{d}{dt} \right) u(t) + a_{c,0} u(t) \end{aligned} \quad (11.14)$$

Da eine kontinuierliche Ableitung mit einem diskontinuierlich arbeitenden Mikroprozessor nicht möglich ist, benötigt man eine Approximation des Ableitungsoperators. Diese Approximation ist der *Differenzquotientenoperator*

$$\left(\frac{d}{dt} \right) e(t) \rightarrow \Delta e(t) := \frac{e(t) - e(t-T)}{T} = \frac{e(kT) - e((k-1)T)}{T} \quad (11.15)$$

mit Abtastwerten zu diskreten Zeitpunkten $t = kT$. Dieser Differenzquotientenoperator kann mehrfach angewendet werden, so dass man auch für Ableitungen höherer Ordnung eine Näherung erhält:

$$\begin{aligned} \left(\frac{d}{dt} \right)^2 e(t) \rightarrow \Delta^2 e(t) &:= \Delta(\Delta e(t)) = \Delta \left(\frac{e(kT) - e((k-1)T)}{T} \right) \\ &= \Delta \left(\frac{e(kT)}{T} \right) - \Delta \left(\frac{e((k-1)T)}{T} \right) \\ &= \frac{e(kT) - e((k-1)T) - e((k-1)T) + e((k-2)T)}{T^2} \\ &= \frac{e(kT) - 2e((k-1)T) + e((k-2)T)}{T^2} \end{aligned} \quad (11.16)$$

Diese Resultate werden nun auf Gleichung (11.14) angewandt:

$$\begin{aligned} b_{c,2} \frac{e(kT) - 2e((k-1)T) + e((k-2)T)}{T^2} + \\ b_{c,1} \frac{e(kT) - e((k-1)T)}{T} + b_{c,0} e(kT) = \\ a_{c,2} \frac{u(kT) - 2u((k-1)T) + u((k-2)T)}{T^2} + \\ a_{c,1} \frac{u(kT) - u((k-1)T)}{T} + a_{c,0} u(kT) \end{aligned} \quad (11.17)$$

Umsortieren und Zusammenfassen der Terme führt auf

$$\begin{aligned} (b_{c,2} + Tb_{c,1} + T^2 b_{c,0}) e(kT) + (-2b_{c,2} - Tb_{c,1}) e((k-1)T) \\ + b_{c,2} e((k-2)T) = \\ (a_{c,2} + Ta_{c,1} + T^2 a_{c,0}) u(kT) + (-2a_{c,2} - Ta_{c,1}) u((k-1)T) + \\ a_{c,2} u((k-2)T) \end{aligned} \quad (11.18)$$

Nach einer Neubenennung der Koeffizienten erhält man

$$\begin{aligned} b_2 e(kT) + b_1 e((k-1)T) + b_0 e((k-2)T) = \\ a_2 u(kT) + a_1 u((k-1)T) + a_0 u((k-2)T) \end{aligned} \quad (11.19)$$

wobei die Koeffizienten ohne Index c nun für ein diskretes System stehen. Es ist zu erkennen, dass diese Differenzengleichung die gleiche Struktur hat wie die entsprechende Differentialgleichung (11.12), jedoch mit anderen Koeffizienten. Man sollte sich der Tatsache bewusst sein, dass die Abtastzeit T in die Berechnung der neuen Koeffizienten eingeht. Wenn sich also die Abtastzeit ändert, ändern sich auch die Koeffizienten.

Zur Vervollständigung sei die allgemeine Gleichung n -ter Ordnung angegeben:

$$\sum_{i=0}^n b_i e((k+i-n)T) = \sum_{i=0}^n a_i u((k+i-n)T) \quad (11.20)$$

wobei zur Vereinfachung auf die Abtastzeit T verzichtet werden kann:

$$\sum_{i=0}^n b_i e(k+i-n) = \sum_{i=0}^n a_i u(k+i-n) \quad (11.21)$$

Dies ist nun die endgültige und allgemeine Version einer Differenzengleichung. Sie beschreibt die notwendigen Berechnungen im Mikroprozessor, um ein ähnliches Verhalten zu erreichen wie der ursprüngliche, kontinuierliche Regler, der durch $K(s)$ gegeben ist. Die Berechnungsformel wird viel klarer, wenn man diese Gleichung umstellt und $a_n := 1$ definiert, was nach einer Division der gesamten Gleichung (11.21) leicht erreicht werden kann:

$$u(k) = \sum_{i=0}^n b_i e(k+i-n) - \sum_{i=0}^{n-1} a_i u(k+i-n) \quad (11.22)$$

Diese Formel beschreibt anschaulich, wie der aktuelle Ausgang des Reglers auf der Grundlage der in den vorherigen Zeitschritten berechneten Ausgangswerte $u(k+i-n)$ und der früheren (und aktuellen) Eingangswerte $e(k+i-n)$ zu berechnen ist. Es handelt sich um eine rekursive Formel, die zur Berechnung des aktuellen Wertes immer die Werte der Vergangenheit benötigt.

In einem realen Mikroprozessor ist es nicht realistisch, einen aktuellen Eingangswert $e(k)$ zu erhalten und den darauf aufbauenden, aktuellen Ausgangswert $u(k)$ ohne Zeitverzögerung zu berechnen. Da die Berechnung immer Zeit benötigt, kann der aktuelle Eingangswert nicht verwendet werden, und manchmal überspringt man sogar nicht nur den aktuellen Eingangswert, sondern auch die Eingangswerte davor. Daher muss die obere Grenze der ersten Summe in dieser Gleichung durch einen Wert $q < n$ ersetzt werden:

$$u(k) = \sum_{i=0}^q b_i e(k+i-n) - \sum_{i=0}^{n-1} a_i u(k+i-n) \quad (11.23)$$

Die Differenz zwischen n und q wird als *Differenzengrad* bezeichnet: $r := n - q$. Er gibt die Zeitverzögerung zwischen einer Änderung des Eingangssignals und der entsprechenden Reaktion des Ausgangssignals an. Bei einer Reglerdifferenzengleichung ist r immer größer als Null, bei einer Differenzengleichung, die das Verhalten einer kontinuierlichen Strecke beschreibt, könnte r auch Null sein, da es keinen Mikroprozessor gibt, der Zeit für die Verarbeitung von Eingangswerten benötigt.

Wenn die zweite Summe ganz verschwindet, ergibt sich ein sehr spezieller Fall der Differenzengleichung:

$$u(k) = \sum_{i=0}^q b_i e(k + i - n) \quad (11.24)$$

Dabei haben die früheren Ausgangswerte keinen Einfluss auf den aktuellen Ausgangswert. Dieses Verhalten wird als *gleitender Durchschnitt* bzw. *moving average* bezeichnet. Ein einfaches Beispiel für ein solches System lässt sich erstellen, wenn alle b_i den gleichen Wert $b_i = 1/(q + 1)$ haben. In diesem Fall ist $u(k)$ immer der Durchschnitt der entsprechenden $q + 1$ Eingabewerte, wobei sich diese Menge von Eingabewerten in jedem Zeitschritt ändert, indem der nächste Eingabewert aufgenommen und der älteste gelöscht wird. Offensichtlich ist der Name *gleitender Durchschnitt* somit sinnvoll.

11.4 Diskretes Zustandsraummodell

11.4.1 Einleitung

Man kann das Konzept der Differenzengleichung von SISO- auf MIMO-Systeme ausdehnen, indem man eine Form verwendet, die auf der gleichen Idee beruht wie die kontinuierliche Zustandsraumbeschreibung. Wieder benötigt man einen Zustandsvektor \mathbf{x} , aber jetzt definiert die Systemgleichung nicht die Ableitung dieses Vektors. Stattdessen zeigt sie, wie der zukünftige Zustandsvektor im nächsten Zeitschritt aus dem Zustand und dem Eingangsvektor des vorhergehenden Zeitschrittes berechnet wird. Die Ausgangsgleichung dieses diskreten Zustandsraummodells ist genau dieselbe wie im kontinuierlichen Fall, da die Ausgangsgleichung keine Dynamik enthält:

$$\begin{aligned} \mathbf{x}(k + 1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \quad (11.25)$$

Für dieses Modell kann man ein Blockdiagramm erstellen, wie in Abb. 11.10 gezeigt. Es hat natürlich die gleiche Struktur wie das Blockdiagramm eines kontinuierlichen Zustandsmodells. Lediglich der komponentenweise Integratorblock wurde durch einen Block mit dem Namen $1/z$ ersetzt. Der Grund für diesen Namen wird später erklärt, aber die Funktion ist klar:

Dieser Block muss eine komponentenweise Verzögerung für einen Zeitschritt darstellen. Zur Erklärung sei angenommen, dass $t = kT$ der aktuelle Zeitpunkt ist. Die eingehenden Signale des Blocks sind $\mathbf{Ax}(k)$ und $\mathbf{Bu}(k)$. Aus der Zustandsgleichung geht hervor, dass am Eingang dieses Blocks $\mathbf{x}(k+1)$ anliegt, was bedeutet, dass dieser Eingang der zukünftige Zustandsvektor im nächsten Zeitschritt sein wird, während der aktuelle Zeitpunkt ja $t = kT$ ist. Einen Zeitschritt später erscheint das Eingangssignal am Ausgang des Blockes, und da sich die aktuelle Zeit nun von k auf $(k+1)$ erhöht hat, ist dieser Vektor nun der neue, aktuelle Zustandsvektor.

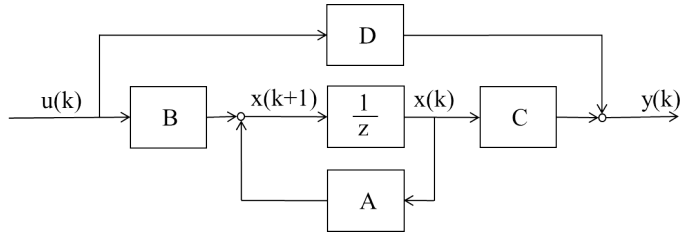


Abb. 11.10. Diskretes Zustandsraummodell

11.4.2 Normalformen

Mit Hilfe des Verzögerungsblocks $1/z$ lässt sich die Differenzengleichung (11.22) auch in einem Blockdiagramm darstellen. Ausgangspunkt sei diese Gleichung, jetzt aber für eine Strecke mit der Eingangsgröße u und der Ausgangsgröße y , um einen Zusammenhang mit Abb. 11.10 herzustellen:

$$y(k) = \sum_{i=0}^n b_i u(k+i-n) - \sum_{i=0}^{n-1} a_i y(k+i-n) \quad (11.26)$$

Um nur die erste Summe der rechten Seite darzustellen, benötigt man n Verzögerungsblöcke in Reihenschaltung (siehe Abb. 11.11). Der letzte Term dieser Summe $b_n u(k)$ beschreibt eine direkte Verbindung vom Eingang zum Ausgang über den Koeffizienten b_n , der im Blockdiagramm leicht zu finden ist. Der vorletzte Term $b_{n-1} u(k-1)$ beinhaltet das Eingangssignal des letzten Zeitschritts, multipliziert mit b_{n-1} . Um diese Beziehung im Blockdiagramm herzustellen, muss das Eingangssignal einen Koeffizienten b_{n-1} durchlaufen und dann um einen Verzögerungsblock verzögert werden, da der Eingang aus dem Zeitschritt $k-1$ einen Zeitschritt lang gehalten werden muss, bevor er einen Beitrag zum aktuellen Ausgangssignal des aktuellen Zeitpunkts $t = kT$ leisten kann. Genau dieser Pfad ist im Blockdiagramm zu finden. Für die anderen Terme der Eingangssumme gilt das gleiche Schema, nur mit einer Verzögerung höherer Ordnung.

Der untere Teil des Blockdiagramms stellt die zweite Summe von Gl. (11.26) dar. Der erste Term der Summe, $-a_0 y(k-n)$, ist das Ausgangssignal

vor n Zeitschritten, multipliziert mit $-a_0$. Dies wird durch die Rückkopplung über den Koeffizienten a_0 dargestellt. Der Ausgangswert wird an den ganz linken Punkt des Blockdiagramms zurückgeführt und erscheint erst wieder am Ausgang (und trägt als Teil der zweiten Summe zum Ausgangssignal bei), nachdem er n Verzögerungsblöcke durchlaufen hat, also n Zeitschritte später. Diese Rückkopplung stellt damit genau den ersten Term der zweiten Summe dar. Für alle anderen Terme ist die Erklärung die gleiche. Diese spezielle Form eines SISO-Blockdiagramms wird als *Beobachternormalform* bezeichnet.

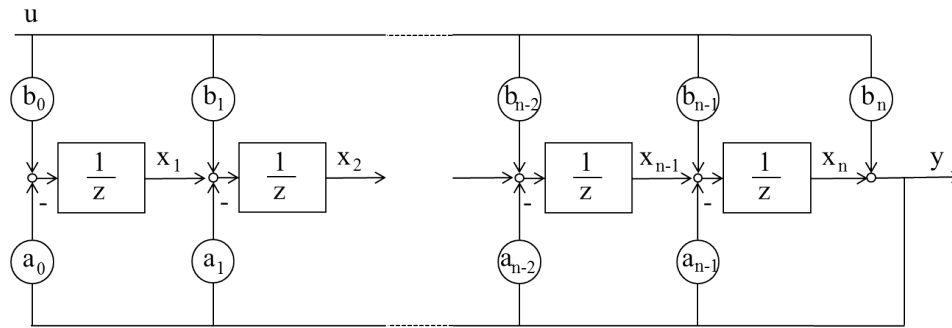


Abb. 11.11. Beobachternormalform eines diskreten Systems

Aus diesem Blockdiagramm kann man nun das diskrete Zustandsmodell dieses Systems erstellen. Zunächst müssen Zustandsgrößen definiert werden. Von der kontinuierlichen Zustandsdarstellung ist bekannt, dass die Zustandsgrößen immer Ausgänge von Integratoren sind. Hier, in diskreten Systemen, sind die Zustandsgrößen immer Ausgänge von $1/z$ -Blöcken. Nach der Festlegung dieser Größen wie im Blockdiagramm gezeigt ist es einfach, die Zustandsgleichung zu entwickeln. Man muss sich nur klarmachen, dass man am Eingang eines Verzögerungsblocks immer den Wert hat, der einen Zeitschritt später am Ausgang dieses Blocks erscheint. Daher steht jeder Verzögerungsblock für eine Zustandsgleichung.

Für den ersten Verzögerungsblock zum Beispiel ist der Ausgang $x_1(k)$, und der Eingang muss $x_1(k+1)$ sein. Am Eingang dieses Blocks müssen nun alle Eingänge aufsummiert werden, um die Zustandsgleichung für diese erste Zustandsgröße zu erhalten:

$$\begin{aligned} b_0 u(k) - a_0 y(k) &= b_0 u(k) - a_0 (x_n(k) + b_n u(k)) \\ &= -a_0 x_n(k) + (b_0 - a_0 b_n) u(k) \end{aligned} \quad (11.27)$$

Für den zweiten Verzögerungsblock ergibt sich:

$$\begin{aligned} b_1 u(k) - a_1 y(k) + x_1(k) &= b_1 u(k) + x_1(k) - a_1 (x_n(k) + b_n u(k)) \\ &= -a_1 x_n(k) + x_1(k) + (b_1 - a_1 b_n) u(k) \end{aligned} \quad (11.28)$$

Beide Ergebnisse sind auch in der Zustandsgleichung der Beobachter-Normalform zu finden:

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & & -a_0 \\ 1 & & -a_1 \\ 0 & 1 & -a_2 \\ & \ddots & \\ & & 1 & -a_{n-1} \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \vdots \\ x_n(k) \end{pmatrix} + \begin{pmatrix} b_0 - b_n a_0 \\ b_1 - b_n a_1 \\ b_2 - b_n a_2 \\ \vdots \\ b_n - b_n a_{n-1} \end{pmatrix} u(k)$$

$$y(k) = (0, \dots, 0, 1) \mathbf{x}(k) + b_n u(k) \quad (11.29)$$

Alternativ kann man als gleichwertige Darstellung auch die *Regelungsnormalform* verwenden, analog zu einem kontinuierlichen System (Abb. 11.12).

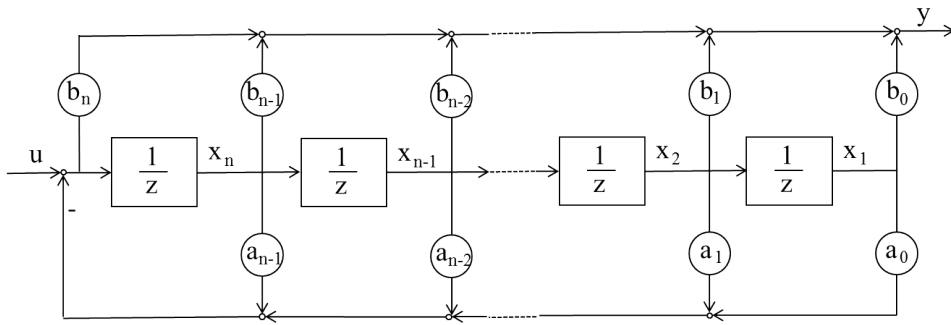


Abb. 11.12. Regelungsnormalform für ein diskretes System

Hier lautet die Zustandsgleichung

$$\mathbf{x}(k+1) = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -a_{n-1} \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix} \mathbf{x}(k) + \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} u(k)$$

$$y(k) = (b_0 - b_n a_0, b_1 - b_n a_1, \dots, b_{n-1} - b_n a_{n-1}) \mathbf{x}(k) + b_n u(k) \quad (11.30)$$

11.4.3 Diskretes und kontinuierliches Zustandsmodell

In diesem Abschnitt soll der Zusammenhang zwischen den Systemmatrizen der kontinuierlichen und der diskreten Zustandsdarstellung erarbeitet werden. Ausgangspunkt ist das kontinuierliche Zustandsmodell einer Strecke:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}_c \mathbf{x}(t) + \mathbf{D}_c \mathbf{u}(t) \end{aligned} \quad (11.31)$$

Vor und hinter dieser Strecke befinden sich Sample+Hold-Blöcke im Regelkreis (Abb. 11.5). Daraus folgt:

1. $\mathbf{u}(t) = \mathbf{u}(k)$ für $kT \leq t < (k+1)T$
2. Die Werte des Ausgangsvektors $\mathbf{y}(t)$ müssen nur für die Abtastzeitpunkte berechnet werden, da alle anderen Werte durch die Abtastung sowieso nicht erfasst werden.

Nun sollen die Matrizen des diskreten Zustandsmodells aus den Matrizen des kontinuierlichen Zustandsmodells berechnet werden:

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)\end{aligned}\quad (11.32)$$

Für die Ausgangsgleichung ist diese Aufgabe trivial, da in beiden Ausgangsgleichungen nur Multiplikationen der Signalwerte mit konstanten Werten und keine dynamischen Operationen vorliegen. Man erhält:

$$\mathbf{C} = \mathbf{C}_c \quad \mathbf{D} = \mathbf{D}_c \quad (11.33)$$

Aber was ist mit den beiden anderen Matrizen? Ausgangspunkt der Berechnung ist die allgemeine Lösung der kontinuierlichen Zustandsgleichung (7.94):

$$\mathbf{x}(t) = e^{\mathbf{A}_c(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}_c(t-\tau)}\mathbf{B}_c\mathbf{u}(\tau)d\tau \quad (11.34)$$

mit der im Anhang definierten Matrixexponentialfunktion (A.3)

$$e^{\mathbf{A}t} := \mathbf{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \mathbf{A}^3\frac{t^3}{3!} + \dots = \sum_{k=0}^{\infty} \mathbf{A}^k \frac{t^k}{k!} \quad (11.35)$$

Für den aktuellen Zeitpunkt $t := (k+1)T$ und den Anfangszeitpunkt $t_0 := kT$ erhält man

$$\mathbf{x}((k+1)T) = e^{\mathbf{A}_c T}\mathbf{x}(kT) + \int_{kT}^{(k+1)T} e^{\mathbf{A}_c((k+1)T-\tau)}\mathbf{B}_c\mathbf{u}(\tau)d\tau \quad (11.36)$$

Unter Verwendung der Tatsache, dass $\mathbf{u}(\tau)$ im Integrationszeitintervall seinen konstanten Wert $\mathbf{u}(kT)$ behält, ergibt sich

$$\mathbf{x}((k+1)T) = e^{\mathbf{A}_c T}\mathbf{x}(kT) + \int_{kT}^{(k+1)T} e^{\mathbf{A}_c((k+1)T-\tau)}\mathbf{B}_c\mathbf{u}(kT)d\tau \quad (11.37)$$

Um das Integral zu vereinfachen, definiert man $\alpha := (k+1)T - \tau$. Damit muss $d\tau$ durch $-d\alpha$ ersetzt werden, die obere Integrationsgrenze durch $\alpha = (k+1)T - \tau = (k+1)T - (k+1)T = 0$ und die untere Integrationsgrenze durch $\alpha = (k+1)T - \tau = (k+1)T - kT = T$:

$$\mathbf{x}((k+1)T) = e^{\mathbf{A}_c T} \mathbf{x}(kT) - \int_T^0 e^{\mathbf{A}_c \alpha} \mathbf{B}_c \mathbf{u}(kT) d\alpha \quad (11.38)$$

Das Vertauschen der Integrationsgrenzen und das Herausnehmen der konstanten Terme aus dem Integral führt zu

$$\mathbf{x}((k+1)T) = e^{\mathbf{A}_c T} \mathbf{x}(kT) + \int_0^T e^{\mathbf{A}_c \alpha} d\alpha \mathbf{B}_c \mathbf{u}(kT) \quad (11.39)$$

Der Vergleich mit der ursprünglichen diskreten Zustandsgleichung liefert die Ergebnisse für die diskreten Matrizen \mathbf{A} und \mathbf{B} :

$$\begin{aligned} \mathbf{A} &= e^{\mathbf{A}_c T} \\ \mathbf{B} &= \int_0^T e^{\mathbf{A}_c \alpha} d\alpha \mathbf{B}_c \end{aligned} \quad (11.40)$$

Man erkennt, dass in beiden Formeln die Abtastzeit enthalten ist, d.h. eine Änderung der Abtastzeit führt zu anderen diskreten Systemmatrizen.

11.4.4 Eigenwerte von diskreten Systemen

In diesem Abschnitt soll die Beziehung zwischen den Eigenwerten der Systemmatrizen des kontinuierlichen und des diskreten Systems betrachtet werden. Zur Vereinfachung sei angenommen, dass das kontinuierliche System nur einzelne Eigenwerte hat und in Jordan-Normalform gegeben ist:

$$\mathbf{A}_c = \begin{pmatrix} \lambda_{c,1} & & \\ & \dots & \\ & & \lambda_{c,n} \end{pmatrix} \quad (11.41)$$

Für die diskrete Systemmatrix ergibt sich aus Gleichung (11.40) und aus der Definition der Matrizenexponentialfunktion (11.35)

$$\begin{aligned} \mathbf{A} &= e^{\mathbf{A}_c T} = \sum_{k=0}^{\infty} \mathbf{A}_c^k \frac{T^k}{k!} = \sum_{k=0}^{\infty} \begin{pmatrix} \lambda_{c,1} & & \\ & \dots & \\ & & \lambda_{c,n} \end{pmatrix}^k \frac{T^k}{k!} \\ &= \sum_{k=0}^{\infty} \begin{pmatrix} \lambda_{c,1}^k & & \\ & \dots & \\ & & \lambda_{c,n}^k \end{pmatrix} \frac{T^k}{k!} \\ &= \sum_{k=0}^{\infty} \begin{pmatrix} \frac{(\lambda_{c,1} T)^k}{k!} & & \\ & \dots & \\ & & \frac{(\lambda_{c,n} T)^k}{k!} \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{(\lambda_{c,1}T)^k}{k!} & & \dots & \\ & & & \sum_{k=0}^{\infty} \frac{(\lambda_{c,n}T)^k}{k!} \end{pmatrix} \\
&= \begin{pmatrix} e^{\lambda_{c,1}T} & & & \\ & \dots & & \\ & & e^{\lambda_{c,n}T} & \end{pmatrix}
\end{aligned} \tag{11.42}$$

Somit erhält man den einfachen Zusammenhang

$$\lambda_i = e^{\lambda_{c,i}T} \tag{11.43}$$

Interessant ist, dass die Eigenwerte eines diskreten Systems, solange es durch Diskretisierung aus einem kontinuierlichen System entwickelt wird, niemals gleich Null sein können. Und weil $\det \mathbf{A} = \prod \lambda_i \neq 0$ ist, folgt daraus, dass eine solche diskrete Systemmatrix immer regulär und damit auch invertierbar ist.

11.4.5 Totzeiten in diskreter Zustandsdarstellung

Dieser Unterabschnitt befasst sich mit einem Thema, das zwar klein, aber sehr interessant ist. Angenommen sei eine SISO-Strecke, bei der die Dynamik nur aus einem Totzeitelement besteht: $y(t) = u(t - T_L)$, und die Totzeit soll ein ganzzahliges Vielfaches der Abtastzeit betragen: $T_L = aT$ mit $a \in \mathbb{N}$. Daraus folgt, dass diese Strecke durch n Verzögerungsblöcke der Dauer T dargestellt werden kann, wie in Abb. gezeigt. 11.13.

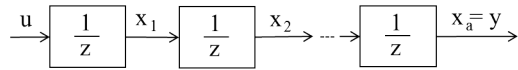


Abb. 11.13. Totzeitstrecke mit $y(t) = u(t - aT)$

Die Zustandsgrößen sind wie folgt definiert:

$$\begin{aligned}
x_1(k+1) &= u(k) \\
x_2(k+1) &= x_1(k) \\
x_3(k+1) &= x_2(k) \\
x_a(k+1) &= x_{a-1}(k) \\
y(k) &= x_a(k)
\end{aligned} \tag{11.44}$$

Somit ergibt sich die Zustandsdarstellung

$$\mathbf{x}(k+1) = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ 0 & 1 & & \\ & & \dots & \\ & & & 1 & 0 \end{pmatrix} \mathbf{x}(k) + \begin{pmatrix} 1 \\ 0 \\ & \\ & \\ 0 \end{pmatrix} u(k)$$

$$y(k) = (0, \dots, 0, 1)\mathbf{x}(k) \quad (11.45)$$

Das Bemerkenswerte an diesem Ergebnis ist, dass es möglich ist, Totzeiten im diskreten Zustandsraum zu beschreiben, während dies im kontinuierlichen Zustandsraum nicht möglich ist.

11.4.6 Lösung der diskreten Zustandsgleichung

In diesem Abschnitt soll die allgemeine Lösung der diskreten Zustandsgleichung berechnet werden, d.h. der Verlauf der Vektorfolge $\mathbf{x}(k)$, welche die Systemgleichung (11.25) erfüllt:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \quad (11.46)$$

Aus der ersten Gleichung ergibt sich mit dem Anfangszustand $\mathbf{x}(0)$ direkt

$$\begin{aligned} \mathbf{x}(1) &= \mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0) \\ \mathbf{x}(2) &= \mathbf{A}\mathbf{x}(1) + \mathbf{B}\mathbf{u}(1) = \mathbf{A}(\mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0)) + \mathbf{B}\mathbf{u}(1) \\ &= \mathbf{A}^2\mathbf{x}(0) + \mathbf{A}\mathbf{B}\mathbf{u}(0) + \mathbf{B}\mathbf{u}(1) \\ \mathbf{x}(3) &= \mathbf{A}\mathbf{x}(2) + \mathbf{B}\mathbf{u}(2) = \mathbf{A}(\mathbf{A}^2\mathbf{x}(0) + \mathbf{A}\mathbf{B}\mathbf{u}(0) + \mathbf{B}\mathbf{u}(1)) + \mathbf{B}\mathbf{u}(2) \\ &= \mathbf{A}^3\mathbf{x}(0) + \mathbf{A}^2\mathbf{B}\mathbf{u}(0) + \mathbf{A}\mathbf{B}\mathbf{u}(1) + \mathbf{B}\mathbf{u}(2) \\ \mathbf{x}(k) &= \mathbf{A}^k\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{A}^{k-j-1}\mathbf{B}\mathbf{u}(j) \end{aligned} \quad (11.47)$$

Die endgültige Gleichung erscheint mit einem Blick auf das Blockdiagramm des diskreten Systems in Abb.11.10 plausibel. Zum Zeitschritt $k=0$ befindet sich der Anfangszustandsvektor $\mathbf{x}(0)$ in der Schleife am Ausgang des $1/z$ -Blocks und erscheint daher gleichzeitig als $\mathbf{A}\mathbf{x}(0)$ am Eingang des $1/z$ -Blocks. Der erste Wert des Eingangsvektors $\mathbf{u}(0)$ erscheint nach Multiplikation mit \mathbf{B} als $\mathbf{B}\mathbf{u}(0)$ ebenfalls am Eingang des $1/z$ -Blocks. Einen Zeitschritt später, bei $k=1$, erscheinen beide Signale am *Ausgang* des $1/z$ -Blocks und bilden den neuen Zustandsvektor $\mathbf{x}(1) = \mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0)$. Am Eingang des Blocks $1/z$ erhält man nun $\mathbf{A}\mathbf{x}(1)$ und $\mathbf{B}\mathbf{u}(1)$. Bei $k=2$ erscheinen diese beiden Terme am Ausgang des Blocks $1/z$ und bilden den neuen Zustandsvektor $\mathbf{x}(2)$. Auf diese Weise wird jeder neue Eingangsvektor mit \mathbf{B} multipliziert und dann zu allen anderen Vektoren addiert, die sich bereits in der Schleife befinden und bei jedem Zeitschritt mit \mathbf{A} multipliziert werden.

Gleichung (11.47) kann mit der allgemeinen Lösung des kontinuierlichen Zustandsraums verglichen werden, die bereits in Gleichung (11.34) verwendet wurde:

$$\mathbf{x}(t) = e^{\mathbf{A}_c(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}_c(t-\tau)}\mathbf{B}_c\mathbf{u}(\tau)d\tau \quad (11.48)$$

Offensichtlich hat diese Gleichung die gleiche Struktur wie die diskrete Version. Daher gibt es auch eine Übergangsmatrix für das diskrete System \mathbf{A}^k , die den Einfluss des Anfangszustands auf den aktuellen Zustand zeigt.

Durch Einsetzen von Gl. (11.47) in die Ausgangsgleichung von (11.25) erhält man die Lösung für die Ausgangsfolge:

$$\begin{aligned}\mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \\ &= \mathbf{C}\mathbf{A}^k\mathbf{x}(0) + \mathbf{C}\sum_{j=0}^{k-1}\mathbf{A}^{k-j-1}\mathbf{B}\mathbf{u}(j) + \mathbf{D}\mathbf{u}(k)\end{aligned}\quad (11.49)$$

11.4.7 Stabilität diskreter Systeme

Aus Gleichung (11.47) lässt sich leicht eine Bedingung für die Stabilität eines diskreten Systems ableiten. In Anlehnung an die Lyapunov-Definition der Stabilität (Definition 7.108) muss sichergestellt werden, dass der Zustandsvektor des Systems für $k \rightarrow \infty$ von einem beliebigen Anfangszustand ohne äußere Anregung gegen Null konvergiert. Die Anwendung dieser Definition auf Gl. (11.47) ergibt (mit $\mathbf{u}(k) = \mathbf{0}$) die Stabilitätsbedingung für ein diskretes System:

$$\lim_{k \rightarrow \infty} \mathbf{x}(k) = \lim_{k \rightarrow \infty} \mathbf{A}^k \mathbf{x}(0) = \mathbf{0} \quad (11.50)$$

Da der Anfangszustand von Null verschieden sein kann, folgt für die Systemmatrix des Systems:

$$\lim_{k \rightarrow \infty} \mathbf{A}^k = \mathbf{0} \quad (11.51)$$

Nun sein angenommen, dass diese Systemmatrix in Jordan-Normalform vorliegt:

$$\mathbf{A} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad (11.52)$$

Für diese Matrix erhält man die Stabilitätsbedingung

$$\lim_{k \rightarrow \infty} \mathbf{A}^k = \lim_{k \rightarrow \infty} \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} = \mathbf{0} \quad (11.53)$$

Offensichtlich ist diese Gleichung erfüllt für

$$|\lambda_i| < 1 \quad (11.54)$$

Satz 11.2 *Ein diskretes System ist stabil im Sinne von Ljapunow, wenn alle Eigenwerte der Systemmatrix einen Absolutwert kleiner als eins haben.*

Lässt man die Gleichheit mit Eins zu,

$$|\lambda_i| \leq 1 \quad (11.55)$$

würde die Folge $\mathbf{x}(k)$ nicht gegen Null konvergieren, aber ihr Absolutwert würde nicht größer werden als der Anfangszustand. Nach der Ljapunov-Stabilitätsdefinition 13.2 für *nichtlineare* Systeme bedeutet dies nur einfache Stabilität, während die Bedingung (11.54) mit Konvergenz gegen Null *asymptotische* Stabilität bedeutet. Da es hier aber nur um lineare Systeme geht, soll nur Satz 11.2 gelten.

Die Überlegungen ändern sich nicht für Systeme in einer anderen Form als der Jordan-Normalform, da sich die Eigenwerte durch eine Systemtransformation nicht ändern. Das bedeutet, dass Satz 11.2 für *alle* diskreten Systeme gilt.

Es ist interessant, die Beziehung zwischen den Stabilitätsbedingungen für diskrete und kontinuierliche Systeme zu betrachten. Aus Satz 7.3 ergibt sich, dass ein kontinuierliches System im Sinne von Lyapunov stabil ist, wenn alle Eigenwerte der Systemmatrix in der linken Hälfte der komplexen Ebene liegen: $\operatorname{Re}\{\lambda_{c,i}\} < 0$. Aus Gl. (11.43) lautet die Beziehung zwischen den Eigenwerten der diskreten und kontinuierlichen Systemmatrizen: $\lambda_i = e^{\lambda_{c,i}T}$. Aus $\operatorname{Re}\{\lambda_{c,i}\} < 0$ folgt $|\lambda_i| < 1$. Obwohl also die Stabilitätsbedingung für diskrete Systeme unabhängig von den Überlegungen zu kontinuierlichen Systemen abgeleitet wurde, spiegeln beide Ergebnisse auch den Zusammenhang aus Gleichung (11.43) wider.

11.4.8 Steuerbarkeit und Beobachtbarkeit

Für diskrete Systeme ist die Idee der Steuerbarkeit und Beobachtbarkeit genau die gleiche wie für kontinuierliche Systeme (siehe Kapitel 7.10.1 und 7.10.2).

Die Definition der Steuerbarkeit lautet:

Definition 11.3 *Ein diskretes System heißt (zustands-)steuerbar, wenn es durch eine geeignete Wahl zeitdiskreter Eingangsvektoren $\mathbf{u}(k)$ in endlicher Zeit von einem beliebigen Anfangszustand $\mathbf{x}(0)$ in einen Endzustand $\mathbf{x}(k_f) = \mathbf{0}$ gebracht werden kann.*

Diese Definition ist der Definition 7.4 für kontinuierliche Systeme sehr ähnlich, aber natürlich muss über eine zeitdiskrete Folge von Eingangsvektoren $\mathbf{u}(k)$ mit $0 \leq k \leq k_f$ und $k_f < \infty$ gesprochen werden. Das gleiche gilt für die Definition der Beobachtbarkeit (siehe Def. 7.8):

Definition 11.4 *Ein diskretes System ist beobachtbar, wenn jeder Anfangszustand $\mathbf{x}(0)$ aus den gemessenen Eingangs- und Ausgangsvektoren $\mathbf{u}(k), \mathbf{y}(k)$ eines endlichen Zeitintervalls $0 \leq k \leq k_f$ und $k_f < \infty$ rekonstruiert werden kann.*

Die Kalman-Kriterien für Steuerbarkeit und Beobachtbarkeit sind genau die gleichen wie für kontinuierliche Systeme (siehe Satz 7.5 und Satz 7.9):

Satz 11.5 *Kalman-Kriterium: Ein diskretes System ist genau dann vollständig steuerbar, wenn die Steuerbarkeitsmatrix*

$$\mathbf{Q}_s = [\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}] \quad (11.56)$$

n linear unabhängige Spaltenvektoren enthält, also den Rang n hat.

Satz 11.6 *Kalman-Kriterium: Ein diskretes System ist genau dann vollständig beobachtbar, wenn die Beobachtbarkeitsmatrix*

$$\mathbf{Q}_b = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \dots \\ \mathbf{CA}^{n-1} \end{bmatrix} \quad (11.57)$$

n linear unabhängige Zeilenvektoren enthält, also den Rang n hat.

Die Beweise für beide Kriterien sind hier allerdings viel einfacher als im kontinuierlichen Fall. Den Beginn soll die Steuerbarkeit machen. Ausgangspunkt ist die allgemeine Lösung der Zustandsgleichung (11.47):

$$\mathbf{x}(k) = \mathbf{A}^k \mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{A}^{k-j-1} \mathbf{B} \mathbf{u}(j) \quad (11.58)$$

Da der Endzustand gemäß der Definition gleich Null sein soll, ist die linke Seite gleich Null. Umstellen liefert

$$-\mathbf{A}^k \mathbf{x}(0) = \sum_{j=0}^{k-1} \mathbf{A}^{k-j-1} \mathbf{B} \mathbf{u}(j) \quad (11.59)$$

Betrachtet man die Summe auf der rechten Seite, so stellt man fest, dass jeder Eingangsvektor $\mathbf{u}(j)$ zu jedem Zeitschritt j mit einer Matrix vom Typ $\mathbf{A}^{k-j-1} \mathbf{B}$ multipliziert wird. Daraus folgt, dass jede Komponente eines beliebigen Eingangsvektors mit einer Spalte dieser Matrizen multipliziert wird. Daher kann die Summe als Linearkombination der Spalten der Matrizen $(\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}, \dots)$ angesehen werden, wobei die Linearkoeffizienten von der Wahl der Eingangssequenz abhängen.

Die Frage der Steuerbarkeit ist nun, ob diese Gleichung für jeden beliebigen Anfangszustand durch eine geeignete Wahl einer Eingangsvektorfolge erfüllt werden kann. Um sicherzustellen, dass eine solche Lösung für jedes $\mathbf{x}(0)$ existiert, muss es n linear unabhängige Vektoren unter den Spalten der Matrizen $(\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \dots)$ geben, die den gesamten n -dimensionalen Raum aufspannen. Das System kann also nur steuerbar sein, wenn die Matrix $(\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}, \dots)$ linear unabhängige Spalten n hat, was gleichbedeutend mit dem Rang n ist.

Da diese Matrix unendlich viele Spalten hat, kann die Bedingung nicht direkt geprüft werden. Das Caley-Hamilton-Theorem (siehe. Kapitel A.4) bietet hier jedoch Abhilfe. Aus diesem Theorem folgt, dass alle höheren Potenzen von \mathbf{A} (mit $k \geq n$) als Linearkombinationen der ersten $(n-1)$ Potenzen von \mathbf{A} dargestellt werden können.

Daraus folgt, dass $\mathbf{A}^n, \mathbf{A}^{n+1}, \mathbf{A}^{n+2}$, usw. keine Spalten haben, die linear unabhängig von den Spalten der ersten $(n-1)$ Potenzen von \mathbf{A} sind. Und deshalb können auch $\mathbf{A}^n \mathbf{B}, \mathbf{A}^{n+1} \mathbf{B}, \mathbf{A}^{n+2} \mathbf{B}$, usw. keine Spalten haben, die linear unabhängig von den Spalten von $\mathbf{A}^k \mathbf{B}$ (mit $k < n$) sind.

Für den Beweis der Steuerbarkeit kann die Suche nach n linear unabhängigen Spalten daher auf die Matrix $(\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B})$ reduziert werden. Damit ist der Beweis des Steuerbarkeitskriteriums erbracht.

Bemerkung: Bisher wurde nur gezeigt, dass der maximale Rang von \mathbf{Q}_s Steuerbarkeit induziert, während der Satz auch die gegenteilige Aussage enthält. Der Beweis dieser Aussage soll hier aber entfallen.

Während der Beweis der Steuerbarkeit für diskrete Systeme sehr eng an den Beweis der Steuerbarkeit für kontinuierliche Systeme anschließt, ist der Beweis für die Beobachtbarkeit anders:

Er beginnt mit der Ausgangsgleichung für $k=0$ (mit $\mathbf{D} = \mathbf{0}$), in die die Zustandsgleichung schrittweise eingesetzt wird:

$$\begin{aligned}
 \mathbf{y}(0) &= \mathbf{C}\mathbf{x}(0) \\
 \mathbf{y}(1) &= \mathbf{C}\mathbf{x}(1) = \mathbf{C}(\mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0)) \\
 &= \mathbf{C}\mathbf{A}\mathbf{x}(0) + \mathbf{C}\mathbf{B}\mathbf{u}(0) \\
 \mathbf{y}(2) &= \mathbf{C}\mathbf{x}(2) = \mathbf{C}(\mathbf{A}\mathbf{x}(1) + \mathbf{B}\mathbf{u}(1)) = \mathbf{C}(\mathbf{A}(\mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0)) + \mathbf{B}\mathbf{u}(1)) \\
 &= \mathbf{C}\mathbf{A}^2\mathbf{x}(0) + \mathbf{C}\mathbf{A}\mathbf{B}\mathbf{u}(0) + \mathbf{C}\mathbf{B}\mathbf{u}(1) \\
 \mathbf{y}(n-1) &= \mathbf{C}\mathbf{x}(n-1) = \mathbf{C}(\mathbf{A}\mathbf{x}(n-2) + \mathbf{B}\mathbf{u}(n-2)) \\
 &= \mathbf{C}\mathbf{A}^{n-1}\mathbf{x}(0) + \mathbf{C}\mathbf{A}^{n-2}\mathbf{B}\mathbf{u}(0) + \dots + \mathbf{C}\mathbf{B}\mathbf{u}(n-2)
 \end{aligned} \tag{11.60}$$

Zusammenfassen all dieser Gleichungen in einer Matrix ergibt

$$\begin{pmatrix} \mathbf{y}(0) \\ \mathbf{y}(1) \\ \dots \\ \mathbf{y}(n-1) \end{pmatrix} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \dots \\ \mathbf{C}\mathbf{A}^{n-1} \end{pmatrix} \mathbf{x}(0) + \begin{pmatrix} \dots \\ \dots \\ \dots \\ \dots \end{pmatrix} \begin{pmatrix} \mathbf{u}(0) \\ \mathbf{u}(1) \\ \dots \\ \mathbf{u}(n-2) \end{pmatrix} \tag{11.61}$$

Nun sei $[0, (n-1)T]$ das Zeitintervall der Beobachtung, und alle Messwerte und Streckenparameter, also alle bekannten Informationen, werden auf die linke Seite der Gleichung gebracht:

$$\begin{pmatrix} \mathbf{y}(0) \\ \mathbf{y}(1) \\ \dots \\ \mathbf{y}(n-1) \end{pmatrix} - \begin{pmatrix} \dots \\ \dots \\ \dots \\ \dots \end{pmatrix} \begin{pmatrix} \mathbf{u}(0) \\ \mathbf{u}(1) \\ \dots \\ \mathbf{u}(n-2) \end{pmatrix} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \dots \\ \mathbf{C}\mathbf{A}^{n-1} \end{pmatrix} \mathbf{x}(0) \tag{11.62}$$

Offensichtlich handelt es sich um ein Gleichungssystem zur Berechnung des unbekannten Anfangszustands $\mathbf{x}(0)$. Auf der linken Seite der Gleichung befindet sich, wenn man alles zusammenfasst, ein Vektor bekannter Werte der Länge nq , wobei n die Anzahl der Zustände und q die Anzahl der Ausgangsvariablen ist. Auf der rechten Seite hat die Koeffizientenmatrix des unbekannten Vektors $\mathbf{x}(0)$ die Dimension $nq \times n$. Somit gibt es nq Gleichungen zur Berechnung der n unbekannten Koeffizienten von $\mathbf{x}(0)$.

Eine eindeutige Lösung existiert nur dann, wenn es unter diesen Gleichungen n linear unabhängige Gleichungen gibt, was bedeutet, dass die Matrix auf der rechten Seite n linear unabhängige Zeilen haben muss. Aber genau das ist es, was das Kalman-Kriterium besagt.

Wenn es zu wenig linear unabhängige Zeilen in dieser Matrix gibt, kann man das Zeitintervall der Beobachtung vergrößern, und die Matrix auf der rechten Seite wird zu

$$\begin{pmatrix} \mathbf{C} \\ \mathbf{CA} \\ \dots \\ \mathbf{CA}^{n-1} \\ \mathbf{CA}^n \\ \mathbf{CA}^{n+1} \\ \dots \end{pmatrix} \quad (11.63)$$

Aber auch hier kann das Caley-Hamilton-Theorem A.4 angewandt werden, um zu beweisen, dass die Matrizen $\mathbf{CA}^n, \mathbf{CA}^{n+1}, \dots$ keine zusätzliche linear unabhängige Zeile zur Matrix beitragen können und es daher sinnlos ist, das Zeitintervall der Beobachtung zu vergrößern.

11.5 Impuls- und Sprungantwort von diskreten Systemen

In diesem Abschnitt geht es um die Impuls- und Sprungantwort von diskreten SISO-Systemen. Den Beginn macht die Definition einer diskreten Impulsfunktion gemäß Abb. 11.14:

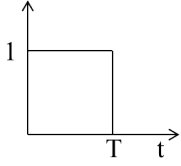
$$\delta_d(k) = \begin{cases} 1 & \text{für } k = 0 \\ 0 & \text{sonst} \end{cases} \quad (11.64)$$

Der Index d zeigt an, dass diese Impulsfunktion nicht mit dem aus dem Kontinuierlichen bekannten Dirac-Impuls zu verwechseln ist.

Die Sprungfunktion entspricht der Definition aus dem Kontinuierlichen:

$$s(k) = \begin{cases} 1 & \text{für } k \geq 0 \\ 0 & \text{sonst} \end{cases} \quad (11.65)$$

Mit Hilfe von Gleichung (11.49) können die Impuls- und Sprungantworten eines diskreten Systems leicht berechnet werden. Mit einem Anfangszustand

**Abb. 11.14.** Diskrete Impulsfunktion

von Null ($\mathbf{x}(0) = \mathbf{0}$), $y(k) = g_d(k)$ und $u(k) = \delta_d(k)$ lautet die Impulsantwort eines SISO-Systems

$$y(k) = \mathbf{c}^T \mathbf{A}^k \mathbf{x}(0) + \mathbf{c}^T \sum_{j=0}^{k-1} \mathbf{A}^{k-j-1} \mathbf{b} u(j) + d u(k)$$

$$g_d(k) = \begin{cases} d & \text{für } k = 0 \\ \mathbf{c}^T \mathbf{A}^{k-1} \mathbf{b} & \text{für } k \geq 1 \end{cases} \quad (11.66)$$

Und die Sprungantwort (mit $y(k) = w(k)$ und $u(k) = s(k)$)

$$w(k) = \begin{cases} d & \text{für } k = 0 \\ \sum_{j=0}^{k-1} \mathbf{c}^T \mathbf{A}^{k-j-1} \mathbf{b} + d & \text{für } k \geq 1 \end{cases}$$

$$= \begin{cases} d & \text{für } k = 0 \\ \sum_{j=0}^{k-1} \mathbf{c}^T \mathbf{A}^j \mathbf{b} + d & \text{für } k \geq 1 \end{cases} \quad (11.67)$$

Einfügen von Impulsfunktion und Impulsantwort in die Differenzengleichung (11.22) liefert

$$g_d(k) = \sum_{i=0}^n b_i \delta_d(k+i-n) - \sum_{i=0}^{n-1} a_i g_d(k+i-n) \quad (11.68)$$

Für $k > n$ verschwindet die erste Summe, da die Impulsfunktion gleich Null ist für alle $k > 0$:

$$g_d(k) = - \sum_{i=0}^{n-1} a_i g_d(k+i-n) \quad (11.69)$$

Wenn die Rückkopplungskoeffizienten a_i nicht bekannt sind, kann diese Gleichung verwendet werden, um sie aus Messungen der Impulsantwort zu berechnen: Man muss die Impulsantwort im Zeitintervall $0 \leq kT \leq 2nT$ messen, wobei n die Ordnung des Systems ist, die vorher bekannt sein muss. Dann kann für jeden Zeitpunkt von $k = n+1$ bis $k = 2n$ eine Gleichung (11.69) aufgestellt werden, indem die entsprechenden Messwerte von g_d eingesetzt werden. Dadurch entsteht ein Gleichungssystem von n Gleichungen für die n unbekannten Koeffizienten a_i , aus dem diese (theoretisch) berechnet werden können.

Bei diesem Ansatz gibt es jedoch zwei Probleme. Das erste ist, dass es vom Verlauf der Impulsantwort abhängt, ob n linear unabhängige Gleichungen entstehen, die benötigt werden, um eine eindeutige Lösung zu erhalten. Wäre z.B. g_d ein konstantes Signal, so erhielte man n mal die gleiche Gleichung, und es gäbe keine Chance für eine eindeutige Lösung. Leider hat man üblicherweise auf diesen Verlauf keinen direkten Einfluss, da die Impulsantwort von der Systemdynamik abhängt.

Die zweite Einschränkung sind Störungen. Wenn es System- oder Messstörungen gibt, erhält man verfälschte Messwerte, und es besteht natürlich keine Möglichkeit, korrekte Rückkopplungskoeffizienten zu berechnen. In Abschnitt 12.2 werden Ansätze zur Überwindung dieser Probleme diskutiert.

Ein weiterer interessanter Punkt ist die Beziehung zwischen Impuls- und Sprungantwort. Zunächst muss man sich klar machen, dass die Sprungfunktion als eine Folge diskreter Impulse dargestellt werden kann, wie in Abb. 11.15 gezeigt. Der erste Impuls ist $\delta_d(k)$ wie in Gl. (11.64) definiert. Der zweite Impuls, der bei $t = T$ beginnt, kann streng nach der Definition durch $\delta_d(k - 1)$ beschrieben werden, und so weiter für die weiteren Impulse. Die Sprungfunktion lässt sich also darstellen als

$$s(k) = \sum_{j=0}^k \delta_d(k - j) \quad (11.70)$$

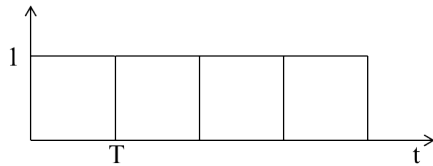


Abb. 11.15. Sprungfunktion als Sequenz von diskreten Impulsen

Umgekehrt kann ein diskreter Impuls als Differenz zweier Sprungfunktionen dargestellt werden, eines positiven Sprungs bei $k = 0$ und eines negativen bei $k = 1$:

$$\delta_d(k) = s(k) - s(k - 1) \quad (11.71)$$

Im kontinuierlichen Bereich besteht die gleiche Beziehung zwischen Dirac-Impuls und Sprungfunktion, nur muss die Summe durch das Integral und die Differenz durch die Ableitung ersetzt werden.

Aufgrund der Linearität des Systems müssen nun die Beziehungen (11.70) und (11.71) auch für die Systemantworten der beiden Signale gelten. Um dies zu zeigen, müssen nur die Gleichungen (11.66) und (11.67) angewendet werden:

$$\begin{aligned}
w(k) &= \begin{cases} d & \text{für } k = 0 \\ \sum_{j=0}^{k-1} \mathbf{c}^T \mathbf{A}^{k-j-1} \mathbf{b} + d & \text{für } k \geq 1 \end{cases} \\
&= \begin{cases} g_d(0) & \text{für } k = 0 \\ \sum_{j=0}^{k-1} g_d(k-j) + g_d(k-k) & \text{für } k \geq 1 \end{cases} \\
&= \begin{cases} g_d(0) & \text{für } k = 0 \\ \sum_{j=0}^k g_d(k-j) & \text{für } k \geq 1 \end{cases} \quad (11.72)
\end{aligned}$$

und

$$\begin{aligned}
g_d(k) &= w(k) - w(k-1) = \sum_{j=0}^{k-1} \mathbf{c}^T \mathbf{A}^j \mathbf{b} + d - \left(\sum_{j=0}^{k-2} \mathbf{c}^T \mathbf{A}^j \mathbf{b} + d \right) \\
&= \mathbf{c}^T \mathbf{A}^{k-1} \mathbf{b} \quad (11.73)
\end{aligned}$$

Das letzte Thema, das hier diskutiert werden soll, ist die *Faltung*. Ausgangspunkt der Betrachtungen ist Gl. (11.49), die den Verlauf des Ausgangssignals in Abhängigkeit vom Anfangszustand und dem Eingangssignal beschreibt. Für ein SISO-System und einen Anfangszustand $\mathbf{x}(0) = \mathbf{0}$ erhält man

$$y(k) = \sum_{j=0}^{k-1} \mathbf{c}^T \mathbf{A}^{k-j-1} \mathbf{b} u(j) + du(k) \quad (11.74)$$

Einsetzen von Gleichung (11.66) für die Impulsantwort liefert:

$$y(k) = \sum_{j=0}^{k-1} g_d(k-j) u(j) + g_d(k-k) u(k) = \sum_{j=0}^k g_d(k-j) u(j) \quad (11.75)$$

Diese Summe hat die gleiche Struktur wie das Faltungsintegral für kontinuierliche Systeme (siehe Abschnitt 3.2):

$$\int_0^t f_1(t-\tau) f_2(\tau) d\tau \quad (11.76)$$

Daher wird diese Summe als *Faltungssumme* bezeichnet. Sie kann wie folgt interpretiert werden: Man muss die Eingangsfolge $u(k)$ als eine Folge von diskreten Impulsen mit der Amplitude $u(k)$ zum Zeitschritt k sehen. Jeder dieser diskreten Impulse verursacht eine Impulsantwort, die mit der Amplitude des jeweiligen Eingang-Impulses multipliziert wird. Um den Beitrag dieser Impulsantwort zum Ausgangswert im Zeitschritt k zu berechnen, muss man sich klar machen, welcher Wert der Impulsantwort, die durch den Impuls $u(j)$ verursacht wurde, zum betrachteten Zeitpunkt gerade am Ausgang erscheint.

Zum Zeitpunkt $t = k$ ist dies $g_d(k - j)$. Der Beitrag des Impulses $u(j)$ zum Ausgangswert im Zeitschritt k ist daher $g_d(k - j)u(j)$. Um schließlich den aktuellen Wert des Ausgangssignals im Zeitschritt k zu erhalten, müssen die Beiträge aller verschiedenen Impulse abschließend aufsummiert werden.

11.6 z-Transformation

11.6.1 Introduction

The z-transformation is the Laplace transformation for discrete systems. Starting point is a sequence of sampled values (not hold!), which is a sequence of Dirac impulses, each of them weighted with the sampled value of the input signal at that time-point (see section 11.1):

$$f^*(t) = \sum_{k=0}^{\infty} f(kT)\delta(t - kT) \quad (11.77)$$

where δ is a Dirac impulse as defined in eq. (3.12), with $s(t)$ being the step function:

$$\delta(t) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (s(t) - s(t - \varepsilon)) \quad (11.78)$$

Obviously, the Dirac impulse has got infinite amplitude for $t = 0$ and 0 else. The integral of that function can be easily calculated, it is 1. The Laplace transform of a Dirac impulse was calculated in section 3.2, it is also equal 1. Following the definition, $\delta(t - kT)$ is a Dirac impulse at time-point $t = kT$, so that the sum in eq. (11.77) describes a sequence of Dirac impulses.

For the Laplace transform of $y^*(t)$ we get

$$\begin{aligned} f^*(s) = \mathcal{L}\{f^*(t)\} &= \mathcal{L}\left\{\sum_{k=0}^{\infty} f(kT)\delta(t - kT)\right\} \\ &= \sum_{k=0}^{\infty} f(kT)\mathcal{L}\{\delta(t - kT)\} = \sum_{k=0}^{\infty} f(kT)e^{-skT} \\ &= \sum_{k=0}^{\infty} f(kT)(e^{sT})^{-k} \end{aligned} \quad (11.79)$$

Now we define:

$$\begin{aligned} z &:= e^{sT} \\ f(z) &:= f^*(s) \\ f(k) &:= f(kT) \end{aligned} \quad (11.80)$$

and get

$$f(z) = \sum_{k=0}^{\infty} f(k)z^{-k} \quad (11.81)$$

We can define $f(z)$ as z-transform of the sequence $f(k)$. Alternatively, we can see $f(z)$ as Laplace transform of $f^*(t)$. Both views are commonly used.

As a practice, we calculate the z-transforms of some specific signals. The first signal shall be the discrete impulse $\delta_d(k)$ as defined in (11.64):

$$f(z) = \sum_{k=0}^{\infty} f(k)z^{-k} = 1 \cdot z^{-0} = 1 \quad (11.82)$$

Also relatively simple is the step function with $s(t) = 1$ for $t \geq 0$ or $f(k) = 1$ for $k \geq 0$. Following eq. (11.81), we get

$$f(z) = \sum_{k=0}^{\infty} f(k)z^{-k} = \sum_{k=0}^{\infty} 1z^{-k} = \sum_{k=0}^{\infty} z^{-k} \quad (11.83)$$

For $|z| > 1$ we can calculate the analytical result of this infinite sum:

$$f(z) = \sum_{k=0}^{\infty} z^{-k} = \frac{1}{1 - \frac{1}{z}} = \frac{z}{z - 1} \quad (11.84)$$

The next shall be the exponential function $f(t) = e^{\lambda t}$ resp. $f(k) = e^{\lambda kT}$:

$$f(z) = \sum_{k=0}^{\infty} f(k)z^{-k} = \sum_{k=0}^{\infty} e^{\lambda kT} z^{-k} = \sum_{k=0}^{\infty} \left(\frac{z}{e^{\lambda T}}\right)^{-k} \quad (11.85)$$

Here, for $|\frac{z}{e^{\lambda T}}| > 1$ resp. $|z| > e^{\lambda T}$, we can again calculate the analytical result of this sum:

$$f(z) = \sum_{k=0}^{\infty} \left(\frac{z}{e^{\lambda T}}\right)^{-k} = \frac{1}{1 - \frac{e^{\lambda T}}{z}} = \frac{z}{z - e^{\lambda T}} \quad (11.86)$$

From (11.84) and (11.86) we can see that we can replace the infinite sum by an analytical expression, if the complex values of z are outside of a circle around the origin, in the first case the unit circle, and in the second case a circle with radius $e^{\lambda T}$ (see fig. 11.16).

Inside of that circle, the infinite sum does not converge, so that the analytical expression is not a valid result for $f(z)$. But on the other hand, the analytical expression itself works also inside of the circle, as it is defined in the entire complex plane. Besides that, the handling of an analytical expression is much easier than the handling of an infinite sum. Therefore, we continue the analytical solution for $f(z)$, which is originally only defined for values of z outside of the circle, to the area inside of the circle. That means that from

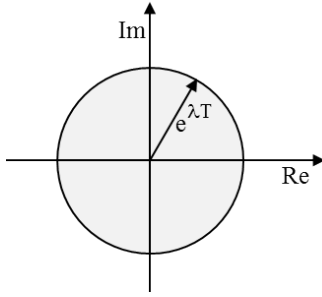


Abb. 11.16. Convergence Radius of the z-transformation

now on, we use the results of eqs. (11.84) and (11.86) without any restriction regarding z .

The final signal to be z-transformed shall be the Dirac impulse. We can either calculate the z-transform directly from eq. (11.81), or we can first create the sampled signal $f^*(t)$ and then Laplace-transform it. Here, we will do the second. Following the block diagram in fig. 11.4, sampling means the multiplication of the input signal with a sequence of Dirac impulses. As a result, we will get here $(\delta(t))^2$ at time-point $t = 0$ and nothing else. The product of two Dirac impulses will have infinite amplitude, as $(\infty)^2 = \infty$ and an integral value still equal one. It follows that $(\delta(t))^2 = \delta(t)$, and as $\mathcal{L}\{\delta(t)\} = 1$, we get $f(z) = f^*(s) = 1$. The z-transform and the Laplace transform of a Dirac impulse therefore have the same value, and also the z-transforms of a discrete impulse and a Dirac impulse (see eq. (11.82)).

11.6.2 Theorems about z-Transformation

For the z-transformation we have theorems very similar to the theorems of the Laplace transformation (see section A.1) in the appendix, but here, we can easily prove them:

Superposition Principle.

$$\begin{aligned}
 f(k) &:= a_1 f_1(k) + a_2 f_2(k) \\
 f(z) &= \sum_{k=0}^{\infty} f(k) z^{-k} = \sum_{k=0}^{\infty} (a_1 f_1(k) + a_2 f_2(k)) z^{-k} \\
 &= a_1 \sum_{k=0}^{\infty} f_1(k) z^{-k} + a_2 \sum_{k=0}^{\infty} f_2(k) z^{-k} \\
 &= a_1 f_1(z) + a_2 f_2(z)
 \end{aligned} \tag{11.87}$$

Delay Theorem.

$$h(k) := f(k - m) \text{ where } f(i) = 0 \text{ for } i < 0$$

$$h(z) = \sum_{k=0}^{\infty} h(k)z^{-k} = \sum_{k=0}^{\infty} f(k-m)z^{-k} \quad (11.88)$$

With $i := k - m$ and $-k = -i - m$ we get

$$\begin{aligned} h(z) &= \sum_{i=-m}^{\infty} f(i)z^{-i-m} = \sum_{i=0}^{\infty} f(i)z^{-i-m} \\ &= z^{-m} \sum_{i=0}^{\infty} f(i)z^{-i} = z^{-m} f(z) \end{aligned} \quad (11.89)$$

Damping Theorem.

$$\begin{aligned} h(k) &:= a^{-k} f(k) \\ h(z) &= \sum_{k=0}^{\infty} h(k)z^{-k} = \sum_{k=0}^{\infty} a^{-k} f(k)z^{-k} \\ &= \sum_{k=0}^{\infty} f(k)(za)^{-k} = f(az) \end{aligned} \quad (11.90)$$

Difference Theorem.

$$\begin{aligned} h(k) &:= f(k) - f(k-1) \\ h(z) &= \sum_{k=0}^{\infty} (f(k) - f(k-1))z^{-k} \\ &= \sum_{k=0}^{\infty} f(k)z^{-k} - \sum_{k=0}^{\infty} f(k-1)z^{-k} \end{aligned} \quad (11.91)$$

With the application of the Delay Theorem we get

$$h(z) = f(z) - \frac{1}{z}f(z) = \frac{z-1}{z}f(z) \quad (11.92)$$

Sum Theorem.

$$h(k) := \sum_{j=0}^k f(j) \quad (11.93)$$

Obviously, it follows

$$h(k) - h(k-1) = f(k) \quad (11.94)$$

The z-transformation of this equation, applying the Difference Theorem, gives

$$\frac{z-1}{z}h(z) = f(z) \quad (11.95)$$

and therefore

$$h(z) = \frac{z}{z-1} f(z) \quad (11.96)$$

From the proof it is clear that the sum operator $\frac{z}{z-1}$ must be the inverse of the difference operator $\frac{z-1}{z}$. This relationship is the same as the relationship between the derivative and integral operator for Laplace transformation, which is s and $\frac{1}{s}$. The same similarity between z -transformation and Laplace transformation holds regarding the equality of the transform of the step function $\frac{z}{z-1}$ in the discrete domain with the sum operator resp. the equality of the step function $\frac{1}{s}$ in the continuous domain with the integration operator.

Convolution Theorem.

$$\begin{aligned} h(k) &:= \sum_{j=0}^k f_1(k-j) f_2(j) \\ h(z) &= \sum_{k=0}^{\infty} \left(\sum_{j=0}^k f_1(k-j) f_2(j) \right) z^{-k} \end{aligned} \quad (11.97)$$

After thinking a while it gets clear that this infinite double sum contains all possible products of $f_1(k)z^{-k}$ with $f_2(k)z^{-k}$, so that it can be written as follows:

$$h(z) = \left(\sum_{k=0}^{\infty} f_1(k) z^{-k} \right) \left(\sum_{k=0}^{\infty} f_2(k) z^{-k} \right) = f_1(z) f_2(z) \quad (11.98)$$

Limit Theorem I.

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} f(k) z^{-k} = f(0) + f(1) \frac{1}{z} + f(2) \frac{1}{z^2} + f(3) \frac{1}{z^3} + \dots \\ \lim_{|z| \rightarrow \infty} f(z) &= f(0) \end{aligned} \quad (11.99)$$

Limit Theorem II. From the Sum Theorem it follows from

$$h(k) := \sum_{j=0}^k f(j) \quad (11.100)$$

for the z -transforms

$$h(z) = \frac{z}{z-1} f(z) \quad (11.101)$$

resp.

$$f(z) = \frac{z-1}{z} h(z) \quad (11.102)$$

Calculating the limes for eq. (11.100), we get

$$\begin{aligned}
\lim_{k \rightarrow \infty} h(k) &= \lim_{k \rightarrow \infty} \sum_{j=0}^k f(j) = \sum_{j=0}^{\infty} f(j) = \sum_{j=0}^{\infty} f(j) 1^{-j} \\
&= \lim_{z \rightarrow 1} \sum_{j=0}^{\infty} f(j) z^{-j} = \lim_{z \rightarrow 1} f(z) = \lim_{z \rightarrow 1} \frac{z-1}{z} h(z) \\
&= \lim_{z \rightarrow 1} (z-1) h(z)
\end{aligned} \tag{11.103}$$

11.6.3 Discrete Transfer Function

Starting point is eq. (11.75)

$$y(k) = \sum_{j=0}^k g_d(k-j)u(j) \tag{11.104}$$

where y is the output signal of the system, u the input, and g_d the discrete impulse response. Applying the Convolution Theorem, we get

$$y(z) = G(z)u(z) \tag{11.105}$$

so that we can see $G(z)$, which is the z-transform of the discrete impulse response, as the discrete transfer function of the system. We can calculate it from the block diagram in fig. 11.11:

$$\begin{aligned}
y(z) &= b_n u(z) + b_{n-1} \frac{1}{z} u(z) + \dots + b_1 \frac{1}{z^{n-1}} u(z) + b_0 \frac{1}{z^n} u(z) \\
&\quad - a_{n-1} \frac{1}{z} y(z) - \dots - a_1 \frac{1}{z^{n-1}} y(z) - a_0 \frac{1}{z^n} y(z) \\
\frac{y(z)}{u(z)} &= G(z) = \frac{b_n z^n + b_{n-1} z^{n-1} + \dots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0}
\end{aligned} \tag{11.106}$$

and of course we see that the coefficients of the transfer function are exactly the same as the coefficients of the difference equation.

For MIMO systems, we can calculate the transfer function resp. the transfer matrix from the discrete state space equation:

Transforming the state space equations (11.25)

$$\begin{aligned}
\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\
\mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)
\end{aligned} \tag{11.107}$$

into the frequency domain using the z-transformation, we get

$$\begin{aligned}
\mathbf{x}(z) &= \frac{1}{z} \mathbf{A}\mathbf{x}(z) + \frac{1}{z} \mathbf{B}\mathbf{u}(z) \\
\mathbf{y}(z) &= \mathbf{C}\mathbf{x}(z) + \mathbf{D}\mathbf{u}(z)
\end{aligned} \tag{11.108}$$

Rearranging the first equation, we get

$$\begin{aligned} z\mathbf{x}(z) &= \mathbf{A}\mathbf{x}(z) + \mathbf{B}\mathbf{u}(z) \\ z\mathbf{x}(z) - \mathbf{A}\mathbf{x}(z) &= \mathbf{B}\mathbf{u}(z) \\ \mathbf{x}(z) &= (z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(z) \end{aligned} \quad (11.109)$$

and inserting this result into the system output equation gives

$$\mathbf{y}(z) = [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}] \mathbf{u}(z) \quad (11.110)$$

Obviously,

$$G(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (11.111)$$

must be the transfer matrix of the system. You can see that the formula is practically the same as for continuous systems (see eq. (7.107)), just s has been replaced by z .

Easily we can calculate the *static gain* of the system, which is the gain in the steady state. In the continuous frequency domain, the steady state is defined by $s = 0$, and therefore, here we define it by $z = e^{sT} = e^0 = 1$. For the static gain, it follows

$$V = G(z = 1) \quad (11.112)$$

11.6.4 Discrete Eigenvalues and Shannon Theorem

We know from section 11.4.4 and eq. (11.43) the relationship between continuous and discrete eigenvalues:

$$\lambda_i = e^{\lambda_{c,i}T} \quad (11.113)$$

where λ_i is an eigenvalue of the discrete system while $\lambda_{c,i}$ is the corresponding eigenvalue of the continuous system. From section 11.4.7 and theorem 11.2 we know the stability condition for discrete systems: The absolute value of each discrete eigenvalue must be smaller than one: $|\lambda_i| < 1$

Unfortunately, the exponential function in eq. (11.113) is not bijective for complex values, so that two different continuous eigenvalues might have the same corresponding discrete eigenvalue. For example, let two different continuous eigenvalues be

$$\begin{aligned} \lambda_{c,1} &= a + jb \\ \lambda_{c,2} &= a + j(b + \frac{2\pi}{T}) \end{aligned} \quad (11.114)$$

According to eq. (11.113), the corresponding discrete eigenvalues are

$$\begin{aligned} \lambda_1 &= e^{\lambda_{c,1}T} = e^{aT} e^{jbT} \\ \lambda_2 &= e^{\lambda_{c,2}T} = e^{aT} e^{j(b + \frac{2\pi}{T})T} = e^{aT} e^{j(bT + 2\pi)} \\ &= e^{aT} e^{jbT} e^{j2\pi} = e^{aT} e^{jbT} = \lambda_1 \end{aligned} \quad (11.115)$$

Fig. 11.17 illustrates the relations. Let's consider an area in the negative half-plane of the continuous s-domain, which is bounded by the lines 1-4, which are defined as follows:

- Line 1: $s = a - j\frac{\pi}{T}$ with $-\infty \leq a \leq 0$
- Line 2: $s = jb$ with $-\frac{\pi}{T} \leq b \leq 0$
- Line 3: $s = jb$ with $0 \leq b \leq \frac{\pi}{T}$
- Line 4: $s = a + j\frac{\pi}{T}$ with $0 \leq a \leq -\infty$

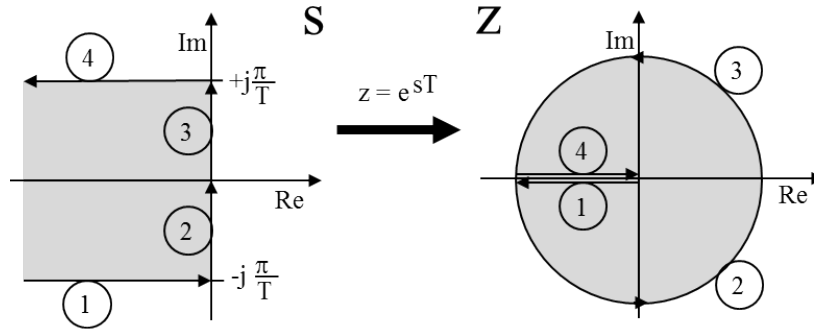


Abb. 11.17. About the non-bijectivity of the mapping $z = e^{sT}$

Now, we want to find the mapping of this area in the discrete z-domain using $z = e^{sT}$. For that, we map each of the four lines into the z-domain. The mappings are:

- Line 1: $z = e^{(a-j\frac{\pi}{T})T} = e^{aT}e^{-j\pi}$ with $-\infty \leq a \leq 0$. The starting point of this line is $z = e^{-\infty}e^{-j\pi}$, which is zero. For increasing a , the absolute value of z gets larger, while the angle is always $-\pi$. When $a = 0$, the absolute value is equal one. The resulting mapping of line 1 is shown in the right half of fig. 11.17.
- Line 2: $z = e^{jbT}$ with $-\frac{\pi}{T} \leq b \leq 0$. The absolute value of this mapping is always equal one, while the angle goes from $-\pi$ up to zero. The result is a half-circle as shown in fig. 11.17.
- Line 3: $z = e^{jbT}$ with $0 \leq b \leq \frac{\pi}{T}$. Again, we get a half-circle, where the angle now goes from zero to π .
- Line 4: $z = e^{(a+j\frac{\pi}{T})T} = e^{aT}e^{j\pi}$ with $0 \leq a \leq -\infty$. Compared to the mapping of line 1, now the absolute value is decreasing from one to zero, while the angle is always π (see fig. 11.17).

Following the rules of analysis, as the area in the s-domain is surrounded by the four lines, the mapping of the area in the z-domain must be surrounded by the mappings of the four lines, so that we get the unit circle in the z-domain. Obviously, all poles, which lie in the surrounded area in the s-domain, must lie inside of the unit circle.

Now we consider another area in the s-domain as shown in fig. 11.18, which is directly neighboring to the first one of fig. 11.17, and after the same steps as before, we get the same mapping in the z-domain, which is the unit circle. Further areas in the s-domain would obviously all have the same mapping in the z-domain. This shows clearly again the non-bijectivity of the mapping $z = e^{sT}$.

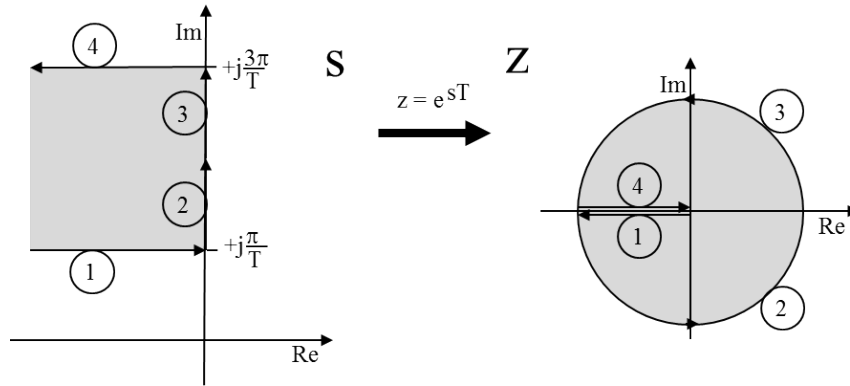


Abb. 11.18. About the non-bijectivity of the mapping $z = e^{sT}$, part II

It is obvious that the two different continuous eigenvalues of eq. (11.114) lie in two different areas, but have the same mappings inside the unit circle in the z-domain. This is a consequence of the Shannon Theorem 11.1. We can see that the upper boundary of the basic area in fig. 11.17 is at

$$j\frac{\pi}{T} = j\pi f_T = j\frac{1}{2}2\pi f_T = j\frac{1}{2}\omega_T \quad (11.116)$$

When we leave the area and cross this line, which represents the half of the circular sampling frequency, we get ambiguity. But this is exactly what the Shannon Theorem says: Different oscillations look the same after sampling.

11.6.5 Damping of Discrete Systems

As we know now that the left half-plane of the s-domain is mapped into the unit circle, the question arises, where to place the eigenvalues of a discrete system exactly to guarantee not only stability, but also a sufficient damping. For this purpose, we will discuss two different approaches. The first one is easier, but the second is more precise.

For both approaches, we should first explain the damping again (see section 2.5). Each complex-conjugated pair of poles causes an oscillation in the step response of a linear system, where several pairs of poles cause several oscillations which have to be added to each other. Each pair of poles forms a second-order lag as a partial transfer function

$$G(s) = \frac{1}{\frac{1}{\omega_0^2}s^2 + \frac{2D}{\omega_0}s + 1} \quad (11.117)$$

where D is the damping. We assume that the poles lie in the left half of the complex plane, which means that the system is stable. The larger D is, the faster the oscillation converges to zero. For $D = 0$ there is no damping at all and the oscillation does not converge to zero, and for $D \geq 1$ there will be no oscillations in the step response. $D < 0$ is not possible in real plants.

We know that the damping of a continuous system gets worse, the closer a complex-conjugated pair of poles, being in the left half of the complex plane, gets to the imaginary axis. Therefore, a simple approach to guarantee good damping is to keep a minimum distance d between the poles and the imaginary axis, as shown in the left half of fig. 11.19.

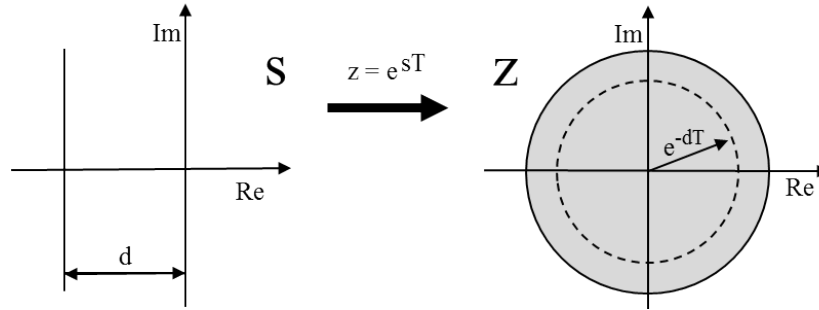


Abb. 11.19. Damping of a discrete system

From the condition $\text{Re}\{s_\lambda\} < -d$ for the continuous poles it follows for the discrete poles in the z-domain:

$$\begin{aligned} z_\lambda &= e^{s_\lambda T} = e^{\text{Re}\{s_\lambda\}T} e^{j\text{Im}\{s_\lambda\}T} \\ |z_\lambda| &= e^{\text{Re}\{s_\lambda\}T} < e^{-dT} \end{aligned} \quad (11.118)$$

This means that to guarantee the same damping for a discrete system, all poles must lie in a circle with the radius e^{-dT} .

For the second approach, we refer to eq. (2.48), which says that $D = \cos \alpha$, where α is the angle as shown in the left half of fig. 11.20.

All pairs of poles, which are right of the dotted lines, have a lower damping. Therefore, we would like to define the area left of the dotted lines, where all poles must lie inside to guarantee a minimum damping $D > \cos \alpha$. Now the task is to define the dotted lines in the s-domain and to map them into the z-domain. With a being the curve parameter going from zero to infinity the dotted lines can be defined by

$$\begin{aligned} s &= -a \cos \alpha \pm ja \sin \alpha = -a \cos \alpha \pm ja \sqrt{1 - \cos^2 \alpha} \\ &= -aD \pm ja \sqrt{1 - D^2} \end{aligned} \quad (11.119)$$

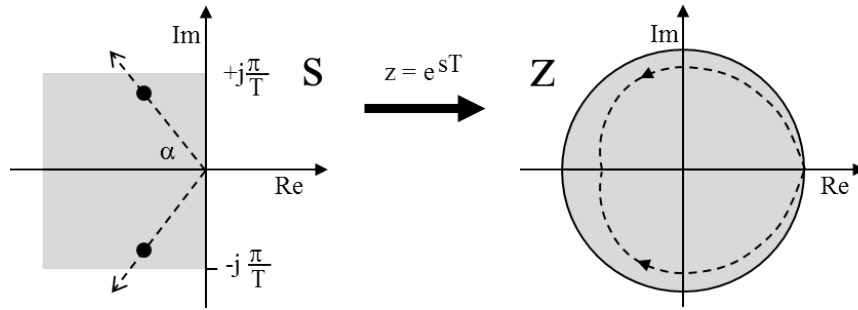


Abb. 11.20. Damping of a discrete system

The mapping of the dotted lines in the z -domain is

$$z = e^{Ts} = e^{-aTD} e^{\pm jTa\sqrt{1-D^2}} \quad (11.120)$$

For $a = 0$ we get $z = 1$ as starting point of the curves. With increasing a we get a decreasing absolute value and increasing angle for the positive case and decreasing angle for the negative case. The result would be two spiral curves, both starting at $z = 1$, one with increasing positive angle and the other one with negative increasing angle. But we have to keep in mind the limit given by the Shannon Theorem. This means that in the s -domain we can stop the curve, when the imaginary part reaches $j\frac{\pi}{T}$. For the z -domain this means that the spiral curves stop, when the angles reach $\pm\pi$. The result is shown in the right half of fig. 11.20.

The area, where the poles must be to have guaranteed minimum damping is bounded by the two curves, which means, they must be inside of the area which is bounded by the dotted lines.

11.6.6 Discrete and Continuous Transfer Functions

For the state-space description, we already have developed the formulas in section 11.4.3, how to calculate the discrete state-space matrices from the continuous ones. Now, we will do the same for transfer functions.

Let's assume a continuous transfer function

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (11.121)$$

with $m < n$, which is the normal case in physical systems. For simplicity we assume that this transfer function has got only poles s_λ , which are different from each other. In that case we can write

$$G(s) = \sum_{\lambda=1}^n \frac{R_\lambda}{s - s_\lambda} \quad (11.122)$$

where R_λ are the residuals of the partial fractions. As the impulse response $g(t)$ is the corresponding function of $G(s)$ in the time-domain, we get

$$g(t) = \sum_{\lambda=1}^n R_\lambda e^{s_\lambda t} \quad (11.123)$$

and for the sampled impulse response the sequence

$$g(kT) = \sum_{\lambda=1}^n R_\lambda e^{s_\lambda kT} \quad (11.124)$$

resp. in shorter form

$$g(k) = \sum_{\lambda=1}^n R_\lambda e^{s_\lambda kT} \quad (11.125)$$

Using eq. (11.81) we get for the z-transform of this signal

$$\begin{aligned} G(z) &= \sum_{k=0}^{\infty} \left(\sum_{\lambda=1}^n R_\lambda e^{s_\lambda kT} \right) z^{-k} \\ &= \sum_{\lambda=1}^n R_\lambda \left(\sum_{k=0}^{\infty} \left(\frac{z}{z_\lambda} \right)^{-k} \right) \quad \text{with} \quad z_\lambda = e^{s_\lambda T} \quad \left| \frac{z}{z_\lambda} \right| < 1 \\ &= \sum_{\lambda=1}^n R_\lambda \frac{1}{1 - \frac{z_\lambda}{z}} = \sum_{\lambda=1}^n R_\lambda \frac{z}{z - z_\lambda} \end{aligned} \quad (11.126)$$

Unfortunately, this result is not complete. If we look at fig. 11.5, we can see the (theoretical) controller function block $K(s)$, followed by the Sample + Hold blocks. These three blocks form the discrete controller, which gets a step function control error and delivers a step function controller output. The same holds for the transfer function. $G(s)$ is representing the physical, continuous plant. If we add Sample + Hold block, we get the discrete transfer function $G(z)$. Transferring this to our calculation so far, we have to recognize that in our calculation the Sample block appears, but not the Hold. Therefore, we must add the Hold block to our calculation.

The transfer function for the Hold block we get from fig. 11.4:

$$H(s) = \frac{1}{s} (1 - e^{-Ts}) \quad (11.127)$$

Now, we first add only the integrator to the plant transfer function and do exactly the same steps as before:

$$G(s) \frac{1}{s} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s(s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0)}$$

$$\begin{aligned}
&= \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s(s-s_1)(s-s_2)\dots(s-s_n)} \\
&= \frac{R_0^*}{s} + \frac{R_1^*}{s-s_1} + \frac{R_2^*}{s-s_2} + \dots \\
&= \frac{R_0^*}{s} + \sum_{\lambda=1}^n \frac{R_\lambda^*}{s-s_\lambda}
\end{aligned} \tag{11.128}$$

As we have added one pole to the function, the values of the residuals have changed, so that we call them R^* here. To calculate these residuals, we apply a well-known trick, as we just have to delete the corresponding pole factor from the denominator of the transfer function and then insert the value of the pole into the rest of the transfer function. The resulting value is the residual value. Following this scheme, we get $R_\lambda^* = \frac{R_\lambda}{s_\lambda}$ for $1 \leq \lambda \leq n$, and for the added pole $s = 0$ we get

$$R_0^* = G(0) = \left(\sum_{\lambda=1}^n \frac{R_\lambda}{-s_\lambda} \right) \tag{11.129}$$

It follows

$$G(s) \frac{1}{s} = \sum_{\lambda=1}^n \frac{R_\lambda}{-s_\lambda} \frac{1}{s} + \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \frac{1}{s-s_\lambda} \tag{11.130}$$

This result is now transformed into the z-domain, where we can apply the result of eq. (11.126):

$$(G(s) \frac{1}{s})(z) = \sum_{\lambda=1}^n \frac{R_\lambda}{-s_\lambda} \frac{z}{z-1} + \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \frac{z}{z-z_\lambda} \tag{11.131}$$

As a final step, already in the z-domain, we now add the other part of the Hold block $(1 - e^{-Ts})$, but directly in z-description $(1 - \frac{1}{z})$, to get the complete z-transfer function:

$$\begin{aligned}
(HG)(z) &= \left[\sum_{\lambda=1}^n \frac{R_\lambda}{-s_\lambda} \frac{z}{z-1} + \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \frac{z}{z-z_\lambda} \right] \left(1 - \frac{1}{z}\right) \\
&= \sum_{\lambda=1}^n \frac{R_\lambda}{-s_\lambda} + \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \frac{z-1}{z-z_\lambda} = \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \left(-1 + \frac{z-1}{z-z_\lambda}\right) \\
&= \sum_{\lambda=1}^n \frac{R_\lambda}{s_\lambda} \frac{z_\lambda - 1}{z-z_\lambda}
\end{aligned} \tag{11.132}$$

One might ask now, why we had to include the *Hold* block explicitly here, while we didn't have to do that, when we developed the relationships between the continuous and discrete state space matrices in section 11.4.3. The answer is that in the derivation of that section we already started our

considerations with the step function, while here we started with the impulse sequence. Therefore, the *Hold* block was included there from the beginning, so that we didn't have to worry about it at the end.

A very simple example shall demonstrate the benefit of this result. Let us apply it to a first-order lag:

$$G(s) = \frac{V}{T_1 s + 1} = \frac{V}{T_1} \frac{1}{s + \frac{1}{T_1}} \quad (11.133)$$

We have $R_1 = \frac{V}{T_1}$, $s_1 = -\frac{1}{T_1}$, $z_1 = e^{Ts_1} = e^{-\frac{T}{T_1}}$ and get

$$(HG)(z) = \frac{R_1}{s_1} \frac{z_1 - 1}{z - z_1} = -V \frac{z_1 - 1}{z - z_1} = V \frac{1 - z_1}{z - z_1} \quad (11.134)$$

To check this result, we try to do the same via the state-space, using the formulas derived in section 11.4.3. To do so, we first need to describe the first-order lag in state-space description:

$$\begin{aligned} \frac{y(s)}{u(s)} &= G(s) = \frac{V}{T_1 s + 1} \\ T_1 s y(s) + y(s) &= V u(s) \\ T_1 \frac{dy(t)}{dt} + y(t) &= V u(t) \end{aligned} \quad (11.135)$$

Now, we define $x(t) := y(t)$ and get the state-space description

$$\begin{aligned} \dot{x}(t) &= -\frac{1}{T_1} x(t) + \frac{V}{T_1} u(t) \\ y(t) &= x(t) \end{aligned} \quad (11.136)$$

with $\mathbf{A}_c = -\frac{1}{T_1}$, $\mathbf{B}_c = \frac{V}{T_1}$, $\mathbf{C}_c = 1$, $\mathbf{D}_c = 0$. From eq. (11.40) we get

$$\begin{aligned} \mathbf{A} &= e^{\mathbf{A}_c T} = e^{-\frac{T}{T_1}} = z_1 \\ \mathbf{B} &= \int_0^T e^{\mathbf{A}_c \alpha} d\alpha \mathbf{B}_c = \int_0^T e^{-\frac{\alpha}{T_1}} d\alpha \frac{V}{T_1} \\ &= -T_1 e^{-\frac{\alpha}{T_1}} \Big|_0^T \frac{V}{T_1} = -T_1 (e^{-\frac{T}{T_1}} - 1) \frac{V}{T_1} = V(1 - z_1) \\ \mathbf{C} &= \mathbf{C}_c = 1 \\ \mathbf{D} &= \mathbf{D}_c = 0 \end{aligned} \quad (11.137)$$

With these results, we can calculate the z-transfer function from eq. (11.111):

$$\begin{aligned}
G(z) &= \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \\
&= (z - z_1)^{-1}V(1 - z_1) = V\frac{1 - z_1}{z - z_1}
\end{aligned}
\tag{11.138}$$

which is the same result as in eq. (11.134).

11.6.7 Discrete Controller Design

In the most cases the sampling time T is very small compared to the smallest time constant of the plant. For this case, the effects of discretization can be neglected during the controller design procedure, and we can calculate the controller parameters with any classical continuous controller design algorithm. Just at the end, we have to implement the quasi-continuous controller in a microprocessor, i.e. in software code.

For a state-space controller, this is trivial, as the state-feedback matrix \mathbf{F} represents just multiplications with constant coefficients and summations. More interesting is a PID-controller, and therefore a typical source code for a PID-controller shall be given.

The following function *PIDcontroller* is assumed to be called at every time-step with the input value *error* and the controller output value *output*. The global variables *integrator* and *OldError* shall keep their last calculated value independently from the time-step. The controller parameters P , I , D , and *limit* have to be chosen adequately during controller design. The source code is just a pseudo code, not following any real programming language syntax. The code also includes an anti-windup mechanism, which stops integration, if the controller output runs out of bounds.

```

integrator := 0
OldError := 0
T := 0.01      (Sampling Time in seconds)
I := 0.1
P := 1.0
D := 0.5
limit := 10.0
function PIDcontroller(error; output)
BEGIN
    OldIntegrator := integrator
    integrator := integrator + T * error
    output := P * error + I * integrator + D/T * (error - OldError)
    OldError := error

```

```

IF output > limit THEN BEGIN
    output := limit
    integrator := OldIntegrator
END
IF output < -limit THEN BEGIN
    output := -limit
    integrator := OldIntegrator
END
END (function PIDcontroller)

```

(11.139)

It should be mentioned that although the sampling time can be neglected during the design of the controller parameters, it cannot be neglected when writing the program code. One can see that the sampling time T has to be considered in the calculation of the integration as well as in the calculation of the difference quotient.

If the sampling time cannot be neglected, we have to do a discrete controller design, and for that purpose we have two different possibilities. The first one is only for the SISO case: We transform the continuous transfer function $G(s)$ into the z -domain with the help of eq. (11.132) to get $G(z)$ (resp. $(HG)(z)$). Then we have to look at the closed-loop transfer function

$$T(z) = \frac{G(z)K(z)}{1 + G(z)K(z)} \quad (11.140)$$

and choose the parameters of $K(z)$, so that all poles of the closed-loop system lie inside of the unit circle resp. inside of the areas given by the damping considerations of section 11.6.5.

The second discrete controller design approach is in the state space, which is obviously possible for MIMO and SISO systems. First, we must describe the plant dynamics in state-space form (see eq. (11.25)), where the discrete system matrices can be calculated from the continuous ones using eq. (11.40):

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \quad (11.141)$$

Then, we have to design the controller feedback matrix: $\mathbf{u}(k) = \mathbf{F}\mathbf{x}(k)$. With this feedback, the closed-loop system matrix is $(\mathbf{A} + \mathbf{B}\mathbf{F})$ as in the continuous state-space, and therefore, the controller design algorithms are practically the same here as in the continuous domain.

For the *Pole Placement*, the Ackermann formula (8.64) can be used directly, but of course we must place the poles of the closed-loop system now inside of the unit circle, and not in the left half-plane:

$$\mathbf{f}_x^T = -\mathbf{t}_1^T(p_0 + p_1\mathbf{A} + \dots + p_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n) \quad (11.142)$$

For the *Riccati Controller Design* we give the discrete version of theorem 8.8:

Satz 11.7 *Riccati Controller: Given is a system in state-space representation eq. (11.141) and a quality function*

$$J = \sum_{k=0}^{\infty} (\mathbf{x}^T(k) \mathbf{Q} \mathbf{x}(k) + \mathbf{u}^T(k) \mathbf{R} \mathbf{u}(k)) \quad (11.143)$$

where the matrices \mathbf{R} and \mathbf{Q} have to be chosen symmetrical and positive definite. Then J is minimal for a controller of the form

$$\mathbf{u}(k) = \mathbf{F} \mathbf{x}(k) \quad (11.144)$$

with

$$\mathbf{F} = -(\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{P} \mathbf{A} \quad (11.145)$$

where \mathbf{P} is the symmetrical, positive definite solution of the discrete matrix Riccati equation

$$\mathbf{A}^T \mathbf{P} \mathbf{A} - \mathbf{P} - \mathbf{A}^T \mathbf{P} \mathbf{B} (\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{P} \mathbf{A} = -\mathbf{Q} \quad (11.146)$$

12. Adaptive Control

12.1 Motivation and Overview

Many plants change their dynamics after a while because of aging: Motors get increased friction, pipes get dirty or rusty, which increases their flow resistance, etc. Other plants might change their dynamics regularly, for example aeroplanes that have a full tank at the beginning of the flight, and an empty tank at the end of the flight. In all these cases it is helpful to have an adaptive controller that changes its parameters regarding to the changes of the plant.

Another field of application are plants with unknown parameters. While the parameters of electro-mechanical systems are well known and possible to calculate with reasonable accuracy, the parameters of chemical or biological plants are rather unknown. But if the model of the plant is not accurate, the controller cannot be accurate, too. The general structure (=order) of the controller can be estimated, but the precise parameters not. In such a case it can be also helpful if the controller parameters are adapted automatically after starting.

We can distinguish between three general approaches for adaptive control. The first approach can be called *Indirect Adaptation* (see fig. 12.1). This approach is a two-stage approach. In the first stage, based on the plant input and output signals, the model of the plant is calculated in the *plant identification block*. How the identification block works in detail will be explained in the following section. The result of this block is a model of the plant, given either as a transfer function or as a state space model. The parameters of this model form the model parameter vector \mathbf{p} .

This vector \mathbf{p} is given to the *controller design block*, which is one of the algorithms treated in this book before, e.g. PID controller design (for transfer function models), pole placement or Riccati controller design (for state space models). The controller parameters, that means the coefficients of the controller transfer function or the coefficients of the controller feedback matrix, form the vector \mathbf{f} .

This vector \mathbf{f} is given to the plant controller to update the controller parameters after each new calculation. The identification resp. the controller calculation can be done at every single time-step or alternatively in larger time intervals after collection of a sufficient amount of new signal values \mathbf{u} and \mathbf{y} . Due to the two-stage approach, there is a good chance to establish a

safety routine by checking the identified plant model for plausibility before starting the controller design block.

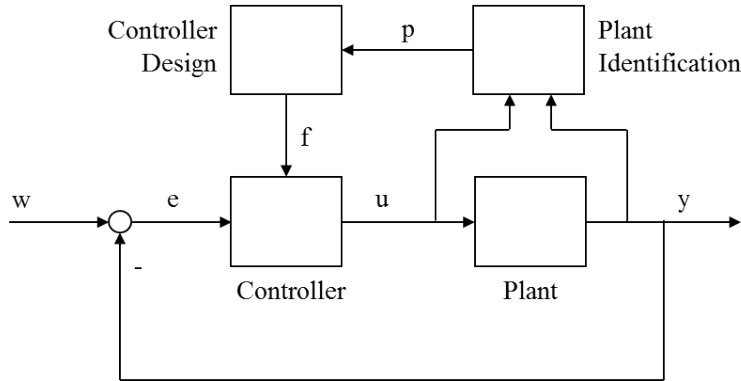


Abb. 12.1. Indirect adaptation

The second approach can be called *Direct Adaptation* (see fig. 12.2). This approach is a single-stage approach. Starting point is a reference model of the **closed-loop system**, not of the plant. This model gets the same input signal w as the closed loop, and it creates an output signal y_d , which is compared to the output y of the closed-loop system. The task of the adaptation algorithm is to design a controller, that minimizes the difference between reference model output and closed-loop system output, where the algorithm uses the reference signal w and the controller output u as additional information. Again, the controller may be given in form of a controller transfer function (PID) or in form of a state feedback matrix, and the calculated controller parameters form the vector f that is given to the controller in certain time intervals, as explained for the indirect approach.

How this algorithm works will be shown in section 12.3. As this is only a single-step approach, it is hard to establish a safety routine. The only chance is to think about some upper and lower bounds for each controller parameter that have to be checked before uploading f to the online controller.

The final approach is the most common approach and normally used in practical applications, it is the so-called *Gain Scheduling* shown in fig. 12.3. First, different set points have to be defined, depending on input and output values of the closed-loop system, the controller and the plant. Then, before the system gets online, for each set point a controller is calculated using any known method for controller design, and the calculated controller parameters are stored in a controller parameter set storage. When the system gets online, the *Gain Scheduling* block just has to identify the current set point from the current measurements and upload the corresponding controller parameter set f to the running system.

Obviously, this approach offers the best safety conditions, as every controller can be calculated and checked for stability and performance in detail

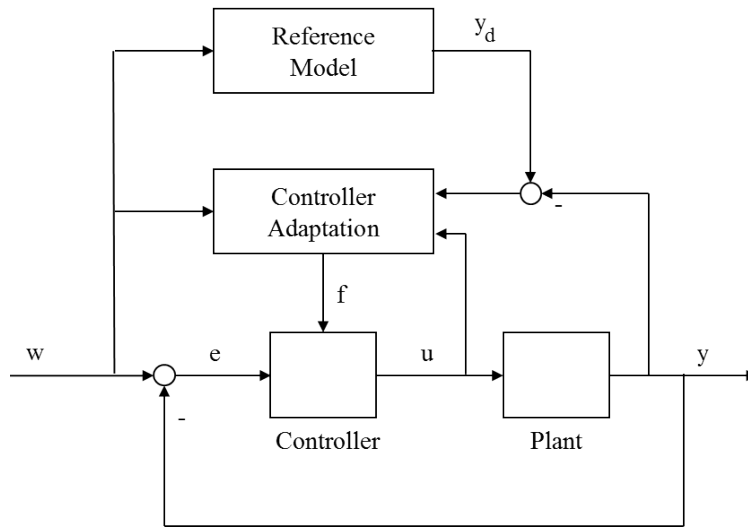


Abb. 12.2. Direct adaptation

off line during design phase, while under operation the Gain Scheduling block is just measuring and choosing a controller parameter set.

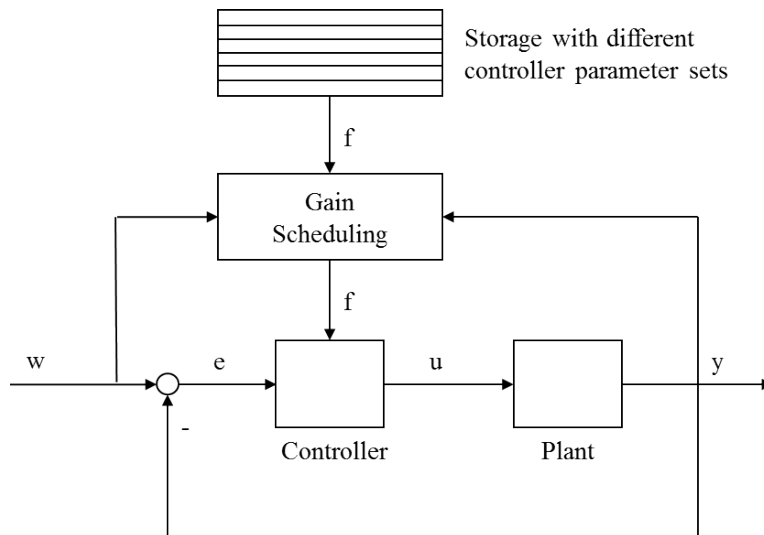


Abb. 12.3. Gain Scheduling

12.2 Identification

12.2.1 Least Squares Method

Goal: Identification of plant parameters a_i and b_i from measured input-output signals:

$$G(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + \dots + b_0}{a_n z^n + a_{n-1} z^{n-1} + \dots + a_0} = \frac{Y(z)}{U(z)} \quad (12.1)$$

Or in the time domain:

$$y(k) + a_{n-1}y(k-1) + \dots + a_0y(k-n) = b_mu(k-n+m) + \dots + b_0u(k-n) \quad (12.2)$$

Rearranging this equation leads to

$$y(k) = -a_{n-1}y(k-1) - \dots - a_0y(k-n) + b_mu(k-n+m) + \dots + b_0u(k-n) \quad (12.3)$$

For N different measurements $k = 1, \dots, N$ we get this equation for N times. In vector notation, we get

$$\mathbf{y}_N = \mathbf{H}_N \Theta \quad (12.4)$$

with

$$\begin{aligned} \mathbf{y}_N &= (y(1), y(2), \dots, y(N))^T \\ \Theta &= (a_{n-1}, a_{n-2}, \dots, a_0, b_m, \dots, b_0)^T \\ \mathbf{H}_N &= \begin{pmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_N^T \end{pmatrix} \\ \mathbf{h}_k^T &= (-y(k-1), \dots, -y(k-n), u(k-n+m), \dots, u(k-n)) \end{aligned} \quad (12.5)$$

As N is normally much larger than $n+m$, the equation system (12.4) is overdetermined, as we have more equations than variables. Because the equations result from measured values with measuring errors, it is impossible to find one solution Θ that fits all equations. To overcome this problem we introduce an unknown error $e(k)$ in each equation:

$$\begin{aligned} y(k) &= -a_{n-1}y(k-1) - \dots - a_0y(k-n) \\ &\quad + b_mu(k-n+m) + \dots + b_0u(k-n) + e(k) \end{aligned} \quad (12.6)$$

resp.

$$\mathbf{y}_N = \mathbf{H}_N \Theta + \mathbf{e}_N \quad (12.7)$$

with

$$\mathbf{e}_N = (e(1), e(2), \dots, e(N))^T \quad (12.8)$$

This unknown vector gives us the chance to reformulate the task: Instead of trying to find a precise solution for (12.4), we will now look for a vector Θ , which minimizes

$$J(\Theta) = \mathbf{e}_N^T \mathbf{e}_N = \sum_{k=1}^N (e(k))^2 \quad (12.9)$$

This idea gives the name *least squares method*. Inserting (12.7) into (12.9) we get

$$\begin{aligned} J(\Theta) &= \mathbf{e}_N^T \mathbf{e}_N = (\mathbf{y}_N - \mathbf{H}_N \Theta)^T (\mathbf{y}_N - \mathbf{H}_N \Theta) \\ &= (\mathbf{y}_N^T - \Theta^T \mathbf{H}_N^T) (\mathbf{y}_N - \mathbf{H}_N \Theta) \\ &= \mathbf{y}_N^T \mathbf{y}_N - \mathbf{y}_N^T \mathbf{H}_N \Theta - \Theta^T \mathbf{H}_N^T \mathbf{y}_N + \Theta^T \mathbf{H}_N^T \mathbf{H}_N \Theta \end{aligned}$$

As $\mathbf{y}_N^T \mathbf{H}_N \Theta$ is a scalar, it is equal to its transpose:

$$\mathbf{y}_N^T \mathbf{H}_N \Theta = \Theta^T \mathbf{H}_N^T \mathbf{y}_N \quad (12.10)$$

and so

$$J(\Theta) = \mathbf{y}_N^T \mathbf{y}_N - 2\Theta^T \mathbf{H}_N^T \mathbf{y}_N + \Theta^T \mathbf{H}_N^T \mathbf{H}_N \Theta \quad (12.11)$$

A minimum value for $J(\Theta)$ requires

$$\frac{\delta J(\Theta)}{\delta \Theta} = 0 \quad (12.12)$$

Inserting (12.11) gives

$$\begin{aligned} \frac{\delta J(\Theta)}{\delta \Theta} &= -2\mathbf{H}_N^T \mathbf{y}_N + 2\mathbf{H}_N^T \mathbf{H}_N \Theta \\ 0 &= -2\mathbf{H}_N^T \mathbf{y}_N + 2\mathbf{H}_N^T \mathbf{H}_N \Theta \\ \mathbf{H}_N^T \mathbf{H}_N \Theta &= \mathbf{H}_N^T \mathbf{y}_N \\ \Theta &= (\mathbf{H}_N^T \mathbf{H}_N)^{-1} \mathbf{H}_N^T \mathbf{y}_N \end{aligned} \quad (12.13)$$

The critical point is the matrix $(\mathbf{H}_N^T \mathbf{H}_N)$. If it is not regular, it cannot be inverted. This depends on the measurements. And this shows the problem of automatic plant identification: If the measurements do not contain enough information, this matrix will be irregular, and there will be no (useful) result.

12.2.2 Recursive Least Squares Method

The classical least squares algorithm requires a large number of measurements to identify the plant parameters Θ . If the plant is varying over the time, we need to start the collection of measurements again after a while and calculate a new solution.

It would be a better way to calculate a first solution by the classical least squares algorithm and then to improve the solution continuously at each time step using the latest measurements. This improvement at every time step is called *recursive least squares algorithm*. To develop the formulas, we first need to analyze the structure of the matrices and vectors of eq. (12.13).

Following (12.5) we get

$$\mathbf{H}_N = \begin{pmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_N^T \end{pmatrix} = \begin{pmatrix} -y(0) & \dots & -y(1-n) & u(1-n+m) & \dots & u(1-n) \\ -y(1) & \dots & -y(2-n) & u(2-n+m) & \dots & u(2-n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -y(N-1) & \dots & -y(N-n) & u(N-n+m) & \dots & u(N-n) \end{pmatrix} \quad (12.14)$$

Rewriting shorter leads to

$$\mathbf{H}_N = (-\mathbf{y}(0), \dots, -\mathbf{y}(1-n), \mathbf{u}(1-n+m), \dots, \mathbf{u}(1-n)) \quad (12.15)$$

with new vectors

$$\mathbf{y}(k) = (y(k), y(k+1), y(k+2), \dots, y(k+N-1))^T \quad (12.16)$$

$$\mathbf{u}(k) = (u(k), u(k+1), u(k+2), \dots, u(k+N-1))^T \quad (12.17)$$

Using these new vectors, we can write

$$\begin{aligned} \mathbf{H}_N^T \mathbf{H}_N &= \begin{pmatrix} -\mathbf{y}^T(0) \\ \vdots \\ -\mathbf{y}^T(1-n) \\ -\mathbf{u}^T(1-n+m) \\ \vdots \\ -\mathbf{u}^T(1-n) \end{pmatrix} (-\mathbf{y}(0), \dots, -\mathbf{y}(1-n), \mathbf{u}(1-n+m), \dots, \mathbf{u}(1-n)) \\ &= \begin{pmatrix} \mathbf{y}^T(0)\mathbf{y}(0) & \mathbf{y}^T(0)\mathbf{y}(-1) & \mathbf{y}^T(0)\mathbf{y}(-2) & \dots \\ \mathbf{y}^T(-1)\mathbf{y}(0) & \mathbf{y}^T(-1)\mathbf{y}(-1) & \mathbf{y}^T(-1)\mathbf{y}(-2) & \dots \\ \mathbf{y}^T(-2)\mathbf{y}(0) & \mathbf{y}^T(-2)\mathbf{y}(-1) & \mathbf{y}^T(-2)\mathbf{y}(-2) & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \end{aligned} \quad (12.18)$$

Each coefficient of this matrix is a vector product and therefore a sum of the form

$$a_{ij} = \sum_{l=1}^N y(l-i)y(l-j) \quad (12.19)$$

where there can also be u instead of y .

The last term of this sum for each coefficient is

$$\begin{pmatrix} y(N-1)y(N-1) & y(N-1)y(N-2) & y(N-1)y(N-3) & \dots \\ y(N-2)y(N-1) & y(N-2)y(N-2) & y(N-2)y(N-3) & \dots \\ y(N-3)y(N-1) & y(N-3)y(N-2) & y(N-3)y(N-3) & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad (12.20)$$

The same matrix we get from the following vector product (using the definition in (12.5)):

$$\begin{aligned}
\mathbf{h}_N \mathbf{h}_N^T &= \begin{pmatrix} -y(N-1) \\ \dots \\ -y(N-n) \\ u(N-n+m) \\ \dots \\ u(N-n) \end{pmatrix} (-y(N-1), \dots, -y(N-n), u(N-n+m), \dots, u(N-n)) \\
&= \begin{pmatrix} y(N-1)y(N-1) & y(N-1)y(N-2) & y(N-1)y(N-3) & \dots \\ y(N-2)y(N-1) & y(N-2)y(N-2) & y(N-2)y(N-3) & \dots \\ y(N-3)y(N-1) & y(N-3)y(N-2) & y(N-3)y(N-3) & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad (12.21)
\end{aligned}$$

So we can write

$$\mathbf{H}_N^T \mathbf{H}_N = \mathbf{H}_{N-1}^T \mathbf{H}_{N-1} + \mathbf{h}_N \mathbf{h}_N^T \quad (12.22)$$

$\mathbf{H}_{N-1}^T \mathbf{H}_{N-1}$ contains only measurements up to time step $N-1$, and after adding $\mathbf{h}_N \mathbf{h}_N^T$ we get $\mathbf{H}_N^T \mathbf{H}_N$. This means, that $\mathbf{H}_N^T \mathbf{H}_N$ can be calculated by taking the same matrix from the time step before with $N-1$ measurements and adding the term $\mathbf{h}_N \mathbf{h}_N^T$, which contains older but also the latest measurements of time step N .

Next, we analyze

$$\begin{aligned}
\mathbf{H}_N^T \mathbf{y}_N &= \begin{pmatrix} -\mathbf{y}^T(0) \\ \dots \\ -\mathbf{y}^T(1-n) \\ -\mathbf{u}^T(1-n+m) \\ \dots \\ -\mathbf{u}^T(1-n) \end{pmatrix} \begin{pmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{pmatrix} \quad (12.23) \\
&= \begin{pmatrix} -y(0)y(1) - y(1)y(2) - \dots - y(N-1)y(N) \\ \dots \\ -y(1-n)y(1) - y(2-n)y(2) - \dots - y(N-n)y(N) \\ u(1-n+m)y(1) + \dots + u(N-n+m)y(N) \\ \dots \\ u(1-n)y(1) + u(2-n)y(2) + \dots + u(N-n)y(N) \end{pmatrix}
\end{aligned}$$

Furthermore, we have

$$\mathbf{h}_N y(N) = \begin{pmatrix} -y(N-1) \\ \dots \\ -y(N-n) \\ u(N-n+m) \\ \dots \\ u(N-n) \end{pmatrix} y(N) = \begin{pmatrix} -y(N-1)y(N) \\ \dots \\ -y(N-n)y(N) \\ u(N-n+m)y(N) \\ \dots \\ u(N-n)y(N) \end{pmatrix} \quad (12.24)$$

Each coefficient of this vector is exactly the last term in each sum of each coefficient of the vector in (12.23). So we can write

$$\mathbf{H}_N^T \mathbf{y}_N = \mathbf{H}_{N-1}^T \mathbf{y}_{N-1} + \mathbf{h}_N y(N) \quad (12.25)$$

This result is similar to that of equation (12.22). Again, we can use the old term from the last time step and add some other term containing the latest measurements to get the new term for the current time step.

Using the equations (12.13), (12.22) and (12.25), we can now develop the final formula for the *recursive least squares algorithm*. We start with equation (12.13):

$$\Theta_N = (\mathbf{H}_N^T \mathbf{H}_N)^{-1} \mathbf{H}_N^T \mathbf{y}_N \quad (12.26)$$

Inserting (12.25) leads to

$$\Theta_N = (\mathbf{H}_N^T \mathbf{H}_N)^{-1} (\mathbf{H}_{N-1}^T \mathbf{y}_{N-1} + \mathbf{h}_N y(N)) \quad (12.27)$$

Applying (12.13) with index $(N-1)$ instead of N gives

$$\Theta_N = (\mathbf{H}_N^T \mathbf{H}_N)^{-1} (\mathbf{H}_{N-1}^T \mathbf{H}_{N-1} \Theta_{N-1} + \mathbf{h}_N y(N)) \quad (12.28)$$

Now we insert equation (12.22):

$$\Theta_N = (\mathbf{H}_N^T \mathbf{H}_N)^{-1} (\mathbf{H}_N^T \mathbf{H}_N \Theta_{N-1} - \mathbf{h}_N \mathbf{h}_N^T \Theta_{N-1} + \mathbf{h}_N y(N)) \quad (12.29)$$

Multiplying and rearranging gives the final formula:

$$\Theta_N = \Theta_{N-1} + (\mathbf{H}_N^T \mathbf{H}_N)^{-1} \mathbf{h}_N (y(N) - \mathbf{h}_N^T \Theta_{N-1}) \quad (12.30)$$

Comparing this result with equation (12.3) and definition (12.5) we can see, that the last term $\mathbf{h}_N^T \Theta_{N-1}$ is exactly the right-hand side of eq. (12.3) for the last measurement $k = N$ with the parameter vector Θ_{N-1} estimated at the time step before. So we can write for this term:

$$\hat{y} := \mathbf{h}_N^T \Theta_{N-1} \quad (12.31)$$

where \hat{y} is the calculated output value using the latest measurements with the parameter vector Θ_{N-1} estimated at the time step before.

Using this result, we can now interpret eq. (12.30). At each time step, the new parameter vector Θ_N can be calculated by using the old parameter vector Θ_{N-1} estimated at the time step before and adding a correction term. This correction term is the model error $(y(N) - \mathbf{h}_N^T \Theta_{N-1})$, which is equal to $(y(N) - \hat{y})$, multiplied with the correction gain factor $(\mathbf{H}_N^T \mathbf{H}_N)^{-1} \mathbf{h}_N$.

Some final remarks:

1. Equation (12.30) is only one possible solution for recursive identification. There exist other alternatives.
2. The quality of an automatic identification, recursive or non-recursive, depends strongly on the quality of the measurements. If the measurements do not contain enough information about the plant behavior, the identified model cannot be a full representation of the plant. It represents the plant correctly only in that set points, which are represented sufficiently by the given measurements.

3. The recursive method shown here always adds the information of a new measurement to the given old model, as discussed for eq. (12.30). This means, that no information will ever be forgotten. Each measurement of the past will always give a contribution to the estimated parameter set. This is perfect for a non-varying plant, as with each time step we add new information to improve our plant model. But this is not useful, if the plant is changing over the time. In that case, old information should be replaced by new information. For this purpose, the algorithm can be extended by identifying the contribution of the oldest measurement and subtract it in the same way, as we have here identified the contribution of the latest measurement and added it to the result of the previous time step (equations (12.22) and (12.25)).

12.3 Direct Adaptation

12.3.1 SISO systems

Idea: We choose a stability criterion and design our adaptation law for the controller in a way, so that the stability criterion is guaranteed.

Here: We choose the direct method of Lyapunov. Alternatively, the Hyperstability criterion or others could be chosen.

Motivating example: 1st-order lag, given by the differential equation

$$\dot{y}(t) = ay(t) + bu(t) \quad (12.32)$$

with unknown parameters a, b , but known system order ($= 1$) and known sign of b . This restriction is not hard, as we just need to know, if a positive controller output $u(t)$ leads to an increasing or decreasing output variable $y(t)$. Now we are looking for a time-variant P-controller with prefilter:

$$u(t) = r(t)w(t) - k(t)y(t) \quad (12.33)$$

This controller has the classical state-space controller structure. In a 1st-order lag, the output y is equal to the state variable x . Therefore, k is equivalent to the state-feedback matrix \mathbf{F} (with just a different sign), and r is equivalent to the reference weighting matrix \mathbf{R} . These assumptions and definitions lead to the overall block diagram in fig. 12.4.

Be aware of the fact, that here $e(t)$ is not the control error but the modelling error $e(t) = y_M(t) - y(t)$. The model (which is the model of the closed-loop system!) is a 1st-order lag, too:

$$\dot{y}_M(t) = a_M y_M(t) + b_M w(t) \quad (12.34)$$

where a_M and b_M are known and $a_M < 0$, because the model must be stable.

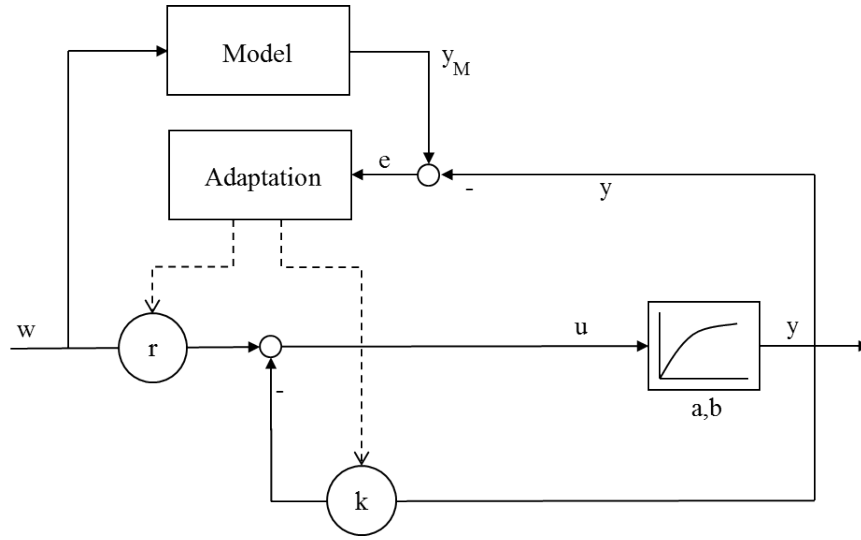


Abb. 12.4. Simple adaptive P-controller with 1st-order lag

The goal of the adaptation is to make the closed-loop system behavior equal to the model behavior, e.g. getting the same reference input w , the closed-loop system output shall be equal to the model output:

$$\lim_{t \rightarrow \infty} e(t) = \lim_{t \rightarrow \infty} (y_M(t) - y(t)) = 0 \quad (12.35)$$

Inserting the controller equation (12.33) into the plant equation (12.32) leads to

$$\begin{aligned} \dot{y}(t) &= ay(t) + b(r(t)w(t) - k(t)y(t)) \\ &= [a - bk(t)]y(t) + br(t)w(t) \\ &= a_R(t)y(t) + b_R(t)w(t) \end{aligned} \quad (12.36)$$

with $a_R(t) = a - bk(t)$ and $b_R(t) = br(t)$ being the coefficients of the regulated system.

The optimal solution and perfect adaptation would be $a_R = a_M$ and $b_R = b_M$, so that

$$k_o = \frac{a - a_M}{b} \quad r_o = \frac{b_M}{b} \quad (12.37)$$

but unfortunately, this cannot be calculated, as a, b are unknown. We need another approach, starting with the definition of the error between the model and the closed-loop system:

$$\begin{aligned} e(t) &= y_M(t) - y(t) \\ \dot{e}(t) &= \dot{y}_M(t) - \dot{y}(t) \end{aligned} \quad (12.38)$$

Inserting (12.34) and (12.36) gives

$$\begin{aligned}
\dot{e}(t) &= a_M y_M(t) + b_M w(t) - a_R(t)y(t) - b_R(t)w(t) \\
&= a_M y_M(t) - a_M y(t) + a_M y(t) - a_R(t)y(t) + (b_M - b_R(t))w(t) \\
&= a_M(y_M(t) - y(t)) + (a_M - a_R(t))y(t) + (b_M - b_R(t))w(t) \\
&= a_M e(t) + \Delta a(t)y(t) + \Delta b(t)w(t)
\end{aligned} \tag{12.39}$$

with the error between the output of the model and the closed-loop system $e(t) = y_M(t) - y(t)$, and the coefficient errors $\Delta a(t) = a_M - a_R(t)$ and $\Delta b(t) = b_M - b_R(t)$ between the closed-loop system and the model.

Now, following the Direct Method of Lyapunov, we define a Lyapunov function

$$V(t) := \frac{1}{2}(q_1 e^2(t) + q_2 (\Delta a(t))^2 + q_3 (\Delta b(t))^2) \tag{12.40}$$

with positive constant values $q_i > 0$. V shall be minimized and, if possible, converge even to zero. This implies, that the model error $e(t)$ as well as the coefficient errors $\Delta a(t)$ and $\Delta b(t)$ will be zero at the end.

To apply the Direct Method we have to ensure that the derivative of V over the time is negative and therefore, V will get smaller and smaller. Starting with (12.40) and inserting (12.39) we get

$$\begin{aligned}
\dot{V}(t) &= q_1 e(t)\dot{e}(t) + q_2 \Delta a(t)\dot{\Delta a}(t) + q_3 \Delta b(t)\dot{\Delta b}(t) \\
&= q_1 e(t)(a_M e(t) + \Delta a(t)y(t) + \Delta b(t)w(t)) + \\
&\quad q_2 \Delta a(t)\dot{\Delta a}(t) + q_3 \Delta b(t)\dot{\Delta b}(t) \\
&= q_1 a_M e^2(t) + \Delta a(t)(q_2 \dot{\Delta a}(t) + q_1 e(t)y(t)) + \\
&\quad \Delta b(t)(q_3 \dot{\Delta b}(t) + q_1 e(t)w(t))
\end{aligned} \tag{12.41}$$

If we can guarantee

$$\begin{aligned}
q_2 \Delta a(t) + q_1 e(t)y(t) &= 0 \\
q_3 \Delta b(t) + q_1 e(t)w(t) &= 0
\end{aligned} \tag{12.42}$$

the second and third term in (12.41) will disappear, and we get

$$\dot{V}(t) = q_1 a_M e^2(t) < 0 \tag{12.43}$$

for $e(t) \neq 0$, because $a_M < 0$, $q_1 > 0$ and $e^2(t) > 0$ for $e(t) \neq 0$. Therefore, the Lyapunov function V will always get smaller and smaller until $e(t) = 0$, so that at the end the error will be equal zero, which is our main adaptation goal. But as soon as $e(t) = 0$, V will stop decreasing, independently from the values of $\Delta a(t)$ and $\Delta b(t)$. Therefore it could happen, that at the end of the adaptation $e(t)$ will be completely disappeared, but there will still be modelling errors $\Delta a(t)$ and $\Delta b(t)$. It means, that the closed-loop system is behaving like the given model, but with different system coefficients. Anyway, from the control point of view this is no problem. The goal was to behave like the model, and not to find the precise system coefficients.

As a final step, we have to think about how to guarantee the conditions of eq. (12.42). From these conditions, it follows

$$\begin{aligned}\Delta\dot{a}(t) &= -\frac{q_1}{q_2}e(t)y(t) \\ \Delta\dot{b}(t) &= -\frac{q_1}{q_3}e(t)w(t)\end{aligned}\quad (12.44)$$

Using the definitions of $\Delta a(t)$ and $\Delta b(t)$ as well as the definitions of a_R and b_R we get

$$\begin{aligned}\Delta\dot{a}(t) &= \frac{d}{dt}(a_M - a_R(t)) = -\dot{a}_R(t) \\ &= -\frac{d}{dt}(a - bk(t)) = b\dot{k}(t) \\ \Delta\dot{b}(t) &= \frac{d}{dt}(b_M - b_R(t)) = -\dot{b}_R(t) = -b\dot{r}(t)\end{aligned}\quad (12.45)$$

From (12.44) and (12.45) it follows

$$\begin{aligned}\dot{k}(t) &= -\frac{q_1}{q_2b}e(t)y(t) \\ \dot{r}(t) &= \frac{q_1}{q_3b}e(t)w(t)\end{aligned}\quad (12.46)$$

The sign of b is known, and all q_i are positive parameters, that can be chosen by the user. Therefore, we can replace these parameters by selectable adaptation parameters α_i with the restriction $\text{sgn}(\alpha_1) = \text{sgn}(\alpha_2) = \text{sgn}(b)$:

$$\begin{aligned}\dot{k}(t) &= -\alpha_1 e(t)y(t) \\ \dot{r}(t) &= \alpha_2 e(t)w(t)\end{aligned}\quad (12.47)$$

This is the adaptation law we have been looking for. It defines how to change the two controller parameters $k(t)$ and $r(t)$ from time-step to time-step. Following this law guarantees decreasing of the Lyapunov function V and therefore, convergence of $e(t)$ towards zero.

Let us give some final remarks:

1. The overall system with the adaptation law of (12.47) is globally asymptotically stable in the sense of Lyapunov.
2. In different setpoints the behavior can be different (but still stable), as the overall system is nonlinear due to the adaptation.
3. Large adaptation parameters α_i mean that the system will adapt very fast, but the adaptation will tend to overshooting. Too small adaptation parameters will avoid overshooting, but the system will converge very slowly.
4. As said before, this method will guarantee convergence from $e(t)$ towards zero, but it does not guarantee $a_R = a_M$ and $b_R = b_M$ in the final state. It depends on the choice of $w(t)$, if the correct coefficients will be met.

12.3.2 MIMO systems

Now we have the same structure as in fig. 12.4, but we replace the SISO plant by a state-space MIMO system. Of course, all variables will be vectors and matrices now:

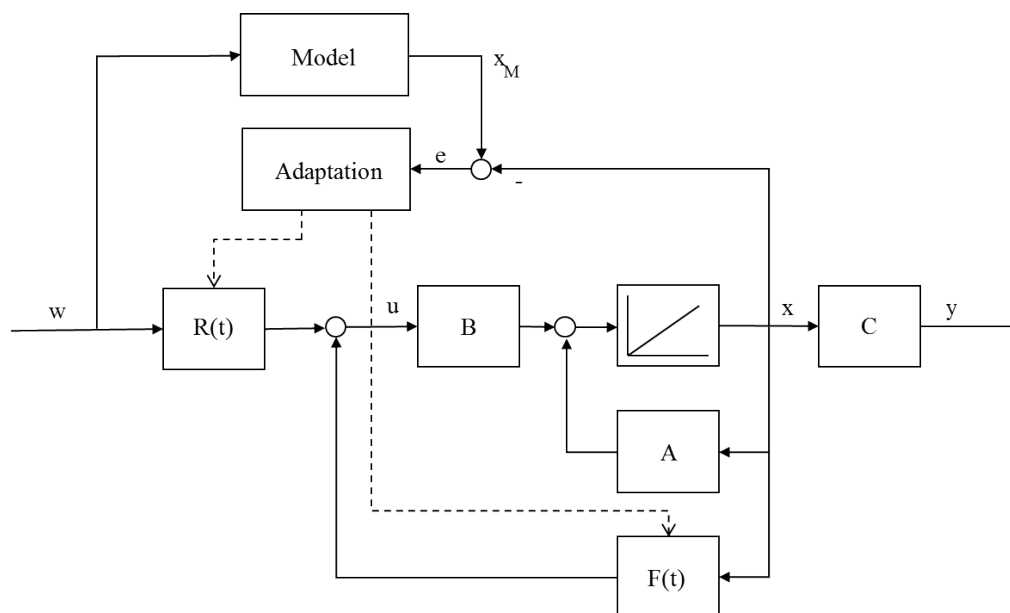


Abb. 12.5. Adaptive state-space controller

The steps will be the same as for the SISO system before. Again we have a model (compare to (12.34)), but here it is given in state-space representation, and the model output is not the output vector, but the state vector:

$$\dot{\mathbf{x}}_M(t) = \mathbf{A}_M \mathbf{x}_M(t) + \mathbf{B}_M \mathbf{w}(t) \quad (12.48)$$

with known matrices \mathbf{A}_M and \mathbf{B}_M and all eigenvalues of \mathbf{A}_M being in the left half-plane, as it would make no sense to define an unstable model.

The regulated closed-loop system (compare to (12.36)) is given here by

$$\begin{aligned}\dot{\mathbf{x}}(t) &= (\mathbf{A} + \mathbf{B}\mathbf{F}(t))\mathbf{x}(t) + \mathbf{B}\mathbf{R}(t)\mathbf{w}(t) \\ &= \mathbf{A}_R(t)\mathbf{x}(t) + \mathbf{B}_R(t)\mathbf{w}(t)\end{aligned}\quad (12.49)$$

with $\mathbf{A}_R(t) = \mathbf{A} + \mathbf{B}\mathbf{F}(t)$ and $\mathbf{B}_R(t) = \mathbf{B}\mathbf{R}(t)$. Now, we try to minimize the state vector error, with all calculation steps being very similar to the ones before for the SISO system (compare to (12.39)):

$$\mathbf{e}(t) := \mathbf{x}_M - \mathbf{x}(t)$$

$$\dot{\mathbf{e}}(t) = \dot{\mathbf{x}}_M - \dot{\mathbf{x}}(t)$$

$$\begin{aligned}
&= \mathbf{A}_M \mathbf{x}_M(t) + \mathbf{B}_M \mathbf{w}(t) - (\mathbf{A}_R(t) \mathbf{x}(t) + \mathbf{B}_R(t) \mathbf{w}(t)) \\
&= \mathbf{A}_M (\mathbf{x}_M(t) - \mathbf{x}(t)) + (\mathbf{A}_M - \mathbf{A}_R(t)) \mathbf{x}(t) + (\mathbf{B}_M - \mathbf{B}_R(t)) \mathbf{w}(t) \\
&= \mathbf{A}_M \mathbf{e}(t) + \Delta \mathbf{A}(t) \mathbf{x}(t) + \Delta \mathbf{B}(t) \mathbf{w}(t)
\end{aligned} \tag{12.50}$$

where $\Delta \mathbf{A}(t) = \mathbf{A}_M - \mathbf{A}_R(t)$ is the system matrix error and $\Delta \mathbf{B}(t) = \mathbf{B}_M - \mathbf{B}_R(t)$ the input matrix error.

Obviously, the Lyapunov function must have more terms than in eq. (12.40), as we have to include the error vector as well as all coefficients of the error matrices:

$$V(t) := \mathbf{e}^T(t) \mathbf{P} \mathbf{e}(t) + \sum_{i,j} \alpha_{i,j} (\Delta a_{i,j}(t))^2 + \sum_{i,j} \beta_{i,j} (\Delta b_{i,j}(t))^2 \tag{12.51}$$

with $\alpha_{i,j} > 0$ and $\beta_{i,j} > 0$ can be chosen, $\Delta a_{i,j}(t)$ are the coefficients of $\Delta \mathbf{A}(t)$ and $\Delta b_{i,j}(t)$ the coefficients of $\Delta \mathbf{B}(t)$. \mathbf{P} is the solution of the Lyapunov equation

$$\mathbf{A}_M^T \mathbf{P} + \mathbf{P} \mathbf{A}_M = -\mathbf{Q} \tag{12.52}$$

where \mathbf{Q} can be chosen arbitrarily, but must be positive definite. As \mathbf{A}_M has only eigenvalues in the left half of the complex plane and \mathbf{Q} is positive definite, the solution \mathbf{P} will be symmetrical and positive definite (see chap. A.5).

Again, the derivative of the Lyapunov function must be negative:

$$\begin{aligned}
\dot{V}(t) &= \dot{\mathbf{e}}^T(t) \mathbf{P} \mathbf{e}(t) + \mathbf{e}^T(t) \mathbf{P} \dot{\mathbf{e}}(t) + 2 \sum_{i,j} \alpha_{i,j} \Delta a_{i,j}(t) \Delta \dot{a}_{i,j}(t) + \\
&\quad 2 \sum_{i,j} \beta_{i,j} \Delta b_{i,j}(t) \Delta \dot{b}_{i,j}(t) < 0
\end{aligned} \tag{12.53}$$

For the first two terms of this equation we can insert (12.50) and get

$$\begin{aligned}
\dot{\mathbf{e}}^T(t) \mathbf{P} \mathbf{e}(t) + \mathbf{e}^T(t) \mathbf{P} \dot{\mathbf{e}}(t) &= \\
&= (\mathbf{e}^T(t) \mathbf{A}_M^T + \mathbf{x}^T(t) \Delta \mathbf{A}^T(t) + \mathbf{w}^T(t) \Delta \mathbf{B}^T(t)) \mathbf{P} \mathbf{e}(t) + \\
&\quad \mathbf{e}^T(t) \mathbf{P} (\mathbf{A}_M \mathbf{e}(t) + \Delta \mathbf{A}(t) \mathbf{x}(t) + \Delta \mathbf{B}(t) \mathbf{w}(t))
\end{aligned} \tag{12.54}$$

As $\mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{A}(t) \mathbf{x}(t)$ and $\mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{B}(t) \mathbf{w}(t)$ are scalars, we have

$$\begin{aligned}
\mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{A}(t) \mathbf{x}(t) &= \mathbf{x}^T(t) \Delta \mathbf{A}^T(t) \mathbf{P} \mathbf{e}(t) \\
\mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{B}(t) \mathbf{w}(t) &= \mathbf{w}^T(t) \Delta \mathbf{B}^T(t) \mathbf{P} \mathbf{e}(t)
\end{aligned} \tag{12.55}$$

Using this result and the Lyapunov equation (12.52) we can write

$$\begin{aligned}
&\dot{\mathbf{e}}^T(t) \mathbf{P} \mathbf{e}(t) + \mathbf{e}^T(t) \mathbf{P} \dot{\mathbf{e}}(t) \\
&= \mathbf{e}^T(t) (\mathbf{A}_M^T \mathbf{P} + \mathbf{P} \mathbf{A}_M) \mathbf{e}(t) + 2 \mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{A}(t) \mathbf{x}(t) + \\
&\quad 2 \mathbf{e}^T(t) \mathbf{P} \Delta \mathbf{B}(t) \mathbf{w}(t)
\end{aligned}$$

$$\begin{aligned}
&= \mathbf{e}^T(t)(\mathbf{A}_M^T \mathbf{P} + \mathbf{P} \mathbf{A}_M) \mathbf{e}(t) + 2\mathbf{h}^T(t) \Delta \mathbf{A}(t) \mathbf{x}(t) + \\
&\quad 2\mathbf{h}^T(t) \Delta \mathbf{B}(t) \mathbf{w}(t) \\
&= -\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t) + 2 \sum_{i,j} h_i(t) \Delta a_{i,j}(t) x_j(t) + \\
&\quad 2 \sum_{i,j} h_i(t) \Delta b_{i,j}(t) w_j(t)
\end{aligned} \tag{12.56}$$

with $\mathbf{h}^T(t) = (h_1(t), h_2(t), \dots) := \mathbf{e}^T(t) \mathbf{P}$. Inserting (12.56) into (12.53) we get

$$\begin{aligned}
\dot{V}(t) &= -\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t) + 2 \sum_{i,j} \Delta a_{i,j}(t) (h_i(t) x_j(t) + \alpha_{i,j} \Delta \dot{a}_{i,j}(t)) + \\
&\quad 2 \sum_{i,j} \Delta b_{i,j}(t) (h_i(t) w_j(t) + \beta_{i,j} \Delta \dot{b}_{i,j}(t))
\end{aligned} \tag{12.57}$$

As \mathbf{Q} is positive definite, $\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t) > 0$ for any $\mathbf{e}(t) \neq \mathbf{0}$, and therefore $-\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t) < 0$.

Now, we choose the adaptation laws

$$\begin{aligned}
\Delta \dot{a}_{i,j}(t) &:= -\frac{1}{\alpha_{i,j}} h_i(t) x_j(t) \\
\Delta \dot{b}_{i,j}(t) &:= -\frac{1}{\beta_{i,j}} h_i(t) w_j(t)
\end{aligned} \tag{12.58}$$

Inserting this into (12.57) gives

$$\dot{V}(t) = -\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t) < 0 \tag{12.59}$$

for $\mathbf{e}(t) \neq \mathbf{0}$. It is therefore guaranteed, that V will get smaller and converge towards zero, as long as $\mathbf{e}(t) \neq \mathbf{0}$. When $\mathbf{e}(t) = \mathbf{0}$, V will stop decreasing. As for the SISO example we have the fact, that adaptation will eliminate the state vector error $\mathbf{e}(t)$, but not automatically the coefficient error matrices $\Delta \mathbf{A}(t)$ and $\Delta \mathbf{B}(t)$.

From $\Delta \mathbf{A}(t) = \mathbf{A}_M - \mathbf{A}_R(t)$ it follows $\Delta \dot{\mathbf{A}}(t) = -\dot{\mathbf{A}}_R(t)$, and from $\Delta \mathbf{B}(t) = \mathbf{B}_M - \mathbf{B}_R(t)$ it follows $\Delta \dot{\mathbf{B}}(t) = -\dot{\mathbf{B}}_R(t)$. So we get from (12.58) the adaptation laws for the closed-loop system matrices parameters:

$$\begin{aligned}
\dot{a}_{R,i,j}(t) &= \frac{1}{\alpha_{i,j}} h_i(t) x_j(t) \\
\dot{b}_{R,i,j}(t) &= \frac{1}{\beta_{i,j}} h_i(t) w_j(t)
\end{aligned} \tag{12.60}$$

where $a_{R,i,j}(t)$ is a coefficient of $\mathbf{A}_R(t)$, and $b_{R,i,j}(t)$ is a coefficient of $\mathbf{B}_R(t)$.

From these equations, we must now find the equations how to tune the controller matrices $\mathbf{F}(t)$ and $\mathbf{R}(t)$. From $\mathbf{A}_R(t) = \mathbf{A} + \mathbf{B} \mathbf{F}(t)$ it follows $\dot{\mathbf{A}}_R(t) = \mathbf{B} \dot{\mathbf{F}}(t)$, and from $\mathbf{B}_R(t) = \mathbf{B} \mathbf{R}(t)$ it follows $\dot{\mathbf{B}}_R(t) = \mathbf{B} \dot{\mathbf{R}}(t)$.

We get

$$\begin{aligned}\dot{a}_{R,i,j}(t) &= \sum_{l=1}^p b_{i,l} \dot{f}_{l,j}(t) = \frac{1}{\alpha_{i,j}} h_i(t) x_j(t) \\ \dot{b}_{R,i,j}(t) &= \sum_{l=1}^p b_{i,l} \dot{r}_{l,j}(t) = \frac{1}{\beta_{i,j}} h_i(t) w_j(t)\end{aligned}\quad (12.61)$$

with $f_{l,j}(t)$ being the coefficients of the controller feedback matrix $\mathbf{F}(t)$, and $r_{l,j}(t)$ being the coefficients of the reference filter matrix $\mathbf{R}(t)$. These coefficients are modified by the adaptation and therefore time-depending.

As a final step, we multiply the first equation with $\alpha_{i,j}$ and the second one with $\beta_{i,j}$:

$$\sum_{l=1}^p \alpha_{i,j} b_{i,l} \dot{f}_{l,j}(t) = h_i(t) x_j(t) \quad (12.62)$$

$$\sum_{l=1}^p \beta_{i,j} b_{i,l} \dot{r}_{l,j}(t) = h_i(t) w_j(t) \quad (12.63)$$

These two equations represent an equation system for the unknown changes of the controller parameters $\dot{f}_{l,j}(t)$ and $\dot{r}_{l,j}(t)$. For each combination (i, j) we get one equation, so that at the end from equation (12.62) we get n^2 equations for the $p * n$ unknown parameters of $\mathbf{F}(t)$ respectively $\dot{\mathbf{F}}(t)$, and from equation (12.63) we get $n * p$ equations for the $p * p$ unknown parameters of $\mathbf{R}(t)$ respectively $\dot{\mathbf{R}}(t)$.

But we have to face the fact that the plant input matrix \mathbf{B} is unknown, and therefore the coefficients $b_{i,l}$ are also unknown. We can use the large number of equations to calculate some of these coefficients, too. And for the remaining rest of the coefficients $b_{i,l}$ we can use the same assumption as for the SISO case: We don't know the value of the coefficient, but at least we know the sign. And as the coefficients are multiplied with the free parameters $\alpha_{i,j}$ and $\beta_{i,j}$ anyway, we can skip the values of the $b_{i,l}$ and just use their signs to calculate the products $\alpha_{i,j} \text{sgn}(b_{i,l})$ and $\beta_{i,j} \text{sgn}(b_{i,l})$.

The exact algorithm depends on the dimensions p and n , i.e. how many coefficients $b_{i,l}$ remain uncalculated after solving the equation system given by equations (12.62) and (12.63).

The final remarks for the MIMO case are exactly the same as for the SISO case. But there is one additional problem here: To calculate $\dot{f}_{l,j}(t)$ from equation (12.61), we need the values of the state vector $x_j(t)$, which can not be measured and normally have to be calculated with the help of an observer. For such an observer, on the other hand, we need a precise model of the plant, while the assumption for our adaptive controller is that we don't know the plant's matrices \mathbf{A} and \mathbf{B} .

This problem fits to the block diagram in figure 12.5, as it shows that we feed back the state vector to the adaptation block, not the output vector. It is possible to transfer the complete algorithm's idea to the case of plant output feedback. But as the equations would get more complicated, we have restricted ourselves here to the more simple case of state feedback, accepting the fact that this algorithm is not complete for using it in reality.

13. Analyse nichtlinearer Systeme

13.1 Eigenschaften nichtlinearer Systeme

In den vorangegangenen Kapiteln wurden ausschließlich lineare Strecken und die zugehörigen Regler behandelt, also Systeme, die durch lineare Differentialgleichungen beschrieben werden können. Reale Systeme enthalten aber fast immer ein oder mehrere nichtlineare Übertragungsglieder. Schon ein nichtlinearer Zusammenhang macht jedoch aus einem linearen ein nichtlineares Gleichungssystem, zu dessen Beschreibung man von der vereinfachten, linearen Form der Zustandsdarstellung

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx} + \mathbf{Du}\end{aligned}\tag{13.1}$$

wieder zur allgemeinen Form (vgl. (7.2)) zurückkehren muss:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}, \mathbf{u})\end{aligned}\tag{13.2}$$

Denkbar ist auch noch eine *Zeitvarianz* des Systems, was dadurch ausgedrückt werden kann, dass die Funktionen \mathbf{f} und \mathbf{g} als zusätzlichen Parameter die Zeit t enthalten. Solche zeitvarianten Systeme sollen aber in diesem Kapitel nicht behandelt werden.

Für nichtlineare Systeme sind viele der aus der linearen Regelungstechnik bekannten Werkzeuge nicht mehr anwendbar. So muss jetzt beispielsweise auf die Laplace-Transformation verzichtet werden, die nur für lineare Systeme eingeführt wurde. Ebenso verhält es sich mit Ortskurven, die nur für lineare Systeme Auskunft darüber geben, wie das Ausgangssignal gegenüber einer Sinusschwingung am Eingang in Amplitude und Phase verändert ist. Auch das Überlagerungsprinzip (Satz 2.1) hat bei nichtlinearen Systemen seine Gültigkeit verloren, d.h. für gleichzeitig angreifende Eingangsgrößen können die Ausgangsgrößen nicht mehr zunächst unabhängig voneinander berechnet und dann überlagert werden. Ebenfalls nicht aufzuweisen haben die nichtlinearen Systeme die Proportionalitätseigenschaft (ebenfalls Satz 2.1), die besagt, dass sich bei Vergrößerung des Eingangssignales um einen bestimmten Faktor das Ausgangssignal um denselben Faktor vergrößert. Wenn

aber damit beispielsweise Sprungfunktionen verschiedener Höhe am Eingang möglicherweise völlig unterschiedliche Systemantworten hervorrufen, wird es auch sinnlos, Sprungantworten zur Charakterisierung des Systemverhaltens zu verwenden.

13.2 Schaltende Übertragungsglieder

13.2.1 Ideales Zweipunktglied

Um ein Gefühl dafür zu vermitteln, welche Effekte bei Nichtlinearitäten überhaupt auftreten können, sollen nun die in der Praxis sehr häufig vorkommenden, schaltenden Übertragungsglieder etwas genauer untersucht werden. Das einfachste Übertragungsglied ist das *ideale Zweipunktglied*, das man auch als einen idealen Schalter auffassen kann. Abb. 13.1 zeigt die Kennlinie. Für eine positive Eingangsgröße hat die Ausgangsgröße den Wert 1 und für eine negative Eingangsgröße den Wert -1 . Um Zweideutigkeiten zu vermeiden, wird für $x = 0$ die Ausgangsgröße $y(0) = 1$ definiert. Neben der Kennlinie ist eine Schaltung gezeigt, mit der diese Kennlinie unter idealen Bedingungen zu realisieren ist. Beide Schalter A und B sind miteinander gekoppelt und werden durch das Feld der Induktivität angesteuert, die ihrerseits als Eingangsgröße den Strom i_e erhält. Sobald der Strom sein Vorzeichen wechselt, ändern sich die Richtung des Feldes und damit die als masselos angenommenen Schalter ihre Position. Als Folge davon ändert auch die Spannung u_R ihr Vorzeichen. Betrachtet man i_e als Eingangs- und u_R als Ausgangsgröße, so hat die Anordnung gerade das in der Kennlinie abgebildete Übertragungsverhalten.

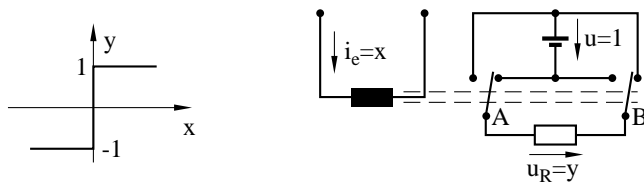


Abb. 13.1. Ideales Zweipunktglied

Schaltende Übertragungsglieder kommen in der Praxis sehr häufig als Stellglieder vor, wobei die Ausgangsgröße nicht unbedingt zwischen den Werten -1 und $+1$ hin- und herschaltet, sondern vielleicht auch zwischen 0 und 1 . Ein Transistor ist beispielsweise ein solcher Schalter, oder auch eine Ventilklappe. Mit diesen Beispielen kann auch schon die Frage beantwortet werden, warum solche Übertragungsglieder überhaupt eingesetzt werden, wenn dadurch die Theorie so erschwert wird: Ein Schalter ist einfacher und billiger als ein kontinuierliches Übertragungsglied. Sein Nachteil ist natürlich, dass seine Ausgangsgröße keinen kontinuierlichen Wertebereich abdecken kann.

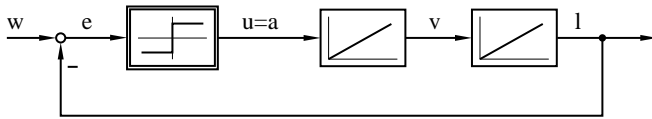


Abb. 13.2. Regelkreis mit idealem Zweipunktglied als Regler

Aber auch dieser Nachteil kann vielfach durch das später noch beschriebene Verfahren der Pulsweitenmodulation aufgehoben werden.

Interessant ist, dass Zweipunktglieder nicht nur als Stellglieder im Regelkreis auftauchen, sondern dass man sie auch als Regler selbst verwenden kann. Ideale Zweipunktglieder sind im Prinzip die einfachsten denkbaren Regler überhaupt. Abb. 13.2 zeigt ein Beispiel mit einem idealen Zweipunktglied als Regler und einem doppelten Integrator als Strecke.

Die beiden Integratoren lassen sich als Beschleunigungsstrecke auffassen (vgl. Abb. 7.2). Die Stellgröße u des Reglers entspricht damit gleichzeitig der Beschleunigung a des Körpers und die Regelgröße seiner Lage l . Die Ausgangsgröße u des Reglers ist entweder auf ihrem positiven oder negativen Maximalwert, und der Körper wird mit maximaler Kraft in positiver oder negativer Richtung beschleunigt. Der zeitliche Verlauf eines Regelvorgangs ist aus Abb. 13.3 ersichtlich. Der Istwert der Lage l sei zunächst kleiner als der Sollwert w . Die Regelabweichung e ist demnach positiv, worauf der Regler mit maximaler positiver Stellgröße u bzw. Beschleunigung a reagiert. Der Körper wird nun zum Lagesollwert hin beschleunigt. Bei Erreichen des Lagesollwertes ist die Geschwindigkeit v aber größer als Null, so dass der Körper über das Ziel hinaus schießt und eine Regelabweichung zur anderen Seite erfährt. Der Regler antwortet mit maximaler negativer Stellgröße, was den Körper zunächst abbremst und dann in die andere Richtung beschleunigt. Der ganze Vorgang wiederholt sich dann mit entgegengesetztem Vorzeichen. Offensichtlich kommt der Körper nie zur Ruhe.

Dieses Verhalten lässt sich auch in der Zustandsebene beschreiben. Wie in Abschnitt 7.2 bereits erläutert, sind Lage l und Geschwindigkeit v Zustandsgrößen des Systems. Der Regelvorgang lässt sich daher auch durch eine Trajektorie in der $v-l$ -Zustandsebene beschreiben. Da eine Dauerschwingung vorliegt, ist diese Trajektorie eine geschlossene Kurve, die immer wieder durchlaufen wird, und die Ausgangsgröße l führt Schwingungen um den Sollwert w aus.

Je weiter der Anfangswert der Lage vom Sollwert entfernt liegt, desto länger braucht der Körper, um den Sollwert zu erreichen, und desto größer ist auch seine Geschwindigkeit, wenn er den Sollwert erreicht. Dadurch fällt dann aber wiederum die Auslenkung in die entgegengesetzte Richtung umso größer aus, was insgesamt zu einer Vergrößerung sowohl der Amplitude der Schwingung als auch des Zeitintervalles zwischen zwei Nulldurchgängen führt. Demnach sind sowohl die Amplitude als auch die Frequenz dieser *Dauerschwingung* von den Anfangsbedingungen abhängig. Ein solches Verhalten

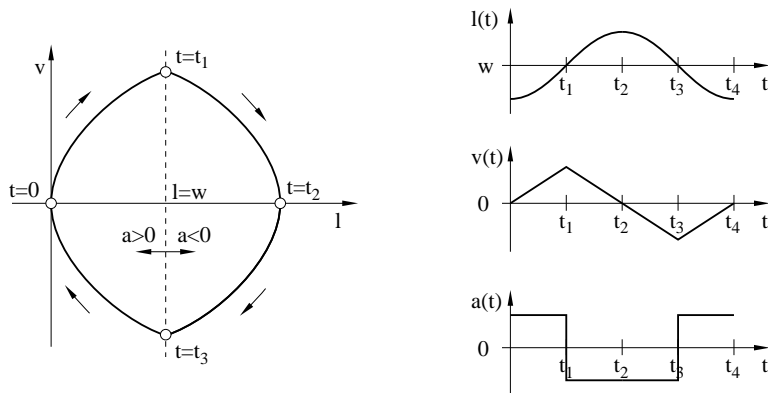


Abb. 13.3. Zustandskurve und zeitlicher Verlauf beim idealem Zweipunktglied mit doppeltem Integrator

kann es bei linearen Systemen nicht geben. Die Frequenz einer Schwingung ist dort immer durch das entsprechende konjugiert komplexe Polpaar der Übertragungsfunktion festgelegt. Nur die Amplitude hängt von den Anfangsbedingungen ab.

Abklingende Schwingungen ergeben sich dagegen, wenn die Strecke aus einem Verzögerungsglied und einem Integrator besteht, wie Abb. 13.4 zeigt. Als Zustandsgrößen können hier x_1 und x_2 gewählt werden. Die Trajektorie strebt offenbar immer weiter dem Endwert $(x_1, x_2) = (w, 0)$ zu. Unabhängig vom Anfangszustand erreicht das System immer diesen Endzustand und ist damit stabil.

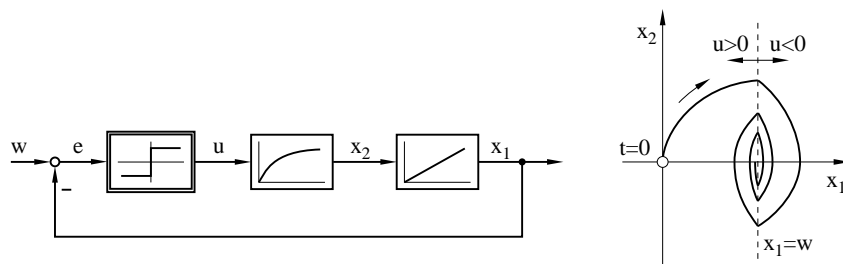


Abb. 13.4. Ideales Zweipunktglied mit IT_1 -Strecke

13.2.2 Zweipunktglied mit Hysterese

Das ideale Zweipunktglied ist, wie der Name schon sagt, praktisch aber gar nicht zu realisieren. So weisen beispielsweise die beiden Schalter in Abb. 13.1 selbstverständlich eine Masse und auch eine Haftreibung auf. Das hat aber zur Folge, dass sie nicht schon bei einem Vorzeichenwechsel des Feldes bzw. des Stromes i_e ihre Position ändern, sondern erst, wenn die Feldstärke eine

gewisse Mindestschwelle überschreitet. Das Übertragungsglied verharrt demnach auch bei einem Vorzeichenwechsel der Eingangsgröße zunächst noch auf seinem alten Wert. Erst bei Überschreiten eines Schwellwertes durch die Eingangsgröße springt die Ausgangsgröße um. Die Kennlinie eines solchen Übertragungsgliedes ist damit im Bereich um den Nullpunkt zweideutig (Abb. 13.5). Welcher Zweig der Kennlinie gerade gültig ist, hängt vom vorhergehenden Zustand ab. Insofern kann man dieses Übertragungsglied als eine Art Zweipunktglied mit Gedächtnis ansehen. Einen solchen Effekt bezeichnet man als *Hysterese*. Praktisch weisen alle schaltenden Übertragungsglieder eine mehr oder minder große Hysterese auf.

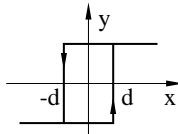


Abb. 13.5. Kennlinie mit Hysterese

Abb. 13.6 zeigt ein hysteresebehaftetes Zweipunktglied als Regler mit einem doppelten Integrator als Strecke. Die Umschaltung der Stellgröße bzw. der Beschleunigung a erfolgt verzögert gegenüber dem Nulldurchgang der Regelabweichung e . Dieses Verhalten lässt sich in der Zustandsebene durch eine Parallelverschiebung der Schaltgeraden berücksichtigen. Denn das Umschalten z.B. vom positiven auf den negativen Wert erfolgt beim Regler nicht dann, wenn die Regelabweichung Null ist, sondern erst für $e = -d$ bzw. $l = w + d$. Dies ist aber gerade auf der um d nach rechts verschobenen Schaltgeraden der Fall. Analog dazu ist die Schaltgerade für das Umschalten vom negativen auf den positiven Wert bei $l = w - d$. Durch dieses verzögerte Umschalten klingt die Schwingung aber immer weiter auf, und das System ist instabil.

Schaltet man das Zweipunktglied mit Hysterese dagegen mit einem IT_1 -Glied zusammen, so führt das System unabhängig vom Anfangszustand nach einer gewissen Zeit immer die gleiche Schwingung aus. In Abb. 13.7 ist deutlich zu erkennen, wie das System aus zwei verschiedenen Anfangszuständen in die gleiche Schwingung hineinläuft. Eine solche Schwingung bezeichnet man als *Grenzzyklus*. Im Gegensatz zur Dauerschwingung, bei der Frequenz und Amplitude vom Anfangszustand abhängig waren, sind beim Grenzzyklus sowohl die Frequenz als auch die Amplitude durch die Systemparameter vorgegeben und vom Anfangszustand völlig unabhängig. Auch Grenzzyklen können bei linearen Systemen nicht auftreten, da dort die Amplitude immer vom Anfangszustand abhängig ist. Dauerschwingungen und Grenzzyklen als spezielle nichtlineare Phänomene werden später noch exakt definiert.

Trotz ihres gegenüber idealen Zweipunktgliedern schon verschlechterten Stabilitätsverhaltens sind die realisierbaren Zweipunktglieder mit Hysterese in der Praxis häufig eingesetzte Regler. Ihr Vorteil ist ein einfacher Aufbau und der damit verbundene, niedrige Preis sowie ein schnelles Regelverhalten,

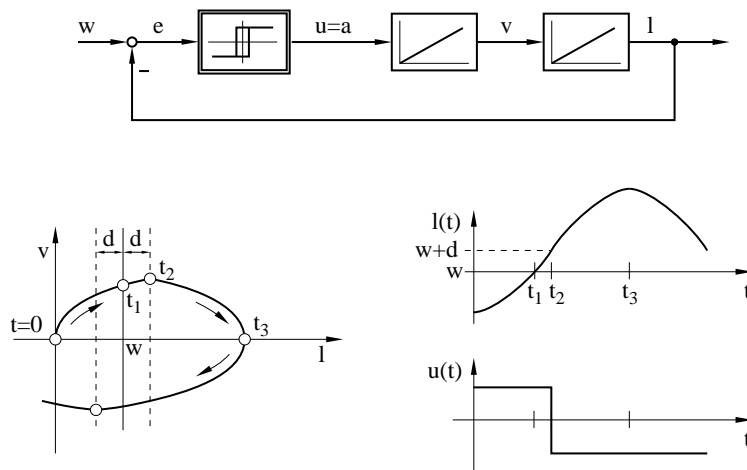
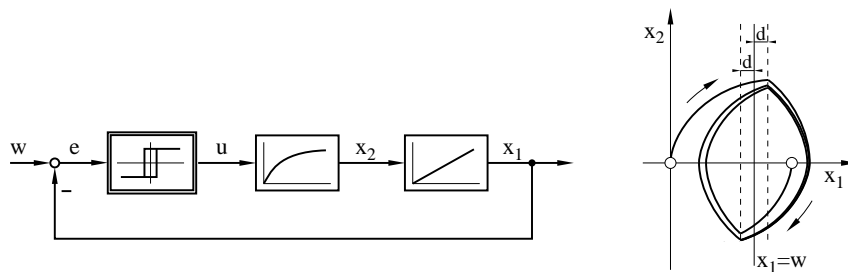


Abb. 13.6. Zweipunktglied mit Hysterese und doppeltem Integrator

Abb. 13.7. Zweipunktglied mit Hysterese und IT_1 -Strecke

was darin begründet liegt, dass die Stellgröße immer den positiven oder negativen Maximalwert annimmt. Problematisch ist aber die Tatsache, dass ein System mit Zweipunktregler nie zur Ruhe kommt, sondern auch im stabilen Zustand immer Schwingungen ausführt. Denn die Ausgangsgröße des Reglers alterniert ständig zwischen zwei Extremwerten und ist damit immer entweder zu groß oder zu klein. Solange die Amplituden dieser Schwingungen innerhalb eines vorgegebenen Toleranzbereiches bleiben, können sie akzeptiert werden. Wenn der Toleranzbereich aber nicht eingehalten wird, sollte anstatt eines Zweipunktreglers der im folgenden Abschnitt vorgestellte Dreipunktregler eingesetzt werden.

13.2.3 Dreipunktglied

Wie der Zweipunktregler kann auch der Dreipunktregler in idealer Form oder mit Hysterese vorkommen (Abb. 13.8). Als Regelstrecke sei hier ein Integrator mit einem nachfolgenden Verzögerungsglied beliebiger Ordnung angenommen. Solange sich die Regelabweichung e und damit die Differenz

$w - y$ außerhalb des Intervalles $[-\varepsilon, \varepsilon]$ befinden, ist auch $u \neq 0$. Da damit die Eingangsgröße des Integrators von Null verschieden ist, verändert sich seine Ausgangsgröße y' und mit ihr auch y . Erst für $u = 0$ bleibt der Integrator auf dem erreichten Wert stehen, und das System kann zur Ruhe kommen.

$u = 0$ ist aber gleichbedeutend damit, dass sich die Regelabweichung $e = w - y$ innerhalb des Intervalles $[-\varepsilon, \varepsilon]$ befindet. Das System erreicht also gerade dann seinen Ruhezustand, wenn sich die Ausgangsgröße y in einem Toleranzbereich $[-\varepsilon, \varepsilon]$ um den Sollwert w befindet. Dies ist ein für viele praktische Anwendungen ausreichendes Ergebnis. Stationäre Genauigkeit (d.h. $w = y$) wird allerdings nicht erzielt.

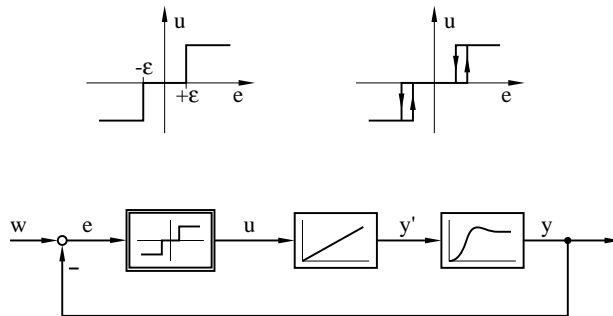


Abb. 13.8. Dreipunktglied ohne und mit Hysterese

Zu beachten ist, dass ein Dreipunktregler nur mit nachgeschaltetem Integrator sinnvoll ist, wie dies in Abb. 13.8 auch skizziert wurde. Damit das System nämlich tatsächlich zur Ruhe kommen kann, muss am Eingang des hinteren Streckenteiles die Größe y' einen Wert annehmen, für den die Regelgröße y am Ausgang im Toleranzbereich $[-\varepsilon, \varepsilon]$ von w liegt. Ein solcher Wert für y' kann aber vom Dreipunktregler selbst, der nur drei verschiedene Ausgangswerte kennt, in den meisten Fällen gar nicht erzeugt werden. Mit einem Dreipunktregler ohne nachgeschalteten Integrator würde der Regelkreis genauso in eine stationäre Schwingung übergehen wie mit einem Zweipunktregler. Dagegen kann mit Hilfe des Integrators wegen seines kontinuierlichen Ausgangsgrößenbereiches genau der passende Wert für y' bereitgestellt werden.

Dies heißt aber nicht, dass der in Abb. 13.8 gezeigte Kreis in jedem Fall zur Ruhe kommt. Wenn nämlich die Integrationszeit sehr kurz im Verhältnis zu den nachfolgenden Streckenzeitkonstanten ist, wird die Ausgangsgröße y dem Integrator-Ausgang y' nicht schnell genug folgen können. Der Integrator durchläuft dann den *richtigen* Wertebereich, d.h. die Werte, die am Streckenausgang einen im Toleranzbereich liegenden Wert von y hervorrufen würden, während die Ausgangsgröße y aber noch außerhalb des Toleranzbereiches von w liegt. Wenn sie diesen dann endlich erreicht, hat sich y' schon wieder aus dem richtigen Wertebereich entfernt. y wird dem neuen Wert von y' folgen

und daher den Toleranzbereich wieder verlassen. Bei ungünstig gewählten Parametern entstehen also auch hier Schwingungen, die nicht abklingen.

Für die Kombination aus Dreipunkt-Regler und Integrator spricht aber, dass ein Integrator bei technischen Systemen relativ häufig am Anfang der Strecke vorkommt, so dass für diese Kombination nicht extra ein Integrator in die Strecke eingefügt werden muss. Ein Beispiel ist die Druckregelung in einem Kessel mit einem Motorventil, das den Ablauf aus diesem Kessel regelt. Das Schließen des Ventils hat also einen Druckanstieg im Kessel zur Folge, und das Öffnen einen Druckabfall. Der Ventilmotor wird über einen Wahlschalter mit den drei Möglichkeiten *auf*, *zu* und *stop* angesteuert. Bei *auf* dreht sich der Motor in die eine Richtung und der Ventil-Öffnungsquerschnitt vergrößert sich, bei *zu* erfolgt eine Drehung in die andere Richtung und der Querschnitt verkleinert sich. Der Wahlschalter bildet demnach das Dreipunkt-Glied, das Übertragungsverhalten zum Ventilöffnungsquerschnitt lässt sich durch einen Integrator beschreiben, und das dynamische Verhalten des Kessels durch das Verzögerungsglied höherer Ordnung.

13.2.4 Vorzeitiges Umschalten und Sliding Mode

Wie die obigen Beispiele gezeigt haben, führt ein System, das einen Zweipunktregler enthält, immer Schwingungen aus. Enthält der Zweipunktregler Hysterese, kann das System durch das verzögerte Umschalten sogar instabil werden. Die Überlegung liegt deshalb nahe, durch vorzeitiges Umschalten das Systemverhalten zu verbessern. Die Umschaltgerade eines idealen Zweipunktgliedes müsste also in der oberen Hälfte der Zustandsebene nach links und in der unteren Hälfte nach rechts verschoben werden. Ähnliche Auswirkungen hat auch eine Verdrehung der Schaltgeraden in positiver Richtung. Diese Verdrehung lässt sich für das System in Abb. 13.2 bzw. 13.3 beispielsweise erzielen, wenn man die Gleichung $e = 0$ bzw. $l = w$ für die Schaltgerade in der Zustandsebene durch $l = w - kv$ mit $k > 0$ ersetzt. Für die Stellgröße folgt die Definitionsgleichung:

$$u = \begin{cases} 1 & : 0 < w - kv - l \\ -1 & : 0 > w - kv - l \end{cases} \quad (13.3)$$

Dieses Verhalten lässt sich offenbar erzielen, wenn man als Eingangsgröße des Zweipunktgliedes statt e bzw. $w - l$ gerade $w - kv - l$ wählt. Man erhält dann die in Abb. 13.9 gezeichnete Struktur mit einer zusätzlichen Rückführung. Anhand der Zustandskurve ist deutlich zu sehen, dass sich das Stabilitätsverhalten des Systems verbessert hat. Die Amplituden der Schwingung nehmen immer weiter ab, bis schließlich der Sollwert $w = l$ erreicht ist.

Betrachtet man das Einschwingverhalten aber etwas genauer, so stellt man fest, dass das System nicht ganz so ideal ist, wie es auf den ersten Blick erscheint (Abb. 13.10). Die Zustandskurve nähert sich beispielsweise irgendwann von rechts mit $u = -1$ der geneigten Schaltgeraden. Bei Erreichen der

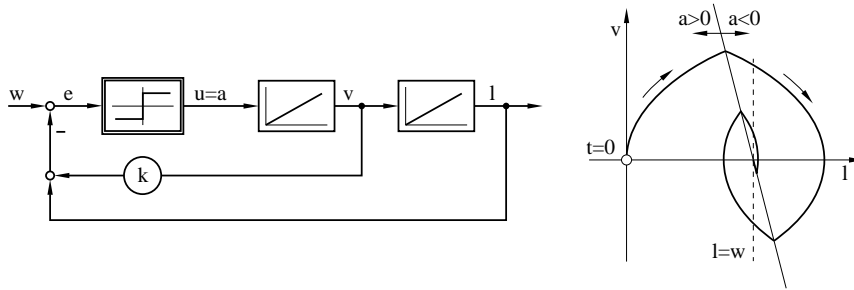


Abb. 13.9. Zweipunktregler mit zusätzlicher Rückführung

Schaltgeraden wird auf $u = 1$ umgeschaltet. Der Systemzustand sollte sich jetzt eigentlich nach links von der Schaltgeraden entfernen. Da die Neigung der Zustandskurve aber größer ist als die Neigung der Schaltgeraden, entfernt sich das System nach rechts. Da in diesem Bereich jedoch die Bedingung für $u = -1$ gilt, wird sofort wieder umgeschaltet, und das System nähert sich wiederum der Schaltgeraden. Auf diese Art und Weise gleitet das System bei einer theoretisch unendlich hohen Umschaltfrequenz in den Endzustand hinein. Ein solches Verhalten wird als *sliding mode* bezeichnet. Eine unendlich hohe Umschaltfrequenz kann es dabei in der Realität natürlich nicht geben, weshalb die vorangegangene Erklärung auch eher als Erklärungsansatz denn als Beweis zu verstehen ist. Dennoch ist bei einer solchen Anordnung in der Tat eine außerordentlich hohe Schaltfrequenz kurz vor Erreichen des Sollwertes zu beobachten. Eine exakte Herleitung für einen Sliding Mode-Regler findet sich in Kapitel 14.1.

Bei hysteresebehafteten Zweipunktreglern weicht das Systemverhalten durch die jeweils verzögerten Umschaltungen zwar etwas vom hier beschriebenen Idealzustand ab, ist im Prinzip aber das gleiche. Die Verwendung eines Dreipunktreglers bringt den Vorteil mit sich, dass der Regler abschaltet, sobald sich die Ausgangsgröße ausreichend nahe am Sollwert befindet. So können die hochfrequenten Umschaltvorgänge in der Endphase vermieden werden.

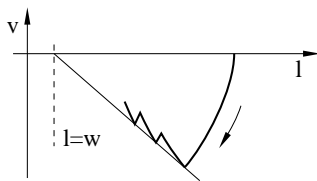


Abb. 13.10. Sliding mode mit einem schaltenden Übertragungsglied

13.2.5 Zeitoptimale Regelung

Bisher war es immer so, dass ein System mit einem schaltenden Übertragungsglied bei Vorgabe eines neuen Sollwertes zunächst über das Ziel hinausgeschossen ist. Erst dann näherte es sich nach mehreren Umschaltvorgängen - wenn überhaupt - dem gewünschten Endzustand. Dieses Verhalten soll für einen doppelten Integrator, d.h. einen beschleunigten Körper kurz analysiert werden. Der Anfangszustand sei $(l, v) = (0, 0)$, und der Endzustand sei $(w, 0)$. Um den Endzustand zu erreichen, muss der Körper zunächst beschleunigt werden. Ein Hinausschießen über das Ziel bedeutet demnach, dass die anfängliche Beschleunigungsphase zu lange gedauert hat und es nicht mehr möglich war, den Körper bis zum Erreichen des Zielpunktes abzubremesen. Ein verbessertes Regelverhalten ergibt sich demnach auf jeden Fall, wenn man rechtzeitig mit dem Abbremsen beginnt. Und ein zeitoptimales Verhalten liegt vor, wenn der Körper so lange wie möglich beschleunigt und dann im letztmöglichen Augenblick mit dem Abbremsen begonnen wird.

Ein Blick auf die Zustandsebene zeigt, welches Vorgehen dazu notwendig ist (Abb. 13.11). Zunächst wird eine Schaltkurve berechnet. Dies ist die Zustandskurve, auf der das System für $v > 0$ bei maximal möglicher negativer Beschleunigung exakt in den Zielpunkt überführt wird (bzw. für $v < 0$ bei maximal möglicher positiver Beschleunigung). Befindet sich der Systemzustand unterhalb dieser Schaltkurve (Punkt 1), so kann das System zunächst noch so lange positiv beschleunigt werden, bis die Schaltkurve erreicht wird. Dann wird auf maximal mögliche negative Beschleunigung umgeschaltet, und das System bewegt sich auf der Schaltkurve exakt in den Zielpunkt. Liegt der Systemzustand dagegen oberhalb der Schaltkurve (Punkt 2), so bedeutet dies, dass das System nicht mehr so abgebremst werden kann, dass der Zielpunkt noch erreicht wird. Stattdessen befindet sich das System nach dem Abbremsen im Zustand 3. Um von dort aus in den Zielzustand zu gelangen, ist es zunächst weiterhin in negativer Richtung zu beschleunigen. Die Geschwindigkeit wird nun negativ, bis der untere Ast der Schaltkurve erreicht wird. Von dort aus kann das System dann durch Abbremsen mit maximaler positiver Beschleunigung in den Zielzustand überführt werden.

Eine Regelungsstruktur, die solche Verstellvorgänge ermöglicht, wird *zeitoptimale Regelung* genannt und ist in Abb. 13.12 gezeichnet. Im Regler 1 wird zunächst der Wert der Schaltkurve v_S für die jeweilige Regeldifferenz e berechnet. Da l in diese Regeldifferenz mit negativem Vorzeichen eingeht, erscheint die Schaltkurve im zugehörigen Block gerade seitenverkehrt. Dies ist auch anschaulich leicht zu erklären. Beispielsweise befindet man sich für eine positive Regeldifferenz, wenn also der Sollwert größer als der Istwert ist, in Abb. 13.11 links vom Zielpunkt. Daher muss in diesem Fall der zugehörige Wert der Schaltkurve positiv sein, was die im Regler 1 eingezeichnete Schaltkurve auch widerspiegelt.

Der so berechnete Wert der Schaltkurve v_S wird dann mit der tatsächlichen Geschwindigkeit v verglichen. Ist die Differenz positiv, so befindet sich

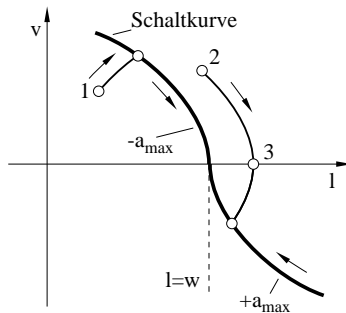


Abb. 13.11. Zeitoptimaler Verstellvorgang

das System unterhalb der Schaltkurve. Entsprechend erzeugt der Zweipunktregler 2 die maximal mögliche positive Stellgröße. Sobald die Differenz $v_S - v$ kleiner als Null wird, schaltet der Regler auf die negative Stellgröße um, und das System bewegt sich auf der Schaltkurve in den Zielpunkt.

Im Zielpunkt ist das Verhalten dann allerdings undefiniert. Da der Regler 2 immer eine von Null verschiedene Ausgangsgröße liefert, kann das System nicht zur Ruhe kommen. Es wird mit theoretisch unendlich hoher Schaltfrequenz um den Zielpunkt herum schwingen. Hier sollte die Möglichkeit einer Abschaltung vorgesehen werden, beispielsweise, indem man das Zweipunktglied durch ein Dreipunktglied ersetzt.

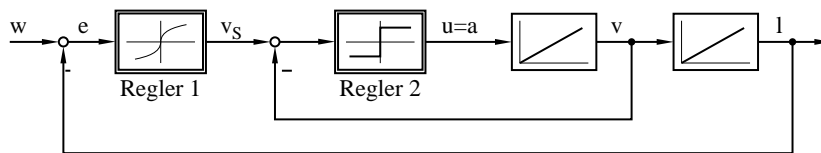


Abb. 13.12. Zeitoptimale Regelung

Anhand der Variationsrechnung lässt sich beweisen, dass die durch diese Reglerstruktur erzeugten Regelvorgänge immer zeitoptimale Regelvorgänge sind. Dies ist auch anschaulich sofort klar, denn die Strecke wird zunächst, abhängig vom Anfangszustand, mit maximaler positiver oder negativer Kraft beschleunigt und dann im letztmöglichen Augenblick mit der entgegengerichteten Kraft abgebremst. Ganz offensichtlich kann man ein solches Verhalten nur mit einem schaltenden Regler erzeugen.

Ist der Regler hysteresebehaftet, so kann das System wegen der verspäteten Umschaltung nicht exakt auf der Schaltkurve in den Zielpunkt laufen. Stattdessen bewegt es sich neben der Schaltkurve und schießt deshalb etwas über das Ziel hinaus. In der Umgebung des Zielpunktes stellen sich dann Schwingungen mit kleiner Amplitude und sehr hoher Frequenz ein.

In der Praxis ergibt sich außerdem noch das Problem, dass man die Strecke zur Berechnung der Schaltkurve genau kennen muss, was meistens

nicht gegeben ist. Aber auch bei nicht exakt berechneter Schaltkurve ergibt sich noch ein recht gutes Regelverhalten.

Eine zeitoptimale Regelung ist auch für Systeme höherer Ordnung möglich. In einem System dritter Ordnung ist beispielsweise die Schaltkurve durch eine Schaltebene im Zustandsraum zu ersetzen. Durch eine maximale positive oder negative Ausgangsgröße u des Reglers wird zunächst diese Schaltebene erreicht. Dort muss das Vorzeichen von u gewechselt werden. Das System bewegt sich dann im Zustandsraum auf der Schaltebene bis zu einer Schaltkurve, die in der Ebene verläuft. Dort wird dann erneut das Vorzeichen von u gewechselt, und das System strebt in den Endzustand. Es lässt sich zeigen, dass für ein System der Ordnung n , das keine Pole mit positivem Realteil aufweist, genau $n - 1$ Vorzeichenwechsel der Stellgröße für einen Regelvorgang erforderlich sind.

13.2.6 Pulsweitenmodulation

Zum Abschluss soll auf das für die Praxis äußerst wichtige Verfahren der Pulsweitenmodulation (PWM) eingegangen werden. Der Nachteil schaltender Übertragungsglieder besteht darin, dass ihre Ausgangsgröße nur wenige, diskrete Werte annehmen kann. Damit sind sie zunächst als Stellglieder für eine hochwertige Regelung nicht zu gebrauchen. Andererseits ist man wegen ihres geringen Preises aber trotzdem daran interessiert, sie innerhalb einer Regelung einzusetzen. Hier stellt die PWM ein geeignetes Verfahren dar, um einem Schalter durch intelligente Ansteuerung ein quasi-kontinuierliches Verhalten aufzuprägen. Bei der PWM wird ein schaltendes Übertragungsglied nach einem speziellen Schaltmuster hochfrequent umgeschaltet, so dass sich an seinem Ausgang eine hochfrequente Rechteckschwingung mit variabler Pulsweite einstellt. Gibt man diese Rechteckschwingung als Stellgröße auf eine Tiefpassstrecke, so werden die hochfrequenten Anteile aus dem Signal herausgefiltert. Die Ausgangsgröße der Strecke ist demnach nur vom Mittelwert der Rechteckschwingung abhängig.

Damit kann man den Mittelwert näherungsweise als Stellgröße ansehen. Da dieser Mittelwert andererseits aber in Abhängigkeit vom Schaltmuster stetig veränderlich ist, ist es nur eine Frage des Schaltmusters, um dem schaltenden Übertragungsglied die Eigenschaften eines linearen Reglers aufzuprägen. Ein sehr anschauliches Verfahren, ein solches Schaltmuster und damit einen quasilinearen Regler zu erzeugen, bildet die *Linearisierung durch eine Rückführung* (Abb. 13.13).

Der aus Zweipunktglied und Rückföhrfunktion $G_R(s)$ bestehende, interne Kreis wird unter der Voraussetzung, dass er stabil ist, Schwingungen ausführen. Und zwar führt u eine Rechteckschwingung aus, und y_R schwingt um die Eingangsgröße e . Je kleiner die Hysteresebreite des Zweipunktgliedes ist, desto hochfrequenter ist die Rechteckschwingung und desto weniger entfernt sich auch y_R von e . Wenn weiterhin die Strecke $G(s)$ Tiefpasseigenschaften besitzt, wird nur der Mittelwert der Rechteckschwingung am Ausgang y

wirksam. Das gleiche gilt auch für die Auswirkungen der Stellgröße auf die interne Rückkopplungsgröße y_R . Man kann sich daher auf eine Betrachtung der Mittelwerte beschränken. Es gilt $\bar{y}_R(s) = G_R(s)\bar{u}(s)$ und bei ausreichend kleiner Hysteresebreite auch $e \approx \bar{y}_R$. Daraus folgt

$$\bar{u}(s) \approx \frac{1}{G_R(s)} e(s) \quad (13.4)$$

d.h. hinsichtlich der Mittelwerte entspricht das gesamte Übertragungsverhalten des Reglers mit Rückführung in etwa dem Kehrwert der internen Übertragungsfunktion. Je nach Wahl dieser Funktion lassen sich so näherungsweise die verschiedensten linearen Regler realisieren.

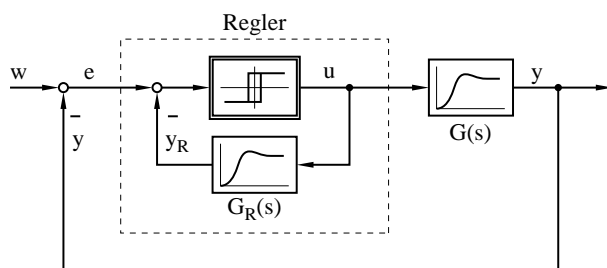


Abb. 13.13. Linearisierung durch interne Rückführung

Neben der Linearisierung durch eine Rückführung gibt es noch viele andere Verfahren zur Pulsweitenmodulation. Häufig werden optimierte Schaltmuster in Tabellen abgelegt und dann je nach zu erzeugendem Mittelwert \bar{u} ausgelesen (vgl. [?]). Die gesamte Pulsweitenmodulation wird dabei von einem einzigen IC übernommen. Insbesondere im Bereich der elektrischen Antriebstechnik ist die Pulsweitenmodulation von großer Bedeutung.

Damit ist die Darstellung der verschiedenen Aspekte schaltender Übertragungsglieder abgeschlossen. Anhand dieser Übertragungsglieder sollten die wichtigsten nichtlinearen Effekte anschaulich erläutert und ein gewisses Grundverständnis für die Problematik nichtlinearer Regelkreise vermittelt werden. Von zentraler Bedeutung für den Regelungstechniker ist dabei die Stabilitätsfrage. Um ein nichtlineares System hinsichtlich seiner Stabilität analysieren zu können, ist es aber erforderlich, dass man von der bisher praktizierten, eher intuitiven Sichtweise zu einer exakten Formulierung des Problems übergeht. Die folgenden Definitionen und Sätze sind dabei nur für zeitkontinuierliche Systeme angegeben, gelten aber in analoger Weise auch für zeitdiskrete Systeme. Anschließend werden dann die für nichtlineare Systeme wichtigsten Stabilitätskriterien vorgestellt.

13.3 Definition der Stabilität bei nichtlinearen Systemen

13.3.1 Ruhelage

Um den Begriff der Stabilität für nichtlineare Systeme exakt definieren zu können, muss zunächst auf den Begriff der *Ruhelage* eingegangen werden (vgl. [13, 14, 17]):

Definition 13.1 *Ein dynamisches System befindet sich für einen gegebenen konstanten Eingangsvektor \mathbf{u}_0 genau dann in der durch den Zustandsvektor \mathbf{x}_R bezeichneten Ruhelage, wenn sich die Zustandsgrößen nicht mehr verändern, d.h. wenn gilt*

$$\mathbf{0} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}_R, \mathbf{u}_0) \quad (13.5)$$

Dabei ist die Festlegung eines konstanten Eingangsvektors notwendig, da das System sonst offensichtlich nie zur Ruhe kommen könnte. Bei einem linearen System ergeben sich die Ruhelagen aus

$$\mathbf{0} = \dot{\mathbf{x}}_R = \mathbf{A}\mathbf{x}_R + \mathbf{B}\mathbf{u}_0 \quad (13.6)$$

Für $|\mathbf{A}| \neq 0$ ergibt sich genau eine Lösung bzw. Ruhelage $\mathbf{x}_R = -\mathbf{A}^{-1}\mathbf{B}\mathbf{u}_0$. Andernfalls treten keine oder unendlich viele Lösungen auf. Ein Beispiel ist ein einfacher Integrator, der sich durch die Zustandsgleichung

$$\dot{x} = 0x + 1u \quad (13.7)$$

darstellen lässt. Offenbar ist $|\mathbf{A}| = 0$. Für $u \neq 0$ gibt es keine Ruhelage, während die Gleichung für $u = 0$ unendlich viele Lösungen besitzt. Dieses Ergebnis ist einsichtig, wenn man sich klarmacht, dass ein Integrator für eine von Null verschiedene Eingangsgröße immer weiter auf- oder abintegriert, während er für $u = 0$ an der Stelle stehenbleibt, wo er sich gerade befindet.

Während ein lineares System also entweder keine, eine oder unendlich viele Ruhelagen besitzt, können bei einem nichtlinearen System auch endlich viele, und zwar mehr als eine Ruhelage auftreten. Ein Beispiel ist das Pendel aus Abb. 2.7. Falls der Körper mittels einer starren Stange aufgehängt ist, existieren offensichtlich Ruhelagen für $\alpha = 0$ und $\alpha = \pi$.

13.3.2 Stabilitätsdefinition nach Ljapunov

Dabei existiert für beide Ruhelagen des Pendels ein wesentlicher qualitativer Unterschied, der mit Hilfe der Stabilitätsdefinition nach *Ljapunov* [36] präzise angegeben werden kann:

Definition 13.2 *Eine Ruhelage \mathbf{x}_R heißt genau dann stabil für eine gegebene konstante Eingangsgröße \mathbf{u}_0 , wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $|\mathbf{x}(0) - \mathbf{x}_R| < \delta$ die Bedingung $|\mathbf{x}(t) - \mathbf{x}_R| < \varepsilon$ mit $t \geq 0$ erfüllt ist.*

Eine Ruhelage heißt also genau dann stabil, wenn der Zustand $\mathbf{x}(t)$ des Systems für alle $t > 0$ in einer beliebig engen Umgebung (ε) der Ruhelage bleibt, sofern der Anfangszustand ausreichend nahe (δ) bei der Ruhelage liegt (Abb. 13.14).

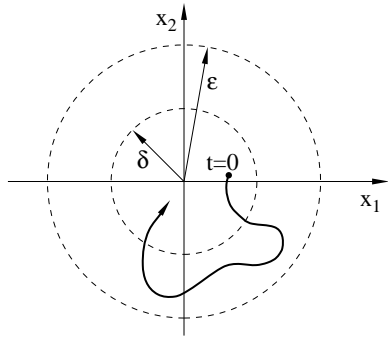


Abb. 13.14. Zur Stabilitätsdefinition nach Ljapunov

Laut dieser Definition ist die obere Ruhelage des Pendels instabil, während die untere Ruhelage stabil ist. Lenkt man beispielsweise das Pendel etwas aus der unteren Ruhelage aus und betrachtet diese Stellung als Startzustand, so wird das Pendel zwar schwingen, sich aber nie weiter von der Ruhelage entfernen als beim Startzustand. Hier existiert also zu jeder beliebigen ε -Umgebung, die für $t > 0$ nicht mehr verlassen werden soll, gerade ein Abstand $\delta = \varepsilon$, in dem der Anfangszustand liegen muss, um diese Bedingung einzuhalten.

Dies ist bei der oberen Ruhelage offenbar nicht der Fall. Angenommen, es ist gefordert, dass eine ε -Umgebung beispielsweise von einigen Winkelgraden um die obere Ruhelage nicht mehr verlassen werden darf. Der einzige Anfangszustand, für den diese Bedingung erfüllt ist, ist die Ruhelage selber. Falls der Anfangszustand nur ganz leicht von der Ruhelage abweicht, kippt das Pendel nach unten, und die geforderte Umgebung wird verlassen. Andererseits ist aber in der Definition gefordert, dass man zu jedem beliebigen ε eine δ -Umgebung für den Startzustand mit $\delta > 0$ angeben können muss, damit die Ruhelage stabil ist. Da dies offenbar für die obere Ruhelage nicht erfüllt wird, ist sie instabil.

Ein anderes anschauliches Beispiel ist das ideale Zweipunktglied mit doppeltem Integrator (Abb. 13.3). Das System führt um die Ruhelage $(l, v) = (w, 0)$ eine Dauerschwingung aus, deren Amplitude vom Anfangszustand abhängig ist. Verlangt man hier, dass das System für $t > 0$ innerhalb einer ganz bestimmten ε -Umgebung um die Ruhelage bleibt, so muss man nur den Anfangszustand entsprechend wählen. Daher ist dieses System stabil im Sinne von Ljapunov.

Man muss sich aber darüber im klaren sein, dass die Stabilität nach Ljapunov nur gewährleistet, dass eine vorgegebene Umgebung um die Ruhelage

nicht mehr verlassen wird. Dies ist in vielen Anwendungsfällen jedoch nicht ausreichend. Dort wird darüber hinaus auch verlangt, dass eine vorgegebene Ruhelage früher oder später tatsächlich erreicht wird. Diese Forderung führt auf den Begriff der *asymptotischen Stabilität*:

Definition 13.3 Eine Ruhelage \mathbf{x}_R heißt *asymptotisch stabil*, wenn sie für eine konstante Anregung \mathbf{u}_0 stabil ist und außerdem eine β -Umgebung besitzt mit $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}_R$ für $|\mathbf{x}(0) - \mathbf{x}_R| < \beta$, d.h. eine Umgebung, aus der alle Zustände in die Ruhelage streben. Die Gesamtheit aller Punkte des Zustandsraumes, aus denen die Trajektorien gegen \mathbf{x}_R streben, heißt *Einzugsbereich* der Ruhelage. Umfasst der Einzugsbereich alle Anfangszustände, die unter gegebenen, technischen Beschränkungen auftreten können, so heißt die Ruhelage *asymptotisch stabil im Großen*. Umfasst der Einzugsbereich den gesamten Zustandsraum, so heißt die Ruhelage *global asymptotisch stabil*.

Als anschauliches Beispiel kann wieder das Pendel dienen, und zwar seine untere Ruhelage. Wenn ein ideales Pendel vorliegt und als Anfangszustand eine Auslenkung vorliegt, so schwingt das Pendel immer weiter und kommt nie zur Ruhe. Die Ruhelage ist stabil nach Ljapunov, aber nicht asymptotisch stabil. Berücksichtigt man dagegen beispielsweise den Luftwiderstand, so nimmt die Amplitude der Schwingung immer weiter ab und die Ruhelage wird - wenn auch theoretisch nach unendlich langer Zeit - erreicht. Damit ist die Ruhelage asymptotisch stabil. Globale asymptotische Stabilität liegt aber nicht vor, denn es existiert genau ein Punkt im Zustandsraum, aus dem keine Trajektorie in die untere Ruhelage verläuft, und zwar die obere Ruhelage. Setzt man jedoch voraus, dass das Pendel an einer Decke aufgehängt ist und die obere Ruhelage damit sowieso nie erreicht werden kann, so kann man das System als asymptotisch stabil im Großen bezeichnen.

Den Ljapunovschen Stabilitätsbegriff kann man auch auf zeitvariante Systeme anwenden. Da sich hier das System aber mit der Zeit verändern kann, muss die obige Definition nicht nur für einen einzigen Anfangszeitpunkt $t = 0$ sondern für alle beliebigen Anfangszeitpunkte erfüllt sein. Damit ist δ möglicherweise nicht nur eine Funktion von ε , sondern auch von der Zeit t . Falls δ aber bei einem zeitvarianten System trotz der Zeitvarianz weiterhin nur eine Funktion von ε ist, spricht man von *gleichmäßiger Stabilität*.

13.3.3 Stabilität von Trajektorien

Gegenstand der bisherigen Stabilitätsbetrachtungen waren die Ruhelagen. Die vorgestellten Stabilitätsbegriffe lassen sich aber auch auf Trajektorien anwenden. In der jeweiligen Definition ist dann lediglich die Ruhelage durch eine Trajektorie zu ersetzen. Ebenso wie eine Ruhelage kann auch eine Trajektorie instabil, stabil oder asymptotisch stabil sein.

Als Beispiel sei eine Schwingung betrachtet, wie sie bei einem doppelten Integrator mit idealem Zweipunktglied auftritt (Abb. 13.3). Bei dieser

Schwingung hängen Amplitude und Frequenz, d.h. der Verlauf der Schwingung, vom jeweiligen Anfangszustand ab. Ein veränderter Anfangszustand führt auf einen anderen Zyklus. Liegt der veränderte Anfangszustand beispielsweise etwas rechts vom ursprünglichen Anfangszustand in der Zustandsebene, so bedeutet dies eine kleinere Anfangsauslenkung von der Ruhelage und daher auch eine kleinere Schwingungsamplitude. Es ergibt sich eine ähnliche Trajektorie wie im ursprünglichen Fall, allerdings näher zur Ruhelage als die erste Trajektorie. Offensichtlich lässt sich zu jeder ε -Umgebung um die ursprüngliche Trajektorie auch eine δ -Umgebung angeben, in der ein Anfangszustand liegen muss, damit die daraus resultierende Trajektorie in der ε -Umgebung der ursprünglichen Schwingungstrajektorie bleibt. Dies bedeutet aber gerade Stabilität der ursprünglichen Schwingung im Ljapunovschen Sinne. Eine solche Schwingung bezeichnet man als Dauerschwingung.

Asymptotische Stabilität liegt aber nicht vor, denn die aus einem veränderten Anfangszustand resultierende Schwingung wird nie in die ursprünglich vorgegebene Schwingung übergehen. Aber nur wenn dies gilt, kann man von asymptotischer Stabilität sprechen. Ein Beispiel für diesen Fall liegt beim Zweipunktglied mit Hysterese und IT_1 -Glied vor (Abb. 13.7). Aus jedem beliebigen Anfangszustand geht die Trajektorie früher oder später in die Trajektorie der gegebenen Schwingung über. Eine solche Schwingung mit asymptotischem Einschwingverhalten bezeichnet man als Grenzzzyklus. Damit ist der Unterschied zwischen Dauerschwingung und Grenzzzyklus mit Hilfe des Ljapunovschen Stabilitätsbegriffs noch einmal präzisiert worden.

Dabei müssen Grenzzyklen nicht unbedingt asymptotisch stabil, sondern können auch instabil sein. Ein instabiler Grenzzyklus ist dadurch definiert, dass sich die von einem dem Grenzzyklus benachbarten Anfangszustand ausgehende Trajektorie von der Trajektorie des Grenzzykus entfernt. Ein Grenzzyklus kann sogar stabil und instabil zugleich sein. Bei einem System zweiter Ordnung unterteilt der Zyklus die Zustandsebene in zwei Gebiete, ein inneres und ein äußeres. Nun kann es vorkommen, dass alle Trajektorien im Innengebiet zum Grenzzyklus hinstreben, während alle Trajektorien außerhalb von ihm wegstreben. Ein solcher Grenzzyklus ist dann nach innen stabil und nach außen instabil. Dies ist allerdings nur eine rein theoretische Konstruktion, denn auch ein nur einseitig instabiler Grenzzyklus kann nicht von langer Lebensdauer sein. Eine kleine Störung reicht aus, damit das System den Zyklus nach außen verlässt und nie wieder zu ihm zurückkehrt. Dennoch sollte man sich über die Möglichkeit solcher Grenzzyklen mit unterschiedlichem Stabilitätsverhalten im klaren sein, da sie beim später behandelten Verfahren der Beschreibungsfunktion noch einmal auftauchen werden.

Zu beachten ist, dass, um die Stabilität einer Schwingung zu untersuchen, der Ljapunovsche Stabilitätsbegriff nur auf die zugehörigen Trajektorien im Zustandsraum angewendet wurde. Würde man den Verlauf der Zustandsgrößen über die Zeit betrachten, ergäbe sich ein ganz anderes Bild. Als Beispiel soll wieder das Pendel dienen. Unter Vernachlässigung des Luftwi-

derstandes führt es eine vom Anfangszustand abhängige Dauerschwingung aus. Auf eine bestimmte Anfangsauslenkung folgt eine Dauerschwingung mit einer ganz bestimmten Frequenz und Amplitude, während auf eine etwas größere Anfangsauslenkung eine Dauerschwingung mit etwas kleinerer Frequenz und etwas größerer Amplitude folgt. Zeichnet man die Trajektorien der beiden Schwingungen, so werden sie eine ähnliche Form aufweisen und in unmittelbarer Nachbarschaft zueinander verlaufen, wobei die Trajektorie der Schwingung mit der kleineren Amplitude innerhalb der anderen Trajektorie verläuft. Daraus folgt die einfache Stabilität der Schwingung nach Ljapunov. Zeichnet man aber den Verlauf der Position des Pendels als Funktion der Zeit für beide Fälle auf, so werden sich die Kurven wegen der unterschiedlichen Frequenzen der Schwingungen immer weiter auseinander bewegen. Würde man die Stabilität anhand dieser Kurven definieren, wäre das System nicht stabil.

In [41] und [48] wird deshalb eine Schwingung nur dann als asymptotisch stabil bezeichnet, wenn die Ljapunovsche Stabilitätsdefinition auf den zeitlichen Verlauf der Zustandsgrößen zutrifft. Wenn dagegen nur Stabilität hinsichtlich der Trajektorien vorliegt, so wird von *orbitaler Stabilität* gesprochen. In der Praxis ist dieser Unterschied allerdings nicht relevant, weil im allgemeinen nicht der explizite zeitliche Verlauf der Zustandsgrößen sondern nur die prinzipielle Form einer Schwingung interessiert. Deshalb soll hier die Stabilität einer Schwingung weiterhin anhand der Trajektorien beurteilt werden.

13.3.4 Stabilität von linearen Systemen

Nachdem nun die Ljapunovsche Stabilitätsdefinition für nichtlineare Systeme ausführlich erörtert wurde, soll noch einmal die Verbindung zu den linearen Systemen hergestellt werden. Ein lineares System ist nach Def. 7.2 genau dann asymptotisch stabil, wenn seine Zustandsgrößen ohne äußere Anregung aus jedem beliebigen Anfangszustand gegen Null streben. Die Frage ist nun, wie man diese Definition mit Def. 13.2 und 13.3 in Einklang bringt.

Zunächst fällt auf, dass in Def. 7.2 von der Stabilität des Systems die Rede ist, während sich 13.2 und 13.3 nur auf die Stabilität einer einzigen Ruhelage beziehen. Zur Erklärung sei ein lineares System mit konstanter Anregung \mathbf{u}_0 betrachtet:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}_0 \quad (13.8)$$

Eine sich bei dieser Anregung einstellende Ruhelage \mathbf{x}_R erfüllt die Differentialgleichung

$$\dot{\mathbf{x}}_R = \mathbf{A}\mathbf{x}_R + \mathbf{B}\mathbf{u}_0 \quad (13.9)$$

Mit $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_R$ ergibt sich nach Subtraktion der beiden Gleichungen

$$\dot{\Delta\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} \quad (13.10)$$

Nun gilt aber doch, dass für die Stabilitätsanalyse der Ruhelage ausschließlich der Abstand des Zustandsvektors zur Ruhelage relevant ist. Die Untersuchung kann damit anhand von Gl. (13.10) durchgeführt werden. In dieser Gleichung tauchen aber sowohl die Anregung als auch die Ruhelage selbst gar nicht mehr auf. Das Ergebnis, das man erhält, wird daher für alle Ruhelagen und alle Anregungen das gleiche sein, d.h. wenn eine Ruhelage für eine Anregung asymptotisch stabil ist, so sind alle Ruhelagen für alle Anregungen asymptotisch stabil. Man spricht aus dem Grund bei einem linearen System nicht von der Stabilität einer Ruhelage, sondern von der Stabilität des Systems. Dies ist ein ganz wesentlicher Unterschied zu einem nichtlinearen System, bei dem verschiedene Ruhelagen ein völlig unterschiedliches Stabilitätsverhalten aufweisen können.

Wenn daher eine einzige Ruhelage eines linearen Systems für eine Anregung asymptotisch stabil nach Definition 13.2 und 13.3 ist, so gilt dies für alle Ruhelagen und insbesondere auch für die Ruhelage $\mathbf{x} = \mathbf{0}$ und $\mathbf{u} = \mathbf{0}$. Damit ist das System aber auch nach Definition 7.2 asymptotisch stabil. Analog gilt die Umkehrung für Instabilität. Aus Def. 13.2 und 13.3 folgt für lineare Systeme also Def. 7.2.

Für die Herleitung der Äquivalenz der Definitionen ist nun noch zu zeigen, dass bei linearen Systemen aus Def. 7.2 auch die beiden anderen Definitionen folgen. Wegen des gleichen Stabilitätsverhaltens aller Ruhelagen bei einem linearen System ist naheliegend, diesen Nachweis für den einfachsten Fall, nämlich für die Ruhelage $\mathbf{x} = \mathbf{0}$ und $\mathbf{u} = \mathbf{0}$ zu führen und das Ergebnis auf das gesamte System zu erweitern. Aus (13.8) ergibt sich wegen $\mathbf{u} = \mathbf{0}$ zunächst

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad (13.11)$$

Nun sei vorausgesetzt, dass das System nach Def. 7.2 asymptotisch stabil ist, d.h. seine Zustandsgrößen streben ohne äußere Anregung aus jedem beliebigen Anfangszustand gegen Null. Dies ist aber nach Satz 7.3 genau dann der Fall, wenn alle Eigenwerte von \mathbf{A} einen negativen Realteil aufweisen. In dem Fall sind aber auch eventuell auftretende Schwingungen abklingend. Demnach lässt sich für jede ε -Umgebung um die Ruhelage $\mathbf{x} = \mathbf{0}$, die für $t > 0$ nicht verlassen werden soll, eine δ -Umgebung angeben, in der der Anfangszustand liegen muss: $\delta = \varepsilon$. Damit ist die Ruhelage nach Def. 13.2 stabil. Und die asymptotische Stabilität nach Def. 13.3 ist gewährleistet, weil alle Zustandsgrößen gegen die Ruhelage Null streben. Auch hier gilt die Umkehrung für Instabilität analog.

Darüber hinaus ist die Ruhelage $\mathbf{x} = \mathbf{0}$ und $\mathbf{u} = \mathbf{0}$ global asymptotisch stabil, d.h. aus allen Zuständen des Zustandsraumes streben die Trajektorien in diese Ruhelage. Als Beweis ist es ausreichend zu zeigen, dass keine weitere Ruhelage existiert. Dies ist aber der Fall, denn wenn \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist, gilt $|\mathbf{A}| \neq 0$, und Gleichung (13.11) kann für $\dot{\mathbf{x}} = 0$ nur die Lösung $\mathbf{x} = \mathbf{0}$ besitzen.

13.3.5 Steuerbarkeit und Beobachtbarkeit

Zwei weitere wichtige Systemeigenschaften neben der Stabilität sind die Steuer- und Beobachtbarkeit. Wie schon in Kapitel 7.10 angesprochen, sollte man vor dem Entwurf der Regelung sicherstellen, ob man auf das System überhaupt den gewünschten Einfluss nehmen kann, d.h. ob Steuerbarkeit vorliegt. Falls für die Regelung Zustandsgrößen verwendet werden, muss auch sichergestellt sein, dass man diese Zustandsgrößen aus den messbaren Ausgangsgrößen überhaupt berechnen kann. Diese Systemeigenschaft entspricht der Beobachtbarkeit. Zwei Möglichkeiten bieten sich hier für nichtlineare Systeme an. Die eine ist, das System am Arbeitspunkt zu linearisieren und auf das lineare Modell die Steuer- und Beobachtbarkeitskriterien linearer Systeme anzuwenden. Hier tritt aber wieder das Problem auf, dass ein lineares Modell das nichtlineare Systemverhalten nur in einem engen Bereich um den Arbeitspunkt ausreichend gut approximiert und die Aussagen hinsichtlich Steuer- und Beobachtbarkeit dementsprechend auf einen kleinen Bereich des Zustandsraumes beschränkt sind.

Der andere Ansatz ist, die Definitionen und Kriterien für nichtlineare Systeme geeignet abzuändern. So gibt es beispielsweise in [55] Definitionen für Erreichbarkeit und Unterscheidbarkeit von Zuständen. Hinreichende und leicht handzuhabende Kriterien wie die Steuerbarkeits- und Beobachtbarkeitskriterien von Kalman (Satz 7.5 und 7.9) existieren aber nicht für allgemeine nichtlineare Systeme. Nur für spezielle Klassen von nichtlinearen Systemen wie beispielsweise bilineare Systeme existieren vergleichbare Kriterien.

13.3.6 Definition der Ruhelage Null

Im weiteren Verlauf wird der Einfachheit halber immer vorausgesetzt, dass in der betrachteten Ruhelage alle Systemgrößen den Wert Null annehmen. Ist dies nicht der Fall, so muss das System umdefiniert werden. Diese Maßnahme kann man auch so interpretieren, dass man von den tatsächlichen Größen \mathbf{x} zu deren Abweichungen von der Ruhelage $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_R$ übergeht. Der Vektor $\Delta\mathbf{x}$ wird dann als neue Systemgröße definiert und erfüllt gerade die Forderung, dass er in der Ruhelage den Wert $\Delta\mathbf{x} = \mathbf{0}$ annimmt. Es sei ausdrücklich darauf hingewiesen, dass es sich bei diesem Schritt um eine exakte Umdefinition des Systems und nicht um eine Linearisierung am Arbeitspunkt handelt.

Für ein lineares System ist dieser Übergang nicht notwendig. Denn da das Systemverhalten von allen Ruhelagen gleich ist, kann man immer von vornherein ausschließlich die Ruhelage $\mathbf{x} = \mathbf{0}$ betrachten. Relevant ist diese Umdefinition daher nur für nichtlineare Systeme. Als Beispiel zeigt Abb. 13.15 die dafür erforderlichen Schritte für einen aus einem linearen und einem nichtlinearen Teil bestehenden Standardregelkreis. Der nichtlineare Teil sei dabei durch eine von \mathbf{e} und $\dot{\mathbf{e}}$ abhängige Funktion $\mathbf{u} = \mathbf{f}(\mathbf{e}, \dot{\mathbf{e}})$ gegeben.

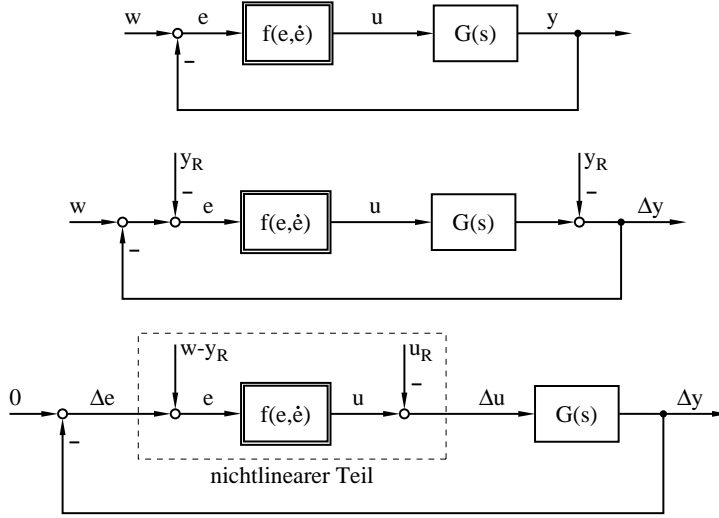


Abb. 13.15. Umformen eines nichtlinearen Standardregelkreises

Die zu betrachtende Ruhelage sei durch die Vektoren \mathbf{w} , \mathbf{u}_R , \mathbf{y}_R und $\mathbf{e}_R = \mathbf{w} - \mathbf{y}_R$ charakterisiert. Der Übergang von den tatsächlichen Größen \mathbf{u} , \mathbf{y} und \mathbf{e} zu ihren Abweichungen von der Ruhelage $\Delta \mathbf{u}$, $\Delta \mathbf{y}$ und $\Delta \mathbf{e}$ geschieht nun folgendermaßen: Im ersten Schritt wird an zwei Stellen des Regelkreises die Größe \mathbf{y}_R subtrahiert. Beide Subtraktionen heben sich in ihrer Wirkung wegen der negativen Rückkopplung gerade auf, d.h. das System wird durch diese Maßnahme nicht verändert.

Im zweiten Schritt soll die Subtraktion von \mathbf{y}_R am Ausgang des linearen Teiles durch die Subtraktion eines Vektors \mathbf{u}_R am Eingang des linearen Teiles ersetzt werden, wobei \mathbf{u}_R gerade der zur Ruhelage gehörende Stellgrößenvektor ist. Das Gleichsignal \mathbf{u}_R kann dabei aus dem Gleichsignal \mathbf{y}_R mit Hilfe des Zusammenhanges $\mathbf{y}_R = \mathbf{G}(s=0)\mathbf{u}_R$ berechnet werden. $\Delta \mathbf{u} = \mathbf{u} - \mathbf{u}_R$ kennzeichnet dann die Abweichung von der Ruhelage. Außerdem werden die Anregungsgrößen \mathbf{w} und \mathbf{y}_R vor dem Eingang des nichtlinearen Teiles zusammengefasst. Insgesamt lässt sich damit ein neues nichtlineares Übertragungsverhalten

$$\Delta \mathbf{u} = \bar{\mathbf{f}}(\Delta \mathbf{e}, \dot{\Delta \mathbf{e}}) = \mathbf{f}(\Delta \mathbf{e} + \mathbf{w} - \mathbf{y}_R, \frac{d}{dt}(\Delta \mathbf{e} + \mathbf{w} - \mathbf{y}_R)) - \mathbf{u}_R \quad (13.12)$$

definieren. Das dadurch entstandene, neue Gesamtsystem mit seinen Systemgrößen $\Delta \mathbf{e}$, $\Delta \mathbf{u}$ und $\Delta \mathbf{y}$ erfüllt die Bedingung, dass alle Systemgrößen in der Ruhelage den Wert Null annehmen. Im Folgenden wird nun immer vorausgesetzt, dass vor der Stabilitätsanalyse eine derartige Umdefinition erfolgt ist. Es wird deshalb immer die Ruhelage Null bei der Anregung Null betrachtet.

13.4 Direkte Methode von Ljapunov

Damit kann nun auf die Frage eingegangen werden, wie denn bei einem gegebenen System die Stabilität einer Ruhelage zu bestimmen ist. Würde man streng nach Def. 13.2 und 13.3 vorgehen, so müsste man für jeden möglichen Anfangszustand die Lösung der nichtlinearen Differentialgleichung für die zu untersuchende Ruhelage ermitteln. Bei unendlich vielen Anfangszuständen ist dies offensichtlich nicht möglich. Es sind daher Kriterien oder Methoden erforderlich, die auch ohne aufwändige Rechnungen eine Stabilitätsaussage für die betreffende Ruhelage zulassen.

Für ein System zweiter Ordnung lässt sich eine Betrachtung in der Zustandsebene durchführen, wie dies bei den schaltenden Übertragungsgliedern bereits gezeigt wurde. Für Systeme höherer Ordnung und auch Mehrgrößensysteme sind dagegen andere Kriterien erforderlich, die im Folgenden vorgestellt werden sollen.

Das erste dieser Kriterien stammt von Ljapunov selbst und wird als *direkte Methode* von Ljapunov bezeichnet. Folgende Idee liegt dem Verfahren zu Grunde: Es wird eine vom Zustandsvektor abhängige, skalare *Ljapunov-Funktion* $V(\mathbf{x})$ definiert, die im Nullpunkt den Wert Null haben muss und ansonsten mit zunehmender Entfernung vom Nullpunkt ansteigt. Man kann V auch als eine Art verallgemeinerten Abstand zur Ruhelage auffassen. Abb. 13.16 zeigt als Beispiel die Höhenlinien einer solchen Funktion in der Zustandsebene.

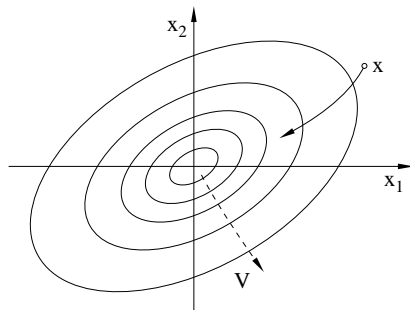


Abb. 13.16. Direkte Methode nach Ljapunov

Weiterhin hat ein Zustandsvektor und mit ihm die Funktion V entsprechend der Zustandsgleichung des Systems einen bestimmten zeitlichen Verlauf. Wenn man nun zeigen kann, dass die zeitliche Ableitung der Funktion V für beliebige Zustandsvektoren \mathbf{x} negativ ist, so bedeutet dies doch, dass die Zustandskurve früher oder später alle Höhenlinien von außen nach innen überschreitet und der Zustandsvektor damit zwangsläufig gegen Null strebt. Das System ist in dem Fall offenbar asymptotisch stabil.

Satz 13.4 Das dynamische System $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ besitze die Ruhelage $\mathbf{x} = \mathbf{0}$. Es gebe eine in der Umgebung der Ruhelage samt ihren partiellen Ableitungen erster Ordnung stetige Funktion $V(\mathbf{x})$, die dort positiv definit ist, d.h. $V(\mathbf{x}) > 0$ für $\mathbf{x} \neq \mathbf{0}$ und $V(\mathbf{x}) = 0$ für $\mathbf{x} = \mathbf{0}$. Weiterhin sei die zeitliche Ableitung

$$\dot{V} = \sum_{i=1}^n \frac{\partial V}{\partial x_i} \dot{x}_i = \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i \quad (13.13)$$

in der Umgebung der Ruhelage negativ definit. Dann ist die Ruhelage asymptotisch stabil und die Umgebung ihr Einzugsbereich.

G sei ein Gebiet innerhalb der Umgebung, in dem $V < c$ gilt (mit $c > 0$) und dessen Rand durch $V = c$ gebildet wird. Wenn G darüber hinaus beschränkt ist und die Ruhelage enthält, so gehört G zum Einzugsbereich der Ruhelage.

Wenn der Einzugsbereich der gesamte Zustandsraum ist und darüber hinaus mit zunehmender Entfernung von der Ruhelage $|\mathbf{x}| = \sqrt{x_1^2 + \dots + x_n^2} \rightarrow \infty$ auch $V(\mathbf{x}) \rightarrow \infty$ gilt, so ist die Ruhelage global asymptotisch stabil.

Falls \dot{V} negativ semidefinit ist ($\dot{V}(\mathbf{x}) \leq 0$), so kann nur die einfache Stabilität gewährleistet werden. Falls aber die Punktmenge, auf der $\dot{V} = 0$ ist, außer $\mathbf{x} = \mathbf{0}$ keine andere Trajektorie enthält, so liegt auch hier asymptotische Stabilität vor.

Der Beweis für diesen Satz findet sich beispielsweise in [30]. Der erste Teil des Satzes bedarf wegen der vorangegangenen Betrachtung keiner weiteren Erklärung, wohl aber die letzten drei Absätze.

Die Überlegungen zum Einzugsbereich der Ruhelage gestalten sich am einfachsten, wenn man sich V anhand von Höhenlinien in einer Zustandsebene gegeben denkt. Zum zweiten Absatz des Satzes zeigt Abb. 13.17 ein Beispiel, in dem die Funktion \dot{V} nicht im gesamten Zustandsraum, sondern nur zwischen den beiden gestrichelten Linien negativ definit und sonst positiv definit ist.

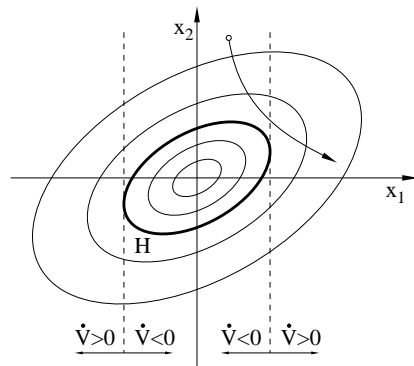


Abb. 13.17. Einzugsbereich einer stabilen Ruhelage

Damit ist eine Zustandskurve möglich, wie sie in der Abbildung eingezeichnet ist. Solange sich die Kurve zwischen den gestrichelten Linien befindet, überschreitet sie die Höhenlinien von außen nach innen, im übrigen Bereich aber von innen nach außen. Der Einzugsbereich der Ruhelage muss damit keinesfalls den gesamten Bereich zwischen den gestrichelten Linien umfassen, in dem $\dot{V} < 0$ gilt. Sicher zum Einzugsbereich gehört nur das Gebiet G innerhalb der Höhenlinie H . Denn diese kann auf keinen Fall von innen nach außen überschritten werden, weil sie vollständig im Gebiet mit $\dot{V} < 0$ verläuft. Offensichtlich muss man als Begrenzung eines Einzugsbereiches daher immer eine geschlossene Höhenlinie angeben. Dies ist aber genau die Bedingung, die im zweiten Absatz des Satzes gefordert wird.

Mit der Vorstellung, dass V durch Höhenlinien gegeben ist, lässt sich auch die Forderung nach dem unendlichen Wachstum von V mit unendlicher Entfernung vom Nullpunkt erklären. Wenn V nämlich nicht mit zunehmender Entfernung immer weiter wachsen würde, so gäbe es Höhenlinien, die bis ins Unendliche reichten und doch nie geschlossen wären. Zwei aufeinander folgende Höhenlinien könnten daher im Unendlichen unendlich weit auseinanderliegen. Wenn dann bei $\dot{V} < 0$ die Zustandskurve die Höhenlinie mit dem größeren Wert von V überschritten hat, so müsste eine unendlich lange Zeit bis zum Überschreiten der Höhenlinie mit dem kleineren Wert vergehen. Der Zustandspunkt würde demnach unendlich lange Zeit zwischen beiden Höhenlinien in möglicherweise unendlicher Entfernung vom Ruhepunkt verweilen, und das System wäre nicht stabil.

Der letzte Absatz des Satzes ist wieder recht einfach zu verstehen. Wenn V und damit der Abstand zum Nullpunkt mit der Zeit nicht kleiner wird, sondern gleich bleibt ($\dot{V} = 0$), so liegt offensichtlich nur einfache Stabilität vor. Wenn aber andererseits die Punkte des Zustandsraumes mit $\dot{V} = 0$ keine zusammenhängenden Trajektorien bilden, so muss das System (sofern es noch nicht den Nullpunkt erreicht hat) immer wieder auch Zustände annehmen, in denen $\dot{V} < 0$ gilt. Damit ist die Funktion V im zeitlichen Verlauf zwar nicht streng monoton, aber doch monoton fallend. Der Nullpunkt wird früher oder später erreicht, und das System ist deshalb asymptotisch stabil.

Mit Hilfe der direkten Methode lässt sich auch die Instabilität einer Ruhelage nachweisen. In völliger Analogie zu Satz 13.4 ist hier die positive Definitheit von \dot{V} zu zeigen. Sowohl für den Nachweis der Instabilität als auch der Stabilität einer Ruhelage existieren für verschiedene Randbedingungen zahlreiche Varianten von Satz 13.4 [17], darunter auch die sogenannten Instabilitätstheoreme von Ljapunov selber. Selbst für zeitvariante Systeme (z.B. [5, 48]) und sogar für Systeme mit äußerer Anregung [30] existieren Theoreme. Die zu Grunde liegende Idee, nämlich die Verwendung einer Ljapunov-Funktion, ist aber in allen Fällen dieselbe, weshalb hier auf nähere Erläuterungen verzichtet wird.

Stattdessen soll noch kurz auf das entscheidende Problem bei der Anwendung der direkten Methode eingegangen werden. Offensichtlich hängen

doch Form und Größe des nachweisbaren Einzugsbereiches einer Ruhelage ganz wesentlich von der gewählten Ljapunov-Funktion V ab. Eine anders gewählte Funktion V kann einen völlig anderen Einzugsbereich liefern. Die Wahl der Ljapunov-Funktion entscheidet sogar darüber, ob überhaupt die Stabilität der Ruhelage nachgewiesen werden kann. Und falls keine geeignete Funktion gefunden wird, so heißt dies nicht, dass die Ruhelage instabil ist, sondern lediglich, dass die Suche erfolglos war. Zum Nachweis der Instabilität müsste, wie oben erwähnt, eine Ljapunov-Funktion gefunden werden, deren Ableitung immer positiv definit ist.

Aus dem Grund sind über die Jahre verschiedene Ansätze entstanden, um die Suche nach einer Ljapunov-Funktion zu systematisieren [5, 13, 14, 16, 17, 48, 53]. Den entscheidenden Makel des Verfahrens, dass nämlich im Falle des Nicht-Auffindens einer Ljapunov-Funktion keine Stabilitätsaussage möglich ist, konnten aber auch sie nicht beseitigen.

Erst in jüngster Zeit konnte dieser Mangel durch den Einsatz von LMI-Algorithmien (vgl. Anhang A.7) für die sehr umfassende Klasse von GS-Systemen (vgl. Kap. 14.2.1) behoben werden. Mit Hilfe dieser Algorithmen ist es nämlich möglich, die generelle Frage nach der Existenz einer Ljapunov-Funktion mit negativ definiter Ableitung für das gegebene System zu beantworten. Und diese ist, wie sich zeigen lässt, gleichbedeutend mit der Frage nach der Stabilität des Systems. Behandelt wird dieser Ansatz in Kapitel 14.2.2.

Hier soll stattdessen noch ein Beispiel für die Anwendung der direkten Methode behandelt werden, und zwar die schwingende Masse aus Abb. 2.1. Dabei macht die Stabilitätsanalyse eines linearen Systems mit Hilfe der direkten Methode eigentlich keinen Sinn, weil eine Untersuchung der Eigenwerte der Systemmatrix wesentlich einfacher wäre. Andererseits ist dieses Beispiel anschaulich und erfordert auch keinen großen Rechenaufwand.

Die Zustandsgleichung des von außen nicht angeregten Systems lautet nach Gleichung (7.18)

$$\begin{bmatrix} \dot{v} \\ \dot{l} \end{bmatrix} = \begin{bmatrix} -\frac{c_r}{m} & -\frac{c_f}{m} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ l \end{bmatrix} \quad (13.14)$$

Der Gesamt-Energieinhalt des Systems ist die Summe aus der in der bewegten Masse enthaltenen kinetischen Energie und der in der Feder gespeicherten potentiellen Energie. Durch die Reibung verliert dieses System Energie, bis die Schwingung schließlich zum Erliegen kommt. Da somit die Energie monoton abnehmend ist, liegt es nahe, den Energieinhalt des Systems als Ljapunov-Funktion zu definieren und auf diese Art und Weise die Stabilität der Ruhelage $(v, l) = (0, 0)$ zu beweisen:

$$V = E = E_{\text{kin}} + E_{\text{pot}} = \frac{1}{2}mv^2 + \int_0^l f_f dx = \frac{1}{2}mv^2 + \int_0^l c_f x dx$$

$$= \frac{1}{2}mv^2 + \frac{1}{2}c_f l^2 \quad (13.15)$$

V ist stetig und stetig differenzierbar. Außerdem ist die Funktion im gesamten Zustandsraum außer im Ursprung $(v, l) = (0, 0)$ positiv und wächst mit $|\mathbf{x}| = |(v, l)^T| \rightarrow \infty$ über alle Maßen. Damit sind die Voraussetzungen aus Satz 13.4 für globale Stabilität erfüllt. Zu untersuchen ist jetzt noch die negative Definitheit von \dot{V} . Die Ableitung von V nach der Zeit ergibt unter Berücksichtigung der Zustandsgleichung

$$\begin{aligned} \dot{V} &= mv\dot{v} + c_f l\dot{l} \\ &= mv\left[-\frac{c_r}{m}v - \frac{c_f}{m}l\right] + c_f lv \\ &= -v^2 c_r \end{aligned} \quad (13.16)$$

Offensichtlich ist diese Funktion negativ semidefinit, da sie nicht nur im Ursprung den Wert Null annimmt, sondern in allen Zuständen mit $v = 0$. Dies ist leicht zu erklären. Ein Energieverlust und damit eine Abnahme von V wird durch Reibung verursacht. Diese tritt genau dann auf, wenn die Geschwindigkeit von Null verschieden ist. In den Punkten maximaler Auslenkung der Feder sind aber die Geschwindigkeit und damit auch die Reibung und \dot{V} gleich Null. Zunächst ist also global nur die einfache Stabilität, nicht aber asymptotische Stabilität gewährleistet. Untersucht man jedoch die Punkte des Zustandsraumes, in denen $\dot{V} = 0$ gilt, so stellt man fest, dass diese (außer im Ursprung) keine zusammenhängende Trajektorie bilden. Ein Zustand $(v = 0, l \neq 0)$ bedeutet, dass die Feder maximal ausgelenkt ist und die Amplitude der Schwingung gerade den Maximalwert erreicht hat. Durch die Feder wird die Masse aber sofort wieder beschleunigt, und das System nimmt einen Zustand mit $v \neq 0$ und $\dot{V} < 0$ an. Insgesamt ist V daher monoton abnehmend und die globale asymptotische Stabilität des Systems gemäß dem vierten Absatz von Satz 13.4 bewiesen.

13.5 Harmonische Balance / Beschreibungsfunktion

13.5.1 Herleitung

Ein völlig anderer Ansatz liegt der Methode der *Beschreibungsfunktion* oder auch Methode der *harmonischen Balance* zu Grunde. Bei diesem Verfahren, das hier nur für Eingrößensysteme betrachtet werden soll, wird zunächst davon ausgegangen, dass die Ausgangsgröße y des Systems um die Ruhelage $y = 0$ eine Schwingung ausführt. Damit schwingen dann natürlich auch die Regelabweichung e und die Stellgröße u . Die Entstehung der Schwingung wird nicht betrachtet. Die Analyse der Schwingung lässt dann Rückschlüsse auf das Stabilitätsverhalten des Systems hinsichtlich der betrachteten Ruhelage zu. Die zu Grunde gelegte Schwingung kann dabei eine Dauerschwingung oder ein Grenzzyklus sein, was im Folgenden nicht unterschieden werden soll.

Voraussetzung ist die Unterteilbarkeit des Regelkreises in einen linearen und einen nichtlinearen Teil wie beim Standardregelkreis gemäß Abb. 13.15. Die gesamte Dynamik des Systems wie z.B. Integratoren, Laufzeitglieder, usw. soll dabei im linearen Teil enthalten sein, während der nichtlineare Teil *momentan wirkend* sein muss:

$$u(t) = f(e, \operatorname{sgn}(\dot{e})) \quad (13.17)$$

Dies bedeutet, dass sich die Ausgangsgröße u des nichtlinearen Teiles im Prinzip aus der momentan anliegenden Eingangsgröße e ohne Kenntnis früherer Werte von e oder u berechnen lässt. So kann man beispielsweise bei Kennliniengliedern direkt aus dem Momentanwert der Eingangsgröße e die Ausgangsgröße $u = f(e)$ berechnen. Sie sind damit momentan wirkend. Als momentan wirkend gelten aber auch die hysteresebehafteten Übertragungsglieder, obwohl dort eine gewisse Kenntnis der Vorgeschichte erforderlich ist, weil man sonst nicht weiß, in welchem Zweig der Hystereseschleife sich das System gerade befindet. Diese Vorgeschichte wird durch den Term $\operatorname{sgn}(\dot{e})$ ausgedrückt.

Darüber hinaus muss die auftretende Kennlinie des nichtlinearen Teiles monoton steigend sein und eine ungerade Funktion darstellen (Nullpunktsymmetrie). Dies ist beispielsweise bei den schaltenden Übertragungsgliedern gegeben. Die Übertragungsfunktion des linearen Teiles muss dagegen ein ausgeprägtes Tiefpassverhalten aufweisen, wobei auf die Bedeutung dieser Eigenschaft im Verlauf der folgenden Herleitung noch näher eingegangen wird. Auch diese Forderung ist in der Praxis in vielen Fällen erfüllt, so dass es für das Verfahren der Beschreibungsfunktion einen großen Anwendungsbereich gibt.

Für die Herleitung geht man davon aus, dass am Ausgang des Systems eine harmonische Schwingung $y(t) = -A \sin(\omega t)$ vorliegt, deren Amplitude A und Frequenz ω bestimmt werden sollen. Da das System vor Anwendung des Verfahrens entsprechend Abb. 13.15 undefiniert wurde und somit die Führungsgröße w gleich Null ist, liegt am Eingang des nichtlinearen Gliedes die Größe $e(t) = A \sin(\omega t)$ an. Dann ergibt sich als Ausgangsgröße des nichtlinearen Gliedes ebenfalls ein periodisches Signal, das sich als Fourierreihe mit der Grundfrequenz ω darstellen lässt und wegen der Nullpunktsymmetrie der nichtlinearen Kennlinie keinen Gleichanteil enthält:

$$u(t) = \sum_{k=1}^{\infty} A_k \cos k\omega t + B_k \sin k\omega t$$

$$\text{mit } A_k = \frac{2}{T} \int_0^T u(t) \cos(k\omega t) dt$$

$$B_k = \frac{2}{T} \int_0^T u(t) \sin(k\omega t) dt$$

$$T = \frac{2\pi}{\omega} \quad (13.18)$$

Dieses Signal bildet wiederum die Eingangsgröße für den linearen Teil. Nach Satz 4.1 erzeugt jede Teilschwingung am Eingang des linearen Teiles eine Ausgangsschwingung mit derselben Frequenz. Wenn nun die Tiefpasswirkung des linearen Teiles ausreichend ausgeprägt ist, so werden aber alle Schwingungen mit einer Frequenz, die größer als die Grundschwingung ω ist, aus dem Signal weitgehend herausgefiltert, und übrig bleibt nur der Grundschwingungsanteil. Die ausreichende Tiefpasswirkung ist dabei eine formal schwer zu beschreibende Eigenschaft. Als Faustregel gilt, dass in der Übertragungsfunktion der Grad des Nennerpolynoms den des Zählerpolynoms um mindestens 2 übersteigen sollte. Aber auch eine Graddifferenz von 1 kann schon ausreichend sein. Auf jeden Fall sollte man am Ende des Verfahrens, wenn die Parameter der Schwingung und damit auch ω berechnet sind, noch einmal überprüfen, ob durch den linearen Teil die höherfrequenten Signalanteile $2\omega, 3\omega, \dots$ tatsächlich ausreichend unterdrückt werden können. Andernfalls ist eine wesentliche Voraussetzung des Verfahrens nicht erfüllt und die gesamte Rechnung ungültig.

Die am Ausgang des linearen Teiles übriggebliebene Grundschwingung stellt gerade das anfangs vorgegebene Signal $y(t) = -A \sin(\omega t)$ dar. Alle anderen Schwingungsanteile, die am Ausgang des nichtlinearen Teiles erzeugt wurden, konnten den linearen Teil nicht passieren. Aber nur Signalanteile, die in der Lage sind, alle Teile des Regelkreises zu passieren, können zu einer sich selbst aufrecht erhaltenden oder sogar aufklingenden Schwingung des Gesamtsystems beitragen und damit dessen Stabilität gefährden. Für die Stabilitätsanalyse ist es daher zulässig, alle höherfrequenten Anteile am Ausgang des nichtlinearen Teiles zu vernachlässigen. Es bleibt

$$u(t) = A_1 \cos \omega t + B_1 \sin \omega t = C_1 \sin(\omega t + \varphi_1) \quad (13.19)$$

mit $C_1 = \sqrt{A_1^2 + B_1^2}$ und $\varphi_1 = \arctan \frac{A_1}{B_1}$. $u(t)$ geht damit aus dem Eingangssignal $e(t) = A \sin(\omega t)$ durch eine Multiplikation mit dem Faktor $\frac{C_1}{A}$ und eine Phasenverzögerung um $-\varphi_1$ hervor. Dies entspricht aber doch gerade dem Verhalten eines linearen Laufzeitgliedes (vgl. (3.3)) mit einem konstanten Faktor. Man kann daher eine quasi-lineare Übertragungsfunktion entsprechend einem Laufzeitglied definieren, die das Verhalten des nichtlinearen Teiles beschreibt. Eine solche Funktion bezeichnet man als Beschreibungsfunktion:

$$\frac{u}{e} = N(A, \omega) = \frac{C_1(A, \omega)}{A} e^{j\varphi_1(A, \omega)} \quad (13.20)$$

Dabei sei angemerkt, dass diese Art der Linearisierung nichts mit der Linearisierung am Arbeitspunkt zu tun hat (Gleichung (7.13)). Gemäß der Definition von A_1 und B_1 hängen C_1 und φ_1 sowohl von der Amplitude A als auch von der Frequenz ω des Eingangssignales ab.

Bei einer momentan wirkenden Nichtlinearität entfällt die ω -Abhängigkeit. Denn ein solches Übertragungselement ist durch eine (nichtlineare) Kennlinie definiert, die das Verhältnis zwischen Ein- und Ausgangssignal beschreibt. Zu einem bestimmten Wert am Eingang wird damit immer der gleiche Wert am Ausgang ausgegeben, unabhängig von zeitlichen Verläufen und damit von Signalfrequenzen am Eingang. Dies wird auch durch das Beispiel im nachfolgenden Abschnitt noch einmal verdeutlicht.

Damit sind für diesen Fall die Parameter der Beschreibungsfunktion ausschließlich von der Amplitude des Eingangssignales abhängig:

$$N(A) = \frac{C_1(A)}{A} e^{j\varphi_1(A)} \quad (13.21)$$

Dies ist ein ganz entscheidender Unterschied zwischen einer solchen quasi-linearen und einer echten linearen Übertragungsfunktion, deren Laufzeit und Verstärkung ausschließlich von der Frequenz des Eingangssignales abhängig sind. Zudem gibt die Beschreibungsfunktion nur das Übertragungsverhalten des nichtlinearen Gliedes hinsichtlich der Grundschiwingung wieder. Die Beschreibungsfunktion darf daher nur dann wie eine lineare Übertragungsfunktion benutzt werden, wenn gewährleistet ist, dass das Eingangssignal des nichtlinearen Teiles tatsächlich $e(t) = A \sin(\omega t)$ ist. Eine Anwendung beispielsweise zur Berechnung der Sprungantwort ist damit ausgeschlossen.

Im vorliegenden Fall seien jedoch die Voraussetzungen erfüllt, und die Beschreibungsfunktion darf demnach wie eine lineare Übertragungsfunktion verwendet werden. Die Kreisübertragungsfunktion des Systems setzt sich nun zusammen aus der Beschreibungsfunktion und der Übertragungsfunktion des linearen Teiles: $N(A)G(j\omega)$. Damit sich eine gleichbleibende Schwingung einstellt, muss das Ausgangssignal y , nachdem es einmal den geschlossenen Kreis durchlaufen hat, am Ausgang in unveränderter Form wieder erscheinen. Die Bedingung für eine solche Schwingung lautet damit:

$$y = -N(A)G(j\omega)y \quad (13.22)$$

oder

$$-1 = N(A)G(j\omega) \quad (13.23)$$

bzw.

$$N(A)G(j\omega) + 1 = 0 \quad (13.24)$$

Die Zerlegung dieser komplexen Gleichung in Real- und Imaginärteil liefert zwei Gleichungen für die beiden Unbekannten, nämlich die Amplitude A und die Frequenz ω der Schwingung. Wenn eine Lösung dieser Gleichung existiert, so ist auch eine entsprechende Schwingung im System möglich, wobei dies eine Dauerschwingung oder ein Grenzyklus sein kann. Es können auch mehrere Lösungen existieren, was bedeutet, dass verschiedene Schwingungen möglich sind. Falls keine Lösung existiert, so bedeutet dies, dass keine harmonische Schwingung im Regelkreis existieren kann. Nichtharmonische Schwingungen

sind dann immer noch möglich, doch im allgemeinen recht unwahrscheinlich. Wie oben schon erwähnt, sollte für jede mögliche Schwingung am Ende noch einmal überprüft werden, ob durch den linearen Teil tatsächlich eine ausreichende Tiefpassfilterung der höherfrequenten Schwingungsanteile erfolgt, da dies eine ganz wesentliche Voraussetzung für das Verfahren ist.

Das Stabilitätsverhalten einer möglichen Schwingung kann im Rahmen einer graphischen Lösung durch Hinzuziehen des Nyquist-Kriteriums (Satz 5.6) ermittelt werden. Dieses Kriterium schreibt die erforderliche Phasendrehung der Ortskurve der Kreisübertragungsfunktion um den kritischen Punkt -1 vor.

Ausgangspunkt dieses Kriteriums ist der Nenner der Übertragungsfunktion des geschlossenen Kreises: $G(s)K(s) + 1$. Vergleicht man diesen Ausdruck mit Gleichung (13.24), so fällt die Ähnlichkeit sofort auf. Dividiert man nun Gleichung (13.24) durch $N(A)$, so erhält man

$$G(j\omega) + \frac{1}{N(A)} = 0 \quad (13.25)$$

und folgende Interpretation ist möglich: Die gemäß Nyquist-Kriterium zu untersuchende Kreisübertragungsfunktion ist hier $G(j\omega)$, und der kritische Punkt -1 des Nyquist-Kriteriums ist durch eine von der Amplitude A abhängige Kurve $-\frac{1}{N(A)}$ zu ersetzen.

Für eine Stabilitätsuntersuchung wird demnach zunächst die Beschreibungsfunktion des nichtlinearen Teiles $N(A)$ berechnet oder gemessen. Dann wird die Kurve $-\frac{1}{N(A)}$ in der komplexen Ebene dargestellt. Diese Kurve entspricht dem kritischen Punkt des Nyquist-Kriteriums. Anschließend misst oder berechnet man den Frequenzgang $G(j\omega)$ und stellt dessen Ortskurve ebenfalls in der komplexen Ebene dar. Jeder Schnittpunkt der beiden Kurven bildet dann eine Lösung der Gleichung (13.25), steht also für eine mögliche Schwingung, deren Stabilitätsverhalten mit Hilfe des Nyquist-Kriteriums ermittelt werden kann, wie in den folgenden Beispielen gezeigt wird.

13.5.2 Beispiele

Im ersten Beispiel besteht der nichtlineare Teil aus einem idealen Zweipunktglied. Zunächst wird dessen Beschreibungsfunktion berechnet. Die Parameter C_1 und φ_1 der Beschreibungsfunktion resultieren aus den Koeffizienten A_1 und B_1 , die demnach zuerst zu berechnen sind. Das Ausgangssignal $u(t)$ ist bei sinusförmigem Eingangssignal eine Rechteckschwingung (Abb. 13.18). Mit $T = \frac{2\pi}{\omega}$ ergibt sich:

$$B_1 = \frac{2}{T} \int_0^T u(t) \sin(\omega t) dt = \frac{2}{T} 2 \int_0^{\frac{T}{2}} K \sin(\omega t) dt$$

$$\begin{aligned}
&= \frac{4K}{T} \int_0^{\frac{T}{2}} \sin(\omega t) dt = -\frac{4K}{T\omega} \cos(\omega t) \Big|_0^{\frac{T}{2}} \\
&= \frac{4K}{T\omega} (-\cos(\frac{\omega T}{2}) + \cos(0)) = \frac{4K}{2\pi} (-\cos(\pi) + \cos(0)) \\
&= \frac{2K}{\pi} (1 + 1) = \frac{4K}{\pi} \\
A_1 &= \frac{2}{T} \int_0^T u(t) \cos(\omega t) dt = 0
\end{aligned} \tag{13.26}$$

und daraus

$$\begin{aligned}
C_1 &= \sqrt{A_1^2 + B_1^2} = B_1 = \frac{4K}{\pi} \\
\varphi_1 &= \arctan \frac{A_1}{B_1} = \arctan 0 = 0
\end{aligned} \tag{13.27}$$

bzw. die Beschreibungsfunktion

$$N(A) = \frac{C_1(A)}{A} e^{j\varphi_1(A)} = \frac{4K}{A\pi} \tag{13.28}$$

Die Phasenverzögerung $-\varphi_1$ der Beschreibungsfunktion beträgt damit Null und die Verstärkung $\frac{C_1}{A} = \frac{4K}{A\pi}$. Dies ist auch anschaulich sofort einsichtig. Die am Ausgang des Zweipunktgliedes anliegende Rechteckschwingung ist phasengleich zu der am Eingang anliegenden Sinusschwingung, weshalb die Phasenverzögerung offensichtlich Null sein muss. Weiterhin bleibt die Amplitude der am Ausgang anliegenden Rechteckschwingung immer gleich. Da die Verstärkung aber als das Verhältnis der Ausgangs- zur Eingangsamplitude definiert ist, muss sie gerade umgekehrt proportional zur Amplitude des Eingangssignales sein.

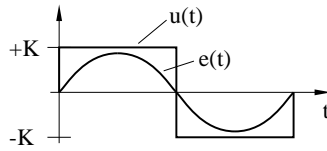


Abb. 13.18. Ein- und Ausgangssignal beim idealen Zweipunktglied

Tabellarische Auflistungen weiterer Beschreibungsfunktionen finden sich unter anderem in [5, 13, 14, 64]. Die ausführlichsten Informationen zur Beschreibungsfunktion bietet [15], das man durchaus als Standardwerk zu dieser Thematik bezeichnen kann.

Wenn die Beschreibungsfunktion bekannt ist, kann die eigentliche Stabilitätsanalyse durchgeführt werden. Dazu muss man zunächst die Funktion

$-\frac{1}{N(A)}$ als Kurve in Abhängigkeit von der Amplitude in die komplexe Ebene eintragen. Für das 2-Punkt-Glied ergibt sich nach (13.28) $-\frac{1}{N(A)} = -\frac{A\pi}{4K}$, also eine Kurve auf der negativ-reellen Achse, die sich mit wachsendem A immer weiter vom Nullpunkt entfernt. In dasselbe Bild wird dann die Ortskurve des linearen Teiles eingetragen. Anhand der entstehenden Schnittpunkte bzw. der Lage der Kurven zueinander sind dann Aussagen über die Stabilität des Systems möglich. Abb. 13.19 zeigt verschiedene Beispiele für den Fall, dass der nichtlineare Teil des Standardregelkreises (Abb. 13.15) aus einem idealen Zweipunktglied besteht.

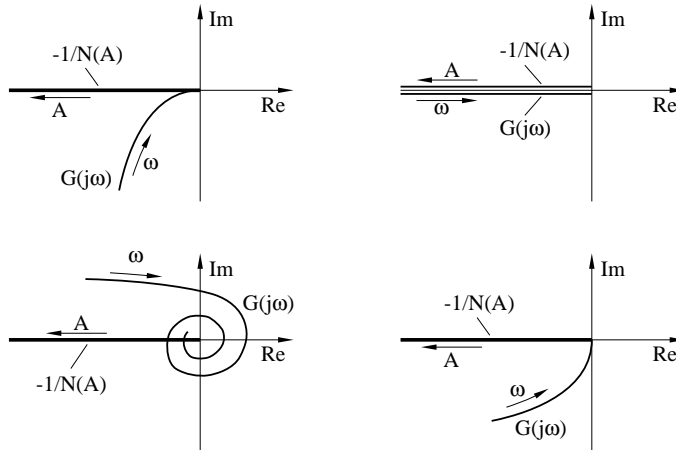


Abb. 13.19. Stabilitätsanalyse mittels Beschreibungsfunktion beim idealen 2-Punkt-Glied

Im Beispiel links oben besteht der lineare Teil aus einem Integrator und einem Verzögerungsglied mit der Übertragungsfunktion

$$G(s) = \frac{1}{s(Ts + 1)} \quad (13.29)$$

Der gesamte Kreis entspricht damit dem System in Abb. 13.4. Die Ortskurve und die Kurve der Beschreibungsfunktion schneiden sich nur im Ursprung, d.h. für $A = 0$, $\omega = \infty$. Damit besitzt aber auch Gleichung (13.23) nur diese eine Lösung, was bedeutet, dass in diesem System nur eine Schwingung mit der Amplitude $A = 0$, d.h. keine Schwingung möglich ist.

Man kann auch entsprechend der Interpretation der Gleichung (13.25) die Kurve $-\frac{1}{N(A)}$ als amplitudenabhängigen kritischen Punkt deuten. Dann ist $G(s)$ die Kreisübertragungsfunktion, deren Ortskurve laut Nyquistkriterium bezüglich des kritischen Punktes eine ganz bestimmte Phasendrehung ausführen muss, damit das System stabil ist. Da $G(s)$ einen Integrator enthält, beträgt diese Phasendrehung $+\frac{\pi}{2}$. Das ist aber gerade gegeben, wenn man den Schnittpunkt im Ursprung außer Acht lässt. Denn bezüglich jedes

anderen Punktes der Kurve $-\frac{1}{N(A)}$ hat die Phasendrehung genau diesen Wert. Dies kann man leicht feststellen, wenn man einen Vektor von einem Punkt der Kurve $-\frac{1}{N(A)}$ zur Ortskurve des linearen Teiles einzeichnet und seine Phasendrehung mit wachsendem ω betrachtet. Damit ist das System stabil.

Im Beispiel rechts oben besteht der lineare Teil aus einem doppelten Integrator, und man erhält das in Abb. 13.2 gezeigte System. Die Kurve der Beschreibungsfunktion und die Ortskurve des linearen Teiles liegen genau übereinander. Es existieren also unendlich viele Schnittpunkte und damit auch unendlich viele Lösungen der Gleichung (13.23). Dabei weist die Ortskurve des linearen Teiles umso größere Werte für ω auf, je weiter sie sich dem Ursprung nähert, während die Kurve der Beschreibungsfunktion umso größere Werte für A aufweist, je weiter sie sich vom Ursprung entfernt. Für einen Schnittpunkt und damit für eine mögliche Lösung bzw. Schwingung gilt also, dass die Amplitude umso größer ist, je kleiner die Frequenz ist. Das entspricht aber auch genau den bereits gemachten Untersuchungen zu diesem System. Wie man anhand von Abb. 13.3 erkennen kann, hängt die sich einstellende Dauerschwingung vom Anfangszustand des Systems ab. Je größer die Amplitude, desto langsamer die Schwingung bzw. desto kleiner die Frequenz.

Unten links besteht der lineare Teil aus einem zweifachen Integrator mit Laufzeit. Da sich die Ortskurve des linearen Teiles spiralförmig immer weiter dem Ursprung nähert, existieren unendlich viele Schnittpunkte zwischen beiden Kurven. Die Frage ist nun, welche Schwingung sich tatsächlich einstellen wird. Hier bietet sich eine Erklärung an, die zwar nicht ganz exakt, dafür aber anschaulich ist und letztendlich zum richtigen Ergebnis führt. Zunächst sei angenommen, dass sich das System in einem Schnittpunkt befindet und eine Schwingung ausführt. Wenn nun eine kleine Störung auftritt und die Amplitude möglicherweise etwas verkleinert wird, bewegt sich das System auf der Kurve der Beschreibungsfunktion ein wenig nach rechts. Dieser Punkt ist aber, wie alle anderen Punkte der Kurve $-\frac{1}{N(A)}$ auch, ein kritischer Punkt. Die Phasendrehung der Ortskurve um diesen Punkt ist sicherlich negativ, während sie laut Nyquist-Kriterium wegen der beiden Integratoren im linearen Teil $+\pi$ betragen müsste. Hinsichtlich dieses Punktes ist das System also instabil, und die Schwingung klingt auf. Das System bewegt sich auf der Kurve der Beschreibungsfunktion nach links zurück in den Schnittpunkt. Gegenüber einer Verkleinerung der Amplitude ist die Schwingung daher stabil. Tritt nun durch eine Störung eine Vergrößerung der Amplitude auf, so bewegt sich das System auf der Kurve der Beschreibungsfunktion nach links. Dieser Punkt ist ebenfalls ein kritischer Punkt, um den die Phasendrehung der Ortskurve wegen der bis ins Unendliche fortgesetzten Spiralförmigkeit sicherlich negativ ist, d.h. es liegt laut Nyquist-Kriterium wieder Instabilität vor. Die Schwingung klingt deshalb weiter auf und läuft in den nächsten, weiter vom Ursprung entfernt liegenden Schnittpunkt hinein. Dieselben Überlegungen gelten für alle Schnittpunkte, d.h. alle Grenzzyklen sind, in der Zustands-

ebene betrachtet, nach innen stabil und nach außen instabil. Daher wird das System im Laufe der Zeit mit jeder Störung zu immer weiter vom Ursprung entfernt liegenden Schnittpunkten wandern, was eine ständige Zunahme der Schwingungsamplitude bedeutet. Damit ist das System instabil.

Für das letzte Beispiel ist der Zweipunktregler mit doppeltem Integrator um eine Rückführung nach Abb. 13.9 ergänzt. Hier stellt sich vor der Anwendung des Verfahrens zunächst das Problem, das gegebene System so umzuformen, dass seine Struktur der des Standardregelkreises (Abb. 13.15) entspricht. Dazu wird das Zweipunktglied als nichtlinearer Teil definiert und alles andere als linearer Teil des Regelkreises. Für diesen linearen Teil muss nun die Übertragungsfunktion bestimmt werden. Sie ergibt sich dadurch, dass man den Zusammenhang zwischen Ausgangsgröße u und Eingangsgröße e des nichtlinearen Teiles herstellt. Im Standardregelkreis lautet dieser Zusammenhang $e = -G(s)u$. Im vorliegenden System gilt nach Abb. 13.9

$$e(s) = -u(s)\left(\frac{k}{s} + \frac{1}{s^2}\right) \quad (13.30)$$

und damit

$$G(s) = -\frac{e(s)}{u(s)} = \frac{ks + 1}{s^2} \quad (13.31)$$

Die Ortskurve dieser Funktion ist unten rechts in Abb. 13.19 eingezeichnet. Wie im ersten Beispiel liegt der einzige Schnittpunkt zwischen beiden Kurven wieder im Ursprung bei $A = 0$, was bedeutet, dass es keine harmonische Schwingung geben kann. Die Phasendrehung der Ortskurve bezüglich des kritischen Punktes, d.h. bezüglich der Kurve der Beschreibungsfunktion beträgt π . Genau dieser Wert ist aber wegen der beiden Integratoren in der Übertragungsfunktion laut Nyquist-Kriterium auch erforderlich, damit das Gesamtsystem stabil ist. Das hier vorliegende System ist also stabil. Allerdings ist dieses Ergebnis mit Vorsicht zu genießen, da die Graddifferenz zwischen Nenner- und Zählerpolynom nur 1 beträgt und von daher die ausreichende Tiefpasswirkung des linearen Teiles, die vorausgesetzt werden muss, fraglich ist. Da man aber bei der zu Abb. 13.9 bereits durchgeführten Betrachtung in der Zustandsebene zu demselben Ergebnis kommt, kann es hier ebenfalls akzeptiert werden.

Beim Zweipunktglied mit Hysterese hat die Kurve der Beschreibungsfunktion eine etwas andere Form. Wegen der Hysterese erfolgt die Umschaltung vom einen auf den anderen Ausgangswert gegenüber dem idealen Zweipunktglied und damit auch gegenüber einer Sinusschwingung am Eingang verzögert. Deshalb ist der Winkel φ ungleich Null und die Beschreibungsfunktion nicht mehr rein reell. Auf die Berechnung der Funktion soll verzichtet werden, aus Abb. 13.20 ist die Form der Kurve $-\frac{1}{N(A)}$ ersichtlich.

Im Beispiel links besteht der lineare Teil wieder aus einem Verzögerungsglied und einem Integrator (vgl. Abb. 13.7). Beide Kurven weisen einen Schnittpunkt bei einer von Null verschiedenen Amplitude auf. Dies deutet

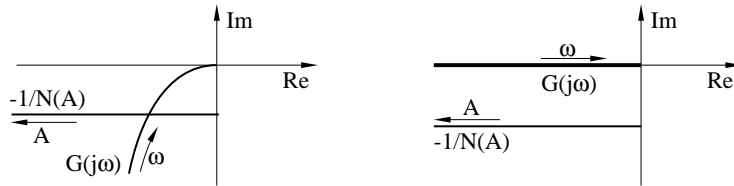


Abb. 13.20. Beschreibungsfunktion beim 2-Punkt-Glied mit Hysterese

auf eine Grenzschwingung hin. Zu untersuchen ist allerdings noch das Stabilitätsverhalten dieser Grenzschwingung, wobei wieder eine zwar nicht exakte, dafür aber anschauliche Erklärung versucht werden soll. Das System befinde sich zunächst in diesem Schnittpunkt. Nun tritt eine Störung auf, die die Amplitude der Schwingung etwas verkleinert. Das System nimmt einen Punkt auf der Kurve der Beschreibungsfunktion rechts vom Schnittpunkt ein. Die Phasendrehung der Ortskurve bezüglich dieses Punktes beträgt ungefähr $-\pi$, während laut Nyquist-Kriterium die für Stabilität erforderliche Phasendrehung wegen des einen Integrators in der linearen Übertragungsfunktion $+\frac{\pi}{2}$ betragen müsste. Es liegt demnach Instabilität vor, die Schwingung klingt auf, die Amplitude steigt an, und das System läuft wieder in den Schnittpunkt der beiden Kurven. Wandert das System infolge einer Störung auf der Kurve der Beschreibungsfunktion dagegen nach links, so beträgt die Phasendrehung ungefähr $+\frac{\pi}{2}$. Hier liegt Stabilität vor, die Schwingung klingt ab, und das System nähert sich ebenfalls wieder dem Schnittpunkt. Insgesamt ergibt sich, dass das System den Schnittpunkt nicht verlassen kann. Die Schwingung ist daher ein stabiler Grenzzyklus.

Im Beispiel rechts besteht der lineare Teil aus einem doppelten Integrator. Die Phasendrehung der Ortskurve hinsichtlich des kritischen Punktes müsste laut Nyquist-Kriterium $+\pi$ betragen, weist aber stattdessen hinsichtlich der gesamten Kurve der Beschreibungsfunktion negative Werte zwischen $-\frac{\pi}{2}$ und $-\pi$ auf. Das System ist daher instabil. Dasselbe Ergebnis lieferte auch die Betrachtung zu Abb. 13.6.

Nach diesen Beispielen ist wohl einsichtig, dass das Verfahren für einen geübten Anwender eine sehr einfache und übersichtliche Möglichkeit der Stabilitätsanalyse bietet. Dabei ist die benötigte Information leicht zu beschaffen. Sowohl die Beschreibungsfunktion als auch die Ortskurve können gemessen werden, wenn die Darstellung mit Hilfe von Formeln nicht möglich oder zu schwierig ist. Darüber hinaus ist die graphische Darstellung so anschaulich, dass man sich auf dieser Basis auch Möglichkeiten zur Stabilisierung eines Systems überlegen kann. Denn die Aufgabe besteht lediglich darin, die Ortskurve des linearen Teiles durch Einfügen linearer Korrekturglieder so zu verändern, dass kein Schnittpunkt zwischen der Ortskurve und der Kurve der Beschreibungsfunktion mehr auftritt. Der Phantasie des Anwenders sind hier keine Grenzen gesetzt. Der einzige Nachteil ist, dass das Verfahren in der bisher vorgestellten Form nur auf eine bestimmte Klasse von Systemen

anwendbar ist. Hier existieren aber verschiedene Erweiterungsmöglichkeiten, die im Folgenden kurz vorgestellt werden sollen.

13.5.3 Erweiterungen

Eine wichtige Einschränkung des bisher vorgestellten Verfahrens ist die Forderung nach einer ausreichend ausgeprägten Tiefpasseigenschaft des linearen Teiles. Dazu wird in [15] vorgeschlagen, bei nicht ausreichender Tiefpasswirkung die der Beschreibungsfunktion zu Grunde liegende Fourierreihe erst nach einem späteren Glied abzuberechnen. Der Charme des Verfahrens, nämlich die Darstellung der Nichtlinearität durch eine lineare Übertragungsfunktion und damit die einfache Handhabbarkeit, geht durch diese Maßnahme allerdings verloren.

Ebenfalls in [15] wird die Möglichkeit diskutiert, auch für andere Signalformen als harmonische Schwingungen Beschreibungsfunktionen nichtlinearer Übertragungsglieder zu berechnen, beispielsweise für gaußsches Rauschen oder Gleichsignale. Dann kann die Beschreibungsfunktion aber nicht mehr aus einer nach dem ersten Glied abgebrochenen Fourier-Reihe berechnet werden. Stattdessen wird eine lineare Übertragungsfunktion mit zunächst unbekannten Parametern angesetzt. Dann werden die Parameter so bestimmt, dass der quadratische Fehler zwischen dem Ausgangssignal dieses linearen und dem des realen, nichtlinearen Übertragungsgliedes bei gegebenem Eingangssignal möglichst klein wird. Setzt man als lineares Übertragungsglied ein Laufzeitglied mit variabler Verstärkung an, so liefert diese Vorgehensweise bei sinusförmigem Eingangssignal dieselbe Beschreibungsfunktion wie die nach dem ersten Glied abgebrochene Fourier-Reihe.

In [13, 14] wird erläutert, wie das Verfahren der harmonischen Balance anzuwenden ist, wenn der Regelkreis nicht nur aus einem linearen und einem nichtlinearen Teil wie im Standardregelkreis besteht, sondern mehrere nichtlineare Teile aufweist, die durch lineare Teile voneinander getrennt sind.

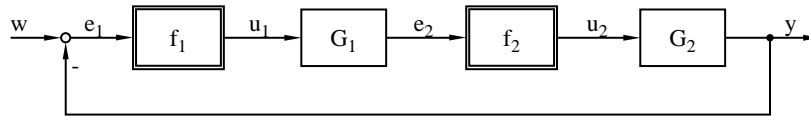


Abb. 13.21. Erweiterter Regelkreis für die Methode der Beschreibungsfunktion

Ein einfaches Beispiel zeigt Abb. 13.21. Unter der Voraussetzung, dass die linearen Teile ausreichende Tiefpasseigenschaft aufweisen, können für e_1 und e_2 harmonische Schwingungen angesetzt werden:

$$\begin{aligned} e_1 &= A_1 \sin \omega t \\ e_2 &= A_2 \sin(\omega t + \varphi_2) \end{aligned} \quad (13.32)$$

bzw. in der Darstellung als komplexe Zeiger

$$\begin{aligned} e_1 &= A_1 e^{j\omega t} \\ e_2 &= A_2 e^{j(\omega t + \varphi_2)} \end{aligned} \quad (13.33)$$

Anschließend werden die nichtlinearen Teile durch quasilineare Beschreibungsfunktionen $N_1(A_1, \omega)$ und $N_2(A_2, \omega)$ ersetzt. Für das Schwingungsgleichgewicht gilt dann:

$$-1 = N_1(A_1, \omega) G_1(\omega) N_2(A_2, \omega) G_2(\omega) \quad (13.34)$$

Eine Zerlegung in Real- und Imaginärteil liefert zwei Gleichungen. Hier gibt es aber drei Unbekannte, nämlich A_1 , A_2 und ω . Da die Beschreibungsfunktionen jedoch wie lineare Übertragungsfunktionen behandelt werden können, lässt sich ein weiterer Zusammenhang, und zwar zwischen den Eingangssignalen der nichtlinearen Glieder aufstellen:

$$\begin{aligned} e_2 &= N_1(A_1, \omega) G_1(\omega) e_1 \\ A_2 e^{j(\omega t + \varphi_2)} &= N_1(A_1, \omega) G_1(\omega) A_1 e^{j\omega t} \\ A_2 e^{j\varphi_2} &= N_1(A_1, \omega) G_1(\omega) A_1 \end{aligned} \quad (13.35)$$

Eine Betrachtung der Beträge liefert dann die notwendige, dritte Gleichung:

$$A_2 = |N_1(A_1, \omega)| |G_1(\omega)| A_1 \quad (13.36)$$

Das so erhaltene Gleichungssystem ist leider nur noch in Sonderfällen graphisch zu lösen. Es bleibt aber die Möglichkeit einer numerischen Lösung.

Wichtiger für die Praxis ist die Möglichkeit, das Verfahren für Nichtlinearitäten zu erweitern, die nicht mehr momentan wirkend sind, sondern eine interne Dynamik aufweisen. Damit ist die Ausgangsgröße des nichtlinearen Teiles u nicht mehr nur vom Eingangssignal e bzw. dem Ausgang des linearen Teiles y abhängig, sondern auch von dessen Ableitungen: $u = f(e, \dot{e}, \dots)$. Eine solche Abhängigkeit tritt offenbar ebenfalls auf, wenn der nichtlineare Teil zwar keine Dynamik aufweist, dafür aber als Eingangsgrößen nicht nur die Regelabweichung bzw. die Ausgangsgröße der Strecke, sondern auch ihre Ableitungen erhält. Dies ist eine Konstellation, wie sie beispielsweise beim Fuzzy-Regler gegeben ist.

So sei der nichtlineare Teil jetzt statt durch $u = f(e)$ durch ein Übertragungsverhalten erster Ordnung $u = f(e, \dot{e})$ definiert. Weiterhin sei diese Funktion ungerade: $f(-e, -\dot{e}) = -f(e, \dot{e})$. Und schließlich muss für jedes $\dot{e} > 0$ die Funktion $f(e, \dot{e})$ mit e monoton steigen. Der ausreichende Tiefpasscharakter des linearen Teiles wird ebenfalls vorausgesetzt. Dann kann man genau wie im Fall momentan wirkender Nichtlinearitäten die Oberschwingungen am Ausgang des nichtlinearen Teiles vernachlässigen. Für die Koeffizienten der Grundschiwingung gilt jetzt:

$$A_1 = \frac{2}{T} \int_0^T f(e, \dot{e}) \cos(\omega t) dt = \frac{2}{T} \int_0^T f(A \sin(\omega t), A\omega \cos(\omega t)) \cos(\omega t) dt$$

(13.37)

$$B_1 = \frac{2}{T} \int_0^T f(e, \dot{e}) \sin(\omega t) dt = \frac{2}{T} \int_0^T f(A \sin(\omega t), A\omega \cos(\omega t)) \sin(\omega t) dt$$

mit $T = \frac{2\pi}{\omega}$. Nach denselben Formeln wie für momentan wirkende Nichtlinearitäten ergibt sich wieder eine Beschreibungsfunktion $N(A, \omega)$, die jetzt aber nicht mehr nur von der Amplitude A , sondern auch von der Frequenz ω der Schwingung abhängig ist. Dies führt dazu, dass die Darstellung dieser Beschreibungsfunktion nicht nur eine Kurve $-\frac{1}{N(A)}$, sondern eine ganze Kurvenschar $-\frac{1}{N(A, \omega)}$ mit ω als Parameter erfordert, d.h. für jede Frequenz ω_1 existiert eine amplitudenabhängige Kurve $-\frac{1}{N(A, \omega_1)}$.

Für eine Stabilitätsanalyse werden diese Kurvenschar und die Ortskurve des linearen Teiles in der komplexen Ebene aufgetragen. Die Kurvenschar wird dann als kritischer Punkt des Nyquist-Kriteriums gedeutet. Aus der Lage der Ortskurve zur Kurvenschar lassen sich auch hier Rückschlüsse auf das Stabilitätsverhalten ziehen. Als Beispiel zeigt Abb. 13.22 die Ortskurve eines PT_3 -Gliedes und eine Kurvenschar, wie sie bei einem Fuzzy-Regler entstehen könnte.

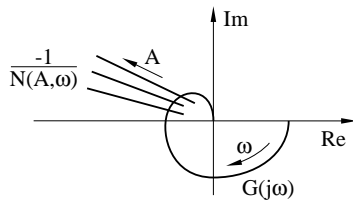


Abb. 13.22. Harmonische Balance mit frequenzabhängiger Beschreibungsfunktion

Sollte überhaupt einer der Schnittpunkte eine Schwingung im System kennzeichnen, so wird es sich auf jeden Fall um einen stabilen Grenzzyklus handeln. Zur Erklärung sei angenommen, dass sich das System in einem Schnittpunkt befindet und eine Schwingung ausführt. Falls durch eine Störung die Amplitude verkleinert wird, so nähert sich das System auf der entsprechenden Kurve der Kurvenschar dem Ursprung und befindet sich dadurch in einem Punkt, der von der Ortskurve umschlossen wird. Die Phasendrehung der Ortskurve um diesen Punkt ist negativ, während sie laut Nyquist-Kriterium für ein stabiles System Null betragen müsste. Es liegt Instabilität vor, die Schwingung klingt wieder auf, und das System wandert zurück in den Schnittpunkt. Falls sich das System dagegen im Schnittpunkt befindet und die Amplitude durch eine Störung vergrößert wird, so entfernt sich das System auf einer Kurve der Kurvenschar vom Ursprung. Die Phasendrehung der Ortskurve um den dann vom System eingenommenen Punkt beträgt Null. Es liegt Stabilität vor, die Schwingung klingt ab, und das System wandert auch hier wieder zurück in den Schnittpunkt.

Im Gegensatz zu vorher repräsentiert jetzt aber nicht mehr jeder Schnittpunkt eine mögliche Schwingung. Bisher wurde in einem Schnittpunkt durch die Kurve der Beschreibungsfunktion die Amplitude und durch die Ortskurve des linearen Teiles die Frequenz der Schwingung definiert, und jeder Schnittpunkt entsprach einer möglichen Lösung der Gleichung (13.25). Jetzt ist die Kurve der Beschreibungsfunktion dagegen zusätzlich noch frequenzabhängig. Damit ein Schnittpunkt eine mögliche Schwingung repräsentiert, müssen die durch die Kurve der Beschreibungsfunktion und die durch die Ortskurve des linearen Teiles im Schnittpunkt gegebenen Frequenzen übereinstimmen. Nur dann ist diese Frequenz auch die Frequenz einer möglichen Schwingung. Die Amplitude ergibt sich nach wie vor aus der Kurve der Beschreibungsfunktion.

Offenbar ist eine graphische Lösung unter diesen Bedingungen reiner Zufall, so dass zur Ermittlung der Werte für Amplitude und Frequenz der Schwingung von vornherein nur eine numerische Lösung der (13.25) entsprechenden Gleichung

$$G(j\omega) = -\frac{1}{N(A, \omega)} \quad (13.38)$$

in Frage kommt.

Auf numerischem Wege lässt sich auch die Beschreibungsfunktion $N(A, \omega)$ selbst grundsätzlich immer bestimmen. Dies bietet sich an, wenn vom nichtlinearen Teil überhaupt keine analytische Beschreibung vorliegt, wie dies vor allem bei einem Fuzzy-Regler der Fall ist. Dazu wird ein bestimmtes Wertepaar (A_1, ω_1) vorgegeben und die entsprechende Sinusschwingung am Eingang des nichtlinearen Teiles aufgeschaltet. An seinem Ausgang wird sich eine periodische Schwingung einstellen, die aber nicht unbedingt einer Sinusschwingung entspricht. Mit der Methode der kleinsten Fehlerquadrate kann sie jedoch durch eine Sinusschwingung approximiert werden. Ein Vergleich dieser approximierenden Schwingung mit der Eingangsschwingung liefert dann die Verstärkung V und die Phasenverzögerung $-\varphi$ des nichtlinearen Teiles für das Wertepaar (A_1, ω_1) . Dies führt aber auch sofort auf den (komplexen) Wert $N(A_1, \omega_1)$ der Beschreibungsfunktion. Auf diese Art und Weise kann die Beschreibungsfunktion punktweise ermittelt werden.

In [62] wird sogar eine Erweiterung des Verfahrens auf Mehrgrößensysteme diskutiert. Diese Erweiterung erfordert aber Voraussetzungen beim System, die im Anwendungsfall nicht nachzuprüfen sind. Eine Stabilitätsanalyse mit diesem Verfahren steht damit auf recht unsicherem Fundament, so dass hier auf eine Darstellung von vornherein verzichtet werden soll.

13.6 Popov-Kriterium

Damit kann zu einem anderen Verfahren übergegangen werden, das auf dem Stabilitätskriterium von Popov basiert. Im Gegensatz zur Methode der harmonischen Balance ist es ein exaktes Kriterium. Allerdings kann es in Einzelfällen zu sehr konservativen Ergebnissen führen, da es zwar hinreichend,

aber nicht notwendig ist. Das bedeutet, dass die Stabilität eines stabilen Systems möglicherweise nicht nachgewiesen werden kann. Andererseits ist es einfach anzuwenden.

Voraussetzung ist wieder, dass das System in einen momentan wirkenden, nichtlinearen Teil und einen linearen Teil unterteilt werden kann. Die Kennlinie des nichtlinearen Teiles und die Ortskurve des linearen Teiles müssen bekannt sein.

Um die Formulierung des Kriteriums möglichst einfach zu halten, ist für den nichtlinearen Teil eine zusätzliche Definition erforderlich (vgl. Abb. 13.23):

Definition 13.5 Eine Kennlinie $f(e)$ verläuft im Sektor $[k_1, k_2]$, wenn gilt

$$k_1 \leq \frac{f(e)}{e} \leq k_2 \quad \text{und} \quad f(0) = 0 \quad (13.39)$$

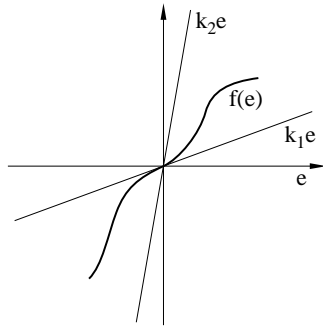


Abb. 13.23. Sektor einer Kennlinie

Damit kann das Popov-Kriterium formuliert werden, dessen Beweis mit Hilfe der direkten Methode von Ljapunov erfolgen kann, auf den hier aber verzichtet werden soll (siehe [2]):

Satz 13.6 (Popov-Kriterium) Gegeben sei ein geschlossener Kreis, bestehend aus einem linearen und einem nichtlinearen Teil. Die Übertragungsfunktion $G(s)$ des linearen Teiles sei rein rational, habe ausschließlich Polstellen mit negativem Realteil und einen Verstärkungsfaktor $V_l = 1$. Weiterhin sei die Ordnung des Nennerpolynoms größer als die des Zählerpolynoms. Der nichtlineare Teil sei durch eine eindeutige und stückweise stetige Kennlinie $u = f(e)$ gegeben. Wenn dann die Ungleichung

$$\operatorname{Re}((1 + jq\omega)G(j\omega)) > -\frac{1}{k} \quad (13.40)$$

mit $0 < k \leq \infty$ und beliebigem, endlichem q für alle Frequenzen $0 \leq \omega < \infty$ erfüllt ist, so besitzt der Regelkreis für jede Kennlinie, die im Sektor $[0, k]$

verläuft, eine global asymptotisch stabile Ruhelage für $u = y = 0$. Er wird dann auch als absolut stabil im Sektor $[0, k]$ bezeichnet. Wenn die rechte Seite der Ungleichung auch gleich Null (d.h. $k = \infty$) gesetzt werden kann, so ergibt sich ein Sektor $[0, \infty]$.

Weist die Funktion $G(s)$ nicht nur Polstellen mit negativem Realteil, sondern auch eine Polstelle auf der imaginären Achse auf, so ist zusätzlich zu zeigen, dass die lineare Übertragungsfunktion $\frac{\varepsilon G(s)}{1 + \varepsilon G(s)}$ mit irgendeinem $\varepsilon > 0$ stabil ist (Grenzstabilität). In diesem Fall ist der zulässige Sektor für die nichtlineare Kennlinie $(0, k]$.

Treten mehrere Polstellen auf der imaginären Achse auf, so ergibt sich neben der Forderung nach Grenzstabilität als weitere Verschärfung, dass für k nur noch endliche Werte zugelassen sind und sich der Sektor absoluter Stabilität auf $[\varepsilon, k]$ reduziert. Weiterhin darf keine Frequenz existieren, die das Gleichungssystem

$$\operatorname{Re}(G(j\omega)) = -\frac{1}{k} \quad \text{und} \quad \omega \operatorname{Im}(G(j\omega)) = 0 \quad (13.41)$$

erfüllt. Dafür ist in (13.40) aber das Gleichheitszeichen zugelassen.

Als sehr starke Einschränkung mag zunächst die Forderung nach einer Verstärkung $V_l = 1$ des linearen Teiles erscheinen. Dies ist aber nicht so, denn ein Verstärkungsfaktor $V_l \neq 1$ kann ohne Probleme dem nichtlinearen Teil hinzugerechnet werden. Statt $u = f(e)$ erhält man dann die Kennlinie $\tilde{u} = V_l f(e) = \tilde{f}(e)$.

Auch die Voraussetzungen für die lineare Übertragungsfunktion sollen kurz erläutert werden. Wenn eine Kennlinie tatsächlich auf der im Satz angegebenen unteren Sektorgrenze $k_1 = 0$ verläuft, so bedeutet dies doch, dass die Stellgröße u und damit die Eingangsgröße des linearen Teiles immer gleich Null sind. Damit ist aber der lineare Systemteil ohne äußere Einwirkung, also gewissermaßen sich selbst überlassen. Wenn dann asymptotische Stabilität des Gesamtsystems gefordert ist, so kann dies nur dadurch gewährleistet werden, dass der lineare Teil auch ohne Einwirkung von außen aus jedem Anfangszustand zur Ruhe kommen kann. Daraus resultiert wiederum die im Satz formulierte Forderung nach dem negativen Realteil sämtlicher Polstellen (vgl. Satz 7.3).

Wenn nun die lineare Übertragungsfunktion auch rein imaginäre Polstellen aufweist (also beispielsweise einen Integralanteil), würde der lineare Systemteil ohne äußere Einwirkung nicht in den Nullzustand laufen. Deshalb muss in diesem Fall Null als untere Sektorgrenze ausgeschlossen werden. Diese Einschränkung des zulässigen Sektors ist aber noch nicht ausreichend. Zusätzlich muss noch gezeigt werden, dass der geschlossene Kreis überhaupt stabilisierbar ist, und zwar durch die Kennlinie $f(e) = \varepsilon e$. Der nichtlineare Teil muss demnach durch einen linearen Verstärkungsfaktor ε ersetzt und für den so entstandenen, linearen Kreis

$$\frac{\varepsilon G(s)}{1 + \varepsilon G(s)} \quad (13.42)$$

die Stabilität nachgewiesen werden. Diese Eigenschaft bezeichnet man als *Grenzstabilität*. Ihr Nachweis ist aber nicht weiter schwierig, da es sich um ein rein lineares Problem handelt.

Schließlich bleiben noch die Verschärfungen im letzten Absatz des Satzes zu diskutieren. Die Reduzierung auf endliche Werte von k bedeutet, dass beispielsweise ein ideales Zweipunktglied nicht mehr die Voraussetzungen für eine Anwendung erfüllt, da die Steigung seiner Kennlinie im Nullpunkt unendlich groß ist. Und die Bedingung, dass für keine Frequenz das Gleichungssystem (13.41) erfüllt sein darf, ist gleichbedeutend mit der Forderung, dass die im Folgenden noch vorgestellte Popov-Ortskurve nicht durch den Punkt $(-\frac{1}{k}, 0)$ laufen darf.

Erweitert werden kann der obige Satz auch für den Fall, dass der lineare Teil eine Laufzeit enthält. Die Voraussetzungen des Satzes sind dann dahingehend zu verschärfen, dass die nichtlineare Kennlinie nicht nur stückweise stetig, sondern stetig sein muss und weiterhin q jetzt nicht mehr beliebig gewählt werden kann, sondern $q > 0$ gelten muss.

Verschiedene andere Spezialfälle, die aber für die Praxis nicht mehr so relevant sind, finden sich in [2]. Man muss sich aber immer darüber im klaren sein, dass das Popov-Kriterium keine Aussage für den Fall macht, dass eine Kennlinie den Sektor verlässt. Instabilität kann mit dem Popov-Kriterium nicht nachgewiesen werden.

Es stellt sich noch die Frage nach der Vorgehensweise bei der Anwendung auf ein praktisches Problem. Gegeben sind beispielsweise eine nichtlineare Kennlinie und die Ortskurve des linearen Teiles, der wiederum die Voraussetzungen des Satzes erfüllt. Die Frage ist, ob der geschlossene Kreis stabil ist. Dazu ist mit Hilfe der Ungleichung (13.40) der zulässige Sektor $[0, k]$ zu ermitteln und zu überprüfen, ob die gegebene Kennlinie in diesem Sektor liegt. Zunächst wird ein beliebiger Wert q festgelegt und aus der Ungleichung (13.40) der zugehörige Wert k berechnet, mit dem diese Ungleichung für alle ω erfüllt ist. Wenn dann die Kennlinie im durch k definierten Sektor liegt, so ist die Stabilität des Systems nachgewiesen.

Ein Problem entsteht aber, wenn eine gegebene Kennlinie in diesem Sektor nicht enthalten ist. Da Satz 13.6 nur ein hinreichendes Stabilitätskriterium darstellt, lässt sich in diesem Fall keine Aussage machen. Die Frage ist dann, ob ein q existiert, mit dem man einen größeren Sektor erhalten hätte. Eine ähnliche Frage stellt sich auch, wenn die nichtlineare Kennlinie (z.B. beim Reglerentwurf) erst noch festgelegt werden soll. In dem Fall ist man natürlich daran interessiert, einen möglichst großen Sektor für die Kennlinie zur Verfügung zu haben. Grundsätzlich sollte man q also nicht beliebig festlegen, sondern versuchen, q so zu bestimmen, dass k maximal wird.

Für diese Aufgabe existiert eine sehr elegante, graphische Lösung. Dazu ist zunächst die Popov-Ungleichung (13.40) umzuschreiben in

$$\operatorname{Re}(G(j\omega)) - q\omega \operatorname{Im}(G(j\omega)) > -\frac{1}{k} \quad (13.43)$$

Nun definiert man eine neue Ortskurve $\tilde{G}(j\omega) = \tilde{x} + j\tilde{y}$ mit dem Realteil $\tilde{x} = \operatorname{Re}(G(j\omega))$ und dem Imaginärteil $\tilde{y} = \omega \operatorname{Im}(G(j\omega))$. Dies ist die sogenannte *Popov-Ortskurve*. Die Popov-Ungleichung lautet mit den Koordinaten dieser Ortskurve

$$\tilde{x} - q\tilde{y} > -\frac{1}{k} \quad (13.44)$$

oder umgestellt

$$\tilde{x} > q\tilde{y} - \frac{1}{k} \quad (13.45)$$

Diese Ungleichung muss für alle Werte von ω , also für jeden Punkt der Ortskurve, erfüllt sein. Der Grenzfall dieser Ungleichung ist

$$\tilde{x}_G = q\tilde{y} - \frac{1}{k} \quad (13.46)$$

bzw.

$$\tilde{y} = \frac{1}{q}(\tilde{x}_G + \frac{1}{k}) \quad (13.47)$$

also eine Gerade mit der Steigung $\frac{1}{q}$ und dem \tilde{x} -Achsenabschnitt $-\frac{1}{k}$. Durch diese Grenzgerade wird zu jedem Imaginärteil \tilde{y} der Popov-Ortskurve ein Realteil \tilde{x}_G vorgegeben. Andererseits muss aber der Realteil \tilde{x} der Popov-Ortskurve nach Gleichung (13.45) größer sein als der durch die Grenzgerade vorgegebene Realteil. Die Ungleichungen (13.45) und damit (13.40) sind daher nur dann für alle Werte von ω erfüllt, wenn die Popov-Ortskurve rechts von der Grenzgeraden, d.h. im Bereich größerer Realteile verläuft (Abb. 13.24).

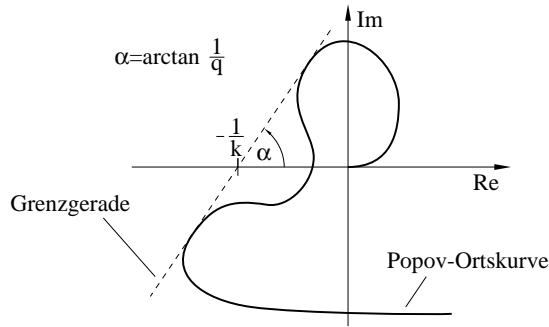


Abb. 13.24. Graphische Bestimmung des maximalen Sektors

Die Vorgehensweise zur Bestimmung der größtmöglichen Sektorgrenze k wird anhand von Abb. 13.24 deutlich. Anhand der gemessenen oder berechneten Ortskurve des linearen Teiles $G(j\omega)$ ist zunächst die Popov-Ortskurve

zu zeichnen. Dann muss eine Grenzgerade eingezeichnet werden. Ihre Steigung $\frac{1}{q}$ ist beliebig, da auch q laut Satz 13.6 beliebig gewählt werden kann. Allerdings muss sie links von der Ortskurve verlaufen, damit die Ungleichung (13.40) erfüllt ist. Durch den Schnittpunkt $-\frac{1}{k}$ der Grenzgeraden mit der reellen Achse wird dann die obere Sektorgrenze k festgelegt. Je weiter dieser Schnittpunkt rechts liegt, desto größer ist k . Das maximale k ergibt sich offensichtlich, wenn die Grenzgerade wie eingezeichnet annähernd eine Tangente an die Popov-Ortskurve darstellt. Eine echte Tangente darf sie nicht sein, da sonst in der Ungleichung (13.40) auch die Gleichheit zugelassen sein müsste. Diese Unterscheidung kann aber in der Praxis ruhigen Gewissens vernachlässigt werden, da dort wegen der Ungenauigkeit beim Messen und Zeichnen sowieso keine exakten Werte ermittelt werden.

Interessant für die Anwendung ist auch die Möglichkeit einer *Sektortransformation*. Satz 13.6 geht immer von einer unteren Sektorgrenze 0 bzw. ε aus. Falls nun die Kennlinie in einem beliebigen Sektor $[k_1, k_2]$ mit $k_1 < 0$ liegt, so ist der Satz zunächst einmal nicht anwendbar. In einem solchen Fall ist die Kennlinie $u = f(e)$ zu ersetzen durch die transformierte Kennlinie $u_t = f_t(e) = f(e) - k_1 e$, wie es in Abb. 13.25 dargestellt ist. Die neue Kennlinie liegt dann in einem Sektor $[0, k]$ mit $k = k_2 - k_1$. Diese Maßnahme kann man auch so deuten, dass man in den geschlossenen Kreis parallel zur Nichtlinearität ein Proportionalglied mit dem Verstärkungsfaktor $-k_1$ einfügt, so dass dann die Nichtlinearität $f(e)$ zusammen mit dem Proportionalglied gerade die transformierte Nichtlinearität $f_t(e)$ bildet.

Eine solche Veränderung des Systems würde natürlich das Ergebnis der Stabilitätsanalyse verfälschen. Deshalb ist vor Beginn der Rechnung die durch die Kennlinientransformation erfolgte Veränderung an anderer Stelle wieder aufzuheben. Es bietet sich an, parallel zur transformierten Nichtlinearität $f_t(e)$ ein weiteres Proportionalglied mit der Verstärkung k_1 einzufügen, das die Wirkung des ersten Proportionalgliedes gerade wieder aufhebt. Dieses Proportionalglied wird dann für die Analyse allerdings dem linearen Systemteil hinzugerechnet. Die Frage ist jetzt, wie die Übertragungsfunktion des veränderten linearen Systemteiles aussieht. In einem nicht transformierten System gilt die Beziehung

$$\frac{e}{u} = -G(s) \quad (13.48)$$

Im transformierten System nach Abb. 13.25 ergibt sich

$$\begin{aligned} (u_t + k_1 e)G(s) &= -e \\ \frac{e}{u_t} &= -\frac{G(s)}{1 + k_1 G(s)} = -G_t(s) \end{aligned} \quad (13.49)$$

und damit für die lineare Übertragungsfunktion des transformierten Systems

$$G_t(s) = \frac{G(s)}{1 + k_1 G(s)} \quad (13.50)$$

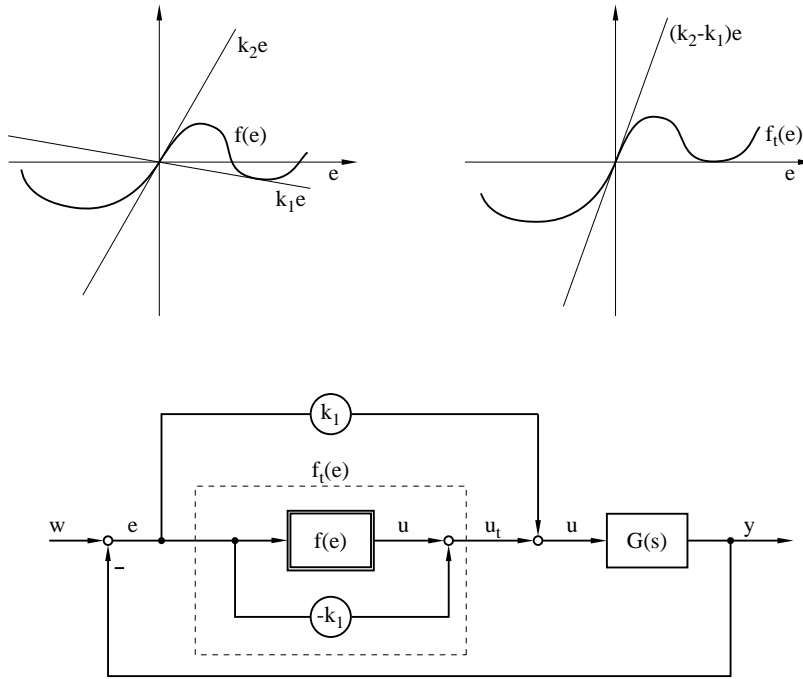


Abb. 13.25. Sektortransformation

Eine Sektortransformation umfasst also zwei Schritte: Der Sektor $[k_1, k_2]$ wird durch den Sektor $[0, k]$ ersetzt mit $k = k_2 - k_1$, und die Übertragungsfunktion $G(s)$ des linearen Teiles durch $G_t(s)$ nach (13.50). Auf dieses transformierte System wird dann das Popov-Kriterium angewendet, was bedeutet, dass G_t die Voraussetzungen des Kriteriums für den linearen Systemteil erfüllen muss. Kann dann für das transformierte System Stabilität nachgewiesen werden, so gilt dies auch für das Originalsystem.

Angemerkt sei, dass eine Sektortransformation auch für $k_1 > 0$ Vorteile bringen kann. Wenn man beispielsweise weiß, dass die Kennlinie im Sektor $[k_1, k_2]$ mit $k_1 > 0$ verläuft, so verkleinert sich durch die Sektortransformation die obere Sektorgrenze $k = k_2 - k_1$ und damit auch der Sektor $[0, k]$, für den Stabilität nachzuweisen ist. Die Bedingung für den linearen Systemteil fällt dadurch offenbar weniger streng aus.

Es gibt auch eine Version des Popov-Kriteriums für Mehrgrößensysteme. Voraussetzung für dessen Anwendung ist jedoch, dass jede Ausgangsgröße des nichtlinearen Systemteils nur von einer einzigen Eingangsgröße abhängig sein darf ($u_i = f_i(e_i)$). Da diese Einschränkung in der Praxis unrealistisch ist, soll hier auf eine Darstellung verzichtet werden.

Nicht unerwähnt bleiben soll jedoch die berühmte Vermutung von Aizermann. Sie lautet: Wenn man den nichtlinearen Systemteil $u = f(e)$ durch ein Proportionalglied mit dem Verstärkungsfaktor k_1 ersetzt und das so entstan-

dene Gesamtsystem stabil ist und dasselbe auch für einen Verstärkungsfaktor $k_2 > k_1$ gilt, dann ist das System auch für jede beliebige nichtlineare Kennlinie im Sektor $[k_1, k_2]$ stabil. Obwohl diese Vermutung plausibel erscheint, so ist sie doch nicht allgemeingültig. Ein Gegenbeweis findet sich in [17] und Gegenbeispiele gibt es schon für Systeme zweiter Ordnung. Man kann die Stabilität eines Systems mit einer nichtlinearen Kennlinie eben nicht dadurch abschätzen, dass man die nichtlineare Kennlinie mit linearen Kennlinien vergleicht. Leider findet sich dieses Vorgehen in der Praxis aber relativ häufig, weshalb hier ausdrücklich davor gewarnt werden soll.

Abschließend ist zu sagen, dass das Popov-Kriterium im Eingrößenfall recht einfach anzuwenden und damit für die Praxis gut geeignet ist. Die benötigten Informationen über das System sind leicht zu beschaffen. Für den linearen Systemteil reicht der gemessene Frequenzgang aus, während für den nichtlinearen Teil nur der Sektor bekannt sein muss, in dem die Kennlinie verläuft. Dafür liefert das Popov-Kriterium, da es nur ein hinreichendes Kriterium ist, sehr konservative Ergebnisse, d.h. oft kann mit dem Popov-Kriterium kein Stabilitätsnachweis erbracht werden, obwohl das System stabil ist.

13.7 Kreiskriterium

Das nächste vorgestellte Stabilitätskriterium ist das *Kreiskriterium*. Es basiert auf genau denselben Voraussetzungen wie das Popov-Kriterium. Auch hier wird von einer Unterteilung des Systems in einen linearen und einen nichtlinearen Teil ausgegangen, wobei das Übertragungsverhalten des nichtlinearen Teiles aber nicht unbedingt durch eine statische Kennlinie darstellbar sein muss. Für den Eingrößenfall mit einer statischen Kennlinie lässt sich das Verfahren relativ einfach herleiten, indem man in der Popov-Ungleichung (13.40) den freien Parameter q zu Null setzt, einige Umformungen vornimmt und das Ergebnis graphisch interpretiert (vgl. [13, 14]). Geradliniger auf Mehrgrößensysteme erweiterbar ist aber eine Herleitung, die auf der Verwendung von Normen basiert (vgl. [5]).

Normen werden im Anhang 10.2 ausführlich behandelt. So lässt sich die Norm einer Übertragungsmatrix als eine Art maximaler Verstärkungsfaktor vom Ein- zum Ausgangssignalvektor interpretieren. Es gilt beispielsweise für die ∞ -Norm einer linearen Übertragungsmatrix \mathbf{G} mit $\mathbf{y} = \mathbf{G}\mathbf{u}$ gemäß Gleichung (10.25)

$$\|\mathbf{G}(j\omega)\|_{\infty} = \sup_{\omega} \sup_{\mathbf{u} \neq \mathbf{0}} \frac{|\mathbf{G}(j\omega)\mathbf{u}|}{|\mathbf{u}|} \quad (13.51)$$

und für die ∞ -Norm einer nichtlinearen Übertragungsfunktion mit $\mathbf{f}(\mathbf{e}, \dot{\mathbf{e}}, \dots) = \mathbf{u}$ laut Gleichung (10.28)

$$\|\mathbf{f}\|_{\infty} = \sup_{\mathbf{e} \neq \mathbf{0}} \frac{|\mathbf{u}|}{|\mathbf{e}|} \quad (13.52)$$

wobei \mathbf{e} , \mathbf{u} und \mathbf{y} die Größen des Regelkreises gemäß Abb. 13.15 darstellen. Für Eingrößensysteme wird daraus (vgl. Gleichung (10.27))

$$\begin{aligned} \|G(j\omega)\|_{\infty} &= \sup_{\omega} |G(j\omega)| \\ \|f\|_{\infty} &= \sup_{e \neq 0} \frac{|u|}{|e|} \end{aligned} \quad (13.53)$$

In einem Regelkreis, der aus einem linearen und einem nichtlinearen Teil besteht, gilt mit diesen Definitionen $\mathbf{y} = \mathbf{G}(j\omega)\mathbf{f}(\mathbf{e}, \dot{\mathbf{e}}, \dots)$. Wäre \mathbf{f} eine lineare Übertragungsmatrix \mathbf{F} , so könnte man schreiben $\mathbf{y} = \mathbf{G}\mathbf{F}\mathbf{e}$, und $\mathbf{G}\mathbf{F}$ wäre die Matrix der Kreisübertragungsfunktion, anhand der eine Stabilitätsanalyse erfolgen kann. Da \mathbf{f} aber nur als nichtlineare Vektorfunktion von \mathbf{e} definiert ist, existiert keine explizite Kreisübertragungsfunktion. Es lässt sich lediglich, entsprechend der Definition der einzelnen Normen, der maximale Übertragungsfaktor von $|\mathbf{e}|$ nach $|\mathbf{y}|$ abschätzen, und zwar durch das Produkt der einzelnen Normen $\|\mathbf{G}\| \|\mathbf{f}\|$.

Weiterhin gilt das small gain theorem (Anhang 10.2). Es besagt, dass der geschlossene Kreis aus linearem und nichtlinearem Teil sicherlich dann stabil ist, wenn der maximale Übertragungsfaktor von $\|\mathbf{e}\|$ nach $\|\mathbf{y}\|$ kleiner als Eins und außerdem der lineare Teil für sich genommen stabil ist.

Mit dem small gain theorem und der obigen Abschätzung für den maximalen Übertragungsfaktor von $\|\mathbf{e}\|$ nach $\|\mathbf{y}\|$ ergibt sich als hinreichende Bedingung für die Stabilität des aus nichtlinearem und linearem Teil bestehenden Regelkreises zum einen die Forderung nach der Stabilität von \mathbf{G} sowie die Bedingung

$$\|\mathbf{G}(j\omega)\| \|\mathbf{f}\| < 1 \quad (13.54)$$

Wählt man für die Normen jeweils die ∞ -Norm, so ergibt sich für ein Mehrgrößensystem

$$\|\mathbf{G}(j\omega)\|_{\infty} \|\mathbf{f}\|_{\infty} < 1 \quad (13.55)$$

und für ein Eingrößensystem mit (13.53)

$$\sup_{\omega} |G(j\omega)| \sup_{e \neq 0} \frac{|u|}{|e|} < 1 \quad (13.56)$$

Die Norm des linearen Teiles ist gerade der maximale Abstand der Ortskurve zum Ursprung, während die Norm des nichtlinearen Teiles dem betragsmäßig größtmöglichen Verstärkungsfaktor vom Eingang zum Ausgang des nichtlinearen Übertragungsgliedes entspricht.

Kann man für das nichtlineare Übertragungsverhalten eine obere und untere Sektorgrenze angeben wie beispielsweise für eine Kennlinie nach Abb. 13.23, so ist dieser Verstärkungsfaktor sicherlich kleiner als der maximale Betrag einer Sektorgrenze

$$\sup_{e \neq 0} \frac{|u|}{|e|} \leq \max \{|k_1|, |k_2|\} \quad (13.57)$$

Einsetzen in Gleichung (13.56) ergibt als neue, verschärfte Bedingung für die Stabilität des geschlossenen Kreises

$$\sup_{\omega} |G(j\omega)| \max\{|k_1|, |k_2|\} < 1 \quad (13.58)$$

bzw.

$$\sup_{\omega} |G(j\omega)| < \frac{1}{\max\{|k_1|, |k_2|\}} \quad (13.59)$$

Das System ist also stabil, wenn der Abstand der Ortskurve des stabilen, linearen Teiles vom Ursprung immer kleiner ist als der Kehrwert des maximalen Betrages einer Sektorgrenze. Demnach ist nur die Sektorgrenze ausschlaggebend, die den größeren Betrag aufweist. Dann kann man aber doch, ohne das Ergebnis der Ungleichung zu beeinflussen, die andere Sektorgrenze dahingehend verändern, dass gilt: $|k_1| = |k_2|$ und $k_1 < 0 < k_2$. Durch diese Maßnahme vergrößert sich der zulässige Sektor für das nichtlineare Übertragungsverhalten, ohne dass die Stabilitätsbedingung für den linearen Teil verschärft wird.

Die gleiche Überlegung lässt sich anstellen, wenn das nichtlineare Übertragungsverhalten bereits vorgegeben ist und durch einen Sektor $[k_1, k_2]$ mit $|k_1| \neq |k_2|$ begrenzt wird. Durch eine Sektortransformation von $[k_1, k_2]$ auf $[-k_d, k_d]$ mit $k_d = \frac{1}{2}|k_2 - k_1|$ (Abb. 13.26) ändert sich die rechte Seite der Ungleichung (13.59) zu $\frac{1}{k_d}$. Wegen $k_d < \max\{|k_1|, |k_2|\}$ ist sie größer geworden und die Bedingung für den linearen Teil damit nicht mehr so streng. Diese Bedingung soll im Folgenden hergeleitet werden.

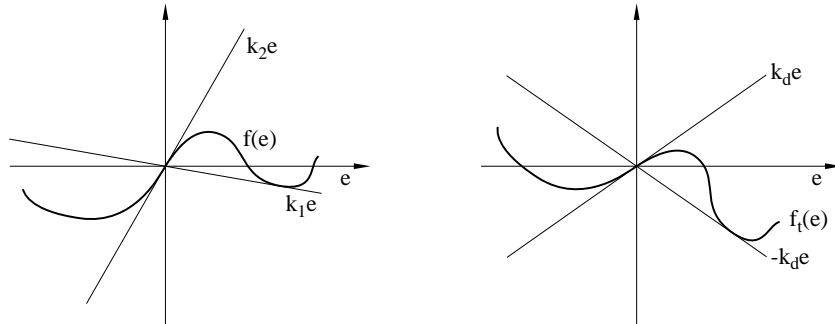


Abb. 13.26. Sektortransformation für das Kreiskriterium

Wie beim Popov-Kriterium erfolgt die Sektor-Transformation durch Einfügen zusätzlicher Proportionalglieder (vgl. Abb. 13.25), wobei hier der Sektor aber nicht um die untere Sektorgrenze k_1 , sondern um den Mittelwert $k_m = \frac{1}{2}(k_1 + k_2)$ verdreht wird. Mit $k_d = \frac{1}{2}|k_2 - k_1|$ wird das nichtlineare Übertragungsverhalten dann durch den symmetrischen Sektor $[-k_d, k_d]$ begrenzt, und der lineare Systemteil verändert sich (vgl. (13.50)) zu

$$G_t(s) = \frac{G(s)}{1 + k_m G(s)} \quad (13.60)$$

Aus Gleichung (13.59) wird dann die Bedingung

$$|G_t(j\omega)| < \frac{1}{k_d} \quad (13.61)$$

für alle ω . Zu beachten ist, dass der lineare Teil für das small gain theorem nun nicht mehr $G(s)$, sondern $G_t(s)$ ist und G_t daher eine stabile Übertragungsfunktion sein muss, während für G zunächst keine Vorgaben mehr bestehen. Weiterhin ist eine Berechnung von k_d und k_m nur für $k_2 < \infty$ möglich, weshalb der Fall $k_2 = \infty$ auszuschließen ist. Einsetzen für G_t und umstellen liefert dann

$$k_d |G(j\omega)| < |1 + k_m G(j\omega)| \quad (13.62)$$

Diese Ungleichung wird nun quadriert, wobei die Betragsquadrate durch Produkte der komplexen Größen mit ihren konjugiert komplexen Werten dargestellt werden:

$$0 < (k_m^2 - k_d^2) G(j\omega) \bar{G}(j\omega) + k_m (G(j\omega) + \bar{G}(j\omega)) + 1 \quad (13.63)$$

Mit $k_m^2 - k_d^2 = k_1 k_2$ ergibt sich

$$0 < k_1 k_2 G(j\omega) \bar{G}(j\omega) + k_m (G(j\omega) + \bar{G}(j\omega)) + 1 \quad (13.64)$$

Nun sind in Abhängigkeit der Vorzeichen von k_1 und k_2 verschiedene Fälle zu unterscheiden. Im ersten Fall sei $k_1 k_2 > 0$, d.h. beide Sektorgrenzen haben das gleiche Vorzeichen. Dann lässt sich die Ungleichung mit den Abkürzungen $r = \frac{1}{2} |\frac{1}{k_1} - \frac{1}{k_2}|$ und $m = -\frac{1}{2} (\frac{1}{k_1} + \frac{1}{k_2})$ umformen zu

$$|G(j\omega) - m| > r \quad (13.65)$$

Die Ortskurve muss also außerhalb eines Kreises mit dem Radius r und dem Mittelpunkt m verlaufen (Abb. 13.27 oben links). Für $k_1 < 0 < k_2$ erhält man mit denselben Abkürzungen

$$|G(j\omega) - m| < r \quad (13.66)$$

Die Ortskurve muss hier innerhalb des Kreises verlaufen (Abb. 13.27 oben rechts). Für $k_1 = 0, k_2 > 0$ entfällt der erste Term in (13.64), und es ergibt sich

$$\operatorname{Re}(G(j\omega)) > -\frac{1}{k_2} \quad (13.67)$$

Die Ortskurve muss also rechts von der durch $-\frac{1}{k_2}$ definierten Geraden verlaufen (Abb. 13.27 unten links). In analoger Weise ergibt sich für $k_1 < 0, k_2 = 0$ eine Gerade durch $-\frac{1}{k_1}$, von der aus gesehen die Ortskurve links verlaufen muss (Abb. 13.27 unten rechts).

In den letzten drei Fällen tritt aber noch ein weiteres Problem hinzu. Denn prinzipiell enthält wegen $0 \in [k_1, k_2]$ jeder von ihnen auch die Möglichkeit einer Kennlinie $f(e) = 0$. Wie schon für das Popov-Kriterium und das

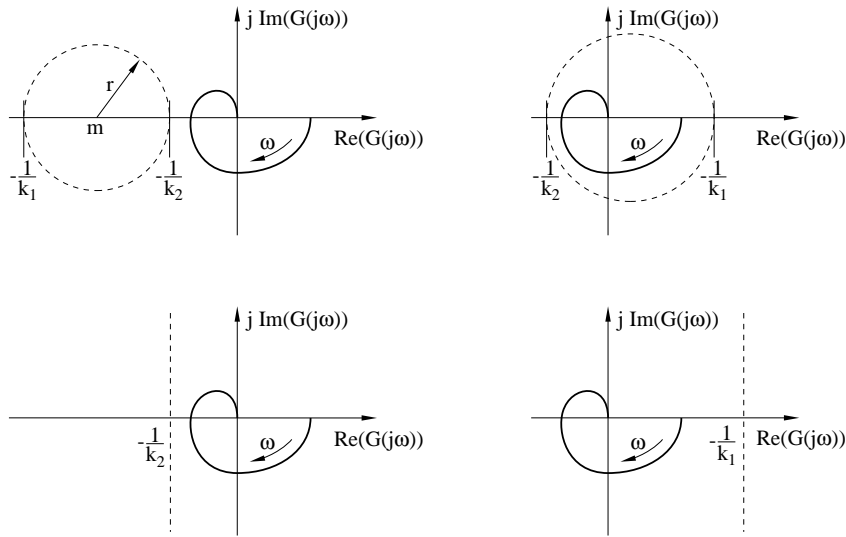


Abb. 13.27. Zum Kreiskriterium

small gain theorem diskutiert, würde damit der lineare Systemteil sich selbst überlassen bleiben. Stabilität des Gesamtsystems kann daher nur dann erreicht werden, wenn der lineare Systemteil für sich genommen stabil ist. Zur Forderung, dass G_t stabil ist und nur Pole mit negativem Realteil aufweist, tritt daher in diesen Fällen die Forderung, dass dies auch für G selbst gilt.

Für jede Konstellation von Sektorgrenzen k_1, k_2 lässt sich also ein verbotenes Gebiet $V(k_1, k_2)$ angeben, in dem die Ortskurve des linearen Teiles nicht verlaufen darf, damit der geschlossene Kreis stabil ist. Wenn man eine Gerade als einen Kreis mit unendlichem Radius ansieht, so ist dieses verbotene Gebiet immer kreisförmig. Daraus resultiert der Name des Kreiskriteriums:

Satz 13.7 (Kreiskriterium) Gegeben sei ein geschlossener Kreis, bestehend aus einem linearen und einem nichtlinearen Teil. Das nichtlineare Übertragungsverhalten sei durch einen Sektor $[k_1, k_2]$ mit $k_2 < \infty$ beschränkt. Falls k_1 und k_2 verschiedene Vorzeichen haben oder eine der beiden Sektorgrenzen gleich Null ist, muss die Übertragungsfunktion des linearen Teiles $G(s)$ stabil sein. Die Funktion $G_t(s) = \frac{G(s)}{1+k_m G(s)}$ mit $k_m = \frac{1}{2}(k_1 + k_2)$ muss immer stabil sein und darüber hinaus auch die Ordnung ihres Nennerpolynoms größer als die Ordnung des Zählerpolynoms. Wenn dann die Ortskurve $G(j\omega)$ für alle $\omega > 0$ außerhalb des durch k_1 und k_2 gegebenen, verbotenen Gebietes $V(k_1, k_2)$ verläuft, so besitzt der geschlossene Kreis eine global asymptotisch stabile Ruhelage für $u = y = 0$.

Interessant ist es, kurz den Zusammenhang zwischen dem Kreiskriterium und dem Popov-Kriterium sowie der Methode der harmonischen Balance aufzuzeigen. Wie bereits erwähnt, lässt sich das Kreiskriterium für statische Nichtlinearitäten auch aus dem Popov-Kriterium herleiten, indem man in

der Popov-Ungleichung (13.40) den freien Parameter q zu Null setzt, einige Umformungen vornimmt und das Ergebnis graphisch interpretiert (vgl. [13, 14]). Dieser Verzicht auf einen frei wählbaren Parameter bedeutet aber eine Verschärfung einer hinreichenden Stabilitätsbedingung, so dass das Kreiskriterium offenbar eine noch konservativere Aussage als das Popov-Kriterium darstellt. Es kann daher durchaus vorkommen, dass man die Stabilität eines Systems mit dem Kreiskriterium nicht nachweisen kann, wohl aber mit dem Popov-Kriterium. Instabilität kann man mit beiden Kriterien nicht nachweisen, da beide nur hinreichend, aber nicht notwendig sind.

Ähnliches gilt auch für den Zusammenhang zwischen dem Kreiskriterium und der Methode der harmonischen Balance. Hier lässt sich nachweisen, dass die für eine gegebene Kennlinie berechnete Kurve der Beschreibungsfunktion $-\frac{1}{N(A)}$ vollständig in dem mit dem Kreiskriterium ermittelten, verbotenen Gebiet $V(k_1, k_2)$ liegt [5]. Wird daher mit dem Kreiskriterium Stabilität nachgewiesen, d.h. verläuft die Ortskurve des linearen Teiles außerhalb des verbotenen Gebietes, so würde man auch mit der Methode der Beschreibungsfunktion Stabilität nachweisen, da die lineare Ortskurve und die Kurve der Beschreibungsfunktion offensichtlich keinen Schnittpunkt aufweisen können. In der anderen Richtung gilt diese Folgerung aber nicht. Denn wenn die Ortskurve die Kurve der Beschreibungsfunktion nicht schneidet bzw. die durch diese Kurve abgedeckte Fläche nicht berührt, so bedeutet dies noch lange nicht, dass sie auch außerhalb des wesentlich größeren, verbotenen Gebietes des Kreiskriteriums bleibt.

Das Kreiskriterium ist also das konservativste der drei Kriterien, dafür aber, da es im Gegensatz zu den beiden anderen Kriterien auch für dynamische Nichtlinearitäten gilt, das Kriterium mit dem größten Anwendungsbereich, wenn man von einigen Spezialfällen absieht, die im Popov-Kriterium noch enthalten sind. Darüber hinaus ist es offensichtlich von allen drei Kriterien am einfachsten anzuwenden. Die Sektorgrenzen eines nichtlinearen Übertragungsgliedes sind einfach zu bestimmen, und die Ortskurve des linearen Systemteiles kann man durch eine Messung erhalten. Es bietet sich daher im Anwendungsfall an, den Stabilitätsnachweis zunächst mit dem Kreiskriterium zu versuchen und nur im Falle eines Misserfolges die anderen Kriterien heranzuziehen.

Eine Übertragung des mit Hilfe des small gain theorem hergeleiteten Kreiskriteriums auf Mehrgrößensysteme ist nun kein Problem mehr, obwohl sich hier bei weitem kein so gut handhabbares Verfahren zur Überprüfung der Stabilität ergibt wie im Eingrößenfall. Ausgangspunkt ist die Ungleichung (13.55)

$$\|\mathbf{G}(j\omega)\|_\infty \|\mathbf{f}\|_\infty < 1 \quad (13.68)$$

Die Norm des nichtlinearen Systemteiles wird entsprechend Gleichung (13.52) bestimmt:

$$\|\mathbf{f}\|_\infty = \sup_{\mathbf{e} \neq \mathbf{0}} \frac{|\mathbf{u}|}{|\mathbf{e}|} \quad (13.69)$$

Eine relativ einfache und trotzdem genaue Abschätzung lässt sich durchführen, wenn für jede Komponente des Vektors \mathbf{u} gilt: $u_i = f_i(e_i)$. Jede dieser nichtlinearen Funktionen verlaufe in einem Sektor $[k_{i1}, k_{i2}]$. Dann fügt man entsprechend Abb. 13.28 zunächst eine Diagonalmatrix \mathbf{M} parallel zur Nichtlinearität ein, um die Sektoren in den einzelnen Komponenten jeweils für sich zu symmetrieren. Für die Elemente von \mathbf{M} muss damit gelten

$$m_{ii} = -\frac{1}{2}(k_{i1} + k_{i2}) \quad (13.70)$$

Anschließend wird noch eine Diagonalmatrix \mathbf{H} eingefügt, mit deren Komponenten die neu entstandenen Kennlinien und damit auch die symmetrischen Sektorgrenzen multipliziert werden. Wählt man

$$h_{ii} = \frac{2}{|k_{i2} - k_{i1}|} \quad (13.71)$$

so verläuft jede Kennlinie der neu entstandenen Nichtlinearität \mathbf{f}' im Sektor $[-1, 1]$. Das Verhältnis $\left|\frac{u'_i}{e_i}\right|$ ist damit für jedes i maximal gleich Eins, weshalb sich die Norm der Nichtlinearität nach (13.69) durch $\|\mathbf{f}'\|_\infty \leq 1$ abschätzen lässt.

Die Erweiterung der Nichtlinearität um \mathbf{M} und \mathbf{H} darf natürlich nicht erfolgen, ohne außerhalb von \mathbf{f}' , also im linearen Teil des Regelkreises, entsprechende Matrizen einzufügen, die die Wirkung von \mathbf{M} und \mathbf{H} gerade kompensieren. Denn sonst würde die Stabilitätsanalyse mit einem veränderten Regelkreis erfolgen, und die resultierenden Stabilitätsaussagen wären für das Originalsystem unbrauchbar. Abb. 13.28 zeigt, wie dies geschieht. \mathbf{H} wird durch die inverse Matrix \mathbf{H}^{-1} kompensiert, und \mathbf{M} durch eine andere Matrix \mathbf{M} , die mit entgegengesetztem Vorzeichen parallel geschaltet wird. Insgesamt sind damit die beiden Regelkreise in Abb. 13.28 äquivalent.

Für den unteren, erweiterten Regelkreis ergibt sich für das lineare Übertragungsverhalten von \mathbf{u}' nach \mathbf{e}

$$\mathbf{G}' = (\mathbf{I} - \mathbf{GM})^{-1} \mathbf{GH}^{-1} \quad (13.72)$$

und für die Stabilitätsforderung (13.68)

$$\|\mathbf{G}'(j\omega)\|_\infty \|\mathbf{f}'\|_\infty < 1 \quad (13.73)$$

Mit $\|\mathbf{f}'\|_\infty < 1$ wird daraus die Forderung

$$\|(\mathbf{I} - \mathbf{GM})^{-1} \mathbf{GH}^{-1}\|_\infty < 1 \quad (13.74)$$

Diese Ungleichung ist sicherlich erfüllt, wenn

$$\|(\mathbf{I} - \mathbf{GM})^{-1}\|_\infty \|\mathbf{GH}^{-1}\|_\infty < 1 \quad (13.75)$$

gilt. Aus (10.23) folgt sofort

$$\|(\mathbf{I} - \mathbf{GM})^{-1}\|_{\infty} = \frac{1}{\|\mathbf{I} - \mathbf{GM}\|_{\infty}} \quad (13.76)$$

und damit

$$\|\mathbf{GH}^{-1}\|_{\infty} < \|\mathbf{I} - \mathbf{GM}\|_{\infty} \quad (13.77)$$

Da die Berechnung der ∞ -Norm mittlerweile in jedem regelungstechnischen Software-Tool enthalten ist, lässt sich diese Bedingung quasi auf Knopfdruck überprüfen. Falls eine algebraische Lösung angestrebt wird, kann man die ∞ -Norm auch durch andere, leichter zu berechnende Normen abschätzen (vgl. [5]). Eine solche Abschätzung kann allerdings sehr grob sein. Abschließend muss dann noch wie im Eingrößenfall die Stabilität von \mathbf{G} und \mathbf{G}' nachgewiesen werden, was aber ein rein lineares Problem und somit nicht besonders schwierig ist.

13.8 Hyperstabilität

13.8.1 Positiv reelle Systeme

Bevor auf den Begriff der Hyperstabilität eingegangen wird, ist zunächst ein anderer Begriff zu klären:

Definition 13.8 Ein lineares Eingrößensystem ist genau dann streng positiv reell, wenn seine Übertragungsfunktion nur Pole mit negativem Realteil aufweist und $\operatorname{Re}(G(j\omega)) > 0$ für $\omega \geq 0$ gilt.

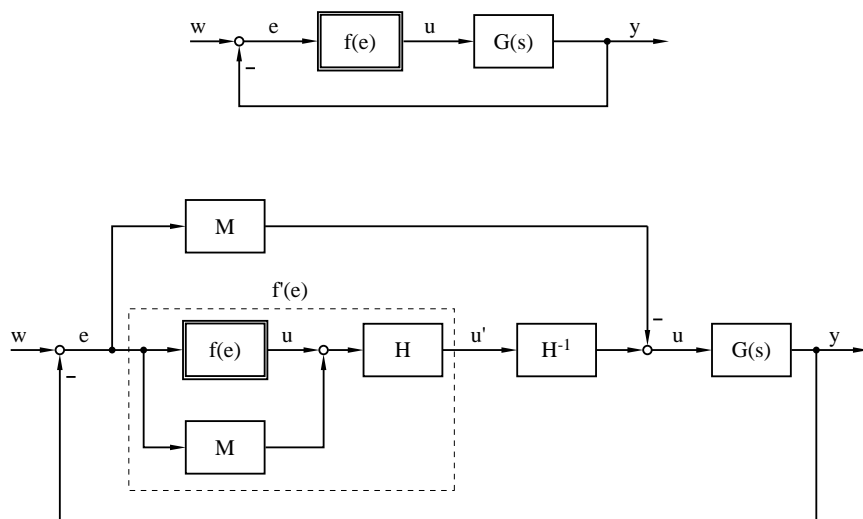


Abb. 13.28. Sektortransformation im Mehrgrößenfall

Ein Mehrgrößensystem mit quadratischer Übertragungsmatrix $\mathbf{G}(s)$ ist genau dann streng positiv reell, wenn die Elemente $G_{ij}(s)$ der Matrix ausschließlich Polstellen mit negativem Realteil aufweisen und außerdem die hermitesche Matrix

$$\mathbf{H}(j\omega) = \frac{1}{2}(\mathbf{G}(j\omega) + \bar{\mathbf{G}}^T(j\omega)) \quad (13.78)$$

für alle $\omega \geq 0$ positiv definit ist, d.h. ausschließlich positive Eigenwerte aufweist.

Die Eigenschaft “positiv reell” ist demnach eine Steigerung der Eigenschaft “stabil” und zudem nur für Systeme mit quadratischer Übertragungsmatrix definiert, d.h. für Systeme mit gleicher Anzahl an Eingangs- und Ausgangsgrößen.

Ohne Beweis sei der folgende Satz angegeben, der die Überprüfung der Eigenschaft erleichtert, wenn das System in Zustandsdarstellung angegeben ist:

Satz 13.9 *Ein durch*

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{aligned} \quad (13.79)$$

gegebenes lineares System ist genau dann streng positiv reell, wenn die folgenden Bedingungen erfüllt sind:

- Das lineare System muss vollständig steuer- und beobachtbar sein.
- Es muss Matrizen \mathbf{L} , \mathbf{P} und \mathbf{V} geeigneter Dimension geben mit

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{L} \mathbf{L}^T \quad (13.80)$$

$$\mathbf{L} \mathbf{V} = \mathbf{C}^T - \mathbf{P} \mathbf{B} \quad (13.81)$$

$$\mathbf{D} + \mathbf{D}^T = \mathbf{V}^T \mathbf{V} \quad (13.82)$$

- $\text{grad}(\mathbf{L}) = \text{grad}(\mathbf{A}) = n$
- \mathbf{P} ist symmetrisch und positiv definit: $\mathbf{P} = \mathbf{P}^T$ und $\mathbf{P} > 0$

Anmerkung: Aus $\text{grad}(\mathbf{L}) = n$ folgt hier, dass $\mathbf{L} \mathbf{L}^T$ eine symmetrische, positiv definite Matrix ist. Weiterhin ist \mathbf{P} nach Voraussetzung ebenfalls positiv definit. Damit folgt aus (13.80) und Satz A.2, dass \mathbf{A} nur Eigenwerte mit negativem Realteil aufweist. Auch aus diesem Satz lässt sich also ableiten, dass die Stabilität eine Voraussetzung dafür ist, dass das System positiv reell ist.

13.8.2 Herleitung

Um das nächste Stabilitätskriterium vorstellen zu können, muss zunächst ein neuer, strengerer Stabilitätsbegriff als der von Ljapunov eingeführt werden, und zwar die *Hyperstabilität* [51, 52]:

Definition 13.10 Gegeben sei ein lineares, steuer- und beobachtbares System mit dem Zustandsvektor \mathbf{x} , dessen Eingangsvektor $\mathbf{u}(t)$ und Ausgangsvektor $\mathbf{y}(t)$ dieselbe Dimension haben. $\mathbf{x}(0)$ ist der Anfangszustand. Damit ist $\mathbf{y}(t)$ von $\mathbf{x}(0)$ und $\mathbf{u}(t)$ abhängig. Das System heißt hyperstabil, wenn für jeden Anfangszustand, jeden Eingangsvektor und jedes $\beta_0 > 0$ aus der Integralungleichung

$$\int_0^T \mathbf{u}^T \mathbf{y} dt \leq \beta_0^2 \quad (13.83)$$

für alle $T > 0$ die Ungleichung $|\mathbf{x}(t)| \leq \beta_0 + \beta_1 |\mathbf{x}(0)|$ für alle $0 < t < T$ mit einer beliebigen positiven Konstanten β_1 folgt. Konvergiert darüber hinaus der Zustandsvektor gegen Null, $\lim_{t \rightarrow \infty} \mathbf{x}(t) = 0$, so heißt das System asymptotisch hyperstabil.

Die Idee dieser Definition lautet: Wenn das Produkt aus Ein- und Ausgangsgrößen eines hyperstabilen Systems in einem gewissen Sinne beschränkt ist, so bleiben auch die Zustandsgrößen beschränkt. Die Voraussetzung gleicher Dimension für die Ein- und Ausgangsgröße ist notwendig, weil das Produkt $\mathbf{u}^T \mathbf{y}$ sonst nicht gebildet werden kann.

Interessant ist ein Vergleich dieser Definition mit den bisher verwendeten Stabilitätsdefinitionen. Die ersten beiden Stabilitätsdefinitionen 5.1 (endliche Sprungantwort) und 5.2 (BIBO-Stabilität) bezogen sich auf die Reaktion des Systemausgangs auf eine Eingangsgröße, während die Definition nach Ljapunov 13.2 das interne Verhalten des Systems (Zustandsgrößen) ohne äußere Anregung als Reaktion auf einen Anfangszustand betrachtete. Dagegen werden bei der Hyperstabilität sowohl der Anfangszustand als auch eine äußere Anregung in Betracht gezogen.

Offensichtlich ist ein (asymptotisch) hyperstabiles System auch (asymptotisch) stabil nach Ljapunov. Denn die Ungleichung (13.83) ist sicherlich erfüllt für einen Eingangsvektor $\mathbf{u}(t) = \mathbf{0}$, d.h. für ein System ohne äußere Anregung. In einem hyperstabilen System ist dann auch der Zustandsvektor beschränkt durch $|\mathbf{x}(t)| \leq \beta_0 + \beta_1 |\mathbf{x}(0)|$. Damit ist das System aber auch stabil nach Ljapunov. Und die Verschärfung hinsichtlich asymptotischer Stabilität ist sowieso in beiden Definitionen gleich. Ein lineares System, das stabil nach Ljapunov ist, ist aber auch stabil nach den Definitionen 5.1 und 5.2, wie bereits früher gezeigt wurde. Von allen vorgestellten Stabilitätsdefinitionen ist daher die Hyperstabilität die strengste. Dies kann man auch daran erkennen, dass beispielsweise die Rückkopplung zweier hyperstabiler Systeme \mathbf{H}_1 und \mathbf{H}_2 gemäß Abb. 13.29 wieder ein hyperstabiles System mit der Eingangsgröße \mathbf{w} und der Ausgangsgröße \mathbf{y} ergibt, wie sich beweisen lässt. Die Rückkopplung zweier Ljapunov-stabiler Systeme muss dagegen nicht zwangsläufig wieder auf ein Ljapunov-stabiles System führen.

Anhand der Integralungleichung sieht man, dass die Hyperstabilität in gewissem Sinne eine Erweiterung der im Popov-Kriterium erwähnten absoluten Stabilität ist. Absolute Stabilität nach Satz 13.6, beispielsweise im

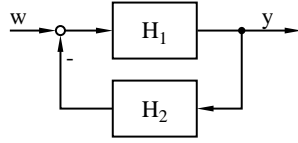


Abb. 13.29. Rückkopplung zweier hyperstabiler Systeme

Sektor $[0, \infty)$, setzt voraus, dass die Kennlinie $f(e)$ in genau diesem Sektor verläuft. Diese Bedingung lässt sich aber auch ausdrücken durch die Forderung $f(e)e \geq 0$. Mit $f(e) = u$ und $e = -y$ wird daraus $uy \leq 0$. Der Zusammenhang mit der Integralungleichung ist deutlich zu erkennen.

Der Begriff der Hyperstabilität lässt sich auch energetisch deuten. So lassen sich u und y beispielsweise als Strom und Spannung am Eingang einer elektrischen Schaltung interpretieren und die Zustandsgröße x als interner Energiespeicher, zum Beispiel die Spannung an einem Kondensator. Das Produkt aus u und y entspricht dann der zugeführten elektrischen Leistung, und das Integral über diesem Produkt der zugeführten elektrischen Energie. Wenn diese gemäß der Integralungleichung beschränkt ist, so muss, sofern die Schaltung hyperstabil ist, auch die intern gespeicherte Energie und damit x beschränkt sein. Für passive elektrische Schaltungen trifft dieser Sachverhalt immer zu, sie sind demnach hyperstabil. Enthält eine Schaltung aber aktive Bauteile wie z.B. Verstärker, so ist die Hyperstabilität nicht unbedingt gegeben.

Anzumerken ist noch, dass die hier angegebene Definition eine stark vereinfachte und enger gefasste Version der allgemeinen Definition (vgl. [49, 52]) ist, die sich auf nichtlineare, zeitvariante Systeme bezieht, wobei dort zudem noch die Beträge durch verallgemeinerte Funktionen ersetzt sind. Für solch allgemeine Systeme ergeben sich dann aber keine praktisch anwendbaren Stabilitätskriterien mehr.

Stattdessen soll hier, ausgehend von der eng gefassten Definition für lineare Systeme, ein Stabilitätskriterium für den Standardregelkreis nach Abb. 13.15 entwickelt werden. Die betrachtete Ruhelage sei $\mathbf{w} = \mathbf{y} = \mathbf{0}$, andernfalls ist das System geeignet umzudefinieren. Erfüllt nun der nichtlineare Teil die Ungleichung

$$\int_0^T \mathbf{u}^T \mathbf{e} \, dt \geq -\beta_0^2 \quad (13.84)$$

für alle $T > 0$, dann erfüllen wegen $\mathbf{e} = -\mathbf{y}$ offensichtlich die Ein- und Ausgangsgrößen \mathbf{u} und \mathbf{y} des linearen Systemteiles auch die Voraussetzung (13.83) aus Definition 13.10:

$$-\int_0^T \mathbf{u}^T \mathbf{y} \, dt \geq -\beta_0^2$$

$$\int_0^T \mathbf{u}^T \mathbf{y} \, dt \leq \beta_0^2 \quad (13.85)$$

Wenn dann noch gezeigt werden kann, dass der lineare Systemteil hyperstabil ist, so ist garantiert, dass seine Zustandsgrößen beschränkt bleiben, und zwar unabhängig vom internen Verhalten des nichtlinearen Systemteiles. Im Falle asymptotischer Hyperstabilität konvergieren die Zustandsgrößen sogar gegen Null. Wenn man darüber hinaus fordert, dass der nichtlineare Systemteil keine internen Zustandsgrößen enthält, so kann es keine Zustandsgrößen im Gesamtsystem geben, die nicht gegen Null konvergieren. Das bedeutet aber doch, dass das Gesamtsystem in der Ruhelage $\mathbf{x} = \mathbf{0}$ asymptotisch stabil im Ljapunovschen Sinne ist. Damit gilt

Satz 13.11 *Der nichtlineare Standardregelkreis hat für den Sollwert $\mathbf{w} = \mathbf{0}$ die (asymptotisch) stabile Ruhelage $\mathbf{x} = \mathbf{0}$, wenn das lineare Teilsystem (asymptotisch) hyperstabil ist und das statische, nichtlineare Teilsystem für alle $T > 0$ die Integralungleichung*

$$\int_0^T \mathbf{u}^T \mathbf{e} \, dt \geq -\beta_0^2 \quad (13.86)$$

erfüllt.

Was ist aber zu tun, wenn der nichtlineare Teil nicht statisch ist, sondern ebenfalls interne Zustandsgrößen enthält? Wie schon gesagt, haben interne Vorgänge im nichtlinearen Teil keinen Einfluss auf die Beschränktheit der Zustandsgrößen des linearen Teiles, sofern nur die Ungleichung (13.86) eingehalten wird. Damit ein Standardregelkreis mit einem dynamischen nichtlinearen Teil asymptotisch stabil ist, muss daher lediglich neben den Bedingungen aus Satz 13.11 sichergestellt sein, dass die Zustandsgrößen des nichtlinearen Teiles gegen Null konvergieren. Dafür gibt es aber kein einfach anzuwendendes, allgemeingültiges Kriterium, oft jedoch ermöglicht eine vergleichsweise einfache dynamische Struktur des nichtlinearen Teiles eine Abschätzung des Zustandsgrößenverlaufes gewissermaßen von Hand. Und falls die Zustandsgrößen des nichtlinearen Teiles unter technischen Gesichtspunkten sowieso nicht von Interesse sind, kann man auf diese Betrachtung auch völlig verzichten. Man darf dann allerdings nicht mehr von der asymptotischen Stabilität des gesamten Systems sprechen, sondern nur noch davon, dass die Zustandsgrößen des linearen Teiles für den gegebenen Sollwert gegen Null konvergieren.

Es stellt sich nun die Frage, wie im Anwendungsfall vorzugehen ist. Oft wird schon die Forderung nach gleicher Dimension der Vektoren \mathbf{u} und \mathbf{y} das erste Problem darstellen, weil dies in vielen Fällen nicht von vornherein gegeben ist. Meist weist der nichtlineare Teil (z.B. ein Fuzzy-Regler) mehr Ein- als Ausgangsgrößen auf. Um hier gleiche Dimension zu gewährleisten,

müssen für den nichtlinearen Systemteil zusätzliche Ausgangsgrößen mit dem konstanten Wert Null definiert werden. Entsprechend ist die Anzahl der Eingangsgrößen des linearen Systemteiles zu erhöhen und dessen Übertragungsmatrix zu verändern.

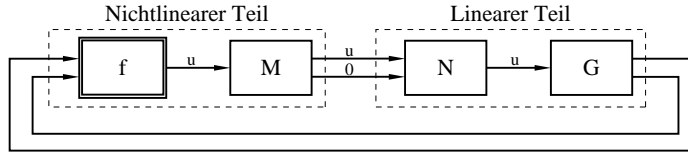


Abb. 13.30. Einfügen zusätzlicher Matrizen zur Herstellung gleicher Dimension von Ein- und Ausgangsvektoren

Die Definition zusätzlicher Ausgangsgrößen entspricht dem Einfügen zweier Matrizen \mathbf{M} und \mathbf{N} in den geschlossenen Kreis (Abb. 13.30). Um dabei das System nicht zu verändern, muss die Bedingung $\mathbf{NM} = \mathbf{I}$ erfüllt sein. In Abb. 13.30 gilt für die Matrizen \mathbf{N} und \mathbf{M}

$$\mathbf{N} = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \mathbf{M} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (13.87)$$

Dadurch wird aus der Ausgangsgröße u des nichtlinearen Systemteiles der Ausgangsvektor $\mathbf{u} = [u, 0]^T$ und aus der Übertragungsmatrix

$$\mathbf{G}(s) = \begin{pmatrix} G_1(s) \\ G_2(s) \end{pmatrix} \quad (13.88)$$

die quadratische Übertragungsmatrix

$$\mathbf{G}(s)\mathbf{N} = \begin{pmatrix} G_1(s) & 0 \\ G_2(s) & 0 \end{pmatrix} \quad (13.89)$$

Damit weisen beide Systemteile die gleiche Anzahl an Ein- und Ausgangsgrößen auf. Im Folgenden wird auf eine explizite Darstellung der Matrizen \mathbf{M} und \mathbf{N} verzichtet, d.h. sowohl \mathbf{f} als auch \mathbf{G} gelten als entsprechend erweiterte Systemteile.

Nun soll zunächst das lineare Teilsystem auf Hyperstabilität überprüft werden. Dazu wird der folgende Satz benötigt, der hier aber nicht bewiesen werden soll:

Satz 13.12 *Ein lineares, zeitinvariantes, steuer- und beobachtbares System ist genau dann asymptotisch hyperstabil, wenn es streng positiv reell ist (vgl. Kap. 13.8.1).*

Wie im Anschluss an Definition 13.8 schon erwähnt, ist damit für die Hyperstabilität des linearen Systemteiles zunächst einmal Voraussetzung, dass dieser stabil ist. Sollte dies nicht der Fall sein, so kann man versuchen, durch

eine Transformation ein stabiles System zu erzeugen. Zu diesem Zweck wird für den linearen Teil eine Rückkopplungsmatrix \mathbf{K} eingefügt, wie es in Abb. 13.31 dargestellt ist. Die Matrix \mathbf{D} sei zunächst Null. Dann ergibt sich durch die Rückkopplung ein neues, lineares System $\mathbf{G}' = (\mathbf{I} + \mathbf{G}\mathbf{K})^{-1}\mathbf{G}$, dessen Eigenwerte unter gewissen Voraussetzungen bei geeigneter Wahl von \mathbf{K} alle einen negativen Realteil aufweisen.

In der Zustandsdarstellung wird die Wirkung von \mathbf{K} noch etwas deutlicher:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{u} - \mathbf{K}\mathbf{y}) \\ &= (\mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C})\mathbf{x} + \mathbf{B}\mathbf{u} \quad \text{mit } \mathbf{y} = \mathbf{C}\mathbf{x} \quad \text{und } \mathbf{D} = \mathbf{0} \\ \mathbf{A}' &= \mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C}\end{aligned}\tag{13.90}$$

Verändert wird also die Systemmatrix, und die Stabilisierung kann nur bei einer bestimmten Struktur von \mathbf{A} , \mathbf{B} und \mathbf{C} gelingen.

Vorausgesetzt wurde bei dieser Darstellung, dass im ursprünglichen System kein direkter Durchgriff von der Stell- zur Ausgangsgröße besteht ($\mathbf{D} = \mathbf{0}$). Grundsätzlich kann das Verfahren aber auch bei direktem Durchgriff angewendet werden. Die Gleichungen werden dann lediglich etwas aufwendiger.

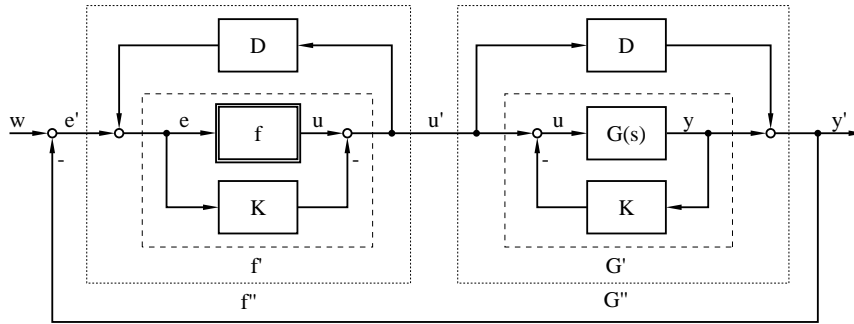


Abb. 13.31. Erweiterung des linearen Systemteiles zur Gewährleistung der Hyperstabilität

Die durch das Hinzufügen von \mathbf{K} erfolgte Veränderung des Gesamtsystems muss aber an anderer Stelle wieder rückgängig gemacht werden, da sonst die Stabilitätsanalyse anhand eines veränderten Regelkreises erfolgen würde (Abb. 13.31). Man kann sich leicht klarmachen, dass durch das Hinzufügen von \mathbf{K} die ursprüngliche Eingangsgröße \mathbf{u} des linearen Teiles um den additiven Term $-\mathbf{K}\mathbf{y}$ verändert wird. Eine zusätzlich über den nichtlinearen Systemteil \mathbf{f} parallel geschaltete Matrix \mathbf{K} hebt diese Wirkung wegen $-\mathbf{K}\mathbf{e} = +\mathbf{K}\mathbf{y}$ aber gerade wieder auf, so dass das abgebildete, erweiterte System (mit $\mathbf{D} = \mathbf{0}$) gerade dem ursprünglichen Originalsystem entspricht. Für die Stabilitätsanalyse wird demnach der lineare Teil \mathbf{G} durch \mathbf{G}' und der nichtlineare Teil \mathbf{f} durch \mathbf{f}' ersetzt. Diese Systemtransformation ist der Sek-

tortransformation beim Popov- und Kreiskriterium vergleichbar. Man passt einen gegebenen Standardregelkreis durch Transformation an die Voraussetzungen des anzuwendenden Stabilitätskriteriums an. Kann dann für das transformierte System Stabilität nachgewiesen werden, so gilt dies auch für das Originalsystem.

Es liege nun ein stabiles, lineares System $\mathbf{G}'(s)$ vor. Nun ist nach Definition 13.8 zu prüfen, ob die Matrix

$$\mathbf{H}'(j\omega) = \frac{1}{2}(\mathbf{G}'(j\omega) + \bar{\mathbf{G}}'^T(j\omega)) \quad (13.91)$$

für alle Frequenzen ω ausschließlich positive Eigenwerte aufweist. Dies kann numerisch durchgeführt werden. Wegen der Frequenzabhängigkeit von \mathbf{H}' ergibt sich für jeden Eigenwert eine frequenzabhängige Kurve. Sollte diese Kurve für jeden Eigenwert im Positiven verlaufen, so ist das lineare System \mathbf{G}' streng positiv reell.

Andernfalls ist wiederum eine Systemtransformation notwendig (Abb. 13.31). Ziel dieser Transformation ist, den linearen Systemteil durch Parallelschaltung einer Diagonalmatrix streng positiv reell zu machen, wobei diese Diagonalmatrix aber möglichst kleine Elemente haben soll. Denn je kleiner die Elemente, desto größer sind die zulässigen Sektoren für das Übertragungsverhalten des nichtlinearen Systemteiles, wie später noch gezeigt wird.

Zu ermitteln ist zunächst der kleinste auftretende Wert $d < 0$ aller Eigenwerte von \mathbf{H}' über ω . Die Addition einer Matrix $\mathbf{D} = |d|\mathbf{I}$ zu \mathbf{G}' führt dann auf das System $\mathbf{G}'' = \mathbf{G}' + \mathbf{D}$ mit der zugeordneten Matrix

$$\begin{aligned} \mathbf{H}''(j\omega) &= \frac{1}{2}(\mathbf{G}''(j\omega) + \bar{\mathbf{G}}''^T(j\omega)) = \frac{1}{2}(\mathbf{G}'(j\omega) + \mathbf{D} + \bar{\mathbf{G}}'^T(j\omega) + \mathbf{D}) \\ &= \mathbf{H}'(j\omega) + |d|\mathbf{I} \end{aligned} \quad (13.92)$$

Offensichtlich sind die Eigenwerte von \mathbf{H}'' gegenüber denen von \mathbf{H}' um $|d|$ nach rechts verschoben und deshalb alle positiv. Da \mathbf{G}'' zudem dieselben, stabilen Pole aufweist wie \mathbf{G}' , ist das erweiterte System \mathbf{G}'' damit streng positiv reell.

Möglich ist auch, die Erweiterung mit einer beliebigen, positiv semidefiniten Diagonalmatrix \mathbf{D} durchzuführen, deren Elemente nicht alle gleich sind. Doch in dem Fall kann kein direkter Zusammenhang zwischen diesen Elementen und der Verschiebung der Eigenwerte von \mathbf{H}' angegeben werden. Dies kann wiederum die Bestimmung der Matrix \mathbf{D} sehr schwierig und zeitaufwendig machen. Um größere zulässige Sektoren für einzelne nichtlineare Kennlinien zu erhalten, kann eine unterschiedliche Wahl der Diagonalelemente jedoch manchmal notwendig sein.

Die Diagonalmatrix \mathbf{D} , durch deren Einfügen der lineare Systemteil streng positiv reell wird, lässt sich auch anhand der Zustandsdarstellung des Systems und Satz 13.9 berechnen. Dazu wird zunächst eine Matrix \mathbf{L} mit $\text{grad}(\mathbf{L}) = n$ beliebig festgelegt. Mit \mathbf{L} und gegebener Systemmatrix \mathbf{A}' lässt sich dann aus der Ljapunov-Gleichung (13.80) eine Matrix \mathbf{P} berechnen.

Da es sich bei \mathbf{A}' um die Systemmatrix des stabilen Systems \mathbf{G}' handelt, sind sämtliche Eigenwerte von \mathbf{A}' negativ. Aus $\text{grad}(\mathbf{L}) = n$ folgt, wie bereits im Anschluss an Satz 13.9 skizziert, dass $\mathbf{L}\mathbf{L}^T$ eine symmetrische, positiv definite Matrix ist. Damit folgt aus Satz A.2, dass \mathbf{P} positiv definit ist und die Voraussetzung aus Satz 13.9 erfüllt.

Wegen der Regularität ist \mathbf{L} invertierbar, und \mathbf{V} ergibt sich aus Gleichung (13.81) zu

$$\mathbf{V} = \mathbf{L}^{-1}(\mathbf{C}'^T - \mathbf{P}\mathbf{B}') \quad (13.93)$$

Da schließlich \mathbf{D} eine Diagonalmatrix sein soll, kann ihre Symmetrie vorausgesetzt werden: $\mathbf{D} = \mathbf{D}^T$. Damit lässt sich aber Gleichung (13.82) zur Bestimmung von \mathbf{D} umformen:

$$\mathbf{D} = \frac{1}{2}\mathbf{V}^T\mathbf{V} \quad (13.94)$$

Gemäß Abb. 13.31 wird diese Matrix zum stabilen linearen Systemteil $\mathbf{G}'(s)$ parallel geschaltet. Das entstehende System $\mathbf{G}''(s)$ erfüllt wegen der Anwendung der Gleichungen (13.80) - (13.82) zur Berechnung von \mathbf{D} sicherlich die Voraussetzungen aus Satz 13.9 und ist damit streng positiv reell.

Im Gegensatz zum vorherigen Ansatz ist bei diesem Verfahren aber nicht gewährleistet, dass die Diagonalelemente von \mathbf{D} so klein wie möglich sind. Im Hinblick auf die weitere Verwendung von \mathbf{D} ist der vorherige Ansatz daher vorzuziehen.

Die durch das Hinzufügen von \mathbf{D} erfolgte Veränderung des Gesamtsystems muss nun an anderer Stelle wieder kompensiert werden. Jetzt ist es so, dass durch das Hinzufügen von \mathbf{D} die ursprüngliche Eingangsgröße des nichtlinearen Teiles \mathbf{e} um den additiven Term $-\mathbf{D}\mathbf{u}'$ verändert wird. Diese Wirkung kann durch eine Rückkopplung mit \mathbf{D} über den nichtlinearen Teil \mathbf{f}' aufgehoben werden. Damit entspricht das abgebildete, erweiterte System gerade wieder dem ursprünglichen Originalsystem. Insgesamt wird also für die Stabilitätsanalyse der lineare Teil \mathbf{G} durch \mathbf{G}'' und der nichtlineare Teil \mathbf{f} durch \mathbf{f}'' ersetzt. Falls für das transformierte System Stabilität nachgewiesen werden kann, so gilt dies auch für das Originalsystem.

Bevor der letzte und entscheidende Schritt der Stabilitätsanalyse vorgestellt wird, sollen zunächst noch einmal alle bisherigen Transformationen aufgelistet werden:

- Hinzufügen zweier Matrizen \mathbf{N} und \mathbf{M} , um gleiche Dimension der Vektoren \mathbf{u} und \mathbf{e} bzw. \mathbf{y} zu erreichen.
- Hinzufügen einer Rückführmatrix \mathbf{K} zur Stabilisierung des linearen Teiles.
- Hinzufügen einer parallelgeschalteten Diagonalmatrix \mathbf{D} , um den linearen Teil positiv reell zu machen.

Nach den Transformationen ist der lineare Teil \mathbf{G}'' des transformierten Systems sicher asymptotisch hyperstabil. Somit muss zum Nachweis der Stabilität des geschlossenen Kreises jetzt noch gezeigt werden, dass der erweiterte nichtlineare Systemteil \mathbf{f}'' die Ungleichung

$$\int_0^T \mathbf{u}'^T \mathbf{e}' dt \geq -\beta_0^2 \quad (13.95)$$

bzw. mit $\mathbf{u}' = \mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e}$ und $\mathbf{e}' = \mathbf{e} - \mathbf{D}\mathbf{u}'$

$$\int_0^T [\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e}]^T [\mathbf{e} - \mathbf{D}(\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e})] dt \geq -\beta_0^2 \quad (13.96)$$

erfüllt.

Hinreichend dafür ist auf jeden Fall, wenn jeweils die i -te Komponente beider Vektoren im Integranden dasselbe Vorzeichen aufweist. Zunächst soll der Fall untersucht werden, dass beide positiv sind. Aus

$$\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e} \geq 0 \quad (13.97)$$

muss demnach folgen

$$e_i - d_{ii}(\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}) \geq 0 \quad (13.98)$$

bzw.

$$e_i \geq d_{ii}(\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}) \quad (13.99)$$

Da sowohl d_{ii} als auch der Ausdruck in der Klammer auf der rechten Seite immer positiv sind, ist e_i ebenfalls positiv. Damit kann die Ungleichung durch e_i und d_{ii} dividiert werden, und man erhält

$$\frac{1}{d_{ii}} \geq \frac{\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}}{e_i} \quad (13.100)$$

Nun sei angenommen, dass die i -te Komponente beider Vektoren im Integranden negativ ist. Aus

$$\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e} < 0 \quad (13.101)$$

muss jetzt folgen

$$e_i - d_{ii}(\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}) < 0 \quad (13.102)$$

bzw.

$$e_i < d_{ii}(\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}) \quad (13.103)$$

Da nun d_{ii} positiv und der Ausdruck in der Klammer auf der rechten Seite negativ ist, ist e_i ebenfalls negativ. Bei Division der Ungleichung durch e_i und d_{ii} dreht sich nun das Ungleichheitszeichen um, und man erhält wieder Ungleichung (13.100). Insgesamt bedeutet dies, dass zum einen Ungleichung (13.100) gelten muss, zum anderen aber auch e_i und $(\mathbf{f}_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e})$ das gleiche Vorzeichen haben müssen.

Dies führt auf die Sektorbedingung

$$\begin{aligned}
0 \leq \frac{f_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e}}{e_i} &\leq \frac{1}{d_{ii}} && \text{falls } e_i \neq 0 \\
f_i(\mathbf{e}) - \mathbf{k}_i^T \mathbf{e} &= 0 && \text{falls } e_i = 0
\end{aligned} \tag{13.104}$$

für alle i mit $d_{ii} > 0$. \mathbf{k}_i^T ist der i -te Zeilenvektor von \mathbf{K} . Für $d_{ii} = 0$ ergibt sich als obere Sektorgrenze ∞ . Man sieht, dass der zulässige Sektor umso größer ist, je kleiner d_{ii} gewählt wurde, und wie wichtig es daher ist, \mathbf{D} so zu wählen, dass ihre Elemente möglichst klein sind.

Falls das lineare System von vornherein stabil ist, entfällt die Transformation mit der Matrix \mathbf{K} , und die Sektorbedingungen lauten

$$\begin{aligned}
0 \leq \frac{f_i(\mathbf{e})}{e_i} &\leq \frac{1}{d_{ii}} && \text{falls } e_i \neq 0 \\
f_i(\mathbf{e}) &= 0 && \text{falls } e_i = 0
\end{aligned} \tag{13.105}$$

In der Praxis werden die Bedingungen (13.104) bzw. (13.105) nur in sehr einfachen Fällen analytisch überprüft werden können. Im Normalfall geht dies nur auf numerischem Wege. Man wird dann eine ausreichend große und repräsentative Menge aus der Menge aller Fehlervektoren \mathbf{e} festlegen und für jeden einzelnen Vektor die Bedingungen überprüfen müssen. Noch besser ist aber, statt der konservativen Abschätzung (13.104) bzw. (13.105) direkt den Integranden aus (13.96) auszuwerten. Wenn dieser Integrand für jeden Vektor \mathbf{e} aus der repräsentativen Menge von Fehlervektoren positiv ist, dann ist auch Bedingung (13.96) sicher erfüllt.

Ein ganz einfaches Beispiel soll nun die Anwendung des Hyperstabilitätskriteriums verdeutlichen: Gegeben sei ein Eingrößensystem, dessen nicht-linearer Teil aus einer Multiplikation von $e(t)$ mit einer zeitabhängigen Verstärkung $k(t)$ besteht (Abb. 13.32 oben). Die Übertragungsfunktion $G(s)$ des linearen Teiles sei streng positiv reell und damit asymptotisch hyperstabil. Um die asymptotische Stabilität des Regelkreises in der Ruhelage $w = u = y = 0$ nachzuweisen, muss daher nur noch die Integralungleichung (13.86) betrachtet werden. Mit $u = ke$ ergibt sich

$$\int_0^T u(t)e(t)dt = \int_0^T k(t)e^2(t)dt \geq -\beta_0^2 \tag{13.106}$$

Die Ungleichung ist sicher dann erfüllt, wenn der Integrand positiv ist, d.h. wenn $k(t) \geq 0$ gilt.

Nun sei die Voraussetzung dahingehend abgeändert, dass die lineare Übertragungsfunktion zwar asymptotisch stabil, aber nicht hyperstabil ist. Sie besitzt demnach ausschließlich Pole mit negativem Realteil, ist aber nicht streng positiv reell. Laut Definition 13.8 bedeutet dies, dass der Realteil des Frequenzganges $G(j\omega)$ nicht nur positive Werte aufweist, d.h. ein Teil der Ortskurve verläuft in der linken Hälfte der komplexen Ebene. Abb. 13.33 und der untere Teil der Abb. 13.32 verdeutlichen, wie in diesem Fall vorzugehen

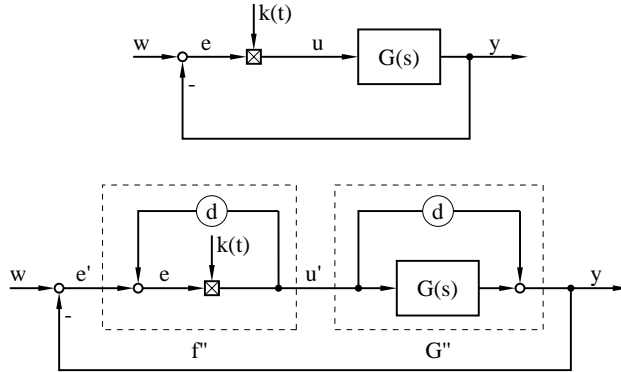


Abb. 13.32. Beispiel zur Anwendung des Hyperstabilitätskriteriums

ist. Die Ortskurve ist so weit nach rechts zu verschieben, dass sie vollständig rechts von der imaginären Achse verläuft. Dies kann durch das Einfügen eines Proportionalgliedes parallel zum linearen Systemteil erreicht werden. Gleichzeitig muss diese Veränderung des Gesamtsystems durch eine Rückkopplung über den nichtlinearen Teil wieder kompensiert werden. Man erhält schließlich einen transformierten Regelkreis mit dem linearen Teil $G''(s)$ und dem nichtlinearen Teil f'' . $G''(s)$ ist streng positiv reell und damit asymptotisch hyperstabil. Es ist demnach noch die Integralungleichung für den nichtlinearen Teil zu betrachten. Mit

$$u'(t) = \frac{k(t)}{1 - dk(t)} e'(t) \quad (13.107)$$

ergibt sich

$$\int_0^T u'(t) e'(t) dt = \int_0^T \frac{k(t)}{1 - dk(t)} e'^2(t) dt \geq -\beta_0^2 \quad (13.108)$$

Die Ungleichung ist sicher erfüllt für $\frac{k(t)}{1 - dk(t)} \geq 0$. Dazu müssen Zähler und Nenner das gleiche Vorzeichen aufweisen und der Nenner darüber hinaus ungleich Null sein. Zunächst sei angenommen, dass $k(t) \geq 0$ gilt. Damit muss auch $1 - dk(t) > 0$ bzw. $\frac{1}{d} > k(t)$ gelten, also insgesamt: $0 \leq k(t) < \frac{1}{d}$.

Bei der Annahme $k(t) \leq 0$ muss $1 - dk(t) < 0$ bzw. $\frac{1}{d} < k(t) \leq 0$ gelten, was zu einem Widerspruch führt, da d nur positiv sein kann.

Damit bleibt als Bedingung für Stabilität: $0 \leq k(t) < \frac{1}{d}$. Die Bedingung für $k(t)$ ist gegenüber dem ersten Fall eingeschränkt, da $G(s)$ nicht streng positiv reell ist.

Abschließend soll noch eine Variante des Hyperstabilitätskriteriums (vgl. [45]) vorgestellt werden, die auf Satz 13.9 basiert. Außerdem werden die möglicherweise notwendigen Erweiterungen des linearen Systemteils dazu genutzt, um zusätzliche Freiheitsgrade für die Stabilitätsanalyse zu erhalten und sie damit weniger konservativ zu machen.

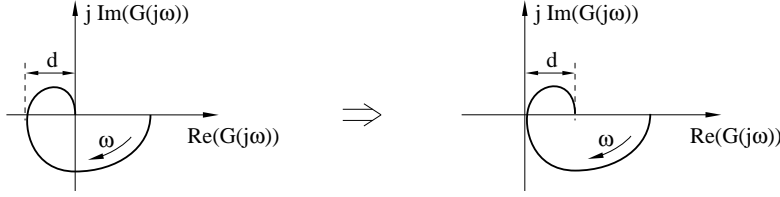


Abb. 13.33. Verschiebung der Ortskurve zur Erzielung von Hyperstabilität

Im Gegensatz zur obigen Darstellung der Erweiterungen wird bei diesem Verfahren zunächst die Stabilisierungsmatrix \mathbf{K} eingefügt, anschließend die Erweiterung \mathbf{N} für gleiche Anzahl von Ein- und Ausgangsgrößen der beiden Systemteile (\mathbf{M} ergibt sich direkt aus $\mathbf{NM} = \mathbf{I}$) und schließlich die Diagonalmatrix \mathbf{D} , um den linearen Systemteil positiv reell zu machen. \mathbf{D} wird entsprechend der zweiten oben aufgeführten Variante berechnet, d.h. \mathbf{D} ergibt sich hier mit den Gleichungen (13.80), (13.93) und (13.94) aus einer frei wählbaren, regulären Matrix \mathbf{L} . \mathbf{K} , \mathbf{N} und \mathbf{L} lassen sich demnach als Parameter betrachten, die für die Stabilitätsanalyse frei unter Einhaltung von Nebenbedingungen gewählt werden können.

Bei Berücksichtigung von \mathbf{N} und \mathbf{M} und unter Beachtung der geänderten Reihenfolge der Systemerweiterungen ergibt sich statt Gleichung (13.96) jetzt

$$\int_0^T (\mathbf{M}(\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e}))^T (\mathbf{e} - \mathbf{D}\mathbf{M}(\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e})) dt \geq -\beta_0^2 \quad (13.109)$$

als Stabilitätsbedingung. Diese Ungleichung ist leicht nachzuvollziehen, wenn man in Abb. 13.31 die Blöcke \mathbf{M} und \mathbf{N} einfügt, die bei dieser Variante explizit berücksichtigt werden. \mathbf{M} wird direkt nach der Subtraktionsstelle von \mathbf{u} und $\mathbf{K}\mathbf{e}$ eingefügt, und \mathbf{N} direkt vor der Subtraktionsstelle von \mathbf{u}' und $\mathbf{K}\mathbf{y}$. Diese Position ergibt sich aus der Reihenfolge der Erweiterungen.

Ein hinreichendes Kriterium für die Erfüllung der Bedingung ist, wenn der Integrand für beliebige Werte von \mathbf{e} positiv ist:

$$(\mathbf{M}(\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e}))^T (\mathbf{e} - \mathbf{D}\mathbf{M}(\mathbf{f}(\mathbf{e}) - \mathbf{K}\mathbf{e})) \geq 0 \quad (13.110)$$

Iterativ wird nun ein Optimierungsverfahren durchlaufen: Die Matrizen \mathbf{N} , \mathbf{K} und \mathbf{L} werden in einem ersten Schritt beliebig festgelegt, natürlich unter Beachtung der Randbedingungen, dass \mathbf{L} regulär ist, eine Matrix \mathbf{M} mit $\mathbf{NM} = \mathbf{I}$ existiert und \mathbf{K} das lineare System stabilisiert. Aus \mathbf{L} ergibt sich mit (13.80), (13.93) und (13.94) die Matrix \mathbf{D} . Dann wird für eine geeignete Menge an Werten \mathbf{e} die Ungleichung (13.110) überprüft. Falls sie für alle Werte erfüllt ist, gilt das System als stabil. Falls nicht, werden im Rahmen des Optimierungsverfahrens andere Matrizen \mathbf{N} , \mathbf{K} und \mathbf{L} gewählt und die gesamte Berechnung erneut durchgeführt. Im Laufe des Verfahrens werden \mathbf{N} , \mathbf{K} und \mathbf{L} so optimiert, dass die linke Seite von Gleichung (13.110) für alle

Werte von ϵ möglichst groß wird. Das Verfahren wird abgebrochen, sobald sie keine negativen Werte mehr annimmt.

Problematisch ist allerdings, dass kein Gradientenfeld für die Abhängigkeit der linken Seite der Ungleichung (13.110) von den Koeffizienten der drei zu optimierenden Matrizen existiert. Daher kann die Suche nach den optimalen Koeffizienten nicht systematisch, sondern nur mit Hilfe eines evolutionären Algorithmus erfolgen. Trotzdem ist eine solche Optimierung als sinnvoll zu werten, da das Resultat der Stabilitätsanalyse mit optimierten Matrizen sicherlich weniger konservativ ausfallen wird als mit nicht optimierten Matrizen, auch wenn die Optimierung nicht auf das absolute Optimum führt.

14. Entwurf nichtlinearer Regelungen

14.1 Sliding Mode-Regler

Nachdem bisher Stabilitätskriterien vorgestellt wurden, die auch für den Entwurf von Reglern herangezogen werden können, soll nun ein echtes Regler-Entwurfsverfahren vorgestellt werden. Es handelt sich um den *Sliding Mode-Regler* [47]. Voraussetzung ist ein Streckenmodell mit der Zustandsgleichung

$$\dot{x}^{(n)}(t) = f(\mathbf{x}(t)) + u(t) + d(t) \quad (14.1)$$

mit dem Zustandsvektor $\mathbf{x} = (x, \dot{x}, \dots, x^{(n-1)})^T$, der Stellgröße $u(t)$ und einer unbekannten Störgröße $d(t)$. Ein solches Modell entspricht im linearen Eingrößenfall der Regelungsnormalform (Abb. 7.5).

Der Sollwert x_d muss nicht unbedingt konstant sein, so dass stattdessen ein Sollvektor $\mathbf{x}_d = (x_d, \dot{x}_d, \dots, x_d^{(n-1)})^T$ eingeführt wird. Damit ist auch der Regelfehler $e = x_d - x$ durch den Fehlervektor $\mathbf{e} = \mathbf{x}_d - \mathbf{x} = (e, \dot{e}, \dots, e^{(n-1)})^T$ zu ersetzen. Ziel der Regelung ist $\mathbf{e} = \mathbf{0}$, d.h. der Regelfehler und seine sämtlichen Ableitungen sollen verschwinden.

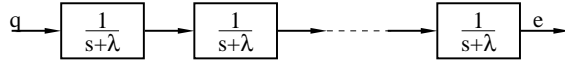
Nun soll dieses Regelziel aber nicht direkt verfolgt, sondern zunächst durch ein anderes Regelziel ersetzt werden, das durch die Differentialgleichung

$$0 = q(\mathbf{e}) = \left(\frac{\partial}{\partial t} + \lambda\right)^{n-1} e \quad (14.2)$$

$$\begin{aligned} &= e^{(n-1)} + \binom{n-1}{1} \lambda e^{(n-2)} + \binom{n-1}{2} \lambda^2 e^{(n-3)} + \dots + \lambda^{n-1} e \\ &= e^{(n-1)} + g_\lambda(\mathbf{e}) \end{aligned} \quad (14.3)$$

mit $\lambda > 0$ beschrieben wird. Genügt der Fehlervektor dieser Differentialgleichung, so wird er, ausgehend von jedem beliebigen Anfangszustand, immer gegen Null gehen. Verfolgt man also das Regelziel $q = 0$, so wird sich das ursprüngliche Regelziel von ganz allein einstellen.

Dies kann man sich auch anschaulich klarmachen. Gleichung (14.2) entspricht doch der Differentialgleichung von $n - 1$ hintereinandergeschalteten PT_1 -Gliedern (Abb. 14.1). Wenn die Eingangsgröße q Null wird, so gilt dies in der Folge auch für e und seine $n - 1$ Ableitungen.

**Abb. 14.1.** Anschauliche Deutung von q

Das neue Regelziel $q(\mathbf{e}) = 0$ soll nun wiederum durch eine andere Bedingung ersetzt werden. Und zwar ist doch die Funktion $q^2(\mathbf{e})$ sicher überall positiv außer im Regelziel $q(\mathbf{e}) = 0$. Wenn man daher gewährleisten kann, dass für die Ableitung dieser Funktion immer die Bedingung

$$\frac{\partial}{\partial t}(q^2(\mathbf{e})) < -2\eta|q(\mathbf{e})| \quad (14.4)$$

mit $\eta \geq 0$ gilt, so wird $q^2(\mathbf{e})$ von jedem beliebigen Anfangswert gegen $q^2(\mathbf{e}) = 0$ gehen, womit dann auch $q(\mathbf{e}) = 0$ gilt. q^2 lässt sich damit als Ljapunov-Funktion deuten. Die Einhaltung der Ungleichung (14.4) führt also dazu, dass irgendwann auch Gleichung (14.2) erfüllt ist. Und daraus resultiert wiederum, wie schon gesagt, früher oder später die Erfüllung des ursprünglichen Regelzieles $\mathbf{e} = \mathbf{0}$.

Es stellt sich die Frage, welchen Vorteil das zweimalige Ersetzen des Regelzieles gebracht hat. Gleichung (14.4) lässt sich zunächst etwas einfacher formulieren. Die Berechnung der Ableitung liefert nämlich

$$q\dot{q} < -\eta|q| \quad (14.5)$$

und damit

$$\dot{q} \operatorname{sgn}(q) < -\eta \quad (14.6)$$

Mit dieser Formulierung des Regelzieles lässt sich eine Antwort auf die eben gestellte Frage geben. Hatte die ursprüngliche Regelaufgabe, das anfangs gegebene System auf einen gegebenen Sollvektor zu regeln, wegen der $n - 1$ Ableitungen von \mathbf{e} noch die Ordnung n , so hat die durch das neue Regelziel (14.6) gegebene Aufgabe offenbar nur noch die Ordnung Eins. Denn die betrachtete, zu regelnde Größe kommt in dieser Gleichung nur in der ersten Ableitung vor.

Interessant ist auch eine geometrische Interpretation der verschiedenen Regelziele. Die Bedingung $q(\mathbf{e}) = 0$ definiert im durch \mathbf{e} aufgespannten, n -dimensionalen Raum eine Hyperfläche. Das System wird bei Einhaltung der Ungleichung (14.4) bzw. (14.6) gezwungen, sich dieser Hyperfläche zu nähern, und kann sie nach ihrem Erreichen nicht mehr verlassen. Auf der Hyperfläche gleitet das System dann von allein in den Punkt $\mathbf{e} = \mathbf{0}$ hinein. $q(\mathbf{e}) = 0$ wird daher auch als *sliding surface* bezeichnet. Abb. 14.2 verdeutlicht dies für den Fall $n = 2$. Die Hyperfläche besteht hier wegen $0 = q(\mathbf{e}) = \dot{e} + \lambda e$ aus einer Geraden durch den Ursprung der $e - \dot{e}$ -Ebene. Die anderen eingezeichneten Geraden sowie die Variable Φ werden später erläutert.

Die Frage ist nun, wie die Stellgröße beschaffen sein muss, damit Ungleichung (14.6) immer erfüllt ist. Nach (14.3) ergibt sich zunächst

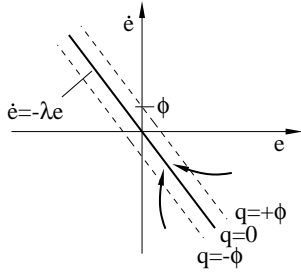


Abb. 14.2. Zur Sliding Mode-Regelung

$$\begin{aligned} q &= e^{(n-1)} + g_\lambda(\mathbf{e}) \\ \dot{q} &= e^{(n)} + g_\lambda(\dot{\mathbf{e}}) = x_d^{(n)} - x^{(n)} + g_\lambda(\dot{\mathbf{e}}) \end{aligned} \quad (14.7)$$

und mit (14.1)

$$\dot{q} = g_\lambda(\dot{\mathbf{e}}) - f(\mathbf{x}) - u - d + x_d^{(n)} \quad (14.8)$$

Daraus folgt für die Ungleichung (14.6)

$$(g_\lambda(\dot{\mathbf{e}}) - f(\mathbf{x}) - u - d + x_d^{(n)}) \operatorname{sgn}(q) < -\eta \quad (14.9)$$

Die Funktion f lässt sich zerlegen in

$$f = f_0 + \Delta f \quad (14.10)$$

Dabei kennzeichnet f_0 das nominale Streckenmodell, d.h. den Anteil des Streckenmodells, von dem man weiß, dass er richtig ist, während Δf die Modellunsicherheit darstellt. Wählt man dann für die Stellgröße

$$u = -f_0(\mathbf{x}) + g_\lambda(\dot{\mathbf{e}}) + x_d^{(n)} + U \operatorname{sgn}(q) \quad (14.11)$$

mit einem konstanten, später zu bestimmenden Wert U , so wird aus Ungleichung (14.9)

$$(-\Delta f(\mathbf{x}) - d) \operatorname{sgn}(q) - U < -\eta \quad (14.12)$$

Die Modellunsicherheit Δf und die Störgröße d sollen durch obere Grenzen abschätzbar sein:

$$|\Delta f| < F \quad |d| < D \quad (14.13)$$

Dann ist die Ungleichung (14.12) und damit auch (14.4) sicher erfüllt, wenn

$$U = F + D + \eta \quad (14.14)$$

gilt, und aus (14.11) folgt für die Stellgröße

$$u = -f_0(\mathbf{x}) + g_\lambda(\dot{\mathbf{e}}) + x_d^{(n)} + (F + D + \eta) \operatorname{sgn}(q) \quad (14.15)$$

Damit ist der Sliding Mode-Regler definiert. Die ersten drei Summanden kann man als inverses Streckenmodell auffassen, während der letzte Summand im

wesentlichen durch die Modellunsicherheiten und Störungen hervorgerufen wird. Weiterhin lässt sich ablesen, dass für eine solche Regelung zunächst die Strecke in der Form (14.1) darstellbar und das zugehörige Streckenmodell f_0 auch (weitgehend) bekannt sein muss. Dabei ist eine Modellunsicherheit Δf zugelassen, deren maximaler Wert durch F abzuschätzen ist. Diese Anforderung an einen robusten Regler ist aber normal. Damit der Regler die Strecke innerhalb eines gewissen Modellunsicherheitsbereichs stabilisieren kann, muss dieser angegeben werden können.

Ebenfalls abschätzbar sein muss die maximale Amplitude der Störgröße durch den Wert D . Darüber hinaus muss der Zustandsvektor \mathbf{x} gemessen werden können. Mit dem sowieso bekannten Verlauf des Sollwertvektors \mathbf{x}_d lassen sich daraus dann aber sofort $x_d^{(n)}$, der noch benötigte Fehlervektor $\mathbf{e} = \mathbf{x}_d - \mathbf{x}$ und damit auch $g_\lambda(\dot{\mathbf{e}})$ sowie $q(\mathbf{e})$ bestimmen.

Die Bestimmung des Fehlervektors birgt allerdings ein Problem. Gemessen wird am Ausgang der Strecke zunächst nur $e = x_d - x$, benötigt werden aber n Ableitungen für den Fehlervektor $\dot{\mathbf{e}}$. Durch einfache diskrete Ableitungen sind diese Größen jedoch nicht zu gewinnen, da sich das unvermeidliche Messrauschen auf e so stark auf die höheren Ableitungen auswirken würde, dass diese für Rechnungen nicht mehr zu gebrauchen wären. $\dot{\mathbf{e}}$ kann daher nur mit Hilfe eines nichtlinearen Beobachters bestimmt werden, der aber wiederum ein relativ präzises Streckenmodell f_0 erfordert.

Festzulegen sind schließlich noch die Parameter λ und η . Durch η wird nach Ungleichung (14.4) die Annäherungsgeschwindigkeit des Systems an die Hyperfläche vorgegeben. Je größer η gewählt wird, desto schneller nähert sich das System der Hyperfläche. Dies erfordert aber auch, wie Gleichung (14.15) zeigt, eine umso größere Stellgröße, so dass bei der Festlegung von η technische Gesichtspunkte zu berücksichtigen sind.

Währenddessen wird durch λ entsprechend Gleichung (14.3) die Hyperfläche definiert. Befindet sich das System auf der Hyperfläche, so wird durch diese das dynamische Verhalten des Systems vorgegeben. Und zwar geht der Fehler umso schneller gegen Null, je größer λ gewählt wird. Wie schon für η sind offenbar auch bei der Wahl von λ die technischen Gegebenheiten des Systems zu berücksichtigen. Ist λ jedoch erst einmal festgelegt, so bestimmt allein dieser Parameter das Systemverhalten auf der Hyperfläche, und zwar unabhängig von den Streckenparametern, Störungen oder Änderungen der Streckenparameter. Dies kennzeichnet aber doch gerade eine robuste Regelung, denn bei einer robusten Regelung ist das beabsichtigte Regelverhalten auch dann gewährleistet, wenn sich die Streckenparameter verändern. Und das Maß für die Robustheit, d.h. die zulässigen Abweichungen der realen von der nominalen Strecke, ist durch F gegeben.

Grundsätzlich kann $q(\mathbf{e})$ anstelle von (14.2) auch durch ein allgemeines Polynom

$$q(\mathbf{e}) = \sum_{i=0}^{n-2} c_i e^{(i)} + e^{(n-1)} \quad (14.16)$$

definiert werden, dessen Koeffizienten c_i so zu bestimmen sind, dass alle Nullstellen des Polynoms

$$c(s) = s^{n-1} + c_{n-2}s^{n-2} + \dots + c_1s + c_0 \quad (14.17)$$

einen negativen Realteil aufweisen. Dies bedeutet Stabilität des entsprechenden linearen Übertragungsgliedes. Und damit ist sichergestellt, dass mit $q(e) = 0$ auch e gegen Null konvergiert. Im Gegensatz zu vorher einem freien Parameter λ hat man nun $n - 1$ freie Parameter, mit denen man die Hyperfläche besser an die Erfordernisse des Systems anpassen kann. Im konkreten Fall ist diese Anpassung allerdings nicht trivial.

Unangenehm an einem Sliding Mode-Regler nach Gleichung (14.15) ist der unstetige Stellgrößenverlauf bei jedem Vorzeichenwechsel von q . Und zwar fällt der Sprung umso größer aus, je größer die Modellungenauigkeit F und die Abschätzung für die Störung D sind. η dagegen kann zur Verkleinerung der Sprunghöhe auch zu Null gesetzt werden, da sich dadurch nur die Regelschwindigkeit verändert. Für ein exaktes Modell und eine ungestörte Strecke ließe sich daher ein stetiger Stellgrößenverlauf erzielen. Da dies aber in der Praxis nie gegeben ist, muss man mit anderen Mitteln versuchen, die Unstetigkeit zu vermeiden. Hier bietet es sich an, die Signumfunktion durch die Funktion

$$h(q) = \begin{cases} \frac{1}{\Phi}q & : |q| < \Phi \\ \text{sgn}(q) & : |q| \geq \Phi \end{cases} \quad (14.18)$$

zu ersetzen (Abb. 14.3). Andererseits war die Signumfunktion im Regelgesetz (14.15) aber zur Einhaltung der Ungleichung (14.4) erforderlich. Ersetzt man sie daher durch $h(q)$, so wird für $|q| < \Phi$ die Ungleichung möglicherweise nicht mehr eingehalten. Als Folge davon kann das mit der Sliding Mode-Regelung eigentlich beabsichtigte Systemverhalten für solche Werte von q nicht mehr garantiert werden. Durch die Regelung wird nur noch gewährleistet, dass das System in die durch $|q| < \Phi$ gegebene Zone um die Hyperfläche $q = 0$ eintritt und auch dort verbleibt (Abb. 14.2). Es nähert sich dem Zielpunkt, wird ihn jedoch in Anwesenheit von Modellunsicherheiten und Störungen nicht exakt erreichen. Andererseits wird die Zone aber auch nicht mehr verlassen. Insofern kann sie als Toleranzbereich der Regelung angesehen werden. Je größer Φ , d.h. je weicher der Verlauf der Stellgröße, desto größer ist auch der zu akzeptierende Toleranzbereich.

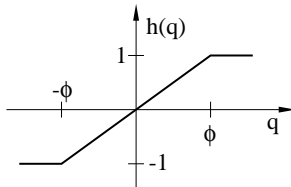


Abb. 14.3. Ersatzfunktion für die Signumfunktion

14.2 Gain-Scheduling-Regler

14.2.1 GS-Regler und GS-Modelle

Gain-Scheduling-Regler sind verallgemeinerte Kennfeldregler, und aus dem Grund bildet eine Beschreibung der Kennfeldregler den Anfang dieses Kapitels.

Bei einem Kennfeldregler ergibt sich die Stellgröße aus einer arbeitspunkt-abhängigen, additiven Überlagerung von konstanten Werten oder Vektoren:

$$\mathbf{u} = \sum_i k_i(\mathbf{z}(t)) \mathbf{u}_i \quad \text{mit} \quad \sum_i k_i(\mathbf{z}(t)) = 1 \quad (14.19)$$

Dabei sind die \mathbf{u}_i die konstanten Vektoren, die sich im Eingrößenfall auf Skalare reduzieren, $k_i(\mathbf{z}(t))$ die Gewichtungsfunktionen und $\mathbf{z}(t)$ die Eingangsgrößen der Gewichtungsfunktionen, die oft die Zustandsgrößen der Strecke sind, aber grundsätzlich zunächst einmal jede beliebige Größe sein können.

Diejenigen Vektoren $\mathbf{z}(t)$, für die eine Gewichtungsfunktion k_i den Wert Eins annimmt und alle anderen den Wert Null, werden im Folgenden als Stützstellen bezeichnet. Der Vektor \mathbf{u}_i ist demnach der Ausgangsvektor des Reglers an der i -ten Stützstelle.

Ein solches Kennfeld ist in Abb. 14.4 für einen Regler mit zwei Eingangsgrößen z_1 und z_2 und einer Ausgangsgröße u dargestellt. Für die Stützstellen sind die Ausgangsgrößen u_i vorgegeben. Liegt ein Eingangsvektor $\mathbf{z} = (z_1, z_2)^T$ nicht exakt auf einer Stützstelle, so ist zwischen den benachbarten Stützstellen zu interpolieren. Wie dies zu geschehen hat, zeigt die rechte Zeichnung. Vorausgesetzt wird, dass die Abstände zwischen den Stützstellen jeweils Eins betragen, was durch eine geeignete Normierung der Eingangsgrößen leicht zu erreichen ist. Für den Eingangsvektor \mathbf{z} berechnet sich dann die Ausgangsgröße nach

$$u(\mathbf{z}) = (1 - a)[(1 - b)u_1 + bu_2] + a[(1 - b)u_3 + bu_4] \quad (14.20)$$

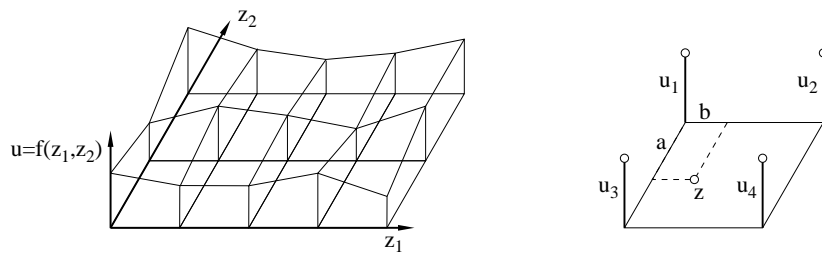


Abb. 14.4. Kennfeld-Regler

Ersetzt man die konstanten Vektoren \mathbf{u}_i durch einen Regler der Form $\mathbf{F}_i \mathbf{x}$ mit der Koeffizientenmatrix \mathbf{F}_i und einem Vektor von Systemgrößen \mathbf{x} , so erhält man den Gain-Scheduling-Regler (GS-Regler):

$$\mathbf{u} = \sum_i k_i(\mathbf{z}) \mathbf{F}_i \mathbf{x} \quad (14.21)$$

An dieser Formel ist ersichtlich, dass die Regler-Eingangsgrößen \mathbf{x} nicht unbedingt dieselben Größen sein müssen, die in die Gewichtungsfunktion eingehen (\mathbf{z}). Üblicherweise ist \mathbf{x} ein Zustandsvektor, jedoch sind auch GS-Regler mit anderen Eingangsgrößen denkbar.

Es lässt sich leicht zeigen, dass der Kennfeldregler ein Spezialfall des GS-Reglers ist: Für den GS-Regler mit den Eingangsgrößen $\mathbf{x} = [x_1, \dots, x_n]^T$ muss lediglich eine zusätzliche, künstliche Eingangsgröße x_{n+1} definiert werden, die immer den konstanten Wert Eins aufweist. Setzt man dann alle Elemente der Matrizen \mathbf{F}_i gleich Null und nur die $(n+1)$ -te Spalte von \mathbf{F}_i jeweils gleich \mathbf{u}_i aus Gleichung (14.19), so wird aus dem GS-Regler gerade der Kennfeldregler aus (14.19):

$$\mathbf{F}_i = (\mathbf{0} \ \mathbf{u}_i) \quad \text{und} \quad \mathbf{x}' = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \quad \mathbf{u}_i = \mathbf{F}_i \mathbf{x}' \quad (14.22)$$

$\mathbf{0}$ ist hier die Nullmatrix der Dimension $p \times n$, wobei p die Anzahl der Eingangsgrößen der Strecke ist.

GS-Systeme lassen sich auch zur Modellierung gegebener Strecken-Übertragungsglieder heranziehen. Mit Hilfe eines solchen GS-Modelles lässt sich jedes beliebige lineare oder nichtlineare Übertragungsverhalten mit oder ohne interne Dynamik mit steigender Anzahl von Stützstellen beliebig genau approximieren, sofern es keine Hysterese oder Laufzeiten enthält. Das GS-Modell besteht aus einzelnen linearen Modellen, deren Ausgangsgrößen mit wechselnden Gewichtungsfaktoren je nach dem Momentanwert der Eingangsgrößen überlagert werden. Für das Zustandsmodell einer Strecke ergibt sich beispielsweise

$$\dot{\mathbf{x}}(t) = \sum_i k_i(\mathbf{z}(t)) [\mathbf{A}_i \mathbf{x}(t) + \mathbf{B}_i \mathbf{u}(t)] \quad (14.23)$$

In der Praxis gewinnt man ein solches Modell, wenn man zunächst verschiedene Stützstellen als Arbeitspunkte auswählt und anschließend für jeden Arbeitspunkt ein lineares Modell bestimmt, welches das Streckenverhalten am Arbeitspunkt beschreibt. Das Modell kann mittels eines Identifikationsverfahrens ermittelt werden oder auch direkt durch Aufstellen von das Systemverhalten beschreibenden, linearen Differentialgleichungen. Das gleiche wird für alle anderen Stützstellen durchgeführt. Damit ist das GS-Modell der Strecke festgelegt. Es ist allerdings darauf zu achten, dass an jedem Arbeitspunkt dieselbe Systemstruktur vorausgesetzt wird, d.h. insbesondere dieselbe Anzahl an Zustands- und Eingangsgrößen, damit die Matrizen \mathbf{A}_i und \mathbf{B}_i an jedem Arbeitspunkt dieselbe Dimension aufweisen.

Nach der Modellbildung kann für jede Stützstelle des GS-Modells ein linearer Zustandsregler ausgelegt werden, d.h. es wird für jedes lineare Teilstreckenmodell $(\mathbf{A}_i, \mathbf{B}_i)$ der Zustandsregler \mathbf{F}_i entworfen. Diese einzelnen

Regler werden dann zu einem GS-Regler zusammengefasst, der die gleichen Stützstellen aufweist wie das GS-Streckenmodell. Es ergibt sich ein GS-Regler entsprechend Gleichung (14.21), wobei der Vektor \mathbf{x} dort ein Vektor beliebiger Systemgrößen war, während er jetzt ausschließlich aus Zustandsgrößen besteht. Inwieweit ein so entworfener Regler das System dann tatsächlich stabilisiert, wird in Kapitel 14.2.2 behandelt.

Einsetzen der Gleichung (14.21) des GS-Reglers in (14.23) liefert die Zustandsgleichung eines von außen nicht angeregten, geschlossenen Kreises in Form eines GS-Modelles

$$\dot{\mathbf{x}}(t) = \sum_i \sum_j k_i(\mathbf{z}(t)) k_j(\mathbf{z}(t)) [\mathbf{A}_i + \mathbf{B}_i \mathbf{F}_j] \mathbf{x}(t) \quad (14.24)$$

$$\dot{\mathbf{x}}(t) = \sum_i \sum_j k_i(\mathbf{z}(t)) k_j(\mathbf{z}(t)) \mathbf{G}_{ij} \mathbf{x}(t) \quad (14.25)$$

mit $\mathbf{G}_{ij} = \mathbf{A}_i + \mathbf{B}_i \mathbf{F}_j$, und nach einer Umindizierung und Zusammenfassung mit $\mathbf{A}_{g,l} = \mathbf{G}_{ij}$ und $k_l(\mathbf{z}(t)) = k_i(\mathbf{z}(t)) k_j(\mathbf{z}(t))$

$$\dot{\mathbf{x}}(t) = \sum_l k_l(\mathbf{z}(t)) \mathbf{A}_{g,l} \mathbf{x}(t) \quad (14.26)$$

Der Index g an der Systemmatrix soll verdeutlichen, dass es sich hierbei um ein Modell des geschlossenen Kreises handelt und nicht um ein Modell der Strecke.

Unter Berücksichtigung einer äußeren Anregung ergibt sich daraus die allgemeine Form des GS-Modelles eines geschlossenen Kreises mit äußerer Anregung (vgl. [57]):

$$\dot{\mathbf{x}}(t) = \sum_i k_i(\mathbf{z}(t)) [\mathbf{A}_{g,i} \mathbf{x}(t) + \mathbf{B}_{g,i} \mathbf{w}(t)] \quad (14.27)$$

Weil heutige Regler im allgemeinen mit einem Mikroprozessor realisiert werden, ist gegebenenfalls auch die diskrete Form dieser Zustandsgleichung erforderlich:

$$\mathbf{x}(k+1) = \sum_i k_i(\mathbf{z}(k)) \left[\mathbf{A}_{g,i}^{(d)} \mathbf{x}(k) + \mathbf{B}_{g,i}^{(d)} \mathbf{w}(k) \right] \quad (14.28)$$

Dabei sind $\mathbf{x}(k)$ der Zustandsvektor des Systems und $\mathbf{w}(k)$ die Anregung zu einem Zeitpunkt $t = kT$, wobei T das Abtastintervall der Regelung darstellt. Die hier auftretenden Matrizen $\mathbf{A}_{g,i}^{(d)}$ und $\mathbf{B}_{g,i}^{(d)}$ sind die Matrizen der diskreten Zustandsdarstellung und nicht identisch mit denen in Gleichung (14.27).

Der GS-Regler als Zustandsregler wirft allerdings die Frage auf, wie die Zustandsgrößen als Eingangsgrößen des Reglers bereitgestellt werden können, da Zustandsgrößen in vielen Fällen nicht direkt messbar sind. Wie schon bei klassischen Zustandsreglern ist daher auch hier ein Beobachter erforderlich. Und da sowohl das Streckenmodell als auch der Regler als GS-Systeme vorliegen, bietet es sich an, auch den Beobachter als GS-System auszuführen.

Dabei ist es naheliegend, für den Beobachter die gleichen Stützstellen zu wählen wie für Streckenmodell und Regler. Für jede Stützstelle wird dann zunächst ein linearer Beobachter (vgl. Abbildung 8.6) definiert:

$$\hat{\mathbf{x}}(t) = \mathbf{A}_i \hat{\mathbf{x}}(t) + \mathbf{B}_i \mathbf{u}(t) + \mathbf{H}_i (\mathbf{y}(t) - \hat{\mathbf{y}}(t)) \quad (14.29)$$

\mathbf{A}_i und \mathbf{B}_i entsprechen den Matrizen des Streckenmodells aus Gleichung (14.23). $\hat{\mathbf{x}}(t)$ ist der geschätzte Zustandsvektor und $\hat{\mathbf{y}}(t)$ der Ausgangsvektor des Beobachters, der einen Schätzwert für den realen Ausgangsvektor $\mathbf{y}(t)$ darstellt. \mathbf{H}_i schließlich ist die Korrekturmatrix, mit der die Differenz zwischen realem und geschätztem Ausgangsvektor multipliziert und als Korrekturterm in den Beobachter zurückgeführt wird.

Diese einzelnen Beobachter werden dann zu einem GS-Beobachter zusammengefügt:

$$\hat{\mathbf{x}}(t) = \sum_i k_i(\mathbf{z}(t)) [\mathbf{A}_i \hat{\mathbf{x}}(t) + \mathbf{B}_i \mathbf{u}(t) + \mathbf{H}_i (\mathbf{y}(t) - \hat{\mathbf{y}}(t))] \quad (14.30)$$

Spezielle Beachtung verdienen bei einem derartigen Beobachter die Eingangsgrößen der Prämissen \mathbf{z} . Denn sofern Zustandsgrößen als Eingangsgrößen verwendet werden, kann es sich dabei auch nur um die geschätzten Zustandsgrößen handeln. Diese Tatsache muss in allen Verfahren, in denen es um GS-Beobachter geht, geeignet berücksichtigt werden.

14.2.2 Stabilitätsanalyse nach Ljapunov

Stabilitätsbeweis. Die Stabilitätsanalyse kann basierend auf dem GS-Modell des geschlossenen Kreises (14.27) bzw. der diskreten Version (14.28) durchgeführt werden.

Mit der diskreten Version soll begonnen werden. Hier gilt der Satz (vgl. [57]):

Satz 14.1 *Gegeben sei ein diskretes System in der Form*

$$\mathbf{x}(k+1) = \sum_i k_i(\mathbf{z}(k)) \mathbf{A}_i \mathbf{x}(k) \quad (14.31)$$

Dieses System besitzt eine global asymptotisch stabile Ruhelage $\mathbf{x} = \mathbf{0}$, wenn eine gemeinsame, positiv definite Matrix \mathbf{P} für alle Teilsysteme \mathbf{A}_i existiert, so dass

$$\mathbf{M}_i = \mathbf{A}_i^T \mathbf{P} \mathbf{A}_i - \mathbf{P} \quad (14.32)$$

für alle i negativ definit ($\mathbf{M}_i < 0$) ist.

Für den Beweis sei angenommen, dass eine positiv definite Matrix \mathbf{P} existiert. Mit dieser wird die Ljapunov-Funktion $V = \mathbf{x}^T(k) \mathbf{P} \mathbf{x}(k)$ angesetzt. Dann gilt:

$$\begin{aligned}
\Delta V(\mathbf{x}(k)) &= V(\mathbf{x}(k+1)) - V(\mathbf{x}(k)) \\
&= \mathbf{x}^T(k+1)\mathbf{P}\mathbf{x}(k+1) - \mathbf{x}^T(k)\mathbf{P}\mathbf{x}(k) \\
&= \left(\sum_i k_i \mathbf{A}_i \mathbf{x}(k) \right)^T \mathbf{P} \left(\sum_j k_j \mathbf{A}_j \mathbf{x}(k) \right) - \mathbf{x}^T(k)\mathbf{P}\mathbf{x}(k) \\
&= \mathbf{x}^T(k) \left[\left(\sum_i k_i \mathbf{A}_i^T \right) \mathbf{P} \left(\sum_j k_j \mathbf{A}_j \right) - \mathbf{P} \right] \mathbf{x}(k) \\
&= \sum_{i,j} k_i k_j \mathbf{x}^T(k) (\mathbf{A}_i^T \mathbf{P} \mathbf{A}_j - \mathbf{P}) \mathbf{x}(k) \\
&= \sum_i k_i^2 \mathbf{x}^T(k) (\mathbf{A}_i^T \mathbf{P} \mathbf{A}_i - \mathbf{P}) \mathbf{x}(k) \\
&\quad + \sum_{i < j} k_i k_j \mathbf{x}^T(k) (\mathbf{A}_i^T \mathbf{P} \mathbf{A}_j + \mathbf{A}_j^T \mathbf{P} \mathbf{A}_i - 2\mathbf{P}) \mathbf{x}(k) < 0 \quad (14.33)
\end{aligned}$$

Die Matrizen in der ersten Summe sind laut Voraussetzung negativ definit, so dass jeder einzelne Summand sicher kleiner als Null ist. Die Matrizen in der zweiten Summe lassen sich folgendermaßen umformen:

$$\begin{aligned}
\mathbf{A}_i^T \mathbf{P} \mathbf{A}_j + \mathbf{A}_j^T \mathbf{P} \mathbf{A}_i - 2\mathbf{P} &= -(\mathbf{A}_i - \mathbf{A}_j)^T \mathbf{P} (\mathbf{A}_i - \mathbf{A}_j) \\
&\quad + \mathbf{A}_i^T \mathbf{P} \mathbf{A}_i + \mathbf{A}_j^T \mathbf{P} \mathbf{A}_j - 2\mathbf{P} \\
&= -(\mathbf{A}_i - \mathbf{A}_j)^T \mathbf{P} (\mathbf{A}_i - \mathbf{A}_j) \\
&\quad + (\mathbf{A}_i^T \mathbf{P} \mathbf{A}_i - \mathbf{P}) + (\mathbf{A}_j^T \mathbf{P} \mathbf{A}_j - \mathbf{P}) \quad (14.34)
\end{aligned}$$

Wegen der positiven Definitheit von \mathbf{P} ist der erste Summand sicher negativ definit, während dies für die beiden anderen Summanden laut Voraussetzung gilt. Damit sind aber auch alle Matrizen in der zweiten Summe von Gleichung (14.33) negativ definit und demnach auch hier alle Summanden kleiner als Null. Die Ableitung bzw. Differenz der Ljapunov-Funktion ist deshalb immer negativ, woraus die Stabilität des Systems folgt.

Für kontinuierliche GS-Systeme nach (14.27) sind die Verhältnisse noch einfacher. Der Stabilitätssatz lautet hier:

Satz 14.2 *Gegeben sei ein kontinuierliches System in der Form*

$$\dot{\mathbf{x}} = \sum_i k_i(\mathbf{z}(t)) \mathbf{A}_i \mathbf{x}(t) \quad (14.35)$$

Dieses System besitzt eine global asymptotisch stabile Ruhelage $\mathbf{x} = \mathbf{0}$, wenn eine gemeinsame, positiv definite Matrix \mathbf{P} für alle Teilsysteme \mathbf{A}_i existiert, so dass

$$\mathbf{M}_i = \mathbf{A}_i^T \mathbf{P} + \mathbf{P} \mathbf{A}_i \quad (14.36)$$

für alle i negativ definit ($\mathbf{M}_i < 0$) ist.

Für den Beweis wird wieder eine Ljapunov-Funktion $V = \mathbf{x}^T \mathbf{P} \mathbf{x}$ mit positiv definiter Matrix \mathbf{P} gewählt. Für die Ableitung dieser Funktion nach der Zeit gilt:

$$\begin{aligned}\dot{V} &= \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} \\ &= \sum_i k_i \mathbf{x}^T \mathbf{A}_i^T \mathbf{P} \mathbf{x} + \sum_i k_i \mathbf{x}^T \mathbf{P} \mathbf{A}_i \mathbf{x} \\ &= \sum_i k_i \mathbf{x}^T (\mathbf{A}_i^T \mathbf{P} + \mathbf{P} \mathbf{A}_i) \mathbf{x} < 0\end{aligned}\quad (14.37)$$

Nach Voraussetzung sind wieder alle Matrizen der Summe negativ definit und damit jeder einzelne Summand kleiner als Null. Das System ist demnach stabil.

Beide Sätze können direkt für die Stabilitätsanalyse eines GS-Systems verwendet werden. Die Analyse gestaltet sich besonders einfach, wenn man die Frage nach der Existenz der Matrix \mathbf{P} als LMI-Problem (Lineare Matrix-Ungleichung) formuliert. Im Anhang A.7 wird ausführlich beschrieben, wie ein System aus Ungleichungen

$$\mathbf{M}_i = \mathbf{A}_i^T \mathbf{P} + \mathbf{P} \mathbf{A}_i < \mathbf{0} \quad (14.38)$$

in die Form (A.59) gebracht werden kann. Auf diese Form lässt sich dann ein LMI-Lösungs-Algorithmus anwenden, der die Frage nach der Existenz einer Lösung \mathbf{P} und damit die Frage nach der Stabilität des Systems beantwortet.

Erleichterung der Stabilitätsbedingungen. Zu beachten ist, dass die negative Definitheit aller \mathbf{M}_i zwar ein hinreichendes, aber kein notwendiges Kriterium für die Stabilität des Systems ist. Die negative Definitheit aller \mathbf{M}_i bewirkt doch, dass jeder Summand in den Gleichungen (14.33) bzw. (14.37) für beliebige \mathbf{x} negativ ist, obwohl doch eigentlich nur die gesamte Summe negativ sein müsste, um Stabilität zu gewährleisten. Einzelne Summanden dürften also durchaus positiv sein, ohne dass das System instabil wäre. Eine entscheidende Abschwächung bzw. Vereinfachung der Stabilitätsbedingungen lässt sich daher erzielen, wenn die Koeffizienten k_i und ihre Abhängigkeit vom Eingangsvektor \mathbf{z} in den Gleichungen (14.33) bzw. (14.37) für ein Stabilitätskriterium mit berücksichtigt werden. Derartige Ansätze existieren aber bisher nur für GS-Systeme mit Reglern (vgl. (14.25)) und werden in Kapitel 14.2.2 noch vorgestellt.

Eine andere Möglichkeit besteht darin, auf numerischem Wege für eine vorgegebene Matrix \mathbf{P} und eine ausreichend große Anzahl an Vektoren \mathbf{x} und \mathbf{z} die Ungleichung (14.33) bzw. (14.37) direkt überprüfen. Dieser Ansatz weist aber gegenüber der Lösung mittels LMI-Algorithmen den entscheidenden Nachteil auf, dass er nur für eine einzelne, gegebene Matrix \mathbf{P} die Stabilität untersucht. Ist das Ergebnis negativ, so wird eine mühsame, unstrukturierte und möglicherweise erfolglose Suche nach einer geeigneten Matrix \mathbf{P} erforderlich, mit der sich die Stabilität des Systems nachweisen lässt, ohne zu wissen, ob eine solche Matrix überhaupt existiert.

Dagegen beantwortet ein LMI-Algorithmus gerade diese grundsätzliche Frage nach der Existenz einer Matrix \mathbf{P} , mit der das Ungleichungssystem (14.38) erfüllt ist. Von daher ist die Vorgehensweise mittels LMI-Algorithmen auf jeden Fall vorzuziehen, auch wenn es einzelne Systeme geben wird, deren Stabilität mit dieser Methode nicht nachgewiesen werden kann. Im weiteren Verlauf dieses Kapitels wird grundsätzlich von der Verwendung eines LMI-Algorithmus ausgegangen. Erst am Ende werden noch einmal kurz andere Ansätze skizziert.

Zur Abschwächung der Stabilitätsbedingungen (14.32) und (14.36) lässt sich in diesem Zusammenhang aber auch die Methode nach Aisermann einsetzen (vgl. [72]). Da der grundsätzliche Ansatz durch diese Methode nur leicht variiert wird, führt sie ebenfalls auf lineare Matrix-Ungleichungen, und die Lösung kann auch hier wieder mit einem LMI-Lösungsalgorithmus gewonnen werden. Im Folgenden soll die Methode für kontinuierliche Modelle demonstriert werden. Dazu ist in

$$\dot{\mathbf{x}}(t) = \sum_i k_i(\mathbf{z}(t)) \mathbf{A}_i \mathbf{x}(t) \quad (14.39)$$

jedes Teilsystem \mathbf{A}_i zu zerlegen in einen gemeinsamen, stabilen Anteil \mathbf{A} und einen Rest $\Delta \mathbf{A}_i$, so dass sich als neue Systemdarstellung ergibt:

$$\dot{\mathbf{x}} = \left[\mathbf{A} + \sum_i k_i \Delta \mathbf{A}_i \right] \mathbf{x} \quad (14.40)$$

Mit der positiv definiten Lösung \mathbf{P} der Ljapunov-Gleichung (vgl. Satz A.2)

$$\mathbf{P}\mathbf{A} + \mathbf{A}^T \mathbf{P} = -\mathbf{I} \quad (14.41)$$

wird dann die positiv definite Ljapunov-Funktion $V = \mathbf{x}^T \mathbf{P} \mathbf{x}$ gebildet, für deren Ableitung gilt

$$\begin{aligned} \dot{V} &= \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} \\ &= \mathbf{x}^T \left[\mathbf{A} + \sum_i k_i \Delta \mathbf{A}_i \right]^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \left[\mathbf{A} + \sum_i k_i \Delta \mathbf{A}_i \right] \mathbf{x} \\ &= \mathbf{x}^T \left[\mathbf{A}^T \mathbf{P} + \left(\sum_i k_i \Delta \mathbf{A}_i \right)^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{P} \left(\sum_i k_i \Delta \mathbf{A}_i \right) \right] \mathbf{x} \end{aligned} \quad (14.42)$$

Mit (14.41) wird daraus die Stabilitätsbedingung

$$\dot{V} = \mathbf{x}^T \left[-\mathbf{I} + \sum_i k_i(\mathbf{z}) (\Delta \mathbf{A}_i^T \mathbf{P} + \mathbf{P} \Delta \mathbf{A}_i) \right] \mathbf{x} < 0 \quad (14.43)$$

und unter Berücksichtigung von $\sum_i k_i(\mathbf{z}) = 1$

$$\dot{V} = \mathbf{x}^T \left[\sum_i k_i(\mathbf{z}) (\Delta \mathbf{A}_i^T \mathbf{P} + \mathbf{P} \Delta \mathbf{A}_i - \mathbf{I}) \right] \mathbf{x} < 0 \quad (14.44)$$

Damit lässt sich die Bedingung (14.36) durch die Forderung

$$\Delta \mathbf{A}_i^T \mathbf{P} + \mathbf{P} \Delta \mathbf{A}_i - \mathbf{I} < \mathbf{0} \quad (14.45)$$

ersetzen, die bei entsprechender Struktur des Systems, z.B. nur schwach ausgeprägten Nichtlinearitäten, sicherlich weniger konservative Ergebnisse liefern wird.

Systeme mit Reglern. In den meisten Anwendungsfällen wird das System nicht direkt in der Form (14.31) bzw. (14.35) vorliegen, sondern in der Form (14.25), denn die Stabilitätsanalyse erfolgt normalerweise zusammen mit oder direkt nach dem Reglerentwurf. Daher müsste \mathbf{A}_i in Gleichung (14.36) durch \mathbf{G}_{ij} ersetzt und die Bedingung dann für sämtliche Indexpaare (i, j) überprüft werden. Prinzipiell ist dies natürlich möglich, es führt aber wegen der großen Anzahl zu überprüfender Ungleichungen auf ein sehr umfangreiches LMI-Problem. Eleganter ist dagegen sicherlich, die Darstellung (14.25) des Systems zunächst umzuformulieren und erst dann den Satz 14.2 auf das System anzuwenden (vgl. [59]).

Ausgangspunkt für den kontinuierlichen Fall ist die Gleichung (14.25), deren rechte Seite in zwei Teilsummen zerlegt wird:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \sum_i \sum_j k_i(\mathbf{z}(t)) k_j(\mathbf{z}(t)) \mathbf{G}_{ij} \mathbf{x}(t) \\ \dot{\mathbf{x}}(t) &= \sum_i k_i(\mathbf{z}(t)) k_i(\mathbf{z}(t)) \mathbf{G}_{ii} \mathbf{x}(t) \\ &\quad + 2 \sum_{i < j} k_i(\mathbf{z}(t)) k_j(\mathbf{z}(t)) \left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right] \mathbf{x}(t) \end{aligned} \quad (14.46)$$

Dieselbe Ljapunovfunktion und die gleiche Rechnung wie in (14.37) liefern dann

$$\begin{aligned} \dot{V} &= \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} \\ &= \sum_i k_i^2 \mathbf{x}^T \mathbf{G}_{ii}^T \mathbf{P} \mathbf{x} + 2 \sum_{i < j} k_i k_j \mathbf{x}^T \left[\frac{\mathbf{G}_{ij}^T + \mathbf{G}_{ji}^T}{2} \right] \mathbf{P} \mathbf{x} \\ &\quad + \sum_i k_i^2 \mathbf{x}^T \mathbf{P} \mathbf{G}_{ii} \mathbf{x} + 2 \sum_{i < j} k_i k_j \mathbf{x}^T \mathbf{P} \left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right] \mathbf{x} \end{aligned} \quad (14.47)$$

$$\begin{aligned} &= \sum_i k_i^2 \mathbf{x}^T (\mathbf{G}_{ii}^T \mathbf{P} + \mathbf{P} \mathbf{G}_{ii}) \mathbf{x} \\ &\quad + 2 \sum_{i < j} k_i k_j \mathbf{x}^T \left(\left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right]^T \mathbf{P} + \mathbf{P} \left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right] \right) \mathbf{x} \end{aligned} \quad (14.48)$$

< 0

und damit die folgenden beiden Stabilitätsbedingungen, die beide für sämtliche i und $j > i$ erfüllt sein müssen:

$$\mathbf{G}_{ii}^T \mathbf{P} + \mathbf{P} \mathbf{G}_{ii} < \mathbf{0} \quad (14.49)$$

$$\left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right]^T \mathbf{P} + \mathbf{P} \left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right] < \mathbf{0} \quad \text{für } i < j \quad (14.50)$$

Offensichtlich hat sich durch diese Umformulierung die Anzahl der zu überprüfenden Ungleichungen von $i \times j$ auf rund die Hälfte reduziert. Die Schritte für den diskreten Fall sind analog.

Das LMI-Problem lässt sich noch weiter reduzieren, wenn man berücksichtigt, dass sowohl jedes Teilmodell der Strecke $(\mathbf{A}_i, \mathbf{B}_i)$ als auch jeder Teilregler F_j nur in der Umgebung der jeweiligen Stützstelle i bzw. j aktiv sind. Da die Indizes i und j dieselben Stützstellen beschreiben, folgt daraus, dass das Produkt $k_i(\mathbf{z}(t))k_j(\mathbf{z}(t))$ für weit voneinander entfernt liegende Indizes i und j immer Null ist. Damit müssen aber auch in Bedingung (14.50) nur diejenigen \mathbf{G}_{ij} überprüft werden, deren Indizes zueinander benachbart sind.

Eine andere Abschwächung bzw. Vereinfachung der Stabilitätsbedingungen (14.49) und (14.50) lässt sich erzielen, wenn, wie in Kapitel 14.2.2 schon erwähnt, die Koeffizienten k_i für die Stabilitätsbedingung mit berücksichtigt werden und die Summe in (14.48) als Ganzes betrachtet wird. Bedingung (14.50) resultiert aus der Forderung, dass jeder Summand in (14.48) kleiner als Null sein muss. Dagegen wird im folgenden Ansatz berücksichtigt, dass die einzelnen Summanden durch die k_i gewichtet werden und positive Summanden durch negative Summanden durchaus kompensiert werden können. Denn entscheidend für die Stabilität des Systems ist nur, dass die gesamte Summe negativ ist, nicht jeder einzelne Summand.

Ausgangspunkt der Überlegungen ist Gleichung (14.48), die lediglich in Matrizenform darzustellen ist (vgl. [26]):

$$\dot{V} = \begin{pmatrix} k_1 \mathbf{x} \\ k_2 \mathbf{x} \\ \vdots \\ k_r \mathbf{x} \end{pmatrix}^T \mathbf{X} \begin{pmatrix} k_1 \mathbf{x} \\ k_2 \mathbf{x} \\ \vdots \\ k_r \mathbf{x} \end{pmatrix} < 0 \quad (14.51)$$

mit

$$\mathbf{X} = \begin{pmatrix} \mathbf{L}_{11}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{11} & 2(\mathbf{L}_{12}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{12}) & \cdots & 2(\mathbf{L}_{1r}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{1r}) \\ \mathbf{0} & \mathbf{L}_{22}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{22} & \cdots & 2(\mathbf{L}_{2r}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{2r}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{L}_{rr}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{rr} \end{pmatrix} \quad (14.52)$$

und $\mathbf{L}_{ij} = \frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2}$. Offensichtlich ist diese Ungleichung immer erfüllt, wenn die Matrix \mathbf{X} negativ definit ist. Da die Matrix linear von \mathbf{P} abhängt, kann

auch hier mit einem LMI-Algorithmus überprüft werden, ob überhaupt ein \mathbf{P} existiert, für das die Matrix negativ definit und das System damit stabil ist.

Es stellt sich die Frage nach einem Vergleich zwischen dieser Stabilitätsbedingung und den beiden Bedingungen (14.49) und (14.50). Zunächst ist festzustellen, dass die negative Definitheit der hier entwickelten Matrix insbesondere erfordert, dass sämtliche Matrizen auf ihrer Hauptdiagonalen negativ definit sind, also die Bedingung $\mathbf{L}_{ii}^T \mathbf{P} + \mathbf{P} \mathbf{L}_{ii} < \mathbf{0}$ für alle i erfüllt sein muss. Dies entspricht aber gerade der Bedingung (14.49). Durch die Forderung nach negativer Definitheit der Matrix \mathbf{X} entfällt also nur Bedingung (14.50), während Bedingung (14.49) implizit enthalten ist.

Da aber für die negative Definitheit der Matrix \mathbf{X} nicht ihre sämtlichen Einträge außerhalb der Hauptdiagonalen negativ definit sein müssen, was gleichbedeutend mit Bedingung (14.50) wäre, ist die negative Definitheit dieser Matrix offenbar die weniger strenge Bedingung und demnach für den Stabilitätsnachweis günstiger.

Als weitere Option wird in [59] vorgeschlagen, eine weitere, positiv semidefinite Matrix \mathbf{Q} einzuführen, um zusätzliche Freiheitsgrade bei der Suche nach einer gemeinsamen, positiv definiten Matrix \mathbf{P} zu gewinnen. Aus (14.49) und (14.50) wird dann

$$\mathbf{G}_{ii}^T \mathbf{P} + \mathbf{P} \mathbf{G}_{ii} + (s - 1) \mathbf{Q} < \mathbf{0} \quad (14.53)$$

$$\left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right]^T \mathbf{P} + \mathbf{P} \left[\frac{\mathbf{G}_{ij} + \mathbf{G}_{ji}}{2} \right] - \mathbf{Q} < \mathbf{0} \quad \text{für } i < j \quad (14.54)$$

Dabei ist s die maximale Anzahl zueinander benachbarter Stützstellen, die in die Berechnung der Ausgangsgröße des GS-Systems eingehen. \mathbf{Q} geht wie \mathbf{P} als Unbekannte in den LMI-Algorithmus ein, und dieser liefert dann als Resultat eine Antwort auf die Frage, ob Matrizen \mathbf{P} und \mathbf{Q} existieren, für die das System aus Ungleichungen (14.53) und (14.54) erfüllt ist.

Die untersuchte Lösungsmenge enthält auch die Lösungsmenge des Ungleichungssystems (14.49) und (14.50). Denn da \mathbf{Q} nicht positiv definit, sondern nur positiv semidefinit sein muss, kann \mathbf{Q} auch die Nullmatrix sein. Der Fall $\mathbf{Q} = \mathbf{0}$ und \mathbf{P} beliebig ist demnach implizit in der Untersuchung enthalten. Dieser Fall entspricht aber gerade den Gleichungen (14.49) und (14.50).

Der Gedanke der Erweiterung des Ungleichungssystems zur Gewinnung zusätzlicher Freiheitsgrade wird auch in [23] aufgegriffen. Das Resultat ist ähnlich wie (14.53) und (14.54), so dass hier auf eine Darstellung verzichtet werden kann.

Sämtliche hier vorgestellten Methoden sind auch auf Systeme mit Beobachtern anwendbar. Der Zustandsvektor eines solchen Gesamtsystems enthält dann nicht nur die Zustandsgrößen der Strecke, sondern auch die des Beobachters. Durch geeignetes Zusammenfassen der Zustandsgleichungen kann man das System dann wieder auf die Form (14.25) bringen und unmittelbar

die Stabilitätsbedingungen angeben (vgl. [59, 27, 8]). Die Gleichungen werden dann allerdings sehr umfangreich.

14.2.3 Normenbasierte Stabilitätsanalyse

Die Verwendung von Normen zur Stabilitätsanalyse ist sowohl für kontinuierliche (Gl. (14.27)) als auch diskrete (Gl. (14.28)) Systeme möglich. Den einfacheren Fall stellt dabei der diskrete Fall dar, der deshalb auch zuerst behandelt werden soll (vgl. [7]).

Ausgangspunkt ist das GS-Modell eines diskreten Systems (14.28) ohne äußere Anregung

$$\mathbf{x}(k+1) = \sum_i k_i(\mathbf{x}(k)) \mathbf{A}_i \mathbf{x}(k) \quad (14.55)$$

Der Übergang zu den Normen liefert

$$\|\mathbf{x}(k+1)\| = \left\| \sum_i k_i(\mathbf{x}(k)) \mathbf{A}_i \mathbf{x}(k) \right\| \leq \sum_i k_i(\mathbf{x}(k)) \|\mathbf{A}_i\| \|\mathbf{x}(k)\| \quad (14.56)$$

Das System ist stabil im Ljapunovschen Sinne, wenn der Zustandsvektor gegen Null konvergiert. Als Forderung für Stabilität ergibt sich daher

$$\sum_i k_i(\mathbf{x}(k)) \|\mathbf{A}_i\| < 1 \quad (14.57)$$

Wegen

$$\sum_i k_i(\mathbf{x}(k)) = 1 \quad (14.58)$$

ist die Ungleichung (14.57) erfüllt, wenn

$$\|\mathbf{A}_i\| < 1 \quad \text{für alle } i \quad (14.59)$$

gilt. Verwendet man beispielsweise die ∞ -Norm, so gilt mit (10.26)

$$\|\mathbf{A}_i\|_\infty = \sqrt{\lambda_{\max} \left\{ \bar{\mathbf{A}}_i^T \mathbf{A}_i \right\}} = \bar{\sigma} \{ \mathbf{A}_i \} \quad (14.60)$$

Dabei ist λ_{\max} der maximale Eigenwert oder auch Spektralradius einer Matrix. Die Supremumbildung über ω in Gleichung (10.26) entfällt, da \mathbf{A}_i nur konstante Koeffizienten besitzt. Damit ist mit (10.24) die ∞ -Norm gleich der Spektralnorm $\bar{\sigma} \{ \mathbf{A}_i \}$. Als Stabilitätsforderung erhält man

$$\bar{\sigma} \{ \mathbf{A}_i \} < 1 \quad \text{für alle } i \quad (14.61)$$

Da diese Bedingung sehr einfach mit einem entsprechenden Software-Paket zu überprüfen ist, besteht das einzige Problem bei diesem Analyseverfahren im Aufstellen des zu Grunde liegenden GS-Modells des geschlossenen Kreises.

Nun soll auf ein Verfahren eingegangen werden, das ein kontinuierliches GS-Modell (14.27) des geschlossenen Kreises erfordert (vgl. [58]) und auf dem folgenden Satz basiert:

Satz 14.3 Gegeben sei ein System

$$\dot{\mathbf{x}} = (\mathbf{A} + \mathbf{D}\mathbf{F}(t)\mathbf{E})\mathbf{x} \quad (14.62)$$

mit den gegebenen, reellen Matrizen \mathbf{A} , \mathbf{D} und \mathbf{E} und einer reellen, zeitvarianten Unsicherheit \mathbf{F} , von der nur bekannt ist, dass ihre Norm kleiner als Eins ist: $\|\mathbf{F}\|_\infty \leq 1$. Dieses System ist stabil, wenn \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist und darüber hinaus

$$\|\mathbf{E}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{D}\|_\infty < 1 \quad (14.63)$$

gilt.

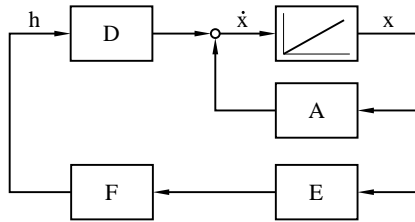


Abb. 14.5. Blockschaltbild des Systems

Der Beweis ist recht einfach anhand des Blockschaltbildes 14.5 des Systems durchzuführen. Der geschlossene Kreis wird bei der Größe \mathbf{h} aufgetrennt. Für die Kreisübertragungsfunktion ergibt sich $\mathbf{F}\mathbf{E}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{D}$ und für ihre Norm die Abschätzung

$$\|\mathbf{F}\mathbf{E}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{D}\|_\infty \leq \|\mathbf{F}\|_\infty \|\mathbf{E}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{D}\|_\infty \quad (14.64)$$

Wegen $\|\mathbf{F}\|_\infty \leq 1$ und (14.63) ist die Norm der Kreisübertragungsfunktion demnach sicher kleiner als Eins. Daraus folgt aber mit dem small gain theorem sofort die Stabilität des Systems.

Die Ungleichung (14.63) ist numerisch leicht zu überprüfen. Es stellt sich aber die Frage, wie das GS-Modell (14.27)

$$\dot{\mathbf{x}}(t) = \sum_i^r k_i(\mathbf{x}(t))\mathbf{A}_i\mathbf{x}(t) \quad (14.65)$$

auf die Form (14.62) zu bringen ist. Zunächst ist jede Matrix \mathbf{A}_i zu zerlegen in einen gemeinsamen Anteil \mathbf{A}_g , der ausschließlich Eigenwerte mit negativem Realteil besitzt, und einen möglichst kleinen Rest $\Delta\mathbf{A}_i$. Man erhält

$$\dot{\mathbf{x}}(t) = \left(\mathbf{A}_g + \sum_i^r k_i(\mathbf{x}(t))\Delta\mathbf{A}_i \right) \mathbf{x}(t) \quad (14.66)$$

Dann lässt sich mit den Matrizen $\Delta\mathbf{A}_i$ eine Singulärwertzerlegung durchführen, d.h.

$$\Delta \mathbf{A}_i = \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T \quad (14.67)$$

mit orthogonalen Matrizen \mathbf{U}_i und \mathbf{V}_i und einer Diagonalmatrix \mathbf{S}_i , die als Diagonalelemente die singulären Werte der Matrix $\Delta \mathbf{A}_i$ enthält. Eine solche Singulärwertzerlegung ist numerisch unproblematisch. Aus (14.66) wird dann

$$\dot{\mathbf{x}} = (\mathbf{A}_g + \mathbf{U} \mathbf{S}(t) \mathbf{V}) \mathbf{x} \quad (14.68)$$

mit

$$\begin{aligned} \mathbf{U} &= [\mathbf{U}_1, \dots, \mathbf{U}_r] \\ \mathbf{V} &= [\mathbf{V}_1, \dots, \mathbf{V}_r]^T \end{aligned} \quad (14.69)$$

und der Diagonalmatrix

$$\mathbf{S}(t) = \text{diag}[k_1(t) \mathbf{S}_1 \dots k_r(t) \mathbf{S}_r] \quad (14.70)$$

\mathbf{S} enthält also die Matrizen \mathbf{S}_i , multipliziert mit $k_i(t)$, auf der Hauptdiagonalen.

Diese Form entspricht schon der im Satz geforderten Form (14.62). Allerdings ist noch nicht gewährleistet, dass $\|\mathbf{S}(t)\|_\infty \leq 1$ gilt. Aus dem Grund wird eine Normierungsmatrix eingeführt mit

$$\mathbf{N} = \frac{1}{2} \text{diag}[\mathbf{S}_1 \dots \mathbf{S}_r] \quad (14.71)$$

Gleichung (14.68) lässt sich damit umschreiben zu

$$\dot{\mathbf{x}} = (\mathbf{A}_g + \mathbf{U} \mathbf{N} \mathbf{V} + \mathbf{U} \mathbf{N} \mathbf{N}^{-1} (\mathbf{S}(t) - \mathbf{N}) \mathbf{V}) \mathbf{x} \quad (14.72)$$

Setzt man

$$\begin{aligned} \mathbf{A} &= \mathbf{A}_g + \mathbf{U} \mathbf{N} \mathbf{V} = \mathbf{A}_g + \sum_i^r \frac{1}{2} \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T = \mathbf{A}_g + \frac{1}{2} \sum_i^r \Delta \mathbf{A}_i \\ \mathbf{D} &= \mathbf{U} \mathbf{N} \\ \mathbf{F}(t) &= \mathbf{N}^{-1} (\mathbf{S}(t) - \mathbf{N}) \\ \mathbf{E} &= \mathbf{V} \end{aligned} \quad (14.73)$$

so erhält man die geforderte Darstellung (14.62), wobei jetzt auch $\|\mathbf{F}\|_\infty \leq 1$ gesichert ist. Denn ein beliebiges Diagonalelement von \mathbf{S} besteht doch aus dem Produkt eines singulären Wertes σ und einem Faktor $k_i(t)$, wobei $0 \leq k_i \leq 1$ gilt. Da singuläre Werte nicht negativ sein können, liegt das betrachtete Diagonalelement damit in einem Intervall $[0, \sigma]$. Durch die Subtraktion $\mathbf{S} - \mathbf{N}$ wird daraus $[-\frac{\sigma}{2}, \frac{\sigma}{2}]$ und durch die Multiplikation mit \mathbf{N}^{-1} gerade $[-1, 1]$. Demnach weisen alle Elemente der Diagonalmatrix $\mathbf{F}(t)$ einen Betrag auf, der maximal gleich Eins ist. Aus dem Grund kann der Betrag des Ausgangsvektors von \mathbf{F} nie größer sein als der Betrag des Eingangsvektors. Wegen (10.25) ist damit die ∞ -Norm von \mathbf{F} nicht größer als Eins.

Auf die in (14.73) aufgeführten Matrizen kann dann Satz 14.3 angewendet werden. Die Wahrscheinlichkeit, dass \mathbf{A} nur Eigenwerte mit negativem Realteil aufweist, ist umso größer, je kleiner die Matrizen $\Delta\mathbf{A}_i$ gewählt wurden. \mathbf{F} erfüllt sicher die geforderte Bedingung, womit dann noch Gleichung (14.63) zu überprüfen ist. Die dazu notwendige Berechnung der Norm lässt sich ebenso wie die Singulärwertzerlegung in (14.67) mit der entsprechenden Software ohne Probleme durchführen.

Ein ähnliches Verfahren für diskrete Systeme wird in [9] vorgestellt. Damit stellt die Verwendung von Normen eine sowohl für den diskreten als auch für den kontinuierlichen Fall einfache und elegante Möglichkeit der Stabilitätsanalyse dar. Voraussetzung ist aber die Darstellung bzw. Darstellbarkeit des geschlossenen Kreises als GS-System.

15. Fuzzy-Systeme

15.1 Fuzzy-Mengen

Die klassische Mathematik basiert auf der Grundannahme, dass allen formallogischen Aussagen immer einer der beiden Wahrheitswerte *wahr* oder *falsch* zugeordnet werden kann. Auf dieser Basis stellt die klassische Mathematik mächtige Werkzeuge zur Problemlösung bereit. Voraussetzung ist, dass sich für eine zu bearbeitende Aufgabe ein formales Modell angeben lässt. Die Beschreibung eines solchen Modells geschieht in einer Terminologie, die sehr viel strikteren Regeln folgt als die natürliche Umgangssprache und aus diesem Grund häufig mit großem Aufwand verbunden ist. Der Lohn für diesen Aufwand besteht dann aber darin, dass sich durch eine solche formale Spezifikation Fehlinterpretationen vermeiden lassen und im Rahmen eines formalen Modells Vermutungen bewiesen oder bisher unbekannte Zusammenhänge abgeleitet werden können.

Trotzdem spielen im alltäglichen Leben formale Modelle bei der Kommunikation zwischen Menschen im Prinzip keine Rolle. Der Mensch ist in der Lage, umgangssprachliche Informationen hervorragend zu verarbeiten, ohne überhaupt an eine Formalisierung der Gegebenheiten zu denken. Beispielsweise kann ein Mensch den Rat, beim langsamen Anfahren nur wenig Gas zu geben, direkt in die Praxis umsetzen. Soll das langsame Anfahren dagegen automatisiert werden, so wäre eine konkrete Anweisung der Art *Drücke das Gaspedal mit einer Geschwindigkeit von einem Zentimeter pro Sekunde herunter* erforderlich. Im Gegensatz zu Anweisungen an einen Menschen muss also eine Automatisierung auf ein formales Modell des technischen oder physikalischen Systems gestützt sein.

Nun dringt aber die Regelungs- und Automatisierungstechnik zunehmend in Bereiche vor, in denen die Erstellung eines formalen Modells nicht oder nur mit sehr großem Aufwand möglich ist. Als Beispiel kann man das *sanfte Umschalten* beim Automatikgetriebe nennen oder die Berücksichtigung der verschiedensten Einflussfaktoren beim Umschaltvorgang. In solchen Fällen liegt tatsächlich kein formales Modell für die Automatisierung vor, sondern nur umgangssprachliche Anweisungen. Hier bleibt keine andere Wahl, als die umgangssprachlichen Anweisungen in irgendeiner Art und Weise in Handlungsanweisungen für ein Automatisierungssystem umzusetzen.

Dabei entsteht das erste Problem schon dadurch, dass ein Mensch in seiner umgangssprachlichen Beschreibung unscharfe oder vage Konzepte wie *schnell*, *sehr groß* oder *kurz* verwendet. Beispielsweise lautet eine umgangssprachliche Handlungsanweisung: *Wenn der Druck im Kessel groß ist, muss das Ventil schnell geöffnet werden*. Bei einem solchen unscharfen Konzept ist aber eine eindeutige Entscheidung, ob ein gegebener Wert das entsprechende Attribut verdient, nicht mehr möglich. Dies hängt zum einen damit zusammen, dass die Attribute eine kontextabhängige Bedeutung haben. Wenn man *schnell* auf Flugzeuge bezieht, so versteht man darunter sicherlich andere Geschwindigkeiten, als wenn man dabei an Autos denkt. Aber selbst wenn der Kontext klar ist, fällt es normalerweise schwer, Autos in *schnell* und *nicht-schnell* aufzuteilen. Und dies liegt nicht daran, dass man nicht den richtigen Wert finden kann, ab dem ein Auto als *schnell* bezeichnet werden soll. Als problematisch wird eher empfunden, sich überhaupt auf einen einzigen Wert festlegen zu müssen. Denn obwohl es natürlich Geschwindigkeiten gibt, die man eindeutig als *schnell* bzw. *nicht schnell* charakterisieren kann, gibt es auch eine Grauzone der mehr oder weniger schnellen Autos.

Als Lösung für dieses Problem bieten sich Fuzzy-Mengen an. Denn die grundlegende Idee der Fuzzy-Mengen besteht darin, die scharfe, zweiwertige Unterscheidung der klassischen Mengenlehre, bei der ein Element entweder vollständig oder gar nicht zur Menge gehört, aufzugeben. Stattdessen lässt man bei Fuzzy-Mengen graduelle *Zugehörigkeitsgrade* zu. Bei einer Fuzzy-Menge muss daher für jedes Element angegeben werden, mit welchem Grad es zur Fuzzy-Menge gehört. Dieser Grad ist ein Wert zwischen Null und Eins, wobei Null dem Attribut *gehört nicht zur Menge* und Eins dem Attribut *gehört zur Menge* entspricht.

Definition 15.1 Eine Fuzzy-Menge oder Fuzzy-Teilmenge A einer Grundmenge X ist charakterisiert durch eine Abbildung $\mu_A : X \rightarrow [0, 1]$, die jedem Element $x \in X$ seinen Zugehörigkeitsgrad $\mu_A(x)$ zu A zuordnet. Dieser liegt im Intervall $[0, 1]$.

Damit kann man eine klassische Menge K als spezielle Fuzzy-Menge auffassen, für die nur die Zugehörigkeitsgrade Null oder Eins existieren:

$$\mu_K : X \rightarrow \{0, 1\}, \quad x \rightarrow \begin{cases} 1 & \text{falls } x \in K \\ 0 & \text{sonst} \end{cases} \quad (15.1)$$

Zu lesen ist diese Schreibweise, die auch im weiteren Verlauf beibehalten werden soll, folgendermaßen: Die Abbildung μ_K ordnet jedem Element x der Menge X einen Wert aus der Menge $\{0, 1\}$, also Null oder Eins zu. Und zwar nach der Funktionsvorschrift $\mu_K(x) = 1$, falls x in K enthalten ist, und Null sonst.

Als Beispiel zeigt Abbildung 15.1 links die *Zugehörigkeitsfunktion* der Menge der Geschwindigkeiten, die größer als 170 km/h sind. Diese Menge ist eine klassische Menge, da für sie nur die Zugehörigkeitsgrade Null oder Eins

existieren. Sie ist aber offenbar keine adäquate Modellierung der Menge aller *hohen* Geschwindigkeiten. Denn nach dieser Definition wäre $169,9 \text{ km/h}$ keine hohe Geschwindigkeit, $170,1 \text{ km/h}$ dagegen doch. Offenbar widerspricht diese Tatsache der umgangssprachlichen Verwendung des Begriffs *hohe Geschwindigkeit*, wonach eine Differenz von $0,2 \text{ km/h}$ aus einer *nichthohen* noch keine *hohe* Geschwindigkeit macht.

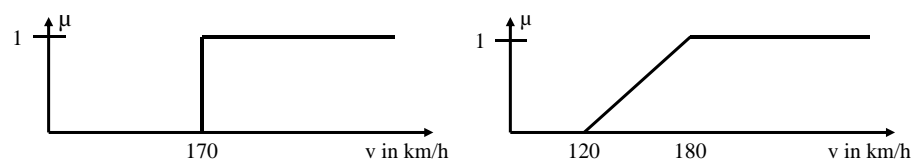


Abb. 15.1. Modellierung der Menge *Hohe Geschwindigkeiten* für Autos

Als Alternative ist rechts die Fuzzy-Menge der *hohen* Geschwindigkeiten dargestellt. Geschwindigkeiten unter 120 km/h gelten als *nicht hoch*, während Geschwindigkeiten über 180 km/h voll und ganz als *hohe* Geschwindigkeiten eingestuft werden. Alle Geschwindigkeiten zwischen diesen beiden Werten gehören nur graduell zur Menge der *hohen* Geschwindigkeiten, und zwar umso mehr, je größer sie sind. Zwischen $170,1 \text{ km/h}$ und $169,9 \text{ km/h}$ besteht demnach, wie es ja auch dem umgangssprachlichen Verständnis entspricht, kaum ein Unterschied. $170,1 \text{ km/h}$ gehört lediglich etwas mehr zur Menge der *hohen* Geschwindigkeiten als $169,9 \text{ km/h}$.

Ein anderes Beispiel ist das Konzept der *Ähnlichkeit*. Wie lässt sich beispielsweise die Menge aller Zahlen beschreiben, die ähnlich zu einer gegebenen reellen Zahl a sind? Bei klassischen Mengen müsste man einen Wert ϵ auswählen und dann festlegen, dass die Zahlen im Intervall $[a - \epsilon, a + \epsilon]$ zu a ähnlich sind und alle anderen Zahlen nicht (vgl. Abb. 15.2, links).

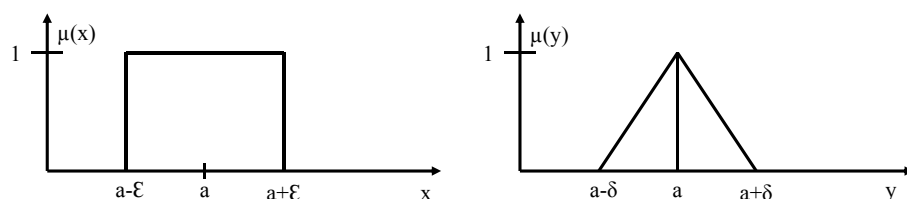


Abb. 15.2. Die Menge aller Zahlen, die a ähnlich sind

Wieder ist rechts als Alternative die Modellierung dieses Konzeptes mit einer Fuzzy-Menge dargestellt. Man sieht, dass a mit dem Zugehörigkeitsgrad Eins in dieser Menge enthalten ist, denn offenbar ist a zu sich selbst ähnlich. Für alle anderen Zahlen sinkt die Ähnlichkeit mit zunehmendem Abstand von a , was auch genau dem entspricht, was man umgangssprachlich unter *ähnlich* versteht. Wird der Abstand größer als δ , so gilt die Zahl als nicht mehr ähnlich zu a .

Bis hier wurden die Fuzzy-Mengen rein intuitiv eingeführt, um ein möglichst gutes Verständnis für die gesamte Idee zu vermitteln. Es bleiben aber einige Punkte, die noch zu klären sind. So stellt sich beispielsweise die Frage, warum als Zugehörigkeitsgrade nur Werte aus dem Einheitsintervall zugelassen sind. Wenn man die klassische Mengenlehre mit ihren Zugehörigkeitsgraden Null oder Eins schon erweitert, hätte doch auch eine andere linear geordnete Menge zur Charakterisierung der Zugehörigkeitsgrade herangezogen werden können. In der Tat finden sich Vorschläge mit anderen Ansätzen. Diese spielen jedoch in den Anwendungen fast keine Rolle, so dass hier weiterhin nur das intuitiv eingeführte Einheitsintervall verwendet werden soll.

Der zweite wichtige Punkt ist die Quantifizierung und Interpretation eines bestimmten Zugehörigkeitsgrades. Aufgrund welcher Kriterien ist beispielsweise festzulegen, ob ein Element mit dem Wert 0,5 oder 0,6 zur Fuzzy-Menge gehört? Und wie soll ein Zugehörigkeitsgrad interpretiert werden? Gehört ein Element mit dem Zugehörigkeitsgrad 1,0 doppelt so sehr zur Fuzzy-Menge wie ein Element mit dem Zugehörigkeitsgrad 0,5?

Eine Antwort auf diese Fragen lässt sich gewinnen, wenn man noch einmal auf die Fuzzy-Menge der *hohen* Geschwindigkeiten zurückgreift: Wie ist hier die Tatsache zu verstehen, dass eine Geschwindigkeit von 180 km/h mit ihrem Zugehörigkeitsgrad 1 einen doppelt so großen Zugehörigkeitsgrad zur Menge der *hohen* Geschwindigkeiten aufweist wie die Geschwindigkeit von 150 km/h mit ihrem Zugehörigkeitsgrad 0,5? Ist damit 180 km/h doppelt so sehr eine *hohe* Geschwindigkeit wie 150 km/h? Eine derartige Interpretation wäre offensichtlich absurd. Sinn macht nur die Feststellung, dass ein Wert von 180 km/h sicherlich mehr zur Menge der *hohen* Geschwindigkeiten gehört als ein Wert von 150 km/h. Jede weitergehende Interpretation würde an der umgangssprachlichen Bedeutung des Attributs *hohe Geschwindigkeit* vorbeigehen.

Formal betrachtet wird also von den verschiedenen Strukturen, die das Einheitsintervall besitzt, lediglich die lineare Ordnung ausgenutzt. Bei zwei gegebenen Zugehörigkeitsgraden kann nur geprüft werden, welcher der größere ist, und damit, welches der zugehörigen Elemente mehr zur Menge gehört. Dagegen wird die kanonische Metrik des Einheitsintervalles, die den Abstand zweier Zahlen quantifiziert, nicht verwendet. Ob sich zwei Zugehörigkeitsgrade um den Wert 0,1 oder 0,001 voneinander unterscheiden, lässt keinen weiteren Schluss zu.

Ein Zugehörigkeitsgrad hat demnach auch keine absolut interpretierbare Bedeutung. Daher kann die Festlegung eines Zugehörigkeitsgrades aufgrund subjektiver Kriterien erfolgen, völlig ungehindert durch irgendwelche Berechnungsvorschriften. Dies ist einerseits natürlich ein Vorteil, weil dem Anwender bei der Definition einer Fuzzy-Menge freie Hand gelassen wird. Andererseits führt diese Tatsache aber auch zu einer Unvergleichbarkeit von Zugehörigkeitsgraden, die von verschiedenen Personen festgelegt wurden. Ein ähnliches

Problem liegt in der Unvergleichbarkeit von Schulnoten, die von verschiedenen Lehrern gegeben wurden.

Aus diesem Dilemma versucht sich die Schule dadurch zu befreien, indem sie objektive Kriterien für die Notengebung vorgibt, wie beispielsweise die Fehlerquote in einer Klausur. Auch im Bereich der Fuzzy-Mengen gibt es Bestrebungen, die Subjektivität bei der Festlegung der Fuzzy-Mengen zu eliminieren. Naheliegend ist beispielsweise die Idee, einen Zugehörigkeitsgrad mit einer berechenbaren Wahrscheinlichkeit zu verknüpfen. Ob dies sinnvoll ist, hängt jedoch ganz entscheidend von der gegebenen Problemstellung ab.

Festzustellen ist aber auf jeden Fall, dass ein Zugehörigkeitsgrad nicht als Wahrscheinlichkeit angesehen werden darf, mit der das Element zur Menge gehört. Dies folgt schon daraus, dass die Zugehörigkeitsfunktion nicht der wahrscheinlichkeitstheoretischen Bedingung für eine Wahrscheinlichkeitsdichte

$$\sum_{x \in X} p(x) = 1 \quad \text{bzw.} \quad \int_X p(x) dx = 1 \quad (15.2)$$

genügt.

Um den Unterschied zwischen Zugehörigkeitsgrad und Wahrscheinlichkeit aber auch anschaulich zu verdeutlichen, sei das folgende Beispiel aus [3] angeführt: U bezeichne die Menge der ungiftigen Flüssigkeiten. Ein Verdurstender erhält zwei Flaschen a und b und die Information, dass die Flasche a mit der Wahrscheinlichkeit 0,9 zu U gehört, während b einen Zugehörigkeitsgrad von 0,9 zu U besitzt. Aus welcher der beiden Flaschen sollte der Verdurstende trinken? Die Wahrscheinlichkeit von 0,9 für a könnte daher stammen, dass die Flasche einem Raum mit zehn Flaschen, von denen neun mit Mineralwasser gefüllt sind und eine mit einer Zyankalilösung, zufällig entnommen wurde. Der Zugehörigkeitsgrad von 0,9 dagegen bedeutet, dass die Flüssigkeit *einigermaßen* trinkbar ist. Beispielsweise könnte sich in b ein Fruchtsaft mit abgelaufenem Haltbarkeitsdatum oder ein mit Glykol versetzter Wein befinden. Es ist daher ratsam, die Flasche b zu wählen, denn schließlich befindet sich in a mit immerhin zehnprozentiger Wahrscheinlichkeit eine giftige Substanz.

Abschließend sollte betont werden, dass Fuzzy-Mengen innerhalb der klassischen Mathematik formalisiert werden, genauso wie die Wahrscheinlichkeitstheorie im Rahmen der herkömmlichen Mathematik formuliert wird. In diesem Sinne eröffnen Fuzzy-Mengen nicht eine neue Mathematik, sondern lediglich einen neuen Zweig der Mathematik.

15.2 Fuzzy-Logik

Dieses Kapitel enthält einen kurzen Ausflug in die Fuzzy-Logik. Behandelt werden sollen die logischen Verknüpfungen *Negation* (NICHT), *Disjunktion* (ODER) und *Konjunktion* (UND) im Rahmen der klassischen Logik so-

wie ihre Erweiterung in der Fuzzy-Logik. Diese Verknüpfungen bilden dann anschließend die Grundlage der mengentheoretischen Operationen *Komplement*, *Vereinigung* und *Durchschnitt* von Fuzzy-Mengen. Diese Operationen sind wiederum die Basis aller Operationen und Schlussfolgerungen auf Fuzzy-Mengen und somit auch die Basis eines Fuzzy-Reglers.

Gegeben seien die Aussagen

$$\begin{aligned}\varphi &: \text{Vier ist eine gerade Zahl} \\ \psi &: \text{Fünf ist eine gerade Zahl}\end{aligned}\tag{15.3}$$

Offenbar ist die erste Aussage wahr und die zweite falsch. Man kann der ersten Aussage demnach den Wahrheitswert 1 zuordnen und der zweiten den Wahrheitswert 0.

Eine Aussage lässt sich negieren:

$$\bar{\varphi} : \text{Vier ist nicht eine gerade Zahl}\tag{15.4}$$

Offenbar ist die negierte Aussage falsch, wenn die Original-Aussage wahr ist und umgekehrt. Entsprechend wechselt der Wahrheitswert von 1 auf 0 oder umgekehrt.

Zwei Aussagen lassen sich auch konjunktiv durch ein logisches UND verknüpfen:

$$\varphi \wedge \psi : \text{Vier ist eine gerade Zahl UND fünf ist eine gerade Zahl}\tag{15.5}$$

Das logische UND ist nun so definiert, dass eine derart zusammengesetzte Aussage nur dann wahr ist, wenn beide Teilaussagen wahr sind. Sind eine oder beide Teilaussagen falsch, so ist auch die zusammengesetzte Aussage falsch. Die obige Aussage ist demnach falsch. Bezogen auf Wahrheitswerte bedeutet dies, dass eine durch UND verknüpfte Aussage nur dann den Wahrheitswert 1 besitzt, wenn beide Teilaussagen den Wahrheitswert 1 besitzen. In allen anderen Fällen besitzt sie den Wahrheitswert 0.

Schließlich bleibt noch die disjunktive Verknüpfung durch ein logisches ODER:

$$\varphi \vee \psi : \text{Vier ist eine gerade Zahl ODER fünf ist eine gerade Zahl}\tag{15.6}$$

Die Definition lautet hier, dass die zusammengesetzte Aussage wahr ist, wenn mindestens eine der beiden Teilaussagen wahr ist. Damit ist die obige Aussage wahr. Der Wahrheitswert einer disjunktiv verknüpften Aussage ist also 1, wenn eine der beiden Teilaussagen den Wahrheitswert 1 aufweist. Nur wenn beide Teilaussagen den Wahrheitswert 0 aufweisen, hat auch die verknüpfte Aussage den Wahrheitswert 0.

Die Annahme, dass eine Aussage nur entweder wahr oder falsch sein kann, erscheint bei der Betrachtung mathematischer Sachverhalte durchaus sinnvoll. In Analogie zu dem, was in Kap. 15.1 bereits über Fuzzy-Mengen

gesagt wurde, lassen sich aber in der Umgangssprache sofort beliebig viele Beispiele angeben, in denen eine strenge Unterteilung in wahre und falsche Aussagen nicht mit dem in Einklang zu bringen wäre, was diese Aussagen tatsächlich bedeuten. Wie soll es beispielsweise bewertet werden, wenn jemand ankündigt, gegen 17.00 Uhr zu erscheinen und dann um 17.05 Uhr eintrifft? War die Ankündigung nun wahr oder falsch? Hier wäre es doch sinnvoll, graduelle Wahrheitswerte für Aussagen zuzulassen, d.h. man kann einer Aussage auch einen Wahrheitswert zwischen 0 und 1 zuweisen. Je näher der Zeitpunkt des Eintreffens an 17.00 Uhr liegt, desto größer war der Wahrheitswert der Ankündigung.

Die Aufgabe besteht nun darin, für solche *unscharfen* Aussagen, die nur graduell wahr sind, die Negation und die logischen Verknüpfungen zu definieren. Doch wie kann man vorgehen? Zunächst einmal ist rein formal festzustellen: Im klassischen Fall entstammen die Wahrheitswerte der Aussagen jeweils der Wertemenge $\{0, 1\}$, bei den unscharfen Aussagen dagegen dem Intervall $[0, 1]$. Im klassischen Fall gibt es demnach beispielsweise für die Negation nur zwei Fälle, die in Betracht zu ziehen sind: Entweder ist die zu negierende Aussage wahr oder falsch. Bei unscharfen Aussagen kann die zu negierende Aussage dagegen jeden beliebigen Wahrheitswert aus dem Intervall $[0, 1]$ aufweisen, so dass hier unendlich viele Fälle zu berücksichtigen sind. Analog verhält es sich bei der logischen Verknüpfung zweier Aussagen. Im klassischen Fall gibt es nur vier mögliche Kombinationen für Wahrheitswerte der beiden Aussagen, während bei unscharfen Aussagen auch hier natürlich wieder unendlich viele Fälle auftreten können.

Es bietet sich daher an, für die Definition der logischen Operationen, also Negation und Verknüpfungen, Funktionen heranzuziehen. Die Negation lässt sich beispielsweise durch eine Funktion $y = n(a)$ festlegen. a ist dabei der Wahrheitswert der gegebenen Aussage und y der Wahrheitswert der negierten Aussage. n ist die Funktion, die die logische Operation *Negation* beschreibt.

Die Definition der konjunktiven Verknüpfung (UND) erfordert eine Funktion mit zwei Argumenten: $y = t(a, b)$. a und b sind die Wahrheitswerte der zu verknüpfenden Aussagen und y der Wahrheitswert der resultierenden Aussage. Die Funktion t beschreibt demnach die logische Operation *konjunktive Verknüpfung*. Und in analoger Weise beschreibt die Funktion $s(a, b)$ mit ebenfalls zwei Argumenten die disjunktive Verknüpfung (ODER).

Nun müssen die drei Funktionen n , t und s nur noch näher spezifiziert werden. Aber welche Anforderungen muss diese Spezifikation erfüllen? Festzuhalten ist zunächst einmal folgendes: Bei unscharfen Aussagen können alle Wahrheitswerte aus dem Intervall $[0, 1]$ auftreten, also auch die Werte 1 und 0. Die klassischen, scharfen Aussagen sind demnach eine Teilmenge der unscharfen Aussagen. Daraus ergibt sich aber sofort die wesentlichste Anforderung an die Spezifikation: Eine logische Operation, definiert für unscharfe Aussagen, muss bei Anwendung auf scharfe Aussagen zu demselben Ergebnis führen wie die klassische Operation.

Die allgemein übliche Definition der Negation unscharfer Aussagen erfüllt diese Forderung:

$$n(a) := 1 - a \quad (15.7)$$

D.h. der Wahrheitswert der negierten Aussage ergibt sich aus Eins minus dem Wahrheitswert der gegebenen Aussage. Hat die gegebene Aussage den Wahrheitswert 0, so hat die negierte Aussage den Wahrheitswert 1 und umgekehrt. Der klassische Fall ist also in dieser Definition enthalten. Darüber hinaus gilt aber auch: Hat die gegebene Aussage den Wahrheitswert 0,5, d.h. ist sie *halb wahr*, so hat auch die negierte Aussage den Wahrheitswert 0,5, was mit der umgangssprachlichen Auffassung durchaus übereinstimmt.

An die Definition der konjunktiven Verknüpfung sind über die obige Forderung hinaus noch einige weitere Forderungen zu stellen, die zwar zunächst trivial erscheinen, aber andererseits notwendig sind, um zu gewährleisten, dass mögliche Definitionen auch mit der umgangssprachlichen Auffassung des logischen UND übereinstimmen.

Zunächst muss gelten, dass der Wahrheitswert der zusammengesetzten Aussage nicht von der Reihenfolge abhängen darf, in der man die Aussagen verknüpft:

$$T1: \quad t(a, b) = t(b, a) \quad (\text{Kommutativität}) \quad (15.8)$$

In ähnlicher Weise darf es bei der Verknüpfung von mehr als zwei Teilaussagen keine Rolle spielen, mit welchen zwei Teilaussagen man die Verknüpfung beginnt:

$$T2: \quad t(a, t(b, c)) = t(t(a, b), c) \quad (\text{Assoziativität}) \quad (15.9)$$

Weiterhin wird gefordert: Wenn der Wahrheitswert a einer Aussage größer ist als der Wahrheitswert b einer anderen Aussage, so darf die Verknüpfung der ersten Aussage mit einer dritten Aussage mit dem Wahrheitswert c auf keinen Fall einen kleineren Wahrheitswert liefern als die Verknüpfung der zweiten Aussage mit der dritten:

$$T3: \quad \text{Aus } a > b \text{ folgt } t(a, c) \geq t(b, c) \quad (\text{Monotonie}) \quad (15.10)$$

Und schließlich darf sich der Wahrheitswert nicht ändern, wenn man eine Aussage mit einer definitiv wahren Aussage verknüpft. Also z.B. *Das Wetter ist schön* ist genauso wahr wie *Das Wetter ist schön* UND $2 > 1$. Formal bedeutet dies:

$$T4: \quad t(a, 1) = a \quad (15.11)$$

Eine Funktion, die alle Forderungen T1–T4 erfüllt, bezeichnet man in der Mathematik als *t-Norm*. Neben den Eigenschaften T1–T4 haben t-Normen aber noch eine andere Eigenschaft. Definiert man nämlich die konjunktive Verknüpfung durch eine t-Norm, so ist automatisch auch gewährleistet, dass sie bei Anwendung auf klassische Aussagen dasselbe Resultat liefert wie die klassische Verknüpfung.

Dies lässt sich leicht beweisen: Denkbar für die klassische konjunktive Verknüpfung zweier Aussagen mit den Wahrheitswerten a und b sind doch nur vier Fälle:

1. $a = 0, b = 0$
2. $a = 0, b = 1$
3. $a = 1, b = 0$
4. $a = 1, b = 1$

Eine t-Norm wird für die Fälle 2 und 4 wegen der Eigenschaft T4 den Wahrheitswert a liefern, bei Fall 2 also 0 und bei Fall 4 den Wert 1. Weiterhin müssen wegen der Kommutativeigenschaft T1 die Fälle 2 und 3 dasselbe Ergebnis liefern, nämlich 0. Und schließlich darf wegen der Monotonieeigenschaft T3 der Wahrheitswert im ersten Fall nicht größer sein als im dritten Fall, womit aber auch hier als Resultat nur 0 in Frage kommt. Damit liefert die t-Norm für den klassischen Fall nur dann den Wahrheitswert 1, wenn beide Teilaussagen den Wahrheitswert 1 aufweisen. Dies entspricht aber gerade der klassischen Definition einer konjunktiven Verknüpfung.

Nach all diesen Betrachtungen ist offensichtlich, dass eine t-Norm eine gute Wahl für die Definition einer konjunktiven Verknüpfung darstellt. So sind auch die beiden am häufigsten verwendeten Definitionen für die konjunktive Verknüpfung unscharfer Aussagen t-Normen:

1. Minimum: $\mu_{\varphi \wedge \psi} := \min\{\mu_{\varphi}, \mu_{\psi}\}$
2. Produkt: $\mu_{\varphi \wedge \psi} := \mu_{\varphi} \cdot \mu_{\psi}$

Dabei ist μ_{φ} der Wahrheitswert der Aussage φ , μ_{ψ} der Wahrheitswert der Aussage ψ und $\mu_{\varphi \wedge \psi}$ der Wahrheitswert der Aussage $\varphi \wedge \psi$.

Die vier Eigenschaften T1 bis T4 lassen sich für beide Beispiele leicht verifizieren.

Analog zur konjunktiven Verknüpfung kann die disjunktive Verknüpfung behandelt werden. Für diese gelten wie für die konjunktive Verknüpfung die Forderungen T1 – T3, während die Forderung T4 durch die Forderung T4' ersetzt wird. Sie lautet: Der Wahrheitswert darf sich nicht ändern, wenn man eine Aussage mit einer definitiv falschen Aussage verknüpft. Also z.B. *Das Wetter ist schön* ist genauso wahr wie *Das Wetter ist schön* ODER $2 < 1$. Formal bedeutet dies:

$$T4' : \quad s(a, 0) = a \quad (15.12)$$

Eine Abbildung bzw. Verknüpfung, die die Forderungen T1 – T3 und T4' erfüllt, bezeichnet man als *t-Conorm*. Analog zur t-Norm beinhaltet die t-Conorm für den klassischen Fall die klassische Definition des logischen ODER. Auf den Beweis soll hier aber verzichtet werden.

Zwei Beispiele für t-Conormen sind:

1. Maximum: $\mu_{\varphi \vee \psi} = \max\{\mu_{\varphi}, \mu_{\psi}\}$
2. Algebraische Summe: $\mu_{\varphi \vee \psi} = \mu_{\varphi} + \mu_{\psi} - \mu_{\varphi} \cdot \mu_{\psi}$

Zwischen t-Normen und t-Conormen besteht ein dualer Zusammenhang. Jede t-Norm t induziert eine t-Conorm s mittels

$$s(a, b) = 1 - t(1 - a, 1 - b) \quad (15.13)$$

d.h. bei gegebener t-Norm ergibt sich die induzierte t-Conorm aus eben dieser Formel.

Ebenso erhält man aus einer t-Conorm s durch

$$t(a, b) = 1 - s(1 - a, 1 - b) \quad (15.14)$$

die entsprechende t-Norm zurück. Mit der Definition der Negation (15.7) entsprechen diese beiden Gleichungen gerade den aus der klassischen Logik bekannten *DeMorganschen Gesetzen*

$$\mu_{\varphi \vee \psi} = \mu_{\overline{\varphi} \wedge \overline{\psi}} \quad \text{und} \quad \mu_{\varphi \wedge \psi} = \mu_{\overline{\varphi} \vee \overline{\psi}} \quad (15.15)$$

Ein zusammengehöriges Paar bilden beispielsweise das *Minimum* als logisches UND sowie das *Maximum* als logisches ODER. Dieses wurde auch schon von L.A. Zadeh in dem Aufsatz [71] angegeben, der die Fuzzy-Logik begründete. Ebenso gehören das *Produkt* und die *algebraische Summe* zusammen.

Beides lässt sich leicht nachweisen: Verwendet man als t-Norm den Minimum-Operator, so ergibt sich aus Gleichung (15.13)

$$s(a, b) = 1 - t(1 - a, 1 - b) = 1 - \min(1 - a, 1 - b) \quad (15.16)$$

Im Falle von $a > b$ wird daraus

$$s(a, b) = 1 - (1 - a) = a \quad (15.17)$$

und für den Fall $b > a$ ergibt sich analog dazu der Wert b . Insgesamt erhält man für die korrespondierende t-Conorm also

$$s(a, b) = \max(a, b) \quad (15.18)$$

Noch einfacher ist die Rechnung für den zweiten Zusammenhang aus Produkt und algebraischer Summe:

$$s(a, b) = 1 - t(1 - a, 1 - b) = 1 - (1 - a)(1 - b) = a + b - ab \quad (15.19)$$

Die Korrespondenz der Dualität von t-Norm und t-Conorm mit den DeMorganschen Gesetzen legt nahe, dass die aus der klassischen Aussagenlogik bekannten Gesetze auch für die unscharfen Aussagen gelten. Dies ist aber nicht immer der Fall, wie am Beispiel der *Idempotenz* erläutert werden soll.

Definition 15.2 Eine t-Norm bzw. die zugehörige konjunktive Verknüpfung ist idempotent, wenn $t(a, a) = a$ für alle $a \in [0, 1]$ gilt.

Dies bedeutet: Wenn man eine Aussage mit sich selbst verknüpft, muss der Wahrheitswert der zusammengesetzten Aussage dem der Aussage selbst entsprechen: *Das Wetter ist schön* UND *Das Wetter ist schön* muss genauso wahr sein wie *Das Wetter ist schön*.

Für klassische Aussagen ist diese Eigenschaft selbstverständlich, nicht aber für unscharfe Aussagen. Ist die t-Norm durch das Minimum definiert, so ist die Verknüpfung idempotent, denn das Minimum aus a und a ist offenbar a . Wenn man die Verknüpfung dagegen durch das Produkt definiert, so ist sie nicht mehr idempotent. Für $a = 0,5$ ergibt sich beispielsweise für die zusammengesetzte Aussage ein Wahrheitswert von $0,25$. Es gilt sogar, dass das Minimum die einzige idempotente t-Norm ist, ebenso wie das Maximum die einzige idempotente t-Conorm darstellt.

Nun könnte man natürlich fordern, dass eine logische Verknüpfung immer idempotent sein muss, damit eine möglichst große Analogie zwischen den neuen und den klassischen Definitionen besteht. Damit würde man sich auf Minimum und Maximum als einzige logische Verknüpfungen beschränken. Dies wäre aber keine glückliche Entscheidung, wie das folgende Beispiel verdeutlicht.

Gegeben seien zwei Häuser, und ein Käufer soll sich für eines der beiden entscheiden. Da sich die Häuser in fast allen Punkten stark ähneln, will er die Wahl aufgrund der beiden Kriterien *niedriger Preis* und *gute Lage* fällen. Er stellt die folgende Tabelle auf, die als Entscheidungsgrundlage dienen soll:

Aussage	Wahrheitswert
1: Der Preis für Haus A ist niedrig.	0,9
2: Die Lage von Haus A ist gut.	0,6
3: Der Preis für Haus B ist niedrig.	0,6
4: Die Lage von Haus B ist gut.	0,6

Beide Häuser haben demnach eine gleich gute Lage, während der Preis von Haus A niedriger ist. Offenbar ist Haus A die bessere Wahl. Zu diesem Ergebnis müsste man also auch gelangen, wenn man streng formal vorgeht. Dazu ist nun für beide Häuser jeweils die verknüpfte Aussage *Der Preis für das Haus ist niedrig* UND *Die Lage des Hauses ist gut* aufzustellen und ihr Wahrheitswert zu ermitteln. Dasjenige Haus, bei dem diese Aussage den größeren Wahrheitswert aufweist, soll gekauft werden.

Die Verknüpfung der beiden Teilaussagen mit der idempotenten Minimum-Norm liefert für Haus A den Wahrheitswert $\min(0,9, 0,6) = 0,6$, und für Haus B $\min(0,6, 0,6) = 0,6$. Beide Wahrheitswerte sind gleich, beide Häuser erscheinen als gleichwertig, und eine Entscheidung ist nicht möglich.

Verwendet man aber eine nicht-idempotente t-Norm für die Verknüpfung, beispielsweise das Produkt, so erhält man für Haus A $0,9 \cdot 0,6 = 0,54$ und für Haus B $0,6 \cdot 0,6 = 0,36$. Hier besitzt die Aussage für Haus A den größeren Wahrheitswert, und eine Entscheidung ist möglich.

Anhand dieses Beispiels wird klar, dass man sich die Auswahl zwischen verschiedenen logischen Verknüpfungen erhalten sollte. Die Idempotenz darf daher keine grundsätzliche Anforderung an die Definition einer logischen Verknüpfung sein. Sinnvoll ist nur die Forderung, dass eine konjunktive Verknüpfung durch eine t-Norm und eine disjunktive Verknüpfung durch die zugehörige t-Conorm definiert werden müssen. Und damit ist ja, wie oben gezeigt wurde, auch gewährleistet, dass eine so definierte Verknüpfung für klassische Aussagen dasselbe Resultat liefert wie die klassische Verknüpfung.

15.3 Operationen auf Fuzzy-Mengen

Mit dem, was bisher über Fuzzy-Mengen und Fuzzy-Logik gesagt wurde, lassen sich nun sehr einfach die mengentheoretischen Operationen *Durchschnitt*, *Vereinigung* und *Komplement* auf Fuzzy-Mengen herleiten. Diese sind, wie schon gesagt, notwendig zum Operieren und Schlussfolgern mit unscharfen Mengen und damit für die Funktion eines Fuzzy-Reglers.

Der Schritt von den logischen Verknüpfungen zu den Mengenoperationen lässt sich am einfachsten anhand eines Beispiels vollziehen: Gegeben sei die Fuzzy-Menge H der *hohen Geschwindigkeiten*, die durch die Zugehörigkeitsfunktion $\mu(v)$ definiert ist (vgl. Abb. 15.1). Man kann sagen, dass diese Menge die Eigenschaft oder das Prädikat *hohe Geschwindigkeit* repräsentiert. Der Zugehörigkeitsgrad einer konkreten Geschwindigkeit v zur Fuzzy-Menge der *hohen Geschwindigkeiten* gibt demnach den Wahrheitswert an, der der Aussage v ist eine *hohe Geschwindigkeit* zuzuordnen ist. In diesem Sinne legt eine Fuzzy-Menge für eine Menge von Aussagen die jeweiligen Wahrheitswerte fest. In unserem Beispiel sind dies alle Aussagen, die man erhält, wenn man für v einen konkreten Geschwindigkeitswert einsetzt.

Damit kann unmittelbar zu den mengentheoretischen Operationen übergegangen werden, und zwar zuerst zur Durchschnittsbildung. Für zwei klassische, scharfe Mengen M_1 und M_2 gilt doch, dass ein Element x genau dann zum Durchschnitt D der beiden Mengen gehört, wenn es sowohl zu M_1 als auch zu M_2 gehört. Oder etwas präziser ausgedrückt: Die Aussage x *gehört zum Durchschnitt D* ist äquivalent zu der zusammengesetzten Aussage x *gehört zu M_1 UND x gehört zu M_2* :

$$x \in D \Leftrightarrow x \in M_1 \wedge x \in M_2 \quad \text{mit } D = M_1 \cap M_2 \quad (15.20)$$

Dies gilt für Fuzzy-Mengen genauso. Nach dem oben Gesagten wird der Wahrheitswert der Aussage x *gehört zu M_j* durch den Zugehörigkeitsgrad $\mu_{M_j}(x)$ ausgedrückt. Und für das logische UND verwendet man gemäß dem letzten Kapitel eine t-Norm, also beispielsweise das Minimum. Schließlich entspricht der Wahrheitswert der Aussage x *gehört zum Durchschnitt D* dem Zugehörigkeitsgrad eines Elementes x zur Menge D . Insgesamt ergibt sich also:

$$\mu_D(x) = \mu_{M_1 \cap M_2}(x) = t(\mu_{M_1}(x), \mu_{M_2}(x)) = \min(\mu_{M_1}(x), \mu_{M_2}(x)) \quad (15.21)$$

Im Folgenden wird der Durchschnittsbildung, wie auch in der Literatur im Allgemeinen üblich, immer das Minimum als t-Norm zugrunde gelegt. Anzumerken bleibt weiterhin, dass sich durch die Definition des Durchschnitts von Fuzzy-Mengen mit Hilfe einer t-Norm die Eigenschaften der t-Norm auf den Durchschnittsoperator übertragen, also insbesondere die Kommutativität und die Assoziativität. Eine graphische Darstellung der Durchschnittsbildung findet sich in Abbildung 15.3.

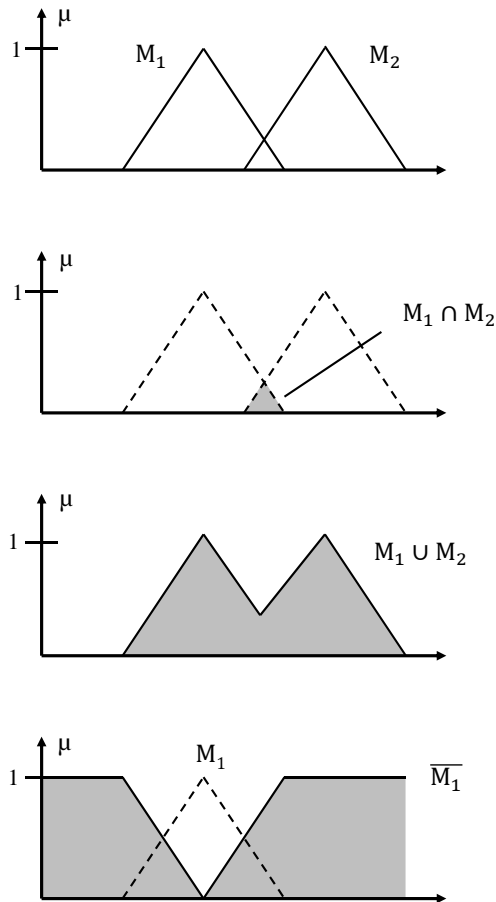


Abb. 15.3. Durchschnitt, Vereinigung und Komplement von Fuzzy-Mengen

In ganz analoger Weise lässt sich auch die Vereinigung von Fuzzy-Mengen herleiten. Ein Element x gehört genau dann zur Vereinigung V zweier klassischer, scharfer Mengen M_1 und M_2 wenn es zu M_1 oder zu M_2 gehört. Oder etwas präziser ausgedrückt: Die Aussage x gehört zur Vereinigung V ist äquivalent zur zusammengesetzten Aussage x gehört zu M_1 ODER x gehört zu

M_2 :

$$x \in V \Leftrightarrow x \in M_1 \vee x \in M_2 \quad \text{mit } V = M_1 \cup M_2 \quad (15.22)$$

Für den Zugehörigkeitsgrad der Vereinigung gilt damit

$$\mu_V(x) = \mu_{M_1 \cup M_2}(x) = s(\mu_{M_1}(x), \mu_{M_2}(x)) = \max(\mu_{M_1}(x), \mu_{M_2}(x)) \quad (15.23)$$

Dabei ist s eine t-Conorm, für die in Zukunft immer das Maximum angenommen werden soll. Auch hier ist noch anzumerken, dass sich die Eigenschaften der t-Conorm auf die Vereinigung übertragen, also insbesondere wieder die Kommutativität und die Assoziativität. Die Vereinigung zweier Mengen ist ebenfalls in Abb. 15.3 dargestellt.

Schließlich bleibt noch das Komplement einer Menge. In der klassischen Mengenlehre besteht das Komplement K einer Menge M aus genau den Elementen, die nicht in der Menge selbst enthalten sind, oder präziser ausgedrückt: Die Aussage x gehört zum Komplement K ist äquivalent zur Aussage x gehört NICHT zu M :

$$x \in K \Leftrightarrow x \notin M \quad \text{mit } K = \bar{M} \quad (15.24)$$

Mit der Definition der Negation einer unscharfen Aussage ergibt sich für den Zugehörigkeitsgrad des Komplementes:

$$\mu_K(x) = 1 - \mu_M(x) \quad (15.25)$$

In Abb. 15.3 findet sich wieder die graphische Darstellung. Zum Komplement ist aber noch etwas sehr Wichtiges anzumerken: Wie beim Komplement für klassische Mengen ist auch die Komplementbildung für Fuzzy-Mengen involutorisch, d.h. das Komplement vom Komplement ist wieder die Ausgangsmenge. Im Gegensatz zu den klassischen Mengen gilt aber nicht, dass der Durchschnitt einer Fuzzy-Menge mit ihrem Komplement immer die leere Menge und die Vereinigung immer die Grundmenge ergibt, wie aus Abbildung 15.4 deutlich zu ersehen ist. Diese Tatsache ist beim Schlussfolgern mit Fuzzy-Mengen unbedingt zu berücksichtigen.

15.4 Fuzzy-Relationen

Relationen eignen sich zur Beschreibung von Zusammenhängen zwischen verschiedenen Variablen, Größen oder Attributen. Eine zweistellige Relation über den Grundmengen X und Y beschreibt beispielsweise den Zusammenhang zwischen jeweils einem Element $x \in X$ und einem Element $y \in Y$. Sinnvollerweise fasst man die Relation selbst ebenfalls als Menge auf, deren Elemente dann diejenigen Wertepaare (x, y) sind, die zusammengehören. Eine Relation ist also im Grunde nichts anderes als eine Menge, deren Elemente Vektoren statt Zahlen sind.

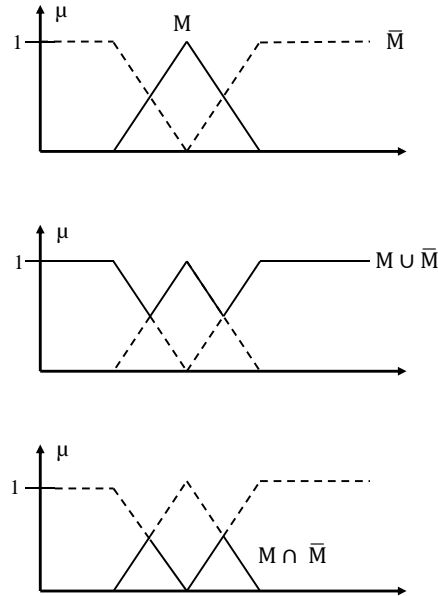


Abb. 15.4. Vereinigung und Durchschnitt einer Fuzzy-Menge mit ihrem Komplement

Beispiel 1: Zur Menge X gehören alle Türen eines Hauses und zur Menge Y alle Schlüssel. Die Menge X enthält die Türen x_1, x_2, x_3 und x_4 und die Menge Y die Schlüssel y_1 bis y_3 . Der Schlüssel y_1 ist ein Generalschlüssel, y_2 passt zur Tür x_2 und y_3 zur Tür x_3 . Jetzt lässt sich formal eine Relation *passt zu* definieren, die den Zusammenhang zwischen den Türen und den Schlüsseln angibt. Diese Relation R ist eine Menge, deren Elemente alle Paare von Türen und Schlüsseln sind, die zueinander passen:

$$R = \{(x_1, y_1), (x_2, y_1), (x_2, y_2), (x_3, y_1), (x_3, y_3), (x_4, y_1)\} \quad (15.26)$$

Beispiel 2: Es sei ein statisches Übertragungsglied gegeben, dessen Ausgangsgröße $y \in Y$ entsprechend der Funktion $y = f(x)$ von der Eingangsgröße $x \in X$ abhängig ist. Wieder lässt sich eine Relation R definieren, die jetzt den Zusammenhang zwischen Ein- und Ausgangswerten beschreibt. Es ist diejenige Menge, die alle zusammengehörigen Wertepaare aus Ein- und Ausgangswerten enthält. Oder anders formuliert, die genau diejenigen Wertepaare enthält, die an diesem Übertragungsglied auftreten können:

$$R = \{(x, y) \mid x \in X \text{ und } y = f(x)\} \quad (15.27)$$

Diese Schreibweise liest sich: Die Menge R enthält alle Wertepaare (x, y) , für die gilt: x ist in X enthalten und $y = f(x)$.

Für diese Menge lässt sich natürlich auch eine Zugehörigkeitsfunktion angeben:

$$\mu_R(x, y) = \begin{cases} 1 & \text{falls } y = f(x) \\ 0 & \text{sonst} \end{cases} \quad (15.28)$$

In Abbildung 15.5 sind die Funktion $f(x)$ und die zugehörige Relation graphisch dargestellt. Man sieht, dass die Relation das Übertragungsverhalten genauso gut beschreibt wie die Kennlinie selbst, da sie exakt dieselbe Information enthält.

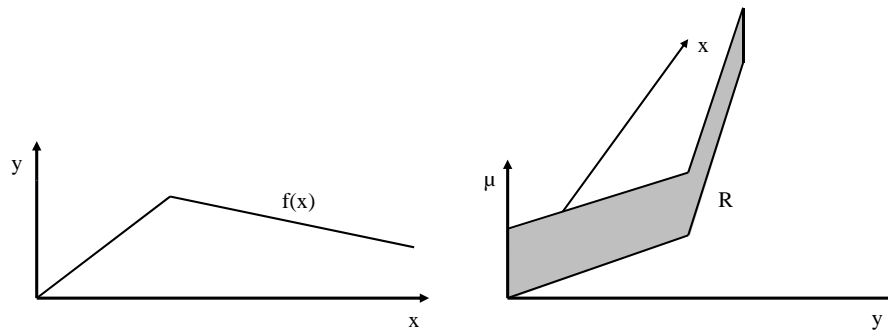


Abb. 15.5. Relation, die ein statisches Übertragungsglied beschreibt

Nun sollen, ebenfalls anhand eines Beispiels, die Fuzzy-Relationen eingeführt werden:

Beispiel 3: Gegeben sei wieder ein statisches Übertragungsglied, dessen Übertragungsverhalten aber nicht durch eine Funktion wie im vorhergehenden Beispiel, sondern nur durch eine umgangssprachliche Beschreibung gegeben ist. Diese Beschreibung lautet:

Wenn die Eingangsgröße x *kalt* ist, dann ist die Ausgangsgröße y *groß*.

Diese Beschreibung nützt wenig, wenn nicht zusätzlich definiert wird, was wir denn unter *kalt* und *groß* zu verstehen haben. Es muss also eine Fuzzy-Menge gegeben sein, die für alle $x \in X$ die Menge aller *kalten* Temperaturen beschreibt, und ebenso für alle $y \in Y$ eine Fuzzy-Menge der *großen* Ausgangswerte. Beide Mengen sind in Abbildung 15.6 dargestellt. Dabei ist zu berücksichtigen, dass beide Fuzzy-Mengen über unterschiedlichen Grundmengen definiert sind. Ihr Durchschnitt oder ihre Vereinigung können daher nicht gebildet werden.

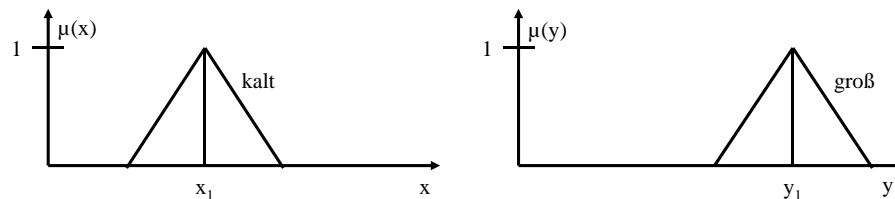


Abb. 15.6. Fuzzy-Mengen der *kalten* Temperaturen und *großen* Ausgangswerte

Nun soll, wie schon im vorhergehenden Beispiel, die Relation R ermittelt werden, die den Zusammenhang zwischen Ein- und Ausgangswerten beschreibt, durch die also ausgedrückt wird, dass die *kalten* Eingangswerte und die *großen* Ausgangswerte zusammengehören. R ist demnach auch hier wieder die Menge aller zusammengehörigen Wertepaare (x, y) , die an diesem Übertragungsglied auftreten können, und beschreibt damit dessen Übertragungsverhalten. In Analogie zu den klassischen Relationen wird es sich um eine Fuzzy-Menge handeln, deren Elemente keine Zahlen, sondern (in diesem Fall zweistellige) Vektoren sind. Umgangssprachlich kann diese Menge sofort angegeben werden: Sie enthält alle Wertepaare (x, y) , bei denen x *kalt* und y *groß* ist. Formal ausgedrückt bedeutet dies:

$$R = \{(x, y) \mid x \text{ ist } \textit{kalt} \text{ UND } y \text{ ist } \textit{gross}\} \quad (15.29)$$

Nun werden aber doch die Wahrheitswerte der Aussagen gerade durch die Zugehörigkeitsgrade zu den entsprechenden Mengen beschrieben. Und für das logische UND soll die t-Norm *Minimum* verwendet werden. Damit ergibt sich für den Zugehörigkeitsgrad eines Wertepaares (x, y) zur Fuzzy-Relation R :

$$\mu_R(x, y) = \min\{\mu_{\textit{kalt}}(x), \mu_{\textit{gross}}(y)\} \quad (15.30)$$

Hierbei ist zu berücksichtigen, dass diese Formel **nicht** eine Durchschnittsmenge beschreibt, da ja x und y zwei völlig verschiedenen Wertemengen entstammen. Stattdessen handelt es sich um eine echte zweidimensionale Fuzzy-Menge, deren Zugehörigkeitsfunktion in Abbildung 15.7 skizziert ist.

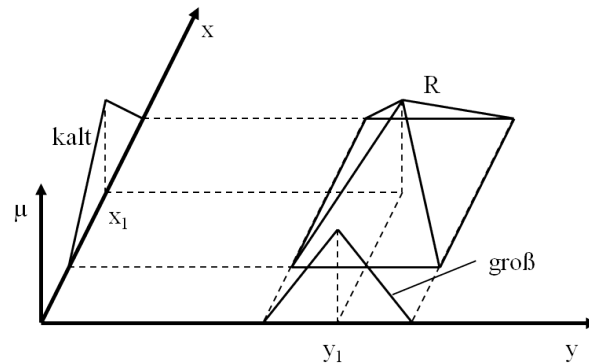


Abb. 15.7. Fuzzy-Relation

Die Konstruktion der Pyramide ist recht einfach. Man projiziert die Fuzzy-Mengen der Ein- und Ausgangsgröße in die $x - y$ -Ebene hinein, und wo sich beide Projektionen überlagern, entsteht die Relation R .

Was drückt nun diese Fuzzy-Relation aus? Wie eine klassische Relation beschreibt sie den durch die obige umgangssprachliche Beschreibung gegebenen Zusammenhang zwischen x - und y -Werten. Oder anders formuliert: Sie

ist eine Menge, die alle Wertepaare enthält, die an diesem Übertragungsglied auftreten können. Allerdings mit einem Unterschied. Bei einer klassischen Relation gibt es doch nur die Zugehörigkeitsgrade Eins oder Null, d.h. ein Vektor gehört entweder zur Menge oder nicht. Oder anders ausgedrückt, zwei Elemente x und y gehören entweder zusammen oder nicht. Bei einer Fuzzy-Relation kann es dagegen, wie bei jeder Fuzzy-Menge, graduelle Zugehörigkeiten der Vektoren zur Menge geben und damit aber auch eine graduelle Zusammengehörigkeit zwischen den Elementen x und y . Wie mit dieser graduellen Zusammengehörigkeit umzugehen ist, wird später noch erläutert. An dieser Stelle sollten nur der Begriff und die Berechnung einer Fuzzy-Relation eingeführt werden.

15.5 Fuzzy-Regler

In diesem Kapitel sollen der Entwurf und die Funktionsweise eines Fuzzy-Reglers anhand eines sehr einfachen Beispiels, und zwar einer Eingrößenregelung, demonstriert werden. Die Ergebnisse sind aber direkt auf beliebig dimensionale Mehrgrößensysteme von beliebiger Komplexität übertragbar.

Geregelt werden soll eine Heizungsanlage. Als Messgröße dient eine gemessene Temperatur, und die Ausgangsgröße des Reglers regelt die Brennstoffzufuhr in den Heizungskessel.

Zuerst wird ein Satz von Fuzzy-Regeln aufgestellt, wobei es sich in unserem Fall nur um zwei Regeln handelt:

1. WENN die Temperatur t kalt ist DANN ist die Ausgangsgröße u groß
2. WENN die Temperatur t heiß ist DANN ist die Ausgangsgröße u klein

Dabei wird der WENN-Teil einer Regel als *Prämisse* und der DANN-Teil als *Konklusion* bezeichnet.

Wenn die Regeln aufgestellt sind, muss definiert werden, was unter *kalt*, *heiß*, *gross* und *klein* zu verstehen ist. Diese vagen Konzepte werden durch Fuzzy-Mengen definiert, wie sie in Abbildung 15.8 dargestellt sind. Das Problematische an diesem Schritt ist die Subjektivität dieser Festlegung. Unterschiedliche Personen haben unterschiedliche Auffassungen von *heiß* oder *kalt*, und eine andere Definition führt offenbar zu einem anderen Regelverhalten. Die geeignete Festlegung der Fuzzy-Mengen kostet daher im allgemeinen mehr Zeit als die Festlegung der Regeln selber.

Jetzt kann die Relation bzw. die Menge aller Wertepaare (t, u) aus Ein- und Ausgangsgröße des Reglers bestimmt werden, die laut Festlegung der Regeln und der Fuzzy-Mengen *kalt*, *klein* usw. auftreten können. Oder einfacher formuliert: R ist die Menge aller erlaubten bzw. möglichen Wertepaare des Fuzzy-Reglers. Somit beschreibt R das Übertragungsverhalten des Reglers.

Zunächst wird jede Regel einzeln betrachtet. Da der Regler ein statisches Übertragungsglied darstellt, lässt sich nämlich für jede Regel entsprechend Gleichung (15.30) eine Teilrelation bilden:

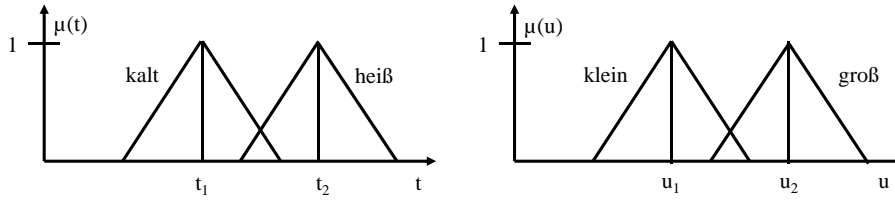


Abb. 15.8. Fuzzy-Mengen für einen Regler

$$\begin{aligned}\mu_{R_1}(t, u) &= \min\{\mu_{kalt}(t), \mu_{gross}(u)\} \\ \mu_{R_2}(t, u) &= \min\{\mu_{heiss}(t), \mu_{klein}(u)\}\end{aligned}\quad (15.31)$$

Jede dieser beiden Teilrelationen stellt eine Menge erlaubter Wertepaare des Reglers dar. Da aber das Übertragungsverhalten des Reglers durch beide Regeln gleichzeitig beschrieben wird, müssen die Relationen bzw. Mengen R_1 und R_2 zu einer Menge R vereinigt werden. Dies ist ohne weiteres möglich, denn beide Relationen sind auf derselben Grundmenge definiert, und zwar auf der Menge aller Wertepaare (t, u) . Sie können also durch eine Mengenoperation miteinander verknüpft werden, weil es für eine solche Operation keine Rolle spielt, ob die Elemente einzelne Werte oder Wertepaare sind:

$$R = R_1 \cup R_2 \quad (15.32)$$

Verwendet man für die Vereinigung bzw. das logische ODER den Maximum-Operator, so folgt für die Zugehörigkeitsfunktion von R :

$$\mu_R(t, u) = \max\{\mu_{R_1}(t, u), \mu_{R_2}(t, u)\} \quad (15.33)$$

Dies ist die Zugehörigkeitsfunktion der Relation bzw. Menge R aller Wertepaare (t, u) aus Ein- und Ausgangsgröße, die beim obigen Fuzzy-Regler per Definition auftreten können d.h. die Menge aller erlaubten Wertepaare. Es soll aber noch einmal daran erinnert werden, dass es sich um eine Fuzzy-Menge handelt und somit die Zugehörigkeit der Wertepaare zu dieser Menge auch graduell sein kann.

Dies wird auch aus der graphischen Darstellung der Relation in Abbildung 15.9 deutlich. Sie besteht aus zwei Pyramiden, wobei jede Pyramide aus einer Regel resultiert. Jede der beiden Pyramiden wurde wie die Pyramide in Abbildung 15.7 konstruiert, und anschließend wurden beide Pyramiden disjunktiv entsprechend Gleichung (15.32) überlagert.

Offensichtlich beinhaltet die so entstandene Fuzzy-Relation R die gesamte Information, die den Fuzzy-Regler beschreibt. Also zum einen die Fuzzy-Regeln, zum anderen aber auch die Form der Zugehörigkeitsfunktionen der Fuzzy-Mengen *heiß*, *kalt*, *groß* und *klein*. Sie werden daher von nun an nicht weiter benötigt.

Bevor sie aber völlig in Vergessenheit geraten, bietet sich noch eine interessante und wichtige Interpretationsmöglichkeit an. Vergleicht man nämlich

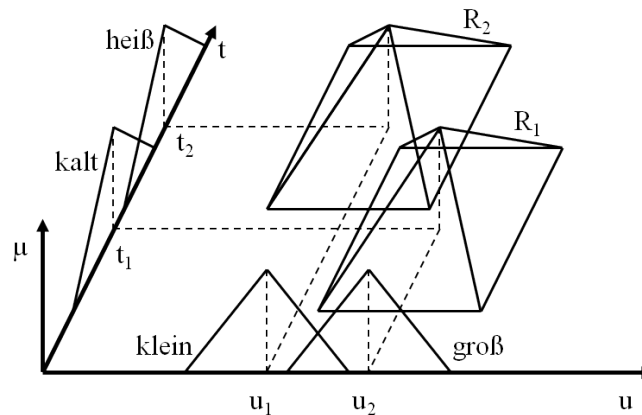


Abb. 15.9. Menge der erlaubten Wertepaare des Fuzzy-Reglers

diese Fuzzy-Mengen mit der in Abbildung 15.2 dargestellten Fuzzy-Menge der zum Wert a ähnlichen Zahlen, so wird offenbar, was die Definition eines Fuzzy-Reglers tatsächlich bedeutet:

Die Menge der *kalten* Temperaturen ist doch nichts weiter als die Menge aller Werte, die *ähnlich* zum Wert t_1 sind. Und analog dazu sind die *großen* Ausgangsgrößen gerade die Werte, die ähnlich zu u_2 sind. Übertragen auf den zweidimensionalen Zusammenhang ist demnach die Teilrelation R_1 die Menge der Wertepaare, die ähnlich zum Wertepaar (t_1, u_2) sind. Die erste Fuzzy-Regel ist also eine **scharfe** Zuordnung des Ausgangswertes u_2 zum Eingangswert t_1 , und durch die Fuzzy-Mengen wird nur ausgedrückt, dass ähnlichen Eingangswerten ein ähnlicher Ausgangswert zugewiesen wird. Was dabei unter *ähnlich* verstanden wird, ergibt sich aus der Definition der Fuzzy-Mengen *kalt* und *groß*, und die ist wiederum kontextabhängig. Dasselbe gilt natürlich für die zweite Regel. Es wird sich im Folgenden zeigen, dass diese Interpretation genau den Kern dessen trifft, was ein Fuzzy-Regler tatsächlich ist.

Bisher haben wir uns eingehend mit der Relation beschäftigt, die das Übertragungsverhalten des Fuzzy-Reglers beschreibt. Doch wie arbeitet nun ein Fuzzy-Regler? Wie wird zu einer gegebenen Eingangsgröße eine Ausgangsgröße ermittelt?

Zur Erklärung sei angenommen, dass der Regler fertig ausgelegt ist und nun zu einem bestimmten Zeitpunkt eine gemessene Temperatur t_m vorliegt. Unter Zuhilfenahme der Relation wird dann zunächst die Menge aller in Frage kommenden Werte der Ausgangsgröße bestimmt. Um diesen Schritt zu verstehen, sollte man sich vorher noch einmal vergegenwärtigen, dass die Relation R die Menge aller erlaubten Wertepaare des Reglers darstellt.

Wenn man nun in Gleichung (15.33) bzw. (15.31) für t den gemessenen Wert t_m einsetzt, so erhält man eine Funktion $\mu_R(t_m, u)$, die nur noch von u abhängig ist. Sie gibt für jeden Wert u den Zugehörigkeitsgrad des Wer-

tepaares (t_m, u) zur Menge der erlaubten Wertepaare des Reglers an. Daher kann man auch direkt eine Fuzzy-Menge U_m definieren mit

$$\mu_{U_m}(u) = \mu_R(t_m, u) \quad (15.34)$$

die man als Menge aller erlaubten Werte der Ausgangsgröße u deuten kann wenn die Eingangsgröße $t = t_m$ ist.

Graphisch interpretiert legt man bei diesem Schritt an der Stelle $t = t_m$ einen Schnitt durch die Relation R parallel zur u -Achse und projiziert diesen auf die u -Achse (Abb. 15.10). Die so entstandene Fuzzy-Menge U_m ist die Menge aller erlaubten Werte für die Ausgangsgröße.

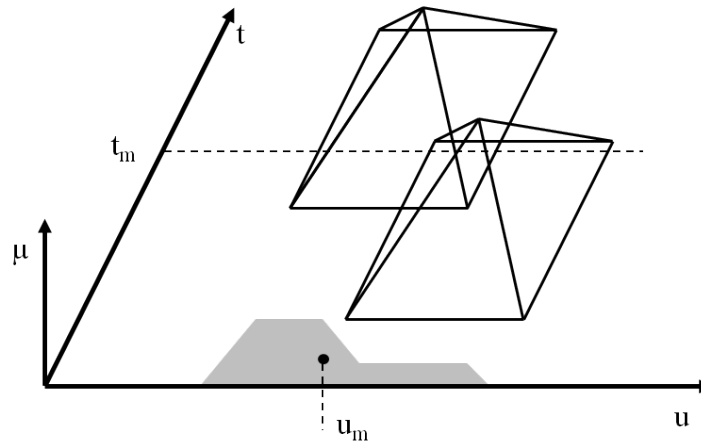


Abb. 15.10. Berechnung der Ausgangsgröße eines Fuzzy-Reglers

Nachdem man die Menge aller erlaubten Werte für die Ausgangsgröße bestimmt hat, muss man sich noch für einen dieser Werte entscheiden. Diesen Schritt bezeichnet man als *Defuzzifizierung*. Die gängigste Methode ist, den Flächenschwerpunkt unter der Kurve der Zugehörigkeitsfunktion zu ermitteln und den zugehörigen Wert u_m der Ausgangsgröße als Stellgröße des Reglers auszugeben. Es werden in der Literatur noch andere Defuzzifizierungsstrategien diskutiert, doch sind diese für den praktischen Einsatz nicht relevant.

Anhand von Abbildung 15.10 ist leicht zu erkennen, dass die Ausgangsgröße u_m umso näher am Wert u_1 liegt, je näher die Eingangsgröße am Wert t_2 liegt. Dasselbe gilt analog für u_2 und t_l . Ist die Eingangsgröße größer als t_2 , so bleibt der Ausgangswert konstant auf u_1 . Für zu große Werte der Eingangsgröße ist der Ausgangswert dann nicht mehr definiert. Analoge Überlegungen gelten für Eingangswerte, die kleiner als t_l sind. Das Übertragungsverhalten des Fuzzy-Reglers lässt sich demnach auch durch eine Kennlinie beschreiben, wie sie in Abbildung 15.11 dargestellt ist.

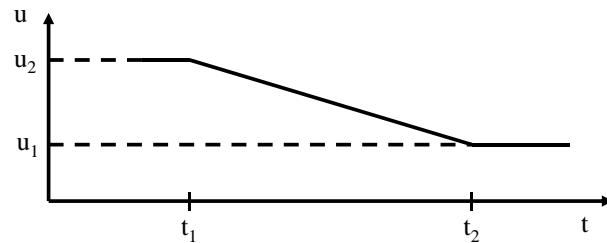


Abb. 15.11. Übertragungsverhalten des Fuzzy-Reglers

Und damit wird auch die obige Interpretation bestätigt: Festgelegt sind die Wertepaare (t_1, u_2) und (t_2, u_1) . Alle anderen Werte der Ausgangsgröße ergeben sich durch Interpolation, also durch Ähnlichkeitsüberlegungen. Die Interpolation muss dabei nicht unbedingt, wie in der Abbildung, linear verlaufen, sondern kann in Abhängigkeit von den Fuzzy-Mengen und der Defuzzifizierungsstrategie auch anders aussehen.

Nach dieser Herleitung lässt sich folgendes festhalten:

- Ein Fuzzy-Regler ist ein Kennlinienregler.
- Durch die Regeln und die Mittelpunkte der Fuzzy-Mengen werden Wertepaare (t_l, u_2) und (t_2, u_1) definiert, die die Stützpunkte der Kennlinie bilden.
- Durch die Form der Fuzzy-Mengen wird festgelegt, was unter *ähnlich* zu verstehen ist. Dies wirkt sich auf die Form der Kennlinie zwischen den Stützpunkten aus. Sie verläuft also nicht unbedingt linear zwischen den Stützpunkten.
- Durch die Defuzzifizierungsstrategie lässt sich ebenfalls die Form der Kennlinie zwischen den Stützpunkten beeinflussen.

Die gesamte Herleitung lässt sich auch auf den Mehrgrößen-Fall mit beliebig vielen Regeln übertragen. Hier ist lediglich *Kennlinie* durch *Kennfeld* zu ersetzen, wobei das Kennfeld auch eine Hyperfläche sein kann.

Interne Dynamik, also beispielsweise integrierende Wirkung, besitzt ein Fuzzy-Regler nicht. Stationäre Regelgenauigkeit lässt sich nur dadurch erzielen, indem man einem Fuzzy-Regler nicht den Regelfehler, sondern von vornherein das Integral des Regelfehlers als Eingangsgröße zur Verfügung stellt. Im Allgemeinen werden Fuzzy-Regler auch im Eingrößenfall sogar mit mehreren Eingangsgrößen verwendet, wobei dem Fuzzy-Regler sowohl der Regelfehler als auch sein Integral und möglicherweise auch noch der differenzierte Regelfehler als Eingangsgrößen zugeführt werden. Alternativ kann auch die Integration des Regelfehlers entfallen und stattdessen die Ausgangsgröße des Fuzzy-Reglers integriert werden.

In der allgemeinen Fassung muss man damit von einer Struktur nach Abbildung 15.12 ausgehen, wobei aber nur einer der beiden Integratoren Verwendung findet. Die Regeln des Fuzzy-Reglers lauten in dieser allgemeinen Fassung

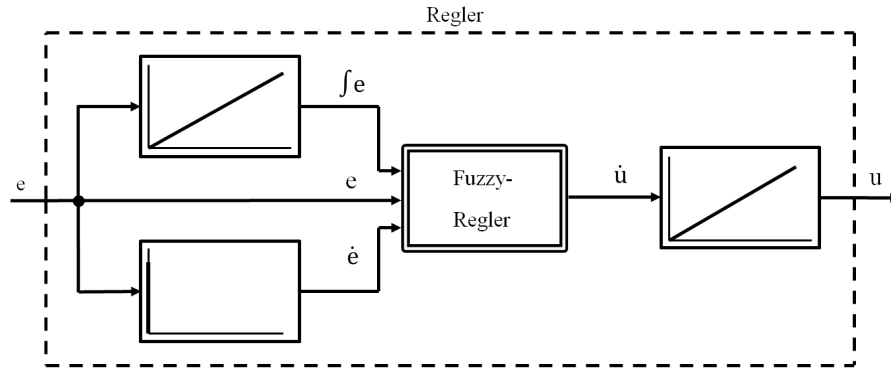


Abb. 15.12. Allgemeine Struktur eines Reglers mit Fuzzy-Kern

$$\text{WENN } e = \dots \text{ UND } \dot{e} = \dots \text{ UND } \int e = \dots \text{ DANN } \dot{u} = \dots \quad (15.35)$$

Für Mehrgrößensysteme gilt diese Darstellung analog.

15.6 Diskussion der Fuzzy-Regler

In diesem Kapitel hat sich gezeigt, dass auch beim Fuzzy-Regler die nach der Defuzzifizierung gewonnenen Ausgangsgrößen wie bei einem klassischen Regler scharf und eindeutig von den Eingangsgrößen abhängig sind. Der Fuzzy-Regler stellt demnach ein nichtlineares Kennfeld ohne interne Dynamik dar. Diesem Kennfeld sind dann im Regelkreis lineare, dynamische Übertragungsglieder zur Integration oder Differentiation vor- oder nachgeschaltet.

Als Kennfeld mit zusätzlichen vor- oder nachgeschalteten linearen Übertragungsgliedern lässt sich aber auch jeder klassisch entworfene Regler darstellen. Hinsichtlich des Verhaltens kann daher zwischen einem Fuzzy-Regler und einem klassischen Regler prinzipiell kein Unterschied bestehen. Der Unterschied zwischen beiden Reglern besteht ausschließlich in der Darstellung und der Entwurfsmethodik. Und ausschließlich hinsichtlich dieser beiden Gesichtspunkte macht es demnach Sinn, die Vor- oder Nachteile eines Fuzzy-Reglers gegenüber dem klassischen Reglerentwurf zu diskutieren.

Nach diesen grundsätzlichen Feststellungen können die Vor- und Nachteile und die daraus resultierenden Anwendungsbereiche von Fuzzy-Reglern diskutiert werden. Zunächst muss bemerkt werden, dass der klassische Reglerentwurf gewisse Vorteile aufweist. Sowohl die Modellbildung der Strecke als auch der darauf aufbauende Reglerentwurf können systematisch erfolgen. Stabilität und gegebenenfalls gewünschte Dämpfungseigenschaften beim Einschwingverhalten sind implizit gewährleistet. Modellunsicherheiten oder -ungenauigkeiten durch Linearisierung können mit Hilfe robuster, insbesondere normoptimaler Regler berücksichtigt werden, so dass auch in diesen Fällen die Stabilität des Regelkreises garantiert werden kann.

Im Gegensatz dazu erfolgt der Entwurf eines Fuzzy-Reglers im allgemeinen auf heuristischem Wege, was sich natürlich im benötigten Zeitaufwand niederschlägt und den Entwurf bei komplexen Strecken sogar unmöglich machen kann. Darüber hinaus kann zunächst einmal keine Gewähr für die Stabilität des entstehenden geschlossenen Kreises übernommen werden. Diese Aussagen sind allerdings insofern zu relativieren, als dass sowohl der systematische Entwurf als auch die Stabilitätsanalyse von Fuzzy-Reglern seit Ende der achtziger Jahre Gegenstand intensiver Forschungstätigkeit geworden sind und mittlerweile einige brauchbare Ansätze auf diesen Gebieten vorliegen.

Ein weiteres zu diskutierendes Argument ist die Robustheit. Fuzzy-Reglern wird wegen der ihnen zu Grunde liegenden Unschärfe eine besonders große Robustheit nachgesagt. Wie oben aber schon erwähnt wurde, weist ein Fuzzy-Regler ein ebenso klar definiertes Übertragungsverhalten auf wie ein klassischer Regler. Daher ist seine Robustheit nichts Geheimnisvolles und lässt sich ebenso diskutieren wie die von klassischen Reglern. Hier ist aber zunächst folgendes klarzustellen: Wie schon in Kapitel 10.5 ausgeführt wurde, macht die Verwendung des Begriffes *Robustheit* nur dann Sinn, wenn man auch quantifizieren kann, wie groß die Abweichungen zwischen nominaler und realer Strecke denn sein dürfen, ohne dass der geschlossene Kreis instabil wird. Denn das unquantifizierte Attribut *robust* trifft praktisch auf jeden Regler zu und ist daher nicht aussagekräftig. Bei Fuzzy-Reglern, die für eine Strecke entworfen wurden, von der kein Modell existiert, ist eine Quantifizierung der Robustheit aber unmöglich. Und auch wenn ein solches Modell zur Verfügung steht, kann die Robustheit normalerweise nur in Form von Simulationsläufen mit veränderten Streckenparametern nachgewiesen werden, also lediglich anhand einiger ausgewählter Beispiele, was natürlich kein echter Beweis für die Robustheit ist. Dagegen existieren im Bereich der klassischen, linearen Regelungstechnik Verfahren, bei denen man für jeden beliebigen Streckenparameter den erwarteten Unsicherheitsbereich vorgeben kann und der mit diesen Verfahren berechnete Regler dann für jede damit mögliche Streckenkonstellation einen stabilen geschlossenen Kreis garantiert.

Schließlich bleibt die Übersichtlichkeit und Anschaulichkeit des Fuzzy-Reglers. Fraglos ist eine Fuzzy-Regel anschaulicher und insbesondere für Nicht-Regelungstechniker leichter zu verstehen als die Übertragungsfunktion eines PI-Reglers oder gar die Koeffizientenmatrix eines Zustandsreglers. Wenn die Strecke und damit der Regler aber eine gewisse Komplexität aufweisen, so besteht der Fuzzy-Regler nicht mehr nur aus einigen wenigen, sondern aus bis zu einigen hundert Regeln. Die Anschaulichkeit jeder einzelnen Regel bleibt dann zwar erhalten, das gesamte Regelwerk ist aber nicht mehr zu überblicken. Die Wirkung der Veränderung einer bestimmten Eingangsgröße kann nur noch vorhergesagt werden, indem man den gesamten Regler durchrechnet. Dagegen lässt sich beispielsweise anhand der Koeffizientenmatrix eines Zustandsreglers noch recht gut ablesen, wie sich die Ausgangsgrößen bei der Veränderung einer Eingangsgröße vergrößern oder verkleinern.

Zusammenfassend ist daher für den sinnvollen Einsatz eines Fuzzy-Reglers festzustellen: Wenn ein Modell der Strecke in Form von Differential- oder Differenzengleichungen vorliegt und es auch möglich ist, anhand dieses Modells (ggf. nach einer Linearisierung) mit klassischen Methoden einen Regler auszulegen, so sollte dies auch versucht werden. Der Einsatz eines Fuzzy-Reglers bietet sich jedoch an, wenn

- kein Streckenmodell in Form von Differenzen- oder Differentialgleichungen zur Verfügung steht.
- die Strecke aufgrund von Nichtlinearitäten eine Struktur aufweist, die den Einsatz klassischer Verfahren unmöglich macht.
- die Regelziele nur unscharf formuliert sind, wie z.B. die Forderung nach *weichem* Umschalten beim Automatikgetriebe eines Kraftfahrzeuges.
- die Strecke und damit die notwendige Regelstrategie so einfach zu überblicken sind, dass der Entwurf eines Fuzzy-Reglers weniger Zeit erfordert als die Modellbildung der Strecke und der Entwurf eines klassischen Reglers.

Daneben besteht auch noch die Möglichkeit, den Fuzzy-Regler nicht im geschlossenen Kreis als echten Regler zu betreiben, sondern auf einer übergeordneten Ebene, beispielsweise zur Vorgabe geeigneter Sollwerte (Bahnplanung), zur Adaption eines klassischen Reglers, zur Prozessüberwachung oder zur Fehlerdiagnose. In diesem Bereich sind der Phantasie des Anwenders keine Grenzen gesetzt. Da hier der Fuzzy-Regler nicht im geschlossenen Kreis arbeitet, existieren auch keine spezifisch regelungstechnischen Probleme wie beispielsweise das Stabilitätsproblem. Bei diesen Anwendungen sind eher sicherheitstechnische Fragen relevant, beispielsweise die Ausfallsicherheit oder ob durch den Fuzzy-Regler tatsächlich alle möglichen Fälle erfasst sind.

Obwohl gerade bei dieser Art von Anwendungen die Vorteile von Fuzzy-Methoden erst richtig zum Tragen kommen, sind sie nicht Gegenstand dieses Buches. Da es sich hierbei um dem Regelkreis übergeordnete Strukturen handelt, sind diese zwangsläufig sehr fachspezifisch, so daß ihre Darstellung den Rahmen dieses Buches sprengen würde. In den verschiedenen Fachzeitschriften finden sich jedoch vielfältige Anwendungsbeispiele aus allen Bereichen der Technik.

15.7 Praktische Aspekte

Abschließend sollen noch kurz einige praktische Aspekte des Einsatzes von Fuzzy-Reglern diskutiert werden. Einer dieser Aspekte ist die Bereitstellung der benötigten Eingangsgrößen für den Fuzzy-Regler. Sofern diese Größen direkt messbar sind, gibt es kein Problem. Wie schon gesagt, sind die Eingangsgrößen von Fuzzy-Reglern aber oft der Regelfehler, sein Integral und eventuell auch seine Ableitung (vgl. die allgemeine Form einer Fuzzy-Regel (15.35)). Die Berechnung des Regelfehlers und seines Integrals bzw. seiner

Summe bereiten keine Probleme. Eine Differenzbildung, um beispielsweise aus zwei aufeinander folgenden Messwerten des Fehlers e die Fehlerdifferenz $\Delta e(k) = e(k) - e(k-1)$ zu erhalten, kann aber schon problematisch sein. Der Grund liegt darin, dass die Differenzbildung einer Hochpassfilterung erster Ordnung entspricht. Eine Hochpassfilterung bedeutet aber eine Verstärkung hochfrequenter Signalanteile, und damit insbesondere auch des Messrauschens. Daher kann eine Differenzbildung nur dann ein brauchbares Ergebnis liefern, wenn das Messsignal praktisch rauschfrei ist.

Erfolgversprechender ist hier sicherlich die Verwendung eines Beobachters. Dieser beinhaltet keine Differenzbildungen oder Differentiationen, sondern lediglich Integrationen, so dass hier keine Rauschsignalverstärkung erfolgen kann. Dafür erfordert ein Beobachter aber ein relativ präzises Streckenmodell in Form von Differenzen- oder Differentialgleichungen. Bei Vorliegen eines solchen Modells stellt sich aber wiederum die Frage, ob ein klassisches Entwurfsverfahren nicht sinnvoller wäre als der Entwurf eines Fuzzy-Reglers.

Die gleiche Frage stellt sich auch, wenn der Entwurf und die Optimierung eines Fuzzy-Reglers anhand von Simulationen erfolgen. Denn für eine Simulation, die brauchbare Rückschlüsse auf das zu erwartende reale Systemverhalten zulässt, wird ebenfalls ein präzises Streckenmodell benötigt.

Abschließend ist der Rechenzeitbedarf zu diskutieren. Grundsätzlich müssen beim Fuzzy-Regler in jedem Abtastschritt sämtliche Regeln abgearbeitet, die Ausgangs-Fuzzy-Menge gebildet und anschließend defuzzifiziert werden. Diese aufwändigen Berechnungen übersteigen aber meistens die Leistungsfähigkeit der zur Verfügung stehenden Prozessoren, insbesondere wenn das Abtastintervall wegen der Dynamik der Strecke klein sein muss. Zur Abhilfe bietet es sich an, für eine ausreichend große Anzahl an Eingangsgrößen die Ausgangsgrößen a priori zu berechnen und in einem Kennfeld abzuspeichern. Zwischen den Werten des Kennfeldes wird dann im laufenden Betrieb interpoliert. Damit entspricht zwar das Übertragungsverhalten eines so implementierten Fuzzy-Reglers normalerweise nicht exakt dem eigentlichen Entwurf, die Unterschiede können jedoch bei ausreichend hoher Auflösung des Kennfeldes vernachlässigt werden.

16. Neural Networks

16.1 Basics

The basic task of Neural Networks is classification. They get trained to compare new inputs with the things they have learned to classify the new input.

A Neural Network consists of neurons and connections between the neurons. Each neuron can be seen as an automat, which has an internal state and an output value, both of them depending on the actual input values and the former internal state value.

The Neural Network normally has different layers of neurons, where in standard Neural Networks each neuron of one layer is connected to all neurons of the next layer. The first layer is the input layer. Each neuron of the input layer gets one input signal, e.g. a pixel information of an image. The last layer is the output layer, where each neuron of the output layer gives one output signal e.g. that classifies an object in the image. The layers in between are called hidden layers. Small Neural Networks have only one or a few hidden layers, while Deep Neural Networks can have up to 100 hidden layers.

Figure 16.1 shows a small Neural Network with 3 inputs, two hidden layers, 3 outputs and a feed-forward-structure. That means that all signals go only from left to right. In Recurrent Neural Networks signals can go in different directions so that internal feedback loops are possible. These Networks can learn to behave like a dynamical system, but they can also get unstable. Here, we will discuss only feed-forward Neural Networks. The left-to-right calculation is called *Forward Propagation*. The learning process, that will be discussed later, goes from right to left, it is called *Back Propagation Learning*.

Now, we will discuss how a Neural Network calculates output values from input values. To get a formal basis, we need a definition first:

Definition 16.1.1. A Neural Network is a tuple (U, W, A, O, NET) where

- U is a finite set of processing units (neurons = units)
- W is the structure of the network: $U \times U \rightarrow \mathbb{R}$ (the connection of each two neurons gets a real value, which is the weight of the connection)
- A is a mapping, that defines an activation function for each neuron (state value of the neuron)

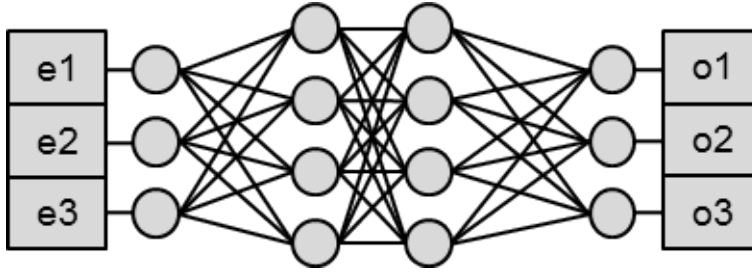


Abb. 16.1. Structure of a feed-forward Neural Network with two hidden layers

- O is a mapping, that defines an output function for each neuron (output value of the neuron)
- NET is a mapping, that defines a network input function for each neuron (a function, that takes all output values from the neurons of the layer before to calculate the network input value of a neuron)

To calculate the output value of a single neuron, in the first step the output values from the neurons of the layer before get collected, and then one network input value is calculated:

$$net_i = NET_i(o_1, o_2, o_3, \dots) \quad (16.1)$$

Usually NET_i is just a weighted sum of the output values from the layer before

$$net_i(k) = \sum_j w_{ji}(k) o_j(k) \quad (16.2)$$

where i is the index of the actual neuron, and j are the indices of the neurons of the layer before. w_{ji} is the weight of the connection between neuron j and the current neuron i . If neuron i is an input layer neuron, the NET_i function is just equal to the network input value for that neuron (see fig. 16.1): $net_i = e_i$.

We will see later, that the learning process affects the weights of the connections. With each training step, the weights of the network connections are modified to improve the final Neural Network output.

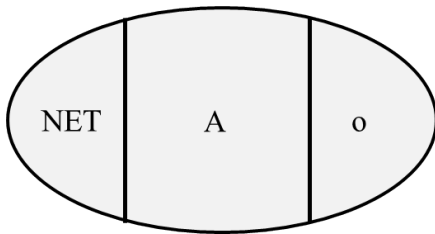


Abb. 16.2. Internal structure of a neuron

In the second step (see fig. 16.2) the internal state of the neuron is calculated, based on the new incoming signal values and on the state value of the last time step:

$$a_i(k) = A_i(net_i(k), a_i(k-1)) \quad (16.3)$$

In the simple case, the last state value is not used to calculate the new one, so that the activation function is only depending on the value net_i . Normally, a sigmoid function is used, which is a monotonically increasing, continuously differentiable function with

$$\lim_{net_i \rightarrow -\infty} a(net_i) = k_1 \quad \text{and} \quad \lim_{net_i \rightarrow \infty} a(net_i) = k_2 \quad \text{with} \quad k_2 > k_1 \quad (16.4)$$

Fig. 16.3 shows a typical activation function of type

$$a_i(k) = \frac{1}{1 + e^{net_i(k)}} \quad (16.5)$$

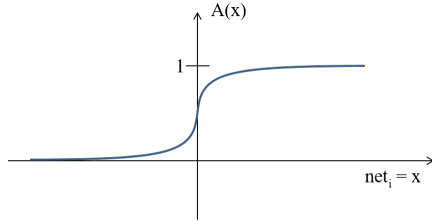


Abb. 16.3. Example of a neuron activation function

The third and final step is the calculation of an output value from the activation value. In this step anything is possible, but the normal case is very simple: $o_i(k) = O_i(a_i(k)) = a_i(k)$.

From this description it is obvious that a Neural Network acts like a characteristic field: For certain inputs certain outputs are calculated, and there are no internal dynamics. The special feature, which differs a Neural Network from a characteristic field, is the training of a Neural Network, which enables us to design or tune it based on data input, while classical characteristic fields have to be designed by hand resp. analytically for each single input value.

There exist different approaches to train a neural network. Generally, it is possible to tune everything, i.e. the number of connections (by adding or deleting connections between neurons), the weights (w_{ij}) of these connections, the number of neurons (by adding or deleting neurons), or the parameters of the activation functions. Here, we will focus on the modification of the weights, which seems to be that type of learning being most similar to the learning of a human brain.

The strategy is to provide a large number of training data sets, consisting of input and reference output data values. The Neural Network gets an input

data set, calculates some output data, this output data is compared to the reference output data, and depending on the difference between the calculated and the reference output data the Neural Network is modified, i.e. the weights of the corresponding connections are increased or decreased. After that, the next training data set is used for the next step.

The error E can be calculated by summing up the square differences of each output value o_i of the output layer neurons with the given reference output r_i for that neuron:

$$E = \sum_i (r_i - o_i)^2 \quad (16.6)$$

Based on this definition, the required change of each connection weight can be calculated from the principle of descending gradients:

$$\Delta w_{ij} = -c \frac{\delta E}{\delta w_{ij}} \quad (16.7)$$

with some adaptation parameter c , that can be chosen arbitrarily.

As an example, we assume a very simple Neural Network with just 6 neurons, as shown in fig. 16.4. For neuron 6 we assume the following functions:

$$\begin{aligned} net_6 &= w_{36}o_3 + w_{46}o_4 + w_{56}o_5 \\ a_6 &= \frac{1}{1 + e^{-net_6}} \\ o_6 &= a_6 n_6 \end{aligned} \quad (16.8)$$

where n_6 is just a normalization factor. For the error we get for this example

$$E = \sum_i (r_i - o_i)^2 = (y - o_6)^2 \quad (16.9)$$

With these definitions, we get for the required change of Δw_{36}

$$\begin{aligned} \Delta w_{36} &= -c \frac{\delta E}{\delta w_{36}} = -2c(y - o_6) \frac{\delta}{\delta w_{36}} (-o_6) \\ &= -2c(y - o_6) \frac{\delta}{\delta w_{36}} (-a_6 n_6) \\ &= 2c(y - o_6) \frac{\delta}{\delta w_{36}} \left(\frac{1}{1 + e^{-net_6}} \right) n_6 \\ &= 2c(y - o_6) \frac{-\frac{\delta}{\delta w_{36}} (e^{-net_6})}{(1 + e^{-net_6})^2} n_6 \\ &= 2c(y - o_6) \left(-\frac{\delta}{\delta w_{36}} (-net_6) \right) e^{-net_6} \frac{1}{(1 + e^{-net_6})^2} n_6 \\ &= 2c(y - o_6) o_3 e^{-net_6} a_6^2 n_6 \end{aligned} \quad (16.10)$$

The weights of the other connections have to be modified in the same way. Of course, this is just one simple approach for the tuning of the weights w_{ij} . Especially for larger Neural Networks, such an approach would require too much calculation time so that more efficient approaches are used.

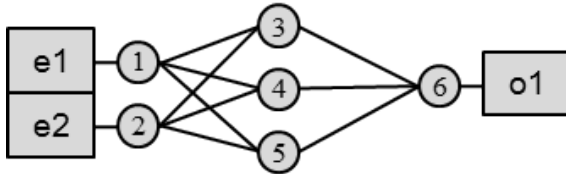


Abb. 16.4. Example Neural Network to demonstrate the Back Propagation Learning

For the training of a Neural Network, the data must get split into a training data set and a (smaller) test data set, that is used to apply the trained Neural Network on formerly unknown data and to check the results. The size of the required training data set must fit to the size of the Neural Network. Small Neural Networks require only a few input-output data examples, while deep-learning Neural Networks normally require many thousand input-output data samples.

A special type of Neural Networks are the so-called Convolutional Neural Networks (CNN). In these networks, there exist layers, which don't have connections from each neuron of one layer to each neuron of the next layer. Instead, there is only a small number of neurons of the preceding layer connected to each neuron of the next layer. That makes sense for image recognition to detect edges and basic structures in the images. Fig. 16.5 shows an example.

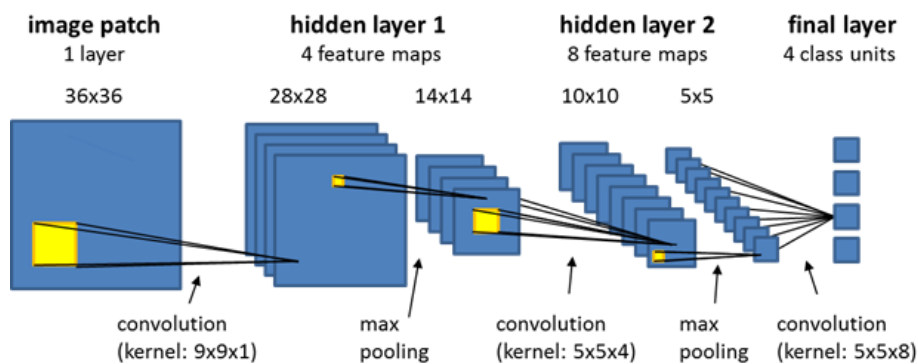


Abb. 16.5. Example of a Convolutional Neural Network

A. Anhang

A.1 Laplace-Transformation

Die Laplace-Transformation ([6, 11, 63]) kann angewendet werden auf eine komplexwertige Funktion $f(t)$ der reellen Variablen t , wenn sie die folgenden Eigenschaften erfüllt:

- $f(t)$ ist für $t \geq 0$ definiert.
- $f(t)$ ist über $(0, \infty)$ integrierbar.
- $f(t)$ unterliegt einer exponentiellen Wachstumsbeschränkung:

$$|f(t)| \leq K e^{ct} \quad c \in \mathbb{R} \quad (\text{A.1})$$

Damit ist die Laplace-Transformation auf Signalverläufe in einem Regelkreis in den meisten Fällen anwendbar. Selbst die Sprungantwort eines Integrators, die ja rampenförmig bis ins Unendliche ansteigt, erfüllt die Wachstumsbeschränkung.

Besonderes Augenmerk verdient die erste der drei Bedingungen. Neben der Tatsache, dass Systeme, in denen für $t < 0$ sämtliche Signale den Wert Null aufweisen, deutlich leichter zu betrachten und analysieren sind, ist es diese Bedingung der Laplace-Transformation, die die Einhaltung dieser Konvention unbedingt erfordert.

Mit der komplexen Variablen s ist die Laplace-Transformierte $f(s)$ der Funktion $f(t)$ definiert als

$$f(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t) dt \quad (\text{A.2})$$

Der Betrag der Laplace-Transformierten lässt sich abschätzen:

$$\begin{aligned} |f(s)| &= \int_0^{\infty} |e^{-st} f(t)| dt \\ &= \int_0^{\infty} |e^{-\operatorname{Re}(s)t} e^{-j\operatorname{Im}(s)t} f(t)| dt \end{aligned} \quad (\text{A.3})$$

Wegen $|e^{-jIm(s)t}| = 1$ und Gleichung (A.1) ergibt sich

$$|f(s)| \leq \int_0^{\infty} |e^{-Re(s)t} K e^{ct}| dt \quad (\text{A.4})$$

Für $Re(s) > c$ gilt $c - Re(s) = \delta < 0$ und damit

$$|f(s)| \leq \int_0^{\infty} |e^{\delta t} K| dt = \left[K \frac{1}{\delta} e^{\delta t} \right]_0^{\infty} = -\frac{K}{\delta} < \infty \quad (\text{A.5})$$

Das Integral in der Definitionsgleichung (A.2) konvergiert demnach absolut für $Re(s) > c$ mit c aus (A.1). In dieser Konvergenz-Halbebene ist $f(s)$ erklärt und eine analytische Funktion von s .

Da t bei einer Anwendung der Transformation auf Signalverläufe die Dimension *Zeit* hat, muss s wegen der Exponentialfunktion in (A.2) die Dimension $Zeit^{-1}$ haben. s ist damit eine komplexe Frequenz. In der Regelungstechnik wird deshalb auch der Bildbereich der Laplace-Transformation als *Frequenz-* und der Originalbereich als *Zeitbereich* bezeichnet.

Unter bestimmten Voraussetzungen lässt sich eine Rücktransformation durchführen:

$$\begin{aligned} f(t) &= \mathcal{L}^{-1} \{f(s)\} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} e^{st} f(s) ds \quad \text{für } t \geq 0 \\ f(t) &= 0 \quad \text{für } t < 0 \end{aligned} \quad (\text{A.6})$$

Der Parameter c ist so zu wählen, dass der Integrationsweg innerhalb der Konvergenz-Halbebene verläuft und c größer als die Realteile aller singulären Punkte von $f(s)$ ist.

Es gelten die folgenden Sätze:

1. **Additionssatz** (Überlagerungssatz)

$$\mathcal{L} \{a_1 f_1(t) + a_2 f_2(t)\} = a_1 \mathcal{L} \{f_1(t)\} + a_2 \mathcal{L} \{f_2(t)\} \quad (\text{A.7})$$

2. **Integrationssatz**

$$\mathcal{L} \left\{ \frac{1}{T} \int_0^t f(\tau) d\tau \right\} = \frac{1}{Ts} \mathcal{L} \{f(t)\} \quad (\text{A.8})$$

3. **Differentiationssatz**

$$\mathcal{L} \left\{ T^n \frac{d^n f(t)}{dt^n} \right\} = s^n T^n \mathcal{L} \{f(t)\} - \sum_{i=1}^n T^{n-1} s^{n-i} \lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{d^{i-1} f(t)}{dt^{i-1}} \quad (\text{A.9})$$

bzw. für die erste Ableitung

$$\begin{aligned}
 \mathcal{L} \left\{ T^1 \frac{d^1 f(t)}{dt^1} \right\} &= T^1 s^1 \mathcal{L} \{ f(t) \} - \sum_{i=1}^1 T^{1-1} s^{1-i} \lim_{\substack{t \rightarrow 0 \\ t > 0}} f(t) \\
 &= Ts \mathcal{L} \{ f(t) \} - T^0 s^0 \lim_{\substack{t \rightarrow 0 \\ t > 0}} f(t) \\
 &= Ts \mathcal{L} \{ f(t) \} - f(0)
 \end{aligned} \tag{A.10}$$

4. Verschiebungssatz

$$\mathcal{L} \{ f(t - T_L) \} = e^{-T_L s} \mathcal{L} \{ f(t) \} \tag{A.11}$$

5. Faltungssatz

$$\mathcal{L} \left\{ \int_0^t f_1(t - \tau) f_2(\tau) d\tau \right\} = \mathcal{L} \{ f_1(t) \} \mathcal{L} \{ f_2(t) \} \tag{A.12}$$

6. Grenzwertsätze

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} s f(s) \quad \text{falls } \lim_{t \rightarrow \infty} f(t) \text{ existiert} \tag{A.13}$$

$$\lim_{\substack{t \rightarrow 0 \\ t > 0}} f(t) = \lim_{s \rightarrow \infty} s f(s) \quad \text{falls } \lim_{\substack{t \rightarrow 0 \\ t > 0}} f(t) \text{ existiert} \tag{A.14}$$

A.2 Korrespondenztabelle zur Laplace-Transformation

$f(t)$ mit $f(t < 0) = 0$	$f(s) = \mathcal{L}\{f(t)\}$
Impuls $\delta(t)$	1
Sprung $s(t)$	$\frac{1}{s}$
$\frac{t^n}{n!}$	$\frac{1}{s^{n+1}}$
$\sin \omega_0 t$	$\frac{\omega_0}{s^2 + \omega_0^2}$
$\cos \omega_0 t$	$\frac{s}{s^2 + \omega_0^2}$
$h(t)e^{-at}$	$h(s + a)$

A.3 Matrizen-e-Funktion

Die Matrizen-e-Funktion ist definiert durch

$$e^{\mathbf{A}t} := \mathbf{I} + \mathbf{A}t + \mathbf{A}^2 \frac{t^2}{2!} + \mathbf{A}^3 \frac{t^3}{3!} + \dots = \sum_{k=0}^{\infty} \mathbf{A}^k \frac{t^k}{k!} \quad (\text{A.15})$$

Dieser Ausdruck beschreibt - ebenso wie die Systemmatrix \mathbf{A} - eine $n \times n$ -Matrix, deren Koeffizienten unendliche Reihen sind.

Zunächst ist das Aussehen einer solchen Reihe zu klären, um die wesentlichen Eigenschaften der Matrizen-e-Funktion herleiten zu können. Offensichtlich gilt für die Koeffizienten von \mathbf{A} , \mathbf{A}^2 und \mathbf{A}^3

$$[\mathbf{A}]_{ik} = a_{ik} \quad (\text{A.16})$$

$$[\mathbf{A}^2]_{ik} = \sum_{\lambda=1}^n a_{i\lambda} a_{\lambda k} \quad (\text{A.17})$$

$$[\mathbf{A}^3]_{ik} = \sum_{\mu=1}^n a_{i\mu} \sum_{\lambda=1}^n a_{\mu\lambda} a_{\lambda k} = \sum_{\mu=1}^n \sum_{\lambda=1}^n a_{i\mu} a_{\mu\lambda} a_{\lambda k} \quad (\text{A.18})$$

Die Koeffizienten der weiteren Matrizen \mathbf{A}^i ergeben sich entsprechend.

Für einen Koeffizienten von $e^{\mathbf{A}t}$ gilt damit

$$[e^{\mathbf{A}t}]_{ik} = e_{ik} + \frac{t}{1!} a_{ik} + \frac{t^2}{2!} \sum_{\lambda=1}^n a_{i\lambda} a_{\lambda k} + \frac{t^3}{3!} \sum_{\mu=1}^n \sum_{\lambda=1}^n a_{i\mu} a_{\mu\lambda} a_{\lambda k} + \dots$$

mit $e_{ik} = \begin{cases} 1 & : i = k \\ 0 & : i \neq k \end{cases} \quad (\text{A.19})$

Mit diesem Ergebnis lässt sich nun untersuchen, ob die Koeffizienten der Matrizen-e-Funktion überhaupt endliche Werte annehmen, d.h. ob die Reihe für jeden Koeffizienten konvergiert. Nur wenn dies der Fall ist, kann mit der Matrizen-e-Funktion sinnvoll gerechnet werden.

Ausgangspunkt der Überlegungen sind die Koeffizienten a_{ik} der Matrix \mathbf{A} . Da diese endlich sind, lassen sie sich sicherlich durch eine obere Schranke abschätzen, d.h. es existiert ein Wert M mit $M > a_{ik}$ für alle Koeffizienten a_{ik} . Mit dieser Schranke ergibt sich für die Koeffizienten der Matrizen-e-Funktion die folgende Abschätzung:

$$\begin{aligned} [e^{\mathbf{A}t}]_{ik} &\leq e_{ik} + \frac{t}{1!} M + \frac{t^2}{2!} n M^2 + \frac{t^3}{3!} n^2 M^3 + \dots \\ &= e_{ik} + \sum_{k=1}^{\infty} \frac{1}{n} \frac{(tnM)^k}{k!} \\ &= e_{ik} + \frac{1}{n} \sum_{k=0}^{\infty} \frac{(tnM)^k}{k!} - \frac{1}{n} \\ &= e_{ik} + \frac{1}{n} e^{tnM} - \frac{1}{n} < \infty \quad \text{für} \quad t < \infty \end{aligned} \quad (\text{A.20})$$

Damit ist gezeigt, dass die Reihe für alle Koeffizienten der Matrizen-e-Funktion konvergiert und diese endliche Werte annehmen.

Nachdem die Konvergenz der Reihe für jeden Koeffizienten sichergestellt ist, können Rechenregeln für die Matrizen-e-Funktion hergeleitet werden. Für das Produkt zweier Matrizen-e-Funktionen gilt zunächst

$$e^{\mathbf{A}t_1} e^{\mathbf{A}t_2} = \left[\sum_{\mu=0}^{\infty} \mathbf{A}^{\mu} \frac{t_1^{\mu}}{\mu!} \right] \left[\sum_{\nu=0}^{\infty} \mathbf{A}^{\nu} \frac{t_2^{\nu}}{\nu!} \right] \quad (\text{A.21})$$

Multipliziert man diese beiden Reihen aus, so gibt es für jeden gegebenen Exponenten k von \mathbf{A} genau k Möglichkeiten, wie sich dieser aus einem Summanden aus der ersten Reihe und einem Summanden aus der zweiten Reihe zusammensetzen kann. Man erhält

$$\begin{aligned} e^{\mathbf{A}t_1} e^{\mathbf{A}t_2} &= \sum_{k=0}^{\infty} \left[\sum_{j=0}^k \mathbf{A}^k \frac{t_1^j}{j!} \frac{t_2^{k-j}}{(k-j)!} \right] \\ &= \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} \left[\sum_{j=0}^k \frac{k!}{j!(k-j)!} t_1^j t_2^{k-j} \right] \\ &= \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} \left[\sum_{j=0}^k \binom{k}{j} t_1^j t_2^{k-j} \right] \\ &= \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} (t_1 + t_2)^k = e^{\mathbf{A}(t_1+t_2)} \end{aligned} \quad (\text{A.22})$$

Daraus folgt auch direkt

$$e^{\mathbf{A}t} e^{-\mathbf{A}t} = e^{\mathbf{A}(t-t)} = e^{\mathbf{0}} = I \quad (\text{A.23})$$

und für die Inversion

$$(e^{\mathbf{A}t})^{-1} = e^{-\mathbf{A}t} \quad (\text{A.24})$$

Für die Ableitung der Matrizen-e-Funktion gilt

$$\begin{aligned} \frac{d}{dt} e^{\mathbf{A}t} &= \mathbf{A} + \mathbf{A}^2 t + \mathbf{A}^3 \frac{t^2}{2!} + \mathbf{A}^4 \frac{t^3}{3!} + \dots \\ &= \mathbf{A} e^{\mathbf{A}t} \quad \text{oder} \quad = e^{\mathbf{A}t} \mathbf{A} \end{aligned} \quad (\text{A.25})$$

Interessant ist der daraus folgende Zusammenhang

$$\mathbf{A} e^{\mathbf{A}t} = e^{\mathbf{A}t} \mathbf{A} \quad (\text{A.26})$$

Ersetzt man in dieser Gleichung \mathbf{A} durch $-\mathbf{A}$, so erhält man

$$\begin{aligned} -\mathbf{A} e^{-\mathbf{A}t} &= e^{-\mathbf{A}t} (-\mathbf{A}) \\ \mathbf{A} e^{-\mathbf{A}t} &= e^{-\mathbf{A}t} \mathbf{A} \end{aligned} \quad (\text{A.27})$$

Schließlich liefert die Multiplikation von Gleichung (A.26) sowohl von rechts als auch von links mit \mathbf{A}^{-1} ein weiteres interessantes Ergebnis:

$$\begin{aligned}\mathbf{A}^{-1}\mathbf{A}e^{\mathbf{A}t}\mathbf{A}^{-1} &= \mathbf{A}^{-1}e^{\mathbf{A}t}\mathbf{A}\mathbf{A}^{-1} \\ e^{\mathbf{A}t}\mathbf{A}^{-1} &= \mathbf{A}^{-1}e^{\mathbf{A}t}\end{aligned}\tag{A.28}$$

A.4 Theorem von Caley-Hamilton

Satz A.1 Die quadratische Matrix \mathbf{A} mit der Dimension $n \times n$ und dem Rang n habe die n Eigenwerte λ_i ($i = 1, \dots, n$). Das charakteristische Polynom dieser Matrix ist definiert als

$$\prod_{i=1}^n (s - \lambda_i) = s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \dots + a_0 \quad (\text{A.29})$$

Dann erfüllt die Matrix \mathbf{A} ihr eigenes charakteristisches Polynom:

$$\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + a_{n-2}\mathbf{A}^{n-2} + \dots + a_0\mathbf{I} = \mathbf{0} \quad (\text{A.30})$$

Darüber hinaus sind die Matrizen \mathbf{A}^i für $1 < i < n - 1$ voneinander linear unabhängig. Und alle Matrizen \mathbf{A}^i mit $i \geq n$ sind als Linearkombination der ersten $n - 1$ Potenzen von \mathbf{A} darstellbar:

$$\mathbf{A}^i = f(\mathbf{A}^{n-1}, \mathbf{A}^{n-2}, \dots, \mathbf{A}) \quad \text{für} \quad i \geq n \quad (\text{A.31})$$

Gleichung (A.30) soll hier nicht bewiesen werden, ebenso nicht die lineare Unabhängigkeit der ersten Potenzen von \mathbf{A} . Gleichung (A.31) lässt sich jedoch relativ einfach aus (A.30) herleiten:

Nach Umstellen von (A.30) ergibt sich

$$\mathbf{A}^n = -a_{n-1}\mathbf{A}^{n-1} - a_{n-2}\mathbf{A}^{n-2} - \dots - a_0\mathbf{I} \quad (\text{A.32})$$

\mathbf{A}^n ist demnach eine Linearkombination der ersten $n - 1$ Potenzen von \mathbf{A} .

Die Multiplikation dieser Gleichung mit \mathbf{A} liefert

$$\mathbf{A}^{n+1} = -a_{n-1}\mathbf{A}^n - a_{n-2}\mathbf{A}^{n-1} - \dots - a_0\mathbf{A} \quad (\text{A.33})$$

und es folgt, dass \mathbf{A}^{n+1} eine Linearkombination der vorangegangenen \mathbf{A}^k (mit $k < n + 1$) ist. Da aber \mathbf{A}^n eine Linearkombination der ersten $n - 1$ Potenzen von \mathbf{A} ist, lässt sich folglich auch \mathbf{A}^{n+1} als Linearkombination dieser ersten $n - 1$ Potenzen von \mathbf{A} darstellen. Der Beweis für alle weiteren Potenzen kann offensichtlich analog erfolgen.

A.5 Die Ljapunov-Gleichung

Satz A.2 *Es sei \mathbf{Q} eine symmetrische, positiv definite Matrix. Dann ist die Lösung \mathbf{P} der Ljapunov-Gleichung*

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q} \quad (\text{A.34})$$

genau dann positiv definit, wenn die Matrix \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist.

Beweis (vgl. [25]): Zunächst soll vorausgesetzt werden, dass \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist. Daraus ist die positive Definitheit von \mathbf{P} zu folgern. \mathbf{x} sei ein Zustandsvektor des Systems

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x} \quad (\text{A.35})$$

Weiterhin gilt unter Verwendung von (A.34):

$$\begin{aligned} \frac{\partial}{\partial t}(\mathbf{x}^T \mathbf{P} \mathbf{x}) &= \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{P} \mathbf{A} \mathbf{x} = -\mathbf{x}^T \mathbf{Q} \mathbf{x} \end{aligned} \quad (\text{A.36})$$

Die Integration dieser Gleichung liefert

$$\mathbf{x}^T(\infty) \mathbf{P} \mathbf{x}(\infty) - \mathbf{x}^T(0) \mathbf{P} \mathbf{x}(0) = - \int_0^\infty \mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) dt \quad (\text{A.37})$$

Da \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist, ist das System (A.35) stabil im Ljapunovschen Sinne, d.h. der Zustandsvektor \mathbf{x} konvergiert aus jedem Anfangszustand gegen Null: $\mathbf{x}(\infty) = \mathbf{0}$. Zudem gilt wegen der positiven Definitheit von \mathbf{Q}

$$\mathbf{x}^T \mathbf{Q} \mathbf{x} > 0 \quad (\text{A.38})$$

für alle \mathbf{x} . Insgesamt wird damit aus (A.37)

$$\mathbf{x}^T(0) \mathbf{P} \mathbf{x}(0) = \int_0^\infty \mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) dt > 0 \quad (\text{A.39})$$

Da diese Ungleichung für jeden Anfangszustand $\mathbf{x}(0)$ erfüllt ist, muss auch \mathbf{P} positiv definit sein.

In anderer Richtung wird nun die positive Definitheit von \mathbf{P} vorausgesetzt. Daraus ist zu folgern, dass \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil besitzt. Zunächst lässt sich für das System (A.35) eine Ljapunov-Funktion

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x} \quad (\text{A.40})$$

angeben, die wegen der positiven Definitheit von \mathbf{P} sicher ebenfalls positiv definit ist. Für die Ableitung dieser Ljapunov-Funktion gilt mit (A.36)

$$\dot{V}(\mathbf{x}) = -\mathbf{x}^T \mathbf{Q} \mathbf{x} \quad (\text{A.41})$$

Diese Ableitung ist wegen der positiven Definitheit von \mathbf{Q} sicherlich negativ definit. Daraus folgt die Stabilität des Systems (A.35) und damit wiederum die Tatsache, dass \mathbf{A} ausschließlich Eigenwerte mit negativem Realteil aufweist.

A.6 Statistische Grundbegriffe

In diesem Abschnitt sollen einige statistische Grundbegriffe erläutert werden, jedoch nur in dem Umfang, wie es das Verständnis einiger Passagen in diesem Buch erfordert.

Der *Erwartungswert* oder *Mittelwert* eines Signales $x(t)$ ist definiert durch

$$\bar{x} := E(x(t)) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt \quad (\text{A.42})$$

Die *Varianz* des Signales ist definiert durch

$$\text{Var}(x) := E((x(t) - \bar{x})^2) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (x(t) - \bar{x})^2 dt \quad (\text{A.43})$$

und ist ein Maß für die durchschnittliche Abweichung der Signalwerte vom Mittelwert.

Die *Standardabweichung* des Signales ist die positive Quadratwurzel der Varianz:

$$\sigma_x := \sqrt{\text{Var}(x)} \quad (\text{A.44})$$

Diese ist nicht zu verwechseln mit dem *Effektivwert* eines Signales. Dieser ist definiert durch

$$X_{eff} := \sqrt{E(x^2(t))} := \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x^2(t) dt} \quad (\text{A.45})$$

Bei der Varianz bzw. der Standardabweichung wird der Mittelwert abgezogen, bevor quadriert wird. Dies ist bei der Berechnung des Effektivwertes nicht der Fall. Für mittelwertfreie Signale entspricht die Standardabweichung aber dem Effektivwert.

Die *Kovarianz* zweier Signale $x(t)$ und $y(t)$ ist definiert durch

$$\text{Cov}(x, y) := E((x(t) - \bar{x})(y(t) - \bar{y})) \quad (\text{A.46})$$

Je größer der Zusammenhang zwischen zwei Signalen ist, desto größer ist ihre Kovarianz. Offensichtlich ist

$$\text{Var}(x) = \text{Cov}(x, x) \quad (\text{A.47})$$

Diese Definitionen lassen sich auch auf Signalvektoren

$$\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T \quad (\text{A.48})$$

übertragen. Der Erwartungswert wird zu einem Erwartungsvektor, dessen einzelne Koeffizienten die Erwartungswerte der einzelnen Signalkomponenten von $\mathbf{x}(t)$ sind:

$$\bar{\mathbf{x}} := E(\mathbf{x}(t)) := (E(x_1(t)), \dots, E(x_n(t)))^T = (\bar{x}_1, \dots, \bar{x}_n)^T \quad (\text{A.49})$$

Die Varianz eines Signales wird für Signalvektoren zu einer Matrix

$$\text{Var}(\mathbf{x}) := E((\mathbf{x}(t) - \bar{\mathbf{x}})(\mathbf{x}(t) - \bar{\mathbf{x}})^T) \quad (\text{A.50})$$

wobei die Erwartungswerte wie bei Vektoren für jeden Matrizenkoeffizienten einzeln gebildet werden. In der Hauptdiagonale findet sich jeweils die Varianz $\text{Var}(x_i)$, und in den übrigen Matrizenelementen jeweils an der Stelle (i, j) die Kovarianz $\text{Cov}(x_i, x_j)$. Wegen $\text{Cov}(x_i, x_j) = \text{Cov}(x_j, x_i)$ ist die Varianzmatrix nicht nur quadratisch, sondern auch symmetrisch.

Analog dazu lässt sich für zwei Vektoren die *Kovarianzmatrix* angeben:

$$\text{Cov}(\mathbf{x}, \mathbf{y}) := E((\mathbf{x}(t) - \bar{\mathbf{x}})(\mathbf{y}(t) - \bar{\mathbf{y}})^T) \quad (\text{A.51})$$

Natürlich gilt auch hier

$$\text{Var}(\mathbf{x}) = \text{Cov}(\mathbf{x}, \mathbf{x}) \quad (\text{A.52})$$

Oft wird auch die obige Varianzmatrix schon als Kovarianzmatrix bezeichnet, da sie in jedem Element (i, j) die Kovarianz der Signalkomponenten $x_i(t)$ und $x_j(t)$ enthält.

Aufbauend auf Varianz und Kovarianz kann nun die *Autokorrelationsfunktion* (AKF) eingeführt werden. Sie ist für stationäre Prozesse definiert durch

$$\psi_{xx}(\tau) := E(x(t)x(t+\tau)) := \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau)dt \quad (\text{A.53})$$

Da die hier verwendeten Signalverläufe für negative Zeitpunkte immer den Wert Null aufweisen, muss die Definition reduziert werden auf das Intervall $[0, \infty]$:

$$\psi_{xx}(\tau) := E(x(t)x(t+\tau)) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt \quad (\text{A.54})$$

Die AKF gibt den inneren Zusammenhang eines Signales an: Für jeden Wert von τ zeigt die AKF an, wie groß der Zusammenhang zwischen den Werten des Signals und den im Abstand τ vorhergehenden Werten ist. Offensichtlich ist der innere Zusammenhang des Signales mit sich selbst am größten, so dass die AKF an der Stelle $\tau = 0$ ihr Maximum aufweist. Sofern der Mittelwert des Signales Null ist, entspricht dieser Wert der Varianz des Signales, wie

ein Vergleich mit (A.43) sofort zeigt. Ebenso offensichtlich ist die AKF eine gerade Funktion ($\psi_{xx}(\tau) = \psi_{xx}(-\tau)$) und weist auch dieselbe Periode auf, wie das eigentliche Signal, sofern dieses periodisch ist.

Entsprechend der AKF lässt sich die *Kreuzkorrelationsfunktion* (KKF) definieren. Sie lautet

$$\psi_{xy}(\tau) := E(x(t)y(t+\tau)) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t+\tau)dt \quad (\text{A.55})$$

Die Kreuzkorrelationsfunktion kennzeichnet den Zusammenhang zwischen den Signalen $x(t)$ und $y(t)$. Offenbar ist - wie bei Varianz und Kovarianz - die AKF ein Spezialfall der KKF.

Auch die Definitionen für AKF und KKF lassen sich auf Signalvektoren übertragen. Für die AKF eines Signalvektors entsteht hier wie bei der Varianz eine Matrix:

$$\psi_{\mathbf{x}\mathbf{x}}(\tau) := E(\mathbf{x}(t) \mathbf{x}^T(t+\tau)) \quad (\text{A.56})$$

wobei die Erwartungswerte wieder für jeden Matrizenkoeffizienten einzeln gebildet werden. Die Elemente der AKF-Matrix sind die KKF der einzelnen Komponenten des Signalvektors. Das Element (i, j) enthält beispielsweise die KKF $\psi_{x_i x_j}(\tau)$. Für $\tau = 0$ entsteht bei mittelwertfreien Signalen gerade die Varianzmatrix. Da die AKF eine gerade Funktion ist, ist die AKF-Matrix symmetrisch.

Analog dazu lässt sich für zwei Vektoren die KKF-Matrix angeben

$$\psi_{\mathbf{x}\mathbf{y}}(\tau) := E(\mathbf{x}(t) \mathbf{y}^T(t+\tau)) \quad (\text{A.57})$$

die für $\tau = 0$ und mittelwertfreie Signale der Kovarianzmatrix entspricht. Sowohl AKF-Matrix als auch KKF-Matrix werden in der Literatur teilweise als Kovarianzmatrix bezeichnet.

Für ein spezielles Signal, das *weiße Rauschen*, sollen die obigen Werte und Matrizen angegeben werden. Das weiße Rauschen ist ein stationärer Prozess, der alle Frequenzen gleichermaßen enthält, mit sich selbst unkorreliert und mittelwertfrei ist. Für die nachfolgend angegebenen Zusammenhänge sei angenommen, dass sämtliche Signale und Signalkomponenten aus weißem Rauschen bestehen. Es gilt:

$$\begin{aligned} \bar{\mathbf{x}} &= E(\mathbf{x}(t)) = \mathbf{0} \\ \text{Var}(\mathbf{x}) &= \mathbf{K} \\ \psi_{xx}(\tau) &= E(x(t) x(t+\tau)) = k \delta(\tau) \\ \psi_{\mathbf{x}\mathbf{x}}(\tau) &= E(\mathbf{x}(t) \mathbf{x}^T(t+\tau)) = \mathbf{K} \delta(\tau) \end{aligned} \quad (\text{A.58})$$

Dabei ist k eine konstante Zahl und \mathbf{K} eine konstante Matrix. Wenn das weiße Rauschen in allen Komponenten von \mathbf{x} zueinander unkorreliert ist, dann sind alle Koeffizienten von \mathbf{K} außerhalb der Hauptdiagonalen Null. $\delta(\tau)$ bezeichnet den Dirac-Impuls, eine Funktion, die an der Stelle $\tau = 0$ einen unendlich hohen Wert aufweist und ansonsten Null ist.

A.7 Lineare Matrixungleichungen

Die Darstellung in diesem Abschnitt folgt im wesentlichen der Darstellung in [56].

Das Grundproblem in der Theorie der linearen Matrixungleichungen (linear matrix inequalities, LMI's) kann folgendermaßen formuliert werden: Gegeben sei eine symmetrische Matrix, deren Koeffizienten affin von gewissen freien Parametern abhängen. Kann man diese freien Parameter dann so wählen, dass die symmetrische Matrix negativ definit wird, d.h. ausschließlich negative Eigenwerte besitzt?

Dabei ist eine affine Funktion eines Parameters α definiert durch $f(\alpha) = a\alpha + b$, wobei a und b Konstanten sind. Eine affine Funktion ist demnach eine lineare Funktion, erweitert um einen konstanten Anteil.

Mit \mathbf{x} als Vektor der freien Parameter und $\mathbf{F}(\mathbf{x})$ als symmetrische Matrix, deren Koeffizienten affin von den freien Parametern abhängen, lässt sich das Grundproblem auch anders definieren: Existiert ein Vektor \mathbf{x} , für den

$$\mathbf{F}(\mathbf{x}) < 0 \quad (\text{A.59})$$

gilt?

Wegen der Affinität der Matrixfunktion $\mathbf{F}(\mathbf{x})$ ist die Lösungsmenge für \mathbf{x} immer konvex. Und dies ist der entscheidende Grund dafür, dass die Lösbarkeitsfrage mittels passender numerischer Algorithmen vollständig behandelt werden kann. Ein Beispiel dafür ist die Matlab LMI-Toolbox, mit deren Hilfe zum einen die Frage beantwortet werden kann, ob überhaupt eine Lösung \mathbf{x} des Problems existiert, und die dann, sofern eine Lösung existiert, eine solche Lösung auch berechnet.

Die Aufgabe reduziert sich mit einem derartigen Tool darauf, ein vorhandenes Problem in die Form der Gleichung (A.59) zu bringen. Dabei sollen die folgenden Bemerkungen helfen:

- Eine LMI der Form $\mathbf{G}(\mathbf{x}) > 0$ ist zu (A.59) mit $\mathbf{G} = -\mathbf{F}$ äquivalent.
- Bei der Matrixungleichung $\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{Q} < 0$ mit der gesuchten Matrix \mathbf{P} lässt sich $\mathbf{F}(\mathbf{x}) = \mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{Q}$ setzen. Die Koeffizienten von \mathbf{P} bilden dabei den Vektor \mathbf{x} der gesuchten Parameter. Offensichtlich sind die Koeffizienten von \mathbf{F} affin von diesen Parametern abhängig. Um die Symmetrie von \mathbf{F} zu gewährleisten, müssen aber \mathbf{P} und \mathbf{Q} symmetrisch sein.
- Die Negativität einer Blockmatrix lässt sich mittels der Negativität ihrer Blöcke charakterisieren:

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{pmatrix} < 0 &\Leftrightarrow \mathbf{A} < 0 \quad \text{und} \quad \mathbf{B} - \mathbf{C}^T \mathbf{A}^{-1} \mathbf{C} < 0 \\ &\Leftrightarrow \mathbf{B} < 0 \quad \text{und} \quad \mathbf{A} - \mathbf{C} \mathbf{B}^{-1} \mathbf{C}^T < 0 \end{aligned} \quad (\text{A.60})$$

Die Ausdrücke $\mathbf{B} - \mathbf{C}^T \mathbf{A}^{-1} \mathbf{C}$ und $\mathbf{A} - \mathbf{C} \mathbf{B}^{-1} \mathbf{C}^T$ werden als *Schur-Komplemente* der Blockmatrix bezüglich der Blöcke \mathbf{A} und \mathbf{B} bezeichnet.

Mit Hilfe der Schur-Komplemente lässt sich beispielsweise die *Riccati-Ungleichung*

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} - \mathbf{P} \mathbf{B} \mathbf{B}^T \mathbf{P} + \mathbf{Q} < \mathbf{0}, \quad (\text{A.61})$$

die keine affine, sondern wegen des dritten Summanden eine quadratische Ungleichung für die gesuchte Matrix \mathbf{P} darstellt und damit zunächst nicht als lineare Matrixungleichung behandelt werden kann, in eine (A.59) entsprechende Form bringen:

$$\begin{pmatrix} \mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{Q} & \mathbf{P} \mathbf{B} \\ \mathbf{B}^T \mathbf{P} & -\mathbf{I} \end{pmatrix} < \mathbf{0} \quad (\text{A.62})$$

Bei dieser Form hängen die einzelnen Koeffizienten der Matrix offensichtlich nur noch affin von den Koeffizienten von \mathbf{P} ab.

Setzt man in Gleichung (A.60) $\mathbf{C} = \mathbf{0}$, so folgt daraus, dass eine blockdiagonale Matrix genau dann negativ definit ist, wenn dies für jeden einzelnen ihrer Blöcke zutrifft.

Damit lässt sich wiederum das System endlich vieler linearer Matrixungleichungen

$$\mathbf{F}_1(\mathbf{x}) < \mathbf{0} \quad , \dots , \quad \mathbf{F}_n(\mathbf{x}) < \mathbf{0} \quad (\text{A.63})$$

mit $\mathbf{F}(\mathbf{x}) = \text{diag}(\mathbf{F}_1(\mathbf{x}), \dots, \mathbf{F}_n(\mathbf{x}))$ auf die Form (A.59) bringen.

Insbesondere das System aus Ungleichungen

$$\mathbf{A}_i^T \mathbf{P} + \mathbf{P} \mathbf{A}_i < \mathbf{0} \quad \text{mit } i = 1, \dots, n \quad (\text{A.64})$$

für eine gesuchte Matrix \mathbf{P} lässt sich sehr einfach auf die Form (A.59) bringen:

$$\begin{pmatrix} \mathbf{A}_1^T \mathbf{P} + \mathbf{P} \mathbf{A}_1 & & \\ & \dots & \\ & & \mathbf{A}_n^T \mathbf{P} + \mathbf{P} \mathbf{A}_n \end{pmatrix} < \mathbf{0} \quad (\text{A.65})$$

Die Lösungsmengen von (A.64) und (A.65) sind nach dem vorher gesagten offensichtlich äquivalent.

Literaturverzeichnis

1. J. Ackermann. *Abtastregelung, Band I: Analyse und Synthese*. Springer-Verlag, Berlin, 1983.
2. M. A. Aisermann und F. R. Gantmacher. *Die absolute Stabilität von Regelsystemen*. Oldenbourg-Verlag, München, 1965.
3. J. C. Bezdek. Fuzzy Models - What Are They, and Why? *IEEE Transactions on Fuzzy Systems*, 1:1–5, 1993.
4. H. W. Bode. *Network Analysis and Feedback Amplifier Design*. D. van Nostrand, Princeton/New Jersey, 1945.
5. J. Böcker, I. Hartmann und C. Zwanzig. *Nichtlineare und adaptive Regelungssysteme*. Springer-Verlag, Berlin, 1986.
6. I. N. Bronstein und K. A. Semendjajew. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt/Main, 1983.
7. S. G. Cao, N. W. Rees und G. Feng. Stability Analysis of Fuzzy Control Systems. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 26:201–204, 1996.
8. Y.-Y. Cao und P. M. Frank. Analysis and Synthesis of Nonlinear Time-Delay Systems via Fuzzy Control Approach. *IEEE Transactions on Fuzzy Systems*, 8:200–211, 2000.
9. Y.-Y. Cao und P. M. Frank. Robust H_∞ Disturbance Attenuation for a Class of Uncertain Discrete-Time Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*, 8:406–415, 2000.
10. L. Cremer. Ein neues Verfahren zur Beurteilung der Stabilität linearer Regelsysteme. *Zeitschrift für angewandte Mathematik und Mechanik*, 25(27):161, 1947.
11. O. Föllinger. *Laplace- und Fourier-Transformation*. Elitera-Verlag, Berlin, 1977.
12. O. Föllinger. *Regelungstechnik*. Hüthig-Verlag, Heidelberg, 1992.
13. O. Föllinger. *Nichtlineare Regelungen, Band I*. Oldenbourg-Verlag, München, 1993.
14. O. Föllinger. *Nichtlineare Regelungen, Band II*. Oldenbourg-Verlag, München, 1993.
15. A. Gelb und W. E. V. Velde. *Multiple-Input Describing Functions and Nonlinear System Design*. McGraw-Hill, New York, 1968.
16. K. Göldner und S. Kubik. *Mathematische Grundlagen der Systemanalyse*. Verlag Harri Deutsch, Frankfurt/Main, 1983.
17. W. Hahn. *Stability of Motion*. Springer-Verlag, Berlin, 1967.
18. A. Hurwitz. Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt. *Math. Annalen*, 46:273, 1895.
19. R. Isermann. *Digitale Regelsysteme, Band I*. Springer-Verlag, Berlin, 1988.
20. R. Isermann. *Digitale Regelsysteme, Band II*. Springer-Verlag, Berlin, 1988.

21. R. Isermann. *Identifikation dynamischer Systeme, Band I*. Springer-Verlag, Berlin, 1992.
22. R. Isermann. *Identifikation dynamischer Systeme, Band II*. Springer-Verlag, Berlin, 1992.
23. M. Johansson, A. Rantzer und K.-E. Arzen. Piecewise Quadratic Stability of Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*, 7:713–723, 1999.
24. R. E. Kalman. On the General Theory of Control Systems. In *Proc. 1st International Congress on Automatic Control 1960, Bd. 1*, pages 481–492, London. Butterworths, 1961.
25. H. Kiendl. Totale Stabilität von linearen Regelungssystemen bei ungenau bekannten Parametern der Regelstrecke. *Automatisierungstechnik*, 33:379–386, 1985.
26. E. Kim und H. Lee. New Approaches to Relaxed Quadratic Stability Condition of Fuzzy Control Systems. *IEEE Transactions on Fuzzy Systems*, 8:523–534, 2000.
27. K. Kiriakidis. Fuzzy Model-Based Control of Complex Plants. *IEEE Transactions on Fuzzy Systems*, 6:517–530, 1998.
28. H.-J. Kowalsky und G. O. Michler. *Lineare Algebra, 10. Auflage*. de Gruyter, Berlin, 1995.
29. R. Krishnapuram. Fitting an Unknown Number of Lines and Planes to Image Data Through Compatible Cluster Merging. *Pattern Recognition*, 25:385–400, 1992.
30. J. LaSalle und S. Lefschetz. *Die Stabilitätstheorie von Ljapunow*. Bibliographisches Institut, Mannheim, 1967.
31. A. J. Laub. A Schur Method for Solving Algebraic Riccati Equation. *IEEE Transactions on Automatic Control*, 24:913–921, 1979.
32. A. Leonhard. Ein neues Verfahren zur Stabilitätsuntersuchung. *Archiv für Elektrotechnik*, 38:17, 1944.
33. W. Leonhard. *Statistische Analyse linearer Regelsysteme*. Teubner-Verlag, Stuttgart, 1973.
34. W. Leonhard. *Einführung in die Regelungstechnik*. Vieweg-Verlag, Braunschweig, 1985.
35. W. Leonhard. *Digitale Signalverarbeitung in der Meß- und Regelungstechnik*. Teubner-Verlag, Stuttgart, 1989.
36. M. A. Ljapunov. Problème général de la stabilité du mouvement (Übersetzung aus dem Russischen). *Ann. Fac. Sci.*, 9:203, 1907.
37. L. Ljung. *System Identification - Theory for the User*. Prentice Hall, Englewood Cliffs, New Jersey, 1987.
38. D. G. Luenberger. Observing the State of a Linear System. *IEEE Transactions on Military Electronics*, 8:74–80, 1964.
39. D. G. Luenberger. An Introduction to Observers. *IEEE Transactions on Automatic Control*, 16:596–602, 1971.
40. A. W. Michailow. Die Methode der harmonischen Analyse in der Regelungstheorie (russ.). *Automat. Telemek.*, 3:27, 1938.
41. R. R. Mohler. *Nonlinear Systems, Vol. I, Dynamics and Control*. Prentice Hall, New Jersey, 1991.
42. R. R. Mohler. *Nonlinear Systems, Vol. II, Applications to Bilinear Control*. Prentice Hall, New Jersey, 1991.
43. K. Müller. *Entwurf robuster Regelungen*. Teubner-Verlag, Stuttgart, 1996.
44. N. N. IEEE Transactions on Automatic Control: Special Issue on Linear-Quadratic-Gaussian Estimation and Control Problem. *IEEE Transactions on Automatic Control*, 16:527–869, 1971.

45. R. Noisser und E. Bodenstorfer. Zur Stabilitätsanalyse von Fuzzy-Regelungen mit Hilfe der Hyperstabilitätstheorie. *Automatisierungstechnik*, 45:76–83, 1997.
46. H. Nyquist. Regeneration Theory. *Bell System Technical Journal*, 11:126, 1932.
47. R. Palm. Sliding Mode Fuzzy Control. In *Proc. of IEEE International Conference on Fuzzy Systems*, San Diego, CA. IEEE, 1992.
48. P. C. Parks und V. Hahn. *Stabilitätstheorie*. Springer-Verlag, Berlin, 1981.
49. H.-P. Opitz. *Entwurf robuster, strukturvariabler Regelungssysteme mit der Hyperstabilitätstheorie*. VDI-Verlag GmbH, Düsseldorf, 1984.
50. H.-P. Opitz. Die Hyperstabilitätstheorie - Eine systematische Methode zur Analyse und Synthese nichtlinearer Systeme. *Automatisierungstechnik*, 34:221–230, 1986.
51. V. M. Popov. The Solution of a New Stability Problem for Controlled Systems. *Automatic and Remote Control*, 24:1–23, 1963.
52. V. M. Popov. *Hyperstability of Control Systems*. Springer-Verlag, Berlin, 1973.
53. E. P. Popow. *Dynamik automatischer Regelsysteme*. Akademie-Verlag, Berlin, 1958.
54. E. J. Routh. *Stability of a Given State of Motion*. Adams Prize Essay, London, 1877.
55. H. Schwarz. *Nichtlineare Regelungssysteme*. Oldenbourg-Verlag, München, 1991.
56. C. W. Scherer. Lineare Matrixungleichungen in der Theorie der robusten Regelung. *Automatisierungstechnik*, 45:306–318, 1997.
57. K. Tanaka und M. Sugeno. Stability analysis and design of fuzzy control systems. *Fuzzy Sets and Systems*, 45:135–156, 1992.
58. K. Tanaka, T. Ikeda und H. O. Wang. Robust Stabilization of a Class of Uncertain Nonlinear Systems via Fuzzy Control: Quadratic Stabilizability, H_∞ Control, and Linear Matrix Inequalities. *IEEE Transactions on Fuzzy Systems*, 4:1–13, 1996.
59. K. Tanaka, T. Ikeda und H. O. Wang. Fuzzy Regulators and Fuzzy Observers: Relaxed Stability Conditions and LMI-Based Designs. *IEEE Transactions on Fuzzy Systems*, 6:250–265, 1998.
60. M. C. M. Teixeira und S. H. Zak. Stabilizing Controller Design for Uncertain Nonlinear Systems Using Fuzzy Models. *IEEE Transactions on Fuzzy Systems*, 7:133–142, 1999.
61. M. A. L. Thathachar und P. Viswanath. On the Stability of Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*, 5:145–151, 1997.
62. H. Tolle. *Mehrgrößen-Regelkreissynthese, Band I: Grundlagen und Frequenzbereichsverfahren*. Oldenbourg-Verlag, München, 1983.
63. H. Unbehauen. *Regelungstechnik I, Lineare Kontinuierliche Systeme*. Vieweg-Verlag, Braunschweig, 1992.
64. H. Unbehauen. *Regelungstechnik II, Zustandsregelungen, digitale und nicht-lineare Systeme*. Vieweg-Verlag, Braunschweig, 1993.
65. H. Unbehauen. *Regelungstechnik III, Identifikation, Adaption, Optimierung*. Vieweg-Verlag, Braunschweig, 1993.
66. R. Unbehauen. *Systemtheorie*. Oldenbourg-Verlag, München, 1971.
67. A. Weinmann. *Regelungen - Analyse und technischer Entwurf, Band 2*. Springer-Verlag, Wien, 1987.
68. D. C. Youla, H. A. Jabr und J. J. Bongiorno, Jr. Modern Wiener-Hopf Design of Optimal Controllers - Part I: The Single Input-Output Case. *IEEE Transactions on Automatic Control*, 21:3–13, 1976.

69. D. C. Youla, H. A. Jabr und J. J. Bongiorno, Jr. *Modern Wiener-Hopf Design of Optimal Controllers - Part II: The Multivariable Case. IEEE Transactions on Automatic Control*, 21:319–338, 1976.
70. L. A. Zadeh und C. A. Desoer. *Linear System Theory: The State Space Approach*. McGraw Hill, New York, 1963.
71. L. A. Zadeh. Fuzzy Sets. *Information and Control*, 8:338–353, 1965.
72. K. Zhou und P. P. Khargonekar. Stability Robustness Bounds for Linear State-Space Models with Structured Uncertainty. *IEEE Transactions on Automatic Control*, 32:621–623, 1987.

Sachverzeichnis

- Allpass 84
 - erster Ordnung 85
- Anstiegsfunktion 34
- Aperiodisches Einschwingen 19, 94
- Arbeitspunkt 22
- Ausgangsmatrix 146

- Bang Bang Control 23
- Basistransformation 153
- Begrenzung 13
- Beobachtbarkeit 162
- Beobachter
 - linear 174
- Beobachtungsnormalform 149
- Blockschaltbild 5
- Bode-Diagramm 44

- Dämpfung 18, 84
- Differenzierer
 - Bode-Diagramm 52
- Dividierer 10
- Durchgangsmatrix 146
- Durchtrittsfrequenz 109
- Durchtrittswinkel 109

- Eigenkreisfrequenz 19
- Eingangsmatrix 146
- Entkopplung 138
- Ersatzzeitkonstante 40

- Faltung 33
- Feder-Masse-System 5, 147
- Frequenzbereich 28
- Frequenzgang 43
- Führungs-Übertragungsfunktion 73
- Führungsgröße 2
- Führungsgrößengenerator 135

- Führungsverhalten 4

- Genauigkeit 4, 91
 - stationär 92
- Gütemaß 94

- Hurwitz-Kriterium 84

- Impulsantwort 33
- Impulsfunktion 32
- Instabilität 4
- Integrator 7, 30
 - Bode-Diagramm 50
 - Ortskurve 65
- Integrierzeit 98
- Invertiertes Pendel 23
- IT_1 -Glied
 - Ortskurve 66

- Jordansche Normalform 151, 155

- Kalman-Filter 176
- Kaskadenregelung 131
- Kaskadenschaltung 131
- Kennlinienglied 11
- Kompensationsregler 121
- Koordinatentransformation 153
- Kreisfrequenz
 - natürlich 19, 55
- Kreisübertragungsfunktion 74
- Kreisverstärkung 74

- Laplace-Transformation 27
 - Differentiationssatz 29
 - Faltungssatz 29

- Grenzwertsatz 29
- Integrationssatz 29
- Korrespondenztabelle 180
- Überlagerungssatz 28
- Verschiebungssatz 29
- Laufzeit 10, 30
- Bode-Diagramm 53
- Ortskurve 65
- Linearisierung 22
- Luenberger-Beobachter 176

- Matrix
 - diagonalähnlich 156
- Matrizen-Exponentialfunktion 181
- Messglied 3
- MIMO-System 141
- Minimalphasenglied 90
- Mittenfrequenz 109
- Modalkoordinaten 156
- Multiplizierer 10

- Normierung 6
- Nutzfrequenzbereich 92
- Nyquist-Kriterium 77

- Optimaler Zustandsregler 173
- Ortskurve 61

- Parameter
 - konzentriert 21
- Pendel 12, 23
 - invertiert 23
- Phasenkorrekturglied 115
- Phasenreserve 109
- Proportionalglied 7, 31
- PT_1 -Glied 14
 - Bode-Diagramm 46
 - Ortskurve 62
 - Übertragungsfunktion 31
- PT_2 -Glied 15
 - Bode-Diagramm 54
 - Ortskurve 63
 - Übertragungsfunktion 31

- Rampenfunktion 34
- Regelabweichung 2
- Regelfehler 2
- Regelfläche 41
- Regelgeschwindigkeit 94
- Regelgröße 2
- Regelung 2
- Regelungsnormalform 148, 157
- Regler
 - I-Regler 97
 - P-Regler 95
 - PD-Regler 115
 - PI-Regler 101
 - PID-Regler 112
- Rekonstruierbarkeit 165
- Resonanzüberhöhung 55
- Riccati-Gleichung 174
- Riccati-Regler 174
- Routh-Kriterium 84
- Rückführung, ergänzend 129

- Separationstheorem 177
- Smith-Prädiktor 124
- Sollgröße 1
- Sollwert 1
- Sprungantwort 8, 35
 - Endwert 39
- Stabilisierbarkeit 172
- Stabilität 4
 - BIBO-stabil 71
 - linearer SISO-Systeme 69
 - Ljapunov 165
 - Regelkreis 73
 - Sprungantwort 69
 - Übertragungsfunktion 72
- Stabilitätsgrenzfall 4
- Stellglied 2
- Stellgröße 1
- Steuerbarkeit 162
 - Ausgangssteuerbarkeit 165
 - Zustandssteuerbarkeit 165
- Steuerbarkeitsmatrix 158
- Steuerung 2
- Stör-Übertragungsfunktion 74
- Störgröße 1
- Störgrößenaufschaltung 127
- Störverhalten 4
- Strecke 1
- Streckenmodell 5
- Summierer 7
- Symmetrisches Optimum 111
- System
 - kontinuierlich 21
 - linear 10
 - nicht-minimale Phase 84
 - zeitdiskret 21

– zeitinvariant 21
Systemmatrix 146

Transfermatrix 161
Transitionsmatrix 160

Übertragungsfunktion 27
– rational 31
– Vereinfachung 39
Übertragungsglied 7
– linear 10
– nichtlinear 10

Verkopplung 138
Verstärkungsfaktor 96
Verzögerungsglied
– erster Ordnung 14
– zweiter Ordnung 15
Vorfilter 126
Vorsteuerung

– dynamisch 135
– statisch 133

Wurzelortsverfahren 119

Youla-Parametrierung 123

Zeitbereich 28
Ziegler-Nichols Einstellregeln 120
Zustandsdarstellung 143
Zustandsgrößen 142
Zustandskurve 145
Zustandsraum 145
Zustandsregelung 162
Zustandsregler
– Entwurf 169
– Polvorgabe 171
– Stationäre Genauigkeit 177
Zustandsvariablen 142
Zweipunktregler 23