

Control Systems I



Institut für Regelungs- und
Automatisierungstechnik



(c) Institut für Regelungs- und Automatisierungstechnik

Technische Universität Graz

Hinweis: Dieses Skriptum dient als Lernbehelf und ersetzt den Besuch der Lehrveranstaltung *nicht!* Prüfungsrelevant sind ausschließlich die in der Lehrveranstaltung behandelten Themen.

Version vom 22. 6. 2017

Inhaltsverzeichnis

I	Systemtheorie	7
1	Grundlagen und Begriffe	9
1.1	Einführung	9
1.2	Grundlegende Begriffe	9
1.3	Linearität	14
1.4	Beispiele für dynamische Systeme	17
1.4.1	Elektrische Systeme	17
1.4.2	Mechanische Systeme	18
1.4.3	Räuber-Beute Modell	20
1.4.4	Epidemiemodell („SIR-Modell“)	21
2	Zeitkontinuierliche Systeme	25
2.1	Einführung	25
2.2	Lösung der Systemgleichungen	25
2.2.1	Freie Lösung	26
2.2.2	Erzwungene Lösung	27
2.2.3	Übertragungsfunktion	28
2.3	Ruhelagen	28
2.4	Linearisierung um eine Ruhelage	30
2.5	Asymptotische Stabilität	34
2.6	Übertragungsfunktion	36
2.7	Deutung von $G(s)$ im Zeitbereich	38
2.8	Pole und Nullstellen	38
2.8.1	PN-Plan	41
2.9	BIBO-Stabilität	41
2.10	Das Routh-Schema	43
2.11	Zusammenschaltung von Übertragungssystemen	47
2.11.1	Reihenschaltung	47
2.11.2	Parallelschaltung	48
2.11.3	Rückkopplung	49
2.11.4	Allgemeine Strukturen	50
2.12	Wichtige Übertragungsglieder	52
2.12.1	Proportionalglied	53
2.12.2	Verzögerungsglied erster Ordnung	53

2.12.3	Verzögerungsglied zweiter Ordnung	54
2.12.4	Integrator	55
2.12.5	Differenzierer	55
2.12.6	Vorhaltglied	55
2.13	Frequenzgang	56
3	Zeitdiskrete Systeme	59
3.1	Einführung	59
3.2	z -Übertragungsfunktion	59
3.3	Pole und Nullstellen	61
3.4	BIBO-Stabilität	62
II	Regelungstechnik	65
3.5	Einführung	67
3.6	Begriffe der Regelungstechnik	67
3.7	Beispiele	70
3.7.1	Drehzahlregelung mit Hilfe des Fliehkraftreglers	70
3.7.2	Der Pupillenapparat des Menschen	70
4	Standardregler	73
4.1	Einführung	73
4.2	Parallelrealisierung - die Lehrbuchform	73
4.3	Windup-Effekt und Gegenmaßnahmen	76
4.3.1	Anti-Windup-Maßnahmen	77
4.4	Einstellregeln für PID-Regler	78
4.4.1	Einstellregeln nach Ziegler-Nichols	79
4.4.2	Methode der Summenzeitkonstante („T-Summen Regel“)	81
5	Diskretisierung zeitkontinuierliche Regelgesetze	83
5.1	Einführung	83
5.2	Vorwärts-Euler-Integration	84
5.3	Rückwärts-Euler-Integration	85
5.4	Integration mittels Trapez-Regel - Tustin Formel	86
6	Analytische Reglersynthese	89
6.1	Einführung	89
6.2	Implementierbarkeit	89
6.3	Entwurf für den Standardregelkreis	90
6.3.1	Direkte Reglerberechnung	90
6.3.2	Polvorgabe	92
6.4	Entwurf für eine erweiterte Regelkreisstruktur	94
6.5	Wahl von $T(s)$	98

7	Frequenzkennlinien und Ortskurven	101
7.1	Einführung	101
7.2	Frequenzkennlinien	102
7.2.1	Normierte Darstellung	102
7.2.2	Zerlegung von $G(s)$ in „elementare Bestandteile“	103
7.3	Frequenzgangs-Ortskurve	110
8	Das Nyquist-Kriterium	113
8.1	Einführung	113
8.2	Stetige Winkeländerung einer Ortskurve	114
8.3	Formulierung des Nyquist-Kriteriums	116
8.4	Vereinfachtes Schnittpunktkriterium	119
8.4.1	Durchtrittsfrequenz	119
8.4.2	Phasenreserve und Amplitudenrand	120
8.4.3	Formulierung des vereinfachten Schnittpunktkriteriums	120
9	Das Frequenzkennlinien-Verfahren	123
9.1	Einführung	123
9.2	Allgemeine Überlegungen zum Reglerentwurf	124
9.2.1	Einschränkungen beim Entwurf	126
9.3	Ein typisches Entwurfsverfahren	127
9.3.1	Systeme mit dominantem Polpaar	127
9.3.2	Spezifikation des Wunschverhaltens	127
9.3.3	Faustformeln für den Reglerentwurf	129
9.3.4	Durchführung des Reglerentwurfes	129
9.4	Korrekturglieder	129
10	Englische Fachbegriffe	135
A	Laplace-Transformation	139
A.1	Linearität	139
A.2	Dämpfungsregel	140
A.3	Verschiebungssatz	140
A.4	Faltungsregel	140
A.5	Differentiation im Zeitbereich	141
A.6	Differentiation im Bildbereich	141
A.7	Integration im Zeitbereich	141
A.8	Anfangswertsatz	141
A.9	Endwertsatz	142
A.10	Korrespondenztabelle für reelle Funktionen	143
B	z-Transformation	145
B.1	Linearität	145
B.2	Dämpfungsregel	146
B.3	Verschiebung nach rechts	146

B.4	Verschiebung nach links	146
B.5	Faltungsregel	146
B.6	Differentiation im Bildbereich	147
B.7	Anfangswertsatz	147
B.8	Endwertsatz	147
B.9	Korrespondenztabelle	148

Teil I

Systemtheorie

Kapitel 1

Grundlagen und Begriffe

1.1 Einführung

Die Vorlesung „Control Systems 1“ verfolgt das Ziel, Studierenden die Grundlagen der Systemtheorie und der Regelungstechnik zu vermitteln. Eine zentrale Rolle spielt dabei der Systembegriff, der in zahlreichen Fachdisziplinen von elementarer Bedeutung ist. Ganz allgemein versteht man unter einem System die Verbindung von interagierenden Komponenten zu einer zweckgebundenen Einheit. Im Bereich der Elektrotechnik kann dies beispielweise eine aus elektronischen Bauteilen bestehende elektrische Schaltung sein. Ein Beispiel aus der Medizin ist das kardiovaskuläre System, das aus Herz und Blutgefäßen besteht und für die Aufrechterhaltung des Blutkreislaufs verantwortlich ist.

Mit Hilfe geeigneter mathematischer Modelle kann das Verhalten von Systemen nachgebildet werden. Solche Modelle können z.B. bei technischen Systemen durch Anwendung physikalischer Gesetzmäßigkeiten, wie etwa den Newtonschen Axiomen hergeleitet oder aus experimentell gewonnenen Daten abgeleitet werden. Die resultierende mathematische Beschreibung ermöglicht die Analyse und Simulation der betrachteten Systeme und bildet die Grundlage für zahlreiche Methoden der System- und Regelungstheorie.

1.2 Grundlegende Begriffe

Die Wechselwirkung eines Systems mit seiner Umgebung erfolgt mittels seiner so genannten **Eingangs- und Ausgangsgrößen**, siehe hierzu Bild 1.1. Die Eingangsgrößen u_1, \dots, u_p wirken von der Umgebung auf das System ein und beeinflussen somit dessen Verhalten. Je nachdem, ob diese Beeinflussung gewollt oder ungewollt ist, spricht man von **Stell- oder Störgrößen**. Die Ausgangsgrößen y_1, \dots, y_q wirken vom System auf die Umgebung ein und beeinflussen so diese. In vielen Fällen besitzen Systeme *eine* Eingangsgröße u und *eine* Ausgangsgröße y , man nennt sie **Eingrößensysteme**¹.

¹In Anlehnung an die englische Bezeichnung „single input - single output system“ spricht man auch von **SISO-Systemen**.

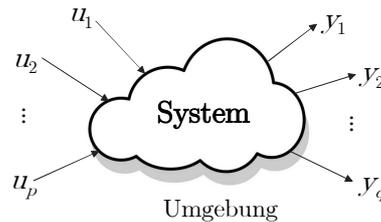


Bild 1.1: Interaktion zwischen einem System und seiner Umgebung

Im vorliegenden Skriptum wird das zeitliche Verhalten von Systemen untersucht. Das bedeutet, dass die Eingangs- und Ausgangsgrößen Funktionen des Zeitparameters t sind.

Man nennt ein System **kausal**, wenn die Werte der Ausgangsgrößen zu einem beliebigen Zeitpunkt t_1 unabhängig von den zukünftigen Verläufen der Eingangsgrößen sind. Das bedeutet, dass die Werte $y_1(t_1), \dots, y_q(t_1)$ ausschließlich von den Verläufen $u_1(t), \dots, u_p(t)$ für $t \leq t_1$ abhängen. Bei realen technischen Systemen trifft dies zu, d.h. sie sind kausal.

Systeme, bei denen die Werte $y_1(t_1), \dots, y_q(t_1)$ ausschließlich von $u_1(t_1), \dots, u_p(t_1)$, also von den Momentanwerten der Eingangsgrößen abhängen, werden **statische Systeme** genannt.

B Gegeben sei das mathematische Modell

$$y(t) = K u(t)$$

eines System mit der Eingangsgröße u und der Ausgangsgröße y . Der Parameter K sei eine (reelle) Konstante. Offensichtlich handelt es sich hier um ein statisches System. ■

Im Gegensatz dazu sind kausale **dynamische Systeme** dadurch charakterisiert, dass die Werte der Ausgangsgrößen zum Zeitpunkt t_1 nicht nur von den Momentanwerten der Eingangsgrößen abhängen, sondern auch von deren Verläufen in der Vergangenheit, also für $t < t_1$.

B Gegeben sei das mathematische Modell

$$\frac{dy(t)}{dt} = u(t)$$

eines zeitkontinuierlichen Systems mit der Eingangsgröße u und der Ausgangsgröße y . Zur Ermittlung des Wertes der Ausgangsgröße zu einem beliebigen Zeitpunkt t_1 ist die Kenntnis des Wertes der Ausgangsgröße zu einem Anfangszeitpunkt t_0 sowie der Verlauf der Eingangsgröße im Intervall $t_0 \leq t \leq t_1$ erforderlich, was unmittelbar aus

$$y(t_1) = y(t_0) + \int_{t_0}^{t_1} u(t) dt$$

ersichtlich ist. Der Anfangswert $y(t_0)$ repräsentiert gewissermaßen die gesamte „Vorgeschichte“ des Systems für $t < t_0$.

So genannte **Zustandsmodelle** spielen in der System- und Regelungstechnik eine besondere Rolle. Sie stellen eine einheitliche, von der Natur des Systems unabhängige, mathematische Beschreibung dar und bilden die Grundlage vieler regelungstechnischer Methoden. Für das Verständnis dieser Beschreibungsform ist der Begriff der so genannten **Zustandsgrößen** essentiell. ■

Wenn es möglich ist, für ein dynamisches System die Zeitfunktionen $x_1(t), \dots, x_n(t)$ so anzugeben, dass die Werte der Ausgangsgrößen $y_1(t), \dots, y_q(t)$ zu einem beliebigen Zeitpunkt t_1 aus den Verläufen der Eingangsgrößen $u_1(t), \dots, u_p(t)$ im Intervall $t_0 \leq t \leq t_1$ und den (konstanten) Werten $x_1(t_0), \dots, x_n(t_0)$ berechnet werden können, so bezeichnet man x_1, \dots, x_n als Zustandsgrößen des Systems.

Die natürliche Zahl n , d.h. die Anzahl der Zustandsgrößen wird auch **Systemordnung** genannt. Man beachte, dass die Wahl der Zustandsgrößen für ein gegebenes System *nicht* eindeutig ist. Vielmehr gibt es für ein und dasselbe System *unendlich viele* Möglichkeiten, die benötigten n Zustandsgrößen zu wählen. Diese Freiheiten bei der Festlegung der Zustandsgrößen werden bei zahlreichen Verfahren der System- und Regelungstechnik vorteilhaft ausgenützt.

B Wählt man im vorangegangenen Beispiel exemplarisch

$$x := 3y$$

so gilt für die Systembeschreibung

$$\frac{dx}{dt} = 3u, \quad y = \frac{1}{3}x.$$

Der Verlauf der Ausgangsgröße $y(t)$ kann somit über die Beziehung

$$y(t) = \frac{1}{3}x(t_0) + \int_{t_0}^t u(\tau) d\tau.$$

ermittelt werden. Offensichtlich kann $y(t)$ aus dem Verlauf von $u(\tau)$ im Intervall $t_0 \leq \tau \leq t$ und dem Wert $x(t_0)$ eindeutig berechnet werden, d.h. x ist eine Zustandsgröße des Systems. Auf analoge Weise kann gezeigt werden, dass sich im vorliegenden Beispiel jede beliebige Wahl $x = \alpha y$ mit $\alpha \neq 0$ als Zustandsgröße qualifiziert. ■

In weiterer Folge wird vorausgesetzt, dass die betrachteten Systeme eine *endliche* Ordnung n besitzen und durch gewöhnliche **Differentialgleichungen** der Form

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, \dots, x_n, u_1, \dots, u_p) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(x_1, \dots, x_n, u_1, \dots, u_p) \end{aligned} \tag{1.1}$$

sowie q algebraische Ausgangsgleichungen

$$\begin{aligned} y_1 &= g_1(x_1, \dots, x_n, u_1, \dots, u_p) \\ &\vdots \\ y_q &= g_q(x_1, \dots, x_n, u_1, \dots, u_p) \end{aligned} \quad (1.2)$$

beschrieben werden können. Man sagt, dass durch (1.1) und (1.2) ein **Zustandsmodell** des betrachteten Systems gegeben ist. Der Zustand des Systems zum Anfangszeitpunkt t_0 , d.h. die „Vorgeschichte“ des Systems, wird durch die Anfangswerte $x_1(t_0), x_2(t_0), \dots, x_n(t_0)$ repräsentiert.

Fasst man die Zustands-, Eingangs- und Ausgangsgrößen zu Vektoren

$$\mathbf{x} := \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{u} := \begin{bmatrix} u_1(t) \\ \vdots \\ u_p(t) \end{bmatrix} \quad \text{und} \quad \mathbf{y} := \begin{bmatrix} y_1(t) \\ \vdots \\ y_q(t) \end{bmatrix}$$

zusammen, sowie die (skalarwertigen) Funktionen f_1, \dots, f_n bzw. g_1, \dots, g_q zu den Vektoren

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) := \begin{bmatrix} f_1(\mathbf{x}, \mathbf{u}) \\ \vdots \\ f_n(\mathbf{x}, \mathbf{u}) \end{bmatrix}, \quad \mathbf{g}(\mathbf{x}, \mathbf{u}) := \begin{bmatrix} g_1(\mathbf{x}, \mathbf{u}) \\ \vdots \\ g_q(\mathbf{x}, \mathbf{u}) \end{bmatrix},$$

so erhält man die Systembeschreibung (1.1) und (1.2) in kompakter Matrixschreibweise

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad (1.3)$$

$$\mathbf{y} = \mathbf{g}(\mathbf{x}, \mathbf{u}), \quad (1.4)$$

wobei für den Anfangszustand

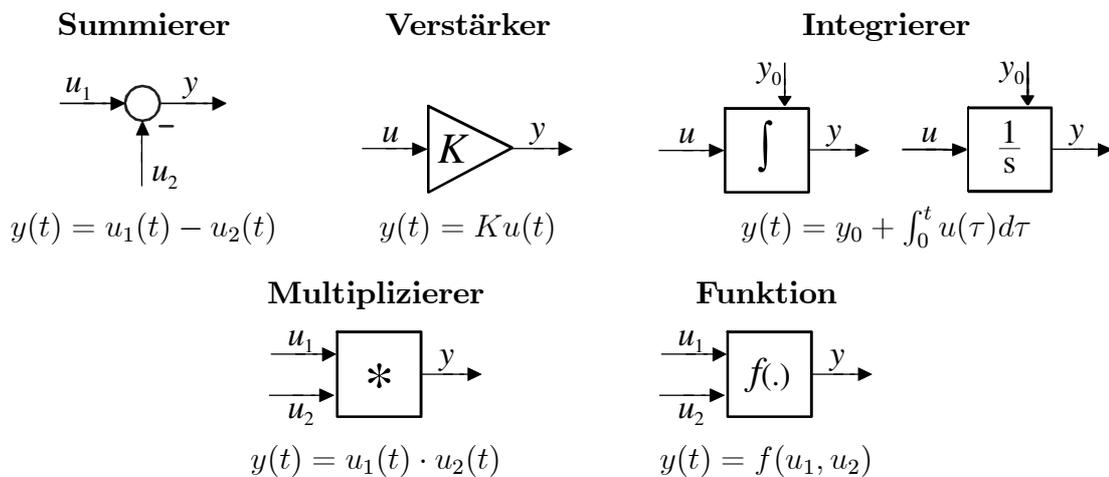
$$\mathbf{x}_0 := \mathbf{x}(t_0) = \begin{bmatrix} x_1(t_0) \\ \vdots \\ x_n(t_0) \end{bmatrix} \quad (1.5)$$

gilt.

Man beachte, dass die rechte Seite der Differentialgleichung (1.3) sowie die Ausgangsgleichung (1.4) nicht explizit vom Zeitparameter t abhängen. Solche Systeme nennt man **zeitinvariant**. Die Dynamik des Systems, d.h. seine Reaktion auf Eingangsgrößen und Anfangszustände ist somit *unabhängig* vom Zeitpunkt der Durchführung eines Experimentes. Das bedeutet, dass bei gleichzeitiger zeitlicher Verschiebung der Verläufe der Eingangsgrößen sowie des Anfangszustands auch die Ausgangsgröße die gleiche zeitliche Verschiebung erfährt. Der Anfangszeitpunkt kann somit ohne Einschränkung der Allgemeinheit zu $t_0 = 0$ gesetzt werden. Hängt die rechte Seite der Differentialgleichungen und/oder die Ausgangsgleichung in (1.3)

hingegen explizit vom Zeitparameter t ab, so spricht man von einem **zeitvarianten** bzw. **zeitvariablen System**.

Häufig werden so genannte **Strukturbilder** zur Visualisierung von mathematischen Modellen eingesetzt. Dabei werden mathematische Operationen, wie z.B. Integration, Multiplikation und Addition durch entsprechende Blöcke dargestellt, die gemäß der Modellbeschreibung kombiniert werden. Da auch Simulationswerkzeuge, wie z.B. Simulink² auf einer solchen „block-orientierten“ Eingabe basieren, stellen Strukturbilder oft die Grundlage für eine numerische Simulation des Systemverhaltens dar. Exemplarisch werden an dieser Stelle einige häufig verwendete Funktionsblöcke angegeben:



Aus (1.3) folgt, dass der Zustandsvektor $\mathbf{x}(t)$ prinzipiell durch Integration der rechten Seite der Differentialgleichung (1.3) ermittelt werden kann, d.h.

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau. \quad (1.6)$$

Hierbei ist die Integration elementweise auf $\mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau))$ anzuwenden. Aus (1.6) kann auch der wichtige Schluss gezogen werden, dass die Zustandsvariablen³ *stetige* Funktionen der Zeit sind. In Bild 1.2 ist das zu (1.3) gehörige Strukturbild dargestellt, wobei vektorielle Größen überlicherweise durch fett gezeichnete Verbindungslinien dargestellt werden.

Die Lösungen $\mathbf{x}(t)$ der Systemgleichungen (1.3) hängen offensichtlich vom Anfangszustand \mathbf{x}_0 und vom Verlauf der Eingangsgrößen $\mathbf{u}(\tau)$ im Intervall $t_0 \leq \tau \leq t$ ab. Diese Abhängigkeit wird im Folgenden durch

$$\mathbf{x}(t) = \Gamma \left(\begin{array}{c} \mathbf{x}_0 \\ \mathbf{u}(\tau), t_0 \leq \tau \leq t \end{array} \right) \quad \text{oder kurz} \quad \mathbf{x} = \Gamma \left(\begin{array}{c} \mathbf{x}_0 \\ \mathbf{u} \end{array} \right) \quad (1.7)$$

symbolisiert. Systeme, deren rechte Seite ausschließlich vom Zustandsvektor \mathbf{x} abhängt, werden **autonom** genannt. Nach einer anfänglichen Anregung durch den Anfangszustand \mathbf{x}_0

²www.mathworks.de

³hierbei werden (praktisch nicht realisierbare) Eingangsgrößen, die so genannte Dirac-Impulse enthalten, explizit ausgeschlossen.

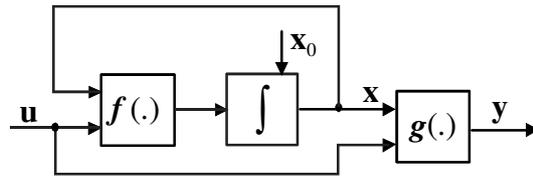


Bild 1.2: Strukturbild zu System (1.3), (1.4)

verläuft die „Bewegung“ eines autonomen Systems ohne weitere Beeinflussung von außen. Das System ist „sich selbst überlassen“. Die Beschreibung von autonomen Systemen ergibt sich unter der Annahme von $\mathbf{u}(t) = \mathbf{0}$ aus (1.3) zu

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{0}), \quad \text{d.h.} \quad \mathbf{x} = \Gamma \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{pmatrix} \quad (1.8)$$

Die Zustandsvariablen können als Koordinaten eines n -dimensionalen Koordinatensystems, dem so genannten **Zustandsraum** interpretiert werden. Im Falle $n = 2$ spricht man von der **Zustandsebene**. Die Kurve, die eine Lösung $\mathbf{x}(t)$ im Zustandsraum beschreibt, wenn der Zeitparameter t variiert, wird **Bahnkurve** oder **Trajektorie** genannt. Der Richtungssinn von Trajektorien für wachsende Zeiten t wird bei der graphischen Darstellung durch Pfeile gekennzeichnet. In Bild 1.3 ist exemplarisch eine Trajektorie eines Systems dritter Ordnung im Zustandsraum dargestellt, der Anfangszustand ist durch \mathbf{x}_0 gekennzeichnet.

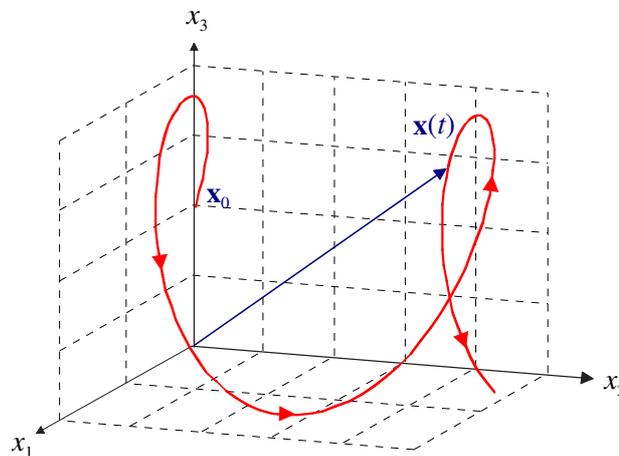


Bild 1.3: Trajektorie im dreidimensionalen Zustandsraum

1.3 Linearität

Ein System (1.3) heißt **linear**, wenn es bezüglich seiner Anfangszustände \mathbf{x}_0 und seiner Eingangsgrößen \mathbf{u} dem **Superpositionsprinzip** genügt. Das bedeutet, dass die zu beliebigen

Anfangszuständen und Eingangsgrößen gehörigen Lösungen (1.7) für beliebige Konstanten α und β den folgenden Bedingungen genügen:

$$\begin{aligned}
 \text{(i)} \quad & \Gamma \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{pmatrix} = \Gamma \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{pmatrix} + \Gamma \begin{pmatrix} \mathbf{0} \\ \mathbf{u} \end{pmatrix} \\
 \text{(ii)} \quad & \Gamma \begin{pmatrix} \alpha \mathbf{x}_{0,1} + \beta \mathbf{x}_{0,2} \\ \mathbf{0} \end{pmatrix} = \alpha \Gamma \begin{pmatrix} \mathbf{x}_{0,1} \\ \mathbf{0} \end{pmatrix} + \beta \Gamma \begin{pmatrix} \mathbf{x}_{0,2} \\ \mathbf{0} \end{pmatrix} \\
 \text{(iii)} \quad & \Gamma \begin{pmatrix} \mathbf{0} \\ \alpha \mathbf{u}_1 + \beta \mathbf{u}_2 \end{pmatrix} = \alpha \Gamma \begin{pmatrix} \mathbf{0} \\ \mathbf{u}_1 \end{pmatrix} + \beta \Gamma \begin{pmatrix} \mathbf{0} \\ \mathbf{u}_2 \end{pmatrix}
 \end{aligned} \tag{1.9}$$

Außerdem muss gegebenenfalls die Ausgangsgleichung $\mathbf{y} = \mathbf{g}(\mathbf{x}, \mathbf{u})$ entsprechende Eigenschaften besitzen. In Analogie zu (1.9) wird die Abhängigkeit der Ausgangsgrößen \mathbf{y} von \mathbf{x} und \mathbf{u} durch

$$\mathbf{y} = \Upsilon \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} \tag{1.10}$$

symbolisiert. Es muss dann gelten:

$$\begin{aligned}
 \text{(iv)} \quad & \Upsilon \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} = \Upsilon \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} + \Upsilon \begin{pmatrix} \mathbf{0} \\ \mathbf{u} \end{pmatrix} \\
 \text{(v)} \quad & \Upsilon \begin{pmatrix} \alpha \mathbf{x}_1 + \beta \mathbf{x}_2 \\ \mathbf{0} \end{pmatrix} = \alpha \Upsilon \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{pmatrix} + \beta \Upsilon \begin{pmatrix} \mathbf{x}_2 \\ \mathbf{0} \end{pmatrix} \\
 \text{(vi)} \quad & \Upsilon \begin{pmatrix} \mathbf{0} \\ \alpha \mathbf{u}_1 + \beta \mathbf{u}_2 \end{pmatrix} = \alpha \Upsilon \begin{pmatrix} \mathbf{0} \\ \mathbf{u}_1 \end{pmatrix} + \beta \Upsilon \begin{pmatrix} \mathbf{0} \\ \mathbf{u}_2 \end{pmatrix}
 \end{aligned} \tag{1.11}$$

Ein System, das den Bedingungen (1.9) und / oder (1.11) *nicht* genügt, nennt man **nichtlinear**.

B Gegeben sei das autonome System

$$\frac{dx}{dt} = -x, \quad x(0) =: x_0.$$

Es soll untersucht werden, ob das System linear ist. Da auf das System keine Eingangsgröße wirkt, muss lediglich Bedingung (1.9,ii) überprüft werden. Die Lösung der Differentialgleichung lautet

$$x(t) = e^{-t} x_0.$$

Daraus ist ersichtlich, dass die Linearitätsbedingung erfüllt ist. ■

B Gegeben sei das System

$$\begin{aligned}
 \frac{dx}{dt} &= x - u, & x(0) &=: x_0 \\
 y &= x^2
 \end{aligned}$$

Die Ausgangsgleichung erfüllt die Bedingung (1.11) nicht, das System ist somit nichtlinear. ■

Eine große Klasse von Systemmodellen, die den Bedingungen (1.9) und (1.11) genügen, besitzt die Form

$$\begin{aligned}\frac{d\mathbf{x}}{dt} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u},\end{aligned}\tag{1.12}$$

wobei \mathbf{A} , \mathbf{B} , \mathbf{C} und \mathbf{D} konstante Matrizen passender Dimensionen sind. Systeme der Form (1.12) repräsentieren eine wichtige Klasse von **linearen, zeitinvarianten Systemen**, die oft auch **LZI-Systeme** genannt werden. Für ein System der Ordnung n ist die **System- oder Dynamikmatrix** \mathbf{A} eine $n \times n$ Matrix⁴, die **Eingangsmatrix** \mathbf{B} ist eine $n \times p$ Matrix, wobei p die Zahl der Eingangsgrößen ist. Unter der Annahme von q Ausgangsgrößen ist die **Ausgangsmatrix** \mathbf{C} eine $q \times n$ Matrix und die **Durchgriffsmatrix** \mathbf{D} hat die Dimension $q \times p$. Im Falle eines Eingrößensystems, d.h. für $p = q = 1$ lautet die Systembeschreibung

$$\begin{aligned}\frac{dx}{dt} &= \mathbf{A}\mathbf{x} + \mathbf{b}u, \\ y &= \mathbf{c}^T\mathbf{x} + du.\end{aligned}\tag{1.13}$$

Der **Eingangsvektor** \mathbf{b} ist ein n -dimensionaler Spaltenvektor, der **Ausgangsvektor** \mathbf{c}^T ein n -dimensionaler Zeilenvektor und der **Durchgriff** d eine skalare Größe. Das zu (1.13) gehörige Strukturbild ist in Bild 1.4 dargestellt.

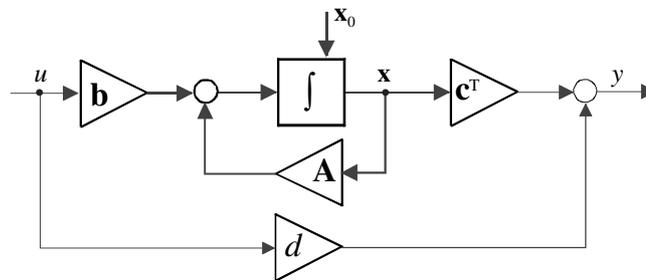


Bild 1.4: Strukturbild eines zeitkontinuierlichen Zustandsmodells der Form (1.13)

Die Gründe, warum diese Systemklasse in der System- und Regelungstechnik eine herausragende Rolle spielt, sind vielfältig:

- die Systeme sind aus mathematischer Sicht „gutmütig“, d.h. die Existenz und Eindeutigkeit der Lösungen sind gesichert
- viele reale Systeme können hinreichend genau durch Systeme dieser Form beschrieben werden

⁴oft wird auch die Schreibweise (n, n) -Matrix verwendet

- es existieren viele bewährte Verfahren, die für diese Systemklasse maßgeschneidert sind

Aus den oben genannten Gründen wird die Klasse der linearen, zeitinvarianten Systeme auch im Mittelpunkt aller weiteren Ausführungen in diesem Skriptum stehen. Wie im folgenden Abschnitt gezeigt wird, können nichtlineare Systeme durch lineare Systeme approximiert werden, was in weiterer Folge die Anwendung von „linearen“ Methoden erlaubt.

1.4 Beispiele für dynamische Systeme

1.4.1 Elektrische Systeme

Die Herleitung von Zustandsmodellen zur Beschreibung von idealisierten elektrischen Netzwerken, bestehend aus Widerständen, Kondensatoren und Spulen ist ziemlich geradlinig. Die Systemordnung n wird im Allgemeinen durch die Anzahl der „Energiespeicher“ bestimmt, also durch die Zahl der Kapazitäten und Induktivitäten. Es bietet sich an, als Zustandsvariablen die Spannungsabfälle an Kapazitäten und Ströme durch Induktivitäten zu wählen. Das Zustandsmodell ergibt sich dann aus der Anwendung der Kirchhoffschen⁵ Regeln.

B Gegeben sei das in Bild 1.5 dargestellte ideale elektrische Netzwerk, bestehend aus einer Spannungsquelle, zwei Ohmschen Widerständen R_1 und R_2 , der Kapazität C und der Induktivität L . Die Eingangsgröße sei die Spannung u , die Ausgangsgröße ist der Spannungsabfall am Widerstand R_2 .

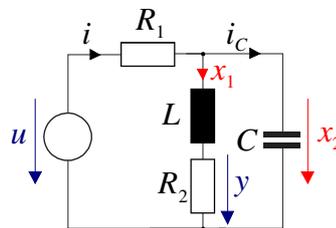


Bild 1.5: Elektrisches Netzwerk mit zwei Energiespeichern

Als Zustandsvariablen werden - wie oben erläutert - der Strom x_1 und die Spannung x_2 gewählt. Offensichtlich gilt dann

$$x_2 = L \frac{dx_1}{dt} + R_2 x_1,$$

d.h. die erste Zustandsdifferentialgleichung lautet

$$\frac{dx_1}{dt} = -\frac{R_2}{L} x_1 + \frac{1}{L} x_2.$$

Weiters gilt

$$u = R_1 i + x_2,$$

⁵benannt nach dem deutschen Physiker Gustav Robert Kirchhoff (1824 - 1887)

wobei sich der Strom i gemäß

$$i = i_C + x_1 = C \frac{dx_2}{dt} + x_1$$

errechnet. Daraus folgt

$$u = R_1 C \frac{dx_2}{dt} + R_1 x_1 + x_2,$$

und für die zweite Differentialgleichung gilt

$$\frac{dx_2}{dt} = -\frac{1}{C}x_1 - \frac{1}{R_1 C}x_2 + \frac{1}{R_1 C}u.$$

Die Ausgangsgleichung ist durch

$$y = R_2 x_1$$

gegeben. Zusammenfassend lautet somit das Zustandsmodell des elektrischen Systems

$$\begin{aligned} \frac{dx_1}{dt} &= -\frac{R_2}{L}x_1 + \frac{1}{L}x_2, \\ \frac{dx_2}{dt} &= -\frac{1}{C}x_1 - \frac{1}{R_1 C}x_2 + \frac{1}{R_1 C}u, \\ y &= R_2 x_1. \end{aligned}$$

■

Das vorangegangene Beispiel verdeutlicht, dass die Komponenten des Zustandsvektors \mathbf{x} verschiedene physikalische Dimensionen besitzen können. Im Beispiel ist x_1 ein elektrischer Strom (Einheit Ampère) während x_2 eine elektrische Spannung (Einheit Volt) darstellt.

1.4.2 Mechanische Systeme

Bei der translatorischen Bewegung von Massen ist es zweckmäßig, deren Positionen und Geschwindigkeiten als Zustandsvariablen einzuführen. Analog dazu können bei rotatorischen Bewegungen von Massen deren Drehwinkel und Winkelgeschwindigkeiten gewählt werden.

B Gegeben sei das in Bild 1.6, links dargestellte Masse-Feder System, bestehend aus einem Körper mit der Masse m und einer Feder mit linearer Federkennlinie, charakterisiert durch die Federkonstante c . Die Position der Masse wird mit y bezeichnet, wobei $y = 0$ der Lage bei entspannter Feder entspricht. Für die Reibung zwischen Körper und Untergrund wird Coulombsche⁶ Reibung angenommen, der Reibungskoeffizient wird mit μ bezeichnet. Weiters wirkt auf den Körper eine äußere Kraft F . Die Anwendung des zweiten Newtonschen⁷ Gesetzes liefert die Differentialgleichungen

$$\frac{dy}{dt} = v, \quad m \frac{dv}{dt} = -cy - \mu mg \operatorname{sign} v + F$$

⁶benannt nach dem französischen Physiker Charles Augustin de Coulomb (1736 - 1806)

⁷benannt nach dem englischen Forscher Isaac Newton (1642 - 1727)

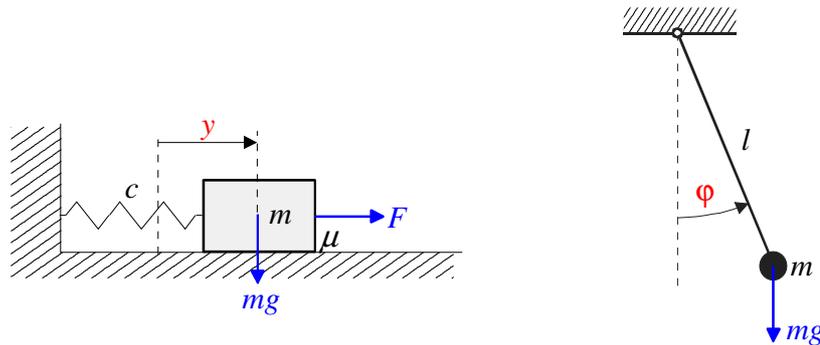


Bild 1.6: Mechanische Systeme: Masse-Feder System und Pendel

wobei v die Geschwindigkeit des Körpers ist und g die Erdbeschleunigung repräsentiert. Führt man nun die Zustandsvariablen

$$x_1 = y \quad \text{und} \quad x_2 = v$$

ein, so erhält man mit der Schreibweise $u := F$ die Systembeschreibung

$$\begin{aligned} \frac{dx_1}{dt} &= x_2, \\ \frac{dx_2}{dt} &= -\frac{c}{m} x_1 - \mu g \operatorname{sign} x_2 + \frac{1}{m} u, \\ y &= x_1. \end{aligned}$$

■

B Das mathematische Modell des in Bild 1.6, rechts dargestellten Pendels soll aufgestellt werden. Dabei wird idealisierend vorausgesetzt, dass das Pendel aus einem drehbar gelagerten, masselosen Stab der Länge l und einer punktförmigen Masse m am Stabende besteht. Das der Pendelbewegung entgegenwirkende Reibungsmoment (Lagerreibung) sei der Winkelgeschwindigkeit proportional (Proportionalitätsfaktor k). Mit Hilfe des Drallsatzes ergeben sich die Differentialgleichungen

$$\frac{d\varphi}{dt} = \omega, \quad ml^2 \frac{d\omega}{dt} = -mgl \sin \varphi - k\omega \quad \text{mit } k \geq 0,$$

wobei ω die Winkelgeschwindigkeit des Pendels ist. Führt man als Zustandsvariablen den Pendelwinkel und die Winkelgeschwindigkeit ein, d.h.

$$x_1 = \varphi, \quad x_2 = \omega,$$

so ergibt sich das folgende Zustandsmodell:

$$\begin{aligned} \frac{dx_1}{dt} &= x_2, \\ \frac{dx_2}{dt} &= -\frac{g}{l} \sin x_1 - \frac{k}{ml^2} x_2, \\ y &= x_1. \end{aligned}$$

Die zum Pendelmodell gehörigen Trajektorien sind in Bild 1.7 für $k = 0$, d.h. für den reibungsfreien Fall, dargestellt. Die ungedämpften Pendelbewegungen sind hier deutlich zu erkennen. In Bild 1.8 sind die entsprechenden Trajektorien für $k > 0$ dargestellt.

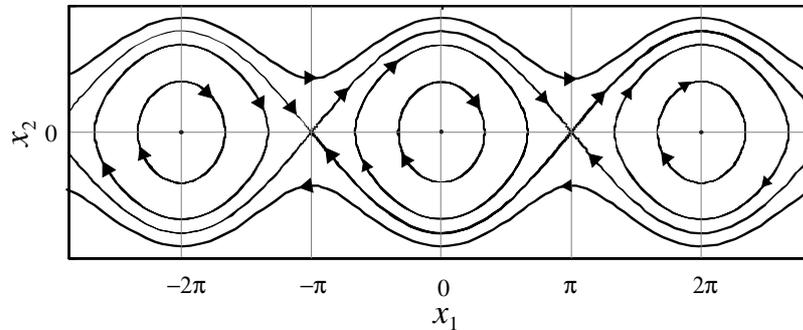


Bild 1.7: Trajektorien des Pendelmodells für $k = 0$

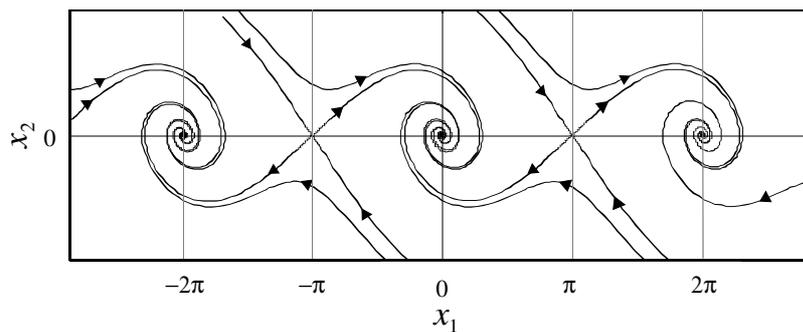


Bild 1.8: Trajektorien des Pendelmodells für $k > 0$

■

1.4.3 Räuber-Beute Modell

Nach Volterra und Lotka kann die Räuber-Beute Beziehung zweier (hinreichend großer) Populationen mit Hilfe der gekoppelten Differentialgleichungen

$$\begin{aligned}\frac{dx_1}{dt} &= ax_1 - bx_1x_2, \\ \frac{dx_2}{dt} &= -cx_2 + dx_1x_2\end{aligned}$$

mit den positiven Konstanten $a, b, c, d > 0$ beschrieben werden. Hierbei repräsentiert x_1 die „Anzahl“ der Beutetiere und x_2 stellt die „Anzahl“ der Räuber dar. Das Modell beruht auf der Annahme, dass sich ohne natürliche Feinde, d.h. für $x_2 = 0$, die Beutetiere exponentiell mit der Rate a vermehren. Umgekehrt nimmt die Population der Räuber für $x_1 = 0$ wegen

Nahrungsmangels exponentiell mit der Rate c ab. Die Wechselwirkungen zwischen den beiden Populationen werden durch die Terme bx_1x_2 und dx_1x_2 charakterisiert. In Bild 1.9 sind die zeitlichen Verläufe von x_1 und x_2 für $x_1(0) = 40$ und $x_2(0) = 20$ und die Modellparameter $a = 0.5$, $b = d = 0.01$, $c = 0.8$ dargestellt. Weiters ist für die gleichen Anfangswerte der Verlauf der Trajektorie in der Zustandsebene eingezeichnet. Man erkennt, dass sich eine geschlossene Trajektorie ergibt. Das bedeutet, dass x_1 und x_2 periodische Funktionen der Zeit sind.

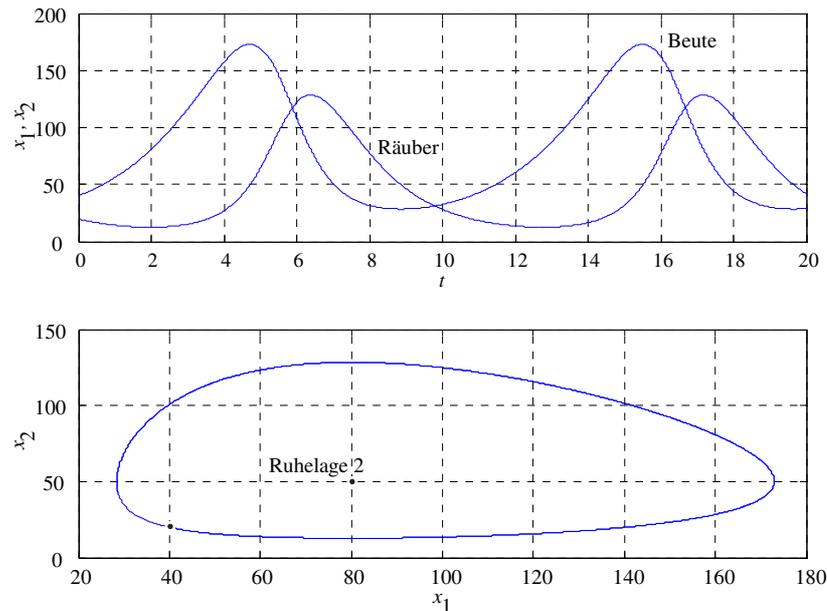


Bild 1.9: Räuber - Beute Beziehung

Weiters ist in Bild 1.9 eine so genannte Ruhelage des Systems eingezeichnet. Sie ist durch $x_{1,R}(t) = x_1(0) = \text{konst.}$ und $x_{2,R}(t) = x_2(0) = \text{konst.}$ charakterisiert. Das bedeutet, dass sich die beiden Populationen in einem Gleichgewicht befinden. Dieser spezielle Zustand ergibt sich aus den algebraischen Gleichungen

$$\frac{dx_{1,R}}{dt} = 0 = ax_{1,R} - bx_{1,R}x_{2,R} \quad \text{und} \quad \frac{dx_{2,R}}{dt} = 0 = -cx_{2,R} + dx_{1,R}x_{2,R}.$$

Die Ruhelage $x_{1,R} = x_{2,R} = 0$ stellt hier den „trivialen Fall“ dar, für die praktisch relevante Ruhelage ergibt sich

$$x_{1,R} = \frac{c}{d}, \quad x_{2,R} = \frac{a}{b}.$$

1.4.4 Epidemiemodell („SIR-Modell“)

Bei vielen klassischen Epidemiemodellen, wie auch bei dem hier vorgestellten Kermack-McKendrick Modell wird eine Gesamtpopulation in drei unterschiedliche Klassen eingeteilt. Die erste Klasse „S“ (=susceptibles) umfasst gesunde und infizierbare, also nicht-immune, Personen. Die zweite Klasse „I“ (=infectives) umfasst infektiöse Personen und die dritte Klasse

„R“ (=removals) wird durch Personen gebildet, die (nach einer Erkrankung) dauerhaft immun sind, wie es zum Beispiel bei Kinderkrankheiten wie Masern der Fall ist. Die „Anzahl“ der Individuen in „S“ wird mit x_1 bezeichnet, zu „I“ gehören x_2 und zu „R“ gehören x_3 Personen. Es wird vorausgesetzt, dass die Gesamtpopulation konstant ist, d.h.

$$x_1(t) + x_2(t) + x_3(t) = N = \text{konstant.}$$

Da Elemente der Klasse „S“ durch Ansteckung zu Elementen der Klasse „I“ werden und nach Gesundung schließlich zur Klasse „R“ gehören, spricht man auch vom SIR-Modell. Es besteht aus den drei gekoppelten gewöhnlichen Differentialgleichungen

$$\begin{aligned}\frac{dx_1}{dt} &= -ax_1x_2, \\ \frac{dx_2}{dt} &= ax_1x_2 - bx_2, \\ \frac{dx_3}{dt} &= bx_2.\end{aligned}$$

Die positive Konstante a wird Infektionsrate genannt, die positive Konstante b ist die Immunisierungsrate. Man beachte, dass sich im obigen mathematischen Modell die Tatsache widerspiegelt, dass die Gesamtpopulation konstant ist, denn es gilt

$$\frac{dx_1}{dt} + \frac{dx_2}{dt} + \frac{dx_3}{dt} = 0.$$

Das Strukturbild zum mathematischen Modell ist in Bild 1.10 dargestellt. Wenn ausgehend

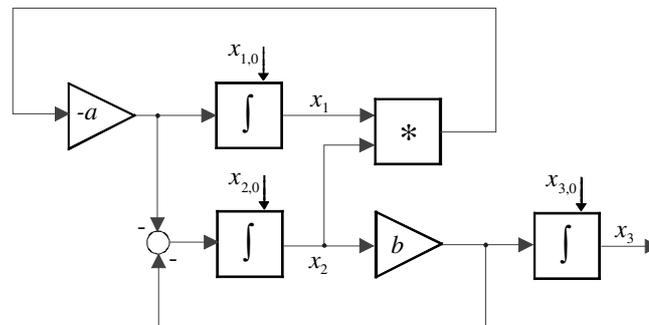


Bild 1.10: Strukturbild zum Epidemiemodell

von den Anfangswerten $x_1(0) + x_2(0) + x_3(0) = N$ die Zahl der Infizierten, also x_2 , zunimmt, d.h.

$$\left. \frac{dx_2}{dt} \right|_{t=0} > 0,$$

so spricht man vom Ausbruch einer Epidemie. Dies ist offensichtlich dann der Fall, wenn $x_1(0)$ größer als der Schwellwert

$$\rho := \frac{b}{a}$$

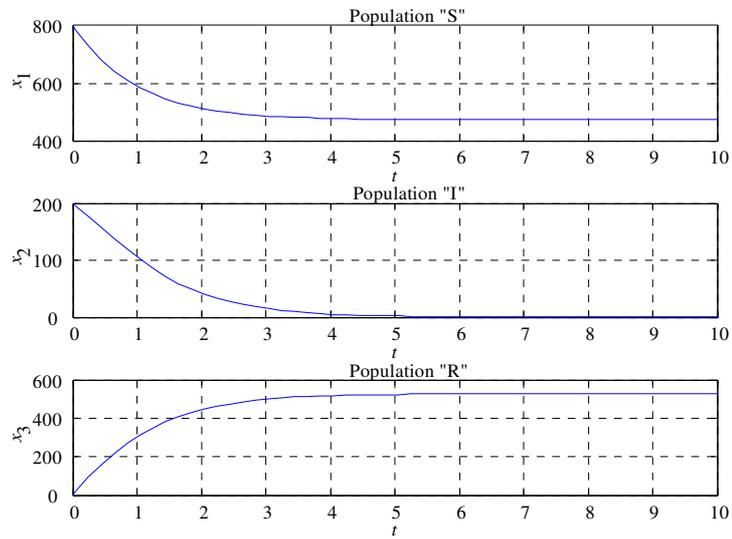


Bild 1.11: Krankheitsverlauf ohne Epidemie

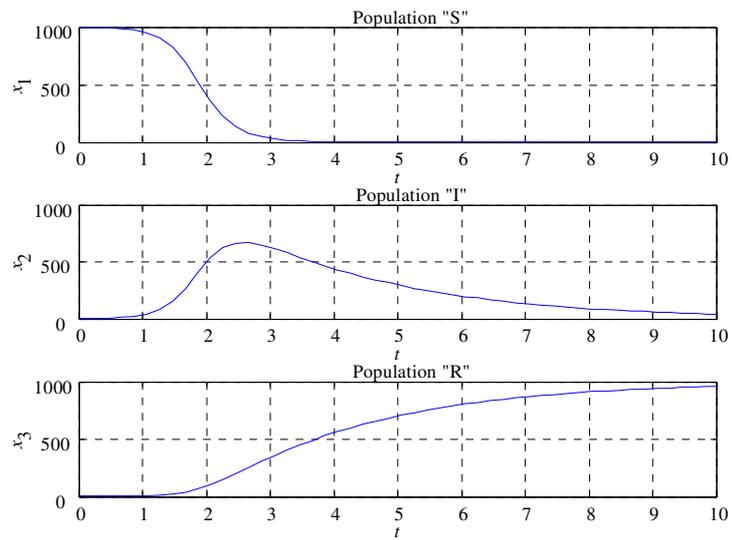


Bild 1.12: Ausbruch einer Epidemie

ist. In Bild 1.11 sind die Verläufe der Populationsgrößen dargestellt, wobei für die Gesamtpopulation $N = 1000$ gilt, die Infektionsrate a beträgt 0.002, die Immunisierungsrate b ist 2. Wie man leicht überprüfen kann, bricht für $x_1(0) = 800$ und $x_2(0) = 200$ keine Epidemie aus. Im Gegensatz dazu bricht für $N = 1000$, $a = 0.004$, $b = 0.4$ und $x_1(0) = 999$, $x_2(0) = 1$ eine Epidemie aus, was auch aus Bild 1.12 hervorgeht.

Kapitel 2

Zeitkontinuierliche Systeme

2.1 Einführung

In diesem Kapitel werden lineare, zeitinvariante Eingrößensysteme der Form

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{b}u \quad (2.1)$$

$$y = \mathbf{c}^T \mathbf{x} + d u \quad (2.2)$$

mit dem Anfangszustand

$$\mathbf{x}_0 := \mathbf{x}(t_0 = 0)$$

analysiert. Mit \mathbf{x} wird der n -dimensionale Zustandsvektor bezeichnet, u ist die Eingangsgröße und y die Ausgangsgröße des Systems. Die Zeitinvarianz von (2.1) impliziert, dass die Systemmatrix \mathbf{A} , der Eingangsvektor \mathbf{b} , der Ausgangsvektor \mathbf{c} und der Durchgriff d konstante Größen passender Dimensionen sind.

2.2 Lösung der Systemgleichungen

Da es sich bei den Zustandsgleichungen (2.1) um lineare Differentialgleichungen mit konstanten Koeffizienten handelt, kann die Laplace-Transformation eingesetzt werden. Wendet man (A.9) auf (2.1) an, so erhält man

$$s\bar{\mathbf{x}}(s) - \mathbf{x}_0 = \mathbf{A}\bar{\mathbf{x}}(s) + \mathbf{b}\bar{u}(s),$$

wobei

$$\bar{\mathbf{x}}(s) = \mathcal{L}\{\mathbf{x}(t)\} \quad \text{und} \quad \bar{u}(s) = \mathcal{L}\{u(t)\}.$$

Daraus ergibt sich unmittelbar

$$(s\mathbf{E} - \mathbf{A})\bar{\mathbf{x}}(s) = \mathbf{x}_0 + \mathbf{b}\bar{u}(s),$$

wobei \mathbf{E} die $n \times n$ -Einheitsmatrix repräsentiert. Nach einer Multiplikation mit $(s\mathbf{E} - \mathbf{A})^{-1}$ von links findet man

$$\bar{\mathbf{x}}(s) = (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{x}_0 + (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b}\bar{u}(s). \quad (2.3)$$

Definiert man die $n \times n$ - Matrix

$$\bar{\phi}(s) := (s\mathbf{E} - \mathbf{A})^{-1} \quad \text{d.h.} \quad \phi(t) = \mathcal{L}^{-1} \{ (s\mathbf{E} - \mathbf{A})^{-1} \}, \quad (2.4)$$

so gilt - nach Anwendung des Faltungssatzes (A.8) - für die Lösung

$$\mathbf{x}(t) = \phi(t) \mathbf{x}_0 + \int_0^t \phi(t - \tau) \mathbf{b} u(\tau) d\tau. \quad (2.5)$$

Der Lösungsvektor $\mathbf{x}(t)$ repräsentiert das zeitliche Verhalten des Systems (2.1) als Reaktion auf einen Anfangswert \mathbf{x}_0 und die Eingangsgröße $u(t)$. Das Systemverhalten (2.5) kann in zwei additive Anteile zerlegt werden. Der erste Anteil rührt von der „Vorgeschichte“ des Systems her, also vom Anfangszustand \mathbf{x}_0 , man spricht von der **freien Lösung** $\mathbf{x}_{frei}(t)$. Im Gegensatz dazu wird der zweite Anteil durch den Verlauf der Eingangsgröße $u(\tau)$ im Intervall $0 \leq \tau \leq t$ geprägt, man spricht von der **erzwungenen Lösung** $\mathbf{x}_{erzw}(t)$. Diese Erkenntnis kann man - unter Verwendung der in (1.7) eingeführten Notation - folgendermaßen zusammenfassen:

$$\mathbf{x} = \Gamma \begin{pmatrix} \mathbf{x}_0 \\ u \end{pmatrix} = \mathbf{x}_{frei} + \mathbf{x}_{erzw} = \Gamma \begin{pmatrix} \mathbf{x}_0 \\ 0 \end{pmatrix} + \Gamma \begin{pmatrix} \mathbf{0} \\ u \end{pmatrix} \quad (2.6)$$

Aufgrund der Struktur von (2.5) ist unmittelbar einzusehen, dass auch die restlichen Eigenschaften der Linearitätsbedingung (1.9) erfüllt sind.

Die Ausgangsgröße kann mittels Relation (2.2) berechnet werden, im Bildbereich gilt mit (2.3)

$$\bar{y}(s) = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{x}_0 + [\mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d] \bar{u}(s), \quad (2.7)$$

bzw. unter Verwendung von (2.5) im Zeitbereich

$$y(t) = \mathbf{c}^T \phi(t) \mathbf{x}_0 + \int_0^t \mathbf{c}^T \phi(t - \tau) \mathbf{b} u(\tau) d\tau + d u(t). \quad (2.8)$$

Analog zu (2.5) kann die Ausgangsgröße also in einen Anteil, der vom Anfangszustand herrührt und einen Anteil, der vom Verlauf der Eingangsgröße geprägt wird, zerlegt werden. Unter Verwendung der Notation (1.10) kann man schreiben

$$y = \Upsilon \begin{pmatrix} \mathbf{x}_0 \\ u \end{pmatrix} = \Upsilon \begin{pmatrix} \mathbf{x}_0 \\ 0 \end{pmatrix} + \Upsilon \begin{pmatrix} \mathbf{0} \\ u \end{pmatrix} \quad (2.9)$$

Wie man sich überzeugen kann, sind auch die weiteren Eigenschaften von (1.11) gegeben.

2.2.1 Freie Lösung

Unter der freien Lösung eines Systems versteht man diejenige Lösung $\mathbf{x}(t)$, die sich ergibt, wenn das System *ausschließlich* durch den Anfangszustand \mathbf{x}_0 angeregt wird. Somit ist hier das autonome System

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} \quad (2.10)$$

zu untersuchen. Gemäß (2.5) gilt für die Lösung dann

$$\mathbf{x}(t) = \phi(t) \mathbf{x}_0. \quad (2.11)$$

Die Matrix $\phi(t)$ wird **Transitionsmatrix**¹ genannt, sie beschreibt den Übergang des Zustandsvektors von seinem Anfangswert zu seinem Wert zum Zeitpunkt t . Ihre Berechnung kann mit Hilfe von Formel (2.4) erfolgen.

B Für das autonome System zweiter Ordnung

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \mathbf{x} \quad \text{mit} \quad \mathbf{x}_0 = \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}.$$

ergibt sich die Laplace-Transformierte der Transitionsmatrix gemäß (2.4) zu

$$\bar{\phi}(s) = (s\mathbf{E} - \mathbf{A})^{-1} = \frac{1}{s(s-3)} \begin{bmatrix} s-2 & 1 \\ 2 & s-1 \end{bmatrix}.$$

Eine Partialbruchzerlegung und anschließende Rücktransformation in den Zeitbereich ergibt

$$\phi(t) = \frac{1}{3} \mathcal{L}^{-1} \left\{ \begin{bmatrix} \frac{2}{s} + \frac{1}{s-3} & -\frac{1}{s} + \frac{1}{s-3} \\ -\frac{2}{s} + \frac{2}{s-3} & \frac{1}{s} + \frac{2}{s-3} \end{bmatrix} \right\} = \begin{bmatrix} \frac{2}{3} + \frac{1}{3}e^{3t} & -\frac{1}{3} + \frac{1}{3}e^{3t} \\ -\frac{2}{3} + \frac{2}{3}e^{3t} & \frac{1}{3} + \frac{2}{3}e^{3t} \end{bmatrix}.$$

Für den zeitlichen Verlauf des Zustandsvektors gilt somit

$$\mathbf{x}(t) = \phi(t) \mathbf{x}_0 = \begin{bmatrix} \left(\frac{2}{3} + \frac{1}{3}e^{3t} \right) x_{1,0} + \left(-\frac{1}{3} + \frac{1}{3}e^{3t} \right) x_{2,0} \\ \left(-\frac{2}{3} + \frac{2}{3}e^{3t} \right) x_{1,0} + \left(\frac{1}{3} + \frac{2}{3}e^{3t} \right) x_{2,0} \end{bmatrix}.$$

■

2.2.2 Erzwungene Lösung

Die erzwungene Lösung oder Bewegung eines Systems ist diejenige Lösung $\mathbf{x}(t)$, die sich ergibt, wenn das System bei *verschwindendem* Anfangszustand² $\mathbf{x}_0 = \mathbf{0}$ durch die Eingangsgröße $u(t)$ angeregt wird. Nach (2.5) gilt unter diesen Umständen

$$\mathbf{x}(t) = \int_0^t \phi(t-\tau) \mathbf{b} u(\tau) d\tau. \quad (2.12)$$

B Gegeben sei das mathematische Modell

$$\frac{dx}{dt} = -x + u$$

¹lat. *transire* = übergehen

²man sagt: "Das System befindet sich zum Zeitpunkt $t = 0$ in Ruhe"

eines linearen, zeitinvarianten Systems erster Ordnung mit der Eingangsgröße u . Mit Hilfe von (2.12) findet man mit $\phi(t) = e^{-t}$ für die erzwungene Lösung

$$x(t) = \int_0^t e^{-(t-\tau)} \cdot 1 \cdot u(\tau) d\tau = e^{-t} \int_0^t e^{\tau} u(\tau) d\tau.$$

Wählt man beispielsweise als Eingangsgröße einen Einheitssprung, d.h. $u(t) = \sigma(t)$, so gilt

$$x(t) = 1 - e^{-t}.$$

■

2.2.3 Übertragungsfunktion

Für die Ausgangsgröße folgt unter der Annahme $\mathbf{x}_0 = \mathbf{0}$ aus (2.7) im Bildbereich unmittelbar

$$\bar{y}(s) = [\mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d] \bar{u}(s). \quad (2.13)$$

Der Ausdruck

$$G(s) := \frac{\bar{y}(s)}{\bar{u}(s)} = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d \quad (2.14)$$

ist die so genannte **Übertragungsfunktion** des Systems. Sie beschreibt das Übertragungsverhalten eines linearen, zeitinvarianten Systems im Bildbereich.

B Gegeben sei das Zustandsmodell

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u, \quad y = [0 \quad 1] \mathbf{x} + 2u.$$

Für die Übertragungsfunktion $G(s)$ des Systems gilt dann

$$G(s) = \frac{\bar{y}(s)}{\bar{u}(s)} = [0 \quad 1] \begin{bmatrix} s-1 & -2 \\ -3 & s-4 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 = \frac{2s^2 - 8s - 3}{s^2 - 5s - 2}.$$

■

Weiterführende Informationen über die Übertragungsfunktion finden sich in Kapitel ??

2.3 Ruhelagen

Ruhelagen sind spezielle Lösungen der Systemgleichungen (2.1), die dadurch charakterisiert sind, dass für den Zustandsvektor gilt:

$$\mathbf{x}(t) = \mathbf{x}_R = \text{konst.} \quad (2.15)$$

Das bedeutet, dass sich das System in einem „Gleichgewichtszustand“ befindet, der auch **Ruhelage** genannt wird. In einer Ruhelage nehmen alle Systemgrößen, d.h. Eingangsgröße und Zustandsgrößen konstante Werte an, d.h. es gilt

$$\frac{d\mathbf{x}_R}{dt} = \mathbf{0}. \quad (2.16)$$

Da auch die Eingangsgröße als konstant angenommen wird, d.h.

$$u(t) = u_R = \text{konst.}$$

führt die Bedingung (2.16) auf das lineare Gleichungssystem

$$-\mathbf{A}\mathbf{x}_R = \mathbf{b} u_R \quad (2.17)$$

zur Ermittlung der Ruhelagen. Bezeichnet man mit $\mathbf{a}_1, \dots, \mathbf{a}_n$ die Spalten der $n \times n$ -Matrix \mathbf{A} , d.h.

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n]$$

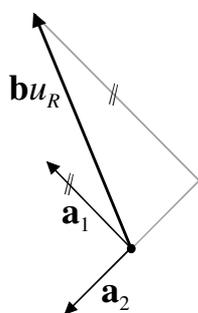
und mit $x_{1,R}, \dots, x_{n,R}$ die Komponenten der gesuchten Ruhelage \mathbf{x}_R , also

$$\mathbf{x}_R = [x_{1,R} \quad \dots \quad x_{n,R}]^T,$$

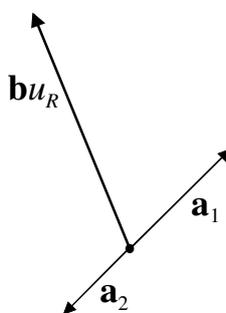
so kann (2.17) folgendermaßen dargestellt werden:

$$-(\mathbf{a}_1 x_{1,R} + \mathbf{a}_2 x_{2,R} + \dots + \mathbf{a}_n x_{n,R}) = \mathbf{b} u_R$$

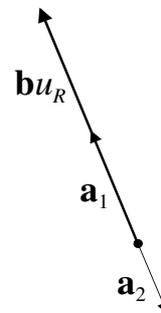
Die Ermittlung möglicher Ruhelagen reduziert sich also auf die Beantwortung der Frage, ob der Vektor $(\mathbf{b} u_R)$ als Linearkombination der Spaltenvektoren von \mathbf{A} dargestellt werden kann. Ist das der Fall, so gibt es mindestens eine Ruhelage, anderenfalls existiert keine Ruhelage. Eine geometrische Interpretation für $n = 2$ liefern die folgenden Abbildungen:



1 Ruhelage



keine Ruhelagen

 ∞ viele Ruhelagen

Über den so genannten Rang einer Matrix kann die vorliegende Problemstellung prägnant formuliert werden. Er gibt an, wieviele linear unabhängige Spalten (bzw. Zeilen) eine Matrix besitzt. Offensichtlich kann der Vektor $(\mathbf{b} u_R)$ genau dann als Linearkombination der Spalten von \mathbf{A} dargestellt werden, wenn gilt:

$$\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A} : \mathbf{b} u_R)$$

Das bedeutet, dass die $n \times (n + 1)$ -Matrix, die sich durch das Hinzufügen des Vektors $(\mathbf{b} u_R)$ als zusätzliche Spalte ergibt, den gleichen Rang besitzt wie die Matrix \mathbf{A} . Dies trifft auf jeden

Fall zu, wenn \mathbf{A} regulär ist, also den Höchststrang, d.h. $\text{rang}(\mathbf{A}) = n$ besitzt. Dann gibt es genau *eine* Ruhelage, nämlich

$$\mathbf{x}_R = -\mathbf{A}^{-1}\mathbf{b} u_R$$

Im Falle einer singulären Matrix \mathbf{A} mit $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A}:\mathbf{b} u_R) < n$ gibt es unendlich viele Ruhelagen, man spricht dann von einer *Ruhezone*.

Die gefundenen Ergebnisse sind in der folgenden Tabelle zusammengefasst.

rang(\mathbf{A}) = n	rang(\mathbf{A}) < n	
$\mathbf{x}_R = -\mathbf{A}^{-1}\mathbf{b} u_R$	rang(\mathbf{A}) = rang($\mathbf{A}:\mathbf{b} u_R$)	rang(\mathbf{A}) < rang($\mathbf{A}:\mathbf{b} u_R$)
genau <i>eine</i> Ruhelage	∞ viele Ruhelagen	keine Ruhelage

B Es sollen die Ruhelagen des linearen, zeitinvarianten Systems

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

ermittelt werden. Es gilt:

$$(\mathbf{A}:\mathbf{b} u_R) = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & u_R \end{bmatrix}$$

Für den Fall $u_R = 0$, d.h. $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A}:\mathbf{b} u_R) = 1$ gibt es unendlich viele Ruhelagen, nämlich

$$\mathbf{x}_R = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

wobei α eine beliebige reelle Konstante ist. Für $u_R \neq 0$, also $\text{rang}(\mathbf{A}) < \text{rang}(\mathbf{A}:\mathbf{b} u_R) = 2$ existiert keine Ruhelage. ■

2.4 Linearisierung um eine Ruhelage

In vielen Fällen führt die mathematische Modellierung von Systemen auf nichtlineare Modelle der Form

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, u), \quad (2.18)$$

$$y = g(\mathbf{x}, u). \quad (2.19)$$

Eine naheliegende Idee besteht darin, das Verhalten des nichtlinearen Systems durch ein lineares System zu approximieren. Eine solche Approximation wird im Allgemeinen nicht global, d.h. im gesamten Zustandsraum gültig sein. Eine in der Praxis bewährte Methode ist die so genannte **Linearisierung um einen Arbeitspunkt**. Ein Arbeitspunkt des nichtlinearen

Systems ist hier dadurch charakterisiert, dass alle Systemgrößen, also Eingangsgröße, Zustandsgrößen und damit auch die Ausgangsgröße konstante Werte annehmen, d.h.

$$u(t) = u_R = \text{konst.}, \quad \mathbf{x}(t) = \mathbf{x}_R = \text{konst.} \quad \Rightarrow \quad y(t) = y_R = \text{konst.}$$

Das System befindet sich also in einem „Gleichgewichtszustand“, der auch **Ruhelage** genannt wird. Die konstanten Systemgrößen müssen die aus (2.18) resultierenden Bedingungen

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_R, u_R) \quad (2.20)$$

erfüllen, für die zugehörige Ausgangsgröße gilt dann

$$y_R = g(\mathbf{x}_R, u_R). \quad (2.21)$$

Da das Systemverhalten „in der Nähe“ des Arbeitspunktes beschrieben werden soll, ist es zweckmäßig, mit Abweichungen aus der betrachteten Ruhelage zu operieren. Aus diesem Grund setzt man

$$\mathbf{x} = \mathbf{x}_R + \Delta\mathbf{x}, \quad u = u_R + \Delta u, \quad (2.22)$$

wobei $\Delta\mathbf{x}$, Δu Auslenkungen aus der Ruhelage repräsentieren. Damit lautet die Systembeschreibung (2.18) nun

$$\frac{d\Delta\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}_R + \Delta\mathbf{x}, u_R + \Delta u).$$

Mit dem Ziel der Linearisierung um den Arbeitspunkt wird die nichtlineare Funktion \mathbf{f} in eine Taylor-Reihe³ entwickelt, d.h.

$$\mathbf{f}(\mathbf{x}_R + \Delta\mathbf{x}, u_R + \Delta u) = \mathbf{f}(\mathbf{x}_R, u_R) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R} \Delta\mathbf{x} + \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}_R, u_R} \Delta u + \text{T.h.O.}$$

Unter Vernachlässigung der Terme höherer Ordnung (T.h.O.) und unter Berücksichtigung von (2.20) erhält man somit das vereinfachte System

$$\frac{d\Delta\mathbf{x}}{dt} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R} \Delta\mathbf{x} + \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}_R, u_R} \Delta u. \quad (2.23)$$

Dabei gilt

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial \mathbf{x}} \\ \frac{\partial f_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial f_n}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad \text{und} \quad \frac{\partial \mathbf{f}}{\partial u} = \begin{bmatrix} \frac{\partial f_1}{\partial u} \\ \frac{\partial f_2}{\partial u} \\ \vdots \\ \frac{\partial f_n}{\partial u} \end{bmatrix}, \quad (2.24)$$

³benannt nach dem britischen Mathematiker Brook Taylor (1685 - 1731)

wobei die $n \times n$ Matrix $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)$ die so genannte Jacobi⁴-Matrix ist. Für die Ausgangsgröße y folgt auf analoge Weise nach einer Taylor-Reihenentwicklung, also

$$y = g(\mathbf{x}_R + \Delta \mathbf{x}, u_R + \Delta u) = g(\mathbf{x}_R, u_R) + \left. \frac{\partial g}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R} \Delta \mathbf{x} + \left. \frac{\partial g}{\partial u} \right|_{\mathbf{x}_R, u_R} \Delta u + \text{T.h.O.},$$

nach Vernachlässigung von Termen höherer Ordnung und unter Berücksichtigung von (2.21)

$$\Delta y = \left. \frac{\partial g}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R} \Delta \mathbf{x} + \left. \frac{\partial g}{\partial u} \right|_{\mathbf{x}_R, u_R} \Delta u. \quad (2.25)$$

Man erkennt, dass es sich bei (2.23), (2.25) um ein lineares, zeitinvariantes System der Form (1.13) handelt, wobei

$$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R}, \quad \mathbf{b} = \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}_R, u_R}, \quad \mathbf{c}^T = \left. \frac{\partial g}{\partial \mathbf{x}} \right|_{\mathbf{x}_R, u_R} \quad \text{und} \quad d = \left. \frac{\partial g}{\partial u} \right|_{\mathbf{x}_R, u_R}$$

gilt. Dieses System approximiert das Verhalten des nichtlinearen Systems (2.18) in einer Umgebung der betrachteten Ruhelage.

B Es wird wieder das in Bild (1.6) dargestellte Pendel untersucht. Wie bereits gezeigt, lautet das zugehörige mathematische Modell

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{g}{l} \sin x_1 - \frac{k}{ml^2} x_2 \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix}.$$

Die Ruhelagen des (autonomen) Systems sind durch

$$x_{2,R} = 0 \quad \text{und} \quad \sin x_{1,R} = 0$$

charakterisiert. Für

$$x_{1,R} = \pm 2i\pi, \quad i = 0, 1, 2, \dots$$

befindet sich das Pendel in der „hängenden“ Ruhelage, während

$$x_{1,R} = \pm(2i + 1)\pi, \quad i = 0, 1, 2, \dots$$

die „aufrechte“ Ruhelage beschreibt. Für die Linearisierung um diese Arbeitspunkte berechnet man die Jacobi-Matrix

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{l} \cos x_1 & -\frac{k}{ml^2} \end{bmatrix},$$

⁴benannt nach dem deutschen Mathematiker Carl Gustav Jacob Jacobi (1804 - 1851)

d.h. das linearisierte mathematische Modell lautet

$$\frac{d\Delta\mathbf{x}}{dt} = \left[\begin{array}{cc} 0 & 1 \\ -\frac{g}{l} \cos x_1 & -\frac{k}{ml^2} \end{array} \right] \Bigg|_{\mathbf{x}_R} \Delta\mathbf{x}.$$

Daraus folgen unmittelbar die linearisierten Systembeschreibungen

$$\frac{d\Delta\mathbf{x}}{dt} = \left[\begin{array}{cc} 0 & 1 \\ -\frac{g}{l} & -\frac{k}{ml^2} \end{array} \right] \Delta\mathbf{x} \quad \text{„hängende“ Position} \quad \frac{d\Delta\mathbf{x}}{dt} = \left[\begin{array}{cc} 0 & 1 \\ \frac{g}{l} & -\frac{k}{ml^2} \end{array} \right] \Delta\mathbf{x} \quad \text{„aufrechte“ Position}$$

In Bild 2.1 sind die Verläufe des Pendelwinkels von nichtlinearem und linearisiertem Modell dargestellt. Als Anfangsauslenkungen wurde $x_{1,0} = 0.35 \text{ rad}$ ($=20^\circ$) bzw. $x_{1,0} = 1.05 \text{ rad}$ ($=60^\circ$) gewählt, das Pendel befindet sich zum Zeitpunkt $t = 0$ im Stillstand, d.h. $x_{2,0} = 0 \text{ rad s}^{-1}$. ■

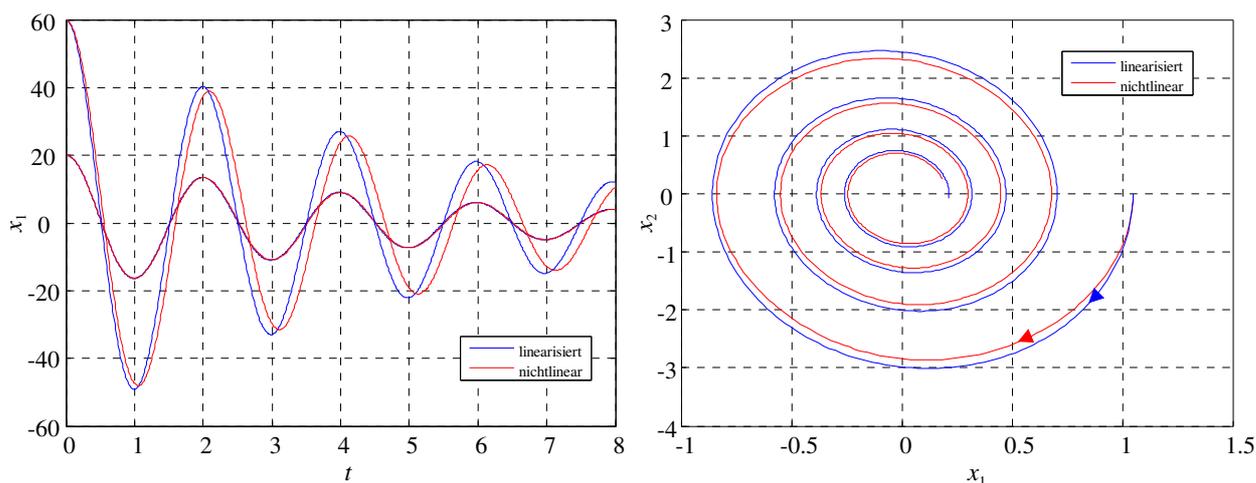


Bild 2.1: Gegenüberstellung des linearisierten Pendelmodells und des nichtlinearen Modells (rot). linkes Bild: Pendelwinkel über Zeit rechtes Bild: Verlauf der Trajektorien

B Gegeben sei das mathematische Modell

$$\frac{dx}{dt} = -(x-2)^2 + 9u^2, \quad y = x^2 + u$$

eines Systems mit der Eingangsgröße u , der Ausgangsgröße y und der Zustandsgröße x . Ruhelagen müssen die Bedingung $(x_R - 2)^2 = 9u_R^2$ erfüllen. Für einen vorgegebenen Wert $u_R = 1$ ergeben sich somit die beiden Ruhelagen

$$x_{R,1} = 5 \quad \text{und} \quad x_{R,2} = -1.$$

Die zugehörigen Ausgangsgrößen nehmen die Werte $y_{R,1} = 26$ bzw. $y_{R,2} = 2$ an. Mit $f(x, u) = -(x - 2)^2 + 9u^2$ und $g(x, u) = x^2 + u$ folgt unmittelbar

$$\frac{\partial f}{\partial x} = -2x + 4, \quad \frac{\partial f}{\partial u} = 18u, \quad \frac{\partial g}{\partial x} = 2x, \quad \frac{\partial g}{\partial u} = 1.$$

Für die (willkürlich gewählte) erste Ruhelage folgt somit die linearisierte Systembeschreibung

$$\frac{d\Delta x}{dt} = -6\Delta x + 18\Delta u, \quad \Delta y = 10\Delta x + \Delta u.$$

■

2.5 Asymptotische Stabilität

Zur Beurteilung der so genannten **asymptotischen Stabilität** wird das autonome System

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} \quad \text{mit} \quad \mathbf{x}_0 := \mathbf{x}(t=0) \quad (2.26)$$

betrachtet. Das System (2.26) wird **asymptotisch stabil** genannt, wenn für *jeden* Anfangszustand \mathbf{x}_0 die Bedingung

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{0} \quad (2.27)$$

erfüllt ist. Das bedeutet, dass jede Trajektorie - unabhängig von ihrem Startpunkt \mathbf{x}_0 - für $t \rightarrow \infty$ in den Ursprung des Zustandsraumes einläuft.

Aus obiger Definition geht auch hervor, dass ein asymptotisch stabiles System (2.26) als einzige Ruhelage \mathbf{x}_R den Koordinatenursprung des Zustandsraumes besitzt.

Da (2.27) für jeden beliebigen Anfangszustand \mathbf{x}_0 gelten muss, folgt aus (2.11) unmittelbar

$$\lim_{t \rightarrow \infty} \phi(t) = \mathbf{0}. \quad (2.28)$$

Das bedeutet, dass jedes Element der Transitionsmatrix eines asymptotisch stabilen Systems für $t \rightarrow \infty$ verschwindet. Dies kann mit Hilfe des Endwertsatzes der Laplace-Transformation überprüft werden. Dazu wird die Laplace-Transformierte der Transitionsmatrix

$$\bar{\phi}(s) = (s\mathbf{E} - \mathbf{A})^{-1} = \frac{1}{\det(s\mathbf{E} - \mathbf{A})} (s\mathbf{E} - \mathbf{A})_{\text{adj}}.$$

betrachtet. Dabei ist $\Delta(s) = \det(s\mathbf{E} - \mathbf{A})$ das charakteristische Polynom der Matrix \mathbf{A} , das den Polynomgrad n besitzt. Die Adjunkte $(s\mathbf{E} - \mathbf{A})_{\text{adj}}$ ist eine Polynommatrix, deren Elemente Polynomgrade kleiner oder gleich $(n - 1)$ besitzen. Die Elemente von $\bar{\phi}(s)$ sind also gebrochen rationale Funktionen. Gemäß den Ausführungen über den Endwertsatz im Abschnitt ?? ist (2.28) genau dann erfüllt, wenn die Nennerpolynome aller Elemente von $\bar{\phi}(s)$ Hurwitzpolynome sind.

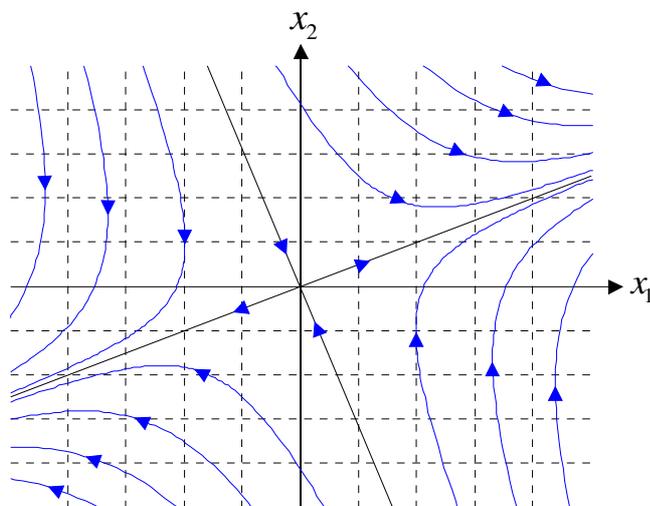


Bild 2.2: Trajektorien zu (2.29) - man spricht in diesem Zusammenhang von einem (instabilen) Sattelpunkt

B Gegeben sei

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \mathbf{x}. \quad (2.29)$$

Für $\bar{\phi}(s)$ gilt dann

$$\bar{\phi}(s) = \frac{1}{\underbrace{s^2 + 3s - 10}_{\Delta(s)}} \underbrace{\begin{bmatrix} s + 4 & 2 \\ 3 & s - 1 \end{bmatrix}}_{(s\mathbf{E} - \mathbf{A})_{\text{adj}}}.$$

Die Nennerpolynome der Elemente von $\bar{\phi}(s)$ sind sicher Hurwitzpolynome, wenn $\Delta(s)$ ein Hurwitzpolynom ist. Hierbei handelt es sich sogar um eine notwendige und hinreichende Bedingung. Da die Nullstellen von $\Delta(s)$ die Eigenwerte von \mathbf{A} sind, gilt zusammenfassend:

Das System (2.26) ist genau dann asymptotisch stabil, wenn alle Eigenwerte von \mathbf{A} einen negativen Realteil besitzen. Eine solche Matrix nennt man auch **Hurwitzmatrix**.

B Die Eigenwerte der Matrix \mathbf{A} aus (2.29) lauten

$$s_1 = 2 \quad \text{und} \quad s_2 = -5.$$

Das bedeutet, dass (2.29) nicht asymptotisch stabil ist, was auch den in Bild 2.2 dargestellten Trajektorien zu entnehmen ist.

B Das charakteristische Polynom der Dynamikmatrix des Systems

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \mathbf{x} \quad (2.30)$$

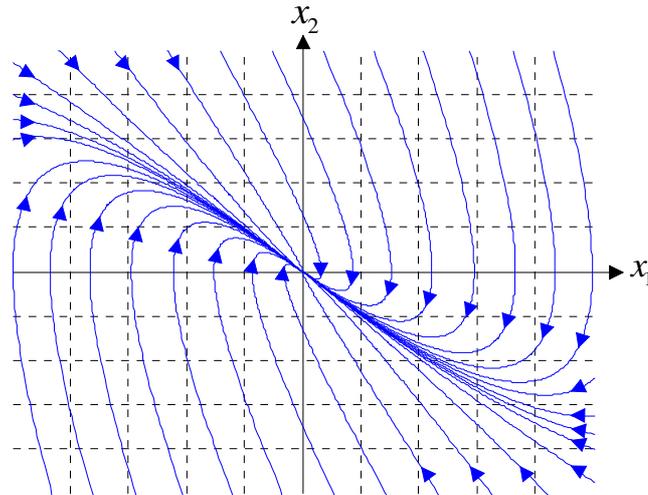


Bild 2.3: Trajektorien zu (2.30) - man spricht in diesem Zusammenhang auch von einem stabilen Knoten

lautet

$$\Delta(s) = s^2 + 3s + 2,$$

d.h. für die Eigenwerte gilt

$$s_1 = -1 \quad \text{und} \quad s_2 = -2.$$

Das bedeutet, dass das System asymptotisch stabil ist. Dies wird durch die in Bild 2.3 dargestellten Trajektorien bestätigt. ■

2.6 Übertragungsfunktion

Die **Übertragungsfunktion** $G(s)$ beschreibt das Übertragungsverhalten von linearen, zeitinvarianten Systemen im Bildbereich. Sie ist definiert als der Quotient der Laplace-Transformierten von Ausgangs- und Eingangsgröße, d.h.

$$G(s) = \frac{\bar{y}(s)}{\bar{u}(s)}. \quad (2.31)$$

Dabei wird vorausgesetzt, dass sich das System zum Zeitpunkt $t = 0$ in Ruhe befindet. Im Rahmen dieses Skriptums wird davon ausgegangen, dass die Übertragungsfunktion stets eine gebrochene rationale Funktion ist. Sie kann somit als Quotient zweier Polynome $\beta(s)$ und $\alpha(s)$ dargestellt werden, d.h.

$$G(s) = \frac{\beta(s)}{\alpha(s)}, \quad (2.32)$$

wobei die Koeffizienten der Polynome $\alpha(s)$ und $\beta(s)$ als reell vorausgesetzt werden. Systeme, bei denen die Polynomgrade von Zähler- und Nennerpolynom der Bedingung

$$\text{Grad } \beta(s) \leq \text{Grad } \alpha(s) \quad (2.33)$$

genügen, nennt man **realisierbar**, siehe auch Abschnitt ???. Alle praktisch relevanten Systeme erfüllen die Realisierbarkeitsbedingung. Im Falle

$$\text{Grad } \beta(s) = \text{Grad } \alpha(s)$$

spricht man von einem **sprungfähigen System**. Sprunghafte Änderungen der Eingangsgröße u haben hier auch eine sprunghafte Änderung der Ausgangsgröße y zur Folge.

Für Systeme der Form (2.1), (2.2) kann die Übertragungsfunktion unter der Annahme $\mathbf{x}_0 = \mathbf{0}$ gemäß

$$G(s) = \frac{\bar{y}(s)}{\bar{u}(s)} = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d \quad (2.34)$$

berechnet werden, vgl. (2.14). Aus dieser Berechnungsvorschrift kann unmittelbar gefolgert werden, dass $G(s)$ eine gebrochene rationale Funktion ist und (2.33) gilt. Man beachte, dass das System genau dann sprungfähig ist, wenn für den Durchgriff $d \neq 0$ gilt. Dann gibt es nämlich eine direkte Verbindung zwischen Eingangs- und Ausgangsgröße, siehe hierzu auch Bild ??.

In Strukturbildern wird eine Übertragungsfunktion $G(s)$ üblicherweise durch einen Block, wie er in Bild 2.4 dargestellt ist, repräsentiert.

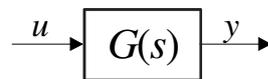


Bild 2.4: Blockdarstellung einer Übertragungsfunktion $G(s)$

B Gegeben sei ein System mit der Übertragungsfunktion

$$G(s) = \frac{\bar{y}(s)}{\bar{u}(s)} = \frac{(s-1)}{(s+1)(s+2)}.$$

Die **Sprungantwort** des Systems, d.h. seine Reaktion auf $u(t) = \sigma(t)$ kann mittels

$$\bar{y}(s) = G(s) \bar{u}(s) = G(s) \frac{1}{s} = \frac{(s-1)}{s(s+1)(s+2)}$$

berechnet werden. Eine Partialbruchzerlegung liefert

$$\bar{y}(s) = -\frac{1/2}{s} + \frac{2}{s+1} - \frac{3/2}{s+2},$$

was im Zeitbereich der Funktion

$$y(t) = -\frac{1}{2} + 2e^{-t} - \frac{3}{2}e^{-2t}$$

entspricht. ■

2.7 Deutung von $G(s)$ im Zeitbereich

Gemäß (2.31) gilt für den zeitlichen Verlauf der Ausgangsgröße y nach Anwendung des Faltungssatzes

$$y(t) = \mathcal{L}^{-1} \{G(s) \bar{u}(s)\} = \int_0^t g(t - \tau) u(\tau) d\tau, \quad (2.35)$$

wobei

$$g(t) = \mathcal{L}^{-1} \{G(s)\}. \quad (2.36)$$

Die Ausgangsgröße y entspricht also dem Integral über die gewichtete Eingangsgröße u . Aus diesem Grund nennt man $g(t)$ auch **Gewichtsfunktion**. Wählt man als Eingangsgröße des Systems einen Dirac-Impuls, also $u(t) = \delta(t)$, so gilt

$$y(t) = \mathcal{L}^{-1} \{G(s) \mathcal{L} \{\delta(t)\}\} = \mathcal{L}^{-1} \{G(s)\} = g(t),$$

die Ausgangsgröße entspricht dann $g(t)$. Deshalb wird $g(t)$ auch **Impulsantwort** des Systems genannt.

B Die Impulsantwort des Systems mit der Übertragungsfunktion

$$G(s) = \frac{1}{(s+1)(s-1)} = \frac{1}{2} \left(\frac{1}{s-1} - \frac{1}{s+1} \right)$$

lautet

$$g(t) = \frac{1}{2} (e^t - e^{-t}).$$

■

2.8 Pole und Nullstellen

Man beachte, dass in (2.32) die Polynome $\alpha(s)$ und $\beta(s)$ nicht notwendigerweise teilerfremd sind. Das bedeutet, dass unter Umständen Kürzungen durchgeführt werden können, wie das folgende Beispiel veranschaulicht.

B Gegeben sei die Übertragungsfunktion

$$G(s) = \frac{\beta(s)}{\alpha(s)} = \frac{s+1}{s^2+3s+2}.$$

Gibt man $G(s)$ in der faktorisierten Darstellung an, d.h.

$$G(s) = \frac{s+1}{(s+1)(s+2)} = \frac{1}{s+2},$$

so erkennt man, dass eine Kürzung möglich ist.

■

Geht man nun von einer teilerfremden Übertragungsfunktion

$$G(s) = \frac{\mu(s)}{\nu(s)} \quad \text{mit} \quad \mu(s), \nu(s) \dots \text{teilerfremd} \quad (2.37)$$

aus, so sind diejenigen Werte von s , für die $\mu(s) = 0$ gilt, die so genannten **Nullstellen** von $G(s)$. Die **Pole** oder **Polstellen** der Übertragungsfunktion sind durch $\nu(s) = 0$ charakterisiert. In der *faktorierten* Darstellung der Übertragungsfunktion

$$G(s) = K \frac{\prod_{i=1}^m (s - n_i)}{\prod_{k=1}^n (s - p_k)} \quad \text{mit} \quad m \leq n \quad (2.38)$$

ist $s = n_i$ eine Nullstelle und $s = p_k$ eine Polstelle von $G(s)$. Tritt der Faktor $(s - n_i)$ bzw. $(s - p_k)$ mehrfach auf, so besitzt das System eine mehrfache Nullstelle bzw. einen mehrfachen Pol. Da die Koeffizienten der Polynome $\mu(s)$ und $\nu(s)$ reell sind, treten Pole bzw. Nullstellen reell und/oder paarweise konjugiert komplex auf. Aus (2.38) folgt unmittelbar.

$$G(s) = 0 \quad \text{für} \quad s = n_i \quad \text{und} \quad |G(s)| \rightarrow \infty \quad \text{für} \quad s \rightarrow p_k \quad (2.39)$$

B Die Übertragungsfunktion

$$G(s) = \frac{s - 0.5}{s^2 + 1}$$

besitzt eine Nullstelle bei $s = 0.5$ und ein konjugiert komplexes Polpaar bei $s = \pm j$. In Bild 2.5 ist $|G(s)|$ über der komplexen s -Ebene dargestellt. Die in (2.39) angeführten Eigenschaften sind deutlich zu erkennen.

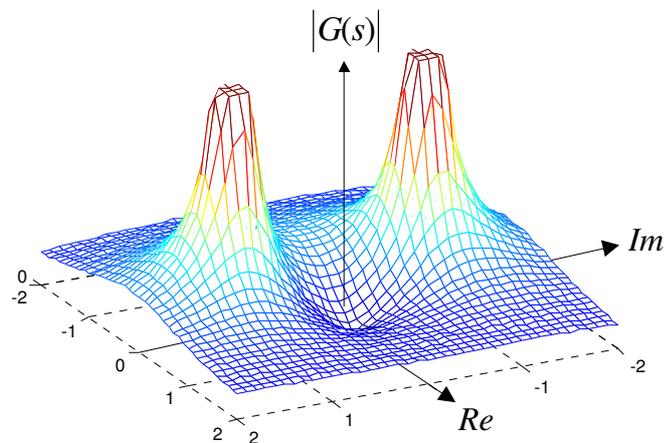


Bild 2.5: Graphische Darstellung von $|G(s)|$ über der komplexen s -Ebene zur Illustration von Pol- und Nullstellen.

B Die faktorisierte Darstellung der Übertragungsfunktion

$$G(s) = \frac{2s^2 + 4s + 4}{s(s+1)^2}$$

ist gegeben durch

$$G(s) = 2 \frac{(s+1-j)(s+1+j)}{s(s+1)(s+1)}.$$

Daraus resultiert ein konjugiert komplexes Nullstellenpaar an der Stelle

$$s = -1 \pm j,$$

ein reeller Pol bei

$$s = 0$$

sowie ein doppelter Pol an der Stelle

$$s = -1.$$

■

Man beachte, dass die Eigenwerte der Systemmatrix \mathbf{A} von (2.1) und die Pole der zugehörigen Übertragungsfunktion $G(s)$ in enger Relation zueinander stehen. Die Eigenwerte sind die Lösungen der charakteristischen Gleichung

$$\Delta(s) = \det(s\mathbf{E} - \mathbf{A}) = 0,$$

die Pole sind die Nullstellen des Nennerpolynoms der gekürzten(!) rationalen Funktion

$$G(s) = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} + d = \frac{1}{\Delta(s)} \mathbf{c}^T (s\mathbf{E} - \mathbf{A})_{adj} \mathbf{b} + d.$$

Daraus kann die Schlußfolgerung gezogen werden, dass die Pole von $G(s)$ eine Teilmenge der Eigenwerte von \mathbf{A} darstellen.

B Gegeben sei das mathematische Modell

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \\ y &= \begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{x} + u. \end{aligned}$$

Die zwei Eigenwerte lauten

$$s_1 = -1 \quad \text{und} \quad s_2 = -2.$$

Aus der Übertragungsfunktion

$$G(s) = \frac{s+2}{s+1}$$

ist zu erkennen, dass es einen Pol bei

$$s = -1$$

gibt.

■

2.8.1 PN-Plan

Die graphische Darstellung der Pole und Nullstellen einer Übertragungsfunktion in der komplexen Ebene wird PN-Plan genannt. Hierbei werden Pole durch ein **X** und Nullstellen durch ein **O** symbolisiert.

B Gegeben sei das System

$$G(s) = \frac{s^3 - 4s^2 + 6s - 4}{s^5 - 4s^4 + 3s^3 + 14s^2 + 26s} = \frac{(s-2)(s^2 - 2s + 2)}{s(s^2 + 2s + 2)(s^2 - 6s + 13)} \quad (2.40)$$

mit

$$\text{Nullstellen} \quad \{2, 1 \pm j\}$$

$$\text{Pole} \quad \{0, -1 \pm j, 3 \pm j2\}$$

Der zugehörige PN-Plan ist in Bild 2.6 zu sehen. Man beachte, dass aufgrund der Symmetrie des PN-Plans prinzipiell die Darstellung des grau hinterlegten Bereiches ausreicht. ■

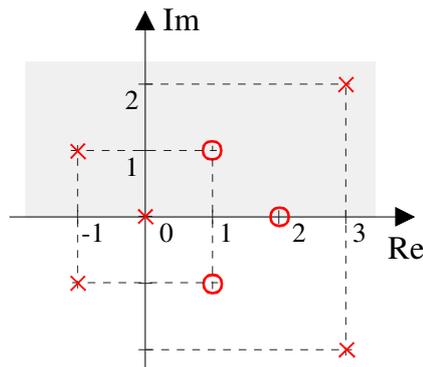


Bild 2.6: PN-Plan zu der Übertragungsfunktion (2.40)

2.9 BIBO-Stabilität

Ein lineares, zeitinvariantes System mit der Übertragungsfunktion $G(s)$ wird **BIBO-stabil** („bounded input - bounded output“) genannt, wenn es auf *jede* beschränkte Eingangsgröße $u(t)$ mit einer beschränkten Ausgangsgröße $y(t)$ reagiert, d.h.

$$|u(t)| \leq M < \infty \quad \Rightarrow \quad |y(t)| \leq N < \infty \quad \forall t \geq 0. \quad (2.41)$$

Mit (2.35) gilt nun unter Annahme einer beschränkten Eingangsgröße

$$|y(t)| = \left| \int_0^t g(\tau) u(t-\tau) d\tau \right| \leq \int_0^t |g(\tau)| |u(t-\tau)| d\tau \stackrel{(2.41)}{\leq} M \int_0^t |g(\tau)| d\tau.$$

Offensichtlich tritt der größtmögliche Wert von $|y(t)|$ für $t \rightarrow \infty$ auf, d.h. zur Erfüllung von (2.41) muss

$$M \int_0^{\infty} |g(\tau)| d\tau \leq N$$

gelten. Das bedeutet, dass die Impulsantwort absolut integrierbar sein muss, d.h.

$$\int_0^{\infty} |g(t)| dt < \infty. \quad (2.42)$$

Man beachte, dass es sich bei (2.42) um eine notwendige und hinreichende Bedingung handelt!

B Gegeben sei die Übertragungsfunktion

$$G(s) = \frac{1}{s - \kappa} \quad \text{mit} \quad \kappa \in \mathbb{R},$$

die zugehörige Impulsantwort lautet

$$g(t) = \begin{cases} 0 & \text{für } t < 0 \\ e^{\kappa t} & \text{für } t \geq 0 \end{cases}.$$

Das Integral über den Absolutbetrag von $g(t)$, also

$$\int_0^{\infty} |g(t)| dt = \int_0^{\infty} |e^{\kappa t}| dt = \int_0^{\infty} e^{\kappa t} dt$$

existiert offensichtlich nur für negative Werte von κ . Daraus kann gefolgert werden, dass das betrachtete System genau dann BIBO-stabil ist, wenn sein Pol $s = \kappa$ in der linken offenen Halbebene der komplexen s -Ebene liegt, d.h. $\operatorname{Re}\{s\} < 0$. ■

Die Erkenntnis aus obigem Beispiel kann auf den allgemeinen Fall ausgeweitet werden. Um dies zu zeigen wird zunächst der Zusammenhang zwischen der Impulsantwort $g(t)$ und der Übertragungsfunktion betrachtet, d.h.

$$G(s) = \mathcal{L}\{g(t)\} = \int_0^{\infty} g(t)e^{-st} dt.$$

Daraus kann folgende Abschätzung abgeleitet werden

$$|G(s)| = \left| \int_0^{\infty} g(t)e^{-st} dt \right| \leq \int_0^{\infty} |g(t)| |e^{-st}| dt.$$

Gilt $\operatorname{Re}\{s\} \geq 0$, so kann weiter geschrieben werden

$$|G(s)| \leq \int_0^{\infty} |g(t)| dt.$$

Das bedeutet, dass im Falle eines BIBO-stabilen Systems - (2.42) ist also erfüllt - der Ausdruck $|G(s)|$ für $\operatorname{Re}\{s\} \geq 0$ *endliche* Werte annimmt. Daraus kann mit Hilfe von (2.39) gefolgert werden, dass $G(s)$ ausschließlich Pole in der linken, offenen komplexen Halbebene besitzt.

Ein System mit der Übertragungsfunktion $G(s)$ ist somit genau dann BIBO-stabil, wenn *alle* Pole von $G(s)$ einen negativen Realteil besitzen. Das bedeutet, dass in der gekürzten Darstellung (2.37) das Nennerpolynom von $G(s)$ ein Hurwitzpolynom sein muss.

B Gegeben sei ein System mit der Übertragungsfunktion

$$G(s) = \frac{s + 4}{s(s + 1)(s + 2)}.$$

Das System ist nicht BIBO-stabil, da es einen Pol mit nicht-negativem Realteil besitzt. ■

B Gegeben sei ein System mit der Übertragungsfunktion

$$G(s) = \frac{s - 1}{(s + 1)^2(s + 2)(s + 4)}.$$

Das System ist BIBO-stabil, da alle Pole von $G(s)$ einen negativen Realteil besitzen. ■

2.10 Das Routh-Schema

Die asymptotische Stabilität erfordert es, dass alle Eigenwerte der Dynamikmatrix einen negativen Realteil besitzen. Weiters ist ein Übertragungssystem genau dann BIBO-stabil, wenn seine Übertragungsfunktion ausschließlich Pole mit negativen Realteilen besitzen. Man beachte, dass die beiden Stabilitätskriterien die Kenntnis der Lage der Eigenwerte bzw. Pole *nicht* erfordern. Vielmehr ist zu überprüfen, ob das charakteristische Polynom der Dynamikmatrix bzw. das Nennerpolynom der Übertragungsfunktion ein Hurwitzpolynom ist. Hierfür gibt es eine Reihe von Verfahren, die ohne die explizite Berechnung der Polynomnullstellen auskommen. Exemplarisch wird an dieser Stelle das so genannte **Routh-Schema**⁵ vorgestellt, wobei auf den Beweis verzichtet wird.

Ausgangspunkt der Betrachtungen ist das zu untersuchende Polynom

$$p(s) = a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0. \quad (2.43)$$

Eine notwendige Bedingung dafür, dass $p(s)$ ein Hurwitzpolynom ist, besteht darin, dass kein Polynomkoeffizient verschwindet und alle Polynomkoeffizienten das gleiche Vorzeichen besitzen. Ist diese Bedingung verletzt, so ist $p(s)$ sicher kein Hurwitzpolynom. Man beachte, dass für Polynome mit $\operatorname{Grad} p(s) \leq 2$ die angegebene Bedingung sogar *notwendig und hinreichend* ist.

⁵benannt nach dem englischen Mathematiker Edward Routh (1831 - 1907)

B Gegeben seien die Polynome

$$\begin{aligned} p_1(s) &= 2s^4 + 2s^2 + s + 1, \\ p_2(s) &= s^5 + 3s^4 - 2s^3 + s^2 + 1, \\ p_3(s) &= 5s^6 + 3s^5 + 2s^4 + 7s^3 + s^2 + 9s + 1, \\ p_4(s) &= -2s^2 - 2 - 1, \\ p_5(s) &= s^2 + 4s - 3, \\ p_6(s) &= s^2 + s + 1, \\ p_7(s) &= s + 8. \end{aligned}$$

Die Polynome $p_1(s)$, $p_2(s)$, $p_5(s)$ erfüllen die angegebene notwendige Bedingung nicht, sie sind also keine Hurwitzpolynome. Das Polynom $p_3(s)$ erfüllt die Bedingung, der Nachweis ob es sich um ein Hurwitzpolynom handelt, erfordert allerdings weitere Schritte (siehe weiter unten). Die Polynome $p_4(s)$, $p_6(s)$ und $p_7(s)$ können direkt als Hurwitzpolynome klassifiziert werden, da sie die Bedingung erfüllen und einen Polynomgrad kleiner oder gleich 2 besitzen. ■

Zunächst werden beim Routh-Schema die Koeffizienten des Polynoms (2.43) in folgendem „Zick-Zack-Muster“ angeordnet

$$\begin{array}{ccccccc} a_n & & a_{n-2} & & a_{n-4} & & \cdots \\ \downarrow & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \\ a_{n-1} & & a_{n-3} & & a_{n-5} & & \cdots \end{array}, \quad (2.44)$$

wobei bei geradem Polynomgrad n die letzte Spalte mit einem Nullelement aufgefüllt bzw. bei ungeradem Polynomgrad eine Nullspalte hinzugefügt wird.

B Für das Polynom

$$p(s) = s^4 + 2s^3 + 3s^2 + 4s + 5,$$

ergibt sich nach (2.44)

$$\begin{array}{cccc} 1 & 3 & 5 & \\ 2 & 4 & 0 & \end{array}.$$

Hingegen ergibt sich für

$$p_2(s) = 6s^5 + 7s^4 + 8s^3 + 9s^2 + 10s + 11$$

die Koeffizientenanzordnung

$$\begin{array}{cccc} 6 & 8 & 10 & 0 \\ 7 & 9 & 11 & 0 \end{array}.$$

■

Ausgehend von der 2-zeiligen Anordnung (2.44) werden weitere Zeilen nach folgendem Schema berechnet bis insgesamt $(n + 1)$ Zeilen vorliegen: Der Koeffizient in der k -ten Spalte einer neuen Zeile entspricht der zweireihigen Determinante aus den Elementen der ersten und

$(k + 1)$ -ten Spalte der beiden darüberliegenden Zeilen, geteilt durch das negative erste Element der darüberliegenden Zeile. Mit Hilfe des so entstehenden **Routh-Tableaus** kann nun entschieden werden, ob ein Hurwitzpolynom vorliegt:

Das Polynom (2.43) ist *genau dann* ein Hurwitzpolynom, wenn alle $(n + 1)$ Elemente der ersten Spalte sich von Null unterscheiden und das gleiche Vorzeichen besitzen.

B Gegeben sei das Polynom

$$p(s) = s^4 + 2s^3 + 3s^2 + 4s + 5.$$

Die oben angegebene notwendige Bedingung ist erfüllt. Ob das Polynom tatsächlich ein Hurwitzpolynom ist, kann nun mit dem Routh-Schema überprüft werden. Zunächst werden die ersten beiden Zeilen gemäß (2.44) aufgestellt, d.h.

$$\begin{array}{ccc} 1 & 3 & 5 \\ 2 & 4 & 0 \end{array} \cdot$$

Der oben angegebenen Berechnungsvorschrift folgend ermittelt man nun die dritte und vierte Zeile, d.h. es ergibt sich

$$\begin{array}{ccc} 1 & 3 & 5 \\ 2 & 4 & 0 \\ \frac{1 \cdot 4 - 2 \cdot 3}{(-2)} = 1 & \frac{1 \cdot 0 - 2 \cdot 5}{(-2)} = 5 & \\ 2 \cdot 5 - 1 \cdot 4 & & \\ \frac{(-1)}{(-1)} = -6 & & \end{array} \cdot$$

Da das erste Element der vierten Zeile ein anderes Vorzeichen hat als die übrigen Elemente der ersten Spalte, ist $p(s)$ kein Hurwitzpolynom. Das Routh-Tableau muss hier also nicht vollständig berechnet werden. ■

B Gegeben sei das Polynom

$$p(s) = s^3 + 3s^2 + 2s + 1.$$

Es soll mit dem Routh-Schema überprüft werden, ob ein Hurwitzpolynom vorliegt. Zunächst werden wieder die ersten beiden Zeilen gemäß (2.44) aufgestellt, d.h.

$$\begin{array}{ccc} 1 & 2 & 0 \\ 3 & 1 & 0 \end{array} \cdot$$

Gemäß obiger Berechnungsvorschrift erhält man für die dritte Zeile

$$\begin{array}{ccc} 1 & 2 & 0 \\ 3 & 1 & 0 \\ \frac{1 \cdot 1 - 3 \cdot 2}{(-3)} = \frac{5}{3} & \frac{1 \cdot 0 - 3 \cdot 0}{(-3)} = 0 & \end{array} \cdot$$

Da das erste Element der dritten Zeile das gleiche Vorzeichen hat wie die übrigen Elemente, wird die nächste (und hier auch zugleich letzte) Zeile ermittelt und man findet

$$\begin{array}{ccc} 1 & 2 & 0 \\ 3 & 1 & 0 \\ \frac{5}{3} & 0 & \end{array} .$$

$$\frac{3 \cdot 0 - \frac{5}{3} \cdot 1}{(-\frac{5}{3})} = 1$$

Alle Elemente der ersten Spalte sind von Null verschieden und haben das gleiche Vorzeichen, d.h. $p(s)$ ist ein Hurwitzpolynom. ■

Besonders von Vorteil ist der Einsatz des Routh-Schemas, wenn die Koeffizienten des zu untersuchenden Polynoms von freien Parametern abhängen.

B Gegeben sei das Polynom

$$p(s) = s^3 + Ks^2 + s + 1.$$

Es soll mit Hilfe des Routh-Schemas überprüft werden, für welche Werte des reellen Parameters K das Polynom $p(s)$ ein Hurwitzpolynom ist. Aus der Forderung, dass alle Polynomkoeffizienten das gleiche Vorzeichen besitzen, folgt unmittelbar die notwendige Bedingung, dass K positiv sein muss. Die ersten beiden Zeilen des Routh-Tableaus ergeben sich zu

$$\begin{array}{ccc} 1 & 1 & 0 \\ K & 1 & 0 \end{array} .$$

Für die weiteren Zeilen erhält man

$$\begin{array}{ccc} 1 & 1 & 0 \\ K & 1 & 0 \\ \frac{1 \cdot 1 - K \cdot 1}{(-K)} = \frac{K-1}{K} & \frac{1 \cdot 0 - K \cdot 0}{(-K)} = 0 & \\ \frac{K \cdot 0 - \frac{K-1}{K} \cdot 1}{-\frac{K-1}{K}} = 1 & & \end{array} .$$

Damit alle Elemente der ersten Spalte gleiches Vorzeichen besitzen, muss

$$\frac{K-1}{K} > 0$$

gelten. Da K sicher positiv ist, folgt daraus unmittelbar der gesuchte Wertebereich

$$K > 1.$$

2.11 Zusammenschaltung von Übertragungssystemen

Die Übertragungsfunktion eines aus mehreren Übertragungssystemen rückwirkungsfrei(!) zusammengesetzten Gesamtsystems kann durch Anwendung der folgenden „Rechenregeln für Übertragungsfunktionen“ systematisch ermittelt werden.

2.11.1 Reihenschaltung

Für die in Bild 2.7 dargestellte Serienschaltung zweier Systeme mit den Übertragungsfunktionen $A(s)$ und $B(s)$ gilt

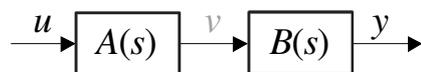


Bild 2.7: Reihenschaltung von Übertragungsfunktionen

$$\bar{y}(s) = B(s) \bar{v}(s) = B(s) A(s) \bar{u}(s).$$

Für die Gesamtübertragungsfunktion gilt somit

$$G(s) = A(s) B(s). \quad (2.45)$$

Die Übertragungsfunktion einer Serienschaltung entspricht also dem *Produkt* der einzelnen Übertragungsfunktionen. Man beachte, dass $G(s)$ ausschließlich Pole bzw. Nullstellen besitzt, die auch Pole bzw. Nullstellen von $A(s)$ und/oder $B(s)$ sind.

B Aus der Serienschaltung der Systeme mit den Übertragungsfunktionen

$$A(s) = \frac{1}{s+1} \quad \text{und} \quad B(s) = \frac{s+2}{s}$$

ergibt sich ein Gesamtsystem mit der Übertragungsfunktion

$$G(s) = A(s) B(s) = \frac{s+2}{s(s+1)}.$$

Die Pole von $G(s)$ liegen bei $s_1 = 0$ und $s_2 = -1$, die Nullstelle bei $s = -2$. ■

Das folgende Beispiel soll verdeutlichen, dass eine „Stabilisierung“ mittels einer Serienschaltung nicht möglich ist.

B Gegeben sei ein instabiles⁶ System mit der Übertragungsfunktion

$$B(s) = \frac{1}{s-5}.$$

⁶unter einem instabilen System verstehen wir hier ein System, das nicht BIBO-stabil ist.

Mit der Wahl

$$A(s) = \frac{s - 5}{s + 1}$$

gelingt es, ein BIBO-stabiles Gesamtsystem mit der Übertragungsfunktion

$$G(s) = A(s)B(s) = \frac{(s - 5)}{(s + 1)} \frac{1}{(s - 5)} = \frac{1}{s + 1}$$

zu erzeugen. Der „instabile“ Pol von $B(s)$ wird durch eine entsprechende Nullstelle von $A(s)$ kompensiert, man spricht von einer „**instabilen Kürzung**“. In der Realität muss allerdings davon ausgegangen werden, dass das Signal v mit einer Störung überlagert ist, was in Bild 2.8 durch eine Störgröße d angedeutet ist. Damit gilt

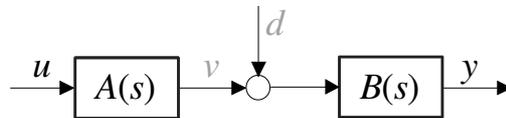


Bild 2.8: Zur Serienschaltung von Systemen

$$\bar{y}(s) = B(s) [\bar{v}(s) + \bar{d}(s)] = B(s) A(s) \bar{u}(s) + B(s) \bar{d}(s).$$

Das bedeutet, dass der Zusammenhang zwischen Störung und Ausgangsgröße durch die instabile Übertragungsfunktion $B(s)$ beschrieben wird, das Gesamtsystem ist somit praktisch unbrauchbar. Diese Problematik tritt auch bei der numerischen Simulation der Serienschaltung auf. Aufgrund der numerischen Darstellung von Zahlen mit endlicher Genauigkeit ist die Größe v mit einem Fehler überlagert, der als Störung d interpretiert werden kann. ■

Das obige Beispiel veranschaulicht, dass instabile Kürzungen *unter keinen Umständen* durchgeführt werden dürfen. Ein weiterer Grund hierfür besteht darin, dass Übertragungsfunktionen stets nur Modelle der Realität darstellen und daher immer mit Unsicherheiten behaftet sind. *Exakte* Kürzungen sind somit grundsätzlich nicht möglich.

2.11.2 Parallelschaltung

Die Übertragungsfunktion $G(s)$ der in Bild 2.9 dargestellten Parallelschaltung der Systeme mit den Übertragungsfunktionen $A(s)$ und $B(s)$ kann mit Hilfe der Zusammenhänge

$$\bar{v}_1(s) = A(s) \bar{u}(s) \quad \text{und} \quad \bar{v}_2(s) = B(s) \bar{u}(s)$$

sowie

$$\bar{y}(s) = A(s) \bar{u}(s) + B(s) \bar{u}(s)$$

ermittelt werden, es gilt

$$G(s) = \frac{\bar{y}(s)}{\bar{u}(s)} = A(s) + B(s). \quad (2.46)$$

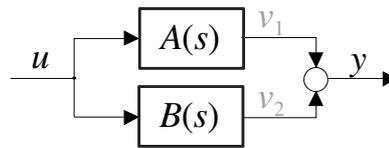


Bild 2.9: Parallelschaltung von Übertragungsfunktion

Die Übertragungsfunktion der Parallelschaltung entspricht also der *Summe* der einzelnen Übertragungsfunktionen. Man beachte, dass $G(s)$ ausschließlich Pole besitzt, die auch Pole von $A(s)$ und/oder $B(s)$ sind.

B Aus der Parallelschaltung der Systeme mit den Übertragungsfunktionen

$$A(s) = \frac{1}{s+1} \quad \text{und} \quad B(s) = \frac{s+2}{s}$$

ergibt sich ein Gesamtsystem mit der Übertragungsfunktion

$$G(s) = A(s) + B(s) = \frac{s^2 + 4s + 2}{s(s+1)}.$$

Die Übertragungsfunktion besitzt Pole bei $s = 0$ und $s = -1$ sowie Nullstellen bei $s = 2 + \sqrt{2} = 3.4142$ und $s = 2 - \sqrt{2} = 0.5858$. ■

2.11.3 Rückkopplung

Die in Bild 2.10 dargestellte rückgekoppelte Struktur spielt vor allem in der Regelungstechnik eine bedeutende Rolle.

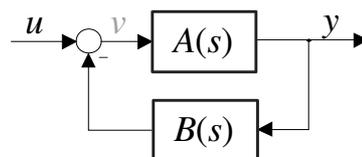


Bild 2.10: Rückgekoppelte Struktur

Es gilt

$$\bar{y}(s) = A(s) \bar{v}(s) \quad \text{und} \quad \bar{v}(s) = \bar{u}(s) - B(s) \bar{y}(s),$$

d.h.

$$\bar{y}(s) = A(s) \bar{u}(s) - A(s) B(s) \bar{y}(s).$$

Daraus folgt die gesuchte Übertragungsfunktion

$$G(s) = \frac{A(s)}{1 + A(s)B(s)}. \quad (2.47)$$

Als Hilfestellung für die Ermittlung der Übertragungsfunktion von rückgekoppelten Strukturen kann die Merkregel

$$G_{\text{gesamt}} = \frac{G_{\text{Vorwärtszweig}}}{1 + G_{\text{Schleife}}} \quad (2.48)$$

dienen, siehe hierzu auch Bild 2.11.



Bild 2.11: Zur Ermittlung der Übertragungsfunktion rückgekoppelter Systeme

B Aus der Zusammenschaltung der Systeme mit den Übertragungsfunktionen

$$A(s) = \frac{1}{s+2} \quad \text{und} \quad B(s) = \frac{1}{s}$$

gemäß Bild 2.10 ergibt sich ein Gesamtsystem mit der Übertragungsfunktion

$$G(s) = \frac{s}{(s+1)^2}.$$

Bemerkenswert ist, dass die Pole $s_{1,2} = -1$ von $G(s)$ weder Pole von $A(s)$ noch von $B(s)$ sind. Diese Möglichkeit, die Pole des „geschlossenen Kreises“ zu beeinflussen, wird in der Regelungstechnik beim Entwurf von Regelgesetzen ausgenutzt. ■

2.11.4 Allgemeine Strukturen

Ein Algorithmus zu Berechnung der Übertragungsfunktion von komplexen Zusammenschaltungen von Übertragungssystemen ist die so genannte Mason-Formel⁷, siehe z.B. [?]. Durch geschickte Anwendung der bereits eingeführten Rechenregeln für Übertragungsfunktionen (Serien-, Parallelschaltung und Rückkopplung) können komplizierte Strukturen auch ohne den etwas schwerfälligen Algorithmus von Mason problemlos aufgelöst werden.

B Die Übertragungsfunktion des in Bild 2.12 dargestellten Systems soll ermittelt werden. Es handelt sich um die Reihenschaltung der Übertragungsfunktion $A(s)$ und einer Rückkopplung mit

$$G_{\text{Vorwärtszweig}} = B(s) + C(s) \quad \text{und} \quad G_{\text{Schleife}} = [B(s) + C(s)] D(s) E(s).$$

Für die Gesamtübertragungsfunktion gilt somit

$$G(s) = A(s) \frac{B(s) + C(s)}{1 + [B(s) + C(s)] D(s) E(s)}.$$

⁷benannt nach dem amerikanischen Ingenieur Samuel Jefferson Mason (1921 - 1974)

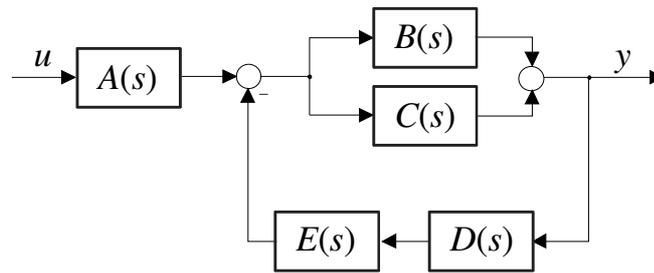


Bild 2.12: Zusammenschaltung mehrerer Übertragungsfunktionen

Bei der Ermittlung der Übertragungsfunktion eines komplexen Systems aus den Übertragungsfunktionen der Teilsysteme ist unbedingt darauf zu achten, dass die Teilsysteme rückwirkungsfrei gekoppelt sind. Im folgenden Beispiel wird diese Problematik illustriert.

B Gegeben sei das in Bild 2.13 dargestellte Netzwerk, bestehend aus idealen Bauelementen. Die Eingangsgröße des Systems ist die Spannung u , die Ausgangsgröße ist die Spannung y . Diese Schaltung kann als Serienschaltung zweier RC-Tiefpass-Schaltungen interpretiert werden. Für den RC-Tiefpass erster Ordnung mit der Eingangsgröße u und der Ausgangsgröße

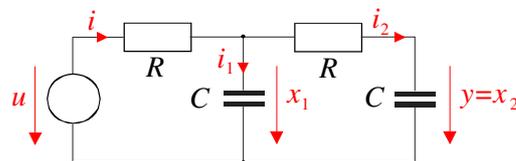


Bild 2.13: Netzwerk, bestehend aus zwei RC-Tiefpass-Schaltungen erster Ordnung

$x = x_1$ gilt

$$u = Ri + x = RC \frac{dx}{dt} + x$$

und es kann die Übertragungsfunktion

$$G(s) = \frac{\bar{x}(s)}{\bar{u}(s)} = \frac{1}{RC} \frac{1}{s + \frac{1}{RC}}$$

ermittelt werden. Die Anwendung der Relation (2.45) über die Serienschaltung zweier Übertragungssysteme liefert dann die (vermeintliche) Übertragungsfunktion der Schaltung aus Bild

2.13, nämlich

$$H(s) = \frac{\bar{y}(s)}{\bar{u}(s)} = G(s)G(s) = \frac{\left(\frac{1}{RC}\right)^2}{\left(s + \frac{1}{RC}\right)^2} = \frac{\left(\frac{1}{RC}\right)^2}{s^2 + \frac{2}{RC}s + \left(\frac{1}{RC}\right)^2}. \quad (2.49)$$

Zur Kontrolle wird nun die Übertragungsfunktion der Schaltung aus Bild 2.13 direkt berechnet. Es gilt

$$x_1 = Ri_2 + x_2 = RC \frac{dx_2}{dt} + x_2$$

sowie

$$u = Ri + x_1 = RC \frac{dx_1}{dt} + RC \frac{dx_2}{dt} + x_1.$$

Aus diesen beiden Relationen ergibt sich das mathematische Modell

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} -\frac{2}{RC} & \frac{1}{RC} \\ \frac{1}{RC} & -\frac{1}{RC} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{RC} \\ 0 \end{bmatrix} u, \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

Die zugehörige Übertragungsfunktion lautet

$$T(s) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} s + \frac{2}{RC} & -\frac{1}{RC} \\ -\frac{1}{RC} & s + \frac{1}{RC} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{RC} \\ 0 \end{bmatrix} = \frac{1}{s^2 + s\frac{3}{RC} + \frac{1}{(RC)^2}}.$$

Vergleicht man dieses Ergebnis mit (2.49), so erkennt man, dass

$$T(s) \neq H(s),$$

d.h. die Formel (2.45) darf hier offenbar nicht angewandt werden! Der Grund hierfür liegt darin, dass (2.49) impliziert, dass zwei unbelastete(!) RC-Tiefpassschaltungen in Serie geschaltet sind. Dies trifft aber bei der betrachteten Schaltung *nicht* zu. Wird zur „Entkopplung“ der beiden Tiefpass-Schaltungen ein Spannungsfolger (Impedanzwandler) eingesetzt, wie in Bild 2.14 dargestellt, so ist die vorausgesetzte Rückwirkungsfreiheit gegeben und es gilt

$$T(s) = H(s).$$

2.12 Wichtige Übertragungsglieder

An dieser Stelle werden einige wichtige Übertragungsglieder vorgestellt. In den Blöcken, welche die einzelnen Übertragungssysteme in Strukturbildern repräsentieren, werden oft auch die zugehörigen Sprungantworten graphisch angedeutet.

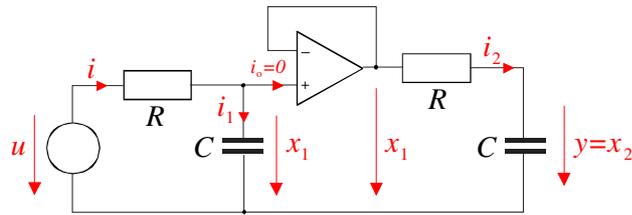


Bild 2.14: Modifikation des Schaltung durch Einbau eines Spannungsfolgers

2.12.1 Proportionalglied

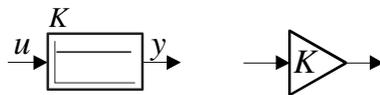
Beim Proportionalglied (**P-Glied**) ist die Ausgangsgröße y proportional zur Eingangsgröße u , d.h.

$$y(t) = K u(t) \quad (2.50)$$

Den (reellen) Proportionalitätsfaktor K bezeichnet man auch als Verstärkung oder Verstärkungsfaktor. Aus (2.50) folgt unmittelbar die zugehörige Übertragungsfunktion

$$G(s) = K. \quad (2.51)$$

In Strukturbildern repräsentieren üblicherweise die folgenden Symbole ein Proportionalglied:



Wie bereits erwähnt, ist im linken Block die Sprungantwort des P-Gliedes angedeutet.

2.12.2 Verzögerungsglied erster Ordnung

Das Übertragungsverhalten eines Verzögerungsgliedes erster Ordnung (**VZ₁-Glied** oder **PT₁-Glied**) mit der Eingangsgröße u und der Ausgangsgröße y wird durch die Differentialgleichung

$$T \frac{dy}{dt} + y(t) = K u(t) \quad (2.52)$$

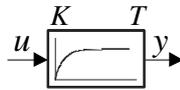
beschrieben. Hierbei ist der positive Parameter T die so genannte Zeitkonstante des Systems und K ist der Verstärkungsfaktor. Die Anwendung der Laplace-Transformation liefert die Übertragungsfunktion

$$G(s) = \frac{K}{1 + sT}. \quad (2.53)$$

Für die Sprungantwort eines Verzögerungsgliedes erster Ordnung gilt

$$h(t) = K \left(1 - e^{-\frac{t}{T}} \right), \quad (2.54)$$

siehe auch Bild 2.15. Daher werden solche Systeme üblicherweise durch folgendes Symbol dargestellt:



Aus Bild 2.15 geht auch hervor, dass zum Zeitpunkt $t = T$ der Wert der Sprungantwort ca. 63% des stationären Endwertes $h_\infty = K$ beträgt. Weiters kann aus dem dargestellten Verlauf von $h(t)$ die Zeitkonstante T als derjenige Zeitpunkt abgelesen werden, bei dem die rot eingezeichnete Tangente an $h(t)$ den Wert $h_\infty = K$ annimmt, d.h. es gilt

$$T = \frac{K}{\left. \frac{dh}{dt} \right|_{t=0}}.$$

In der Praxis wird häufig auf die Faustformel zurückgegriffen, die besagt, dass der stationäre

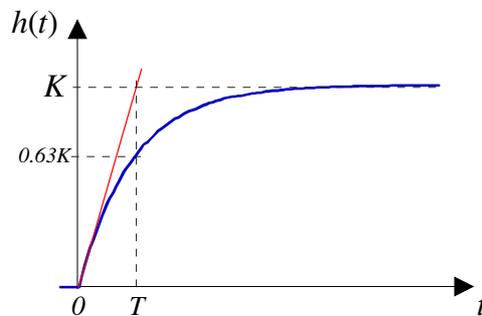


Bild 2.15: Sprungantwort eines PT_1 Elementes mit der Zeitkonstante T

Endwert h_∞ für $t \geq 5T$ erreicht ist. Durch Einsetzen in (2.54) erkennt man, dass der durch diese Faustformel entstehende Fehler kleiner als 1% von h_∞ ist.

2.12.3 Verzögerungsglied zweiter Ordnung

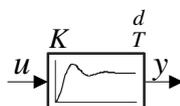
Verzögerungsglieder zweiter Ordnung (**VZ₂-G**lied oder **PT₂-G**lied) werden durch die Differentialgleichung

$$T^2 \frac{d^2 y}{dt^2} + 2dT \frac{dy}{dt} + y(t) = Ku(t) \quad (2.55)$$

beschrieben. Hierbei sind die positiven Parameter d und T der Dämpfungsfaktor und die Zeitkonstante, K ist der Verstärkungsfaktor. Aus obiger Differentialgleichung ergibt sich die Übertragungsfunktion

$$G(s) = \frac{K}{1 + 2dT s + (sT)^2}. \quad (2.56)$$

Für $d < 1$ besitzt $G(s)$ ein konjugiert komplexes Polpaar, was zu dem im Block



angedeuteten „oszillierenden“ Verlauf der Sprungantwort führt. Für $d \geq 1$ besitzt $G(s)$ zwei reelle Pole und kann somit als Reihenschaltung zweier Verzögerungsglieder erster Ordnung interpretiert werden.

2.12.4 Integrator

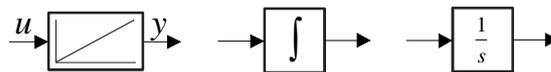
Der Integrator (**I-Glied**) wird durch die Differentialgleichung

$$\frac{dy}{dt} = u \quad (2.57)$$

beschrieben, für die zugehörige Übertragungsfunktion gilt

$$G(s) = \frac{1}{s}. \quad (2.58)$$

Folgende Symbole



werden in Strukturbildern für den Integrator verwendet, wobei im linken Block die Rampe als Sprungantwort des Systems dargestellt ist

2.12.5 Differenzierer

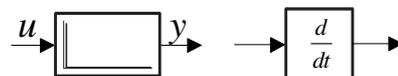
Ein idealer Differenzierer (**D-Glied**) bildet die zeitliche Ableitung des Eingangssignales, d.h.

$$y = \frac{du}{dt}, \quad (2.59)$$

für die (nicht realisierbare) Übertragungsfunktion folgt

$$G(s) = s. \quad (2.60)$$

Üblicherweise werden in Strukturbildern die Symbole



verwendet, wobei im linken Block der Dirac-Impuls als Sprungantwort des Systems angedeutet ist.

2.12.6 Vorhalteglied

Das dynamische Verhalten des realen Differenzierers (**DT₁-Glied**), auch Vorhalteglied bezeichnet, wird durch die Differentialgleichung

$$T \frac{dy}{dt} + y(t) = \frac{du}{dt} \quad (2.61)$$

beschrieben. Hierbei ist der positive Parameter T die Zeitkonstante. Für die Übertragungsfunktion des Systems ergibt sich

$$G(s) = \frac{s}{1 + sT}. \quad (2.62)$$

Beim DT_1 -Glied handelt es sich also um ein D-Glied mit Verzögerungsverhalten erster Ordnung. Seine Sprungantwort lautet

$$h(t) = \frac{1}{T} e^{-\frac{t}{T}},$$

sie ist in Bild 2.16 graphisch dargestellt. Wie man leicht überprüfen kann, gilt unabhängig von

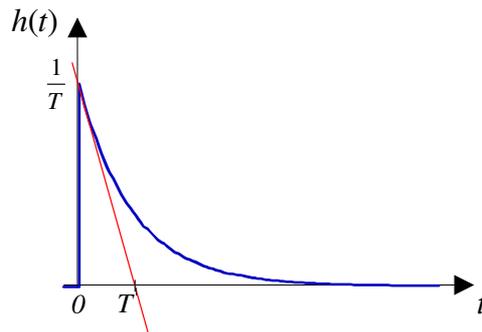


Bild 2.16: Sprungantwort DT_1 -Glied

T die Relation $\int_0^\infty h(t) dt = 1$. Daraus kann man folgern, dass für $T \rightarrow 0$ die Sprungantwort $h(t)$ dem Dirac-Impuls $\delta(t)$ entspricht, das DT_1 -Glied geht dann in einen idealen Differenzierer über.

2.13 Frequenzgang

Der so genannte Frequenzgang beschreibt den eingeschwungenen Zustand eines linearen, zeitinvarianten Systems bei harmonischer Erregung. Dabei wird hier vorausgesetzt, dass das betrachtete System BIBO-stabil ist, d.h. seine Übertragungsfunktion $G(s)$ besitzt ausschließlich Pole mit negativem Realteil. Weiters wird vereinfachend vorausgesetzt, dass alle Pole von $G(s)$ *einfach* sind, die gefundenen Ergebnisse gelten allerdings auch für den allgemeinen Fall mehrfacher Pole. Als Eingangsgröße $u(t)$ wird nun die komplexe harmonische Funktion

$$u(t) = e^{j\omega t} \quad \bullet \quad \bar{u}(s) = \frac{1}{s - j\omega} \quad (2.63)$$

gewählt⁸. Für die zugehörige Ausgangsgröße $y(t)$ gilt im Bildbereich

$$\bar{y}(s) = G(s) \frac{1}{s - j\omega} = K \frac{\prod_{i=1}^m (s - n_i)}{\prod_{k=1}^n (s - p_k)} \frac{1}{s - j\omega},$$

⁸Man beachte, dass es sich hier um ein reines Gedankenexperiment handelt.

eine Partialbruchzerlegung liefert

$$\bar{y}(s) = \frac{c_0}{s - j\omega} + \sum_{i=1}^n \frac{c_i}{s - p_i}.$$

Für den Koeffizienten c_0 findet man

$$c_0 = \lim_{s \rightarrow j\omega} (s - j\omega) \bar{y}(s) = \lim_{s \rightarrow j\omega} \left[(s - j\omega) G(s) \frac{1}{(s - j\omega)} \right] = G(j\omega),$$

d.h.

$$\bar{y}(s) = \frac{G(j\omega)}{s - j\omega} + \sum_{i=1}^n \frac{c_i}{s - p_i}.$$

Für die zugehörige Zeitfunktion gilt nun

$$y(t) = G(j\omega) e^{j\omega t} + \sum_{i=1}^n c_i e^{p_i t}.$$

Für „sehr große Werte von t “, also im so genannten *eingeschwungenen Zustand* gilt

$$y(t) \approx G(j\omega) e^{j\omega t}, \quad (2.64)$$

d.h. die Ausgangsgröße entspricht der mit $G(j\omega)$ gewichteten Eingangsgröße. Diese komplexe Funktion

$$G(j\omega) = |G(j\omega)| e^{j \arg G(j\omega)} \quad (2.65)$$

ist der **Frequenzgang** des Systems. Salopp formuliert beschreibt er, wie sich im eingeschwungenen Zustand die Amplitude und die Phasenlage der harmonischen Eingangsgröße beim Durchlaufen des Systems verändern. Man beachte, dass die Eingangsgröße (2.63) als Linearkombination einer Sinus- und einer Cosinusfunktion dargestellt werden kann, d.h.

$$u(t) = e^{j\omega t} = \cos \omega t + j \sin \omega t.$$

Nach (2.64) und (2.65) gilt im eingeschwungenen Zustand für die zugehörige Ausgangsgröße

$$y(t) = |G(j\omega)| e^{j(\omega t + \arg G(j\omega))} = |G(j\omega)| \cos(\omega t + \arg G(j\omega)) + j |G(j\omega)| \sin(\omega t + \arg G(j\omega)).$$

Aus der Linearität der Laplace-Transformation und des Systems kann daher unmittelbar gefolgert werden, dass im eingeschwungenen Zustand gilt:

$$u(t) = \hat{u} \cos(\omega t + \phi) \quad \Rightarrow \quad y(t) = \hat{u} |G(j\omega)| \cos(\omega t + \phi + \arg G(j\omega))$$

$$u(t) = \hat{u} \sin(\omega t + \phi) \quad \Rightarrow \quad y(t) = \hat{u} |G(j\omega)| \sin(\omega t + \phi + \arg G(j\omega))$$

B Gegeben sei das System mit der Übertragungsfunktion

$$G(s) = \frac{1}{s + 1}.$$

Für eine Eingangsgröße $u(t) = 3 \sin t$ ergibt sich für den Verlauf von $y(t)$ im eingeschwungenen Zustand

$$y(t) \approx 3 |G(j)| \sin(t + \arg G(j)) = \frac{3}{\sqrt{2}} \sin\left(t - \frac{\pi}{4}\right).$$

■

Für jeden Wert von ω ergibt sich also eine komplexe Zahl $G(j\omega)$. Die graphische Darstellung dieser komplexen Zahlen in der komplexen Ebene als Funktion von ω wird **Frequenzgangs-Ortskurve**, **Nyquist-Ortskurve**⁹ oder kurz **Ortskurve** genannt. In den so genannten (logarithmischen) **Frequenzkennlinien** werden jeweils Betrag und Phase von $G(j\omega)$ über ω dargestellt, man spricht in diesem Zusammenhang auch von den **Bode**¹⁰-**Diagrammen**.

⁹benannt nach dem (in Schweden geborenen) amerikanischen Physiker Harry Nyquist (1889-1976).

¹⁰benannt nach dem amerikanischen Wissenschaftler Hendrik Wade Bode (1905-1982).

Kapitel 3

Zeitdiskrete Systeme

3.1 Einführung

Bei den bisherigen Betrachtungen wurde davon ausgegangen, dass die zeitlich veränderlichen Systemgrößen Funktionen des kontinuierlichen Zeitparameters $t \in \mathbb{R}$ sind. Man spricht in diesem Zusammenhang auch von **zeitkontinuierlichen Systemen**. Im Gegensatz dazu wird bei einem **zeitdiskreten System** die Zeitabhängigkeit von Systemgrößen durch einen ganzzahligen Index $k \in \mathbb{Z}$ charakterisiert. Die zeitlichen Verläufe der zeitabhängigen Systemgrößen werden durch Zahlenfolgen beschrieben. Der Wert der Systemgrößen wird also zu diskreten Zeitpunkten betrachtet. In vielen Fällen handelt es sich um äquidistante Zeitpunkte, d.h. der Abstand zweier aufeinanderfolgender Zeitpunkte ist konstant. Das zeitliche Verhalten von zeitdiskreten System kann mit Hilfe von **Differenzgleichungen** beschrieben werden. Begriffe, wie Linearität und Zeitinvarianz können vom zeitkontinuierlichen Fall geradlinig auf zeitdiskrete Systeme übertragen werden.

In weiterer Folge werden die Betrachtungen auf lineare, zeitinvariante Eingrößensysteme, die durch lineare Differenzgleichungen der Ordnung n mit konstanten Koeffizienten der Form

$$\alpha_n y_{k+n} + \alpha_{n-1} y_{k+n-1} + \dots + \alpha_0 y_k = \beta_m u_{k+m} + \beta_{m-1} u_{k+m-1} + \dots + \beta_0 u_k \quad (3.1)$$

beschrieben werden können, beschränkt. Dabei ist $(u_k) = (u_0, u_1, \dots)$ die Eingangsfolge und $(y_k) = (y_0, y_1, \dots)$ ist die zugehörige Ausgangsfolge. Es wird vorausgesetzt, dass durch den Index k Zeitpunkte beschrieben werden, die ganzzahligen Vielfachen einer konstanten **Diskretisierungszeit** T_d entsprechen.

3.2 z -Übertragungsfunktion

Die z -Übertragungsfunktion $G(z)$ beschreibt das Übertragungsverhalten von zeitdiskreten linearen, zeitinvarianten Systemen. Sie ist definiert als der Quotient der z -Transformierten von Ausgangs- und Eingangsgröße, d.h.

$$G(z) = \frac{\mathfrak{Z}\{(y_k)\}}{\mathfrak{Z}\{(u_k)\}} = \frac{\tilde{y}(z)}{\tilde{u}(z)}. \quad (3.2)$$

Dabei wird vorausgesetzt, dass sich das System zum Zeitpunkt $k = 0$ in Ruhe befindet, d.h. in (3.1) sind alle Anfangswerte Null zu setzen. Im Rahmen der vorliegenden Betrachtungen handelt es sich bei der z -Übertragungsfunktion stets um eine gebrochen rationale Funktion, d.h. $G(z)$ kann als Quotient zweier Polynome in der Form

$$G(z) = \frac{\beta(z)}{\alpha(z)}, \quad (3.3)$$

dargestellt werden, wobei die Polynomkoeffizienten von $\alpha(z)$ und $\beta(z)$ als reell vorausgesetzt werden. Systeme, bei denen die Polynomgrade der Bedingung

$$\text{Grad } \beta(z) \leq \text{Grad } \alpha(z) \quad (3.4)$$

genügen, werden als realisierbar bzw. **kausal** bezeichnet. Im Falle

$$\text{Grad } \beta(z) = \text{Grad } \alpha(z)$$

spricht man von einem sprungfähigen System.

Die Anwendung des Verschiebungssatzes der z -Transformation auf (3.1) führt auf die z -Übertragungsfunktion

$$G(z) = \frac{\mathfrak{Z}\{(y_k)\}}{\mathfrak{Z}\{(u_k)\}} = \frac{\tilde{y}(z)}{\tilde{u}(z)} = \frac{\beta_m z^m + \beta_{m-1} z^{m-1} + \dots + \beta_1 z + \beta_0}{\alpha_n z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_1 z + \alpha_0}.$$

Die Kausalität ist gemäß (3.4) für $m \leq n$ gegeben.

B Gegeben sei die Differenzgleichung

$$y_{k+2} - y_{k+1} + y_k = u_{k+1} - u_k.$$

Die entsprechende z -Übertragungsfunktion lautet

$$G(z) = \frac{z - 1}{z^2 - z + 1}.$$

■

Oft sind Differenzgleichungen so angeschrieben, dass man den Wert der Ausgangsgröße zum Zeitpunkt k aus *vergangenen* Werten der Ausgangsgröße und dem Verlauf der Eingangsfolge berechnen kann. Die Ermittlung der zugehörigen z -Übertragungsfunktion kann natürlich analog erfolgen.

B Gegeben sei die Differenzgleichung

$$y_k = y_{k-1} - y_{k-2} + u_{k-1} - u_{k-2}.$$

Die Anwendung der z -Transformation liefert

$$\tilde{y}(z) = z^{-1}\tilde{y}(z) - z^{-2}\tilde{y}(z) + z^{-1}\tilde{u}(z) - z^{-2}\tilde{u}(z).$$

Daraus ergibt sich unmittelbar

$$G(z) = \frac{\tilde{y}(z)}{\tilde{u}(z)} = \frac{z^{-1} - z^{-2}}{1 - z^{-1} + z^{-2}} = \frac{z - 1}{z^2 - z + 1},$$

das System besitzt also die gleiche z-Übertragungsfunktion wie im vorigen Beispiel. ■

B Gegeben sei die Differenzgleichung

$$y_k = y_{k-1} + u_{k+1}.$$

Hierbei handelt es sich offenbar um ein nicht-kausales System, denn zur Ermittlung der Ausgangsgröße zum Zeitpunkt k wird der *zukünftige* Wert der Eingangsgröße benötigt. Dies manifestiert sich in der z-Übertragungsfunktion

$$G(z) = \frac{\tilde{y}(z)}{\tilde{u}(z)} = \frac{z^2}{z - 1},$$

durch die Verletzung der Kausalitätsbedingung (3.4). ■

Die Ermittlung einer Differenzgleichung aus einer z-Übertragungsfunktion beruht ebenso auf der Anwendung des Verschiebungssatzes, wie das nachfolgende Beispiel demonstriert.

B Gegeben sei die Übertragungsfunktion

$$G(z) = \frac{\tilde{y}(z)}{\tilde{u}(z)} = \frac{z + 0.2}{z^2 + 0.1z + 0.15}.$$

Daraus folgt

$$\tilde{y}(z) (z^2 + 0.1z + 0.15) = \tilde{u}(z) (z + 0.2),$$

woraus sich die Differenzgleichung

$$y_{k+2} + 0.1y_{k+1} + 0.15y_k = u_{k+1} + 0.2u_k$$

bzw.

$$y_k = -0.1y_{k-1} - 0.15y_{k-2} + u_{k-1} + 0.2u_{k-2}$$

ergibt. ■

3.3 Pole und Nullstellen

Man beachte, dass in (3.3) die Polynome $\alpha(z)$ und $\beta(z)$ nicht notwendigerweise teilerfremd sind. Das bedeutet, dass unter Umständen Kürzungen durchgeführt werden können, wie das folgende Beispiel veranschaulicht.

B Gegeben sei die Übertragungsfunktion

$$G(z) = \frac{\beta(z)}{\alpha(z)} = \frac{z + 0.1}{z^2 + 0.3z + 0.02}.$$

Gibt man $G(z)$ in der faktorisierten Darstellung an, d.h.

$$G(z) = \frac{z + 0.1}{(z + 0.1)(z + 0.2)} = \frac{1}{z + 0.2}$$

so erkennt man, dass eine Kürzung möglich ist. ■

Geht man nun von einer aus teilerfremden Polynomen gebildeten, d.h. nicht weiter kürzbaren, Übertragungsfunktion

$$G(z) = \frac{\mu(z)}{\nu(z)} \quad \text{mit} \quad \mu(z), \nu(z) \dots \text{teilerfremd} \quad (3.5)$$

aus, so sind diejenigen Werte von z , für die $\mu(z) = 0$ gilt, die so genannten **Nullstellen** von $G(z)$. Die **Pole** oder **Polstellen** der Übertragungsfunktion sind durch $\nu(z) = 0$ charakterisiert. In der *faktorisierten* Darstellung der Übertragungsfunktion

$$G(z) = K \frac{\prod_{i=1}^m (z - n_i)}{\prod_{k=1}^n (z - p_k)} \quad \text{mit} \quad m \leq n \quad (3.6)$$

ist $z = n_i$ eine Nullstelle und $z = p_k$ eine Polstelle von $G(z)$. Tritt der Faktor $(z - n_i)$ bzw. $(z - p_k)$ mehrfach auf, so besitzt das System eine mehrfache Nullstelle bzw. einen mehrfachen Pol. Da die Koeffizienten der Polynome $\mu(z)$ und $\nu(z)$ reell sind, treten Pole bzw. Nullstellen reell und/oder paarweise konjugiert komplex auf. Aus (3.6) folgt unmittelbar

$$G(z) = 0 \quad \text{für} \quad z = n_i \quad \text{und} \quad |G(z)| \rightarrow \infty \quad \text{für} \quad z \rightarrow p_k. \quad (3.7)$$

3.4 BIBO-Stabilität

Ein zeitdiskretes lineares, zeitinvariantes System mit der z -Übertragungsfunktion $G(z)$ wird **BIBO-stabil** („bounded input - bounded output“) genannt, wenn es auf *jede* beschränkte Eingangsfolge mit einer beschränkten Ausgangsfolge reagiert. Unter einer beschränkten Folge versteht man in diesem Kontext eine Folge, deren Elemente betragsmäßig beschränkt sind. Somit muss bei BIBO-Stabilität

$$|u_k| \leq M < \infty \quad \Rightarrow \quad |y_k| \leq N < \infty \quad \forall k \geq 0. \quad (3.8)$$

gelten. Aus

$$\tilde{y}(z) = G(z)\tilde{u}(z) \quad \text{d.h.} \quad y_k = \sum_{i=0}^k g_{k-i} u_i \quad \text{wobei} \quad G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$$

folgt nun unter Annahme einer beschränkten Eingangsgröße

$$|y_k| = \left| \sum_{i=0}^k g_{k-i} u_i \right| \leq \sum_{i=0}^k |g_{k-i}| |u_i| \stackrel{(3.8)}{\leq} M \sum_{i=0}^k |g_{k-i}| = M \sum_{i=0}^k |g_i|.$$

Offensichtlich tritt der größtmögliche Wert von $|y_k|$ für $k \rightarrow \infty$ auf, d.h. zur Erfüllung von (3.8) muss

$$M \sum_{i=0}^{\infty} |g_i| \leq N$$

gelten. Das bedeutet, dass die Impulsantwort (g_k) absolut summierbar sein muss, d.h.

$$\sum_{k=0}^{\infty} |g_k| < \infty. \quad (3.9)$$

Man beachte, dass es sich bei (3.9) um eine notwendige und hinreichende Bedingung handelt!

B Gegeben sei die z-Übertragungsfunktion

$$G(z) = \frac{z}{z-p} \quad \text{mit} \quad p \in \mathbb{R},$$

mit der zugehörigen Impulsantwort

$$(g_k) = (p^k).$$

Die Summe der Absolutbeträge der Elemente von (g_k) , also

$$\sum_{k=0}^{\infty} |g_k| = \sum_{k=0}^{\infty} |p^k| = \sum_{k=0}^{\infty} |p|^k$$

existiert offensichtlich nur für Werte von p , die betragsmäßig kleiner als eins sind. Daraus kann gefolgert werden, dass das betrachtete System genau dann BIBO-stabil ist, wenn sein Pol $z = p$ einen Betrag besitzt, der kleiner als eins ist. ■

Die Erkenntnis aus obigem Beispiel kann auf den allgemeinen Fall ausgeweitet werden. Um dies zu zeigen wird zunächst der Zusammenhang zwischen der Impulsantwort (g_k) und der z-Übertragungsfunktion betrachtet, d.h.

$$G(z) = \mathfrak{Z}\{(g_k)\} = \sum_{k=0}^{\infty} g_k z^{-k}.$$

Daraus kann folgende Abschätzung abgeleitet werden

$$|G(z)| = \left| \sum_{k=0}^{\infty} g_k z^{-k} \right| \leq \sum_{k=0}^{\infty} |g_k| |z^{-k}| = \sum_{k=0}^{\infty} |g_k| |z|^{-k}$$

Gilt $|z| \geq 1$, so kann weiter geschrieben werden

$$|G(z)| \leq \sum_{k=0}^{\infty} |g_k|.$$

Das bedeutet, dass im Falle eines BIBO-stabilen Systems - (3.9) ist also erfüllt - der Ausdruck $|G(z)|$ für $|z| \geq 1$ *endliche* Werte annimmt. Daraus kann gefolgert werden, dass $G(z)$ ausschließlich Pole *im Inneren des Einheitskreises* besitzt.

Ein System mit der Übertragungsfunktion $G(z)$ ist somit genau dann BIBO-stabil, wenn *alle* Pole von $G(z)$ im Einheitskreis liegen. Das bedeutet, dass in der gekürzten Darstellung (3.5) das Nennerpolynom von $G(z)$ ein Einheitskreispolynom sein muss.

B Gegeben sei ein System mit der Übertragungsfunktion

$$G(z) = \frac{z + 0.1}{z(z - 0.2)(z + 1)}.$$

Das System ist nicht BIBO-stabil, da es einen Pol besitzt, der nicht im Einheitskreis liegt. ■

B Gegeben sei ein System mit der Übertragungsfunktion

$$G(z) = \frac{z - 1}{(z + 0.1)^2(z - 0.5)(z + 0.9)}.$$

Das System ist BIBO-stabil, da alle Pole von $G(z)$ im Einheitskreis liegen. ■

Teil II

Regelungstechnik

3.5 Einführung

In allen Bereichen des täglichen Lebens wird vorausgesetzt, dass sich Systeme bzw. Prozesse genau so verhalten, wie man es von ihnen erwartet. Flugzeuge sind in der Lage, vollautomatisch Kurs und Höhe zu halten und können komplizierte Manöver völlig ohne menschliche Einflussnahme fehlerfrei ausführen. Industrieroboter führen mit höchster Präzision und Wiederholgenauigkeit komplexe Bewegungsabläufe aus und verrichten dabei Schweiß- oder Schneidarbeiten. In Geschäftslokalen herrscht, völlig unabhängig von den äußeren Witterungsbedingungen, ein angenehmes Raumklima. Leistungsfähige Assistenzsysteme, wie z.B. Spurhaltesysteme oder Tempomat verbessern Sicherheit und Komfort von Fahrzeugen, autonom bzw. automatisiert fahrende Autos sind keine Zukunftsmusik mehr.

In den angegebenen Beispielen wird das Verhalten der Systeme „Flugzeug“, „Roboter“, „Geschäftslokal“ und „Fahrzeug“ über vorgebbare Größen, wie z. B. die Ruderstellungen beim Flugzeug, in gewünschter Weise beeinflusst. Die hierfür erforderlichen Ruderstellungen werden aus der messtechnisch erfassten Abweichung des Flugzeugs von seinem vorgegebenen Kurs berechnet. Es entsteht also ein **geschlossener Wirkungskreis**, in dem das „Ist-Verhalten“ fortlaufend mit dem „Soll-Verhalten“ verglichen wird und das System im Sinne einer Angleichung an das „Soll-Verhalten“ beeinflusst wird. Dieses Prinzip der **Rückkopplung** ist charakteristisch für eine **Regelung**. Im Gegensatz dazu zeichnet sich eine **Steuerung** durch eine **offene Wirkungskette** aus, bei der ein System auf eine vom aktuellen Systemverhalten unabhängige Weise beeinflusst wird. Ein typisches Beispiel hierfür ist ein Bewässerungssystem, bei dem Pflanzen völlig unabhängig von der Witterung nach einem zeitlich fest vorgegebenem Schema mit Wasser versorgt werden.

3.6 Begriffe der Regelungstechnik

In Bild 3.1 ist eine sehr allgemeine Darstellung eines **Regelkreises** zu sehen. Das zu bee-

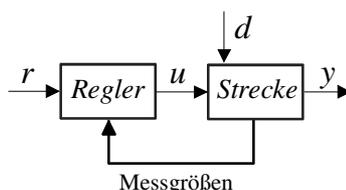


Bild 3.1: Regelkreis, allgemeine Darstellung

influssende System wird **Regelstrecke** oder kurz **Strecke** genannt. Diejenigen Größen der Regelstrecke, deren Verläufe kontrolliert werden sollen, sind die so genannten **Regelgrößen**. Diese werden als messbar vorausgesetzt. Die vorgebbaren Eingangsgrößen der Regelstrecke, also diejenigen Größen, die eine Beeinflussung der Regelgrößen erlauben, sind die so genannten **Stellgrößen**. Im Rahmen dieses Skriptums wird vorausgesetzt, dass es genau *eine* Stellgröße $u(t)$ und *eine* Regelgröße $y(t)$ gibt, d.h. die Strecke ist ein Eingrößensystem. Der Wunschverlauf der Regelgröße wird in Form der **Führungsgröße** oder **Referenzgröße** $r(t)$ vorgegeben.

Der **Regelfehler** $e(t)$ repräsentiert die Abweichung der Regelgröße von der Führungsgröße, d.h. $e(t) = r(t) - y(t)$. Soll $y(t)$ dem Verlauf von $r(t)$ folgen, so spricht man von einer **Folgeregelung** oder **Nachführung**. Wenn $y(t)$ hingegen auf einem konstanten Wert gehalten werden soll - was natürlich auch als Vorgabe einer konstanten Führungsgröße interpretiert werden kann - handelt es sich um eine **Festwertregelung**. Der **Regler** bzw. das **Regelgesetz** generiert aus gegebenen bzw. messbaren Signalen die Stellgröße $u(t)$ so, dass die Regelgröße $y(t)$ der Führungsgröße $r(t)$ trotz des Einwirkens einer **Störgröße** $d(t)$ „möglichst gut“ entspricht. Diese unscharfe Ausdrucksweise soll verdeutlichen, dass die Güte bzw. Qualität einer Regelung, die so genannte **Regelgüte**, nicht eindeutig definiert werden kann. In späteren Kapiteln wird sich zeigen, dass es verschiedene Möglichkeiten gibt, die Regelgüte sinnvoll zu definieren. Allen Definitionen ist allerdings gemeinsam, dass sie die Stabilität des Regelkreises voraussetzen.

In weiterer Folge wird angenommen, dass die Regelstrecke im relevanten Arbeitsbereich durch ein lineares, zeitinvariantes System der Ordnung n hinreichend genau approximiert werden kann. Das bedeutet, dass der Zusammenhang zwischen y und u bzw. d im Bildbereich durch die Relation

$$\bar{y}(s) = P(s) \bar{u}(s) + P_d(s) \bar{d}(s) \quad (3.10)$$

dargestellt werden kann, siehe hierzu auch Bild 3.2, links. Die beiden Übertragungsfunktionen

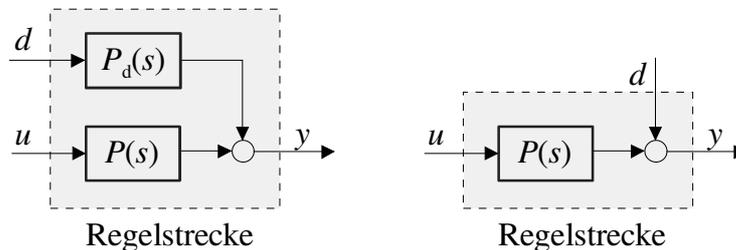


Bild 3.2: Einfluss von d und u auf die Regelgröße y

$P(s)$ und $P_d(s)$ sind gebrochen rationale Funktionen, d.h. sie können als Quotienten von Polynomen in s dargestellt werden. Dabei wird im Folgenden vorausgesetzt, dass die so genannte „Streckenübertragungsfunktion“ $P(s)$, die das Verhalten der störungsfreien Strecke ($d = 0$) beschreibt, nicht sprunghaft ist, d.h. der Nennerpolynomgrad n der Übertragungsfunktion ist größer als der Zählerpolynomgrad. Im Gegensatz dazu wird angenommen, dass d einen *unmittelbaren* Einfluss auf y besitzt, d.h. $P_d(s)$ repräsentiert ein sprunghaftes System, bei dem der Zählerpolynomgrad dem Nennerpolynomgrad entspricht. Sehr häufig wird auch vereinfachend angenommen, dass die Störung d *direkt* auf die Regelgröße y wirkt, d.h. $P_d(s) = 1$, siehe Bild 3.2, rechts. Mit dieser Annahme wird der Tatsache Rechnung getragen, dass der Einfluss einer unbekannt und nicht messbaren „inneren“ Störung ohnehin nur mittels des gemessenen Verlaufs von y erfasst werden kann. Die tatsächliche Störgröße wird also in eine äquivalente Störung am Streckenausgang umgerechnet.

Die Streckenbeschreibung kann auch in Form eines Zustandsmodells der Form

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u + \mathbf{h}d \\ y &= \mathbf{c}^T\mathbf{x} + g d\end{aligned}\quad (3.11)$$

erfolgen, wobei \mathbf{A} eine $n \times n$ Matrix ist, \mathbf{b} , \mathbf{h} und \mathbf{c} sind n -dimensionale Spaltenvektoren und g ist eine reelle Konstante. Der Zustand der Regelstrecke wird durch den Vektor

$$\mathbf{x}(t) = [x_1 \ x \ \dots \ x_n]^T$$

beschrieben. Der Zusammenhang zwischen den Darstellungen (3.10) und (3.11) ist durch die Relation

$$P(s) = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{b} \quad \text{und} \quad P_d(s) = \mathbf{c}^T (s\mathbf{E} - \mathbf{A})^{-1} \mathbf{h} + g \quad (3.12)$$

gegeben.

In Bild 3.3 ist exemplarisch der so genannte **Standardregelkreis** dargestellt, er repräsentiert den klassischen Regelkreis schlechthin. Die Eingangsgröße des Reglers ist hier der Regelfehler, d.h. die Abweichung zwischen gewünschtem und tatsächlichem Verlauf der Regelgröße. Es wird sich später zeigen, dass es oft auch sinnvoll ist, andere Regelkreis-Strukturen zu wählen. Die in weiterer Folge für den Standardregelkreis eingeführten Begriffe sind allgemeingültig, die zugehörigen Berechnungsvorschriften müssen allerdings an die jeweilige Regelkreisstruktur angepasst werden.

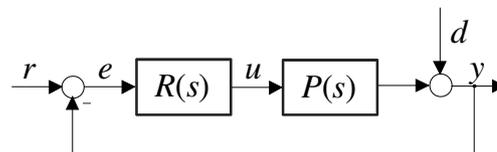


Bild 3.3: Standardregelkreis

Der Regler wird als lineares, zeitinvariantes System angesetzt. Im Fall des Standardregelkreises bedeutet dies, dass der Regler durch eine Übertragungsfunktion

$$R(s) = \frac{\bar{u}(s)}{\bar{e}(s)} \quad (3.13)$$

dargestellt werden kann. Der Regler wird im Allgemeinen als sprunghfähiges System angesetzt, da es wünschenswert ist, dass sich sprunghafte Änderungen am Reglereingang unmittelbar am Reglerausgang auswirken. Der Zusammenhang zwischen Regelgröße y und Führungsgröße r sowie Störgröße d kann im Bildbereich durch die Relation

$$\bar{y}(s) = \frac{R(s)P(s)}{1 + R(s)P(s)} \bar{r}(s) + \frac{1}{1 + R(s)P(s)} \bar{d}(s) \quad (3.14)$$

ausgedrückt werden. Wie man erkennen kann, wird das **Führungsverhalten** durch die so genannte **Führungsübertragungsfunktion**

$$T(s) = \left. \frac{\bar{y}(s)}{\bar{r}(s)} \right|_{d=0} = \frac{R(s)P(s)}{1 + R(s)P(s)} \quad (3.15)$$

beschrieben, während das **Störverhalten** durch die **Störübertragungsfunktion**

$$S(s) = \left. \frac{\bar{y}(s)}{\bar{d}(s)} \right|_{r=0} = \frac{1}{1 + R(s)P(s)} \quad (3.16)$$

charakterisiert ist.

3.7 Beispiele

3.7.1 Drehzahlregelung mit Hilfe des Fliehkraftreglers

Im Jahr 1788 wurde von Watt und Boulton¹ der in Bild 3.4 dargestellte Fliehkraftregler zur Drehzahlregelung von Dampfmaschinen eingesetzt. Ursprünglich wurde dieses Regelungskonzept zur Drehzahlregelung bei Windmühlen mittels mechanischer Flügelverstellung entwickelt. Der Mechanismus besteht im Wesentlichen aus einem Fliehkraftpendel durch dessen Bewegung die Dampfzufuhr zur Dampfmaschine über einen Hebel verstellt wird. Im Ruhezustand ist die Dampfleitung vollständig geöffnet, mit der Kolbenbewegung der Dampfmaschine beginnt sich das Fliehkraftpendel zu drehen. Mit sich ändernder Drehzahl n werden die zwei rotierenden Gewichte aufgrund der Fliehkraft angehoben bzw. abgesenkt. Über einen Hebelmechanismus wird in der Dampfleitung eine Drossel (Klappe oder Schieber) betätigt, bis sich ein Gleichgewichtszustand von Dampfzufuhr und Drehzahl einstellt. Die Struktur des Regelkreises ist in Bild 3.5 dargestellt. Der Fliehkraftregler, bestehend aus Fliehkraftpendel und Hebelmechanismus fungiert hier als Drehzahlsensor und Regler, die auf diese Weise generierte Stellgröße wird durch das so genannte Stellglied, im vorliegenden Fall durch den Schieber mechanisch umgesetzt. Der gewünschte Soll-Wert für die Drehzahl kann über die Verschiebung des Angelpunktes am Hebel eingestellt werden.

3.7.2 Der Pupillenapparat des Menschen

Der so genannte Pupillarapparat des Menschen hat die Aufgabe, die Intensität des auf die Netzhaut des Auges einfallenden Lichtes konstant zu halten². Dies wird durch eine gezielte Verstellung der Pupillenfläche mittels der Irismuskulatur erreicht. Fotorezeptoren in der Netzhaut fungieren als Sensoren, die nicht nur die Lichtintensität, sondern auch deren zeitliche Änderung detektieren. Aus den Sensorinformationen werden im Zentralnervensystem (ZNS) entsprechende Maßnahmen zur Verstellung der Pupillenfläche abgeleitet. Bei der Pupillenregelung handelt es sich somit um eine Festwertregelung, Lichtintensitätsschwankungen in der

¹James Watt (1736 - 1819), Matthew Boulton (1728 - 1809)

²bei besonders starken Intensitätsschwankungen erfolgt eine zusätzliche Adaption, die auf chemischen Mechanismen beruht.

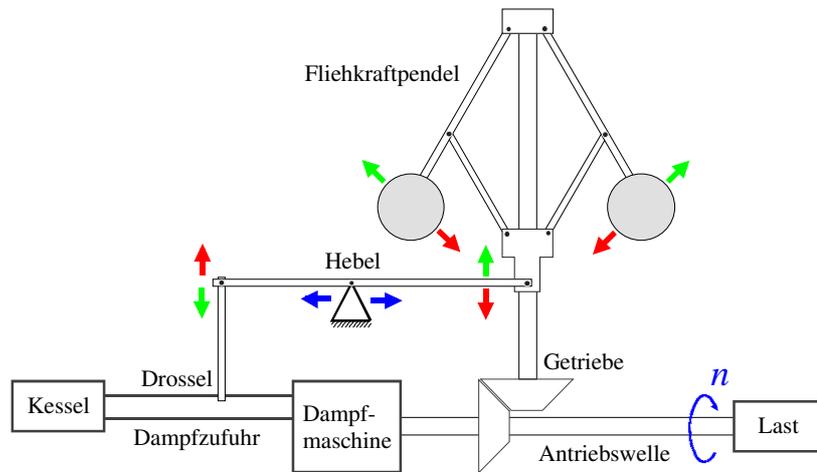


Bild 3.4: Drehzahlregelkreis mit Flihkraftregler

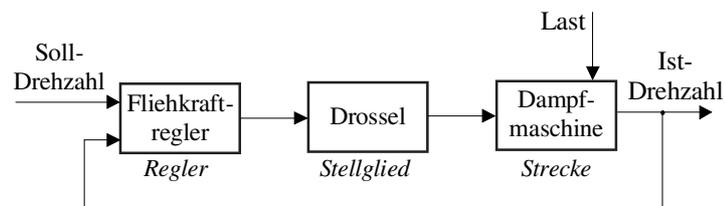


Bild 3.5: Struktur der Drehzahlregelung

Außenwelt wirken als Störungen. Die Struktur dieses wichtigen biologischen Regelkreises ist in Bild 3.6 dargestellt.

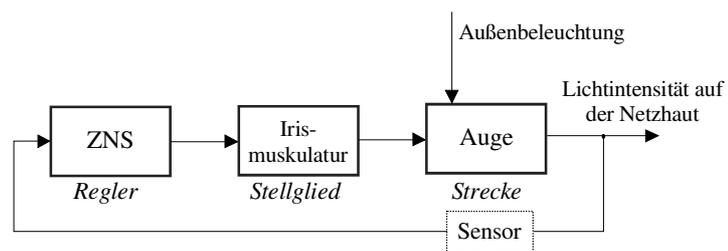


Bild 3.6: Pupillarapparat als Regelkreis

Kapitel 4

Standardregler

4.1 Einführung

PID-Regler sind dadurch charakterisiert, dass sich ihre Wirkung aus drei Komponenten zusammensetzt, nämlich einem proportionalen Anteil („P-Anteil“), einem integrierenden Anteil („I-Anteil“) sowie einem differenzierenden Anteil („D-Anteil“). Aufgrund ihrer strukturellen Einfachheit (und auch aus traditionellen Gründen) sind PID-Regler im industriellen Einsatz oft die erste - wenngleich in vielen Fällen nicht die beste - Wahl. Die Einstellung der Regler erfolgt meistens empirisch, unter Zuhilfenahme von Expertenwissen, oder auf Basis so genannter Einstellregeln, also „Kochrezepten“ zur Reglerauslegung.

In den folgenden Abschnitten wird davon ausgegangen, dass die zu regelnde Strecke ein Eingrößensystem ist, die Regelkreisstruktur ist der Standardregelkreis. Zunächst werden die Grundlagen der PID-Regelung erläutert, einige wichtige Begriffe erklärt und verschiedene Realisierungsformen vorgestellt. Danach wird der so genannte Windup-Effekt beschrieben und einfache Gegenmaßnahmen werden vorgestellt. Den Abschluss des Kapitels bilden einige Einstellregeln, mit deren Hilfe geeignete Reglerparameter ermittelt werden können.

4.2 Parallelrealisierung - die Lehrbuchform

Diese Reglerform zeichnet sich dadurch aus, dass die Anteile (P, I und D) durch drei parallele Zweige realisiert sind (siehe Bild 4.1), man spricht daher von der Parallelrealisierung eines PID-Reglers. Da in dieser Form Modifikationen bzw. Erweiterungen fehlen, die in industriellen PID-Reglern zu finden sind, spricht man auch von der Lehrbuchform des PID-Reglers.

Die Stellgröße $u(t)$ setzt sich aus drei Komponenten zusammen, die proportional zum Regelfehler $e(t)$, zum zeitlichen Integral über den Regelfehler sowie zur zeitlichen Ableitung des Regelfehlers sind. Es gilt somit

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau + K_D \frac{de(t)}{dt}, \quad (4.1)$$

wobei K_P , K_I und K_D reelle Konstanten sind. Durch Veränderung dieser Gewichtungsfaktoren können die einzelnen Anteile unabhängig voneinander eingestellt werden, sie sind also

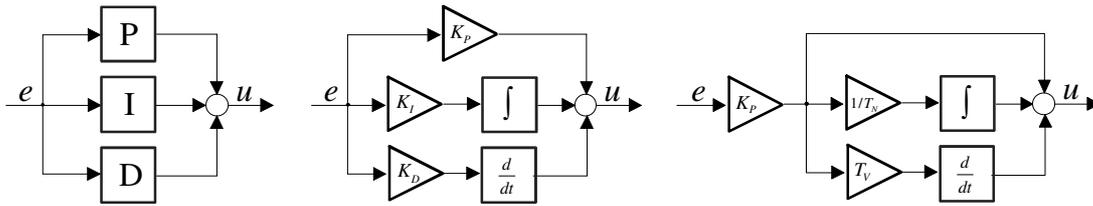


Bild 4.1: Parallelrealisierung eines PID-Reglers

entkoppelt¹.

Häufig wird auch die in der DIN 19226 vorgeschlagene Realisierung verwendet, nämlich

$$u(t) = K_P \left[e(t) + \frac{1}{T_N} \int_0^t e(\tau) d\tau + T_V \frac{de(t)}{dt} \right]. \quad (4.2)$$

Hierbei bezeichnen die Konstanten K_P , T_N und T_V den Proportionalbeiwert, die Nachstellzeit und die Vorhaltezeit, der Ursprung dieser Bezeichnungen wird später erläutert. Man beachte, dass bei der Realisierung (4.2) die oben erwähnte Entkopplung der einzelnen Anteile verloren geht. Vergleicht man nämlich (4.1) und (4.2) so findet man die Zusammenhänge

$$K_I = \frac{K_P}{T_N} \quad \text{bzw.} \quad K_D = K_P T_V, \quad (4.3)$$

d.h. bei einer Änderung von K_P verändern sich auch die Konstanten K_I und K_D . Eine Vergrößerung der Nachstellzeit T_N entspricht einer Reduktion der Gewichtung des I-Anteils, d.h. für $T_N \rightarrow \infty$ wird der I-Anteil deaktiviert. In Tabelle 1 sind die üblicherweise eingesetzten Konfigurationen von PID-Reglern dargestellt.

Reglertyp	K_P	K_I	K_D	K_P	T_N	T_V
P-Regler	*	0	0	*	∞	0
PI-Regler	*	*	0	*	*	0
PD-Regler	*	0	*	*	∞	*
PID-Regler	*	*	*	*	*	*
	Realisierung (4.1)			Realisierung (4.2)		

Tabelle 1: Konfiguration von Standardreglern

¹engl.: non-interacting

Weiters ist zu beachten, dass in (4.1) bzw. (4.2) die - praktisch nicht mögliche - Durchführung einer *idealen* zeitlichen Differentiation erforderlich ist. Das bedeutet, dass die zugehörige Reglerübertragungsfunktion nicht realisierbar ist, d.h. ihr Zählergrad ist größer als ihr Nennergrad. Für einen PID-Regler gemäß (4.2) ergibt sich beispielsweise die Übertragungsfunktion

$$R(s) = \frac{\bar{u}(s)}{\bar{e}(s)} = K_P \left(1 + \frac{1}{sT_N} + sT_V \right) = K_P \frac{1 + sT_N + s^2T_N T_V}{sT_N}. \quad (4.4)$$

Ersetzt man das D-Element durch ein so genanntes DT₁-Element, also durch einen Differenzierer mit Verzögerungsverhalten, so lautet die nunmehr realisierbare Reglerübertragungsfunktion

$$R(s) = \frac{\bar{u}(s)}{\bar{e}(s)} = K_P \left(1 + \frac{1}{sT_N} + \frac{sT_V}{1 + sT_R} \right) = K_P \frac{1 + sT_N(1 + T_R/T_N) + s^2T_N T_V(1 + T_R/T_V)}{sT_N(1 + sT_R)}. \quad (4.5)$$

In diesem Zusammenhang spricht man auch von einem realen PID-Regler bzw. von einem PID-T₁ Regler. Hierbei ist zu beachten, dass die Konstante T_R hinreichend klein gewählt werden muss, d.h.

$$\frac{T_R}{T_V} \ll 1.$$

Für die Sprungantwort, d.h. $e(t) = \sigma(t)$, des realen PID-Reglers gilt

$$u(t) = K_P \left(1 + \frac{1}{T_N} t + \frac{T_V}{T_R} e^{-\frac{t}{T_R}} \right),$$

in Bild 4.2 ist der zugehörige Verlauf der Stellgröße graphisch dargestellt.

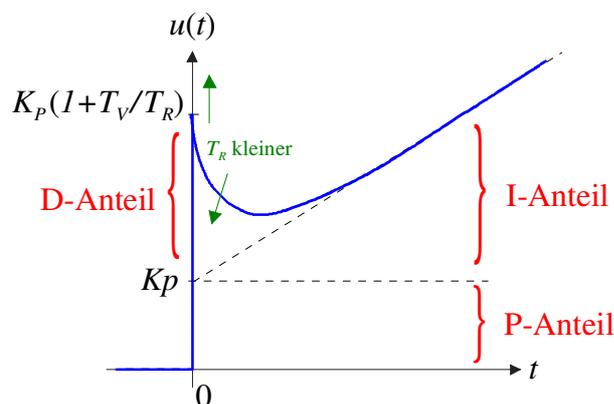


Bild 4.2: Sprungantwort eines realen PID-Reglers

Nachstellzeit

Der Begriff der Nachstellzeit T_N eines PI-Reglers (d.h. „D-Anteil=0“) kann mittels seiner Sprungantwort interpretiert werden. Dazu wird als Eingangsgröße des Reglers der Ein-

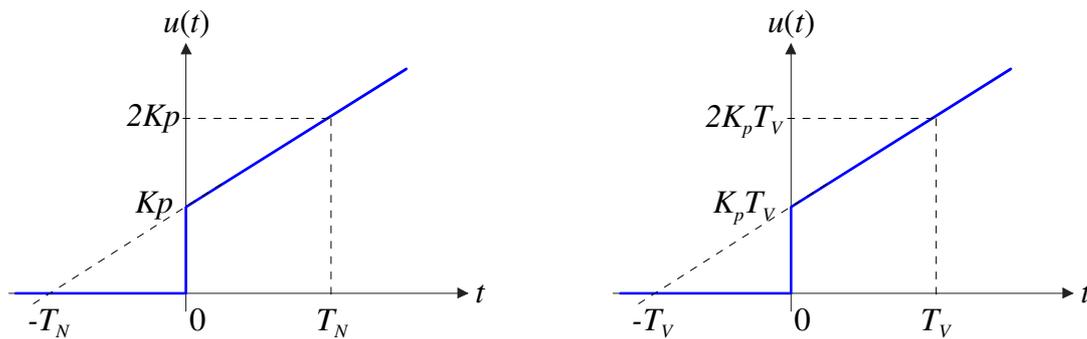


Bild 4.3: Sprungantwort PI-Regler (links) bzw. Rampenantwort PD-Regler (rechts)

heittsprung gewählt, also $e(t) = \sigma(t)$. Für den zugehörigen Verlauf der Ausgangsgröße $u(t)$ gilt dann offensichtlich

$$u(t) = K_P \left(1 + \frac{1}{T_N} t \right).$$

Das bedeutet, dass sich die Stellgröße aus zwei Komponenten zusammensetzt. Während der P-Anteil des Reglers einen konstanten Stellgrößenanteil K_P bewirkt, liefert der I-Anteil einen rampenförmigen Anteil $\frac{K_P}{T_N} t$. Wie leicht zu erkennen ist, siehe hierzu auch Bild 4.3, sind die beiden Stellgrößenanteile zum Zeitpunkt $t = T_N$ genau gleich groß. Somit kann die Nachstellzeit als diejenige Zeit interpretiert werden, die der I-Anteil bei einer sprungförmigen Änderung des Regelfehlers braucht, um einen gleich großen Beitrag zur Stellgröße zu liefern wie der P-Anteil.

Vorhaltezeit

Die Vorhaltezeit T_V kann auf ähnliche Art und Weise gedeutet werden, wie die Nachstellzeit. Wählt man nämlich als Eingangsgröße eines idealen PD-Reglers eine Rampe, d.h. $e(t) = t$, so ergibt sich für den entsprechenden Verlauf der Stellgröße

$$u(t) = K_P (t + T_V).$$

Wieder setzt sich die Stellgröße aus zwei Anteilen zusammen, die offensichtlich zum Zeitpunkt $t = T_V$ den gleichen Wert annehmen. Die Vorhaltezeit ist somit diejenige Zeit, die der P-Anteil bei einer rampenförmigen Änderung des Regelfehlers benötigt, um einen gleich großen Beitrag zur Stellgröße zu liefern wie der D-Anteil, siehe auch Bild 4.3.

4.3 Windup-Effekt und Gegenmaßnahmen

Grundsätzlich ist beim Entwurf von Regelkreisen immer zu beachten, dass in praktischen Anwendungen die Stellgröße betragsmäßig beschränkt ist, also nicht beliebige Werte annehmen kann. Dieser Effekt kann mit Hilfe der in Bild 4.4 in den Standardregelkreis eingefügten Sättigungsfunktion nachgebildet werden, es gilt

$$u^*(t) = \text{sat } u(t) = \begin{cases} u(t) & \text{für } |u(t)| \leq u_{\max} \\ u_{\max} \text{ sign } u(t) & \text{für } |u(t)| \geq u_{\max} \end{cases},$$

wobei die positive, reelle Konstante u_{\max} den maximal zulässigen Betrag der Stellgröße repräsentiert. Man beachte, dass durch das „Ansprechen“ der Sättigungsfunktion die Regel-

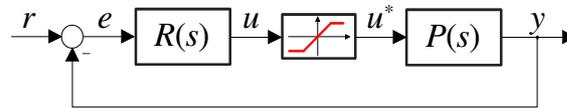


Bild 4.4: Standardregelkreis mit Stellgrößenbeschränkung

güte des Regelkreises dramatisch verschlechtert werden kann, im schlimmsten Fall kann es sogar zur Destabilisierung kommen.

Bei Reglern mit Integralanteil, also auch bei PID-Reglern, kann der so genannte „Windup-Effekt“ auftreten. Dieser ist dadurch gekennzeichnet, dass es beispielsweise bei einem abrupten Arbeitspunktwechsel zu inakzeptablem Über- bzw. Unterschwingen der Regelgröße y kommt bevor der gewünschte stationäre Wert erreicht wird. Die Ursache hierfür besteht darin, dass bei aktiver Stellgrößenbeschränkung, also $|u(t)| \geq u_{\max}$, die Regelstrecke unabhängig vom Verlauf von y mit einer konstanten Eingangsgröße ($\pm u_{\max}$) gespeist wird. Die Rückkopplung ist somit unwirksam und der Regelkreis wird „offen“ betrieben. Obwohl der Betrag von u^* nicht weiter anwachsen kann, kommt es aufgrund der fortwährenden Integration des Regelfehlers zu einem weiteren betragsmäßigen Ansteigen des I-Anteils u_i der Stellgröße. Man spricht in diesem Zusammenhang vom „Aufwickeln“ oder „Aufziehen“ des Integrators oder kurz vom „Windup-Effekt“. Erst wenn der Regelfehler sein Vorzeichen ändert, kommt es zu einer betragsmäßigen Verkleinerung von u_i . Wenn der Betrag von u_i schließlich hinreichend klein ist, spricht die Stellgrößenbeschränkung nicht mehr an und der Regelkreis funktioniert wieder in gewünschter Weise. Aufgrund der unnötig langen zeitlichen Phase, während der $u^* = \pm u_{\max}$ gilt, kommt es zu dem oben erwähnten ungünstigen Verlauf der Regelgröße y .

4.3.1 Anti-Windup-Maßnahmen

Eine naheliegende Maßnahme zur Bekämpfung des Windup-Effekts besteht darin, das betragsmäßige Ansteigen von u_i im Falle einer aktiven Stellgrößenbeschränkung, d.h. $|u(t)| \geq u_{\max}$, zu vermeiden. Das bedeutet, dass die Integration des Regelfehlers bei aktiver Stellgrößenbeschränkung unterbunden wird, also

$$\frac{du_i}{dt} = 0 \quad \text{für} \quad |u(t)| \geq u_{\max}. \quad (4.6)$$

Eine andere häufig angewandte Anti-Windup Strategie ist in Bild 4.5 dargestellt. Sie besteht darin, dass bei aktiver Stellgrößenbeschränkung die Integration nicht vollständig angehalten wird, sondern dass dem Aufwickeln des Integrators proportional zur „Verletzung“ der Stellgrößenbeschränkung entgegengewirkt wird. Für den in Bild 4.5 dargestellten PI-Regler gilt also

$$\frac{du_i}{dt} = K_I e - \frac{1}{T_a} (u - u^*), \quad (4.7)$$

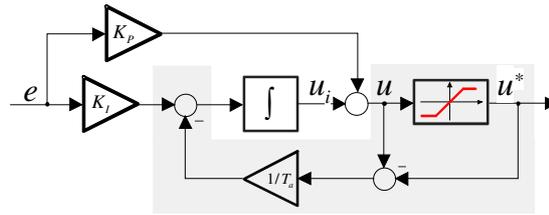


Bild 4.5: PI-Regler mit Anti-Windup Maßnahme (grau hinterlegt)

wobei T_a eine positive Konstante² ist. Die Einstellung von T_a erfolgt üblicherweise empirisch, in der Literatur existieren aber auch Berechnungsvorschläge, wie z.B.

$$T_a = \sqrt{T_N T_V} \quad \text{oder} \quad T_a = \frac{1}{2} (T_N + T_V)$$

für PID-Regler.

B Gegeben sei eine Regelstrecke, deren dynamisches Verhalten im interessierenden Betriebsbereich durch die Übertragungsfunktion

$$P(s) = \frac{s + 0.9}{s^2 + 0.5s + 1}$$

beschrieben werden kann. Für die betragsmäßig beschränkte Stellgröße u gilt

$$|u(t)| \leq u_{\max} = 1.5.$$

Als Regler wurde ein PI-Regler mit der Übertragungsfunktion

$$R(s) = K_P \left(1 + \frac{1}{s T_N} \right) \quad \text{mit} \quad K_P = 3, \quad T_N = 0.5$$

ausgewählt. Weiters wurde die in Bild 4.5 dargestellte Anti-Windup-Maßnahme realisiert, wobei

$$T_a = 0.1$$

gesetzt wurde. In Bild 4.6 sind die Verläufe der Regelgröße y und der Stellgröße u graphisch dargestellt. Es ist deutlich zu erkennen, dass es ohne Anti-Windup-Maßnahme (blaue Kurven) zu einem starken Überschwingen der Regelgröße kommt, was auf das oben beschriebene „Aufwickeln“ des Integrierers zurück zu führen ist. ■

4.4 Einstellregeln für PID-Regler

Mit Hilfe von so genannten Einstellregeln können die Reglerparameter von PID-Reglern relativ geradlinig ermittelt werden. Die hierfür benötigten charakteristischen Streckenparameter

² T_a wird auch als "tracking time constant" bezeichnet.

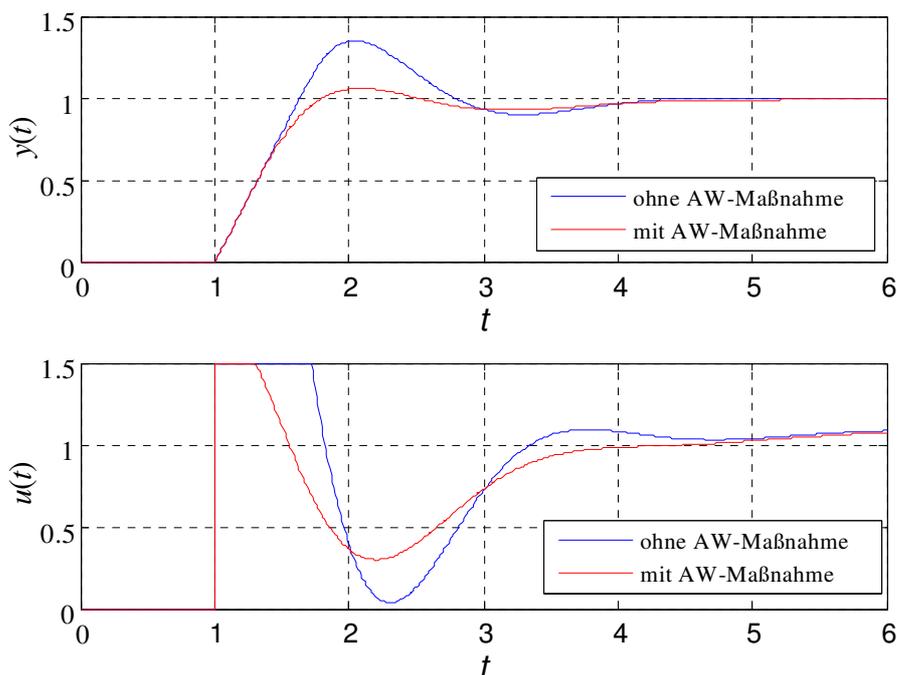


Bild 4.6: Verläufe von Regelgröße (Bild oben) und Stellgröße (Bild unten) ohne und mit Anti-Windup Maßnahme

werden experimentell ermittelt. Die gesuchten Reglerparameter werden dann als Funktion dieser Streckenparameter angegeben, die entsprechenden Relationen sind tabellarisch zusammengefasst. Man beachte, dass hier nur einige der unzähligen Einstellregeln für PID-Regler vorgestellt werden. Weiters ist zu beachten, dass die mittels der Einstellregeln gefundenen Reglerparameter nur als Ausgangspunkt für die Reglerauslegung dienen, die endgültigen Werte der Parameter werden durch nachfolgendes „Feintuning“ ermittelt.

4.4.1 Einstellregeln nach Ziegler-Nichols

Die beiden Methoden nach Ziegler und Nichols wurden im Jahr 1942 vorgestellt und gehören seitdem zu den klassischen Einstellregeln für PID-Regler. Man unterscheidet zwischen der Wendetangenten-Methode („open-loop method“) und der Stabilitätsrand-Methode („closed loop method“). Erfahrungsgemäß führt die Anwendung der Ziegler-Nichols Regeln im Allgemeinen zu schwach gedämpften Regelkreisen.

Wendetangenten-Methode

Bei diesem Verfahren wird vorausgesetzt, dass die Sprungantwort der Regelstrecke gefahrlos experimentell ermittelt werden kann. Wie in Bild 4.7 angedeutet, wird die Wendetangente der Sprungantwort eingezeichnet und die Streckenverstärkung K_S , die Verzugszeit T_v und die Ausgleichszeit T_g werden abgelesen. Die Reglerdimensionierung erfolgt dann mit Hilfe von

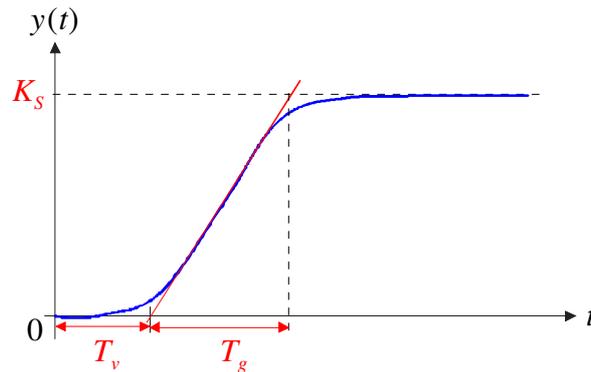


Bild 4.7: Zur Wendetangenten-Methode nach Ziegler-Nichols

Tabelle 2.

Reglertyp	K_P	T_N	T_V
P-Regler	$\frac{T_g}{K_S T_v}$	∞	0
PI-Regler	$0.9 \frac{T_g}{K_S T_v}$	$3.33 T_v$	0
PID-Regler	$1.2 \frac{T_g}{K_S T_v}$	$2 T_v$	$0.5 T_v$

Tabelle 2: Reglerparameter, Wendetangenten-Methode

Stabilitätsrand-Methode

Bei diesem Ansatz werden die benötigten Streckeninformationen aus dem dynamischen Verhalten des geschlossenen Regelkreises ermittelt, wobei als Regler zunächst ein P-Glied verwendet wird. Der Verstärkungsfaktor des P-Reglers wird solange variiert, bis die Regelgröße y bei sprunghörmiger Änderung der Referenzgröße eine ungedämpfte Schwingung vollführt, das System wird also bis an den „Stabilitätsrand“ gebracht. Die Periodendauer T_k der Schwingung wird kritische Periode genannt, der zugehörige Verstärkungsfaktor des Reglers wird mit K_k

bezeichnet. Die Reglerdimensionierung basiert auf den in Tabelle 3 angegebenen Relationen.

Reglertyp	K_P	T_N	T_V
P-Regler	$0.5 K_k$	∞	0
PI-Regler	$0.4 K_k$	$0.8 T_k$	0
PID-Regler	$0.6 K_k$	$0.5 T_k$	$0.12 T_k$

Tabelle 3: Reglerparameter, Stabilitätsrand-Methode

Bei "trägen" Regelstrecken, also Systemen mit sehr großer dominanter Zeitkonstante, kann die Durchführung der Stabilitätsrand-Methode mit einem großen zeitlichen Aufwand verbunden sein. Durch leichte Modifikationen kann die Methode auch zum "Autotuning", d.h. zur selbständigen Einstellung, von Reglerparametern eingesetzt werden. Dabei wird der P-Regler durch ein geeignetes nichtlineares Element ersetzt. Dies hat zur Folge, dass sich die Dauerschwingung der Regelgröße automatisch einstellt.

4.4.2 Methode der Summenzeitkonstante („T-Summen Regel“)

Aus der Sprungantwort der Regelstrecke wird die Streckenverstärkung K_S und die so genannte Summenzeitkonstante T_Σ abgelesen. Wie in Bild 4.8 zu erkennen ist, sind für $t = T_\Sigma$ die beiden

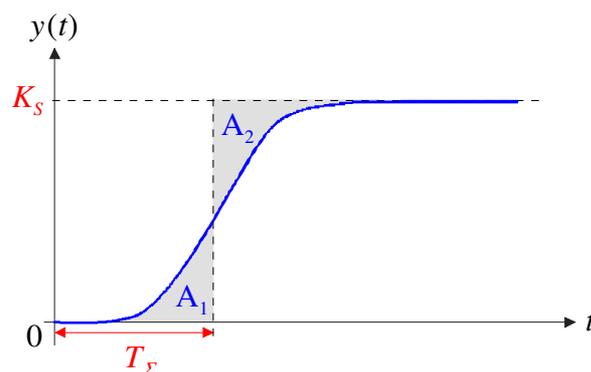


Bild 4.8: Zur Methode der Summenzeitkonstante

Flächen A_1 und A_2 gleich groß. Die Erfahrung zeigt, dass die Ermittlung von T_Σ auch bei stark verrauschter Sprungantwort hinreichend genau durchgeführt werden kann - ganz im Gegensatz zur Wendetangente beim Ziegler-Nichols Verfahren. Die Summenzeitkonstante ist

offensichtlich ein Maß für die „Reaktionsfreudigkeit“ des Systems, d.h. je kleiner T_Σ ist, umso schneller reagiert der Streckenausgang auf sprunghafte Änderungen am Streckeneingang. In Tabelle 4 sind die vorgeschlagenen Reglereinstellungen zusammengefasst.

Reglertyp	K_P	T_N	T_V
P-Regler	$\frac{1}{K_S}$	∞	0
PI-Regler	$\frac{1}{2 K_S}$	$0.5 T_\Sigma$	0
PD-Regler	$\frac{1}{K_S}$	∞	$0.33 T_\Sigma$
PID-Regler	$\frac{1}{K_S}$	$0.66 T_\Sigma$	$0.17 T_\Sigma$

Tabelle 4: Reglerparameter, T-Summen Regel

Kapitel 5

Diskretisierung zeitkontinuierliche Regelgesetze

5.1 Einführung

Sehr häufig müssen zeitkontinuierlich entworfene Regelgesetze zeitdiskret realisiert werden. Das dynamische Verhalten des resultierenden zeitdiskreten Regelkreises soll dann dem Verhalten des ursprünglich entworfenen zeitkontinuierlichen Regelkreises „möglichst nahe“ kommen. Diese „Güte“ der Diskretisierung wird dabei üblicherweise durch numerische Simulation des zeitdiskreten Regelkreises bewertet. Eine notwendige Voraussetzung für eine zufriedenstellende Diskretisierung ist natürlich die sinnvolle Wahl der Diskretisierungszeit T_d , wovon in weiterer Folge auch ausgegangen wird.

Die grundlegenden Ideen der Reglerdiskretisierung werden im Folgenden anhand eines PI-Reglers mit der Zeitbereichsbeschreibung

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau \quad (5.1)$$

und der Übertragungsfunktion

$$R(s) = \frac{\bar{u}(s)}{\bar{e}(s)} = \frac{\mathcal{L}\{u(t)\}}{\mathcal{L}\{e(t)\}} = \frac{K_P s + K_I}{s} \quad (5.2)$$

demonstriert. Hierbei stellen K_P und K_I die reellen Reglerparameter dar, die beispielsweise mit Hilfe der vorgestellten Einstellregeln ermittelt werden können. Das Regelgesetz (5.1) wird nun zu den äquidistanten Zeitpunkten

$$t = k T_d \quad \text{wobei} \quad k = 0, 1, 2, 3, \dots \quad (5.3)$$

ausgewertet, d.h.

$$u(k T_d) = K_P e(k T_d) + K_I \int_0^{k T_d} e(\tau) d\tau. \quad (5.4)$$

Mit den Abkürzungen

$$u_k := u(k T_d) \quad \text{and} \quad e_k := e(k T_d)$$

erhält man

$$u_k = K_P e_k + K_I \int_0^{kT_d} e(\tau) d\tau = K_P e_k + K_I \int_0^{(k-1)T_d} e(\tau) d\tau + K_I \int_{(k-1)T_d}^{kT_d} e(\tau) d\tau, \quad (5.5)$$

was schlussendlich auf die Beziehung

$$u_k = u_{k-1} - K_P e_{k-1} + K_P e_k + K_I \int_{(k-1)T_d}^{kT_d} e(\tau) d\tau. \quad (5.6)$$

führt. Die nachfolgend angeführten Methoden zur Diskretisierung des PI-Reglers unterscheiden sich durch die Art der numerischen Berechnung des in (5.6) auftretenden Integrals.

5.2 Vorwärts-Euler-Integration

Hier wird die Approximation

$$\int_{(k-1)T_d}^{kT_d} e(\tau) d\tau \approx T_d e_{k-1} \quad (5.7)$$

verwendet, d.h. das Integral wird durch ein Rechteck, das in Bild 5.1 schraffiert dargestellt ist, angenähert. Die Differenzengleichung, die das zeitliche Verhalten des zeitdiskreten PI-Reglers

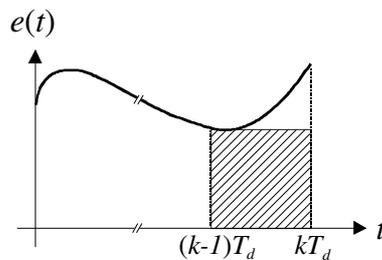


Bild 5.1: Prinzip der Vorwärts-Euler-Integration

beschreibt, lautet somit

$$u_k = u_{k-1} + K_P e_k + (K_I T_d - K_P) e_{k-1}. \quad (5.8)$$

Wendet man auf (5.8) die z-Transformation an, so erhält man die z-Übertragungsfunktion

$$R_d(z) = \frac{\tilde{u}(z)}{\tilde{e}(z)} = \frac{K_P z + (K_I T_d - K_P)}{z - 1} = \frac{K_P (z - 1) + K_I T_d}{z - 1} = \frac{K_P \frac{z-1}{T_d} + K_I}{\frac{z-1}{T_d}}. \quad (5.9)$$

Der direkte Vergleich von (5.2) mit dem zeitdiskreten Regler (5.9) zeigt, dass

$$R_d(z) = R(s)|_{s=\frac{z-1}{T_d}}. \quad (5.10)$$

gilt. Der Zusammenhang zwischen Punkten der komplexen s -Ebene und Punkten in der komplexen z -Ebene wird also durch die Relation

$$z = 1 + sT_d \quad (5.11)$$

beschrieben. Offensichtlich wird die imaginäre Achse der s -Ebene, d.h. $s = j\omega$ auf $z = 1 + j\omega T_d$ abgebildet. Das bedeutet, dass die linke offene Halbebene $\operatorname{Re}\{s\} < 0$ abgebildet wird auf die Halbebene $\operatorname{Re}\{z\} < 1$, siehe Bild 5.2. Ein stabiler Regler $R(s)$ kann also prinzipiell durch diese Art der Diskretisierung (5.10) in einen instabilen zeitdiskreten Regler $R_d(z)$ übergehen.

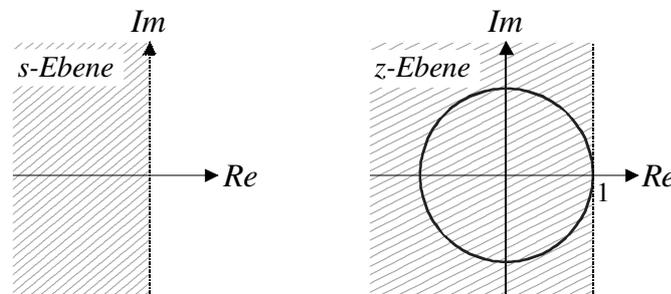


Bild 5.2: Abbildung von $\operatorname{Re}\{s\} \leq 0$ auf $\operatorname{Re}\{z\} \leq 1$ bei der Vorwärts-Euler-Integration

5.3 Rückwärts-Euler-Integration

Bei dieser Art der Diskretisierung wird die Approximation

$$\int_{(k-1)T_s}^{kT_s} e(\tau) d\tau \approx T_d e_k \quad (5.12)$$

eingesetzt, das Integral in (5.6) wird also durch das in Bild 5.3 schraffiert dargestellte Rechteck angenähert. Für das zeitdiskrete Regelgesetz ergibt sich somit die Differenzgleichung

$$u_k = u_{k-1} + (K_P + K_I T_d) e_k - K_P e_{k-1} \quad (5.13)$$

mit der zugehörigen z -Übertragungsfunktion

$$R_d(z) = \frac{\tilde{u}(z)}{\tilde{e}(z)} = \frac{(K_P + K_I T_d) z - K_P}{z - 1} = \frac{K_P (z - 1) + K_I T_d z}{z - 1} = \frac{K_P \frac{z-1}{zT_d} + K_I}{\frac{z-1}{zT_d}}. \quad (5.14)$$

Der Zusammenhang zwischen $R(s)$ und $R_d(z)$ wird also durch

$$R_d(z) = R(s)|_{s=\frac{z-1}{zT_d}} \quad (5.15)$$

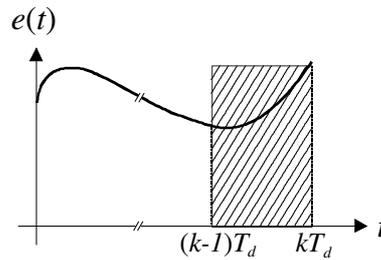


Bild 5.3: Prinzip der Rückwärts-Euler-Integration

beschrieben. Punkte in der komplexen s -Ebene stehen mit Punkten in der komplexen z -Ebene über die Relation

$$z = \frac{1}{1 - sT_d}$$

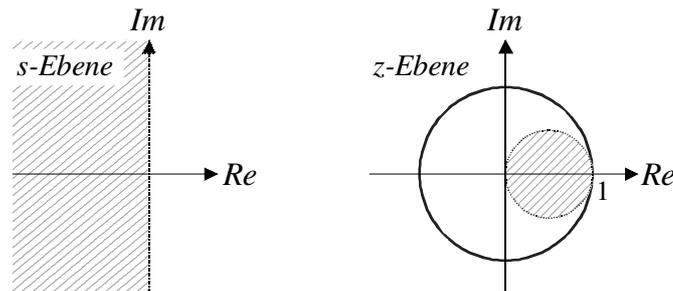
in Beziehung. Die imaginäre Achse der s -Ebene, d.h. $s = j\omega$ geht in den in Bild 5.4 dargestellten Kreis

$$\left| z - \frac{1}{2} \right| = \frac{1}{2}$$

über. Die linke offene s -Ebene geht in das Innere des Kreises über, d.h.

$$\operatorname{Re}\{s\} < 0 \quad \Longrightarrow \quad \left| z - \frac{1}{2} \right| < \frac{1}{2}. \quad (5.16)$$

Daraus kann unmittelbar gefolgert werden, dass durch diese Art der Diskretisierung die BIBO-Stabilität der Reglerübertragungsfunktion erhalten bleibt.

Bild 5.4: Abbildung von $\operatorname{Re}\{s\} \leq 0$ auf $\left| z - \frac{1}{2} \right| \leq \frac{1}{2}$ bei der Rückwärts-Euler-Integration

5.4 Integration mittels Trapez-Regel - Tustin Formel

Bei der Integration nach der Trapez-Regel wird, wie in Bild 5.5 dargestellt, das Integral in (5.6) durch das schraffierte Trapez approximiert, d.h.

$$\int_{(k-1)T_s}^{kT_s} e(\tau) d\tau = \frac{T_d}{2} (e_{k-1} + e_k). \quad (5.17)$$

Für die Differenzengleichung ergibt sich somit

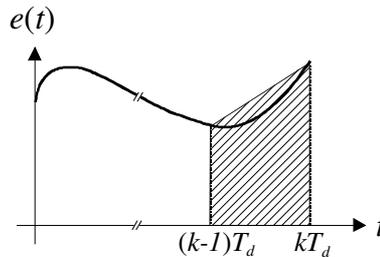


Bild 5.5: Prinzip der Trapez-Integration

$$\begin{aligned} u_k &= K_P e_k - K_P e_{k-1} + u_{k-1} + K_I \frac{T_d}{2} (e_{k-1} + e_k) = \\ &= u_{k-1} + (K_I \frac{T_d}{2} + K_P) e_k + (K_I \frac{T_d}{2} - K_P) e_{k-1}, \end{aligned} \quad (5.18)$$

d.h. z-Übertragungsfunktion des Reglers lautet

$$R_d(z) = \frac{\tilde{u}(z)}{\tilde{e}(z)} = \frac{(K_I \frac{T_d}{2} + K_P)z + (K_I \frac{T_d}{2} - K_P)}{z - 1} = \frac{K_P(z - 1) + K_I \frac{T_d}{2}(z + 1)}{z - 1} = \frac{K_P \frac{2}{T_d} \frac{z-1}{z+1} + K_I}{\frac{2}{T_d} \frac{z-1}{z+1}}.$$

Es gilt somit

$$R_d(z) = R(s)|_{s=\frac{2}{T_d} \frac{z-1}{z+1}}. \quad (5.19)$$

Den in (5.19) angegebenen Zusammenhang

$$s = \frac{2}{T_d} \frac{z - 1}{z + 1} \quad (5.20)$$

zwischen den komplexen Variablen s und z nennt man auch **Tustin-Formel**. Die Umkehrrelation zu (5.20) lautet

$$z = \frac{1 + s \frac{T_d}{2}}{1 - s \frac{T_d}{2}}, \quad (5.21)$$

d.h. die imaginäre Achse der s -Ebene geht in den Einheitskreis der z -Ebene über bzw. die linke offene s -Ebene geht in das Innere des Einheitskreises über, siehe auch Bild 5.6.

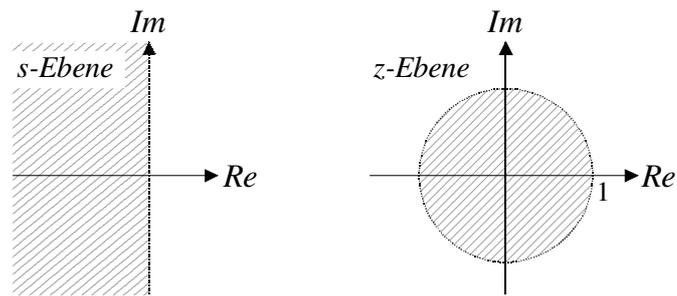


Bild 5.6: Abbildung von $\text{Re}\{s\} \leq 0$ auf $|z| \leq 1$ bei der Trapez-Integration

Kapitel 6

Analytische Reglersynthese

6.1 Einführung

Die Idee der so genannten „Analytischen Reglersynthese“ besteht darin, dem Regelkreis eine gewünschte Führungsübertragungsfunktion $T(s)$ vorzugeben. Aus $T(s)$ und der gegebenen Streckenübertragungsfunktion $P(s)$ wird dann der gesuchte Regler berechnet. Damit unterscheidet sich der Entwurfsprozess grundlegend von anderen Entwurfstechniken.

6.2 Implementierbarkeit

Eine Übertragungsfunktion $T(s)$ wird implementierbar genannt, wenn es eine Regelkreisstruktur mit folgenden Eigenschaften gibt:

- (i) Alle Teilsysteme des Regelkreises sind realisierbar
- (ii) Der Regelkreis ist intern stabil
- (iii) $T(s)$ ist die Führungsübertragungsfunktion des Regelkreises

Die Bedingung (i) besagt, dass jeder „Block“ im Regelkreis ein realisierbares System sein muss, d.h. jedes Teilsystem kann in Form eines Zustandsmodells realisiert werden. Die Erfüllung von (ii) hat zur Folge, dass es im Regelkreis zu keinen unerlaubten „instabilen“ Kürzungen kommt, d.h. jede Übertragungsfunktion des geschlossenen Regelkreises ist BIBO-stabil. Ist $T(s)$ also nicht implementierbar, so gibt es keinen praktisch sinnvollen Regelkreis mit dieser Führungsübertragungsfunktion.

Die Implementierbarkeit einer Führungsübertragungsfunktion

$$T(s) = \frac{\mu_T(s)}{\nu_T(s)} \quad (6.1)$$

bei gegebener teilerfremder (d.h. gekürzter) Streckenübertragungsfunktion

$$P(s) = \frac{\mu(s)}{\nu(s)} \quad (6.2)$$

kann leicht überprüft werden. $T(s)$ ist nämlich genau dann implementierbar, wenn gilt

- (a) Das Polynom $\nu_T(s)$ ist ein Hurwitzpolynom
- (b) Nullstellen von $\mu(s)$ in der geschlossenen rechten Halbebene sind Nullstellen von $\mu_T(s)$
- (c) $\text{Grad } \nu_T(s) - \text{Grad } \mu_T(s) \geq \text{Grad } \nu(s) - \text{Grad } \mu(s)$

B Gegeben sei die Regelstrecke

$$P(s) = \frac{s-1}{s(s+2)}.$$

Es ist zu überprüfen, welche der angegebenen Übertragungsfunktionen implementierbar sind:

	(a)	(b)	(c)	
$T_1(s) = \frac{s-1}{s+1}$	✓	✓	×	nicht implementierbar
$T_2(s) = \frac{s-1}{s^2-2s+1}$	×	✓	✓	nicht implementierbar
$T_3(s) = \frac{s+1}{s^2+2s+5}$	✓	×	✓	nicht implementierbar
$T_4(s) = \frac{s-1}{s^2+2s+5}$	✓	✓	✓	implementierbar

Das bedeutet, dass nur $T_4(s)$ eine sinnvolle Wahl für die Führungsübertragungsfunktion eines Regelkreises mit der vorgegebenen Regelstrecke $P(s)$ darstellt. ■

6.3 Entwurf für den Standardregelkreis

Zunächst wird die wohl naheliegendste Vorgangsweise zur Ermittlung von $R(s)$ aus $T(s)$ und $P(s)$ vorgestellt. Es zeigt sich, dass diese direkte Art der Reglerberechnung allerdings nur eingeschränkt anwendbar ist und mit Hilfe des Standardregelkreises nicht jede implementierbare Führungsübertragungsfunktion realisiert werden kann. Daher beschränkt man sich bei der so genannten Polvorgabe auf die Vorgabe der Pole von $T(s)$.

6.3.1 Direkte Reglerberechnung

Aus der vorgegebenen Führungsübertragungsfunktion $T(s)$ errechnet sich die Übertragungsfunktion des offenen Kreises gemäß

$$L(s) = \frac{T(s)}{1-T(s)}. \quad (6.3)$$

Man beachte, dass der resultierende Regelkreis genau dann intern stabil ist, wenn alle „instabilen“ Pole bzw. Nullstellen von $P(s)$ auch Pole bzw. Nullstellen von $L(s)$ sind. Anderenfalls kommt es zu unerlaubten instabilen Kürzungen bei der Berechnung von $L(s)$. Der gesuchte Regler $R(s)$ kann über die Relation

$$R(s) = \frac{L(s)}{P(s)} \quad (6.4)$$

berechnet werden.

B Gegeben sei die Regelstrecke

$$P(s) = \frac{s-1}{s(s+2)}$$

und die gewünschte (implementierbare) Führungsübertragungsfunktion

$$T(s) = \frac{-3(s-1)}{s^2 + 2.8s + 3}.$$

Für die Übertragungsfunktion des offenen Kreises ergibt sich

$$L(s) = \frac{-3(s-1)}{s(s+5.8)},$$

d.h. $L(s)$ besitzt den instabilen Streckenpol bei $s = 0$ und die instabile Streckennullstelle bei $s = 1$. Das bedeutet, dass der resultierende Regelkreis intern stabil ist und die Reglerübertragungsfunktion lautet

$$R(s) = \frac{-3(s+2)}{s+5.8}.$$

■

B Mit der Streckenübertragungsfunktion

$$P(s) = \frac{1}{s^2 - 1} = \frac{1}{(s+1)(s-1)},$$

und der gewünschten (implementierbaren) Führungsübertragungsfunktion

$$T(s) = \frac{3}{s^2 + 2.8s + 3}$$

findet man

$$L(s) = \frac{3}{s(s+2.8)}.$$

Man erkennt, dass der instabile Streckenpol bei $s = 1$ kein Pol von $L(s)$ ist, d.h. es kommt zu einer instabilen Kürzung im offenen Kreis. Diese Tatsache erkennt man auch an der Reglerübertragungsfunktion

$$R(s) = \frac{3(s+1)(s-1)}{s(s+2.8)}.$$

■

so ist die Resultante eine reguläre $2n \times 2n$ -Matrix, d.h. das Gleichungssystem (6.9) kann eindeutig gelöst werden.

B Gegeben sei die Streckenübertragungsfunktion ($n = 2$)

$$P(s) = \frac{1}{s^2 + 2}.$$

Es soll ein Regler der Ordnung $\rho = n - 1 = 1$ so ermittelt werden, dass alle Pole von $T(s)$ an der Stelle $s = -1$ liegen, d.h.

$$\nu_T(s) = (s + 1)^3 = s^3 + 3s^2 + 3s + 1.$$

Für das Gleichungssystem (6.9) ergibt sich somit

$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix},$$

als Lösung findet man

$$a_0 = 3, a_1 = 1, b_0 = -5 \quad \text{und} \quad b_1 = 1,$$

d.h.

$$R(s) = \frac{s - 5}{s + 3}.$$

Die resultierende Führungsübertragungsfunktion lautet

$$T(s) = \frac{s - 5}{s^3 + 3s^2 + 3s + 1} = \frac{s - 5}{(s + 1)^3}.$$

Sie besitzt eine Nullstelle bei $s = 5$, die sich gemäß (6.7) aus dem Entwurf ergibt. ■

Wählt man

$$\rho > n - 1, \tag{6.11}$$

so ist das Gleichungssystem (6.9) unterbestimmt, d.h. die Resultante besitzt weniger Zeilen als Spalten. Durch die Erhöhung der Reglerordnung können also zusätzlich zu den gewünschten Polen von $T(s)$ weitere $(\rho - n + 1)$ Bedingungen für die Reglerkoeffizienten vorgegeben werden. Dadurch ergeben sich zusätzliche Freiheitsgrade für den Reglerentwurf.

B Fortsetzung. Gleich wie im vorigen Beispiel sollen alle Pole von $T(s)$ an der Stelle $s = -1$ liegen. Zusätzlich soll der Regler integrierendes Verhalten zeigen, d.h. eine Polstelle bei $s = 0$ besitzen. Dieser „Zusatzwunsch“ kann durch die Erhöhung der Reglerordnung auf $\rho = 2$ im Entwurf berücksichtigt werden. Dazu muss lediglich die Gleichung $a_0 = 0$ zum Gleichungssystem (6.9) hinzugefügt werden. Damit folgt unter Berücksichtigung von

$$\nu_T(s) = (s + 1)^4 = s^4 + 4s^3 + 6s^2 + 4s + 1$$

unmittelbar das Gleichungssystem

$$\begin{bmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \\ 0 \end{bmatrix},$$

aus dem sich

$$a_0 = 0, a_1 = 4, a_2 = 1, b_0 = 1, b_1 = -4 \text{ und } b_2 = 4$$

ergibt, d.h.

$$R(s) = \frac{4s^2 - 4s + 1}{s^2 + 4s}.$$

Die resultierende Führungsübertragungsfunktion lautet

$$T(s) = \frac{4s^2 - 4s + 1}{s^4 + 4s^3 + 6s^2 + 4s + 1} = \frac{4(s - 0.5)^2}{(s + 1)^4}.$$

■

Mit der beschriebenen Methode lassen sich zwar die Pole der Übertragungsfunktion $T(s)$, nicht jedoch ihre Nullstellen vorgeben. Aus diesem Grund wird im folgenden Abschnitt eine erweiterte Regelkreisstruktur betrachtet, mit deren Hilfe *jede* implementierbare Führungsübertragungsfunktion realisiert werden kann.

6.4 Entwurf für eine erweiterte Regelkreisstruktur

Für die Führungsübertragungsfunktion der in Bild 6.1 dargestellten erweiterten Regelkreis-

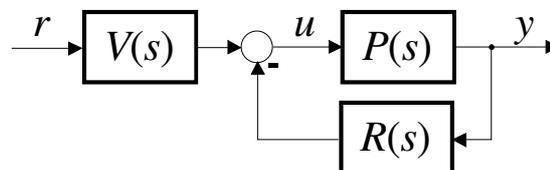


Bild 6.1: Erweiterte Regelkreisstruktur

struktur findet man

$$T(s) = \frac{V(s)P(s)}{1 + R(s)P(s)}. \quad (6.12)$$

Man beachte, dass der Regler im vorliegenden Fall *ein* System mit zwei Eingangsgrößen y und r und einer Ausgangsgröße u ist. Daraus resultiert, dass die beiden Übertragungsfunktionen

$R(s)$ und $V(s)$ dasselbe Nennerpolynom $a(s)$ besitzen, d.h.

$$R(s) = \frac{b(s)}{a(s)} = \frac{b_\rho s^\rho + \dots + b_1 s + b_0}{a_\rho s^\rho + \dots + a_1 s + a_0} \quad \text{und} \quad V(s) = \frac{c(s)}{a(s)} = \frac{c_\rho s^\rho + \dots + c_1 s + c_0}{a_\rho s^\rho + \dots + a_1 s + a_0}. \quad (6.13)$$

Für $T(s)$ folgt nun

$$T(s) = \frac{c(s)\mu(s)}{a(s)\nu(s) + b(s)\mu(s)}. \quad (6.14)$$

Wählt man nun eine beliebige implementierbare Führungsübertragungsfunktion gemäß (6.1), so müssen die Bedingungen

$$c(s)\mu(s) = \mu_T(s) \quad \text{und} \quad a(s)\nu(s) + b(s)\mu(s) = \nu_T(s) \quad (6.15)$$

gelten. Daraus ist unmittelbar ersichtlich, dass $c(s)$ genau dann, wie gewünscht, ein Polynom ist, wenn $\mu(s)$ ein Teiler von $\mu_T(s)$ ist.

B Mit der Streckenübertragungsfunktion

$$P(s) = \frac{1}{s^2 - 1} = \frac{1}{(s+1)(s-1)} = \frac{\mu(s)}{\nu(s)},$$

und der gewünschten (implementierbaren) Führungsübertragungsfunktion

$$T(s) = \frac{1}{s^3 + 3s^2 + 3s + 1} = \frac{\mu_T(s)}{\nu_T(s)}$$

findet man mittels (6.15)

$$c(s) \cdot 1 = 1 \quad \Rightarrow \quad c(s) = 1$$

und

$$a(s) \cdot (s^2 - 1) + b(s) \cdot 1 = s^3 + 3s^2 + 3s + 1.$$

Die Polynome $a(s)$ und $b(s)$ können aus dem Gleichungssystem (6.9) berechnet werden. Für den Regler findet man

$$R(s) = \frac{4(s+1)}{s+3} \quad \text{und} \quad V(s) = \frac{1}{s+3}.$$

■

Im nachfolgenden Beispiel ist die Voraussetzung, dass $\mu(s)$ ein Teiler von $\mu_T(s)$ ist, verletzt, d.h. die beschriebene Methodik kann nicht angewandt werden.

B Mit der Streckenübertragungsfunktion

$$P(s) = \frac{s+2}{s^2-1} = \frac{s+2}{(s+1)(s-1)} = \frac{\mu(s)}{\nu(s)},$$

und der gewünschten (implementierbaren) Führungsübertragungsfunktion

$$T(s) = \frac{1}{s^3 + 3s^2 + 3s + 1} = \frac{\mu_T(s)}{\nu_T(s)}$$

führt (6.15) auf die Bedingung

$$c(s)(s + 2) = 1,$$

d.h. es gibt kein Polynom $c(s)$ als Lösung dieser Gleichung. ■

Abhilfe schafft eine geringfügige Modifikation der Vorgangsweise. Zunächst wird die Übertragungsfunktion $T(s)$ durch das Zählerpolynom der Strecke dividiert, also

$$\frac{1}{\mu(s)}T(s) = \frac{1}{\mu(s)}\frac{\mu_T(s)}{\nu_T(s)} = \frac{\tilde{\mu}_T(s)}{\tilde{\nu}_T(s)}.$$

Dabei ist darauf zu achten, dass bei der Berechnung von $\tilde{\mu}_T(s)$ bzw. $\tilde{\nu}_T(s)$ gegebenenfalls alle instabile Kürzungen durchzuführen sind, d.h. $\tilde{\nu}_T(s)$ ist jedenfalls ein Hurwitzpolynom. Aus (6.14) folgt dann unmittelbar

$$\frac{\tilde{\mu}_T(s)}{\tilde{\nu}_T(s)} = \frac{c(s)}{a(s)\nu(s) + b(s)\mu(s)}, \quad (6.16)$$

d.h. es ergeben sich die Entwurfsgleichungen

$$c(s) = \tilde{\mu}_T(s) \quad \text{und} \quad a(s)\nu(s) + b(s)\mu(s) = \tilde{\nu}_T(s). \quad (6.17)$$

B Mit der Streckenübertragungsfunktion

$$P(s) = \frac{s + 2}{(s + 1)(s - 1)} = \frac{\mu(s)}{\nu(s)},$$

und der gewünschten (implementierbaren) Führungsübertragungsfunktion

$$T(s) = \frac{1}{(s + 1)^2} = \frac{\mu_T(s)}{\nu_T(s)}$$

führt (6.16) zu

$$\frac{1}{\mu(s)}T(s) = \frac{1}{(s + 2)(s + 1)^2} = \frac{1}{s^3 + 4s^2 + 5s + 2} = \frac{\tilde{\mu}_T(s)}{\tilde{\nu}_T(s)}.$$

Aus (6.17) folgt nun unmittelbar

$$R(s) = \frac{6(s + 1)}{s + 4} \quad \text{und} \quad V(s) = \frac{1}{s + 4}.$$

■

Man beachte, dass der Polynomgrad von $\tilde{\nu}_T(s)$ gemäß (6.17) genau $(n + \rho)$ betragen muss. Gegebenenfalls kann die Erfüllung dieser Bedingung durch eine Erweiterung von $T(s)$ erreicht werden, wie das nachfolgende Beispiel verdeutlicht.

B Gegeben sei die Streckenübertragungsfunktion

$$P(s) = \frac{s + 2}{(s + 1)(s - 1)} = \frac{\mu(s)}{\nu(s)},$$

und die gewünschte (implementierbare) Führungsübertragungsfunktion

$$T(s) = \frac{1}{s + 1} = \frac{\mu_T(s)}{\nu_T(s)}.$$

Es soll ein Regler so entworfen werden, dass die Führungsübertragungsfunktion des Regelkreises $T(s)$ ist und der Regler integrierendes Verhalten hat. Daher wird die Reglerordnung $\rho = 2$ gewählt, der Polynomgrad von $\tilde{\nu}_T(s)$ muss also 4 betragen.

Aus (6.16) folgt

$$\frac{1}{\mu(s)}T(s) = \frac{1}{(s + 2)(s + 1)} = \frac{1}{s^2 + 3s + 2} = \frac{\tilde{\mu}_T(s)}{\tilde{\nu}_T(s)},$$

d.h. der Polynomgrad von $\tilde{\nu}_T(s)$ beträgt nur 2. Eine Erhöhung des Polynomgrads erreicht man, indem $T(s)$ als

$$T(s) = \frac{1}{s + 1} \frac{w(s)}{w(s)}$$

angeschrieben wird, wobei $w(s)$ ein Hurwitzpolynom vom Grade 2 ist. Die spezielle Wahl von $w(s)$ hat keinen Einfluss auf das Führungsverhalten des Regelkreises, wohl aber auf sein Störverhalten. In Form von $w(s)$ ergeben sich also zusätzliche Freiheiten beim Reglerentwurf. Im vorliegenden Fall wird willkürlich

$$w(s) = (s + 3)^2$$

gewählt, d.h.

$$T(s) = \frac{1}{(s + 1)} \frac{(s + 3)^2}{(s + 3)^2}.$$

Aus (6.16) folgt nun

$$\frac{1}{\mu(s)}T(s) = \frac{1}{(s + 2)(s + 1)} \frac{(s + 3)^2}{(s + 3)^2} = \frac{(s + 3)^2}{s^4 + 9s^3 + 29s^2 + 39s + 18} = \frac{\tilde{\mu}_T(s)}{\tilde{\nu}_T(s)},$$

der Regler kann also mittels (6.17) berechnet werden, d.h.

$$c(s) = (s + 3)^2 = s^2 + 6s + 9$$

und

$$\begin{bmatrix} -1 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 1 & 2 & 0 \\ 1 & 0 & -1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 18 \\ 39 \\ 29 \\ 9 \\ 1 \\ 0 \end{bmatrix}.$$

Daraus ergibt sich

$$R(s) = \frac{7s^2 + 16s + 9}{s(s+2)} \quad \text{und} \quad V(s) = \frac{s^2 + 6s + 9}{s(s+2)}.$$

■

6.5 Wahl von $T(s)$

Häufig wird für $T(s)$ der Ansatz

$$T(s) = \frac{\omega_n^2}{s^2 + 2d\omega_n s + \omega_n^2} \quad \text{mit} \quad \omega_n > 0 \quad \text{und} \quad 0 < d < 1, \quad (6.18)$$

gewählt, wobei ω_n die so genannte Kennkreisfrequenz und d der Dämpfungsgrad ist. Die angegebenen Wertebereiche für ω_n und d gewährleisten, dass $T(s)$ BIBO-stabil ist. In Bild 6.2 ist der typische Verlauf der Sprungantwort eines Systems (6.18) dargestellt. Mit M_p wird die

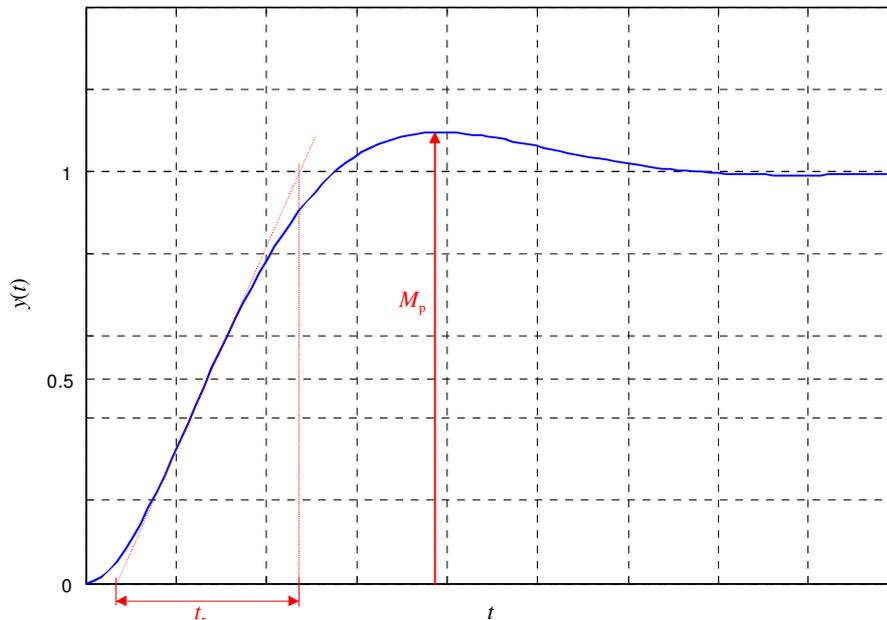


Bild 6.2: Sprungantwort eines Systems der Form (6.18)

Überschwingweite („maximum peak“) bezeichnet, sie gibt den maximalen Wert der Regelgröße an, wenn die Referenzgröße der Einheitssprung ist. Aus der Überschwingweite M_p kann das entsprechende prozentuale Überschwingen gemäß

$$\ddot{u} = 100 (M_p - 1) \quad \text{in \%} \quad (6.19)$$

ermittelt werden. Mit t_r wird die so genannte Anstiegszeit („rise time“) bezeichnet. Man beachte, dass in der Literatur verschiedene Definitionen für die Anstiegszeit zu finden sind, die sich bei genauerer Untersuchung jedoch nur geringfügig voneinander unterscheiden.

Bei vorgegebenem M_p und t_r können die Parameter ω_n und d mit Hilfe der in Bild 6.3 dargestellten Kurven leicht ermittelt werden.

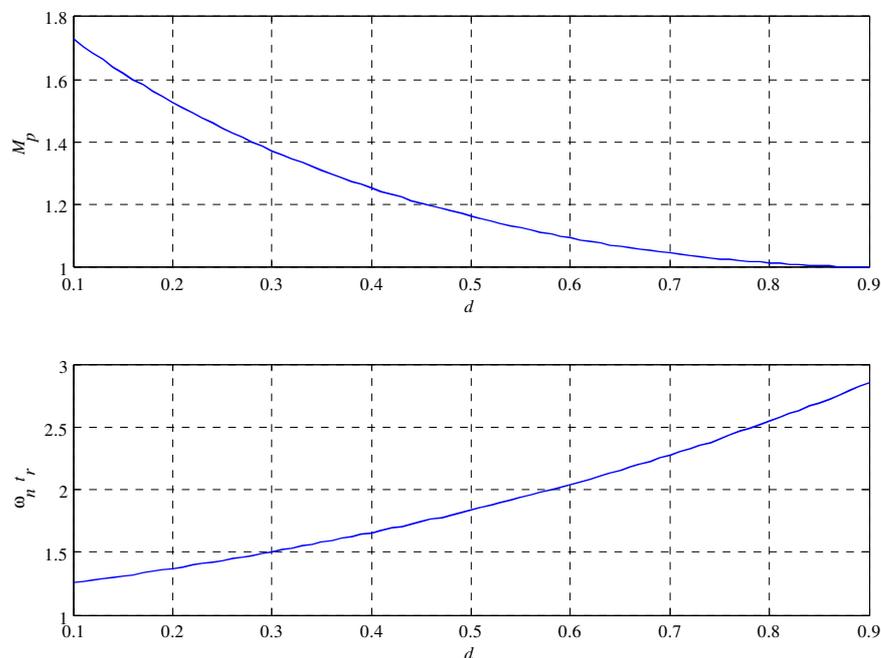


Bild 6.3: Ermittlung von ω_n und d aus M_p und t_r

B Soll die Sprungantwort des Regelkreises ein prozentuales Überschwingen von $\ddot{u} = 10\%$ und eine Anstiegszeit $t_r = 1$ besitzen, so folgt aus Bild 6.3 unmittelbar

$$M_p = 1.1 \quad \Rightarrow \quad d = 0.6 \quad \Rightarrow \quad \omega_n t_r = 2 \quad \Rightarrow \quad \omega_n = 2.$$

Die gesuchte Führungsübertragungsfunktion lautet somit

$$T(s) = \frac{4}{s^2 + 2.4s + 4}.$$

Man beachte, dass eine Führungsübertragungsfunktion der Form (6.18) prinzipiell nur dann implementierbar sein kann, wenn die Streckenübertragungsfunktion $P(s)$ keine „instabilen“ Nullstellen, d.h. Nullstellen mit nichtnegativem Realteil besitzt. Weiters muss für Implementierbarkeit von (6.18) die Bedingung

$$\text{Grad } \nu(s) - \text{Grad } \mu(s) \leq 2 \quad (6.20)$$

erfüllt sein. Die Bedingung (6.20) kann gelockert werden, wenn der Ansatz für $T(s)$ gemäß

$$T(s) = \frac{\omega_n^2}{s^2 + 2d\omega_n s + \omega_n^2} \cdot \frac{1}{\left(1 + \frac{s}{\alpha}\right)^k} \quad (6.21)$$

modifiziert wird. Dabei ist α eine hinreichend große positive Konstante und für den ganzzahligen Parameter k gilt

$$k \geq \text{Grad } \nu(s) - \text{Grad } \mu(s) - 2. \quad (6.22)$$

B Die im vorigen Beispiel gefundene Führungsübertragungsfunktion soll so modifiziert werden, dass sie bei gegebener Streckenübertragungsfunktion

$$P(s) = \frac{1}{(s-1)(s-2)(s-3)}$$

implementierbar ist. Um die Implementierbarkeit zu gewährleisten, muss der Nennergrad von $T(s)$ gemäß (6.21) erweitert werden, d.h.

$$T(s) = \frac{4}{s^2 + 2.4s + 4} \cdot \frac{1}{1 + \frac{s}{\alpha}}$$

Der reelle Parameter α muss groß genug gewählt werden, damit die Sprungantwort der modifizierten Führungsübertragungsfunktion die Vorgaben $M_p = 1.1$ und $t_r = 1$ hinreichend gut erfüllt. Es ist allerdings zu beachten, dass eine zu große Wahl von α zu großen Werten der Stellgröße u führt. Berechnet man den Wert der Stellgröße zum Zeitpunkt $t = 0$ für eine sprungförmige Referenzgröße, so findet man

$$u_0 = u(0) = \lim_{s \rightarrow \infty} s \bar{u}(s) = \lim_{s \rightarrow \infty} s \frac{T(s)}{P(s)} \bar{r}(s) = \lim_{s \rightarrow \infty} s \frac{4(s-1)(s-2)(s-3)}{(s^2 + 2.4s + 4)\left(1 + \frac{s}{\alpha}\right)} \frac{1}{s} = 4\alpha,$$

d.h. der Wert u_0 steigt proportional zu α . ■

Kapitel 7

Frequenzkennlinien und Ortskurven

7.1 Einführung

In diesem Abschnitt werden in aller Kürze die mathematischen Grundlagen wiederholt, die zum Zeichnen von Frequenzkennlinien und in weiterer Folge von Ortskurven erforderlich sind.

Komplexe Zahlen

Eine komplexe Zahl z kann in der kartesischen Darstellung

$$z = x + jy \quad \text{mit} \quad x, y \in \mathbb{R} \quad (7.1)$$

oder als komplexer Zeiger, also in der Polardarstellung

$$z = |z| e^{j \arg z} \quad \text{mit} \quad |z| = \sqrt{x^2 + y^2} \quad \text{und} \quad \tan(\arg z) = \frac{y}{x} \quad (7.2)$$

angegeben werden. Hierbei ist $x = \operatorname{Re}\{z\}$ der Realteil, $y = \operatorname{Im}\{z\}$ der Imaginärteil, $|z|$ ist der Betrag und $\arg z$ ist die Phase der komplexen Zahl z . Für das Produkt bzw. den Quotienten zweier komplexer Zahlen z_1 und z_2 gilt nun

$$z_1 z_2 = |z_1 z_2| e^{j \arg(z_1 z_2)} = |z_1| e^{j \arg z_1} |z_2| e^{j \arg z_2} = |z_1| |z_2| e^{j(\arg z_1 + \arg z_2)}, \quad (7.3)$$

bzw.

$$\frac{z_1}{z_2} = \left| \frac{z_1}{z_2} \right| e^{j \arg \frac{z_1}{z_2}} = \frac{|z_1| e^{j \arg z_1}}{|z_2| e^{j \arg z_2}} = \frac{|z_1|}{|z_2|} e^{j(\arg z_1 - \arg z_2)}. \quad (7.4)$$

Rechnen mit Logarithmen

Gegeben seien zwei positive, reelle Zahlen x und y . Berechnet man den (dekadischen) Logarithmus¹ des Produktes bzw. des Quotienten der beiden Zahlen, so gilt

$$\lg(xy) = \lg x + \lg y \quad \text{bzw.} \quad \lg \frac{x}{y} = \lg x - \lg y. \quad (7.5)$$

Für zwei beliebige reelle Zahlen x und y gilt weiters

$$\lg |x|^y = y \lg |x|. \quad (7.6)$$

¹Natürlich kann auch $\log_{10} x$ anstelle von $\lg x$ geschrieben werden.

7.2 Frequenzkennlinien

Die Frequenzkennlinien eines Systems mit der Übertragungsfunktion $G(s)$ erhält man, indem man Betrag und Phase des Frequenzganges $G(j\omega)$ jeweils über ω darstellt. Die Darstellung von $|G(j\omega)|$ über ω bezeichnet man als die *Amplitudenkennlinie* bzw. als den *Amplitudengang*. Üblicherweise wird dabei auf der Abszisse die Kreisfrequenz ω mit einer logarithmischen Skala oder $\lg \omega$ und auf der Ordinate der Betrag des Frequenzganges in Dezibel (dB), d.h.

$$|G(j\omega)|_{\text{dB}} = 20 \lg |G(j\omega)|. \quad (7.7)$$

aufgetragen. In der folgenden Tabelle sind wichtige Betragswerte in dB umgerechnet.

$ G(j\omega) $	0.01	0.1	1	10	100	2
$ G(j\omega) _{\text{dB}}$	-40	-20	0	20	40	6

Daraus können problemlos weitere Werte bestimmt werden, beispielsweise gilt

$$\begin{aligned} |\sqrt{2}|_{\text{dB}} &= |2^{\frac{1}{2}}|_{\text{dB}} \stackrel{(7.6)}{=} \frac{1}{2} |2|_{\text{dB}} = 3 \text{ dB} \\ |0.5|_{\text{dB}} &= \left| \frac{1}{2} \right|_{\text{dB}} \stackrel{(7.5)}{=} |1|_{\text{dB}} - |2|_{\text{dB}} = -6 \text{ dB}. \end{aligned}$$

Die Darstellung von $\arg G(j\omega)$ über ω wird *Phasenkennlinie* bzw. *Phasengang* des Systems genannt. Hierbei wird auf der Abszisse die Kreisfrequenz ω mit einer logarithmischen Skala oder $\lg \omega$ und auf der Ordinate die Phase von $G(j\omega)$ in Grad oder Radiant dargestellt. Aus den genannten Gründen spricht man oft auch von den *logarithmischen Frequenzkennlinien* eines Systems.

7.2.1 Normierte Darstellung

Der erste Schritt beim händischen Zeichnen der Frequenzkennlinien ist immer die Normierung der Übertragungsfunktion $G(s)$. Es wird dabei - so wie bisher immer - vorausgesetzt, dass $G(s)$ der Quotient zweier Polynome in s mit reellen Polynomkoeffizienten ist. Eine solche gebrochen rationale Funktion kann immer in die normierte Darstellung

$$G(s) = \frac{V p(s)}{s^\lambda q(s)} \quad \text{mit} \quad p(0) = q(0) = 1 \quad (7.8)$$

gebracht werden. Hierbei ist die reelle Konstante V der so genannte *Verstärkungsfaktor* des Systems, λ ist eine ganzzahlige Konstante und $p(s)$ bzw. $q(s)$ sind Polynome in s , bei denen die Koeffizienten zu s^0 auf 1 normiert sind.

B Gegeben seien die Übertragungsfunktionen

$$G(s) = \frac{s+3}{s^3+2s^2+2s} \quad \text{und} \quad H(s) = \frac{s-1}{s^2-3s+2}.$$

Für die zugehörigen normierten Darstellungen gilt offensichtlich

$$G(s) = \frac{3}{2} \frac{\left(1 + \frac{s}{3}\right)}{s \left(1 + s + \frac{s^2}{2}\right)} \quad \Rightarrow \quad V = \frac{3}{2}, \quad \lambda = 1, \quad p(s) = 1 + \frac{s}{3}, \quad q(s) = 1 + s + \frac{s^2}{2}$$

und

$$H(s) = -\frac{1}{2} \frac{(1-s)}{\left(1 - \frac{3}{2}s + \frac{s^2}{2}\right)} \quad \Rightarrow \quad V = -\frac{1}{2}, \quad \lambda = 0, \quad p(s) = 1 - s, \quad q(s) = 1 - \frac{3}{2}s + \frac{s^2}{2}.$$

■

Man beachte, dass $p(s)$ und $q(s)$ reelle Polynomkoeffizienten besitzen und somit rein reelle und/oder paarweise konjugiert komplexe Nullstellen haben. Aus diesem Grund können beide Polynome als Produkte von *Linearfaktoren* (=Polynome ersten Grades mit reellen Koeffizienten zur Berücksichtigung von reellen Nullstellen) und *quadratischen Faktoren* (=Polynome zweiten Grades mit reellen Koeffizienten zur Berücksichtigung von konjugiert komplexen Nullstellen) dargestellt werden.

B Gegeben seien die beiden (normierten) Polynome

$$p(s) = 1 - s^2 \quad \text{und} \quad q(s) = 1 + \frac{7}{13}s - \frac{5}{13}s^2 + \frac{1}{13}s^3.$$

Das Polynom $p(s)$ besitzt offensichtlich zwei rein reelle Nullstellen bei $s_1 = +1$ und $s_2 = -1$ und kann daher als Produkt zweier Linearfaktoren dargestellt werden, d.h.

$$p(s) = (1+s)(1-s).$$

Im Gegensatz dazu hat $q(s)$ eine reelle Nullstelle $s_1 = -1$ sowie ein Paar konjugiert komplexer Nullstellen $s_{2,3} = 3 \pm j2$ und kann somit als Produkt eines Linearfaktors und eines quadratischen Faktors dargestellt werden, d.h.

$$q(s) = (1+s)\left(1 - \frac{6}{13}s + \frac{1}{13}s^2\right).$$

■

7.2.2 Zerlegung von $G(s)$ in „elementare Bestandteile“

Für den Frequenzgang des Systems mit der Übertragungsfunktion $G(s)$ in normierter Darstellung (7.8) gilt nun

$$G(j\omega) = \frac{V}{(j\omega)^\lambda} \frac{p(j\omega)}{q(j\omega)}.$$

Aufgrund der Tatsache, dass der Betrag in dB angegeben wird, gilt nun gemäß (7.5) für den Betrag des Frequenzganges

$$|G(j\omega)|_{\text{dB}} = \left| \frac{V}{(j\omega)^\lambda} \frac{p(j\omega)}{q(j\omega)} \right|_{\text{dB}} = |V|_{\text{dB}} - |(j\omega)^\lambda|_{\text{dB}} + |p(j\omega)|_{\text{dB}} - |q(j\omega)|_{\text{dB}},$$

für die Phase gilt gemäß (7.3) und (7.4)

$$\arg G(j\omega) = \arg V - \arg (j\omega)^\lambda + \arg p(j\omega) - \arg q(j\omega).$$

Zur Ermittlung von $|G(j\omega)|_{\text{dB}}$ bzw. $\arg G(j\omega)$ müssen also die Frequenzgänge der einzelnen „Bestandteile“ von $G(j\omega)$ entsprechend zusammengesetzt werden. Dabei ist zu beachten, dass die Polynome $p(s)$ und $q(s)$, wie bereits erläutert, in Linearfaktoren und quadratische Faktoren zerlegt werden können. Somit sind insgesamt vier verschiedene „elementare Bestandteile“ zu berücksichtigen.

I Proportionalfaktor

Die Ermittlung von Betrag und Phase des Frequenzganges eines Proportionalgliedes

$$G(s) = V \quad \text{mit} \quad V \in \mathbb{R} \quad (7.9)$$

ist sehr geradlinig. Gemäß (7.7) ergibt sich der Betrag in dB zu

$$|G(j\omega)|_{\text{dB}} = |V|_{\text{dB}} = 20 \lg |V|, \quad (7.10)$$

für die Phase des Frequenzganges gilt

$$\arg G(j\omega) = \begin{cases} 0^\circ & \text{für } V > 0 \\ -180^\circ & \text{für } V < 0 \end{cases}. \quad (7.11)$$

II Term der Form $\left(\frac{1}{s^\lambda}\right)$

Der Frequenzgang der Übertragungsfunktion

$$G(s) = \frac{1}{s^\lambda} \quad \text{mit} \quad \lambda \dots \text{ganzzahlig} \quad (7.12)$$

ist durch

$$G(j\omega) = \frac{1}{(j\omega)^\lambda}$$

gegeben. Für den Betrag des Frequenzganges in dB ergibt sich somit

$$|G(j\omega)|_{\text{dB}} = 20 \lg \left| \frac{1}{(j\omega)^\lambda} \right| \stackrel{(7.5)}{=} -20 \lg (j\omega)^\lambda \stackrel{(7.6)}{=} -20\lambda \lg |(j\omega)| \stackrel{(7.2)}{=} -20\lambda \lg \omega. \quad (7.13)$$

Trägt man den Betrag (7.13) über $\lg \omega$ auf, so erhält man eine Gerade, die die 0 dB-Linie bei $\omega = 1$ schneidet und eine Steigung von (-20λ) dB pro Dekade, also pro Frequenz-Verzehnfachung besitzt. In Bild 7.1 sind Betragskennlinien für verschiedene Werte von λ dargestellt, wobei die Steigungen der Geraden zusätzlich angegeben sind.

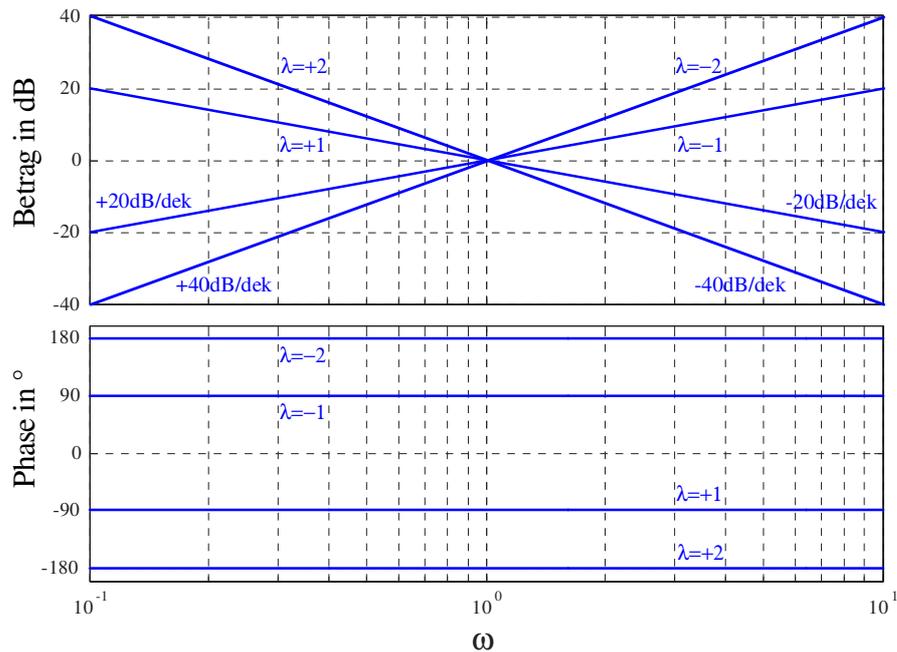


Bild 7.1: Frequenzkennlinien von $\frac{1}{s^\lambda}$ für $\lambda = -2, -1, +1, +2$

Für die Phase des Frequenzganges gilt

$$\arg G(j\omega) = \arg \frac{1}{(j\omega)^\lambda} \stackrel{(7.3)}{=} -\lambda \arg(j\omega) = -\lambda \cdot 90^\circ, \quad (7.14)$$

d.h. die Phase besitzt einen frequenzunabhängigen, konstanten Wert, was auch Bild 7.1 zu entnehmen ist.

III Linearfaktor

Der Frequenzgang eines Linearfaktors der Form

$$G(s) = 1 + \frac{s}{\omega_k} \quad \text{mit} \quad \omega_k \in \mathbb{R} \quad (7.15)$$

ist gegeben durch

$$G(j\omega) = 1 + j \frac{\omega}{\omega_k}. \quad (7.16)$$

Hierbei ist ω_k ein konstanter, reeller Parameter. Mit (7.2) und (7.7) findet man für den Betrag des Frequenzganges

$$|G(j\omega)|_{\text{dB}} = 20 \lg \sqrt{1 + \left(\frac{\omega}{\omega_k}\right)^2} = \begin{cases} 0 & \text{für } \omega \ll |\omega_k| \\ 20 \lg \sqrt{2} \approx 3 & \text{für } \omega = |\omega_k| \\ 20 \lg \frac{\omega}{|\omega_k|} = 20 \lg \omega - 20 \lg |\omega_k| & \text{für } \omega \gg |\omega_k| \end{cases}, \quad (7.17)$$

siehe hierzu auch Bild 7.2, wo $\omega_k = 1$ gesetzt wurde. Häufig - vor allem aber beim händischen Zeichnen der Frequenzkennlinien - wird vereinfachend der *asymptotische* Verlauf des Amplitudenganges dargestellt. Dabei wird der Betrag für $\omega \leq |\omega_k|$ gleich 0 dB gesetzt, ab $\omega = |\omega_k|$ entspricht die Betragskennlinie einer Gerade mit einer Steigung von 20 dB pro Dekade. Dadurch entsteht bei $\omega = |\omega_k|$ ein Knick in der Betragskennlinie, weshalb man ω_k auch *Knickfrequenz* nennt. In Bild 7.2 ist der asymptotische Verlauf des Amplitudenganges rot strichliert dargestellt. Man erkennt, dass man bei der asymptotischen Darstellung des Betrages an der Stelle $\omega = |\omega_k|$ einen Fehler von 3 dB in Kauf nehmen muss.

Für den Verlauf der Phasenkennlinie ist, im Gegensatz zur Betragskennlinie, das Vorzeichen von ω_k zu berücksichtigen, gemäß (7.2) gilt nämlich

$$\arg G(j\omega) = \arctan \frac{\omega}{\omega_k}.$$

Daraus folgt unmittelbar

$$\arg G(j\omega) = \begin{cases} 0^\circ & \text{für } \omega \ll |\omega_k| \\ 45^\circ \operatorname{sgn} \omega_k & \text{für } \omega = |\omega_k| \\ 90^\circ \operatorname{sgn} \omega_k & \text{für } \omega \gg |\omega_k| \end{cases}. \quad (7.18)$$

Der exakte Verlauf der Phasenkennlinie ist in Bild 7.2 für $\omega_k = 1$ dargestellt. Man beachte, dass es auch die Möglichkeit gibt, einen linear interpolierten Verlauf der Phasenkennlinie zu zeichnen. Hierbei wird für $\omega \leq \frac{1}{10} |\omega_k|$ die Phase gleich 0° gesetzt, für $\omega \geq 10 |\omega_k|$ wird die Phase gleich $90^\circ \operatorname{sgn} \omega_k$ gesetzt. Im Bereich zwischen $\omega = \frac{1}{10} |\omega_k|$ und $\omega = 10 |\omega_k|$ wird die Phase linear interpoliert, der entsprechende Verlauf ist in Bild 7.2 rot strichliert dargestellt. Im Gegensatz zur asymptotischen Darstellung des Betragsganges wird die interpolierte Darstellung des Phasenganges allerdings nur eher selten angewandt.

Bei einem Linearfaktor im Nenner einer Übertragungsfunktion ändern sich gemäß (7.5) die Vorzeichen von Betrags- und Phasenkennlinie.

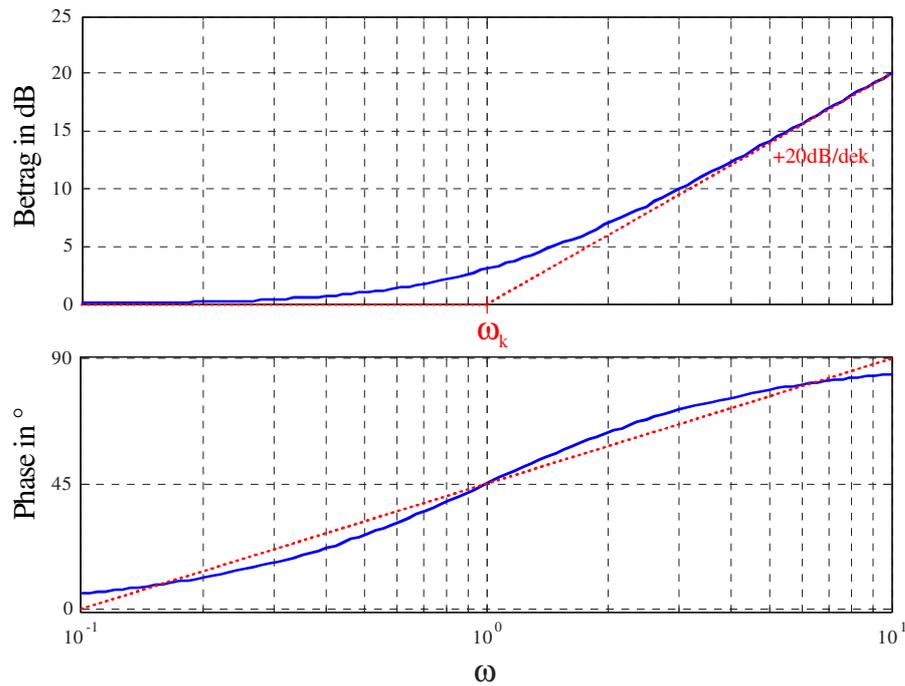


Bild 7.2: Frequenzkennlinien eines Linearfaktors $\left(1 + \frac{s}{\omega_k}\right)$ mit der Knickfrequenz $\omega_k = 1$.

IV Quadratischer Faktor

Der Frequenzgang eines quadratischen Faktors

$$G(s) = 1 + 2\zeta \frac{s}{\omega_k} + \left(\frac{s}{\omega_k}\right)^2 \quad \text{mit} \quad \omega_k, \zeta \in \mathbb{R} \quad \text{und} \quad 0 \leq \zeta < 1 \quad (7.19)$$

ist gegeben durch

$$G(j\omega) = \left[1 - \left(\frac{\omega}{\omega_k}\right)^2\right] + j 2\zeta \frac{\omega}{\omega_k}. \quad (7.20)$$

Für den Betrag des Frequenzganges gilt somit

$$|G(j\omega)|_{\text{dB}} = 20 \log \sqrt{\left[1 - \left(\frac{\omega}{\omega_k}\right)^2\right]^2 + \left(2\zeta \frac{\omega}{\omega_k}\right)^2} = \begin{cases} 0 & \text{für } \omega \ll |\omega_k| \\ 20 \lg 2\zeta & \text{für } \omega = |\omega_k| \\ 20 \lg \left(\frac{\omega}{\omega_k}\right)^2 = 40 \lg \omega - 40 \lg |\omega_k| & \text{für } \omega \gg |\omega_k| \end{cases}, \quad (7.21)$$

d.h. in der asymptotischen Darstellung beträgt der Betrag für „niedrige“ Frequenzen 0 dB, für „hohe“ Frequenzen nimmt der Betrag um 40 dB pro Dekade zu. Der Wert des Betrages an der Stelle $\omega = |\omega_k|$ hängt offensichtlich vom Wert des Parameters ζ ab, siehe auch Bild 7.3, wo die Betragskennlinie für verschiedene Werte von ζ dargestellt ist.

Für die Phase des Frequenzganges (7.20) ist nicht nur das Vorzeichen von ω_k , sondern auch der Wert des Parameters ζ von entscheidender Bedeutung. Analog zum Linearfaktor gilt für die Phase

$$\arg G(j\omega) = \begin{cases} 0^\circ & \text{für } \omega \ll |\omega_k| \\ 90^\circ \operatorname{sgn} \omega_k & \text{für } \omega = |\omega_k| \\ 180^\circ \operatorname{sgn} \omega_k & \text{für } \omega \gg |\omega_k| \end{cases} . \quad (7.22)$$

Wie man Bild 7.3 entnehmen kann, ergibt sich für kleine Werte von ζ ein „schärferer“ Übergang der Phase von 0° zu $180^\circ \operatorname{sgn} \omega_k$, für $\zeta = 0$ ergibt sich ein Phasensprung an der Stelle $\omega = |\omega_k|$.

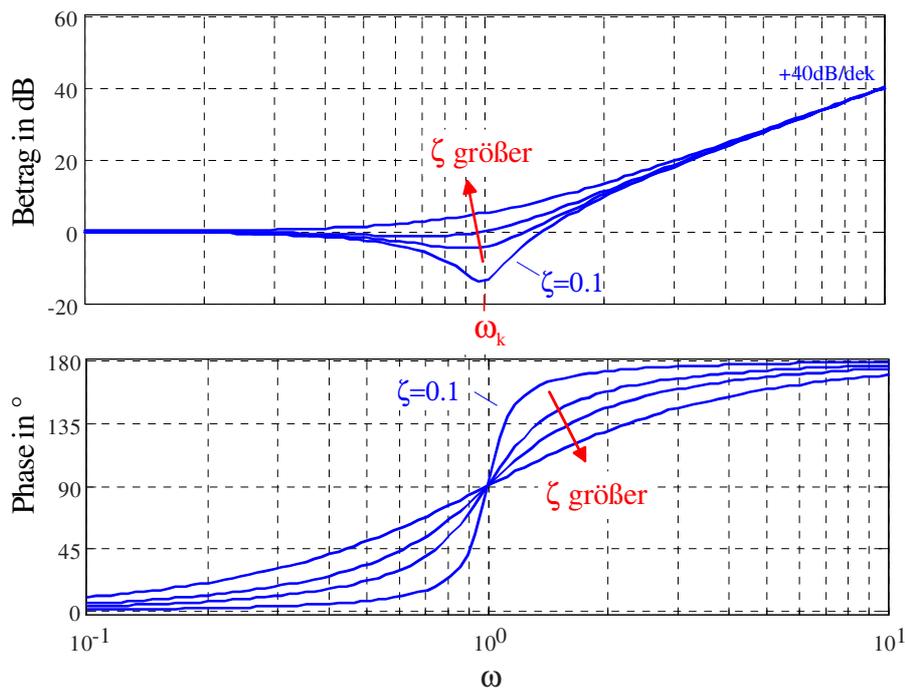


Bild 7.3: Frequenzkennlinien eines quadratischen Faktors $\left(1 + 2\zeta \frac{s}{\omega_k} + \frac{s^2}{\omega_k^2}\right)$ für $\zeta = 0.1, 0.5, 0.7, 0.9$ und $\omega_k = 1$

Bei einem quadratischen Faktor im Nenner einer Übertragungsfunktion ändern sich gemäß (7.5) natürlich die Vorzeichen von Betrags- und Phasenkennlinie.

Vorgangsweise beim Zeichnen der (asymptotischen) Frequenzkennlinien

- Normierung von $G(s)$.
- Einzeichnen der Betrags- und Phasenkennlinie von $\frac{1}{s^\lambda}$.
- Aufspaltung von $p(s)$ und $q(s)$ in Linearfaktoren und/oder quadratische Faktoren.
- Die Betragsgänge der zu $p(s)$ gehörigen Linearfaktoren "knicken" bei der Knickfrequenz $|\omega_k|$ um $+20$ dB/Dekade nach „oben“, Linearfaktoren von $q(s)$ nach „unten“. Ausgehend von 0° für $\omega \ll |\omega_k|$ strebt die Phase von zu $p(s)$ gehörigen Linearfaktoren für $\omega \gg |\omega_k|$ zum Wert $90^\circ \cdot \text{sgn} \omega_k$, an der Stelle $\omega = |\omega_k|$ lautet die Phase $45^\circ \cdot \text{sgn} \omega_k$. Linearfaktoren von $q(s)$ sind analog, allerdings mit umgekehrten Vorzeichen zu berücksichtigen.
- Quadratische Faktoren sind analog zu den Linearfaktoren zu berücksichtigen, die Steigung beträgt jedoch 40 dB/Dekade und die Phase strebt zum Wert $180^\circ \cdot \text{sgn} \omega_k$, an der Stelle $\omega = |\omega_k|$ lautet die Phase $90^\circ \cdot \text{sgn} \omega_k$.
- Betrags- und Phasenkennlinien der einzelnen „Bestandteile“ von $G(s)$ werden additiv zusammengefügt.
- Die Betragskennlinie ist um den Wert $|V|_{dB}$ zu verschieben. Bei negativem V ist die Phasenkennlinie ebenfalls zu verschieben bzw. die Beschriftung anzupassen.

Die oben beschriebene Prozedur kann etwas vereinfacht werden, indem man im Betragsgang zunächst bei $\omega = 1$ den Wert $|V|_{dB}$ einzeichnet und mit derjenigen Asymptote „anvisiert“, die für $\omega \rightarrow 0$ die Betragskennlinie von $G(s)$ dominiert². Dadurch kann man die sonst erforderliche Verschiebung der Betragskennlinien (siehe letzter Punkt oben) um $|V|_{dB}$ umgehen. Die Phasenkennlinie kann man prinzipiell mit einem so genannten Phasenlineal auch händisch sehr genau ermitteln.

B Es sollen die zur Übertragungsfunktion

$$G(s) = \frac{s+1}{s(s+10)}$$

gehörigen Frequenzkennlinien gezeichnet werden. Hierzu wird $G(s)$ zunächst in die normierte Form gebracht, d.h.

$$G(s) = \frac{1}{10} \frac{(1+s)}{s \left(1 + \frac{s}{10}\right)}$$

Das bedeutet, dass der Verstärkungsfaktor V , ein Term der Form $\frac{1}{s}$ und zwei Linearfaktoren zu berücksichtigen sind. Die entsprechenden Betrags- und Phasenkennlinien sind in den Bildern 7.4 und 7.5 dargestellt. ■

²Man beachte, dass $\omega = 0$ aufgrund der logarithmischen Skalierung der Frequenz nicht dargestellt werden kann.

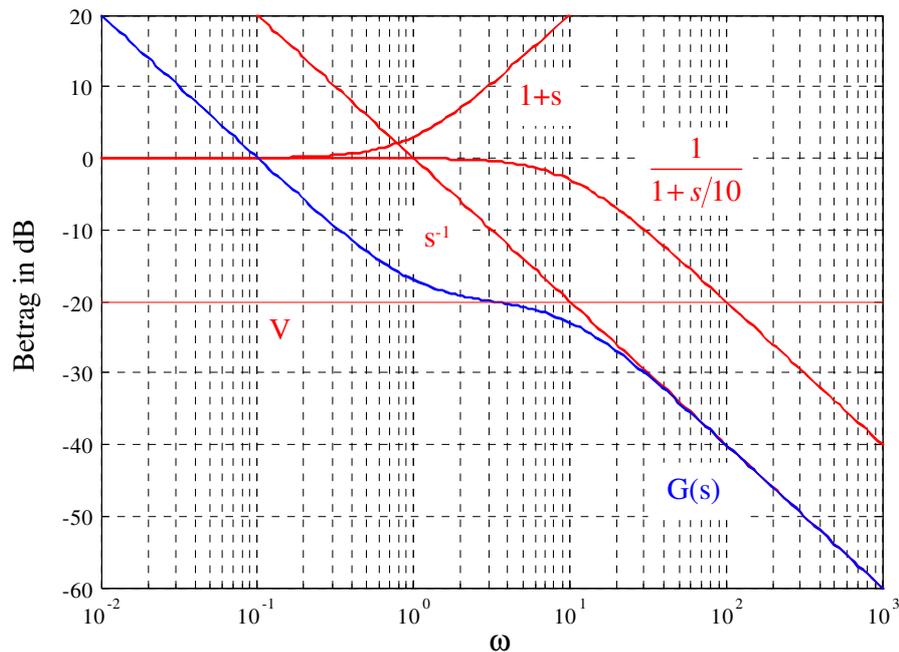


Bild 7.4: Betragskennlinie von $G(s) = \frac{s+1}{s(s+10)}$

7.3 Frequenzgangs-Ortskurve

Die Frequenzgangs-Ortskurve oder kurz Ortskurve eines Systems mit der Übertragungsfunktion $G(s)$ ist die graphische Darstellung des Frequenzganges $G(j\omega)$ in der komplexen Ebene. Für einfache Übertragungsfunktionen kann der Verlauf der Ortskurve direkt aus $G(j\omega)$ ermittelt werden.

B Es soll die Ortskurve zur Übertragungsfunktion

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

gezeichnet werden. Hierzu wird zunächst der Frequenzgang

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

angeschrieben, offensichtlich gilt

$$G(j0) = 1 \quad \text{und} \quad G(j\omega) = 0 \quad \text{für} \quad \omega \rightarrow \infty,$$

d.h. die Ortskurve beginnt für $\omega = 0$ bei 1 und endet für $\omega \rightarrow \infty$ im Koordinatenursprung. Berechnet man nun den Abstand der Ortskurve zum Punkt $\frac{1}{2}$, so erhält man

$$\left| \frac{1}{2} - G(j\omega) \right| = \left| \frac{1}{2} - \frac{1-j\omega}{1+\omega^2} \right| = \frac{1}{2}.$$

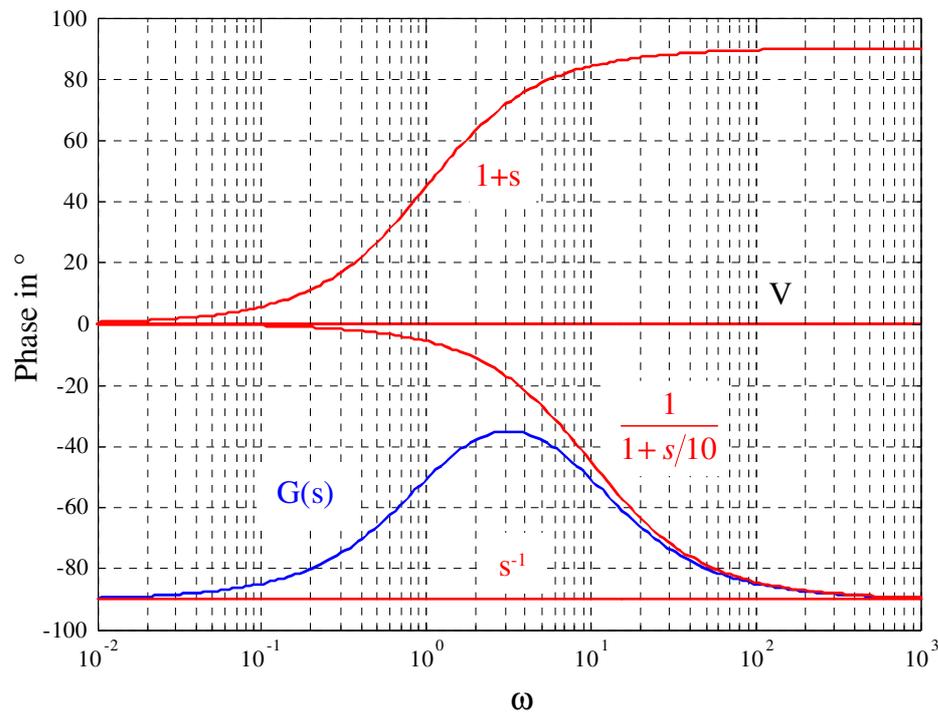


Bild 7.5: Phasenkenmlinie von $G(s) = \frac{s+1}{s(s+10)}$

Das bedeutet, dass die Ortskurve dem in Bild 7.6 dargestellten Halbkreis entspricht. Der eingezeichnete Pfeil zeigt dabei in Richtung wachsender Werte von ω .

■

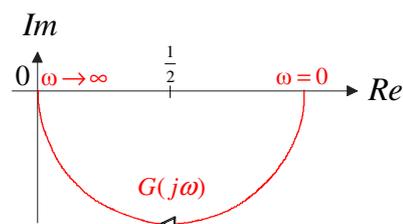


Bild 7.6: Ortskurve zu $G(s) = \frac{1}{s+1}$

Bei „komplizierteren“ Übertragungsfunktionen werden üblicherweise zunächst die Frequenzkennlinien des Systems gezeichnet und daraus die Ortskurve skizziert. Hierzu liest man für interessante Frequenzwerte Betrag und Phase des Frequenzganges aus den Frequenzkennlinien ab und zeichnet die entsprechenden komplexen Zeiger in die komplexe Ebene ein. Daraus ergibt sich eine Approximation des Ortskurvenverlaufes.

Man beachte, dass die Ortskurve manchmal für $-\infty < \omega < \infty$, also auch für negative Werte von ω gezeichnet wird. Dabei ist zu beachten, dass aufgrund von

$$|G(-j\omega)| = |G(j\omega)| \quad \text{und} \quad \arg G(-j\omega) = -\arg G(j\omega)$$

der Teil der Ortskurve für $-\infty < \omega \leq 0$ aus dem Teil der Ortskurve für $0 \leq \omega < \infty$ durch Spiegelung an der reellen Achse hervorgeht.

B Es soll die zur Übertragungsfunktion

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

gehörige Ortskurve gezeichnet werden. Wie man aus den in Bild 7.7, links dargestellten Frequenzkennlinien erkennt, beträgt die Phase von $G(j\omega)$ für sehr kleine Frequenzwerte -90° , der zugehörige Betrag ist sehr groß und geht für $\omega = 0$ gegen Unendlich. Mit wachsenden Frequenzen wird der Betrag immer kleiner und die Phase strebt gegen -180° , bei $\omega = 1$ beträgt $|G(j\omega)|_{dB} = -3$, d.h. $|G(j\omega)| = \frac{1}{\sqrt{2}}$ und $\arg G(j\omega) = -135^\circ$. Mit diesen Informationen kann man die gesuchte Ortskurve skizzieren, sie ist in Bild 7.7, rechts dargestellt.

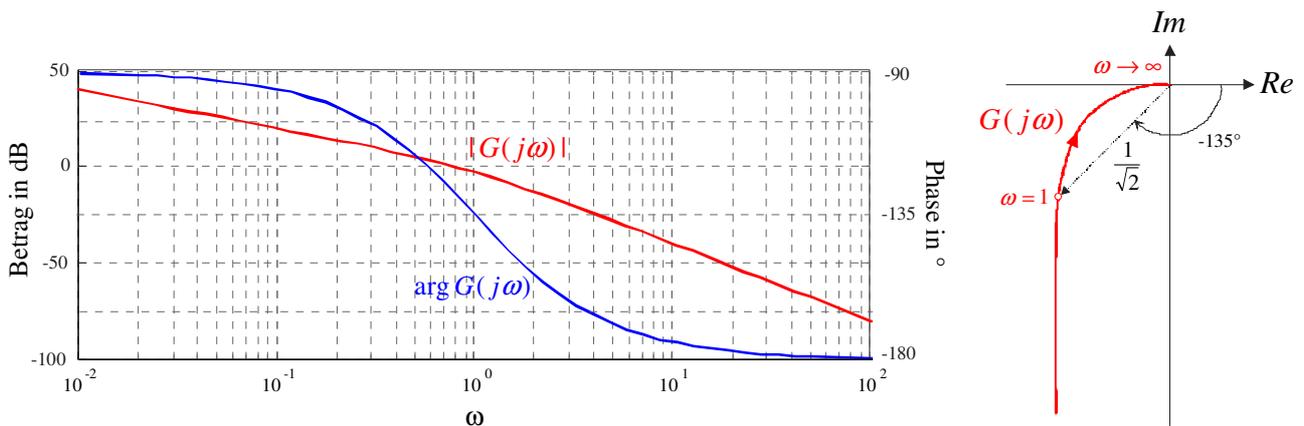


Bild 7.7: Frequenzkennlinien zu $G(s) = \frac{1}{s(s+1)}$ und zugehörige Ortskurve

■

Kapitel 8

Das Nyquist-Kriterium

8.1 Einführung

Ausgangspunkt der Überlegungen ist der in Bild 8.1 dargestellte Standardregelkreis. Will man die Führungsübertragungsfunktion

$$T(s) = \frac{R(s)P(s)}{1 + R(s)P(s)} \quad (8.1)$$

des Regelkreises hinsichtlich ihrer BIBO-Stabilität untersuchen, so ist zu überprüfen, ob das Nennerpolynom von $T(s)$ ein Hurwitzpolynom ist. Zur Kontrolle, ob alle Nullstellen des

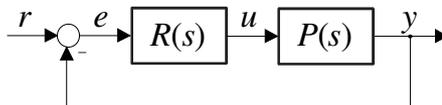


Bild 8.1: Standardregelkreis

Nennerpolynoms von $T(s)$ in der linken, offenen komplexen Ebene liegen, gibt es natürlich mehrere Möglichkeiten. Die naheliegendste Möglichkeit besteht darin, die Pole von $T(s)$ explizit zu berechnen, was meistens nur auf numerischem Weg möglich ist. Dieser Zugang zur Beantwortung des Stabilitätsproblems scheidet allerdings meistens, wenn die Koeffizienten des Nennerpolynoms von $T(s)$ beispielsweise Funktionen von vorgebbaren Reglerparametern sind. In diesem Fall können Stabilitätskriterien, wie z.B. das Routh-Schema oder das Hurwitz-Kriterium angewandt werden, eine explizite Ermittlung der Polstellen ist dort nicht erforderlich.

Im Gegensatz zu den genannten numerischen Kriterien ist das Nyquist-Kriterium¹ ein *graphisches Stabilitätskriterium*, bei dem aus dem Verlauf der Ortskurve des *offenen Kreises*, also

$$L(j\omega) = R(j\omega)P(j\omega) \quad (8.2)$$

¹benannt nach dem in Schweden geborenen, amerikanischen Physiker Harry Nyquist (1889 - 1976)

auf die *Stabilität des geschlossenen Kreises* geschlossen werden kann. Ein wesentlicher Vorzug des Nyquist-Kriteriums besteht darin, dass neben dem Verlauf der Ortskurve $L(j\omega)$, z.B. in Form von Messwerten, nur wenige Informationen über die Übertragungsfunktion $L(s)$ benötigt werden.

Man beachte, dass der deutsche Elektrotechniker Felix Strecker (1892-1951) bereits 1930, also 2(!) Jahre vor Harry Nyquist, ein ähnliches Stabilitätskriterium vorgeschlagen hat. Aus diesem Grund findet man in einigen deutschen Literaturstellen auch die Bezeichnung **Strecker-Nyquist-Kriterium**, siehe z.B. [?].

Für die Herleitung des Kriteriums wird der im nächsten Abschnitt erläuterte Begriff der stetigen Winkeländerung einer Ortskurve benötigt.

8.2 Stetige Winkeländerung einer Ortskurve

Zunächst wird vereinfachend die zur Übertragungsfunktion

$$F(s) = s - \beta \quad \text{mit} \quad \beta \in \mathbb{C} \quad (8.3)$$

gehörige Ortskurve genauer untersucht. Hierfür wird der Frequenzgang

$$F(j\omega) = j\omega - \beta$$

in der komplexen Ebene dargestellt, wobei der Frequenzparameter ω hier ausnahmsweise Werte von $-\infty$ bis $+\infty$ durchläuft. Aus der resultierenden Ortskurve wird nun die so genannte *stetige Winkeländerung* von $F(j\omega)$, also

$$\Delta_{-\infty}^{+\infty} \arg F(j\omega) = \Delta_{-\infty}^{+\infty} \arg (j\omega - \beta)$$

ermittelt. Darunter versteht man die Änderung der Phasenlage des komplexen Zeigers $F(j\omega)$, wenn ω Werte von $-\infty$ bis $+\infty$ durchläuft. Phasensprünge, also Unstetigkeiten im Phasenverlauf, werden *nicht* mitgezählt. In Bild 8.2 ist die Ortskurve $F(j\omega)$ graphisch dargestellt, wobei die drei Fälle $\operatorname{Re} \beta < 0$, $\operatorname{Re} \beta = 0$ und $\operatorname{Re} \beta > 0$ unterschieden werden. Weiters ist der komplexe Zeiger $F(j\omega_*)$ eingezeichnet, wobei ω_* ein beliebiger Frequenzwert ist. Man kann leicht erkennen, dass für die stetige Winkeländerung von $F(j\omega)$ gilt:

$$\Delta_{-\infty}^{+\infty} \arg F(j\omega) = \Delta_{-\infty}^{+\infty} \arg (j\omega - \beta) = \begin{cases} -\pi & \text{für } \operatorname{Re} \beta > 0 \\ 0 & \text{für } \operatorname{Re} \beta = 0 \\ +\pi & \text{für } \operatorname{Re} \beta < 0 \end{cases} \quad (8.4)$$

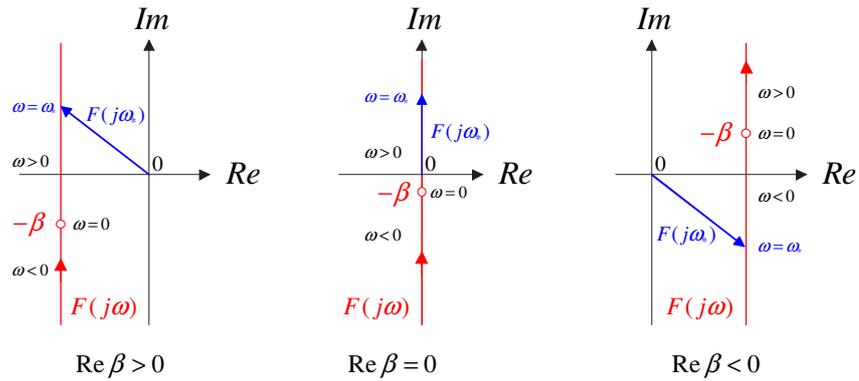


Bild 8.2: Zur Bestimmung der stetigen Winkeländerung von $F(j\omega) = j\omega - \beta$

Diese Erkenntnisse können nun sehr geradlinig auf eine (teilerfremde, also gekürzte) Übertragungsfunktion der Form

$$F(s) = K \frac{\prod_{i=1}^m (s - \beta_i)}{\prod_{i=1}^n (s - \alpha_i)} \quad \text{mit} \quad K \in \mathbb{R}, \quad \alpha_i, \beta_i \in \mathbb{C} \quad (8.5)$$

übertragen werden. Hierbei wird vorausgesetzt, dass die Übertragungsfunktion m_l Nullstellen mit negativem Realteil, m_a Nullstellen mit verschwindendem Realteil und m_r Nullstellen mit positivem Realteil besitzt, d.h.

$$m = m_l + m_a + m_r. \quad (8.6)$$

Analoges gilt für die n Pole von $F(s)$, d.h.

$$n = n_l + n_a + n_r. \quad (8.7)$$

Weiters wird wieder vorausgesetzt, dass Nullstellen bzw. Pole reell und/oder paarweise konjugiert komplex auftreten. Für die Phase des zugehörigen Frequenzganges

$$F(j\omega) = K \frac{\prod_{i=1}^m (j\omega - \beta_i)}{\prod_{i=1}^n (j\omega - \alpha_i)}$$

gilt bekanntlich

$$\arg F(j\omega) = \arg K + \sum_{i=1}^m \arg (j\omega - \beta_i) - \sum_{i=1}^n \arg (j\omega - \alpha_i)$$

und in weiterer Folge auch

$$\Delta_{-\infty}^{+\infty} \arg F(j\omega) = \Delta_{-\infty}^{+\infty} \arg K + \sum_{i=1}^m \Delta_{-\infty}^{+\infty} \arg (j\omega - \beta_i) - \sum_{i=1}^n \Delta_{-\infty}^{+\infty} \arg (j\omega - \alpha_i).$$

Mit (8.4) kann nun unmittelbar

$$\Delta_{-\infty}^{+\infty} \arg F(j\omega) = 0 + m_a \cdot 0 + m_l \pi - m_r \pi - n_a \cdot 0 - n_l \pi + n_r \pi$$

gefolgert werden. Eliminiert man nun mittels (8.6) und (8.7) m_l und n_l , so findet man nach kurzer Rechnung

$$\Delta_{-\infty}^{+\infty} \arg F(j\omega) = (m - n) \pi - (m_a + 2m_r) \pi + (n_a + 2n_r) \pi. \quad (8.8)$$

Darüber hinaus kann man die Tatsache ausnützen, dass der „negative Ast“ der Ortskurve $F(j\omega)$ (d.h. für negative Werte von ω) aus dem „positiven Ast“ von $F(j\omega)$ durch Spiegelung an der reellen Achse hervorgeht. Daraus folgt unmittelbar, dass die stetige Winkeländerung von $F(j\omega)$, wenn ω (nichtnegative) Werte von 0 bis $+\infty$ durchläuft genau die Hälfte von $\Delta_{-\infty}^{+\infty} \arg F(j\omega)$ beträgt, d.h.

$$\Delta \arg F(j\omega) := \Delta_0^{+\infty} \arg F(j\omega) = \frac{1}{2} \Delta_{-\infty}^{+\infty} \arg F(j\omega). \quad (8.9)$$

Damit gilt aber auch

$$\Delta \arg F(j\omega) = (m - n) \frac{\pi}{2} - (m_a + 2m_r) \frac{\pi}{2} + (n_a + 2n_r) \frac{\pi}{2}. \quad (8.10)$$

B Gegeben sei die Übertragungsfunktion

$$F(s) = \frac{s^2 + 0.1s + 1}{s^3 - s}$$

mit den Nullstellen und Polen

$$\beta_{1,2} = -0.05 \pm j 0.9987, \quad \alpha_1 = 0, \quad \alpha_2 = +1, \quad \alpha_3 = -1,$$

d.h. $m = 2$, $n = 3$ sowie $m_a = m_r = 0$ und $n_a = n_r = 1$. Gemäß (8.10) gilt für die stetige Winkeländerung

$$\Delta \arg F(j\omega) = (2 - 3) \frac{\pi}{2} - (0 + 0) \frac{\pi}{2} + (1 + 2) \frac{\pi}{2} = \pi. \quad \blacksquare$$

8.3 Formulierung des Nyquist-Kriteriums

Wie bereits in der Einführung angedeutet wurde, wird beim Nyquist-Kriterium mit der Ortskurve des offenen Kreises operiert. Die Übertragungsfunktion $L(s) = R(s)P(s)$ des offenen Kreises ist der Quotient der teilerfremden Polynome $\mu(s)$ und $\nu(s)$, d.h.

$$L(s) = \frac{\mu(s)}{\nu(s)}, \quad (8.11)$$

wobei vorausgesetzt wird, dass $L(s)$ eine realisierbare Übertragungsfunktion ist, d.h.

$$\text{Grad } \mu(s) \leq \text{Grad } \nu(s). \quad (8.12)$$

Zur Überprüfung der Stabilität des geschlossenen Kreises muss das Nennerpolynom von

$$T(s) = \frac{L(s)}{1 + L(s)} = \frac{\mu(s)}{\mu(s) + \nu(s)}$$

untersucht werden. Der geschlossene Kreis ist genau dann BIBO-stabil, wenn das Polynom $\mu(s) + \nu(s)$ ein Hurwitzpolynom ist. Man beachte, dass die Übertragungsfunktion

$$F(s) := 1 + L(s) = \frac{\mu(s) + \nu(s)}{\nu(s)} \quad (8.13)$$

das zu untersuchende Polynom als Zählerpolynom besitzt, d.h. der geschlossene Regelkreis ist genau dann BIBO-stabil, wenn das *Zählerpolynom* von $F(s)$ ein Hurwitzpolynom ist. Wegen (8.12) und (8.13) gilt sicher

$$m = n,$$

und aus (8.10) folgt nun unmittelbar

$$\Delta \arg F(j\omega) = \Delta \arg \{1 + L(j\omega)\} = -(m_a + 2m_r) \frac{\pi}{2} + (n_a + 2n_r) \frac{\pi}{2}.$$

Der geschlossene Regelkreis ist genau dann BIBO-stabil, wenn $m_a = m_r = 0$ gilt, d.h. wenn die Bedingung (**Nyquist-Kriterium**)

$$\Delta \arg \{1 + L(j\omega)\} = (n_a + 2n_r) \frac{\pi}{2} \quad (8.14)$$

erfüllt ist.

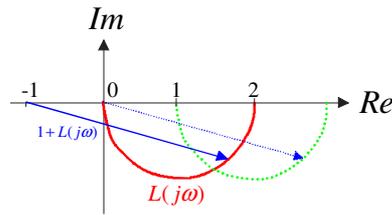
Das bedeutet, dass $T(s)$ dann, und nur dann BIBO-stabil ist, wenn die stetige Winkeländerung der Ortskurve $\{1 + L(j\omega)\}$ dem (nichtnegativen) Ausdruck auf der rechten Seite von (8.14) entspricht. Es ist also die stetige Winkeländerung des komplexen Zeigers $\{1 + L(j\omega)\}$ zu untersuchen. Diesen Zeiger erhält man auch, wenn man den Punkt (-1) mit $L(j\omega)$ verbindet, eine Verschiebung der Ortskurve $L(j\omega)$ nach „rechts“ ist somit nicht erforderlich, siehe hierzu auch Bild 8.3. Man beachte auch, dass aufgrund von (8.13) die Polstellen von $F(s)$ identisch sind mit den Polstellen von $L(s)$, d.h. die nichtnegativen ganzen Zahlen n_a bzw. n_r geben an, wieviele Pole des offenen Kreises auf bzw. rechts der imaginären Achse liegen.

B Gegeben sei ein Standardregelkreis mit

$$L(s) = \frac{K}{s(s+1)} \quad \text{wobei} \quad K \in \mathbb{R},$$

d.h. $n_a = 1$ und $n_r = 0$. Es soll mit Hilfe des Nyquist-Kriteriums derjenige Wertebereich von K ermittelt werden, für den der geschlossene Regelkreis BIBO-stabil ist, d.h. gemäß (8.14) muss für die stetige Winkeländerung von $\{1 + L(j\omega)\}$ die Bedingung

$$\Delta \arg \{1 + L(j\omega)\} \stackrel{!}{=} \frac{\pi}{2}$$

Bild 8.3: Zur Bildung des komplexen Zeigers $1 + L(j\omega)$

gelten. Zunächst wird die in Bild 8.4, links dargestellte Ortskurve $L(j\omega)$ für $K = 1$ skizziert, Details dazu findet man im Kapitel über Ortskurven. Unter Annahme positiver Werte für K bewirkt eine Variation von K eine Skalierung der Ortskurve, d.h. die Form der Ortskurve bleibt gleich, die Ortskurve wird aber für $K > 1$ „aufgeblasen“ und für $K < 1$ „geschrumpft“. Offensichtlich gilt für beliebige positive Werte von K tatsächlich

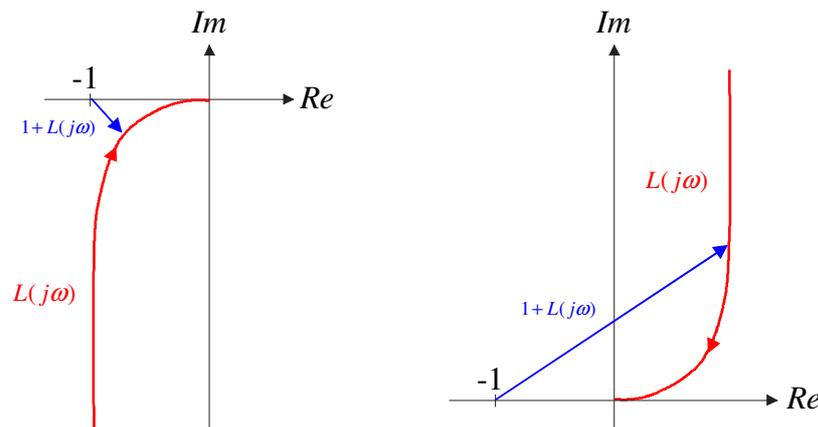


Bild 8.4: Ortskurven von $L(s) = \frac{K}{s(s+1)}$ für $K = +1$ (linkes Bild) und $K = -1$ (rechtes Bild)

$$\Delta \arg \{1 + L(j\omega)\} = \frac{\pi}{2},$$

d.h. das Nyquist-Kriterium ist erfüllt. Für negative Werte von K ändert sich die Phase von $L(j\omega)$ um 180° , die entsprechende Ortskurve für $K = -1$ ist in Bild 8.4, rechts dargestellt. Für beliebige negative Werte von K kann nun die stetige Winkeländerung

$$\Delta \arg \{1 + L(j\omega)\} = -\frac{\pi}{2},$$

abgelesen werden, d.h. das Nyquist-Kriterium ist *nicht* erfüllt. Der zulässige Wertebereich für den reellen Parameter K lautet somit

$$0 < K < \infty.$$

Dieses Ergebnis kann auch analytisch bestätigt werden. Für den geschlossenen Regelkreis gilt

$$T(s) = \frac{L(s)}{1 + L(s)} = \frac{K}{s^2 + s + K},$$

d.h. das Nennerpolynom ist genau für den oben angegebenen Wertebereich von K ein Hurwitzpolynom. ■

8.4 Vereinfachtes Schnittpunktkriterium

Das Nyquist-Kriterium (8.14) wird besonders einfach, wenn die (realisierbare) Übertragungsfunktion $L(s)$ vom so genannten *einfachen Typ* ist, d.h.

1. Der Verstärkungsfaktor V des offenen Kreises ist positiv.
2. Alle Pole von $L(s)$ haben einen negativen Realteil bis auf möglicherweise einen Pol bei $s = 0$.
3. Die Betragskennlinie von $L(j\omega)$ besitzt genau einen Schnittpunkt mit der 0 dB Linie und verläuft für $\omega \rightarrow \infty$ unter dieser.

Ist $L(s)$ vom einfachen Typ, dann gilt auf jeden Fall $n_r = 0$ und $n_a = 0$ bzw. $n_a = 1$, d.h. aus (8.14) folgt

$$\Delta \arg \{1 + L(j\omega)\} = \begin{cases} 0 & \text{wenn } L(s) \text{ keinen Pol bei } s = 0 \text{ besitzt } (n_a = 0) \\ \frac{\pi}{2} & \text{wenn } L(s) \text{ einen Pol bei } s = 0 \text{ besitzt } (n_a = 1) \end{cases}$$

Dies bedeutet, dass die Phase von $L(j\omega)$ für diejenige (eindeutige, s.o.) Frequenz ω_c , bei der $L(j\omega)$ den Einheitskreis schneidet, größer sein muss als -180° , d.h.

$$\arg L(j\omega_c) > -180^\circ. \quad (8.15)$$

Zur endgültigen, prägnanten Formulierung dieses so genannten vereinfachten Schnittpunktkriteriums ist die Einführung der Begriffe „Durchtrittsfrequenz“ und „Phasenreserve“ erforderlich.

8.4.1 Durchtrittsfrequenz

Die Durchtrittsfrequenz ω_c ist diejenige Kreisfrequenz², bei der die Ortskurve $L(j\omega)$ des offenen Kreises den Einheitskreis schneidet, bzw. bei der die Betragskennlinie $|L(j\omega)|_{dB}$ die 0 dB Linie schneidet, d.h.

$$|L(j\omega_c)| = 1 \quad \text{bzw.} \quad |L(j\omega_c)|_{dB} = 0. \quad (8.16)$$

Die Durchtrittsfrequenz ist in Bild 8.5 im oberen Diagramm beim Schnittpunkt der Betragskennlinie mit der 0 dB Linie eingezeichnet, in Bild 8.6 schneidet die Ortskurve $L(j\omega)$ den Einheitskreis (EHK) für den Frequenzparameter $\omega = \omega_c$.

²Die Bezeichnung ω_c ist auf die englische Übersetzung „crossover-frequency“ zurück zu führen.

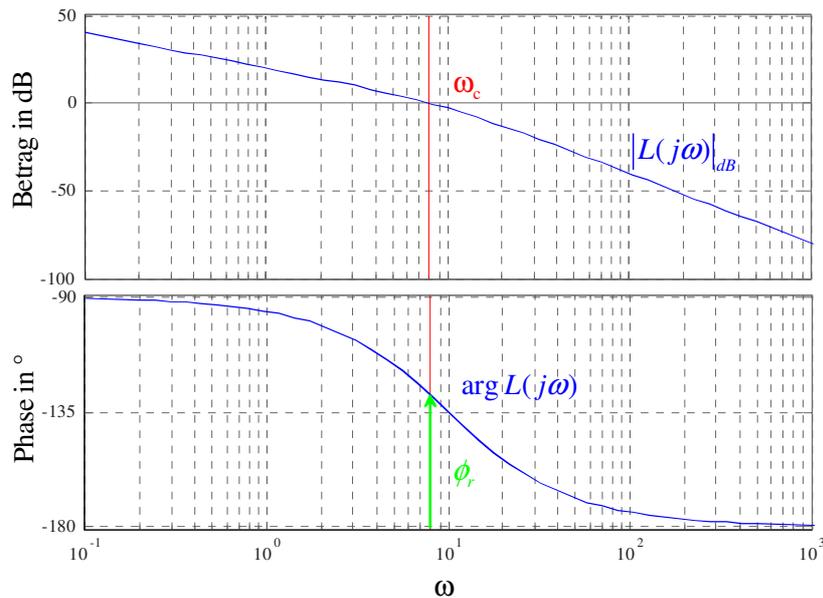


Bild 8.5: Illustration von Durchtrittsfrequenz und Phasenreserve im Bode-Diagramm

8.4.2 Phasenreserve und Amplitudenrand

Die Phasenreserve ϕ_r ist der „Abstand“ der Phasenkennlinie von $L(j\omega)$ an der Stelle $\omega = \omega_c$ zu (-180°) , d.h.

$$\phi_r = \arg L(j\omega_c) + 180^\circ. \quad (8.17)$$

In den Bildern 8.5 und 8.6 ist die Phasenreserve in den Frequenzkennlinien bzw. in der Ortskurve des offenen Kreises dargestellt.

Bezeichnet man mit ω^- diejenige Frequenz, bei der $L(j\omega^-) = -180^\circ$ gilt, so gilt für den so genannten Amplitudenrand

$$A_r = \frac{1}{|L(j\omega^-)|}. \quad (8.18)$$

Aus der in Bild 8.6 dargestellten Ortskurve ist der Amplitudenrand leicht ablesbar, bei den in Bild 8.5 dargestellten Frequenzkennlinien gilt $A_r \rightarrow \infty$, da die Phasenkennlinie den Wert -180° für endliche Werte von ω nicht erreicht.

8.4.3 Formulierung des vereinfachten Schnittpunktkriteriums

Fasst man die Erkenntnisse und Begriffe der letzten Abschnitte zusammen, so kann das **vereinfachte Schnittpunktkriterium** folgendermaßen formuliert werden:

Ist der offene Kreis $L(s)$ vom einfachen Typ, so ist der geschlossene Regelkreis genau dann BIBO-stabil, wenn die Phasenreserve (8.17) positiv ist.

Aus dem vereinfachten Schnittpunktkriterium folgen anschauliche Interpretationen für ϕ_r und A_r . Die Phasenreserve ϕ_r gibt den „Spielraum“ an, in dem sich die Phase des offenen Kreises

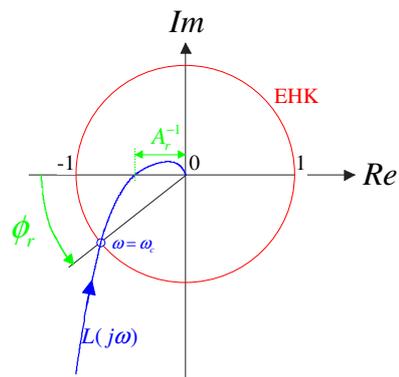


Bild 8.6: Zur Erläuterung von Durchtrittsfrequenz, Phasenreserve und Amplitudenrand

an der Stelle ω_c verändern darf, ohne dass der geschlossene Regelkreis instabil wird. Der Amplitudenrand A_r hingegen kennzeichnet den „Spielraum“, in dem sich der Betrag des offenen Kreises an der Stelle ω_c ändern darf, ohne dass der geschlossene Regelkreis instabil wird.

Kapitel 9

Das Frequenzkennlinien-Verfahren

9.1 Einführung

Der Entwurf von Regelgesetzen mit Hilfe von Frequenzkennlinien gehört zu den klassischen Verfahren der Regelungstechnik. Der Ausgangspunkt der Betrachtungen ist der in Bild 9.1 dargestellte Standardregelkreis, wobei $P(s)$ die *gegebene* Übertragungsfunktion der Regel-

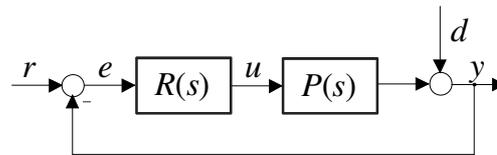


Bild 9.1: Standardregelkreis

strecke ist und $R(s)$ die *gesuchte* Reglerübertragungsfunktion repräsentiert. Diese muss so bestimmt werden, dass Führungs- und Störverhalten, ausgedrückt durch die Übertragungsfunktionen

$$r \rightarrow y: T(s) = \frac{R(s)P(s)}{1 + R(s)P(s)} \quad \text{und} \quad d \rightarrow y: S(s) = \frac{1}{1 + R(s)P(s)} \quad (9.1)$$

so beeinflusst werden, dass vorgegebene Spezifikationen möglichst gut erfüllt werden. Aus (9.1) geht hervor, dass die Auswirkungen von Modifikationen des Reglers $R(s)$ auf die Übertragungsfunktionen $T(s)$ und $S(s)$ im Allgemeinen nur sehr schwer abzuschätzen sind. Dies erschwert die systematische Vorgangsweise bei der Bestimmung des Regelgesetzes. Erfreulicherweise können gewisse Spezifikationen, die der geschlossene Regelkreis erfüllen muss - wenigstens näherungsweise - in entsprechende Bedingungen für den offenen Kreis

$$L(s) = R(s)P(s) \quad (9.2)$$

übersetzt werden. Wie sich zeigen wird, ergeben sich aus den *Vorgaben für den geschlossenen Kreis im Zeitbereich* mittels einfacher Relationen *Vorgaben für den offenen Kreis im Frequenzbereich*. Der Regler $R(s)$ wird also dazu verwendet, den Frequenzgang $L(j\omega)$ des offenen Kreises entsprechend zu „formen“, oft wird dafür der Begriff „loop shaping“ verwendet.

In weiterer Folge wird davon ausgegangen, dass die Übertragungsfunktion des offenen Regelkreises $L(s)$ vom einfachen Typ ist, d.h. das vereinfachte Schnittpunktkriterium kann eingesetzt werden.

9.2 Allgemeine Überlegungen zum Reglerentwurf

Wie man aus (9.1) leicht erkennen kann, gilt

$$S(s) + T(s) = 1, \quad (9.3)$$

d.h. in einem Standardregelkreis ergibt die Summe aus Führungsübertragungsfunktion und Störübertragungsfunktion immer den Wert 1. Daraus folgt unmittelbar, dass in einem (unrealistischen) Regelkreis mit idealem Führungsverhalten, d.h. $T(s) = 1$, ganz automatisch auch $S(s) = 0$ gilt und der Regelkreis somit auch ideales Störverhalten aufweist. In einem realen Regelkreis ist $T(s) = 1$ aufgrund diverser Einschränkungen (siehe weiter unten) überhaupt nicht erzielbar und, wie sich zeigen wird, auch gar nicht erwünscht. Der Grund hierfür sind bisher unberücksichtigte Störeinflüsse, wie z.B. das in Bild 9.2 eingezeichnete Messrauschen $n(t)$. Wie man leicht überprüfen kann, gilt

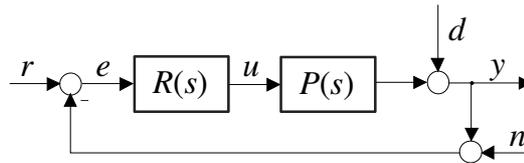


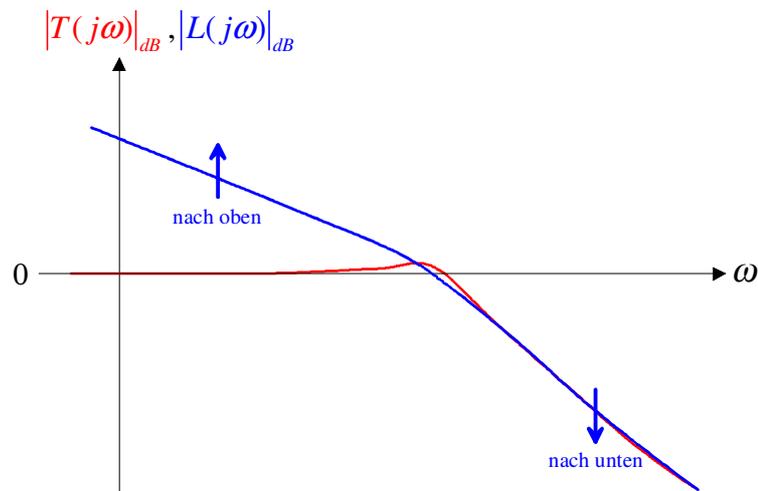
Bild 9.2: Standardregelkreis

$$n \rightarrow y: \quad N(s) = \frac{-R(s)P(s)}{1 + R(s)P(s)} = -T(s),$$

d.h. das Messrauschen wird für $T(s) = 1$ in keinsten Weise unterdrückt. In den meisten praktischen Anwendungen sind r , d und n in verschiedenen Frequenzbereichen wirksam. Verglichen mit r und d ist n üblicherweise ein hochfrequentes Signal. Die daraus resultierenden geforderten Eigenschaften der Führungsübertragungsfunktion $T(s)$ können somit besonders geradlinig mittels des Frequenzganges $T(j\omega)$ angegeben werden, es muss gelten:

$$\begin{aligned} |T(j\omega)| &\approx 1 && \text{im „niederfrequenten Bereich“} \\ |T(j\omega)| &\ll 1 && \text{im „hochfrequenten Bereich“} \end{aligned} \quad (9.4)$$

Diese Erkenntnisse sind in Bild 9.3 graphisch zusammengefasst. Für „niedrige“ Frequenzen soll der Betrag von $T(j\omega)$ möglichst längs der $0dB$ -Linie verlaufen, für „hohe“ Frequenzen möglichst weit unter der $0dB$ -Linie.

Bild 9.3: Typische Verläufe von $|T(j\omega)|$ und $|L(j\omega)|$

Mit Hilfe der bekannten Relation

$$T(s) = \frac{L(s)}{1 + L(s)}, \quad \text{d.h.} \quad |T(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|} \quad (9.5)$$

kann der wünschenswerte Verlauf von $|T(j\omega)|$ leicht auf den offenen Kreis umgerechnet werden, es gilt

$$\begin{aligned} |T(j\omega)| \approx 1 & \Leftrightarrow |L(j\omega)| \gg 1, \\ |T(j\omega)| \ll 1 & \Leftrightarrow |L(j\omega)| \approx |T(j\omega)| \ll 1. \end{aligned} \quad (9.6)$$

D.h. bei niedrigen Frequenzen muss $|L(j\omega)|$ möglichst weit über der 0dB -Linie verlaufen, bei hohen Frequenzen möglichst weit unter der 0dB -Linie, siehe Bild 9.3.

Von besonderer Bedeutung für das Verhalten des Regelkreises ist der Frequenzbereich nahe der Durchtrittsfrequenz ω_c , die bekanntlich durch $|L(j\omega_c)| = 1$ charakterisiert ist. Aus Bild 9.3 kann gefolgert werden, dass die Durchtrittsfrequenz ω_c ungefähr der Bandbreite ω_B des geschlossenen Kreises entspricht. Diese ist definiert durch

$$|T(j\omega_B)|_{dB} = -3 \quad \text{bzw.} \quad |T(j\omega_B)| = \frac{1}{\sqrt{2}}.$$

Wie leicht einzusehen ist, bedeutet eine große Bandbreite, dass der geschlossene Regelkreis auf schnelle Änderungen der Eingangsgrößen ohne nennenswerte Verzögerung reagieren kann. Somit stellen ω_B bzw. ω_c ein Maß für die „Reaktionsfreudigkeit“ des Regelkreises dar. Die Phasenreserve ϕ_r gibt den „Abstand“ der Phasenkennlinie des offenen Kreises an der Stelle $\omega = \omega_c$ zu -180° an. Bei kleiner Phasenreserve ist der geschlossene Regelkreis nahe an der Stabilitätsgrenze, was sich durch eine verstärkte Schwingneigung manifestiert. Weiters können Abweichungen des Streckenmodells von der Realität bis zur Instabilität des Regelkreises

führen. Somit ist die Phasenreserve ein Maß für die Robustheit des Regelkreises und charakterisiert dessen Schwingneigung.

9.2.1 Einschränkungen beim Entwurf

Wesentlich für den erfolgreichen Reglerentwurf ist die Vorgabe von sinnvollen Spezifikationen. Aus diesem Grund ist bei der Wahl der Spezifikationen genau darauf zu achten, ob diese überhaupt erfüllbar sind. Anhand einiger charakteristischer Streckeneigenschaften können mögliche Einschränkungen bei der Vorgabe von Spezifikationen erkannt werden. Dazu zählen Streckenpole und -nullstellen mit positivem Realteil¹ sowie Beschränkungen der Stellgröße.

Instabile Pole

Instabile Streckenpole erfordern eine Mindestbandbreite ω_B , woraus auch unmittelbar eine untere Grenze für die Durchtrittsfrequenz ω_c folgt. Diese Tatsache kann leicht anhand des Balancierens eines Stabes auf einem Finger plausibel gemacht werden. Die Stabilisierung des Stabes ist prinzipiell nur dann möglich, wenn der Finger hinreichend schnell bewegt werden kann, der Regelkreis also eine gewisse Mindestbandbreite besitzt.

Instabile Nullstellen

Im Gegensatz zu instabilen Polen begrenzen instabile Streckennullstellen die erzielbare Bandbreite ω_B bzw. die maximale Durchtrittsfrequenz ω_c . Es ist zu beachten, dass instabile Streckennullstellen immer auch Nullstellen von $L(s)$ und $T(s)$ sind. Anderenfalls kommt es im Regelkreis zu einer verbotenen „instabilen“ Kürzung. Als Beispiel kann wieder das Balancieren eines Stabes auf dem Finger herangezogen werden. Will man beispielsweise den balancierenden Finger nach rechts bewegen ohne dass der Stab hinunterfällt, so muss der Finger zunächst kurz nach links, also in die Gegenrichtung bewegt werden. Dieses „Systemverhalten“ ist typisch für Systeme mit einer instabilen Nullstelle. Es veranschaulicht sehr schön, warum die gewünschte Endposition des balancierenden Fingers nicht beliebig schnell erreicht werden kann.

Stellgrößenbeschränkung

Bei praktischen Anwendungen sind de facto immer Beschränkungen der Stellgröße zu berücksichtigen. Eine gezielte Einhaltung von Stellgrößenbeschränkungen ist allerdings bei vielen gängigen Entwurfsverfahren nur schwer oder gar nicht möglich. Aus diesem Grund werden oft sehr „vorsichtige“ Reglereinstellungen verwendet. Dies führt zu einer schlechten Ausnutzung des Stellbereiches und in weiterer Folge zu einem „trägen“ Verhalten des Regelkreises. Häufige und länger andauernde Verletzungen der Stellgrößenbeschränkung aufgrund zu „aggressiver“ Reglereinstellungen können hingegen zu unerwünschtem dynamischen Verhalten des geschlossenen Kreises führen. Ein Beispiel hierfür ist der bereits ausführlich diskutierte Windup-Effekt.

¹In diesem Zusammenhang spricht man auch von „instabilen“, Polen bzw. Nullstellen.

9.3 Ein typisches Entwurfsverfahren

Es wird davon ausgegangen, dass der geschlossene Regelkreis ein System mit dominantem Polpaar ist. Das bedeutet, dass das dynamische Verhalten des Regelkreises durch ein konjugiert komplexes Polpaar dominiert wird. Wie sich zeigen wird, ergeben sich aus dieser Annahme einfache Richtlinien für den Reglerentwurf.

9.3.1 Systeme mit dominantem Polpaar

Aufgrund obiger Annahme kann davon ausgegangen werden, dass das Verhalten des Regelkreises näherungsweise dem eines System mit der Übertragungsfunktion

$$T(s) = \frac{\omega_n^2}{s^2 + 2d\omega_n s + \omega_n^2} \quad \text{mit} \quad \omega_n > 0 \quad \text{und} \quad 0 < d < 1, \quad (9.7)$$

entspricht, wobei ω_n die Kennkreisfrequenz und d der Dämpfungsgrad ist. Die angegebenen Wertebereiche für ω_n und d gewährleisten, dass die BIBO-stabile Übertragungsfunktion $T(s)$ ein konjugiert komplexes Polpaar

$$s_{1,2} = -d\omega_n \pm j\omega_n\sqrt{1-d^2}$$

besitzt. Für die zugehörige Übertragungsfunktion des offenen Kreises gilt

$$L(s) = \frac{\omega_n^2}{s(s + 2d\omega_n)}. \quad (9.8)$$

Wie bereits erwähnt wurde, haben Kenngrößen des offenen Kreises, nämlich die Durchtrittsfrequenz ω_c , bei der bekanntlich $|L(j\omega_c)| = 1$ gilt, und die Phasenreserve $\phi_r = \arg L(j\omega_c) + 180^\circ$ einen wesentlichen Einfluss auf das dynamische Verhalten des geschlossenen Regelkreises. Diese Tatsache wird beim vorliegenden Entwurfsverfahren ausgenutzt, indem das Wunschverhalten des geschlossenen Regelkreises in entsprechende Werte für ω_c und ϕ_r umgerechnet wird.

9.3.2 Spezifikation des Wunschverhaltens

Sehr oft wird das Wunschverhalten eines Regelkreises im Zeitbereich spezifiziert. Hierfür gibt es eine Vielzahl verschiedener Möglichkeiten. Im vorliegenden Fall wird das dynamische Verhalten des Regelkreises durch den Verlauf seiner Sprungantwort mit Hilfe der Anstiegszeit und der Überschwingweite (siehe Abschnitt 6.5) charakterisiert, das asymptotische Verhalten wird über die bleibende Regelabweichung spezifiziert.

Bleibende Regelabweichung

Das asymptotische Verhalten des Regelkreises wird durch die bleibende Regelabweichung e_∞ charakterisiert, wobei für die Führungsgröße spezielle „Testfunktionen“ der Form

$$r(t) = \frac{t^{\nu-1}}{(\nu-1)!} \sigma(t) \quad \text{mit} \quad \bar{r}(s) = \frac{1}{s^\nu} \quad \text{mit} \quad \nu = 1, 2, 3, \dots \quad (9.9)$$

gewählt werden. Geht man von einer normierten Darstellung der Übertragungsfunktion des offenen Kreises aus, also

$$L(s) = \frac{Vp(s)}{s^\lambda q(s)} \quad \text{mit} \quad p(0) = q(0) = 1,$$

so ergibt sich für die Laplace-Transformierte des Regelfehlers

$$\bar{e}(s) = \mathcal{L}\{e(t)\} = \frac{1}{1+L(s)} \bar{r}(s) = \frac{s^\lambda q(s)}{s^\lambda q(s) + Vp(s)} \bar{r}(s),$$

wobei in einem sinnvoll entworfenen Regelkreis das Polynom $[s^\lambda q(s) + Vp(s)]$ natürlich ein Hurwitzpolynom ist. Setzt man (9.9) in den obigen Ausdruck ein, so erhält man

$$\bar{e}(s) = \frac{s^\lambda q(s)}{s^\lambda q(s) + Vp(s)} \frac{1}{s^\nu} = \frac{s^{\lambda-\nu} q(s)}{s^\lambda q(s) + Vp(s)}.$$

Vom Endwertsatz der Laplace-Transformation weiß man, dass der Grenzwert

$$e_\infty = \lim_{t \rightarrow \infty} e(t)$$

genau dann existiert, wenn das Nennerpolynom von

$$s \bar{e}(s) = \frac{s^{\lambda-\nu+1} q(s)}{s^\lambda q(s) + Vp(s)}$$

ein Hurwitzpolynom ist. Dies ist offensichtlich dann der Fall, wenn die Bedingung

$$\lambda - \nu + 1 \geq 0 \quad \Rightarrow \quad \lambda \geq \nu - 1$$

erfüllt ist. Es darf dann der Endwertsatz

$$e_\infty = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s \bar{e}(s) = \lim_{s \rightarrow 0} \frac{s^{\lambda-\nu+1} q(s)}{s^\lambda q(s) + Vp(s)} = \lim_{s \rightarrow 0} \frac{s^{\lambda-\nu+1}}{s^\lambda + V}$$

angewandt werden. Aus obigem Ausdruck für die bleibende Regelabweichung kann die unten angegebene Tabelle unmittelbar abgeleitet werden.

e_∞	$\lambda = 0$	$\lambda = 1$	$\lambda = 2$
$\nu = 1 \Rightarrow r(t) = \sigma(t)$	$\frac{1}{1+V}$	0	0
$\nu = 2 \Rightarrow r(t) = t\sigma(t)$	-	$\frac{1}{V}$	0
$\nu = 3 \Rightarrow r(t) = \frac{t^2}{2!}\sigma(t)$	-	-	$\frac{1}{V}$

Aus der Tabelle ist beispielsweise zu erkennen, dass bei sprungförmiger Führungsgröße r die bleibende Regelabweichung e_∞ verschwindet, wenn der offene Kreis mindestens einfach integrierend ist.

9.3.3 Faustformeln für den Reglerentwurf

Es kann gezeigt werden, dass bei Systemen mit dominantem Polpaar die beiden Relationen

$$\phi_r \text{ [in Grad]} + \ddot{u} \text{ [in \%]} \approx 70 \quad \text{und} \quad \omega_c t_r \approx 1.5 \quad (9.10)$$

näherungsweise gelten. Diese beiden „Faustformeln“ spiegeln die bereits bekannten Tatsachen wider, dass eine kleine Phasenreserve eine große Schwingneigung des Regelkreises zur Folge hat und dass eine kleine Anstiegszeit eine große Bandbreite erfordert.

9.3.4 Durchführung des Reglerentwurfes

Der Reglerentwurf gliedert sich typischerweise in drei Schritte, nämlich

I Umrechnung der gegebenen Regelkreisspezifikationen auf den offenen Kreis

Das gewünschte Überschwingen und die gewünschte Anstiegszeit werden mittels (9.10) in ω_c und ϕ_r umgerechnet. Weiters werden aus der vorgegebenen bleibenden Regelabweichung λ und gegebenenfalls V ermittelt.

II Entsprechende Modifikation der Frequenzkennlinien des offenen Kreises

Durch geschicktes Einfügen von Korrekturgliedern wird der Frequenzgang des offenen Kreises sukzessive so verändert, dass er gewünschte Eigenschaften besitzt.

III Simulation des Regelkreises

Da der Reglerentwurf auf einigen vereinfachenden Annahmen beruht, ist eine Simulation des geschlossenen Kreises unumgänglich. Hierbei muss überprüft werden, ob die vorgegebenen Spezifikationen zufriedenstellend erfüllt werden.

9.4 Korrekturglieder

Proportionalglieder

Mit Hilfe eines Proportionalgliedes, d.h.

$$R(s) = K$$

kann die Betragskennlinie des offenen Kreises um $|K|_{dB}$ angehoben ($|K| > 1$) bzw. abgesenkt ($|K| < 1$) werden, wobei die Phasenkennlinie nicht beeinflusst bzw. bei negativem K um 180° abgesenkt wird.

Integrierer

Werden m Integrierer in den offenen Kreis eingefügt, d.h.

$$R(s) = \frac{1}{s^m},$$

so wird die Phasenkennlinie um $m \cdot 90^\circ$ Grad abgesenkt wobei auch die Betragskennlinie entsprechend verändert wird. Üblicherweise werden Integrierer in den offenen Kreis eingefügt, um das asymptotische Verhalten des Regelkreises den vorgegebenen Spezifikationen anzupassen.

Lead/Lag-Glied

Das so genannte Lead / Lag-Glied besitzt die Übertragungsfunktion

$$R(s) = \frac{1 + \frac{s}{\omega_Z}}{1 + \frac{s}{\omega_N}} \quad \text{mit} \quad \omega_Z, \omega_N > 0,$$

ist also der Quotient zweier Linearfaktoren mit den (positiven) Knickfrequenzen ω_Z und ω_N . Für $\omega_Z < \omega_N$ spricht man von einem Lead-Glied (phasenanhebend), für $\omega_N < \omega_Z$ von einem Lag-Glied (phasenabsenkend). Definiert man die Größe m gemäß

$$m := \frac{\omega_N}{\omega_Z},$$

so gilt offensichtlich

$$m > 1 : \quad \text{Lead-Glied} \quad \text{bzw.} \quad m < 1 : \quad \text{Lag-Glied.}$$

In Bild 9.4 sind die Frequenzkennlinien eines Lead-Gliedes graphisch dargestellt. Man erkennt, dass die Betragskennlinie angehoben wird, für die maximale Anhebung ΔA gilt

$$\Delta A_{dB} = m_{dB} = 20 \log m \quad (\Delta A_{dB} > 0 \text{ wegen } m > 1).$$

Das bedeutet, man kann an einer vorgegebenen Frequenz $\omega \gg \omega_N$ die Betragskennlinie des offenen Kreises um ΔA anheben, ohne die Phasenkennlinie des offenen Kreises nennenswert zu beeinflussen. Typischerweise wird aber ein Lead-Glied dazu eingesetzt, die Phasenkennlinie des offenen Kreises anzuheben. Wie man Bild 9.4 entnehmen kann, tritt die größte Phasenanhebung $\Delta\varphi$ bei der so genannten Mittenfrequenz

$$\omega_m = \sqrt{\omega_Z \omega_N}$$

auf, sie kann gemäß

$$\Delta\varphi = \arcsin \frac{m-1}{m+1} \quad (9.11)$$

berechnet werden.

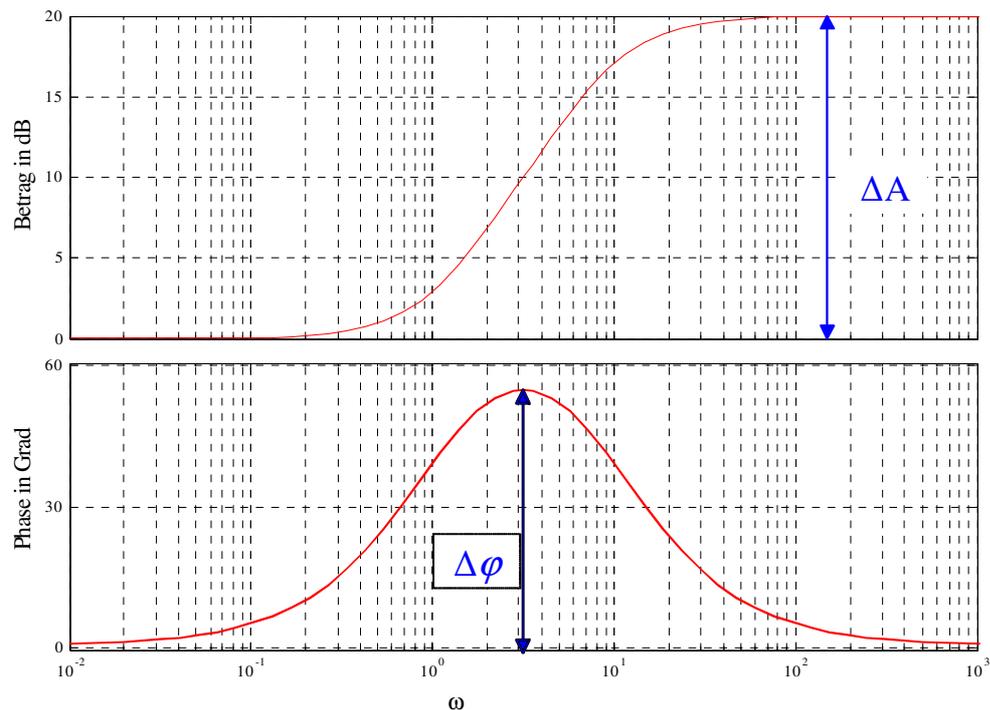


Bild 9.4: Frequenzkennlinien, Lead-Glied

In Bild 9.5 sind die Frequenzkennlinien eines Lag-Gliedes graphisch dargestellt. Üblicherweise wird ein Lag-Glied dazu eingesetzt, die Betragskennlinie an einer vorgegebenen Frequenz $\omega \gg \omega_Z$ um

$$\Delta A_{dB} = m_{dB} = 20 \log m \quad (\Delta A_{dB} < 0 \text{ wegen } m < 1).$$

abzusenken, ohne die Phasenkennlinie nennenswert zu beeinflussen.

B Gegeben sei die Übertragungsfunktion einer Regelstrecke

$$P(s) = \frac{10}{(s+1)^2}.$$

Für die Sprungantwort des Regelkreises sind die Spezifikationen

$$t_r = 3s, \quad M_p = 1.1 \quad \text{und} \quad e_\infty = 0$$

vorgegeben. Die Anwendung der Faustformeln (9.10) liefert die Vorgaben

$$\omega_c = 0.5 \text{ rads}^{-1} \quad \text{und} \quad \phi_r = 60^\circ$$

für den Frequenzgang des offenen Kreises. Zur Erzielung von $e_\infty = 0$ für eine sprungförmige Referenzgröße ist mindestens ein Integrator im offenen Kreis erforderlich. Aus diesem Grund wird als Ausgangsbasis für den Regler die Übertragungsfunktion

$$R_1(s) = \frac{1}{s}$$

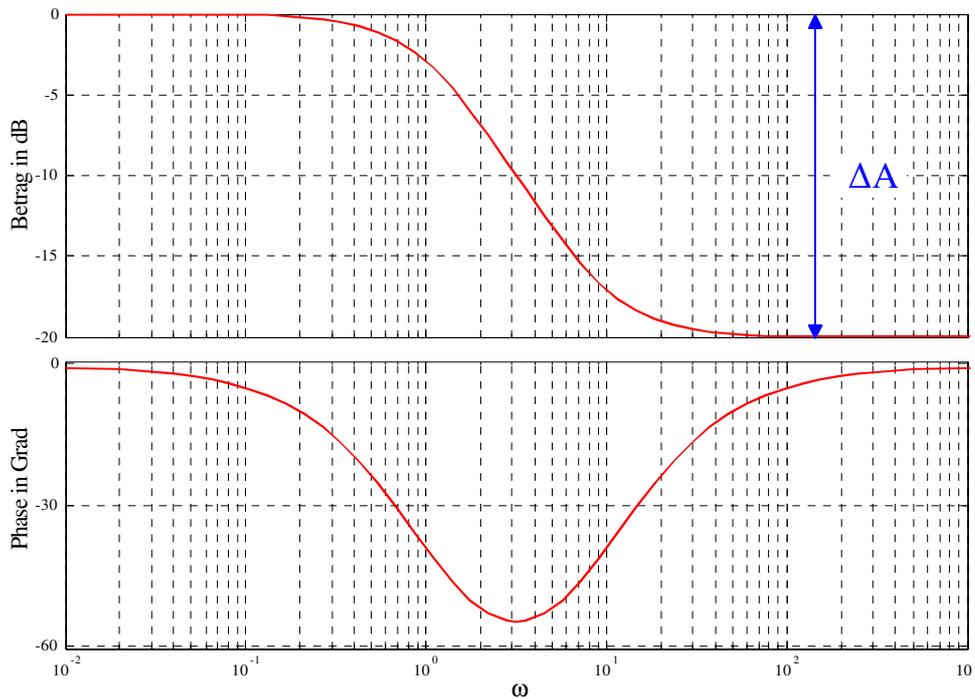


Bild 9.5: Frequenzkennlinien, Lag-Glied

gewählt, für den offenen Kreis gilt dann

$$L_1(s) = R_1(s)P(s) = \frac{10}{s(1+s)^2}.$$

Wie in Bild 9.6 zu erkennen ist, gilt

$$\arg L_1(j\omega_c) = -143^\circ,$$

d.h. zur Erzielung einer Phasenreserve von 60° muss die Phasenkennlinie des offenen Kreises an der Stelle ω_c um $\Delta\varphi = 23^\circ = 0.4 \text{ rad}$ angehoben werden. Hierzu wird ein Lead-Glied eingesetzt. Mit Hilfe von (9.11) findet man

$$m = \frac{1 + \sin 0.4}{1 - \sin 0.4} = 2.28$$

und man erhält mit $\omega_m = \omega_c$ die Parameter

$$\omega_Z = \frac{0.5}{\sqrt{2.28}} = 0.33 \text{ rads}^{-1}, \quad \omega_N = 0.5\sqrt{2.28} = 0.76 \text{ rads}^{-1}$$

Das gesuchte Lead-Glied lautet somit

$$R_2(s) = \frac{1 + \frac{s}{0.33}}{1 + \frac{s}{0.76}}.$$

Damit lautet die Übertragungsfunktion des offenen Kreises nun

$$L_2(s) = R_1(s)R_2(s)P(s) = 10 \frac{\left(1 + \frac{s}{0.33}\right)}{s \left(1 + \frac{s}{0.76}\right) (1 + s)^2}.$$

In Bild 9.6 erkennt man, dass die gewünschte Phase an der Stelle ω_c nun gewährleistet ist, für die Betragskennlinie gilt

$$|L_2(j\omega_c)|_{dB} = 27.$$

Die entsprechende Absenkung der Betragskennlinie kann mit dem Proportionalelement

$$R_3(s) = 10^{-\frac{27}{20}} = 0.041$$

bewerkstelligt werden. Damit lautet der endgültige Regler

$$R(s) = R_1(s)R_2(s)R_3(s) = 0.041 \frac{\left(1 + \frac{s}{0.33}\right)}{s \left(1 + \frac{s}{0.76}\right)}.$$

In Bild 9.7 ist die Sprungantwort des Regelkreises dargestellt, Überschwingweite und Anstiegszeit entsprechen in guter Näherung den vorgegebenen gewünschten Werten.

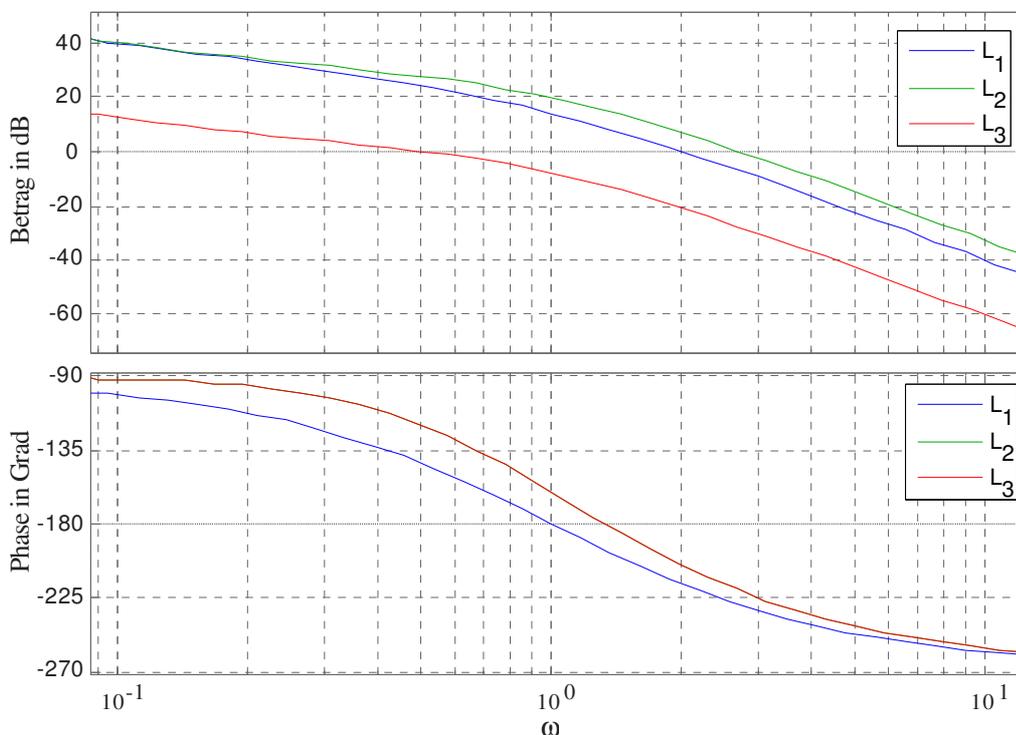


Bild 9.6: Frequenzkennlinien des offenen Kreises

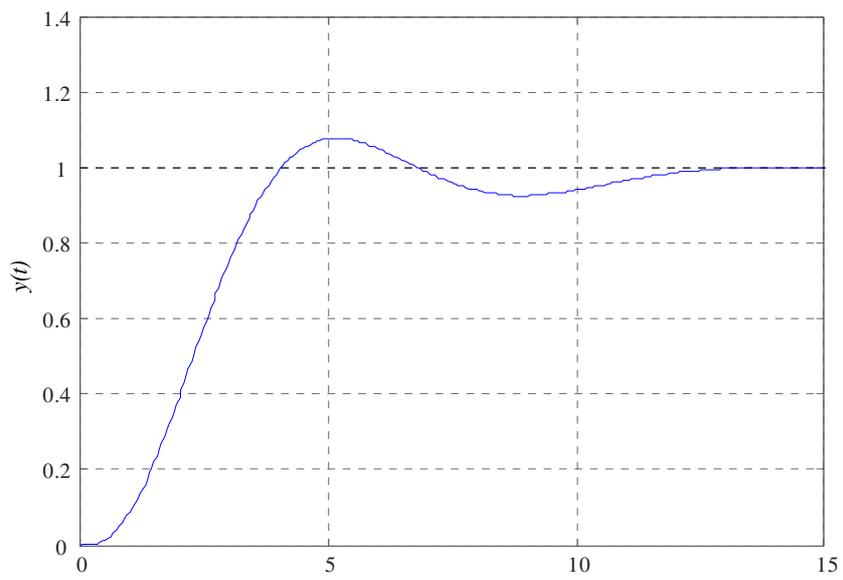


Bild 9.7: Sprungantwort des Regelkreises



Kapitel 10

Englische Fachbegriffe

Anfangswertsatz	initial value theorem
Anfangszustand	initial state
asymptotisch stabil	asymptotically stable
asymptotische Stabilität	asymptotic stability
autonom	autonomous
Begleitform	companion form
beobachtbar	observable
Beobachtbarkeit	observability
Bebachtbarkeitsnormalform	observable-form realization, observable canonical form
BIBO Stabilität	BIBO stability
Determinante	determinant
Diagonalform	diagonal form
Differentialgleichung	differential equation
Dirac Impuls	Dirac delta function
Durchgriffsterm	direct transmission term
Eigenwert	eigenvalue
Eigenvektor	eigenvector
Eingrößensystem	single input - single output system
Einheitskreis	unit circle
Einheitsmatrix	identity matrix
Einheitssprung	unit step
Endwertsatz	final value theorem
Erreichbarkeit	reachability
Faltung	convolution
Faltungsintegral	convolution integral
Faltungssumme	convolution sum
Frequenzgang	frequency response
Impuls	impulse
instabil	unstable
Jordan-Form	Jordan canonical form

Jordanblock	Jordan block
Koeffizient	coefficient
konjugiert komplex	conjugate complex
Knotenpunkt	node
stabil	nodal sink, stable node
instabil	nodal source, unstable node
kürzen	to cancel
Kürzung	cancellation
Laplace Transformation	Laplace transform
linear, zeitinvariant (LZI)	linear, time-invariant (LTI)
linear unabhängig	linearly independent
Linearisierung	linearization
Lösung, freie	zero-input response
erzwungene	zero-state response
Matrix	matrix, <i>pl.</i> matrices
adjungierte	adjoint matrix
diagonale	diagonal matrix
Hurwitz-	Hurwitz matrix
in Begleitform	companion matrix
inverse	inverse matrix
quadratische	square matrix
reguläre	nonsingular matrix
schiefsymmetrische	skew symmetric matrix
singuläre	singular matrix
symmetrische	symmetric matrix
transponierte	transpose matrix, transposed matrix
Minimalrealisierung	minimal realization
Nenner	denominator
Normalform	canonical form
Nullstelle	zero
Parallelschaltung	parallel connection
Phasenportrait	phase portrait
PN plan	pole-zero map
Pol, Polstelle	pole
Polynom	polynomial
charakteristisches	characteristic polynomial
Hurwitz-	Hurwitz polynomial
monisches	monic polynomial
Nenner-	denominator polynomial
Zähler-	numerator polynomial
Rang	rank
realisierbar	proper
Reihenschaltung	series connection
Routh Schema	Routh test

Rückkopplung	feedback
Ruhelage	equilibrium point
Sattelpunkt	saddle point
Sprungantwort	step response
sprungfähig	biproper
Sprungfunktion	step function
Stabilität	stability
asymptotische	asymptotic stability
BIBO	BIBO stability
steuerbar	controllable
Steuerbarkeit	controllability
Steuerbarkeitsmatrix	controllability matrix
Steuerbarkeitsnormalform	controllable-form realization, controllable canonical form
Systemordnung	system order
teilerfremd	coprime
Trajektorie	trajectory
Transitionsmatrix	state transition matrix
Übertragungsfunktion	transfer function
Vektor	vector
Spalten-	column vector
Zeilen-	row vector
Vielfachheit	multiplicity
Wirbelpunkt	spiral point
stabil	spiral sink
instabil	spiral source
Zähler	numerator
zeitdiskret	discrete-time
zeitinvariant	time-invariant
zeitvariant	time-variant
Zustandsebene	state plane
-modell	state model
-raum	state space
-variable	state variable
-transformation	state transformation
-vektor	state vector

Anhang A

Laplace-Transformation

Die (einseitige) **Laplace-Transformation** ordnet einer Zeitfunktion $f(t)$ über die Berechnungsvorschrift

$$\bar{f}(s) = \int_0^{+\infty} f(t) e^{-st} dt \quad (\text{A.1})$$

eine Funktion $\bar{f}(s)$ der komplexen Variable s zu. Dabei wird an dieser Stelle vereinfachend angenommen, dass $f(t)$ für negative Werte des Parameters t verschwindet, d.h.

$$f(t) = 0 \quad \text{für } t < 0.$$

Weiters wird vorausgesetzt, dass $f(t)$ auf jedem endlichen Zeitintervall stückweise stetig ist und - für geeignet gewählte reelle Konstanten M und γ - der Ungleichung

$$|f(t)| \leq M e^{\gamma t} \quad (\text{A.2})$$

genügt. Dann ist das Integral (A.1) für alle s mit $\text{Re}\{s\} > \gamma$ absolut konvergent¹, man spricht auch vom Existenzbereich von $\bar{f}(s)$. Jeder transformierbaren Funktion $f(t)$ im **Zeitbereich** wird durch (A.1) eine Funktion $\bar{f}(s)$ im **Bildbereich** zugeordnet. Diese Korrespondenz zwischen der **Originalfunktion** $f(t)$ und der zugehörigen **Bildfunktion** $\bar{f}(s)$ wird durch

$$\bar{f}(s) = \mathcal{L}\{f(t)\} \quad \text{und} \quad f(t) = \mathcal{L}^{-1}\{\bar{f}(s)\} \quad (\text{A.3})$$

bzw. durch das „Hantelsymbol“

$$f(t) \circ \bullet \bar{f}(s) \quad (\text{A.4})$$

symbolisiert.

A.1 Linearität

Die Laplace-Transformation ist eine lineare Transformation, d.h. es gilt

$$\mathcal{L}\{\alpha_1 f_1(t) + \alpha_2 f_2(t)\} = \alpha_1 \mathcal{L}\{f_1(t)\} + \alpha_2 \mathcal{L}\{f_2(t)\} \quad (\text{A.5})$$

für beliebige komplexe Konstanten α_1 und α_2 .

¹das Laplace-Integral ist also in einer rechten Halbebene absolut konvergent. In den meisten Fällen kann die Funktion $\bar{f}(s)$ in die gesamte komplexe Ebene analytisch fortgesetzt werden.

A.2 Dämpfungsregel

Sei $f(t)$ eine Zeitfunktion mit zugehöriger Laplace-Transformierter $\bar{f}(s)$. Es gilt dann

$$\mathcal{L}\{f(t)e^{\alpha t}\} = \bar{f}(s - \alpha), \quad (\text{A.6})$$

wobei α eine beliebige komplexe Konstante ist.

A.3 Verschiebungssatz

Sei $f(t)$ eine Zeitfunktion mit zugehöriger Laplace-Transformierter $\bar{f}(s)$. Für die Laplace-Transformierte der auf der Zeitachse um $\tau > 0$ (nach rechts) verschobenen Funktion $f(t - \tau)$ gilt

$$\mathcal{L}\{f(t - \tau)\} = e^{-s\tau} \mathcal{L}\{f(t)\}. \quad (\text{A.7})$$

Man beachte hierzu auch die Darstellung in Bild A.1, die nochmals verdeutlichen soll, dass die verschobene Funktion $f(t - \tau)$ für $t < \tau$ identisch Null ist.

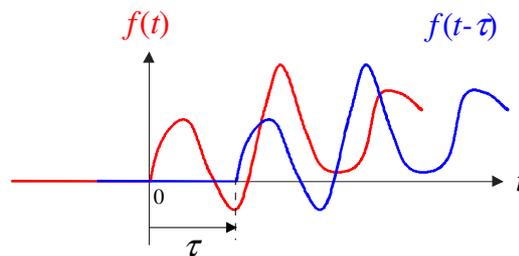


Bild A.1: zur Illustration des Verschiebungssatzes

A.4 Faltungsregel

Gegeben seien zwei Zeitfunktionen mit den zugehörigen Laplace-Transformierten, also $\bar{f}(s) = \mathcal{L}\{f(t)\}$ und $\bar{g}(s) = \mathcal{L}\{g(t)\}$. Es gilt dann

$$\mathcal{L}^{-1}\{\bar{f}(s)\bar{g}(s)\} = \int_0^t f(t - \tau)g(\tau) d\tau = \int_0^t f(\tau)g(t - \tau) d\tau. \quad (\text{A.8})$$

Man spricht in diesem Zusammenhang von der Faltung der Funktionen $f(t)$ und $g(t)$ und verwendet häufig die Kurzschreibweise

$$f(t) * g(t) := \int_0^t f(t - \tau)g(\tau) d\tau.$$

A.5 Differentiation im Zeitbereich

Sei $f(t)$ eine Zeitfunktion mit zugehöriger Laplace-Transformierter $\bar{f}(s)$. Für die Laplace-Transformierte der zeitlichen Ableitung von $f(t)$ gilt dann

$$\mathcal{L} \left\{ \frac{df(t)}{dt} \right\} = s\bar{f}(s) - f(0). \quad (\text{A.9})$$

Eine wiederholte Anwendung des Differentiationssatzes liefert für die n -te Ableitung von $f(t)$ die Korrespondenz

$$\mathcal{L} \left\{ \frac{d^n f(t)}{dt^n} \right\} = s^n \bar{f}(s) - \sum_{i=0}^{n-1} s^{n-1-i} \left. \frac{d^i f}{dt^i} \right|_{t=0}. \quad (\text{A.10})$$

Man beachte, dass sich im Falle verschwindender Anfangswerte obige Relation zu

$$\mathcal{L} \left\{ \frac{d^n f(t)}{dt^n} \right\} = s^n \bar{f}(s) \quad (\text{A.11})$$

vereinfacht. Eine n -fache Differentiation im Zeitbereich entspricht einer Multiplikation im Bildbereich mit s^n .

A.6 Differentiation im Bildbereich

Sei $f(t)$ eine Zeitfunktion mit zugehöriger Laplace-Transformierter $\bar{f}(s)$. Es gilt dann die Relation

$$-\mathcal{L} \{ t f(t) \} = \frac{d\bar{f}(s)}{ds}. \quad (\text{A.12})$$

Eine n -fache Differentiation nach s führt auf das Ergebnis

$$\mathcal{L} \{ t^n f(t) \} = (-1)^n \frac{d^n \bar{f}(s)}{ds^n}. \quad (\text{A.13})$$

A.7 Integration im Zeitbereich

Sei $f(t)$ eine Zeitfunktion mit zugehöriger Laplace-Transformierter $\bar{f}(s)$. Es gilt dann

$$\mathcal{L} \left\{ \int_0^t f(\tau) d\tau \right\} = \frac{1}{s} \bar{f}(s). \quad (\text{A.14})$$

A.8 Anfangswertsatz

Gegeben sei die Laplace-Transformierte $\bar{f}(s)$ einer Zeitfunktion $f(t)$. Der Wert der Zeitfunktion $f(t)$ zum Zeitpunkt $t = 0$ kann mit Hilfe des Anfangswertsatzes

$$f(0) = \lim_{s \rightarrow \infty} s \bar{f}(s) \quad (\text{A.15})$$

berechnet werden.

A.9 Endwertsatz

Gegeben sei die Laplace-Transformierte $\bar{f}(s)$ einer Zeitfunktion $f(t)$. Der Wert der Zeitfunktion $f(t)$ für $t \rightarrow \infty$ kann mit Hilfe des Endwertsatzes

$$f_\infty := \lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} s \bar{f}(s) \quad (\text{A.16})$$

ermittelt werden. Man beachte, dass die Anwendung von (A.16) nur dann zulässig ist, wenn der Grenzwert von $f(t)$ für $t \rightarrow \infty$ existiert. Dies kann ebenfalls im Bildbereich überprüft werden. Ist die Funktion $\bar{g}(s) = s \bar{f}(s)$ eine gebrochen rationale Funktion, d.h.

$$\bar{g}(s) = \frac{\kappa(s)}{\lambda(s)},$$

wobei $\kappa(s)$ und $\lambda(s)$ teilerfremde (gekürzte) Polynome sind, so existiert f_∞ genau dann, wenn das Polynom $\lambda(s)$ ausschließlich Nullstellen mit negativem Realteil besitzt. Ein solches Polynom wird auch **Hurwitzpolynom**² genannt.

²benannt nach dem deutschen Mathematiker Adolf Hurwitz (1859-1919)

A.10 Korrespondenztabelle für reelle Funktionen

$f(t)$	$\circ \text{---} \bullet$	$\bar{f}(s)$
$\delta(t)$		1
$\sigma(t)$		$\frac{1}{s}$
t		$\frac{1}{s^2}$
$\frac{t^n}{n!}, n \geq 0$		$\frac{1}{s^{n+1}}$
$e^{\alpha t}$		$\frac{1}{s - \alpha}$
$te^{\alpha t}$		$\frac{1}{(s - \alpha)^2}$
$\frac{t^n}{n!}e^{\alpha t}, n \geq 0$		$\frac{1}{(s - \alpha)^{n+1}}$
$\cos \omega t$		$\frac{s}{s^2 + \omega^2}$
$\sin \omega t$		$\frac{\omega}{s^2 + \omega^2}$
$e^{\alpha t} \cos \omega t$		$\frac{(s - \alpha)}{(s - \alpha)^2 + \omega^2}$
$e^{\alpha t} \sin \omega t$		$\frac{\omega}{(s - \alpha)^2 + \omega^2}$
$t \cos \omega t$		$\frac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$
$t \sin \omega t$		$\frac{2\omega s}{(s^2 + \omega^2)^2}$

Anhang B

z-Transformation

Die (einseitige) **z-Transformation** ordnet einer Zahlenfolge (f_k) über die Berechnungsvorschrift

$$\tilde{f}(z) = f_0 + f_1 z^{-1} + f_2 z^{-2} + \dots = \sum_{k=0}^{\infty} f_k z^{-k} \quad (\text{B.1})$$

eine Funktion $\tilde{f}(z)$ der komplexen Variable z zu. Dabei wird an dieser Stelle vereinfachend angenommen, dass die Folgeelemente von (f_k) für negative Werte von k verschwinden, d.h.

$$f_k = 0 \quad \text{für} \quad k < 0.$$

Für Zahlenfolgen (f_k) , deren Elemente - für geeignet gewählte reelle Konstanten M und γ - der Ungleichung

$$|f_k| \leq M \gamma^k$$

genügen, konvergiert die Summe (B.1) im Gebiet $|z| > \gamma$. Jeder transformierbaren Folge $(f_k) = (f_0, f_1, f_2, \dots)$ wird somit durch (B.1) eine Funktion $\tilde{f}(z)$ im **Bildbereich** zugeordnet. Diese Korrespondenz wird durch

$$\tilde{f}(z) = \mathfrak{Z}\{(f_k)\} \quad \text{und} \quad (f_k) = \mathfrak{Z}^{-1}\{\tilde{f}(z)\} \quad (\text{B.2})$$

bzw. durch das „Hantelsymbol“

$$(f_k) \circ \bullet \tilde{f}(z) \quad (\text{B.3})$$

symbolisiert¹.

B.1 Linearität

Die z-Transformation ist eine lineare Transformation, d.h. es gilt

$$\mathfrak{Z}\{\alpha_1 (f_{1,k}) + \alpha_2 (f_{2,k})\} = \alpha_1 \mathfrak{Z}\{(f_{1,k})\} + \alpha_2 \mathfrak{Z}\{(f_{2,k})\} \quad (\text{B.4})$$

für beliebige komplexe Konstanten α_1 und α_2 .

¹Häufig wird auch die Schreibweise $\tilde{f}(z) = \mathcal{Z}\{(f_k)\}$ und $(f_k) = \mathcal{Z}^{-1}\{\tilde{f}(z)\}$ verwendet.

B.2 Dämpfungsregel

Sei (f_k) eine Zahlenfolge mit zugehöriger z-Transformierter $\tilde{f}(z)$. Es gilt dann

$$\mathfrak{Z}\{(\alpha^k f_k)\} = \tilde{f}\left(\frac{z}{\alpha}\right), \quad (\text{B.5})$$

wobei α eine beliebige komplexe Konstante ist.

B.3 Verschiebung nach rechts

Sei (f_k) eine Zahlenfolge mit zugehöriger Transformierter $\tilde{f}(z)$. Die z-Transformierte der um $n > 0$ Schritte (nach rechts) verschobenen Folge (f_{k-n}) lautet

$$\mathfrak{Z}\{(f_{k-n})\} = z^{-n} \tilde{f}(z). \quad (\text{B.6})$$

Man beachte, dass die Elemente der verschobenen Folge für $k < n$ identisch Null sind.

B.4 Verschiebung nach links

Sei (f_k) eine Zahlenfolge mit zugehöriger Transformierter $\tilde{f}(z)$. Die z-Transformierte der um $n > 0$ Schritte (nach links) verschobenen Folge (f_{k+n}) lautet

$$\mathfrak{Z}\{(f_{k+n})\} = z^n \tilde{f}(z) - z^n \sum_{i=0}^{n-1} f_i z^{-i}. \quad (\text{B.7})$$

Sind alle Elemente der Folge (f_k) für $k < n$ identisch Null, so vereinfacht sich (B.7) zu

$$\mathfrak{Z}\{(f_{k+n})\} = z^n \tilde{f}(z). \quad (\text{B.8})$$

B.5 Faltungsregel

Gegeben seien zwei Zahlenfolgen mit den zugehörigen z-Transformierten $\tilde{f}(z) = \mathfrak{Z}\{(f_k)\}$ und $\tilde{g}(z) = \mathfrak{Z}\{(g_k)\}$. Es gilt dann

$$\mathfrak{Z}^{-1}\{\tilde{f}(z)\tilde{g}(z)\} = \left(\sum_{i=0}^k f_{k-i}g_i\right) = \left(\sum_{i=0}^k g_{k-i}f_i\right). \quad (\text{B.9})$$

Man spricht in diesem Zusammenhang von der Faltung der Folgen (f_k) und (g_k) und verwendet die Kurzschreibweise

$$\mathfrak{Z}^{-1}\{\tilde{f}(z)\tilde{g}(z)\} = (f_k) * (g_k) = (g_k) * (f_k).$$

B.6 Differentiation im Bildbereich

Sei (f_k) eine Zahlenfolge mit zugehöriger Transformierter $\tilde{f}(z)$. Es gilt dann

$$\mathfrak{Z}\{(k f_k)\} = -z \frac{d\tilde{f}(z)}{dz}. \quad (\text{B.10})$$

B.7 Anfangswertsatz

Gegeben sei die z-Transformierte $\tilde{f}(z)$ einer Folge $(f_k) = (f_0, f_1, f_2, \dots)$. Das Element f_0 kann mit Hilfe des Anfangswertsatzes

$$f_0 = \lim_{z \rightarrow \infty} \tilde{f}(z) \quad (\text{B.11})$$

berechnet werden.

B.8 Endwertsatz

Gegeben sei die z-Transformierte $\tilde{f}(z)$ einer Folge (f_k) . Das Folgeelement f_k für $k \rightarrow \infty$ kann mit Hilfe des Endwertsatzes

$$f_\infty := \lim_{k \rightarrow \infty} f_k = \lim_{z \rightarrow 1} (z - 1) \tilde{f}(z) \quad (\text{B.12})$$

ermittelt werden. Man beachte, dass die Anwendung von (B.12) nur dann zulässig ist, wenn der Grenzwert f_∞ existiert. Dies kann ebenfalls im Bildbereich überprüft werden. Ist die Funktion $\tilde{g}(z) = (z - 1) \tilde{f}(z)$ eine gebrochen rationale Funktion, d.h.

$$\tilde{g}(z) = \frac{\kappa(z)}{\lambda(z)},$$

wobei $\kappa(z)$ und $\lambda(z)$ teilerfremde (gekürzte) Polynome sind, so existiert f_∞ genau dann, wenn das Polynom $\lambda(z)$ ausschließlich Nullstellen besitzt, die betragsmäßig kleiner als 1 sind. Das bedeutet, dass alle Nullstellen von $\lambda(z)$ im Einheitskreis der komplexen z-Ebene liegen. Ein solches Polynom wird auch **Einheitskreispolynom** oder **Schurpolynom** genannt.

B.9 Korrespondenztabelle

(f_k)	$\circ \rightarrow \bullet$	$\tilde{f}(z)$
$(\delta_k) = (1, 0, \dots)$		1
$(\sigma_k) = (1, 1, \dots)$		$\frac{z}{z-1}$
(k)		$\frac{z}{(z-1)^2}$
(k^2)		$\frac{z(z+1)}{(z-1)^3}$
(α^k)		$\frac{z}{z-\alpha}$
$(k\alpha^k)$		$\frac{z\alpha}{(z-\alpha)^2}$
$(k^2\alpha^k)$		$\frac{z(z+\alpha)\alpha}{(z-\alpha)^3}$
$(\cos \omega k T_d)$		$\frac{z(z - \cos \omega T_d)}{z^2 - 2z \cos \omega T_d + 1}$
$(\sin \omega k T_d)$		$\frac{z \sin \omega T_d}{z^2 - 2z \cos \omega T_d + 1}$
$(\alpha^k \cos \omega k T_d)$		$\frac{z(z - \alpha \cos \omega T_d)}{z^2 - 2z\alpha \cos \omega T_d + \alpha^2}$
$(\alpha^k \sin \omega k T_d)$		$\frac{z\alpha \sin \omega T_d}{z^2 - 2z\alpha \cos \omega T_d + \alpha^2}$
$(k \cos \omega k T_d)$		$\frac{z[(z^2 + 1) \cos \omega T_d - 2z]}{(z^2 - 2z \cos \omega T_d + 1)^2}$
$(k \sin \omega k T_d)$		$\frac{z(z^2 - 1) \sin \omega T_d}{(z^2 - 2z \cos \omega T_d + 1)^2}$

Literatur

- [1] ACKERMANN J.: Abtastregelung, 3. Auflage, Springer Verlag, 1988
- [2] CHEN C.T.: Analog and Digital Control System Design: Transfer-Function, State-Space, and Algebraic Methods, Saunders College Publishing, 1993
- [3] CHEN C.T.: Linear System Theory and Design, Saunders College Publishing, 1984
- [4] FÖLLINGER _ O.: Regelungstechnik, 6.Auflage, Hüthig Verlag, 1990
- [5] FÖLLINGER _ O.: Lineare Abtastsysteme, 4. Auflage, Oldenbourg Verlag, 1990
- [6] HORN M., DOURDOUMAS N.: Regelungstechnik, Pearson Verlag, 2004
- [7] KAILATH T.: Linear Systems, Prentice Hall, 1980
- [8] LANDGRAF C., SCHNEIDER G.: Elemente der Regelungstechnik, Springer Verlag, 1970
- [9] LUNZE J.: Regelungstechnik 1, 3. Auflage, Springer Verlag, 2001