

REGELUNGSTECHNIK

Professor Dr.-Ing. Ferdinand Svaricek

Institut für Steuer- und Regelungstechnik
Fakultät für Luft- und Raumfahrttechnik
Universität der Bundeswehr München

Vorwort

Diese Arbeitsblätter beschreiben den wesentlichen Stoffinhalt des MA-Moduls "Regelungstechnik", (V 2, SÜ 2), für Studentinnen und Studenten des Masterr-Studiengangs Luft- und Raumfahrttechnik an der *Universität der Bundeswehr München* ab WT 2010. Das Modul beginnt jedes Studienjahr jeweils im Wintertrimester. Als Startzeitpunkt ist das Wintertrimester im 1. MA-Studienjahr vorgesehen.

Mit dieser als Studienbegleittext zu Vorlesung und Übung verfaßten Schrift soll einmal die für die Fachprüfung verlangte Stoffauswahl abgegrenzt und zum anderen den Studentinnen und Studenten die Mitschrift erleichtert werden. Darüber hinaus sind wichtige mathematische Grundlagen zur Wiederholung und zum Nachschlagen im Anhang zusammengestellt.

Diese Studienhilfe ersetzt kein Lehrbuch, das zum Selbststudium geeignet wäre, vielmehr sind zum Verständnis des Stoffes Erläuterungen und Beispiele der Vorlesungsveranstaltungen sowie die aktive Mitarbeit bei den Übungen notwendig. Für die Nacharbeit und die Prüfungsvorbereitung sind die in der Literaturzusammenstellung genannten Bücher geeignet.

Allen Mitarbeitern des Instituts für Steuer- und Regelungstechnik, die bei den Lehrveranstaltungen mitwirken, möchte ich für die Unterstützung herzlich danken.

Neubiberg, im Januar 2010

F. Svaricek

Inhaltsverzeichnis

Literatur	4
1 Einleitung	5
2 Beschreibung und Analyse dynamischer Systeme im Bild- und Frequenzbereich	6
2.1 Die Frequenzgangdarstellung	6
2.2 Ortskurvendarstellung des Frequenzgangs	8
2.3 Frequenzkennliniendiagramm (Bode-Diagramm)	8
2.4 Eigenschaften wichtiger Übertragungsglieder	10
2.5 Allpaß- und Phasenminimumsysteme	26
2.6 Beispiele für die Konstruktion des Bode-Diagramms	33
2.7 Das Nyquistverfahren (Frequenzgangverfahren)	35
2.8 Stabilitätsreserve	40
2.9 Das Wurzelortskurvenverfahren (WOK-Verfahren)	42
3 Beschreibung und Analyse dynamischer Systeme im Zustandsraum	51
3.1 Zustandsraumdarstellung linearer zeitinvarianter Systeme	52
3.2 Graphische Darstellung des Zustandsraummodells	54
3.3 Lösung der Zustandsgleichung	55
3.4 Lösung der Zustandsgleichung im Frequenzbereich	58
3.5 Numerische Lösung der Zustandsgleichung mit Hilfe der Transitionsmatrix	58
3.6 Stabilitätsanalyse im Zustandsraum	59
3.6.1 Beziehungen zwischen asymptotischer Stabilität und BIBO-Stabilität	62
3.7 Steuerbarkeit und Beobachtbarkeit	63
3.7.1 Dualität von Steuerbarkeit und Beobachtbarkeit	68
3.8 Kalman-Zerlegung des Zustandsraummodells	68
3.9 Normalformen für Zustandsraummodelle	69
3.9.1 Jordannormalform	70
3.9.2 Regelungsnormalform	71
3.9.3 Beobachtungsnormalform	73
4 Entwurf linearer Zustandsregler	74
4.1 Einleitung	74
4.2 Entwurf einer Zustandsrückführung	74
4.3 Entwurf eines Zustandsbeobachters	77
4.3.1 Konvergenz des Beobachters	79
4.3.2 Berechnung der Beobachterrückführung	80
4.3.3 Separationstheorem	82
4.4 Zusammenfassung	82

A	Mathematische Grundlagen	84
A.1	Definition einer Matrix	84
A.2	Transponierte einer Matrix	84
A.3	Quadratische Matrix	84
A.4	Vektoren	85
A.5	Determinante einer Matrix	86
A.5.1	Determinanten von 2×2 -Matrizen	86
A.5.2	Determinanten von 3×3 -Matrizen mit Hilfe der Sarrusschen Regel	86
A.5.3	Determinanten beliebiger Matrizen	87
A.5.4	Rechenregeln für Determinanten	87
A.6	Inverse Matrix	87
A.7	Eigenwerte und Eigenvektoren	88
A.7.1	Cayley–Hamilton–Theorem	89
A.7.2	Ähnlichkeitstransformation	89
A.8	Rang einer Matrix	90
A.8.1	Rang von Matrizenprodukten	90
A.8.2	Term–Rang und generischer Rang	91
A.9	Orthogonale–Transformationen	93
A.10	Hessenberg–Form einer Matrix	96
A.11	Singulärwertzerlegung und Anwendungen	96
A.11.1	Norm und Konditionszahl einer Matrix	98
A.11.2	Numerische Berechnung der Singulärwertzerlegung	99
A.11.3	Numerischer Rang einer Matrix	100
B	Numerische Überprüfung der Zustandssteuerbarkeit	101
C	Stabilitätsdefinitionen	110
	Stichwortverzeichnis	113

Literatur

- Dickmanns, E.D.** 1985. *Systemanalyse und Regelkreissynthese*. Stuttgart: B.G. Teubner.
- Föllinger, O.** 1985. *Regelungstechnik*. Heidelberg: Hüthig.
- Lunze, J.** 2006. *Regelungstechnik 1: Systemtheoretische Grundlagen, Analyse und Entwurf einschleifiger Regelungen*. Berlin: Springer-Verlag. (5. Auflage).
- Lunze, J.** 2005. *Regelungstechnik 2: Mehrgrößensysteme, Digitale Regelung*. Berlin: Springer-Verlag. (3. Auflage).
- Ludyk, G.** 1990. *CAE von Dynamischen Systemen: Analyse, Simulation, Entwurf von Regelungssystemen*. Berlin: Springer-Verlag.
- Schwarz, H.** 1976. *Frequenzgang und Wurzelortskurvenverfahren*. Mannheim: Bibliographisches Institut. (2. Auflage. BI-Hochschultaschenbücher).
- Svaricek, F.** 1995. *Zuverlässige numerische Analyse linearer Regelungssysteme*. Stuttgart: B.G. Teubner.
- Unbehauen, H.** 2005. *Regelungstechnik. Band I*. Braunschweig: Vieweg und Sohn (13. Auflage).
- Unbehauen, H.** 2000. *Regelungstechnik. Band II*. Braunschweig: Vieweg und Sohn (8. Auflage).
- Voigt, C.** und **J. Adamy.** 2007. *Formelsammlung der Matrizenrechnung*. München: Oldenbourg-Verlag.

1 Einleitung

Im Modul Regelungstechnik wird das in der BA–Grundlagenvorlesung ”Steuer– und Regelungstechnik” erworbene Grundwissen zur Analyse und Synthese linearer kontinuierlicher Regelungs- und Steuerungssysteme vertieft und erweitert.

Zunächst wird die Beschreibung und Analyse des Verhaltens von dynamischen Systemen im Bild- und Frequenzbereich fortgesetzt, wobei folgende Begriffe und Methoden erläutert werden:

- Frequenzgang,
- Ortskurve und Bodediagramm,
- Phasenminimum- und Allpaßsysteme,
- Wurzelortskurvenverfahren (WOK-Verfahren),
- Stabilitätsprüfung mit Hilfe des Nyquistverfahrens,
- Stabilitätsrand (Amplituden- und Phasenrand).

Im zweiten Teil wird die Beschreibung und Analyse dynamischer Eingrößensysteme im Zustandsraum behandelt. Neben der Vorstellung der grundlegenden Begriffe und Definitionen (Zustandsvariablen, Zustandsgleichung, Zustandstrajektorie, Steuer- und Beobachtbarkeit, Stabilität) werden auch elementare Methoden zur Analyse des dynamischen Verhaltens (Eigenwerte, Eigenbewegung, Stabilität) und der Steuer- und Beobachtbarkeit sowie wichtige Normalformen (Diagonal-, Regelungs- und Beobachtungsnormalform, Kalman-Zerlegung) vermittelt.

Abschließend werden dann die Grundlagen zur Synthese linearer Zustandsregler vermittelt. Dabei werden im Einzelnen die folgenden Aspekte behandelt:

- Entwurf linearer Zustandsrückführungen,
- Entwurf von Beobachtern zur Schätzung des Verlaufs der Zustandsgrößen.

2 Beschreibung und Analyse dynamischer Systeme im Bild- und Frequenzbereich

Die im BA-Modul "Steuer- und Regelungstechnik" eingeführte Übertragungsfunktion $G(s)$ stellt eine abstrakte, nicht meßbare Beschreibungsform zur mathematischen Behandlung linearer Systeme dar. Setzt man in $s = \sigma + j\omega$ den Realteil σ zu Null, so geht die komplexe Übertragungsfunktion $G(s)$ in den komplexen *Frequenzgang* $G(j\omega) = \text{Re}(\omega) + j\text{Im}(\omega)$ über, der physikalisch interpretiert und gemessen werden kann.

2.1 Die Frequenzgangdarstellung

Zur Ermittlung des Frequenzgangs wird ein System zunächst mit einem sinusförmigen Eingangssignal

$$u(t) = \hat{u} \sin(\omega t - \varphi_0) \quad (2.1)$$

erregt (Bild 2.1).

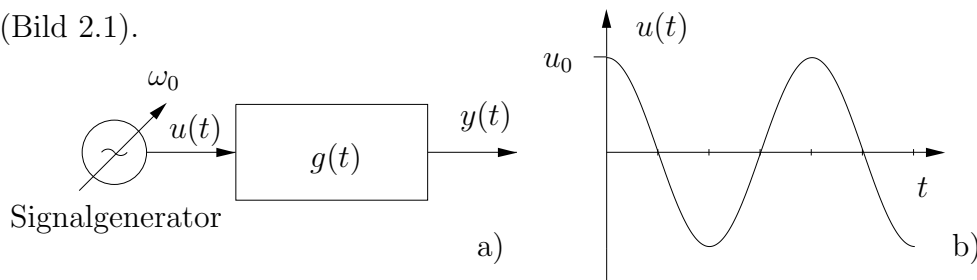


Bild 2.1: Zur Einführung des Frequenzgangverfahrens

- a) Blockschaltbild eines Systems mit harmonischer Erregung
- b) harmonisches Eingangssignal

Betrachtet man das Ausgangssignal $y(t)$ nach Abklingen des Einschwingvorgangs, so wird die Ausgangsgröße mit derselben Frequenz ω , jedoch mit anderer Amplitude \hat{y} und mit einer gewissen Phasenverschiebung $\varphi = \omega\Delta t$, ebenfalls sinusförmige Schwingungen ausführen (Bild 2.2).

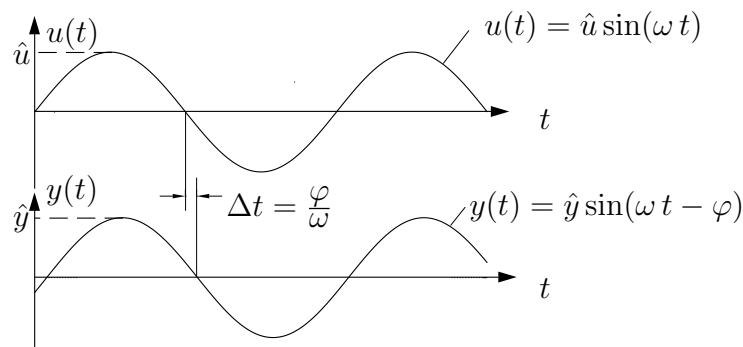


Bild 2.2: Sinusförmiges Eingangssignal $u(t)$ und zugehöriges Ausgangssignal $y(t)$ eines linearen Übertragungsgliedes

Wird dieses Experiment für verschiedene Frequenzen $\omega_i, (i = 0, 1, 2, \dots)$ mit $\hat{u} = \text{const}$ durchgeführt, so stellt man eine Frequenzabhängigkeit der Ausgangsamplitude \hat{y} und der Phasenverschiebung φ fest. Aus dem Verhältnis der Amplituden $\hat{y}(\omega)$ und \hat{u} läßt sich der **Amplitudengang** des Frequenzgangs

$$\frac{\hat{y}(\omega)}{\hat{u}} = |G(j\omega)| = \sqrt{\text{Re}(G(j\omega))^2 + \text{Im}(G(j\omega))^2} \quad (2.2)$$

als frequenzabhängige Größe definieren. Die frequenzabhängige Phasenverschiebung $\varphi(\omega)$ bezeichnet man als **Phasengang** des Frequenzgangs:

$$\varphi(\omega) = \arg G(j\omega) = \arctan \frac{\text{Im} G(j\omega)}{\text{Re} G(j\omega)} . \quad (2.3)$$

Man erhält damit folgende Zeigerdarstellung (Polarkoordinaten) des Frequenzgangs

$$G(j\omega) = |G(j\omega)|e^{j\varphi(\omega)} . \quad (2.4)$$

Der Frequenzgang $G(j\omega)$ für $\omega = 0$ bis $\omega \rightarrow \infty$ beschreibt wie die Übertragungsfunktion $G(s)$ oder die Gewichtsfunktion $g(t)$ das Übertragungsverhalten eines linearen, zeitinvarianten Systems vollständig.

Die Systemkenngröße $G(j\omega)$ – der Frequenzgang eines Systems – kann auf unterschiedliche Weise gedeutet werden:

- a) $G(j\omega)$ bestimmt die Signalübertragungseigenschaften eines Systems für harmonische Signale und kann beispielsweise experimentell bestimmt werden.
- b) $G(j\omega)$ ist die Fouriertransformierte der Gewichtsfunktion $G(j\omega) = \mathcal{F}\{g(t)\}$.
- c) $G(j\omega)$ ist eine Randfunktion der komplexen Übertragungsfunktion $G(s)$:

$$G(j\omega) = G(s)|_{s=j\omega} . \quad (2.5)$$

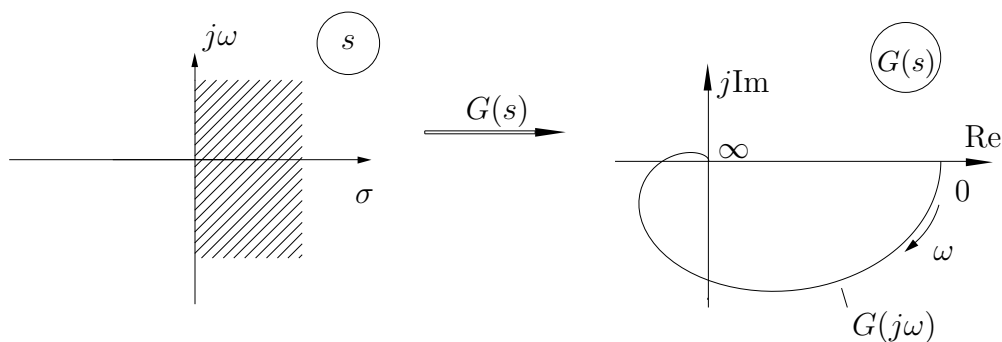


Bild 2.3: Deutung des Frequenzganges als Abbildung der (positiven) imaginären Achse der s -Ebene in die $G(s)$ -Ebene

Die s -Ebene wird durch die imaginäre Achse in zwei Teilgebiete geteilt. Die $j\omega$ -Achse stellt den Rand z. B. der rechten s -Halbebene dar (Bild 2.3).

2.2 Ortskurvendarstellung des Frequenzgangs

Die Ortskurve des Frequenzgangs ist die grafische Darstellung von $G(j\omega)$ in der komplexen Ebene für $\omega = 0 \dots \infty$ und wird auch *Nyquist-Ortskurve* genannt. Faßt man den Frequenzgang als Zeiger mit der Länge $|G(j\omega)|$ und dem Winkel $\varphi(j\omega)$ auf, so beschreibt die Zeigerspitze die Ortskurve des Systems, wenn man ω von Null an vergrößert.

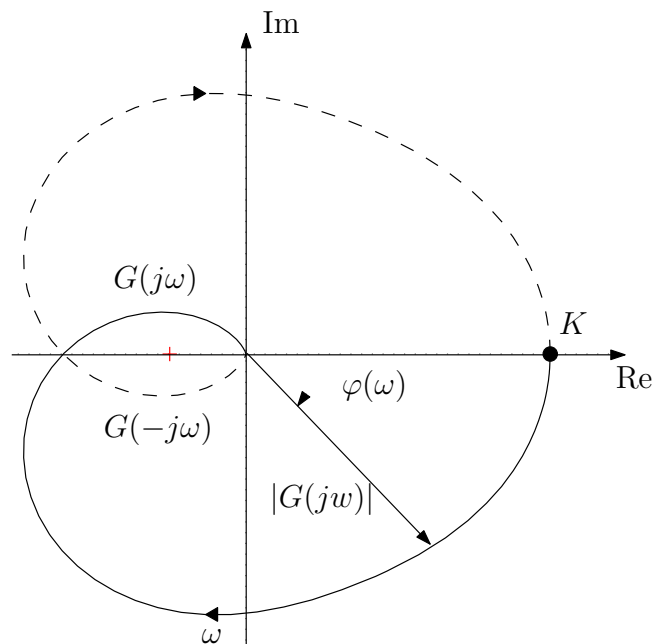


Bild 2.4: Nyquist-Ortskurve eines Systems mit Ausgleich.

Für Systeme mit Ausgleich beginnt die Ortskurve für $\omega = 0$ aufgrund der Eigenschaft

$$\lim_{t \rightarrow \infty} h(t) = \lim_{s \rightarrow 0} sH(s) = \lim_{s \rightarrow 0} G(s) = \lim_{j\omega \rightarrow 0} G(j\omega) \quad (2.6)$$

bei der Systemverstärkung K . Für hohe Frequenzen nähert sich die Ortskurve dem Ursprung der komplexen Ebene. Gestrichelt dargestellt ist in Bild 2.4 der Verlauf der Ortskurve für $\omega = -\infty \dots 0$, den man durch Spiegelung der Ortskurve für $\omega = 0 \dots \infty$ an der reellen Achse erhält.

2.3 Frequenzkennliniendiagramm (Bode-Diagramm)

Die getrennte Darstellung des Betrages und der Phase des Frequenzganges $G(j\omega) = |G(j\omega)|e^{j\varphi(\omega)}$ in Abhängigkeit von der Kreisfrequenz ω wird *Frequenzkennlinien-Darstellung* von $G(j\omega)$ genannt. Die beiden Kennlinien werden als *Betragskennlinie* (Amplitudengang) bzw. *Phasenkennlinie* (Phasengang) bezeichnet (Bild 2.5).

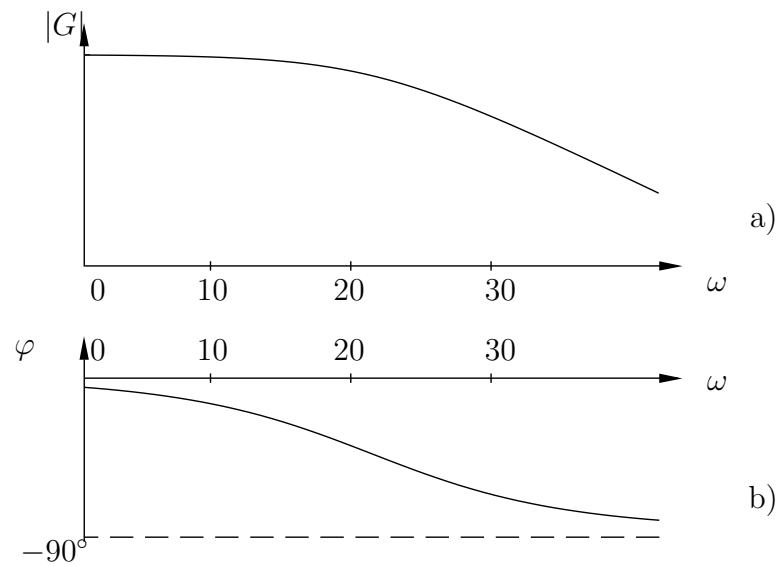


Bild 2.5: Darstellung des Frequenzganges nach a) Betrag b) Phase in Abhängigkeit der Frequenz ω

Eine besonders aussagekräftige Darstellung des Frequenzganges erhält man in Form des **BODE-Diagramms** mit einer logarithmischen Skalierung der Frequenzachse.

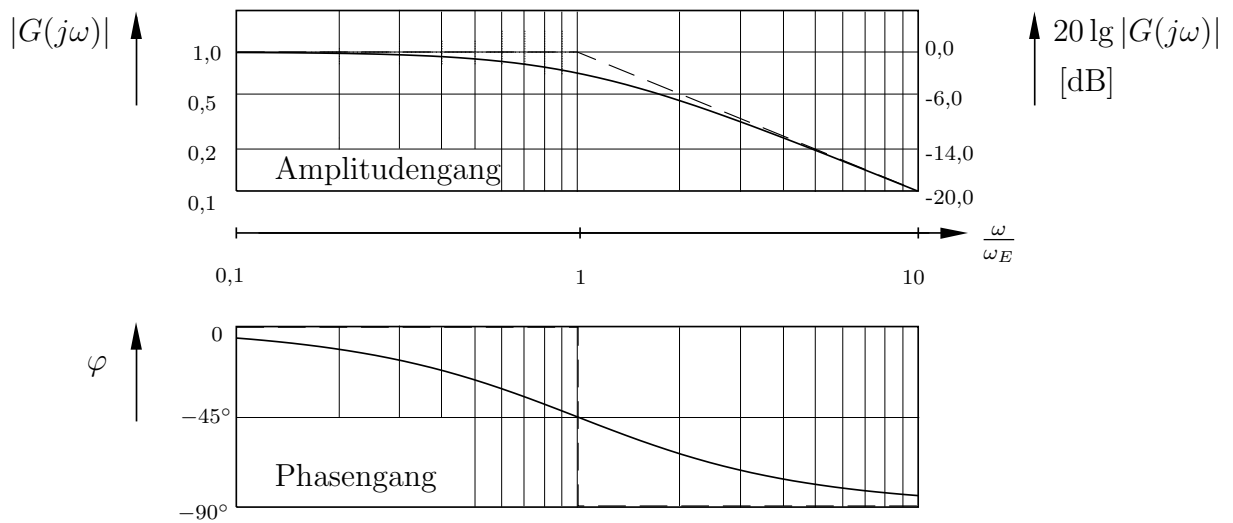


Bild 2.6: Frequenzkennlinie des PT_1 -Gliedes im Bode-Diagramm mit der Zeitkonstante T , der Eckfrequenz $\omega_E = 1/T$ und der Verstärkung $K = 1$

Darüber hinaus wird der Betrag des Amplitudengangs $|G(j\omega)|$ im Bode-Diagramm als $20\log_{10} |G(j\omega)| = 20 \lg |G(j\omega)|$ in Dezibel (dB) aufgetragen (Bild 2.6).

Die Darstellung der logarithmischen Frequenzkennlinie hat den großen Vorteil, daß der Amplitudengang und der Phasengang zweier in Reihe geschalteter Systeme durch (graphische) Addition der einzelnen Amplituden- bzw. Phasengänge gefunden werden kann.

Aus

$$G(j\omega) = G_1(j\omega) \cdot G_2(j\omega) \quad \text{mit} \quad G_i(j\omega) = |G_i(j\omega)| e^{j\varphi_i(\omega)}, \quad i = 1, 2 \quad (2.7)$$

folgt:

$$G(j\omega) = |G_1(j\omega)| |G_2(j\omega)| e^{j(\varphi_1(\omega) + \varphi_2(\omega))} . \quad (2.8)$$

Für den Amplitudengang der Reihenschaltung erhält man also

$$20 \lg |G(j\omega)| = 20 \lg |G_1(j\omega)| + 20 \lg |G_2(j\omega)| \quad (2.9)$$

sowie für den Phasengang

$$\varphi(\omega) = \varphi_1(\omega) + \varphi_2(\omega) . \quad (2.10)$$

2.4 Eigenschaften wichtiger Übertragungsglieder

In diesem Abschnitt werden systematisch die wichtigsten Regelkreisglieder parallel im Zeit- und im Bildbereich besprochen und ihre wesentlichen Eigenschaften und deren Kennzeichnung zusammenfassend und übersichtlich dargestellt.

a) PT_1 -System: (System mit Verzögerung 1. Ordnung)

Das PT_1 -System wird durch die Differentialgleichung

$$T\dot{y}(t) + y(t) = K u(t) \quad (2.11)$$

beschrieben, aus der mittels Laplace-Transformation die komplexe Übertragungsfunktion gewonnen werden kann:

$$\left. \begin{aligned} (Ts + 1)Y(s) &= K U(s) \\ G(s) &= \frac{K}{1 + sT} \end{aligned} \right\} \quad (2.12)$$

Die Sprungantwort $h(t)$ kann dann mittels Rücktransformation (siehe Tabelle A2 im SRT-Skript) gefunden werden, wenn das Sprungsignal als Eingangsgröße – bzw. genauer dessen Laplace-Transformierte $\mathcal{L}\{1(t)\} = \frac{1}{s}$ – berücksichtigt wird:

$$\left. \begin{aligned} h(t) &= \mathcal{L}^{-1} \left\{ \frac{1}{s} \cdot \frac{K}{1 + sT} \right\} \\ &= K (1 - e^{-t/T}) \cdot 1(t) \end{aligned} \right\} \quad (2.13)$$

Für den Frequenzgang erhalten wir zunächst aus

$$G(j\omega) = G(s)|_{s=j\omega} \quad (2.14)$$

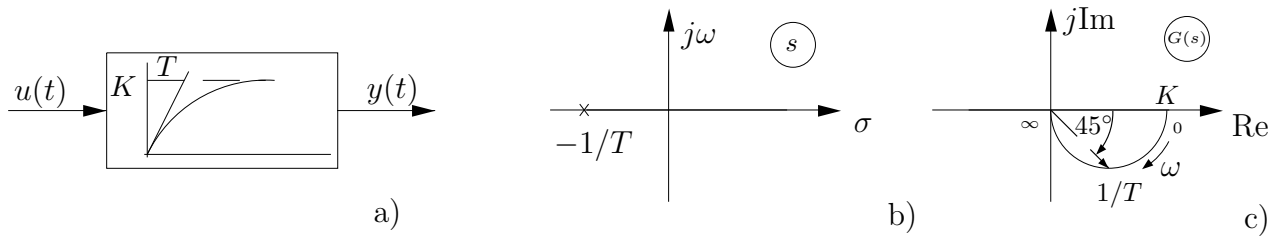


Bild 2.7: PT₁-System

- a) Blockschaltbild mit Sprungantwort
- b) Pole in der s -Ebene
- c) Frequenzgang in der $G(s)$ -Ebene

$$\begin{aligned}
 G(j\omega) &= \frac{K}{1 + j\omega T} = \frac{K(1 - j\omega T)}{1 + \omega^2 T^2} = \frac{K}{1 + \omega^2 T^2} - j \frac{K\omega T}{1 + \omega^2 T^2} \\
 &= \operatorname{Re}(G) + j\operatorname{Im}(G) ,
 \end{aligned} \tag{2.15}$$

der im Bild 2.7c) dargestellt ist.

Durch Logarithmieren erhält man hierzu die Darstellung im Bode-Diagramm:

$$\ln G(j\omega) = \ln |G| + j\varphi(\omega) = \ln \sqrt{\operatorname{Re}^2(G) + \operatorname{Im}^2(G)} + j \arctan \frac{\operatorname{Im}(G)}{\operatorname{Re}(G)} \tag{2.16}$$

oder durch Übergang auf Dezibel

$$20 \lg |G| = 20 \lg K - 20 \lg \sqrt{1 + \omega^2 T^2} \begin{cases} \approx 20 \lg K \text{ für } \omega T \ll 1 \\ = 20 \lg K - 3 \text{ für } \omega T = 1 \\ \approx 20 \lg \frac{K}{\omega T} \text{ für } \omega T \gg 1 \end{cases} \tag{2.17}$$

$$\varphi(\omega) = \arctan(-\omega T) = -\arctan(\omega T) \tag{2.18}$$

Ein Blick auf Bild 2.8 zeigt, daß der Amplitudengang in sehr guter Näherung durch zwei Asymptoten dargestellt werden kann, wobei die zweite Asymptote mit 20 dB/Dekade abfällt. Die beiden Asymptoten schneiden sich bei der Frequenz $\omega_E = 1/T$, die deshalb auch *Knickfrequenz* genannt wird. Als *Bandbreite* eines Systems bezeichnet man die Frequenz bei der der logarithmische Amplitudengang gegenüber der horizontalen Asymptote um 3 dB abgefallen ist. Aus Gl. (2.17) folgt, daß bei einem PT₁-System die Eckfrequenz auch die Bandbreite des Systems bestimmt. Da ein PT₁-System niederfrequente Signale gut überträgt und hochfrequente schlecht, wird es auch als *Tiefpaß* bezeichnet.

Aus (2.18) ergeben sich für den Phasengang folgende Näherungen: Für $\omega T \ll 1$ ist die Phase etwa 0 und für $\omega T \gg 1$ etwa -90° . Für einfache Überschlagsbetrachtungen kann sogar der Phasengang, wie angedeutet, grob durch eine Treppenkurve approximiert werden.

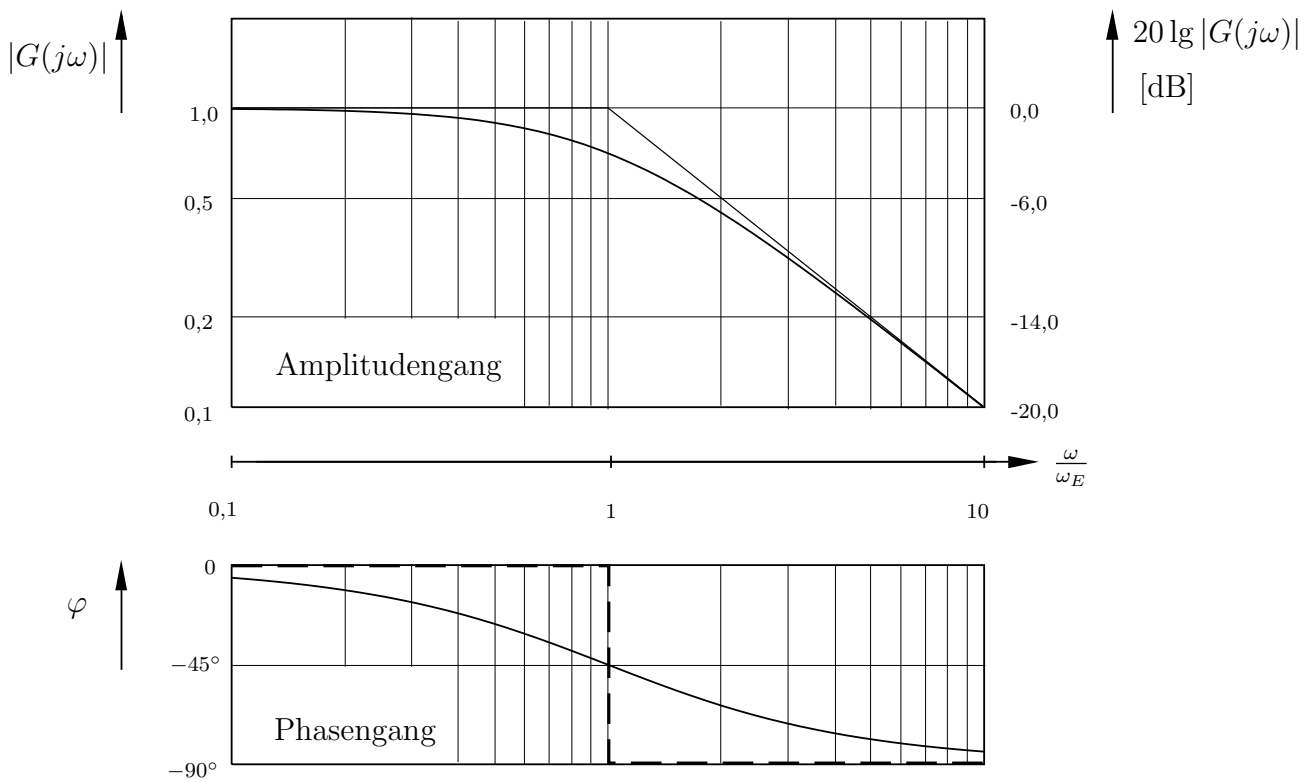


Bild 2.8: Frequenzkennlinie des PT_1 -Gliedes im Bode-Diagramm mit der Zeitkonstante T und der Eckfrequenz $\omega_E = 1/T$ und der Verstärkung $K = 1$

In Bild 2.9 sind die Abweichungen der Verläufe von $|G|$ und φ von den Asymptoten dargestellt:

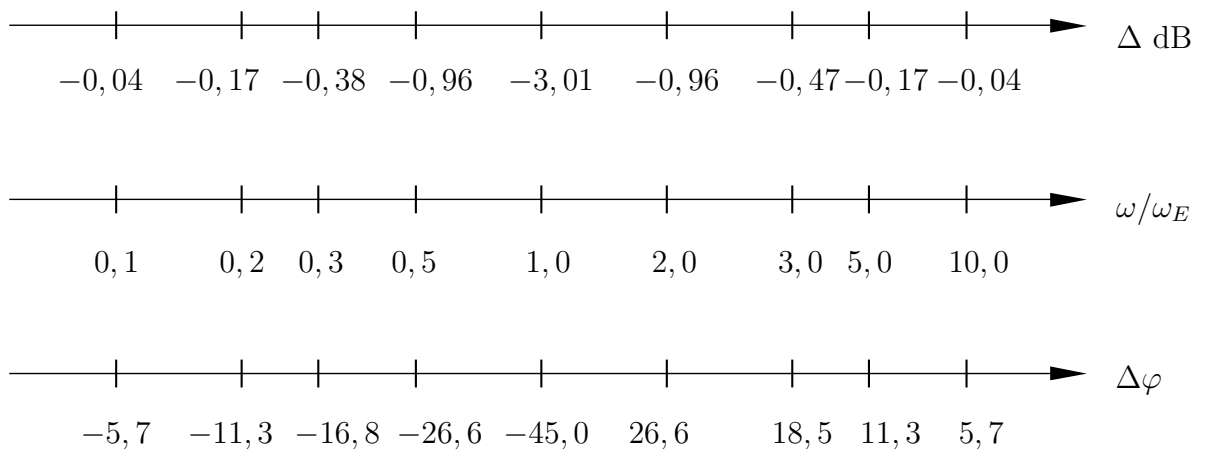


Bild 2.9: Abweichung des Amplituden- und Phasenganges des PT_1 -Systems von der Asymptote

b) **PT₂-System (reelle Pole)**

Das PT₂-System wird durch die Differentialgleichung

$$T_1 T_2 \ddot{y}(t) + (T_1 + T_2) \dot{y}(t) + y(t) = K u(t) \quad (2.19)$$

beschrieben und hat im Falle reeller Eigenwerte diese Übertragungsfunktion $G(s)$ bzw. Frequenzgangfunktion $G(j\omega)$ (Reihenschaltung zweier PT₁-Glieder):

$$\left. \begin{aligned} G(s) &= \frac{K}{(1 + sT_1)(1 + sT_2)} ; \\ G(j\omega) &= \frac{K}{(1 + j\omega T_1)(1 + j\omega T_2)} \end{aligned} \right\} \quad (2.20)$$

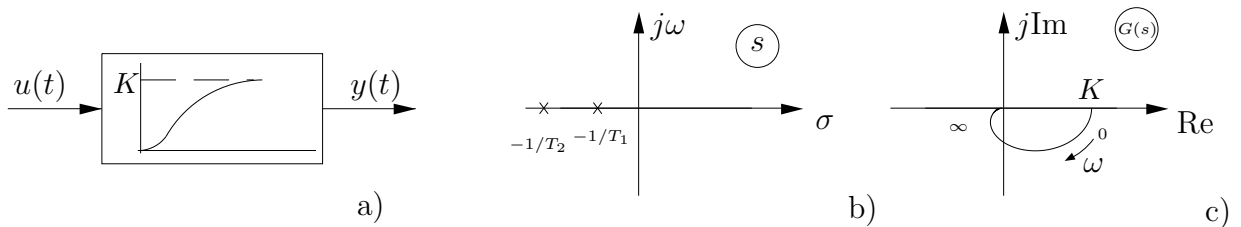


Bild 2.10: PT₂-System

- a) Blockschaltbild mit Übergangsfunktion
- b) Pole in der s -Ebene
- c) Frequenzgang in der $G(s)$ -Ebene

c) **PT₂-System (komplexe-Pole)**

Ist ein PT₂-System schwingungsfähig, ist es zweckmäßig, die Differentialgleichung in dieser normierten Form

$$\begin{aligned} \frac{1}{\omega_0^2} \ddot{y}(t) + \frac{2D}{\omega_0} \dot{y}(t) + y(t) &= K u(t) \\ \omega_0 &= \text{Eigenfrequenz des ungedämpften Schwingers;} \\ D &= \text{Dämpfungsgrad} \end{aligned} \quad (2.21)$$

zu schreiben, aus der dann auch die Übertragungsfunktion bzw. der Frequenzgang folgen zu:

$$G(s) = \frac{K}{\frac{s^2}{\omega_0^2} + \frac{2D}{\omega_0} s + 1} \quad (2.22)$$

$$G(j\omega) = \frac{K}{1 + j \frac{2D}{\omega_0} \omega - \left(\frac{\omega}{\omega_0}\right)^2} = \frac{K \omega_0^2}{\omega_0^2 - \omega^2 + 2jD\omega_0\omega} \quad (2.23)$$

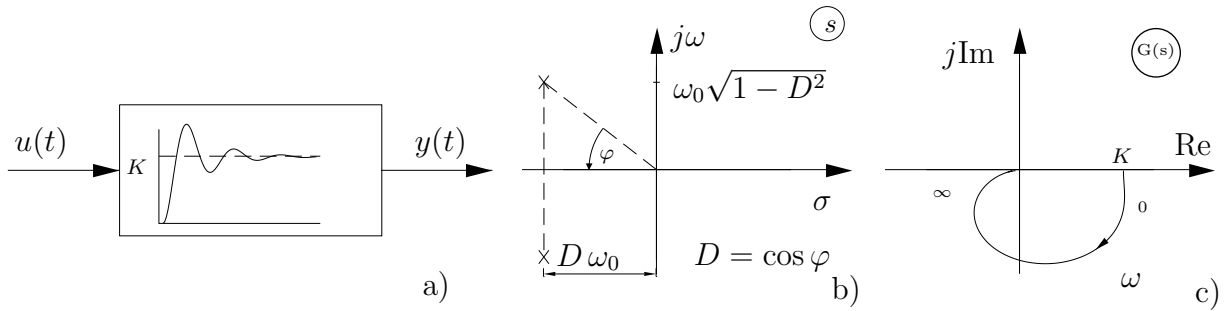


Bild 2.11: Schwingungsfähiges PT₂-System

- a) Blockschaubild mit Übergangsfunktion
- b) Pole in der s -Ebene
- c) Frequenzgang in der $G(s)$ -Ebene

In Bild 2.12 ist das Kennlinienfeld des PT₂-Systems im Bode-Diagramm für die auf eins normierte Eckfrequenz und Verstärkung in Abhängigkeit des Dämpfungsgrades dargestellt. Diese Diagramme sind für die Auslegung von Regelungssystemen wichtig, da häufig das Schwingungsverhalten eines Regelkreises mit dem eines Schwingers 2. Ordnung verglichen wird.

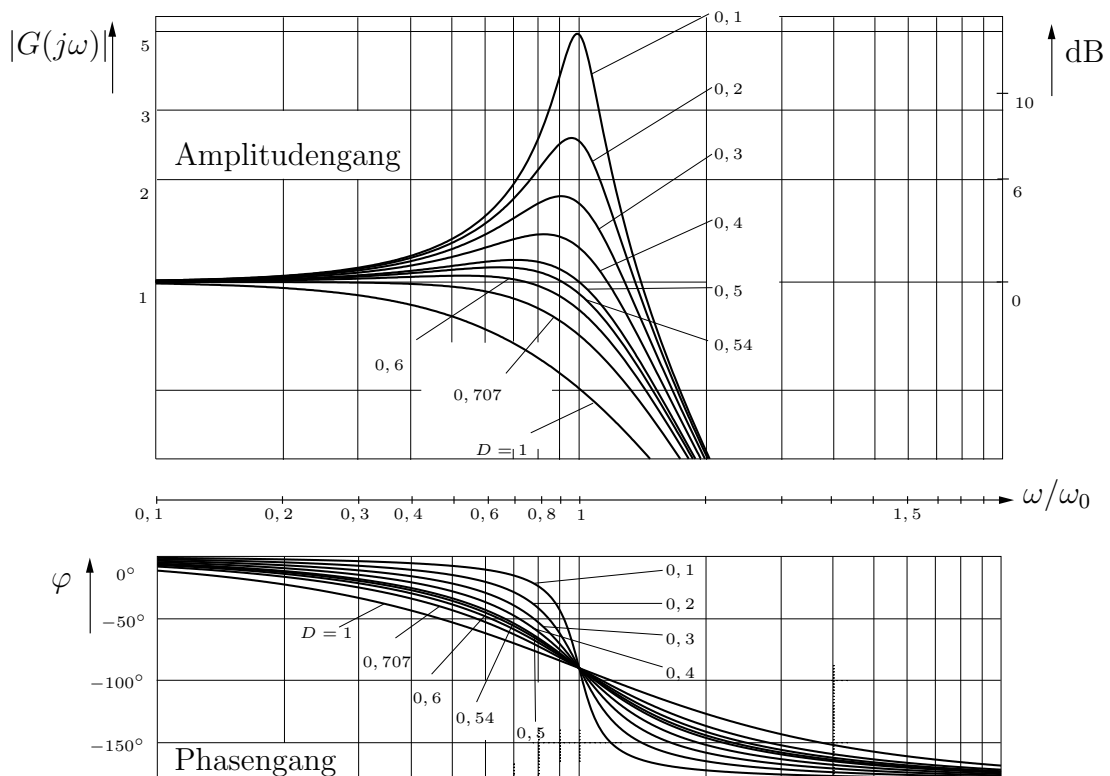


Bild 2.12: Bode-Diagramm des Verzögerungssystems 2. Ordnung

Der logarithmische Amplitudengang des PT₂-Systems kann für große bzw. kleine

Frequenzen ähnlich wie beim PT_1 -System durch Geraden approximiert werden. Es gilt

$$|G(j\omega)|_{\text{dB}} = 20 \lg K - 20 \lg \sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \left(2D \frac{\omega}{\omega_0}\right)^2},$$

woraus für $\frac{\omega}{\omega_0} \ll 1$

$$|G(j\omega)|_{\text{dB}} \approx 20 \lg K \quad (2.24)$$

und für $\frac{\omega}{\omega_0} \gg 1$

$$|G(j\omega)|_{\text{dB}} \approx 20 \lg K - 20 \lg \left(\frac{\omega}{\omega_0}\right)^2 = 20 \lg K - 40 \lg \frac{\omega}{\omega_0} \quad (2.25)$$

folgt. Das bedeutet, daß die zweite Asymptote bei einem schwingungsfähigen PT_2 -System mit 40 dB/Dekade abfällt. Der Schnittpunkt liegt dann bei der Eckfrequenz $\omega_E = \omega_0$.

Dämpfung	Systemeigenschaft	Pole
$D > 1$	überkritisch gedämpft	negative reelle Pole
$D = 1$	kritisch gedämpft	negativer reeller Doppelpol
$\frac{1}{\sqrt{2}} < D < 1$	gedämpft ohne Resonanzüberhöhung	konjugiert komplexe Pole mit negativen Realteilen
$0 < D < \frac{1}{\sqrt{2}}$	gedämpft mit Resonanzüberhöhung	konjugiert komplexe Pole mit negativen Realteilen
$D = 0$	ungedämpft	konjugiert komplexe Pole mit verschwindenden Realteilen
$-1 < D < 0$	instabil	konjugiert komplexe Pole mit positiven Realteilen
$D = -1$	instabil	positiv reeller Doppelpol
$D < -1$	instabil	positive reelle Pole

Tabelle 2.1: Verhalten von PT_2 -Systemen

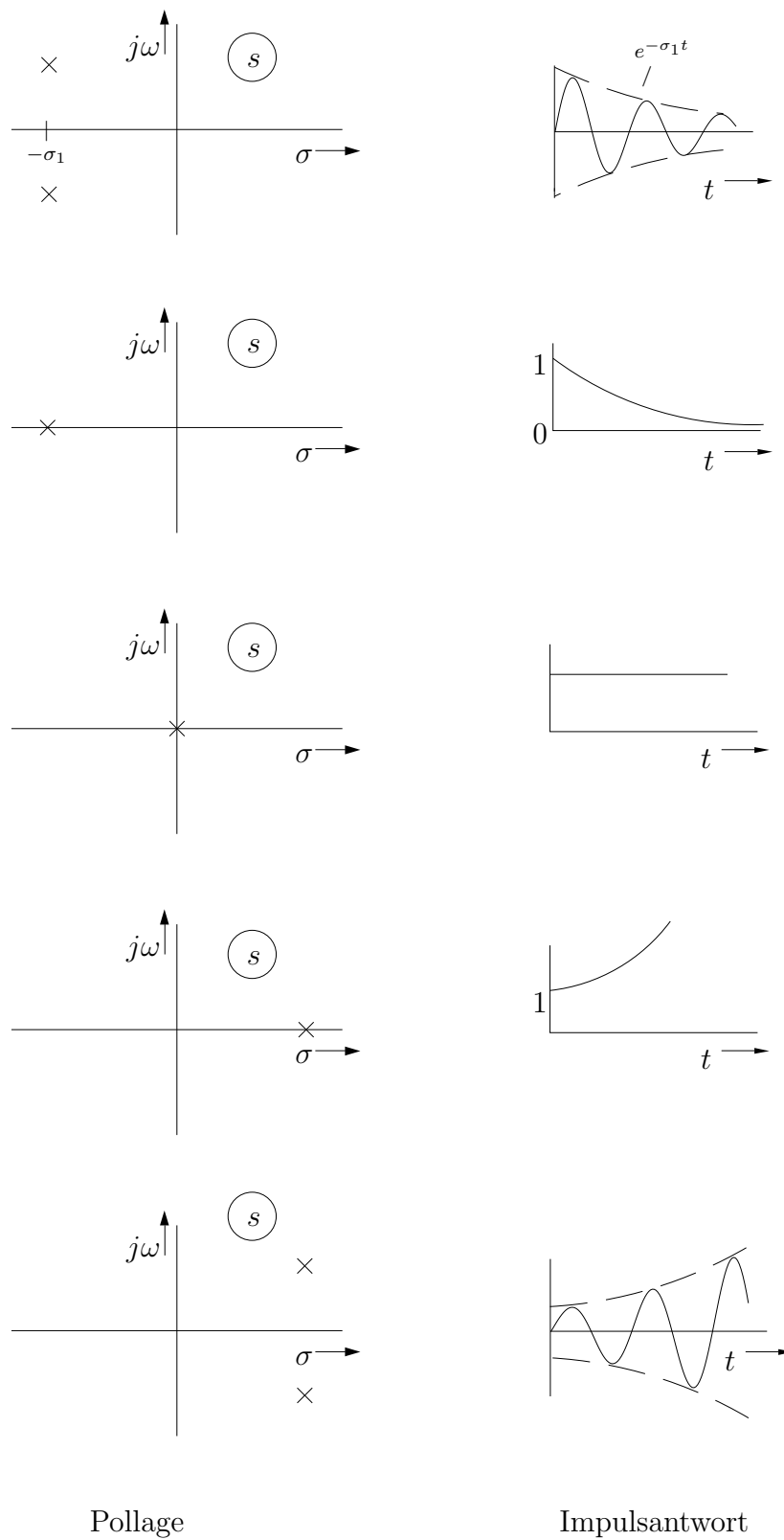


Bild 2.13: Abhängigkeit der Impulsantwort von der Lage der Pole in der s -Ebene

Der Zusammenhang zwischen Dämpfungsgrad, Lage der Pole und dem Verhalten der Impulsantwort sind in Tabelle 2.1 und Bild 2.13 dargestellt.

d) **PT_n-Systeme**

Die PT_n-Systeme (Reihenschaltung von n PT₁-Gliedern) werden durch eine Differentialgleichung n -ter Ordnung beschrieben:

$$a_n y^{(n)}(t) + \dots + a_2 \ddot{y}(t) + a_1 \dot{y}(t) + y(t) = K u(t) \quad . \quad (2.26)$$

Die komplexe Übertragungsfunktion

$$G(s) = \frac{K}{1 + a_1 s + a_2 s^2 + \dots + a_n s^n} \quad (2.27)$$

hat n Pole, die in die s -Ebene eingetragen werden können, und der Frequenzgang

$$G(j\omega) = \frac{K}{1 + a_1 j\omega - a_2 \omega^2 + \dots + a_n (j\omega)^n} \quad (2.28)$$

verläuft durch n Quadranten der $G(s)$ -Ebene (Bild 2.14c).

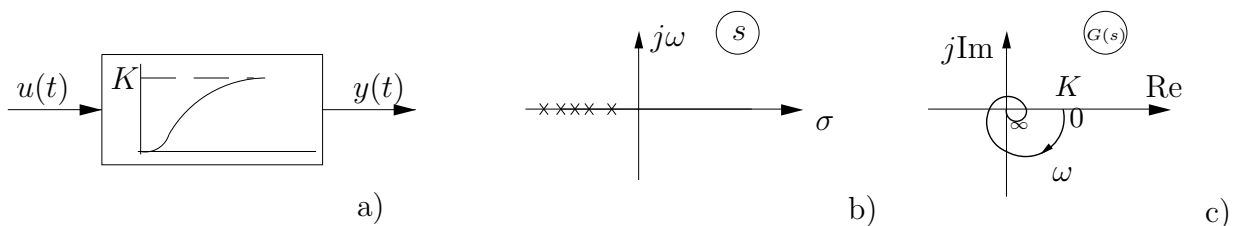


Bild 2.14: PT_n-System

- a) Blockschaltbild mit Übergangsfunktion
- b) Pole in der s -Ebene
- c) Frequenzgang in der $G(s)$ -Ebene

e) **I-Systeme**

Systeme, deren Ausgang das Integral des Systemeinganges ist, treten als Regelstrecken in Form der Strecken ohne Selbstausgleich (Füllstand, Antriebsregelstrecke, Lageregelung von Fahrzeugen) und auch als Regler auf. Ein I-Glied mit zusätzlichen zeitlichen Verzögerungen wird im Zeitbereich durch die Differentialgleichung

$$a_n y^{(n)}(t) + \dots + a_2 \ddot{y}(t) + \dot{y}(t) = K_I u(t) \quad (2.29)$$

beschrieben. Die zugehörige Übertragungsfunktion ist durch einen Pol im Ursprung ($s = 0$) ausgezeichnet:

$$\left. \begin{aligned}
 G(s) &= \frac{K_I}{s + a_2 s^2 + \dots + a_n s^n} \\
 &= \frac{K_I}{s} \cdot \frac{1}{(1 + \bar{a}_1 s + \dots + \bar{a}_{n-1} s^{n-1})} \\
 \text{mit } \bar{a}_v &= a_{v+1}, v = 1, 2, \dots, n-1
 \end{aligned} \right\} \quad (2.30)$$

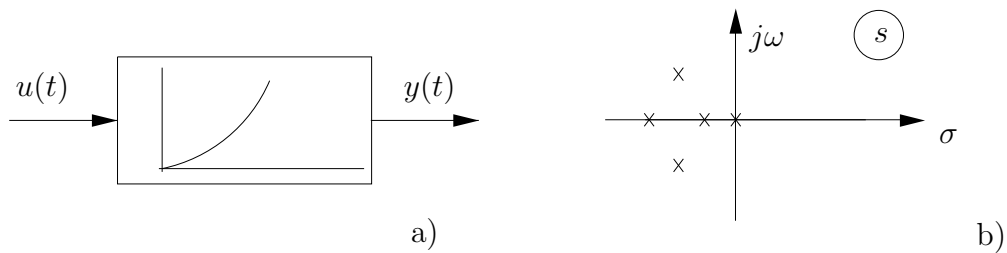


Bild 2.15: I-Systeme a) Blockschaltbild mit Sprungantwort b) Polstellen in der s -Ebene

Die Frequenzgänge der I-Glieder kommen für $\omega = 0$ aus dem Unendlichen (Bild 2.16).

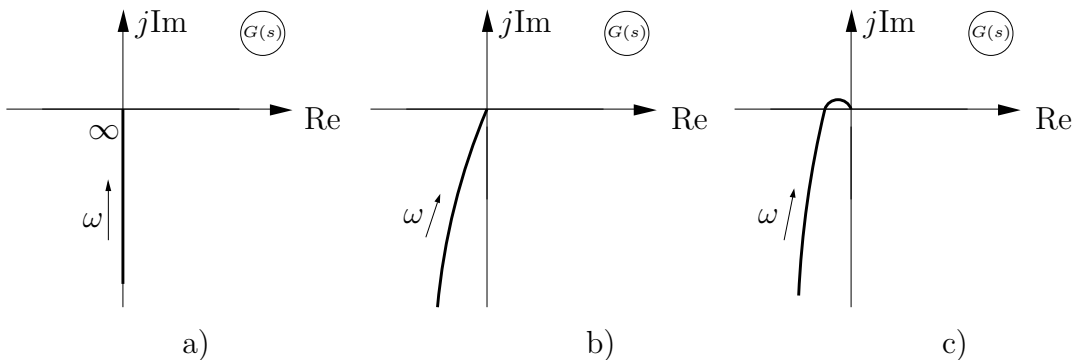


Bild 2.16: Frequenzgänge von I-Systemen a) ohne Verzögerung b) IT₁-System c) IT₂-System

Die Ortskurve des Frequenzganges

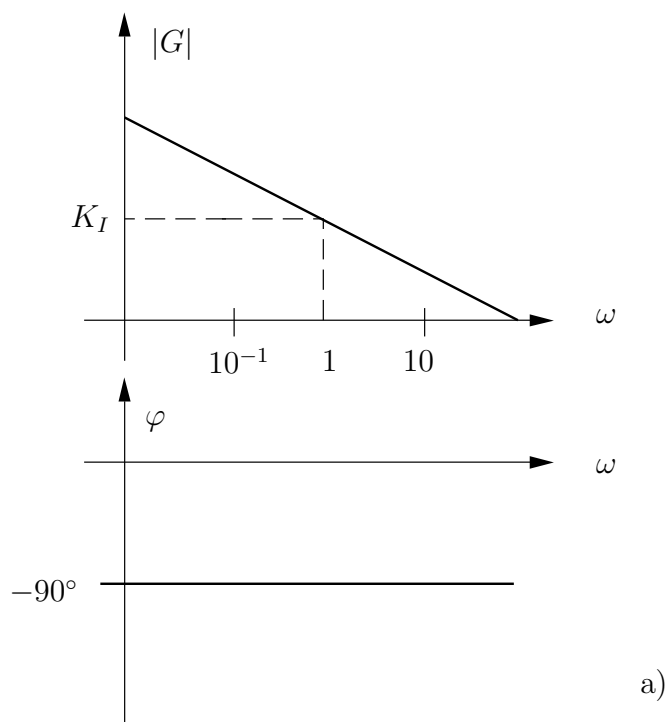
$$G(j\omega) = \frac{K_I}{j\omega}$$

eines reinen I-Gliedes liegt auf der negativen Imaginärachse, der Amplitudengang ist eine Gerade

$$|G(j\omega)|_{\text{dB}} = 20 \lg K_I - 20 \lg \omega$$

mit einer Neigung von -20 dB/Dekade, die die Frequenzachse bei $\omega = K_I$ schneidet.

$$G(j\omega) = \frac{K_I}{j\omega}$$



$$G(j\omega) = \frac{K_I}{j\omega(1 + j\omega T)}$$

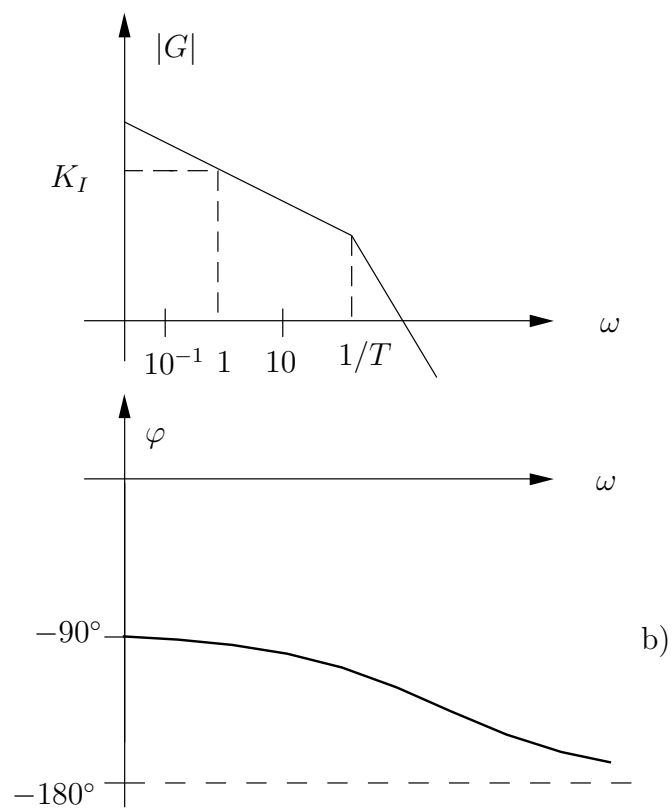


Bild 2.17: Skizzen im Bode-Diagramm a) I-System b) IT₁-System

f) D-Systeme

Das verzögerungsfreie *D-Glied* wird durch die Gleichung

$$y(t) = T_D \dot{u}(t) \tag{2.31}$$

beschrieben und hat die Übertragungsfunktion

$$G(s) = T_D s . \tag{2.32}$$

Die Ortskurve des zugehörigen Frequenzganges liegt auf der positiven imaginären Achse der komplexen Zahlenebene. Der Amplitudengang

$$|G(j\omega)|_{dB} = 20 \lg T_D \omega = 20 \lg T_D + 20 \lg \omega$$

stellt im Bodediagramm eine Gerade mit einer Steigung von 20 dB/Dekade dar. Der Schnittpunkt mit der Abszisse liegt bei $\omega = 1/T_D$.

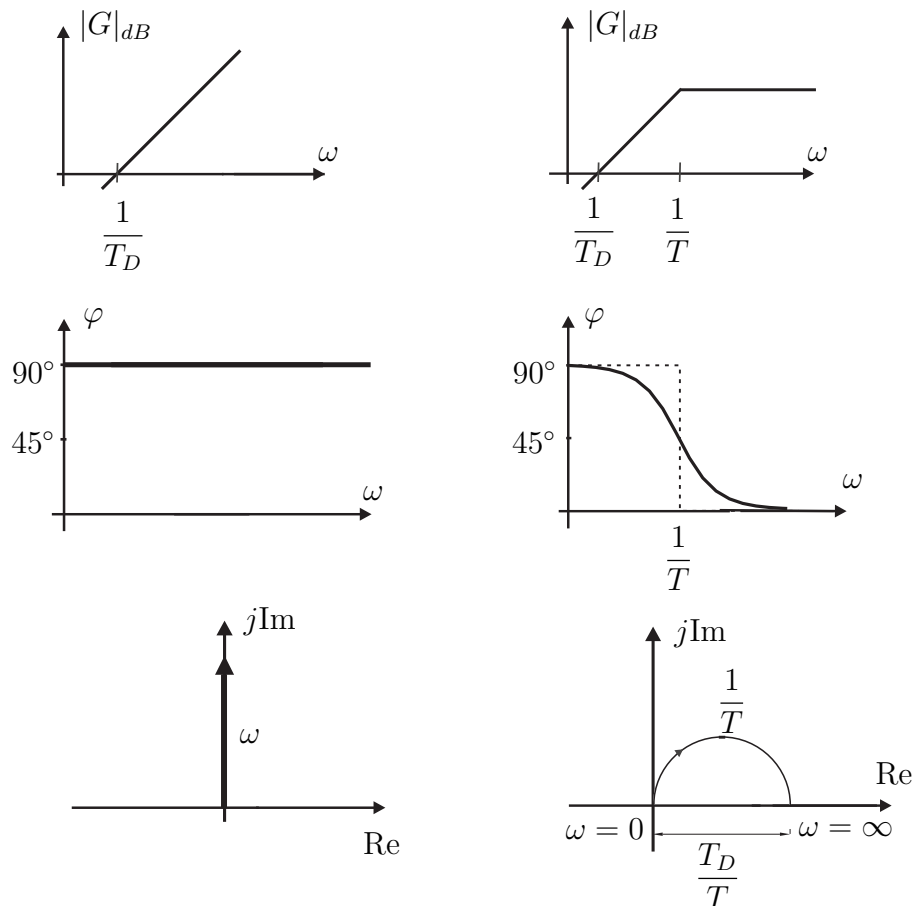


Bild 2.18: Bode-Diagramm und Ortskurve eines D- und eines DT_1 -Gliedes

Aus der Differentialgleichung eines D-Gliedes erster Ordnung mit Verzögerung

$$T \dot{y}(t) + y(t) = T_D \dot{u}(t) \tag{2.33}$$

erhält man die Übertragungsfunktion

$$G(s) = \frac{T_D s}{T s + 1} . \quad (2.34)$$

Der Frequenzgang dieses Systems ist in Bild 2.18 dargestellt. Ein solches System wird auch *Hochpaß-Filter* genannt, da es Signale hoher Frequenz besser überträgt als Signale niedriger Frequenz.

g) **PD-Systeme**

PD-Glieder weisen sowohl ein proportionales als auch ein differenzierendes Übertragungsverhalten auf:

$$a_n y^{(n)}(t) + \dots + a_1 \dot{y}(t) + a_0 y(t) = K (u(t) + T_D \dot{u}(t)) \quad . \quad (2.35)$$

Die Übertragungsfunktion bzw. Frequenzgangfunktion haben die Form

$$G(s) = \frac{K (1 + s T_D)}{(1 + s T_1) (1 + s T_2) \dots (1 + s T_n)} \quad (2.36)$$

bzw.

$$G(j\omega) = \frac{K (1 + j\omega T_D)}{(1 + j\omega T_1) \dots (1 + j\omega T_n)} \quad (2.37)$$

Ein PD-Glied ohne Verzögerung mit der Übertragungsfunktion

$$G(s) = K(1 + s T_D) \quad (2.38)$$

kann aus der Parallelschaltung eines P- und eines D-Gliedes gebildet werden (Bild 2.19)

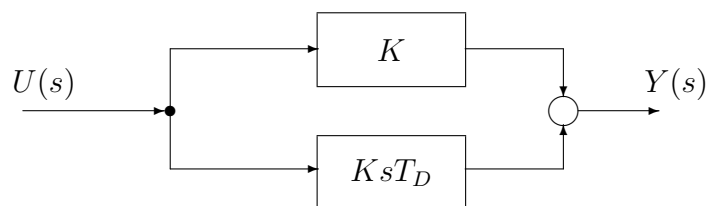


Bild 2.19: PD-Glied als Parallelschaltung eines P- und eines D-Gliedes

und stellt abgesehen vom Verstärkungsfaktor K das zum PT_1 -Glied inverse Glied dar.

Für $K = 1$ ergibt sich daher der Verlauf des logarithmischen Amplitudenganges sowie des Phasenganges (Bild 2.20) durch entsprechende Spiegelung an der 0-dB-Achse bzw. an der Linie $\varphi(\omega) = 0$.

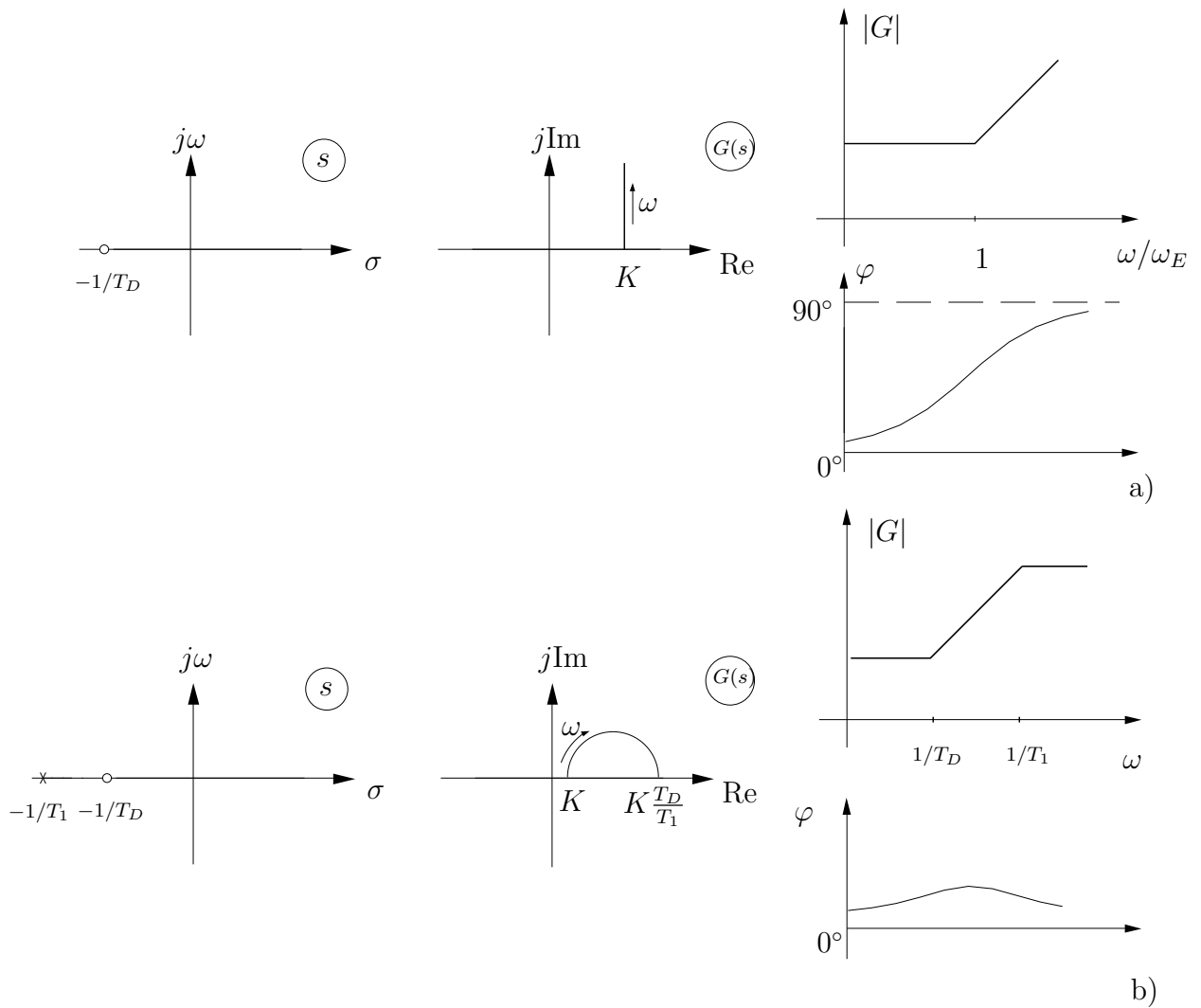


Bild 2.20: Pol-/Nullstellen, Frequenzgang und Bode-Diagramm von PD-Systemen
a) ohne Verzögerung b) mit Verzögerung 1. Ordnung

h) PI-Systeme

Diese Systeme genügen der Differential-Integralgleichung

$$a_n y^{(n)}(t) + \dots + a_1 \dot{y}(t) + y(t) = K \left(u(t) + \frac{1}{T_I} \int_0^t u(\tau) d\tau \right) \quad (2.39)$$

bzw. nach Differentiation der Differentialgleichung

$$a_n y^{(n+1)}(t) + \dots + a_1 \ddot{y}(t) + \dot{y}(t) = K \left(\dot{u}(t) + \frac{1}{T_I} u(t) \right) \quad (2.40)$$

mit der Übertragungsfunktion

$$G(s) = K \left(1 + \frac{1}{sT_I} \right) \frac{1}{(1 + sT_1)(1 + sT_2) \cdots (1 + sT_n)} \quad (2.41)$$

$$= \frac{K}{sT_I} \frac{(1 + sT_I)}{(1 + sT_1)(1 + sT_2) \cdots (1 + sT_n)}$$

Diese Systeme treten vor allem in der Form des idealen PI-Reglers (ohne Verzögerung)

$$G(s) = K \left(1 + \frac{1}{sT_I} \right)$$

und des PI-Reglers mit Verzögerung 1. Ordnung auf (Bild 2.21).

$$G(s) = K \left(1 + \frac{1}{sT_I} \right) \frac{1}{(1 + sT_1)}$$

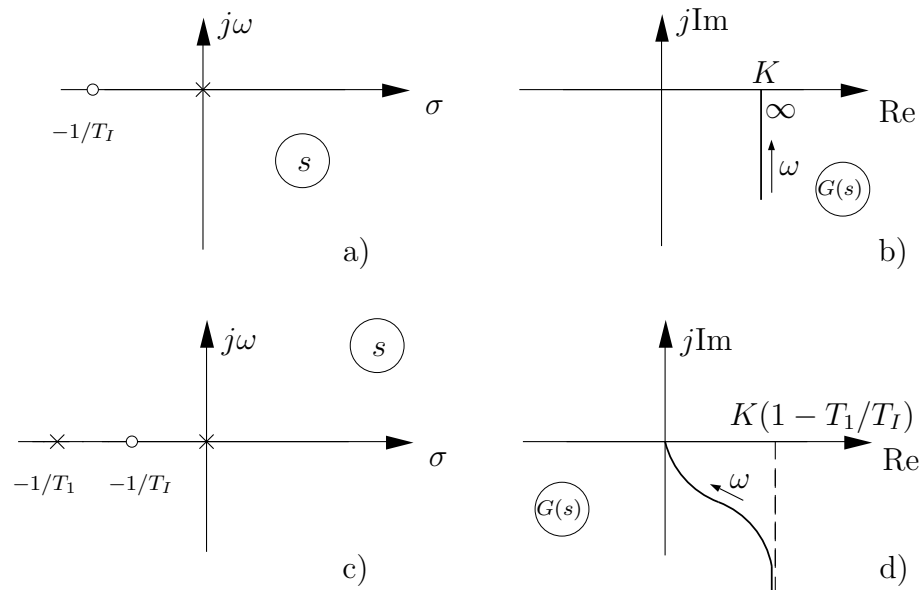


Bild 2.21: Pol-Nullstellen bzw. Frequenzgang von PI-Systemen a) bzw. b) ohne Verzögerung c) bzw. d) mit Verzögerung 1. Ordnung ($T_1 < T_I$)

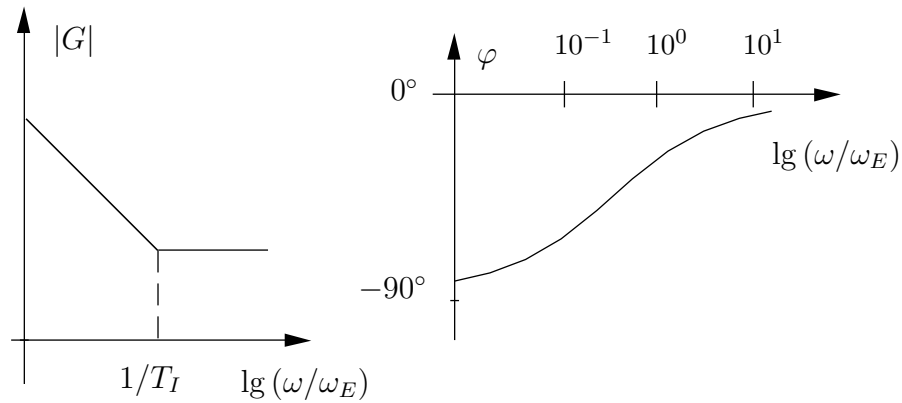


Bild 2.22: Bode-Diagramm des PI-Systems ohne Verzögerung

i) **PID-Systeme**

Das PID-System tritt vorwiegend in der Form des PID-Reglers ohne und mit Verzögerung 1. Ordnung auf.

$$T_1 \dot{y}(t) + y(t) = K \left(u(t) + \frac{1}{T_I} \int_0^t u(\tau) d\tau + T_D \dot{u}(t) \right) \quad (2.42)$$

$$T_1 \ddot{y}(t) + \dot{y}(t) = K \left(T_D \ddot{u}(t) + \dot{u}(t) + \frac{1}{T_I} u(t) \right) \quad (2.43)$$

$$\begin{aligned} \text{mit } G(s) &= K \left(1 + \frac{1}{sT_I} + sT_D \right) \frac{1}{1 + sT_1} \\ &= \frac{K}{sT_I} \frac{(1 + T_a s)(1 + T_b s)}{(1 + sT_1)}, \end{aligned} \quad (2.44)$$

wobei T_a und T_b mit T_I und T_D wie folgt zusammenhängen:

$$T_{a/b} = \frac{T_I}{2} \left(1 \pm \sqrt{1 - \frac{4T_D}{T_I}} \right) \quad (2.45)$$

Ein verzögerungsfreies PID-Glied mit der Übertragungsfunktion

$$G(s) = K \left(1 + \frac{1}{sT_I} + sT_D \right)$$

läßt sich wiederum als Parallelschaltung je eines P-, I- und D-Gliedes darstellen (Bild 2.23).

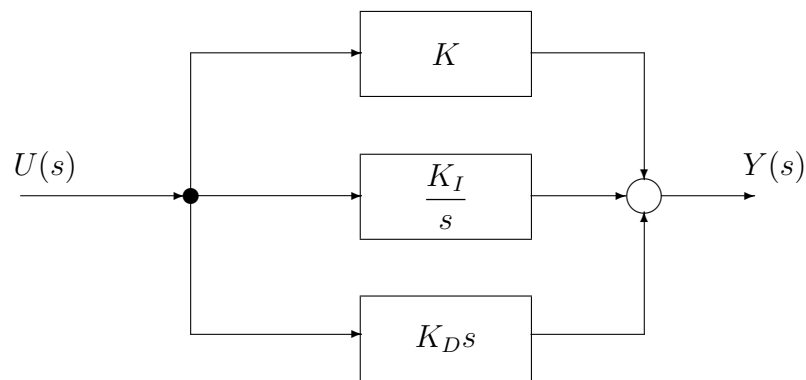


Bild 2.23: Blockschaltbild eines PID-Systems

Hierbei gilt

$$K_I = \frac{K}{T_I}$$

und

$$K_D = K T_D .$$

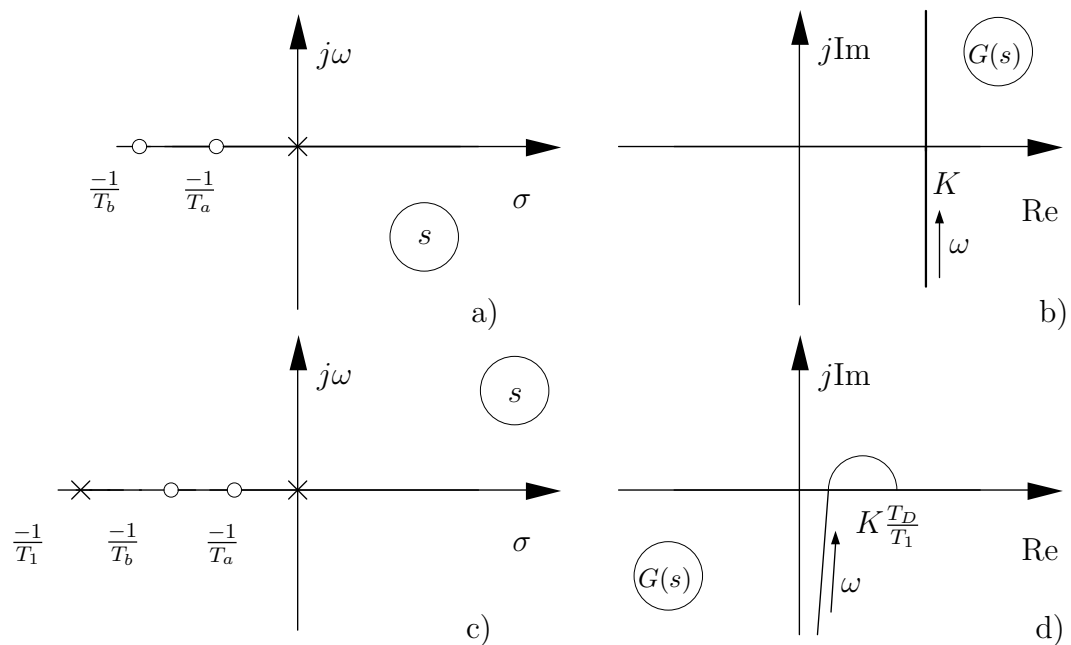


Bild 2.24: Pol-Nullstellen bzw. Frequenzgang von PID-Systemen a) bzw. b) ohne Verzögerung c) bzw. d) mit Verzögerung 1. Ordnung ($0 \leq T_D \leq \frac{1}{4}T_I$)

2.5 Allpaß- und Phasenminimumsysteme

In der Klasse der linearen, zeitinvarianten Systeme kann folgende Einteilung vorgenommen werden:

- 1) Stabile und instabile Systeme
- 2) Innerhalb der stabilen Systeme:
 - a) Phasenminimumsysteme
 - b) Nichtphasenminimumsysteme (Allpaßhaltige Systeme)

Definition 2.1

Phasenminimumsysteme sind Systeme ohne Totzeit, deren rationale Übertragungsfunktionen $G(s) = \frac{Z(s)}{N(s)}$ nur Pole und Nullstellen in der linken s -Halbebene haben. \square

Die Phasenminimumsysteme haben die Eigenschaft – und daher ihre Bezeichnung –, daß sie bei vorgegebenem Amplitudengang den kleinstmöglichen zugehörigen Phasengang haben, der in einem berechenbaren festen Zusammenhang zu ihrem Amplitudengang steht.

Ein System hat ein *nichtminimales Phasenverhalten*, wenn die zugehörige Übertragungsfunktion Pole und/oder Nullstellen in der rechten s -Halbebene aufweist. Der Phasenverlauf ist betragsmäßig stets größer als bei dem entsprechenden System mit Minimalphasenverhalten, das denselben Amplitudengang besitzt.

Beispiel:

Gegeben seien zwei stabile Systeme mit den Übertragungsfunktionen

$$G_1(s) = \frac{1 + sT}{1 + sT_1}$$

und

$$G_2(s) = \frac{1 - sT}{1 + sT_1},$$

wobei $0 < T < T_1$ gilt.

Der Amplitudengang der zugehörigen Frequenzgänge $G_1(j\omega)$ (Minimalphasensystem) und $G_2(j\omega)$ (Nichtminimalphasensystem) ist für beide Systeme gleich, da

$$|G_1(j\omega)| = |G_2(j\omega)| = \sqrt{\frac{1 + (\omega T)^2}{1 + (\omega T_1)^2}}$$

gilt. Für die Phasengänge ergibt sich mit

$$\varphi_1(\omega) = -\arctan \frac{\omega(T_1 - T)}{1 + \omega^2 T_1 T}$$

und

$$\varphi_2(\omega) = -\arctan \frac{\omega(T_1 + T)}{1 - \omega^2 T_1 T}$$

jedoch ein unterschiedlicher Verlauf (Bild 2.25).

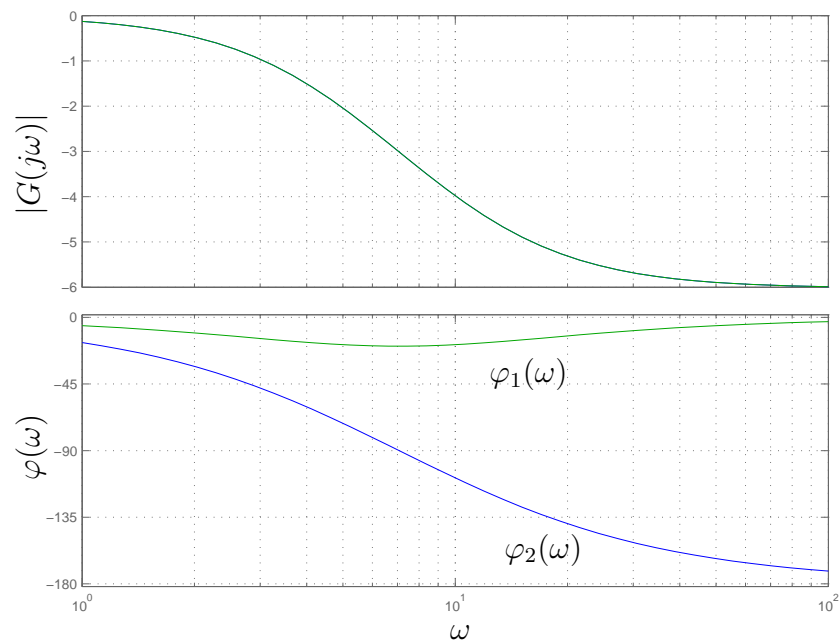


Bild 2.25: Beispiel für minimales und nichtminimales Phasenverhalten

Jedes Nichtphasenminimumsystem kann als Reihenschaltung eines Phasenminimumsystems und eines reinen phasendrehenden Gliedes dargestellt werden:

$$G_2(s) = G_A(s)G_1(s) . \tag{2.46}$$

Ein solches phasendrehendes Glied wird *Allpaßglied* genannt und ist wie folgt definiert:

Definition 2.2

Allpaßsysteme sind Systeme, die für alle Frequenzen den **konstanten** Amplitudengang $|G(j\omega)| = 1$ haben. □

Allpässe mit konzentrierten Energiespeichern sind stabile Systeme, deren Nullstellen spiegelbildlich zu den stabilen Polen in der rechten s -Halbebene liegen:

$$G_A(s) = \frac{N(-s)}{N(s)} \quad (2.47)$$

In der Tabelle 2.2 sind die Kenngrößen der Allpässe 1. und 2. Ordnung zusammengestellt, die auch das typische Verhalten (zunächst Reaktion in die falsche Richtung) eines Allpaßgliedes für eine sprungförmige Erregung zeigen.

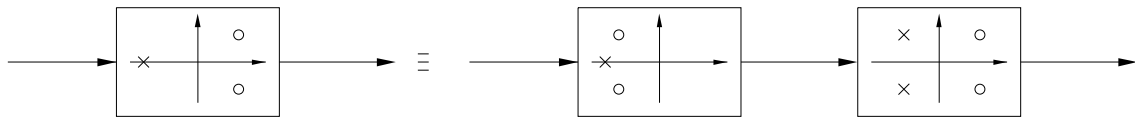


Bild 2.26: Darstellung eines Nichtphasenminimumsystems als Reihenschaltung eines Phasenminimumsystems und eines Allpasses

Diese Eigenschaft überträgt sich auch auf Systeme die aus einer Reihenschaltung eines Allpaßgliedes und einem Phasenminimumsystem bestehen, d.h. auf alle nichtminimalphasigen Systeme.

Eine wichtige Klasse der in der Regelungstechnik auftretenden Allpässe sind die Totzeitsysteme, die durch die Übertragungsgleichung im Zeitbereich

$$y(t) = K u(t - T_t) \quad (2.48)$$

beschrieben werden, und deren Frequenzgangfunktion die transzendente Funktion

$$G(j\omega) = K e^{-j\omega T_t} \quad (2.49)$$

ist, die mittels des Verschiebungssatzes der Fourier- oder Laplace-Transformation gefunden werden kann.

Das Totzeitsystem repräsentiert einen idealen **Allpaß**, da es alle harmonischen Signale ohne Amplitudendämpfung mit einer Phasenverschiebung proportional der Frequenz überträgt (Bild 2.27).

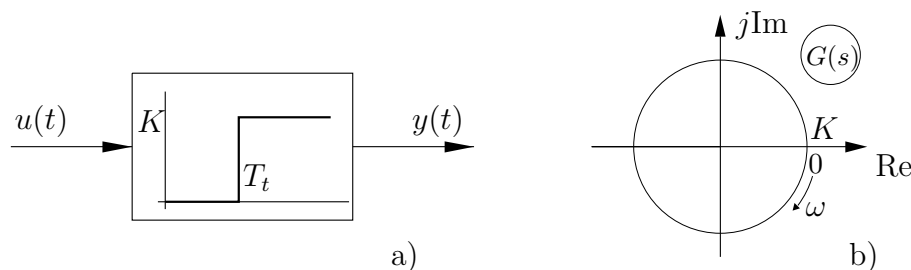


Bild 2.27: Totzeitsystem als Allpaß

a) Blockschaltbild b) Frequenzgang in der $G(s)$ -Ebene

Tabelle 2.2: Allpaßsysteme

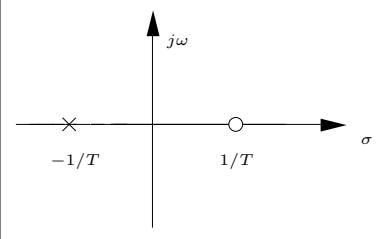
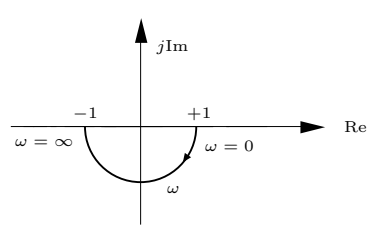
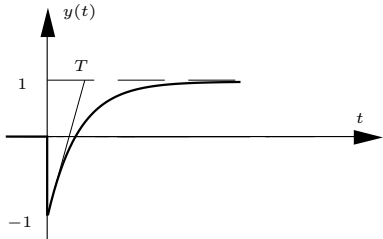
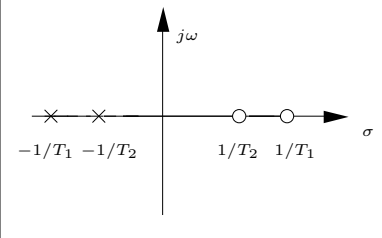
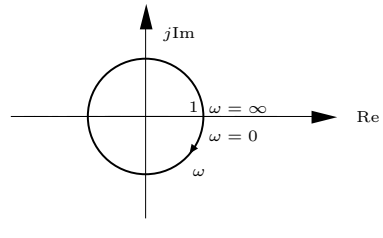
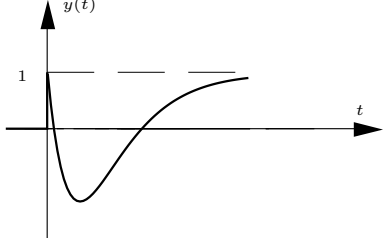
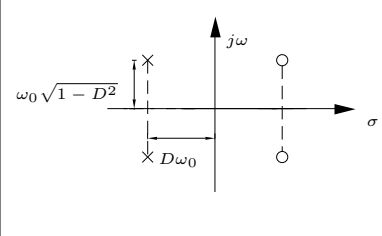
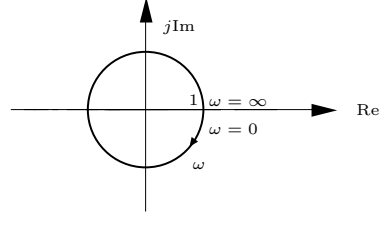
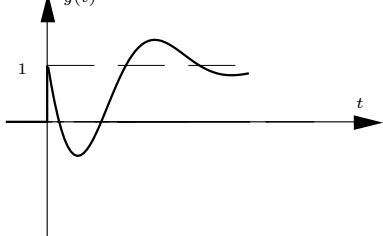
System	Übertragungsfunktion	s-Ebene	Ortskurve	Übergangsfunktion
Allpaß 1. Ordnung	$G(s) = \frac{1 - sT}{1 + sT}$			
Allpaß 2. Ordnung mit reellen Wurzeln	$G(s) = \frac{(1 - sT_1)(1 - sT_2)}{(1 + sT_1)(1 + sT_2)}$			
Allpaß 2. Ordnung mit komplexen Wurzeln	$G(s) = \frac{1 - \frac{2D}{\omega_0}s + \frac{s^2}{\omega_0^2}}{1 + \frac{2D}{\omega_0}s + \frac{s^2}{\omega_0^2}}$			

Tabelle 2.3: Verhalten der wichtigsten Regelkreisglieder

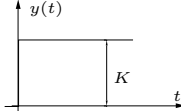
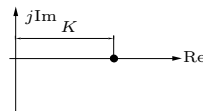
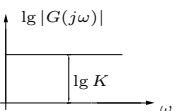
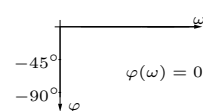
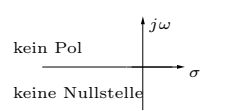
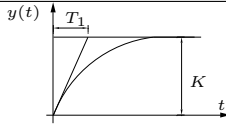
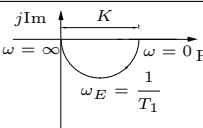
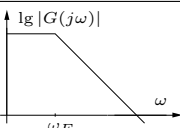
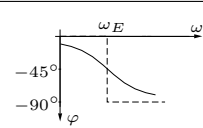
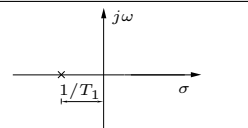
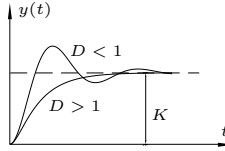
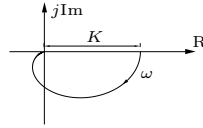
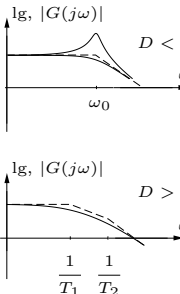
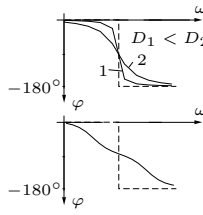
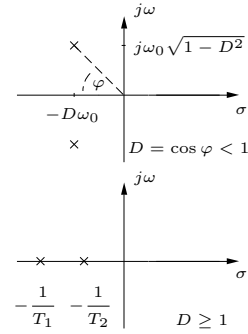
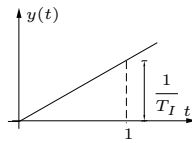
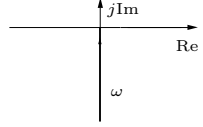
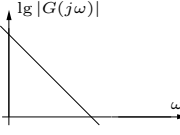
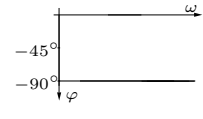
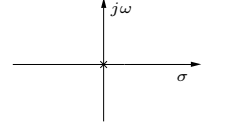
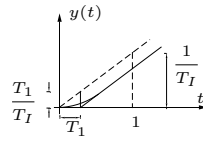
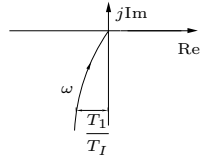
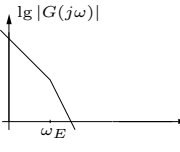
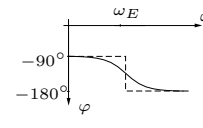
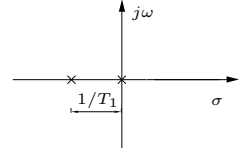
System	Zeitbereich Bildbereich	Übergangsfunktion	Ortskurve	Bode-Diagramm (Amplitudengang) (Phasengang)		s-Ebene × Pol ○ Nullstelle
P	$y(t) = K u(t)$ $G(s) = K$					
PT_1	$T_1 \dot{y}(t) + y(t) = K u(t)$ $G(s) = K \frac{1}{1 + T_1 s}$					
PT_2	$\frac{1}{\omega_0^2} \ddot{y}(t) + \frac{2D}{\omega_0} \dot{y}(t) + y(t) = K u(t)$ $G(s) = K \frac{1}{\frac{1}{\omega_0^2} s^2 + \frac{2D}{\omega_0} s + 1}$ $D < 1$: konjugiert komplexe Wurzeln der char. Gleichung $\lambda_{1,2} = -\omega_0(D \pm j\sqrt{1-D^2})$ $D \geq 1$: reelle Wurzeln der char. Gleichung $\lambda_{1,2} = -\omega_0(D \pm \sqrt{D^2-1}) = -1/T_{1,2}$					
I	$y(t) = \frac{1}{T_I} \int u dt$ $G(s) = \frac{1}{T_I s}$					
IT_1	$T_1 \dot{y}(t) + y(t) = \frac{1}{T_I} \int u(t) dt$ $G(s) = \frac{1}{T_I s(1 + T_1 s)}$					

Tabelle 2.3: Fortsetzung

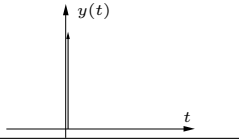
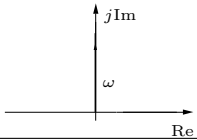
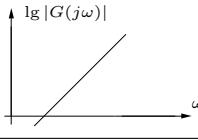
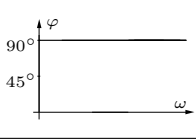
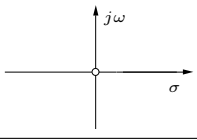
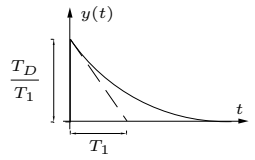
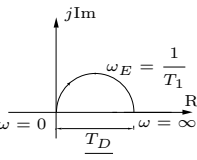
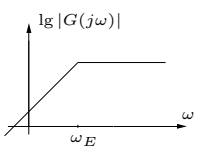
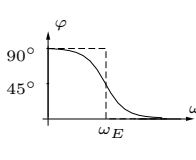
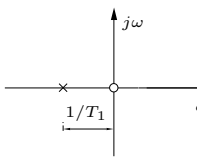
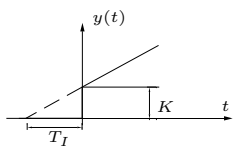
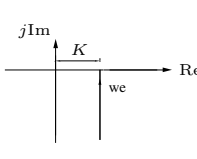
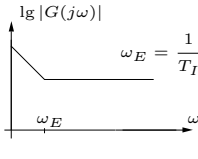
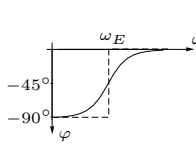
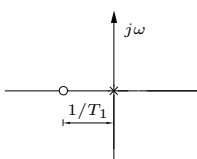
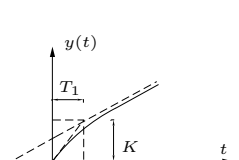
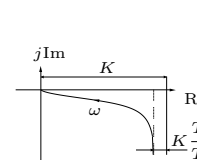
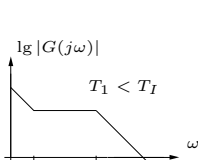
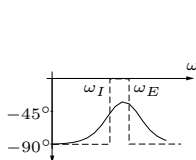
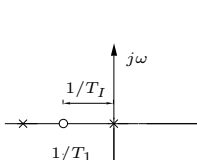
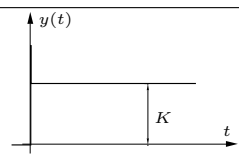
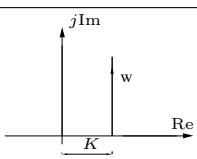
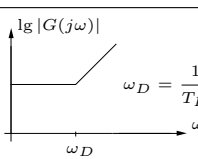
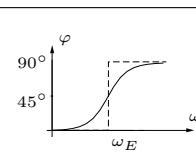
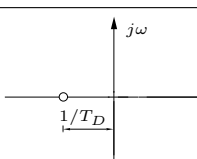
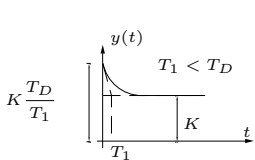
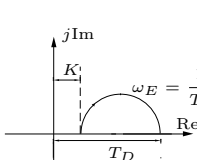
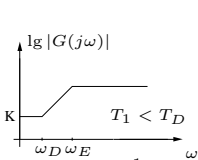
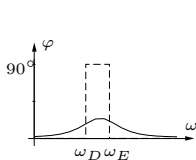
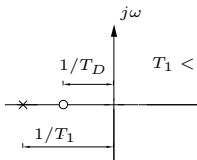
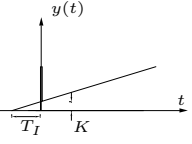
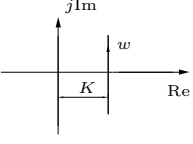
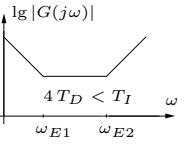
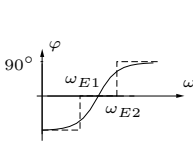
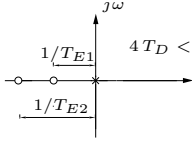
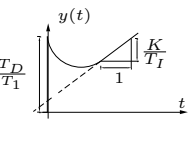
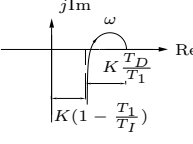
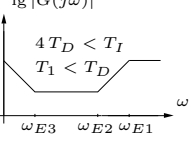
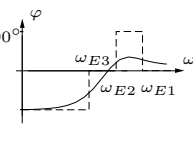
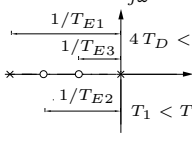
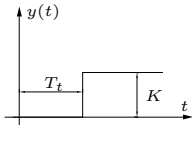
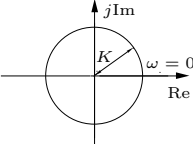
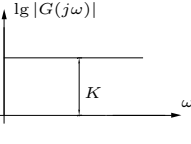
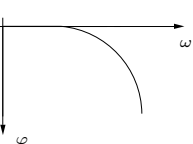
System	Zeitbereich Bildbereich	Übergangsfunktion	Ortskurve	Bode-Diagramm		s-Ebene × Pol ○ Nullstelle
				(Amplitudengang)	(Phasengang)	
D	$y(t) = T_D \frac{du}{dt}$ $G(s) = T_D s$					
DT_1	$T_1 \dot{y}(t) + y(t) = T_D \frac{du}{dt}$ $G(s) = T_D \frac{s}{1 + T_1 s}$					
PI	$y(t) = K \left[u(t) + \frac{1}{T_I} \int u(t) dt \right]$ $G(s) = K \left[1 + \frac{1}{T_I s} \right]$					
PIT_1	$T_1 \dot{y}(t) + y(t) = K \left[u(t) + \frac{1}{T_I} \int u dt \right]$ $G(s) = K \frac{1 + \frac{1}{T_I s}}{1 + T_1 s}$					
PD	$y(t) = K [u(t) + T_D \dot{u}(t)]$ $G(s) = K [1 + T_D s]$					
PDT_1	$T_1 \dot{y}(t) + y(t) = K [u(t) + T_D \dot{u}(t)]$ $G(s) = K \frac{1 + T_D s}{1 + T_1 s}$					

Tabelle 2.3: Fortsetzung

System	Zeitbereich Bildbereich	Übergangsfunktion	Ortskurve	Bode-Diagramm		s-Ebene × Pol ○ Nullstelle
				(Amplitudengang)	(Phasengang)	
<i>PID</i>	$y(t) = K \left[u(t) + \frac{1}{T_I} \int u dt + T_D \frac{du}{dt} \right]$ $G(s) = K \left[1 + T_D s + \frac{1}{T_I s} \right]$					
<i>PIDT₁</i>	$T_1 \dot{y}(t) + y(t) = K \left[u(t) + \frac{1}{T_I} \int u dt + T_D \frac{du}{dt} \right]$ $G(s) = K \frac{1 + T_D s + \frac{1}{T_I s}}{1 + T_1 s}$					
<i>T_t</i>	$y(t) = K u(t - T_t)$ $G(s) = K e^{-s T_t}$					<p>Pole bei $-\infty$ Nullstellen bei $+\infty$</p>

2.6 Beispiele für die Konstruktion des Bode-Diagramms

$$G(j\omega) = 40 \frac{1 + 0,5j\omega}{(1 + 2,5j\omega)(1 + j\omega)(1 + 0,2j\omega)}$$

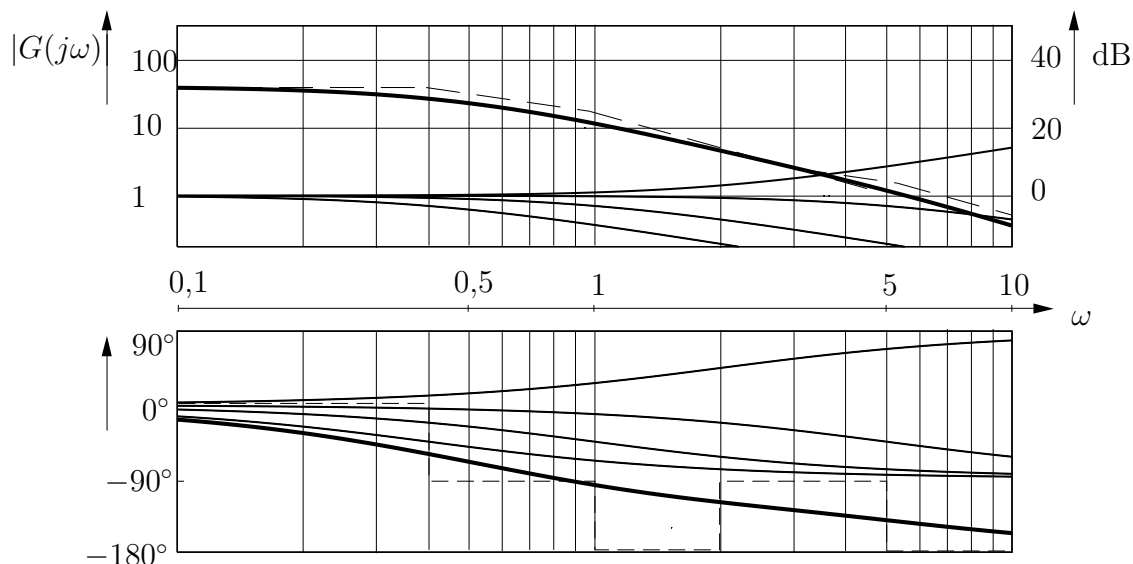


Bild 2.28: Beispiel 1 eines Bode-Diagrammes

$$G(j\omega) = \frac{3}{j\omega} (1 + 5j\omega) \frac{1}{(1 + 1,25j\omega)} \frac{1}{(1 + 0,33j\omega)}$$

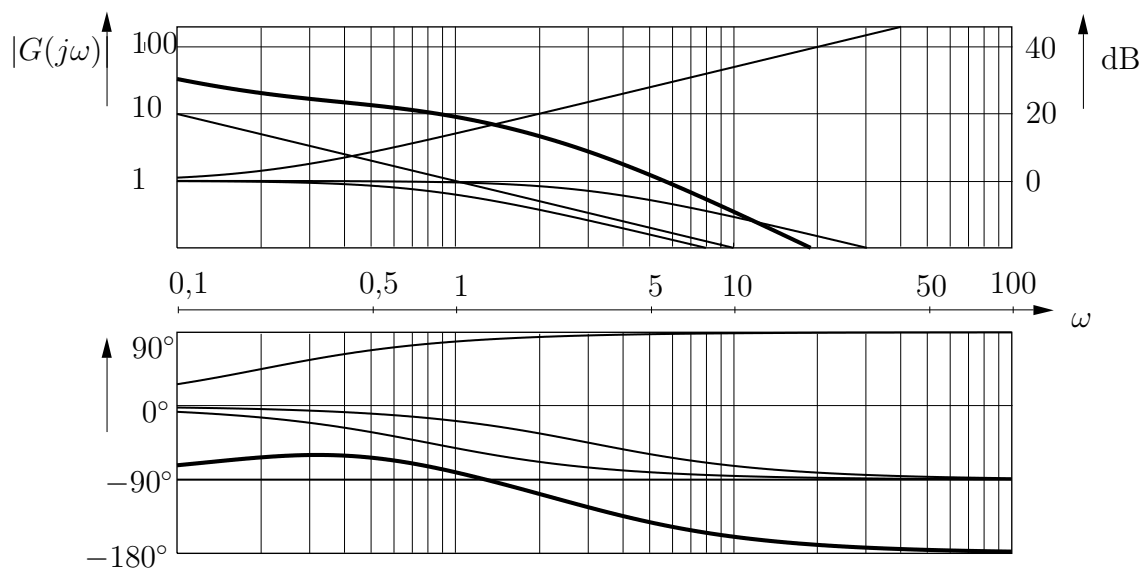


Bild 2.29: Beispiel 2 eines Bode-Diagrammes

$$G(j\omega) = 10 \frac{1 + 2,5j\omega}{j\omega \left(1 + 0,06j\omega - \frac{\omega^2}{9}\right)}$$

$$D = 0,1$$

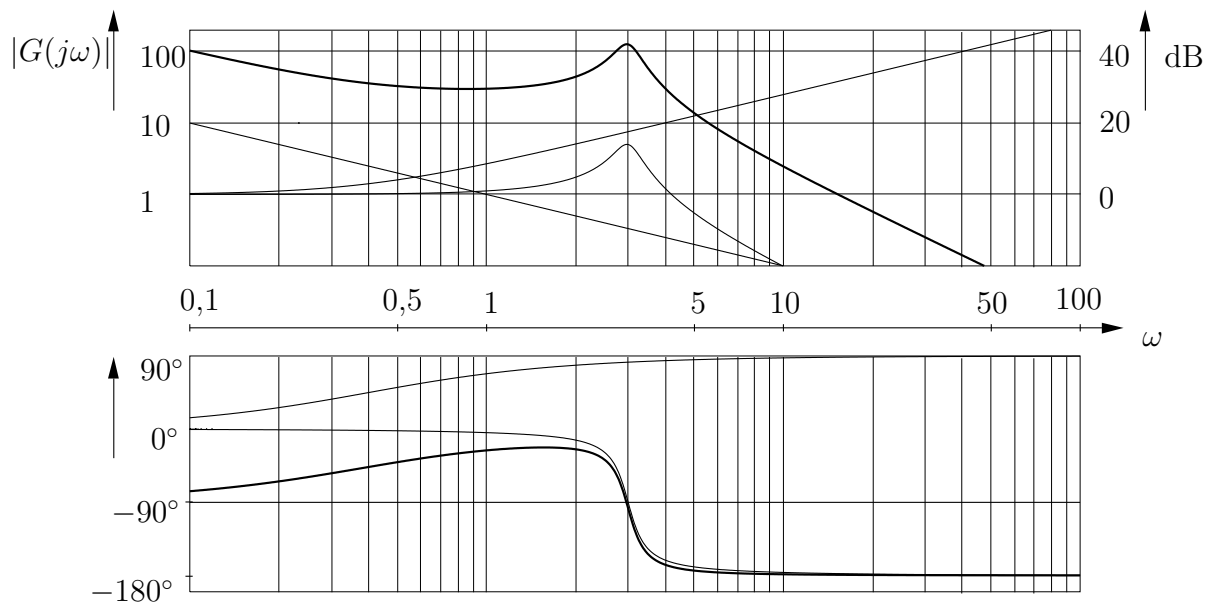


Bild 2.30: Beispiel 3 eines Bode-Diagrammes

2.7 Das Nyquistverfahren (Frequenzgangverfahren)

Das nun zu besprechende Verfahren — das 1932 von Nyquist für Stabilitätsprobleme rückgekoppelter Verstärker entwickelt wurde — basiert auf dem Frequenzgang $G_0(j\omega)$ eines dynamischen Systems. Mit diesem Verfahren kann die Stabilität des geschlossenen Regelkreises geprüft werden, ohne die Pole des Regelkreises zu berechnen.

Die Frequenzgangfunktion $G_0(j\omega)$ kann auf mehreren Wegen gewonnen werden:

- a) experimentell durch Messungen an einem System und insbesondere Auswertung mittels Bode-Diagramm (Bestimmung der Eckfrequenzen aus den Schnittpunkten der Asymptoten).
- b) analytisch, z. B. aus Übertragungsfunktionen

$$G(s) = \frac{\sum_{i=0}^m b_i \cdot s^i}{\sum_{l=0}^n a_l \cdot s^l} \cdot e^{-sT_t} \quad , \quad (2.50)$$

$$G(j\omega) = G(s)|_{s=j\omega} = \frac{\sum_{i=0}^m b_i \cdot (j\omega)^i}{\sum_{l=0}^n a_l \cdot (j\omega)^l} \cdot e^{-j\omega T_t} \quad . \quad (2.51)$$

Das Nyquistverfahren wird in der regelungstechnischen Praxis sehr häufig angewendet, da anhand der Ortskurve $G_0(j\omega)$ des offenen Kreises eine Aussage über die Stabilität des geschlossenen Regelkreises getroffen werden kann. Für die praktische Anwendung genügt es, daß der Frequenzgang $G_0(j\omega)$ graphisch vorliegt. Aus folgenden Gründen gehört das Nyquistverfahren zu den Standardverfahren der Regelungstechnik:

- Die Ortskurve läßt sich aus einer Reihenschaltung der einzelnen Regelkreisglieder ermitteln, wenn deren Kennwerte bekannt sind.
- Experimentell durch Messungen ermittelte Frequenzgänge der Regelkreisglieder können direkt berücksichtigt werden.
- Das Verfahren ermöglicht nicht nur die Untersuchung von Systemen mit konzentrierten Parametern, sondern auch von solchen mit verteilten Parametern (z.B. Totzeit-Systeme).
- Mit Hilfe der Frequenzkennlinien-Darstellung von $G_0(j\omega)$ läßt sich nicht nur die Stabilitätsanalyse, sondern auch der Entwurf (Synthese) stabiler Regelsysteme durchführen.

Grundlage des Nyquist-Kriteriums bildet die Abbildung der komplexen s -Ebene in die komplexe $F(s)$ -Ebene. Hierzu werden zunächst rationale Funktionen $F(s) = Z(s)/N(s)$ in der Pol-Nullstellendarstellung

$$F(s) = k_0 \frac{\prod_{i=1}^m (s - n_i)}{\prod_{i=1}^n (s - p_i)} \quad (2.52)$$

betrachtet. Zerlegt man $F(s)$ in Betrag und Phase

$$F(s) = k_0 |G(s)| e^{j\varphi(s)}, \quad (2.53)$$

so ist das Argument $\varphi(s)$ von $F(s)$ der Winkel, den dieser Zeiger mit der positiven reellen Achse einschließt. Dieser Winkel setzt sich aus den Phasen der einzelnen Linearfaktoren in Zähler und Nenner zusammen:

$$\varphi(s) = \sum_{i=1}^m \varphi_{n_i}(s) - \sum_{i=1}^n \varphi_{p_i}(s). \quad (2.54)$$

Desweiteren sei eine beliebige geschlossene Kurve Γ_s (vgl. Bild 2.31) in der s -Ebene gegeben, die einige Pole und Nullstellen von $F(s)$ umschließt. Unter der Voraussetzung, daß keine Pole und Nullstellen von $F(s)$ exakt auf der Kurve Γ_s liegen, bildet eine rationale Funktion $F(s)$ eine geschlossene Kurve Γ_s in der s -Ebene in eine geschlossene Kurve Γ_F in der $F(s)$ -Ebene ab. Dabei hängt die Form der Kurve Γ_F sowohl von der gewählten Kurve Γ_s als auch von der rationalen Funktion $F(s)$ und der Anzahl der umschlossenen Pole und Nullstellen ab:

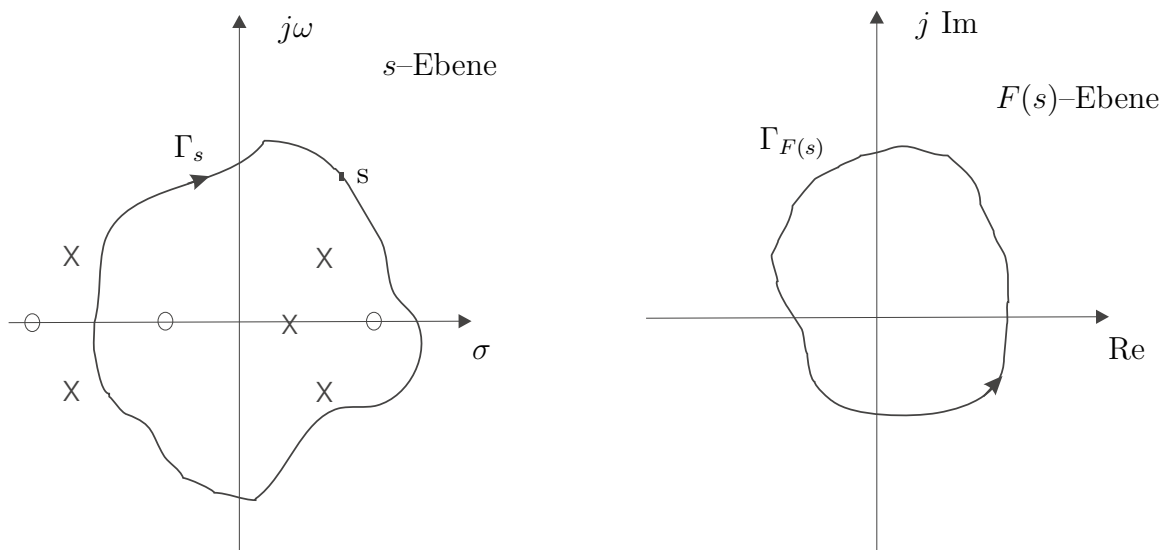


Bild 2.31: Abbildung einer geschlossenen Kurve Γ_s in der s -Ebene durch $F(s)$ in der $F(s)$ -Ebene.

Satz 2.1 : Satz zur konformen Abbildung

Gegeben sei eine rationale Funktion $F(s)$, die in allen Punkten der geschlossenen Kurve Γ_s analytisch ist (alle Ableitungen existieren). Wird die Kurve Γ_s nun im (mathematisch negativen Sinn) Uhrzeigersinn durchlaufen und liegen innerhalb der Kurve P -Pol- und N -Nullstellen von $F(s)$ mit $(N - P) > 0$, dann wird die Kurve Γ_s durch $F(s)$ in eine Kurve Γ_F in der $F(s)$ -Ebene abgebildet, die den Ursprung der $F(s)$ -Ebene $(N - P)$ -mal im Uhrzeigersinn umschlingt. \square

Eine Erläuterung des Satzes 2.1 ist anhand des Bildes 2.31 möglich. Hierzu betrachtet man die Phasenänderungen der Linearfaktoren von $F(s)$ für einen Umlauf des Punktes s im Uhrzeigersinn auf der Kurve Γ_s . Man erkennt, daß ein Linearfaktor $(s - p_i)$ bzw. $(s - n_k)$ nur dann einen Beitrag von 360° zur Abbildung in der $F(s)$ -Ebene liefert, wenn er innerhalb der geschlossenen Kurve Γ_s liegt. Aus Gl.(2.54) folgt dann sofort der in Satz 2.1 angegebene Zusammenhang.

Das Nyquistkriterium ist eine direkte Anwendung des Satzes 2.1 auf einen geschlossenen Regelkreis. Da es um die Stabilität von linearen Regelkreisen geht, kann die Kurve Γ_s nicht mehr beliebig gewählt werden.

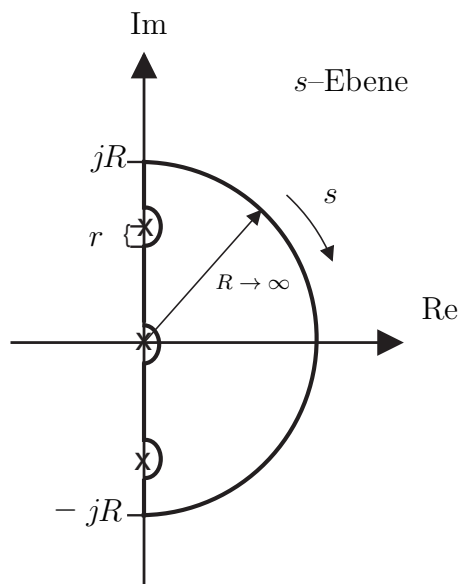


Bild 2.32: Nyquistkurve Γ_N .

Die nach Nyquist bezeichnete Nyquistkurve Γ_N wird jetzt so gelegt, daß sie die rechte s -Halbebene umschließt. Wenn die rationale Funktion $F(s)$ Pole auf der imaginären Achse besitzt, hat die Abbildung $F(s)$ an dieser Stelle einen unendlich großen Betrag. Deshalb wird die Nyquistkurve Γ_N an dieser Stelle durch einen Halbkreis mit dem Radius $r \rightarrow 0$ um die Polstellen herumgeführt (siehe Bild 2.32). Damit bleibt die Funktion $F(s)$ auf der Nyquistkurve analytisch und eine direkte Anwendung des Satzes 2.1 ist auch in diesen Fällen möglich.

Ausgangspunkt der Stabilitätsuntersuchung mit Hilfe der Nyquistkurve ist die Rückführdifferenzfunktion

$$D(s) = 1 + G_0(s) = \frac{Z_0(s) + N_0(s)}{N_0(s)} . \quad (2.55)$$

Offensichtlich ist das Zählerpolynom von $D(s)$ identisch mit dem Nennerpolynom der Stör-

$$G_z(s) = \frac{1}{1 + G_0(s)} = \frac{N_0(s)}{Z_0(s) + N_0(s)}$$

bzw. der Führungsübertragungsfunktion

$$G_w(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{Z_0(s)}{Z_0(s) + N_0(s)}$$

des geschlossenen Kreises und das Nennerpolynom von $D(s)$ gerade das Nennerpolynom des offenen Kreises.

Die Abbildungsfunktion $D(s)$ habe nun $P_{R,G}$ Nullstellen und $P_{R,O}$ Pole in der rechten s -Halbebene. Wird nun die Nyquistkurve Γ_N durch $D(s)$ abgebildet, dann liefert der Satz 2.1, daß die Kurve Γ_D den Ursprung der komplexen Ebene $(P_{R,G} - P_{R,O})$ -mal im Uhrzeigersinn umschließt.

Die Anwendung dieses Stabilitätskriteriums kann vereinfacht werden, wenn man den Zusammenhang

$$G_0(s) = D(s) - 1 \quad (2.56)$$

nutzt. Der Übergang von $D(s)$ zu $G_0(s)$ bedeutet lediglich eine Verschiebung der Ortskurve von $D(j\omega)$ um den Wert 1 nach links. Der *kritische Punkt*, dessen Umschließung durch die Kurve untersucht wird, verschiebt sich daher vom Koordinatenursprung $(0, j0)$ zum Punkt $(-1, j0)$ in der $G_0(j\omega)$ -Ebene. Da ein geschlossener Regelkreis dann und nur dann stabil ist, wenn die Übertragungsfunktion keine Pole in der rechten s -Halbebene hat, also $(P_{R,G} = 0)$ gilt, erhält man das von Nyquist angegebene Stabilitätskriterium:

Satz 2.2 : Nyquistkriterium

Für die Stabilität des geschlossenen Regelkreises (Gegenkopplung) ist notwendig und hinreichend, daß die Ortskurve $G_0(j\omega)$ bei Änderung der Frequenz ω von $-\infty$ bis $+\infty$ den kritischen Punkt $(-1; j0)$ im mathematisch positiven Sinn (Gegenuhrzeigersinn) so oft umfährt, wie instabile Pole in $G_0(s)$ enthalten sind. \square

Ist der offene Regelkreis stabil, hat also $G_0(s)$ keine Pole in der rechten s -Halbebene, dann ist auch $P_{R,O} = 0$ und man erhält:

Satz 2.3 : Vereinfachtes Nyquistkriterium

Für die Stabilität des geschlossenen Regelkreises ist bei **stabilem** $G_0(s)$ notwendig und hinreichend, daß die Ortskurve $G_0(j\omega)$ bei Änderung der Frequenz ω von $-\infty$ bis $+\infty$ den kritischen Punkt $(-1; j0)$ weder umschließt noch durchdringt. \square

Eine andere Fassung des vereinfachten Nyquistkriteriums, die sogar auch noch angewandt werden kann, wenn $G_0(s)$ Pole bei $s = 0$ besitzt, ist die sogenannte *Linke-Hand-Regel*:

Linke-Hand-Regel. Wenn die Übertragungsfunktion $G_0(s)$ des offenen Regelkreises nur Pole in der linken s -Halbebene hat, ausgenommen einem 1- oder 2-fachen Pol bei $s = 0$, ist der geschlossene Kreis genau dann asymptotisch stabil, wenn bei Änderung der Frequenz ω von $+0$ bis $+\infty$ der kritische Punkt $(-1; j0)$ immer *links* von der Ortskurve $G_0(j\omega)$ liegt (Bild 2.33).

Sowohl bei der experimentellen Ermittlung eines Frequenzgangs als auch bei der analytischen Berechnung ist eine Darstellung des Frequenzgangs in der logarithmischen Darstellung des Bode-Diagrammes besonders übersichtlich.

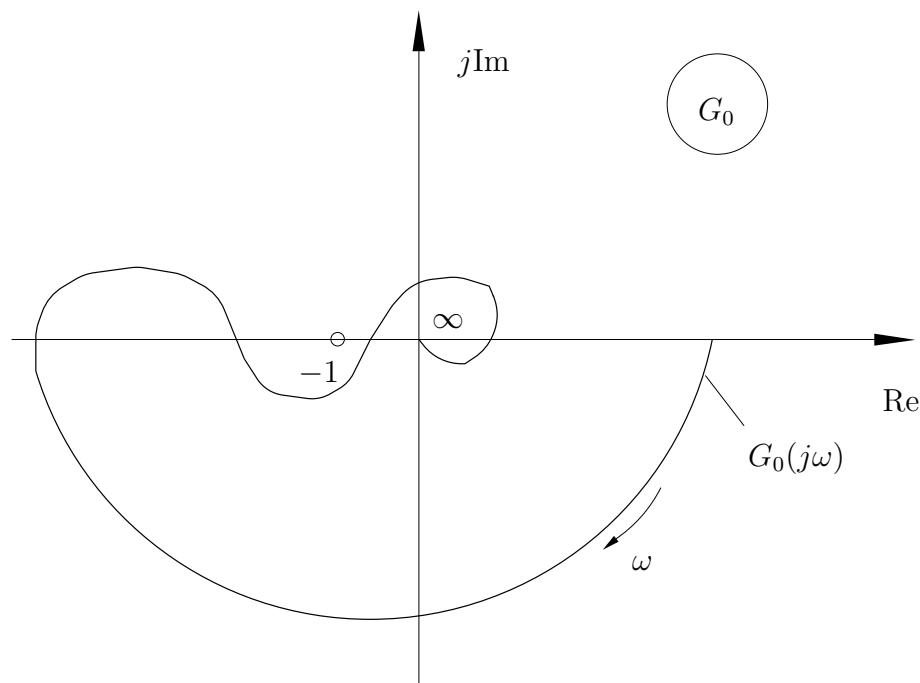


Bild 2.33: Beispiel zum vereinfachten Nyquistkriterium

Da beim Bode-Diagramm eine getrennte Darstellung des Amplituden- und des Phasengangs erfolgt, muß bei der Anwendung des Nyquist-Stabilitätsnachweises für Frequenzgänge in Bode-Diagramm-Darstellung das Nyquistkriterium umformuliert werden:

Satz 2.4 : Vereinfachtes Nyquistkriterium im Bode-Diagramm

Wenn die Übertragungsfunktion $G_0(s)$ des offenen Regelkreises nur Pole in der linken s -Halbebene hat, ausgenommen einem 1- oder 2-fachen Pol bei $s = 0$, ist der geschlossene Kreis genau dann asymptotisch stabil, wenn $G_0(j\omega)$ für die Durchtrittsfrequenz ω_D einen Phasenwinkel $\varphi_0(\omega_D) > -180^\circ$ hat (Bild 2.34). □

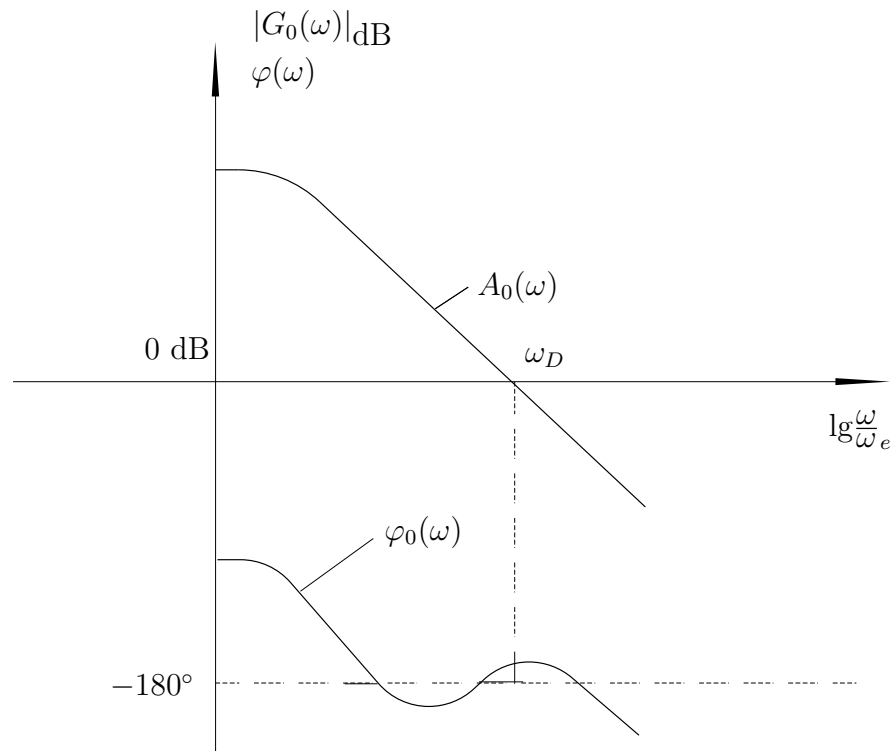


Bild 2.34: Vereinfachtes Nyquistkriterium im Bode-Diagramm

2.8 Stabilitätsreserve

Bei der Auslegung von Regelkreisen muß immer eine gewisse Stabilitätsreserve vorgesehen werden, um Parameterunsicherheiten in den Systemmodellen zu berücksichtigen. Diese Stabilitätsreserve kann auch mittels des Nyquistverfahrens an Hand der Frequenzgangskennlinien berücksichtigt werden. Je größer der Abstand der Ortskurve vom kritischen Punkt ist, desto weiter ist der geschlossene Regelkreis vom Stabilitätsrand entfernt. Als Maß hierfür benutzt man die Begriffe Phasenrand und Amplitudenrand, die in Bild 2.35 erklärt sind.

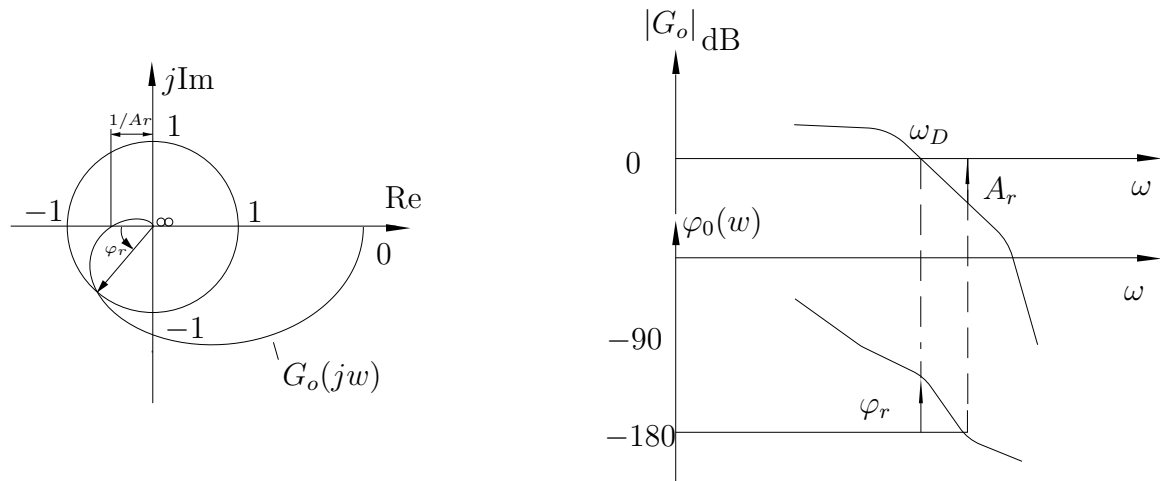


Bild 2.35: Definition der Stabilitätsreserve

a) in der Ortskurvendarstellung b) im Bodediagramm

Definition 2.3 : Amplituden- und Phasenrand

1. Der Phasenrand

$$\varphi_R = 180^\circ + \varphi(\omega_D) \tag{2.57}$$

ist der Abstand der Phasenkennlinie von der -180° -Geraden bei der Durchtrittsfrequenz ω_D , d. h. beim Durchgang der Amplitudenkennlinie durch die 0-dB-Linie ($|G_0| = 1$).

2. Als Amplitudenrand

$$A_R = \frac{1}{|G_0|} \Big|_{\varphi_0=-180^\circ} ; A_{R_{dB}} = -|G_0|_{dB} \Big|_{\varphi_0=-180^\circ} \tag{2.58}$$

wird der Abstand der Amplitudenkennlinie von der 0-dB-Linie beim Winkel $\varphi_0 = -180^\circ$ bezeichnet. \square

Für eine gut gedämpfte Regelung sollten etwa folgende Werte eingehalten werden:

$$A_{R_{dB}} = \begin{cases} 12 \text{ dB} & \text{bis } 20 \text{ dB} & \text{bei Führungsverhalten} \\ 3,5 \text{ dB} & \text{bis } 9,5 \text{ dB} & \text{bei Störverhalten} \end{cases}$$

$$\varphi_R = \begin{cases} 40^\circ & \text{bis } 60^\circ & \text{bei Führungsverhalten} \\ 20^\circ & \text{bis } 50^\circ & \text{bei Störverhalten} \end{cases}$$

Die Durchtrittsfrequenz ω_D stellt ein Maß für die dynamische Güte des Regelkreises dar. Je größer ω_D , desto größer ist die Bandbreite des geschlossenen Regelkreises, und desto schneller die Reaktion auf Sollwertänderungen oder Störungen. Als Bandbreite ist dabei jene Frequenz ω_B zu betrachten, bei der der Betrag des Frequenzganges des geschlossenen Regelkreises auf den Wert -3 dB abgefallen ist.

2.9 Das Wurzelortskurvenverfahren (WOK–Verfahren)

Die Stabilität eines linearen Systems wird durch die Lage der Pole der Übertragungsfunktion in der komplexen s -Ebene vollständig bestimmt. Sind die Pol- und Nullstellen des offenen Systems bekannt, so ist die Frage von Interesse, wie sich die Lage der Pole ändert, wenn der Kreis geschlossen wird. Für den geschlossenen Regelkreis berechnen sich die Pole als Wurzeln der charakteristischen Gleichung

$$1 + G_0(s) = 1 + G_S(s)G_R(s) = 0. \quad (2.59)$$

Sie verändern sich mit den Reglerparametern, die in die Übertragungsfunktion $G_R(s)$ eingehen. Variiert man beispielsweise einen Parameter des Reglers (z.B. die Reglerverstärkung), so verändert sich die Lage der Wurzeln der charakteristischen Gleichung (2.59) des geschlossenen Regelkreises in der s -Ebene. Die Wurzeln beschreiben somit Bahnen in der s -Ebene, die man die *Wurzelortskurve* (WOK) des geschlossenen Regelkreises nennt. Die Kenntnis der Wurzelortskurve ermöglicht nicht nur eine Aussage über die Stabilität des geschlossenen Regelkreises, sondern auch eine Beurteilung der Stabilitätsgüte, z.B. durch den Abstand der Pole von der Imaginärachse.

Definition 2.4 Wurzelortskurve

Der Wurzelort ist der geometrische Ort der Wurzeln der charakteristischen Gleichung (2.59) in der komplexen Ebene. Die Wurzelortskurve (WOK) stellt die Abhängigkeit der Wurzelorte von einem Parameter (vielfach der Verstärkung K_0) des offenen Regelkreises dar. \square

Die Begriffe Wurzelort und Wurzelortskurve werden durch das folgende Beispiel verdeutlicht.

Beispiel: Wurzelortskurve eines Folgeregelkreises

Es wird der in Bild 2.36 dargestellte Folgeregelkreis betrachtet.

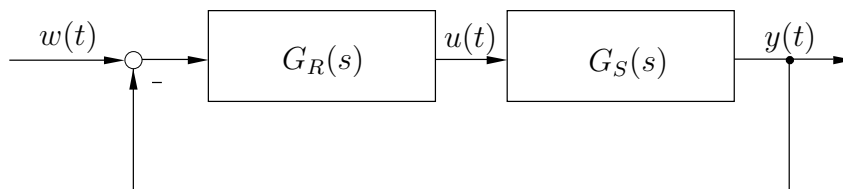


Bild 2.36: Blockdiagramm eines Folgeregelkreises

Die Regelstrecke genüge der Differentialgleichung

$$\Theta \ddot{y}(t) + B \dot{y}(t) = u(t) \quad ,$$

womit für $G_S(s)$ folgt:

$$G_S(s) = \frac{Y(s)}{U(s)} = \frac{1/B}{s} \frac{1}{\left(1 + \frac{\Theta}{B}s\right)}$$

Es werde nun ein P-Regler mit der Verstärkung K_R eingesetzt. Damit folgt mit den Abkürzungen $K_S = 1/B$ und $T = \Theta/B$ für $G_0(s)$:

$$G_0(s) = K_R \cdot \frac{K_S}{s} \frac{1}{1 + sT} = K_0 \frac{1}{s(1 + sT)}$$

und $G_w(s)$ zu:

$$G_w(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{K_0}{K_0 + s(1 + sT)} .$$

Die charakteristische Gleichung lautet:

$$C(s) = N_w(s) = Ts^2 + s + K_0 = 0$$

und hat die Wurzeln

$$s_{1/2} = -\frac{1}{2T} \pm \sqrt{\frac{1}{4T^2} - \frac{K_0}{T}}$$

- i) für $K_0 = 0 : s_1 = 0 ; s_2 = -\frac{1}{T}$
- ii) für $K_0 = 1/4T : s_{1/2} = -\frac{1}{2T}$
- iii) für $K_0 > 1/4T : s_{1/2} = -\frac{1}{2T} \pm j\frac{1}{2T}\sqrt{4K_0T - 1} .$

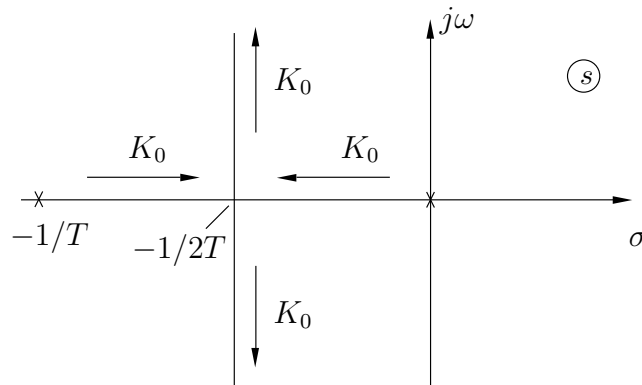


Bild 2.37: Wurzelortskurve des Beispiels

Bild 2.37 zeigt die Polbewegung des geschlossenen Regelkreises für das Beispiel in Abhängigkeit der Kreisverstärkung $K_0 = K_R \cdot K_S$. Die Lage der Pole des geschlossenen Kreises für ein gegebenes K_0 wird *Wurzelort* genannt. Für $K_0 = 0 \dots +\infty$ entstehen zwei Kurven, die als Äste der *Wurzelortskurve* bezeichnet werden.

Aus der Wurzelortskurve als Ganzes kann man erkennen, wie sich die Eigenschaften des geschlossenen Kreises bei Erhöhung der Reglerverstärkung verändern. Das System des Beispiels ist für $K_0 = 0$ aufgrund des Pols im Ursprung grenzstabil. Für $K_0 > 0$ ist der geschlossene Kreis immer stabil, doch nimmt mit zunehmender Verstärkung $K_0 > \frac{1}{2T}$ die Eigenfrequenz ω_0 zu und die Dämpfung D ab. \square

Bei komplizierteren Systemen ist die Polbewegung in Abhängigkeit eines Parameters nicht mehr ohne weiteres berechenbar. Hier hilft das von Evans (1948) angegebene Wurzelortskurvenverfahren, das mittels einfach anzuwendender Regeln WOKn zu konstruieren gestattet.

Zur Bestimmung der WOKn wird von einer Übertragungsfunktion $G_0(s)$ des offenen Kreises in folgender Darstellungsform (Pol–Nullstellen–Form) ausgegangen:

$$G_0(s) = k_0 \hat{G}_0(s) = k_0 \frac{\prod_{i=1}^m (s - n_i)}{\prod_{i=1}^n (s - p_i)} \quad . \quad (2.60)$$

Die charakteristische Gleichung des geschlossenen Regelkreises ergibt sich dann zu

$$C(s) = 1 + k_0 \hat{G}_0(s) = 0 \quad , \quad (2.61)$$

die zunächst umgestellt

$$\hat{G}_0(s) = -\frac{1}{k_0} \quad (2.62)$$

und in *Polarkoordinaten* geschrieben wird:

$$|\hat{G}_0(s)| e^{j\varphi_0(s)} = -\frac{1}{k_0} \quad . \quad (2.63)$$

Diese Beziehung läßt sich für $k_0 \geq 0$ in zwei Gleichungen zerlegen:

$$\text{Amplitudenbedingung:} \quad |\hat{G}_0(s)| = \frac{\prod_{i=1}^m |s - n_i|}{\prod_{i=1}^n |s - p_i|} = \frac{1}{k_0} \quad (2.64)$$

$$\text{Phasenbedingung:} \quad \varphi_0(s) = \sum_{i=1}^m \varphi_{n_i} - \sum_{i=1}^n \varphi_{p_i} = \pm(2k + 1)\pi \quad (2.65)$$

mit $k = 0, 1, 2, \dots$

Die Winkelbedingung ist offensichtlich unabhängig von k_0 . Alle Punkte der komplexen s -Ebene, die die Phasenbedingung erfüllen, stellen also den geometrischen Ort aller möglichen Pole des geschlossenen Regelkreises dar, die durch die Variation von k_0 entstehen können. Die Parametrierung dieser WOK, d.h. die Zuordnung zwischen den Kurvenpunkten und den Werten von k_0 , erhält man durch Auswertung der Amplitudenbedingung (2.64). Zur Berechnung von k_0 wird diese Bedingung in

$$k_0 = \frac{\prod_{i=1}^n |s - p_i|}{\prod_{i=1}^m |s - n_i|} \quad (2.66)$$

umgeformt. Offensichtlich kann k_0 aus den Abständen des betrachteten Punktes von allen Polen und Nullstellen bestimmt werden. Für Systeme ohne Nullstellen ($m = 0$) wird der Nenner gleich eins gesetzt. Bei der Anwendung dieser Methode (vgl. Bild 2.38) müssen reelle und imaginäre Achse mit gleichem Maßstab gezeichnet sein.

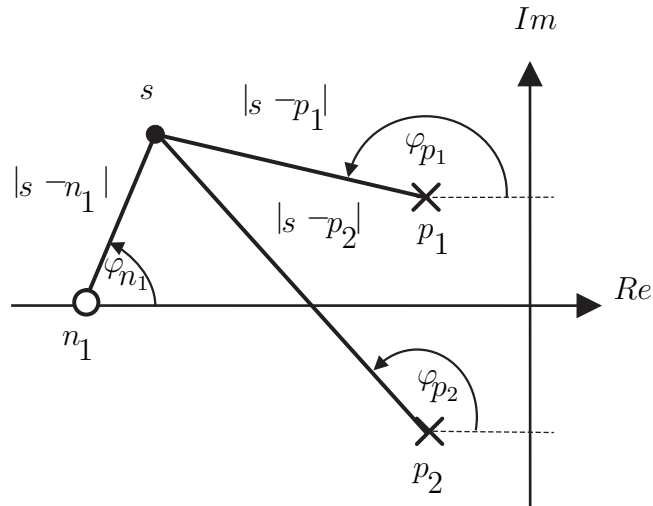


Bild 2.38: Analyse der Wurzelortskurve

Für die Konstruktion der WOK ist dieser Weg allerdings aufwendig. Es werden deshalb im folgenden einfache, von Evans abgeleitete Regeln angegeben, mit denen der qualitative Verlauf der WOK auch bei Systemen höherer Ordnung leicht zu bestimmen ist.

Satz 2.5

Die Wurzelortskurve eines Systems ist der geometrische Ort aller Punkte der s -Ebene, für die die Summe der Pol- und Nullstellenwinkel der Übertragungsfunktion des offenen Systems $\pm 180^\circ(2k + 1)$ beträgt.

$$\sum_{i=1}^m \varphi_{n_i} - \sum_{i=1}^n \varphi_{p_i} = \pm 180^\circ(2k + 1) \quad (2.67)$$

Alle Pole der Übertragungsfunktionen des geschlossenen Regelkreises liegen auf den WOKn. □

Start- und Endpunkte der WOK

Die n Äste der WOK beginnen für $k_0 = 0$ in den Polen p_i des offenen Systems $G_0(s)$. m Äste enden für $k_0 \rightarrow \infty$ in den endlichen Nullstellen n_i von $G_0(s)$, während die verbleibenden $(n - m)$ Äste gegen Unendlich laufen (man kann auch sagen, daß diese Äste in der Nullstelle im Unendlichen der Vielfachheit oder Ordnung $(n - m)$ enden). \square

Winkel der Asymptoten

Die *Asymptoten* der $(n - m)$ nach Unendlich strebenden Äste der WOK sind Geraden, die sich für $(n - m) \geq 2$ alle im *Wurzelschwerpunkt*

$$\sigma_W = \frac{\sum_{i=1}^n p_i - \sum_{i=1}^m n_i}{n - m}$$

auf der reellen Achse schneiden. Sie haben zur reellen Achse die Neigungswinkel

$$\varphi_k = \frac{(2k - 1)180^\circ}{n - m}, \quad k = 1, 2, \dots, n - m.$$

\square

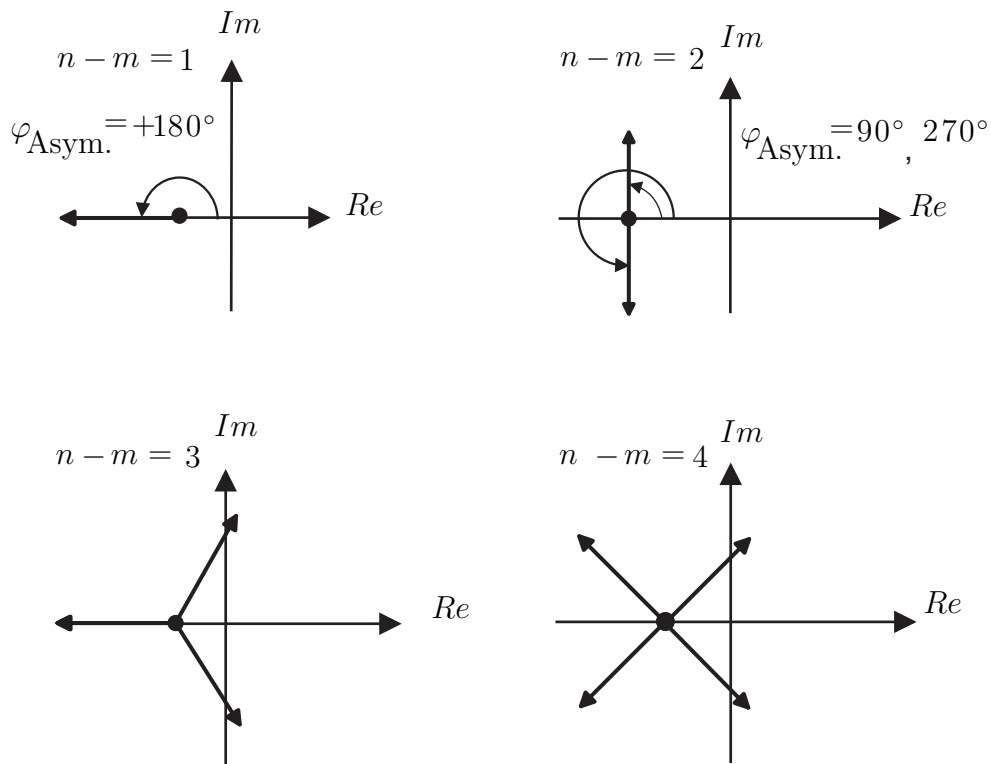


Bild 2.39: Asymptoten der WOK für unterschiedliche Differenzgrade $(n - m)$ von $G_0(s)$

Wurzelorte auf der reellen Achse

Zur WOK gehören genau diejenigen reellen Werte s , für die die Anzahl der von diesem Punkt s aus gesehen rechts liegenden Pole und Nullstellen des offenen Kreises ungerade ist. □

Symmetrie der WOK

Die WOKn verlaufen symmetrisch zur reellen Achse. □

Mit Hilfe dieser Regeln lassen sich die WOKn in vielen Fällen bereits grob skizzieren. Für eine exakte Konstruktion von WOK sind weitere Regeln notwendig, die in Tabelle 3.2 zusammengestellt sind, wobei $G_0(s)$ in dieser Form vorliegen muß (Bild 2.40):

$$G_0(s) = k_0 \cdot \frac{\prod_{i=1}^m (s - n_i)}{\prod_{i=1}^n (s - p_i)} \quad \text{mit} \quad k_0 \geq 0$$

$m =$ Anzahl der im Endlichen liegenden Nullstellen

$n =$ Anzahl der im Endlichen liegenden Pole.

} (2.68)

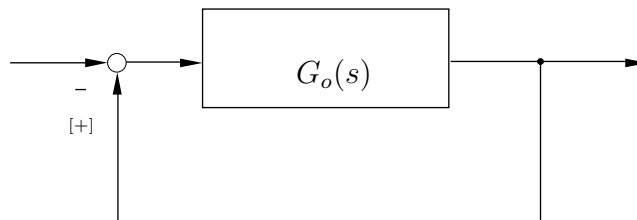


Bild 2.40: Regelkreis mit Einheitsrückführung

1	Anfangspunkte	Die WOK beginnt für $k_0 = 0$ in den Polen von G_0 .
2	Endpunkte	Die WOK endet für $k_0 \rightarrow \infty$ in den Nullstellen von G_0 .
3	WOK auf der reellen Achse	Jeder Ort auf der reellen Achse, auf dessen rechter Seite die Summe von Polen und Nullstellen ungerade [gerade] ² ist, ist ein Wurzelort.
4	Anzahl der separaten Äste	Die WOK besteht aus n Ästen.
5	Anzahl der Äste im Unendlichen	$n - m$ Äste enden im Unendlichen, d. h. es existieren auch $n - m$ Asymptoten.
6	Symmetrie der WOK	Die WOK ist symmetrisch zur reellen Achse.

Tabelle 2.4: Regeln zum WOK–Verfahren

² Die Angaben in eckigen Klammern gelten für positive Rückführung.

7	Winkel der Asymptoten	Der Winkel der Asymptoten zur reellen Achse folgt aus $\varphi_k = \frac{(2k-1)\pi}{n-m} \quad \left[\varphi_k = \frac{2k\pi}{n-m} \right], \quad k = 1, 2, \dots, n-m.$
8	Schnittpunkt der Asymptoten (Wurzelschwerpunkt σ_W)	Der Schnittpunkt liegt auf der reellen Achse. $\sigma_W = \frac{\sum_{i=1}^n p_i - \sum_{i=1}^m n_i}{n-m} \quad \text{für } n-m \geq 2$
9	Austritts- und Eintrittswinkel	Der Austrittswinkel Θ der WOK aus einem Pol (oder Eintrittswinkel in eine Nullstelle) ergibt sich, indem man einen Punkt s_1 , der auf der WOK liegen soll, in unmittelbarer Nähe des Poles (bzw. der Nullstelle) annimmt und dann die Winkelbedingung ansetzt: $\sum_{i=1}^m \arg(s_1 - n_i) - \sum_{i=1}^n \arg(s_1 - p_i) = \pm(2k-1)\pi \quad [\pm 2k\pi]$
10	Verzweigungspunkte der WOK auf der reellen Achse	Allgemein gilt am Verzweigungspunkt a : $\left. \frac{dG_0(s)}{ds} \right _{s=a} = 0$ Für $a \neq p_i$ und $a \neq n_i$ gilt damit: a) reelle Pole und Nullstellen $\sum_{i=1}^n \frac{1}{a-p_i} = \sum_{i=1}^m \frac{1}{a-n_i}$ b) konjugiert komplexe Pole und Nullstellen $n_i = \alpha_i \pm j \cdot \beta_i; \quad p_i = \alpha_l \pm j \cdot \beta_l$ $\sum_{i=1}^{m/2} \frac{2(a-\alpha_i)}{(a-\alpha_i)^2 + \beta_i^2} = \sum_{l=1}^{n/2} \frac{2(a-\alpha_l)}{(a-\alpha_l)^2 + \beta_l^2}$
11	Schnittpunkt mit der imaginären Achse	Stabilitätsgrenze des geschlossenen Kreises: k_{0krit} und ω_{krit} , z. B. mit Hurwitz-Kriterium
12	k_0 -Parametrierung der WOK	Die Verstärkung für einen Punkt s_1 der WOK ergibt sich aus: $k_0 = \frac{\prod_{i=1}^n s_1 - p_i }{\prod_{i=1}^m s_1 - n_i }$ (Falls keine Nullstellen auftreten, ist der Nenner gleich 1 zu setzen.)
13	Potential-Analogie	Die WOK verläuft beim Austritt aus der reellen Achse konkav zu benachbarten Nullstellen von $G_0(s)$, bzw. konvex zu benachbarten Polen.

Tabelle 2.4: Fortsetzung

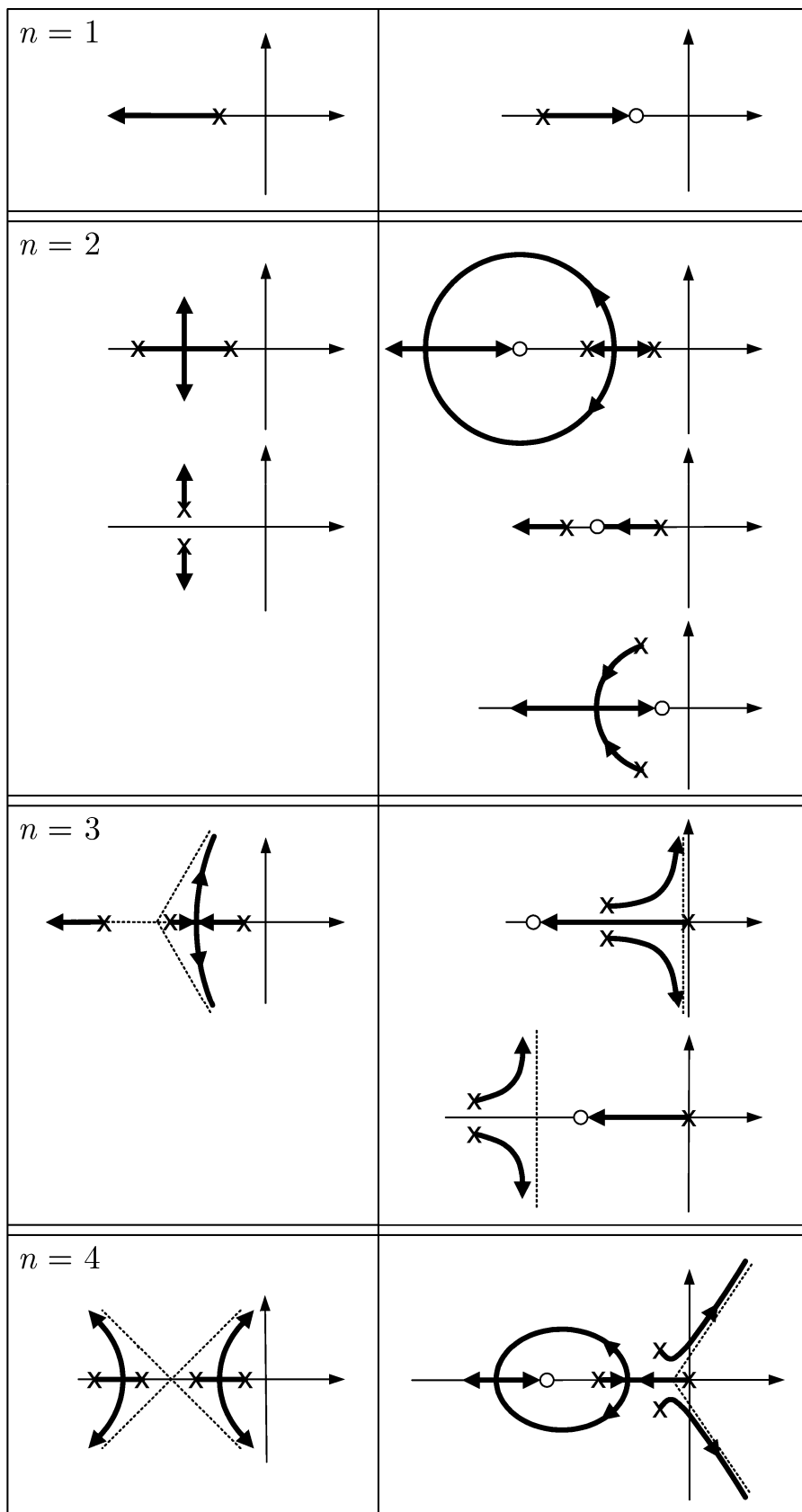


Bild 2.41: Typische Beispiele für Wurzelortskurven

Aus den Beispielen in Bild 2.41 lassen sich unmittelbar zwei wichtige Konsequenzen ableiten:

- Besitzt die Übertragungsfunktion $G_0(s)$ des offenen Regelkreises Nullstellen in der rechten s -Halbebene, so wird der geschlossene Regelkreis für große Kreisverstärkungen immer instabil. Ein Allpaßanteil in der Regelstrecke beschränkt also die mögliche Reglerverstärkung und damit die Möglichkeit, durch den Regler das dynamische Verhalten der Regelstrecke zu beeinflussen.
- Entsprechende Beschränkungen gelten auch dann, wenn der Differenzgrad $(n - m)$ der Übertragungsfunktion $G_0(s)$ größer als 2 ist.

3 Beschreibung und Analyse dynamischer Systeme im Zustandsraum

In der Systemtheorie und Regelungstechnik lassen sich zwei Modelle zur Beschreibung des dynamischen Verhaltens unterscheiden:

- a) Signalübertragungsmodell (Klemmenmodell)
- b) Zustandsmodell .

Bisher stand das Klemmenmodell zur Beschreibung des Ursache–Wirkungs–Zusammenhangs im Vordergrund; im Falle der linearen Systeme wird das Systemverhalten in bezug auf seine Signalübertragungsfunktion $G(s)$ bzw. die Frequenzgangfunktion $G(j\omega)$ jeweils eindeutig charakterisiert (Bild 3.1).

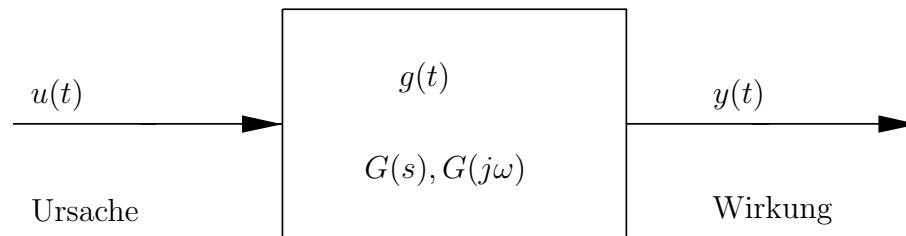


Bild 3.1: Blockschaltbild–Darstellung des linearen Systems

Der Darstellung dynamischer Systeme im Zustandsraum liegt aus mathematischer Sicht die Umwandlung einer Differentialgleichung n -ter Ordnung in ein äquivalentes System von n Differentialgleichungen erster Ordnung zu Grunde. Die Anwendung dieser Darstellung auf regelungstechnische Probleme führte seit Ende der fünfziger Jahre zu einer beträchtlichen Erweiterung der Regelungstheorie. Besonders für die Behandlung von Systemen mit mehreren Ein- und Ausgangsgrößen sowie von nichtlinearen und zeitvarianten Systemen ist die Zustandsraumdarstellung hervorragend geeignet. Diese Art der Systemdarstellung erlaubt darüber hinaus im Zeitbereich eine einfache Formulierung dynamischer Optimierungsprobleme, die in den sechziger Jahren eine Reihe von Raumfahrtprojekten erst ermöglicht hat.

Für den Einsatz der Zustandsraumdarstellung sprechen u.a. folgende Gründe:

- Ein- und Mehrgrößensysteme (Systeme mit mehr als einer Ein- bzw. Ausgangsgröße bezeichnet man als *Mehrgrößensysteme*) können formal gleich behandelt werden.
- Die Zustandsraumdarstellung ist sowohl für die theoretische Behandlung (analytische Lösungen, Optimierung) als auch für die numerische Analyse und Berechnung gut geeignet.

- Die Berechnung des Verhaltens des homogenen Systems unter Verwendung der Anfangsbedingung $\mathbf{x}(t_0)$ ist sehr einfach.
- Diese Darstellung erlaubt einen besseren Einblick in das innere Verhalten eines Systems. Hier spielen insbesondere Systemeigenschaften wie die Steuer- und Beobachtbarkeit des Systems eine besondere Rolle.

3.1 Zustandsraumdarstellung linearer zeitinvarianter Systeme

Ein weiterer wichtiger Grund für den Einsatz dieser Darstellungsform ist die grundsätzliche Bedeutung des Begriffs *Zustand eines dynamischen Systems*. Physikalisch betrachtet ist der Zustand eines dynamischen Systems durch den Energiegehalt der im System vorhandenen Energiespeicher bestimmt. Ist der Zustand zu einem beliebigen Zeitpunkt $t = t_0$ bekannt, so kann das Verhalten des Systems bei einer gegebenen Eingangsgröße $u(t)$ für alle Zeitpunkte $t \geq t_0$ einfach berechnet werden. Der Zustand eines Systems mit n Energiespeichern wird dann durch n Zustandsgrößen $x_1(t), x_2(t), \dots, x_n(t)$ beschrieben, die zu einem Zustandsvektor $\mathbf{x}(t)$ zusammengefaßt werden.

Definition 3.1 (*Zustand eines dynamischen Systems*)

Der Zustand eines Systems wird durch Variablen charakterisiert, die zu jedem Zeitpunkt die gesamte Vorgeschichte des Systems repräsentieren. Bei Kenntnis des Systemzustandes und aller auf das System einwirkenden Größen kann das zukünftige Systemverhalten vorausbestimmt werden. \square

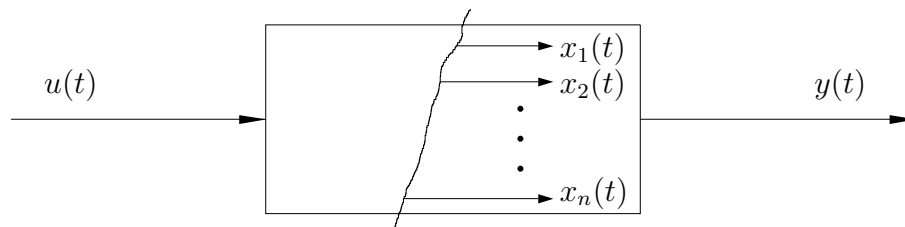


Bild 3.2: Zur Definition der Systemzustände

In Bild 3.2 ist dazu angedeutet, daß die Zustandsvariablen $x_i(t), i = 1, \dots, n$ im allgemeinen nur für das innere Systemverhalten definiert sind und normalerweise nicht (alle) von außen direkt gemessen werden können. Vielmehr ist das meßbare Ausgangssignal von den (einigen) Systemzuständen abhängig. Bei den linearen Systemen ist $y(t)$ eine Linearkombination der Systemzustände, so daß das Zustandsmodell der linearen zeitinvarianten Systeme diese Form hat:

$$\left. \begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}^T \mathbf{x}(t) + du(t) \quad ; \quad \mathbf{x}_0 = \mathbf{x}(t_0) . \end{aligned} \right\} (3.1)$$

Die erste Gleichung wird auch als *Zustandsgleichung* und die zweite auch als *Ausgangs- oder Beobachtungsgleichung* bezeichnet:

- Die Zustandsgleichung ist eine *Vektordifferentialgleichung 1. Ordnung*, die die Systemdynamik beschreibt und angibt, wie das Eingangssignal die einzelnen Speicher beeinflusst und wie diese Speicher miteinander verkoppelt sind.
- Die algebraische Ausgangsgleichung beschreibt, wie das Ausgangssignal mit den Systemzuständen verbunden ist.

Der *Zustandsvektor* $\mathbf{x}(t)$ in Gl. (3.1) ist ein n -dimensionaler Vektor mit den zeitabhängigen Elementen $x_i(t)$

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} .$$

Die *Systemmatrix* \mathbf{A} ist eine konstante $(n \times n)$ -Matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} ,$$

\mathbf{b} ein n -dimensionaler Spaltenvektor

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

und \mathbf{c}^T ein n -dimensionaler Zeilenvektor mit konstanten Elementen

$$\mathbf{c}^T = [c_1 \quad c_2 \quad \dots \quad c_n] .$$

Der Skalar d ist bei nicht sprunghafigen Systemen, d.h. praktisch bei allen realen Systemen und Prozessen gleich Null. Die Anfangsbedingungen aller Komponenten von \mathbf{x} sind in dem n -dimensionalen Vektor \mathbf{x}_0 zusammengefaßt und werden als bekannt vorausgesetzt. Die Dimension n der Vektoren und der Matrix wird *Ordnung des Systems* genannt.

Zu einem Zustandsraummodell eines Systems kann man auf zwei Wegen gelangen:

- a) durch Aufstellen des physikalisch-technischen Wirkungszusammenhanges, wenn dieser durch Differentialgleichungen 1. Ordnung und/oder algebraische Beziehungen beschrieben wird,

- b) durch Umwandlung einer Differentialgleichung höherer Ordnung in ein System gekoppelter Differentialgleichungen 1. Ordnung.

Die zeitliche Abhängigkeit des n -dimensionalen Zustandsvektors \mathbf{x} kann man als Bewegung eines Punktes im n -dimensionalen Vektorraum \mathbf{R}^n (Zustandsraum) darstellen (Bild 3.3).

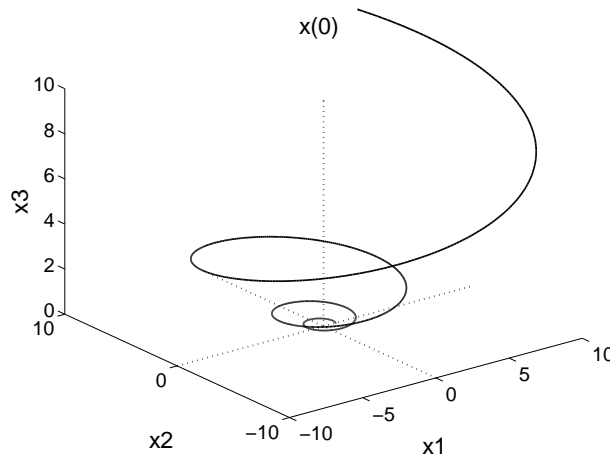


Bild 3.3: Trajektorie eines Systems dritter Ordnung im Zustandsraum

Der durch die Koordinaten von \mathbf{x} beschriebene Punkt verändert sich mit der Zeit und beschreibt eine Kurve im Zustandsraum, die als *Zustandskurve* oder *Trajektorie* des Systems bezeichnet wird.

3.2 Graphische Darstellung des Zustandsraummodells

Das durch die Gln. (3.1) beschriebene Zustandsraummodell ist in Bild 3.4 durch ein Blockschaltbild veranschaulicht, wobei Doppelpfeile vektorielle Größen darstellen.

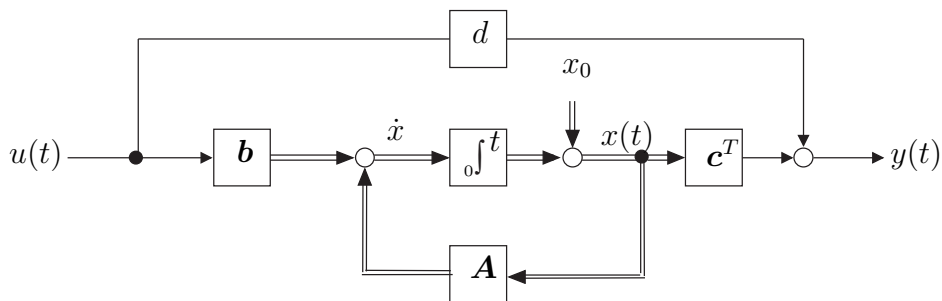


Bild 3.4: Blockschaltbild eines linearen Zustandsraummodells

Vier der fünf Blöcke sind statische Übertragungsglieder mit den Übertragungsfaktoren \mathbf{A} , \mathbf{b} , \mathbf{c}^T und d . Der mittlere Block enthält n Integratoren und kann — in Anlehnung an die in der Analogrechentechnik verwendeten Symbole — auch als Dreieck dargestellt werden.

Einen detaillierten Einblick in die Struktur eines Systems erlaubt die Darstellung als *Signalflußdiagramm* (Bild 3.5). Hierbei werden alle Signale durch Knoten repräsentiert. Die gerichteten Kanten zeigen, welches Signal direkt auf andere Signale einwirkt. Die Elemente der Matrix \mathbf{A} bzw. der Vektoren \mathbf{b} und \mathbf{c}^T sowie der Skalar d treten als Kantengewichte auf. Sind die entsprechenden Elemente gleich Null, so wird keine Kante in den Signalflußgraph eingetragen.

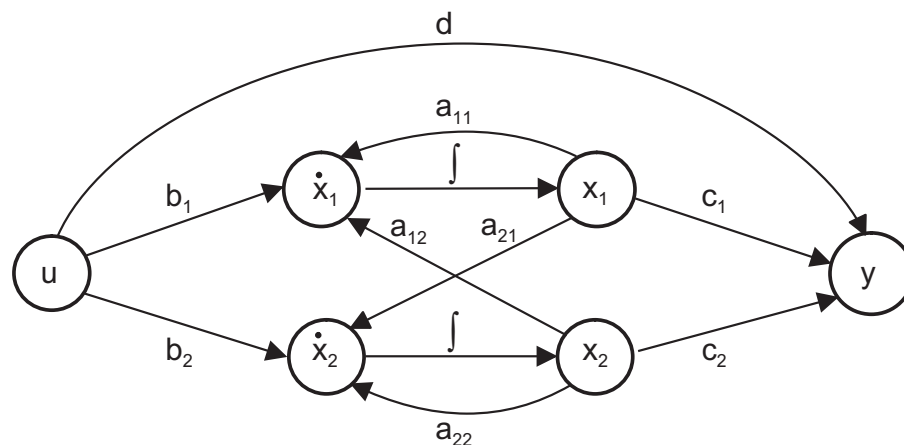


Bild 3.5: Signalflußdiagramm eines Systems zweiter Ordnung

3.3 Lösung der Zustandsgleichung

Die die Systemdynamik des Zustandsmodells (3.1) beschreibende Vektordifferentialgleichung 1. Ordnung

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \quad ; \quad \mathbf{x}_0 = \mathbf{x}(t_0) \quad (3.2)$$

muß in Abhängigkeit von den Anfangsbedingungen und der auf das System einwirkenden Eingangssignale $u(t)$ gelöst werden. Die Gesamtlösung läßt sich auch bei den linearen Vektordifferentialgleichungen aus zwei Teillösungen überlagern:

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_p(t) \quad . \quad (3.3)$$

Die homogene Lösung $\mathbf{x}_h(t)$ wird mittels der $(n \times n)$ *Matrixexponentialfunktion* $e^{\mathbf{A}t}$ aus den Anfangsbedingungen berechnet:

$$\mathbf{x}_h(t) = e^{\mathbf{A}t} \mathbf{x}_0 \quad , \quad (3.4)$$

wobei die Matrixexponentialfunktion wie folgt definiert ist:

$$e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \frac{\mathbf{A}^i t^i}{i!} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2}{2!} t^2 + \frac{\mathbf{A}^3}{3!} t^3 + \dots \quad (3.5)$$

Diese Reihe ist analog der Reihenentwicklung der e -Funktion aufgebaut:

$$e^{at} = \sum_{i=0}^{\infty} \frac{a^i t^i}{i!} = 1 + at + \frac{a^2}{2!} t^2 + \frac{a^3}{3!} t^3 + \dots \quad (3.6)$$

Die partikuläre Lösung läßt sich als Faltungsintegral darstellen:

$$\mathbf{x}_p(t) = \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{b}u(\tau) d\tau \quad . \quad (3.7)$$

So lautet die Lösung zu (3.2) mit (3.4) und (3.7):

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{b}u(\tau) d\tau \quad , \quad (3.8)$$

wobei wegen der Zeitinvarianz des betrachteten Systems $t_0 = 0$ gesetzt wurde.

Nach Einführung der Abkürzung

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} \quad . \quad (3.9)$$

erhält man als Lösung

$$\mathbf{x}(t) = \mathbf{\Phi}(t) \mathbf{x}_0 + \int_0^t \mathbf{\Phi}(t - \tau) \mathbf{b}u(\tau) d\tau \quad . \quad (3.10)$$

Die Gleichung (3.10) nennt man *Bewegungsgleichung* des Systems. Die darin vorkommende Matrix $\mathbf{\Phi}(t)$ wird als *Transitions-* oder *Übergangsmatrix* bezeichnet, da diese Matrix den Anfangszustand \mathbf{x}_0 für $u(t) = 0$ in den aktuellen Zustand $\mathbf{x}(t)$ überführt.

Mit

$$\mathbf{y}(t) = \mathbf{c}^T \mathbf{x}(t) \quad (3.11)$$

folgt dann auch für die Systemantwort:

$$\mathbf{y}(t) = \mathbf{c}^T \mathbf{\Phi}(t) \mathbf{x}_0 + \mathbf{c}^T \int_0^t \mathbf{\Phi}(t - \tau) \mathbf{b}u(\tau) d\tau \quad . \quad (3.12)$$

Setzt man hier $\mathbf{x}_0 = \mathbf{0}$, dann verbleibt eine Faltungsbeziehung zwischen $u(t)$ und $\mathbf{c}^T \mathbf{\Phi}(t) \mathbf{b}$. Wird nun ferner $u(t) = \delta(t)$ gesetzt, d. h. das System wird durch einen Dirac-Impuls

erregt, dann antwortet es nach Definition mit der Gewichtsfunktion $g(t)$. Ein Koeffizientenvergleich liefert den Zusammenhang zwischen den Systemmatrizen \mathbf{A} , \mathbf{b} , \mathbf{c}^T und der Gewichtsfunktion zu:

$$g(t) = \mathbf{c}^T \Phi(t) \mathbf{b} \quad . \quad (3.13)$$

Die Transitionsmatrix $\Phi(t)$ ist eine transzendente Funktion und hat folgende Eigenschaften:

$$(1) \quad \Phi(0) = \mathbf{I}_n \quad (3.14a)$$

$$(2) \quad \Phi(t_2 - t_0) = \Phi(t_2 - t_1) \Phi(t_1 - t_0) \quad (3.14b)$$

$$(3) \quad \Phi(t)^{-1} = \Phi(-t) \quad , \quad \Phi(t)^k = \Phi(kt) \quad (3.14c)$$

$$(4) \quad \dot{\Phi}(t) = \mathbf{A} \Phi(t) \quad . \quad (3.14d)$$

Sie kann für ein gegebenes System $\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t)$ auf verschiedene Art und Weise berechnet werden:

1. Durch numerische Integration der der Systemdifferentialgleichung zugeordneten Matrixdifferentialgleichung

$$\dot{\mathbf{X}}(t) = \mathbf{A} \mathbf{X}(t) \quad \mathbf{X}_0 = \mathbf{I}_n \quad . \quad (3.15)$$

2. Analog zu dem skalaren Fall als unendliche Reihe:

$$e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \frac{(\mathbf{A}t)^i}{i!} \quad . \quad (3.16)$$

3. Wenn die $n \times n$ Systemmatrix \mathbf{A} n linear unabhängige Eigenvektoren hat, kann sie mit Hilfe der aus den n Eigenvektoren gebildeten Transformationsmatrix \mathbf{T} auf Diagonalform transformiert werden:

$$\mathbf{T}^{-1} \mathbf{A} \mathbf{T} = \mathbf{J} = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix} \quad , \quad (3.17)$$

wobei $\lambda_1, \lambda_2, \dots, \lambda_n$ die Eigenwerte der Matrix \mathbf{A} sind. Die zur Diagonalmatrix \mathbf{J} gehörende Transitionsmatrix $e^{\mathbf{J}t}$ hat die leicht berechenbare Form

$$e^{\mathbf{J}t} = \begin{bmatrix} e^{\lambda_1 t} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & e^{\lambda_n t} \end{bmatrix} \quad . \quad (3.18)$$

Die gesuchte Transitionsmatrix $\Phi(t)$ ergibt sich dann zu:

$$\Phi(t) = e^{\mathbf{A}t} = e^{\mathbf{T} \mathbf{J} \mathbf{T}^{-1} t} = \mathbf{T} e^{\mathbf{J}t} \mathbf{T}^{-1} \quad . \quad (3.19)$$

4. Durch inverse Laplacetransformation

$$e^{\mathbf{A}t} = \mathcal{L}^{-1} \{ [s\mathbf{I} - \mathbf{A}]^{-1} \} = \mathcal{L}^{-1} \left\{ \frac{\text{adj}(s\mathbf{I} - \mathbf{A})}{\det(s\mathbf{I} - \mathbf{A})} \right\} \quad . \quad (3.20)$$

3.4 Lösung der Zustandsgleichung im Frequenzbereich

Wird auf das Zustandsmodell (3.1) die Laplace-Transformation angewendet:

$$\left. \begin{aligned} s\mathbf{X}(s) - \mathbf{x}_0 &= \mathbf{A}\mathbf{X}(s) + \mathbf{b}U(s) \\ Y(s) &= \mathbf{c}^T \mathbf{X}(s) \end{aligned} \right\} \quad (3.21)$$

und dann $\mathbf{X}(s)$ eliminiert, wird für das Übertragungsverhalten gefunden:

$$Y(s) = \mathbf{c}^T [\mathbf{I}s - \mathbf{A}]^{-1} \mathbf{x}_0 + \mathbf{c}^T [\mathbf{I}s - \mathbf{A}]^{-1} \mathbf{b}U(s) . \quad (3.22)$$

Durch Koeffizientenvergleich mit Gl. (3.12) erhält man die Übertragungsfunktion zu:

$$\left. \begin{aligned} G(s) &= \mathbf{c}^T [\mathbf{I}s - \mathbf{A}]^{-1} \mathbf{b} \\ &= \mathcal{L}\{g(t)\} = \mathcal{L}\{\mathbf{c}^T e^{\mathbf{A}t} \mathbf{b}\} \end{aligned} \right\} \quad (3.23)$$

3.5 Numerische Lösung der Zustandsgleichung mit Hilfe der Transitionsmatrix

Für die Ermittlung der numerischen Lösung der Zustandsgleichung (3.1) geht man von der allgemeinen Lösung

$$\mathbf{x}(t) = \Phi(t - t_0) \mathbf{x}_0 + \int_{t_0}^t \Phi(t - \tau) \mathbf{b}u(\tau) d\tau \quad (3.24)$$

aus. Soll das System für das Zeitintervall $[t_0, t_e]$ simuliert werden, so wird dieses Zeitintervall zunächst hinreichend fein unterteilt: $t_0 < t_1 < t_2 < \dots < t_e$. Wird die Eingangsgröße $u(t)$ zwischen zwei Zeitpunkten t_i und t_{i+1} als annähernd konstant angesehen, so erhält man aus (3.24) für den Anfangszeitpunkt t_i und den Endzeitpunkt t_{i+1} die Lösung

$$\mathbf{x}(t_{i+1}) = \Phi(t_{i+1} - t_i) \mathbf{x}(t_i) + \int_{t_i}^{t_{i+1}} \Phi(t_{i+1} - \tau) \mathbf{b}u(t_i) d\tau .$$

Da $u(t_i)$ im Intervall $\Delta t = t_{i+1} - t_i$ konstant sein soll, kann die mit $\mathbf{H}(\Delta t)$ bezeichnete Lösung des Integrals für eine reguläre Systemmatrix \mathbf{A} direkt angegeben werden:

$$\begin{aligned} \mathbf{H}(\Delta t) &= \int_{t_i}^{t_{i+1}} \Phi(t_{i+1} - \tau) \mathbf{b} d\tau = \int_0^{\Delta t} \Phi(\tau) \mathbf{b} d\tau \\ &= \mathbf{A}^{-1} \Phi(\Delta t) \mathbf{b} . \end{aligned} \quad (3.25)$$

Damit lautet die rekursive Lösung der Zustandsgleichung für stückweise konstante Eingangssignale:

$$\boxed{\mathbf{x}_{i+1} = \Phi(\Delta t) \mathbf{x}_i + \mathbf{H}(\Delta t) u_i.} \quad (3.26)$$

Für eine konstante Schrittweite Δt brauchen die konstanten Matrizen $\Phi(\Delta t)$ und $\mathbf{H}(\Delta t)$ nur einmal berechnet werden.

3.6 Stabilitätsanalyse im Zustandsraum

Die Definition 3.2 der *asymptotischen Stabilität* im Abschnitt 3.3 läßt sich wie folgt auf Systeme in einer Zustandsraumdarstellung (3.1) übertragen:

Definition 3.2 (*Stabilität eines linearen Systems*)

Ein lineares, zeitinvariantes System, das durch die Zustandsgleichungen (3.1) beschrieben wird, heißt **asymptotisch stabil**, wenn die Lösung $\mathbf{x}(t)$ der homogenen Zustandsdifferentialgleichung

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

für einen beliebigen Anfangszustand \mathbf{x}_0 für $t \rightarrow \infty$ gegen Null geht. \square

Aus der Bewegungsgleichung (3.10) des Systems erhält man für $u(t) = 0$ die Beziehung

$$\mathbf{x}(t) = \Phi(t)\mathbf{x}_0. \quad (3.27)$$

Für asymptotische Stabilität muß offensichtlich

$$\lim_{t \rightarrow \infty} \|\Phi(t)\| = 0 \quad (3.28)$$

gelten. Aus der Beziehung (3.19) folgt unmittelbar, daß Gl. (3.28) genau dann erfüllt ist, wenn alle Terme $e^{\lambda_i t}$ in Gl. (3.18) abklingen, d.h., wenn

$$\operatorname{Re}\{\lambda_i\} < 0 \quad \text{für} \quad i = 1, 2, \dots, n \quad (3.29)$$

gilt. Alle Eigenwerte der Systemmatrix \mathbf{A} , d.h. alle Wurzeln der charakteristischen Gleichung

$$C(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0 \quad (3.30)$$

müssen also in der offenen linken komplexen Halbebene liegen.

Die charakteristische Gleichung (3.30) ist für die Analyse eines linearen Systems im Zustandsraum von grundlegender Bedeutung:

1. Die charakteristische Gleichung bestimmt in erster Linie das Systemverhalten.
2. $C(\lambda)$ gibt Aufschluß über das Stabilitätsverhalten des Systems.
3. $C(\lambda)$ für ein System mit n konzentrierten Speichern ist ein Polynom n -ten Grades.
4. Die Werte λ_i für λ , die $C(\lambda) = 0$ erfüllen, heißen die *Eigenwerte* des Systems.
5. Die Eigenwerte λ_i bestimmen die *Eigenbewegungen* des Systems, die durch $x_i(t) = e^{\lambda_i t}$ beschrieben werden.
6. Für ein Polynom mit *reellen Koeffizienten* können die Wurzeln nur *reell* oder *paarweise konjugiert komplex* sein.

7. Damit ein System nur abklingende Eigenbewegungen hat, also asymptotisch stabil ist, ist notwendig und hinreichend, daß

- a) alle *reellen* Wurzeln negativ sind,
- b) alle komplexen Wurzeln negative Realteile haben.

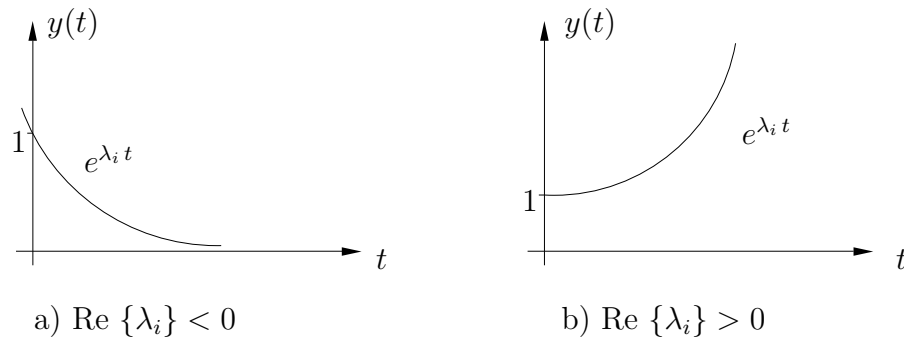


Bild 3.6: Stabilität von Eigenbewegungen

- a) abklingende Eigenbewegung
- b) aufklingende Eigenbewegung

8. Hat ein System außer Wurzeln mit negativen Realteilen auch rein imaginäre Wurzeln (Realteile sind Null), befindet sich das System am *Stabilitätsrand*.

9. Komplexe Wurzeln haben die Form:

$$\lambda_{1/2} = a \pm jb$$

Die zugehörigen Eigenbewegungen werden beschrieben durch:

$$x_{1/2}(t) = e^{(a \pm jb)t} = e^{at} e^{\pm jbt} \tag{3.31}$$

Mit der Euler'schen Relation:

$$e^{jz} = \cos z + j \sin z \tag{3.32}$$

erhält man

$$x_{1/2}(t) = e^{at} (\cos bt \pm j \sin bt) \tag{3.33}$$

Um die Stabilität eines Systems zu untersuchen, muß die Lage der Wurzeln der charakteristischen Gleichung in der komplexen λ -Ebene ermittelt werden. Bei Systemen höherer Ordnung ($n \geq 4$) ist dies nicht mehr in geschlossener Form, sondern nur noch (näherungsweise) numerisch möglich.

In Bild 3.7 sind die Impulsantworten für verschiedene Wurzellagen für zwei dynamische Systeme dargestellt:

3.6.1 Beziehungen zwischen asymptotischer Stabilität und BIBO–Stabilität

Die Definition der Eingangs– Ausgangs–Stabilität (BIBO–Stabilität, Definition 3.1 in SRT I) kann unmittelbar auf ein lineares Zustandsmodell übertragen werden:

Definition 3.3 (*BIBO–Stabilität eines linearen Zustandsmodells*)

Ein lineares, zeitinvariantes System, das durch die Zustandsgleichungen (3.1) beschrieben wird, heißt **BIBO–stabil**, wenn für verschwindende Anfangsauslenkungen, d.h.

$$\mathbf{x}_0 = \mathbf{0}$$

und ein beliebiges beschränktes Eingangssignal

$$|u(t)| < u_{\max} \quad \text{für alle } t > 0$$

das Ausgangssignal beschränkt bleibt:

$$|y(t)| < y_{\max} \quad \text{für alle } t > 0 .$$

□

Ein Übertragungssystem mit der Gewichtsfunktion $g(t)$ ist genau dann BIBO–stabil, wenn

$$\int_0^{\infty} |g(t)| dt < \infty \tag{3.34}$$

gilt (vgl. Abschnitt 3.3 in SRT I). Aus den Gleichungen (3.13) und (3.28) erhält man sofort den Zusammenhang, daß die asymptotische Stabilität des Zustandsmodells

$$\lim_{t \rightarrow \infty} g(t) = 0$$

sicherstellt und somit auch das Integral (3.34) existiert:

Satz 3.1 (*Asymptotische Stabilität und BIBO–Stabilität*)

Ist ein System (3.1) asymptotisch stabil, so ist es auch BIBO–stabil. □

Die Umkehrung dieses Satzes gilt nur dann, wenn alle Eigenbewegungen des Systems (3.1) in die Gewichtsfunktion eingehen und die Gewichtsfunktion $g(t)$ für $t \rightarrow \infty$ nicht verschwindet, sobald mindestens ein Eigenwert der Systemmatrix \mathbf{A} einen nichtnegativen Realteil hat. Für solche Systeme sind alle Eigenwerte der Systemmatrix \mathbf{A} auch Pole der Übertragungsfunktion.

Ist in der Gewichtsfunktion eine Eigenbewegung $e^{\lambda_i t}$ des Systems nicht enthalten, so bedeutet dies, daß der Eigenwert λ_i entweder nicht steuerbar oder nicht beobachtbar ist. Diese Eigenbewegung kann dann durch die Eingangsgröße nicht angeregt werden oder die Ausgangsgröße nicht beeinflussen.

3.7 Steuerbarkeit und Beobachtbarkeit

Ein grundlegendes und alle Zustandsraummethoden durchdringendes Konzept ist das der *Steuer- und Beobachtbarkeit*, mit dessen Hilfe Kalman Ende der fünfziger Jahre erstmalig anschaulich erklären konnte, warum ein instabiles System auch durch eine *perfekte* Kompensation der instabilen Pole nicht stabilisiert werden kann. Kalman zeigte, daß das System nach einer in der rechten s -Halbebene durchgeführten Pol-/Nullstellenkompensation immer noch instabil ist, und dann lediglich über eine *stabile* Übertragungsfunktion verfügt. Die Ordnung der Übertragungsfunktion ist dann jedoch kleiner als die Systemordnung (die Anzahl der konzentrierten Energiespeicher) und die instabilen Eigenbewegungen können entweder mit Hilfe von Stellsignalen nicht mehr beeinflußt werden (nicht steuerbar) oder/(und) sind im Ausgangssignal nicht mehr sichtbar (nicht beobachtbar). Die Steuerbarkeit eines linearen Systems $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$ kann dabei nach Kalman wie folgt definiert werden:

Definition 3.4 (*Zustandssteuerbarkeit*)

Ein dynamisches System (\mathbf{A}, \mathbf{b}) heißt *vollständig zustandssteuerbar* (*z-steuerbar*), wenn der Zustandsvektor $\mathbf{x}(t)$ durch eine geeignete Steuerfunktion $u(t)$ in einer endlichen Zeitspanne $[t_0, t_e]$ von jedem beliebigen Anfangszustand $\mathbf{x}(t_0) = \mathbf{x}_0$ in den Nullzustand $\mathbf{x}(t_e) = \mathbf{0}$ überführt werden kann.

Diese Eigenschaft der Steuerbarkeit ist eine notwendige Voraussetzung für viele Reglerentwurfverfahren und somit auch bei praktischen Aufgabenstellungen von entscheidender Bedeutung. Eng verwandt mit der Steuerbarkeit ist der Begriff der *Erreichbarkeit*:

Definition 3.5 (*Zustandserreichbarkeit*)

Ein dynamisches System (\mathbf{A}, \mathbf{b}) heißt *vollständig erreichbar*, wenn der Zustandsvektor $\mathbf{x}(t)$ durch eine geeignete Steuerfunktion $u(t)$ in einer endlichen Zeitspanne $[t_0, t_e]$ aus dem Nullzustand $\mathbf{x}(t_0) = \mathbf{0}$ in jeden gewünschten Endzustand $\mathbf{x}(t_e)$ überführt werden kann.

Im allgemeinen stellt die Erreichbarkeit im Vergleich zur Steuerbarkeit höhere Anforderungen an die Eigenschaften eines dynamischen Systems, da jedes asymptotisch stabile System für $u(t) = 0$ und $\mathbf{x}_0 \neq \mathbf{0}$ nach einer entsprechenden Zeit immer von selbst in den Nullzustand zurückkehrt. Die Erreichbarkeit und die Steuerbarkeit sind daher auch nur bei der in dieser Vorlesung betrachteten Klasse der linearen, zeitinvarianten Systeme äquivalente Eigenschaften. Bereits bei linearen Systemen die zeitvariabel oder zeitdiskret sind, muß zwischen der Erreichbarkeit und der Steuerbarkeit unterschieden werden. Dies gilt natürlich erst recht bei allen nichtlinearen Systemen. Da es in der Literatur zur linearen Systemtheorie üblich ist, die für lineare, zeitinvariante Systeme äquivalenten Eigenschaften Steuer- und Erreichbarkeit unter dem Oberbegriff Steuerbarkeit zusammenzufassen, wird im weiteren auch nur noch die Steuerbarkeit betrachtet.

In der Praxis stellt allerdings häufig nicht die Erreichung eines bestimmten Systemzustandes das primäre Ziel der Regelungsaufgabe dar. Vielmehr ist es erwünscht, nur dem *Systemausgang* einen vorgegebenen Wert (oder Funktion) zu erteilen. In diesem Zusammenhang spricht man dann von *Ausgangssteuerbarkeit* und es gilt analog zur Zustandssteuerbarkeit:

Definition 3.6 (*Ausgangssteuerbarkeit*)

Ein dynamisches System $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$; $y(t) = \mathbf{c}^T \mathbf{x}(t)$ heißt vollständig ausgangsteuerbar, wenn die Ausgangsgröße $y(t)$ durch eine geeignete Steuerfunktion $u(t)$ in einer endlichen Zeitspanne $[t_0, t_e]$ von einem beliebigen Anfangswert $y(t_0) = y_0$ in irgendeinen Endwert $y(t_e)$ überführt werden kann.

Die Eigenschaft der Steuerbarkeit ist eine Systemeigenschaft und daher von der Art der Stellgrößensignale unabhängig, d. h. über die Art und insbesondere den Energieinhalt von $u(t)$ werden keine Einschränkungen gemacht. Die z -Steuerbarkeit des Systems ist ein den beiden Matrizen \mathbf{A} und \mathbf{b} in (3.1) gemeinsames Strukturmerkmal, das angibt, ob die Lösung der Systemgleichung den vollständigen Vektorraum des Zustandsvektors aufspannt.

Denn nur wenn der Teilausdruck

$$\mathbf{m}(t_e) = \int_{t_0}^{t_e} e^{\mathbf{A}(t_e-\tau)} \mathbf{b}u(\tau) d\tau \tag{3.35}$$

in der Lösung (3.10) zu (3.1):

$$\mathbf{x}(t_e) = e^{\mathbf{A}t_e} \mathbf{x}_0 + \int_{t_0}^{t_e} e^{\mathbf{A}(t_e-\tau)} \mathbf{b}u(\tau) d\tau \tag{3.36}$$

ein vollständiges System von n linear unabhängigen Basisvektoren für den n -dimensionalen Vektorraum \mathbb{R}^n des Zustandsvektors $\mathbf{x}(t)$ enthält, kann, ausgehend von dem Anfangszustand $\mathbf{x}(t_0)$, ein beliebig in \mathbb{R}^n vorgegebener Endzustand erreicht werden. Kriterien zur Überprüfung der z -Steuerbarkeit eines Systems überprüfen letztlich mit Methoden der Algebra der Vektorräume, ob der Zustandsraum eines Systems gegebenenfalls (bei nicht vollständiger Steuerbarkeit) in orthogonale Unterräume von vollständig steuerbaren und vollständig nicht steuerbaren Zuständen zerlegt werden kann.

Satz 3.2 (*Steuerbarkeitskriterien*)

Ein zeitinvariantes System (3.1) der Ordnung n ist

a1) (Kalman-Kriterium)

dann und nur dann vollständig z -steuerbar, wenn für die $(n \times n)$ Steuerbarkeitsmatrix \mathbf{Q}_S gilt:

$$\text{Rang } \mathbf{Q}_S = \text{Rang} [\mathbf{b} \ \mathbf{A}\mathbf{b} \ \cdots \ \mathbf{A}^{n-1}\mathbf{b}] = n \tag{3.37}$$

bzw.

a2) (Hautus-Kriterium)

Überprüfung der einzelnen Eigenwerte auf Steuerbarkeit

$$\text{Rang } \mathbf{Q}_S^* = \text{Rang} [\lambda \mathbf{I} - \mathbf{A} \mathbf{b}]_{\lambda=\lambda_i} = n \quad (3.38)$$

mit $\lambda_i, i = 1, 2, \dots, n$ den Eigenwerten von \mathbf{A} .

b) Der Index $n_1 = \text{Rang } \mathbf{Q}_S$ gibt die Dimension des vollständig z-steuerbaren Unterraumes $\mathbb{R}^{n_1} \subset \mathbb{R}^n$ an.

c) Jedes im Intervall $[t_0, t_e]$ vollständig z-steuerbare (zeitinvariante) System (\mathbf{A}, \mathbf{b}) ist auch für alle t_1 mit $t_0 < t_1 < \infty$ vollständig z-steuerbar. \square

Ein zu der Kalman-Bedingung (3.37) analoges Kriterium zur Überprüfung der Ausgangssteuerbarkeit lautet:

Satz 3.3 (*Kriterium zur Überprüfung der Ausgangssteuerbarkeit*)

Ein dynamisches System $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ ist genau dann vollständig ausgangssteuerbar, wenn für die Ausgangs-Steuerbarkeitsmatrix \mathbf{Q}_A gilt:

$$\text{Rang } \mathbf{Q}_A = \text{Rang} [\mathbf{c}^T \mathbf{b} \mathbf{c}^T \mathbf{A} \mathbf{b} \dots \mathbf{c}^T \mathbf{A}^{n-1} \mathbf{b}] = 1 . \quad (3.39)$$

Obwohl eng mit der Zustandssteuerbarkeit verknüpft, ist die vollständige Zustandssteuerbarkeit eines Systems weder notwendig noch hinreichend für die vollständige Steuerbarkeit der Ausgangsgrößen. So ist beispielsweise ein vollständig zustandssteuerbares System nur dann vollständig ausgangssteuerbar, wenn $\text{Rang } \mathbf{c} = 1$ gilt. Andererseits kann aber auch ein nicht vollständig zustandssteuerbares System durchaus vollständig ausgangssteuerbar sein.

Die Eigenschaft der Zustandsbeobachtbarkeit eines Systems ist dual zu der z-Steuerbarkeit und ist für praktische Fragestellungen deshalb wichtig, da bei einem Zustandsmodell die Systemzustände im allgemeinen nicht direkt meßbar sind, sondern nur über die Ausgangssignale als Linearkombinationen zur Verfügung stehen. Es erhebt sich also die Frage, ob aus der Kenntnis der Ausgangssignale $y(t)$ auf den Systemzustand zu einem bestimmten Zeitpunkt, z. B. $\mathbf{x}(t_0)$, geschlossen werden kann.

Definition 3.7 *Beobachtbarkeit*

Das dynamische System (3.1) heißt *vollständig beobachtbar* im Intervall $[t_0, t_e]$, wenn für gegebene t_0 und t_e jeder Systemzustand $\mathbf{x}(t_0) \in \mathbb{R}^n$ aus der Kenntnis der Eingangsgröße $u(t)$ und der Ausgangsgröße $y(t)$ in $[t_0, t_e]$ ermittelt werden kann. \square

Da mit $u(t) = 0$ aus Gl. 3.12 für den Systemausgang

$$y(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{x}_0 \quad (3.40)$$

folgt, bedeutet die Eigenschaft der vollständigen Beobachtbarkeit eines Systems, daß auf jeden Fall alle Elemente des Zeilenvektors

$$\mathbf{c}^T e^{\mathbf{A}t} \quad (3.41)$$

von Null verschieden sein müssen. Darüber hinaus muß noch die folgende Bedingung erfüllt sein:

Satz 3.4 (Beobachtbarkeitskriterien)

Ein zeitinvariantes System (3.1) der Ordnung n ist

a1) (Kalman Kriterium)

dann und nur dann vollständig beobachtbar, wenn gilt:

$$\text{Rang } \mathbf{Q}_B = \text{Rang} \left[\mathbf{c} \mathbf{A}^T \mathbf{c} \dots \dots (\mathbf{A}^T)^{n-1} \mathbf{c} \right] = n \quad (3.42)$$

bzw.

a2) (Hautus-Kriterium)

Überprüfung der einzelnen Eigenwerte auf Beobachtbarkeit

$$\text{Rang } \mathbf{Q}_B^* = \text{Rang} \left[\lambda \mathbf{I} - \mathbf{A}^T \mathbf{c} \right]_{\lambda=\lambda_i} = n \quad (3.43)$$

mit $\lambda_i, i = 1, 2, \dots, n$ den Eigenwerten von \mathbf{A} .

b) Der Index $n_2 = \text{Rang } \mathbf{Q}_B$ gibt die Ordnung des vollständig beobachtbaren Unterraumes $\mathbb{R}^{n_2} \subset \mathbb{R}^n$ an.

c) Ein im Intervall $[t_0, t_e]$ vollständig beobachtbares System ist für alle t_1 mit $t_0 < t_1 < \infty$ vollständig beobachtbar. \square

Die Eigenschaften der vollständigen z-Steuerbarkeit und Beobachtbarkeit lassen sich bei zeitinvarianten Systemen dann besonders einfach bestimmen, wenn die Systemmatrix eine reine Diagonalmatrix der (*durchweg verschiedenen*) Eigenwerte der Systeme ist. Bei durchweg verschiedenen Eigenwerten kann jedes Zustandsraummodell $(\mathbf{A}, \mathbf{b}, \mathbf{c}^T)$ auf eine derartige kanonische Normalform transformiert werden.

Hierzu transformiert man den Zustandsvektor $\mathbf{x}(t)$ mit einer regulären $(n \times n)$ Matrix \mathbf{V}

$$\tilde{\mathbf{x}}(t) = \mathbf{V}^{-1} \mathbf{x}(t), \quad (3.44)$$

die man aus den n Eigenvektoren v_i der Systemmatrix \mathbf{A} bildet:

$$\mathbf{V} = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n) . \quad (3.45)$$

Die transformierten Systemgleichungen (3.1) lauten dann

$$\left. \begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \lambda_2 & \\ & & \ddots \\ \mathbf{0} & & & \lambda_n \end{bmatrix} \tilde{\mathbf{x}}(t) + \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{bmatrix} u(t) , \\ y(t) &= [\tilde{c}_1 \ \tilde{c}_2 \ \dots \ \tilde{c}_n] \tilde{\mathbf{x}}(t) . \end{aligned} \right\} (3.46)$$

wobei

$$\text{diag } \lambda_i = \mathbf{V}^{-1} \mathbf{A} \mathbf{V} \quad (3.47)$$

$$\tilde{\mathbf{b}} = \mathbf{V}^{-1} \mathbf{b} \quad (3.48)$$

$$\tilde{\mathbf{c}}^T = \mathbf{c}^T \mathbf{V} \quad (3.49)$$

$$\tilde{\mathbf{x}}_0 = \mathbf{V}^{-1} \mathbf{x}_0 \quad (3.50)$$

gilt.

Das Signalflußdiagramm des transformierten Modells ist in Bild 3.8 angegeben. Wie die einzelnen transformierten Zustandsvariablen durch die Eingangsgröße angeregt werden und wie sie die Ausgangsgröße beeinflussen, wird durch die Vektoren $\tilde{\mathbf{b}}$ und $\tilde{\mathbf{c}}^T$ beschrieben.

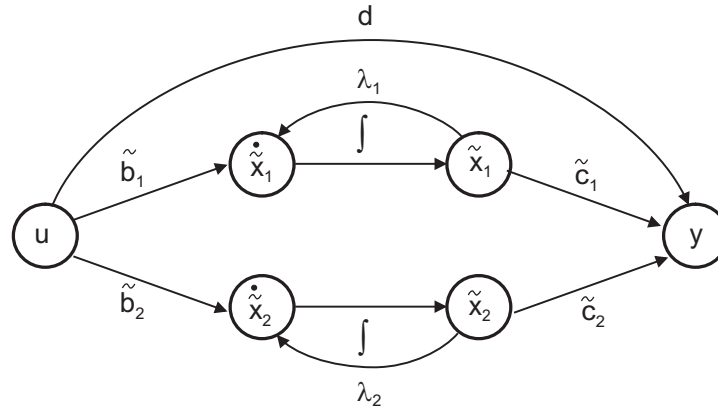


Bild 3.8: Signalflußdiagramm eines Systems zweiter Ordnung in Diagonalforn

Ein solches System ist dann vollständig steuerbar, wenn $\lambda_i \neq \lambda_j$ für $i, j = 1, 2, \dots, n$ und alle $b_i \neq 0$ und vollständig beobachtbar, wenn mit $\lambda_i \neq \lambda_j$ alle $c_i \neq 0$ sind.

Das Konzept der vollständigen Steuerbarkeit und Beobachtbarkeit hat für sehr viele regelungstechnische Fragen eine fundamentale Bedeutung. So können nur solche instabilen Systemteile stabilisiert werden, die auch vollständig steuerbar sind. Ganz allgemein können nur vollständig steuerbare Systeme durch ein geeignetes Regelgesetz in ihrer Dynamik beliebig verändert werden. Ferner sind nur solche Systemzustände für ein Regelgesetz direkt oder indirekt über Beobachtersysteme zugänglich, die vollständig beobachtbar sind.

3.7.1 Dualität von Steuerbarkeit und Beobachtbarkeit

Die Kriterien zur Überprüfung der Steuerbarkeit und Beobachtbarkeit sind sehr ähnlich. Wendet man das Steuerbarkeitskriterium auf das System $(\mathbf{A}^T, \mathbf{c})$ an, so erhält man das Beobachtbarkeitskriterium. Man bezeichnet die Steuer- und Beobachtbarkeit daher auch als *duale* Eigenschaften.

Dementsprechend nennt man das System

$$\left. \begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \mathbf{A}^T \tilde{\mathbf{x}}(t) + \mathbf{c}\tilde{u}(t) \\ \tilde{\mathbf{y}}(t) &= \mathbf{b}^T \tilde{\mathbf{x}}(t) \end{aligned} \right\} \quad (3.51)$$

als das zu (3.1) *duale System*. Die Matrizen des dualen Systems $(\mathbf{A}^T, \mathbf{c}, \mathbf{b}^T)$ sind gerade die transponierten Matrizen des Originalsystems $(\mathbf{A}, \mathbf{b}, \mathbf{c}^T)$. Die Dualität von Steuerbarkeit und Beobachtbarkeit stellt sich dann wie folgt dar:

Satz 3.5 (Dualität von Steuer- und Beobachtbarkeit)

Ein zeitinvariantes System (3.1) ist vollständig zustandssteuerbar (beobachtbar), wenn sein duales System (3.51) vollständig beobachtbar (zustandssteuerbar) ist.

3.8 Kalman–Zerlegung des Zustandsraummodells

Im allgemeinen kann der Zustandsvektor \mathbf{x} eines Zustandsraummodells durch eine Transformation $\tilde{\mathbf{x}} = \mathbf{T}^{-1}\mathbf{x}$ in einen Vektor überführt werden, der aus vier Teilvektoren besteht:

$$\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_1^T \quad \tilde{\mathbf{x}}_2^T \quad \tilde{\mathbf{x}}_3^T \quad \tilde{\mathbf{x}}_4^T]^T .$$

Dabei enthält

- $\tilde{\mathbf{x}}_1$ die vollständig steuerbaren, aber nicht beobachtbaren Zustandsgrößen,
- $\tilde{\mathbf{x}}_2$ die vollständig steuerbaren und vollständig beobachtbaren Zustandsgrößen,
- $\tilde{\mathbf{x}}_3$ die vollständig beobachtbaren, aber nicht steuerbaren Zustandsgrößen,
- $\tilde{\mathbf{x}}_4$ die nicht steuerbaren und nicht beobachtbaren Zustandsgrößen.

Diese Transformation führt auf ein Zustandsmodell, das folgende Struktur hat:

$$\begin{bmatrix} \dot{\tilde{\mathbf{x}}}_1(t) \\ \dot{\tilde{\mathbf{x}}}_2(t) \\ \dot{\tilde{\mathbf{x}}}_3(t) \\ \dot{\tilde{\mathbf{x}}}_4(t) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{A}}_{13} & \tilde{\mathbf{A}}_{14} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} & \tilde{\mathbf{A}}_{23} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{A}}_{33} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{A}}_{43} & \tilde{\mathbf{A}}_{44} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_1(t) \\ \tilde{\mathbf{x}}_2(t) \\ \tilde{\mathbf{x}}_3(t) \\ \tilde{\mathbf{x}}_4(t) \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} u(t) \quad (3.52)$$

$$y(t) = [\mathbf{0}^T \quad \tilde{\mathbf{c}}_2^T \quad \tilde{\mathbf{c}}_3^T \quad \mathbf{0}^T] \tilde{\mathbf{x}}(t) . \quad (3.53)$$

Diese sogenannte Kalman–Zerlegung kann als Blockschaltbild, wie in Bild 3.9 dargestellt, angegeben werden.

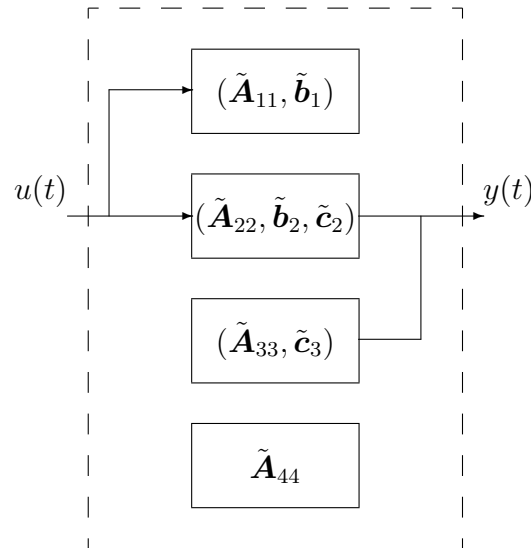


Bild 3.9: Kalman-Zerlegung des Zustandsraummodells

Alleine der vollständig steuerbare und beobachtbare Systemteil $(\tilde{\mathbf{A}}_{22}, \tilde{\mathbf{b}}_2, \tilde{\mathbf{c}}_2)$ wird durch die Übertragungsfunktion

$$G(s) = \mathbf{c}^T [s\mathbf{I} - \mathbf{A}]^{-1} \mathbf{b} = \mathbf{c}^T \frac{\text{adj}(s\mathbf{I} - \mathbf{A})}{|s\mathbf{I} - \mathbf{A}|} \mathbf{b} = \frac{Z(s)}{N(s)} \quad (3.54)$$

mit

$$N(s) = a_0 + a_1 s + \cdots + a_{n-1} s^{n-1} + s^n \quad (3.55)$$

beschrieben. Die Übertragungsfunktion $G(s)$ ist eine Ein-/ Ausgangsbeschreibung und daher auf den steuer- und beobachtbaren Systemteil beschränkt. Nur für ein vollständig steuer- und beobachtbares System $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ ist die Zahl der Pole von $G(s)$, d. h. der Grad des Nennerpolynoms, gleich der Anzahl der Eigenwerte der Systemmatrix \mathbf{A} .

Ist das System nicht vollständig steuer- und/oder beobachtbar, dann haben das Zähler- und Nennerpolynom von $G(s)$ gemeinsame Wurzeln, d. h. durch entsprechende Linearfaktoren kann gekürzt werden. Damit hat dann das System weniger Pole als Eigenwerte.

3.9 Normalformen für Zustandsraummodelle

In dem vorhergehenden Abschnitt wurde die Diagonalform als kanonische Normalform eines Zustandsraummodells bereits vorgestellt. Außer der Diagonalform gibt es weitere Normalformen, die wichtige Strukturmerkmale leicht erkennen lassen. Es sind dies unter anderem die Jordan-, Regelungs- und Beobachtungsnormalformen, die im weiteren für nicht-sprungfähige Systeme erläutert werden.

3.9.1 Jordannormalform

Bisher wurde angenommen, daß die Systemmatrix \mathbf{A} genau n verschiedene Eigenwerte hat. Unter dieser Voraussetzung ist die Matrix \mathbf{A} diagonalähnlich und kann mit Hilfe der Eigenvektoren in die Diagonalmatrix $\text{diag } \lambda_i$ transformiert werden (vgl. Gl. (3.46)). Wenn die Systemmatrix mehrfache Eigenwerte besitzt, kann \mathbf{A} nur noch unter bestimmten Voraussetzungen auf Diagonalform gebracht werden.

Eine Transformation auf Diagonalform ist nur dann möglich, wenn für einen l_i -fachen Eigenwert λ_i der Matrix \mathbf{A}

$$\text{Rang } (\lambda_i \mathbf{I} - \mathbf{A}) = n - l_i \tag{3.56}$$

gilt und die Matrix $(\lambda_i \mathbf{I} - \mathbf{A})$ gerade $n - l_i$ unabhängige Eigenvektoren besitzt.

Im Fall, daß die Bedingung (3.56) nicht erfüllt ist, gibt es für einen l_i -fachen Eigenwert weniger als l_i linear unabhängige Eigenvektoren und die Matrix \mathbf{A} kann dann nicht auf Diagonalform transformiert werden. Existiert für einen l_i -fachen Eigenwert nur ein einziger linear unabhängiger Eigenvektor, dann müssen zur Bildung der Transformationsmatrix \mathbf{V} $l_i - 1$ weitere Vektoren gebildet werden, die als *verallgemeinerte Eigenvektoren* oder *Hauptvektoren* der Matrix \mathbf{A} bezeichnet werden. Man erhält dann durch eine Ähnlichkeitstransformation mit der Matrix \mathbf{V} die neue Matrix

$$\mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \mathbf{J} = \left[\begin{array}{ccccc|ccc} \lambda_i & 1 & 0 & \dots & 0 & & & \\ 0 & \lambda_i & 1 & \dots & 0 & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & & & \\ 0 & 0 & \dots & \lambda_i & 1 & & & \\ 0 & 0 & \dots & 0 & \lambda_i & & & \\ \hline & & & & & \lambda_{l_i+1} & & \\ & & & & & \dots & & \\ & & & & & & & \lambda_n \end{array} \right] \left. \vphantom{\begin{array}{c} \lambda_i \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}} \right\} l_i \text{ Zeilen} \tag{3.57}$$

die man *Jordannormalform* der Matrix \mathbf{A} nennt.

Die zugehörige Matrixexponentialfunktion der Jordannormalform (3.57) hat dann folgendes Aussehen

$$e^{\mathbf{J}t} = \left[\begin{array}{ccccc|ccc} e^{\lambda_i t} & t e^{\lambda_i t} & t^2 e^{\lambda_i t} / 2! & \dots & t^{l_i-1} e^{\lambda_i t} / (l_i - 1)! & & & \\ 0 & e^{\lambda_i t} & t e^{\lambda_i t} & \dots & t^{l_i-2} e^{\lambda_i t} / (l_i - 2)! & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & & & \\ 0 & 0 & \dots & e^{\lambda_i t} & t e^{\lambda_i t} & & & \\ 0 & 0 & \dots & 0 & e^{\lambda_i t} & & & \\ \hline & & & & & e^{\lambda_{l_i+1} t} & & \\ & & & & & \dots & & \\ & & & & & & & e^{\lambda_n t} \end{array} \right] \tag{3.58}$$

3.9.2 Regelungsnormalform

Neben der Diagonalfom und der Jordannormalform gibt es eine Reihe weiterer Normalformen des Zustandsraummodells, von denen als nächste die *Regelungsnormalform* behandelt wird.

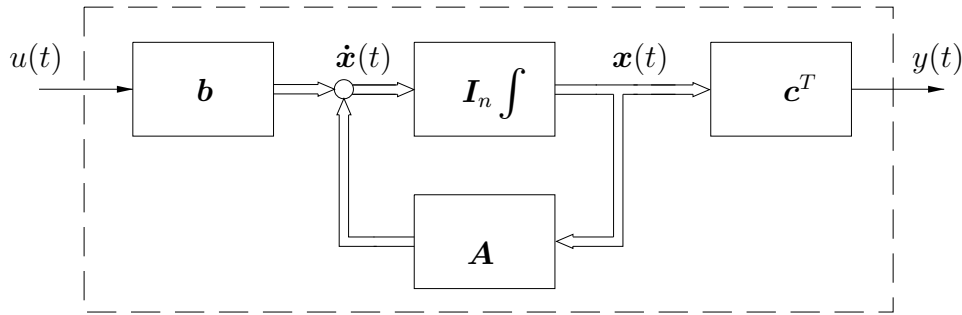


Bild 3.10: Blockschaltbild des Zustandsmodells eines Eingrößensystems

Definition 3.8 Regelungsnormalform

Ein Eingrößensystem (Bild 3.10) hat regelungskanonische Normalform, wenn das Systemmodell diese Form hat:

$$\left. \begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u(t) \quad , \\ y(t) &= [\quad b_0 \quad b_1 \quad b_2 \quad \cdots \quad b_{n-1}] \mathbf{x}(t) \end{aligned} \right\} (3.59)$$

oder abgekürzt

$$\left. \begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_R \mathbf{x}(t) + \mathbf{b}_R u(t) \quad , \\ y(t) &= \mathbf{c}_R^T \mathbf{x}(t) \quad . \end{aligned} \right\} (3.60)$$

□

Die Systemmatrix \mathbf{A}_R ist von Frobenius–kanonischer Form und liefert das charakteristische Polynom des Systems in den Koeffizienten a_i . Diese Polynomkoeffizienten sowie auch die Elemente des Zeilenvektors \mathbf{c}_R^T sind so bezeichnet, daß der Zusammenhang mit den Zähler– und Nennerpolynomen der zugehörigen Übertragungsfunktionen sogleich erkennbar wird:

$$G(s) = \frac{Z(s)}{N(s)} = \frac{b_0 + b_1 s + \cdots + b_{n-1} s^{n-1}}{a_0 + a_1 s + \cdots + a_{n-1} s^{n-1} + s^n} = \mathbf{c}_R^T [\mathbf{I}s - \mathbf{A}_R]^{-1} \mathbf{b}_R . \quad (3.61)$$

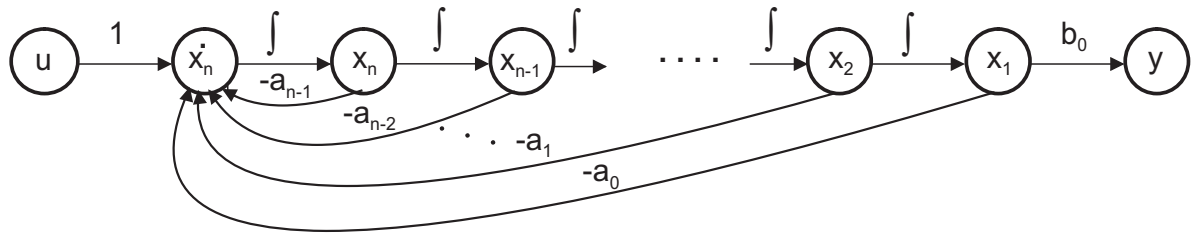


Bild 3.11: Signalflußdiagramm eines Systems in Regelungsnormform

In Bild 3.11 ist das Signalflußdiagramm der Steuerungsnormalform eines Zustandsraummodells dargestellt, dessen Übertragungsfunktion keine Nullstellen besitzt, d.h. es gilt $Z(s) = b_0$.

Man erkennt, daß die Eingangsgröße $u(t)$ nur direkt auf \dot{x}_n wirkt und alle Zustandsgrößen $x_i(t)$, $i = 1, 2, \dots, n - 1$ durch Integration von $x_{i+1}(t)$ hervorgehen. Der Ausgang ist proportional zu Zustandsgröße $x_1(t)$.

Jedes vollständig steuerbare System mit einer invertierbaren Steuerbarkeitsmatrix (3.37) kann auf die Regelungsnormform transformiert werden. Bezeichnet man die letzte Zeile der Inversen der Steuerbarkeitsmatrix mit \mathbf{q}'_S

$$\mathbf{q}'_S = [0 \ 0 \ \dots \ 0 \ 1] \mathbf{Q}_S^{-1}, \quad (3.62)$$

dann liefert eine Transformation

$$\mathbf{x}_R(t) = \mathbf{T}_R^{-1} \mathbf{x}(t) \quad (3.63)$$

mit

$$\mathbf{T}_R = \begin{bmatrix} \mathbf{q}'_S \\ \mathbf{q}'_S \mathbf{A} \\ \mathbf{q}'_S \mathbf{A}^2 \\ \vdots \\ \mathbf{q}'_S \mathbf{A}^{n-1} \end{bmatrix} \quad (3.64)$$

die gesuchte Regelungsnormform. Für die Matrizen der Regelungsnormform gelten dann die folgenden Beziehungen:

$$\mathbf{A}_R = \mathbf{T}_R^{-1} \mathbf{A} \mathbf{T}_R \quad (3.65)$$

$$\mathbf{b}_R = \mathbf{T}_R^{-1} \mathbf{b} \quad (3.66)$$

$$\mathbf{c}_R^T = \mathbf{c}^T \mathbf{T}_R \quad (3.67)$$

$$\mathbf{x}_{R0} = \mathbf{T}_R^{-1} \mathbf{x}_0. \quad (3.68)$$

3.9.3 Beobachtungsnormalform

Abschließend sei hier noch die zu der oben erklärten regelungskanonischen Struktur duale Form, die *Beobachtungsnormalform*, erwähnt.

Definition 3.9 *Beobachtungsnormalform*

Ein Eingrößensystem (Bild 3.10) ist in *Beobachtungsnormalform*, wenn das Zustandsraummodell $(\mathbf{A}_B, \mathbf{b}_B, \mathbf{c}_B)$ die Form

$$\left. \begin{aligned} \dot{\mathbf{x}}_B(t) &= \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{bmatrix} \mathbf{x}_B(t) + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ \vdots \\ b_{n-2} \\ b_{n-1} \end{bmatrix} u(t) \\ y(t) &= \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \mathbf{x}_B(t) \end{aligned} \right\} (3.69)$$

mit

$$G(s) = \frac{Z(s)}{N(s)} = \frac{b_0 + b_1 s + \cdots + b_{n-1} s^{n-1}}{a_0 + a_1 s + \cdots + a_{n-1} s^{n-1} + s^n} \quad (3.70)$$

hat.

□

4 Entwurf linearer Zustandsregler

4.1 Einleitung

Eine der wichtigsten Aufgaben für den Regelungstechniker stellt der Entwurf oder die Synthese eines Regelkreises dar, wobei für eine vorgegebene Regelstrecke ein geeigneter Regler zu entwerfen ist, der die an den Regelkreis gestellten Anforderungen möglichst gut oder bei geringstem technischen Aufwand erfüllt.

Der Entwurf eines linearen Regelkreises kann sowohl im Zeit- als auch im Frequenzbereich erfolgen. Unabhängig vom gewählten Entwurfsverfahren fallen bei der Lösung einer Regelungsaufgabe folgende Teilaufgaben an:

Lösung einer Regelungsaufgabe

1. Aufstellung eines linearen Modells der Regelstrecke.
2. Analyse des dynamischen Verhaltens (Stabilität, Dämpfung, Minimalphasigkeit, Steuer- und Beobachtbarkeit ...).
3. Anpassung der Güteforderungen an das jeweilige Entwurfsverfahren.
4. Wahl der Regelkreisstruktur.
5. Wahl des Reglers
6. Festlegung der Reglerparameter.
7. Simulation des Verhaltens des geschlossenen Regelkreises.
8. Bewertung der Regelgüte anhand der Güteanforderungen.

Für den Reglerentwurf ist typisch, daß einige dieser Teilaufgaben im allgemeinen mehrfach zu lösen sind.

4.2 Entwurf einer Zustandsrückführung

Das Verhalten eines dynamischen Systems wurde bisher durch Rückkopplung der Ausgangsgröße verändert. Bei der Art der Rückkopplung kann zwischen

- a) Ausgangsrückführung und
- b) Zustandsrückführung

unterschieden werden. Für die Ausgangssignalarückführung nach Bild 4.1 gilt dann im Bildbereich der Zusammenhang zwischen Stör- und Führungsgrößen als Eingangssignal und der Regelgröße als Ausgangssignal:

$$Y(s) = \frac{G_S(s)}{1 + G_S(s)G_R(s)}Z(s) + \frac{G_S(s)G_R(s)}{1 + G_S(s)G_R(s)}W(s) \quad . \quad (4.1)$$

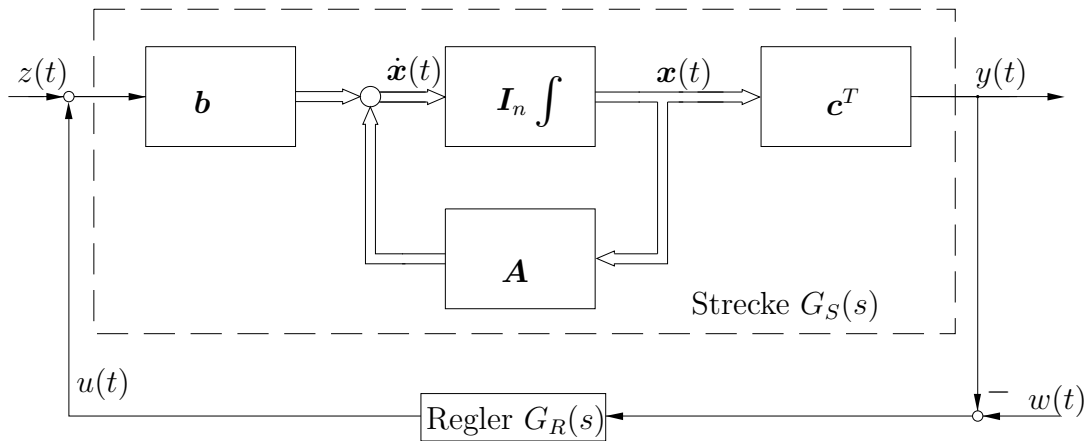


Bild 4.1: Ausgangssignalsrückführung einer Regelstrecke in Zustandsmodelldarstellung

Wird eine Regelstrecke durch eine Zustandsrückführung geregelt (Bild 4.2), dann gelten folgende Beziehungen im Zeitbereich:

$$\text{Regelstrecke: } \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}[z(t) + u(t)] \quad (4.2a)$$

$$y(t) = \mathbf{c}^T \mathbf{x}(t) \quad (4.2b)$$

$$\text{Regelgesetz: } u(t) = -\mathbf{k}^T \mathbf{x}(t) \quad (4.3)$$

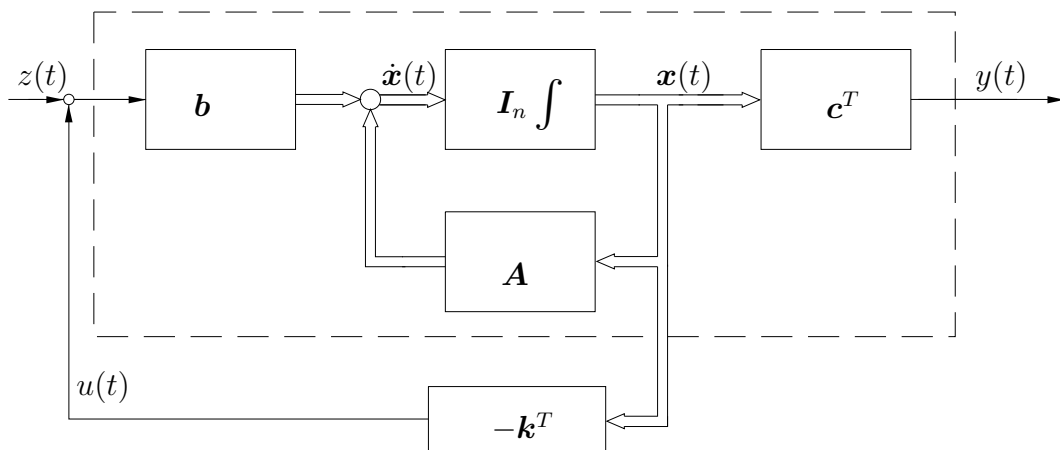


Bild 4.2: Blockschaltbild einer Zustandsregelung

Durch Einsetzen des Regelgesetzes in die Zustandsgleichung (4.2a) der Regelstrecke wird das Zustandsmodell des Regelkreises gewonnen zu

$$\dot{\mathbf{x}}(t) = [\mathbf{A} - \mathbf{b}\mathbf{k}^T] \mathbf{x}(t) + \mathbf{b}z(t) \quad (4.4a)$$

$$y(t) = \mathbf{c}^T \mathbf{x}(t) \quad (4.4b)$$

mit der charakteristischen Gleichung des geschlossenen Regelkreises

$$C(\lambda) = |\lambda \mathbf{I} - (\mathbf{A} - \mathbf{b}\mathbf{k}^T)| = 0 \quad . \quad (4.5)$$

Die Zustandsregelung hat den Vorteil, daß (im Falle der vollständigen Steuerbarkeit des Systems) jede Eigenwertkonfiguration erzeugt werden kann. Umgekehrt werden durch eine Vorgabe der Eigenwerte des geschlossenen Regelkreises die Koeffizienten einer Zustandsrückführung vollständig festgelegt.

Seien die gewünschten Eigenwerte λ_{g_i} , $i = 1, 2, \dots, n$, dann ergibt sich das charakteristische Polynom $C_g(\lambda)$ des geschlossenen Regelkreises zu

$$C_g(\lambda) = (\lambda - \lambda_{g_1})(\lambda - \lambda_{g_2}) \cdots (\lambda - \lambda_{g_n}) \quad (4.6)$$

$$= \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 \quad . \quad (4.7)$$

Wenn das System (\mathbf{A}, \mathbf{b}) vollständig steuerbar ist, existiert eine Ähnlichkeitstransformation $\mathbf{x} = \mathbf{T}_R \mathbf{x}_R$ mit \mathbf{T}_R nach Gleichung (3.64), die das System (3.1) auf Regulationsnormalform transformiert:

$$\left. \begin{aligned} \dot{\mathbf{x}}_R(t) &= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix} \mathbf{x}_R(t) + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u(t) \quad , \\ y(t) &= [b_0 \quad b_1 \quad b_2 \quad \cdots \quad b_{n-1}] \mathbf{x}_R(t) . \end{aligned} \right\} \quad (4.8)$$

Die Koeffizienten des charakteristischen Polynoms können in der letzten Zeile der transformierten Systemmatrix \mathbf{A}_R direkt abgelesen werden:

$$\det(\lambda \mathbf{I} - \mathbf{A}_R) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 \quad . \quad (4.9)$$

Die Rückführung des Zustandsvektors \mathbf{x}_R über den Zeilenvektor $\tilde{\mathbf{k}}$

$$u = -\tilde{\mathbf{k}}^T \mathbf{x}_R = -[\tilde{k}_1 \quad \tilde{k}_2 \quad \dots \quad \tilde{k}_n] \mathbf{x}_R \quad (4.10)$$

liefert

$$\tilde{\mathbf{A}}_g = \mathbf{A}_R - \mathbf{b}_R \tilde{\mathbf{k}}^T \quad (4.11)$$

$$= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -(a_0 + \tilde{k}_1) & -(a_1 + \tilde{k}_2) & \cdots & \cdots & -(a_{n-1} + \tilde{k}_n) \end{bmatrix} \quad (4.12)$$

als Systemmatrix des geschlossenen Regelkreises.

Die letzte der Zeile der Matrix \mathbf{A}_g enthält jetzt die Koeffizienten der charakteristischen Gleichung des rückgeführten Systems. Ein Koeffizientenvergleich der Gleichung (4.7) mit der letzten Zeile der Matrix (4.12) ergibt die gesuchten Parameter des Rückführvektors $\tilde{\mathbf{k}}$ zu

$$\tilde{k}_i = \alpha_{i-1} - a_{i-1} \quad \text{für } i = 1, 2, \dots, n. \quad (4.13)$$

Den Rückführvektor für das Ausgangssystem erhält man mit $\mathbf{x}_R = \mathbf{T}_R^{-1} \mathbf{x}$ aus

$$u = -\tilde{\mathbf{k}}^T \mathbf{x}_R = -\tilde{\mathbf{k}}^T \mathbf{T}_R^{-1} \mathbf{x} = -\mathbf{k}^T \mathbf{x} \quad (4.14)$$

zu

$$\mathbf{k}^T = \left[\alpha_0 - a_0 : \alpha_1 - a_1 : \dots : \alpha_{n-1} - a_{n-1} \right] \mathbf{T}_R^{-1}. \quad (4.15)$$

Dem Vorteil der beliebigen Festlegung der Eigenwerte im geschlossenen Regelkreis steht als Nachteil gegenüber, daß durch einen zweiten Synthese- und dann Realisierungsschritt die Systemzustände aus den verfügbaren Ausgangssignalen beobachtet (geschätzt) werden müssen, da per Definition die Systemzustände interne Variablen und normalerweise nicht alle als Ausgangssignal verfügbar sind.

4.3 Entwurf eines Zustandsbeobachters

Für die Realisierung einer Zustandsregelung oder -rückführung werden die Zustandsvariablen $\mathbf{x}(t) = [x_1, \dots, x_n]^T$ explizit benötigt. Dies kann dadurch erreicht werden, daß für jede Variable $x_i(t)$ ein Meßgerät eingesetzt wird, womit aber tatsächlich $\mathbf{y}(t) = \mathbf{x}(t)$ eine vektorielle Ausgangsgröße wird.

Es kann gezeigt werden, daß bei einem vollständig beobachtbaren System der Zustand $\mathbf{x}(t)$ auch indirekt aus einer Ausgangsgröße $y(t)$ und der auf das System einwirkenden Stellgröße $u(t)$ geschätzt werden kann. Solche Zustandsschätzer heißen auch „Beobachter“ (vgl. Bild 4.3).

Definition 4.1 (Luenberger-Beobachter)

Ein System mit dem n -dimensionalen Zustandsvektor $\hat{\mathbf{x}}(t) \in \mathbb{R}^n$

$$\dot{\hat{\mathbf{x}}} = \mathbf{F} \hat{\mathbf{x}}(t) + \mathbf{l}y(t) + \mathbf{b}u(t) \quad (4.16)$$

heißt ein Beobachter der Ordnung n für das System

$$\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{b}u(t) \quad (4.17a)$$

$$y(t) = \mathbf{c}^T \mathbf{x}(t) \quad , \quad (4.17b)$$

wenn aus

$$\hat{\mathbf{x}}(t_0) = \mathbf{x}(t_0) \quad (4.18)$$

für alle $u(t)$ folgt:

$$\hat{\mathbf{x}}(t) = \mathbf{x}(t) \quad , \quad t \geq t_0 \quad \square \quad (4.19)$$

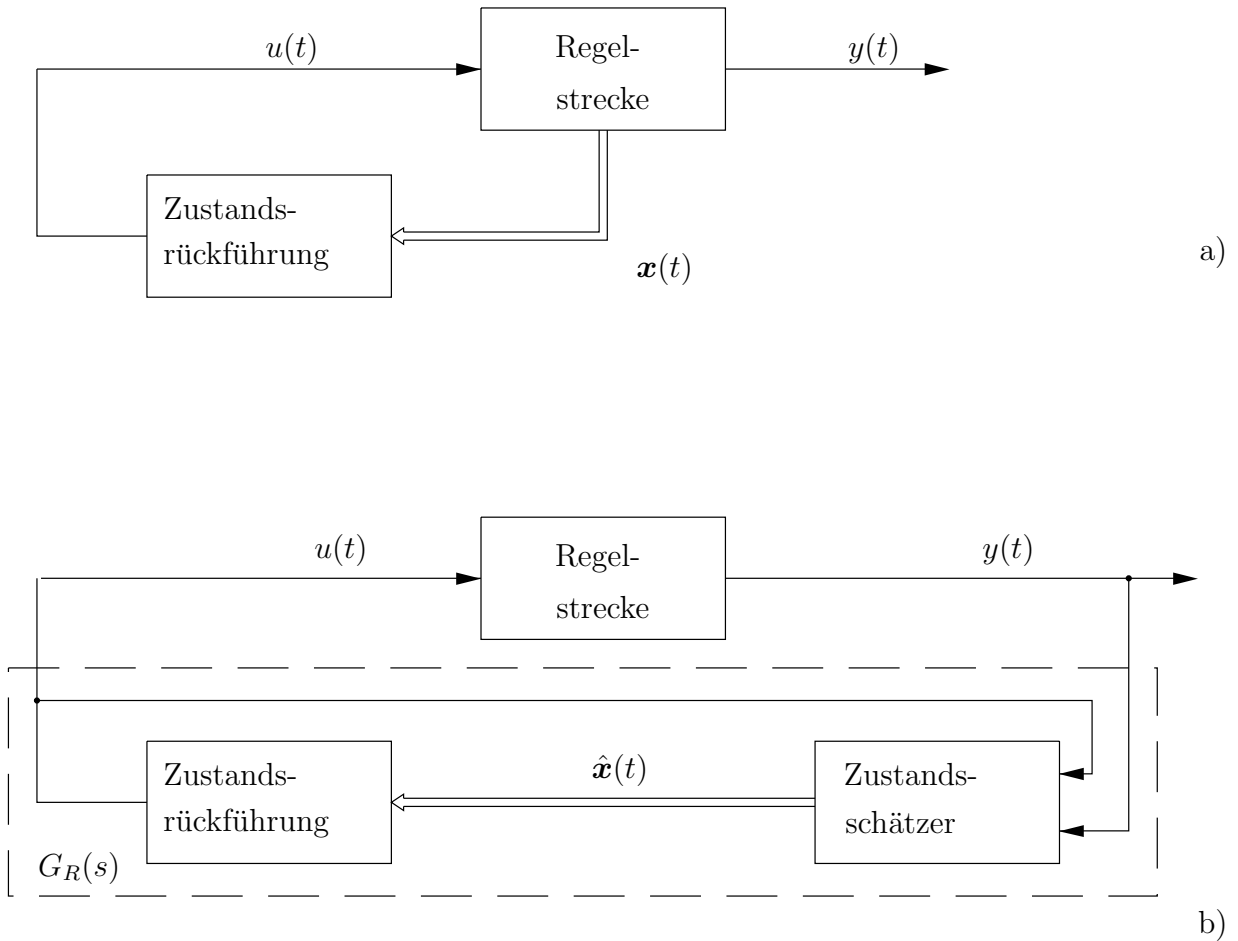


Bild 4.3: a) Regelstrecke mit Zustandsrückführung

b) System wie unter a) aber nun mit dem durch einen Beobachter geschätzten Zustandsvektor $\hat{x}(t)$

Das System Gl. (4.16) ist dann und nur dann ein Beobachter (vollständiger) Ordnung zu dem System nach Gl. (4.17a), wenn für seine Systemmatrix gilt:

$$\mathbf{F} = \mathbf{A} - \mathbf{l}\mathbf{c}^T \quad (4.20)$$

wobei \mathbf{l} ein Verstärkungsvektor der Dimension n derart ist, daß die Systemmatrix \mathbf{F} des Beobachters nur stabile Eigenwerte hat.

Bild 4.4 zeigt die Struktur des (vollständigen) Beobachters der Ordnung n .

Der Beobachter vollständiger Ordnung schätzt (unnötigerweise) auch Zustände, die direkt meßbar sind. Es kann deshalb ausgehend von der Beobachtungsnormalform ein Beobachter der Ordnung $(n - 1)$ konstruiert werden, dessen Blockschaltbild und Systemgleichungen aber weniger übersichtlich sind. Häufig werden zur Lösung einer regelungstechnischen Aufgabe auch nicht alle Zustände benötigt — z.B., wenn nur einige Eigenwerte verändert, aber andere unverändert übernommen werden sollen —, dann genügt eine Teilzustandsschätzung und Teilzustandsrückführung.

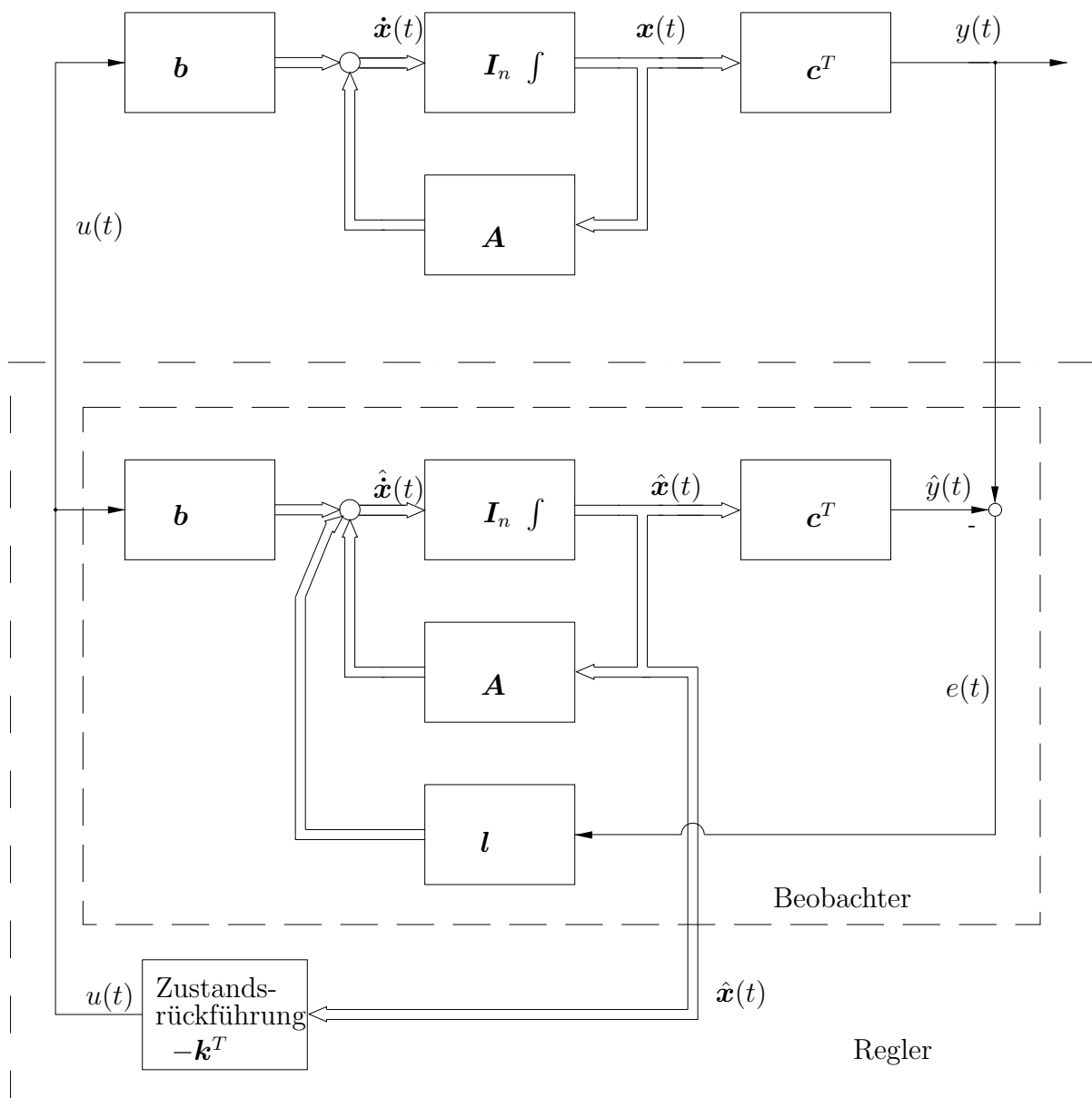


Bild 4.4: Struktur eines linearen Regelgesetzes bestehend aus Zustandsbeobachter und Rückführung des (geschätzten) Systemzustandes

Zusammen mit einem Zustandsbeobachter ergibt die Regelungsstruktur (Bild 4.5) einen dynamischen Regler, denn der Zustandsschätzer enthält die Dynamik der Regelstrecke.

4.3.1 Konvergenz des Beobachters

Betrachtet man die Gleichungen (4.20), so ist in diesen Gleichungen lediglich der Verstärkungsvektor l , die *Beobachterrückführung* nicht bekannt. Um die Bedeutung dieser Beob-

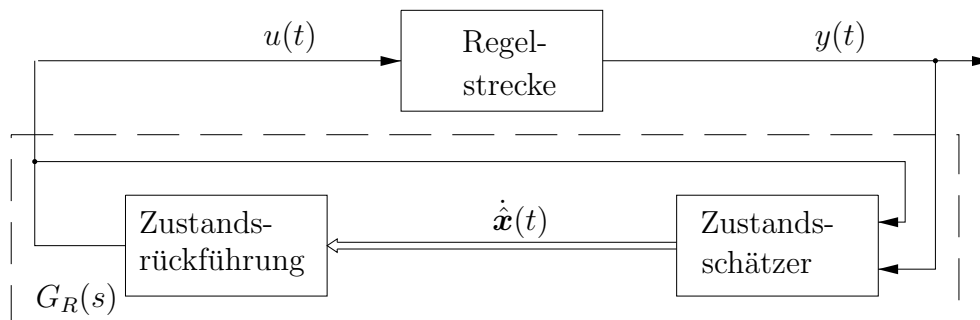


Bild 4.5: Zustandsregler und Zustandsschätzer

achterrückführung zu erkennen, wird der Beobachtungsfehler

$$\tilde{\mathbf{e}}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t) \quad (4.21)$$

eingeführt. Mit Hilfe der Gleichungen für das Modell (4.17a) und den Beobachter (4.16) erhält man für den Fehler die Differentialgleichung

$$\dot{\tilde{\mathbf{e}}}(t) = \dot{\mathbf{x}}(t) - \dot{\hat{\mathbf{x}}}(t) \quad (4.22)$$

$$= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) - \mathbf{A}\hat{\mathbf{x}}(t) - \mathbf{b}u(t) - \mathbf{l}\mathbf{c}^T(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) \quad (4.23)$$

$$= [\mathbf{A} - \mathbf{l}\mathbf{c}^T][\mathbf{x}(t) - \hat{\mathbf{x}}(t)] \quad (4.24)$$

$$= [\mathbf{A} - \mathbf{l}\mathbf{c}^T] \tilde{\mathbf{e}}(t) . \quad (4.25)$$

Aus Gleichung (4.25) folgt unmittelbar der folgende Satz:

Satz 4.1 *Konvergenz der Beobachtungsfehlers*

Für den Beobachtungsfehler

$$\tilde{\mathbf{e}}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$$

eines vollständigen Zustandsbeobachters gilt

$$\lim_{t \rightarrow \infty} |\tilde{\mathbf{e}}(t)| = 0$$

für beliebige Anfangszustände des Systems und des Beobachters genau dann, wenn alle Eigenwerte der Matrix $(\mathbf{A} - \mathbf{l}\mathbf{c}^T)$ negativen Realteil haben. \square

4.3.2 Berechnung der Beobachterrückführung

Die Matrix $[\mathbf{A} - \mathbf{l}\mathbf{c}^T]$ in Gleichung (4.25) hat dieselben Eigenwerte wie die transponierte Matrix

$$[\mathbf{A} - \mathbf{l}\mathbf{c}^T]^T = \mathbf{A}^T - \mathbf{c}\mathbf{l}^T . \quad (4.26)$$

Die Festlegung der Eigenwerte der Matrix $[\mathbf{A} - \mathbf{l}\mathbf{c}^T]$ mit Hilfe des Vektors \mathbf{l} kann auf die Berechnung einer Zustandsrückführung überführt werden, da die transponierte Matrix (4.26) die Systemmatrix des dualen Systems

$$\dot{\tilde{\mathbf{x}}} = \mathbf{A}^T \tilde{\mathbf{x}}(t) + \mathbf{c}\tilde{u}(t) \quad (4.27)$$

mit der Zustandsrückführung

$$\tilde{u}(t) = -\mathbf{l}^T \tilde{\mathbf{x}}(t) \quad (4.28)$$

ist. Aus dieser Analogie zum Entwurf einer Zustandsrückführung ergeben sich sofort die folgenden Konsequenzen:

- Die Eigenwerte der Matrix $(\mathbf{A} - \mathbf{l}\mathbf{c}^T)$ können nur dann beliebig festgelegt werden, wenn das System (\mathbf{A}, \mathbf{c}) vollständig beobachtbar ist.
- Damit der Beobachtungsfehler schneller abklingt als das Übergangsverhalten des zu beobachtenden Systems, müssen die Eigenwerte der Matrix $(\mathbf{A} - \mathbf{l}\mathbf{c}^T)$ — auch Beobachtereigenwerte genannt — links der dominierenden Eigenwerte von \mathbf{A} bzw. von $(\mathbf{A} - \mathbf{b}\mathbf{k}^T)$ liegen.

Besonders einfach kann die Beobachterrückführung \mathbf{l} dann berechnet werden, wenn das Modell der Regelstrecke in Beobachtungsnormalform (3.69) vorliegt. Der Vektor

$$\mathbf{l}_B = [l_1 \ l_2 \ \dots \ l_n]^T$$

muß dann so gewählt werden, daß die Matrix

$$\mathbf{A}_B^T - \mathbf{c}_B \mathbf{l}_B^T = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -(a_0 + l_1) & -(a_1 + l_2) & \dots & \dots & -(a_{n-1} + l_n) \end{bmatrix} \quad (4.29)$$

die gewünschten Beobachtereigenwerte λ_{B_i} , $i = 1, 2, \dots, n$ besitzt. In Analogie zu Gleichung (4.15) ergibt sich die Beobachterrückführung zu

$$\mathbf{l}_B^T = [a_{B_0} \ a_{B_1} \ \dots \ a_{B_n}] - [a_0 \ a_1 \ \dots \ a_{n-1}] \quad (4.30)$$

mit a_{B_i} , $i = 0, 1, 2, \dots, n$ den Koeffizienten des charakteristischen Polynoms der Matrix (4.29).

4.3.3 Separationstheorem

Die Anordnung des Zustandsbeobachters im geschlossenen Regelkreis in Bild 4.4 zeigt, daß anstelle des Zustandsvektors $\mathbf{x}(t)$ der geschätzte Vektor $\hat{\mathbf{x}}(t)$ über \mathbf{k} auf $u(t)$ zurückgeführt wird. Zur Aufstellung der Zustandsraumdarstellung des Gesamtsystems können folgende Zustandsgleichungen für die Regelstrecke und den Beobachter anhand von Bild 4.4 direkt angegeben werden:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{b}\mathbf{k}^T \hat{\mathbf{x}} \quad (4.31)$$

und

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{l}\mathbf{c}^T[\mathbf{x} - \hat{\mathbf{x}}] - \mathbf{b}\mathbf{k}^T \hat{\mathbf{x}}. \quad (4.32)$$

Ersetzt man in Gl. (4.31) $\hat{\mathbf{x}}$ durch $\mathbf{x} - \mathbf{e}$, so erhält man

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{b}\mathbf{k}^T \mathbf{x} + \mathbf{b}\mathbf{k}^T \mathbf{e}. \quad (4.33)$$

Damit können die Gleichungen (4.33) und (4.25) zu einer einzigen Gleichung

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\mathbf{e}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} - \mathbf{b}\mathbf{k}^T & \mathbf{b}\mathbf{k}^T \\ \mathbf{0} & \mathbf{A} - \mathbf{l}\mathbf{c}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{e}(t) \end{bmatrix} \quad (4.34)$$

zusammengefaßt werden.

Da die Systemmatrix des Gesamtsystems eine Blockdreiecksmatrix ist, setzen sich ihre Eigenwerte aus denen der Matrizen $\mathbf{A} - \mathbf{b}\mathbf{k}^T$ und $\mathbf{A} - \mathbf{l}\mathbf{c}^T$ zusammen und man erhält folgendes Ergebnis:

Satz 4.2 Separationstheorem

Sofern das offene System $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ vollständig steuer- und beobachtbar ist, können die n Eigenwerte der charakteristischen Gleichung des geschlossenen Regelkreises (ohne Beobachter) und die n Beobachtereigenwerte separat vorgegeben werden. \square

4.4 Zusammenfassung

- Mit Hilfe einer Zustandsrückführung kann die Eigendynamik der Regelstrecke beliebig verändert werden, wenn das System (\mathbf{A}, \mathbf{b}) vollständig steuerbar ist.
- Der Zustandsregler ist ein rein proportionaler Regler und garantiert bei sprunghaftigen Führungs- und Störsignalen keine verschwindende bleibende Regelabweichung.
- Zur Realisierung einer Zustandsrückführung müssen alle Zustandsgrößen $x_i(t)$, $i = 1, 2, \dots, n$ zur Verfügung stehen.

- In der Regel werden nicht alle Zustandsgrößen meßbar sein, so daß der Einsatz eines Beobachters erforderlich ist, der den Zustand $\mathbf{x}(t)$ aus dem Verlauf der Eingangsgröße $u(t)$ und der Ausgangsgröße $y(t)$ rekonstruiert.
- Im geschlossenen Regelkreis können die Dynamik der Regelstrecke und des Beobachters unabhängig voneinander mit Hilfe der Vektoren \mathbf{k} und \mathbf{l} eingestellt werden.
- Die Eigenwerte des Beobachters sollten in der linken komplexen Ebene deutlich links von den Eigenwerten der Regelstrecke bzw. des geschlossenen Kreises liegen. Liegen die Beobachtereigenwerte allerdings zu weit links, so wird das Meßrauschen im Beobachter stark verstärkt.

Der Entwurf einer Zustandsrückführung, die mit Hilfe eines Beobachters realisiert werden soll, kann dann wie folgt zusammengefaßt werden:

Entwurf einer Zustandsrückführung und eines Beobachters

Voraussetzung: Die Regelstrecke $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ ist vollständig steuer- und beobachtbar.

1. Überprüfung, daß die Regelstrecke vollständig steuer- und beobachtbar ist.
2. Entwurf einer Zustandsrückführung

$$u(t) = -\mathbf{k}^T \mathbf{x}(t) ,$$

die die an den Regelkreis gestellten Güteforderungen erfüllt.

3. Anhand der Eigenwerte der Matrix $\mathbf{A} - \mathbf{b}\mathbf{k}^T$ werden die Beobachtereigenwerte festgelegt.
4. Bestimmung des Vektors \mathbf{l} der Beobachterrückführung als Zustandsrückführung des dualen Systems $(\mathbf{A}_B^T, \mathbf{c}_B)$.
5. Bewertung des Verhaltens des geschlossenen Regelkreises anhand von Simulationsuntersuchungen.

Ergebnis: Zustandsrückführung \mathbf{k} , Beobachter mit Rückführvektor \mathbf{l} .

A Mathematische Grundlagen

A.1 Definition einer Matrix

Unter einer $m \times n$ Matrix (gesprochen "m Kreuz n Matrix") versteht man eine rechteckig angeordnete Menge von Zahlen mit m Zeilen und n Spalten. Die Angabe $m \times n$ bezeichnet die Ordnung bzw. den Typ einer Matrix. Zur Kennzeichnung von Matrizen werden fette Großbuchstaben verwendet:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \quad (\text{A.1})$$

Die Zahl a_{ij} ist das Element der i -ten Zeile und der j -ten Spalte der Matrix \mathbf{A} . Der erste Index i wird als Zeilenindex und der zweite Index j als Spaltenindex bezeichnet.

A.2 Transponierte einer Matrix

Eine oft verwendete Operation ist die Transposition, die darin besteht, daß die Zeilen und Spalten einer Matrix vertauscht werden. Sei \mathbf{A} eine $n \times m$ -Matrix so entsteht durch Transposition eine $m \times n$ -Matrix. Um die Transponierte einer reellen Matrix zu bezeichnen wird ein hochgestelltes "T" verwendet. Es gilt folgende wichtige Regel:

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T. \quad (\text{A.2})$$

Symmetrische Matrizen sind durch

$$\mathbf{A}^T = \mathbf{A} \quad (\text{A.3})$$

definiert.

A.3 Quadratische Matrix

Als *quadratisch* bezeichnet man Matrizen, die genauso viele Zeilen wie Spalten besitzen. Sie haben die Dimension $n \times n$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \quad (\text{A.4})$$

Die Elemente $a_{11}, a_{22}, \dots, a_{nn}$ bilden die Hauptdiagonale der quadratischen Matrix \mathbf{A} . Die Summe der Hauptdiagonalelemente bezeichnet man als Spur:

$$\text{Spur } \mathbf{A} = a_{11} + a_{22} + \dots + a_{nn}. \quad (\text{A.5})$$

Eine quadratische $n \times n$ -Matrix \mathbf{A} wird eine *Diagonalmatrix* genannt, wenn alle Elemente außerhalb der Hauptdiagonalen Null sind. Zur Definition von Diagonalmatrizen wird oft die Schreibweise

$$\text{diag}(a_1, a_2, \dots, a_n) = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} \quad (\text{A.6})$$

verwendet.

Ein wichtiges Beispiel für eine Diagonalmatrix ist die *Einheitsmatrix*, die zu

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (\text{A.7})$$

definiert ist.

A.4 Vektoren

Vektoren sind Matrizen, die nur aus einer Zeile oder einer Spalte bestehen. $1 \times n$ -Matrizen werden als Zeilenvektoren, $n \times 1$ -Matrizen werden als Spaltenvektoren bezeichnet. Im Kontext der Matrizenrechnung werden 1×1 -Matrizen als Skalare bezeichnet. Die Zahl 5 kann beispielsweise als 1×1 -Matrix bzw. Skalar aufgefaßt werden. Zur Bezeichnung von Vektoren werden fettgedruckte Kleinbuchstaben verwendet. Aus einen Spaltenvektor

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad (\text{A.8})$$

erhält man durch Transponieren einen Zeilenvektoren

$$\mathbf{b} = \mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_n] . \quad (\text{A.9})$$

Beim *Nullvektor* sind sämtliche Elemente Null:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \quad (\text{A.10})$$

Ein *Einheitsvektor* unterscheidet sich vom Nullvektor dadurch, daß eine Komponente Eins ist. Für den i -ten Einheitsvektor gilt:

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad - \quad i\text{-tes Element.} \quad (\text{A.11})$$

A.5 Determinante einer Matrix

Eine Determinante ist eine charakteristische Zahl einer quadratischen Matrix. Eine Determinante wird entweder durch zwei senkrechte Striche $|\mathbf{A}|$ oder $\det \mathbf{A}$ gekennzeichnet.

A.5.1 Determinanten von 2×2 -Matrizen

Die Determinante einer 2×2 -Matrix berechnet sich, indem das Produkt der Nebendiagonalelemente von dem Produkt der Hauptdiagonalelemente abgezogen wird:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{21} \cdot a_{12} . \quad (\text{A.12})$$

A.5.2 Determinanten von 3×3 -Matrizen mit Hilfe der Sarrusschen Regel

Betrachtet wird eine erweiterte Matrix $\tilde{\mathbf{A}}$ der Matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad (\text{A.13})$$

die man erhält, indem die Matrix \mathbf{A} um die ersten beiden Spalten erweitert wird:

$$\tilde{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \vdots & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & \vdots & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & \vdots & a_{31} & a_{32} \end{bmatrix} . \quad (\text{A.14})$$

Die Determinante von \mathbf{A} berechnet sich als Differenz der aufsummierten Produkte der Hauptdiagonalelemente und der aufsummierten Produkte der Nebendiagonalelemente von $\tilde{\mathbf{A}}$:

$$\tilde{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \vdots & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & \vdots & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & \vdots & a_{31} & a_{32} \end{bmatrix}, \quad (\text{A.15})$$

$$|\mathbf{A}| = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - (a_{13} \cdot a_{22} \cdot a_{31} + a_{11} \cdot a_{23} \cdot a_{32} + a_{12} \cdot a_{21} \cdot a_{33}) . \quad (\text{A.16})$$

A.5.3 Determinanten beliebiger Matrizen

Unter der Voraussetzung, daß die Determinante einer beliebigen $(n-1) \times (n-1)$ -Matrix bekannt ist, läßt sich die Determinante einer beliebigen $n \times n$ -Matrix wie folgt berechnen
Laplacescher Entwicklungssatz :

$$\det \mathbf{A} = \sum_{j=1}^n a_{ij} \mathbf{A}_{ij} , \quad (\text{A.17})$$

wobei i eine beliebige fest gewählte Zeilennummer ist¹. Ferner bezeichnet \mathbf{A}_{ij} die sogenannte *Adjunkte*, auch *Kofaktor* genannt, zu dem Element a_{ij} . Die Adjunkte \mathbf{A}_{ij} besteht aus derjenigen Unterdeterminante von \mathbf{A} , die durch Streichung der i -ten Zeile und j -ten Spalte entsteht, multipliziert mit dem Vorzeichen $(-1)^{i+j}$.

A.5.4 Rechenregeln für Determinanten

Vertauschen von Zeilen und Spalten

Eine Determinante ändert sich nicht, wenn man Zeilen und Spalten miteinander vertauscht. Damit gilt auch $|\mathbf{A}^T| = |\mathbf{A}|$.

Vertauschen von Zeilen oder Spalten

Eine Determinante ändert ihr Vorzeichen, wenn man zwei Zeilen oder zwei Spalten miteinander vertauscht.

Addieren des Vielfachen einer Zeile oder Spalte

Eine Determinante ändert sich nicht, wenn man zu einer Zeile (oder Spalte) das Vielfache einer anderen Zeile (oder Spalte) addiert.

Multiplikation mit einer Zahl

Eine Determinante multipliziert man mit einer Zahl, indem man eine fest gewählte Zeile (oder Spalte) mit dieser Zahl multipliziert.

Determinante eines Produktes

Sind \mathbf{A} und \mathbf{B} zwei beliebige quadratische Matrizen, dann gilt für die Determinante des Produkts:

$$|\mathbf{A} \cdot \mathbf{B}| = |\mathbf{A}| \cdot |\mathbf{B}| = |\mathbf{B} \cdot \mathbf{A}| . \quad (\text{A.18})$$

A.6 Inverse Matrix

Unter einer inversen Matrix zu einer $n \times n$ -Matrix \mathbf{A} versteht man eine $n \times n$ -Matrix \mathbf{B} mit

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I} . \quad (\text{A.19})$$

¹ Eine analoge Aussage gilt für die Entwicklung nach der k -ten Spalte.

Die inverse Matrix existiert genau dann, wenn $\det \mathbf{A} \neq 0$ gilt. In diesem Fall ist \mathbf{B} eindeutig bestimmt und wird mit \mathbf{A}^{-1} bezeichnet. Sie kann mit Hilfe der Formel

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{\det \mathbf{A}} \quad (\text{A.20})$$

berechnet werden. Hierbei ist die *adjungierte* Matrix ($\text{adj } \mathbf{A}$) die Transponierte der Matrix die entsteht, wenn man die Elemente a_{ij} durch ihre Kofaktoren (Adjunkten) \mathbf{A}_{ij} ersetzt. Der Kofaktor \mathbf{A}_{ij} ist dabei die Unterdeterminante von \mathbf{A} , die durch Streichung der i -ten Zeile und j -ten Spalte entsteht, multipliziert mit dem Vorzeichen $(-1)^{i+j}$.

Für zwei reguläre quadratische Matrizen \mathbf{A} und \mathbf{B} gilt

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1} . \quad (\text{A.21})$$

A.7 Eigenwerte und Eigenvektoren

Ein Skalar λ heißt *Eigenwert* einer $n \times n$ -Matrix \mathbf{A} , wenn es einen Vektor $\mathbf{x} \neq \mathbf{0}$ gibt, so daß die Gleichung

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (\text{A.22})$$

erfüllt ist. Die (A.22) wird auch als *Eigenwertgleichung* bezeichnet und der Vektor \mathbf{x} heißt der zum Eigenwert λ gehörige *Eigenvektor*.

Eine Umformung der Gleichung (A.22) liefert

$$[\lambda\mathbf{I} - \mathbf{A}]\mathbf{x} = \mathbf{0} . \quad (\text{A.23})$$

Diese Gleichung hat nur dann eine nichttriviale Lösung $\mathbf{x} \neq \mathbf{0}$, wenn die folgende Determinante verschwindet:

$$|\lambda\mathbf{I} - \mathbf{A}| = 0 . \quad (\text{A.24})$$

Die Ermittlung der Determinanten der Gleichung (A.24) führt zu einer Polynomgleichung n -ter Ordnung in λ

$$\begin{aligned} P(\lambda) &= |\lambda\mathbf{I} - \mathbf{A}| = 0 \\ &= \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0 . \end{aligned} \quad (\text{A.25})$$

Diese Gleichung nennt man *charakteristische Gleichung* der Matrix \mathbf{A} und das Polynom das *charakteristische Polynom*. Die n Nullstellen (Wurzeln) λ_i sind die *Eigenwerte* der Matrix \mathbf{A} und können einfach oder mehrfach, reell oder konjugiert komplex sein. Ist \mathbf{A} symmetrisch, so sind alle Eigenwerte reell. Die Eigenwerte von Dreiecksmatrizen sind gleich den Hauptdiagonalelementen.

Wenn die n Eigenwerte λ_i der Matrix \mathbf{A} verschieden sind, dann gehört zu jedem Eigenwert λ_i ein Eigenvektor \mathbf{x}_i . Diese Eigenvektoren sind die Lösungen der Gleichung

$$[\lambda_i\mathbf{I} - \mathbf{A}]\mathbf{x}_i = \mathbf{0} , \quad i = 1, 2, \dots, n . \quad (\text{A.26})$$

Wenn \mathbf{A} n unterschiedliche Eigenwerte besitzt, dann sind die zugehörigen Eigenvektoren linear unabhängig.

Nach dem Fundamentalsatz der Algebra kann das charakteristische Polynom wie folgt faktorisiert werden:

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) . \quad (\text{A.27})$$

Das Produkt der Eigenwerte einer quadratischen Matrix ist gleich der Determinanten der Matrix:

$$\lambda_1 \lambda_2 \cdots \lambda_n = |\mathbf{A}| , \quad (\text{A.28})$$

und die Summe der Eigenwerte ist gleich der *Spur* der Matrix:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{Spur } \mathbf{A} = \sum_{i=1}^n a_{ii} . \quad (\text{A.29})$$

A.7.1 Cayley–Hamilton–Theorem

Das Cayley–Hamilton–Theorem besagt, daß jede quadratische Matrix ihre eigene charakteristische Gleichung erfüllt. D.h., ersetzt man in dem charakteristischen Polynom (A.25) λ durch \mathbf{A} so gilt

$$P(\mathbf{A}) = \mathbf{A}^n + a_{n-1} \mathbf{A}^{n-1} + \cdots + a_1 \mathbf{A} + a_0 \mathbf{I} = \mathbf{0} . \quad (\text{A.30})$$

Damit lassen sich die n -te Potenz \mathbf{A}^n und alle höheren Potenzen von \mathbf{A} als Linearkombinationen der Potenzen $\mathbf{A}^0, \mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^{n-1}$ darstellen. Dasselbe gilt auch für die inverse Matrix \mathbf{A}^{-1} , sofern diese existiert.

A.7.2 Ähnlichkeitstransformation

Durch eine Ähnlichkeitstransformation der Matrix \mathbf{A} mit der regulären Matrix \mathbf{T} entsteht die Matrix

$$\tilde{\mathbf{A}} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} . \quad (\text{A.31})$$

Die Matrizen \mathbf{A} und $\tilde{\mathbf{A}}$ haben dieselben Eigenwerte. Sind die Eigenvektoren \mathbf{x}_i von \mathbf{A} linear unabhängig, so kann man mit ihnen die Transformationsmatrix

$$\mathbf{T} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \quad (\text{A.32})$$

bilden. Aus der Ähnlichkeitstransformation mit dieser Matrix erhält man dann die Diagonalmatrix

$$\mathbf{T}^{-1} \mathbf{A} \mathbf{T} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} , \quad (\text{A.33})$$

deren Hauptdiagonalelemente gerade die Eigenwerte der Matrix \mathbf{A} sind. Die Bedingung, daß alle Eigenvektoren linear unabhängig sind, ist stets erfüllt, wenn alle Eigenwerte nur einmal auftreten.

A.8 Rang einer Matrix

Bei der Analyse und Synthese linearer Systeme im Zustandsraum sind viele Systemeigenschaften mit dem Rang von entsprechenden Matrizen verknüpft. Ganz allgemein versteht man unter dem Rang einer Matrix:

Definition A.1

Der *Rang* einer Matrix ist identisch mit der Ordnung der größten von Null verschiedenen Unterdeterminante der Matrix.

Zur übersichtlichen Darstellung einer Unterdeterminante wird folgende Beschreibung eingeführt:

Definition A.2

Zu einer beliebigen $n \times m$ Matrix \mathbf{A} bezeichnet

$$\mathbf{A}_{j_1, j_2, \dots, j_r}^{i_1, i_2, \dots, i_r} \quad (\text{A.34})$$

eine Unterdeterminante der Ordnung r , nämlich die Determinante der $r \times r$ Submatrix, die aus \mathbf{A} hervorgeht, indem alle Zeilen außer i_1, i_2, \dots, i_r und alle Spalten außer j_1, j_2, \dots, j_r gestrichen werden.

Der Rang einer Matrix ist dabei auch identisch mit der maximalen Anzahl der linear unabhängigen Zeilen bzw. Spalten einer Matrix. Besitzt eine $n \times n$ Matrix den vollen Rang n , so wird sie *regulär* genannt. Entsprechend spricht man von einer zeilen- oder spaltenregulären Matrix, wenn die $n \times m$ Matrix \mathbf{A} den Rang $r = n$ bzw. $r = m$ hat.

A.8.1 Rang von Matrizenprodukten

Viele Ergebnisse und Zusammenhänge beruhen auf der exakten Vorhersage des Ranges eines Matrizenproduktes \mathbf{AB} . Sind die Rangzahlen r_A und r_B der beiden Faktoren \mathbf{A} und \mathbf{B} bekannt, so sind allerdings nur in bestimmten Fällen genaue und verbindliche Aussagen möglich.

Satz A.1

Ist \mathbf{A} eine $n \times m$ Matrix vom Range r und \mathbf{B} regulär und verkettbar mit \mathbf{A} , so hat auch die Produktmatrix \mathbf{AB} bzw. \mathbf{BA} den gleichen Rang r .

Sind beide Faktoren eines Matrizenproduktes nicht regulär, d.h. singulär, so erhält man solch eindeutige Aussagen nur noch in Sonderfällen.

Satz A.2

Das Produkt einer spaltenregulären $m \times r$ Matrix \mathbf{A} mit einer zeilenregulären $r \times n$ Matrix \mathbf{B} , die $m \times n$ Matrix $\mathbf{C} = \mathbf{AB}$, hat den Rang r .

Der Beweis zu dem folgenden Satz zeigt einen Weg auf, wie man weitere Fälle selbst untersuchen kann.

Satz A.3

Das Produkt einer zeilenregulären $r \times m$ Matrix \mathbf{A} mit einer zeilenregulären $m \times n$ Matrix \mathbf{B} , die $r \times n$ Matrix $\mathbf{C} = \mathbf{AB}$, hat den Rang r .

Beweis:

Für jede zeilenreguläre $m \times n$ Matrix \mathbf{B} existiert eine reguläre $n \times n$ Matrix \mathbf{T} , so daß

$$\mathbf{BT} = [\tilde{\mathbf{B}} \mid \mathbf{0}] \tag{A.35}$$

gilt, wobei $\tilde{\mathbf{B}}$ eine reguläre $m \times m$ Matrix ist. Wendet man die gleiche Transformation auf das Produkt \mathbf{AB} an, so liefert dies

$$\mathbf{ABT} = \mathbf{A}[\tilde{\mathbf{B}} \mid \mathbf{0}] \tag{A.36}$$

$$= [\mathbf{A}\tilde{\mathbf{B}} \mid \mathbf{0}]. \tag{A.37}$$

Entsprechend Satz A.1 ändert die Multiplikation mit einer regulären Matrix den Rang einer Matrix nicht, so daß der Rang von \mathbf{AB} gleich dem Rang r von \mathbf{A} ist. \square

A.8.2 Term-Rang und generischer Rang

Die Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C}$ zur Beschreibung des dynamischen Verhaltens eines realen Systems zeichnen sich u.a. dadurch aus, daß häufig eine Reihe ihrer Elemente exakt Null sind. Der prozentuale Anteil dieser Nullelemente steigt i.a. mit der Größe der Matrizen deutlich an. Die Matrizen komplexerer Systeme sind daher oft sogenannte *schwach besetzte Matrizen* (engl. *sparse matrices*), deren Rangzahlen bereits durch die Struktur der Matrizen beschränkt werden. Diese obere Schranke ist für eine $n \times m$ Matrix mit $m < n$, die *keine* festen Nullelemente enthält, offenbar mit der kleineren Dimension m identisch.

Die Determinante einer $n \times n$ Matrix \mathbf{A} setzt sich aus vorzeichengewichteten Termen der Form

$$a_{1t_1} a_{2t_2} \cdots a_{nt_n} \tag{A.38}$$

zusammen, wobei (t_1, t_2, \dots, t_n) eine Permutation von $(1, 2, \dots, n)$ ist. Ein Term (A.38) ist dadurch gekennzeichnet, daß die n Elemente sowohl zu n verschiedenen Zeilen als auch zu n verschiedenen Spalten der Matrix gehören. Dementsprechend besteht eine Unterdeterminante der Ordnung r einer beliebigen $n \times m$ Matrix \mathbf{A} aus Termen

$$a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_r j_r}, \tag{A.39}$$

wobei die Elemente $a_{i_i j_i}$ aus verschiedenen Zeilen i_1, i_2, \dots, i_r und Spalten j_1, j_2, \dots, j_r auszuwählen sind.

Definition A.3

Der größte Wert k , für den ein Term

$$a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_k j_k} \neq 0 \tag{A.40}$$

existiert, wird als *Term-Rang* der $n \times m$ Matrix \mathbf{A} bezeichnet.

Existieren für eine Matrix mehrere Terme der Form (A.40), so kann der tatsächliche Rang der Matrix nur dann kleiner als der Term-Rang sein, wenn sich diese Determinanten-Terme exakt kompensieren. Können die von Null verschiedenen Elemente einer Matrix unabhängig voneinander variiert werden, so gibt es immer Zahlenrealisierungen bei denen Rang (im üblichen numerischen Sinn) und Term-Rang übereinstimmen. Die Voraussetzung, daß die von Null verschiedenen Elemente in den Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C}$ unabhängig voneinander variiert werden können, ist allerdings für viele technische Systeme nicht erfüllt. So treten beispielsweise in der Matrix \mathbf{A} neben festen Nullen auch feste Einsen auf, wenn eine Zustandsgröße (Geschwindigkeit) die zeitliche Ableitung einer anderen Zustandsgröße (Weg) ist. Darüber hinaus ist die Anzahl der physikalischen Parameter, wie z.B. Massen, Längen, Kapazitäten usw., oft kleiner als die Anzahl der von Null verschiedenen Elemente in den Matrizen \mathbf{A} , \mathbf{B} und \mathbf{C} , so daß gewisse Abhängigkeiten zwischen diesen Elementen vorliegen. Das kann dazu führen, daß der maximal mögliche Rang einer Matrix unter Berücksichtigung dieser Abhängigkeiten kleiner als ihr Term-Rang ist. Dieser maximal mögliche Rang einer Matrix (unter Berücksichtigung etwa vorhandener Abhängigkeiten) wird als *generischer Rang* bezeichnet, und ist offensichtlich immer kleiner oder gleich dem Term-Rang. Oft ist der Fall von besonderem Interesse, bei dem diese beiden Größen übereinstimmen. Wenn die von Null verschiedenen Elemente der Matrix unabhängig voneinander variiert werden können, ist dieser Fall der Übereinstimmung sicherlich gegeben. Sind der generische und der Term-Rang nicht identisch, so ist für größere Matrizen eine Bestimmung des generischen Ranges nur sehr umständlich (z.B. mit Hilfe einer symbolischen Determinantenberechnung) möglich. Im Gegensatz dazu, existieren für die Berechnung des Term-Ranges eine Reihe einfacher und zuverlässiger Verfahren.

In der Literatur wird häufig nicht zwischen generischen Rang und Term-Rang unterschieden, so daß dort für den Term-Rang einer Matrix auch die Bezeichnungen generischer und struktureller Rang verwendet werden.

Der Unterschied zwischen dem generischen und dem Term-Rang einer Matrix soll abschließend anhand eines Beispiels verdeutlicht werden.

Beispiel

Gegeben sei eine 3×3 Matrix

$$\mathbf{A} = \begin{bmatrix} \lambda_1 \lambda_3 & \lambda_5 & \lambda_3 \\ 0 & \lambda_3 & 0 \\ \lambda_2 \lambda_1 & \lambda_4 & \lambda_2 \end{bmatrix},$$

deren 7 von Null verschiedenen Elemente Funktionen der unbestimmten Parameter $\lambda_1, \lambda_2, \dots, \lambda_5$ sind. Gesucht werden der generische und der Term-Rang der Matrix.

Ein Term der Form (A.40) mit 3 Elementen kann leicht zu

$$a_{11}a_{22}a_{33} = \lambda_1 \lambda_3 \lambda_3 \lambda_2 \neq 0$$

gefunden werden. Die Matrix besitzt demnach einen Term-Rang von 3. Eine Berechnung der Determinante liefert allerdings:

$$\begin{aligned} |\mathbf{A}| &= \begin{vmatrix} \lambda_1 \lambda_3 & \lambda_5 & \lambda_3 \\ 0 & \lambda_3 & 0 \\ \lambda_2 \lambda_1 & \lambda_4 & \lambda_2 \end{vmatrix} = \lambda_1 \lambda_3 \begin{vmatrix} \lambda_3 & 0 \\ \lambda_4 & \lambda_2 \end{vmatrix} + \lambda_2 \lambda_1 \begin{vmatrix} \lambda_5 & \lambda_3 \\ \lambda_3 & 0 \end{vmatrix} \\ &= \lambda_1 \lambda_3 \lambda_3 \lambda_2 + \lambda_2 \lambda_1 \cdot (-\lambda_3 \lambda_3) = 0. \end{aligned}$$

Das bedeutet, der generische Rang dieser Matrix ist kleiner als der Term-Rang von 3, da die Determinante für beliebige λ_i -Werte verschwindet. Betrachtet man beispielsweise die Unterdeterminante \mathbf{A}_{12}^{12} , so ist diese für λ_i -Werte $\neq 0$ immer von Null verschieden, so daß die Matrix einen generischen Rang von 2 aufweist.

A.9 Orthogonale-Transformationen

Bei der numerisch stabilen Berechnung der Eigenwerte oder des Ranges einer Matrix spielen unitäre bzw. orthogonale Matrizen eine herausragende Rolle. Unter unitären Matrizen versteht man dabei:

Definition A.4

Eine komplexe Matrix \mathbf{U} heißt *unitär*, wenn

$$\mathbf{U}^H \mathbf{U} = \mathbf{I}, \tag{A.41}$$

ist, wobei \mathbf{U}^H die konjugiert komplexe, transponierte Matrix zu \mathbf{U} ist.

Aus Gl. (A.41) folgt, daß die Inverse einer unitären Matrix mit ihrer konjugiert komplex Transponierten identisch ist:

$$\mathbf{U}^{-1} = \mathbf{U}^H. \tag{A.42}$$

Die Inverse einer unitären Matrix kann daher ohne jede numerische Schwierigkeit sofort angegeben werden. Ist die unitäre Matrix \mathbf{U} außerdem noch *hermitisch*, d.h. ist $\mathbf{U}^H = \mathbf{U}$, dann folgt aus (A.41) sogar

$$\mathbf{U}^2 = \mathbf{I}. \tag{A.43}$$

Eine orthogonale Matrix kann als eine reelle Version der unitären Matrix angesehen werden und ist wie folgt definiert:

Definition A.5

Wenn eine reelle Matrix \mathbf{U} die Gleichung

$$\mathbf{U}^T \mathbf{U} = \mathbf{I} \tag{A.44}$$

erfüllt, dann heißt sie *orthogonal*.

Diese Matrizen sind für numerische Berechnungen besonders gut geeignet, da die Norm einer Matrix durch unitäre bzw. orthogonale Transformationen nicht verändert wird. So sind gewisse orthogonale Transformationen für eine Reihe numerischer Verfahren sehr nützlich. Die einfachste orthogonale Matrix ist

$$\mathbf{G} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \quad (\text{A.45})$$

und wird im Rahmen einer sogenannten *Givens-Transformation* eingesetzt. Eine weitere orthogonale Transformation, die in zuverlässigen Standard-Softwarepaketen (z.B. EISPACK: Lösung von Eigenwertproblemen, LINPACK: Lösung linearer Gleichungssysteme vielfach verwendet wird, kann mit Hilfe einer modifizierten Einheitsmatrix

$$\mathbf{U} = \mathbf{I} - \frac{2\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \quad (\text{A.46})$$

realisiert werden. Diese sogenannte *Householder-Matrix* ist nicht nur orthogonal sondern auch symmetrisch, da

$$\mathbf{U}^T = \left(\mathbf{I} - \frac{2\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right)^T = \mathbf{I} - \frac{2}{\mathbf{v}^T\mathbf{v}} (\mathbf{v}\mathbf{v}^T)^T = \mathbf{I} - \frac{2\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = \mathbf{U} \quad (\text{A.47})$$

gilt. Für jeden beliebigen Vektor $\mathbf{a} \neq \mathbf{0}$ existiert eine Householder-Matrix, die

$$\mathbf{U}\mathbf{a} = -\alpha\mathbf{e}_1, \quad (\text{A.48})$$

erfüllt. Dabei ist α eine von Null verschiedene Konstante und \mathbf{e}_1 der erste Einheitsvektor. Wird

$$\alpha = \text{sign}(a_1)\|\mathbf{a}\| \quad (\text{A.49})$$

und

$$\mathbf{v} = \mathbf{a} + \alpha\mathbf{e}_1 \quad (\text{A.50})$$

gewählt, dann ergibt sich

$$\mathbf{v}^T\mathbf{v} = (\mathbf{a} + \alpha\mathbf{e}_1)^T(\mathbf{a} + \alpha\mathbf{e}_1) \quad (\text{A.51})$$

mit

$$\mathbf{a}^T\mathbf{a} = \|\mathbf{a}\|^2 = (\text{sign}(a_1)\|\mathbf{a}\|)^2 = \alpha^2 \quad (\text{A.52})$$

zu

$$\mathbf{v}^T\mathbf{v} = 2\alpha(a_1 + \alpha). \quad (\text{A.53})$$

Die Anwendung dieser Transformation auf einen Vektor \mathbf{a} liefert dann

$$\begin{aligned}
 \mathbf{U}\mathbf{a} &= \left(\mathbf{I} - \frac{2\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) \mathbf{a} \\
 &= \mathbf{a} - \mathbf{v} \frac{2(\mathbf{a}^T\mathbf{a} + \alpha\mathbf{e}_1^T\mathbf{a})}{2\alpha(a_1 + \alpha)} \\
 &= \mathbf{a} - \mathbf{v} \frac{2(\alpha^2 + \alpha a_1)}{2\alpha(a_1 + \alpha)} \\
 &= \mathbf{a} - \mathbf{v} \\
 &= \mathbf{a} - \mathbf{a} - \alpha\mathbf{e}_1 = \alpha\mathbf{e}_1
 \end{aligned} \tag{A.54}$$

das gewünschte Ergebnis.

Die einfache Berechnung des Vektors \mathbf{v} ist ein weiterer Vorteil dieser Transformation, die nun auch eingesetzt werden kann, um Nullelemente in den Zeilen und Spalten einer Matrix zu erzeugen. Insbesondere kann mit Hilfe derartiger Householder–Matrizen eine beliebige $n \times m$ Matrix auf eine Dreiecksform transformiert werden, die sofort den Rang der Matrix erkennen läßt. Betrachtet man ein Matrix \mathbf{A} mit $n > m$, so wird zunächst mittels einer Householder–Transformation \mathbf{U}_1 die erste Spalte der Matrix verdichtet:

$$\mathbf{U}_1\mathbf{A} = \left[\begin{array}{c|c|c} a_1 & * & \\ - & + & - \\ 0 & & \\ \vdots & & \mathbf{A}_2 \\ 0 & & \end{array} \right]. \tag{A.55}$$

Ist \mathbf{U}_2 eine entsprechende Householder–Matrix zur Verdichtung der ersten Spalte der Matrix \mathbf{A}_2 , so ergibt sich

$$\left[\begin{array}{c|c|c} 1 & & 0 \\ - & + & - \\ 0 & & \mathbf{U}_2 \end{array} \right] \mathbf{U}_1\mathbf{A} = \left[\begin{array}{c|c|c|c|c} a_1 & * & * & * & \\ - & + & - & - & - \\ 0 & & a_2 & & * \\ \vdots & & 0 & & \\ \vdots & & \vdots & & \mathbf{A}_3 \\ 0 & & 0 & & \end{array} \right]. \tag{A.56}$$

Eine Fortsetzung dieses Prozesses liefert dann eine Matrix

$$\mathbf{U}\mathbf{A} = \left[\begin{array}{ccc|c|c} a_1 & * & & & \\ & a_2 & * & & \\ 0 & \ddots & & & * \\ & 0 & a_r & & * \\ - & - & - & + & - \\ & & 0 & & 0 \end{array} \right] \tag{A.57}$$

mit r linear unabhängigen Zeilen (somit ist $r = \text{Rang } \mathbf{A}$). Eine derartige Matrizen-Transformation wird daher auch als *Zeilenverdichtung* bezeichnet.

A.10 Hessenberg-Form einer Matrix

Viele numerische Verfahren zur Analyse und Synthese linearer Regelungssysteme werden sehr vereinfacht, wenn die zu untersuchende Matrix durch eine Ähnlichkeitstransformation $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ auf eine besondere Form gebracht wird. Eine solche besondere Form ist bei un-symmetrischen Matrizen die sogenannte *Hessenberg-Form* (A.58):

$$\begin{bmatrix} * & * & * & * & \cdots & * \\ * & * & * & * & \cdots & * \\ 0 & * & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & * & * \end{bmatrix}. \quad (\text{A.58})$$

Diese Hessenberg-Form kann durch die zuvor beschriebenen Householder-Matrizenoperationen leicht erzeugt werden. Das besondere an dieser Form ist, daß sie auch durch eine weitere Multiplikation von rechts mit der Householder-Matrix \mathbf{U} nicht zerstört wird. Eine derartige Transformation $\mathbf{U}\mathbf{A}\mathbf{U}$ stellt dann eine auch numerisch leicht zu realisierende Ähnlichkeitstransformation dar, weil aufgrund der Symmetrie und Orthogonalität der Householder-Matrix $\mathbf{U}^{-1} = \mathbf{U}$ gilt. Neben dem Rang bleiben durch eine Transformation auf die Hessenberg-Form auch die Eigenwerte unverändert.

A.11 Singulärwertzerlegung und Anwendungen

Für jede $n \times m$ Matrix² \mathbf{A} existieren unitäre Matrizen $\mathbf{U} \in \mathbb{C}^{n \times n}$ und $\mathbf{V} \in \mathbb{C}^{m \times m}$ in der Form, daß

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = [\mathbf{U}_1 \mid \mathbf{U}_2] \begin{bmatrix} \tilde{\mathbf{\Sigma}} & | & \mathbf{0} \\ - & + & - \\ \mathbf{0} & | & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^H \\ - \\ \mathbf{V}_2^H \end{bmatrix} \quad (\text{A.59})$$

mit

$$\tilde{\mathbf{\Sigma}} = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & & \end{bmatrix} \quad (\text{A.60})$$

² Ohne Einschränkung der Allgemeinheit kann hier vorausgesetzt werden, daß $n \geq m$ ist. Ist dies nicht der Fall, wird statt \mathbf{A} die konjugiert transponierte Matrix \mathbf{A}^H betrachtet.

und

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_m = 0 \quad (\text{A.61})$$

gilt. Die Zerlegung (A.59) wird *Singulärwertzerlegung* der Matrix \mathbf{A} genannt. Die Diagonalelemente $\sigma_1, \dots, \sigma_m$ der Matrix sind die *Singulärwerte* der Matrix \mathbf{A} und mit den positiven Wurzeln der Eigenwerte der Matrix $\mathbf{A}\mathbf{A}^H$ identisch. Die Spalten der Matrizen \mathbf{U} bzw. \mathbf{V} werden linke bzw. rechte *Singulärvektoren* genannt und sind die Eigenvektoren der hermiteschen Matrix $\mathbf{A}\mathbf{A}^H$ bzw. $\mathbf{A}^H\mathbf{A}$. Multipliziert man (A.59) von links mit \mathbf{U}^H , ergibt das

$$\mathbf{U}^H\mathbf{A} = \mathbf{\Sigma}\mathbf{V}^H, \quad (\text{A.62})$$

also für den i -ten Zeilenvektor ($i = 1, 2, \dots, r$)

$$\mathbf{u}_i^H\mathbf{A} = \sigma_i\mathbf{v}_i^H \quad (\text{A.63})$$

oder, konjugiert transponiert,

$$\mathbf{A}^H\mathbf{u}_i = \sigma_i\mathbf{v}_i \quad (\text{A.64})$$

und für $i > r$

$$\mathbf{A}^H\mathbf{u}_i = \mathbf{0}. \quad (\text{A.65})$$

Multipliziert man entsprechend (A.59) von rechts mit \mathbf{V} , so erhält man

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}, \quad (\text{A.66})$$

also für den i -ten Spaltenvektor ($i = 1, 2, \dots, r$)

$$\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i \quad (\text{A.67})$$

und für $i > r$

$$\mathbf{A}\mathbf{v}_i = \mathbf{0}. \quad (\text{A.68})$$

Das bedeutet, daß an einer Matrix \mathbf{A} mit Hilfe der unitären Matrizen \mathbf{U}, \mathbf{V} der Singulärwertzerlegung (A.59) eine Verdichtung der Zeilen

$$\mathbf{U}^H\mathbf{A} = \left[\begin{array}{c} \tilde{\mathbf{\Sigma}}\mathbf{V}_1^H \\ \mathbf{0} \end{array} \right] \left. \begin{array}{l} \} r \\ \} n-r \end{array} \right. \quad (\text{A.69})$$

bzw. der Spalten

$$\mathbf{A}\mathbf{V} = \left[\underbrace{\mathbf{U}_1\tilde{\mathbf{\Sigma}}}_r \mid \underbrace{\mathbf{0}}_{m-r} \right] \quad (\text{A.70})$$

vorgenommen werden kann.

A.11.1 Norm und Konditionszahl einer Matrix

Zur Beurteilung der numerischen Eigenschaften von Matrizenalgorithmen werden häufig bestimmte Matrixnormen verwendet. Diese Matrixnormen müssen dabei mit den entsprechenden Vektornormen eines Vektors $\mathbf{x} \in \mathbf{R}^n$, wie z.B. der *Maximumnorm*

$$\|\mathbf{x}\|_\infty := \max_i |x_i|, \quad (\text{A.71})$$

der *Summennorm*

$$\|\mathbf{x}\|_1 := |x_1| + |x_2| + \cdots + |x_n| \quad (\text{A.72})$$

und der *Euklidischen Norm*

$$\|\mathbf{x}\|_2 := \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \sqrt{\mathbf{x}^T \mathbf{x}} \quad (\text{A.73})$$

verträglich sein. Eine Matrixnorm $\|\mathbf{A}\|_p$ einer Matrix $\mathbf{A} \in \mathbf{R}^{n \times m}$ ist mit der Vektornorm $\|\mathbf{x}\|_p$ verträglich, wenn sie für alle $\mathbf{x} \in \mathbf{R}^m$ die Ungleichung

$$\|\mathbf{A}\mathbf{x}\|_p \leq \|\mathbf{A}\|_p \cdot \|\mathbf{x}\|_p. \quad (\text{A.74})$$

erfüllt.

Die folgenden drei Matrixnormen sind jeweils nur mit den entsprechenden Vektornormen verträglich:

$$\|\mathbf{A}\|_\infty := \max_i \sum_{j=1}^m |a_{ij}| \quad (\text{Zeilennorm}), \quad (\text{A.75})$$

$$\|\mathbf{A}\|_1 := \max_j \sum_{i=1}^n |a_{ij}| \quad (\text{Spaltennorm}), \quad (\text{A.76})$$

$$\|\mathbf{A}\|_2 := \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})} \quad (\text{Spektralnorm}). \quad (\text{A.77})$$

Mit allen drei Vektornormen ist allerdings die *Gesamtnorm*

$$\|\mathbf{A}\|_G := n \max |a_{ij}| \quad (\text{A.78})$$

verträglich. Eine leichter zu berechnende Norm, die ebenfalls mit der euklidischen Vektornorm verträglich ist, ist die *euklidische Matrixnorm*

$$\|\mathbf{A}\|_E := \sqrt{\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2} = \sqrt{\text{Spur } \mathbf{A}^T \mathbf{A}}. \quad (\text{A.79})$$

Aus den Definitionen dieser Matrixnormen ergibt sich folgender interessanter Zusammenhang zu den Singulärwerten einer Matrix:

$$\|\mathbf{A}\|_2 = \sigma_1 \quad (\text{A.80})$$

und

$$\|\mathbf{A}\|_E = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}, \quad (\text{A.81})$$

d.h. die Spektralnorm einer Matrix ist gleich dem größten Singulärwert und für die Berechnung der Euklidischen Norm wird die Summe aller Singulärwerte benötigt.

Vielfach hilfreich sind auch die in der Literatur angegebenen Beziehungen zwischen diesen Normen:

$$\text{i)} \quad \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_E \leq \sqrt{n} \|\mathbf{A}\|_2, \quad (\text{A.82})$$

$$\text{ii)} \quad \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_G, \quad (\text{A.83})$$

$$\text{iii)} \quad \frac{1}{\sqrt{m}} \|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{n} \|\mathbf{A}\|_\infty, \quad (\text{A.84})$$

$$\text{iv)} \quad \frac{1}{\sqrt{m}} \|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{n} \|\mathbf{A}\|_1. \quad (\text{A.85})$$

Mit Hilfe der Matrizenormen sind auch Aussagen über die Kondition einer $n \times n$ Matrix \mathbf{A} möglich. Der Faktor

$$\kappa(\mathbf{A}) := \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \quad (\text{A.86})$$

heißt *Konditionszahl* der Matrix \mathbf{A} . Eine Matrix mit einer sehr großen Konditionszahl $\kappa(\mathbf{A})$ wird als *schlecht konditionierte* Matrix bezeichnet. Wird die Konditionszahl (A.86) mit Hilfe der Spektralnorm (A.77) gebildet, so erhält man:

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}. \quad (\text{A.87})$$

Mit anderen Worten ergibt sich die Konditionszahl κ_2 aus dem Quotienten des größten und des kleinsten Singulärwertes, so daß eine singuläre Matrix die Konditionszahl ∞ aufweist. Im Gegensatz dazu sind orthogonale Matrizen mit $\kappa_2 = 1$ ideal konditioniert.

A.11.2 Numerische Berechnung der Singulärwertzerlegung

Zuverlässige Programme zur *numerischen* Berechnung der Singulärwertzerlegung einer *reellen* Matrix sind sowohl in der EISPACK-Bibliothek als auch in der LINPACK-Bibliothek und damit auch in MATLAB enthalten. Die numerisch berechneten Singulärwerte einer reellen Matrix \mathbf{A} sind bedingt durch die auftretenden Rundungsfehler gleich den Singulärwerte einer Matrix $\mathbf{A} + \mathbf{F}$, wobei für die Fehlermatrix \mathbf{F} folgende Abschätzungen gelten:

$$|\sigma_i(\mathbf{A} + \mathbf{F}) - \sigma_i(\mathbf{A})| \leq \|\mathbf{F}\| \leq \sigma_1(\mathbf{F}) \quad (\text{A.88})$$

mit

$$\|\mathbf{F}\| \leq N \cdot \epsilon \cdot \sigma_1(\mathbf{A}), \quad (\text{A.89})$$

wobei N ein Produkt aus der Anzahl der benötigten Rechenoperationen und einem Polynom niedriger Ordnung der Matrixdimension n ist. Darüber hinaus steht ϵ für die Rechengenauigkeit der verwendeten Gleitpunktarithmetik und $\sigma_1(\mathbf{A})$ für den größten Singulärwert der Matrix \mathbf{A} . Die erste Abschätzung (A.88) besagt nun, daß die Singulärwerte einer schwach gestörten Matrix $\mathbf{A} + \mathbf{F}$ nur wenig von den Werten der Matrix \mathbf{A} abweichen werden. Aufgrund dieser Eigenschaft nennt man die Singulärwerte einer Matrix auch gut konditioniert. In diesem Zusammenhang ist es erwähnenswert, daß eine der Abschätzung (A.88) entsprechende Beziehung für die Eigenwerte einer Matrix nicht angegeben werden kann. Ist die Norm der untersuchten Matrix klein, so werden aufgrund von (A.89) auch die absoluten Fehler der numerisch berechneten Singulärwerte klein sein.

A.11.3 Numerischer Rang einer Matrix

Viele Analyseverfahren im Zustandsraum erfordern die exakte Bestimmung des Ranges einer Matrix. Bedingt durch die endliche Wortlänge der Zahlendarstellung im Digitalrechner kann numerisch bestenfalls der wahre Rang einer schwach gestörten Matrix $\mathbf{A} + \Delta\mathbf{A}$ ermittelt werden. Mit anderen Worten wird nur ein sogenannter *numerischer* Rang bestimmt, während der „wahre“ Rang der Matrix als bloße Fiktion unbekannt bleibt.

Aufgrund der günstigen Fehlerabschätzung (A.88) liegt es nahe, die numerisch berechneten Singulärwerte zur Definition des *numerischen* Ranges einer Matrix heranzuziehen. Hierzu wird eine Toleranz τ eingeführt und festgelegt, daß die Matrix \mathbf{A} einen *numerischen* Rang r genau dann besitzt, wenn für die numerisch berechneten Singulärwerte gilt:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \tau \geq \sigma_{r+1} \geq \dots \geq \sigma_n. \quad (\text{A.90})$$

Die Toleranz τ muß mit der relativen Genauigkeit ϵ der verwendeten Gleitpunktarithmetik konsistent sein und sollte daher größer als $\epsilon \cdot \|\mathbf{A}\|_2$ sein. Sind die Werte der Matrixelemente z.B. aufgrund von fehlerbehafteten Messungen nicht exakt bekannt und ist dieser Fehler größer als ϵ , so ist ein entsprechend größerer Toleranzwert (z.B. $\tau = 10^{-2} \cdot \|\mathbf{A}\|_2$, wenn die Matrixelemente bis auf zwei Dezimalstellen korrekt sind) vorzugeben.

Die Festlegung des Ranges einer Matrix anhand der numerisch bestimmten Singulärwertzerlegung ist anerkanntermaßen die beste und zuverlässigste konventionelle Methode zur numerischen Rangbestimmung. Dies gilt auch unter dem Gesichtspunkt, daß die singulären Werte nicht nur eine reine Ja/Nein-Entscheidung bezüglich des vollen Ranges einer Matrix zulassen, sondern auch anzeigen, wie weit eine gegebene Matrix von Matrizen mit einem kleinerem Rang entfernt ist. Als ein derartiges Maß kann der kleinste Singulärwert σ_r mit $\sigma_r > \tau$ angesehen werden, wobei σ_r allerdings von der Norm der Matrix abhängt. Ein besseres Maß stellt daher der Wert $\sigma_r/\|\mathbf{A}\|$ dar, der bei einer Verwendung der Spektralnorm mit σ_r/σ_1 identisch ist.

B Numerische Überprüfung der Zustandssteuerbarkeit

Zur Überprüfung der Steuerbarkeit stehen eine Reihe verschiedener Kriterien zur Verfügung, die für eine zuverlässige numerische Analyse allerdings unterschiedlich gut geeignet sind. Wie die folgenden Untersuchungen aufzeigen werden, kann auch mit Hilfe der numerisch stabil umsetzbaren Kriterien nicht eindeutig festgestellt werden, wann ein System *nicht* vollständig steuerbar ist. Aufgrund der beschränkten Genauigkeit eines Digitalrechners wird bestenfalls ermittelt, daß ein System (\mathbf{A}, \mathbf{b}) sehr nahe an einem nicht steuerbaren System $(\mathbf{A} + \delta\mathbf{A}, \mathbf{b} + \delta\mathbf{b})$ liegt.

Die einfache Umsetzbarkeit der Zustandsraummethoden in effiziente Berechnungsalgorithmen wird als eine ihrer Stärken angesehen. Allerdings kann bei der Anwendung dieser Algorithmen nur in den einfachsten Fällen auf eine geeignete Rechnerunterstützung verzichtet werden. Galt diese Eigenschaft bis Anfang der achtziger noch als Handikap, so hat sich dies in den letzten Jahren durch die immens gestiegene Verfügbarkeit und Leistungsfähigkeit von Personalcomputer und der entsprechenden Software (z.B. MATLAB, SCILAB) ins Gegenteil verkehrt.

Die bei dem Einsatz dieser Verfahren im Bereich der Forschung und Lehre gemachten Erfahrungen zeigen allerdings immer wieder, daß es leicht zu fehlerhaften — numerisch gewonnenen — Aussagen kommen kann, die, begünstigt durch eine allgemein zunehmenden Computergläubigkeit, häufig auf eine Mißachtung von elementaren Grundregeln aus dem Bereich der numerischen Mathematik zurückzuführen sind. Darüber hinaus tragen im Bereich der Analyse und Synthese von Regelungssystemen mit Hilfe von Zustandsraumverfahren folgende Umstände zu einer Vergrößerung dieser Problematik bei:

- Die ein reales physikalisches System beschreibenden Zustandsmodelle sind häufig numerisch schlecht konditioniert, d.h. die Elemente der Systemmatrizen weisen große Betragsunterschiede auf.
- Die in der Standardliteratur angegebenen Analyse- und Syntheseverfahren für Systeme in einer Zustandsraumdarstellung sind aus einer numerisch orientierten Sicht oft für eine numerische Untersuchung weniger gut geeignet.

Diese Problematik und einige grundlegenden Möglichkeiten zur Erhöhung der Zuverlässigkeit numerisch ermittelter Ergebnisse im Bereich der Analyse und Synthese von Regelungssystemen sollen daher im folgenden für das Kalman-Kriterium anhand des Modelles eines elektrischen Antriebes einer Zahnradschleifmaschine näher erläutert werden.

Zur numerischen Überprüfung der Zustandssteuerbarkeit stehen, wie zuvor beschrieben, u.a. folgende zwei Verfahren zur Verfügung:

1. Überprüfung der Rangbedingung

$$\text{Rang } \mathbf{Q}_S = \text{Rang } [\mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}] = n. \quad (\text{B.1})$$

2. Überprüfung, ob für alle Eigenwerte λ_i , $i = 1, \dots, n$ der Matrix \mathbf{A}

$$\text{Rang } [\lambda\mathbf{I} - \mathbf{A}, \mathbf{b}] = n \quad (\text{B.2})$$

erfüllt ist.

Die aufgeführte Reihenfolge dieser beiden Verfahren gibt in etwa deren Bekanntheitsgrad an. Das bedeutet, das erste Verfahren ist das Standardverfahren, das man praktisch in jedem Lehrbuch findet. Im Gegensatz dazu wird das zweite Verfahren nur in neueren Werken angesprochen.

Bezüglich der numerischen Zuverlässigkeit muß diese Reihenfolge allerdings genau umgekehrt gesehen werden. Das erste Verfahren ist dann — wie im weiteren noch gezeigt wird — das unzuverlässigste und das zweite das numerisch zuverlässigere Verfahren.

Das Problem der Überprüfung der Steuerbarkeit eines linearen Systems (\mathbf{A}, \mathbf{b}) muß entsprechend seiner Natur zu den sehr sensiblen Problemstellungen gezählt werden, d.h., daß geringe relative Änderungen der Eingangsdaten große relative Fehler der Lösung zur Folge haben können. Diese Sensibilität des Steuerbarkeitsproblems hat ihre Ursache in der Tatsache, daß jedes nicht steuerbare System (\mathbf{A}, \mathbf{b}) durch eine geringfügige Änderungen der Elemente der Matrizen \mathbf{A}, \mathbf{b} in ein steuerbares System überführt werden kann. Diese Tatsache ist insbesondere für die numerische Überprüfung der Steuerbarkeit von besonderer Bedeutung, da bedingt durch die Rundungsfehler der mit einer endlichen Wortlänge arbeitenden Gleitkommaarithmetik die gefundene Lösung bestenfalls eine exakte Lösung für ein schwach gestörtes System $(\mathbf{A} + \delta\mathbf{A}, \mathbf{b} + \delta\mathbf{b})$ darstellt.

Diese Sensibilität oder *Kondition* eines Problems ist dabei zunächst unabhängig von dem gewählten Lösungsverfahren. So wie zwischen mehr oder weniger sensiblen Problemen muß auch zwischen gut oder schlecht konditionierten numerischen Verfahren unterschieden werden. Ein numerisches Verfahren wird dabei als gut konditioniert bezeichnet, wenn seine Lösung die Lösung eines mathematischen Problems ist, das durch geringe relative Änderungen der Eingabedaten aus dem ursprünglich vorgelegten mathematischen Problem entsteht.

So gesehen ist das in der Literatur am häufigsten anzutreffende Verfahren (1) zu den schlecht konditionierten numerischen Verfahren zur Überprüfung der Steuerbarkeit zu zählen, da die zunächst zu berechnende Steuerbarkeitsmatrix (B.1) sehr sensibel auf kleine Änderungen der Eingangsdaten reagieren kann, so daß die Sensibilität dieses Verfahrens oft größer ist als die Sensibilität des Ausgangsproblems.

Bei der numerischen Lösung eines Problems sind daher zunächst zwei Gesichtspunkte zu berücksichtigen:

- i) Die Sensibilität der vorliegenden Problemstellung muß bei der Interpretation der numerischen Lösung im Auge behalten werden.
- ii) Es sollte ein numerisches Verfahren ausgewählt werden, das gut konditioniert ist.

Um die weiteren Möglichkeiten zur Erhöhung der numerischen Zuverlässigkeit eines Rechnerprogrammes zu demonstrieren, wird im weiteren die Überprüfung der Steuerbarkeit mit Hilfe des schlecht konditionierten Standardverfahrens (1) untersucht.

Die weiteren Betrachtungen sollen exemplarisch an einem praktischen Beispiel erfolgen. Die folgenden Matrizen $\mathbf{A}, \mathbf{b}, \mathbf{c}$ beschreiben das um einen Arbeitspunkt linearisierte Zustandsmodell eines elektrischen Vertikaltriebes einer Zahnradschleifmaschine auf der Basis der physikalischen Einheiten kg, m, s.

$$\mathbf{A} = \begin{bmatrix} .000\text{D}+00 & .100\text{D}+01 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \\ -.545\text{D}+06 & -.689\text{D}+00 & .545\text{D}+06 & .000\text{D}+00 & .000\text{D}+00 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .100\text{D}+01 & .000\text{D}+00 \\ .167\text{D}+06 & .000\text{D}+00 & -.167\text{D}+06 & -.791\text{D}+00 & \underline{.985\text{D}-03} \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & \underline{-.281\text{D}+08} & -.167\text{D}+03 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} .000\text{D}+00 \\ .320\text{D}-02 \\ .000\text{D}+00 \\ .000\text{D}+00 \\ .000\text{D}+00 \end{bmatrix}$$

$$\mathbf{c}^T = \begin{bmatrix} .100\text{D}+01 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \end{bmatrix}$$

Für eine Rechengenauigkeit von 16 Dezimalstellen ist dieses Modell nicht gut konditioniert (vgl. Betragsunterschiede der unterstrichenen Elemente). Zur Überprüfung der Steuerbarkeit mittels des Verfahrens (1) ist jetzt zunächst die Berechnung der Steuerbarkeitsmatrix

$$\mathbf{Q}_S = [\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \mathbf{A}^3\mathbf{b}, \mathbf{A}^4\mathbf{b}] \tag{B.3}$$

$$= \begin{bmatrix} .000\text{D}+00 & .320\text{D}-02 & \underline{-.221\text{D}-02} & -.174\text{D}+04 & .240\text{D}+04 \\ .320\text{D}-02 & -.221\text{D}-02 & -.174\text{D}+04 & .240\text{D}+04 & .124\text{D}+10 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .534\text{D}+03 & -.791\text{D}+03 \\ .000\text{D}+00 & .000\text{D}+00 & .534\text{D}+03 & -.791\text{D}+03 & -.395\text{D}+09 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & -.150\text{D}+11 & \underline{.253\text{D}+13} \end{bmatrix}$$

notwendig. Diese Steuerbarkeitsmatrix ist für eine Rechengenauigkeit von 16 Dezimalstellen bereits sehr schlecht konditioniert.

Zur Überprüfung der Steuerbarkeit ist dann eine numerische Bestimmung des Ranges dieser schlecht konditionierten Matrix erforderlich. Hierbei muß jetzt noch berücksichtigt werden, daß bedingt durch die endliche Wortlänge der Zahlendarstellung im Digitalrechner nur eine Bestimmung des *numerischen* und nicht des wahren Ranges einer Matrix (vgl. Anhang) möglich ist.

Das zuverlässigste Verfahren zur Bestimmung des numerischen Ranges einer Matrix ist eine Berechnung der *Singulärwerte* mit den entsprechenden EISPACK- oder LINPACK-Programmen, die auch in MATLAB verfügbar sind. Die Singulärwerte σ_i , $i = 1, \dots, n$ einer $n \times n$ Matrix \mathbf{A} sind dabei die nach ihrer Größe geordneten ($\sigma_1 > \sigma_2 > \dots > \sigma_n$) positiven Wurzeln der Eigenwerte von $\mathbf{A}^T \mathbf{A}$. Die Anzahl der numerisch bestimmten Singulärwerte mit

$$\sigma_i > \epsilon = \sigma_1 \cdot \epsilon_m \cdot n \quad (\text{B.4})$$

gibt den numerisch Rang der Matrix an, wobei ϵ_m die Rechengenauigkeit der verwendeten Gleitkommaarithmetik ist. Dieser numerische Rang der Matrix \mathbf{A} ist der wahre Rang einer Matrix $\mathbf{A} + \mathbf{F}$, wobei für die Fehlermatrix \mathbf{F} folgende Abschätzungen angegeben werden können:

$$|\sigma_i(\mathbf{A} + \mathbf{F}) - \sigma_i(\mathbf{A})| \leq \|\mathbf{F}\| \leq \sigma_1(\mathbf{F}) \quad (\text{B.5})$$

mit

$$\|\mathbf{F}\| \leq N \cdot \epsilon_m \cdot \sigma_1(\mathbf{A}), \quad (\text{B.6})$$

wobei N ein Produkt aus der Anzahl der benötigten Rechenoperationen und einem Polynom niedriger Ordnung der Matrixdimension n ist. Die Abschätzungen (B.5) besagt, daß die Singulärwerte einer *schwach* gestörten Matrix $\mathbf{A} + \mathbf{F}$ nur wenig von den Werten der Matrix \mathbf{A} abweichen werden. Aufgrund dieser Eigenschaft, die die Eigenwerte einer Matrix beispielsweise nicht besitzen, werden die Singulärwerte auch als gut konditioniert bezeichnet. Ist die Norm der untersuchten Matrix klein, so wird aufgrund von (B.6) auch der absolute Fehler der numerisch berechneten Singulärwerte klein sein.

Die folgenden numerischen Untersuchungen wurden mit Hilfe von MATLAB mit doppelter Genauigkeit durchgeführt, so daß für die Rechengenauigkeit

$$\epsilon_m = 0.222 \cdot 10^{-15}. \quad (\text{B.7})$$

gilt.

Für die Steuerbarkeitsmatrix (B.3) ergeben sich dann folgende Singulärwerte:

$$\begin{aligned} \sigma_1 &= 0.253 \cdot 10^{13}, & \sigma_2 &= 0.774 \cdot 10^7, & \sigma_3 &= 0.199 \cdot 10^2, \\ \sigma_4 &= 0.321 \cdot 10^{-2}, & \sigma_5 &= 0.589 \cdot 10^{-5}. \end{aligned} \quad (\text{B.8})$$

Mit Hilfe des größten singulären Wertes σ_1 und der Rechengenauigkeit ϵ_m ergibt sich eine Nullschranke ϵ von

$$\begin{aligned}\epsilon &= \sigma_1 \cdot n \cdot \epsilon_m \\ &= 0.253 \cdot 10^{13} \cdot 5 \cdot 0.222 \cdot 10^{-15} \\ &= 0.281 \cdot 10^{-2}\end{aligned}\tag{B.9}$$

Aus

$$\sigma_4 = 0.321 \cdot 10^{-2} > \epsilon\tag{B.10}$$

und

$$\sigma_5 = 0.589 \cdot 10^{-5} < \epsilon\tag{B.11}$$

folgt, daß die Steuerbarkeitsmatrix (B.3) einen numerischen Rang von $4 < n$ besitzt, und demnach das Zustandsmodell des Schleifmaschinenantriebes als nicht vollständig steuerbar anzusehen ist.

Wenn dieses Ergebnis korrekt wäre, müßte bei der Bildung der Übertragungsfunktion eine Pol-/Nullstellenkürzung auftreten. Eine Berechnung der Pole und Nullstellen liefert allerdings folgendes Ergebnis:

$$\begin{aligned}\text{Pole:} & \quad 0.0, \quad -1.13 \pm 847.6j, \quad -83.11 \pm 119.2j, \\ \text{Nullstellen:} & \quad -145.6, \quad -11.11 \pm 437.6j.\end{aligned}$$

Das bedeutet, daß dieses System noch nicht einmal eng beieinander liegende Pole und Nullstellen besitzt, und daß daher das zuvor ermittelte Ergebnis nicht korrekt sein kann.

Im weiteren wird sich zeigen, daß ein korrektes Ergebnis auch mittels des schlecht konditionierten Standardverfahrens (1) erzielt werden kann, wenn das Ausgangssystem entsprechend skaliert wird. Unter Skalierung versteht man dabei eine Angleichung der Beträge einer Matrix zur Verbesserung ihrer numerischen Kondition.

Für ein System in einer Zustandsraumdarstellung $\mathbf{A}, \mathbf{b}, \mathbf{c}^T$ ermöglicht die Ausführung einer Zustandstransformation

$$\tilde{\mathbf{x}}(t) = \mathbf{D}\mathbf{x}(t)\tag{B.12}$$

mit einer Diagonalmatrix \mathbf{D} eine Skalierung der Systemmatrizen. Die Systemmatrizen des transformierten Systems ergeben sich dann zu:

$$\tilde{\mathbf{A}} = \mathbf{D}\mathbf{A}\mathbf{D}^{-1}, \quad \tilde{\mathbf{b}} = \mathbf{D}\mathbf{b}, \quad \tilde{\mathbf{c}}^T = \mathbf{c}^T\mathbf{D}^{-1}.\tag{B.13}$$

Bei der Bestimmung einer geeigneten Transformationsmatrix kann zwischen zwei Skalierungsverfahren unterschieden werden:

- i) Verbesserung der Kondition der Matrizen $\mathbf{A}, \mathbf{b}, \mathbf{c}$ durch eine günstige Wahl der physikalischen Einheiten des Zustandsmodelles.
- ii) Balancieren der Systemmatrizen.

Bei dem ersten Verfahren versucht man die physikalischen Einheiten des Zustandsvektors in der Form zu verändern, daß sich die Betragsunterschiede in den Systemmatrizen verringern. Diese Art der Skalierung bewirkt nur eine Änderungen der Einheiten an den Koordinatenachsen, so daß alle Ergebnisse, die anhand eines derartig skalierten Modelles gewonnen werden, ohne Rücktransformation physikalisch interpretierbar bleiben. So, wie die Normierung einer Analogrechenhaltung selbstverständlich ist, sollte eine derartige Skalierung eines linearen Zustandsmodelles immer vor weiteren numerischen Untersuchungen vorgenommen werden.

Bei dem Zustandsmodell des Werkzeugmaschinenantriebes repräsentiert die Zustandsvariable x_5 die Kraft in [N], die auf die Spindel des Werkzeugschlittens einwirkt. Wird diese Kraft in [kN] angesetzt, so entspricht dies einer Transformation (B.13) des Zustandsvektors mit

$$\mathbf{D} = \text{diag} [1, 1, 1, 1, 10^{-3}]. \quad (\text{B.14})$$

Für unser transformiertes System (B.13) erhalten wir dann:

$$\tilde{\mathbf{A}} = \begin{bmatrix} .000\text{D}+00 & .100\text{D}+01 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \\ -.545\text{D}+06 & \underline{-.689\text{D}+00} & \underline{.545\text{D}+06} & .000\text{D}+00 & .000\text{D}+00 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .100\text{D}+01 & .000\text{D}+00 \\ .167\text{D}+06 & .000\text{D}+00 & -.167\text{D}+06 & -.791\text{D}+00 & .985\text{D}+00 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & -.281\text{D}+05 & -.167\text{D}+03 \end{bmatrix},$$

$$\tilde{\mathbf{b}} = \begin{bmatrix} .000\text{D}+00 \\ .320\text{D}-02 \\ .000\text{D}+00 \\ .000\text{D}+00 \\ .000\text{D}+00 \end{bmatrix},$$

$$\tilde{\mathbf{c}}^T = \begin{bmatrix} .100\text{D}+01 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \end{bmatrix}.$$

Die Kondition der Systemmatrix $\tilde{\mathbf{A}}$ konnte durch diese Änderung der physikalischen Einheit der auf die Spindel des Werkzeugschlittens einwirkenden Kraft erheblich verbessert werden (vgl. unterstrichene Elemente). Diese Verbesserung der Kondition der Ausgangsdaten spiegelt sich auch in der zugehörigen Steuerbarkeitsmatrix

$$\mathbf{Q}_S = [\tilde{\mathbf{b}}, \tilde{\mathbf{A}}\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^2\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^3\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^4\tilde{\mathbf{b}}] \quad (\text{B.15})$$

$$= \begin{bmatrix} .000\text{D}+00 & .320\text{D}-02 & \underline{-.221\text{D}-02} & -.174\text{D}+04 & .240\text{D}+04 \\ .320\text{D}-02 & -.221\text{D}-02 & -.174\text{D}+04 & .240\text{D}+04 & .124\text{D}+10 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .534\text{D}+03 & -.791\text{D}+03 \\ .000\text{D}+00 & .000\text{D}+00 & .534\text{D}+03 & -.791\text{D}+03 & -.395\text{D}+09 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & -.150\text{D}+08 & \underline{.253\text{D}+10} \end{bmatrix}$$

wieder. Für diese besser konditionierte Steuerbarkeitsmatrix (B.15) ergeben sich dann die Singulärwerte zu:

$$\begin{aligned} \sigma_1 &= 0.284 \cdot 10^{10}, & \sigma_2 &= 0.686 \cdot 10^7, & \sigma_3 &= 0.198 \cdot 10^2, \\ \sigma_4 &= 0.321 \cdot 10^{-2}, & \sigma_5 &= 0.590 \cdot 10^{-5}. \end{aligned} \quad (\text{B.16})$$

Mit Hilfe des größten Singulärwertes σ_1 und der Rechengenauigkeit ϵ_m ergibt sich eine zugehörige Nullschranke von

$$\begin{aligned} \epsilon &= \sigma_1 \cdot n \cdot \epsilon_m \\ &= 0.285 \cdot 10^{10} \cdot 5 \cdot 0.222 \cdot 10^{-15} \\ &= 0.316 \cdot 10^{-5} \end{aligned} \quad (\text{B.17})$$

Aus

$$\sigma_5 = 0.589 \cdot 10^{-5} > \epsilon \quad (\text{B.18})$$

folgt, daß diese Steuerbarkeitsmatrix einen numerischen Rang von 5 besitzt, und das System somit vollständig steuerbar ist.

Eine weitere Möglichkeit zur Verbesserung der Kondition der Ausgangsdaten besteht in deren Balancierung. Dieses Balancieren verringert mit Hilfe einer Zustandstransformation (B.12) die Norm der Matrizen und gleicht die Größenordnungen der Matrizenelemente an. Bei der in MATLAB enthaltenen numerischen Realisierung SSBAL werden zur Transformation ausschließlich Diagonalmatrizen verwendet, deren Elemente ganzzahlige Potenzen der Rechnerbasis sind. Dadurch werden zusätzliche Rundungsfehler, sowohl bei der Transformation als auch bei der Rücktransformation ausgeschlossen.

Dieses Balancieren kann nun praktisch bei allen Problemstellungen, die gegenüber einer Zustandstransformation invariant sind (z.B. Pol- und Nullstellenberechnungen), angewendet werden.

Das Balancierungsprogramm SSBAL in MATLAB transformiert das Modell des Werkzeugschlittenantriebes mittels

$$\begin{aligned} \mathbf{D}^{-1} &= \text{diag} [2^{-9}, 2^1, 2^{-10}, 2^{-1}, 2^{+16}] \\ &= \text{diag} [0.20 \cdot 10^{-2}, 2.0, 0.977 \cdot 10^{-3}, 0.5, 0.655 \cdot 10^5] \end{aligned} \quad (\text{B.19})$$

auf folgende Form:

$$\tilde{\mathbf{A}} = \begin{bmatrix} .000\text{D}+00 & \underline{.102\text{D}+04} & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \\ -.532\text{D}+03 & \underline{-.689\text{D}+00} & .266\text{D}+03 & .000\text{D}+00 & .000\text{D}+00 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .512\text{D}+03 & .000\text{D}+00 \\ .654\text{D}+03 & .000\text{D}+00 & -.326\text{D}+03 & -.791\text{D}+00 & .129\text{D}+03 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & -.214\text{D}+03 & \underline{-.167\text{D}+03} \end{bmatrix},$$

$$\tilde{\mathbf{b}} = \begin{bmatrix} .000\text{D}+00 \\ .160\text{D}-02 \\ .000\text{D}+00 \\ .000\text{D}+00 \\ .000\text{D}+00 \end{bmatrix},$$

$$\tilde{\mathbf{c}}^T = \begin{bmatrix} .195\text{D}-02 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 \end{bmatrix}.$$

Die Kondition der Systemmatrix $\tilde{\mathbf{A}}$ konnte durch das Balancieren nochmals erheblich verbessert werden (vgl. unterstrichene Elemente). Bei der Steuerbarkeitsmatrix

$$\mathbf{Q}_S = [\tilde{\mathbf{b}}, \tilde{\mathbf{A}}\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^2\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^3\tilde{\mathbf{b}}, \tilde{\mathbf{A}}^4\tilde{\mathbf{b}}] \quad (\text{B.20})$$

$$= \begin{bmatrix} .000\text{D}+00 & .164\text{D}+01 & -.113\text{D}+01 & -.891\text{D}+06 & .123\text{D}+07 \\ .160\text{D}-02 & \underline{-.110\text{D}-02} & -.872\text{D}+03 & .120\text{D}+04 & .621\text{D}+9 \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & .547\text{D}+06 & -.810\text{D}+06 \\ .000\text{D}+00 & .000\text{D}+00 & .107\text{D}+04 & -.158\text{D}+04 & \underline{-.791\text{D}+9} \\ .000\text{D}+00 & .000\text{D}+00 & .000\text{D}+00 & -.229\text{D}+06 & .386\text{D}+08 \end{bmatrix}$$

hat dies zwar zu keiner weiteren Reduktion der Betragsunterschiede geführt, aber es erfolgte eine sichtbare Angleichung der Zeilennormen. Eine Beurteilung dieser Skalierungsmethode kann jetzt leichter anhand der zugehörigen numerisch berechneten Singulärwerten durchgeführt werden.

$$\begin{aligned} \sigma_1 &= 0.100 \cdot 10^{10}, & \sigma_2 &= 0.107 \cdot 10^7, & \sigma_3 &= 0.575 \cdot 10^2, \\ \sigma_4 &= 0.864 \cdot 10^0, & \sigma_5 &= 0.109 \cdot 10^{-2}. \end{aligned} \quad (\text{B.21})$$

Mit Hilfe des größten singulären Wertes σ_1 und der Rechengenauigkeit ϵ_m ergibt sich eine zugehörige Nullschranke von

$$\begin{aligned} \epsilon &= \sigma_1 \cdot n \cdot \epsilon_m \\ &= 0.100 \cdot 10^{10} \cdot 5 \cdot 0.222 \cdot 10^{-16} \\ &= 0.111 \cdot 10^{-5} \end{aligned} \quad (\text{B.22})$$

Aus

$$\sigma_5 = 0.109 \cdot 10^{-2} > \epsilon \quad (\text{B.23})$$

folgt, daß diese Steuerbarkeitsmatrix ebenfalls einen numerischen Rang von 5 besitzt, und das System somit vollständig steuerbar ist. Die Auswirkungen des Balancierens werden jetzt durch den vergrößerten Abstand des kleinsten singulären Wertes von der Nullschranke ϵ verdeutlicht. Obwohl das Balancieren die Betragsunterschiede in der Steuerbarkeitsmatrix (B.20) gegenüber (B.15) nicht weiter verringern konnte, besitzt die Matrix (B.20) einen kleinsten Singulärwert, der um drei Zehnerpotenzen größer ist als der entsprechende Wert von (B.15) bei identischer Nullschranke ϵ .

Der Einfluß der beiden Skalierungsmethoden wird besonders deutlich anhand der jeweiligen Konditionszahlen der einzelnen Steuerbarkeitsmatrizen. Die Konditionszahl κ einer regulären Matrix \mathbf{Q} ist dabei durch

$$\kappa = \|\mathbf{Q}\| \cdot \|\mathbf{Q}^{-1}\|. \quad (\text{B.24})$$

definiert.

Wird in (B.24) die Spektralnorm (vgl. Anhang) der Berechnung der Konditionszahl zugrunde gelegt, so erhält man mit den Singulärwerten $\sigma_i(\mathbf{Q})$

$$\kappa = \max \sigma_i(\mathbf{Q}) \cdot \max \left(\frac{1}{\sigma_i(\mathbf{Q})} \right) = \frac{\max \sigma_i(\mathbf{Q})}{\min \sigma_i(\mathbf{Q})}. \quad (\text{B.25})$$

Mit anderen Worten ergibt sich die Konditionszahl aus dem Quotienten des größten und des kleinsten singulären Wertes, so daß eine singuläre Matrix die Konditionszahl ∞ aufweist.

Die folgende Tabelle enthält die entsprechenden Konditionszahlen der jeweiligen Steuerbarkeitsmatrizen für skalierte bzw. nicht skalierte Ausgangsdaten.

Kraft in N	$\kappa = \frac{\sigma_1}{\sigma_5} = \frac{0.253 \cdot 10^{13}}{0.589 \cdot 10^{-5}} = 0.430 \cdot 10^{18}$
Kraft in kN	$\kappa = \frac{\sigma_1}{\sigma_5} = \frac{0.285 \cdot 10^{10}}{0.589 \cdot 10^{-5}} = 0.483 \cdot 10^{15}$
Balanciert	$\kappa = \frac{\sigma_1}{\sigma_5} = \frac{0.101 \cdot 10^{10}}{0.109 \cdot 10^{-2}} = 0.927 \cdot 10^{12}$

Tabelle A.1. Konditionszahlen der untersuchten Steuerbarkeitsmatrizen

Diese Tabelle zeigt, daß das Balancieren der Systemdaten zu einer erheblichen Verbesserung (≈ 6 Zehnerpotenzen) der Kondition der Steuerbarkeitsmatrix geführt hat. Aber bereits die günstigere Wahl der physikalischen Einheiten ergab eine so deutlich verbesserte Konditionszahl, daß eine erfolgreiche Rangbestimmung möglich wurde.

Als letztes soll jetzt ein Gesichtspunkt angesprochen werden, der für jeden, der die Grenzen numerischer Berechnungsverfahren kennt und berücksichtigt, selbstverständlich ist, der aber für viele, die vielleicht zum ersten Mal mit digitalen Lösungsverfahren in Berührung kommen, nicht unbedingt selbstverständlich sein muß. Es geht um eine, wie auch immer geartete, Verifikation der numerisch berechneten Lösung. Bei dem weiter oben diskutierten Beispiel der Überprüfung der Steuerbarkeit wurde bereits eine derartige Kontrollmöglichkeit angegeben und eingesetzt.

Abschließend sollen die Hauptaussagen dieses Abschnittes noch einmal wie folgt zusammengefaßt werden:

- Unter der Berücksichtigung der Sensibilität des Problems sollte nur ein entsprechend gut konditioniertes Verfahren ausgewählt werden.
- Anschließend sollte nur eine numerisch stabile Implementierung des gewählten Verfahrens unter weitgehender Verwendung bewährter Standardsoftware (MATLAB, EISPACK, LINPACK) eingesetzt werden.
- Die physikalischen Einheiten sollten in der Form gewählt werden, daß die Eingangsdaten möglichst gut konditioniert sind.
- Wenn die Problemstellung es zuläßt, sollten die Eingangsdaten zur weiteren Verbesserung der Kondition balanciert werden.
- Die numerisch berechnete Lösung ist mit Hilfe anderer Verfahren (z.B. zuverlässigen parameterunabhängigen Verfahren zur Überprüfung der strukturellen Steuerbarkeit) zu verifizieren.

C Stabilitätsdefinitionen

Die folgende Tabelle gibt einen Überblick über die Definitionen, die in den Vorlesungen SRT und RT verwendeten Stabilitätsbegriffe.

Tabelle: Stabilitätsdefinitionen für lineare, zeitinvariante Systeme, Teil 1

Begriff	Definition	Stabilitätsbedingung		
		Zeitbereich (Gewichtsfunktion)	Bildbereich (Übertragungsfunktion)	Zeitbereich (Zustandsraumdarstellung)
asymptotische Stabilität (Übertragungsverhalten)	Das Ausgangssignal $y(t)$ des nicht angeregten Systems klingt nach einer beliebigen Anfangsauslenkung $y_0 \neq 0, \dot{y}_0 \neq 0, \dots$ asymptotisch auf Null ab: $\lim_{t \rightarrow \infty} y(t) = 0$.	Die Gewichtsfunktion klingt asymptotisch auf Null ab: $\lim_{t \rightarrow \infty} g(t) = 0$.	All Pole der Übertragungsfunktion $G(s)$ liegen in der linken s -Halbebene.	Alle steuer- und beobachtbaren Eigenwerte der Systemmatrix \mathbf{A} liegen in der linken komplexen Halbebene.
BIBO-Stabilität	Für beschränkte Eingangssignale $ u(t) \leq N < \infty$ ist auch das Ausgangssignal beschränkt: $ y(t) \leq M < \infty$.	Die Gewichtsfunktion klingt asymptotisch auf Null ab: $\lim_{t \rightarrow \infty} g(t) = 0$.	All Pole der Übertragungsfunktion $G(s)$ liegen in der linken s -Halbebene.	Alle steuer- und beobachtbaren Eigenwerte der Systemmatrix \mathbf{A} liegen in der linken komplexen Halbebene.
Grenzstabilität (Übertragungsverhalten)	Das Ausgangssignal $y(t)$ des nicht angeregten Systems bleibt nach einer beliebigen Anfangsauslenkung $y_0 \neq 0, \dot{y}_0 \neq 0, \dots$ für wachsende Zeiten t in endlichen Grenzen: $ \lim_{t \rightarrow \infty} y(t) \leq C \leq \infty$.	Die Gewichtsfunktion bleibt für wachsende t in endlichen Grenzen: $ \lim_{t \rightarrow \infty} g(t) \leq D \leq \infty$.	All Pole der Übertragungsfunktion $G(s)$ liegen in der linken s -Halbebene oder auf der imaginären Achse, wobei die Pole auf der imaginären Achse alle einfach sind.	Alle steuer- und beobachtbaren Eigenwerte der Systemmatrix \mathbf{A} liegen in der linken komplexen Halbebene oder auf der imaginären Achse, wobei die Eigenwerte auf der imaginären Achse alle einfach sind.

Tabelle: Stabilitätsdefinitionen für lineare, zeitinvariante Systeme, Teil 2

Begriff	Definition	Stabilitätsbedingung		
		Zeitbereich (Gewichtsfunktion) (Transitionsmatrix)	Bildbereich (Übertragungsfunktion)	Zeitbereich (Zustandsraumdarstellung)
Instabilität (Übertragungsverhalten)	Das Ausgangssignal $y(t)$ des nicht angeregten Systems strebt nach einer beliebigen Anfangsauslenkung $y_0 \neq 0, \dot{y}_0 \neq 0, \dots$ mit wachsender Zeit t gegen Unendlich: $\lim_{t \rightarrow \infty} y(t) \rightarrow \infty$.	Die Gewichtsfunktion strebt mit wachsender Zeit t gegen Unendlich: $\lim_{t \rightarrow \infty} g(t) \rightarrow \infty$.	Mindestens ein Pole der Übertragungsfunktion $G(s)$ liegt in der rechten s -Halbebene oder auf der imaginären Achse ist ein mehrfacher Pol vorhanden.	Mindestens ein steuer- und beobachtbaren Eigenwert der Systemmatrix \mathbf{A} liegt in der rechten komplexen Halbebene oder auf der imaginären Achse, wobei der Eigenwert auf der imaginären Achse ein mehrfacher Eigenwert ist.
asymptotische Stabilität des Zustands	Der Zustandsvektor $\mathbf{x}(t)$ des nicht angeregten Systems geht nach einer beliebigen Anfangsauslenkung $\mathbf{x}_0 \neq \mathbf{0}$ asymptotisch gegen Null: $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{0}$.	Die Transitionsmatrix $\Phi(t)$ klingt asymptotisch auf Null ab: $\lim_{t \rightarrow \infty} \Phi(t) = \mathbf{0}$.	—	Alle Eigenwerte der Systemmatrix \mathbf{A} liegen in der linken komplexen Halbebene.
Instabilität des Zustands	Der Zustandsvektor $\mathbf{x}(t)$ des nicht angeregten Systems strebt nach einer beliebigen Anfangsauslenkung $\mathbf{x}_0 \neq \mathbf{0}$ gegen Unendlich: $\lim_{t \rightarrow \infty} \mathbf{x}(t) \rightarrow \infty$.	Die Transitionsmatrix $\Phi(t)$ strebt für wachsende t gegen Unendlich: $\lim_{t \rightarrow \infty} \Phi(t) \rightarrow \infty$.	—	Mindestens ein Eigenwert der Systemmatrix \mathbf{A} liegt in der rechten komplexen Halbebene oder auf der imaginären Achse, wobei der Eigenwert auf der imaginären Achse ein mehrfacher Eigenwert ist.

Index

- Abbildung
 - konforme, 37
- Ähnlichkeitstransformation, 70, 76
- Allpaß
 - ideal, 28
- Allpaßglied, 27
- Allpaßhaltiges System, 26
- Allpaßsystem, 27
- Amplitudenrand, 40, 41
- Antriebsregelstrecke, 17
- Ausgangs-Steuerbarkeitsmatrix, 65
- Ausgangsgleichung, 53
- Ausgangsrückführung, 74
- Ausgangssignalarückführung, 75
- Ausgangssteuerbarkeit, 64
 - Kriterium zur Überprüfung, 65
- Bandbreite, 11
- Beobachtbarkeit, 52, 63, 65
- Beobachtbarkeitskriterien
 - Hautus-Kriterium, 66
 - Kalman Kriterium, 66
- Beobachter, 77, 82
 - vollständiger, 78
- Beobachtereigenwerte, 81–83
- Beobachterrückführung, 79, 80, 83
- Beobachtungsfehler, 80, 81
- Beobachtungsgleichung, 53
- Beobachtungsnormalform, 73, 78, 81
- Bode-Diagramm, 8, 9
 - Beispiel, 33
- charakteristische Gleichung, 42–44, 59, 77
- D-Glied, 20
 - Bode-Diagramm, 20
 - Ortskurve, 20
 - Übertragungsfunktion, 20
- D-System, 20
- Differenzgrad, 50
- DT₁-Glied
 - Bode-Diagramm, 20
 - Ortskurve, 20
- Durchtrittsfrequenz, 39, 41
- Eigenbewegung, 62
 - abklingende, 60
 - aufklingende, 60
 - instabile, 63
- Eigenwert
 - nicht beobachtbar, 62
 - nicht steuerbar, 62
- Eigenwerte, 59, 61
 - beliebige Festlegung, 77
 - dominierende, 81
 - mehrfache, 70
 - verschiedene, 66
- Eingrößensysteme, 51
- Energiespeicher, 52
- Folgeregelkreis, 42
- Frequenzgang, 6
 - Amplitudengang, 7
 - Ortskurve, 8
 - Ortskurvendarstellung, 8
 - Phasengang, 7
 - transzendent, 28
- Frequenzgangdarstellung, 6
- Frequenzgangverfahren, 35
- Frequenzkennlinie
 - logarithmisch, 9
- Frequenzkennlinien-Darstellung, 8
 - Amplitudengang, 8
 - Betragskennlinie, 8
 - Phasengang, 8
 - Phasenkennlinie, 8
- Frequenzkennliniendiagramm, 8
- Führungsübertragungsfunktion, 38
- geschlossener Regelkreis

- Bandbreite, 41
- Systemmatrix, 76
- charakteristisches Polynom, 76
- Gewichtsfunktion, 57, 62
- Gleichung
 - charakteristische, 76
- Glied
 - phasendrehend, 27
- Hochpaß-Filter, 21
- I-System, 17
 - Bode-Diagramm, 19
 - Frequenzgang, 18
 - Pol, 18
 - Sprungantwort, 18
- Impulsantwort, 61
- IT₁-System
 - Bode-Diagramm, 18
- Jordannormalform, 70
- Knickfrequenz, 11
- Kreisverstärkung, 44
 - große, 50
- Lageregelung, 17
- Laplace-Transformation
 - Verschiebungssatz, 28
- Laplacetransformation
 - inverse, 57
- Luenberger-Beobachter, 77
- Matrixexponentialfunktion, 55, 56
- Mehrgrößensysteme, 51
- Minimalphasensystem, 26
- Nichtminimalphasensystem, 26
- Nichtphasenminimumsystem, 26, 27
- Nyquist-Ortskurve, 8
- Nyquistkriterium, 37, 38
 - Linke-Hand-Regel, 39
 - vereinfachtes, 38
- Nyquistkurve, 37
- Nyquistverfahren, 35, 40
 - kritischer Punkt, 38
- PD-System, 21
 - Bode-Diagramm, 22
 - Frequenzgang, 21
 - Pol-/Nullstelle, 22
 - Übertragungsfunktion, 21
- Phasenminimumsystem, 26, 27
- Phasenrand, 40, 41
- Phasenverhalten
 - nichtminimal, 26
- PI-Regler, 23
- PI-System, 22
 - Bode-Diagramm, 24
 - Frequenzgang, 23
 - Pol-Nullstelle, 23
 - Übertragungsfunktion, 23
- PID-Glied
 - Übertragungsfunktion, 24
- PID-Regler, 24
 - Übertragungsfunktion, 24
- PID-System, 24
 - Blockschaltbild, 25
 - Frequenzgang, 25
 - Pol-Nullstelle, 25
- Polbewegung, 44
- Polynom
 - charakteristisches, 76, 81
- PT₁-Glied
 - Bode-Diagramm, 9, 12
 - Differentialgleichung, 10
 - Frequenzgang, 10
 - Pole, 11
 - Sprungantwort, 10
 - Übertragungsfunktion, 10
- PT₂-System
 - Bode-Diagramm, 14
 - Frequenzgang, 14
 - komplexe-Pole, 13
 - Pole, 14
 - reelle Pole, 13

- Übergangsfunktion, 14
- PT_n-System, 17
 - Frequenzgang, 17
 - Pole, 17
 - Übergangsfunktion, 17
- Regelungsaufgabe
 - Lösung, 74
- Regelungsnormalform, 71, 72, 76
- Reglerverstärkung, 42, 44
- Rückführdifferenzfunktion, 38
- Rückführvektor, 77
- Separationstheorem, 82
- Signalflußdiagramm, 55, 67, 72
- Stabilität
 - Ausgangs-, 62
 - BIBO-, 62
 - Eingangs-, 62
 - asymptotische, 59, 62
- Stabilitätsgüte, 42
- Stabilitätsreserve, 40, 41
- Steuerbarkeit, 52, 63
- Steuerbarkeitskriterien
 - Hautus-Kriterium, 65
 - Kalman-Kriterium, 64
- Steuerbarkeitsmatrix, 64, 72
- Störübertragungsfunktion, 38
- System
 - Bewegungsgleichung, 56
 - duales, 68
 - duales System, 81
 - Eigenbewegungen, 59, 60
 - konzentrierte Parameter, 35
 - lineares, 52
 - mit Ausgleich, 8
 - ohne Nullstellen, 45
 - Ordnung, 53
 - PT₁, 10
 - Stabilität, 60
 - Stabilitätsrand, 60
 - Verstärkung, 8
 - verteilte Parameter, 35
 - Zustand, 52
- Systemmatrix, 53, 62, 76
 - Eigenwerte, 62
 - Frobenius-kanonischer Form, 71
- Systemverstärkung, 8
- Tiefpaß, 11
- Totzeit-Systeme, 35
- Totzeitsystem, 28
 - Blockschaltbild, 28
 - Frequenzgang, 28
- Trajektorie, 54
- Transitionsmatrix, 56
 - Eigenschaften, 57
- Übergangsmatrix, 56
- Übertragungsfunktion, 44, 69
 - Pol-Nullstellen-Form, 44
 - Pole, 62
- Vereinfachtes Nyquistkriterium
 - Bode-Diagramm, 39
- Verstärkung, 8
- WOK-Verfahren, 42, 44
 - Amplitudenbedingung, 44, 45
 - Phasenbedingung, 44
 - Regeln, 48
 - Winkelbedingung, 45
- Wurzelort, 42, 44
- Wurzelortskurve, 42, 44, 45
 - Nullstelle im Unendlichen, 46
 - Start- und Endpunkte, 46
 - Symmetrie, 47
 - Winkel der Asymptoten, 46
 - Wurzelorte auf der reellen Achse, 47
 - Wurzelschwerpunkt, 46, 48
- Zustandsbeobachter, 77, 80
- Zustandserreichbarkeit, 63
- Zustandsgleichung, 53, 75
 - Frequenzbereich, 58

- Lösung, 55
- Numerische Lösung, 58
- Zustandsgleichungen
 - Beobachter, 82
 - Regelstrecke, 82
- Zustandsgrößen, 52
- Zustandskurve, 54
- Zustandsmodell, 51, 52
- Zustandsraum, 51
- Zustandsraumdarstellung, 51, 59
- Zustandsraummodell, 53
 - Blockschaltbild, 54
 - Kalman–Zerlegung, 68
 - Normalformen, 69
- Zustandsregelung, 75–77
- Zustandsregler, 79
 - Entwurf, 74
- Zustandsrückführung, 74, 75, 77, 81, 83
- Zustandsschätzer, 77, 79
- Zustandssteuerbarkeit, 63
- Zustandsvektor, 52, 53