

Vorlesungsskript

Erweiterte Methoden der Regelungstechnik

Prof. Dr.-Ing. W. Schumacher

Stand : 9. Juli 2019

Inhaltsverzeichnis

1	Modellbasierte Regler	1
1.1	Interne Stabilität	1
1.2	Kompensationsregler	2
1.3	Parametrierter Regler für stabile Eingrößenstrecken	4
1.4	Smith-Prädiktor	5
2	Störgrößenaufschaltung und Führungsgrößenvorsteuerung	11
2.1	Regelung mit Störgrößenaufschaltung	11
2.2	Regelung mit Vorsteuerung	15
2.3	Modellfolgeregelung	16
2.4	Führungsgrößengenerator	17
3	Beobachtbarkeit und Steuerbarkeit	19
3.1	Normalformen	19
3.2	Transformation von Zustandsgrößen	20
3.3	Jordan-Normalform	21
3.4	Definition von Steuer- und Beobachtbarkeit	26
3.5	Realisierbarkeit und minimale Realisierung	27
4	Zustandsregelung	31
4.1	Definition von Steuerbarkeit	32
4.2	Zustandsrückführung für Mehrgrößensysteme	33
4.3	Ausgangsvektorrückführung	34
4.4	Praktische Anwendung der Polvorgabe	35
5	Beobachter	37
5.1	Zustandsrekonstruktion durch Parallelmodell	37
5.2	Luenberger-Beobachter	38
5.3	Zustandsrückführung mit Beobachter	41
5.4	Störgrößenbeobachter	43
5.5	Polvorgabe für Beobachter	46
6	Optimale Regelung	48
6.1	Lösung des LQ-Problems	49
6.2	Anwendung der optimalen Regelung	51
7	Zustandskurven	55
7.1	Trajektorien im zweidimensionalen Zustandsraum	55

7.2	Zustandskurven linearer Systeme 2. Ordnung	59
7.2.1	Trajektorien eines zweifachen Integrators	59
7.2.2	Zustandskurven eines verzögerten Integrators	61
7.2.3	Zustandskurven eines PT_2 -Gliedes	64
8	Nichtlineare Systeme	68
8.1	Nichtlineare Systeme in der Zustandsebene	69
8.1.1	Grenzyklen	69
8.1.2	Existenz von Grenzyklen	71
8.2	Nichtlineare Übertragungselemente	73
8.2.1	Zweipunktglied	73
8.2.2	Dreipunktglied	74
8.2.3	Begrenzung	75
8.2.4	Totzone oder Lose	76
8.2.5	Kennlinienglieder	77
8.3	Schaltende Regler	77
8.3.1	Linearisierung durch periodisches Schalten	78
8.3.2	Typen von Regelkreisen mit schaltenden Reglern	80
8.3.3	Zweipunktregler an PT_1 mit Laufzeitglied	83
8.4	Analyse schaltender Regler im Zustandsraum	87
8.4.1	Hysterese freier Zweipunktregler	88
8.4.2	Beschleunigungsstrecke mit Zweipunktregler und Rückführung	90
8.4.3	Beschleunigungsstrecke mit Zweipunktregler mit Hysterese	92
8.4.4	Beschleunigungsstrecke mit Zweipunktregler mit Hysterese und Rückführung	94
8.5	Beschleunigungsstrecke mit Dreipunktregler	96
8.5.1	Dreipunktregler mit Hysterese	96
8.5.2	Dreipunktregler mit Hysterese und Rückführung	98
9	Zeitoptimale Steuerung/Regelung	100
9.1	Zeitoptimale Steuerung eines PT_1 -Gliedes	101
9.2	Zeitoptimale Steuerung eines doppelten Integrators	102
9.3	Zeitoptimale Vorsteuerung mit Führungsgrößengenerator	108
10	Beschreibungs-Funktion	110
10.1	Begrenzer	113
10.2	Zweipunktglied	116
10.3	Grafische Auswertung	117
10.3.1	Begrenzerkennlinie mit Tot-Bereich	118
10.3.2	Zweipunktglied mit Hysterese	120

10.3.3 Dreipunktglied mit Hysterese	123
11 Stabilitätsuntersuchung nach Ljapunov	125
11.1 Beispiel zu Ruhelagen	126
11.2 Beispiel I: Mathematisches Pendel	128
11.3 Ljapunov-Stabilität	129
11.4 Beispiel II: Nichtlinearer Schwinger	131
11.5 Auffinden einer Ljapunov-Funktion	133
11.6 Beispiel III: Anwendung eines Ljapunov-Ansatzes	134
12 Popov-Kriterium	136
Literatur	139

1 Modellbasierte Regler

Die Auswahl von Reglern ist nicht auf die bekannten Standardregler beschränkt. Die Reglerübertragungsfunktion kann auch freizügig an die Strecke angepasst werden, ohne eine bestimmte Reglerordnung und/oder -struktur vorzugeben. Man kann sogar fordern, dass der geschlossene Regelkreis eine bestimmte Sollübertragungsfunktion aufweist. Dabei ist zu berücksichtigen, dass der geschlossene Kreis stabil ist und der Regler eine physikalisch realisierbare Übertragungsfunktion erhält. Doch zuvor soll der Begriff der internen Stabilität eingeführt werden.

1.1 Interne Stabilität

Interne Stabilität ist eine Erweiterung des Stabilitätsbegriffs. Gegeben sei der folgende Regelkreis ($N(s)$ sei dabei das Messrauschen):

Die Übertragungsfunktionen $K(s)$, $G(s)$ und $G_M(s)$ bilden den Regelkreis. Damit ergeben sich die Zustandsgrößen zu:

$$\begin{aligned} X_1(s) &= W(s) - G_M(s) X_3(s) \\ X_2(s) &= D(s) + K(s) X_1(s) \\ X_3(s) &= N(s) + G(s) X_2(s) \end{aligned}$$

oder in Matrixform:

$$\begin{bmatrix} 1 & 0 & G_M(s) \\ -K(s) & 1 & 0 \\ 0 & -G(s) & 1 \end{bmatrix} \begin{bmatrix} X_1(s) \\ X_2(s) \\ X_3(s) \end{bmatrix} = \begin{bmatrix} W(s) \\ D(s) \\ N(s) \end{bmatrix} . \quad (1.1)$$

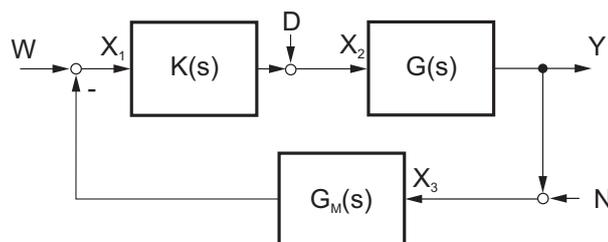


Bild 1.1: Geschlossener Regelkreis

Nach X_1, X_2 und X_3 aufgelöst ergibt sich:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \frac{1}{1 + G_M K G} \begin{bmatrix} 1 & -G G_M & -G_M \\ K & 1 & -K G_M \\ G K & G & 1 \end{bmatrix} \begin{bmatrix} W \\ D \\ N \end{bmatrix}. \quad (1.2)$$

Hierbei wird der Faktor $\frac{1}{1+G_M K G}$ als Empfindlichkeitsfunktion $S(s)$ definiert.

Es gilt nun folgendes:

Der Regelkreis ist *intern stabil*, wenn jede der neun Übertragungsfunktionen stabil ist

Interne Stabilität ist eine schärfere Stabilitätsbedingung als Stabilität beispielsweise der Führungsübertragungsfunktion. Das wird daran deutlich, dass es rein rechnerisch möglich ist, in einer der Übertragungsfunktionen instabile Pole durch Allpassnullstellen zu kompensieren. Dadurch treten bei einer praktischen Realisierung entsprechende Stabilitätsprobleme auf, die durch den Begriff der internen Stabilität aufgedeckt werden. Dies zeigt der folgende Abschnitt.

1.2 Kompensationsregler

Gibt man für die Übertragungsfunktion des geschlossenen Regelkreises, Bild 1.2, eine *Modellfunktion* $M(s)$ vor,

$$M(s) \stackrel{!}{=} \frac{K(s) G(s)}{1 + K(s) G(s)}, \quad (1.3)$$

so kann der Regler durch einfache Gleichungsumformung bestimmt werden:

$$K(s) = \frac{M(s)}{G(s) - M(s) G(s)} = \frac{1}{G(s)} \cdot \frac{M(s)}{1 - M(s)}. \quad (1.4)$$

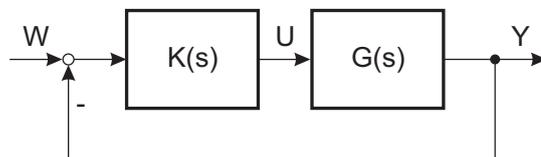


Bild 1.2: Geschlossener Regelkreis

In der Regel steht nur ein *Modell der Strecke* $\widehat{G}(s)$ zur Verfügung, so dass sich der Regler aus (1.4) zu

$$K(s) = \frac{1}{\widehat{G}(s)} \cdot \frac{M(s)}{1 - M(s)} \quad (1.5)$$

bestimmt. Zeichnet man den so gewonnenen Regler in das Blockschaltbild der Regelung ein, siehe Bild 1.3, so erkennt man, wo die zu beachtenden Punkte des Verfahrens liegen:

Um die gewünschte Modellübertragungsfunktion des geschlossenen Kreises zu erzeugen, ist im Regler ein multiplikativer Anteil $1/\hat{G}(s)$ enthalten, der das dynamische Verhalten der Strecke idealerweise vollständig kompensiert. Daher auch der Name *Kompensationsregler*.

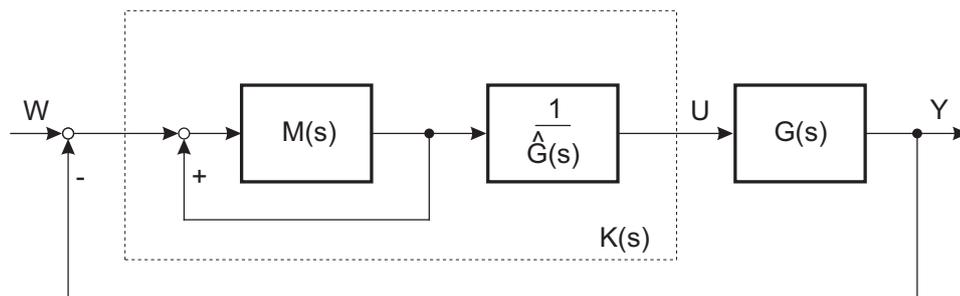


Bild 1.3: Kompensationsregler

Die Reihenschaltung des *inversen Modells der Regelstrecke* $1/\hat{G}(s)$ im Regler und $G(s)$ liefert die Übertragungsfunktion $G(s) \cdot 1/\hat{G}(s) \approx 1$. Übrig bleibt die gewünschte Modellübertragungsfunktion mit einer inneren Mit- und einer äußeren Gegenkopplung, die sich ebenfalls aufheben; es ergibt sich als Gesamtübertragungsfunktion des geschlossenen Kreises die gewünschte Modellübertragungsfunktion $M(s)$.

Folgende Bedingungen sind an die Strecke und die Modellübertragungsfunktion zu stellen, damit das Verfahren anwendbar ist:

- Die Regelstrecke muss *stabil* sein. Bei diesem einfachen Verfahren würden instabile Pole durch Allpass-Nullstellen mit positivem Realteil gekürzt. In der Praxis wird man aber die Pole nie exakt treffen, das Gesamtsystem bleibt folglich instabil.
- Die Regelstrecke muss aus dem gleichen Grunde *minimalphasig* sein, sonst würde eine Allpass-Nullstelle der Strecke wiederum durch einen instabilen Pol des Reglers kompensiert, was ebenfalls bei Anregungen ein unbeschränktes Wachsen der Stellgröße zur Folge hat.
- Damit ein realisierbarer Regler entsteht, muss der *Polüberschuss* des geforderten Modells mindestens so groß sein wie der der Strecke, so dass der Regler $K(s)$ aufgrund der eingeschlossenen Invertierung der Strecke noch ein kausales System ist.

Die Ordnung des Reglers $K(s)$ ergibt sich maximal aus der Summe der Zählerordnung der Strecke $G(s)$ und der Ordnung des Modells $M(s)$, wenn keine Kürzung möglich ist.

Das Modell wird sinnvollerweise für stationär genaues Führungsverhalten ausgelegt, d. h. $M(0) = 1$. Bei einer proportional wirkenden Strecke ergibt sich aus der Rechnung dann automatisch ein integrierender Regler.

1.3 Parametrierter Regler für stabile Eingrößenstrecken

Im Gegensatz zu den Kompensationsreglern, die ein inverses Modell der Regelstrecke beinhalten, ist in den *parametrierten Reglern* ein Modell der Strecke in einer Rückführung verwendet. Die Grundidee besteht darin, eine Regelung für eine stabile Strecke mit den Überlegungen für eine Steuerung zu entwerfen. Durch den Ansatz einer Steuerung ist bei Übereinstimmung zwischen Modell und Strecke Stabilität garantiert und der Entwurf des dynamischen *Vorsteuerfilters*, vergleiche Kapitel 2, ist wesentlich übersichtlicher.

Beim Ansatz einer modellbasierten Regelung wird dem Regler eine interne Struktur vorgegeben, die in Bild 1.4 dargestellt ist. Der Block $\hat{G}(s)$ innerhalb des Reglers enthält ein Modell der Regelstrecke. Durch die positive Rückführung der Stellgröße über das Modell $\hat{G}(s)$ wird die Regelschleife bei Übereinstimmung von $G(s)$ und $\hat{G}(s)$ für $K'(s)$ aufgehoben.

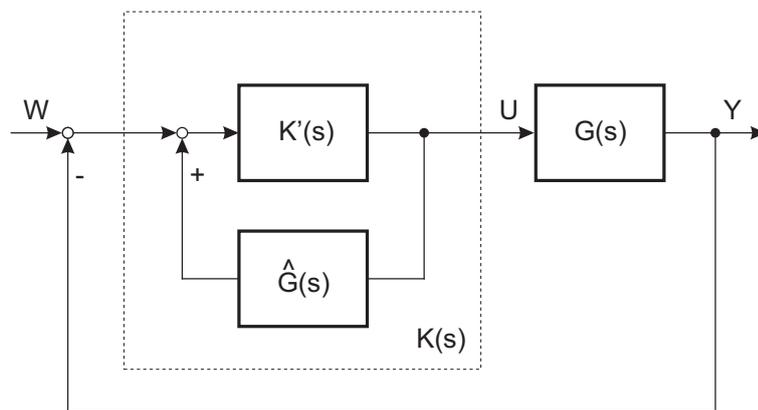
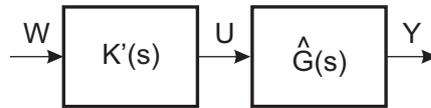


Bild 1.4: Ansatz einer modellbasierten Regelung

Die Übertragungsfunktion $K'(s)$ kann in der offenen Kette aus $K'(s)$ und einem stabilen Modell der Strecke $\hat{G}(s)$ entworfen werden, vergleiche Bild 1.5.

Bild 1.5: Entwurf von $K'(s)$ für $G(s) = \hat{G}(s)$

Die Berechnung des Reglers ist in diesem Fall ebenso einfach wie beim Kompensationsregler, Bild 1.4,

$$K(s) = \frac{K'(s)}{1 - \hat{G}(s) K'(s)}. \quad (1.6)$$

$K'(s)$ wird so gewählt, dass die Kette $K'(s) \hat{G}(s)$ ein vorgegebenes, gewünschtes dynamisches Verhalten aufweist. Vorzugsweise wird man auch hier stationäre Genauigkeit anstreben, also $K'(0) \hat{G}(0) = 1$ wählen. Wie schon beim Kompensationsregler führt diese Wahl bei proportionalen Strecken auf einen integrierenden Regler.

Da jedoch $\hat{G}(s)$ in der Regel kein exaktes Modell der Strecke $G(s)$ ist, wird der Regler $K'(s)$ auf Modellabweichungen entsprechend reagieren. Ist die Auslegung des Reglers nicht *robust*, d.h. empfindlich gegenüber Parametervariationen der Strecke, so kann dies zu Stabilitätsproblemen führen.

Gleichzeitig ist diese Art den Regler zu beschreiben eine Möglichkeit, die Menge aller Reglerübertragungsfunktionen zu spezifizieren, die die nominale Strecke stabilisieren: Alle Übertragungsfunktionen $K(s)$, die sich bei einem Streckenmodell $\hat{G}(s)$ schreiben lassen als Gleichung (1.6) mit einer stabilen Übertragungsfunktion $K'(s)$ stabilisieren die Strecke. Diese Art der Reglerdefinition wird auch als *Q-Parametrierung* bezeichnet – ein Verfahren nach Youla von 1976.

1.4 Smith-Prädiktor

Ein weiteres modellbasiertes Regelverfahren ist unter der Bezeichnung *Smith-Prädiktor* bekannt. Es wurde hauptsächlich für Regelstrecken mit großen Laufzeitanteilen entwickelt, wie sie in der Verfahrenstechnik häufig vorkommen.

Die Übertragungsfunktion der Regelstrecke $G(s)$ wird als eine Reihenschaltung eines rationalen Teils $G_R(s)$ (nur konzentrierte Energiespeicher) mit einem Laufzeitelement aufgefasst,

$$G(s) = G_R(s) e^{-T_L s} \quad \text{mit z.B.} \quad G_R(s) = G_{R1}(s) \cdot G_{R2}(s), \quad (1.7)$$

vergleiche auch Bild 1.6.

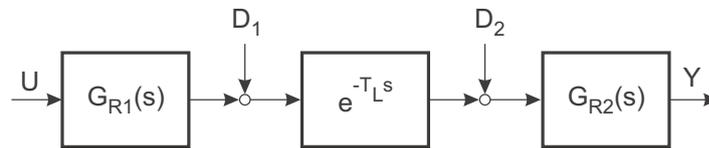


Bild 1.6: Aufteilung der Regelstrecke in rationale Anteile und Laufzeit

Bei der Führungsantwort der geregelten Strecke tritt bei einer konventionellen Reglerauslegung ein sehr langsames Einschwingen auf, weil der Regler auf die große Laufzeit als dominierende Zeitkonstante ausgelegt werden muss. Die Idee des Smith-Prädiktors beruht darauf, zunächst einen Regler nur für den schnellen rationalen Anteil der Übertragungsfunktion zu entwerfen, Bild 1.7. Die Laufzeit wird als nachgeschaltetes Element mit der Verstärkung 1 aufgefasst, obwohl dies nicht der physikalischen Struktur der Strecke entsprechen muss.

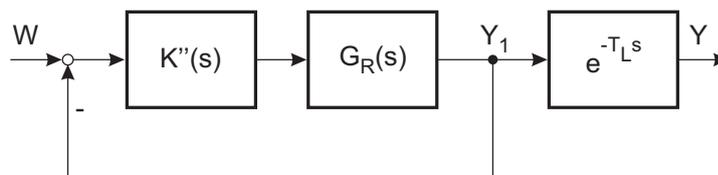
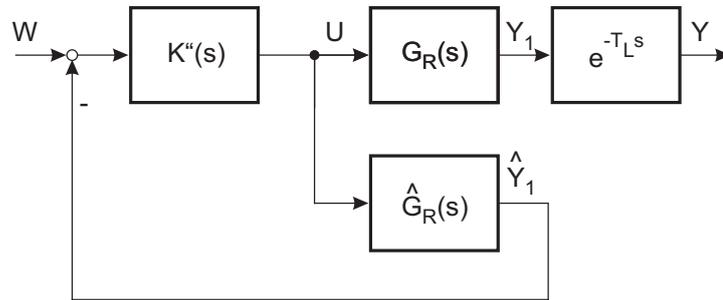


Bild 1.7: Idee des Smith-Prädiktors

Wie bereits erwähnt, wird der Regler nur auf den rationalen Teil der Strecke ausgelegt und kann entsprechend schneller reagieren. Allerdings gehen in dieser Variante die Störungen nicht in die Regelung ein und werden nach Art einer Steuerung nicht berücksichtigt. Nachteilig an diesem Konzept ist auch der Abgriff y_1 der Regelstrecke, der unter Umständen nicht vorhanden ist. Daher muss die Schätzgröße \hat{y}_1 durch ein Modell $\hat{G}_R(s)$ berechnet werden, Bild 1.8.

Bild 1.8: Berechnung von \hat{y}_1 durch ein Modell $\hat{G}_R(s)$

Dieses Schema lässt sich nach Art einer modellbasierten Regelung so ergänzen, dass der Ausgang der Strecke auch berücksichtigt wird. Es entsteht eine erste Darstellung der Regelung mit Smith-Prädiktor, Bild 1.9.

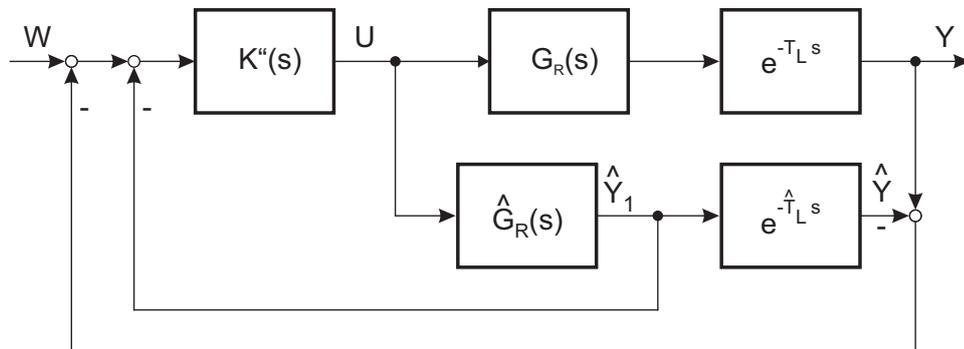


Bild 1.9: Regelung mit Smith-Prädiktor

Es wird durch Berechnung einer *geschätzten Ausgangsgröße* \hat{y} auf Basis der reinen Führungsanregungen ein Vergleich mit der tatsächlichen Ausgangsgröße y durchgeführt. Die Differenz kann vor dem Regler als *synthetischer Istwert* aufgeschaltet werden. Die Stabilität des ursprünglichen Reglers wird durch diese Maßnahme nicht gefährdet, solange die Modelle $\hat{G}_R(s)$ und $e^{-\hat{T}_L s}$ der Strecke in Bild 1.9 gut mit der tatsächlichen Strecke übereinstimmen. Die gute Übereinstimmung zeigt sich im verschwindenden synthetischen Istwert. Nach einer weiteren Umzeichnung kann der Smith-Prädiktor wie ein Reglerblock dargestellt werden, vgl. Bild 1.10.

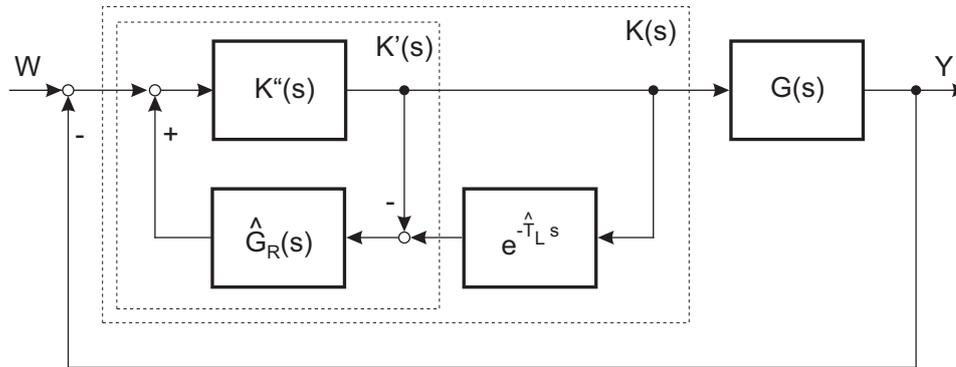


Bild 1.10: Smith-Prädiktor als Reglerblock

Man kann den Smith-Prädiktor in dieser Darstellung auch als eine spezielle Art der modellbasierten Regelung auffassen. Die Reglerübertragungsfunktion lautet

$$K(s) = \frac{K''(s)}{1 + K''(s) \hat{G}_R(s) (1 - e^{-\hat{T}_L s})}. \quad (1.8)$$

Bild 1.10 lässt sich umzeichnen, so dass Bild 1.11 entsteht. Hier wird die Analogie zur modellbasierten Regelung deutlich.

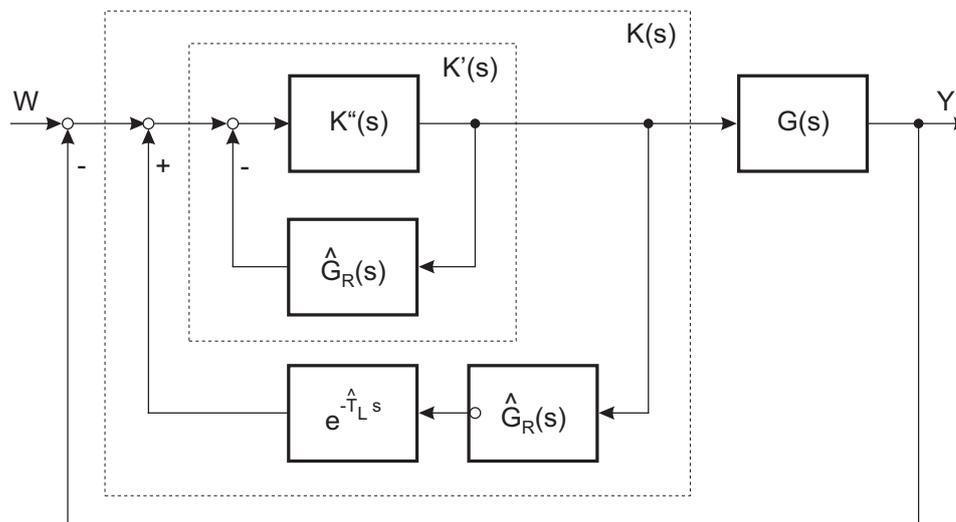


Bild 1.11: Smith-Prädiktor als modellbasierter Regler

Der Smith-Prädiktor bietet eine sehr gute Möglichkeit, lauffzeitbehaftete Strecken mit einer kurzen Einschwingzeit der Führungsantwort zu regeln. Zur Demonstration des Ver-

haltens einer Regelung mit Smith-Prädiktor wird an einer laufzeitbehafteten Strecke

$$G(s) = \frac{1}{s+1} \cdot \frac{1}{0.2s+1} \cdot e^{-2s} \quad (1.9)$$

eine Auslegung mit Smith-Prädiktor und ein einfacher PI-Regler gegenübergestellt. Im Smith-Prädiktor wird für die Regelung des rationalen Streckenteils wiederum ein PI-Regler eingesetzt, bei dem jedoch wegen der kleineren Streckenzeitkonstanten eine sehr viel höhere Verstärkung eingestellt werden kann. Bild 1.12 zeigt das Einschwingverhalten für Führungs- und Störانregung, $w(t) = \sigma(t)$, $d(t) = \sigma(t - 20)$. Man erkennt in beiden Anregungen ein deutlich besseres Einschwingen des Smith-Prädiktors.

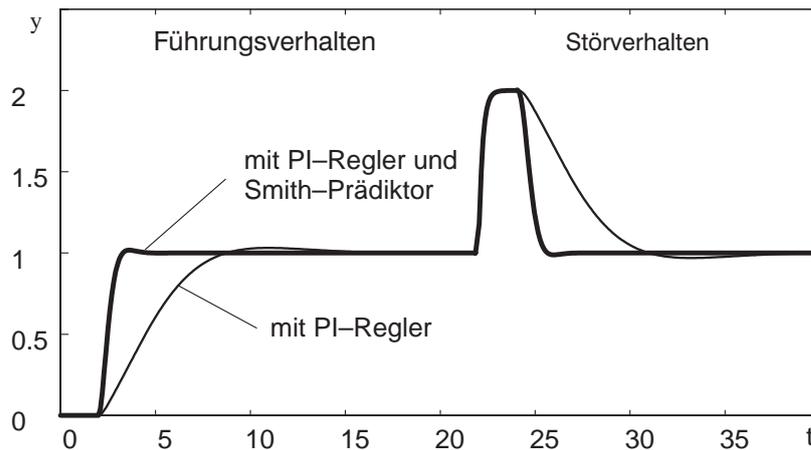


Bild 1.12: Vergleich Smith-Prädiktor / PI-Regler

Je größer die Laufzeit gegenüber der Einschwingzeit des rationalen Anteils ist, desto deutlicher wird der Vorteil des Smith-Prädiktors gegenüber einem Standardregler. Allerdings nimmt gleichzeitig auch die Robustheit gegenüber Fehleinstellungen der Zeitkonstanten ab. Dieses zeigt Bild 1.13. Hier wurde im Modell des Reglers die Laufzeit fehlerhaft abgeglichen gegenüber der angeschlossenen Strecke mit $T_L = 2$. Abweichungen von \hat{T}_L um 10 Prozent führen in diesem Beispiel schon zu deutlich schlechteren Einschwingvorgängen.

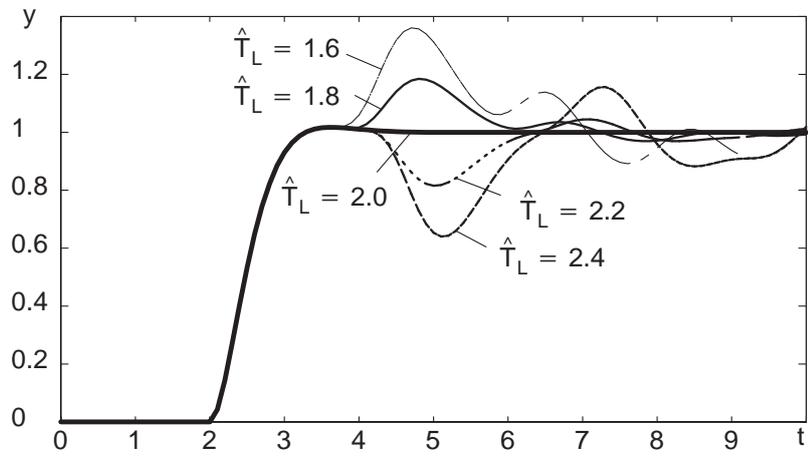


Bild 1.13: Einschwingvorgänge bei Fehleinstellung der Zeitkonstanten

2 Störgrößenaufschaltung und Führungsgrößenvorsteuerung

2.1 Regelung mit Störgrößenaufschaltung

In der Praxis vorkommende Regelkreise weisen häufig eine dominierende Störgröße d auf, die wie in Bild 2.1 dargestellt, innerhalb der Strecke angreift. Der Regler ist in der Lage, diese Störgröße zu bekämpfen. Jedoch ist zum wirkungsvollen Eingreifen des Reglers zunächst eine nennenswerte Regelabweichung notwendig, damit der Regler auf die Störung anspricht. Der Streckenteil $G_2(s)$ weist häufig die größten Zeitkonstanten auf. Dies führt dann zu verhältnismäßig lang andauernden Regelabweichungen als Reaktion auf die Störung.

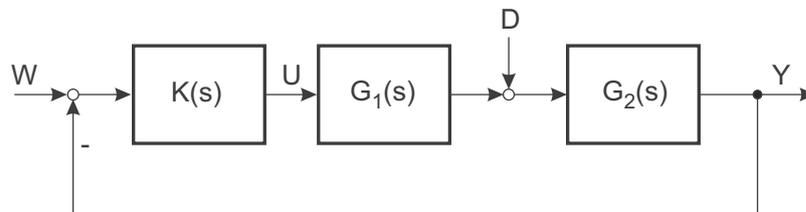


Bild 2.1: Regelkreis mit dominierender Störgröße

Wenn die Störung ohne großen Aufwand messbar ist, so bietet sich die Möglichkeit, mit einer dynamisch geformten Aufschaltung die Wirkung der Störung schneller zu neutralisieren. Dazu wird in der offenen Kette nach Art einer Steuerung eine direkte Kompensation der Störgröße durchgeführt. Hierbei ist zu berücksichtigen, dass als einziger Eingriffspunkt eine zusätzliche Aufschaltung nur an der ohnehin vorhandenen Stellgröße u möglich ist, Bild 2.2.

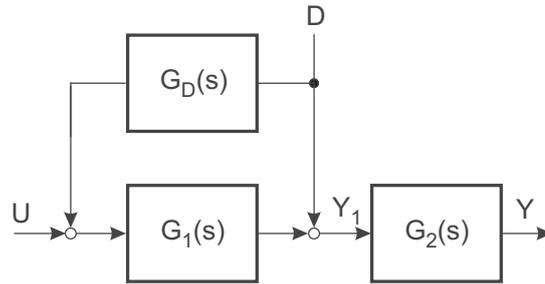


Bild 2.2: Störgrößenaufschaltung

Es gilt:

$$\begin{aligned}
 Y(s) &= G_2(s) D(s) + G_1(s) G_2(s) [U(s) + G_D(s) D(s)] \\
 &= G_1(s) G_2(s) U(s) + G_2(s) [1 + G_1(s) G_D(s)] D(s).
 \end{aligned} \tag{2.1}$$

Offensichtlich kann der Einfluss der Störung d vollständig unterdrückt werden, wenn

$$G_1(s) G_D(s) + 1 \stackrel{!}{=} 0 \tag{2.2}$$

gilt, also

$$G_D(s) = -\frac{1}{G_1(s)} \tag{2.3}$$

gewählt wird.

Da jedoch der reale Streckenteil $G_1(s)$ nicht ideal invertiert werden kann, ist die Funktion $G_D(s)$ nur angenähert realisierbar. Die Kompensation nach dieser Methode ist nur durchführbar, wenn $G_1(s)$ *stabil* und *minimalphasig* ist.

Beispiel: Für einen Streckenteil

$$G_1(s) = \frac{V_1}{T_1 s + 1} \tag{2.4}$$

lässt sich die Invertierung nicht ideal umsetzen, weil ein ideales PD-Element entsteht. Mit

$$G_D(s) = -\frac{1}{V_1} \frac{T_1 s + 1}{a T_1 s + 1} \quad (\text{PDT}) \tag{2.5}$$

lässt sich eine gute Kompensation erzielen. Mit dem Parameter a lässt sich ein Kompromiss zwischen der Geschwindigkeit der Kompensation und der eingesetzten Stellamplitude einstellen. $a = 0.1 \dots 0.2$ stellt eine gute Wahl dar. Auch eine statische *Vorsteuerung* $a = 1$ ist schon sehr wirkungsvoll für den gesamten Regelkreis, da dem Regler durch die Vorsteuerung die Regelung der stationären Größen erspart bleibt.

Die Übertragungsfunktion von d nach y_1 ist

$$\frac{Y_1(s)}{D(s)} = 1 - \frac{1}{a T_1 s + 1} = \frac{a T_1 s}{a T_1 s + 1}. \quad (2.6)$$

Die Fläche unter der Sprungantwort dieser Übertragungsfunktion ist ein Maß für die verbleibende Signalintensität. Die Fläche ist als Maßzahl für die Qualität des Systemverhaltens zu nehmen und das Systemverhalten als umso besser anzusehen, je kleiner die Fläche ist. Sofern die Störgröße d unabhängig von anderen Größen des Regelkreises ist und keine weiteren Verkopplungen bestehen, hat die Einführung einer *Störgrößenaufschaltung* keine Auswirkung auf die Stabilität des Regelkreises.

Die Bilder 2.3 und 2.4 zeigen eine Simulation zur Verdeutlichung der Störgrößenaufschaltung. Dabei werden zwei in Reihe geschaltete PT2-Glieder mit jeweils reellen Polen mit einem PI-Regler betrieben. Ein Störgrößensprung wird bei $t=20\text{s}$ zwischen den beiden Streckengliedern eingeleitet. Bild 2.3 zeigt die Ausgangsgröße. Es fällt vor allem der langsame Abfall bei nichtvorhandener Störaufschaltung auf. Der Stellgrößenverlauf ist in Bild 2.4 dargestellt. Hier ist die stationäre Abweichung der Stellgröße ohne Störaufschaltung zu beachten, während diese Abweichung in den anderen beiden Fällen gerade von der Störaufschaltung übernommen wird.

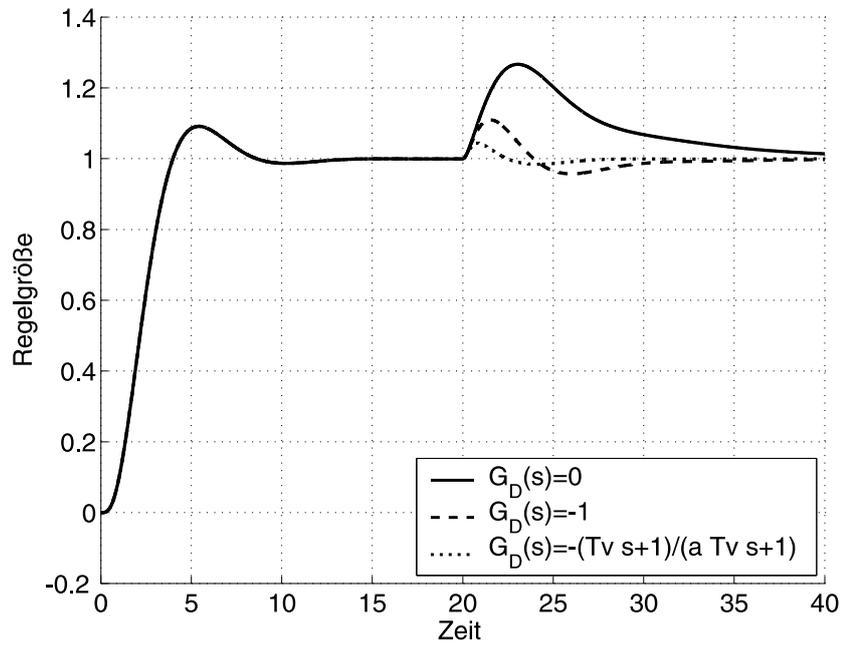


Bild 2.3: Verlauf der Regelgröße bei Führungs- und Störgrößensprung

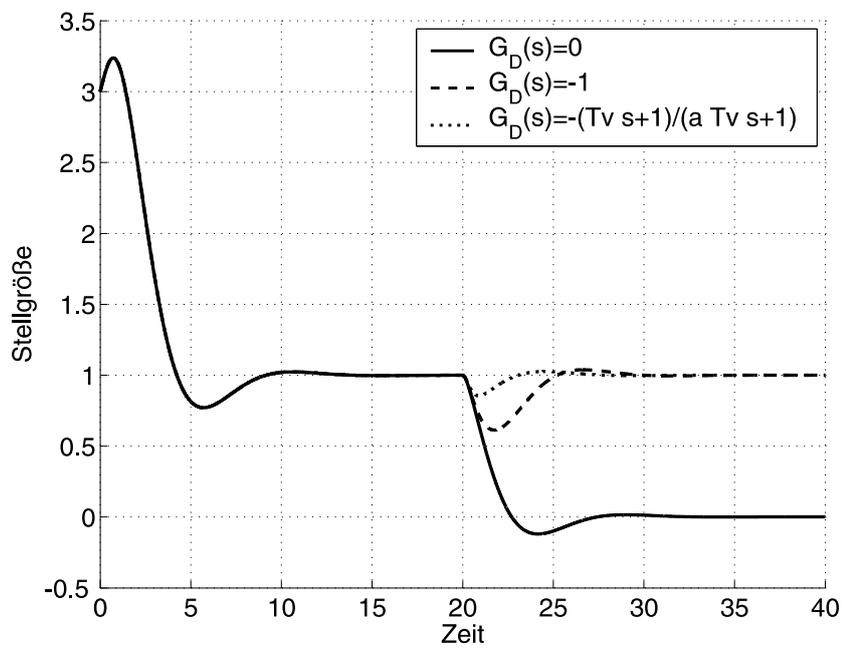


Bild 2.4: Verlauf des Reglerausgangs mit und ohne Störgrößenaufschaltung

2.2 Regelung mit Vorsteuerung

Genauso wie sich das Störverhalten durch eine Umschaltung nach Abschnitt 2.1 verbessern lässt, kann das Führungsverhalten durch eine *Führungsgrößenvorsteuerung* verbessert werden, Bild 2.5.

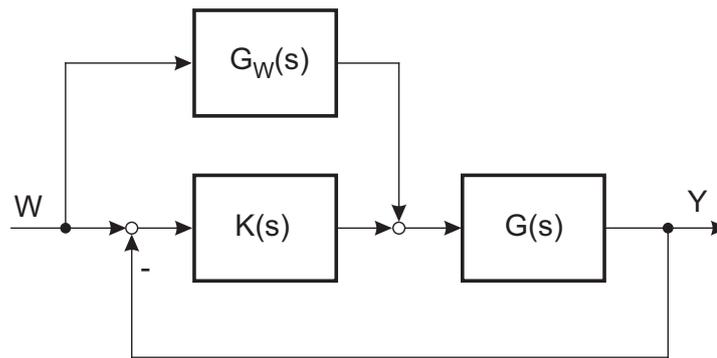


Bild 2.5: Führungsgrößenvorsteuerung

In diesem Fall liegt die Idee auch darin, den Regler von den dynamischen Eingriffen zu entlasten. Wenn es möglich ist, die invertierte Streckenübertragungsfunktion (Voraussetzung für $G(s)$: stabil und minimalphasig) als Vorsteuerung $G_W(s)$ zu realisieren, ist der Regler bei richtiger Einstellung der Vorsteuerung für Führungsanregung nicht notwendig (ideal: $G_W(s) \approx 1/G(s)$).

In vielen Fällen, zum Beispiel proportional wirkender Regler an proportionaler Strecke, ist auch eine rein statische Führungsvorsteuerung geeignet, den statischen Regelfehler zu reduzieren und das Einschwingen der Führungsanregung zu beschleunigen.

Bei einer praktischen Realisierung der Vorsteuerung muss man eine nicht vollständige Kompensation aller Pole der Strecke sowie zusätzliche „parasitäre“, zur Implementierung notwendige Pole in $G_W(s)$ in Kauf nehmen. Dies führt dann dazu, dass die Ausgangsgröße in der offenen Kette aus $G_W(s)$ und $G(s)$ der Führungsgröße nur verzögert folgt, aber bei korrekter Verstärkung auf den richtigen Endwert einschwingt. Je nach Ausnutzung des Betriebsbereichs des Stellgliedes kann der Regler diesen Vorgang nicht mehr wesentlich verbessern, wird aber trotzdem, da eine Regelabweichung auftritt, entsprechend reagieren. Dies kann verhindert werden, wenn ein weiteres Filter in den Führungsweig eingefügt wird, wie in Bild 2.6 gezeigt.

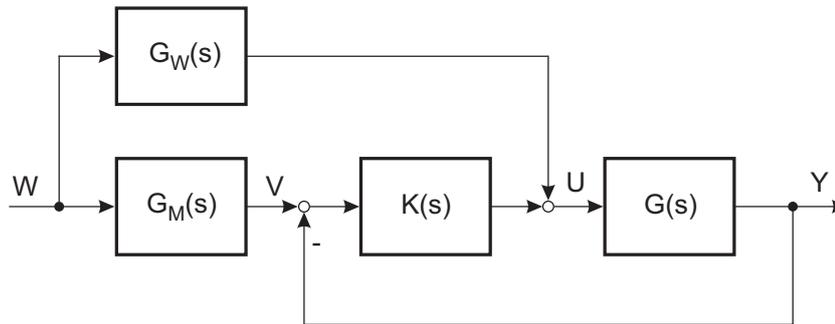


Bild 2.6: Führungsfilter

Wird jetzt das *Führungsfilter*

$$G_M(s) = G_W(s) \cdot \hat{G}(s), \quad G_M(0) = 1 \quad (2.7)$$

gewählt, so wird bei Führungsanregung und korrekter Wahl $\hat{G}(s) \equiv G(s)$ in Abwesenheit von Störungen der Verlauf von v und y identisch sein. Der Regler wird durch diese Struktur für Führungsanregungen nicht verwendet, ist aber für Störungen weiterhin voll wirksam. Bei Modellfehlern, falls $\hat{G}(s) \neq G(s)$ ist, wird der Regler so eingreifen, dass der Verlauf von y dem von v angenähert wird.

2.3 Modellfolgeregelung

In den bisher vorgestellten Vorsteuervarianten wird immer versucht, eine Übertragungsfunktion zumindest angenähert zu invertieren, um das Verhalten der Regelung zu verbessern. Die dabei auftretenden Ungenauigkeiten werden entweder ignoriert oder in speziellen Filtern wie in der zuletzt beschriebenen Struktur berücksichtigt.

Die explizite Invertierung kann jedoch auch vermieden werden, in dem ein eigener Regler an einer simulierten Strecke diese Aufgabe übernimmt, Bild 2.7.

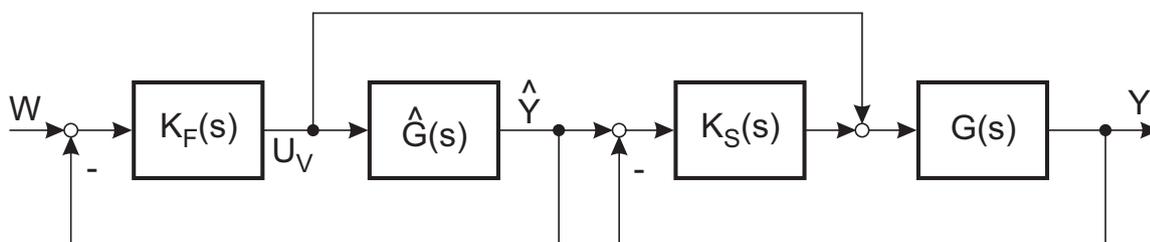


Bild 2.7: Modellfolgeregelung

Der Führungsregler $K_F(s)$ regelt ein Modell der Strecke $\hat{G}(s)$ mit der Ausgangsgröße \hat{y} . Dabei erzeugt der Regler die Stellgröße u_V , die gleichzeitig auf die zu regelnde Strecke $G(s)$ als Stellgröße aufgeschaltet wird. Falls das Modell $\hat{G}(s)$ identisch mit der Strecke $G(s)$ ist und keine Störgrößen angreifen, wird der Ausgang y unmittelbar gleich dem Ausgang \hat{y} sein. Das Signal \hat{y} kann als innerer Sollwert für einen *Störgrößenregler* $K_S(s)$ dienen. Dieser Regler reagiert nur bei Störungen oder Modellabweichungen zwischen $\hat{G}(s)$ und $G(s)$.

Das Führungsverhalten kann durch Einstellen von $K_F(s)$ unabhängig vom Störverhalten vorgegeben werden. Der Führungsregler $K_F(s)$ muss nicht unbedingt eigene Zustandsgrößen enthalten, da innerhalb des Reglers auch direkt auf die Zustände von $\hat{G}(s)$ zugegriffen werden kann. Der Vorteil der modellgestützten Vorsteuerung liegt in der Möglichkeit, Begrenzungen der Stellgröße unabhängig für Führung und Störung vorzugeben.

In einer Verallgemeinerung kann der Regler auch eine statische (evt. nichtlineare) Funktion der Zustandsgrößen von $\hat{G}(s)$ sein.

2.4 Führungsgrößengenerator

Das Verfahren der Modellfolgeregelung lässt sich für Kaskadenregelungen so weiterentwickeln, dass neben dem gewünschten Verlauf der Regelgröße bei Führungsanregung und der dazugehörigen Stellgröße auch die Verläufe der inneren Sollwerte der anderen Kaskadenstufen erzeugt werden, Bild 2.8.

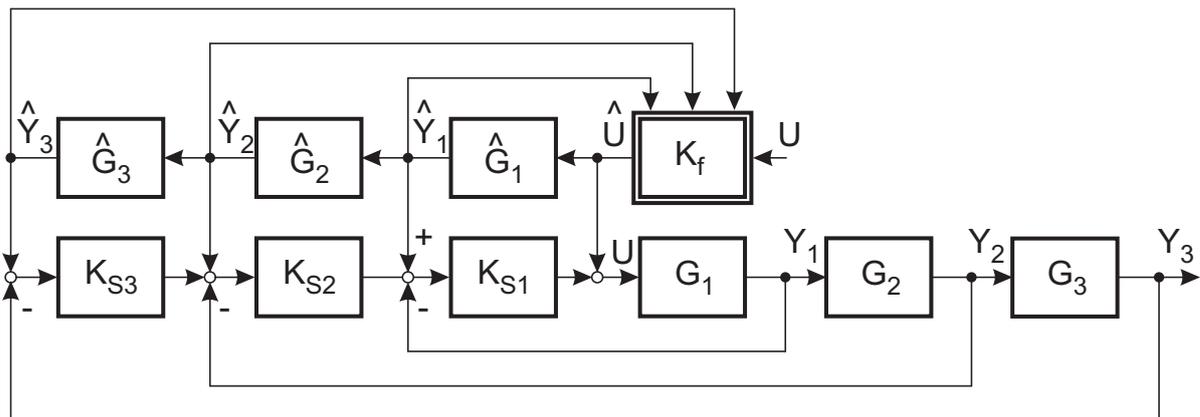


Bild 2.8: Führungsgrößengenerator

Die in Form einer als Kaskade zu regelnde Strecke besteht aus drei Teilstrecken $G_1(s)$, $G_2(s)$ und $G_3(s)$, denen die Reglerstufen $K_{S1}(s)$, $K_{S2}(s)$ und $K_{S3}(s)$ einer Kaskadenre-

gelung zugeordnet sind. Ein häufig nichtlinearer Reglerblock $K_F(s)$ ohne eigene Dynamik erzeugt ein Stellsignal \hat{u} , das eine Kette von Modellen der Teilübertragungstrecken $\hat{G}_1(s)$, $\hat{G}_2(s)$ und $\hat{G}_3(s)$ anregt. Die nur aufgrund der Führungsanregung berechneten Ausgangssignale \hat{y}_1 , \hat{y}_2 und \hat{y}_3 der Modelle der Teilübertragungstrecken $\hat{G}_1(s)$, $\hat{G}_2(s)$ und $\hat{G}_3(s)$ werden als Zusatzsollwerte auf die Reglerkaskade aufgeschaltet. Gleichzeitig dienen Sie zusammen mit dem Sollwert dem Führungsregler zur Erzeugung der Stellgröße. Bei Übereinstimmung der Teilübertragungsfunktionen mit ihren Modellen und Abwesenheit von Störungen wird die gesamte Übertragungstrecke von \hat{u} angesteuert. Für die Regler heben sich die Vorsteuersollwerte und die Istwerte der Strecken jeweils auf, so dass sie keine Regelabweichung erhalten und auch keine Stellgröße erzeugen.

Die Regler $K_{S_i}(s)$ der Kaskade übernehmen wie bei der Modellfolgeregelung die Funktion von Störreglern, die nur auf Störungen und Modellfehler reagieren. Das Thema Führungsgrößengenerator wird im Zusammenhang mit nichtlinearen Reglerfunktionen und zeitoptimalen Regelungen in einem späteren Kapitel nochmals aufgegriffen.

3 Beobachtbarkeit und Steuerbarkeit

Bereits im Grundlagen-Skript wurde im Kapitel 3.4 die *Zustandsdarstellung* dynamischer Systeme vorgestellt und erläutert.

3.1 Normalformen

Als eine wichtige Normalform für Eingrößensysteme (SISO – **S**ingle **I**nterface **S**ingle **O**utput) wurde die *Regelungsnormalform* dort im Kapitel 5.6 beschrieben, die nochmals in Bild 3.1 dargestellt ist.

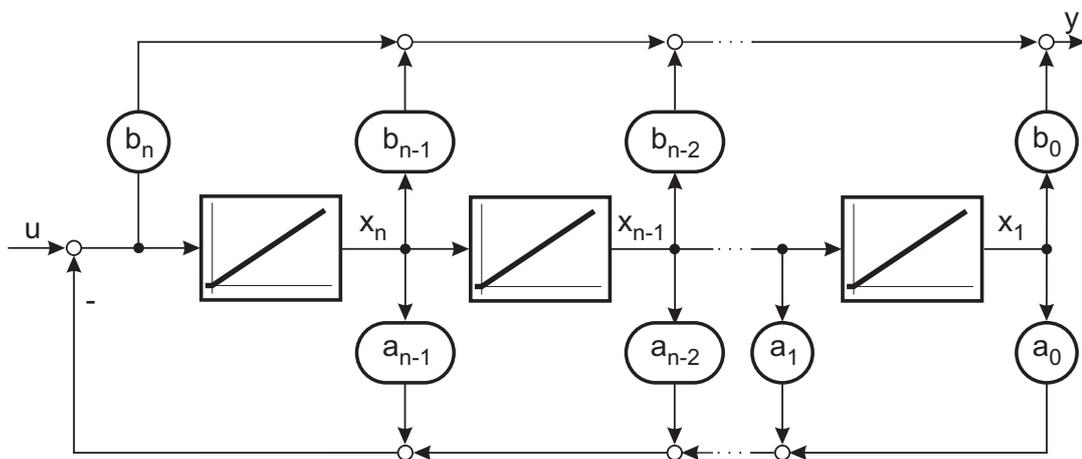


Bild 3.1: Regelungsnormalform

Als weitere gebräuchliche Normalform für Eingrößensysteme wird die *Beobachtungsnormalform* häufig verwendet, Bild 3.2.

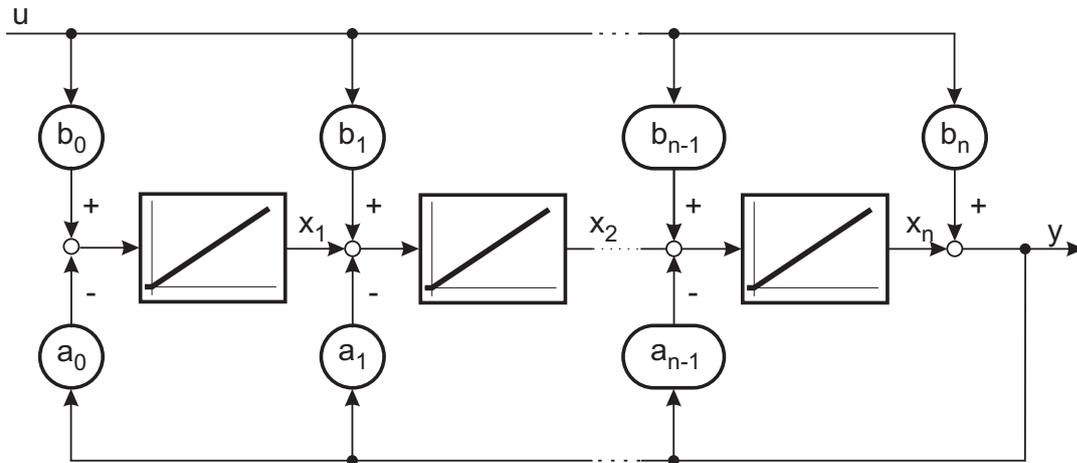


Bild 3.2: Beobachtungsnormalform

Von der Richtigkeit des Blockschaltbildes überzeugt man sich sehr einfach durch schrittweises Auflösen, indem alle Zwischengrößen auf die Summationsstelle am ersten Integrator verschoben werden. Man findet als zugehörige Übertragungsfunktion:

$$G(s) = \frac{b_n s^n + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (3.1)$$

Falls der Zählergrad kleiner ist als der Nennergrad, sind die höchsten Koeffizienten b_i entsprechend zu Null zu setzen.

Mathematisch gesehen handelt es sich bei der Umformung eines Systems aus einer beliebigen Zustandsraumdarstellung in eine andere Zustandsraumdarstellung um eine *Transformation der Zustandsgrößen*. Während die ursprüngliche Zustandsraumdarstellung eines Systems in der Wahl der Zustandsgrößen in der Regel physikalisch motiviert ist, haben die transformierten Zustandsgrößen meist keinerlei physikalische Bedeutung mehr.

3.2 Transformation von Zustandsgrößen

Gegeben sei ein dynamisches System in Zustandsraumdarstellung der Ordnung n :

$$\dot{\underline{x}} = \underline{A} \underline{x} + \underline{B} u, \quad \underline{x}(0) = \underline{x}_0 \quad (3.2)$$

$$\underline{y}(t) = \underline{C} \underline{x} + \underline{D} u \quad (3.3)$$

Dann erhält man mit einer nichtsingulären Matrix \underline{T} der Dimension $(n \times n)$ einen neuen Satz von Zustandsgrößen durch:

$$\underline{T} \tilde{x} = x \quad \text{bzw.} \quad \tilde{x} = \underline{T}^{-1} x \quad (3.4)$$

Die Transformation gilt gleichermaßen für die zeitliche Ableitung der neuen Zustandsgrößen:

$$\underline{T} \dot{\tilde{x}} = \dot{x} \quad \text{bzw.} \quad \dot{\tilde{x}} = \underline{T}^{-1} \dot{x}. \quad (3.5)$$

Diese Beziehungen ergeben eingesetzt in das Zustandsraummodell:

$$\dot{\tilde{x}} = \underline{T}^{-1} \dot{x} = \underline{T}^{-1} \underline{A} x + \underline{T}^{-1} \underline{B} u \quad (3.6)$$

$$\dot{\tilde{x}} = \underbrace{\underline{T}^{-1} \underline{A} \underline{T}}_{\tilde{\underline{A}}} \tilde{x} + \underbrace{\underline{T}^{-1} \underline{B}}_{\tilde{\underline{B}}} u \quad (3.7)$$

mit der Ausgangsgleichung

$$y = \underbrace{\underline{C} \underline{T}}_{\tilde{\underline{C}}} \tilde{x} + \underbrace{\underline{D}}_{\tilde{\underline{D}}} u. \quad (3.8)$$

Es entstehen damit neue Systemmatrizen:

$$\tilde{\underline{A}} = \underline{T}^{-1} \underline{A} \underline{T} \quad (3.9)$$

$$\tilde{\underline{B}} = \underline{T}^{-1} \underline{B} \quad (3.10)$$

$$\tilde{\underline{C}} = \underline{C} \underline{T} \quad (3.11)$$

$$\tilde{\underline{D}} = \underline{D} \quad (3.12)$$

3.3 Jordan-Normalform

Die Systemmatrix \underline{A} kann als Abbildung aufgefasst werden, die den n -dimensionalen Vektorraum auf sich selbst abbildet. Dabei gibt es ausgezeichnete Vektoren v_i , die bei der Abbildung ihre Richtung beibehalten und nur in der Länge geändert werden. Zu jedem

dieser so genannten *Eigenvektoren* gehört ein Streckungsfaktor λ_i , der als *Eigenwert* bezeichnet wird. Es gilt:

$$\underline{\underline{A}} \underline{v}_i = \lambda_i \underline{v}_i. \quad (3.13)$$

Es existieren maximal n verschiedene Eigenvektoren. Löst man (3.13) auf, so erhält man:

$$(\lambda_i \underline{\underline{I}} - \underline{\underline{A}}) \underline{v}_i = \underline{0}. \quad (3.14)$$

Sucht man nichttriviale Lösungen dieser Gleichung, so muss gelten:

$$\det(\lambda_i \underline{\underline{I}} - \underline{\underline{A}}) = 0. \quad (3.15)$$

Dieses ist die bereits bekannte Gleichung zur Bestimmung der Eigenwerte einer Systemmatrix $\underline{\underline{A}}$.

Hat diese Gleichung n verschiedene Eigenwerte als Lösung, so gibt es n verschiedene, linear unabhängige Eigenvektoren. Im Falle n linear unabhängiger Eigenvektoren können diese für die weitere Lösung direkt verwendet werden.

Sind unter den n Eigenwerten j gleiche vorhanden, so können zu den gleichen Eigenwerten noch gleiche oder unterschiedliche Eigenvektoren gehören. Lässt sich zu einem Eigenwert, der mit der Vielfachheit j auftritt nur ein linear unabhängiger Eigenvektor finden, so muss für die j vielfachen Eigenwerte auf *verallgemeinerte Eigenvektoren* zurückgegriffen werden. Tritt der Eigenwert λ_i mit der Vielfachheit j auf, so bestimmt man verallgemeinerte Eigenvektoren zu:

$$\begin{aligned} (\lambda_i \underline{\underline{I}} - \underline{\underline{A}}) \underline{v}_{i,1} &= \underline{0} \\ (\lambda_i \underline{\underline{I}} - \underline{\underline{A}}) \underline{v}_{i,2} &= \underline{v}_{i,1} \\ &\vdots \\ (\lambda_i \underline{\underline{I}} - \underline{\underline{A}}) \underline{v}_{i,j} &= \underline{v}_{i,j-1}. \end{aligned} \quad (3.16)$$

Nach diesen Überlegungen stehen n linear unabhängige (verallgemeinerte) Eigenvektoren \underline{v}_i zur Verfügung.

Schreibt man diese Vektoren spaltenweise in eine Matrix

$$\underline{\underline{V}} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = (\underline{v}_1, \dots, \underline{v}_n) \quad (3.17)$$

als Spaltenvektoren, so entsteht eine nichtsinguläre Matrix, die als Transformationsmatrix für eine Transformation des Zustandsraumes verwendet werden kann.

Wendet man diese Transformation auf das System an, so erhält man für den Fall, dass sämtliche Eigenvektoren verschieden sind (nicht verallgemeinerte Eigenvektoren) als transformierte Systemmatrix $\underline{\tilde{A}}$ eine Diagonalmatrix mit den Eigenwerten in der Hauptdiagonalen:

$$\underline{\tilde{A}} = \underline{V}^{-1} \underline{A} \underline{V} = \text{diag } \lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad (3.18)$$

Wendet man eine Transformation mit verallgemeinerten Eigenvektoren an, so entstehen im Bereich der Eigenwerte, die zu verallgemeinerten Eigenvektoren gehören in der Nebendiagonalen noch Werte von 1:

$$\underline{\tilde{A}} = \underline{V}^{-1} \underline{A} \underline{V} = \begin{pmatrix} \lambda_1 & 1 & & & 0 \\ & \lambda_1 & 1 & & \\ & & \lambda_1 & 1 & \\ & & & \lambda_2 & \\ & & & & \ddots & \\ 0 & & & & & \lambda_{n-j} \end{pmatrix} \quad (3.19)$$

Die weiteren Systemmatrizen werden ebenfalls gemäß (3.10) und (3.11) transformiert. Treten keine verallgemeinerten Eigenvektoren auf, so lässt sich zusammenfassend das Blockschaltbild in Bild 3.3 zeichnen.

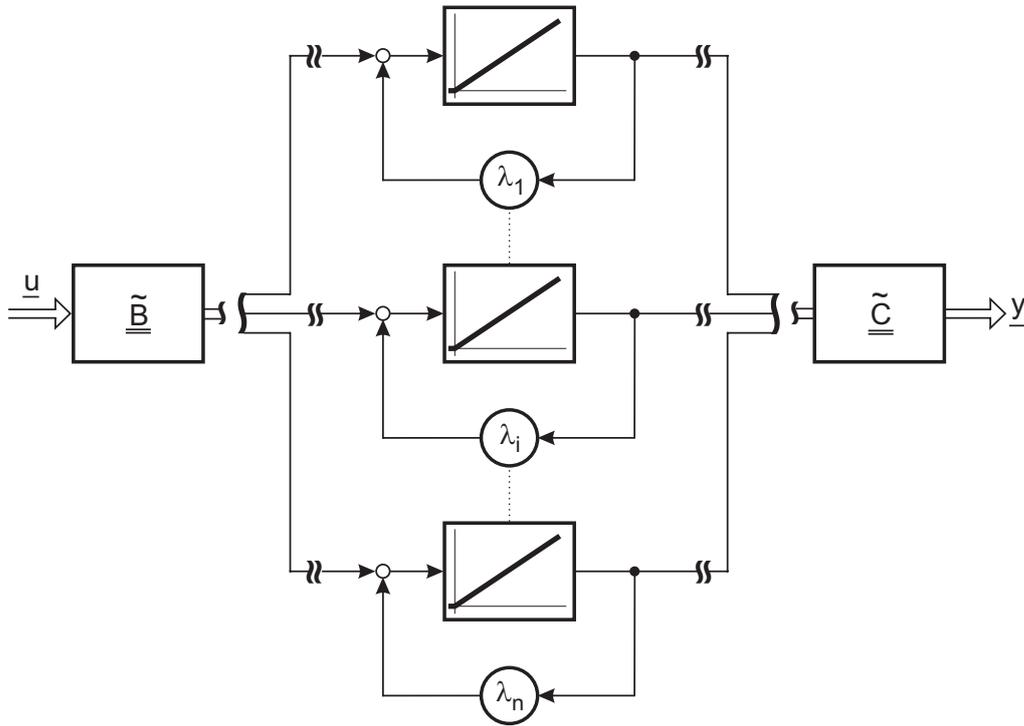


Bild 3.3: Mehrgrößensystem in Modalform

Die übersichtlichsten Verhältnisse ergeben sich für ein Eingrößen-System.

$$\dot{\underline{x}} = \underline{\tilde{A}} \underline{x} + \underline{\tilde{b}} u \quad (3.20)$$

$$y = \underline{\tilde{c}}^T \underline{x}. \quad (3.21)$$

Die Matrizen $\underline{\tilde{B}}$ und $\underline{\tilde{C}}$ werden einspaltig $\underline{\tilde{b}}$ bzw. einzeilig $\underline{\tilde{c}}^T$. Sind zusätzlich alle Eigenwerte reell und die Eigenvektoren verschieden, so entsteht ein Blockschaltbild wie in Bild 3.4.

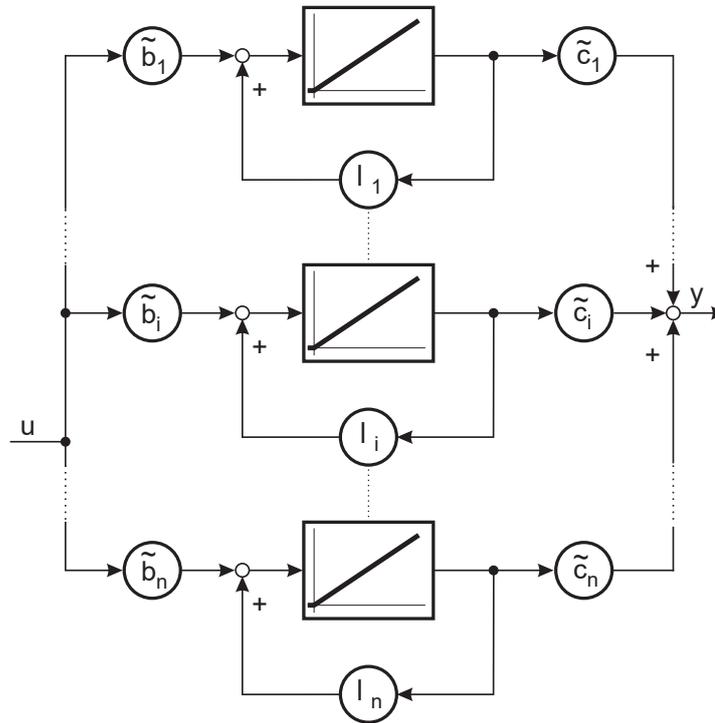


Bild 3.4: Eingrößensystem in Modalform

Anhand dieser Ausführung lassen sich jetzt die Eigenschaften *Steuerbarkeit* und *Beobachtbarkeit* veranschaulichen. Im Zuge der Eigenvektortransformation können entweder einzelne Werte \tilde{b}_i oder \tilde{c}_i zu Null werden. Dann ist in diesem Blockschaltbild offensichtlich, dass ein Ausgleichsvorgang eines Eigenwertes nur dann über den Eingang beeinflusst werden kann, wenn der zugehörige Koeffizient \tilde{b}_i ungleich Null berechnet wurde. Ebenso liefert ein *Eigenvorgang* keinen Beitrag zum Ausgangssignal y , wenn der zugehörige Ausgangskoeffizient Null ist. Unter *Eigenvorgang* oder auch *Eigenbewegung* versteht man den zu einem Eigenwert gehörigen Einschwingvorgang. Man bezeichnet das System als *steuerbar*, wenn der Eingang auf alle Eigenvorgänge wirkt. Entsprechend heißt das System *beobachtbar*, wenn alle Eigenvorgänge zum Ausgangssignal beitragen.

Diese knappe anschauliche Begründung muss noch ergänzt werden, wenn gleiche Eigenwerte vorhanden sind. Dann muss zusätzlich gefordert werden, dass die zugehörigen Koeffizienten verschieden sind. In der Erweiterung auf ein Mehrgrößensystem mit verschiedenen Eigenvektoren entsteht ein Blockschaltbild gemäß Bild 3.3.

Die Bedingungen für Steuerbarkeit und Beobachtbarkeit müssen dann so erweitert werden, dass die Matrix $\underline{\tilde{B}}$ keine Nullzeilen enthält und dass die Zeilen, die zu gleichen Eigenwerten gehören, linear unabhängig sind. Entsprechendes muss für Beobachtbarkeit für die $\underline{\tilde{C}}$ -Matrix gelten.

3.4 Definition von Steuer- und Beobachtbarkeit

Nach der mehr anschaulichen Begründung der Begriffe Steuerbarkeit und Beobachtbarkeit müssen Kriterien für beliebige Systemdarstellungen eingeführt werden. Die Definition nach Kalman lautet:

Ein System heißt *steuerbar*, wenn es zu einem beliebigen Anfangszustand einen Verlauf der Eingangssignale gibt, die den Systemzustand in endlicher Zeit zum Zustand $\underline{0}$ führen.

Aus dieser Definition der Steuerbarkeit können Anforderungen an die Systemmatrix \underline{A} und die Eingangsmatrix \underline{B} abgeleitet werden, die die oben dargestellten speziellen Bedingungen an ein System in Modalform einschließen.

Ein System ist steuerbar, wenn die Steuerbarkeitsmatrix \underline{S}_s den Rang n hat.

$$\underline{S}_s = (\underline{B} \quad \underline{A}\underline{B} \quad \underline{A}^2\underline{B} \quad \dots \quad \underline{A}^{n-1}\underline{B}) \quad (3.22)$$

Die Eigenschaft der Steuerbarkeit heißt bei linearen Systemen auch, dass jeder von $\underline{0}$ verschiedene Zustand transient angesteuert werden kann, jedoch nicht, dass dieser stationär gehalten werden kann.

Eine entsprechende Definition und ein Satz zur Prüfung lassen sich für die Beobachtbarkeit angeben:

Ein System heißt *beobachtbar*, wenn aus dem zeitlichen Verlauf der Ausgangs- und Eingangssignale in einem endlichen Zeitabschnitt der Anfangswert des Systemzustandes zu Beginn des Zeitintervalls berechnet werden kann.

Der zugehörige Satz zur Prüfung der Beobachtbarkeit lautet: Ein System ist beobachtbar, wenn die Beobachtbarkeitsmatrix \underline{S}_B den Rang n hat.

$$\underline{S}_B = \begin{pmatrix} \underline{C} \\ \underline{C}\underline{A} \\ \underline{C}\underline{A}^2 \\ \vdots \\ \underline{C}\underline{A}^{n-1} \end{pmatrix} \quad (3.23)$$

Steuerbarkeit und Beobachtbarkeit sind sofort erfüllt, wenn bereits die \underline{B} bzw. \underline{C} Matrix den Rang n aufweist. Im allgemeinen Fall geht jedoch die \underline{A} -Matrix mit ein. Eine Matrix \underline{A} der Dimension $n \times n$ besitzt maximal n verschiedene Eigenvektoren. Damit in unmittelbarem Zusammenhang steht der *Satz von Cayley-Hamilton*, dass durch mehrfaches Anwenden der Transformation \underline{A} nur eine Überlagerung von n verschiedenen Transformationen entsteht, d. h. jede k -te Potenz von \underline{A} lässt sich als Linearkombination der Potenzen 0 bis $n - 1$ darstellen.

$$\underline{A}^k = \sum_{i=0}^{n-1} l_i \underline{A}^i = \underline{I} + l_1 \underline{A} + \dots + l_{n-1} \underline{A}^{n-1} \quad (3.24)$$

Der Begriff der Beobachtbarkeit offenbart die Tatsache, dass in einem zeitlichen Verlauf eines Signals wesentlich mehr Information über ein System steckt als in einem Augenblickswert. Die Kriterien der Steuerbarkeit und Beobachtbarkeit können verwendet werden, um ein dynamisches System in Teilsysteme zu untergliedern. Zunächst transformiert man die Zustandsgrößen in Jordanform oder Modalform – wenn möglich. Dann stellt man anhand der Besetzung der transformierten Systemmatrizen fest, welche Zustandsgrößen beobachtbar bzw. steuerbar sind. Es entstehen vier Gruppen von Zustandsgrößen:

- \underline{x}_{SB} steuerbar, beobachtbar;
- \underline{x}_{SN} steuerbar, nicht beobachtbar;
- \underline{x}_{NB} nicht steuerbar, beobachtbar;
- \underline{x}_{NN} nicht steuerbar, nicht beobachtbar.

3.5 Realisierbarkeit und minimale Realisierung

Bei linearen dynamischen Systemen stehen zwei Systembeschreibungen nebeneinander. Das System kann einerseits durch die Systemmatrizen in der Zustandsdarstellung beschrieben werden oder durch eine Übertragungsmatrix, die die Kopplung zwischen Eingängen und Ausgängen im Frequenzbereich beschreibt. Bleibt die Frage, inwieweit diese beiden Beschreibungen dieselbe Information über ein System beinhalten und ohne Informationsverlust ineinander umrechenbar sind. Wird ein System in Zustandsform beschrieben, so kann die Wahl der Zustandsgrößen physikalisch motiviert sein. Aus der Zustandsform

$$\dot{\underline{x}} = \underline{A} \underline{x} + \underline{B} \underline{u} \quad (3.25)$$

$$\underline{y} = \underline{C} \underline{x} + \underline{D} \underline{u} \quad (3.26)$$

kann sofort durch die Anwendung der Laplacetransformation bei verschwindenden Anfangsbedingungen die Übertragungsmatrix berechnet werden (vgl. Kapitel 4.3 des Grundlagen-Skripts):

$$\underline{G}(s) = \underline{C}(s \underline{I} - \underline{A})^{-1} \underline{B} + \underline{D}. \quad (3.27)$$

Bei dieser Vorgehensweise kann es vorkommen, dass man beim Aufstellen des Zustandsmodells Zustandsgrößen einführt, die zwar physikalisch begründbar sind, die aber nicht unabhängig voneinander in die Systemdynamik eingehen. Der einfachste Fall dieser Art besteht aus zwei parallelen Zweigen mit jeweils einem Integrator, Bild 3.5. Als Zustandsgrößen sind die stetigen Ausgangsgrößen der Integratoren anzusehen.

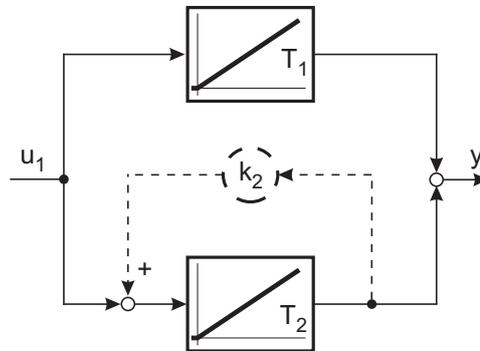


Bild 3.5: Abhängige Zustandsgrößen

$$\frac{1}{T_1 s} + \frac{1}{T_2 s} = \frac{T_2 + T_1}{T_1 T_2 s}. \quad (3.28)$$

Jede beliebige Kombination von Zeitkonstanten und Anfangsbedingungen lässt sich auf einen einzelnen Integrator umrechnen. Insofern ist diese Systembeschreibung redundant. Der Fall ist grundsätzlich anders gelagert, wenn einer der Integratoren mit seinem Ausgang unabhängig vom anderen auf weitere Systemteile wirkt oder sein Eingang eine weitere Aufschaltung erhält.

Das Vorhandensein einer solchen Abhängigkeit lässt sich bei einem SISO-System daran erkennen, dass die Ordnung der Übertragungsfunktion nicht der Anzahl der Zustandsgrößen entspricht. Gleichzeitig entsteht ein Defekt in der Steuerbarkeits- oder Beobachtbarkeitsmatrix oder in beiden.

Ein nicht so offensichtlicher Fall liegt vor, wenn in einer Reihenfolge von Übertragungsgliedern ein Pol gegen eine Nullstelle gekürzt werden kann und keine Signalpfade bestehen, die diese Kompensation nicht enthalten, Bild 3.6.

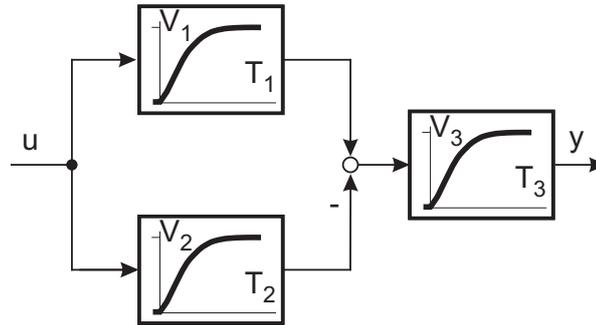


Bild 3.6: Zustandsgrößen mit komplizierter Abhängigkeit

Die Übertragungsfunktion des obigen Beispiels lautet:

$$\begin{aligned}
 G(s) &= V_3 \frac{V_1(T_2 s + 1) - (T_1 s + 1) V_2}{(T_1 s + 1)(T_2 s + 1)(T_3 s + 1)} \\
 &= V_3 \frac{(V_1 T_2 - V_2 T_1) s + V_1 - V_2}{(T_1 s + 1)(T_2 s + 1)(T_3 s + 1)} \\
 &= V_3 (V_1 - V_2) \frac{\frac{V_1 T_2 - V_1 T_1}{V_1 - V_2} s + 1}{(T_1 s + 1)(T_2 s + 1)(T_3 s + 1)}. \tag{3.29}
 \end{aligned}$$

Für den Fall, dass $\frac{V_1 T_2 - V_2 T_1}{V_1 - V_2}$ zufällig mit einer der drei Zeitkonstanten im Nenner zusammenfällt, kann eine Nullstelle gegen einen Pol gekürzt werden, und die Ordnung der Übertragungsfunktion fällt um eins. Auch hier ist das System nicht beobachtbar oder steuerbar.

Generell ist die Beschreibung durch eine Zustandsdarstellung detailreicher, weil auch ein bestimmter Satz von Zustandsgrößen mit festgelegt wird. Dadurch ist die innere Struktur des Systems bekannt. Die Übertragungsmatrix beschreibt nur das Ein-/Ausgangsverhalten des Systems und nicht den inneren Aufbau. Daher lässt sich aus der Übertragungsmatrix zwar ein Satz von Matrizen rekonstruieren, der dem Ein-/Ausgangsverhalten entspricht. Dieser muss aber nichts mit der ursprünglichen Realisierung bzw. der tatsächlichen physikalischen Struktur zu tun haben. Die Konstruktionsvorschriften für Eingrößensysteme wurden bereits vorgestellt.

Wird eine Systembeschreibung als gebrochen-rationale Übertragungsfunktion vorgegeben:

$$G(s) = \frac{b_n s^n + \dots + b_1 s + b_0}{s^n + \dots + a_1 s + a_0}, \tag{3.30}$$

so ist dieser Darstellung erst nach einer Faktorisierung von Zähler und Nenner die Lage der Pole und Nullstellen zu entnehmen. Dabei kann sich herausstellen, dass auch die ursprüngliche Angabe der Übertragungsfunktion nicht minimal war. Wird trotzdem die nicht vereinfachte Übertragungsfunktion durch die Regelungsnormalform oder Beobachtungsnormalform mit einem Satz von Zustandsgrößen realisiert, so stellt sich heraus, dass bei gemeinsamen Nullstellen in Zähler und Nenner das System in Regelungsnormalform steuerbar aber nicht beobachtbar ist, während das System in Beobachtungsnormalform beobachtbar aber nicht steuerbar ist. Dies begründet im Nachhinein die Wahl der Namen sehr gut.

4 Zustandsregelung

Das Prinzip der Zustandsregelung lässt sich am anschaulichsten an einer Eingrößenstrecke erläutern. Man nimmt an, dass alle Zustandsgrößen direkt messbar sind. Stellt man die Strecke in Regelungsnormalform dar, so ergibt sich bei einer Strecke 3. Ordnung ein Blockschaltbild, wie in Bild 4.1 gezeigt.

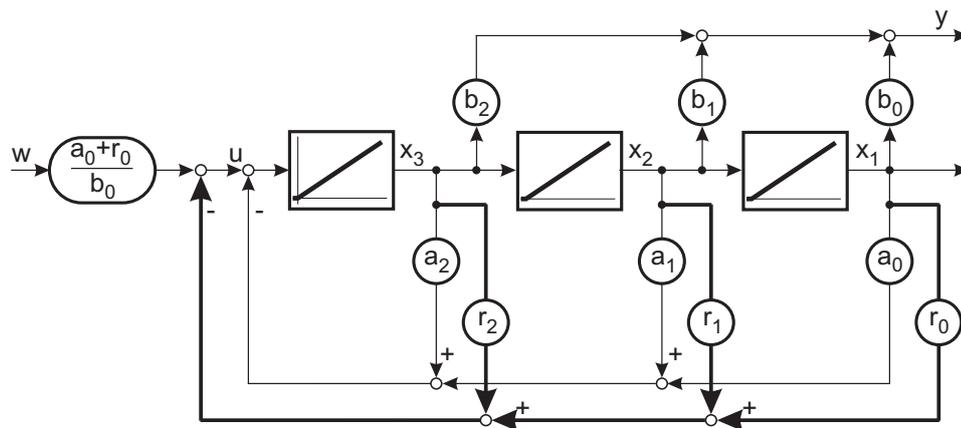


Bild 4.1: Eingrößenzustandsregelung in Regelungsnormalform

Die Übertragungsfunktion der Eingrößenstrecke lässt sich schreiben als:

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0} \quad (4.1)$$

Bei den Überlegungen zu Zustandreglern wird zunächst angenommen, dass die Strecke nicht sprunghaft ist ($b_3 = 0$). Das Nennerpolynom ist gleichzeitig das charakteristische Polynom und bestimmt die Eigenwerte und damit die Stabilität des Systems. In Bild 4.1 ist weiterhin grau hinterlegt eine Zustandsrückführung ($r_2 \dots r_0$) eingezeichnet. Aus dem Blockschaltbild ist ersichtlich, dass die Koeffizienten der Rückführung direkt parallel zu den Koeffizienten des Nennerpolynoms liegen. Die Übertragungsfunktion des Systems mit Rückführung lautet also:

$$\tilde{G}(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + (a_2 + r_2) s^2 + (a_1 + r_1) s + a_0 + r_0} \quad (4.2)$$

Über die Koeffizienten der Rückführung lassen sich somit die Eigenwerte des Systems beliebig verschieben oder durch entsprechende Vorgabe zuweisen. Einzige Voraussetzung ist die vollständige Steuerbarkeit des Systems. Allerdings muss man berücksichtigen, dass

eine starke Verschiebung zu großen Stellamplituden führen kann, die in Abhängigkeit des Systemzustandes auftritt.

Bemerkenswert ist die Tatsache, dass man neben den messbaren Zustandsgrößen nur einen Stelleingriff benötigt, um Zugriff auf alle Eigenwerte zu erlangen. Für die freie Festlegung aller Eigenwerte benötigt man alle Koeffizienten a_0 bis a_{n-1} einschließlich a_0 . Selbst wenn die Strecke vor der Polverschiebung eine stationäre Verstärkung von 1 hatte, kann nach der Beschaltung mit den Rückführkoeffizienten die stationäre Führungsgenauigkeit abhanden gekommen sein. Die Korrektur erfolgt durch einen Steuerfaktor $V = \frac{a_0+r_0}{b_0}$, der die stationäre Führungsgenauigkeit sicherstellt. Für stationäre Störungen findet im Allgemeinen keine Ausregelung statt, da die Zustandsrückführung nur die Lage der Pole modifiziert. Die Störung ist nicht modelliert und kann deshalb auch nicht kompensiert werden. An der Übertragungsfunktion $\tilde{G}(s)$ sieht man auch, dass die Zustandsrückführung keinen Einfluss auf die Nullstellen der Übertragungsfunktion hat. Ein Eingang und n Messstellen reichen aus, um alle Eigenwerte festzulegen.

4.1 Definition von Steuerbarkeit

Voraussetzung für die Anwendung eines Zustandsreglers ist die Steuerbarkeit des zu regelnden Systemes. Die Steuerbarkeit ist dabei wie folgt definiert:

Ein System heißt *steuerbar*, wenn es zu einem beliebigen Anfangszustand einen Verlauf der Eingangssignale gibt, die den Systemzustand in endlicher Zeit zum Zustand $\underline{0}$ führen.

Aus dieser Definition der Steuerbarkeit können Anforderungen an die Systemmatrix \underline{A} und die Eingangsmatrix \underline{B} abgeleitet werden, die die oben dargestellten speziellen Bedingungen an ein System in Modalform einschließen.

Ein System ist steuerbar, wenn die Steuerbarkeitsmatrix \underline{S}_s den Rang n hat.

$$\underline{S}_s = (\underline{B} \quad \underline{A}\underline{B} \quad \underline{A}^2\underline{B} \quad \dots \quad \underline{A}^{n-1}\underline{B}) \quad (4.3)$$

Die Eigenschaft der Steuerbarkeit heißt bei linearen Systemen auch, dass jeder von $\underline{0}$ verschiedene Zustand transient angesteuert werden kann, jedoch nicht, dass dieser stationär gehalten werden kann.

4.2 Zustandsrückführung für Mehrgrößensysteme

Das an einem Eingrößensystem gezeigte Prinzip der Zustandsrückführung lässt sich auch auf Mehrgrößensysteme übertragen. Auch in diesem Fall wird von einem nicht sprungfähigen System ausgegangen ($\underline{D} = \underline{0}$).

$$\dot{\underline{x}} = \underline{A} \underline{x} + \underline{B} \underline{u} \quad (4.4)$$

$$\underline{y} = \underline{C} \underline{x} \quad (4.5)$$

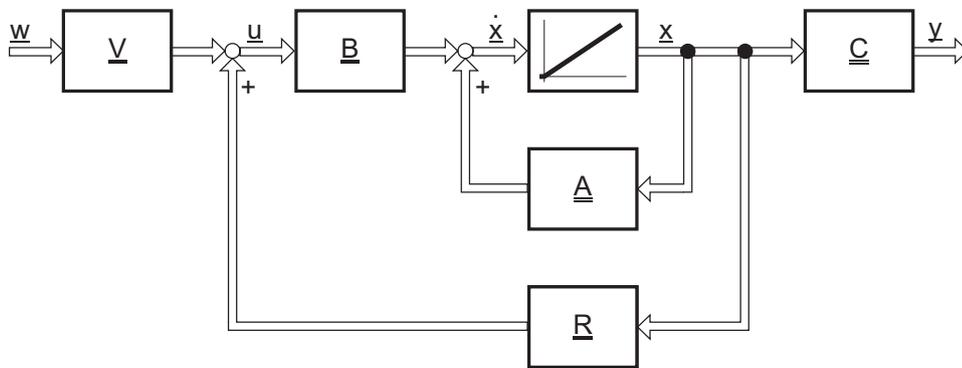


Bild 4.2: Mehrgrößensystem

Unter der Annahme, dass die Zustandsgrößen \underline{x} wiederum vollständig messbar sind, kann mit einer Rückführungsmatrix \underline{R} der Zustand auf den Stellgrößenvektor zurückgeführt werden. Auch im Mehrgrößenfall muss die stationäre Führungsgenauigkeit durch eine Matrix \underline{V} angepasst werden. Das zustandsgeregelte System wird durch folgende Gleichungen beschrieben:

$$\dot{\underline{x}} = (\underline{A} + \underline{B} \underline{R}) \underline{x} + \underline{B} \underline{V} \underline{w} \quad (4.6)$$

$$\underline{y} = \underline{C} \underline{x} \quad (4.7)$$

Die neue Systemmatrix

$$\tilde{\underline{A}} = \underline{A} + \underline{B} \underline{R} \quad (4.8)$$

bestimmt die Eigenwerte.

Unter der Voraussetzung der Steuerbarkeit können wiederum die Eigenwerte vorgegeben werden. Es stehen im Allgemeinen aber mehr Stellgrößen zur Verfügung als zur Gewährleistung der Steuerbarkeit erforderlich werden. Daher gibt es für einen geforderten Satz von Eigenwerten oft mehrere Lösungen zur Wahl der Matrix \underline{R} . Die Eigenwerte des rückgeführten Systems bestimmen sich als Nullstellen des charakteristischen Polynoms

$$\det (s \underline{I} - \underline{A} - \underline{B} \underline{R}) = 0. \quad (4.9)$$

Die Matrix-Übertragungsfunktion von \underline{w} nach \underline{y} berechnet sich zu:

$$\underline{G}(s) = \underline{C} (s \underline{I} - \underline{A} - \underline{B} \underline{R})^{-1} \underline{B} \underline{V} \quad (4.10)$$

Die Bedingung stationärer Genauigkeit ermittelt man aus der Matrix-Übertragungsfunktion und der Grenzwertbetrachtung für $s = 0$, die den stationären Endwert ergibt:

$$\underline{G}(0) = \underline{I} \quad (4.11)$$

Dieses liefert eine Gleichung zur Bestimmung des Führungsfilters:

$$\underline{V} = - [\underline{C} (\underline{A} + \underline{B} \underline{R})^{-1} \underline{B}]^{-1} \quad (4.12)$$

Bleibt also nur noch zu klären, ob die oben angesprochenen Freiheitsgrade in der Wahl von \underline{R} sinnvoll zur Dimensionierung des Regelkreises genutzt werden können.

Wenn das System vollständig steuerbar ist, genügt eigentlich ein Eingang, um die Eigenwerte vollständig festzulegen. Stehen mehrere Eingänge zur Verfügung, so können zusätzliche Eigenschaften des Systems bestimmt werden. Damit können nicht nur die Eigenwerte, sondern auch mehrere oder alle Eigenvektoren vorgegeben werden.

4.3 Ausgangsvektorrückführung

Für den Fall, dass die Zustandsgrößen \underline{x} nicht messbar sind, kann in speziellen Fällen (z. B. \underline{C} invertierbar) anstelle der Zustandsrückführung gemäß Bild 4.2 eine Ausgangsrückführung nach Bild 4.3 verwendet werden.

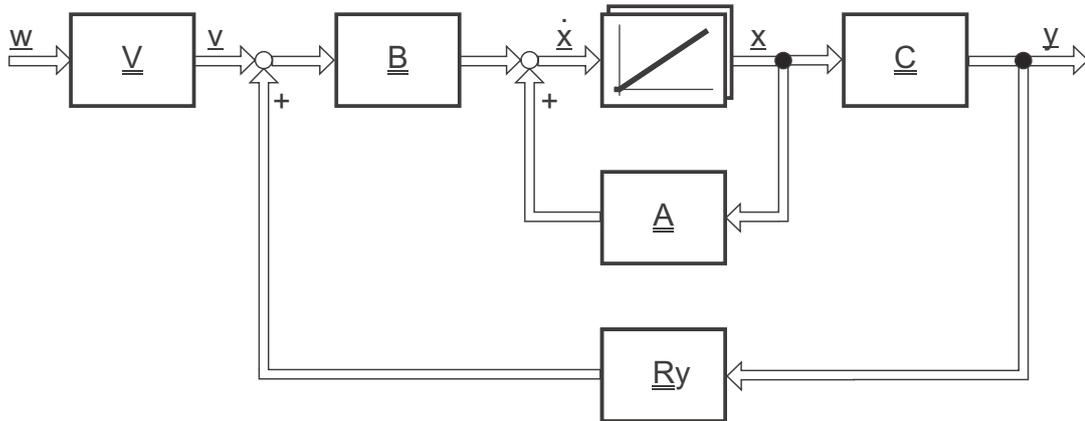


Bild 4.3: Mehrgrößensystem

Durch die Invertierbarkeit von \underline{C} kann zu jedem Zeitpunkt der aktuelle Zustand \underline{x} des Systems berechnet werden. Damit kann mit den Ergebnissen aus Abschnitt 4.2 weitergerechnet werden.

In der Regel ist es allerdings nicht möglich \underline{C} zu invertieren, da hierzu die \underline{C} Matrix natürlich zum einen quadratisch sein muss und zum anderen nicht singular sein darf. Dies ist aber nur in Ausnahmefällen erfüllt, da meist die Anzahl der Ausgangsgrößen kleiner als die Anzahl der Zustandsgrößen ist. Dies hat zur Folge, dass eine Ausgangsvektorrückführung nur in seltenen Fällen möglich ist. Die Forderung nach Invertierbarkeit der Matrix \underline{C} ist allerdings zu weitreichend. Falls eine Zustandsrückführung \underline{R} gefunden wurde, muss lediglich gelten, dass eine Matrix \underline{R}_y existiert, für die gilt:

$$\underline{R}_y \underline{C} = \underline{R} \quad (4.13)$$

Dies ist der Fall wenn gilt:

$$\text{Rang} \left\{ \begin{bmatrix} \underline{C} \\ \underline{R} \end{bmatrix} \right\} = \text{Rang} \{ \underline{C} \} \quad (4.14)$$

4.4 Praktische Anwendung der Polvorgabe

Der im vorherigen Kapitel sehr theoretische Ansatz ist allgemein gültig. Im Anwendungsfall gestaltet sich die Polvorgabe jedoch wesentlich einfacher. Das charakteristische

Polynom eines Mehrgrößensystems ist:

$$P_{char} = \det(s\underline{I} - \underline{A} - \underline{B}\underline{R}) \quad (4.15)$$

Die Matrizen \underline{A} und \underline{B} sind über das System gegeben. Das Polynom weist als Variablen nur die Reglerparameter auf, welche über die Polvorgabe bestimmt werden sollen. Die gewünschten Pole seien s_1 bis s_n . Für die Polvorgabe ergibt sich damit:

$$P_{soll} = \prod_{i=1}^n (s - s_i) \quad (4.16)$$

Durch Koeffizientenvergleich der beiden Polynome und Auflösen nach den Reglerparametern erhält man das Ergebnis. Die Wahl der Pole selber kann nach verschiedenen Mustern erfolgen:

- Festlegung aller Pole als *reeller Mehrfachpol*. Damit ist garantiert, dass kein Überschwingen stattfindet.
- Eine Auslegung nach *Butterworth* weist geringes Überschwingen auf und hat die kleinste Phasennacheilung.
- Eine Auslegung nach *Bessel* hat eine lineare Phase und die geringste Pulsverzerrung.

In jedem Falle muss der letzte verbleibende Parameter bei der Polplatzierung, die Eckfrequenz des Standardfilters, geeignet in Relation zur *Summenzeitkonstante* d. h. *Ersatzzeitkonstante* der Strecke gewählt werden. Bei einer Verschiebung zur sehr großen Frequenzen besteht die Gefahr, dass bei Führungsanregung zu große Stellamplituden erforderlich werden.

5 Beobachter

Für die Reglersynthese des Zustandsreglers wurde bisher davon ausgegangen, dass alle Zustandsgrößen x_i der Regelstrecke messbar sind. Bei realen Anlagen ist dies allerdings häufig nicht der Fall, da Zustandsgrößen nicht zwangsläufig physikalisch messbare Größen sein müssen oder aber eventuell einer Messung nur sehr schwer zugänglich sind.

Als Beispiel sei hier die Biomassebestimmung bei einem bioverfahrenstechnischen Prozess genannt. Der Anteil der Biomasse kann eventuell durch eine Trübungsmessung bestimmt werden. Diese Messung kann allerdings leicht durch Quereinflüsse verfälscht werden. Als Alternative bietet sich an, die Biomasse auszuwiegen. Diese Messung erfordert aber das Ziehen von Proben und eine zeitaufwändige Auswertung im Labor und erfolgt somit nicht on-line. Man kann daher nicht von einer Messung im eigentlichen Sinne sprechen. Es ist deshalb einfacher, die Biomasse über eine *Schätzung* zu bestimmen.

Die Zustandsgrößen x_i lassen sich nicht einfach aus den Ausgangsgrößen y_i durch eine algebraische Auflösung bestimmen, weil in fast allen Fällen die Anzahl der Ausgangsgrößen wesentlich geringer als die Anzahl der Zustandsgrößen ist.

5.1 Zustandsrekonstruktion durch Parallelmodell

Die Schätzung erfolgt dadurch, dass der realen Strecke ein Modell der Regelstrecke parallel geschaltet wird, wie in Bild 5.1 gezeigt. Die Zustandsgrößen können dann aus dem Modell abgegriffen werden.

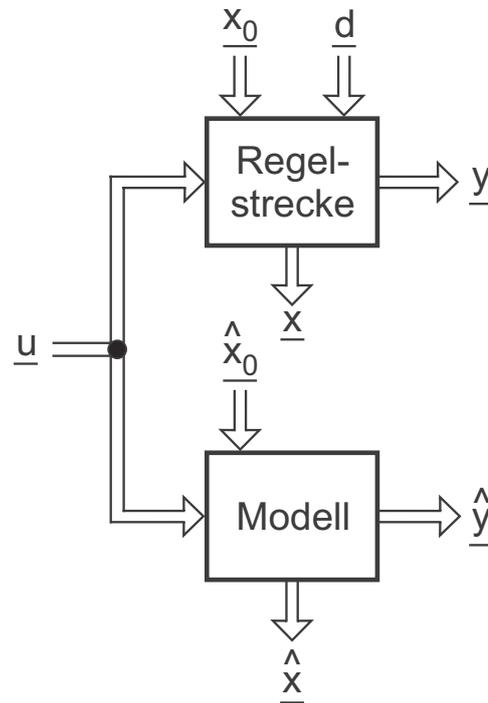


Bild 5.1: Parallelschaltung von Regelstrecke und Streckenmodell

Entspricht das Modell exakt der Strecke, so führt es aufgrund der gleichen Anregung \underline{u} dieselbe Zwangsbewegung wie die Originalstrecke aus. Jedoch ergeben sich aus den folgenden Gründen Unterschiede zwischen dem Ausgangsvektor \underline{y} und dem geschätzten Ausgangsvektor $\hat{\underline{y}}$:

- Unterschiedliche Anfangszustände des Modells $\hat{\underline{x}}_0$ und der realen Strecke \underline{x}_0 führen zu Abweichungen zwischen den Zustandsgrößen \underline{x} und $\hat{\underline{x}}$. Für den Fall einer stabilen Regelstrecke klingen die Eigenbewegungen ab, so dass nach dem Einschwingvorgang $\underline{y} = \hat{\underline{y}}$ und somit auch $\underline{x} = \hat{\underline{x}}$ gilt.
- Störungen \underline{d} , die an der Regelstrecke angreifen, führen zu Veränderungen des Zustandes \underline{x} , jedoch nicht zu Beeinflussungen des geschätzten Zustandes $\hat{\underline{x}}$.

Das Vorgehen nach Bild 5.1 ist daher nur auf stabile Systeme beschränkt und außerdem nur auf ungestörte Systeme anwendbar.

5.2 Luenberger-Beobachter

Als Erweiterung der im vorherigen Abschnitt 5.1 vorgestellten Parallelschaltung eines Modells zur Strecke, bietet sich die Möglichkeit einer Rückkopplung der Differenz der bei-

den Ausgangsvektoren $(\underline{y} - \hat{\underline{y}})$ über eine Matrix \underline{H} auf den Modelleingang. Man spricht hier von einem sogenannten *Luenberger-Beobachter*, der auch *Identitätsbeobachter* genannt wird. Die Struktur zeigt Bild 5.2.

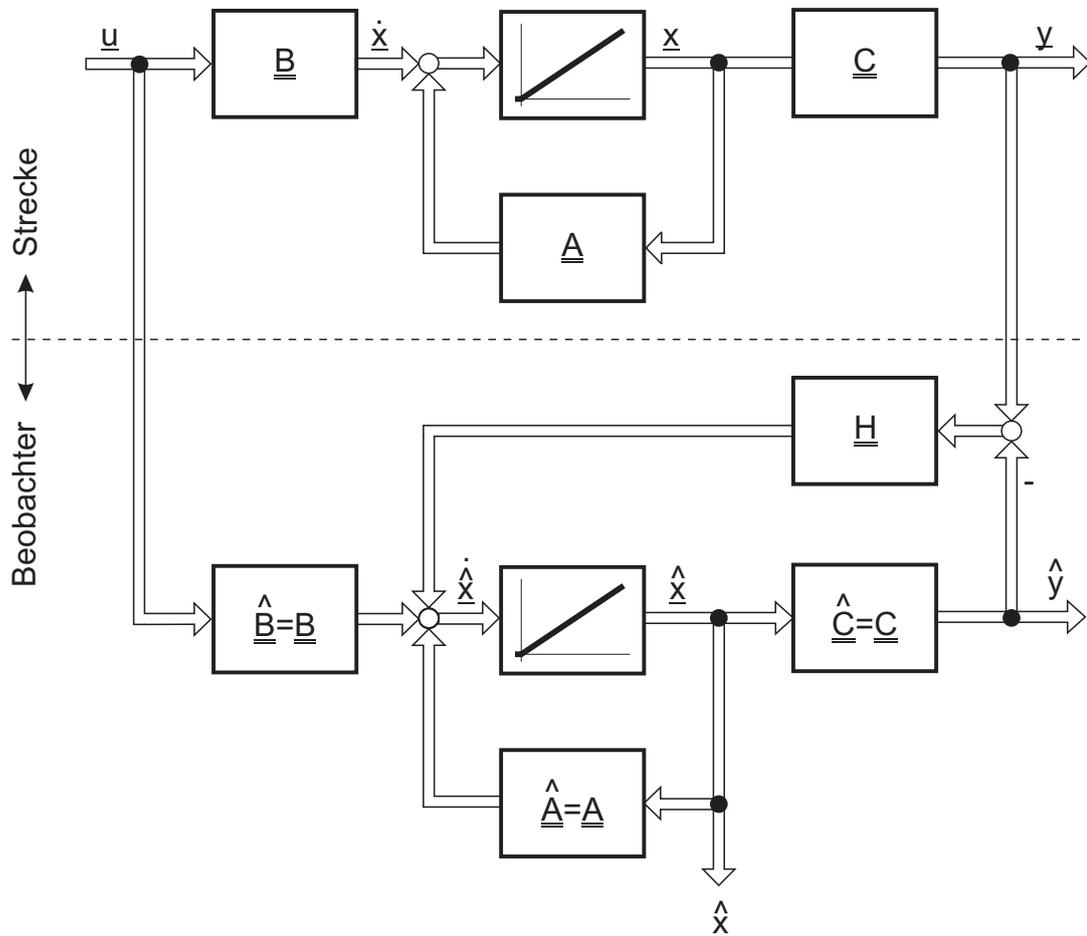


Bild 5.2: Luenberger-Beobachter

Aufgrund der Rückkopplung der Differenz der Ausgangsvektoren gilt im eingeschwungenen Zustand des Beobachters

$$\hat{\underline{y}} \approx \underline{y}. \quad (5.1)$$

Setzt man weiterhin voraus, dass die Matrizen $\hat{\underline{A}} = \underline{A}$, $\hat{\underline{B}} = \underline{B}$ und $\hat{\underline{C}} = \underline{C}$ identisch sind, dann gilt auch

$$\hat{\underline{x}} \approx \underline{x}. \quad (5.2)$$

Aus Bild 5.2 lässt sich unmittelbar die Zustandsgleichung des Beobachters ersehen:

$$\dot{\hat{\underline{x}}} = \underline{A} \hat{\underline{x}} + \underline{B} u + \underline{H} (\underline{y} - \hat{\underline{y}}) \quad (5.3)$$

Mit der Ausgangsgleichung des Beobachters

$$\underline{\hat{y}} = \underline{C} \underline{\hat{x}} \quad (5.4)$$

folgt hiermit

$$\dot{\underline{\hat{x}}} = \underline{A} \underline{\hat{x}} + \underline{B} u + \underline{H} \underline{C} (\underline{x} - \underline{\hat{x}}). \quad (5.5)$$

Anhand dieser Gleichung erkennt man, dass bei Gleichheit von den tatsächlichen und den geschätzten Zustandsgrößen $\underline{x} = \underline{\hat{x}}$ die Beobachtergleichung in die Systemgleichung übergeht. Die *Rückführungsverstärkung* \underline{H} ist also nur wirksam, wenn $\underline{x} \neq \underline{\hat{x}}$ gilt, was im Allgemeinen die Folge von Störungen ist.

Die Differenz zwischen dem wahren und dem geschätzten Zustandsvektor wird *Schätzfehler* genannt und ergibt sich zu:

$$\underline{e} = \underline{x} - \underline{\hat{x}} \quad (5.6)$$

Somit gilt unter obiger Annahme der Übereinstimmung der Matrizen:

$$\begin{aligned} \dot{\underline{e}} &= \frac{d}{dt} (\underline{x} - \underline{\hat{x}}) = \dot{\underline{x}} - \dot{\underline{\hat{x}}} \\ &= \underline{A} \underline{x} + \underline{B} u - \underline{A} \underline{\hat{x}} - \underline{B} u - \underline{H} \underline{C} (\underline{x} - \underline{\hat{x}}) \\ &= (\underline{A} - \underline{H} \underline{C}) (\underline{x} - \underline{\hat{x}}) \\ &= (\underline{A} - \underline{H} \underline{C}) \underline{e}, \quad \underline{e}(0) = \underline{x}_0 - \underline{\hat{x}}_0 \end{aligned} \quad (5.7)$$

Dieser *Beobachtungsfehler* \underline{e} kann als Zustand eines nicht steuerbaren Systems gedeutet werden, da die Gleichung (5.7) unabhängig von der Stellgröße u ist. Besitzen alle Eigenwerte der Matrix $(\underline{A} - \underline{H} \underline{C})$ negativen Realteil, dann gilt

$$\lim_{t \rightarrow \infty} \underline{e}(t) = \underline{0} \quad (5.8)$$

und man erhält im stationären Zustand eine Übereinstimmung des geschätzten Zustandsvektors $\underline{\hat{x}}$ mit dem Zustandsvektor \underline{x} .

In der Matrix $(\underline{A} - \underline{H} \underline{C})$ ist die Verstärkungsmatrix \underline{H} noch frei wählbar, und sie sollte unter folgenden Gesichtspunkten bestimmt werden:

Die Eigenwerte der Matrix $(\underline{A} - \underline{H} \underline{C})$ sollten in der s-Ebene links der Eigenwerte von \underline{A} liegen. Damit wird der Beobachter schneller als die Regelstrecke. Allerdings ist zu

beachten, dass zu weit links gewählte Beobachterpole sich ungünstig auswirken, da sie das Messrauschen von y im Beobachter stark vergrößern. Der Beobachter versucht dann, die Zustandsgrößen durch Differenzieren der Ausgangsgrößen zu bestimmen.

Ist das System $(\underline{A}, \underline{C})$ beobachtbar, können durch eine geeignete Wahl von \underline{H} die Eigenwerte der Matrix $(\underline{A} - \underline{H}\underline{C})$ beliebig verschoben werden, und \underline{H} kann mit den bekannten Entwurfsverfahren der Zustandsrückführung entworfen werden.

5.3 Zustandsrückführung mit Beobachter

Verwendet man einen Beobachter zur Ermittlung der geschätzten Zustandsgrößen und führt man anstelle des wahren Zustandsvektors \underline{x} den Zustandsvektor $\hat{\underline{x}}$ zurück, so ergibt sich folgendes Blockschaltbild:

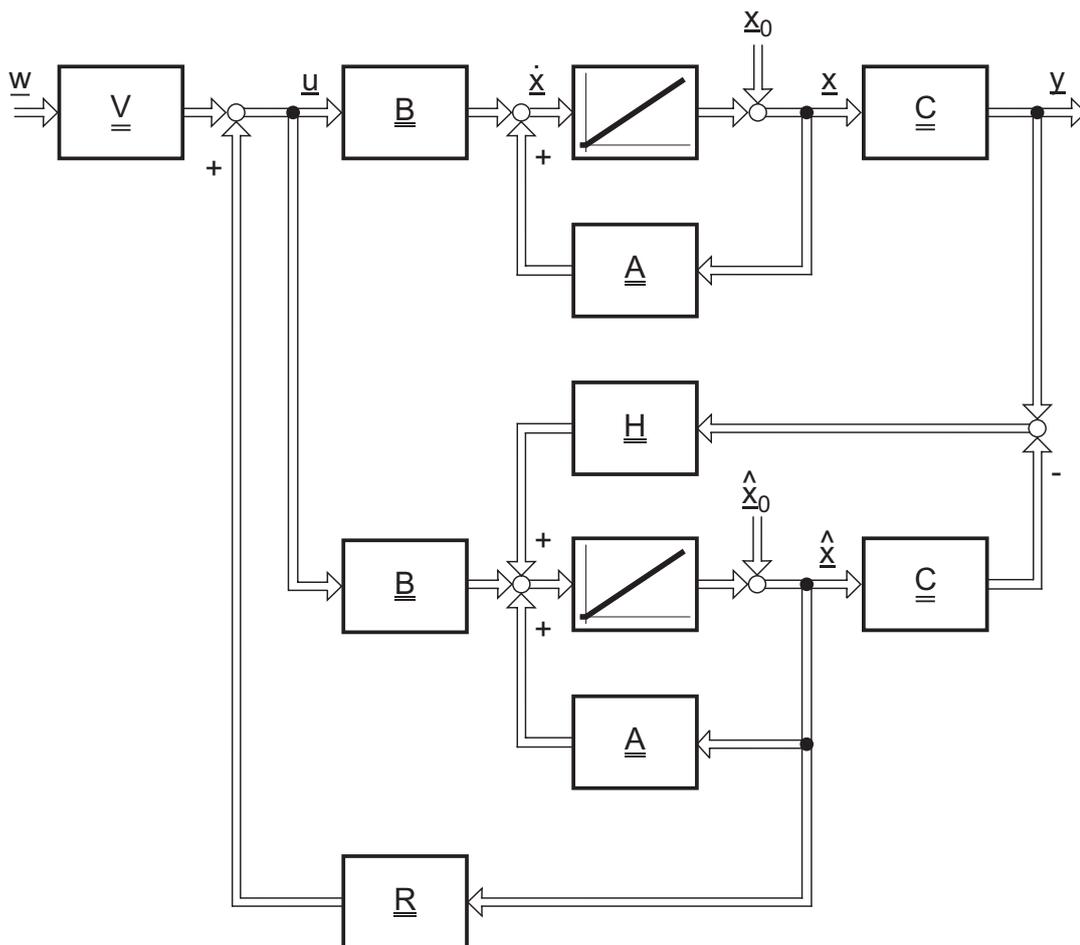


Bild 5.3: Zustandsrückführung mit Beobachter

Anstelle des Regelgesetzes

$$\underline{u}(t) = \underline{R}\underline{x} + \underline{V}\underline{w} \quad (5.9)$$

mit dem realen Zustandsvektor \underline{x} ergibt sich mit dem geschätzten Zustandsvektor $\hat{\underline{x}}$

$$\underline{u}(t) = \underline{R}\hat{\underline{x}} + \underline{V}\underline{w} \quad . \quad (5.10)$$

Das Gesamtsystem besitzt nun die Ordnung $2n$. Die Zustandsraumdarstellung des Gesamtsystems ergibt sich direkt aus Bild 5.3:

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{R}\hat{\underline{x}} + \underline{B}\underline{V}\underline{w} \quad (5.11)$$

$$\dot{\hat{\underline{x}}} = \underline{A}\hat{\underline{x}} + \underline{H}\underline{C}(\underline{x} - \hat{\underline{x}}) + \underline{B}\underline{R}\hat{\underline{x}} + \underline{B}\underline{V}\underline{w} \quad (5.12)$$

Mit dem Schätzfehler

$$\dot{\underline{e}} = \dot{\underline{x}} - \dot{\hat{\underline{x}}} = \underline{A}(\underline{x} - \hat{\underline{x}}) - \underline{H}\underline{C}(\underline{x} - \hat{\underline{x}}) = (\underline{A} - \underline{H}\underline{C})\underline{e} \quad (5.13)$$

und Gleichung (5.11) folgt mit $\hat{\underline{x}} = \underline{x} - \underline{e}$

$$\dot{\underline{x}} = (\underline{A} + \underline{B}\underline{R})\underline{x} - \underline{B}\underline{R}\underline{e} + \underline{B}\underline{V}\underline{w}. \quad (5.14)$$

Die Zustandsgleichung in Matrixform lautet folglich:

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{e}} \end{bmatrix} = \begin{bmatrix} \underline{A} + \underline{B}\underline{R} & -\underline{B}\underline{R} \\ \underline{0} & \underline{A} - \underline{H}\underline{C} \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{e} \end{bmatrix} + \begin{bmatrix} \underline{B}\underline{V} \\ \underline{0} \end{bmatrix} \underline{w} \quad (5.15)$$

Und die Ausgangsgleichung ergibt sich (ohne Durchgriff) zu:

$$\underline{y} = \begin{bmatrix} \underline{C} & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{e} \end{bmatrix} \quad (5.16)$$

Da die angegebene Systemmatrix $\begin{bmatrix} \underline{A} + \underline{B}\underline{R} & -\underline{B}\underline{R} \\ \underline{0} & \underline{A} - \underline{H}\underline{C} \end{bmatrix}$ eine Blockdreiecksmatrix ist, setzen sich ihre Eigenwerte aus denen der Matrizen $(\underline{A} + \underline{B}\underline{R})$ und $(\underline{A} - \underline{H}\underline{C})$ zusammen, denn für die charakteristische Gleichung gilt:

$$\det \left\{ s\underline{I} - \begin{bmatrix} \underline{A} + \underline{B}\underline{R} & -\underline{B}\underline{R} \\ \underline{0} & \underline{A} - \underline{H}\underline{C} \end{bmatrix} \right\} = 0 \quad (5.17)$$

Formt man diese Gleichung um, so erkennt man die Zusammensetzung der charakteristischen Gleichung aus den beiden Teilsystemen:

$$\underbrace{\det \{s\underline{I} - \underline{A} - \underline{B}\underline{R}\}}_{\substack{\text{System ohne} \\ \text{Beobachter} \\ \text{mit Rückführung}}} \cdot \underbrace{\det \{s\underline{I} - \underline{A} + \underline{H}\underline{C}\}}_{\text{Beobachter}} = 0 \quad (5.18)$$

Dies lässt sich im sogenannten *Separationstheorem* zusammenfassen:

Sofern das durch die Matrizen \underline{A} , \underline{B} und \underline{C} vorgegebene offene System vollständig steuerbar und beobachtbar ist, können die n Eigenwerte der charakteristischen Gleichung des Beobachters und die n Eigenwerte der charakteristischen Gleichung des geschlossenen Systems (ohne Beobachter) separat vorgegeben werden.

Unter der Voraussetzung der Steuerbarkeit und Beobachtbarkeit der Regelstrecke, kann die Zustandsrückführung völlig unabhängig entworfen werden, unabhängig davon, ob sie mit oder ohne Beobachter realisiert wird. Bei diesem Entwurfsschritt werden die Eigenwerte von $(\underline{A} + \underline{B}\underline{R})$ festgelegt.

Unabhängig von der Rückführung kann man den Beobachter entwerfen, sofern die Regelstrecke vollständig beobachtbar ist. Man legt hierbei die Eigenwerte von $(\underline{A} - \underline{H}\underline{C})$ fest.

Aus dem Separationstheorem ist auch ersichtlich, dass die Eigenwerte eines Beobachters, der zusammen mit einer Zustandsrückführung betrieben wird, nicht in Bezug auf die Eigenwerte der Regelstrecke, sondern in Relation zu den Eigenwerten der Systemmatrix $(\underline{A} + \underline{B}\underline{R})$ des zustandsrückgeführten Systems gewählt werden müssen.

5.4 Störgrößenbeobachter

Im Kapitel 2 wurde gezeigt, dass mit einer Störgrößenaufschaltung ein günstigeres Systemverhalten erzielt werden kann, für den Fall, dass die Störungen messbar sind.

Hieraus leitet sich der Wunsch ab, einerseits auch in einer Zustandsregelung eine Verbesserung des dynamischen Verhaltens durch eine Störgrößenaufschaltung zu erzielen und andererseits auch nicht messbare Störungen mittels eines sogenannten *Störgrößenbeobachters* zu rekonstruieren.

Für die allgemeine Zustandsbeschreibung erhält man beim Vorhandensein einer Störung unter der Annahme, dass kein Durchgriff der Stellgröße vorhanden ist, die Zustandsgleichung

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{u} + \underline{d} \quad (5.19)$$

und die Ausgangsgleichung

$$\underline{y} = \underline{C}\underline{x} \quad . \quad (5.20)$$

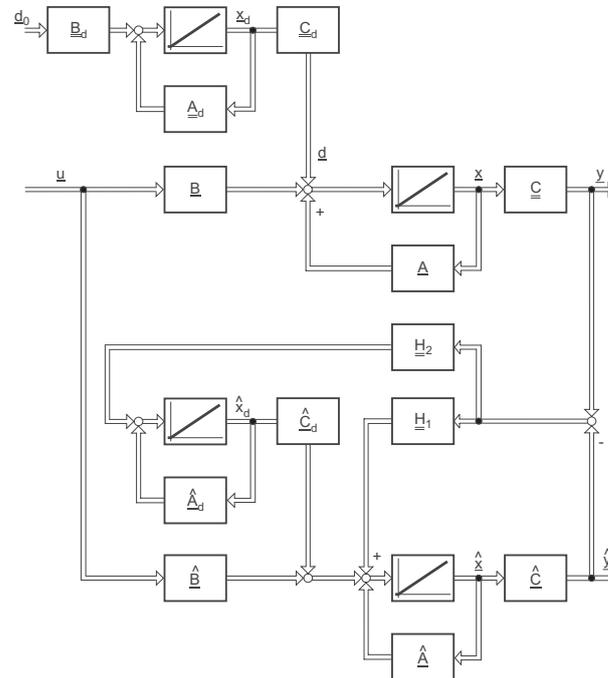


Bild 5.4: Beobachter mit integriertem Störgrößenbeobachter

Der *Störgrößenbeobachter* schätzt aus dem Verlauf der Stellgröße \underline{u} und der Messgröße \underline{y} die unbekannte Störung \underline{d} . Die Störgrößen müssen beobachtbar sein, sonst hätten sie auch keinen Einfluss auf die Regelgröße. Daher kann diese Eigenschaft direkt vorausgesetzt werden. Jedoch muss zusätzlich bekannt sein, um welche Klasse von Störungen es sich handelt. Dabei werden konstante Störungen als offene Integratoren modelliert, und sinusförmige Störungen als ungedämpfte Schwinger der entsprechenden Frequenz. Der Störgrößenbeobachter schätzt im ersten Fall die Amplitude der konstanten Störung, dieses natürlich auch mit einer gewissen Verzögerung, wenn sich die Höhe der Störung im Betrieb ändert. Bei einer Dauerschwingung entsteht ein Teilmodell zweiter Ordnung, damit ermittelt der Beobachter dann Betrag und Phase der Schwingung. Die Frequenz der Schwingung ist dabei im Modell sozusagen hart festgelegt und strenggenommen kann der Beobachter nur Störungen mit genau dieser Frequenz korrekt schätzen. Falls nur ein gewisser Frequenzbereich bekannt ist, in dem eine sinusförmige Störung auftritt, so muss das Modell mit einer Dämpfung versehen werden, mit dem Effekt, dass die Schätzung stationäre Fehler aufweist und die Störung nicht exakt kompensiert werden kann. Das Störmodell wird eingangsseitig nicht angeschlossen. Dadurch geht man implizit davon aus, dass nur die Anfangszustände vom Beobachter geschätzt werden müssen. Das Stör-

größenmodell

$$\dot{\underline{x}}_d = \underline{A}_d \underline{x}_d(t), \quad \underline{x}_d(0) = \underline{x}_{d0} \quad (5.21)$$

$$\underline{d} = \underline{C}_d \underline{x}_d \quad (5.22)$$

generiert das Signal \underline{d} .

Das Systemmodell wird um das Störmodell erweitert

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{x}}_d \end{bmatrix} = \begin{bmatrix} \underline{A} & \underline{C}_d \\ \underline{0} & \underline{A}_d \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{x}_d \end{bmatrix} + \begin{bmatrix} \underline{B} \\ \underline{0} \end{bmatrix} u \quad (5.23)$$

Mit den Anfangswerten

$$\begin{bmatrix} \underline{x}(0) \\ \underline{x}_d(0) \end{bmatrix} = \begin{bmatrix} \underline{x}_0 \\ \underline{x}_{d0} \end{bmatrix} \quad (5.24)$$

und der Ausgangsgleichung

$$\underline{y}(t) = \begin{bmatrix} \underline{C} & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{x}_d \end{bmatrix} \quad (5.25)$$

lässt sich das Gesamtverhalten der Regelstrecke beschreiben.

Führt man den durch den Zustandsbeobachter geschätzten Zustandsvektor zurück und schaltet man zusätzlich die geschätzte Störgröße des Störgrößenbeobachters auf den Eingang des Systems, so erhält man das in Bild 5.5 dargestellte System.

Man erhält damit folgende Zustandsgleichungen des Gesamtsystems:

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{x}}_d \end{bmatrix} = \begin{bmatrix} \underline{A} & \underline{C}_d \\ \underline{0} & \underline{A}_d \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{x}_d \end{bmatrix} + \begin{bmatrix} \underline{B} \\ \underline{0} \end{bmatrix} u + \begin{bmatrix} \underline{H} \underline{C} \\ \underline{0} \end{bmatrix} (\underline{x} - \hat{\underline{x}}) \quad (5.26)$$

$$\hat{\underline{d}} = \begin{bmatrix} \underline{0} & \underline{C}_d \end{bmatrix} \begin{bmatrix} \hat{\underline{x}} \\ \hat{\underline{x}}_d \end{bmatrix} \quad (5.27)$$

wobei die um die Störgrößenaufschaltung erweiterte Zustandsrückführung durch

$$\underline{u} = \underline{R} \hat{\underline{x}} + \underline{R}_d \hat{\underline{x}}_d \quad (5.28)$$

in die obige Gleichung eingeht.

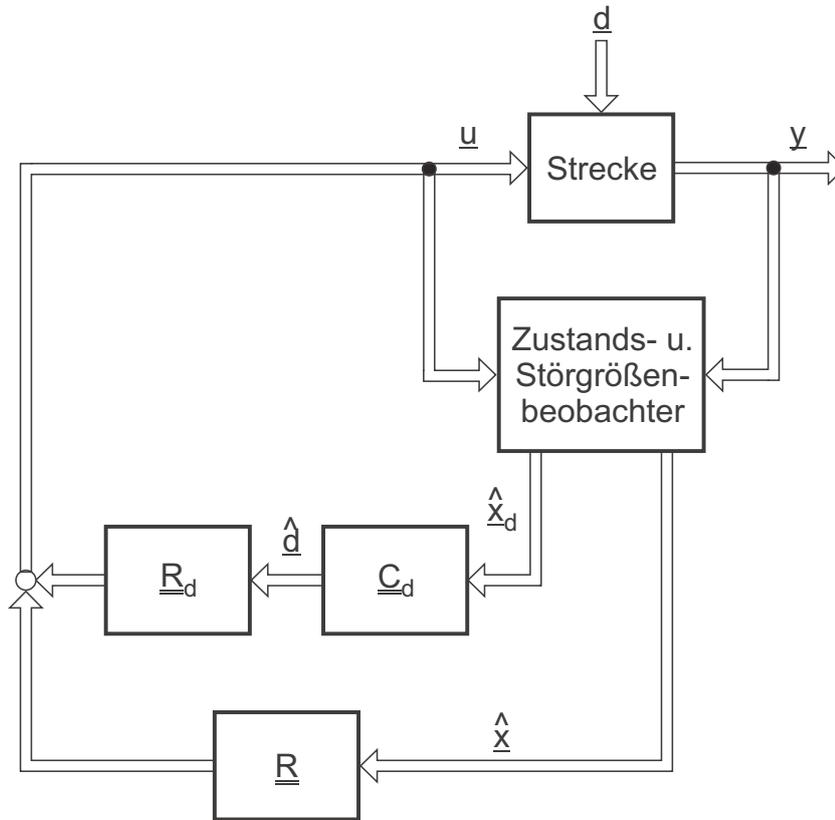


Bild 5.5: Zustandsrückführung und Störgrößenaufschaltung über einen Zustands- und Störgrößenbeobachter

Die Matrix \underline{R}_d muss so gewählt werden, dass die Wirkung der Störung exakt kompensiert wird, also gilt:

$$\underline{B}\underline{R}_d = -\underline{C}_d \quad (5.29)$$

5.5 Polvorgabe für Beobachter

Entsprechend der Vorgehensweise aus Kap. 4.4, können auch die Pole des Beobachters vorgegeben werden. Betrachten wir folgendes Beobachter-System ohne Rückführung (siehe Luenberger-Beobachter):

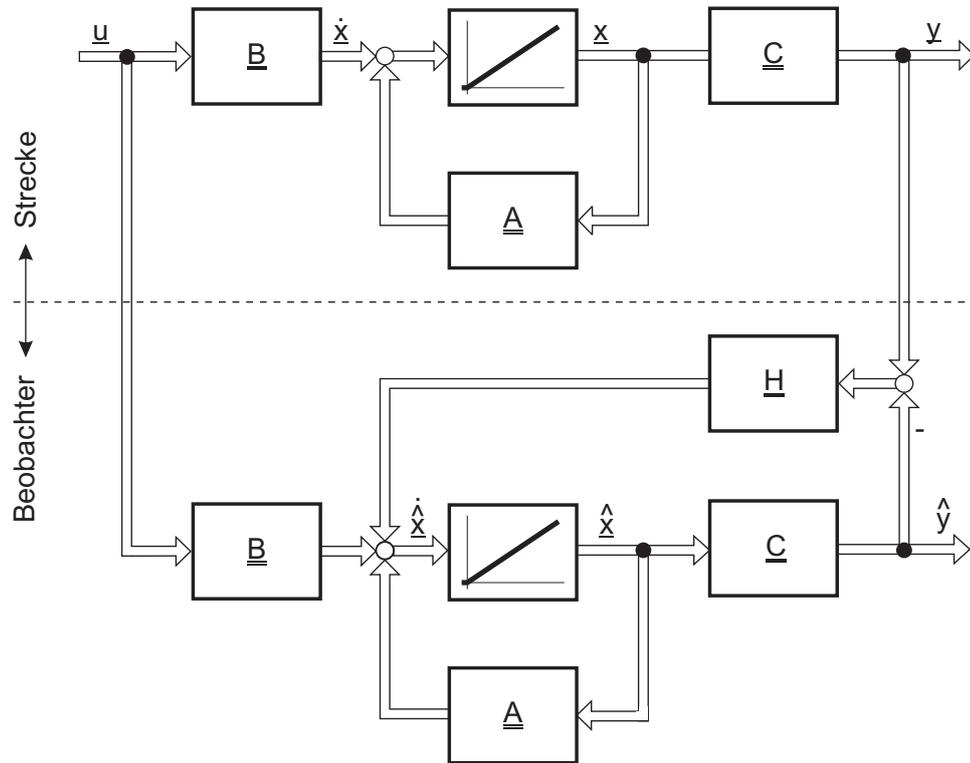


Bild 5.6: Luenberger-Beobachter ohne Rückführung

Aufgrund des Separationsansatzes gilt, dass das Beobachtersystem immer dann unabhängig vom Reglersystem betrachtet werden kann, wenn das Modell der Strecke exakt im Beobachter implementiert ist. Das charakteristische Polynom des Beobachters ergibt sich zu:

$$P_{char} = \det(s\underline{I} - \underline{A} + \underline{H}\underline{C}) \quad (5.30)$$

Dieses kann nun, genau wie im Fall des Reglers im vorherigen Kapitel, durch Koeffizientenvergleich mit der Polvorgabe abgeglichen werden, woraus sich die gesuchten Werte für \underline{H} ergeben.

6 Optimale Regelung

Man erwartet von einer Regelung im Allgemeinen, dass sie einen vorgegebenen Sollwert in kurzer Zeit genau einstellt und Störungen weitgehend unterdrückt. Daher ist es naheliegend, die zu bestimmenden Parameter im Rahmen einer *Optimierung* so zu bestimmen, dass dieses Ziel bestmöglich erreicht wird. Dazu ist die mathematische Formulierung eines Gütekriteriums erforderlich, das die Annäherung an den optimalen Zustand quantifiziert. Die quadratische Abweichung zwischen dem stationären Endwert der Regelgröße und dem Verlauf der Sprungantwort wird als Gütemaß verwendet. Dabei stellt sich jedoch schnell heraus, dass das Problem so noch nicht wohldefiniert ist, da durch beliebig große Stellamplituden bei einer linearen Strecke eine beliebig schnelle Annäherung erreicht werden kann. Es ist ein Gegenspieler im Gütekriterium J erforderlich, der verhindert, dass übermäßig große Stellamplituden verwendet werden. Dabei kann wegen der Linearität der Strecke statt eines endlichen Sollwertes immer ein Regelvorgang von einem Anfangswert \underline{x}_0 zum Sollwert $\underline{0}$ betrachtet werden. Wird mit J der Verlauf der Sprungantwort bis zu einem willkürlich gewählten Endzeitpunkt t_e gemäß dieser Überlegungen quadratisch bewertet, so entsteht folgender Ausdruck:

$$J_e(\underline{x}_0, \underline{u}(t)) = \underline{y}^T(t_e) \underline{S} \underline{y}(t_e) + \int_0^{t_e} [\underline{y}^T(t) \underline{Q}_y \underline{y}(t) + \underline{u}^T(t) \underline{Q}_u \underline{u}(t)] dt . \quad (6.1)$$

Dabei sind \underline{S} , \underline{Q}_y und \underline{Q}_u frei wählbare positive definite Wichtungsmatrizen, die in der Regel nur auf der Hauptdiagonalen besetzt sind. Neben dem Verlauf der Sprungantwort muss auch die Regelabweichung am Ende des Zeitraums extra bewertet werden.

Ziel der *optimalen Regelung* ist es, den Verlauf der Stellgröße $\underline{u}(t)$ zu finden, bei dem das Gütefunktional minimal wird. Man spricht auch von der Lösung des *Optimierungsproblems*: $\min_{\underline{u}(t)} J_e(\underline{x}_0, \underline{u})$.

Wird das betrachtete Zeitintervall bis unendlich ausgedehnt, so kann auf die Bewertung des Endwertes $\underline{y}^T(t_e)$ verzichtet werden, weil das Gütefunktional nur konvergiert, wenn der Endwert $\underline{0}$ exakt erreicht wird.

Dadurch schreibt sich das Funktional deutlich kompakter und man benötigt eine Wichtungsmatrix weniger:

$$J(\underline{x}_0, \underline{u}) = \int_0^{\infty} [\underline{y}^T(t) \underline{Q}_y \underline{y}(t) + \underline{u}^T(t) \underline{Q}_u \underline{u}(t)] dt . \quad (6.2)$$

In der jetzt vorliegenden Form ist das Problem eigentlich gar kein Reglerentwurf mehr, sondern eine Steuerungsaufgabe, bei der es darum geht, für eine gegebene lineare Strecke,

die sich im Anfangszustand $\underline{0}$ befindet, einen Verlauf der Stellgröße zu finden der die Regelgröße optimal zu $\underline{0}$ führt. Man bezeichnet es auch als LQ-Problem (*linear-quadratisches Optimierungsproblem*):

$$\min_{\underline{u}(t)} J(\underline{x}_0, \underline{u}) . \quad (6.3)$$

In Bild 6.1 ist der Aufbau des Problems dargestellt. Eine optimale Steuereinheit generiert für einen gegebenen Anfangszustand den Verlauf der Stellgröße $u(t)$, der das Gütefunktional minimiert.

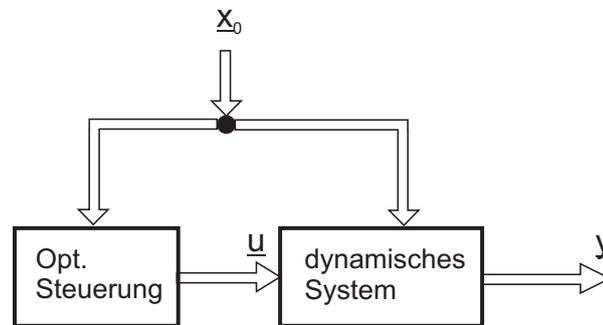


Bild 6.1: Steuerungssystem

6.1 Lösung des LQ-Problems

(nach Lunze [4]) Die Lösung des Steuerungsproblems kann verwendet werden, um ein Problem der optimalen Regelung zu lösen. Daher wird das Steuerungssystem aus Bild 6.1 in ein Regelungssystem umgeformt, indem die Stellgröße als Ergebnis einer Zustandsrückführung aus den Zustandsgrößen des Systems dargestellt wird:

$$\underline{u}(t) = \underline{R} \underline{x}(t) \quad (6.4)$$

Die Stellgröße $\underline{u}(t)$ wird aus dem zurückgeführten Zustand und nicht mehr allein aus dem Anfangswert berechnet. Weil das Funktional (6.2) quadratisch und die Regelstrecke (6.4) linear ist, spricht man auch von *linear-quadratischer Regelung* (LQ-Regelung) oder *Optimaler Regelung*.

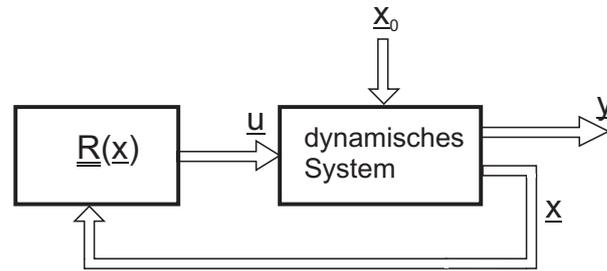


Bild 6.2: Regelungssystem

Damit vereinfacht sich das Gütefunktional, weil die Ausgangsgröße $\underline{y} = \underline{C}\underline{x}$ wie auch die Stellgröße $u = \underline{R}\underline{x}$ aus dem Zustand berechnet werden. Damit hängt das Integral J nur noch vom Anfangszustand ab:

$$J(x_0) = \int_0^{\infty} [\underline{x}^T \underline{C}^T \underline{Q}_y \underline{C} \underline{x} + \underline{x}^T \underline{R}^T \underline{Q}_u \underline{R} \underline{x}] dt . \quad (6.5)$$

mit Einführung der positiv definiten Matrix \underline{Q}_x

$$\underline{Q}_x = \underline{C}^T \underline{Q}_y \underline{C} + \underline{R}^T \underline{Q}_u \underline{R}. \quad (6.6)$$

ergibt sich das vereinfachte Funktional J :

$$J = \int_0^{\infty} \underline{x}^T(t) \underline{Q}_x \underline{x}(t) dt . \quad (6.7)$$

Dieses Integral soll zunächst für ein stabiles System mit der Systemmatrix \underline{A} ohne Anregung ausgewertet werden. Das Ergebnis lässt sich dann auf das geregelte System übertragen. Die Zustandsgröße verläuft für homogene Einschwingvorgänge nach der Fundamentalfunktion:

$$\underline{x}(t) = \underline{x}_0 e^{\underline{A}t} . \quad (6.8)$$

Hierbei ist \underline{x}_0 der Anfangszustand.

$$J = \int_0^{\infty} \underline{x}_0^T e^{\underline{A}^T t} \underline{Q} e^{\underline{A} t} \underline{x}_0 dt \quad (6.9)$$

$$= \underline{x}_0^T \underline{P} \underline{x}_0 \quad (6.10)$$

$$\text{mit } \underline{P} = \int_0^{\infty} e^{\underline{A}^T t} \underline{Q} e^{\underline{A} t} dt \quad (6.11)$$

Über partielle Integration ergibt sich:

$$\underline{P} = \left[e^{\underline{A}^T t} \underline{Q} \underline{A}^{-1} e^{\underline{A} t} \right]_0^{\infty} - \int_0^{\infty} \underline{A}^T e^{\underline{A}^T t} \underline{Q} \underline{A}^{-1} e^{\underline{A} t} dt . \quad (6.12)$$

Da die Matrix \underline{A} ein stabiles System beschreibt und somit nur Eigenwerte mit negativem Realteil aufweist, gelten die Grenzwerte:

$$\lim_{t \rightarrow \infty} e^{\underline{A} t} = \underline{0} \quad \text{und} \quad \lim_{t \rightarrow 0} e^{\underline{A} t} = \underline{I} \quad (6.13)$$

Damit berechnet sich die Matrix \underline{P} zu:

$$\begin{aligned} \underline{P} &= -\underline{Q} \underline{A}^{-1} - \underline{A}^T \int_0^{\infty} e^{\underline{A}^T t} \underline{Q} e^{\underline{A} t} dt \underline{A}^{-1} \\ &= -\underline{Q} \underline{A}^{-1} - \underline{A}^T \underline{P} \underline{A}^{-1} \end{aligned} \quad (6.14)$$

Diese Beziehung von rechts mit \underline{A} multipliziert und anschließend umgestellt ergibt die Ljapunov Gleichung:

$$\underline{A}^T \underline{P} + \underline{P} \underline{A} = -\underline{Q} \quad (6.15)$$

Die Stabilitätseigenschaft eines Systems ist äquivalent mit der Tatsache dass für die Ljapunov Gleichung mit einer positiv definiten Matrix \underline{Q} eine Lösung \underline{P} existiert.

6.2 Anwendung der optimalen Regelung

Betrachten wir nun die Anwendung auf ein geregeltes System inkl. Rückführung:

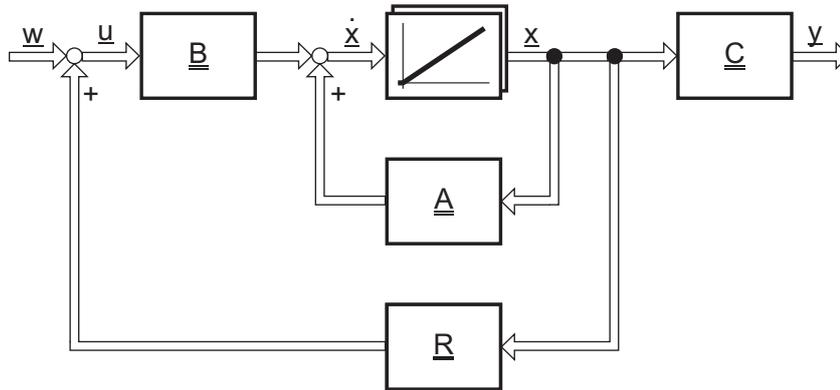


Bild 6.3: Mehrgrößensystem mit Rückführung

In unserem Fall sei die Eingangsgröße $w = 0$. Mit $\tilde{\underline{A}} := \underline{A} + \underline{B}\underline{R}$ stellt sich das System folgendermaßen dar, und entspricht damit Bild 6.2 aus dem vorherigen Kapitel:

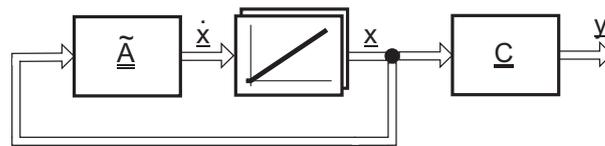


Bild 6.4: Vereinfachtes Mehrgrößensystem

Für das Gütefunktional

$$J = \int_0^{\infty} (\underline{x}^T(t) \underline{Q} \underline{x}(t) + \underline{u}^T(t) \underline{Q} \underline{u}(t)) dt \quad (6.16)$$

gilt mit $\underline{u} = \underline{R}\underline{x}$ (siehe Bild 6.3):

$$J = \int_0^{\infty} \underline{x}^T(t) \underbrace{[\underline{Q} + \underline{R}^T \underline{Q} \underline{R}]}_{\tilde{\underline{Q}}} \underline{x}(t) dt . \quad (6.17)$$

Daraus folgt bei äquivalentem Vorgehen wie zuvor:

$$J = \underline{x}_0^T \underline{P} \underline{x}_0 \quad \text{mit} \quad \underline{P} = \int_0^{\infty} e^{\tilde{\underline{A}}^T t} \tilde{\underline{Q}} e^{\tilde{\underline{A}} t} dt \quad (6.18)$$

sowie die Ljapunow-Gleichung

$$\underline{\tilde{A}}^T \underline{P} + \underline{P} \underline{\tilde{A}} = -\underline{\tilde{Q}}_{\underline{x}} . \quad (6.19)$$

Da sowohl $\underline{\tilde{A}} = \underline{A} + \underline{B} \underline{R}$ als auch $\underline{\tilde{Q}}_{\underline{x}} = \underline{Q}_{\underline{x}} + \underline{R}^T \underline{Q}_{\underline{u}} \underline{R}$ gelten, ergibt sich:

$$(\underline{A}^T + \underline{B}^T \underline{R}^T) \underline{P} + \underline{P} (\underline{A} + \underline{B} \underline{R}) = -\underline{Q}_{\underline{x}} - \underline{R}^T \underline{Q}_{\underline{u}} \underline{R} . \quad (6.20)$$

Die Reglermatrix \underline{R} ist die zu bestimmende Größe. Sie muss so gewählt werden, dass J minimal wird. Die notwendige Bedingung hierzu lautet:

$$\frac{dJ}{dr_{ij}} \stackrel{!}{=} 0 . \quad (6.21)$$

Nach (6.10) ergibt sich:

$$\frac{dJ}{dr_{ij}} = \frac{d}{dr_{ij}} [x_0^T \underline{P} x_0] \quad (6.22)$$

was äquivalent ist zu

$$\frac{d\underline{P}}{dr_{ij}} = 0 \quad (6.23)$$

da die Optimale Regelung nach Voraussetzung unabhängig von \underline{x}_0 ist. Die Ableitung der Ljapunow-Gleichung stellt sich somit dar als:

$$\frac{d\underline{R}^T}{dr_{ij}} (\underline{Q}_{\underline{u}} \underline{R} + \underline{B}^T \underline{P}) + (\underline{R}^T \underline{Q}_{\underline{u}} + \underline{P} \underline{B}) \frac{d\underline{R}}{dr_{ij}} = \underline{0} . \quad (6.24)$$

Wird \underline{R} nach r_{ij} abgeleitet, so ergibt sich eine $(m \times n)$ -Matrix, die fast nur aus Nullen besteht. Lediglich an der Position (i, j) ist eine Eins zu finden. Es gilt:

$$\underline{Q}_{\underline{u}} \underline{R} + \underline{B}^T \underline{P} \stackrel{!}{=} \underline{0} \quad (6.25)$$

$$\Rightarrow \underline{R} = -\underline{Q}_{\underline{u}}^{-1} \underline{B}^T \underline{P} \quad (6.26)$$

Letzteres eingesetzt in (6.20) führt zu:

$$\underline{A}^T \underline{P} + \underline{P} \underline{A} - \underline{P} \underline{B} \underline{Q}_{\underline{u}}^{-1} \underline{B}^T \underline{P} + \underline{Q}_{\underline{x}} = \underline{0} . \quad (6.27)$$

Dies ist eine Matrix-Riccati-Gleichung, welche sich nach \underline{P} auflösen lässt. Das Ergebnis für \underline{P} wird wiederum in (6.26) eingesetzt, wodurch sich die gesuchte R -Matrix berechnen lässt.

Zusammengefasst ergibt sich zur Auslegung eines *optimalen Reglers* folgende Vorgehensweise:

1. Wahl der Wichtungsmatrizen \underline{Q}_u und \underline{Q}_x . Die Matrix \underline{Q}_u muss positiv definit, die Matrix \underline{Q}_x positiv-semidefinit sein, in der Regel sind sie nur auf der Hauptdiagonalen besetzt.
2. Mittels der Wichtungsmatrizen muss die Riccati-Gleichung (6.27) gelöst werden, um die Matrix \underline{P} zu erhalten
3. Über (6.26) wird schließlich die Reglermatrix \underline{R} bestimmt.

Mit der Vorgabe der Wichtungsmatrizen kann ein für die gewählte Parametrierung optimaler Regler berechnet werden. Die Übersetzung der Vorgaben für die Regelung in die Parametrierung der Matrizen ist nicht einfach herzuleiten. Insbesondere die Wichtungsfaktoren abseits der Hauptdiagonalen sind in ihrem Einfluss auf das Regelverhalten schwer einzuschätzen. Daher wird man in vielen Fällen eine iterative Vorgehensweise mit Überprüfung durch Simulationen nicht vermeiden können.

7 Zustandskurven

Zustandsgleichungen und die dazugehörigen *Zustandskurven* (auch *Trajektorien* genannt) im n -dimensionalen Raum sind von großer Bedeutung für regelungstheoretische Untersuchungen.

Übersichtliche, verallgemeinerte und nicht nur numerisch verwertbare Ergebnisse lassen sich bis $n = 2$ gewinnen. Man spricht dann von *ebenen Zustandskurven* oder der *zweidimensionalen Phasenebene*. Diese bereits von Poincaré (Mathematiker, 1854-1912) eingeführte Beschreibungsform erlaubt eine anschauliche graphische Darstellung des dynamischen Verhaltens *linearer* und *nichtlinearer* Systeme *zweiter Ordnung*.

Sie dient neben der Beschreibung des Eigenverhaltens auch zur Ermittlung des Übergangsverhaltens und ist stets anwendbar, wenn die das System beschreibende Differentialgleichung zweiter Ordnung in zwei Differentialgleichungen erster Ordnung überführt werden kann. Dies ist möglich, wenn die Zeit t nicht explizit auftritt.

7.1 Trajektorien im zweidimensionalen Zustandsraum

Im Folgenden betrachtet man ein System, das durch eine Differentialgleichung zweiter Ordnung beschrieben wird, wobei vereinfachend angenommen wird, dass die Zustandsintegratoren in einer Kette angeordnet sind ohne zwischengeschaltete Additionsstellen. Eine derartige Struktur ist in Bild 7.1 gezeichnet.

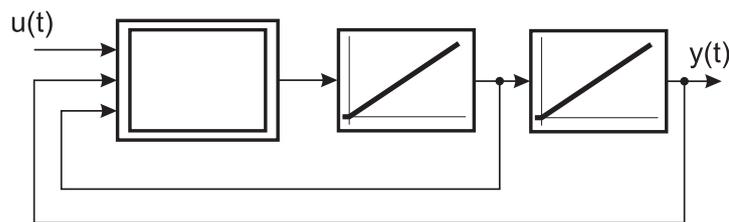


Bild 7.1: vereinfachte Struktur eines nichtlinearen Systems

Diese Struktur kann wie folgt beschrieben werden:

$$\ddot{y} - f(y, \dot{y}, u) = 0, \quad (7.1)$$

wobei $f(y, \dot{y}, u)$ eine lineare oder nichtlineare Funktion sei. Die Differentialgleichung soll nicht explizit von der Zeit t abhängen.

Führt man die Zustandsvariablen $x_1 = y$ und $x_2 = \dot{y}$ ein, so kann man folgende Zustandsgleichung aufstellen:

$$\dot{x}_1 = x_2 \quad (7.2)$$

$$\dot{x}_2 = f(x_1, x_2, u) \quad (7.3)$$

Diese Differenzialgleichung kann man als ein Vektorfeld visualisieren, wie in Bild 7.2 gezeigt. Über der Ebene der Zustandsgrößen (x_1, x_2) ist an jedem Punkt ein Vektor aufgezichnet, der Betrag und Richtung der Ableitung (\dot{x}_1, \dot{x}_2) anzeigt. Eine Zustandskurve (wie als Linie in derselben Grafik eingezeichnet) ergibt sich dann durch Nachzeichnen der Pfeile.

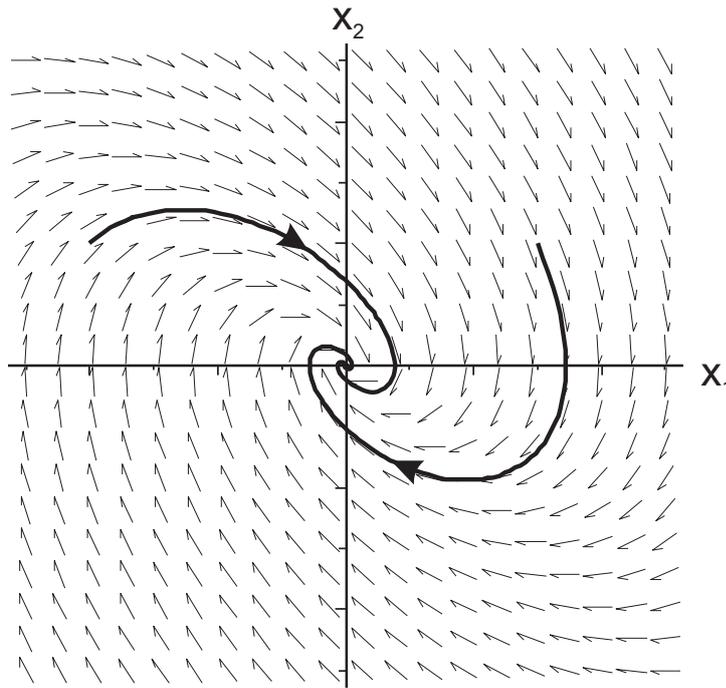


Bild 7.2: Vektorfeld einer Differentialgleichung zweiter Ordnung

Wie bekannt beschreiben die Zustandsgrößen x_1 und x_2 den Zustand des Systems in jedem Zeitpunkt vollständig. Trägt man nun in einem kartesischen Koordinatensystem x_2 über x_1 auf, so stellt jede Lösung $y(t)$ der Systemgleichung (7.1) eine Kurve in dieser Zustandsebene dar, die der Zustandspunkt (x_1, x_2) mit einer bestimmten Geschwindigkeit durchläuft. Dies ist durch die Zeitkodierung in Bild 7.3 dargestellt.

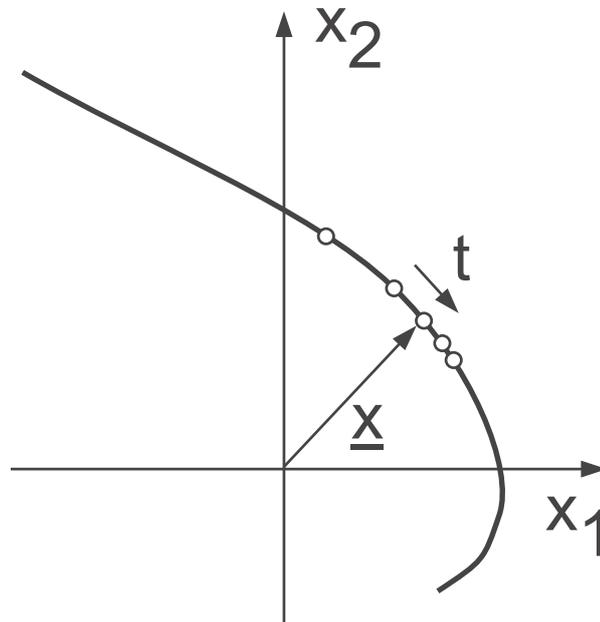


Bild 7.3: Trajektorie mit Zeitkodierung in der Zustandsebene

Kennzeichnend für die Wahl der Zustandsvariablen ist hierbei, dass die zweite Zustandsvariable die zeitliche Ableitung der ersten ist. Zu jedem $u(t)$ gehört eine eindeutige Trajektorie in der Zustandsebene. Zeichnet man für verschiedene Anfangsbedingungen (x_{10}, x_{20}) die Trajektorien, so erhält man eine Kurvenschar, das sogenannte *Phasenporträt*, vgl. Bild 7.4.

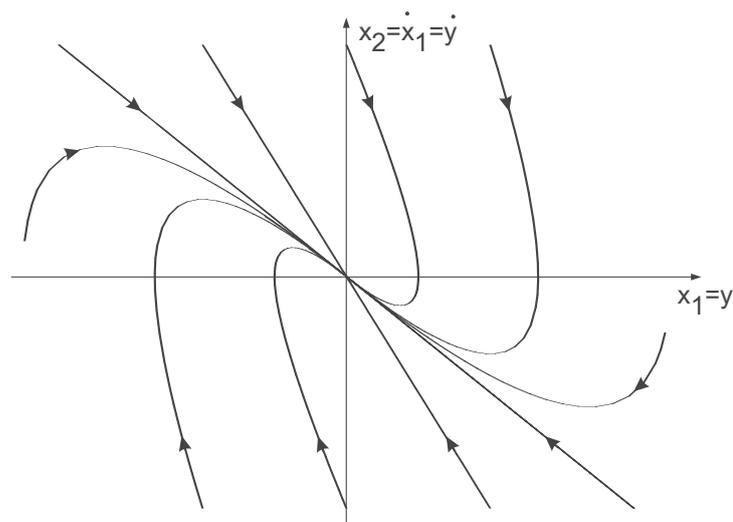


Bild 7.4: Phasenporträt

Nunmehr werde angenommen, dass $u(t)$ eine stückweise konstante Funktion sei. Dies ist z. B. immer der Fall, wenn u die Ausgangsgröße eines *Kennliniengliedes mit Relais-Charakter* (Zweipunkt- oder Dreipunktglied) ist. Damit hängen die ZustandsDifferenzialgleichungen nicht explizit von t ab und es ist möglich, die beiden Zustandsgleichungen zusammenzufassen:

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{f(x_1, x_2, u)}{x_2}, \quad u = \text{const.} \quad (7.4)$$

Das ist eine Differenzialgleichung erster Ordnung, die eine eindeutige Lösung bei gegebenem Anfangspunkt (x_{10}, x_{20}) besitzt. Die Untersuchung der ursprünglichen Differenzialgleichung zweiter Ordnung reduziert und vereinfacht sich damit auf eine Behandlung einer Differenzialgleichung erster Ordnung.

Diese Vereinfachung hat jedoch zur Folge, dass der zeitliche Ablauf des Vorgangs aus der dargestellten Trajektorie nicht mehr unmittelbar zu entnehmen ist, sie lässt sich aber aus

$$dt = \frac{dx_1}{x_2} \quad (7.5)$$

durch Integration bestimmen.

$$t = t_0 + \int_{x_{10}}^{x_{20}} \frac{dx_1}{x_2} \quad (7.6)$$

Hierbei ist x_{10} die Anfangsbedingung im Zeitpunkt t_0 und für x_2 muss die Gleichung der Trajektorie $x_2 = x_2(x_1)$ eingesetzt werden.

Abschließend werden einige allgemeine Eigenschaften von Zustandskurven dargestellt. (Gilt für die eingangs des Kapitels dargestellte Struktur des nichtlinearen Systems, bei der die Nichtlinearität am Eingang einer Integratorkette ohne Zwischenadditionsstellen angebracht ist.):

1. Jede Trajektorie verläuft in der *oberen* Zustandshalbebene ($x_2 > 0$) *von links nach rechts*, da wegen $x_2 = \dot{x}_1$ und $\dot{x}_1 > 0$ der Wert von x_1 zunimmt.
2. Jede Trajektorie verläuft in der *unteren* Halbebene der Phasenebene ($x_2 < 0$) *von rechts nach links*, da wegen $x_2 = \dot{x}_1$ und $\dot{x}_1 < 0$ der Wert von x_1 abnimmt.
3. Trajektorien schneiden die x_1 -Achse senkrecht. Dies ergibt sich unmittelbar aus den beiden obigen Bedingungen aufgrund der Stetigkeit von Trajektorien. Damit folgt auch, dass diese Schnittpunkte gewöhnlich Extremwerte von x_1 darstellen, und dass in der oberen oder der unteren Phasenhälfte keine Bahnpunkte mit vertikaler Tangente existieren.

Ausnahme: Erfolgt der Schnitt der Trajektorie mit der x_1 -Achse nicht senkrecht, dann liegt ein *singulärer Punkt* vor.

4. Die *Gleichgewichtslagen* eines dynamischen Systems werden stets durch *singuläre Punkte* gebildet. Wird wie eingangs dieses Kapitels angenommen, dass die Nichtlinearität am Eingang einer Integratorkette konzentriert ist, so liegen die Gleichgewichtspunkte auf der x_1 -Achse, im allgemeinen Fall, wenn zwischen den Integratoren Additionsstellen liegen, können Gleichgewichtslagen in der gesamten Zustande ebene auftreten. Dabei unterscheidet man verschiedene singuläre Punkte: Wirbelpunkte, Strudelpunkte, Knotenpunkte und Sattelpunkte (sie werden im Kapitel 7.2 näher erläutert).
5. Im Phasenporträt stellen die in sich geschlossenen Zustandskurven Dauerschwingungen dar.
6. Trajektorien für verschiedene Anfangsbedingungen mit gleicher Anregung schneiden sich nicht.

7.2 Zustandskurven linearer Systeme zweiter Ordnung

Bevor im Kapitel 8 die Einführung und Untersuchung nichtlinearer Systeme beginnt, soll zunächst die lineare Differenzialgleichung zweiter Ordnung mit konstanten Koeffizienten betrachtet werden. Einerseits soll damit gezeigt werden, wie sich die bekannten Verhältnisse in der Zustande ebene abbilden, und andererseits werden die Ergebnisse für die spätere Betrachtung der nichtlinearen Systeme benötigt.

7.2.1 Trajektorien eines zweifachen Integrators

Die Reihenschaltung zweier I-Glieder, die durch eine konstante Eingangsgröße u angesteuert wird, stellt die einfachste lineare Strecke zweiter Ordnung dar (vgl. Bild 7.5).

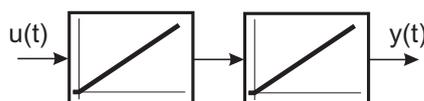


Bild 7.5: Zweifacher Integrator

Die Differenzialgleichung lautet:

$$\ddot{y}(t) = u(t) \quad (7.7)$$

Führt man die Zustandsvariablen $x_1 = y$ und $x_2 = \dot{y} = \dot{x}_1$ ein, so erhält man die folgenden Zustandsdifferenzialgleichungen:

$$\dot{x}_1 = x_2 \quad (7.8)$$

$$\dot{x}_2 = u \quad (7.9)$$

Durch die Elimination der Zeit t erhält man die zusammengefasste Zustandsgleichung

$$\frac{d x_2}{d x_1} = \frac{d x_2 / d t}{d x_1 / d t} = \frac{u}{x_2} \quad (7.10)$$

Durch die Trennung der Veränderlichen in Gleichung (7.10) erhält man:

$$\int x_2 dx_2 = \int u dx_1 \quad (7.11)$$

$$\frac{x_2^2}{2} = \int u dx_1 + C_1 \quad (7.12)$$

Mit $u = u_0 = \text{const.}$ erhält man die Parabelgleichung:

$$x_2^2 = 2 u_0 (x_1 - C), \quad (7.13)$$

deren Achse in die x_1 -Achse fällt und deren Scheitel bei $x_1 = C$ liegt, da dort $x_2 = 0$ ist. Die Integrationszeitkonstante C ist durch den Anfangswert (x_{10}, x_{20}) bestimmt. Dabei ist $C = -\frac{C_1}{u_0}$. Es ergeben sich die in Bild 7.6 gezeichneten Verläufe der Zustandstrajektorien.

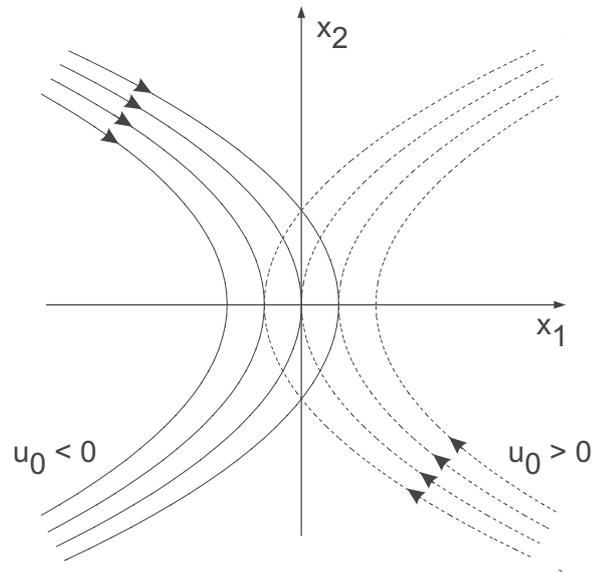


Bild 7.6: Zustandskurven eines doppelten Integrators

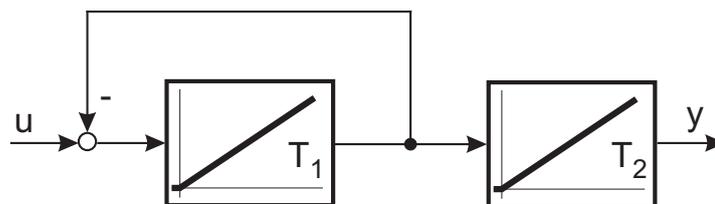
Für den Fall, dass $u_0 > 0$ ist, sind die Parabeln nach rechts, für den Fall, dass $u_0 < 0$ ist, sind die Parabeln nach links geöffnet. Der Zustandspunkt (x_1, x_2) durchläuft die Zustandskurven in Pfeilrichtung.

7.2.2 Zustandskurven eines verzögerten Integrators

Die Differenzialgleichung eines verzögerten Integrators IT_1 lautet bekanntlich:

$$\ddot{y}(t) + a_1 \dot{y}(t) = b_0 u(t) \quad (7.14)$$

Das Bild 7.7 zeigt das dazugehörige Blockschaltbild.

Bild 7.7: Blockschaltbild einer IT_1 -Strecke

Geht man wiederum von einem konstanten Steuersignal $u(t)$ aus, so kann man den konstanten Faktor b_0 multipliziert mit $u(t)$ als konstantes Steuersignal u_0 auffassen.

$$\ddot{y}(t) + a_1 \dot{y}(t) = u_0 \quad (7.15)$$

Man erhält mit der Wahl der Zustandsvariablen $x_1 = y$ und $x_2 = \dot{y}$ die Zustandsgleichungen

$$\dot{x}_1 = x_2 \quad (7.16)$$

$$\dot{x}_2 = u_0 - a_1 x_2 \quad (7.17)$$

und somit die auf die erste Ordnung reduzierte Zustandsgleichung

$$\frac{d x_2}{d x_1} = \frac{u_0 - a_1 x_2}{x_2} \quad (7.18)$$

Die Lösung der Zustandsgleichung erfolgt auch hier wieder durch die Trennung der Veränderlichen:

$$d x_1 = \frac{x_2}{u_0 - a_1 x_2} d x_2 \quad (7.19)$$

Mit der Substitution

$$z := u_0 - a_1 x_2, \quad x_2 = \frac{z - u_0}{-a_1}, \quad d x_2 = -\frac{d z}{a_1}$$

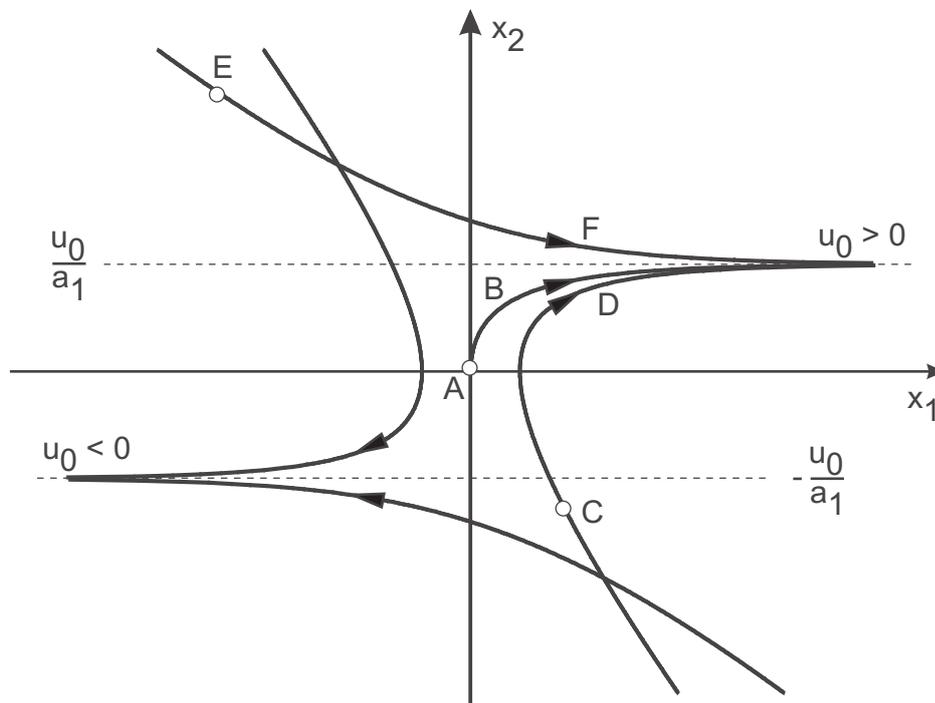
erhält man

$$d x_1 = \left(\frac{1}{a_1^2} - \frac{u_0}{a_1^2 z} \right) d z \quad (7.20)$$

und schließlich nach der Integration und Rücksubstitution

$$x_1 = -\frac{x_2}{a_1} - \frac{u_0}{a_1^2} \ln|u_0 - a_1 x_2| + C_1 \quad (7.21)$$

Das Bild 7.8 zeigt den prinzipiellen Verlauf der Zustandskurven. Für beide Vorzeichen der Stellgröße $u = \pm u_0$ existieren je zwei Äste der Zustandskurve, die entlang den horizontalen Asymptoten $x_2 = \pm u_0/a_1$ ineinander münden.

Bild 7.8: Zustandskurven des IT_1 -Gliedes

In Bild 7.9 sind den Zustandskurven aus Bild 7.8 einige typische zeitliche Verläufe gegenübergestellt. Die Kurve A–B entspricht dabei der Sprungantwort eines verzögerten Integrators, während bei C–D und E–F nichtverschwindende Anfangsbedingungen zugrunde gelegt wurden.

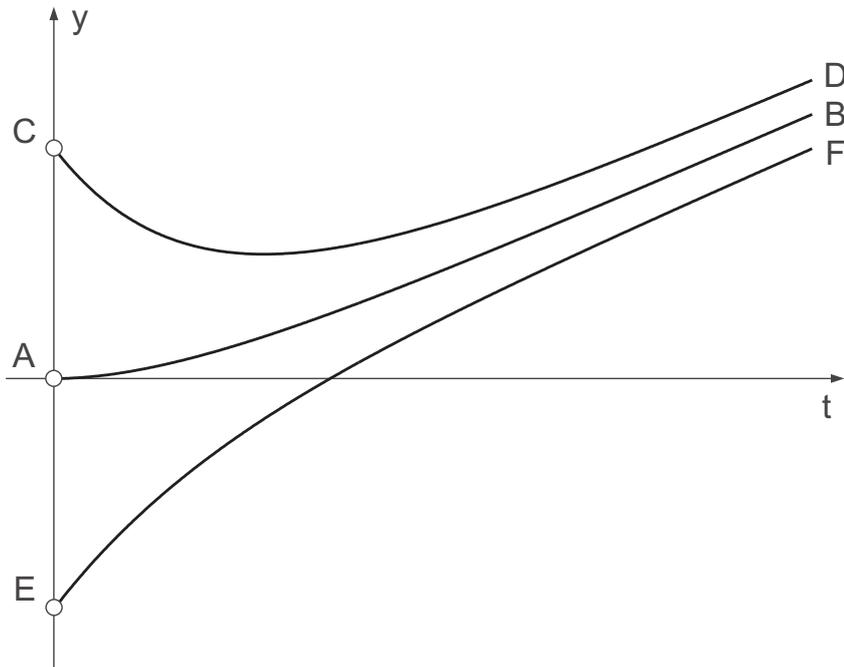


Bild 7.9: Zeitliche Verläufe

7.2.3 Zustandskurven eines PT_2 -Gliedes

Die Differentialgleichung eines Verzögerungsgliedes zweiter Ordnung lautet:

$$\ddot{y} + a_1 \dot{y} + a_0 y = 0 \quad (7.22)$$

Der Einfachheit halber sei hier nur ein System ohne äußere Anregung $u(t) = 0$ betrachtet. Die Auslenkung des Systems erfolge rein durch die Anfangswerte der Zustandsgrößen. Löst man diese Differentialgleichung im Zeitbereich, dann erhält man in Abhängigkeit von den Eigenwerten die beiden folgenden Ansätze:

$$y(t) = k_1 e^{s_1 t} + k_2 e^{s_2 t} \quad \text{für } s_1 \neq s_2 \quad (7.23)$$

$$y(t) = k_1 e^{s_1 t} + k_2 t e^{s_1 t} \quad \text{für } s_1 = s_2 \quad (7.24)$$

Hierbei stellen s_1 und s_2 die Lösungen der charakteristischen Gleichung, also die Pole des Systems dar.

$$s^2 + a_1 s + a_0 = (s - s_1)(s - s_2) = 0 \quad (7.25)$$

Mit der Lösungsformel für quadratische Gleichungen erhält man:

$$s_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_0}}{2} \quad (7.26)$$

$$s_2 = \frac{-a_1 - \sqrt{a_1^2 - 4a_0}}{2} \quad (7.27)$$

Lineare Systeme, die durch die Gleichung (7.22) beschrieben werden, besitzen nur einen *singulären Punkt*, nämlich den Ursprung. Trotzdem ergeben sich in Abhängigkeit von a_1 und a_0 , und damit der Eigenwerte des Systems, sehr unterschiedliche Formen der Trajektorien um den Ursprung. Es können folgende vier Fälle auftreten. Mit der bekannten Wahl der Zustandsgrößen $x_1 = y$ und $x_2 = \dot{y}$ erhält man die angegebenen Zustandstrajektorien.

- s_1 und s_2 sind beide reell und besitzen das gleiche Vorzeichen. Im Phasenporträt ergeben sich je nach dem, ob die Pole positive oder negative Vorzeichen haben, die folgenden Verläufe:

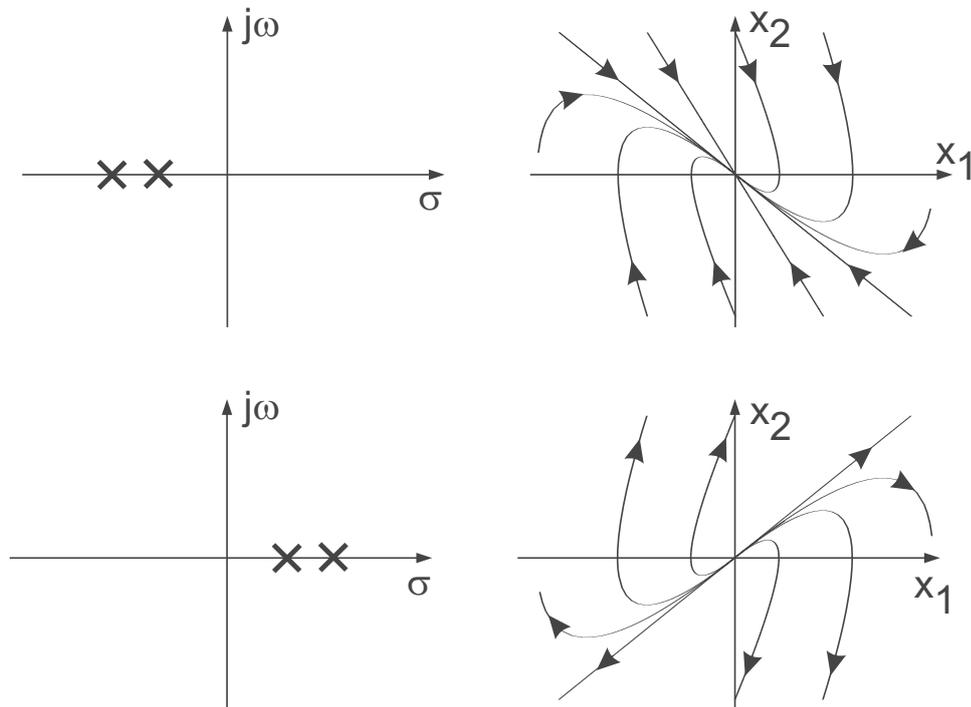


Bild 7.10: Zustandskurven eines PT_2 -Gliedes, reelle Pole, gleiches Vorzeichen

Wie man aus Bild 7.10 erkennt, streben die Trajektorien gegen die *Ruhelage* $(x_1, x_2) = (0, 0)$ für Pole mit negativem Vorzeichen (stabil) und von ihr weg für Pole mit po-

sitivem Vorzeichen (instabil). Man bezeichnet die Ruhelage $(x_1, x_2) = (0, 0)$ als einen *Knotenpunkt*.

- s_1 und s_2 sind beide reell und besitzen unterschiedliche Vorzeichen. In diesem Fall erhält man den folgenden Verlauf der Zustandskurven:

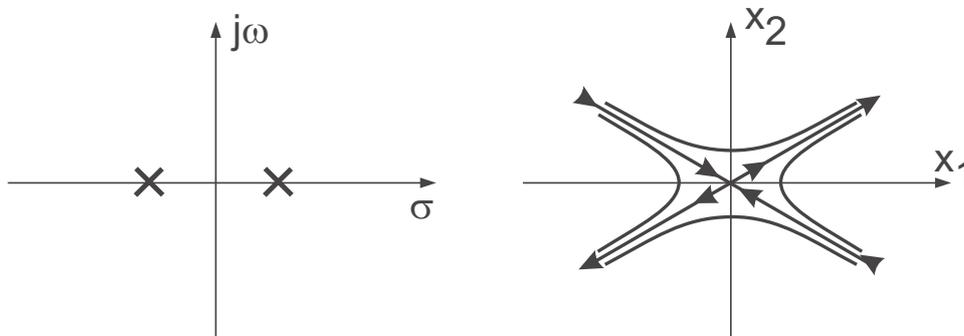
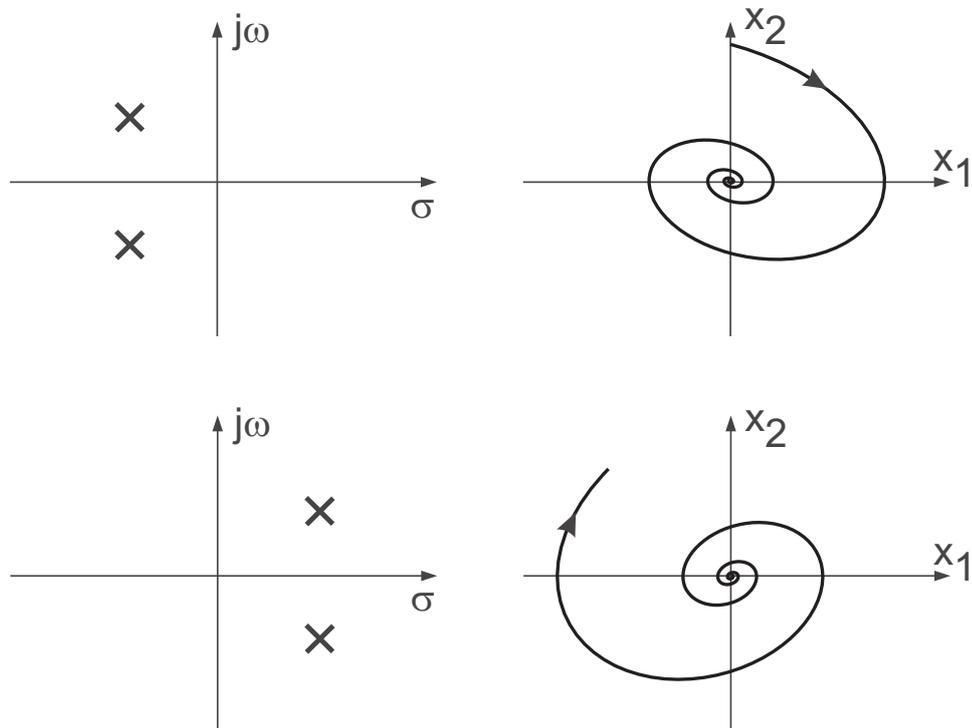


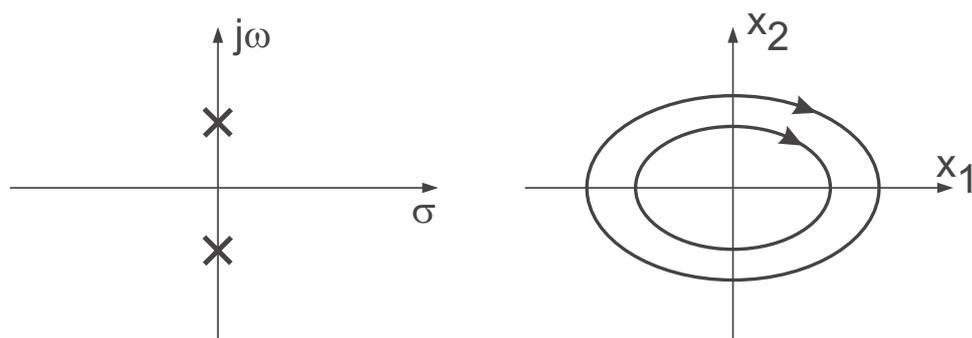
Bild 7.11: Zustandskurven eines PT_2 -Gliedes, reelle Pole, unterschiedliches Vorzeichen

Man bezeichnet die Ruhelage $\underline{0}$ hier als *Sattelpunkt*. Sie ist instabil, da eine geringste Auslenkung aus der Ruhelage zu einer Zustandskurve führt, die sich vom Sattelpunkt entfernt.

- s_1 und s_2 sind konjugiert komplex, der Realteil ist ungleich Null. Besitzt das System konjugiert komplexe Pole mit einem nichtverschwindenden Realteil, so erhält man Spiralen um den Ursprung. Bei negativem Realteil ziehen sich die Spiralen auf den Ursprung zusammen, bei positivem Realteil bewegen sie sich vom Ursprung weg. Man nennt die Ruhelage *Strudelpunkt* und erhält in Abhängigkeit der Lage der Pole folgende Zustandskurven:

Bild 7.12: Zustandskurven eines PT₂-Gliedes, konjugiert komplexe Pole, $\sigma \neq 0$

- s_1 und s_2 sind konjugiert komplex, der Realteil ist Null, d. h. sie stellen Pole auf der imaginären Achse dar. Wie das Bild 7.13 zeigt, sind die Zustandskurven in diesem Fall Ellipsen bzw. Kreise (je nach Skalierung der Achsen) um den Ursprung. Man spricht hierbei von einem *Wirbelpunkt*. Das System ist *grenzstabil*.

Bild 7.13: Zustandskurven eines PT₂-Gliedes, konjugiert komplexe Pole, $\sigma = 0$

8 Nichtlineare Systeme

Bisher wurden überwiegend lineare Systeme betrachtet. *Nichtlineare Elemente* in Regelkreisen wurden um den Arbeitspunkt *linearisiert*, so dass anschließend das System wieder mit den Methoden der „linearen Regelungstechnik“ untersucht werden konnte. Bei einer Linearisierung von Nichtlinearitäten lassen sich die Eigenschaften des Systems jedoch nur näherungsweise erfassen. Bei vielen nichtlinearen Systemen, z. B. mit schaltenden Gliedern, ist eine lineare Approximation nicht möglich, da es sich um sogenannte „harte Nichtlinearitäten“ handelt, für die die Methoden der „linearen Regelungstechnik“ nicht oder nur sehr eingeschränkt anwendbar sind.

Lineare Übertragungssysteme oder auch nur einzelne lineare Übertragungselemente sind dadurch gekennzeichnet, dass das Verstärkungs- und Überlagerungsprinzip gilt. Gilt dieses Superpositionsprinzip nicht, spricht man von *nichtlinearen Übertragungsgliedern*.

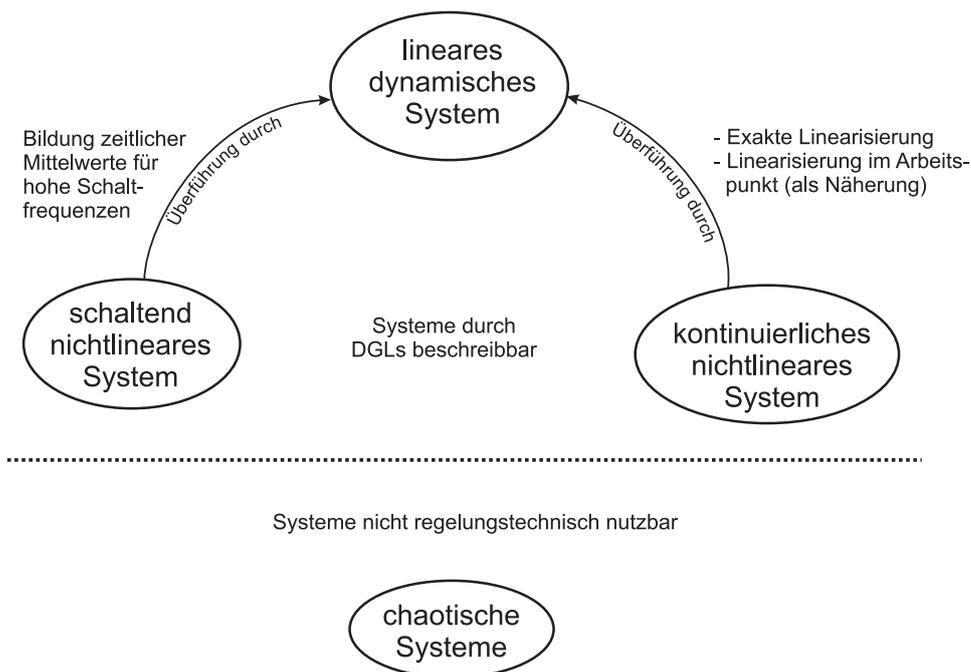


Bild 8.1: Zusammenhänge unterschiedlicher Systemarten

Die Gründe für nichtlineares Verhalten können vielfältige Ursachen besitzen, denkbar wären:

- Physikalische Phänomene sind im Arbeitsbereich nichtlinear, z. B. Diodenkennli-

nie, Magnetisierung ferromagnetischer Stoffe oder nichtlineare Reibung bei mechanischen Systemen.

- Die technische Realisierung verursacht ein nichtlineares Verhalten, z. B. ist bei einem mehrgelenkigen Roboterarm der Zusammenhang zwischen Drehwinkeln und Positionierung nichtlinear.
- An lineare Bereiche schließen sich Sättigungen oder Begrenzungen an. Diese können beabsichtigt, z. B. Schutzbegrenzungen für Ströme oder Spannungen, oder unbeabsichtigt aufgrund physikalischer Rahmenbedingungen sein, z. B. Maximalwert einer Spannungsquelle, Ventilhub.

Für nichtlineare Systeme existiert keine einheitliche Theorie, die universell anwendbar ist. Allerdings haben sich für Stabilitätsuntersuchungen nichtlinearer Systeme verschiedene Verfahren etabliert, die in späteren Abschnitten vorgestellt werden. Auf Basis eines guten Systemmodells lassen sich auf dem Rechner aber auch durchaus mit heuristischen Ansätzen gute Ergebnisse erzielen. Eine besondere Klasse stellen auch die Ereignis-diskreten Systeme dar, die im weitesten Sinne unter den schaltenden nichtlinearen Systemen eingeordnet werden können. Hier ist der Zustandsraum diskretisiert und Zustandsübergänge finden augenblicklich statt.

8.1 Nichtlineare Systeme in der Zustandsebene

In Kapitel 7 wurde gezeigt, dass sich für Regelstrecken mit zwei Energiespeichern und konstanter Anregung das dynamische Verhalten anschaulich in der Zustandsebene darstellen lässt. Diese Erkenntnisse sollen nun dazu verwendet werden, die Wirkungsweise einfacher Regelkreise zu untersuchen, bei denen ein Zwei- oder Dreipunktregler über das nachfolgende Schaltglied die Regelstrecke mit einer abschnittsweise konstanten Stellgröße ansteuert.

Zunächst sollen allerdings noch einige Ergänzungen zur Darstellung von Zustandskurven für nichtlineare Systeme eingeführt werden.

8.1.1 Grenzyklen

Existieren für die Zustandsdifferenzialgleichung

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t), u(t)) \quad (8.1)$$

mit dem Anfangszustand $\underline{x}(t_0) = \underline{x}_0$ und $u(t) = \text{const}$ Lösungen mit der Eigenschaft

$$\underline{x}(t, \underline{x}_0) = \underline{x}(t + T, \underline{x}_0), \quad \text{für alle } t \geq t_0, \quad (8.2)$$

so nennt man diese *periodisch* mit der Periodendauer T . Sie stellen im Zustandsraum *geschlossene Trajektorien* dar und werden *Dauerschwingungen* genannt. Allerdings ist nicht jede geschlossene Trajektorie eine Dauerschwingung, da z. B. auf ihr Ruhelagen liegen können. Dann bewegen sich die Zustände nur von einer Ruhelage auf eine andere zu, ohne den Umlauf zu vollenden.

Liegt in einer hinreichend kleinen Umgebung einer Dauerschwingung keine weitere periodische Lösung, so nennt man die Dauerschwingung eine *Grenzschiwingung* bzw. bei Systemen 2. Ordnung einen *Grenzyklus*.

Ruhelagen, auch *stationäre Zustände*, *Beharrungszustände* oder *Gleichgewichtszustände* genannt, sind dadurch charakterisiert, dass sich die Zustandsvariablen nicht ändern:

$$\dot{x} \equiv 0$$

Außerdem sind die Grenzyklen dadurch gekennzeichnet, dass benachbarte Trajektorien in den Grenzyklus einlaufen oder sich von ihm entfernen. Entsprechend Bild 8.2 unterscheidet man *stabile* (beidseitig stabil) *semistabile* (einseitig stabil) und *instabile Grenzyklen*.

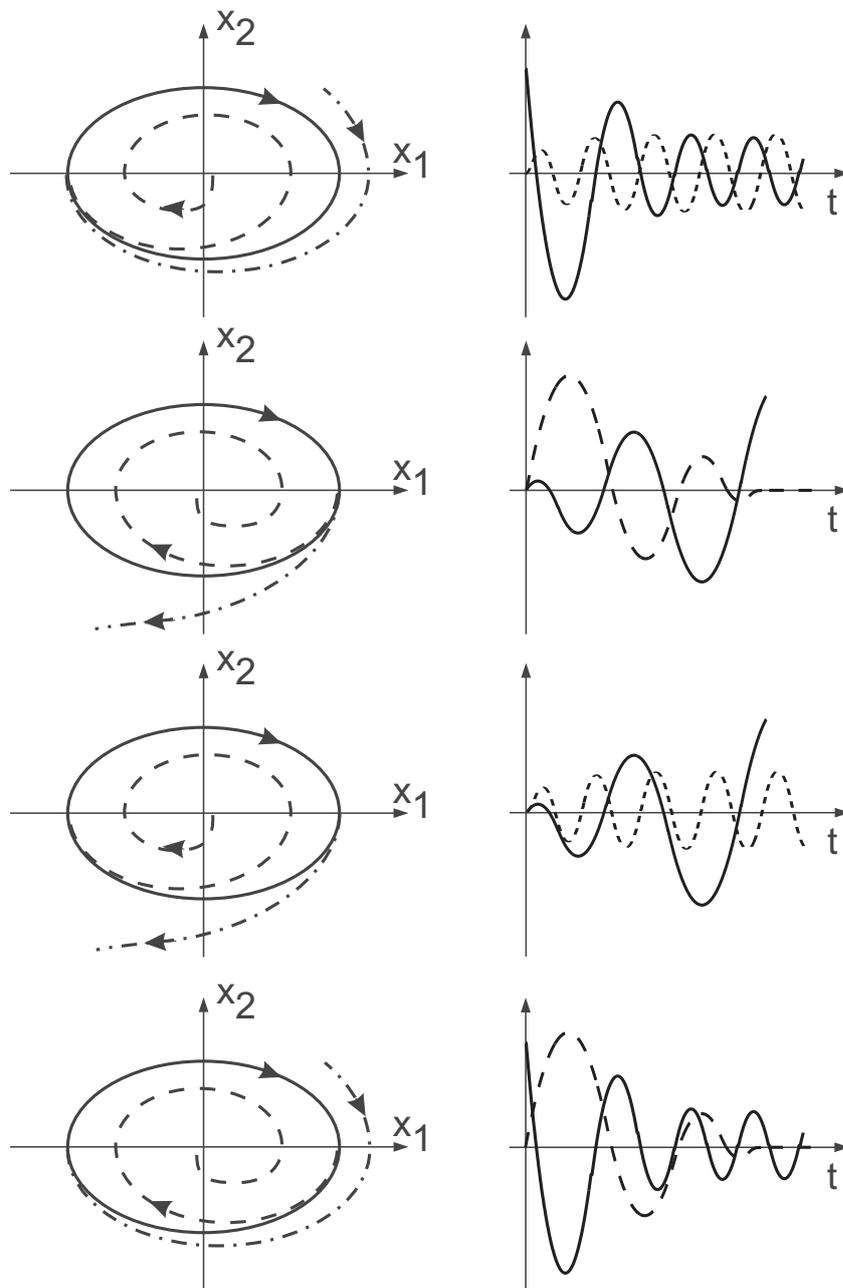


Bild 8.2: Grenzyklen

8.1.2 Existenz von Grenzyklen

Zum Nachweis der Existenz von Grenzyklen gibt es mehrere Ansätze, die hier nur kurz vorgestellt werden sollen.

Poincaré stellt hierfür beispielsweise eine einfache Beziehung auf. Hierbei bezeichnet N

die Anzahl der Summe der Strudel-, Knoten- und Wirbelpunkte, S die Anzahl der Sattelpunkte innerhalb eines Grenzyklus.

Satz 8.1 (Poincaré) *Existiert ein Grenzyklus in einem System zweiter Ordnung, dann muss $N = S + 1$ sein.*

Betrachtet man das Asymptotenverhalten der Trajektorien, lässt sich folgender Satz postulieren:

Satz 8.2 (Poincaré-Bendixson) *Bleibt eine Trajektorie eines Systems zweiter Ordnung innerhalb eines endlichen Gebiets Ω , dann gilt einer der folgenden Zusammenhänge:*

- *Die Trajektorie bewegt sich zu einem singulären Punkt.*
- *Die Trajektorie läuft in einen Grenzyklus ein.*
- *Die Trajektorie selbst ist ein Grenzyklus.*

Bendixson gibt weiterhin eine hinreichende Bedingung für die Nichtexistenz einer Grenzschwingung an.

Satz 8.3 (Bendixson) *Für ein nichtlineares System mit*

$$\dot{x}_1 = f_1(x_1, x_2) \quad \text{und} \quad \dot{x}_2 = f_2(x_1, x_2)$$

existieren keine Grenzyklen innerhalb eines Gebiets Ω der Zustandsebene, wenn

$$\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2}$$

nicht verschwindet und nicht das Vorzeichen ändert.

Beispiel: Betrachtet man das nichtlineare System

$$\dot{x}_1 = g(x_2) + 4x_1x_2^2$$

$$\dot{x}_2 = h(x_1) + 4x_1^2x_2,$$

dann hat dieses System keine Grenzyklen, da

$$\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = 4(x_1^2 + x_2^2)$$

mit Ausnahme im Ursprung immer positiv ist und das Vorzeichen nicht ändert.

8.2 Nichtlineare Übertragungselemente

Im Folgenden werden einige idealisierte nichtlineare Übertragungselemente, die in realen Systemen auftreten können, kurz vorgestellt.

8.2.1 Zweipunktglied

Das in Bild 8.3 dargestellte *ideale Zweipunktglied* wird mathematisch durch

$$x_a = b \operatorname{sign} x_e = \begin{cases} -b & \text{für } x_e < 0 \\ b & \text{für } x_e \geq 0 \end{cases} \quad (8.3)$$

beschrieben.

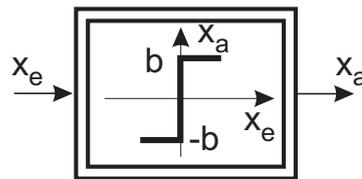


Bild 8.3: Zweipunktglied ohne Hysterese

Das folgende Bild 8.4 zeigt ein *ideales Zweipunktglied mit Nullpunkt*. Es wird mathematisch beschrieben durch:

$$x_a = b \operatorname{sign} x_e = \begin{cases} -b & \text{für } x_e < 0 \\ 0 & \text{für } x_e = 0 \\ b & \text{für } x_e > 0 \end{cases} \quad (8.4)$$

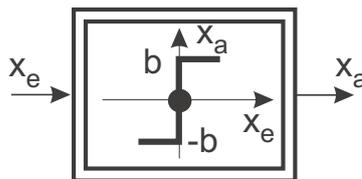


Bild 8.4: Ideales Zweipunktglied mit Nullpunkt

Das folgende Bild 8.5 zeigt ein *Zweipunktglied mit Hysterese*.

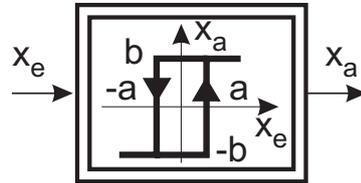


Bild 8.5: Zweipunktglied mit Hysterese

Das Verhalten lässt sich durch

$$x_a = \begin{cases} -b & \text{für } x_e < -a \\ b \operatorname{sign}(x_e - a \operatorname{sign} \dot{x}_e) & \text{für } -a < x_e < a \\ b & \text{für } x_e > a \end{cases} \quad (8.5)$$

mathematisch beschreiben. Diese Beschreibung ist gut geeignet für sinusförmige Anregungen, wie sie beispielsweise bei Untersuchungen mittels *Beschreibungsfunktion* (siehe Kap. 10) vorkommen. Eine exaktere Beschreibung geht von einer internen binären Zustandsgröße aus, die bei Überschreiten der Hysteresegrenzen gesetzt wird und im Bereich dazwischen den Ausgangszustand speichert.

8.2.2 Dreipunktglied

Auch *Dreipunktglieder* gibt es *mit* und *ohne Hysterese*. Die mathematische Abhängigkeit des Ausgangssignals x_a vom Eingangssignal x_e für ein Dreipunktglied ohne Hysterese lautet:

$$x_a = \begin{cases} -b & \text{für } x_e < -a \\ 0 & \text{für } -a < x_e < a \\ b & \text{für } x_e > a \end{cases} \quad (8.6)$$

Man erhält folgendes Blockschaltbildsymbol:

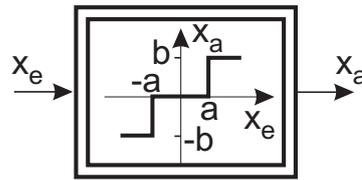


Bild 8.6: Dreipunktglied ohne Hysterese

Die mathematische Beschreibung für ein *Dreipunktglied mit Hysterese* ist relativ aufwendig und wird daher hier nicht angegeben. Das Ein-/Ausgangsverhalten des hysteresebehafteten Dreipunktglieds stellt sich wie folgt dar:

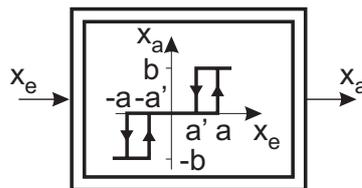


Bild 8.7: Dreipunktglied mit Hysterese

Wie auch schon beim Zweipunktglied mit Hysterese müsste für eine exakte Beschreibung auch hier von einem inneren Zustand ausgegangen werden, der den Wert der Ausgangsgröße speichert, wenn das Eingangssignal sich innerhalb eines Hysteresebereiches befindet und hart gesetzt wird außerhalb dieses Bereiches.

8.2.3 Begrenzung

Begrenzungs- oder *Sättigungskennlinien* treten in realen Anlagen häufig auf, da man praktisch bei jedem Entwurf eines Regelkreises die Begrenzung der Stellgröße berücksichtigen muss. Ursachen hierfür können z. B. mechanische Anschläge oder die Eisensättigung bei elektrischen Maschinen sein. Das Element wird häufig in Regelungen bewußt eingesetzt, um Sollwerte zu begrenzen.

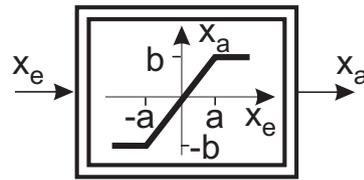


Bild 8.8: Begrenzung

Aus dem Blockschaltbildsymbol 8.8 lässt sich direkt die mathematische Beschreibung dieses Verhaltens ableiten:

$$x_a = \begin{cases} -b & \text{für } x_e \leq -a \\ \frac{b}{a} x_e & \text{für } -a \leq x_e \leq a \\ b & \text{für } x_e \geq a \end{cases} \quad (8.7)$$

Auch Begrenzungen können mit Hysterese existieren, als Beispiel sei hier nur der Reversierbetrieb eines Getriebes mit Zahnspiel genannt (*Lose*).

8.2.4 Totzone oder Lose

Die *Totzone* hat die Gleichung

$$x_a = \begin{cases} (x_e + a) \tan \alpha & \text{für } x_e < -a \\ 0 & \text{für } -a \leq x_e \leq a \\ (x_e - a) \tan \alpha & \text{für } x_e > a \end{cases} \quad (8.8)$$

und stellt sich als Blockschaltbildsymbol nach Bild 8.9 dar.

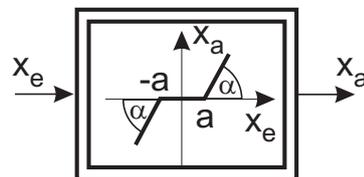


Bild 8.9: Totzone

Sie tritt beispielsweise bei elektrischen Systemen als Schwellwert auf, kann sich bei mechanischen Systemen aus dem Übergang Haft-/Gleitreibung bemerkbar machen oder stellt

bei Messeinrichtungen die Ansprechzone dar. Auch dieses Element stellt häufig eine grobe Vereinfachung dar. Für die Lose beispielsweise wird durch die übertragenen Kräfte oder Momente festgelegt, in welcher Richtung der Offset zur Ausgangsgröße addiert wird.

8.2.5 Kennlinienglieder

Neben den in den vorherigen Abschnitten vorgestellten *Kennliniengliedern* mit *stückweise linearen Kennlinien* gibt es noch beliebige „krumme“ Kennlinien. Sie werden im Blockschaltbild als Nachbildung der realen Kennlinie angegeben. Als Beispiel sei hier nur die Federkennlinie mit einem progressiven oder degressiven Verlauf genannt.

Der allgemeine Zusammenhang lautet:

$$x_a = f(x_e), \tag{8.9}$$

wobei f eine beliebige mathematische Funktion beschreibt. Häufig können diese allgemeinen Kennlinienglieder auch noch hysteresebehaftet sein.

8.3 Schaltende Regler

Schaltende Regler sind aus dem industriellen Einsatz nicht wegzudenken, weil schaltende Stellglieder im Vergleich zu kontinuierlichen Stelleinheiten zwei wesentliche Vorteile bieten: Sie sind deutlich preiswerter und effizienter. Ein kontinuierliches Stellelement erzeugt immer prinzipbedingt Verluste, deren Wärmeabfuhr nicht unerhebliche Kosten verursacht. Das Schalten als diskontinuierlicher Vorgang hat aber auch eine Reihe von Nachteilen. Zunächst entsteht außer an den Stellbereichsgrenzen im normalen Betrieb immer eine *Schalterschwingung* entweder durch eine bewusst eingeführte Puls-Weiten-Modulation oder der Regelkreis gerät von sich aus in einen periodischen Grenzyklus mit einer bestimmten Frequenz und Amplitude auf der Regelgröße. Daher ist zunächst abzuschätzen, wie groß die Amplitude dieser Grenzschiwingung ausfällt.

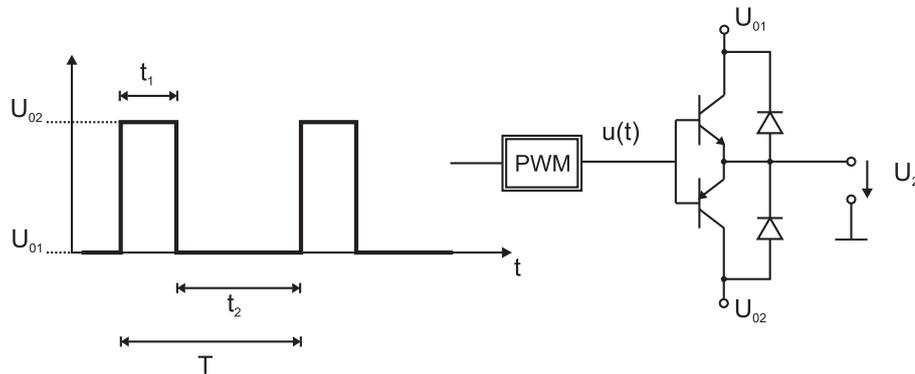


Bild 8.10: PWM-Signal an einem bidirektionalem Emitterfolger

8.3.1 Linearisierung durch periodisches Schalten

Puls Weiten Modulation wird gerne eingesetzt, um kompakte, preiswerte Stellglieder aufbauen zu können. Allerdings ist PWM eine nichtlineare Operation, weil sich die zeitliche Gestalt der Impulse aussteuerungsabhängig ändert. Dies ist ein fundamentaler Unterschied zur zeitdiskreten Signalverarbeitung, die als Z-Transformation eine lineare Modellierung erlaubt. Die in der zeitdiskreten Signalverarbeitung auftretenden Impulsreihen oder Stufenfunktionen haben in jedem Abtastschritt dieselbe zeitliche Gestalt; es ändert sich nur die Amplitude des Impulses. Bei der PWM ändert sich jedoch mit der Aussteuerung der zeitliche Verlauf innerhalb einer Periode, was zu den nichtlinearen Effekten führt. Solange die Schaltfrequenz deutlich höher ist als die dominierende Streckenzeitkonstante ist eine kleine Schaltschwingung auf der Regelgröße zu erwarten und die nichtlinearen Effekte bleiben klein, so dass man sie in vielen Fällen vernachlässigen kann. Man kann dann die PWM vernachlässigen und die Regler mit den Verfahren der linearen Regelungstechnik auslegen, indem man einfach mit den zeitlichen Mittelwerten der pulswertenmodulierten Signale rechnet.

$$\bar{u}_2 = u_2 = \frac{1}{T} \int_t^{t+T} u_2(\tau) d\tau = \frac{t_1}{T} U_{01} + \frac{t_2}{T} U_{02} = U_{02} + \frac{t_1}{T} (U_{01} - U_{02}) \quad (8.10)$$

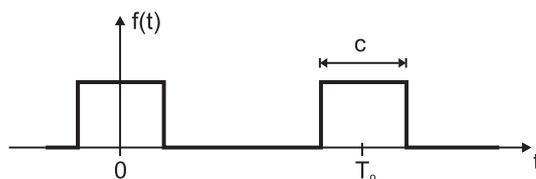


Bild 8.11: PWM-Signal

Auf der Regelgröße verbleibt aber ein Rest der Schaltschwingung, dessen Gestalt und Amplitude vom Verhältnis der Schaltfrequenz zu den Polen und Nullstellen der Regelstrecke und vom Aussteuergrad der PWM abhängt.

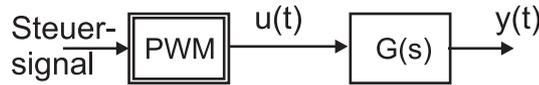


Bild 8.12: PWM an Strecke

Für das Ausgangssignal des Pulsweitenmodulator wie in Bild 8.11, das als Eingangssignal der Strecke dient, lässt sich eine Fourierreihe aufstellen:

$$f(t) = \frac{c}{T_0} + \frac{2}{\pi} \sum_{\nu=1}^{\infty} \frac{1}{\nu} \cdot \sin\left(\frac{\nu\pi c}{T_0}\right) \cdot \cos\left(\nu \frac{2\pi t}{T_0}\right) \quad \text{mit} \quad \omega = \frac{2\pi}{T_0} \quad (8.11)$$

Die Stellgröße ergibt sich zu:

$$u(t) = U_0 \cdot f(t) = U_0 \left[\frac{c_o}{T_0} + \frac{2}{\pi} \sum_{\nu=1}^{\infty} \frac{1}{\nu} \cdot \sin\left(\nu\omega_0 \frac{c}{2}\right) \cdot \cos(\nu\omega_0 t) \right] \quad (8.12)$$

einsetzen der si-Funktion $si(x) = \frac{\sin(x)}{x}$ führt auf

$$u(t) = U_0 \left[\frac{c_o}{T_0} + \frac{2}{\pi} \sum_{\nu=1}^{\infty} \frac{\pi c}{T_0} si\left(\nu\omega_0 \frac{c}{2}\right) \cdot \cos(\nu\omega_0 t) \right] \quad (8.13)$$

$$= U_0 \frac{c_o}{T_0} \sum_{\nu=-\infty}^{\infty} si\left(\nu\omega_0 \frac{c}{2}\right) \cdot e^{j(\nu\omega_0 t)} \quad (8.14)$$

Im stationären Zustand kann jede Spektrallinie des Anregungssignals mit dem Frequenzgang $G(j\omega)$ des linearen Systems bewertet werden und zum gesamten Signal der Schaltschwingungen im Ausgangssignal überlagert werden.

$$y(t) = U_0 \frac{c_o}{T_0} \sum_{\nu=-\infty}^{\infty} si\left(\nu\omega_0 \frac{c}{2}\right) \cdot e^{j(\nu\omega_0 t)} \cdot G(j\nu\omega_0) \quad (8.15)$$

8.3.2 Typen von Regelkreisen mit schaltenden Reglern

Für die Anordnung von schaltenden Reglern lassen sich vier grundsätzliche Topologien unterscheiden.

1) Der schaltende Regler wird ohne weitere dynamische Beschaltung mit der Regelstrecke zu einem Regelkreis verschaltet. In diesem Bild wurde ein Dreipunktregler mit Hysterese als Regler angedeutet, es ist im Prinzip auch jeder andere schaltende Reglertyp denkbar.

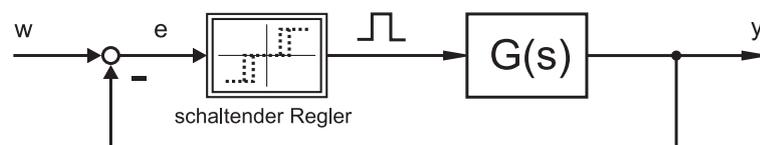


Bild 8.13: Schaltender Regler mit einfacher Rückführung

Diese Anordnung bietet keinerlei Möglichkeiten, mit einer Reglerdynamik auf die Dynamik der Strecke einzugehen. Die Stabilität des geschlossenen Kreises ist praktisch nur für Strecken 1. Ordnung zu erreichen. Die Amplitude und Frequenz der Schaltschwingung auf der Regelgröße ist nur durch die Streckeneigenschaften und die Wahl der Stellamplitude zu beeinflussen. Daher ist die Anwendbarkeit auf Strecken erster Ordnung beschränkt.

2) Anstelle eines schaltenden Reglers wird ein Modulator für Pulsweitenmodulation einem linearen kontinuierlichen Regler verwendet. Als Modulator wird entweder in einem Mikrorechner eine entsprechende Peripherieeinheit verwendet oder ein Modulator mit Operationsverstärkern und Komparatoren aufgebaut, der die Stellgröße des analogen Reglers mit einem dreieckförmigen Trägersignal vergleicht und über den Komparator das PWM-Signal erzeugt.

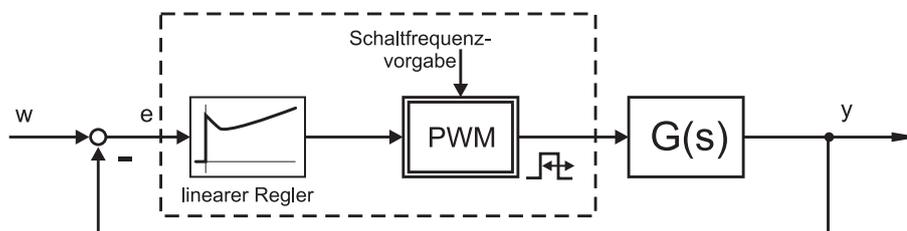


Bild 8.14: linearer Regler mit nachgeschalteter PWM

Als linearer Regler in dieser Schaltung können alle üblichen Reglertypen oder eigene Entwürfe verwendet werden. Die Schaltfrequenz kann durch die PWM-Einheit unabhängig

vorgegeben und ist nicht von den anderen Parametern der Strecke oder dem Arbeitspunkt des Regelkreises abhängig. Die Reglerauslegung kann bei geeigneter Wahl der Schaltfrequenz rein mit den Mitteln der linearen Regelungstechnik anhand eines linearen Streckenmodells erfolgen.

3) In dieser Variante wird der schaltende Regler mit einer Übertragungsfunktion $G_R(s)$, die Tiefpassverhalten oder DT_n -Verhalten aufweist, gegengekoppelt. Dadurch entsteht ein schaltender Regler mit dynamischem Verhalten für die Mittelwerte der Ein- / Ausgangssignale.

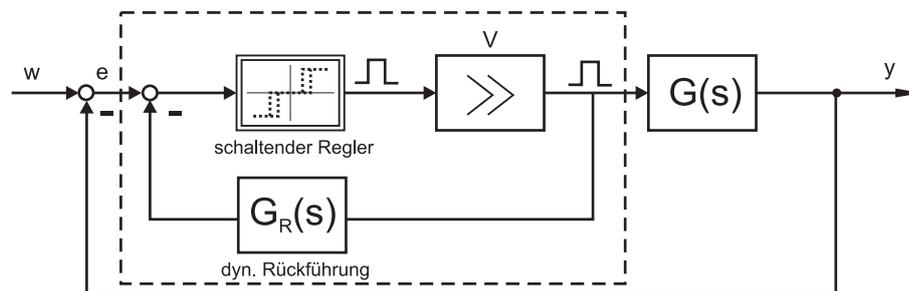


Bild 8.15: Schaltender Regler mit dynamischer Rückführung

Die äquivalente Übertragungsfunktion des Reglers für die Mittelwerte von Regelgröße und Stellgröße kann man nach Bild 8.16 dadurch berechnen, dass man die nichtlineare Reglerkennlinie einfach als ein lineares Element sehr hoher Verstärkung auffasst.

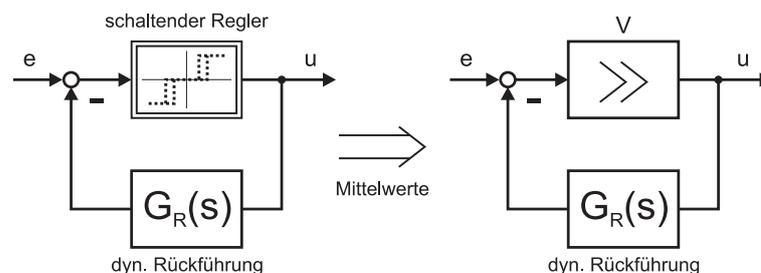


Bild 8.16: Äquivalente Übertragungsfunktion für rückgeführten Regler

Damit berechnet sich die äquivalente Übertragungsfunktion des Reglers zu:

$$\frac{\bar{u}(s)}{e(s)} = \frac{V}{1 + VG_R(s)} \stackrel{V \rightarrow \infty}{\approx} \frac{1}{G_R(s)} \quad (8.16)$$

Die Vorgehensweise nach Variante 3 ermöglicht schon eine recht gute Parametrierung des Reglers im Hinblick auf das Streckenmodell. Als Nachteil ist jedoch anzusehen, dass die

mittlere Schaltfrequenz des Reglers im Wesentlichen durch die Rückführung des Reglers bestimmt wird und damit nicht unabhängig von der Reglerauslegung eingestellt werden kann. Weiterhin ist die Schaltfrequenz natürlich auch vom Arbeitspunkt der Regelung und damit vom stationären und dynamischen Mittelwert der Stellgröße abhängig. Die variable Schaltfrequenz kann aber auch als Vorteil angesehen werden, da sie dazu führt, dass an den Rändern des Aussteuerbereiches der Stellgröße die Schaltfrequenz reduziert wird. Wegen der kleinen Pulsbreiten der Stellgröße in diesem Bereich führt dies im Allgemeinen nicht zu störenden Schaltüberschwingungen auf der Regelgröße und ist sogar erwünscht um die Wechselbelastung und damit den Verschleiß des schaltenden Stellgliedes zu reduzieren.

4) Mit der letzten Variante wird gerade der Punkt der unabhängig einstellbaren Schaltfrequenz adressiert. Der schaltende Reglerblock erhält neben der Gegenkopplung durch eine Übertragungsfunktion zur Einstellung der Reglerdynamik eine dynamische Mitkopplung mit höheren Eigenfrequenzen, um eine hochfrequente Schaltschwingung in Relation zu den Streckenzeitkonstanten zu erzeugen. Das Schaltelement erzeugt selbständig eine hochfrequente Schaltschwingung mit einer mittleren Frequenz, die von der Aussteuerung des Ausgangs abhängt. Diese Schaltschwingung wird dann durch die Gegenkopplung und die Regelabweichung in ihrem Mittelwert beeinflusst, um die Reglereigenschaften zu erzielen.

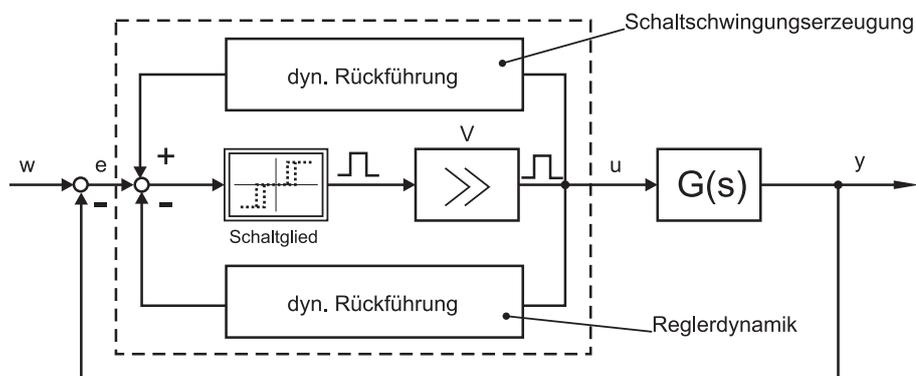


Bild 8.17: Schaltregler mit unabhängiger Schaltschwingungserzeugung und Reglerdynamik

Diese Vorgehensweise ist in der Industrietechnik recht verbreitet. Nicht nur analog aufgebaute Regler verwenden diese Systematik, sondern durchaus auch digital realisierte Regler. Die beschriebenen Rückführungen lassen sich ohne Weiteres in einer digitalen Signalverarbeitung implementieren.

8.3.3 Zweipunktregler an $PT1$ mit Laufzeitglied

Im Folgenden wird als Beispiel eine laufzeitbehaftete Verzögerungsstrecke erster Ordnung mit einem einfachen *Zweipunktregler mit unsymmetrischer Kennlinie* zusammengeschaltet. Die Struktur des Regelkreises ist in Bild 8.18 erkennbar.

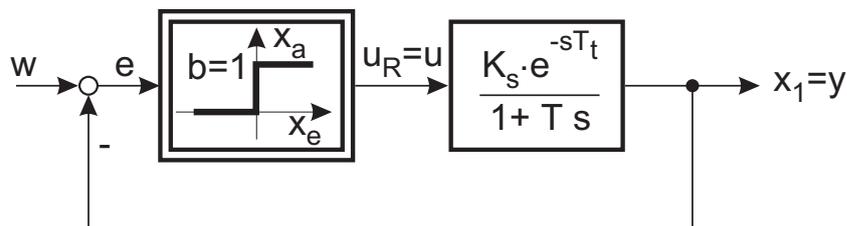


Bild 8.18: Regelkreis mit Zweipunktregler

Bei der Aufschaltung eines Sollwertsprunges $w(t) = w_0 \cdot \sigma(t)$ und unter der Annahme, dass $y(t) = 0$ für $t < 0$ ist, ergibt sich eine Regelabweichung:

$$e(t) = w_0 - y(t) \quad (8.17)$$

Die *Schaltbedingung* für den Zweipunktregler lautet:

$$u_R(t) = \begin{cases} 1 & \text{für } e(t) > 0 \\ 0 & \text{für } e(t) < 0 \end{cases} \quad (8.18)$$

Der qualitative Verlauf der Regelkreissignale $w(t)$, $y(t)$, $e(t)$ und $u(t) = u_R(t)$ ist in Bild 8.19 dargestellt.

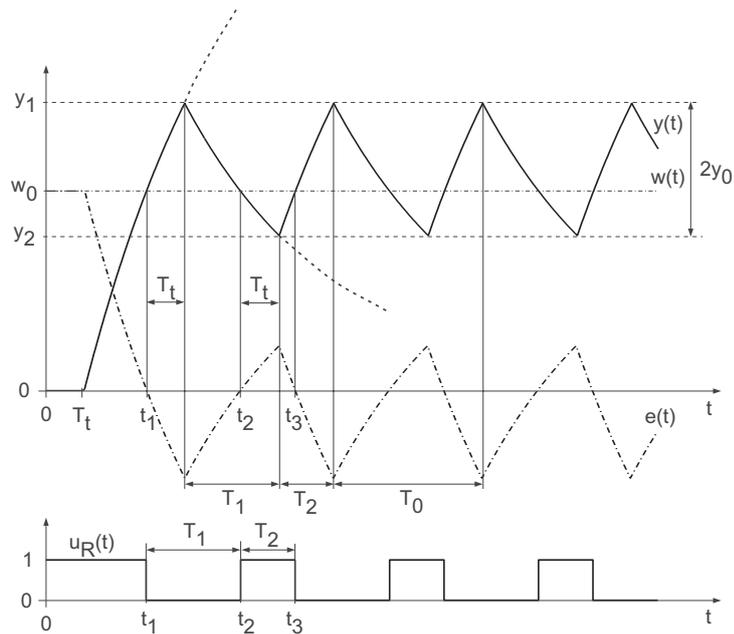


Bild 8.19: Signalverläufe des Regelkreises gemäß Bild 8.18

Die periodische Schwingung, die der Regelkreis vom Zeitpunkt $t = t_1$ ab ausführt, wird als *Arbeitsbewegung* bezeichnet. Dabei kennzeichnet diese Arbeitsbewegung nicht, wie bei linearen Systemen, ein grenzstabiles Systemverhalten, sondern ist typisch für nichtlineare Regelkreise.

Die Laufzeit T_t der Regelstrecke bewirkt hierbei, dass die Umschaltung des Zweipunktreglers um die Laufzeit verzögert erfolgt. Eine kleine Laufzeit T_t führt zu einem schnellen Einschwingen, allerdings steigt die *Schalzhäufigkeit*. Für $T_t \rightarrow 0$ geht die Schalzhäufigkeit gegen unendlich.

Im Weiteren werden die Kenndaten für die Arbeitsbewegung abgeleitet.

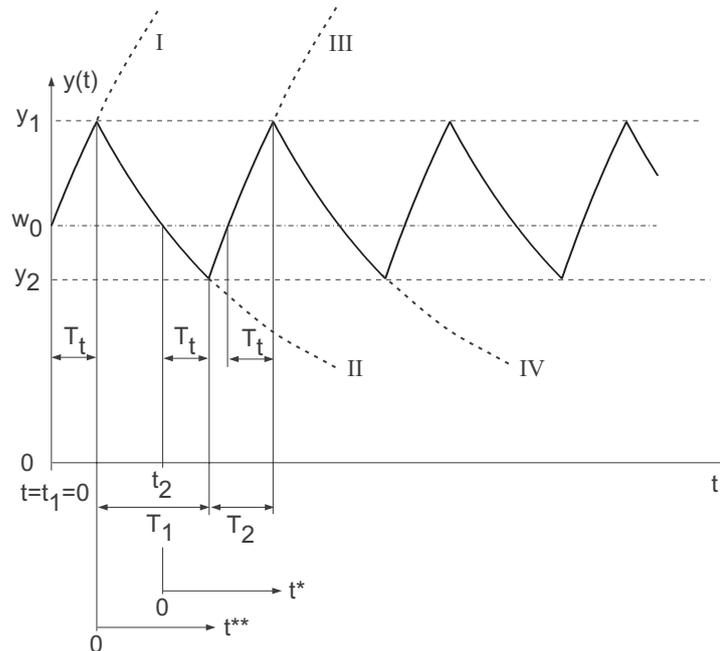


Bild 8.20: Arbeitsbewegung des Regelkreises mit Zweipunktregler

Die periodische Arbeitsbewegung lässt sich durch stückweise ansteigende (I, III, ...) und stückweise abfallende Zeitfunktionen (II, IV, ...) beschreiben. Für die Zeitfunktion I mit $t = t_1 = 0$ gilt:

$$\begin{aligned} y_I(t) &= w_0 + (K_S - w_0)(1 - e^{-t/T}) \\ &= K_S + (w_0 - K_S)e^{-t/T} \end{aligned} \quad (8.19)$$

Nach der Zeit T_t wird der Umkehrpunkt mit dem Wert

$$y_1 = y_I(T_t) = K_S + (w_0 - K_S)e^{-T_t/T} \quad (8.20)$$

erreicht.

Für die Zeitfunktion y_{II} erhält man in einem um die Zeit t_2 verschobenen Koordinatensystem

$$y_{II}(t^*) = y_1 e^{-t^*/T} \quad (8.21)$$

und zum Zeitpunkt $t^* = T_t$ den unteren Umkehrpunkt

$$y_2 = y_{II}(T_t) = w_0 e^{-T_t/T} \quad . \quad (8.22)$$

Durch eine weitere Koordinatenverschiebung lässt sich die abfallende Zeitfunktion II auch durch

$$y_{II}(t^{**}) = y_1 e^{-t^{**}/T} \quad (8.23)$$

beschreiben. Für $t^{**} = T_1$ erreicht man gerade den Umkehrpunkt y_2 .

Durch Einsetzen der obigen Gleichungen in Gleichung (8.23) errechnet sich T_1 zu

$$T_1 = T \ln \left(\frac{K_S + (w_0 - K_S) e^{-T_t/T}}{w_0 e^{-T_t/T}} \right) \quad . \quad (8.24)$$

In analoger Weise erhält man

$$T_2 = T \ln \left(\frac{w_0 e^{-T_t/T} - K_S}{(w_0 - K_S) e^{-T_t/T}} \right) \quad (8.25)$$

und somit die *Schwingungsdauer* T_0 der Arbeitsbewegung

$$\begin{aligned} T_0 &= T_1 + T_2 \\ &= T \ln \left(\frac{\left(K_S e^{T_t/T} + w_0 - K_S \right) \left(-w_0 + K_S e^{T_t/T} \right)}{w_0 (K_S - w_0)} \right) \\ &= T \ln \left(\frac{K_S^2 e^{T_t/T} \left(1 - e^{T_t/T} \right) + w_0 (w_0 - K_S)}{w_0 (w_0 - K_S)} \right) \quad . \quad (8.26) \end{aligned}$$

Der Verlauf der Frequenz über dem Sollwert ist in Bild8.21 gezeigt.

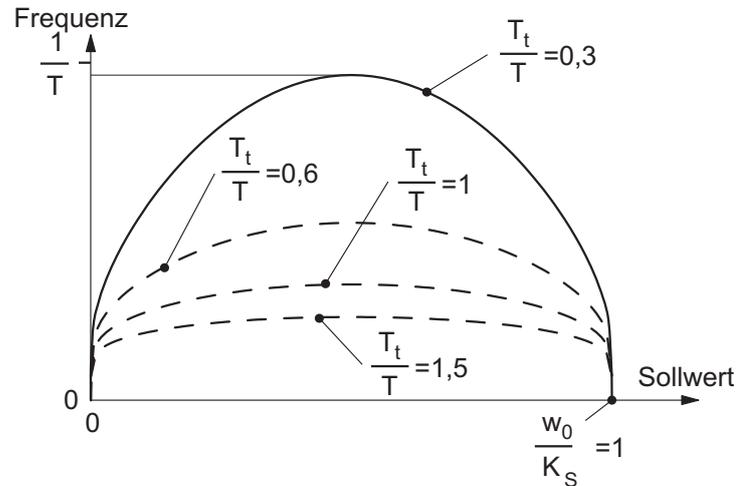


Bild 8.21: Schaltfrequenz

Die *Schwingungsamplitude* der Arbeitsbewegung errechnet sich zu:

$$\begin{aligned}
 y_0 &= \frac{1}{2}(y_1 - y_2) \\
 &= K_S \left(1 - e^{-T_t/T}\right)
 \end{aligned} \tag{8.27}$$

Die *mittlere Regelabweichung* beträgt:

$$\begin{aligned}
 \bar{e} &= \frac{y_1 + y_2}{2} - w_0 \\
 &= \left(\frac{1}{2}K_S - w_0\right) \left(1 - e^{-T_t/T}\right)
 \end{aligned} \tag{8.28}$$

Die mittlere Regelabweichung kann je nach Größe von K_S positiv oder negativ sein, dagegen ist sie für $K_S = 2w_0$ gleich Null. Aus Bild 8.19 wird auch ersichtlich, dass für das Zustandekommen einer Arbeitsbewegung $K_S > w_0$ gelten muss.

8.4 Analyse schaltender Regler im Zustandsraum

Wie aus Abschnitt 7.2.1 bekannt ist, ergeben Beschleunigungsstrecken (zwei Integratoren in Reihe) sehr übersichtliche Zustandskurven. Da diese Regelstrecken auch in der Praxis vorkommen, sollen sie in Verbindung mit verschiedenen Zweipunktreglern untersucht werden.

8.4.1 Hysterese freier Zweipunktregler

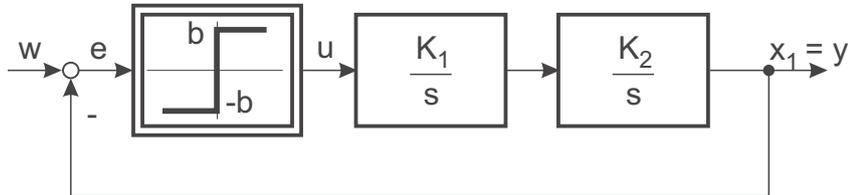


Bild 8.22: Regelkreis mit Zweipunktglied mit doppeltem Integrator

Aus der Differenzialgleichung

$$\ddot{y} = K_1 K_2 u$$

ergibt sich, da u nur die beiden Werte b und $-b$ annehmen kann,

$$u = \varepsilon b, \quad \varepsilon = \pm 1,$$

$$\ddot{y} = K_1 K_2 b \varepsilon = \varepsilon K \quad .$$

Dies ist aber die bekannte Differenzialgleichung aus Abschnitt 7.2.1. Als ihre Lösungen ergeben sich die bereits bekannten Parabelscharen (Anfangszustandswerte $x_i \neq 0$), die im folgenden Bild 8.23 dem Leser in Erinnerung gerufen werden sollen.

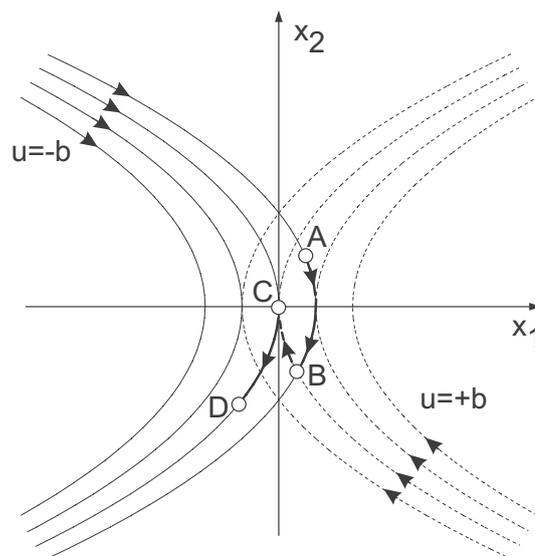


Bild 8.23: Trajektorien des linearen Teilsystems

Da nur Stellgrößen mit $\pm b$ auftreten können, bewegen sich die Zustandspunkte der Regelstrecke nur auf Kurven der beiden Scharen. Ändert sich die Stellfunktion $u(t)$ bei offenem Regelkreis zu bestimmten Zeitpunkten, so wechselt die Trajektorie von einer Parabel der einen Schar auf eine andere Parabel der anderen Schar (vgl. Verlauf A – B – C – D in Bild 8.23).

Betrachtet man nun den geschlossenen Kreis, so ist

$$u = \begin{cases} -b & \text{falls } e < 0, \text{ also } y > 0 \\ b & \text{falls } e > 0, \text{ also } y < 0 \end{cases}$$

Wenn $x_1 = y$ durch Null geht, erfolgt die Umschaltung des Zweipunktgliedes. Die Gerade $x_1 = 0$, also die x_2 -Achse, wird als *Schaltlinie* oder *Schaltgerade* des geschlossenen Kreises bezeichnet.

Ist die Führungsgröße $w \neq 0$, dann verschieben sich die Trajektorien entsprechend dem Betrag von w längs der x_1 -Achse.

Der Zustand durchläuft eine geschlossene Kurve, wie sie in Bild 8.24 dargestellt ist. Sie setzt sich aus der Bewegung auf 2 Parabelästen zusammen. Im Zeitbereich entspricht dies einer Dauerschwingung.

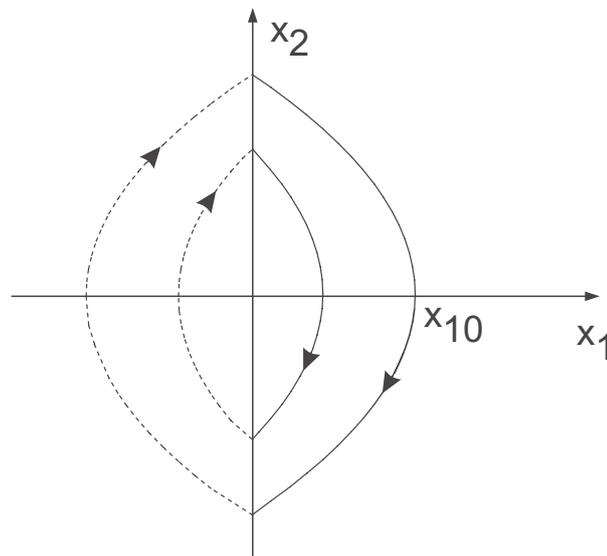


Bild 8.24: Trajektorien des geschlossenen Regelkreises

Die Periode der Dauerschwingung hängt vom Anfangszustand ab.

Aus dem Trajektorienverlauf erkennt man auch das Stabilitätsverhalten der Ruhelage $(x_1, x_2) = \underline{0}$. Sie ist *stabil*, aber *nicht asymptotisch stabil*: Die Trajektorien bleiben in einer beliebig engen Umgebung der Ruhelage, wenn sie hinreichend nahe bei ihr beginnen, ohne jedoch für $t \rightarrow \infty$ gegen die Ruhelage zu streben.

Dieses Verhalten ist für eine Anwendung nicht wünschenswert. Im folgenden Abschnitt wird gezeigt, wie dennoch mit einer einfachen Rückführung der Regelkreis stabilisiert werden kann, damit er ein asymptotisch stabiles Verhalten aufweist.

8.4.2 Beschleunigungsstrecke mit Zweipunktregler und Rückführung

Um die Schwingung zu dämpfen, kann man daran denken, ähnlich wie bei einem linearen Regelkreis eine differenzierende Rückführung zu verwenden. Ist die Zustandsgröße x_2 direkt verfügbar, erübrigt sich die sonst nur näherungsweise verwirklichtbare Differenziation der Regelgröße. Man erhält die folgende Anordnung:

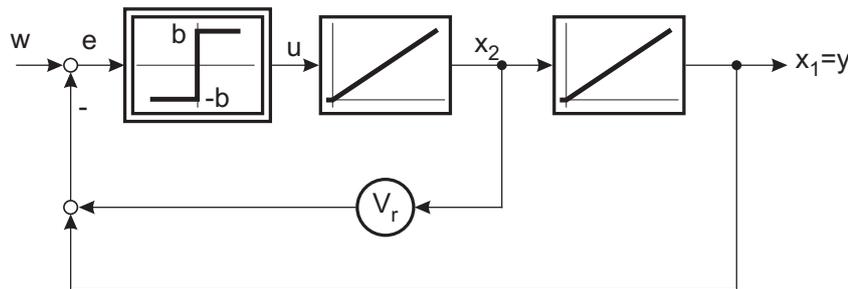


Bild 8.25: Idealer Zweipunktregler mit Rückführung

Bei der Annahme eines symmetrischen Zweipunktreglers mit $x_a = \pm b$ lautet die Differentialgleichung

$$\ddot{y}(t) = b \operatorname{sign}(w - y - V_r \cdot x_2), \quad (8.29)$$

die Umschaltung der Stellgröße erfolgt jetzt bei

$$y = w - V_r x_2, \quad (8.30)$$

d. h. auf einer unter dem Winkel $\alpha = \arctan V_r$ gegen die x_2 -Achse geneigten Schaltgerade, vgl. Bild 8.26.

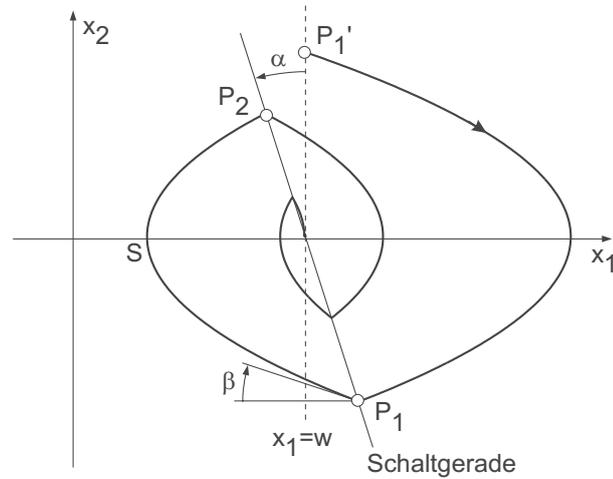
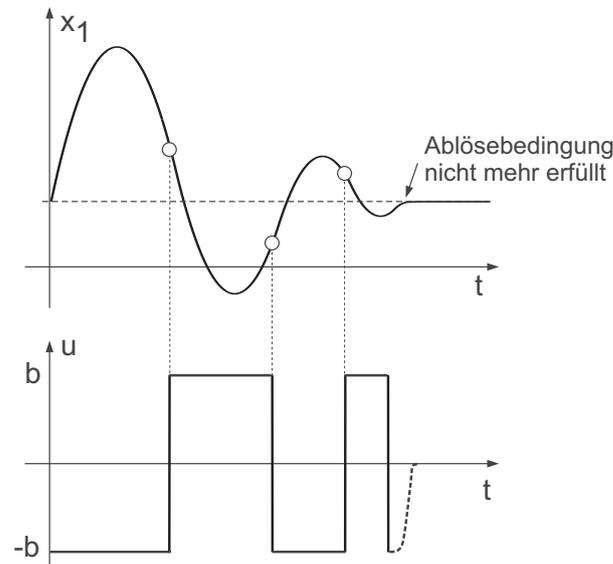


Bild 8.26: Zustandstrajektorien mit Rückführung

Der Vorhalt bewirkt also eine Umschaltung vor Erreichen des Abgleichs der Regelgröße y . Das Bild 8.26 lässt auch erkennen, dass der erzielte Dämpfungseffekt mit der Erhöhung der Verstärkung V_r zunimmt. Mit kleiner werdender Schwingungsamplitude erhöht sich gleichzeitig die Frequenz, da zunehmend kürzere Abschnitte der gleichen Zustandskurven durchlaufen werden. Der zum Punkt P_1 symmetrisch liegende Punkt P_1' wird nicht mehr erreicht, da bereits bei P_2 umgeschaltet wird.

In Bild 8.27 ist der aus Parabelbögen zusammengesetzte gedämpfte Verlauf von y über der Zeit aufgetragen. Der Anstieg der Frequenz ist auch hier zu erkennen.

Bild 8.27: Zeitlicher Verlauf von y und u_r

Aufgrund der senkrechten Tangente der Zustandsparabeln im jeweiligen Scheitelpunkt nimmt der Neigungswinkel

$$\beta = \arctan \left(-\frac{dx_2}{dx_1} \right)$$

der Zustandskurve nach der Umschaltung in P_1 mit sinkender Amplitude ständig zu (analog in P_2).

Nach einer endlichen Anzahl von Schwingungen tritt somit der Fall ein, dass $\beta \geq \frac{\pi}{2} - \alpha$ ist und die Zustandskurve sich nicht mehr von der Schaltgeraden löst. Der Zustandspunkt wird dann nach jeder Umschaltung sofort wieder auf die ursprüngliche Seite der Schaltkurve zurückgetrieben. Der Zustandspunkt *gleitet* entlang der Schaltgeraden aufgrund der hohen Umschaltfrequenz in den Ursprung. Effekte dieser Art werden im Englischen als *Sliding Mode* oder auch als *Variabla Structure Control* bezeichnet. Da dieser Gleitvorgang die Stellfunktion u zwischen den Werten b und $-b$ oszillieren lässt, wird die Stelleinrichtung in dieser Phase stark belastet. Der Gleitvorgang sollte sich daher nicht zu lange ausdehnen, d. h. die Schaltgerade darf nicht zu stark gegen die x_2 -Achse geneigt sein.

8.4.3 Beschleunigungsstrecke mit Zweipunktregler mit Hysterese

Reale Zweipunktregler haben immer eine mehr oder weniger stark ausgeprägte Hysterese. Daher ersetzt man das ideale Zweipunktglied in Bild 8.28 durch ein Zweipunktglied mit

Hysterese. Man erhält das folgende Blockschaltbild.

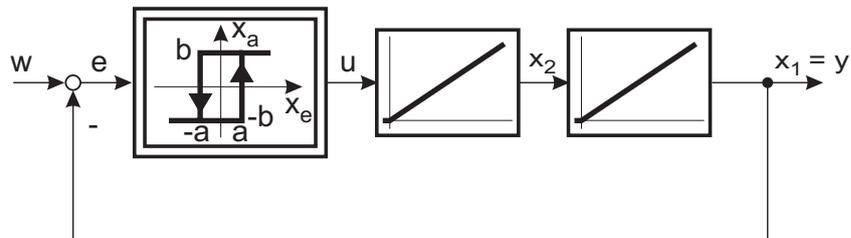


Bild 8.28: Hysteresebehafteter Zweipunktregler mit einer doppelt integrierenden Strecke

Das Besondere der Hysteresefunktion ist die Zweideutigkeit im Bereich $-a < x_a < a$. Die Ausgangsgröße des Zweipunktgliedes x_a ist dort nicht von $x_e = e$, sondern von der Richtung abhängig, aus welcher $e(t)$ zuletzt in diesen Bereich eingetreten ist. Die Hysteresefunktion hat also die Eigenschaft eines elementaren Gedächtnisses.

Die Umschaltbedingung lautet:

- $w - y = e = a \Rightarrow u = +b$
- $w - y = e = -a \Rightarrow u = -b$

Man erhält also, wie in Bild 8.29 eingezeichnet, zwei Schaltgeraden parallel zur x_2 -Achse. Die Hysterese hat zur Folge, dass die Umschaltung verspätet erfolgt. Dadurch entsteht eine Schwingung, deren Amplitude in jeder Halbwellen um den Wert $2a$ anwächst, andererseits nimmt die Frequenz der Schwingung monoton ab.

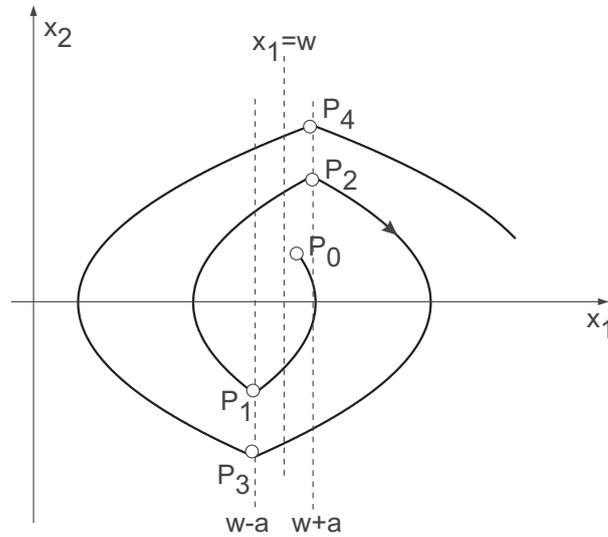


Bild 8.29: Trajektorien eines doppelten Integrators mit hysteresebefahetem Zweipunktregler

Eine solche Anordnung ist für die Anwendung aufgrund der aufklingenden Schwingungsamplitude völlig unbrauchbar.

8.4.4 Beschleunigungsstrecke mit Zweipunktregler mit Hysterese und Rückführung

Ergänzt man den Regelkreis mit dem hysteresebefaheten Zweipunktglied nun zusätzlich um eine differenzierende Rückführung, unter der Voraussetzung natürlich, dass die Zustandsgröße x_2 überhaupt gemessen werden kann, so erhält man das in Bild 8.30 dargestellte Blockschaltbild und den in Bild 8.31 dazugehörigen Trajektorienverlauf.

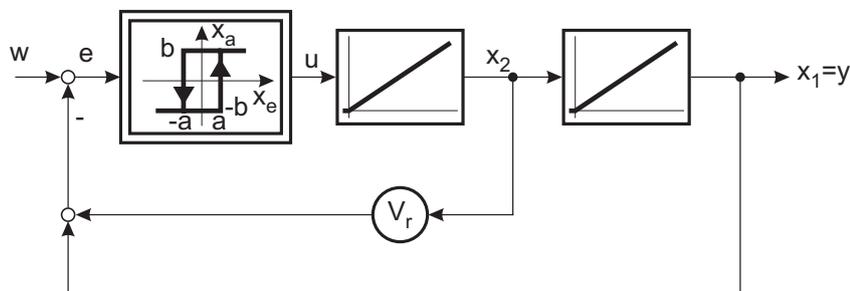


Bild 8.30: Doppelter Integrator mit hysteresebefahetem Zweipunktregler und differenzierender Rückführung

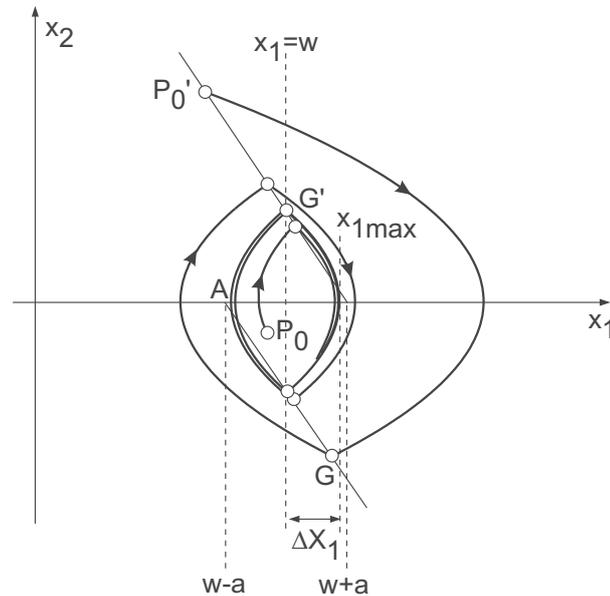


Bild 8.31: Trajektorienverlauf zum Blockschaltbild 8.30

Die Umschaltbedingung lautet $y = w \mp a - V_r x_2$. Dies sind zwei um den Winkel $\alpha = \arctan V_r$ gegen die x_2 -Achse gedrehte parallele Schaltgeraden.

Bild 8.31 zeigt die Zustandskurven für zwei verschiedene Anfangszustände. Dabei ist zu beachten, dass bei kleinen Schwingungsamplituden die Umschaltung später erfolgt, so dass die Schwingung angefacht verläuft. Hier überwiegt also der Einfluss der Hysterese. Bei großen Amplituden ist der Verlauf umgekehrt. Hier schaltet der Regler infolge des Vorhalts vorzeitig um, die Schwingung wird gedämpft.

Beide Vorgänge konvergieren gegen eine stationäre Grenzschiwingung, die dann erreicht ist, wenn die beiden Umschaltpunkte G und G' symmetrisch zur x_1 -Achse liegen. Nur die Schaltgerade $x_1 = y = w$ erfüllt diese Bedingung. Die Wirkungen von Hysterese und Vorhalt heben sich hier gerade auf.

Der Grenzyklus trennt somit Bereiche der Zustandsebene, in denen die Vorgänge angefacht bzw. gedämpft verlaufen. Mögliche zeitliche Verläufe solcher Schwingungsvorgänge sind im Bild 8.32 dargestellt.

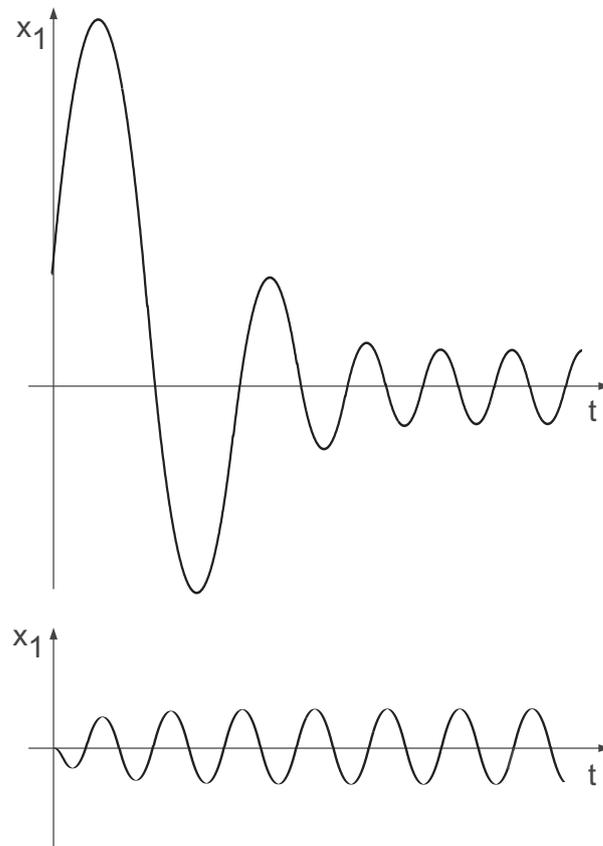


Bild 8.32: Ausbildung von Grenzwahlungen für zwei verschiedene Anfangszustände

8.5 Beschleunigungsstrecke mit Dreipunktregler

Bei der Verwendung von Dreipunktreglern kann die Stellgröße $u(t)$ neben den Werten $\pm b$ auch den Wert Null annehmen. Dies äußert sich im Zustandsdiagramm durch das zusätzliche Auftreten eines geradlinigen Astes. Die wesentlichen Eigenschaften des Regelkreises bleiben jedoch unverändert. Dies wird im folgenden gezeigt.

8.5.1 Dreipunktregler mit Hysterese

Ersetzt man in Bild 8.28 das Zweipunktglied durch ein Dreipunktglied mit Hysterese erhält man folgendes Blockschaltbild.

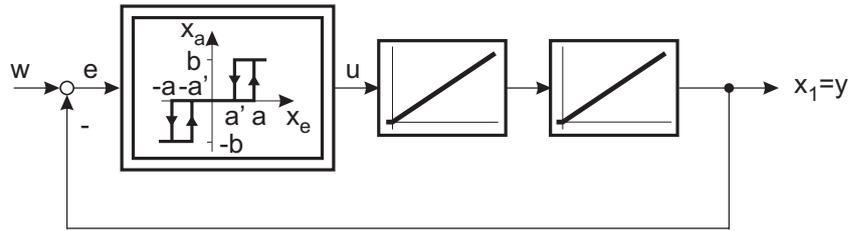


Bild 8.33: Beschleunigungsstrecke mit hysteresebefahertem Dreipunktregler

Die Einschaltbedingung lautet jetzt

$$e = w - y = \pm a \quad (8.31)$$

und die Ausschaltbedingung

$$e = w - y = \pm a' \quad (8.32)$$

Man erhält also 4 Schaltgeraden parallel zur x_2 -Achse, die in die Zustandsebene eingetragen werden, vgl. Bild 8.34.

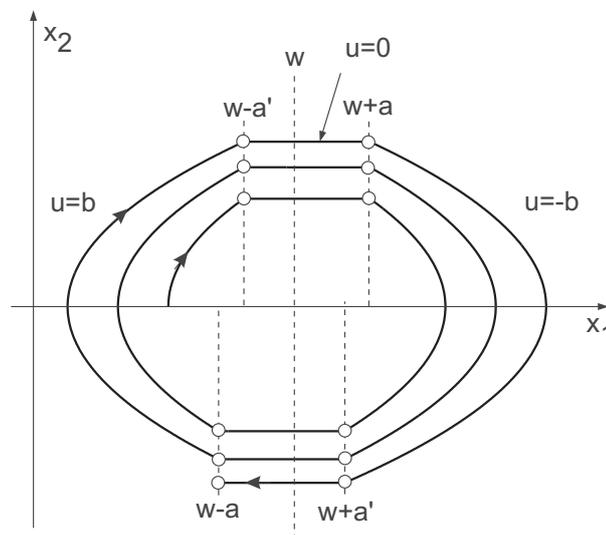


Bild 8.34: Trajektorien einer Beschleunigungsstrecke mit Dreipunktregler

Für $u = \pm b$ werden die bekannten Parabelbögen durchlaufen, während für $u = 0$ die Lösung

$$x_2 = \text{const.}, \quad (8.33)$$

d. h. die Gleichung einer zur x_1 -Achse parallelen Geraden entsteht. Die Ausgangsgröße bewegt sich im Unempfindlichkeitsbereich, also mit konstanter Geschwindigkeit. Wie bild 8.34 zeigt, entsteht wieder eine aufklingende Schwingung, deren Amplitude je Halperiode um den Wert $\Delta x_1 = a - a'$ zunimmt.

8.5.2 Dreipunktregler mit Hysterese und Rückführung

Führt man jetzt wiederum eine Rückführung von x_2 ein, so erhält man das Blockschaltbild 8.35.

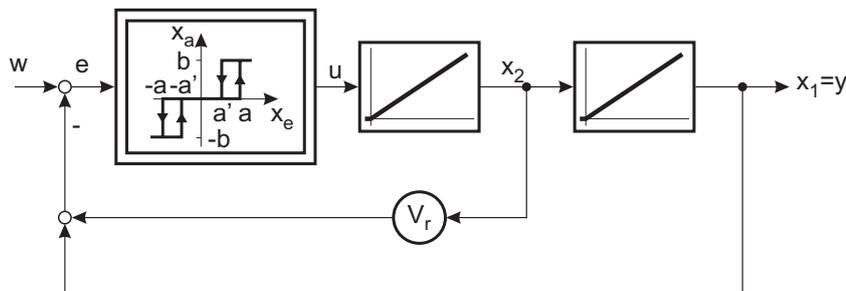


Bild 8.35: Beschleunigungsstrecke mit Dreipunktregler und Rückführung

Für den symmetrischen Dreipunktregler mit Hysterese lautet die Einschaltbedingung jetzt

$$e = w - y - V_r x_2 = \pm a \quad (8.34)$$

und die Ausschaltbedingung

$$e = w - y - V_r x_2 = \pm a' \quad . \quad (8.35)$$

Dem entsprechen 4 parallele und unter dem Winkel $\alpha = \arctan V_r$ gegen die x_2 -Achse geneigte Schaltgeraden, vgl. Bild 8.36.

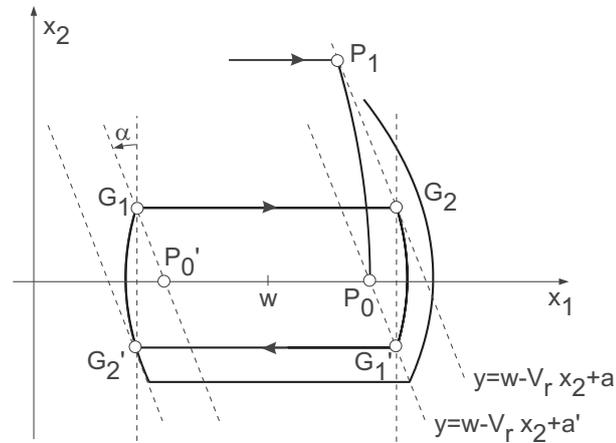


Bild 8.36: Trajektorien zum Regelkreis 8.35

Infolge der Hysterese und des Vorhalts gibt es auch hier eine stationäre Grenzschwingung, die durch die Umschaltpunkte $G_1 - G_2 - G_1' - G_2'$ gekennzeichnet wird. Schwingungen mit größerer Amplitude verlaufen wieder gedämpft, solche mit kleinerer Amplitude angefacht, so dass der Vorgang bei beliebigem Anfangszustand schließlich in den Grenzzyklus einmündet.

Mit einer Beschleunigungsstrecke kommt also auch bei Verwendung eines Dreipunktreglers die stationäre Schwingung normalerweise nicht zum Erliegen. Ein Ruhezustand ist nur unter speziellen Bedingungen möglich, nämlich wenn der Regler die Stellgröße in einem Punkt P_0 oder P_0' , d. h. bei $y = w \pm a'$, $x_2 = 0$ abschaltet.

9 Zeitoptimale Steuerung/Regelung

In technischen Systemen ist es oftmals gewünscht, ein dynamisches System in kürzest möglicher Zeit von einem Anfangszustand zu einem Zielzustand zu bewegen. Bei einem linearen System lässt sich durch Einsatz großer Stellgrößen diese Zeit beliebig kurz halten. Daher ist die Aufgabe nur in Verbindung mit einer begrenzten Stellgröße sinnvoll gestellt. Als Zusatzbedingung tritt hinzu, dass der Endzustand nicht vorübergehend (im Vorbeiflug) erreicht wird, sondern dass das System im Zielzustand stationär eingeschwungen und damit in Ruhe ist. Das zeitoptimale Steuerungs-Problem lässt sich damit folgendermaßen formulieren: Gesucht ist der beschränkte Verlauf der Eingangsgröße eines dynamischen Systems, der die Ausgangsgröße in möglichst kurzer Zeit von einem gegebenen Anfangszustand seinen stationären Zielzustand überführt.

Einen Ansatz zur Lösung dieses Problems kann man am Besten anhand einer reinen Tiefpassstrecke, die keine Zählernullstellen aufweist, veranschaulichen.

Eine reine Tiefpassstrecke hat die Übertragungsfunktion:

$$G(s) = \frac{b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} \quad (9.1)$$

und das Blockschaltbild:

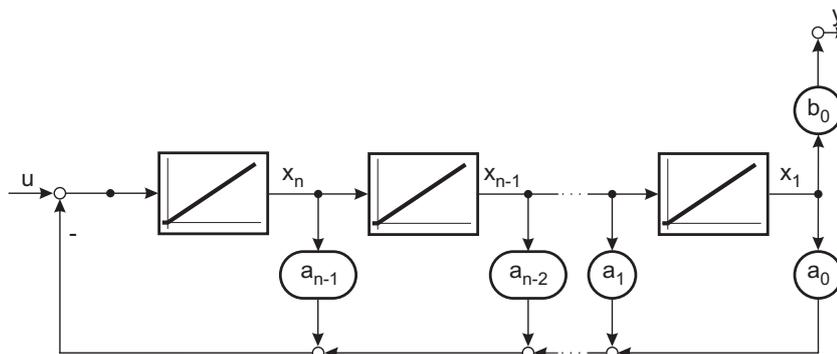


Bild 9.1: reine Tiefpass-Strecke

Wird auf den Eingang ein Sprung gegeben, so schwingt die letzte Zustandsgröße, die ja zugleich der Ausgangsgröße entspricht, auf einen stationären Endwert ein, der von Null verschieden ist. Die einzelnen Zustandsgrößen bilden die Ableitungen der Ausgangsgröße und kehren nach einer transienten Bewegung wieder auf den Wert Null zurück.

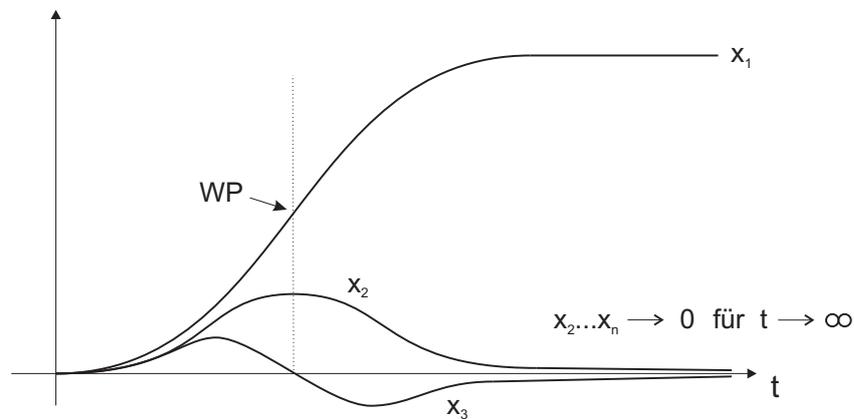


Bild 9.2: Sprungantworten der Tiefpass-Strecke

Der optimale Verlauf der Stellgröße muss gleichzeitig mehrere Bedingungen erfüllen: Die Ausgangsgröße muss den gewünschten stationären Endwert erreichen und alle inneren Zustandsgrößen müssen wieder zu Null geführt werden. Bei der Suche nach diesem Verlauf der Eingangsgröße hilft das *Pontryaginsche Maximum-Prinzip*. Danach ist der zeitliche Verlauf der Eingangsgröße so gestaltet, dass er zu jedem Zeitpunkt den maximalen Betrag der Stellgröße ausnutzt. Bei der *zeitoptimalen Steuerung* ist der Verlauf der Stellgröße in n Abschnitte gemäß der Streckenordnung n unterteilt, in denen jeweils abwechselnd der maximale positive oder negative Betrag verwendet wird. Am Ende des letzten Abschnitts wird die Stellgröße vom maximalen Betrag auf den stationären Endwert umgeschaltet. Der Vorgang startet mit dem positiven Maximalwert der Stellgröße, dann folgen $(n - 1)$ Umschaltungen zwischen dem positiven und negativen Maximalwert der Stellgröße und anschließend eine letzte Umschaltung auf den stationär notwendigen Endwert.

9.1 Zeitoptimale Steuerung eines PT_1 -Gliedes

Im folgenden Diagramm ist die zeitoptimale Steuerung am Beispiel eines PT_1 dargestellt.

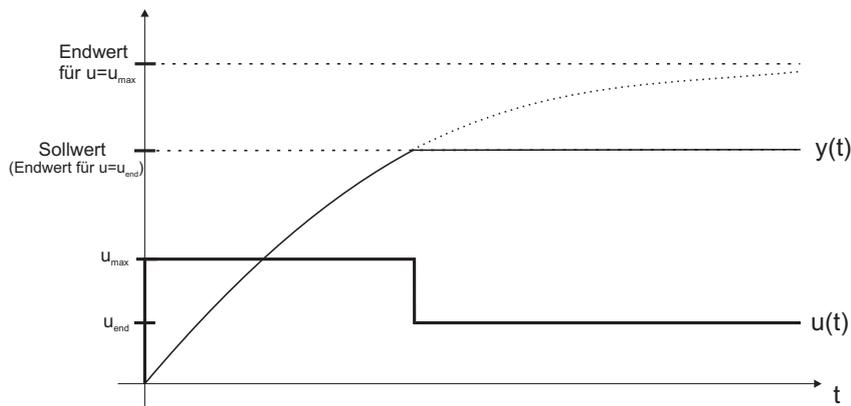


Bild 9.3: Zeitoptimale Steuerung einer PT1-Strecke

Die zu steuernde Strecke ist erster Ordnung; daher ist nur ein Schaltblock notwendig. Vom Einschalten an wird die maximale positive Stellgröße so lange gehalten, bis die gewünschte Ausgangsgröße erreicht ist. Dann wird auf den stationären Endwert umgeschaltet. Damit wird die Eingangsgröße des einzigen Integrators Null und die Ausgangsgröße bleibt konstant. Der stationär notwendige Wert der Stellgröße beträgt $u_{\text{end}} = \frac{\text{Sollwert}}{V_{PT1}}$.

9.2 Zeitoptimale Steuerung eines doppelten Integrators

Auch für einen doppelten Integrator kann man den optimalen Verlauf der Stellgröße erraten. Die Bezeichnungen der Signale ergeben sich aus dem nachfolgenden Blockschaltbild:

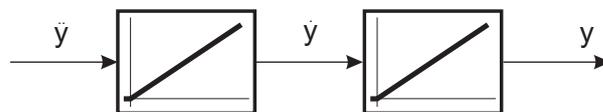


Bild 9.4: Doppelter Integrator

Für die Stellgröße u sind lediglich die Werte u_{\max} , u_{\min} und Null zulässig. Das System hat zwei Zustandsgrößen, die Ausgangsgrößen der beiden Integratoren, y und \dot{y} .

Die Stellgröße wird in der ersten Hälfte des Bewegungsvorganges auf den positiven Maximalwert gesetzt. Dies führt zu einem zeitlich linearen Anstieg der Geschwindigkeit \dot{y} . In der zweiten Hälfte wird die Eingangsgröße auf den negativen Maximalwert gesetzt; dies

führt die Geschwindigkeit linear wieder zu Null, so dass der Zielpunkt genau beim Stillstand erreicht wird. Der Verlauf der Stellgröße und Zustandsgrößen sind in der folgenden Grafik dargestellt:

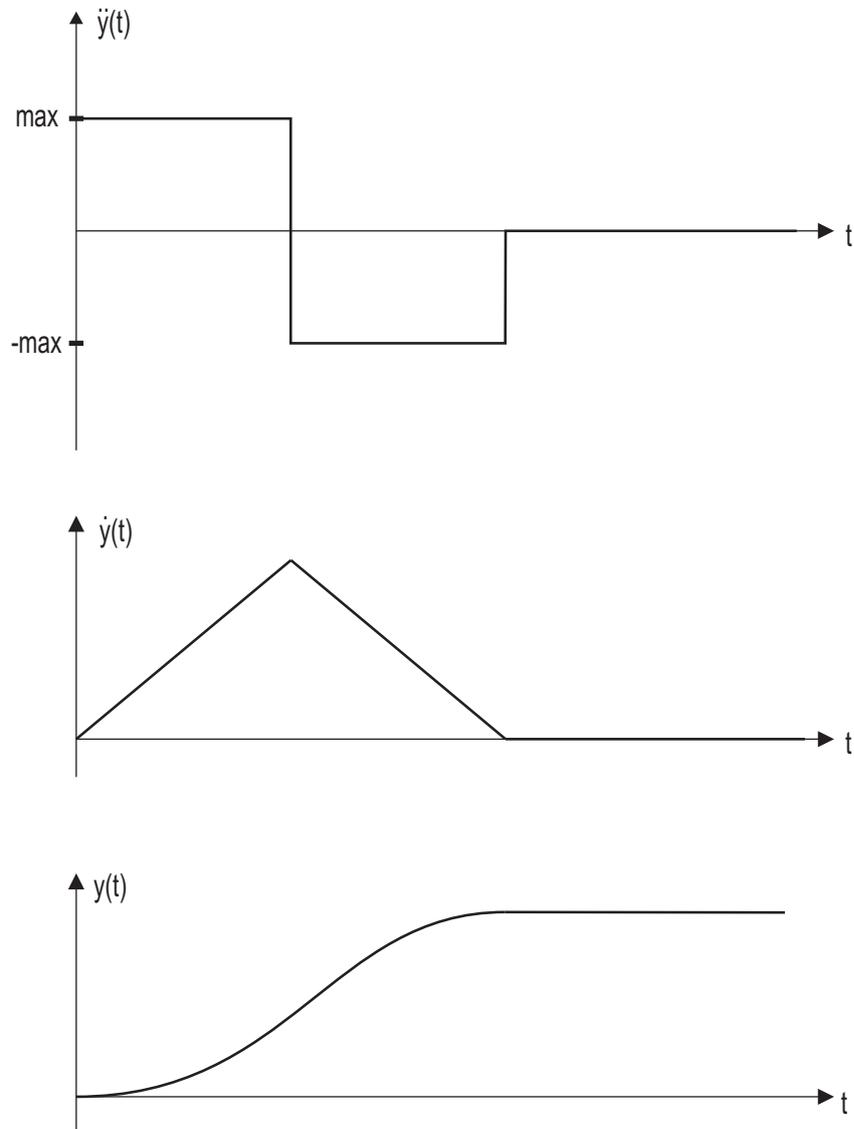
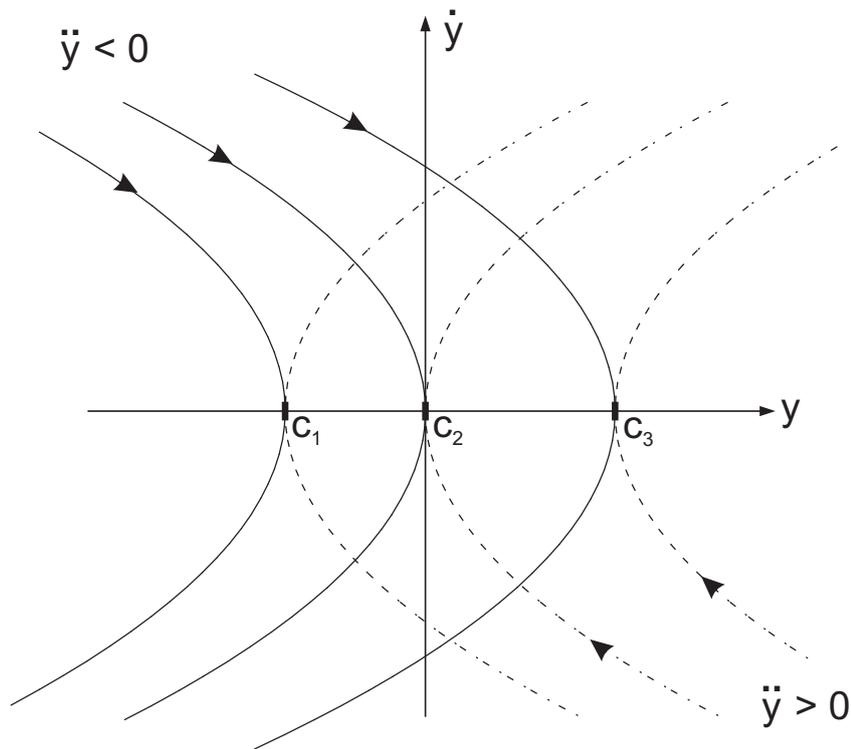
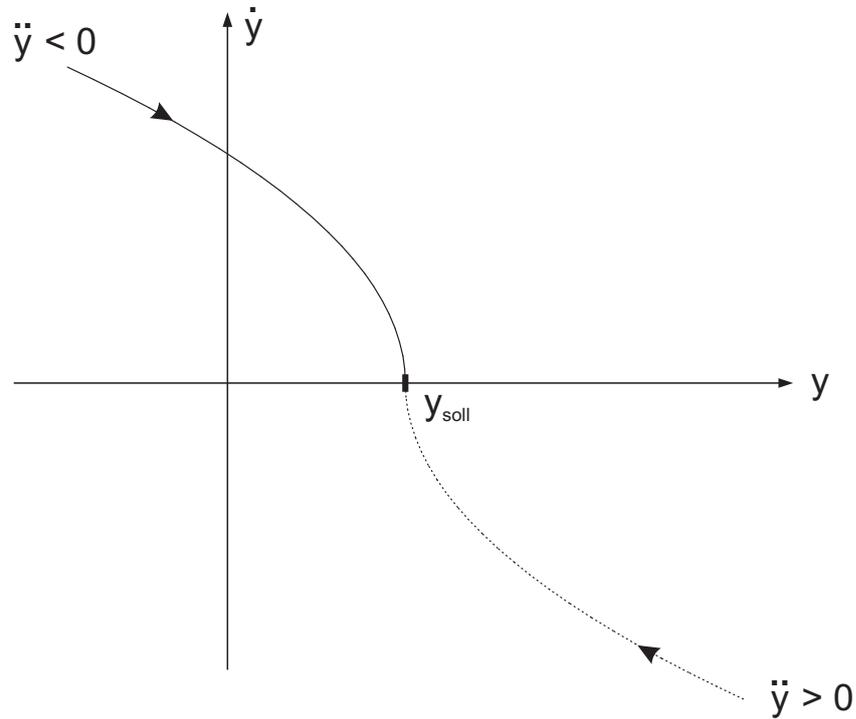


Bild 9.5: Zustandsgrößen einer I_2 -Strecke

Diese Verläufe lassen sich genauso im Zustandsraum darstellen (Bild 9.6). Dabei ergeben sich die Größen c_1 , c_2 und c_3 als Umkehrpunkte der Ausgangsgröße y . Sie stellen gleichzeitig unterschiedliche Integrationskonstanten dar, die wiederum aus unterschiedlichen Anfangsbedingungen resultieren. Diese Zusammenhänge wurden bereits in Kap. 7.2.1) hergeleitet.

Bild 9.6: Trajektorien einer I_2 -Strecke

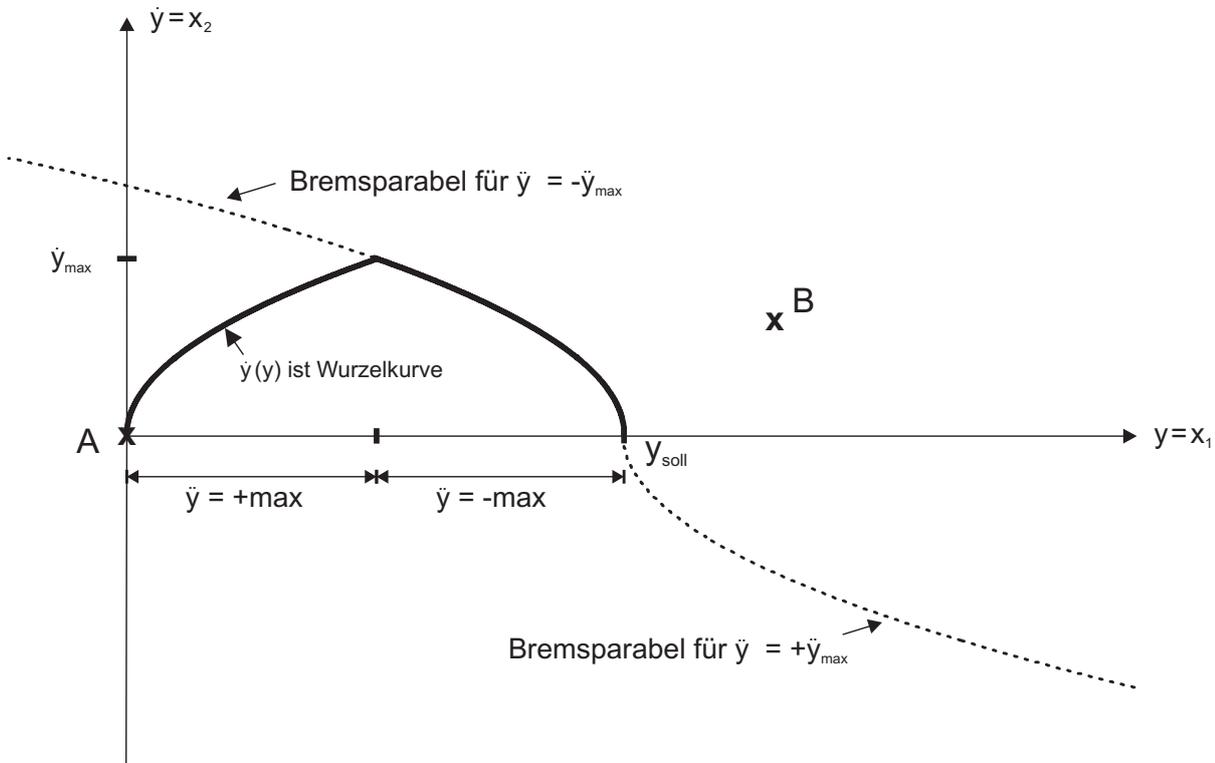
An den Umkehrpunkten c_1 , c_2 und c_3 ist jeweils ein Gleichgewichtspunkt des Systems erreicht, der stationär gehalten werden kann, wenn die Stellgröße auf Null geschaltet wird. Man kann sie daher als Zielpunkte y_{soll} oder Sollwerte interpretieren. Für jeden Zielpunkt gibt es für einen gegebenen Wert des Betrages der Stellgröße genau eine Trajektorie für konstante positive und negative Beschleunigung, die genau dort enden. Diese Trajektorien sind in Bild 9.7 veranschaulicht. Will man also von einem beliebigen Startpunkt y_0, \dot{y}_0 aus das Ziel $y_{soll}, \dot{y} = 0$ erreichen, kann dies in zwei Schritten erfolgen. Zunächst muss das Vorzeichen der maximalen Stellgröße so gewählt werden, dass das System auf eine der beiden Zieltrajektorien hin läuft, dann wird das Vorzeichen beim Erreichen der Zieltrajektorie umgeschaltet, um direkt auf der Zieltrajektorie in den Zielpunkt zu laufen. Dann muss nur noch bei Geschwindigkeit Null im Zielpunkt auch die Beschleunigung auf Null umgeschaltet werden. Man kann also beide Zieltrajektorien zusammen als eine Schaltlinie auffassen, auf der die zugehörige Stellgröße für den Zieleinlauf angelegt wird.

Bild 9.7: optimale Schaltlinie für Endwert y_{soll}

Die beiden Äste der Schaltlinie werden auch *Bremssparabeln* genannt. Von einem beliebigen Startpunkt aus muss immer erst die nächstliegende dieser beiden Bremssparabeln angesteuert werden, um dann darauf in den Zielpunkt einzulaufen. In diesem Fall muss für alle Startpunkte in der Zustandsebene, die links der Bremssparabeln liegen zunächst die positive Stellgröße verwendet werden. Das führt das System auf die obere Bremssparabel, auf der dann mit negativer Beschleunigung ins Ziel eingelaufen wird. Für Startpunkte rechts der Bremssparabeln wird mit negativer Beschleunigung die untere Bremssparabel erreicht, die dann mit positiver Beschleunigung in den Zielpunkt einbremst.

Das vorliegende System eines doppelten Integrators ist zweiter Ordnung; also benötigt man nach Pontryagin zwei Signal-Zeit-Blöcke maximaler Amplitude, um das Ziel zu erreichen.

Der Vorgang ist in Bild 9.8 dargestellt.

Bild 9.8: I_2 -Strecke im Zustandsraum

Falls der Anfangszustand sich im Bereich links unterhalb der Parabeln befindet, so muss zuerst $\ddot{y} = +u_{max}$ als Stellgröße angelegt werden. In Bild 9.8 mit Punkt A als Anfangszustand ist dies der Fall. Die Zustandsgröße \dot{y} läuft entlang ihrer Zustandskurve für $\ddot{y} = +u_{max}$ (vgl. Bild 9.6). Dies wird solange beibehalten, bis die obere Bremsparabel (und damit die Schaltlinie) erreicht ist. Dort wird auf $\ddot{y} = u_{min}$ umgeschaltet, so dass \dot{y} auf der Bremsparabel in den gewünschten stationären Endwert läuft. Für den oberen Bereich (z.B. ab Punkt B) gilt dieselbe Vorgehensweise mit umgekehrten Vorzeichen.

Die dargestellte Strategie kann in einen Signalfussplan für eine zeitoptimale Regelung umgesetzt werden. Dabei wird allerdings von einigen stark idealisierten Annahmen ausgegangen, die in der Praxis in der Regel nicht erfüllt sind, so dass das erste vorgestellte Konzept lediglich als prinzipielle Möglichkeit betrachtet werden muss. Es werden zwei nichtlineare Blöcke innerhalb der Regelung verwendet. Zunächst wird durch die erste Summationsstelle der Abstand der Istposition vom Ziel berechnet und auf die optimale Schaltkennlinie gegeben, die daraus die Geschwindigkeit berechnet, die bei diesem Abstand auf der Bremsparabel vorliegt. Diese wird mit einem Zweipunktglied mit der tatsächlich vorliegenden Geschwindigkeit verglichen. Das Vorzeichen der Differenz bestimmt das Vorzeichen der maximalen Stellgröße. Im Zielpunkt sollte die Geschwindigkeit genau

Null werden, so dass das verwendete Zweipunktglied mit Nullpunkt auch genau Null als Stellgröße ausgibt.

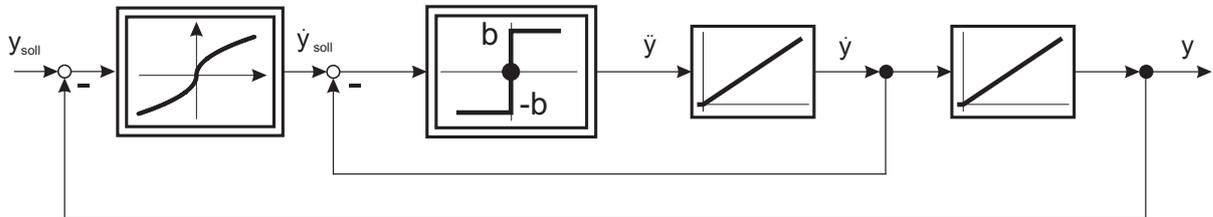


Bild 9.9: zeitoptimale Positionierung als Regelung

Dieser Ansatz vernachlässigt die Tatsachen, dass die Beschleunigung nicht geschaltet werden kann und dadurch immer Verzögerungen entstehen, so dass nicht auf der Schaltlinie geschaltet wird, sondern immer zu spät. Damit wird über den Zielpunkt hinaus geschossen und ein Grenzzyklus entsteht. Weiterhin ist natürlich ein ideales Zweipunktglied mit Nullpunkt technisch nicht darstellbar.

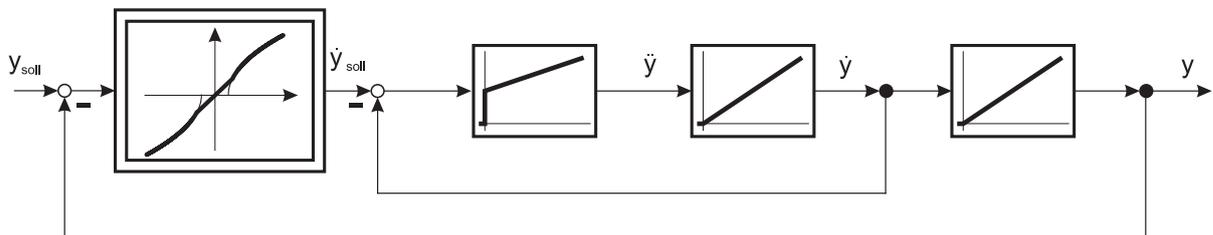


Bild 9.10: zeitoptimale Positionierung als Regelung (verbesserte Version)

Bild 9.10 zeigt eine verbesserte Version der zeitoptimalen Regelung. Aufgrund der Modellfehler (weitere Verzögerungen gegenüber reinem doppeltem Integrator) wird der einfache Ansatz nach Bild 9.9 nicht zuverlässig funktionieren. Es ergeben sich in der Praxis immer Schwingungen, weil im Zielpunkt sowohl die Bremsparabel als auch das Zweipunktglied unendliche Verstärkungen aufweisen. Daher wird in diesem Ansatz die Bremsparabel mit einem linearen P-Regler so erweitert, dass die eigentliche Bremsparabel auf einen Punkt zielt, der vor dem Zielpunkt liegt. Kurz vor Erreichen dieses Punktes findet ein stoßfreies Umschalten auf den linearen P-Regler statt, der die letzte Differenz zum Zielpunkt überbrückt. Die Geschwindigkeit wird in dieser Variante durch einen linearen PI-Regler eingestellt, der dann auch für zusätzliche Verzögerungen (hier nicht dargestellt) nach dem symmetrischen Optimum ausgelegt werden kann.

9.3 Zeitoptimale Vorsteuerung mit Führungsgrößengenerator

Das vorgestellte Verfahren wird üblicherweise nicht direkt innerhalb eines Regelkreises eingesetzt, sondern vielmehr an einer simulierten Strecke verwendet. Dann kann damit ein zusammengehöriger Satz von Bahngrößen, bestehend aus y , \hat{y} und $\hat{\dot{y}}$, eine *Trajektorie* erzeugt werden, die zur Vorsteuerung auf eine Lageregelungskaskade geschaltet wird. Dies ist in Bild 9.11 gezeigt. Die zusätzlichen Verzögerungen sind hier als PT1-Glied dargestellt. Es kann zur Verbesserung der Vorsteuerung in alle Pfade der Vorsteuerung außer der Beschleunigung verwendet werden.

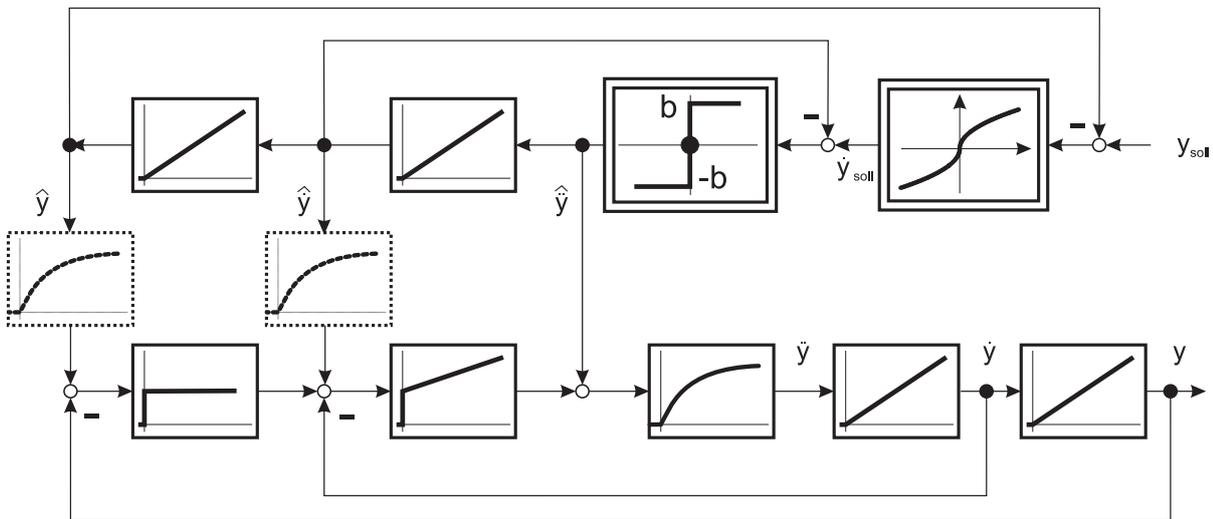


Bild 9.11: Lageregelung mit Führungsgrößengenerator

Da die Maschinen aus Stahl gebaut werden, die eine sehr geringe Eigendämpfung aufweist, kommt es bei Verwendung von Bahnprofilen 2. Ordnung für die Lage, wie bisher angenommen, immer zu relativ starken Anregungen von Maschinenschwingungen. Daher werden in der Praxis meist ruckbegrenzte Trajektorien verwendet. Als Ruck bezeichnet man die Ableitung der Beschleunigung. Dann ist die Beschleunigung als künstliche Zustandsgröße aufzufassen, die die Ordnung um eins erhöht. Die Berechnung einer ruckbegrenzten Trajektorie benötigt sehr viele Fallunterscheidungen und soll an dieser Stelle nicht im Detail angesprochen werden. Das Grundprinzip ist in Bild 9.12 gezeigt.

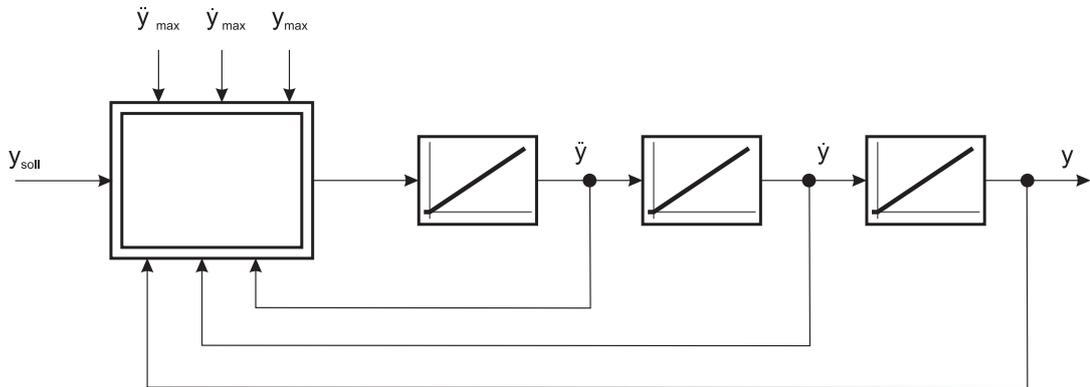


Bild 9.12: Führungsgrößengenerator für ruckbegrenzte Trajektorien

Die bei verschiedenen Fahrbewegungen entstehenden Zeitverläufe zeigt Bild 9.13.

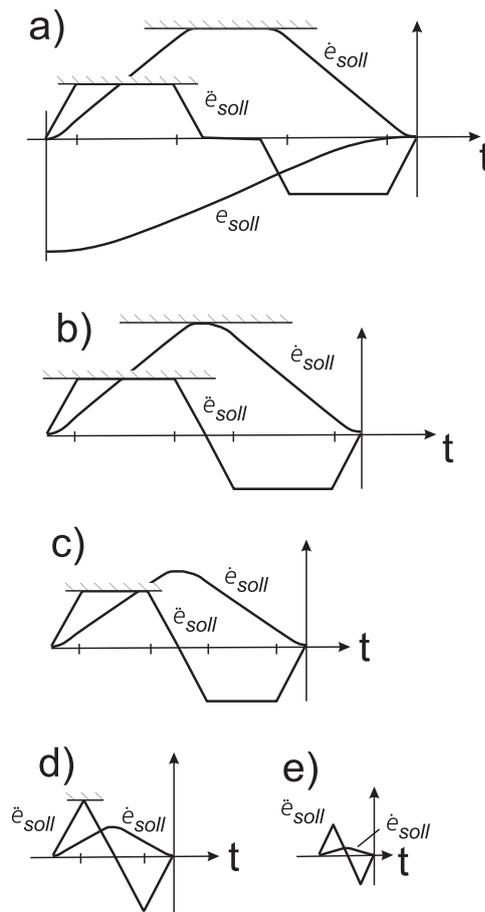


Bild 9.13: Trajektorien für ruckbegrenztes Verfahren

10 Beschreibungs-Funktion

Das Hilfsmittel der *Beschreibungs-Funktion* dient dazu, die Stabilität von nichtlinearen Systemen abzuschätzen. Diese Systeme können, abhängig von ihrem inneren Aufbau, stabil bzw. instabil sein, wie man es von linearen Systemen gewohnt ist, oder als neue Systemeigenschaft Grenzzyklen aufweisen. Voraussetzung für die Anwendung des Verfahrens der *Beschreibungs-Funktion* ist es, dass das nichtlineare System als Reihenschaltung zweier Systemteile darstellbar ist: Eine nichtlineare Kennlinie K , die keine Dynamik aufweist und augenblicklich wirkt und einen linearen Systemteil mit Tiefpasscharakter. Durch das Verfahren der *Beschreibungs-Funktion* wird die Kennlinie durch einen äquivalenten Verstärkungsfaktor für diese Grundschwingung ersetzt. Ausgegangen wird zunächst von einem System der Form in Bild 10.1.

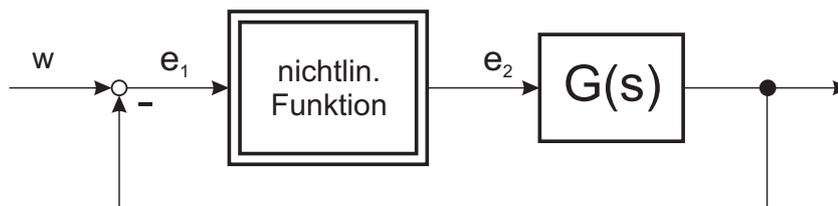


Bild 10.1: Nichtlineares System

Der Sollwert w wird für die Untersuchung auf 0 gesetzt. Für die Stabilitätsuntersuchung wird zunächst der Kreis aufgetrennt und an der Stelle e_1 mit einer reinen Sinus-Funktion angeregt: $e_1(t) := \hat{e}_1 \sin(\omega t)$.

Der Augenblickswert $e_1(t)$ wird durch die Kennlinie K auf einen Augenblickswert $e_2(t)$ abgebildet.

$$e_2(t) = K(e_1(t)) \quad (10.1)$$

Da angenommen wird, dass die lineare Strecke nach der konzentrierten nichtlinearen Kennlinie K starke Tiefpasseigenschaften aufweist, so reicht es aus, am Ausgang der nichtlinearen Strecke nur den Anteil der Grundschwingung zu betrachten. Man definiert anstelle der Kennlinie einen äquivalenten komplexen Verstärkungsfaktor, der die Übertragung der Grundschwingung nach Betrag und Phasenverschiebung beschreibt. Er wird *Beschreibungsfunktion* N genannt und ist definiert als Verhältnis der Grundschwingungsamplitude und -phase zwischen e_1 und e_2 : $N(\hat{e}_1) = \frac{e_2}{e_1}$. Die *Beschreibungsfunktion* ist aufgrund der Nichtlinearität abhängig von der Anregungsamplitude \hat{e}_1

Die *Beschreibungsfunktion* wird wie ein linearer Verstärkungsfaktor verwendet. Damit ergibt sich für die Übertragungsfunktion des geschlossenen Kreises:

$$G_g(s) = \frac{N(\hat{e}_1)G(s)}{1 + N(\hat{e}_1)G(s)} = \frac{G(s)}{\frac{1}{N(\hat{e}_1)} + G(s)} \quad (10.2)$$

Damit das Verfahren der *Beschreibungs-Funktion* angewendet werden kann, müssen vier Bedingungen erfüllt sein:

1. $G(s)$ muss Tiefpasscharakter haben. Der Betrag des Frequenzganges muss monoton fallend sein. Dies ist bei den meisten praktisch vorhandenen linearen Strecken gegeben.
2. Die Kennlinie der nichtlinearen Funktion muss punktsymmetrisch zum Ursprung sein und darf nur im 1. und 3. Quadranten verlaufen.
3. Die Kennlinie muss momentan wirkend sein, darf also keinerlei Dynamik enthalten.
4. Der Mittelwert des Eingangssignal des nichtlinearen Elements (hier e_1) muss Null sein, denn sonst wäre indirekt wiederum die zweite Voraussetzung verletzt. Diese Bedingung erfordert entweder $w = 0$ oder dass $G(s)$ einen Integrator-Anteil aufweist, der dazu führt, dass bei Stabilität des geschlossenen Kreises der Gleichanteil durch den Integrator kompensiert wird..

Nimmt man zur Analyse wiederum an, dass die offene Kette aus der Nichtlinearität und dem lineareren Tiefpass an der Stelle e_1 durch einen Sinus angeregt wird

$$e_1(t) = \hat{e}_1 \sin(\omega t), \quad (10.3)$$

so lässt sich $e_2(t)$ in eine Fourier-Reihe zerlegen:

$$e_2(t) = \sum_{\nu=1,3,\dots}^{\infty} \hat{a}_\nu \sin(\nu\omega t) + \hat{b}_\nu \cos(\nu\omega t) \quad (10.4)$$

Da die Kennlinie ungerade ist, entstehen auch nur ungerade Oberschwingungen. Dabei berechnen sich die Fourierkoeffizienten mit $T = \frac{2\pi}{\omega}$ zu:

$$\hat{a}_\nu = \frac{2}{T} \int_0^T e_2(t) \sin(\nu\omega t) dt \quad (10.5)$$

$$\hat{b}_\nu = \frac{2}{T} \int_0^T e_2(t) \cos(\nu\omega t) dt \quad (10.6)$$

Diese Darstellung kann man auch in Betrag und Phase, d. h. Polarkoordinaten, für die einzelnen Oberschwingungen umschreiben. Es gilt:

$$e_2(t) = \sum_{\nu=1,3,\dots}^{\infty} \hat{e}_{2\nu} \sin(\nu\omega t + \varphi_\nu) \quad (10.7)$$

$$\text{mit } \hat{e}_{2\nu} = \sqrt{a_\nu^2 + b_\nu^2} \quad \text{und} \quad \tan(\varphi_\nu) = \frac{b_\nu}{a_\nu}$$

Die hier gewählten Voraussetzungen für die Anwendung des Verfahrens erzwingen, dass $e_1(t)$ und damit wegen der ungeraden Kennlinie auch $e_2(t)$ keinen Gleichanteil haben kann. Somit wird $b_0 = 0$. Weiterhin wird nur noch die Grundschiwingung der Fourierzerlegung von $e_2(t)$ weiter betrachtet, da $G(s)$ Tiefpasscharakter aufweist und damit alle Oberschwingungen gegenüber der Grundschiwingung stark bedämpft werden. Es ergibt sich angenähert für $e_2(t)$:

$$e_2(t) \approx a_1 \sin(\omega t) + b_1 \cos(\omega t) \quad (10.8)$$

Darin sind a_1 und b_1 die Fourier-Koeffizienten der Grundschiwingung:

$$\hat{a}_1 = \frac{2}{T} \int_0^T e_2(t) \sin(\omega t) dt \quad (10.9)$$

$$\hat{b}_1 = \frac{2}{T} \int_0^T e_2(t) \cos(\omega t) dt \quad (10.10)$$

In der Darstellung nach Betrag und Phase ergibt sich:

$$e_2(t) \approx \hat{e}_{21} \sin(\omega t + \varphi_1) \quad (10.11)$$

Werden die Zeitfunktionen $e_1(t)$ und $e_2(t)$ mit Hilfe der komplexen Exponentialfunktion geschrieben, so lassen sie sich folgendermaßen darstellen:

$$e_1(t) = \frac{1}{2j} (\hat{e}_1 e^{j\omega t} - \hat{e}_1 e^{-j\omega t}) \quad (10.12)$$

$$e_2(t) = \frac{1}{2j} (\hat{e}_{21} e^{j(\omega t + \varphi_1)} - \hat{e}_{21} e^{-j(\omega t + \varphi_1)}) \quad (10.13)$$

Die Beschreibungsfunktion $N(\hat{e}_1)$ wird als komplexer Verstärkungsfaktor für das Amplitudenverhältnis und die Phasenverschiebung der Grundschiwingung definiert. Sie ist abhängig von der Amplitude \hat{e}_1 aber nicht von der Frequenz ω . $N(\hat{e}_1)$ berechnet sich als Verhältnis der komplexen Zeiger $\hat{e}_2 e^{j(\omega t + \varphi_1)}$ und $\hat{e}_1 e^{j\omega t}$:

$$\begin{aligned} N(\hat{e}_1) &= \frac{\hat{e}_{21} e^{j(\omega t + \varphi_1)}}{\hat{e}_1 e^{j\omega t}} = \frac{\hat{e}_{21}}{\hat{e}_1} e^{j\varphi_1} \\ &= \frac{\hat{e}_{21}}{\hat{e}_1} \cos(\varphi_1) + j \frac{\hat{e}_{21}}{\hat{e}_1} \sin(\varphi_1) \end{aligned} \quad (10.14)$$

$$= \frac{\hat{a}_1}{\hat{e}_1} + j \frac{\hat{b}_1}{\hat{e}_1} = a_1 + j b_1 \quad (10.15)$$

$$\text{da } \cos(\varphi_1) = \frac{a_1}{\hat{e}_2} \quad \text{und} \quad \sin(\varphi_1) = \frac{b_1}{\hat{e}_2}$$

Falls die Kennlinie der nichtlinearen Funktion eindeutig ist und keine Hysterese aufweist, so ist $b_1 = 0$ und damit $N(\hat{e}_1)$ reell.

Setzt man die Kennlinie $K(e_1)$ in die Berechnung für die Fourierkoeffizienten der Beschreibungsfunktion ein, berechnen diese sich zu:

$$a_1 = \frac{2}{T\hat{e}_1} \int_0^T K(\hat{e}_1 \sin(\omega t)) \sin(\omega t) dt \quad (10.16)$$

$$b_1 = \frac{2}{T\hat{e}_1} \int_0^T K(\hat{e}_1 \sin(\omega t)) \cos(\omega t) dt \quad (10.17)$$

Mit der Substitution $\omega t = \xi$ vereinfacht sich die Gleichung zu:

$$a_1 = \frac{1}{\pi\hat{e}_1} \int_0^{2\pi} K(\hat{e}_1 \sin(\xi)) \sin(\xi) d\xi \quad (10.18)$$

$$b_1 = \frac{1}{\pi\hat{e}_1} \int_0^{2\pi} K(\hat{e}_1 \sin(\xi)) \cos(\xi) d\xi \quad (10.19)$$

10.1 Begrenzer

Betrachten wir das Beispiel eines Begrenzers für das nichtlineare Element in Bild 10.1. Dessen Kennlinie, sowie die Ein- und Ausgangskennlinien sind im folgenden Bild dargestellt. Solange die Amplitude die Begrenzungsschwelle nicht erreicht, ergibt sich eine lineare Verstärkung von $V = \frac{e_{20}}{e_{10}}$.

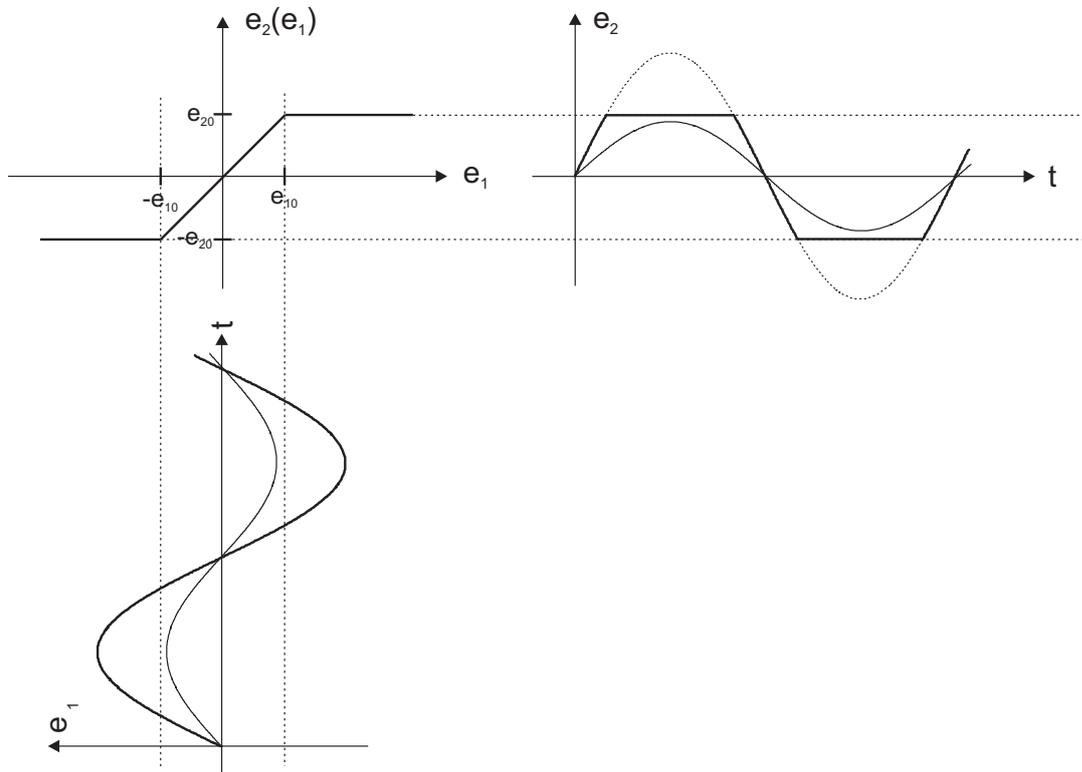


Bild 10.2: Ein- und Ausgangssignale bei der Begrenzerkennlinie

Bei der Berechnung der Beschreibungs-Funktion N müssen aus Symmetriegründen nur zwei Bereiche betrachtet werden. Da die Kennlinie eindeutig ist und keine Hysterese aufweist, gilt auf jeden Fall $b_1 = 0$.

- Bereich 1: $\hat{e}_1 < e_{10}$

$$K(e_1) = \frac{e_{20}}{e_{10}} e_1 = V \hat{e}_1 \sin(\omega t) \quad (10.20)$$

$$\Rightarrow a_1 = \frac{2}{T} \int_0^T V \hat{e}_1 \sin(\omega t) \sin(\omega t) dt \quad \text{Subst.: } \omega t = \xi \quad (10.21)$$

$$\Rightarrow a_1 = \frac{1}{\pi} \int_0^{2\pi} V \hat{e}_1 \sin^2(\xi) d\xi \quad (10.22)$$

$$\text{mit } \int \sin^2(\xi) d\xi = \frac{1}{2}\xi - \frac{1}{4}\sin(2\xi)$$

$$\Rightarrow a_1 = V \hat{e}_1 \quad (10.23)$$

$$\Rightarrow N(\hat{e}_1) = \frac{a_1}{\hat{e}_1} = V \quad (10.24)$$

- Bereich 2: $\hat{e}_1 > e_{10}$

Hier wird der Zeitverlauf von $e_2(t)$ in 6 Abschnitte aufgeteilt, in denen jeweils eine geschlossene Formel für die Teilintegrale des Fourierkoeffizienten aufgestellt werden kann:

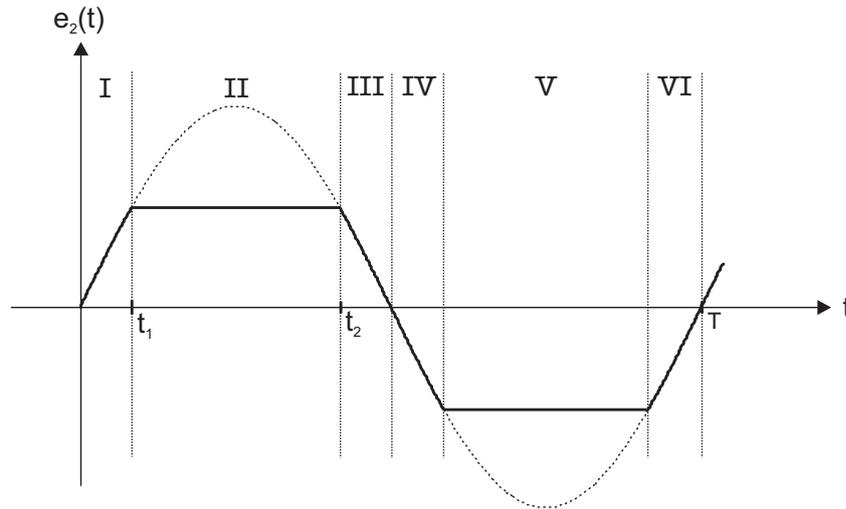


Bild 10.3: Kennlinie für $e_1 > e_{10}$

In den Bereichen I, III, IV und VI gilt:

$$K(e_1(t)) = V\hat{e}_1 \sin(\omega t) \quad (10.25)$$

In II und V befindet sich die Kennlinie in der Berggrenzung und $e_2(t)$ ist konstant:

$$K(e_1(t)) = e_{20} \quad (10.26)$$

Damit ergibt sich für den Fourierkoeffizienten folgende Gleichung:

$$a_1 = \frac{2}{T} \left[4 \int_0^{t_1} V\hat{e}_1 \sin^2(\omega t) dt + 2 \int_{t_1}^{\frac{T}{2}-t_1} e_{20} \sin(\omega t) dt \right] \quad (10.27)$$

Mit der Substitution $\omega t = \xi$ vereinfacht:

$$a_1 = \frac{1}{\pi} \left[4 \int_0^{\xi_1} V\hat{e}_1 \sin^2(\xi) d\xi + 2 \int_{\xi_1}^{\pi-\xi_1} e_{20} \sin(\xi) d\xi \right] \quad (10.28)$$

Dabei ist $\xi_1 = \arcsin\left(\frac{e_{10}}{\hat{e}_1}\right)$.

Als Endergebnis für $N(\hat{e}_1)$ ergibt sich nach Auswertung der Integrale:

$$N(\hat{e}_1) = \frac{2}{\pi} \left[\frac{e_{20}}{e_{10}} \arcsin \frac{e_{10}}{\hat{e}_1} + \frac{e_{20}}{\hat{e}_1} \cos \arcsin \frac{e_{10}}{\hat{e}_1} \right] \quad (10.29)$$

An der Grenze zwischen Bereich 1 und 2 ist ein stetiger Übergang der beiden Gleichungen. Beide liefern den Wert $N(\hat{e}_1) = V$.

Somit ergibt sich für die Beschreibungsfunktion über beide Bereiche:

$$N(\hat{e}_1) = \begin{cases} V = \frac{e_{20}}{e_{10}} & \text{für } \hat{e}_1 \leq e_{10}, \\ \frac{2}{\pi} \left[\frac{e_{20}}{e_{10}} \arcsin \frac{e_{10}}{\hat{e}_1} + \frac{e_{20}}{\hat{e}_1} \cos \arcsin \frac{e_{10}}{\hat{e}_1} \right] & \hat{e}_1 \geq e_{10} \end{cases} \quad (10.30)$$

Bild 10.4 zeigt den Verlauf in grafischer Form. Zum Vergleich ist der Verlauf der Beschreibungsfunktion des idealen Zweipunktgliedes gestrichelt dargestellt.

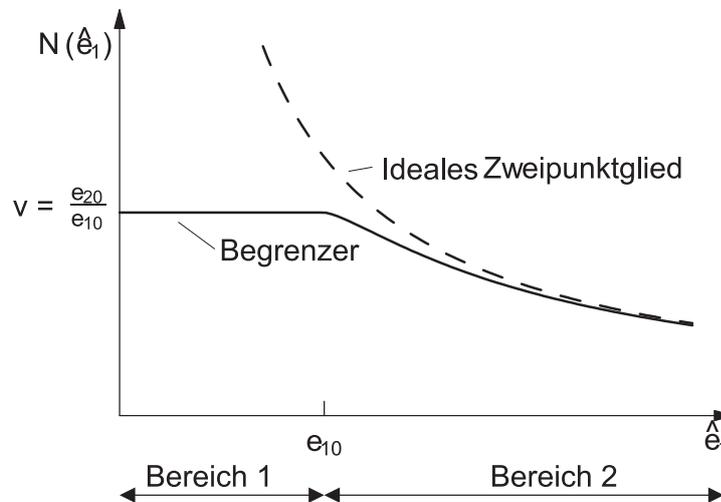


Bild 10.4: Verlauf von $N(\hat{e}_1)$ für die Begrenzerfunktion und das Zweipunktglied

10.2 Zweipunktglied

Das Zweipunktglied kann als Grenzfall des Begrenzers bei sehr kleinem Linearverstärkungsbereich $e_{10} \rightarrow 0$ bei konstanter Ausgangsamplitude $e_{20} = const$ aufgefasst werden. Mit der weiterhin eindeutigen Kennlinie bleibt die Beschreibungsfunktion reell und damit $b_1 = 0$:

$$\begin{aligned} N(\hat{e}_1) &= a_1 = \frac{2}{T} \int_0^T e_2(t) \sin(\omega t) dt \\ &= \frac{1}{\pi \hat{e}_1} \left[\int_0^\pi e_{20} \sin(\xi) d\xi + \int_\pi^{2\pi} -e_{20} \sin(\xi) d\xi \right] = \frac{4e_{20}}{\pi \hat{e}_1} \end{aligned} \quad (10.31)$$

Der grafische Verlauf ist bereits in Bild 10.4 als einfache Hyperbel gestrichelt dargestellt.

10.3 Grafische Auswertung

Für die grafische Interpretation der Stabilitätsuntersuchung kommen wir auf den Ausdruck (10.2) vom Anfang des Kapitels zurück. Gibt es ein Paar s und \hat{e}_1 , für das die Gleichung $\frac{1}{N(\hat{e}_1)} + G(s) = 0$ gilt, so ist das untersuchte System grenzstabil. Dieser sogenannte Gleichgewichtspunkt, auch *Harmonische Balance* genannt, lässt sich grafisch ermitteln, indem die beiden komplexen Funktionen $-\frac{1}{N(\hat{e}_1)}$ und $G(s)$ in der komplexen Ebene in einem Diagramm aufgetragen werden. In Bild 10.5 ist dies für den Begrenzer und zwei IT_2 -Glieder mit unterschiedlicher Verstärkung dargestellt.

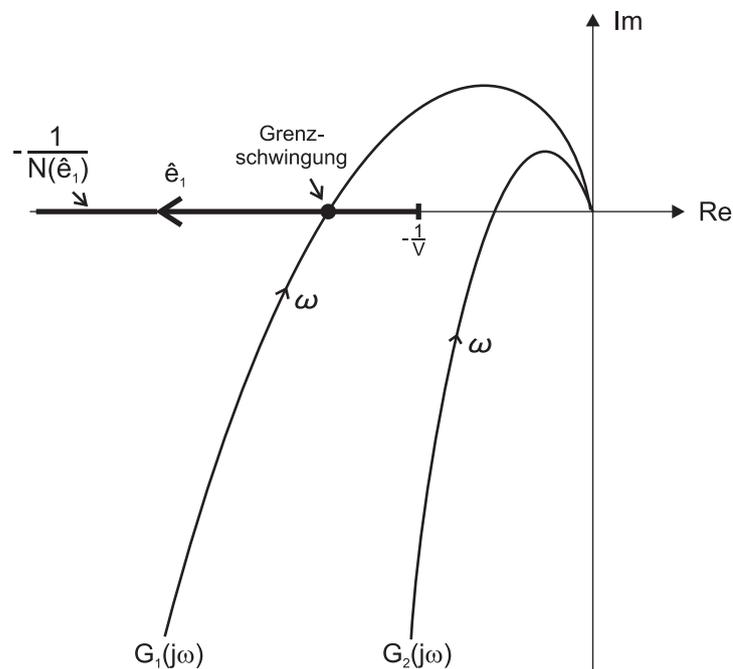


Bild 10.5: G - und $-\frac{1}{N}$ -Ebene

Die $G(s)$ sind von folgender Form:

$$G(s) = \frac{V}{T_1 s (T_2 s + 1) (T_3 s + 1)} \quad (10.32)$$

Bei $G_1(s)$ ist V größer als bei $G_2(s)$. Für G_2 gibt es keine Grenzschiwingung, da sich die beiden Ortskurven nicht schneiden. D.h., dass die Strecke für alle ω und \hat{e}_1 stabil ist. Die Ortskurve von G_1 wiederum hat einen Schnittpunkt mit der Ortskurve von $-\frac{1}{N}$. Für die an diesem Punkt geltenden Variablen ω_s und \hat{e}_s kommt es im Regelkreis zu einer Dauerschwingung, auch *Harmonische Balance* genannt. Für alle Punkte, die links des Schnittpunktes liegen, ist das System stabil. Der Bereich der $-\frac{1}{N}$ -Linie, der sich

rechts vom Punkt der Grenzwinkel befindet, wird von der Ortskurve des linearen Streckenabschnitts umschlossen, und ist somit instabil (vgl. Nyquist-Kriterium). Liegt ein Startzustand in diesem Abschnitt, so klingen die Amplituden solange auf, bis sie den Zustand der Grenzwinkel erreicht haben.

Bei einem Zweipunktglied (Beschreibungs-Funktion siehe Bild 10.4) und einem doppelten Integrator als Strecke ergibt sich Bild 10.6.

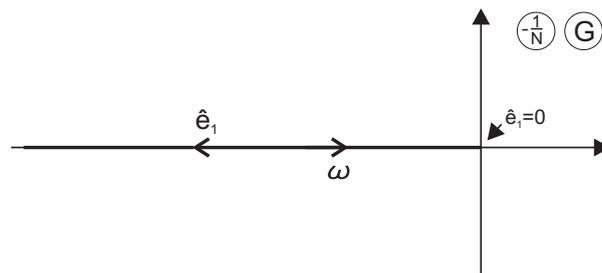


Bild 10.6: Zweipunktglied und I_2

Beide Kennlinien verlaufen auf der reellen Achse. Es kommt für jedes beliebige Variablenpaar ω_s und \hat{e}_s zur Dauerschwingung. Für hohe Frequenzen ergeben sich dabei kleine Amplituden. Dieses Ergebnis wurde auch schon vorher durch die Analyse im Zustandsraum erzielt.

10.3.1 Begrenzerkennlinie mit Tot-Bereich

Die bereits untersuchte Begrenzerkennlinie lässt sich erweitern, indem ein Totbereich vorgesehen wird. Er gilt für eine Amplitude von a für das Eingangssignal. Der Begrenzeinsatz ist wiederum als e_{10} definiert. Mit dieser Parametrierung können die Ergebnisse auch für das Dreipunktglied ohne Hysterese verwendet werden. In diesem Fall muss nur $a = e_{10}$ gesetzt werden. Die Kennlinie und die Zeitverläufe des Ausgangssignals sind in Bild 10.7 dargestellt. Die Beschreibungsfunktion ist aufgrund der eindeutigen Kennlinie wiederum reell und muss in zwei Abschnitten integriert werden, wobei aus Symmetriegründen nur eine viertel Periode integriert wird.

$$\begin{aligned}
 N(\hat{e}_1) &= a_1 = \frac{8}{T\hat{e}_1} \int_0^{\frac{T}{4}} e_2(t) \sin(\omega t) dt \\
 &= \frac{4}{\pi\hat{e}_1} \frac{e_{20}}{e_{10} - a} \int_{\xi_1}^{\xi_2} [(\hat{e}_1 \sin \xi - a) \sin(\xi) d\xi + \int_{\xi_2}^{\frac{\pi}{2}} -e_{20} \sin(\xi) d\xi \quad (10.33)
 \end{aligned}$$

$$\text{mit } \xi_1 = \arcsin\left(\frac{a}{\hat{e}_1}\right), \quad \xi_2 = \arcsin\left(\frac{e_{10}}{\hat{e}_1}\right) \quad (10.34)$$

Für kleinere Eingangsamplituden müssen noch einige Fälle unterschieden werden. Für $\hat{e}_1 < a$ gilt $N(\hat{e}_1) = 0$, für $a < \hat{e}_1 < e_{10}$ ist die Obergrenze des ersten Integrals $\frac{\pi}{2}$ und das zweite entfällt.

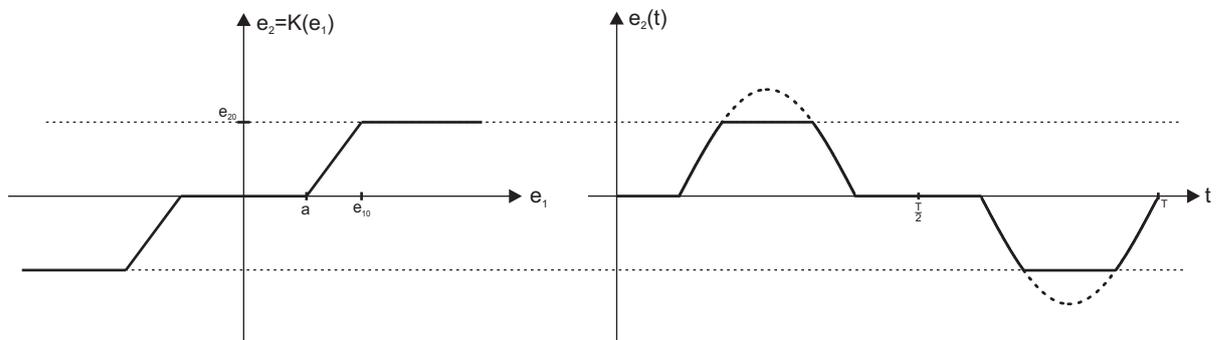


Bild 10.7: Begrenzerkennlinie mit Tot-Bereich und zugehöriges Ausgangssignal

$$N(\hat{e}_1) = \frac{e_{20}}{(e_{10} - a)\pi} [2\xi_2 - 2\xi_1 + \sin 2\xi_1 - \sin 2\xi_2] + \frac{4 a e_{20}}{\pi \hat{e}_1 (e_{10} - a)} [\cos \xi_2 - \cos \xi_1] + \frac{4 e_{20}}{\pi \hat{e}_1} \cos \xi_2 \quad (10.35)$$

Die zugehörige Beschreibungs-Funktion ist in Bild 10.8 zusammen mit den Verläufen der bereits früher berechneten Elemente (Begrenzer, ideales Zweipunktglied) dargestellt. Wählt man $a \rightarrow e_{10}$ so erhält man im Grenzfall die Beschreibungs-funktion des idealen Dreipunktgliedes, die ebenfalls mit dargestellt ist.

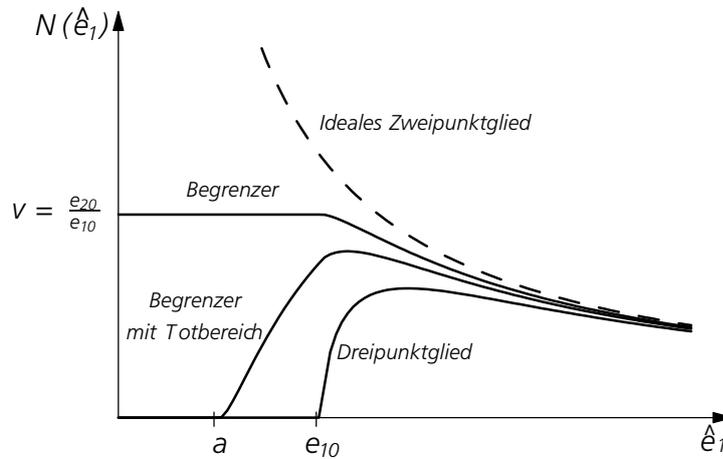


Bild 10.8: Beschreibungs-Funktion des Begrenzers mit Tot-Bereich

Die Ortskurven ergeben folgendes Bild (Strecke ist IT_2):

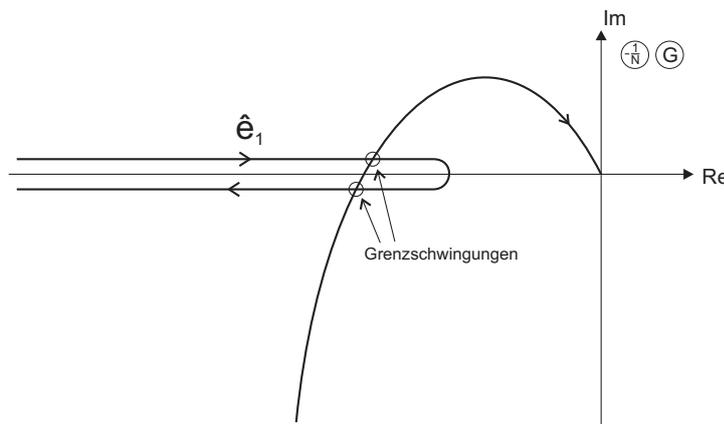


Bild 10.9: Ortskurven für obiges Beispiel

Hierbei ist das System für Werte, die zwischen den beiden Grenzschnittpunkten liegen, instabil und stabil für die restlichen Abschnitte.

10.3.2 Zweipunktglied mit Hysterese

Hysteresebehaftete Elemente weisen eine Besonderheit auf: Die Grundschnittpunkte in e_2 ist gegenüber der sinusförmigen Anregung e_1 phasenverschoben, weil eine gewisse Amplitude überwunden werden muss, bevor der Ausgang reagiert. Die Phasenverschiebung

entsteht rein aufgrund von Amplitudenverhältnissen und ist nicht frequenzabhängig. Damit kommt eine weitere einschränkende Bedingung zum Tragen: Bei hysteresebehafteten Elementen muss als Ergebnis der Analyse immer eine stabile Grenzwinkelung entstehen, von der dann Amplitude und Frequenz abgeschätzt werden kann oder Instabilität nachgewiesen werden. Entsteht keine Grenzwinkelung, die die Hysterese symmetrisch durchläuft, so kann mit dem Verfahren der Beschreibungsfunktion zunächst keine Aussage gemacht werden und es muss überprüft werden, ob weitere Voraussetzungen zur Anwendung des Verfahrens verletzt sind. Es sei folgendes Element gegeben:

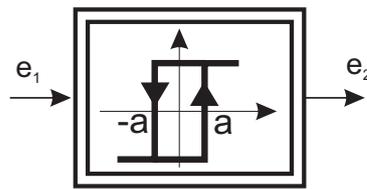


Bild 10.10: Zweipunktglied mit Hysterese

Bei Anregung mit einer Sinusfunktion ergeben sich folgende Zeitverläufe des Eingangs- und Ausgangssignals (Bild 10.11).

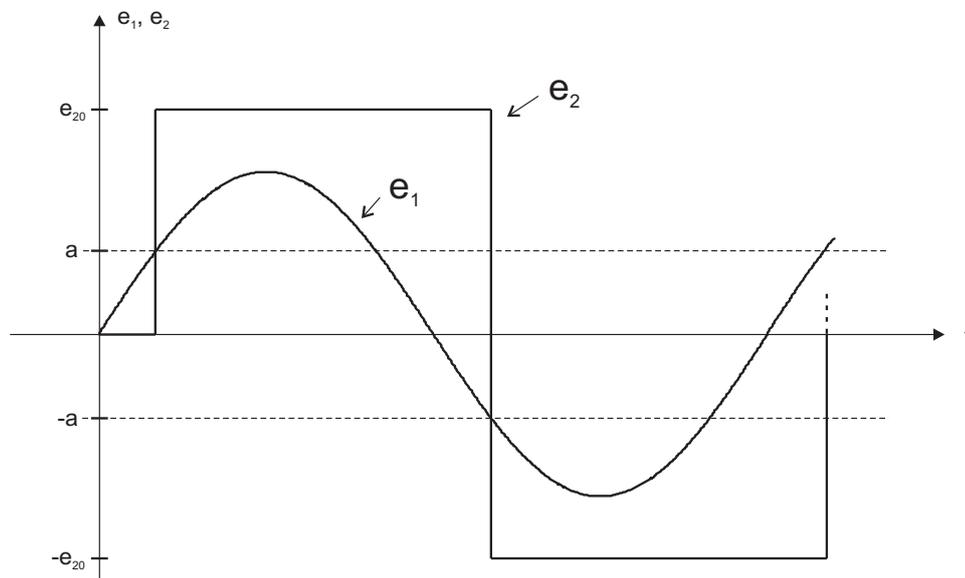


Bild 10.11: Eingangs- und Ausgangssignale beim Zweipunktglied mit Hysterese

Es ist deutlich erkennbar, dass das Ausgangssignal gegenüber dem Eingangssignal phasenverschoben ist. Diese Verschiebung wird für $\hat{e}_1 = a$ maximal, unmittelbar vor dem

Aussetzen des zyklischen Durchlaufs der Hysterese schleife. Das Ausgangssignal ist in seiner Grundschiwingung dann genau 90° ($\frac{T}{4}$) zur Anregung phasenverschoben.

Die Beschreibungs-Funktion lautet:

$$N(\hat{e}_1) = \frac{4e_{20}}{\pi\hat{e}_1} e^{-j \arcsin(\frac{a}{\hat{e}_1})} \quad (10.36)$$

Um die Ortskurve zeichnen zu können, wird dieser Ausdruck nochmals umgeformt:

$$\begin{aligned} -\frac{1}{N(\hat{e}_1)} &= -\frac{\hat{e}_1\pi}{4e_{20}} e^{j \arcsin(\frac{a}{\hat{e}_1})} \\ &= -\frac{\hat{e}_1\pi}{4e_{20}} \left[\cos(\arcsin(\frac{a}{\hat{e}_1})) + j \sin(\arcsin(\frac{a}{\hat{e}_1})) \right] \\ &= -\frac{\hat{e}_1\pi}{4e_{20}} \cos(\arcsin(\frac{a}{\hat{e}_1})) - j \frac{a\pi}{4e_{20}} \end{aligned} \quad (10.37)$$

Der Imaginärteil ist also unabhängig von \hat{e}_1 und damit konstant. Im folgenden Bild ist die Ortskurve dargestellt.

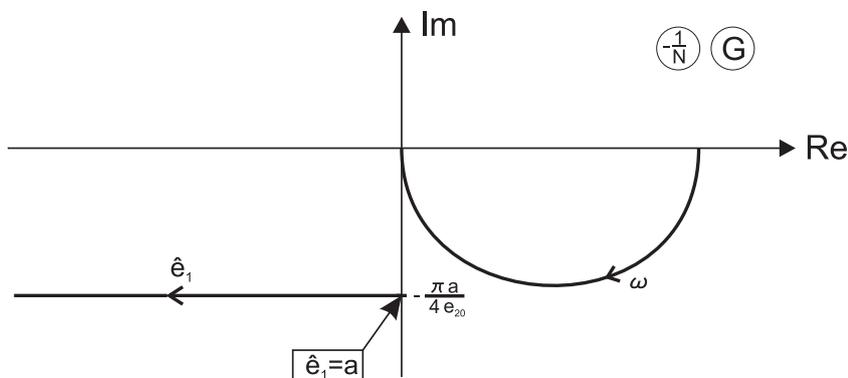


Bild 10.12: Zweipunktglied mit Hysterese und PT_1

In Bild 10.12 ist die Ortskurve der Strecke, eines PT_1 -Glieds, eingetragen. Sie hat keinen Schnittpunkt mit der Ortskurve des Zweipunktgliedes mit Hysterese, die ebenfalls eingetragen ist. Daraus könnte man vorschnell auf eine Konvergenz und damit einen stabilen Endzustand schließen. Dabei wird aber vernachlässigt, dass bei einem Hysteresebehafteten Element die Beschreibungs-funktion nur für Amplituden \hat{e}_1 definiert ist, die die Hysterese überwinden. Somit ist in diesem Fall mit der Methode der Beschreibungs-funktion keine Aussage über die Stabilität möglich, da sich die Definitionsbereiche der Ortskurven nicht überschneiden.

10.3.3 Dreipunktglied mit Hysterese

Das Dreipunktglied mit Hysterese weist zwei Hystereseschwellen auf. Es wird meist in Verbindung mit einem nachgeschalteten Integrator, z. B. ein Stellmotor, der auf Vorwärts- oder Rückwärtslauf angesteuert wird. Eine andere typische Anwendung ist bei Temperaturregelstrecken mit Heizen / Kühlen.

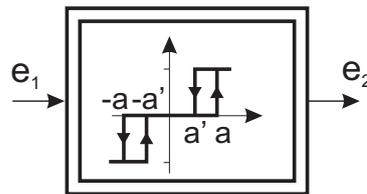


Bild 10.13: Dreipunktglied mit Hysterese

Beim Dreipunktglied sehen Zeitverläufe der Eingangs- und Ausgangssignale bei sinusförmiger Anregung folgendermaßen aus:

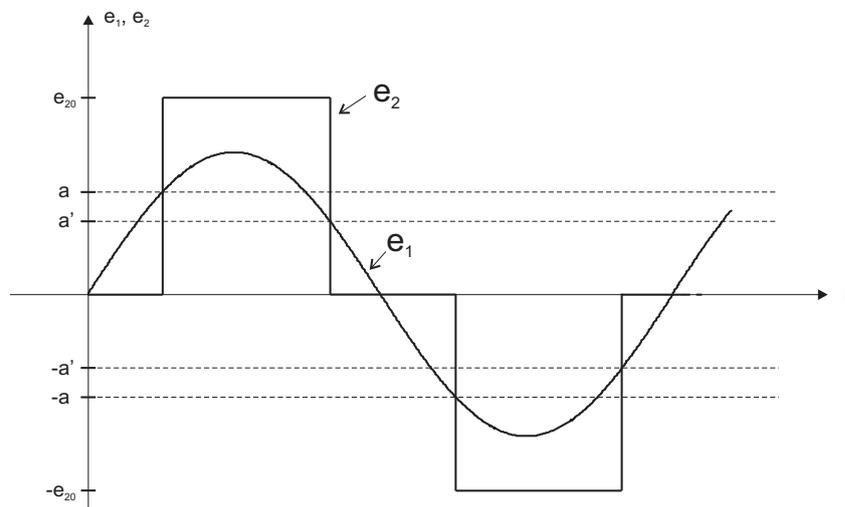


Bild 10.14: Eingangs- und Ausgangssignale beim Dreipunktglied

Als Ortskurvendarstellung ergibt sich Bild 10.15.

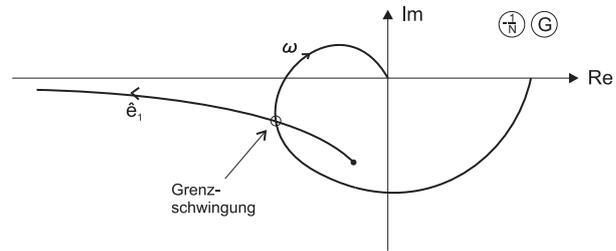


Bild 10.15: Dreipunktglied mit Hysterese und PT_3

Wie schon in einem der vorherigen Beispiele gibt es auch hier Bereiche der Stabilität und der Instabilität. Der gestrichelt dargestellte Abschnitt ist in diesem Fall instabil.

11 Stabilitätsuntersuchung nach Ljapunov

Die Stabilitätsanalyse nach Ljapunov geht von einem nichtlinearen dynamischen System mit konzentrierten Zustandsgrößen aus, wie es in Bild 11.1 gezeigt ist.

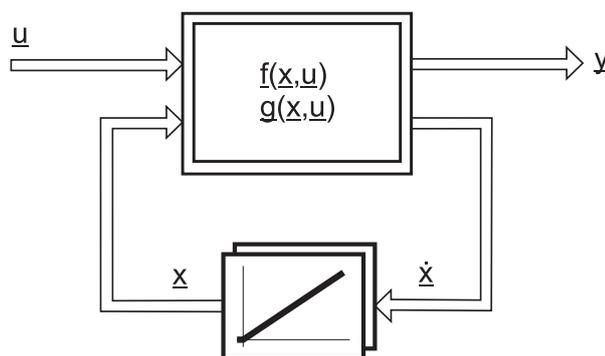


Bild 11.1: nichtlineares System

Es wird beschrieben durch die nichtlinearen Vektorfunktionen $\underline{f}(\underline{x}, \underline{u})$ und $\underline{g}(\underline{x}, \underline{u})$, die das Eingangssignal \underline{u} mit dem inneren Systemzustand \underline{x} und dem Ausgangssignal \underline{y} verknüpfen:

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}) \quad (11.1)$$

$$\underline{y} = \underline{g}(\underline{x}, \underline{u}) \quad (11.2)$$

In einer *Ruhelage* $\underline{u}_R, \underline{x}_R$ hat die Funktion $\underline{f}(\underline{x}, \underline{u})$ eine Nullstelle, so dass die Ableitung des Systemzustandes $\dot{\underline{x}}$ verschwindet. Sie ist folgendermaßen definiert:

$$\underline{f}(\underline{x}_R, \underline{u}_R) = \underline{0} \quad (11.3)$$

Jeder Ruhelage ist ein Paar aus Systemzustand \underline{x}_R und Eingangsvektor \underline{u}_R zugeordnet, für die die Gleichgewichtsbedingung erfüllt ist. Die Menge der Ruhelagen eines nichtlinearen Systems kann geschlossene Bereiche bilden oder diskrete Punkte umfassen. Es kann auch vorkommen, dass Bahnen oder Teilräume des Zustandsraumes existieren, in denen die Gleichgewichtsbedingung Gl. 11.3 erfüllt ist. Der Begriff Gleichgewicht oder Ruhelage darf nicht mit Stabilität verwechselt werden, denn es existieren sowohl stabile wie auch instabile Ruhelagen, wie im Folgenden gezeigt wird.

11.1 Beispiel zu Ruhelagen

Als Beispiel zur Erläuterung der Ruhelagen wird ein synthetisches mathematisches Beispiel erster Ordnung in Gleichung 11.4 betrachtet. Ein Integrator besitzt eine lineare Gegenkopplung und außerdem eine Mitkopplung mit einer Funktion y^3 .

$$\frac{dy}{dt} - k \cdot y^3 + y = u \quad (11.4)$$

Diese Gleichung ist als Signalflussplan in Bild 11.2 dargestellt.

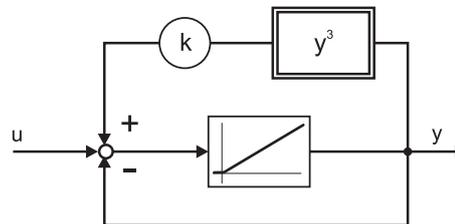


Bild 11.2: System mit Nichtlinearität in der Rückführung

Die Menge der Gleichgewichtspunkte (u_0, y_0) wird beschrieben durch Nullsetzen der Ableitung in der Differentialgleichung bestimmt.

$$u_0 = -k \cdot y_0^3 + y_0 \quad \text{mit} \quad \frac{dy}{dt} \stackrel{!}{=} 0 \quad (11.5)$$

Die im Gleichgewichtspunkt linearisierte Differentialgleichung lautet:

$$\frac{d}{dy}(-k \cdot y^3 + y) \Big|_{y_0} = -3k y_0^2 + 1. \quad (11.6)$$

Die Stabilität oder Instabilität im Gleichgewichtspunkt wird durch die Stabilität der linearisierten Differentialgleichung im betreffenden Arbeitspunkt entschieden.

$$-3k y^2 + 1 \geq 0 \quad (11.7)$$

$$y_0^2 \leq \frac{1}{3}k \quad (11.8)$$

$$|y_0| \leq \sqrt{\frac{k}{3}} \quad (11.9)$$

Einsetzen des Gleichgewichtspunktes führt auf die Menge der stabilen Gleichgewichtspunkte:

$$|u_0| \leq k \cdot \sqrt{\frac{k}{3}}^3 + \sqrt{\frac{k}{3}} \quad (11.10)$$

$$|u_0| \leq k \cdot k^{3/2} \cdot \frac{1}{3^{3/2}} + \sqrt{\frac{k}{3}} \quad (11.11)$$

Die Verhältnisse sind in der folgenden Grafik darstellt.

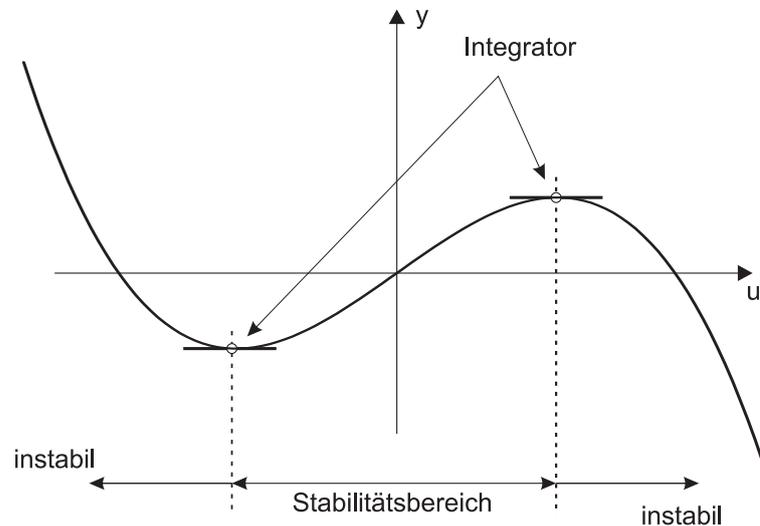


Bild 11.3: Übertragungsfunktion zeigt stabile und instabile Punkte

Betrachten wir zuerst die Stabilität innerhalb kleiner Bereiche um die Ruhelage. Dazu wird der Systemzustand minimal variiert: $\underline{x} = \underline{x}_R + \Delta\underline{x}$. Der Ansteuervektor \underline{u}_R bleibe konstant. Die Ruhelage wird als stabil bezeichnet, wenn der Wert der Ableitung $\dot{\underline{x}} = \underline{f}(\underline{x} + \Delta\underline{x}, \underline{u}_R)$ der Bewegung $\Delta\underline{x}$ entgegen gerichtet ist:

$$\underline{f}(\underline{x}_R + \Delta\underline{x}, \underline{u}_R) \rightarrow \underline{0} \quad \text{für} \quad \Delta\underline{x} \rightarrow \underline{0} \quad (11.12)$$

Für kleine Änderungen des Systemzustands läuft das System immer wieder in die Ruhelage zurück. Das im Punkt der Ruhelage linearisierte System ist stabil. Dies führt direkt auf die schwächste von drei Definitionen für die Stabilität von Ruhelagen:

- Schwächste Stabilitätsbedingung:

Eine Ruhelage heißt *stabil*, wenn zu jeder Auslenkung $|\underline{x}_0 - \underline{x}_R| < \epsilon$ ein δ existiert, so dass $|\underline{x}(t) - \underline{x}_R| < \delta$ gilt für alle $t > t_0$

- Schärfere Stabilitätsbedingung:

Eine Ruhelage heißt *asymptotisch stabil*, wenn sie *stabil* ist und eine Umgebung γ um die Ruhelage \underline{x}_R existiert, so dass aus $|\underline{x}_0 - \underline{x}_R| < \gamma$ folgt, dass der Systemzustand $\underline{x}(t)$ gegen \underline{x}_R konvergiert.

- Schärfste Stabilitätsbedingung:

Eine *asymptotisch stabile* Ruhelage heißt *global asymptotisch stabil*, wenn besagte Umgebung γ unendlich groß ist. Damit ist der gesamte Zustandsraum Einzugsbereich der Ruhelage.

11.2 Beispiel I: Mathematisches Pendel

Eine Punktmasse ist an einem Fixpunkt über eine feste Länge l drehbar gelagert. Hier werden nur ebene Bewegungen betrachtet: Die Auslenkung kann also durch einen Winkel beschrieben werden.

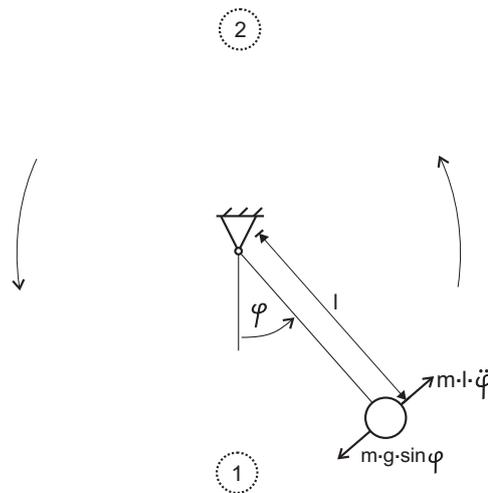


Bild 11.4: Mathematisches Pendel

Das mathematische Pendel besitzt keine Anregung und es ist keine Reibung modelliert. Somit verzehrt es während der Bewegung keine Energie und schwingt einmal angeregt unendlich lange nach. Daher sind die Trajektorien des Systemzustandes ungedämpfte Schwingungen, die auch für $t \rightarrow \infty$ nicht abklingen.

Es gibt in dem System nach Bild 11.4 zwei diskrete Ruhelagen. Die erste befindet sich bei $\varphi = 0$, die zweite bei $\varphi = \pi$. Hierauf lässt sich die o.g. schwächste Stabilitätsbedingung anwenden. Danach ist Ruhelage 1 *stabil*, da das Pendel für kleine Auslenkungen

niemals einen eingeschränkten Bereich verlässt. Bei Ruhelage 2 jedoch reicht eine kleine Abweichung, um große Auslenkungen, die außerhalb der unmittelbaren Umgebung der Ruhelage liegen, hervorzurufen. Die Ruhelage 2 ist damit *instabil*.

Nach Bild 11.4 gilt folgendes Momentengleichgewicht (Es wird von einem System ohne Anregung ausgegangen):

$$\sum_{\text{Momente}} = 0 \quad \Rightarrow \quad ml^2\ddot{\varphi} + mgl \sin(\varphi) = 0 \quad (11.13)$$

$$\Leftrightarrow \quad \ddot{\varphi} + \frac{g}{l} \sin(\varphi) = 0$$

$$\Leftrightarrow \quad \ddot{\varphi} = -\frac{g}{l} \sin(\varphi) \quad (11.14)$$

Hierbei sind φ und $\dot{\varphi}$ die Zustandsgrößen des Systems und es ergibt sich die Zustandsgleichung:

$$\dot{\underline{x}} = \begin{bmatrix} \dot{\varphi} \\ \ddot{\varphi} \end{bmatrix} = \begin{bmatrix} \dot{\varphi} \\ -\frac{g}{l} \sin(\varphi) \end{bmatrix} \quad (11.15)$$

Die Zustandsdarstellung zeigt Bild 11.5

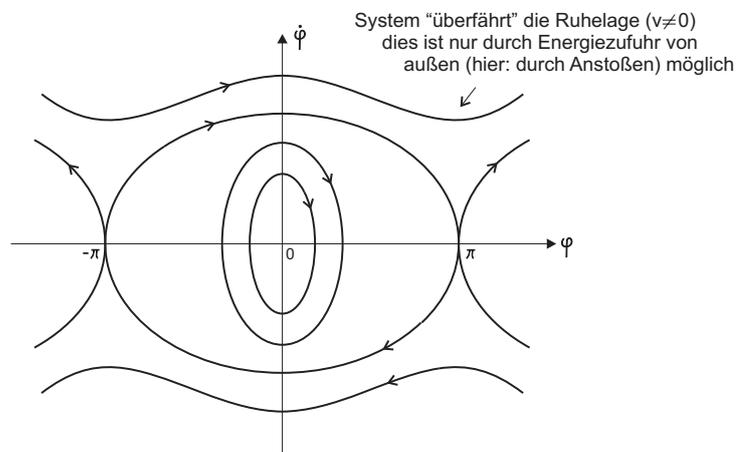


Bild 11.5: Zustandskurven beim mathematischen Pendel

11.3 Ljapunov-Stabilität

Eine mathematische Methode zur Stabilitätskontrolle von Ruhelagen nichtlinearer Systeme ist die Untersuchung nach *Ljapunov*. Sie ist durch eine Energiebetrachtung motiviert: Nimmt die Gesamtenergie eines dynamischen Systems kontinuierlich ab, so wird

der Systemzustand gegen eine Ruhelage konvergieren und diese Ruhelage ist eine stabile Ruhelage. Als allgemeinere Betrachtung kann mathematisch eine skalare Potenzialfunktion $V(\underline{x})$ definiert werden. Damit sie gleiche Eigenschaften wie die Energiebetrachtung aufweist, muss gelten:

1. $V(\underline{x}) > 0$ für alle $\underline{x} \neq \underline{x}_0$ ($\hat{=}$ V positiv definit)
2. $V(\underline{x}) = 0 \Leftrightarrow \underline{x} = \underline{x}_0$

Sind diese Bedingungen erfüllt, so kann mittels $V(\underline{x})$ die Stabilität der Ruhelage überprüft werden. Hierbei sei \underline{x} der Zustandsvektor und \underline{x}_0 die Ruhelage des zu untersuchenden Systems.

Wird V als der Energie eines mathematischen Pendels aufgefasst, so lassen sich die Bedingungen sehr anschaulich erklären. Die Gesamtenergie des Pendels setzt sich aus potentieller und kinetischer Energie zusammen. Betrachten wir die Ruhelage 2 aus Bild 11.4. Dort gilt, dass sowohl die potentielle, als auch die kinetische Energie gleich Null ist. Als Umkehrschluss lässt sich sagen, dass die Gesamtenergie größer als Null ist, für alle Zustände, die von der Ruhelage abweichen (1. Bedingung). Befindet sich das Pendel in der genannten Ruhelage, so ist logischerweise die Gesamtenergie gleich Null (2. Bedingung).

Das Theorem von *Ljapunov* besagt:

1. ist $\dot{V}(\underline{x})$ positiv definit, so ist die Ruhelage *instabil*.
2. ist $\dot{V}(\underline{x}) \leq 0$ für alle $\underline{x} \in M$, wobei M eine fest definierte Umgebung von \underline{x}_0 sei, so ist die Ruhelage *lokal stabil*.
3. gilt $\dot{V}(\underline{x}) < 0$ für alle $\underline{x} \in M$, $\underline{x} \neq \underline{x}_0$, so ist die Ruhelage *lokal asymptotisch stabil*.

Auch diese Aussagen lassen sich am Beispiel der Energiefunktion sehr gut verdeutlichen: Wenn die Ableitung der Funktion V nach der Zeit größer als Null ist, so bedeutet das einen Anstieg der Energie. Die Ruhelage eines System mit ansteigender Energie ist somit instabil (1. Aussage). Sinkt die Energie jedoch, wie es in den drei anderen Aussagen der Fall ist, so läuft das System unweigerlich auf eine Ruhelage zu. Dies ist äquivalent mit Stabilität.

Das Theorem von *Ljapunov* stellt eine hinreichende Bedingung für Stabilität dar. Leider ist die Methode nach *Ljapunov* nicht konstruktiv. Wird keine Potenzialfunktion gefunden, mit der die Stabilität bewiesen werden kann, heißt das nicht zwangsläufig, dass überhaupt keine existiert. Es kann daraus also nicht auf Instabilität geschlossen werden. Im folgenden soll das Vorgehen an einem Beispiel erläutert werden.

11.4 Beispiel II: Nichtlinearer Schwinger

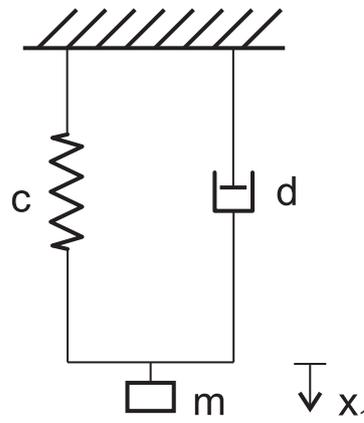


Bild 11.6: Feder-Dämpfer-System

In Bild 11.6 ist ein schwingfähiges System skizziert. Die Kennlinie der Federkraft F_c kann dabei beliebig nichtlinear sein. Es gibt drei Kraftgleichungen in diesem System:

1. Federkraft: $F_c = f(x_1)$
2. Dämpfungskraft: $F_d = d\dot{x}_1$
3. Trägheitskraft: $F_t = m\ddot{x}_1$

Die Zustandsgrößen seien x_1 und $x_2 = \dot{x}_1$.

$$\begin{aligned}
 \sum_{\text{Kräfte}} &= 0 & \Rightarrow & m\ddot{x}_1 + d\dot{x}_1 + f(x_1) = 0 \\
 & & \Leftrightarrow & \ddot{x}_1 = -\frac{d}{m}\dot{x}_1 - \frac{1}{m}f(x_1) \\
 & & \Leftrightarrow & \dot{x}_2 = -\frac{d}{m}x_2 - \frac{1}{m}f(x_1)
 \end{aligned} \tag{11.16}$$

Für die Ruhelage gilt: $x_1 = 0, x_2 = 0 \rightarrow \dot{x}_1 = 0, \dot{x}_2 = 0$

Betrachten wir nun die Energien des Systems:

- Kinetische Energie der trägen Masse:

$$E_{kin} = \frac{1}{2}m\dot{x}_1^2 = \frac{1}{2}mx_2^2 \tag{11.17}$$

- Potentielle Energie der Feder:

$$E_{pot} = \int_0^{x_1} f(x) dx \quad (11.18)$$

Eine lineare Feder mit der Federkonstanten c_F hätte die Energie $E_{pot} = \frac{1}{2}c_F x_1^2$.

Daraus ergibt sich die Gesamtenergie:

$$\begin{aligned} E &= E_{kin} + E_{pot} \\ &= \frac{1}{2} m x_2^2 + E_{pot} \geq 0 \end{aligned} \quad (11.19)$$

Für die Gesamtenergie sind sämtliche Voraussetzungen für eine Ljapunov-Funktion erfüllt. E ist also eine gültige Potenzialfunktion. Deswegen kann mit der Stabilitätsuntersuchung fortgefahren werden. Die zeitliche Änderung der Gesamtenergie ist:

$$\begin{aligned} \dot{E} &= \frac{dE}{dt} = \frac{dE_{kin}}{dt} + \frac{dE_{pot}}{dt} = \underbrace{m\dot{x}_2 x_2}_{\text{Kettenregel}} + \frac{dE_{pot}}{dx_1} \frac{dx_1}{dt} \\ &= m x_2 \dot{x}_2 + f(x_1) \dot{x}_1 \end{aligned} \quad (11.20)$$

Werden hier die Zustandsgleichungen eingesetzt, ergibt sich:

$$\begin{aligned} \dot{E} &= m x_2 \left(-\frac{d}{m} x_2 - \frac{1}{m} f(x_1) \right) + f(x_1) x_2 \\ &= -d x_2^2 \leq 0 \quad \text{für alle } x_1, x_2 \in \mathbb{R} \end{aligned} \quad (11.21)$$

Es ist zu beachten, dass $\dot{E}(\underline{x}) = 0$ auch für $\underline{x} \neq 0$ auftritt, z.B. für $x_2 = 0$ und x_1 beliebig. Daher ist die 2. Aussage von *Ljapunov* erfüllt und

⇒ Die Ruhelage $x_1 = 0, x_2 = 0$ ist lokal stabil.

Wir haben damit mittels der Stabilitätsuntersuchung nach Ljapunov die Stabilität der Ruhelage nachgewiesen, ohne eine einzige DGL lösen zu müssen.

11.5 Auffinden einer Ljapunov-Funktion

Bisher waren sämtliche betrachteten Ljapunov-Funktionen Energiefunktionen. Das muss jedoch nicht so sein, denn die Aussagen aus Kap. 11.3 sind allgemein gültig. In der Regel werden Ljapunov-Funktionen so gewählt, dass sie verallgemeinerte quadratische Terme enthalten. Drei Ansätze sind dabei besonders verbreitet:

1. $V = x_1^2 + x_2^2 + \dots$ (kreisförmige Höhenlinien)
2. $V = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots$ (achsenparallele Ellipsen)
3. $V = \sum_{i,k=1}^n P_{ik} x_i x_k$ mit $P_{ik} = P_{ki}$ (gedrehte Ellipsen)

Der letzte Ansatz soll noch etwas näher erläutert werden:

für $n = 3$ ergibt sich:

$$V = P_{11}x_1^2 + P_{22}x_2^2 + P_{33}x_3^2 + 2P_{12}x_1x_2 + P_{13}x_1x_3 + 2P_{23}x_2x_3$$

als Matrix geschrieben:

$$\underline{\underline{P}} = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}$$

$$\Rightarrow V = \underline{x}^T \underline{\underline{P}} \underline{x} \quad (\underline{\underline{P}} \text{ positiv definit, damit } V \geq 0) \quad (11.22)$$

Ansatz Nr. 3 kann auch auf lineare Systeme angewendet werden:

$$\begin{aligned} \dot{\underline{x}} &= \underline{\underline{A}} \underline{x} \\ V &= \underline{x}^T \underline{\underline{P}} \underline{x} \\ \Rightarrow \dot{V} &= \underline{x}^T \underline{\underline{P}} \dot{\underline{x}} + \dot{\underline{x}}^T \underline{\underline{P}} \underline{x} \\ &= \underline{x}^T \underline{\underline{P}} \underline{\underline{A}} \underline{x} + \underline{x}^T \underline{\underline{A}} \underline{\underline{P}} \underline{x} = \underline{x}^T (\underline{\underline{P}} \underline{\underline{A}} + \underline{\underline{A}} \underline{\underline{P}}) \underline{x} := \underline{x}^T (-\underline{\underline{Q}}) \underline{x} \end{aligned}$$

Allgemein gilt:

Zu einer Matrix $\underline{\underline{Q}} = -\underline{\underline{P}} \underline{\underline{A}} - \underline{\underline{A}} \underline{\underline{P}}$, positiv definit, existiert eine Matrix $\underline{\underline{P}}$, ebenfalls positiv definit, wenn alle Eigenwerte von $\underline{\underline{A}}$ einen negativen Realteil haben.

Wenn diese Bedingung erfüllt ist, so ist gilt $\dot{V} \leq 0$ für alle \underline{x} und das System ist somit nach Ljapunov stabil.

11.6 Beispiel III: Anwendung eines Ljapunov-Ansatzes

Es sei das System in Bild 11.7 gegeben.

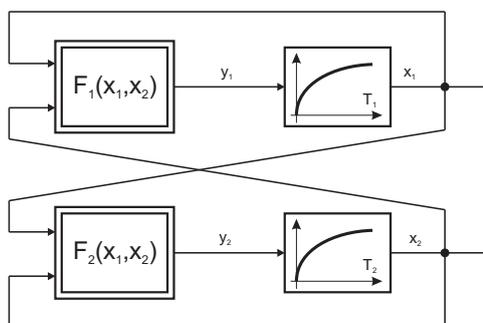


Bild 11.7: Beispielsystem

Außerdem sei $F_1(\underline{0}) = \underline{0}$ und $F_2(\underline{0}) = \underline{0}$. Damit ist $\underline{0}$ die Ruhelage. Aus den Übertragungsfunktionen der PT_1 -Glieder ergeben sich folgende Zustandsgleichungen:

$$\frac{1}{T_1} \dot{x}_1 + x_1 = y_1 \quad \text{mit} \quad y_1 = F_1(x_1, x_2) \quad (11.23)$$

$$\frac{1}{T_2} \dot{x}_2 + x_2 = y_2 \quad \text{mit} \quad y_2 = F_2(x_1, x_2) \quad (11.24)$$

$$\begin{aligned} \dot{x}_1 &= -T_1 x_1 + T_1 F_1(x_1, x_2) \\ \Rightarrow \dot{x}_2 &= -T_2 x_2 + T_2 F_2(x_1, x_2) \end{aligned} \quad (11.25)$$

Hierauf wird nun der Ljapunov-Ansatz der achsenparallelen Ellipsen angewendet:

$$V = \lambda_1 x_1^2 + \lambda_2 x_2^2 \quad \text{mit} \quad \lambda_1, \lambda_2 > 0 \quad (11.26)$$

$$\Rightarrow \dot{V} = \frac{dV}{dt} = 2\lambda_1 x_1 \dot{x}_1 + 2\lambda_2 x_2 \dot{x}_2 \quad (11.27)$$

Nach einsetzen der DGLs aus (11.25) ergibt sich:

$$\dot{V} = -2\lambda_1 T_1 x_1 [x_1 - F_1(x_1, x_2)] - 2\lambda_2 T_2 x_2 [x_2 - F_2(x_1, x_2)] \quad (11.28)$$

$$= -2\lambda_1 T_1 x_1^2 \left[1 - \frac{F_1(x_1, x_2)}{x_1}\right] - 2\lambda_2 T_2 x_2^2 \left[1 - \frac{F_2(x_1, x_2)}{x_2}\right] \quad (11.29)$$

Die geringste Anforderung für Stabilität nach Ljapunov ist, dass \dot{V} negativ semidefinit ist. Deswegen müssen in (11.29) die Ausdrücke in den Klammern ≥ 0 sein.

$$\Rightarrow \frac{F_1(x_1, x_2)}{x_1} \stackrel{!}{\leq} 1 \quad \text{und} \quad \frac{F_2(x_1, x_2)}{x_2} \stackrel{!}{\leq} 1 \quad (11.30)$$

Wenn diese Bedingungen erfüllt sind, so ist das System nach Ljapunov stabil.

12 Popov-Kriterium

Ein weiteres Kriterium zur Stabilitätskontrolle ist das *Popov-Kriterium*. Nehmen wir das folgende System als gegeben an:

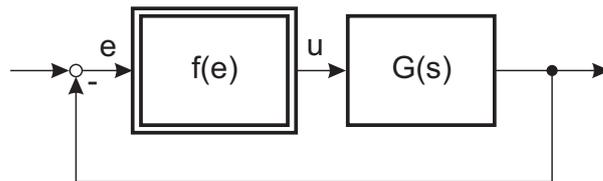


Bild 12.1: nichtlineares System

Es setzt sich aus einem nichtlinearen Teil, der keine Dynamik aufweist, und einem linearen dynamischen Teil zusammen. Ein Beispiel für die nichtlineare Kennlinie $f(e)$ ist in Bild 12.2 skizziert.

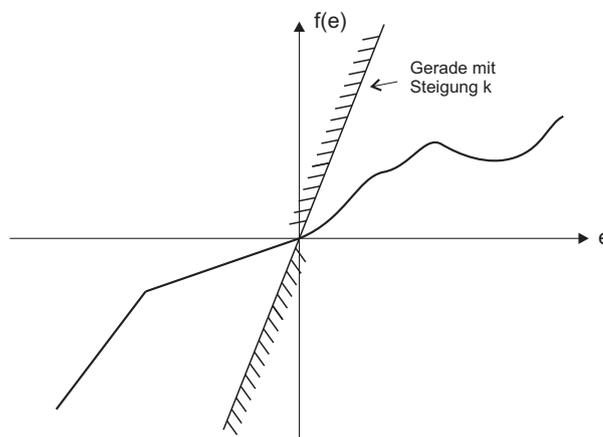


Bild 12.2: Beispiel für $f(e)$

Die Kennlinie darf sich nur im 1. und 3. Quadranten befinden. Ihre Steigung muss insgesamt kleiner sein als die Steigung k einer beliebigen Geraden. k sollte derart gewählt werden, dass $f(e)$ durch ke beschränkt wird. Die beliebige Gerade hat damit eine begrenzte Funktion. Da $f(e)$ zeitvariant sein darf, sind die einzigen Bedingungen, die erfüllt

werden müssen:

$$0 \stackrel{!}{<} f(e) \stackrel{!}{<} ke \quad \text{im positiven Bereich} \quad (12.1)$$

$$ke \stackrel{!}{<} f(e) \stackrel{!}{<} 0 \quad \text{im negativen Bereich} \quad (12.2)$$

Die Übertragungsfunktion $G(s)$ aus Bild 12.1 muss stabil sein ($Re\{s_i\} < 0$) und darf keinen Integrator beinhalten. Sind all diese Voraussetzungen erfüllt, so kann das *Popov-Kriterium* zur Feststellung der Stabilität angewendet werden:

$$\boxed{\text{Popov-Kriterium: } Re\{(1 + qj\omega)G(j\omega)\} + \frac{1}{k} \stackrel{!}{>} 0} \quad (12.3)$$

q ist dabei ein beliebiger aber reeller Faktor, dessen Funktion später noch erläutert wird. Gl. (12.3) kann folgendermaßen umgeformt werden:

$$Re\{G(j\omega)\} + qRe\{j\omega G(j\omega)\} + \frac{1}{k} > 0 \quad (12.4)$$

$$\text{mit } j\omega G(j\omega) = j\omega Re\{G(j\omega)\} - \omega Im\{G(j\omega)\}$$

$$Re\{G(j\omega)\} - q\omega Im\{G(j\omega)\} + \frac{1}{k} > 0 \quad (12.5)$$

Nun wird die Ortskurve von $G(j\omega)$ leicht modifiziert zur Popov-Ortskurve:

$$G_P(j\omega) = Re\{G(j\omega)\} + j\omega Im\{G(j\omega)\} = X + jY \quad (12.6)$$

Damit lässt sich das Popov-Kriterium umformulieren zu:

$$X - qY + \frac{1}{k} > 0 \quad (12.7)$$

Für den Grenzfall ergibt sich:

$$X - qY + \frac{1}{k} = 0 \quad (12.8)$$

$$\Rightarrow Y = \frac{1}{q} \left(X + \frac{1}{k} \right) \quad (12.9)$$

Dies ist die Gleichung der Geraden, die die Ortskurve begrenzen muss, damit Stabilität gewährleistet ist. Die Steigung der Geraden lässt sich durch den Faktor q einstellen.

Mathematisch formuliert ergeben sich aus (12.7) die Stabilitätsbedingungen:

$$Y \stackrel{!}{<} \frac{1}{q} \left(X + \frac{1}{k} \right) \quad \text{für } q > 0 \quad (12.10)$$

$$Y \stackrel{!}{>} \frac{1}{q} \left(X + \frac{1}{k} \right) \quad \text{für } q < 0 \quad (12.11)$$

Grafisch ist der Zusammenhang in Bild 12.3 dargestellt.

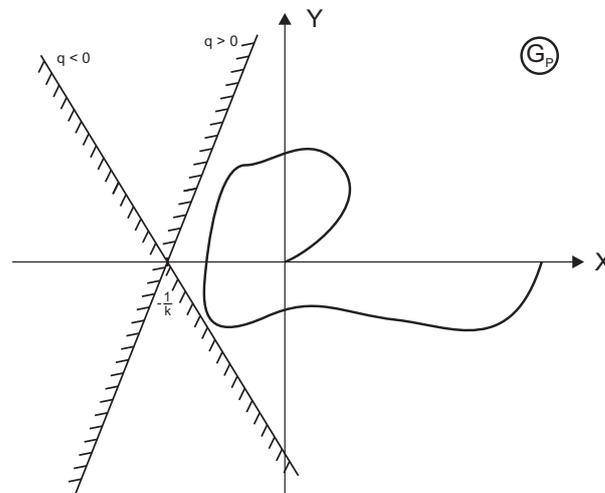


Bild 12.3: Stabilität nach Popov

Stabilität nach Popov bedeutet in dieser Darstellung, dass die Ortskurve G_P die begrenzende Gerade nicht berühren darf.

Literatur

- [1] Böcker, J., Hartmann, I. & Zwanzig, Ch. *Nichtlineare und adaptive Regelungssysteme*. Springer Verlag 1986.
- [2] Föllinger O. *Nichtlineare Regelungen I*. Oldenbourg Verlag 1993.
- [3] Leonhard, W. *Einführung in die Regelungstechnik*. Vieweg Verlag 1990.
- [4] Lunze, J. *Regelungstechnik 2*. Springer Verlag 1997.
- [5] Müller, K. *Entwurf robuster Regelungen*. Teubner Verlag 1996.
- [6] Slotine, E. & Li, W. *Applied Nonlinear Control*. Prentic-Hall 1991.
- [7] Unbehauen, H. *Regelungstechnik II*. Vieweg Verlag 1993.