

Regelungstechnik

Wintersemester 2003

Claus Hillermeier

Kapitelübersicht :

1. Aufgabe der Regelungstechnik und Grundlagen
2. Modellierung der Struktur dynamischer Systeme
3. Beschreibung dynamischer Systeme im Zeitbereich
4. Beschreibung dynamischer Systeme im Frequenzbereich
5. Der Regelkreis
6. Entwurf einschleifiger Regelkreise

Vorbemerkung:

Aufbau der Vorlesung und Stoffauswahl lehnen sich eng an das Springer-Lehrbuch
„Regelungstechnik 1“ von Prof. Dr.-Ing. Jan Lunze an.

Regelungstechnik

Wintersemester

Claus Hillermeier

Kapitel 1 : Aufgaben der Regelungstechnik und Grundlagen

Regelungstechnik befasst sich mit der Aufgabe, einen dynamischen (d.h. sich zeitlich verändernden Prozess von außen gezielt zu beeinflussen).

Beispiel : (siehe [Schmidt 2001]):

Ein Fahrzeug (z.B. Auto) ist auf gerader Straße in der Spurmitte zu halten.

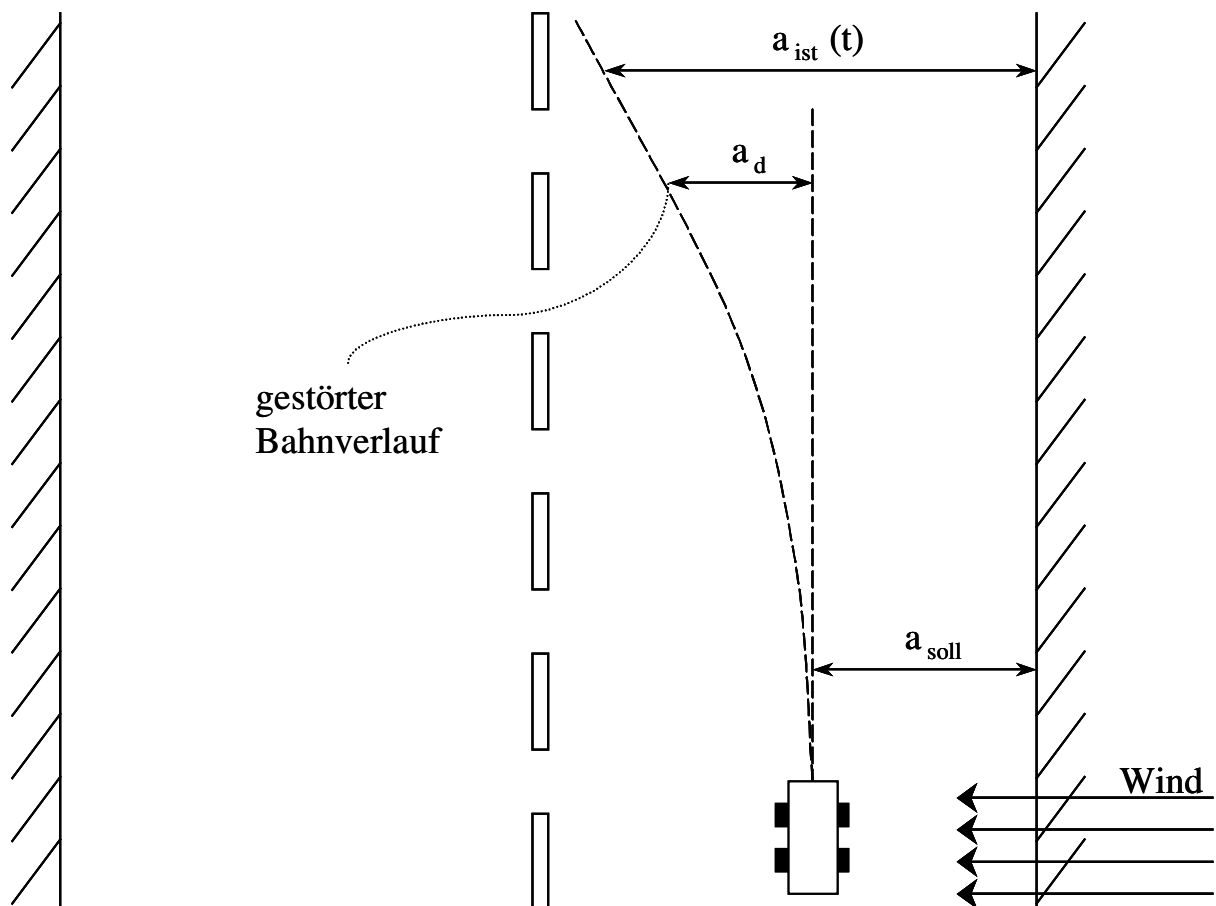


Abb. 1 : Spurmittenführung eines Fahrzeugs

Dabei wirken Störungen, wie z.B. Seitenwind und Radschlupf, auf das Fahrzeug ein (sog. Störgrößen).

Übernimmt ein Mensch die Lenkfunktion, so handelt es sich um eine manuelle Steuerung oder Regelung: Aufgabe des Fahrers ist es, den Ist-Abstand $a_{\text{ist}}(t)$ des Fahrzeugs vom Straßenrand auf dem gewünschten Sollabstand a_{soll} (z.B. in Spurmitte) zu halten, also den Abstandsfehler $a_d(t) = a_{\text{soll}} - a_{\text{ist}}(t)$ zu Null zu machen. Dabei gehen wir davon aus, dass sich das Fahrzeug zum Anfangszeitpunkt $t = 0$ in der Spurmitte befindet, d.h. dass $a_{\text{ist}}(0) = a_{\text{soll}}$ und damit $a_d(0) = 0$. Prinzipiell kann die Aufgabe durch Steuerung oder durch Regelung angegangen werden.

1.1 Manuelle Steuerung

Der manuellen Steuerung entspricht im Beispiel das Fahren mit geschlossenen Augen. Da laut Newtonschen Trägheitsgesetz ein Fahrzeug ohne äußere Einwirkung geradlinig weiterfährt, wird der „Blind-Fahrer“ als einzige sinnvolle Steuerungsmaßnahme das Lenkrad zu Beginn in die „Geradeausfahr-Position“ bringen. Der Istwert des Abstands wird bei dieser Vorgehensweise **nicht** zum laufenden Ausgleich eines Abstandsfehlers herangezogen, was bereits das wesentliche Merkmal einer Steuerung ist.

Da $a_d(t)$ im Verlauf des Steuervorgangs nicht laufend überwacht wird, führen unvorhersagbare Einflüsse, wie etwa Seitenwind oder Radschlupf, schnell zu einem Anwachsen von $a_d(t)$ für $t > 0$, vergl. Abb. 1. Somit kann durch Steuerung die gestellte Fahrzeugführungsaufgabe nicht gelöst werden. Vielmehr zeigt das vom Fahrer gesteuerte System (= Fahrzeug) das gleiche zeitlich instabile Verhalten wie ein völlig ungesteuertes Fahrzeug.

1.2 Manuelle Regelung

Fahren mit offenen Augen entspricht dagegen der manuellen Regelung: Die Größe a_d wird vom Fahrer ständig beobachtet und über die Lenkung zur Bahnkorrektur herangezogen, so dass $a_d(t)$ stets in der Nähe von Null bleibt. a_d wird auch als **Regelabweichung** bezeichnet. Bei der hier betrachteten Regelung spricht man von einer **Festwertregelung**, da $a_{\text{soll}} = \text{const}$ ist.

Wesentliches Merkmal einer Regelung ist also das Auftreten eines **geschlossenen Wirkungskreislaufs**, was bedeutet, dass die vom Regler zu beeinflussende Aufgabengröße (die sog. Regelgröße), hier $a_{\text{ist}}(t)$, laufend gemessen und zur Korrektur herangezogen wird.

Der geschlossene Wirkungskreislauf zeigt sich deutlich im sog. **Blockschaltbild** der soeben beispielhaft skizzierten Regelung, siehe Abb. 2. Blockschaltbilder, auch Strukturbilder genannt, bauen sich aus den Teilsystem-Blöcken auf. Gerichtete Kanten sowie Summations- und Verteilstellen beschreiben die Ursache-Wirkungs-Beziehungen zwischen den Teilsystemen.

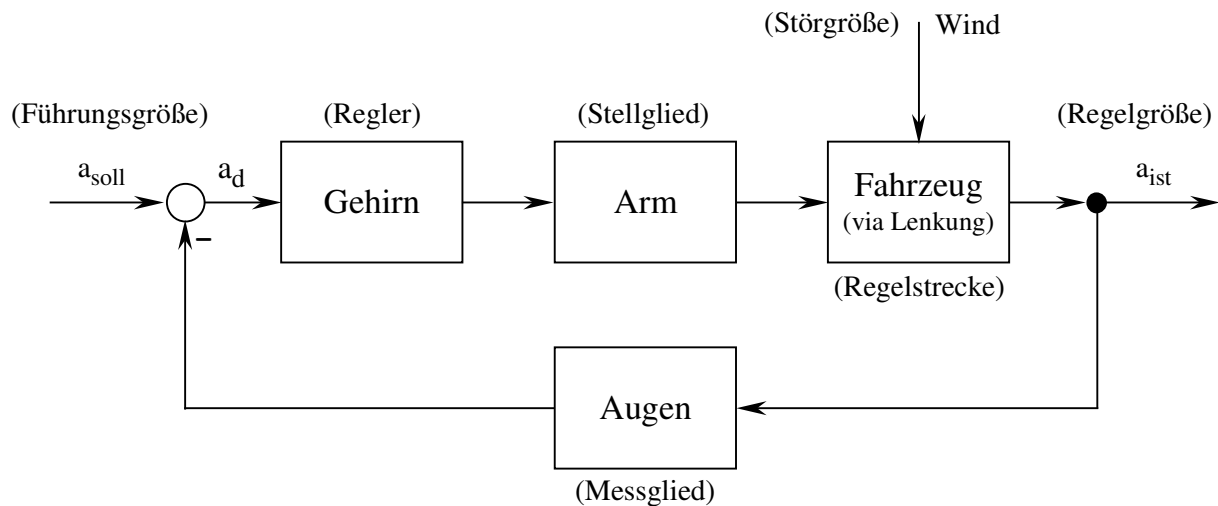


Abb. 2 : Blockschaltbild der manuellen Fahrzeugregelung

Streicht man in Abb. 2 die Rückmeldung der Regelgröße a_{ist} über die Augen ans Gehirn, so bleibt nur die offene Wirkungskette Gehirn – Arm – Kfz/Lenkung übrig. Dies ist genau das Blockschaltbild der manuellen (Fahrzeug-)Steuerung.

1.3 Generelle Regelungsaufgabe

Verallgemeinern wir das obige Beispiel, so gelangen wir zur generellen Regelungsaufgabe. Als zu regelndes Objekt (oben: Fahrzeug) ist uns im allgemeinen Fall ein **dynamisches System** vorgegeben. Dies ist eine Funktionseinheit, deren wichtigste Kenngrößen sich zeitlich ändern und die deshalb als Funktionen der Zeit dargestellt werden. Beispiele für technische dynamische Systeme: elektrischer Antriebsmotor, Flugzeug, Satellit, industrielle Fertigungsanlage; Beispiele für nichttechnische dynamische Systeme: Wirtschaftsunternehmen, Organ im menschlichen Körper.

Die Kopplungen zwischen dem dynamischen System und seiner Systemumgebung lassen sich unterteilen in **Eingangsgrößen**, die auf das System einwirken und die zeitlichen Veränderungen innerhalb des Systems verursachen, und **Ausgangsgrößen**, die das Verhalten des Systems als Reaktion auf die Eingangsgrößen beschreiben.

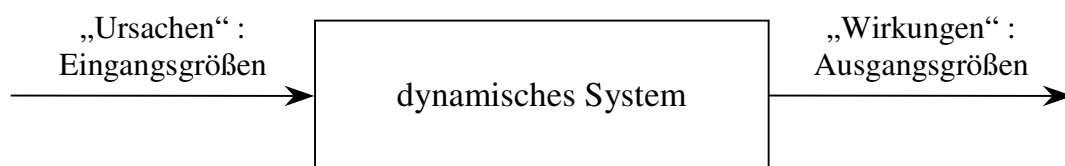


Abb. 3 : Blockschaltbild eines dynamischen Systems

Die **Regelungsaufgabe** lautet nun:

Gegeben ist ein dynamisches System (**Regelstrecke** oder kurz **Strecke**) mit von außen beeinflussbaren Eingangsgrößen (**Stellgrößen**) und einer oder mehreren messbaren Ausgangsgrößen (**Regelgrößen**). Als Regelungsziel ist typischerweise die Aufgabe vorgegeben, die Regelgröße(n) auf vorgegebenen konstanten Werten zu halten (\rightarrow Festwertregelung, siehe Fahrzeug-Beispiel) oder in einer vorgegebenen Weise zeitlich zu verändern. Gleichzeitig soll die Wirkung äußerer Störungen unterdrückt werden.

Gesucht ist eine Regeleinrichtung (kurz: **Regler**), die unter Nutzung der gemessenen Werte die Stellgröße(n) so vorgibt, dass das geregelte System das Regelungsziel erfüllt.

In dieser Vorlesung werden wir uns auf Systeme mit nur einer Stell- und nur einer Regelgröße (englisch: single-input single-output (SISO) control systems) beschränken. Für SISO-Systeme illustriert Abb. 4 die Regelungsaufgabe.

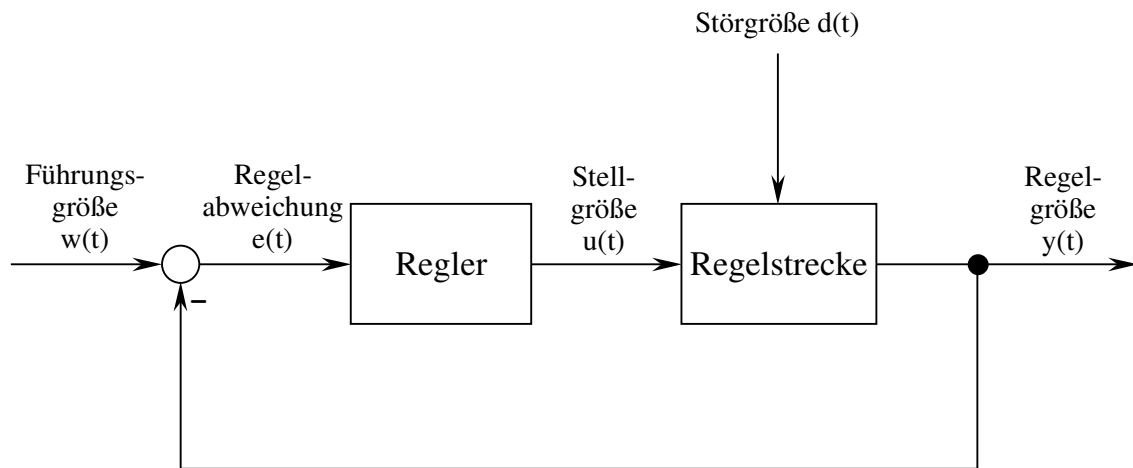


Abb. 4 : Blockschaltbild des Regelkreises (Grundstruktur)

Die Regelgröße $y(t)$ hängt von der Stellgröße $u(t)$ und einer Störgröße $d(t)$ ab, die nicht beeinflussbar ist. Ziel der Regelung ist es, die Regelgröße einer vorgegebenen Führungsgröße $w(t)$ nachzuführen, so dass idealerweise $y(t) = w(t)$ für alle Zeitpunkte t gilt. Der Regler muss deshalb die Stellgröße $u(t)$ so vorgeben, dass die Regelgröße der Führungsgröße angepasst und der Einfluss der Störgröße auf die Regelgröße kompensiert wird. Dabei steht dem Regler die Information zur Verfügung, um wieviel der aktuelle Wert $y(t)$ der Regelgröße vom gewünschten Werte $w(t)$ abweicht, also die sog. **Regelabweichung**

$$e(t) := w(t) - y(t) \quad (1.1)$$

Das **Reglergesetz** hat deshalb die Form

$$\underbrace{u(t)}_{\text{gemeint : \{aktueller Wert\}}} = \underbrace{k_R(e(t))}_{\text{\{gesamte Funktion } e: [0, t] \rightarrow \mathbb{R} \text{ \}}}} \quad (1.2)$$

Der Vergleich mit realen Regelungen enthüllt noch einige Vereinfachungen, die in Abb. 4 vorgenommen wurden:

- (A) Häufig muss zwischen der tatsächlich vorliegenden Regelgröße $y(t)$ und der gemessenen Regelgröße $y_m(t)$ unterschieden werden, weil das Messglied (d.h. der Sensor) selbst dynamische Eigenschaften besitzt, aufgrund derer die Messgröße u.U. erheblich von der tatsächlich im Prozess vorhandenen Größe abweicht.
- (B) Der vom Regler vorgegebene Wert $u(t)$ für die Stellgröße wird durch das Stellglied in den am Prozess wirksamen Wert $u_R(t)$ umgesetzt. Stellglieder weisen i.a. ebenfalls ein eigenes dynamisches Verhalten auf und besitzen häufig nichtlineare Eigenschaften, z.B. obere und untere Schranken für die Stellgröße $u_R(t)$.
- (C) Im Hinblick auf (A) kann es sinnvoll sein, das Reglergesetz (1.2) so zu erweitern, dass der Regler anstelle der Regelabweichung die Führungsgröße $w(t)$ und die gemessene Regelgröße $y_m(t)$ als getrennte Eingangsgrößen erhält. Dies eröffnet z.B. die Möglichkeit, das Messrauschen aus der Messgröße zu filtern, bevor Regel- und Führungsgröße verglichen werden.

Berücksichtigt man (A) – (C), so gelangt man zur erweiterten Struktur eines Regelkreises:

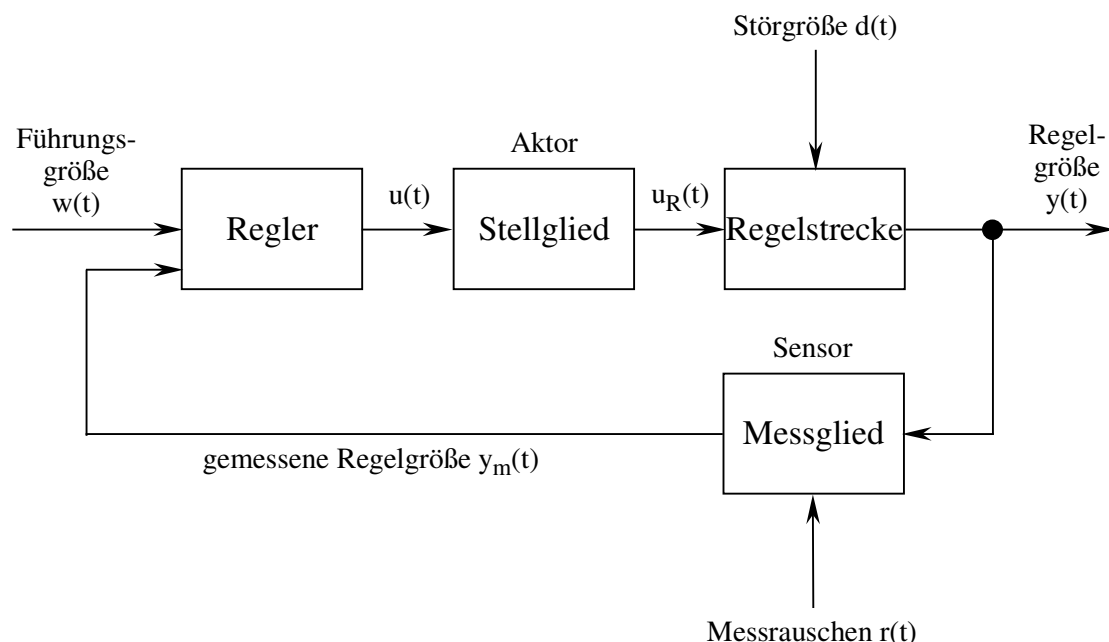


Abb. 5 : Erweiterte Struktur eines Regelkreises

Fasst man in Abb. 5 das Stellglied, die Regelstrecke und das Messglied zur „erweiterten Regelstrecke“ zusammen und vernachlässigt man den Unterschied zwischen $y_m(t)$ und $y(t)$, so erhält man wieder die Grundstruktur des Regelkreises (Abb. 4).

Die drei wichtigen Schritte jeder Regelung seien nochmals explizit notiert:

- 1) **Messen.** Die Regelgröße wird entweder direkt gemessen oder – bei nicht messbaren Regelgrößen wie z.B. Qualitätskennwerten – aus anderen Messgrößen berechnet.
- 2) **Vergleichen** des Wertes der Regelgröße mit dem Wert der Führungsgröße.
Die Differenz ist die Regelabweichung $e(t) = w(t) - y(t)$.
- 3) **Stellen.** Aus der Regelabweichung wird gemäß des Reglergesetzes die Stellgröße bestimmt.
Zentrale Aufgabe des Regelungstechnikers ist es, ein für die dynamischen Eigenschaften der Regelstrecke maßgeschneidertes Reglergesetz zu entwerfen.

1.4 Vorteile der Rückkopplung

Wie bei der Lenkungs Aufgabe (Abschnitt 1.1) gesehen, kann ein dynamisches System prinzipiell auch ohne Rückführung gesteuert werden.

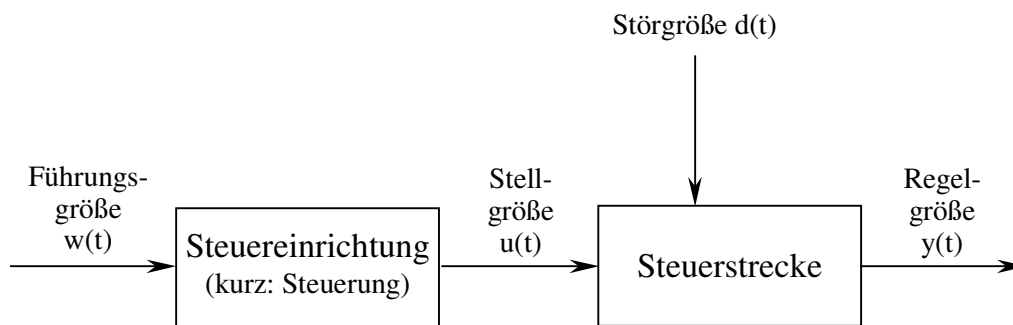


Abb. 6 : Grundstruktur einer Steuerung: offene Wirkungskette

Damit in einer Steuerkette das gewünschte Ziel $y(t) = w(t)$ erreicht werden kann, muss

- erstens die dynamischen Eigenschaften der Steuerstrecke genau bekannt sein. In Vorwegnahme der dynamischen Reaktion der Steuerstrecke auf die Eingangsgröße $u(t)$ wird dann $u(t)$ so gewählt, dass die Ausgangsgröße $y(t)$ dem zeitlichen Verlauf von $w(t)$ möglichst nahe kommt.
- Zweitens darf die Steuerstrecke nicht gestört sein. Da die Steuereinrichtung keine Informationen über die Störgröße erhält, kann sie nicht auf die Störung reagieren. Die Wirkungen der Störung übertragen sich in der Steuerkette ungemindert auf die Regelgröße.

Ersetzt man die Steuerung durch eine Regelung, so fallen beide Voraussetzungen weg. Das Prinzip des geschlossenen Wirkungskreises erlaubt es nämlich, das Regelungsziel $y(t) \approx w(t)$ auch dann gut zu erreichen, wenn

- (a) die dynamischen Eigenschaften der Regelstrecke nicht genau bekannt sind oder sich zeitlich ändern, und/oder wenn
- (b) die Regelstrecke durch nichtmessbare Störungen beeinflusst wird.

Eigenschaft (a) wird als **Robustheit** gegenüber Modellierungs-Unsicherheiten und Eigenschaft (b) als **Störkompensation** bezeichnet. Aufgrund der Robustheit von Regelungen kann ein am genauen Regelstreckenmodell entworfener Regler auch dann noch eingesetzt werden, wenn sich die Eigenschaften der Regelstrecke wesentlich verändert haben. Andererseits kann man Regler auch dann entwerfen, wenn nur ein Näherungsmodell der Strecke verfügbar ist.

Beide Vorzüge der Regelung beruhen auf dem **Prinzip der Rückkopplung** (engl. feedback). Informationen über das aktuelle Verhalten der Regelstrecke werden zur Regeleinrichtung zurückgeführt. Der Regler kann deshalb die Stellgröße in Abhängigkeit von den sich tatsächlich einstellenden Werten der Regelgröße – und nicht nur in Abhängigkeit von dem mit einem Modell vorhergesagten Verhalten der Regelstrecke – festlegen.

1.5 Prozessebene und Modellebene

Bei der Lösung einer Regelungsaufgabe in der Praxis können zwei Phasen unterschieden werden.

- **Vorbereitungsphase:** Für eine gestellte Regelungsaufgabe wird das Reglergesetz (siehe Gl. (1.2)) ermittelt und durch eine Funktionseinheit, z.B. als Reglerbaustein, technisch realisiert. Der Reglerbaustein wird mit der Regelstrecke zu einem Kreis verschaltet.
- **Arbeitsphase:** Der gerätetechnisch realisierte Regler bestimmt selbsttätig und kontinuierlich aus dem aktuellen Wert $e(t)$ der Regelabweichung den aktuellen Wert $u(t)$ der Stellgröße.

In dieser Vorlesung wollen wir lernen, wie man für eine gegebene Regelungsaufgabe ein geeignetes Reglergesetz bestimmt bzw. wie man gegebenenfalls erkennt, dass und warum die Regelungsaufgabe unter Beachtung aller in ihr enthaltenen Randbedingungen nicht lösbar ist.

Der prinzipielle Lösungsweg fusst darauf, dass wir neben der eigentlichen Prozessebene noch eine weitere Ebene, die Modellebene, ins Spiel bringen.

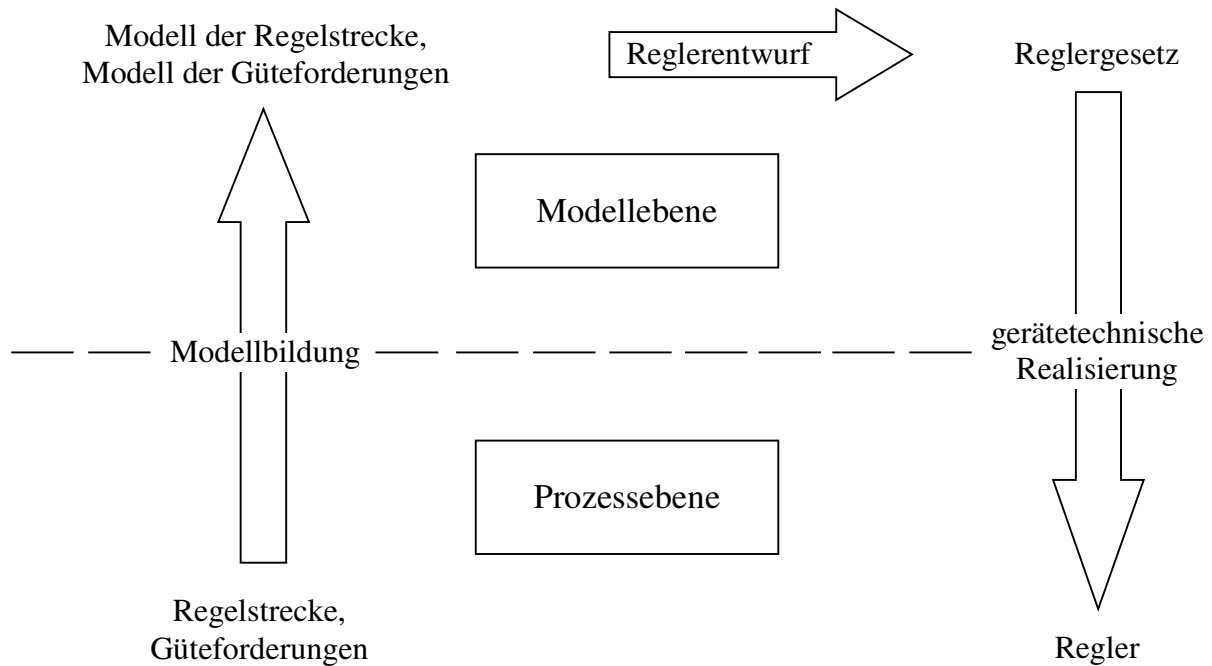


Abb. 7 : Prinzipieller Lösungsweg für Regelungsaufgaben

Ausgangs- und Endpunkt für Regelungsaufgaben ist die Prozessebene, auf der die Regelstrecke als „Gerät“ vor dem Betrachter steht und auf der der Regler als neues Gerät in Betrieb genommen werden soll.

Als erster Lösungsschritt wird ein quantitatives Modell der Regelstrecke und der Güteforderungen erstellt. Die Analyse der Regelstrecke, die Auswahl der Regelungsstruktur und die Festlegung des Reglergesetzes erfolgen nun auf der Modellebene, d.h. unter Verwendung des erstellten Modells.

Der skizzierte Weg zur Lösung einer Regelungsaufgabe spiegelt sich in den drei Schwerpunkten dieser Vorlesung wider:

(A) Modellbildung

Das Blockschaltbild (Kap. 2) zeigt die Gliederung des Systems in seine Elemente und deren Verkopplung, gibt jedoch die dynamischen Eigenschaften der Elemente noch nicht genau wieder.

Modelle in Form von Differentialgleichungen (siehe Kap. 3) beschreiben das zeitliche Verhalten dynamischer Systeme quantitativ exakt. Neben der Differentialgleichung ist das Zustandsraummodell (Kap. 3) eine weitere „standardisierte“ Modellform, auf der viele Analyse- und Entwurfsverfahren für Regelkreise beruhen.

(B) Analyse dynamischer Systeme

Mit Hilfe von Zustandsraummodellen dynamischer Systeme kann für eine gegebene Eingangsgröße $u(t)$ die durch das System erzeugte Ausgangsgröße $y(t)$ berechnet wer-

den (Kap. 3). Rechentechnisch sehr elegant ist die Beschreibung und Analyse dynamischer Systeme im Frequenzbereich (Kap. 4). Ein weiteres Analysewerkzeug ist die Zerlegung des betrachteten Systems in elementare Übertragungsglieder, deren Eigenschaften in Kap. 3 und 4 diskutiert werden.

(C) **Reglerentwurf**

Kapitel 5 untersucht, wie das stationäre und das Übergangsverhalten von Regelkreisen durch die Wahl des Reglers beeinflusst werden kann, welche Regelgüte erreicht werden kann bzw. welche Entwurfskompromisse gemacht werden müssen. Da die Stabilität des Regelkreises ein zentrales Entwurfskriterium ist, werden wir die Stabilität rückgekoppelter Systeme eingehend untersuchen.

Kapitel 6 stellt schließlich verschiedene Verfahren zum Reglerentwurf vor.

Kapitel 2 : *Modellierung der Struktur dynamischer Systeme*

Zur Analyse dynamischer Systeme und zum Reglerentwurf eignen sich besonders Modelle, die das Verhalten dynamischer Systeme durch **Signale** (= zeitveränderliche Größen) und **Signalumformungen** beschreiben.

Ein erster Schritt einer solchen Modellbildung ist es, das zu betrachtende System von seiner Umgebung abzugrenzen und die Kopplungen zwischen System und Umgebung durch (ein- bzw. auslaufende) Signale zu beschreiben (siehe Abb. 3). Das als Block dargestellte System läßt sich i.a. in Teilsysteme zerlegen, und die Ursache-Wirkungs-Beziehungen zwischen diesen Teilsystemen lassen sich wiederum durch ausgetauschte Signale modellieren. Diese sukzessive Modell-Verfeinerung läßt sich so lange fortsetzen, bis als Teilsysteme gewisse elementare Standard-Übertragungsglieder auftauchen bzw. bis der gewünschte Detaillierungsgrad erreicht ist. Das entstandene **Bockschaltbild** (auch Blockschaltplan, Wirkungsschema oder Strukturbild genannt) stellt die Zerlegung des Systems in seine Elemente sowie die kausalen Kopplungen zwischen diesen Elementen über Signale dar. Strukturbilder lassen erkennen, wo im System Rückkopplungen auftreten und ob das System eventuell in unabhängige oder schwach gekoppelte Teilsysteme zerlegt werden kann, und tragen so entscheidend zum Verständnis des Systemverhaltens bei.

In der Praxis entstehen Strukturbilder oft aus einer umgangssprachlichen Beschreibung des Systems samt seiner Elemente, Funktionen und strukturellen Verknüpfungen („Wortmodell“) und werden i.a. **vor** der quantitativen Modellierung des Systemverhaltens in Form von Differentialgleichungen erstellt.

„Grammatik“ von Blockschaltbildern:

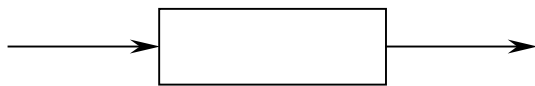
- 1) Pfeile stellen zeitlich veränderliche Größen, also Signale, dar.
- 2) Blöcke stellen Signalverarbeitungs-Einheiten, also dynamische Systeme, dar. Sie können prinzipiell mehrere Eingangs- und mehrere Ausgangssignale besitzen. Zu einem Vektor zusammengefasste Signale werden oft durch einen Doppelpfeil \Rightarrow gekennzeichnet. Blöcke mit nur einer Eingangs- und einer Ausgangsgröße heißen auch Übertragungsglieder.
- 3) Signale haben eine eindeutige Wirkungsrichtung, die durch die Pfeilrichtung beschrieben wird. Der Pfeilanzfang gibt an, wo das Signal entsteht. Die Pfeilspitze weist auf den Block, auf den das Signal wirkt, d.h. in dem es als Ursache anderer Vorgänge auftritt.

Blöcke sind stets rückwirkungsfrei, d.h. die Vorgänge innerhalb eines Blockes verändern die Eingangssignale nicht (es sei denn, ein Ausgangssignal wird zum Eingang desselben Blockes zurückgeführt).

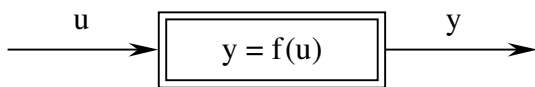
In einer ersten Modellierungsstufe wird der Inhalt der Blöcke zunächst verbal beschrieben (z.B. Gerätename wie „Kühler“ oder „Gleichstrommotor“, → Geräte-Blockschaltbild).

In einer zweiten Modellierungs- bzw. Abstraktionsstufe können diese verbalen Beschreibungen durch quantitative Modelle (z.B. Differentialgleichungen, aus denen sich dann die Funktionalbeziehungen zwischen Eingangs- und Ausgangssignalen berechnen lassen) ersetzt werden. → **mathematisches Blockschaltbild**. Die quantitativen Beziehungen können entweder aus den in den Elementen wirkenden physikalischen Gesetzmäßigkeiten abgeleitet werden (= theoretische Modellbildung) oder aus Messdaten bestimmt werden (= experimentelle Modellbildung oder Identifikation).

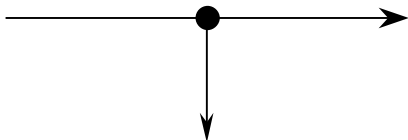
Mathematische Blockschaltbilder verwenden folgende Symbole:



Übertragungsglied (häufig als linear vorausgesetzt)

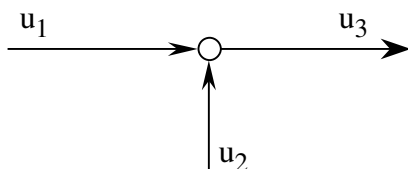


nichtlineares Übertragungsglied



Signalverzweigung

(beachte: Die beiden Pfeilenden stellen dasselbe Signal dar (und nicht etwa Teilströme eines von links zum Knoten fließenden Stromes!))



Summationsstelle $u_3 = u_1 + u_2$

(Minuszeichen rechts von der Pfeilspitze → Subtraktion)

Beispiel 1: Regelung eines Abwasser-Reaktors

Abb. 8 zeigt den verfahrenstechnischen Aufbau eines Reaktors, in dem Abwasser durch Bioorganismen gereinigt wird.

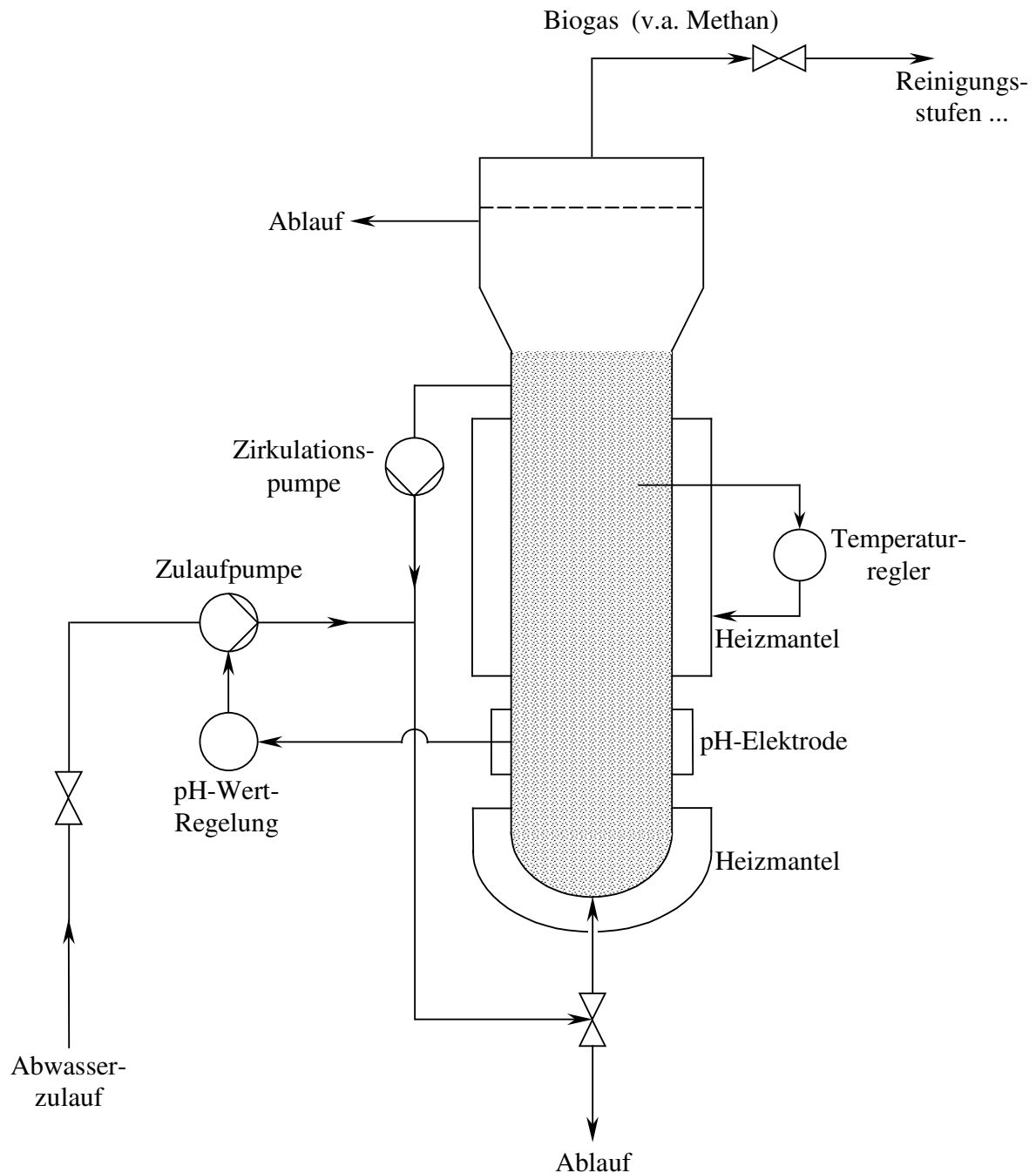


Abb. 8 : Abwasser-Reaktor aus verfahrenstechnischer Sicht

Das Abwasser wird durch eine Zulaufpumpe in den Zirkulationskreislauf des Reaktors geführt. Der Zirkulationskreis entnimmt Abwasser aus dem oberen Bereich des Reaktors und pumpt es unten wieder in den Reaktor hinein. So wird eine gute Durchmischung von Abwas-

ser und Bioorganismen erreicht und die Biomasse über die gesamte Höhe des Reaktors verteilt. Das geklärte Wasser wird aus dem Reaktorkopf abgezogen, wobei ein Sedimeter verhindert, dass dabei Biomasse mit abfließt. Das Biogas (hauptsächlich Methan) wird vor der weiteren Verwendung gereinigt. Durch den Heizmantel wird der Reaktor auf einer Temperatur von 38°C gehalten.

Betrachten wir nun aus regelungstechnischer Sicht die pH-Wert-Regelung dieser Anlage. Im Hinblick auf diese Regelungsaufgabe kann von vielen Einzelheiten des Anlagenschemas abstrahiert werden.

Aufgrund des Zirkulationskreises können wir davon ausgehen, dass der Reaktor homogen durchmischt und folglich der pH-Wert des Reaktorinhalts überall derselbe ist. Ansonsten muss der Zirkulationskreis nicht weiter betrachtet werden.

Aufgrund der Temperaturregelung können wir annehmen, dass der Reaktor eine konstante Temperatur besitzt, so dass wechselnde Temperaturen des zugeführten Abwassers und der Reaktorumgebung nicht als Störgröße betrachtet werden müssen.

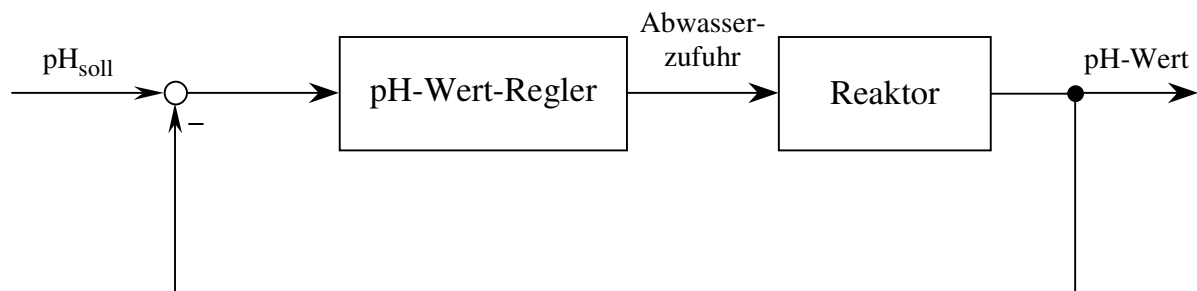


Abb. 9 : Blockschaltbild der pH-Wert-Regelung

Regelgröße ist der pH-Wert, Stellgröße die Abwasserzufuhr (beachte: Abwasser = „Nährstoff“ für die Bioorganismen). Die Abwasserzufuhr wird durch die Drehzahl der Zulaufpumpe bestimmt. Streng genommen gibt der Regler die Spannung für den Motor der Zulaufpumpe vor. Wenn man aber davon ausgehen kann, dass eine Spannungsänderung verzögerungsfrei eine Drehzahländerung nach sich zieht, so unterscheiden sich Abwasserzufuhr und Spannung nur um einen konstanten Faktor (siehe Abb. 9).

Die Regelgröße wird durch den Abbauprozess ständig verändert (dynamisches Verhalten der Regelstrecke!). Die Regelung soll dieses dynamische Verhalten ausgleichen und den pH-Wert dem vorgegebenen Sollwert angleichen.

Das Beispiel macht deutlich, wie der Regelungstechniker aus einem häufig komplizierten Anlagenschema durch problemangepasste Abstraktion ein oft recht einfaches Blockschaltbild

macht. Es zeigt, welche Teile der Anlage dynamisch modelliert und beim Reglerentwurf berücksichtigt werden müssen.

Beispiel 2: Modell einer Lagerhaltung

Vom Blickwinkel des Lagers eines Kaufhauses aus sollen die dynamischen Prozesse „Verkauf, Lagerhaltung und Bestellung“ durch ein Blockschaltbild modelliert werden. Die Tagesbestellung, die von der Einkaufsabteilung des Kaufhauses an den Hersteller übermittelt wird, ist die Stellgröße. Ziel ist es, einen vorgegebenen, evtl. zeitlich veränderlichen Bestand im Lager zu erreichen (Regelgröße).

Der – kaum zu beeinflussende und schwer vorhersagbare – Tagesverkauf im Kaufhaus wirkt dabei im regelungstechnischen Sinn als Störgröße. Entsprechend dieser Abgrenzung der Regelstrecke ist die Tagesbestellung die Eingangsgröße und der Lagerbestand die Ausgangsgröße der Regelstrecke.

(Lagerbestand = Vortagesbestand + angelieferte Stückzahl – verkaufte Stückzahl).

In Abhängigkeit vom Lagerdefizit wird die Ware nachbestellt. Der Bestellvorgang hat also die Funktion eines Reglers.

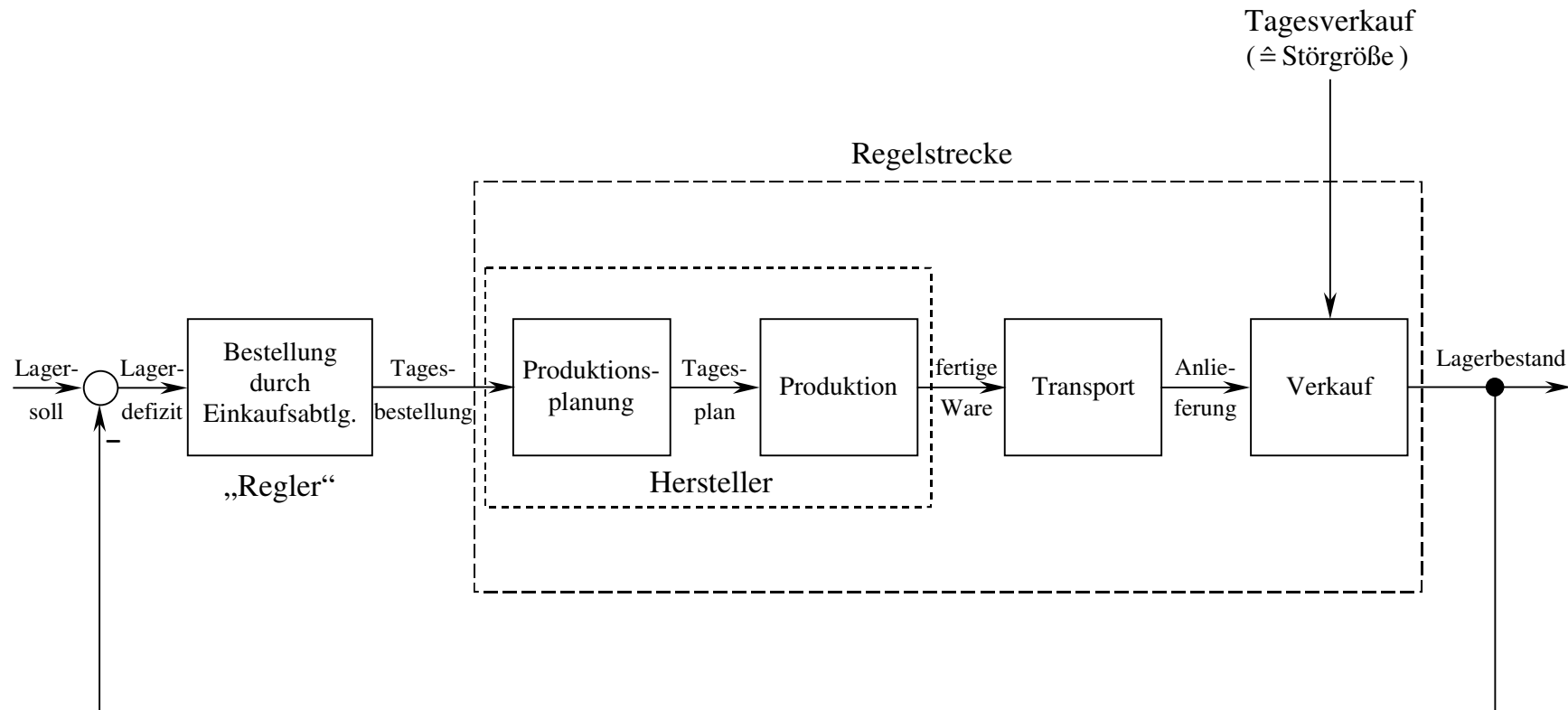


Abb. 10 : Blockschaltbild für das Modell einer Lagerhaltung

Für die Dynamik des Systems ist wesentlich, dass von der Tagesbestellung bis zur Anlieferung der bestellten Ware eine erhebliche Zeit („Totzeit“) vergeht. Das Blockschaltbild in Abb. 10 lässt diese Totzeit(en) unschwer erkennen und macht darüber hinaus auch die im System enthaltene Rückkopplungsschleife sichtbar. Allgemein gilt: Totzeiten in rückgekoppelten Systemen sind kritisch für die Stabilität des Systems (siehe Kapitel 5). Aus dem Blockschaltbild lässt sich also bereits ablesen, dass der Bestellvorgang auf die vorhandene Totzeit Rücksicht nehmen muss. Die Bestellung darf keinesfalls so vorgenommen werden, dass jeweils das aktuelle Lagerdefizit als Tagesbestellung verwendet wird (sondern: bereits abgegebene, aber noch nicht realisierte Tagesbestellungen sind zu berücksichtigen).

Das Beispiel lehrt, dass

- auch nichttechnische Prozesse häufig durch Blockschaltbilder modellierbar sind, und dass
- wichtige Eigenschaften dynamischer Systeme aus dem Blockschaltbild erkennbar sind, ohne dass die einzelnen Blöcke quantitativ exakt beschrieben werden müssen.

Kapitel 3 : Beschreibung dynamischer Systeme im Zeitbereich

3.1 Beschreibung linearer Systeme durch Differentialgleichungen

3.1.1 Beispiel: Modellierung eines Reihenschwingkreises durch eine Differentialgleichung

Im Reihenschwingkreis von Abb. 11 wird die Spannung u_1 von außen vorgegeben und die Spannung u_2 , mit der der Schwingkreis auf u_1 reagiert, gemessen. Die dynamische Struktur des Systems ist rechts dargestellt.

- R : Ohmscher Widerstand
- L : Induktivität („Spule“)
- C : Kapazität (Kondensator)

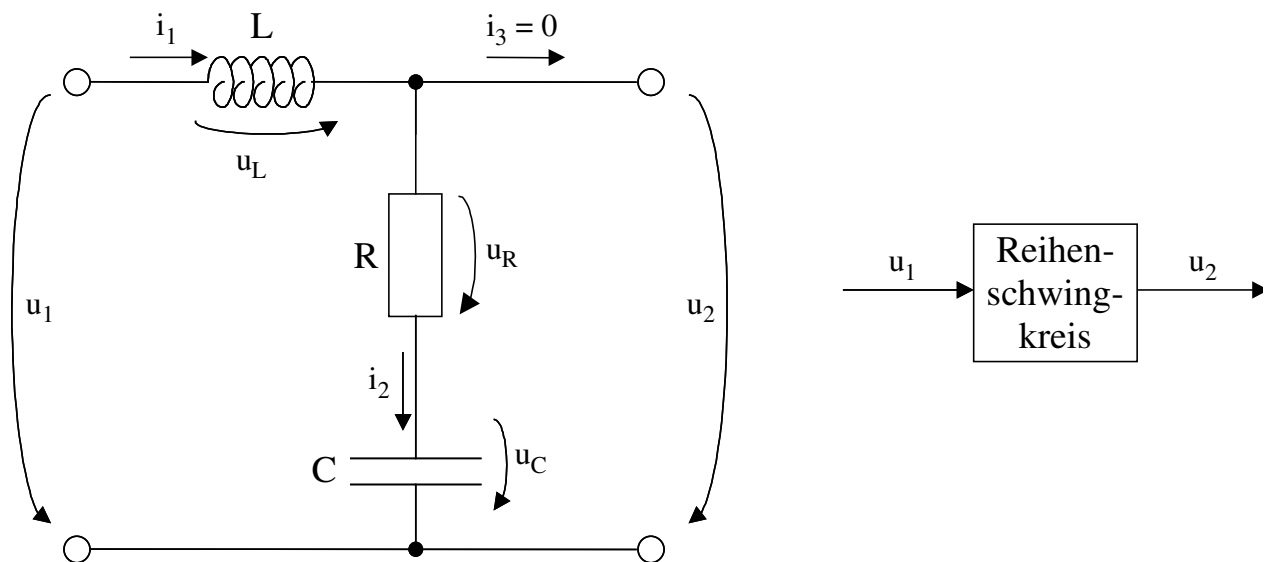


Abb. 11 : Schaltung und Blockschaltbild eines Reihenschwingkreises

Wir gehen von folgendem Anfangszustand aus:

- $i_1(0) = 0$, d.h. zur Zeit $t = 0$ fließt kein Strom durch die Induktivität,
- $u_C(0) = u_o$, d.h. die Kondensatorspannung habe zur Zeit $t = 0$ einen bekannten Wert u_o .

Weitere Annahme : Keine Last am Schwingkreis-Ausgang $\Rightarrow i_3 = 0$ und somit $i_1 = i_2$.

Physikalische Beziehungen:

A) Zusammenhang zwischen Spannung und Strom für die Bauteile („**Bauelemente-Gleichungen**“):

$$u_R = R \cdot i_1 \quad (\text{Ohmsches Gesetz}) \quad (3.1)$$

$$u_L = L \cdot \frac{di_1}{dt} \quad (\text{selbstinduzierte Spannung}) \quad (3.2)$$

$$u_C = \underbrace{u_C(0)}_{\frac{Q(0)}{C}} + \frac{1}{C} \underbrace{\int_0^t i_1(\tau) d\tau}_{\text{Ladungsänderung } \Delta Q} \quad (3.3)$$

B) **Kopplungen zwischen den Bauelementen** mittels Kirchhoffscher Gesetze (Maschenregel: Summe der Spannungen innerhalb einer Masche gleich Null):

$$u_1 = u_L + u_R + u_C \quad (\text{linke Masche}) \quad (3.4)$$

$$u_2 = u_R + u_C \quad (\text{rechte Masche}) \quad (3.5)$$

Ziel: Herleitung einer (Differential-)Gleichung, in der nur noch die Eingangsgröße $u = u_1$ und die Ausgangsgröße $y = u_2$ sowie deren Ableitungen vorkommen (denn wir wollen ja die Funktionalbeziehung zwischen Eingangs- und Ausgangsgröße berechnen).

⇒ zu eliminieren sind u_L , u_R , u_C und i_1 .

Schritt 1: Eliminiere u_L , u_R und u_C mit Hilfe der Bauelementgleichungen.

$$\begin{array}{ccc} (3.4) & & (3.2) \\ \downarrow & & \downarrow \\ u_1 = u_L + \underbrace{u_R + u_C}_{u_2 \text{ laut (3.5)}} & = & L \cdot \frac{di_1}{dt} + u_2 \end{array} \quad (3.6)$$

$$\begin{array}{ccc} (3.5) & & (3.1), (3.3) \\ \downarrow & & \downarrow \\ u_2 = u_L + u_C & = & R i_1 + u_C(0) + \frac{1}{C} \int_0^t i_1(\tau) d\tau \end{array} \quad (3.7)$$

Gl. (3.6) erlaubt uns, $\frac{di_1}{dt}$ und $\frac{d^2 i_1}{dt^2}$ durch u_1 , \dot{u}_1 und u_2 , \dot{u}_2 auszudrücken:

$$\frac{di_1}{dt} = \frac{1}{L} (u_1 - u_2) \quad (3.8)$$

$$\frac{d^2 i_1}{dt^2} = \frac{1}{L} (\dot{u}_1 - \dot{u}_2) \quad \left[\text{mit } \dot{u}_1 := \frac{d}{dt} u_1 \right] \quad (3.9)$$

Wenn es also gelingt, Gl. (3.7) so umzuformen, dass statt

$$i_1 \text{ und } \int_0^t i_1(\tau) d\tau$$

die erste und zweite Ableitung von i_1 vorkommen, sind wir fertig.

Methode: Leite (3.7) zweimal nach t ab.

$$\dot{u}_2 = R \frac{di_1}{dt} + \frac{1}{C} i_1(t) \quad (3.10)$$

$$\ddot{u}_2 = R \frac{d^2 i_1}{dt^2} + \frac{1}{C} \frac{di_1}{dt} \stackrel{(3.8), (3.9)}{=} \frac{R}{L} (\dot{u}_1 - \dot{u}_2) + \frac{1}{LC} (u_1 - u_2) \quad (3.11)$$

Bringe u_2 samt Ableitungen auf linke Gleichungsseite,
 u_1 samt Ableitungen auf rechte Gleichungsseite :

$$\boxed{CL \ddot{u}_2(t) + CR \dot{u}_2(t) + u_2(t) = CR \dot{u}_1(t) + u_1(t)} \quad (3.12)$$

Differentialgleichung 2. Ordnung für $u_2(t)$.

Damit eine Differentialgleichung 2. Ordnung **eindeutig** lösbar ist, müssen die Anfangswerte
(= Werte zum Zeitpunkt $t = 0$) von $u_2(t)$ und $\dot{u}_2(t)$ bekannt sein.

$$\text{Gl. (3.7): } u_2(t) = R \underbrace{i_1(t)}_{i_1(0)=0} + \underbrace{u_C(0)}_{u_0}$$

$$\Rightarrow \boxed{u_2(0) = u_0} \quad (3.13)$$

$$\text{Gl. (3.10): } \dot{u}_2(t) = R \underbrace{\frac{di_1}{dt}}_{i_1(0)=0} + \frac{1}{C} \underbrace{i_1(t)}_{i_1(0)=0}$$

laut (3.8):

$$\frac{1}{L} (u_1 - u_2)$$

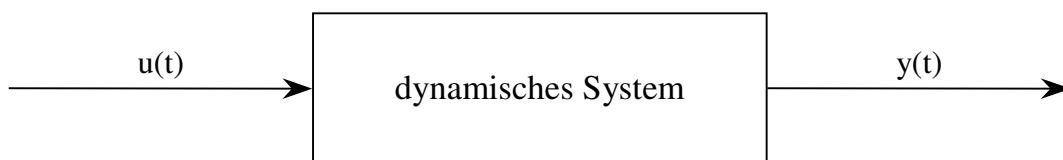
$$\Rightarrow \dot{u}_2(0) = \frac{R}{L} (u_1(0) - u_2(0))$$

Das Eingangssignal $u_1(t)$ soll – wie stets bei Eingangssignalen – für alle Zeiten $t < 0$ verschwinden. Wenn u_1 als stetig angenommen wird, gilt damit auch $u_1(0) = 0$.

$$\Rightarrow \dot{u}_2(0) = -\frac{R}{L} u_2(0) \stackrel{(3.13)}{=} -\frac{R}{L} u_0 . \quad (3.14)$$

3.1.2 Lineare Differentialgleichung n-ter Ordnung

Da die physikalischen Grundgesetze häufig zeitliche Ableitungen enthalten, führt die mathematische Modellierung eines dynamischen Systems oft „ganz organisch“ auf eine Differentialgleichung (siehe obiges Beispiel).



Die Differentialgleichung (kurz: DGL) beschreibt den dynamischen Zusammenhang zwischen der Eingangsgröße $u(t)$ und der Ausgangsgröße $y(t)$, also die Übertragungseigenschaften des dynamischen Systems. (Wie aus der DGL das explizite Eingangs-Ausgangs-Verhalten des Systems bestimmt werden kann, ist Thema von Abschnitt 3.3.)

Allgemeine Form der DGL für ein lineares dynamisches System:

$$\begin{aligned} a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y(t) &= \\ &= b_q \frac{d^q u}{dt^q} + b_{q-1} \frac{d^{q-1} u}{dt^{q-1}} + \dots + b_1 \frac{du}{dt} + b_0 u(t) \end{aligned} \quad (3.15)$$

Gl. (3.15) ist eine lineare DGL n-ter Ordnung. Die a_i und b_i sind reellwertige Koeffizienten, die aus den physikalischen Parametern des betrachteten Systems berechnet werden können. Für eine gegebene Eingangsgröße $u(t)$, $t \geq 0$, hat die DGL (3.15) genau dann eine eindeutige Lösung $y(t)$, $t \geq 0$, wenn die n Anfangswerte (= Werte zum Zeitpunkt $t = 0$)

$$y(0) = y_{01}, \quad \dot{y}(0) = y_{02}, \quad \dots, \quad \frac{d^{n-1} y}{dt^{n-1}}(0) = y_{0n} \quad (3.16)$$

gegeben sind.

Bemerkungen:

- Schreibweise: $\frac{du}{dt} =: \dot{u}$, $\frac{d^2u}{dt^2} =: \ddot{u}$
- Häufig wird die DGL (3.15) so umgeformt, dass $a_n = 1$ gilt.
- Es wird angenommen, dass für die Grade der höchsten Ableitungen von y und u die Beziehung $q \leq n$ gilt, weil nur Systeme, die diese Bedingung erfüllen, technisch realisierbar sind.
- Da wir uns nur für die **zukünftige** Bewegung, also für $y(t)$ für $t \geq 0$ (oder allgemeiner $t \geq t_0$) interessieren, muss auch die Eingangsgröße $u(t)$ nur für $t \geq 0$ bekannt sein. Die gesamte Wirkung, die die Vergangenheit der Eingangsgröße (also $u(t)$ für $t < 0$) auf das System ausübt, ist in den Anfangswerten (3.16) der DGL enthalten. Ein durch eine DGL beschreibbares dynamisches System ist also quasi gedächtnislos. Aus diesem Grunde werden im folgenden nur Funktionen der Zeit t betrachtet, die für $t < 0$ konstant den Wert Null annehmen („verschwinden“). Diese Verabredung gilt für die Funktionen $u(t)$ und $y(t)$ wie für alle später eingeführten Signale.

Analog zum Schwingkreis-Beispiel besteht die Modellierung eines linearen dynamischen Systems durch eine DGL der Form (3.15) i.a. aus drei Schritten:

- A) Aufstellen der Gleichungen, die aus den physikalischen Grundgesetzen folgen („Bauelementgleichungen“).
- B) Aufstellen der Gleichungen, die die Kopplungen zwischen den „Bauelementen“ beschreiben.
- C) Umformen und Verknüpfen dieser Gleichungen zur DGL vom Typ (3.15) (sprich: Außer dem Eingangs- und dem Ausgangssignal sowie deren Ableitungen sind alle Signale zu eliminieren).

Nur Schritt C hängt davon ab, welches Signal als Eingangsgröße und welches als Ausgangsgröße betrachtet wird. So könnte z.B. beim Reihenschwingkreis anstelle der Spannung $u_2(t)$ auch der Strom $i_2(t)$ als Ausgangsgröße angesehen werden. Bei der Modellierung würden dann dieselben Bauelementebeschreibungen und Maschengleichungen verwendet, diese jedoch zu einer DGL umgeformt, in die $u_1(t)$ als Eingangsgröße und $i_2(t)$ [anstelle von $u_2(t)$] als Ausgangsgröße eingeht.

Welche physikalischen Gesetzmäßigkeiten für die Schritte A und B heranzuziehen sind, hängt vom Charakter des zu modellierenden Systems ab.

- Elektrotechnische Systeme, die sich als Schaltbild beschreiben lassen, werden durch die Kirchhoffschen Gesetze modelliert.

- Verfahrenstechnische Anlagen werden zunächst in Bilanzräume unterteilt. Für diese Räume werden dann Energie- und Stoffbilanzen aufgestellt. Die Koppelgleichungen für die Bilanzräume ergeben sich aus dem Energie- bzw. Stoffaustausch zwischen diesen Räumen.
- Mechanische Systeme werden durch Kräfte- und Drehmoment-Gleichungen beschrieben.

3.1.3 Linearität dynamischer Systeme

Wir nehmen an, dass ein durch Gl. (3.15) beschriebenes dynamisches System auf die Eingangsgröße $u_1(t)$ mit der Ausgangsgröße $y_1(t)$ antwortet. $y_1(t)$ ist also die Lösung der DGL (3.15), wenn $u = u_1(t)$ gesetzt wird.

Suggestive Schreibweise hierfür: $u_1(t) \mapsto y_1(t)$.

Gelte weiterhin: $u_2(t) \mapsto y_2(t)$,

d.h. $y_2(t)$ ist die Lösung von (3.15), wenn $u = u_2(t)$ gesetzt wird.

Wenn wir nun als Eingangssignal eine beliebige Linearkombination von $u_1(t)$ und $u_2(t)$ vorgeben, also

$$u(t) = k \cdot u_1(t) + l \cdot u_2(t), \quad k, l \in \mathbb{R}, \quad (3.17)$$

so ergibt sich am Ausgang des Systems die entsprechende Linearkombination von $y_1(t)$ und $y_2(t)$:

Aus $u_1(t) \mapsto y_1(t)$ und $u_2(t) \mapsto y_2(t)$	(3.18a)
folgt : $\underbrace{k \cdot u_1(t) + l \cdot u_2(t)}_{u(t)} \mapsto \underbrace{k \cdot y_1(t) + l \cdot y_2(t)}_{y(t)}$	(3.18b)

Dynamische Systeme mit dieser Eigenschaft heißen **linear**.

Weil sich bei linearen Systemen die Wirkungen zweier linear überlagerter Eingangssignale am Ausgang des Systems in gleicher Weise linear überlagern („superponieren“), sagt man, dass lineare dynamische Systeme dem ***Superpositionsprinzip*** gehorchen.

Die Linearität dynamischer Systeme, die durch Gl. (3.15) beschrieben werden, ist offensichtlich eine dieser Differentialgleichung innewohnende Eigenschaft und kann mit Hilfe der Differentialgleichung leicht nachgewiesen werden.

Die Linearität (3.18) ist gezeigt, wenn wir nachgewiesen haben, dass $y(t) = ky_1(t) + ly_2(t)$ die Lösung der DGL (3.15) für $u(t) = ku_1(t) + lu_2(t)$ darstellt.

Hierfür setzen wir einfach auf der linken Seite von (3.15) $ky_1 + ly_2$ ein:

$$\begin{aligned}
 & a_n \frac{d^n \overbrace{(ky_1 + ly_2)}^y}{dt^n} + \dots + a_0 \overbrace{(ky_1 + ly_2)}^y = \\
 & \quad \text{(Linearität der Ableitung)} \\
 & \quad = \underbrace{k \left\{ a_n \frac{d^n y_1}{dt^n} + \dots + a_0 y_1 \right\}}_{\text{wegen (3.18a): } b_q \frac{d^q u_1}{dt^q} + \dots + b_0 u_1} + \underbrace{l \left\{ a_n \frac{d^n y_2}{dt^n} + \dots + a_0 y_2 \right\}}_{b_q \frac{d^q u_2}{dt^q} + \dots + b_0 u_2} \\
 & \quad \text{(Linearität der Ableitung)} \\
 & \quad = b_q \frac{d^q (ku_1 + lu_2)}{dt^q} + \dots + b_0 (ku_1 + lu_2) \\
 & \quad = b_q \frac{d^q u}{dt^q} + \dots + b_0 u \quad (= \text{rechte Seite von (3.15)}), \text{ was zu zeigen war.}
 \end{aligned}$$

Bemerkungen:

- 1) Damit das dynamische System wirklich linear ist, muss Gl. (3.18b), also $y(t) = ky_1(t) + ly_2(t)$, natürlich für alle Zeiten $t \geq 0$ gelten, insbesondere auch für $t = 0$.

Die Linearitätseigenschaft gilt also nur unter der Bedingung, dass die Anfangswerte des Signals y durch entsprechende Überlagerung aus den Anfangswerten der Signale y_1 und y_2 hervorgehen:

$$\frac{d^i y}{dt^i}(0) = k \frac{d^i y_1}{dt^i}(0) + l \frac{d^i y_2}{dt^i}(0), \quad i = 0, 1, \dots, n-1 \quad (3.19)$$

Diese Bedingung ist insbesondere dann erfüllt, wenn sich das System zur Zeit $t = 0$ in der Ruhelage befindet und folglich alle Anfangswerte verschwinden.

- 2) Die Linearität dynamischer Systeme bezieht sich auf die Eingangs-Ausgangs-Beziehung.

Die Mathematik gewöhnlicher Differentialgleichungen kennt noch einen anderen Linearitätsbegriff:

Gewöhnliche Differentialgleichungen heißen linear, wenn sie folgende Form haben:

$$f_n(t) \frac{d^n y(t)}{dt^n} + f_{n-1}(t) \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + f_0(t) y(t) + f(t) = 0, \quad (3.20)$$

wobei $f(t), f_0(t), \dots, f_n(t)$ gegebene Funktionen sind.

Konsequenz der Linearität im letztgenannten Sinne (Näheres siehe Abschnitt 3.3):

Sämtliche Lösungen der zu Gl. (3.20) gehörigen homogenen Gleichung – die aus (3.20) durch Nullsetzen von $f(t)$ entsteht – bilden einen linearen Raum (= Vektorraum). (Legt man zusätzlich zu (3.20) die n Anfangsbedingungen fest, so wählt man aus diesem Lösungsraum genau eine Lösung aus.) Hat man eine (beliebige) Lösung der inhomogenen (= vollen) Gl. (3.20) gefunden, so erhält man daraus den gesamten Lösungsraum der inhomogenen Gl. (3.20) – wenn keine zusätzlichen Anfangsbedingungen vorgegeben sind –, indem man zu den Lösungen der homogenen Gleichung jeweils diese spezielle inhomogene Lösung addiert.

Die DGL (3.15) für ein lineares dynamisches System ist ein Spezialfall einer linearen DGL (im Sinne der Mathematik gewöhnlicher Differentialgleichungen), denn die Konstanten a_0, a_1, \dots, a_n in (3.15) spielen die Rolle der Funktionen f_0, f_1, \dots, f_n in (3.20), und die rechte Gleichungsseite von (3.15) spielt die Rolle der Funktion $-f$.

Bei linearen dynamischen Systemen muss die Funktion f aus Gl. (3.20) eine spezielle Gestalt besitzen

$$\left(\text{nämlich die einer Linearkombination von } u, \dot{u}, \dots, \frac{d^q u}{dt^q} \right)$$

und darf insbesondere keine Konstante enthalten. Eine solche würde die Eingangs-Ausgangs-Linearität des dynamischen Systems zerstören.

- 3) Eingangs-Ausgangs-Linearität ist eine fundamentale Eigenschaft für die Beschreibung dynamischer Systeme im Frequenzbereich. Sie erlaubt es, beliebige Signale $u(t)$ zunächst in ihre sinusförmigen Bestandteile zu zerlegen, die Übertragung der einzelnen sinusförmigen Komponenten durch das lineare dynamische System getrennt voneinander zu untersuchen, und schließlich die übertragenen Signalkomponenten durch lineare Superposition wieder zur Gesamt-Systemantwort zusammenzusetzen (siehe Kapitel 4).

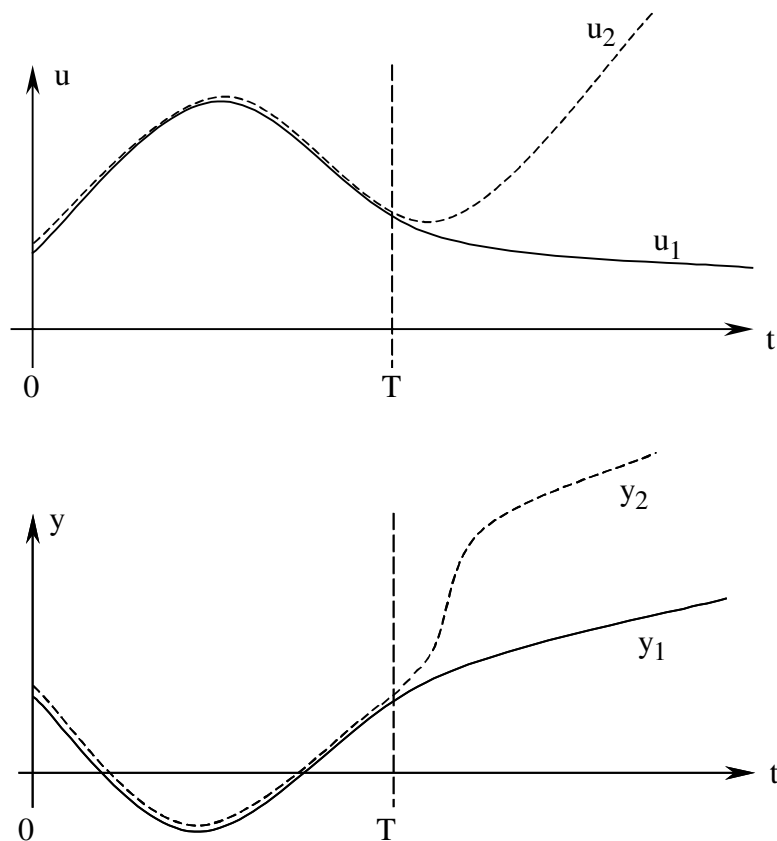
3.1.4 Kausalität dynamischer Systeme

Eine wichtige Eigenschaft dynamischer Systeme ist ihre Kausalität. Die Werte des Eingangssignals können das Verhalten des Systems nur für zukünftige Zeitpunkte beeinflussen.

Für ein dynamisches System, das reale physikalische Vorgänge korrekt modelliert, ist Kausalität „automatisch“ erfüllt. Beim Reglerentwurf können jedoch Reglergesetze konstruiert werden, die die Kausalität verletzen. Solche Reglergesetze sind technisch nicht realisierbar. Wir müssen also imstande sein, zu beurteilen, ob ein System kausal ist. Hierfür müssen wir zunächst Kausalität präzise definieren.

Wir betrachten das Eingangs-Ausgangs-Verhalten (kurz: E/A-Verhalten) des Systems. Da das Ausgangssignal y und damit auch das E/A-Verhalten von den Anfangswerten abhängt, setzen wir bei unserer Betrachtung voraus, dass alle Anfangswerte der DGL (3.15) verschwinden. Ein System heißt kausal, wenn es zwei beliebige Eingangssignale u_1 und u_2 , die im Zeitintervall $[0, T]$ (T beliebig) übereinstimmen, in Ausgangssignale y_1 und y_2 verwandelt, die ebenfalls im Zeitintervall $[0, T]$ übereinstimmen:

<p>Aus $u_1(t) = u_2(t)$ für $t \in [0, T]$ folgt: $y_1(t) = y_2(t)$ für $t \in [0, T]$</p>	(3.21)
--	--------



Abweichungen im Eingangssignal ab dem Zeitpunkt T dürfen also bei einem kausalen System keine Abweichungen im Ausgangssignal **vor** diesem Zeitpunkt T bewirken.

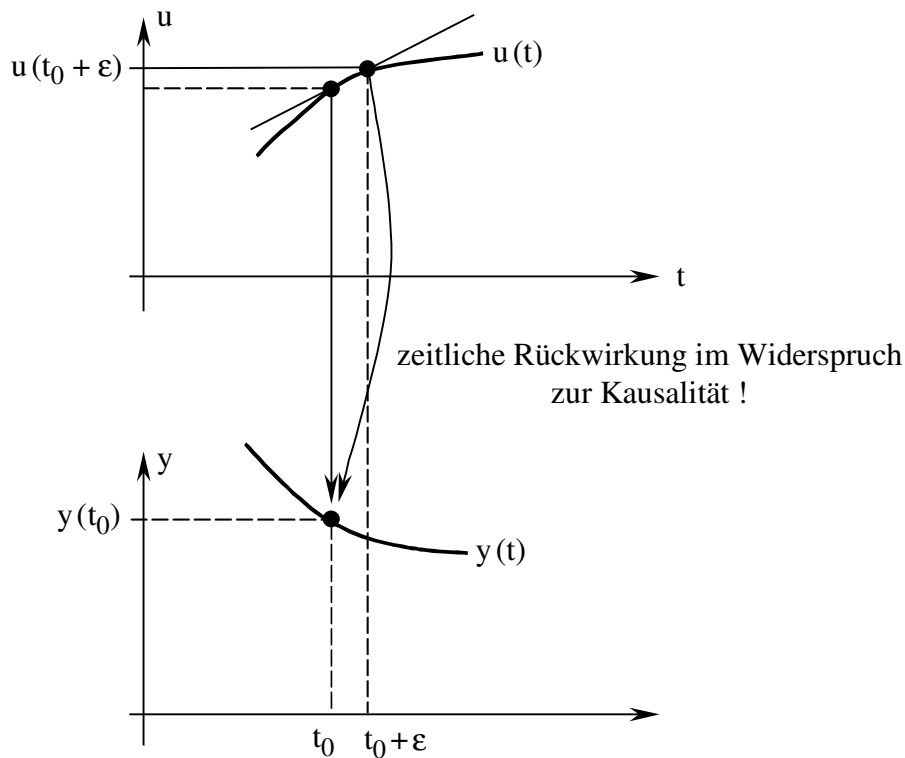
Ein lineares dynamisches System, das in Form einer DGL vom Typ (3.15) gegeben ist, erfüllt genau dann die Kausalität, wenn

$$q \leq n \quad (3.22)$$

gilt, d.h. wenn die Eingangsgröße u nicht mit höherer Ableitung als die Ausgangsgröße y erscheint. Wäre diese Bedingung verletzt, so müßte das System verzögerungsfreie Differentia-

tionen der Eingangsgröße ausführen, die die Kausalität verletzen würden und daher physikalisch nicht realisierbar sind.

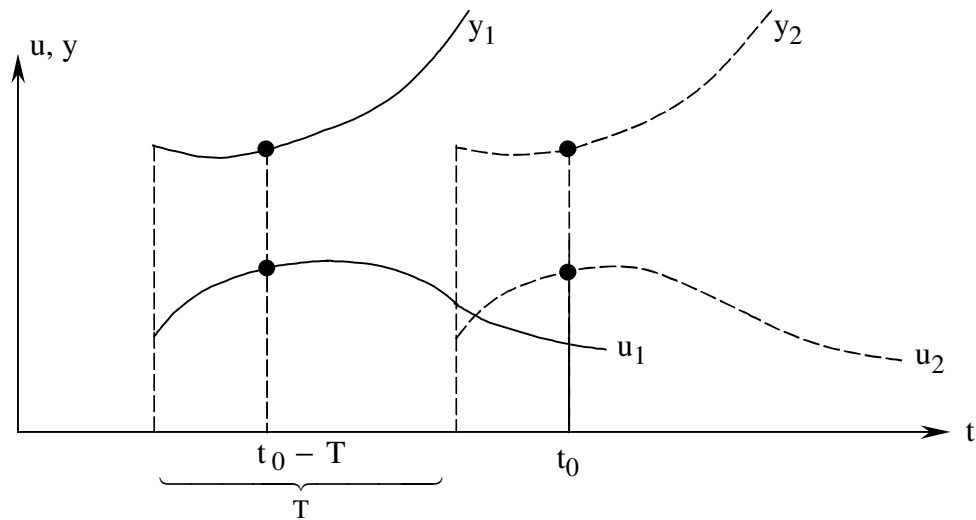
Um uns dies plausibel zu machen, betrachten wir ein fiktives System, das der Gl. $y(t) = b_1 \cdot \dot{u}(t)$ gehorcht. $\dot{u}(t)$ enthält Information über das Verhalten der Funktion u in der nahen Zukunft (siehe Taylor-Entwicklung!).



Wenn also $\dot{u}(t_0)$ den Ausgangs-Signalwert $y(t_0)$ bestimmen würde, so wäre $y(t_0)$ damit – im Widerspruch zur Kausalität – von zukünftigen Werten $u(t_0 + \epsilon)$ der Eingangsgröße u abhängig. Ein solches fiktives System kann es also real nicht geben.

3.1.5 Zeitinvarianz

Die Koeffizienten unserer DGL (3.15) sind Konstante, also von der Zeit unabhängig. Durch (3.15) beschriebene dynamische Systeme sind folglich **zeitinvariant**. Sie reagieren auf ein Eingangssignal (also einen vorgegebenen Verlauf von $u(t)$) unabhängig davon, wann dieses Signal eintrifft. Verschiebt man also ein Eingangssignal auf der Zeitachse um T nach rechts, so verschiebt sich auch das Ausgangssignal um T nach rechts, bleibt aber in der Form gleich. (Hierbei wieder vorausgesetzt, dass alle Anfangswerte verschwinden.)



<p>Zeitinvarianz : Aus $u_2(t) = u_1(t - T)$</p> <p>folgt: $y_2(t) = y_1(t - T)$</p>
--

(3.23)

Die Modellierung eines Systems als zeitinvariant ist dann gerechtfertigt, wenn Parameterdrift oder Veränderungen in der Systemstruktur entweder ganz vernachlässigt werden können oder sich – gemessen an der Zeitskala der Systemdynamik – sehr langsam vollziehen.

3.2 Zustandsraumdarstellung linearer Systeme

1.1.1 3.2.1 Beispiel: Zustandsraummodell für einen Reihenschwingkreis

Der Reihenschwingkreis aus Abb. 11 lässt sich statt durch eine DGL zweiter Ordnung auch mühelos durch zwei Differentialgleichungen erster Ordnung modellieren. Als erste Gleichung wird Gl. (3.8) übernommen,

$$\frac{di_1}{dt} = \frac{1}{L}(u_1 - u_2), \quad (3.24)$$

als zweite Gleichung dient (3.10) :

$$\begin{aligned} \dot{u}_2 &= R \frac{di_1}{dt} + \frac{1}{C} i_1 \\ (3.24) \\ &= \frac{R}{L} u_1 - \frac{R}{L} u_2 + \frac{1}{C} i_1 \end{aligned}$$

Umordnen der rechten Gleichungsseiten so, dass die Summanden mit der Eingangsgröße u_1 ganz rechts stehen, führt zu folgendem Differentialgleichungssystem erster Ordnung:

$$\begin{aligned} \frac{di_1}{dt} &= -\frac{1}{L} u_2 + \frac{1}{L} u_1 \\ \frac{du_2}{dt} &= \frac{1}{C} i_1 - \frac{R}{L} u_2 + \frac{R}{L} u_1 \end{aligned}$$

Führt man den Vektor

$$\mathbf{x} := \begin{pmatrix} i_1 \\ u_2 \end{pmatrix} \quad (3.25)$$

ein, so lässt sich obiges Gleichungssystem übersichtlich in Matrixschreibweise darstellen:

$$\begin{pmatrix} \frac{di_1}{dt} \\ \frac{du_2}{dt} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{L} \\ \frac{1}{C} & -\frac{R}{L} \end{pmatrix} \begin{pmatrix} i_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{L} \\ \frac{R}{L} \end{pmatrix} u_1 \quad (3.26)$$

Gibt man für beide Komponenten von \mathbf{x} Anfangswerte vor (laut der im Beispiel gemachten Annahmen $i_1(0) = 0$ und $u_2(0) = u_0$), so hat das Differentialgleichungssystem erster Ordnung (3.26) eine eindeutige Lösung.

Die in $\mathbf{x}(0)$ enthaltenen Informationen reichen also aus, um für eine beliebig gegebene Eingangsgröße $u_1(t)$, $t \geq 0$, die Ausgangsgröße $u_2(t)$ für $t \geq 0$ eindeutig zu berechnen. $\mathbf{x}(0)$ beschreibt also den Zustand, in dem sich das System zum Zeitpunkt $t = 0$ befindet. Diese Eigenschaft gilt für jeden beliebigen Zeitpunkt \bar{t} . Das heißt, wenn $\mathbf{x}(\bar{t})$ bekannt und die Eingangsgröße $u_1(t)$ für $t \geq \bar{t}$ gegeben ist, so kann mit dem Modell (3.26) die Ausgangsgröße $u_2(t)$ für alle Zeiten $t \geq \bar{t}$ eindeutig berechnet werden. $\mathbf{x}(t)$ wird deshalb als **Zustand** des Systems zum Zeitpunkt t bezeichnet.

Die allgemeine Zustandsdefinition greift genau diese Eigenschaften auf:

Ein Vektor $\mathbf{x}(t)$ (genauer: vektorwertige Funktion der Zeit) heißt Zustandsvektor oder kurz **Zustand eines dynamischen Systems**, wenn die Werte $x_i(0)$ der Komponenten von \mathbf{x} zum Zeitpunkt 0 zusammen mit dem Verlauf der Eingangsgröße $u(\tau)$ für das Zeitintervall $0 \leq \tau \leq t$ die Ausgangsgröße $y(t)$ zur Zeit $t \geq 0$ eindeutig bestimmen. Die Komponenten $x_i(t)$ heißen **Zustandsvariable** oder **Zustandsgrößen**.

3.2.2 Allgemeines Zustandsraummodell

Gl. (3.26) für den Reihenschwingkreis ist von der allgemeinen Form

$$\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{b} u(t), \quad (3.27)$$

wobei

$$\mathbf{x}(t) := \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

den n -dimensionalen Zustandsvektor,

$$\mathbf{A} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & \\ \vdots & & & \\ a_{n1} & & \dots & a_{nn} \end{pmatrix}$$

eine konstante (n, n) -Matrix (genannt: **Systemmatrix**), und

$$\mathbf{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

einen konstanten n -dimensionalen Spaltenvektor bezeichnen.

In (3.27) taucht die Ausgangsgröße $y(t)$ nicht explizit auf. Im Schwingkreis-Beispiel war sie gleich der zweiten Komponente des Zustandsvektors \mathbf{x} , d.h. es galt

$$y(t) = \begin{pmatrix} 0 & 1 \end{pmatrix} \mathbf{x} \quad [\text{Verknüpfung rechts: Matrixmultiplikation}] .$$

Im allgemeinen ist die Ausgangsgröße eine Linearkombination der Zustandsgrößen x_i und der Eingangsgröße u und kann deshalb in der Form

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

geschrieben werden, wobei

$$\mathbf{c}^T := \begin{pmatrix} c_1 & \dots & c_n \end{pmatrix}$$

ein konstanter n -dimensionaler Zeilenvektor und d ein konstanter Skalar ist.

Nimmt man noch die Anfangsbedingungen für alle Zustandsgrößen x_i hinzu (in Vektorschreibweise: $\mathbf{x}(0) = \mathbf{x}_0$, \mathbf{x}_0 : vorgegebener n -dimensionaler Vektor der Anfangswerte), so erhält man das allgemeine **Zustandsraummodell**:

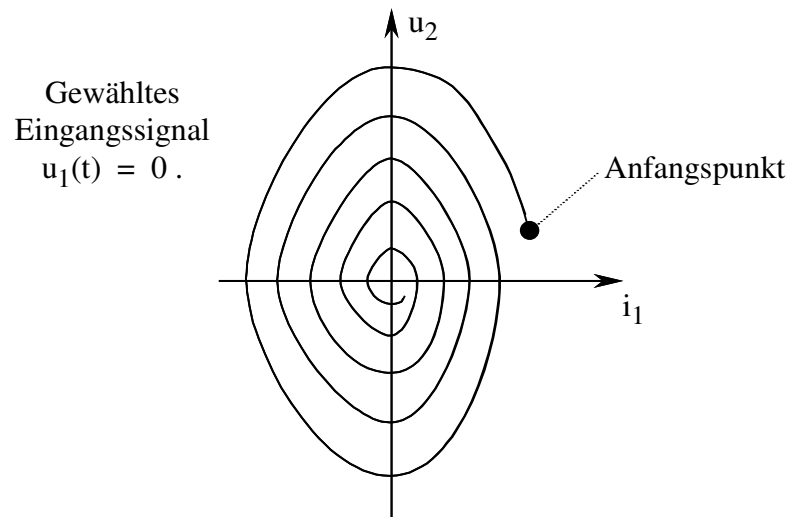
Zustandsgleichung: $\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{b} u(t), \mathbf{x}(0) = \mathbf{x}_0$	(3.28)
Ausgabegleichung: $y(t) = \mathbf{c}^T \mathbf{x}(t) + d u(t)$	(3.29)

Komponentenweise betrachtet, ist die Zustandsgleichung ein System von Differentialgleichungen erster Ordnung. Da jede lineare DGL n -ter Ordnung, also auch das Differentialgleichungsmodell (3.15) eines linearen dynamischen Systems, äquivalenterweise als System von n Differentialgleichungen 1. Ordnung dargestellt werden kann, ist das Zustandsraummodell eine zur DGL (3.15) äquivalente mathematische Beschreibung des gegebenen Systems. Da die Einführung des Systemzustandes die Analyse dynamischer Systeme erleichtert, wird in der Regelungstechnik meist von einem Zustandsraummodell ausgegangen.

Die Zustandsvariablen sind in der Regel physikalische Größen [Beispiele: Ströme und Spannungen (bei elektrischen Systemen), Auslenkungen, Geschwindigkeiten und Beschleunigungen (bei mechanischen Systemen)], die das Verhalten von Speicherelementen, wie z.B. Kapazitäten, Induktivitäten, Massen oder Federn, beschreiben. Die Systemordnung n stimmt deshalb in der Regel mit der Zahl der im System enthaltenen Speicherelemente überein. (Im Schwingkreis-Beispiel: 2 Speicherelemente, nämlich Kapazität und Induktivität, \rightarrow Ordnung 2.) Es können aber auch Größen als Zustandsvariable verwendet werden, die physikalisch nicht interpretierbar sind (siehe Abschnitt 3.2.4).

Die zeitliche Abhängigkeit des n -dimensionalen Vektors \mathbf{x} kann man sich als Bewegung eines Punktes im n -dimensionalen Vektorraum \mathbb{R}^n vorstellen. \mathbb{R}^n wird daher als **Zustandsraum** bezeichnet. Die vom Zustand \mathbf{x} mit der Zeit durchlaufene Bahnkurve heißt **Trajektorie** oder **Zustandskurve** des Systems.

Beispiel : Trajektorie des Schwingkreises im Zustandsraum, also in der i_1/u_2 -Ebene:



Diese Darstellung zeigt, welche Werte die Zustandsgrößen i_1 und u_2 gleichzeitig annehmen. Die Zeit ist jedoch nicht mehr explizit zu erkennen. Dass die Zustandstrajektorie eine Spirale ist, ist typisch für schwingende Systeme.

3.2.3 Graphische Veranschaulichung durch ein Simulationsdiagramm

Die durch Zustands- und Ausgabegleichung beschriebenen Zusammenhänge zwischen den Signalen $u(t)$, $x_i(t)$ und $y(t)$ lassen sich durch das folgende mathematische Blockschaltbild veranschaulichen:

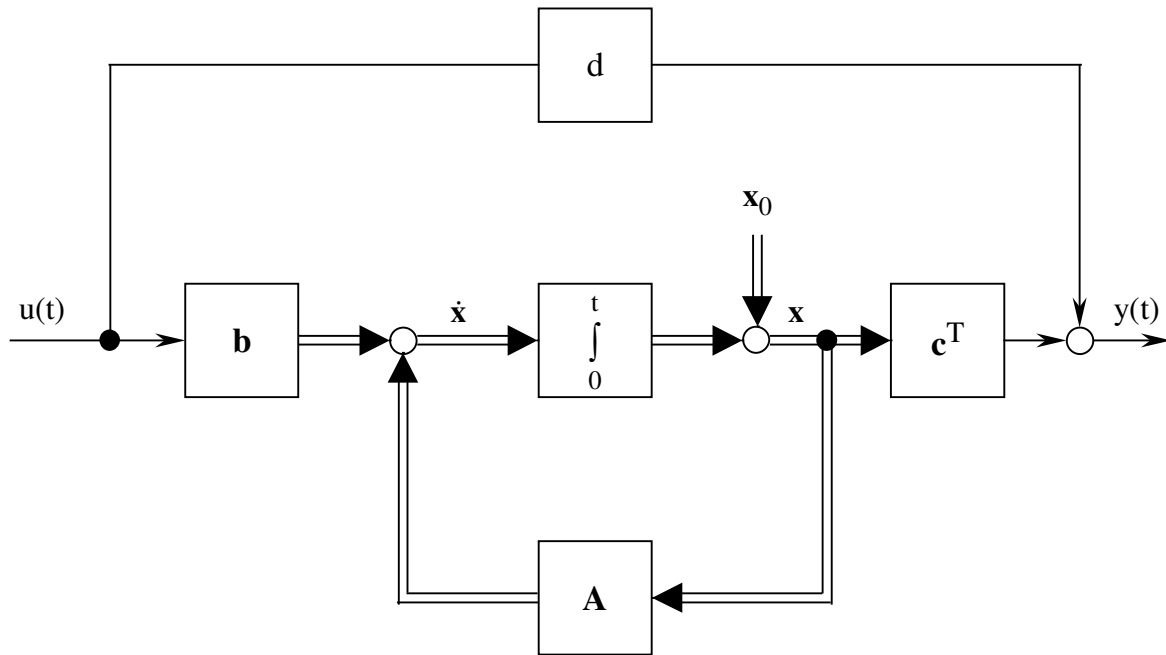
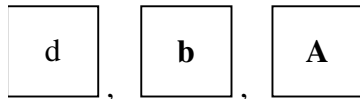


Abb. 12 : Blockschaltbild des Zustandsraummodells

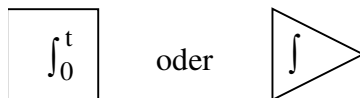
Symbolkonvention:

Doppelpfeil \Rightarrow

vektorielle Größe



Statisches Übertragungsglied / Multiplizierer mit skalarem, vektor- bzw. matrixwertigem Übertragungsfaktor d , b bzw. A



Integrator: Eingangssignal wird integriert (ggf. komponentenweise)

Wenn man die vektoriell zusammengefassten Signale separat als Einzelsignale darstellt, so zeigt das resultierende Diagramm die Kopplungsstruktur der n einzelnen Differentialgleichungen, aus denen die Zustandsgleichung (3.28) besteht.

Beispiel : Gegeben sei das System zweiter Ordnung

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} u$$

$$y = (c_1 \quad c_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + d \cdot u$$

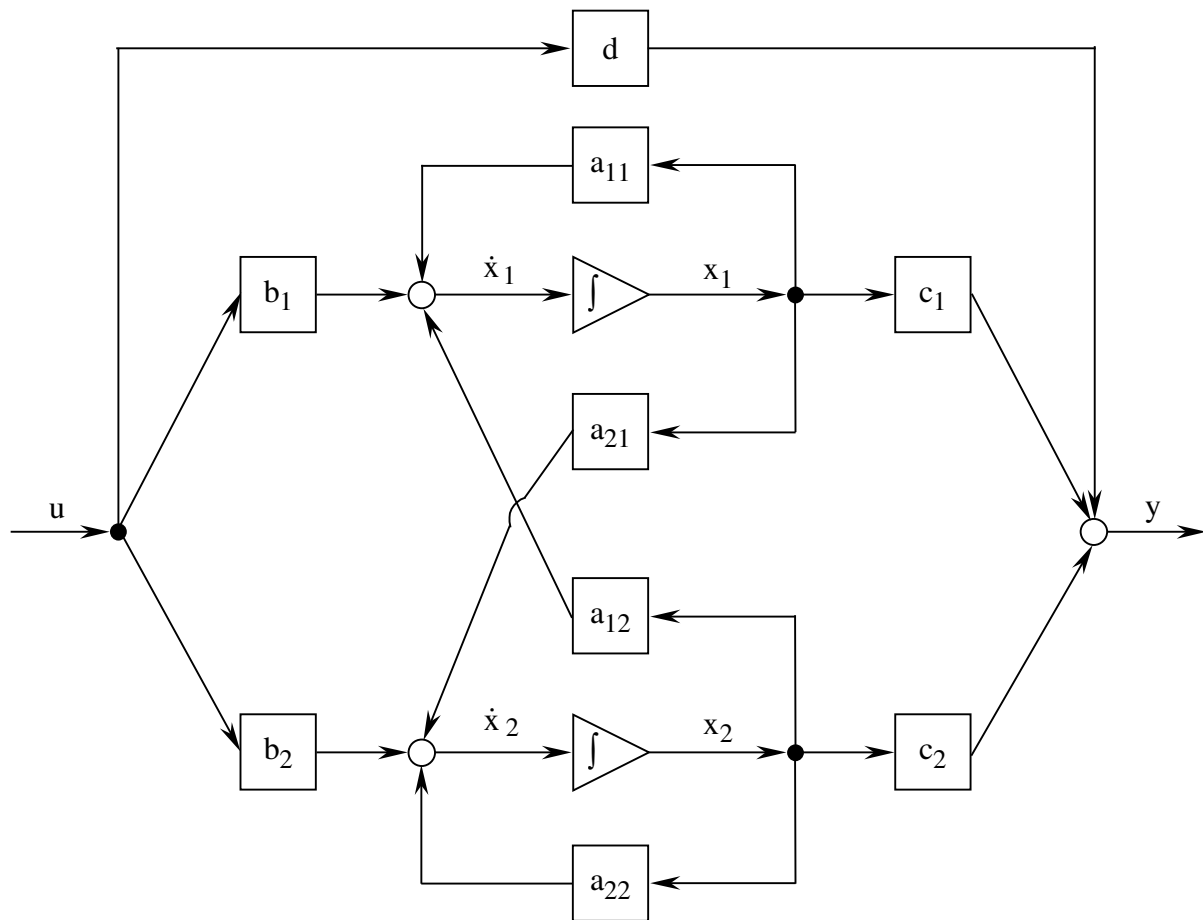


Abb. 13 : Simulationsdiagramm eines Systems zweiter Ordnung

Derart verfeinerte Blockschaltbilder heißen oft auch Signalflussdiagramme, dynamische Diagramme oder, weil sie sich als Grundlage für die rechnergestützte Simulation des Systemverhaltens eignen, Simulationsdiagramme. Beispielsweise ist im Simulations-Programmpaket Matlab/Simulink das zu simulierende Modell in Form eines dynamischen Diagramms einzugeben.

3.2.4 Ableitung des Zustandsraummodells aus der Differentialgleichung

Wir geben nun einen Weg an, um aus der DGL (3.15) ein Zustandsraummodell abzuleiten. Dabei gehen wir zunächst von einer DGL aus, in der keine Ableitungen der Eingangsgröße vorkommen (d.h. $q = 0$) und die bereits so umgeformt ist, dass $a_n = 1$ gilt:

$$\frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \dot{y} + a_0 y(t) = u(t) \quad (3.30)$$

$\uparrow \qquad \qquad \qquad \uparrow \qquad \qquad \uparrow$

werden als Zustandsvariable verwendet, d.h.

$$\mathbf{x}(t) := \begin{pmatrix} y \\ \dot{y} \\ \ddot{y} \\ \vdots \\ \frac{d^{n-1}y}{dt^{n-1}} \end{pmatrix} \quad (3.31)$$

Für die ersten (n-1) Komponenten gilt per definitionem von \mathbf{x} : Die Ableitung ergibt gerade die nächsthöhere Komponente, also

$$\begin{aligned} \dot{x}_1 &= \dot{y} = x_2 \\ \dot{x}_2 &= \ddot{y} = x_3 \\ &\vdots \\ \dot{x}_{n-1} &= \frac{d^{n-1}y}{dt^{n-1}} = x_n \end{aligned}$$

Für \dot{x}_n erhält man aus Gl. (3.30):

$$\begin{aligned} \dot{x}_n &\stackrel{(3.31)}{=} \frac{d^n y}{dt^n} \\ &\stackrel{(3.30)}{=} -a_{n-1} \frac{d^{n-1}y}{dt^{n-1}} - \dots - a_1 \dot{y} - a_0 y + u \\ &\stackrel{(3.31)}{=} -a_{n-1} x_n - \dots - a_1 x_2 - a_0 x_1 + u \end{aligned}$$

Fasst man diese Gleichungen vektoriell zusammen, so resultiert ein Zustandsraummodell (3.28), (3.29) mit

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \\ & & & & 0 & 1 \\ -a_0 & -a_1 & \dots & & & -a_{n-1} \end{pmatrix} \quad (3.32)$$

$$\mathbf{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (3.33)$$

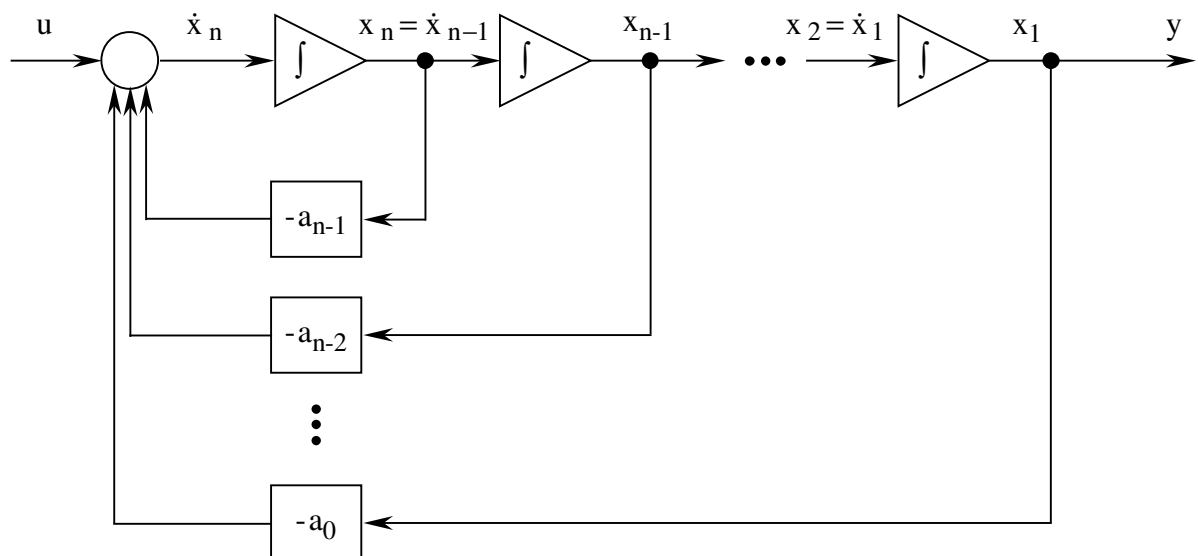
$$\mathbf{c}^T = (1 \quad 0 \quad \dots \quad 0) \quad (3.34)$$

$$d = 0 \quad (3.35)$$

Eine Matrix \mathbf{A} der Form (3.32) wird als *Begleitmatrix* oder *Frobeniusmatrix* bezeichnet. Die Anfangswerte (3.16) der DGL lassen sich aufgrund der Konstruktion des Zustands direkt zum Vektor \mathbf{x}_0 zusammenfassen:

$$\mathbf{x}(0) = \begin{pmatrix} y_{01} \\ y_{02} \\ \vdots \\ y_{0n} \end{pmatrix} \quad (3.36)$$

Das Signalflussdiagramm dieses Zustandsraummodells zeigt, dass es aufgrund der speziellen Struktur von \mathbf{A} und \mathbf{b} nur wenige direkte Kopplungen zwischen den Signalen gibt. Der Eingang u wirkt nur auf \dot{x}_n direkt. Der Ausgang ist gleich der Zustandsgröße x_1 .



Die Verallgemeinerung auf DGL mit $q \neq 0$, d.h. mit Ableitungen der Eingangsgröße u , wird ermöglicht durch folgende Eigenschaft:

Wenn man in der DGL (3.30) anstelle der Funktion $u(t)$ deren Ableitung $\dot{u}(t)$ als äußere Erregung verwendet, so erhält man als Lösung dieser neuen DGL die Ableitung der für u erhaltenen Lösung y (vorausgesetzt: verschwindende Anfangswerte für $y, \dot{y} \dots$).

$$\text{Kurz: } u(t) \mapsto y(t) \Rightarrow \dot{u}(t) \mapsto \dot{y}(t) \quad (3.37)$$

Begründung:

Angenommen, wir kennen die Lösung $y(t)$ für

$$\frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_0 y(t) = u(t), \quad (3.38)$$

$$\text{wobei } \frac{d^{n-1} y}{dt^{n-1}}(0) = 0, \dots, y(0) = 0. \quad (3.39)$$

Leiten wir beide Gleichungen nach t ab, so folgt daraus:

$$\frac{d^n \dot{y}}{dt^n} + a_{n-1} \frac{d^{n-1} \dot{y}}{dt^{n-1}} + \dots + a_0 \dot{y}(t) = \dot{u}(t) \quad (3.40)$$

$$\text{mit } \frac{d^{n-1} \dot{y}}{dt^{n-1}}(0) = 0, \dots, \dot{y}(0) = 0. \quad (3.41)$$

Wir setzen nun in (3.38) als äußere Erregung $\bar{u}(t) := \dot{u}(t)$ (also die Ableitung des „alten“ u) und nennen die zugehörige Lösung \bar{y} . \bar{y} erfüllt also per definitionem die Gleichungen

$$\frac{d^n \bar{y}}{dt^n} + a_{n-1} \frac{d^{n-1} \bar{y}}{dt^{n-1}} + \dots + a_0 \bar{y}(t) = \dot{u}(t) \quad (3.42)$$

$$\text{mit } \frac{d^{n-1} \bar{y}}{dt^{n-1}}(0) = 0, \dots, \bar{y}(0) = 0 \quad (3.43)$$

\bar{y} und \dot{y} lösen also dieselbe DGL mit denselben Anfangsbedingungen. Da es nur eine einzige Lösung geben kann („Eindeutigkeit der Lösung“), folgt somit

$$\bar{y}(t) = \frac{dy}{dt}, \quad (3.44)$$

und (3.37) ist gezeigt.

Konsequenz von (3.37):

Wenn wir – für verschwindende Anfangswerte – in Gl. (3.30) die äußere Erregung $u(t)$ durch $\dot{u}(t)$ ersetzen, so können wir die Zustandsdefinition (3.31) sowie die Zustandsgleichung, also \mathbf{A} aus (3.32) und \mathbf{b} aus (3.33), unverändert übernehmen. Einzige Änderung: Die Ausgangsgleichung greift statt y deren Ableitung \dot{y} ab, also statt der ersten Zustandskomponente die zweite: $\mathbf{c}^T = (0 \quad 1 \quad 0 \quad \dots \quad 0)$.

Die allgemeine rechte Seite der DGL (3.15),

$$\tilde{u} := b_q \frac{d^q u}{dt^q} + \dots + b_1 \dot{u} + b_0 u, \quad (3.45)$$

können wir als Summe von q unterschiedlichen Eingangssignalen auffassen. Wenn wir die Lösung von DGL (3.30), also mit u als alleiniger Erregung, wiederum y nennen, so ergibt sich die Lösung \tilde{y} zur Erregung \tilde{u} nach dem Superpositionsprinzip als Summe der durch diese q Eingangssignale hervorgerufenen Ausgangssignale, d.h.

$$\tilde{y} = b_q \frac{d^q y}{dt^q} + \dots + b_1 \dot{y} + b_0 y. \quad (3.46)$$

Wiederum können wir also die Zustandsdefinition und die Zustandsgleichung, d.h. \mathbf{A} und \mathbf{b} , unverändert vom $(q=0)$ -Fall, also aus Gl. (3.32) und (3.33), übernehmen.

$\tilde{y}(t)$ lässt sich gemäß

$$\tilde{y}(t) = b_q x_{q+1} + b_{q-1} x_q + \dots + b_0 x_1$$

aus den ersten $q+1$ Zustandsgrößen bilden, falls $q < n$ gilt $\left\{ \text{d.h. } \mathbf{c}^T = (b_0 \ b_1 \ \dots \ b_q \ 0 \ \dots \ 0) \right\}$. Ist $q = n$, so hängt \tilde{y} auch von der n -ten Ableitung von y ab:

$$\begin{aligned} \tilde{y}(t) &= b_n \frac{d^n y}{dt^n} + b_{n-1} x_n + \dots + b_0 x_1 \\ (3.30) \quad &= b_n (-a_0 x_1 - \dots - a_{n-1} x_n + u) + b_0 x_1 + \dots + b_{n-1} x_n \\ &= (b_0 - b_n a_0) x_1 + \dots + (b_{n-1} - b_n a_{n-1}) x_n + b_n u. \end{aligned} \quad (3.47)$$

Wir fassen zusammen:

Das zur allgemeinen DGL (3.15) äquivalente Zustandsraummodell hat die durch \mathbf{A} , \mathbf{b} aus (3.32), (3.33) gegebene Zustandsgleichung für den gemäß (3.31) gebildeten Zustand $\mathbf{x}(t)$. Die Parameter \mathbf{c}^T und d der Ausgabegleichung lauten für $q < n$

$$\mathbf{c}^T = (b_o \ b_1 \ \dots \ b_q \ 0 \ \dots \ 0), \quad (3.48)$$

$$d = 0, \quad (3.49)$$

für $q = n$ haben sie die Form (siehe (3.47)) :

$$\mathbf{c}^T = (b_o - b_n a_o \ b_1 - b_n a_1 \ \dots \ b_{n-1} - b_n a_{n-1}) \quad (3.50)$$

$$d = b_n. \quad (3.51)$$

Ein Zustandsraummodell mit \mathbf{A} in Frobeniusform (3.32) und \mathbf{b} in der Form (3.33) heißt Modell in **Regelungsnormalform**. Die speziellen Strukturen von \mathbf{A} und \mathbf{b} erleichtern die Analyse des Systems und den Reglerentwurf.

In der obigen Herleitung hatten wir stets verschwindende Anfangswerte für $y, \dot{y} \dots$ vorausgesetzt. Nichtverschwindende Anfangswerte

$$\frac{d^{n-1}y}{dt^{n-1}}(0) = y_{0n}, \ \dots, \ \dot{y}(0) = y_{02}, \ y(0) = y_{01}$$

werden im obigen Zustandsraummodell einfach durch entsprechende Anfangswerte für den Zustandsvektor \mathbf{x} berücksichtigt:

$$\mathbf{x}(0) = \begin{pmatrix} y_{01} \\ \vdots \\ y_{0n} \end{pmatrix} \quad (3.52)$$

3.2.5 Linearisierung nichtlinearer Systeme

Die physikalischen Gleichungen der meisten dynamischen Prozesse sind nichtlinear. Für viele reale Systeme entsteht also bei der Modellierung anstatt der linearen DGL (3.15) eine nichtlineare DGL bzw. anstatt des linearen Zustandsraummodells (3.28), (3.29) ein nichtlineares Zustandsraummodell der Form

$$\text{Zustandsgleichung: } \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), u(t)), \ \mathbf{x}(0) = \mathbf{x}_o \quad (3.53)$$

$$\text{Ausgabegleichung: } y(t) = g(\mathbf{x}(t), u(t)) \quad (3.54)$$

Dabei stellt \mathbf{f} eine vektorwertige Funktion dar:

$$\mathbf{f}(\mathbf{x}, u) = \begin{pmatrix} f_1(x_1, \dots, x_n, u) \\ \vdots \\ f_n(x_1, \dots, x_n, u) \end{pmatrix}$$

Trotz dieses nichtlinearen Charakters der realen Welt verwendet man in der Regelungstechnik meist lineare Modelle. Da Regelungen häufig die Aufgabe haben, ein System trotz der Einwirkungen äußerer Störungen in einem vorgegebenen Arbeitspunkt zu halten, ist die Verwendung linearer Modelle leicht zu rechtfertigen: Aufgrund der Regelung bleibt das System in der Umgebung des Arbeitspunktes, und es genügt, wenn das Modell die Systembewegung um diesen Arbeitspunkt beschreibt. Für kleine Abweichungen kann man nichtlineare Modelle linearisieren.

Ausgangspunkt dieser sog. Kleinsignal-Linearisierung ist die Festlegung eines stationären Punktes als Arbeitspunkt. Ein stationärer (oder Gleichgewichts-)Punkt ist ein Zustand $\bar{\mathbf{x}}$ zusammen mit einem konstanten Eingangssignalwert $u(t) = \bar{u}$, die dadurch charakterisiert sind, dass sich das System nicht mehr bewegt, sobald $(\bar{\mathbf{x}}, \bar{u})$ einmal erreicht ist, d.h.

$$\mathbf{f}(\bar{\mathbf{x}}, \bar{u}) = \mathbf{0} \quad (= \dot{\mathbf{x}}) \quad (3.55)$$

Der Wert der Ausgabegröße in diesem stationären Zustand sei mit \bar{y} bezeichnet:

$$\bar{y} = g(\bar{\mathbf{x}}, \bar{u}) \quad (3.56)$$

Wir entwickeln nun die nichtlineare Zustandsgleichung (3.53) um diesen Arbeitspunkt herum, d.h. wir führen die Abweichungen („Kleinsignale“)

$$\delta \mathbf{x}(t) := \mathbf{x}(t) - \bar{\mathbf{x}} \quad (3.57)$$

$$\text{und} \quad \delta u(t) := u(t) - \bar{u} \quad (3.58)$$

ein und ersetzen $\mathbf{f}(\mathbf{x}, u)$ approximativ durch die ersten beiden Terme seiner Taylor-Entwicklung (also bis zum linearen Glied):

$$\begin{aligned} \frac{d\delta \mathbf{x}}{dt} &\stackrel{(3.57)}{=} \dot{\mathbf{x}} \stackrel{(3.53)}{=} \mathbf{f}(\bar{\mathbf{x}} + \delta \mathbf{x}(t), \bar{u} + \delta u(t)) \\ &\approx \underbrace{\mathbf{f}(\bar{\mathbf{x}}, \bar{u})}_{= \mathbf{0}} + \left. \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \right|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}} \delta \mathbf{x}(t) + \left. \left(\frac{\partial \mathbf{f}}{\partial u} \right) \right|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}} \delta u(t) \\ &= 0 \quad \text{wegen Gleichgewichtsbedingung (3.55)} \end{aligned}$$

Hierbei sind die (n, n) -Matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ bzw. der Spaltenvektor $\frac{\partial \mathbf{f}}{\partial u}$ identisch mit den ersten n Spalten bzw. mit der letzten Spalte der Jakobimatrix von \mathbf{f} :

$$\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & & & \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} =: \mathbf{A} \quad (3.59)$$

$$\left(\frac{\partial \mathbf{f}}{\partial u} \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}} = \begin{pmatrix} \frac{\partial f_1}{\partial u} \\ \vdots \\ \frac{\partial f_n}{\partial u} \end{pmatrix} =: \mathbf{b} \quad (3.60)$$

Mit den soeben eingeführten Bezeichnungen \mathbf{A} und \mathbf{b} erhält man die lineare DGL

$$\frac{d \delta \mathbf{x}}{dt} = \mathbf{A} \delta \mathbf{x}(t) + \mathbf{b} \delta u(t), \quad \delta \mathbf{x}(0) = \mathbf{x}_0 - \bar{\mathbf{x}}$$

Meist lässt man das δ vor den Signalen weg, ersetzt das „ungefähr gleich“ durch ein „=“ und schreibt einfach

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x}(t) + \mathbf{b} u(t), \quad \mathbf{x}(0) = \mathbf{x}_0 - \bar{\mathbf{x}}. \quad (3.61)$$

Bei der Interpretation der Ergebnisse, die mit diesem linearisierten Modell berechnet werden, muss man sich natürlich wieder daran erinnern, dass \mathbf{x} und u die Abweichung(en) vom Gleichgewichtszustand/Arbeitspunkt bedeuten.

In gleicher Weise kann auch die nichtlineare Ausgabegleichung (3.54) linearisiert werden.

Mit $\delta y(t) := y(t) - \bar{y}$ ergibt sich

$$\begin{aligned}
 y(t) &= \bar{y} + \delta y(t) \\
 \text{Taylor-Entwicklung von (3.54):} \quad y(t) &\approx \underbrace{g(\mathbf{x}, u) \Big|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}}}_{\parallel} + \underbrace{\left(\frac{\partial g}{\partial \mathbf{x}} \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}}}_{=: \mathbf{c}^T} \delta \mathbf{x} + \underbrace{\left(\frac{\partial g}{\partial u} \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}, u=\bar{u}}}_{=: d} \delta u
 \end{aligned}$$

und in vereinfachter Schreibweise schließlich

$$y(t) = \mathbf{c}^T \mathbf{x}(t) + d u(t). \quad (3.62)$$

3.3 Lösung der Zustandsgleichung

3.3.1 Lösung für skalaren Zustand

Für den Fall einer skalarwertigen Zustandsgröße x lautet die Zustandsgleichung

$$\dot{x} = a x(t) + b u(t), \quad x(0) = x_0. \quad (3.63)$$

Wir wollen zunächst den Raum *aller* Lösungen von (3.63) ohne Berücksichtigung der Anfangsbedingung bestimmen und dann aus diesem Lösungsraum diejenige – eindeutig bestimmte – Lösung auswählen, die die Anfangsbedingung erfüllt. Der Raum aller Lösungen – oft auch als „allgemeine Lösung“ bezeichnet – setzt sich additiv aus der allgemeinen Lösung der homogenen Gleichung und einer beliebigen speziellen (wie auch immer gefundenen) Lösung der inhomogenen (d.h. „kompletten“) Gleichung zusammen. Deshalb betrachten wir zuerst die homogene DGL

$$\dot{x} = a x(t). \quad (3.64)$$

Die Ableitung der Funktion nach t reproduziert also bis auf eine Konstante die Funktion selber. Da wir dieses Verhalten von der Exponentialfunktion her kennen, machen wir den Lösungsansatz:

$$x(t) = k \cdot e^{\lambda t}, \quad k, \lambda \in \mathbb{R} \quad (3.65)$$

Einsetzen in (3.64) liefert

$$\begin{aligned} k \cdot \lambda \cdot e^{\lambda t} &= a \cdot k \cdot e^{\lambda t} \quad \forall t \geq 0 \\ \Leftrightarrow k \cdot \lambda &= a \cdot k \quad (\text{für } k \neq 0) \quad \Leftrightarrow \lambda = a. \\ x(t) &= k \cdot e^{a t} \end{aligned} \quad (3.66)$$

ist also die allgemeine Lösung der homogenen DGL (3.64). (Beachte: k ist hier eine frei wählbare reelle Konstante, d.h. der Raum aller Lösungen ist ein durch k parametrisierter eindimensionaler Vektorraum.)

Eine Lösung der inhomogenen DGL erhalten wir, wenn wir in den Ansatz (3.65) einen zusätzlichen Freiheitsgrad einführen und die Konstante k durch eine Funktion der Zeit $k(t)$ ersetzen (sog. Methode der Variation der Konstanten).

$$\rightarrow \text{Lösungsansatz: } x(t) = k(t) \cdot e^{a t} \quad (3.67)$$

Einsetzen in (3.63) ergibt

$$\begin{aligned} a \cdot e^{at} k(t) + e^{at} \dot{k}(t) &= a e^{at} k(t) + b u(t) \\ \Leftrightarrow \dot{k}(t) &= e^{-at} \cdot b \cdot u(t) \end{aligned}$$

Integration über das Intervall $[0, t]$ ergibt

$$\int_0^t \dot{k}(\tau) d\tau = k(t) - k(0) = \int_0^t e^{-a\tau} \cdot b \cdot u(\tau) d\tau,$$

und Einsetzen dieser Funktion $k(t)$ in den Ansatz (3.67) liefert

$$x(t) = k(0) \cdot e^{at} + \int_0^t e^{a(t-\tau)} \cdot b \cdot u(\tau) d\tau. \quad (3.68)$$

Da dieses $x(t)$ die inhomogene DGL für jede (reelle) Wahl des Parameters $k(0)$ erfüllt, haben wir somit gleich die allgemeine Lösung der inhomogenen DGL erhalten. (Eine spezielle Lösung wäre beispielsweise der zweite Summand, der der Wahl $k(0) = 0$ entspricht.)

Die vorgegebene Anfangsbedingung $x(0) = x_0$ erzwingt nun die Auswahl genau einer Lösung aus diesem (durch $k(0)$ parametrisierten) Lösungsraum, nämlich

$$k(0) = x_0.$$

Mit der Bezeichnung $e^{at} =: \phi(t)$ erhält die Lösung von (3.63) die Form

$$x(t) = \underbrace{\phi(t) x_0}_{x_{\text{frei}}(t)} + \underbrace{\int_0^t \phi(t-\tau) \cdot b \cdot u(\tau) d\tau}_{x_{\text{erzw.}}(t)} \quad (3.69)$$

Gl. (3.69) ist die **Bewegungsgleichung** des durch die Zustandsgleichung (3.63) beschriebenen Systems. Der erste Summand beschreibt die Lösung der homogenen DGL, also diejenige Bewegung im Zustandsraum, die das System ohne Erregung von außen, allein aufgrund der Anfangsauslenkung x_0 ausführt. $x_{\text{frei}}(t)$ heißt daher auch **Eigenbewegung** oder **freie Bewegung** des Systems. Der zweite Summand beschreibt die durch die äußere Erregung $u(t)$ **erzwungene Bewegung** des Systems.

Die Eigenbewegung hat, je nach Vorzeichen des Parameters a , drei typische Formen (siehe Abb. 14):

- $a < 0$: Eigenbewegung klingt exponentiell ab, System geht asymptotisch in den Ruhezustand $x = 0$ über.
- $a = 0$: System verharrt im Anfangszustand x_0 .
- $a > 0$: Eigenbewegung klingt auf, d.h. die Zustandsgröße wächst exponentiell über alle Grenzen und das System entfernt sich immer weiter von der Ruhelage $x = 0$.

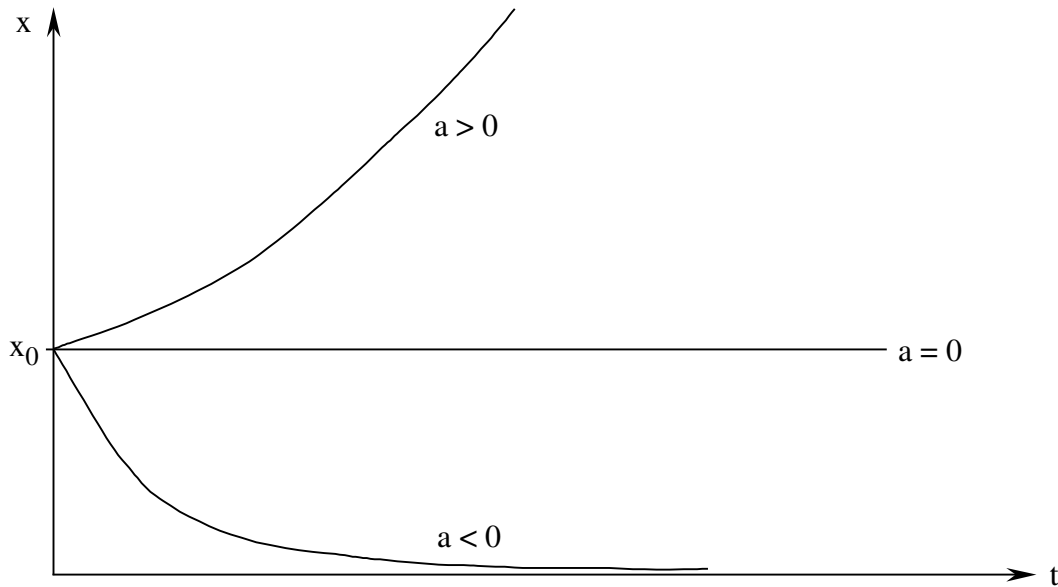
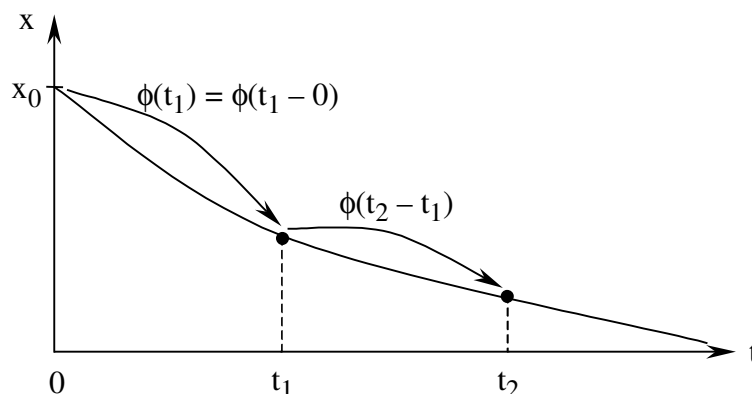


Abb 14 : Eigenbewegung des linearen Systems (3.63)

In Abwesenheit einer äußeren Erregung u wird der Übergang von $x(0)$ nach $x(t)$ allein durch $\phi(t) = e^{at}$ bestimmt. Für drei zeitliche Momentaufnahmen $x(0)$, $x(t_1)$, $x(t_2)$ mit $0 \leq t_1 \leq t_2$ gilt die Beziehung

$$\begin{aligned}
 (3.69) \quad x(t_2) &= e^{at_2} x(0) = e^{a(t_2 - t_1 + t_1)} \cdot x(0) = \\
 &= e^{a(t_2 - t_1)} \cdot e^{at_1} \cdot x(0) = e^{a(t_2 - t_1)} \cdot x(t_1) = \phi(t_2 - t_1) x(t_1)
 \end{aligned}
 \tag{3.70}$$



Allgemein wird also der Übergang des Systemzustands von $x(t_1)$ nach $x(t_2)$ (für $t_1 \leq t_2$) nur durch $\phi(t_2 - t_1)$ bestimmt (freie Bewegung vorausgesetzt). Zusammen mit dem Superpositionsprinzip ermöglicht diese Erkenntnis ein tieferes Verständnis der Bewegungsgleichung (3.69).

Laut Superpositionsprinzip überlagern sich – bei verschwindendem Anfangswert – die Systemantworten, die durch einzelne „Inputs“ hervorgerufen werden, additiv.

Der erste Summand in (3.69) kann als Systemantwort auf die Anfangsauslenkung $x(0)$ aufgefasst werden. Diese Antwort wird einfach durch den „Übergangs-Operator“ $\phi(t)$ für das freie System bestimmt.

Um den zweiten Term zu interpretieren, zerlegen wir das Eingangssignal $u(t)$ gedanklich in eine Sequenz einzelner Rechteckpulse der Breite Δ (siehe Abb. 15).

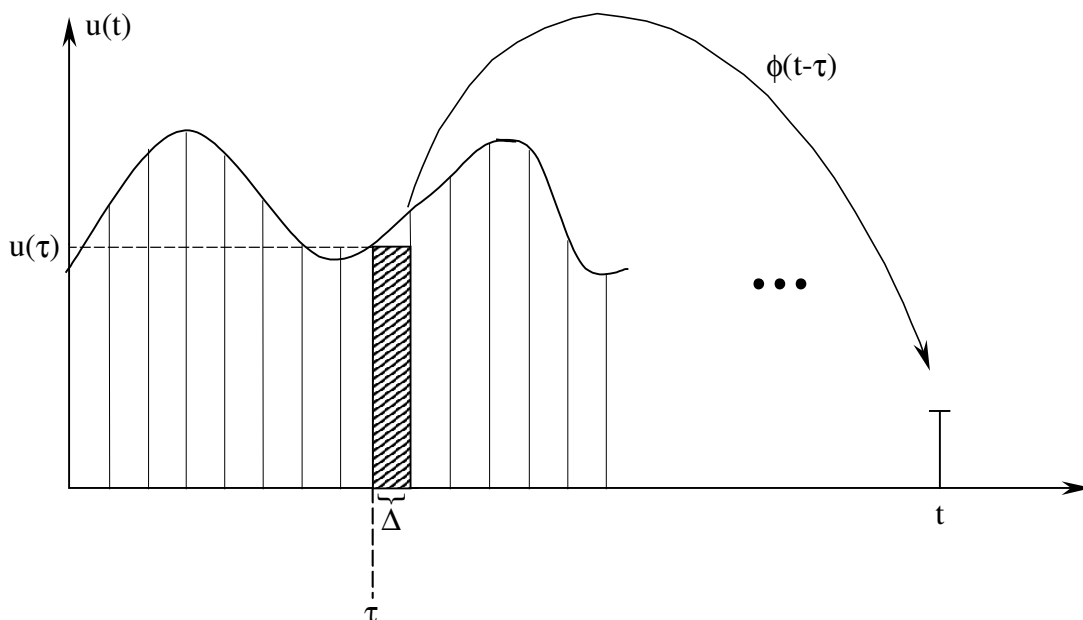


Abb. 15 : Zerlegung des Eingangssignals in Rechteckimpulse

Wir greifen nun den Rechteckimpuls zur Zeit τ heraus (Pulshöhe also $u(\tau)$) und nehmen an, dieser Puls sei – bei verschwindendem Anfangswert – das einzige Eingangssignal. Dann wäre der Zustandswert zu Beginn des Pulses immer noch 0, d.h. $x(\tau) = 0$, und am Ende des Pulses wäre der Zustand wegen $\dot{x}(\tau) = \underbrace{a x(\tau)}_0 + b u(\tau)$ gleich

$$x(\tau + \Delta) = x(\tau) + \dot{x}(\tau) \cdot \Delta = b \cdot u(\tau) \cdot \Delta .$$

Bis zum Zeitpunkt t würde nun wieder freie Bewegung herrschen, d.h. die Systemantwort zum Zeitpunkt t auf den Rechteckpuls $u(\tau)$ lautet

$$\phi(t - \tau) \cdot b \cdot u(\tau) \cdot \Delta .$$

Da die Gesamtantwort auf die ganze Pulssequenz die Summe der Antworten auf die Einzelpulse ist, ergibt sich im Grenzwert $\Delta \rightarrow 0$ genau das Integral

$$\int_0^t \phi(t-\tau) \cdot b \cdot u(\tau) d\tau.$$

3.3.2 Lösung für vektoriellen Zustand

Die vektorielle Zustandsgleichung

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x}(t) + \mathbf{b} u(t), \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (3.71)$$

wird völlig analog zum skalaren Fall gelöst. Homogene Gleichung:

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x}(t) \quad (3.72)$$

Lösungsansatz:

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{k}, \quad (3.73)$$

wobei der Exponent $\mathbf{A}t$ aufgrund der Lösung $\lambda = a$ für den skalaren Fall eingesetzt wurde.

$e^{\mathbf{A}t}$ ist, wie \mathbf{A} selbst, eine (n, n) -Matrix, heißt **Matrixexponentialfunktion** und ist analog zur reellwertigen e-Funktion als (Grenzwert einer) Reihe definiert:

$$\text{reelle e-Funktion: } e^{at} = \sum_{i=0}^{\infty} \frac{a^i t^i}{i!} = 1 + at + \frac{a^2}{2!} t^2 + \dots$$

$$\text{Matrixexponentialfunktion: } e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \frac{\mathbf{A}^i t^i}{i!} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2}{2!} t^2 + \dots \quad (3.74)$$

Die Reihe konvergiert für alle quadratischen Matrizen \mathbf{A} . Daher kann die Differentiation mit der Summenbildung vertauscht werden, und man erhält

$$(\text{in Analogie zum skalaren Fall, d.h. zu } \frac{d}{dt} e^{at} = a \cdot e^{at})$$

$$\begin{aligned} \frac{d}{dt} e^{\mathbf{A}t} &= \mathbf{A} + \mathbf{A}^2 t + \frac{\mathbf{A}^3}{2!} t^2 + \dots \\ &= \mathbf{A} \left(\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2}{2!} t^2 + \dots \right) \\ &= \mathbf{A} e^{\mathbf{A}t} = e^{\mathbf{A}t} \mathbf{A} \end{aligned}$$

⇒ Ansatz (3.73) ist die durch \mathbf{k} parametrisierte allgemeine Lösung der homogenen DGL (3.72).

Methode der Variation der Konstanten (d.h. ersetze Konstante \mathbf{k} durch zeitabhängigen Vektor)

→ Lösungsansatz für inhomogene DGL (3.71) ohne Anfangsbedingung

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{k}(t) \quad (3.75)$$

Einsetzen in die DGL (3.71) ergibt

$$\dot{\mathbf{x}}(t) = \underbrace{\mathbf{A} \cdot e^{\mathbf{A}t} \mathbf{k}(t)} + e^{\mathbf{A}t} \cdot \dot{\mathbf{k}}(t) = \underbrace{\mathbf{A} \cdot e^{\mathbf{A}t} \mathbf{k}(t)} + \mathbf{b} u(t) \quad (3.76)$$

Die Matrix $e^{\mathbf{A}t}$ ist für beliebige Matrizen \mathbf{A} (und für beliebige t) regulär und damit invertierbar, und es gilt

$$\left(e^{\mathbf{A}t}\right)^{-1} = e^{-\mathbf{A}t}.$$

Also ist (3.76) gleichbedeutend mit

$$\dot{\mathbf{k}}(t) = e^{-\mathbf{A}t} \mathbf{b} u(t).$$

Integration über das Intervall $[0, t]$ liefert

$$\int_0^t \dot{\mathbf{k}}(\tau) d\tau = \mathbf{k}(t) - \mathbf{k}(0) = \int_0^t e^{-\mathbf{A}\tau} \cdot \mathbf{b} \cdot u(\tau) d\tau.$$

Einsetzen dieser vektorwertigen Funktion $\mathbf{k}(t)$ in den Ansatz (3.75) ergibt schließlich

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{k}(0) + \int_0^t e^{\mathbf{A}(t-\tau)} \cdot \mathbf{b} \cdot u(\tau) d\tau$$

als allgemeine, durch $\mathbf{k}(0)$ parametrisierte Lösung der inhomogenen DGL. Die vorgegebene Anfangsbedingung $\mathbf{x}(0) = \mathbf{x}_0$ erzwingt die Festlegung

$$\mathbf{k}(0) = \mathbf{x}_0.$$

Wenn wir noch die matrixwertige Funktion

$$\Phi(t) = e^{\mathbf{A}t} \quad (3.77)$$

eingeführen, so lautet die Lösung der Zustandsgleichung (3.71), genannt **Bewegungsgleichung** des Systems:

$$\mathbf{x}(t) = \underbrace{\Phi(t) \cdot \mathbf{x}_0}_{\mathbf{x}_{\text{frei}}(t)} + \underbrace{\int_0^t \Phi(t-\tau) \cdot \mathbf{b} \cdot u(\tau) d\tau}_{\mathbf{x}_{\text{erzw.}}(t)} \quad (3.78)$$

Analog zum skalaren Fall bestimmt die Matrix $\Phi(t-\tau)$ bei freier Bewegung den Übergang des Systemzustands von $\mathbf{x}(\tau)$ nach $\mathbf{x}(t)$ (für $\tau \leq t$) und heißt deshalb **Übergangsmatrix** oder **Transitionsmatrix** des dynamischen Systems.

Die Eigenbewegung des Systems entfällt, wenn $\mathbf{x}_0 = \mathbf{0}$ ist. In diesem Fall hat das System zum Zeitpunkt $t = 0$ keine Energie gespeichert und führt deshalb aus eigener Kraft keine Bewegung aus.

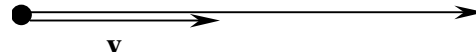
Vielfach beschreibt \mathbf{x} ja nicht den absoluten Wert der Zustandsgrößen, sondern die Abweichung $\delta \mathbf{x}$ vom Gleichgewichts-Arbeitspunkt $\bar{\mathbf{x}}$. $\mathbf{x}_0 = \mathbf{0}$ bedeutet dann, dass das System keine Energie über die Energie des Gleichgewichts-Punktes hinaus gespeichert hat und keine freie Bewegung um den Arbeitspunkt ausführt.

3.3.3 Diagonalisierung der Systemmatrix

Die Zustandsgleichung (3.71) ist nichts anderes als die vektorielle Schreibweise für n gekoppelte skalare Differentialgleichungen. Wenn wir die Zustandsvektoren \mathbf{x} aus der Sicht einer geeignet gewählten Basis des Zustandsraums beschreiben, können wir diese Differentialgleichungen **entkoppeln** und somit das System in n unabhängige Subsysteme zerlegen. Die hierfür benötigte, für das System maßgeschneiderte Basis des Zustandsraums \mathbb{R}^n besteht aus den Eigenvektoren der Systemmatrix \mathbf{A} .

Zur Erinnerung:

Wenn für einen Vektor $\mathbf{v} \in \mathbb{R}^n$ die Wirkung der Matrix \mathbf{A} nur in einer Streckung (oder Stauchung) um den Faktor λ besteht,

$$\mathbf{A} \mathbf{v} = \lambda \cdot \mathbf{v}$$


so heißt λ **Eigenwert** und \mathbf{v} der zu λ gehörige **Eigenvektor** der Matrix \mathbf{A} .

Aus der linearen Algebra wissen wir, dass eine (möglicherweise komplexe) Zahl λ genau dann Eigenwert von \mathbf{A} ist, wenn sie die **charakteristische Gleichung**

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (3.79)$$

löst. $\det(\mathbf{A} - \lambda \mathbf{I})$ ist ein Polynom vom Grad n in der Variablen λ ,

$$\det(\mathbf{A} - \lambda \mathbf{I}) = P_A(\lambda),$$

genannt **charakteristisches Polynom** des Systems bzw. der (System-)Matrix \mathbf{A} . Es gilt also:

$$\lambda \in \mathbb{C} \text{ Eigenwert von } \mathbf{A} \Leftrightarrow \lambda \text{ Nullstelle des charakteristischen Polynoms } P_A$$

Nach dem Fundamentalsatz der Algebra kann das charakteristische Polynom faktorisiert werden in der Form

$$P_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda),$$

wobei die – nicht notwendig verschiedenen – Zahlen $\lambda_i \in \mathbb{C}$ die Nullstellen des Polynoms und damit die Eigenwerte des Systems sind.

Wenn die Eigenwerte $\lambda_1, \dots, \lambda_n$ paarweise verschieden sind, so sind die zugehörigen Eigenvektoren $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$ stets linear unabhängig und bilden somit eine Basis des Zustandsraums \mathbb{R}^n , die sog. Eigenbasis zu \mathbf{A} .

Wir nehmen im folgenden an, dass die n Eigenvektoren linear unabhängig sind, und beschreiben den Zustandsvektor \mathbf{x} durch seine Koordinaten z_1, \dots, z_n bezüglich der Eigenbasis $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$, d.h.

$$\mathbf{x} = z_1 \mathbf{v}^{(1)} + \dots + z_n \mathbf{v}^{(n)}.$$

Ausgeschrieben bedeutet dies

$$\begin{aligned} \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= z_1 \overbrace{\begin{pmatrix} \mathbf{v}_1^{(1)} \\ \vdots \\ \mathbf{v}_n^{(1)} \end{pmatrix}}^{\mathbf{v}^{(1)}} + \dots + z_n \overbrace{\begin{pmatrix} \mathbf{v}_1^{(n)} \\ \vdots \\ \mathbf{v}_n^{(n)} \end{pmatrix}}^{\mathbf{v}^{(n)}} \\ &= \underbrace{\left(\begin{array}{c|c} \mathbf{v}_1^{(1)} & \mathbf{v}_1^{(n)} \\ \vdots & \vdots \\ \mathbf{v}_n^{(1)} & \mathbf{v}_n^{(n)} \end{array} \right)}_{(\mathbf{v}^{(1)} | \dots | \mathbf{v}^{(n)}) =: \mathbf{V}} \underbrace{\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}}_{=: \mathbf{z}} = \mathbf{V} \cdot \mathbf{z} \end{aligned}$$

Da die Transformationsmatrix \mathbf{V} regulär ist (linear unabhängige Spalten!), erhält man also den Koordinatenvektor \mathbf{z} (und somit den „neuen Zustandsvektor“) gemäß

$$\mathbf{z} = \mathbf{V}^{-1} \mathbf{x} \quad (3.80)$$

aus dem „alten“ Zustandsvektor \mathbf{x} .

Wenn wir nun die Zustandsgleichung (3.71) von links mit der Matrix \mathbf{V}^{-1} multiplizieren, so erhalten wir die Zustandsgleichung für den neuen/transformierten Zustandsvektor \mathbf{z} :

$$\begin{aligned} \underbrace{\mathbf{V}^{-1} \dot{\mathbf{x}}}_{\dot{\mathbf{z}}} &= \mathbf{V}^{-1} \mathbf{A} \underbrace{\mathbf{x}}_{\mathbf{V} \mathbf{z}} + \underbrace{\mathbf{V}^{-1} \mathbf{b}}_{=: \tilde{\mathbf{b}}} u(t), & \underbrace{\mathbf{V}^{-1} \mathbf{x}(0)}_{\mathbf{z}(0)} &= \mathbf{V}^{-1} \mathbf{x}_0 \\ \Leftrightarrow \quad \dot{\mathbf{z}} &= \underbrace{\mathbf{V}^{-1} \mathbf{A} \mathbf{V}}_{=: \mathbf{\Lambda}} \mathbf{z} + \tilde{\mathbf{b}} u(t) \\ \Leftrightarrow \quad \dot{\mathbf{z}} &= \mathbf{\Lambda} \mathbf{z} + \tilde{\mathbf{b}} u(t), & \mathbf{z}(0) &= \mathbf{V}^{-1} \mathbf{x}_0, \end{aligned} \quad (3.81)$$

wobei $\tilde{\mathbf{b}}$ die Koordinatendarstellung von \mathbf{b} bezüglich der Eigenvektor-Basis ist und $\mathbf{\Lambda}$ die „neue“ Systemmatrix (d.h. die Darstellung von \mathbf{A} bezüglich des neuen Koordinatensystems).

$\mathbf{\Lambda}$ hat eine ganz spezielle Gestalt, nämlich Diagonalstruktur (beachte: genau diese Diagonalstruktur war es, was wir mit unserer Koordinatentransformation des Zustandsraums erreichen wollten!), was leicht zu zeigen ist:

$$\begin{aligned} \mathbf{\Lambda} &= \mathbf{V}^{-1} \mathbf{A} \mathbf{V} \\ &= \mathbf{V}^{-1} \mathbf{A} \left(\underbrace{\mathbf{v}^{(1)} \mid \dots \mid \mathbf{v}^{(n)}}_{\left(\lambda_1 \mathbf{v}^{(1)} \mid \dots \mid \lambda_n \mathbf{v}^{(n)} \right)} \right) \\ &= \underbrace{\left(\mathbf{v}^{(1)} \mid \dots \mid \mathbf{v}^{(n)} \right)}_{\mathbf{V}^{-1} \mathbf{V}} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ &= \underbrace{\mathbf{V}^{-1} \mathbf{V}}_{\mathbf{I}} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ \Rightarrow \quad \mathbf{\Lambda} &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \end{aligned} \quad (3.82)$$

Aufgrund der erzielten Diagonalgestalt der transformierten Matrix \mathbf{A} wird der Übergang zur Eigenvektor-Basis auch als **Diagonalisierung der Systemmatrix** bezeichnet.

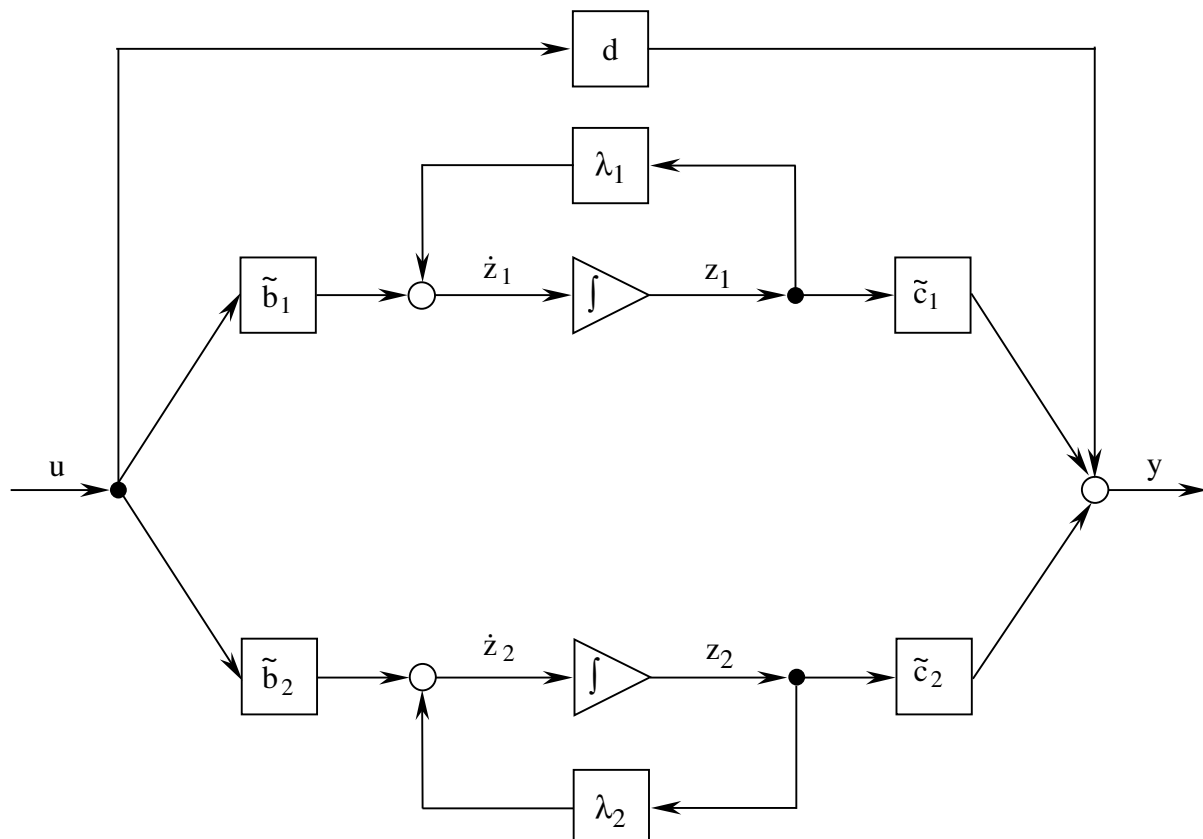
Schreiben wir die Zustandsgleichung (3.81) etwas ausführlicher hin, also

$$\begin{pmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{pmatrix} u(t)$$

Beziehungsweise

$$\begin{aligned} \dot{z}_1 &= \lambda_1 z_1(t) + \tilde{b}_1 u(t) \\ &\vdots \\ \dot{z}_n &= \lambda_n z_n(t) + \tilde{b}_n u(t), \end{aligned}$$

so sehen wir, dass die Kopplungen zwischen den Zustandsvariablen in dieser (Koordinaten-) Darstellung (bezüglich der Eigenbasis von \mathbf{A}) völlig wegfallen. Beispielhaft illustriert dies das Signalflussdiagramm eines Systems zweiter Ordnung (vergl. Abb. 13!):



Obiges Signalflussdiagramm nimmt bereits die Ausgabegleichung in den neuen Zustandsvariablen vorweg. Sie lautet:

$$\begin{aligned} y(t) &= \mathbf{c}^T \mathbf{x}(t) + d u(t) \\ &= \underbrace{\mathbf{c}^T \mathbf{V}}_{=: \tilde{\mathbf{c}}^T} \mathbf{z}(t) + d u(t) \end{aligned}$$

$$\Rightarrow \text{Ausgabegleichung: } y(t) = \tilde{\mathbf{c}}^T \mathbf{z}(t) + d u(t) \quad \text{mit} \quad \tilde{\mathbf{c}}^T := \mathbf{c}^T \mathbf{V} \quad (3.83)$$

Betrachtet man nur die Eigenbewegung des Systems, so zerfällt die Vektordifferentialgleichung in n voneinander unabhängige skalare Differentialgleichungen, die quasi n vollständig entkoppelte Subsysteme mit jeweils skalaren Zustandsgrößen beschreiben:

$$\begin{aligned} \dot{z}_1 &= \lambda_1 z_1(t), & z_1(0) &= z_{01} \\ &\vdots \\ \dot{z}_n &= \lambda_n z_n(t), & z_n(0) &= z_{0n} \end{aligned}$$

$$\text{wobei} \quad \begin{pmatrix} z_{01} \\ \vdots \\ z_{0n} \end{pmatrix} \equiv \mathbf{z}_0 := \mathbf{V}^{-1} \mathbf{x}_0.$$

Die durch den Basiswechsel entstandenen Zustandsvariablen z_i werden auch als **kanonische Zustandsvariable** bezeichnet. Sie sind i.a. keine physikalisch interpretierbaren Größen mehr, sondern reine Rechengrößen.

Wie eine homogene DGL der Form

$$\dot{z}_i = \lambda_i z_i(t)$$

zu lösen ist, wissen wir bereits:

$$\dot{z}_i(t) = e^{\lambda_i t} \cdot z_i(0). \quad (3.84)$$

Auch für die zugehörige inhomogene DGL

$$\dot{z}_i = \lambda_i z_i + \tilde{\mathbf{b}}_i \cdot u(t)$$

kennen wir die Lösung (siehe Formel (3.68)):

$$\dot{z}_i(t) = e^{\lambda_i t} z_i(0) + \int_0^t e^{\lambda_i(t-\tau)} \tilde{b}_i u(\tau) d\tau. \quad (3.85)$$

Wenn wir nun die n Übergangsfaktoren $e^{\lambda_i t}$ in Form einer Diagonalmatrix anordnen,

$$\tilde{\Phi} := \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}, \quad (3.86)$$

so können wir die n einzelnen Bewegungsgleichungen für die kanonischen Zustandsvariablen z_1, \dots, z_n in vektorieller Schreibweise zusammenfassen zu

$$\mathbf{z}(t) = \underbrace{\tilde{\Phi}(t) \mathbf{z}_0}_{\mathbf{z}_{\text{frei}}(t)} + \underbrace{\int_0^t \tilde{\Phi}(t-\tau) \tilde{\mathbf{b}} u(\tau) d\tau}_{\mathbf{z}_{\text{erzw.}}(t)}. \quad (3.87)$$

Die Diagonalmatrix $\tilde{\Phi}(t)$ ist also die Übergangsmatrix in der Eigenbasis-Koordinatendarstellung.

Da das dynamische Verhalten linearer Systeme von der Übergangsmatrix und damit von der freien Bewegung bestimmt wird, wollen wir uns die freie Bewegung nun genauer ansehen. Um aus den kanonischen Zustandsvariablen den ursprünglichen, physikalisch interpretierbaren Zustand \mathbf{x} zurückzugewinnen, müssen wir uns nur der Bedeutung von $z_i(t)$ als Koordinate von \mathbf{x} bezüglich des Eigenvektors $\mathbf{v}^{(i)}$ erinnern:

$$\begin{aligned} \mathbf{x}_{\text{frei}}(t) &= z_{\text{frei},1}(t) \cdot \mathbf{v}^{(1)} + \dots + z_{\text{frei},n}(t) \cdot \mathbf{v}^{(n)} \\ &= e^{\lambda_1 t} z_1(0) \cdot \mathbf{v}^{(1)} + \dots + e^{\lambda_n t} z_n(0) \cdot \mathbf{v}^{(n)} \end{aligned} \quad (3.88)$$

Die freie Bewegung ist also eine lineare Überlagerung von n Funktionen („eindimensionalen Bewegungen“) der Form $e^{\lambda_i t} \cdot z_i(0) \cdot \mathbf{v}^{(i)}$, wobei $e^{\lambda_i t}$ die zeitliche Abhängigkeit und der Skalar $z_i(0)$ die „Amplitude“ darstellen. ($z_i(0)$ gibt ja den Anteil der Anfangsauslenkung \mathbf{x}_0 in Richtung von $\mathbf{v}^{(i)}$ an.) Die Funktionen $e^{\lambda_i t}$ heißen **Modi** oder **Eigenvorgänge** des Systems.

Die saloppe Bezeichnung der Funktionen $e^{\lambda_i t} \cdot z_i(0) \cdot \mathbf{v}^{(i)}$ als „eindimensionale Bewegungen“ ist allerdings nicht immer in eine räumliche Vorstellung umzusetzen, da sowohl $e^{\lambda_i t}$ eine komplexe Zahl als auch $\mathbf{v}^{(i)}$ ein komplexer Vektor sein kann.

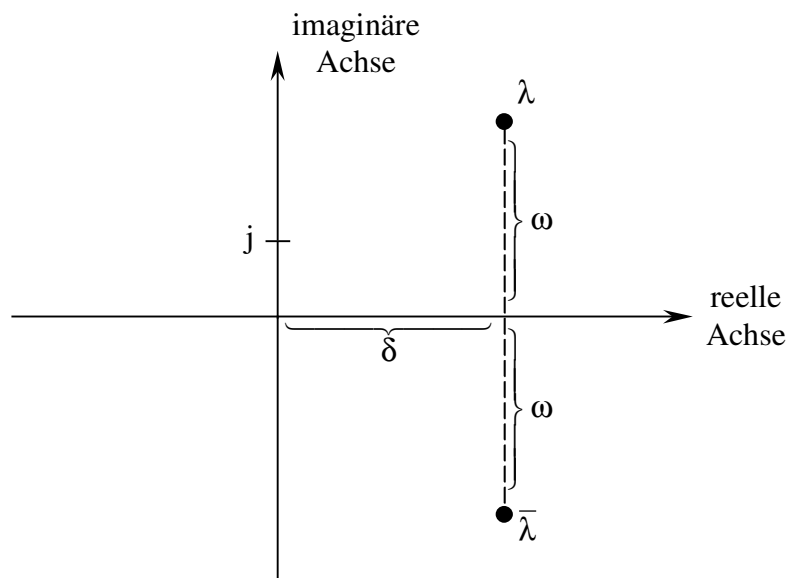
Um die Modi besser zu verstehen, müssen wir die Eigenwerte λ_i genauer betrachten. Sei λ also ein Eigenwert der Systemmatrix \mathbf{A} .

Als Nullstelle des charakteristischen Polynoms ist λ eine komplexe Zahl und kann daher als Summe von Real- und Imaginärteil geschrieben werden:

$$\lambda = \delta + j \cdot \omega, \quad (3.89)$$

wobei $j := \sqrt{-1}$ die imaginäre Einheit bezeichnet.

Da das charakteristische Polynom $P_A(\lambda)$ nur reelle Koeffizienten besitzt, muss, wenn die komplexe Zahl λ Nullstelle ist, ihre komplex Konjugierte $\bar{\lambda} = \delta - j\omega$ ebenfalls Nullstelle sein.



Im wesentlichen sind damit 2 Fälle zu unterscheiden:

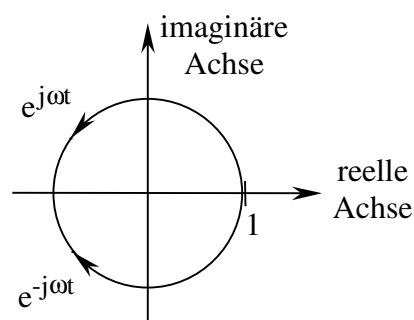
Fall a) λ ist reell, d.h. $\lambda = \delta$ (und $\omega = 0$).

Die zugehörige Mode $e^{\delta t}$ hat für alle $t \geq 0$ dasselbe Vorzeichen, ein reeller Eigenwert führt also **nicht** zu einer Schwingung.

Fall b) λ ist komplex, d.h. $\lambda = \delta + j\omega$.

Da die zugehörige Mode

$$e^{\lambda t} = e^{(\delta + j\omega)t} = \underbrace{e^{\delta t}}_{\text{reell}} \cdot \underbrace{e^{j\omega t}}_{\text{komplex}}$$



komplex ist, der physikalisch interpretierbare Zustand \mathbf{x} aber reell sein muss, darf die Mode $e^{\lambda t}$ nicht isoliert betrachtet werden. Vielmehr müssen wir das zu λ und $\bar{\lambda}$ gehörige Modenpaar gemeinsam betrachten. Die imaginären Beiträge dieses Modenpaars zu den physikalischen Zustandskomponenten x_i heben sich gerade weg. Übrig bleiben reelle Beiträge der Form

$$e^{\delta t} \cdot (\alpha \sin \omega t + \beta \cos \omega t) .$$

Das zu $(\lambda, \bar{\lambda})$ zugehörige Modenpaar liefert also als Beitrag zur Eigenbewegung eine **Schwingung** mit der Frequenz ω und einer Amplitude, die gemäß $e^{\delta t}$ entweder exponentiell abklingt, falls $\delta < 0$, oder exponentiell anwächst, falls $\delta > 0$.

Die Lage eines Eigenwertes (bzw. Eigenwert-Paares) in der komplexen Ebene verrät uns also die Rate des exponentiellen Anwachsens bzw. Abklingens der zugehörigen Mode sowie, ob es sich um eine Schwingungsmode handelt und, falls ja, welche Frequenz sie hat.

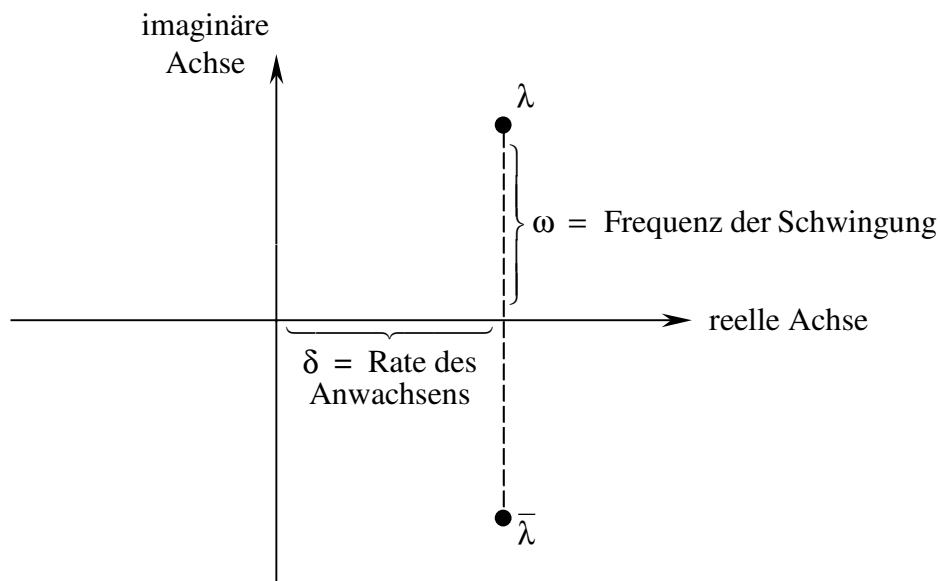


Abb. 16 : Bedeutung der Lage eines komplexen Eigenwertes

Das Langzeitverhalten eines linearen Systems wird im wesentlichen durch ein oder zwei seiner Eigenwerte bzw. Modi bestimmt. Diese **dominanten Modi** sind daher von speziellem Interesse. Um den dominanten Modus / die dominanten Modi zu finden, ordnen wir die Eigenwerte des Systems entsprechend der Größe ihrer Realteile und numerieren sie so, dass gilt

$$\text{Re}(\lambda_1) \geq \text{Re}(\lambda_2) \geq \dots \geq \text{Re}(\lambda_n) .$$

Der Eigenwert λ_1 [bzw. das Eigenwertpaar (λ_1, λ_2) , falls $\lambda_2 = \bar{\lambda}_1$ und damit $\text{Re}(\lambda_1) = \text{Re}(\lambda_2)$] mit dem größten Realteil heißt **dominanter Eigenwert**.

Seine Bedeutung wird klar, wenn wir die Modenzerlegung des freien Zustands \mathbf{x}_{frei} betrachten:

$$\begin{aligned}\mathbf{x}_{\text{frei}}(t) &= e^{\lambda_1 t} \mathbf{z}_1(0) \cdot \mathbf{v}^{(1)} + \dots + e^{\lambda_n t} \mathbf{z}_n(0) \cdot \mathbf{v}^{(n)} \\ &= e^{\delta_1 t} e^{j\omega_1 t} \mathbf{z}_1(0) \cdot \mathbf{v}^{(1)} + \dots + e^{\delta_n t} e^{j\omega_n t} \mathbf{z}_n(0) \cdot \mathbf{v}^{(n)}\end{aligned}\quad (3.90)$$

Wegen $|e^{j\omega t}| = 1$ wird das zeitliche Anwachsen bzw. Abklingen der einzelnen Terme (genauer: ihres Betrags) allein durch die Faktoren $e^{\delta_i t}$ bestimmt.

Falls beispielsweise λ_1 allein dominant ist, gilt für $t \rightarrow \infty$:

$$|e^{\delta_1 t}| \gg |e^{\delta_i t}| \quad \text{für } i = 2, 3, \dots, n$$

\Rightarrow Für große Zeiten wird der erste Term in (3.90) alle anderen Terme dominieren, und der Zustand \mathbf{x}_{frei} wird sich eng an die Richtung des Eigenvektors $\mathbf{v}^{(1)}$ anschmiegen.

Ausschlaggebend dafür, **wie schnell** sich das System gegen den Eigenvektor der dominanten Mode bewegt, ist im wesentlichen der Eigenwert mit dem zweitgrößten Eigenwert. ($\text{Re}(\lambda_1) \gg \text{Re}(\lambda_2) \Rightarrow$ schneller Übergang zum Langzeitverhalten.)

Die Lage der Eigenwerte in der komplexen Ebene gibt auch Aufschluss über die **Stabilität** des Systems. Die sog. asymptotische Stabilität eines (linearen) dynamischen Systems ist definiert im Hinblick auf einen Gleichgewichtspunkt $(\bar{\mathbf{x}}, \bar{u})$. Vorbemerkung: Das das System im Gleichgewichtspunkt ruht, muss hier die rechte Seite der Zustands-DGL

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x} + \mathbf{b} u \quad \text{verschwinden, d.h.}$$

$$\mathbf{A} \bar{\mathbf{x}} = -\mathbf{b} \bar{u} . \quad (3.91)$$

Ein Gleichgewichtspunkt $(\bar{\mathbf{x}}, \bar{u})$ eines linearen, zeitinvarianten, dynamischen Systems heißt **asymptotisch** (Zustands-) **stabil**, wenn mit $t \rightarrow \infty$ der Zustandsvektor \mathbf{x} für **jeden beliebigen Anfangswert** gegen $\bar{\mathbf{x}}$ strebt (bei konstantem Eingangssignal $u(t) \equiv \bar{u}$). Gibt es dagegen Anfangswerte des Zustands, für die der Zustand betragsmäßig gegen ∞ läuft, so heißt der Gleichgewichtspunkt $(\bar{\mathbf{x}}, \bar{u})$ **instabil**.

Um Kriterien für asymptotische Stabilität zu gewinnen, bemerken wir zunächst die Äquivalenz folgender Gleichungen (wobei $u(t) \equiv \bar{u}$ vorausgesetzt ist):

$$\underbrace{\dot{(\mathbf{x} - \bar{\mathbf{x}})}}_{=: \dot{\tilde{\mathbf{x}}}} = \dot{\mathbf{x}} = \mathbf{A} \mathbf{x} + \mathbf{b} u = \mathbf{A} \underbrace{(\mathbf{x} - \bar{\mathbf{x}})}_{=: \tilde{\mathbf{x}}} + \underbrace{\mathbf{A} \bar{\mathbf{x}} + \mathbf{b} u}_{= \mathbf{0}} \quad \text{wegen (3.91)}$$

$$\Leftrightarrow \dot{\tilde{\mathbf{x}}} = \mathbf{A} \tilde{\mathbf{x}} .$$

Die inhomogene DGL für den Zustand \mathbf{x} ist also (bei $u(t) \equiv \bar{u}$) äquivalent mit der homogenen DGL für die Abweichung $\tilde{\mathbf{x}}$ vom Gleichgewichtszustand. Folglich ist auch das Langzeitverhalten der jeweiligen Lösungen analog, d.h.

$$\mathbf{x}(t) \rightarrow \bar{\mathbf{x}} \quad \text{für alle } \mathbf{x}_0 \quad \Leftrightarrow \quad \tilde{\mathbf{x}} \rightarrow \mathbf{0} \quad \text{für alle } \tilde{\mathbf{x}}_0 := (\mathbf{x}_0 - \bar{\mathbf{x}}) \quad (3.92)$$

Ob ein lineares System asymptotisch stabil oder instabil ist, hängt also – anders, als die Definition der asymptotischen Stabilität vermuten lässt – nicht vom betrachteten Gleichgewichtspunkt ab, sondern ist eine Eigenschaft der homogenen Zustandsgleichung und damit der Eigenbewegung. Alle Gleichgewichtspunkte des Systems zeigen also dasselbe Stabilitätsverhalten.

Das laut (3.92) zu untersuchende Langzeitverhalten der freien Bewegung können wir mittels der Modenzerlegung (3.90) mühelos anhand der Eigenwerte beurteilen:

$$\begin{aligned} \text{System asymptotisch stabil} & \Leftrightarrow \text{alle Faktoren } e^{\delta_i t} \xrightarrow[t \rightarrow \infty]{} 0 \\ & \Leftrightarrow \operatorname{Re} \lambda_i = \delta_i < 0 \quad \text{für } i = 1, \dots, n \end{aligned} \quad (3.93)$$

$$\begin{aligned} \text{System instabil} & \Leftrightarrow \text{mindestens eine Funktion } e^{\delta_i t} \xrightarrow[t \rightarrow \infty]{} \infty \\ & \Leftrightarrow \text{für mindestens einen Eigenwert } \lambda_i \text{ gilt:} \\ & \operatorname{Re} \lambda_i = \delta_i > 0 \end{aligned} \quad (3.94)$$

Liegen also alle Eigenwerte des Systems in der linken Hälfte der komplexen Ebene, so ist das System asymptotisch stabil:

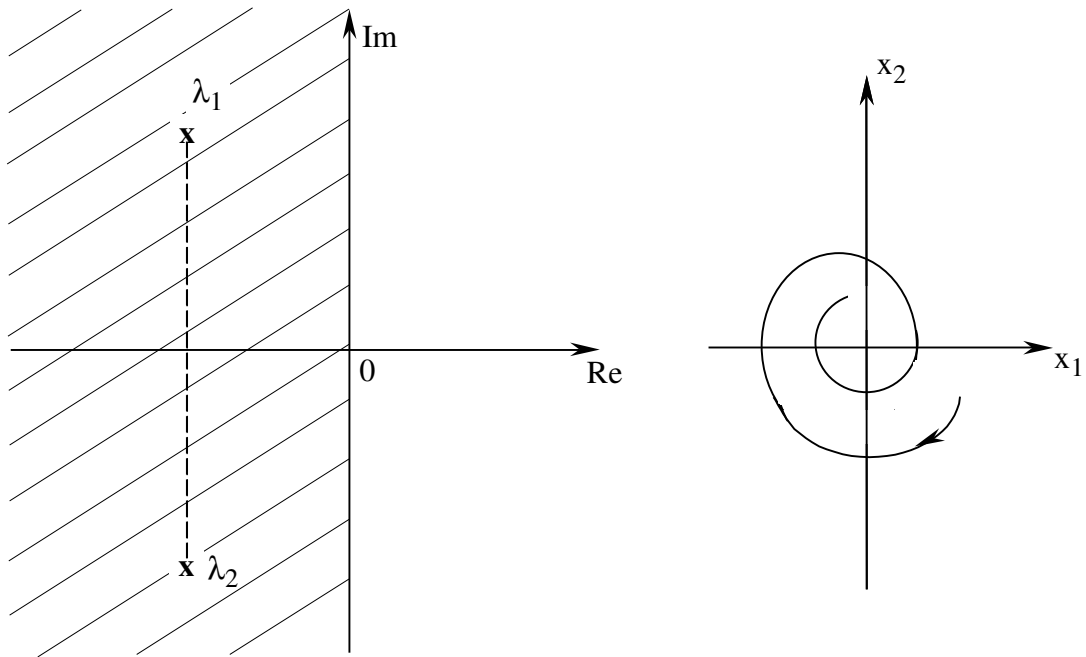


Abb. 17 :

asymptotisch stabiles System

zugehörige Zustandstrajektorie:
Strudel (in der Projektion:
gedämpfte Schwingung)

Liegt dagegen mindestens ein Eigenwert in der rechten Halbebene, so ist das System instabil:

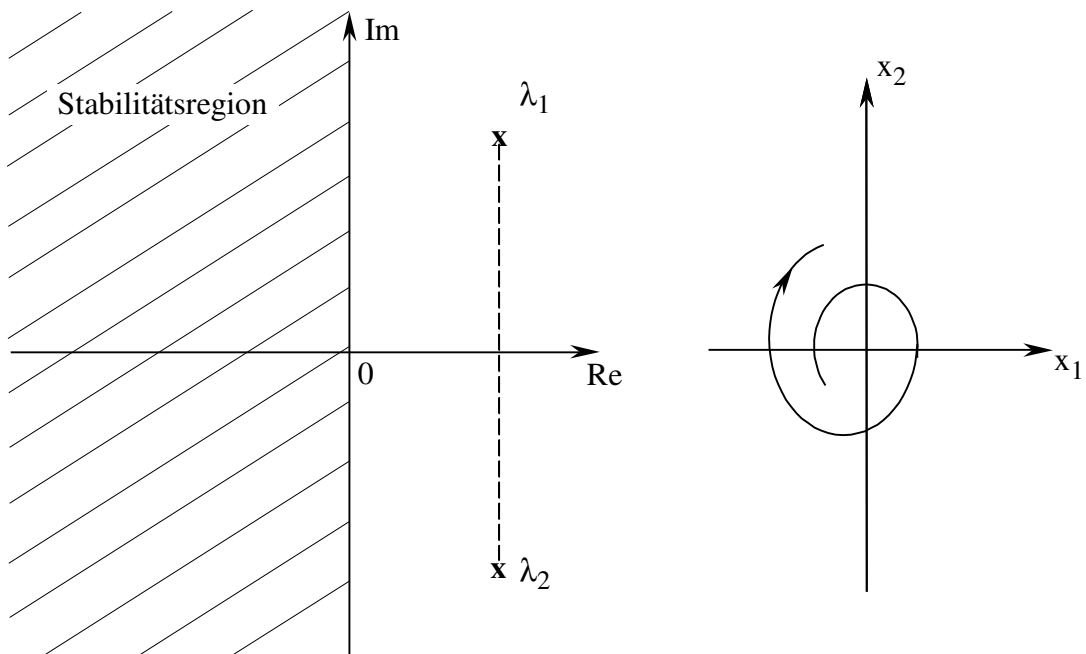


Abb. 18 :

Instabiles System

Zustandstrajektorie : Strudel
(in der Projektion: exponentiell
anwachsende Schwingung)

Wenn einer oder zwei Eigenwerte genau auf der imaginären Achse, alle anderen Eigenwerte aber links davon liegen, so ist das System „marginal stabil“: Der zu dieser Mode gehörige Term in $\mathbf{x}_{\text{frei}}(t)$ bleibt betragsmäßig stets gleich:

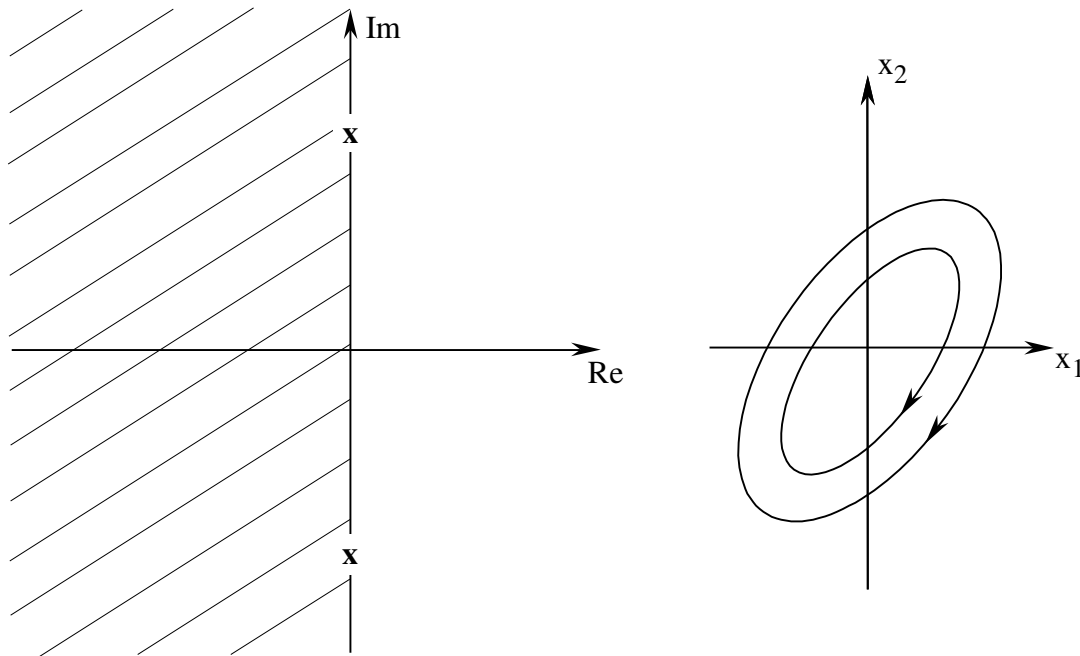


Abb. 19 :

Marginal stabiles System

Zustandstrajektorie : Wirbel
(in der Projektion: ungedämpfte
Schwingung)

Übrigens kann die asymptotische Stabilität eines Systems auch dann noch wie oben anhand der Lage der Eigenwerte in der komplexen Ebene abgelesen werden, wenn die Systemmatrix nicht diagonalisierbar ist, weil gewisse Eigenwerte identisch sind („zusammenfallen“ bzw. „übereinander liegen“).

Die Berechnung der Eigenwerte der Systemmatrix ist also ein wichtiges Hilfsmittel zur Analyse dynamischer Systeme. Häufig ist uns das zu betrachtende dynamische System in seiner Regelungsnormform, d.h. der aus der DGL n-ter Ordnung (in y) hervorgegangenen Zustandsbeschreibung, gegeben.

Die besondere (Frobenius-)Struktur der zugehörigen Systemmatrix [siehe (3.32)]

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

ermöglicht es, das charakteristische Polynom sofort hinzuschreiben:

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 \quad (3.95)$$

(Tipp zum Nachrechnen: Man entwickle $(\lambda \mathbf{I} - \mathbf{A})$ nach der letzten Zeile!)

3.4 Eingangs-Ausgangs-Verhalten

Aus der Bewegungsgleichung (3.78) für den – noch undiagonalisierten – Zustand \mathbf{x} kann mit Hilfe der Ausgabegleichung (3.29) sofort die Bewegungsgleichung für die Ausgangsgröße y bestimmt werden:

$$y(t) = \underbrace{\mathbf{c}^T \Phi(t) \mathbf{x}_0}_{y_{\text{frei}}(t)} + \underbrace{\int_0^t \mathbf{c}^T \Phi(t-\tau) \mathbf{b} u(\tau) d\tau + d \cdot u(t)}_{y_{\text{erzw.}}(t)} \quad (3.96)$$

Die Reaktion des Systemausgangs auf eine gegebene Eingangsgröße $u(t)$ wird als **Eingangs-Ausgangs-Verhalten** (kurz: E/A-Verhalten) oder auch **dynamisches Übertragungsverhalten** des Systems bezeichnet.

Gl. (3.96) macht deutlich, dass der [durch $u(t)$] erzwungenen Bewegung die Eigenbewegung des Systems überlagert ist, die vom Anfangswert \mathbf{x}_0 des Systemzustands abhängt. Bei Betrachtungen des E/A-Verhaltens müssen wir also von stets gleichen Anfangswerten \mathbf{x}_0 ausgehen.

Die natürliche Wahl ist

$$\mathbf{x}_0 = \mathbf{0}, \quad (3.97)$$

d.h. wir setzen bei E/A-Betrachtungen stets voraus, dass sich das System zur Zeit $t = 0$ im Ruhezustand $\mathbf{x}(0) = \mathbf{0}$ befindet. Das E/A-Verhalten eines durch das Zustandsraummodell (3.28), (3.29) beschriebenen Systems ist also direkt aus der Bewegungsgleichung für die erzwungene Bewegung $y_{\text{erzw.}}(t)$ ablesbar [wobei die Kennzeichnung „erzw.“ i.a. weggelassen wird]:

$$\text{E/A - Verhalten: } y(t) = \int_0^t \mathbf{c}^T \Phi(t-\tau) \mathbf{b} u(\tau) d\tau + d \cdot u(t) \quad (3.98)$$

Bemerkungen:

- (α) Wenn der Parameter d in der Ausgabegleichung nicht verschwindet, gibt der zweite Summand in (3.98) für $t = 0$ einen Beitrag

$$y(0) = d \cdot u(0).$$

Für $u(0) \neq 0$ folgt also das System / der Systemausgang der [im Zeitpunkt $t = 0$ vom Wert 0 auf $u(0)$ springenden] Eingabegröße ohne Verzögerung auf die mit dem Faktor d multiplizierte „Höhe“. Ein System mit $d \neq 0$ heißt daher **sprungfähig**. Bei der Herleitung des Zustandsraummodells aus der DGL n -ter Ordnung für y haben wir gesehen, dass $d \neq 0$

$$\left[\text{nämlich } d = b_n = \text{Koeffizient von } \frac{d^n u}{dt^n} \right]$$

genau dann gilt, wenn $u(t)$ und $y(t)$ in die DGL mit derselben höchsten Ableitung eingehen ($q = n$). Differentialgleichungen (3.15) mit $q = n$ beschreiben also sprunghafte Systeme.

- (β) Weil alle Transformationen, denen das Eingangssignal in (3.98) unterworfen wird, linear sind, kann die E/A-Linearität des Systems aus (3.98) unmittelbar abgelesen werden (vgl. Abschnitt 3.1.3).

Notiz zur **freien Bewegung des Systemausgangs** (System sei also anfänglich ausgelenkt!):

Von der Modenzerlegung der freien Zustandsbewegung (3.90) gelangt man mittels der Ausgabegleichung sofort zur Modenzerlegung der freien Bewegung des Systemausgangs:

$$y_{\text{frei}}(t) = e^{\lambda_1 t} z_1(0) \mathbf{c}^T \mathbf{v}^{(1)} + \dots + e^{\lambda_n t} z_n(0) \mathbf{c}^T \mathbf{v}^{(n)} \quad (3.99)$$

Die Bedingungen für das Abklingen der Eigenbewegung (nämlich: Lage der Eigenwerte in linker Halbebene) gelten also für y_{frei} in gleicher Weise wie für \mathbf{x}_{frei} .

3.5 Kennfunktionen des dynamischen Übertragungsverhaltens im Zeitbereich

Wir erregen nun das System durch ganz bestimmte Eingangssignale (nämlich sprungförmig bzw. Dirac-Impuls) und notieren die zugehörige Antwort des Systemausgangs. Die entstehenden **Kennfunktionen** beschreiben das dynamische Übertragungsverhalten auf eine graphisch sehr anschauliche Weise.

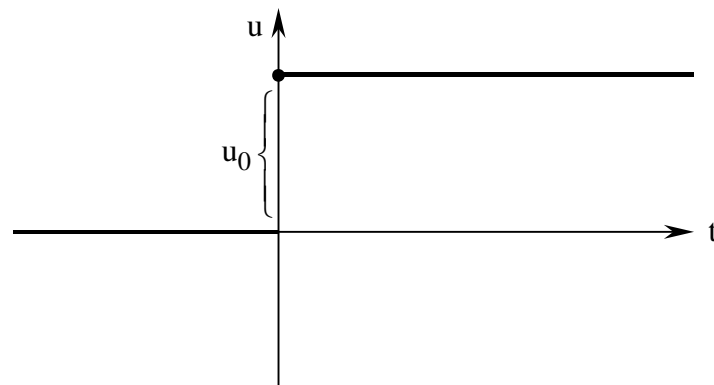
3.5.1 Übergangsfunktion

Die Übergangsfunktion beschreibt, wie das System auf eine sprungförmige Eingangsgröße reagiert.

$$\text{Eingangsgröße : } u(t) = \underset{\substack{\uparrow \\ \text{Sprunghöhe}}}{u_0} \cdot \sigma(t), \quad (3.100)$$

mit der Sprung- oder Heaviside-Funktion

$$\sigma(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases} \quad (3.101)$$



Setzt man dieses Eingangssignal in die E/A-Beziehung (3.98) ein, so erhält man die sog. **Sprungantwort**

$$y(t) = \int_0^t \underbrace{\mathbf{c}^T \Phi(t-\tau)}_{=: \tau'} \mathbf{b} u_0 d\tau + d \cdot u_0 \cdot \sigma(t) = \left(\int_0^t \mathbf{c}^T \Phi(\tau') \mathbf{b} d\tau' + d \cdot \sigma(t) \right) u_0$$

($\sigma(t)$ wird im folgenden aufgrund der Vereinbarung weggelassen, dass alle Signale für $t < 0$ verschwinden).

Für den Fall $u_0 = 1$ heißt die Sprungantwort **Übergangsfunktion** $h(t)$:

$$\text{Übergangsfunktion : } h(t) = \int_0^t \mathbf{c}^T \underset{\substack{\uparrow \\ e^{\mathbf{A}\tau}}}{\boldsymbol{\Phi}(\tau)} \mathbf{b} d\tau + d \quad (3.102)$$

Falls die Systemmatrix \mathbf{A} invertierbar ist, d.h. falls $\det \mathbf{A} \neq 0$, gilt

$$\int_0^t e^{\mathbf{A}\tau} d\tau = \mathbf{A}^{-1} e^{\mathbf{A}\tau} \bigg|_{\tau=0}^{\tau=t}$$

(Check: Man differenziere die Stammfunktion auf der rechten Gleichungsseite) und somit

$$h(t) = \mathbf{c}^T \mathbf{A}^{-1} e^{\mathbf{A}\tau} \mathbf{b} \bigg|_{\tau=0}^{\tau=t} + d$$

$$\Leftrightarrow h(t) = d - \mathbf{c}^T \mathbf{A}^{-1} \mathbf{b} + \mathbf{c}^T \mathbf{A}^{-1} e^{\mathbf{A}t} \mathbf{b}. \quad (3.103)$$

Wir interessieren uns nun für das Langzeitverhalten von $h(t)$. Die gesamte Zeitabhängigkeit von $h(t)$ steckt in der Übergangsmatrix $\boldsymbol{\Phi}(t) = e^{\mathbf{A}t}$. Um $\boldsymbol{\Phi}(t)$ in eine besser interpretierbare Form zu bringen, erinnern wir uns, dass wir die „Schwestermatrix“ $\tilde{\boldsymbol{\Phi}}(t)$, also die Übergangsmatrix in der Eigenbasis-Koordinatendarstellung, bereits aus Abschnitt 3.3.3 kennen:

$$\tilde{\boldsymbol{\Phi}}(t) = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}.$$

Und wie eine Matrix aus der Eigenbasis-Darstellung in die entsprechende Matrix in der Standardbasis-Darstellung umzurechnen ist, können wir folgendem Umrechnungsbeispiel [nämlich: Umrechnung zwischen der Systemmatrix \mathbf{A} (Standardbasis-Darstellung) und der „diagonalisierten Systemmatrix“ $\boldsymbol{\Lambda}$ (Eigenbasis-Darstellung)] entnehmen:

$$\mathbf{A} = \mathbf{V}^{-1} \boldsymbol{\Lambda} \mathbf{V} \quad \Leftrightarrow \quad \boldsymbol{\Lambda} = \mathbf{V} \mathbf{A} \mathbf{V}^{-1}. \quad (3.104)$$

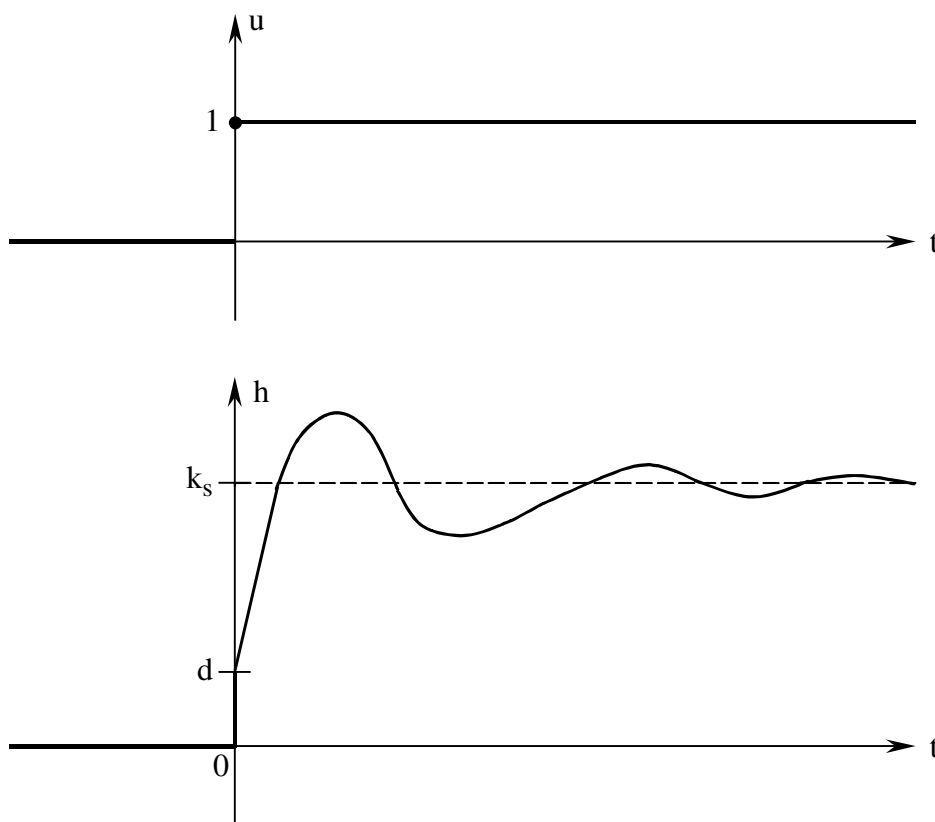
Also gilt in Analogie zu (3.104):

$$e^{\mathbf{A}t} = \boldsymbol{\Phi}(t) = \mathbf{V} \tilde{\boldsymbol{\Phi}}(t) \mathbf{V}^{-1} = \mathbf{V} \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} \mathbf{V}^{-1} \quad (3.105)$$

Nachdem es uns mit Hilfe der Diagonalisierung gelungen ist, die Übergangsmatrix e^{At} explizit zu berechnen, können wir aus Gl. (3.103) mühelos folgern: Für asymptotisch stabile Systeme verschwinden im Langzeitlimit $t \rightarrow \infty$ alle Terme $e^{\lambda_i t}$, so dass in (3.103) der Term $\propto e^{At}$ verschwindet. Die Übergangsfunktion $h(t)$ eines stabilen Systems nähert sich also asymptotisch einem konstanten Endwert

$$k_s := h(\infty) = -\mathbf{c}^T \mathbf{A}^{-1} \mathbf{b} + d, \quad (3.106)$$

genannt *statische Verstärkung* des Systems.



3.5.2 Impulsantwort oder Gewichtsfunktion

Die Impulsantwort oder Gewichtsfunktion gibt an, wie das System auf einen sehr kurzen Impuls am Systemeingang reagiert. „Sehr kurz“ bedeutet hier, dass die Impulsdauer ε klein ist gegenüber den maßgebenden Zeitkonstanten des Systems. Zur mathematischen Beschreibung solcher sehr kurzen Impulse gehen wir folgendermaßen vor:

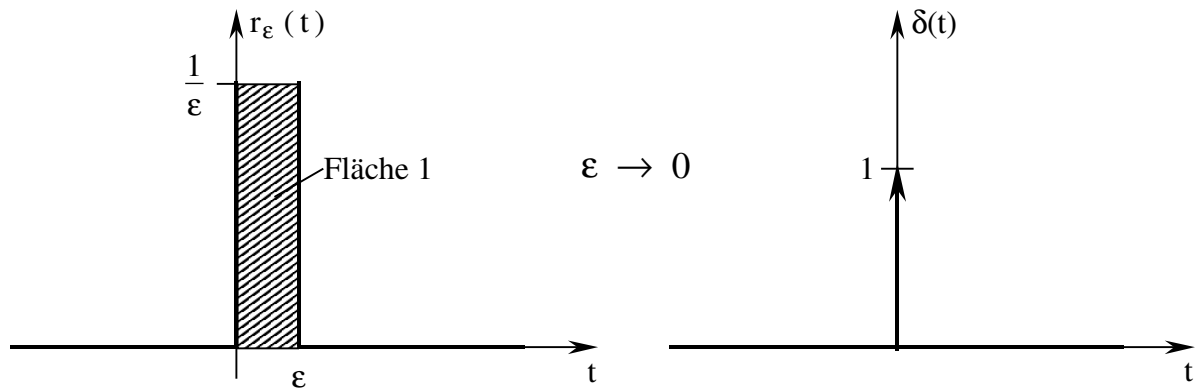


Abb. 20 : Definition und graphische Darstellung eines δ -Impulses

Abb. 20 zeigt einen Rechteckimpuls

$$r_\varepsilon(t) := \begin{cases} \frac{1}{\varepsilon} & \text{für } t \in [0, \varepsilon] \\ 0 & \text{sonst,} \end{cases}$$

der eine Fläche der Größe 1 einschließt. Würde man das System durch einen solchen Rechteckimpuls erregen, wäre freilich die Ausgangsgröße vom Parameter ε abhängig. Um diese Abhängigkeit loszuwerden, arbeiten wir mit einer mathematischen Idealisierung, dem sog. Dirac-Impuls (auch Einheitsimpuls oder δ -Impuls genannt). Dieser entsteht aus r_ε durch den Grenzübergang $\varepsilon \rightarrow 0$ (also als Grenzwert immer kürzerer und gleichzeitig höherer Rechteckimpulse):

$$\delta(t) := \lim_{\varepsilon \rightarrow 0} r_\varepsilon(t) \quad (3.107)$$

Streng genommen ist der δ -Impuls keine Funktion im klassischen Sinne, sondern eine verallgemeinerte Funktion oder Distribution. (Dennoch spricht man oft salopp von der „ δ -Funktion“.) Für das Rechnen mit δ -Impulsen sind folgende Eigenschaften, die unmittelbar aus der Definition (3.107) folgen, wichtig:

$$\delta(t) = 0 \quad \text{für } t \neq 0 \quad (3.108)$$

$$\int_{-\infty}^t \delta(\tau) d\tau = \sigma(t) \quad (\text{Sprungfunktion}) \quad (3.109)$$

Das graphische Symbol eines δ -Impulses, ein Pfeil der Länge 1 (siehe Abb. 20), erinnert daran, dass gilt [folgt sofort aus (3.109)]

$$\int_{-\infty}^{+\infty} \delta(\tau) d\tau = 1.$$

Als Antwort des Systemausgangs auf eine impulsförmige Erregung $u(t) = \delta(t)$ liefert die E/A-Beziehung (3.98):

$$y(t) = \int_0^t \mathbf{c}^T \Phi(t-\tau) \mathbf{b} \delta(\tau) d\tau + d \cdot \delta(t)$$

Da der δ -Impuls für alle Zeiten außerhalb eines winzigen, die Null umschließenden Intervalls gleich Null ist, können wir statt

$$\int_0^t \dots \text{ schreiben } \int_{-0}^{+0} \dots$$

In diesem Intervall gilt $\Phi(t-\tau) = \Phi(t)$, d.h. $\Phi(t-\tau)$ ist eine (bezüglich der Integrationsvariablen τ) konstante Größe und kann daher vor das Integral gezogen werden

$$\begin{aligned} \Rightarrow y(t) &= \mathbf{c}^T \Phi(t) \mathbf{b} \cdot \underbrace{\int_{-0}^{+0} \delta(\tau) d\tau}_{=1} + d \cdot \delta(t) \\ &= 1 \text{ laut (3.109)} \end{aligned}$$

Die **Impulsantwort** wird auch **Gewichtsfunktion** (Name wird binnen kurzem klar!) genannt und durch $g(t)$ symbolisiert:

Gewichtsfunktion: $g(t) = \mathbf{c}^T \Phi(t) \mathbf{b} + d \delta(t) = \mathbf{c}^T e^{At} \mathbf{b} + d \delta(t)$

 (3.110)

Falls die Systemmatrix \mathbf{A} diagonalisierbar ist, können wir die explizite Formel (3.105) für die Übergangsmatrix e^{At} einsetzen und erhalten

$$g(t) = \underbrace{\mathbf{c}^T \mathbf{V}}_{\tilde{\mathbf{c}}^T} \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} \underbrace{\mathbf{V}^{-1} \mathbf{b}}_{\tilde{\mathbf{b}}} + d \delta(t),$$

wobei uns $\tilde{\mathbf{c}}$ und $\tilde{\mathbf{b}}$ bereits aus Abschnitt 3.3.3 bekannt sind. Komponentenweise ausgeschrieben (\tilde{c}_i sei die i -te Komponente des Vektors $\tilde{\mathbf{c}}$) ergibt obige Gleichung die **Modenzerlegung** der **Gewichtsfunktion**:

$$g(t) = \sum_{i=1}^n \underbrace{\tilde{c}_i \tilde{b}_i}_{=: g_i} e^{\lambda_i t} + d \delta(t) \quad (3.111)$$

Eine typische Impulsantwort eines stabilen, nicht sprungfähigen (d.h. $d = 0$) Systems zweiter Ordnung ist in Abb. 21 skizziert:

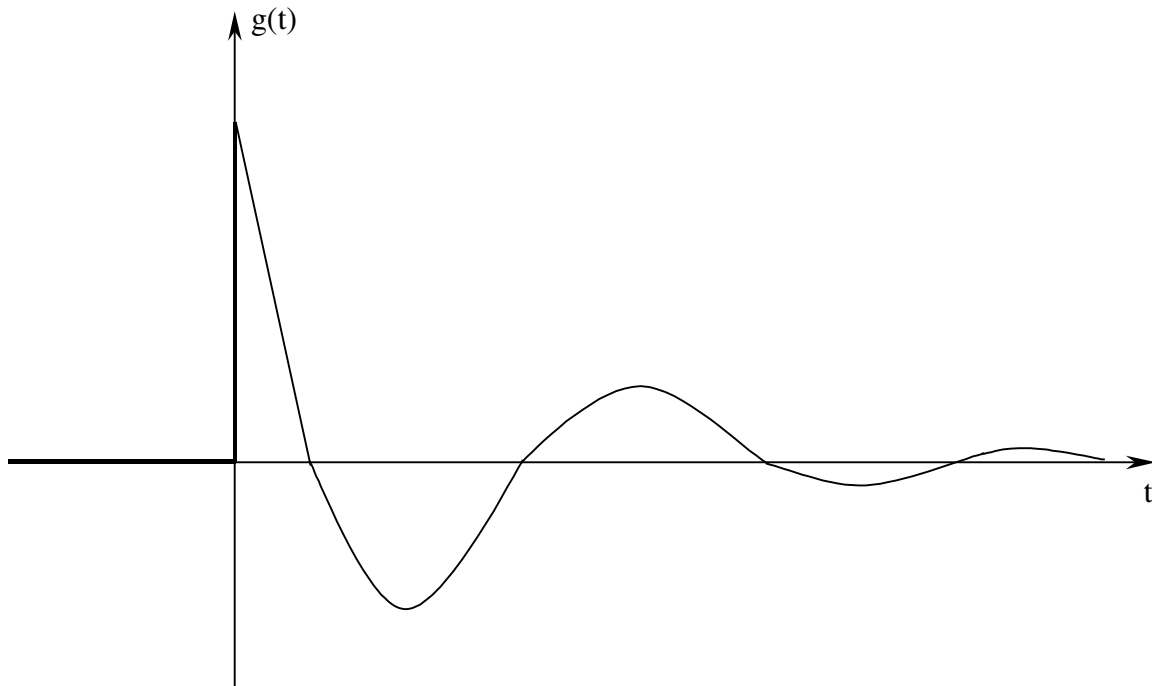


Abb. 21 : Impulsantwort eines Systems zweiter Ordnung

3.5.3 Zusammenhang zwischen Sprung- und Impulsantwort

Ein expliziter Vergleich von Übergangs- und Gewichtsfunktion zeigt:

$$\begin{array}{ccc}
 h(t) = \int_0^t \mathbf{c}^T e^{\mathbf{A}\tau} \mathbf{b} d\tau & + & d \cdot \sigma(t) \\
 \begin{array}{c} \uparrow \\ \int_0^t \dots d\tau \end{array} & \begin{array}{c} \downarrow \frac{d}{dt} \end{array} & \begin{array}{c} \uparrow \\ \int_0^t \dots d\tau \end{array} & \begin{array}{c} \downarrow \frac{d''}{dt} \end{array} \\
 & & \text{[siehe auch (3.109)]} & \\
 g(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{b} & + & d \cdot \delta(t) &
 \end{array}$$

Also lassen sich Übergangs- und Gewichtsfunktion entsprechend

$$h(t) = \int_0^t g(\tau) d\tau \quad (3.112)$$

$$g(t) = \frac{d h(t)}{dt} \quad (3.113)$$

ineinander umrechnen.

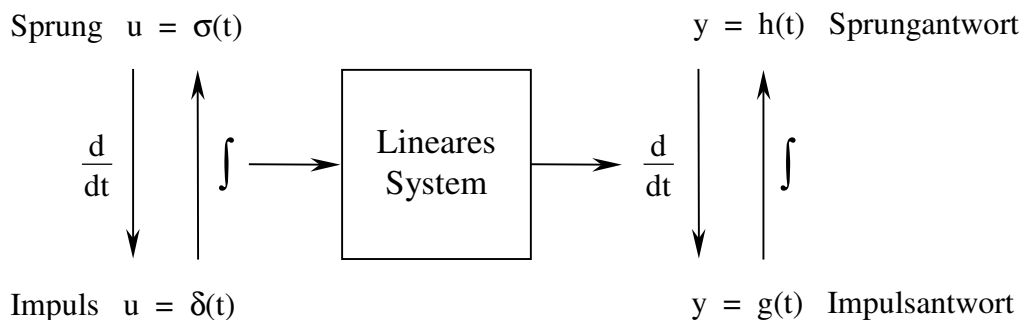
Gl. (3.112) liefert uns eine zweite Beziehung für die statische Verstärkung des Systems:

$$k_s = h(\infty) = \int_0^{\infty} g(\tau) d\tau \quad (3.114)$$

Dass die Impulsantwort $g(t)$ die Ableitung der Sprungantwort $h(t)$ ist, überrascht uns letztlich nicht. Schließlich ist ja auch das Impuls-Eingangssignal die zeitliche Ableitung (streng genommen im Sinne der verallgemeinerten Funktionen) des sprungförmigen Eingangssignals, und wir wissen ja bereits [siehe (3.37)], dass gilt

$$u(t) \mapsto y(t) \quad \Rightarrow \quad \dot{u}(t) \mapsto \dot{y}(t)$$

Die folgende Skizze fasst den Zusammenhang zwischen Impuls- und Sprungantwort (bzw. zwischen Gewichtsfunktion und Übergangsfunktion) zusammen.



3.5.4 Darstellung des E/A-Verhaltens mit Hilfe der Gewichtsfunktion

Bei einem nicht sprungfähigen System (d.h. $d = 0$) entdecken wir in der E/A-Beziehung (3.98) mühelos die Gewichtsfunktion:

$$y(t) = \int_0^t \underbrace{\mathbf{c}^T \boldsymbol{\Phi}(t-\tau) \mathbf{b}}_{g(t-\tau)} u(\tau) d\tau$$

Um zu sehen, dass diese Beziehung auch für sprungfähige Systeme gilt, nutzen wir die Eigenschaften (3.108), (3.109) des δ -Impulses aus:

$$\begin{aligned}
 \int_0^t d \cdot \underbrace{\delta(t-\tau)}_{\tau'} u(\tau) d\tau & \stackrel{[\text{Substitution } t-\tau=\tau']}{=} \int_0^t d u(t-\tau') \delta(\tau') d\tau' \\
 & \quad \quad \quad \uparrow \\
 & \quad \quad \quad \text{Beiträge nur für } \int_{-0}^{+0} \dots \\
 & = d u(t) \underbrace{\int_{-0}^{+0} \delta(\tau') d\tau'}_1 \\
 \Rightarrow y &= \int_0^t \mathbf{c}^T \boldsymbol{\Phi}(t-\tau) \mathbf{b} u(\tau) d\tau + \underbrace{d \cdot u(t)}_{\int_0^t d \cdot \delta(t-\tau) u(\tau) d\tau} \\
 y &= \int_0^t \underbrace{\left[\mathbf{c}^T \boldsymbol{\Phi}(t-\tau) \mathbf{b} + d \cdot \delta(t-\tau) \right]}_{g(t-\tau)} u(\tau) d\tau
 \end{aligned}$$

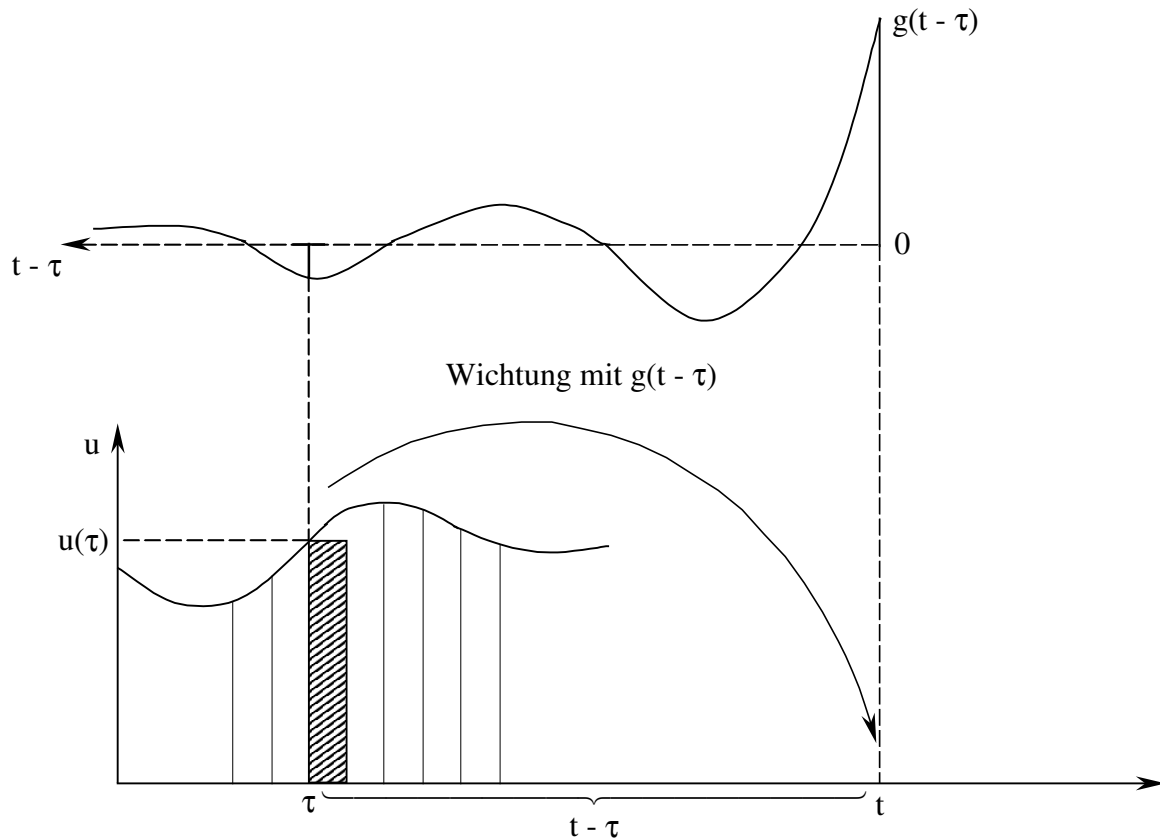
Die rechte Seite stellt ein **Faltungintegral** dar, das durch das Operatorsymbol $*$ abgekürzt wird:

$$g(t) * u(t) := \int_0^t g(t-\tau) u(\tau) d\tau \stackrel{\left[\begin{array}{c} \text{Substitution} \\ t-\tau=\tau' \end{array} \right]}{=} \int_0^t g(\tau) u(t-\tau) d\tau \quad (3.115)$$

Damit können wir das E/A-Verhalten auch schreiben als

$\text{E/A-Verhalten: } y(t) = \int_0^t g(t-\tau) u(\tau) d\tau = g(t) * u(t)$	(3.116)
--	---------

Die Gewichtsfunktion charakterisiert also das dynamische Übertragungsverhalten vollständig. Ist $g(t)$ bekannt, so kann das Ausgangssignal für beliebige Eingangsgrößen berechnet werden (vorausgesetzt, dass das System zur Zeit $t = 0$ in Ruhe war). Formel (3.116) lässt sich leicht interpretieren, wenn wir (wie in Abschnitt 3.3.1) das Eingangssignal $u(t)$ wieder in einzelne Rechteckimpulse zerlegen:



Die aktuelle (d.h. zur Zeit t auftretende) Systemantwort auf einen vergangenen Eingabeimpuls (Zeit τ , Dauer Δ) lautet:

$$\underbrace{g(t - \tau)}_{\text{Antwort auf Impuls der Fläche 1}} \cdot \underbrace{u(\tau) \cdot \Delta}_{\text{Fläche des Rechteckimpulses}}$$

Mit $\Delta \rightarrow 0$ gibt die Summe der Systemantworten auf alle Rechteckimpulse der Zerlegung genau das obige Faltungsintegral. Die Impulsantwort $g(t - \tau)$ gibt laut (3.116) das **Gewicht** an, mit der die vergangene Eingangsgröße $u(\tau)$ in die (zur Zeit t) aktuelle Ausgangsgröße $y(t)$ eingeht \rightarrow Name „Gewichtsfunktion“!

Da die Gewichtsfunktion typischerweise für kleine Zeiten groß und für große Zeiten klein ist, wird die Ausgangsgröße stark von der „jüngsten Vergangenheit“ und wenig von den sehr weit zurückliegenden Werten der Eingangsgröße bestimmt.

3.6 *Übergangsverhalten und stationäres Verhalten*

Nach dem Vorbild der Modenzerlegung der freien Bewegung des Systemausgangs, $y_{\text{frei}}(t)$, wollen wir nun eine ähnliche Darstellung für die erzwungene Bewegung ableiten. Wir setzen wieder $\mathbf{x}_0 = \mathbf{0}$ voraus, so dass $y(t) = y_{\text{erzw}}(t)$ gilt.

Die betrachteten Eingangsgrößen seien von der Form

$$u(t) = \sum_{j=1}^m u_j e^{\mu_j t}, \quad (3.117)$$

wobei die Faktoren u_j und die Parameter μ_j im Exponenten reell sind oder als komplex konjugierte Paare auftreten. Damit ist $u(t)$ eine reellwertige Funktion. Obige Klasse von Eingangsgrößen schließt beispielsweise die sprungförmige Eingangsgröße

$$u(t) = u_1 e^{0 \cdot t} = u_1 \cdot \sigma(t)$$

und die cosinusförmige Eingangsgröße

$$u(t) = e^{j\omega t} + e^{-j\omega t} = 2 \cos(\omega t)$$

ein. Mit Hilfe der Fourieranalyse können alle praktisch interessanten Eingangssignale in der Form (3.117) dargestellt werden, wobei möglicherweise unendlich viele Summanden auftreten.

Zur Vereinfachung der Rechnung nehmen wir weiterhin an, dass $\lambda_i \neq \mu_j$ für alle i, j und dass das System nicht sprungfähig ist (also $d = 0$).

Für die Eingangsgröße (3.117) untersuchen wir nun die erzwungene Ausgangsbewegung $y(t) = g(t) * u(t)$, wobei wir von der Modenzerlegung (3.111) der Gewichtsfunktion Gebrauch machen.

$$\begin{aligned}
 y(t) &= \int_0^t \left(\sum_{i=1}^n g_i e^{\lambda_i(t-\tau)} \right) \left(\sum_{j=1}^m u_j e^{\mu_j \tau} \right) d\tau \\
 &= \sum_{i=1}^n \sum_{j=1}^m g_i u_j e^{\lambda_i t} \underbrace{\int_0^t e^{(\mu_j - \lambda_i) \tau} d\tau}_{\frac{1}{\mu_j - \lambda_i} e^{(\mu_j - \lambda_i) \tau} \Big|_{\tau=0}^{\tau=t}} \\
 &= \sum_{i=1}^n \sum_{j=1}^m \frac{g_i u_j}{\lambda_i - \mu_j} e^{\lambda_i t} + \sum_{i=1}^n \sum_{j=1}^m \frac{g_i u_j}{\mu_j - \lambda_i} e^{\mu_j t} \\
 \Rightarrow y_{\text{erzw}}(t) &= \underbrace{\sum_{i=1}^n g_i \left(\sum_{j=1}^m \frac{u_j}{\lambda_i - \mu_j} \right) e^{\lambda_i t}}_{y_{\ddot{u}}(t)} + \underbrace{\sum_{j=1}^m u_j \left(\sum_{i=1}^n \frac{g_i}{\mu_j - \lambda_i} \right) e^{\mu_j t}}_{y_s(t)} \\
 (3.118)
 \end{aligned}$$

Bei asymptotisch stabilen Systemen klingen alle Moden $e^{\lambda_i t}$ für $t \rightarrow \infty$ ab, so dass im Langzeitlimit nur noch der Term y_s übrigbleibt:

$$y_{\text{erzw}}(t) \rightarrow y_s(t) \quad \text{für } t \rightarrow \infty. \quad (3.119)$$

$y_s(t)$ bezeichnet also das **stationäre Verhalten** (= den eingeschwungenen Zustand), $y_{\ddot{u}}(t)$ das **Übergangsverhalten** (= den Einschwingvorgang).

Zu $y_{\ddot{u}}(t)$: Das Übergangsverhalten entsteht, weil die Eigenvorgänge des Systems durch die Eingangsgröße $u(t)$ angeregt werden. Je ähnlicher die in $u(t)$ vorkommenden Terme $e^{\mu_j t}$ den Eigenmoden $e^{\lambda_i t}$ des Systems sind, desto stärker die Anregung $\left[\text{siehe Faktor } \sum_j \frac{1}{\lambda_i - \mu_j} \right]$.

Falls $g_i = 0$ für einen Index i , so wird die entsprechende Mode überhaupt nicht angeregt, welches Eingangssignal wir auch immer anlegen. $g_i = 0$ kommt zustande, wenn die Mode entweder generell nicht durch die Eingangsgröße anregbar ist ($\tilde{b}_i = 0$) oder wenn sie keinen Einfluss auf die Ausgangsgröße hat ($\tilde{c}_i = 0$).

Zu $y_s(t)$: Das stationäre Verhalten setzt sich aus e-Funktionen $e^{\mu_j t}$ zusammen, die ausschließlich durch die Eingangsgröße vorgegeben sind. Das System verändert jedoch die „Wichtungen“, mit denen diese e-Funktionen in der Summe auftreten. Anstelle von u_j steht in $y(t)$ die Wichtung $y_{sj} \cdot u_j$.

Von besonderem Interesse ist der Fall, dass ein Term $e^{\mu_j t}$ zwar im Eingangssignal $u(t)$, nicht jedoch im Ausgangssignal $y_{\text{erzw}}(t)$ auftritt. Laut Gl. (3.118) ist dies genau dann der Fall, wenn $y_{sj} = 0$ gilt, wenn also für μ_j die Bedingung

$$\text{Nullstelle } \mu_j : \sum_{i=1}^n \frac{g_i}{\mu_j - \lambda_i} = 0 \quad (3.120)$$

erfüllt ist. Ein solcher Wert μ_j beschreibt eine Funktion $e^{\mu_j t}$, die vom System nicht übertragen wird. Er wird deshalb als **Nullstelle** bezeichnet. Laut Gl. (3.120) hängt eine Nullstelle ausschließlich von den durch g_i und λ_i beschriebenen Systemeigenschaften, nicht von der Amplitude u_j der Eingangsgröße ab.

Die oben diskutierte Zerlegung der erzwungenen Bewegung in Übergangsverhalten und stationäres Verhalten ist für den Reglerentwurf sehr wichtig. Der Sollwert wird nur dann ohne bleibende Regelabweichung erreicht, wenn das stationäre Verhalten $y_s(t)$ des Regelkreises genau mit dem Führungssignal $w(t)$ übereinstimmt [$y_s(t) = w(t)$]. Für die Bewertung des dynamischen Verhaltens des Regelkreises ist das Übergangsverhalten $y_{\ddot{u}}(t)$ maßgebend, denn dieses beschreibt, wie lange es dauert, bis der Sollwert erreicht ist.

Kapitel 4 : *Beschreibung dynamischer Systeme im Frequenzbereich*

Fast alle Rechnungen der Regelungstechnik lassen sich mit Hilfe der Laplace-Transformation stark vereinfachen. Ein Signal Laplace-zu-transformieren heißt, es in auf- oder abklingende Sinusschwingungen zu zerlegen. Die Übertragung dieser „Elementarsignale“ durch das lineare System ist sehr einfach zu berechnen. Und das Superpositionsprinzip erlaubt es, aus den Einzelantworten auf die Elementarsignale durch Superposition die Systemantwort auf das gesamte Eingangssignal zu gewinnen.

Wir können die Systemantwort also in drei Schritten berechnen:

- 1) Zerlegung der Eingangsgröße $u(t)$ in sinusförmige Elementarsignale (durch Laplace-Transformation).
- 2) Getrennte Berechnung der Systemantworten für jeden einzelnen sinusförmigen Anteil von $u(t)$.
- 3) Bestimmung der Ausgangsgröße $y(t)$ durch Überlagerung aller berechneten Systemantworten (durch Laplace-Rücktransformation).

4.1 *Übertragungsfunktion*

Betrachten wir die Übertragung einer Exponentialfunktion

$$u(t) = e^{st} = e^{(\delta + j\omega)t} = e^{\delta t} \cdot \underbrace{e^{j\omega t}}_{\cos \omega t + j \sin \omega t}$$

\uparrow
auf- oder abklingende
Amplitude

Aus der E/A-Beschreibung mittels des Faltungsintegrals erhalten wir

$$\begin{aligned} y(t) &= \int_0^t g(t-\tau) u(\tau) d\tau = \int_0^t g(\tau) u(t-\tau) d\tau \\ &= \int_0^t g(\tau) e^{s(t-\tau)} d\tau \\ &= \left(\int_0^t g(\tau) e^{-s\tau} d\tau \right) \cdot e^{st} \\ &= \underbrace{\left(\int_0^\infty g(\tau) e^{-s\tau} d\tau \right)}_{=: G(s)} \cdot e^{st} + \underbrace{\left(- \int_t^\infty g(\tau) e^{-\delta\tau} d\tau \right)}_{y_{\ddot{u}}(t)} e^{st} \end{aligned}$$

Der zweite Term verschwindet für $t \rightarrow \infty$ wegen des immer kleiner werdenden Integrationsintervalls und beschreibt somit das Übergangsverhalten des Systems.

Für unsere Betrachtungen ist also nur der erste Summand, der das stationäre Verhalten beschreibt, maßgebend:

$$y(t) = \left(\int_0^\infty g(\tau) e^{-s\tau} d\tau \right) \cdot e^{st} = G(s) \cdot e^{st}$$

Eine Exponentialfunktion (= auf- oder abklingende Sinusschwingung) am Systemeingang führt also zu derselben Exponentialfunktion am Systemausgang, lediglich multipliziert mit einem komplexen Faktor $G(s)$. Die funktionelle Abhängigkeit $G(s)$ dieses Übertragungsfaktors vom Exponenten $s = \delta + j\omega$ des übertragenen Elementarsignals heißt **Übertragungsfunktion**.

4.2 Berechnung der Systemantwort durch Laplace-Transformation

4.2.1 Definition und Eigenschaften der Laplace-Transformation

Der soeben gefundene Zusammenhang zwischen der Gewichtsfunktion $g(t)$ und der Übertragungsfunktion $G(s)$ ist genau die Beziehung zwischen einer Zeitfunktion und ihrer Laplace-Transformierten. Die Laplace-Transformation ordnet einer Zeitfunktion $f(t)$ für $t \geq 0$ umkehrbar eindeutig – d.h. ohne Informationsverlust – eine Frequenzbereichsfunktion $F(s)$ zu. $F(s)$ ist eine komplexwertige Funktion der komplexen Frequenz

$$s = \delta + j\omega.$$

Die Deutung der komplexen Variablen s als „Frequenz“ wird anhand der Exponentialfunktion

$$e^{st} = e^{\delta t} \cdot e^{j\omega t} = e^{\delta t} (\cos \omega t + j \sin \omega t)$$

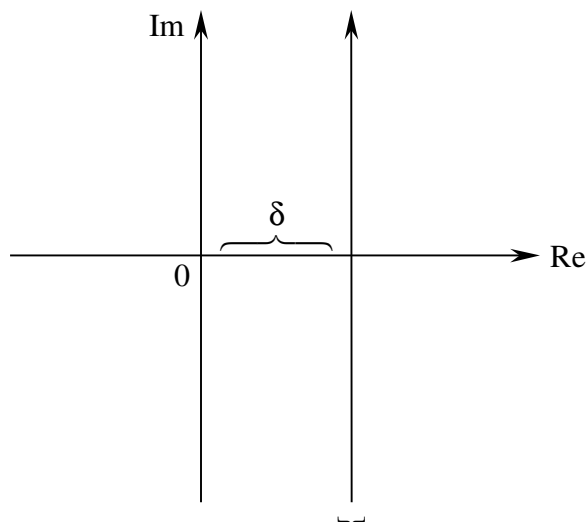
anschaulich: Der Realteil von e^{st} ist eine mit $e^{\delta t}$ auf- bzw. abklingende Kosinusschwingung der Frequenz ω .

$$\text{Laplace-Transformation: } \mathcal{L}[f(t)] := \int_{-0}^{\infty} f(t) e^{-st} dt = F(s)$$

 (4.1)

$$\text{Laplace-Rücktransformation: } \mathcal{L}^{-1}[F(s)] := \frac{1}{2\pi j} \int_{\delta-j\infty}^{\delta+j\infty} F(s) e^{st} ds = f(t)$$

 (4.2)



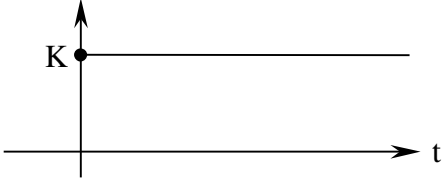
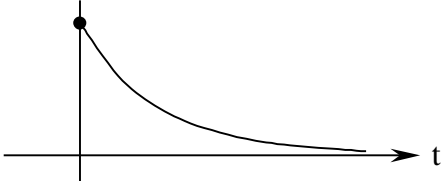
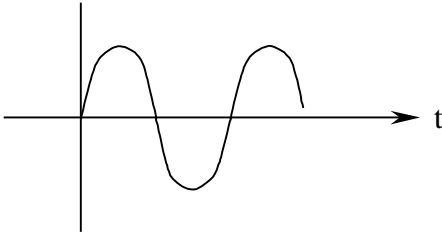
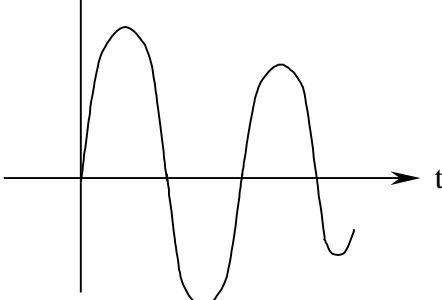
Integrationsweg bei der Laplace-Rücktransformation

Man schreibt für diese Zuordnung von Zeit- und Frequenzbereichsfunktion symbolisch auch

$$f(t) \quad \circ \longrightarrow \bullet \quad F(s) \quad (4.3)$$

Wenn man zur Laplace-(Rück-)Transformation jedesmal die Definitionsintegrale berechnen müßte, ginge der Gewinn an Rechenvereinfachung größtenteils wieder verloren. Zum Glück kann man für die gängigen Funktionen die zusammengehörigen Funktionspaare aus **Korrespondenztabelle** entnehmen.

Beispiele solcher Funktionenpaare :

Zeitbereich $f(t)$	$\xrightarrow{\mathcal{L}}$ $\xleftarrow{\mathcal{L}^{-1}}$	Frequenzbereich $F(s)$
	Sprungfunktion $K \cdot \sigma(t)$	$\frac{K}{s}$
	$e^{-at} \cdot \sigma(t)$	$\frac{1}{s + a}$
	$\sin \omega t \cdot \sigma(t)$	$\frac{\omega}{s^2 + \omega^2}$
	$e^{-at} \cdot \sin \omega t \cdot \sigma(t)$	$\frac{\omega}{(s + a)^2 + \omega^2}$

Viele der Rechenoperationen zur Bestimmung der Systemantwort sind im Frequenzbereich um Grade einfacher durchzuführen. Dies liegt im wesentlichen an folgenden Rechenregeln zur Laplace-Transformation:

- (i) Die Laplace-Transformation ist eine **lineare** Transformation, d.h.

$$c_1 f_1(t) + c_2 f_2(t) \quad \longleftrightarrow \quad c_1 F_1(s) + c_2 F_2(s) \quad (4.4)$$

- (ii) **Differentiationssatz:**

$$\dot{f}(t) \quad \longleftrightarrow \quad s \cdot F(s) - f(0_-) \quad (4.5)$$

und für höhere Ableitungen

$$\frac{d^k f}{dt^k} \quad \longleftrightarrow \quad s^k F(s) - s^{k-1} f(0_-) - s^{k-2} \dot{f}(0_-) - \dots - \frac{d^{k-1} f}{dt^{k-1}} \Big|_{t=0_-} \quad (4.6)$$

Sind alle Anfangsbedingungen $f(0_-)$, $\dot{f}(0_-)$... gleich Null, so entspricht der Differentiation im Zeitbereich eine Multiplikation der Laplacetransformierten mit s . Eine gewöhnliche DGL n-ter Ordnung geht somit über in eine Polynomgleichung n-ter Ordnung, was eine bedeutsame Vereinfachung mit sich bringt.

- (iii) **Faltungssatz:**

Die Faltung zweier Zeitfunktionen wird im Frequenzbereich zu einer simplen Multiplikation:

$$f_1(t) * f_2(t) \quad \longleftrightarrow \quad F_1(s) \cdot F_2(s) \quad (4.7)$$

4.2.2 Berechnung der Übertragungsfunktion

Wenn wir auf die bekannte E/A-Beziehung (3.116)

$$y(t) = \underbrace{g(t)}_{\text{Gewichtsfunktion}} * u(t)$$

den Faltungssatz der Laplace-Transformation anwenden, so ergibt sich

$$\underbrace{Y(s)}_{\text{Laplacetransformierte der Ausgangsgröße}} = G(s) \cdot \underbrace{U(s)}_{\text{Laplacetransformierte der Eingangsgröße}} \quad (4.8)$$

mit

$$G(s) \quad \longleftrightarrow \quad g(t), \quad (4.9)$$

$G(s)$ als die Laplace-Transformierte der Gewichtsfunktion ist nicht anderes als die bereits in Abschnitt 4.1 eingeführte Übertragungsfunktion des Systems.

Da die E/A-Beziehung (3.116) nur die erzwungene Bewegung des Systems berücksichtigt, bedeutet Gl. (4.8)

$$y_{\text{erzw}}(t) \quad \text{O} \longrightarrow \bullet \quad G(s) \cdot U(s) \quad (4.10)$$

Das Systemverhalten kann also sehr elegant im Frequenzbereich berechnet werden: Für ein gegebenes Eingangssignal $u(t)$ bestimmt man durch Laplace-Trafo $U(s)$, daraus $Y(s)$ und durch Rücktransformation schließlich $y_{\text{erzw}}(t)$.

Der skizzierte Rechenweg setzt voraus, dass wir die Übertragungsfunktion $G(s)$ des Systems kennen. $G(s)$ kann auf mehrere Weisen berechnet werden.

(A) Berechnung der Übertragungsfunktion aus der Gewichtsfunktion

Die Berechnung von $G(s)$ anhand der definierenden Gleichung

$$G(s) = \mathcal{L} [g(t)]$$

ist nur dann von praktischer Relevanz, wenn die Gewichtsfunktion in Form ihrer Modenzerlegung vorliegt,

$$g(t) = \sum_{i=1}^n g_i e^{\lambda_i t} + d \delta(t) .$$

Da laut Korrespondenztabelle gilt

$$e^{\lambda_i t} \quad \text{O} \longrightarrow \bullet \quad \frac{1}{s - \lambda_i}$$

$$\delta(t) \quad \text{O} \longrightarrow \bullet \quad 1 ,$$

folgt aus der Linearität der Laplace-Transformation

$$\boxed{G(s) = \sum_{i=1}^n \frac{g_i}{s - \lambda_i} + d} \quad (4.11)$$

Hier ist die Übertragungsfunktion durch ihre Partialbrüche

$$\frac{g_i}{s - \lambda_i}$$

dargestellt. Eine solche Darstellung ist genau dann möglich, wenn die Systemmatrix **A** des Zustandsraummodells **diagonalisierbar** ist, denn genau dann gibt es die oben verwendete Modenzerlegungs-Darstellung der Gewichtsfunktion.

(B) *Berechnung der Übertragungsfunktion aus der Differentialgleichung*

Betrachten wir nun den Fall, dass das dynamische System in Form einer DGL n-ter Ordnung vorliegt

$$a_n \frac{d^n y}{dt^n} + \dots + a_1 \dot{y} + a_0 y(t) = b_q \frac{d^q u}{dt^q} + \dots + b_1 \dot{u} + b_0 u(t) \quad (4.12)$$

Da wir uns für die E/A-Beziehung und damit für die erzwungene Bewegung interessieren, gehen wir von verschwindenden Anfangsbedingungen aus:

$$y(0) = 0, \quad \dot{y}(0) = 0, \quad \dots, \quad \frac{d^{n-1} y}{dt^{n-1}}(0) = 0.$$

Nun unterziehen wir beide Seiten der DGL (4.12) der Laplace-Transformation. Aufgrund der Linearität der Laplace-Trafo und mit Hilfe des Differentiationssatzes ergibt sich

$$Y(s) (a_n s^n + \dots + a_1 s + a_0) = U(s) (b_q s^q + \dots + b_1 s + b_0) \quad (4.13)$$

und damit für die Übertragungsfunktion

$$G(s) \stackrel{(4.8)}{=} \frac{Y(s)}{U(s)} \stackrel{(4.13)}{=} \frac{b_q s^q + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0}$$

(4.14)

Da Zähler und Nenner der Übertragungsfunktion (4.14) Polynome in s sind, heißt diese Darstellung auch **Polynomform** der Übertragungsfunktion. Eine andere Form ergibt sich, wenn man die Polynome im Zähler und Nenner als Produkte von Linearfaktoren schreibt, was aufgrund des Fundamentalsatzes der Algebra stets möglich ist. Es gilt

$$b_q s^q + \dots + b_1 s + b_0 = b_q \prod_{i=1}^q (s - s_{0i})$$

$$a_n s^n + \dots + a_1 s + a_0 = a_n \prod_{i=1}^n (s - s_i),$$

wobei die Zahlen s_{0i} die Nullstellen des Zählerpolynoms und die Zahlen s_i die Nullstellen des Nennerpolynoms von $G(s)$ darstellen, also die Lösungen der Gleichungen

$$b_q s^q + \dots + b_1 s + b_0 = 0 \quad (4.15)$$

$$a_n s^n + \dots + a_1 s + a_0 = 0 \quad (4.16)$$

sind. Weil die Übertragungsfunktion für die komplexen Frequenzen $s = s_{0i}$ gleich Null ist und für $s = s_i$ unendlich groß wird, heißen die s_{0i} **Nullstellen** und die s_i **Pole** der Übertragungsfunktion.

Gl. (4.16), aus der die Pole bestimmt werden, heißt **charakteristische Gleichung** des Systems, das Polynom auf der linken Gleichungsseite **charakteristisches Polynom**.

Bei der Überführung der DGL (4.12) in die Regelungsnormalform (des Zustandsraummodells) hatten wir zunächst durch Division durch a_n erzwungen, dass

$$a_n = 1 \text{ (genauer: } \tilde{a}_n = \frac{a_n}{a_n} = 1 \text{)}$$

gilt. Wenn wir dies hier auch tun, so ist die entstehende charakteristische Gleichung (4.16) identisch mit der charakteristischen Gleichung

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

der Matrix \mathbf{A} in Regelungsnormalform [siehe Gl. (3.95)]. Folglich stimmen die jeweiligen Lösungen dieser Gleichungen ebenfalls überein. Die Pole der Übertragungsfunktion sind also nichts anderes als die Eigenwerte der Systemmatrix \mathbf{A} :

$$\lambda_i = s_i \quad (4.17)$$

Stabile Systeme haben deshalb Pole mit negativen Realteilen. Aus der Partialbruchzerlegung (4.11) ist als Bedingung für eine Nullstelle s_{0j} (für Systeme mit $d = 0$) abzulesen:

$$\sum_{i=1}^n \frac{g_i}{s_{0j} - \lambda_i} = 0 \quad [\text{vergl. Gl. (3.120)}]$$

Die Nullstellen s_{0j} stimmen also genau mit denjenigen komplexen Frequenzen μ_j überein, für die die Elementarfunktionen $e^{\mu_j t}$ der Eingangsgröße durch das System nicht übertragen werden (genauer: die im stationären Verhalten der Ausgangsgröße „ausgeblendet“ sind).

Unter Verwendung der Pole und Nullstellen kann die Übertragungsfunktion in der sog. **Pol-Nullstellen-Form**

$$G(s) = k \frac{\prod_{i=1}^q (s - s_{0i})}{\prod_{i=1}^n (s - s_i)} \quad (4.18)$$

geschrieben werden, wobei $k = \frac{b_q}{a_n}$ gilt.

Da die Polynome, aus denen s_i und s_{0i} berechnet werden, reelle Koeffizienten haben, sind die Pole und Nullstellen entweder reell oder treten als konjugiert komplexe Paare auf. Im **Pol-Nullstellen-Bild** (PN-Bild) werden Pole und Nullstellen in der komplexen Ebene graphisch dargestellt, wobei Pole durch „x“ und Nullstellen durch „o“ gekennzeichnet werden.

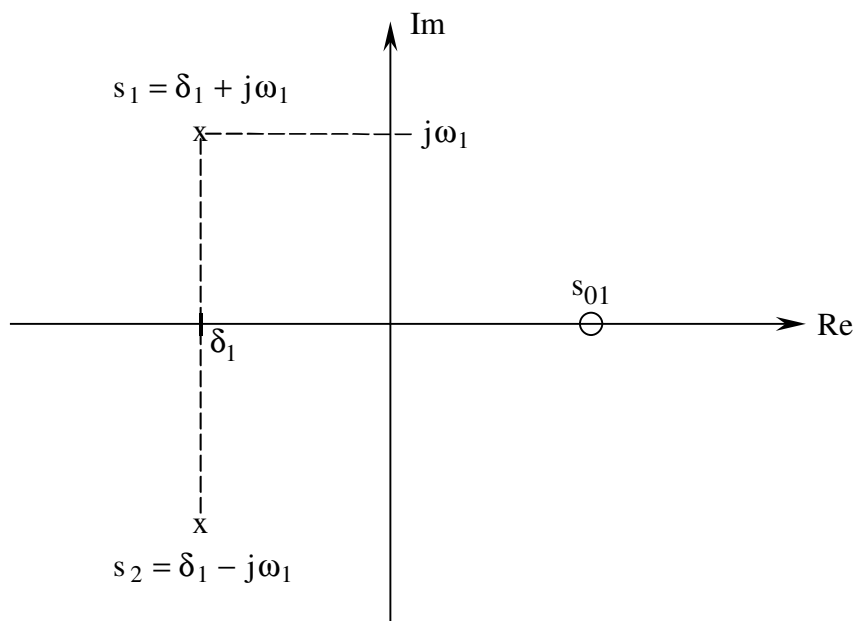


Abb. 22 : Pol-Nullstellen-Bild der Übertragungsfunktion

4.2.3 Berechnung der Systemantwort

Bei bekannter Übertragungsfunktion $G(s)$ können wir die Systemantwort elegant durch Übergang in den Frequenzbereich berechnen. Folgende Rechenschritte sind durchzuführen:

Gegeben: Übertragungsfunktion $G(s)$, Eingangsgröße $u(t)$.

- 1) Berechnung der Laplacetransformierten des Eingangssignals $u(t)$

$$U(s) = \mathcal{L} [u(t)]$$

- 2) Berechnung der Laplacetransformierten der Ausgangsgröße gemäß

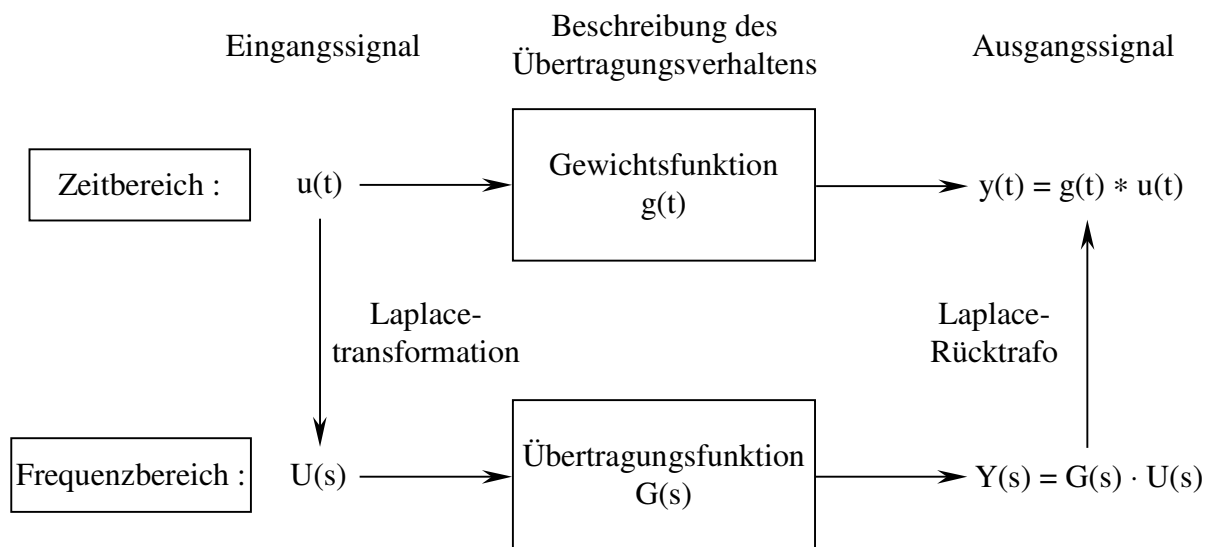
$$Y(s) = G(s) \cdot U(s)$$

- 3) Bestimmung des Ausgangssignals $y(t)$ [genauer: $y_{\text{erzw}}(t)$] durch Laplace-Rücktransformation

$$Y(t) = \mathcal{L}^{-1} [Y(s)]$$

Ergebnis: Ausgangssignal $y(t)$

Die folgende Abbildung illustriert dieses Vorgehen:



zu Schritt 1 : Unter Verwendung der Korrespondenztabelle i.a. einfach.

zu Schritt 2 : Simple Multiplikation !

zu Schritt 3 : Die Rücktransformation der Ausgangsgröße ist i.a. etwas problematischer, da $Y(s)$ i.a. als gebrochen rationale Funktion

$$Y(s) = \frac{d_r s^r + \dots + d_1 s + d_0}{c_m s^m + \dots + c_1 s + c_0} \quad (4.19)$$

vorliegt. Derartige gebrochen rationale Funktionen sind in dieser Form in Korrespondenztafeln nicht zu finden.

Ausweg: Man zerlegt die gebrochen rationale Funktion $Y(s)$ durch **Partialbruchzerlegung** in eine Summe einfacher Elementarfunktionen, für die sich in der Korrespondenztabelle direkt die Zeitfunktionen finden lassen.

1. Schritt: Bestimme die Nullstellen s_i des Nennerpolynoms von $Y(s)$, also die Pole von $Y(s)$. Für $Y(s) = G(s) \cdot U(s)$ sind dies die Eigenwerte des Systems und die Pole von $U(s)$. Da das Nennerpolynom nur reelle Koeffizienten c_i enthält, sind die Pole entweder reell oder treten als konjugiert-komplexe Paare auf.
2. Schritt: Nun zerlegt man die rechte Seite von (4.19) für die reellen Pole in Partialbrüche und für die konjugiert komplexen Pole in Terme mit quadratischem Nenner gemäß folgendem Ansatz:

$$Y(s) = k_0 + \frac{k_1}{s-s_1} + \frac{k_2}{s-s_2} + \dots + \frac{A_i s + B_i}{(s-\delta_i)^2 + \omega_i^2} + \dots$$

\uparrow
reelle

\uparrow
Pole

\uparrow
komplex konjugierte Pole

(4.20)

$s_i = \delta_i + j\omega_i$
 $s_{i+1} = \delta_i - j\omega_i$

3. Schritt: Bestimmung der Koeffizienten k_i und A_i, B_i .

- Bestimmung von k_0 :

k_0 verschwindet für gebrochen rationale Funktionen $Y(s)$ mit $r < m$ und lässt sich für $r = m$ durch den Grenzübergang $s \rightarrow \infty$ berechnen, denn aus Gl. (4.20) folgt

$$k_0 = \lim_{s \rightarrow \infty} Y(s).$$

Für diesen Grenzübergang erhält man aus Gl. (4.19)

$$k_0 = \frac{d_r}{c_m}. \quad (4.21)$$

- Bestimmung von k_j für einfachen, reellen Pol s_j :

Wir multiplizieren beide Seiten von Gl. (4.20) mit $(s-s_j)$ und setzen dann $s = s_j$:

$$\begin{aligned} [(s-s_j)Y(s)]_{s=s_j} &= \left[k_0(s-s_j) + \frac{k_1(s-s_j)}{s-s_1} + \dots + \frac{k_j(s-s_j)}{s-s_j} + \dots \right. \\ &\quad \left. + \dots + \frac{(A_i s + B)(s-s_j)}{(s-\delta_i)^2 + \omega_i^2} + \dots \right]_{s=s_j} \\ &= k_j \end{aligned}$$

(Alle Summanden verschwinden bis auf k_j .)

Wenn wir das Nennerpolynom von $Y(s)$ in seine Linearfaktoren zerlegen (die Pole kennen wir ja bereits!),

$$c_m s^m + \dots + c_1 s + c_0 = c_m \cdot \prod_{i=1}^m (s - s_i),$$

so können wir $[(s - s_j) Y(s)]_{s=s_j}$ direkt hinschreiben:

$$k_j = \frac{d_r (s_j)^r + \dots + d_1 s_j + d_0}{c_m \cdot \prod_{\substack{i=1 \\ i \neq j}}^m (s_j - s_i)} \quad (4.22)$$

- Bestimmung von A_i, B_i für konjugiert komplexe Pole s_i, s_{i+1} :

Wegen

$$(s - \delta_i)^2 + \omega_i^2 = (s - \delta_i - j\omega_i)(s - \delta_i + j\omega_i) = (s - s_i)(s - s_{i+1})$$

können die reellen Koeffizienten A_i und B_i analog zu oben bestimmt werden, indem wir Gl. (4.20) mit $(s - s_i)(s - s_{i+1})$ multiplizieren und anschließend $s = s_i$ setzen:

$$\begin{aligned} \underbrace{[(s - s_i)(s - s_{i+1}) Y(s)]_{s=s_i}}_{\frac{d_r (s_i)^r + \dots + d_1 s_i + d_0}{c_m \cdot \prod_{\substack{j=1 \\ j \neq i, j \neq i+1}}^m (s_i - s_j)}} &= \left[\dots + \frac{k_1 (s - s_i)(s - s_{i+1})}{s - s_1} + \dots \right. \\ &\quad \left. \dots + \frac{(A_i s + B_i)(s - s_i)(s - s_{i+1})}{(s - s_i)(s - s_{i+1})} + \dots \right]_{s=s_i} \\ &= A_i s_i + B_i \end{aligned}$$

Beide Seiten dieser Gleichung sind komplex. A_i und B_i (beide reell!) erhält man aus einem Vergleich beider Real- bzw. Imaginärteile.

Die Summanden von $Y(s)$ in der Partialbruchzerlegung (4.20) können einzeln gemäß der Korrespondenztabelle in den Zeitbereich transformiert werden.

Für reelle Pole $s_j = \delta_j$ gilt:

$$\frac{k_j}{s - \delta_j} \quad \bullet \longrightarrow \quad k_j e^{\delta_j t}$$

Partialbrüche der Form $\frac{A_i s + B_i}{(s - \delta_i)^2 + \omega_i^2}$ müssen in Summen der Gestalt

$$\frac{A_i (s - \delta_i)}{(s - \delta_i)^2 + \omega_i^2} + \frac{\tilde{B}_i \omega_i}{(s - \delta_i)^2 + \omega_i^2}$$

zerlegt werden. Dann können die Korrespondenzen

$$\frac{A_i (s - \delta_i)}{(s - \delta_i)^2 + \omega_i^2} \quad \bullet \longrightarrow \circ \quad A_i e^{\delta_i t} \cos \omega_i t$$

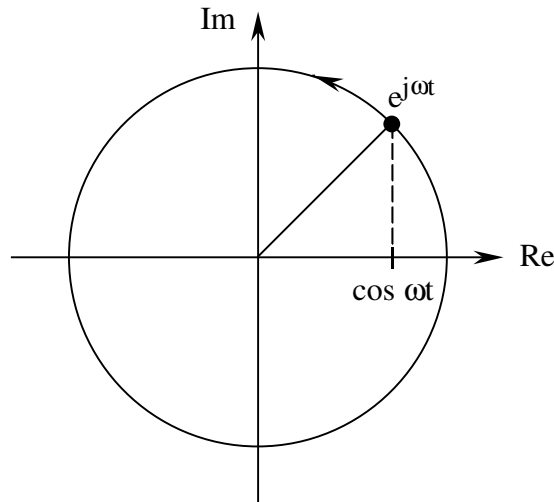
$$\frac{\tilde{B}_i \omega_i}{(s - \delta_i)^2 + \omega_i^2} \quad \bullet \longrightarrow \circ \quad \tilde{B}_i e^{\delta_i t} \sin \omega_i t$$

nachgeschlagen werden.

Summanden mit reellen Polen führen also zu Teilvorgängen mit ab- oder aufklingenden „reinen“ e-Funktionen, während für konjugiert komplexe Pole ab- oder aufklingende harmonische Schwingungen entstehen.

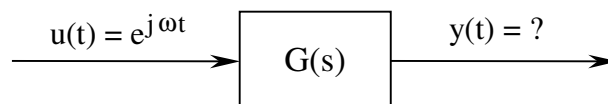
4.3 Übertragungsfunktion und Frequenzgang

Wir wollen nun – eingehender als in Abschnitt 4.1 – die Antwort des Systems auf eine harmonische Schwingung, die bei $t = 0$ beginnt, betrachten.



Wegen $e^{j\omega t} = \cos \omega t + j \sin \omega t$ können wir ein kosinusförmiges Eingangssignal dadurch beschreiben, dass wir für die Rechnungen $e^{j\omega t}$ einsetzen und am Ende unserer Rechnungen von $u(t)$ und $y(t)$ jeweils den Realteil nehmen.

Wir interessieren uns also für



Wegen der Korrespondenz

$$e^{j\omega t} \quad \longleftrightarrow \quad \frac{1}{s - j\omega}$$

lautet die Systemantwort im Frequenzbereich

$$Y(s) = G(s) \cdot U(s) = G(s) \frac{1}{s - j\omega}.$$

Seien die s_1, \dots, s_n die Pole von $G(s)$, die alle einen negativen Realteil haben mögen (d.h. das betrachtete System sei stabil). Die Partialbruchzerlegung von $Y(s)$ ergibt dann

$$Y(s) = \frac{k_1}{s - s_1} + \frac{k_2}{s - s_2} + \dots + \frac{A_i s + B_i}{(s - \delta_i)^2 + \omega_i^2} + \dots + \frac{C}{s - j\omega}.$$

Den Koeffizienten C erhält man wie oben gezeigt:

$$C = \left[\overbrace{(s - j\omega) G(s) \frac{1}{s - j\omega}}^{Y(s)} \right]_{s=j\omega} = G(j\omega)$$

$$\text{Mit } \frac{A_i s + B_i}{(s - \delta_i)^2 + \omega_i^2} = \frac{A_i (s - \delta_i)}{(s - \delta_i)^2 + \omega_i^2} + \frac{\left(\frac{B_i + A_i \delta_i}{\omega_i} \right) \omega_i}{(s - \delta_i)^2 + \omega_i^2}$$

ergibt die Rücktransformation in den Zeitbereich dann

$$y(t) = k_1 e^{s_1 t} + \dots + A_i e^{\delta_i t} \cos \omega_i t + \frac{B_i + A_i \delta_i}{\omega_i} e^{\delta_i t} \sin \omega_i t + \dots + G(j\omega) e^{j\omega t} \quad (4.23)$$

Da die Eigenbewegungen gemäß der Annahme, das System sei stabil, alle abklingen, bleibt nach hinreichend langer Zeit nur die erzwungene Schwingung als stationäre Systemantwort übrig:

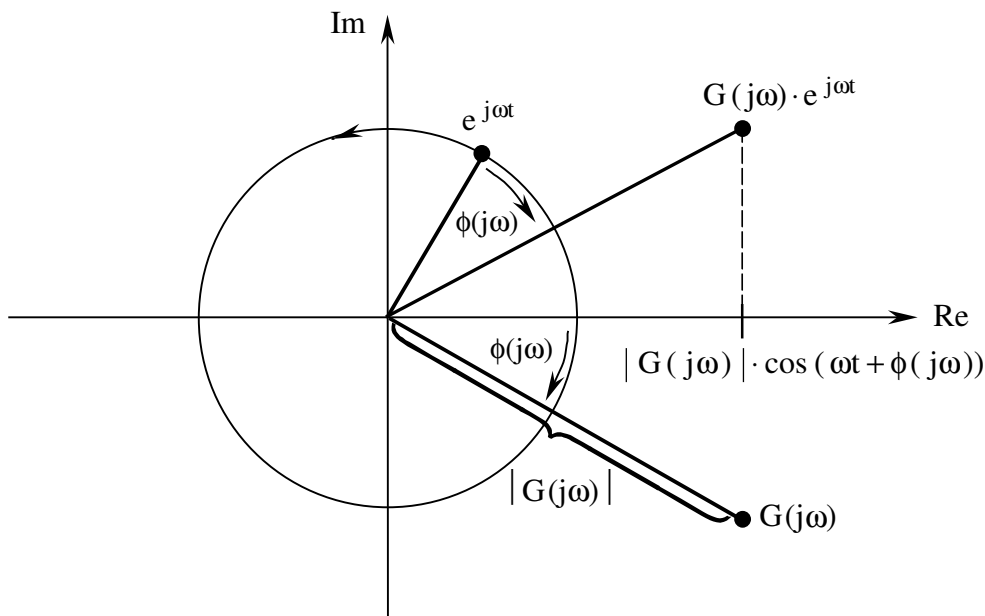
$$y_s(t) = G(j\omega) e^{j\omega t}$$

Wenn man nun zum Realteil – und damit zur Anregung durch eine reelle Kosinusschwingung – übergeht, erhält man

$$y_s(t) = \operatorname{Re} \left\{ G(j\omega) e^{j\omega t} \right\} = |G(j\omega)| \cdot \cos(\omega t + \phi(j\omega)) \quad (4.24)$$

wobei $|G(j\omega)|$ die Amplitude und $\phi(j\omega)$ das Argument von $G(j\omega)$ bezeichnen:

$$G(j\omega) = |G(j\omega)| \cdot e^{j\phi(j\omega)}$$



Die durch eine harmonische Anregung am Eingang erzwungene Bewegung am Ausgang ist also (vgl. Abschnitt 4.1!) ebenfalls eine harmonische Schwingung gleicher Frequenz ω , deren Amplitude um den Faktor $|G(j\omega)|$ verstärkt (bzw. abgeschwächt) ist und die um den Phasenwinkel $\phi(j\omega)$ gegenüber der anregenden Schwingung versetzt ist. Betrachtet man Eingangssignale mit unterschiedlicher Kreisfrequenz ω , so nehmen $|G|$ und ϕ i.a. verschiedene Werte an. Beide Größen können deshalb als Funktion der Kreisfrequenz ω aufgefasst werden. Man spricht dann vom **Amplitudengang** $|G(j\omega)|$ und vom **Phasengang** $|\phi(j\omega)|$. Beide Funktionen zusammen stellen den **Frequenzgang** des Systems dar. Um ihn in Abhängigkeit von ω graphisch darzustellen, gibt es verschiedene Möglichkeiten.

Eine kompakte Darstellung ist die als **Ortskurve** in der komplexen Ebene, die dadurch entsteht, dass man $G(j\omega)$ als Punkt in der komplexen Ebene mit ω wandern lässt und die Punkte dabei zu einer Kurve verbindet. Die Kreisfrequenz ω tritt hierbei als Parameter des Kurvenzugs auf.

Für $\omega = 0$ ist $G(j\omega)$ eine reelle Zahl, nämlich die statische Verstärkung k_s des Systems:

$$G(0) \stackrel{(4.1)}{=} \int_0^{\infty} g(\tau) \underbrace{e^{-0 \cdot \tau}}_1 d\tau = \int_0^{\infty} g(\tau) d\tau \stackrel{(3.114)}{=} h(\infty) = k_s \quad (4.25)$$

Für $\omega \rightarrow \infty$ ist $G(j\omega)$ ebenfalls reellwertig, wie man aus der Polynomform (4.14) der Übertragungsfunktion ablesen kann:

$$G(j\omega) = \frac{b_q \cdot (j\omega)^q + \dots + b_1 \cdot (j\omega) + b_0}{a_n \cdot (j\omega)^n + \dots + a_1 \cdot (j\omega) + a_0}$$

$$\Rightarrow \lim_{\omega \rightarrow \infty} G(j\omega) = \begin{cases} 0 & \text{für } q < n \quad (\text{nicht sprunghafte Systeme}) \\ \frac{b_n}{a_n} = d & \text{für } q = n \quad (\text{sprunghafte Systeme}) \end{cases} \quad (4.26)$$

[Beachte, dass wir bei der Herleitung der Formel (3.51), $d = b_n$, den Koeffizienten a_n zu 1 gesetzt hatten !]

Aus (4.26) folgt, dass nicht-sprunghafte Systeme Signale sehr hoher Frequenzen nicht übertragen können.

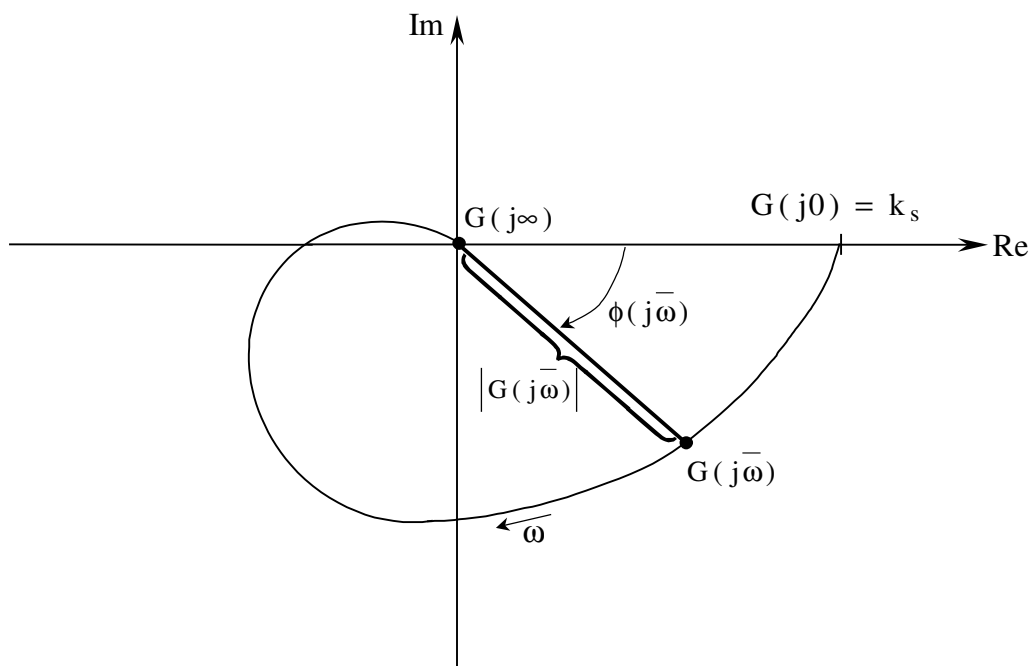


Abb. 23 : Ortskurve eines nicht sprunghaften Systems

Die Phasenverschiebung ist im allgemeinen negativ (vgl. Abb. 23). Das negative Vorzeichen der Phasenverschiebung ist Ausdruck der Verzögerung, mit der ein System das Eingangssignal überträgt. Das Ausgangssignal als Wirkung folgt dem Eingangssignal als Ursache der Systembewegung mit einiger „Verspätung“.

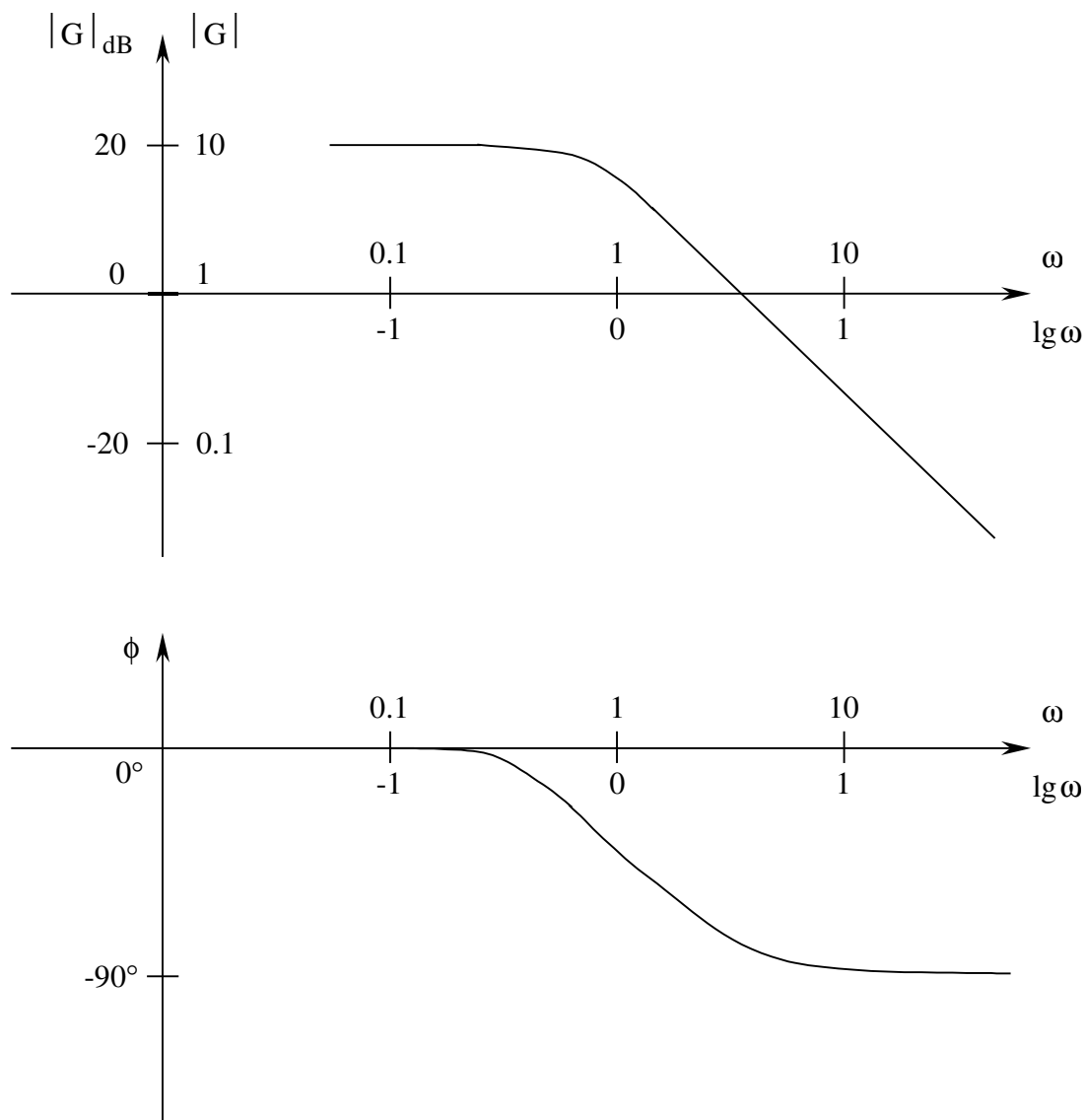
Der Nachteil der Ortskurven-Darstellung ist, dass nicht zu ersehen ist, wie „schnell“ die Ortskurve durchlaufen wird. Eine Auftragungsart, die diese Information anzeigt, ist das **Frequenzkennlinien-Diagramm** oder **Bode-Diagramm**. Hierbei werden Betrag $|G(j\omega)|$ und Phase $\phi(j\omega)$ des Frequenzganges jeweils separat in Abhängigkeit von der Kreisfrequenz ω dargestellt. Die beiden Kennlinien werden als Amplitudenkennlinie (Amplitudengang) bzw. Phasenkennlinie (Phasengang) bezeichnet.

Der für praktische Aufgaben interessante Frequenzbereich erstreckt sich über mehrere Größenordnungen. Da sich auch der Betrag $|G(j\omega)|$ des Frequenzganges um mehrere Größenordnungen verändert, werden die Maßstäbe bei der Amplitudenkennlinie logarithmiert. Auf der Abszissenachse wird die Kreisfrequenz ω mit einer logarithmischen Skala oder $\lg \omega$ (\lg = dekadischer Logarithmus) mit einer linearen Skala aufgetragen. Die Ordinatenachse wird i.a. linear geteilt und der Amplitudengang in Dezibel (dB) aufgetragen. Die in Dezibel angegebene Amplitude $|G|_{\text{dB}}$ berechnet sich aus der dimensionslosen Amplitude $|G|$ gemäß

$$|G|_{\text{dB}} = 20 \cdot \lg |G| \quad (4.27)$$

Man beachte, dass bei diesem $\lg |G| - \lg \omega$ -Diagramm der Punkt $(0, 0)$ ins negativ Unendliche rückt und sich folglich der Darstellung entzieht.

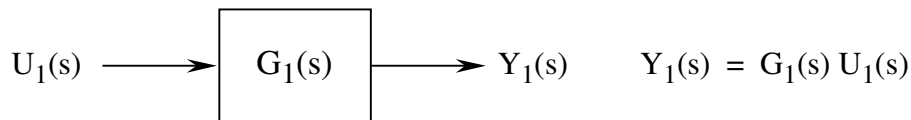
Bei der Phasenkennlinie wird die Phase ϕ als Funktion des Logarithmus der Kreisfrequenz dargestellt ($\rightarrow \phi - \lg \omega$ -Diagramm).



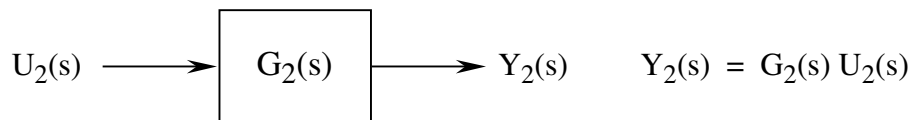
4.4 Übertragungsfunktion zusammenschalteter Übertragungsglieder

Die Übertragungsfunktion eignet sich sehr gut für die Berechnung von Systemen, die aus mehreren Teilsystemen zusammengesetzt sind. Die Übertragungsfunktion des Gesamtsystems kann dann sehr einfach aus den Übertragungsfunktionen der Teilsysteme gebildet werden.

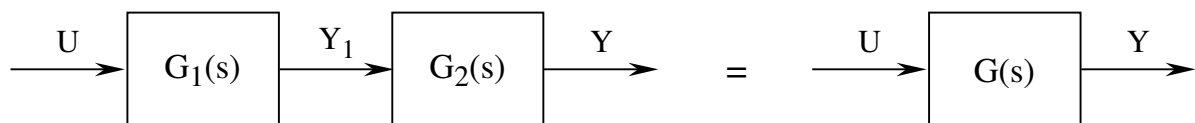
Gegeben seien zwei Teilsysteme



und



Reihenschaltung :



Gegeben : $G_1(s)$, $G_2(s)$

Gesucht : Übertragungsfunktion $G(s)$ der Reihenschaltung

Ein- und Ausgangsgrößen der Übertragungsglieder und Ein- und Ausgangsgrößen des Gesamtsystems hängen wie folgt zusammen:

$$Y(s) = Y_2(s) \quad (\text{Gesamt-Ausgang} = \text{Ausgang von } \ddot{\text{U}}\text{-Glieder 2})$$

$$U_2(s) = Y_1(s) \quad (\text{Eingang von } \ddot{\text{U}}\text{-Glieder 2} = \text{Ausgang von } \ddot{\text{U}}\text{-Glieder 1})$$

$$U_1(s) = U(s) \quad (\text{Gesamt-Eingang} = \text{Eingang von } \ddot{\text{U}}\text{-Glieder 1})$$

$$\begin{aligned} \Rightarrow Y(s) &= Y_2(s) = G_2(s) \cdot U_2(s) = G_2(s) \cdot Y_1(s) \\ &= G_2(s) \cdot G_1(s) \cdot U_1(s) = \underbrace{G_2(s) G_1(s)}_{G(s)} U(s) \end{aligned}$$

$$\Rightarrow \boxed{\text{Reihenschaltung : } G(s) = G_1(s) \cdot G_2(s)} \quad (4.28)$$

Schreiben wir $G_1(s)$ und $G_2(s)$ jeweils in der Polynomform als

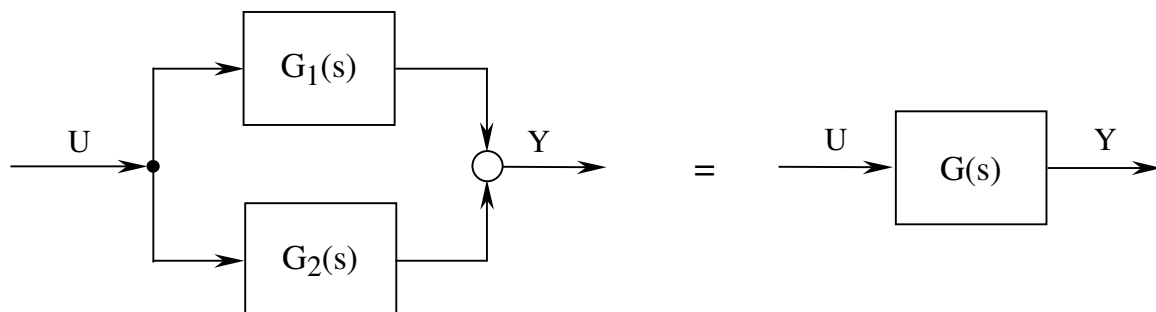
$$G_1(s) = \frac{Z_1(s)}{N_1(s)}, \quad G_2(s) = \frac{Z_2(s)}{N_2(s)},$$

mit den Zählerpolynomen $Z_1(s)$, $Z_2(s)$ und den Nennerpolynomen $N_1(s)$, $N_2(s)$, so ergibt sich

$$G(s) = \frac{Z_1(s) \cdot Z_2(s)}{N_1(s) \cdot N_2(s)}.$$

Das charakteristische Polynom des Gesamtsystems, d.h. der Nenner der (ungekürzten) Gesamtübertragungsfunktion, ist also das Produkt der charakteristischen Polynome der Teilsysteme. Damit sind die Pole der (ungekürzten) Gesamtübertragungsfunktion (d.h. die Eigenwerte des Gesamtsystems) die Vereinigung der Pole der Teilübertragungsfunktionen (d.h. der Eigenwerte der Teilsysteme). Das Gesamtsystem ist also genau dann asymptotisch stabil, wenn die Teilsysteme stabil sind. Da der Grad des Nennerpolynoms von $G(s)$ die Systemordnung (= Zahl der Zustandsvariablen im Zustandsraummodell) widerspiegelt, ist die Ordnung des Gesamtsystems die Summe der Ordnungen der Teilsysteme.

Parallelschaltung:



Bei der Parallelschaltung wirkt die Eingangsgröße auf beide Übertragungsglieder, und die Ausgangsgrößen summieren sich:

$$Y(s) = Y_1(s) + Y_2(s)$$

$$U_1(s) = U_2(s) = U(s)$$

$$\Rightarrow Y(s) = G_1(s) \underbrace{U_1(s)}_{U(s)} + G_2(s) \underbrace{U_2(s)}_{U(s)} = \underbrace{(G_1(s) + G_2(s))}_{G(s)} U(s)$$

$$\Rightarrow \boxed{\text{Parallelschaltung: } G(s) = G_1(s) + G_2(s)} \quad (4.29)$$

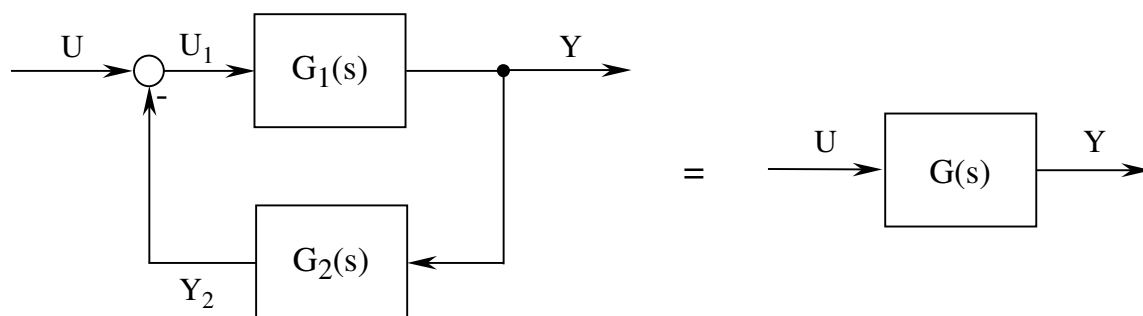
Daraus ergibt sich die Polynomform

$$G(s) = \frac{Z_1(s)}{N_1(s)} + \frac{Z_2(s)}{N_2(s)} = \frac{Z_1(s) N_2(s) + Z_2(s) \cdot N_1(s)}{N_1(s) \cdot N_2(s)}$$

Da also auch bei der Parallelschaltung der Nenner der (ungekürzten) Übertragungsfunktion (d.h. das charakteristische Polynom des Gesamtsystems) das Produkt der Nenner der Teilübertragungsfunktionen (bzw. der charakteristischen Polynome der Teilsysteme) ist, gilt auch für die Parallelschaltung:

- Die Ordnung des Gesamtsystems ist die Summe der Ordnungen der Teilsysteme.
- Die Pole der Gesamtübertragungsfunktion sind die Vereinigung der Pole der Teilübertragungsfunktionen.
- Das Gesamtsystem ist genau dann stabil, wenn die Teilsysteme stabil sind.

Rückführschaltung:



$$Y(s) = Y_1(s) \quad (\text{Gesamt-Ausgang} = \text{Ausgang von } \ddot{U}\text{-Glieder 1})$$

$$U_1(s) = U(s) - Y_2(s) \quad (\text{Ausgang von } \ddot{U}\text{-Glieder 2 wird mit negativem Vorzeichen zum Gesamt-Eingang } U \text{ addiert})$$

$$\begin{aligned} \Rightarrow Y(s) &= G_1(s) U_1(s) = G_1(s) [U(s) - Y_2(s)] \\ &= G_1(s) U(s) - G_1(s) G_2(s) \underbrace{Y(s)}_{Y(s)} \end{aligned}$$

$$\Leftrightarrow Y(s) [1 + G_1(s) G_2(s)] = G_1(s) \cdot U(s)$$

$$\Leftrightarrow Y(s) = \underbrace{\frac{G_1(s)}{1 + G_1(s) G_2(s)}}_{G(s)} U(s)$$

$$\Rightarrow \boxed{\text{Rückführschaltung : } G(s) = \frac{G_1(s)}{1 + G_1(s) G_2(s)}} \quad (4.30)$$

In Polynomform:

$$G(s) = \frac{\frac{Z_1}{N_1}}{1 + \frac{Z_1 Z_2}{N_1 N_2}} = \frac{\frac{Z_1}{N_1}}{\frac{N_1 N_2 + Z_1 Z_2}{N_1 N_2}} = \frac{Z_1(s) \cdot N_2(s)}{N_1(s) N_2(s) + Z_1(s) Z_2(s)} \quad (4.31)$$

Aus (4.31) können wir folgendes ablesen:

- Die Ordnung des Gesamtsystems ist der Grad des Nenners der ungekürzten Übertragungsfunktion. Wegen $\text{Grad}(Z_1) \leq \text{Grad}(N_1)$, $\text{Grad}(Z_2) \leq \text{Grad}(N_2)$ ergibt sich die Ordnung des Gesamtsystems aus der Summe der Ordnungen der Teilsysteme.
- Die Pole der Gesamtübertragungsfunktion sind hier von den Polen der Teilübertragungsfunktionen verschieden, denn diese machen nur den ersten Summanden im Nenner von (4.31) zu Null. Die Eigenwerte des geschlossenen Kreises sind also von den Eigenwerten der offenen Kette (d.h. der an der Rückführ-Summutationsstelle aufgeschnittenen Kreisschleife) verschieden!
- Hieraus folgt: Aus der Stabilität (bzw. Instabilität) der Teilsysteme kann nicht auf die Stabilität (Instabilität) des Gesamtsystems geschlossen werden oder umgekehrt.

4.5 Elementare lineare Übertragungsglieder

Da sich die Übertragungsfunktion zusammengesetzter Systeme gut berechnen lässt, ist es häufig sinnvoll, Regelkreise und andere Übertragungssysteme als Zusammenschaltungen elementarer Systembausteine aufzufassen (bzw. in ihre Komponenten zu zerlegen). Die wichtigsten Standard-Übertragungsglieder wollen wir uns nun näher ansehen.

4.5.1 Proportionalglieder

Als **Proportionalglieder** (abgekürzt: **P-Glieder**) werden dynamische Übertragungsglieder bezeichnet, die auf eine konstante Eingangsgröße $u(t) = \bar{u}$ im stationären (= eingeschwungenen) Zustand mit einer konstanten, zu \bar{u} proportionalen Ausgangsgröße antworten:

$$y_s(t) \propto \bar{u}$$

Das **verzögerungsfreie P-Glied** ist ein statisches Übertragungsglied, das einen rein proportionalen Zusammenhang zwischen Eingangs- und Ausgangssignal modelliert:

$$y(t) = k_s \cdot u(t), \quad (4.32)$$

wobei k_s eine beliebige positive oder negative Konstante darstellt und **Verstärkungsfaktor** des P-Gliedes genannt wird. Die Sprungantwort des P-Gliedes lautet folglich

$$h(t) = k_s.$$

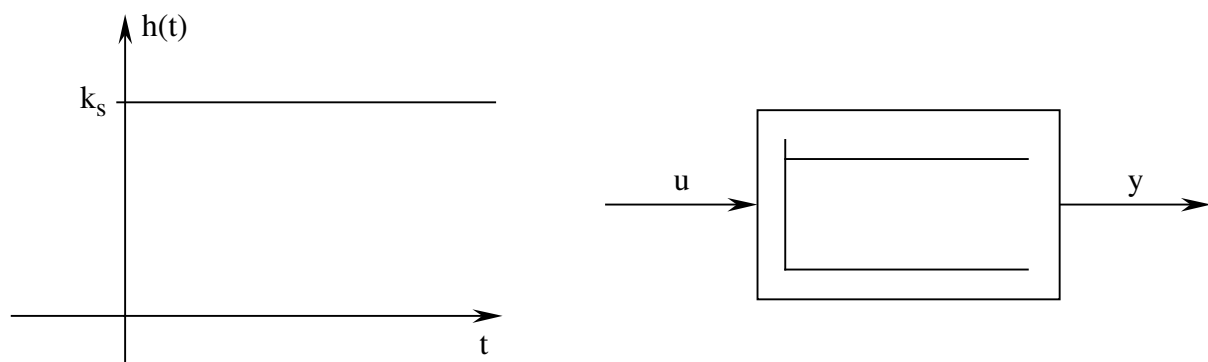


Abb. 24 : Sprungantwort und Symbol eines P-Gliedes

Die Ausgangsgröße folgt der Eingangsgröße ohne Verzögerung. Das Symbol für ein P-Glied zeigt die stilisierte Sprungantwort (Abb. 24 rechts).

Da die Laplace-Transformation linear ist, folgt aus (4.32)

$$Y(s) = k_s \cdot U(s),$$

woraus wir die Übertragungsfunktion $G(s)$ zu

$$G(s) = k_s$$

und den Frequenzgang zu

$$G(j\omega) = k_s \quad \text{ablesen.}$$

Die Ortskurve des Frequenzganges stellt damit für sämtliche Frequenzen einen Punkt auf der reellen Achse dar, d.h. der Phasengang $\phi(j\omega)$ ist Null.

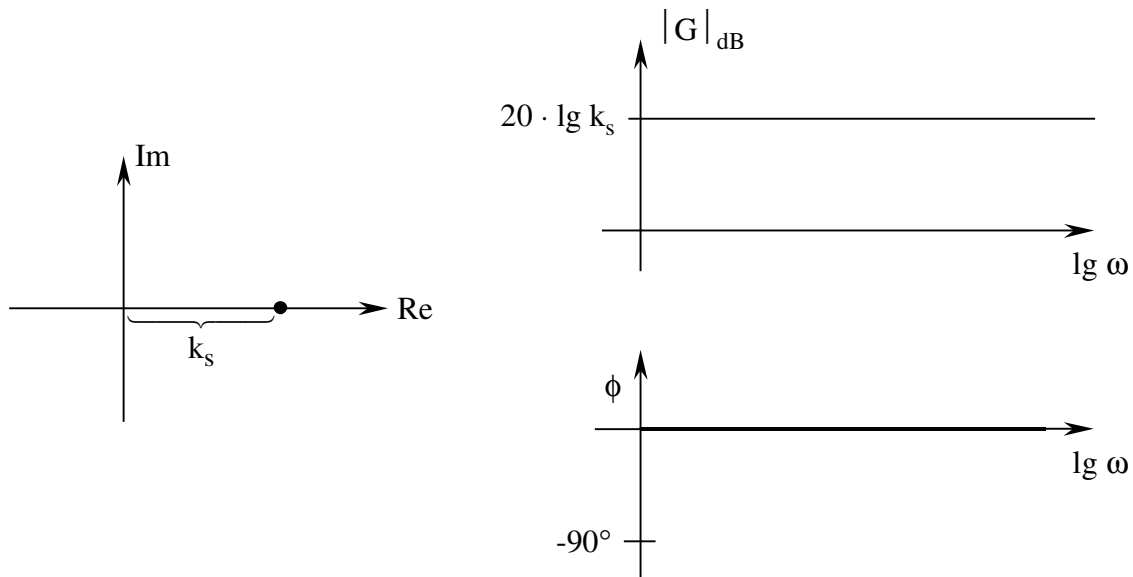


Abb. 25 : Ortskurve und Bodediagramm des verzögerungsfreien P-Gliedes

PT₁ – Glied:

Übertragungsglieder, die – wie das verzögerungsfreie P-Glied – im stationären Zustand proportional zur Eingangsgröße reagieren, aber im Gegensatz zum verzögerungsfreien P-Glied ein ausgeprägtes Übergangsverhalten aufweisen, heißen PT_n -Glieder, wobei n die Systemordnung angibt. Ein PT_1 -Glied ist also ein proportional wirkendes Verzögerungsglied erster Ordnung.

Als nicht sprungfähiges (da es ja mit Verzögerung reagieren soll) Teilsystem 1. Ordnung gehorcht das PT_1 -Glied der DGL 1. Ordnung

$$T \dot{y} + y(t) = k_s u(t), \quad y(0) = y_0. \quad (4.33)$$

[Dies ist im Prinzip die allgemeine DGL 1. Ordnung für nicht sprungfähige Systeme, d.h. ohne Vorkommen von \ddot{u} ; durch Division durch a_0 wurde $a_0 = 1$ erreicht; die Namensgebung $a_1 = T$ und $b_0 = k_s$ für die Koeffizienten spielt auf ihre dynamische Bedeutung an, die wir uns binnen kurzem erschließen werden.]

Indem wir die Zustandsvariable

$$x(t) = \frac{1}{k_s} y(t)$$

eingeführen, können wir Gl. (4.33) in die Form des Zustandsraummodells

$$\dot{x} = \underbrace{-\frac{1}{T}}_A x(t) + \underbrace{\frac{1}{T}}_b u(t), \quad x(0) = \frac{1}{k_s} y_0 \quad (4.34)$$

$$y(t) = \underbrace{k_s}_c \cdot x(t)$$

bringen. [Beachte: Die Systemmatrix A ist hier eine (1,1)-Matrix, also ein Skalar.]

Nun können wir die Sprungantwort $h(t)$ des PT_1 -Gliedes gemäß der allgemeinen Formel (3.103) sofort hinschreiben:

$$\begin{aligned} h(t) &= \underbrace{d}_0 - c A^{-1} b + c A^{-1} e^{At} b \\ &= -k_s (-T) \frac{1}{T} + k_s (-T) e^{-\frac{t}{T}} \frac{1}{T} \\ &= \underbrace{k_s}_{\text{stationäres Verhalten}} - \underbrace{k_s \cdot e^{-\frac{t}{T}}}_{\text{Übergangsverhalten}} \end{aligned} \quad (4.35)$$

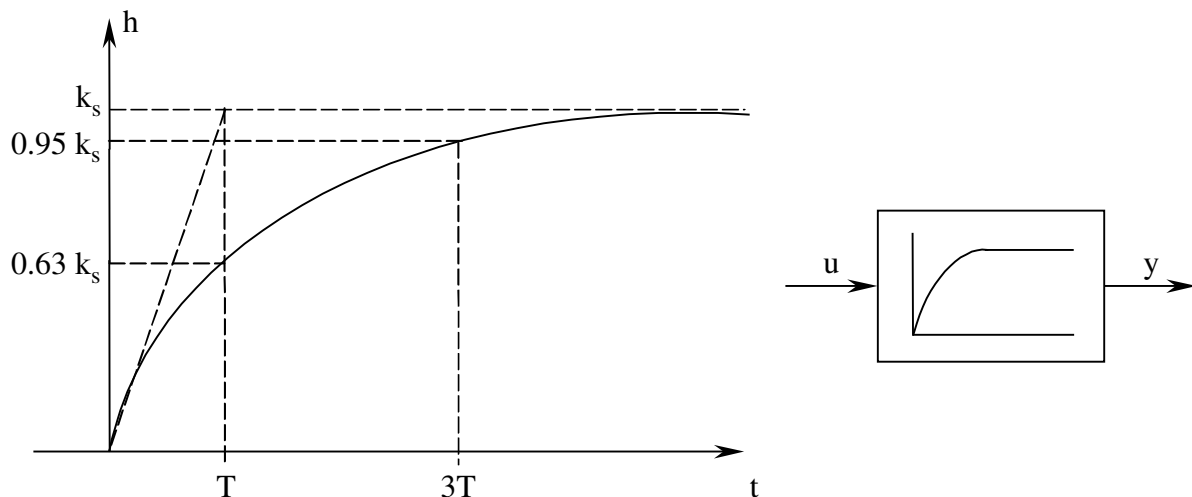


Abb. 26 : Sprungantwort und Symbol eines PT_1 -Gliedes

Aus der Sprungantwort – die in stilisierter Form auch als Symbol eines PT_1 -Gliedes verwendet wird, siehe Abb. 26 rechts – können wir einige Schlussfolgerungen ziehen:

- Das PT_1 -Glied ist in der Tat ein Proportionalglied, mit der statischen Verstärkung k_s (auch als Verstärkungsfaktor des PT_1 -Gliedes bezeichnet).
- T ist eine Zeitkonstante, aus der sich die durch das PT_1 -Glied bewirkte Verzögerung zwischen Eingangs- und Ausgangssignal ablesen lässt:

Zur Zeit $t = T$ nimmt die Übergangsfunktion den Wert

$$h(T) = k_s (1 - e^{-1}) \approx 0.63 k_s \quad (4.36)$$

an (d.h. zur Zeit $t = T$ hat die Sprungantwort 63 % ihres Endwertes $h(\infty) = k_s$ erreicht), zur Zeit $t = 3 \cdot T$ den Wert

$$h(3 \cdot T) = k_s (1 - e^{-3}) \approx 0.95 k_s .$$

Wenn wir zur Zeit $t = 0$ an die Übergangsfunktion die Tangente legen, so lautet diese

$$h_{\text{Tang}} = \dot{h}(t) \Big|_{t=0} \cdot t = \underbrace{\frac{k_s}{T} e^{-\frac{t}{T}} \Big|_{t=0}}_1 \cdot t = \frac{k_s}{T} \cdot t$$

Die Tangente erreicht den Wert k_s also gerade zur Zeit $t = T$ (siehe Abb. 26).

Aus der DGL (4.33) können wir nach bekanntem Muster [siehe Gl. (4.14)] mühelos die Übertragungsfunktion des PT_1 -Gliedes

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

und daraus (indem wir $s = j\omega$ setzen) den Frequenzgang

$$G(j\omega) = \frac{k_s}{j\omega T + 1} \quad (4.37)$$

bestimmen. Um daraus Amplituden- und Phasengang ansehen zu können, müssen wir $G(j\omega)$ in die Polardarstellung bringen. Wir machen zunächst den Nenner reell,

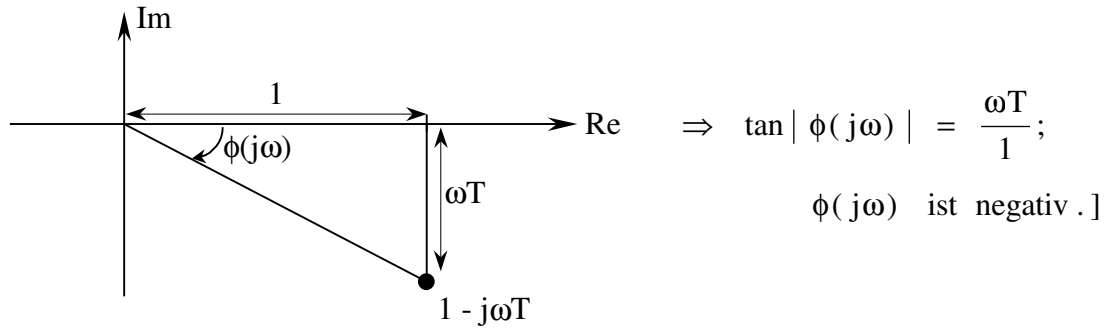
$$G(j\omega) = \frac{k_s (1 - j\omega T)}{(1 + j\omega T)(1 - j\omega T)} = \frac{k_s}{1 + \omega^2 T^2} (1 - j\omega T)$$

und notieren nun

$$\left[\text{zur Erinnerung: } x + i y = r e^{i\theta} \Rightarrow r = \sqrt{x^2 + y^2} \right]$$

$$|G(j\omega)| = \frac{|k_s|}{1 + \omega^2 T^2} \sqrt{1 + \omega^2 T^2} = \frac{|k_s|}{\sqrt{1 + \omega^2 T^2}} \quad (4.38)$$

sowie [Zwischenbetrachtung:



$$\phi(j\omega) = - \arctan(\omega T) \quad (4.39)$$

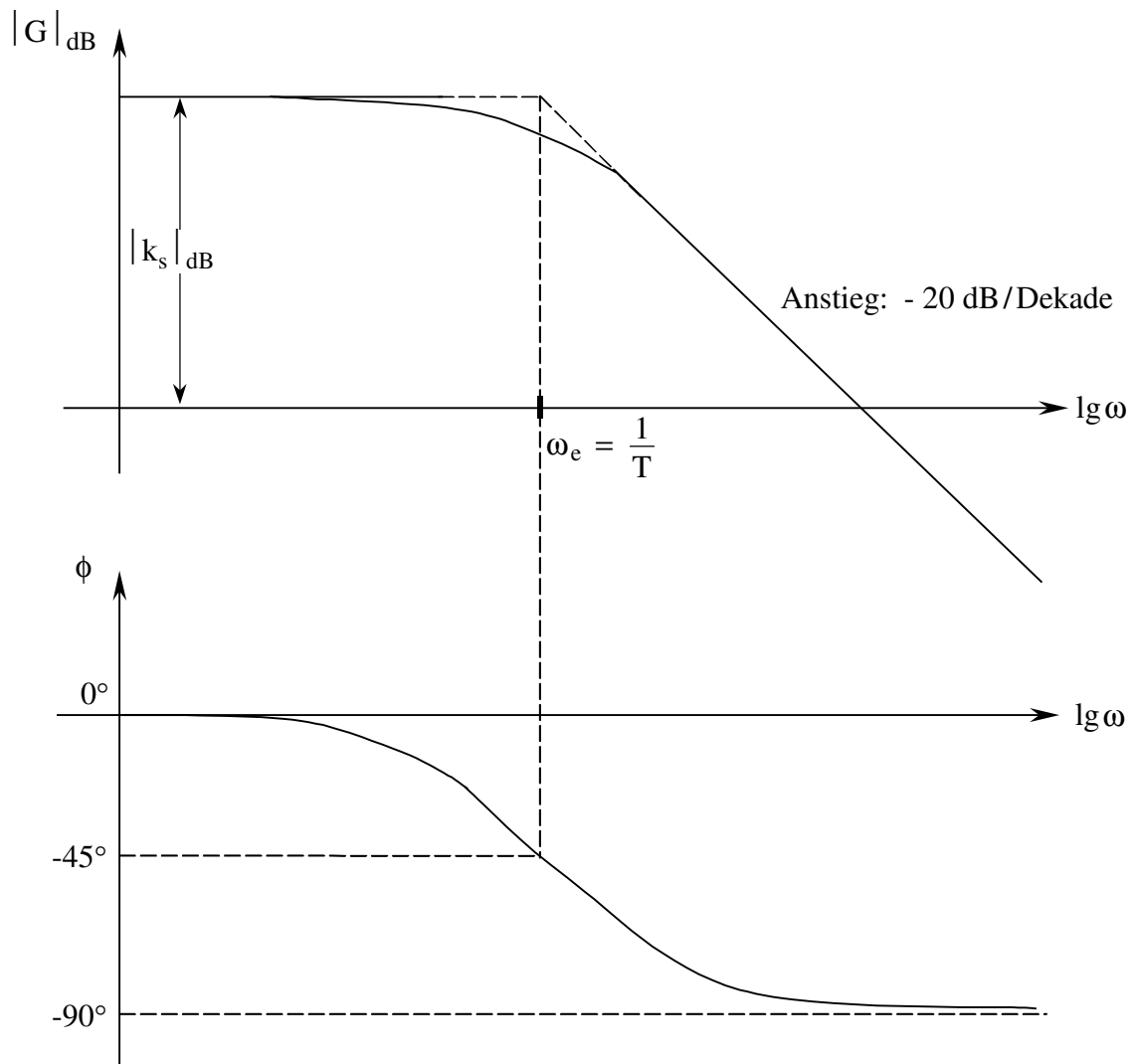


Abb. 27 : Bodediagramm des PT₁-Gliedes

Der Amplitudengang kann asymptotisch durch zwei Geraden approximiert werden. Umgerechnet in Dezibel erhält man für $|G(j\omega)|$

$$\begin{aligned} |G(j\omega)|_{\text{dB}} &= 20 \cdot \lg \left[|k_s| \frac{1}{\sqrt{\omega^2 T^2 + 1}} \right] \\ &= 20 \lg |k_s| - 20 \lg \sqrt{\omega^2 T^2 + 1} \\ &= 20 \lg |k_s| - 20 \lg \sqrt{\left(\frac{\omega}{\omega_e}\right)^2 + 1} \\ &\quad \uparrow \\ &\quad \text{mit } \omega_e := \frac{1}{T}, \text{ sog. Knickfrequenz.} \end{aligned}$$

Für $\frac{\omega}{\omega_e} \ll 1$ ist der Amplitudengang nahezu konstant

$$|G(j\omega)|_{\text{dB}} \approx 20 \lg |k_s| = |k_s|_{\text{dB}},$$

für $\frac{\omega}{\omega_e} \gg 1$ fällt der Amplitudengang linear mit 20 dB/Dekade:

$$|G(j\omega)|_{\text{dB}} \approx 20 \lg |k_s| - 20 \lg \frac{\omega}{\omega_e}$$

Wie dem Amplitudengang zu entnehmen ist, überträgt ein PT_1 -Glied niedrige Frequenzen gut und hohe schlecht und wird daher auch als **Tiefpass** bezeichnet.

Da der Amplitudengang bei

$$\frac{\omega}{\omega_e} = \frac{1}{T}$$

abknickt (\rightarrow Name „Knickfrequenz“!), ist der Durchlassbereich (sog. Bandbreite des Systems) reziprok zur Verzögerungszeit T . PT_1 -Glieder mit großer Verzögerungszeit T wirken also als Tiefpässe mit geringer Bandbreite.

Die Ortskurve des PT_1 -Frequenzganges [siehe Gl. (4.37)] ist ein Halbkreis, der für $\omega = 0$ beim reellen Wert k_s beginnt und für $\omega \rightarrow \infty$ im Ursprung der komplexen Ebene endet.

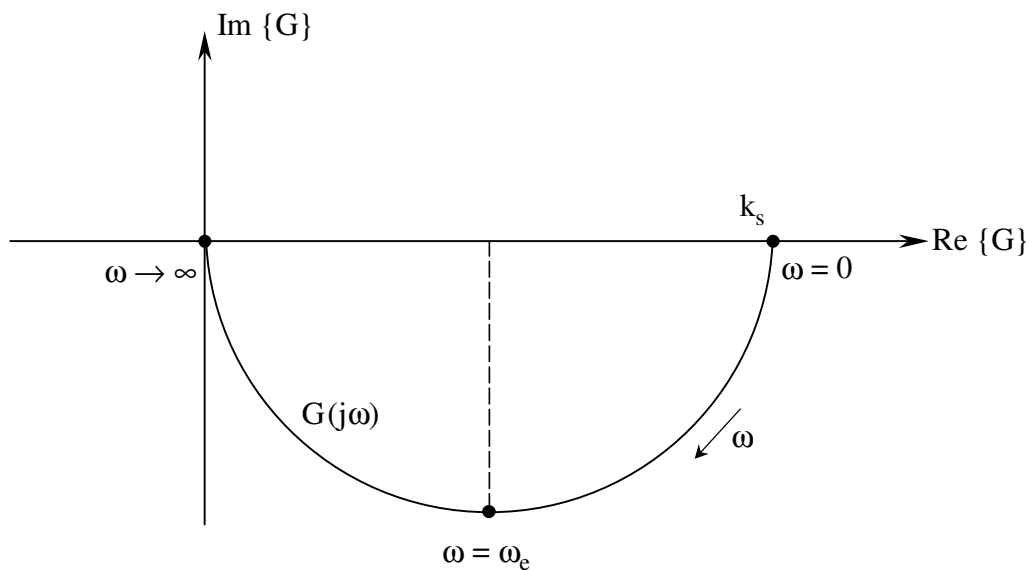


Abb. 28 : Ortskurve eines PT_1 -Gliedes

PT₂ – Glied:

Jedes System 2. Ordnung, das durch eine DGL gegeben ist, in der keine Ableitungen der Eingangsgröße u vorkommen, wirkt – sofern es stabil ist – als verzögerndes Proportionalglied und kann daher als PT_2 -Glied aufgefasst werden. Um dies nachzuweisen, nehmen wir die DGL des oben erwähnten Typs genauer unter die Lupe:

$$T^2 \ddot{y} + 2dT \dot{y} + y(t) = k_s u(t) \quad (4.40)$$

[Allgemeine DGL 2. Ordnung ohne Ableitungen von u ; durch Division durch das „alte“ a_0 wurde $a_0 = 1$ erreicht; die Benennung der Koeffizienten, $a_2 = T^2$, $a_1 = 2dT$, $b_0 = k_s$ spielt auf deren dynamische Bedeutung an.]

In der zu (4.40) äquivalenten Schreibweise [dividiere (4.40) durch T^2 und setze $\omega_0 := \frac{1}{T}$]

$$\ddot{y} + 2d\omega_0 \dot{y} + \omega_0^2 y(t) = (k_s \omega_0^2) u(t)$$

entpuppt sich (4.40) als die aus der Physik wohlbekannte Schwingungsgleichung. ω_0 hat also die Bedeutung der (ungedämpften) Eigenfrequenz des Systems, T als Kehrwert von ω_0 ist eine Zeitkonstante und d ist ein Dämpfungsfaktor.

Ein eleganter Weg zur Berechnung der Sprungantwort des PT_2 -Gliedes führt über die Übertragungsfunktion. $G(s)$ lässt sich (in Polynomform) anhand von (4.40) ablesen zu

$$G(s) = \frac{k_s}{T^2 s^2 + 2dT s + 1} = \frac{k_s}{\frac{1}{\omega_0^2} s^2 + \frac{2d}{\omega_0} s + 1}. \quad (4.41)$$

Da die Sprungantwort $h(t)$ die Reaktion des Systems auf eine Sprungfunktion $\sigma(t)$ am Systemeingang darstellt, gilt

$$h(t) \quad \text{O} \longrightarrow \bullet \quad H(s) = G(s) \cdot \underbrace{\frac{1}{s}}_{\text{Laplace-Transformierte des Eingangssignals } \sigma(t)} \quad [\text{vgl. } Y(s) = G(s) \cdot U(s)].$$

Um die Partialbruchzerlegung für $\frac{1}{s} G(s)$ ausführen zu können, müssen die Pole von $\frac{1}{s} G(s)$ bestimmt werden, die sich aus dem Pol $s_3 = 0$ der Eingangsgröße $\frac{1}{s}$ und den beiden Polen s_1, s_2 des PT_2 -Gliedes zusammensetzen.

Die charakteristische Gleichung des PT_2 -Gliedes lautet [siehe (4.41)]

$$\frac{1}{\omega_0^2} s^2 + \frac{2d}{\omega_0} s + 1 = 0,$$

woraus man die Pole

$$s_{1,2} = \frac{-\frac{2d}{\omega_0} \pm \sqrt{\frac{4d^2}{\omega_0^2} - \frac{4}{\omega_0^2}}}{2 \cdot \frac{1}{\omega_0^2}} = -\omega_0 d \pm \omega_0 \sqrt{d^2 - 1}$$

erhält.

Entsprechend der Größe von d , also des Dämpfungsfaktors, sind 5 Fälle zu unterscheiden (siehe Abb. 29):

A) Für $d > 1$ ergeben sich zwei negative reelle Pole

$$s_1 = -\omega_0 d + \omega_0 \sqrt{d^2 - 1}, \quad s_2 = -\omega_0 d - \omega_0 \sqrt{d^2 - 1}.$$

Nach Einführung der Zeitkonstanten $T_1 := -\frac{1}{s_1}$ und $T_2 := -\frac{1}{s_2}$ (beide positiv, da s_1, s_2 negativ !) erhält man die Übertragungsfunktion

$$\begin{aligned}
 & \text{Linearfaktorzerlegung} \\
 & \text{des charakter. Polynoms} \\
 & \downarrow \\
 G(s) &= \frac{k_s}{\frac{1}{\omega_0^2} (s-s_1)(s-s_2)} = \frac{k_s \cdot \omega_0^2}{\left(s + \frac{1}{T_1}\right) \left(s + \frac{1}{T_2}\right)} \\
 &= \frac{k_s \cdot \omega_0^2}{(T_1 s + 1)(T_2 s + 1) \cdot \underbrace{\frac{1}{T_1 T_2}}_{s_1 s_2 = \omega_0^2 d^2 - \omega_0^2 (d^2 - 1) = \omega_0^2}} \\
 &= \frac{k_s}{T_1 s + 1} \cdot \frac{1}{T_2 s + 1},
 \end{aligned}$$

die als Reihenschaltung zweier PT₁-Glieder mit unterschiedlichen Zeitkonstanten (T₁ und T₂ mit T₁ > T₂) gedeutet werden kann.

Partialbruchzerlegung von H(s):

$$\begin{aligned}
 H(s) &= \frac{k_s \cdot \omega_0^2}{\left[s - \left(-\frac{1}{T_1}\right)\right] \left[s - \left(-\frac{1}{T_2}\right)\right] (s-0)} \stackrel{!}{=} \frac{k_1}{s - \left(-\frac{1}{T_1}\right)} + \frac{k_2}{s - \left(-\frac{1}{T_2}\right)} + \frac{k_3}{s-0} \\
 & \qquad \qquad \qquad \frac{1}{T_1 T_2} = \omega_0^2 \\
 \Rightarrow k_1 &\stackrel{(4.22)}{=} \frac{k_s \cdot \omega_0^2}{\left[\left(-\frac{1}{T_1}\right) - \left(-\frac{1}{T_2}\right)\right] \left(-\frac{1}{T_1}\right)} = \frac{k_s \cdot \omega_0^2}{\frac{T_1 - T_2}{T_1 \cdot T_2}} \cdot (-T_1) \stackrel{\downarrow}{=} -\frac{k_s \cdot T_1}{T_1 - T_2} \\
 k_2 &\stackrel{(4.22)}{=} \frac{k_s \cdot \omega_0^2}{\left[\left(-\frac{1}{T_2}\right) - \left(-\frac{1}{T_1}\right)\right] \left(-\frac{1}{T_2}\right)} = \frac{k_s T_2}{T_1 - T_2} \\
 k_3 &\stackrel{(4.22)}{=} \frac{k_s \cdot \omega_0^2}{\left[0 - \left(-\frac{1}{T_1}\right)\right] \left[0 - \left(-\frac{1}{T_2}\right)\right]} = k_s
 \end{aligned}$$

Mit $\frac{k_j}{s-s_j} \longrightarrow k_j \cdot e^{s_j t}$ können wir nun das Partialbruch-zerlegte H(s) (Laplace-) rücktransformieren zur Sprungantwort

$$h(t) = k_s \left[\underbrace{1}_{e^{0t}} - \frac{T_1}{T_1 - T_2} e^{-\frac{t}{T_1}} + \frac{T_2}{T_1 - T_2} e^{-\frac{t}{T_2}} \right] \quad (4.42)$$

Wegen $\dot{h}(0) = 0$ beginnt die Sprungantwort in Richtung der Zeitachse (siehe Abb. 29). Je größer die Dämpfung d ist, um so langsamer erreicht $h(t)$ den statischen Endwert k_s .

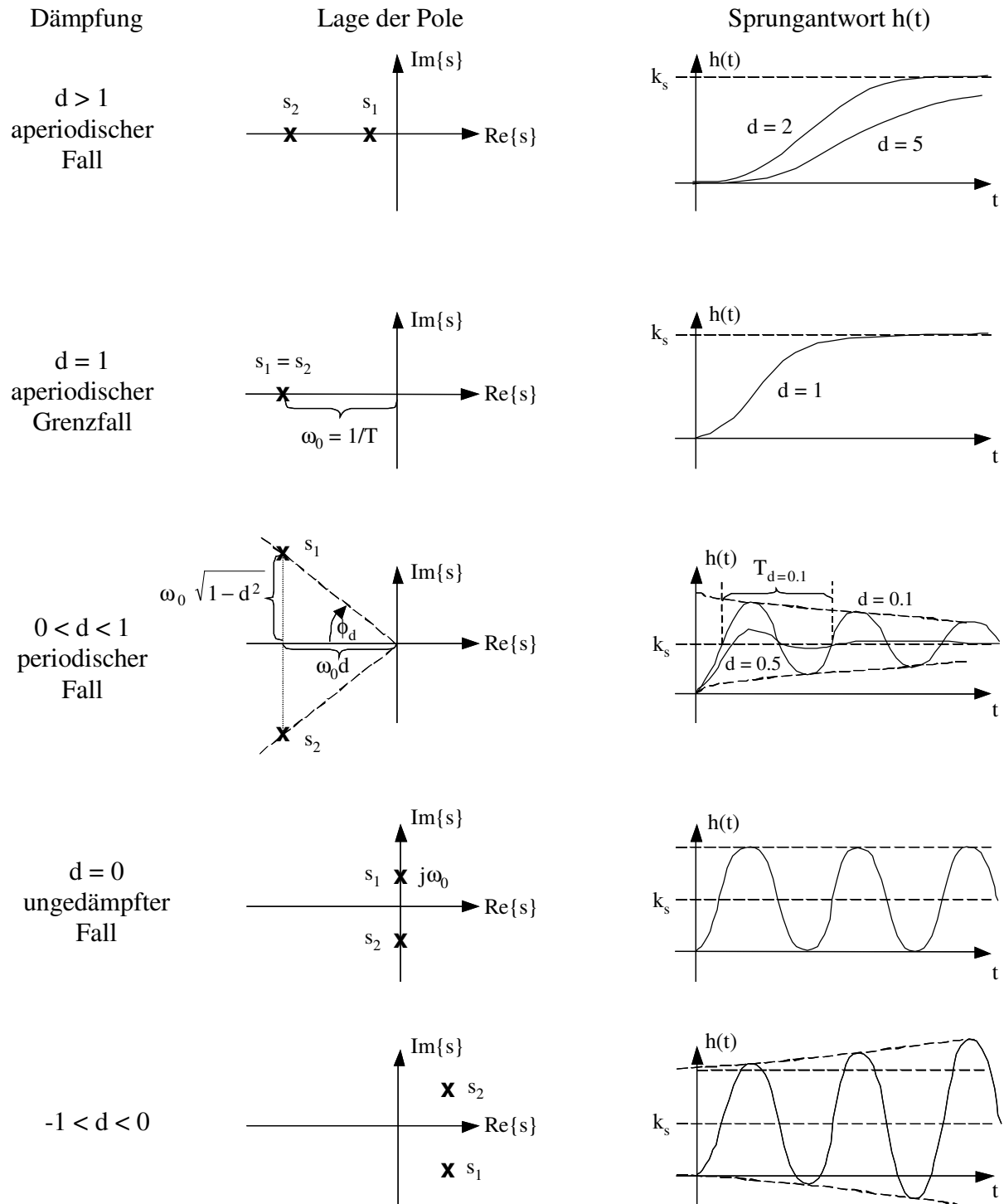


Abb. 29 : Lage der Pole und Verlauf der Sprungantwort eines PT_2 -Gliedes in Abhängigkeit der Dämpfung d

B) Für $d = 1$ (aperiodischer Grenzfall) tritt ein zweifacher Pol

$$s_{1,2} = -\omega_0$$

auf. Setzt man $T = \frac{1}{\omega_0}$, so erhält man für die Übertragungsfunktion die Darstellung

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

die als Reihenschaltung zweier PT_1 -Glieder mit derselben Zeitkonstanten T interpretiert werden kann. Nach Rücktransformation von $\frac{1}{s} G(s)$ erhält man die Sprungantwort (siehe Abb. 29)

$$h(t) = k_s [1 - e^{-\omega_0 t} (1 + \omega_0 t)] \quad (4.43)$$

C) Für $0 < d < 1$ ergibt sich ein konjugiert komplexes Polpaar

$$s_{1,2} = -\omega_0 d \pm j \omega_0 \sqrt{1-d^2}$$

mit negativem Realteil $-\omega_0 d$.

Die aus $\frac{1}{s} G(s)$ durch Partialbruchzerlegung und anschließende Rücktransformation gewonnene Sprungantwort

$$h(t) = k_s \left[1 - e^{-d\omega_0 t} \left\{ \frac{d}{\sqrt{1-d^2}} \sin(\omega_0 \sqrt{1-d^2} t) + \cos(\omega_0 \sqrt{1-d^2} t) \right\} \right] \quad (4.44)$$

$\begin{matrix} a \sin x + b \cos x = r \sin(x + \varphi) \\ \text{mit } r = \sqrt{a^2 + b^2}, \varphi = \arccos \frac{a}{r} \text{ (falls } b \geq 0) \\ \downarrow \\ = \end{matrix}$

$$k_s \left[1 - \frac{1}{\sqrt{1-d^2}} e^{-d\omega_0 t} \sin(\omega_0 \sqrt{1-d^2} t + \arccos d) \right]$$

ist im wesentlichen eine Sinusfunktion mit der Kreisfrequenz (= „gedämpfte Eigenfrequenz“) $\omega_d = \omega_0 \sqrt{1-d^2} < \omega_0$ $\left[\rightarrow \text{Periodendauer } T_d = \frac{2\pi}{\omega_0 \sqrt{1-d^2}} \right]$.

Die Amplitude der Schwingung klingt exponentiell ab, so dass das System schließlich den statischen Endwert k_s erreicht. Die relative Dämpfung der Schwingungen, die die Abnahme der Amplitude von einer zur nächsten Periode beschreibt, lautet

$$e^{-d\omega_0 T_d} = e^{-\frac{d}{\sqrt{1-d^2}} 2\pi}$$

und hängt somit vom Quotienten $\frac{d}{\sqrt{1-d^2}}$ ab. Wegen

$$\frac{d}{\sqrt{1-d^2}} = \left| \frac{\operatorname{Re}\{s_i\}}{\operatorname{Im}\{s_i\}} \right| = \cot \phi_d$$

kann dieser „Dämpfungsquotient“ direkt in Form des Winkels ϕ_d aus dem Pol-Nullstellen-Bild abgelesen werden (siehe Abb. 29). Die Pole $s_{1,2}$ für PT_2 -Glieder mit derselben relativen Dämpfung, aber unterschiedlicher Periodendauer T_d liegen auf Geraden, die nur durch d bzw. den Winkel ϕ_d bestimmt sind. Je größer die Dämpfung d ist, um so näher liegen die Eigenwerte an der reellen Achse und um so schneller klingt das Übergangsverhalten ab.

D) Für $d = 0$ ist das System ungedämpft und hat die Pole

$$s_{1,2} = \pm j \omega_0$$

sowie die Sprungantwort [setze in (4.44) $d = 0$]

$$h(t) = k_s (1 - \cos \omega_0 t) .$$

Das System schwingt mit gleichbleibender Amplitude um den Mittelwert k_s .

E) Für $d < 0$ liegen die beiden Pole von $G(s)$ in der rechten Halbebene:

$$s_{1,2} = |d| \omega_0 \pm \omega_0 \sqrt{d^2 - 1} .$$

Sie können rein reell (falls $d \leq -1$) oder konjugiert komplex (falls $-1 < d < 0$) sein. Entsprechend treffen wir auch hier wieder auf die Sprungantworten der Fälle A) bis C), nur dass der Exponent der Funktion(en) jetzt positiv ist und die Sprungantwort daher für $t \rightarrow \infty$ über alle Grenzen anwächst (\rightarrow instabiler Fall).

Zusammenfassend sehen wir, dass ein System, das durch eine DGL vom Typ (4.40) beschrieben wird, immer dann, wenn es stabil ist – d.h. für $d > 0$, also in den Fällen A) bis C) bzw. im aperiodischen Fall, dem aperiodischen Grenzfall und dem Schwingungsfall – als Proportionalglied mit Übergangsverhalten wirkt und somit als PT_2 -Glieder aufgefasst werden kann.

Indem wir den Frequenzgang [vgl. (4.41)]

$$G(j\omega) = \frac{k_s}{1 + j2d\frac{\omega}{\omega_0} - \frac{\omega^2}{\omega_0^2}} = k_s \cdot \frac{\left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right] - j2d\frac{\omega}{\omega_0}}{\left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right]^2 + \left[2d\frac{\omega}{\omega_0}\right]^2}$$

in seine Polardarstellung bringen (Vorgehen wie beim PT₁-Glied !), erhalten wir den Amplitudengang

$$|G(j\omega)| = \frac{|k_s|}{\sqrt{\left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right]^2 + \left[2d\frac{\omega}{\omega_0}\right]^2}} \quad (4.45)$$

und den Phasengang

$$\phi(j\omega) = -\arctan \frac{2d\frac{\omega}{\omega_0}}{1 - \left(\frac{\omega}{\omega_0}\right)^2}.$$

Der logarithmische Amplitudengang des PT₂-Gliedes

$$|G(j\omega)|_{dB} = 20 \lg |k_s| - 20 \lg \sqrt{\left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right]^2 + \left(2d\frac{\omega}{\omega_0}\right)^2}$$

kann für große bzw. kleine Frequenzen ähnlich wie beim PT₁-Glied durch Geraden approximiert werden.

a) Für $\frac{\omega}{\omega_0} \ll 1$ durch

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

mit

$$\phi(j\omega) \approx 0.$$

b) Für $\frac{\omega}{\omega_0} \gg 1$ durch

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

(= Endasymptote) mit

$$\phi(j\omega) \approx -\pi.$$

Die Endasymptote stellt im Bode-Diagramm eine Gerade mit der Steigung -40 dB/Dekade dar. Als Schnittpunkt beider Asymptoten folgt aus

$$20 \lg |k_s| = 20 \lg |k_s| - 40 \lg \underbrace{\frac{\omega}{\omega_0}}_{= 0 \text{ für } \frac{\omega}{\omega_0} = 1}$$

die ungedämpfte Eigenfrequenz ω_0 .

Die Approximationsgenauigkeit in der Umgebung von ω_0 ist allerdings von der Dämpfung abhängig (siehe Abb. 30).

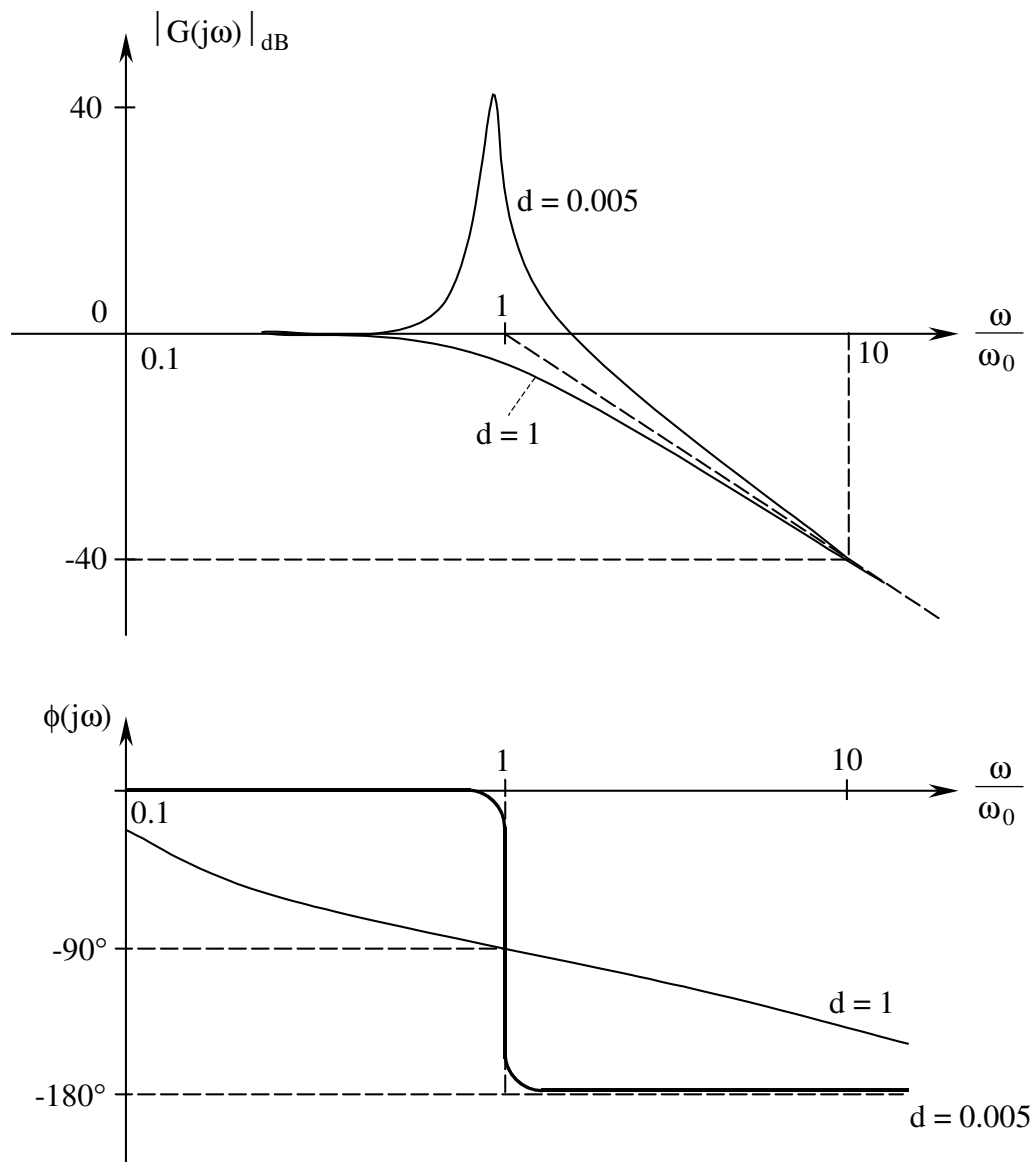


Abb. 30 : Bode-Diagramm eines PT₂-Glieds ($k_s = 1$)

Für große Dämpfung, z.B. $d = 1$, geht der Amplitudengang „glatt“ von einer Geraden in die andere über. Je kleiner die Dämpfung ist, um so größer wird der Amplitudengang in der Nähe von ω_0 . Wir haben gesehen, dass für $0 < d < 1$ das PT₂-Glieder als Antwort auf einen Sprung am Systemeingang eine gedämpfte Schwingung ausführt, die um so langsamer abklingt, je kleiner d ist. Dies macht sich im Amplitudengang durch eine Spitze („Resonanzüberhöhung“) bemerkbar, die um so höher ist, je kleiner d ist. Der Maximalwert tritt für die einzelnen d -Werte bei der sogenannten **Resonanzfrequenz** ω_r auf. Der Maximalwert $|G(j\omega)|_{\max} = |G(j\omega_r)|$ des Amplitudengangs sowie die Resonanzfrequenz ω_r lassen sich aus Formel (4.45) einfach ermitteln. $|G(j\omega)|$ wird maximal, wenn der Radikand im Nenner von (4.45), also

$$a(\omega) = \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 \right]^2 + \left(2d \frac{\omega}{\omega_0} \right)^2,$$

ein Minimum wird. Nullsetzen der Ableitung liefert für $\omega = \omega_r$

$$\frac{da(\omega)}{d\omega} \Big|_{\omega=\omega_r} = \left[-1 + \left(\frac{\omega_r}{\omega_0} \right)^2 + 2d^2 \right] \cdot \omega_r \stackrel{!}{=} 0.$$

Daraus ergibt sich die Resonanzfrequenz

$$\omega_r = \omega_0 \cdot \underbrace{\sqrt{1-2d^2}}_{< 1} \left(\text{für } 1-2d^2 > 0 \text{ bzw. } d < \frac{1}{\sqrt{2}} \right) \quad (4.46)$$

sowie der Maximalwert des Amplitudengangs zu

$$|G(j\omega)|_{\max} = |G(j\omega_r)| = \frac{k_s}{2d\sqrt{1-d^2}}.$$

Aus Gl. (4.46) folgt, dass ein Maximum nur für $d < \frac{1}{\sqrt{2}}$ existiert. Für $d = \frac{1}{\sqrt{2}} \approx 0.707$ wird $\omega_r = 0$ und $|G(j\omega_r)| = 1$ bzw. $|G(j\omega_r)|_{\text{dB}} = 0$, und für $d = 0$ wird $\omega_r = \omega_0$ und $|G(j\omega_r)| = \infty$.

4.5.2 Integralglieder

Integralglieder (kurz: I-Glieder) sind Übertragungsglieder, deren Ausgang bei konstanter Eingangsgröße $u(t) = \bar{u}$ für große Zeiten in eine Rampenfunktion übergeht (siehe Abb. 31).

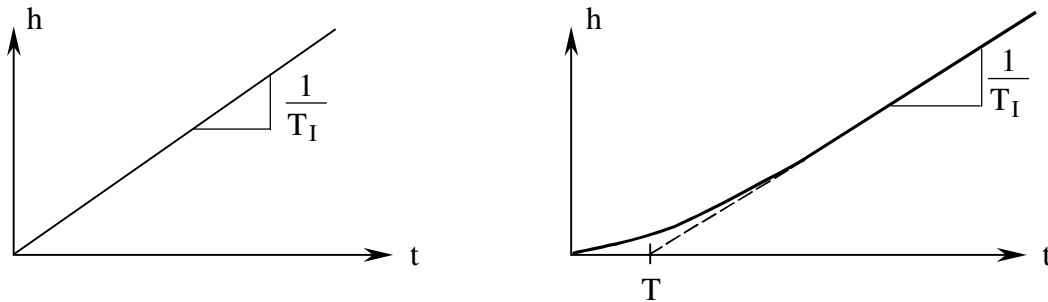


Abb. 31 : Sprungantwort eines verzögerungsfreien I-Gliedes (links) und eines IT_1 -Gliedes (rechts)

Der Ausgang von I-Gliedern wird also im wesentlichen durch Integration der Eingangsgröße gebildet und strebt nur dann einem konstanten Wert zu, wenn die Eingangsgröße gleich Null ist.

Das *verzögerungsfreie I-Glied* wird durch die DGL

$$T_I \dot{y} = u(t), \quad y(0) = y_0 \quad (4.47)$$

beschrieben. Die Lösung dieser DGL lautet

$$y(t) = \frac{1}{T_I} \int_0^t u(\tau) d\tau + y(0), \quad (4.48)$$

so dass dieses Übertragungsglied in der Tat als verzögerungsfreier Integrator wirkt. Die aus (4.48) sofort ablesbare Sprungantwort

$$h(t) = \frac{1}{T_I} \cdot t$$

ist eine Gerade („Rampenfunktion“) mit Steigung $\frac{1}{T_I}$ (siehe Abb. 31 links).

Aus der definierenden DGL (4.47) lässt sich die Übertragungsfunktion des reinen I-Gliedes mühelos zu

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

ersehen. Die Ortskurve des Frequenzganges

$$G(j\omega) = \frac{1}{j\omega T_I} = \frac{1}{\omega T_I} \cdot e^{-j\frac{\pi}{2}}$$

liegt auf der negativen Imaginärachse, der Amplitudengang ist eine Gerade

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

mit Neigung -20 dB/Dekade , die die Frequenzachse bei $\omega = \frac{1}{T_I}$ schneidet (siehe Abb.

32). Die Phase liegt konstant bei -90° .

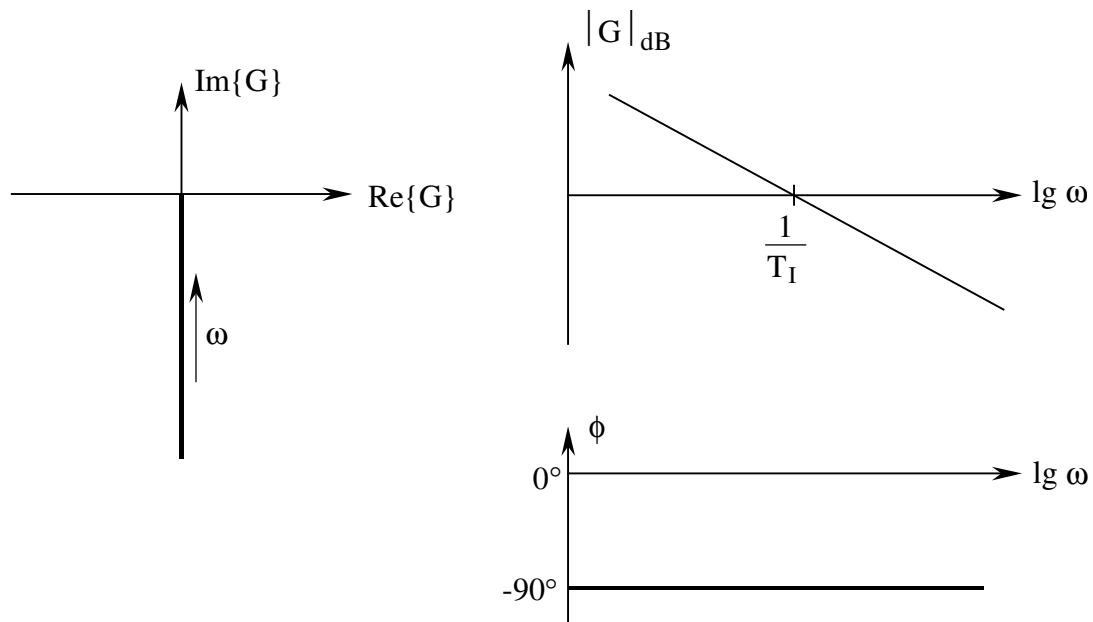


Abb. 32 : Ortskurve und Bodediagramm eines reinen I-Gliedes

Ein I-Glied mit „Verzögerung erster Ordnung“ heißt **IT₁-Glieder**. Es wird durch die DGL

$$T_I T \ddot{y} + T_I \dot{y} = u(t)$$

beschrieben. [Beachte: Das IT₁-Glieder ist also ein System 2. Ordnung, d.h. der Index 1 bei IT₁ bezieht sich auf „Verzögerung erster Ordnung“.]

Seine Übertragungsfunktion

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

zeigt, dass das IT_1 -Glieder als Reihenschaltung eines verzögerungsfreien I-Gliedes und eines PT_1 -Gliedes interpretiert werden kann. Seine in Abb. 31 (rechts) skizzierte Sprungantwort erreicht erst mit einer durch die Zeitkonstante T festgelegten Verzögerung die Rampenfunktion mit Steigung $\frac{1}{T_I}$.

4.5.3 Differenzierglieder

Differenzierglieder (kurz: D-Glieder) sind Übertragungsglieder, deren Ausgang im wesentlichen durch **Änderungen** der Eingangsgröße bestimmt wird und bei konstanter Eingangsgröße dem Wert Null zustrebt.

Das **verzögerungsfreie D-Glied** wird durch die DGL

$$y(t) = T_D \cdot \dot{u} \quad (4.49)$$

beschrieben, führt also eine reine Differentiation der Eingangsgröße durch. Da es die Kausalitätsbedingung $q \leq n$ [siehe (3.22)] verletzt, können reale physikalische Systeme kein unverzögertes D-Verhalten haben.

Aus der DGL (4.49) entnehmen wir die Übertragungsfunktion

$$G(s) = s \cdot T_D, \quad (4.50)$$

woraus sich die Laplace-Transformierte $H(s)$ der Sprungantwort $h(t)$ zu

$$H(s) = G(s) \cdot \frac{1}{s} = T_D$$

und wegen $\delta(t) \xrightarrow{0 \rightarrow 1}$ die Sprungantwort des verzögerungsfreien D-Gliedes zu

$$h(t) = T_D \cdot \delta(t)$$

ergibt (siehe Abb. 33).

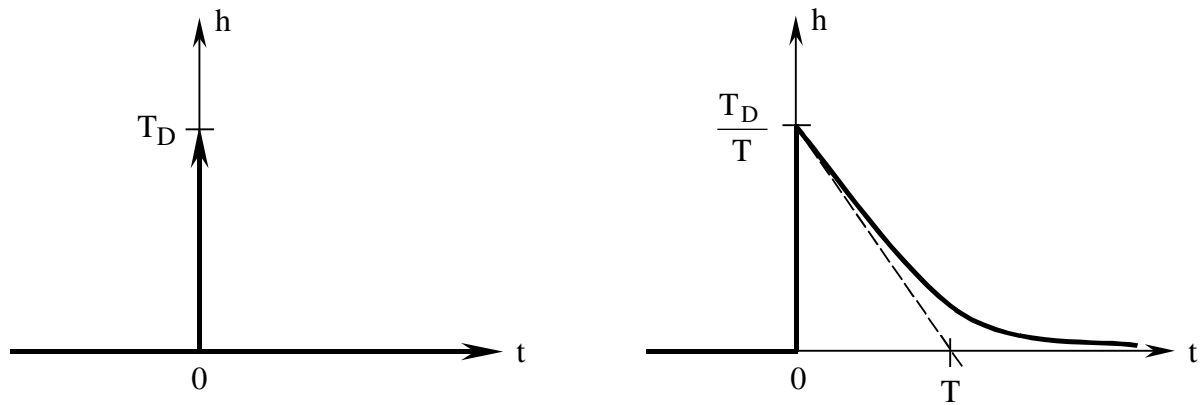


Abb. 33 : Sprungantwort des unverzögerten D-Gliedes (links)
und des DT_1 -Gliedes (rechts)

Aus (4.50) ergibt sich der Frequenzgang des reinen D-Gliedes zu

$$G(j\omega) = j\omega T_D = \omega T_D \cdot e^{j\frac{\pi}{2}}.$$

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

$$\phi(j\omega) = \frac{\pi}{2}$$

Die Ortskurve des Frequenzgangs ist also die positive imaginäre Achse, und der Amplitudengang bildet im Bodediagramm eine Gerade mit dem Anstieg $+ 20 \text{ dB/Dekade}$ und dem Schnittpunkt mit der Abszisse bei $\omega = \frac{1}{T_D}$ (siehe Abb. 34 links). Die Phase des reinen D-Gliedes ist konstant $+ 90^\circ$.

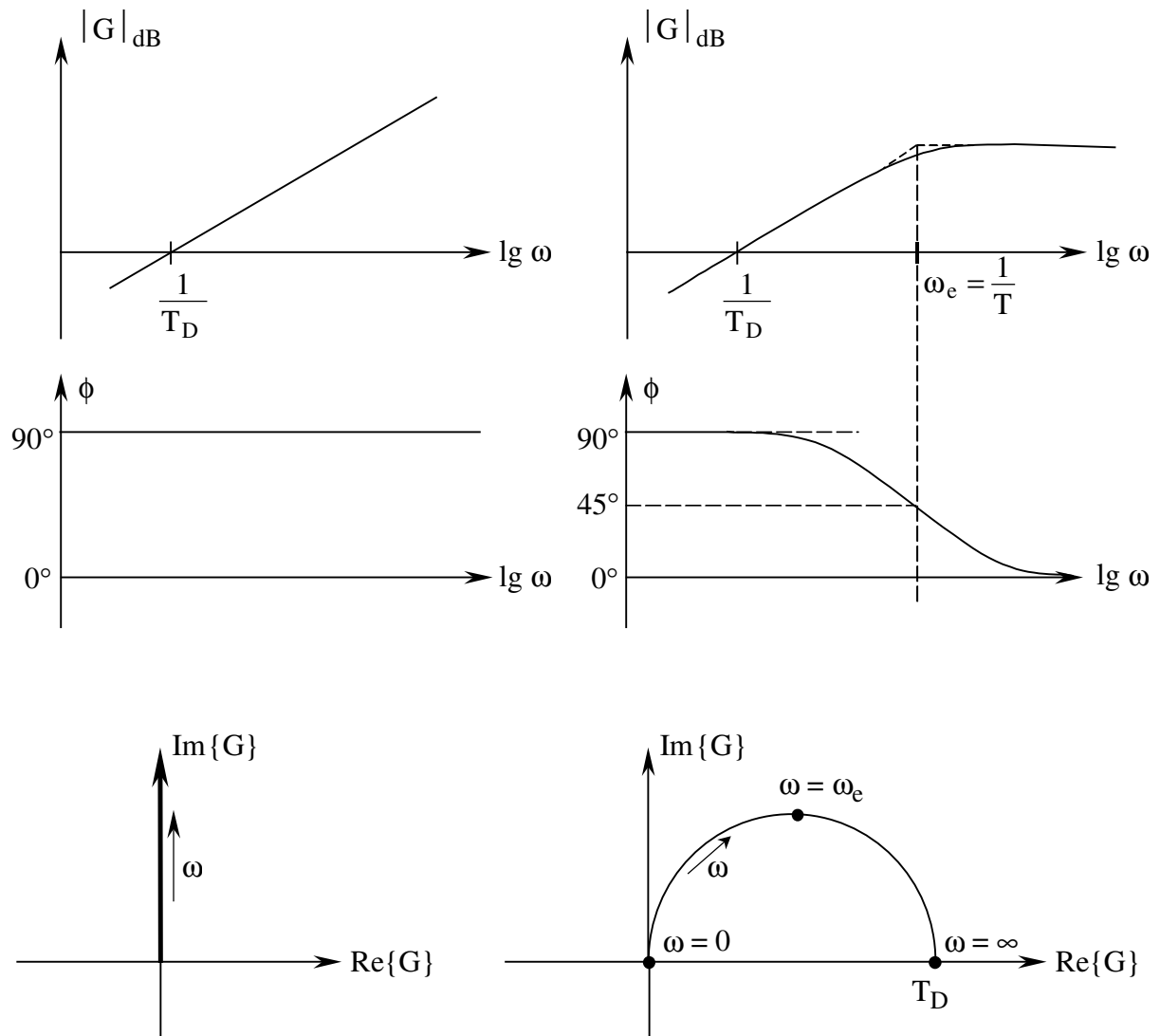
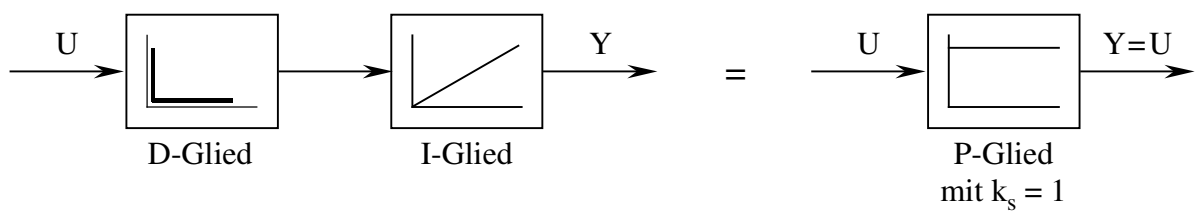


Abb. 34 : Ortskurve und Bodediagramm eines reinen D-Gliedes (links) und eines DT_1 -Gliedes (rechts)

Bemerkung:

Da Differentiation und Integration zueinander inverse Operationen sind, führt die Reihenschaltung eines reinen D-Gliedes und eines reinen I-Gliedes (mit gleichen Zeitkonstanten, d.h. $T_D = T_I$) das Eingangssignal in sich selbst über.



$$\Rightarrow G_D(s) \cdot G_I(s) = 1 \quad \Rightarrow \quad G_D(j\omega) \cdot G_I(j\omega) = 1$$

⇒ Die Beträge (= Amplituden) von D- und I-Glied sind reziprok zueinander, die Phasen sind jeweils das Negative des anderen:

$$|G_D| = \frac{1}{|G_I|} \quad \text{bzw.} \quad |G_D|_{\text{dB}} = -|G_I|_{\text{dB}}, \quad \phi_D = -\phi_I$$

Der Vergleich der Bodediagramme von D- und I-Glied bestätigt genau diese Beziehung.

Das **verzögerungsbehaftete D-Glied** wird DT_n -Glied genannt, wobei n die Zahl der Verzögerungselemente angibt. Die DGL des **DT_1 -Gliedes** (auch **Vorhalteglied** genannt) lautet

$$T \dot{y} + y(t) = T_D \dot{u} \quad (4.51)$$

und führt auf die Übertragungsfunktion

$$G(s) = \frac{T_D s}{Ts + 1} = \underbrace{\frac{T_D s}{Ts + 1}}_{\text{D-Glied}} \cdot \underbrace{\frac{1}{Ts + 1}}_{\text{PT}_1\text{-Glied}} \quad (4.52)$$

Das DT_1 -Glied kann also tatsächlich als Reihenschaltung eines reinen Differenzierers und eines PT_1 -Gliedes (mit statischer Verstärkung 1) aufgefasst werden, d.h. der differenzierende Charakter wird durch ein Verzögerungsglied erster Ordnung „gedämpft“.

Die Sprungantwort ergibt sich mit

$$H(s) = G(s) \cdot \frac{1}{s} = \frac{T_D s}{Ts + 1} \cdot \frac{1}{s} = \frac{T_D}{T} \frac{1}{s - (-\frac{1}{T})}$$

sofort zu

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

(siehe Abb. 33 rechts: sprungartiger Anstieg auf den Wert $\frac{T_D}{T}$ und dann exponentielles Abklingen mit der charakteristischen Zeitkonstanten T).

Laut (4.52) lässt sich der Frequenzgang des DT_1 -Gliedes als Produkt der Frequenzgänge eines reinen D-Gliedes und eines PT_1 -Gliedes mit Verstärkung 1 schreiben:

$$G_{DT_1}(j\omega) = G_D(j\omega) \cdot G_{PT_1}(j\omega).$$

Beim Produkt zweier komplexer Zahlen multiplizieren sich die Beträge und addieren sich die Phasenwinkel

$$\left[z_1 = r_1 e^{j\varphi_1}, \quad z_2 = r_2 e^{j\varphi_2} \Rightarrow z_1 \cdot z_2 = (r_1 \cdot r_2) e^{j(\varphi_1 + \varphi_2)} \right],$$

so dass wir für den Amplitudengang des DT_1 -Gliedes

$$| G_{DT_1}(j\omega) | = | G_D(j\omega) | \cdot | G_{PT_1}(j\omega) |$$

bzw. für den logarithmischen Amplitudengang

$$\begin{aligned} | G_{DT_1}(j\omega) |_{dB} &= | G_D(j\omega) |_{dB} + | G_{PT_1}(j\omega) |_{dB} \\ &= 20 \lg T_D + 20 \lg \omega - 20 \lg \sqrt{\left(\frac{\omega}{\omega_e} \right)^2 + 1} \end{aligned}$$

[mit $\omega_e := \frac{1}{T}$] und für den Phasengang

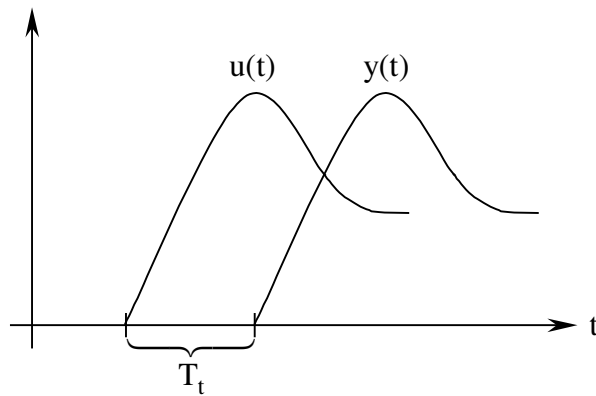
$$\phi_{DT_1}(j\omega) = \phi_D(j\omega) + \phi_{PT_1}(j\omega) = \frac{\pi}{2} - \arctan \left(\frac{\omega}{\omega_e} \right)$$

erhalten (siehe Abb. 34 rechts). Ein DT_1 -System überträgt also Signale hoher Frequenzen besser als Signale mit niedriger Frequenz und ist somit ein **Hochpass** (Bandbreite : $\omega_{Grenz} = \frac{1}{T} \dots \infty$).

4.5.4 Totzeitglieder

Totzeitglieder reproduzieren das Eingangssignal am Ausgang, allerdings mit einer Zeitverschiebung T_t (genannt Totzeit), und modellieren so beispielsweise Transportzeiten in industriellen Produktionsprozessen:

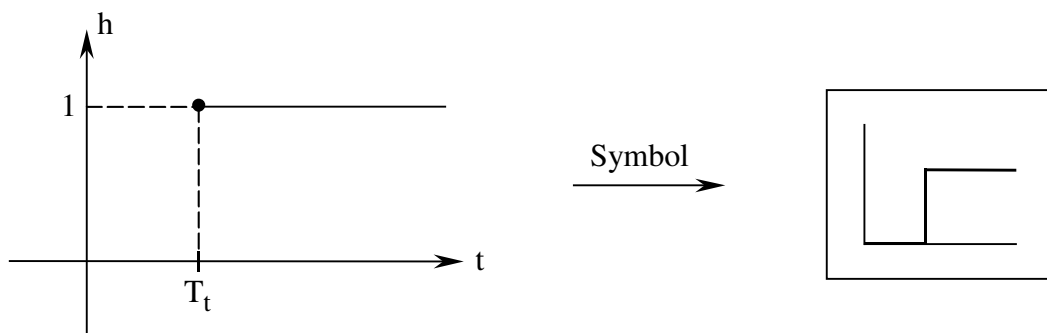
$$y(t) = u(t - T_t) \quad (4.53)$$



Die Sprungantwort ergibt sich aus (4.53) sofort zu

$$h(t) = \sigma(t - T_t)$$

und wird, in schematisierter Form, als Symbol des Totzeitgliedes benutzt (siehe Skizze).



Gemäß dem **Verschiebungssatz der Laplace-Transformation** ergibt sich die Laplace-Transformierte einer um T zeitverschobenen Funktion zu

$$f(t - T) \quad \text{---} \bullet \quad e^{-sT} \quad \underbrace{F(s)}_{F(s) \quad \bullet \text{---} \circ \quad f(t)} .$$

Anwendung auf (4.53) führt zu

$$y(t) = u(t - T_t) \quad \text{---} \bullet \quad e^{-sT_t} U(s)$$

$$\Rightarrow Y(s) = \underbrace{e^{-sT_t}}_{G(s)} U(s)$$

Das Totzeitglied hat also die Übertragungsfunktion

$$G(s) = e^{-sT_t}$$

und den Frequenzgang

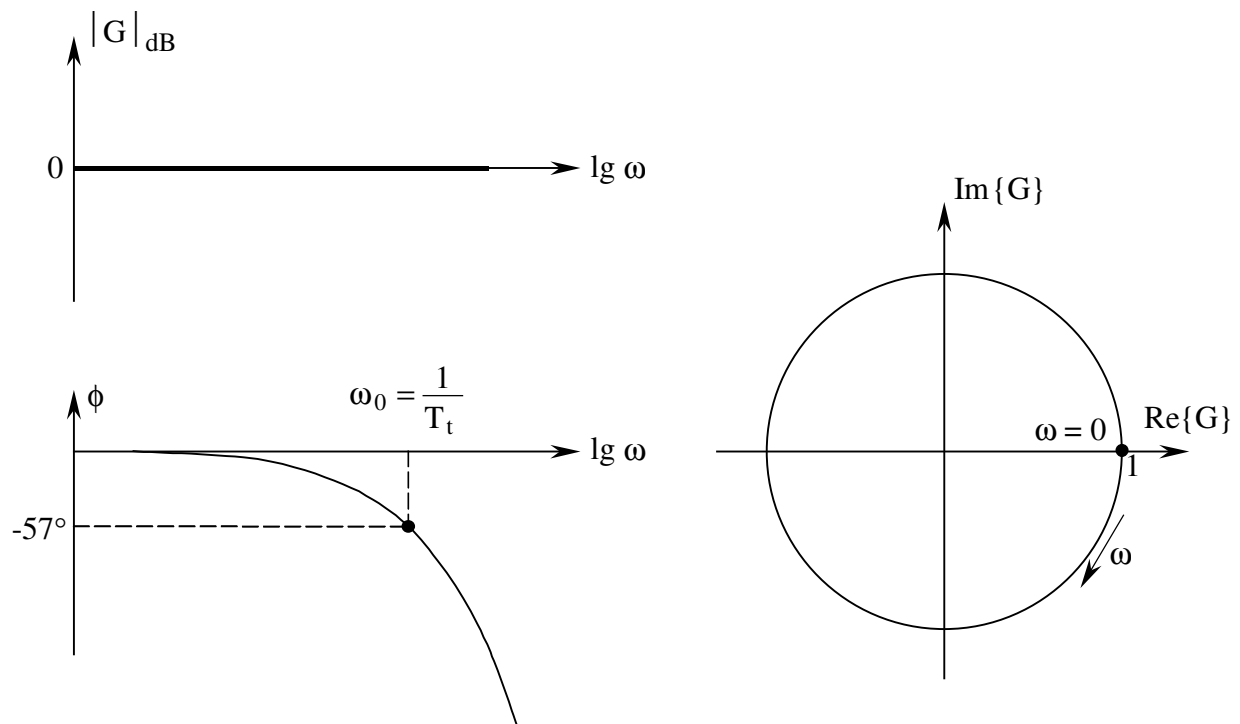
$$G(j\omega) = e^{-j\omega T_t} .$$

Wegen

$$|G(j\omega)| = 1$$

„verstärkt“ ein Totzeitglied alle Frequenzen in gleicher Weise (mit Faktor 1) und verändert nur die Phasenlage sinusförmiger Signale gemäß

$$\phi(j\omega) = -\omega T_t = -\frac{\omega}{\omega_0} \quad \text{mit} \quad \omega_0 := \frac{1}{T_t} .$$



Im Bodediagramm (siehe oben) liegt der Amplitudengang auf der 0dB-Achse, während die Phase, von 0° ausgehend, beständig abnimmt. Die Ortskurve verläuft auf dem Einheitskreis, der für steigende Frequenzen beliebig oft umrundet wird.

Kapitel 5 : Der Regelkreis

4.6 Modell des Standardregelkreises

Das Modell des Standardregelkreises im Frequenzbereich zeigt Abb. 35:

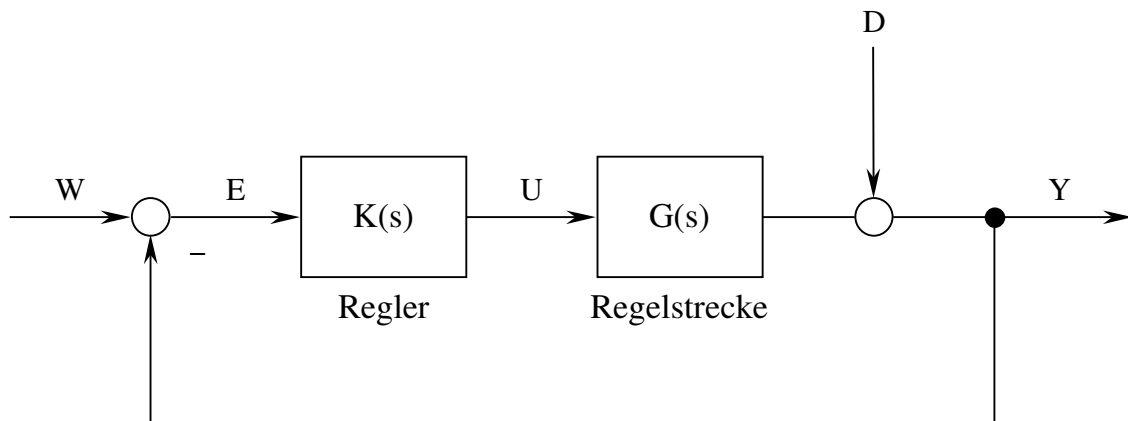
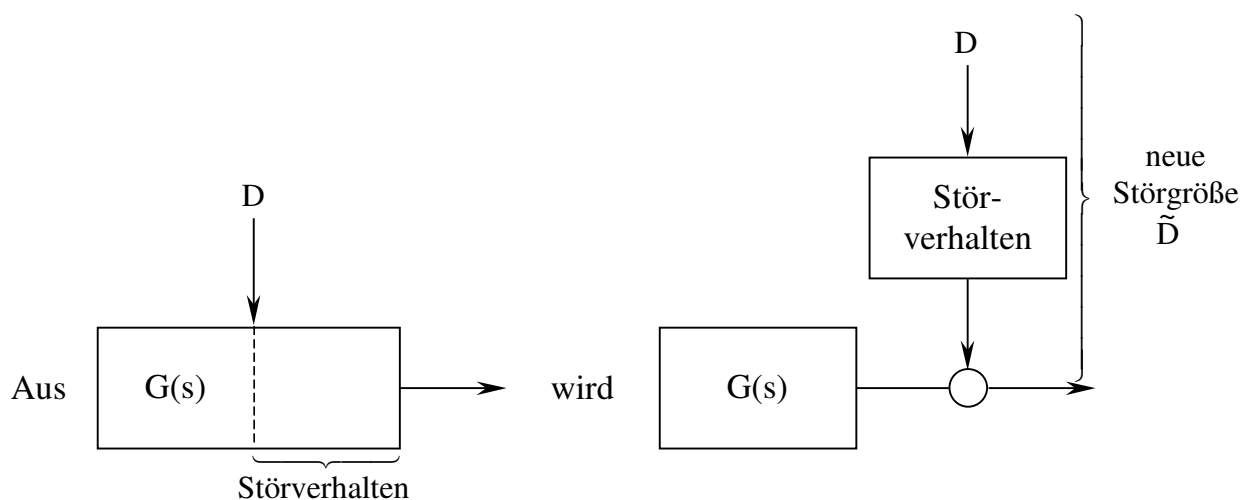


Abb. 35 : Standardregelkreis

Hierbei bezeichnen $K(s)$ die Übertragungsfunktion des Reglers, $G(s)$ die der Regelstrecke, W die (Laplace-transformierte der) Führungsgröße und D die (Laplace-transformierte der) Störgröße. In diesem Modell greift die Störung am Ausgang der Regelstrecke ein. Störungen, die an einem anderen Ort der Regelstrecke eingreifen, können „an den Streckenausgang transformiert“ werden. Das Übertragungsverhalten desjenigen Teils der Regelstrecke, der zwischen (realem) Angriffspunkt der Störgröße und Streckenausgang liegt, wird dann als Übertragungsglied (genannt „Störverhalten“) eingefügt:



Je nachdem, ob die Führungsgröße W oder die Störgröße D als Eingangsgröße des Regelkreises angesehen wird – die jeweils andere Größe wird dann zu Null gesetzt –, erhält man die **Führungsübertragungsfunktion**

$$G_w(s) := \frac{Y(s)}{W(s)} \quad [\text{d.h. } Y(s) = G_w(s) \cdot W(s)]$$

bzw. die **Störübertragungsfunktion**

$$G_d(s) := \frac{Y(s)}{D(s)} \quad [\text{d.h. } Y(s) = G_d(s) \cdot D(s)] .$$

Aus Abb. 35 können folgende Gleichungen abgelesen werden:

$$Y(s) = G(s) \cdot U(s) + D(s)$$

$$U(s) = K(s) \cdot E(s)$$

$$E(s) = W(s) - Y(s)$$

Daraus erhält man für $D(s) = 0$:

$$\begin{aligned} Y(s) &= G(s) \cdot U(s) = G(s) \cdot K(s) \cdot E(s) = \\ &= G(s) K(s) W(s) - G(s) K(s) Y(s) \end{aligned}$$

$$\Leftrightarrow Y(s) [1 + G(s) K(s)] = G(s) K(s) W(s)$$

$$\Leftrightarrow Y(s) = \underbrace{\frac{G(s) K(s)}{1 + G(s) K(s)}}_{G_w(s)} \cdot W(s)$$

Führungsübertragungsfunktion : $G_w(s) = \frac{G(s) K(s)}{1 + G(s) K(s)}$	(5.1)
---	-------

Für $W(s) = 0$ ergibt sich

$$\begin{aligned} Y(s) &= G(s) \cdot U(s) + D(s) = G(s) \cdot K(s) \cdot E(s) + D(s) \\ &= - G(s) K(s) Y(s) + D(s) \end{aligned}$$

$$\Leftrightarrow Y(s)[1 + G(s)K(s)] = D(s)$$

$$\Leftrightarrow Y(s) = \underbrace{\frac{1}{1 + G(s)K(s)}}_{G_d(s)} \cdot D(s)$$

$$\text{Störübertragungsfunktion: } G_d(s) = \frac{1}{1 + G(s)K(s)} \quad (5.2)$$

Beispiel einer Störung (aus der Antriebstechnik) :

Ein drehzahl geregelter Gleichstrommotor wird durch die Ankerspannung u_A angesteuert und erzeugt das Antriebs-Drehmoment M_A (siehe Abb. 36).

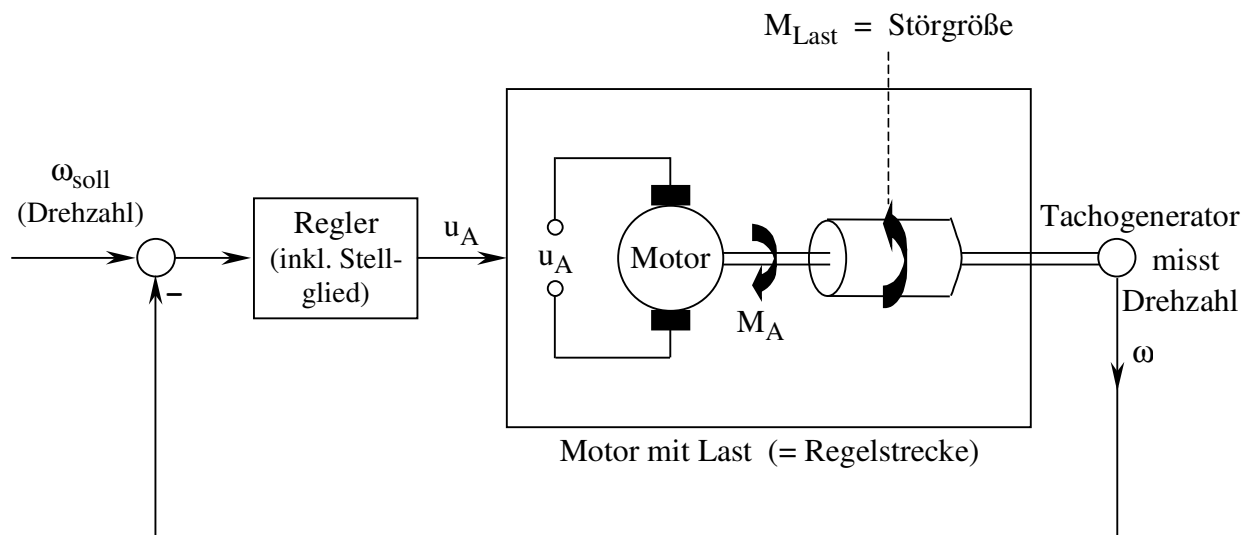


Abb. 36 : Blockschaltbild eines drehzahl geregelten Gleichstrommotors

Hiermit treibt er ein Objekt an, das als „Last“ bezeichnet wird, z.B. eine Pumpe oder die Seilscheibe eines Personenaufzugs. Das von der Last erzeugte Drehmoment M_{Last} wirkt in diesem Regelkreis als Störgröße.

In Verallgemeinerung derartiger Beispiele aus der Antriebstechnik wird $G_d(s)$ auch als Störübertragungsfunktion für **Laststörungen** bezeichnet.

Im Nenner der Führungs- wie der Störübertragungsfunktion kommt das Produkt $G(s) \cdot K(s)$ vor, das als **Übertragungsfunktion der** – aus Regler und Regelstrecke bestehenden – **offenen Kette** mit

$$G_0(s) := G(s) \cdot K(s) \quad (5.3)$$

abgekürzt wird. Ist die offene Kette stabil, so wird ihre statische Verstärkung [siehe Gl. (4.25)]

$$G_0(0) = G(0) \cdot K(0) = k_0 \quad (5.4)$$

als ***Kreisverstärkung*** bezeichnet.

4.7 Forderungen an die Regelung

Die Güte einer Regelung wird i.a. anhand von vier Forderungen beurteilt:

- I) Der Regelkreis muss *stabil* sein.
- II) Für vorgegebene Klassen von Führungs- und Störsignalen soll die Regelgröße der Führungsgröße asymptotisch (d.h. im Langzeitlimes) folgen:

$$\lim_{t \rightarrow \infty} \left[\underbrace{w(t) - y(t)}_{e(t)} \right] \stackrel{!}{=} 0 \quad (5.5)$$

Forderung (5.5) bedeutet, dass der Regelkreis keine bleibende Regelabweichung $e(\infty)$ besitzt: $e(\infty) = 0$. Ein solcher Regelkreis heißt *stationär genau*.

Stationäre Genauigkeit wird meist für sprungförmige Führungs- und Störsignale gefordert, da sprungförmige Signale eine gute Näherung für sich zeitlich ändernde Signale darstellen, die für lange Zeit auf bestimmten Werten verbleiben.

- III) Die Antwort auf einen Führungs- bzw. Störgrößensprung soll hinreichend *gedämpft* sein, d.h. keine zu starken Oszillationen aufweisen. (Wenn beispielsweise nach einem Führungssprung die Regelabweichung zwar abklänge, aber nur sehr langsam, so dass eine lang anhaltende Schwingung entstünde, so wäre der Regelkreis in der Praxis fast unbrauchbar.)
- IV) Der Regelkreis soll genügend schnell sein, d.h. die Antwort auf einen Führungs- bzw. Störgrößensprung soll in hinreichend kurzer Zeit dem stationären Wert genügend nahekommen.

4.8 Stationäres Verhalten

Wir setzen voraus, dass der Regelkreis stabil ist (siehe Abschnitt 5.4), so dass seine Eigenbewegungen mit wachsender Zeit abklingen.

Zur Untersuchung der stationären Genauigkeit wollen wir 2 Fälle betrachten:

$\alpha)$ **Führungssprung** $w(t) = \sigma(t) \left[\rightarrow W(s) = \frac{1}{s} \right]$ bei verschwindender Störung, d.h. $d(t) = 0$.

Beispiele sprungförmiger Führungsgrößen:

- Temperatur soll durch Regelkreis auf Sollwert gehoben werden.
- Greifer eines Roboters soll über einer Bohrung positioniert werden.

Beide Male werden entsprechende Sollwerte \bar{w} vorgegeben.

$\beta)$ Sprungförmige Störung / **Lastsprung** $d(t) = \sigma(t) \left[\rightarrow D(s) = \frac{1}{s} \right]$

Beispiel für Lastsprung:

Zuschalten einer Last in einem Energieversorgungsnetz, deren Wirkung auf die Netzfrequenz durch eine Regelung ausgeglichen werden muss.

Zum Führungssprung $\alpha)$:

Betrachten wir den Regelfehler $e(t)$ in Reaktion auf eine Führungsgröße $w(t)$, so ergibt sich im Frequenzbereich der allgemein gültige Zusammenhang (siehe Abb. 35)

$$\begin{aligned} E(s) &= W(s) - Y(s) = W(s) - G(s)U(s) - D(s) \\ \Rightarrow E(s) &= W(s) - G(s)K(s)E(s) - D(s) \end{aligned} \quad (5.6)$$

Nullsetzen von $D(s)$, wie in Szenario $\alpha)$ vorgesehen, liefert

$$\begin{aligned} E(s)[1 + G(s)K(s)] &= W(s) \\ \Rightarrow E(s) &= \underbrace{\frac{1}{1 + G(s)K(s)}}_{G_d(s)} W(s) \end{aligned} \quad (5.7)$$

Um aus $E(s)$ das Langzeitverhalten von $e(t)$ zu gewinnen, benötigen wir den **Grenzwertsatz der Laplace-Transformation** :

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} s \cdot F(s) \quad (5.8)$$

(Voraussetzung: Die auf beiden Seiten der Gleichung stehenden Grenzwerte existieren wirklich.)

Anwendung auf Gl. (5.7) $\left[\text{mit } W(s) = \frac{1}{s} \right]$ bringt uns

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s \cdot E(s) = \lim_{s \rightarrow 0} s \cdot \frac{1}{1 + G_0(s)} \cdot \frac{1}{s} = \frac{1}{1 + \lim_{s \rightarrow 0} G_0(s)}, \quad (5.9)$$

sofern $G_0(0) \neq -1$ ist.

Zum Lastsprung β :

Wir verfahren analog zu oben.

Regelfehler in Reaktion auf Störgröße (bei zu Null gesetzter Führungsgröße):

$$(5.6) \quad \begin{matrix} W(s)=0 \\ \Rightarrow \end{matrix} E(s) = -G(s)K(s)E(s) - D(s)$$

$$\Rightarrow E(s) = \frac{-1}{1 + G_0(s)} D(s) \quad (5.10)$$

$$\begin{matrix} D(s)=\frac{1}{s} \\ \Rightarrow \end{matrix} \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s \cdot \frac{-1}{1 + G_0(s)} \cdot \frac{1}{s} = \frac{-1}{1 + \lim_{s \rightarrow 0} G_0(s)} \quad (5.11)$$

Die bleibende Regelabweichung $e(\infty)$ ist also beim Führungs- und beim Lastsprung (= sprungförmige Störung) betragsmäßig gleich groß und von der Übertragungsfunktion der offenen Kette für $s \rightarrow 0$ abhängig.

Sehr häufig lässt sich das Übertragungsverhalten der offenen Kette durch eine allgemeine Standardübertragungsfunktion der Form

$$G_0(s) = \frac{k}{s^m} \cdot \frac{b_q s^q + \dots b_1 s + b_0}{a_{n-m} s^{n-m} + \dots + a_1 s + a_0} \cdot e^{-T_t s}, \quad \text{mit } n \geq q, \quad (5.12)$$

beschreiben. (Die offene Kette kann also als Reihenschaltung von m reinen Integralgliedern, einem Proportionalglied der Ordnung $(n-m)$ und einem Totzeitglied interpretiert werden. In der Darstellung (5.12) steckt die Annahme, dass $a_0 \neq 0$, dass sich also kein weiterer Faktor s aus dem Nennerpolynom „herausziehen“ lässt.)

Der Exponent m bestimmt nun den Charakter der Übertragungsfunktion $G_0(s)$:

A) Für $m = 0$ hat die offene Kette **proportionales Verhalten**. Es gilt

$$G_0(0) \equiv k_0 = k \cdot \frac{b_0}{a_0}$$

und (beim Führungssprung)

$$\lim_{t \rightarrow \infty} e(t) = \frac{1}{1 + G_0(0)} = \frac{1}{1 + k_0} . \quad (5.13)$$

Der Regelkreis hat eine bleibende Regelabweichung, deren Größe von der statischen Verstärkung k_0 der offenen Kette abhängt. Je größer k_0 , um so kleiner die bleibende Regelabweichung.

B) Für $m = 1$ (und allgemeiner für $m \geq 1$) hat die offene Kette **integrales Verhalten**, und es gilt

$$\lim_{s \rightarrow 0} G_0(s) = \infty$$

und (beim Führungssprung)

$$\lim_{t \rightarrow \infty} e(t) = \frac{1}{1 + \lim_{s \rightarrow 0} G_0(s)} = 0 . \quad (5.14)$$

Für $m \geq 1$ hat der Regelkreis also keine bleibende Regelabweichung.

Fassen wir diese Ergebnisse zur **Störkompensation** und **Sollwertfolge bei sprungförmigen Signalen** zusammen:

Damit in einem stabilen Regelkreis die Regelgröße einem sprungförmigen Führungssignal ohne bleibende Regelabweichung nachgeführt wird und sprungförmige, auf den Ausgang der Regelstrecke wirkende Störungen ohne bleibende Regelabweichung (d.h. ohne bleibende Auslenkung) kompensiert werden, muss die offene Kette integrales Verhalten (kurz: I-Verhalten) besitzen.

Das heißt, in der Übertragungsfunktion $G_0(s)$ der offenen Kette muss man den Faktor $\frac{1}{s}$ ausklammern können.

Wenn die Regelstrecke selbst keinen I-Anteil besitzt, muss man den Regler mit einem I-Anteil versehen, um stationäre Genauigkeit zu erzielen.

Regelstrecken ohne I-Anteil, die also stationär (d.h. im Langzeitlimit) Proportional-(kurz: P-)Verhalten aufweisen, sind i.a. in Anlagen zu finden, die im Bereich eines gegebenen Arbeitspunktes arbeiten. Beispiele:

- Heizungsregelung mit der Raumtemperatur als Regelgröße, dem Ventilhub als Stellgröße und der Außentemperatur als Störgröße.
- Elektrische Antriebsregelung mit der Drehzahl als Regelgröße, der Ankerspannung als Stellgröße und dem Lastmoment als Störgröße (siehe Abb. 36).

Regelstrecke mit (einfachem) I-Verhalten, bei denen also im eingeschwungenen Zustand zu einer konstanten Stellgröße u eine konstante Geschwindigkeit \dot{y} der Regelgröße gehört, findet man meistens in Positionier- und Servosystemen. Beispiel: Elektrischer Stellmotor mit der Ankerspannung als Stellgröße und dem Winkel der Antriebswelle als Regelgröße.

Falls eine Regelstrecke mit P-Verhalten vorliegt, so lautet das Resümee der obigen Untersuchung:

Um genügende stationäre Genauigkeit zu erzielen, muss man entweder die Kreisverstärkung k_0 genügend groß machen oder in die offene Kette ein I-Glied einfügen. Ein Grundproblem der Regelkreissynthese besteht nun darin, dass die Forderungen nach Stabilität und stationärer Genauigkeit gegensätzlich sind. Erhöhung von k_0 oder Einführung eines I-Gliedes bringt die Tendenz zur Instabilität mit sich. Um dies zu verstehen, beschäftigen wir uns nun mit der Stabilität von Regelkreisen.

4.9 Stabilität

5.4.1 E/A - Stabilität

In Abschnitt 3.3.3 haben wir Systeme, deren Eigenbewegungen alle asymptotisch abklingen – was daran abzulesen ist, dass alle Eigenwerte bzw. Pole negative Realteile haben –, als asymptotisch stabil bezeichnet. Ein System mit dieser Eigenschaft reagiert auf eine „Erregung“ in Form einer Auslenkung \mathbf{x}_0 des Zustandes aus der Gleichgewichtslage damit, dass es von diesem Anfangszustand wieder in die Gleichgewichtslage zurückkehrt.

Bei einem Übertragungssystem – und insbesondere bei einem Regelkreis – interessiert uns aber i.a. weniger die „Erregung“ in Form einer Auslenkung des Systemzustands, sondern vielmehr die Erregung von außen durch ein Eingangssignal. Stabilität heißt dann, dass das System auf ein beschränktes Eingangssignal mit einem ebenfalls beschränkten Ausgangssignal antwortet.

Konkret definieren wir:

Ein lineares System, das durch eine DGL vom Typ (3.15) oder in Form eines Zustandsraummodells (3.28), (3.29) gegeben ist, heißt **eingangs-ausgangs-stabil (E/A-stabil)**, wenn für verschwindende Anfangsauslenkung $\mathbf{x}_0 = \mathbf{0}$ und ein beliebiges beschränktes Eingangssignal

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

das Ausgangssignal ebenfalls beschränkt bleibt:

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

(Im Angelsächsischen wird dieser Stabilitätsbegriff auch BIBO-Stabilität genannt für „bounded input – bounded output“.)

Ein beschränktes Eingangssignal $u(t)$ schlägt sich wegen

$$y(t) = \int_0^t g(t-\tau) u(\tau) d\tau \quad (\text{für } \mathbf{x}_0 = \mathbf{0})$$

in folgender Abschätzung des Ausgangssignals $y(t)$ nieder:

$$\begin{aligned} |y(t)| &\leq \int_0^t |g(t-\tau)| |u(\tau)| d\tau \\ &\leq |u_{\max}| \int_0^t |g(t-\tau)| d\tau = |u_{\max}| \int_0^t |g(\tau)| d\tau \end{aligned}$$

Folglich ist das System E/A - stabil, wenn das Integral

$$\int_0^t |g(\tau)| d\tau$$

für alle Zeiten t beschränkt bleibt, d.h. wenn

$$\int_0^{\infty} |g(\tau)| d\tau < \infty \quad (5.15)$$

gilt. Da auch der Umkehrschluss gilt, haben wir nun ein Kriterium, um ein System auf E/A-Stabilität hin zu untersuchen: Ein lineares System (das in Form seiner Zustandsraumdarstellung gegeben ist) ist genau dann E/A-stabil, wenn seine Gewichtsfunktion $g(t)$ die Bedingung (5.15) erfüllt.

Die Gewichtsfunktion muss also für große Zeiten hinreichend schnell dem Wert Null zustreben. Aus der Modenzerlegung der Gewichtsfunktion

$$g(t) = \sum_{i=1}^n g_i e^{\lambda_i t} + d \delta(t)$$

ersehen wir, dass die Bedingung (5.15) auf jeden Fall dann erfüllt ist, wenn alle e-Funktionen abklingen, wenn also alle Eigenwerte λ_i negativen Realteil haben. Somit haben wir ein für den Entwurf E/A-stabiler Systeme (v.a.: E/A-stabiler Regelkreise) geeignetes **Stabilitätskriterium** gefunden:

Wenn alle Eigenwerte und damit alle Pole der ungekürzten Übertragungsfunktion einen negativen Realteil haben, so ist das System E/A-stabil.	(5.16)
---	--------

Bemerkungen:

- 1) Dem Regelungstechniker kommt es vor allem darauf an, beim Entwurf eines Regelkreises durch Erfüllung bestimmter Bedingungen die E/A-Stabilität sicherzustellen. Daher wird ihn die Tatsache, dass Bedingung (5.16) zwar hinreichend, aber nicht notwendig für E/A-Stabilität ist, i.a. nicht stören.

- 2) Die Beziehung zwischen asymptotischer Stabilität und E/A-Stabilität soll noch kurz beleuchtet werden.

Aus (5.16) folgt sofort: Ist ein System asymptotisch stabil, so ist es auch E/A-stabil.

Umgekehrt können wir aus der E/A-Stabilität eines Systems auf seine asymptotische Stabilität schließen, sofern alle Eigenbewegungen des Systems in die Gewichtsfunktion $g(t)$ eingehen, d.h. sofern $g_i \neq 0$ für alle $i = 1, \dots, n$. [Im Falle $g_i = 0$ ist die zu λ_i gehörige Eigenbewegung entweder prinzipiell nicht durch ein Eingangssignal anregbar oder hat keinen Einfluss auf die Ausgangsgröße, siehe Abschnitt 3.6. Im Frequenzbereich äußert sich $g_i = 0$ darin, dass ein Pol s_i und eine Nullstelle der Übertragungsfunktion $G(s)$ zusammenfallen, so dass der in Zähler und Nenner von $G(s)$ gemeinsame Linearfaktor $(s - s_i)$ gekürzt werden kann.]

- 3) Das Stabilitätskriterium (5.16), das wir hier für den Fall einer rationalen Übertragungsfunktion hergeleitet haben, gilt auch dann, wenn das betrachtete System ein Regelkreis ist, der ein Totzeitglied (oder mehrere davon) enthält.

5.4.2 Das Nyquist-Kriterium

Wir wollen nun die E/A-Stabilität des Standard-Regelkreises (siehe Abb. 35) untersuchen. Das E/A-Verhalten dieses Kreises wird durch die Führungsübertragungsfunktion $G_w(s)$ und die Störübertragungsfunktion $G_d(s)$ beschrieben:

$$Y(s) = G_w(s) \cdot W(s) + G_d(s) \cdot D(s) = \frac{G_0(s)}{1 + G_0(s)} W(s) + \frac{1}{1 + G_0(s)} D(s)$$

Die Übertragungsfunktionen $G_w(s)$ und $G_d(s)$ haben dasselbe Nennerpolynom

$$F(s) = 1 + G_0(s)$$

und folglich auch dieselben Pole. Diese Pole (die Pole des geschlossenen Kreises!) können als Nullstellen von $F(s)$ berechnet werden, d.h. als Lösungen der Gleichung (= sog. **charakteristische Gleichung des geschlossenen Regelkreises**):

$$1 + G_0(s) = 0 \tag{5.17}$$

Ein solches Vorgehen, um abzuklären, ob der Regelkreis E/A-stabil ist, hat 2 Nachteile:

- a) Erheblicher numerischer Aufwand bei Systemen höherer Ordnung.

- b) Oft ist der Regler nur in seiner Struktur festgelegt, seine Parameter müssen hingegen noch bestimmt werden. Gl. (5.17) stellt nun leider einen unüberschaubaren Zusammenhang zwischen den Polen des geschlossenen Regelkreises und den Reglerparametern dar. Wenn die Rechnung z.B. für einen Parametersatz ergibt, dass der Kreis instabil ist, dann erhält man keinen Hinweis, wie der Regler zu modifizieren ist, um den Kreis zu stabilisieren.

Die Nullstellen der charakteristischen Gleichung (5.17) zu berechnen, um den Regelkreis auf E/A-Stabilität hin zu prüfen, schießt über das Ziel hinaus. Die **Werte** der Pole sind ja für die E/A-Stabilität gar nicht maßgebend. Es muss lediglich geprüft werden, ob alle Pole in der linken komplexen Halbebene liegen. Dieser Gedanke liegt dem sog. **Nyquist-Kriterium** zugrunde, mit dem die E/A-Stabilität des geschlossenen Regelkreises geprüft werden kann, ohne die Pole zu berechnen.

Eine einfache Version des Nyquist-Kriteriums soll nun erarbeitet werden.

Hierfür betrachten wir zunächst die komplexwertige Übertragungsfunktion $G_0(s)$: Sie bildet die komplexe „s-Ebene“ in die komplexe „z-Ebene“ ab (siehe Abb. 37). Dabei geht die positive imaginäre Achse (also die Punkte $s = j\omega$) in die Ortskurve des Frequenzgangs der offenen Kette über: $j\omega \mapsto G_0(j\omega)$.

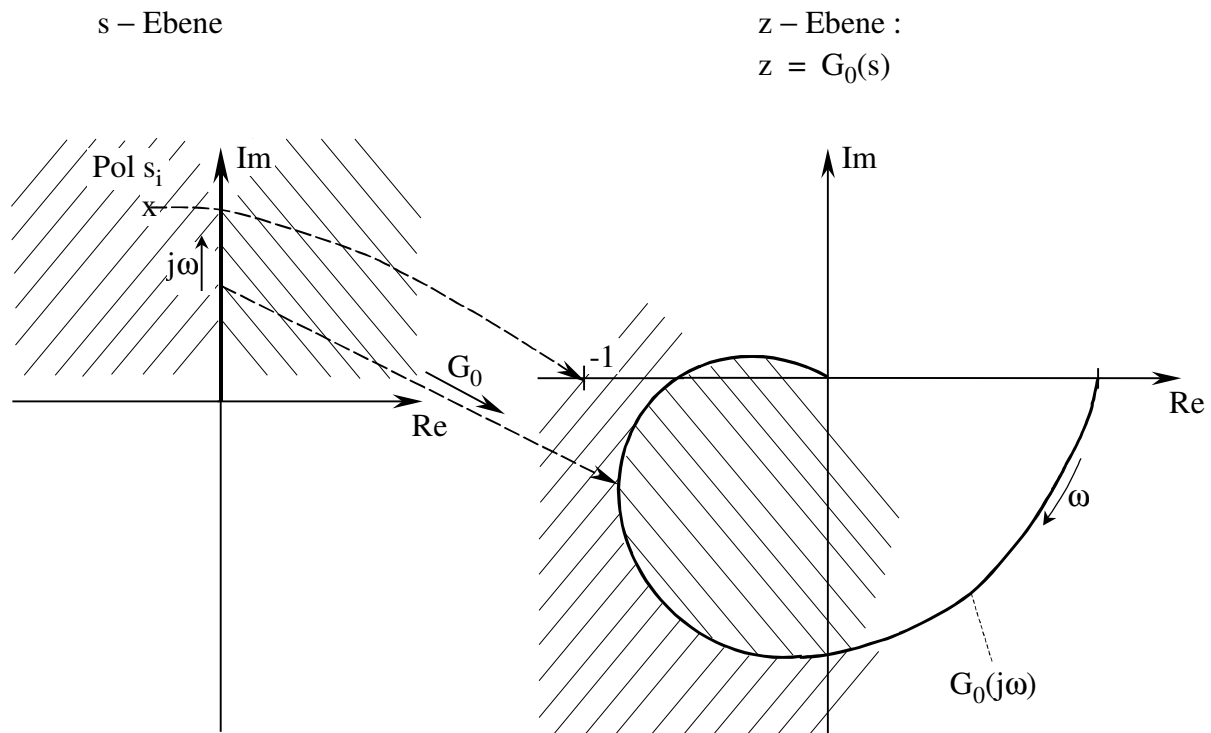


Abb. 37 : Abbildung der komplexen „s-Ebene“ in die komplexe „z-Ebene“ durch die Übertragungsfunktion $G_0(s)$ der offenen Kette

Aus der linken Seite der (positiven) j -Achse wird die linke Seite der Ortskurve, und umgekehrt stammt das, was auf der rechten Seite der Ortskurve liegt, von der rechten Seite der j -Achse.

Dieses Wissen wenden wir nun auf den Punkt $(-1 + j \cdot 0)$ der z -Ebene an: Genau dann, wenn dieser Punkt – wie in Abb. 37 – links von der Ortskurve liegt, werden seine s -Urbilder [d.h. die Punkte der s -Ebene mit $G_0(s) = -1$] in der linken Hälfte der s -Ebene zu finden sein.

Wegen

$$1 + G(s) = 0 \quad \Leftrightarrow \quad G_0(s) = -1$$

sind die s -Urbilder des Punktes $(-1 + j \cdot 0)$ aber nichts anderes als die Nullstellen der charakteristischen Gleichung (5.17) und damit die Pole des geschlossenen Regelkreises. Der Regelkreis wird also genau dann asymptotisch stabil (\Rightarrow insbesondere : E/A-stabil) sein, wenn der Punkt $(-1 + j \cdot 0)$ links von der Ortskurve liegt.

Wenn man obige Plausibilitätsbetrachtungen mathematisch präzisiert, bestätigt sich die soeben getroffene Aussage – vorausgesetzt, die offene Kette ist stabil –, und wir erhalten eine einfache Version des Nyquist-Kriteriums, die sog.

Linke-Hand-Regel :

Eine stabile offene Kette bleibt bei Schließung zum Regelkreis genau dann asymptotisch stabil, wenn der „kritische Punkt“ $(-1 + j \cdot 0)$ links von der in Richtung wachsender Frequenzen durchlaufenen Ortskurve von $G_0(j\omega)$ liegt (siehe Abb. 37).

Bemerkung:

Das Nyquist-Kriterium beruht auf der Frequenzgang-Ortskurve $G_0(j\omega)$ der offenen Kette. Wenn, wie es oft der Fall ist, die offene Kette als Reihen-(und/oder Parallel-)schaltung einzelner Übertragungsglieder, deren Kennwerte bekannt sind, gegeben ist, so lässt sich $G_0(j\omega)$ leicht berechnen. Im Rahmen des Nyquist-Kriteriums können aber auch experimentell ermittelte Frequenzgänge einzelner Regelkreisglieder oder der gesamten offenen Kette Verwendung finden.

5.4.3 Das Phasenrandkriterium

Man betrachte die in Abb. 38 (oben) skizzierte Ortskurve eines Regelkreises mit stabiler offener Kette.

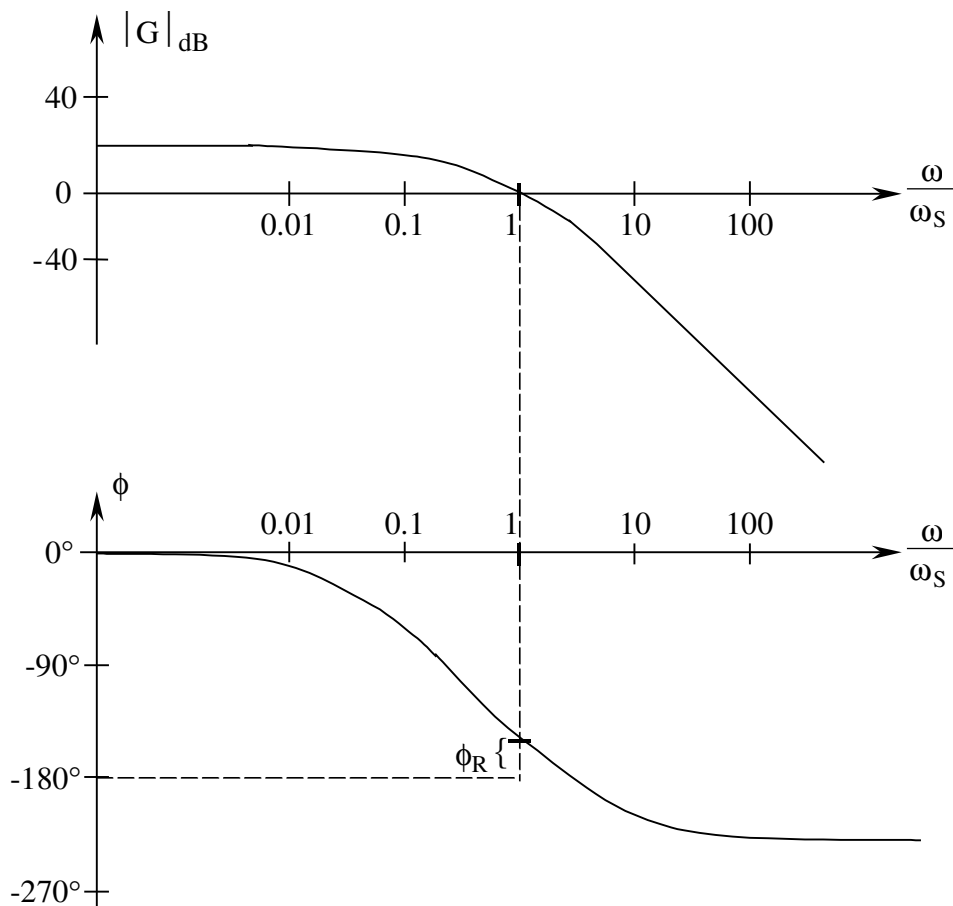
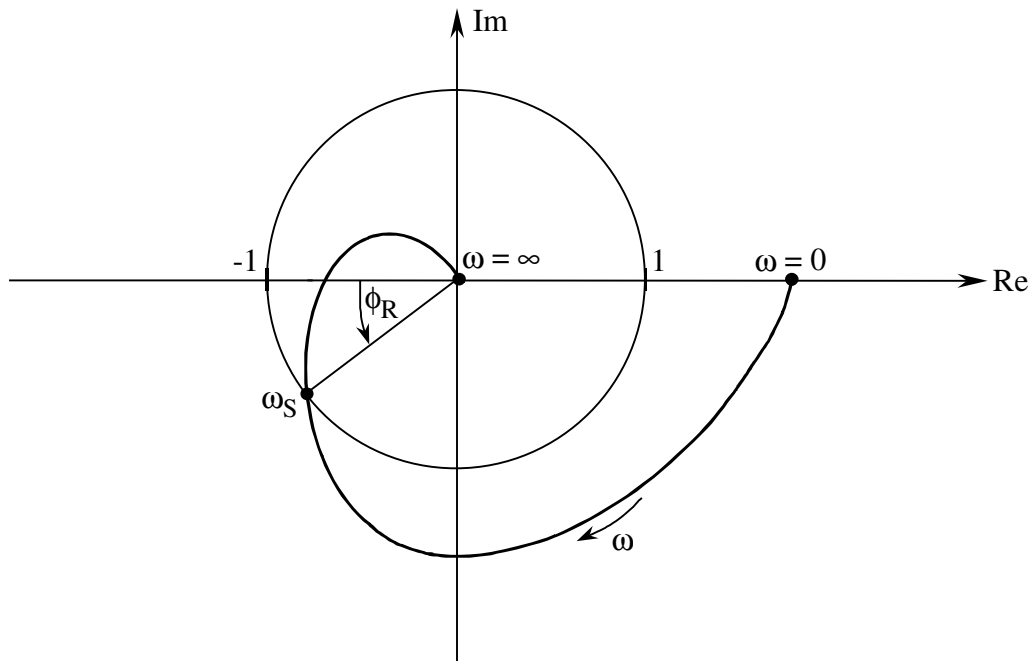


Abb. 38 : Phasenrand in Ortskurvendarstellung (oben) und Bodediagramm

Ob entsprechend des Nyquist-Kriteriums (in Form der Linken-Hand-Regel) der kritische Punkt $(-1 + j \cdot 0)$ von der Ortskurve umschlungen wird oder nicht, lässt sich dort ablesen, wo die Ortskurve den Einheitskreis schneidet. Die Frequenz ω dieses Schnittpunktes wird als Schnittfrequenz ω_S bezeichnet, d.h. für ω_S gilt

$$|G_0(j\omega_S)| = 1 \quad \text{bzw.} \quad |G_0(j\omega_S)|_{\text{dB}} = 0.$$

Der **Phasenrand** ϕ_R (auch Phasenreserve genannt) bezeichnet den Abstand der Phase $\phi(\omega_S)$ von -180° (siehe Abb. 38):

$$\phi_R = 180^\circ - |\phi(\omega_S)|.$$

Der Phasenrand ist also genau dann positiv, wenn die Schnittphase $\phi(\omega_S)$ betragsmäßig kleiner als 180° ist.

Betrachtet man eine offene Kette, deren Ortskurve den Einheitskreis nur einmal schneidet, so muss nach dem Nyquist-Kriterium der Schnittpunkt im dritten Quadranten der komplexen Ebene liegen (siehe Abb. 38).

⇒ **Phasenrandkriterium:**

Eine stabile offene Kette, deren Phasengang nur eine Schnittfrequenz ω_S besitzt, bleibt bei Schließung zum Regelkreis genau dann asymptotisch stabil, wenn der Phasenrand ϕ_R positiv ist. Im Bodediagramm heißt das, dass bei der Frequenz ω_S , bei der der Amplitudengang die 0dB-Achse schneidet, der Phasengang oberhalb der „ (-180°) – Marke“ liegen muss (siehe Abb. 38 unten).

Die Bedingung des Phasenrandkriteriums, dass der Amplitudengang die 0dB-Achse nur einmal schneidet, ist bei I- und P-Ketten meistens erfüllt.

Anhand des Phasenrandkriteriums können wir uns die Wirkung von Totzeitgliedern auf die Stabilität von Regelkreisen klarmachen. Die offene Kette sei beschrieben durch

$$G_0(j\omega) = \underbrace{\bar{G}_0(j\omega)}_{\substack{\text{totzeitfreies} \\ \text{System mit} \\ \text{Phasenrand } \phi_R}} \cdot \underbrace{e^{-j\omega T_t}}_{\text{Totzeitglied}}.$$

Da das Totzeitglied nur die Phase von $G_0(j\omega)$, nicht jedoch die Amplitude beeinflusst, verändert es den Phasengang „nach unten“, ohne den Amplitudengang zu beeinflussen. Die Schnittfrequenz ω_S bleibt also dieselbe, aber die zu dieser Schnittfrequenz gehörige Phase ist kleiner. Totzeitglieder verkleinern den Phasenrand und führen zu instabilen Regelkreisen, wenn

die (durch das Totzeitglied bei ω_S bewirkte) Phasenveränderung größer als der Phasenrand $\bar{\phi}_R$ ist.

Grobe Abschätzung:

Schneidet der Amplitudengang die 0dB-Achse mit -20 dB/Dekade (wie bei der zweiten asymptotischen Geraden des PT_1 -Amplitudengangs), so können wir annehmen, dass auch der Phasengang des totzeitfreien Anteils $\bar{G}_0(j\omega)$ dem des PT_1 -Gliedes ähnelt und bei ω_S eine Phasenverschiebung $\bar{\phi}_R$ von etwa -90° aufweist (siehe Abb. 39). Liegt die Schnittfrequenz bei $\omega_S = \frac{1}{T_t}$ (Annahme !), so verschiebt das Totzeitglied die Phase dort um -57° , so dass ein Phasenrand von $\phi_R = 180^\circ - 90^\circ - 57^\circ = 33^\circ$ entsteht. Einen Phasenrand von 33° sollte man aber für einen Regelkreis mindestens fordern, um neben der Stabilität auch noch eine hinreichende Dämpfung des Einschwingvorgangs zu erzielen. Tritt also in der offenen Kette die Totzeit T_t auf, so sollte die Schnittfrequenz des totzeitfreien Anteils $\bar{G}_0(j\omega)$ der offenen Kette die Bedingung

$$\omega_S \stackrel{!}{<} \frac{1}{T_t}$$

erfüllen (siehe Abb. 39).

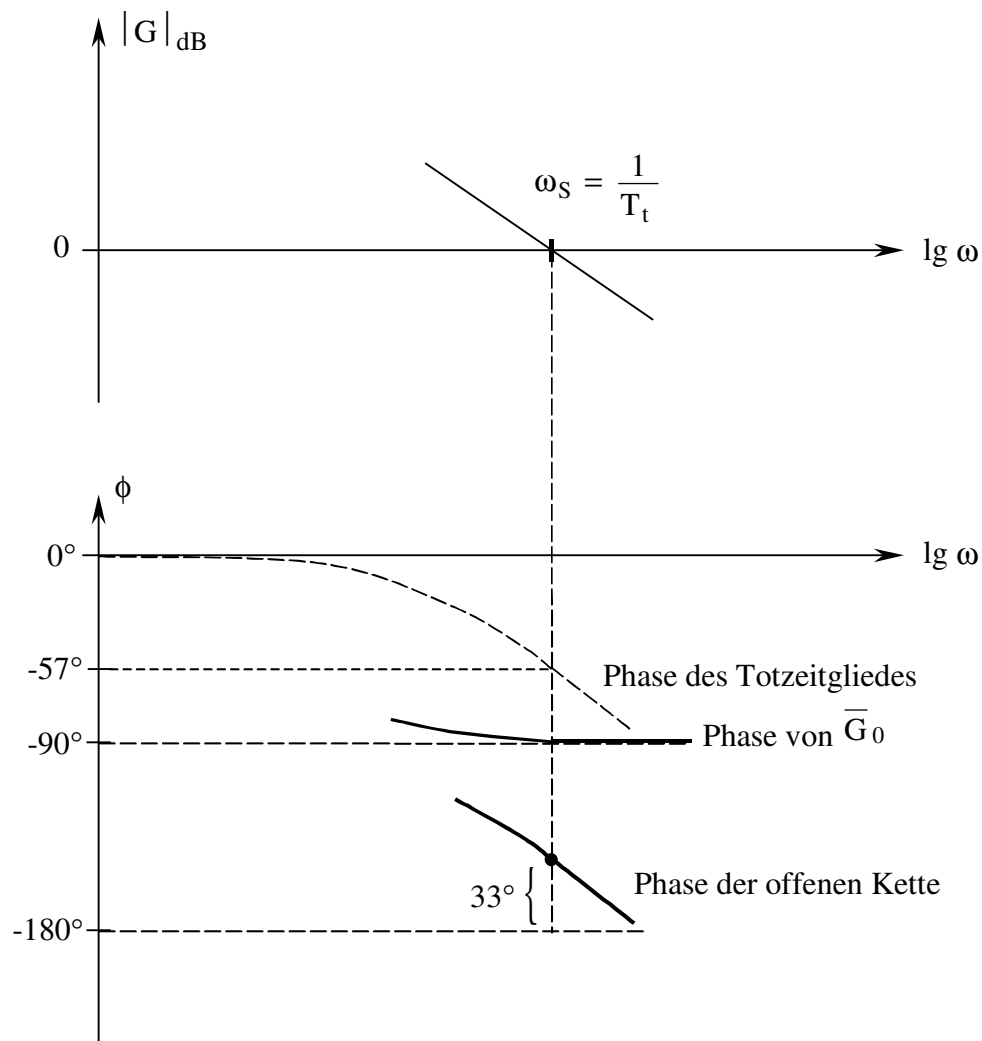


Abb. 39 : Größte mögliche Schnittfrequenz bei offener Kette mit Totzeit

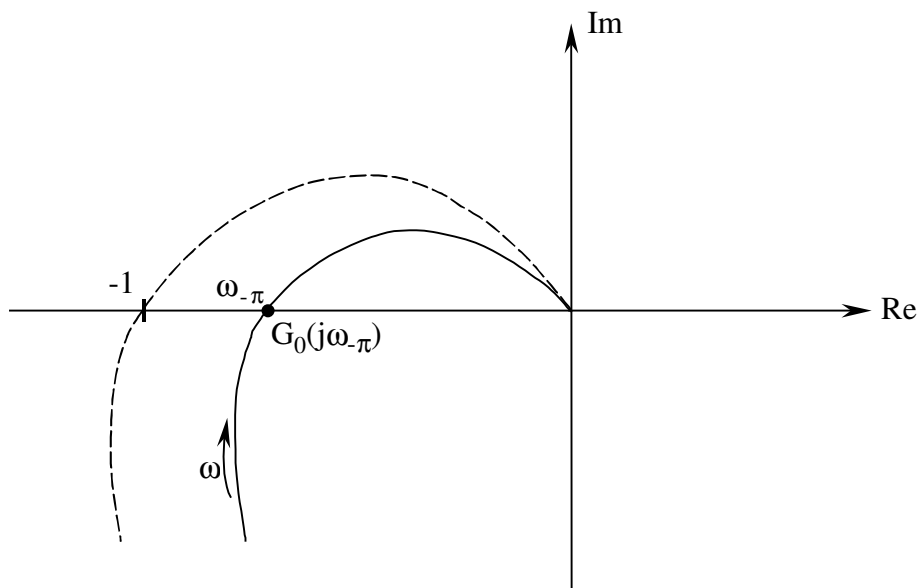
Diese Bedingung zeigt, wie eine Totzeit in der Regelstrecke die Freiheiten einschränkt, die offene Kette durch die Auslegung des Reglers zu gestalten.

5.4.4 Amplitudenrand

Der Frequenzgang $G_0(j\omega)$ einer stabilen offenen Kette lässt sich faktorisieren als

$$G_0(j\omega) = \underbrace{k_0}_{\text{statische Verstärkung}} \cdot \underbrace{\tilde{G}_0(j\omega)}_{\text{Frequenzgang mit Verstärkung 1}}.$$

Wenn wir k_0 vergrößern (bzw. verkleinern), so wird jeder Punkt auf der Ortskurve dem Betrag nach um denselben Faktor vergrößert (bzw. verkleinert), und die Ortskurve weitet sich maßstäblich aus (bzw. schrumpft maßstäblich).



Der Schnittpunkt $G_0(j\omega_{-\pi})$ der Ortskurve mit der negativen reellen Achse [$\omega_{-\pi}$ ist diejenige Frequenz, für die die Phasenverschiebung $-180^\circ (\hat{=} -\pi)$ beträgt;] wandert, wenn wir k_0 vergrößern, nach links. Wenn dabei der Schnittpunkt den kritischen Punkt $(-1 + j0)$ passiert, so wird der Regelkreis laut Nyquist-Kriterium instabil. Der **Amplitudenrand** k_R (auch Amplitudenreserve genannt) bezeichnet nun genau denjenigen Faktor, um den die Kreisverstärkung k_0 angehoben werden darf, ohne dass der Regelkreis instabil wird:

$$G_0(j\omega_{-\pi}) \cdot k_R = -1 \quad \Rightarrow \quad k_R = \frac{1}{|G_0(j\omega_{-\pi})|} \quad (5.18)$$

Bemerkungen:

- Die Kreisverstärkung k_0 können wir über die Verstärkung des Reglers beeinflussen. Das Nyquist-Kriterium (in Form der Betrachtung der Amplitudenreserve) gibt uns also den Spielraum an, innerhalb dessen wir die Reglerverstärkung variieren können.
- Obige Betrachtungen belegen die frühere Aussage, dass der Erhöhung der Kreisverstärkung k_0 – um die stationäre Genauigkeit des Regelkreises zu verbessern – aus Stabilitätsgründen Grenzen gesetzt sind.

4.10 PID - Regler

5.5.1 Das Übertragungsverhalten

Die meisten industriell eingesetzten linearen Reglertypen sind Standardregler, deren Übertragungsverhalten sich auf die drei linearen idealisierten Grundformen des P-, I- und D-Glieds zurückführen lässt. Wichtigster Standardregler im industriellen Bereich: Der PID-Regler, in dem je ein P-, I- und D-Glied parallel geschaltet sind.

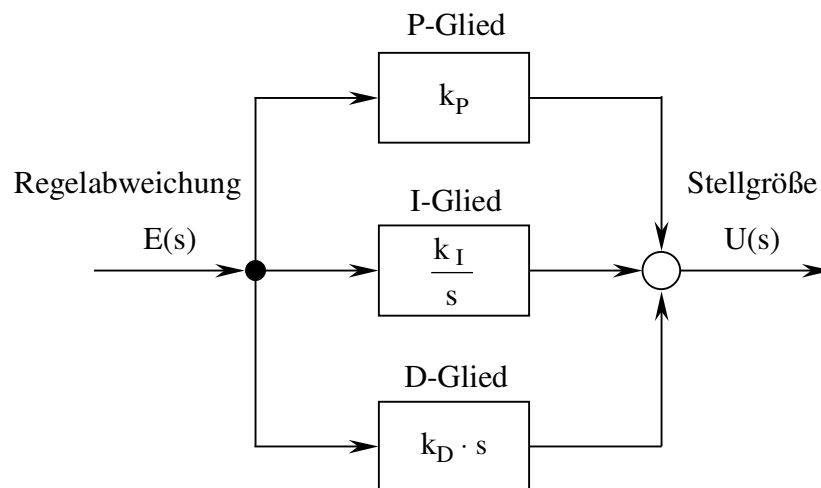


Abb. 40 : Blockschaltbild eines PID-Reglers.

Also lautet die Übertragungsfunktion $K(s)$ des PID-Reglers

$$K_{\text{PID}}(s) = k_P + \frac{k_I}{s} + k_D \cdot s = k_P \left(1 + \frac{1}{T_I s} + T_D s \right), \quad (5.19)$$

wobei wir die sog. **Nachstellzeit** $T_I := \frac{k_P}{k_I}$ und die sog. **Vorhaltezeit** $T_D := \frac{k_D}{k_P}$ eingeführt haben.

Die drei Größen k_P (Verstärkungsfaktor), T_I und T_D sind gewöhnlich in bestimmten Wertebereichen einstellbar und heißen daher Reglerparameter oder Einstellwerte. Durch geeignete Wahl dieser Parameter lässt sich der Regler dem Verhalten der Regelstrecke so anpassen, dass ein möglichst günstiges Regelverhalten entsteht.

Aus (5.19) berechnen wir nun durch Laplace-Rücktransformation, wie die Reglerausgangsgröße $u(t)$ (= Stellgröße !) beim PID-Regler von der Regelabweichung $e(t)$ abhängt:

$$U(s) = k_P E(s) + \frac{k_P}{T_I} \frac{1}{s} E(s) + k_P T_D s \cdot E(s)$$



Faltungs-
satz



Differentiations-
satz

$$u(t) = k_P e(t) + \frac{k_P}{T_I} [\sigma(t) * e(t)] + k_P T_D \frac{de(t)}{dt}$$

$$\Rightarrow u(t) = k_P e(t) + \frac{k_P}{T_I} \int_0^t e(\tau) d\tau + k_P T_D \dot{e}(t) \quad (5.20)$$

Interpretation dieser Formel:

Der P-Anteil arbeitet nach dem Prinzip:

„Je größer die Regelabweichung ist, um so größer muss die Stellgröße sein.“

Nach dieser Regel kann man Regelabweichungen zwar schnell abbauen, aber man kann i.a. nicht sichern, dass die Regelabweichung vollständig verschwindet. Dafür ist der I-Anteil notwendig (siehe Abschnitt 5.3 zur stationären Genauigkeit), der nach der Regel

„Solange eine Regelabweichung auftritt, muss die Stellgröße **verändert** werden.“

arbeitet. Ist $e = 0$, so wird die Stellgröße durch den I-Anteil nicht verändert, was so interpretiert werden kann, dass derjenige Wert der Stellgröße, mit dem die Störung kompensiert bzw. der vorgegebene Sollwert erreicht wird, gefunden ist. Der D-Anteil reagiert nur auf Veränderungen der Regelabweichung nach dem Prinzip:

„Je stärker sich die Regelabweichung verändert, um so stärker muss die Regelung eingreifen.“

Über den D-Anteil reagiert der Regler bereits dann mit einer großen Stellgröße, wenn die Regelabweichung stark zunimmt, selbst wenn sie noch keine großen Werte angenommen hat.

Setzt man in Gl. (5.20) $e(t) = \sigma(t)$, so erhält man die Sprungantwort des PID-Reglers

$$h_{PID}(t) = k_P + \frac{k_P}{T_I} t + k_P T_D \delta(t), \quad (5.21)$$

wie in Abb. 41 (links) skizziert.

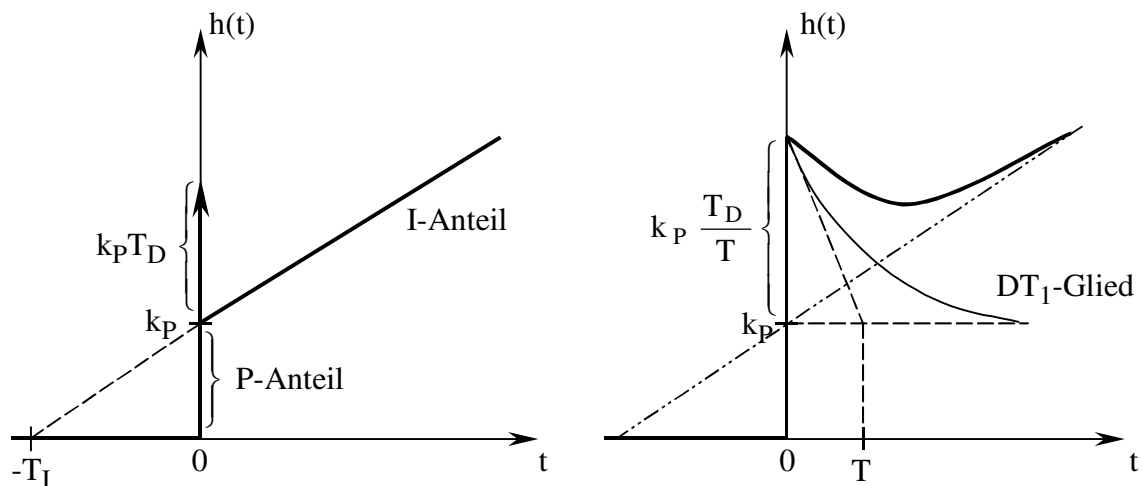


Abb. 41 : Sprungantwort des idealen (links) und des realen (rechts) PID-Reglers

Das reine D-Verhalten verletzt die Kausalitätsbedingung und kann daher gerätetechnisch nicht verwirklicht werden. Bei tatsächlich ausgeführten Reglern ist das D-Verhalten stets mit einer gewissen Verzögerung behaftet. An die Stelle des D-Gliedes in der Parallelschaltung von Abb. 40 tritt daher in der Realisierung ein DT_1 -Glied mit der Übertragungsfunktion

$$G_D(s) = k_P \frac{T_D s}{T s + 1} = \underbrace{k_P T_D s}_{\text{D-Glied}} \cdot \underbrace{\frac{1}{T s + 1}}_{\substack{\text{verzögerndes} \\ \text{PT}_1\text{-Glied}}},$$

wobei die Zeitkonstante T der Verzögerung sehr klein gewählt wird. Für $T \ll T_D$ hat das zusätzliche Verzögerungsglied $\frac{1}{T s + 1}$ praktisch keine Wirkung, so dass man beim Entwurf davon ausgehen kann, dass der Regler einen reinen D-Anteil besitzt.

Aus dem PID-Regler entstehen durch Weglassen einzelner Anteile die Spezialfälle des PI-, des PD-, des P- und des I-Reglers. Beim Einsatz industrieller Standardregler sind P-, I- und D-Komponenten beliebig zusammenschaltbar, so dass die aufgeführten Spezialfälle problemlos realisiert werden können.

Als erster Schritt zur Lösung einer industriellen Regelungsaufgabe ist festzulegen, ob mit einem PID-Regler gearbeitet werden muss oder ob einfachere Reglertypen für die Lösung der Aufgabe ausreichen. Diese Wahl der Reglerstruktur hängt von den gestellten Güteforderungen und den Eigenschaften der Regelstrecke ab.

Faustregeln:

- Um zu verhindern, dass bei sprungförmigen Führungs- und Störsignalen eine bleibende Regelabweichung entsteht, muss die offene Kette I-Verhalten aufweisen (siehe Abschnitt 5.3). Bei Strecken mit I-Anteil reicht für die Verhinderung einer bleibenden Regelabweichung ein P-Regler aus. Zeigt die Strecke jedoch proportionales Verhalten, so muss der Regler einen I-Anteil aufweisen.
- Bei reinen I-Reglern steigt die Stellgröße langsam an und verändert sich nicht sprungförmig. Der I-Regler kann also nicht schnell auf große Regelabweichungen reagieren. Der Regelkreis hat daher ein langsames Übergangsverhalten.
- P- und D-Anteile beschleunigen das Übergangsverhalten des Regelkreises, da der Regler sehr schnell auf Veränderungen der Regelabweichung reagiert. Der Regelkreis neigt aber aufgrund des D-Anteils, insbesondere bei großen Werten von k_D , zu großen Schwingungen oder wird instabil. Mit dem D-Anteil sollte auch deshalb vorsichtig umgegangen werden, weil stochastische Störungen wie z.B. das Messrauschen durch das Differenzierglied verstärkt werden. Der D-Anteil darf nur bei gut gefilterten Messgrößen verwendet werden.

5.5.2 Einstellregeln für PID-Regler

In vielen Anwendungsfällen, insbesondere aus der Verfahrenstechnik, ist eine mathematische Modellierung der Regelstrecke sehr aufwendig. Oftmals ist dabei das dynamische Verhalten der Regelstrecke unkritisch und hinreichend gedämpft. Für solche Fälle wurden empirisch bzw. durch Simulationsuntersuchungen Einstellregeln entwickelt, die es ermöglichen, einen Standardregler vom P-, PI- oder PID-Typ ohne detaillierte Kenntnis der Strecke und ohne ein aufwendiges Syntheseverfahren zu dimensionieren und damit bereits recht gute Regelungsergebnisse zu erzielen.

A) *Methode des Stabilitätsrandes* (nach Ziegler-Nichols)

Dieser Weg zur Reglereinstellung beruht darauf, dass zunächst die Stabilitätsgrenze des proportional geregelten Systems experimentell ermittelt wird. Die Methode kommt vollkommen ohne Regelstreckenmodell aus, ist jedoch nur bei Prozessen einsetzbar, die bis an die Stabilitätsgrenze gefahren werden können.

Voraussetzung: Die Regelstrecke ist stabil, hat Proportionalverhalten und kann zeitweise im grenzstabilen Bereich betrieben werden.

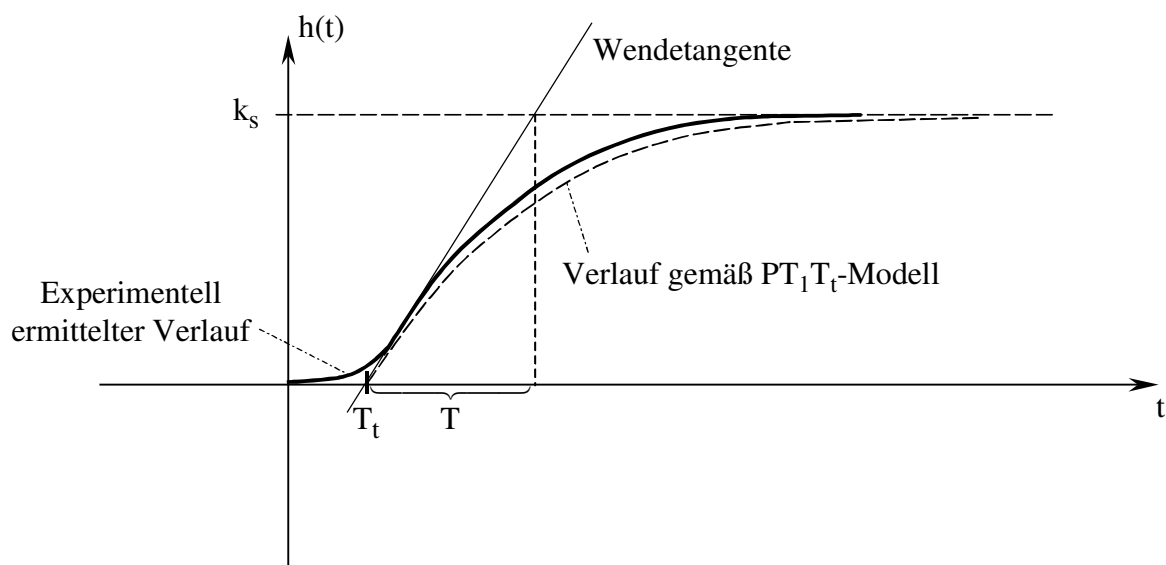
1. Schritt: Der Regelkreis wird mit Hilfe eines P-Reglers geschlossen.
2. Schritt: Die Reglerverstärkung wird solange erhöht, bis der geschlossene Kreis nach einer Sollwertänderung eine Dauerschwingung ausführt. Die dabei eingestellte Reglerverstärkung heißt kritische Verstärkung k_{krit} , die Periodendauer der Schwingung T_{krit} .
3. Schritt: Die Reglerparameter werden gemäß folgender Tabelle festgelegt:

Reglertyp	Verstärkung k_P	Nachstellzeit T_I	Vorhaltezeit T_D
P – Regler	$0.5 k_{krit}$	–	–
PI – Regler	$0.45 k_{krit}$	$0.85 T_{krit}$	–
PID – Regler	$0.6 k_{krit}$	$0.5 T_{krit}$	$0.12 T_{krit}$

B) *Wendetangenten – Methode* (nach Ziegler-Nichols)

Diese Methode ist anwendbar, wenn die Regelstrecke stabil ist und P-Verhalten mit einer aperiodischen Übergangsfunktion aufweist.

1. Schritt: Man gebe eine Sprungfunktion $u(t) = u_0 \cdot \sigma(t)$ auf den Streckeneingang und nehme (durch Messung) die Sprungantwort auf. Division der Sprungantwort durch u_0 ergibt die Übergangsfunktion $h(t)$:



Den experimentell ermittelten aperiodischen Verlauf der Übergangsfunktion kann man näherungsweise durch ein PT_1T_t -Modell, also eine Hintereinanderschaltung eines Totzeitgliedes mit Totzeit T_t und eines PT_1 -Gliedes mit Verzögerungszeit T und Verstärkungsfaktor k_s , beschreiben (vgl. obige Abbildung):

$$G(s) \approx \frac{k_s}{T s + 1} \cdot e^{-s T_t}$$

2. Schritt: Die Kennwerte T_t , T und k_s des PT_1T_t -Modells werden mit Hilfe der Wendetangente aus der aufgezeichneten Übergangsfunktion ermittelt (siehe obige Abb.).

3. Schritt: Die Reglerparameter werden gemäß folgender Tabelle festgelegt:

Reglertyp	Verstärkung k_P	Nachstellzeit T_I	Vorhaltezeit T_D
P – Regler	$k_P = \frac{1}{k_s} \frac{T}{T_t}$	–	–
PI – Regler	$k_P = \frac{0.9}{k_s} \frac{T}{T_t}$	$3.33 T_t$	–
PID – Regler	$k_P = \frac{1.2}{k_s} \frac{T}{T_t}$	$2 T_t$	$0.5 T_t$